云图
YUN TU

2021版

张宇数学教育系列丛书·一

张宇考研数学基础30讲

主编　张宇

北京理工大学出版社

张宇数学教育系列丛书 · 一

张宇考研数学基础30讲

○主编　张宇

张宇数学教育系列丛书编委（按姓氏拼音排序）

毕泗真　蔡燧林　陈静静　崔巧莲　高昆轮　胡金德　贾建厂　雷会娟
史明洁　王成富　王慧珍　王燕星　谢珩　徐兵　严守权　亦一（笔名）　于吉霞
曾凡（笔名）　曾熊　张聪聪　张乐　张青云　张婷婷　张宇　郑光玉　郑利娜
朱杰　朱坤颇

云图
YUN TU

北京理工大学出版社

图书在版编目(CIP)数据

张宇考研数学基础30讲/张宇主编. —北京：北京理工大学出版社，2019.8
ISBN 978-7-5682-7466-1

Ⅰ.①张…　Ⅱ.①张…　Ⅲ.①高等数学-研究生-入学考试-自学参考资料　Ⅳ.①O13

中国版本图书馆CIP数据核字(2019)第181980号

出版发行／北京理工大学出版社有限责任公司
社　　址／北京市海淀区中关村南大街5号
邮　　编／100081
电　　话／(010)68914775(总编室)
　　　　(010)82562903(教材售后服务热线)
　　　　(010)68948351(其他图书服务热线)
网　　址／http://www.bitpress.com.cn
经　　销／全国各地新华书店
印　　刷／天津市蓟县宏图印务有限公司
开　　本／787毫米×1092毫米　1/16
印　　张／33.75
字　　数／842千字
版　　次／2019年8月第1版　2019年8月第1次印刷
定　　价／85.80元

责任编辑／多海鹏
文案编辑／多海鹏
责任校对／周瑞红
责任印制／李志强

前 言

这本书是专门供学生考研数学基础复习之用的。之所以叫《张宇考研数学基础 30 讲》，是因为将考研数学中的全部基础知识系统化和科学化地分成了 30 个部分，希望考生一讲一讲地学，一关一关地过，最终建立起考研数学的基础知识结构，实现真正意义上的夯实基础。

一、这是基础课笔记

这是我在基础课上讲出来的笔记，学生可以听着课跟我一页一页地学，我把笔记写好了，你可以集中精力认真听，不需再记大量笔记；不上课的学生也完全可以一页一页地自学，我几乎把要说的话一句一句都写出来了，请务必搞懂吃透。

二、这是课后作业

我会选择书中某些好题作为课后作业，所有题目均有详细解答，课后务必及时落实。

三、这是真正意义上的考研数学基础教材

考研数学命题并没有指定教材，学生可以自行选择市面上的各种教材进行复习，但有一个专业问题：市面上的数学教材大多是为大学数学教学而编写，依据的是《本科教学基本要求》，鲜见真正意义上按照《全国硕士研究生招生考试数学考试大纲》（简称《考试大纲》）编写的数学教材，尤其是基础教材，本书就是在多年一线考研辅导基础上做出的最新成果。

四、这是答疑解惑

起步阶段的复习，很多学生会遇到各种问题和疑惑：知识理解上的问题，思路方法上的疑惑。本书集中回答并希望能够切实解决学生的各种问题和疑惑。

五、这是减负不是增负

不论你在读哪本数学教材，本书都可以作为思考总结的笔记，放在手边随时翻阅，基础阶段的知识、思路、题型和方法，皆会以清晰的结构呈现在你面前，把握在你手中。你若能再添砖加瓦，画龙点睛，将其内化为你自己的，那将是极妙的。

六、看到什么程度

一遍当然不够。反复研读两遍甚至以上，直至字字搞懂、句句通透并熟稔于心。

感谢命题专家们给予的支持、帮助与指导，感谢编辑老师们的辛勤工作和无私奉献，感谢学生们的努力和信任。

本书是我多年基础阶段教学经验的总结，愿助潜心研读者打好地基、夯实基础，勇攀考研数学高峰。

张宇

目录

第一部分 高等数学 …… 1

第 1 讲 高等数学预备知识 …… 3
第 2 讲 数列极限 …… 25
第 3 讲 函数极限与连续性 …… 32
第 4 讲 一元函数微分学的概念与计算 …… 52
第 5 讲 一元函数微分学的几何应用 …… 71
第 6 讲 中值定理 …… 85
第 7 讲 零点问题与微分不等式 …… 95
第 8 讲 一元函数积分学的概念与计算 …… 104
第 9 讲 一元函数积分学的几何应用 …… 142
第 10 讲 积分等式与积分不等式 …… 150
第 11 讲 多元函数微分学 …… 157
第 12 讲 二重积分 …… 174
第 13 讲 常微分方程 …… 184
第 14 讲 无穷级数(仅数学一、数学三要求) …… 196
第 15 讲 数学一、数学二专题内容 …… 218
第 16 讲 数学三专题内容 …… 239
第 17 讲 多元函数积分学的基础知识(仅数学一要求) …… 248
第 18 讲 三重积分、曲线曲面积分(仅数学一要求) …… 260

第二部分 线性代数 …… 287

第 1 讲 行列式 …… 289
第 2 讲 矩阵 …… 308
第 3 讲 向量组 …… 335
第 4 讲 线性方程组 …… 354
第 5 讲 特征值与特征向量 …… 373
第 6 讲 二次型 …… 401

第三部分　概率论与数理统计（仅数学一、数学三要求） ······ 419

第 1 讲　随机事件与概率 ······ 421
第 2 讲　一维随机变量及其分布 ······ 442
第 3 讲　多维随机变量及其分布 ······ 466
第 4 讲　随机变量的数字特征 ······ 494
第 5 讲　大数定律与中心极限定理 ······ 509
第 6 讲　数理统计 ······ 514

第一部分

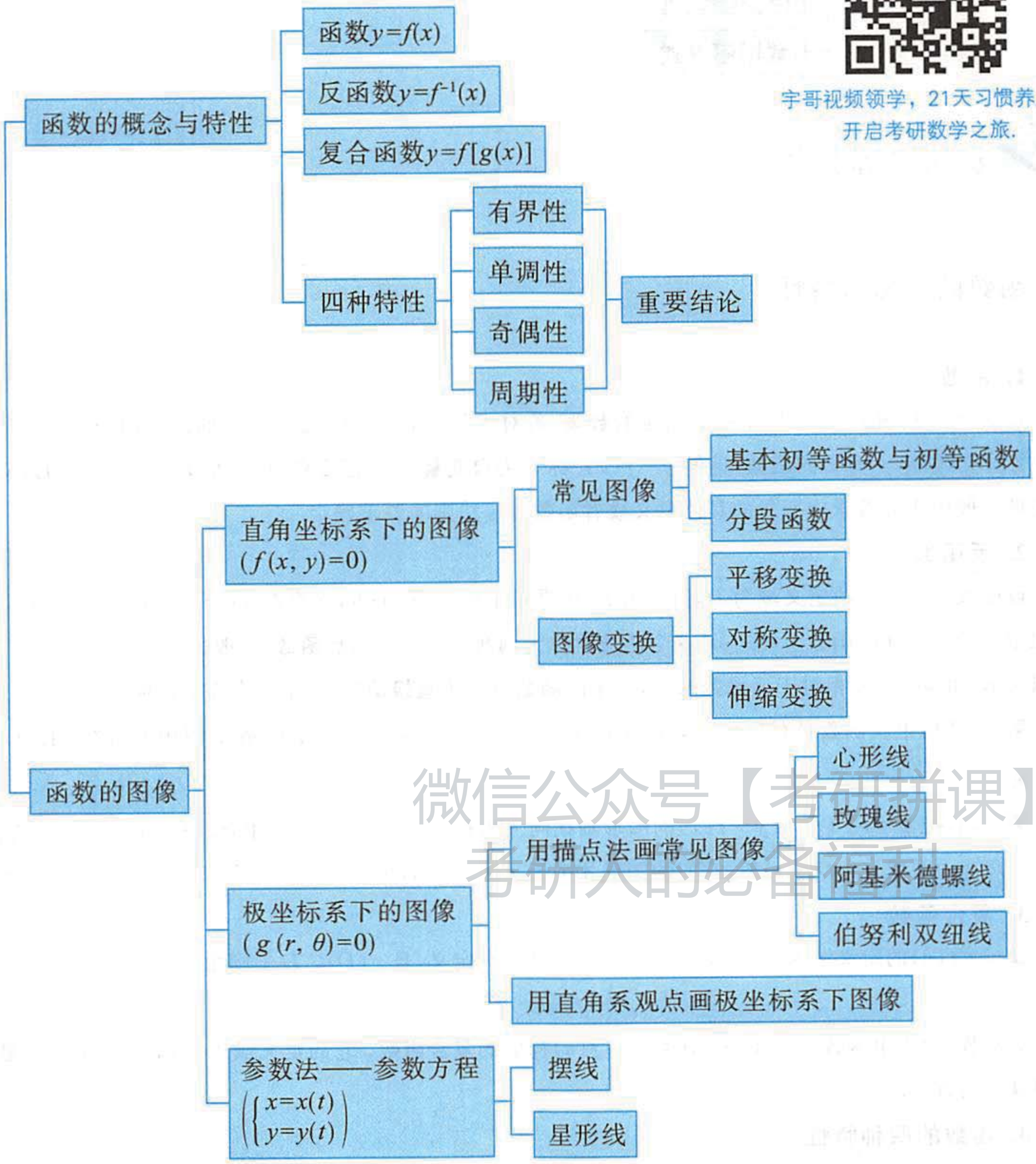
高等数学

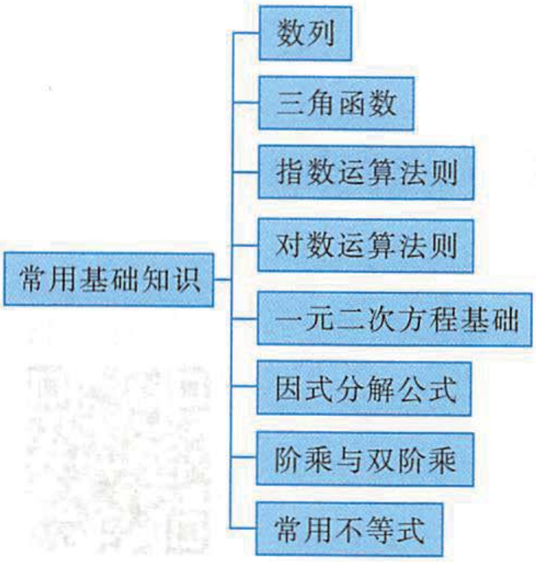
第1讲 高等数学预备知识

基础知识结构

“宇”你同行

宇哥视频领学，21天习惯养成，
开启考研数学之旅.





基础内容精讲

一、函数的概念与特性

1. 函数

设 x 与 y 是两个变量，D 是一个给定的数集，若对于每个值 $x\in D$，按照一定的法则，有一个确定的值 y 与之对应，则称 y 为 x 的**函数**，记作 $y=f(x)$. 称 x 为**自变量**，y 为**因变量**. 称数集 D 为此函数的**定义域**，定义域一般由实际背景中变量的具体意义或者函数对应法则的要求确定.

2. 反函数

设函数 $y=f(x)$ 的定义域为 D，值域为 R. 如果对于每一个 $y\in R$，必存在 $x\in D$ 使得 $y=f(x)$ 成立，则由此定义了一个新的函数 $x=\varphi(y)$. 这个函数就称为函数 $y=f(x)$ 的**反函数**，一般记作 $x=f^{-1}(y)$，它的定义域为 R，值域为 D. 相对于反函数来说，原来的函数也称为**直接函数**. 以下两点需要说明.

第一，严格单调函数必有反函数，比如函数 $y=x^2\,(x\in[0,+\infty))$ 是严格单调函数，故它有反函数 $x=\sqrt{y}$.

第二，若把 $x=f^{-1}(y)$ 与 $y=f(x)$ 的图形画在同一坐标系中，则它们完全重合. 只有把 $y=f(x)$ 的反函数 $x=f^{-1}(y)$ 写成 $y=f^{-1}(x)$ 后，它们的图形才关于 $y=x$ 对称，事实上这也是字母 x 与 y 互换的结果.

3. 复合函数

设 $y=f(u)$ 的定义域为 D_1，函数 $u=g(x)$ 在 D 上有定义，且 $g(D)\subset D_1$，则由

$$y=f[g(x)]\,(x\in D)$$

确定的函数，称为由函数 $u=g(x)$ 和函数 $y=f(u)$ 构成的**复合函数**，它的定义域为 D，u 称为**中间变量**，要掌握复合的方法.

4. 函数的四种特性

(1)有界性.

设 $f(x)$ 的定义域为 D，数集 $I\subset D$. 如果存在某个正数 M，使对任一 $x\in I$，有 $|f(x)|\leqslant M$，则称 $f(x)$

在 I 上有界；如果这样的 M 不存在，则称 $f(x)$ 在 I 上无界.

【注】 (1)从几何上看，如果在给定的区间，函数 $y=f(x)$ 的图形能够被直线 $y=-M$ 和 $y=M$ “完全包起来”，则为有界；从解析上说，找到某个正数 M，使得 $|f(x)|\leqslant M$，则为有界.

(2)有界还是无界的讨论首先是**指明区间** I，不知区间，无法谈论有界性. 比如 $y=\frac{1}{x}$ 在 $(2,+\infty)$ 内有界，但在 $(0,2)$ 内无界.

(3)事实上，只要在区间 I 上存在点 x_0，使得函数 $\lim\limits_{x\to x_0}f(x)$ 的值为无穷大，则没有任何两条直线 $y=-M$ 和 $y=M$ 可以把 I 上的 $f(x)$“包起来”，这就叫无界. 考研中常出这样的题目，比如例 1.3.2.

(2)单调性.

设 $f(x)$ 的定义域为 D，区间 $I\subset D$. 如果对于区间 I 上任意两点 x_1,x_2，当 $x_1<x_2$ 时，恒有 $f(x_1)<f(x_2)$，则称 $f(x)$ 在区间 I 上**单调增加**. 如果对于区间 I 上任意两点 x_1,x_2，当 $x_1<x_2$ 时，恒有 $f(x_1)>f(x_2)$，则称 $f(x)$ 在区间 I 上**单调减少**.

【注】 后面会看到，在考研试题中常常用求导来讨论函数在某个区间上的单调性，但是定义法不可以忘记. 试题中也用到如下定义法的判别形式，请读者留意.

对任何 $x_1,x_2\in D, x_1\neq x_2$，则

$$f(x)\text{是单调增函数}\Leftrightarrow(x_1-x_2)[f(x_1)-f(x_2)]>0;$$

$$f(x)\text{是单调减函数}\Leftrightarrow(x_1-x_2)[f(x_1)-f(x_2)]<0;$$

$$f(x)\text{是单调不减函数}\Leftrightarrow(x_1-x_2)[f(x_1)-f(x_2)]\geqslant 0;$$

$$f(x)\text{是单调不增函数}\Leftrightarrow(x_1-x_2)[f(x_1)-f(x_2)]\leqslant 0.$$

(3)奇偶性.

设 $f(x)$ 的定义域 D 关于原点对称(即若 $x\in D$，则 $-x\in D$). 如果对于任一 $x\in D$，恒有 $f(-x)=f(x)$，则称 $f(x)$ 为**偶函数**. 如果对于任一 $x\in D$，恒有 $f(-x)=-f(x)$，则称 $f(x)$ 为**奇函数**. 我们熟知的是，偶函数的图形关于 y 轴对称，奇函数的图形关于原点对称.

【注】 设 $f(x)$ 是定义在 $[-l,l]$ 上的任意函数，则

$$F_1(x)=f(x)-f(-x)\text{必为}\textbf{奇函数}；\quad F_2(x)=f(x)+f(-x)\text{必为}\textbf{偶函数}.$$

显然 $u(x)=\frac{1}{2}[f(x)+f(-x)]$ 是偶函数，$v(x)=\frac{1}{2}[f(x)-f(-x)]$ 是奇函数，而

$$f(x)=\frac{1}{2}[f(x)+f(-x)]+\frac{1}{2}[f(x)-f(-x)]=u(x)+v(x).$$

(1)奇函数 $y=f(x)$ 的图形关于坐标原点对称，当 $f(x)$ 在 $x=0$ 处有定义时，必有 $f(0)=0$.

(2)偶函数 $y=f(x)$ 的图形关于 y 轴对称，且当 $f'(0)$ 存在时，必有 $f'(0)=0$.

(3)函数 $y=f(x)$ 与 $y=-f(x)$ 的图形关于 x 轴对称；函数 $y=f(x)$ 与 $y=f(-x)$ 的图形关于 y 轴对称；函数 $y=f(x)$ 与 $y=-f(-x)$ 的图形关于原点对称.

(4)函数 $y=f(x)$ 的图形关于直线 $x=T$ 对称的充分必要条件是

$$f(x)=f(2T-x)\text{或}f(x+T)=f(T-x).$$

(4)周期性.

设 $f(x)$ 的定义域为 D，如果存在一个正数 T，使得对于任一 $x\in D$，有 $x\pm T\in D$，且 $f(x+T)=f(x)$，则称 $f(x)$ 为**周期函数**，T 称为 $f(x)$ 的**周期**. 从几何图形上看，在周期函数的定义域内，相邻两个长度为 T 的区间上，函数的图形完全一样.

(5)重要结论.

事实上，关于 $f'(x)$ 和 $\int_a^x f(t)\mathrm{d}t$ 的性质才是这个知识的落脚点，先提前总结在这里：

① **若 $f(x)$ 是可导的偶函数，则 $f'(x)$ 是奇函数，见例** 1.4.1(1)；

② **若 $f(x)$ 是可导的奇函数，则 $f'(x)$ 是偶函数，见例** 1.4.1(2)；

③ **若 $f(x)$ 是可导的周期为 T 的周期函数，则 $f'(x)$ 也是以 T 为周期的周期函数，见例** 1.4.2；

④ **连续的奇函数的一切原函数都是偶函数，见例** 1.8.6；

⑤ **连续的偶函数的原函数中仅有一个原函数是奇函数，见例** 1.8.6；

⑥ **若连续函数 $f(x)$ 以 T 为周期且 $\int_0^T f(x)\mathrm{d}x=0$，则 $f(x)$ 的一切原函数也以 T 为周期. 见例** 1.8.8；

⑦ **若 $f(x)$ 在有限区间 (a,b) 内可导且 $f'(x)$ 有界，则 $f(x)$ 在 (a,b) 内有界.**

二、函数的图像

(一)直角坐标系下的图像($f(x,y)=0$)

1. 常见图像

(1)基本初等函数与初等函数.

基本初等函数：常数函数、幂函数、指数函数、对数函数、三角函数、反三角函数.

①常数函数.

$y=A$，A 为常数，其图形为平行于 x 轴的水平直线(如图 1-1-1).

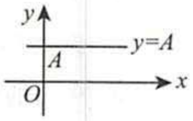


图 1-1-1

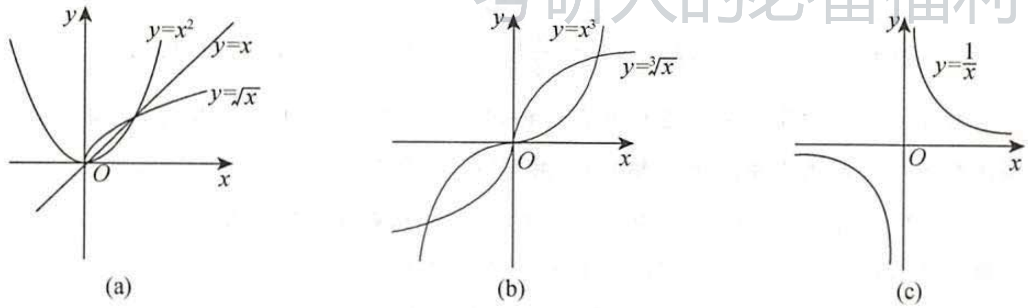
②幂函数.

$y=x^{\mu}$(μ 是实数).

【注】 (1) $y=x^{\mu}$ 的定义域和值域取决于 μ 的值. 当 $x>0$ 时，$y=x^{\mu}$ 都有定义.

(2)常用的幂函数(如图 1-1-2(a)—(c)).

$$y=x,\quad y=x^2,\quad y=\sqrt{x},\quad y=x^3,\quad y=\sqrt[3]{x},\quad y=\frac{1}{x}.$$

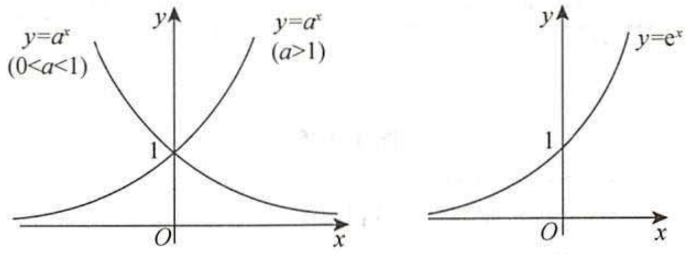


(a) (b) (c)

图 1-1-2

③指数函数.

$y=a^x(a>0,a\neq1)$(如图 1-1-3(a)).



(a)　　(b)

图 1-1-3

【注】 (1)定义域：$(-\infty,+\infty)$. 值域：$(0,+\infty)$.

(2)单调性：当 $a>1$ 时，$y=a^x$ 单调增加；当 $0<a<1$ 时，$y=a^x$ 单调减少.

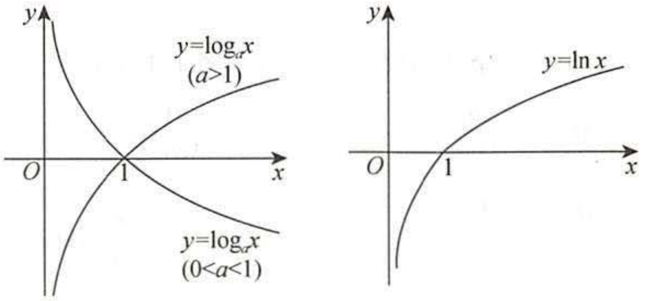
(3)常用的指数函数：$y=e^x$(如图 1-1-3(b)).

(4)极限：$\lim\limits_{x\to-\infty} e^x=0$, $\lim\limits_{x\to+\infty} e^x=+\infty$.

(5)特殊函数值：$a^0=1, e^0=1$.

④对数函数.

$y=\log_a x(a>0,a\neq1)$(如图 1-1-4(a))是 $y=a^x$ 的反函数.



(a)　　(b)

图 1-1-4

【注】 (1)定义域：$(0,+\infty)$. 值域：$(-\infty,+\infty)$.

(2)单调性：当 $a>1$ 时，$y=\log_a x$ 单调增加；当 $0<a<1$ 时，$y=\log_a x$ 单调减少.

(3)常用的对数函数：$y=\ln x$(自然对数：$\ln x=\log_e x$, $e=2.71828\cdots$)(如图 1-1-4(b)).

(4)特殊函数值：$\log_a 1=0, \log_a a=1, \ln 1=0, \ln e=1$.

(5)极限：$\lim\limits_{x\to0^+}\ln x=-\infty$, $\lim\limits_{x\to+\infty}\ln x=+\infty$.

(6)常用公式：$x=e^{\ln x}, u^v=e^{\ln u^v}=e^{v\ln u}(x>0,u>0)$.

⑤三角函数.

(ⅰ)正弦函数与余弦函数.

正弦函数 $y=\sin x$(如图 1-1-5(a))，余弦函数 $y=\cos x$(如图 1-1-5(b)).

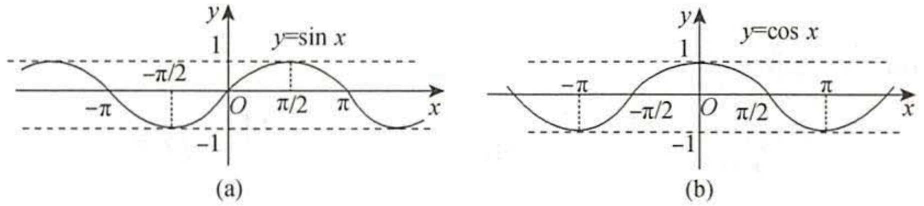


图 1-1-5

【注】 (1)定义域：$(-\infty,+\infty)$. 值域：$[-1,1]$.

(2)奇偶性：$y=\sin x$ 是奇函数，$y=\cos x$ 是偶函数，$x\in(-\infty,+\infty)$.

(3)周期性：$y=\sin x$ 和 $y=\cos x$ 均以 2π 为最小正周期，$x\in(-\infty,+\infty)$.

(4)有界性：$|\sin x|\leqslant 1$，$|\cos x|\leqslant 1$.

(5)特殊函数值： $\sin 0=0$， $\sin\frac{\pi}{6}=\frac{1}{2}$， $\sin\frac{\pi}{4}=\frac{\sqrt{2}}{2}$， $\sin\frac{\pi}{3}=\frac{\sqrt{3}}{2}$，

$\sin\frac{\pi}{2}=1$， $\sin\pi=0$， $\sin\frac{3\pi}{2}=-1$， $\sin 2\pi=0$，

$\cos 0=1$， $\cos\frac{\pi}{6}=\frac{\sqrt{3}}{2}$， $\cos\frac{\pi}{4}=\frac{\sqrt{2}}{2}$， $\cos\frac{\pi}{3}=\frac{1}{2}$，

$\cos\frac{\pi}{2}=0$， $\cos\pi=-1$， $\cos\frac{3\pi}{2}=0$， $\cos 2\pi=1$.

(ⅱ)正切函数与余切函数.

正切函数 $y=\tan x$(如图 1-1-6(a))，余切函数 $y=\cot x$(如图 1-1-6(b)).

$$\tan x=\frac{\sin x}{\cos x},\quad \cot x=\frac{\cos x}{\sin x}=\frac{1}{\tan x}.$$

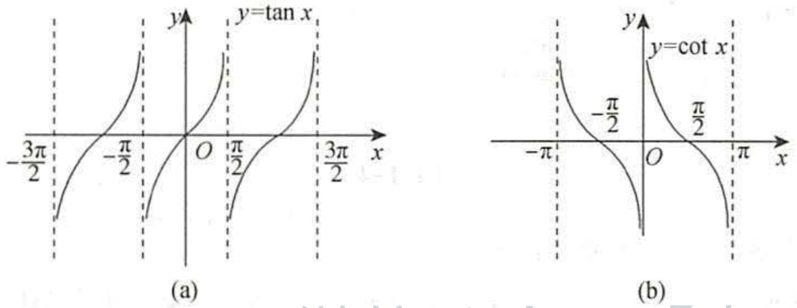


图 1-1-6

【注】 (1)定义域：$y=\tan x$ 的定义域为 $x\neq k\pi+\frac{\pi}{2}(k\in\mathbf{Z})$ 的一切实数 x；

$y=\cot x$ 的定义域为 $x\neq k\pi(k\in\mathbf{Z})$ 的一切实数 x.

值域：$(-\infty,+\infty)$.

(2)奇偶性：$y=\tan x$ 和 $y=\cot x$ 均为奇函数(在其定义域内).

(3)周期性：$y=\tan x$ 和 $y=\cot x$ 均以 π 为最小正周期(在其定义域内).

(4)特殊函数值： $\tan 0=0$， $\tan\frac{\pi}{6}=\frac{\sqrt{3}}{3}$， $\tan\frac{\pi}{4}=1$， $\tan\frac{\pi}{3}=\sqrt{3}$，

$\lim\limits_{x\to\frac{\pi}{2}}\tan x=\infty$，　$\tan \pi=0$，　$\lim\limits_{x\to\frac{3\pi}{2}}\tan x=\infty$，　$\tan 2\pi=0$，

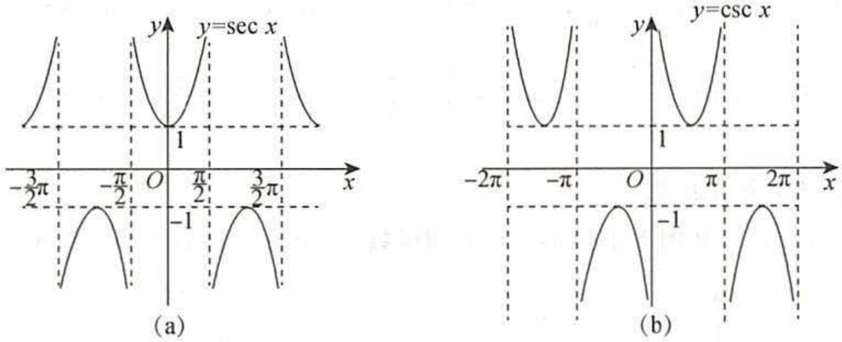
$\lim\limits_{x\to 0}\cot x=\infty$，　$\cot\frac{\pi}{6}=\sqrt{3}$，　$\cot\frac{\pi}{4}=1$，　$\cot\frac{\pi}{3}=\frac{\sqrt{3}}{3}$，

$\cot\frac{\pi}{2}=0$，　$\lim\limits_{x\to\pi}\cot x=\infty$，　$\cot\frac{3\pi}{2}=0$，　$\lim\limits_{x\to 2\pi}\cot x=\infty$.

(ⅲ)正割函数与余割函数.

正割函数 $y=\sec x$(如图 1-1-7(a))，余割函数 $y=\csc x$(如图 1-1-7(b)).

$$\sec x=\frac{1}{\cos x},\quad \csc x=\frac{1}{\sin x}.$$



(a)　　(b)

图 1-1-7

【注】 (1)定义域：$y=\sec x$ 的定义域为 $x\neq k\pi+\frac{\pi}{2}(k\in\mathbf{Z})$的一切实数 x；

$y=\csc x$ 的定义域为 $x\neq k\pi(k\in\mathbf{Z})$的一切实数 x.

值域：$(-\infty,-1]\cup[1,+\infty)$.

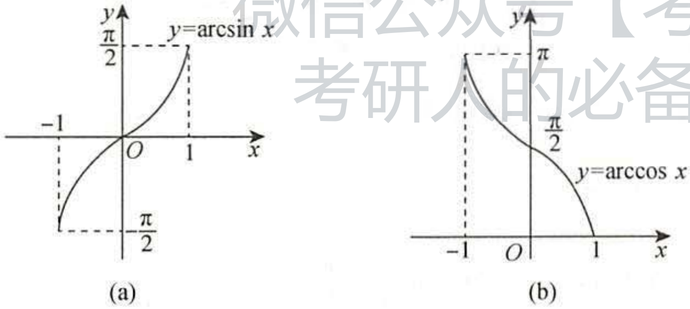
(2)奇偶性：$y=\sec x$ 为偶函数，$y=\csc x$ 为奇函数(在其定义域内).

(3)周期性：$y=\sec x$ 和 $y=\csc x$ 均以 2π 为最小正周期(在其定义域内).

⑥反三角函数.

(ⅰ)反正弦函数与反余弦函数.

反正弦函数 $y=\arcsin x$(如图 1-1-8(a))，反余弦函数 $y=\arccos x$(如图 1-1-8(b)).



(a)　　(b)

图 1-1-8

$y=\arcsin x$ 是 $y=\sin x\left(-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}\right)$的反函数，$y=\arccos x$ 是 $y=\cos x(0\leqslant x\leqslant\pi)$的反函数.

【注】 (1)定义域：$[-1,1]$.

值域：$y=\arcsin x$ 的值域为 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$，$y=\arccos x$ 的值域为 $[0,\pi]$.

(2)单调性：$y=\arcsin x$ 单调增加，$y=\arccos x$ 单调减少.

(3)奇偶性：$y=\arcsin x$ 为奇函数(在其定义域内).

(4)有界性：两个函数在其定义域内有界，$-\frac{\pi}{2}\leqslant\arcsin x\leqslant\frac{\pi}{2}$，$0\leqslant\arccos x\leqslant\pi$.

(5)性质：$\arcsin x+\arccos x=\frac{\pi}{2}(-1\leqslant x\leqslant 1)$.

(6)特殊函数值：

$$\arcsin 0=0,\quad \arcsin\frac{1}{2}=\frac{\pi}{6},\quad \arcsin\frac{\sqrt{2}}{2}=\frac{\pi}{4},\quad \arcsin\frac{\sqrt{3}}{2}=\frac{\pi}{3},\quad \arcsin 1=\frac{\pi}{2},$$

$$\arccos 1=0,\quad \arccos\frac{\sqrt{3}}{2}=\frac{\pi}{6},\quad \arccos\frac{\sqrt{2}}{2}=\frac{\pi}{4},\quad \arccos\frac{1}{2}=\frac{\pi}{3},\quad \arccos 0=\frac{\pi}{2}.$$

(ⅱ)反正切函数与反余切函数.

反正切函数 $y=\arctan x$(如图 1-1-9(a))，反余切函数 $y=\operatorname{arccot} x$(如图 1-1-9(b)).

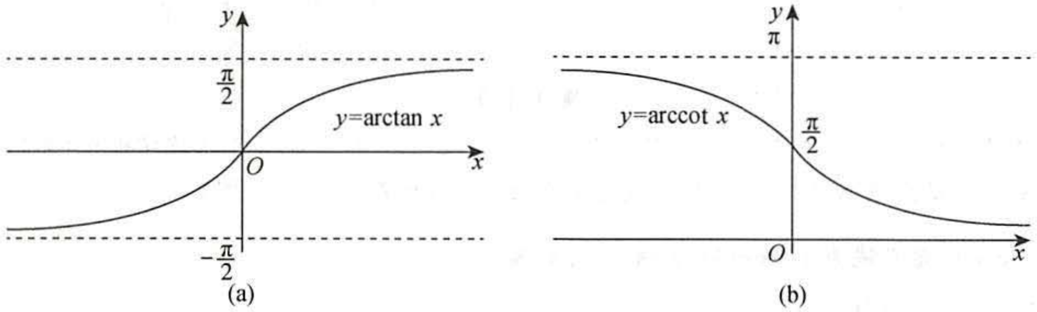


图 1-1-9

$y=\arctan x$ 是 $y=\tan x\left(-\frac{\pi}{2}<x<\frac{\pi}{2}\right)$的反函数，$y=\operatorname{arccot} x$ 是 $y=\cot x(0<x<\pi)$的反函数.

【注】 (1)定义域：$(-\infty,+\infty)$. 值域：$y=\arctan x$ 的值域为 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$，$y=\operatorname{arccot} x$ 的值域为 $(0,\pi)$.

(2)单调性：$y=\arctan x$ 单调增加，$y=\operatorname{arccot} x$ 单调减少.

(3)奇偶性：$y=\arctan x$ 为奇函数(在其定义域内).

(4)有界性：两个函数在其定义域内有界，$-\frac{\pi}{2}<\arctan x<\frac{\pi}{2}$，$0<\operatorname{arccot} x<\pi$.

(5)性质：$\arctan x+\operatorname{arccot} x=\frac{\pi}{2}(-\infty<x<+\infty)$.

(6)特殊函数值：$\arctan 0=0,\quad \arctan\frac{\sqrt{3}}{3}=\frac{\pi}{6},\quad \arctan 1=\frac{\pi}{4},\quad \arctan\sqrt{3}=\frac{\pi}{3},$

$$\operatorname{arccot} 0=\frac{\pi}{2},\quad \operatorname{arccot}\sqrt{3}=\frac{\pi}{6},\quad \operatorname{arccot} 1=\frac{\pi}{4},\quad \operatorname{arccot}\frac{\sqrt{3}}{3}=\frac{\pi}{3}.$$

(7)极限：$\lim\limits_{x\to-\infty}\arctan x=-\frac{\pi}{2}$，$\lim\limits_{x\to+\infty}\arctan x=\frac{\pi}{2}$，$\lim\limits_{x\to-\infty}\operatorname{arccot} x=\pi$，$\lim\limits_{x\to+\infty}\operatorname{arccot} x=0$.

⑦初等函数.

由基本初等函数经有限次的四则运算，以及有限次的复合步骤所构成的并且可以由一个式子所表示的函数称为**初等函数**.

【注】 (1)初等函数的定义域可以是一个区间，也可以是几个区间的并集，甚至可以是一些孤立的点. 例如，$y=\sqrt{\cos \pi x-1}$的定义域是 $x=0,\pm2,\pm4,\cdots$.

(2)幂指函数 $u(x)^{v(x)}=\mathrm{e}^{v(x)\ln u(x)}$ 也是初等函数.

(2)分段函数.

在自变量的不同变化范围中，对应法则用不同式子来表示的函数称为**分段函数**. 需要强调一句，分段函数是用几个式子来表示的一个(不是几个)函数，一般说来它不是初等函数. 分段函数的典型形式如下：

$$f(x)=\begin{cases}\varphi_1(x), & x>x_0,\\ a, & x=x_0,\\ \varphi_2(x), & x<x_0\end{cases} \text{或 } f(x)=\begin{cases}\varphi(x), & x\neq x_0,\\ a, & x=x_0.\end{cases}$$

分段函数很重要，原因在于其形式的复杂性所带来的命题的丰富性. 后面会看到，不论是求极限、求导数，还是求积分，出现最多的研究对象之一便是分段函数.

下面列出三个重要的分段函数.

①$y=|x|=\begin{cases}x, & x\geqslant0,\\ -x, & x<0\end{cases}$ 称为**绝对值函数**，如图 1-1-10 所示.

②$y=\operatorname{sgn} x=\begin{cases}1, & x>0,\\ 0, & x=0,\\ -1, & x<0\end{cases}$ 称为**符号函数**，如图 1-1-11 所示. 对于任何实数 x，有 $x=|x|\operatorname{sgn} x$.

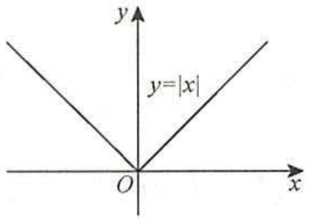


图 1-1-10

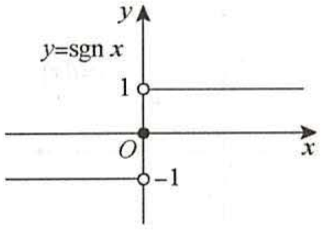


图 1-1-11

③$y=[x]$称为**取整函数**. 先给出定义：设 x 为任一实数，不超过 x 的最大整数称为 x 的整数部分，记作$[x]$. 如

$$[0.99]=0,[\pi]=3,[-1]=-1,[-1.99]=-2.$$

因此，取整函数$y=[x]$的定义域为 $\mathbf{R}$，值域为 $\mathbf{Z}$. 它的图形如图 1-1-12 所示，在 x 为整数值处图形发生跳跃.

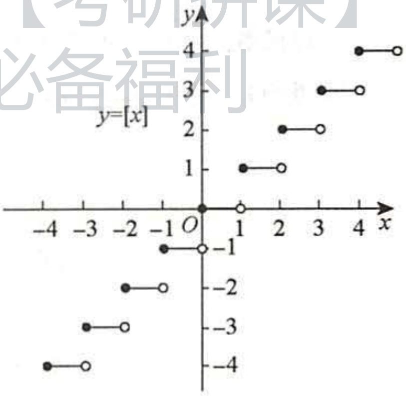


图 1-1-12

从定义出发，以下两点需要读者注意.

(ⅰ)$x-1<[x]\leqslant x$.

(ⅱ)$\lim\limits_{x\to0^+}[x]=0$；$\lim\limits_{x\to0^-}[x]=-1$.

2. 图像变换

图形变换方式一般有如下三种.

(1)平移变换.

①将函数 $y=f(x)$ 的图像沿 x 轴向左平移 $x_0(x_0>0)$ 个单位长度，得到函数 $y=f(x+x_0)$ 的图像(如图 1-1-13)；将函数 $y=f(x)$ 的图像沿 x 轴向右平移 $x_0(x_0>0)$ 个单位长度，得到函数 $y=f(x-x_0)$ 的图像(如图 1-1-14).

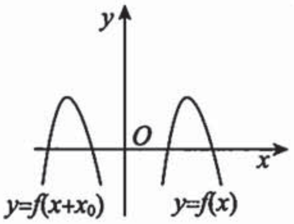


图 1-1-13

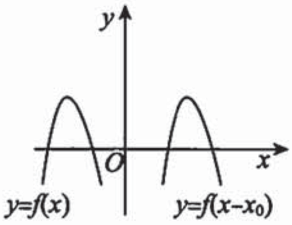


图 1-1-14

②将函数 $y=f(x)$ 的图像沿 y 轴向上平移 $y_0(y_0>0)$ 个单位长度，得到函数 $y=f(x)+y_0$ 的图像(如图 1-1-15)；将 $y=f(x)$ 的图像沿 y 轴向下平移 $y_0(y_0>0)$ 个单位长度，得到函数 $y=f(x)-y_0$ 的图像(如图 1-1-16).

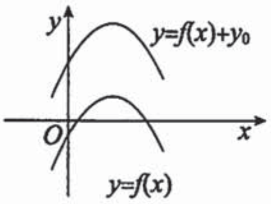


图 1-1-15

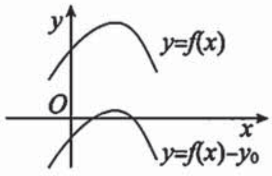


图 1-1-16

(2)对称变换.

①将函数 $y=f(x)$ 的图像关于 x 轴对称，得到函数 $y=-f(x)$ 的图像(如图 1-1-17).

②将函数 $y=f(x)$ 的图像关于 y 轴对称，得到函数 $y=f(-x)$ 的图像(如图 1-1-18).

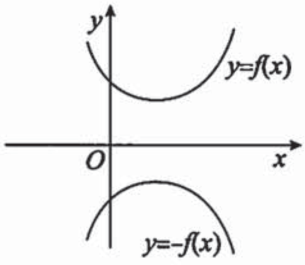


图 1-1-17

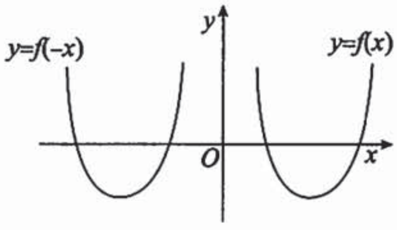


图 1-1-18

③将函数 $y=f(x)$ 的图像关于原点对称，得到函数 $y=-f(-x)$ 的图像(如图 1-1-19).

④将函数 $y=f(x)$ 的图像关于直线 $y=x$ 对称，得到函数 $y=f^{-1}(x)$ 的图像(如图 1-1-20).

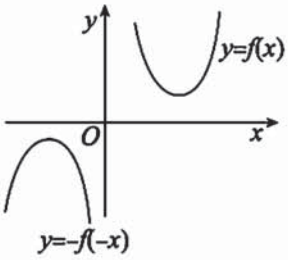


图 1-1-19

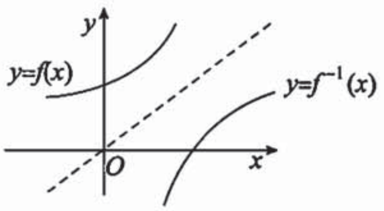


图 1-1-20

⑤保留函数 $y=f(x)$ 在 x 轴及 x 轴上方的部分，把 x 轴下方的部分关于 x 轴对称到 x 轴上并去掉原

来下方的部分，得到函数 $y=|f(x)|$ 的图像（如图 1-1-21）.

⑥保留函数 $y=f(x)$ 在 y 轴及 y 轴右侧的部分，去掉 y 轴左侧的部分，再将 y 轴右侧图像对称到 y 轴左侧，得到函数 $y=f(|x|)$ 的图像（如图 1-1-22）.

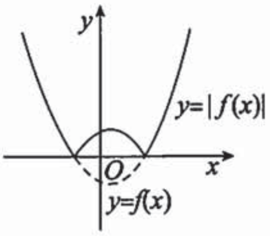


图 1-1-21

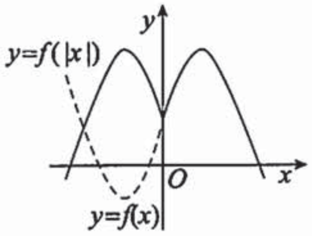


图 1-1-22

(3)伸缩变换.

①水平伸缩：$y=f(kx)(k>1)$ 的图像，可由 $y=f(x)$ 的图像上每点的横坐标缩短到原来的 $\frac{1}{k}$ 倍且纵坐标不变得到（如图 1-1-23）；$y=f(kx)(0<k<1)$ 的图像，可由 $y=f(x)$ 的图像上每点的横坐标伸长到原来 $\frac{1}{k}$ 倍且纵坐标不变得到.

②垂直伸缩：$y=kf(x)(k>1)$ 的图像，可由 $y=f(x)$ 的图像上每点的纵坐标伸长到原来的 k 倍且横坐标不变得到（如图 1-1-24）；$y=kf(x)(0<k<1)$ 的图像，可由 $y=f(x)$ 的图像上每点的纵坐标缩短到原来的 k 倍且横坐标不变得到.

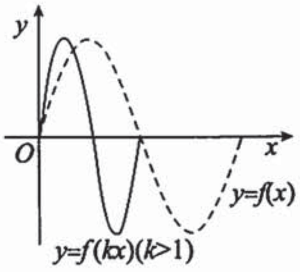


图 1-1-23

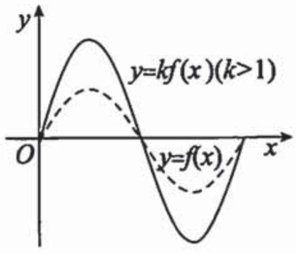


图 1-1-24

(二)极坐标系下的图像（$g(r,\theta)=0$）

1. 用描点法画常见图像

(1)心形线.

下面画出心形线 $r=a(1-\cos\theta)(a>0)$ 的图形.

其表达式的右端是以 2π 为周期的周期函数，作图时只要考虑 $0\leqslant\theta\leqslant2\pi$ 就可以了. 并且，对于方程的右端，θ 换作 $(2\pi-\theta)$ 时，其值不变，也就是说，如 (θ,r) 是曲线上的一个点，则 $(2\pi-\theta,r)$ 也是曲线上的一个点，因此图形以极轴为对称轴，从而只需先考虑 $0\leqslant\theta\leqslant\pi$.

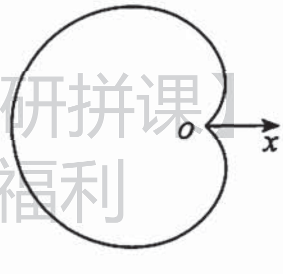


图 1-1-25

当 θ 由 0 增大到 π，$\cos\theta$ 的值由 1 逐渐减小到 -1，从而，r 由 0 逐渐增大到 $2a$. 计算出曲线上的若干个点，列表如下：

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π
r	0	$\frac{2-\sqrt{3}}{2}a$	$\frac{2-\sqrt{2}}{2}a$	$\frac{1}{2}a$	a	$\frac{3}{2}a$	$\frac{2+\sqrt{2}}{2}a$	$\frac{2+\sqrt{3}}{2}a$	$2a$

描出这些点，连接成一条光滑曲线，然后利用它对极轴的对称性画出全部图形．这条曲线叫作心形线（如图 1-1-25）．

(2)玫瑰线．

下面画出三叶玫瑰线 $r=a\sin 3\theta(a>0)$ 的图形．

其表达式的右端是以 $\frac{2\pi}{3}$ 为周期的周期函数，作图时应先考虑 $0\leqslant\theta\leqslant\frac{2\pi}{3}$，然后仿照在这一范围内曲线上的点的变化规律，画出 $\frac{2\pi}{3}\leqslant\theta\leqslant\frac{4\pi}{3}$，$\frac{4\pi}{3}\leqslant\theta\leqslant 2\pi$ 等范围内的曲线．对于在 $0\leqslant\theta\leqslant\frac{2\pi}{3}$ 范围内的作图方法，仍是采用描点法．计算出曲线上若干个点：

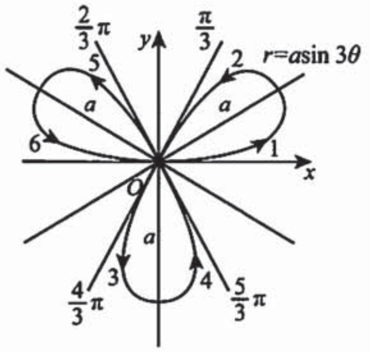


图 1-1-26

θ	0	$\frac{\pi}{12}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$	$\frac{7\pi}{12}$	$\frac{2\pi}{3}$
r	0	$\frac{\sqrt{2}}{2}a$	a	$\frac{\sqrt{2}}{2}a$	0	$-\frac{\sqrt{2}}{2}a$	$-a$	$-\frac{\sqrt{2}}{2}a$	0

描出这些点，连接成一条光滑曲线．这段曲线由弧段 1，2，3，4 构成（如图 1-1-26）．在 $\frac{2\pi}{3}\leqslant\theta\leqslant\frac{4\pi}{3}$ 的范围内，可以按照同样的规律画出由弧段 5，6，1，2 所构成的曲线，在 $\frac{4\pi}{3}\leqslant\theta\leqslant 2\pi$ 的范围内，同样可画出由弧段 3，4，5，6 所构成的曲线．这样就得到了曲线的全部．这曲线叫作三叶玫瑰线．

(3)阿基米德螺线．

下面画出 $r=a\theta(a>0,\theta\geqslant 0)$ 的图形．

当 $\theta(\theta\geqslant 0)$ 由 0 增大时，r 亦逐渐增大，这曲线称为阿基米德螺线（如图 1-1-27）．

(4)伯努利双纽线．

设定线段 AB 长度为 $2a$，动点 M 满足 $MA\cdot MB=a^2$，那么 M 的轨迹称为双纽线．

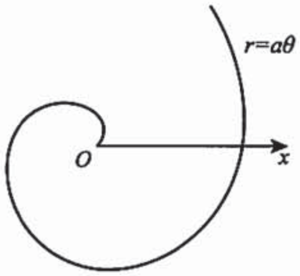


图 1-1-27

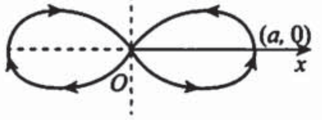
取 AB 为 x 轴，中点为原点，那么 A，B 的坐标分别为 $(-a,0)$，$(a,0)$．设 $M(x,y)$，则有

$$\sqrt{(x+a)^2+y^2}\cdot\sqrt{(x-a)^2+y^2}=a^2,$$

整理得 $(x^2+y^2)^2=2a^2(x^2-y^2)$．

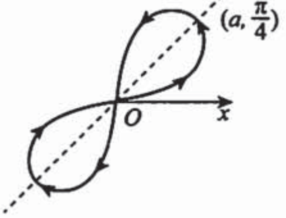
在极坐标系中，可化简得 $r^2=2a^2\cos 2\theta$．

在极坐标系中，双纽线的极坐标方程常常写成 $r^2=a^2\cos 2\theta$（如图 1-1-28），或 $r^2=a^2\sin 2\theta$（如图 1-1-29）．



$r^2=a^2\cos 2\theta$

图 1-1-28

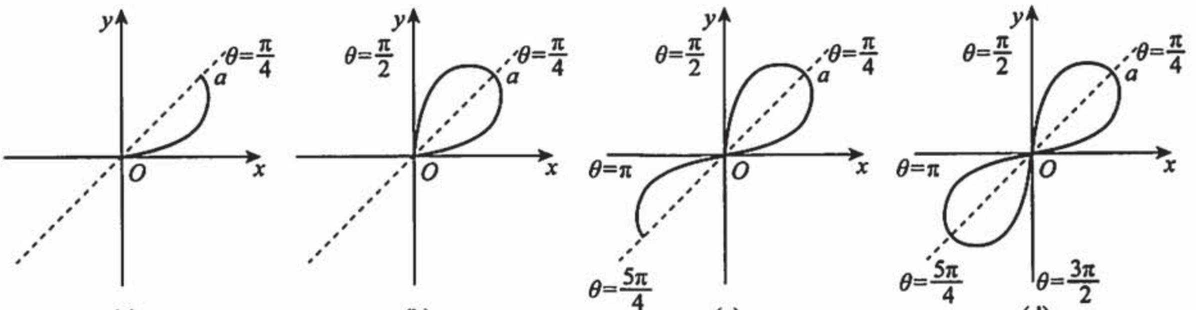


$r^2=a^2\sin 2\theta$

图 1-1-29

比如下面画出 $r^2=a^2\sin 2\theta(a>0)$ 的图像，由 $r=a\sqrt{\sin 2\theta}$，知 θ 的取值范围是 $\left[0,\frac{\pi}{2}\right]\cup\left[\pi,\frac{3\pi}{2}\right]$.

当 θ 从 0 增加到 $\frac{\pi}{4}$ 时，r 从 0 增加到 a. 故在图 1-1-30(a)中画出相应的部分；当 θ 从 $\frac{\pi}{4}$ 增加到 $\frac{\pi}{2}$ 时，r 从 a 减少到 0. 在图 1-1-30(b)中画出相应的部分；当 θ 从 π 增加到 $\frac{5\pi}{4}$ 时，r 从 0 增加到 a，在图 1-1-30(c)中画出相应的部分；当 θ 从 $\frac{5\pi}{4}$ 增加到 $\frac{3\pi}{2}$ 时，r 从 a 减少到 0，在图 1-1-30(d)中画出相应的部分. 最后形成一个“∞”形图像，这曲线称为伯努利双纽线.



(a)　(b)　(c)　(d)

图 1-1-30　伯努利双纽线的作图过程

2. 用直角系观点画极坐标系下图像

比较直角坐标方程 $y=x$，它表示平面上的一条直线，而极坐标方程 $r=\theta$ 表示螺线. 以方程的角度看问题，两个方程的形式相同，只是表示变量的字母不同而已，但是由于坐标系不同，它们表示的曲线完全不同(如图 1-1-31，图 1-1-32). 这里启发我们，若较易画出直角坐标系观点下 $r=f(\theta)$ 的图像，可转化为极坐标系下的曲线图像.

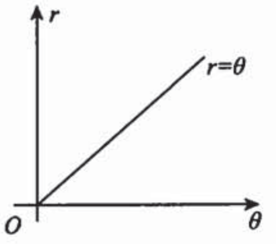


图 1-1-31

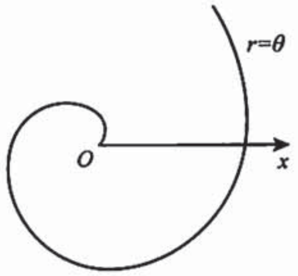


图 1-1-32

比如 $r=2(1+\cos\theta)$，较易画出其在直角坐标系观点下 $r=f(\theta)$ 的图像. 如图 1-1-33 所示，可转化为极坐标系下的图像，如图 1-1-34 所示，若读者掌握此种方法，不失为一个妙招.

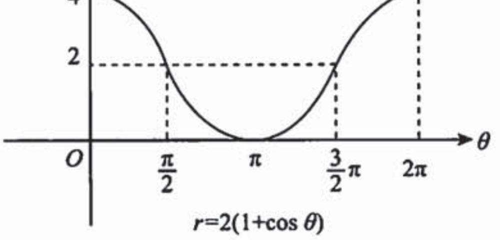


图 1-1-33

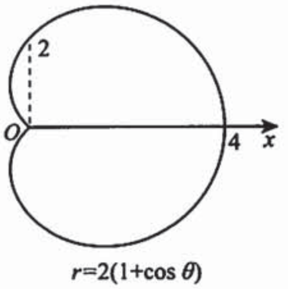


图 1-1-34

（三）参数法——参数方程$\left(\begin{cases}x=x(t)\\y=y(t)\end{cases}\right)$

前面的（一）与（二）介绍了如何在直角坐标系或极坐标系内用动点坐标(x,y)或(r,t)来表示平面内一些曲线的方程. 但在实际问题中，有些曲线用以上的方法来表示比较困难，也就是说很难找到曲线所满足的$f(x,y)=0$或$g(r,t)=0$的式子. 这个节目将引入一个新变量（叫作参数）来表示曲线方程，即参数方程.

（1）摆线.

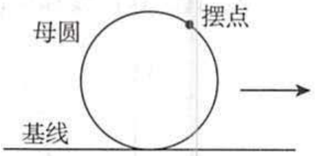


图 1-1-35

设自行车外胎上粘了一点红色的油漆，当你骑车向前直行时，这个油漆红点就在平面上形成一条轨迹，这轨迹就是摆线. 用数学语言描述如下：

当一个圆沿一条定直线作无滑动的滚动时，动圆圆周上一个定点的轨迹叫作摆线（如图 1-1-35）.

现取已给的这条定直线为x轴，其正方向就是圆滚动的方向，当这圆与直线在圆上的定点A相切时，就取这点为原点O. 取半径$\overline{CO}$旋转的角度t为参数. 对于所求运动轨迹上的任何一点$A(x,y)$，由图1-1-36容易看出

$$x=OP=|OQ|-|PQ|,$$
$$|OQ|=\text{圆弧}\overset{\frown}{QA}\text{的长度}=rt,$$
$$|PQ|=|AC'|\sin t=r\sin t,$$

故得

$$x=rt-r\sin t.$$

由图也容易看出

$$y=PA=|QC'|-|DC'|=r-r\cos t.$$

因此，所求定点A的运动轨迹的参数方程为

$$\begin{cases}x=r(t-\sin t),\\y=r(1-\cos t).\end{cases}\qquad (*)$$

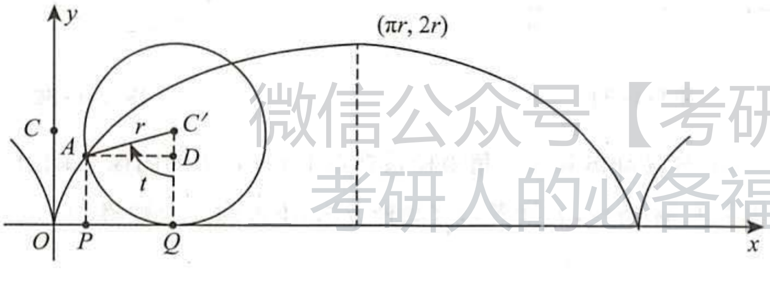


图 1-1-36

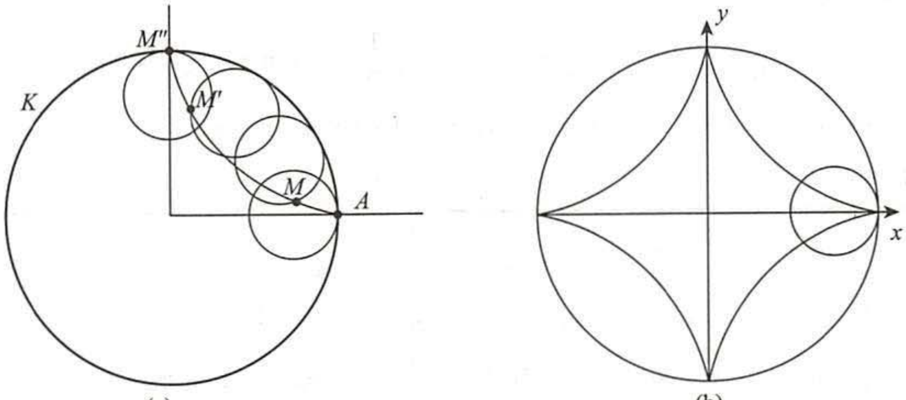
【注】 上面的推导过程只适用于$0\leqslant t<\frac{\pi}{2}$的情况，当$t$取其他任何值时，推导的方法是相仿的，所得结果与（＊）式完全一样. 因此（＊）式中没有写出t的变化范围，这就意味着t可取任何实数值.

摆线的图形具有周期性，当 t 增加 2π 时，也就是说，圆滚动一周时，摆线上的点的横坐标增加了 $2\pi r$，纵坐标不变. 圆继续滚动，圆上的定点 A 就描绘出一拱接一拱的图形. 容易看出，从原点开始的第一拱以直线 $x=\pi r$ 为对称轴，拱顶的坐标是 $(\pi r,2r)$.

要从(＊)式消去参数 t 是不困难的，但所得 x,y 间的函数表达式较复杂. 因此我们常通过(＊)式来直接研究摆线.

(2)星形线.

如图 1-1-37(a)所示，一个小圆 J 在一个固定的大圆 K 内部作纯滚动，如果大圆半径 r 是小圆半径的 4 倍，那么小圆圆周上任一点 M 的轨迹称为星形线，如图 1-1-37(b)所示.



(a) (b)

图 1-1-37

此轨迹方程的推导过程要用到较繁杂的几何知识与三角公式，不作要求，读者记住它的参数方程表达式即可. 则表达式为

$$\begin{cases}x=r\cos^3 t,\\ y=r\sin^3 t,\end{cases}$$

若消去 t，可得 $x^{\frac{2}{3}}+y^{\frac{2}{3}}=r^{\frac{2}{3}}$，此为直角坐标方程.

三、常用基础知识

1. 数列

(1)等差数列.

首项为 a_1，公差为 $d(d\neq0)$的数列 $a_1,a_1+d,a_1+2d,\cdots,a_1+(n-1)d,\cdots$.

①通项公式 $a_n=a_1+(n-1)d$.

②前 n 项的和 $S_n=\frac{n}{2}[2a_1+(n-1)d]=\frac{n}{2}(a_1+a_n)$.

(2)等比数列.

首项为 a_1，公比为 $r(r\neq0)$的数列 $a_1,a_1r,a_1r^2,\cdots,a_1r^{n-1},\cdots$.

①通项公式 $a_n=a_1r^{n-1}$.

②前 n 项的和 $S_n=\begin{cases}na_1, & r=1,\\ \frac{a_1(1-r^n)}{1-r}, & r\neq1.\end{cases}$

③常用 $1+r+r^2+\cdots+r^{n-1}=\dfrac{1-r^n}{1-r}(r\neq1)$.

(3)一些常见数列前 n 项的和.

① $\sum\limits_{k=1}^{n}k=1+2+3+\cdots+n=\dfrac{n(n+1)}{2}$.

② $\sum\limits_{k=1}^{n}k^2=1^2+2^2+3^2+\cdots+n^2=\dfrac{n(n+1)(2n+1)}{6}$.

③ $\sum\limits_{k=1}^{n}\dfrac{1}{k(k+1)}=\dfrac{1}{1\times2}+\dfrac{1}{2\times3}+\dfrac{1}{3\times4}+\cdots+\dfrac{1}{n(n+1)}=\dfrac{n}{n+1}$.

2. 三角函数

(1)三角函数基本关系.

$$\csc\alpha=\frac{1}{\sin\alpha},\quad \sec\alpha=\frac{1}{\cos\alpha},\quad \cot\alpha=\frac{1}{\tan\alpha},\quad \tan\alpha=\frac{\sin\alpha}{\cos\alpha},\quad \cot\alpha=\frac{\cos\alpha}{\sin\alpha},$$

$$\sin^2\alpha+\cos^2\alpha=1,\quad 1+\tan^2\alpha=\sec^2\alpha,\quad 1+\cot^2\alpha=\csc^2\alpha.$$

(2)诱导公式.

角 θ 函数	$\frac{\pi}{2}-\alpha$	$\frac{\pi}{2}+\alpha$	$\pi-\alpha$	$\pi+\alpha$	$\frac{3}{2}\pi-\alpha$	$\frac{3}{2}\pi+\alpha$	$2\pi-\alpha$
	$90°-\alpha$	$90°+\alpha$	$180°-\alpha$	$180°+\alpha$	$270°-\alpha$	$270°+\alpha$	$360°-\alpha$
$\sin\theta$	$\cos\alpha$	$\cos\alpha$	$\sin\alpha$	$-\sin\alpha$	$-\cos\alpha$	$-\cos\alpha$	$-\sin\alpha$
$\cos\theta$	$\sin\alpha$	$-\sin\alpha$	$-\cos\alpha$	$-\cos\alpha$	$-\sin\alpha$	$\sin\alpha$	$\cos\alpha$
$\tan\theta$	$\cot\alpha$	$-\cot\alpha$	$-\tan\alpha$	$\tan\alpha$	$\cot\alpha$	$-\cot\alpha$	$-\tan\alpha$
$\cot\theta$	$\tan\alpha$	$-\tan\alpha$	$-\cot\alpha$	$\cot\alpha$	$\tan\alpha$	$-\tan\alpha$	$-\cot\alpha$

【注】 如上表所示，奇变偶不变，符号看象限(因任一角度均可表示为 $\frac{k\pi}{2}+\alpha,k\in\mathbf{Z},|\alpha|<\frac{\pi}{4}$，故 k 为奇数时得角 α 的异名函数值，k 为偶数时得角 α 的同名函数值，然后在前面加上一个把角 α 看作锐角时原来函数值的符号).

三角函数在四个象限中的符号如下表所示.

角 θ 所在象限 函数	第一象限	第二象限	第三象限	第四象限
$\sin\theta$	+	+	−	−
$\cos\theta$	+	−	−	+
$\tan\theta$	+	−	+	−
$\cot\theta$	+	−	+	−

(3)特殊的三角函数值如下表所示.

α	0°	30°	45°	60°	90°	120°	135°	150°	180°	270°	360°
	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{3\pi}{4}$	$\frac{5\pi}{6}$	π	$\frac{3\pi}{2}$	2π
$\sin\alpha$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1	0
$\cos\alpha$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{3}}{2}$	-1	0	1
$\tan\alpha$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	∞	$-\sqrt{3}$	-1	$-\frac{\sqrt{3}}{3}$	0	∞	0
$\cot\alpha$	∞	$\sqrt{3}$	1	$\frac{\sqrt{3}}{3}$	0	$-\frac{\sqrt{3}}{3}$	-1	$-\sqrt{3}$	∞	0	∞

【注】 (1) $\sec\alpha$ 和 $\csc\alpha$ 的函数值可由 $\frac{1}{\cos\alpha}$ 和 $\frac{1}{\sin\alpha}$ 得出.

(2)表格中的“∞”均是指极限结果，如 $\tan 90^\circ$ 处的“∞”，是指 $\lim\limits_{x\to 90^\circ}\tan x=\infty$.

(4)重要公式.

①倍角公式.

$$\sin 2\alpha=2\sin\alpha\cos\alpha,\quad \cos 2\alpha=\cos^2\alpha-\sin^2\alpha=1-2\sin^2\alpha=2\cos^2\alpha-1,$$

$$\sin 3\alpha=-4\sin^3\alpha+3\sin\alpha,\quad \cos 3\alpha=4\cos^3\alpha-3\cos\alpha,$$

$$\tan 2\alpha=\frac{2\tan\alpha}{1-\tan^2\alpha},\quad \cot 2\alpha=\frac{\cot^2\alpha-1}{2\cot\alpha}.$$

②半角公式.

$$\sin^2\frac{\alpha}{2}=\frac{1}{2}(1-\cos\alpha),\quad \cos^2\frac{\alpha}{2}=\frac{1}{2}(1+\cos\alpha),\text{(降幂公式)}$$

$$\sin\frac{\alpha}{2}=\pm\sqrt{\frac{1-\cos\alpha}{2}},\quad \cos\frac{\alpha}{2}=\pm\sqrt{\frac{1+\cos\alpha}{2}},$$

$$\tan\frac{\alpha}{2}=\frac{1-\cos\alpha}{\sin\alpha}=\frac{\sin\alpha}{1+\cos\alpha}=\pm\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}},$$

$$\cot\frac{\alpha}{2}=\frac{\sin\alpha}{1-\cos\alpha}=\frac{1+\cos\alpha}{\sin\alpha}=\pm\sqrt{\frac{1+\cos\alpha}{1-\cos\alpha}}.$$

③和差公式.

$$\sin(\alpha\pm\beta)=\sin\alpha\cos\beta\pm\cos\alpha\sin\beta,\quad \cos(\alpha\pm\beta)=\cos\alpha\cos\beta\mp\sin\alpha\sin\beta,$$

$$\tan(\alpha\pm\beta)=\frac{\tan\alpha\pm\tan\beta}{1\mp\tan\alpha\tan\beta},\quad \cot(\alpha\pm\beta)=\frac{\cot\alpha\cot\beta\mp 1}{\cot\beta\pm\cot\alpha}.$$

④积化和差与和差化积公式.

(ⅰ)积化和差公式.

$$\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)],\quad \cos\alpha\sin\beta=\frac{1}{2}[\sin(\alpha+\beta)-\sin(\alpha-\beta)],$$

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)],\quad \sin\alpha\sin\beta=\frac{1}{2}[\cos(\alpha-\beta)-\cos(\alpha+\beta)].$$

(ⅱ)和差化积公式.

$$\sin\alpha+\sin\beta=2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2},\quad \sin\alpha-\sin\beta=2\sin\frac{\alpha-\beta}{2}\cos\frac{\alpha+\beta}{2},$$

$$\cos\alpha+\cos\beta=2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2},\quad \cos\alpha-\cos\beta=-2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}.$$

⑤万能公式.

若 $u=\tan\frac{x}{2}(-\pi<x<\pi)$，则 $\sin x=\frac{2u}{1+u^2}$，$\cos x=\frac{1-u^2}{1+u^2}$.

3. 指数运算法则

$$a^\alpha\cdot a^\beta=a^{\alpha+\beta},\quad \frac{a^\alpha}{a^\beta}=a^{\alpha-\beta},\quad (a^\alpha)^\beta=a^{\alpha\beta},\quad (ab)^\alpha=a^\alpha b^\alpha,\quad \left(\frac{a}{b}\right)^\alpha=\frac{a^\alpha}{b^\alpha},$$

其中 a,b 是正实数，α,β 是任意实数.

4. 对数运算法则

①$\log_a(MN)=\log_aM+\log_aN$（积的对数＝对数的和）.

②$\log_a\frac{M}{N}=\log_aM-\log_aN$（商的对数＝对数的差）.

③$\log_aM^n=n\log_aM$（幂的对数＝对数的倍数）.

④$\log_a\sqrt[n]{M}=\frac{1}{n}\log_aM$.

5. 一元二次方程基础

①一元二次方程 $ax^2+bx+c=0(a\neq0)$.

②根的公式 $x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$.

③根与系数的关系（韦达定理）$x_1+x_2=-\frac{b}{a}$，$x_1x_2=\frac{c}{a}$.

④判别式 $\Delta=b^2-4ac$.

$\Delta>0$，方程有两个不等的实根；$\Delta=0$，方程有两个相等的实根；$\Delta<0$，方程有两个共轭的复根.

⑤抛物线 $y=ax^2+bx+c$ 的顶点 $\left(-\frac{b}{2a},c-\frac{b^2}{4a}\right)$.

6. 因式分解公式

①$(a+b)^2=a^2+2ab+b^2$.

②$(a-b)^2=a^2-2ab+b^2$.

③$(a+b)^3=a^3+3a^2b+3ab^2+b^3$.

④$(a-b)^3=a^3-3a^2b+3ab^2-b^3$.

⑤$a^2-b^2=(a+b)(a-b)$.

⑥$a^3-b^3=(a-b)(a^2+ab+b^2)$.

⑦$a^3+b^3=(a+b)(a^2-ab+b^2)$.

⑧$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1})$（$n$ 是正整数）.

⑨n 是正偶数时，$a^n-b^n=(a+b)(a^{n-1}-a^{n-2}b+\cdots+ab^{n-2}-b^{n-1})$.

⑩n 是正奇数时，$a^n+b^n=(a+b)(a^{n-1}-a^{n-2}b+\cdots-ab^{n-2}+b^{n-1})$.

⑪二项式定理

$$(a+b)^n=\sum_{k=0}^{n}C_n^ka^{n-k}b^k=a^n+na^{n-1}b+\frac{n(n-1)}{2!}a^{n-2}b^2+\cdots+$$

$$\frac{n(n-1)\cdots(n-k+1)}{k!}a^{n-k}b^k+\cdots+nab^{n-1}+b^n.$$

7. 阶乘与双阶乘

①$n!=1\cdot2\cdot3\cdot\cdots\cdot n$，规定 $0!=1$.

②$(2n)!!=2\cdot4\cdot6\cdot\cdots\cdot(2n)=2^n\cdot n!$.

③$(2n-1)!!=1\cdot3\cdot5\cdot\cdots\cdot(2n-1)$.

8. 常用不等式

(1)设 a,b 为实数,则① $|a\pm b|\leqslant|a|+|b|$;② $\big||a|-|b|\big|\leqslant|a-b|$.

【注】 可以将上述不等式①推广为

离散情况:设 $a_1,a_2,\cdots,a_n$ 为实数,则

$$|a_1\pm a_2\pm\cdots\pm a_n|\leqslant|a_1|+|a_2|+\cdots+|a_n|.$$

连续情况:设 $f(x)$ 在 $[a,b](a<b)$ 上可积,则

$$\left|\int_a^b f(x)\mathrm{d}x\right|\leqslant\int_a^b|f(x)|\mathrm{d}x.$$

(2)① $\sqrt{ab}\leqslant\frac{a+b}{2}\leqslant\sqrt{\frac{a^2+b^2}{2}}(a,b>0)$;

② $\sqrt[3]{abc}\leqslant\frac{a+b+c}{3}\leqslant\sqrt{\frac{a^2+b^2+c^2}{3}}(a,b,c>0)$.

(3)设 $a>b>0$,则 $\begin{cases}\text{当 } n>0 \text{ 时}, a^n>b^n,\\ \text{当 } n<0 \text{ 时}, a^n<b^n.\end{cases}$

(4)若 $0<a<x<b,0<c<y<d$,则 $\frac{c}{b}<\frac{y}{x}<\frac{d}{a}$.

(5) $\sin x<x<\tan x\left(0<x<\frac{\pi}{2}\right)$.

(6) $\sin x<x(x>0)$.

(7) $\arctan x\leqslant x\leqslant\arcsin x(0\leqslant x\leqslant1)$.

(8) $\mathrm{e}^x\geqslant x+1(\forall x)$.

(9) $x-1\geqslant\ln x(x>0)$.

(10) $\frac{1}{1+x}<\ln\left(1+\frac{1}{x}\right)<\frac{1}{x}(x>0)$.

【注】 **证明** 令 $f(x)=\ln x$,并在区间 $[x,x+1]$ 上对其应用拉格朗日中值定理,有

$$\ln\left(1+\frac{1}{x}\right)=\ln(1+x)-\ln x=\frac{1}{\xi},$$

其中 $0<x<\xi<x+1$. 因此,对任意的 $x>0$,有 $\frac{1}{1+x}<\ln\left(1+\frac{1}{x}\right)=\frac{1}{\xi}<\frac{1}{x}$.

基础例题精解

例 1.1.1 设 $f(x)=x^2,f[\varphi(x)]=-x^2+2x+3$,且 $\varphi(x)\geqslant0$,求 $\varphi(x)$ 及其定义域与值域.

解 由题设条件知,$f[\varphi(x)]=\varphi^2(x)=-x^2+2x+3$,于是 $\varphi(x)=\sqrt{-x^2+2x+3}$.

由 $-x^2+2x+3\geqslant0$,即 $(x-3)(x+1)\leqslant0$,知 $\varphi(x)$ 的定义域为 $[-1,3]$.

又 $\sqrt{-x^2+2x+3}=\sqrt{-(x-1)^2+4}$，当 $x=1$ 时，$\varphi(x)=2$ 为最大值，显然 $\varphi(-1)=\varphi(3)=0$ 为最小值，故 $\varphi(x)$ 的值域为 $[0,2]$.

例 1.1.2 求函数 $y=f(x)=\ln(x+\sqrt{x^2+1})$ 的反函数 $f^{-1}(x)$ 的表达式及其定义域.

解 直接由 $y=\ln(x+\sqrt{x^2+1})$ 解出 $x=f^{-1}(y)$ 会很麻烦，现采用下述方法.

$$-y=-\ln(x+\sqrt{x^2+1})=\ln\frac{1}{x+\sqrt{x^2+1}}$$

$$=\ln\frac{\sqrt{x^2+1}-x}{(\sqrt{x^2+1}+x)(\sqrt{x^2+1}-x)}=\ln(\sqrt{x^2+1}-x),$$

所以 $$e^{-y}=\sqrt{x^2+1}-x,\qquad ①$$

再由 $y=f(x)$ 的表达式有 $$e^{y}=\sqrt{x^2+1}+x,\qquad ②$$

②－①，得 $$x=\frac{1}{2}(e^{y}-e^{-y}),$$

交换上式中 x,y 的位置后就是 $y=f(x)$ 的反函数，即

$$y=f^{-1}(x)=\frac{1}{2}(e^{x}-e^{-x}),-\infty<x<+\infty.$$

【注】 函数 $y=\ln(x+\sqrt{x^2+1})$ 的图像如图 1-1-38 所示. 函数 $y=\frac{e^{x}-e^{-x}}{2}$ 的图像如图 1-1-39 所示.

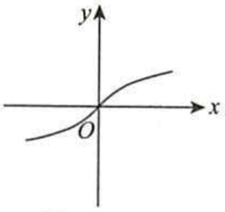


图 1-1-38

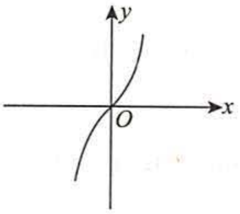


图 1-1-39

它们都是奇函数，其图像也应记住.

例 1.1.3 将下列各组函数 $f(x)$ 与 $\varphi(x)$ 复合，求复合函数 $f[\varphi(x)]$.

(1) $f(x)=\begin{cases}2-x, & x\leqslant 0,\\ x+2, & x>0,\end{cases}\quad \varphi(x)=\begin{cases}x^2, & x<0,\\ -x, & x\geqslant 0.\end{cases}$

(2) $f(x)=\varphi(x)=\begin{cases}x, & x\geqslant 0,\\ \dfrac{1}{x}, & x<0.\end{cases}$

解 (1) $f[\varphi(x)]=\begin{cases}2-\varphi(x), & \varphi(x)\leqslant 0,\\ \varphi(x)+2, & \varphi(x)>0.\end{cases}$

又当 $\varphi(x)\leqslant 0$ 时，$x\geqslant 0$，且 $\varphi(x)=-x$；当 $\varphi(x)>0$ 时，$x<0$，且 $\varphi(x)=x^2$，于是

$$f[\varphi(x)]=\begin{cases}2-(-x), & x\geqslant 0,\\ 2+x^2, & x<0\end{cases}=\begin{cases}2+x, & x\geqslant 0,\\ 2+x^2, & x<0.\end{cases}$$

(2) $f[\varphi(x)]=f[f(x)]=\begin{cases}f(x), & f(x)\geqslant 0,\\ \dfrac{1}{f(x)}, & f(x)<0.\end{cases}$

又当 $f(x)\geqslant 0$ 时，$x\geqslant 0$，且 $f(x)=x$；当 $f(x)<0$ 时，$x<0$，且 $f(x)=\frac{1}{x}$，于是

$$f[\varphi(x)]=f[f(x)]=\begin{cases}x, & x\geqslant 0,\\ \dfrac{1}{\frac{1}{x}}, & x<0\end{cases}=\begin{cases}x, & x\geqslant 0,\\ x, & x<0\end{cases}=x(-\infty<x<+\infty).$$

例 1.1.4　证明函数 $f(x)=\frac{x}{1+x^2}$ 在 $(-\infty,+\infty)$ 内有界.

证明　当 $x\neq 0$ 时，
$$|f(x)|=\frac{|x|}{1+x^2}=\frac{1}{\frac{1}{|x|}+|x|},$$

由不等式 $\frac{a+b}{2}\geqslant\sqrt{ab}(a,b>0)$，有 $\frac{1}{|x|}+|x|\geqslant 2\sqrt{\frac{1}{|x|}|x|}=2$，即 $|f(x)|\leqslant\frac{1}{2}$.

当 $x=0$ 时，$f(0)=0$. 综上，函数 $f(x)$ 在 $(-\infty,+\infty)$ 内有界.

基础习题精练

习题

1.1.1　设 $f(x)=\begin{cases}\ln\sqrt{x}, & x\geqslant 1,\\ 2x-1, & x<1.\end{cases}$ 求 $f[f(x)]$.

1.1.2　设 $f(x)=\frac{x}{x-1}$，试验证 $f\{f[f(f(x))]\}=x$，并求 $f\left[\frac{1}{f(x)}\right]$，这里 $x\neq 0,x\neq 1$.

1.1.3　求函数 $y=2x+|2-x|$，$x\in(-\infty,+\infty)$ 的反函数.

1.1.4　画出 $r=a\mathrm{e}^{k\theta}(a>0,k>0)$ 的图形.（这里 e 是以后常要用到的一个常数，以 e 为底的对数称为自然对数，e≈2.71828.）

解答

1.1.1　解　$f[f(x)]=\begin{cases}\ln\sqrt{f(x)}, & f(x)\geqslant 1,\\ 2f(x)-1, & f(x)<1.\end{cases}$

由题设可得：①当 $f(x)\geqslant 1$ 时，$x\geqslant\mathrm{e}^2$，此时 $f(x)=\ln\sqrt{x}$；

②当 $f(x)<1$ 时，或者 $1\leqslant x<\mathrm{e}^2$，此时 $f(x)=\ln\sqrt{x}$，或者 $x<1$，此时 $f(x)=2x-1$.

综合以上各种情况有

$$f[f(x)]=\begin{cases}\ln\sqrt{\ln\sqrt{x}}, & x\geqslant\mathrm{e}^2,\\ 2\ln\sqrt{x}-1, & 1\leqslant x<\mathrm{e}^2,\\ 2(2x-1)-1, & x<1\end{cases}=\begin{cases}\frac{1}{2}\ln(\ln\sqrt{x}), & x\geqslant\mathrm{e}^2,\\ \ln x-1, & 1\leqslant x<\mathrm{e}^2,\\ 4x-3, & x<1.\end{cases}$$

1.1.2 解 令 $f_1(x)=f(x), f_2(x)=f[f_1(x)],\cdots,f_n(x)=f[f_{n-1}(x)]$. 现只需验证 $f_4(x)=x$ 即可.

因 $f(x)=\dfrac{x}{x-1}=\dfrac{1}{1-\dfrac{1}{x}}$，故 $\dfrac{1}{f(x)}=1-\dfrac{1}{x}$.

而
$$f_2(x)=f[f(x)]=\frac{1}{1-\dfrac{1}{f(x)}}=\frac{1}{1-\left(1-\dfrac{1}{x}\right)}=x.$$

又 $f_3(x)=f[f_2(x)]=f(x)$，则 $f_4(x)=f[f_3(x)]=f[f(x)]=x$.

下面求 $f\left[\dfrac{1}{f(x)}\right]$. 由 $\dfrac{1}{f(x)}=1-\dfrac{1}{x}$，则
$$f\left[\frac{1}{f(x)}\right]=f\left(1-\frac{1}{x}\right)=1-x, x\neq 0, x\neq 1.$$

1.1.3 解 为了去掉绝对值，将方程改写为
$$y=\begin{cases}x+2, & x\leqslant 2,\\ 3x-2, & x>2.\end{cases}$$

当 $x\leqslant 2$ 时，$y=x+2\Rightarrow x=y-2, y\leqslant 4$；当 $x>2$ 时，$y=3x-2\Rightarrow x=\dfrac{y+2}{3}, y>4$.

综上所述即得
$$x=\begin{cases}y-2, & y\leqslant 4,\\ \dfrac{y+2}{3}, & y>4\end{cases}\xrightarrow{x,y\text{互换}} y=\begin{cases}x-2, & x\leqslant 4,\\ \dfrac{x+2}{3}, & x>4.\end{cases}$$

于是，若记 $f(x)=2x+|2-x|, x\in(-\infty,+\infty)$，则有 $f^{-1}(x)=\begin{cases}x-2, & x\leqslant 4,\\ \dfrac{x+2}{3}, & x>4.\end{cases}$

1.1.4 解 当 θ 由 0 逐渐增大时，r 由 a 逐渐增大，当 θ 无限制地增大，r 也随之无限制地增大. 当 θ 由 0 逐渐减小，r 由 a 逐渐减小，但 r 却永远保持正值，当 θ 由 0 无限制地减小，r 也随之无限制地接近于 0. 如图 1-1-40 所示，是取 $a=2, k=0.2$ 时画出来的. 这曲线叫作对数螺线(如图 1-1-40).

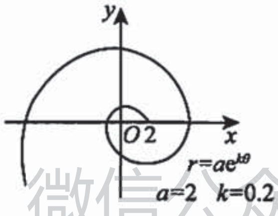
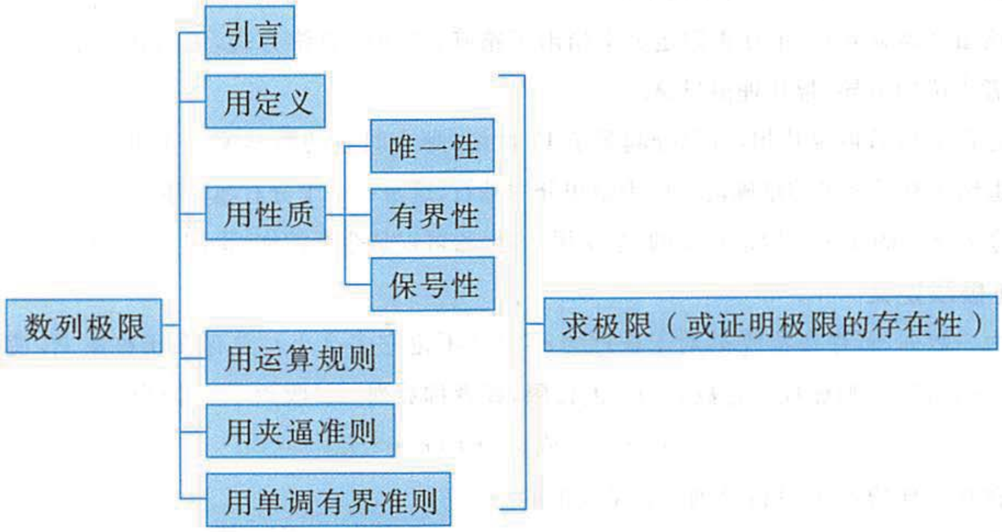


图 1-1-40

第2讲 数列极限

基础知识结构



基础内容精讲

1. 引言

极限，从通俗、直观的意义上讲，是一个“**无限趋近的过程**”．

给一个简单的数列$\left\{\frac{n}{n+1}\right\}$，若从第1项开始一直写下去，那就是$\frac{1}{2},\frac{2}{3},\frac{3}{4},\cdots,\frac{n}{n+1},\cdots$，不难发现，随着$n$无限增大，$n$与$n+1$的比值无限趋近于1，这就写成了

$$\lim_{n\to\infty}\frac{n}{n+1}=1.$$

但若从专业角度讲，可没这么简单，不妨通过下面这段两个学生之间的对话来琢磨．

学生甲：给出一个数列$u_1,u_2,\cdots,u_n,\cdots$，其通项$u_n=\frac{n}{n+1}$，当$n$无限增大时，这个数列趋向于一个值1，这叫作极限值，可写成

$$\lim_{n\to\infty}\frac{n}{n+1}=1.$$

学生乙：嗯，当n充分大时，我可以理解$\frac{n}{n+1}$与1会非常非常接近，此数列应该会趋向于常数1，但数学是讲究精确的学科，你能给我指出一个数$N>0$，使得这个数列中下标大于N的项与你说的极限值1之间的距离最终保持在$(0,\varepsilon)$之内吗？

学生甲：好，我来尝试一下，这个数列的第 n 项与数 1 之间的距离是 $\left|\frac{n}{n+1}-1\right|=\frac{1}{n+1}$，让此距离最终保持在 $(0,\varepsilon)$ 之内，也就是 $\frac{1}{n+1}<\varepsilon$，反解出 n 来，即 $n>\frac{1}{\varepsilon}-1$，这样一来，随便你给我什么正数 ε，只要我从数列的第 N 项开始，其中 $N>\frac{1}{\varepsilon}-1$，在此之后的所有项都会有 $\left|\frac{n}{n+1}-1\right|<\varepsilon$ 成立，即保持与 1 之间的距离小于 ε.

学生乙：我明白了，数列 $\left\{\frac{n}{n+1}\right\}$ 一定会出现这样的局面——不论我指定多么小的正数 ε，你总可以找到某项，从此项开始往后，所有项都与 1"亲密无间"地保持着距离始终小于那个要多小有多小的尺度 ε. 于是，我们可以理直气壮地认为数列 $\left\{\frac{n}{n+1}\right\}$ 在某项之后的所有项会**更接近**数"1"，而不是除了 1 之外的任何其他数. 因此必须承认，此数列趋向于极限值 1.

这段对话虽然略显冗长，但在极限定义上给出了精确、本质的思维方式，能给在此定义上偶尔陷入思维混乱的读者正确的引导，帮其理清思路.

我想在这部分最后再说几句，微积分起源于 17 世纪，那个时候的数学家是不可能讲出上面这段对话的，直到 19 世纪才有了极限的精确定义作为微积分的基石，这是一个十分有趣的事实：历史上微积分知识的出现与我们今天学习的次序是恰恰相反的. 学生甲、学生乙可以是今天的你、我、他，但不是牛顿、莱布尼茨.

2. 数列极限定义

设 $\{x_n\}$ 为一数列，若存在常数 a，对于任意的 $\varepsilon>0$（不论它多么小），总存在正整数 N，使得当 $n>N$ 时，$|x_n-a|<\varepsilon$ 恒成立，则称数 a 是数列 $\{x_n\}$ 的极限，或者称数列 $\{x_n\}$ 收敛于 a，记为

$$\lim_{n\to\infty}x_n=a \text{ 或 } x_n\to a\ (n\to\infty).$$

如果不存在这样的数 a，就说数列 $\{x_n\}$ 是发散的.

常用的语言：$\lim\limits_{n\to\infty}x_n=a\Leftrightarrow\forall\varepsilon>0,\ \exists N\in\mathbf{N}_+$，当 $n>N$ 时，恒有 $|x_n-a|<\varepsilon$.

【注 1】 这是用"$\varepsilon-N$ 语言"来描述数列极限. 符号"$\forall$"是英文 Arbitrary（任意的）的首字母上下方向倒着写出来的；符号"$\exists$"是英文 Exist（存在）的首字母左右方向倒着写出来的.

【注 2】 数列收敛与其子列收敛的关系.

定义 从数列 $\{a_n\}$：$a_1,a_2,\cdots,a_n,\cdots$ 中选取无穷多项，并按原来的先后顺序组成新的数列，称新数列为原数列的**子列**，记为

$$\{a_{n_k}\}:a_{n_1},a_{n_2},\cdots,a_{n_k},\cdots,$$

其中下标 $n_1,n_2,\cdots,n_k,\cdots$ 为正整数.

例如，若 $n_k(k=1,2,\cdots)$ 分别取为 $2k$ 和 $2k-1$，则得到数列 $\{a_n\}$ 的两个子列

$$\{a_{2k}\}:a_2,a_4,\cdots,a_{2k},\cdots;\{a_{2k-1}\}:a_1,a_3,\cdots,a_{2k-1},\cdots,$$

这两个子列的项在原数列中交错出现.

定理 若数列 $\{a_n\}$ 收敛，则其任何子列 $\{a_{n_k}\}$ 也收敛，且 $\lim\limits_{k\to\infty}a_{n_k}=\lim\limits_{n\to\infty}a_n$.

此定理为我们提供了一个判断数列发散的方法：对于一个数列 $\{a_n\}$，如果能找到一个发散的子列，则原数列一定发散；如果能找到两个收敛的子列 $\{a_{n_k}\}$ 和 $\{a_{n_k'}\}$，但它们收敛到不同极限，则原数列也一定发散.

例如,对于数列$\{(-1)^n\}$:$-1,1,-1,1,\cdots,(-1)^n,\cdots$,我们找到其收敛的子列

$$\{(-1)^{2k}\}:1,1,\cdots,1,\cdots;\{(-1)^{2k-1}\}:-1,-1,\cdots,-1,\cdots,$$

它们的极限分别为 1 和-1,所以原数列发散.

由定理可得一个重要推论:$\lim\limits_{n\to\infty}a_n=a\Leftrightarrow\lim\limits_{k\to\infty}a_{2k}=a$,且$\lim\limits_{k\to\infty}a_{2k-1}=a$.

3. 收敛数列的性质

定理 1(唯一性) 给出数列$\{x_n\}$,若$\lim\limits_{n\to\infty}x_n=a$(存在),则 a 是唯一的.

定理 2(有界性) 若数列$\{x_n\}$极限存在,则数列$\{x_n\}$有界.

定理 3(保号性) 设数列$\{a_n\}$存在极限 a,且 $a>0$(或 $a<0$),则存在正整数 N,当 $n>N$ 时,有 $a_n>0$(或 $a_n<0$).

推论 如果数列$\{a_n\}$从某项起有 $a_n\geqslant0$,且$\lim\limits_{n\to\infty}a_n=a$,则 $a\geqslant0$.

4. 极限运算规则

设$\lim\limits_{n\to\infty}x_n=a$,$\lim\limits_{n\to\infty}y_n=b$,则

①$\lim\limits_{n\to\infty}(x_n\pm y_n)=a\pm b$;

②$\lim\limits_{n\to\infty}x_ny_n=ab$;

③若 $b\neq0$,$y_n\neq0$,则$\lim\limits_{n\to\infty}\dfrac{x_n}{y_n}=\dfrac{a}{b}$.

运算规则可以推广至有限个数列情形.

5. 夹逼准则

如果数列$\{x_n\}$,$\{y_n\}$及$\{z_n\}$满足下列条件

①$y_n\leqslant x_n\leqslant z_n(n=1,2,3,\cdots)$;②$\lim\limits_{n\to\infty}y_n=a$,$\lim\limits_{n\to\infty}z_n=a$.

则数列$\{x_n\}$的极限存在,且$\lim\limits_{n\to\infty}x_n=a$.

6. 单调有界准则

单调有界数列必有极限,即若数列$\{x_n\}$单调增加(减少)且有上界(下界),则$\lim\limits_{n\to\infty}x_n$存在.

基础例题精解

例 1.2.1 用定义证明$\lim\limits_{n\to\infty}\left[1+\dfrac{(-1)^n}{n}\right]=1$.

证明 按照前面引言的思路,不难写出证明过程.因$\left|1+\dfrac{(-1)^n}{n}-1\right|=\dfrac{1}{n}$,则对任意正数 ε,欲使

$$\left|1+\frac{(-1)^n}{n}-1\right|<\varepsilon,$$

即

$$\frac{1}{n}<\varepsilon,$$

只要

$$n>\frac{1}{\varepsilon}.$$

取 N 为$\dfrac{1}{\varepsilon}$的整数部分加 1,即 $N=\left[\dfrac{1}{\varepsilon}\right]+1$,则当 $n>N$ 时,就有 $n>\dfrac{1}{\varepsilon}$,从而有

$$\left|1+\frac{(-1)^n}{n}-1\right|<\varepsilon.$$

故$\lim\limits_{n\to\infty}\left[1+\frac{(-1)^n}{n}\right]=1$.

例 1.2.2 证明：若$\lim\limits_{n\to\infty}a_n=A$，则$\lim\limits_{n\to\infty}|a_n|=|A|$.

证明 因$\lim\limits_{n\to\infty}a_n=A$，对任意正数 ε，存在正整数 N，当 $n>N$ 时，有

$$|a_n-A|<\varepsilon.$$

又由不等式$||a|-|b||\leqslant|a-b|$，有

$$||a_n|-|A||\leqslant|a_n-A|<\varepsilon.$$

故$\lim\limits_{n\to\infty}|a_n|=|A|$.

> **【注】** 在本题中若 $A=0$，则$||a_n|-|A||=||a_n|-0|=|a_n-0|$，即有
>
> $$\lim_{n\to\infty}a_n=0\Leftrightarrow\lim_{n\to\infty}|a_n|=0,$$
>
> 这结论常用，即若要证$\lim\limits_{n\to\infty}a_n=0$，可转化为证$\lim\limits_{n\to\infty}|a_n|=0$，由于$|a_n|\geqslant0$，若使用夹逼准则，便省了一半的气力，只需证$|a_n|\leqslant0$ 即可.

例 1.2.3 证明数列$\{n^{(-1)^n}\}$极限不存在.

证明 从数列

$$\{n^{(-1)^n}\}:\frac{1}{1},2,\frac{1}{3},4,\frac{1}{5},6,\cdots,\frac{1}{2n-1},2n,\cdots$$

中选取一个子列$\{2n\}$：$2,4,\cdots,2n,\cdots$，该数列不是有界数列，由定理 2 的逆否命题知该子列发散，因此，由"收敛数列的任何子列也收敛"的逆否命题知，原数列极限不存在.

> **【注】** 该数列存在收敛的子列$\left\{\frac{1}{2n-1}\right\}$：$1,\frac{1}{3},\frac{1}{5},\cdots,\frac{1}{2n-1},\cdots$，但原数列发散. 这说明一个数列的某个子列收敛并不能保证原数列收敛.

例 1.2.4 设$\lim\limits_{n\to\infty}(a_n+b_n)=1$，$\lim\limits_{n\to\infty}(a_n-b_n)=3$，证明$\{a_n\}$和$\{b_n\}$的极限存在，并求出它们的极限值.

证明 令 $u_n=a_n+b_n$，$v_n=a_n-b_n$，则$\lim\limits_{n\to\infty}u_n=1$，$\lim\limits_{n\to\infty}v_n=3$. 由极限运算规则①知$\{u_n+v_n\}$和$\{u_n-v_n\}$均存在极限，且有

$$\lim_{n\to\infty}(u_n+v_n)=\lim_{n\to\infty}u_n+\lim_{n\to\infty}v_n=1+3=4,$$

$$\lim_{n\to\infty}(u_n-v_n)=\lim_{n\to\infty}u_n-\lim_{n\to\infty}v_n=1-3=-2.$$

另一方面，$a_n=\frac{1}{2}(u_n+v_n)$，$b_n=\frac{1}{2}(u_n-v_n)$，所以$\{a_n\}$和$\{b_n\}$的极限存在，且有

$$\lim_{n\to\infty}a_n=\frac{1}{2}\times4=2,\quad \lim_{n\to\infty}b_n=\frac{1}{2}\times(-2)=-1.$$

例 1.2.5 求极限$\lim\limits_{n\to\infty}\left(\frac{n}{n^2+1}+\frac{n}{n^2+2}+\cdots+\frac{n}{n^2+n}\right)$.

解 $\lim\limits_{n\to\infty}\left(\frac{n}{n^2+1}+\frac{n}{n^2+2}+\cdots+\frac{n}{n^2+n}\right)=\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\frac{n}{n^2+i}$，对和式$\sum\limits_{i=1}^{n}\frac{n}{n^2+i}$作适当的放缩有

$$n\cdot\frac{n}{n^2+n}\leqslant\sum_{i=1}^{n}\frac{n}{n^2+i}\leqslant n\cdot\frac{n}{n^2+1},$$

由于$\lim\limits_{n\to\infty}\frac{n^2}{n^2+n}=1$，$\lim\limits_{n\to\infty}\frac{n^2}{n^2+1}=1$，根据夹逼准则，原式$=1$.

例 1.2.6　设数列$\{a_n\}$满足$a_1=a(a>0)$，$a_{n+1}=\frac{1}{2}\left(a_n+\frac{2}{a_n}\right)$，证明极限$\lim\limits_{n\to\infty}a_n$存在，并求其值.

证明　由题设可知$a_n>0$，由不等式$\frac{a+b}{2}\geqslant\sqrt{ab}(a,b>0)$，有

$$a_{n+1}=\frac{1}{2}\left(a_n+\frac{2}{a_n}\right)\geqslant\sqrt{a_n\cdot\frac{2}{a_n}}=\sqrt{2},$$

即数列$\{a_n\}$有下界. 又

$$a_{n+1}-a_n=\frac{1}{2}\left(a_n+\frac{2}{a_n}\right)-a_n=\frac{2-a_n^2}{2a_n}\leqslant 0(n\geqslant 2),$$

故数列$\{a_n\}$单调减少.

由单调有界准则，极限$\lim\limits_{n\to\infty}a_n\xlongequal[记为]{存在}A$，由保号性，有$A\geqslant\sqrt{2}$.

在等式$a_{n+1}=\frac{1}{2}\left(a_n+\frac{2}{a_n}\right)$两端取极限，得

$$A=\frac{1}{2}\left(A+\frac{2}{A}\right),$$

解得$A=\sqrt{2}$，即$\lim\limits_{n\to\infty}a_n=\sqrt{2}$.

例 1.2.7　设数列$\{x_n\}$满足$0<x_1<\pi$，$x_{n+1}=\sin x_n(n=1,2,\cdots)$. 证明$\lim\limits_{n\to\infty}x_n$存在，并求该极限.

证明　由于当$0<x<\pi$时，$0<\sin x<x$，所以当$0<x_n<\pi$时，$0<x_{n+1}=\sin x_n<x_n<\pi$. 已知$0<x_1<\pi$，故由数学归纳法知对一切$n=1,2,\cdots$，有

$$0<x_{n+1}=\sin x_n<x_n,$$

即$\{x_n\}$单调减少且$x_n>0$.

由单调有界准则知$\lim\limits_{n\to\infty}x_n$存在，记为$a$，则$a\geqslant 0$. 令$n\to\infty$，将$x_{n+1}=\sin x_n$两边取极限，得$a=\sin a$，易见$a=0$是它的解. 故$\lim\limits_{n\to\infty}x_n=0$.

基础习题精练

习题

1.2.1　用定义证明$\lim\limits_{n\to\infty}q^n=0$($q$为常数且$|q|<1$).

1.2.2　求$\lim\limits_{n\to\infty}\left(\frac{1}{\sqrt{n^2+1}}+\frac{1}{\sqrt{n^2+2}}+\cdots+\frac{1}{\sqrt{n^2+n}}\right)$.

1.2.3　求极限$\lim\limits_{n\to\infty}\left(\frac{1}{n^2+n+1}+\frac{2}{n^2+n+2}+\cdots+\frac{n}{n^2+n+n}\right)$.

1.2.4　设$a_n=\frac{1}{1^2}+\frac{1}{2^2}+\cdots+\frac{1}{n^2}(n=1,2,\cdots)$. 证明数列$\{a_n\}$收敛.

1.2.5 设 $x_1=2, x_n+(x_n-4)x_{n-1}=3(n=2,3,\cdots)$，试求 $\lim\limits_{n\to\infty}x_n$.

1.2.6 设数列 $\{a_n\}$ 满足 $a_0=0, a_1=1, 2a_{n+1}=a_n+a_{n-1}, n=1,2,\cdots$.

(1)证明 $a_{n+1}-a_n=\left(-\frac{1}{2}\right)^n, n=1,2,\cdots$;

(2)求 $\lim\limits_{n\to\infty}a_n$.

解答

1.2.1 证明 对任意正数 ε，不妨设 ε<1，欲使 $|q^n|<\varepsilon$ 成立，只要

$$n\ln|q|<\ln\varepsilon,$$

即

$$n>\frac{\ln\varepsilon}{\ln|q|}.$$

取 $N=\left[\frac{\ln\varepsilon}{\ln|q|}\right]+1$，则当 $n>N$ 时，就有 $n>\frac{\ln\varepsilon}{\ln|q|}$，则

$$|q^n|<\varepsilon.$$

故 $\lim\limits_{n\to\infty}q^n=0$（$q$ 为常数且 $|q|<1$）.

【注】 (1)当 $\{a_n\}$ 是首项为 a_1，公比为 $q(q\neq1)$ 的等比数列时，其前 n 项和为

$$S_n=a_1\frac{1-q^n}{1-q}.$$

当 $|q|<1$ 时，$S=\lim\limits_{n\to\infty}S_n=\frac{a_1}{1-q}$.

(2)要强调 q 为常数且 $|q|<1$，若没有 q 为常数这个条件，如 $q=1-\frac{1}{n}, n=2,3,\cdots$，则 $|q|<1$，但

$$\lim_{n\to\infty}q^n=\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n=e^{\lim\limits_{n\to\infty}n\ln\left(1-\frac{1}{n}\right)}=e^{\lim\limits_{n\to\infty}-\frac{n}{n}}=e^{-1}\neq0.$$

1.2.2 解 由题设得，

$$\frac{n}{\sqrt{n^2+n}}<\sum_{i=1}^{n}\frac{1}{\sqrt{n^2+i}}<\frac{n}{\sqrt{n^2+1}},$$

又 $\lim\limits_{n\to\infty}\frac{n}{\sqrt{n^2+n}}=\lim\limits_{n\to\infty}\frac{n}{\sqrt{n^2+1}}=1$，根据夹逼准则，原式=1.

1.2.3 解 分子不动，将分母改成相同的，则

$$\frac{n(n+1)}{2(n^2+n+n)}\leqslant\frac{1}{n^2+n+1}+\frac{2}{n^2+n+2}+\cdots+\frac{n}{n^2+n+n}\leqslant\frac{n(n+1)}{2(n^2+n+1)},$$

又 $\lim\limits_{n\to\infty}\frac{n(n+1)}{2(n^2+n+1)}=\lim\limits_{n\to\infty}\frac{n(n+1)}{2(n^2+n+n)}=\frac{1}{2}$，所以根据夹逼准则，原式 $=\frac{1}{2}$.

1.2.4 证明 因

$$a_{n+1}-a_n=\frac{1}{(n+1)^2}>0,$$

故数列 $\{a_n\}$ 单调增加.

当 $n>1$ 时，因为

$$\begin{aligned} a_n &= \frac{1}{1^2}+\frac{1}{2^2}+\cdots+\frac{1}{n^2} \\ &< 1+\frac{1}{1\cdot 2}+\cdots+\frac{1}{(n-1)\cdot n} \\ &= 1+\left(1-\frac{1}{2}\right)+\cdots+\left(\frac{1}{n-1}-\frac{1}{n}\right) \\ &= 2-\frac{1}{n}<2, \end{aligned}$$

故数列$\{a_n\}$有上界. 由单调有界准则，数列$\{a_n\}$收敛.

1.2.5 解 先证单调性. 由 $x_n+(x_n-4)x_{n-1}=3$，得 $x_n=\frac{3+4x_{n-1}}{1+x_{n-1}}$，又 $x_1=2$，所以 $x_2=\frac{3+4\times 2}{1+2}=\frac{11}{3}>x_1>0$，假设 $x_k>x_{k-1}>0$ 成立，则

$$x_{k+1}-x_k=\frac{3+4x_k}{1+x_k}-\frac{3+4x_{k-1}}{1+x_{k-1}}=\frac{x_k-x_{k-1}}{(1+x_k)(1+x_{k-1})}>0,$$

故 $x_{k+1}>x_k$，即数列$\{x_n\}$单调增加.

再证明其有界. 又 $x_n=\frac{3+4x_{n-1}}{1+x_{n-1}}=3+\frac{x_{n-1}}{1+x_{n-1}}<3+1=4$，所以数列$\{x_n\}$有上界.

由单调有界准则知$\lim\limits_{n\to\infty}x_n$存在. 设$\lim\limits_{n\to\infty}x_n=A$，令 $n\to\infty$，由 $x_n=\frac{3+4x_{n-1}}{1+x_{n-1}}$，得 $A=\frac{3+4A}{1+A}$，解得 $A=\frac{3\pm\sqrt{21}}{2}$，由题设，$x_n>0$，根据极限保号性可知 $A\geqslant 0$，故$\lim\limits_{n\to\infty}x_n=\frac{3+\sqrt{21}}{2}$.

1.2.6 (1)**证明** 由 $a_{n+1}=\frac{a_n+a_{n-1}}{2}$，则

$$\begin{aligned} a_{n+1}-a_n &= \frac{a_n+a_{n-1}}{2}-a_n=\frac{-a_n+a_{n-1}}{2} \\ &= \left(-\frac{1}{2}\right)\cdot(a_n-a_{n-1})=\left(-\frac{1}{2}\right)^2\cdot(a_{n-1}-a_{n-2}) \\ &= \cdots=\left(-\frac{1}{2}\right)^n(a_1-a_0)=\left(-\frac{1}{2}\right)^n, n=1,2,\cdots. \end{aligned}$$

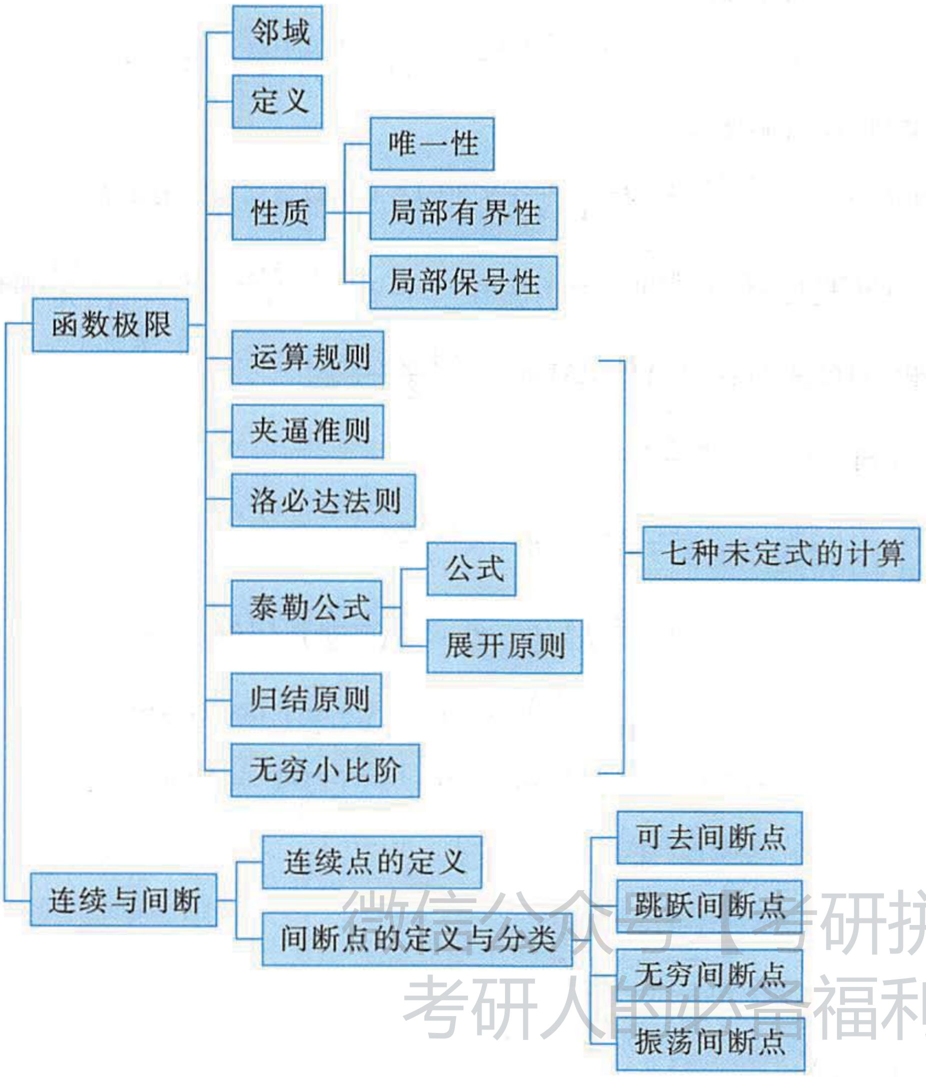
(2)**解**

$$\begin{aligned} a_n &= (a_n-a_{n-1})+(a_{n-1}-a_{n-2})+\cdots+(a_1-a_0)+a_0 \\ &= \left(-\frac{1}{2}\right)^{n-1}+\left(-\frac{1}{2}\right)^{n-2}+\cdots+1 \\ &= \frac{1-\left(-\frac{1}{2}\right)^n}{1-\left(-\frac{1}{2}\right)}, \end{aligned}$$

故$\lim\limits_{n\to\infty}a_n=\lim\limits_{n\to\infty}\frac{1-\left(-\frac{1}{2}\right)^n}{\frac{3}{2}}=\frac{2}{3}$.

第3讲 函数极限与连续性

基础知识结构



基础内容精讲

一、函数极限

1. 邻域

(1)一维的情形.

邻域 以点 x_0 为中心的任何开区间称为点 x_0 的邻域，记作 $U(x_0)$.

δ 邻域 设 δ 是一正数，则称开区间 $(x_0-\delta,x_0+\delta)$ 为点 x_0 的 δ 邻域，记作 $U(x_0,\delta)$，即

$$U(x_0,\delta)=\{x\mid x_0-\delta<x<x_0+\delta\}=\{x\mid |x-x_0|<\delta\},$$

其中点 x_0 称为邻域的中心，δ 称为邻域的半径.

去心 δ 邻域 定义去心邻域 $\mathring{U}(x_0,\delta)$：$\mathring{U}(x_0,\delta)=\{x\mid 0<|x-x_0|<\delta\}$.

左、右 δ 邻域 $\{x\mid 0<x-x_0<\delta\}$ 称为点 x_0 的右 δ 邻域，记作 $U^+(x_0,\delta)$；$\{x\mid 0<x_0-x<\delta\}$ 称为点 x_0 的左 δ 邻域，记作 $U^-(x_0,\delta)$.

(2)二维的情形.

δ 邻域 设 $P_0(x_0,y_0)$ 是 xOy 平面上的一个点，δ 是某一正数. 与点 $P_0(x_0,y_0)$ 的距离小于 δ 的点 $P(x,y)$ 的全体，称为点 P_0 的 δ 邻域，记为 $U(P_0,\delta)$，即

$$U(P_0,\delta)=\{P\mid |PP_0|<\delta\} \text{或} U(P_0,\delta)=\{(x,y)\mid \sqrt{(x-x_0)^2+(y-y_0)^2}<\delta\}.$$

去心 δ 邻域 点 P_0 的去心 δ 邻域，记作 $\mathring{U}(P_0,\delta)$，即 $\mathring{U}(P_0,\delta)=\{P\mid 0<|PP_0|<\delta\}$. 需要指出，如果不需要强调邻域的半径 δ，则用 $U(P_0)$ 表示点 P_0 的某个邻域，点 P_0 的去心邻域记作 $\mathring{U}(P_0)$.

δ 邻域的几何意义 $U(P_0,\delta)$ 表示 xOy 平面上以点 $P_0(x_0,y_0)$ 为中心，$\delta>0$ 为半径的圆的内部的点 $P(x,y)$ 的全体.

邻域与区间(区域) 邻域当然属于区间(区域)的范畴，但事实上，邻域通常表示“一个局部位置”，比如“点 x_0 的 δ 邻域”，就可以称为“点 x_0 的附近”. 于是，函数 $f(x)$ 在点 x_0 的某 δ 邻域内有定义也就是函数 $f(x)$ 在点 x_0 的附近有定义，这个“附近”到底有多近多远，既难以说明也没有必要说明. 有例为证，2007 年有一道考研数学题如下：设函数 $y=y(x)$ 由方程 $y\ln y-x+y=0$ 确定，试判断曲线 $y=y(x)$ 在点(1,1)附近的凹凸性.

【注】 关于邻域的一组概念非常重要，因为这涉及我们将要“在一个局部位置”细致地研究问题.

2. 函数极限的定义

设函数 $f(x)$ 在点 x_0 的某一去心邻域内有定义. 若存在常数 A，对于任意给定的 $\varepsilon>0$(不论它多么小)，总存在正数 δ，使得当 $0<|x-x_0|<\delta$ 时，对应的函数值 $f(x)$ 都满足不等式 $|f(x)-A|<\varepsilon$，则 A 就叫作函数 $f(x)$ 当 $x\to x_0$ 时的极限，记为

$$\lim_{x\to x_0}f(x)=A \text{ 或 } f(x)\to A\ (x\to x_0).$$

写成“$\varepsilon-\delta$ 语言”：$\lim\limits_{x\to x_0}f(x)=A\Leftrightarrow\forall\varepsilon>0$，$\exists\delta>0$，当 $0<|x-x_0|<\delta$ 时，有 $|f(x)-A|<\varepsilon$.

【注1】 这里 x 的趋向方式要比数列问题多得多，对于 $x\to x_0$，既要考虑 x 从 x_0 的左侧（小于 x_0）无限接近 x_0，即 $x\to x_0^-$，也要考虑 x 从 x_0 的右侧（大于 x_0）无限接近 x_0，即 $x\to x_0^+$；对于 $x\to\infty$，既包括 $x\to-\infty$，也包括 $x\to+\infty$，不再一一列出. 读者应学会写出函数极限的精确定义，提示一下：对于$x\to\infty$时的极限，其“$\varepsilon-X$语言”为

$$\lim_{x\to\infty}f(x)=A\Leftrightarrow\forall\varepsilon>0,\ \exists X>0,\ \text{当}\ |x|>X\ \text{时，有}\ |f(x)-A|<\varepsilon.$$

【注2】 (1)函数的单侧极限.

若当 $x\to x_0^-$ 时，$f(x)$无限接近于某常数 A，则常数 A 叫作函数 $f(x)$ 当 $x\to x_0$ 时的**左极限**，记为

$$\lim_{x\to x_0^-}f(x)=A\ \text{或}\ f(x_0^-)=A.$$

若当 $x\to x_0^+$ 时，$f(x)$无限接近于某常数 A，则常数 A 叫作函数 $f(x)$ 当 $x\to x_0$ 时的**右极限**，记为

$$\lim_{x\to x_0^+}f(x)=A\ \text{或}\ f(x_0^+)=A.$$

(2)函数极限存在的充要条件.

$$\lim_{x\to x_0}f(x)=A\Leftrightarrow\lim_{x\to x_0^-}f(x)=A,\text{且}\lim_{x\to x_0^+}f(x)=A,$$

$$\lim_{x\to x_0}f(x)=A\Leftrightarrow f(x)=A+\alpha(x),\lim_{x\to x_0}\alpha(x)=0.$$

3. 函数极限的性质

唯一性 如果极限$\lim\limits_{x\to x_0}f(x)$存在，那么极限唯一.

局部有界性 如果$\lim\limits_{x\to x_0}f(x)=A$，则存在正常数 M 和 δ，使得当 $0<|x-x_0|<\delta$ 时，有 $|f(x)|\leqslant M$.

局部保号性 如果 $f(x)\to A(x\to x_0)$，且 $A>0$（或 $A<0$），那么存在常数 $\delta>0$，使得当 $0<|x-x_0|<\delta$ 时，有 $f(x)>0$（或 $f(x)<0$）.

【注】 **推论** 如果 $f(x)\geqslant0$（或$\leqslant0$）$(x\to x_0)$且$\lim\limits_{x\to x_0}f(x)=A$，则 $A\geqslant0$（或$\leqslant0$）.

4. 极限运算规则

若 $\lim f(x)=A,\lim g(x)=B$，那么

①$\lim[kf(x)\pm lg(x)]=k\lim f(x)\pm l\lim g(x)=kA\pm lB$，其中 k,l 为常数；

②$\lim[f(x)\cdot g(x)]=\lim f(x)\cdot\lim g(x)=A\cdot B$，特别地，若 $\lim f(x)$存在，n 为正整数，则

$$\lim[f(x)]^n=[\lim f(x)]^n;$$

③$\lim\dfrac{f(x)}{g(x)}=\dfrac{\lim f(x)}{\lim g(x)}=\dfrac{A}{B}(B\neq0)$.

5. 夹逼准则

如果函数 $f(x)$，$g(x)$及 $h(x)$满足下列条件：

①$g(x)\leqslant f(x)\leqslant h(x)$；

②$\lim g(x)=A,\lim h(x)=A$.

则 $\lim f(x)$存在，且 $\lim f(x)=A$.

【注】 常见的一个问题：设任意的 x，总有 $\varphi(x)\leqslant f(x)\leqslant g(x)$，且 $\lim[g(x)-\varphi(x)]=0$，则 $\lim f(x)$ 是否一定存在？答案是否定的. $\lim[g(x)-\varphi(x)]$ 存在并不能说明 $\lim g(x)$，$\lim\varphi(x)$ 都存在，从而也不能保证 $\lim f(x)$ 存在.

例如，当 $x>0$ 时，取 $\varphi(x)=x+\frac{1}{x+1}$，$f(x)=x+\frac{2}{x+1}$，$g(x)=x+\frac{3}{x+1}$，则 $\varphi(x)<f(x)<g(x)$，且 $\lim\limits_{x\to+\infty}[g(x)-\varphi(x)]=0$，但 $\lim\limits_{x\to+\infty}f(x)$ 不存在.

6. 洛必达法则

洛必达

(1661—1704)

法则一 设①当 $x\to a$（或 $x\to\infty$）时，函数 $f(x)$ 及 $F(x)$ 都趋于零；

② $f'(x)$ 及 $F'(x)$ 在点 a 的某去心邻域内（或当 $|x|>X$，此时 X 为充分大的正数）存在，且 $F'(x)\neq0$；

③ $\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$（或 $\lim\limits_{x\to\infty}\frac{f'(x)}{F'(x)}$）存在或无穷大，

则 $\lim\limits_{x\to a}\frac{f(x)}{F(x)}=\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$（或 $\lim\limits_{x\to\infty}\frac{f(x)}{F(x)}=\lim\limits_{x\to\infty}\frac{f'(x)}{F'(x)}$）.

法则二 设①当 $x\to a$（或 $x\to\infty$）时，函数 $f(x)$ 及 $F(x)$ 都趋于无穷大；

② $f'(x)$ 及 $F'(x)$ 在点 a 的某去心邻域内（或当 $|x|>X$，此时 X 为充分大的正数）存在，且 $F'(x)\neq0$；

③ $\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$（或 $\lim\limits_{x\to\infty}\frac{f'(x)}{F'(x)}$）存在或无穷大，

则 $\lim\limits_{x\to a}\frac{f(x)}{F(x)}=\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$（或 $\lim\limits_{x\to\infty}\frac{f(x)}{F(x)}=\lim\limits_{x\to\infty}\frac{f'(x)}{F'(x)}$）.

【注】 (1)一般说来，洛必达法则是用来计算"$\frac{0}{0}$"型或者"$\frac{\infty}{\infty}$"型未定式极限的，不是"$\frac{0}{0}$"型或者"$\frac{\infty}{\infty}$"型，就不能用洛必达法则.

(2)如果极限 $\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$ 仍属于"$\frac{0}{0}$"型或者"$\frac{\infty}{\infty}$"型，且 $f'(x)$，$F'(x)$ 继续满足洛必达法则的条件，则可以继续使用洛必达法则，即 $\lim\limits_{x\to a}\frac{f(x)}{F(x)}=\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}=\lim\limits_{x\to a}\frac{f''(x)}{F''(x)}$.

(3)如果 $\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$ 不存在也不为 ∞，不能推出 $\lim\limits_{x\to a}\frac{f(x)}{F(x)}$ 不存在也不为 ∞，简单一点说就是：

对于 $\lim\limits_{x\to a}\frac{f(x)}{F(x)}=\lim\limits_{x\to a}\frac{f'(x)}{F'(x)}$，**"右存在，则左存在；但左存在，并不意味着右一定存在"**. 比如说，极限

$$\lim_{x\to0}\frac{x^2\cdot\sin\frac{1}{x}}{x}=\lim_{x\to0}x\cdot\sin\frac{1}{x}=0$$

存在，而如果使用洛必达法则，会有

$$\lim_{x\to0}\frac{x^2\cdot\sin\frac{1}{x}}{x}=\lim_{x\to0}\left(2x\cdot\sin\frac{1}{x}-\cos\frac{1}{x}\right),$$

这个极限显然不存在. 这是一个很细致、很隐蔽的问题，稍不注意就可能出错.

7. 泰勒公式

泰勒公式是极限计算的重要工具(其来源见第 6 讲“基础内容精讲”的定理 9).

泰勒

(1685—1731)

第一，要将以下几个重要函数的泰勒公式熟稔于心($x\to0$)：

$\sin x=x-\frac{x^3}{3!}+o(x^3)$，　　$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}+o(x^4)$，

$\arcsin x=x+\frac{x^3}{3!}+o(x^3)$，　　$\tan x=x+\frac{x^3}{3}+o(x^3)$，

$\arctan x=x-\frac{x^3}{3}+o(x^3)$，　　$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}+o(x^3)$，

$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+o(x^3)$，　　$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+o(x^2)$.

【注】 从数学命题的角度，对以上公式进行处理，可得到一组“差函数”的等价无穷小代换式：如 $x-\sin x=\frac{1}{6}x^3+o(x^3)$，则 $x-\sin x\sim\frac{1}{6}x^3\ (x\to0)$，同理有 $\arcsin x-x\sim\frac{1}{6}x^3\ (x\to0)$，$\tan x-x\sim\frac{1}{3}x^3\ (x\to0)$，$x-\arctan x\sim\frac{x^3}{3}(x\to0)$等. 并可将这些公式广义化，如第一个公式广义化为狗$-\sin$狗$\sim\frac{1}{6}$(狗)^3(狗$\to0$)，其余类似.

第二，要掌握高阶无穷小的计算规则，详见本讲的“基础内容精讲”部分.

第三，也是最关键的一点，用泰勒公式求极限时，函数应展开到 x 的几次幂？

(1)$\frac{A}{B}$型，适用于“上下同阶”原则.

具体说来，如果分母(或分子)是 x 的 k 次幂，则应把分子(或分母)展开到 x 的 k 次幂，可称为“上下同阶”原则.

例如，为了计算$\lim\limits_{x\to0}\frac{x-\sin x}{x^3}$，把 $\sin x$ 泰勒展开，一般可写成如下三种形式：

①$\sin x=x+o(x)$；②$\sin x=x-\frac{1}{6}x^3+o(x^3)$；③$\sin x=x-\frac{1}{6}x^3+\frac{1}{5!}x^5+o(x^5)$.

根据“上下同阶”原则，①式展开得“不够”，③式展开得“过多”，②式正符合要求. 于是，

$$\lim_{x\to0}\frac{x-\sin x}{x^3}=\lim_{x\to0}\frac{x-\left[x-\frac{1}{6}x^3+o(x^3)\right]}{x^3}=\lim_{x\to0}\frac{\frac{1}{6}x^3+o(x^3)}{x^3}=\frac{1}{6}.$$

(2)$A-B$ 型，适用于“幂次最低”原则.

具体说来，即将 A,B 分别展开到它们的系数不相等的 x 的最低次幂为止.

例如，已知当 $x\to0$ 时，$\cos x-e^{-\frac{x^2}{2}}$与 ax^b为等价无穷小，求 a,b.

用泰勒公式，$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}+o(x^4)$，$e^{-\frac{x^2}{2}}=1-\frac{x^2}{2}+\frac{1}{2!}\frac{x^4}{4}+o(x^4)$.

显然，将 $\cos x$，$e^{-\frac{x^2}{2}}$展开到 x^4时，其系数就不一样了，使用“幂次最低”原则，展开到此项后，进行运算，得

$$\cos x-e^{-\frac{x^2}{2}}=\left[1-\frac{x^2}{2!}+\frac{x^4}{4!}+o(x^4)\right]-\left[1-\frac{x^2}{2}+\frac{1}{2!}\frac{x^4}{4}+o(x^4)\right]=-\frac{1}{12}x^4+o(x^4),$$

于是可知 $\cos x-\mathrm{e}^{-\frac{x^2}{2}}\sim-\frac{1}{12}x^4(x\to0)$，故 $a=-\frac{1}{12},b=4$.

8. 海涅定理(归结原则)

设 $f(x)$ 在 $\mathring{U}(x_0,\delta)$ 内有定义，则

$\lim\limits_{x\to x_0}f(x)=A$ 存在⇔对任何 $\mathring{U}(x_0,\delta)$ 内以 x_0 为极限的数列 $\{x_n\}(x_n\neq x_0)$，极限 $\lim\limits_{n\to\infty}f(x_n)=A$ 存在.

【注】 众所周知，数列极限与函数极限是分别独立定义的，但是海涅定理是联系数列极限与函数极限的桥梁. 它指出：在极限存在的条件下，函数极限和数列极限可以相互转化. 虽然有些读者可能没有听说过这个定理，但是我们都在不知不觉地使用它，比如，试证明 $\lim\limits_{x\to0}\frac{1}{x}\sin\frac{1}{x}$ 不存在. 解答如下：

若取 $x_n=\frac{1}{n\pi}\to0,n\to\infty$，则有 $f(x_n)=0$；若取 $x_n'=\frac{1}{\left(2n+\frac{1}{2}\right)\pi}\to0,n\to\infty$，则有 $f(x_n')=\left(2n+\frac{1}{2}\right)\pi\to\infty,n\to\infty$，根据海涅定理，极限不存在. 证毕.

事实上，$\frac{1}{x}\sin\frac{1}{x}$ 是当 $x\to0$ 时的无界量，但不是无穷大量.

9. 无穷小比阶

(1)无穷小定义.

如果当 $x\to x_0$(或 $x\to\infty$)时，函数 $f(x)$ 的极限为零，那么称函数 $f(x)$ 为当 $x\to x_0$(或 $x\to\infty$)时的**无穷小**，记为

$$\lim_{x\to x_0}f(x)=0(\text{或}\lim_{x\to\infty}f(x)=0).$$

特别地，以零为极限的数列 $\{x_n\}$ 称为 $n\to\infty$ 时的无穷小.

【注】 (1)无穷大定义.

如果当 $x\to x_0$(或 $x\to\infty$)时，函数 $|f(x)|$ 无限增大，那么称函数 $f(x)$ 为当 $x\to x_0$(或 $x\to\infty$)时的**无穷大**，记为

$$\lim_{x\to x_0}f(x)=\infty(\text{或}\lim_{x\to\infty}f(x)=\infty).$$

(2)无穷小与无穷大的关系.

在自变量的同一变化过程中，如果 $f(x)$ 为无穷大，则 $\frac{1}{f(x)}$ 为无穷小；反之，如果 $f(x)$ 为无穷小，且 $f(x)\neq0$，则 $\frac{1}{f(x)}$ 为无穷大.

(2)无穷小的比阶.

设在自变量的同一变化过程中，$\lim\alpha(x)=0,\lim\beta(x)=0$，且 $\beta(x)\neq0$，则

①若 $\lim\frac{\alpha(x)}{\beta(x)}=0$，则称 $\alpha(x)$ 是比 $\beta(x)$ **高阶的无穷小**，记为 $\alpha(x)=o(\beta(x))$；

②若 $\lim\frac{\alpha(x)}{\beta(x)}=\infty$，则称 $\alpha(x)$ 是比 $\beta(x)$ **低阶的无穷小**；

③若 $\lim\frac{\alpha(x)}{\beta(x)}=c\neq0$，则称 $\alpha(x)$ 与 $\beta(x)$ 是**同阶无穷小**；

④若 $\lim \frac{\alpha(x)}{\beta(x)}=1$，则称 $\alpha(x)$ 与 $\beta(x)$ 是**等价无穷小**，记为 $\alpha(x)\sim\beta(x)$；

⑤若 $\lim \frac{\alpha(x)}{[\beta(x)]^k}=c\neq0,k>0$，则称 $\alpha(x)$ 是 $\beta(x)$ 的 **k 阶无穷小**.

【注】 **并不是任意两个无穷小都可进行比阶的.** 例如，当 $x\to0$ 时，$x\sin\frac{1}{x}$ 与 x^2 虽然都是无穷小，但是却不可以比阶，也就是说既无高低阶之分，也无同阶可言，因为 $\lim\limits_{x\to0}\frac{x\sin\frac{1}{x}}{x^2}=\lim\limits_{x\to0}\frac{1}{x}\sin\frac{1}{x}$ 不存在.

(3)无穷小运算规则.

①**有限个**无穷小的和是无穷小.

②有界函数与无穷小的乘积是无穷小.

③**有限个**无穷小的乘积是无穷小.

④无穷小的运算.

设 m,n 为正整数，则

a. $o(x^m)\pm o(x^n)=o(x^l),l=\min\{m,n\}$（加减法时低阶"吸收"高阶）；

b. $o(x^m)\cdot o(x^n)=o(x^{m+n}),x^m\cdot o(x^n)=o(x^{m+n})$（乘法时阶数"累加"）；

c. $o(x^m)=o(kx^m)=k\cdot o(x^m),k\neq0$ 且为常数（非零常数相乘不影响阶数）.

【注】 在后面的泰勒公式中，会对上述高阶无穷小的运算提出要求，请读者学会正确书写.

(4)常用的等价无穷小.

当 $x\to0$ 时，常用的等价无穷小有：

$$\sin x\sim x,\quad \tan x\sim x,\quad \arcsin x\sim x,\quad \arctan x\sim x,\quad \ln(1+x)\sim x,\quad e^x-1\sim x,$$

$$a^x-1\sim x\ln a,\quad 1-\cos x\sim\frac{1}{2}x^2,\quad (1+x)^\alpha-1\sim\alpha x.$$

【注】 使用时一般都要做广义化：可将 x 替换为趋向于 0 的函数，请灵活使用.

二、函数的连续与间断

1. 连续点的定义

设函数 $f(x)$ 在点 x_0 的某一邻域内有定义，且有 $\lim\limits_{x\to x_0}f(x)=f(x_0)$，则称函数 $f(x)$ 在点 x_0 处连续.

2. 间断点的定义与分类

以下设函数 $f(x)$ 在点 x_0 的某去心邻域内有定义.

(1)可去间断点.

若 $\lim\limits_{x\to x_0}f(x)=A\neq f(x_0)$（$f(x_0)$ 甚至可以无定义），则这类间断点称为**可去间断点**.

【注】 只要修改或者补充 $f(x_0)$，使得 $\lim\limits_{x\to x_0}f(x)=A=f(x_0)$，就会使得函数在点 x_0 处连续，于是，这个点叫作可去间断点，也叫作**可补间断点**.

(2)跳跃间断点.

若 $\lim\limits_{x\to x_0^-}f(x)$ 与 $\lim\limits_{x\to x_0^+}f(x)$ 都存在，但 $\lim\limits_{x\to x_0^-}f(x)\neq\lim\limits_{x\to x_0^+}f(x)$，则这类间断点称为**跳跃间断点**.

可去间断点和跳跃间断点统称为**第一类间断点**.

(3)无穷间断点.

若 $\lim\limits_{x\to x_0}f(x)=\infty$，则这类间断点称为**无穷间断点**，如函数 $y=\dfrac{1}{x}$ 的点 $x=0$ 处为无穷间断点.

(4)振荡间断点.

若 $\lim\limits_{x\to x_0}f(x)$ 振荡不存在，则这类间断点称为**振荡间断点**，如函数 $y=\sin\dfrac{1}{x}$ 在点 $x=0$ 处没有定义，且当 $x\to 0$ 时，函数值在 -1 与 1 这两个数之间交替振荡取值，极限不存在，故点 $x=0$ 为函数 $y=\sin\dfrac{1}{x}$ 的振荡间断点.

无穷间断点和振荡间断点都属于**第二类间断点**，除此之外，还有不属于上述定义的第二类间断点，比如下面的"注"的(a)，(b)两种情形.

【注】 (1)大学教材(如同济大学《高等数学》第七版)中一般均写明"在点 x_0 的某去心邻域有定义"的前提下，才讨论间断点. 故如图 1-3-1 所示：显然，$f(x)$ 在 $x=x_0$ 处只有右侧邻域有定义，故不讨论 $x=x_0$ 处是否为间断点.(但 $x=x_0$ 是 $f(x)$ 的铅垂渐近线)

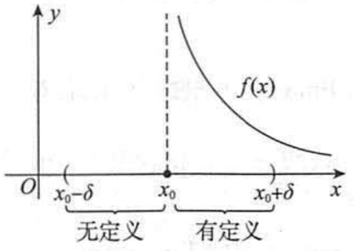


图 1-3-1

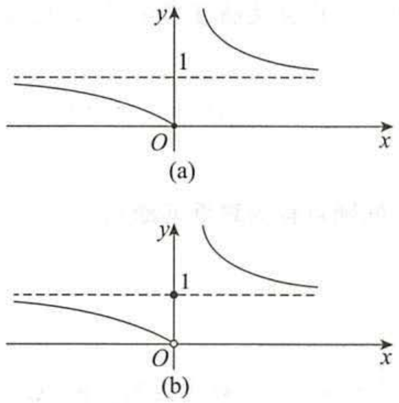
(2)进一步地，"在点 x_0 处某去心邻域均有定义"的函数 $f(x)$，如图1-3-2所示.

如图 1-3-2(a)，(b)所示的情形从未考过，因为一般中国大陆的数学教材中没有单侧间断的定义，这种定义可见国际上的一些教材，如菲氏《微积分学教程》有如下定义.

函数 $f(x)$ 在点 x_0 处是右(或左)连续的，只需满足极限关系式

$$f(x_0^+)=\lim_{x\to x_0^+}f(x)=f(x_0)\text{(或 }f(x_0^-)=\lim_{x\to x_0^-}f(x)=f(x_0)\text{)},$$

若这关系式不成立，则称函数 $f(x)$ 在点 x_0 处有**右间断(或左间断)**，如此说来，对于图 1-3-2(a)，$\lim\limits_{x\to 0^-}f(x)=f(0)$，故 $x=0$ 处左连续，但 $\lim\limits_{x\to 0^+}f(x)=+\infty$，故 $x=0$ 处叫右侧无穷间断，读者可类似分析图 1-3-2(b)的情形.



(a)

$f(x)=\begin{cases} e^{\frac{1}{x}}, & x\neq 0, \\ 0, & x=0 \end{cases}$ 说 $\begin{cases} x=0 \text{ 处右侧无穷间断}, \\ x=0 \text{ 处左侧连续}. \end{cases}$

(b)

$f(x)=\begin{cases} e^{\frac{1}{x}}, & x\neq 0, \\ 1, & x=0 \end{cases}$ 说 $\begin{cases} x=0 \text{ 处右侧无穷间断}, \\ x=0 \text{ 处左侧跳跃间断}. \end{cases}$

图 1-3-2

基础例题精解

1. 函数极限的性质

本部分主要考查函数极限的唯一性、局部有界性和局部保号性.

(1)对于唯一性，

① 对于 $x\to\infty$，意味着 $x\to+\infty$，且 $x\to-\infty$；

② 对于 $x\to x_0$，意味着 $x\to x_0^+$，且 $x\to x_0^-$.

我们称这个细节的问题为自变量取值的“双向性(有正有负)”，基于此，我们看几个重要的函数极限问题.

$\lim\limits_{x\to\infty}e^x$不存在，因为 $\lim\limits_{x\to+\infty}e^x=+\infty$，$\lim\limits_{x\to-\infty}e^x=0$，根据“极限若存在，必唯一”，得原极限不存在；

$\lim\limits_{x\to0}\frac{\sin x}{|x|}$不存在，因为$\lim\limits_{x\to0^+}\frac{\sin x}{|x|}=\lim\limits_{x\to0^+}\frac{\sin x}{x}=1$，$\lim\limits_{x\to0^-}\frac{\sin x}{|x|}=\lim\limits_{x\to0^-}\frac{\sin x}{-x}=-1$；

$\lim\limits_{x\to\infty}\arctan x$ 不存在，因为 $\lim\limits_{x\to+\infty}\arctan x=\frac{\pi}{2}$，$\lim\limits_{x\to-\infty}\arctan x=-\frac{\pi}{2}$；

$\lim\limits_{x\to0}[x]$不存在，因为$\lim\limits_{x\to0^+}[x]=0$，$\lim\limits_{x\to0^-}[x]=-1$.

(2)对于局部有界性，

①设$\lim\limits_{x\to\cdot}f(x)$存在，则当“$x\to\cdot$”时，$f(x)$有界. 其中“$x\to\cdot$”是指 $x\to x_0$，$x\to x_0^-$，$x\to x_0^+$，$x\to\infty$，$x\to-\infty$，$x\to+\infty$六种情形. 值得注意的是极限存在只是函数局部有界的充分条件，并非必要条件；

②设 $f(x)$在$[a,b]$上连续，则 $f(x)$在$[a,b]$上有界；

③有界函数与有界函数的和、差、积仍为有界函数；

④若 $f'(x)$在有限区间(a,b)内有界，则 $f(x)$在该区间内有界.

(3)对于局部保号性，主要和导数的几何应用结合来命题，参见第 5 讲的例题.

例 1.3.1 设 a 为常数，且$\lim\limits_{x\to0}\left(\frac{e^{\frac{1}{x}}-\pi}{e^{\frac{2}{x}}+1}+a\cdot\arctan\frac{1}{x}\right)$存在，求 a 的值，并计算此极限.

解 $\lim\limits_{x\to 0^-}\left(\dfrac{e^{\frac{1}{x}}-\pi}{e^{\frac{2}{x}}+1}+a\cdot\arctan\dfrac{1}{x}\right)=-\pi-\dfrac{\pi}{2}a$, $\lim\limits_{x\to 0^+}\left(\dfrac{e^{\frac{1}{x}}-\pi}{e^{\frac{2}{x}}+1}+a\cdot\arctan\dfrac{1}{x}\right)=0+\dfrac{\pi}{2}a$.

由于原极限存在，则左、右极限相等，即 $-\pi-\dfrac{\pi}{2}a=\dfrac{\pi}{2}a$，则 $a=-1$，所以原极限 $=-\dfrac{\pi}{2}$.

例 1.3.2 函数 $f(x)=\dfrac{|x|\sin(x-2)}{x(x-1)(x-2)^2}$ 在下列哪个区间内有界（　　）.

(A) $(-1,0)$　　(B) $(0,1)$　　(C) $(1,2)$　　(D) $(2,3)$

分析 有如下两个重要结论：

①若 $f(x)$ 在闭区间 $[a,b]$ 上连续，则 $f(x)$ 在闭区间 $[a,b]$ 上有界；

②若 $f(x)$ 在开区间 (a,b) 内连续，且极限 $\lim\limits_{x\to a^+}f(x)$ 与 $\lim\limits_{x\to b^-}f(x)$ 存在，则函数 $f(x)$ 在开区间 (a,b) 内有界.

解 应选(A).

因为当 $x\neq 0,1,2$ 时，$f(x)$ 连续，而

$$\lim_{x\to(-1)^+}f(x)=-\frac{\sin 3}{18},\quad \lim_{x\to 0^-}f(x)=-\frac{\sin 2}{4},$$

$$\lim_{x\to 0^+}f(x)=\frac{\sin 2}{4},\quad \lim_{x\to 1}f(x)=\infty,\quad \lim_{x\to 2}f(x)=\infty,$$

所以，函数 $f(x)$ 在 $(-1,0)$ 内有界，故选(A).

2. 七种未定式的计算

本节内容极其重要，它是高等数学计算的基础，读者要高度重视，充分训练.

对于 $\lim\limits_{x\to\cdot}f(x)$，自变量 x 的变化趋势共有六种情形：$x\to x_0^+\,(x>x_0)$，$x\to x_0^-\,(x<x_0)$，$x\to x_0$，$x\to+\infty$，$x\to-\infty$ 及 $x\to\infty$. 在没必要对其进行区别时，统一记成 $x\to\cdot$.

(1) 化简是第一步，切记.

化简的方法：①提出极限不为 0 的因式；②等价无穷小代换；③恒等变形（基本的恒等变形法如提公因式、拆项、合并、分子分母同除变量的最高次幂等，高级的恒等变形法如变量代换，也叫换元法等）. 需要强调的是很多问题如果不化简就计算，可能计算会很复杂，甚至可能计算不出结果.

(2) 判断类型，七种："$\dfrac{0}{0}$""$\dfrac{\infty}{\infty}$""$0\cdot\infty$""$\infty-\infty$""∞^0""0^0""1^∞".

(3)选择相应的方法进行计算（包括运算规则、夹逼准则、洛必达法则、泰勒公式、归结原则等）.

(1)"$\dfrac{0}{0}$" "$\dfrac{\infty}{\infty}$" "$0\cdot\infty$".

例 1.3.3 求极限 $\lim\limits_{x\to 0}\dfrac{e^{-\frac{1}{x^2}}}{x^{100}}$.

分析 这是"$\dfrac{0}{0}$"型未定式，看起来好像很简单，很多读者拿到这样一个问题，就直接用洛必达法则求导去了，试试看？

$$\lim_{x\to 0}\frac{e^{-\frac{1}{x^2}}}{x^{100}}\xlongequal{\text{洛必达法则}}\lim_{x\to 0}\frac{e^{-\frac{1}{x^2}}\cdot 2x^{-3}}{100x^{99}}=\frac{1}{50}\lim_{x\to 0}\frac{e^{-\frac{1}{x^2}}}{x^{102}}\xlongequal{\text{洛必达法则}}\cdots.$$

我们看到，越求导反而越复杂了，为什么？问题出在"$\dfrac{1}{x^2}$"这个形式上，注意，我们可以将"$\dfrac{1}{x^2}=\dfrac{x^0}{x^2}$"称为"头轻脚重"的"正三角形状△"，常识告诉我们，"正三角形状△"是稳定的，生活和工程中我们喜欢这种稳定吗？

当然，金字塔就是很好的例子. 可在数学上，这种稳定就不妙了——稳定就意味着不能动，不能动就做不下去了！比如：$\int \frac{1}{x^3+2x^2+x+1}\mathrm{d}x$，你怎么积分？而 $\int \frac{x^3+2x^2+x+1}{1}\mathrm{d}x$，积分多么简单！这例子很极端，但很能说明问题. 所以此题的正确做法是先作“倒代换”，将“正三角形状△”改成“倒三角形状▽”，头重脚轻根底浅，垮了……搞定！请看：

解

$$\lim_{x\to 0}\frac{\mathrm{e}^{-\frac{1}{x^2}}}{x^{100}} \xlongequal{令\frac{1}{x^2}=t} \lim_{t\to+\infty}\frac{\mathrm{e}^{-t}}{t^{-50}} = \lim_{t\to+\infty}\frac{t^{50}}{\mathrm{e}^t} \xlongequal{洛必达法则} \lim_{t\to+\infty}\frac{50t^{49}}{\mathrm{e}^t}$$

$$\xlongequal{洛必达法则} \lim_{t\to+\infty}\frac{50\cdot 49\cdot t^{48}}{\mathrm{e}^t} \xlongequal{洛必达法则} \cdots = \lim_{t\to+\infty}\frac{50!}{\mathrm{e}^t} = 0.$$

【注】 本题给我们的启示：①有些极限问题看起来下手很简单，但事实上如果不化简就计算，可能根本算不出结果；②可以通过变量倒代换将“正三角形状△”改成“倒三角形状▽”进行化简，这是一个经典思路，在后面的积分学和微分方程各章中，我们都会再次提到，请读者留心总结.

例 1.3.4 求极限 $\lim\limits_{x\to 0}\frac{\sqrt{1+\tan x}-\sqrt{1+\sin x}}{x\ln(1+x)-x^2}$.

解 这是“$\frac{0}{0}$”型未定式，分子是根号差“$\sqrt{a}-\sqrt{b}$”的形式，请记住一句话，一般说来，**“见根号差，用有理化”**，这是一种重要的化简手段，则

$$原式 = \lim_{x\to 0}\left\{\frac{\tan x-\sin x}{x[\ln(1+x)-x]}\cdot\frac{1}{\sqrt{1+\tan x}+\sqrt{1+\sin x}}\right\}$$

$$= \frac{1}{2}\lim_{x\to 0}\frac{\tan x\cdot(1-\cos x)}{x[\ln(1+x)-x]} = \frac{1}{2}\lim_{x\to 0}\frac{\sin x}{-\frac{x}{1+x}} = -\frac{1}{2}.$$

例 1.3.5 求极限 $\lim\limits_{x\to-\infty}\frac{\sqrt{4x^2+x-1}+x+1}{\sqrt{x^2+\sin x}}$.

解 这是“$\frac{\infty}{\infty}$”型未定式，注意到 $x\to-\infty$，也就意味着 $x<0$，并且 $\sqrt{4x^2+x-1}+x+1$ 是根号差“$\sqrt{a}-\sqrt{b}$”的形式，于是先用有理化做化简，再用 $t=-x$ 代换化为我们熟悉的式子，然后再计算.

$$原式 = \lim_{x\to-\infty}\frac{3x^2-x-2}{\sqrt{x^2+\sin x}(\sqrt{4x^2+x-1}-x-1)} = \lim_{t\to+\infty}\frac{3t^2+t-2}{\sqrt{t^2-\sin t}(\sqrt{4t^2-t-1}+t-1)}$$

$$= \lim_{t\to+\infty}\frac{3+\frac{1}{t}-\frac{2}{t^2}}{\sqrt{1-\frac{\sin t}{t^2}}\left(\sqrt{4-\frac{1}{t}-\frac{1}{t^2}}+1-\frac{1}{t}\right)} = 1.$$

例 1.3.6 求极限 $\lim\limits_{x\to-\infty}x(\sqrt{x^2+100}+x)$.

解 这是“$\infty\cdot 0$”型未定式，理由同上，则有

$$\lim_{x\to-\infty}x(\sqrt{x^2+100}+x) = \lim_{x\to-\infty}\frac{100x}{\sqrt{x^2+100}-x} \xlongequal{令t=-x} \lim_{t\to+\infty}\frac{-100}{\sqrt{1+\frac{100}{t^2}}+1} = -50.$$

【注】 顺便指出，本题和上题都有一个要注意的细节，就是当 $x<0$ 时，要使用代换 $t=-x$ 化为常见的情况，或用 $-x$ 同时除分子、分母，这样才不会出现正负上的错误.

例 1.3.7 求极限 $\lim\limits_{x\to1^-}\ln x\ln(1-x)$.

解 这是“$\infty\cdot0$”型未定式，注意一个事实：当 $x\to0$ 时，$\ln(1+x)\sim x$，将其广义化，得

$$\ln(1+u)\sim u\ (u\to0),$$

于是在考研中常考的一个式子是 $\ln x=\ln(1+x-1)\sim x-1\ (x\to1)$，则

$$\begin{aligned}\lim_{x\to1^-}\ln x\ln(1-x)&=\lim_{x\to1^-}\ln(1+x-1)\ln(1-x)\\&=\lim_{x\to1^-}(x-1)\ln(1-x)\\&\xlongequal{1-x=t}-\lim_{t\to0^+}t\ln t=0.\end{aligned}$$

【注】 在实考中，很多考生对于这种问题处理得不理想，事实上，读者一定要明白一个道理：不能只是简单地记住那几个基本的等价无穷小代换公式，而应该把它们广义化，并通过好的题目载体灵活应用，像本题一样. 这再一次说明：数学解题能力必须在做题实践中才能提高.

例 1.3.8 求 $\lim\limits_{x\to0}\dfrac{\arcsin x-\arctan x}{\sin x-\tan x}$.

解 当 $x\to0$ 时，由于 $\arcsin x-\arctan x=\left[\dfrac{1}{6}-\left(-\dfrac{1}{3}\right)\right]x^3+o(x^3)=\dfrac{1}{2}x^3+o(x^3)$，故 $\arcsin x-\arctan x\sim\dfrac{1}{2}x^3$；同理，当 $x\to0$ 时，由于 $\sin x-\tan x=\left(-\dfrac{1}{6}-\dfrac{1}{3}\right)x^3+o(x^3)=-\dfrac{1}{2}x^3+o(x^3)$，故 $\sin x-\tan x\sim-\dfrac{1}{2}x^3$，所以原式 $=\lim\limits_{x\to0}\dfrac{\dfrac{1}{2}x^3}{-\dfrac{1}{2}x^3}=-1$.

【注】 如果用洛必达法则，计算起来会相当麻烦.

例 1.3.9 求 $I=\lim\limits_{x\to0}x\left[\dfrac{10}{x}\right]$，其中[·]为取整符号.

解 当 $x\to0$ 时，$\dfrac{10}{x}\to\infty$，对于[∞]，此时想到极限计算的利器——夹逼准则(当常规求极限的方法，比如等价无穷小代换、泰勒公式、洛必达法则无法使用时，一定要能够想得起这个“两边夹击”的重要方法).

根据 $x-1<[x]\leqslant x$，有

$$\frac{10}{x}-1<\left[\frac{10}{x}\right]\leqslant\frac{10}{x},$$

于是

$$\begin{cases}x>0\Rightarrow10-x<x\cdot\left[\dfrac{10}{x}\right]\leqslant10,\\x<0\Rightarrow10-x>x\cdot\left[\dfrac{10}{x}\right]\geqslant10.\end{cases}$$

可见，无论 $x>0$，还是 $x<0$，不等式两边均可趋于同一极限，故 $I=\lim\limits_{x\to0}x\left[\dfrac{10}{x}\right]=10$.

(2)“$\infty-\infty$”.

对于“$\infty-\infty$”型未定式,一般有两种思路.

(1)如果函数中有分母,则通分,将加减法变形为乘除法,以便于使用其他计算工具(比如洛必达法则),见例 1.3.10.

(2)如果函数中没有分母,则可以通过提取公因式,或者作倒代换,出现分母后,再利用通分等恒等变形的方法,将加减法变形为乘除法,见例 1.3.11.

例 1.3.10 求极限$\lim\limits_{x\to0}\left(\dfrac{1}{\sin^2x}-\dfrac{\cos^2x}{x^2}\right)$.

解

$$\lim_{x\to0}\left(\frac{1}{\sin^2x}-\frac{\cos^2x}{x^2}\right)=\lim_{x\to0}\frac{x^2-\sin^2x\cos^2x}{x^2\sin^2x}=\lim_{x\to0}\frac{x^2-\frac{1}{4}\sin^22x}{x^4}$$

$$=\lim_{x\to0}\frac{2x-\frac{1}{2}\sin4x}{4x^3}=\lim_{x\to0}\frac{1-\cos4x}{6x^2}=\lim_{x\to0}\frac{\frac{1}{2}(4x)^2}{6x^2}=\frac{4}{3}.$$

例 1.3.11 求极限 $\lim\limits_{x\to+\infty}[x^2(\mathrm{e}^{\frac{1}{x}}-1)-x]$.

解

$$\text{原式}\xlongequal{令\ u=\frac{1}{x}}\lim_{u\to0^+}\frac{\mathrm{e}^u-1-u}{u^2}=\lim_{u\to0^+}\frac{\mathrm{e}^u-1}{2u}=\frac{1}{2}.$$

(3)“∞^0” “0^0” “1^∞”.

例 1.3.12 求极限 $\lim\limits_{x\to+\infty}(x+\sqrt{1+x^2})^{\frac{1}{x}}$.

解 这是“∞^0”型未定式,是幂指函数的极限,对于“∞^0”和“0^0”型这两种未定式,一般说来,我们都用恒等变形

$$\lim u^v=\mathrm{e}^{\lim v\ln u}\xlongequal{记}\exp\{\lim v\ln u\},$$

将其化成“$\frac{0}{0}$”“$\frac{\infty}{\infty}$”“$0\cdot\infty$”这三种类型,然后计算. 故原式$=\exp\left\{\lim\limits_{x\to+\infty}\dfrac{1}{x}\ln(x+\sqrt{1+x^2})\right\}$.

因 $\lim\limits_{x\to+\infty}\dfrac{1}{x}\ln(x+\sqrt{1+x^2})=\lim\limits_{x\to+\infty}\dfrac{1}{x+\sqrt{1+x^2}}\left(1+\dfrac{x}{\sqrt{1+x^2}}\right)=\lim\limits_{x\to+\infty}\dfrac{1}{\sqrt{1+x^2}}=0$,所以

$$\lim_{x\to+\infty}(x+\sqrt{1+x^2})^{\frac{1}{x}}=\mathrm{e}^0=1.$$

例 1.3.13 求极限$\lim\limits_{x\to0}\left(\dfrac{\mathrm{e}^x+\mathrm{e}^{2x}+\cdots+\mathrm{e}^{nx}}{n}\right)^{\frac{\mathrm{e}}{x}}$,其中 n 是给定的正整数.

解 这是“1^∞”型未定式,是幂指函数的极限,如果 $\lim u^v$ 属于“1^∞”型,则有一个重要且简单的计算方法:$\lim u^v=\mathrm{e}^{\lim(u-1)v}$.

推导如下:利用重要极限$\lim\limits_{x\to\infty}\left(1+\dfrac{1}{x}\right)^x=\mathrm{e}$,得

$$\lim u^v=\lim\{[1+(u-1)]^{\frac{1}{u-1}}\}^{(u-1)v}=\mathrm{e}^{\lim(u-1)v},$$

故原式$=\exp\left\{\lim\limits_{x\to0}\left(\dfrac{\mathrm{e}^x+\mathrm{e}^{2x}+\cdots+\mathrm{e}^{nx}}{n}-1\right)\dfrac{\mathrm{e}}{x}\right\}=\exp\left\{\lim\limits_{x\to0}\dfrac{\mathrm{e}}{x}\dfrac{\mathrm{e}^x+\mathrm{e}^{2x}+\cdots+\mathrm{e}^{nx}-n}{n}\right\}$

$$=\exp\left\{\lim_{x\to0}\frac{\mathrm{e}}{x}\frac{(\mathrm{e}^x-1)+(\mathrm{e}^{2x}-1)+\cdots+(\mathrm{e}^{nx}-1)}{n}\right\}=\exp\left\{\frac{\mathrm{e}}{n}\left(\lim_{x\to0}\frac{\mathrm{e}^x-1}{x}+\lim_{x\to0}\frac{\mathrm{e}^{2x}-1}{x}+\cdots+\lim_{x\to0}\frac{\mathrm{e}^{nx}-1}{x}\right)\right\}$$

$$=\exp\left\{\frac{e(1+2+\cdots+n)}{n}\right\}=\exp\left\{\frac{n+1}{2}e\right\}.$$

例 1.3.14 求$\lim\limits_{n\to\infty}\left(n\tan\frac{1}{n}\right)^{n^2}$（$n$为正整数）.

解 这是“1^∞”型未定式,解法同上,于是

$$\lim_{x\to0^+}\left(\frac{\tan x}{x}\right)^{\frac{1}{x^2}}=\exp\left\{\lim_{x\to0^+}\frac{1}{x^2}\left(\frac{\tan x}{x}-1\right)\right\}=\exp\left\{\lim_{x\to0^+}\frac{1}{x^2}\left(\frac{\tan x-x}{x}\right)\right\}$$

$$=\exp\left\{\lim_{x\to0^+}\frac{\tan x-x}{x^3}\right\}=\exp\left\{\lim_{x\to0^+}\frac{\sec^2 x-1}{3x^2}\right\}=e^{\frac{1}{3}}.$$

根据归结原则,取$x=\frac{1}{n}$,则原式$=e^{\frac{1}{3}}$.

3. 已知某一极限,求另一极限

此类问题是考研的常考题,关键要抓住已知极限与未知极限的联系,实现问题的转化.

例 1.3.15 如果$\lim\limits_{x\to0}\frac{x-\sin x+f(x)}{x^4}$存在,则$\lim\limits_{x\to0}\frac{x^3}{f(x)}=$(　　).

(A)-36　　(B)36　　(C)6　　(D)-6

解 应选(D).

记
$$\lim_{x\to0}\frac{x-\sin x+f(x)}{x^4}=A(\text{常数}),$$

则
$$0=\lim_{x\to0}\frac{x-\sin x+f(x)}{x^4}\cdot x=\lim_{x\to0}\frac{x-\sin x+f(x)}{x^3}$$

$$=\lim_{x\to0}\frac{x-\sin x}{x^3}+\lim_{x\to0}\frac{f(x)}{x^3}=\lim_{x\to0}\frac{x-\left[x-\frac{x^3}{3!}+o(x^3)\right]}{x^3}+\lim_{x\to0}\frac{f(x)}{x^3}$$

$$=\lim_{x\to0}\frac{\frac{x^3}{3!}+o(x^3)}{x^3}+\lim_{x\to0}\frac{f(x)}{x^3}=\frac{1}{6}+\lim_{x\to0}\frac{f(x)}{x^3},$$

所以$\lim\limits_{x\to0}\frac{f(x)}{x^3}=-\frac{1}{6}$,则$\lim\limits_{x\to0}\frac{x^3}{f(x)}=-6$,故选(D).

4. 已知极限反求参数

此类问题也是考研的常考题,事实上就是带着参数求极限,只是由于表达式中含着未知参数,要注意慎用洛必达法则(详见“基础内容精讲”“一”中的“6”).

例 1.3.16 设$\lim\limits_{x\to0}\frac{\ln(1+x)-(ax+bx^2)}{x^2}=2$,求常数$a,b$.

解 原极限$=\lim\limits_{x\to0}\frac{x-\frac{1}{2}x^2+o(x^2)-ax-bx^2}{x^2}=\lim\limits_{x\to0}\frac{(1-a)x-\left(\frac{1}{2}+b\right)x^2+o(x^2)}{x^2}=2$,从而

$$1-a=0,-\left(\frac{1}{2}+b\right)=2\Rightarrow a=1,b=-\frac{5}{2}.$$

【注】 本题还有一个值得借鉴的解法.

根据泰勒公式容易得：$x-\ln(1+x)\sim\frac{1}{2}x^2(x\to0)$（请读者记住此式），故想办法把分子凑出这形式.

$$2=\lim_{x\to0}\frac{\ln(1+x)-(ax+bx^2)}{x^2}=-\lim_{x\to0}\frac{x-\ln(1+x)-x+(ax+bx^2)}{x^2}$$

$$=-\lim_{x\to0}\frac{x-\ln(1+x)}{x^2}-\lim_{x\to0}\frac{(a-1)x}{x^2}-\lim_{x\to0}\frac{bx^2}{x^2}=-\frac{1}{2}-b-\lim_{x\to0}\frac{(a-1)x}{x^2},$$

故 $a-1=0\Rightarrow a=1$. 因而 $-\frac{1}{2}-b=2\Rightarrow b=-\frac{5}{2}$.

5. 无穷小比阶

例 1.3.17 设当 $x\to0$ 时，$e^x-(ax^2+bx+1)$是比 x^2 高阶的无穷小，则（　　）.

(A)$a=\frac{1}{2},b=1$　　(B)$a=1,b=1$　　(C)$a=-\frac{1}{2},b=-1$　　(D)$a=-1,b=1$

解　应选(A).

方法一　由泰勒公式可知　$e^x=1+x+\frac{x^2}{2!}+o(x^2)$.

由题设可知
$$\lim_{x\to0}\frac{e^x-(ax^2+bx+1)}{x^2}=0,$$

即
$$\lim_{x\to0}\frac{\left(\frac{1}{2}-a\right)x^2+(1-b)x+o(x^2)}{x^2}=0.$$

则 $a=\frac{1}{2},b=1$.

方法二　由洛必达法则可知
$$\lim_{x\to0}\frac{e^x-(ax^2+bx+1)}{x^2}=\lim_{x\to0}\frac{e^x-2ax-b}{2x},$$

若 $b\neq1$，上式右端趋于无穷，从而左端也趋于无穷，与原题设矛盾，所以 $b=1$. 因此
$$\lim_{x\to0}\frac{e^x-2ax-b}{2x}=\lim_{x\to0}\frac{e^x-1}{2x}-a=\frac{1}{2}-a=0,$$

故 $a=\frac{1}{2}$，所以应选(A).

例 1.3.18 设 $x\to0$ 时，$e^{\tan x}-e^x$ 与 x^n 是同阶无穷小，则 n 为（　　）.

(A)1　　(B)2　　(C)3　　(D)4

解　应选(C).

当 $x\to0$ 时，
$$e^{\tan x}-e^x=e^x(e^{\tan x-x}-1)\sim e^x(\tan x-x)\sim\tan x-x,$$

而
$$\tan x=x+\frac{x^3}{3}+o(x^3),$$

所以

$$\tan x-x=\frac{x^3}{3}+o(x^3),$$

因此选(C).

【注】 可以仿照解答来验证如下结论：

当 $x\to x_0$ 时，若 $f(x)\to0,g(x)\to0$，则 $e^{f(x)}-e^{g(x)}\sim f(x)-g(x)$.

例 1.3.19 设 $\alpha_1=x(\cos\sqrt{x}-1),\alpha_2=\sqrt{x}\ln(1+\sqrt[3]{x}),\alpha_3=\sqrt[3]{x+1}-1$. 当 $x\to0^+$ 时，以上 3 个无穷小量按照从低阶到高阶的排序是(　　).

(A)$\alpha_1,\alpha_2,\alpha_3$　　(B)$\alpha_2,\alpha_3,\alpha_1$　　(C)$\alpha_2,\alpha_1,\alpha_3$　　(D)$\alpha_3,\alpha_2,\alpha_1$

解　应选(B).

当 $x\to0^+$ 时，

$$\alpha_1=x(\cos\sqrt{x}-1)\sim-\frac{1}{2}x^2,$$

$$\alpha_2=\sqrt{x}\ln(1+\sqrt[3]{x})\sim x^{\frac{5}{6}},$$

$$\alpha_3=\sqrt[3]{x+1}-1\sim\frac{1}{3}x.$$

则从低阶到高阶排序是 $\alpha_2,\alpha_3,\alpha_1$，故选(B).

6. 函数的连续与间断

例 1.3.20 设 $f(x)=\begin{cases}2x+a, & x\leqslant0,\\ e^x(\sin x+\cos x), & x>0\end{cases}$ 在$(-\infty,+\infty)$内连续，则 $a=$________.

解　应填 1.

由于　　$f(0^-)=\lim\limits_{x\to0^-}(2x+a)=a,\quad f(0^+)=\lim\limits_{x\to0^+}e^x(\sin x+\cos x)=1.$

要使 $f(x)$在$(-\infty,+\infty)$上连续，只需 $a=1=f(0)$，即 $a=1$.

例 1.3.21 设 $f(x)=\begin{cases}a+bx^2, & x\leqslant0,\\ \dfrac{\sin bx}{x}, & x>0\end{cases}$ 在 $x=0$ 处连续，则常数 a 与 b 应满足的关系是________.

解　应填 $a=b$.

由于

$$f(0^+)=\lim_{x\to0^+}\frac{\sin bx}{x}=b,\quad f(0^-)=\lim_{x\to0^-}(a+bx^2)=a,\quad f(0)=a,$$

要使 $f(x)$在$x=0$ 处连续，则应有 $f(0^+)=f(0^-)=f(0)$，即 $a=b$.

例 1.3.22 设函数 $f(x)=\dfrac{\ln|x|}{|x-1|}\sin x$，则 $f(x)$有(　　).

(A)1 个可去间断点，1 个跳跃间断点　　(B)1 个可去间断点，1 个无穷间断点

(C)2 个跳跃间断点　　(D)2 个无穷间断点

解　应选(A).

$x=0,x=1$ 是间断点. 因为

$$\lim_{x\to0}\frac{\ln|x|}{|x-1|}\sin x=\lim_{x\to0}\frac{\ln|x|}{\csc x}=\lim_{x\to0}\frac{\frac{1}{x}}{-\csc x\cdot\cot x}=\lim_{x\to0}\frac{-\sin^2x}{x\cos x}=0,$$

故 $x=0$ 是可去间断点；而

$$\lim_{x\to1^-}\frac{\ln|x|}{|x-1|}\sin x=\sin 1\cdot\lim_{x\to1^-}\frac{\ln[1+(x-1)]}{1-x}=\sin 1\cdot\lim_{x\to1^-}\frac{x-1}{1-x}=-\sin 1,$$

$$\lim_{x\to1^+}\frac{\ln|x|}{|x-1|}\sin x=\sin 1\cdot\lim_{x\to1^+}\frac{\ln[1+(x-1)]}{x-1}=\sin 1\cdot\lim_{x\to1^+}\frac{x-1}{x-1}=\sin 1,$$

故 $x=1$ 是跳跃间断点.

【注】 $\lim\limits_{x\to0}\frac{\ln|x|}{|x-1|}\sin x=\lim\limits_{x\to0}x\ln|x|=0.$

例 1.3.23 函数 $f(x)=\frac{x^2-x}{x^2-1}\sqrt{1+\frac{1}{x^2}}$ 的无穷间断点的个数为(　　).

(A)0　　(B)1　　(C)2　　(D)3

解 应选(B).

$f(x)$为非分段函数，只需讨论函数的无定义点：$x=0,1,-1$，又

$$\lim_{x\to0^-}f(x)=\lim_{x\to0^-}\left(-\frac{\sqrt{x^2+1}}{x+1}\right)=-1,\quad \lim_{x\to0^+}f(x)=\lim_{x\to0^+}\frac{\sqrt{x^2+1}}{x+1}=1,$$

$$\lim_{x\to1}f(x)=\lim_{x\to1}\frac{\sqrt{x^2+1}}{x+1}=\frac{\sqrt{2}}{2},\quad \lim_{x\to-1}f(x)=\lim_{x\to-1}\left(-\frac{\sqrt{x^2+1}}{x+1}\right)=\infty.$$

故 $x=0$ 是 $f(x)$的跳跃间断点，$x=1$ 是 $f(x)$的可去间断点，$x=-1$ 是 $f(x)$的无穷间断点，故选择(B).

例 1.3.24 求极限 $\lim\limits_{t\to x}\left(\frac{\sin t}{\sin x}\right)^{\frac{x}{\sin t-\sin x}}$，记此极限为 $f(x)$，求函数 $f(x)$的间断点并指出其类型.

解 此极限是“1^∞”型未定式，用公式 $\lim u^v=\mathrm{e}^{\lim(u-1)v}$，有

$$\begin{aligned}f(x)&=\exp\left\{\lim_{t\to x}\frac{x}{\sin t-\sin x}\left(\frac{\sin t}{\sin x}-1\right)\right\}\\&=\exp\left\{\lim_{t\to x}\frac{x}{\sin t-\sin x}\cdot\frac{\sin t-\sin x}{\sin x}\right\}=\mathrm{e}^{\frac{x}{\sin x}}.\end{aligned}$$

函数无分段，只讨论无定义点即可：

①$x=0$ 是函数的无定义点，由于 $\lim\limits_{x\to0}f(x)=\lim\limits_{x\to0}\mathrm{e}^{\frac{x}{\sin x}}=\mathrm{e}$，故 $x=0$ 是 $f(x)$的第一类间断点；

②$x=k\pi(k=\pm1,\pm2,\cdots)$也是函数的无定义点，由于 $f(x)$在 $x=k\pi$ 时的左极限和右极限中总有一个为$+\infty$，故$x=k\pi(k=\pm1,\pm2,\cdots)$是第二类间断点(详见“基础内容精讲”“二”中的“2”).

基础习题精练

习　题

1.3.1 已知 $I=\lim\limits_{x\to0}\left(\frac{\ln(1+\mathrm{e}^{\frac{2}{x}})}{\ln(1+\mathrm{e}^{\frac{1}{x}})}+a[x]\right)$存在，[·]为取整函数，求 I,a.

1.3.2 已知 $a>0,b>0$，则 $\lim\limits_{x\to+\infty}x(a^{\frac{1}{x}}-b^{\frac{1}{x}})=$ ________.

1.3.3　求极限$\lim\limits_{x\to 0}\left(\dfrac{e^x+xe^x}{e^x-1}-\dfrac{1}{x}\right)$.

1.3.4　设$a\neq\dfrac{1}{2}$，计算$\lim\limits_{n\to\infty}\ln\left[\dfrac{n-2na+1}{n(1-2a)}\right]^n$.

1.3.5　求极限$\lim\limits_{x\to 0^+}\dfrac{x^x-(\sin x)^x}{x^2\ln(1+x)}$.

1.3.6　求极限$\lim\limits_{x\to 0}\left(\dfrac{a_1^x+a_2^x+\cdots+a_n^x}{n}\right)^{\frac{n}{x}}$，其中$a_i>0,i=1,2,\cdots,n$.

1.3.7　设$\lim\limits_{x\to 0}\dfrac{\ln\left[1+\dfrac{f(x)}{\sin x}\right]}{2^x-1}=2$，求$\lim\limits_{x\to 0}\dfrac{f(x)}{x^2}$.

1.3.8　设$f(x)$在$(-\infty,+\infty)$内有定义，且$\lim\limits_{x\to\infty}f(x)=a$，$g(x)=\begin{cases}f\left(\dfrac{1}{x}\right), & x\neq 0,\\ 0, & x=0,\end{cases}$则(　　).

(A)$x=0$必是$g(x)$的第一类间断点　　(B)$x=0$必是$g(x)$的第二类间断点

(C)$x=0$必是$g(x)$的连续点　　(D)$g(x)$在点$x=0$处的连续性与a的取值有关

1.3.9　设函数$f(x)=\lim\limits_{n\to\infty}\dfrac{1+x}{1+x^{2n}}$，讨论函数的间断点，其结论为(　　).

(A)不存在间断点　　(B)存在间断点$x=1$　　(C)存在间断点$x=0$　　(D)存在间断点$x=-1$

解　答

1.3.1　解

$$\lim_{x\to 0^-}\frac{\ln(1+e^{\frac{2}{x}})}{\ln(1+e^{\frac{1}{x}})}\xlongequal{u=\frac{1}{x}}\lim_{u\to-\infty}\frac{\dfrac{2e^{2u}}{1+e^{2u}}}{\dfrac{e^u}{1+e^u}}=0,\quad \lim_{x\to 0^-}a[x]=-a,$$

$$\lim_{x\to 0^+}\frac{\ln(1+e^{\frac{2}{x}})}{\ln(1+e^{\frac{1}{x}})}\xlongequal{u=\frac{1}{x}}\lim_{u\to+\infty}\frac{\dfrac{2e^{2u}}{1+e^{2u}}}{\dfrac{e^u}{1+e^u}}=\lim_{u\to+\infty}\frac{2(e^u+e^{2u})}{1+e^{2u}}=\lim_{u\to+\infty}\frac{2(e^{-u}+1)}{1+e^{-2u}}=2,\quad \lim_{x\to 0^+}a[x]=0,$$

所以当且仅当$a=-2$时，原极限存在，此时$I=2$.

1.3.2　$\ln\dfrac{a}{b}$　**解**　利用变量代换．令$\dfrac{1}{x}=t$，则

$$\lim_{x\to+\infty}x(a^{\frac{1}{x}}-b^{\frac{1}{x}})=\lim_{t\to 0^+}\frac{a^t-b^t}{t}=\lim_{t\to 0^+}(a^t\ln a-b^t\ln b)=\ln\frac{a}{b}(a>0,b>0).$$

1.3.3　解　方法一　这是“$\infty-\infty$”型未定式极限，首先通分变成“$\dfrac{0}{0}$”或“$\dfrac{\infty}{\infty}$”型未定式，然后使用洛必达法则求极限.

$$\lim_{x\to 0}\left(\frac{e^x+xe^x}{e^x-1}-\frac{1}{x}\right)=\lim_{x\to 0}\frac{xe^x(1+x)+1-e^x}{x(e^x-1)}=\lim_{x\to 0}\frac{3xe^x+x^2e^x}{e^x+xe^x-1}$$
$$=\lim_{x\to 0}\frac{3e^x+5xe^x+x^2e^x}{2e^x+xe^x}=\frac{3}{2}.$$

方法二　利用等价无穷小量代换，$e^x-1\sim x$(当$x\to 0$时)，则

$$\lim_{x\to0}\left(\frac{\mathrm{e}^x+x\mathrm{e}^x}{\mathrm{e}^x-1}-\frac{1}{x}\right)=\lim_{x\to0}\frac{x\mathrm{e}^x(1+x)+1-\mathrm{e}^x}{x(\mathrm{e}^x-1)}=\lim_{x\to0}\frac{x\mathrm{e}^x(1+x)+1-\mathrm{e}^x}{x^2}$$

$$=\lim_{x\to0}\frac{3x\mathrm{e}^x+x^2\mathrm{e}^x}{2x}=\lim_{x\to0}\frac{3\mathrm{e}^x+x\mathrm{e}^x}{2}=\frac{3}{2}.$$

【注】 下面的方法是错误的：

$$\lim_{x\to0}\left(\frac{\mathrm{e}^x+x\mathrm{e}^x}{\mathrm{e}^x-1}-\frac{1}{x}\right)=\lim_{x\to0}\left(\frac{\mathrm{e}^x+x\mathrm{e}^x}{x}-\frac{1}{x}\right)$$

$$=\lim_{x\to0}\frac{\mathrm{e}^x+x\mathrm{e}^x-1}{x}=\lim_{x\to0}(2\mathrm{e}^x+x\mathrm{e}^x)=2.$$

1.3.4 分析 $\lim\limits_{n\to\infty}\left[\frac{n-2na+1}{n(1-2a)}\right]^n$ 是“1^∞”型未定式极限，可以使用洛必达法则，也可以凑成第二个重要极限，还可以利用等价无穷小量代换.

解 方法一 $\lim\limits_{n\to\infty}\ln\left[\frac{n-2na+1}{n(1-2a)}\right]^n=\lim\limits_{x\to+\infty}\ln\left[\frac{x-2ax+1}{x(1-2a)}\right]^x=\lim\limits_{x\to+\infty}x\ln\left[1+\frac{1}{x(1-2a)}\right]$

$$\xlongequal{令\frac{1}{x}=t}\lim_{t\to0^+}\frac{\ln\left(1+\frac{t}{1-2a}\right)}{t}\xlongequal{洛必达法则}\lim_{t\to0^+}\frac{1}{1+\frac{t}{1-2a}}\cdot\frac{1}{1-2a}=\frac{1}{1-2a}.$$

方法二 $\lim\limits_{n\to\infty}\ln\left[\frac{n-2na+1}{n(1-2a)}\right]^n=\lim\limits_{n\to\infty}\ln\left[1+\frac{1}{n(1-2a)}\right]^n=\ln\left\{\lim\limits_{n\to\infty}\left[1+\frac{1}{n(1-2a)}\right]^n\right\}$

$$=\ln\left\{\lim_{n\to\infty}\left[1+\frac{1}{n(1-2a)}\right]^{n(1-2a)\frac{1}{1-2a}}\right\}=\ln\mathrm{e}^{\frac{1}{1-2a}}=\frac{1}{1-2a}.$$

方法三 因为 $\ln\left[1+\frac{1}{n(1-2a)}\right]\sim\frac{1}{n(1-2a)}(n\to\infty)$，所以

$$\lim_{n\to\infty}\ln\left[\frac{n-2na+1}{n(1-2a)}\right]^n=\lim_{n\to\infty}n\cdot\frac{1}{n(1-2a)}=\frac{1}{1-2a}.$$

1.3.5 解 原式 $=-\lim\limits_{x\to0^+}\frac{x^x\left[\left(\frac{\sin x}{x}\right)^x-1\right]}{x^3}=-\lim\limits_{x\to0^+}\frac{\mathrm{e}^{x\ln\frac{\sin x}{x}}-1}{x^3}(\lim\limits_{x\to0^+}x^x=1)$

$$=-\lim_{x\to0^+}\frac{\ln\frac{\sin x}{x}}{x^2}=-\lim_{x\to0^+}\frac{\ln\left(1+\frac{\sin x-x}{x}\right)}{x^2}$$

$$=-\lim_{x\to0^+}\frac{\sin x-x}{x^3}=-\lim_{x\to0^+}\frac{x-\frac{1}{3!}x^3+o(x^3)-x}{x^3}=\frac{1}{6}.$$

1.3.6 解 因为 $\lim\limits_{x\to0}a_i^x=1$，所以原极限为“$1^\infty$”型未定式.

方法一 使用洛必达法则求极限.

$$\lim_{x\to0}\left(\frac{a_1^x+a_2^x+\cdots+a_n^x}{n}\right)^{\frac{n}{x}}=\exp\left\{\lim_{x\to0}\frac{n}{x}\ln\left(\frac{a_1^x+a_2^x+\cdots+a_n^x}{n}\right)\right\}$$

$$=\exp\left\{\lim_{x\to0}n\cdot\frac{a_1^x\ln a_1+a_2^x\ln a_2+\cdots+a_n^x\ln a_n}{a_1^x+a_2^x+\cdots+a_n^x}\right\}=a_1a_2\cdots a_n.$$

方法二 凑成第二个重要极限（“1^∞”型未定式极限都可以凑成第二个重要极限），在计算过程中使用洛必达法则.

$$\lim_{x\to0}\left(\frac{a_1^x+a_2^x+\cdots+a_n^x}{n}\right)^{\frac{n}{x}}=\lim_{x\to0}\left(1+\frac{a_1^x+a_2^x+\cdots+a_n^x-n}{n}\right)^{\frac{n}{a_1^x+a_2^x+\cdots+a_n^x-n}\cdot\frac{a_1^x+a_2^x+\cdots+a_n^x-n}{x}},$$

其中

$$\lim_{x\to0}\frac{a_1^x+a_2^x+\cdots+a_n^x-n}{x}=\ln(a_1a_2\cdots a_n),\quad \lim_{x\to0}\left(1+\frac{a_1^x+a_2^x+\cdots+a_n^x-n}{n}\right)^{\frac{n}{a_1^x+a_2^x+\cdots+a_n^x-n}}=\mathrm{e},$$

所以原式$=a_1a_2\cdots a_n$.

1.3.7　**解**　$\lim\limits_{x\to0}(2^x-1)=0\Rightarrow\lim\limits_{x\to0}\ln\left[1+\dfrac{f(x)}{\sin x}\right]=0\Rightarrow\lim\limits_{x\to0}\dfrac{f(x)}{\sin x}=0\Rightarrow\ln\left[1+\dfrac{f(x)}{\sin x}\right]\sim\dfrac{f(x)}{\sin x}\sim\dfrac{f(x)}{x}$(当$x\to0$时).

又$2^x-1=\mathrm{e}^{x\ln 2}-1\sim x\ln 2$(当$x\to0$时)，所以原式变为$\lim\limits_{x\to0}\dfrac{f(x)}{x^2\ln 2}=2$，故$\lim\limits_{x\to0}\dfrac{f(x)}{x^2}=\ln 4$.

1.3.8　(D)　**解**　因为$\lim\limits_{x\to0}g(x)=\lim\limits_{x\to0}f\left(\dfrac{1}{x}\right)=\lim\limits_{x\to\infty}f(x)=a$，而$g(0)=0$，所以，当$a=0$时，函数$g(x)$在点$x=0$处连续；当$a\neq0$时，函数$g(x)$在点$x=0$处间断，故应选(D).

1.3.9　(B)　**分析**　函数$f(x)=\lim\limits_{n\to\infty}\dfrac{1+x}{1+x^{2n}}$是以$x$为自变量的函数，但是，在$n\to\infty$求极限时，$x$则被看成一个常数(参数)，根据$x$的不同取值求出极限，极限完成后$x$又恢复变量的本来身份.

因为分式中有x^{2n}，所以应该把$x=-1$和$x=1$作为分段点将函数变成分段函数，然后讨论函数的间断点.

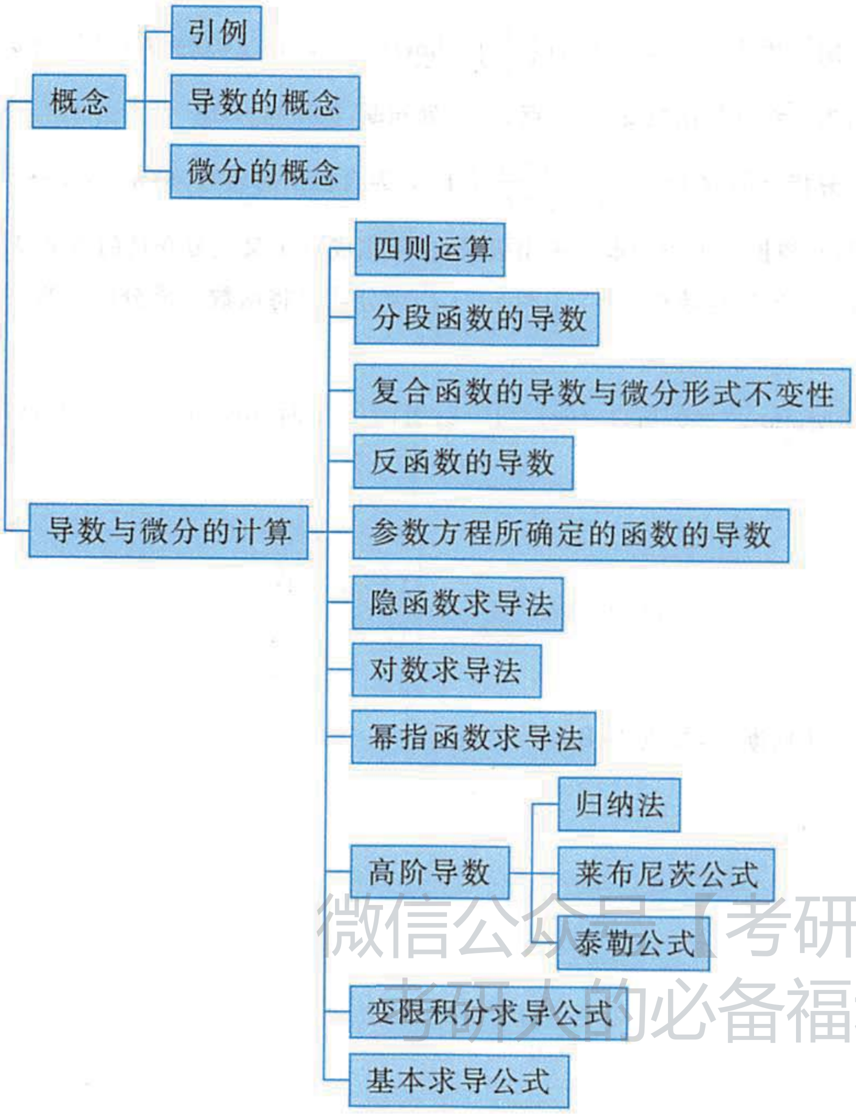
解　当$|x|<1$时，$\lim\limits_{n\to\infty}x^{2n}=0$，所以$f(x)=1+x$；当$|x|>1$时，$\lim\limits_{n\to\infty}\dfrac{1+x}{1+x^{2n}}=0$. 又$f(1)=1$，$f(-1)=0$，所以

$$f(x)=\lim_{n\to\infty}\frac{1+x}{1+x^{2n}}=\begin{cases}0, & x\leqslant-1,\\ 1+x, & -1<x<1,\\ 1, & x=1,\\ 0, & x>1.\end{cases}$$

由此可知$x=1$为间断点，故应选(B).

第4讲 一元函数微分学的概念与计算

基础知识结构



基础内容精讲

一、概念

1. 引例

让我们进入一个普通教室，注意，“普通”二字意味着这不是一个特殊的（比如恒温）教室．假设 $t=9:00$ 时教室里的温度 $u=20$ ℃，$\tilde{t}=9:05$ 时教室里的温度 $\tilde{u}=25$ ℃，问教室里的温度在这 5 min 里的**平均变化率**是多少？

这是一个极其简单的问题，其平均变化率为

$$\frac{\Delta u}{\Delta t}=\frac{\tilde{u}-u}{\tilde{t}-t}=1(℃/\text{min}).$$

不过，若只研究“平均”变化率，有时会出现意外．再次回到刚才的教室，假设现在这个时刻 t，教室里的温度 $u=20$ ℃，一年之前的时刻 $\tilde{t}$，教室里的温度 $\tilde{u}=20$ ℃，则其平均变化率为

$$\frac{\Delta u}{\Delta t}=\frac{\tilde{u}-u}{\tilde{t}-t}=0(℃/\text{min}).$$

显然，这是一个令人失望的结果．一般来说，普通教室里的温度随时间变化而不断变化，试图研究教室温度变化的这个目的落空了．究其原因，是因为平均变化率这个概念是“粗糙”的——只研究起点和终点的信息，完全没有考虑其过程，当然达不到精确研究温度变化率的目的，如何补救？这时，我们让 $\Delta t\to 0$，也就是让时间差成为无穷小的数，写成

$$\lim_{\Delta t\to 0}\frac{\Delta u}{\Delta t}\overset{\text{记}}{=\!=}u'(t),$$

这里的 t 被抽象为某一时刻，则 $u'(t)$ 这个符号所表达的极限 $\lim\limits_{\Delta t\to 0}\dfrac{\Delta u}{\Delta t}$ 称为教室温度 u 在 t 时刻的**瞬时变化率**．由于 t 为泛指的时刻点，则 $u'(t)$ 即可表达“时时刻刻”教室温度的（瞬时）变化率．这就达到了精确研究温度在不同时刻的变化率的目的．

按照上面这个思路，我们还可以研究质点在直线运动中的速度问题，给出质点的运动方程 $S=S(t)$，我们要求出质点位移在任意时刻 t 的瞬时变化率（这就是读者熟知的速度），也即

$$\lim_{\Delta t\to 0}\frac{\Delta S}{\Delta t}\overset{\text{记}}{=\!=}S'(t)=v(t).$$

历史上，以牛顿为代表的数学家们饶有兴趣地研究了上面的速度问题，虽然他们的表述方式与符号记法与今天已大相径庭，但是对数学的发展做出了巨大贡献．不过，相对于牛顿，莱布尼茨对几何问题更感兴趣．

接下来我们就进入莱布尼茨的领域——作曲线的切线——给定曲线 l 及其上一点 P（如图 1-4-1），建立曲线 l 上点 P 处的切线的概念．

在曲线 l 上另取一个异于点 P 的点 Q，作割线 PQ，如图 1-4-1 所示．当点 Q 沿着曲线 l 移动时，割线 PQ 将绕点 P 旋转，现让点 Q 沿曲线 l 趋向于点 P，则割线 PQ 绕点 P 旋转而趋向于**极限位置** PT．直线 PT 称为曲线 l 在点 P 处的**切线**，此处

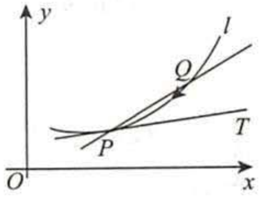


图 1-4-1

的极限位置的含义为只要弦长$|PQ|\to 0$，则$\angle QPT\to 0$.

进一步地，记曲线$y=f(x)$，设$P(x_0,y_0)$，$Q(x,y)$，割线PQ的斜率为

$$\tan\theta=\frac{y-y_0}{x-x_0}=\frac{f(x)-f(x_0)}{x-x_0},$$

其中θ为割线PQ的倾角.

当点Q沿曲线$y=f(x)$趋向于点P时，自然有$x\to x_0$，于是

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}\overset{\text{记}}{=\!=}k,$$

即为割线斜率的极限，即**切线的斜率**. 如图1-4-2所示，$k=\tan\alpha$，其中α为切线PT的倾角.

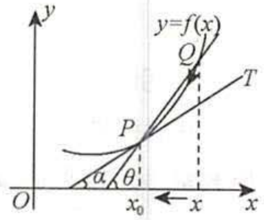


图 1-4-2

这里有两点需要说明.

①前面谈到，极限位置的含义是只要弦长$|PQ|\to 0$，则$\angle QPT\to 0$. 我们不妨验证一下，在图1-4-2中，$\angle QPT=\theta-\alpha$，当$x\to x_0$时，$\theta\to\alpha$. 显然，点Q沿曲线趋向点P时，弦长$|PQ|\to 0$，此时有$x\to x_0$，即$\theta\to\alpha$，则$\angle QPT\to 0$，故直线PT为曲线$y=f(x)$在点P处的切线.

②$\lim\limits_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$并不一定存在，图1-4-2当然是为方便读者理解而画出的.

2. 导数的概念

数学工作者们总希望透过形形色色的具体问题看到数学本质，这在技术上通常的做法是撇开问题中各种变量的具体意义，归纳出抽象化、统一化的数学表述，从而建立相应的数学概念.

事实上，上面的引例本质上是做了一件事：计算“$\frac{\text{函数的增量}}{\text{自变量的增量}}$”的极限，这就会得出导数的概念.

设$y=f(x)$定义在区间I上，让自变量在$x=x_0$处加一个增量Δx(可正可负)，其中$x_0\in I$，$x_0+\Delta x\in I$，则可得函数的增量$\Delta y=f(x_0+\Delta x)-f(x_0)$. 若函数增量$\Delta y$与自变量增量$\Delta x$的比值在$\Delta x\to 0$时的极限存在，即$\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}$存在，则称函数$y=f(x)$在点$x_0$处**可导**，并称这个极限为$y=f(x)$在点$x_0$处的**导数**，记作$f'(x_0)$，即

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}. \tag{1-4-1}$$

当然，$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=x_0}$，$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0}$，$y'(x_0)$或$\left.y'\right|_{x=x_0}$这些符号记法与$f'(x_0)$等价. 顺便交代一下，“导数”这个名词被认为是拉格朗日最先使用的，记号$f'(x_0)$，$\left.y'\right|_{x=x_0}$多次出现在拉格朗日的文章中，而莱布尼茨则喜欢写作$\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=x_0}$，$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0}$.

【注】 这里有几点需要说明.

(1)在考题中，增量Δx一般会被命题人广义化为“狗”：

$$f'(x_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\overset{\text{广义化}}{=\!=\!=}\lim_{\text{狗}\to 0}\frac{f(x_0+\text{狗})-f(x_0)}{\text{狗}}. \tag{1-4-2}$$

(2)若在(1-4-1)式中，令$x_0+\Delta x=x$，则可将导数定义式写成

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}, \tag{1-4-3}$$

(1-4-1),(1-4-3)两式等价,读者将会在各种场合见到上面这两种等价写法.

(3)下面这三种提法是等价的.

(ⅰ)$y=f(x)$在点 x_0 处可导;

(ⅱ)$y=f(x)$在点 x_0 处导数存在;

(ⅲ)$f'(x_0)=A$(A 为有限数).

(4)让我们回顾引例中的例子.定义 $u=u(t)$为教室的温度函数,则

$$u'(t)=\lim_{\Delta t\to 0}\frac{\Delta u}{\Delta t}$$

表示 u 在 t 时刻的瞬时变化率.再仔细一些,当 t 取为 9:00 时,读者是否想到这样一个问题:对于 9 点整温度的瞬时变化率,这个"瞬时"到底是 9 **点整之前**的瞬时,**还是** 9 **点整之后**的瞬时呢? 请读者放心,这在数学上,是有概念上的完美对应的:

若 $\Delta t<0$,则 $t+\Delta t$ 在 $\Delta t\to 0^-$ 时,表达了 t 时刻之前的瞬时;

若 $\Delta t>0$,则 $t+\Delta t$ 在 $\Delta t\to 0^+$ 时,表达了 t 时刻之后的瞬时.

于是便有了**单侧导数**的概念,

$$\lim_{\Delta x\to 0^-}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\overset{\text{记}}{=\!=\!=}f'_-(x_0),$$

$$\lim_{\Delta x\to 0^+}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\overset{\text{记}}{=\!=\!=}f'_+(x_0),$$

这里,$f'_-(x_0)$,$f'_+(x_0)$分别是 $f(x)$在点 x_0 处的**左导数**、**右导数**,统称为**单侧导数**.

我们要说,函数 $f(x)$在点 x_0 处可导的**充分必要条件**是其左导数 $f'_-(x_0)$与右导数 $f'_+(x_0)$均存在且相等,这一点当然是与极限存在的充分必要条件(左、右极限均存在且相等)对应.因为从本质上来说,导数的定义就是一个极限问题.有了上面这些说法,读者就会回答刚才提出的关于引例的更为细致的问题了.

$u'_-(t)=\lim\limits_{\Delta t\to 0^-}\dfrac{\Delta u}{\Delta t}$表示 t 时刻之前的温度瞬时变化率;

$u'_+(t)=\lim\limits_{\Delta t\to 0^+}\dfrac{\Delta u}{\Delta t}$表示 t 时刻之后的温度瞬时变化率.

而温度 u 在时刻 t 的瞬时变化率存在的充要条件为 $u'_-(t)=u'_+(t)=A$(A 为有限数).

下面的例子将从几何上进一步加深我们对导数的理解.

注例 1 研究 $y=f(x)=|x|$在 $x=0$ 处的切线问题.

解 从 $x=0$ 出发,取增量 Δx,有 $\Delta y=f(0+\Delta x)-f(0)=|\Delta x|$.

当 $\Delta x>0$ 时,$\Delta y=\Delta x$,则 $f'_+(0)=\lim\limits_{\Delta x\to 0^+}\dfrac{\Delta y}{\Delta x}=1\overset{\text{记}}{=\!=\!=}k_+$;

当 $\Delta x<0$ 时,$\Delta y=-\Delta x$,则 $f'_-(0)=\lim\limits_{\Delta x\to 0^-}\dfrac{\Delta y}{\Delta x}=-1\overset{\text{记}}{=\!=\!=}k_-$.

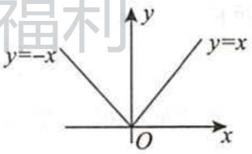


图 1-4-3

如图 1-4-3 所示,函数曲线 $y=f(x)=|x|$在原点 O 处出现了两条**单侧切线**,这两条单侧切线形成了一个角,数学上称这里的原点 O 为一个**角点**.不过,虽然此曲线在角点 O 处有两条单侧切线,但按照前面讲到的$f(x)$在点 x_0 处可导的充要条件,这里的 $f'_+(0)=k_+\neq k_-=f'_-(0)$,显然 $f'(0)$不存在,所以我们说 $y=f(x)=|x|$在原点 O 处不可导,也就没有切线.

注例 2 研究 $y=f(x)=x^{\frac{1}{3}}$ 在 $x=0$ 处的切线问题.

解 显然，在 $x=0$ 处 $\frac{\Delta y}{\Delta x}=\frac{f(0+\Delta x)-f(0)}{\Delta x}=\frac{(\Delta x)^{\frac{1}{3}}}{\Delta x}=\frac{1}{(\Delta x)^{\frac{2}{3}}}$.

当 $\Delta x>0$ 时，$f'_{+}(0)=\lim\limits_{\Delta x\to 0^{+}}\frac{1}{(\Delta x)^{\frac{2}{3}}}=+\infty$；

当 $\Delta x<0$ 时，$f'_{-}(0)=\lim\limits_{\Delta x\to 0^{-}}\frac{1}{(\Delta x)^{\frac{2}{3}}}=+\infty$.

这样的结果称为**无穷导数**. 又 $\pm\infty$ 被叫作广义的数，所以无穷导数在有些数学场合也可被视为导数存在的特殊情形. 不过要强调的是，学习“高等数学”这门课程的读者，还是将无穷导数视为导数不存在为好，因为这是“高等数学”里的“规矩”.

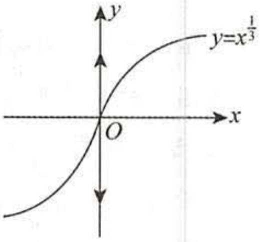


图 1-4-4

还要指出，如图 1-4-4 所示，$y=f(x)=x^{\frac{1}{3}}$ 在 $x=0$ 处是有垂直于 x 轴的切线 $x=0$ 的. 我们说，若曲线 $y=f(x)$ 在点 $P(x_0,y_0)$ 处有垂直于 x 轴的切线，则等价于

$$f'(x_0)=+\infty \text{或} -\infty(\text{为无穷导数}),$$

在“高等数学”中，$\pm\infty$ 被视为不存在的情形下，读者就不要把“可导”与“光滑”完全对等起来了. 因为这个例子中，显然曲线 $y=x^{\frac{1}{3}}$ 在点 $x=0$ 处有切线，且为“光滑”的，但在这一点处却不可导.

(5) 导数的几何意义.

$y=f(x)$ 在点 x_0 处的导数值 $f'(x_0)$ 就是曲线 $y=f(x)$ 在点 (x_0,y_0) 处切线的斜率 k，即 $k=f'(x_0)$，于是曲线 $y=f(x)$ 在点 (x_0,y_0) 处的切线方程为 $y-y_0=f'(x_0)(x-x_0)$；

法线方程为
$$y-y_0=-\frac{1}{f'(x_0)}(x-x_0)\ (f'(x_0)\neq 0).$$

(6) 高阶导数的概念.

$$f^{(n)}(x_0)=\lim_{\Delta x\to 0}\frac{f^{(n-1)}(x_0+\Delta x)-f^{(n-1)}(x_0)}{\Delta x},$$

其中 $n\geqslant 2$，为整数，且 $f^{(n-1)}(x)$ 在 x_0 的某邻域内有定义，$x_0+\Delta x$ 也在该邻域内.

3. 微分的概念

先说个引例. 如图 1-4-5 所示，设正方形边长为 x，当其边长增加 Δx 时，它的面积 S 增加了

$$\Delta S=(x+\Delta x)^2-x^2=2x\Delta x+(\Delta x)^2.$$

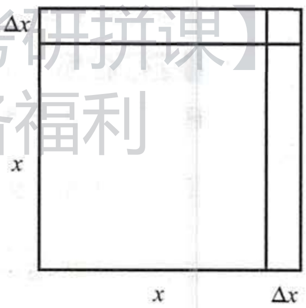


图 1-4-5

上述面积的增量 ΔS 由两部分组成，一部分是 $2x\Delta x$（图 1-4-5 中两个长方形面积），它是 Δx 的一次项；另一部分是 $(\Delta x)^2$（图 1-4-5 中右上角小正方形的面积），它满足 $\lim\limits_{\Delta x\to 0}\frac{(\Delta x)^2}{\Delta x}=0$，即 $(\Delta x)^2=o(\Delta x)$. 故 $\Delta S=2x\Delta x+o(\Delta x)$，$2x\Delta x$ 为增量的主要部分，$o(\Delta x)$ 为高阶无穷小，是误差，当 Δx 足够小时，有 $\Delta S\approx 2x\Delta x$. 下面给出微分的定义.

设函数 $y=f(x)$ 在点 x_0 的某邻域内有定义，且 $x_0+\Delta x$ 在该邻域内，对于函数增量

$$\Delta y=f(x_0+\Delta x)-f(x_0),$$

若存在与 Δx 无关的常数 A，使得 $$\Delta y=A\Delta x+o(\Delta x), \tag{1-4-4}$$

其中 $o(\Delta x)$ 是在 $\Delta x\to 0$ 时比 Δx 更高阶的无穷小，则称 $f(x)$ 在点 x_0 处**可微**，并称 $A\Delta x$ 为 $f(x)$ 在点 x_0 处的**微分**，记作 $\mathrm{d}y\Big|_{x=x_0}=A\Delta x$ 或者 $\mathrm{d}f(x)\Big|_{x=x_0}=A\Delta x$. 又 $\Delta x=\mathrm{d}x$，故 $\mathrm{d}y\Big|_{x=x_0}=A\mathrm{d}x$.　(1-4-5)

【注】　(1)可微的判别.

①写增量 $\Delta y=f(x_0+\Delta x)-f(x_0)$；

②写线性增量 $A\Delta x=f'(x_0)\Delta x$；

③作极限 $\lim\limits_{\Delta x\to 0}\dfrac{\Delta y-A\Delta x}{\Delta x}$.

若该极限等于 0，则 $y=f(x)$ 在点 x_0 处可微，否则不可微.

(2)从上述判别步骤可以看出，用形式简单的"线性增量 $A\Delta x$"去代替形式复杂的"增量 Δy"，且其误差"$\Delta y-A\Delta x$"是 $o(\Delta x)$，这就是说，用"简单的量"代替了"复杂的量"，且产生的误差又可以忽略不计，这就是可微的含义.

(3)由于"$f(x)$ 在点 x_0 处可微"与"$f(x)$ 在点 x_0 处可导"互为充要条件，判别 $f(x)$ 在点 x_0 处是否可微可以转化为判别其在点 x_0 处是否可导，这样的问题考生会比较熟悉.

(4)可微的几何意义.

若 $f(x)$ 在点 x_0 处可微，则在点 (x_0,y_0) 附近可以用切线段近似代替曲线段，这是可微的几何意义.

二、导数与微分的计算

1. 四则运算

若以下函数均可导，则

和、差的导数(微分)　$[u(x)\pm v(x)]'=u'(x)\pm v'(x)$，　$\mathrm{d}[u(x)\pm v(x)]=\mathrm{d}u(x)\pm \mathrm{d}v(x)$.

积的导数(微分)　$[u(x)v(x)]'=u'(x)v(x)+u(x)v'(x)$，　$\mathrm{d}[u(x)v(x)]=u(x)\mathrm{d}v(x)+v(x)\mathrm{d}u(x)$.

【注】　$[u(x)v(x)w(x)]'=u'(x)v(x)w(x)+u(x)v'(x)w(x)+u(x)v(x)w'(x)$，如果遇到因式超过三个的式子，一般不要直接求导，而要另谋他法，见例 1.4.8.

商的导数(微分)　$\left[\dfrac{u(x)}{v(x)}\right]'=\dfrac{u'(x)v(x)-u(x)v'(x)}{[v(x)]^2}, v(x)\neq 0,$

$$\mathrm{d}\left[\frac{u(x)}{v(x)}\right]=\frac{v(x)\mathrm{d}u(x)-u(x)\mathrm{d}v(x)}{[v(x)]^2}, v(x)\neq 0.$$

2. 分段函数的导数

设 $f(x)=\begin{cases} f_1(x), & x\geqslant x_0, \\ f_2(x), & x<x_0, \end{cases}$ 则 $f'_+(x_0)=\lim\limits_{x\to x_0^+}\dfrac{f_1(x)-f(x_0)}{x-x_0}$，$f'_-(x_0)=\lim\limits_{x\to x_0^-}\dfrac{f_2(x)-f(x_0)}{x-x_0}$. 根据 $f'_+(x_0)\overset{?}{=\!=}f'_-(x_0)$ 来判定 $f'(x_0)$.

3. 复合函数的导数与微分形式不变性

设 $u=g(x)$ 在点 x(没有下标是泛指的点，下同)处可导，$y=f(u)$ 在点 $u=g(x)$ 处可导，则

$$\{f[g(x)]\}'=f'[g(x)]g'(x),$$

$$\mathrm{d}\{f[g(x)]\}=f'[g(x)]g'(x)\mathrm{d}x. \tag{1-4-6}$$

(1-4-6)式就是微分形式的不变性——无论 u 是中间变量还是自变量，$\mathrm{d}y=f'(u)\mathrm{d}u$ 都成立.

【注】 $\{f[g(x)]\}'=\dfrac{\mathrm{d}\{f[g(x)]\}}{\mathrm{d}x}$，而 $f'[g(x)]=\dfrac{\mathrm{d}\{f[g(x)]\}}{\mathrm{d}g(x)}$，要看清楚求导符号的位置，不要弄错了.

4. 反函数的导数

设 $y=f(x)$ 可导，且 $f'(x)\neq 0$，则存在反函数 $x=\varphi(y)$，且 $\dfrac{\mathrm{d}x}{\mathrm{d}y}=\dfrac{1}{\dfrac{\mathrm{d}y}{\mathrm{d}x}}$，即 $\varphi'(y)=\dfrac{1}{f'(x)}$.

5. 参数方程所确定的函数的导数

设函数 $y=y(x)$ 由参数方程 $\begin{cases}x=\varphi(t),\\ y=\psi(t)\end{cases}$ 确定，其中 t 是参数，且 $\varphi(t)$，$\psi(t)$ 均对 t 可导，$\varphi'(t)\neq 0$，则

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\frac{\psi'(t)}{\varphi'(t)}.$$

6. 隐函数求导法

设函数 $y=y(x)$ 是由方程 $F(x,y)=0$ 确定的可导函数，则

①方程 $F(x,y)=0$ 两边对自变量 x 求导，注意 $y=y(x)$，即将 y 看作中间变量，得到一个关于 y' 的方程；

②解该方程便可求出 y'.

7. 对数求导法

对于多项相乘、相除、开方、乘方的式子，一般先取对数再求导. 设 $y=f(x)(f(x)>0)$，则

①等式两边取对数，得 $\ln y=\ln f(x)$；

②两边对自变量 x 求导（同样注意 $y=f(x)$，即将 y 看作中间变量），得

$$\frac{1}{y}y'=[\ln f(x)]'\Rightarrow y'=y[\ln f(x)]'.$$

8. 幂指函数求导法

对于 $u(x)^{v(x)}(u(x)>0,u(x)\neq 1)$，除了用上面的对数求导法外，还可以先化成指数函数

$$u(x)^{v(x)}=\mathrm{e}^{v(x)\ln u(x)},$$

然后求导，得

$$[u(x)^{v(x)}]'=[\mathrm{e}^{v(x)\ln u(x)}]'=u(x)^{v(x)}\left[v'(x)\ln u(x)+v(x)\cdot\frac{u'(x)}{u(x)}\right].$$

9. 高阶导数

高阶导数主要是三种方法.

(1)用归纳法.

逐次求导，探索规律，得出通式.

比如，设 $y=2^x$，则

$$y'=2^x\ln 2,\quad y''=2^x(\ln 2)^2,\quad \cdots.$$

得出通式

$$y^{(n)}=2^x(\ln 2)^n,n=0,1,2,\cdots.$$

(2)用高阶求导公式.

设 $u=u(x),v=v(x)$，均 n 阶可导，则

$$[u\pm v]^{(n)}=u^{(n)}\pm v^{(n)},$$

$$\begin{aligned}(uv)^{(n)}&=u^{(n)}v+\mathrm{C}_n^1u^{(n-1)}v'+\mathrm{C}_n^2u^{(n-2)}v''+\cdots+\mathrm{C}_n^ku^{(n-k)}v^{(k)}+\cdots+\mathrm{C}_n^{n-1}u'v^{(n-1)}+uv^{(n)}\\&=\sum_{k=0}^{n}\mathrm{C}_n^ku^{(n-k)}v^{(k)}.\end{aligned}\tag{1-4-7}$$

(1-4-7)式就是求乘积的高阶导数的**莱布尼茨公式**，其中 $u^{(0)}=u,v^{(0)}=v$.

见到两个函数乘积的高阶导数，一般用莱布尼茨公式即可.

(3)用泰勒公式.

先写出 $y=f(x)$ 的泰勒公式或者麦克劳林公式，再通过比较系数来获得 $f^{(n)}(x_0)$. 具体说来，

① 任何一个无穷阶可导的函数(在收敛的条件下)都可写成 $y=f(x)=\sum\limits_{n=0}^{\infty}\dfrac{f^{(n)}(x_0)}{n!}(x-x_0)^n$，

或者

$$y=f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n.$$

②题目给出一个具体的无穷阶可导函数 $y=f(x)$，可以通过已知公式展开成幂级数. 这些已知公式为

(ⅰ) $\mathrm{e}^x=\sum\limits_{n=0}^{\infty}\dfrac{x^n}{n!}=1+x+\dfrac{x^2}{2!}+\cdots+\dfrac{x^n}{n!}+\cdots,-\infty<x<+\infty$.

(ⅱ) $\dfrac{1}{1+x}=\sum\limits_{n=0}^{\infty}(-1)^nx^n=1-x+x^2-x^3+\cdots+(-1)^nx^n+\cdots,-1<x<1$.

(ⅲ) $\dfrac{1}{1-x}=\sum\limits_{n=0}^{\infty}x^n=1+x+x^2+\cdots+x^n+\cdots,-1<x<1$.

(ⅳ) $\ln(1+x)=\sum\limits_{n=1}^{\infty}(-1)^{n-1}\dfrac{x^n}{n}=x-\dfrac{x^2}{2}+\dfrac{x^3}{3}-\dfrac{x^4}{4}+\cdots+(-1)^{n-1}\dfrac{x^n}{n}+\cdots,-1<x\leqslant 1$.

(ⅴ)
$$\begin{aligned}\sin x&=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}\\&=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\cdots,-\infty<x<+\infty.\end{aligned}$$

(ⅵ)
$$\begin{aligned}\cos x&=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}\\&=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots+(-1)^n\frac{x^{2n}}{(2n)!}+\cdots,-\infty<x<+\infty.\end{aligned}$$

(ⅶ)
$$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+\cdots,\begin{cases}x\in(-1,1), & 当\ \alpha\leqslant-1,\\x\in(-1,1], & 当\ -1<\alpha<0,\\x\in[-1,1], & 当\ \alpha>0.\end{cases}$$

③根据函数展开式的唯一性，比较①，②中公式的系数，就可以获得 $f^{(n)}(x_0)$ 或者 $f^{(n)}(0)$.

需要指出，数学二的考生虽然不考无穷级数，但是却必须掌握这个内容，因为数学二对于高阶导数的考查频率较高.

【注】 除上述情形，下面两个内容也是重要的.

(1)参数方程确定的函数的二阶导数.

设函数 $y=y(x)$ 由参数方程 $\begin{cases}x=\varphi(t),\\y=\psi(t)\end{cases}$ 确定，且 $\varphi(t),\psi(t)$ 均二阶可导，$\varphi'(t)\neq0$，其中 t 是参数，则

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\frac{\psi'(t)}{\varphi'(t)},\quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)}{\mathrm{d}x}=\frac{\mathrm{d}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=\frac{\psi''(t)\varphi'(t)-\psi'(t)\varphi''(t)}{[\varphi'(t)]^3}.$$

(2)反函数的二阶导数.

在 $y=f(x)$ 二阶可导的情况下，记 $f'(x)=y'_x$，$\varphi'(y)=x'_y\,(x'_y\neq0)$，则有

$$y'_x=\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}y}}=\frac{1}{x'_y},\quad y''_{xx}=\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)}{\mathrm{d}x}=\frac{\mathrm{d}\left(\frac{1}{x'_y}\right)}{\mathrm{d}x}=\frac{\mathrm{d}\left(\frac{1}{x'_y}\right)}{\mathrm{d}y}\cdot\frac{1}{x'_y}=\frac{-x''_{yy}}{(x'_y)^3}.$$

反过来，则有

$$x'_y=\frac{1}{y'_x},\quad x''_{yy}=\frac{-y''_{xx}}{(y'_x)^3}.$$

10. 变限积分求导公式

设 $F(x)=\int_{\varphi_1(x)}^{\varphi_2(x)}f(t)\mathrm{d}t$，其中 $f(x)$ 在 $[a,b]$ 上连续，可导函数 $\varphi_1(x)$ 和 $\varphi_2(x)$ 的值域在 $[a,b]$ 上，则在函数 $\varphi_1(x)$ 和 $\varphi_2(x)$ 的公共定义域上，有

$$F'(x)=\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{\varphi_1(x)}^{\varphi_2(x)}f(t)\mathrm{d}t\right]=f[\varphi_2(x)]\varphi_2'(x)-f[\varphi_1(x)]\varphi_1'(x).$$

11. 基本求导公式

$(x^\alpha)'=\alpha x^{\alpha-1}$（$\alpha$ 为常数），　$(a^x)'=a^x\ln a\,(a>0,a\neq1)$，　$(\mathrm{e}^x)'=\mathrm{e}^x$，　$(\log_a x)'=\frac{1}{x\ln a}\,(a>0,a\neq1)$，

$(\ln x)'=\frac{1}{x}$，　$(\sin x)'=\cos x$，　$(\cos x)'=-\sin x$，　$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$，

$(\arccos x)'=-\frac{1}{\sqrt{1-x^2}}$，　$(\tan x)'=\sec^2x$，　$(\cot x)'=-\csc^2x$，　$(\arctan x)'=\frac{1}{1+x^2}$，

$(\operatorname{arccot} x)'=-\frac{1}{1+x^2}$，　$(\sec x)'=\sec x\tan x$，　$(\csc x)'=-\csc x\cot x$，

$[\ln(x+\sqrt{x^2+1})]'=\frac{1}{\sqrt{x^2+1}}$，　$[\ln(x+\sqrt{x^2-1})]'=\frac{1}{\sqrt{x^2-1}}$.

基础例题精解

一、概念

例 1.4.1　证明：(1)若 $f(x)$ 是可导的偶函数，则 $f'(x)$ 是奇函数；

(2)若 $f(x)$ 是可导的奇函数，则 $f'(x)$ 是偶函数.

证明　(1)设 x 是定义域中任指定的一点(也叫泛指的点)，由导数定义，得

$$f'(-x)=\lim_{\Delta x\to0}\frac{f(-x+\Delta x)-f(-x)}{\Delta x}=\lim_{\Delta x\to0}\frac{f(x-\Delta x)-f(x)}{\Delta x}$$

$$=(-1)\lim_{-\Delta x\to0}\frac{f(x-\Delta x)-f(x)}{-\Delta x}=-f'(x).$$

故 $f'(x)$是奇函数.

(2)同理可证. 若 $f(x)$是可导的奇函数，则 $f'(x)$是偶函数.

例 1.4.2 证明：若 $f(x)$是可导的周期为 T 的周期函数，则 $f'(x)$也是以 T 为周期的周期函数.

证明 由导数定义，得

$$\begin{aligned} f'(x+T) &= \lim_{\Delta x\to 0}\frac{f(x+T+\Delta x)-f(x+T)}{\Delta x} \\ &\xlongequal{f(x+T)=f(x)}\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x} \\ &= f'(x). \end{aligned}$$

故 $f'(x)$也是以 T 为周期的周期函数.

例 1.4.3 设 $f(x)$是二阶可导的以 2 为周期的奇函数，且 $f\left(\frac{1}{2}\right)>0$，$f'\left(\frac{1}{2}\right)>0$，比较$f\left(-\frac{1}{2}\right)$，$f'\left(\frac{3}{2}\right)$与 $f''(0)$的大小，并说明理由.

解 由 $f(x)$为奇函数，则 $f\left(-\frac{1}{2}\right)=-f\left(\frac{1}{2}\right)<0$. 根据例 1.4.1(2)，知 $f'(x)$为偶函数，再用一次例 1.4.1(1)的结论，则 $f''(x)$为奇函数(事实上，若 $f(x)$无穷阶可导，则求导一次，奇偶性即互换一次)，由 $f''(x)$存在，故 $f''(0)=0$.

又 $f'\left(\frac{1}{2}\right)>0$，由例 1.4.2 的结论，知 $f'(x)$也是以 2 为周期的周期函数，则$f'\left(\frac{3}{2}\right)=f'\left(\frac{3}{2}-2\right)=f'\left(-\frac{1}{2}\right)=f'\left(\frac{1}{2}\right)>0$，故 $f\left(-\frac{1}{2}\right)<f''(0)<f'\left(\frac{3}{2}\right)$.

例 1.4.4 设 $\delta>0$，$f(x)$在$[-\delta,\delta]$上有定义，$f(0)=1$，且满足

$$\lim_{x\to 0}\frac{\ln(1-2x)+2xf(x)}{x^2}=0,$$

证明：$f(x)$在 $x=0$ 处可导，并求 $f'(0)$.

证明 使用泰勒公式，有

$$\lim_{x\to 0}\frac{\ln(1-2x)+2xf(x)}{x^2}=\lim_{x\to 0}\frac{-2x-\frac{1}{2}\cdot 4x^2+o(x^2)+2xf(x)}{x^2}=2\lim_{x\to 0}\frac{f(x)-1}{x}-2+0=0,$$

于是极限 $\lim\limits_{x\to 0}\frac{f(x)-f(0)}{x-0}=1$，即为 $f'(0)$，于是函数 $f(x)$在 $x=0$ 处可导，且 $f'(0)=1$.

例 1.4.5 设函数 $f(x)$在 $x=0$ 处连续，则下列命题错误的是( ).

(A)若$\lim\limits_{x\to 0}\frac{f(x)}{x}$存在，则 $f(0)=0$

(B)若$\lim\limits_{x\to 0}\frac{f(x)+f(-x)}{x}$存在，则 $f(0)=0$

(C)若$\lim\limits_{x\to 0}\frac{f(x)}{x}$存在，则 $f'(0)$存在

(D)若$\lim\limits_{x\to 0}\frac{f(x)-f(-x)}{x}$存在，则 $f'(0)$存在

解 应选(D).

对于选项(A)，(C)，事实上有如下命题：设 $f(x)$在 $x=0$ 处连续，且$\lim\limits_{x\to 0}\frac{f(x)}{x}=A$ 存在，则

$$f(0)=\lim_{x\to 0}f(x)=0,\quad f'(0)=\lim_{x\to 0}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0}\frac{f(x)}{x}=A,$$

故(A),(C)均正确.

对于选项(B),由$\lim\limits_{x\to 0}\frac{f(x)+f(-x)}{x}$存在,则$\lim\limits_{x\to 0}[f(x)+f(-x)]=0$,即$2f(0)=0$,故$f(0)=0$,(B)正确.

对于选项(D),可取反例$f(x)=|x|$,则$f(x)$在$x=0$处连续,且

$$\lim_{x\to 0}\frac{f(x)-f(-x)}{x}=\lim_{x\to 0}\frac{|x|-|-x|}{x}=0$$

存在,但是$f(x)$在$x=0$处不可导,故选择(D).

例 1.4.6 设函数$y=f(x)$,$f'(x_0)=\frac{1}{2}$,则当$\Delta x\to 0$时,该函数在$x=x_0$处的微分$\mathrm{d}y$(　　).

(A)与Δx等价无穷小　　(B)与Δx同阶非等价无穷小

(C)比Δx低阶的无穷小　　(D)比Δx高阶的无穷小

解 应选(B).

按照微分的定义,在$x=x_0$处,$\mathrm{d}y=f'(x_0)\Delta x=\frac{1}{2}\Delta x$,当$\Delta x\to 0$时,$\mathrm{d}y$与$\Delta x$为同阶非等价无穷小,故选择(B).

例 1.4.7 设函数$f(u)$可导,且$y=f(x^2)$,当自变量x在$x=-1$处取得增量$\Delta x=-0.1$时,相应的函数增量Δy的线性主部为0.1,则$f'(1)=$(　　).

(A)-1　　(B)0.1　　(C)1　　(D)0.5

解 应选(D).

本题依然是考查微分的定义.函数的微分是函数增量的线性主部,且$\mathrm{d}y=y'\mathrm{d}x=y'\Delta x$,而

$$\mathrm{d}y=f'(x^2)\mathrm{d}x^2=2xf'(x^2)\mathrm{d}x=2xf'(x^2)\Delta x,$$

因此,由$0.1=-2f'(1)\cdot(-0.1)$,可得$f'(1)=0.5$,故选(D).

二、计算

例 1.4.8 设$f(x)=\prod\limits_{n=1}^{100}\left(\tan\frac{\pi x^n}{4}-n\right)$,则$f'(1)=$________.

解 应填$-\frac{\pi\cdot 99!}{2}$.

本题考查一元函数微分学的基本计算,是一道比较新颖的试题.

本题的研究对象$f(x)$是多因式相乘,如果直接对其使用导数定义或者先求导再代值,都比较麻烦.本题希望考生发现,当把$x=1$代入每个因式后,只有第一项$\tan\frac{\pi}{4}-1=0$,而其余所有项都不等于0,抓住第一项这个"特立独行"的主要条件,记$g(x)=\prod\limits_{n=2}^{100}\left(\tan\frac{\pi x^n}{4}-n\right)$,于是

$$f(x)=\left(\tan\frac{\pi x}{4}-1\right)\cdot g(x),$$

故

$$f'(1)=\frac{\pi}{4}\sec^2\frac{\pi x}{4}\bigg|_{x=1}\cdot g(1)=-\frac{\pi\cdot 99!}{2}.$$

例 1.4.9 设 $y=\ln|x|$，求 y'.

解
$$y=\ln|x|=\begin{cases}\ln x, & x>0,\\ \ln(-x), & x<0.\end{cases}$$

当 $x>0$ 时，
$$y'=(\ln x)'=\frac{1}{x};$$

当 $x<0$ 时，
$$y'=[\ln(-x)]'=\frac{1}{-x}\cdot(-x)'=\frac{1}{-x}\cdot(-1)=\frac{1}{x},$$

因此
$$y'=(\ln|x|)'=\frac{1}{x}(x\neq 0).$$

例 1.4.10 已知函数 $f(x)=\begin{cases}x^2, & x\geqslant 0,\\ ax+b, & x<0\end{cases}$ 在 $x=0$ 处可导，求 a,b 的值.

解 因为 $f(x)$在 $x=0$ 处可导，所以 $f(x)$在 $x=0$ 处连续.

因为
$$\lim_{x\to 0^-}f(x)=\lim_{x\to 0^-}(ax+b)=b,\quad f(0)=0,$$
所以由$\lim\limits_{x\to 0^-}f(x)=f(0)$，得 $b=0$.

又因为
$$f'_+(0)=\lim_{x\to 0^+}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0^+}\frac{x^2-0}{x}=0,$$
$$f'_-(0)=\lim_{x\to 0^-}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0^-}\frac{ax+b-0}{x}=\lim_{x\to 0^-}\frac{ax}{x}=a,$$
所以由 $f'_-(0)=f'_+(0)$，得 $a=0$.

例 1.4.11 求函数 $f(x)=2^{|x-a|}$ 的导数.

解 因为
$$f(x)=\begin{cases}2^{-(x-a)}, & x<a,\\ 1, & x=a,\\ 2^{x-a}, & x>a,\end{cases}$$

所以，当 $x<a$ 时，
$$f'(x)=[2^{-(x-a)}]'=-2^{a-x}\ln 2,$$

当 $x>a$ 时，
$$f'(x)=(2^{x-a})'=2^{x-a}\ln 2,$$

当 $x=a$ 时，由于
$$f'_-(a)=\lim_{x\to a^-}\frac{f(x)-f(a)}{x-a}=\lim_{x\to a^-}\frac{2^{-(x-a)}-1}{x-a}$$
$$=\lim_{x\to a^-}\frac{-(x-a)\ln 2}{x-a}=-\ln 2,$$
$$f'_+(a)=\lim_{x\to a^+}\frac{f(x)-f(a)}{x-a}=\lim_{x\to a^+}\frac{2^{x-a}-1}{x-a}$$
$$=\lim_{x\to a^+}\frac{(x-a)\ln 2}{x-a}=\ln 2,$$

所以 $f'_-(a) \neq f'_+(a)$，故 $f'(a)$ 不存在.

因此，

$$f'(x)=\begin{cases}-2^{a-x}\ln 2, & x<a,\\ 2^{x-a}\ln 2, & x>a.\end{cases}$$

例 1.4.12 设 $y=\ln(x+\sqrt{x^2+a^2})(a\neq0)$，求 $y'\Big|_{x=0}$.

解 因为

$$\begin{aligned}y'&=\frac{1}{x+\sqrt{x^2+a^2}}\cdot(x+\sqrt{x^2+a^2})'\\&=\frac{1}{x+\sqrt{x^2+a^2}}\left[1+\frac{1}{2\sqrt{x^2+a^2}}\cdot(x^2+a^2)'\right]\\&=\frac{1}{x+\sqrt{x^2+a^2}}\left(1+\frac{x}{\sqrt{x^2+a^2}}\right)=\frac{1}{\sqrt{x^2+a^2}}.\end{aligned}$$

所以

$$y'\Big|_{x=0}=\frac{1}{\sqrt{x^2+a^2}}\Bigg|_{x=0}=\frac{1}{|a|}(a\neq0).$$

例 1.4.13 设函数 $f(x)=\begin{cases}\ln\sqrt{x}, & x\geqslant1,\\ 2x-1, & x<1,\end{cases}$ 且 $y=f[f(x)]$，则 $\dfrac{\mathrm{d}y}{\mathrm{d}x}\Big|_{x=\mathrm{e}}=$________.

解 应填$\dfrac{1}{\mathrm{e}}$.

$$\frac{\mathrm{d}y}{\mathrm{d}x}\Big|_{x=\mathrm{e}}=f'[f(x)]f'(x)\Big|_{x=\mathrm{e}}=f'[f(\mathrm{e})]f'(\mathrm{e}),$$

其中 $f(\mathrm{e})=\ln\sqrt{x}\Big|_{x=\mathrm{e}}=\dfrac{1}{2}$， $f'[f(\mathrm{e})]=f'\left(\dfrac{1}{2}\right)=(2x-1)'\Big|_{x=\frac{1}{2}}=2$， $f'(\mathrm{e})=(\ln\sqrt{x})'\Big|_{x=\mathrm{e}}=\dfrac{1}{2\mathrm{e}}$，

所以
$$\frac{\mathrm{d}y}{\mathrm{d}x}\Big|_{x=\mathrm{e}}=2\cdot\frac{1}{2\mathrm{e}}=\frac{1}{\mathrm{e}}.$$

例 1.4.14 设 $y=\mathrm{e}^{\sin(\ln x)}$，求 $\mathrm{d}y$ 及$\dfrac{\mathrm{d}y}{\mathrm{d}x}$.

解 由一阶微分形式的不变性，得

$$\begin{aligned}\mathrm{d}[\mathrm{e}^{\sin(\ln x)}]&=\mathrm{e}^{\sin(\ln x)}\mathrm{d}[\sin(\ln x)]\\&=\mathrm{e}^{\sin(\ln x)}\cos(\ln x)\mathrm{d}(\ln x)\\&=\mathrm{e}^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}\mathrm{d}x,\end{aligned}$$

所以

$$\mathrm{d}y=\mathrm{e}^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}\mathrm{d}x,\quad \frac{\mathrm{d}y}{\mathrm{d}x}=\mathrm{e}^{\sin(\ln x)}\cos(\ln x)\frac{1}{x}.$$

例 1.4.15 求下列函数的导数.

(1) $y=\arcsin x, -1<x<1$；

(2) $y=\arctan x$.

解 (1)由 $y=\arcsin x$，得反函数 $x=\sin y, y\in\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)$. 根据反函数求导公式，得

$$(\arcsin x)'=\frac{1}{(\sin y)'}=\frac{1}{\cos y}=\frac{1}{\sqrt{1-\sin^2 y}}=\frac{1}{\sqrt{1-x^2}}.$$

(2)由 $y=\arctan x$，得反函数 $x=\tan y,y\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$. 根据反函数求导公式，得

$$(\arctan x)'=\frac{1}{(\tan y)'}=\frac{1}{\sec^2 y}=\frac{1}{1+\tan^2 y}=\frac{1}{1+x^2}.$$

【注】 类似地，可以求得 $(\arccos x)'=-\frac{1}{\sqrt{1-x^2}}$，$(\text{arccot}\, x)'=-\frac{1}{1+x^2}$.

例 1.4.16 设 $y=f(x)$ 的反函数是 $x=\varphi(y)$，且 $f(x)=\int_1^{2x}e^{t^2}\mathrm{d}t+1$，则 $\varphi''(1)=$________.

解 应填 $-\frac{1}{e^2}$.

由题设可得
$$\varphi'(y)=\frac{1}{f'(x)},$$
$$\varphi''(y)=\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{f'(x)}\right]\cdot\frac{\mathrm{d}x}{\mathrm{d}y}=-\frac{f''(x)}{[f'(x)]^2}\cdot\frac{1}{f'(x)}=-\frac{f''(x)}{[f'(x)]^3}.$$

由 $f(x)=\int_1^{2x}e^{t^2}\mathrm{d}t+1$ 知，$x=\frac{1}{2}$时 $y=1$，且
$$f'(x)=2e^{4x^2},\quad f''(x)=16xe^{4x^2},$$

则
$$\varphi''(1)=-\frac{f''\left(\frac{1}{2}\right)}{\left[f'\left(\frac{1}{2}\right)\right]^3}=-\frac{8e}{8e^3}=-\frac{1}{e^2}.$$

例 1.4.17 设 $y=y(x)$ 由方程 $\begin{cases}x=\sin t,\\ y=t\sin t+\cos t\end{cases}$ （t 为参数）确定，则 $\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{t=\frac{\pi}{4}}=$________.

解 应填 $\sqrt{2}$.

因为
$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y/\mathrm{d}t}{\mathrm{d}x/\mathrm{d}t}=t,$$
$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{\cos t},$$

从而有
$$\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{t=\frac{\pi}{4}}=\frac{1}{\cos\frac{\pi}{4}}=\sqrt{2}.$$

例 1.4.18 设 $y=y(x)$ 是由方程 $\sin(xy)=\ln\frac{x+e}{y}+1$ 确定的隐函数，求 $y'(0)$ 的值.

解 在方程中令 $x=0$ 可得 $0=\ln\frac{e}{y(0)}+1\Rightarrow y(0)=e^2$. 将方程两边对 x 求导，得
$$\cos(xy)(y+xy')=\frac{1}{x+e}-\frac{y'}{y}.$$

将 $x=0,y(0)=e^2$ 代入，有 $e^2=\frac{1}{e}-\frac{y'(0)}{e^2}$，即 $y'(0)=e-e^4$.

例 1.4.19 设 $y=\sqrt[3]{\frac{(x+1)(2x-1)^2}{(4-3x)^5}}$，求 y'.

解 两边取对数，并化简，得

$$\ln|y|=\frac{1}{3}(\ln|x+1|+2\ln|2x-1|-5\ln|4-3x|),$$

两边对 x 求导，得

$$\frac{1}{y}\cdot y'=\frac{1}{3}\left(\frac{1}{x+1}+2\cdot\frac{2}{2x-1}-5\cdot\frac{-3}{4-3x}\right),$$

于是得到

$$y'=\frac{1}{3}\sqrt[3]{\frac{(x+1)(2x-1)^2}{(4-3x)^5}}\left(\frac{1}{x+1}+\frac{4}{2x-1}+\frac{15}{4-3x}\right).$$

例 1.4.20 求函数 $y=x^x(x>0)$ 的导数.

解 两边取对数得

$$\ln y=x\ln x.$$

两边关于变量 x 求导，将 y 看作是中间变量，得

$$\frac{1}{y}y'=\ln x+1.$$

所以 $y'=x^x(\ln x+1)(x>0)$.

例 1.4.21 求函数 $y=x^{\frac{1}{x}}(x>0)$ 的导数.

解 两边取对数得

$$\ln y=\frac{\ln x}{x}.$$

两边关于变量 x 求导，将 y 看作是中间变量，得

$$\frac{1}{y}y'=\frac{1-\ln x}{x^2}.$$

所以 $y'=x^{\frac{1}{x}}\frac{1-\ln x}{x^2}(x>0)$.

例 1.4.22 求 $y=\frac{1}{x}$ 的 n 阶导数.

解

$$y'=\left(\frac{1}{x}\right)'=(x^{-1})'=(-1)x^{-1-1}=(-1)x^{-2},$$

$$y''=[(-1)x^{-2}]'=(-1)\cdot(-2)x^{-3},$$

$$y'''=[(-1)\cdot(-2)x^{-3}]'=(-1)\cdot(-2)\cdot(-3)x^{-4},$$

于是

$$y^{(n)}=(-1)\cdot(-2)\cdot\cdots\cdot(-n)x^{-n-1},$$

即

$$\left(\frac{1}{x}\right)^{(n)}=(-1)^n n!\,x^{-n-1}=(-1)^n\frac{n!}{x^{n+1}},n=1,2,\cdots.$$

【注】 由此可得 $\ln x$ 的 n 阶导数公式. 因为 $(\ln x)'=\frac{1}{x}$，所以

$$(\ln x)^{(n)}=[(\ln x)']^{(n-1)}=\left(\frac{1}{x}\right)^{(n-1)}=(-1)^{n-1}\frac{(n-1)!}{x^n},n=2,3,\cdots.$$

例 1.4.23 求 $y=\sin x$ 的 n 阶导数.

解

$$y'=(\sin x)'=\cos x,$$

$$y''=(\cos x)'=-\sin x,$$

$$y'''=(-\sin x)'=-\cos x,$$

……

但这样算下去，很难找到规律. 想到 $\cos x=\sin\left(x+\dfrac{\pi}{2}\right)$，则有

$$y'=(\sin x)'=\cos x=\sin\left(x+\frac{\pi}{2}\right),$$

$$y''=\left[\sin\left(x+\frac{\pi}{2}\right)\right]'=\cos\left(x+\frac{\pi}{2}\right)=\sin\left(x+\frac{\pi}{2}+\frac{\pi}{2}\right)=\sin\left(x+2\cdot\frac{\pi}{2}\right),$$

$$y'''=\left[\sin\left(x+2\cdot\frac{\pi}{2}\right)\right]'=\cos\left(x+2\cdot\frac{\pi}{2}\right)=\sin\left(x+3\cdot\frac{\pi}{2}\right),$$

于是

$$y^{(n)}=\sin\left(x+n\cdot\frac{\pi}{2}\right),$$

即

$$(\sin x)^{(n)}=\sin\left(x+n\cdot\frac{\pi}{2}\right),n=1,2,\cdots.$$

【注】 按此方法，我们可以得出几个重要的公式：

$(a^x)^{(n)}=a^x(\ln a)^n(a>0,a\neq1)$；　$(e^x)^{(n)}=e^x$；　$(\sin kx)^{(n)}=k^n\sin\left(kx+n\cdot\dfrac{\pi}{2}\right)$；

$(\cos kx)^{(n)}=k^n\cos\left(kx+n\cdot\dfrac{\pi}{2}\right)$；　$(\ln x)^{(n)}=(-1)^{n-1}\dfrac{(n-1)!}{x^n}(x>0)$；

$[\ln(1+x)]^{(n)}=(-1)^{n-1}\dfrac{(n-1)!}{(1+x)^n}(x>-1)$；

$[(x+x_0)^m]^{(n)}=m(m-1)(m-2)\cdots(m-n+1)(x+x_0)^{m-n}$；　$\left(\dfrac{1}{x+a}\right)^{(n)}=\dfrac{(-1)^n n!}{(x+a)^{n+1}}$.

读者若能记住这些式子，那是最好；若记不住，学会推导的方式，在考试中快速计算出来，也是可以的.

例 1.4.24　已知函数 $f(x)$ 具有任意阶导数，且 $f'(x)=[f(x)]^2$，求 $f^{(n)}(x)$，其中 n 为正整数.

解　$f'(x)=[f(x)]^2$ 两边同时对 x 求导，得

$$f''(x)=2f(x)f'(x)=2[f(x)]^3,$$

$$f'''(x)=2\cdot3[f(x)]^2\cdot f'(x)=2\cdot3[f(x)]^4,$$

$$f^{(4)}(x)=2\cdot3\cdot4[f(x)]^3\cdot f'(x)=2\cdot3\cdot4[f(x)]^5,$$

于是得到

$$f^{(n)}(x)=n![f(x)]^{n+1}.$$

【注】 可用数学归纳法严格证明，但考试中不必给出.

设 $n=k$ 时有 $f^{(k)}(x)=k![f(x)]^{k+1}$，则 $n=k+1$ 时，将上式两边再对 x 求导，有

$$f^{(k+1)}(x)=(k+1)\cdot k![f(x)]^k f'(x)=(k+1)![f(x)]^{k+2},$$

故对正整数 n，有 $f^{(n)}(x)=n![f(x)]^{n+1}$.

例 1.4.25　设函数 $y=e^x\cos x$，求 $y^{(4)}$.

解　因为 $y=e^x\cos x$，所以

$$y^{(4)}=(\mathrm{e}^x)^{(4)}\cos x+4(\mathrm{e}^x)'''(\cos x)'+\frac{4\cdot 3}{2!}(\mathrm{e}^x)''(\cos x)''+$$

$$\frac{4\cdot 3\cdot 2}{3!}(\mathrm{e}^x)'(\cos x)'''+\mathrm{e}^x(\cos x)^{(4)}$$

$$=\mathrm{e}^x(\cos x-4\sin x-6\cos x+4\sin x+\cos x)$$

$$=-4\mathrm{e}^x\cos x.$$

例 1.4.26 设 $y=x^3\sin x$，求 $y^{(6)}(0)$.

解 严格按照“基础内容精讲”中所述的三个步骤来解题.

①由于 $y=x^3\sin x$ 无穷阶可导，则可以先将其抽象展开为 $y=\sum\limits_{n=0}^{\infty}\frac{y^{(n)}(0)}{n!}x^n$；

②又由于 $y=x^3\sin x=x^3\left(x-\frac{1}{6}x^3+\cdots\right)=x^4-\frac{1}{6}x^6+\cdots$；

③根据函数展开式的唯一性，比较①，②中公式的系数，则 $\frac{y^{(6)}(0)}{6!}=-\frac{1}{6}$，于是

$$y^{(6)}(0)=-\frac{6!}{6}=-120.$$

基础习题精练

习 题

1.4.1 设函数 $f(x)$ 在 $x=0$ 处连续，且 $\lim\limits_{x\to 0}\frac{f(x^2)}{x^2}=1$，则（ ）.

(A) $f(0)=0$ 且 $f'_-(0)$ 存在　　(B) $f(0)=1$ 且 $f'_-(0)$ 存在

(C) $f(0)=0$ 且 $f'_+(0)$ 存在　　(D) $f(0)=1$ 且 $f'_+(0)$ 存在

1.4.2 设函数 $f(x)=|x^3-1|\varphi(x)$，其中 $\varphi(x)$ 在 $x=1$ 处连续，则 $\varphi(1)=0$ 是 $f(x)$ 在 $x=1$ 处可导的（ ）.

(A)充分必要条件　　(B)充分但非必要条件

(C)必要但非充分条件　　(D)既非充分又非必要条件

1.4.3 设 $y=f(\ln^2 x)\mathrm{e}^{f^2(x)}$，其中 f 可微，计算 $\frac{\mathrm{d}y}{\mathrm{d}x}$.

1.4.4 设函数 $f(x)$ 在 $x=2$ 的某邻域内具有任意阶导数，且 $f'(x)=\mathrm{e}^{f(x)}$，$f(2)=1$，计算 $f^{(n)}(2)$，其中 n 为正整数.

1.4.5 设函数 $y=f(x)$ 由方程 $y=x\ln y$ 所确定，求 $\frac{\mathrm{d}y}{\mathrm{d}x}$.

1.4.6 已知 $f'(x)=A\mathrm{e}^x$（A 为正常数），求 $f(x)$ 的反函数的二阶导数.

1.4.7 设可导函数 $y=f(x)$ 有反函数 $g(x)$，且 $f(a)=3$，$f'(a)=1$，$f''(a)=2$，求 $g''(3)$.

1.4.8 设 $f(x)=\begin{cases}x^2\sin\frac{\pi}{x}, & x<0,\\ A, & x=0,\\ ax^2+b, & x>0,\end{cases}$ 求常数 A,a,b 的值，使 $f(x)$ 在 $x=0$ 处可导，并求 $f'(0)$.

1.4.9　设函数 $y=f(x)$ 由 $\begin{cases} x=\ln(1+t^2)+1, \\ y=2\arctan t-(t+1)^2 \end{cases}$ 确定，求 $\dfrac{\mathrm{d}y}{\mathrm{d}x}, \dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}$.

1.4.10　设 $f(x)$ 满足 $f(0)=0$，且 $f'(0)$ 存在，求 $\lim\limits_{x\to0}\dfrac{f(1-\sqrt{\cos x})}{\ln(1-x\sin x)}$.

解答

1.4.1　(C)　**解**　因为 $f(x)$ 在 $x=0$ 处连续，且 $\lim\limits_{x\to0}\dfrac{f(x^2)}{x^2}=1$，所以 $\lim\limits_{x\to0}f(x^2)=0$，即 $f(0)=0$. 从而有

$$\lim_{x\to0}\frac{f(x^2)}{x^2}=\lim_{x\to0}\frac{f(x^2)-f(0)}{x^2-0}\overset{t=x^2}{=\!=\!=}\lim_{t\to0^+}\frac{f(t)-f(0)}{t-0}=f'_+(0),$$

应选(C).

1.4.2　(A)　**解**　由 $\varphi(1)=0$ 可知

$$f'_+(1)=\lim_{x\to1^+}\frac{f(x)-f(1)}{x-1}=\lim_{x\to1^+}\frac{|x^3-1|\varphi(x)}{x-1}=\lim_{x\to1^+}(x^2+x+1)\varphi(x)=0,$$

$$f'_-(1)=\lim_{x\to1^-}\frac{f(x)-f(1)}{x-1}=\lim_{x\to1^-}\frac{|x^3-1|\varphi(x)}{x-1}=-\lim_{x\to1^-}(x^2+x+1)\varphi(x)=0,$$

即 $f'_+(1)=f'_-(1)=0$，所以 $f'(1)=0$.

设 $f(x)$ 在 $x=1$ 处可导，因为 $f(1)=0$，所以

$$f'_+(1)=\lim_{x\to1^+}\frac{f(x)-f(1)}{x-1}=\lim_{x\to1^+}\frac{|x^3-1|\varphi(x)}{x-1}=\lim_{x\to1^+}(x^2+x+1)\varphi(x)=3\varphi(1),$$

$$f'_-(1)=\lim_{x\to1^-}\frac{f(x)-f(1)}{x-1}=\lim_{x\to1^-}\frac{|x^3-1|\varphi(x)}{x-1}=-\lim_{x\to1^-}(x^2+x+1)\varphi(x)=-3\varphi(1),$$

由 $f'_+(1)=f'_-(1)$ 可得，$3\varphi(1)=-3\varphi(1)$，故 $\varphi(1)=0$，应选(A).

1.4.3　**解**

$$\begin{aligned}\frac{\mathrm{d}y}{\mathrm{d}x}&=[f(\ln^2x)]'\cdot\mathrm{e}^{f^2(x)}+f(\ln^2x)\cdot[\mathrm{e}^{f^2(x)}]'\\&=f'(\ln^2x)\cdot2\ln x\cdot\frac{1}{x}\cdot\mathrm{e}^{f^2(x)}+f(\ln^2x)\cdot\mathrm{e}^{f^2(x)}\cdot2f(x)\cdot f'(x)\\&=2\mathrm{e}^{f^2(x)}\left[\frac{\ln x}{x}\cdot f'(\ln^2x)+f(\ln^2x)\cdot f(x)\cdot f'(x)\right].\end{aligned}$$

1.4.4　**解**　对 $f'(x)=\mathrm{e}^{f(x)}$ 两边关于 x 求导，得

$$f''(x)=\mathrm{e}^{f(x)}\cdot f'(x)=\mathrm{e}^{2f(x)},$$

两边再对 x 求导，得　$f'''(x)=\mathrm{e}^{2f(x)}\cdot2f'(x)=2!\,\mathrm{e}^{3f(x)}$，

两边再对 x 求导，得　$f^{(4)}(x)=2\mathrm{e}^{3f(x)}\cdot3f'(x)=3!\,\mathrm{e}^{4f(x)}$.

由以上导数规律可得　$f^{(n)}(x)=(n-1)!\,\mathrm{e}^{nf(x)}$，

所以　$f^{(n)}(2)=(n-1)!\,\mathrm{e}^{n}$.

1.4.5　**解**　方程两边同时对 x 求导，得 $y'=\ln y+x\cdot\dfrac{1}{y}\cdot y'$，解得 $y'=\dfrac{y\ln y}{y-x}$.

或令 $F(x,y)=y-x\ln y$，则 $F'_x=-\ln y, F'_y=1-\dfrac{x}{y}$，故 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{F'_x}{F'_y}=\dfrac{y\ln y}{y-x}$.

1.4.6 解 设 $y=f(x)$，则

$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{1}{\frac{\mathrm{d}y}{\mathrm{d}x}}=\frac{1}{f'(x)},$$

$$\frac{\mathrm{d}^2x}{\mathrm{d}y^2}=\frac{\mathrm{d}\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)}{\mathrm{d}y}=\frac{\mathrm{d}\left[\frac{1}{f'(x)}\right]}{\mathrm{d}x}\cdot\frac{\mathrm{d}x}{\mathrm{d}y}=-\frac{f''(x)}{[f'(x)]^2}\cdot\frac{1}{f'(x)}$$

$$=-\frac{f''(x)}{[f'(x)]^3}=-\frac{A\mathrm{e}^x}{(A\mathrm{e}^x)^3}=-\frac{1}{A^2\mathrm{e}^{2x}}.$$

1.4.7 解 为利用反函数求导公式，需改 $g(x)$ 为 $g(y)$，则

$$f'(x)\cdot g'(y)=1,$$

两边同时对 x 求导，有

$$f''(x)g'(y)+f'(x)g''(y)\cdot y'=0,$$

即

$$f''(x)g'(y)+[f'(x)]^2\cdot g''(y)=0, \tag{$*$}$$

由 $g'(3)=\frac{1}{f'(a)}=1$，将 $x=a,y=3$，及已知条件代入（$*$）式得 $g''(3)=-2$.

1.4.8 解 由可导与连续的关系有

$$\lim_{x\to0^-}x^2\sin\frac{\pi}{x}=\lim_{x\to0^+}(ax^2+b)=A,$$

所以 $A=b=0$.

又

$$f'_-(0)=\lim_{x\to0^-}\frac{x^2\sin\frac{\pi}{x}-0}{x-0}=0,\quad f'_+(0)=\lim_{x\to0^+}\frac{ax^2-0}{x-0}=0,$$

所以，a 可以为任意常数，且 $f'(0)=0$.

1.4.9 解 $\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}}=\frac{\frac{2}{1+t^2}-2(t+1)}{\frac{2t}{1+t^2}}=-(t^2+t+1)$,

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\frac{\mathrm{d}x}{\mathrm{d}t}\cdot\frac{\mathrm{d}^2y}{\mathrm{d}t^2}-\frac{\mathrm{d}^2x}{\mathrm{d}t^2}\cdot\frac{\mathrm{d}y}{\mathrm{d}t}}{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^3}=\frac{\frac{2t}{1+t^2}\cdot\left[-\frac{4t}{(1+t^2)^2}-2\right]-\frac{2-2t^2}{(1+t^2)^2}\cdot\left[\frac{2}{1+t^2}-2(t+1)\right]}{\frac{8t^3}{(1+t^2)^3}}$$

$$=-\frac{(1+2t)(1+t^2)}{2t},$$

或

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}}{\mathrm{d}t}[-(t^2+t+1)]\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{\frac{\mathrm{d}}{\mathrm{d}t}[-(t^2+t+1)]}{\frac{\mathrm{d}x}{\mathrm{d}t}}=-\frac{(1+2t)(1+t^2)}{2t}.$$

1.4.10 解

$$原式=\lim_{x\to0}\frac{f(1-\sqrt{\cos x})-f(0)}{(1-\sqrt{\cos x})-0}\cdot\lim_{x\to0}\frac{1-\sqrt{\cos x}}{\ln(1-x\sin x)}$$

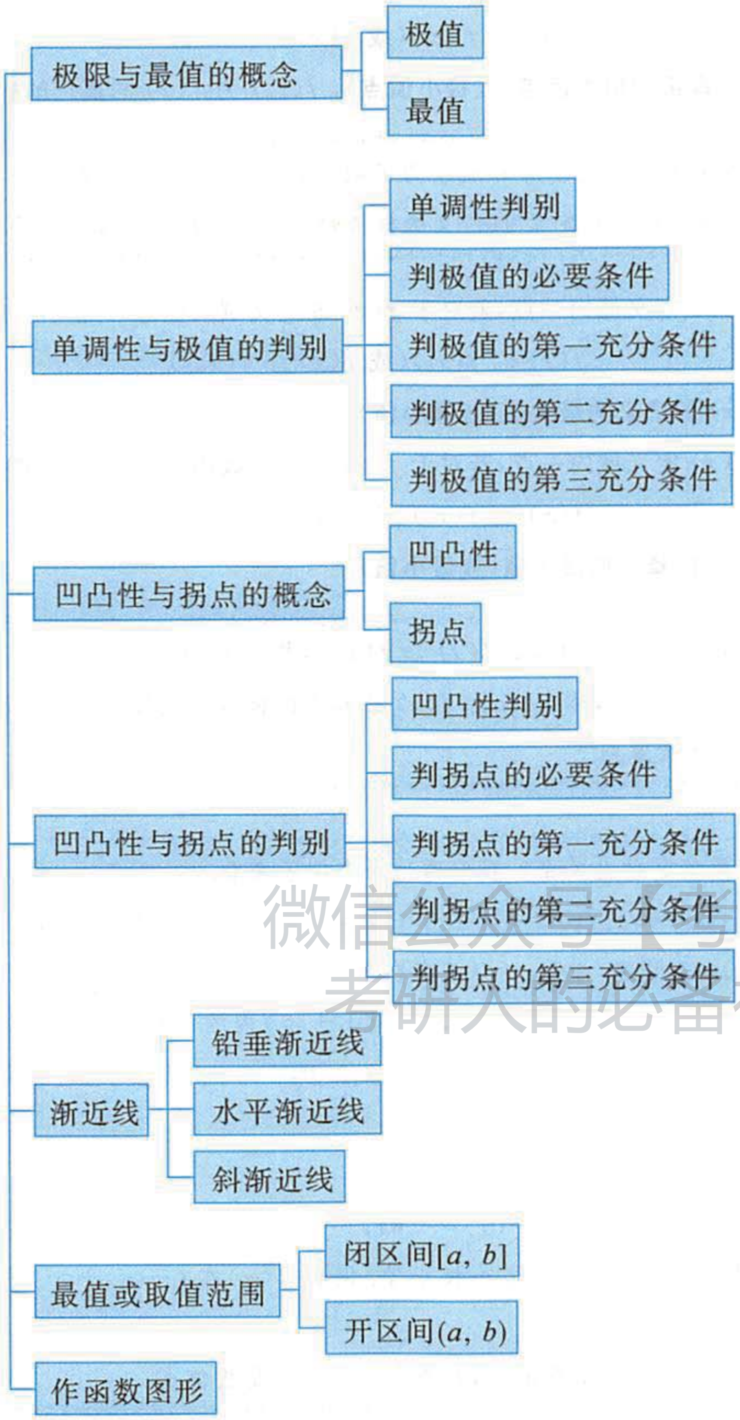
$$=f'(0)\lim_{x\to0}\frac{1-\sqrt{\cos x}}{\ln(1-x\sin x)}$$

$$\xlongequal{等价无穷小替换}f'(0)\lim_{x\to0}\frac{1-\cos x}{-x\sin x}\cdot\frac{1}{1+\sqrt{\cos x}}$$

$$=-\frac{1}{2}f'(0)\lim_{x\to0}\frac{1-\cos x}{x^2}=-\frac{1}{4}f'(0).$$

第5讲 一元函数微分学的几何应用

基础知识结构



基础内容精讲

一、极值与最值的概念

定义 1 若存在 x_0 的某个邻域，使得在该邻域内任意一点 x，均有

$$f(x) \leqslant f(x_0)（或 f(x) \geqslant f(x_0)）$$

成立，则称 x_0 为 $f(x)$ 的**广义的极大值点（或极小值点）**，$f(x_0)$ 为 $f(x)$ 的**广义的极大值（或极小值）**.

定义 2 若存在 x_0 的某个去心邻域，使得对于该邻域内任一异于 x_0 的点 x，均有

$$f(x) < f(x_0)（或 f(x) > f(x_0)）$$

成立，则称 x_0 为 $f(x)$ 的**真正的极大值点（或极小值点）**，$f(x_0)$ 为 $f(x)$ 的**真正的极大值（或极小值）**.

【注】 若在上述定义 2 中把去心邻域改为邻域，则当然有 $f(x) \leqslant f(x_0)$，但这个等号指的是仅在 $x = x_0$ 处取到，这一点是与定义 1（广义的极值）的一个主要区别，请注意区分.

定义 3 设 x_0 为 $f(x)$ 定义域内一点，若对于 $f(x)$ 的定义域内任意一点 x，均有

$$f(x) \leqslant f(x_0)（或 f(x) \geqslant f(x_0)）$$

成立，则称 $f(x_0)$ 为 $f(x)$ 的**广义的最大值（或最小值）**.

定义 4 设 x_0 为 $f(x)$ 定义域内一点，若对于 $f(x)$ 的定义域内任一异于 x_0 的点 x，均有

$$f(x) < f(x_0)（或 f(x) > f(x_0)）$$

成立，则称 $f(x_0)$ 为 $f(x)$ 的**真正的最大值（或最小值）**.

【注 1】 通俗来讲，带着等号的，如 $f(x) \geqslant f(x_0)$，称点 x_0 为 $f(x)$ 的广义的极值点或最值点；不带等号的，如 $f(x) > f(x_0)$，称点 x_0 为 $f(x)$ 的真正的极值点或最值点. 一般教材中用"广义"的较多，读者在读书中应注意鉴别.

【注 2】 极值和最值是什么关系？我们通过两个例子来看.

设 $f(x)=e^x, x\in[0,+\infty)$. $f(0)=e^0=1$ 为 $f(x)$ 在 $[0,+\infty)$ 内的最小值：$f(x)\geqslant f(0)$. 但 $f(x)$ 在 $[0,+\infty)$ 内没有极值. 细致说来，首先，$x=0$ 不存在哪种邻域 $U(0)$，使 $x\in U(0)$ 时，$f(x)\geqslant f(0)$，因为 $x=0$ 的左半邻域已超出 $f(x)$ 的定义域. 其次，对于 $(0,+\infty)$ 内的任意 x_0，不论 $U(x_0)$ 取得多么小，对于 $x\in U(x_0)$，并不总有 $f(x)\geqslant f(x_0)$，所以此 $f(x)$ 在 $[0,+\infty)$ 内无极小值，易见也无极大值. 所以，$f(x)$ 在 $[0,+\infty)$ 内无极值.

设 $f(x)=3x-x^3$，有

$$f'(x)=3(1-x^2),\quad f''(x)=-6x,\quad f'(\pm1)=0,\quad f''(\pm1)=\mp6,$$

所以 $f(1)=2$ 为极大值，$f(-1)=-2$ 为极小值. 但该 $f(x)$ 在 $(-\infty,+\infty)$ 内无最大值，也无最小值.

由此可见，极值点并不一定是最值点，最值点也不一定是极值点.

但是,下面这个结论是正确的:

如果 $f(x)$ 在区间 I 上有最值点 x_0,并且此最值点 x_0 不是区间 I 的端点而是 I 内部的点,那么此 x_0 必是 $f(x)$ 的一个极值点.

事实上,设 $f(x_0)$ 为 $f(x)$ 在 I 上的最大值,则对一切 $x\in I$,有 $f(x)\leqslant f(x_0)$,又因为 x_0 为 I 内部的点,故存在一个邻域 $U(x_0)\subset I$,当 $x\in U(x_0)$ 时,$f(x)\leqslant f(x_0)$. 按极大值的定义,$f(x_0)$ 为 $f(x)$ 的一个极大值.

【注 3】 结合第 3 讲的知识,一个常见的问题是:间断点可以是极值点吗? 答案是肯定的. 举四个例子供考生分析.

(1) $f(x)=\begin{cases}1, & x=0,\\ |x|, & x\neq 0,\end{cases}$ $x=0$ 是 $f(x)$ 的可去间断点,但它是 $f(x)$ 的极大值点.

(2) $f(x)=\begin{cases}1, & x>0,\\ -x, & x\leqslant 0,\end{cases}$ $x=0$ 是 $f(x)$ 的跳跃间断点,但它是 $f(x)$ 的极小值点.

(3) $f(x)=\begin{cases}-1, & x=0,\\ \dfrac{1}{x^2}, & x\neq 0,\end{cases}$ $x=0$ 是 $f(x)$ 的无穷间断点,但它是 $f(x)$ 的极小值点.

(4) $f(x)=\begin{cases}2, & x=0,\\ \sin\dfrac{1}{x}, & x\neq 0,\end{cases}$ $x=0$ 是 $f(x)$ 的振荡间断点,但它是 $f(x)$ 的极大值点.

二、单调性与极值的判别

1. 单调性的判别

若用导数工具,则若 $y=f(x)$ 在区间 I 上有 $f'(x)>0$,则 $y=f(x)$ 在 I 上**严格单调增加**;相应地,若 $y=f(x)$ 在区间 I 上有 $f'(x)<0$,则 $y=f(x)$ 在 I 上**严格单调减少**.

2. 一阶可导点是极值点的必要条件

设 $f(x)$ 在 $x=x_0$ 处可导,且在点 x_0 处取得极值,则必有 $f'(x_0)=0$.

3. 判别极值的第一充分条件

设 $f(x)$ 在 $x=x_0$ 处连续,且在 x_0 的某去心邻域 $\mathring{U}(x_0,\delta)$ 内可导.

①若 $x\in(x_0-\delta,x_0)$ 时,$f'(x)<0$,而 $x\in(x_0,x_0+\delta)$ 时,$f'(x)>0$,则 $f(x)$ 在 $x=x_0$ 处取得**极小值**;

②若 $x\in(x_0-\delta,x_0)$ 时,$f'(x)>0$,而 $x\in(x_0,x_0+\delta)$ 时,$f'(x)<0$,则 $f(x)$ 在 $x=x_0$ 处取得**极大值**;

③若 $f'(x)$ 在 $(x_0-\delta,x_0)$ 和 $(x_0,x_0+\delta)$ 内不变号,则点 x_0 不是极值点.

4. 判别极值的第二充分条件

设 $f(x)$ 在 $x=x_0$ 处二阶可导,且 $f'(x_0)=0,f''(x_0)\neq 0$.

①若 $f''(x_0)<0$,则 $f(x)$ 在 x_0 处取得**极大值**;②若 $f''(x_0)>0$,则 $f(x)$ 在 x_0 处取得**极小值**.

上述第二充分条件可以推广为第三充分条件.

5. 判别极值的第三充分条件

设 $f(x)$ 在 x_0 处 n 阶可导,且 $f^{(m)}(x_0)=0(m=1,2,\cdots,n-1),f^{(n)}(x_0)\neq 0(n\geqslant 2)$,则

①当 n 为偶数且 $f^{(n)}(x_0)<0$ 时，$f(x)$ 在 x_0 处取得**极大值**；

②当 n 为偶数且 $f^{(n)}(x_0)>0$ 时，$f(x)$ 在 x_0 处取得**极小值**.

【注】 上述第三充分条件的证明如下：由于 n 为偶数，令 $n=2k$，构造极限

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{(x-x_0)^{2k}}\overset{\text{洛必达法则}}{=\!=\!=}\lim_{x\to x_0}\frac{f'(x)}{2k(x-x_0)^{2k-1}}\overset{\text{洛必达法则}}{=\!=\!=}\cdots\overset{\text{洛必达法则}}{=\!=\!=}\lim_{x\to x_0}\frac{f^{(2k-1)}(x)}{(2k)!(x-x_0)}$$

$$=\lim_{x\to x_0}\frac{f^{(2k-1)}(x)-f^{(2k-1)}(x_0)}{(2k)!(x-x_0)}=\frac{1}{(2k)!}f^{(2k)}(x_0)\neq 0.$$

上述洛必达法则成立的依据是，最后的结果 $\frac{1}{(2k)!}f^{(2k)}(x_0)$ 是存在的.

当 $f^{(2k)}(x_0)<0$ 时，由函数极限的局部保号性 $\Rightarrow\frac{f(x)-f(x_0)}{(x-x_0)^{2k}}<0\Rightarrow f(x)<f(x_0)$，故 x_0 为极大值点；

当 $f^{(2k)}(x_0)>0$ 时，由函数极限的局部保号性 $\Rightarrow\frac{f(x)-f(x_0)}{(x-x_0)^{2k}}>0\Rightarrow f(x)>f(x_0)$，故 x_0 为极小值点.

三、凹凸性与拐点的概念

1. 凹凸性的定义

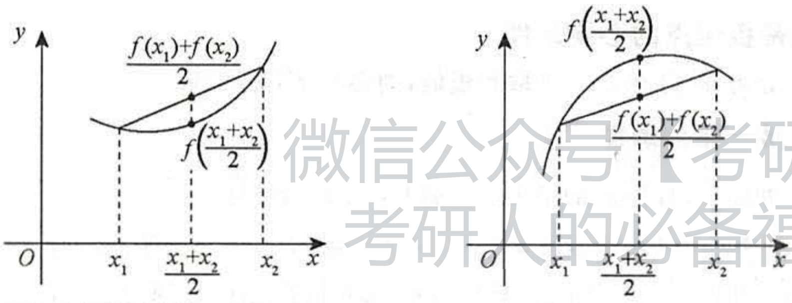
设函数 $f(x)$ 在区间 I 上连续. 如果对 I 上任意 x_1,x_2 两点，恒有

$$f\left(\frac{x_1+x_2}{2}\right)<\frac{f(x_1)+f(x_2)}{2},$$

则称 $y=f(x)$ 在 I 上的图形是(向上)凹的(或凹弧)，如图 1-5-1(a)所示；如果恒有

$$f\left(\frac{x_1+x_2}{2}\right)>\frac{f(x_1)+f(x_2)}{2},$$

则称 $y=f(x)$ 在 I 上的图形是(向上)凸的(或凸弧)，如图 1-5-1(b)所示.



图形上任意弧段位于弦的下方

$\frac{f(x_1)+f(x_2)}{2}>f\left(\frac{x_1+x_2}{2}\right)$

(a)

图形上任意弧段位于弦的上方

$\frac{f(x_1)+f(x_2)}{2}<f\left(\frac{x_1+x_2}{2}\right)$

(b)

图 1-5-1

2. 拐点定义

连续曲线的凹弧与凸弧的分界点称为该曲线的**拐点**.

四、凹凸性与拐点的判别

1. 判别凹凸性

设函数 $f(x)$ 在 I 上二阶可导.

①若在 I 上 $f''(x)>0$，则 $f(x)$ 在 I 上的图形是凹的；

②若在 I 上 $f''(x)<0$，则 $f(x)$ 在 I 上的图形是凸的.

2. 二阶可导点是拐点的必要条件

设 $f''(x_0)$ 存在，且点 $(x_0,f(x_0))$ 为曲线上的拐点，则 $f''(x_0)=0$.

3. 判别拐点的第一充分条件

设 $f(x)$ 在点 $x=x_0$ 处连续，在点 $x=x_0$ 的某去心邻域 $\mathring{U}(x_0,\delta)$ 内二阶导数存在，且在该点的左、右邻域内 $f''(x)$ 变号（无论是由正变负，还是由负变正），则点 $(x_0,f(x_0))$ 为曲线上的拐点.

【注】 $(x_0,f(x_0))$ 为曲线 $y=f(x)$ 上的拐点时，并不要求 $f(x)$ 在点 x_0 的导数存在，如 $y=\sqrt[3]{x}$ 在 $x=0$ 的情形.

4. 判别拐点的第二充分条件

设 $f(x)$ 在 $x=x_0$ 的某邻域内三阶可导，且 $f''(x_0)=0$，$f'''(x_0)\neq0$，则 $(x_0,f(x_0))$ 为拐点.

5. 判别拐点的第三充分条件

设 $f(x)$ 在 x_0 处 n 阶可导，且 $f^{(m)}(x_0)=0\,(m=2,\cdots,n-1)$，$f^{(n)}(x_0)\neq0\,(n\geqslant3)$，则当 n 为奇数时，$(x_0,f(x_0))$ 为拐点.

【注1】 上述第三充分条件的证明如下：由于 n 为奇数，令 $n=2k+1$，构造极限

$$\lim_{x\to x_0}\frac{f''(x)}{(x-x_0)^{2k-1}}\overset{\text{洛必达法则}}{=\!=\!=\!=}\cdots\overset{\text{洛必达法则}}{=\!=\!=\!=}\lim_{x\to x_0}\frac{f^{(2k)}(x)}{(2k-1)!(x-x_0)}$$

$$=\lim_{x\to x_0}\frac{f^{(2k)}(x)-f^{(2k)}(x_0)}{(2k-1)!(x-x_0)}=\frac{1}{(2k-1)!}f^{(2k+1)}(x_0)\neq0.$$

上述洛必达法则成立的依据是，最后的结果 $\frac{1}{(2k-1)!}f^{(2k+1)}(x_0)$ 是存在的.

当 $f^{(2k+1)}(x_0)>0$ 时，由函数极限的局部保号性，得

$$\frac{f''(x)}{(x-x_0)^{2k-1}}>0,$$

当 $x\to x_0^+$ 时，$f''(x)>0$；当 $x\to x_0^-$ 时，$f''(x)<0$，故 $(x_0,f(x_0))$ 为拐点.

【注2】 由上述证明过程可知，第三充分条件不需要 $f'(x_0)=0$ 这个条件，有些教材和辅导书写出这个条件是不必要的.

五、渐近线

1. 铅垂渐近线

若$\lim\limits_{x\to x_0^+}f(x)=\infty$（或$\lim\limits_{x\to x_0^-}f(x)=\infty$），则 $x=x_0$为一条铅垂渐近线.

【注】 此处的 x_0一般是函数的无定义点.

2. 水平渐近线

若$\lim\limits_{x\to+\infty}f(x)=y_1$，则 $y=y_1$为一条水平渐近线；若$\lim\limits_{x\to-\infty}f(x)=y_2$，则 $y=y_2$为一条水平渐近线；

若$\lim\limits_{x\to+\infty}f(x)=\lim\limits_{x\to-\infty}f(x)=y_0$，则 $y=y_0$为一条水平渐近线.

3. 斜渐近线

若$\lim\limits_{x\to+\infty}\dfrac{f(x)}{x}=k_1$，$\lim\limits_{x\to+\infty}[f(x)-k_1x]=b_1$，则 $y=k_1x+b_1$是曲线 $y=f(x)$的一条斜渐近线；

若$\lim\limits_{x\to-\infty}\dfrac{f(x)}{x}=k_2$，$\lim\limits_{x\to-\infty}[f(x)-k_2x]=b_2$，则 $y=k_2x+b_2$是曲线 $y=f(x)$的一条斜渐近线；

若$\lim\limits_{x\to+\infty}\dfrac{f(x)}{x}=\lim\limits_{x\to-\infty}\dfrac{f(x)}{x}=k$，$\lim\limits_{x\to+\infty}[f(x)-kx]=\lim\limits_{x\to-\infty}[f(x)-kx]=b$，则 $y=kx+b$ 是曲线 $y=f(x)$的一条斜渐近线.

六、最值或取值范围

1. 求闭区间$[a,b]$上连续函数 $f(x)$的最大值 M 和最小值 m

①求出 $f(x)$在(a,b)内的可疑点——驻点与不可导点，并求出这些可疑点处的函数值；

②求出端点的函数值 $f(a)$和 $f(b)$；

③比较以上所求得的所有函数值，其中最大者为 $f(x)$在$[a,b]$上的最大值 M，最小者为 $f(x)$在$[a,b]$上的最小值 m.

【注】 有时这类问题也可命制为“求连续函数 $f(x)$在闭区间$[a,b]$上的值域$[m,M]$”.见例1.5.8.

2. 求开区间(a,b)内连续函数 $f(x)$的最值或者取值范围

①求出 $f(x)$在(a,b)内的可疑点——驻点与不可导点，并求出这些可疑点处的函数值.

②求(a,b)两端的单侧极限：若 a,b 为有限常数，则求$\lim\limits_{x\to a^+}f(x)$与$\lim\limits_{x\to b^-}f(x)$；若 a 为$-\infty$，则求$\lim\limits_{x\to-\infty}f(x)$；若 b 为$+\infty$，则求$\lim\limits_{x\to+\infty}f(x)$. 记以上所求左端极限为 A，右端极限为 B.

③比较①，②所得结果，确定最值或取值范围.

【注】 这类问题有时没有最大值、最小值.

七、作函数图形

给出函数 $f(x)$，作图的一般步骤：

①确定函数 $f(x)$ 的定义域，并考查它是否有奇偶对称性；

②求出 $f'(x)$，$f''(x)$，用 $f(x)$ 的无定义点，$f'(x)=0$ 的点，$f'(x)$ 不存在的点，$f''(x)=0$ 的点，$f''(x)$ 不存在的点，将定义域划分为若干子区间，确定函数图形在各个子区间上的单调性与凹凸性，进而确定函数的极值点和拐点；

③确定渐近线（如果有的话）；

④作出函数图形.

这是基本功，一定要重视.

基础例题精解

导数的几何应用主要指："三点、两性、一线"."三点"是指：极值点、最值点、拐点."两性"是指：单调性、凹凸性."一线"是指：渐近线.这部分内容知识点众多，各知识点之间又联系紧密，故很容易命制综合题.

例 1.5.1 设 $f(x)$ 在 $[0,+\infty)$ 内连续，$f'(x)$ 为 $(0,+\infty)$ 内的单调增加函数，$f(0)=0$，试讨论函数 $\frac{f(x)}{x}$ 在 $(0,+\infty)$ 内的增减性.

解 令 $g(x)=\frac{f(x)}{x}$，则

$$g'(x)=\frac{f'(x)}{x}-\frac{f(x)}{x^2}=\frac{1}{x^2}[xf'(x)-xf'(\xi)],$$

这里使用了拉格朗日中值定理，其中 $0<\xi<x$. 因为 $f'(x)$ 为 $(0,+\infty)$ 内的单调增加函数，于是

$$f'(x)>f'(\xi),\quad g'(x)>0,$$

所以 $g(x)=\frac{f(x)}{x}$ 在 $(0,+\infty)$ 内单调增加.

例 1.5.2 求函数 $f(x)=x-\frac{3}{2}x^{\frac{2}{3}}$ 的单调区间和极值.

解 定义域为 $(-\infty,+\infty)$，因为

$$f'(x)=1-x^{-\frac{1}{3}}=\frac{\sqrt[3]{x}-1}{\sqrt[3]{x}},$$

令 $f'(x)=0$，得驻点 $x_1=1$，$f'(x)$ 不存在的点为 $x_2=0$. 列表如下：

x	$(-\infty,0)$	0	$(0,1)$	1	$(1,+\infty)$
$f'(x)$	+	不存在	−	0	+
$f(x)$	↗	极大值	↘	极小值	↗

由此可知函数 $f(x)$ 在 $(-\infty,0)$，$(1,+\infty)$ 内单调增加，在 $(0,1)$ 内单调减少；在 $x=0$ 处取得极大值 $f(0)=0$，在 $x=1$ 处取得极小值 $f(1)=-\dfrac{1}{2}$.

例 1.5.3 试问 a 为何值时，函数 $f(x)=a\sin x+\dfrac{1}{3}\sin 3x$ 在 $x=\dfrac{\pi}{3}$ 处取得极值？并求此极值.

解 因 $f'(x)=a\cos x+\cos 3x$. 令 $f'\left(\dfrac{\pi}{3}\right)=0$，即 $\dfrac{a}{2}-1=0$，得 $a=2$.

又当 $a=2$ 时，

$$f''(x)=-2\sin x-3\sin 3x,\quad f''\left(\frac{\pi}{3}\right)=-\sqrt{3}<0,$$

所以 $f(x)$ 在 $x=\dfrac{\pi}{3}$ 处取得极大值 $f\left(\dfrac{\pi}{3}\right)=\sqrt{3}$.

例 1.5.4 设函数 $y=y(x)$ 由方程 $y^3+xy^2+x^2y+6=0$ 确定，求 $y(x)$ 的极值.

解 在 $y^3+xy^2+x^2y+6=0$ 两边关于 x 求导，得

$$3y^2y'+y^2+2xyy'+2xy+x^2y'=0,$$

令 $y'=0$，得 $y=-2x$ 或 $y=0$（不适合方程，舍去）.

将 $y=-2x$ 代入方程得 $-6x^3+6=0$，解得 $x=1$，$y(1)=-2$，且 $y'(1)=0$.

在 $3y^2y'+y^2+2xyy'+2xy+x^2y'=0$ 两边关于 x 求导，得

$$(3y^2+2xy+x^2)y''+2(3y+x)(y')^2+4(y+x)y'+2y=0,$$

求得 $y''(1)=\dfrac{4}{9}>0$. 所以 $x=1$ 是函数 $y(x)$ 的极小值点，极小值为 $y(1)=-2$.

例 1.5.5 曲线 $y=(x-1)(x-2)^2(x-3)^3(x-4)^4$ 的一个拐点是（　　）.

(A)(1,0)　　(B)(2,0)　　(C)(3,0)　　(D)(4,0)

解 应选(C).

由于 $(x-3)^3$ 的二阶导数为 $6(x-3)$，它在点 $x=3$ 的两侧改变正、负号，故该因式为解决问题的关键.

设 $g(x)=(x-1)(x-2)^2(x-4)^4$，则 $y=(x-3)^3g(x)$，于是

$$\begin{aligned} y'&=3(x-3)^2g(x)+(x-3)^3g'(x), \\ y''&=6(x-3)g(x)+3(x-3)^2g'(x)+3(x-3)^2g'(x)+(x-3)^3g''(x) \\ &=(x-3)\cdot[6g(x)+6(x-3)g'(x)+(x-3)^2g''(x)]. \end{aligned}$$

再令 $h(x)=6g(x)+6(x-3)g'(x)+(x-3)^2g''(x)$，由 $h(3)=6g(3)=12>0$ 及连续函数的保号性可知，存在 $\delta>0$，使得当 $3-\delta<x<3+\delta$ 时，$h(x)>0$，这样，当 $3-\delta<x<3$ 时，$y''<0$，而当 $3<x<3+\delta$ 时，$y''>0$. 即 $(3,0)$ 是拐点. 故选择(C).

【注】 要善于在纷繁复杂的表达式中，抓住主要矛盾，从而找到解决问题的突破口. 与此类似的问题还有前面的第 4 讲的例 1.4.8.

例 1.5.6 设函数 $y=y(x)$ 由参数方程 $\begin{cases} x=\dfrac{1}{3}t^3+t+\dfrac{1}{3}, \\ y=\dfrac{1}{3}t^3-t+\dfrac{1}{3} \end{cases}$ 确定，求函数 $y=y(x)$ 的极值和曲线 $y=y(x)$ 的凹凸区间及拐点.

解 $$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y'(t)}{x'(t)}=\frac{t^2-1}{t^2+1},\quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^2-1}{t^2+1}\right)\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{4t}{(t^2+1)^3}.$$

令$\frac{\mathrm{d}y}{\mathrm{d}x}=0$，得 $t=\pm1$. 当 $t=1$ 时，$x=\frac{5}{3}$，$y=-\frac{1}{3}$；当 $t=-1$ 时，$x=-1$，$y=1$.

令$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=0$，得 $t=0$，此时 $x=\frac{1}{3}$，$y=\frac{1}{3}$. 列表如下：

t	$(-\infty,-1)$	-1	$(-1,0)$	0	$(0,1)$	1	$(1,+\infty)$
x	$(-\infty,-1)$	-1	$\left(-1,\frac{1}{3}\right)$	$\frac{1}{3}$	$\left(\frac{1}{3},\frac{5}{3}\right)$	$\frac{5}{3}$	$\left(\frac{5}{3},+\infty\right)$
y'	$+$	0	$-$	$-$	$-$	0	$+$
y''	$-$	$-$	$-$	0	$+$	$+$	$+$

由此可知，函数 $y(x)$的极大值为 $y(-1)=1$，极小值为 $y\left(\frac{5}{3}\right)=-\frac{1}{3}$；

曲线 $y=y(x)$的凹区间为$\left(\frac{1}{3},+\infty\right)$，凸区间为$\left(-\infty,\frac{1}{3}\right)$；曲线 $y=y(x)$的拐点为$\left(\frac{1}{3},\frac{1}{3}\right)$.

例 1.5.7 求曲线 $y=\frac{1}{x}+\ln(1+\mathrm{e}^x)$的渐近线.

解 因为

$$\lim_{x\to0}\left[\frac{1}{x}+\ln(1+\mathrm{e}^x)\right]=\infty,$$

所以直线 $x=0$ 是曲线 $y=\frac{1}{x}+\ln(1+\mathrm{e}^x)$的一条铅垂渐近线.

因为

$$\lim_{x\to-\infty}\left[\frac{1}{x}+\ln(1+\mathrm{e}^x)\right]=0,$$

所以直线 $y=0$ 是曲线 $y=\frac{1}{x}+\ln(1+\mathrm{e}^x)$在 $x\to-\infty$时的一条水平渐近线.

因为

$$\lim_{x\to+\infty}\frac{\frac{1}{x}+\ln(1+\mathrm{e}^x)}{x}=\lim_{x\to+\infty}\left[\frac{1}{x^2}+\frac{x+\ln(1+\mathrm{e}^{-x})}{x}\right]=1,$$

且

$$\lim_{x\to+\infty}\left[\frac{1}{x}+\ln(1+\mathrm{e}^x)-x\right]=\lim_{x\to+\infty}\left[\frac{1}{x}+\ln(1+\mathrm{e}^{-x})\right]=0,$$

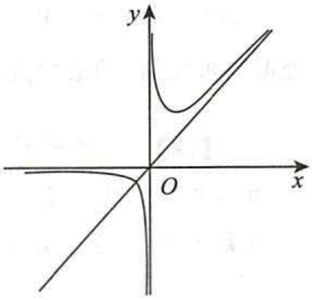


图 1-5-2

所以直线 $y=x$ 是曲线 $y=\frac{1}{x}+\ln(1+\mathrm{e}^x)$在 $x\to+\infty$时的一条斜渐近线. 如图 1-5-2 所示.

例 1.5.8 求 $f(x)=|x|\mathrm{e}^x$ 在区间$[-2,1]$上的值域(即求最大值和最小值).

解 将 $f(x)$写成分段函数为 $f(x)=\begin{cases}x\mathrm{e}^x, & 0\leqslant x\leqslant1,\\ -x\mathrm{e}^x, & -2\leqslant x<0,\end{cases}$ 因此

当 $x\neq0$ 时，$f'(x)=\begin{cases}(x+1)\mathrm{e}^x, & 0<x\leqslant1,\\ -(x+1)\mathrm{e}^x, & -2\leqslant x<0.\end{cases}$

当 $x=0$ 时，

$$f'_-(0)=\lim_{x\to0^-}\frac{f(x)-f(0)}{x-0}=\lim_{x\to0^-}\frac{-x\mathrm{e}^x-0}{x}=-1,$$

$$f'_+(0)=\lim_{x\to0^+}\frac{f(x)-f(0)}{x-0}=\lim_{x\to0^+}\frac{xe^x-0}{x}=1.$$

因为 $f'_-(0)\neq f'_+(0)$，所以 $f(x)$ 在 $x=0$ 处不可导.

令 $f'(x)=0$，解得 $x=-1$. 所以 $f(x)$ 在 $(-2,1)$ 内的驻点和不可导点分别为 $x=-1$ 和 $x=0$. 比较它们及端点的函数值：

$$f(-2)=\frac{2}{e^2},\quad f(-1)=\frac{1}{e},\quad f(0)=0,\quad f(1)=e.$$

知 $f(1)=e$ 为最大值，$f(0)=0$ 为最小值，即 $f(x)$ 在区间 $[-2,1]$ 上的值域为 $[0,e]$.

例 1.5.9 求函数 $f(x)=\int_0^{x^2}(2-t)e^{-t}dt$ 的最大值和最小值.

分析 本题是一道综合题，既考查了变上限定积分函数的导数、反常积分的计算，又考查了导数应用. 由于 $f(x)$ 是 $(-\infty,+\infty)$ 内连续的偶函数，求最大值、最小值时，只需求出驻点处的函数值，然后与 $f(0)$，$\lim\limits_{x\to+\infty}f(x)$ 比较.

解 因 $f(x)$ 是偶函数，故只需讨论 $f(x)$ 在 $[0,+\infty)$ 上的最大值与最小值.

令 $f'(x)=2x(2-x^2)e^{-x^2}=0$，则 $x=\sqrt{2}$ 是 $f(x)$ 在 $(0,+\infty)$ 内的唯一驻点，所对应的函数值为

$$f(\sqrt{2})=\int_0^2(2-t)e^{-t}dt=-(2-t)e^{-t}\Big|_0^2-\int_0^2e^{-t}dt=1+e^{-2}.$$

因为 $\int_0^{+\infty}(2-t)e^{-t}dt=-(2-t)e^{-t}\Big|_0^{+\infty}+e^{-t}\Big|_0^{+\infty}=2-1=1$，以及 $f(0)=0$，故 $x=0$ 是最小值点，$f(x)$ 的最小值为 $f(0)=0$；而 $f(\sqrt{2})=1+e^{-2}$ 为其最大值.

例 1.5.10 求数列 $\{\sqrt[n]{n}\}$ 的最大项.

解 设 $f(x)=x^{\frac{1}{x}}(x>0)$，则

$$f'(x)=x^{\frac{1}{x}}\cdot\frac{1-\ln x}{x^2},$$

令 $f'(x)=0$，得唯一驻点 $x=e$，当 $x\in(0,e)$ 时，$f'(x)>0$，当 $x\in(e,+\infty)$ 时，$f'(x)<0$，所以 $f(x)$ 在 $x=e$ 处取得极大值，即最大值为 $f(e)=e^{\frac{1}{e}}$，又因为 $\sqrt{2}<\sqrt[3]{3}$，故 $\sqrt[3]{3}$ 是数列 $\{\sqrt[n]{n}\}$ 的最大项.

> **【注】** 如果遇到最小值和最大值的实际问题，则首先应建立目标函数(即欲求其最值的那个函数)，并确定其定义区间，将它转化为函数的最值问题. 特别地，如果所考虑的实际问题存在最小值或最大值，并且所建立的目标函数 $f(x)$ 有唯一的驻点 x_0，则 $f(x_0)$ 即为所求的最小值或最大值.

例 1.5.11 作函数 $f(x)=(2+x)e^{\frac{1}{x}}$ 的图像.

解 函数 $f(x)$ 的定义域为 $(-\infty,0)\cup(0,+\infty)$.

令 $f'(x)=\frac{(x+1)(x-2)}{x^2}e^{\frac{1}{x}}=0$，得驻点 $x_1=-1$，$x_2=2$，及 $f'(x)$ 不存在的点 $x_3=0$.

因此可知，函数 $f(x)$ 的单调增区间为 $(-\infty,-1)$ 和 $(2,+\infty)$，单调减区间为 $(-1,0)$ 和 $(0,2)$；极大值为 $f(-1)=\frac{1}{e}$，极小值为 $f(2)=4\sqrt{e}$.

令 $f''(x)=\frac{5x+2}{x^4}e^{\frac{1}{x}}=0$，得 $x_4=-\frac{2}{5}$ 及 $f''(x)$ 不存在的点 $x_3=0$.

所以函数 $f(x)$ 的凸区间为 $\left(-\infty,-\frac{2}{5}\right)$，凹区间为 $\left(-\frac{2}{5},0\right)$ 和 $(0,+\infty)$；曲线 $y=f(x)$ 的拐点为 $\left(-\frac{2}{5},\frac{8}{5}e^{-\frac{5}{2}}\right)$.

由

$$\lim_{x\to 0^-} f(x)=\lim_{x\to 0^-}(2+x)\mathrm{e}^{\frac{1}{x}}=0,$$

$$\lim_{x\to 0^+} f(x)=\lim_{x\to 0^+}(2+x)\mathrm{e}^{\frac{1}{x}}=+\infty,$$

故直线 $x=0$ 是曲线 $y=f(x)$ 的铅垂渐近线.

又

$$\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\frac{(2+x)\mathrm{e}^{\frac{1}{x}}}{x}=1,$$

$$\begin{aligned}\lim_{x\to\infty}[f(x)-x]&=\lim_{x\to\infty}[(2+x)\mathrm{e}^{\frac{1}{x}}-x]\\&=\lim_{x\to\infty}x\left[\left(1+\frac{2}{x}\right)\mathrm{e}^{\frac{1}{x}}-1\right]\\&=\lim_{x\to\infty}\frac{\left(1+\frac{2}{x}\right)\mathrm{e}^{\frac{1}{x}}-1}{\frac{1}{x}}\\&\xlongequal{t=\frac{1}{x}}\lim_{t\to 0}\frac{(1+2t)\mathrm{e}^t-1}{t}\\&=\lim_{t\to 0}\frac{2\mathrm{e}^t+(1+2t)\mathrm{e}^t}{1}=3,\end{aligned}$$

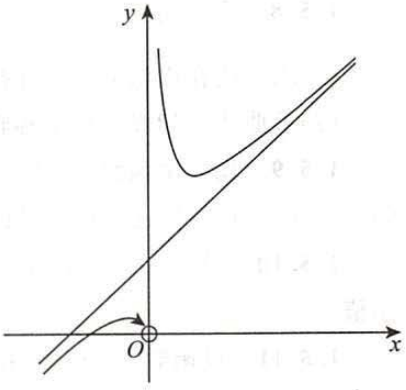


图 1-5-3

故直线 $y=x+3$ 是曲线 $y=f(x)$ 在 $x\to\infty$ 时的斜渐近线.

作图如图 1-5-3 所示.

基础习题精练

习题

1.5.1　已知 $f(x)$ 在 $x=0$ 的某个邻域内连续，且 $\lim\limits_{x\to 0}\frac{f(x)}{1-\cos x}=2$，则在点 $x=0$ 处 $f(x)$（　　）.

(A)不可导　　(B)可导，且 $f'(0)\neq 0$

(C)取得极大值　　(D)取得极小值

1.5.2　设函数 $f(x)$ 在 $(-\infty,+\infty)$ 内有定义，$x_0(x_0\neq 0)$ 是函数 $f(x)$ 的极大值点，则（　　）.

(A) x_0 必是 $f(x)$ 的驻点　　(B) $-x_0$ 必是 $-f(-x)$ 的极小值点

(C) $-x_0$ 必是 $-f(x)$ 的极小值点　　(D)对一切 x 都有 $f(x)\leqslant f(x_0)$

1.5.3　设函数 $y=f(x)$ 具有二阶导数，且 $f'(x)>0$，$f''(x)>0$，Δx 为自变量 x 在点 x_0 处的增量，Δy 与 $\mathrm{d}y$ 分别为 $f(x)$ 在点 x_0 处对应的增量与微分. 若 $\Delta x>0$，则（　　）.

(A) $0<\mathrm{d}y<\Delta y$　　(B) $0<\Delta y<\mathrm{d}y$　　(C) $\Delta y<\mathrm{d}y<0$　　(D) $\mathrm{d}y<\Delta y<0$

1.5.4　已知曲线 $y=x^3-3a^2x+b$ 与 x 轴相切，则 b^2 可通过 a 表示为 $b^2=$________.

1.5.5　求曲线 $y=f(x)=x\mathrm{e}^{\frac{1}{x^2}}$ 的渐近线.

1.5.6　设 $f(x)=x\mathrm{e}^x$，求 $f^{(n)}(x)$ 的极值点和极值.

1.5.7 设 $y=\dfrac{x^3+4}{x^2}$，求

(1)函数的增减区间及极值；

(2)函数图像的凹凸区间及拐点；

(3)渐近线；

(4)作出其图形.

1.5.8 给定曲线 $y=\dfrac{1}{x^2}$.

(1)求曲线在横坐标为 x_0 的点处的切线方程；

(2)求曲线的切线被两坐标轴所截线段的最短长度.

1.5.9 设三次函数 $f(x)=x^3+3ax^2+3bx+c$ 在 $x=2$ 处有极值，其图形在 $x=1$ 处的切线与直线 $6x+2y+5=0$ 平行，试问极大值比极小值大多少？

1.5.10 设 $a>1$，$f(t)=a^t-at$ 在 $(-\infty,+\infty)$ 内的驻点为 $t(a)$，问 a 为何值时，$t(a)$ 最小，并求出最小值.

1.5.11 设函数 $y=y(x)$ 由方程 $y\ln y-x+y=0$ 确定，试判断曲线 $y=y(x)$ 在点 $(1,1)$ 附近的凹凸性.

1.5.12 设曲线 $f(x)=x^3+ax$ 与 $g(x)=bx^2+c$ 都通过点 $(-1,0)$，且在点 $(-1,0)$ 有公共切线，求 a，b，c 的值.

解 答

1.5.1 (D) **解** 由于 $\lim\limits_{x\to0}\dfrac{f(x)}{1-\cos x}=2$，且 $\lim\limits_{x\to0}(1-\cos x)=0$，故 $\lim\limits_{x\to0}f(x)=0=f(0)$.

又 $\lim\limits_{x\to0}\dfrac{f(x)}{1-\cos x}=\lim\limits_{x\to0}\dfrac{2f(x)}{x^2}=2$，故 $\lim\limits_{x\to0}\dfrac{f(x)}{x^2}=1>0$，由极限的局部保号性可知在 $x=0$ 的去心邻域内有 $\dfrac{f(x)}{x^2}>0$，即 $f(x)>0$，故 $f(x)$ 在 $x=0$ 处取极小值，从而选(D).

【注】 易知 $f(0)=0$，$\lim\limits_{x\to0}\dfrac{f(x)}{1-\cos x}=\lim\limits_{x\to0}\dfrac{f(x)-f(0)}{x}\cdot\dfrac{x}{1-\cos x}=2$. 由于 $\lim\limits_{x\to0}\dfrac{x}{1-\cos x}=\infty$，则必有 $\lim\limits_{x\to0}\dfrac{f(x)-f(0)}{x}=0$，即 $f'(0)=0$. 排除(A)，(B).

1.5.2 (B) **解** 对任意的 $x\in\mathring{U}(x_0,\delta)$，有 $f(x)<f(x_0)$，故 $-f(x)>-f(x_0)=-f[-(-x_0)]$，故 $-x_0$ 必是 $-f(-x)$ 的极小值点，故选(B).

1.5.3 (A) **解** 由 $f'(x)>0$，$f''(x)>0$ 知，曲线 $y=f(x)$ 单调增加且曲线图形是向上凹的，当 $\Delta x>0$ 时，可画出图形(如图 1-5-4).

由微分 $\mathrm{d}y$ 及函数增量 Δy 的几何意义可从图形中得到正确的答案为(A).

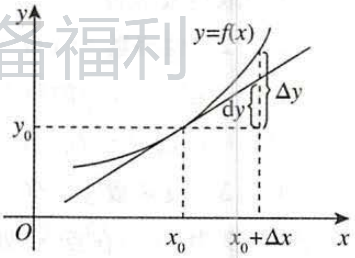


图 1-5-4

1.5.4 $4a^6$ **解** 因为曲线 $y=x^3-3a^2x+b$ 与 x 轴相切，所以，在切点处

$$y'=3x^2-3a^2=0,$$

由此可解出切点横坐标 $x=a$ 或 $x=-a$，且在切点处有

$$y(a)=a^3-3a^3+b=0\text{ 或 }y(-a)=-a^3+3a^3+b=0,$$

所以 $b=2a^3$ 或 $b=-2a^3$，即 $b^2=4a^6$.

1.5.5 分析 当极限 $\lim\limits_{x\to+\infty} y$ 存在或极限 $\lim\limits_{x\to-\infty} y$ 存在⇒曲线 $y=f(x)$ 存在水平渐近线；

当 $\lim\limits_{x\to x_0^+} y=\infty$ 或 $\lim\limits_{x\to x_0^-} y=\infty$⇒曲线 $y=f(x)$ 存在铅垂渐近线；

当极限 $\lim\limits_{x\to+\infty}\dfrac{f(x)}{x}=k\neq0$ 或极限 $\lim\limits_{x\to-\infty}\dfrac{f(x)}{x}=k\neq0$，$\lim\limits_{x\to\infty}[f(x)-kx]=b$ 存在⇒曲线 $y=f(x)$ 存在斜渐近线. 即斜渐近线方程为 $y=kx+b$.

所以，求曲线的渐近线要讨论三种极限.

解 因为 $\lim\limits_{x\to\infty} xe^{\frac{1}{x^2}}=\infty$，所以曲线 $y=f(x)$ 无水平渐近线.

因为 $\lim\limits_{x\to0} xe^{\frac{1}{x^2}} \xlongequal{x=\frac{1}{t}} \lim\limits_{t\to\infty}\dfrac{e^{t^2}}{t}=\lim\limits_{t\to\infty}2te^{t^2}=\infty$，所以曲线 $y=f(x)$ 有铅垂渐近线 $x=0$.

又
$$\lim_{x\to\infty}\frac{xe^{\frac{1}{x^2}}}{x}=\lim_{x\to\infty}e^{\frac{1}{x^2}}=1,\quad \lim_{x\to\infty}\left(xe^{\frac{1}{x^2}}-x\right)=\lim_{x\to\infty}\frac{e^{\frac{1}{x^2}}-1}{\frac{1}{x}}\xlongequal{x=\frac{1}{t}}\lim_{t\to0}\frac{e^{t^2}-1}{t}=\lim_{t\to0}2te^{t^2}=0,$$

所以曲线 $y=f(x)$ 有斜渐近线 $y=x$.

1.5.6 解 由 $f(x)=xe^x$ 求得
$$f'(x)=(1+x)e^x,\quad f''(x)=(2+x)e^x,\quad f'''(x)=(3+x)e^x,$$

从而可得
$$f^{(n)}(x)=(n+x)e^x,\quad f^{(n+1)}(x)=(n+1+x)e^x,\quad f^{(n+2)}(x)=(n+2+x)e^x.$$

令 $f^{(n+1)}(x)=(n+1+x)e^x=0$，得函数 $f^{(n)}(x)=(n+x)e^x$ 的驻点 $x=-(n+1)$. 又
$$f^{(n+2)}[-(n+1)]=e^{-n-1}>0,$$

所以 $x=-(n+1)$ 是函数 $f^{(n)}(x)=(n+x)e^x$ 的极小值点，极小值为
$$f^{(n)}[-(n+1)]=-e^{-(n+1)}.$$

1.5.7 解 定义域为 $(-\infty,0)\cup(0,+\infty)$. 当 $x=-\sqrt[3]{4}$ 时，$y=0$.

(1) $y'=1-\dfrac{8}{x^3}$，故驻点为 $x=2$. 又

x	$(-\infty,0)$	$(0,2)$	2	$(2,+\infty)$
y'	+	−	0	+
y	↗	↘	3	↗

所以，$(-\infty,0)$ 及 $(2,+\infty)$ 为增区间，$(0,2)$ 为减区间，$x=2$ 为极小值点，极小值为 $y=3$.

(2) 由 $y''=\dfrac{24}{x^4}>0$，故 $(-\infty,0)$，$(0,+\infty)$ 均为凹区间，无拐点.

(3) 因 $\lim\limits_{x\to0}\dfrac{x^3+4}{x^2}=+\infty$，$\lim\limits_{x\to\infty}\dfrac{y}{x}=\lim\limits_{x\to\infty}\dfrac{x^3+4}{x^3}=1=k$，
$$\lim_{x\to\infty}(y-kx)=\lim_{x\to\infty}\left(\frac{x^3+4}{x^2}-x\right)=0=b,$$

所以，$x=0$ 为铅垂渐近线，$y=x$ 为斜渐近线.

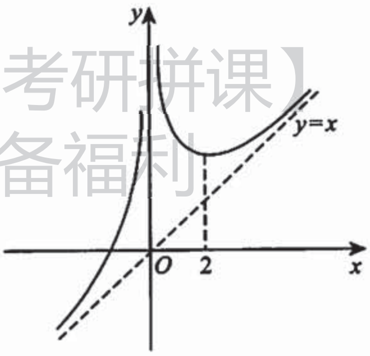


图 1-5-5

(4) 函数的图形如图 1-5-5 所示.

1.5.8 解 (1) 由 $y'=-\dfrac{2}{x^3}$ 可知曲线 $y=\dfrac{1}{x^2}$ 在横坐标为 x_0 的点处的切线方程为 $y-\dfrac{1}{x_0^2}=-\dfrac{2}{x_0^3}(x-x_0)$.

(2) 由切线方程 $y-\dfrac{1}{x_0^2}=-\dfrac{2}{x_0^3}(x-x_0)$ 分别令 $y=0$，$x=0$ 可求得该切线在 x 轴、y 轴上的截距分别为

$$X=\frac{3}{2}x_0,\quad Y=\frac{3}{x_0^2}.$$

设该切线被两坐标轴所截线段长度为 L，则 $L^2=X^2+Y^2=\frac{9}{4}x_0^2+\frac{9}{x_0^4}$.

令 $\frac{\mathrm{d}L^2}{\mathrm{d}x_0}=\frac{9}{2}x_0-\frac{36}{x_0^5}=0$，得驻点 $x_0=\pm\sqrt{2}$.

又 $\frac{\mathrm{d}^2L^2}{\mathrm{d}x_0^2}=\frac{9}{2}+\frac{180}{x_0^6}$，显然 $\left.\frac{\mathrm{d}^2L^2}{\mathrm{d}x_0^2}\right|_{x_0=\pm\sqrt{2}}>0$，由此可知，$L^2$ 在 $x_0=\pm\sqrt{2}$处取得极小值，即最小值，$L_{\min}^2=\frac{27}{4}$，故 $L_{\min}=\frac{3}{2}\sqrt{3}$.

1.5.9 解 根据所给条件可以确定函数中参数的值.

因为函数 $f(x)=x^3+3ax^2+3bx+c$ 在 $x=2$ 处有极值，所以 $f'(2)=\left.(3x^2+6ax+3b)\right|_{x=2}=0$，即

$$4+4a+b=0. \qquad ①$$

又因为 $f(x)=x^3+3ax^2+3bx+c$ 的图形在 $x=1$ 处的切线与直线 $6x+2y+5=0$ 平行，所以

$$f'(1)=\left.(3x^2+6ax+3b)\right|_{x=1}=-3,$$

即

$$3(1+2a+b)=-3. \qquad ②$$

由①，②两式联立可得 $a=-1,b=0$，故所求三次函数是 $f(x)=x^3-3x^2+c$（c 为任意值）.

令其一阶导数 $f'(x)=3x(x-2)=0$，可求出极大值 $M=c$，极小值 $m=-4+c$，于是 $M-m=4$.

1.5.10 解 本题是求驻点函数 $t(a)$ 的最小值点和最小值，所以，首先由原函数 $f(t)=a^t-at$ 求出驻点函数 $t(a)$，然后再对 $t(a)$ 进行运算.

$$f'(t)=a^t\ln a-a,$$

令 $f'(t)=a^t\ln a-a=0$，得唯一驻点函数 $t(a)=1-\frac{\ln\ln a}{\ln a}$，$t'(a)=-\frac{\frac{1}{a}-\frac{1}{a}\ln\ln a}{(\ln a)^2}$.

令 $t'(a)=-\frac{\frac{1}{a}-\frac{1}{a}\ln\ln a}{(\ln a)^2}=0$，得唯一驻点 $a=\mathrm{e}^{\mathrm{e}}$.

当 $a<\mathrm{e}^{\mathrm{e}}$ 时，$t'(a)<0$；当 $a>\mathrm{e}^{\mathrm{e}}$ 时，$t'(a)>0$，所以，$t(\mathrm{e}^{\mathrm{e}})=1-\frac{1}{\mathrm{e}}$为极小值，也是最小值.

1.5.11 解 在 $y\ln y-x+y=0$ 两边对 x 求导得 $y'\ln y+2y'-1=0$，解得

$$y'=\frac{1}{2+\ln y},$$

两边再对 x 求导得

$$y''=\frac{-\frac{y'}{y}}{(2+\ln y)^2}\xlongequal{将 y' 代入}\frac{-1}{y(2+\ln y)^3},$$

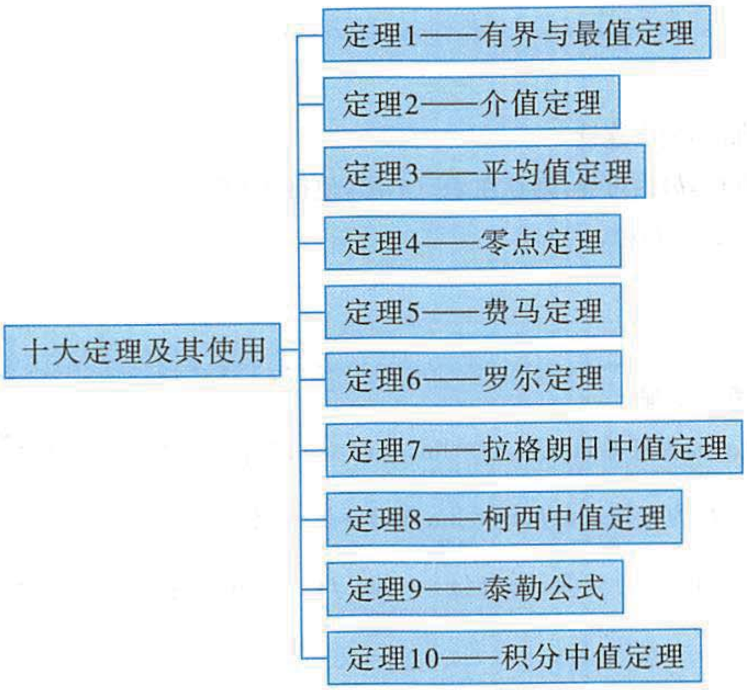
将 $x=1,y=1$ 代入得 $y''(1)=-\frac{1}{8}$.

由于二阶导数 y''在 $x=1$ 附近是连续函数，所以由 $y''(1)=-\frac{1}{8}$可知，在 $x=1$ 附近 $y''<0$，故曲线 $y=y(x)$ 的图形在点$(1,1)$附近是凸的.

1.5.12 解 由题意知 $f(-1)=0$，且 $g(-1)=0$，$f'(-1)=g'(-1)$，即 $-1-a=0$，$b+c=0$，$3+a=-2b$，解之得 $a=-1,b=-1,c=1$.

第6讲 中值定理

基础知识结构



基础内容精讲

以下列出需要考生掌握的 9 个基本定理. 其中，定理 1 至定理 4 是涉及函数的中值定理；定理 5 至定理 9 是涉及导数(微分)的中值定理；涉及积分的中值定理见例 1.6.2.

1. 涉及函数的中值定理

设 $f(x)$在$[a,b]$上连续，则

定理 1(有界与最值定理) $m\leqslant f(x)\leqslant M$，其中，$m,M$ 分别为 $f(x)$在$[a,b]$上的最小值与最大值.

定理 2(介值定理) 当 $m\leqslant\mu\leqslant M$ 时，存在 $\xi\in[a,b]$，使得 $f(\xi)=\mu$.

定理 3(平均值定理) 当 $a<x_1<x_2<\cdots<x_n<b$ 时，在$[x_1,x_n]$内至少存在一点 ξ，使

$$f(\xi)=\frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}.$$

定理 4 (零点定理) 当 $f(a)\cdot f(b)<0$ 时，存在 $\xi\in(a,b)$，使得 $f(\xi)=0$.

2. 涉及导数(微分)的中值定理

定理 5(费马定理) 设 $f(x)$满足在 x_0点处$\begin{cases}\text{①可导，}\\\text{②取极值，}\end{cases}$则 $f'(x_0)=0$.

费马

(1601—1665)

【注】 (1)要求考生掌握费马定理的证明.

不妨假设 $f(x)$ 在点 x_0 处取得极大值，则存在 x_0 的邻域 $U(x_0)$，对任意的 $x\in U(x_0)$，都有 $\Delta f=f(x)-f(x_0)\leqslant 0$，于是根据导数的定义与极限的保号性，有

$$f'_-(x_0)=\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0}\geqslant 0,$$

$$f'_+(x_0)=\lim_{x\to x_0^+}\frac{f(x)-f(x_0)}{x-x_0}\leqslant 0.$$

又 $f(x)$ 在点 x_0 处可导，于是 $f'_-(x_0)=f'_+(x_0)$，故 $f'(x_0)=0$.

(2)当一个人跑到最远处时，他的速度为零；当一个人跑得最快时，他的加速度为零. 这些都是费马定理在生活中的通俗应用.

定理 6(罗尔定理)

$$\text{设 } f(x)\text{ 满足}\begin{cases}\text{①在}[a,b]\text{上连续,}\\ \text{②在}(a,b)\text{内可导,则存在 }\xi\in(a,b)\text{,使得 }f'(\xi)=0.\\ \text{③}f(a)=f(b),\end{cases}$$


罗尔
(1652—1719)

【注】 推广的罗尔定理.

(1)设 $f(x)$ 在 (a,b) 内可导，$\lim\limits_{x\to a^+}f(x)=\lim\limits_{x\to b^-}f(x)=A$，则在 (a,b) 内至少存在一点 ξ，使 $f'(\xi)=0$.

(2)设 $f(x)$ 在 (a,b) 内可导，$\lim\limits_{x\to a^+}f(x)=\lim\limits_{x\to b^-}f(x)=\pm\infty$，则在 (a,b) 内至少存在一点 ξ，使 $f'(\xi)=0$.

(3)设 $f(x)$ 在 $(a,+\infty)$ 内可导，$\lim\limits_{x\to a^+}f(x)=\lim\limits_{x\to+\infty}f(x)=A$，则在 $(a,+\infty)$ 内至少存在一点 ξ，使 $f'(\xi)=0$.

(4)设 $f(x)$ 在 $(-\infty,+\infty)$ 内可导，$\lim\limits_{x\to-\infty}f(x)=\lim\limits_{x\to+\infty}f(x)=\pm\infty$，则在 $(-\infty,+\infty)$ 内至少存在一点 ξ，使 $f'(\xi)=0$.

定理 7(拉格朗日中值定理)

$$\text{设 } f(x)\text{ 满足}\begin{cases}\text{①在}[a,b]\text{上连续,}\\ \text{②在}(a,b)\text{内可导,}\end{cases}\text{则存在 }\xi\in(a,b)\text{,使得}$$

$$f(b)-f(a)=f'(\xi)(b-a),$$

或者写成

$$f'(\xi)=\frac{f(b)-f(a)}{b-a}.$$


拉格朗日
(1736—1813)

定理 8(柯西中值定理)

$$\text{设 } f(x),g(x)\text{ 满足}\begin{cases}\text{①在}[a,b]\text{上连续,}\\ \text{②在}(a,b)\text{内可导,则存在 }\xi\in(a,b)\text{,使得}\\ \text{③}g'(x)\neq 0,\end{cases}$$

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}.$$


柯西
(1789—1857)

定理 9（泰勒公式）

（1）带拉格朗日余项的 n 阶泰勒公式.

设 $f(x)$ 在点 x_0 的某个邻域内 $n+1$ 阶导数存在，则对该邻域内的任意点 x，有

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\cdots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1},$$

其中 ξ 介于 x，x_0 之间.

（2）带佩亚诺余项的 n 阶泰勒公式.

设 $f(x)$ 在点 x_0 处 n 阶可导，则存在 x_0 的一个邻域，对于该邻域内的任意点，有

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\cdots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+o((x-x_0)^n).$$

【注 1】 当 $x_0=0$ 时的泰勒公式称为**麦克劳林公式**.

（1）$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}$，其中 ξ 介于 0 和 x 之间.

（2）$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+o(x^n)$.

【注 2】 几个重要函数的麦克劳林展开式.

（1）$e^x=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+o(x^n)$.

（2）$\sin x=x-\frac{x^3}{3!}+\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+o(x^{2n+1})$.

（3）$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots+(-1)^n\frac{x^{2n}}{(2n)!}+o(x^{2n})$.

（4）$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+o(x^n)$.

（5）$\frac{1}{1+x}=1-x+x^2-\cdots+(-1)^nx^n+o(x^n)$.

（6）$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots+(-1)^{n-1}\frac{x^n}{n}+o(x^n)$.

（7）$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+o(x^n)$.

基础例题精解

1. 介值定理的使用

例 1.6.1 （平均值定理）设 $f(x)$ 在 $[a,b]$ 上连续，$a<x_1<x_2<\cdots<x_n<b$，证明存在 $\xi\in[x_1,x_n]$，使得

$$f(\xi)=\frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}.$$

证明 由题意可知，$f(x)$ 在 $[x_1,x_n]$ 上连续，所以 $m\leqslant f(x)\leqslant M$（$m$，$M$ 分别为 $f(x)$ 在 $[x_1,x_n]$ 上的

最小值和最大值)，于是

$$m \leqslant f(x_1) \leqslant M, \quad ①$$

$$m \leqslant f(x_2) \leqslant M, \quad ②$$

$$\cdots\cdots \quad \vdots$$

$$m \leqslant f(x_n) \leqslant M. \quad ⓝ$$

由①+②+…+ⓝ，有 $nm \leqslant f(x_1)+f(x_2)+\cdots+f(x_n) \leqslant nM$，故

$$m \leqslant \frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n} \leqslant M.$$

由介值定理可知，存在 $\xi \in [x_1, x_n]$，使得

$$f(\xi)=\frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}.$$

例 1.6.2 （**定理 10　积分中值定理**）设 $f(x)$ 在 $[a,b]$ 上连续，证明存在 $\xi \in [a,b]$，使得

$$\int_a^b f(x)\mathrm{d}x = f(\xi)(b-a).$$

证明　因为 $f(x)$ 在 $[a,b]$ 上连续，所以 $f(x)$ 在 $[a,b]$ 上存在最大值 M 与最小值 m，使得

$$m(b-a) \leqslant \int_a^b f(x)\mathrm{d}x \leqslant M(b-a),$$

故

$$m \leqslant \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \leqslant M,$$

由介值定理可知，存在 $\xi \in [a,b]$，使得 $f(\xi) = \dfrac{1}{b-a}\displaystyle\int_a^b f(x)\mathrm{d}x$，得证.

【注】　如何证明 $\xi \in (a,b)$？

设 $f(x)$ 在 $[a,b]$ 上连续，则存在 $\xi \in (a,b)$，使 $\displaystyle\int_a^b f(x)\mathrm{d}x = f(\xi)(b-a)$.

证明　令 $F(x) = \displaystyle\int_a^x f(t)\mathrm{d}t$，在 $[a,b]$ 上用拉格朗日中值定理 $\Rightarrow F(b)-F(a) = F'(\xi)(b-a)$，

即

$$\int_a^b f(x)\mathrm{d}x - 0 = f(\xi)(b-a), \xi \in (a,b),$$

得证.

考研真题中已经考过，可直接在大题中使用，不必证明再用.

例 1.6.3　设 $f(x)$ 在 $[0,1]$ 上具有一阶连续导数，$f(0)=0$，证明：存在 $\xi \in [0,1]$，使得

$$f'(\xi) = 2\int_0^1 f(x)\mathrm{d}x.$$

证明　$f'(x)$ 在 $[0,1]$ 上连续，故 $m \leqslant f'(x) \leqslant M$，其中 m, M 分别是 $f'(x)$ 在 $[0,1]$ 上的最小值和最大值.

由 $f(x)-f(0)=f'(\eta)(x-0)$ $(0<\eta<x)$，有 $f(x)=xf'(\eta)$，因 $m \leqslant f'(\eta) \leqslant M$，所以

$$mx \leqslant f'(\eta)x = f(x) \leqslant Mx.$$

因此 $$\int_0^1 mx\mathrm{d}x \leqslant \int_0^1 f(x)\mathrm{d}x \leqslant \int_0^1 Mx\mathrm{d}x, \quad \int_0^1 2mx\mathrm{d}x \leqslant 2\int_0^1 f(x)\mathrm{d}x \leqslant \int_0^1 2Mx\mathrm{d}x,$$

故 $$m = 2m \cdot \frac{1}{2} \leqslant 2\int_0^1 f(x)\mathrm{d}x \leqslant 2M \cdot \frac{1}{2} = M,$$

由介值定理可知，存在 $\xi\in[0,1]$，使得 $f'(\xi)=2\int_0^1 f(x)\mathrm{d}x$.

2. 罗尔定理的使用

(1) 常用乘积求导公式 $(uv)'=u'v+uv'$ 的逆用来制造辅助函数.

①$[f(x)f(x)]'=[f^2(x)]'=2f(x)\cdot f'(x)$.

见到 $f(x)f'(x)$，作 $F(x)=f^2(x)$.

②$[f(x)\cdot f'(x)]'=[f'(x)]^2+f(x)f''(x)$.

见到 $[f'(x)]^2+f(x)f''(x)$，作 $F(x)=f(x)f'(x)$.

③$[f(x)\mathrm{e}^{\varphi(x)}]'=f'(x)\mathrm{e}^{\varphi(x)}+f(x)\mathrm{e}^{\varphi(x)}\cdot\varphi'(x)=[f'(x)+f(x)\varphi'(x)]\mathrm{e}^{\varphi(x)}$.

见到 $f'(x)+f(x)\varphi'(x)$，作 $F(x)=f(x)\mathrm{e}^{\varphi(x)}$.

【注】 常考以下情形.

(1)$\varphi(x)=x\Rightarrow$ 见到 $f'(x)+f(x)$，作 $F(x)=f(x)\mathrm{e}^x$.

(2)$\varphi(x)=-x\Rightarrow$ 见到 $f'(x)-f(x)$，作 $F(x)=f(x)\mathrm{e}^{-x}$.

(3)$\varphi(x)=kx\Rightarrow$ 见到 $f'(x)+kf(x)$，作 $F(x)=f(x)\mathrm{e}^{kx}$.

例 1.6.4 设函数 $f(x)$ 可导，且 $f(x)f'(x)>0$，则(　　).

(A)$f(1)>f(-1)$　　(B)$f(1)<f(-1)$

(C)$|f(1)|>|f(-1)|$　　(D)$|f(1)|<|f(-1)|$

解 应选(C).

由 $f(x)f'(x)>0$ 知

$$\left[\frac{1}{2}f^2(x)\right]'=f(x)f'(x)>0,$$

则 $\frac{1}{2}f^2(x)$ 单调增加，从而 $f^2(x)$ 单调增加，由此可知

$$f^2(1)>f^2(-1),$$

两端开方得

$$|f(1)|>|f(-1)|.$$

例 1.6.5 设函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 内可导，且 $f(a)=f(b)=0$. 求证：对任意实数 α，都存在 $\xi\in(a,b)$，使得 $f'(\xi)+\alpha f(\xi)=0$.

证明 令 $F(x)=f(x)\mathrm{e}^{\alpha x}$，则 $F(x)$ 在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 内可导，且

$$F(a)=F(b)=0.$$

根据罗尔定理可知，存在 $\xi\in(a,b)$，使得 $F'(\xi)=0$，即

$$f'(\xi)\mathrm{e}^{\alpha\xi}+\alpha f(\xi)\mathrm{e}^{\alpha\xi}=0.$$

所以 $f'(\xi)+\alpha f(\xi)=0$.

例 1.6.6 设函数 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，且 $f(0)=f(1)=0$，$f\left(\frac{1}{2}\right)=1$，试证：

(1)存在 $\eta\in\left(\frac{1}{2},1\right)$，使得 $f(\eta)=\eta$；

(2)对于任意实数 λ,必存在 $\xi\in(0,\eta)$,使得 $f'(\xi)-\lambda[f(\xi)-\xi]=1$.

证明 (1)令 $F(x)=f(x)-x$,则函数 $F(x)=f(x)-x$ 在 $\left[\frac{1}{2},1\right]$ 上连续,且有

$$F\left(\frac{1}{2}\right)=f\left(\frac{1}{2}\right)-\frac{1}{2}=\frac{1}{2}>0,\quad F(1)=f(1)-1=-1<0,$$

由于 $F\left(\frac{1}{2}\right)\cdot F(1)<0$,根据零点定理可知,存在 $\eta\in\left(\frac{1}{2},1\right)$,使得 $F(\eta)=0$,即 $f(\eta)=\eta$.

(2)令 $F(x)=\mathrm{e}^{-\lambda x}[f(x)-x]$,则函数 $F(x)$ 在 $[0,\eta]$ 上连续,在 $(0,\eta)$ 内可导,$F(0)=F(\eta)=0$,即函数 $F(x)$ 在 $[0,\eta]$ 上满足罗尔定理的条件,于是存在 $\xi\in(0,\eta)$,使得

$$F'(\xi)=\mathrm{e}^{-\lambda\xi}[f'(\xi)-1]-\lambda\mathrm{e}^{-\lambda\xi}[f(\xi)-\xi]=0,$$

即

$$f'(\xi)-\lambda[f(\xi)-\xi]=1.$$

(2)上面是证明一阶导数为 0,也就是使用一次罗尔定理的问题,但有些题目涉及二阶导数为 0,即要多次使用罗尔定理,这种问题难点一般不在辅助函数的构造,而是**要找到函数值相等的三个不同点**,即 $f(a)=f(b)=f(c)$(不妨设 $a<b<c$),分别在 $[a,b]$,$[b,c]$ 上使用罗尔定理,有 $f'(\xi_1)=f'(\xi_2)=0$,$\xi_1\in(a,b)$,$\xi_2\in(b,c)$,进而在 $[\xi_1,\xi_2]$ 上再对 $f'(x)$ 使用罗尔定理,得 $f''(\xi)=0$,$\xi\in(\xi_1,\xi_2)\subset(a,c)$.

例 1.6.7 设函数 $f(x)$ 在 $[0,3]$ 上连续,在 $(0,3)$ 内有二阶导数,且

$$2f(0)=\int_0^2 f(x)\mathrm{d}x=f(2)+f(3),$$

证明:(1)存在 $\eta\in(0,2)$,使得 $f(\eta)=f(0)$;(2)存在 $\xi\in(0,3)$,使得 $f''(\xi)=0$.

证明 (1) 设 $F(x)=\int_0^x f(t)\mathrm{d}t(0\leqslant x\leqslant 2)$,则 $\int_0^2 f(x)\mathrm{d}x=F(2)-F(0)$.

根据拉格朗日中值定理可知,存在 $\eta\in(0,2)$,使 $F(2)-F(0)=2F'(\eta)=2f(\eta)$,即

$$\int_0^2 f(x)\mathrm{d}x=2f(\eta),$$

由题设 $\int_0^2 f(x)\mathrm{d}x=2f(0)$,从而 $f(\eta)=f(0)$.

(2)由于 $f(x)$ 在 $[2,3]$ 上连续,则其在 $[2,3]$ 上必有最大值 M 和最小值 m,于是

$$m\leqslant f(2)\leqslant M,\quad m\leqslant f(3)\leqslant M,$$

故

$$m\leqslant\frac{f(2)+f(3)}{2}\leqslant M.$$

根据连续函数的介值定理可知,存在 $\tau\in[2,3]$,使 $f(\tau)=\frac{f(2)+f(3)}{2}$. 由题设,$\frac{f(2)+f(3)}{2}=f(0)$,故 $f(\tau)=f(0)$,由(1)的结果可知 $f(0)=f(\eta)=f(\tau)$,且 $0<\eta<\tau\leqslant 3$. 根据罗尔定理可知,存在 $\xi_1\in(0,\eta)$,$\xi_2\in(\eta,\tau)$,使 $f'(\xi_1)=f'(\xi_2)=0$,从而存在 $\xi\in(\xi_1,\xi_2)\subset(0,3)$,使得 $f''(\xi)=0$.

3. 费马定理的使用

证明某点导数为 0 除了构造辅助函数使用罗尔定理外,**切记不可忘记还有定理 5(费马定理)**,使用费马定理只需说明**可导函数的最值在区间内部取到**,这类问题的题目往往带有不等式关系的条件(因为最值就是通过不等式关系体现的).

例 1.6.8 （导数零点定理）设 $f(x)$ 在 $[a,b]$ 上可导，证明当 $f'_+(a)\cdot f'_-(b)<0$ 时，存在 $\xi\in(a,b)$，使得 $f'(\xi)=0$.

证明 不妨设 $f'_+(a)>0, f'_-(b)<0$，于是

$$f'_+(a)=\lim_{x\to a^+}\frac{f(x)-f(a)}{x-a}>0\Rightarrow 存在\ \xi_1>0，在\ (a,a+\xi_1)\ 内，f(x)>f(a)，$$

$$f'_-(b)=\lim_{x\to b^-}\frac{f(x)-f(b)}{x-b}<0\Rightarrow 存在\ \xi_2>0，在\ (b-\xi_2,b)\ 内，f(x)>f(b).$$

故 $f(a)$ 与 $f(b)$ 均不是 $f(x)$ 在 $[a,b]$ 上的最大值，则 $f(x)$ 在 (a,b) 内取得最大值，根据费马定理可知，存在 $\xi\in(a,b)$，使得 $f'(\xi)=0$.

4. 拉格朗日中值定理的使用

例 1.6.9 设 $a>b>0, n>1$，证明：$nb^{n-1}(a-b)<a^n-b^n<na^{n-1}(a-b)$.

证明 设 $f(x)=x^n$，显然 $f(x)$ 在区间 $[b,a]$ 上连续，在 (b,a) 内可导，即 $f(x)$ 在区间 $[b,a]$ 上满足拉格朗日中值定理的条件，于是得

$$f(a)-f(b)=f'(\xi)(a-b), \xi\in(b,a),$$

由于 $f'(x)=nx^{n-1}$，因此上式即为

$$a^n-b^n=n\xi^{n-1}(a-b),$$

又由 $0<b<\xi<a, n>1$，有

$$nb^{n-1}(a-b)<a^n-b^n=n\xi^{n-1}(a-b)<na^{n-1}(a-b).$$

例 1.6.10 设 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，证明：至少存在一点 $\xi\in(0,1)$，使得 $f(1)=3\xi^2f(\xi)+\xi^3f'(\xi)$.

证明 注意到 $3x^2f(x)+x^3f'(x)$ 是 $x^3f(x)$ 的导函数，故令 $F(x)=x^3f(x)$，易知 $F(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，即 $F(x)$ 在 $[0,1]$ 上满足拉格朗日中值定理的条件，于是得

$$F(1)-F(0)=F'(\xi), \xi\in(0,1),$$

即

$$f(1)=3\xi^2f(\xi)+\xi^3f'(\xi).$$

例 1.6.11 设函数 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，且 $f(0)=0, f(1)=1$，证明存在不同的 $\xi_1,\xi_2\in(0,1)$，使得 $\frac{1}{f'(\xi_1)}+\frac{1}{f'(\xi_2)}=2$.

证明 用 ξ 将 $[0,1]$ 划分为 $[0,\xi]$，$[\xi,1]$. 在这两个区间上分别对 $f(x)$ 使用拉格朗日中值定理，得

$$f(\xi)-f(0)=f'(\xi_1)(\xi-0)\Rightarrow\frac{1}{f'(\xi_1)}=\frac{\xi}{f(\xi)}, \xi_1\in(0,\xi),$$

$$f(1)-f(\xi)=f'(\xi_2)(1-\xi)\Rightarrow\frac{1}{f'(\xi_2)}=\frac{1-\xi}{1-f(\xi)}, \xi_2\in(\xi,1),$$

与欲证等式比较，只需证 $\frac{\xi}{f(\xi)}+\frac{1-\xi}{1-f(\xi)}=2$ 即可，于是可取 $f(\xi)=\frac{1}{2}$，则

$$\frac{\xi}{f(\xi)}+\frac{1-\xi}{1-f(\xi)}=2(\xi+1-\xi)=2,$$

命题得证.

【注】 这种反推思想值得借鉴.

5. 柯西中值定理的使用

例 1.6.12 设 $f(x)$在$[a,b]$上连续，在(a,b)内可导，$0<a<b$，证明：至少存在一点 $\xi\in(a,b)$，使得 $f(b)-f(a)=\xi\ln\frac{b}{a}f'(\xi)$.

证明 因为 $f(x)$与 $g(x)=\ln x$ 在$[a,b]$上连续，在(a,b)内可导，且 $g'(x)=\frac{1}{x}\neq0(0<a<x<b)$，即 $f(x)$与$g(x)=\ln x$ 在$[a,b]$上满足柯西中值定理的条件，则至少存在一点 $\xi\in(a,b)$，使得

$$\frac{f(b)-f(a)}{\ln b-\ln a}=\frac{f'(\xi)}{\frac{1}{\xi}},$$

即

$$f(b)-f(a)=\xi\ln\frac{b}{a}f'(\xi).$$

6. 泰勒公式的使用

例 1.6.13 设 $f(x)$在区间$[-a,a](a>0)$上具有二阶连续导数，$f(0)=0$.

(1)写出 $f(x)$的带拉格朗日余项的一阶麦克劳林公式；

(2) 证明：存在 $\eta\in[-a,a]$，使得 $a^3f''(\eta)=3\int_{-a}^{a}f(x)\mathrm{d}x$.

(1)**解** 对任意的 $x\in[-a,a]$，$f(x)=f(0)+f'(0)x+\frac{1}{2}f''(\xi)x^2=f'(0)x+\frac{f''(\xi)}{2}x^2$，$\xi$ 介于 x 与 0 之间.

(2) **证明** $\int_{-a}^{a}f(x)\mathrm{d}x=\int_{-a}^{a}f'(0)x\mathrm{d}x+\frac{1}{2}\int_{-a}^{a}f''(\xi)x^2\mathrm{d}x=\frac{1}{2}\int_{-a}^{a}f''(\xi)x^2\mathrm{d}x.$

因为 $f''(x)$在$[-a,a]$上连续，由最值定理：$m\leqslant f''(x)\leqslant M, x\in[-a,a]$，其中 m,M 分别是 $f''(x)$在$[-a,a]$上的最小值和最大值，有

$$mx^2\leqslant f''(\xi)x^2\leqslant Mx^2,$$

$$\frac{2}{3}ma^3=m\int_{-a}^{a}x^2\mathrm{d}x\leqslant\int_{-a}^{a}f''(\xi)x^2\mathrm{d}x\leqslant M\int_{-a}^{a}x^2\mathrm{d}x=\frac{2}{3}Ma^3,$$

$$\frac{a^3}{3}\cdot m\leqslant\frac{1}{2}\int_{-a}^{a}f''(\xi)x^2\mathrm{d}x=\int_{-a}^{a}f(x)\mathrm{d}x\leqslant\frac{a^3}{3}\cdot M,$$

$$m\leqslant\frac{3}{a^3}\int_{-a}^{a}f(x)\mathrm{d}x\leqslant M,$$

由介值定理知，存在 $\eta\in[-a,a]$，使得 $f''(\eta)=\frac{3}{a^3}\int_{-a}^{a}f(x)\mathrm{d}x$，得证.

基础习题精练

习 题

1.6.1 设在$[0,1]$上 $f''(x)>0$，则 $f'(0)$，$f'(1)$，$f(1)-f(0)$的大小顺序为(　　).

(A)$f'(1)>f'(0)>f(1)-f(0)$　　(B)$f'(1)>f(1)-f(0)>f'(0)$

(C) $f(1)-f(0)>f'(1)>f'(0)$　　　　(D) $f(1)-f(0)>f'(0)>f'(1)$

1.6.2 设 $f(x)$ 在 $[0,3]$ 上连续，在 $(0,3)$ 内可导，且 $f(0)+f(1)+f(2)=3$，$f(3)=1$. 试证必存在 $\xi\in(0,3)$，使得 $f'(\xi)=0$.

1.6.3 设 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，且 $f(1)=k\int_0^{\frac{1}{k}} xe^{1-x}f(x)\mathrm{d}x\,(k>1)$. 证明至少存在一点 $\xi\in(0,1)$，使得 $f'(\xi)=(1-\xi^{-1})f(\xi)$.

1.6.4 设函数 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内二阶可导，过点 $A(0,f(0))$ 与 $B(1,f(1))$ 的直线与曲线 $y=f(x)$ 相交于点 $C(c,f(c))$，其中 $0<c<1$，证明存在 $\xi\in(0,1)$，使得 $f''(\xi)=0$.

1.6.5 （导数介值定理）设 $f(x)$ 在 $[a,b]$ 上可导，若 $f'_+(a)\neq f'_-(b)$，证明对于任意的介于 $f'_+(a)$ 与 $f'_-(b)$ 之间的 μ，存在 $\xi\in(a,b)$，使得 $f'(\xi)=\mu$.

1.6.6 已知 $f(x)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，且 $f(0)=0$，$f(1)=1$. 证明：

(1) 存在 $\xi\in(0,1)$，使得 $f(\xi)=1-\xi$；

(2) 存在 $\eta,\tau\in(0,1)$，$\eta\neq\tau$，使得 $f'(\eta)f'(\tau)=1$.

解 答

1.6.1 (B) **解** 由 $f''(x)>0$ 可知 $f'(x)$ 在 $[0,1]$ 上单调增加. 又根据拉格朗日中值定理可知，$f(1)-f(0)=f'(\xi)(0<\xi<1)$，从而有 $f'(0)<f'(\xi)<f'(1)$，即 $f'(1)>f(1)-f(0)>f'(0)$. 故选(B).

1.6.2 **分析** 根据题设条件，只需再补充一个条件：存在一点 $c\in[0,3)$，使得 $f(c)=1$. 则函数 $f(x)$ 就在 $[c,3]$ 上满足罗尔定理.

证明 因为 $f(x)$ 在 $[0,3]$ 上连续，所以 $f(x)$ 在 $[0,2]$ 上连续，且在 $[0,2]$ 上必有最大值 M 和最小值 m，于是

$$m\leqslant f(0)\leqslant M,\quad m\leqslant f(1)\leqslant M,\quad m\leqslant f(2)\leqslant M,$$

故

$$m\leqslant\frac{f(0)+f(1)+f(2)}{3}=1\leqslant M.$$

由介值定理可知，至少存在一点 $c\in[0,2]$，使得 $f(c)=\frac{f(0)+f(1)+f(2)}{3}=1$.

所以函数 $f(x)$ 在 $[c,3]$ 上满足罗尔定理，于是存在 $\xi\in(c,3)\subset(0,3)$，使得 $f'(\xi)=0$.

1.6.3 **分析** 注意到被积函数 $xe^{1-x}f(x)$ 的导数

$$[xe^{1-x}f(x)]'=xe^{1-x}[f'(x)-(1-x^{-1})f(x)]$$

中含有 $f'(x)-(1-x^{-1})f(x)$，这正是欲证结论中的一部分.

证明 令 $F(x)=xe^{1-x}f(x)$. 由积分中值定理，可知

$$f(1)=k\int_0^{\frac{1}{k}} xe^{1-x}f(x)\mathrm{d}x=\eta e^{1-\eta}f(\eta)=1\cdot e^{1-1}f(1),\eta\in\left(0,\frac{1}{k}\right)\subset(0,1),$$

又 $F(x)=xe^{1-x}f(x)$ 在 $[\eta,1]$ 上连续，在 $(\eta,1)$ 内可导，且 $F(\eta)=F(1)$，则 $F(x)=xe^{1-x}f(x)$ 在 $[\eta,1]$ 上满足罗尔定理的条件，于是存在 $\xi\in(\eta,1)\subset(0,1)$，使得 $F'(\xi)=0$，即 $f'(\xi)=(1-\xi^{-1})f(\xi)$.

1.6.4 **证明** 按题意，点 A 与 B 的连线的方程为 $y_1=[f(1)-f(0)]x+f(0)$.

令 $F(x)=f(x)-y_1=f(x)-[f(1)-f(0)]x-f(0)$，则 $F(x)$ 在 $[0,c]$ 与 $[c,1]$ 上均满足罗尔定理的条件，于是可得 $F'(\xi_1)=F'(\xi_2)=0$，$\xi_1\in(0,c)$，$\xi_2\in(c,1)$.

再由罗尔定理可得 $F''(\xi)=f''(\xi)=0$，$\xi\in(\xi_1,\xi_2)\subset(0,1)$，命题得证.

1.6.5 证明 因 $f'_{+}(a)\neq f'_{-}(b)$，不妨设 $f'_{+}(a)<f'_{-}(b)$. 并设 $F(x)=f(x)-\mu x$，则函数 $F(x)$ 在 $[a,b]$ 上可导，且 $F'_{+}(a)=f'_{+}(a)-\mu<0$，$F'_{-}(b)=f'_{-}(b)-\mu>0$，于是，

$$\begin{cases} F'_{+}(a)=\lim\limits_{x\to a^{+}}\dfrac{F(x)-F(a)}{x-a}<0, \\ F'_{-}(b)=\lim\limits_{x\to b^{-}}\dfrac{F(x)-F(b)}{x-b}>0, \end{cases}$$

根据极限的保号性，知：

在点 $x=a$ 的某个右邻域内，$\dfrac{F(x)-F(a)}{x-a}<0$，即 $F(x)<F(a)$；

在点 $x=b$ 的某个左邻域内，$\dfrac{F(x)-F(b)}{x-b}>0$，即 $F(x)<F(b)$.

故 $F(a)$ 和 $F(b)$ 均不是函数 $F(x)$ 在 $[a,b]$ 上的最小值，又因 $F(x)$ 一定可以取得最小值，则其最小值必在 (a,b) 内取到，设函数 $F(x)$ 在 (a,b) 内的最小值点是 ξ，根据费马定理，得 $F'(\xi)=0$，即 $f'(\xi)=\mu$.

1.6.6 证明 (1)令 $F(x)=f(x)-1+x$，可得 $\begin{cases} F(0)=f(0)-1+0=-1<0, \\ F(1)=f(1)-1+1=1>0, \end{cases}$ 由零点定理可知，存在 $\xi\in(0,1)$，使得 $F(\xi)=0$，即 $f(\xi)=1-\xi$.

(2)用 ξ 将 $[0,1]$ 划分为 $[0,\xi]$，$[\xi,1]$，再用拉格朗日中值定理有

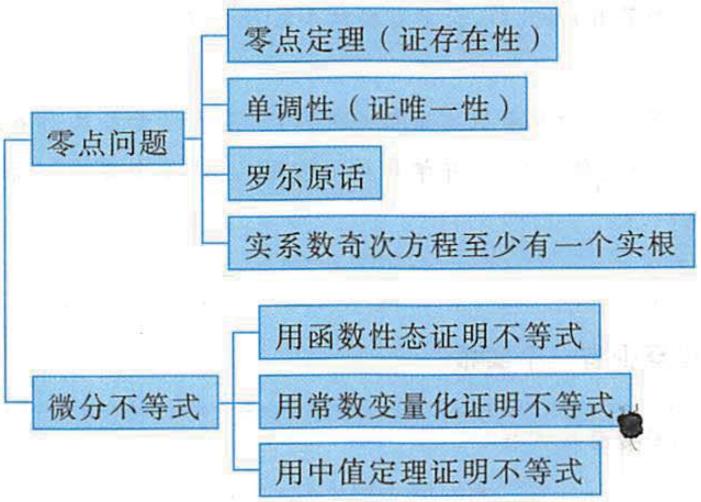
$$f(\xi)-f(0)=f'(\eta)(\xi-0),\eta\in(0,\xi),$$
$$f(1)-f(\xi)=f'(\tau)(1-\xi),\tau\in(\xi,1),$$

则
$$f'(\eta)=\frac{f(\xi)-f(0)}{\xi}=\frac{1-\xi}{\xi},\quad f'(\tau)=\frac{f(1)-f(\xi)}{1-\xi}=\frac{\xi}{1-\xi},$$

故 $f'(\eta)f'(\tau)=1$.

第7讲 零点问题与微分不等式

基础知识结构



基础内容精讲

一、零点问题

方程 $f(x)=0$ 的根就是函数 $f(x)$ 的零点. 从几何上讲，方程的根作为两条曲线的交点，代数语言“$f(x)=g(x)$的根”与几何语言“曲线 $f(x)$与 $g(x)$的交点”，两者概念不同，但描述的是同一件事. 基于此，为讨论方程的根，有时可改为讨论曲线的交点. 讨论方程根的问题（也称为函数的零点问题）通常可以考虑下面这些方法.

1. 零点定理（主要用于证明根的存在性）

设 $f(x)$在$[a,b]$上连续，且 $f(a)f(b)<0$，则 $f(x)=0$ 在(a,b)内至少有一个根.

【注】 推广的零点定理：若 $f(x)$在(a,b)内连续，$\lim\limits_{x\to a^+}f(x)=\alpha$，$\lim\limits_{x\to b^-}f(x)=\beta$，且 $\alpha\cdot\beta<0$，则 $f(x)=0$在(a,b)内至少有一个根，这里 a,b,α,β 可以是有限数，也可以是无穷大.

2. 单调性（主要用于证明根的唯一性）

若 $f(x)$在(a,b)内单调，则 $f(x)=0$ 在(a,b)内至多有一个根，这里 a,b 可以是有限数，也可以是无穷大.

3. 罗尔原话（罗尔定理的推论）

若 $f^{(n)}(x)=0$ 至多有 k 个根，则 $f(x)=0$ 至多有 $k+n$ 个根.

【注】 读者可能有所不知，罗尔是研究方程的权威，可是他并不研究微积分. 由于牛顿在表达无穷小量时，出现过很多错误，罗尔还公开指责牛顿的微积分是谬论，故在微积分教材中罗尔定理并非是罗尔所创造，而是后人为了纪念罗尔，以罗尔的名字命名的. 那为什么微积分的研究者会用一位反对微积分的人的名字来命名这么有名的定理呢？原来，罗尔曾经讲过一段话：若 $f(x)=0$ 有两个根，则 $f'(x)=0$ 至少有一个根(当然，这个符号和概念是我们写的，罗尔当年只是用具体函数来表达 $f'(x)$). 于是可得其逆否命题：

“若 $f'(x)=0$ 无实根，则 $f(x)=0$ 至多有一个根”.

读者可根据现在的罗尔定理，利用归纳法得出：“若 $f^{(n-1)}(x)=0$ 至多有一个根，则 $f^{(n-2)}(x)=0$ 至多有 2 个根.”依次类推，$f^{(n-3)}(x)=0$ 至多有 3 个根，…，$f(x)=0$ 至多有 n 个根. 读者可以看出，方程阶数(即 $f(x)$ 的导数阶数)降几阶，至多有根的个数就会多几个，于是便可得到上述结论. 此结论可在考试时直接使用.

以上所述与现在的罗尔定理有着极为密切的关系，故后人以昔日反对者的名字命名了著名的罗尔定理，读者应将这段描述熟记于心，并学会灵活使用. 且让作者多一句嘴：真正的科学，一定是对事不对人的，你说呢？

怎么用？见例 1.7.2.

4. 实系数奇次方程至少有一个实根

【注】 试证任何实系数奇次方程

$$x^{2n+1}+a_1x^{2n}+\cdots+a_{2n}x+a_{2n+1}=0$$

至少有一个实根.

证明 设

$$f(x)=x^{2n+1}+a_1x^{2n}+\cdots+a_{2n}x+a_{2n+1},$$

则 $\lim\limits_{x\to+\infty}f(x)=+\infty$，故 $\forall M_1>0$，$\exists X_1>0$，当 $x>X_1$ 时，$f(x)>M_1>0$. 故 $\exists x_1>X_1$，使 $f(x_1)>0$；

又 $\lim\limits_{x\to-\infty}f(x)=-\infty$，故 $\forall M_2>0$，$\exists X_2>0$，当 $x<-X_2$ 时，$f(x)<-M_2<0$，故 $\exists x_2<-X_2$，使 $f(x_2)<0$.

由 $f(x)$ 的连续性及零点定理，知 $\exists\xi\in(x_2,x_1)\subset(-\infty,+\infty)$，使 $f(\xi)=0$，即至少有一个实根.

怎么用？见例 1.7.3.

二、微分不等式

1. 用函数性态(包括单调性、凹凸性和最值等)证明不等式

一般地，使用如下依据.

(1)若有 $f'(x)\geqslant0$，$a<x<b$，则有 $f(a)\leqslant f(x)\leqslant f(b)$.

(2)若有 $f''(x)\geqslant0$，$a<x<b$，则有 $f'(a)\leqslant f'(x)\leqslant f'(b)$.

① 当 $f'(a)>0$ 时，$f'(x)>0\Rightarrow f(x)$ 单调增加；

② 当 $f'(b)<0$ 时，$f'(x)<0 \Rightarrow f(x)$ 单调减少.

(3)设 $f(x)$ 在 I 内有唯一的极值点 x_0，则 $\begin{cases} \text{当 } x_0 \text{ 为极大值点时}, f(x_0) \geqslant f(x), \\ \text{当 } x_0 \text{ 为极小值点时}, f(x_0) \leqslant f(x), \end{cases} \forall x \in I.$

(4)若有 $f''(x)>0$，$a<x<b$，$f(a)=f(b)=0$，则有 $f(x)<0$.

2. 用常数变量化证明不等式

如果欲证的不等式中都是常数，则可以将其中一个或者几个常数变量化，再利用上面所述的导数工具去证明.

3. 用中值定理证明不等式

主要用拉格朗日中值定理或者泰勒公式.

具体请看后面的例题.

基础例题精解

一、方程根的问题（函数的零点问题）

例 1.7.1　求证方程 $x+p+q\cos x=0$ 恰有一个实根，其中 p,q 为常数，且 $0<q<1$.

证明　令 $f(x)=x+p+q\cos x$，由 $\lim\limits_{x\to+\infty} f(x)=+\infty$，知存在一点 b，使得 $f(b)>0$.

由 $\lim\limits_{x\to-\infty} f(x)=-\infty$，知存在一点 a，使 $f(a)<0$. 故由零点定理可知，在 (a,b) 内至少存在一点 c，使得 $f(c)=0$，即方程 $f(x)=0$ 在 (a,b) 内至少有一实根.

又因为 $f'(x)=1-q\sin x>0$，故 $f(x)$ 在 $(-\infty,+\infty)$ 内单调增加，所以 $f(x)=0$ 在 $(-\infty,+\infty)$ 内至多有一个实根.

综上所述，方程 $x+p+q\cos x=0$ 恰有一个实根.

例 1.7.2　证明方程 $2^x-x^2=1$ 有且仅有 3 个实根.

证明　令 $f(x)=2^x-x^2-1$，则 $f(x)$ 在 $(-\infty,+\infty)$ 内连续，由于 $f'''(x)=2^x(\ln 2)^3\neq 0$，即 $f'''(x)=0$ 无实根（至多 0 个实根），则由罗尔原话（罗尔定理的推论），$f(x)=0$ 至多有 3 个实根. 又

$$f(0)=0,\quad f(1)=0,\quad f(2)=-1<0,\quad f(5)=6>0,$$

故由零点定理，知 $f(x)=0$ 至少有 3 个实根. 综上所述，方程 $2^x-x^2=1$ 有且仅有 3 个实根.

例 1.7.3　若 $3a^2-5b<0$，则方程 $x^5+2ax^3+3bx+4c=0$（　　）.

(A) 无实根　　(B) 有唯一实根　　(C) 有三个不同实根　　(D) 有五个不同实根

解　应选(B).

由于 $f(x)=x^5+2ax^3+3bx+4c$ 是奇次的，故方程 $f(x)=0$ 至少有一实根，又 $f'(x)=0$，即 $5x^4+6ax^2+3b=0$ 的判别式

$$\Delta=12(3a^2-5b)<0,$$

则方程 $f'(x)=0$ 无实根，所以方程 $f(x)=0$ 只有唯一实根. 故答案选(B).

例 1.7.4　设常数 $k>0$，函数 $f(x)=\ln x-\dfrac{x}{\mathrm{e}}+k$ 在 $(0,+\infty)$ 内的零点个数为（　　）.

(A)3　　(B)2　　(C)1　　(D)0

解　应选(B).

易知 $f'(x)=\frac{1}{x}-\frac{1}{e}$. 当 $x>e$ 时，$f'(x)<0$；当 $0<x<e$ 时，$f'(x)>0$，故在 $(0,e)$ 和 $(e,+\infty)$ 内 $f(x)$ 分别至多有一个零点，又 $f(e)=k>0$，$\lim\limits_{x\to0^+}f(x)=-\infty$，$\lim\limits_{x\to+\infty}f(x)=-\infty$，所以 $f(x)$ 在 $(0,+\infty)$ 内有两个零点. 故答案选(B).

例 1.7.5 在区间 $(-\infty,+\infty)$ 内，方程 $|x|^{\frac{1}{4}}+|x|^{\frac{1}{2}}-\cos x=0$（ ）.

(A)无实根 (B)有且仅有一个实根

(C)有且仅有两个实根 (D)有无穷多个实根

解 应选(C).

记 $f(x)=|x|^{\frac{1}{4}}+|x|^{\frac{1}{2}}-\cos x$，易知 $f(x)$ 为偶函数，又因为当 $x>1$ 时，$f(x)>0$，故只需判断在 $(0,1)$ 内 $f(x)=0$ 有无实根即可. 因为 $f(0)=-1<0$，$f(1)=2-\cos 1>0$，且当 $x\in(0,1)$ 时，$f'(x)=\frac{1}{4}x^{-\frac{3}{4}}+\frac{1}{2}x^{-\frac{1}{2}}+\sin x>0$，所以在 $(0,+\infty)$ 内 $f(x)=0$ 有且仅有一个实根. 故在 $(-\infty,+\infty)$ 内有且仅有两个实根. 故答案选(C).

例 1.7.6 讨论曲线 $y=x\ln x$ 与直线 $y=-A$ 的交点个数.

解 令 $f(x)=x\ln x+A$，讨论曲线 $y=x\ln x$ 与直线 $y=-A$ 的交点个数等价于讨论方程 $f(x)=0$ 的不同实根的个数. 因 $f'(x)=\ln x+1$，令 $f'(x)=0$，得 $x=\frac{1}{e}$.

x	$\left(0,\frac{1}{e}\right)$	$\frac{1}{e}$	$\left(\frac{1}{e},+\infty\right)$
$f'(x)$	$-$	0	$+$
$f(x)$	↘	极小值 $f\left(\frac{1}{e}\right)=A-\frac{1}{e}$	↗

$$\lim_{x\to0^+}f(x)=A,\quad \lim_{x\to+\infty}f(x)=+\infty.$$

若 $A\leqslant0$，极小值 $f\left(\frac{1}{e}\right)<0$，则在 $\left(0,\frac{1}{e}\right)$ 内方程无实根，在 $\left(\frac{1}{e},+\infty\right)$ 内方程有一实根；

若 $0<A<\frac{1}{e}$，极小值 $f\left(\frac{1}{e}\right)<0$，则方程在 $\left(0,\frac{1}{e}\right)$，$\left(\frac{1}{e},+\infty\right)$ 内各有一实根；

若 $A=\frac{1}{e}$，则极小值 $f\left(\frac{1}{e}\right)=0$，在 $\left(0,\frac{1}{e}\right)$ 内，$f(x)>0$，在 $\left(\frac{1}{e},+\infty\right)$ 内，$f(x)>0$，即方程在 $(0,+\infty)$ 内有且仅有一个实根；

若 $A>\frac{1}{e}$，则极小值 $f\left(\frac{1}{e}\right)>0$，从而在 $(0,+\infty)$ 内，$f(x)>0$，方程没有实根.

例 1.7.7 求方程 $k\arctan x-x=0$ 的不同实根的个数，其中 k 为参数.

分析 本题主要考查利用函数单调性、函数极值及函数的零点定理与渐近性态，讨论方程实根的存在性问题，由于方程中含有参数 k，故要根据其不同取值范围进行讨论，是一道综合题.

解 令 $f(x)=k\arctan x-x$，则 $f(x)$ 是 $(-\infty,+\infty)$ 内的奇函数，且 $f(0)=0$，$f'(x)=\frac{k-1-x^2}{1+x^2}$.

当 $k-1\leqslant0$，即 $k\leqslant1$ 时，$f'(x)<0(x\neq0)$，$f(x)$ 在区间 $(-\infty,+\infty)$ 内单调减少，方程 $f(x)=0$ 只有一个实根 $x=0$；

当 $k-1>0$，即 $k>1$ 时，在区间 $(0,\sqrt{k-1})$ 内，$f'(x)>0$，$f(x)$ 单调增加，在区间 $(\sqrt{k-1},+\infty)$ 内，$f'(x)<0$，$f(x)$ 单调减少，所以 $f(\sqrt{k-1})$ 是 $f(x)$ 在 $(0,+\infty)$ 内的最大值，从而 $f(\sqrt{k-1})>f(0)=0$.

又因为 $\lim\limits_{x\to+\infty}f(x)=-\infty$，所以由函数的零点定理知，存在 $\xi\in(\sqrt{k-1},+\infty)$，使得 $f(\xi)=0$.

由 $f(x)$ 是奇函数及其单调性可知：当 $k>1$ 时，方程 $f(x)=0$ 有且仅有 3 个不同的实根

$$x=-\xi,\quad x=0,\quad x=\xi.$$

【注】 上面的例 1.7.6 与例 1.7.7 都是含参数的零点问题，一种是导数中不含参数，其特点是在结果中讨论参数（曲线与 x 轴的位置关系）；一种是导数中含有参数，其特点是在求导过程中讨论参数（确定函数性态）.

二、微分不等式

例 1.7.8 证明当 $x>0$ 时，$\ln\left(1+\frac{1}{x}\right)<\frac{1}{\sqrt{x(x+1)}}$.

证明 令 $t=\frac{1}{x}$，则 $x>0$ 时，

$$\ln\left(1+\frac{1}{x}\right)<\frac{1}{\sqrt{x(x+1)}}\Leftrightarrow\ln(1+t)<\frac{t}{\sqrt{1+t}}\Leftrightarrow t-\sqrt{1+t}\ln(1+t)>0(t>0).$$

令 $F(t)=t-\sqrt{1+t}\ln(1+t)$，则 $F'(t)=\frac{2\sqrt{1+t}-\ln(1+t)-2}{2\sqrt{1+t}}$，无法直接判断正负.

再令 $G(t)=2\sqrt{1+t}-\ln(1+t)-2$，则 $G'(t)=\frac{\sqrt{1+t}-1}{1+t}>0\Rightarrow G(t)>G(0)=0$，故 $F'(t)=\frac{G(t)}{2\sqrt{1+t}}>0$，因此 $F(t)>F(0)=0$，得证.

例 1.7.9 证明当 $0<x<\frac{\pi}{2}$ 时，有 $\sin x>\frac{2x}{\pi}$.

证明 设 $f(x)=\sin x-\frac{2x}{\pi}$，则

$$f'(x)=\cos x-\frac{2}{\pi},\quad f''(x)=-\sin x<0,x\in\left(0,\frac{\pi}{2}\right),$$

所以曲线 $f(x)=\sin x-\frac{2x}{\pi}$ 在 $\left(0,\frac{\pi}{2}\right)$ 内是凸的，又 $f(0)=f\left(\frac{\pi}{2}\right)=0$，所以 $f(x)=\sin x-\frac{2x}{\pi}>0$，即

$$\sin x>\frac{2x}{\pi}.$$

例 1.7.10 证明 $1+x\ln(x+\sqrt{1+x^2})\geqslant\sqrt{1+x^2},-\infty<x<+\infty$.

证明 设 $f(x)=1+x\ln(x+\sqrt{1+x^2})-\sqrt{1+x^2}$，则

$$f'(x)=\ln(x+\sqrt{1+x^2})+x\frac{1+\frac{x}{\sqrt{1+x^2}}}{x+\sqrt{1+x^2}}-\frac{x}{\sqrt{1+x^2}}=\ln(x+\sqrt{1+x^2}).$$

令 $f'(x)=0$ 得驻点为 $x=0$，由于 $f''(x)=\frac{1}{\sqrt{1+x^2}}>0$，知 $x=0$ 为唯一极小值点，即最小值点. $f(x)$ 的最小值为 $f(0)=0$，于是，对一切 $x\in(-\infty,+\infty)$，有 $f(x)\geqslant0$，即有

$$1+x\ln(x+\sqrt{1+x^2})\geqslant\sqrt{1+x^2},-\infty<x<+\infty.$$

例 1.7.11 设 $0<a<b$，证明：$\ln\frac{b}{a}>2\frac{b-a}{a+b}$.

分析 我们可以想到三种思路：

①$b\to x\Rightarrow \ln\dfrac{x}{a}>2\dfrac{x-a}{a+x}(x>a>0)$；

②$a\to x\Rightarrow \ln\dfrac{b}{x}>2\dfrac{b-x}{x+b}(0<x<b)$；

③先化成“齐次式”：$\ln\dfrac{b}{a}>2\dfrac{\dfrac{b}{a}-1}{1+\dfrac{b}{a}}$，$\dfrac{b}{a}\to x\Rightarrow \ln x>2\dfrac{x-1}{1+x}$，其中 $x>1$.

当然第三种思路最为简单，$\ln x>2\dfrac{x-1}{1+x}\Leftrightarrow(1+x)\ln x-2(x-1)>0(x>1)$.

证明 令 $F(x)=(1+x)\ln x-2(x-1)(x>1)$，则

$$F'(x)=\frac{1}{x}+\ln x-1,\quad F''(x)=-\frac{1}{x^2}+\frac{1}{x}=\frac{1}{x}\left(1-\frac{1}{x}\right)>0.$$

因此，$F'(x)>F'(1)=0$，故 $F(x)>F(1)=0$. 再令 $x=\dfrac{b}{a}(0<a<b)$，即得证.

例 1.7.12 设 $0<a<b<1$，证明不等式 $\arctan b-\arctan a<\dfrac{b-a}{2ab}$.

证明 只需证明$\dfrac{\arctan b-\arctan a}{b-a}<\dfrac{1}{2ab}$，在$[a,b]$上对 $f(x)=\arctan x$ 应用拉格朗日中值定理，

$$\frac{\arctan b-\arctan a}{b-a}=\frac{1}{1+\xi^2}<\frac{1}{1+a^2}<\frac{1}{a^2+b^2}<\frac{1}{2ab}(0<a<\xi<b<1).$$

例 1.7.13 设 $f(x)$在闭区间$[0,c]$上连续，其导数 $f'(x)$在开区间$(0,c)$内存在且单调减少，又 $f(0)=0$，试应用拉格朗日中值定理证明不等式 $f(a+b)\leqslant f(a)+f(b)$，其中常数 a,b 满足条件 $0\leqslant a\leqslant b\leqslant a+b\leqslant c$.

证明 当 $a=0$ 时，由 $f(0)=0$ 知 $f(a+b)=f(b)=f(a)+f(b)$；

当 $a>0$ 时，在$[0,a]$和$[b,a+b]$上分别应用拉格朗日中值定理，有

$$f'(\xi_1)=\frac{f(a)-f(0)}{a-0}=\frac{f(a)}{a},\xi_1\in(0,a),$$

$$f'(\xi_2)=\frac{f(a+b)-f(b)}{(a+b)-b}=\frac{f(a+b)-f(b)}{a},\xi_2\in(b,a+b).$$

显然 $0<\xi_1<a\leqslant b<\xi_2<a+b\leqslant c$，因 $f'(x)$在$(0,c)$内单调减少，故 $f'(\xi_2)\leqslant f'(\xi_1)$，从而有$\dfrac{f(a+b)-f(b)}{a}\leqslant\dfrac{f(a)}{a}$，因为 $a>0$，所以有 $f(a+b)\leqslant f(a)+f(b)$.

例 1.7.14 设 $f(x)$在(a,b)内二阶可导，且 $f''(x)>0$，证明对于任意的 $x_1,x_2\in(a,b)$，且 $x_1\neq x_2$ 及$\lambda(0<\lambda<1)$，恒有 $f[\lambda x_1+(1-\lambda)x_2]<\lambda f(x_1)+(1-\lambda)f(x_2)$.

证明 **方法一** 记 $x=\lambda x_1+(1-\lambda)x_2$，由 $a<x_1<b,a<x_2<b,0<\lambda<1$ 可知 $a<x<b$.

由泰勒公式

$$\begin{aligned}f(x_1)&=f(x)+f'(x)(x_1-x)+\frac{f''(\xi_1)}{2!}(x_1-x)^2\\&=f(x)+f'(x)(1-\lambda)(x_1-x_2)+\frac{f''(\xi_1)}{2!}(x_1-x)^2,\xi_1\text{ 介于 }x\text{ 与 }x_1\text{ 之间},\\f(x_2)&=f(x)+f'(x)(x_2-x)+\frac{f''(\xi_2)}{2!}(x_2-x)^2\\&=f(x)+f'(x)\lambda(x_2-x_1)+\frac{f''(\xi_2)}{2!}(x_2-x)^2,\xi_2\text{ 介于 }x\text{ 与 }x_2\text{ 之间},\end{aligned}$$

于是 $$\lambda f(x_1)+(1-\lambda)f(x_2)=f(x)+\lambda\frac{f''(\xi_1)}{2}(x_1-x)^2+(1-\lambda)\frac{f''(\xi_2)}{2}(x_2-x)^2>f(x),$$

其中 $\lambda>0,1-\lambda>0,f''(\xi_1)>0,f''(\xi_2)>0$.

所以有 $$\lambda f(x_1)+(1-\lambda)f(x_2)>f[\lambda x_1+(1-\lambda)x_2].$$

方法二 记 $x=\lambda x_1+(1-\lambda)x_2$，不妨设 $x_1<x_2$，则有 $x_1<x<x_2$.

由拉格朗日中值定理

$$f(x)-f(x_1)=f'(\xi_1)(x-x_1)=f'(\xi_1)(1-\lambda)(x_2-x_1),x_1<\xi_1<x, \quad ①$$

$$f(x_2)-f(x)=f'(\xi_2)(x_2-x)=f'(\xi_2)\lambda(x_2-x_1),x<\xi_2<x_2, \quad ②$$

由 $\lambda\times$①$-(1-\lambda)\times$②，得

$$\begin{aligned}f(x)-\lambda f(x_1)-(1-\lambda)f(x_2)&=\lambda(1-\lambda)(x_2-x_1)[f'(\xi_1)-f'(\xi_2)]\\&=\lambda(1-\lambda)(x_2-x_1)f''(\xi)(\xi_1-\xi_2),\end{aligned}$$

其中 $\xi_1<\xi<\xi_2$，又 $f''(\xi)>0,\lambda(1-\lambda)>0,x_2-x_1>0,\xi_1-\xi_2<0$，所以上式有

$$f(x)-\lambda f(x_1)-(1-\lambda)f(x_2)<0,$$

即 $$f[\lambda x_1+(1-\lambda)x_2]<\lambda f(x_1)+(1-\lambda)f(x_2).$$

基础习题精练

习 题

1.7.1 设当 $x>0$ 时，方程 $kx+\frac{1}{x^2}=1$ 有且仅有一个根，求 k 的取值范围.

1.7.2 证明对任意常数 a,b，且 $a<b$，都有 $\sin b-\sin a\leqslant b-a$.

1.7.3 证明对任意 $x\in(-\infty,+\infty)$，都有 $x-x^2<\frac{1}{e}$.

1.7.4 设 $b>a>e$，证明 $a^b>b^a$.

1.7.5 设 $0<x<1$，求证 $xe^{-x}>\frac{1}{x}e^{-\frac{1}{x}}$.

1.7.6 证明：$\left(\ln\frac{1+x}{x}-\frac{1}{1+x}\right)^2<\frac{1}{x(1+x)^2}(x>0)$.

解 答

1.7.1 解 设 $f(x)=kx+\frac{1}{x^2}-1$，则 $f'(x)=k-\frac{2}{x^3},f''(x)=\frac{6}{x^4}>0$.

当 $k\leqslant0$ 时，$f'(x)<0$，$f(x)$ 为单调减函数. 又

$$\lim_{x\to0^+}f(x)=+\infty,\quad \lim_{x\to+\infty}f(x)=\begin{cases}-\infty, & k<0,\\ -1, & k=0.\end{cases}$$

所以，当 $k\leqslant0$ 时，$f(x)=kx+\frac{1}{x^2}-1=0$ 在 $(0,+\infty)$ 内仅有一个根.

当 $k>0$ 时，令 $f'(x)=0$，得唯一驻点 $x=\sqrt[3]{\frac{2}{k}}$. 又 $f''(x)>0$，所以 $x=\sqrt[3]{\frac{2}{k}}$ 为极小值点，且 $y=f(x)$ 的图形在 $(0,+\infty)$ 内是凹的，所以，当极小值为零，即

$$f\left(\sqrt[3]{\frac{2}{k}}\right)=k\cdot\sqrt[3]{\frac{2}{k}}+\frac{1}{\left(\sqrt[3]{\frac{2}{k}}\right)^2}-1=0$$

时，原方程有且仅有一个根. 由上式解得 $k=\frac{2}{9}\sqrt{3}$.

综上，当 $k\leqslant 0$ 或 $k=\frac{2}{9}\sqrt{3}$ 时，方程有且仅有一个根.

1.7.2 **证明** $\sin b-\sin a\leqslant b-a\Leftrightarrow\frac{\sin b-\sin a}{b-a}\leqslant 1$. 令 $f(x)=\sin x$，则 $f(x)$ 在 $[a,b]$ 上满足拉格朗日中值定理，所以存在一点 $\xi\in(a,b)$，使 $\frac{\sin b-\sin a}{b-a}=f'(\xi)=\cos\xi\leqslant 1$，得证.

1.7.3 **证明** 令 $F(x)=\frac{1}{e}-x+x^2$，由 $F'(x)=-1+2x=0$，得唯一驻点 $x=\frac{1}{2}$，且 $F''(x)=2>0$，所以 $F\left(\frac{1}{2}\right)$ 为函数 $F(x)$ 的最小值，又 $F\left(\frac{1}{2}\right)=\frac{1}{e}-\frac{1}{2}+\frac{1}{4}=\frac{1}{e}-\frac{1}{4}>0$，故对任意的 $x\in(-\infty,+\infty)$，都有

$$F(x)\geqslant F\left(\frac{1}{2}\right)>0,$$

故

$$\frac{1}{e}-x+x^2>0,\text{即 } x-x^2<\frac{1}{e}.$$

1.7.4 **分析** $a^b>b^a\Leftrightarrow b\ln a>a\ln b\Leftrightarrow\frac{\ln a}{a}>\frac{\ln b}{b}$. 可用函数的单调性证明.

证明 设 $f(x)=\frac{\ln x}{x}$，$x\in[a,b]$，其中 $b>a>e$，则

$$f'(x)=\frac{1-\ln x}{x^2},$$

其中 $\ln x>\ln e=1$，所以 $f'(x)<0$，即函数 $f(x)$ 单调减少. 所以，当 $b>a>e$ 时，$\frac{\ln a}{a}>\frac{\ln b}{b}$，即 $b\ln a>a\ln b$，则 $\ln a^b>\ln b^a$，即 $a^b>b^a$.

1.7.5 **证明** 令 $f(x)=2\ln x-x+\frac{1}{x}$，$x\in(0,1)$，则有

$$f'(x)=\frac{2}{x}-1-\frac{1}{x^2}=\frac{2x-x^2-1}{x^2}=\frac{-(x-1)^2}{x^2}<0,$$

故 $f(x)$ 在 $(0,1)$ 内单调减少，从而 $f(x)>f(1)=0$，即在 $(0,1)$ 内 $f(x)>0$，从而 $\ln x-x>-\ln x-\frac{1}{x}$，

即

$$xe^{-x}>\frac{1}{x}e^{-\frac{1}{x}}.$$

1.7.6 **证明** 只要证明当 $x>0$ 时，$\left|\ln\frac{1+x}{x}-\frac{1}{1+x}\right|-\frac{1}{\sqrt{x}(1+x)}<0$ 即可.

令

$$f(x)=\ln\frac{1+x}{x}-\frac{1}{1+x}=\ln(1+x)-\ln x-\frac{1}{1+x},$$

则

$$f'(x)=\frac{1}{1+x}-\frac{1}{x}+\frac{1}{(1+x)^2}$$

$$=\frac{(1+x)x-(1+x)^2+x}{x(1+x)^2}$$

$$=\frac{-1}{x(1+x)^2}<0,$$

又 $\lim\limits_{x\to+\infty}f(x)=0$，所以 $f(x)>0$.

令
$$g(x)=\ln\frac{x+1}{x}-\frac{1}{1+x}-\frac{1}{\sqrt{x}(1+x)},$$

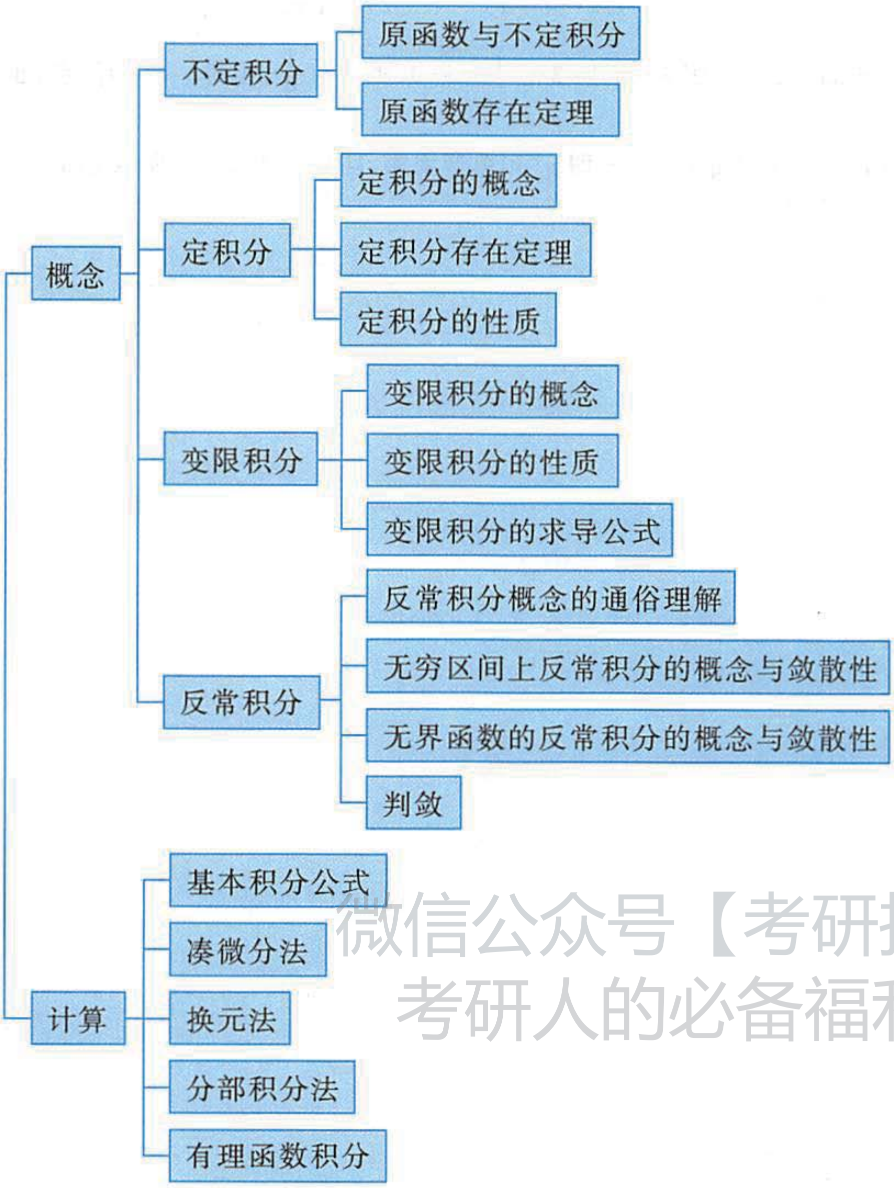
则
$$g'(x)=\frac{1}{1+x}-\frac{1}{x}+\frac{1}{(1+x)^2}+\frac{\frac{1}{2\sqrt{x}}\cdot(1+x)+\sqrt{x}}{x(1+x)^2}$$

$$=\frac{-2\sqrt{x}+1+x+2x}{2x^{\frac{3}{2}}(1+x)^2}=\frac{1+3x-2\sqrt{x}}{2x^{\frac{3}{2}}(1+x)^2}.$$

再令 $h(x)=1+3x-2\sqrt{x}$，则 $h'(x)=3-\frac{1}{\sqrt{x}}\overset{令}{=\!=}0$，得 $x=\frac{1}{9}$，为唯一极小值点，也即最小值点，且 $h_{\min}=\frac{2}{3}>0$，故 $h(x)>0$，于是 $g'(x)>0$，即 $g(x)$ 单调增加，且 $\lim\limits_{x\to+\infty}g(x)=0$，所以 $g(x)<0$，证毕.

第8讲 一元函数积分学的概念与计算

基础知识结构



基础内容精讲

一、概念

(一)不定积分

1. 原函数与不定积分

设函数 $f(x)$ 定义在某区间 I 上，若存在可导函数 $F(x)$，对于该区间上任意一点都有 $F'(x)=f(x)$ 成立，则称 $F(x)$ 是 $f(x)$ 在区间 I 上的**一个原函数**. 称 $\int f(x)\mathrm{d}x=F(x)+C$ 为 $f(x)$ 在区间 I 上的**不定积分**，其中 C 为任意常数.

> **【注】** 谈到函数 $f(x)$ 的原函数与不定积分，必须指明 $f(x)$ 所定义的区间.

2. 原函数(不定积分)存在定理

(1)连续函数 $f(x)$ 必有原函数 $F(x)$.

(2)含有第一类间断点、无穷间断点的函数 $f(x)$ 在包含该间断点的区间内必没有原函数 $F(x)$.

> **【注】** 在考研中多次考到上面这个结论，但一般教材中并没有给出. 它的证明见例 1.8.2.

(二)定积分

1. 定积分的概念

若函数 $f(x)$ 在区间 $[a,b]$ 上有界，在 (a,b) 上任取 $n-1$ 个分点 $x_k(k=1,2,3,\cdots,n-1)$，定义 $x_0=a$ 和 $x_n=b$，且 $a=x_0<x_1<x_2<x_3<\cdots<x_{n-1}<x_n=b$，记 $\Delta x_k=x_k-x_{k-1}$，$k=1,2,3,\cdots,n$. 并任取 $\xi_k\in[x_{k-1},x_k]$，记 $\lambda=\max\limits_{1\leqslant k\leqslant n}\{\Delta x_k\}$，当 $\lambda\to0$ 时，极限 $\lim\limits_{\lambda\to0}\sum\limits_{k=1}^{n}f(\xi_k)\Delta x_k$ 存在且与分点 x_k 及 ξ_k 的取法无关，则称函数 $f(x)$ 在区间 $[a,b]$ 上可积，即 $\int_a^b f(x)\mathrm{d}x=\lim\limits_{\lambda\to0}\sum\limits_{k=1}^{n}f(\xi_k)\Delta x_k$.

> **【注】** (1)若 $f(x)<0$，曲边梯形就在 x 轴下方，定积分的绝对值仍等于曲边梯形的面积，但定积分的值是负的.
>
> (2)当我们说到"a 到 b 上的定积分"时，不要总认为 $a<b$，事实上，$a>b$ 的情形是完全可以的，不过注意，$a<b$ 时，$\mathrm{d}x>0$；$a>b$ 时，$\mathrm{d}x<0$.
>
> (3)定积分的定义由德国数学家波恩哈德·黎曼(Bernhard Riemann)给出，故这种积分又被称为**黎曼积分**.
>
> **(4)定积分的精确定义(重点)**
>
> $$\int_a^b f(x)\mathrm{d}x=\lim_{n\to\infty}\sum_{i=1}^{n}f\left(a+\frac{b-a}{n}i\right)\frac{b-a}{n}.$$
>
> 使用定积分的精确定义，主要是为了计算一些特殊形式的数列极限，其详细步骤和例题参见"基础例题精解"的"三、定积分的精确定义"部分.

黎曼

(1826—1866)

2. 定积分存在定理

定积分的存在性，也称之为**一元函数的（常义）可积性**. 这里的“常义”是指“区间有限，函数有界”，也有人称为“黎曼”可积性，与后面要谈到的“区间无穷，函数无界”的“反常”积分有所区别. 在本讲中所谈到的可积性都是指常义可积性.

按照《考试大纲》，定积分存在定理包括下面两个方面.

(1)定积分存在的充分条件.

① 若 $f(x)$ 在 $[a,b]$ 上连续，则 $\int_a^b f(x)\mathrm{d}x$ 存在.

② 若 $f(x)$ 在 $[a,b]$ 上单调，则 $\int_a^b f(x)\mathrm{d}x$ 存在.

③ 若 $f(x)$ 在 $[a,b]$ 上有界，且只有有限个间断点，则 $\int_a^b f(x)\mathrm{d}x$ 存在.

(2)定积分存在的必要条件.

可积函数必有界，即若定积分 $\int_a^b f(x)\mathrm{d}x$ 存在，则 $f(x)$ 在 $[a,b]$ 上必有界.

【注1】 关于定积分存在的必要条件，不妨这样理解：当我们任意分割图形底边为若干小段时，若 $f(x)$ 在区间 $[a,b]$ 上无界，则至少存在一个小段 Δx，在 Δx 上，$f(x)$ 可以任意大，于是一个“小竖条”的面积 $f(x)\Delta x$ 便可以无穷大，这样整个曲边梯形的面积就是无穷大，于是极限就不存在了，所以可积函数必有界.

【注2】 函数不定积分存在定理与定积分存在定理的区别与联系见例 1.8.3.

3. 定积分的性质（以下假设所写积分均存在）

性质1（求区间长度） 假设 $a<b$，则 $\int_a^b \mathrm{d}x = b-a = L$，其中 L 为区间 $[a,b]$ 的长度.

性质2（积分的线性性质） 设 k_1,k_2 为常数，则 $\int_a^b [k_1 f(x) \pm k_2 g(x)]\mathrm{d}x = k_1\int_a^b f(x)\mathrm{d}x \pm k_2\int_a^b g(x)\mathrm{d}x$.

性质3（积分的可加（拆）性） 无论 a,b,c 的大小如何，总有 $\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x$.

性质4（积分的保号性） 若在区间 $[a,b]$ 上 $f(x)\leqslant g(x)$，则有 $\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x$.

特殊地，有

$$\left|\int_a^b f(x)\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\mathrm{d}x.$$

【注】 事实上，设 $f(x)$ 是 $[a,b]$ 上非负的连续函数，只要 $f(x)$ 不恒等于零，则必有

$$\int_a^b f(x)\mathrm{d}x > 0.$$

在有些积分不等式的证明与定积分值的估计中，要求获得严格的不等式结果，便需要用到这个结论. 其证明见例 1.8.9.

性质5（估值定理） 设 M,m 分别是 $f(x)$ 在 $[a,b]$ 上的最大值和最小值，L 为区间 $[a,b]$ 的长度，则有

$$mL \leqslant \int_a^b f(x)\mathrm{d}x \leqslant ML.$$

性质 6(中值定理)　设 $f(x)$在闭区间$[a,b]$上连续，则在$[a,b]$上至少存在一点 ξ，使得

$$\int_a^b f(x)\mathrm{d}x = f(\xi)(b-a).$$

(三)变限积分

1. 变限积分的概念

当 x 在$[a,b]$上变动时，对应于每一个 x 值，积分$\int_a^x f(t)\mathrm{d}t$ 就有一个确定的值，因此$\int_a^x f(t)\mathrm{d}t$ 是一个变上限的函数，记作 $\Phi(x)=\int_a^x f(t)\mathrm{d}t(a\leqslant x\leqslant b)$，称函数 $\Phi(x)$为**变上限的定积分**. 同理可以定义变下限的定积分和上、下限都变化的定积分，这些都称为**变限积分**. 事实上，变限积分就是定积分的推广.

2. 变限积分的性质

(1) 函数 $f(x)$ 在$[a,b]$ 上可积，则函数 $F(x)=\int_a^x f(t)\mathrm{d}t$ 在$[a,b]$ 上连续.

(2) 函数 $f(x)$ 在$[a,b]$ 上连续，则函数 $F(x)=\int_a^x f(t)\mathrm{d}t$ 在$[a,b]$ 上可导.

【注】　这两个性质的证明见例 1.8.4 和例 1.8.1. 考研中常用到函数 $f(x)$ 的原函数的一个具体形式：如果函数 $f(x)$ 在区间$[a,b]$上连续，则 $F(x)=\int_a^x f(t)\mathrm{d}t$ 是 $f(x)$ 在区间$[a,b]$ 上的一个原函数. 由上述性质可以得出一个重要结论：对于变限积分 $F(x)=\int_a^x f(t)\mathrm{d}t$，只要它存在，就必然是连续的. 记住这一点，在有些考研题中可以起到重要作用，见例 1.8.5.

3. 变限积分的求导公式

设 $F(x)=\int_{\varphi_1(x)}^{\varphi_2(x)} f(t)\mathrm{d}t$，其中 $f(x)$ 在$[a,b]$上连续，可导函数 $\varphi_1(x)$ 和 $\varphi_2(x)$ 的值域在$[a,b]$上，则在函数 $\varphi_1(x)$ 和 $\varphi_2(x)$ 的公共定义域上，有 $F'(x)=\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(t)\mathrm{d}t\right]=f[\varphi_2(x)]\varphi'_2(x)-f[\varphi_1(x)]\varphi'_1(x)$.

【注】　我们称上面公式中的 x 为"求导变量"，t 为"积分变量"."求导变量"x 只出现在积分的上、下限时才能使用变限积分求导公式，若"求导变量"x 出现在被积函数中，必须通过恒等变形(比如变量代换等)，将其移出被积函数，才能使用变限积分求导公式.

注例　设函数 $f(x)$ 具有连续导数，求$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_a^x (x-t)f'(t)\mathrm{d}t\right]$.

解　首先将"求导变量"x 移出被积函数，

$$\int_a^x (x-t)f'(t)\mathrm{d}t=\int_a^x xf'(t)\mathrm{d}t-\int_a^x tf'(t)\mathrm{d}t=x\int_a^x f'(t)\mathrm{d}t-\int_a^x tf'(t)\mathrm{d}t,$$

然后再使用变限积分求导公式，于是

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_a^x (x-t)f'(t)\mathrm{d}t\right]=\int_a^x f'(t)\mathrm{d}t+xf'(x)-xf'(x)=f(x)-f(a).$$

(四)反常积分

1. 反常积分概念的通俗理解

反常积分的概念很容易从定积分的概念中引出. 前面已经指出，定积分存在有两个必要条件：一是积

分区间有限，二是被积函数有界．如果破坏了积分区间的有限性，就引出无穷区间上的反常积分；如果破坏了被积函数的有界性，就引出无界函数的反常积分．我们以无穷区间上的反常积分为例，来通俗地解释一下到底什么是反常积分．

定积分的几何背景是一个曲边梯形的面积，现在我们假设曲边梯形的底边长$l\to+\infty$，高为$h=f(x)$，则面积$S=l\cdot h$，问题是，什么情况下此面积才会存在呢？我们自然想到极限理论中的未定式“$\infty\cdot 0$”型，当$x\to+\infty$，也就是底边长$l\to+\infty$时，只有$f(x)\to 0$，S才可能存在，故一般来说，

$$\int_a^{+\infty} f(x)\mathrm{d}x \text{ 收敛} \Rightarrow \lim_{x\to+\infty} f(x)=0.$$

反过来说，若$a>0$，高$h=\frac{1}{\sqrt{x}}$，当$x\to+\infty$时，$\frac{1}{\sqrt{x}}\to 0$，虽然高$h\to 0$，但面积$S=\int_a^{+\infty}\frac{1}{\sqrt{x}}\mathrm{d}x\to+\infty$，发散．很好理解，这是因为高$h$“无穷小的程度”小于底边长$l$“无穷大的程度”，结果还会发散．

所以，高$h=f(x)$的“无穷小的程度”就成为反常积分敛散性判别中的关键．一般来说，

$f(x)$**“越小”**，则$\int_a^{+\infty} f(x)\mathrm{d}x$**“越容易收敛”**．

这个结论对于理解后面的敛散性判别法有着重要的意义，即便我在下面会给出一个令人遗憾的特例，它也不会影响上述结论的重要地位．

【注】 事实上，$\int_a^{+\infty} f(x)\mathrm{d}x$收敛不一定能推出$\lim\limits_{x\to+\infty} f(x)=0$．比如，设

$$f(x)=\begin{cases} n, & x\in\left[n,n+\frac{1}{n\cdot 2^n}\right],n=1,2,\cdots, \\ 0, & \text{其他}, \end{cases}$$

则$\int_0^{+\infty} f(x)\mathrm{d}x=\int_0^1 0\mathrm{d}x+\int_1^{\frac{3}{2}}\mathrm{d}x+\int_{\frac{3}{2}}^2 0\mathrm{d}x+\int_2^{\frac{17}{8}} 2\mathrm{d}x+\cdots=\sum\limits_{n=1}^{\infty}\frac{1}{2^n}$收敛，但$\lim\limits_{x\to+\infty} f(x)\neq 0$．

这是一个“特例”，为什么会出现这种情形？你看$f(x)$的正值区间是$\left[n,n+\frac{1}{n\cdot 2^n}\right]$，其区间长度为$\frac{1}{n\cdot 2^n}$，而其“高”为$n$，故此面积为$\frac{1}{2^n}$，依然可以收敛是因为正值区间上底边长“无穷小的程度”超过了高“无穷大的程度”．

2．无穷区间上反常积分的概念与敛散性

(1)$\int_a^{+\infty} f(x)\mathrm{d}x$的定义为$\int_a^{+\infty} f(x)\mathrm{d}x=\lim\limits_{b\to+\infty}\int_a^b f(x)\mathrm{d}x$．

若上述极限存在，则称反常积分$\int_a^{+\infty} f(x)\mathrm{d}x$**收敛**，否则称为**发散**．

(2)$\int_{-\infty}^b f(x)\mathrm{d}x$的定义为$\int_{-\infty}^b f(x)\mathrm{d}x=\lim\limits_{a\to-\infty}\int_a^b f(x)\mathrm{d}x$．

若上述极限存在，则称反常积分$\int_{-\infty}^b f(x)\mathrm{d}x$**收敛**，否则称为**发散**．

(3)$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$的定义为$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{-\infty}^c f(x)\mathrm{d}x+\int_c^{+\infty} f(x)\mathrm{d}x$．

若右边两个反常积分都收敛，则称反常积分$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$**收敛**，否则称为**发散**．

【注】 在反常积分中，一般把“∞”和使得函数极限为无穷的点（瑕点）统称为奇点．

3. 无界函数的反常积分的概念与敛散性

(1) 若 b 是 $f(x)$ 的唯一瑕点，则无界函数 $f(x)$ 的反常积分 $\int_a^b f(x)\mathrm{d}x$ 定义为

$$\int_a^b f(x)\mathrm{d}x=\lim_{\varepsilon\to 0^+}\int_a^{b-\varepsilon} f(x)\mathrm{d}x.$$

若上述极限存在，则称反常积分 $\int_a^b f(x)\mathrm{d}x$ **收敛**，否则称为**发散**.

> **【注】** 当 $x=b$ 为 $f(x)$ 的无穷间断点，这时 $f(x)$ 便是一个无界函数了，积分 $\int_a^b f(x)\mathrm{d}x$ 也可能存在. 细心的考生可能会联想到，前面我们不是说"$\int_a^b f(x)\mathrm{d}x$ 存在的必要条件是 $f(x)$ 有界"吗？这不是矛盾了吗？事实上，前面所说的 $\int_a^b f(x)\mathrm{d}x$ 是**定积分(黎曼积分)**，而这里的 $\int_a^b f(x)\mathrm{d}x$ 是**反常积分**，它们并不是一个概念，所以没有任何矛盾. 只是当考生读完这一段后，最好今后在提到积分存在时，特别强调一下，是定积分存在(黎曼可积，常义可积)，还是反常积分存在(广义可积).

(2) 若 a 是 $f(x)$ 的唯一瑕点，则无界函数 $f(x)$ 的反常积分 $\int_a^b f(x)\mathrm{d}x$ 定义为

$$\int_a^b f(x)\mathrm{d}x=\lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^{b} f(x)\mathrm{d}x.$$

若上述极限存在，则称反常积分 $\int_a^b f(x)\mathrm{d}x$ **收敛**，否则称为**发散**.

(3) 若 $c\in(a,b)$是 $f(x)$的唯一瑕点，则无界函数 $f(x)$ 的反常积分 $\int_a^b f(x)\mathrm{d}x$ 定义为

$$\int_a^b f(x)\mathrm{d}x=\int_a^c f(x)\mathrm{d}x+\int_c^b f(x)\mathrm{d}x.$$

若上述右边两个反常积分都收敛，则称反常积分 $\int_a^b f(x)\mathrm{d}x$ **收敛**，否则称为**发散**.

关于敛散性的判别，具体参看"基础例题精解"的"六、反常积分的计算与敛散性判别"部分.

二、计算

(一)不定积分的积分法

1. 基本积分公式

①$\int x^k\mathrm{d}x=\dfrac{1}{k+1}x^{k+1}+C,k\neq -1;\begin{cases}\int\dfrac{1}{x^2}\mathrm{d}x=-\dfrac{1}{x}+C,\\ \int\dfrac{1}{\sqrt{x}}\mathrm{d}x=2\sqrt{x}+C.\end{cases}$

②$\int\dfrac{1}{x}\mathrm{d}x=\ln|x|+C.$

③$\int \mathrm{e}^x\mathrm{d}x=\mathrm{e}^x+C;\int a^x\mathrm{d}x=\dfrac{a^x}{\ln a}+C,a>0$ 且 $a\neq 1.$

④ $\int\sin x\mathrm{d}x=-\cos x+C;\int\cos x\mathrm{d}x=\sin x+C;$

$\int\tan x\mathrm{d}x=-\ln|\cos x|+C;\int\cot x\mathrm{d}x=\ln|\sin x|+C;$

$$\int \frac{\mathrm{d}x}{\cos x} = \int \sec x\mathrm{d}x = \ln|\sec x + \tan x| + C;$$

$$\int \frac{\mathrm{d}x}{\sin x} = \int \csc x\mathrm{d}x = \ln|\csc x - \cot x| + C;$$

$$\int \sec^2 x\mathrm{d}x = \tan x + C; \int \csc^2 x\mathrm{d}x = -\cot x + C;$$

$$\int \sec x\tan x\mathrm{d}x = \sec x + C; \int \csc x\cot x\mathrm{d}x = -\csc x + C.$$

⑤$$\begin{cases} \displaystyle\int \frac{1}{1+x^2}\mathrm{d}x = \arctan x + C, \\ \displaystyle\int \frac{1}{a^2+x^2}\mathrm{d}x = \frac{1}{a}\arctan\frac{x}{a} + C(a>0). \end{cases}$$

⑥$$\begin{cases} \displaystyle\int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \arcsin x + C, \\ \displaystyle\int \frac{1}{\sqrt{a^2-x^2}}\mathrm{d}x = \arcsin\frac{x}{a} + C(a>0). \end{cases}$$

⑦$$\begin{cases} \displaystyle\int \frac{1}{\sqrt{x^2+a^2}}\mathrm{d}x = \ln(x+\sqrt{x^2+a^2}) + C(\text{常见 } a=1), \\ \displaystyle\int \frac{1}{\sqrt{x^2-a^2}}\mathrm{d}x = \ln|x+\sqrt{x^2-a^2}| + C(|x|>|a|). \end{cases}$$

⑧$$\int \frac{1}{x^2-a^2}\mathrm{d}x = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C\left(\int \frac{1}{a^2-x^2}\mathrm{d}x = \frac{1}{2a}\ln\left|\frac{x+a}{x-a}\right| + C\right).$$

⑨$$\int \sqrt{a^2-x^2}\mathrm{d}x = \frac{a^2}{2}\arcsin\frac{x}{a} + \frac{x}{2}\sqrt{a^2-x^2} + C(a>|x|\geqslant 0).$$

⑩$$\int \sin^2 x\mathrm{d}x = \frac{x}{2} - \frac{\sin 2x}{4} + C\left(\sin^2 x = \frac{1-\cos 2x}{2}\right);$$

$$\int \cos^2 x\mathrm{d}x = \frac{x}{2} + \frac{\sin 2x}{4} + C\left(\cos^2 x = \frac{1+\cos 2x}{2}\right);$$

$$\int \tan^2 x\mathrm{d}x = \tan x - x + C(\tan^2 x = \sec^2 x - 1);$$

$$\int \cot^2 x\mathrm{d}x = -\cot x - x + C(\cot^2 x = \csc^2 x - 1).$$

2. 凑微分法

(1) 基本思想$\int f[g(x)]g'(x)\mathrm{d}x = \int f[g(x)]\mathrm{d}[g(x)] = \int f(u)\mathrm{d}u$.

当被积函数比较复杂时，拿出一部分放到 d 后面去，若能凑成$\int f(u)\mathrm{d}u$ 的形式，则凑微分成功. 比如，

$$\int \frac{\ln^5 x}{x}\mathrm{d}x = \int \ln^5 x \cdot \frac{1}{x}\mathrm{d}x = \int \ln^5 x\mathrm{d}(\ln x) = \frac{\ln^6 x}{6} + C.$$

(2)熟练掌握基本积分公式及常用的凑微分公式. 比如能熟练计算下面这种题目：

$$\int \frac{\sqrt{x}}{\sqrt{4-x^3}}\mathrm{d}x = \frac{2}{3}\int \frac{\mathrm{d}\left(x^{\frac{3}{2}}\right)}{\sqrt{4-\left(x^{\frac{3}{2}}\right)^2}} = \frac{2}{3}\int \frac{\mathrm{d}\left(x^{\frac{3}{2}}\right)}{2\cdot\sqrt{1-\left(\frac{1}{2}x^{\frac{3}{2}}\right)^2}}$$

$$= \frac{2}{3}\int \frac{\mathrm{d}\left(\frac{1}{2}x^{\frac{3}{2}}\right)}{\sqrt{1-\left(\frac{1}{2}x^{\frac{3}{2}}\right)^2}} = \frac{2}{3}\arcsin\left(\frac{1}{2}x^{\frac{3}{2}}\right) + C.$$

3. 换元法

(1) 基本思想 $\int f(x)\mathrm{d}x \xlongequal{x=g(u)} \int f[g(u)]\mathrm{d}[g(u)]=\int f[g(u)]g'(u)\mathrm{d}u.$

当被积函数不容易积分(比如含有根式，含有反三角函数)时，可以通过换元的方法从 d 后面拿出一部分放到前面来，就成为 $\int f[g(u)]g'(u)\mathrm{d}u$ 的形式，若 $f[g(u)]g'(u)$ 容易积分，则换元成功.

【注】 $x=g(u)$ 须是单调可导函数，且不要忘记计算结束后用反函数 $u=g^{-1}(x)$ 回代.

(2) 归纳总结换元法的思维结构.

①三角函数代换——当被积函数含有如下根式时，可作三角代换，这里 $a>0$.

$$\begin{cases} \sqrt{a^2-x^2} \xrightarrow{\text{令}} x=a\sin t, \quad |t|<\dfrac{\pi}{2}, \\ \sqrt{a^2+x^2} \xrightarrow{\text{令}} x=a\tan t, \quad |t|<\dfrac{\pi}{2}, \\ \sqrt{x^2-a^2} \xrightarrow{\text{令}} x=a\sec t, \quad \begin{cases} \text{若 } x>0, \quad \text{则 } 0<t<\dfrac{\pi}{2}, \\ \text{若 } x<0, \quad \text{则 } \dfrac{\pi}{2}<t<\pi. \end{cases} \end{cases}$$

②恒等变形后作三角函数代换——当被积函数含有根式 $\sqrt{ax^2+bx+c}$ 时，可化为以下三种形式 $\sqrt{\varphi^2(x)+k^2}$，$\sqrt{\varphi^2(x)-k^2}$，$\sqrt{k^2-\varphi^2(x)}$，再作三角函数代换.

③根式代换——当被积函数含有根式 $\sqrt[n]{ax+b}$，$\sqrt{\dfrac{ax+b}{cx+d}}$，$\sqrt{ae^{bx}+c}$ 等时，一般令根式 $\sqrt{*}=t$(因为事实上，很难通过根号内换元的办法凑成平方，所以根号无法去掉). 对既含有 $\sqrt[n]{ax+b}$，也含有 $\sqrt[m]{ax+b}$ 的函数，一般取 m，n 的最小公倍数 l，令 $\sqrt[l]{ax+b}=t$.

④倒代换——当被积函数分母的幂次比分子高两次及两次以上时，作倒代换，令 $x=\dfrac{1}{t}$.

⑤复杂函数的直接代换——当被积函数中含有 a^x，e^x，$\ln x$，$\arcsin x$，$\arctan x$ 等时，可考虑直接令复杂函数等于 t，值得指出的是，当 $\ln x$，$\arcsin x$，$\arctan x$ 与 $P_n(x)$ 或 e^{ax} 作乘法时(其中 $P_n(x)$ 为 x 的 n 次多项式)，优先考虑分部积分法.

比如能熟练计算这种题目：计算 $\int e^{\sqrt{2x-1}}\mathrm{d}x$.

解 令 $t=\sqrt{2x-1}$，则 $x=\dfrac{t^2+1}{2}$，$\mathrm{d}x=t\mathrm{d}t$，于是

$$\begin{aligned}\int e^{\sqrt{2x-1}}\mathrm{d}x &= \int e^t t\mathrm{d}t=\int t\mathrm{d}(e^t) \\ &= te^t-\int e^t\mathrm{d}t=te^t-e^t+C \\ &= (t-1)e^t+C \\ &= (\sqrt{2x-1}-1)e^{\sqrt{2x-1}}+C.\end{aligned}$$

4. 分部积分法

(1) 基本思想 $\int u\mathrm{d}v=uv-\int v\mathrm{d}u.$

一目了然，这个方法主要适用于求 $\int u\mathrm{d}v$ 比较困难，而 $\int v\mathrm{d}u$ 比较容易的情形.

什么函数积分后会“简单”些？宜取作 v；什么函数微分后会“简单”些？宜取作 u. 选取的一般原则：设 $P_n(x)$ 为 x 的 n 次多项式，

①被积函数为 $P_n(x)e^{kx}$，$P_n(x)\sin ax$，$P_n(x)\cos ax$ 等形式时，一般来说选取 $u=P_n(x)$；

②被积函数为 $e^{ax}\sin bx$，$e^{ax}\cos bx$ 等形式时，u 可以取其中两因子中的任意一个；

③被积函数为 $P_n(x)\ln x$，$P_n(x)\arcsin x$，$P_n(x)\arctan x$ 等形式时，一般分别选取

$$u=\ln x,\quad u=\arcsin x,\quad u=\arctan x.$$

(2)分部积分法的推广公式与 $\int P_n(x)e^{kx}dx$，$\int P_n(x)\sin ax\,dx$，$\int P_n(x)\cos bx\,dx$.

设函数 u 与 v 具有直到第 $(n+1)$ 阶的连续导数，并根据分部积分公式

$$\int u\mathrm{d}v = uv-\int v\mathrm{d}u\ ,$$

则有
$$\int uv^{(n+1)}\mathrm{d}x = uv^{(n)}-u'v^{(n-1)}+u''v^{(n-2)}-\cdots+(-1)^n u^{(n)}v+(-1)^{n+1}\int u^{(n+1)}v\mathrm{d}x.$$

【注】 **证明** 在公式 $\int u\mathrm{d}v = uv-\int v\mathrm{d}u$ 中以 $v^{(n)}$ 代替 v，则

$$\int uv^{(n+1)}\mathrm{d}x = \int u\mathrm{d}[v^{(n)}] = uv^{(n)}-\int v^{(n)}\mathrm{d}u = uv^{(n)}-\int u'v^{(n)}\mathrm{d}x.$$

同理可得
$$\int u'v^{(n)}\mathrm{d}x = u'v^{(n-1)}-\int u''v^{(n-1)}\mathrm{d}x,$$

$$\int u''v^{(n-1)}\mathrm{d}x = u''v^{(n-2)}-\int u'''v^{(n-2)}\mathrm{d}x,$$

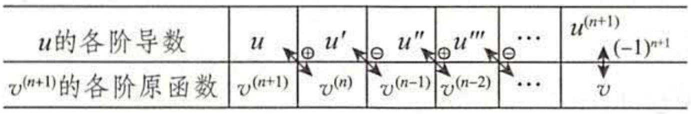
……

$$\int u^{(n)}v'\mathrm{d}x = u^{(n)}v-\int u^{(n+1)}v\mathrm{d}x.$$

联立以上式子，并保留第一个和最后一个积分，便可得到**分部积分法的推广公式**：

$$\int uv^{(n+1)}\mathrm{d}x = uv^{(n)}-u'v^{(n-1)}+u''v^{(n-2)}-\cdots+(-1)^n u^{(n)}v+(-1)^{n+1}\int u^{(n+1)}v\mathrm{d}x.$$

事实上，可写成如下表格



u的各阶导数	u	⊕	u'	⊖	u''	⊕	u'''	⊖	$\cdots$	$u^{(n+1)}$ $(-1)^{n+1}$
$v^{(n+1)}$的各阶原函数	$v^{(n+1)}$		$v^{(n)}$		$v^{(n-1)}$		$v^{(n-2)}$		$\cdots$	v

计算方法：以 u 作起点左上、右下错位相乘，各项符号“＋”“－”相间，最后一项为 $(-1)^{n+1}\int u^{(n+1)}v\mathrm{d}x$.

对于 $\int P_n(x)e^{kx}dx$，$\int P_n(x)\sin ax\,dx$，$\int P_n(x)\cos bx\,dx$ 三种积分，其中 $P_n(x)$ 是 x 的 n（n 为正整数）次多项式，令 $u=P_n(x)$，则 $u^{(n+1)}=0$，于是积分便可顺利算出.

比如，求不定积分 $\int(x^3+2x+6)e^{2x}dx$.

则

x^3+2x+6	$3x^2+2$	$6x$	6	0
e^{2x}	$\frac{1}{2}e^{2x}$	$\frac{1}{4}e^{2x}$	$\frac{1}{8}e^{2x}$	$\frac{1}{16}e^{2x}$

利用上述表格，可得

$$原式=(x^3+2x+6)\left(\frac{1}{2}e^{2x}\right)-(3x^2+2)\left(\frac{1}{4}e^{2x}\right)+6x\left(\frac{1}{8}e^{2x}\right)-6\left(\frac{1}{16}e^{2x}\right)+\int 0\cdot\left(\frac{1}{16}e^{2x}\right)dx$$

$$=\left(\frac{1}{2}x^3-\frac{3}{4}x^2+\frac{7}{4}x+\frac{17}{8}\right)e^{2x}+C.$$

5. 有理函数的积分

(1) 定义　形如$\int\frac{P_n(x)}{Q_m(x)}dx(n<m)$的积分称为**有理函数的积分**，其中$P_n(x)$，$Q_m(x)$分别是$x$的$n$次多项式和$m$次多项式.

(2) 方法　先将$Q_m(x)$因式分解，再把$\frac{P_n(x)}{Q_m(x)}$拆成若干最简有理分式之和.

(3)分解的基本原则.

①$Q_m(x)$的一次单因式$ax+b$产生一项$\frac{A}{ax+b}$；

②$Q_m(x)$的k重一次因式$(ax+b)^k$产生k项，分别为$\frac{A_1}{ax+b}+\frac{A_2}{(ax+b)^2}+\cdots+\frac{A_k}{(ax+b)^k}$；

③$Q_m(x)$的二次单因式px^2+qx+r产生一项$\frac{Ax+B}{px^2+qx+r}$；

④$Q_m(x)$的k重二次因式$(px^2+qx+r)^k$产生k项

$$\frac{A_1x+B_1}{px^2+qx+r}+\frac{A_2x+B_2}{(px^2+qx+r)^2}+\cdots+\frac{A_kx+B_k}{(px^2+qx+r)^k}.$$

【注】 具体见例1.8.21，例1.8.22.

(二)定积分的计算

定积分的计算，主要依赖于牛顿-莱布尼茨公式.

若$f(x)$在$[a,b]$上连续，则

$$\int_a^b f(x)dx\xlongequal{F'(x)=f(x)}F(x)\Big|_a^b=F(b)-F(a).$$

【注】 前面已经指出，含有间断点的函数也可能存在原函数，例如，

$$f(x)=\begin{cases}2x\sin\frac{1}{x}-\cos\frac{1}{x}, & x\neq 0,\\ 0, & x=0\end{cases}$$

在$x=0$处不连续，但显然$F(x)=\begin{cases}x^2\sin\frac{1}{x}, & x\neq 0,\\ 0, & x=0\end{cases}$是$f(x)$的一个原函数，因为$f(x)$是在$[0,1]$上只有一个间断点$(x=0)$的有界函数，所以可积，从而

$$\int_0^1 f(x)dx=F(x)\Big|_0^1=\sin 1.$$

这里用到了**牛顿-莱布尼茨公式的推广**：在积分区间$[a,b]$上只有有限个间断点的被积函数$f(x)$，只要其在$[a,b]$上存在原函数，牛顿-莱布尼茨公式依然成立.

牛顿
(1642—1726)

莱布尼茨
(1646—1716)

由牛顿-莱布尼茨公式结合不定积分的计算方法，有定积分的换元积分法和分部积分法，分别如下.

1. 定积分的换元积分法

设 $f(x)$ 在 $[a,b]$ 上连续，函数 $x=\varphi(t)$ 满足① $\varphi(\alpha)=a,\varphi(\beta)=b$；② $x=\varphi(t)$ 在 $[\alpha,\beta]$（或 $[\beta,\alpha]$）上有连续的导数，且其值域为 $R_\varphi=[a,b]$，则有

$$\int_a^b f(x)\mathrm{d}x=\int_\alpha^\beta f[\varphi(t)]\varphi'(t)\mathrm{d}t.$$

【注】 当 $\varphi(t)$ 的值域 R_φ 超出 $[a,b]$，但 $\varphi(t)$ 满足其余条件时，只要 $f(x)$ 在 R_φ 上连续，则上述结论仍成立.

2. 定积分的分部积分法

$$\int_a^b u(x)v'(x)\mathrm{d}x=u(x)v(x)\Big|_a^b-\int_a^b v(x)u'(x)\mathrm{d}x,$$

这里要求 $u'(x),v'(x)$ 在 $[a,b]$ 上连续.

用牛顿-莱布尼茨公式或用分部积分法计算定积分，方法与求不定积分一致（只是要代入上、下限而已），而用换元法计算定积分时，有一些需特别注意的地方，这些注意的地方，在定理条件中已写明白，但要真正掌握，还需通过后面的例 1.8.25～例 1.8.28 来练习.

【注】 在计算定积分时，下面这些结论是很有用的.

(1)设 $f(x)$ 为连续的偶函数，则

$$\int_{-a}^a f(x)\mathrm{d}x=2\int_0^a f(x)\mathrm{d}x.$$

(2)设 $f(x)$ 为连续的奇函数，则

$$\int_{-a}^a f(x)\mathrm{d}x=0.$$

(3)设 $f(x)$ 是以 T 为周期的连续函数，则

$$\int_a^{a+T} f(x)\mathrm{d}x=\int_0^T f(x)\mathrm{d}x,$$

即在长度为一个周期的区间上的定积分，与该区间的起点位置无关，见例 1.8.7.

(4)设 $f(x)$ 为连续函数，则

$$\int_a^b f(x)\mathrm{d}x=\int_a^b f(a+b-x)\mathrm{d}x,$$

这叫“区间再现公式”，其证明见例 1.8.29.

(5) $$\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x=\begin{cases}\dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{2}{3}\cdot 1, & n\text{ 为大于 1 的奇数},\\[2ex] \dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{1}{2}\cdot\dfrac{\pi}{2}, & n\text{ 为正偶数}.\end{cases}$$

这个式子的证明见例 1.8.33.

基础例题精解

一、一元函数积分学的概念与性质

1. 不定积分、定积分、变限积分和反常积分的概念与存在性

这是考生比较容易混淆、不好把握的考试内容.

例 1.8.1 试证明如果函数 $f(x)$ 在 $[a,b]$ 上连续，则函数 $F(x)=\int_a^x f(t)\mathrm{d}t$ 在 $[a,b]$ 上可导，且 $F'(x)=f(x)$.

证明 若 $x\in(a,b)$，取 Δx 使 $x+\Delta x\in(a,b)$，则

$$\Delta F=F(x+\Delta x)-F(x)=\int_a^{x+\Delta x}f(t)\mathrm{d}t-\int_a^x f(t)\mathrm{d}t$$
$$=\int_a^x f(t)\mathrm{d}t+\int_x^{x+\Delta x}f(t)\mathrm{d}t-\int_a^x f(t)\mathrm{d}t=\int_x^{x+\Delta x}f(t)\mathrm{d}t,$$

使用积分中值定理，有 $\int_x^{x+\Delta x}f(t)\mathrm{d}t=f(\xi)\Delta x$，其中 ξ 介于 x 与 $x+\Delta x$ 之间，$\Delta x\to 0$ 时，$\xi\to x$，于是

$$F'(x)=\lim_{\Delta x\to 0}\frac{\Delta F}{\Delta x}=\lim_{\Delta x\to 0}f(\xi)=\lim_{\xi\to x}f(\xi)=f(x).$$

若 $x=a$，取 $\Delta x>0$，则同理可证 $F'_+(x)=f(a)$；若 $x=b$，取 $\Delta x<0$，则同理可证 $F'_-(x)=f(b)$.

【注 1】 本题是一般教材中的定理，在考研中出现了对该定理的考查，希望引导考生重视数学基础.

【注 2】 本题事实上就是证明了原函数（不定积分）存在定理——连续函数 $f(x)$ 必有原函数 $F(x)$.

例 1.8.2 试证明含有第一类间断点、无穷间断点的函数 $f(x)$ 在包含该间断点的区间内必没有原函数 $F(x)$.

证明 设 $F(x)$ 为 $f(x)$ 在 I 内的一个原函数，则 $F(x)$ 在 I 内可导，且 $F'(x)=f(x)$，并设 $x=x_0\in I$ 为 $F'(x)$ 的间断点，我们讨论如下三种情况：

①$x=x_0$ 为可去间断点，即 $\lim\limits_{x\to x_0}F'(x)$ 存在且为 A，但 $A\neq F'(x_0)$，而

$$F'(x_0)=\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0}\xlongequal{\text{洛必达法则}}\lim_{x\to x_0}F'(x)=A,$$

矛盾；

②$x=x_0$ 为跳跃间断点，即 $\lim\limits_{x\to x_0^+}F'(x)$ 存在且为 A_+，$\lim\limits_{x\to x_0^-}F'(x)$ 存在且为 A_-，但 $A_+\neq A_-$，而由

$$F'_+(x_0)=\lim_{x\to x_0^+}\frac{F(x)-F(x_0)}{x-x_0}\xlongequal{\text{洛必达法则}}\lim_{x\to x_0^+}F'(x)=A_+,$$

$$F'_-(x_0)=\lim_{x\to x_0^-}\frac{F(x)-F(x_0)}{x-x_0}\xlongequal{\text{洛必达法则}}\lim_{x\to x_0^-}F'(x)=A_-,$$

又 $F'(x_0)$是存在的，则 $F'_+(x_0)=F'_-(x_0)$，即 $A_+=A_-$，矛盾；

③$x=x_0$ 为无穷间断点，即$\lim\limits_{x\to x_0}F'(x)=\infty$，而

$$F'(x_0)=\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0}\xlongequal{\text{洛必达法则}}\lim_{x\to x_0}F'(x)=\infty,$$

又 $F'(x_0)$是存在的，矛盾.

综上所述，①，②和③均不存在原函数，即导函数 $F'(x)$在 I 内必定没有第一类间断点和无穷间断点，也即含有第一类间断点和无穷间断点的函数 $f(x)$在包含该间断点的区间内没有原函数 $F(x)$.

【注 1】 含有振荡间断点的函数是否有原函数呢？举例说来，对于

$$f(x)=\begin{cases}2x\sin\dfrac{1}{x}-\cos\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

其在$(-\infty,+\infty)$上不连续，它有一个振荡间断点 $x=0$，但是它在$(-\infty,+\infty)$上存在原函数

$$F(x)=\begin{cases}x^2\sin\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

即对于$(-\infty,+\infty)$上任一点都有$F'(x)=f(x)$成立.

当然，对于

$$f(x)=\begin{cases}\dfrac{1}{x}\sin\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

其在$(-\infty,+\infty)$上也有一个振荡间断点 $x=0$，且其在$(-\infty,+\infty)$上没有原函数.

【注 2】 综合以上几点，可以得出重要结论：可导函数 $F(x)$求导后的函数 $F'(x)=f(x)$不一定是连续函数，但是如果有间断点，一定是第二类间断点（在考研的范畴内，只能是振荡间断点）.

例 1.8.3 在区间$[-1,2]$上，以下四个结论：

①$f(x)=\begin{cases}2, & x>0,\\ 1, & x=0,\\ -1, & x<0\end{cases}$有原函数，但其定积分不存在；

②$f(x)=\begin{cases}2x\sin\dfrac{1}{x^2}-\dfrac{2}{x}\cos\dfrac{1}{x^2}, & x\neq 0,\\ 0, & x=0\end{cases}$有原函数，其定积分也存在；

③$f(x)=\begin{cases}\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0\end{cases}$没有原函数，其定积分也不存在；

④$f(x)=\begin{cases}2x\cos\dfrac{1}{x}+\sin\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0\end{cases}$有原函数，其定积分也存在.

正确结论的个数为（　　）.

(A)1　　　　(B)2　　　　(C)3　　　　(D)4

解　应选(B).

本题通过具体的例子考查考生是否能够明确区分不定积分与定积分的存在性. 逐个分析即可.

对于 $f(x)=\begin{cases}2, & x>0,\\ 1, & x=0,\\ -1, & x<0,\end{cases}$ 由于 $x=0$ 是其跳跃间断点，根据不定积分存在定理，在任意一个包含 $x=0$ 在其内部的区间 $[a,b]$ 上，$f(x)$ 一定不存在原函数，但由于 $f(x)$ 满足定积分存在定理，故定积分 $\int_a^b f(x)\mathrm{d}x$ 存在，所以 ① 错误.

对于 $f(x)=\begin{cases}2x\sin\dfrac{1}{x^2}-\dfrac{2}{x}\cos\dfrac{1}{x^2}, & x\neq 0,\\ 0, & x=0,\end{cases}$ $x=0$ 是其振荡间断点，但是容易验证：

若 $F(x)=\begin{cases}x^2\sin\dfrac{1}{x^2}, & x\neq 0,\\ 0, & x=0,\end{cases}$ 则 $F'(x)=f(x)(-\infty<x<+\infty)$，所以 $f(x)$ 存在原函数，但在任意一个包含 $x=0$ 在其内部的区间 $[a,b]$ 上，定积分 $\int_a^b f(x)\mathrm{d}x$ 不存在，因为在 $x=0$ 的邻域 $f(x)$ 无界，所以 ② 错误.

对于 $f(x)=\begin{cases}\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$ 由于 $x=0$ 是其无穷间断点，所以 $f(x)$ 在包含 $x=0$ 在内的区间 $[a,b]$ 上总不存在原函数，定积分 $\int_a^b f(x)\mathrm{d}x$ 也不存在，所以 ③ 正确.

对于 $f(x)=\begin{cases}2x\cos\dfrac{1}{x}+\sin\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$ 它在 $(-\infty,+\infty)$ 内存在原函数 $F(x)=\begin{cases}x^2\cos\dfrac{1}{x}, & x\neq 0,\\ 0, & x=0.\end{cases}$ 并且在任意一个包含 $x=0$ 在其内部的区间 $[a,b]$ 上，定积分 $\int_a^b f(x)\mathrm{d}x$ 也存在，因为 $f(x)$ 有界且只有一个振荡间断点，所以 ④ 正确.

综上所述，故答案选择(B).

例 1.8.4　试证明如果函数 $f(x)$ 在 $[a,b]$ 上可积，则 $F(x)=\int_a^x f(t)\mathrm{d}t$ 在 $[a,b]$ 上连续.

证明　任意 $x,x+\Delta x\in[a,b]$（当 $x=a$ 时，$0<\Delta x<b-a$；当 $x=b$ 时，$a-b<\Delta x<0$），则

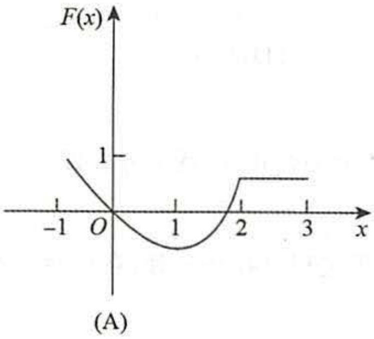
$$F(x+\Delta x)-F(x)=\int_x^{x+\Delta x}f(t)\mathrm{d}t.$$

由可积的必要条件可知，存在 $M>0$，在 $[a,b]$ 上有 $|f(x)|\leqslant M$，所以有

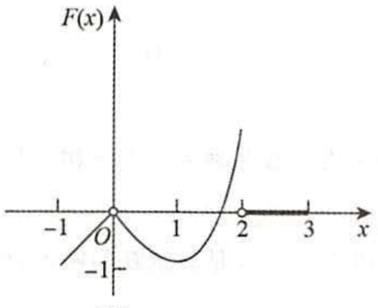
$$0\leqslant|F(x+\Delta x)-F(x)|\leqslant\int_x^{x+|\Delta x|}|f(t)|\mathrm{d}t\leqslant M|\Delta x|,$$

则 $\lim\limits_{\Delta x\to 0}|F(x+\Delta x)-F(x)|=0$，即 $\lim\limits_{\Delta x\to 0}[F(x+\Delta x)-F(x)]=0$ 或 $\lim\limits_{\Delta x\to 0}F(x+\Delta x)=F(x)$，得证.

例 1.8.5　设函数 $y=f(x)$ 在区间 $[-1,3]$ 上的图形如图 1-8-1 所示，则函数 $F(x)=\int_0^x f(t)\mathrm{d}t$ 的图形为（　　）.



(A)



(B)

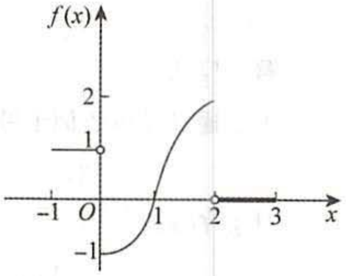
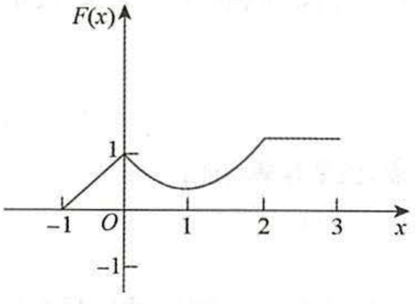
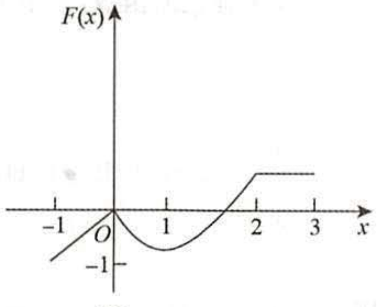


图 1-8-1



(C)



(D)

解　应选(D).

本题有三个要点：第一，由例 1.8.4 可知，变限积分只要存在就必连续，故排除(B)；第二，$f(x)$ 有两个第一类间断点，由例 1.8.2 可知，$F(x)$ 应有两个不可导点，排除(A)；第三，$F(0)=\int_0^0 f(t)\mathrm{d}t=0$，所以 $F(x)$ 的图像过原点，排除(C). 故答案选择(D).

【注】 本题是考研真题，如果考生能够熟练掌握一元函数积分学的有关概念和性质，便可轻松解决这个问题，而无须进行烦琐的计算，从这个角度说，本题是概念题.

2. 积分函数的奇偶性、有界性、单调性、周期性

考生需要能够判别以积分形式定义的函数的奇偶性、有界性、单调性、周期性等，这是重点.

例 1.8.6　证明连续的奇函数的一切原函数都是偶函数；连续的偶函数的原函数中仅有一个原函数是奇函数.

证明　设 $f(x)$ 是连续函数，则其一个原函数可以表示为 $F(x)=\int_a^x f(t)\mathrm{d}t$.

若 $f(x)$ 是连续的奇函数，即有 $f(x)=-f(-x)$，且 $\int_{-a}^{a} f(t)\mathrm{d}t=0$，则

$$F(-x)=\int_a^{-x} f(t)\mathrm{d}t \xlongequal{令\ t=-u} -\int_{-a}^{x} f(-u)\mathrm{d}u=\int_{-a}^{a} f(u)\mathrm{d}u+\int_a^x f(u)\mathrm{d}u=0+F(x)=F(x),$$

所以连续的奇函数的一切原函数都是偶函数.

若 $f(x)$ 是连续的偶函数，即有 $f(-x)=f(x)$，且 $\int_{-a}^{a} f(t)\mathrm{d}t=2\int_0^a f(t)\mathrm{d}t$，则

$$F(-x)=\int_a^{-x} f(t)\mathrm{d}t \xlongequal{令\ t=-u} -\int_{-a}^{x} f(-u)\mathrm{d}u=-\int_{-a}^{a} f(u)\mathrm{d}u-\int_a^x f(u)\mathrm{d}u=-2\int_0^a f(u)\mathrm{d}u-F(x),$$

只有当 $\int_0^a f(u)\mathrm{d}u=0$ 时，$F(-x)=-F(x)$，即连续的偶函数 $f(x)$ 的原函数中仅有一个原函数为奇函数.

【注】 趁热打铁，请读者做一个练习：设函数 $f(x)$ 连续，则在下列函数中，必为偶函数的是(　　).

(A) $\int_0^x f(t^2)\mathrm{d}t$　(B) $\int_0^x f^2(t)\mathrm{d}t$　(C) $\int_0^x t[f(t)-f(-t)]\mathrm{d}t$　(D) $\int_0^x t[f(t)+f(-t)]\mathrm{d}t$

解　应选(D).

直接推导，令 $F(x)=\int_0^x t[f(t)+f(-t)]\mathrm{d}t$，则

$$F(-x)=\int_0^{-x} t[f(t)+f(-t)]\mathrm{d}t \xlongequal{t=-u} \int_0^x (-u)[f(-u)+f(u)](-\mathrm{d}u)$$

$$=\int_0^x u[f(u)+f(-u)]\mathrm{d}u=F(x).$$

也可以根据例 1.8.6 的结论，直接看哪个被积函数是奇函数，显然答案选择(D).

例 1.8.7　证明：若函数 $f(x)$ 是以 T 为周期的连续函数，则对任意的实数 a，都有

$$\int_a^{a+T} f(x)\mathrm{d}x=\int_0^T f(x)\mathrm{d}x.$$

证明　$\int_a^{a+T} f(x)\mathrm{d}x=\int_a^0 f(x)\mathrm{d}x+\int_0^T f(x)\mathrm{d}x+\int_T^{a+T} f(x)\mathrm{d}x.$

设 $t=x-T$，则

$$\int_T^{a+T} f(x)\mathrm{d}x=\int_0^a f(t+T)\mathrm{d}t=\int_0^a f(t)\mathrm{d}t=\int_0^a f(x)\mathrm{d}x,$$

所以

$$\int_a^{a+T} f(x)\mathrm{d}x=\int_a^0 f(x)\mathrm{d}x+\int_0^T f(x)\mathrm{d}x+\int_0^a f(x)\mathrm{d}x=\int_0^T f(x)\mathrm{d}x.$$

例 1.8.8　若连续函数 $f(x)$ 以 T 为周期，证明 $f(x)$ 的一切原函数也以 T 为周期 $\Leftrightarrow \int_0^T f(x)\mathrm{d}x=0$.

证明　“$\Rightarrow$”已知连续函数 $f(x)$ 及任一原函数 $F(x)$ 都以 T 为周期，则有

$$f(x+T)=f(x),\quad F(x+T)=F(x).$$

由牛顿-莱布尼茨公式得

$$\int_0^T f(x)\mathrm{d}x=F(x)\Big|_0^T=F(T)-F(0)=0.$$

“$\Leftarrow$”由例 1.8.7 的结论，有

$$\int_a^{a+T} f(x)\mathrm{d}x=\int_0^T f(x)\mathrm{d}x.$$

由于 $\int_0^T f(x)\mathrm{d}x=0$，则

$$\int_a^{a+T} f(x)\mathrm{d}x=F(x)\Big|_a^{a+T}=F(a+T)-F(a)=0.$$

故 $F(x)$ 是周期为 T 的函数，又 $\int f(x)\mathrm{d}x=F(x)+C$，故 $f(x)$ 的一切原函数均以 T 为周期.

例 1.8.9　设 $f(x)$ 是 $[a,b]$ 上非负的连续函数，且 $f(x)$ 不恒等于零，证明必有 $\int_a^b f(x)\mathrm{d}x>0$.

证明　因函数 $f(x)$ 在 $[a,b]$ 上不恒等于零，且非负，故至少存在一点 $x_0\in(a,b)$，使得 $f(x_0)\neq 0$，即 $f(x_0)>0$.

因函数 $f(x)$ 连续，故有 $\lim\limits_{x\to x_0} f(x)=f(x_0)>0$，由极限的保号性知，存在 $\delta>0$ 与 $\eta>0$，使得当 $x\in[x_0-\delta,x_0+\delta]$ 时恒有 $f(x)\geqslant\eta>0$.

根据定积分的不等式性质，便有 $\int_a^b f(x)\mathrm{d}x \geqslant \int_{x_0-\delta}^{x_0+\delta} f(x)\mathrm{d}x \geqslant \eta\int_{x_0-\delta}^{x_0+\delta}\mathrm{d}x = 2\eta\delta>0$.

【注】 作为该命题的推论，若连续函数 $f(x)$，$g(x)$ 满足 $f(x)\geqslant g(x)$，且 $f(x)$ 不恒等于 $g(x)$，又 $a<b$，则必有严格不等式

$$\int_a^b f(x)\mathrm{d}x > \int_a^b g(x)\mathrm{d}x.$$

二、不定积分的基本计算

考研数学基本都是考常规计算，不涉及太强的技巧性，也就是说题目的门槛不高，考生都容易上手，所以，一定要加强基本计算的训练，如果基础的都算不对，就太可惜了.

例 1.8.10 求 $\int \frac{1}{\sqrt{x(1-x)}}\mathrm{d}x$.

解

$$\int \frac{1}{\sqrt{x(1-x)}}\mathrm{d}x = \int \frac{1}{\sqrt{1-(\sqrt{x})^2}}\cdot\frac{1}{\sqrt{x}}\mathrm{d}x$$

$$= 2\int \frac{1}{\sqrt{1-(\sqrt{x})^2}}\mathrm{d}(\sqrt{x}) = 2\arcsin\sqrt{x}+C.$$

例 1.8.11 求 $\int \frac{1}{x(1+\ln x)}\mathrm{d}x$.

解 $\int \frac{1}{x(1+\ln x)}\mathrm{d}x = \int \frac{\mathrm{d}(\ln x)}{1+\ln x} = \int \frac{\mathrm{d}(1+\ln x)}{1+\ln x} = \ln|1+\ln x|+C.$

例 1.8.12 求 $\int \sqrt{a^2-x^2}\mathrm{d}x\,(a>0)$.

解 设 $x=a\sin t$，则

$$\mathrm{d}x = a\cos t\mathrm{d}t,\ t=\arcsin\frac{x}{a},\ -a\leqslant x\leqslant a,\ -\frac{\pi}{2}\leqslant t\leqslant\frac{\pi}{2}.$$

所以

$$\int\sqrt{a^2-x^2}\mathrm{d}x = \int\sqrt{a^2-a^2\sin^2 t}\cdot a\cos t\mathrm{d}t = a^2\int\cos^2 t\mathrm{d}t = \frac{a^2}{2}\int(1+\cos 2t)\mathrm{d}t$$

$$= \frac{a^2}{2}t+\frac{a^2}{4}\sin 2t+C = \frac{a^2}{2}t+\frac{a^2}{2}\sin t\cos t+C$$

$$= \frac{a^2}{2}\arcsin\frac{x}{a}+\frac{x}{2}\sqrt{a^2-x^2}+C.$$

例 1.8.13 求 $\int \frac{\mathrm{d}x}{\sqrt{x^2+a^2}}\,(a>0)$.

解 设 $x=a\tan t$，则 $\mathrm{d}x=a\sec^2 t\mathrm{d}t$，当 $-\frac{\pi}{2}<t<\frac{\pi}{2}$ 时，$x=a\tan t$ 存在反函数.

所以

$$\int\frac{\mathrm{d}x}{\sqrt{x^2+a^2}} = \int\frac{a\sec^2 t\mathrm{d}t}{a\sqrt{\tan^2 t+1}} = \int\frac{\sec^2 t}{\sec t}\mathrm{d}t = \int\sec t\mathrm{d}t = \ln|\sec t+\tan t|+C_1.$$

由 $\tan t=\frac{x}{a}$ 得 $\sec t=\frac{\sqrt{a^2+x^2}}{a}$，所以

$$\int\frac{\mathrm{d}x}{\sqrt{x^2+a^2}}=\ln|\sec t+\tan t|+C_1=\ln\left(\frac{\sqrt{x^2+a^2}}{a}+\frac{x}{a}\right)+C_1$$

$$=\ln(\sqrt{x^2+a^2}+x)+C(a>0).$$

例 1.8.14　求 $\int\frac{\mathrm{d}x}{\sqrt{x^2-a^2}}(a>0)$.

解　设 $x=a\sec t$，有 $\mathrm{d}x=a\sec t\tan t\mathrm{d}t$. 当 $0<t<\frac{\pi}{2}$ 时，$x=a\sec t>0$，且存在反函数.

$$\int\frac{\mathrm{d}x}{\sqrt{x^2-a^2}}=\int\frac{a\sec t\tan t}{a\tan t}\mathrm{d}t=\int\sec t\mathrm{d}t=\ln(\sec t+\tan t)+C_1=\ln(x+\sqrt{x^2-a^2})+C.$$

当 $\frac{\pi}{2}<t<\pi$ 时，$x=a\sec t<0$，且存在反函数.

$$\int\frac{\mathrm{d}x}{\sqrt{x^2-a^2}}=-\int\sec t\mathrm{d}t=-\ln(-\sec t-\tan t)+C_2$$

$$=-\ln(-x+\sqrt{x^2-a^2})+C_3$$

$$=\ln(-x-\sqrt{x^2-a^2})+C.$$

综上 $\int\frac{\mathrm{d}x}{\sqrt{x^2-a^2}}=\ln|x+\sqrt{x^2-a^2}|+C.$

例 1.8.15　求 $\int\frac{\mathrm{d}x}{\sin 2x+2\sin x}$.

解　本题主要考查三角函数的恒等变形、凑微分法、换元法等.

方法一　原式 $=\int\frac{\mathrm{d}x}{2\sin x(\cos x+1)}=\frac{1}{4}\int\frac{\mathrm{d}\left(\frac{x}{2}\right)}{\sin\frac{x}{2}\cos^3\frac{x}{2}}=\frac{1}{4}\int\frac{\mathrm{d}\left(\tan\frac{x}{2}\right)}{\tan\frac{x}{2}\cos^2\frac{x}{2}}$

$$=\frac{1}{4}\int\frac{1+\tan^2\frac{x}{2}}{\tan\frac{x}{2}}\mathrm{d}\left(\tan\frac{x}{2}\right)=\frac{1}{8}\tan^2\frac{x}{2}+\frac{1}{4}\ln\left|\tan\frac{x}{2}\right|+C.$$

方法二　原式 $=\int\frac{\mathrm{d}x}{2\sin x(\cos x+1)}=\int\frac{\sin x\mathrm{d}x}{2(1-\cos^2x)(1+\cos x)}\xlongequal{\cos x=u}-\frac{1}{2}\int\frac{\mathrm{d}u}{(1-u)(1+u)^2}$

$$\xlongequal{(*)}-\frac{1}{8}\int\left[\frac{1}{1-u}+\frac{3+u}{(1+u)^2}\right]\mathrm{d}u=\frac{1}{8}\left(\ln|1-u|-\ln|1+u|+\frac{2}{1+u}\right)+C$$

$$=\frac{1}{8}\ln\frac{1-\cos x}{1+\cos x}+\frac{1}{4(1+\cos x)}+C.$$

【注】　(1)(＊)处的处理方法叫有理函数积分法，可参考例 1.8.21、例 1.8.22.

(2)由于使用方法不同，得到的不定积分的答案形式可能不唯一. 相同的，还有下面的例1.8.16.

例 1.8.16　求 $\int\frac{\mathrm{d}x}{1+\sin x}$.

解　本题主要考查三角函数的恒等变形、凑微分法.

方法一　原式 $=\int\frac{1-\sin x}{\cos^2x}\mathrm{d}x=\tan x-\frac{1}{\cos x}+C.$

方法二 原式$=\displaystyle\int\frac{\mathrm{d}x}{\left(\cos\frac{x}{2}+\sin\frac{x}{2}\right)^2}=\int\frac{\sec^2\frac{x}{2}}{\left(1+\tan\frac{x}{2}\right)^2}\mathrm{d}x=2\int\frac{\mathrm{d}\left(1+\tan\frac{x}{2}\right)}{\left(1+\tan\frac{x}{2}\right)^2}=-\frac{2}{1+\tan\frac{x}{2}}+C.$

例 1.8.17 求$\displaystyle\int(\arcsin x)^2\mathrm{d}x.$

解 本题主要考查换元法、分部积分法.

方法一

$$\begin{aligned}\int(\arcsin x)^2\mathrm{d}x&=x(\arcsin x)^2-\int\frac{2x\arcsin x}{\sqrt{1-x^2}}\mathrm{d}x\\&=x(\arcsin x)^2+2\int\arcsin x\mathrm{d}(\sqrt{1-x^2})\\&=x(\arcsin x)^2+2\sqrt{1-x^2}\arcsin x-2x+C.\end{aligned}$$

方法二 令$u=\arcsin x$，则$x=\sin u$，$\mathrm{d}x=\cos u\mathrm{d}u$，于是

$$\begin{aligned}\int(\arcsin x)^2\mathrm{d}x&=\int u^2\cos u\mathrm{d}u=\int u^2\mathrm{d}(\sin u)=u^2\sin u-\int 2u\sin u\mathrm{d}u\\&=u^2\sin u+2u\cos u-2\sin u+C\\&=x(\arcsin x)^2+2\sqrt{1-x^2}\arcsin x-2x+C.\end{aligned}$$

例 1.8.18 求$\displaystyle\int\frac{xe^x}{\sqrt{e^x-1}}\mathrm{d}x.$

解 本题主要考查换元法.

令$u=\sqrt{e^x-1}$，则$x=\ln(1+u^2)$，$\mathrm{d}x=\dfrac{2u}{1+u^2}\mathrm{d}u$，从而

$$\begin{aligned}\int\frac{xe^x}{\sqrt{e^x-1}}\mathrm{d}x&=\int\frac{(1+u^2)\ln(1+u^2)}{u}\cdot\frac{2u}{1+u^2}\mathrm{d}u=2\int\ln(1+u^2)\mathrm{d}u\\&=2u\ln(1+u^2)-\int\frac{4u^2}{1+u^2}\mathrm{d}u=2u\ln(1+u^2)-4u+4\arctan u+C\\&=2x\sqrt{e^x-1}-4\sqrt{e^x-1}+4\arctan\sqrt{e^x-1}+C.\end{aligned}$$

例 1.8.19 求$\displaystyle\int\frac{xe^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x.$

解 本题涉及换元法、凑微分法和分部积分法.

方法一 设$x=\tan t$，则

$$\int\frac{xe^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x=\int\frac{e^t\tan t}{(1+\tan^2t)^{\frac{3}{2}}}\sec^2t\mathrm{d}t=\int e^t\sin t\mathrm{d}t,$$

又

$$\begin{aligned}\int e^t\sin t\mathrm{d}t&=-\int e^t\mathrm{d}(\cos t)=-\left(e^t\cos t-\int e^t\cos t\mathrm{d}t\right)\\&=-e^t\cos t+e^t\sin t-\int e^t\sin t\mathrm{d}t,\end{aligned}$$

故原式$=\displaystyle\int e^t\sin t\mathrm{d}t=\frac{1}{2}e^t(\sin t-\cos t)+C=\frac{(x-1)e^{\arctan x}}{2\sqrt{1+x^2}}+C.$

方法二

$$\begin{aligned}\int\frac{xe^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x&=\int\frac{x}{\sqrt{1+x^2}}\mathrm{d}(e^{\arctan x})=\frac{xe^{\arctan x}}{\sqrt{1+x^2}}-\int\frac{e^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x\\&=\frac{xe^{\arctan x}}{\sqrt{1+x^2}}-\int\frac{1}{\sqrt{1+x^2}}\mathrm{d}(e^{\arctan x})=\frac{xe^{\arctan x}}{\sqrt{1+x^2}}-\frac{e^{\arctan x}}{\sqrt{1+x^2}}-\int\frac{xe^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x,\end{aligned}$$

移项整理，得
$$\int \frac{x\mathrm{e}^{\arctan x}}{(1+x^2)^{\frac{3}{2}}}\mathrm{d}x=\frac{(x-1)\mathrm{e}^{\arctan x}}{2\sqrt{1+x^2}}+C.$$

例 1.8.20 求 $\int \frac{\mathrm{d}x}{x^2\sqrt{2x-4}}$.

解 本题是一道综合题，涉及换元法、凑微分法、分部积分法以及有理函数的积分.

方法一 $\int \frac{\mathrm{d}x}{x^2\sqrt{2x-4}}=-\int \frac{\mathrm{d}\left(\frac{1}{x}\right)}{\sqrt{2x-4}}=-\frac{1}{x\sqrt{2x-4}}-\int \frac{1}{x\sqrt{(2x-4)^3}}\mathrm{d}x$,

对上式右边的第二项，令 $\sqrt{2x-4}=t$，则 $x=\frac{t^2+4}{2}$，$\mathrm{d}x=t\mathrm{d}t$，于是

$$\begin{aligned}\int \frac{1}{x\sqrt{(2x-4)^3}}\mathrm{d}x&=\int \frac{2t}{(t^2+4)t^3}\mathrm{d}t=\frac{1}{2}\int \frac{t^2+4-t^2}{(t^2+4)t^2}\mathrm{d}t=\frac{1}{2}\int\left(\frac{1}{t^2}-\frac{1}{t^2+4}\right)\mathrm{d}t\\&=-\frac{1}{2}\left(\frac{1}{t}+\frac{1}{2}\arctan\frac{t}{2}\right)-C\\&=-\frac{1}{2\sqrt{2x-4}}-\frac{1}{4}\arctan\frac{\sqrt{2x-4}}{2}-C,\end{aligned}$$

因此 $\int \frac{\mathrm{d}x}{x^2\sqrt{2x-4}}=\frac{\sqrt{2x-4}}{4x}+\frac{1}{4}\arctan\sqrt{\frac{x-2}{2}}+C.$

方法二 令 $\sqrt{2x-4}=t$，则 $x=\frac{t^2+4}{2}$，$\mathrm{d}x=t\mathrm{d}t$，于是

$$I=\int \frac{\mathrm{d}x}{x^2\sqrt{2x-4}}=\int \frac{4\mathrm{d}t}{(t^2+4)^2},$$

再令 $t=2\tan u$，则 $\mathrm{d}t=2\sec^2 u\mathrm{d}u$，于是

$$\begin{aligned}I&=4\int \frac{2\sec^2 u\mathrm{d}u}{4^2\cdot\sec^4 u}=\frac{1}{2}\int\cos^2 u\mathrm{d}u=\frac{1}{4}\int(1+\cos 2u)\mathrm{d}u=\frac{1}{4}u+\frac{1}{8}\sin 2u+C\\&=\frac{1}{4}\arctan\frac{t}{2}+\frac{1}{4}\cdot\frac{t}{\sqrt{4+t^2}}\cdot\frac{2}{\sqrt{4+t^2}}+C=\frac{1}{4}\arctan\frac{\sqrt{2x-4}}{2}+\frac{\sqrt{2x-4}}{4x}+C,\end{aligned}$$

因此 $\int \frac{\mathrm{d}x}{x^2\sqrt{2x-4}}=\frac{\sqrt{2x-4}}{4x}+\frac{1}{4}\arctan\sqrt{\frac{x-2}{2}}+C.$

例 1.8.21 求 $\int \frac{4x^2-6x-1}{(x+1)(2x-1)^2}\mathrm{d}x$.

解 本题主要考查有理函数的积分.

先将被积函数分解为简单分式之和. 这时应有分解式

$$\frac{4x^2-6x-1}{(x+1)(2x-1)^2}=\frac{A}{x+1}+\frac{B}{2x-1}+\frac{C}{(2x-1)^2},$$

问题是如何定出 A,B,C 这三个数. 右边通分后等号左右的分子应恒等，即

$$4x^2-6x-1\equiv A(2x-1)^2+B(x+1)(2x-1)+C(x+1). \qquad (*)$$

由这恒等式不难定出 A,B,C 来，常用的方法有两种.

一种方法是将右端展开，得到

$$4x^2-6x-1\equiv(4A+2B)x^2+(-4A+B+C)x+(A-B+C),$$

因为这是恒等式，等号左右 x 的同次幂的系数应该相等，故应有

$$\begin{cases}4A+2B=4,\\-4A+B+C=-6,\\A-B+C=-1,\end{cases}$$

解得 $A=1,B=0,C=-2$.

这种方法较死板，且解系数满足的方程组有时较烦琐. 我们希望能求得 A,B,C 应满足的较简单的条件.

另一种方法是根据在恒等式中以变量 x 的任何值代入等号两边应该得到相同的值. 利用这一性质，赋予 x 适当的值，可以得到 A,B,C 应满足的简单条件. 例如，在（＊）式中

令 $x=-1$，有 $9=9A,A=1$；

令 $x=\frac{1}{2}$，有 $-3=\frac{3}{2}C,C=-2$；

令 $x=0$，有 $-1=A-B+C$，可求出 $B=0$.

两种方法求得的结果一致：

$$\frac{4x^2-6x-1}{(x+1)(2x-1)^2}=\frac{1}{x+1}-\frac{2}{(2x-1)^2}.$$

因此可求得

$$\int\frac{4x^2-6x-1}{(x+1)(2x-1)^2}\mathrm{d}x=\int\frac{\mathrm{d}x}{x+1}-\int\frac{2}{(2x-1)^2}\mathrm{d}x$$
$$=\ln|x+1|+\frac{1}{2x-1}+C.$$

【注】 (1) 在上述“另一种方法”中，有读者会问，令 $x=-1$，或令 $x=\frac{1}{2}$，则被积函数 $\frac{4x^2-6x-1}{(x+1)(2x-1)^2}$ 不是无定义了吗？首先，我们给出的恒等式是没有分母的，不涉及 $x=-1,\frac{1}{2}$ 时分母为 0 的问题；其次，请观察第一种方法中，当我们令 x 的同次幂系数相等时，x 的取值不就是任意的吗？记住，我们的最终目的是求出 A,B,C.

(2) 自然，以上两种定系数的方法可以联合起来应用. 见下例.

例 1.8.22 求 $\int\frac{x}{x^3-x^2+x-1}\mathrm{d}x$.

解 本题主要考查有理函数的积分.

因为 $x^3-x^2+x-1=(x-1)(x^2+1)$，设

$$\frac{x}{x^3-x^2+x-1}=\frac{A}{x-1}+\frac{Bx+C}{x^2+1},$$

这时应有

$$x\equiv A(x^2+1)+(Bx+C)(x-1),\qquad(＊)$$

在（＊）式中，令 $x=1$，得 $1=2A,A=\frac{1}{2}$；令 $x=0$，得 $0=A-C,C=\frac{1}{2}$.

比较（＊）式两端 x^2 的系数，有 $0=A+B$,

已求得 $A=\frac{1}{2}$，故有 $B=-\frac{1}{2}$. 于是可得

$$\begin{aligned}\int \frac{x}{x^3-x^2+x-1}\mathrm{d}x &= \frac{1}{2}\int\frac{\mathrm{d}x}{x-1}-\frac{1}{2}\int\frac{x-1}{x^2+1}\mathrm{d}x\\ &= \frac{1}{2}\ln|x-1|-\frac{1}{4}\ln(x^2+1)+\frac{1}{2}\arctan x+C\\ &= \frac{1}{4}\ln\frac{(x-1)^2}{x^2+1}+\frac{1}{2}\arctan x+C.\end{aligned}$$

三、定积分的精确定义

一般说来，考生主要掌握下面这个式子：

$$\int_a^b f(x)\mathrm{d}x=\lim_{n\to\infty}\sum_{i=1}^{n} f\left(a+\frac{b-a}{n}i\right)\frac{b-a}{n},$$

并且将式子中的 a,b 特殊化为 0，1 这两个数，得出的形式最为简单，也最利于解决问题：

$$\int_0^1 f(x)\mathrm{d}x=\lim_{n\to\infty}\sum_{i=1}^{n} f\left(\frac{i}{n}\right)\frac{1}{n}.$$

于是，“凑定积分定义”的步骤如下：

①先提出 $\frac{1}{n}$；②再凑出 $\frac{i}{n}$；③由于 $\frac{i}{n}=0+\frac{1-0}{n}i$，故 $\frac{i}{n}$ 可以读作“0 到 1 上的 x”，且 $\frac{1}{n}=\frac{1-0}{n}$，读作“0 到 1 上的 $\mathrm{d}x$”，于是，“凑定义”完毕. 见例 1.8.24.

例 1.8.23 计算 $\lim\limits_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+n}\right)$.

解
$$\lim_{n\to\infty}\left(\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+n}\right)=\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{n+i},$$

如果对和式 $\sum\limits_{i=1}^{n}\frac{1}{n+i}$ 作放缩

$$n\cdot\frac{1}{n+n}\leqslant\sum_{i=1}^{n}\frac{1}{n+i}\leqslant n\cdot\frac{1}{n+1},$$

则由于 $\lim\limits_{n\to\infty}\frac{n}{n+n}=\frac{1}{2}$，$\lim\limits_{n\to\infty}\frac{n}{n+1}=1$，夹逼准则失效. 此时，我们改用定积分定义，即

$$\text{原式}=\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{n+i}=\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{1+\frac{i}{n}}\cdot\frac{1}{n}=\int_0^1\frac{1}{1+x}\mathrm{d}x=\ln 2.$$

例 1.8.24 计算 $\lim\limits_{n\to\infty}\left(\frac{n+1}{n^2+1}+\frac{n+2}{n^2+4}+\frac{n+3}{n^2+9}+\cdots+\frac{n+n}{n^2+n^2}\right)$.

解 $I=\lim\limits_{n\to\infty}\left(\frac{n+1}{n^2+1}+\frac{n+2}{n^2+4}+\frac{n+3}{n^2+9}+\cdots+\frac{n+n}{n^2+n^2}\right)=\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\frac{n+i}{n^2+i^2}$

$$\overset{①}{=\!=}\lim_{n\to\infty}\sum_{i=1}^{n}\frac{n^2+ni}{n^2+i^2}\cdot\frac{1}{n}\overset{②}{=\!=}\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1+\frac{i}{n}}{1+\left(\frac{i}{n}\right)^2}\cdot\frac{1}{n}\overset{③}{=\!=}\int_0^1\frac{1+x}{1+x^2}\mathrm{d}x=\frac{\pi}{4}+\frac{1}{2}\ln 2.$$

其中①，②，③分别对应“凑定积分定义”的三步.

四、定积分的计算

定积分的计算，除了要用到不定积分的基本积分法，还有自己的特色，这部分内容极其丰富，需要考生狠下功夫.

例 1.8.25 求 $\int_{-1}^{1} x^2\sqrt{1-x^2}\,\mathrm{d}x$.

解 因被积函数为偶函数，故

$$\int_{-1}^{1} x^2\sqrt{1-x^2}\,\mathrm{d}x = 2\int_{0}^{1} x^2\sqrt{1-x^2}\,\mathrm{d}x.$$

再作三角变换，令 $x=\sin t$，当 $x=0$ 时可取 $t=0$；当 $x=1$ 时可取 $t=\frac{\pi}{2}$，且当 $t\in\left[0,\frac{\pi}{2}\right]$时，$x=\sin t$ 不超出原上、下限的区间$[0,1]$，在$\left[0,\frac{\pi}{2}\right]$上，$x'=x'(t)=\cos t$ 连续. 于是

$$\begin{aligned}\text{原式} &= 2\int_{0}^{1} x^2\sqrt{1-x^2}\,\mathrm{d}x = 2\int_{0}^{\frac{\pi}{2}} \sin^2 t\cos^2 t\,\mathrm{d}t \\ &= 2\left(\int_{0}^{\frac{\pi}{2}} \sin^2 t\,\mathrm{d}t - \int_{0}^{\frac{\pi}{2}} \sin^4 t\,\mathrm{d}t\right) \\ &\overset{(*)}{=\!=} 2\left(\frac{1}{2}\times\frac{\pi}{2}-\frac{3}{4}\times\frac{1}{2}\times\frac{\pi}{2}\right)=\frac{\pi}{8}.\end{aligned}$$

【注】 (1)(*)处来自例 1.8.33 的结果.

(2)从原则上讲，作变换 $x=\sin t$ 后，当 $x=0$ 时，可取 $t=0,\pm\pi,\pm2\pi,\cdots$；当 $x=1$ 时，可取 $t=\frac{\pi}{2},\frac{\pi}{2}\pm2\pi,\cdots$，上、下限有多种组合可满足定理条件. 但被积函数中 $\sqrt{1-x^2}=\sqrt{1-\sin^2 t}=|\cos t|$，在不同组合中，此绝对值的处理有简有繁，应引起注意，不要自找麻烦.

例 1.8.26 求 $\int_{1}^{4} \frac{1}{1+\sqrt{x}}\mathrm{d}x$.

解 令$\sqrt{x}=t$，有 $x=t^2$，$\mathrm{d}x=2t\mathrm{d}t$. 当 $x=1$ 时，$t=1$；当 $x=4$ 时，$t=2$. 于是

$$\begin{aligned}\int_{1}^{4} \frac{1}{1+\sqrt{x}}\mathrm{d}x &= \int_{1}^{2} \frac{2t}{1+t}\mathrm{d}t = 2\int_{1}^{2}\left(1-\frac{1}{1+t}\right)\mathrm{d}t \\ &= 2(t-\ln|1+t|)\Big|_{1}^{2} = 2(1-\ln 3+\ln 2).\end{aligned}$$

【注】 若一开始就作变换 $x=t^2$，则当 $x=1$ 时，可取 $t=1$ 或 -1；当 $x=4$ 时，可取 $t=2$ 或 -2. 于是对于 t 的积分区间有 4 种取法：$\int_{1}^{2}$，$\int_{1}^{-2}$，$\int_{-1}^{2}$，$\int_{-1}^{-2}$，其中第 1 种已经用过，第 4 种为

$$\int_{1}^{4} \frac{1}{1+\sqrt{x}}\mathrm{d}x = \int_{-1}^{-2} \frac{2t}{1+|t|}\mathrm{d}t = \int_{-1}^{-2} \frac{2t}{1-t}\mathrm{d}t = 2(1-\ln 3+\ln 2).$$

而第 2 种、第 3 种都超出了原上、下限范围. 例如第 2 种，当 t 在 1 与 -2 之间变化时，$x=t^2$ 的变化范围超出原上、下限的范围$[1,4]$，但计算结果仍然正确. 不过第 2,3,4 种方法，都要处理$|t|$，会比较麻烦.

例 1.8.27　求$\int_0^{2\pi}\frac{1}{1+\cos^2 x}\mathrm{d}x$.

分析　此题若立刻作变换 $\tan x=t$ 或 $\tan\frac{x}{2}=t$，则在 $0\leqslant x\leqslant 2\pi$ 上，不能确定出单值连续的反函数 $x=\varphi(t)$（例如对于变换 $\tan x=t$，反函数 $x=\arctan t$ 的值域为$\left[0,\frac{\pi}{2}\right)$）. 有两个办法来解决：

①将区间 $0\leqslant x\leqslant 2\pi$ 划分成若干个小区间，使在每个小区间上都可以确定出单值连续的反函数；

②利用周期性和奇偶性简化所求积分，使得在简化之后的小区间上，作变换时可确定出单值连续的反函数.

这两条思路实际上是一样的，今用②的办法解之.

解

$$\begin{aligned}\int_0^{2\pi}\frac{1}{1+\cos^2 x}\mathrm{d}x&=\int_{-\pi}^{\pi}\frac{1}{1+\cos^2 x}\mathrm{d}x=2\int_0^{\pi}\frac{1}{1+\cos^2 x}\mathrm{d}x\\&=2\left(\int_0^{\frac{\pi}{2}}\frac{1}{1+\cos^2 x}\mathrm{d}x+\int_{\frac{\pi}{2}}^{\pi}\frac{1}{1+\cos^2 x}\mathrm{d}x\right),\end{aligned}$$

上式右端第二项积分作变换 $x=\pi-t$，即可化成第一项. 于是

$$\text{原式}=4\int_0^{\frac{\pi}{2}}\frac{1}{1+\cos^2 x}\mathrm{d}x,$$

这样就可作变换 $\tan x=t$ 了. 于是有

$$x=\arctan t,\quad \cos^2 x=\frac{1}{1+t^2},\quad \mathrm{d}x=\frac{1}{1+t^2}\mathrm{d}t.$$

当 $x=0$ 时，$t=0$；当 $x=\frac{\pi}{2}$时，$t=+\infty$. 于是

$$\text{原式}=4\int_0^{+\infty}\frac{1}{2+t^2}\mathrm{d}t=\frac{4}{\sqrt{2}}\arctan\frac{t}{\sqrt{2}}\bigg|_0^{+\infty}=\sqrt{2}\pi.$$

【注】　也有读者这样做：

$$\begin{aligned}\int_0^{2\pi}\frac{\mathrm{d}x}{1+\cos^2 x}&=\int_0^{2\pi}\frac{\mathrm{d}x}{(\sec^2 x+1)\cos^2 x}=\int_0^{2\pi}\frac{\mathrm{d}(\tan x)}{\tan^2 x+2}\\&=\frac{1}{\sqrt{2}}\arctan\frac{\tan x}{\sqrt{2}}\bigg|_0^{2\pi}=0.\end{aligned}$$

显然这是错误的，因为 $\tan x$ 在$[0,2\pi]$上有 $x=\frac{\pi}{2},\frac{3\pi}{2}$两个无穷间断点，故牛顿-莱布尼茨公式是不能用的.

例 1.8.28　求$\int_{-1}^{1}\frac{x+1}{1+\sqrt[3]{x^2}}\mathrm{d}x$.

解　如果作变换$\sqrt[3]{x^2}=t$，则 $x=\pm\sqrt{t^3}$. 在 x 的区间$[-1,1]$上的积分，应分两段$[-1,0]$与$[0,1]$来处理（即相应地，在前一段上取 $x=-\sqrt{t^3}$；在后一段上取 $x=\sqrt{t^3}$）. 如果作变换$\sqrt[3]{x}=t$，则 $x=t^3$. 虽然不必将区间$[-1,1]$分两段处理，但也不是最好的办法. 经仔细审题发现

$$\int_{-1}^{1}\frac{x+1}{1+\sqrt[3]{x^2}}\mathrm{d}x=\int_{-1}^{1}\frac{x}{1+\sqrt[3]{x^2}}\mathrm{d}x+\int_{-1}^{1}\frac{1}{1+\sqrt[3]{x^2}}\mathrm{d}x=0+2\int_0^1\frac{1}{1+\sqrt[3]{x^2}}\mathrm{d}x,$$

其中前一项因为被积函数是奇函数，故在对称区间上的积分应为 0；后一项是因为被积函数在对称区间上是偶函数的缘故.

再作积分变量变换,令 $\sqrt[3]{x^2}=t$,有 $x=\sqrt{t^3}$,$\mathrm{d}x=\frac{3}{2}\sqrt{t}\mathrm{d}t$. 当 $x=0$ 时,$t=0$;当 $x=1$ 时,$t=1$. 于是

$$原式=3\int_0^1\frac{\sqrt{t}}{1+t}\mathrm{d}t\xlongequal{\sqrt{t}=u}6\int_0^1\frac{u^2}{1+u^2}\mathrm{d}u=6-6\arctan 1=6-\frac{3}{2}\pi.$$

【注】 可见仔细审题,利用性质,会给解题带来方便.

例 1.8.29 设 $f(x)$ 为连续函数,证明 $\int_a^b f(x)\mathrm{d}x=\int_a^b f(a+b-x)\mathrm{d}x$.

证明 作变量代换,令 $x=a+b-t$,则

$$\begin{aligned}\int_a^b f(x)\mathrm{d}x&=\int_b^a f(a+b-t)(-\mathrm{d}t)\\&=\int_a^b f(a+b-t)\mathrm{d}t=\int_a^b f(a+b-x)\mathrm{d}x,\end{aligned}$$

证毕.

【注】 此结论一般被叫作"区间再现公式",其证明过程较为简单,但其用处很大,可见例 1.8.30~例 1.8.32.

例 1.8.30 求 $\int_0^\pi x\sin^9 x\mathrm{d}x$.

解 设法消除 x. 使用"区间再现公式",令 $x=\pi-t$,则

$$I=\int_0^\pi x\sin^9 x\mathrm{d}x=\int_\pi^0(\pi-t)\sin^9 t(-\mathrm{d}t)=\int_0^\pi \pi\sin^9 t\mathrm{d}t-\int_0^\pi t\sin^9 t\mathrm{d}t,$$

从而

$$\begin{aligned}I&=\frac{\pi}{2}\int_0^\pi\sin^9 t\mathrm{d}t=\frac{\pi}{2}\left(\int_0^{\frac{\pi}{2}}\sin^9 t\mathrm{d}t+\int_{\frac{\pi}{2}}^\pi\sin^9 t\mathrm{d}t\right)\\&=\frac{\pi}{2}\left[\int_0^{\frac{\pi}{2}}\sin^9 t\mathrm{d}t+\int_{\frac{\pi}{2}}^0\sin^9(\pi-t)(-\mathrm{d}t)\right]\\&=\pi\int_0^{\frac{\pi}{2}}\sin^9 t\mathrm{d}t=\pi\times\frac{8}{9}\times\frac{6}{7}\times\frac{4}{5}\times\frac{2}{3}=\frac{128\pi}{315}.\end{aligned}$$

例 1.8.31 求 $\int_0^{\frac{\pi}{4}}\ln(1+\tan x)\mathrm{d}x$.

解 使用"区间再现公式",令 $x=\frac{\pi}{4}-t$,$\tan x=\frac{1-\tan t}{1+\tan t}$,则

$$\begin{aligned}\int_0^{\frac{\pi}{4}}\ln(1+\tan x)\mathrm{d}x&=\int_{\frac{\pi}{4}}^0\ln\frac{2}{1+\tan t}(-\mathrm{d}t)\\&=-\int_0^{\frac{\pi}{4}}\ln(1+\tan t)\mathrm{d}t+\int_0^{\frac{\pi}{4}}\ln 2\mathrm{d}t,\end{aligned}$$

从而得

$$\int_0^{\frac{\pi}{4}}\ln(1+\tan x)\mathrm{d}x=\frac{\pi}{8}\ln 2.$$

例 1.8.32 求 $\int_0^{\frac{\pi}{2}}\frac{\sin x}{\sin x+\cos x}\mathrm{d}x$.

解 此题是三角有理式的积分,可按该类型积分的标准步骤去做,但比较烦琐. 现令 $x=\frac{\pi}{2}-t$,得

$$I=\int_0^{\frac{\pi}{2}}\frac{\sin x}{\sin x+\cos x}\mathrm{d}x=\int_{\frac{\pi}{2}}^0\frac{\cos t}{\cos t+\sin t}(-\mathrm{d}t)$$

$$=\int_0^{\frac{\pi}{2}}\frac{\cos t}{\cos t+\sin t}\mathrm{d}t,$$

所以

$$2I=\int_0^{\frac{\pi}{2}}\frac{\sin x}{\sin x+\cos x}\mathrm{d}x+\int_0^{\frac{\pi}{2}}\frac{\cos x}{\sin x+\cos x}\mathrm{d}x=\int_0^{\frac{\pi}{2}}1\mathrm{d}x=\frac{\pi}{2},$$

从而 $I=\frac{\pi}{4}$.

例 1.8.33 设 n 是非负整数，证明 $\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x(n=0,1,\cdots)$，并求之.

解 先证 $\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x(n=0,1,\cdots)$.

作变量代换. 令 $x=\frac{\pi}{2}-t$，有

$$\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=\int_{\frac{\pi}{2}}^0\sin^n\left(\frac{\pi}{2}-t\right)(-\mathrm{d}t)=\int_0^{\frac{\pi}{2}}\cos^n t\mathrm{d}t=\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x(n=0,1,\cdots).$$

现用分部积分法计算 $\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x$，将它记为 I_n. 于是

$$I_0=\int_0^{\frac{\pi}{2}}1\mathrm{d}x=\frac{\pi}{2},\quad I_1=\int_0^{\frac{\pi}{2}}\sin x\mathrm{d}x=-\cos x\Big|_0^{\frac{\pi}{2}}=1,$$

当 $n\geqslant 2$ 时，

$$I_n=\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=\int_0^{\frac{\pi}{2}}\sin^{n-1}x\cdot\sin x\mathrm{d}x=-\int_0^{\frac{\pi}{2}}\sin^{n-1}x\mathrm{d}(\cos x)$$

$$=-\sin^{n-1}x\cos x\Big|_0^{\frac{\pi}{2}}+\int_0^{\frac{\pi}{2}}\cos x\mathrm{d}(\sin^{n-1}x)=\int_0^{\frac{\pi}{2}}\cos x\cdot(n-1)\sin^{n-2}x\cdot\cos x\mathrm{d}x$$

$$=(n-1)\int_0^{\frac{\pi}{2}}\cos^2 x\sin^{n-2}x\mathrm{d}x=(n-1)\int_0^{\frac{\pi}{2}}(1-\sin^2 x)\sin^{n-2}x\mathrm{d}x$$

$$=(n-1)\int_0^{\frac{\pi}{2}}\sin^{n-2}x\mathrm{d}x-(n-1)\int_0^{\frac{\pi}{2}}\sin^n x\mathrm{d}x=(n-1)I_{n-2}-(n-1)I_n,$$

于是

$$nI_n=(n-1)I_{n-2},\quad I_n=\frac{n-1}{n}I_{n-2}(n=2,3,\cdots).$$

按此公式递推下去有

$$I_n=\begin{cases}\dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{2}{3}\cdot I_1,\\[2ex] \dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{1}{2}\cdot I_0\end{cases}=\begin{cases}\dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{2}{3}\cdot 1,\ n\text{ 为大于 1 的奇数},\\[2ex] \dfrac{n-1}{n}\cdot\dfrac{n-3}{n-2}\cdot\cdots\cdot\dfrac{1}{2}\cdot\dfrac{\pi}{2},\ n\text{ 为正偶数}.\end{cases}$$

上式叫华里士公式.

【注】 利用华里士公式可快速计算某些特殊的定积分，如

$$\int_0^{\frac{\pi}{2}}\sin^8 x\mathrm{d}x=\frac{7}{8}\cdot\frac{5}{6}\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot\frac{\pi}{2}=\frac{35\pi}{256},$$

$$\int_0^{\frac{\pi}{2}}\sin^9 x\mathrm{d}x=\frac{8}{9}\cdot\frac{6}{7}\cdot\frac{4}{5}\cdot\frac{2}{3}\cdot 1=\frac{128}{315}.$$

例 1.8.34 设 n 为正整数,求 $\int_0^{2\pi}\sin^n x\,\mathrm{d}x$.

解 $\sin^n x$ 是以 2π 为周期的周期函数,于是

$$\int_0^{2\pi}\sin^n x\,\mathrm{d}x=\int_{-\pi}^{\pi}\sin^n x\,\mathrm{d}x.$$

若 n 为正奇数,则 $\sin^n x$ 为奇函数;若 n 为正偶数,则 $\sin^n x$ 为偶函数.从而

$$\int_0^{2\pi}\sin^n x\,\mathrm{d}x=\begin{cases}0, & n\text{ 为正奇数},\\ 2\int_0^{\pi}\sin^n x\,\mathrm{d}x, & n\text{ 为正偶数}.\end{cases}$$

当 n 为正偶数时,有

$$\begin{aligned}2\int_0^{\pi}\sin^n x\,\mathrm{d}x&=2\left(\int_0^{\frac{\pi}{2}}\sin^n x\,\mathrm{d}x+\int_{\frac{\pi}{2}}^{\pi}\sin^n x\,\mathrm{d}x\right)(\text{第二个积分令 }x=\pi-t)\\&=2\left[\int_0^{\frac{\pi}{2}}\sin^n x\,\mathrm{d}x+\int_{\frac{\pi}{2}}^{0}\sin^n(\pi-t)(-\mathrm{d}t)\right]\\&=2\left(\int_0^{\frac{\pi}{2}}\sin^n x\,\mathrm{d}x+\int_0^{\frac{\pi}{2}}\sin^n t\,\mathrm{d}t\right)=4\int_0^{\frac{\pi}{2}}\sin^n x\,\mathrm{d}x\\&=4\cdot\frac{n-1}{n}\cdot\frac{n-3}{n-2}\cdot\cdots\cdot\frac{1}{2}\cdot\frac{\pi}{2},\end{aligned}$$

即

$$\int_0^{2\pi}\sin^n x\,\mathrm{d}x=\begin{cases}0, & n\text{ 为正奇数},\\ 4\cdot\frac{n-1}{n}\cdot\frac{n-3}{n-2}\cdot\cdots\cdot\frac{1}{2}\cdot\frac{\pi}{2}, & n\text{ 为正偶数}.\end{cases}$$

例 1.8.35 设 n 为正整数,求 $\int_0^{2\pi}\cos^n x\,\mathrm{d}x$.

解 $\cos^n x$ 是以 2π 为周期的周期函数,于是

$$\begin{aligned}\int_0^{2\pi}\cos^n x\,\mathrm{d}x&=\int_{-\pi}^{\pi}\cos^n x\,\mathrm{d}x=2\int_0^{\pi}\cos^n x\,\mathrm{d}x\\&=2\left(\int_0^{\frac{\pi}{2}}\cos^n x\,\mathrm{d}x+\int_{\frac{\pi}{2}}^{\pi}\cos^n x\,\mathrm{d}x\right)(\text{第二个积分令 }x=\pi-t)\\&=2\left[\int_0^{\frac{\pi}{2}}\cos^n x\,\mathrm{d}x+\int_{\frac{\pi}{2}}^{0}\cos^n(\pi-t)(-\mathrm{d}t)\right]\\&=2\left[\int_0^{\frac{\pi}{2}}\cos^n x\,\mathrm{d}x+\int_0^{\frac{\pi}{2}}(-\cos t)^n\mathrm{d}t\right].\end{aligned}$$

当 n 为正奇数时,$(-\cos t)^n=-\cos^n t$,于是原积分$=0$;

当 n 为正偶数时,原积分$=4\int_0^{\frac{\pi}{2}}\cos^n x\,\mathrm{d}x$,由例 1.8.33 及例 1.8.34,有

$$\int_0^{2\pi}\cos^n x\,\mathrm{d}x=\int_0^{2\pi}\sin^n x\,\mathrm{d}x=\begin{cases}0, & n\text{ 为正奇数},\\ 4\cdot\frac{n-1}{n}\cdot\frac{n-3}{n-2}\cdot\cdots\cdot\frac{1}{2}\cdot\frac{\pi}{2}, & n\text{ 为正偶数}.\end{cases}$$

【注】 上面三个例子的结论,读者应牢记,这对考研很有帮助,请看下面这个例子.

例 1.8.36 求 $\int_0^6 x^2\sqrt{6x-x^2}\,\mathrm{d}x$.

解 先配方,即 $\sqrt{6x-x^2}=\sqrt{9-(x^2-6x+9)}=\sqrt{3^2-(x-3)^2}$,

再令 $x-3=3\sin t$,有

$$\mathrm{d}x=3\cos t\mathrm{d}t,\quad \sqrt{6x-x^2}=\sqrt{9(1-\sin^2 t)}=3|\cos t|,$$

则

$$\begin{aligned}\int_0^6 x^2\sqrt{6x-x^2}\,\mathrm{d}x&=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(3+3\sin t)^2\cdot 3|\cos t|\cdot 3\cos t\mathrm{d}t\\&=81\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(1+2\sin t+\sin^2 t)\cos^2 t\mathrm{d}t\\&=81\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^2 t\mathrm{d}t+162\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin t\cos^2 t\mathrm{d}t+81\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(1-\cos^2 t)\cos^2 t\mathrm{d}t\\&=162\int_0^{\frac{\pi}{2}}\cos^2 t\mathrm{d}t+0+162\int_0^{\frac{\pi}{2}}\cos^2 t\mathrm{d}t-162\int_0^{\frac{\pi}{2}}\cos^4 t\mathrm{d}t\\&=324\cdot\frac{1}{2}\cdot\frac{\pi}{2}-162\cdot\frac{3}{4}\cdot\frac{1}{2}\cdot\frac{\pi}{2}=\frac{405}{8}\pi.\end{aligned}$$

例 1.8.37 设 $f''(u)$在区间$[0,2]$上连续，$f(2)=a,f'(2)=b,\int_0^2 f(u)\mathrm{d}u=c$. 求$\int_0^1 x^2 f''(2x)\mathrm{d}x$.

分析 由于并不知道 $f''(u)$是什么，仅知道 $f(2)=a,f'(2)=b,\int_0^2 f(u)\mathrm{d}u=c$，故宜用分部积分法以消除 $f''(u)$.

解
$$\begin{aligned}\int_0^1 x^2 f''(2x)\mathrm{d}x&=\frac{1}{2}\int_0^1 x^2\mathrm{d}[f'(2x)]=\frac{1}{2}[x^2 f'(2x)]\Big|_0^1-\int_0^1 xf'(2x)\mathrm{d}x\\&=\frac{1}{2}f'(2)-\frac{1}{2}\int_0^1 x\mathrm{d}[f(2x)]=\frac{1}{2}f'(2)-\frac{1}{2}\left[xf(2x)\Big|_0^1-\int_0^1 f(2x)\mathrm{d}x\right]\\&=\frac{1}{2}f'(2)-\frac{1}{2}f(2)+\frac{1}{4}\int_0^2 f(u)\mathrm{d}u=-\frac{1}{2}a+\frac{1}{2}b+\frac{1}{4}c\\&=\frac{1}{4}(c+2b-2a).\end{aligned}$$

例 1.8.38 设 $G(x)=\int_0^x\frac{\sin t}{\pi-t}\mathrm{d}t$，求$\int_0^\pi G(x)\mathrm{d}x$.

分析 由于并不知道 $G(x)$，仅知道 $G'(x)$，故宜用分部积分法以消除 $G(x)$.

解
$$\begin{aligned}\int_0^\pi G(x)\mathrm{d}x&=xG(x)\Big|_0^\pi-\int_0^\pi x\mathrm{d}[G(x)]=\pi G(\pi)-\int_0^\pi x\cdot\frac{\sin x}{\pi-x}\mathrm{d}x\\&=\pi\int_0^\pi\frac{\sin x}{\pi-x}\mathrm{d}x-\int_0^\pi x\cdot\frac{\sin x}{\pi-x}\mathrm{d}x\\&=\int_0^\pi\sin x\mathrm{d}x=-\cos x\Big|_0^\pi=2.\end{aligned}$$

五、变限积分

由于变限积分就是一种函数，所以其考题也很丰富、精彩. 需要注意的已经在“基础内容精讲”“一”中的“(三)3. 变限积分的求导公式”处讲述过.

例 1.8.39 设 $F(x)=\int_x^{x+2\pi}\mathrm{e}^{\sin t}\sin t\mathrm{d}t$，则 $F(x)$ (　　).

(A)为正常数　　(B)为负常数　　(C)恒为零　　(D)不为常数

解 应选(A).

因 $e^{\sin x}\sin x$ 是以 2π 为周期的可积函数，所以

$$\int_x^{x+2\pi} e^{\sin t}\sin t\mathrm{d}t=\int_0^{2\pi} e^{\sin t}\sin t\mathrm{d}t=\int_0^{2\pi} e^{\sin t}\mathrm{d}(-\cos t)$$

$$=-e^{\sin t}\cos t\Big|_0^{2\pi}-\int_0^{2\pi}(-\cos t)e^{\sin t}\cos t\mathrm{d}t$$

$$=\int_0^{2\pi} e^{\sin t}\cos^2 t\mathrm{d}t,$$

又 $e^{\sin x}\cos^2 x\geqslant 0$，故选(A).

【注】 若 $f(x)$ 是以 T 为周期的可积函数，则

$$\int_0^T f(x)\mathrm{d}x=\int_a^{a+T} f(x)\mathrm{d}x,$$

a 为任意常数. 这个结论十分重要，请读者牢记.

例 1.8.40 设 $f(x)$ 连续，则 $\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x tf(x^2-t^2)\mathrm{d}t\right]=$ (　　).

(A) $xf(x^2)$　　(B) $-xf(x^2)$　　(C) $2xf(x^2)$　　(D) $-2xf(x^2)$

解 应选(A).

“求导变量”x 出现在了被积函数中，必须通过变量代换 $x^2-t^2=u$，将其移出被积函数，才能使用变限积分求导公式.

令 $x^2-t^2=u$，则 $-2t\mathrm{d}t=\mathrm{d}u$，当 $t=0$ 时，$u=x^2$；当 $t=x$ 时，$u=0$.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^x tf(x^2-t^2)\mathrm{d}t\right]=\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_0^{x^2} f(u)\mathrm{d}u\right]=xf(x^2).$$

故答案选择(A).

例 1.8.41 设函数 $y=y(x)$ 由方程 $x=\int_1^{y+x} e^{-u^2}\mathrm{d}u$ 所确定，求 $y(0)$，$y'(0)$ 和 $y''(0)$.

解 将 $x=0$ 代入 $x=\int_1^{y+x} e^{-u^2}\mathrm{d}u$ 得 $\int_1^y e^{-u^2}\mathrm{d}u=0$，由于 $g(y)=\int_1^y e^{-u^2}\mathrm{d}u$ 是单调增加函数，且 $g(1)=0$，所以 $\int_1^y e^{-u^2}\mathrm{d}u=0$ 有唯一解 $y=1$，即 $y(0)=1$.

在方程 $x=\int_1^{y+x} e^{-u^2}\mathrm{d}u$ 两边同时对 x 求导，得

$$1=e^{-(y+x)^2}(y'+1),\qquad ①$$

以 $x=0$，$y(0)=1$ 代入①式，得到 $y'(0)=e-1$.

在①式两边再同时对 x 求导，得

$$e^{-(y+x)^2}[-2(y+x)(y'+1)^2+y'']=0,\qquad ②$$

以 $x=0$，$y(0)=1$，$y'(0)=e-1$ 代入②式，得到 $y''(0)=2e^2$.

例 1.8.42 设 xOy 平面上有正方形区域 $D=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}$ 及直线 l：$x+y=t(t\geqslant 0)$. 若 $S(t)$ 表示正方形区域 D 位于直线 l 左下方部分的面积，试求 $\int_0^x S(t)\mathrm{d}t(x\geqslant 0)$.

解 由题设知

$$S(t)=\begin{cases}\frac{1}{2}t^2, & 0\leqslant t\leqslant 1,\\ -\frac{1}{2}t^2+2t-1, & 1<t\leqslant 2,\\ 1, & t>2.\end{cases}$$

所以当 $0\leqslant x\leqslant 1$ 时，$\int_0^x S(t)\mathrm{d}t=\int_0^x \frac{1}{2}t^2\mathrm{d}t=\frac{1}{6}x^3$；

当 $1<x\leqslant 2$ 时，$\int_0^x S(t)\mathrm{d}t=\int_0^1 S(t)\mathrm{d}t+\int_1^x S(t)\mathrm{d}t=-\frac{x^3}{6}+x^2-x+\frac{1}{3}$；

当 $x>2$ 时，$\int_0^x S(t)\mathrm{d}t=\int_0^2 S(t)\mathrm{d}t+\int_2^x S(t)\mathrm{d}t=x-1$.

因此
$$\int_0^x S(t)\mathrm{d}t=\begin{cases}\frac{1}{6}x^3, & 0\leqslant x\leqslant 1,\\ -\frac{x^3}{6}+x^2-x+\frac{1}{3}, & 1<x\leqslant 2,\\ x-1, & x>2.\end{cases}$$

六、反常积分的计算与敛散性判别

1. 计算

关于反常积分的计算有两点要讲.

第一，反常积分是变限积分的极限，也就是在一道积分题目的求解过程中，出现无穷区间或者瑕点，直接代入（做极限计算），这里仅注意计算规则即可，不必考虑其敛散性.

第二，如何识别反常积分？只要一看积分限有 ∞，便知这是无穷区间上的反常积分，所以此类反常积分是容易识别的.

无界函数的反常积分就较难识别了.一般是看被积函数是否有使其分母为零的点.但这句话既不必要，也不充分.例如 $\int_0^1 \ln x\mathrm{d}x$，被积函数没有“分母”，而 $x=0$ 是它的瑕点，所以该积分是反常积分.又如 $\int_0^1 \frac{1}{x^2}\mathrm{e}^{-\frac{1}{x}}\mathrm{d}x$ 的下限 $x=0$ 使分母为 0，但却不是反常积分.这是因为

$$\lim_{x\to0^+}\frac{1}{x^2}\mathrm{e}^{-\frac{1}{x}}=\lim_{x\to0^+}\frac{\frac{1}{x^2}}{\mathrm{e}^{\frac{1}{x}}}=\lim_{x\to0^+}\frac{-\frac{2}{x^3}}{-\frac{1}{x^2}\mathrm{e}^{\frac{1}{x}}}=\lim_{x\to0^+}\frac{\frac{2}{x}}{\mathrm{e}^{\frac{1}{x}}}=\lim_{x\to0^+}\frac{-\frac{2}{x^2}}{-\frac{1}{x^2}\mathrm{e}^{\frac{1}{x}}}=\lim_{x\to0^+}2\mathrm{e}^{-\frac{1}{x}}=0,$$

所以 $\int_0^1 \frac{1}{x^2}\mathrm{e}^{-\frac{1}{x}}\mathrm{d}x$ 不是反常积分.

话虽如此，但被积函数的分母为零仍是重要的识别标志.不但要看积分的上、下限，还要看区间内部是否有使 $f(x)\to\infty$ 的点.**不少考题故意将瑕点埋伏在区间内部，致使考生不注意.**

例 1.8.43 计算积分 $\int_{\frac{1}{2}}^{\frac{3}{2}}\frac{\mathrm{d}x}{\sqrt{|x-x^2|}}$.

解 注意到被积函数含有绝对值符号且 $x=1$ 是其无穷间断点，故

$$原式=\int_{\frac{1}{2}}^{1}\frac{\mathrm{d}x}{\sqrt{x-x^2}}+\int_{1}^{\frac{3}{2}}\frac{\mathrm{d}x}{\sqrt{x^2-x}}.$$

而
$$\int_{\frac{1}{2}}^{1}\frac{\mathrm{d}x}{\sqrt{x-x^2}}=\int_{\frac{1}{2}}^{1}\frac{\mathrm{d}x}{\sqrt{\frac{1}{4}-\left(x-\frac{1}{2}\right)^2}}=\arcsin(2x-1)\Big|_{\frac{1}{2}}^{1}=\arcsin 1=\frac{\pi}{2},$$

$$\int_1^{\frac{3}{2}}\frac{\mathrm{d}x}{\sqrt{x^2-x}}=\int_1^{\frac{3}{2}}\frac{\mathrm{d}x}{\sqrt{\left(x-\frac{1}{2}\right)^2-\frac{1}{4}}}=\ln\left[\left(x-\frac{1}{2}\right)+\sqrt{\left(x-\frac{1}{2}\right)^2-\frac{1}{4}}\right]\Bigg|_1^{\frac{3}{2}}$$

$$=\ln(2+\sqrt{3}),$$

因此
$$\int_{\frac{1}{2}}^{\frac{3}{2}}\frac{\mathrm{d}x}{\sqrt{|x-x^2|}}=\frac{\pi}{2}+\ln(2+\sqrt{3}).$$

例 1.8.44 计算 $I=\int_1^{+\infty}\frac{\mathrm{d}x}{\mathrm{e}^{1+x}+\mathrm{e}^{3-x}}$.

解
$$I=\int_1^{+\infty}\frac{\mathrm{e}^{x-3}}{\mathrm{e}^{2(x-1)}+1}\mathrm{d}x=\mathrm{e}^{-2}\int_1^{+\infty}\frac{\mathrm{d}(\mathrm{e}^{x-1})}{1+\mathrm{e}^{2(x-1)}}=\mathrm{e}^{-2}\cdot\arctan\mathrm{e}^{x-1}\Big|_1^{+\infty}$$

$$=\mathrm{e}^{-2}\left(\frac{\pi}{2}-\frac{\pi}{4}\right)=\frac{\pi}{4}\mathrm{e}^{-2}.$$

例 1.8.45 求 $\int_3^{+\infty}\frac{\mathrm{d}x}{(x-1)^4\sqrt{x^2-2x}}$.

解 原式 $=\int_3^{+\infty}\frac{\mathrm{d}x}{(x-1)^4\sqrt{(x-1)^2-1}}\xlongequal{\text{令 }x-1=\sec\theta}\int_{\frac{\pi}{3}}^{\frac{\pi}{2}}\frac{\sec\theta\tan\theta}{\sec^4\theta\tan\theta}\mathrm{d}\theta$

$$=\int_{\frac{\pi}{3}}^{\frac{\pi}{2}}(1-\sin^2\theta)\cos\theta\,\mathrm{d}\theta=\frac{2}{3}-\frac{3\sqrt{3}}{8}.$$

例 1.8.46 已知 $\lim\limits_{x\to\infty}\left(\frac{x-a}{x+a}\right)^x=\int_a^{+\infty}4x^2\mathrm{e}^{-2x}\mathrm{d}x$，求常数 a 的值.

解 右边 $=-2\int_a^{+\infty}x^2\mathrm{d}(\mathrm{e}^{-2x})=-2x^2\mathrm{e}^{-2x}\Big|_a^{+\infty}+4\int_a^{+\infty}x\mathrm{e}^{-2x}\mathrm{d}x$

$$=2a^2\mathrm{e}^{-2a}-(2x\mathrm{e}^{-2x}+\mathrm{e}^{-2x})\Big|_a^{+\infty}=2a^2\mathrm{e}^{-2a}+2a\mathrm{e}^{-2a}+\mathrm{e}^{-2a}.$$

左边 $=\lim\limits_{x\to\infty}\left(1-\frac{2a}{x+a}\right)^x=\mathrm{e}^{-2a}$.

于是，有 $\mathrm{e}^{-2a}=2a^2\mathrm{e}^{-2a}+2a\mathrm{e}^{-2a}+\mathrm{e}^{-2a}$，得 $a=0$ 或 $a=-1$.

2. 敛散性的判别

这是难点.在《考试大纲》中，对反常积分敛散性判别的要求并不高，但是近几年却经常出现考题，且题目难度还比较大，从这个角度说，我们还是要仔细深入地来研究一下这个问题.

针对此类问题，我们需要掌握**两个重要结论**，并能够熟练地进行无穷小、无穷大比阶.

(1)无穷区间的反常积分 $\int_1^{+\infty}\frac{\mathrm{d}x}{x^p}$：在 $p>1$ 时收敛，在 $p\leqslant 1$ 时发散.

(2)无界函数的反常积分 $\int_0^1\frac{\mathrm{d}x}{x^p}(p>0$，奇点 $x=0)$：在 $0<p<1$ 时收敛，在 $p\geqslant 1$ 时发散.

例 1.8.47 判别反常积分 $\int_1^{+\infty}\frac{\left(\arctan\frac{1}{x}\right)^{\alpha}}{\left[\ln\left(1+\frac{1}{x}\right)\right]^{2\beta}}\mathrm{d}x\ (\alpha,\beta>0)$ 的敛散性.

解 积分存在唯一奇点 $x=+\infty$，当 $x\to+\infty$ 时，有

$$\arctan\frac{1}{x}\sim\frac{1}{x},\quad \ln\left(1+\frac{1}{x}\right)\sim\frac{1}{x},$$

故
$$\frac{\left(\arctan\frac{1}{x}\right)^{\alpha}}{\left[\ln\left(1+\frac{1}{x}\right)\right]^{2\beta}}\sim\frac{1}{x^{\alpha-2\beta}},$$

于是根据前面所说的重要结论(1)，当 $\alpha-2\beta>1$ 时，积分收敛；当 $\alpha-2\beta\leqslant1$ 时，积分发散.

例 1.8.48 已知 $\alpha>0$，则对于反常积分 $\int_0^1\frac{\ln x}{x^\alpha}\mathrm{d}x$ 的敛散性情况判别，正确的是(　　).

(A)当 $\alpha\geqslant1$ 时，积分收敛　　　　(B)当 $\alpha<1$ 时，积分收敛

(C)敛散性与 α 的取值无关，必收敛　　(D)敛散性与 α 的取值无关，必发散

解 应选(B).

首先，考生需要掌握的已知结论是对于无界函数的反常积分 $\int_0^1\frac{\mathrm{d}x}{x^p}$ ($p>0$，奇点 $x=0$)：在 $0<p<1$ 时收敛，在 $p\geqslant1$ 时发散.

根据上述结论，作如下讨论：

当 $\alpha<1$ 时，取正数 ε 充分小，使得 $\alpha+\varepsilon<1$，由于

$$\lim_{x\to0^+}x^{\alpha+\varepsilon}\frac{\ln x}{x^\alpha}=\lim_{x\to0^+}x^\varepsilon\ln x=\lim_{x\to0^+}\frac{\ln x}{x^{-\varepsilon}}=\lim_{x\to0^+}\frac{\frac{1}{x}}{-\varepsilon x^{-\varepsilon-1}}=\lim_{x\to0^+}\left(-\frac{1}{\varepsilon}x^\varepsilon\right)=0,$$

故当 $x\to0^+$ 时，$\frac{1}{x^{\alpha+\varepsilon}}$是比$\frac{\ln x}{x^\alpha}$高阶的无穷大量，于是 $\int_0^1\frac{\ln x}{x^\alpha}\mathrm{d}x$ 收敛，答案(B)正确；

当 $\alpha\geqslant1$ 时，由于$\lim\limits_{x\to0^+}x^\alpha\frac{\ln x}{x^\alpha}=\infty$，故当 $x\to0^+$ 时，$\frac{1}{x^\alpha}$是比$\frac{\ln x}{x^\alpha}$低阶的无穷大量，于是 $\int_0^1\frac{\ln x}{x^\alpha}\mathrm{d}x$ 发散.

例 1.8.49 讨论$\int_2^{+\infty}\frac{1}{x\ln^p x}\mathrm{d}x$ 的敛散性，其中 p 为任意实数.

分析 本题考查反常积分的敛散性的计算判别法(是否收敛，通过计算结果来判别)，同样是历届考生复习比较薄弱的知识点，考生可回顾例 1.8.48，那里使用的是理论判别法(是否收敛，无法通过计算结果来判别，只能用已有结论做比较判别).

解 ① 当 $p=1$ 时，$\int_2^{+\infty}\frac{1}{x\ln x}\mathrm{d}x=\ln(\ln x)\Big|_2^{+\infty}=+\infty$，发散；

② 当 $p\neq1$ 时，
$$\int_2^{+\infty}\frac{\mathrm{d}x}{x\ln^p x}=\frac{1}{1-p}(\ln x)^{1-p}\Big|_2^{+\infty},$$

$p>1$ 时，$\lim\limits_{x\to+\infty}(\ln x)^{1-p}=0$，故收敛，$p<1$ 时，$\lim\limits_{x\to+\infty}(\ln x)^{1-p}=+\infty$，故发散.

综上，$\int_2^{+\infty}\frac{1}{x\ln^p x}\mathrm{d}x\begin{cases}\text{收敛，} & p>1\text{ 时，}\\ \text{发散，} & p\leqslant1\text{ 时.}\end{cases}$

基础习题精练

习 题

1.8.1 设函数 $f(x)$ 与 $g(x)$ 在$[0,1]$上连续，且 $f(x)\leqslant g(x)$，则对任何 $c\in(0,1)$，有(　　).

(A) $\int_{\frac{1}{2}}^{c}f(t)\mathrm{d}t\geqslant\int_{\frac{1}{2}}^{c}g(t)\mathrm{d}t$　　　　(B) $\int_{\frac{1}{2}}^{c}f(t)\mathrm{d}t\leqslant\int_{\frac{1}{2}}^{c}g(t)\mathrm{d}t$

(C) $\int_c^1 f(t)\mathrm{d}t \geqslant \int_c^1 g(t)\mathrm{d}t$　　(D) $\int_c^1 f(t)\mathrm{d}t \leqslant \int_c^1 g(t)\mathrm{d}t$

1.8.2 设 $f'(\ln x)=1+x$，则 $\int_0^1 f'(2x)\mathrm{d}x=$ (　　).

(A) $\frac{\mathrm{e}^2+1}{2}$　　(B) $\frac{1-\mathrm{e}^2}{2}$　　(C) $\frac{\mathrm{e}^2-1}{2}$　　(D) $-\frac{\mathrm{e}^2+1}{2}$

1.8.3 下列反常积分中收敛的是(　　).

(A) $\int_{\mathrm{e}}^{+\infty}\frac{\ln x}{x}\mathrm{d}x$　　(B) $\int_{\mathrm{e}}^{+\infty}\frac{\mathrm{d}x}{x\ln x}$　　(C) $\int_{\mathrm{e}}^{+\infty}\frac{\mathrm{d}x}{x(\ln x)^2}$　　(D) $\int_{\mathrm{e}}^{+\infty}\frac{\mathrm{d}x}{x(\ln x)^{\frac{1}{2}}}$

1.8.4 $\int\frac{\arcsin\sqrt{x}}{\sqrt{x}}\mathrm{d}x=$ ________.

1.8.5 设 $f(x)$ 的一个原函数为 $\ln^2 x$，则 $\int xf'(x)\mathrm{d}x=$ ________.

1.8.6 $f(x)=\frac{1}{1+x^2}+x^3\int_0^1 f(x)\mathrm{d}x$，则 $\int_0^1 f(x)\mathrm{d}x=$ ________.

1.8.7 $\int \mathrm{e}^{\sqrt{2x+1}}\mathrm{d}x=$ ________.

1.8.8 设 $x>1$，则 $\int\frac{x\mathrm{e}^x}{\sqrt{\mathrm{e}^x-2}}\mathrm{d}x=$ ________.

1.8.9 $\int\frac{\mathrm{d}x}{1+\sin x+\cos x}=$ ________.

1.8.10 求不定积分 $\int\frac{x\cos^4\frac{x}{2}}{\sin^3 x}\mathrm{d}x$.

1.8.11 计算 $\int\arcsin\sqrt{\frac{x}{a+x}}\mathrm{d}x$ ($a>0$ 是常数).

1.8.12 求不定积分 $\int\frac{\arctan \mathrm{e}^x}{\mathrm{e}^x}\mathrm{d}x$.

1.8.13 计算不定积分 $\int\frac{\tan x}{a^2\sin^2 x+b^2\cos^2 x}\mathrm{d}x$ (a,b 不同时为零).

1.8.14 求极限 $\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+\frac{1}{i}}$.

1.8.15 $\lim\limits_{n\to\infty}\left(\frac{1}{n}\cdot\frac{2}{n}\cdot\cdots\cdot\frac{n}{n}\right)^{\frac{1}{n}}=$ ________.

1.8.16 计算定积分 $\int_{\frac{1}{2}}^{\frac{3}{2}}\frac{(1-x)\arcsin(1-x)}{\sqrt{2x-x^2}}\mathrm{d}x$.

1.8.17 计算定积分 $\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\frac{\cos x-\sin x}{\sqrt{\cos x}}\mathrm{d}x$.

1.8.18 计算定积分 $I=\int_0^{\pi}\frac{x\sin x}{1+\cos^2 x}\mathrm{d}x$.

1.8.19 计算定积分 $\int_{-1}^{1}(|x|+x)\mathrm{e}^{-|x|}\mathrm{d}x$.

1.8.20 设 $\lim\limits_{x\to\infty}\left(\frac{1+x}{x}\right)^{ax}=\int_{-\infty}^{a}t\mathrm{e}^t\mathrm{d}t$，求常数 a.

1.8.21 设 $f(x)=\begin{cases}\dfrac{1}{1+\sin x}, & x\geqslant 0,\\ \dfrac{1}{1+e^{x}}, & x<0,\end{cases}$ 求 $\int_{-1}^{\frac{\pi}{4}}f(x)\mathrm{d}x$.

解答

1.8.1 (D) **解** 设 $F(x)=g(x)-f(x)$，由已知条件知 $F(x)\geqslant 0$，所以有

$$\int_c^1 F(t)\mathrm{d}t\geqslant 0,\text{即}\int_c^1 f(t)\mathrm{d}t\leqslant\int_c^1 g(t)\mathrm{d}t.$$

故选(D).

1.8.2 (A) **解** 令 $\ln x=t$，则 $x=e^t$，从而 $f'(t)=1+e^t$，$f'(2x)=1+e^{2x}$，故

$$\int_0^1 f'(2x)\mathrm{d}x=\int_0^1(1+e^{2x})\mathrm{d}x=1+\frac{1}{2}e^{2x}\Big|_0^1=\frac{e^2+1}{2}.$$

故选(A).

1.8.3 (C) **解** 选项(A)，$\int_e^{+\infty}\frac{\ln x}{x}\mathrm{d}x=\frac{1}{2}(\ln x)^2\Big|_e^{+\infty}$，发散.

选项(B)，$\int_e^{+\infty}\frac{\mathrm{d}x}{x\ln x}=\int_e^{+\infty}\frac{\mathrm{d}(\ln x)}{\ln x}=\ln(\ln x)\Big|_e^{+\infty}$，发散.

选项(C)，$\int_e^{+\infty}\frac{\mathrm{d}x}{x(\ln x)^2}=\int_e^{+\infty}\frac{\mathrm{d}(\ln x)}{(\ln x)^2}=-\frac{1}{\ln x}\Big|_e^{+\infty}=-(0-1)=1$，收敛.

选项(D)，$\int_e^{+\infty}\frac{\mathrm{d}x}{x(\ln x)^{\frac{1}{2}}}=\int_e^{+\infty}(\ln x)^{-\frac{1}{2}}\mathrm{d}(\ln x)=2(\ln x)^{\frac{1}{2}}\Big|_e^{+\infty}$，发散.

故选(C).

1.8.4 $2\sqrt{x}\arcsin\sqrt{x}+2\sqrt{1-x}+C$ **解** 去掉根号将会使计算变得简单. 令 $\sqrt{x}=t$，$x=t^2$，则

$$\int\frac{\arcsin\sqrt{x}}{\sqrt{x}}\mathrm{d}x=2\int t\cdot\frac{\arcsin t}{t}\mathrm{d}t=2\int\arcsin t\mathrm{d}t=2(t\arcsin t+\sqrt{1-t^2})+C$$

$$=2\sqrt{x}\arcsin\sqrt{x}+2\sqrt{1-x}+C.$$

1.8.5 $2\ln x-\ln^2x+C$ **解** 被积函数中有 $f'(x)$，用分部积分法.

$$\int xf'(x)\mathrm{d}x=\int x\mathrm{d}[f(x)]=xf(x)-\int f(x)\mathrm{d}x=xf(x)-\ln^2x+C,$$

其中
$$f(x)=(\ln^2x)'=\frac{2\ln x}{x},$$

于是
$$\int xf'(x)\mathrm{d}x=2\ln x-\ln^2x+C.$$

1.8.6 $\frac{\pi}{3}$ **解** 定积分 $\int_0^1 f(x)\mathrm{d}x$ 是一个常数，所以等式两端同时积分得

$$\int_0^1 f(x)\mathrm{d}x=\int_0^1\frac{1}{1+x^2}\mathrm{d}x+\int_0^1\left[x^3\int_0^1 f(x)\mathrm{d}x\right]\mathrm{d}x,$$

即
$$\int_0^1 f(x)\mathrm{d}x=\arctan x\Big|_0^1+\int_0^1 f(x)\mathrm{d}x\int_0^1 x^3\mathrm{d}x$$

$$=\frac{\pi}{4}+\frac{1}{4}\int_0^1 f(x)\mathrm{d}x,$$

解得$\int_0^1 f(x)\mathrm{d}x = \frac{\pi}{3}$.

1.8.7 $(\sqrt{2x+1}-1)\mathrm{e}^{\sqrt{2x+1}}+C$ **解 方法一**

$$\int \mathrm{e}^{\sqrt{2x+1}}\mathrm{d}x \xlongequal{\text{令}\sqrt{2x+1}=t} \int t\mathrm{e}^t\mathrm{d}t = \int t\mathrm{d}(\mathrm{e}^t)$$

$$\xlongequal{\text{分部积分法}} t\mathrm{e}^t - \mathrm{e}^t + C = (\sqrt{2x+1}-1)\mathrm{e}^{\sqrt{2x+1}}+C.$$

方法二 $\int \mathrm{e}^{\sqrt{2x+1}}\mathrm{d}x \xlongequal{\text{凑微分法}} \int \sqrt{2x+1}\mathrm{e}^{\sqrt{2x+1}}\mathrm{d}(\sqrt{2x+1})$

$$= \int \sqrt{2x+1}\mathrm{d}(\mathrm{e}^{\sqrt{2x+1}}) \xlongequal{\text{分部积分法}} \sqrt{2x+1}\mathrm{e}^{\sqrt{2x+1}} - \int \mathrm{e}^{\sqrt{2x+1}}\mathrm{d}(\sqrt{2x+1})$$

$$= \sqrt{2x+1}\mathrm{e}^{\sqrt{2x+1}} - \mathrm{e}^{\sqrt{2x+1}} + C = (\sqrt{2x+1}-1)\mathrm{e}^{\sqrt{2x+1}}+C.$$

1.8.8 $2(x-2)\sqrt{\mathrm{e}^x-2}+4\sqrt{2}\arctan\sqrt{\frac{\mathrm{e}^x}{2}-1}+C$ **解**

$$\int \frac{x\mathrm{e}^x}{\sqrt{\mathrm{e}^x-2}}\mathrm{d}x \xlongequal{\text{令}\sqrt{\mathrm{e}^x-2}=t} 2\int \ln(t^2+2)\mathrm{d}t \xlongequal{\text{分部积分法}} 2t\ln(t^2+2) - 4\int \frac{t^2}{t^2+2}\mathrm{d}t$$

$$=2t\ln(t^2+2)-4t+4\sqrt{2}\arctan\frac{t}{\sqrt{2}}+C$$

$$=2(x-2)\sqrt{\mathrm{e}^x-2}+4\sqrt{2}\arctan\sqrt{\frac{\mathrm{e}^x}{2}-1}+C.$$

1.8.9 $\ln\left|1+\tan\frac{x}{2}\right|+C$ **解**

$$\int \frac{\mathrm{d}x}{1+\sin x+\cos x} = \int \frac{\mathrm{d}x}{2\sin\frac{x}{2}\cos\frac{x}{2}+2\cos^2\frac{x}{2}}$$

$$= \frac{1}{2}\int \frac{\mathrm{d}x}{\cos^2\frac{x}{2}\cdot\left(1+\tan\frac{x}{2}\right)} = \int \frac{\mathrm{d}\left(1+\tan\frac{x}{2}\right)}{1+\tan\frac{x}{2}}$$

$$= \ln\left|1+\tan\frac{x}{2}\right|+C.$$

1.8.10 解 当被积函数为幂函数与三角函数的乘积，且不能用凑微分法积分时，一定要用分部积分法积分.

因为分子与分母中三角函数的角度分别为$\frac{x}{2}$和x，所以首先要用三角恒等式变成同角度的三角函数形式.

方法一 $\int \frac{x\cos^4\frac{x}{2}}{\sin^3 x}\mathrm{d}x = \int \frac{x\cos^4\frac{x}{2}}{8\sin^3\frac{x}{2}\cos^3\frac{x}{2}}\mathrm{d}x = \int \frac{x\cos\frac{x}{2}}{8\sin^3\frac{x}{2}}\mathrm{d}x = \frac{1}{4}\int x\sin^{-3}\frac{x}{2}\mathrm{d}\left(\sin\frac{x}{2}\right)$

$$=-\frac{1}{8}\int x\mathrm{d}\left(\sin^{-2}\frac{x}{2}\right) = \frac{-x}{8\sin^2\frac{x}{2}} + \frac{1}{8}\int \frac{\mathrm{d}x}{\sin^2\frac{x}{2}} = \frac{-x}{8\sin^2\frac{x}{2}} - \frac{1}{4}\cot\frac{x}{2}+C.$$

方法二 $\int \frac{x\cos^4\frac{x}{2}}{\sin^3 x}\mathrm{d}x = \int \frac{x\cos\frac{x}{2}}{8\sin^3\frac{x}{2}}\mathrm{d}x = -\frac{1}{4}\int x\cot\frac{x}{2}\mathrm{d}\left(\cot\frac{x}{2}\right) = -\frac{1}{8}\int x\mathrm{d}\left(\cot^2\frac{x}{2}\right)$

$$=-\frac{1}{8}x\cot^2\frac{x}{2}+\frac{1}{8}\int\cot^2\frac{x}{2}\mathrm{d}x=-\frac{1}{8}x\cot^2\frac{x}{2}+\frac{1}{8}\int\left(\csc^2\frac{x}{2}-1\right)\mathrm{d}x$$

$$=\frac{-x}{8\sin^2\frac{x}{2}}-\frac{1}{4}\cot\frac{x}{2}+C.$$

1.8.11 **解** 令 $\arcsin\sqrt{\frac{x}{a+x}}=t,x=\frac{a\sin^2t}{1-\sin^2t}=a\tan^2t$，于是

$$\int\arcsin\sqrt{\frac{x}{a+x}}\mathrm{d}x=\int t\mathrm{d}(a\tan^2t)=at\tan^2t-a\int\tan^2t\mathrm{d}t$$

$$=at\tan^2t+a\int(1-\sec^2t)\mathrm{d}t=at\tan^2t+at-a\tan t+C$$

$$=(a+x)\arcsin\sqrt{\frac{x}{a+x}}-\sqrt{ax}+C.$$

1.8.12 **解** 因为不满足凑微分法的条件，所以要用分部积分法.

方法一
$$\int\frac{\arctan\mathrm{e}^x}{\mathrm{e}^x}\mathrm{d}x=-\int\arctan\mathrm{e}^x\mathrm{d}(\mathrm{e}^{-x})=-\mathrm{e}^{-x}\arctan\mathrm{e}^x+\int\frac{\mathrm{d}x}{1+\mathrm{e}^{2x}}$$

$$=-\mathrm{e}^{-x}\arctan\mathrm{e}^x+\int\left(1-\frac{\mathrm{e}^{2x}}{1+\mathrm{e}^{2x}}\right)\mathrm{d}x$$

$$=-\mathrm{e}^{-x}\arctan\mathrm{e}^x+x-\frac{1}{2}\ln(1+\mathrm{e}^{2x})+C.$$

方法二 令 $\mathrm{e}^x=t$，则

$$\int\frac{\arctan\mathrm{e}^x}{\mathrm{e}^x}\mathrm{d}x=\int\frac{\arctan t}{t^2}\mathrm{d}t=-\int\arctan t\mathrm{d}\left(\frac{1}{t}\right)=-\frac{1}{t}\arctan t+\int\frac{\mathrm{d}t}{t(1+t^2)}$$

$$=-\frac{1}{t}\arctan t+\int\frac{\mathrm{d}t}{t}-\int\frac{t\mathrm{d}t}{1+t^2}=-\frac{1}{t}\arctan t+\ln t-\frac{1}{2}\ln(1+t^2)+C$$

$$=-\frac{1}{\mathrm{e}^x}\arctan\mathrm{e}^x+x-\frac{1}{2}\ln(1+\mathrm{e}^{2x})+C.$$

1.8.13 **解** 当 $a=0,b\neq0$ 时，原式 $=\frac{1}{b^2}\int\frac{\sin x}{\cos^3x}\mathrm{d}x=\frac{1}{2b^2\cos^2x}+C$；

当 $a\neq0,b=0$ 时，
$$\text{原式}=\frac{1}{a^2}\int\frac{1}{\sin x\cos x}\mathrm{d}x=\frac{2}{a^2}\int\frac{1}{\sin 2x}\mathrm{d}x$$

$$=\frac{1}{a^2}\ln|\csc 2x-\cot 2x|+C;$$

当 $ab\neq0$ 时，
$$\text{原式}=\int\frac{\tan x}{\cos^2x(a^2\tan^2x+b^2)}\mathrm{d}x=\int\frac{\tan x\mathrm{d}(\tan x)}{a^2\tan^2x+b^2}$$

$$=\frac{1}{2a^2}\int\frac{\mathrm{d}(a^2\tan^2x+b^2)}{a^2\tan^2x+b^2}=\frac{1}{2a^2}\ln(a^2\tan^2x+b^2)+C.$$

1.8.14 **解** 当各项分母相同且均为 n 时，

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n}=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sin\frac{i}{n}\pi=\int_0^1\sin\pi x\mathrm{d}x,$$

于是，先对 $\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+\frac{1}{i}}$ 进行放缩，有

$$\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+1}\leqslant\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+\frac{1}{i}}\leqslant\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n},$$

由于
$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n}=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sin\frac{i}{n}\pi=\int_0^1\sin\pi x\mathrm{d}x=\frac{2}{\pi},$$

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+1}=\lim_{n\to\infty}\frac{n}{n+1}\cdot\frac{1}{n}\sum_{i=1}^{n}\sin\frac{i}{n}\pi=1\cdot\int_0^1\sin\pi x\mathrm{d}x=\frac{2}{\pi},$$

所以，由夹逼准则即得
$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{\sin\frac{i\pi}{n}}{n+\frac{1}{i}}=\frac{2}{\pi}.$$

1.8.15 $\frac{1}{\mathrm{e}}$ **解** 令 $x_n=\left(\frac{1}{n}\cdot\frac{2}{n}\cdot\cdots\cdot\frac{n}{n}\right)^{\frac{1}{n}}$，则

$$\ln x_n=\frac{1}{n}\left(\ln\frac{1}{n}+\ln\frac{2}{n}+\cdots+\ln\frac{n}{n}\right)=\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n},$$

于是

$$\begin{aligned}\lim_{n\to\infty}\ln x_n&=\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}=\int_0^1\ln x\mathrm{d}x\\&=x\ln x\Big|_0^1-\int_0^1x\cdot\frac{1}{x}\mathrm{d}x\\&=0-1=-1,\end{aligned}$$

故 $\lim\limits_{n\to\infty}x_n=\mathrm{e}^{-1}=\frac{1}{\mathrm{e}}.$

1.8.16 **解** 令 $1-x=\sin t$，则

$$\begin{aligned}\int_{\frac{1}{2}}^{\frac{3}{2}}\frac{(1-x)\arcsin(1-x)}{\sqrt{2x-x^2}}\mathrm{d}x&=\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}}\frac{t\sin t}{\cos t}\cos t\mathrm{d}t=\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}}t\sin t\mathrm{d}t=-\int_{-\frac{\pi}{6}}^{\frac{\pi}{6}}t\mathrm{d}(\cos t)\\&=-(t\cos t-\sin t)\Big|_{-\frac{\pi}{6}}^{\frac{\pi}{6}}=1-\frac{\sqrt{3}\pi}{6}.\end{aligned}$$

1.8.17 **解** 当被积函数为指数函数与三角函数的乘积时，不能使用凑微分法求出原函数，只能使用分部积分法.

$$\begin{aligned}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\frac{\cos x-\sin x}{\sqrt{\cos x}}\mathrm{d}x&=\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\mathrm{d}x-\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\frac{\sin x}{\sqrt{\cos x}}\mathrm{d}x\\&=\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\mathrm{d}x+2\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\mathrm{d}(\sqrt{\cos x})\\&=\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\mathrm{d}x+2\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\Big|_{-\frac{\pi}{4}}^{\frac{\pi}{4}}-\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\mathrm{d}x\\&=\sqrt[4]{8}(\mathrm{e}^{\frac{\pi}{8}}-\mathrm{e}^{-\frac{\pi}{8}}).\end{aligned}$$

积分过程中，$\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}}\mathrm{e}^{\frac{x}{2}}\sqrt{\cos x}\mathrm{d}x$ 被相互抵消，这是分部积分法的一种特殊类型，与循环积分类似.

1.8.18 **解** 被积函数中既有幂函数，又有三角函数，是一种较复杂的分式形式，很难直接求出其原函数. 但是，注意到被积函数为偶函数，其中 x 及 $\sin x$ 为奇函数，可考虑将其化成对称区间上的定积分，为

此，需将积分区间$[0,\pi]$平移成$\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$，计算如下：

设$x=\frac{\pi}{2}+u$，则有

$$I=\int_0^{\pi}\frac{x\sin x}{1+\cos^2x}\mathrm{d}x=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{\left(\frac{\pi}{2}+u\right)\cos u}{1+\sin^2u}\mathrm{d}u=\frac{\pi}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{\cos u}{1+\sin^2u}\mathrm{d}u+\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{u\cos u}{1+\sin^2u}\mathrm{d}u$$

$$=\pi\int_0^{\frac{\pi}{2}}\frac{\mathrm{d}(\sin u)}{1+\sin^2u}+0=\pi\arctan(\sin u)\Big|_0^{\frac{\pi}{2}}=\pi\times\frac{\pi}{4}=\frac{1}{4}\pi^2.$$

1.8.19 解 $\int_{-1}^{1}(|x|+x)\mathrm{e}^{-|x|}\mathrm{d}x=\int_{-1}^{0}(-x+x)\mathrm{e}^{x}\mathrm{d}x+\int_0^1(x+x)\mathrm{e}^{-x}\mathrm{d}x$

$$=2\int_0^1x\mathrm{e}^{-x}\mathrm{d}x=2\left(-x\mathrm{e}^{-x}\Big|_0^1+\int_0^1\mathrm{e}^{-x}\mathrm{d}x\right)$$

$$=2\left(-\mathrm{e}^{-1}-\mathrm{e}^{-x}\Big|_0^1\right)=2(-\mathrm{e}^{-1}-\mathrm{e}^{-1}+1)=2(1-2\mathrm{e}^{-1}).$$

1.8.20 解 由于$\lim\limits_{x\to\infty}\left(\frac{1+x}{x}\right)^{\alpha x}=\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^{\alpha x}=\mathrm{e}^{\alpha}$，

$$\int_{-\infty}^{\alpha}t\mathrm{e}^{t}\mathrm{d}t=(t-1)\mathrm{e}^{t}\Big|_{-\infty}^{\alpha}=(\alpha-1)\mathrm{e}^{\alpha}-\lim_{t\to-\infty}(t-1)\mathrm{e}^{t}=(\alpha-1)\mathrm{e}^{\alpha},$$

故$\alpha-1=1$，即$\alpha=2$.

1.8.21 解 $\int_{-1}^{\frac{\pi}{4}}f(x)\mathrm{d}x=\int_{-1}^{0}f(x)\mathrm{d}x+\int_0^{\frac{\pi}{4}}f(x)\mathrm{d}x=\int_{-1}^{0}\frac{\mathrm{d}x}{1+\mathrm{e}^x}+\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}x}{1+\sin x}$，

$$\int_{-1}^{0}\frac{\mathrm{d}x}{1+\mathrm{e}^x}\xlongequal{\mathrm{e}^x=t}\int_{\mathrm{e}^{-1}}^{1}\frac{1}{1+t}\cdot\frac{1}{t}\mathrm{d}t=\int_{\mathrm{e}^{-1}}^{1}\left(\frac{1}{t}-\frac{1}{1+t}\right)\mathrm{d}t=\ln\frac{t}{1+t}\Big|_{\mathrm{e}^{-1}}^{1}=-\ln 2+\ln(1+\mathrm{e}),$$

$$\int_0^{\frac{\pi}{4}}\frac{\mathrm{d}x}{1+\sin x}=\int_0^{\frac{\pi}{4}}\frac{1-\sin x}{\cos^2x}\mathrm{d}x=\int_0^{\frac{\pi}{4}}\sec^2x\mathrm{d}x-\int_0^{\frac{\pi}{4}}\frac{\sin x}{\cos^2x}\mathrm{d}x$$

$$=\tan x\Big|_0^{\frac{\pi}{4}}-\frac{1}{\cos x}\Big|_0^{\frac{\pi}{4}}=1-(\sqrt{2}-1)=2-\sqrt{2},$$

从而 $$\int_{-1}^{\frac{\pi}{4}}f(x)\mathrm{d}x=-\ln 2+\ln(1+\mathrm{e})+2-\sqrt{2}.$$

第9讲 一元函数积分学的几何应用

基础知识结构

- 平面图形的面积
- 旋转体的体积
- 函数的平均值

基础内容精讲

假设以下曲线都是连续的.

1. 用定积分表达和计算平面图形的面积

(1)曲线 $y=y_1(x)$ 与 $y=y_2(x)$ 及 $x=a,x=b(a<b)$ 所围成的平面图形的面积

$$S=\int_a^b|y_1(x)-y_2(x)|\,\mathrm{d}x.$$

(2)曲线 $r=r_1(\theta)$ 与 $r=r_2(\theta)$ 与两射线 $\theta=\alpha$ 与 $\theta=\beta(0<\beta-\alpha\leqslant 2\pi)$ 所围成的曲边扇形的面积

$$S=\frac{1}{2}\int_\alpha^\beta|r_1^2(\theta)-r_2^2(\theta)|\,\mathrm{d}\theta.$$

2. 用定积分表达和计算旋转体的体积

(1)曲线 $y=y(x)$ 与 $x=a,x=b(a<b)$ 及 x 轴围成的曲边梯形绕 x 轴旋转一周所得到的旋转体的体积

$$V=\int_a^b\pi y^2(x)\,\mathrm{d}x.$$

(2)曲线 $y=y_1(x)\geqslant 0$ 与 $y=y_2(x)\geqslant 0$ 及 $x=a,x=b(a<b)$ 所围成的平面图形绕 x 轴旋转一周所得到的旋转体的体积

$$V=\pi\int_a^b|y_1^2(x)-y_2^2(x)|\,\mathrm{d}x.$$

(3)曲线 $y=y(x)$ 与 $x=a,x=b(0\leqslant a<b)$ 及 x 轴围成的曲边梯形绕 y 轴旋转一周所得到的旋转体的体积

$$V_y=2\pi\int_a^b x|y(x)|\,\mathrm{d}x. \qquad (1\text{-}9\text{-}1)$$

【注】 公式(1-9-1)有时用起来很方便，现简单推导如下：

取$[x,x+\Delta x]$$(\Delta x>0)$，得到一个小竖条，如图 1-9-1 的阴影区域所示，此小竖条绕着 y 轴旋转一周，成为一个“圆柱壳”，将其沿任何一条竖线“切开”，可展开为一个“长方体”，其体积为

$$\mathrm{d}V_y = 2\pi x\mid y(x)\mid \mathrm{d}x,$$

故

$$V_y = 2\pi\int_a^b x\mid y(x)\mid \mathrm{d}x.$$

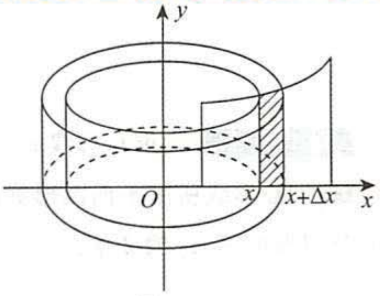


图 1-9-1

(4)曲线 $y=y_1(x)$与 $y=y_2(x)$及 $x=a,x=b(0\leqslant a\leqslant b)$所围成的图形绕 y 轴旋转一周所成的旋转体的体积

$$V = 2\pi\int_a^b x\mid y_1(x)-y_2(x)\mid \mathrm{d}x.$$

3. 用定积分表达和计算函数的平均值

设 $x\in[a,b]$，函数 $y(x)$在$[a,b]$上的平均值为 $\bar{y}=\dfrac{1}{b-a}\displaystyle\int_a^b y(x)\mathrm{d}x$.

基础例题精解

本部分内容大题、小题都常考，考生应熟记基本公式，并加强训练即可，没有什么难点. 考数学一、数学二的考生要注意弧长的计算问题，这是考试的热点. 弧长问题放在第 15 讲进行讲解.

例 1.9.1 求由曲线 $y=\sin x,y=\cos x$ 及直线 $x=0,x=\dfrac{\pi}{2}$ 所围平面图形的面积.

解 如图 1-9-2 所示，由$\begin{cases}y=\sin x,\\ y=\cos x,\end{cases}$知两曲线的交点为$\left(\dfrac{\pi}{4},\dfrac{\sqrt{2}}{2}\right)$，故

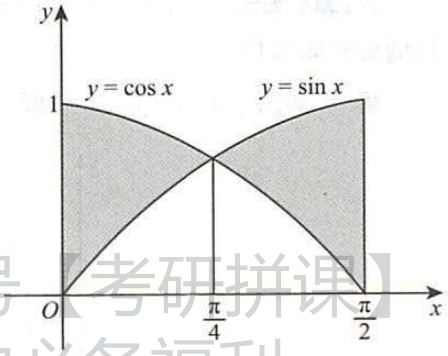


图 1-9-2

$$\begin{aligned}S &= \int_0^{\frac{\pi}{2}}\mid \sin x-\cos x\mid \mathrm{d}x\\ &= \int_0^{\frac{\pi}{4}}(\cos x-\sin x)\mathrm{d}x+\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}(\sin x-\cos x)\mathrm{d}x\\ &= (\sin x+\cos x)\Big|_0^{\frac{\pi}{4}}+(-\cos x-\sin x)\Big|_{\frac{\pi}{4}}^{\frac{\pi}{2}}\\ &= 2(\sqrt{2}-1).\end{aligned}$$

例 1.9.2 求由摆线$\begin{cases}x=a(t-\sin t),\\ y=a(1-\cos t)\end{cases}$$(a>0)$的一拱(如图 1-9-3)与 x 轴所围平面图形的面积.

解 当 $t=0,2\pi$ 时，$y=0$. 故当 t 由 0 变到 2π 时，曲线正好成一拱. 所以

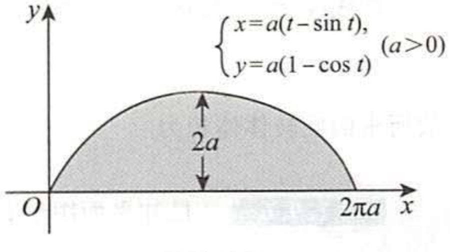


图 1-9-3

$$S=\int_0^{2\pi}a(1-\cos t)[a(t-\sin t)]'\mathrm{d}t$$

$$=\int_0^{2\pi}a^2(1-\cos t)^2\mathrm{d}t=3a^2\pi.$$

例 1.9.3　求心形线 $r=a(1+\cos\theta)(a>0)$ 所围平面图形的面积.

解　心形线所围平面图形如图 1-9-4 所示，此图形关于极轴对称，因此所求图形面积 S 是位于极轴上半部分图形面积 S_1 的 2 倍.

$$S_1=\int_0^{\pi}\frac{1}{2}a^2(1+\cos\theta)^2\mathrm{d}\theta$$

$$=\frac{a^2}{2}\int_0^{\pi}(1+2\cos\theta+\cos^2\theta)\mathrm{d}\theta$$

$$=\frac{a^2}{2}\int_0^{\pi}\left(\frac{3}{2}+2\cos\theta+\frac{1}{2}\cos 2\theta\right)\mathrm{d}\theta$$

$$=\frac{a^2}{2}\left(\frac{3}{2}\theta+2\sin\theta+\frac{1}{4}\sin 2\theta\right)\Big|_0^{\pi}=\frac{3}{4}a^2\pi,$$

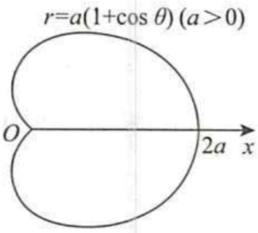


图 1-9-4

所以 $S=2S_1=\dfrac{3}{2}a^2\pi$.

例 1.9.4　求伯努利双纽线 $r^2=a^2\cos 2\theta$ 围成的图形面积.

解　如图 1-9-5 所示，利用对称性，所求面积是阴影部分面积的 4 倍.

阴影部分的图形由射线 $\theta=0,\theta=\dfrac{\pi}{4}$ 与伯努利双纽线 $r^2=a^2\cos 2\theta$ 围成，于是所求的平面图形面积为

$$S=4\int_0^{\frac{\pi}{4}}\frac{1}{2}a^2\cos 2\theta\mathrm{d}\theta=a^2\sin 2\theta\Big|_0^{\frac{\pi}{4}}=a^2.$$

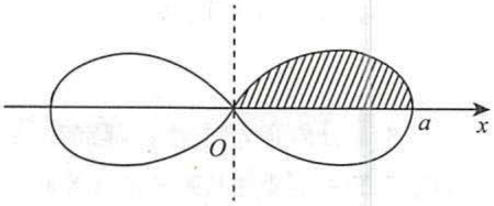


图 1-9-5

例 1.9.5　设平面图形由曲线 $y=x^2$ 与直线 $x=1$ 及 x 轴围成，求此平面图形绕 x 轴旋转一周所得的旋转体体积.

解　如图 1-9-6，图 1-9-7 所示.

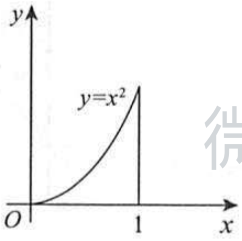


图 1-9-6

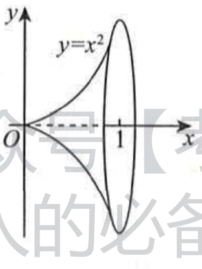


图 1-9-7

曲线 $y=x^2$ 与直线 $x=1$ 及 x 轴所围成的平面图形绕 x 轴旋转所得旋转体的体积微元为

$$\mathrm{d}V=\pi(x^2)^2\mathrm{d}x=\pi x^4\mathrm{d}x,$$

故所求的旋转体体积为

$$V=\int_0^1\pi x^4\mathrm{d}x=\frac{\pi}{5}x^5\Big|_0^1=\frac{\pi}{5}.$$

例 1.9.6　已知平面图形 D 由椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 围成，求 D 绕 x 轴旋转一周所得的旋转体（旋转椭球体）体积.

解　该旋转椭球体也可以看作是由上半椭圆 $y=\frac{b}{a}\sqrt{a^2-x^2}$ 与 x 轴围成的平面图形绕 x 轴旋转一周而成的旋转体，其体积微元为

$$\mathrm{d}V=\pi\left(\frac{b}{a}\sqrt{a^2-x^2}\right)^2\mathrm{d}x,$$

故所求的旋转体体积为

$$V=\int_{-a}^{a}\pi\left(\frac{b}{a}\sqrt{a^2-x^2}\right)^2\mathrm{d}x=\int_{-a}^{a}\frac{\pi b^2}{a^2}(a^2-x^2)\mathrm{d}x$$
$$=\frac{\pi b^2}{a^2}\left(a^2x-\frac{1}{3}x^3\right)\Big|_{-a}^{a}=\frac{4\pi}{3}ab^2.$$

例 1.9.7　曲线 $y=(x-1)(x-2)$ 和 x 轴围成一平面图形，求此平面图形绕 y 轴旋转一周所成的旋转体的体积.

解　在[1,2]上取积分微元，由“基础内容精讲”“2”中的“(3)”，得

$$\mathrm{d}V=2\pi x|y|\mathrm{d}x,$$

所以旋转体的体积为

$$V=\int_1^2 2\pi x\ |y|\ \mathrm{d}x=-2\pi\int_1^2 x(x-1)(x-2)\mathrm{d}x=\frac{1}{2}\pi.$$

例 1.9.8　计算由摆线 $\begin{cases}x=a(t-\sin t),\\ y=a(1-\cos t)\end{cases}$ $(a>0)$ 的一拱与 x 轴所围平面图形分别绕 x 轴及 y 轴旋转一周所得旋转体的体积.

解　平面图形绕 x 轴旋转一周所得旋转体体积为

$$\begin{aligned}V_x&=\pi\int_0^{2\pi a}y^2(x)\mathrm{d}x\\&=\pi\int_0^{2\pi}a^2(1-\cos t)^2a(1-\cos t)\mathrm{d}t\\&=\pi a^3\int_0^{2\pi}(1-3\cos t+3\cos^2 t-\cos^3 t)\mathrm{d}t\\&=5\pi^2a^3.\end{aligned}$$

作平面图形如图 1-9-8 所示. 平面图形绕 y 轴旋转一周所得旋转体体积为

$$\begin{aligned}V_y&=\pi\left[\int_0^{2a}x_2^2(y)\mathrm{d}y-\int_0^{2a}x_1^2(y)\mathrm{d}y\right]\\&=\pi\left[\int_{2\pi}^{\pi}a^2(t-\sin t)^2a\sin t\mathrm{d}t-\int_0^{\pi}a^2(t-\sin t)^2a\sin t\mathrm{d}t\right]\\&=-\pi a^3\int_0^{2\pi}(t-\sin t)^2\sin t\mathrm{d}t\\&=6\pi^3a^3,\end{aligned}$$

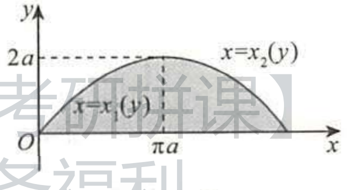


图 1-9-8

或者

$$\begin{aligned}V_y&=2\pi\int_0^{2\pi a}xy(x)\mathrm{d}x\\&=2\pi\int_0^{2\pi}a(t-\sin t)a(1-\cos t)a(1-\cos t)\mathrm{d}t\\&=6\pi^3a^3.\end{aligned}$$

例 1.9.9　过坐标原点作曲线 $y=\mathrm{e}^x$ 的切线，该切线与曲线 $y=\mathrm{e}^x$ 以及 x 轴围成的向 x 轴负向无限伸展的平面图形记为 D. 求

(1)D 的面积 A；

(2)D 绕直线 $x=1$ 旋转一周所成的旋转体的体积 V.

解 设切点坐标为 $P(x_0,y_0)$，于是曲线 $y=e^x$ 在点 P 的切线斜率为

$$y'(x_0)=e^{x_0},$$

切线方程为

$$y-y_0=e^{x_0}(x-x_0).$$

它经过点(0,0)，所以 $-y_0=-x_0e^{x_0}$. 又因 $y_0=e^{x_0}$，代入求得 $x_0=1$，从而 $y_0=e^{x_0}=e$，切线方程为 $y=ex$，如图 1-9-9 所示.

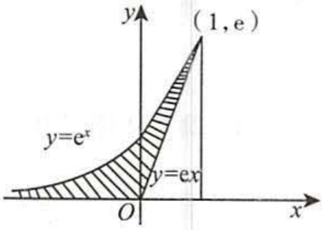


图 1-9-9

(1)取水平条面积微元，则 D 的面积

$$A=\int_0^e\left(\frac{y}{e}-\ln y\right)dy=\left(\frac{y^2}{2e}-y\ln y+y\right)\bigg|_0^e=\frac{e}{2}+\lim_{y\to0^+}y\ln y=\frac{e}{2}.$$

(积分 $\int_0^e \ln y dy$ 为反常积分，$\lim\limits_{y\to0^+}y\ln y=0$ 可由洛必达法则得到)

(2)D 绕直线 $x=1$ 旋转一周所成的旋转体的体积微元为

$$dV=\left[\pi(1-\ln y)^2-\pi\left(1-\frac{y}{e}\right)^2\right]dy,$$

从而

$$\begin{aligned}V&=\pi\int_0^e\left(\ln^2y-2\ln y+\frac{2y}{e}-\frac{y^2}{e^2}\right)dy\\&=\pi\left(y\ln^2y-4y\ln y+4y+\frac{y^2}{e}-\frac{y^3}{3e^2}\right)\bigg|_0^e=\frac{5}{3}\pi e.\end{aligned}$$

例 1.9.10 函数 $y=\dfrac{x^2}{\sqrt{1-x^2}}$在区间$\left[\dfrac{1}{2},\dfrac{\sqrt3}{2}\right]$上的平均值为________.

解 应填$\dfrac{\sqrt3+1}{12}\pi$.

函数 $y=f(x)$在区间$[a,b]$上的平均值是指$\dfrac{1}{b-a}\displaystyle\int_a^b f(x)dx$，故所求的平均值为

$$\frac{2}{\sqrt3-1}\int_{\frac12}^{\frac{\sqrt3}{2}}\frac{x^2}{\sqrt{1-x^2}}dx.$$

令 $x=\sin\theta$，$dx=\cos\theta d\theta$，则

$$上式=\frac{2}{\sqrt3-1}\int_{\frac{\pi}{6}}^{\frac{\pi}{3}}\sin^2\theta d\theta=\frac{2}{\sqrt3-1}\left(\frac12\theta-\frac14\sin2\theta\right)\bigg|_{\frac{\pi}{6}}^{\frac{\pi}{3}}=\frac{\sqrt3+1}{12}\pi.$$

基础习题精练

习 题

1.9.1 已知曲线 $y=a\sqrt{x}(a>0)$与曲线 $y=\ln\sqrt{x}$在点(x_0,y_0)处有公共切线，求：

(1)常数 a 及切点(x_0,y_0)；

(2)两曲线与 x 轴围成的平面图形的面积 S.

1.9.2　设曲线 $y=x^2-2x$，$y=0$，$x=1$，$x=3$ 围成一平面图形 A，求：

(1)A 的面积 S；

(2)该平面图形绕 y 轴旋转一周所得旋转体的体积 V.

1.9.3　已知一抛物线经过 x 轴上两点 $A(1,0)$，$B(3,0)$.

(1)求证两坐标轴与该抛物线所围图形的面积等于 x 轴与该抛物线所围图形的面积；

(2)计算(1)中两个平面图形绕 x 轴旋转一周所产生的两个旋转体的体积之比.

1.9.4　求曲线 $y=\sqrt{x}\mathrm{e}^{-x}$ 与 x 轴所围成的平面图形在 $[0,+\infty)$ 内绕 x 轴旋转一周所成的旋转体的体积.

1.9.5　设 D_1 是由抛物线 $y=2x^2$ 和直线 $x=a$，$x=2$ 及 $y=0$ 所围成的平面区域；D_2 是由抛物线 $y=2x^2$ 和直线 $y=0$，$x=a$ 所围成的平面区域，其中 $0<a<2$.

(1)试求 D_1 绕 x 轴旋转一周而成的旋转体体积 V_1，D_2 绕 y 轴旋转一周而成的旋转体体积 V_2；

(2)问当 a 为何值时，V_1+V_2 取得最大值？试求此最大值.

1.9.6　设直线 $y=ax$ 与抛物线 $y=x^2$ 所围成图形的面积为 S_1，它们与直线 $x=1$ 所围成图形的面积为 S_2，并且 $a<1$.

(1)试确定 a 的值，使 S_1+S_2 达到最小，并求出最小值；

(2)求该最小值所对应的平面图形绕 x 轴旋转一周所得旋转体的体积.

解答

1.9.1　分析　利用两条曲线都经过点 (x_0,y_0) 及两条曲线在点 (x_0,y_0) 处有公共切线可列出三个方程，从而可解出常数 a,x_0,y_0，然后可求出平面图形的面积.

解　(1)由题设条件可得

$$\begin{cases} y_0=a\sqrt{x_0}, \\ y_0=\ln\sqrt{x_0}, \\ \dfrac{a}{2\sqrt{x_0}}=\dfrac{1}{2x_0}, \end{cases}$$

解此方程组可得 $a=\dfrac{1}{\mathrm{e}}$，$x_0=\mathrm{e}^2$，$y_0=1$，于是切点为 $(\mathrm{e}^2,1)$.

(2)**方法一**　画出曲线 $y=\dfrac{1}{\mathrm{e}}\sqrt{x}$ 与曲线 $y=\ln\sqrt{x}$ 的图形，则两曲线与 x 轴围成的平面图形(如图 1-9-10)的面积

$$S=\int_0^1(\mathrm{e}^{2y}-\mathrm{e}^2y^2)\mathrm{d}y=\frac{1}{6}\mathrm{e}^2-\frac{1}{2}.$$

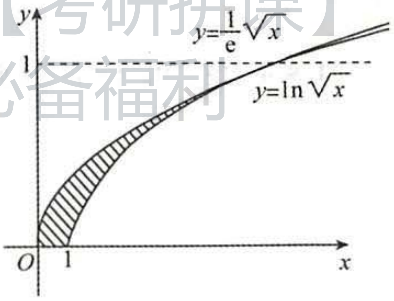


图 1-9-10

方法二　$S=\displaystyle\int_0^1\frac{1}{\mathrm{e}}\sqrt{x}\mathrm{d}x+\int_1^{\mathrm{e}^2}\left(\frac{1}{\mathrm{e}}\sqrt{x}-\ln\sqrt{x}\right)\mathrm{d}x=\frac{1}{6}\mathrm{e}^2-\frac{1}{2}.$

方法三　$S=\displaystyle\int_0^{\mathrm{e}^2}\frac{1}{\mathrm{e}}\sqrt{x}\mathrm{d}x-\int_1^{\mathrm{e}^2}\ln\sqrt{x}\mathrm{d}x=\frac{1}{6}\mathrm{e}^2-\frac{1}{2}.$

1.9.2 **解** (1)画出图形(如图 1-9-11),可见 $S=S_1+S_2$,其中

$$S_1=\int_1^2(2x-x^2)\mathrm{d}x=\frac{2}{3},$$

$$S_2=\int_2^3(x^2-2x)\mathrm{d}x=\frac{4}{3},$$

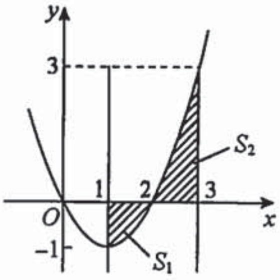


图 1-9-11

所以 $S=2$.

(2)**方法一** $V=V_1+V_2$,其中

$$V_1=\pi\int_{-1}^0(1+\sqrt{1+y})^2\mathrm{d}y-\pi=\frac{11\pi}{6},$$

$$V_2=27\pi-\pi\int_0^3(1+\sqrt{1+y})^2\mathrm{d}y=\frac{43\pi}{6},$$

所以 $V=9\pi$.

方法二 $V_1=2\pi\int_1^2x[0-(x^2-2x)]\mathrm{d}x=\frac{11\pi}{6}$, $V_2=2\pi\int_2^3x(x^2-2x)\mathrm{d}x=\frac{43\pi}{6}$,

所以 $V=9\pi$.

1.9.3 (1)**证明** **方法一** 设过 $A(1,0)$,$B(3,0)$两点的抛物线方程为

$$y=a(x-1)(x-3),$$

则两坐标轴与该抛物线所围图形的面积为

$$S_1=\int_0^1|a(x-1)(x-3)|\mathrm{d}x=|a|\int_0^1(x^2-4x+3)\mathrm{d}x=\frac{4}{3}|a|.$$

x 轴与该抛物线所围图形的面积为

$$S_2=\int_1^3|a(x-1)(x-3)|\mathrm{d}x=-|a|\int_1^3(x^2-4x+3)\mathrm{d}x=\frac{4}{3}|a|,$$

所以 $S_1=S_2$.

方法二 因为 $\int_0^3a(x-1)(x-3)\mathrm{d}x=a\left(\frac{1}{3}x^3-2x^2+3x\right)\Big|_0^3=0$,所以 $S_1=S_2$.

(2)**解** 两个平面图形绕 x 轴旋转一周所产生的两个旋转体的体积分别为

$$V_1=\pi\int_0^1a^2[(x-1)(x-3)]^2\mathrm{d}x=\frac{38}{15}\pi a^2,$$

$$V_2=\pi\int_1^3a^2[(x-1)(x-3)]^2\mathrm{d}x=\frac{16}{15}\pi a^2,$$

所以 $\frac{V_1}{V_2}=\frac{19}{8}$.

1.9.4 **解** 由题意得,该平面图形绕 x 轴旋转一周所成的旋转体体积为

$$\begin{aligned}V_x&=\pi\int_0^{+\infty}(\sqrt{x}\mathrm{e}^{-x})^2\mathrm{d}x=\pi\int_0^{+\infty}x\mathrm{e}^{-2x}\mathrm{d}x\\&=-\frac{1}{2}\pi\int_0^{+\infty}x\mathrm{d}(\mathrm{e}^{-2x})=-\frac{\pi}{2}\left(x\mathrm{e}^{-2x}\Big|_0^{+\infty}-\int_0^{+\infty}\mathrm{e}^{-2x}\mathrm{d}x\right)\\&=-\frac{\pi}{2}\left(0+\frac{1}{2}\mathrm{e}^{-2x}\Big|_0^{+\infty}\right)=\frac{\pi}{4}.\end{aligned}$$

1.9.5 **解** (1)由题意得, $V_1=\pi\int_a^2(2x^2)^2\mathrm{d}x=\frac{4}{5}\pi(32-a^5)$,

$$V_2=\pi a^2\cdot2a^2-\pi\int_0^{2a^2}\frac{y}{2}\mathrm{d}y=\pi a^4.$$

(2) 由(1)得，

$$V=V_1+V_2=\frac{4}{5}\pi(32-a^5)+\pi a^4.$$

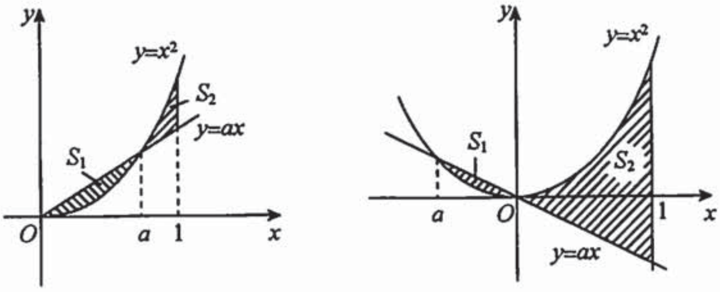
令

$$V'=4\pi a^3(1-a)=0,$$

得区间(0,2)内唯一的驻点 $a=1$，且 $V''(1)=-4\pi<0$，因此 $a=1$ 是极大值点，即最大值点，此时

$$V_{\max}=\frac{129}{5}\pi.$$

1.9.6　解　因为 $a<1$，所以可分成 $0<a<1$，$a\leqslant 0$ 两种情况，分别画出两种情况下的图形（如图 1-9-12），求出 S_1+S_2 的最小值后，即可确定 a 的值.



(a) $0<a<1$　　　(b) $a\leqslant 0$

图 1-9-12

(1)当 $0<a<1$ 时，

$$S=S_1+S_2=\int_0^a(ax-x^2)\mathrm{d}x+\int_a^1(x^2-ax)\mathrm{d}x=\frac{a^3}{3}-\frac{a}{2}+\frac{1}{3},$$

令 $S'(a)=a^2-\frac{1}{2}=0$，求得 $a=\frac{1}{\sqrt{2}}$. 又 $S''\left(\frac{1}{\sqrt{2}}\right)=\sqrt{2}>0$，知 $S\left(\frac{1}{\sqrt{2}}\right)=\frac{2-\sqrt{2}}{6}$是极小值，即最小值；

当 $a\leqslant 0$ 时，

$$S=S_1+S_2=\int_a^0(ax-x^2)\mathrm{d}x+\int_0^1(x^2-ax)\mathrm{d}x=-\frac{a^3}{6}-\frac{a}{2}+\frac{1}{3},$$

因为

$$S'(a)=-\frac{1}{2}a^2-\frac{1}{2}=-\frac{1}{2}(a^2+1)<0,$$

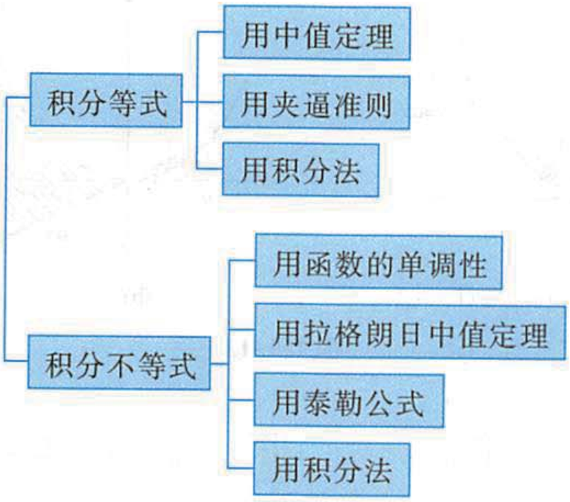
所以 S 单调减少，故 $a=0$ 时，S 取得最小值，此时 $S=\frac{1}{3}$.

比较可知，当 $a=\frac{1}{\sqrt{2}}$时，$S\left(\frac{1}{\sqrt{2}}\right)=\frac{2-\sqrt{2}}{6}$是最小值.

(2) 由(1)可知旋转体体积为 $V_x=\pi\int_0^{\frac{1}{\sqrt{2}}}\left(\frac{1}{2}x^2-x^4\right)\mathrm{d}x+\pi\int_{\frac{1}{\sqrt{2}}}^1\left(x^4-\frac{1}{2}x^2\right)\mathrm{d}x=\frac{\sqrt{2}+1}{30}\pi.$

第10讲 积分等式与积分不等式

基础知识结构



基础内容精讲

积分等式主要涉及积分形式的中值定理(见例 1.10.1,例 1.10.2),用夹逼准则求一类积分的极限(见例 1.10.3,例 1.10.4)与证明某些特殊的积分等式(见例 1.10.5);积分不等式问题主要涉及积分形式的不等式证明,可用函数的单调性(见例 1.10.6,例 1.10.7)、拉格朗日中值定理(见例 1.10.8)、泰勒公式(见例 1.10.9)与积分法(见例 1.10.10)来解决.

基础例题精解

一、积分等式

1. 用中值定理

例 1.10.1 设 $f(x)$,$g(x)$在$[a,b]$上连续且 $g(x)$不变号,证明:至少存在一点 $\xi\in[a,b]$,使得

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi)\int_a^b g(x)\mathrm{d}x\,.$$

证明 当 $g(x)\equiv 0$ 时,有

$$\int_a^b g(x)\mathrm{d}x=0,\quad \int_a^b f(x)g(x)\mathrm{d}x=\int_a^b f(x)\cdot 0\mathrm{d}x=0,$$

此时 ξ 可以是 $[a,b]$ 上任何一点，都会有 $\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x$ 成立 .

当 $g(x)\neq 0$ 时，必存在 $x_0\in[a,b]$，$g(x_0)$ 或者大于零或者小于零. 不妨设 $g(x_0)>0$，此时由 $g(x)$ 不变号且连续，必有 $\int_a^b g(x)\mathrm{d}x>0$. 又 $f(x)$ 在 $[a,b]$ 上连续，必能取到最小值 m 与最大值 M，于是对于一切 $x\in[a,b]$，都有

$$m\leqslant f(x)\leqslant M\Rightarrow mg(x)\leqslant f(x)g(x)\leqslant Mg(x)\Rightarrow$$

$$m\int_a^b g(x)\mathrm{d}x=\int_a^b mg(x)\mathrm{d}x\leqslant\int_a^b f(x)g(x)\mathrm{d}x\leqslant\int_a^b Mg(x)\mathrm{d}x=M\int_a^b g(x)\mathrm{d}x.$$

由于 $\int_a^b g(x)\mathrm{d}x>0$，得 $m\leqslant\dfrac{\int_a^b f(x)g(x)\mathrm{d}x}{\int_a^b g(x)\mathrm{d}x}\leqslant M.$

根据介值定理，至少存在一点 $\xi\in[a,b]$，使得 $\dfrac{\int_a^b f(x)g(x)\mathrm{d}x}{\int_a^b g(x)\mathrm{d}x}=f(\xi)$，即

$$\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x.$$

【注】 本题是推广的积分中值定理. $\xi\in(a,b)$ 怎么证，见下例.

设 $f(x)$，$g(x)$ 在 $[a,b]$ 上连续，且 $g(x)$ 在 $[a,b]$ 上不变号，证明 $\exists\xi\in(a,b)$，使

$$\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x.$$

证明　若 $g(x)\equiv 0$，结论显然成立；

若 $g(x)\not\equiv 0$，由于不变号，不妨设 $g(x)>0$. 令

$$F(x)=\int_a^x f(t)g(t)\mathrm{d}t,\quad G(x)=\int_a^x g(t)\mathrm{d}t,$$

在 $[a,b]$ 上用柯西中值定理，有 $\dfrac{F(b)-F(a)}{G(b)-G(a)}=\dfrac{F'(\xi)}{G'(\xi)}$，即

$$\frac{\int_a^b f(x)g(x)\mathrm{d}x-0}{\int_a^b g(x)\mathrm{d}x-0}=\frac{f(\xi)g(\xi)}{g(\xi)},$$

$$\int_a^b f(x)g(x)\mathrm{d}x=f(\xi)\int_a^b g(x)\mathrm{d}x,\xi\in(a,b),$$

其中 $\int_a^b g(x)\mathrm{d}x>0$. 得证.

例 1.10.2　设函数 $f(x)$ 在 $[a,b]$ 上不恒为 0 且二阶可导，$f(a)>0$，$f(b)>0$，$\int_a^b f(x)\mathrm{d}x=0$，证明：

(1) 存在不同的 $\xi_1,\xi_2\in(a,b)$，使得 $f(\xi_1)=f(\xi_2)=0$；

(2) 存在 $\xi\in(a,b)$，使得 $f''(\xi)>0$.

证明　(1) 依题设，$f(a)>0$，$f(b)>0$.

由$\int_a^b f(x)\mathrm{d}x=0$知，存在$x_0\in(a,b)$，使得$f(x_0)<0$，否则$f(x)$非负，且$f(x)\not\equiv 0$，则$\int_a^b f(x)\mathrm{d}x>0$，矛盾.

由零点定理，存在$\xi_1\in(a,x_0)$，$\xi_2\in(x_0,b)$，使得$f(\xi_1)=0$，$f(\xi_2)=0$.

(2)$f(x)$在$[a,\xi_1]$，$[\xi_2,b]$上分别应用拉格朗日中值定理，则存在$\eta_1\in(a,\xi_1)$，使得

$$f'(\eta_1)=\frac{f(\xi_1)-f(a)}{\xi_1-a}<0,$$

存在$\eta_2\in(\xi_2,b)$，使得

$$f'(\eta_2)=\frac{f(b)-f(\xi_2)}{b-\xi_2}>0.$$

$f'(x)$在$[\eta_1,\eta_2]$上应用拉格朗日中值定理，则存在$\xi\in(\eta_1,\eta_2)\subset(a,b)$，使得

$$f''(\xi)=\frac{f'(\eta_2)-f'(\eta_1)}{\eta_2-\eta_1}>0.$$

2. 用夹逼准则

例 1.10.3 求$\lim\limits_{n\to\infty}\int_0^1\frac{x^n}{1+x}\mathrm{d}x$.

解 因为$0\leqslant\frac{x^n}{1+x}\leqslant x^n(\forall x\in[0,1])$，所以$0\leqslant\int_0^1\frac{x^n}{1+x}\mathrm{d}x\leqslant\frac{1}{1+n}$，故由夹逼准则知

$$\lim_{n\to\infty}\int_0^1\frac{x^n}{1+x}\mathrm{d}x=0.$$

例 1.10.4 (1)比较$\int_0^1|\ln t|[\ln(1+t)]^n\mathrm{d}t$与$\int_0^1 t^n|\ln t|\mathrm{d}t(n=1,2,\cdots)$的大小，说明理由.

(2)记$u_n=\int_0^1|\ln t|[\ln(1+t)]^n\mathrm{d}t(n=1,2,\cdots)$，求$\lim\limits_{n\to\infty}u_n$.

解 (1)当$0\leqslant t\leqslant 1$时，$0\leqslant\ln(1+t)\leqslant t$，所以

$$0\leqslant|\ln t|[\ln(1+t)]^n\leqslant t^n|\ln t|,$$

根据积分的保号性，得

$$\int_0^1|\ln t|[\ln(1+t)]^n\mathrm{d}t\leqslant\int_0^1 t^n|\ln t|\mathrm{d}t.$$

(2)由(1)知，
$$0\leqslant u_n=\int_0^1|\ln t|[\ln(1+t)]^n\mathrm{d}t\leqslant\int_0^1 t^n|\ln t|\mathrm{d}t.$$

又因为
$$\int_0^1 t^n|\ln t|\mathrm{d}t=-\int_0^1 t^n\ln t\,\mathrm{d}t=-\frac{t^{n+1}}{n+1}\ln t\Big|_0^1+\frac{1}{n+1}\int_0^1 t^n\mathrm{d}t=\frac{1}{(n+1)^2},$$

所以$\lim\limits_{n\to\infty}\int_0^1 t^n|\ln t|\mathrm{d}t=0$，于是由夹逼准则得$\lim\limits_{n\to\infty}u_n=0$.

3. 用积分法

例 1.10.5 设$f(x)$是连续的偶函数，且是以T为周期的周期函数.

(1) 证明$\int_0^{nT}xf(x)\mathrm{d}x=\frac{n^2T}{2}\int_0^T f(x)\mathrm{d}x(n=1,2,3,\cdots)$；

(2) 利用(1)的结论计算$I=\int_0^{n\pi}x\mid\sin x\mid\mathrm{d}x$.

(1) **证明**
$$\int_0^{nT}xf(x)\mathrm{d}x\xlongequal{x=nT-t}nT\int_0^{nT}f(t)\mathrm{d}t-\int_0^{nT}tf(t)\mathrm{d}t,$$

于是有

$$\int_0^{nT} xf(x)\mathrm{d}x = \frac{nT}{2}\int_0^{nT} f(x)\mathrm{d}x.$$

又 $f(x+T)=f(x)$，则

$$\int_0^{nT} f(x)\mathrm{d}x = n\int_0^{T} f(x)\mathrm{d}x,$$

故

$$\int_0^{nT} xf(x)\mathrm{d}x = \frac{n^2T}{2}\int_0^{T} f(x)\mathrm{d}x(n=1,2,3,\cdots).$$

(2) **解** $|\sin x|$ 是连续的以 π 为周期的偶函数，故

$$I=\int_0^{n\pi} x|\sin x|\mathrm{d}x = \frac{n^2\pi}{2}\int_0^{\pi}|\sin x|\mathrm{d}x$$
$$=\frac{n^2\pi}{2}\int_0^{\pi}\sin x\mathrm{d}x = n^2\pi.$$

二、积分不等式

1. 用函数的单调性

通常的做法：首先将某一限（通常取上限）变量化，然后移项构造辅助函数，由辅助函数的单调性来证明不等式，此方法多用于所给条件为"$f(x)$在$[a,b]$上连续"的情形.

例 1.10.6 设 $f(x)$在$[a,b]$上连续，且 $f(x)>0$，证明：$\int_a^b f(x)\mathrm{d}x\int_a^b \frac{1}{f(x)}\mathrm{d}x \geqslant (b-a)^2$.

证明 令 $F(x)=\int_a^x f(t)\mathrm{d}t\int_a^x \frac{1}{f(t)}\mathrm{d}t-(x-a)^2(a\leqslant x\leqslant b)$，则

$$F'(x)=f(x)\int_a^x \frac{\mathrm{d}t}{f(t)}+\frac{1}{f(x)}\int_a^x f(t)\mathrm{d}t-2(x-a)$$
$$=\int_a^x\left[\frac{f(x)}{f(t)}+\frac{f(t)}{f(x)}-2\right]\mathrm{d}t \geqslant \int_a^x (2-2)\mathrm{d}t=0.$$

从而，$F(x)$单调增加. 故 $F(b)\geqslant F(a)=0$，得证.

例 1.10.7 设函数 $f(x)$，$g(x)$在区间$[a,b]$上连续，且 $f(x)$单调增加，$0\leqslant g(x)\leqslant 1$. 证明：

(1) $0\leqslant\int_a^x g(t)\mathrm{d}t\leqslant x-a, x\in[a,b]$；

(2) $\int_a^{a+\int_a^b g(t)\mathrm{d}t} f(x)\mathrm{d}x\leqslant\int_a^b f(x)g(x)\mathrm{d}x$.

证明 (1) 因为 $0\leqslant g(x)\leqslant 1$，所以当 $x\in[a,b]$ 时，有 $\int_a^x 0\mathrm{d}t\leqslant\int_a^x g(t)\mathrm{d}t\leqslant\int_a^x 1\mathrm{d}t$，即 $0\leqslant\int_a^x g(t)\mathrm{d}t\leqslant x-a$.

(2) 令 $F(x)=\int_a^{a+\int_a^x g(u)\mathrm{d}u} f(t)\mathrm{d}t-\int_a^x f(t)g(t)\mathrm{d}t, x\in[a,b]$.

因为 $f(x)$，$g(x)$ 在区间$[a,b]$ 上连续，所以 $F(x)$ 在区间$[a,b]$ 上可导，且

$$F'(x)=\left\{f\left[a+\int_a^x g(u)\mathrm{d}u\right]-f(x)\right\}g(x).$$

由(1) 知，$a+\int_a^x g(u)\mathrm{d}u\leqslant x, x\in[a,b]$. 又因为 $f(x)$ 单调增加，且 $g(x)\geqslant 0$，所以 $F'(x)\leqslant 0$，从而

$F(x)$ 在区间$[a,b]$ 上单调减少.

又 $F(a)=0$，故 $F(b)\leqslant 0$，即$\int_a^{a+\int_a^b g(t)\mathrm{d}t} f(x)\mathrm{d}x\leqslant\int_a^b f(x)g(x)\mathrm{d}x$.

2. 用拉格朗日中值定理

此方法多用于所给条件为“$f(x)$一阶可导”且某一端点值较简单(甚至为 0)的题目.

例 1.10.8 设 $f(x)$在$[0,1]$上具有一阶连续导数，且 $f(0)=f(1)=0$. 证明：

$$\left|\int_0^1 f(x)\mathrm{d}x\right|\leqslant\frac{1}{4}\max_{x\in[0,1]}\{|f'(x)|\}.$$

证明 将大区间$[0,1]$分成两个小区间$[0,x]$和$[x,1]$.

在$[0,x]$上对 $f(x)$使用拉格朗日中值定理，得 $f(x)-f(0)=f(x)=f'(\xi_1)x$，其中 $\xi_1\in(0,x)$，于是

$$|f(x)|=|f'(\xi_1)|x,$$

在$[x,1]$上对 $f(x)$使用拉格朗日中值定理，得 $f(1)-f(x)=-f(x)=f'(\xi_2)(1-x)$，其中 $\xi_2\in(x,1)$，于是

$$|f(x)|=|f'(\xi_2)|(1-x),$$

当 $x\in[0,1]$时，记 $M=\max\{|f'(x)|\}$，则

$$|f(x)|\leqslant Mx,\quad |f(x)|\leqslant M(1-x),$$

于是

$$\left|\int_0^1 f(x)\mathrm{d}x\right|=\left|\int_0^x f(t)\mathrm{d}t+\int_x^1 f(t)\mathrm{d}t\right|$$
$$\leqslant M\int_0^x t\mathrm{d}t+M\int_x^1(1-t)\mathrm{d}t=M\left[\frac{x^2}{2}+\frac{(1-x)^2}{2}\right],$$

其中，根据基本不等式，$\min\left\{\frac{x^2}{2}+\frac{(1-x)^2}{2}\right\}=\frac{1}{4}$，故得证.

3. 用泰勒公式

此方法多用于所给条件为“$f(x)$二阶可导”且某一端点值较简单(甚至为 0)的题目.

例 1.10.9 设 $f(x)$在$[0,2]$上二阶导数连续，且 $f(1)=0$. 当 $x\in[0,2]$时，记 $M=\max\{|f''(x)|\}$，证明$\left|\int_0^2 f(x)\mathrm{d}x\right|\leqslant\frac{1}{3}M$.

证明 根据题设，选取基点 $x_0=1$ 展开成泰勒公式，

$$f(x)=f(1)+f'(1)(x-1)+\frac{f''(\xi)}{2}(x-1)^2\text{，其中 }\xi\text{ 介于 }x,1\text{ 之间}$$

$$\Rightarrow\int_0^2 f(x)\mathrm{d}x=f'(1)\int_0^2(x-1)\mathrm{d}x+\int_0^2\frac{f''(\xi)}{2}(x-1)^2\mathrm{d}x=\frac{1}{2}\int_0^2 f''(\xi)(x-1)^2\mathrm{d}x$$

$$\Rightarrow\left|\int_0^2 f(x)\mathrm{d}x\right|\leqslant\frac{1}{2}\int_0^2|f''(\xi)|(x-1)^2\mathrm{d}x\leqslant\frac{1}{2}M\int_0^2(x-1)^2\mathrm{d}x=\frac{1}{3}M,$$

故得证.

4. 用积分法

例 1.10.10 设 $f''(x)$ 在$[0,1]$ 上连续，且 $f(0)=f(1)=0$，求证：

(1)$\int_0^1 f(x)\mathrm{d}x=\frac{1}{2}\int_0^1 x(x-1)f''(x)\mathrm{d}x$；

(2) $\left|\int_0^1 f(x)\mathrm{d}x\right| \leqslant \dfrac{1}{12}\max\limits_{0\leqslant x\leqslant 1}|f''(x)|$.

证明 (1) $\dfrac{1}{2}\displaystyle\int_0^1 x(x-1)f''(x)\mathrm{d}x = \dfrac{1}{2}\int_0^1 x(x-1)\mathrm{d}[f'(x)]$

$$= \frac{1}{2}x(x-1)f'(x)\Big|_0^1 - \frac{1}{2}\int_0^1 f'(x)(2x-1)\mathrm{d}x$$

$$= -\frac{1}{2}\int_0^1 (2x-1)\mathrm{d}[f(x)]$$

$$= -\frac{1}{2}(2x-1)f(x)\Big|_0^1 + \int_0^1 f(x)\mathrm{d}x.$$

由条件 $f(0)=f(1)=0$ 知结论成立.

(2) 记 $M=\max\limits_{0\leqslant x\leqslant 1}\{|f''(x)|\}$，则由(1)有

$$\left|\int_0^1 f(x)\mathrm{d}x\right| \leqslant \frac{M}{2}\int_0^1 x(1-x)\mathrm{d}x = \frac{M}{2}\left(\frac{1}{2}-\frac{1}{3}\right) = \frac{M}{12}.$$

基础习题精练

习 题

1.10.1 设 $f(x)$，$g(x)$在$[a,b]$上连续，证明至少存在一点 $\xi\in(a,b)$，使得

$$f(\xi)\int_\xi^b g(x)\mathrm{d}x = g(\xi)\int_a^\xi f(x)\mathrm{d}x.$$

1.10.2 设 $f(x)$在$[a,b]$上连续且严格单调增加，证明

$$(a+b)\int_a^b f(x)\mathrm{d}x < 2\int_a^b xf(x)\mathrm{d}x.$$

1.10.3 设 $f(x)$在$[0,1]$上连续且单调减少，证明：当 $0<\lambda<1$ 时，$\int_0^\lambda f(x)\mathrm{d}x \geqslant \lambda\int_0^1 f(x)\mathrm{d}x$.

解 答

1.10.1 **分析** 若令 $G(x)=f(x)\int_x^b g(t)\mathrm{d}t - g(x)\int_a^x f(t)\mathrm{d}t$，无法验证其满足零点定理. 所以应求 $G(x)$ 的原函数，使用罗尔定理证明.

证明 求 $G(x)$的原函数 $F(x)$（对 $G(x)$求不定积分，用分部积分法），有

$$F(x) = \int_a^x f(t)\mathrm{d}t\int_x^b g(t)\mathrm{d}t,$$

因为 $f(x)$，$g(x)$在$[a,b]$上连续，所以 $F(x)$在$[a,b]$上连续，在(a,b)内可导，$F(a)=F(b)=0$，即$F(x)$在$[a,b]$上满足罗尔定理的条件，所以至少存在一点 $\xi\in(a,b)$，使得 $F'(\xi)=0$，即

$$f(\xi)\int_\xi^b g(x)\mathrm{d}x = g(\xi)\int_a^\xi f(x)\mathrm{d}x.$$

1.10.2 分析 $(a+b)\int_a^b f(x)\mathrm{d}x<2\int_a^b xf(x)\mathrm{d}x \Leftrightarrow (a+b)\int_a^b f(x)\mathrm{d}x-2\int_a^b xf(x)\mathrm{d}x<0$，可构造辅助函数，用单调性证明.

证明 令
$$F(t)=(a+t)\int_a^t f(x)\mathrm{d}x-2\int_a^t xf(x)\mathrm{d}x,t\in[a,b],$$
则
$$F'(t)=\int_a^t f(x)\mathrm{d}x+(a+t)f(t)-2tf(t)=\int_a^t f(x)\mathrm{d}x-(t-a)f(t)$$
$$=\int_a^t f(x)\mathrm{d}x-\int_a^t f(t)\mathrm{d}x=\int_a^t[f(x)-f(t)]\mathrm{d}x.$$

因为 $f(x)$ 在$[a,b]$上严格单调增加，所以 $f(x)-f(t)<0$，于是有
$$F'(t)=\int_a^t[f(x)-f(t)]\mathrm{d}x<0,$$
即 $F(t)$ 严格单调减少，又 $F(a)=0$，所以 $F(b)<0$，即
$$(a+b)\int_a^b f(x)\mathrm{d}x-2\int_a^b xf(x)\mathrm{d}x<0,$$
即
$$(a+b)\int_a^b f(x)\mathrm{d}x<2\int_a^b xf(x)\mathrm{d}x.$$

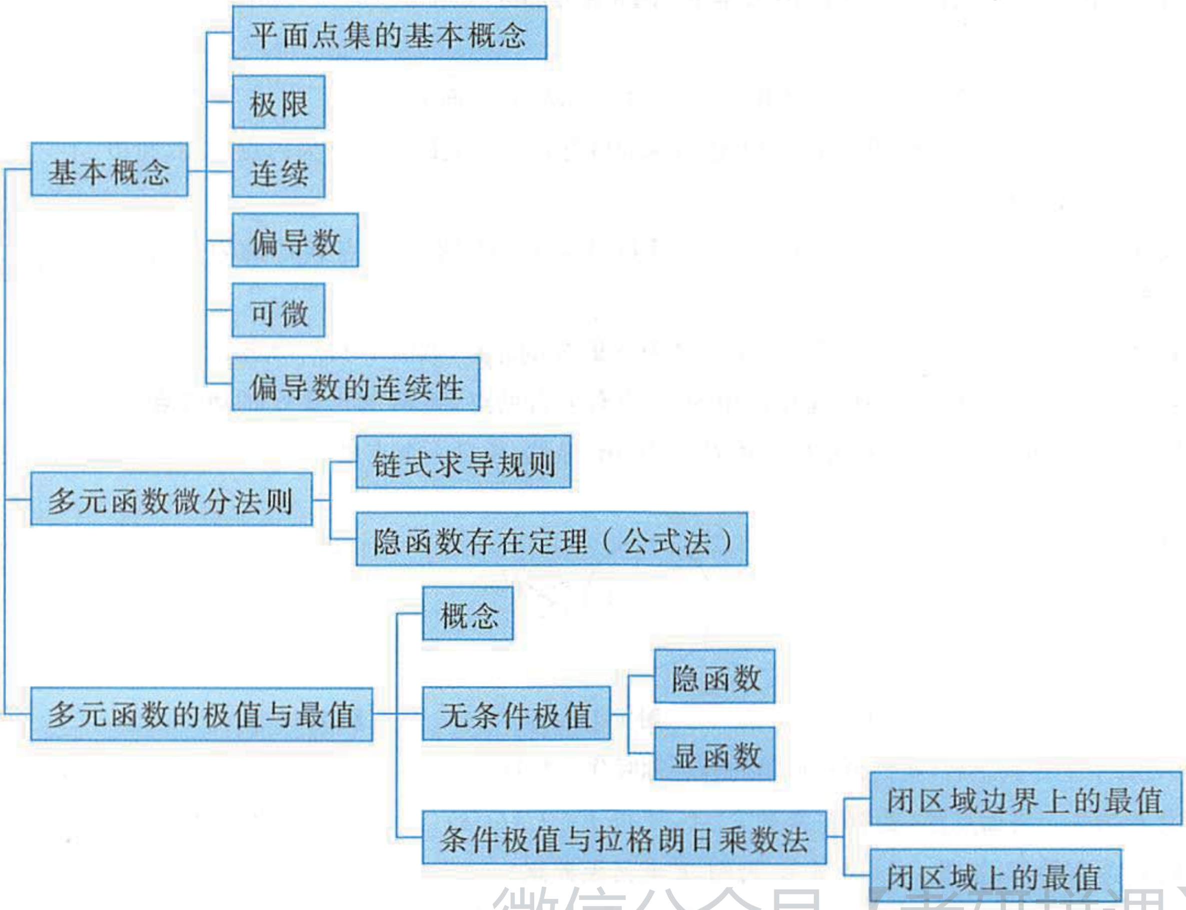
1.10.3 证明 要证原不等式成立，只需证 $\dfrac{\int_0^\lambda f(x)\mathrm{d}x}{\lambda}\geqslant\int_0^1 f(x)\mathrm{d}x$.

令 $F(t)=\dfrac{\int_0^t f(x)\mathrm{d}x}{t}$，由于 $F(\lambda)=\dfrac{\int_0^\lambda f(x)\mathrm{d}x}{\lambda}$，$F(1)=\int_0^1 f(x)\mathrm{d}x$，故只需证当 $\lambda\in(0,1)$ 时，有
$$F(\lambda)\geqslant F(1). \tag{$*$}$$

由 $F(t)$在$(0,1]$内连续，在$(0,1)$内可导，且
$$F'(t)=\frac{f(t)t-\int_0^t f(x)\mathrm{d}x}{t^2}=\frac{f(t)t-f(c)t}{t^2}=\frac{f(t)-f(c)}{t},$$
其中 $0\leqslant c\leqslant t$，知 $f(c)\geqslant f(t)$，有 $F'(t)\leqslant0$，知 $F(t)$ 在$(0,1]$内单调减少. 又 $0<\lambda<1$，有 $F(\lambda)\geqslant F(1)$，即$(*)$式成立，故原不等式成立.

第11讲 多元函数微分学

基础知识结构



基础内容精讲

一、基本概念

1. 平面点集的基本概念

在平面上建立直角坐标系 xOy，则平面上的点就可用两个实数组成的有序数组(x,y)来表示，而二元函数 $f(x,y)$的定义域恰是以两个实数组成的有序数组(x,y)为元素的集合，于是 $f(x,y)$的定义域就是平面上的点集．下面给出平面点集的一些基本概念．

(1)平面上任意两点 $M_1(x_1,y_1)$ 与 $M_2(x_2,y_2)$ 之间的距离定义为

$$\rho(M_1,M_2)=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2},$$

其中记号 $\rho(M_1,M_2)$ 表示点 M_1 与 M_2 的距离，它满足三个要素：

①非负性　$\rho(M_1,M_2)\geqslant 0$；

②对称性　$\rho(M_1,M_2)=\rho(M_2,M_1)$；

③三角不等式　$\rho(M_1,M_3)\leqslant\rho(M_1,M_2)+\rho(M_2,M_3)$.

(2)设 M_0 为平面上的一个点，$\delta>0$，则平面上以点 M_0 为圆心，以 δ 为半径的圆的内部叫作**点 M_0 的 δ 邻域**，记作 $U(M_0,\delta)$（如图 1-11-1），也即

$$U(M_0,\delta)=\{M\mid\rho(M_0,M)<\delta,M\text{ 在平面上}\}.$$

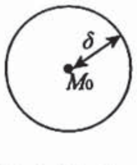


图 1-11-1

若在上述邻域中去掉圆心 M_0，则叫作**点 M_0 的 δ 去心邻域**，记作 $\mathring{U}(M_0,\delta)$（如图 1-11-2），也即

$$\mathring{U}(M_0,\delta)=\{M\mid 0<\rho(M_0,M)<\delta,M\text{ 在平面上}\}.$$

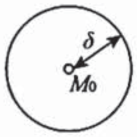


图 1-11-2

(3)给定平面上的一个点集 E，可用上述邻域的概念将平面上的点分类为内点、外点和边界点，下面分别叙述之.

设 M 为平面上的一个点，若存在 $\delta>0$，使得 $U(M,\delta)\subset E$，则 M 为点集 E 的**内点**，如图 1-11-3所示.

若存在 $\delta>0$，使得 $U(M,\delta)\cap E=\varnothing$，则 M 为点集 E 的**外点**，如图 1-11-4 所示.

若对任意的 $\delta>0$，$U(M,\delta)$ 中既有 E 中的点，也有 E 外的点，则 M 为点集 E 的**边界点**，如图 1-11-5 所示. 且 E 的所有边界点的集合称为 E 的**边界**，记作 ∂E. 显然，任意一个点集 E 与它的余集 E^C 有公共边界，即 $\partial E=\partial E^C$.

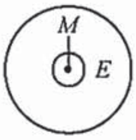


图 1-11-3

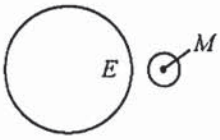


图 1-11-4

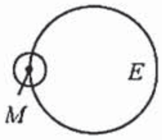


图 1-11-5

(4)有了上述概念后，我们再对点集的各种称呼作一叙述.

设 E 为一个平面点集，若存在常数 $\delta>0$，使得 $E\subset U(O,\delta)$（这里 O 是指坐标原点），则 E 为**有界集**，如图 1-11-6 所示. 否则，E 就是**无界集**.

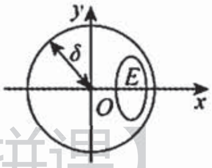


图 1-11-6

若 E 中的每个点都是 E 的内点，则 E 为**开集**；若 E 的边界点都是 E 的点，则 E 为**闭集**. 显然，若一个点集是开集，其余集必是闭集；若一个点集是闭集，其余集必是开集.

设 E 为一个平面点集，若对于 E 中的任意两点，都可用一条完全属于 E 的折线（说成曲线亦可）将这两点连接起来，则这样的 E 为（道路）**连通集**（如图 1-11-7）. 连通的开集叫**开区域**. 一个开区域和它的边界点集的并集叫**闭区域**. 开区域、闭区域统称为**区域**.

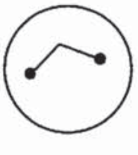


图 1-11-7

若 E 是一个平面区域，且 E 内的任一条**简单闭曲线**的内部还在 E 内，则这样的 E 称为**单连通区域**. 否则就叫**多连通区域**.

(5)还有两个重要概念，放在这部分最后讲.

设 E 是一个平面点集，M_0 为平面上的一个点，若对任意的 $\delta>0$，总有 $\mathring{U}(M_0,\delta)\cap E\neq\varnothing$，即 M_0 的任意邻域中都含有异于 M_0 的 E 中的点，则称 M_0 为 E 的**聚点**. 显然，**非空开集的内点与边界点都是这个点集的聚点，闭区域的任何一点都是它的聚点.**

若存在 $\delta>0$，使得 $U(M_0,\delta)\cap E=\{M_0\}$，即如果 M_0 的某一邻域与点集 E 的交集是一个孤立的点 M_0，则称点 M_0 为 E 的**孤立点**，如图 1-11-8 所示. 显然，**边界点要么是聚点，要么是孤立点**.

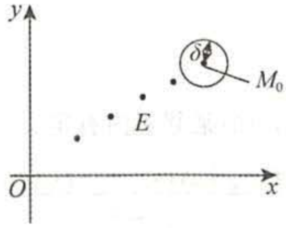


图 1-11-8

以上这些概念细致入微，需要读者潜心品味，反复琢磨，方可掌握牢固.

2. 极限

关于二元函数的极限，有两种定义. 下面第一种定义是大部分数学分析教材从点集角度出发的；第二种定义是大部分高等数学教材从邻域角度出发的.

第一种定义：

定义 1　设二元函数 $f(P)=f(x,y)$ 的定义域为 D，$P_0(x_0,y_0)$ 是 D 的聚点. 如果存在常数 A，对于任意给定的正数 ε，总存在正数 δ，使得当点 $P(x,y)\in D\cap\mathring{U}(P_0,\delta)$ 时，都有

$$|f(x,y)-A|<\varepsilon$$

成立，那么就称常数 A 为函数 $f(x,y)$ 当 $(x,y)\to(x_0,y_0)$ 时的极限，记作

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)=A \text{ 或 } f(x,y)\to A((x,y)\to(x_0,y_0)).$$

以上是按集合论知识（以点集趋向方式）定义多元极限，通俗说来，只要 $f(x,y)$ 是"有定义的"，且满足 $|f(x,y)-A|<\varepsilon$，则 $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)=A$，这里自然"排除"了 (x_0,y_0) 邻域内的无定义点，所以

$$\lim_{(x,y)\to(0,0)}\frac{\sqrt{xy+1}-1}{xy}=\lim_{(x,y)\to(0,0)}\frac{xy+1-1}{xy(\sqrt{xy+1}+1)}=\lim_{(x,y)\to(0,0)}\frac{1}{\sqrt{xy+1}+1}=\frac{1}{2}.$$

第二种定义：

定义 1′　若二元函数 $f(x,y)$ 在 (x_0,y_0) 的去心邻域内有定义，且 (x,y) 以任意方式趋向于 (x_0,y_0) 时，$f(x,y)$ 均趋向于 A，则 $\lim\limits_{\substack{x\to x_0\\y\to y_0}} f(x,y)=A$.

根据定义 1′，由于函数 $f(x,y)=\dfrac{\sqrt{xy+1}-1}{xy}$ 在原点的邻域内的坐标轴上处处无定义，如图 1-11-9 所示，于是 $\lim\limits_{\substack{x\to 0\\y\to 0}}\dfrac{\sqrt{xy+1}-1}{xy}$ 不存在.

按照两种定义，你会得出

$$\lim_{\substack{x\to 0\\y\to 0}}(x^2+y^2)\sin\frac{1}{x+y}=\begin{cases}0, & \text{（第一种定义），}\\ \text{不存在，} & \text{（第二种定义）.}\end{cases}$$

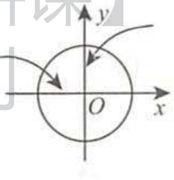


图 1-11-9

所以，为避免这种教材定义不同导致的"矛盾"，命题人目前处理得比较恰当，只考这种无论在哪种定义下极限都存在或都不存在的函数，比如：$\lim\limits_{\substack{x\to 0\\y\to 0}}(x^2+y^2)\sin\dfrac{1}{x^2+y^2}=0$.

3. 连续

如果 $\lim\limits_{\substack{x\to x_0\\y\to y_0}} f(x,y)=f(x_0,y_0)$，则称 $f(x,y)$ 在点 (x_0,y_0) 处**连续**.

【注】 验证二元函数 $f(x,y)$ 在某一点 (x_0,y_0) 处是否连续是考研的重点，但是如果不连续，对于多元函数是不讨论间断点类型的.

4. 偏导数

定义 2 设函数 $z=f(x,y)$ 在点 (x_0,y_0) 的某邻域内有定义. 若极限

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}$$

存在，则称此极限为函数 $z=f(x,y)$ 在点 (x_0,y_0) 处对 x 的**偏导数**，记作

$$\left.\frac{\partial z}{\partial x}\right|_{\substack{x=x_0\\y=y_0}},\quad \left.\frac{\partial f}{\partial x}\right|_{\substack{x=x_0\\y=y_0}},\quad \left.z'_x\right|_{\substack{x=x_0\\y=y_0}}\text{ 或 }f'_x(x_0,y_0).$$

于是

$$f'_x(x_0,y_0)=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}=\lim_{x\to x_0}\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0},$$

$$f'_y(x_0,y_0)=\lim_{\Delta y\to 0}\frac{f(x_0,y_0+\Delta y)-f(x_0,y_0)}{\Delta y}=\lim_{y\to y_0}\frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}.$$

如果函数 $z=f(x,y)$ 在区域 D 内的偏导数 $f'_x(x,y)$，$f'_y(x,y)$ 仍具有偏导数，则它们的偏导数称为函数 $z=f(x,y)$ 的**二阶偏导数**. 按照对变量求导次序的不同，有如下四个二阶偏导数：

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial^2 z}{\partial x^2}=f''_{xx}(x,y),\quad \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=\frac{\partial^2 z}{\partial x\partial y}=f''_{xy}(x,y),$$

$$\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=\frac{\partial^2 z}{\partial y\partial x}=f''_{yx}(x,y),\quad \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)=\frac{\partial^2 z}{\partial y^2}=f''_{yy}(x,y),$$

其中 $f''_{xy}(x,y)$，$f''_{yx}(x,y)$ 称为**混合偏导数**. 同样可得三阶、四阶以及 n 阶偏导数. 二阶及二阶以上的偏导数统称为高阶偏导数.

5. 可微

先看一个引例. 如图 1-11-10 所示，设矩形的长和宽分别为 x 和 y，则此矩形的面积 $S=xy$. 若 x,y 分别增长 $\Delta x,\Delta y$，则该矩形面积的增量为

$$\Delta S=(x+\Delta x)(y+\Delta y)-xy=y\Delta x+x\Delta y+\Delta x\Delta y.$$

上式右端由两部分组成：一部分是 $y\Delta x+x\Delta y$，它是关于 $\Delta x,\Delta y$ 的线性函数；另一部分是 $\Delta x\Delta y$，因为

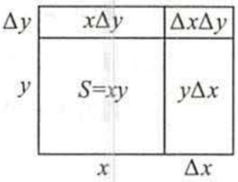


图 1-11-10

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\Delta x\Delta y}{\sqrt{(\Delta x)^2+(\Delta y)^2}}\leqslant\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\frac{(\Delta x)^2+(\Delta y)^2}{2}}{\sqrt{(\Delta x)^2+(\Delta y)^2}}=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\sqrt{(\Delta x)^2+(\Delta y)^2}}{2}=0,$$

所以当 $\Delta x\to 0,\Delta y\to 0$ 时，$\Delta x\Delta y$ 是比 $\rho=\sqrt{(\Delta x)^2+(\Delta y)^2}$ 高阶的无穷小量，即

$$\Delta S=y\Delta x+x\Delta y+o(\rho)\ (\rho\to 0).$$

$y\Delta x+x\Delta y$ 是 ΔS 的主要部分，$o(\rho)$ 是误差，

$$\Delta S\approx y\Delta x+x\Delta y.$$

称 $y\Delta x+x\Delta y$ 为函数 $S=xy$ 在点 (x,y) 处的全微分.

定义 3 如果函数 $z=f(x,y)$ 在点 (x,y) 的全增量 $\Delta z=f(x+\Delta x,y+\Delta y)-f(x,y)$ 可表示为

$$\Delta z=A\Delta x+B\Delta y+o(\rho),$$

其中 $\rho=\sqrt{(\Delta x)^2+(\Delta y)^2}$，$A,B$ 不依赖于 $\Delta x,\Delta y$ 而仅与 x,y 有关，则称函数 $z=f(x,y)$ 在点 (x,y) 可微，而称 $A\Delta x+B\Delta y$ 为函数 $z=f(x,y)$ 在点 (x,y) 的全微分，记作 $\mathrm{d}z$，即 $\mathrm{d}z=A\Delta x+B\Delta y$.

在第 4 讲中，已经详细阐述了一元函数可微的深刻含义，二元函数的可微概念也是如此（请注意对比，加深理解）. 判断函数 $z=f(x,y)$ 是否可微，步骤如下：

①写出全增量 $\Delta z=f(x_0+\Delta x,y_0+\Delta y)-f(x_0,y_0)$；

②写出线性增量 $A\Delta x+B\Delta y$，其中 $A=f'_x(x_0,y_0)$，$B=f'_y(x_0,y_0)$；

③作极限 $\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\dfrac{\Delta z-(A\Delta x+B\Delta y)}{\sqrt{(\Delta x)^2+(\Delta y)^2}}$，若该极限等于 0，则 $z=f(x,y)$ 在点 (x_0,y_0) 可微，否则，就不可微.

用形式简单的"线性增量 $A\Delta x+B\Delta y$"去代替形式复杂的"全增量 Δz"，且其误差"$\Delta z-(A\Delta x+B\Delta y)$"是 $o(\sqrt{(\Delta x)^2+(\Delta y)^2})$. 这就是说，用简单的代替了复杂的，且产生的误差可以忽略不计，这就是可微的真正含义. 在实际问题中，有时直接计算 Δz，有时算出 $A\Delta x+B\Delta y$ 来近似 Δz.

6. 偏导数的连续性

对于 $z=f(x,y)$，讨论其在某特殊点 (x_0,y_0)（比如二元分段函数的分段点）处偏导数是否连续，是考研的重点，其步骤为：

①用定义法求 $f'_x(x_0,y_0)$，$f'_y(x_0,y_0)$；

②用公式法求 $f'_x(x,y)$，$f'_y(x,y)$；

③计算 $\lim\limits_{\substack{x\to x_0\\ y\to y_0}}f'_x(x,y)$，$\lim\limits_{\substack{x\to x_0\\ y\to y_0}}f'_y(x,y)$.

看 $\lim\limits_{\substack{x\to x_0\\ y\to y_0}}f'_x(x,y)=f'_x(x_0,y_0)$，$\lim\limits_{\substack{x\to x_0\\ y\to y_0}}f'_y(x,y)=f'_y(x_0,y_0)$ 是否成立. 若成立，则 $z=f(x,y)$ 在点 (x_0,y_0) 处的偏导数是连续的.

二、多元函数微分法则

1. 链式求导规则

（1）复合函数的中间变量均为一元函数的情形. 复合结构图如图 1-11-11 所示.

设 $z=f(u,v)$，$u=\varphi(t)$，$v=\psi(t)$，则 $z=f[\varphi(t),\psi(t)]$，且 $\dfrac{\mathrm{d}z}{\mathrm{d}t}=\dfrac{\partial z}{\partial u}\dfrac{\mathrm{d}u}{\mathrm{d}t}+\dfrac{\partial z}{\partial v}\dfrac{\mathrm{d}v}{\mathrm{d}t}$.

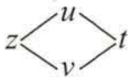


图 1-11-11

（2）复合函数的中间变量均为多元函数的情形. 复合结构图如图 1-11-12 所示.

设 $z=f(u,v)$，$u=\varphi(x,y)$，$v=\psi(x,y)$，则 $z=f[\varphi(x,y),\psi(x,y)]$，且

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x},\quad \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y}.$$

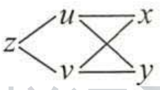


图 1-11-12

（3）复合函数的中间变量既有一元函数，又有多元函数的情形. 复合结构图如图 1-11-13所示.

设 $z=f(u,v)$，$u=\varphi(x,y)$，$v=\psi(y)$，则 $z=f[\varphi(x,y),\psi(y)]$，且

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x},\quad \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}y}.$$

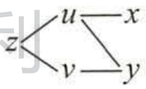


图 1-11-13

【注】 无论 z 对谁求导，也无论 z 已经求了几阶导，求导后的新函数仍然具有与原函数完全相同的复合结构.

2. 隐函数存在定理（公式法）

在本书第 1 讲提出的函数定义中，对每个 $x\in D$，对应的函数值 y 总是唯一的，这样定义的函数称为

单值函数. 如果给定一个对应法则，按这个法则，对每个 $x \in D$，总有确定的 y 值与之对应，但这个 y 不是唯一的，于是，这样的对应法则就不符合函数的定义了，我们称这种法则确定了一个**多值函数**. 在考研中所提到的函数是指单值函数，也就是当自变量 x 取一个值时，这个对应法则 f 要保证因变量 y 有唯一的实数值与之对应，否则就必须分成若干个单值函数去研究，比如下述的“隐函数存在问题”.

一般说来，只要在满足定义域的条件下，形如 $y=f(x)$ 的函数称为显函数，例如 $y=\sin x$；由方程 $F(x,y)=0$ 所确定的函数称为隐函数，例如由方程 $x+y^3-1=0$ 确定的隐函数（其显式表示为 $y=\sqrt[3]{1-x}$）. 但“隐函数存在”是有前提的，大多数教材都是这样的提法：“如果 x 与 y 满足方程 $F(x,y)=0$，那么在一定条件下，当 x 取某区间内的任一值时，相应地总有满足这个方程的唯一的 y 值存在，那么就说方程 $F(x,y)=0$ 在该区间内确定了一个隐函数.”读者是否注意到“在一定条件下”和“唯一的 y 值”这样的语句？请看下面的两个定理.

隐函数存在定理 1 设函数 $F(x,y)$ 在点 $P(x_0,y_0)$ 的某一邻域内具有连续偏导数，$F(x_0,y_0)=0$，$F'_y(x_0,y_0)\neq 0$，则方程 $F(x,y)=0$ 在点 (x_0,y_0) 的某一邻域内能唯一确定一个连续且具有连续导数的函数 $y=f(x)$，它满足条件 $y_0=f(x_0)$，并有 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{F'_x}{F'_y}$.

这里的 $F'_y(x_0,y_0)\neq 0$（也就是 $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ 存在）是定理的关键. 由此看来，所谓的“隐函数存在”，是要求在一个“指定的位置”，方程 $F(x,y)=0$ 能确定一个“不仅有意义，而且要有可导这种良好性质的函数”. 而在一个指定位置处可导的函数必然首先得是单值的.

举个例子供读者理解，给出方程 $x^2+y^2-1=0$，设 $F(x,y)=x^2+y^2-1$，则 $F'_x=2x$，$F'_y=2y$，$F(0,1)=0$，$F'_y(0,1)=2\neq 0$，由上述隐函数存在定理 1 可知，方程 $x^2+y^2-1=0$ 在点 $(0,1)$ 的某一邻域内能确定一个有连续导数的隐函数 $y=f(x)$；而在点 $(-1,0)$ 和点 $(1,0)$ 的邻域内就不存在这样一个有连续导数的隐函数，因为在点 $(-1,0)$ 和点 $(1,0)$ 处的切线都是竖直方向的，显然导数不存在，在这两个点的任何去心邻域中，一个 x 对应着两个 y 的值，这就不符合函数定义了（如图 1-11-14）.

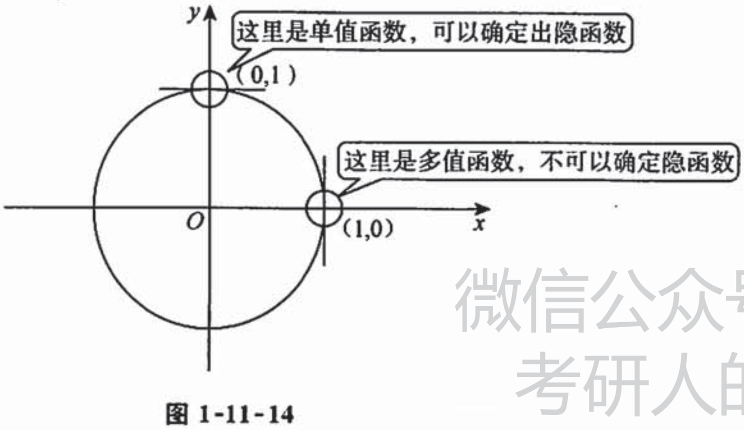


图 1-11-14

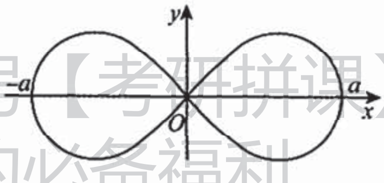


图 1-11-15

再看个典型的例子，对于伯努利双纽线（如图 1-11-15），设 $F(x,y)=(x^2+y^2)^2-a^2(x^2-y^2)=0$，在点 $(0,0)$，$(a,0)$，$(-a,0)$ 处，隐函数不存在；而在其他位置，隐函数都存在.

这里还要指出的是，定理最后给出的 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{F'_x}{F'_y}$，这个公式就是**隐函数求导公式**，证明如下：

将 $y=f(x)$ 代入 $F(x,y)=0$，得恒等式 $F(x,f(x))\equiv 0$，在这个恒等式两边对 x 求导，得 $\dfrac{\partial F}{\partial x}+\dfrac{\partial F}{\partial y}\cdot\dfrac{\mathrm{d}y}{\mathrm{d}x}=0$，由于 F'_y 连续且 $F'_y(x_0,y_0)\neq 0$，所以存在 (x_0,y_0) 的一个邻域，在这个邻域内 $F'_y\neq 0$，于是得 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{F'_x}{F'_y}$.

于是，很自然地，可以把隐函数存在定理推广到多元函数。既然一个二元方程$F(x,y)=0$有可能确定一个一元隐函数，那么一个三元方程$F(x,y,z)=0$也就有可能确定一个二元隐函数。

隐函数存在定理 2　设函数$F(x,y,z)$在点$P(x_0,y_0,z_0)$的某一邻域内具有连续偏导数，且$F(x_0,y_0,z_0)=0$，$F'_z(x_0,y_0,z_0)\neq 0$，则方程$F(x,y,z)=0$在点(x_0,y_0,z_0)的某一邻域内能唯一确定一个连续且具有连续偏导数的函数$z=f(x,y)$，它满足条件$z_0=f(x_0,y_0)$，并有

$$\frac{\partial z}{\partial x}=-\frac{F'_x}{F'_z},\quad \frac{\partial z}{\partial y}=-\frac{F'_y}{F'_z}.$$

此公式的证明同样简单，将$z=f(x,y)$代入$F(x,y,z)=0$，得$F(x,y,f(x,y))\equiv 0$，式子两端分别对x和y求偏导数，得$F'_x+F'_z\cdot\dfrac{\partial z}{\partial x}=0$，$F'_y+F'_z\cdot\dfrac{\partial z}{\partial y}=0$。

因为F'_z连续且$F'_z(x_0,y_0,z_0)\neq 0$，所以存在点(x_0,y_0,z_0)的一个邻域，使$F'_z\neq 0$，于是得

$$\frac{\partial z}{\partial x}=-\frac{F'_x}{F'_z},\quad \frac{\partial z}{\partial y}=-\frac{F'_y}{F'_z}.$$

同理，$F'_z(x_0,y_0,z_0)\neq 0$（也就是$\dfrac{\partial z}{\partial x}$，$\dfrac{\partial z}{\partial y}$都存在且连续）是定理的关键。

请看2005年考研数学一真题第(10)题，也就是下面的注例。

注例　设有三元方程$xy-z\ln y+e^{xz}=1$，根据隐函数存在定理，存在点$(0,1,1)$的一个邻域，在此邻域内该方程（　　）。

(A)只能确定一个具有连续偏导数的隐函数$z=z(x,y)$

(B)可确定两个具有连续偏导数的隐函数$y=y(x,z)$和$z=z(x,y)$

(C)可确定两个具有连续偏导数的隐函数$x=x(y,z)$和$z=z(x,y)$

(D)可确定两个具有连续偏导数的隐函数$x=x(y,z)$和$y=y(x,z)$

解　应选(D)。

令

$$F(x,y,z)=xy-z\ln y+e^{xz}-1,$$

则$F'_x=y+e^{xz}z$，$F'_y=x-\dfrac{z}{y}$，$F'_z=-\ln y+e^{xz}x$。于是，

$$F'_x(0,1,1)=2\neq 0,\quad F'_y(0,1,1)=-1\neq 0,\quad F'_z(0,1,1)=0.$$

因此，在点$(0,1,1)$的某一个邻域U_1内，存在一个连续且具有连续偏导数的隐函数$x=x(y,z)$；在点$(0,1,1)$的某一个邻域U_2内，也存在一个连续且具有连续偏导数的隐函数$y=y(x,z)$。于是，在邻域$U=U_1\cap U_2$内，就存在两个连续且具有连续偏导数的隐函数$x=x(y,z)$和$y=y(x,z)$，故应选(D)。

三、多元函数的极值与最值

1. 概念

定义 4　若存在(x_0,y_0)的某个邻域，使得在该邻域内任意一点(x,y)，均有

$$f(x,y)\leqslant f(x_0,y_0)\,(\text{或 } f(x,y)\geqslant f(x_0,y_0))$$

成立，则称(x_0,y_0)为$f(x,y)$的**广义的极大值点（或极小值点）**，$f(x_0,y_0)$为$f(x,y)$的**广义的极大值（或极小值）**。

定义 5　若存在(x_0,y_0)的某个去心邻域，使得对于该邻域内任意一点(x,y)，均有

$$f(x,y)<f(x_0,y_0)(\text{或 } f(x,y)>f(x_0,y_0))$$

成立，则称(x_0,y_0)为$f(x,y)$的**真正的极大值点**(或**极小值点**)，$f(x_0,y_0)$为$f(x,y)$的**真正的极大值**(或**极小值**).

【注】 与一元函数极值的定义类似，若在上述定义5(真正的极值)中把去心邻域改为邻域，则当然有$f(x,y)\leqslant f(x_0,y_0)$，但这个等号仅在$(x_0,y_0)$处取到，这一点是与定义4(广义的极值)的一个主要区别，请注意区分.

定义6 设(x_0,y_0)为$f(x,y)$定义域内一点，若对于$f(x,y)$的定义域内任意一点(x,y)，均有

$$f(x,y)\leqslant f(x_0,y_0)(\text{或 } f(x,y)\geqslant f(x_0,y_0))$$

成立，则称$f(x_0,y_0)$为$f(x,y)$的**广义的最大值**(或**最小值**).

定义7 设(x_0,y_0)为$f(x,y)$定义域内一点，若对于$f(x,y)$的定义域内任意一个异于(x_0,y_0)的点(x,y)，均有

$$f(x,y)<f(x_0,y_0)(\text{或 } f(x,y)>f(x_0,y_0))$$

成立，则称$f(x_0,y_0)$为$f(x,y)$的**真正的最大值**(或**最小值**).

【注】 若题目只问极值(或最值)，未指明是广义的极值(或最值)还是真正的极值(或最值)，读者按广义的极值(或最值)来解答.

2. 无条件极值

(1)二元函数取极值的必要条件(类比一元函数).

设$z=f(x,y)$在点$(x_0,y_0)\begin{cases}\text{一阶偏导数存在,}\\ \text{取极值,}\end{cases}$则$f'_x(x_0,y_0)=0,f'_y(x_0,y_0)=0$.

【注】 该必要条件同样适用于三元及三元以上函数.

(2)二元函数取极值的充分条件.

$$\text{记}\begin{cases}f''_{xx}(x_0,y_0)=A,\\ f''_{xy}(x_0,y_0)=B,\text{则 }\Delta=B^2-AC\\ f''_{yy}(x_0,y_0)=C,\end{cases}\begin{cases}<0\Rightarrow\text{极值}\begin{cases}A<0\Rightarrow\text{极大值,}\\ A>0\Rightarrow\text{极小值,}\end{cases}\\ >0\Rightarrow\text{非极值,}\\ =0\Rightarrow\text{方法失效，另谋他法.(方法失效怎么办？见例1.11.10)}\end{cases}$$

【注】 该充分条件不适用于三元及三元以上函数.

综合(1)，(2)可用必要条件求出可疑点，用充分条件判别可疑点.

3. 条件极值与拉格朗日乘数法

求目标函数$u=f(x,y,z)$在条件$\begin{cases}\varphi(x,y,z)=0,\\ \psi(x,y,z)=0\end{cases}$下的最值，则

①构造辅助函数$F(x,y,z,\lambda,\mu)=f(x,y,z)+\lambda\varphi(x,y,z)+\mu\psi(x,y,z)$；

②令

$$\begin{cases} F_x' = f_x' + \lambda\varphi_x' + \mu\psi_x' = 0, \\ F_y' = f_y' + \lambda\varphi_y' + \mu\psi_y' = 0, \\ F_z' = f_z' + \lambda\varphi_z' + \mu\psi_z' = 0, \\ F_\lambda' = \varphi(x,y,z) = 0, \\ F_\mu' = \psi(x,y,z) = 0; \end{cases}$$

③解上述方程组得备选点 $P_i, i=1,2,3,\cdots,n$，并求 $f(P_i)$，取其最大值为 $u_{\max}$，最小值为 $u_{\min}$；

④根据实际问题，必存在最值，所得即为所求.

基础例题精解

一、基本概念

例 1.11.1 已知函数 $f(x,y)=x\ln(x+\ln y)$，求 $f_x'(1,\mathrm{e}), f_y'(1,\mathrm{e})$.

解 把 y 看成常数，利用乘积函数的求导法则，对 x 求导，得

$$f_x'(x,y)=\ln(x+\ln y)+\frac{x}{x+\ln y},$$

将$(1,\mathrm{e})$代入，得

$$f_x'(1,\mathrm{e})=\ln(1+\ln \mathrm{e})+\frac{1}{1+\ln \mathrm{e}}=\ln 2+\frac{1}{2}.$$

把 x 看成常数，对 y 求导，得

$$f_y'(x,y)=x\cdot\frac{1}{x+\ln y}\cdot\frac{1}{y}=\frac{x}{xy+y\ln y},$$

将$(1,\mathrm{e})$代入，得

$$f_y'(1,\mathrm{e})=\frac{1}{1\cdot \mathrm{e}+\mathrm{e}\ln \mathrm{e}}=\frac{1}{2\mathrm{e}}.$$

例 1.11.2 设 $z=f(x,y)=\begin{cases}(x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0,\\ 0, & x^2+y^2=0,\end{cases}$ 则下列四个结论中，

①$f(x,y)$在$(0,0)$处连续；　②$f_x'(0,0)$，$f_y'(0,0)$存在；

③$f_x'(x,y)$，$f_y'(x,y)$在$(0,0)$处连续；　④$f(x,y)$在$(0,0)$处可微.

正确结论的个数为(　　).

(A)1　　(B)2　　(C)3　　(D)4

解 应选(C).

对于结论①，根据连续定义，

$$\lim_{\substack{x\to 0\\ y\to 0}}(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}=\lim_{\rho\to 0}\rho^2\sin\frac{1}{\rho}=0,\quad \lim_{\substack{x\to 0\\ y\to 0}}f(x,y)=f(0,0),$$

①正确；

对于结论②，用定义法，$f_x'(0,0)=\lim\limits_{\Delta x\to 0}\dfrac{f(\Delta x,0)-f(0,0)}{\Delta x}=\lim\limits_{\Delta x\to 0}\dfrac{(\Delta x)^2\sin\dfrac{1}{|\Delta x|}-0}{\Delta x}=0$，同理可得

$f'_y(0,0)=0$，②正确；

对于结论③，当$(x,y)\neq(0,0)$时，用公式法，

$$f'_x(x,y)=2x\sin\frac{1}{\sqrt{x^2+y^2}}+(x^2+y^2)\cos\frac{1}{\sqrt{x^2+y^2}}\cdot\left[-\frac{1}{2}\cdot\frac{2x}{\sqrt{(x^2+y^2)^3}}\right]$$

$$=2x\sin\frac{1}{\sqrt{x^2+y^2}}-\frac{x}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}}.$$

因为当$(x,y)\to(0,0)$时，$2x\sin\frac{1}{\sqrt{x^2+y^2}}\to0$，$\frac{x}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}}$不存在，故$\lim\limits_{\substack{x\to0\\y\to0}}f'_x(x,y)$不存在，所以$f'_x(x,y)$在$(0,0)$处不连续．同理，$f'_y(x,y)$在$(0,0)$处也不连续，所以③不正确；

对于结论④，$$f'_x(0,0)=0,\quad f'_y(0,0)=0,$$

$$\Delta z=f(0+\Delta x,0+\Delta y)-f(0,0)=[(\Delta x)^2+(\Delta y)^2]\cdot\sin\frac{1}{\sqrt{(\Delta x)^2+(\Delta y)^2}}=\rho^2\sin\frac{1}{\rho},$$

$$\lim_{\substack{\Delta x\to0\\\Delta y\to0}}\frac{\Delta z-f'_x(0,0)\Delta x-f'_y(0,0)\Delta y}{\sqrt{(\Delta x)^2+(\Delta y)^2}}=\lim_{\rho\to0}\frac{\rho^2\sin\frac{1}{\rho}}{\rho}=\lim_{\rho\to0}\rho\sin\frac{1}{\rho}=0,$$

其中$\rho=\sqrt{(\Delta x)^2+(\Delta y)^2}$．故$f(x,y)$在$(0,0)$处可微，所以④正确，答案选择(C)．

二、多元函数微分法则

1. 复合函数求偏导问题

例 1.11.3 设$z=f(e^x\sin y,x^2+y^2)$，其中f具有二阶连续偏导数，求$\frac{\partial^2 z}{\partial x\partial y}$．

解
$$\frac{\partial z}{\partial x}=e^x\sin yf'_1+2xf'_2,$$

$$\frac{\partial^2 z}{\partial x\partial y}=f''_{11}e^{2x}\sin y\cos y+2e^x(y\sin y+x\cos y)f''_{12}+4xyf''_{22}+f'_1e^x\cos y.$$

例 1.11.4 设$z=x^3f\left(xy,\frac{y}{x}\right)$，$f$具有连续的二阶偏导数，求$\frac{\partial z}{\partial y}$，$\frac{\partial^2 z}{\partial y^2}$及$\frac{\partial^2 z}{\partial x\partial y}$．

解
$$\frac{\partial z}{\partial y}=x^4f'_1+x^2f'_2,$$

$$\frac{\partial^2 z}{\partial y^2}=x^4\left(xf''_{11}+\frac{1}{x}f''_{12}\right)+x^2\left(xf''_{21}+\frac{1}{x}f''_{22}\right)=x^5f''_{11}+2x^3f''_{12}+xf''_{22},$$

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial^2 z}{\partial y\partial x}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=4x^3f'_1+x^4\left(yf''_{11}-\frac{y}{x^2}f''_{12}\right)+2xf'_2+x^2\left(yf''_{21}-\frac{y}{x^2}f''_{22}\right)$$

$$=4x^3f'_1+2xf'_2+x^4yf''_{11}-yf''_{22}.$$

2. 隐函数求偏导问题

例 1.11.5 设函数$y=y(x)$由方程$\sin xy-\frac{1}{y-x}=1$确定，求$\left.\frac{dy}{dx}\right|_{x=0}$．

解 设$F(x,y)=\sin xy-\frac{1}{y-x}-1$，则

$$F'_x=y\cos xy-\frac{1}{(y-x)^2},\quad F'_y=x\cos xy+\frac{1}{(y-x)^2}.$$

将 $x=0$ 代入方程 $\sin xy-\dfrac{1}{y-x}=1$，得 $y=-1$.

又因为

$$F'_x(0,-1)=(-1)\cdot\cos 0-\frac{1}{(-1-0)^2}=-2,$$

$$F'_y(0,-1)=0\cdot\cos 0+\frac{1}{(-1-0)^2}=1,$$

所以 $\left.\dfrac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=0}=-\dfrac{F'_x(0,-1)}{F'_y(0,-1)}=2.$

例 1.11.6 设 $y=y(x)$，$z=z(x)$ 是由方程 $z=xf(x+y)$ 和 $F(x,y,z)=0$ 所确定的函数，其中 f 和 F 分别具有一阶连续导数和一阶连续偏导数，求 $\dfrac{\mathrm{d}z}{\mathrm{d}x}$.

解 分别在 $z=xf(x+y)$ 和 $F(x,y,z)=0$ 的两边对 x 求导，得

$$\begin{cases}\dfrac{\mathrm{d}z}{\mathrm{d}x}=f+x\left(1+\dfrac{\mathrm{d}y}{\mathrm{d}x}\right)f',\\ F'_x+F'_y\dfrac{\mathrm{d}y}{\mathrm{d}x}+F'_z\dfrac{\mathrm{d}z}{\mathrm{d}x}=0\end{cases}\Rightarrow\begin{cases}-xf'\dfrac{\mathrm{d}y}{\mathrm{d}x}+\dfrac{\mathrm{d}z}{\mathrm{d}x}=f+xf',\\ F'_y\dfrac{\mathrm{d}y}{\mathrm{d}x}+F'_z\dfrac{\mathrm{d}z}{\mathrm{d}x}=-F'_x.\end{cases}$$

由此解得 $\dfrac{\mathrm{d}z}{\mathrm{d}x}=\dfrac{(f+xf')F'_y-xf'F'_x}{F'_y+xf'F'_z}(F'_y+xf'F'_z\neq 0).$

例 1.11.7 设 $z=f(x,y)$ 是由方程 $z-y-x+xe^{z-y-x}=0$ 所确定的二元函数，求 $\mathrm{d}z$.

解 **方法一** 偏导数法（视二元函数为一元函数）.

这是由隐函数确定的二元函数，对方程两边求偏导数（求 $\dfrac{\partial z}{\partial x}$ 时，x 是自变量，z 是因变量，y 是常数；求 $\dfrac{\partial z}{\partial y}$ 时，y 是自变量，z 是因变量，x 是常数）得

$$\frac{\partial z}{\partial x}-1+e^{z-y-x}+xe^{z-y-x}\left(\frac{\partial z}{\partial x}-1\right)=0,\quad \frac{\partial z}{\partial y}-1+xe^{z-y-x}\left(\frac{\partial z}{\partial y}-1\right)=0,$$

解得

$$\frac{\partial z}{\partial x}=\frac{1+(x-1)e^{z-y-x}}{1+xe^{z-y-x}},\quad \frac{\partial z}{\partial y}=1,$$

所以

$$\mathrm{d}z=\frac{1+(x-1)e^{z-y-x}}{1+xe^{z-y-x}}\mathrm{d}x+\mathrm{d}y.$$

方法二 全微分法.

等式两边求微分得 $\mathrm{d}z-\mathrm{d}y-\mathrm{d}x+e^{z-y-x}\mathrm{d}x+xe^{z-y-x}(\mathrm{d}z-\mathrm{d}y-\mathrm{d}x)=0$,

整理得

$$\mathrm{d}z=\frac{1+(x-1)e^{z-y-x}}{1+xe^{z-y-x}}\mathrm{d}x+\mathrm{d}y.$$

方法三 公式法.

设

$$F(x,y,z)=z-y-x+xe^{z-y-x},$$

则

$$F'_x=-1+e^{z-y-x}-xe^{z-y-x},\quad F'_y=-1-xe^{z-y-x},\quad F'_z=1+xe^{z-y-x},$$

于是有

$$\frac{\partial z}{\partial x}=-\frac{F'_x}{F'_z}=\frac{1+(x-1)e^{z-y-x}}{1+xe^{z-y-x}},\quad \frac{\partial z}{\partial y}=-\frac{F'_y}{F'_z}=1,$$

所以

$$\mathrm{d}z=\frac{1+(x-1)e^{z-y-x}}{1+xe^{z-y-x}}\mathrm{d}x+\mathrm{d}y.$$

3. 逆问题

例 1.11.8 求方程 $\dfrac{\partial^2 z}{\partial x\partial y}=x+y$ 满足条件 $z(x,0)=x$，$z(0,y)=y^2$ 的解 $z=z(x,y)$.

解 将$\frac{\partial^2 z}{\partial x\partial y}=x+y$两边对 y 积分，得

$$\frac{\partial z}{\partial x}=xy+\frac{1}{2}y^2+\varphi_1(x); \qquad ①$$

①式两边对 x 积分，得

$$z=\frac{1}{2}x^2y+\frac{1}{2}xy^2+\varphi(x)+\psi(y). \qquad ②$$

现在来确定 $\varphi(x)$ 和 $\psi(y)$.

由已知 $z(x,0)=x,z(0,y)=y^2$，代入②式有 $x=\varphi(x)+\psi(0),y^2=\varphi(0)+\psi(y)$，于是

$$z=x+y^2+\frac{1}{2}x^2y+\frac{1}{2}xy^2-[\varphi(0)+\psi(0)].$$

又因 $z(0,0)=0$，故 $\varphi(0)+\psi(0)=0$，故

$$z=z(x,y)=x+y^2+\frac{1}{2}x^2y+\frac{1}{2}xy^2.$$

三、多元函数的极值与最值

1. 无条件极值——隐函数

例 1.11.9 设 $z=z(x,y)$是由

$$x^2-6xy+10y^2-2yz-z^2+18=0 \qquad (*)$$

所确定的. 试求 $z=z(x,y)$的极值点与极值.

解 用必要条件求出可疑点 P，即$\begin{cases}z'_x=0,\\z'_y=0\end{cases}\Rightarrow P.$

对(*)式两边依次求 x,y 的偏导数为

$$\begin{cases}2x-6y-2yz'_x-2zz'_x=0,\\-6x+20y-2z-2yz'_y-2zz'_y=0,\end{cases}$$

整理得

$$\begin{cases}x-3y-(y+z)z'_x=0, & (**)\\-3x+10y-z-(y+z)z'_y=0, & (***)\end{cases}$$

令$\begin{cases}z'_x=0,\\z'_y=0,\end{cases}$得$\begin{cases}x-3y=0,\\-3x+10y-z=0,\end{cases}$即$\begin{cases}x=3y,\\z=y,\end{cases}$将其代入(*)式得到可疑点

$$P_1(9,3),\quad z_1=3;\quad P_2(-9,-3),\quad z_2=-3.$$

用充分条件判别可疑点.

对(**)式再求 x 的偏导数为 $1-(z'_x)^2-(y+z)z''_{xx}=0$,

对(**)式再求 y 的偏导数为 $-3-(1+z'_y)z'_x-(y+z)z''_{xy}=0$,

对(***)式再求 y 的偏导数为 $10-z'_y-(1+z'_y)z'_y-(y+z)z''_{yy}=0$,

又结合$\begin{cases}z'_x=0,\\z'_y=0\end{cases}\Rightarrow z''_{xx}=\frac{1}{y+z},z''_{xy}=\frac{-3}{y+z},z''_{yy}=\frac{10}{y+z}.$

对于 $P_1(9,3),z_1=3$，记$\begin{cases}A_1=z''_{xx}\Big|_{(9,3)}=\frac{1}{6},\\B_1=z''_{xy}\Big|_{(9,3)}=-\frac{1}{2},\\C_1=z''_{yy}\Big|_{(9,3)}=\frac{5}{3},\end{cases}$则 $\Delta_1=B_1^2-A_1C_1=-\frac{1}{36}.$

由$\begin{cases}\Delta_1<0,\\ A_1>0\end{cases}\Rightarrow P_1(9,3)$为极小值点，$z_1=3$为极小值；

对于$P_2(-9,-3)$，$z_2=-3$，同理可得$\begin{cases}\Delta_2=-\dfrac{1}{36}<0,\\ A_2=-\dfrac{1}{6}<0\end{cases}\Rightarrow P_2(-9,-3)$为极大值点，$z_2=-3$为极大值.

【注】 有些读者看到$z_1=3$为极小值，$z_2=-3$为极大值，便认为这是错误的，这反映了他们对极值概念的认识是模糊的.

2. 无条件极值——显函数

例 1.11.10 求函数$f(x,y)=x^4+y^4-(x+y)^2$的极值.

解 $\begin{cases}f'_x=4x^3-2(x+y)=0,\\ f'_y=4y^3-2(x+y)=0\end{cases}\Rightarrow\begin{cases}x=-1,\\ y=-1;\end{cases}\begin{cases}x=0,\\ y=0;\end{cases}\begin{cases}x=1,\\ y=1.\end{cases}$

即得驻点$P_1(-1,-1)$，$P_2(0,0)$，$P_3(1,1)$. 又

$$A=f''_{xx}=12x^2-2,\quad B=f''_{xy}=-2,\quad C=f''_{yy}=12y^2-2.$$

$(B^2-AC)\Big|_{P_1}=(-2)^2-10\cdot 10=-96<0$，且$A\Big|_{P_1}>0$，故$f(-1,-1)=-2$是极小值.

$(B^2-AC)\Big|_{P_2}=0$，故$P_2(0,0)$点的极值情况无法由二元函数极值存在的充分性定理判别. 由于$f(x,y)=x^4+y^4-(x+y)^2$，则在$(0,0)$点的任一邻域$x^2+y^2<\delta^2<1$内，取$(\varepsilon,-\varepsilon)$，得$f(\varepsilon,-\varepsilon)=2\varepsilon^4>0$，而$f(\varepsilon,\varepsilon)=2\varepsilon^4-4\varepsilon^2=2\varepsilon^2(\varepsilon-\sqrt{2})(\varepsilon+\sqrt{2})$，由$0<\varepsilon<1$知$f(\varepsilon,\varepsilon)<0$，故$f(0,0)=0$不是极值.

$(B^2-AC)\Big|_{P_3}=-96<0$，且$A\Big|_{P_3}>0$，故$f(1,1)=-2$也是极小值.

3. 闭区域边界上的最值

例 1.11.11 求函数$u=x^2+y^2+z^2$在约束条件$\begin{cases}z=x^2+y^2,\\ x+y+z=4\end{cases}$下的最大值与最小值.

解 记　　$F(x,y,z,\lambda,\mu)=x^2+y^2+z^2+\lambda(x^2+y^2-z)+\mu(x+y+z-4)$.

令

$$\begin{cases}F'_x=2x+2\lambda x+\mu=0,\\ F'_y=2y+2\lambda y+\mu=0,\\ F'_z=2z-\lambda+\mu=0,\\ F'_\lambda=x^2+y^2-z=0,\\ F'_\mu=x+y+z-4=0,\end{cases}$$

解该方程组得$P_1(1,1,2)$，$P_2(-2,-2,8)$.

根据实际问题，必存在最值，所得即所求，故$u_{\max}=72$，$u_{\min}=6$.

4. 闭区域上的最值

例 1.11.12 求函数$f(x,y)=x^2+2y^2-x^2y^2$在区域$D=\{(x,y)\mid x^2+y^2\leqslant 4,y\geqslant 0\}$上的最大值与最小值.

解 对于区域D的内部，这是无条件极值，求出可疑点并计算可疑点处的函数值.

令$\begin{cases} f'_x=2x-2xy^2=0, \\ f'_y=4y-2x^2y=0 \end{cases} \Rightarrow \begin{cases} P_1(\sqrt{2},1), \\ P_2(-\sqrt{2},1) \end{cases} \Rightarrow f(P_1)=f(P_2)=2.$

对于区域 D 的边界，这是条件极值，考虑用拉格朗日乘数法或者直接代入法，这个题目把边界方程代入 $f(x,y)$ 会比较简单.

设 L_1 为 $x^2+y^2=4(y>0)$，于是有 $f(x,y)=f(x,\sqrt{4-x^2})=x^4-5x^2+8\overset{记}{=\!=}g(x)$，则

$$g'(x)=4x^3-10x=0\Rightarrow x_1=0,x_{2,3}=\pm\sqrt{\frac{5}{2}},y_1=2,y_{2,3}=\sqrt{\frac{3}{2}},$$

即
$$P_3(0,2),P_4\left(\sqrt{\frac{5}{2}},\sqrt{\frac{3}{2}}\right),P_5\left(-\sqrt{\frac{5}{2}},\sqrt{\frac{3}{2}}\right)\Rightarrow f(P_3)=8,f(P_4)=f(P_5)=\frac{7}{4}.$$

设 L_2 为 $y=0(-2\leqslant x\leqslant 2)$，于是 $f(x,y)=f(x,0)=x^2$. 故得

$$P_6(0,0),P_7(-2,0),P_8(2,0)\Rightarrow f(P_6)=0,f(P_7)=f(P_8)=4.$$

比较 $f(P_1)$ 至 $f(P_8)$ 的大小，得 $f_{\max}=8,f_{\min}=0$.

【注】 求函数 $f(x,y)$ 在某区域 D 上的最值的步骤：

①求出 $f(x,y)$ 在 D 内所有可疑点处的函数值；

②求出 $f(x,y)$ 在 D 的边界上的最值；

③比较得到的所有函数值，其中最大者即为最大值，最小者即为最小值.

在实际问题中，如果可以判断出 $f(x,y)$ 的最大值（最小值）一定在 D 的内部取得，且 $f(x,y)$ 在 D 内只有一个驻点，则可以断定该驻点处的函数值就是 $f(x,y)$ 在 D 上的最大值（最小值）.

基础习题精练

习 题

1.11.1 设 $f(x,y)=\begin{cases} (x+y)\sin\dfrac{1}{\sqrt{x^2+y^2}}, & (x,y)\neq(0,0), \\ 0, & (x,y)=(0,0), \end{cases}$ 则在点(0,0)处（　　）.

(A)连续，且偏导数存在　　(B)连续，但偏导数不存在

(C)不连续，但偏导数存在　　(D)不连续，偏导数也不存在

1.11.2 函数 $f(x,y)=\sqrt{|xy|}$ 在点(0,0)处（　　）.

(A)偏导数不存在　　(B)偏导数存在，但不可微

(C)可微，但偏导数不连续　　(D)偏导数连续

1.11.3 设函数 $f(x)$ 可微，且 $f'(0)=\frac{1}{2}$，则 $z=f(4x^2-y^2)$ 在点(1,2)处的全微分 $\mathrm{d}z\Big|_{(1,2)}=$________.

1.11.4 设 $z=f(x^2y^2,\mathrm{e}^{xy})$，其中 $f(u,v)$ 有二阶连续偏导数，求 $z''_{xx},z''_{yy},z''_{xy}$.

1.11.5 设函数 $u=f(x^2,xy,xz)$ 具有一阶连续偏导数，又函数 $y=y(x),z=z(x)$ 分别由

$$\sin xy = y,\quad e^z = \int_0^{xz} \sin t\mathrm{d}t$$

确定，求$\frac{\mathrm{d}u}{\mathrm{d}x}$.

1.11.6 求 $f(x,y)=x^2-y^2+2$ 在椭圆域 $D=\left\{(x,y)\,\middle|\,x^2+\frac{y^2}{4}\leqslant 1\right\}$上的最大值和最小值.

1.11.7 求函数 $u=xyz$ 在约束条件$\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=\frac{1}{a}(x>0,y>0,z>0,a>0)$下的最小值.

1.11.8 已知函数 $z=f(x,y)$的全微分 $\mathrm{d}z=2x\mathrm{d}x+\sin y\mathrm{d}y$，且 $f(1,0)=2$，求 $f(x,y)$.

解答

1.11.1 (B) **解** $\lim\limits_{(x,y)\to(0,0)} f(x,y)=\lim\limits_{(x,y)\to(0,0)}(x+y)\sin\frac{1}{\sqrt{x^2+y^2}}=0=f(0,0)$，

故$f(x,y)$在$(0,0)$处连续. 又

$$\lim_{\Delta x\to 0}\frac{f(0+\Delta x,0)-f(0,0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\Delta x\sin\frac{1}{\sqrt{(\Delta x)^2}}}{\Delta x}=\lim_{\Delta x\to 0}\sin\frac{1}{|\Delta x|}$$

不存在，则 $f_x'(0,0)$不存在，同理 $f_y'(0,0)$不存在，故 $f(x,y)$在点$(0,0)$处的偏导数不存在.

1.11.2 (B) **解** $f_x'(0,0)=\lim\limits_{\Delta x\to 0}\frac{f(0+\Delta x,0)-f(0,0)}{\Delta x}=\lim\limits_{\Delta x\to 0}\frac{\sqrt{|\Delta x\cdot 0|}-0}{\Delta x}=0=A$，

$$f_y'(0,0)=\lim_{\Delta y\to 0}\frac{f(0,0+\Delta y)-f(0,0)}{\Delta y}=0=B,$$

故 $f(x,y)$在$(0,0)$处偏导数存在. 又

$$\Delta z=f(0+\Delta x,0+\Delta y)-f(0,0)=\sqrt{|\Delta x\cdot\Delta y|}-0=\sqrt{|\Delta x\cdot\Delta y|},$$

故有$\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\Delta z-A\Delta x-B\Delta y}{\sqrt{(\Delta x)^2+(\Delta y)^2}}=\lim\limits_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{\sqrt{|\Delta x\cdot\Delta y|}}{\sqrt{(\Delta x)^2+(\Delta y)^2}}$不存在. 从而 $f(x,y)$在点$(0,0)$处不可微.

1.11.3 $4\mathrm{d}x-2\mathrm{d}y$ **解** 令 $u=4x^2-y^2$，则 $z=f(u)$，于是

$$\frac{\partial z}{\partial x}=f'(u)(4x^2-y^2)_x'=8xf'(u),\quad \left.\frac{\partial z}{\partial x}\right|_{(1,2)}=8f'(0)=4,$$

$$\frac{\partial z}{\partial y}=f'(u)(4x^2-y^2)_y'=-2yf'(u),\quad \left.\frac{\partial z}{\partial y}\right|_{(1,2)}=-4f'(0)=-2,$$

所以 $\left.\mathrm{d}z\right|_{(1,2)}=4\mathrm{d}x-2\mathrm{d}y$.

1.11.4 **解** $z_x'=f_1'\cdot 2xy^2+f_2'\cdot ye^{xy}$，$z_y'=f_1'\cdot 2x^2y+f_2'\cdot xe^{xy}$，

$$z_{xx}''=f_{11}''\cdot(2xy^2)^2+f_{12}''\cdot ye^{xy}\cdot 2xy^2+f_1'\cdot 2y^2+f_{21}''\cdot 2xy^2\cdot ye^{xy}+f_{22}''\cdot(ye^{xy})^2+f_2'\cdot y^2e^{xy}$$
$$=f_{11}''\cdot 4x^2y^4+f_{22}''\cdot y^2e^{2xy}+f_{12}''\cdot 4xy^3e^{xy}+f_1'\cdot 2y^2+f_2'\cdot y^2e^{xy}.$$

同理可得

$$z_{yy}''=f_{11}''\cdot 4x^4y^2+f_{22}''\cdot x^2e^{2xy}+f_{12}''\cdot 4x^3ye^{xy}+f_1'\cdot 2x^2+f_2'\cdot x^2e^{xy},$$

$$z_{xy}''=f_{11}''\cdot 2x^2y\cdot 2xy^2+f_{12}''\cdot xe^{xy}\cdot 2xy^2+f_1'\cdot 4xy+f_{21}''\cdot 2x^2y\cdot ye^{xy}+f_{22}''\cdot xe^{xy}\cdot ye^{xy}+f_2'\cdot(e^{xy}+xye^{xy})$$
$$=f_{11}''\cdot 4x^3y^3+f_{22}''\cdot xye^{2xy}+f_{12}''\cdot 4x^2y^2e^{xy}+f_1'\cdot 4xy+f_2'\cdot(1+xy)e^{xy}.$$

1.11.5 解 复合函数 $u=f(x^2,xy,xz)$ 两边对 x 求导数得

$$\frac{\mathrm{d}u}{\mathrm{d}x}=f_1'\cdot 2x+f_2'\cdot\left(y+x\frac{\mathrm{d}y}{\mathrm{d}x}\right)+f_3'\cdot\left(z+x\frac{\mathrm{d}z}{\mathrm{d}x}\right). \quad ①$$

由隐函数 $\sin xy=y$ 求 $\frac{\mathrm{d}y}{\mathrm{d}x}$. 等式 $\sin xy=y$ 两边对 x 求导数得

$$\cos xy\cdot\left(y+x\frac{\mathrm{d}y}{\mathrm{d}x}\right)=\frac{\mathrm{d}y}{\mathrm{d}x},$$

解得

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y\cos xy}{1-x\cos xy}. \quad ②$$

由隐函数 $\mathrm{e}^z=\int_0^{xz}\sin t\mathrm{d}t$ 求 $\frac{\mathrm{d}z}{\mathrm{d}x}$. 等式 $\mathrm{e}^z=\int_0^{xz}\sin t\mathrm{d}t$ 两边对 x 求导数得

$$\mathrm{e}^z\frac{\mathrm{d}z}{\mathrm{d}x}=\sin xz\cdot\left(z+x\frac{\mathrm{d}z}{\mathrm{d}x}\right),$$

解得

$$\frac{\mathrm{d}z}{\mathrm{d}x}=\frac{z\sin xz}{\mathrm{e}^z-x\sin xz}. \quad ③$$

将②,③两式代入①式,得

$$\frac{\mathrm{d}u}{\mathrm{d}x}=f_1'\cdot 2x+f_2'\cdot\left(y+x\frac{y\cos xy}{1-x\cos xy}\right)+f_3'\cdot\left(z+x\frac{z\sin xz}{\mathrm{e}^z-x\sin xz}\right).$$

1.11.6 分析 因为只求函数 $f(x,y)=x^2-y^2+2$ 在椭圆域 D 上的最大值和最小值,而不求极值,所以,只需求出 D 的内部及 D 的边界上的驻点和导数不存在的点(不用判断它们是否为极值点),并求出这些点的函数值,然后比较它们的大小,求出最大值和最小值.

解 先求 D 内部 $f(x,y)$ 的驻点和导数不存在的点.

令 $\begin{cases}f_x'=2x=0,\\ f_y'=-2y=0,\end{cases}$ 得唯一驻点 $(0,0)$, $f(x,y)$ 在 D 内部没有导数不存在的点.

再求 $f(x,y)$ 在 D 的边界上的驻点和导数不存在的点.

这是条件极值问题,边界方程 $x^2+\frac{y^2}{4}=1$ 为约束条件. 构造拉格朗日函数

$$F(x,y,\lambda)=x^2-y^2+2+\lambda\left(x^2+\frac{y^2}{4}-1\right),$$

令

$$\begin{cases}F_x'=2x+2\lambda x=0,\\ F_y'=-2y+\frac{\lambda}{2}y=0,\\ F_\lambda'=x^2+\frac{y^2}{4}-1=0,\end{cases}$$

解上述方程组得四个驻点 $(0,2),(0,-2),(1,0),(-1,0)$, 计算得

$$f(0,0)=2,\quad f(0,2)=f(0,-2)=-2,\quad f(1,0)=f(-1,0)=3,$$

故函数 $f(x,y)=x^2-y^2+2$ 在椭圆域 $D=\left\{(x,y)\middle|x^2+\frac{y^2}{4}\leqslant1\right\}$ 上的最大值和最小值分别为 $f_{\max}=3$, $f_{\min}=-2$.

【注】 若将 $x^2=1-\frac{y^2}{4}$ 代入 $f(x,y)=x^2-y^2+2$, 得

$$f(y)=1-\frac{y^2}{4}-y^2+2=3-\frac{5}{4}y^2,\ -2\leqslant y\leqslant2,$$

令 $f'(y)=-\frac{5}{2}y=0$ 得 $y=0$, 于是得驻点 $(1,0)$ 及 $(-1,0)$, 并考虑两个端点 $(0,-2),(0,2)$ 作比较亦可.

1.11.7　解　构造辅助函数

$$F(x,y,z,\lambda)=xyz+\lambda\left(\frac{1}{x}+\frac{1}{y}+\frac{1}{z}-\frac{1}{a}\right),$$

令
$$\begin{cases}\dfrac{\partial F}{\partial x}=yz-\dfrac{\lambda}{x^2}=0,\\ \dfrac{\partial F}{\partial y}=xz-\dfrac{\lambda}{y^2}=0,\\ \dfrac{\partial F}{\partial z}=xy-\dfrac{\lambda}{z^2}=0,\\ \dfrac{\partial F}{\partial \lambda}=\dfrac{1}{x}+\dfrac{1}{y}+\dfrac{1}{z}-\dfrac{1}{a}=0,\end{cases}$$

解得
$$x=y=z=3a,$$

从而根据 $x>0,y>0,z>0$，可知 $u=xyz$ 有最小值为 $u_{\min}=3a\cdot 3a\cdot 3a=27a^3$.

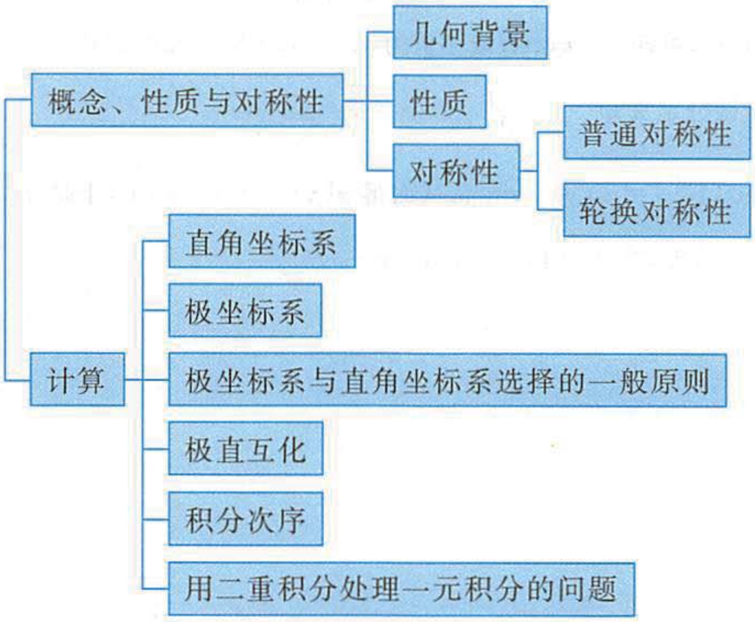
1.11.8　解　由全微分定义知　$\dfrac{\partial f}{\partial x}=2x,\quad \dfrac{\partial f}{\partial y}=\sin y,$

于是 $f(x,y)=x^2+\varphi(y)$，且$\dfrac{\partial f}{\partial y}=\varphi'(y)=\sin y$，从而得 $\varphi(y)=-\cos y+C$，于是 $f(x,y)=x^2-\cos y+C$.

再由$f(1,0)=1-1+C=2$，得 $C=2$，所以 $f(x,y)=x^2-\cos y+2$.

第12讲 二重积分

基础知识结构



基础内容精讲

一、概念、性质与对称性

1. 几何背景

二重积分的几何背景就是**曲顶柱体的体积**.

简要说来，第 8 讲用“分割、近似、求和、取极限”的方法与步骤求出了二维平面上“曲边梯形的面积”，现在我们可以用同样的办法求出“曲顶柱体的体积”，这就是二重积分 $\iint\limits_D f(x,y)\mathrm{d}\sigma$.

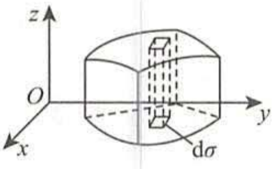


图 1-12-1

具体说来，设 $f(x,y)\geqslant 0$，如图 1-12-1 所示，被积函数 $f(x,y)$ 可作为曲顶柱体在点 (x,y) 处柱体微元的竖坐标(高)，用底面积 $\mathrm{d}\sigma$ 乘以高 $f(x,y)$，得到一个“小竖条”的体积，再在区域 D 上把所有的“小竖条”累加起来，就得到了整个曲顶柱体的体积. 若 $f(x,y)<0$，柱体就在 xOy 面的下方，二重积分的绝对值仍等于柱体的体积，但二重积分的值是负的. 不过，无论高 $f(x,y)$ 是正还是负，底面积 $\mathrm{d}\sigma$ 不能是负

的,即 $d\sigma>0$,这一点要注意,否则就不符合二重积分的定义了.

在考研数学中,一般总假设 $f(x,y)$ 在 D 上可积,即二重积分总是存在的.

2. 性质

性质 1(求区域面积) $\iint\limits_D 1\cdot d\sigma=\iint\limits_D d\sigma=A$,其中 A 为 D 的面积.

性质 2(可积函数必有界) 当 $f(x,y)$ 在有界闭区域 D 上可积时,则 $f(x,y)$ 在 D 上必有界.

性质 3(积分的线性性质) 设 k_1,k_2 为常数,则

$$\iint\limits_D[k_1f(x,y)\pm k_2g(x,y)]d\sigma=k_1\iint\limits_D f(x,y)d\sigma\pm k_2\iint\limits_D g(x,y)d\sigma.$$

性质 4(积分的可加性) 当 $f(x,y)$ 在有界闭区域 D 上可积时,且 $D_1\cup D_2=D,D_1\cap D_2=\varnothing$,则

$$\iint\limits_D f(x,y)d\sigma=\iint\limits_{D_1}f(x,y)d\sigma+\iint\limits_{D_2}f(x,y)d\sigma.$$

性质 5(积分的保号性) 当 $f(x,y),g(x,y)$ 在有界闭区域 D 上可积时,若在 D 上,

$$f(x,y)\leqslant g(x,y),$$

则有

$$\iint\limits_D f(x,y)d\sigma\leqslant\iint\limits_D g(x,y)d\sigma.$$

特殊地有

$$\left|\iint\limits_D f(x,y)d\sigma\right|\leqslant\iint\limits_D|f(x,y)|d\sigma.$$

性质 6(二重积分的估值定理) 设 M,m 分别是 $f(x,y)$ 在有界闭区域 D 上的最大值和最小值,A 为 D 的面积,则有

$$mA\leqslant\iint\limits_D f(x,y)d\sigma\leqslant MA.$$

性质 7(二重积分的中值定理) 设函数 $f(x,y)$ 在有界闭区域 D 上连续,A 为 D 的面积,则在 D 上至少存在一点 (ξ,η),使得

$$\iint\limits_D f(x,y)d\sigma=f(\xi,\eta)A.$$

3. 普通对称性与轮换对称性

前面已经叙述过,二重积分的几何背景是曲顶柱体的体积. 请理解并牢记下面这句话:"用底面积 $d\sigma$ 乘以高 $f(x,y)$,得到一个'小竖条'的体积,再在区域 D 上把所有的'小竖条'累加起来,就得到了整个曲顶柱体的体积."基于这个思路,就可以谈普通对称性和轮换对称性了.

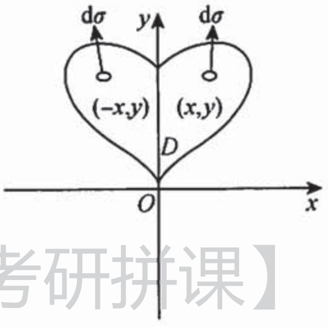


图 1-12-2

第一,先看**普通对称性**. 设区域 D 关于 y 轴对称,如图 1-12-2 所示,取对称的两块小面积 $d\sigma$,对称点分别为 (x,y) 与 $(-x,y)$,则对称点处的高分别为 $f(x,y)$ 与 $f(-x,y)$. 依据定义,对称位置的两个"小竖条"的体积分别为 $f(x,y)d\sigma$ 与 $f(-x,y)d\sigma$. 由于 $d\sigma$ 一样,所以,当 $f(x,y)=f(-x,y)$ 时,$f(x,y)d\sigma=f(-x,y)d\sigma$,体积相同,此时只需计算一半的对称区域,然后再乘以 2,即可得到整个积分值;而当 $f(x,y)=-f(-x,y)$ 时,$f(x,y)d\sigma=-f(-x,y)d\sigma$,对称位置的体积正好相反,这样累加起来的总体积自然就是 0. 于是,我们有

$$\iint\limits_D f(x,y)dxdy=\begin{cases}2\iint\limits_{D_1}f(x,y)dxdy, & f(x,y)=f(-x,y),\\ 0, & f(x,y)=-f(-x,y),\end{cases}$$

其中 D_1 是 D 在 y 轴右边的部分.

你看,这种基于概念的分析方法,不用死记硬背,而且真正理解了本质. 区域 D 关于 x 轴对称,关于原

点对称的问题，都与此类似.

第二，再看**轮换对称性**．这是考研的重点，也是很多考生始终弄不明白的地方．先看下面的一个问题：请问 $\iint\limits_{D_1:\frac{x^2}{4}+\frac{y^2}{3}\leqslant 1}(2x^2+3y^2)\mathrm{d}x\mathrm{d}y$ 是否等于 $\iint\limits_{D_2:\frac{y^2}{4}+\frac{x^2}{3}\leqslant 1}(2y^2+3x^2)\mathrm{d}y\mathrm{d}x$？答案显然是肯定的．因为上述两个积分只是将 x 与 y 这两个字母对调了，之所以相等，是因为**积分值与用什么字母表示是无关的**，这里并无任何对称性可言．抽象来说，便有

$$\iint\limits_{D_{xy}}f(x,y)\mathrm{d}x\mathrm{d}y\equiv\iint\limits_{D_{yx}}f(y,x)\mathrm{d}y\mathrm{d}x.$$

再看一个问题：请问 $\iint\limits_{D:\frac{x^2}{4}+\frac{y^2}{4}\leqslant 1}(2x^2+3y^2)\mathrm{d}x\mathrm{d}y$ 是否等于 $\iint\limits_{D:\frac{x^2}{4}+\frac{y^2}{4}\leqslant 1}(2y^2+3x^2)\mathrm{d}y\mathrm{d}x$？答案当然也是肯定的．理由如上，不再重复．不过，你是否注意到，这个问题中的区域 D 有个特点，就是当你把 x 与 y 对调后，区域 D 不变（事实上，区域 D 关于 $y=x$ 对称）．于是抽象化写出的式子为

$$\iint\limits_{D}f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{D}f(y,x)\mathrm{d}y\mathrm{d}x,$$

令 $\mathrm{d}\sigma=\mathrm{d}x\mathrm{d}y=\mathrm{d}y\mathrm{d}x$，则 $\iint\limits_{D}f(x,y)\mathrm{d}\sigma=\iint\limits_{D}f(y,x)\mathrm{d}\sigma$.

整理一下，我们可以这样来描述：

若把 x 与 y 对调后，区域 D 不变（或区域 D 关于 $y=x$ 对称），则

$$\iint\limits_{D}f(x,y)\mathrm{d}\sigma=\iint\limits_{D}f(y,x)\mathrm{d}\sigma,$$

这就是**轮换对称性**.

轮换对称性在解决某些"棘手"问题时往往很巧妙．看个例子.

注例 设区域 $D=\{(x,y)\mid x^2+y^2\leqslant 1,x\geqslant 0,y\geqslant 0\}$，$f(x)$ 为 D 上的正值连续函数，a,b 为常数，求 $I=\iint\limits_{D}\dfrac{a\sqrt{f(x)}+b\sqrt{f(y)}}{\sqrt{f(x)}+\sqrt{f(y)}}\mathrm{d}\sigma$.

解 被积函数是抽象的，无法直接计算，不过我们发现，若把 x 与 y 对调后，区域 D 不变，也就是说区域 D 关于 $y=x$ 对称，根据轮换对称性，有

$$I=\iint\limits_{D}\frac{a\sqrt{f(x)}+b\sqrt{f(y)}}{\sqrt{f(x)}+\sqrt{f(y)}}\mathrm{d}\sigma=\iint\limits_{D}\frac{a\sqrt{f(y)}+b\sqrt{f(x)}}{\sqrt{f(y)}+\sqrt{f(x)}}\mathrm{d}\sigma,$$

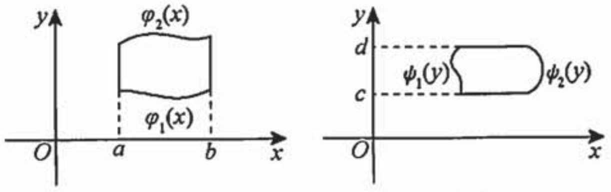
故 $$2I=\iint\limits_{D}\frac{a\sqrt{f(x)}+b\sqrt{f(y)}}{\sqrt{f(x)}+\sqrt{f(y)}}\mathrm{d}\sigma+\iint\limits_{D}\frac{a\sqrt{f(y)}+b\sqrt{f(x)}}{\sqrt{f(y)}+\sqrt{f(x)}}\mathrm{d}\sigma=\iint\limits_{D}(a+b)\mathrm{d}\sigma=(a+b)\frac{1}{4}\pi\cdot 1^2,$$

于是 $I=\dfrac{a+b}{8}\pi$.

二、计算

1. 直角坐标系下的计算法

在直角坐标系下，按照积分次序的不同，一般将二重积分的计算分为两种情况，如图 1-12-3 所示：



(a)　　(b)

图 1-12-3

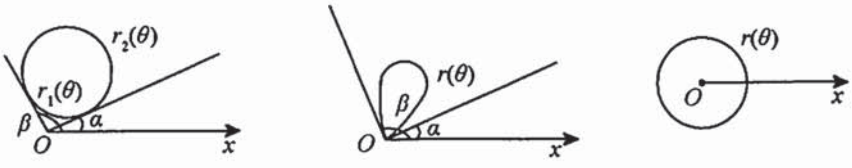
(1) $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_a^b \mathrm{d}x\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\mathrm{d}y$,其中 D 为 X 型区域:$\varphi_1(x)\leqslant y\leqslant\varphi_2(x)$, $a\leqslant x\leqslant b$;

(2) $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_c^d \mathrm{d}y\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\mathrm{d}x$,其中 D 为 Y 型区域:$\psi_1(y)\leqslant x\leqslant\psi_2(y)$, $c\leqslant y\leqslant d$.

有一点需要指出,这里的下限都必须小于等于上限.

2. 极坐标系下的计算法

在极坐标系下,按照积分区域与极点位置关系的不同,一般将二重积分的计算分为三种情况,如图 1-12-4 所示:



(a)　　(b)　　(c)

图 1-12-4

(1) $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_\alpha^\beta \mathrm{d}\theta\int_{r_1(\theta)}^{r_2(\theta)} f(r\cos\theta,r\sin\theta)r\mathrm{d}r$(极点 O 在区域 D 外部);

(2) $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_\alpha^\beta \mathrm{d}\theta\int_0^{r(\theta)} f(r\cos\theta,r\sin\theta)r\mathrm{d}r$(极点 O 在区域 D 边界上);

(3) $\iint\limits_D f(x,y)\mathrm{d}\sigma=\int_0^{2\pi} \mathrm{d}\theta\int_0^{r(\theta)} f(r\cos\theta,r\sin\theta)r\mathrm{d}r$(极点 O 在区域 D 内部).

需要指出的是,有很多同学在复习的过程中提问,为什么在极坐标系下不讨论积分次序的交换问题呢? 问得好.事实上,在极坐标系下,几乎所有的计算都是先积 r,后积 θ,所以一般不讨论.

3. 极坐标系与直角坐标系选择的一般原则

一般来说,给出一个二重积分.

①先看被积函数是否为 $f(x^2+y^2)$,$f\left(\frac{y}{x}\right)$,$f\left(\frac{x}{y}\right)$等形式;

②再看积分区域是否为圆或者圆的一部分.

如果两者兼是,那么优先选用极坐标系.否则,就优先考虑直角坐标系.(这只是一般原则,是大方向,请大家一定不要教条化)

4. 极坐标系与直角坐标系的互相转化

把握两个桥梁就可以,一是用好$\begin{cases}x=r\cos\theta,\\ y=r\sin\theta\end{cases}$这个公式,二是画好区域 D 的图形,确定好上、下限的转化.请看后面例题中具体问题的分析.

基础例题精解

1. 直角坐标系下的计算

例 1.12.1 设 $I_1=\iint\limits_D \cos\sqrt{x^2+y^2}\,\mathrm{d}\sigma, I_2=\iint\limits_D \cos(x^2+y^2)\,\mathrm{d}\sigma, I_3=\iint\limits_D \cos(x^2+y^2)^2\,\mathrm{d}\sigma$，其中 $D=\{(x,y)\mid x^2+y^2\leqslant 1\}$，则(　　).

(A) $I_3>I_2>I_1$　　(B) $I_1>I_2>I_3$　　(C) $I_2>I_1>I_3$　　(D) $I_3>I_1>I_2$

分析　关键在于比较 $\sqrt{x^2+y^2}, x^2+y^2$ 与 $(x^2+y^2)^2$ 在区域 $D=\{(x,y)\mid x^2+y^2\leqslant 1\}$ 上的大小.

解　应选(A).

在区域 $D=\{(x,y)\mid x^2+y^2\leqslant 1\}$ 上，有 $0\leqslant x^2+y^2\leqslant 1$，从而有

$$\frac{\pi}{2}>1\geqslant\sqrt{x^2+y^2}\geqslant x^2+y^2\geqslant(x^2+y^2)^2\geqslant 0.$$

由于 $\cos x$ 在 $\left(0,\frac{\pi}{2}\right)$ 上为单调减函数，于是

$$0<\cos\sqrt{x^2+y^2}\leqslant\cos(x^2+y^2)\leqslant\cos(x^2+y^2)^2,$$

且等号不恒成立，因此，$\iint\limits_D \cos\sqrt{x^2+y^2}\,\mathrm{d}\sigma<\iint\limits_D \cos(x^2+y^2)\,\mathrm{d}\sigma<\iint\limits_D \cos(x^2+y^2)^2\,\mathrm{d}\sigma$，故应选(A).

例 1.12.2 设 D 是由 $y=x^3, y=1, x=-1$ 所围成的平面闭区域，其中 D_1 为 D 在第一象限的部分，则 $\iint\limits_D (xy+\cos x\sin y)\,\mathrm{d}x\mathrm{d}y=$(　　).

(A) $2\iint\limits_{D_1} xy\,\mathrm{d}\sigma$　　(B) $2\iint\limits_{D_1} \cos x\sin y\,\mathrm{d}\sigma$

(C) $2\iint\limits_{D_1} (xy+\cos x\sin y)\,\mathrm{d}\sigma$　　(D) 0

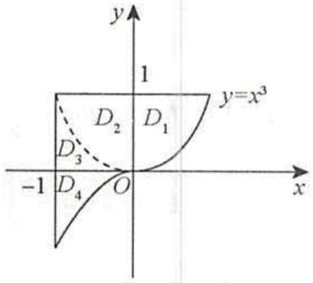


图 1-12-5

解　应选(B).

作辅助线(如图 1-12-5)，然后分别在 D_1, D_2, D_3, D_4 上讨论普通对称性，即可得结果，答案选择(B).

例 1.12.3 计算 $\iint\limits_D y\left(1+x\mathrm{e}^{\frac{x^2+y^2}{2}}\right)\mathrm{d}x\mathrm{d}y$，其中平面区域 D 由直线 $y=x, y=-1$ 及 $x=1$ 所围成.

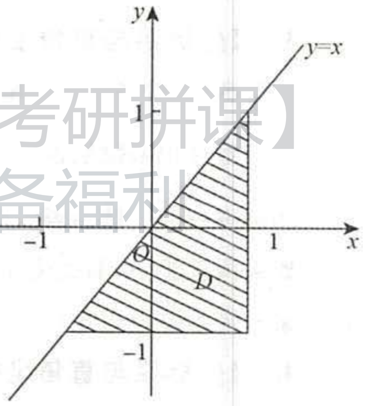


图 1-12-6

解　积分区域 D 如图 1-12-6 所示，

$$\iint\limits_D y\left(1+x\mathrm{e}^{\frac{x^2+y^2}{2}}\right)\mathrm{d}x\mathrm{d}y=\iint\limits_D y\,\mathrm{d}x\mathrm{d}y+\iint\limits_D xy\mathrm{e}^{\frac{x^2+y^2}{2}}\,\mathrm{d}x\mathrm{d}y,$$

其中 $\iint\limits_D xy\mathrm{e}^{\frac{x^2+y^2}{2}}\,\mathrm{d}x\mathrm{d}y$ 可直接由区域对称性和函数奇偶性得其值为 0，或者化为累次积分再根据对称性得 0，如下：

$$\iint\limits_D xy\mathrm{e}^{\frac{x^2+y^2}{2}}\,\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}y\,\mathrm{d}y\int_{y}^{1}x\mathrm{e}^{\frac{x^2+y^2}{2}}\,\mathrm{d}x=\int_{-1}^{1}y\left(\mathrm{e}^{\frac{1+y^2}{2}}-\mathrm{e}^{y^2}\right)\mathrm{d}y=0.$$

而 $\iint\limits_D y\,\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\mathrm{d}y\int_{y}^{1}y\,\mathrm{d}x=\int_{-1}^{1}y(1-y)\,\mathrm{d}y=-\frac{2}{3}$，故 $\iint\limits_D y\left(1+x\mathrm{e}^{\frac{x^2+y^2}{2}}\right)\mathrm{d}x\mathrm{d}y=-\frac{2}{3}$.

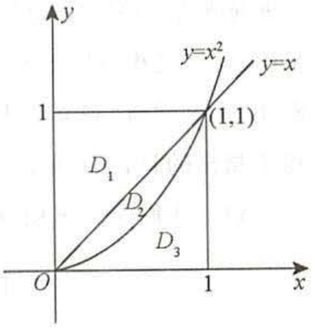


图 1-12-7

例 1.12.4　计算 $I=\iint\limits_D |y-x^2|\max\{x,y\}\mathrm{d}\sigma$,其中

$$D=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant 1\}.$$

解　用曲线 $y=x^2$ 和直线 $y=x$ 将区域 D 分为三块:D_1,D_2,D_3,如图 1-12-7 所示.

在 D_1 上,$\begin{cases} y\geqslant x^2\Rightarrow |y-x^2|=y-x^2, \\ y\geqslant x\Rightarrow \max\{x,y\}=y. \end{cases}$

在 D_2 和 D_3 上可同理分析,故

$$\iint\limits_D |y-x^2|\max\{x,y\}\mathrm{d}x\mathrm{d}y=\iint\limits_{D_1}(y-x^2)y\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}(y-x^2)x\mathrm{d}x\mathrm{d}y+\iint\limits_{D_3}(x^2-y)x\mathrm{d}x\mathrm{d}y$$

$$=\int_0^1\mathrm{d}x\int_x^1(y^2-x^2y)\mathrm{d}y+\int_0^1\mathrm{d}x\int_{x^2}^x(xy-x^3)\mathrm{d}y+\int_0^1\mathrm{d}x\int_0^{x^2}(x^3-xy)\mathrm{d}y$$

$$=\int_0^1\left(\frac{1}{3}-\frac{x^3}{3}-\frac{x^2}{2}+\frac{x^4}{2}\right)\mathrm{d}x+\int_0^1\left(\frac{x^3}{2}-\frac{x^5}{2}-x^4+x^5\right)\mathrm{d}x+\int_0^1\left(x^5-\frac{x^5}{2}\right)\mathrm{d}x=\frac{11}{40}.$$

2. 极坐标系下的计算

例 1.12.5　设区域 $D=\{(x,y)\mid x^2+y^2\leqslant R^2\}$,计算$\iint\limits_D\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)\mathrm{d}x\mathrm{d}y$.

解　$$\iint\limits_D\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\mathrm{d}\theta\int_0^R\left(\frac{r^2\cos^2\theta}{a^2}+\frac{r^2\sin^2\theta}{b^2}\right)r\mathrm{d}r=\int_0^{2\pi}\left(\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}\right)\mathrm{d}\theta\int_0^R r^3\mathrm{d}r$$

$$=\frac{r^4}{4}\bigg|_0^R\cdot 4\int_0^{\frac{\pi}{2}}\left(\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}\right)\mathrm{d}\theta=\frac{\pi R^4(a^2+b^2)}{4a^2b^2}.$$

【注】　本题还可以用轮换对称性来求解.

$$\iint\limits_D\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)\mathrm{d}x\mathrm{d}y=\iint\limits_D\left(\frac{y^2}{a^2}+\frac{x^2}{b^2}\right)\mathrm{d}x\mathrm{d}y=\frac{1}{2}\left(\frac{1}{a^2}+\frac{1}{b^2}\right)\iint\limits_D(x^2+y^2)\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{2}\left(\frac{1}{a^2}+\frac{1}{b^2}\right)\int_0^{2\pi}\mathrm{d}\theta\int_0^R(r^2\cos^2\theta+r^2\sin^2\theta)r\mathrm{d}r=\frac{\pi R^4(a^2+b^2)}{4a^2b^2}.$$

3. 直角坐标系与极坐标系互相转化后的计算

二重积分的考研题有个基本特点,就是命题人所给的坐标系往往是不容易甚至是不能做出积分的.这就出现了两个出题角度:一是给出极坐标系,你需要转化成直角坐标系去计算;二是反过来,给出直角坐标系,你需要转化成极坐标系去计算.如此"伎俩",命题人用心良苦,不过我们早已看透,见怪不怪,请看两个典型的例子.

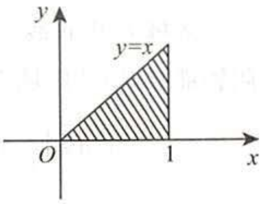


图 1-12-8

例 1.12.6　计算二重积分 $I=\iint\limits_D r^2\sin\theta\sqrt{1-r^2\cos 2\theta}\,\mathrm{d}r\mathrm{d}\theta$,其中

$$D=\left\{(r,\theta)\,\middle|\,0\leqslant r\leqslant\sec\theta,0\leqslant\theta\leqslant\frac{\pi}{4}\right\}.$$

分析　这就是将一个普通的课后练习题改编为 2010 年数学二的考研题,如果题目直接告诉我们,请计算 $I=\iint\limits_D y\sqrt{1-x^2+y^2}\mathrm{d}x\mathrm{d}y$,其中 D 是由$y=x,x=1,y=$

所围成的平面区域(如图 1-12-8),恐怕没有人一上来就用极坐标去计算,因为显然,被积函数中不含 x^2+y^2 的形式,且积分区域不是圆的一部分,所以优先考虑的应该是直角坐标系. 而当命题人“故意选择错误的路子”去做题,做到某一步做不下去的时候,他停下来,请你帮他继续做下去,就出现了如上的考题. 最后不得不指出的是,此题的得分率极低.

解 我们先画出积分区域 D,然后将积分化为直角坐标系下的二重积分,再去计算,即

$$\begin{aligned}I&=\iint_D r^2\sin\theta\sqrt{1-r^2\cos^2\theta+r^2\sin^2\theta}\,\mathrm{d}r\mathrm{d}\theta=\iint_D y\sqrt{1-x^2+y^2}\,\mathrm{d}x\mathrm{d}y\\&=\frac{1}{2}\int_0^1\mathrm{d}x\int_0^x\sqrt{1-x^2+y^2}\,\mathrm{d}(1-x^2+y^2)=\frac{1}{3}\int_0^1(1-x^2+y^2)^{\frac{3}{2}}\Big|_0^x\mathrm{d}x\\&=\frac{1}{3}\int_0^1\left[1-(1-x^2)^{\frac{3}{2}}\right]\mathrm{d}x.\end{aligned}$$

设 $x=\sin t$,则 $I=\frac{1}{3}-\frac{1}{3}\int_0^{\frac{\pi}{2}}\cos^4 t\mathrm{d}t=\frac{1}{3}-\frac{1}{3}\times\frac{3}{4}\times\frac{1}{2}\times\frac{\pi}{2}=\frac{1}{3}-\frac{\pi}{16}$.

例 1.12.7 计算 $I=\int_0^1\mathrm{d}x\int_{1-x}^{\sqrt{1-x^2}}\frac{x+y}{x^2+y^2}\mathrm{d}y$.

解 本题乍一看,也许我们会先考虑题目是否在积分次序上设置了障碍,是否需要交换积分次序再做积分,但是,细致做来,我们会发现不管是先对 x 积分,还是先对 y 积分,都不容易计算. 看来这不是积分次序上的问题,这时想想看是不是选择何种坐标系的问题呢? 被积函数中含有 x^2+y^2 的形式,且积分区域是圆的一部分,显然应该优先考虑极坐标系,题目给出的却是直角坐标系,如图 1-12-9 所示,我们需要改变一下. 于是,有

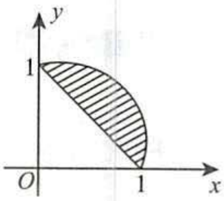


图 1-12-9

$$\begin{aligned}I&=\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_{\frac{1}{\cos\theta+\sin\theta}}^1\frac{r(\cos\theta+\sin\theta)}{r^2}r\mathrm{d}r\\&=\int_0^{\frac{\pi}{2}}(\cos\theta+\sin\theta)\left(\frac{\cos\theta+\sin\theta-1}{\cos\theta+\sin\theta}\right)\mathrm{d}\theta\\&=\int_0^{\frac{\pi}{2}}\cos\theta\mathrm{d}\theta+\int_0^{\frac{\pi}{2}}\sin\theta\mathrm{d}\theta-\int_0^{\frac{\pi}{2}}\mathrm{d}\theta\\&=1+1-\frac{\pi}{2}=2-\frac{\pi}{2}.\end{aligned}$$

4. 积分次序

例 1.12.8 计算 $\int_1^2\mathrm{d}x\int_{\sqrt{x}}^x\sin\frac{\pi x}{2y}\mathrm{d}y+\int_2^4\mathrm{d}x\int_{\sqrt{x}}^2\sin\frac{\pi x}{2y}\mathrm{d}y$.

解 由于 $\sin\frac{\pi x}{2y}$ 作为 y 的函数,其原函数不能用初等函数表示,因此交换积分次序.

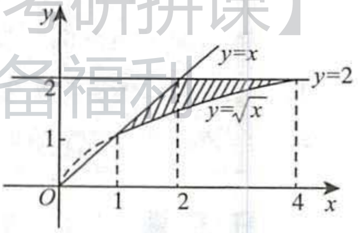


图 1-12-10

区域 D 由直线 $y=x$, $y=2$ 及抛物线 $y=\sqrt{x}$ 所围成,如图1-12-10所示的阴影部分,因此区域 D 可以写成 $D=\{(x,y)\mid 1\leqslant y\leqslant 2,y\leqslant x\leqslant y^2\}$,故

$$\begin{aligned}&\int_1^2\mathrm{d}x\int_{\sqrt{x}}^x\sin\frac{\pi x}{2y}\mathrm{d}y+\int_2^4\mathrm{d}x\int_{\sqrt{x}}^2\sin\frac{\pi x}{2y}\mathrm{d}y=\iint_D\sin\frac{\pi x}{2y}\mathrm{d}x\mathrm{d}y=\int_1^2\mathrm{d}y\int_y^{y^2}\sin\frac{\pi x}{2y}\mathrm{d}x\\&\qquad=\int_1^2\frac{2y}{\pi}\left(\cos\frac{\pi}{2}-\cos\frac{\pi}{2}y\right)\mathrm{d}y=\frac{4}{\pi^3}(2+\pi).\end{aligned}$$

5. 用二重积分处理一元积分的问题

例 1.12.9 设 $f(x)=\int_x^1 \sin(\pi u^2)\mathrm{d}u$，求 $\int_0^1 f(x)\mathrm{d}x$.

解 **方法一**

$$\int_0^1 f(x)\mathrm{d}x = xf(x)\Big|_0^1 - \int_0^1 x\mathrm{d}[f(x)]$$

$$= 0-\int_0^1 x\mathrm{d}\left[\int_x^1 \sin(\pi u^2)\mathrm{d}u\right]$$

$$= \int_0^1 x\sin(\pi x^2)\mathrm{d}x = \frac{1}{\pi}.$$

方法二 化二重积分，交换积分次序.

$$\int_0^1 f(x)\mathrm{d}x = \int_0^1 \mathrm{d}x\int_x^1 \sin(\pi u^2)\mathrm{d}u = \iint\limits_D \sin(\pi u^2)\mathrm{d}u\mathrm{d}x$$

$$= \int_0^1 \mathrm{d}u\int_0^u \sin(\pi u^2)\mathrm{d}x = \int_0^1 u\sin(\pi u^2)\mathrm{d}u = \frac{1}{\pi}.$$

例 1.12.10 设 $f(x)$ 为恒大于零的连续函数，求证 $\int_a^b f(x)\mathrm{d}x\cdot\int_a^b \frac{1}{f(x)}\mathrm{d}x \geqslant (b-a)^2$.

证明 设 $D=\{(x,y)\mid a\leqslant x\leqslant b, a\leqslant y\leqslant b\}$，

$$I=\int_a^b f(x)\mathrm{d}x\cdot\int_a^b \frac{1}{f(x)}\mathrm{d}x = \int_a^b f(x)\mathrm{d}x\cdot\int_a^b \frac{1}{f(y)}\mathrm{d}y = \iint\limits_D \frac{f(x)}{f(y)}\mathrm{d}x\mathrm{d}y,$$

$$I=\int_a^b f(x)\mathrm{d}x\cdot\int_a^b \frac{1}{f(x)}\mathrm{d}x = \int_a^b f(y)\mathrm{d}y\cdot\int_a^b \frac{1}{f(x)}\mathrm{d}x = \iint\limits_D \frac{f(y)}{f(x)}\mathrm{d}x\mathrm{d}y,$$

因此

$$I=\frac{1}{2}\left[\iint\limits_D \frac{f(x)}{f(y)}\mathrm{d}x\mathrm{d}y+\iint\limits_D \frac{f(y)}{f(x)}\mathrm{d}x\mathrm{d}y\right]$$

$$=\frac{1}{2}\iint\limits_D\left[\frac{f(x)}{f(y)}+\frac{f(y)}{f(x)}\right]\mathrm{d}x\mathrm{d}y \geqslant \frac{1}{2}\iint\limits_D 2\sqrt{\frac{f(x)}{f(y)}\cdot\frac{f(y)}{f(x)}}\mathrm{d}x\mathrm{d}y = \frac{1}{2}\iint\limits_D 2\mathrm{d}x\mathrm{d}y=(b-a)^2.$$

例 1.12.11 利用广义二重积分计算 $\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x$.

解 设 $I=\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x$，显然 $I>0$，于是

$$I^2=\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x\cdot\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x\cdot\int_0^{+\infty} \mathrm{e}^{-y^2}\mathrm{d}y$$

$$=\int_0^{+\infty}\mathrm{d}x\int_0^{+\infty} \mathrm{e}^{-(x^2+y^2)}\mathrm{d}y = \iint\limits_{\substack{0\leqslant x<+\infty\\0\leqslant y<+\infty}} \mathrm{e}^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y$$

$$\xlongequal{\text{用极坐标变换}} \int_0^{\frac{\pi}{2}}\mathrm{d}\theta\int_0^{+\infty} \mathrm{e}^{-r^2}\cdot r\mathrm{d}r = \frac{\pi}{2}\cdot\left(-\frac{1}{2}\right)\int_0^{+\infty} \mathrm{e}^{-r^2}\mathrm{d}(-r^2)$$

$$=-\frac{\pi}{4}\mathrm{e}^{-r^2}\Big|_0^{+\infty} = \frac{\pi}{4},$$

故 $I=\int_0^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \frac{\sqrt{\pi}}{2}$.

【注】 这一结果经常用到，最好记住.

基础习题精练

习 题

1.12.1 累次积分$\int_0^{\frac{\pi}{2}} \mathrm{d}\theta \int_0^{\cos\theta} f(r\cos\theta, r\sin\theta) r\mathrm{d}r$可以写成(　　).

(A) $\int_0^1 \mathrm{d}y \int_0^{\sqrt{y-y^2}} f(x,y)\mathrm{d}x$　　(B) $\int_0^1 \mathrm{d}y \int_0^{\sqrt{1-y^2}} f(x,y)\mathrm{d}x$

(C) $\int_0^1 \mathrm{d}x \int_0^1 f(x,y)\mathrm{d}y$　　(D) $\int_0^1 \mathrm{d}x \int_0^{\sqrt{x-x^2}} f(x,y)\mathrm{d}y$

1.12.2 证明：$\iint\limits_D (\mathrm{e}^{\sin y} + \mathrm{e}^{-\sin x})\mathrm{d}x\mathrm{d}y \geqslant 2\pi^2$，其中$D = \{(x,y) \mid 0 \leqslant x \leqslant \pi, 0 \leqslant y \leqslant \pi\}$.

1.12.3 交换积分次序：$\int_0^{\frac{1}{4}} \mathrm{d}y \int_y^{\sqrt{y}} f(x,y)\mathrm{d}x + \int_{\frac{1}{4}}^{\frac{1}{2}} \mathrm{d}y \int_y^{\frac{1}{2}} f(x,y)\mathrm{d}x =$ ________.

1.12.4 计算二重积分$\iint\limits_D \mathrm{e}^{\max\{x^2,y^2\}}\mathrm{d}x\mathrm{d}y$，其中$D = \{(x,y) \mid 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$.

1.12.5 计算$I = \iint\limits_D \dfrac{2\mathrm{e}^{x^2} + 3\mathrm{e}^{y^2}}{\mathrm{e}^{x^2} + \mathrm{e}^{y^2}}\mathrm{d}x\mathrm{d}y$，其中$D = \{(x,y) \mid 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$.

1.12.6 求二重积分$\iint\limits_D \dfrac{1-x^2-y^2}{1+x^2+y^2}\mathrm{d}x\mathrm{d}y$，其中$D$是$\{(x,y) \mid x^2+y^2 \leqslant 1\}$在第一象限的部分.

1.12.7 设$F(t) = \iint\limits_{x+y\leqslant t} f(x,y)\mathrm{d}x\mathrm{d}y$，其中$f(x,y) = \begin{cases} 1, & 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, \\ 0, & \text{其他}, \end{cases}$求$F(t)$的表达式.

解 答

1.12.1 (D) **解** 积分区域$D=\left\{(r,\theta) \,\middle|\, 0 \leqslant \theta \leqslant \dfrac{\pi}{2}, 0 \leqslant r \leqslant \cos\theta\right\}$，化为直角坐标表示为

$$D=\{(x,y) \mid 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant \sqrt{x-x^2}\}=\left\{(x,y) \,\middle|\, 0 \leqslant y \leqslant \frac{1}{2}, \frac{1}{2}-\sqrt{\frac{1}{4}-y^2} \leqslant x \leqslant \frac{1}{2}+\sqrt{\frac{1}{4}-y^2}\right\}.$$

1.12.2 **证明** 积分区域D如图 1-12-11 所示，且关于$y = x$对称.

$$I = \iint\limits_D (\mathrm{e}^{\sin y} + \mathrm{e}^{-\sin x})\mathrm{d}x\mathrm{d}y = \iint\limits_D \mathrm{e}^{\sin y}\mathrm{d}x\mathrm{d}y + \iint\limits_D \mathrm{e}^{-\sin x}\mathrm{d}x\mathrm{d}y,$$

由于$\iint\limits_D \mathrm{e}^{\sin y}\mathrm{d}x\mathrm{d}y \xlongequal{\text{轮换对称性}} \iint\limits_D \mathrm{e}^{\sin x}\mathrm{d}x\mathrm{d}y$，故

$$I = \iint\limits_D \mathrm{e}^{\sin x}\mathrm{d}x\mathrm{d}y + \iint\limits_D \mathrm{e}^{-\sin x}\mathrm{d}x\mathrm{d}y = \iint\limits_D (\mathrm{e}^{\sin x} + \mathrm{e}^{-\sin x})\mathrm{d}x\mathrm{d}y \geqslant \iint\limits_D 2\mathrm{d}x\mathrm{d}y = 2\pi^2,$$

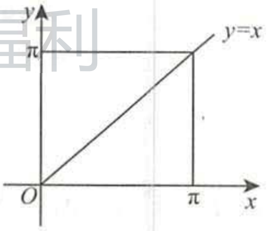


图 1-12-11

其中$\mathrm{e}^{\sin x} + \mathrm{e}^{-\sin x} \geqslant 2\sqrt{\mathrm{e}^{\sin x} \cdot \mathrm{e}^{-\sin x}} = 2$.

1.12.3 $\int_0^{\frac{1}{2}} \mathrm{d}x \int_{x^2}^{x} f(x,y)\mathrm{d}y$　**解**　积分区域为$D = \left\{(x,y) \,\middle|\, 0 \leqslant x \leqslant \dfrac{1}{2}, x^2 \leqslant y \leqslant x\right\}$，故应填

$\int_0^1 dx \int_{x^2}^{x} f(x,y)dy$.

1.12.4　解　设 $D_1=\{(x,y)\mid 0\leqslant x\leqslant 1,0\leqslant y\leqslant x\}$，$D_2=\{(x,y)\mid 0\leqslant x\leqslant 1,x\leqslant y\leqslant 1\}$，则

$$\iint\limits_D e^{\max\{x^2,y^2\}}dxdy=\iint\limits_{D_1} e^{\max\{x^2,y^2\}}dxdy+\iint\limits_{D_2} e^{\max\{x^2,y^2\}}dxdy=\iint\limits_{D_1} e^{x^2}dxdy+\iint\limits_{D_2} e^{y^2}dxdy$$

$$=\int_0^1 dx\int_0^x e^{x^2}dy+\int_0^1 dy\int_0^y e^{y^2}dx=\int_0^1 xe^{x^2}dx+\int_0^1 ye^{y^2}dy=e-1.$$

1.12.5　解　积分区域 D 如图 1-12-12 所示，D 关于直线 $y=x$ 对称.

由轮换对称性，得

$$I=\frac{1}{2}\iint\limits_D\left(\frac{2e^{x^2}+3e^{y^2}}{e^{x^2}+e^{y^2}}+\frac{2e^{y^2}+3e^{x^2}}{e^{y^2}+e^{x^2}}\right)dxdy$$

$$=\frac{5}{2}\iint\limits_D dxdy=\frac{5}{2}\times 1=\frac{5}{2}.$$

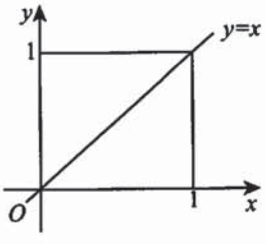


图 1-12-12

1.12.6　解　利用极坐标，得

$$原式=\int_0^{\frac{\pi}{2}}d\theta\int_0^1\frac{1-r^2}{1+r^2}rdr=\frac{\pi}{2}\int_0^1\left(\frac{2}{1+r^2}-1\right)rdr$$

$$=\frac{\pi}{2}\left[\ln(1+r^2)-\frac{1}{2}r^2\right]\Bigg|_0^1=\frac{\pi}{2}\left(\ln 2-\frac{1}{2}\right).$$

1.12.7　解　由被积函数 $f(x,y)$ 的表达式，可知 $F(t)$ 在数值上等于区域 $x+y\leqslant t$ 与正方形 $0\leqslant x\leqslant 1$，$0\leqslant y\leqslant 1$ 的公共部分的面积.

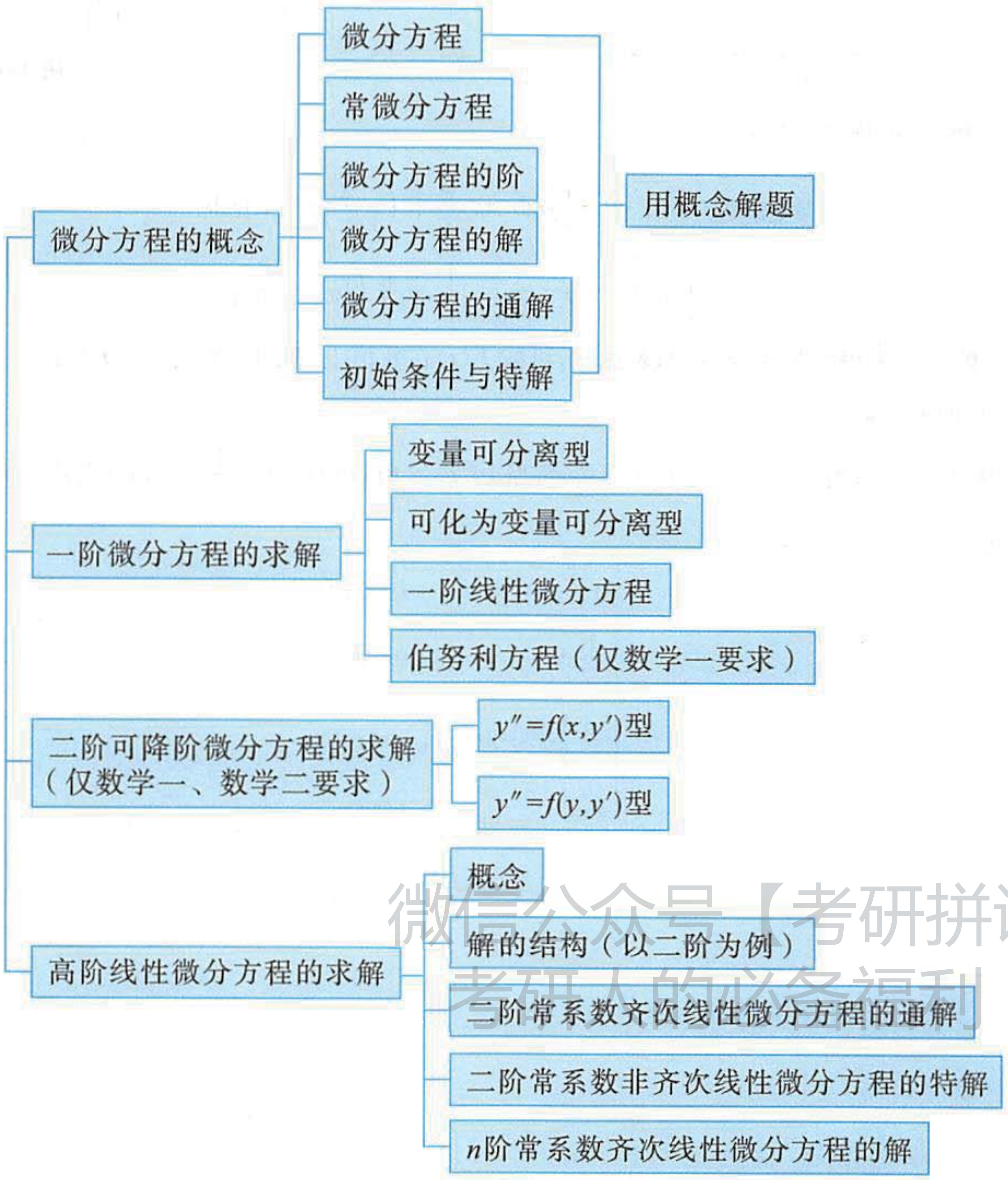
当 $t\leqslant 0$ 时，$F(t)=0$；当 $0<t\leqslant 1$ 时，$F(t)=\frac{1}{2}t^2$；当 $1<t\leqslant 2$ 时，$F(t)=1-\frac{1}{2}(2-t)^2$；当 $t>2$ 时，$F(t)=1$.

综上所述，

$$F(t)=\begin{cases}0, & t\leqslant 0,\\ \frac{1}{2}t^2, & 0<t\leqslant 1,\\ -\frac{1}{2}t^2+2t-1, & 1<t\leqslant 2,\\ 1, & t>2.\end{cases}$$

第13讲 常微分方程

基础知识结构



基础内容精讲

一、微分方程的概念

1. 微分方程

含有未知函数及其导数(或者微分)的方程称为**微分方程**.一般写成

$$F(x,y,y',\cdots,y^{(n)})=0 \text{ 或 } y^{(n)}=f(x,y,y',\cdots,y^{(n-1)}).$$

2. 常微分方程

未知函数是一元函数的微分方程称为**常微分方程**.如：$y'''-y''+6y=0$，$y\mathrm{d}x-(x+\sqrt{x^2+y^2})\mathrm{d}y=0$.

3. 微分方程的阶

方程中未知函数导数的最高阶数称为**微分方程的阶**.如：$y'''-y''+6y=0$就是三阶微分方程.

4. 微分方程的解

若将函数代入微分方程，使方程成为恒等式，则该函数称为**微分方程的解**.设$y=y(x)$在区间I上连续且有直到n阶的导数，使$F(x,y(x),y'(x),\cdots,y^{(n)}(x))\equiv 0$，则称$y=y(x)$为该微分方程在区间$I$上的一个解.

5. 微分方程的通解

若微分方程的解中含有的独立常数的个数等于微分方程的阶数，则该解称为微分方程的**通解**.

也就是说，若$y=y(x,C_1,C_2,\cdots,C_n)$是$n$阶微分方程$F(x,y(x),y'(x),\cdots,y^{(n)}(x))\equiv 0$在区间$I$上的解，其中$C_1,C_2,\cdots,C_n$为$n$个独立的常数，则称它为该微分方程的通解.

6. 初始条件与特解

确定通解中常数的条件就是**初始条件**.如：$y(x_0)=a_0$，$y'(x_0)=a_1$，$\cdots$，$y^{(n-1)}(x_0)=a_{n-1}$，其中a_0，a_1，$\cdots$，a_{n-1}为n个给定的数.确定了通解中的常数后，解就成了**特解**.

二、一阶微分方程的求解

1. 变量可分离型

能写成$y'=f(x)g(y)$形式的方程称为**变量可分离型方程**.其解法为

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x)g(y)\Rightarrow\int\frac{\mathrm{d}y}{g(y)}=\int f(x)\mathrm{d}x.$$

例如：

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\mathrm{e}^{x-y}=\mathrm{e}^x\cdot\mathrm{e}^{-y}\Rightarrow\int\mathrm{e}^y\,\mathrm{d}y=\int\mathrm{e}^x\mathrm{d}x\Rightarrow\begin{cases}\mathrm{e}^y=\mathrm{e}^x+C(\text{隐式解}),\\ y=\ln(\mathrm{e}^x+C)(\text{显式解}).\end{cases}$$

【注】 有时将$g(y)$放到分母中，即要求$g(y)\neq 0$，但$g(y)$本身并没有非零的要求，可能会丢解，丢的什么解？该如何处理？见例1.13.2的"解"与"注".

2. 可化为变量可分离型

(1)形如$\frac{\mathrm{d}y}{\mathrm{d}x}=f(ax+by+c)$的方程，其中常数$a,b$全都不为零.其解法为令$u=ax+by+c$，则$\frac{\mathrm{d}u}{\mathrm{d}x}=a+$

$b\frac{\mathrm{d}y}{\mathrm{d}x}$，代入原方程得$\frac{\mathrm{d}u}{\mathrm{d}x}=a+bf(u)$. 见例 1. 13. 2.

(2)齐次型微分方程.

如，对于微分方程

$$\left(x+y\cos\frac{y}{x}\right)\mathrm{d}x-x\cos\frac{y}{x}\mathrm{d}y=0,$$

将其变形为

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{x+y\cos\frac{y}{x}}{x\cos\frac{y}{x}}=\frac{1+\frac{y}{x}\cos\frac{y}{x}}{\cos\frac{y}{x}}=\varphi\left(\frac{y}{x}\right).$$

我们把形如$\frac{\mathrm{d}y}{\mathrm{d}x}=\varphi\left(\frac{y}{x}\right)$的方程叫作齐次型微分方程. 其解法为令 $u=\frac{y}{x}$，则

$$y=ux\Rightarrow\frac{\mathrm{d}y}{\mathrm{d}x}=u+x\frac{\mathrm{d}u}{\mathrm{d}x},$$

于是原方程变为 $x\frac{\mathrm{d}u}{\mathrm{d}x}+u=\varphi(u)$，即$\frac{\mathrm{d}u}{\varphi(u)-u}=\frac{\mathrm{d}x}{x}$. 见例 1. 13. 3.

3. 一阶线性微分方程

形如 $y'+p(x)y=q(x)$的方程，其中 $p(x)$，$q(x)$为已知的连续函数. 其通解公式为

$$y=\mathrm{e}^{-\int p(x)\mathrm{d}x}\left[\int\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot q(x)\mathrm{d}x+C\right].$$

具体见例 1. 13. 4.

请大家一定要掌握该公式的推导过程，这是一个很好的锻炼机会，不要错过.

推导计算公式 在方程两边同时乘以 $\mathrm{e}^{\int p(x)\mathrm{d}x}$，得

$$\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot y'+\mathrm{e}^{\int p(x)\mathrm{d}x}p(x)\cdot y=\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot q(x),$$

于是

$$\left[\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot y\right]'=\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot q(x),$$

两边积分，得

$$\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot y=\int\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot q(x)\mathrm{d}x+C,$$

则

$$y=\mathrm{e}^{-\int p(x)\mathrm{d}x}\left[\int\mathrm{e}^{\int p(x)\mathrm{d}x}\cdot q(x)\mathrm{d}x+C\right].$$

【注】 在一阶线性微分方程的通解公式 $y=\mathrm{e}^{-\int p(x)\mathrm{d}x}\left[\int\mathrm{e}^{\int p(x)\mathrm{d}x}q(x)\mathrm{d}x+C\right]$中，若

$$\int p(x)\mathrm{d}x=\ln|\varphi(x)|,$$

则

$$\mathrm{e}^{\int p(x)\mathrm{d}x}=|\varphi(x)|=\pm\varphi(x),\quad\mathrm{e}^{-\int p(x)\mathrm{d}x}=\pm\frac{1}{\varphi(x)},$$

代入上述公式中，有

$$\begin{aligned}y&=\pm\frac{1}{\varphi(x)}\left[\int\pm\varphi(x)\cdot q(x)\mathrm{d}x+C\right]\\&=\frac{1}{\varphi(x)}\left[\int\varphi(x)\cdot q(x)\mathrm{d}x\pm C\right]\\&\xlongequal{\text{令}\pm C=D}\frac{1}{\varphi(x)}\left[\int\varphi(x)\cdot q(x)\mathrm{d}x+D\right],\end{aligned}$$

其中 D 依然为任意常数，故 $\mathrm{e}^{\int p(x)\mathrm{d}x}=|\varphi(x)|$ 可不加绝对值.

在其他计算过程中，若出现 $\ln u$，且 u 不知正负，一律加绝对值.

4. 伯努利方程(仅数学一要求)

形如 $y'+p(x)y=q(x)y^n(n\neq 0,1)$ 的方程,其中 $p(x),q(x)$ 为已知的连续函数.其解法具体步骤为

(1)先变形为 $y^{-n}\cdot y'+p(x)y^{1-n}=q(x)$;

(2)令 $z=y^{1-n}$,得 $\frac{\mathrm{d}z}{\mathrm{d}x}=(1-n)y^{-n}\frac{\mathrm{d}y}{\mathrm{d}x}$,则 $\frac{1}{1-n}\cdot\frac{\mathrm{d}z}{\mathrm{d}x}+p(x)z=q(x)$;

(3)解此一阶线性微分方程即可.

见例 1.13.5.

三、二阶可降阶微分方程的求解(仅数学一、数学二要求)

1. $y''=f(x,y')$ 型(方程中不显含未知函数 y)

(1)令 $y'=p(x)$,$y''=p'$,则原方程变为一阶方程 $\frac{\mathrm{d}p}{\mathrm{d}x}=f(x,p)$;

(2)若求得其解为 $p=\varphi(x,C_1)$,即 $y'=\varphi(x,C_1)$,则原方程的通解为 $y=\int\varphi(x,C_1)\mathrm{d}x+C_2$.

见例 1.13.6.

2. $y''=f(y,y')$ 型(方程中不显含自变量 x)

(1)令 $y'=p$,$y''=\frac{\mathrm{d}p}{\mathrm{d}x}=\frac{\mathrm{d}p}{\mathrm{d}y}\cdot\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}p}{\mathrm{d}y}\cdot p$,则原方程变为一阶方程 $p\frac{\mathrm{d}p}{\mathrm{d}y}=f(y,p)$;

(2)若求得其解 $p=\varphi(y,C_1)$,则由 $p=\frac{\mathrm{d}y}{\mathrm{d}x}$ 可得 $\frac{\mathrm{d}y}{\mathrm{d}x}=\varphi(y,C_1)$,分离变量得 $\frac{\mathrm{d}y}{\varphi(y,C_1)}=\mathrm{d}x$;

(3)两边积分得 $\int\frac{\mathrm{d}y}{\varphi(y,C_1)}=x+C_2$,即可求得原方程的通解.

见例 1.13.7.

四、高阶线性微分方程的求解

1. 概念

(1)方程 $y''+p(x)y'+q(x)y=f(x)$ 称为**二阶变系数线性微分方程**,其中 $p(x),q(x)$ 叫系数函数,$f(x)$ 叫自由项,均为已知的连续函数.

当 $f(x)\equiv 0$ 时,$y''+p(x)y'+q(x)y=0$ 为**齐次方程**;

当 $f(x)$ 不恒等于 0 时,$y''+p(x)y'+q(x)y=f(x)$ 为**非齐次方程**.

(2)方程 $y''+py'+qy=f(x)$ 称为**二阶常系数线性微分方程**,其中 p,q 为常数,$f(x)$ 叫自由项.

当 $f(x)\equiv 0$ 时,$y''+py'+qy=0$ 为**齐次方程**;

当 $f(x)$ 不恒等于 0 时,$y''+py'+qy=f(x)$ 为**非齐次方程**.

2. 解的结构(以二阶为例)

(1)若 $y_1(x),y_2(x)$ 是 $y''+p(x)y'+q(x)y=0$ 的两个解,且 $\frac{y_1(x)}{y_2(x)}\neq C$(常数),则称 $y_1(x),y_2(x)$ 是该方程**的两个线性无关的解**,且 $y(x)=C_1y_1(x)+C_2y_2(x)$ 是方程 $y''+p(x)y'+q(x)y=0$ 的通解.

(2)若 $y(x)=C_1y_1(x)+C_2y_2(x)$ 是 $y''+p(x)y'+q(x)y=0$ 的通解,$y^*(x)$ 是

$$y''+p(x)y'+q(x)y=f(x)$$

的一个特解，则 $y(x)+y^*(x)$ 是 $y''+p(x)y'+q(x)y=f(x)$ 的通解.

(3)若 $y_1^*(x)$ 是 $y''+p(x)y'+q(x)y=f_1(x)$ 的解，$y_2^*(x)$ 是 $y''+p(x)y'+q(x)y=f_2(x)$ 的解，则 $y_1^*(x)+y_2^*(x)$ 是 $y''+p(x)y'+q(x)y=f_1(x)+f_2(x)$ 的解.

3. 二阶常系数齐次线性微分方程的通解

对于 $y''+py'+qy=0$，其对应的特征方程为 $\lambda^2+p\lambda+q=0$，求其特征根，有以下三种情况请大家牢记（其中 C_1,C_2 为任意常数）.

(1)若 $p^2-4q>0$，设 λ_1,λ_2 是特征方程的两个不等实根，即 $\lambda_1\neq\lambda_2$，可得其通解为

$$y=C_1e^{\lambda_1 x}+C_2e^{\lambda_2 x}.$$

(2)若 $p^2-4q=0$，设 λ_1,λ_2 是特征方程的两个相等的实根，即二重根，令 $\lambda_1=\lambda_2=\lambda$，可得其通解为

$$y=(C_1+C_2x)e^{\lambda x}.$$

(3)若 $p^2-4q<0$，设 $\alpha\pm\beta i$ 是特征方程的一对共轭复根，可得其通解为

$$y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x).$$

4. 二阶常系数非齐次线性微分方程的特解

对于 $y''+py'+qy=f(x)$，《考试大纲》规定我们需要会解以下两种情况.

设 $P_n(x)$，$P_m(x)$ 分别为 x 的 n 次，m 次多项式.

(1)当自由项 $f(x)=P_n(x)e^{\alpha x}$ 时，特解要设为 $y^*=e^{\alpha x}Q_n(x)x^k$，

$$\text{其中}\quad\begin{cases}e^{\alpha x}\text{照抄},\\ Q_n(x)\text{为 }x\text{ 的 }n\text{ 次多项式},\\ k=\begin{cases}0, & \alpha\text{ 不是特征根},\\ 1, & \alpha\text{ 是单特征根},\\ 2, & \alpha\text{ 是二重特征根}.\end{cases}\end{cases}$$

见例 1.13.8.

(2)当自由项 $f(x)=e^{\alpha x}[P_m(x)\cos\beta x+P_n(x)\sin\beta x]$ 时，特解要设为

$$y^*=e^{\alpha x}[Q_l^{(1)}(x)\cos\beta x+Q_l^{(2)}(x)\sin\beta x]x^k,$$

$$\text{其中}\quad\begin{cases}e^{\alpha x}\text{照抄},\\ l=\max\{m,n\},Q_l^{(1)}(x),Q_l^{(2)}(x)\text{分别为 }x\text{ 的两个不同的 }l\text{ 次多项式},\\ k=\begin{cases}0, & \alpha\pm\beta i\text{ 不是特征根},\\ 1, & \alpha\pm\beta i\text{ 是特征根}.\end{cases}\end{cases}$$

见例 1.13.9.

5. n 阶常系数齐次线性微分方程的解

方程 $y^{(n)}+p_1y^{(n-1)}+\cdots+p_{n-1}y'+p_ny=0$ 称为 n 阶常系数齐次线性微分方程，其中 $p_1,p_2,\cdots,p_n$ 为常数，其对应的特征方程为 $\lambda^n+p_1\lambda^{n-1}+\cdots+p_{n-1}\lambda+p_n=0$，求出其特征根. 有如下情况需要大家牢记（其中大写的英文字母均为任意常数）.

(1)特征根为单实根 λ 时，微分方程通解中对应一项 $Ce^{\lambda x}$；

(2)特征根为 k 重实根 λ 时，微分方程通解中对应 k 项 $(C_1+C_2x+\cdots+C_kx^{k-1})e^{\lambda x}$；

(3)特征根为单复根 $\alpha\pm\beta i(\beta>0)$ 时，微分方程通解中对应两项 $e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x)$；

(4)特征根为 k 重复根 $\alpha\pm\beta i(\beta>0)$ 时，微分方程通解中对应 $2k$ 项

$$e^{\alpha x}[(C_1+C_2x+\cdots+C_kx^{k-1})\cos\beta x+(D_1+D_2x+\cdots+D_kx^{k-1})\sin\beta x].$$

基础例题精解

一、一阶微分方程的求解

1. 变量可分离型

例 1.13.1 求 $y\sin\frac{x}{2}\mathrm{d}x-\cos\frac{x}{2}\mathrm{d}y=0$ 的通解.

解 分离变量，得$\frac{\mathrm{d}y}{y}=\frac{\sin\frac{x}{2}}{\cos\frac{x}{2}}\mathrm{d}x$，两边积分，有 $\ln|y|=-\ln\cos^2\frac{x}{2}+\ln C_1$，两边取指数，可得方程通解 $y=\pm\frac{C_1}{\cos^2\frac{x}{2}}=\frac{C}{1+\cos x}$，其中 C 为任意常数.

2. 可化为变量可分离型

例 1.13.2 求微分方程 $\mathrm{d}y=\sin(x+y+100)\mathrm{d}x$ 的通解.

解 ①方程可写成$\frac{\mathrm{d}y}{\mathrm{d}x}=\sin(x+y+100)$，令 $u=x+y+100$，则$\frac{\mathrm{d}u}{\mathrm{d}x}=1+\frac{\mathrm{d}y}{\mathrm{d}x}$，于是原方程化为$\frac{\mathrm{d}u}{\mathrm{d}x}=1+\sin u$，就得到了变量可分离型方程；

②分离变量，得$\frac{\mathrm{d}u}{1+\sin u}=\mathrm{d}x$，恒等变形，有$\frac{(1-\sin u)\mathrm{d}u}{1-\sin^2 u}=\mathrm{d}x$，即$(\sec^2 u-\tan u\sec u)\mathrm{d}u=\mathrm{d}x$.

两边积分，得 $\tan u-\sec u=x+C$. 将 $u=x+y+100$ 代入，得原方程的通解为

$$\tan(x+y+100)-\sec(x+y+100)=x+C,$$

其中 C 为任意常数.

【注】 事实上，在本题解析过程中的②处，分离变量得$\frac{\mathrm{d}u}{1+\sin u}=\mathrm{d}x$ 时，默认了一件事情，那就是 $\sin u\neq-1$，回避了 $1+\sin u=0$ 的情况，从而丢掉了全部解中的部分解(可称为“奇解”). 当 $\sin u=-1$ 时，得

$$x+y+100=2k\pi-\frac{\pi}{2},$$

其中 $k=0,\pm1,\pm2,\cdots$. 在《考试大纲》中，只要求求通解，并不要求求出全部解.

例 1.13.3 设 L 是一条平面曲线，其上任意一点 $P(x,y)(x>0)$ 到坐标原点的距离恒等于该点处的切线在 y 轴上的截距，且 L 经过点$\left(\frac{1}{2},0\right)$. 试求曲线 L 的方程.

解 设曲线 L 过点 $P(x,y)$ 的切线方程为 $Y-y=y'(X-x)$. 令 $X=0$，得该切线在 y 轴上的截距为 $y-xy'$.

由题设知 $\sqrt{x^2+y^2}=y-xy'$，令 $u=\frac{y}{x}$，则此方程可化为$\frac{\mathrm{d}u}{\sqrt{1+u^2}}=-\frac{\mathrm{d}x}{x}$，解之得

$$y+\sqrt{x^2+y^2}=C.$$

由 L 经过点 $\left(\frac{1}{2},0\right)$，知 $C=\frac{1}{2}$. 于是 L 的方程为 $y+\sqrt{x^2+y^2}=\frac{1}{2}$，即 $y=\frac{1}{4}-x^2(x>0)$.

3. 一阶线性微分方程

例 1.13.4 求微分方程 $xy'+ay=1+x^2$ 满足 $y\Big|_{x=1}=1$ 的解 $y(x,a)$，其中 a 为常数且 $a\neq 0$，$a\neq -2$.

解 微分方程的通解为

$$y=\mathrm{e}^{-\int\frac{a}{x}\mathrm{d}x}\left(\int\frac{1+x^2}{x}\mathrm{e}^{\int\frac{a}{x}\mathrm{d}x}\mathrm{d}x+C\right)$$

$$=x^{-a}\left(\int\frac{1+x^2}{x}x^a\mathrm{d}x+C\right)=x^{-a}\left(\frac{x^a}{a}+\frac{x^{a+2}}{a+2}+C\right)=\frac{1}{a}+\frac{x^2}{a+2}+Cx^{-a}.$$

由 $y(1)=1$，得 $C=\frac{a^2-2}{a(a+2)}$，故 $y(x,a)=\frac{1}{a}+\frac{x^2}{a+2}+\frac{a^2-2}{a(a+2)}x^{-a}$.

4. 伯努利方程（仅数学一要求）

例 1.13.5 求 $y\mathrm{d}x=(1+x\ln y)x\mathrm{d}y(y>0)$ 的通解.

解 方程变形为 $\frac{\mathrm{d}x}{\mathrm{d}y}-\frac{1}{y}x=\frac{\ln y}{y}x^2$，这是以 y 为自变量，x 为未知函数的伯努利方程（如果以 x 为自变量，y 为未知函数来解方程，是极其困难的，所以当我们遇到困难时，要学会“换位思考”——x 与 y，谁作为自变量，谁作为未知函数是可以互换的）.

①两边同时除以 x^2，并令 $z=x^{-1}$，有 $\frac{\mathrm{d}z}{\mathrm{d}y}=-\frac{1}{x^2}\frac{\mathrm{d}x}{\mathrm{d}y}$，于是方程化为 $\frac{\mathrm{d}z}{\mathrm{d}y}+\frac{1}{y}z=-\frac{\ln y}{y}$；

②应用一阶线性微分方程的通解公式，得

$$z=\frac{1}{x}=\mathrm{e}^{-\ln y}\left[\int\left(-\frac{\ln y}{y}\mathrm{e}^{\ln y}\right)\mathrm{d}y+C\right]=\frac{1}{y}[y(1-\ln y)+C],$$

故通解为 $\frac{1}{x}=1-\ln y+\frac{C}{y}(y>0,C$ 为任意常数).

二、二阶可降阶微分方程的求解（仅数学一、数学二要求）

1. $y''=f(x,y')$ 型

例 1.13.6 求 $y''=\frac{2xy'}{1+x^2}$ 的通解.

解 设 $y'=p(x)$，则 $y''=p'$，将其代入方程后可得 $p'=\frac{2xp}{1+x^2}$.

此方程为可分离变量方程，分离变量得 $\frac{\mathrm{d}p}{p}=\frac{2x}{1+x^2}\mathrm{d}x$，两边积分并化简，得 $p=C_1(1+x^2)$，从而有 $y'=C_1(1+x^2)$. 再两边积分可得原方程的通解为 $y=C_1\left(x+\frac{1}{3}x^3\right)+C_2$，其中 C_1,C_2 为任意常数.

2. $y''=f(y,y')$ 型

例 1.13.7 求微分方程 $2yy''+(y')^2=0$ 的通解，其中 $y>0$.

解 ①令 $y'=p$，则 $y''=p\frac{\mathrm{d}p}{\mathrm{d}y}$，于是原方程可化为一阶方程 $2yp\frac{\mathrm{d}p}{\mathrm{d}y}+p^2=0$，即 $p\left(2y\frac{\mathrm{d}p}{\mathrm{d}y}+p\right)=0$，故 $2y\frac{\mathrm{d}p}{\mathrm{d}y}+p=0$，分离变量得 $\frac{\mathrm{d}p}{p}=-\frac{\mathrm{d}y}{2y}$，两边积分后有

$$\ln|p|=-\frac{1}{2}\ln y+\ln C_0,$$

从而 $p=\pm C_0\frac{1}{\sqrt{y}}=C_1\frac{1}{\sqrt{y}}$.

②由 $p=\frac{\mathrm{d}y}{\mathrm{d}x}$，则$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{C_1}{\sqrt{y}}$，分离变量，积分得$\frac{2}{3}y^{\frac{3}{2}}=C_1x+C_2$，即原方程的通解为

$$y=(Cx+D)^{\frac{2}{3}}，其中 C,D 为任意常数，且 C\neq 0.$$

【注】 由 $p=0$ 得 $y'=0$，此时可解得 $y=C$，这又是漏掉的“奇解”.

三、高阶线性微分方程的求解

例 1.13.8 求 $y''-4y'+4y=3xe^{2x}$的通解.

解 特征方程为 $\lambda^2-4\lambda+4=0$，特征根为 $\lambda_1=\lambda_2=2$，于是相应的齐次方程的通解是

$$y=(C_1+C_2x)e^{2x}.$$

由于 $\lambda=2$ 是特征方程的重根，而 $P_n(x)=3x$ 为一次多项式，故应设原方程的特解为 $y^*=x^2(ax+b)e^{2x}$.

将特解代入原式有 $6ax+2b=3x$，比较系数可得 $a=\frac{1}{2}$，$b=0$. 故原方程的一个特解为 $y^*=\frac{1}{2}x^3e^{2x}$，其通解是

$$y=(C_1+C_2x)e^{2x}+\frac{1}{2}x^3e^{2x}，其中 C_1,C_2 为任意常数.$$

例 1.13.9 微分方程 $y''-4y'+3y=e^x\cos x+xe^{3x}$对应齐次微分方程的通解为$\bar{y}$=________，它的特解形式为 y^*=________.

解 应填 $C_1e^x+C_2e^{3x}$，其中 C_1,C_2 为任意常数；$e^x(A\cos x+B\sin x)+x(ax+b)e^{3x}$.

特征方程为 $\lambda^2-4\lambda+3=0$，解得 $\lambda_1=1,\lambda_2=3$，故对应的齐次微分方程的通解为$\bar{y}=C_1e^x+C_2e^{3x}$，其中 C_1,C_2 为任意常数.

非齐次微分方程特解形式的假设，可分为两个方程进行：

$$y''-4y'+3y=e^x\cos x, \qquad ①$$

$$y''-4y'+3y=xe^{3x}. \qquad ②$$

①式的特解形式是 $y_1^*=e^x(A\cos x+B\sin x)$；②式的特解形式应是 $y_2^*=x(ax+b)e^{3x}$，则 $y_1^*+y_2^*=y^*$ 即是原方程的特解形式.

四、根据微分方程的概念解题

并非所有方程都要求解后才能解决问题，这里给出用微分方程的概念而不用求解的例子.

1. 已知微分方程的解，反求系数

例 1.13.10 设 y_1,y_2是一阶非齐次线性微分方程 $y'+p(x)y=q(x)$的两个特解，若常数 λ,μ 使 $\lambda y_1+\mu y_2$是该方程的解，$\lambda y_1-\mu y_2$是该方程对应的齐次方程的解，则(　　).

(A)$\lambda=\frac{1}{2},\mu=\frac{1}{2}$　　(B)$\lambda=-\frac{1}{2},\mu=-\frac{1}{2}$　　(C)$\lambda=\frac{2}{3},\mu=\frac{1}{3}$　　(D)$\lambda=\frac{2}{3},\mu=\frac{2}{3}$

解　应选(A).

将解 $\lambda y_1+\mu y_2$ 代入方程 $y'+p(x)y=q(x)$，得 $\lambda[y_1'+p(x)y_1]+\mu[y_2'+p(x)y_2]=q(x)$.

又 $y_1'+p(x)y_1=q(x)$，$y_2'+p(x)y_2=q(x)$，故

$$\lambda+\mu=1. \qquad ①$$

将解 $\lambda y_1-\mu y_2$ 代入方程 $y'+p(x)y=0$，得 $\lambda[y_1'+p(x)y_1]-\mu[y_2'+p(x)y_2]=0$.

又 $y_1'+p(x)y_1=q(x)$，$y_2'+p(x)y_2=q(x)$，故

$$\lambda-\mu=0. \qquad ②$$

联立①，②两式，得 $\lambda=\frac{1}{2},\mu=\frac{1}{2}$，所以选择(A).

2. 不解微分方程，而利用方程所隐含的信息解题

$F(y,y',y'',\cdots,y^{(n)})=0$ 反映了未知函数 y 及其各阶导数之间的关系，故可以充分利用此关系来解决问题.

例 1.13.11　设 $y=f(x)$ 是方程 $y''-2y'+4y=0$ 的一个解，若 $f(x_0)>0$，且 $f'(x_0)=0$，则函数 $f(x)$ 在点 x_0(　　).

(A)取得极大值　　(B)取得极小值

(C)某个邻域内单调增加　　(D)某个邻域内单调减少

解　应选(A).

乍一看题目，$y''-2y'+4y=0$ 是《考试大纲》要求的一个很简单的微分方程，于是很多同学便去写它的解，然后再去讨论问题. 姑且不论这样做能否解决问题(事实上由于初始条件不够，是解不出特解的)，考场上时间有限，需要的是效率，是最好的解题方法. 希望大家在平时的复习中，努力研究最"恰当"的方法，这种"恰当"也正是命题人想考查你的地方.

由题设，有 $f''(x_0)-2f'(x_0)+4f(x_0)=0$，结合 $f(x_0)>0$，$f'(x_0)=0$，可得 $f''(x_0)=-4f(x_0)<0$，又 $f'(x_0)=0$，由极值判别的第二充分条件知，点 $x=x_0$ 是 $f(x)$ 的极大值点. 答案选择(A).

基础习题精练

习 题

1.13.1　求微分方程 $\left(x\frac{\mathrm{d}y}{\mathrm{d}x}-y\right)\arctan\frac{y}{x}=x$ 的通解.

1.13.2　求微分方程 $x\frac{\mathrm{d}y}{\mathrm{d}x}=y+\sqrt{x^2-y^2}\,(x>0)$ 的通解.

1.13.3　求微分方程 $y'+1=\mathrm{e}^{-y}\sin x$ 的通解.

1.13.4　设函数 $f(t)$ 在 $[0,+\infty)$ 上连续，且满足方程 $f(t)=\mathrm{e}^{4\pi t^2}+\iint\limits_{x^2+y^2\leqslant 4t^2}f\left(\frac{1}{2}\sqrt{x^2+y^2}\right)\mathrm{d}x\mathrm{d}y$，

求 $f(t)$.

1.13.5 求 $(x+2)y''+x(y')^2=y'$ 的通解.

1.13.6 求 $y''-3y'+2y=2e^{-x}\cos x+e^{2x}(4x+5)$ 的通解.

1.13.7 设 $y=y(x)$ 是二阶常系数线性微分方程 $y''+py'+qy=e^{3x}$ 满足初始条件 $y(0)=y'(0)=0$ 的特解，则当 $x\to 0$ 时，函数 $\dfrac{\ln(1+x^2)}{y(x)}$ 的极限（　　）.

(A)不存在　　(B)等于 1　　(C)等于 2　　(D)等于 3

解 答

1.13.1 分析 所求微分方程为齐次型微分方程，只要作代换 $u=\dfrac{y}{x}$ 解之即可.

解 方程变形为 $\dfrac{dy}{dx}=\dfrac{1}{\arctan\dfrac{y}{x}}+\dfrac{y}{x}$，令 $u=\dfrac{y}{x}$，则有 $\arctan u\,du=\dfrac{dx}{x}$，两边积分，得

$$u\arctan u-\int\frac{u}{1+u^2}du=\int\frac{dx}{x},$$

所以有 $u\arctan u-\dfrac{1}{2}\ln(1+u^2)=\ln|x|+\ln C$，即 $u\arctan u=\ln(C\cdot|x|\sqrt{1+u^2})$.

代回 $u=\dfrac{y}{x}$，得 $\dfrac{y}{x}\arctan\dfrac{y}{x}=\ln(C\sqrt{x^2+y^2})$. 即得原方程的通解为 $C\sqrt{x^2+y^2}=e^{\frac{y}{x}\arctan\frac{y}{x}}$，其中 C 为任意正常数.

1.13.2 分析 变形和作适当代换后可变为可分离变量的方程.

解 方程两边同除以 x，得

$$\frac{dy}{dx}=\frac{y}{x}+\frac{1}{x}\sqrt{x^2-y^2}.$$

因 $x>0$，故

$$\frac{dy}{dx}=\frac{y}{x}+\sqrt{1-\left(\frac{y}{x}\right)^2},$$

作变换，令 $\dfrac{y}{x}=u$，有 $u+x\dfrac{du}{dx}=u+\sqrt{1-u^2}$，即 $\dfrac{du}{\sqrt{1-u^2}}=\dfrac{dx}{x}$. 解之，得 $\arcsin u=\ln Cx$.

再将 $u=\dfrac{y}{x}$ 代回，得原方程的通解为 $\arcsin\dfrac{y}{x}=\ln Cx$，即 $y=x\sin(\ln Cx)$，其中 C 为任意正常数.

1.13.3 解 将方程 $y'+1=e^{-y}\sin x$ 变形为 $(e^y)'+e^y=\sin x$，这是关于 e^y 的一阶线性微分方程. 利用一阶非齐次线性微分方程的通解公式，得

$$e^y=e^{-\int dx}\left(C+\int\sin x\cdot e^{\int dx}dx\right)=e^{-x}\left(C+\int e^x\sin x\,dx\right)$$
$$=e^{-x}\left[C+\frac{1}{2}e^x(\sin x-\cos x)\right].$$

【注】 变量替换思想是求解微分方程问题的重要方法.

1.13.4 解 显然 $f(0)=1$，由于

$$\iint\limits_{x^2+y^2\leqslant 4t^2}f\left(\frac{1}{2}\sqrt{x^2+y^2}\right)dxdy=\int_0^{2\pi}d\theta\int_0^{2t}f\left(\frac{1}{2}r\right)rdr=2\pi\int_0^{2t}rf\left(\frac{1}{2}r\right)dr,$$

可见
$$f(t)=e^{4\pi t^2}+2\pi\int_0^{2t} rf\left(\frac{1}{2}r\right)dr,\quad f'(t)=8\pi te^{4\pi t^2}+8\pi tf(t).$$

解上述关于 $f(t)$ 的一阶非齐次线性微分方程，得
$$f(t)=\left(\int 8\pi te^{4\pi t^2}e^{-\int 8\pi t dt}dt+C\right)e^{\int 8\pi t dt}=\left(8\pi\int tdt+C\right)e^{4\pi t^2}=(4\pi t^2+C)e^{4\pi t^2},$$
将 $f(0)=1$ 代入，得 $C=1$. 因此 $f(t)=(4\pi t^2+1)e^{4\pi t^2}$.

1.13.5 **解** 令 $y'=p$，有 $y''=\frac{dp}{dx}$，原式成为 $(x+2)\frac{dp}{dx}+xp^2=p$，即 $\frac{dp}{dx}-\frac{p}{x+2}=-\frac{x}{x+2}p^2$. 两边同除以 $-p^2$，化为 $\frac{dp^{-1}}{dx}+\frac{p^{-1}}{x+2}=\frac{x}{x+2}$，得
$$p^{-1}=e^{-\int\frac{1}{x+2}dx}\left(\int\frac{x}{x+2}e^{\int\frac{1}{x+2}dx}dx+C\right)=\frac{1}{x+2}\left(\int xdx+C\right)=\frac{1}{x+2}\left(\frac{x^2}{2}+C\right).$$

解出
$$p=\frac{2(x+2)}{x^2+2C}\xlongequal{\text{改写为}}\frac{2x+4}{x^2+C_1}(C_1=2C),\tag{*}$$

①当 $C_1>0$ 时，得 $y=\ln(x^2+C_1)+\frac{4}{\sqrt{C_1}}\arctan\frac{x}{\sqrt{C_1}}+C_2$；

②当 $C_1=0$ 时，得 $y=2\ln|x|-\frac{4}{x}+C_2$；

③当 $C_1<0$ 时，得 $y=\ln|x^2+C_1|+\frac{2}{\sqrt{-C_1}}\ln\left|\frac{x-\sqrt{-C_1}}{x+\sqrt{-C_1}}\right|+C_2$，

其中 C_2 为任意常数.

> **【注】** 若题中给出初始条件，例如 $y(0)=1,y'(0)=4$，则应由式（*）先定出 $C_1=1$，然后再积分. 由于此时 $C_1>0$，所以不必再讨论 C_1 的各种情况. 可见，应尽早代入初始条件定出任意常数.

1.13.6 **解** 特征方程为 $\lambda^2-3\lambda+2=0$. 其解为 $\lambda_1=2,\lambda_2=1$. 因此对应的齐次微分方程的通解是
$$\bar{y}=C_1e^x+C_2e^{2x}.$$

为求非齐次微分方程的一个特解，将原方程分解为两个方程：
$$y''-3y'+2y=2e^{-x}\cos x,\tag{①}$$
$$y''-3y'+2y=e^{2x}(4x+5).\tag{②}$$

方程①的一个特解可设为 $Y_1=e^{-x}(A\cos x+B\sin x)$，求得
$$Y_1'=e^{-x}[(B-A)\cos x-(A+B)\sin x],$$
$$Y_1''=e^{-x}(-2B\cos x+2A\sin x).$$

代入方程①解得 $A=\frac{1}{5},B=-\frac{1}{5}$，即 $Y_1=\frac{e^{-x}}{5}(\cos x-\sin x)$.

方程②的一个特解设为 $Y_2=e^{2x}\cdot x(ax+b)$，求得
$$Y_2'=e^{2x}[2ax^2+2(a+b)x+b],\quad Y_2''=e^{2x}[4ax^2+(8a+4b)x+2a+4b].$$

代入方程②解得 $a=2,b=1$，即 $Y_2=e^{2x}(2x^2+x)$.

因此
$$Y=Y_1+Y_2=\frac{e^{-x}}{5}(\cos x-\sin x)+e^{2x}(2x^2+x)$$
是原方程的一个特解，从而原方程的通解为
$$y=\bar{y}+Y=C_1e^x+(C_2+x+2x^2)e^{2x}+\frac{1}{5}e^{-x}(\cos x-\sin x),$$

其中 C_1,C_2 为任意常数.

1.13.7　(C)　**解**　不用去求解微分方程，而是利用方程所反映出来的函数与各阶导数之间的关系来解题.

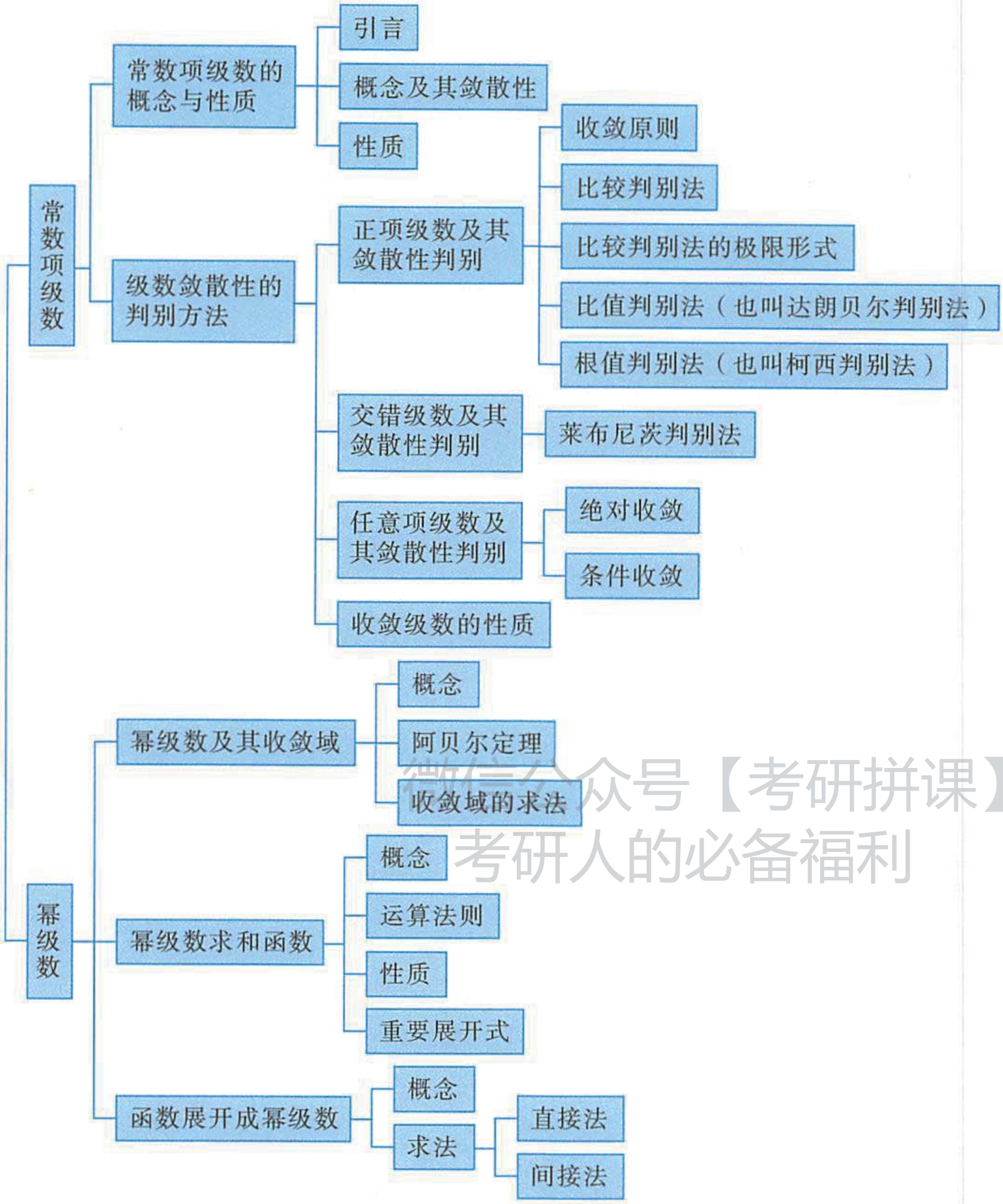
由 $y''+py'+qy=e^{3x}$ 得 $y''(x)$ 连续，且 $y''(0)=-py'(0)-qy(0)+e^0=1$，而

$$\lim_{x\to0}\frac{\ln(1+x^2)}{y(x)}=\lim_{x\to0}\frac{x^2}{y(x)}=\lim_{x\to0}\frac{2x}{y'(x)}=\lim_{x\to0}\frac{2}{y''(x)}=\frac{2}{y''(0)},$$

故所求极限等于 2，应选(C).

第14讲 无穷级数（仅数学一、数学三要求）

基础知识结构



基础内容精讲

一、常数项级数的概念与性质

1. 引言

古希腊数学家芝诺曾经讲了一个有趣的悖论：一个人从 A 点走到 B 点，先走完路程的 $\frac{1}{2}$，然后走完剩下路程的 $\frac{1}{2}$，再走完剩下路程的 $\frac{1}{2}$……依此类推，一直走下去，永远走不到终点.

我们不妨将芝诺的文字描述写成如下的数学表达式.

假设此人匀速前进走完路程的 $\frac{1}{2}$ 要用半个小时，则按照芝诺的描述，从 A 点到 B 点所用总时间 T 可以写成：

$$T=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots.$$

这便出现了无穷多项相加的式子，且越往后的项就越小，这“无穷多个越来越小的项加起来”的表达式合理吗？芝诺认为不对，其核心问题就在于以前人们所接触的都是**有限项相加**，项数再多，也加得完；现在出现了无穷多项相加，用所谓的加法是永远加不到尽头的. 接下来，我们就要开始研究这个总时间 T 的表达式到底是什么，该如何处理它.

2. 概念及其敛散性

给定一个无穷数列 $u_1,u_2,\cdots,u_n,\cdots$，将其各项用加号连起来得到的**记号** $\sum\limits_{n=1}^{\infty}u_n$，即

$$\sum_{n=1}^{\infty}u_n=u_1+u_2+\cdots+u_n+\cdots$$

叫作**无穷级数**，简称**级数**，其中 u_n 叫作该级数的**通项**. 若 u_n 是常数而不是函数，则 $\sum\limits_{i=1}^{\infty}u_n$ 就被更为细致地称为**常数项无穷级数**. 简称**常数项级数**. 正如在引言中所写出的总时间 $T=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots=\sum\limits_{n=1}^{\infty}\frac{1}{2^n}$，这里的通项 $u_n=\frac{1}{2^n}$. 之所以称 $\sum\limits_{n=1}^{\infty}u_n$ 为一个记号，是说这种相加只是**形式上的**，因为无穷多项是不可能逐一相加求和的，那么该如何处理“无穷多项相加”呢？

现将级数中的各项逐一相加，得到下面这些和：

$$S_1=u_1,\quad S_2=u_1+u_2,\quad S_3=u_1+u_2+u_3,\cdots,$$

$$S_n=u_1+u_2+\cdots+u_n,\cdots,$$

这称为级数的**部分和**，$\{S_n\}$ 就是级数的**部分和数列**. 显然，我们愿意去研究当 $n\to\infty$ 时所发生的事情，因为 $\lim\limits_{n\to\infty}S_n=\sum\limits_{n=1}^{\infty}u_n$，这便道出了无穷多项相加的方法：用极限工具来处理. 若 $\lim\limits_{n\to\infty}S_n=S$（一个存在的有限数），则 $\sum\limits_{n=1}^{\infty}u_n$ **收敛**，并称 S 为该收敛级数 $\sum\limits_{n=1}^{\infty}u_n$ 的**和**；若 $\lim\limits_{n\to\infty}S_n$ 不存在或为 $\pm\infty$，则称 $\sum\limits_{n=1}^{\infty}u_n$ **发散**. 雅各布 · 伯努利曾在自己的文章中这样写道：“末项消逝的无穷级数之和**有时有限而有时无限或不定**”. 这里的“有限”与

无限或不定”就分别对应着“收敛”与“发散”.

研究一个级数$\sum_{n=1}^{\infty} u_n$是收敛还是发散,也可以简单说成研究级数$\sum_{n=1}^{\infty} u_n$的**敛散性**.

3. 性质

性质1 若级数$\sum_{n=1}^{\infty} u_n$,$\sum_{n=1}^{\infty} v_n$均收敛,且其和分别为S,T,则任给常数a,b,有$\sum_{n=1}^{\infty}(au_n+bv_n)$也收敛,且其和为$aS+bT$,即

$$\sum_{n=1}^{\infty}(au_n+bv_n)=a\sum_{n=1}^{\infty} u_n+b\sum_{n=1}^{\infty} v_n.$$

这个性质称为收敛级数的线性性质.

在讲下一个重要性质前,先给出一个概念.

在级数$\sum_{n=1}^{\infty} u_n$中**去掉前m项**,得

$$\sum_{n=m+1}^{\infty} u_n=u_{m+1}+u_{m+2}+\cdots+u_{m+k}+\cdots,$$

这叫作级数$\sum_{n=1}^{\infty} u_n$的**m项后余项**.

性质2 若级数$\sum_{n=1}^{\infty} u_n$收敛,则其任意m项后余项$\sum_{n=m+1}^{\infty} u_n$也收敛;反之,若存在$m$项后余项$\sum_{n=m+1}^{\infty} u_n$收敛,则$\sum_{n=1}^{\infty} u_n$也收敛.

【注】 一个级数的前面部分去掉或加上任意有限项,不会改变该级数的敛散性.事实上也可说成**改变级数任意有限项,不会改变该级数的敛散性.**

性质3 若$\sum_{n=1}^{\infty} u_n$收敛,则$\lim\limits_{n\to\infty} u_n=0$.

证明 由于$u_n=S_n-S_{n-1}$,故$\lim\limits_{n\to\infty} u_n=\lim\limits_{n\to\infty}(S_n-S_{n-1})=\lim\limits_{n\to\infty} S_n-\lim\limits_{n\to\infty} S_{n-1}=S-S=0$.证毕.

证明虽简单,但有几句话要说,此性质是**级数收敛的必要条件**,故其逆否命题:若$\lim\limits_{n\to\infty} u_n\neq 0$,则$\sum_{n=1}^{\infty} u_n$必发散.且即使$\lim\limits_{n\to\infty} u_n=0$,也不能保证$\sum_{n=1}^{\infty} u_n$收敛,见例1.14.2.

二、级数敛散性的判别方法

1. 正项级数及其敛散性判别

若通项$u_n\geqslant 0,n=1,2,\cdots$,则称$\sum_{n=1}^{\infty} u_n$为**正项级数**.这种级数在形式上的简单性最容易让读者接受,下面就来讲这种级数的各种经典的敛散性判别方法.

(1) 收敛原则.

正项级数$\sum_{n=1}^{\infty} u_n$收敛的充分必要条件是它的部分和数列$\{S_n\}$有界.

证明 必要性.由于$u_n\geqslant 0$,故$S_n=u_1+u_2+\cdots+u_n\geqslant 0$,且$S_1\leqslant S_2\leqslant\cdots\leqslant S_n\leqslant\cdots$,$\{S_n\}$天生便是一个单调不减且下界为0的数列.当$\sum_{n=1}^{\infty} u_n$收敛时,$\lim\limits_{n\to\infty} S_n$存在,则$\{S_n\}$必然有上界.上界、下界都有,则

$\{S_n\}$ 有界.

充分性. 若 $\{S_n\}$ 有界，不要忘记 $\{S_n\}$ 天生单调不减，则根据单调有界准则，可得 $\{S_n\}$ 收敛，也即 $\lim\limits_{n\to\infty}S_n$ 存在，于是 $\sum\limits_{n=1}^{\infty}u_n$ 收敛. 证毕.

【注】 读者不难发现，由于 $\{S_n\}$ 天生单调不减，$\lim\limits_{n\to\infty}S_n$ 只有两种可能的结果：若 $\{S_n\}$ 有界，则 $\lim\limits_{n\to\infty}S_n=S$（有限数）；若 $\{S_n\}$ 无界，则 $\lim\limits_{n\to\infty}S_n=+\infty$. 除此之外，再无其他结果.

直接使用收敛原则来判别正项级数敛散性的情形并不多见（见习题 1.14.2），因为正项级数部分和有界性的证明往往并不容易. 下面我们将利用收敛原则来给出更为实用的正项级数敛散性判别法.

（2）比较判别法.

给出两个正项级数 $\sum\limits_{n=1}^{\infty}u_n$ 和 $\sum\limits_{n=1}^{\infty}v_n$，如果从某项起有 $u_n\leqslant v_n$ 成立，则

① 若 $\sum\limits_{n=1}^{\infty}v_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}u_n$ 也收敛；

② 若 $\sum\limits_{n=1}^{\infty}u_n$ 发散，则 $\sum\limits_{n=1}^{\infty}v_n$ 也发散.

有人常通俗地念成：**两个正项级数，若大的收敛，则小的必收敛；若小的发散，则大的必发散**. 对于记忆来讲，这种念法未尝不可.

（3）比较判别法的极限形式.

作为比较判别法的推论，这个标题下的方法将更为实用. 给出两个正项级数 $\sum\limits_{n=1}^{\infty}u_n$ 和 $\sum\limits_{n=1}^{\infty}v_n$，$v_n\neq 0$，且

$$\lim_{n\to\infty}\frac{u_n}{v_n}=A.$$

① 若 $A=0$，则当 $\sum\limits_{n=1}^{\infty}v_n$ 收敛时，$\sum\limits_{n=1}^{\infty}u_n$ 也收敛；

② 若 $A=+\infty$，则当 $\sum\limits_{n=1}^{\infty}v_n$ 发散时，$\sum\limits_{n=1}^{\infty}u_n$ 也发散；

③ 若 $0<A<+\infty$，则 $\sum\limits_{n=1}^{\infty}u_n$ 与 $\sum\limits_{n=1}^{\infty}v_n$ 有相同的敛散性.

（4）比值判别法（也叫达朗贝尔判别法）.

给出一正项级数 $\sum\limits_{n=1}^{\infty}u_n$，如果 $\lim\limits_{n\to\infty}\dfrac{u_{n+1}}{u_n}=\rho$，那么

① 若 $\rho<1$，则 $\sum\limits_{n=1}^{\infty}u_n$ 收敛；

② 若 $\rho>1$，则 $\sum\limits_{n=1}^{\infty}u_n$ 发散.

【注】 需要指出，若 $\rho=1$，无法用此法判定 $\sum\limits_{n=1}^{\infty}u_n$ 的敛散性. 比如对 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n}$，$u_n=\dfrac{1}{n}$，则 $\lim\limits_{n\to\infty}\dfrac{u_{n+1}}{u_n}=\lim\limits_{n\to\infty}\dfrac{\dfrac{1}{n+1}}{\dfrac{1}{n}}=\lim\limits_{n\to\infty}\dfrac{n}{n+1}=1$；对 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^2}$，$u_n=\dfrac{1}{n^2}$，则 $\lim\limits_{n\to\infty}\dfrac{u_{n+1}}{u_n}=\lim\limits_{n\to\infty}\dfrac{\dfrac{1}{(n+1)^2}}{\dfrac{1}{n^2}}=\lim\limits_{n\to\infty}\left(\dfrac{n}{n+1}\right)^2=1$. 你看，前者发散，后者收敛，但都有 $\rho=1$.

(5) 根值判别法(也叫柯西判别法).

给出一正项级数$\sum_{n=1}^{\infty} u_n$，如果$\lim_{n\to\infty} \sqrt[n]{u_n}=\rho$，那么

① 若$\rho<1$，则$\sum_{n=1}^{\infty} u_n$收敛；

② 若$\rho>1$，则$\sum_{n=1}^{\infty} u_n$发散.

【注】 同样要指出，若$\rho=1$，根值判别法也会失效. 如$\sum_{n=1}^{\infty}\frac{1}{n}$发散，有$\lim_{n\to\infty}\sqrt[n]{\frac{1}{n}}=\rho=1$；$\sum_{n=1}^{\infty}\frac{1}{n^2}$收敛，也有$\lim_{n\to\infty}\sqrt[n]{\frac{1}{n^2}}=\rho=1$.

2. 交错级数及其敛散性判别

若级数**各项正负相间**出现，称这样的级数为**交错级数**，一般写为

$$\sum_{n=1}^{\infty}(-1)^{n-1}u_n=u_1-u_2+u_3-u_4+\cdots+(-1)^{n-1}u_n+\cdots,$$

其中$u_n>0(n=1,2,\cdots)$，以使各项的正负号明显地呈现出来，从而表达出各项依次正负相间的独特规律.

显然，交错级数在形式上比正项级数复杂，那么它的敛散性判别是否也同样复杂了呢？莱布尼茨回答了这个问题. 1714年，在写给约翰·伯努利的信件中，莱布尼茨给出了下面这个定理，史称**莱布尼茨判别法**：

给出一交错级数$\sum_{n=1}^{\infty}(-1)^{n-1}u_n$，$u_n>0$，$n=1,2,\cdots$，若$\{u_n\}$**单调不增**且$\lim_{n\to\infty}u_n=0$，则该级数收敛.

3. 任意项级数及其敛散性判别

若级数各项可正、可负、亦可为零，称这样的级数为**任意项级数**，写为$\sum_{n=1}^{\infty}u_n$，这里u_n符号不作限制. 对于这种级数敛散性的判别方法，一般认为超出了本科高等数学的教学要求，本书也不予讨论. 在高等数学中，是按下述方式进行研究的.

给任意项级数的每一项加上绝对值，写成$\sum_{n=1}^{\infty}|u_n|$，这样，就使得$|u_n|\geqslant 0$，成了正项级数，它叫作原级数$\sum_{n=1}^{\infty}u_n$**的绝对值级数**.

因此绝对值级数就自然站到了正项级数的队伍中，前面讲过的五种正项级数的敛散性判别法均可派上用场了. 不过，绝对值级数$\sum_{n=1}^{\infty}|u_n|$与原级数$\sum_{n=1}^{\infty}u_n$的敛散性有何关系呢？看下面的定义与定理.

定义1 设$\sum_{n=1}^{\infty}u_n$为任意项级数，若$\sum_{n=1}^{\infty}|u_n|$收敛，则称$\sum_{n=1}^{\infty}u_n$**绝对收敛**.

定义2 设$\sum_{n=1}^{\infty}u_n$为任意项级数，若$\sum_{n=1}^{\infty}u_n$收敛，但$\sum_{n=1}^{\infty}|u_n|$发散，则称$\sum_{n=1}^{\infty}u_n$**条件收敛**.

定理 若$\sum_{n=1}^{\infty}|u_n|$收敛(即任意项级数$\sum_{n=1}^{\infty}u_n$绝对收敛)，则$\sum_{n=1}^{\infty}u_n$必收敛.

4. 收敛级数的性质

在初等数学中，人们都知道，有限多个数相加，满足结合律及交换律. 对于无穷级数$\sum_{n=1}^{\infty}u_n$，显然这里所

有的项绝非简单的相加．它所含有的**极限运算过程**导致我们必须相当警觉加法运算律是否还能天然的成立．如果不能天然的成立，我们还要添加什么前提条件．下面就来研究这个问题．

性质4　收敛级数的项任意加括号后所得的新级数仍收敛，且其和不变．

证明　给出一收敛级数$\sum\limits_{n=1}^{\infty}u_n$，其部分和数列记为$\{S_n\}$：$S_1,S_2,\cdots,S_n,\cdots$，且$\lim\limits_{n\to\infty}S_n=S$（存在）．将此级数的项任意加括号，得新级数

$$(u_1+\cdots+u_{n_1})+(u_{n_1+1}+\cdots+u_{n_2})+\cdots+(u_{n_{k-1}+1}+\cdots+u_{n_k})+\cdots,$$

这个级数的部分和数列记为

$$\{S_{n_k}\}:S_{n_1},S_{n_2},\cdots,S_{n_k},\cdots,k=1,2,\cdots,$$

显然它是$\{S_n\}$的某一个子列，故$\{S_{n_k}\}$亦收敛于S．证毕．

根据性质4，如下两个结论也需注意．

(1) 若加括号后得到的新级数发散，则原级数必然发散（逆否命题）．

(2) 若加括号后得到的新级数收敛，不能断言原级数一定收敛．如级数$\sum\limits_{n=1}^{\infty}(-1)^{n-1}a$，其中常数$a\neq 0$，若从第一项起，两项两项地加括号，可得$(a-a)+(a-a)+\cdots=0$收敛，但原级数是发散的．

性质5　若原级数绝对收敛，**不论将其各项如何重新排列**，所得的新级数也绝对收敛，且其和不变．（这个性质就是在说：绝对收敛的级数具有可交换性，此性质是德国数学家狄利克雷给出的．）

三、幂级数及其收敛域

1. 概念

(1)函数项级数．

设函数列$\{u_n(x)\}$定义在区间I上，称

$$u_1(x)+u_2(x)+u_3(x)+\cdots+u_n(x)+\cdots$$

为定义在区间I上的**函数项级数**，记为$\sum\limits_{n=1}^{\infty}u_n(x)$，当$x$取确定的值$x_0$时，$\sum\limits_{n=1}^{\infty}u_n(x)$成为常数项级数$\sum\limits_{n=1}^{\infty}u_n(x_0)$．

(2)幂级数．

若$\sum\limits_{n=1}^{\infty}u_n(x)$的一般项$u_n(x)$是$n$次幂函数，则称$\sum\limits_{n=1}^{\infty}u_n(x)$为**幂级数**，它是一种特殊且常用的函数项级数，其一般形式为

$$\sum_{n=0}^{\infty}a_n(x-x_0)^n=a_0+a_1(x-x_0)+a_2(x-x_0)^2+\cdots+a_n(x-x_0)^n+\cdots;$$

其标准形式为

$$\sum_{n=0}^{\infty}a_nx^n=a_0+a_1x+a_2x^2+\cdots+a_nx^n+\cdots,$$

其中a_n $(n=0,1,2,\cdots)$为**幂级数的系数**．

(3)收敛点与发散点．

若给定$x_0\in I$，有$\sum\limits_{n=1}^{\infty}u_n(x_0)$收敛，则称点$x_0$为幂级数$\sum\limits_{n=1}^{\infty}u_n(x)$的**收敛点**；若给定$x_0\in I$，有$\sum\limits_{n=1}^{\infty}u_n(x_0)$发

散，则称点 x_0 为幂级数 $\sum\limits_{n=1}^{\infty}u_n(x)$ 的**发散点**.

(4)收敛域.

函数项级数 $\sum\limits_{n=1}^{\infty}u_n(x)$ 的所有收敛点的集合称为它的**收敛域**.

要研究幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$，首要任务是判别敛散性，因为只有收敛了才有继续讨论它的意义. 具体来说，将某个 x_0 代入级数可得 $\sum\limits_{n=0}^{\infty}a_nx_0^n$，判别此数项级数是否收敛，我们的目标当然是找到所有收敛点的集合，即收敛域. 但问题来了，有多少个收敛点呢？如果有无穷多个收敛点，难道要一个一个地去验证么？显然，那样做是不可能的. 当人们正为此犯难时，一位年轻人提出了一个重要定理，请看下文.

2. 阿贝尔定理

当幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 在点 $x=x_1(x_1\neq 0)$ 处收敛时，对于满足 $|x|<|x_1|$ 的一切 x，幂级数**绝对收敛**；当幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 在点 $x=x_2(x_2\neq 0)$ 处发散时，对于满足 $|x|>|x_2|$ 的一切 x，幂级数**发散**.

多么伟大的阿贝尔，多么给力的办法！只是，不得不告诉大家，这样一个天才数学家生于 1802 年，一生的工作不为世人所重视，穷困潦倒，卒于 1829 年，仅仅活了 27 年，太遗憾！

阿贝尔

(1802—1829)

3. 收敛域的求法

①若 $\lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|=\rho$，则 $\sum\limits_{n=0}^{\infty}a_nx^n$ 的收敛半径 R 的表达式为

$$R=\begin{cases}\dfrac{1}{\rho}, & \rho\neq 0,\\ +\infty, & \rho=0,\\ 0, & \rho=+\infty.\end{cases}$$

②开区间 $(-R,R)$ 为幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 的收敛区间；单独考查幂级数在 $x=\pm R$ 处的敛散性就可以确定其收敛域为 $(-R,R)$ 或 $[-R,R)$ 或 $(-R,R]$ 或 $[-R,R]$.

【注】 以上讨论的研究对象都是标准幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$，对于一般幂级数 $\sum\limits_{n=0}^{\infty}a_n(x-x_0)^n$，有完全一致的说法，不再赘述.

四、幂级数求和函数

1. 概念

在收敛域上，记 $S(x)=\sum\limits_{n=1}^{\infty}u_n(x)$，并称 $S(x)$ 为 $\sum\limits_{n=1}^{\infty}u_n(x)$ 的和函数.

2. 运算法则

若幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 与 $\sum\limits_{n=0}^{\infty}b_nx^n$ 的收敛半径分别为 R_a 和 $R_b(R_a\neq R_b)$，则有

①$k\sum\limits_{n=0}^{\infty}a_nx^n=\sum\limits_{n=0}^{\infty}ka_nx^n,|x|<R_a$，$k$ 为常数；

②$\sum\limits_{n=0}^{\infty}a_nx^n\pm\sum\limits_{n=0}^{\infty}b_nx^n=\sum\limits_{n=0}^{\infty}(a_n\pm b_n)x^n,|x|<R=\min\{R_a,R_b\}$.

【注】 在实际运算中，可能出现需要改变通项、下标的问题，现总结求和运算中恒等变形方式如下.

(1)通项、下标一起变.

$$\sum_{n=k}^{\infty}a_nx^n=\sum_{n=k+l}^{\infty}a_{n-l}x^{n-l},$$

其中 l 为整数，可正可负可为 0.

(2) 只变下标，不变通项.

$$\sum_{n=k}^{\infty}a_nx^n=a_kx^k+a_{k+1}x^{k+1}+\cdots+a_{k+l-1}x^{k+l-1}+\sum_{n=k+l}^{\infty}a_nx^n.$$

(3) 只变通项，不变下标.

$$\sum_{n=k}^{\infty}a_nx^n=x^l\sum_{n=k}^{\infty}a_nx^{n-l}.$$

举例.

$$\sum_{n=0}^{\infty}a_nx^{2n}+\sum_{n=0}^{\infty}b_{n+1}x^{2n+2}=a_0x^0+\sum_{n=1}^{\infty}a_nx^{2n}+\sum_{n=1}^{\infty}b_nx^{2n}=a_0+\sum_{n=1}^{\infty}(a_n+b_n)x^{2n}.$$

3. 性质

(1)幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 的和函数 $S(x)$ 在其收敛区间 I 上连续，且如果幂级数在收敛区间的端点 $x=R$(或 $x=-R$)处收敛，则和函数 $S(x)$ 在 $(-R,R]$(或 $[-R,R)$)上连续.

(2)幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 的和函数 $S(x)$ 在其收敛域 I 上可积，且有逐项积分公式

$$\int_0^x S(t)\mathrm{d}t=\int_0^x\left(\sum_{n=0}^{\infty}a_nt^n\right)\mathrm{d}t=\sum_{n=0}^{\infty}a_n\int_0^x t^n\mathrm{d}t=\sum_{n=0}^{\infty}\frac{a_n}{n+1}x^{n+1}\ (x\in I),$$

逐项积分后所得到的幂级数与原级数有相同的收敛半径，但收敛域可能扩大.

(3)幂级数 $\sum\limits_{n=0}^{\infty}a_nx^n$ 的和函数 $S(x)$ 在其收敛区间 $(-R,R)$ 内可导，且有逐项求导公式

$$S'(x)=\left(\sum_{n=0}^{\infty}a_nx^n\right)'=\sum_{n=0}^{\infty}(a_nx^n)'=\sum_{n=1}^{\infty}na_nx^{n-1}\ (|x|<R),$$

逐项求导后所得到的幂级数与原级数有相同的收敛半径，但收敛域可能缩小.

4. 重要展开式

① $\mathrm{e}^x=\sum\limits_{n=0}^{\infty}\dfrac{x^n}{n!}=1+x+\dfrac{x^2}{2!}+\cdots+\dfrac{x^n}{n!}+\cdots,-\infty<x<+\infty$.

② $\dfrac{1}{1+x}=\sum\limits_{n=0}^{\infty}(-1)^nx^n=1-x+x^2-x^3+\cdots+(-1)^nx^n+\cdots,-1<x<1$.

③ $\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n=1+x+x^2+\cdots+x^n+\cdots,-1<x<1.$

④ $\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots+(-1)^{n-1}\frac{x^n}{n}+\cdots,-1<x\leqslant 1.$

⑤ $\sin x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$

$$=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\cdots,-\infty<x<+\infty.$$

⑥ $\cos x=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!}$

$$=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots+(-1)^n\frac{x^{2n}}{(2n)!}+\cdots,-\infty<x<+\infty.$$

⑦ $(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+\cdots,\begin{cases}x\in(-1,1),\text{当 }\alpha\leqslant-1,\\x\in(-1,1],\text{当 }-1<\alpha<0,\\x\in[-1,1],\text{当 }\alpha>0.\end{cases}$

这里,①～⑥右端 x 的取值范围是指收敛域,而对于⑦,问题比较复杂,其收敛区间的端点是否收敛与 α 的取值有关,可以证明(这里不证):

当 $\alpha\leqslant-1$ 时,收敛域为 $(-1,1)$;

当 $-1<\alpha<0$ 时,收敛域为 $(-1,1]$;

当 $\alpha>0$ 时,收敛域为 $[-1,1]$.

比如,当 $\alpha=-1$ 时得到

$$\frac{1}{1+x}=1-x+x^2-x^3+\cdots+(-1)^nx^n+\cdots,-1<x<1;$$

当 $\alpha=-\frac{1}{2}$ 时得到 $\frac{1}{\sqrt{1+x}}=1-\frac{1}{2}x+\frac{1\times3}{2\times4}x^2-\frac{1\times3\times5}{2\times4\times6}x^3+\cdots,-1<x\leqslant1.$

五、函数展开成幂级数

1. 概念

如果函数 $f(x)$ 在点 $x=x_0$ 处存在任意阶导数,则称

$$f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\cdots$$

为函数 $f(x)$ 在 x_0 处的**泰勒级数**,若收敛,则 $f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$.

特别地,当 $x_0=0$ 时,则称

$$f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+\cdots$$

为函数 $f(x)$ 的**麦克劳林级数**,若收敛,则 $f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n$.

它们都被称为函数展开成幂级数.

2. 求法

方法一(直接法) 逐个计算 $a_n=\frac{f^{(n)}(x_0)}{n!}$ $(n=0,1,\cdots)$,并代入

$$f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\cdots.$$

直接法太麻烦了，我们一般都不用.

方法二（间接法） 利用已知的幂级数展开式，通过变量代换、四则运算、逐项求导、逐项积分和待定系数等方法得到函数的展开式.

基础例题精解

一、正项级数 $\left(\sum\limits_{n=1}^{\infty}u_n, u_n\geqslant 0\right)$ 的敛散性判别法

例 1.14.1 判别几何级数(也叫等比级数)

$$\sum_{n=1}^{\infty}aq^{n-1}=a+aq+aq^2+\cdots+aq^{n-1}+\cdots$$

的敛散性，其中 $a\neq 0$.

解 在公比 q 满足 $|q|<1$ 的情形下，初等数学知识告诉我们，部分和

$$S_n=\frac{a(1-q^n)}{1-q},$$

且 $\lim\limits_{n\to\infty}q^n=0$，这样一来，$\lim\limits_{n\to\infty}S_n=\dfrac{a}{1-q}$，这个极限是存在的，故 $\sum\limits_{n=1}^{\infty}aq^{n-1}$ 收敛.

反之，当 $|q|\geqslant 1$ 时，$\sum\limits_{n=1}^{\infty}aq^{n-1}$ 是发散的，这是因为：

若 $q>1$，$\lim\limits_{n\to\infty}q^n=+\infty$，则 $\lim\limits_{n\to\infty}S_n$ 不存在；

若 $q=1$，$\sum\limits_{n=1}^{\infty}aq^{n-1}=a+a+\cdots+a+\cdots$（无穷个 a 相加），则 $\lim\limits_{n\to\infty}S_n$ 亦不存在；

（读者可轻易看出，当 $q\geqslant 1$ 时，$\lim\limits_{n\to\infty}S_n$ 为 $\pm\infty$，到底是 $+\infty$ 还是 $-\infty$，由 a 的正负性来决定.）

若 $q=-1$，$\sum\limits_{n=1}^{\infty}aq^{n-1}=a+(-a)+a+(-a)+\cdots$，这是一个十分著名且有趣的级数，欧拉、莱布尼茨都在这个级数的敛散性判别上犯了错误，今天的我们当然可以清晰地给出正确答案，你看，部分和 $S_1=a$，$S_2=0$，$S_3=a$，$S_4=0$，…，它们交错着等于 a 和 0，显然 $\lim\limits_{n\to\infty}S_n$ 不存在(不唯一自然不存在).

若 $q<-1$，$\lim\limits_{n\to\infty}q^n$ 不存在，则 $\lim\limits_{n\to\infty}S_n$ 不存在.

综上所述，当 $|q|\geqslant 1$ 时，$\lim\limits_{n\to\infty}S_n$ 均不存在，则 $\sum\limits_{n=1}^{\infty}aq^{n-1}$ 发散. 于是有

$$\text{几何级数}\sum_{n=1}^{\infty}aq^{n-1}\begin{cases}\text{发散}, & |q|\geqslant 1,\\ \text{收敛，其和为}\dfrac{a}{1-q}, & |q|<1,\end{cases}\text{其中 }a\neq 0.$$

【注】 显然，在引言中提出的芝诺的问题是本题的特例，$T=\sum\limits_{n=1}^{\infty}\dfrac{1}{2^n}=\sum\limits_{n=1}^{\infty}\dfrac{1}{2}\cdot\left(\dfrac{1}{2}\right)^{n-1}$，其首项 $a=\dfrac{1}{2}$，公比 $q=\dfrac{1}{2}$，是 $|q|<1$ 的情况，则 $\lim\limits_{n\to\infty}S_n=\dfrac{a}{1-q}=\dfrac{\frac{1}{2}}{1-\frac{1}{2}}=1$. 事实上，此人将恰好用一个小时从 A 走到 B.

例 1.14.2 判断级数 $\sum\limits_{n=1}^{\infty}\frac{1}{\sqrt{n}}$ 的敛散性.

解 显然，$\lim\limits_{n\to\infty}\frac{1}{\sqrt{n}}=0$，但其部分和 $S_n=1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\cdots+\frac{1}{\sqrt{n}}>n\cdot\frac{1}{\sqrt{n}}=\sqrt{n}$，由于$\lim\limits_{n\to\infty}\sqrt{n}=+\infty$，自然有 $\lim\limits_{n\to\infty}S_n=+\infty$(不存在)，则 $\sum\limits_{n=1}^{\infty}\frac{1}{\sqrt{n}}$ 发散.

例 1.14.3 判断级数 $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ 的敛散性.

解 $\sum\limits_{n=1}^{\infty}\frac{1}{n}=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\cdots$，这个级数从第二项起，每项都是其左右两个相邻项的调和平均数. 所谓数 c 是数 a 与数 b 的调和平均数是指它们之间满足下面这个式子：

$$\frac{1}{c}=\frac{1}{2}\left(\frac{1}{a}+\frac{1}{b}\right).$$

于是，$\sum\limits_{n=1}^{\infty}\frac{1}{n}$ 叫作调和级数.

下面来研究它的敛散性.

易知当 $x>0$ 时，有 $x>\ln(1+x)$，所以对于任意正整数 n，显然也有$\frac{1}{n}>\ln\left(1+\frac{1}{n}\right)$. 读者可看出我们试图先研究 $\sum\limits_{n=1}^{\infty}\ln\left(1+\frac{1}{n}\right)$ 的敛散性，然后用比较判别法得出结论.

$$\sum_{n=1}^{\infty}\ln\left(1+\frac{1}{n}\right)=\sum_{n=1}^{\infty}\ln\frac{n+1}{n},$$

其部分和

$$S_n=\ln\frac{2}{1}+\ln\frac{3}{2}+\cdots+\ln\frac{n+1}{n}=(\ln 2-\ln 1)+(\ln 3-\ln 2)+\cdots+[\ln(n+1)-\ln n]=\ln(n+1),$$

由于$\lim\limits_{n\to\infty}S_n=\lim\limits_{n\to\infty}\ln(n+1)=+\infty$，则 $\sum\limits_{n=1}^{\infty}\ln\left(1+\frac{1}{n}\right)$ 发散，根据正项级数的比较判别法，有 $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ 发散.

【注 1】 在 18 世纪初，莱布尼茨与雅各布・伯努利分别独立地证明了"调和级数的和是无限的"这个结论.

【注 2】 $\sum\limits_{n=1}^{\infty}\frac{1}{n^p}$ 叫作 p **级数**，调和级数 $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ 是 p 级数中 $p=1$ 时的特殊情况.

这里不加证明地给出重要结论，

$$p\text{ 级数}\sum_{n=1}^{\infty}\frac{1}{n^p}\begin{cases}\text{发散}, & p\leqslant 1,\\ \text{收敛}, & p>1.\end{cases}$$

这个十分重要的结论在以后的多种场合都会派上用场，读者需牢记.

例 1.14.4 判断级数 $\sum\limits_{n=1}^{\infty}\frac{|a|^n n!}{n^n}$的敛散性，其中 a 为非零常数.

解 $\lim\limits_{n\to\infty}\frac{u_{n+1}}{u_n}=\lim\limits_{n\to\infty}\frac{|a|^{n+1}\cdot(n+1)!}{(n+1)^{n+1}}\cdot\frac{n^n}{|a|^n\cdot n!}=\lim\limits_{n\to\infty}\frac{|a|}{\left(1+\frac{1}{n}\right)^n}=\frac{|a|}{\mathrm{e}}.$

这里 $\rho=\frac{|a|}{\mathrm{e}}$，显然当 $0<|a|<\mathrm{e}$ 时，原级数收敛；当 $|a|>\mathrm{e}$ 时，原级数发散；当 $|a|=\mathrm{e}$ 时，由于当 $n\to\infty$ 时，$\left(1+\frac{1}{n}\right)^n$ 单调增加趋向于 e，所以 $\frac{u_{n+1}}{u_n}=\frac{\mathrm{e}}{\left(1+\frac{1}{n}\right)^n}>1, n=1,2,\cdots$，从而知 $\lim\limits_{n\to\infty}u_n\neq0$，由级数收敛的必要条件，知原级数发散.

综上所述，当 $0<|a|<\mathrm{e}$ 时，原级数收敛，当 $|a|\geqslant\mathrm{e}$ 时，原级数发散.

例 1.14.5 判断级数 $\sum\limits_{n=1}^{\infty}\left(n\sin\frac{1}{n}\right)^{n^3}$ 的敛散性.

解 记 $u_n=\left(n\sin\frac{1}{n}\right)^{n^3}$，则 $\sqrt[n]{u_n}=\left(n\sin\frac{1}{n}\right)^{n^2}$，又

$$\lim_{n\to\infty}\sqrt[n]{u_n}=\lim_{n\to\infty}\left(n\sin\frac{1}{n}\right)^{n^2}=\exp\left\{\lim_{n\to\infty}n^2\left(n\sin\frac{1}{n}-1\right)\right\}$$

$$=\exp\left\{\lim_{n\to\infty}n^3\left(\sin\frac{1}{n}-\frac{1}{n}\right)\right\}=\exp\left\{\lim_{n\to\infty}\frac{\sin\frac{1}{n}-\frac{1}{n}}{\frac{1}{n^3}}\right\}=\mathrm{e}^{-\frac{1}{6}}<1,$$

由根值判别法可知，原级数收敛.

例 1.14.6 已知级数 $\sum\limits_{n=1}^{\infty}u_n$ 绝对收敛，判断 $\sum\limits_{n=1}^{\infty}\frac{u_n^2}{1+u_n^2}$ 的敛散性，其中 $u_n\neq0$.

解 由不等式 $a^2+b^2\geqslant2ab$，可知 $1+u_n^2\geqslant2|u_n|$，于是 $\frac{u_n^2}{1+u_n^2}\leqslant\frac{u_n^2}{2|u_n|}=\frac{|u_n|}{2}$，由题设知 $\sum\limits_{n=1}^{\infty}\frac{|u_n|}{2}$ 收敛，又 $\sum\limits_{n=1}^{\infty}\frac{u_n^2}{1+u_n^2}$ 是正项级数，根据正项级数的比较判别法知，原级数收敛.

二、莱布尼茨判别法

例 1.14.7 判断级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\cdot\frac{1}{n}$ 的敛散性.

解 级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\cdot\frac{1}{n}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+(-1)^{n-1}\frac{1}{n}+\cdots$，其中 $u_n=\frac{1}{n}$.

显然有 $\frac{1}{n}>\frac{1}{n+1}$，即 $u_n>u_{n+1}$，则 $\left\{\frac{1}{n}\right\}$ 单调不增，且有 $\lim\limits_{n\to\infty}u_n=\lim\limits_{n\to\infty}\frac{1}{n}=0$，故由莱布尼茨判别法知，该级数收敛.

【注】 级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\cdot\frac{1}{n}$ 叫作**交错调和级数**，这是一个经典且常用的例子，请读者知悉.

例 1.14.8 判断级数 $\sum\limits_{n=1}^{\infty}\sin(\pi\sqrt{n^2+a^2})$ 的敛散性，其中 a 为非零常数.

解 由三角公式 $\sin(\alpha+n\pi)=(-1)^n\sin\alpha$ 可知，

$$\sin(\pi\sqrt{n^2+a^2})=\sin(\pi\sqrt{n^2+a^2}-n\pi+n\pi)=(-1)^n\sin[(\sqrt{n^2+a^2}-n)\pi]=(-1)^n\sin\frac{\pi a^2}{\sqrt{n^2+a^2}+n},$$

故原级数是交错级数，当 $n\to\infty$ 时，$\frac{\pi a^2}{\sqrt{n^2+a^2}+n}$ 是正的无穷小量，且 $\left\{\frac{\pi a^2}{\sqrt{n^2+a^2}+n}\right\}$ 单调减少，所以

$\left\{\sin\dfrac{\pi a^2}{\sqrt{n^2+a^2}+n}\right\}$亦单调减少，且$\lim\limits_{n\to\infty}\sin\dfrac{\pi a^2}{\sqrt{n^2+a^2}+n}=0$，满足莱布尼茨判别法的条件，则原级数收敛.

例 1.14.9 再次考查级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\cdot\dfrac{1}{n}$，它是条件收敛还是绝对收敛？

解 由于$\left|(-1)^{n-1}\cdot\dfrac{1}{n}\right|=\dfrac{1}{n}$，$\sum\limits_{n=1}^{\infty}\dfrac{1}{n}$是发散的；而由例 1.14.7 的答案可知，$\sum\limits_{n=1}^{\infty}(-1)^{n-1}\cdot\dfrac{1}{n}$是收敛的. 故此级数条件收敛.

例 1.14.10 判别 $\sum\limits_{n=1}^{\infty}(-1)^n\dfrac{\ln(1+n)}{1+n}$ 的敛散性.

解 显然，$\lim\limits_{n\to\infty}\dfrac{\ln(1+n)}{1+n}=\lim\limits_{x\to+\infty}\dfrac{\ln(1+x)}{1+x}\overset{\text{“}\frac{\infty}{\infty}\text{”}}{\underset{\text{洛必达法则}}{=\!=\!=\!=\!=}}0$；

$$u_n=\frac{\ln(1+n)}{1+n}\xrightarrow{\text{连续化为}}f(x)=\frac{\ln(1+x)}{1+x}\xrightarrow{\text{当 }x>\mathrm{e}-1\text{ 时}}f'(x)=\frac{1-\ln(1+x)}{(1+x)^2}<0,$$

故当 $n\geqslant 2$ 时有 $u_n\geqslant u_{n+1}$，满足莱布尼茨判别法条件，原级数收敛.

三、任意项级数的判敛问题

例 1.14.11 如果数项级数 $\sum\limits_{n=1}^{\infty}u_n$ 收敛，则以下级数必收敛的是（　　）.

(A) $\sum\limits_{n=1}^{\infty}(-1)^n\dfrac{u_n}{n}$　(B) $\sum\limits_{n=1}^{\infty}u_n^2$　(C) $\sum\limits_{n=1}^{\infty}(u_{2n-1}-u_{2n})$　(D) $\sum\limits_{n=1}^{\infty}(u_n+u_{n+1})$

解 应选(D).

分析详见“注”中的“(5)”“(7)”“(10)”“(11)”.

【注】 很多同学在复习数项级数的判敛问题时感觉摸不着头脑，尤其是对于这种抽象的问题，更是害怕. 我想，做题时的“总结不足”可能是一个重要的原因——有些同学说，我已经做了很多题，可是为什么做新的练习题时还是不会呢？答：很可能是你做题的质量不高，为了做题而做题，做完一题扔一题，只追求数量，这肯定不行. 做完一题，就要停下来总结一下这道题你能学到什么技巧，容易错在哪里，有什么值得借鉴的；做完同一知识点下的几道题，就要停下来好好琢磨这几道题之间有没有什么联系，考虑能从哪些角度去解决这类问题. 这种总结做多了，你的做题质量和效率就会越来越高.

下面给大家做一个示范，做了很多题目以后，对于抽象的数项级数的判敛问题，有如下总结.

设 $\sum\limits_{n=1}^{\infty}u_n$，$\sum\limits_{n=1}^{\infty}v_n$，$\sum\limits_{n=1}^{\infty}w_n$ 是任意项级数（$\sum$ 的上标都是 ∞，而下标并非总是 1）.

(1) 设 a,b,c 为**非零**常数，且 $au_n+bv_n+cw_n=0$，则在 $\sum\limits_{n=1}^{\infty}u_n$，$\sum\limits_{n=1}^{\infty}v_n$ 和 $\sum\limits_{n=1}^{\infty}w_n$ 中只要有两个级数是收敛的，另一个必收敛.

(2) 设 $\sum\limits_{n=1}^{\infty}|u_n|$ 收敛，则 $\sum\limits_{n=1}^{\infty}u_n$ 收敛；设 $\sum\limits_{n=1}^{\infty}u_n$ 发散，则 $\sum\limits_{n=1}^{\infty}|u_n|$ 发散.

(3) 设 $\sum\limits_{n=1}^{\infty} u_n^2$ 收敛，则 $\sum\limits_{n=1}^{\infty} \frac{u_n}{n}$ 绝对收敛（$\left|\frac{u_n}{n}\right| \leqslant \frac{1}{2}\left(u_n^2+\frac{1}{n^2}\right)$）.

(4) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} |u_n|$ 不定（反例：$\sum\limits_{n=1}^{\infty}(-1)^n \frac{1}{n}$ 收敛，但 $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ 发散）.

(5) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} u_n^2$ 不定（反例：$\sum\limits_{n=1}^{\infty}(-1)^n \frac{1}{\sqrt{n}}$ 收敛，但 $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ 发散）.

(6) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}(-1)^n u_n$ 不定（反例：$\sum\limits_{n=1}^{\infty}(-1)^n \frac{1}{n}$ 收敛，但 $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ 发散）.

(7) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}(-1)^n \frac{u_n}{n}$ 不定（反例：$\sum\limits_{n=2}^{\infty}(-1)^n \frac{1}{\ln n}$ 收敛，但 $\sum\limits_{n=2}^{\infty} \frac{1}{n\ln n}$ 发散）.

(8) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} u_{2n}$（偶数项），$\sum\limits_{n=1}^{\infty} u_{2n-1}$（奇数项）不定（反例：$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots=\sum\limits_{n=1}^{\infty}(-1)^{n-1} \frac{1}{n}$ 收敛，但是其奇数项和偶数项都发散）.

(9) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}(u_{2n-1}+u_{2n})$ 收敛（收敛级数任意加括号所得的新级数仍收敛，且和不变）.

(10) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}(u_{2n-1}-u_{2n})$ 不定（反例：$u_1+u_2+u_3+u_4+u_5+u_6+\cdots=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots$ 收敛，但 $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\cdots$ 发散）.

(11) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\begin{cases} \sum\limits_{n=1}^{\infty}(u_n+u_{n+1}) \text{ 收敛}, \sum\limits_{n=1}^{\infty} u_n+\sum\limits_{n=1}^{\infty} u_{n+1} \text{ 收敛}, \\ \sum\limits_{n=1}^{\infty}(u_n-u_{n+1}) \text{ 收敛}, \sum\limits_{n=1}^{\infty} u_n-\sum\limits_{n=1}^{\infty} u_{n+1} \text{ 收敛}. \end{cases}$

(12) 设 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty} u_n u_{n+1}$ 不定（反例：$u_n=(-1)^n \frac{1}{\sqrt{n}}$，$u_n u_{n+1}=(-1)^n \frac{1}{\sqrt{n}}(-1)^{n+1} \frac{1}{\sqrt{n+1}}=-\frac{1}{\sqrt{n(n+1)}}$，级数发散）.

例 1.14.12 设有两个数列 $\{a_n\}$，$\{b_n\}$，若 $\lim\limits_{n \to \infty} a_n=0$，则（　）.

(A) 当 $\sum\limits_{n=1}^{\infty} b_n$ 收敛时，$\sum\limits_{n=1}^{\infty} a_n b_n$ 收敛

(B) 当 $\sum\limits_{n=1}^{\infty} b_n$ 发散时，$\sum\limits_{n=1}^{\infty} a_n b_n$ 发散

(C) 当 $\sum\limits_{n=1}^{\infty} |b_n|$ 收敛时，$\sum\limits_{n=1}^{\infty} a_n^2 b_n^2$ 收敛

(D) 当 $\sum\limits_{n=1}^{\infty} |b_n|$ 发散时，$\sum\limits_{n=1}^{\infty} a_n^2 b_n^2$ 发散

解 应选(C).

我们都知道，如果 $\sum\limits_{n=1}^{\infty} u_n$ 收敛，则 $\lim\limits_{n \to \infty} u_n=0$. 根据极限定义，$\forall \varepsilon>0$，存在 $N>0$，当 $n>N$ 时，必有 $|u_n-0|<\varepsilon$，取 $\varepsilon=1$，则 $|u_n|<1$. 你是否会问，为什么要取 ε 为 1 呢？请你通过下面题目的解答思考一下. 于是，我们可以将一个级数的收敛或者一般项趋向于 0 这样的条件，转化为当 n 充分大时，一般项 $|u_n|<1$，这是一个非常典型的技巧，于是有

由 $\sum_{n=1}^{\infty}|b_n|$ 收敛 $\Rightarrow \lim_{n\to\infty}|b_n|=0\Rightarrow$ 存在 $N_1>0$，当 $n>N_1$ 时，一定有 $|b_n|<1$；

由 $\lim_{n\to\infty}a_n=0\Rightarrow$ 存在 $N_2>0$，当 $n>N_2$ 时，一定有 $|a_n|<1$.

故当 $n>N_1+N_2$ 时，有 $0\leqslant a_n^2b_n^2<1\cdot b_n^2=|b_n|\cdot|b_n|<|b_n|\cdot 1=|b_n|$，由正项级数的比较判别法易知，当 $\sum_{n=1}^{\infty}|b_n|$ 收敛时，$\sum_{n=1}^{\infty}a_n^2b_n^2$ 收敛，正确答案是(C). 当然，本题举反例也可以排除错误选项，对于选项(A)，取 $a_n=b_n=(-1)^n\frac{1}{\sqrt{n}}$，对于选项(B)和(D)，取 $a_n=b_n=\frac{1}{n}$.

四、幂级数的收敛域

我们把它分成两种出题角度，第一种给出的研究对象是具体型的，第二种给出的研究对象是抽象型的，两种类型在考研中均是考查重点.

1. 具体型问题

(1)对于 $\sum u_n(x)$，加绝对值 $\Rightarrow\sum|u_n(x)|$，成为正项级数.

(2)用正项级数的 $\begin{cases}\text{比值}\\\text{根值}\end{cases}$ 判别法，令 $\begin{cases}\lim\limits_{n\to\infty}\dfrac{|u_{n+1}(x)|}{|u_n(x)|}<1,\\ \lim\limits_{n\to\infty}\sqrt[n]{|u_n(x)|}<1\end{cases}\Rightarrow x$ 的取值范围，即得收敛区间.

(3)单独讨论收敛区间的两个端点处级数的敛散性.

综合(2)，(3)即可求得幂级数的收敛域为 (a,b) 或 $(a,b]$ 或 $[a,b)$ 或 $[a,b]$.

例 1.14.13 求幂级数 $\sum_{n=1}^{\infty}\frac{x^n}{n}$ 的收敛域.

解 因为 $\lim\limits_{n\to\infty}\dfrac{\frac{|x^{n+1}|}{n+1}}{\frac{|x^n|}{n}}=\lim\limits_{n\to\infty}\left|\dfrac{nx}{n+1}\right|=|x|$，所以由比值判别法可知：

当 $|x|<1$ 时，幂级数 $\sum_{n=1}^{\infty}\frac{x^n}{n}$ 绝对收敛；当 $|x|>1$ 时，幂级数 $\sum_{n=1}^{\infty}\frac{x^n}{n}$ 发散.

又因为当 $x=1$ 时，幂级数 $\sum_{n=1}^{\infty}\frac{x^n}{n}=\sum_{n=1}^{\infty}\frac{1}{n}$ 发散，当 $x=-1$ 时，幂级数 $\sum_{n=1}^{\infty}\frac{x^n}{n}=\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$ 条件收敛，所以收敛域为 $[-1,1)$.

2. 抽象型问题

有如下几个结论请大家记住.

结论 1 根据阿贝尔定理，已知 $\sum_{n=0}^{\infty}a_n(x-x_0)^n$ 在某点 $x_1(x_1\neq x_0)$ 的敛散性，确定该幂级数的收敛半径可分为以下三种情况.

(1)若在 x_1 处收敛，则收敛半径 $R\geqslant|x_1-x_0|$.

(2)若在 x_1 处发散，则收敛半径 $R \leqslant |x_1 - x_0|$.

(3)若在 x_1 处条件收敛，则 $R = |x_1 - x_0|$.【重要考点】

结论 2 已知 $\sum a_n(x-x_1)^n$ 的敛散性信息，要求讨论 $\sum b_n(x-x_2)^m$ 的敛散性.

(1) $(x-x_1)^n$ 与 $(x-x_2)^m$ 的转化一般通过初等变形来完成，包括①“平移”收敛区间；②提出或者乘以因式 $(x-x_0)^k$ 等.

(2) a_n 与 b_n 的转化一般通过微积分变形来完成，包括①对级数逐项求导；②对级数逐项积分等.

(3)以下三种情况，级数的收敛半径不变，收敛域要具体问题具体分析.

①对级数提出或者乘以因式 $(x-x_0)^k$，或者作平移等，收敛半径不变.

②对级数逐项求导，收敛半径不变，收敛域可能缩小.

③对级数逐项积分，收敛半径不变，收敛域可能扩大.

这样说可能有些抽象，我们来看下面一个具体的例题，请大家结合这个例题来理解和掌握上述理论.

例 1.14.14 设 $\sum\limits_{n=1}^{\infty} a_n(x+1)^n$ 在点 $x=1$ 处条件收敛，则幂级数 $\sum\limits_{n=1}^{\infty} na_n(x-1)^n$ 在点 $x=2$ 处（　　）.

(A)绔对收敛　　(B)条件收敛　　(C)发散　　(D)敛散性不确定

解 应选(A).

根据结论 1 的(3)，由 $\sum\limits_{n=1}^{\infty} a_n(x+1)^n$ 在 $x=1$ 处条件收敛，则 $R = |x_1 - x_0| = |1-(-1)| = 2$，且收敛区间为 $(-3,1)$；

根据结论 2 的(1)和(3)，将 $(x+1)^n$ 转化为 $(x-1)^n$，也就是把级数的中心点由 -1 转移到 1，即将收敛区间平移到 $(-1,3)$，得 $\sum\limits_{n=1}^{\infty} a_n(x-1)^n$，收敛半径不变；

根据结论 2 的(1)，(2)和(3)，对 $\sum\limits_{n=1}^{\infty} a_n(x-1)^n$ 逐项求导，得 $\sum\limits_{n=1}^{\infty} na_n(x-1)^{n-1}$，再逐项乘以 $(x-1)$ 得 $\sum\limits_{n=1}^{\infty} na_n(x-1)^n$，收敛半径不变.

故 $\sum\limits_{n=1}^{\infty} na_n(x-1)^n$ 的收敛区间为 $(-1,3)$，$x=2$ 在收敛区间内部，故在该点处级数绝对收敛，答案选择(A).

五、求和函数

例 1.14.15 求级数 $\sum\limits_{n=1}^{\infty} \dfrac{x^n}{n}$ 的和函数.

解 设 $S(x) = \sum\limits_{n=1}^{\infty} \dfrac{x^n}{n}$，逐项求导，得 $S'(x) = \sum\limits_{n=1}^{\infty} \left(\dfrac{x^n}{n}\right)' = \sum\limits_{n=1}^{\infty} x^{n-1} = \dfrac{1}{1-x}$，$|x|<1$，然后两边积分，得 $S(x) - S(0) = \int_0^x \dfrac{\mathrm{d}t}{1-t} = -\ln(1-x)$，又因 $S(0)=0$，故和函数

$$S(x) = -\ln(1-x) + S(0) = -\ln(1-x), x \in [-1,1).$$

例 1.14.16 求级数 $\sum\limits_{n=1}^{\infty}nx^n$ 的和函数.

解 $\sum\limits_{n=1}^{\infty}nx^n=x\sum\limits_{n=1}^{\infty}nx^{n-1}$，设 $S(x)=\sum\limits_{n=1}^{\infty}nx^{n-1}$，则原级数 $\sum\limits_{n=1}^{\infty}nx^n=xS(x)$.

对 $S(x)=\sum\limits_{n=1}^{\infty}nx^{n-1}$ 两边积分，得 $\int_0^x S(t)\mathrm{d}t=\sum\limits_{n=1}^{\infty}\int_0^x nt^{n-1}\mathrm{d}t=\sum\limits_{n=1}^{\infty}x^n=\dfrac{x}{1-x}$，$|x|<1$.

再两边求导，得 $\left[\int_0^x S(t)\mathrm{d}t\right]'=\left(\dfrac{x}{1-x}\right)'=\dfrac{1}{(1-x)^2}$，所以 $\sum\limits_{n=1}^{\infty}nx^n=xS(x)=\dfrac{x}{(1-x)^2}$，$|x|<1$.

【注】 (1)幂级数求和函数的突破口：

①当 $(an+b)^c$ 在分母上时，先导后积，见例 1.14.15.

②当 $(an+b)^c$ 在分子上时，先积后导，见例 1.14.16.

(2)对于先导后积的处理办法：

这里有个细致的问题需要大家做扎实、做准确，请问 $\int S'(x)\mathrm{d}x$ 是否等于 $S(x)$ 呢？

事实上，由于 $\int_a^x S'(t)\mathrm{d}t=S(t)\Big|_a^x=S(x)-S(a)$，故 $S(x)=\int_a^x S'(t)\mathrm{d}t+S(a)$；

又由于当 a 为中心点时，$\begin{cases}\sum\limits_{n=0}^{\infty}a_nx^n \xlongequal{a=0} a_0, \\ \sum\limits_{n=0}^{\infty}a_n(x-x_0)^n \xlongequal{a=x_0} a_0\end{cases}$ 都收敛，故下限通常取中心点.

(3)始终不要忘记在解题过程中标注收敛域.

(4)建议大家记住由这两个例题所得到的结果：

$$\sum_{n=1}^{\infty}\frac{x^n}{n}=-\ln(1-x),-1\leqslant x<1;\quad \sum_{n=1}^{\infty}nx^{n-1}=\frac{1}{(1-x)^2},-1<x<1,$$

利用这两个公式可直接得到如下常数项级数的和：

$$\sum_{n=1}^{\infty}\frac{1}{n2^n}=\sum_{n=1}^{\infty}\frac{1}{n}\cdot\left(\frac{1}{2}\right)^n=-\ln\left(1-\frac{1}{2}\right)=\ln 2,$$

$$\sum_{n=1}^{\infty}\frac{n}{2^n}=\frac{1}{2}\sum_{n=1}^{\infty}n\left(\frac{1}{2}\right)^{n-1}=\frac{1}{2}\times\frac{1}{\left(1-\frac{1}{2}\right)^2}=2.$$

例 1.14.17 求幂级数 $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n-1}}{2n-1}x^{2n}$ 的收敛域及和函数.

解 先求收敛域.

记 $u_n(x)=\dfrac{(-1)^{n-1}}{2n-1}x^{2n}$，由于 $\lim\limits_{n\to\infty}\left|\dfrac{u_{n+1}(x)}{u_n(x)}\right|=\lim\limits_{n\to\infty}\dfrac{2n-1}{2n+1}x^2=x^2$，故

当 $|x|<1$ 时，$\sum\limits_{n=1}^{\infty}u_n(x)$ 收敛；

当 $x=\pm 1$ 时，$\sum\limits_{n=1}^{\infty}u_n(\pm 1)=\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n-1}}{2n-1}$，根据莱布尼茨判别法知此级数收敛. 所以，幂级数 $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n-1}}{2n-1}x^{2n}$ 的收敛域为 $[-1,1]$.

再求和函数.

设 $S(x)=\sum\limits_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}x^{2n-1}(-1\leqslant x\leqslant 1)$，由于 $S'(x)=\sum\limits_{n=1}^{\infty}(-1)^{n-1}x^{2n-2}=\frac{1}{1+x^2}$，且 $S(0)=0$，所以 $S(x)=\int_0^x\frac{\mathrm{d}t}{1+t^2}+S(0)=\arctan x$，从而

$$\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{2n-1}x^{2n}=xS(x)=x\arctan x(-1\leqslant x\leqslant 1).$$

六、函数展开成幂级数

例 1.14.18 将 $f(x)=\frac{1}{x^2-5x+6}$ 展开为 $x+5$ 的幂级数.

解 本题属于利用“直接恒等变形”的展开问题.

$$\begin{aligned}f(x)&=\frac{1}{(x-2)(x-3)}=\frac{1}{x-3}-\frac{1}{x-2}\\&=\frac{1}{x+5-8}-\frac{1}{x+5-7}=-\frac{1}{8}\cdot\frac{1}{1-\frac{x+5}{8}}+\frac{1}{7}\cdot\frac{1}{1-\frac{x+5}{7}}\\&=-\frac{1}{8}\cdot\sum_{n=0}^{\infty}\left(\frac{x+5}{8}\right)^n+\frac{1}{7}\cdot\sum_{n=0}^{\infty}\left(\frac{x+5}{7}\right)^n\\&=\sum_{n=0}^{\infty}\left(\frac{1}{7^{n+1}}-\frac{1}{8^{n+1}}\right)(x+5)^n(-12<x<2).\end{aligned}$$

例 1.14.19 求函数 $f(x)=\arctan x$ 在 $x=0$ 处的幂级数展开.

解 因为 $(\arctan x)'=\frac{1}{1+x^2}$，且

$$\frac{1}{1+x^2}=1-x^2+x^4-x^6+\cdots,|x|<1,$$

所以

$$\begin{aligned}\arctan x&=\int_0^x\frac{1}{1+t^2}\mathrm{d}t=\int_0^x(1-t^2+t^4-t^6+\cdots)\mathrm{d}t\\&=x-\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^7}{7}+\cdots=\sum_{n=0}^{\infty}(-1)^n\frac{x^{2n+1}}{2n+1},\ |x|\leqslant 1.\end{aligned}$$

【注】 取 $x=\frac{\sqrt{3}}{3}$ 代入上式，得到一个 π 的级数展开，即

$$\begin{aligned}\frac{\pi}{6}&=\frac{\sqrt{3}}{3}-\frac{(\sqrt{3})^3}{3^3\cdot 3}+\frac{(\sqrt{3})^5}{3^5\cdot 5}-\cdots=\frac{\sqrt{3}}{3}\left(1-\frac{1}{3\cdot 3}+\frac{1}{3^2\cdot 5}-\cdots\right)\\&=\frac{\sqrt{3}}{3}\sum_{n=0}^{\infty}(-1)^n\frac{1}{(2n+1)3^n};\end{aligned}$$

取 $x=1$，则有 $\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$.

这是求 π 的两个计算公式.

基础习题精练

习题

1.14.1 求极限$\lim\limits_{n\to\infty}\frac{2^n}{n!}$.

1.14.2 判别级数 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}$ 的敛散性.

1.14.3 设 $\sum\limits_{n=1}^{\infty}u_n$ 和 $\sum\limits_{n=1}^{\infty}v_n$ 都是正项级数，试证：

(1)若 $\sum\limits_{n=1}^{\infty}u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}\sqrt{u_nu_{n+1}}$ 收敛；

(2)若 $\sum\limits_{n=1}^{\infty}\sqrt{u_nu_{n+1}}$ 收敛，且 u_n 单调减少，则 $\sum\limits_{n=1}^{\infty}u_n$ 收敛；

(3)若 $\sum\limits_{n=1}^{\infty}v_n$ 和 $\sum\limits_{n=1}^{\infty}u_n$ 都收敛，则 $\sum\limits_{n=1}^{\infty}u_nv_n$ 收敛；

(4)若 $\sum\limits_{n=1}^{\infty}u_n$ 收敛，则 $\sum\limits_{n=1}^{\infty}\frac{u_n}{n}$ 收敛.

1.14.4 判别级数 $\sum\limits_{n=2}^{\infty}\ln\left[1+\frac{(-1)^n}{\sqrt{n}}\right]$ 的敛散性.

1.14.5 判别级数 $\sum\limits_{n=2}^{\infty}\frac{(-1)^n}{\sqrt{n}+(-1)^n}$ 的敛散性.

1.14.6 求极限$\lim\limits_{n\to\infty}\left(\frac{1}{2}+\frac{3}{2^2}+\frac{5}{2^3}+\cdots+\frac{2n-1}{2^n}\right)$.

1.14.7 求幂级数 $\sum\limits_{n=1}^{\infty}\frac{n}{3^n}(x-1)^n$ 的收敛半径和收敛域.

1.14.8 求级数 $\sum\limits_{n=0}^{\infty}\frac{(-1)^n(n^2-n+1)}{2^n}$ 的和.

1.14.9 设 $f(x)=\sum\limits_{n=0}^{\infty}\frac{a_n}{n!}x^n$，且满足$\begin{cases}f''(x)-f'(x)-2f(x)=0,\\ f(0)=0,f'(0)=1.\end{cases}$ 求：(1)$f(x)$；(2)a_n.

解答

1.14.1 **解** 记 $u_n=\frac{2^n}{n!}>0$，则所求极限$\lim\limits_{n\to\infty}\frac{2^n}{n!}$可视为正项级数 $\sum\limits_{n=1}^{\infty}u_n$ 的一般项的极限. 如果该级数收敛，则由级数收敛的必要条件就可求其极限. 因

$$\lim_{n\to\infty}\frac{u_{n+1}}{u_n}=\lim_{n\to\infty}\frac{\frac{2^{n+1}}{(n+1)!}}{\frac{2^n}{n!}}=\lim_{n\to\infty}\frac{2}{n+1}=0<1,$$

由正项级数的比值判别法知 $\sum\limits_{n=1}^{\infty}\frac{2^n}{n!}$ 收敛，于是$\lim\limits_{n\to\infty}\frac{2^n}{n!}=0$.

【注】 该题与习题1.14.6是关于级数理论的极限问题.

1.14.2 解 首先,我们要确认一般项 $u_n=\frac{1}{n!}>0$,也就是要确认所研究的级数 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}$ 为正项级数.这一点看上去很简单,可在后面引入形形色色的级数类型后,判别类型就显得十分重要了.希望读者打好基础,养成先判断级数类型的习惯.接下来可去掉级数 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}$ 的第一项,也就是当 $n=0$ 时的那一项,即 $\frac{1}{0!}=1$,由前面讲的常数项级数的性质2,这不会改变该级数的敛散性,于是我们来研究 $\sum\limits_{n=1}^{\infty}\frac{1}{n!}$ 的敛散性.其部分和 S_n 满足

$$\begin{aligned}0<S_n&=1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}\\&\leqslant 1+\frac{1}{1\times2}+\frac{1}{2\times3}+\cdots+\frac{1}{(n-1)n}\\&=1+\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{n-1}-\frac{1}{n}\right)=2-\frac{1}{n}<2,\end{aligned}$$

即 $\{S_n\}$ 有界,由收敛原则,有 $\sum\limits_{n=1}^{\infty}\frac{1}{n!}$ 收敛,也即 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}$ 收敛.

【注】 这个级数的和是著名的欧拉数e.可在 $e^x=\sum\limits_{n=0}^{\infty}\frac{x^n}{n!}$ 中,令 $x=1$,则有 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}=e$.

1.14.3 证明 (1) $\sum\limits_{n=1}^{\infty}u_n$ 收敛,且有 $\sqrt{u_nu_{n+1}}\leqslant\frac{1}{2}(u_n+u_{n+1})\Rightarrow\sum\limits_{n=1}^{\infty}\sqrt{u_nu_{n+1}}$ 收敛.

(2) u_n 单调减少 $\Rightarrow u_{n+1}\leqslant u_n\Rightarrow u_{n+1}^2\leqslant u_nu_{n+1}\Rightarrow u_{n+1}\leqslant\sqrt{u_nu_{n+1}}\Rightarrow\sum\limits_{n=1}^{\infty}u_n$ 收敛.

(3) $\sum\limits_{n=1}^{\infty}u_n$ 和 $\sum\limits_{n=1}^{\infty}v_n$ 收敛 $\Rightarrow\sum\limits_{n=1}^{\infty}u_n^2$ 和 $\sum\limits_{n=1}^{\infty}v_n^2$ 收敛,且 $u_nv_n\leqslant\frac{1}{2}(u_n^2+v_n^2)\Rightarrow\sum\limits_{n=1}^{\infty}u_nv_n$ 收敛.

(4) $\sum\limits_{n=1}^{\infty}u_n$ 收敛 $\Rightarrow\sum\limits_{n=1}^{\infty}u_n^2$ 收敛,且 $\frac{u_n}{n}\leqslant\frac{1}{2}\left(u_n^2+\frac{1}{n^2}\right)\Rightarrow\sum\limits_{n=1}^{\infty}\frac{u_n}{n}$ 收敛.

1.14.4 分析 这是交错级数,但不易判别 $\{|u_n|\}$ 的单调性,因此不能使用莱布尼茨判别法.为了能掌握一般项 $u_n=\ln\left[1+\frac{(-1)^n}{\sqrt{n}}\right]$ 的级别,需使用泰勒公式.

解 由泰勒公式,
$$\ln\left[1+\frac{(-1)^n}{\sqrt{n}}\right]=\frac{(-1)^n}{\sqrt{n}}-\frac{1}{2n}+o\left(\frac{1}{n}\right),$$

由比较判别法,$\lim\limits_{n\to\infty}\dfrac{\frac{1}{2n}-o\left(\frac{1}{n}\right)}{\frac{1}{n}}=\frac{1}{2}$,而级数 $\sum\limits_{n=2}^{\infty}\frac{(-1)^n}{\sqrt{n}}$ 是(条件)收敛的.故原级数发散.

1.14.5 解 事实上,其一般项为

$$\frac{(-1)^n}{\sqrt{n}+(-1)^n}=\frac{(-1)^n[\sqrt{n}-(-1)^n]}{n-1}=\frac{(-1)^n\sqrt{n}-1}{n-1}=\frac{(-1)^n}{\sqrt{n}-\frac{1}{\sqrt{n}}}-\frac{1}{n-1},$$

易知交错级数 $\sum\limits_{n=2}^{\infty}\dfrac{(-1)^n}{\sqrt{n}-\frac{1}{\sqrt{n}}}$ 是收敛的,而级数 $\sum\limits_{n=2}^{\infty}\frac{1}{n-1}$ 发散,故原级数发散.

1.14.6 解 $\lim\limits_{n\to\infty}\left(\frac{1}{2}+\frac{3}{2^2}+\frac{5}{2^3}+\cdots+\frac{2n-1}{2^n}\right)=\lim\limits_{n\to\infty}\sum\limits_{i=1}^{n}\frac{2i-1}{2^i}=\sum\limits_{n=1}^{\infty}\frac{2n-1}{2^n}.$

令 $S(x)=\sum\limits_{n=1}^{\infty}\frac{2n-1}{2^n}x^{2n-2}(-\sqrt{2}<x<\sqrt{2})$，用先积分后求导的方法，则

$$S(x)=\left(\sum_{n=1}^{\infty}\int_0^x\frac{2n-1}{2^n}t^{2n-2}\mathrm{d}t\right)'=\left[\frac{1}{x}\sum_{n=1}^{\infty}\left(\frac{x^2}{2}\right)^n\right]'=\left(\frac{1}{x}\cdot\frac{\frac{x^2}{2}}{1-\frac{x^2}{2}}\right)'=\left(\frac{x}{2-x^2}\right)'=\frac{2+x^2}{(2-x^2)^2},$$

故原极限 $=S(1)=3.$

1.14.7 解 用比值判别法，因为

$$\lim_{n\to\infty}\left|\frac{\frac{(n+1)(x-1)^{n+1}}{3^{n+1}}}{\frac{n(x-1)^n}{3^n}}\right|=\lim_{n\to\infty}\frac{n+1}{3n}|x-1|=\frac{1}{3}|x-1|,$$

所以，当 $\frac{1}{3}\mid x-1\mid<1$，即 $\mid x-1\mid<3$ 时，幂级数 $\sum\limits_{n=1}^{\infty}\frac{n}{3^n}(x-1)^n$ 绝对收敛；

当 $\frac{1}{3}\mid x-1\mid>1$，即 $\mid x-1\mid>3$ 时，幂级数 $\sum\limits_{n=1}^{\infty}\frac{n}{3^n}(x-1)^n$ 发散.

根据收敛半径的定义可知收敛半径 $R=3$，收敛区间为 $(-2,4)$.

因为当 $x=4$ 时，幂级数成为 $\sum\limits_{n=1}^{\infty}n$，发散；当 $x=-2$ 时，幂级数成为 $\sum\limits_{n=1}^{\infty}(-1)^n n$，发散，所以收敛域为 $(-2,4)$.

1.14.8 解 $\sum\limits_{n=0}^{\infty}\frac{(-1)^n(n^2-n+1)}{2^n}=\sum\limits_{n=0}^{\infty}n(n-1)\left(-\frac{1}{2}\right)^n+\sum\limits_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n,$

$$\sum_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n=\frac{1}{1+\frac{1}{2}}=\frac{2}{3}.$$

设 $S(x)=\sum\limits_{n=2}^{\infty}n(n-1)x^{n-2}=\left(\sum\limits_{n=0}^{\infty}x^n\right)''=\frac{2}{(1-x)^3}$，$\mid x\mid<1$，则

$$\sum_{n=2}^{\infty}n(n-1)x^n=\frac{2x^2}{(1-x)^3},\quad\sum_{n=2}^{\infty}n(n-1)\left(-\frac{1}{2}\right)^n=\frac{4}{27},$$

所以

$$\sum_{n=0}^{\infty}\frac{(-1)^n(n^2-n+1)}{2^n}=\frac{4}{27}+\frac{2}{3}=\frac{22}{27}.$$

1.14.9 解 (1)特征方程为 $r^2-r-2=0$，得 $r_1=-1,r_2=2$，故通解为

$$f(x)=C_1\mathrm{e}^{-x}+C_2\mathrm{e}^{2x}.$$

由 $f(0)=0,f'(0)=1$，得

$$\begin{cases}C_1+C_2=0,\\-C_1+2C_2=1,\end{cases}$$

故 $C_1=-\frac{1}{3},C_2=\frac{1}{3}$，所以

$$f(x)=-\frac{1}{3}\mathrm{e}^{-x}+\frac{1}{3}\mathrm{e}^{2x}.$$

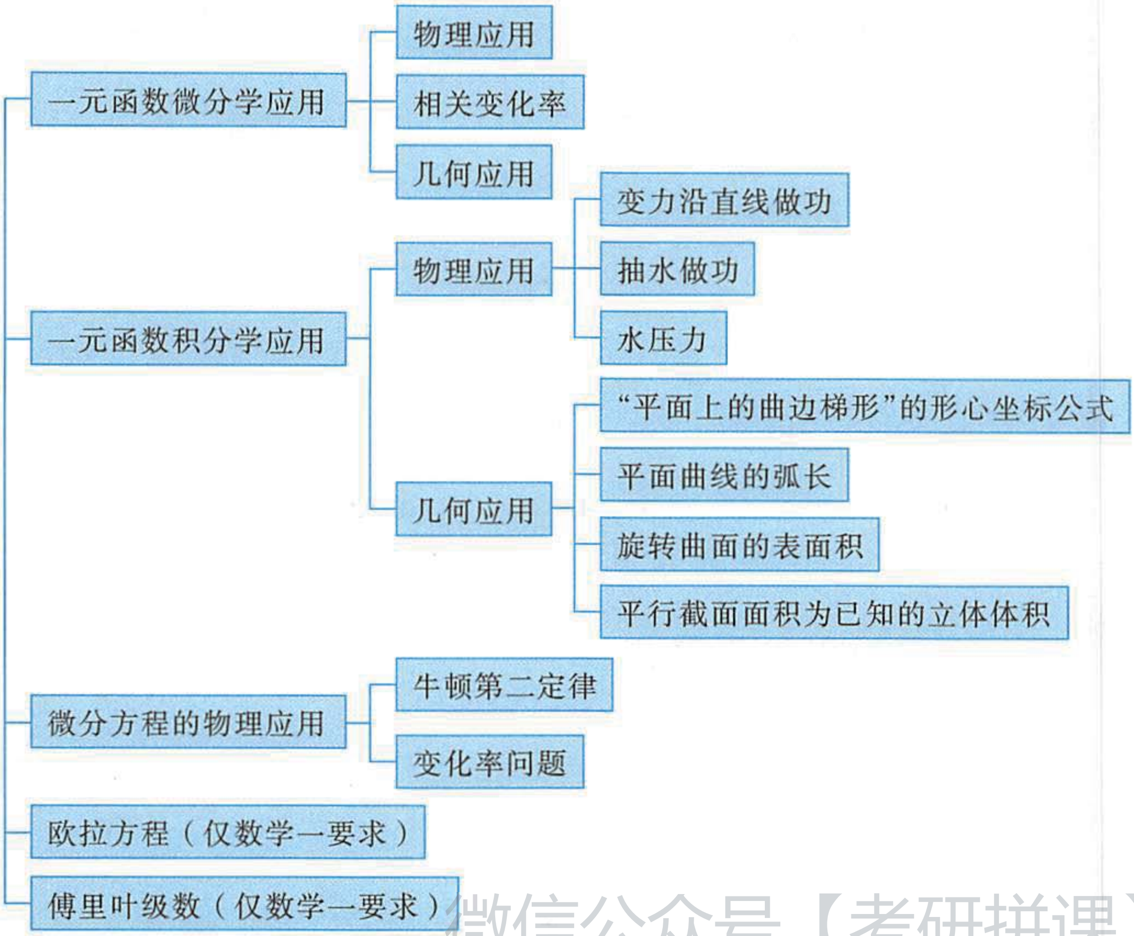
(2)
$$f(x)=\sum_{n=0}^{\infty}\frac{a_n}{n!}x^n=-\frac{1}{3}\mathrm{e}^{-x}+\frac{1}{3}\mathrm{e}^{2x}$$

$$
\begin{aligned}
&=-\frac{1}{3}\sum_{n=0}^{\infty}\frac{(-x)^n}{n!}+\frac{1}{3}\sum_{n=0}^{\infty}\frac{(2x)^n}{n!}\\
&=-\frac{1}{3}\sum_{n=0}^{\infty}\frac{(-1)^n x^n}{n!}+\frac{1}{3}\sum_{n=0}^{\infty}\frac{2^n\cdot x^n}{n!}\\
&=\sum_{n=0}^{\infty}\left[\left(-\frac{1}{3}\right)(-1)^n+\frac{2^n}{3}\right]\frac{x^n}{n!},
\end{aligned}
$$

故 $a_n=\frac{1}{3}[(-1)^{n+1}+2^n]$.

第15讲 数学一、数学二专题内容

基础知识结构



基础内容精讲

一、一元函数微分学应用

1. 物理应用

已知质点运动的位移 s 关于时间 t 的函数为 $s=s(t)$，称它为质点的运动方程(位移方程)，则其速度为

$$v=\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t}=s'(t), \qquad (1\text{-}15\text{-}1)$$

其加速度为

$$a(t)=\lim_{\Delta t\to 0}\frac{\Delta v}{\Delta t}=v'(t)=s''(t). \tag{1-15-2}$$

这就是导数的物理意义.

2. 相关变化率

若 $y=f(x)$，$\begin{cases}x=x(t),\\ y=y(t).\end{cases}$ 这些函数均可导，则 $\frac{\mathrm{d}y}{\mathrm{d}t}=\frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}=f'(x)\frac{\mathrm{d}x}{\mathrm{d}t}$，上式中，$\frac{\mathrm{d}y}{\mathrm{d}t}$ 与 $\frac{\mathrm{d}x}{\mathrm{d}t}$ 由 $f'(x)$ 联系在一起，称这种相互关联的变化率为**相关变化率**.

3. 几何应用

设 $y(x)$ 二阶可导，则曲线 $y=y(x)$ 在其上点 $(x,y(x))$ 处的**曲率公式**为

$$k=\frac{|y''|}{[1+(y')^2]^{\frac{3}{2}}}. \tag{1-15-3}$$

曲率半径的计算公式

$$R=\frac{1}{k}=\frac{[1+(y')^2]^{\frac{3}{2}}}{|y''|}\ (y''\neq 0). \tag{1-15-4}$$

二、一元函数积分学应用

1. 物理应用

(1)变力沿直线做功.

设方向沿 x 轴正向的力函数为 $F(x)(a\leqslant x\leqslant b)$，则物体沿 x 轴从点 a 移动到点 b 时(如图 1-15-1)，变力 $F(x)$ 所做的功为

$$W=\int_a^b F(x)\mathrm{d}x,$$

功的元素 $\mathrm{d}W=F(x)\mathrm{d}x$.

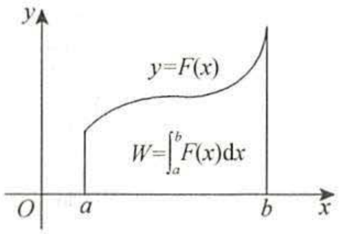


图 1-15-1

【注】 力关于路程的定积分就是功.

(2)抽水做功.

如图 1-15-2 所示，将容器中的水全部抽出所做的功为

$$W=\rho g\int_a^b xA(x)\mathrm{d}x,$$

其中 ρ 为水的密度，g 为重力加速度.

功的元素 $\mathrm{d}W=\rho gxA(x)\mathrm{d}x$ 为位于 x 处厚度为 $\mathrm{d}x$，水平截面面积为 $A(x)$ 的一层水被抽出(路程为 x)所做的功.

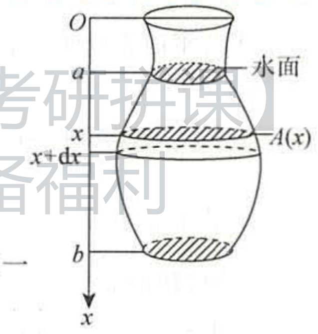


图 1-15-2

【注】 抽水做功的特点：力(重力)不变，路程在变. 求解这类问题的关键是确定 x 处的水平截面面积 $A(x)$，其余的量都是固定的.

(3)水压力.

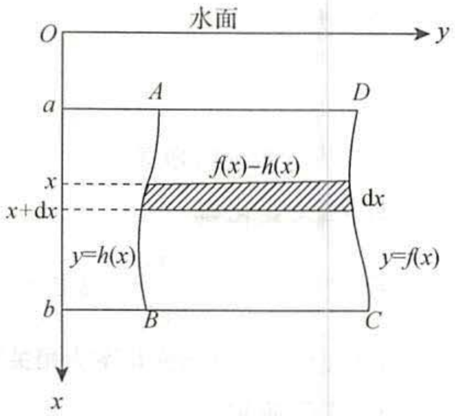


图 1-15-3

垂直浸没在水中的平板 $ABCD$(如图 1-15-3)的一侧受到的水压力为

$$P=\rho g\int_a^b x[f(x)-h(x)]\mathrm{d}x,$$

其中 ρ 为水的密度，g 为重力加速度.

压力元素 $\mathrm{d}P=\rho gx[f(x)-h(x)]\mathrm{d}x$ 是图中矩形条所受到的压力. x 表示水深，$f(x)-h(x)$ 是矩形条的宽度，$\mathrm{d}x$ 是矩形条的高度.

【注】 水压力问题的特点：压强随水的深度的改变而改变. 求解这类问题的关键是确定水深 x 处的平板的宽度 $f(x)-h(x)$.

2. 几何应用

(1)“平面上的曲边梯形”的形心坐标公式.

设平面区域 $D=\{(x,y)\mid 0\leqslant y\leqslant f(x),a\leqslant x\leqslant b\}$，$f(x)$ 在 $[a,b]$ 上连续，如图 1-15-4 所示. 现推导 D 的形心坐标 $\bar{x}$，$\bar{y}$ 的计算公式.

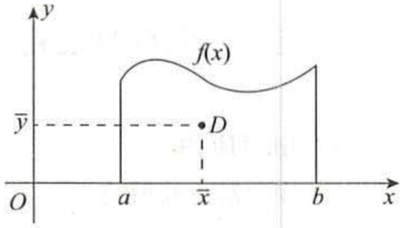


图 1-15-4

$$\bar{x}=\frac{\iint\limits_D x\,\mathrm{d}\sigma}{\iint\limits_D \mathrm{d}\sigma}=\frac{\int_a^b\mathrm{d}x\int_0^{f(x)}x\,\mathrm{d}y}{\int_a^b\mathrm{d}x\int_0^{f(x)}\mathrm{d}y}=\frac{\int_a^b xf(x)\mathrm{d}x}{\int_a^b f(x)\mathrm{d}x};$$

$$\bar{y}=\frac{\iint\limits_D y\,\mathrm{d}\sigma}{\iint\limits_D \mathrm{d}\sigma}=\frac{\int_a^b\mathrm{d}x\int_0^{f(x)}y\,\mathrm{d}y}{\int_a^b\mathrm{d}x\int_0^{f(x)}\mathrm{d}y}=\frac{\frac{1}{2}\int_a^b f^2(x)\mathrm{d}x}{\int_a^b f(x)\mathrm{d}x}.$$

【注】 在考研题中出现过此类考题，见例 1.15.6.

(2)平面曲线的弧长.

①若平面光滑曲线由直角坐标方程 $y=y(x)(a\leqslant x\leqslant b)$ 给出，则 $s=\int_a^b\sqrt{1+[y'(x)]^2}\mathrm{d}x$.

②若平面光滑曲线由参数方程 $\begin{cases}x=x(t),\\y=y(t)\end{cases}(\alpha\leqslant t\leqslant\beta)$ 给出，则 $s=\int_\alpha^\beta\sqrt{[x'(t)]^2+[y'(t)]^2}\mathrm{d}t$.

③若平面光滑曲线由极坐标方程 $r=r(\theta)(\alpha\leqslant\theta\leqslant\beta)$ 给出，则 $s=\int_\alpha^\beta\sqrt{[r(\theta)]^2+[r'(\theta)]^2}\mathrm{d}\theta$.

(3)旋转曲面的表面积.

①曲线 $y=y(x)$ 在区间 $[a,b]$ 上的曲线弧段绕 x 轴旋转一周所得到的旋转曲面的表面积

$$S=2\pi\int_a^b|y(x)|\sqrt{1+[y'(x)]^2}\mathrm{d}x.$$

②曲线 $x=x(t)$，$y=y(t)(\alpha\leqslant t\leqslant\beta,x'(t)\neq0)$ 在区间 $[\alpha,\beta]$ 上的曲线弧段绕 x 轴旋转一周所得到的旋转曲面的表面积

$$S=2\pi\int_\alpha^\beta|y(t)|\sqrt{[x'(t)]^2+[y'(t)]^2}\mathrm{d}t.$$

(4)平行截面面积为已知的立体体积.

在区间$[a,b]$上，垂直于x轴的平面截立体Ω所得到的截面面积为x的连续函数$A(x)$，则Ω的体积为

$$V=\int_a^b A(x)\mathrm{d}x.$$

注例 1 设一个底面半径为 3 的圆柱体，被一个与圆柱的底面相交、夹角为$\frac{\pi}{4}$且过底面直径AB的平面所截，求截下的楔形体的体积.

解 建立如图 1-15-5 所示的坐标系，垂直于x轴的截面是直角三角形，由题设条件，这个直角三角形的底边长为$\sqrt{3^2-x^2}$，对边长为$\sqrt{3^2-x^2}\tan\frac{\pi}{4}=\sqrt{3^2-x^2}$.

故截面面积$S=\frac{1}{2}(3^2-x^2)$，则$V=\int_{-3}^{3}\frac{1}{2}(3^2-x^2)\mathrm{d}x=18.$

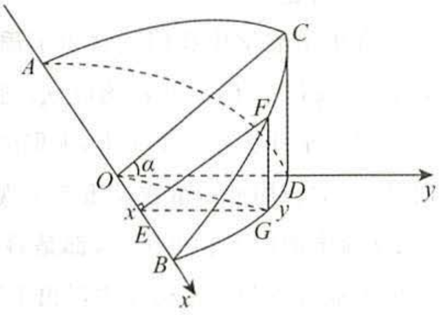


图 1-15-5

三、微分方程的物理应用

主要涉及以下两个方面的问题.

(1)牛顿第二定律.

相关物理量有①物体质量m；②力(包括重力、阻力、浮力)f；③加速度$a=\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\mathrm{d}v}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}=v\frac{\mathrm{d}v}{\mathrm{d}x}$. 见例 1. 15. 11.

(2)变化率问题.

提法多为“t时刻某量y对t的变化率与t时刻某量成正比”，如

①考题 1(事实上此题叫冷却定律).

“t时刻物体温度$T(t)$对时间的变化率与t时刻物体和介质的温差$T-T_0$成正比”，应写成$\frac{\mathrm{d}T}{\mathrm{d}t}=-k(x-x_0)$，负号代表$\frac{\mathrm{d}T}{\mathrm{d}t}<0$(温度随着时间的增加而降低). 见例 1. 15. 12.

②考题 2.

题设总人数N，“t时刻已掌握新技术的人数x的变化率和已掌握新技术与未掌握新技术的人数之积成正比”，应写成$\frac{\mathrm{d}x}{\mathrm{d}t}=+kx(N-x)$，正号代表$\frac{\mathrm{d}x}{\mathrm{d}t}>0$(人数随着时间增加而增多). 见习题 1. 15. 3.

四、欧拉方程(仅数学一要求)

形如$x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+px\frac{\mathrm{d}y}{\mathrm{d}x}+qy=f(x)$的方程称为**欧拉方程**，其中$p$与$q$为常数，$f(x)$为已知函数. 它有固定的解法.

(1)当$x>0$时，令$x=\mathrm{e}^t$，则$t=\ln x$，$\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{x}$，于是

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}t}\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}t},\quad \frac{\mathrm{d}^2y}{\mathrm{d}x^2}=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}t}\right)=-\frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}t}+\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)=-\frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}t}+\frac{1}{x^2}\frac{\mathrm{d}^2y}{\mathrm{d}t^2},$$

方程化为

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+(p-1)\frac{\mathrm{d}y}{\mathrm{d}t}+qy=f(\mathrm{e}^t),$$

即可求解(别忘了用 $t=\ln x$ 回代成 x 的函数).

(2)当 $x<0$ 时,令 $x=-e^t$,同理可得.

五、傅里叶级数(仅数学一要求)

1. 引言

傅里叶

(1768—1830)

在中学数学中我们就认识了**周期函数**:设常数 $T>0$,若对于定义域内的任意 $x,x+T,\varphi(x+T)=\varphi(x)$ 均成立,则称 $\varphi(x)$ 是以 T 为周期的周期函数.这个概念的提出是为了什么呢?事实上,人们的生活中确实常常见到**周期现象**,即每经过一段时间 T(称为周期),事物又重新出现其原先状态的现象.如汽车发动机中的活塞运动、交流电的电流和电压等,都是科学技术中常见的周期现象,每位读者都可从自己的专业背景与生活感受中提出不同的周期现象.在周期现象中的各种数量,不论 x 是从什么时刻开始,经历时间 T 后又重新回到其原先的取值,可用 $\varphi(x+T)=\varphi(x)$ 来表示,这就是周期函数的实际意义.读者若将此处的 x 当作时间 t,甚至直接写成 t,那是更好的.

除去常数函数这种特殊的略显无趣的周期函数,最简单的周期函数是**正弦型函数**:$A\sin(wt+\varphi)$,其中 w 叫角频率,A 叫振幅,φ 叫初相角,且 $w=\frac{2\pi}{T}$.我们给出这种函数的如下序列:

$$A_0,\quad A_1\sin(wt+\varphi_1),\quad A_2\sin(2wt+\varphi_2),\quad \cdots,\quad A_n\sin(nwt+\varphi_n),\quad \cdots,$$

一个著名的问题出现了:给出一个周期为 T 的函数 $\varphi(t)$,它是否可由上面的序列的和表示出来?即

$$\begin{aligned}\varphi(t)&\overset{?}{=\!=}A_0+A_1\sin(wt+\varphi_1)+A_2\sin(2wt+\varphi_2)+\cdots+A_n\sin(nwt+\varphi_n)+\cdots\\&=A_0+\sum_{n=1}^{\infty}A_n\sin(nwt+\varphi_n).\end{aligned}$$

一批杰出的数学家(你至少应该记住傅里叶、欧拉、达朗贝尔和狄利克雷这几个名字,我说的是"至少应该")为此做了一项艰辛且极具价值的工作,他们最终得出的结论是在极为广泛的条件(这条件要留在后面详细讲出来)下,答案是肯定的——**周期函数可由一个序列的正弦型函数叠加得到**——**从几何上讲,周期函数的图像可由一个序列的正弦型曲线合成**,这是一个多么美妙的结果.后面我们就称 $A_0+\sum_{n=1}^{\infty}A_n\sin(nwt+\varphi_n)$ 为**三角级数**.

1744 年,欧拉写信给他的好朋友哥德巴赫,信中提到了一些重要信息,下面我将尽量用容易被你所理解的语言来描述出这些重要信息.

设 $\varphi(x)$ 是以 2π 为周期的周期函数,且其在 $-\pi<x<\pi$ 上的函数表达式为 $\varphi(x)=x$,在 $x=\pm\pi$ 时,$\varphi(x)=0$.现在你可与我一同画出其图像(如图 1-15-6).

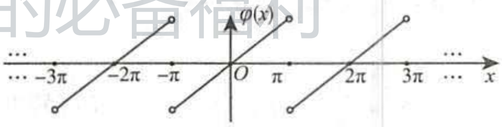


图 1-15-6

这样的一个周期函数是否像前面所言,可由一个序列的正弦型函数叠加得到呢?答案是肯定且精彩的.

事实上,

$$2\left(\sin x-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\frac{\sin 4x}{4}+\cdots\right)=2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}$$

的图像竟然就和图 1-15-6 的 $\varphi(x)$ 重合了！

如图1-15-7所示，n仅仅取到5时，也就是$\varphi_5(x)=2\left(\sin x-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\frac{\sin 4x}{4}+\frac{\sin 5x}{5}\right)$的图像：

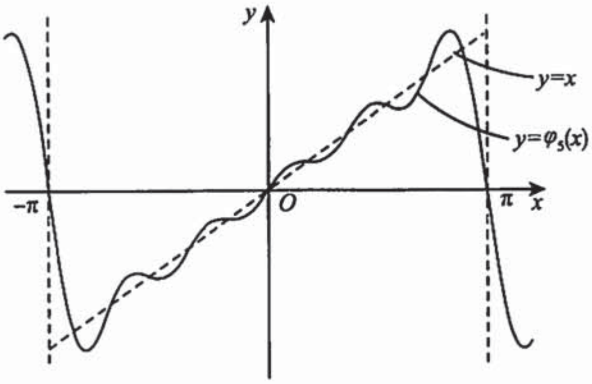


图 1-15-7

你看这 $\varphi_5(x)$ 在$(-\pi,\pi)$ 内不是已经有了“趋向”于 $y=x$ 的势头了吗？

看来，此问题的研究是相当有价值的. 接下来，我们就让欧拉和傅里叶这两位大师一起登场.

2. 欧拉-傅里叶方法与傅里叶级数的产生

如前面所述，若能写出 $\varphi(t)=A_0+\sum\limits_{n=1}^{\infty}A_n\sin(nwt+\varphi_n)$（仍然需要再强调一下，可以写出这种式子的条件在后面会讲到，读者先假设可以这样写，这样思路就不会混乱），用三角公式

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$$

对此式进行处理，得

$$\varphi(t)=A_0+\sum_{n=1}^{\infty}(A_n\sin nwt\cos\varphi_n+A_n\cos nwt\sin\varphi_n),$$

记 $A_n\sin\varphi_n=a_n$，$A_n\cos\varphi_n=b_n$，则 $\varphi(t)=A_0+\sum\limits_{n=1}^{\infty}(a_n\cos nwt+b_n\sin nwt)$.

若再记 $x=wt=\frac{2\pi t}{T}$，则有 $\varphi(t)=\varphi\left(\frac{x}{w}\right)\overset{记}{=\!=}f(x)$，且

$$f(x)=A_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx). \tag{1-15-5}$$

显然这个函数的周期为2π. 下面的问题是，系数 A_0，a_n，b_n 是多少？欧拉和傅里叶给出了求这些系数的方法，称为欧拉-傅里叶方法.

为了更好地讲清楚欧拉-傅里叶方法，首先需要给读者温习一段基础知识，叫三角函数族的正交性. 称$\{1,\cos x,\sin x,\cos 2x,\sin 2x,\cdots,\cos nx,\sin nx,\cdots\}$为三角函数族. 显然有

$$\int_{-\pi}^{\pi}1\cdot\cos nx\,\mathrm{d}x=\left.\frac{\sin nx}{n}\right|_{-\pi}^{\pi}=0,$$

$$\int_{-\pi}^{\pi}1\cdot\sin nx\,\mathrm{d}x=-\left.\frac{\cos nx}{n}\right|_{-\pi}^{\pi}=0,$$

$$\int_{-\pi}^{\pi}\cos nx\cdot\sin mx\,\mathrm{d}x=\frac{1}{2}\int_{-\pi}^{\pi}[\sin(m+n)x+\sin(m-n)x]\mathrm{d}x=0,$$

$$\int_{-\pi}^{\pi}\cos nx\cdot\cos mx\,\mathrm{d}x=\frac{1}{2}\int_{-\pi}^{\pi}[\cos(n+m)x+\cos(n-m)x]\mathrm{d}x=0, m\neq n,$$

$$\int_{-\pi}^{\pi}\sin nx\cdot\sin mx\,\mathrm{d}x=\frac{1}{2}\int_{-\pi}^{\pi}[\cos(n-m)x-\cos(m+n)x]\mathrm{d}x=0, m\neq n.$$

我们发现，这个函数族中任意两个不同函数的乘积在$[-\pi,\pi]$上的积分值为0.这性质被称为三角函数族在$[-\pi,\pi]$上的**正交性**.

顺便，还有

$$\int_{-\pi}^{\pi}\cos^2 nx\,\mathrm{d}x=\int_{-\pi}^{\pi}\frac{1+\cos 2nx}{2}\mathrm{d}x=\pi,\quad \int_{-\pi}^{\pi}\sin^2 nx\,\mathrm{d}x=\int_{-\pi}^{\pi}\frac{1-\cos 2nx}{2}\mathrm{d}x=\pi.$$

好了，有了上面这些准备知识，我们可以着手来求A_0,a_n,b_n了.

假设(1-15-5)式成立，即$f(x)=A_0+\sum\limits_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$，将此式在$[-\pi,\pi]$上逐项积分，有

$$\int_{-\pi}^{\pi}f(x)\mathrm{d}x=2\pi A_0+\sum_{n=1}^{\infty}\left(a_n\int_{-\pi}^{\pi}\cos nx\,\mathrm{d}x+b_n\int_{-\pi}^{\pi}\sin nx\,\mathrm{d}x\right),$$

由三角函数族的正交性，$\int_{-\pi}^{\pi}\cos nx\,\mathrm{d}x=\int_{-\pi}^{\pi}\sin nx\,\mathrm{d}x=0$，从而有

$$A_0=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\mathrm{d}x. \tag{1-15-6}$$

再将(1-15-5)式两边同时乘以$\cos mx$，在$[-\pi,\pi]$上逐项积分，有

$$\int_{-\pi}^{\pi}f(x)\cos mx\,\mathrm{d}x=A_0\int_{-\pi}^{\pi}\cos mx\,\mathrm{d}x+\sum_{n=1}^{\infty}\left(a_n\int_{-\pi}^{\pi}\cos nx\cos mx\,\mathrm{d}x+b_n\int_{-\pi}^{\pi}\sin nx\cos mx\,\mathrm{d}x\right),$$

由三角函数族的正交性，$\int_{-\pi}^{\pi}\cos mx\,\mathrm{d}x=0$，$\int_{-\pi}^{\pi}\cos nx\cos mx\,\mathrm{d}x=0$(其中$m\neq n$)，$\int_{-\pi}^{\pi}\sin nx\cos mx\,\mathrm{d}x=0$，

但当$m=n$时，有

$$\int_{-\pi}^{\pi}\cos^2 mx\,\mathrm{d}x=\pi,$$

故$\int_{-\pi}^{\pi}f(x)\cos mx\,\mathrm{d}x=a_m\int_{-\pi}^{\pi}\cos^2 mx\,\mathrm{d}x=a_m\cdot\pi$，于是$a_m=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos mx\,\mathrm{d}x$，即

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x. \tag{1-15-7}$$

同理，将(1-15-5)式两边同时乘以$\sin mx$，在$[-\pi,\pi]$上逐项积分，便可得出

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x. \tag{1-15-8}$$

现在把(1-15-6)，(1-15-7)，(1-15-8)式放在一起来看.

$$A_0=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\mathrm{d}x,$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x.$$

你会发现，若在(1-15-7)式中取$n=0$，则$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\mathrm{d}x=2A_0$，今后为了写法上的统一，我们干脆记$A_0=\frac{a_0}{2}$，于是便有$f(x)=\frac{a_0}{2}+\sum\limits_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$，其中

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x,n=0,1,2,\cdots,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x,n=1,2,\cdots,$$

称这里的a_n,b_n为函数$f(x)$的**傅里叶系数**，以傅里叶系数构成的三角级数

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

称为 $f(x)$ 在 $[-\pi,\pi]$ 上的傅里叶级数.

至此，我给你交代了两件事情. 第一，周期函数可由一个序列的正弦型函数叠加得到；第二，如何求出这样一个序列的正弦型函数（事实上，只要确定了傅里叶系数 a_n, b_n，这个序列自然就唯一确定下来了）. 可是**这两件事情的成立是要有“前提”的**. 这前提到底是什么？下面来告诉你.

3. 傅里叶级数的收敛性 —— 狄利克雷收敛定理

在历史事实中，关于这个“前提”，傅里叶本人虽在多个场合表达过他的观点，可始终未给出明确的定论，这个工作最终由他的学生、著名的抽象派数学大师 —— 狄利克雷在 1829 年完成，史称**狄利克雷收敛定理**.

设函数 $f(x)$ 在 $[-\pi,\pi]$ 上连续或只有有限个第一类间断点，且至多只有有限个（真正的）极值点，则 $f(x)$ 的傅里叶级数

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

在 $[-\pi,\pi]$ 上处处收敛，记其和函数为 $S(x)$，则

$$S(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

且

$$S(x)=\begin{cases}f(x), & x\text{ 为连续点},\\ \dfrac{f(x-0)+f(x+0)}{2}, & x\text{ 为间断点},\\ \dfrac{f(-\pi+0)+f(\pi-0)}{2}, & x=\pm\pi,\end{cases}$$

其中 $f(x_0+0)$，$f(x_0-0)$ 分别表示 $\lim\limits_{x\to x_0^+}f(x)$，$\lim\limits_{x\to x_0^-}f(x)$.

这个定理（也就是前提）的证明超出了高等数学的要求，我们就不证了. 不过这里有两句重要的话需要说明.

第一，你还记得一个函数能在某点展开为收敛于本身的泰勒级数需要什么条件吗？是的，比较苛刻. 此函数要在该点**无穷阶可导**且在 x 趋于该点时其**泰勒余项要趋向于** 0. 而这里，一个函数要展开为收敛于本身的傅里叶级数，条件很“轻松”地就能满足 —— 狄利克雷说的“$f(x)$ 在 $[-\pi,\pi]$ 上连续或只有有限个第一类间断点，且至多只有有限个（真正的）极值点”，这条件把极为广泛的函数都囊括在内了 —— 你可以随手画出一个“违背”狄利克雷条件的函数吗？那很困难. 正因如此，将一个函数展开成傅里叶级数，在各个领域都有着极为广泛的应用.

第二，这个定理明确指出，只有当 x 为连续点时，才可理直气壮地写 $S(x)=f(x)$，其他情形下并非能使这等式成立，故今后你要注意下面这个科学的写法：

$$f(x)\sim S(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

其中“~”读作“展开为”. 记住，$f(x)$ 与 $S(x)$ 并非处处相等，注例 2 后面的“注意”中会让你更清楚地看到这一点.

注例 2　设 $f(x)=\begin{cases}0, & -\pi<x<0,\\ 1, & 0\leqslant x<\pi,\end{cases}$ 在 $[-\pi,\pi]$ 上将其展开为傅里叶级数.

解　显然题中所给 $f(x)$ 在 $[-\pi,\pi]$ 上是满足狄利克雷收敛定理的条件的. 下面需求出傅里叶系数.

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\mathrm{d}x=\frac{1}{\pi}\int_{0}^{\pi}1\mathrm{d}x=1,$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\,\mathrm{d}x=\frac{1}{\pi}\int_{0}^{\pi}\cos nx\,\mathrm{d}x=0,n=1,2,\cdots,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\,\mathrm{d}x=\frac{1}{\pi}\int_{0}^{\pi}\sin nx\,\mathrm{d}x=-\frac{1}{n\pi}\cos nx\Big|_{0}^{\pi}$$

$$=-\frac{1}{n\pi}\cos n\pi+\frac{1}{n\pi}=[(-1)^{n+1}+1]\cdot\frac{1}{n\pi},$$

其中 $\cos n\pi=(-1)^n,n=1,2,\cdots$，于是当 n 为偶数时，$(-1)^{n+1}+1=0$，所以

$$b_2=b_4=b_6=\cdots=0.$$

故

$$f(x)\sim S(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$$

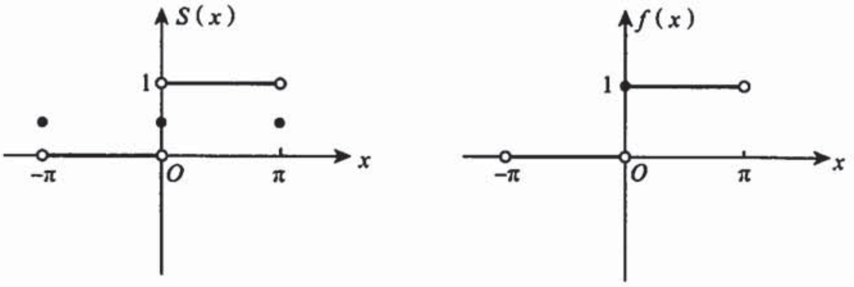
$$=\frac{1}{2}+\frac{2}{\pi}\left(\frac{\sin x}{1}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots\right)$$

$$=\frac{1}{2}+\frac{2}{\pi}\sum_{k=0}^{\infty}\frac{\sin(2k+1)x}{2k+1}.$$

注意 第一，读者与我一起再仔细研究一下这个结果. 根据狄利克雷收敛定理，可知

$$S(x)=\begin{cases}f(x), & -\pi<x<0,0<x<\pi,\\ \dfrac{f(0+0)+f(0-0)}{2}=\dfrac{1}{2}, & x=0,\\ \dfrac{f(-\pi+0)+f(\pi-0)}{2}=\dfrac{1}{2}, & x=\pm\pi.\end{cases}$$

下面分别画出 $S(x)$，$f(x)$ 在 $[-\pi,\pi]$ 上的图像(如图 1-15-8)，供读者直观对比一下.



(a) (h)

图 1-15-8

显然，在 $-\pi<x<0$ 与 $0<x<\pi$ 这样的区间上，有 $S(x)=f(x)$；但在 $x=-\pi,0,\pi$ 这三个点处，$S(x)\neq f(x)$.

第二，由于 $\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx)$ 是以 2π 为周期的，故这样的展开结果不仅在 $[-\pi,\pi]$ 上正确，**在整个数轴上也是确定无疑的**. 所以，虽然题目给出的区间仅限于 $[-\pi,\pi]$，但这结果可直接推广至整个实数轴上. 稍作修改，就可写出下例：设 $f(x)$ 是以 2π 为周期的函数，且其在 $(-\pi,\pi)$ 上的表达式为

$$f(x)=\begin{cases}0, & -\pi<x<0,\\ 1, & 0\leqslant x<\pi.\end{cases}$$

请将 $f(x)$ 在实数轴上展开为傅里叶级数.

其实，我们的展开工作仍然是在 $[-\pi,\pi]$ 上进行，与注例 2 无异，只需在写出结果后"顺便"延拓至整个数轴即可. 写成

$$f(x) \sim \frac{1}{2}+\frac{2}{\pi}\left(\frac{\sin x}{1}+\frac{\sin 3x}{3}+\frac{\sin 5x}{5}+\cdots\right)=\begin{cases} f(x), & x \neq m\pi, \\ \dfrac{1}{2}, & x=m\pi, \end{cases}$$

其中 m 为任意整数，分别画出 $f(x)$ 与 $S(x)$ 的图像（如图 1-15-9），供读者对比.

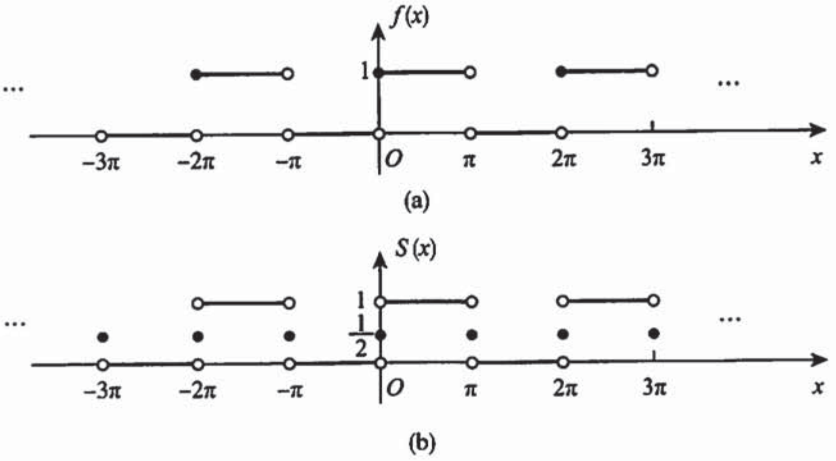


图 1-15-9

4. 推广至任意区间上的傅里叶展开式

上面我们的讨论仅限于$[-\pi,\pi]$这个区间，区间端点定了，区间长度也定了，这未免有些死板，导致使用有一定的局限性 —— 区间端点不是$-\pi,\pi$时，区间长度不是2π时，还能讨论傅里叶展开式吗?回答是肯定的.

第一个问题，区间端点的移动，相当容易.你还记得这个定理么?设 $f(x)$ 是以 T 为周期的可积函数，则对于任意实数 a，有

$$\int_0^T f(x)\mathrm{d}x=\int_a^{a+T} f(x)\mathrm{d}x.$$

这定理用通俗的话讲就是，周期函数 $f(x)$ 在一个周期上的积分值与起点无关.于是，当 $f(x)$ 以 2π 为周期时，$[-\pi,\pi]$ 可随意换成$[a,a+2\pi]$，这个问题迎刃而解.

第二个问题，区间长度的改变，作变量代换可解决.若 $f(x)$ 的周期是 $2l$，而不是 2π，即研究区间$[-l,l]$，则可令 $t=\frac{\pi}{l}x$，当$-l \leqslant x \leqslant l$时，有$-\pi \leqslant t \leqslant \pi$.记 $f(x)=f\left(\frac{l}{\pi}t\right)=g(t)$，$-\pi \leqslant t \leqslant \pi$.于是有

$$g(t) \sim \frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nt+b_n\sin nt),$$

其中
$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos nt\,\mathrm{d}t, n=0,1,2,\cdots,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sin nt\,\mathrm{d}t, n=1,2,\cdots.$$

将变量 t 换回到 x，有

$$f(x) \sim \frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{l}x+b_n\sin\frac{n\pi}{l}x\right),$$

其中

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos nt\,\mathrm{d}t \xlongequal{t=\frac{\pi}{l}x} \frac{1}{l}\int_{-l}^{l} f(x)\cos\frac{n\pi}{l}x\,\mathrm{d}x, n=0,1,2,\cdots,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}g(t)\sin nt\,\mathrm{d}t\xlongequal{t=\frac{\pi}{l}x}\frac{1}{l}\int_{-l}^{l}f(x)\sin\frac{n\pi}{l}x\,\mathrm{d}x,n=1,2,\cdots.$$

这里,对定义在$[-l,l]$上的函数,狄利克雷收敛定理同样成立.读者只需回到前面,把"π"改成"l",并注意自变量的改变即可.

设$f(x)$在$[-l,l]$上连续或只有有限个第一类间断点,且至多只有有限个(真正的)极值点,则$f(x)$的傅里叶级数处处收敛,记其和函数为$S(x)$,有

$$S(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{l}x+b_n\sin\frac{n\pi}{l}x\right),$$

且

$$S(x)=\begin{cases}f(x), & x\text{ 为连续点},\\ \dfrac{f(x-0)+f(x+0)}{2}, & x\text{ 为间断点},\\ \dfrac{f(-l+0)+f(l-0)}{2}, & x=\pm l.\end{cases}$$

5. 正弦级数与余弦级数

若$f(x)$是$[-l,l]$上可积的奇函数,则$\int_{-l}^{l}f(x)\mathrm{d}x=0$;若$f(x)$是$[-l,l]$上可积的偶函数,则

$$\int_{-l}^{l}f(x)\mathrm{d}x=2\int_{0}^{l}f(x)\mathrm{d}x.$$

如此一来,便可得到下面两种特殊情形下的傅里叶展开式.

先看$f(x)$是$[-l,l]$上的奇函数的情形.有

$$a_n=\frac{1}{l}\int_{-l}^{l}f(x)\cos\frac{n\pi}{l}x\,\mathrm{d}x=0,n=0,1,2,\cdots,$$

$$b_n=\frac{2}{l}\int_{0}^{l}f(x)\sin\frac{n\pi}{l}x\,\mathrm{d}x,n=1,2,\cdots.$$

于是,$f(x)\sim\sum\limits_{n=1}^{\infty}b_n\sin\frac{n\pi}{l}x$.由于展开式中只含正弦函数,故称其为**正弦级数**.

再来看$f(x)$是$[-l,l]$上的偶函数的情形.有

$$a_n=\frac{2}{l}\int_{0}^{l}f(x)\cos\frac{n\pi}{l}x\,\mathrm{d}x,n=0,1,2,\cdots,$$

$$b_n=\frac{1}{l}\int_{-l}^{l}f(x)\sin\frac{n\pi}{l}x\,\mathrm{d}x=0,n=1,2,\cdots.$$

于是,$f(x)\sim\frac{a_0}{2}+\sum\limits_{n=1}^{\infty}a_n\cos\frac{n\pi}{l}x$.同样地,展开式中只含余弦函数,故称其为**余弦级数**.

我们知道,任何一个定义在$[-l,l]$上的函数均可表示为一个奇函数与一个偶函数之和:

$$f(x)=\frac{f(x)-f(-x)}{2}+\frac{f(x)+f(-x)}{2},$$

其中$f_1(x)=\dfrac{f(x)-f(-x)}{2}$为奇函数,$f_2(x)=\dfrac{f(x)+f(-x)}{2}$为偶函数.自然可以联系到上面刚刚讲过的知识而得出这样的结论:$f(x)$的傅里叶级数是由$f_1(x)$展开的正弦级数和由$f_2(x)$展开的余弦级数的叠加.

这一部分的最后,再谈一个要紧的问题.若$f(x)$只在$[0,l]$上给出,要将其展开为以$2l$为周期的傅里叶级数,该怎么办?

读者只需再翻回狄利克雷收敛定理,看看便知,我们在$[-l,0]$这个区间上补充$f(x)$的定义而使其满

足狄利克雷收敛定理的条件，即可将补充定义后的 $f(x)$ 在 $[-l,l]$ 上展开成傅里叶级数了，我们将这种补充定义的手段叫作**延拓**. 显然，在 $[-l,0]$ 上给出不同的延拓形式，就会展开成不同的傅里叶级数. 为方便起见，我们常作以下两种延拓.

(1) **奇延拓**，也就是令

$$F(x)=\begin{cases} f(x), & 0<x\leqslant l,\\ -f(-x), & -l\leqslant x<0,\\ 0, & x=0,\end{cases}$$

这个 $F(x)$ 是延拓之后的奇函数.

如图1-15-10 所示. 这里应特别注意 $x=0$ 的情形. 若原先 $f(0)\neq 0$，只能修改此处定义，强行令 $F(0)=0$，以保证 $F(x)$ 为 $[-l,l]$ 上的奇函数.

(2) **偶延拓**，即令

$$F(x)=\begin{cases} f(x), & 0\leqslant x\leqslant l,\\ f(-x), & -l\leqslant x<0.\end{cases}$$

这个 $F(x)$ 是延拓后的偶函数. 如图 1-15-11 所示.

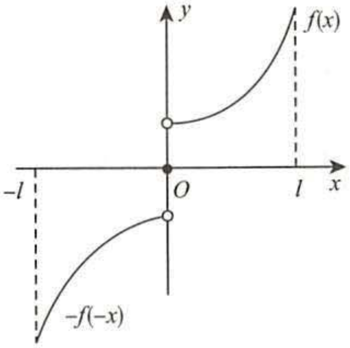


图 1-15-10

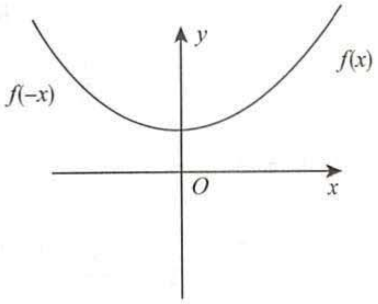


图 1-15-11

两种延拓后的奇、偶函数便可用前面所讲的方法分别展开成正弦、余弦级数了.

至此，基本理论全部讲完了. 下面通过一个例题来欣赏引言中所提到的欧拉写给哥德巴赫的信中的精彩内容.

注例 3　将函数 $f(x)=\frac{\pi}{2}-\frac{x}{2}$ 在区间 $(0,2\pi)$ 内展开成傅里叶级数.

解　此函数显然满足狄利克雷收敛定理的条件. 轻车熟路，我们来求傅里叶系数.

$$a_0 \overset{(*)}{=\!=} \frac{1}{\pi}\int_0^{2\pi} f(x)\mathrm{d}x = \frac{1}{\pi}\int_0^{2\pi}\left(\frac{\pi}{2}-\frac{x}{2}\right)\mathrm{d}x = 0.$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi}\frac{\pi - x}{2}\cos nx\,\mathrm{d}x = 0,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi}\frac{\pi - x}{2}\sin nx\,\mathrm{d}x = \frac{1}{n},$$

其中 $n=1,2,\cdots$.

则 $f(x)$ 的傅里叶级数展开式为

$$f(x)=\frac{\pi}{2}-\frac{x}{2}=S(x)=\sum_{n=1}^{\infty}\frac{\sin nx}{n}, 0<x<2\pi.$$

【注1】 (＊)处的意思是想再次提醒读者，周期为2π的函数$f(x)$的傅里叶系数公式均为$[-\pi,\pi]$上的积分，但由于对任意实数a，均有

$$\int_0^T f(x)\mathrm{d}x=\int_a^{a+T}f(x)\mathrm{d}x,$$

故可直接“平移”积分区间至题中所给的0到2π.

【注2】 (1) 在端点$0,2\pi$处，有$S(0)=S(2\pi)=\dfrac{f(0+0)+f(2\pi-0)}{2}=0.$

并如前所述，将$S(x)$的周期为2π的特性“充分”展示出来，就得到了下面的图像(如图1-15-12).

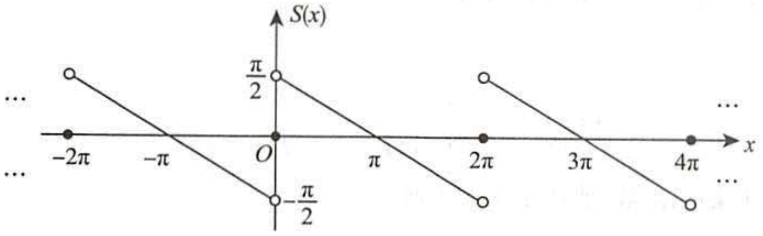


图 1-15-12

(2) 对

$$\frac{\pi}{2}-\frac{x}{2}=\sum_{n=1}^{\infty}\frac{\sin nx}{n}(0<x<2\pi)\qquad(*)$$

做一些有趣的工作. 比如令$2t=x$，则有$\dfrac{\pi}{2}-t=\sum\limits_{n=1}^{\infty}\dfrac{\sin 2nt}{n}(0<2t<2\pi)$，等式两边同时除以2，即

$$\frac{\pi}{4}-\frac{t}{2}=\sum_{n=1}^{\infty}\frac{\sin 2nt}{2n}=\sum_{k=1}^{\infty}\frac{\sin 2kt}{2k}(0<t<\pi),\qquad(**)$$

显然，这里$2k$是**偶数项**的标志.

在(＊＊)式中令$x=t$，用(＊)式减去(＊＊)式，当然也就得到了级数中**奇数项**的和：

$$\frac{\pi}{4}=\sum_{k=1}^{\infty}\frac{\sin(2k-1)x}{2k-1},0<x<\pi.$$

将(＊)式与(＊＊)式写得清楚些：

$$\frac{\pi}{2}-\frac{x}{2}=\frac{\sin x}{1}+\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\frac{\sin 4x}{4}+\cdots,$$

$$\frac{\pi}{4}-\frac{x}{2}=\frac{\sin 2x}{2}+\frac{\sin 4x}{4}+\cdots.$$

若用(＊)式减去2倍的(＊＊)式，有

$$\left(\frac{\pi}{2}-\frac{x}{2}\right)-2\left(\frac{\pi}{4}-\frac{x}{2}\right)=\left(\frac{\sin x}{1}+\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\frac{\sin 4x}{4}+\cdots\right)-2\left(\frac{\sin 2x}{2}+\frac{\sin 4x}{4}+\cdots\right),$$

即

$$\frac{x}{2}=\frac{\sin x}{1}-\frac{\sin 2x}{2}+\frac{\sin 3x}{3}-\frac{\sin 4x}{4}+\cdots,$$

也就是

$$x=2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}(0<x<\pi).$$

由于上式两边均为奇函数，此时定义域可写为$-\pi<x<\pi$. 这便回到了引言中的例子——读者现在可以重新感受到那个例子的精彩——在赞叹欧拉天赋的同时，你掌握傅里叶级数了吗？

基础例题精解

一、一元函数微分学应用

例 1.15.1　已知动点 P 在曲线 $y=x^3$ 上运动，记坐标原点与点 P 间的距离为 l. 若点 P 的横坐标对时间的变化率为常数 v_0，则当点 P 运动到点(1,1)时，l 对时间的变化率是________.

解　应填 $2\sqrt{2}v_0$.

由题设知 $l=\sqrt{x^2+y^2}=\sqrt{x^2+x^6}$，则

$$\frac{\mathrm{d}l}{\mathrm{d}t}=\frac{\mathrm{d}l}{\mathrm{d}x}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}=\frac{2x+6x^5}{2\sqrt{x^2+x^6}}\cdot v_0,$$

$$\left.\frac{\mathrm{d}l}{\mathrm{d}t}\right|_{(1,1)}=\frac{8}{2\sqrt{2}}v_0=2\sqrt{2}v_0.$$

例 1.15.2　曲线 $\begin{cases}x=\cos^3 t,\\ y=\sin^3 t\end{cases}$ 在 $t=\frac{\pi}{4}$ 对应点处的曲率为________.

解　应填 $\frac{2}{3}$.

用参数求导法，有

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{y_t'}{x_t'}=\frac{3\sin^2 t\cos t}{-3\cos^2 t\sin t}=-\tan t,\quad \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{t=\frac{\pi}{4}}=-1,$$

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}=(-\tan t)'\frac{\mathrm{d}t}{\mathrm{d}x}=-\frac{1}{\cos^2 t}\cdot\frac{1}{-3\cos^2 t\sin t},\quad \left.\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right|_{t=\frac{\pi}{4}}=\frac{4}{3}\sqrt{2}.$$

按曲率公式，$t=\frac{\pi}{4}$ 对应点处的曲率为

$$k=\left.\frac{\left|\frac{\mathrm{d}^2 y}{\mathrm{d}x^2}\right|}{\left[1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{3/2}}\right|_{t=\frac{\pi}{4}}=\frac{\frac{4}{3}\sqrt{2}}{2^{3/2}}=\frac{2}{3}.$$

二、一元函数积分学应用

例 1.15.3　设地球的质量为 M，半径为 R. 将地面上的质量为 m 的物体举高 H 米，求克服重力所做的功.

解　如图 1-15-13 所示，考虑将物体从距地心 x 米处举高到距地心 $x+\mathrm{d}x$ 米的情况. 在此过程中，假设物体所受的重力大小近似为

$$G\frac{Mm}{x^2}(R\leqslant x\leqslant R+H),$$

其中 G 是引力常量.

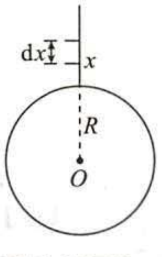


图 1-15-13

将物体从距地心 x 米处举高到距地心 $x+\mathrm{d}x$ 米时，克服重力所做功的微元为

$$\mathrm{d}W=G\frac{Mm}{x^2}\mathrm{d}x.$$

于是所求的功为

$$W=\int_R^{R+H}G\frac{Mm}{x^2}\mathrm{d}x=-\frac{GMm}{x}\Big|_R^{R+H}=GMm\left(\frac{1}{R}-\frac{1}{R+H}\right).$$

例 1.15.4 有一倒圆锥形容器，高 10 m，上底半径 4 m，水面高 8 m. 求将容器中的水全部从容器顶部抽出所做的功（水的密度为 1 000 kg/m^3，重力加速度 g 为 9.8 m/s^2）.

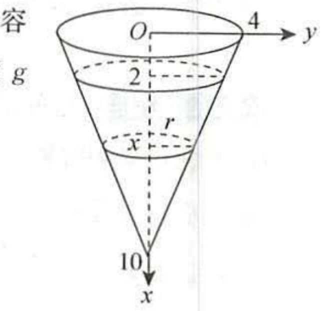


图 1-15-14

解 如图 1-15-14 所示，建立坐标系. 在 x 处的水平截面的半径 r 满足

$$\frac{r}{10-x}=\frac{4}{10}\text{或 }r=\frac{2}{5}(10-x).$$

截面面积 $A(x)=\pi r^2=\pi\cdot\left[\frac{2}{5}(10-x)\right]^2=\frac{4}{25}\pi(10-x)^2.$

所以将水全部抽出所做的功

$$W=1\,000\,g\int_2^{10}xA(x)\mathrm{d}x=1\,000\,g\int_2^{10}x\cdot\frac{4}{25}\pi(10-x)^2\mathrm{d}x$$

$$=160\,g\pi\int_2^{10}x(10-x)^2\mathrm{d}x=160\times9.8\pi\times\frac{2\,048}{3}\approx3\,361(\mathrm{kJ}).$$

例 1.15.5 洒水车上的水箱是一个横放的椭圆柱体，椭圆的尺寸如图 1-15-15 所示. 当水箱装满水时，计算水箱的一个端面所受的压力（水的密度为 1 000 kg/m^3，重力加速度 g 为 9.8 m/s^2）.

解 如图 1-15-16 所示，建立坐标系. 在 x 处椭圆的宽度为 $2y=2\sqrt{1-\frac{x^2}{0.75^2}}$，$x$ 处的深度为 $x+0.75$，所以椭圆面所受到的压力

$$P=1\,000\times9.8\int_{-0.75}^{0.75}(x+0.75)\cdot2\sqrt{1-\frac{x^2}{0.75^2}}\mathrm{d}x$$

$$=1\,000\times2\times9.8\left(\int_{-0.75}^{0.75}x\sqrt{1-\frac{x^2}{0.75^2}}\mathrm{d}x+\int_{-0.75}^{0.75}\sqrt{0.75^2-x^2}\mathrm{d}x\right)$$

$$\overset{(*)}{=\!=}19\,600\left(0+\frac{1}{2}\pi\cdot0.75^2\right)\approx1.73\times10^4(\mathrm{N}).$$

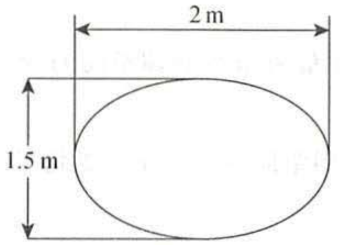


图 1-15-15

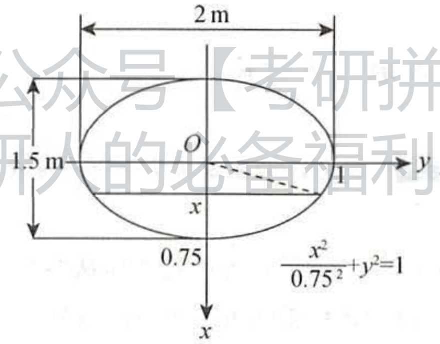


图 1-15-16

【注】 其中（$*$）利用对称性，前一个定积分为零；利用定积分的几何意义，后一个定积分为半圆的面积.

例 1.15.6 设曲线 L 的方程为 $y=\frac{1}{4}x^2-\frac{1}{2}\ln x, 1\leqslant x\leqslant \mathrm{e}$，$D$ 是由曲线 L，直线 $x=1, x=\mathrm{e}$ 及 x 轴围成的平面图形，求 D 的形心的横坐标.

解 平面图形 D 的形心横坐标的计算公式为 $\bar{x}=\dfrac{\int_1^{\mathrm{e}} xy\mathrm{d}x}{\int_1^{\mathrm{e}} y\mathrm{d}x}$，其中

$$\int_1^{\mathrm{e}} xy\mathrm{d}x=\int_1^{\mathrm{e}} x\left(\frac{1}{4}x^2-\frac{1}{2}\ln x\right)\mathrm{d}x=\left(\frac{1}{16}x^4-\frac{1}{4}x^2\ln x+\frac{1}{8}x^2\right)\Bigg|_1^{\mathrm{e}}=\frac{1}{16}(\mathrm{e}^2+1)(\mathrm{e}^2-3),$$

$$\int_1^{\mathrm{e}} y\mathrm{d}x=\int_1^{\mathrm{e}}\left(\frac{1}{4}x^2-\frac{1}{2}\ln x\right)\mathrm{d}x=\left(\frac{1}{12}x^3-\frac{1}{2}x\ln x+\frac{1}{2}x\right)\Bigg|_1^{\mathrm{e}}=\frac{1}{12}\mathrm{e}^3-\frac{7}{12},$$

所以 D 的形心的横坐标为

$$\bar{x}=\frac{\frac{1}{16}(\mathrm{e}^2+1)(\mathrm{e}^2-3)}{\frac{1}{12}\mathrm{e}^3-\frac{7}{12}}=\frac{3(\mathrm{e}^2+1)(\mathrm{e}^2-3)}{4(\mathrm{e}^3-7)}.$$

【注】 平面区域 D 的形心一般表达式：$\bar{x}=\dfrac{\iint_D x\mathrm{d}x\mathrm{d}y}{\iint_D \mathrm{d}x\mathrm{d}y}$，$\bar{y}=\dfrac{\iint_D y\mathrm{d}x\mathrm{d}y}{\iint_D \mathrm{d}x\mathrm{d}y}$.

例 1.15.7 计算曲线 $y=\ln(1-x^2)$ 上相应于 $0\leqslant x\leqslant\frac{1}{2}$ 的一段弧的长度.

解 $$s=\int_0^{\frac{1}{2}}\sqrt{1+(y')^2}\mathrm{d}x=\int_0^{\frac{1}{2}}\sqrt{1+\left(\frac{-2x}{1-x^2}\right)^2}\mathrm{d}x=\int_0^{\frac{1}{2}}\frac{1+x^2}{1-x^2}\mathrm{d}x$$

$$=\int_0^{\frac{1}{2}}\left(\frac{1}{1+x}+\frac{1}{1-x}-1\right)\mathrm{d}x=\ln 3-\frac{1}{2}.$$

例 1.15.8 求心形线 $r=a(1+\cos\theta)$ 的全长，其中 $a>0$ 是常数.

解 $$r'(\theta)=-a\sin\theta,$$

$$\mathrm{d}s=\sqrt{r^2+(r')^2}\mathrm{d}\theta=a\sqrt{(1+\cos\theta)^2+(-\sin\theta)^2}\mathrm{d}\theta=2a\left|\cos\frac{\theta}{2}\right|\mathrm{d}\theta.$$

由对称性得 $$s=2\int_0^{\pi}2a\cos\frac{\theta}{2}\mathrm{d}\theta=8a\sin\frac{\theta}{2}\Bigg|_0^{\pi}=8a.$$

例 1.15.9 求星形线 $x=\cos^3 t, y=\sin^3 t(0\leqslant t\leqslant 2\pi)$ 的弧长.

解 如图 1-15-17 所示，曲线具有对称性，我们只需计算在第一象限的弧段，即 $t\in\left[0,\frac{\pi}{2}\right]$ 对应的部分弧长. 故

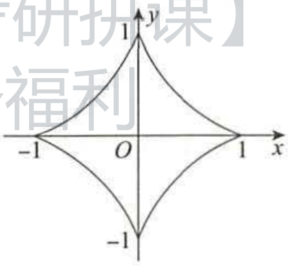


图 1-15-17

$$s=4\int_0^{\frac{\pi}{2}}\sqrt{[x'(t)]^2+[y'(t)]^2}\mathrm{d}t$$

$$=4\int_0^{\frac{\pi}{2}}\sqrt{(-3\cos^2 t\sin t)^2+(3\sin^2 t\cos t)^2}\mathrm{d}t$$

$$=4\int_0^{\frac{\pi}{2}}3|\sin t\cos t|\mathrm{d}t=12\int_0^{\frac{\pi}{2}}\sin t\cos t\mathrm{d}t=6.$$

例 1.15.10 设有曲线 $y=\sqrt{x-1}$，过原点作其切线，求由此曲线、切线及 x 轴围成的平面图形绕 x

轴旋转一周所得到的旋转体的表面积.

解 设切点的横坐标为 x_0，则切点为 $(x_0, \sqrt{x_0-1})$，曲线 $y=\sqrt{x-1}$ 在此点的切线斜率为 $\frac{1}{2\sqrt{x_0-1}}$，于是切线方程为

$$y-\sqrt{x_0-1}=\frac{1}{2\sqrt{x_0-1}}(x-x_0),$$

又因它经过原点，将点 $(0,0)$ 代入，得 $-2(x_0-1)=-x_0$，解得 $x_0=2$，于是切线方程为 $y-1=\frac{1}{2}(x-2)$，即 $y=\frac{1}{2}x$，切点为 $(2,1)$.

由曲线段 $y=\sqrt{x-1}(1\leqslant x\leqslant 2)$ 绕 x 轴旋转一周所得到的旋转面的表面积为

$$S_1=\int_1^2 2\pi y\sqrt{1+(y')^2}\mathrm{d}x=\pi\int_1^2\sqrt{4x-3}\mathrm{d}x=\frac{\pi}{6}(5\sqrt{5}-1).$$

由直线段 $y=\frac{1}{2}x(0\leqslant x\leqslant 2)$ 绕 x 轴旋转一周所得到的旋转面的表面积为

$$S_2=\int_0^2 2\pi\cdot\frac{1}{2}x\cdot\frac{\sqrt{5}}{2}\mathrm{d}x=\sqrt{5}\pi.$$

因此，所求旋转体的表面积为 $S=S_1+S_2=\frac{\pi}{6}(11\sqrt{5}-1)$.

三、微分方程的物理应用

例 1.15.11 某种飞机在机场降落时，为了减少滑行距离，在触地的瞬间，飞机尾部张开减速伞，以增大阻力，使飞机迅速减速并停下.

现有一质量为 9 000 kg 的飞机，着陆时的水平速度为 700 km/h. 经测试，减速伞打开后，飞机所受的总阻力与飞机的速度成正比(比例系数为 $k=6.0\times10^6$). 问从着陆点算起，飞机滑行的最长距离是多少?(kg 表示千克，km/h 表示千米/时)

分析 本题是标准的牛顿第二定律的应用题，列出关系式后再解微分方程即可.

解 由题设，飞机的质量 $m=9\ 000$ kg，着陆时的水平速度 $v_0=700$ km/h. 从飞机接触跑道开始计时，设 t 时刻飞机的滑行距离为 $x(t)$，速度为 $v(t)$.

根据牛顿第二定律，得 $m\frac{\mathrm{d}v}{\mathrm{d}t}=-kv$，又 $\frac{\mathrm{d}v}{\mathrm{d}t}=\frac{\mathrm{d}v}{\mathrm{d}x}\cdot\frac{\mathrm{d}x}{\mathrm{d}t}=v\frac{\mathrm{d}v}{\mathrm{d}x}$，由此可得 $\mathrm{d}x=-\frac{m}{k}\mathrm{d}v$，积分得

$$x(t)=-\frac{m}{k}v+C.$$

由于 $v(0)=v_0$，$x(0)=0$，故得 $C=\frac{m}{k}v_0$，从而

$$x(t)=\frac{m}{k}[v_0-v(t)].$$

当 $v(t)\to 0$ 时，$x(t)\to\frac{mv_0}{k}=\frac{9\ 000\times700}{6.0\times10^6}=1.05(\mathrm{km})$，所以飞机滑行的最长距离为 1.05 km.

【注】 本题求飞机滑行的最长距离，可理解为 $t\to+\infty$ 或 $v(t)\to 0$ 的极限值，这种条件应引起注意.

例 1.15.12　已知高温物体置于低温介质中，任一时刻该物体温度对时间的变化率与该时刻物体和介质的温差成正比．现将一初始温度为 120 ℃的物体在 20 ℃恒温介质中冷却，30 min后该物体温度降至 30 ℃，若要将该物体的温度继续降至 21 ℃，还需冷却多长时间？

解　设该物体在 t 时刻的温度为 $T(t)$ ℃，由题意得

$$\frac{\mathrm{d}T}{\mathrm{d}t}=-k(T-20),$$

其中 k 为比例系数，$k>0$，解得

$$T=C\mathrm{e}^{-kt}+20.$$

将初始条件 $T(0)=120$ 代入上式，解得 $C=100$，故

$$T=100\mathrm{e}^{-kt}+20.$$

将 $t=30$，$T=30$ 代入上式得 $k=\frac{\ln 10}{30}$，所以

$$T=100\mathrm{e}^{-\frac{\ln 10}{30}t}+20.$$

令 $T=21$，得 $t=60$．因此要降至 21 ℃，还需 $60-30=30(\mathrm{min})$．

四、欧拉方程（仅数学一要求）

例 1.15.13　欧拉方程 $x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2}+4x\frac{\mathrm{d}y}{\mathrm{d}x}+2y=0(x>0)$的通解为________．

解　应填 $y=\frac{C_1}{x}+\frac{C_2}{x^2}$，$C_1$，$C_2$ 为任意常数．

由题设，$x>0$，令 $x=\mathrm{e}^t$，则　$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{\mathrm{d}y}{\mathrm{d}t}\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{x}\frac{\mathrm{d}y}{\mathrm{d}t},$$

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}=-\frac{1}{x^2}\frac{\mathrm{d}y}{\mathrm{d}t}+\frac{1}{x}\frac{\mathrm{d}^2y}{\mathrm{d}t^2}\cdot\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{x^2}\left(\frac{\mathrm{d}^2y}{\mathrm{d}t^2}-\frac{\mathrm{d}y}{\mathrm{d}t}\right),$$

代入原方程，得　$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2}+3\frac{\mathrm{d}y}{\mathrm{d}t}+2y=0,$$

解此方程得通解为 $y=C_1\mathrm{e}^{-t}+C_2\mathrm{e}^{-2t}=\frac{C_1}{x}+\frac{C_2}{x^2}$，$C_1$，$C_2$ 为任意常数．

五、傅里叶级数（仅数学一要求）

例 1.15.14　设 $f(x)=\begin{cases}x, & 0\leqslant x\leqslant\frac{1}{2},\\ 2-2x, & \frac{1}{2}<x\leqslant 1,\end{cases}$　$S(x)=\frac{a_0}{2}+\sum\limits_{n=1}^{\infty}a_n\cos n\pi x,-\infty<x<+\infty,$

其中 $a_n=2\int_0^1 f(x)\cos n\pi x\mathrm{d}x(n=0,1,2,\cdots)$，求 $S\left(-\frac{5}{2}\right)$．

解　由余弦级数 $S(x)$ 为函数 $f(x)$ 作偶延拓的傅里叶级数，知

$$S\left(-\frac{5}{2}\right)=S\left(-2-\frac{1}{2}\right)=S\left(-\frac{1}{2}\right)=S\left(\frac{1}{2}\right)$$

$$=\frac{f\left(\frac{1}{2}-0\right)+f\left(\frac{1}{2}+0\right)}{2}=\frac{\frac{1}{2}+\left(2-2\times\frac{1}{2}\right)}{2}=\frac{3}{4}.$$

例 1.15.15 将函数 $f(x)=1-x^2(0\leqslant x\leqslant\pi)$ 展开成余弦级数，并求级数 $\sum\limits_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2}$.

解 将 $f(x)$ 延拓成 $[-\pi,\pi]$ 上的偶函数，则

$$a_0=\frac{2}{\pi}\int_0^{\pi}(1-x^2)\mathrm{d}x=2\left(1-\frac{\pi^2}{3}\right),$$

$$\begin{aligned}a_n&=\frac{2}{\pi}\int_0^{\pi}f(x)\cos nx\,\mathrm{d}x=\frac{2}{\pi}\left(\int_0^{\pi}\cos nx\,\mathrm{d}x-\int_0^{\pi}x^2\cos nx\,\mathrm{d}x\right)\\&=\frac{2}{\pi}\left(0-\int_0^{\pi}x^2\cos nx\,\mathrm{d}x\right)=\frac{-2}{\pi}\left(\frac{x^2\sin nx}{n}\bigg|_0^{\pi}-\int_0^{\pi}\frac{2x\sin nx}{n}\mathrm{d}x\right)\\&=\frac{2}{\pi}\cdot\frac{2\pi\cdot(-1)^{n+1}}{n^2}=\frac{4\cdot(-1)^{n+1}}{n^2},\ n=1,2,\cdots,\end{aligned}$$

$$b_n=0,n=1,2,\cdots,$$

故 $1-x^2=\frac{a_0}{2}+\sum\limits_{n=1}^{\infty}a_n\cos nx=1-\frac{\pi^2}{3}+\sum\limits_{n=1}^{\infty}\frac{4(-1)^{n+1}}{n^2}\cos nx\,(0\leqslant x\leqslant\pi)$.

取 $x=0$，得 $1=1-\frac{\pi^2}{3}+\sum\limits_{n=1}^{\infty}\frac{4(-1)^{n+1}}{n^2}$，于是 $\sum\limits_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2}=\frac{\pi^2}{12}$.

基础习题精练

习题

1.15.1 甲车以 24 km/h 的速度向北行驶，同时正东 10 km 处有乙车以20 km/h的速度向东行驶. 从这一时刻起经过 1 小时后，求两车间的距离对时间的变化率.

1.15.2 设 $\rho=\rho(x)$ 是抛物线 $y=\sqrt{x}$ 上任一点 $M(x,y)(x\geqslant1)$ 处的曲率半径，$s=s(x)$ 是该抛物线上介于点 $A(1,1)$ 与 M 之间的弧长，计算 $3\rho\frac{\mathrm{d}^2\rho}{\mathrm{d}s^2}-\left(\frac{\mathrm{d}\rho}{\mathrm{d}s}\right)^2$ 的值.

1.15.3 在某一人群中推广新技术是通过其中已掌握新技术的人进行的. 设该人群的总人数为 N，在 $t=0$时刻已掌握新技术的人数为 x_0，在任意时刻 t 已掌握新技术的人数为 $x(t)$（将 $x(t)$ 视为连续可微变量），其变化率与已掌握新技术人数和未掌握新技术人数之积成正比，比例常数 $k>0$，求 $x(t)$.

1.15.4 有一半径为 4 m 的半球形水池蓄满了水，现在要将水全部抽到距水池原水面6 m高的水箱中，求需做多少功（水的密度为 1 000 kg/m^3，重力加速度 $g=9.8$ m/s^2）.

1.15.5 求对数螺线 $r=\mathrm{e}^{2\theta}$ 上 $\theta=0$ 到 $\theta=2\pi$ 的一段弧长.

1.15.6 求摆线 $\begin{cases}x=a(t-\sin t),\\y=a(1-\cos t)\end{cases}(a>0)$ 的一拱 $(0\leqslant t\leqslant2\pi)$ 的弧长.

1.15.7 求阿基米德螺线 $r=a\theta(a>0)$ 相应于 θ 从 0 到 2π 一段的弧长.

解答

1.15.1　解　设甲车最初在原点 O 处，乙车在 C 处，$OC=10$ km，在 t 小时后，甲在 A 点，乙在 B 点，如图 1-15-18 所示. 设 $AB=s$，$OA=x$，$CB=y$，则 $s=\sqrt{x^2+(y+10)^2}$，其中 $s=s(t)$，$x=x(t)$，$y=y(t)$ 都是关于 t 的函数. 写成

$$s^2=x^2+(y+10)^2,$$

两边对 t 求导，得 $2s\cdot\frac{\mathrm{d}s}{\mathrm{d}t}=2x\frac{\mathrm{d}x}{\mathrm{d}t}+2(y+10)\frac{\mathrm{d}y}{\mathrm{d}t}$，即

$$\frac{\mathrm{d}s}{\mathrm{d}t}=\frac{x\frac{\mathrm{d}x}{\mathrm{d}t}+(y+10)\frac{\mathrm{d}y}{\mathrm{d}t}}{\sqrt{x^2+(y+10)^2}}.$$

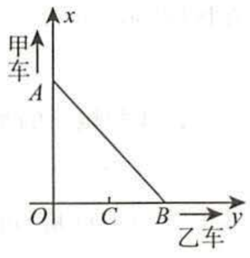


图 1-15-18

上式表达了三个变化率 $\frac{\mathrm{d}s}{\mathrm{d}t}$，$\frac{\mathrm{d}x}{\mathrm{d}t}$，$\frac{\mathrm{d}y}{\mathrm{d}t}$ 之间的关系. 已知 $\frac{\mathrm{d}x}{\mathrm{d}t}=24$，$\frac{\mathrm{d}y}{\mathrm{d}t}=20$；$t=1$ 时，$x=24$，$y=20$，代入上式，得 $\frac{\mathrm{d}s}{\mathrm{d}t}=\frac{196}{\sqrt{41}}\approx30.6$(km/h).

1.15.2　解　由题知 $y'=\frac{1}{2\sqrt{x}}$，$y''=-\frac{1}{4\sqrt{x^3}}$，代入曲率半径公式，得抛物线在点 $M(x,y)$ 处的曲率半径

$$\rho=\rho(x)=\frac{[1+(y')^2]^{\frac{3}{2}}}{|y''|}=\frac{1}{2}(4x+1)^{\frac{3}{2}}.$$

又抛物线上 AM 的弧长

$$s=s(x)=\int_1^x\sqrt{1+(y')^2}\,\mathrm{d}t=\int_1^x\sqrt{1+\frac{1}{4t}}\,\mathrm{d}t.$$

故

$$\frac{\mathrm{d}\rho}{\mathrm{d}s}=\frac{\frac{\mathrm{d}\rho}{\mathrm{d}x}}{\frac{\mathrm{d}s}{\mathrm{d}x}}=\frac{\frac{1}{2}\cdot\frac{3}{2}(4x+1)^{\frac{1}{2}}\cdot4}{\sqrt{1+\frac{1}{4x}}}=6\sqrt{x},$$

$$\frac{\mathrm{d}^2\rho}{\mathrm{d}s^2}=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}\rho}{\mathrm{d}s}\right)\cdot\frac{1}{\frac{\mathrm{d}s}{\mathrm{d}x}}=\frac{6}{2\sqrt{x}}\cdot\frac{1}{\sqrt{1+\frac{1}{4x}}}=\frac{6}{\sqrt{4x+1}},$$

因此

$$3\rho\frac{\mathrm{d}^2\rho}{\mathrm{d}s^2}-\left(\frac{\mathrm{d}\rho}{\mathrm{d}s}\right)^2=3\cdot\frac{1}{2}(4x+1)^{\frac{3}{2}}\cdot\frac{6}{\sqrt{4x+1}}-36x=9.$$

【注】　本题不仅要求记住曲率公式、弧长公式，而且基本运算量比较大，容易算错.

1.15.3　解　由题设可得 $\frac{\mathrm{d}x}{\mathrm{d}t}=kx(N-x)$，$x\Big|_{t=0}=x_0$，分离变量积分得 $x=\frac{NCe^{kNt}}{1+Ce^{kNt}}$.

由初始条件得特解

$$x=\frac{Nx_0e^{kNt}}{N-x_0+x_0e^{kNt}}.$$

【注】　此题十分容易，但实考得分率不高. 题中文字多，考生容易迷糊，甚至不知道这是可分离变量的微分方程，即使知道，有的也不知道如何去积分. 对于简单题的分数，考生一定要拿到手.

1.15.4 解 如图 1-15-19 所示，建立坐标系. 在高度为 y 处的水面面积为

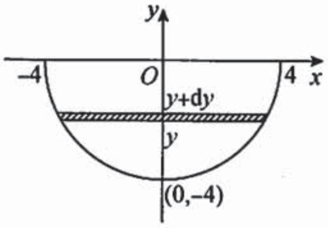


图 1-15-19

$$\pi x^2=\pi(4^2-y^2)=\pi(16-y^2),$$

在区间$[y,y+\mathrm{d}y]$上的体积微元为

$$\pi(16-y^2)\mathrm{d}y,$$

提升此体积微元的水所需要的力的微元为

$$\rho g\pi(16-y^2)\mathrm{d}y,$$

其中 $\rho=1\,000\ \mathrm{kg/m^3}$，$g=9.8\ \mathrm{m/s^2}$. 提升到距原水面 6 m 高处等于提升距离为$(6-y)$，从而提升此微元的水需做功的微元为 $$(6-y)\rho g\pi(16-y^2)\mathrm{d}y,$$

所以将水全部提升至原水面上方 6 m 处需做功为

$$W=\int_{-4}^{0}(6-y)\rho g\pi(16-y^2)\mathrm{d}y=320\pi\rho g\approx 9\,852(\mathrm{kJ}).$$

1.15.5 解 $r'=2\mathrm{e}^{2\theta}$，弧长 $s=\int_0^{2\pi}\sqrt{r^2+(r')^2}\mathrm{d}\theta=\int_0^{2\pi}\sqrt{5}\mathrm{e}^{2\theta}\mathrm{d}\theta=\frac{\sqrt{5}}{2}(\mathrm{e}^{4\pi}-1).$

1.15.6 解 曲线弧图形如图 1-15-20 所示.

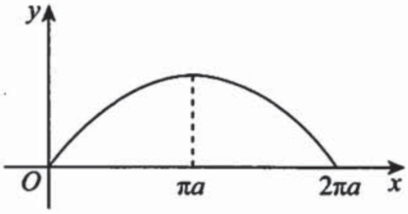


图 1-15-20

由题意，所求弧长为

$$\begin{aligned}s&=\int_0^{2\pi}\sqrt{[x'(t)]^2+[y'(t)]^2}\mathrm{d}t\\&=a\int_0^{2\pi}\sqrt{2(1-\cos t)}\mathrm{d}t\\&=2a\int_0^{2\pi}\sin\frac{t}{2}\mathrm{d}t=2a\left(-2\cos\frac{t}{2}\right)\Big|_0^{2\pi}=8a.\end{aligned}$$

1.15.7 解 曲线弧图形如图 1-15-21 所示.

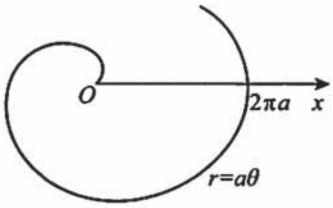


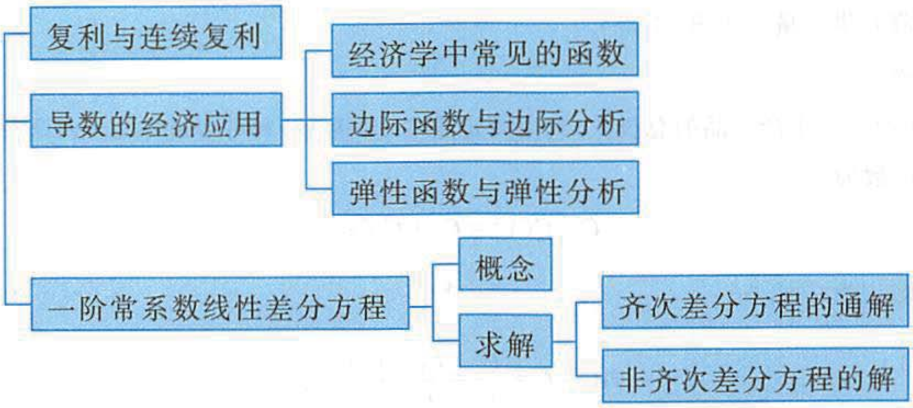
图 1-15-21

由题意，所求弧长为

$$\begin{aligned}s&=\int_0^{2\pi}\sqrt{r^2(\theta)+[r'(\theta)]^2}\mathrm{d}\theta=\int_0^{2\pi}\sqrt{a^2\theta^2+a^2}\mathrm{d}\theta\\&=a\int_0^{2\pi}\sqrt{1+\theta^2}\mathrm{d}\theta\\&=a\left[\frac{\theta}{2}\sqrt{1+\theta^2}+\frac{1}{2}\ln(\theta+\sqrt{1+\theta^2})\right]\Big|_0^{2\pi}\\&=a\left[\pi\sqrt{1+4\pi^2}+\frac{1}{2}\ln(2\pi+\sqrt{1+4\pi^2})\right].\end{aligned}$$

第16讲　数学三专题内容

基础知识结构



基础内容精讲

一、复利与连续复利

复利计算公式

$$A_m = A(1+r)^m .$$

其中 A 表示一开始的本金，r 表示每一期的利率，m 表示复利的总期数，A_m 表示 m 期后的余额.

①如果年利率为 r 的利息一年支付 1 次，那么当初始存款为 A 元时，t 年后余额 A_t 则为

$$A_t = A(1+r)^t .$$

②如果年利率为 r 的利息一年支付 n 次，那么当初始存款为 A 元时，t 年后余额 A_t 则为

$$A_t = A\left(1+\frac{r}{n}\right)^{nt} .$$

③对于②，当 $n \to \infty$ 时，$\lim\limits_{n\to\infty} A_t = \lim\limits_{n\to\infty} A\left(1+\frac{r}{n}\right)^{nt} = Ae^{rt}$，这称为**连续复利**.

【注】 考试时要弄清楚①，②，③三种情况，题目会明确告知.

二、导数的经济应用

1. 经济学中常见的函数

(1)需求函数.

设某产品的需求量为x,价格为p.一般地,需求量x作为价格p的函数$x=\varphi(p)$,称为**需求函数**.并且价格p上升(下降),需求量x下降(上升).需求函数的反函数$p=\varphi^{-1}(x)$称为**价格函数**,也常称为反需求函数.

(2)供给函数.

设某产品的供给量为x,价格为p.一般地,供给量x作为价格p的函数$x=\psi(p)$,称为**供给函数**.并且价格p上升(下降),供给量x上升(下降).

(3)成本函数.

成本$C=C(x)$——生产产品的总投入.它由固定成本C_1(常量)和可变成本$C_2(x)$两部分组成,其中x表示产量.成本函数为

$$C=C(x)=C_1+C_2(x).$$

称$\frac{C}{x}$为**平均成本**,记为$\overline{C}$或AC:

$$AC=\overline{C}=\frac{C}{x}=\frac{C_1}{x}+\frac{C_2(x)}{x}.$$

(4)收益(入)函数.

收益$R=R(x)$——产品售出后所得的收入,是销售量x与销售单价p之积.收益函数为

$$R=R(x)=px.$$

(5)利润函数.

利润$L=L(x)$——收益扣除成本后的余额,由总收益减去总成本得到.利润函数为

$$L=L(x)=R(x)-C(x)(x:销售量).$$

2. 边际函数与边际分析

在经济学中,导函数称为边际函数.若函数$f(x)$可导,则称$f'(x)$为$f(x)$的**边际函数**.$f'(x_0)$称为$f(x)$在x_0点的**边际值**.用边际函数来分析经济量的变化叫**边际分析**.

由$\Delta y\approx \mathrm{d}y$,即$f(x_0+\Delta x)-f(x_0)\approx f'(x_0)\Delta x$,取$\Delta x=1$得

$$f(x_0+1)-f(x_0)\approx f'(x_0).$$

于是,边际值$f'(x_0)$被解释为:在x_0点,当x改变一个单位时,函数$f(x)$近似(实际问题中,经常略去"近似"二字)改变$|f'(x_0)|$个单位.$f'(x_0)$的符号反映自变量的改变与因变量的改变是同向还是反向.

(1)边际成本.

设总成本函数为

$$C=C(q)(q:产量),$$

则**边际成本函数**(记为MC)为

$$MC=C'(q).$$

产量为q_0时的边际成本$C'(q_0)$表示:当产量为q_0时,产量q改变一个单位,总成本$C(q)$将改变$|C'(q_0)|$个单位.$C'(q_0)$的符号反映产量q的改变与成本$C(q)$的改变是同向还是反向.

(2)边际收益.

设总收益函数为

$$R=R(q)(q:\text{销售量}),$$

则边际收益函数(记为 MR)为

$$MR=R'(q).$$

销售量为 q_0 时的边际收益 $R'(q_0)$ 表示:当销售量为 q_0 时,销售量 q 改变一个单位,总收益将改变 $|R'(q_0)|$ 个单位. $R'(q_0)$ 的符号反映销售量 q 的改变与总收益 R 的改变是同向还是反向.

(3)边际利润.

设利润函数为

$$L=L(q)(q:\text{销售量}),$$

则边际利润函数(记为 ML)为

$$ML=L'(q).$$

销售量为 q_0 时的边际利润 $L'(q_0)$ 表示:当销售量为 q_0 时,销售量 q 改变一个单位,利润将改变 $|L'(q_0)|$ 个单位. $L'(q_0)$ 的符号反映销售量 q 的改变与利润 L 的改变是同向还是反向.

3. 弹性函数与弹性分析

在经济学中,把因变量对自变量变化的反应的灵敏度,称为**弹性或弹性系数**. 设函数 $y=f(x)$ 可导,称

$$\eta=\lim_{\Delta x\to 0}\frac{\Delta y}{y}\Big/\frac{\Delta x}{x}=\frac{x}{y}y'=\frac{x}{f(x)}f'(x)$$

为函数 $y=f(x)$ 的**弹性函数**,称

$$\eta\Big|_{x=x_0}=\frac{x_0}{f(x_0)}f'(x_0)$$

为函数 $f(x)$ 在 x_0 处的(**点**)**弹性**.

$\eta\Big|_{x=x_0}$ 表示在 x_0 处,当自变量 x 改变 1%时,因变量 y 将改变 $\left|\eta\Big|_{x=x_0}\right|$%. 其符号表示自变量 x 与因变量 y 的改变是同向还是反向.

用弹性函数来分析经济量的变化叫**弹性分析**.

(1)需求的价格弹性.

设需求函数为 $Q=\varphi(p)$(p:价格,Q:需求量),则需求弹性为

$$\eta_d=\frac{p}{\varphi(p)}\varphi'(p).$$

由于需求函数单调递减,故 $\varphi'(p)<0$,从而 $\eta_d<0$.

其经济意义:当价格为 p 时,若提价(降价)1%,则需求量将减少(增加)$|\eta_d|$%.

【注】 若题设要求 $\eta_d>0$,则取 $\eta_d=-\frac{p}{\varphi(p)}\cdot\varphi'(p)$.

(2)供给的价格弹性.

设供给函数为 $Q=\psi(p)$(p:价格,Q:供给量),则供给弹性为

$$\eta_s=\frac{p}{\psi(p)}\psi'(p).$$

由于供给函数单调增加,故 $\psi'(p)>0$,从而 $\eta_s>0$.

其经济意义：当价格为 p 时，若提价(降价)1%，则供给量将增加(减少)η_s%.

三、一阶常系数线性差分方程

(一)差分方程的基本概念

1. 函数差分的定义

定义 1 函数 $y_t=f(t)$，$t=0,\pm1,\pm2,\cdots$. 函数 $f(t)$在 t 时刻的**一阶差分**定义为

$$\Delta y_t=y_{t+1}-y_t=f(t+1)-f(t);$$

函数 $f(t)$在 t 时刻的二阶差分定义为

$$\Delta^2 y_t=\Delta(\Delta y_t)=\Delta y_{t+1}-\Delta y_t=y_{t+2}-2y_{t+1}+y_t;$$

其余类推，函数 $f(t)$在 t 时刻的 n 阶差分定义为

$$\Delta^n y_t=\Delta(\Delta^{n-1}y_t)=\Delta^{n-1}y_{t+1}-\Delta^{n-1}y_t=\sum_{k=0}^{n}(-1)^k\frac{n!}{k!(n-k)!}y_{t+n-k}.$$

2. 差分方程及其基本概念

(1)差分方程.

定义 2 含有自变量 t，未知函数 y_t，以及 y_t 的差分 $\Delta y_t,\Delta^2 y_t,\cdots$的函数方程，称为(常)**差分方程**.

n 阶差分方程的一般形式为

$$F(t,y_t,\Delta y_t,\cdots,\Delta^n y_t)=0, \tag{1-16-1}$$

这里 F 为已知函数，且 $\Delta^n y_t$ 必定要出现.

定义 2′ 含有自变量 t 和两个或两个以上函数值 $y_t,y_{t+1},\cdots$的函数方程，称为(常)**差分方程**.

n 阶差分方程的另一种一般形式为

$$F(t,y_t,y_{t+1},\cdots,y_{t+n})=0, \tag{1-16-2}$$

这里 F 是已知函数，且 y_t 与 y_{t+n}必定要出现.

(2)差分方程的阶.

定义 3 出现在差分方程(1-16-1)中差分的最高阶数，称为**差分方程的阶**.

定义 3′ 出现在差分方程(1-16-2)中未知函数下标的最大差，称为**差分方程的阶**.

由于经济学中经常遇到的是按形如(1-16-2)给出的差分方程，因此下述将只研究形如(1-16-2)的差分方程. 差分方程的阶也依照定义 3′给定.

(3)差分方程的解.

定义 4 若函数 $y_t=\varphi(t)$代入方程(1-16-2)，使得对一切的 t 均成为恒等式，则称 $y_t=\varphi(t)$为**差分方程(1-16-2)的解**.

含有 n 个任意独立常数 $C_1,C_2,\cdots,C_n$ 的解

$$y_t=\varphi(t,C_1,C_2,\cdots,C_n),$$

称为 n **阶差分方程(1-16-2)的通解**.

在通解中给任意常数 $C_1,C_2,\cdots,C_n$ 以确定值而得出的解，称为 n **阶差分方程(1-16-2)的特解**.

(二)一阶常系数线性差分方程的求解

一阶常系数线性差分方程的一般形式为

$$y_{t+1}+ay_t=f(t), \tag{1-16-3}$$

其中 $f(t)$ 为已知函数，a 为非零常数.

当 $f(t)\equiv 0$ 时，方程(1-16-3)变为

$$y_{t+1}+ay_t=0, \tag{1-16-4}$$

我们称(1-16-3)为一阶常系数**非齐次线性差分方程**，(1-16-4)为其对应的一阶常系数**齐次线性差分方程**.

1. 齐次差分方程的通解

通过迭代，并由数学归纳法可得(1-16-4)的通解为

$$y_C(t)=C\cdot(-a)^t,$$

这里 C 为任意常数.

2. 非齐次差分方程的解

定理 1　若 y_t^* 是非齐次差分方程(1-16-3)的一个特解，$y_C(t)$ 是齐次差分方程(1-16-4)的通解，则非齐次差分方程(1-16-3)的通解为

$$y_t=y_C(t)+y_t^*.$$

定理 2　若 $\bar{y}_t$ 与 $\tilde{y}_t$ 分别是差分方程

$$y_{t+1}+ay_t=f_1(t)$$

和

$$y_{t+1}+ay_t=f_2(t)$$

的解，则

$$y_t=\bar{y}_t+\tilde{y}_t$$

是差分方程

$$y_{t+1}+ay_t=f_1(t)+f_2(t)$$

的解.

非齐次差分方程(1-16-3)的特解 y_t^* 形式的设定见下表.

(1-16-3)中 $f(t)$ 的形式	取待定特解的条件	试取特解的形式
$f(t)=P_m(t)=b_0+b_1t+\cdots+b_mt^m$	$a\neq -1$	$y_t^*=Q_m(t)=B_0+B_1t+\cdots+B_mt^m$ $B_0,B_1,\cdots,B_m$ 为待定常数
	$a=-1$	$y_t^*=tQ_m(t)$
$f(t)=d^t\cdot P_m(t)$ d 为非零常数	$a+d\neq 0$	$y_t^*=d^tQ_m(t)$
	$a+d=0$	$y_t^*=td^tQ_m(t)$
$f(t)=b_1\cos wt+b_2\sin wt$ $w\neq 0$ 且 b_1,b_2 为不同时为零的常数	$D=\begin{vmatrix} a+\cos w & \sin w \\ -\sin w & a+\cos w \end{vmatrix}\neq 0$	$y_t^*=\alpha\cos wt+\beta\sin wt$ α,β 为待定常数
	$D=0$	$y_t^*=t(\alpha\cos wt+\beta\sin wt)$

基础例题精解

例 1.16.1　设某酒厂有一批新酿的好酒，如果现在(假定 $t=0$)就售出，总收入为 R_0 元，如果窖藏起来待来日按陈酒价格出售，t 年末总收入为 $R=R_0e^{\frac{2}{5}\sqrt{t}}$. 假定银行的年利率为 r，并以连续复利计息，试求窖

藏多少年售出可使总收入的现值最大，并求 $r = 6\%$ 时的 t 值.

解 根据连续复利公式，这批酒在窖藏 t 年年末售出总收入 R 的现值为 $A(t)=Re^{-rt}$，而 $R = R_0e^{\frac{2}{5}\sqrt{t}}$，故 $A(t) = R_0e^{\frac{2}{5}\sqrt{t}-rt}$. 令 $\frac{dA}{dt} = R_0e^{\frac{2}{5}\sqrt{t}-rt}\left(\frac{1}{5\sqrt{t}}-r\right) = 0$，得驻点 $t_0 = \frac{1}{25r^2}$.

又 $\frac{d^2A}{dt^2} = R_0e^{\frac{2}{5}\sqrt{t}-rt}\left[\left(\frac{1}{5\sqrt{t}}-r\right)^2-\frac{1}{10\sqrt{t^3}}\right]$，则有 $\left.\frac{d^2A}{dt^2}\right|_{t=t_0} = R_0e^{\frac{1}{25r}}(-12.5r^3) < 0$.

于是，$t_0 = \frac{1}{25r^2}$ 是极大值点亦是最大值点，故窖藏 $t = \frac{1}{25r^2}$ 年售出可使总收入的现值最大. 当 $r = 6\%$ 时，$t = \frac{100}{9} \approx 11$ (年).

例 1.16.2 设商品的需求函数为 $Q = 100-5p$，其中 Q,p 分别表示需求量和价格，如果商品需求弹性的绝对值大于 1，求商品价格的取值范围.

解 需求量 Q 对价格 p 的弹性 $\eta=\frac{p}{Q}Q'=\frac{-5p}{100-5p}$，由于需求弹性的绝对值大于 1，即 $\left|\frac{-5p}{100-5p}\right|>1$，结合 $Q=100-5p>0$，解得 $p\in(10,20)$.

例 1.16.3 设生产某商品的固定成本为 60 000 元，可变成本为 20 元/件，价格函数为 $p=60-\frac{Q}{1\ 000}$（p 是单价，单位：元；Q 是销量，单位：件）. 已知产销平衡，求：

(1)该商品的边际利润；

(2)当 $p=50$ 时的边际利润，并解释其经济意义；

(3)使得利润最大的定价 p.

解 (1)成本函数 $C(Q)=60\ 000+20Q$，收益函数 $R(Q)=pQ=60Q-\frac{Q^2}{1\ 000}$，利润函数

$$L(Q)=R(Q)-C(Q)=-\frac{Q^2}{1\ 000}+40Q-60\ 000,$$

故该商品的边际利润

$$L'(Q)=-\frac{Q}{500}+40.$$

(2)当 $p=50$ 时，销量 $Q=10\ 000$，$L'(10\ 000)=20$.

其经济意义：销售第 10 001 件商品所得的利润为 20 元.

(3)令 $L'(Q)=-\frac{Q}{500}+40=0$，得 $Q=20\ 000$，且 $L''(20\ 000)<0$，故当 $Q=20\ 000$件时利润最大，此时 $p=40$ 元.

例 1.16.4 设某商品从时刻 0 到时刻 t 的销售量为 $x(t) = kt, t\in[0,T](k>0)$. 欲在 T 时将数量为 A 的该商品销售完，试求：

(1)t 时的商品剩余量，并确定 k 的值；

(2)商品在时间段$[0,T]$上的平均剩余量.

分析 在时刻 t 的剩余量 $y(t)$ 可用总量 A 减去销量 $x(t)$ 得到；由于 $y(t)$ 随时间连续变化，因此在时间段$[0,T]$上的平均剩余量，即函数平均值可用积分 $\frac{1}{T}\int_0^T y(t)dt$ 表示.

解 (1)在时刻 t 商品的剩余量为 $y(t) = A-x(t) = A-kt, t\in[0,T]$.

由 $A-kT=0$ 得 $k=\dfrac{A}{T}$，因此 $y(t)=A-\dfrac{A}{T}t,t\in[0,T]$.

(2)依题意有，$y(t)$ 在$[0,T]$上的平均值为

$$\bar{y}=\frac{1}{T}\int_0^T y(t)\mathrm{d}t=\frac{1}{T}\int_0^T\left(A-\frac{A}{T}t\right)\mathrm{d}t=\frac{A}{2},$$

因此商品在时间段$[0,T]$上的平均剩余量为 $\dfrac{A}{2}$.

例 1.16.5　某公司可通过电台及报纸两种方式做销售某种商品的广告，根据统计资料，销售收入 R(万元)与电台广告费 x_1 (万元)及报纸广告费用 x_2 (万元)之间的关系有如下经验公式：

$$R=15+14x_1+32x_2-8x_1x_2-2x_1^2-10x_2^2.$$

(1)在广告费用不限的情况下，求最优广告策略；

(2)若提供的广告费用为 1.5 万元，求相应的最优广告策略.

解　(1)利润函数为 $z=f(x_1,x_2)=R-(x_1+x_2)=15+13x_1+31x_2-8x_1x_2-2x_1^2-10x_2^2$.

令 $\begin{cases}\dfrac{\partial z}{\partial x_1}=13-8x_2-4x_1=0,\\ \dfrac{\partial z}{\partial x_2}=31-8x_1-20x_2=0,\end{cases}$ 解得 $\begin{cases}x_1=0.75,\\ x_2=1.25.\end{cases}$

故利润函数 $z=f(x_1,x_2)$在$(0.75,1.25)$处达到极大值，也是最大值.

(2)若广告费用为 1.5 万元，则需要求利润函数 $z=f(x_1,x_2)$在 $x_1+x_2=1.5$ 时的条件极值点.

构造拉格朗日函数 $F(x_1,x_2,\lambda)=15+13x_1+31x_2-8x_1x_2-2x_1^2-10x_2^2+\lambda(x_1+x_2-1.5)$.

由方程组 $\begin{cases}\dfrac{\partial F}{\partial x_1}=-8x_2-4x_1+13+\lambda=0,\\ \dfrac{\partial F}{\partial x_2}=-8x_1-20x_2+31+\lambda=0,\\ \dfrac{\partial F}{\partial \lambda}=x_1+x_2-1.5=0,\end{cases}$ 解得 $x_1=0,x_2=1.5$.

故将 1.5 万元广告费全部用于报纸广告，可使利润最大.

例 1.16.6　已知某商品的需求量 x 对价格 p 的弹性 $\eta=-3p^3$，而市场对该商品的最大需求量为 1(万件)，求需求函数.

解　根据弹性定义，有

$$\eta=\frac{\mathrm{d}x}{x}\Big/\frac{\mathrm{d}p}{p}=-3p^3,\quad \frac{\mathrm{d}x}{x}=-3p^2\mathrm{d}p.$$

由此得 $x=C\mathrm{e}^{-p^3}$，C 为待定常数，由条件知 $p=0$ 时，$x=1$，从而 $C=1$，于是所求的需求函数为 $x=\mathrm{e}^{-p^3}$.

例 1.16.7　求下列一阶差分方程的通解：

(1)$2y_{t+1}+10y_t-5t=0$；

(2)$y_{t+1}-y_t=4\cos\dfrac{\pi}{3}t$.

解　(1)方程可化简为

$$y_{t+1}+5y_t=\frac{5}{2}t,$$

即 $a=5,f(t)=\dfrac{5}{2}t$，则对应的齐次方程的通解为

$$y_A(t)=A(-5)^t.$$

非齐次方程的特解应具有的形式为

$$y_t^*=B_0+B_1t.$$

代入原方程后可求得 $B_0=-\frac{5}{72},B_1=\frac{5}{12}$. 于是原方程的通解为

$$y_t=A(-5)^t+\frac{5}{12}\left(t-\frac{1}{6}\right)(A\text{ 为任意常数}).$$

(2)设非齐次方程的特解具有形式

$$y_t^*=B_0\cos\frac{\pi}{3}t+B_1\sin\frac{\pi}{3}t.$$

代入原方程后可求得 $B_0=-2,B_1=2\sqrt{3}$. 于是原方程的通解为

$$y_t=A-2\cos\frac{\pi}{3}t+2\sqrt{3}\sin\frac{\pi}{3}t(A\text{ 为任意常数}).$$

基础习题精练

习 题

1.16.1 设某商品的最大需求量为 1 200 件，该商品的需求函数 $Q=Q(p)$，需求弹性 $\eta=\frac{p}{120-p}$ $(\eta>0)$，p 为单价(单位:万元).

(1)求需求函数的表达式；

(2)求 $p=100$ 万元时的边际收益，并说明其经济意义.

1.16.2 假设某企业在两个相互分割的市场上出售同一种产品，两个市场的需求函数分别是

$$p_1=18-2Q_1,\quad p_2=12-Q_2,$$

其中 p_1 和 p_2 分别表示该产品在两个市场的价格(单位:万元/吨)，Q_1 和 Q_2 表示该产品在两个市场的销售量(即需求量，单位:吨)，并且该企业生产这种产品的总成本函数是 $C=2Q+5$，其中 Q 表示该产品在两个市场的销售总量，即 $Q=Q_1+Q_2$.

(1)如果该企业实行价格差别策略，试确定两个市场上该产品的销售量和价格，使该企业获得最大利润；

(2)如果该企业实行价格无差别策略，试确定两个市场上该产品的销售量及其统一的价格，使该企业的总利润最大化，并比较两种价格策略下的总利润大小.

1.16.3 求方程 $y_{t+1}-3y_t=2^t-1,y_0=1$ 的特解.

解 答

1.16.1 解 (1)由题设 $$-\frac{p}{Q}Q'=\frac{p}{120-p},$$

所以 $\int \frac{\mathrm{d}Q}{Q} = -\int \frac{1}{120-p}\,\mathrm{d}p$，可得 $\ln Q=\ln(120-p)+\ln C$，即 $Q=C(120-p)$.

又最大需求量为 1 200，故 $C=10$，所以需求函数 $Q=1\,200-10p$.

(2)由(1)知，收益函数 $R=120Q-\frac{1}{10}Q^2$，边际收益 $R'(Q)=120-\frac{1}{5}Q$.

当 $p=100$ 时，$Q=200$，故当 $p=100$ 万元时的边际收益 $R'(200)=80$. 其经济意义：销售第 201 件商品所得的收益为 80 万元.

1.16.2　解　(1)根据题意，总利润函数为

$$L=R-C=p_1Q_1+p_2Q_2-(2Q+5)=-2Q_1^2-Q_2^2+16Q_1+10Q_2-5.$$

令
$$\begin{cases} L'_{Q_1}=-4Q_1+16=0, \\ L'_{Q_2}=-2Q_2+10=0, \end{cases}$$

解得 $Q_1=4, Q_2=5$，则 $p_1=10$ 万元/吨，$p_2=7$ 万元/吨.

因驻点 (4,5) 唯一，且实际问题一定存在最大值，故最大值必在驻点处达到，最大利润为

$$L=-2\times4^2-5^2+16\times4+10\times5-5=52(\text{万元}).$$

(2)若实行价格无差别策略，则 $p_1=p_2$，于是有约束条件 $2Q_1-Q_2=6$.

构造拉格朗日函数 $F(Q_1,Q_2,\lambda)=-2Q_1^2-Q_2^2+16Q_1+10Q_2-5+\lambda(2Q_1-Q_2-6)$.

令
$$\begin{cases} F'_{Q_1}=-4Q_1+16+2\lambda=0, \\ F'_{Q_2}=-2Q_2+10-\lambda=0, \\ F'_{\lambda}=2Q_1-Q_2-6=0, \end{cases}$$

解得 $Q_1=5, Q_2=4, \lambda=2$，则 $p_1=p_2=8$ 万元/吨.

最大利润 $L=-2\times5^2-4^2+16\times5+10\times4-5=49$ (万元).

由上述结果可知，企业实行差别定价所得总利润要大于统一价格的总利润.

1.16.3　解　由于 $y_{t+1}-3y_t=2^t$ 的特解为 -2^t，而 $y_{t+1}-3y_t=1$ 的特解为 $-\frac{1}{2}$. 由叠加原理可得，原方程的通解为

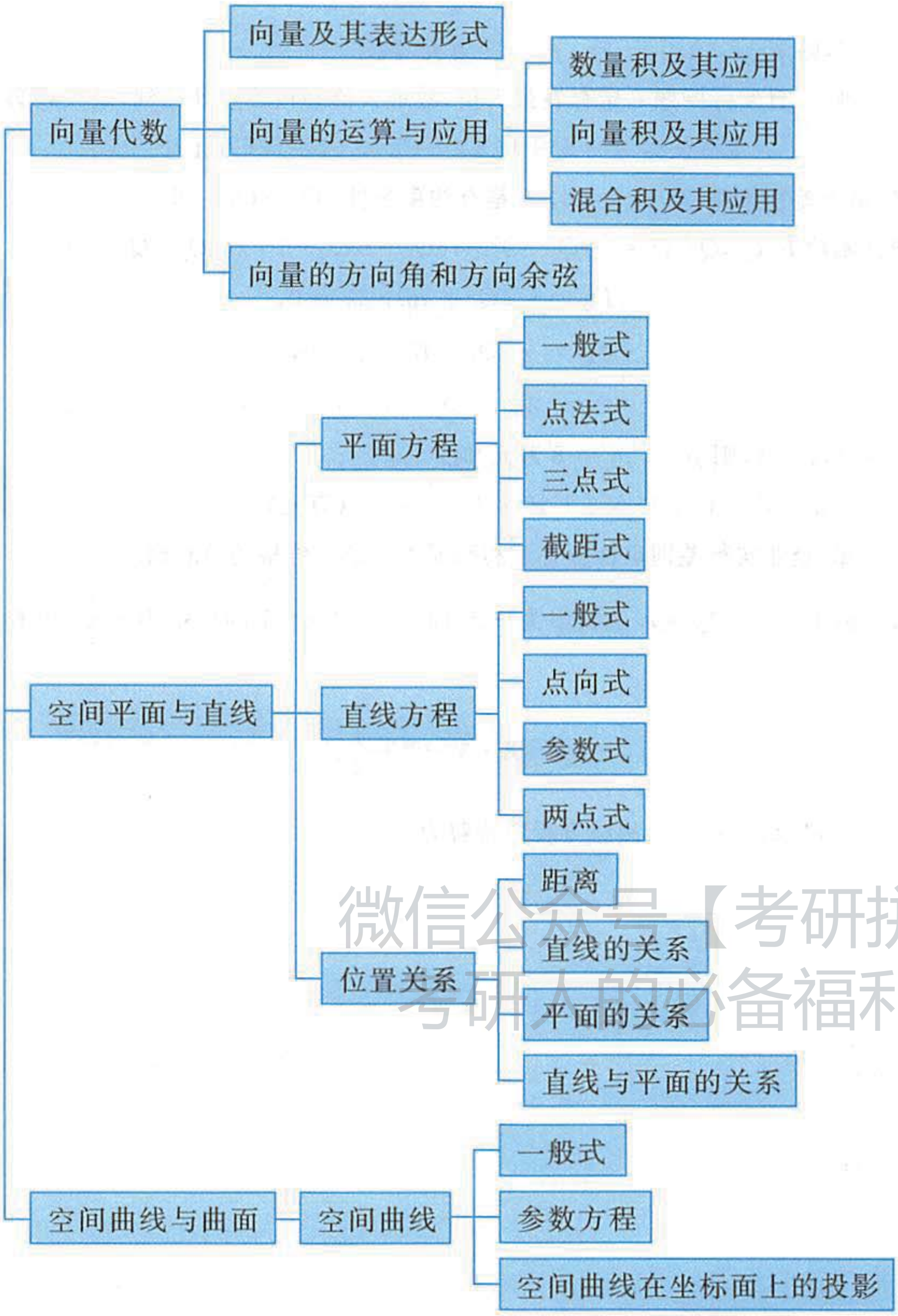
$$y_t=A\cdot3^t-2^t+\frac{1}{2}.$$

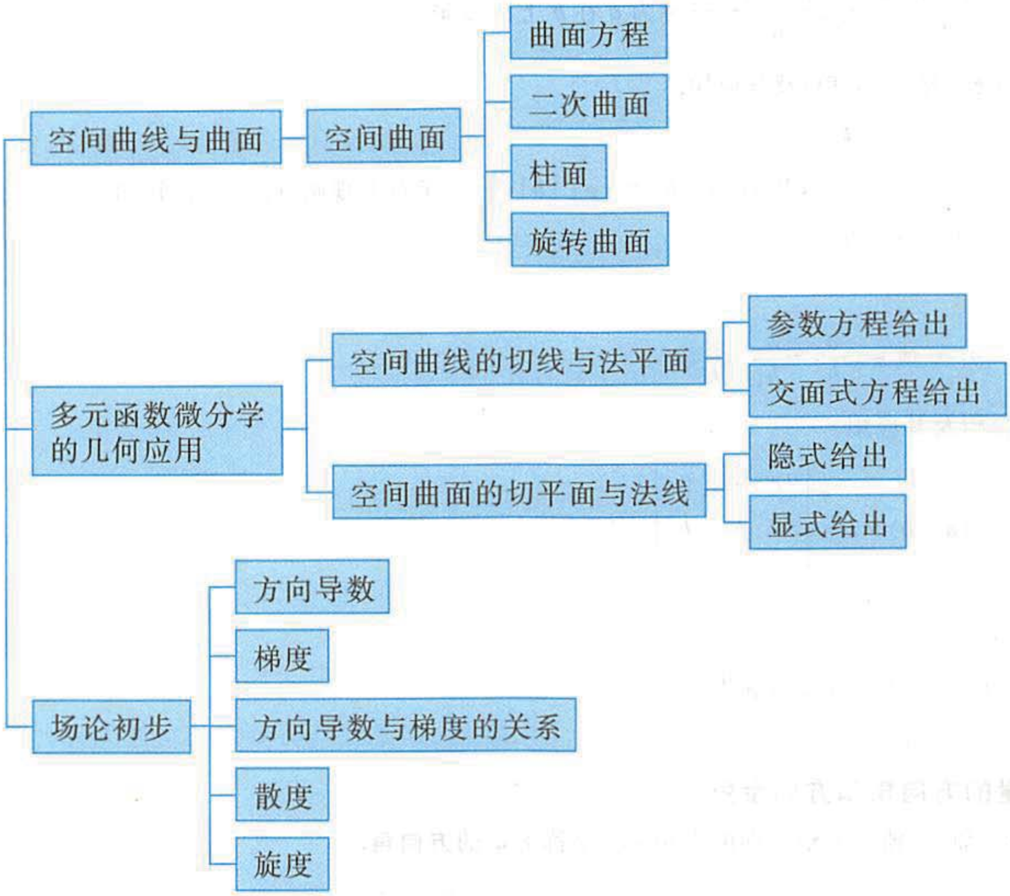
代入定解条件 $y_0=1$，可求得 $A=\frac{3}{2}$，故原方程的特解为

$$y_t=\frac{1}{2}\cdot3^{t+1}-2^t+\frac{1}{2}.$$

第17讲 多元函数积分学的基础知识（仅数学一要求）

基础知识结构





基础内容精讲

一、向量代数

1. 向量及其表达形式

既有大小又有方向的量称为**向量**.

【注】 两个向量，只要它们的大小相等、方向相同，它们就是相等的向量，与它们在空间中的位置无关（这也称为**向量的自由性**）.

向量的表达形式为 $\boldsymbol{a}=(a_x,a_y,a_z)=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$.

2. 向量的运算及其应用

设 $\boldsymbol{a}=(a_x,a_y,a_z)$，$\boldsymbol{b}=(b_x,b_y,b_z)$，$\boldsymbol{c}=(c_x,c_y,c_z)$，$\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 均不是零向量.

（1）数量积（内积、点积）及其应用.

①$\boldsymbol{a}\cdot\boldsymbol{b}=(a_x,a_y,a_z)\cdot(b_x,b_y,b_z)=a_xb_x+a_yb_y+a_zb_z$.

②$\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta$，则 $\cos\theta=\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}=\dfrac{a_xb_x+a_yb_y+a_zb_z}{\sqrt{a_x^2+a_y^2+a_z^2}\cdot\sqrt{b_x^2+b_y^2+b_z^2}}$，其中 θ 为 $\boldsymbol{a},\boldsymbol{b}$ 的夹角.

③$\boldsymbol{a}\perp\boldsymbol{b}\Leftrightarrow\theta=\dfrac{\pi}{2}\Leftrightarrow\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta=0\Leftrightarrow a_xb_x+a_yb_y+a_zb_z=0$.

④$\mathrm{Prj}_{\boldsymbol{b}}\boldsymbol{a}=\dfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{b}|}=\dfrac{a_xb_x+a_yb_y+a_zb_z}{\sqrt{b_x^2+b_y^2+b_z^2}}$称为 $\boldsymbol{a}$ 在 $\boldsymbol{b}$ 上的**投影**.

(2) 向量积(外积、叉积)及其应用.

①$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$,其中 $|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\theta$,用右手规则确定方向(转向角不超过 π),θ 为 $\boldsymbol{a}$,$\boldsymbol{b}$ 的夹角.

②$\boldsymbol{a}/\!/\boldsymbol{b}\Leftrightarrow\theta=0$ 或 $\pi\Leftrightarrow\dfrac{a_x}{b_x}=\dfrac{a_y}{b_y}=\dfrac{a_z}{b_z}$.

(3) 混合积及其应用.

①$[\boldsymbol{abc}]=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}$.

②$\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}=0\Leftrightarrow$三向量共面.

3. 向量的方向角和方向余弦

(1)$\boldsymbol{a}$ 与 x 轴、y 轴和 z 轴正向的夹角 α,β,γ 称为 $\boldsymbol{a}$ 的**方向角**.

(2)$\cos\alpha,\cos\beta,\cos\gamma$ 称为 $\boldsymbol{a}$ 的**方向余弦**,且 $\cos\alpha=\dfrac{a_x}{|\boldsymbol{a}|}$,$\cos\beta=\dfrac{a_y}{|\boldsymbol{a}|}$,$\cos\gamma=\dfrac{a_z}{|\boldsymbol{a}|}$.

(3)$\boldsymbol{a}^{\circ}=\dfrac{\boldsymbol{a}}{|\boldsymbol{a}|}=(\cos\alpha,\cos\beta,\cos\gamma)$称为向量 $\boldsymbol{a}$ 的**单位向量**(表示方向的向量).

(4)任意向量 $\boldsymbol{r}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}=(r\cos\alpha,r\cos\beta,r\cos\gamma)=r(\cos\alpha,\cos\beta,\cos\gamma)$,其中 $\cos\alpha,\cos\beta,\cos\gamma$ 为 $\boldsymbol{r}$ 的方向余弦,r 为 $\boldsymbol{r}$ 的模,$\cos\alpha=\dfrac{x}{\sqrt{x^2+y^2+z^2}}$,$\cos\beta=\dfrac{y}{\sqrt{x^2+y^2+z^2}}$,$\cos\gamma=\dfrac{z}{\sqrt{x^2+y^2+z^2}}$,$r=\sqrt{x^2+y^2+z^2}$.

二、空间平面与直线

1. 平面方程

以下假设平面的法向量 $\boldsymbol{n}=(A,B,C)$.

①一般式:$Ax+By+Cz+D=0$.

②点法式:$A(x-x_0)+B(y-y_0)+C(z-z_0)=0$.

③三点式:$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x-x_2 & y-y_2 & z-z_2 \\ x-x_3 & y-y_3 & z-z_3 \end{vmatrix}=0$(平面过不共线的三点 $P_i(x_i,y_i,z_i)$,$i=1,2,3$).

④截距式:$\dfrac{x}{a}+\dfrac{y}{b}+\dfrac{z}{c}=1$(平面过$(a,0,0)$,$(0,b,0)$,$(0,0,c)$三点).

2. 直线方程

以下假设直线的方向向量 $\boldsymbol{\tau}=(l,m,n)$.

①一般式：$\begin{cases} A_1x+B_1y+C_1z+D_1=0, \boldsymbol{n}_1=(A_1,B_1,C_1), \\ A_2x+B_2y+C_2z+D_2=0, \boldsymbol{n}_2=(A_2,B_2,C_2), \end{cases}$ 其中 $\boldsymbol{n}_1$ 不平行于 $\boldsymbol{n}_2$.

【注】 其几何背景很直观，是两个平面的交线，且该直线的方向向量 $\boldsymbol{\tau}=\boldsymbol{n}_1\times\boldsymbol{n}_2$.

②点向式（标准式、对称式）：$\dfrac{x-x_0}{l}=\dfrac{y-y_0}{m}=\dfrac{z-z_0}{n}$.

③参数式：$\begin{cases} x=x_0+lt, \\ y=y_0+mt, \\ z=z_0+nt, \end{cases}$ $M(x_0,y_0,z_0)$为直线上的已知点，t 为参数.

④两点式：$\dfrac{x-x_1}{x_2-x_1}=\dfrac{y-y_1}{y_2-y_1}=\dfrac{z-z_1}{z_2-z_1}$（直线过不同的两点 $P_i(x_i,y_i,z_i)$，$i=1,2$）.

3. 位置关系

（1）距离.

点 $P_0(x_0,y_0,z_0)$到平面 $Ax+By+Cz+D=0$ 的距离 $d=\dfrac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}$.

（2）直线间的关系.

设 $\boldsymbol{\tau}_1=(l_1,m_1,n_1)$，$\boldsymbol{\tau}_2=(l_2,m_2,n_2)$分别为直线 L_1，L_2 的方向向量.

①$L_1\perp L_2 \Leftrightarrow \boldsymbol{\tau}_1\perp\boldsymbol{\tau}_2 \Leftrightarrow l_1l_2+m_1m_2+n_1n_2=0$.

②$L_1 /\!/ L_2 \Leftrightarrow \boldsymbol{\tau}_1 /\!/ \boldsymbol{\tau}_2 \Leftrightarrow \dfrac{l_1}{l_2}=\dfrac{m_1}{m_2}=\dfrac{n_1}{n_2}$.

（3）平面间的关系.

设平面 π_1，π_2 的法向量分别为 $\boldsymbol{n}_1=(A_1,B_1,C_1)$，$\boldsymbol{n}_2=(A_2,B_2,C_2)$.

①$\pi_1\perp\pi_2 \Leftrightarrow \boldsymbol{n}_1\perp\boldsymbol{n}_2 \Leftrightarrow A_1A_2+B_1B_2+C_1C_2=0$.

②$\pi_1 /\!/ \pi_2 \Leftrightarrow \boldsymbol{n}_1 /\!/ \boldsymbol{n}_2 \Leftrightarrow \dfrac{A_1}{A_2}=\dfrac{B_1}{B_2}=\dfrac{C_1}{C_2}$.

（4）平面与直线的关系.

设直线 L 的方向向量为 $\boldsymbol{\tau}=(l,m,n)$，平面 π 的法向量为 $\boldsymbol{n}=(A,B,C)$.

①$L\perp\pi \Leftrightarrow \boldsymbol{\tau} /\!/ \boldsymbol{n} \Leftrightarrow \dfrac{A}{l}=\dfrac{B}{m}=\dfrac{C}{n}$.

②$L /\!/ \pi \Leftrightarrow \boldsymbol{\tau}\perp\boldsymbol{n} \Leftrightarrow Al+Bm+Cn=0$.

三、空间曲线与曲面

1. 空间曲线

（1）一般式 Γ：$\begin{cases} F(x,y,z)=0, \\ G(x,y,z)=0, \end{cases}$ 其几何背景为两个曲面的交线.

（2）参数方程 Γ：$\begin{cases} x=\varphi(t), \\ y=\psi(t), \\ z=\omega(t), \end{cases}$ $t\in[\alpha,\beta]$.

(3)空间曲线在坐标面上的投影(**重点**).

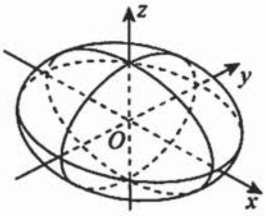
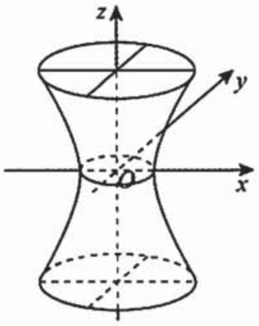
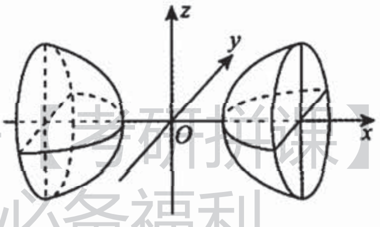
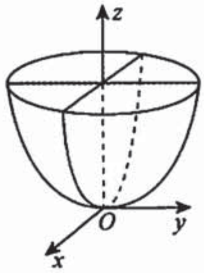
以求曲线 Γ 在 xOy 平面上的投影曲线为例. 将 Γ: $\begin{cases}F(x,y,z)=0,\\G(x,y,z)=0\end{cases}$ 中的 z 消去,得到 $\varphi(x,y)=0$,则曲线 Γ 在 xOy 面上的投影曲线包含于曲线 $\begin{cases}\varphi(x,y)=0,\\z=0.\end{cases}$

曲线 Γ 在其他平面上的投影曲线可类似求得.

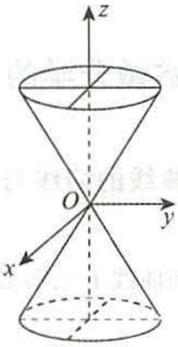
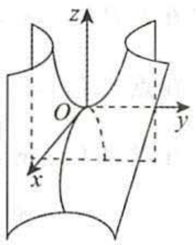
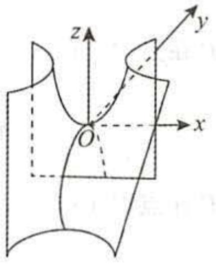
2. 空间曲面

(1)曲面方程:$F(x,y,z)=0$.

(2)二次曲面.

曲面名称	方程	图形
椭球面	$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$	
单叶双曲面	$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	
双叶双曲面	$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$	
椭圆抛物面	$\frac{x^2}{2p}+\frac{y^2}{2q}=z(p,q>0)$	

续表

曲面名称	方程	图形
椭圆锥面	$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{z^2}{c^2}$	
双曲抛物面（马鞍面）	$-\frac{x^2}{2p}+\frac{y^2}{2q}=z(p,q>0)$	
	$z=xy$	

（3）柱面：动直线沿定曲线平行移动所形成的曲面.

椭圆柱面　$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$（当 $a=b$ 时为圆柱面）.

双曲柱面　$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$.

抛物柱面　$y=ax^2$.

【注】　在空间解析几何中，一般认为缺少变量的方程为柱面.

（4）旋转曲面（重点）：曲线 Γ 绕一条定直线旋转一周所形成的曲面.

曲线 Γ：$\begin{cases}F(x,y,z)=0,\\G(x,y,z)=0\end{cases}$ 绕直线 L：$\frac{x-x_0}{m}=\frac{y-y_0}{n}=\frac{z-z_0}{p}$ 旋转形成一个旋转曲面，旋转曲面方程的求法如下.

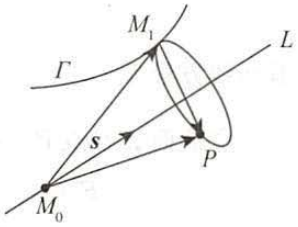


图 1-17-1

如图 1-17-1 所示，已知点 $M_0(x_0,y_0,z_0)$，方向向量 $\boldsymbol{s}=(m,n,p)$. 在母线 Γ 上任取一点 $M_1(x_1,y_1,z_1)$，则过 M_1 的纬圆上任意一点 $P(x,y,z)$ 满足条件

$$\overrightarrow{M_1P}\perp\boldsymbol{s},\quad|\overrightarrow{M_0P}|=|\overrightarrow{M_0M_1}|,$$

即
$$\begin{cases} m(x-x_1)+n(y-y_1)+p(z-z_1)=0, \\ (x-x_0)^2+(y-y_0)^2+(z-z_0)^2=(x_1-x_0)^2+(y_1-y_0)^2+(z_1-z_0)^2, \end{cases}$$

与方程 $F(x_1,y_1,z_1)=0$ 和 $G(x_1,y_1,z_1)=0$ 联立消去 x_1,y_1,z_1，便可得到旋转曲面的方程.

四、多元函数微分学的几何应用

1. 空间曲线的切线与法平面

(1)设空间曲线 Γ 由参数方程 $\begin{cases} x=\varphi(t), \\ y=\psi(t), \\ z=\omega(t) \end{cases}$ 给出，其中 $\varphi(t),\psi(t),w(t)$ 均可导，$P_0(x_0,y_0,z_0)$ 是 Γ 上的点，且当 $t=t_0$ 时，$\varphi'(t_0),\psi'(t_0),\omega'(t_0)$ 都不为 0，则

①曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的切向量为 $\boldsymbol{\tau}=(\varphi'(t_0),\psi'(t_0),\omega'(t_0))$.

②曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的切线方程为 $\dfrac{x-x_0}{\varphi'(t_0)}=\dfrac{y-y_0}{\psi'(t_0)}=\dfrac{z-z_0}{\omega'(t_0)}$.

③曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的法平面(过点 $P_0(x_0,y_0,z_0)$ 且与切线垂直的平面)方程为

$$\varphi'(t_0)(x-x_0)+\psi'(t_0)(y-y_0)+\omega'(t_0)(z-z_0)=0.$$

(2)设空间曲线 Γ 由交面式方程 $\begin{cases} F(x,y,z)=0, \\ G(x,y,z)=0 \end{cases}$ 给出，则在以下表达式有意义的条件下，有

①曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的切向量为

$$\boldsymbol{\tau}=\left(\begin{vmatrix} F_y' & F_z' \\ G_y' & G_z' \end{vmatrix}_{P_0},\begin{vmatrix} F_z' & F_x' \\ G_z' & G_x' \end{vmatrix}_{P_0},\begin{vmatrix} F_x' & F_y' \\ G_x' & G_y' \end{vmatrix}_{P_0}\right).$$

②曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的切线方程为

$$\frac{x-x_0}{\begin{vmatrix} F_y' & F_z' \\ G_y' & G_z' \end{vmatrix}_{P_0}}=\frac{y-y_0}{\begin{vmatrix} F_z' & F_x' \\ G_z' & G_x' \end{vmatrix}_{P_0}}=\frac{z-z_0}{\begin{vmatrix} F_x' & F_y' \\ G_x' & G_y' \end{vmatrix}_{P_0}}.$$

③曲线 Γ 在点 $P_0(x_0,y_0,z_0)$ 处的法平面方程为

$$\begin{vmatrix} F_y' & F_z' \\ G_y' & G_z' \end{vmatrix}_{P_0}(x-x_0)+\begin{vmatrix} F_z' & F_x' \\ G_z' & G_x' \end{vmatrix}_{P_0}(y-y_0)+\begin{vmatrix} F_x' & F_y' \\ G_x' & G_y' \end{vmatrix}_{P_0}(z-z_0)=0.$$

2. 空间曲面的切平面与法线

(1)设空间曲面 Σ 由方程 $F(x,y,z)=0$ 给出，$P_0(x_0,y_0,z_0)$ 是 Σ 上的点，则

①曲面 Σ 在点 $P_0(x_0,y_0,z_0)$ 处的法向量(垂直于该点切平面的向量)为

$$\boldsymbol{n}=(F_x'(x_0,y_0,z_0),F_y'(x_0,y_0,z_0),F_z'(x_0,y_0,z_0)),$$

且法线方程为
$$\frac{x-x_0}{F_x'(x_0,y_0,z_0)}=\frac{y-y_0}{F_y'(x_0,y_0,z_0)}=\frac{z-z_0}{F_z'(x_0,y_0,z_0)}.$$

②曲面 Σ 在点 $P_0(x_0,y_0,z_0)$ 处的切平面方程为

$$F_x'(x_0,y_0,z_0)(x-x_0)+F_y'(x_0,y_0,z_0)(y-y_0)+F_z'(x_0,y_0,z_0)(z-z_0)=0.$$

(2)设空间曲面 Σ 由方程 $z=f(x,y)$ 给出，令 $F(x,y,z)=f(x,y)-z$，则

①曲面 Σ 在点 $P_0(x_0,y_0,z_0)$ 处的法向量为

$$\boldsymbol{n}=(f_x'(x_0,y_0),f_y'(x_0,y_0),-1),$$

得法线方程为

$$\frac{x-x_0}{f'_x(x_0,y_0)}=\frac{y-y_0}{f'_y(x_0,y_0)}=\frac{z-z_0}{-1}.$$

②曲面 Σ 在点 $P_0(x_0,y_0,z_0)$ 处的切平面方程为

$$f'_x(x_0,y_0)(x-x_0)+f'_y(x_0,y_0)(y-y_0)-(z-z_0)=0.$$

【注】 若用 α,β,γ 表示曲面 $z=f(x,y)$ 在点 (x_0,y_0,z_0) 处的法向量的方向角，并假定法向量的方向是向上的，即它与 z 轴正向所成的角 γ 是锐角，则法向量的**方向余弦**为

$$\cos\alpha=\frac{-f_x}{\sqrt{1+f_x^2+f_y^2}},\quad \cos\beta=\frac{-f_y}{\sqrt{1+f_x^2+f_y^2}},\quad \cos\gamma=\frac{1}{\sqrt{1+f_x^2+f_y^2}},$$

其中，$f_x=f'_x(x_0,y_0)$，$f_y=f'_y(x_0,y_0)$.

五、场论初步

1. 方向导数

在许多问题中，不仅要知道函数在坐标轴方向上的变化率(即偏导数)，而且还要设法求得函数在某点沿着其他特定方向上的变化率. 这就是本部分所要讨论的方向导数.

定义 1 设三元函数 $u=u(x,y,z)$ 在点 $P_0(x_0,y_0,z_0)$ 的某空间邻域 $U\subset\mathbf{R}^3$ 内有定义，l 为从点 P_0 出发的射线，$P(x,y,z)$ 为 l 上且在 U 内的任一点，则

$$\begin{cases}x-x_0=\Delta x=t\cos\alpha,\\ y-y_0=\Delta y=t\cos\beta,\\ z-z_0=\Delta z=t\cos\gamma.\end{cases}$$

以 $t=\sqrt{(\Delta x)^2+(\Delta y)^2+(\Delta z)^2}$ 表示 P 与 P_0 之间的距离，如图 1-17-2 所示，若极限

$$\lim_{t\to0^+}\frac{u(P)-u(P_0)}{t}=\lim_{t\to0^+}\frac{u(x_0+t\cos\alpha,y_0+t\cos\beta,z_0+t\cos\gamma)-u(x_0,y_0,z_0)}{t}$$

存在，则称此极限为函数 $u=u(x,y,z)$ 在点 P_0 沿方向 l 的**方向导数**，记作 $\left.\frac{\partial u}{\partial l}\right|_{P_0}$.

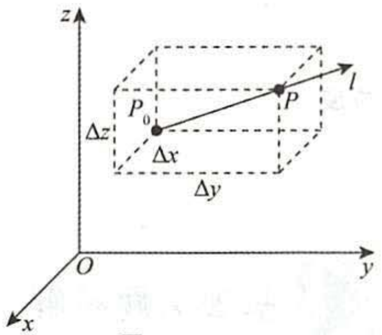


图 1-17-2

定理(方向导数的计算公式) 设三元函数 $u=u(x,y,z)$ 在点 $P_0(x_0,y_0,z_0)$ 处可微分，则 $u=u(x,y,z)$ 在点 P_0 处沿任一方向 l 的方向导数都存在，且

$$\left.\frac{\partial u}{\partial l}\right|_{P_0}=u'_x(P_0)\cos\alpha+u'_y(P_0)\cos\beta+u'_z(P_0)\cos\gamma,$$

其中 $\cos\alpha,\cos\beta,\cos\gamma$ 为方向 l 的方向余弦.

【注】 对于二元函数 $f(x,y)$ 的情况与三元函数类似.

2. 梯度

在一个数量场中，函数在给定点处沿不同的方向，其方向导数一般是不相同的，现在我们所关心的是沿哪一个方向其方向导数最大？最大值是多少？函数在点 P 沿哪一方向增加的速度最快？为此引进一个很重要的概念——梯度.

定义 2 设三元函数 $u=u(x,y,z)$ 在点 $P_0(x_0,y_0,z_0)$ 处具有一阶偏导数，则定义

$$\mathbf{grad}\ u\Big|_{P_0}=(u'_x(P_0),u'_y(P_0),u'_z(P_0))$$

为函数 $u=u(x,y,z)$ 在点 P_0 处的梯度.

3. 方向导数与梯度的关系

由方向导数的计算公式 $\dfrac{\partial u}{\partial l}\Big|_{P_0}=u'_x(P_0)\cos\alpha+u'_y(P_0)\cos\beta+u'_z(P_0)\cos\gamma$ 与梯度的定义

$$\mathbf{grad}\ u\Big|_{P_0}=(u'_x(P_0),u'_y(P_0),u'_z(P_0)),$$

可得到
$$\begin{aligned}\frac{\partial u}{\partial l}\Big|_{P_0}&=(u'_x(P_0),u'_y(P_0),u'_z(P_0))\cdot(\cos\alpha,\cos\beta,\cos\gamma)=\mathbf{grad}\ u\Big|_{P_0}\cdot \boldsymbol{l}^{\circ}\\&=\left|\mathbf{grad}\ u\Big|_{P_0}\right|\,|\boldsymbol{l}^{\circ}|\cos\theta=\left|\mathbf{grad}\ u\Big|_{P_0}\right|\cos\theta,\end{aligned}$$

其中 θ 为 $\mathbf{grad}\ u\Big|_{P_0}$ 与 $\boldsymbol{l}^{\circ}$ 的夹角，当 $\cos\theta=1$ 时，$\dfrac{\partial u}{\partial l}\Big|_{P_0}$ 有最大值.

结论 函数在某点的梯度是一个向量，它的方向与取得最大方向导数的方向一致，而它的模为方向导数的最大值.

4. 散度与旋度

设向量场 $\boldsymbol{A}(x,y,z)=(P(x,y,z),Q(x,y,z),R(x,y,z))$，则

散度
$$\mathrm{div}\ \boldsymbol{A}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z},$$

旋度
$$\mathbf{rot}\ \boldsymbol{A}=\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\\dfrac{\partial}{\partial x}&\dfrac{\partial}{\partial y}&\dfrac{\partial}{\partial z}\\P&Q&R\end{vmatrix}.$$

基础例题精解

例 1.17.1 设 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=2$，则 $[(\boldsymbol{a}+\boldsymbol{b})\times(\boldsymbol{b}+\boldsymbol{c})]\cdot(\boldsymbol{c}+\boldsymbol{a})=$________.

解 应填 4.

$$\begin{aligned}[(\boldsymbol{a}+\boldsymbol{b})\times(\boldsymbol{b}+\boldsymbol{c})]\cdot(\boldsymbol{c}+\boldsymbol{a})&=[\boldsymbol{a}\times(\boldsymbol{b}+\boldsymbol{c})+\boldsymbol{b}\times(\boldsymbol{b}+\boldsymbol{c})]\cdot(\boldsymbol{c}+\boldsymbol{a})\\&=(\boldsymbol{a}\times\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c})\cdot(\boldsymbol{c}+\boldsymbol{a})\\&=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}+(\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}=2(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=4.\end{aligned}$$

例 1.17.2 设直线 $L:\begin{cases}x+y+b=0,\\x+ay-z-3=0\end{cases}$ 在平面 π 上，而平面 π 与曲面 $z=x^2+y^2$ 相切于点 $(1,-2,5)$，求 a,b 的值.

分析 首先求出与曲面 $z=x^2+y^2$ 相切于点 $(1,-2,5)$ 的切平面 π 的方程，进一步可将 L 代入切平面 π 的方程，求出 a,b 的值.

解 在点 $(1,-2,5)$ 处曲面的法向量为 $\boldsymbol{n}=(2,-4,-1)$，故切平面 π 的方程为

$$2(x-1)-4(y+2)-(z-5)=0,$$

即
$$2x-4y-z-5=0. \qquad (*)$$

由直线方程 $L:\begin{cases}x+y+b=0,\\ x+ay-z-3=0,\end{cases}$ 解得 $y=-x-b,z=x-3-a(x+b)$. 将其代入(*)式,得

$$(5+a)x+4b+ab-2=0,$$

因而有

$$5+a=0,\quad 4b+ab-2=0,$$

解得

$$a=-5,\quad b=-2.$$

例 1.17.3 求空间曲线 $L:\begin{cases}x^2+y^2+z^2=2,\\ z=x^2+y^2\end{cases}$ 在三个坐标平面上的投影曲线及投影柱面.

解 ①消去变量 z,得到

$$(x^2+y^2)^2+(x^2+y^2)-2=0,$$

即

$$(x^2+y^2-1)(x^2+y^2+2)=0,$$

所以曲线 L 沿 z 轴的投影柱面为圆柱面:$x^2+y^2=1$,在 xOy 坐标面上的投影曲线为圆:$\begin{cases}x^2+y^2=1,\\ z=0.\end{cases}$

②消去变量 y,得到 $z^2+z-2=0$,即 $z=1(z=-2$ 舍去),且由①可知 x 的变化范围为 $-1\leqslant x\leqslant 1$,所以曲线 L 沿 y 轴的投影柱面为平面块:$z=1(-1\leqslant x\leqslant 1)$.

在 xOz 坐标面上的投影曲线为直线段:$\begin{cases}z=1(-1\leqslant x\leqslant 1),\\ y=0.\end{cases}$

③同理可得曲线 L 沿 x 轴的投影柱面为平面块:$z=1(-1\leqslant y\leqslant 1)$.

在 yOz 坐标面上的投影曲线为直线段:$\begin{cases}z=1(-1\leqslant y\leqslant 1),\\ x=0.\end{cases}$

> **【注】** 本题解法给出了投影曲线及投影柱面的求法,并且揭示了投影曲线及投影柱面可能是曲线段和柱面块.

例 1.17.4 求直线 $L:\begin{cases}x-y+2z-1=0,\\ x-3y-2z+1=0\end{cases}$ 绕 y 轴旋转一周所形成的曲面方程.

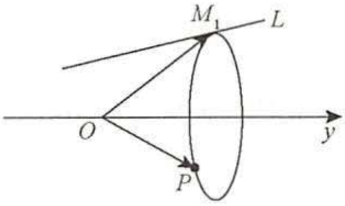


图 1-17-3

解 如图 1-17-3 所示,在直线 L 上任取一点 $M_1(x_1,y_1,z_1)$,则过 M_1 的纬圆上任一点 $P(x,y,z)$满足条件 $|\overrightarrow{OM_1}|=|\overrightarrow{OP}|$,且 $y=y_1$.

于是由 $x^2+y^2+z^2=x_1^2+y_1^2+z_1^2$,得 $x^2+z^2=x_1^2+z_1^2$.

由 L 的方程解出 $x_1=2y_1,z_1=-\frac{1}{2}(y_1-1)$,即 $x_1=2y,z_1=-\frac{1}{2}(y-1)$. 于是旋转曲面方程为

$$x^2+z^2=(2y)^2+\left[-\frac{1}{2}(y-1)\right]^2 \text{ 或 } 4x^2-17y^2+4z^2+2y-1=0.$$

例 1.17.5 求空间曲线 $\Gamma:\begin{cases}x=\int_0^t e^u\cos u\mathrm{d}u,\\ y=2\sin t+\cos t,\\ z=1+e^{3t}\end{cases}$ 在 $t=0$ 处的切线方程和法平面方程.

解 当 $t=0$ 时,$x=0,y=1,z=2;x'=e^t\cos t,y'=2\cos t-\sin t,z'=3e^{3t}$,则 $x'(0)=1,y'(0)=2,z'(0)=3$,于是,切线方程为$\frac{x-0}{1}=\frac{y-1}{2}=\frac{z-2}{3}$,法平面方程为 $x+2(y-1)+3(z-2)=0$,即 $x+2y+3z-8=0$.

例 1.17.6 求函数 $z=xe^{2y}$在点 $P(1,0)$处沿点 $P(1,0)$指向点 $Q(2,-1)$方向的方向导数.

解 由题意可知方向 l 即为 $\overrightarrow{PQ}=(1,-1)$，故 x 轴到方向 l 的转角 $\varphi=-\frac{\pi}{4}$. 因为

$$\left.\frac{\partial z}{\partial x}\right|_{(1,0)}=\left.\mathrm{e}^{2y}\right|_{(1,0)}=1,\quad \left.\frac{\partial z}{\partial y}\right|_{(1,0)}=\left.2x\mathrm{e}^{2y}\right|_{(1,0)}=2,$$

故所求方向导数 $\left.\frac{\partial z}{\partial l}\right|_{(1,0)}=\cos\left(-\frac{\pi}{4}\right)+2\sin\left(-\frac{\pi}{4}\right)=-\frac{\sqrt{2}}{2}$.

例 1.17.7 函数 $u=\ln(x^2+y^2+z^2)$ 在点 $M(1,2,-2)$ 处的梯度 $\left.\mathbf{grad}\ u\right|_M=$________.

解 应填 $\frac{2}{9}(1,2,-2)$.

由已知，得 $\left.u'_x\right|_M=\frac{2}{9}$，$\left.u'_y\right|_M=\frac{4}{9}$，$\left.u'_z\right|_M=-\frac{4}{9}$，则 $\left.\mathbf{grad}\ u\right|_M=\left(\left.u'_x\right|_M,\left.u'_y\right|_M,\left.u'_z\right|_M\right)=\frac{2}{9}(1,2,-2)$.

例 1.17.8 向量场 $\boldsymbol{u}(x,y,z)=xy^2\boldsymbol{i}+y\mathrm{e}^z\boldsymbol{j}+x\ln(1+z^2)\boldsymbol{k}$ 在点 $P(1,1,0)$ 处的散度 $\mathrm{div}\ \boldsymbol{u}=$________.

解 应填 2.

由散度计算公式有

$$\mathrm{div}\ \boldsymbol{u}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}，其中\ \boldsymbol{u}=P\boldsymbol{i}+Q\boldsymbol{j}+R\boldsymbol{k}，$$

得

$$\left.\mathrm{div}\ \boldsymbol{u}\right|_P=\left.\left(y^2+\mathrm{e}^z+\frac{2xz}{1+z^2}\right)\right|_{(1,1,0)}=1+1=2.$$

例 1.17.9 设 $\boldsymbol{F}(x,y,z)=xy\boldsymbol{i}-yz\boldsymbol{j}+zx\boldsymbol{k}$，则 $\mathbf{rot}\ \boldsymbol{F}(1,1,0)=$________.

解 应填 $\boldsymbol{i}-\boldsymbol{k}$.

三元向量函数 $\boldsymbol{F}(x,y,z)=(P,Q,R)$，则

$$\mathbf{rot}\ \boldsymbol{F}(x,y,z)=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ P & Q & R\end{vmatrix},$$

这里 $P=xy,Q=-yz,R=zx$，于是

$$\mathbf{rot}\ \boldsymbol{F}(1,1,0)=\left.\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ xy & -yz & zx\end{vmatrix}\right|_{(1,1,0)}=\left.(y\boldsymbol{i}-z\boldsymbol{j}-x\boldsymbol{k})\right|_{(1,1,0)}$$

$$=\boldsymbol{i}-\boldsymbol{k}.$$

基础习题精练

习 题

1.17.1 求旋转抛物面 $z=x^2+y^2-1$ 在点 $(2,1,4)$ 处的切平面及法线方程.

1.17.2 求曲面 $z-\mathrm{e}^z+2xy=3$ 在点 $(1,2,0)$ 处的切平面及法线方程.

1.17.3　函数 $u=\ln(x+\sqrt{y^2+z^2})$ 在点 $A(1,0,1)$ 处沿点 A 指向点 $B(3,-2,2)$ 方向的方向导数为________.

解答

1.17.1　**解**　令 $F(x,y,z)=x^2+y^2-z-1$，法向量 $\boldsymbol{n}\Big|_{(2,1,4)}=(2x,2y,-1)\Big|_{(2,1,4)}=(4,2,-1)$.

切平面方程为 $4(x-2)+2(y-1)-(z-4)=0$，即 $4x+2y-z-6=0$.

法线方程为
$$\frac{x-2}{4}=\frac{y-1}{2}=\frac{z-4}{-1}.$$

1.17.2　**解**　令 $F(x,y,z)=z-\mathrm{e}^z+2xy-3$，则

$$F_x'\Big|_{(1,2,0)}=2y\Big|_{(1,2,0)}=4,\quad F_y'\Big|_{(1,2,0)}=2x\Big|_{(1,2,0)}=2,\quad F_z'\Big|_{(1,2,0)}=1-\mathrm{e}^z\Big|_{(1,2,0)}=0.$$

则切平面方程为 $4(x-1)+2(y-2)+0\cdot(z-0)=0$，即 $2x+y-4=0$.

法线方程为
$$\frac{x-1}{2}=\frac{y-2}{1}=\frac{z-0}{0}.$$

1.17.3　$\frac{1}{2}$　**解**　由公式得 $\frac{\partial u}{\partial \boldsymbol{l}}\Big|_A=\frac{\partial u}{\partial x}\Big|_A\cos\alpha+\frac{\partial u}{\partial y}\Big|_A\cos\beta+\frac{\partial u}{\partial z}\Big|_A\cos\gamma$，方向 $\boldsymbol{l}=\overrightarrow{AB}=(2,-2,1)$，

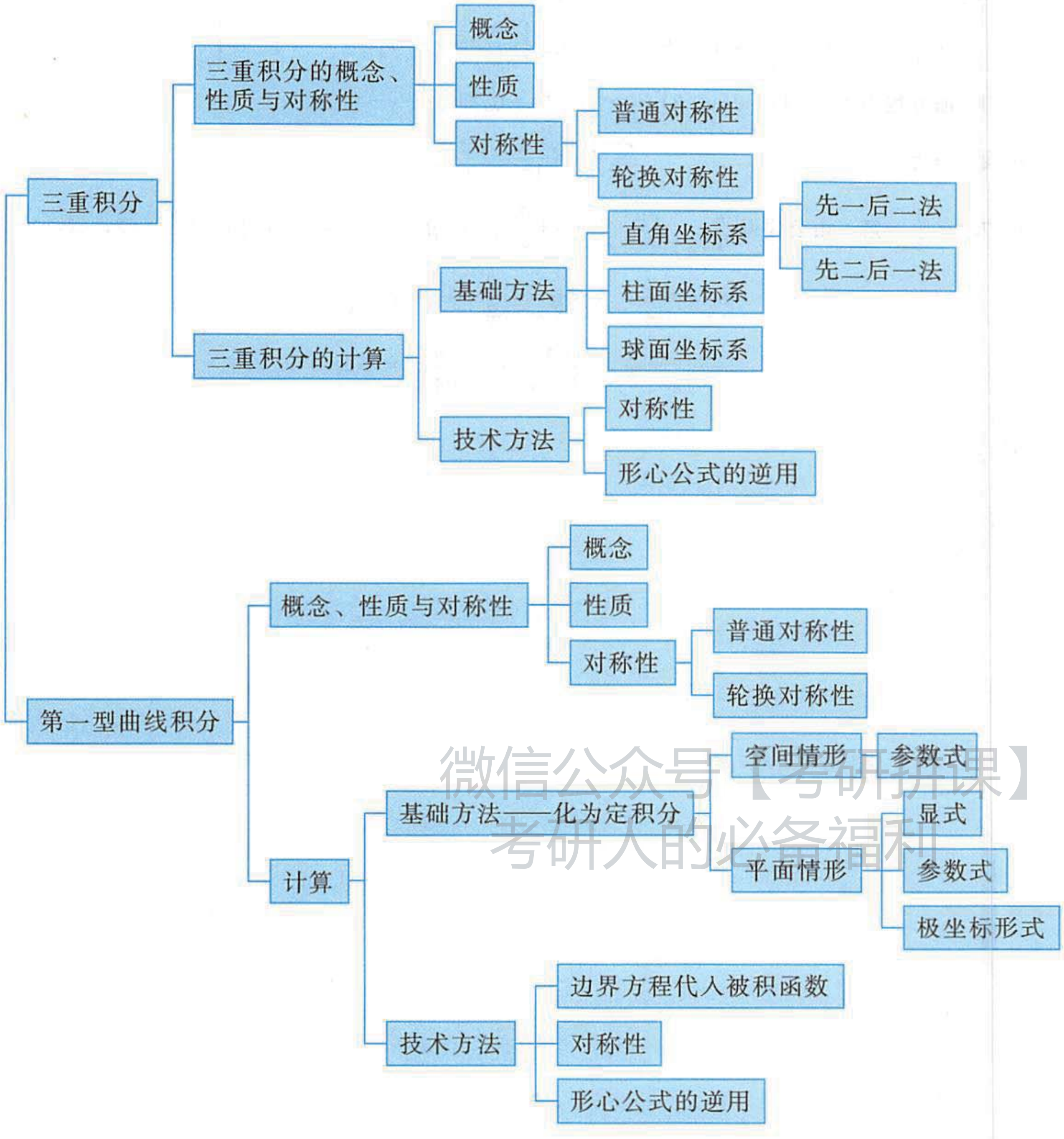
故
$$\cos\alpha=\frac{2}{3},\quad \cos\beta=-\frac{2}{3},\quad \cos\gamma=\frac{1}{3}.$$

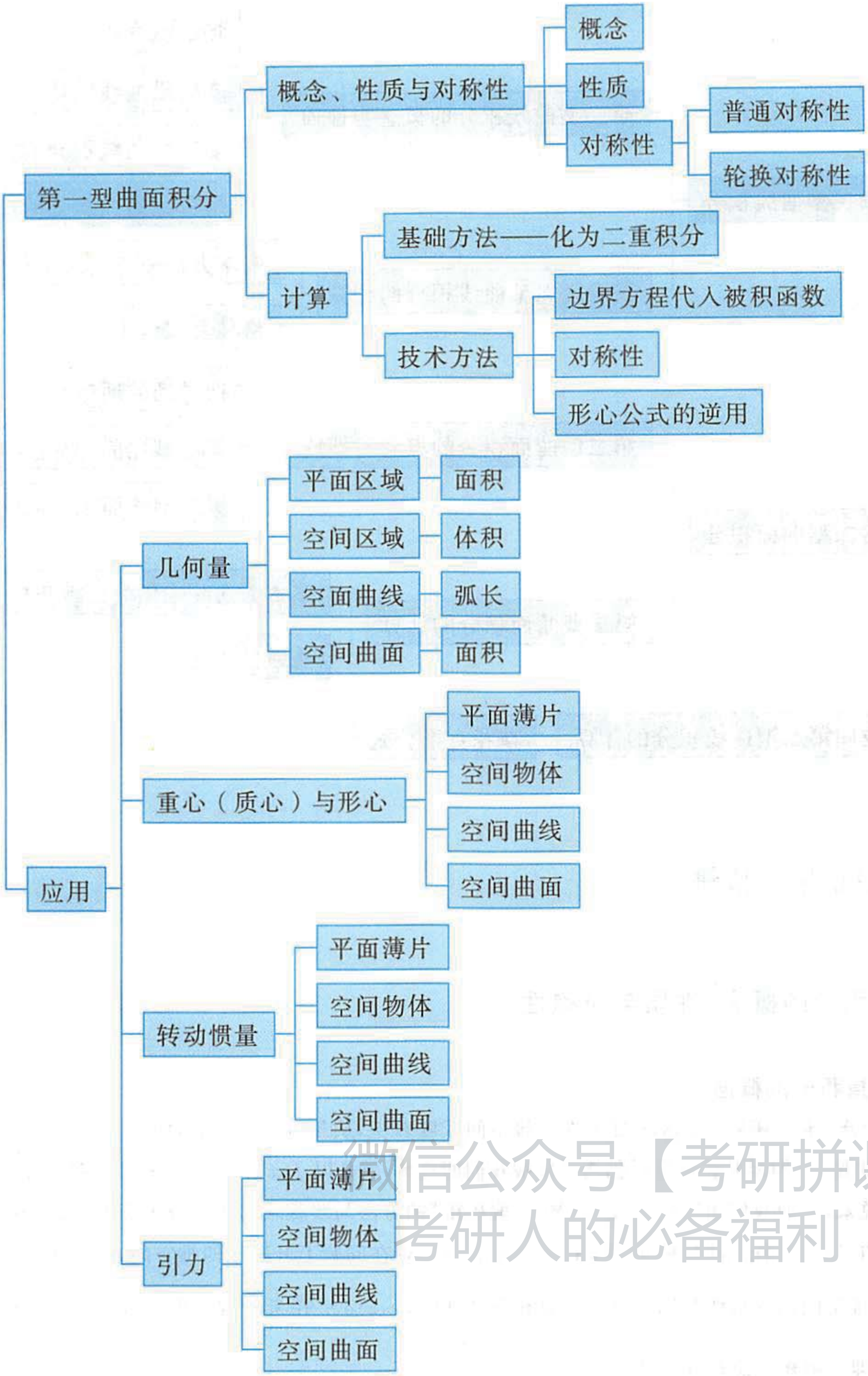
$$\frac{\partial u}{\partial x}\Big|_A=\frac{1}{2},\quad \frac{\partial u}{\partial y}\Big|_A=0,\quad \frac{\partial u}{\partial z}\Big|_A=\frac{1}{2},$$

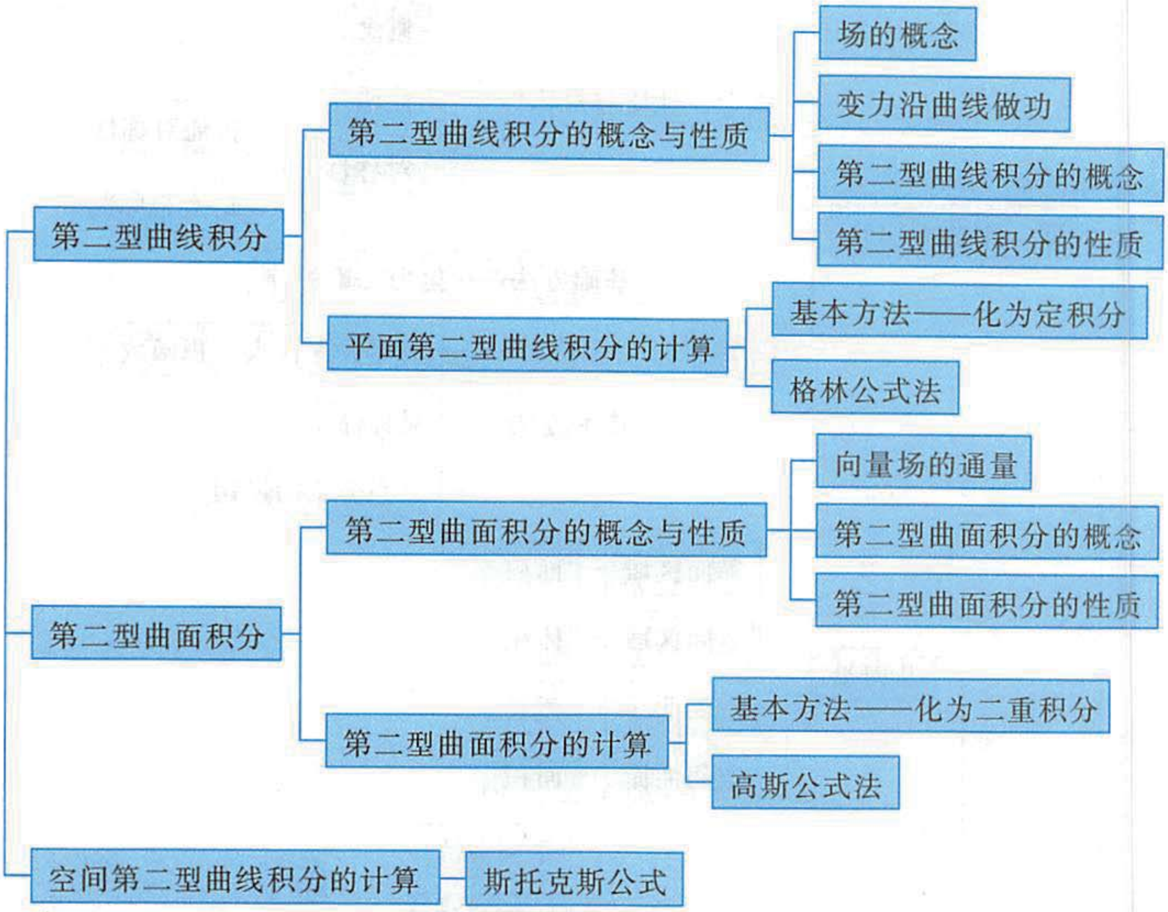
所以 $\frac{\partial u}{\partial \boldsymbol{l}}\Big|_A=\frac{1}{2}$.

第18讲 三重积分、曲线曲面积分（仅数学一要求）

基础知识结构




第一型曲面积分
概念、性质与对称性
概念
性质
对称性
普通对称性
轮换对称性
计算
基础方法——化为二重积分
技术方法
边界方程代入被积函数
对称性
形心公式的逆用
应用
几何量
平面区域
面积
空间区域
体积
空面曲线
弧长
空间曲面
面积
重心（质心）与形心
平面薄片
空间物体
空间曲线
空间曲面
转动惯量
平面薄片
空间物体
空间曲线
空间曲面
引力
平面薄片
空间物体
空间曲线
空间曲面



基础内容精讲

一、三重积分的概念、性质与对称性

1. 三重积分的概念

三重积分的被积函数 $f(x,y,z)$ 定义在三维空间区域 Ω 上，从几何上来说很抽象了，是四维空间的图形体积，无法画出图形；但是其物理背景仍然可以被我们所理解，就是以 $f(x,y,z)$ 为点密度的空间物体的质量.

简要说来，前面我们用“分割、近似、求和、取极限”的方法与步骤求出了二维平面上“曲边梯形的面积”（定积分）和三维空间中“曲顶柱体的体积”（二重积分），现在问题上升到了四维空间，我们可以用同样的办法求出“四维空间的图形体积”，这就是三重积分 $\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v$. 在考研数学中，一般总假设 $f(x,y,z)$ 在 Ω 上可积，即三重积分总是存在的.

2. 三重积分的性质（以下总假设 Ω 为空间有界闭区域）

性质 1（求空间区域的体积） $\iiint\limits_{\Omega} 1\mathrm{d}v = \iiint\limits_{\Omega} \mathrm{d}v = V$，其中 V 为 Ω 的体积.

性质 2（可积函数必有界） 设 $f(x,y,z)$ 在 Ω 上可积，则其在 Ω 上必有界.

性质 3（积分的线性性质） 设 k_1,k_2 为常数，则

$$\iiint\limits_{\Omega}[k_1 f(x,y,z) \pm k_2 g(x,y,z)]\mathrm{d}v = k_1\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v \pm k_2\iiint\limits_{\Omega} g(x,y,z)\mathrm{d}v.$$

性质 4（积分的可加性） 设 $f(x,y,z)$ 在 Ω 上可积，且 $\Omega_1 \cup \Omega_2 = \Omega, \Omega_1 \cap \Omega_2 = \varnothing$，则

$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = \iiint\limits_{\Omega_1} f(x,y,z)\mathrm{d}v + \iiint\limits_{\Omega_2} f(x,y,z)\mathrm{d}v.$$

性质 5（积分的保号性） 设 $f(x,y,z), g(x,y,z)$ 在 Ω 上可积，且在 Ω 上 $f(x,y,z) \leqslant g(x,y,z)$，则有

$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v \leqslant \iiint\limits_{\Omega} g(x,y,z)\mathrm{d}v.$$

特殊地，有
$$\left|\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v\right| \leqslant \iiint\limits_{\Omega} |f(x,y,z)|\mathrm{d}v.$$

性质 6（三重积分的估值定理） 设 M, m 分别是 $f(x,y,z)$ 在 Ω 上的最大值和最小值，V 为 Ω 的体积，则有
$$mV \leqslant \iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v \leqslant MV.$$

性质 7（三重积分的中值定理） 设 $f(x,y,z)$ 在 Ω 上连续，V 为 Ω 的体积，则在 Ω 上至少存在一点 (ξ,η,ζ)，使得
$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = f(\xi,\eta,\zeta)V.$$

3. 普通对称性与轮换对称性

分析方法与二重积分完全一样.

(1)普通对称性.

假设 Ω 关于 yOz 面对称，则

$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = \begin{cases} 2\iiint\limits_{\Omega_1} f(x,y,z)\mathrm{d}v, & f(x,y,z) = f(-x,y,z), \\ 0, & f(x,y,z) = -f(-x,y,z), \end{cases}$$

其中 Ω_1 是 Ω 在 yOz 面前面的部分.

关于其他坐标面对称的情况与此类似.

(2)轮换对称性.

若把 x 与 y 对调后，Ω 不变，则 $\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = \iiint\limits_{\Omega} f(y,x,z)\mathrm{d}v$，这就是**轮换对称性**.

关于其他情况与此类似.

如 $\Omega = \{(x,y,z) \mid x^2+y^2+z^2 \leqslant R^2\}$，则 $\iiint\limits_{\Omega} f(x)\mathrm{d}v = \iiint\limits_{\Omega} f(y)\mathrm{d}v = \iiint\limits_{\Omega} f(z)\mathrm{d}v$ 可以化简计算.

具体应用见后面的例子.

二、三重积分的计算

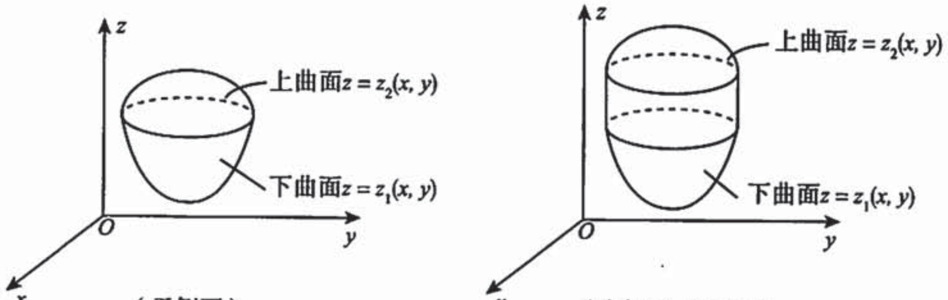
（一）基础方法

1. 直角坐标系

(1)先一后二法（先 z 后 xy 法，也叫投影穿线法）.

①适用场合.

Ω 有下曲面 $z=z_1(x,y)$、上曲面 $z=z_2(x,y)$. 无侧面或侧面为柱面，如图 1-18-1 所示.



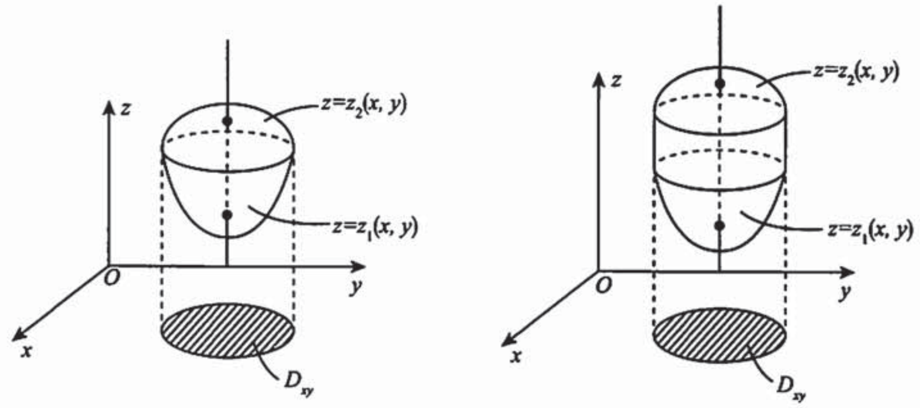
（无侧面）
(a)

（有侧面，为柱面）
(b)

图 1-18-1

②计算方法.

如图 1-18-2 所示，有
$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = \iint\limits_{D_{xy}} \mathrm{d}\sigma \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\mathrm{d}z.$$



(a)

(b)

图 1-18-2

(2)先二后一法（先 xy 后 z 法，也叫定限截面法）.

①适用场合.

Ω 是旋转体，其旋转曲面方程为 $\Sigma: z = z(x,y)$. 如图 1-18-3 所示.

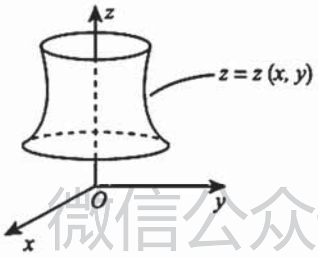


图 1-18-3

② 计算方法.

如图 1-18-4 所示，有
$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}v = \int_a^b \mathrm{d}z \iint\limits_{D_z} f(x,y,z)\mathrm{d}\sigma.$$

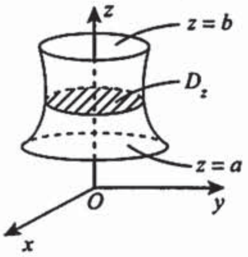


图 1-18-4

2. 柱面坐标系

在直角坐标系的先一后二法中，若$\iint\limits_{D_{xy}} \mathrm{d}\sigma$ 适用于极坐标系，则令$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \end{cases}$便有

$$\iiint\limits_{\Omega} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \iiint\limits_{\Omega} f(r\cos\theta, r\sin\theta, z)r\mathrm{d}r\mathrm{d}\theta\mathrm{d}z,$$

此种计算方法称为柱面坐标系下三重积分的计算.

3. 球面坐标系

(1)适用场合.

何时用球面坐标法？

①被积函数中含$\begin{cases} f(x^2+y^2+z^2), \\ f(x^2+y^2). \end{cases}$

②积分区域为$\begin{cases} \text{球或球的部分}, \\ \text{锥或锥的部分}. \end{cases}$

(2)计算方法.

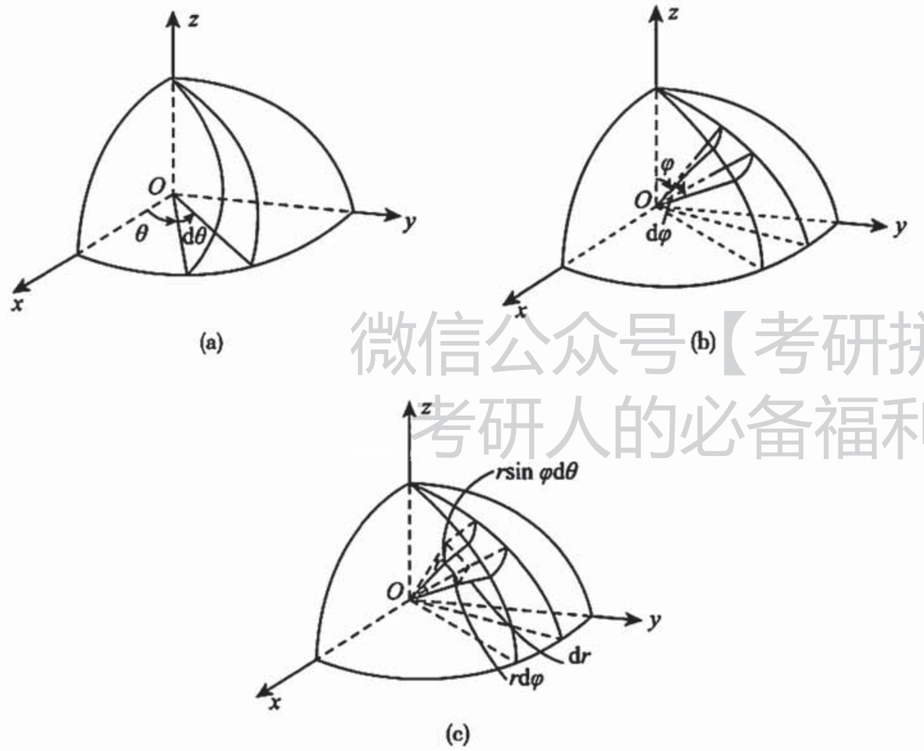
令
$$\begin{cases} x = r\sin\varphi\cos\theta, \\ y = r\sin\varphi\sin\theta, \\ z = r\cos\varphi, \end{cases}$$

用三族面将空间 Ω 切分成一个一个的微元体，如图 1-18-5 所示.

其中①$r=r_0$(常数)，为球心在原点的球面，其半径为 r_0，这里 $0\leqslant r_0<+\infty$；

②$\varphi=\varphi_0$(常数)，为以 z 轴为中心轴的圆锥面. 其顶点为原点，半顶角为 φ_0，$0\leqslant\varphi_0\leqslant\pi$；

③$\theta=\theta_0$(常数)，为过 z 轴的半平面，其与 xOz 面正向夹角为 θ_0，$0\leqslant\theta_0\leqslant 2\pi$.



(a)　(b)

(c)

图 1-18-5

此微元体近似为长方体，其三组边界面分别为：以原点为圆心，半径为 r 与 $r+\mathrm{d}r$ 的球面；以 z 轴为中心轴，半顶角为 φ 与 $\varphi+\mathrm{d}\varphi$ 的圆锥面；过 z 轴且与 xOz 面正向夹角为 θ 与 $\theta+\mathrm{d}\theta$ 的半平面. 它的体积元素即为三个边长 $\mathrm{d}r, r\mathrm{d}\varphi$ 与 $r\sin\varphi\mathrm{d}\theta$ 的乘积，即 $\mathrm{d}v=r^2\sin\varphi\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}r$.

于是球面坐标系下三重积分计算的公式就写成：

$$\iiint_{\Omega} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iiint_{\Omega} f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)r^2\sin\varphi\mathrm{d}\theta\mathrm{d}\varphi\mathrm{d}r.$$

【注】 怎样定限？

①从原点出发画一条半射线（取值范围 $(0,+\infty)$）$\begin{cases}\text{先碰到 }\Omega\text{，记 } r_1(\varphi,\theta),\\ \text{后离开 }\Omega\text{，记 } r_2(\varphi,\theta).\end{cases}$

②顶点在原点，以 z 轴为对称轴的圆锥面半顶角（取值范围 $[0,\pi]$）$\begin{cases}\text{先碰到 }\Omega\text{，记 } \varphi_1(\theta),\\ \text{后离开 }\Omega\text{，记 } \varphi_2(\theta).\end{cases}$

③过 z 轴的半平面（取值范围 $[0,2\pi]$）$\begin{cases}\text{先碰到 }\Omega\text{，记 } \theta_1,\\ \text{后离开 }\Omega\text{，记 } \theta_2,\end{cases}$ 则

$$\begin{aligned}\iiint_{\Omega} f(x,y,z)\mathrm{d}v&=\iiint_{\Omega} f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)r^2\sin\varphi\mathrm{d}r\mathrm{d}\varphi\mathrm{d}\theta\\&=\int_{\theta_1}^{\theta_2}\mathrm{d}\theta\int_{\varphi_1(\theta)}^{\varphi_2(\theta)}\mathrm{d}\varphi\int_{r_1(\varphi,\theta)}^{r_2(\varphi,\theta)} f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)r^2\sin\varphi\,\mathrm{d}r.\end{aligned}$$

（二）技术方法

①对称性（包括普通对称性和轮换对称性）；

② 形心公式的逆用（由 $\bar{x}=\dfrac{\iiint_{\Omega} x\mathrm{d}v}{\iiint_{\Omega}\mathrm{d}v}\Rightarrow\iiint_{\Omega} x\mathrm{d}v=\bar{x}\cdot V$，其中 V 为 Ω 的体积）.

三、第一型曲线积分的概念、性质与对称性

为表述方便，除特殊说明外，以下仅讨论空间情况，平面情况比较简单.

1. 第一型曲线积分的概念

第一型曲线积分的被积函数 $f(x,y)$（或 $f(x,y,z)$）定义在平面曲线 L（或空间曲线 Γ）上，其物理背景是以 $f(x,y)$（或 $f(x,y,z)$）为线密度的**平面（或空间）物质曲线的质量**. 与前面类似，我们仍然可以用“分割、近似、求和、取极限”的方法与步骤写出第一型曲线积分

$$\int_L f(x,y)\mathrm{d}s\text{（或}\int_{\Gamma} f(x,y,z)\mathrm{d}s\text{）}.$$

但事实上，如果仅理解到此，还是不够的，不妨把定积分和第一型曲线积分放在一起做个对比，加深我们对概念的理解.

定积分定义在“直线段”上，而第一型曲线积分定义在“曲线段”上.

在考研数学中，一般总假设 $f(x,y,z)$ 在 Γ 上可积，即第一型曲线积分总是存在的.

2. 第一型曲线积分的性质（以下总假设 Γ 为空间有限长分段光滑曲线）

性质 8（求空间曲线的长度（弧长）） $\int_{\Gamma} 1\mathrm{d}s = l_{\Gamma}$，其中 l_{Γ} 为 Γ 的长度.

性质 9（可积函数必有界） 设 $f(x,y,z)$ 在 Γ 上可积，则其在 Γ 上必有界.

性质 10（积分的线性性质） 设 k_1, k_2 为常数，则

$$\int_{\Gamma}[k_1 f(x,y,z) \pm k_2 g(x,y,z)]\mathrm{d}s = k_1\int_{\Gamma} f(x,y,z)\mathrm{d}s \pm k_2\int_{\Gamma} g(x,y,z)\mathrm{d}s.$$

性质 11（积分的可加性） 设 $f(x,y,z)$ 在 Γ 上可积，且 $\Gamma_1 \cup \Gamma_2 = \Gamma, \Gamma_1 \cap \Gamma_2 = \varnothing$，则

$$\int_{\Gamma} f(x,y,z)\mathrm{d}s = \int_{\Gamma_1} f(x,y,z)\mathrm{d}s + \int_{\Gamma_2} f(x,y,z)\mathrm{d}s.$$

性质 12（积分的保号性） 设 $f(x,y,z), g(x,y,z)$ 在 Γ 上可积，且在 Γ 上 $f(x,y,z) \leqslant g(x,y,z)$，则有

$$\int_{\Gamma} f(x,y,z)\mathrm{d}s \leqslant \int_{\Gamma} g(x,y,z)\mathrm{d}s.$$

特殊地，有

$$\left|\int_{\Gamma} f(x,y,z)\mathrm{d}s\right| \leqslant \int_{\Gamma} |f(x,y,z)|\mathrm{d}s.$$

性质 13（第一型曲线积分的估值定理） 设 M, m 分别是 $f(x,y,z)$ 在 Γ 上的最大值和最小值，l_{Γ} 为 Γ 的长度，则有

$$ml_{\Gamma} \leqslant \int_{\Gamma} f(x,y,z)\mathrm{d}s \leqslant Ml_{\Gamma}.$$

性质 14（第一型曲线积分的中值定理） 设 $f(x,y,z)$ 在 Γ 上连续，l_{Γ} 为 Γ 的长度，则在 Γ 上至少存在一点 (ξ,η,ζ)，使得

$$\int_{\Gamma} f(x,y,z)\mathrm{d}s = f(\xi,\eta,\zeta)l_{\Gamma}.$$

3. 普通对称性与轮换对称性

分析方法与二、三重积分完全一样.

(1)普通对称性.

假设 Γ 关于 yOz 面对称，则

$$\int_{\Gamma} f(x,y,z)\mathrm{d}s = \begin{cases} 2\int_{\Gamma_1} f(x,y,z)\mathrm{d}s, & f(x,y,z) = f(-x,y,z), \\ 0, & f(x,y,z) = -f(-x,y,z), \end{cases}$$

其中 Γ_1 是 Γ 在 yOz 面前面的部分.

关于其他坐标面对称的情况与此类似.

(2)轮换对称性.

若把 x 与 y 对调后，Γ 不变，则 $\int_{\Gamma} f(x,y,z)\mathrm{d}s = \int_{\Gamma} f(y,x,z)\mathrm{d}s$，这就是**轮换对称性**.

关于其他情况与此类似.

具体应用见后面的例子.

四、第一型曲线积分的计算

(一)基础方法——化为定积分

由于第一型曲线积分就是由定积分推广而来的，所以计算第一型曲线积分的基本方法就是将其化为定积分.

1. 对于空间情形

若空间曲线 Γ 由参数式 $\begin{cases} x = x(t), \\ y = y(t), (\alpha \leqslant t \leqslant \beta) \\ z = z(t) \end{cases}$ 给出，则

$$ds = \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2}\,dt,$$

且
$$\int_\Gamma f(x,y,z)\,ds = \int_\alpha^\beta f[x(t),y(t),z(t)]\sqrt{[x'(t)]^2+[y'(t)]^2+[z'(t)]^2}\,dt.$$

2. 对于平面情形

(1)若平面曲线 L 由 $y=y(x)$ $(a\leqslant x\leqslant b)$给出，则 $ds=\sqrt{1+[y'(x)]^2}\,dx$，且

$$\int_L f(x,y)\,ds = \int_a^b f[x,y(x)]\sqrt{1+[y'(x)]^2}\,dx.$$

(2)若平面曲线 L 由参数式 $\begin{cases} x=x(t), \\ y=y(t) \end{cases}$ $(\alpha\leqslant t\leqslant\beta)$给出，则 $ds=\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt$，且

$$\int_L f(x,y)\,ds = \int_\alpha^\beta f[x(t),y(t)]\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt.$$

(3)若平面曲线 L 由极坐标形式 $r=r(\theta)$ $(\alpha\leqslant\theta\leqslant\beta)$给出，则 $ds=\sqrt{[r(\theta)]^2+[r'(\theta)]^2}\,d\theta$，且

$$\int_L f(x,y)\,ds = \int_\alpha^\beta f[r(\theta)\cos\theta, r(\theta)\sin\theta]\sqrt{[r(\theta)]^2+[r'(\theta)]^2}\,d\theta.$$

(二)技术方法

①边界方程代入被积函数(由于被积函数就定义在边界方程上，所以可以把边界方程的表达式代入到被积函数中，从而达到简化计算的目的)；

②对称性(包括普通对称性和轮换对称性)；

③形心公式的逆用 (由 $\bar{x} = \dfrac{\int_\Gamma x\,ds}{\int_\Gamma ds} \Rightarrow \int_\Gamma x\,ds = \bar{x}\cdot l_\Gamma$，其中 l_Γ 为 Γ 的长度).

五、第一型曲面积分的概念、性质与对称性

1. 第一型曲面积分的概念

第一型曲面积分的被积函数 $f(x,y,z)$ 定义在空间曲面 Σ 上，其物理背景是以 $f(x,y,z)$ 为面密度的空间物质曲面的质量. 与前面类似，我们可以用"分割、近似、求和、取极限"的方法与步骤写出第一型曲面积分

$$\iint_\Sigma f(x,y,z)\,dS.$$

如前所述，仅理解到此是不够的，不妨把二重积分和第一型曲面积分放在一起做个对比，加深我们对概念的理解.

二重积分定义在"二维平面"上，而第一型曲面积分则定义在"空间曲面"上.

在考研数学中，一般总假设 $f(x,y,z)$ 在 Σ 上可积，即第一型曲面积分总是存在的.

2. 第一型曲面积分的性质(以下总假设 Σ 为空间有限分片光滑曲面)

性质 15(求空间曲面的面积) $\iint_\Sigma 1\,dS = S$，其中 S 为 Σ 的面积.

性质 16（可积函数必有界）　当 $f(x,y,z)$ 在 Σ 上可积时，则其在 Σ 上必有界.

性质 17（积分的线性性质）　设 k_1,k_2 为常数，则

$$\iint\limits_{\Sigma}[k_1 f(x,y,z)\pm k_2 g(x,y,z)]\mathrm{d}S = k_1\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S \pm k_2\iint\limits_{\Sigma} g(x,y,z)\mathrm{d}S.$$

性质 18（积分的可加性）　设 $f(x,y,z)$ 在 Σ 上可积，且 $\Sigma_1\cup\Sigma_2=\Sigma,\Sigma_1\cap\Sigma_2=\varnothing$，则

$$\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S = \iint\limits_{\Sigma_1} f(x,y,z)\mathrm{d}S + \iint\limits_{\Sigma_2} f(x,y,z)\mathrm{d}S.$$

性质 19（积分的保号性）　当 $f(x,y,z),g(x,y,z)$ 在 Σ 上可积，且在 Σ 上 $f(x,y,z)\leqslant g(x,y,z)$，则有

$$\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S \leqslant \iint\limits_{\Sigma} g(x,y,z)\mathrm{d}S.$$

特殊地，有

$$\left|\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S\right| \leqslant \iint\limits_{\Sigma} |f(x,y,z)|\,\mathrm{d}S.$$

性质 20（第一型曲面积分的估值定理）　设 M,m 分别是 $f(x,y,z)$ 在 Σ 上的最大值和最小值，S 为 Σ 的面积，则有

$$mS \leqslant \iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S \leqslant MS.$$

性质 21（第一型曲面积分的中值定理）　设 $f(x,y,z)$ 在 Σ 上连续，S 为 Σ 的面积，则在 Σ 上至少存在一点 (ξ,η,ζ)，使得

$$\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S = f(\xi,\eta,\zeta)S.$$

3. 普通对称性与轮换对称性

分析方法与二、三重积分和第一型曲线积分完全一样.

（1）普通对称性.

假设 Σ 关于 yOz 面对称，则

$$\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S = \begin{cases} 2\iint\limits_{\Sigma_1} f(x,y,z)\mathrm{d}S, & f(x,y,z)=f(-x,y,z), \\ 0, & f(x,y,z)=-f(-x,y,z), \end{cases}$$

其中 Σ_1 是 Σ 在 yOz 面前面的部分.

关于其他坐标面对称的情况与此类似.

（2）轮换对称性.

若把 x 与 y 对调后，Σ 不变，则 $\iint\limits_{\Sigma} f(x,y,z)\mathrm{d}S = \iint\limits_{\Sigma} f(y,x,z)\mathrm{d}S$，这就是**轮换对称性**.

关于其他情况与此类似.

具体应用见后面的例子.

六、第一型曲面积分的计算

1. 基础方法——化为二重积分

由于第一型曲面积分就是由二重积分推广而来的，所以计算第一型曲面积分的基本方法就是将其化为二重积分.

无论空间曲面 Σ 是由显式 $z=z(x,y)$ 还是隐式 $F(x,y,z)=0$ 给出的，我们都需要做三件事（无逻辑上的先后顺序，哪件事情最利于解题就先做哪件）：

①将 Σ 投影到某一平面（比如 xOy 面）上⇒投影区域为 D（比如 D_{xy}）；

②将 $z=z(x,y)$ 或者 $F(x,y,z)=0$ 代入 $f(x,y,z)$；

③计算 $z'_x,z'_y \Rightarrow \mathrm{d}S=\sqrt{1+(z'_x)^2+(z'_y)^2}\,\mathrm{d}x\mathrm{d}y$.

这就把第一型曲面积分化成二重积分（如化成关于 x,y 的二重积分），得到

$$\iint_{\Sigma} f(x,y,z)\,\mathrm{d}S=\iint_{D_{xy}} f(x,y,z(x,y))\sqrt{1+(z'_x)^2+(z'_y)^2}\,\mathrm{d}x\mathrm{d}y.$$

化成关于其他变量的二重积分与此类似.

【注】 这里有一点需要特别强调，将 Σ 投影到哪个平面上应该是由你自己决定的，但是 **Σ 上的任何两点的投影点不能重合**，换言之，假如你要将 Σ 投向 xOy 面，则 $z=z(x,y)$ 必须是单值函数. 忘记了这一点，就可能算错结果.

如果将 Σ 投向某一平面，但是曲面投影后有重合点，则

①要么将 Σ 转投向另一个平面，使得曲面投影后无重合点；

②要么将 Σ 分成若干曲面 $\Sigma_1,\Sigma_2,\cdots$，使得这些曲面各自投影后无重合点.

2. 技术方法

①边界方程代入被积函数（由于被积函数就定义在边界方程上，所以可以把边界方程的表达式代入到被积函数中，从而达到简化计算的目的）；

②对称性（包括普通对称性和轮换对称性）；

③ 形心公式的逆用（由 $\bar{x}=\dfrac{\iint\limits_{\Sigma} x\,\mathrm{d}S}{\iint\limits_{\Sigma}\mathrm{d}S} \Rightarrow \iint\limits_{\Sigma} x\,\mathrm{d}S=\bar{x}\cdot S$，其中 S 为 Σ 的面积）.

七、重积分与第一型线面积分的应用

1. 几何量

（1）对于平面区域，有两种情况：

① 若 D 是由 $y=f(x)$，$x=a$，$x=b$ 与 x 轴所围成的曲边梯形，则其面积为 $A=\int_a^b |f(x)|\,\mathrm{d}x$，其中 $a<b$；

② 平面区域 D 的面积为 $A=\iint\limits_{D}\mathrm{d}\sigma$.

（2）对于空间区域，有两种情况：

① 以区域 D 为底，曲面 $z=z(x,y)$ 为顶的曲顶柱体的体积为 $V=\iint\limits_{D}|z(x,y)|\,\mathrm{d}\sigma$；

【注】 这种情况是考试重点，常有 $z(x,y)=f(x,y)-g(x,y)$，也就是这个空间区域是由两个不同的曲面 $\Sigma_1:z_1=f(x,y)$ 和 $\Sigma_2:z_2=g(x,y)$（有时也需要添加一些坐标面）所围成的.

② 若 Ω 是所占的空间区域，则其体积为 $V=\iiint\limits_{\Omega}\mathrm{d}v$.

(3)对于空间光滑曲线 Γ，若其由参数式$\begin{cases}x=x(t),\\ y=y(t),(\alpha\leqslant t\leqslant\beta)\\ z=z(t)\end{cases}$给出，则计算空间曲线的长度(弧长)的公式为

$$L=\int_{\alpha}^{\beta}\sqrt{[x'(t)]^2+[y'(t)]^2+[z'(t)]^2}\mathrm{d}t.$$

(4)对于光滑曲面薄片 Σ，若 Σ 由单值函数 $z=z(x,y)$ 给出，D_{xy} 为曲面 Σ 在 xOy 面上的投影区域，则其面积

$$A=\iint_{D_{xy}}\sqrt{1+(z_x')^2+(z_y')^2}\mathrm{d}x\mathrm{d}y.$$

【注】 同理，在同样保证单值函数的情况下，可向另外两个坐标面投影，得

$$A=\iint_{D_{yz}}\sqrt{1+(x_y')^2+(x_z')^2}\mathrm{d}y\mathrm{d}z,$$

其中 $\Sigma:x=x(y,z)$，D_{yz} 是曲面在 yOz 面上的投影区域；

$$A=\iint_{D_{zx}}\sqrt{1+(y_z')^2+(y_x')^2}\mathrm{d}z\mathrm{d}x,$$

其中 $\Sigma:y=y(x,z)$，D_{zx} 是曲面在 zOx 面上的投影区域.

事实上，曲面面积就是第一型曲面积分的被积函数是 1 时，用投影法所得出的积分，请大家注意这个联系.

2. 重心(质心)与形心

以下如无特殊说明，均假设 $\rho(x,y)$ 与 $\rho(x,y,z)$ 在所定义的区域上连续.

(1)对于平面薄片，面密度为 $\rho(x,y)$，D 是薄片所占的平面区域，则计算重心 $(\bar{x},\bar{y})$ 的公式为

$$\bar{x}=\frac{\iint\limits_D x\rho(x,y)\mathrm{d}\sigma}{\iint\limits_D \rho(x,y)\mathrm{d}\sigma},\quad \bar{y}=\frac{\iint\limits_D y\rho(x,y)\mathrm{d}\sigma}{\iint\limits_D \rho(x,y)\mathrm{d}\sigma}.$$

(2)对于空间物体，体密度为 $\rho(x,y,z)$，Ω 是物体所占的空间区域，则计算重心 $(\bar{x},\bar{y},\bar{z})$ 的公式为

$$\bar{x}=\frac{\iiint\limits_\Omega x\rho(x,y,z)\mathrm{d}v}{\iiint\limits_\Omega \rho(x,y,z)\mathrm{d}v},\quad \bar{y}=\frac{\iiint\limits_\Omega y\rho(x,y,z)\mathrm{d}v}{\iiint\limits_\Omega \rho(x,y,z)\mathrm{d}v},\quad \bar{z}=\frac{\iiint\limits_\Omega z\rho(x,y,z)\mathrm{d}v}{\iiint\limits_\Omega \rho(x,y,z)\mathrm{d}v}.$$

(3)对于光滑曲线 L，线密度为 $\rho(x,y,z)$，则计算重心 $(\bar{x},\bar{y},\bar{z})$ 的公式为

$$\bar{x}=\frac{\int_L x\rho(x,y,z)\mathrm{d}s}{\int_L \rho(x,y,z)\mathrm{d}s},\quad \bar{y}=\frac{\int_L y\rho(x,y,z)\mathrm{d}s}{\int_L \rho(x,y,z)\mathrm{d}s},\quad \bar{z}=\frac{\int_L z\rho(x,y,z)\mathrm{d}s}{\int_L \rho(x,y,z)\mathrm{d}s}.$$

(4)对于光滑曲面薄片 Σ，面密度为 $\rho(x,y,z)$，则计算重心 $(\bar{x},\bar{y},\bar{z})$ 的公式为

$$\bar{x}=\frac{\iint\limits_\Sigma x\rho(x,y,z)\mathrm{d}S}{\iint\limits_\Sigma \rho(x,y,z)\mathrm{d}S},\quad \bar{y}=\frac{\iint\limits_\Sigma y\rho(x,y,z)\mathrm{d}S}{\iint\limits_\Sigma \rho(x,y,z)\mathrm{d}S},\quad \bar{z}=\frac{\iint\limits_\Sigma z\rho(x,y,z)\mathrm{d}S}{\iint\limits_\Sigma \rho(x,y,z)\mathrm{d}S}.$$

【注】 (1)在考研的范畴内，重心就是质心.

(2)当密度 $\rho(x,y)$ 或者 $\rho(x,y,z)$ 为常数时，重心就成了**形心**.

3. 转动惯量

(1)对于平面薄片，面密度为 $\rho(x,y)$，D 是薄片所占的平面区域，则计算该薄片对 x 轴、y 轴和原点 O 的转动惯量 I_x, I_y 和 I_O 公式分别为

$$I_x=\iint\limits_D y^2\rho(x,y)\mathrm{d}\sigma,\quad I_y=\iint\limits_D x^2\rho(x,y)\mathrm{d}\sigma,\quad I_O=\iint\limits_D (x^2+y^2)\rho(x,y)\mathrm{d}\sigma.$$

(2)对于空间物体，体密度为 $\rho(x,y,z)$，Ω 是物体所占的空间区域，则计算该物体对 x 轴、y 轴、z 轴和原点 O 的转动惯量 I_x, I_y, I_z 和 I_O 公式分别为

$$I_x=\iiint\limits_\Omega (y^2+z^2)\rho(x,y,z)\mathrm{d}v,\quad I_y=\iiint\limits_\Omega (z^2+x^2)\rho(x,y,z)\mathrm{d}v,$$

$$I_z=\iiint\limits_\Omega (x^2+y^2)\rho(x,y,z)\mathrm{d}v,\quad I_O=\iiint\limits_\Omega (x^2+y^2+z^2)\rho(x,y,z)\mathrm{d}v.$$

(3)对于光滑曲线 L，线密度为 $\rho(x,y,z)$，则计算该曲线对 x 轴、y 轴、z 轴和原点 O 的转动惯量 I_x, I_y, I_z 和 I_O 公式分别为

$$I_x=\int_L (y^2+z^2)\rho(x,y,z)\mathrm{d}s,\quad I_y=\int_L (z^2+x^2)\rho(x,y,z)\mathrm{d}s,$$

$$I_z=\int_L (x^2+y^2)\rho(x,y,z)\mathrm{d}s,\quad I_O=\int_L (x^2+y^2+z^2)\rho(x,y,z)\mathrm{d}s.$$

(4)对于光滑曲面薄片 Σ，面密度为 $\rho(x,y,z)$，则计算该曲面对 x 轴、y 轴、z 轴和原点 O 的转动惯量 I_x, I_y, I_z 和 I_O 公式分别为

$$I_x=\iint\limits_\Sigma (y^2+z^2)\rho(x,y,z)\mathrm{d}S,\quad I_y=\iint\limits_\Sigma (z^2+x^2)\rho(x,y,z)\mathrm{d}S,$$

$$I_z=\iint\limits_\Sigma (x^2+y^2)\rho(x,y,z)\mathrm{d}S,\quad I_O=\iint\limits_\Sigma (x^2+y^2+z^2)\rho(x,y,z)\mathrm{d}S.$$

4. 引力

(1)对于 xOy 面上一平面薄片，面密度为 $\rho(x,y)$，D 是薄片所占的平面区域，则计算该薄片对点 $M_0(x_0,y_0,z_0)$ 处的质量为 m 的质点的引力 (F_x,F_y,F_z) 公式为

$$F_x=Gm\iint\limits_D\frac{\rho(x,y)(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+z_0^2]^{\frac{3}{2}}}\mathrm{d}\sigma,$$

$$F_y=Gm\iint\limits_D\frac{\rho(x,y)(y-y_0)}{[(x-x_0)^2+(y-y_0)^2+z_0^2]^{\frac{3}{2}}}\mathrm{d}\sigma,$$

$$F_z=-z_0Gm\iint\limits_D\frac{\rho(x,y)}{[(x-x_0)^2+(y-y_0)^2+z_0^2]^{\frac{3}{2}}}\mathrm{d}\sigma.$$

其中 G 为引力常量，下同.

(2)对于空间物体，体密度为 $\rho(x,y,z)$，Ω 是物体所占的空间区域，则计算该物体对点 $M_0(x_0,y_0,z_0)$ 处的质量为 m 的质点的引力 (F_x,F_y,F_z) 公式为

$$F_x=Gm\iiint\limits_\Omega\frac{\rho(x,y,z)(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}v,$$

$$F_y=Gm\iiint\limits_{\Omega}\frac{\rho(x,y,z)(y-y_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}v,$$

$$F_z=Gm\iiint\limits_{\Omega}\frac{\rho(x,y,z)(z-z_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}v.$$

(3)对于光滑曲线 L，线密度为 $\rho(x,y,z)$，则计算该曲线对点 $M_0(x_0,y_0,z_0)$ 处的质量为 m 的质点的引力 (F_x,F_y,F_z) 公式为

$$F_x=Gm\int_L\frac{\rho(x,y,z)(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}s,$$

$$F_y=Gm\int_L\frac{\rho(x,y,z)(y-y_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}s,$$

$$F_z=Gm\int_L\frac{\rho(x,y,z)(z-z_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}s.$$

(4)对于光滑曲面薄片 Σ，面密度为 $\rho(x,y,z)$，则计算该曲面对点 $M_0(x_0,y_0,z_0)$ 处的质量为 m 的质点的引力 (F_x,F_y,F_z) 公式为

$$F_x=Gm\iint\limits_{\Sigma}\frac{\rho(x,y,z)(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}S,$$

$$F_y=Gm\iint\limits_{\Sigma}\frac{\rho(x,y,z)(y-y_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}S,$$

$$F_z=Gm\iint\limits_{\Sigma}\frac{\rho(x,y,z)(z-z_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{\frac{3}{2}}}\mathrm{d}S.$$

八、第二型曲线积分的概念与性质

1. 场的概念

什么叫“场”？从数学上说，场就是空间区域 Ω 上的一种对应法则.

(1)如果 Ω 上的每一点 $M(x,y,z)$ 都对应着一个数量 u，则在 Ω 上就确定了一个数量函数 $u=u(x,y,z)$，它表示一个**数量场**. 数量场的例子很多，比如温度场，数量场不讲究方向.

(2)如果 Ω 上的每一点 $M(x,y,z)$ 都对应着一个向量 $\boldsymbol{F}$，则在 Ω 上就确定了一个向量函数

$$\boldsymbol{F}(x,y,z)=P(x,y,z)\boldsymbol{i}+Q(x,y,z)\boldsymbol{j}+R(x,y,z)\boldsymbol{k},$$

它表示一个**向量场**. 向量场的例子也很多，比如引力场，向量场讲究方向.

2. 变力沿曲线做功

在一个向量场——变力场中，设某质点在变力 $\boldsymbol{F}(x,y,z)$ 作用下，沿着有向曲线 Γ 从起点 A 移动到终点 B，总共做了多少功？（这个物理背景请大家熟记，考研中出过基于这种背景的考题.）

设沿着有向曲线 Γ 在 $M(x,y,z)$ 点移动了一个微位移 $\mathrm{d}\boldsymbol{r}=\mathrm{d}x\boldsymbol{i}+\mathrm{d}y\boldsymbol{j}+\mathrm{d}z\boldsymbol{k}$，将变力 $\boldsymbol{F}(x,y,z)$ 近似看作常力，则力在此微位移上的微功 $\mathrm{d}W=\boldsymbol{F}(x,y,z)\cdot\mathrm{d}\boldsymbol{r}$，于是变力 $\boldsymbol{F}(x,y,z)$ 沿着有向曲线 Γ 从起点 A 移动到终点 B 所做的总功为

$$W=\int_{\Gamma}\mathrm{d}W=\int_{\Gamma}\boldsymbol{F}(x,y,z)\cdot\mathrm{d}\boldsymbol{r}=\int_{\Gamma}(P(x,y,z),Q(x,y,z),R(x,y,z))\cdot(\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)$$

$$= \int_{\Gamma} P(x,y,z)\mathrm{d}x+Q(x,y,z)\mathrm{d}y+R(x,y,z)\mathrm{d}z,$$

于是我们就引出了第二型曲线积分的概念.

3. 第二型曲线积分的概念

第二型曲线积分的被积函数 $\boldsymbol{F}(x,y)=P(x,y)\boldsymbol{i}+Q(x,y)\boldsymbol{j}$（或 $\boldsymbol{F}(x,y,z)=P(x,y,z)\boldsymbol{i}+Q(x,y,z)\boldsymbol{j}+R(x,y,z)\boldsymbol{k}$）定义在平面曲线 L（或空间曲线 Γ）上，其物理背景是变力 $\boldsymbol{F}(x,y)$（或 $\boldsymbol{F}(x,y,z)$）在平面曲线 L（或空间曲线 Γ）上从起点移动到终点所做的总功：$\int_L P(x,y)\mathrm{d}x+Q(x,y)\mathrm{d}y$（或 $\int_{\Gamma} P(x,y,z)\mathrm{d}x+Q(x,y,z)\mathrm{d}y+R(x,y,z)\mathrm{d}z$）.

由此可以看出，前面所学到的定积分、二重积分、三重积分和第一型曲线积分有着完全一致的背景，都是一个**数量函数在定义区域上计算几何量**（面积、体积等），但是第二型曲线积分与之不同，它是一个**向量函数沿有向曲线的积分**（无几何量可言），于是，有些性质和计算方法都不一样了，一定要加以对比，理解它们的区别和联系，不要用错或者用混了.

4. 第二型曲线积分的性质（以下总假设 Γ 为空间有限长分段光滑曲线）

性质 1（积分的线性性质） 设 k_1,k_2 为常数，则 $\int_{\Gamma}(k_1\boldsymbol{F}_1 \pm k_2\boldsymbol{F}_2)\cdot\mathrm{d}\boldsymbol{r}=k_1\int_{\Gamma}\boldsymbol{F}_1\cdot\mathrm{d}\boldsymbol{r}\pm k_2\int_{\Gamma}\boldsymbol{F}_2\cdot\mathrm{d}\boldsymbol{r}$.

性质 2（积分的有向性） $\int_{\overset{\frown}{AB}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}=-\int_{\overset{\frown}{BA}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}$.

性质 3（积分的可加性） 当 $\overset{\frown}{AB}+\overset{\frown}{BC}=\overset{\frown}{AC}$ 时，$\int_{\overset{\frown}{AC}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}=\int_{\overset{\frown}{AB}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}+\int_{\overset{\frown}{BC}}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}$.

九、平面第二型曲线积分的计算

先做个说明，由于空间第二型曲线积分的经典计算方法与第二型曲面积分有着密切的联系，故放在后面去讨论，本部分只讨论平面第二型曲线积分的计算问题.

1. 基本方法——化为定积分

如果平面有向曲线 L 由参数方程 $\begin{cases} x=x(t), \\ y=y(t) \end{cases}$ $(t:\alpha\to\beta)$ 给出，其中 $t=\alpha$ 对应着起点 A，$t=\beta$ 对应着终点 B，则可以将平面第二型曲线积分化为定积分：

$$\int_L P(x,y)\mathrm{d}x+Q(x,y)\mathrm{d}y=\int_{\alpha}^{\beta}\{P[x(t),y(t)]x'(t)+Q[x(t),y(t)]y'(t)\}\mathrm{d}t,$$

这里的 α,β 谁大谁小无关紧要，关键是分别和起点与终点对应.

格林
(1793—1841)

2. 格林公式法

格林公式 设平面有界闭区域 D 由分段光滑闭曲线 L 围成，$P(x,y),Q(x,y)$ 在 D 上具有一阶连续偏导数，L 取正向，则

$$\oint_L P(x,y)\mathrm{d}x+Q(x,y)\mathrm{d}y=\iint_D\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}\sigma.$$

如图 1-18-6 所示，所谓 L 取正向，是指当一个人沿着 L 的正向前进时，**左手始终在 L 所围成的区域 D 内**. 试想一下假如你在学校的环形操场上跑步，你的左手始终在草坪中，这就是逆时针，说明你跑的方向是正向.

考试要点 一般来说，考试题目不可能直接满足能够使用格林公式的条件，命题人可以“破坏”两种条件：

(1)L 不是封闭曲线，也就是没有围成一个平面有界闭区域 D；

(2)即使 L 围成了一个平面有界闭区域 D，但是 $P(x,y),Q(x,y),\frac{\partial P}{\partial y},\frac{\partial Q}{\partial x}$ 在 D 上不连续.

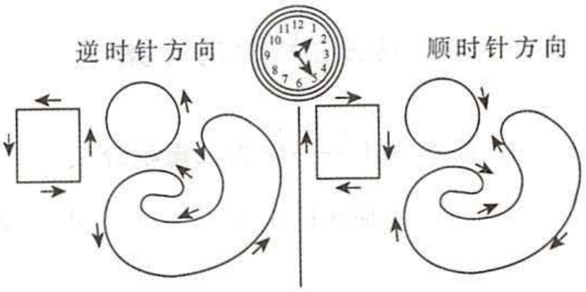


图 1-18-6

这两种情况下，不可以直接使用格林公式.

针对(1)，我们可以采取“补线法”，补上一条或者若干条线，围出一个平面有界闭区域 D，就可以用格林公式了.

针对(2)，我们可以采取“挖去法”，把不连续点(称为“奇点”)挖去，使条件得以满足，从而使用格林公式.

十、第二型曲面积分的概念与性质

1. 向量场的通量

简单回顾一下向量场的概念. 如果 Ω 上的每一点 $M(x,y,z)$ 都对应着一个向量 $\boldsymbol{F}$，则在 Ω 上就确定了一个向量函数 $\boldsymbol{F}(x,y,z)=P(x,y,z)\boldsymbol{i}+Q(x,y,z)\boldsymbol{j}+R(x,y,z)\boldsymbol{k}$，它表示一个**向量场**.

在一个向量场(比如电场、磁场或者某种不可压缩流体的速度场)中，Σ 为该场中的某一有向分片光滑曲面，并指定了曲面的外侧，则向量函数 $\boldsymbol{F}(x,y,z)$ 通过曲面 Σ 的通量(比如电场中的电通量，磁场中的磁通量，或者某流体的流量)为 $\iint\limits_{\Sigma}\boldsymbol{F}\cdot \mathrm{d}\boldsymbol{S}=\iint\limits_{\Sigma}\boldsymbol{F}\cdot\boldsymbol{n}^{\circ}\mathrm{d}S$，其中 $\boldsymbol{n}^{\circ}=(\cos\alpha,\cos\beta,\cos\gamma)$ 是有向曲面 Σ 在指定侧的单位法向量，且由 $\mathrm{d}\boldsymbol{S}=(\mathrm{d}y\mathrm{d}z,\mathrm{d}z\mathrm{d}x,\mathrm{d}x\mathrm{d}y)$，得

$$\iint\limits_{\Sigma}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}=\iint\limits_{\Sigma}P(x,y,z)\mathrm{d}y\mathrm{d}z+Q(x,y,z)\mathrm{d}z\mathrm{d}x+R(x,y,z)\mathrm{d}x\mathrm{d}y.$$

于是我们就引出了第二型曲面积分的概念.

2. 第二型曲面积分的概念

第二型曲面积分的被积函数 $\boldsymbol{F}(x,y,z)=P(x,y,z)\boldsymbol{i}+Q(x,y,z)\boldsymbol{j}+R(x,y,z)\boldsymbol{k}$ 定义在空间曲面 Σ 上，其物理背景是向量函数 $\boldsymbol{F}(x,y,z)$ 通过曲面 Σ 的通量：

$$\iint\limits_{\Sigma}P(x,y,z)\mathrm{d}y\mathrm{d}z+Q(x,y,z)\mathrm{d}z\mathrm{d}x+R(x,y,z)\mathrm{d}x\mathrm{d}y.$$

由此可以看出，第二型曲面积分是一个**向量函数通过某有向曲面的通量(无几何量可言)**，要加强和前面所学积分的横向对比，理解它们的区别和联系，不要用错或者用混了.

3. 第二型曲面积分的性质(以下总假设 Σ 是一有向分片光滑曲面)

性质 1(积分的线性性质) 设 k_1,k_2 为常数，则 $\iint\limits_{\Sigma}(k_1\boldsymbol{F}_1\pm k_2\boldsymbol{F}_2)\cdot\mathrm{d}\boldsymbol{S}=k_1\iint\limits_{\Sigma}\boldsymbol{F}_1\cdot\mathrm{d}\boldsymbol{S}\pm k_2\iint\limits_{\Sigma}\boldsymbol{F}_2\cdot\mathrm{d}\boldsymbol{S}$.

性质 2(积分的方向性) $\iint\limits_{\Sigma^-}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}=-\iint\limits_{\Sigma^+}\boldsymbol{F}\cdot\boldsymbol{n}^{\circ}\mathrm{d}S$，其中 Σ^- 为 Σ^+ 的另一侧.

性质 3(积分的可加性) 当 $\Sigma_1\cup\Sigma_2=\Sigma,\Sigma_1\cap\Sigma_2=\varnothing$ 时，$\iint\limits_{\Sigma}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}=\iint\limits_{\Sigma_1}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}+\iint\limits_{\Sigma_2}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}$.

十一、第二型曲面积分的计算

1. 基本方法——化为二重积分

对于第二型曲面积分 $\iint\limits_{\Sigma} P(x,y,z)\mathrm{d}y\mathrm{d}z+Q(x,y,z)\mathrm{d}z\mathrm{d}x+R(x,y,z)\mathrm{d}x\mathrm{d}y$，可以将其拆成三个积分：$\iint\limits_{\Sigma} P(x,y,z)\mathrm{d}y\mathrm{d}z$，$\iint\limits_{\Sigma} Q(x,y,z)\mathrm{d}z\mathrm{d}x$，$\iint\limits_{\Sigma} R(x,y,z)\mathrm{d}x\mathrm{d}y$，分别投影到相应的坐标面上，化为二重积分计算，然后再相加．直观上，我们习惯投影到 xOy 面上去，所以以 $\iint\limits_{\Sigma} R(x,y,z)\mathrm{d}x\mathrm{d}y$ 为例．

无论空间曲面 Σ 是由显式 $z=z(x,y)$ 还是隐式 $F(x,y,z)=0$ 给出的，我们都需要做三件事（无逻辑上的先后顺序，哪件事情最利于解题就先做哪件）：

(1)将 Σ 投影到某一平面（比如 xOy 面）上⇒投影区域为 D（比如 D_{xy}）；

(2)将 $z=z(x,y)$ 或者 $F(x,y,z)=0$ 代入 $R(x,y,z)$；

(3)将 $\mathrm{d}x\mathrm{d}y$ 写成“$\pm\mathrm{d}x\mathrm{d}y$”，其中 Σ 方向向上（即法向量与 z 轴夹角为锐角）时取“+”，否则取“−”．

这就把第二型曲面积分化为了二重积分，得到

$$\iint\limits_{\Sigma} R(x,y,z)\mathrm{d}x\mathrm{d}y=\pm\iint\limits_{D_{xy}} R(x,y,z(x,y))\mathrm{d}x\mathrm{d}y.$$

同样需要指出的是，投影时 Σ 上的任何两点的投影点不能重合．

2. 高斯公式法

高斯

(1777—1855)

高斯公式 设空间有界闭区域 Ω 由有向分片光滑闭曲面 Σ 围成，$P(x,y,z)$，$Q(x,y,z)$，$R(x,y,z)$ 在 Ω 上具有一阶连续偏导数，则有公式

$$\oint\!\!\!\oint\limits_{\Sigma} P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y=\iiint\limits_{\Omega}\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}v,$$

其中，Σ 是 Ω 的整个边界曲面的外侧．

考试要点 一般来说，考试题目不可能直接满足能够使用高斯公式的条件，命题人可以“破坏”两种条件：

(1)Σ 不是封闭曲面，也就是没有围成一个空间有界闭区域 Ω；

(2)即使 Σ 围成了一个空间有界闭区域 Ω，但是 $P,Q,R,\frac{\partial P}{\partial x},\frac{\partial Q}{\partial y},\frac{\partial R}{\partial z}$ 在 Ω 上不连续．

这两种情况下，不可以直接使用高斯公式．

针对(1)，我们可以采取“补面法”，补上一片或者若干片曲面，围出一个空间有界闭区域 Ω，就可以用高斯公式了．

针对(2)，我们可以采取“挖去法”，把不连续点（称为“奇点”）挖去，使条件得以满足，从而可以使用高斯公式．

十二、空间第二型曲线积分的计算

斯托克斯公式 设 Ω 为空间某区域，Σ 为 Ω 内的分片光滑有向曲面片，l 为逐段光滑的 Σ 的边界，它

的方向与 Σ 的法向量成右手系，函数 $P(x,y,z)$，$Q(x,y,z)$ 与 $R(x,y,z)$ 在 Ω 内具有连续的一阶偏导数，则有斯托克斯公式：

$$\oint_{l} P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z=\iint_{\Sigma}\begin{vmatrix}\cos\alpha & \cos\beta & \cos\gamma\\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z}\\ P & Q & R\end{vmatrix}\mathrm{d}S.$$

【注】 可以证明（这里不证），公式的成立与绷在 l 上的曲面大小、形状无关，如图 1-18-7 所示，有$\oint_{l}=\iint_{\Sigma_1}=\iint_{\Sigma_2}$. 其中 $\boldsymbol{n}^{\circ}=(\cos\alpha,\cos\beta,\cos\gamma)$ 为 Σ 的单位外法线向量.

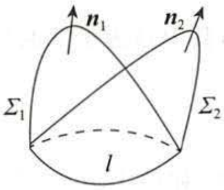


图 1-18-7

基础例题精解

一、三重积分

1. 直角坐标系与柱面坐标系下的计算

例 1.18.1 计算三重积分 $I=\iiint_{\Omega}\dfrac{\mathrm{d}x\mathrm{d}y\mathrm{d}z}{(1+x+y+z)^3}$，其中 Ω 是由平面 $x=0$，$y=0$，$z=0$ 及 $x+y+z=1$ 所围成的四面体.

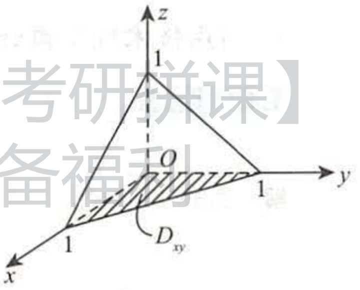


图 1-18-8

解 如图 1-18-8 所示，Ω 有下曲面（平面）$z=0$ 与上曲面（平面）$z=1-x-y$，其侧面为柱面，故可用先一后二法.

$$I=\iint_{D_{xy}}\mathrm{d}\sigma\int_0^{1-x-y}\frac{\mathrm{d}z}{(1+x+y+z)^3}=\int_0^1\mathrm{d}x\int_0^{1-x}\mathrm{d}y\int_0^{1-x-y}\frac{\mathrm{d}z}{(1+x+y+z)^3}$$

$$=\int_0^1\mathrm{d}x\int_0^{1-x}\frac{1}{2}\left[\frac{1}{(1+x+y)^2}-\frac{1}{4}\right]\mathrm{d}y$$

$$=\frac{1}{2}\int_0^1\left(\frac{1}{x+1}+\frac{x-3}{4}\right)\mathrm{d}x=\frac{1}{2}\left(\ln 2-\frac{5}{8}\right).$$

例 1.18.2 计算$\iiint_{\Omega}\mathrm{e}^{|z|}\mathrm{d}v$，其中 $\Omega:x^2+y^2+z^2\leqslant 1$.

解 被积函数仅为 z 的函数，截面 D_z 为圆域 $x^2+y^2\leqslant 1-z^2$，故采用先二后一法. 这里 Ω' 表示上半球面.

$$\iiint_{\Omega}\mathrm{e}^{|z|}\mathrm{d}v=2\iiint_{\Omega'}\mathrm{e}^{z}\mathrm{d}v=2\int_0^1\left(\iint_{D_z}\mathrm{d}x\mathrm{d}y\right)\mathrm{e}^{z}\mathrm{d}z$$

$$= 2\int_0^1 \pi(1-z^2)\mathrm{e}^z\mathrm{d}z = 2\pi.$$

【注】 一般来说，如果被积函数只依赖于一个变量，如 $f(x,y,z)=g(z)$，并假设当 z 等于常数的平面与 Ω 相截时，其截面面积易求，则可采用将三重积分化为先二重积分后定积分的计算方法.

设在 Ω 中 z 的最小值为 c_1，最大值为 c_2，用 $z=k$ 的平面与 Ω 相截，截得曲面在 xOy 坐标平面上的投影为 D_z，其面积为 A_z，则

$$\iiint_\Omega g(z)\mathrm{d}v = \int_{c_1}^{c_2} g(z)\mathrm{d}z\iint_{D_z}\mathrm{d}x\mathrm{d}y = \int_{c_1}^{c_2} g(z)A_z\mathrm{d}z.$$

例 1.18.3 计算 $I=\iiint_\Omega (x^2+y^2)\mathrm{d}v$，其中 Ω 为平面曲线 $\begin{cases} y^2=2z, \\ x=0 \end{cases}$ 绕 z 轴旋转一周形成的曲面与平面 $z=8$ 所围成的区域.

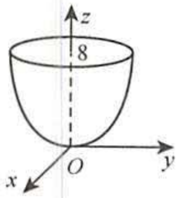


图 1-18-9

解 旋转曲面的方程为 $z=\dfrac{x^2+y^2}{2}$，因此，如图 1-18-9 所示，有

$$I=\int_0^{2\pi}\mathrm{d}\theta\int_0^4 r\mathrm{d}r\int_{\frac{r^2}{2}}^8 r^2\mathrm{d}z = 2\pi\int_0^4 r^3\left(8-\frac{r^2}{2}\right)\mathrm{d}r = \frac{1\,024}{3}\pi.$$

或 $$I=\int_0^8\mathrm{d}z\iint_{x^2+y^2\leqslant 2z}(x^2+y^2)\mathrm{d}x\mathrm{d}y = \int_0^8\mathrm{d}z\int_0^{2\pi}\mathrm{d}\theta\int_0^{\sqrt{2z}} r^3\mathrm{d}r = \frac{1\,024}{3}\pi.$$

2. 利用球面坐标计算三重积分

例 1.18.4 计算三重积分 $\iiint_\Omega (x^2+y^2)\mathrm{d}v$，其中 Ω 是右半球面 $x^2+y^2+z^2=a^2(y\geqslant 0)$ 与 xOz 面所围成的区域.

解 在球面坐标系下，积分区域可以表示为

$$\Omega=\{0\leqslant r\leqslant a, 0\leqslant\theta\leqslant\pi, 0\leqslant\varphi\leqslant\pi\},$$

所以 $$\iiint_\Omega (x^2+y^2)\mathrm{d}v = \iiint_\Omega r^2\sin^2\varphi\cdot r^2\sin\varphi\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi = \int_0^\pi\mathrm{d}\theta\int_0^\pi\mathrm{d}\varphi\int_0^a r^4\sin^3\varphi\mathrm{d}r = \int_0^\pi\mathrm{d}\theta\int_0^\pi\sin^3\varphi\left(\frac{1}{5}r^5\right)\Big|_0^a\mathrm{d}\varphi$$

$$=\frac{\pi}{5}a^5\int_0^\pi\sin^3\varphi\mathrm{d}\varphi = \frac{4}{15}\pi a^5.$$

3. 利用技术性工具计算三重积分

例 1.18.5 设 $\Omega=\{(x,y,z)\mid x^2+y^2+z^2\leqslant 1\}$，求 $\iiint_\Omega z^2\mathrm{d}x\mathrm{d}y\mathrm{d}z$.

解 由轮换对称性可知，$\iiint_\Omega z^2\mathrm{d}x\mathrm{d}y\mathrm{d}z = \iiint_\Omega x^2\mathrm{d}x\mathrm{d}y\mathrm{d}z = \iiint_\Omega y^2\mathrm{d}x\mathrm{d}y\mathrm{d}z$，所以

$$\iiint_\Omega z^2\mathrm{d}x\mathrm{d}y\mathrm{d}z = \frac{1}{3}\iiint_\Omega (x^2+y^2+z^2)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

$$=\frac{1}{3}\int_0^{2\pi}\mathrm{d}\theta\int_0^\pi\mathrm{d}\varphi\int_0^1 r^4\sin\varphi\mathrm{d}r$$

$$=\frac{2\pi}{3}\int_0^\pi\sin\varphi\mathrm{d}\varphi\int_0^1 r^4\mathrm{d}r = \frac{4\pi}{15}.$$

二、第一型曲线积分

例 1.18.6 计算 $\oint_{\Gamma}|y|\,\mathrm{d}s$，其中 Γ 为球面 $x^2+y^2+z^2=2$ 与平面 $x=y$ 的交线.

解 因为球面 $x^2+y^2+z^2=2$ 与平面 $x=y$ 的交线是一个大圆，由对称性可知，

$$\oint_{\Gamma}|y|\,\mathrm{d}s=4\int_{\Gamma_1}|y|\,\mathrm{d}s,$$

其中 Γ_1 是 Γ 在第一卦限的部分.

又由于 Γ 的参数方程为 $x=y=\cos t, z=\sqrt{2}\sin t, 0\leqslant t\leqslant 2\pi$，则

$$\mathrm{d}s=\sqrt{[x'(t)]^2+[y'(t)]^2+[z'(t)]^2}\,\mathrm{d}t=\sqrt{2}\,\mathrm{d}t.$$

故 $\oint_{\Gamma}|y|\,\mathrm{d}s=4\int_{\Gamma_1}|y|\,\mathrm{d}s=4\int_0^{\frac{\pi}{2}}\sqrt{2}\cos t\,\mathrm{d}t=4\sqrt{2}.$

例 1.18.7 计算 $I=\oint_L(x\sin\sqrt{x^2+y^2}+x^2+3y^2-5y)\,\mathrm{d}s$，其中 $L:\frac{x^2}{3}+(y-1)^2=1$，其周长为 a.

解 本题需要充分利用技术性工具化简计算，否则很难得出答案，这类题在考研中一般为填空题形式.

由对称性知，$\oint_L x\sin\sqrt{x^2+y^2}\,\mathrm{d}s=0$. 又 L 的方程可以化为 $x^2+3y^2-6y=0$，即 $x^2+3y^2-5y=y$，将边界方程代入被积函数，于是有

$$\oint_L(x^2+3y^2-5y)\,\mathrm{d}s=\oint_L y\,\mathrm{d}s=\overline{y}\cdot a=a,$$

故原式 $=a$.

三、第一型曲面积分

例 1.18.8 设曲面 $\Sigma:|x|+|y|+|z|=1$，求 $\oiint\limits_{\Sigma}(x+|y|)\,\mathrm{d}S$.

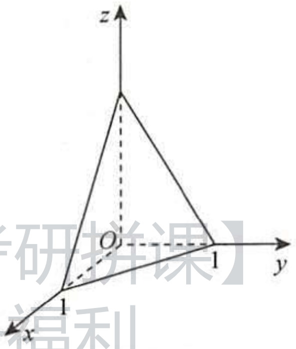


图 1-18-10

解 如图 1-18-10 所示，首先根据普通对称性，可得

$$\oiint\limits_{\Sigma}|y|\,\mathrm{d}S=8\iint\limits_{\Sigma_1}|y|\,\mathrm{d}S,$$

其中 Σ_1 是 Σ 在第一卦限的部分.

因 Σ 关于 yOz 平面对称，由普通对称性，可得 $\oiint\limits_{\Sigma}x\,\mathrm{d}S=0$.

又由轮换对称性可得

$$\oiint\limits_{\Sigma}|y|\,\mathrm{d}S=\oiint\limits_{\Sigma}|x|\,\mathrm{d}S=\oiint\limits_{\Sigma}|z|\,\mathrm{d}S=\frac{1}{3}\oiint\limits_{\Sigma}(|x|+|y|+|z|)\,\mathrm{d}S,$$

于是 $I=\oiint\limits_{\Sigma}(x+|y|)\,\mathrm{d}S=\oiint\limits_{\Sigma}|y|\,\mathrm{d}S=\frac{1}{3}\oiint\limits_{\Sigma}(|x|+|y|+|z|)\,\mathrm{d}S=\frac{8}{3}\iint\limits_{\Sigma_1}(|x|+|y|+|z|)\,\mathrm{d}S$

$$=\frac{8}{3}\iint\limits_{\Sigma_1}1\,\mathrm{d}S=\frac{8}{3}\times\frac{\sqrt{3}}{2}=\frac{4\sqrt{3}}{3}.$$

四、重积分与第一型线面积分的应用

例 1.18.9 求平面$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$被三坐标面所截部分$\Sigma$的面积，其中$a,b,c>0$.

解 平面方程可以改写为$z=c-\frac{c}{a}x-\frac{c}{b}y$，为单值函数，记$D_{xy}$为$\Sigma$在$xOy$面上的投影，则所截部分的面积

$$A=\iint\limits_{D_{xy}}\sqrt{1+(z_x')^2+(z_y')^2}\,\mathrm{d}\sigma=\iint\limits_{D_{xy}}\sqrt{1+\frac{c^2}{a^2}+\frac{c^2}{b^2}}\,\mathrm{d}\sigma$$
$$=\frac{1}{ab}\sqrt{a^2b^2+b^2c^2+a^2c^2}\iint\limits_{D_{xy}}\mathrm{d}\sigma=\frac{1}{2}\sqrt{a^2b^2+b^2c^2+a^2c^2}.$$

例 1.18.10 求锥面$z=\sqrt{x^2+y^2}$被柱面$z^2=2x$所截部分的面积.

解 联立$\begin{cases}z=\sqrt{x^2+y^2},\\z^2=2x,\end{cases}$消去$z$，得到$x^2+y^2=2x$，将其投影到$xOy$面上，得区域$D_{xy}=\{(x,y)\mid x^2+y^2\leqslant 2x\}$，其图形如图 1-18-11 所示，则所截部分的面积

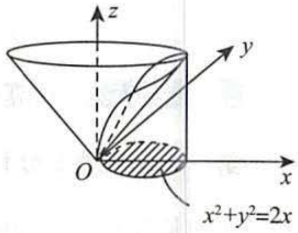


图 1-18-11

$$A=\iint\limits_{D_{xy}}\sqrt{1+(z_x')^2+(z_y')^2}\,\mathrm{d}\sigma,$$

其中$z=\sqrt{x^2+y^2}$，$z_x'=\frac{x}{\sqrt{x^2+y^2}}$，$z_y'=\frac{y}{\sqrt{x^2+y^2}}$，故$A=\iint\limits_{D_{xy}}\sqrt{2}\,\mathrm{d}\sigma=\sqrt{2}\pi$.

例 1.18.11 设空间物体$\Omega=\{(x,y,z)\mid x^2+y^2\leqslant z\leqslant 1\}$，求$\Omega$的形心的竖坐标$\bar{z}$.

分析 在重心(质心)公式中，当密度$\rho(x,y)$或者$\rho(x,y,z)$为常数时(在公式中分子分母均可以提出常数，可以约分)，重心就成为了形心，本题考查空间物体Ω的形心公式及其计算.

解 空间物体Ω在xOy面的投影为$D=\{(x,y)\mid x^2+y^2\leqslant 1\}$，则

$$\iiint\limits_{\Omega}\mathrm{d}v=\iint\limits_{D}\mathrm{d}x\mathrm{d}y\int_{x^2+y^2}^{1}\mathrm{d}z=\iint\limits_{D}(1-x^2-y^2)\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\mathrm{d}\theta\int_0^1(1-r^2)r\mathrm{d}r=\frac{\pi}{2},$$

$$\iiint\limits_{\Omega}z\mathrm{d}v=\iint\limits_{D}\mathrm{d}x\mathrm{d}y\int_{x^2+y^2}^{1}z\mathrm{d}z=\frac{1}{2}\iint\limits_{D}[1-(x^2+y^2)^2]\mathrm{d}x\mathrm{d}y=\frac{1}{2}\int_0^{2\pi}\mathrm{d}\theta\int_0^1(1-r^4)r\mathrm{d}r=\frac{\pi}{3}.$$

故

$$\bar{z}=\frac{\iiint\limits_{\Omega}z\mathrm{d}v}{\iiint\limits_{\Omega}\mathrm{d}v}=\frac{2}{3}.$$

例 1.18.12 由平面图形$1\leqslant x\leqslant 2$，$0\leqslant y\leqslant\sqrt{x}$绕$x$轴旋转所生成的旋转体$\Omega$，其密度$\rho(x,y,z)=1$，求该旋转体$\Omega$对$x$轴的转动惯量.

解 如图 1-18-12 所示，有

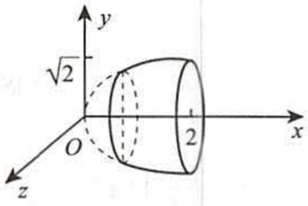


图 1-18-12

$$I_x=\iiint\limits_{\Omega}(y^2+z^2)\rho(x,y,z)\mathrm{d}v=\int_1^2\mathrm{d}x\iint\limits_{D_{yz}}(y^2+z^2)\cdot 1\mathrm{d}y\mathrm{d}z$$
$$=\int_1^2\mathrm{d}x\int_0^{2\pi}\mathrm{d}\theta\int_0^{\sqrt{x}}r^2\cdot r\mathrm{d}r=2\pi\int_1^2\frac{1}{4}(\sqrt{x})^4\mathrm{d}x=\frac{7\pi}{6}.$$

五、平面第二型曲线积分

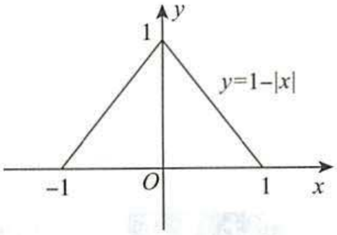


图 1-18-13

例 1.18.13　已知曲线 L 的方程为 $y=1-|x|\ (x\in[-1,1])$，起点是 $(-1,0)$，终点为 $(1,0)$，则曲线积分 $\int_L xy\mathrm{d}x+x^2\mathrm{d}y=$________．

解　应填 0.

如图 1-18-13 所示，记 $\overline{L}: y=0\ (x\in[-1,1])$，起点是 $(1,0)$，终点为 $(-1,0)$，D 是由 L 与 $\overline{L}$ 围成的平面闭区域，利用格林公式及区域 D 关于 y 轴的对称性，得

$$\int_L xy\mathrm{d}x+x^2\mathrm{d}y=\oint_{L+\overline{L}} xy\mathrm{d}x+x^2\mathrm{d}y-\int_{\overline{L}} xy\mathrm{d}x+x^2\mathrm{d}y=-\iint_D(2x-x)\mathrm{d}x\mathrm{d}y-0=0.$$

例 1.18.14　计算曲线积分 $\oint_L \dfrac{x\mathrm{d}y-y\mathrm{d}x}{4x^2+y^2}$，其中 L 是以点 $(1,0)$ 为圆心，$R\ (R>1)$ 为半径的圆周，取逆时针方向.

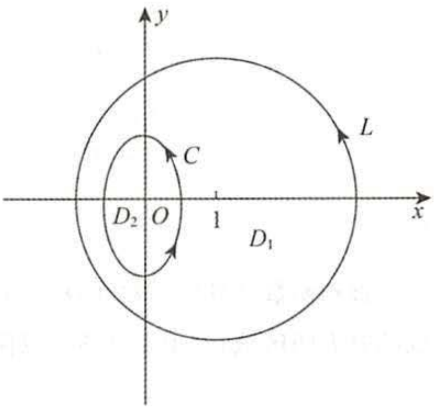


图 1-18-14

解　令 $P=\dfrac{-y}{4x^2+y^2}$，$Q=\dfrac{x}{4x^2+y^2}$，当 $(x,y)\neq(0,0)$ 时，计算可得，$\dfrac{\partial Q}{\partial x}=\dfrac{\partial P}{\partial y}$ 成立.

记曲线 L 围成的区域为 D，如图 1-18-14 所示，D 中包含点 $(0,0)$，所以 $P(x,y)$，$Q(x,y)$，$\dfrac{\partial P}{\partial y}$，$\dfrac{\partial Q}{\partial x}$ 在点 $(0,0)$ 处都不连续，这种情况下，不可以直接使用格林公式.

作足够小的椭圆 $C: 4x^2+y^2=\delta^2$（使其在 D 内部），取逆时针方向.

令 C 与 L 围成的区域为 D_1，C 围成的区域为 D_2，由格林公式，

$$\begin{aligned}I&=\oint_L P\mathrm{d}x+Q\mathrm{d}y=\oint_{L+C^-}P\mathrm{d}x+Q\mathrm{d}y-\oint_{C^-}P\mathrm{d}x+Q\mathrm{d}y\\&=\iint_{D_1}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y-\oint_{C^-}P\mathrm{d}x+Q\mathrm{d}y\\&=0-\oint_{C^-}P\mathrm{d}x+Q\mathrm{d}y=\oint_C P\mathrm{d}x+Q\mathrm{d}y\\&=\frac{1}{\delta^2}\oint_C x\mathrm{d}y-y\mathrm{d}x=\frac{1}{\delta^2}\iint_{D_2}2\mathrm{d}\sigma=\frac{1}{\delta^2}\cdot2\pi\frac{\delta}{2}\cdot\delta=\pi.\end{aligned}$$

六、第二型曲面积分

例 1.18.15　计算曲面积分 $I=\iint_\Sigma(x^3+az^2)\mathrm{d}y\mathrm{d}z+(y^3+ax^2)\mathrm{d}z\mathrm{d}x+(z^3+ay^2)\mathrm{d}x\mathrm{d}y$，其中 Σ 为上半球面 $z=\sqrt{a^2-x^2-y^2}$ 的上侧.

解　如图 1-18-15 所示，记 S 为平面 $z=0\ (x^2+y^2\leqslant a^2)$ 的下侧，Ω 为 Σ 与 S 所围成的空间区域. 则

$$I=\oiint_{\Sigma+S}(x^3+az^2)\mathrm{d}y\mathrm{d}z+(y^3+ax^2)\mathrm{d}z\mathrm{d}x+(z^3+ay^2)\mathrm{d}x\mathrm{d}y$$

$$-\iint\limits_{S}(x^3+az^2)\mathrm{d}y\mathrm{d}z+(y^3+ax^2)\mathrm{d}z\mathrm{d}x+(z^3+ay^2)\mathrm{d}x\mathrm{d}y$$

$$=\iiint\limits_{\Omega}3(x^2+y^2+z^2)\mathrm{d}v+\iint\limits_{x^2+y^2\leqslant a^2}ay^2\mathrm{d}x\mathrm{d}y$$

$$=3\int_0^{2\pi}\mathrm{d}\theta\int_0^{\frac{\pi}{2}}\sin\varphi\mathrm{d}\varphi\int_0^a r^4\mathrm{d}r+\int_0^{2\pi}a\sin^2\theta\mathrm{d}\theta\int_0^a r^3\mathrm{d}r$$

$$=\frac{6}{5}\pi a^5+\frac{1}{4}\pi a^5=\frac{29}{20}\pi a^5.$$

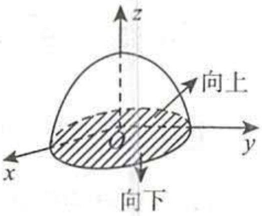


图 1-18-15

例 1.18.16 计算曲面积分 $I=\oiint\limits_{\Sigma}\frac{x\mathrm{d}y\mathrm{d}z+y\mathrm{d}z\mathrm{d}x+z\mathrm{d}x\mathrm{d}y}{(x^2+y^2+z^2)^{\frac{3}{2}}}$，其中 Σ 是曲面 $2x^2+2y^2+z^2=4$ 的外侧(如图 1-18-16).

解 先计算$\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}$.

$$\frac{\partial}{\partial x}\left[\frac{x}{(x^2+y^2+z^2)^{\frac{3}{2}}}\right]=\frac{y^2+z^2-2x^2}{(x^2+y^2+z^2)^{\frac{5}{2}}},$$

$$\frac{\partial}{\partial y}\left[\frac{y}{(x^2+y^2+z^2)^{\frac{3}{2}}}\right]=\frac{x^2+z^2-2y^2}{(x^2+y^2+z^2)^{\frac{5}{2}}},$$

$$\frac{\partial}{\partial z}\left[\frac{z}{(x^2+y^2+z^2)^{\frac{3}{2}}}\right]=\frac{x^2+y^2-2z^2}{(x^2+y^2+z^2)^{\frac{5}{2}}},$$

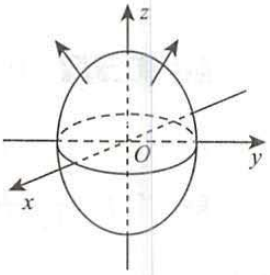


图 1-18-16

故

$$\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}=0.$$

被积函数及其偏导数在点(0,0,0)处不连续，取封闭曲面 $\Sigma_1:x^2+y^2+z^2=\delta^2$ 的外侧，Ω_0 为 Σ 与 Σ_1^- 围成的有界闭区域，其中 δ 足够小以保证其在曲面 $2x^2+2y^2+z^2=4$ 的内部，于是

$$\oiint\limits_{\Sigma}=\oiint\limits_{\Sigma+\Sigma_1^-}-\oiint\limits_{\Sigma_1^-}=\iiint\limits_{\Omega_0}0\mathrm{d}v+\oiint\limits_{\Sigma_1}\frac{x\mathrm{d}y\mathrm{d}z+y\mathrm{d}z\mathrm{d}x+z\mathrm{d}x\mathrm{d}y}{\delta^3}$$

$$=\frac{1}{\delta^3}\iiint\limits_{x^2+y^2+z^2\leqslant\delta^2}3\mathrm{d}v=\frac{3}{\delta^3}\cdot\frac{4\pi\delta^3}{3}=4\pi.$$

七、空间第二型曲线积分

例 1.18.17 计算 $\oint_{\Gamma}y\mathrm{d}x+z\mathrm{d}y+x\mathrm{d}z$，其中 Γ 为圆周 $\begin{cases}x^2+y^2+z^2=a^2,\\x+y+z=0,\end{cases}$ 若从 x 轴正向往负向看去，圆周是依逆时针的方向进行的(如图 1-18-17).

解 利用斯托克斯公式，

$$\oint_{\Gamma}y\mathrm{d}x+z\mathrm{d}y+x\mathrm{d}z=\iint\limits_{\Sigma}\begin{vmatrix}\cos\alpha & \cos\beta & \cos\gamma\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ y & z & x\end{vmatrix}\mathrm{d}S$$

$$=-\iint\limits_{\Sigma}(\cos\alpha+\cos\beta+\cos\gamma)\mathrm{d}S,$$

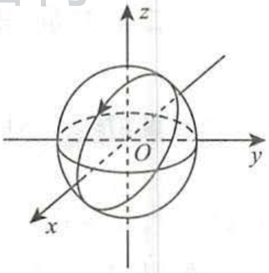


图 1-18-17

其中，Σ 为平面 $x+y+z=0$ 上以圆周 Γ 为边界的圆域，并且 Σ 的法线与 x 轴成锐角，因此

$$\cos\alpha=\cos\beta=\cos\gamma=\frac{1}{\sqrt{3}},$$

故原式 $=-\iint\limits_{\Sigma}\frac{3}{\sqrt{3}}\mathrm{d}S=-\sqrt{3}\iint\limits_{\Sigma}\mathrm{d}S=-\sqrt{3}\pi a^2$.

基础习题精练

习 题

1.18.1 计算三重积分 $\iiint\limits_{\Omega}(x^2+y^2)\mathrm{d}v$，其中 Ω 是由椭圆抛物面 $z=4(x^2+y^2)$ 和平面 $z=4$ 所围成的区域.

1.18.2 在直角坐标系下计算下列三重积分：

(1) $\iiint\limits_{\Omega}(x^2+y^2+z^2)\mathrm{d}x\mathrm{d}y\mathrm{d}z$，其中 Ω：$0\leqslant x\leqslant 1,0\leqslant y\leqslant 1,0\leqslant z\leqslant 1$；

(2) $\iiint\limits_{\Omega}z\mathrm{d}x\mathrm{d}y\mathrm{d}z$，其中 Ω 为三个坐标面及平面 $x+y+z=1$ 所围成的闭区域.

1.18.3 计算 $\int_{\Gamma}(x^2+y^2+z^2)\mathrm{d}s$，其中 Γ：$x=\cos t,y=\sin t,z=t,0\leqslant t\leqslant 2\pi$.

1.18.4 设平面薄片所占的区域 D 由抛物线 $y=x^2$ 及直线 $y=x$ 所围成，它在 (x,y) 处的面密度 $\rho(x,y)=x^2y$，求此薄片的重心.

1.18.5 设 L 为取正向的圆 $x^2+y^2=9$，则曲线积分 $\oint_{L}(2xy-2y)\mathrm{d}x+(x^2-4x)\mathrm{d}y$ 的值是________.

1.18.6 设 Σ 为曲面 $x^2+y^2+z^2=1$ 的外侧，计算曲面积分

$$I=\oiint\limits_{\Sigma}x^3\mathrm{d}y\mathrm{d}z+y^3\mathrm{d}z\mathrm{d}x+z^3\mathrm{d}x\mathrm{d}y.$$

1.18.7 已知曲线 L 的方程为 $\begin{cases}z=\sqrt{2-x^2-y^2},\\ z=x,\end{cases}$ 起点为 $A(0,\sqrt{2},0)$，终点为 $B(0,-\sqrt{2},0)$，计算曲线积分

$$I=\int_{L}(y+z)\mathrm{d}x+(z^2-x^2+y)\mathrm{d}y+x^2y^2\mathrm{d}z.$$

解 答

1.18.1 解 如图 1-18-18 所示，积分区域 Ω 在 xOy 面上的投影是一个圆心在原点的单位圆，所以

$$\Omega=\{(r,\theta,z)\mid 0\leqslant r\leqslant 1,0\leqslant\theta\leqslant 2\pi,4r^2\leqslant z\leqslant 4\}.$$

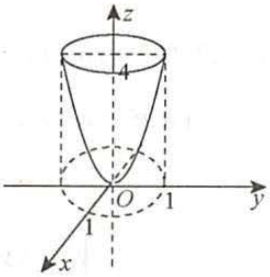


图 1-18-18

于是
$$\iiint\limits_{\Omega}(x^2+y^2)\mathrm{d}v=\iiint\limits_{\Omega}r^2\cdot r\mathrm{d}r\mathrm{d}\theta\mathrm{d}z=\int_0^{2\pi}\mathrm{d}\theta\int_0^1 r^2\cdot r\mathrm{d}r\int_{4r^2}^{4}\mathrm{d}z$$
$$=\int_0^{2\pi}\mathrm{d}\theta\int_0^1(4r^3-4r^5)\mathrm{d}r=\frac{2}{3}\pi.$$

1.18.2　解　(1)由轮换对称性，有
$$\iiint\limits_{\Omega}(x^2+y^2+z^2)\mathrm{d}x\mathrm{d}y\mathrm{d}z=3\int_0^1\mathrm{d}x\int_0^1\mathrm{d}y\int_0^1 z^2\mathrm{d}z=1.$$

(2)
$$\iiint\limits_{\Omega}z\mathrm{d}x\mathrm{d}y\mathrm{d}z=\int_0^1\mathrm{d}x\int_0^{1-x}\mathrm{d}y\int_0^{1-x-y}z\mathrm{d}z$$
$$=\frac{1}{2}\int_0^1\mathrm{d}x\int_0^{1-x}(1-x-y)^2\mathrm{d}y=-\frac{1}{6}\int_0^1(1-x-y)^3\Big|_0^{1-x}\mathrm{d}x$$
$$=\frac{1}{6}\int_0^1(1-x)^3\mathrm{d}x=\frac{1}{24}.$$

1.18.3　解　$x^2+y^2+z^2=\cos^2t+\sin^2t+t^2=1+t^2,\mathrm{d}s=\sqrt{(-\sin t)^2+\cos^2t+1}\mathrm{d}t=\sqrt{2}\mathrm{d}t.$

故
$$\int_{\Gamma}(x^2+y^2+z^2)\mathrm{d}s=\int_0^{2\pi}(1+t^2)\sqrt{2}\mathrm{d}t=\sqrt{2}\left(2\pi+\frac{8\pi^3}{3}\right).$$

1.18.4　解　设此薄片的重心为$(\overline{x},\overline{y})$，则
$$\overline{x}=\frac{\iint\limits_{D}x\rho(x,y)\mathrm{d}\sigma}{\iint\limits_{D}\rho(x,y)\mathrm{d}\sigma}=\frac{\int_0^1\mathrm{d}x\int_{x^2}^{x}x^3y\mathrm{d}y}{\int_0^1\mathrm{d}x\int_{x^2}^{x}x^2y\mathrm{d}y}=\frac{\frac{1}{48}}{\frac{1}{35}}=\frac{35}{48},$$
$$\overline{y}=\frac{\iint\limits_{D}y\rho(x,y)\mathrm{d}\sigma}{\iint\limits_{D}\rho(x,y)\mathrm{d}\sigma}=\frac{\int_0^1\mathrm{d}x\int_{x^2}^{x}x^2y^2\mathrm{d}y}{\frac{1}{35}}=\frac{\frac{1}{54}}{\frac{1}{35}}=\frac{35}{54}.$$

故此薄片的重心为$\left(\frac{35}{48},\frac{35}{54}\right)$.

1.18.5　-18π　**解**　由格林公式可知
$$\text{原式}=\iint\limits_{D}[2x-4-(2x-2)]\mathrm{d}x\mathrm{d}y=-2\iint\limits_{D}\mathrm{d}x\mathrm{d}y=-2\times9\pi=-18\pi,$$
其中$D=\{(x,y)\mid x^2+y^2\leqslant9\}$.

1.18.6　解　由高斯公式，并利用球面坐标计算三重积分，得
$$I=3\iiint\limits_{\Omega}(x^2+y^2+z^2)\mathrm{d}v(\Omega\text{是由}\Sigma\text{所围成的区域})$$
$$=3\int_0^{2\pi}\mathrm{d}\theta\int_0^{\pi}\sin\varphi\mathrm{d}\varphi\int_0^1 r^2\cdot r^2\mathrm{d}r=\frac{12}{5}\pi.$$

1.18.7　解　设L_1是从点B到点A的直线段，Σ为平面$z=x$上由L与L_1围成的半圆面下侧，其法向量的单位向量为$\left(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}}\right)$.

由斯托克斯公式
$$\oint_{L+L_1}(y+z)\mathrm{d}x+(z^2-x^2+y)\mathrm{d}y+x^2y^2\mathrm{d}z=\iint\limits_{\Sigma}\begin{vmatrix}\frac{1}{\sqrt{2}}&0&-\frac{1}{\sqrt{2}}\\\frac{\partial}{\partial x}&\frac{\partial}{\partial y}&\frac{\partial}{\partial z}\\y+z&z^2-x^2+y&x^2y^2\end{vmatrix}\mathrm{d}S$$

$$=\frac{1}{\sqrt{2}}\iint_{\Sigma}(2x^2y+1)\mathrm{d}S.$$

由于曲面 Σ 关于 xOz 平面对称，所以 $\iint_{\Sigma}2x^2y\mathrm{d}S=0$，故

$$\oint_{L+L_1}(y+z)\mathrm{d}x+(z^2-x^2+y)\mathrm{d}y+x^2y^2\mathrm{d}z=\frac{1}{\sqrt{2}}\iint_{\Sigma}\mathrm{d}S=\frac{\sqrt{2}}{2}\pi.$$

又 L_1 的参数方程为 $x=0,y=y,z=0$（y 从 $-\sqrt{2}$ 到 $\sqrt{2}$），所以

$$\int_{L_1}(y+z)\mathrm{d}x+(z^2-x^2+y)\mathrm{d}y+x^2y^2\mathrm{d}z=\int_{-\sqrt{2}}^{\sqrt{2}}y\mathrm{d}y=0.$$

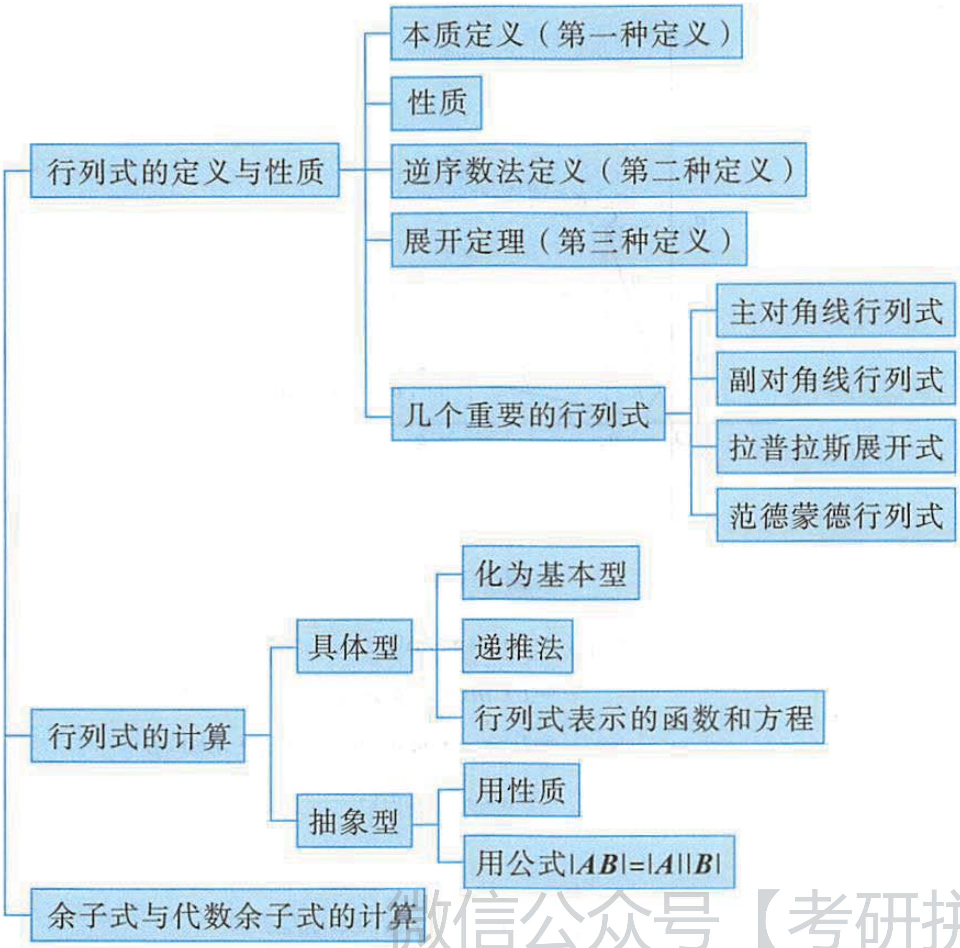
因此 $I=\oint_{L+L_1}-\int_{L_1}=\frac{\sqrt{2}}{2}\pi.$

第二部分

线性代数

第1讲 行列式

基础知识结构



基础内容精讲

一、行列式的本质定义（第一种定义）

大多数教材都是从“纯粹”的代数学角度来定义行列式的，较为抽象，难于理解和接受．我们先详细通俗地给出行列式的本质定义，并且告诉读者，我们将要开始的线性代数这门课程到底要学什么．

先看一个式子：$D_2=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}$．我们称其为 2 阶行列式，其中 a_{ij} 的第一个下标 i 表示此元素所在的行

数，第二个下标 j 表示此元素所在的列数，$i=1,2,j=1,2$，于是此行列式中有四个元素，并且 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$. 这是一个什么样的计算规则？它背后有什么样的意义？

希望读者跟着我的思路走.请你将此行列式的第 1 行的两个元素 a_{11},a_{12} 看成一个 2 维向量 $[a_{11},a_{12}]\overset{\text{记}}{=\!=}\boldsymbol{\alpha}_1$（线性代数中，向量不需要在字母上加箭头写成 $\overrightarrow{\alpha_1}$，只要写 $\boldsymbol{\alpha}_1$ 即可，此后类同，不再重复）. 将此行列式的第 2 行的两个元素 a_{21},a_{22} 看成另一个 2 维向量 $[a_{21},a_{22}]\overset{\text{记}}{=\!=}\boldsymbol{\alpha}_2$. 不失一般性，将其标在直角坐标系中，且以这两个向量为邻边拼出一个 $\square OABC$，则 $S_{\square OABC}=?$

不妨设 $\boldsymbol{\alpha}_1$ 的长度（模）为 l，$\boldsymbol{\alpha}_2$ 的长度（模）为 m，$\boldsymbol{\alpha}_1$ 与 x 轴正向的夹角为 α，$\boldsymbol{\alpha}_2$ 与 x 轴正向的夹角为 β，于是，如图 2-1-1 所示：

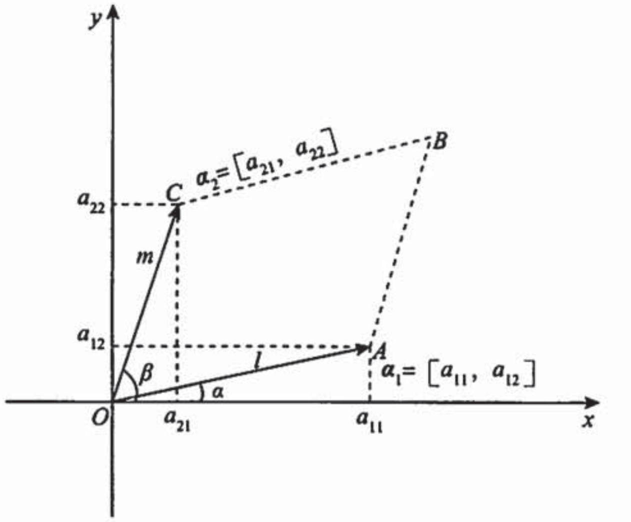


图 2-1-1

则
$$\begin{aligned} S_{\square OABC} &= l\cdot m\cdot \sin(\beta-\alpha) \\ &= l\cdot m(\sin\beta\cos\alpha-\cos\beta\sin\alpha) \\ &= l\cos\alpha\cdot m\sin\beta-l\sin\alpha\cdot m\cos\beta \\ &= a_{11}a_{22}-a_{12}a_{21}, \end{aligned}$$

于是
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}=S_{\square OABC}.$$

我们看到了一个极其直观有趣的结论：**2 阶行列式是由两个 2 维向量组成的，其（运算规则的）结果为以这两个向量为邻边的平行四边形的面积**. 这不仅得出了 2 阶行列式的计算规则，也能够清楚地看到其几何意义.

线性代数这门学问最大的一个特点是"可以作线性推广"——3 阶行列式 $D_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ 是什么？

我希望读者能够仿照上述定义回答出：**3 阶行列式是由三个 3 维向量 $\boldsymbol{\alpha}_1=[a_{11},a_{12},a_{13}]$，$\boldsymbol{\alpha}_2=[a_{21},a_{22},a_{23}]$，$\boldsymbol{\alpha}_3=[a_{31},a_{32},a_{33}]$ 组成的，其（运算规则的）结果为以这三个向量为邻边的平行六面体的体积.** 如图 2-1-2 所示.

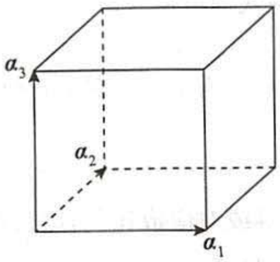


图 2-1-2

依此类推，我们便可以给出 n 阶行列式 $D_n=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$ 的本质定义：

n 阶行列式是由 n 个 n 维向量 $\boldsymbol{\alpha}_1=[a_{11},a_{12},\cdots,a_{1n}]$，$\boldsymbol{\alpha}_2=[a_{21},a_{22},\cdots,a_{2n}]$，$\cdots$，$\boldsymbol{\alpha}_n=[a_{n1},a_{n2},\cdots,a_{nn}]$组成的，其（运算规则的）结果为以这 n 个向量为邻边的 n 维图形的体积.

由此看来，一个重要观点出现了：读者一开始，就应该**把行列式看作是由若干个向量拼成的，并且要把这些向量作运算.** 以 3 阶行列式为例，若 $D_3\neq0$，则意味着体积不为 0，则称组成该行列式的三个向量线性无关；若 $D_3=0$，称线性相关.

二、行列式的性质

性质 1 行列互换，其值不变，即 $|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|$.

性质 2 行列式中某行(列)元素全为零，则行列式为零.

性质 3 行列式中某行(列)元素有公因子 $k(k\neq0)$，则 k 可提到行列式外面，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=k\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

【注】 (1)本书用 kⓘ表示第 i 行乘 k，$k[j]$表示第 j 列乘 k.

(2)以后称上述等式从右到左的运算为“**倍乘**”性质.

性质 4 行列式中某行(列)元素均是两个元素之和，则可拆成两个行列式之和，即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

【注】 等式从右到左是两个行列式相加的运算. 如果两个行列式的其他元素对应相等，只有一行(列)不同时，可以相加，相加时其他元素不变，不同元素的行(列)对应相加即可.

性质 5 行列式中两行(列)互换，行列式的值反号.

【注】 (1)以后用ⓘ↔ⓙ表示第 i 行与第 j 行互换，$[i]\leftrightarrow[j]$表示第 i 列与第 j 列互换.

(2)以后称上述运算为“**互换**”性质.

性质 6 行列式中的两行(列)元素相等或对应成比例，则行列式为零.

性质 7 行列式中某行(列)的 k 倍加到另一行(列)，行列式的值不变.

【注 1】 (1)以后用ⓘ$+k$ⓙ表示第 j 行的 k 倍加到第 i 行，$[i]+k[j]$表示第 j 列的 k 倍加到第 i 列.

(2)以后称上述运算为“**倍加**”性质.

【注 2】 以上七个性质均可由本讲“一、行列式的本质定义(第一种定义)”所介绍的行列式的几何背景直观地得到，而不需复杂抽象的分析. 如性质 6 所说到的“两行(列)元素对应成比例，则行列式为零”，可取 $\begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix}$，因向量[2,3]与向量[4,6]为平行向量，故 $\begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix}=S_{\square}=0$，如图 2-1-3 所示，一目了然.

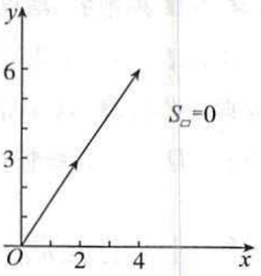


图 2-1-3

三、行列式的逆序数法定义(第二种定义)

1. 排列和逆序

排列 由 n 个数 $1,2,\cdots,n$ 组成的一个有序数组称为一个 n 级排列，如 23145 是一个 5 级排列，41352 也是一个 5 级排列. n 级排列共有 $n!$ 个.

逆序 在一个 n 级排列 $i_1i_2\cdots i_s\cdots i_t\cdots i_n$ 中，若 $i_s>i_t$，且 i_s 排在 i_t 前面，则称这两个数构成一个逆序.

逆序数 一个排列中，逆序的总数称为该排列的逆序数，记作 $\tau(i_1i_2\cdots i_n)$，如 $\tau(231546)=3$，$\tau(621534)=8$. 由小到大顺排的排列称为自然排序，如 12345，显然，自然排序的逆序数为 0.

奇排列和偶排列 排列的逆序数为奇数时，该排列称为奇排列；排列的逆序数为偶数时，该排列称为偶排列.

2. n 阶行列式的定义

$n(n\geqslant 2)$阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

这里 $\sum\limits_{j_1j_2\cdots j_n}$ 表示对所有 n 个列下标排列求和，故为 $n!$ 项之和. 注意到行下标已经顺排，而列下标是任一个 n 级排列，故每项由取自不同行、不同列的 n 个元素的乘积组成，每项的正负号取决于$(-1)^{\tau(j_1j_2\cdots j_n)}$. 当列

下标为奇排列时，应附加负号；当列下标为偶排列时，应附加正号.

【注】 (1)规定 1 阶行列式 $|a_{11}|=a_{11}$.

(2)如：请确定“$a_{12}a_{31}a_{54}a_{43}a_{25}$”这一展开项前的正负号. 答：首先将行下标顺排为 $a_{12}a_{25}a_{31}a_{43}a_{54}$，然后计算 $\tau(25134)=4$，为偶排列，故该项前为正号.

(3)上述 n 阶行列式利用逆序的定义和教材中对于 2,3 阶行列式的定义是完全一致的. 也就是

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}=ad-bc,$$

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=a_1b_2c_3+a_2b_3c_1+a_3b_1c_2-a_1b_3c_2-a_2b_1c_3-a_3b_2c_1.$$

如

$$\begin{vmatrix} 5 & 2 & 1 \\ 1 & 2 & 5 \\ 34 & 1 & 34 \end{vmatrix}=520,\quad \begin{vmatrix} \text{我} & 0 & \text{生} \\ 0 & \text{有} & 0 \\ \text{你} & 0 & \text{幸} \end{vmatrix}=\text{我有幸一生有你},$$

读者若将上式结果中的减号“—”看作汉字“一”，一个浪漫的公式便产生了.

四、行列式的展开定理(第三种定义)

阶数超过 3 的行列式，若还用“一”“三”的方法，就太麻烦了，为此，提出行列式的展开定理.

1. 余子式

在 n 阶行列式中，去掉元素 a_{ij} 所在的第 i 行、第 j 列元素，由剩下的元素按原来的位置与顺序组成的 $n-1$阶行列式称为元素 a_{ij} 的余子式，记作 M_{ij}，即

$$M_{ij}=\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

2. 代数余子式

余子式 M_{ij} 乘$(-1)^{i+j}$后称为 a_{ij} 的代数余子式，记作 A_{ij}，即

$$A_{ij}=(-1)^{i+j}M_{ij},$$

显然也有 $M_{ij}=(-1)^{i+j}A_{ij}$.

3. 行列式按某一行(列)展开的展开公式

行列式的值等于行列式的某行(列)元素分别乘其相应的代数余子式后再求和，即

$$|\boldsymbol{A}|=\begin{cases} a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}=\sum\limits_{j=1}^{n}a_{ij}A_{ij}\,(i=1,2,\cdots,n), \\ a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}=\sum\limits_{i=1}^{n}a_{ij}A_{ij}\,(j=1,2,\cdots,n). \end{cases}$$

但行列式的某行(列)元素分别乘另一行(列)元素的代数余子式后再求和，结果为零，即

$$a_{i1}A_{k1}+a_{i2}A_{k2}+\cdots+a_{in}A_{kn}=0,i\neq k;$$

$$a_{1j}A_{1k}+a_{2j}A_{2k}+\cdots+a_{nj}A_{nk}=0,j\neq k.$$

【注】 余子式与代数余子式是行列式展开定理的核心概念，关于它们的灵活使用请参看例2.1.13.

五、几个重要的行列式

1. 主对角线行列式(上(下)三角形行列式)

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}=\prod_{i=1}^{n}a_{ii}.$$

2. 副对角线行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & 0 & \cdots & 0 & 0 \end{vmatrix}=\begin{vmatrix} 0 & \cdots & 0 & a_{1n} \\ 0 & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix}=\begin{vmatrix} 0 & \cdots & 0 & a_{1n} \\ 0 & \cdots & a_{2,n-1} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & 0 & 0 \end{vmatrix}$$

$$=(-1)^{\frac{n(n-1)}{2}}a_{1n}a_{2,n-1}\cdots a_{n1}.$$

3. 拉普拉斯展开式

设 $\boldsymbol{A}$ 为 m 阶矩阵，$\boldsymbol{B}$ 为 n 阶矩阵，则

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{vmatrix}=\begin{vmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{vmatrix}=\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{vmatrix}=|\boldsymbol{A}||\boldsymbol{B}|,$$

$$\begin{vmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{vmatrix}=\begin{vmatrix} \boldsymbol{C} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{vmatrix}=\begin{vmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{C} \end{vmatrix}=(-1)^{mn}|\boldsymbol{A}||\boldsymbol{B}|.$$

【注】 以后称以上 12 个行列式为“基本形”行列式，加上下面“4. 范德蒙德行列式”简称“12＋1”型行列式.

4. 范德蒙德行列式

记

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

基础例题精解

一、具体型行列式的计算

1. 化基本形法
   - 直接展开
   - 爪形
   - 异爪形
   - 行(列)和相等
   - 消零化基本形
   - 拉普拉斯展开
   - 范德蒙德行列式

用行列式的性质或展开公式，化为“基本形”行列式.

例 2.1.1　n 阶行列式 $\begin{vmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ b & 0 & 0 & \cdots & 0 & a \end{vmatrix}_{n\times n}$ $=$______.

解　应填 $a^n+(-1)^{n+1}b^n$.

按第 1 列展开，得

$$D_n=\begin{vmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ b & 0 & 0 & \cdots & 0 & a \end{vmatrix}_{n\times n}$$

$$=a\begin{vmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ 0 & 0 & 0 & \cdots & 0 & a \end{vmatrix}_{(n-1)\times(n-1)}+(-1)^{n+1}b\begin{vmatrix} b & 0 & \cdots & 0 & 0 \\ a & b & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & b & 0 \\ 0 & 0 & \cdots & a & b \end{vmatrix}_{(n-1)\times(n-1)}$$

$$=a^n+(-1)^{n+1}b^n.$$

例 2.1.2　计算行列式：$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 4 \end{vmatrix}$.

解

$$\begin{vmatrix}1&1&1&1\\1&2&0&0\\1&0&3&0\\1&0&0&4\end{vmatrix}=2\times3\times4\begin{vmatrix}1&1&1&1\\\frac{1}{2}&1&0&0\\\frac{1}{3}&0&1&0\\\frac{1}{4}&0&0&1\end{vmatrix}=24\begin{vmatrix}1-\frac{1}{2}-\frac{1}{3}-\frac{1}{4}&0&0&0\\\frac{1}{2}&1&0&0\\\frac{1}{3}&0&1&0\\\frac{1}{4}&0&0&1\end{vmatrix}$$

$$=24\times\left(1-\frac{1}{2}-\frac{1}{3}-\frac{1}{4}\right)=24-12-8-6=-2.$$

【注】 本题的行列式是"爪形行列式"(只有第 1 列、第 1 行及主对角线元素不为零,其余元素均为零的行列式),这种行列式都可以化为"基本形"行列式.

例 2.1.3 行列式 $\begin{vmatrix}\lambda&-1&0&0\\0&\lambda&-1&0\\0&0&\lambda&-1\\4&3&2&\lambda+1\end{vmatrix}=$________.

解 应填 $4+3\lambda+2\lambda^2+\lambda^3+\lambda^4$.

按第 4 行展开,得

$$D=4(-1)^{4+1}\begin{vmatrix}-1&0&0\\\lambda&-1&0\\0&\lambda&-1\end{vmatrix}+3(-1)^{4+2}\begin{vmatrix}\lambda&0&0\\0&-1&0\\0&\lambda&-1\end{vmatrix}+2(-1)^{4+3}\begin{vmatrix}\lambda&-1&0\\0&\lambda&0\\0&0&-1\end{vmatrix}+$$

$$(\lambda+1)(-1)^{4+4}\begin{vmatrix}\lambda&-1&0\\0&\lambda&-1\\0&0&\lambda\end{vmatrix}=4+3\lambda+2\lambda^2+\lambda^3+\lambda^4.$$

例 2.1.4 计算 n 阶行列式:

$$D_n=\begin{vmatrix}a&b&b&\cdots&b\\b&a&b&\cdots&b\\b&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\b&b&b&\cdots&a\end{vmatrix}.$$

解 将第 2,3,…,n 列加到第 1 列,则可提出公因子,即

$$D_n=\begin{vmatrix}a&b&b&\cdots&b\\b&a&b&\cdots&b\\b&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\b&b&b&\cdots&a\end{vmatrix}\xlongequal{[1]+\sum_{i=2}^{n}[i]}[a+(n-1)b]\begin{vmatrix}1&b&b&\cdots&b\\1&a&b&\cdots&b\\1&b&a&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\1&b&b&\cdots&a\end{vmatrix}$$

$$\overset{\text{ⓘ}-\text{①}}{\underset{i=2,3,\cdots,n}{=\!=\!=\!=\!=}}[a+(n-1)b]\begin{vmatrix}1 & b & b & \cdots & b\\0 & a-b & 0 & \cdots & 0\\0 & 0 & a-b & \cdots & 0\\\vdots & \vdots & \vdots & & \vdots\\0 & 0 & 0 & \cdots & a-b\end{vmatrix}=[a+(n-1)b](a-b)^{n-1}.$$

【注 1】 行和或列和相等的行列式(行和是指每一行元素相加的和，列和同理)

$$\begin{vmatrix}a & b & b & \cdots & b\\b & a & b & \cdots & b\\b & b & a & \cdots & b\\\vdots & \vdots & \vdots & & \vdots\\b & b & b & \cdots & a\end{vmatrix}_{n\times n}=[a+(n-1)b](a-b)^{n-1}.$$

【注 2】 行列式中每行(列)元素之和相等时，将各列(行)加到第 1 列(行)，然后提出公因子是可取的方法.

【注 3】 这类字母抽象型行列式显然具有代表性. 如

(1)当 $a=0,b=1$ 时，$\begin{vmatrix}0 & 1 & 1 & \cdots & 1\\1 & 0 & 1 & \cdots & 1\\1 & 1 & 0 & \cdots & 1\\\vdots & \vdots & \vdots & & \vdots\\1 & 1 & 1 & \cdots & 0\end{vmatrix}_{n\times n}=(n-1)(-1)^{n-1}$;

当 $a=2,b=1$ 时，$\begin{vmatrix}2 & 1 & 1 & \cdots & 1\\1 & 2 & 1 & \cdots & 1\\1 & 1 & 2 & \cdots & 1\\\vdots & \vdots & \vdots & & \vdots\\1 & 1 & 1 & \cdots & 2\end{vmatrix}_{n\times n}=n+1$.

(2)当 $a=x$ 时，$\begin{vmatrix}x & b & b & \cdots & b\\b & x & b & \cdots & b\\b & b & x & \cdots & b\\\vdots & \vdots & \vdots & & \vdots\\b & b & b & \cdots & x\end{vmatrix}_{n\times n}=[x+(n-1)b](x-b)^{n-1}$.

若视 x 为变量，b 是常数，则行列式是 x 的 n 次多项式，其根是 $x_1=x_2=\cdots=x_{n-1}=b,x_n=(1-n)b$.

当 a 取为 $\lambda-a$ 时，$\begin{vmatrix}\lambda-a & b & b & \cdots & b\\b & \lambda-a & b & \cdots & b\\b & b & \lambda-a & \cdots & b\\\vdots & \vdots & \vdots & & \vdots\\b & b & b & \cdots & \lambda-a\end{vmatrix}_{n\times n}=[\lambda-a+(n-1)b](\lambda-a-b)^{n-1}$.

若视 λ 为变量，a,b 为常数，则行列式是 λ 的 n 次多项式，其根是 $\lambda_1=\lambda_2=\cdots=\lambda_{n-1}=a+b,\lambda_n=a-(n-1)b$.

(3)当 a 在副对角线上时，$G_n=\begin{vmatrix} b & b & \cdots & b & a \\ b & b & \cdots & a & b \\ \vdots & \vdots & & \vdots & \vdots \\ b & a & \cdots & b & b \\ a & b & \cdots & b & b \end{vmatrix}\xlongequal{(*)}(-1)^{\frac{n(n-1)}{2}}\begin{vmatrix} a & b & \cdots & b & b \\ b & a & \cdots & b & b \\ \vdots & \vdots & & \vdots & \vdots \\ b & b & \cdots & a & b \\ b & b & \cdots & b & a \end{vmatrix}$

$$=(-1)^{\frac{n(n-1)}{2}}[a+(n-1)b](a-b)^{n-1}.$$

$(*)$处是将最后 1 列和前面相邻列对换，对换 $n-1$ 次到第 1 列，再将最新的行列式的最后 1 列和相邻列对换，对换 $n-2$ 次到第 2 列，……，直到换成 D_n，共交换 $(n-1)+(n-2)+\cdots+1=\frac{n(n-1)}{2}$ 次，故得

$$G_n=(-1)^{\frac{n(n-1)}{2}}[a+(n-1)b](a-b)^{n-1}.$$

注意，上述结果不要当作公式去记忆，而是要学会分析行列式中元素分布的规律性，并掌握相应的计算方法.

例 2.1.5 计算行列式

$$D_4=\begin{vmatrix} a_1 & 0 & 0 & b_1 \\ 0 & a_2 & b_2 & 0 \\ 0 & b_3 & a_3 & 0 \\ b_4 & 0 & 0 & a_4 \end{vmatrix}.$$

解 先将第 2,4 列互换，再将第 2,4 行互换，最后利用拉普拉斯展开式.

$$D_4=\begin{vmatrix} a_1 & 0 & 0 & b_1 \\ 0 & a_2 & b_2 & 0 \\ 0 & b_3 & a_3 & 0 \\ b_4 & 0 & 0 & a_4 \end{vmatrix}=(-1)\begin{vmatrix} a_1 & b_1 & 0 & 0 \\ 0 & 0 & b_2 & a_2 \\ 0 & 0 & a_3 & b_3 \\ b_4 & a_4 & 0 & 0 \end{vmatrix}=(-1)\times(-1)\begin{vmatrix} a_1 & b_1 & 0 & 0 \\ b_4 & a_4 & 0 & 0 \\ 0 & 0 & a_3 & b_3 \\ 0 & 0 & b_2 & a_2 \end{vmatrix}$$

$$=\begin{vmatrix} a_1 & b_1 \\ b_4 & a_4 \end{vmatrix}\begin{vmatrix} a_3 & b_3 \\ b_2 & a_2 \end{vmatrix}=(a_1a_4-b_1b_4)(a_2a_3-b_2b_3)$$

$$=a_1a_2a_3a_4+b_1b_2b_3b_4-a_1b_2b_3a_4-b_1a_2a_3b_4.$$

【注】 注意：

$$D_4=\begin{vmatrix} a_1 & 0 & 0 & b_1 \\ 0 & a_2 & b_2 & 0 \\ 0 & b_3 & a_3 & 0 \\ b_4 & 0 & 0 & a_4 \end{vmatrix}=a_1a_2a_3a_4-b_1b_2b_3b_4$$

是错误的．不要当作 2 阶行列式计算．展开式共有 4 项，且副对角线元素乘积的那一项取正号.

例 2.1.6 计算行列式 $\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ b+c & a+c & a+b \end{vmatrix}$.

解

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ b+c & a+c & a+b \end{vmatrix} = \begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a+b+c & a+b+c & a+b+c \end{vmatrix}$$

$$=(a+b+c)\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ 1 & 1 & 1 \end{vmatrix}$$

$$=(a+b+c)(-1)\begin{vmatrix} a & b & c \\ 1 & 1 & 1 \\ a^2 & b^2 & c^2 \end{vmatrix}$$

$$=(a+b+c)(-1)^2\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}$$

$$=(a+b+c)(b-a)(c-b)(c-a).$$

2. 递推法

例 2.1.7 计算：

$$D_n=\begin{vmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix}_{n\times n}.$$

解 将 D_n 中的第 2,3,…,n 列加到第 1 列，再按第 1 列展开，得

$$D_n=\begin{vmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 1 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix}_{n\times n}$$

$$=D_{n-1}+(-1)^{n+1}\begin{vmatrix} -1 & 0 & 0 & \cdots & 0 & 0 \\ 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & \cdots & 2 & -1 \end{vmatrix}_{(n-1)\times(n-1)}$$

$$=D_{n-1}+(-1)^{n+1}(-1)^{n-1}=D_{n-1}+1.$$

得递推关系 $D_n=D_{n-1}+1$，则 $D_n=D_{n-1}+1=D_{n-2}+1+1=D_{n-2}+2=\cdots$，因此可得

$$D_n=D_1+n-1=2+n-1=n+1.$$

3. 行列式表示的函数和方程

这类问题的行列式元素 a_{ij} 往往不是具体数值，而是含 x 或 λ 等的函数，可能会在计算之外，给读者带来新的困难和麻烦，自然也会给命题人带来新的角度. 请读者重视对此类问题的研究.

例 2.1.8 设 $f(x)=\begin{vmatrix}1&0&x\\1&2&x^2\\1&3&x^3\end{vmatrix}$，求 $f(x+1)-f(x)$.

解

$$\begin{aligned}f(x+1)-f(x)&=\begin{vmatrix}1&0&x+1\\1&2&(x+1)^2\\1&3&(x+1)^3\end{vmatrix}-\begin{vmatrix}1&0&x\\1&2&x^2\\1&3&x^3\end{vmatrix}\\&=\begin{vmatrix}1&0&x+1\\1&2&x^2+2x+1\\1&3&x^3+3x^2+3x+1\end{vmatrix}-\begin{vmatrix}1&0&x\\1&2&x^2\\1&3&x^3\end{vmatrix}\\&=\begin{vmatrix}1&0&1\\1&2&2x+1\\1&3&3x^2+3x+1\end{vmatrix}=\begin{vmatrix}1&0&0\\1&2&0\\1&3&3x^2\end{vmatrix}=6x^2.\end{aligned}$$

【注】 函数可以用含有变量的行列式表示，因此，对这类行列式当然也可以求极限、导数、积分等，如本题 $[f(x+1)-f(x)]'=(6x^2)'=12x$，$\int[f(x+1)-f(x)]\mathrm{d}x=2x^3+C$.

例 2.1.9 设 $f(x)=\begin{vmatrix}x-2&x-1&x-2&x-3\\2x-2&2x-1&2x-2&2x-3\\3x-3&3x-2&4x-5&3x-5\\4x&4x-3&5x-7&4x-3\end{vmatrix}$，则方程 $f(x)=0$ 的根的个数为（　　）.

(A)1　　(B)2　　(C)3　　(D)4

分析 此题实质上是计算行列式，观察计算出的关于 x 的多项式的次数. 在计算过程中要充分运用行列式的性质.

解 应选(B).

$$f(x)\xlongequal{\substack{[2]-[1]\\ {[3]-[1]}\\ {[4]-[1]}}}\begin{vmatrix}x-2&1&0&-1\\2x-2&1&0&-1\\3x-3&1&x-2&-2\\4x&-3&x-7&-3\end{vmatrix}\xlongequal{[4]+[2]}\begin{vmatrix}x-2&1&0&0\\2x-2&1&0&0\\3x-3&1&x-2&-1\\4x&-3&x-7&-6\end{vmatrix}$$

$$=\begin{vmatrix}x-2&1\\2x-2&1\end{vmatrix}\begin{vmatrix}x-2&-1\\x-7&-6\end{vmatrix}=5x(x-1),$$

由此可知 $f(x)$ 是二次多项式，故应选(B).

【注】 不要错误地认为 $f(x)$ 一定是四次多项式.

例 2.1.10 设关于 λ 的方程

$$\begin{vmatrix} \lambda-1 & -2 & 3 \\ 1 & \lambda-4 & 3 \\ -1 & a & \lambda-5 \end{vmatrix}=0$$

有二重根，求参数 a 的值.

解

$$\begin{vmatrix} \lambda-1 & -2 & 3 \\ 1 & \lambda-4 & 3 \\ -1 & a & \lambda-5 \end{vmatrix} \xlongequal{\text{①}-\text{②}} \begin{vmatrix} \lambda-2 & 2-\lambda & 0 \\ 1 & \lambda-4 & 3 \\ -1 & a & \lambda-5 \end{vmatrix}$$

$$\xlongequal{[2]+[1]} \begin{vmatrix} \lambda-2 & 0 & 0 \\ 1 & \lambda-3 & 3 \\ -1 & a-1 & \lambda-5 \end{vmatrix}$$

$$=(\lambda-2)[(\lambda-3)(\lambda-5)-3(a-1)]$$

$$=(\lambda-2)(\lambda^2-8\lambda+18-3a)=0.$$

若 $\lambda=2$ 是二重根，则

$$(\lambda^2-8\lambda+18-3a)\Big|_{\lambda=2}=4-16+18-3a=0,$$

得 $a=2$.

若 $\lambda=2$ 不是二重根，则 $\lambda^2-8\lambda+18-3a=0$ 有两个相等的根，故 $\Delta=(-8)^2-4(18-3a)=0$，得 $a=\dfrac{2}{3}$. 此时，$\lambda=4$ 是二重根.

综上所述，$a=2$ 或 $a=\dfrac{2}{3}$.

二、抽象型行列式的计算

例 2.1.11 已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta},\boldsymbol{\gamma}$ 均为 4 维列向量，且

$$|\boldsymbol{\gamma},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=n,\quad |\boldsymbol{\alpha}_1,\boldsymbol{\beta}+\boldsymbol{\gamma},\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=m,$$

则 $|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,3\boldsymbol{\beta}|=$__________.

解 应填 $3(m+n)$.

$$\begin{aligned} m&=|\boldsymbol{\alpha}_1,\boldsymbol{\beta}+\boldsymbol{\gamma},\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=|\boldsymbol{\alpha}_1,\boldsymbol{\beta},\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|+|\boldsymbol{\alpha}_1,\boldsymbol{\gamma},\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3| \\ &=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}|-|\boldsymbol{\gamma},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3| \\ &=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}|-n, \end{aligned}$$

则

$$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}|=m+n,$$

故

$$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,3\boldsymbol{\beta}|=3|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}|=3(m+n).$$

【注】 利用行列式的性质将未知行列式化成已知行列式是解决此类问题的关键.

例 2.1.12 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 均为 3 维列向量，已知

$$|\boldsymbol{A}|=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=2,\quad |\boldsymbol{B}|=|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2-5\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3|,$$

则 $|\boldsymbol{B}-\boldsymbol{A}|=$__________.

解 应填 10.

$$|\boldsymbol{B}-\boldsymbol{A}|=|-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-5\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_3|\overset{(*)}{=\!=\!=}\left|[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}0&2&1\\-1&2&2\\2&-5&-2\end{bmatrix}\right|$$

$$=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\begin{vmatrix}0&2&1\\-1&2&2\\2&-5&-2\end{vmatrix},$$

其中 $|\boldsymbol{A}|=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|=2$. 又

$$\begin{vmatrix}0&2&1\\-1&2&2\\2&-5&-2\end{vmatrix}=\begin{vmatrix}0&2&1\\-1&2&2\\0&-1&2\end{vmatrix}=\begin{vmatrix}2&1\\-1&2\end{vmatrix}=5,$$

故 $|\boldsymbol{B}-\boldsymbol{A}|=10$.

【注】 能够熟练地将线性组合表示成矩阵乘积的形式，

$$0\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}0\\-1\\2\end{bmatrix},\quad 2\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-5\boldsymbol{\alpha}_3=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}2\\2\\-5\end{bmatrix},$$

$$\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_3=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}1\\2\\-2\end{bmatrix},$$

再合并成矩阵乘积的行列式（*）式是解决此类问题的关键.

三、余子式和代数余子式的线性组合的计算

由
$$a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}=\begin{vmatrix}&*&&\\a_{i1}&a_{i2}&\cdots&a_{in}\\&*&&\end{vmatrix},\qquad ①$$

则
$$k_1A_{i1}+k_2A_{i2}+\cdots+k_nA_{in}=\begin{vmatrix}&*&&\\k_1&k_2&\cdots&k_n\\&*&&\end{vmatrix},\qquad ②$$

其中 * 处表示元素不变，①，②的区别仅仅在于第 i 行的元素 $a_{i1},a_{i2},\cdots,a_{in}$ 换成了 $k_1,k_2,\cdots,k_n$，这样，给出不同的系数 $k_1,k_2,\cdots,k_n$，即得到了不同的行列式. 而若要求 $k_1M_{i1}+k_2M_{i2}+\cdots+k_nM_{in}$，只需用 $M_{ij}=(-1)^{i+j}A_{ij}$ 化为关于 A_{ij} 的线性组合即可.

例 2.1.13 设

$$|\boldsymbol{A}|=\begin{vmatrix}2&1&-3&5\\1&1&1&2\\4&2&3&1\\2&5&1&3\end{vmatrix},$$

计算 $A_{41}+A_{42}+A_{43}+A_{44}$，其中 A_{4j} 是元素 $a_{4j}(j=1,2,3,4)$ 的代数余子式.

分析 A_{4j} 是第 4 行元素 $a_{4j}(j=1,2,3,4)$ 的代数余子式，其值只与第 1,2,3 行元素有关，与第 4 行元素无关.

解 将 $|\boldsymbol{A}|$ 中第 4 行元素依次换为 1,1,1,1，并不改变 $A_{4j}(j=1,2,3,4)$ 的大小，有

$$A_{41}+A_{42}+A_{43}+A_{44}=1\cdot A_{41}+1\cdot A_{42}+1\cdot A_{43}+1\cdot A_{44}$$

$$=\begin{vmatrix}2&1&-3&5\\1&1&1&2\\4&2&3&1\\1&1&1&1\end{vmatrix}\xlongequal{②-④}\begin{vmatrix}2&1&-3&5\\0&0&0&1\\4&2&3&1\\1&1&1&1\end{vmatrix}$$

$$=(-1)^{2+4}\begin{vmatrix}2&1&-3\\4&2&3\\1&1&1\end{vmatrix}\xlongequal[{[3]-[1]}]{[2]-[1]}\begin{vmatrix}2&-1&-5\\4&-2&-1\\1&0&0\end{vmatrix}=1-10=-9.$$

基础习题精练

习 题

2.1.1 行列式

$$D_4=\begin{vmatrix}4&1&2&3\\3&4&1&2\\2&3&4&1\\1&2&3&4\end{vmatrix}=\underline{\qquad\qquad}.$$

2.1.2 行列式

$$D_4=\begin{vmatrix}x&-1&0&0\\0&x&-1&0\\0&0&x&-1\\a_0&a_1&a_2&a_3+x\end{vmatrix}=\underline{\qquad\qquad}.$$

2.1.3 行列式

$$D_4=\begin{vmatrix}1&a&0&0\\-1&2-a&b&0\\0&-2&3-b&c\\0&0&-3&4-c\end{vmatrix}=\underline{\qquad\qquad}.$$

2.1.4 计算

$$D_{n+1}=\begin{vmatrix} a_1^n & a_1^{n-1}b_1 & a_1^{n-2}b_1^2 & \cdots & a_1b_1^{n-1} & b_1^n \\ a_2^n & a_2^{n-1}b_2 & a_2^{n-2}b_2^2 & \cdots & a_2b_2^{n-1} & b_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n+1}^n & a_{n+1}^{n-1}b_{n+1} & a_{n+1}^{n-2}b_{n+1}^2 & \cdots & a_{n+1}b_{n+1}^{n-1} & b_{n+1}^n \end{vmatrix}.$$

2.1.5 计算 n 阶行列式 $D_n=\begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 & 0 \\ a_2 & x & -1 & \cdots & 0 & 0 \\ a_3 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n-1} & 0 & 0 & \cdots & x & -1 \\ a_n & 0 & 0 & \cdots & 0 & x \end{vmatrix}$.

2.1.6 已知 $\begin{vmatrix} \lambda & 2 & 2 \\ -2 & \lambda-2 & 2 \\ 2 & 2 & \lambda-2 \end{vmatrix}=0$，则 $\lambda=$__________.

2.1.7 一元二次方程 $\begin{vmatrix} x & 1 & a \\ 1 & x-1 & 1 \\ 1 & 1 & 1 \end{vmatrix}=0$ 有二重根，求参数 a 的值及该二重根.

2.1.8 已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 3 维列向量，且 $|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\neq 0$，$\boldsymbol{A}$ 是 3 阶矩阵，满足

$$\boldsymbol{A\alpha}_1=\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_3,\quad \boldsymbol{A\alpha}_2=\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,\quad \boldsymbol{A\alpha}_3=2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,$$

则 $|\boldsymbol{A}|=$__________.

2.1.9 已知 n 阶行列式 $|\boldsymbol{A}|=a$，$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & b \end{vmatrix}=0$，其中 $\boldsymbol{\beta}$ 是 n 维列向量，则 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & c \end{vmatrix}=$__________.

2.1.10 设

$$|\boldsymbol{A}|=\begin{vmatrix} -2 & 2 & -2 & 2 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ 2 & -1 & 3 & 5 \end{vmatrix}.$$

计算：(1)$A_{41}+A_{42}+A_{43}+A_{44}$，其中 A_{4j} 是元素 $a_{4j}(j=1,2,3,4)$ 的代数余子式；

(2)$M_{31}+M_{32}+M_{33}+M_{34}$，其中 M_{3j} 是元素 $a_{3j}(j=1,2,3,4)$ 的余子式.

解 答

2.1.1 160 **解** $D_4=10\begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 4 & 1 & 2 \\ 1 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{vmatrix}=10\begin{vmatrix} 1 & 1 & 2 & 3 \\ 0 & 3 & -1 & -1 \\ 0 & 2 & 2 & -2 \\ 0 & 1 & 1 & 1 \end{vmatrix}=10\begin{vmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 0 & -4 \end{vmatrix}=160.$

2.1.2 $a_0+a_1x+a_2x^2+a_3x^3+x^4$ **解** 按第 4 行展开，得

$$D_4=a_0(-1)^{1+4}\begin{vmatrix}-1&0&0\\x&-1&0\\0&x&-1\end{vmatrix}+a_1(-1)^{2+4}\begin{vmatrix}x&0&0\\0&-1&0\\0&x&-1\end{vmatrix}+a_2(-1)^{3+4}\begin{vmatrix}x&-1&0\\0&x&0\\0&0&-1\end{vmatrix}+$$

$$(a_3+x)(-1)^{4+4}\begin{vmatrix}x&-1&0\\0&x&-1\\0&0&x\end{vmatrix}$$

$$=a_0+a_1x+a_2x^2+a_3x^3+x^4.$$

【注】 若此题中的 a_0,a_1,a_2,a_3 取特殊值，如 $\begin{vmatrix}\lambda&-1&0&0\\0&\lambda&-1&0\\0&0&\lambda&-1\\4&3&2&\lambda+1\end{vmatrix}$（此题为2016年考研实考题)，可直接代入.

2.1.3　24　**解**　从第1行开始，依次把每行加到下一行(逐行相加)，则

$$D_4=\begin{vmatrix}1&a&0&0\\0&2&b&0\\0&-2&3-b&c\\0&0&-3&4-c\end{vmatrix}=\begin{vmatrix}1&a&0&0\\0&2&b&0\\0&0&3&c\\0&0&-3&4-c\end{vmatrix}=\begin{vmatrix}1&a&0&0\\0&2&b&0\\0&0&3&c\\0&0&0&4\end{vmatrix}=24.$$

2.1.4　**解**　$D_{n+1}=(a_1a_2\cdots a_{n+1})^n\begin{vmatrix}1&\frac{b_1}{a_1}&\left(\frac{b_1}{a_1}\right)^2&\cdots&\left(\frac{b_1}{a_1}\right)^n\\1&\frac{b_2}{a_2}&\left(\frac{b_2}{a_2}\right)^2&\cdots&\left(\frac{b_2}{a_2}\right)^n\\\vdots&\vdots&\vdots&&\vdots\\1&\frac{b_{n+1}}{a_{n+1}}&\left(\frac{b_{n+1}}{a_{n+1}}\right)^2&\cdots&\left(\frac{b_{n+1}}{a_{n+1}}\right)^n\end{vmatrix}$

$$=(a_1a_2\cdots a_{n+1})^n\begin{vmatrix}1&1&\cdots&1\\\frac{b_1}{a_1}&\frac{b_2}{a_2}&\cdots&\frac{b_{n+1}}{a_{n+1}}\\\vdots&\vdots&&\vdots\\\left(\frac{b_1}{a_1}\right)^n&\left(\frac{b_2}{a_2}\right)^n&\cdots&\left(\frac{b_{n+1}}{a_{n+1}}\right)^n\end{vmatrix}$$

$$=(a_1a_2\cdots a_{n+1})^n\prod_{1\leqslant i<j\leqslant n+1}\left(\frac{b_j}{a_j}-\frac{b_i}{a_i}\right).$$

2.1.5　**解**　按第 n 行展开，得

$$D_n=a_n(-1)^{n+1}\begin{vmatrix}-1&0&0&\cdots&0\\x&-1&0&\cdots&0\\0&x&-1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&-1\end{vmatrix}+x(-1)^{n+n}D_{n-1}$$

$$=a_n(-1)^{n+1}(-1)^{n-1}+xD_{n-1}=a_n+xD_{n-1},$$

于是按递推关系 $D_n=a_n+xD_{n-1}$，有

$$D_{n-1}=a_{n-1}+xD_{n-2},\quad D_{n-2}=a_{n-2}+xD_{n-3},\quad \cdots,\quad D_2=a_2+a_1x,$$

分别用 $x,x^2,\cdots,x^{n-2}$ 乘上述各等式，然后代入 D_n 的表达式中，则

$$D_n=a_n+a_{n-1}x+a_{n-2}x^2+\cdots+a_1x^{n-1}.$$

> **【注】** 对元素有一定规律(如，某行或某列只有两个元素不是 0，而其余元素都是 0)的行列式可考虑利用展开公式建立递推关系来处理.

2.1.6 0 或 4 **解** $\begin{vmatrix}\lambda&2&2\\-2&\lambda-2&2\\2&2&\lambda-2\end{vmatrix}=\begin{vmatrix}\lambda&2&2\\0&\lambda&\lambda\\2&2&\lambda-2\end{vmatrix}=\lambda\begin{vmatrix}\lambda&0&0\\0&1&1\\2&2&\lambda-2\end{vmatrix}=\lambda^2(\lambda-4)=0,$

故 $\lambda=0$ 或 4.

2.1.7 **解** $\begin{vmatrix}x&1&a\\1&x-1&1\\1&1&1\end{vmatrix}=\begin{vmatrix}x&1-x&a-x\\1&x-2&0\\1&0&0\end{vmatrix}=-(x-2)(a-x)$，故当 $a=2$ 时，$x=2$ 是二重根.

2.1.8 -2 **解** $\boldsymbol{A}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=[\boldsymbol{A\alpha}_1,\boldsymbol{A\alpha}_2,\boldsymbol{A\alpha}_3]=[\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2]$，

等式两边取行列式，有

$$\begin{aligned}|\boldsymbol{A}||\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|&=|\boldsymbol{\alpha}_2-2\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2|\\&=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\begin{vmatrix}0&1&2\\1&-1&1\\-2&2&0\end{vmatrix}.\end{aligned}$$

因 $|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\neq0$，故

$$|\boldsymbol{A}|=\begin{vmatrix}0&1&2\\1&-1&1\\-2&2&0\end{vmatrix}=\begin{vmatrix}0&1&2\\1&-1&1\\0&0&2\end{vmatrix}=-2.$$

2.1.9 $a(c-b)$ **解** $\begin{vmatrix}\boldsymbol{A}&\boldsymbol{\beta}\\\boldsymbol{\beta}^{\mathrm{T}}&c\end{vmatrix}=\begin{vmatrix}\boldsymbol{A}&\boldsymbol{\beta}\\\boldsymbol{\beta}^{\mathrm{T}}&b+c-b\end{vmatrix}=\begin{vmatrix}\boldsymbol{A}&\boldsymbol{\beta}+\boldsymbol{0}\\\boldsymbol{\beta}^{\mathrm{T}}&b+(c-b)\end{vmatrix}=\begin{vmatrix}\boldsymbol{A}&\boldsymbol{\beta}\\\boldsymbol{\beta}^{\mathrm{T}}&b\end{vmatrix}+\begin{vmatrix}\boldsymbol{A}&\boldsymbol{0}\\\boldsymbol{\beta}^{\mathrm{T}}&c-b\end{vmatrix}$

$$=0+|\boldsymbol{A}|(c-b)=a(c-b).$$

2.1.10 **解** (1)**方法一** 将 $|\boldsymbol{A}|$ 中第 4 行元素依次改为 1，1，1，1，得

$$A_{41}+A_{42}+A_{43}+A_{44}=\begin{vmatrix}-2&2&-2&2\\1&2&3&4\\1&1&1&1\\1&1&1&1\end{vmatrix}=0.$$

方法二 行列式 $|\boldsymbol{A}|$ 中第 3 行元素全为 1，故

$$A_{41}+A_{42}+A_{43}+A_{44}=a_{31}A_{41}+a_{32}A_{42}+a_{33}A_{43}+a_{34}A_{44},$$

由行列式展开公式知，第 3 行元素与第 4 行元素对应的代数余子式乘积之和为零，故有

$$A_{41}+A_{42}+A_{43}+A_{44}=0.$$

(2)余子式和代数余子式有如下关系：$A_{ij}=(-1)^{i+j}M_{ij}$．两边同乘 $(-1)^{i+j}$，因 $(-1)^{i+j}\cdot(-1)^{i+j}=$

$(-1)^{2i+2j}=1$，故 $M_{ij}=(-1)^{i+j}A_{ij}\,(i,j=1,2,\cdots,n)$，从而有

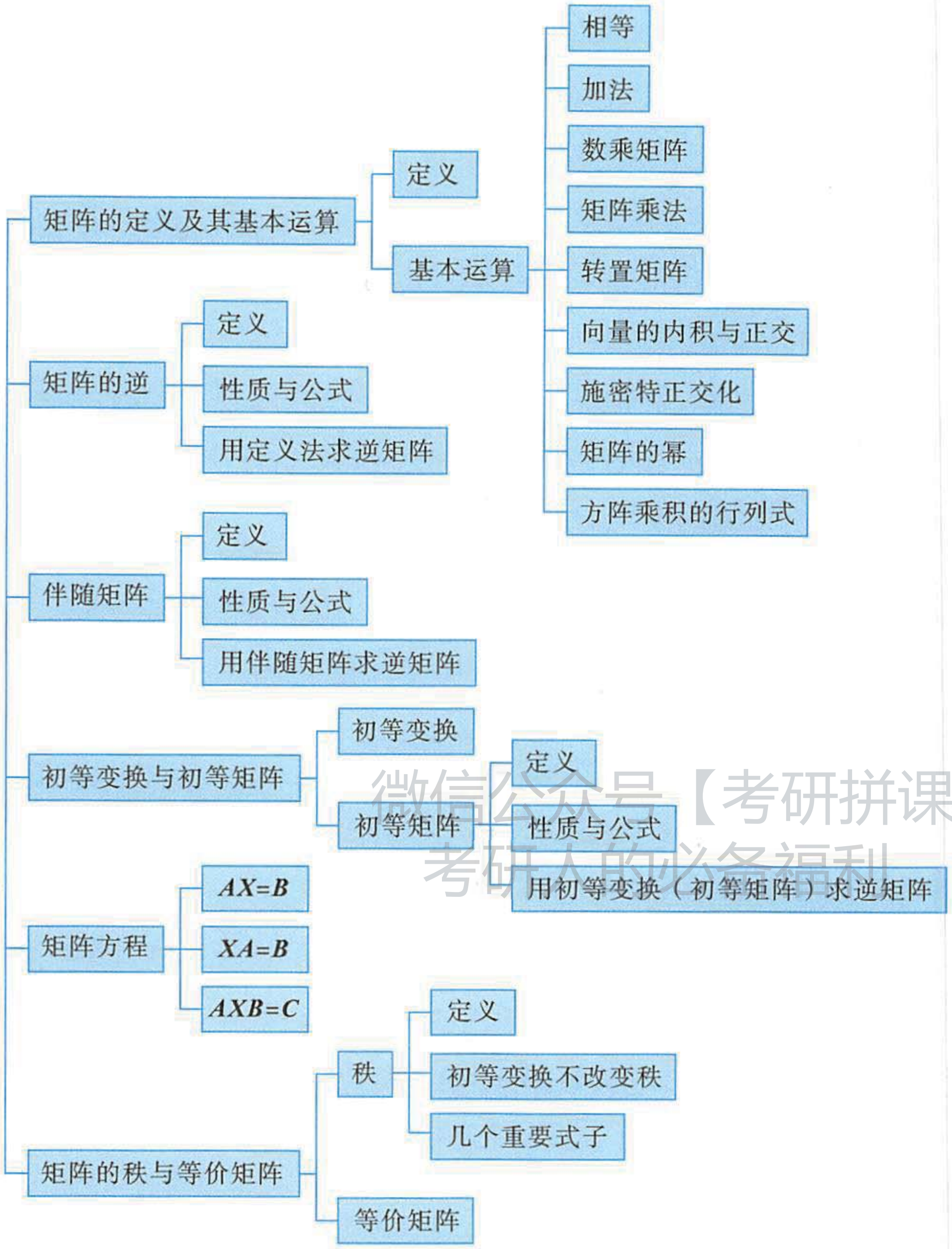
$$M_{31}+M_{32}+M_{33}+M_{34}=A_{31}-A_{32}+A_{33}-A_{34}.$$

将$|\boldsymbol{A}|$中第3行元素依次换成1，-1，1，-1，得

$$M_{31}+M_{32}+M_{33}+M_{34}=A_{31}-A_{32}+A_{33}-A_{34}=\begin{vmatrix}-2&2&-2&2\\1&2&3&4\\1&-1&1&-1\\2&-1&3&5\end{vmatrix}=0.$$

第2讲　矩　阵

基础知识结构



基础内容精讲

一、矩阵的本质

假如英语系有 98 个女生，2 个男生；机械系有 95 个男生，5 个女生. 你能把这个系统信息表达出来吗？这个问题的回答是显然而且简单的：

$$\begin{array}{c|cc|} & \text{英语系} & \text{机械系} \\ \text{男生人数} & 2 & 95 \\ \text{女生人数} & 98 & 5 \end{array}.$$

在这个数表中，我们可以看到：第一列表达了英语系的男、女生人数，第二列表达了机械系的男、女生人数，而第一行表达了不同系的男生人数，第二行表达了不同系的女生人数.

这就是读者对**矩阵的初步认识——表达系统信息**.

但读者不可对矩阵的概念就此罢休. 再看一个矩阵：

$$\begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 9 \\ 2 & 4 & 6 \end{bmatrix}.$$

事实上，我们不必再给这个矩阵赋予具体背景了，总之，可以抽象认为这是一个系统信息的表达即可.

我给读者提供两个重要观点，供你在学习相关知识后参考.

重要观点 1　矩阵也是由若干行(列)向量拼成的——上面那个矩阵可以看作由三个行向量[1,2,3]，[6,7,9]与[2,4,6]组成，也可以看作由三个列向量$[1,6,2]^{\mathrm{T}}$，$[2,7,4]^{\mathrm{T}}$与$[3,9,6]^{\mathrm{T}}$组成.

重要观点 2　矩阵不能运算，但是其若干行(列)向量之间存在着某种联系——你是否看到——[1,2,3]与[2,4,6]这两个向量是平行的(存在线性关系)，而[1,2,3]与[6,7,9]，[2,4,6]与[6,7,9]却不存在这种线性关系. **这种关系反映了矩阵的本质——矩阵的秩**——这个秩是 2.

下面给出矩阵的秩的定义.

设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{A}$ 中最高阶非零子式的阶数称为矩阵 $\boldsymbol{A}$ 的秩，记为 $r(\boldsymbol{A})$.

也可以这样定义：若存在 k 阶子式不为零，而任意 $k+1$ 阶子式(如果有的话)全为零，则 $r(\boldsymbol{A})=k$，且

$$r(\boldsymbol{A}_{n\times n})=n\Leftrightarrow|\boldsymbol{A}|\neq0\Leftrightarrow\boldsymbol{A}\text{ 可逆}.$$

从此定义可以看出，矩阵秩的本质就是组成该矩阵的线性无关的向量的个数.

二、矩阵的定义及其基本运算

1. 矩阵的定义

由 $m\times n$ 个数 $a_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 排成的 m 行 n 列的矩形表格

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为一个 $m\times n$ 矩阵，简记为 $\boldsymbol{A}$ 或 $(a_{ij})_{m\times n}(i=1,2,\cdots,m;j=1,2,\cdots,n)$. 当 $m=n$ 时，称 $\boldsymbol{A}$ 为 n 阶方阵.

两个矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{B}=(b_{ij})_{s\times k}$. 若 $m=s,n=k$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为同型矩阵.

2. 矩阵的基本运算

(1)**相等**　$\boldsymbol{A}=(a_{ij})_{m\times n}=\boldsymbol{B}=(b_{ij})_{s\times k}\Leftrightarrow m=s,n=k$，且 $a_{ij}=b_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，即 $\boldsymbol{A},\boldsymbol{B}$ 是同型矩阵，且对应元素相等.

(2)**加法**　两个矩阵是同型矩阵时，可以相加，即

$$\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}=(a_{ij})_{m\times n}+(b_{ij})_{m\times n}=(c_{ij})_{m\times n},$$

其中，$c_{ij}=a_{ij}+b_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$，即对应元素相加.

(3)**数乘矩阵**　设 k 是一个数，$\boldsymbol{A}$ 是一个 $m\times n$ 矩阵. 数 k 和 $\boldsymbol{A}$ 的乘积称为数乘矩阵，即

$$k\boldsymbol{A}=\boldsymbol{A}k=k\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}=\begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix}$$

$$=(ka_{ij})_{m\times n},$$

即 $\boldsymbol{A}$ 的每个元素都乘以 k.

加法运算和数乘运算统称为矩阵的线性运算，满足下列运算规律：

①**交换律**　$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$；

②**结合律**　$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$；

③**分配律**　$k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B},(k+l)\boldsymbol{A}=k\boldsymbol{A}+l\boldsymbol{A}$；

④**数和矩阵相乘的结合律**　$k(l\boldsymbol{A})=(kl)\boldsymbol{A}=l(k\boldsymbol{A})$.

其中，$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 是同型矩阵，k,l 是任意常数.

当 n 阶方阵 $\boldsymbol{A}$ 计算行列式时，记成 $|\boldsymbol{A}|$.

【注】 (1) $|k\boldsymbol{A}|=k^n|\boldsymbol{A}|\neq k|\boldsymbol{A}|\ (n\geqslant 2,k\neq 0,1)$；

(2)一般，$|\boldsymbol{A}+\boldsymbol{B}|\neq|\boldsymbol{A}|+|\boldsymbol{B}|$；

(3) $\boldsymbol{A}\neq\boldsymbol{O}\nRightarrow|\boldsymbol{A}|\neq 0$；

(4) $\boldsymbol{A}\neq\boldsymbol{B}\nRightarrow|\boldsymbol{A}|\neq|\boldsymbol{B}|$.

(4)**矩阵的乘法**　设 $\boldsymbol{A}$ 是 $m\times s$ 矩阵，$\boldsymbol{B}$ 是 $s\times n$ 矩阵(矩阵 $\boldsymbol{A}$ 的列数必须与矩阵 $\boldsymbol{B}$ 的行数相等)，则 $\boldsymbol{A},\boldsymbol{B}$ 可乘，乘积 $\boldsymbol{AB}$ 是 $m\times n$ 矩阵，记 $\boldsymbol{C}=\boldsymbol{AB}=(c_{ij})_{m\times n}$. $\boldsymbol{C}$ 的第 i 行第 j 列元素 c_{ij} 是 $\boldsymbol{A}$ 的第 i 行的 s 个元素与 $\boldsymbol{B}$ 的第 j 列的 s 个对应元素两两乘积之和，即

$$c_{ij}=\sum_{k=1}^{s}a_{ik}b_{kj}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}\,(i=1,2,\cdots,m;j=1,2,\cdots,n).$$

矩阵乘法满足下列运算规律：

①**结合律**　$(\boldsymbol{A}_{m\times s}\boldsymbol{B}_{s\times r})\boldsymbol{C}_{r\times n}=\boldsymbol{A}_{m\times s}(\boldsymbol{B}_{s\times r}\boldsymbol{C}_{r\times n})$；

②**分配律** $A_{m\times s}(B_{s\times n}+C_{s\times n})=A_{m\times s}B_{s\times n}+A_{m\times s}C_{s\times n}$,

$(A_{m\times s}+B_{m\times s})C_{s\times n}=A_{m\times s}C_{s\times n}+B_{m\times s}C_{s\times n}$;

③**数乘与矩阵乘积的结合律**

$$(kA_{m\times s})B_{s\times n}=A_{m\times s}(kB_{s\times n})=k(A_{m\times s}B_{s\times n}).$$

【注】 (1)矩阵的乘法一般情况下不满足交换律，即 $AB\neq BA$.

例如
$$A=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix},\quad B=\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix},$$

则
$$AB=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix},$$

$$BA=\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}=\begin{bmatrix}2 & 2\\ -2 & -2\end{bmatrix},\quad AB\neq BA.$$

(2)由上面的例子知，存在 $A\neq O,B\neq O$，而 $AB=O$，故 $AB=O\nRightarrow A=O$ 或 $B=O$.

(3)$AB=AC,A\neq O\Rightarrow A(B-C)=O$ 及 $A\neq O\nRightarrow B=C$.

(5)**转置矩阵** 将 $m\times n$ 矩阵 $A=(a_{ij})_{m\times n}$的行与列互换得到的 $n\times m$ 矩阵，称为矩阵 A 的转置矩阵，记为 A^{T}，即

$$A^{\mathrm{T}}=\begin{bmatrix}a_{11} & a_{21} & \cdots & a_{m1}\\ a_{12} & a_{22} & \cdots & a_{m2}\\ \vdots & \vdots & & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{mn}\end{bmatrix}.$$

转置矩阵满足下列运算规律：

①$(A^{\mathrm{T}})^{\mathrm{T}}=A$；②$(kA)^{\mathrm{T}}=kA^{\mathrm{T}}$；③$(A+B)^{\mathrm{T}}=A^{\mathrm{T}}+B^{\mathrm{T}}$；④$(AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}$；⑤当 $m=n$ 时，$|A^{\mathrm{T}}|=|A|$.

(6)**向量的内积与正交**

内积 设 $\alpha=[a_1,a_2,\cdots,a_n]^{\mathrm{T}},\beta=[b_1,b_2,\cdots,b_n]^{\mathrm{T}}$，则称

$$\alpha^{\mathrm{T}}\beta=\sum_{i=1}^{n}a_ib_i=a_1b_1+a_2b_2+\cdots+a_nb_n$$

为向量 α,β 的内积，记作(α,β)，即$(\alpha,\beta)=\alpha^{\mathrm{T}}\beta$.

正交 当 $\alpha^{\mathrm{T}}\beta=0$ 时，称向量 α,β 是正交向量.

模 $\|\alpha\|=\sqrt{\sum_{i=1}^{n}a_i^2}$ 称为向量 α 的模(长度).

$\|\alpha\|=1$ 时，称 α 为单位向量.

标准正交向量组 若列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 满足

$$\alpha_i^{\mathrm{T}}\alpha_j=\begin{cases}0, & i\neq j,\\ 1, & i=j,\end{cases}$$

则称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为标准或单位正交向量组.

(7)**施密特标准正交化(又称正交规范化)过程**

线性无关向量组 α_1,α_2 的标准正交化(又称正交规范化)公式为

$$\beta_1=\alpha_1,$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1,$$

得到的 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是正交向量组.

将 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 单位化，得

$$\boldsymbol{\eta}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|},\quad \boldsymbol{\eta}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|},$$

则 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2$ 是标准正交向量组.

(8)**矩阵的幂** $\boldsymbol{A}$ 是一个 n 阶方阵，$\boldsymbol{A}^m=\overbrace{\boldsymbol{AA}\cdots\boldsymbol{A}}^{m个}$ 称为 $\boldsymbol{A}$ 的 m 次幂.

【注】 (1)因矩阵乘法不满足交换律，故一般，

$$(\boldsymbol{A}+\boldsymbol{B})^2=(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}+\boldsymbol{B})=\boldsymbol{A}^2+\boldsymbol{AB}+\boldsymbol{BA}+\boldsymbol{B}^2\neq\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2,$$

$$(\boldsymbol{A}-\boldsymbol{B})^2=\boldsymbol{A}^2-\boldsymbol{AB}-\boldsymbol{BA}+\boldsymbol{B}^2\neq\boldsymbol{A}^2-2\boldsymbol{AB}+\boldsymbol{B}^2,$$

$$(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2+\boldsymbol{BA}-\boldsymbol{AB}-\boldsymbol{B}^2\neq\boldsymbol{A}^2-\boldsymbol{B}^2,$$

$$(\boldsymbol{AB})^m=\overbrace{(\boldsymbol{AB})(\boldsymbol{AB})\cdots(\boldsymbol{AB})}^{m个}\neq\boldsymbol{A}^m\boldsymbol{B}^m.$$

(2)若 $f(x)=a_0+a_1x+\cdots+a_mx^m$，则

$$f(\boldsymbol{A})=a_0\boldsymbol{E}+a_1\boldsymbol{A}+\cdots+a_m\boldsymbol{A}^m.$$

(9)**方阵乘积的行列式** 设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶方阵，则 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$.

【注】 几种重要矩阵.

(1)**零矩阵** 每个元素均为零的矩阵，记为 $\boldsymbol{O}$.

(2)**单位矩阵** 主对角元素均为 1，其余元素全为零的 n 阶方阵，称为 n 阶单位矩阵，记成 $\boldsymbol{E}$(或 $\boldsymbol{I}$).

(3)**数量矩阵** 数 k 和单位矩阵的乘积称为数量矩阵.

(4)**对角矩阵** 非主对角元素均为零的矩阵称为对角矩阵.

(5)**上(下)三角矩阵** 当 $i>(<)j$ 时，$a_{ij}=0$ 的矩阵称为上(下)三角矩阵.

(6)**对称矩阵** 满足条件 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$ 的矩阵 $\boldsymbol{A}$ 称为对称矩阵，$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\Leftrightarrow a_{ij}=a_{ji}$.

(7)**反对称矩阵** 满足条件 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$ 的矩阵 $\boldsymbol{A}$ 称为反对称矩阵，

$$\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}\Leftrightarrow\begin{cases}a_{ij}=-a_{ji},i\neq j,\\ a_{ii}=0.\end{cases}$$

(8)**正交矩阵** 设 $\boldsymbol{A}$ 是 n 阶方阵，满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$，则称 $\boldsymbol{A}$ 是正交矩阵.

$\boldsymbol{A}$ 是正交矩阵 $\Leftrightarrow\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}\Leftrightarrow\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}\Leftrightarrow\boldsymbol{A}$ 的行(列)向量组是标准正交向量组.

分析 设 $\boldsymbol{A}=\begin{bmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{bmatrix}$，且记

$$\boldsymbol{\alpha}=[a_1,a_2,a_3]^{\mathrm{T}},\quad\boldsymbol{\beta}=[b_1,b_2,b_3]^{\mathrm{T}},\quad\boldsymbol{\gamma}=[c_1,c_2,c_3]^{\mathrm{T}}.$$

则

$$\boldsymbol{AA}^{\mathrm{T}}=\begin{bmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{bmatrix}\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{bmatrix}=\boldsymbol{E}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},$$

则有

$$\begin{cases} a_1^2+a_2^2+a_3^2=1 \Rightarrow \|\boldsymbol{\alpha}\|=1, \\ b_1^2+b_2^2+b_3^2=1 \Rightarrow \|\boldsymbol{\beta}\|=1, \\ c_1^2+c_2^2+c_3^2=1 \Rightarrow \|\boldsymbol{\gamma}\|=1, \\ a_1b_1+a_2b_2+a_3b_3=0 \Rightarrow (\boldsymbol{\alpha},\boldsymbol{\beta})=0,\text{即 }\boldsymbol{\alpha}\text{ 与 }\boldsymbol{\beta}\text{ 正交}, \\ a_1c_1+a_2c_2+a_3c_3=0 \Rightarrow (\boldsymbol{\alpha},\boldsymbol{\gamma})=0,\text{即 }\boldsymbol{\alpha}\text{ 与 }\boldsymbol{\gamma}\text{ 正交}, \\ b_1c_1+b_2c_2+b_3c_3=0 \Rightarrow (\boldsymbol{\beta},\boldsymbol{\gamma})=0,\text{即 }\boldsymbol{\beta}\text{ 与 }\boldsymbol{\gamma}\text{ 正交}, \end{cases}$$

即 $\boldsymbol{A}$ 是由两两正交的单位向量组(称为规范正交基)组成.

(9)分块矩阵

①矩阵的分块.

用几条纵线和横线把一个矩阵分成若干小块，每一小块称为原矩阵的子块. 把子块看作原矩阵的一个元素，就得到了分块矩阵.

如 $\boldsymbol{A}$ 按行分块：

$$\boldsymbol{A}=\left[\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hdashline a_{21} & a_{22} & \cdots & a_{2n} \\ \hdashline \vdots & \vdots & & \vdots \\ \hdashline a_{m1} & a_{m2} & \cdots & a_{mn} \end{array}\right]=\begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \vdots \\ \boldsymbol{A}_m \end{bmatrix},$$

其中，$\boldsymbol{A}_i=[a_{i1},a_{i2},\cdots,a_{in}](i=1,2,\cdots,m)$是 $\boldsymbol{A}$ 的一个子块.

$\boldsymbol{B}$ 按列分块：

$$\boldsymbol{B}=\left[\begin{array}{c:c:c:c} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{array}\right]=[\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_n],$$

其中，$\boldsymbol{B}_j=[b_{1j},b_{2j},\cdots,b_{mj}]^{\mathrm{T}}(j=1,2,\cdots,n)$是 $\boldsymbol{B}$ 的一个子块.

②分块矩阵的基本运算(以 2×2 型分块矩阵为例).

加法：同型，且分法一致，则 $\begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{bmatrix}+\begin{bmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 \\ \boldsymbol{B}_3 & \boldsymbol{B}_4 \end{bmatrix}=\begin{bmatrix} \boldsymbol{A}_1+\boldsymbol{B}_1 & \boldsymbol{A}_2+\boldsymbol{B}_2 \\ \boldsymbol{A}_3+\boldsymbol{B}_3 & \boldsymbol{A}_4+\boldsymbol{B}_4 \end{bmatrix}$.

数乘：$k\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}=\begin{bmatrix} k\boldsymbol{A} & k\boldsymbol{B} \\ k\boldsymbol{C} & k\boldsymbol{D} \end{bmatrix}$.

乘法：$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}\begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \\ \boldsymbol{Z} & \boldsymbol{W} \end{bmatrix}=\begin{bmatrix} \boldsymbol{AX}+\boldsymbol{BZ} & \boldsymbol{AY}+\boldsymbol{BW} \\ \boldsymbol{CX}+\boldsymbol{DZ} & \boldsymbol{CY}+\boldsymbol{DW} \end{bmatrix}$，要可乘、可加.

注意 对于乘法的运算要注意，分块相乘后，左边的仍在左边，右边的仍在右边.

若 $\boldsymbol{A},\boldsymbol{B}$ 分别为 m,n 阶方阵，则分块对角矩阵的幂为

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{bmatrix}^n=\begin{bmatrix} \boldsymbol{A}^n & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}^n \end{bmatrix}.$$

三、矩阵的逆

1. 逆矩阵的定义

(1)定义　A,B 是 n 阶方阵，E 是 n 阶单位矩阵，若 $AB=BA=E$，则称 A 是可逆矩阵，并称 B 是 A 的逆矩阵，且逆矩阵是唯一的，记作 A^{-1}.

(2)A 可逆的充分必要条件是 $|A|\neq 0$. 当 $|A|\neq 0$ 时，A 可逆，且

$$A^{-1}=\frac{1}{|A|}A^{*}.$$

2. 逆矩阵的性质与重要公式

设 A,B 是同阶可逆矩阵，则

(1)$(A^{-1})^{-1}=A$；

(2)若 $k\neq 0$，则 $(kA)^{-1}=\frac{1}{k}A^{-1}$；

(3)AB 也可逆，且 $(AB)^{-1}=B^{-1}A^{-1}$；

(4)A^{T} 也可逆，且 $(A^{\mathrm{T}})^{-1}=(A^{-1})^{\mathrm{T}}$；

【注】　此处可称为“穿脱”原则，即穿衣时先内后外，脱衣时先外后内.

(5)$|A^{-1}|=|A|^{-1}$.

【注】　$A+B$ 不一定可逆，且 $(A+B)^{-1}\neq A^{-1}+B^{-1}$.

3. 用定义求逆矩阵的方法

方法一　依定义，即求一个矩阵 B，使 $AB=E$，则 A 可逆，且 $A^{-1}=B$.

方法二　将 A 分解成若干个可逆矩阵乘积. 因两个可逆矩阵的积仍是可逆矩阵，即若 $A=BC$，其中，B,C 均可逆，则 A 可逆，且

$$A^{-1}=(BC)^{-1}=C^{-1}B^{-1}.$$

方法三　一些简单分块矩阵的逆. 若 A,B 均是可逆方阵，则

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix}^{-1}=\begin{bmatrix} A^{-1} & O \\ O & B^{-1} \end{bmatrix},\quad \begin{bmatrix} O & A \\ B & O \end{bmatrix}^{-1}=\begin{bmatrix} O & B^{-1} \\ A^{-1} & O \end{bmatrix}.$$

四、伴随矩阵

1. 伴随矩阵的定义

伴随矩阵　将行列式 $|A|$ 的 n^2 个元素的代数余子式按如下形式排成的矩阵称为 A 的伴随矩阵，记作 A^{*}，即

$$A^{*}=\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix},$$

且有

$$AA^*=A^*A=|A|E.$$

2. 伴随矩阵的性质与重要公式

(1)对任意 n 阶方阵 A，都有伴随矩阵 A^*，且有公式

$$AA^*=A^*A=|A|E,\quad |A^*|=|A|^{n-1}.$$

当 $|A|\neq 0$ 时，有

$$A^*=|A|A^{-1},\quad A^{-1}=\frac{1}{|A|}A^*,\quad A=|A|(A^*)^{-1};$$

$$(kA)(kA)^*=|kA|E;$$

$$A^{\mathrm{T}}(A^{\mathrm{T}})^*=|A^{\mathrm{T}}|E;$$

$$A^{-1}(A^{-1})^*=|A^{-1}|E;$$

$$A^*(A^*)^*=|A^*|E.$$

(2)$(A^{\mathrm{T}})^*=(A^*)^{\mathrm{T}}$，$(A^{-1})^*=(A^*)^{-1}$，$(AB)^*=B^*A^*$，$(A^*)^*=|A|^{n-2}A$.

【注】 $(A+B)^*\neq A^*+B^*$.

3. 用伴随矩阵求逆矩阵的方法

若 $|A|\neq 0$，则 A 可逆，且 $A^{-1}=\frac{1}{|A|}A^*$.

五、初等变换与初等矩阵

1. 初等变换

(1)一个非零常数乘矩阵的某一行(列)；

(2)互换矩阵中某两行(列)的位置；

(3)将矩阵的某一行(列)的 k 倍加到另一行(列).

以上三种变换称为**矩阵的初等行(列)变换**，且分别称为互换、倍乘、倍加初等行(列)变换.

2. 初等矩阵的定义

由单位矩阵经过一次初等变换得到的矩阵称为**初等矩阵**. 以 3 阶矩阵为例.

(1)$E_2(k)=\begin{bmatrix}1&0&0\\0&k&0\\0&0&1\end{bmatrix}$，$E$ 的第 2 行(或第 2 列)乘 k 倍，称为倍乘初等矩阵.

定义：$E_i(k)(k\neq 0)$表示单位矩阵 E 的第 i 行(或第 i 列)乘以非零常数 k 所得的初等矩阵.

(2)$E_{12}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$，$E$ 的第 1，2 行(或第 1，2 列)互换，称为互换初等矩阵.

定义：E_{ij} 表示单位矩阵 E 交换第 i 行与第 j 行(或交换第 i 列与第 j 列)所得的初等矩阵.

(3)$E_{31}(k)=\begin{bmatrix}1&0&0\\0&1&0\\k&0&1\end{bmatrix}$，$E$ 的第 1 行的 k 倍加到第 3 行(或第 3 列的 k 倍加到第 1 列)，称为倍加初等矩阵.

定义：$\boldsymbol{E}_{ij}(k)$表示单位矩阵$\boldsymbol{E}$的第j行的k倍加到第i行（或第i列的k倍加到第j列）所得的初等矩阵.

3. 初等矩阵的性质与重要公式

（1）初等矩阵的转置仍是初等矩阵.

（2）因

$$|\boldsymbol{E}_i(k)|=k\neq0,\quad |\boldsymbol{E}_{ij}|=-1\neq0,\quad |\boldsymbol{E}_{ij}(k)|=1\neq0,$$

故初等矩阵都是可逆矩阵，且

$$[\boldsymbol{E}_i(k)]^{-1}=\boldsymbol{E}_i\left(\frac{1}{k}\right),\quad \boldsymbol{E}_{ij}^{-1}=\boldsymbol{E}_{ij},\quad [\boldsymbol{E}_{ij}(k)]^{-1}=\boldsymbol{E}_{ij}(-k),$$

其逆矩阵仍是同一类型的初等矩阵.

（3）若$\boldsymbol{A}$是可逆矩阵，则$\boldsymbol{A}$可以表示成有限个初等矩阵的乘积，即$\boldsymbol{A}=\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_s$，其中$\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_s$是初等矩阵.

（4）对n阶矩阵$\boldsymbol{A}$进行初等行变换，相当于矩阵$\boldsymbol{A}$左乘相应的初等矩阵. 同样，对$\boldsymbol{A}$进行初等列变换，相当于矩阵$\boldsymbol{A}$右乘相应的初等矩阵.

4. 用初等变换求逆矩阵的方法

$$[\boldsymbol{A}\,\vdots\,\boldsymbol{E}]\xrightarrow{\text{初等行变换}}[\boldsymbol{E}\,\vdots\,\boldsymbol{A}^{-1}],$$

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{E}\end{bmatrix}\xrightarrow{\text{初等列变换}}\begin{bmatrix}\boldsymbol{E}\\ \boldsymbol{A}^{-1}\end{bmatrix}.$$

六、等价矩阵和矩阵的等价标准形

设$\boldsymbol{A},\boldsymbol{B}$均是$m\times n$矩阵，若存在可逆矩阵$\boldsymbol{P}_{m\times m},\boldsymbol{Q}_{n\times n}$，使得$\boldsymbol{PAQ}=\boldsymbol{B}$，则称$\boldsymbol{A},\boldsymbol{B}$是等价矩阵，记作$\boldsymbol{A}\cong\boldsymbol{B}$.

$\boldsymbol{A}$是一个$m\times n$矩阵，则$\boldsymbol{A}$等价于形如$\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}$的矩阵（$\boldsymbol{E}_r$中的$r$恰是$r(\boldsymbol{A})$），后者称为$\boldsymbol{A}$的等价标准形. 等价标准形是唯一的，即若$r(\boldsymbol{A})=r$，则存在可逆矩阵$\boldsymbol{P},\boldsymbol{Q}$，使得

$$\boldsymbol{PAQ}=\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}.$$

七、矩阵的秩

1. 定义

设$\boldsymbol{A}$是$m\times n$矩阵，若存在k阶子式不为零，而任意$k+1$阶子式全为零（如果有的话），则$r(\boldsymbol{A})=k$，且若$\boldsymbol{A}$为$n\times n$矩阵，则

$$r(\boldsymbol{A}_{n\times n})=n\Leftrightarrow|\boldsymbol{A}|\neq0\Leftrightarrow\boldsymbol{A}\text{ 可逆}.$$

2. 初等变换不改变矩阵的秩

性质　设$\boldsymbol{A}$是$m\times n$矩阵，$\boldsymbol{P},\boldsymbol{Q}$分别是$m$阶、$n$阶可逆矩阵，则

$$r(\boldsymbol{A})=r(\boldsymbol{PA})=r(\boldsymbol{AQ})=r(\boldsymbol{PAQ}).$$

3. 有关秩的几个重要式子

设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{B}$ 是满足有关矩阵运算要求的矩阵，则

①$0\leqslant r(\boldsymbol{A})\leqslant\min\{m,n\}$（由定义）；

②$r(k\boldsymbol{A})=r(\boldsymbol{A})\ (k\neq0)$（由定义）；

③$r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}$（见例 2.3.10）；

④$r(\boldsymbol{A}+\boldsymbol{B})\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$（见例 2.3.11）；

⑤$r(\boldsymbol{A}^*)=\begin{cases}n, & r(\boldsymbol{A})=n,\\ 1, & r(\boldsymbol{A})=n-1,\text{其中 }\boldsymbol{A}\text{ 为 }n\text{ 阶方阵（见例 2.3.12).}\\ 0, & r(\boldsymbol{A})<n-1,\end{cases}$

基础例题精解

一、矩阵的基本运算

例 2.2.1 设 $\boldsymbol{\alpha}=[a_1,a_2,a_3]^{\mathrm{T}},\boldsymbol{\beta}=[b_1,b_2,b_3]^{\mathrm{T}},\boldsymbol{A}=\boldsymbol{\alpha\beta}^{\mathrm{T}}$，则 $\boldsymbol{A}^n=$__________.

解 应填 $\left(\sum\limits_{i=1}^{3}a_ib_i\right)^{n-1}\boldsymbol{A}$.

因 $\boldsymbol{A}^n=(\boldsymbol{\alpha\beta}^{\mathrm{T}})(\boldsymbol{\alpha\beta}^{\mathrm{T}})\cdots(\boldsymbol{\alpha\beta}^{\mathrm{T}})=\boldsymbol{\alpha}(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha})(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha})\cdots(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha})\boldsymbol{\beta}^{\mathrm{T}}$

$$=(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha})^{n-1}\boldsymbol{\alpha\beta}^{\mathrm{T}}=\left([b_1,b_2,b_3]\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}\right)^{n-1}\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}[b_1,b_2,b_3]=\left(\sum_{i=1}^{3}a_ib_i\right)^{n-1}\boldsymbol{A}.$$

【注】 (1)矩阵乘法没有交换律，但结合律、分配律是成立的，要充分运用.利用结合律是本题简化运算的关键，因 $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}$ 是数（一阶矩阵），可提到前面.

(2)本题显然具有代表性.

若 $\boldsymbol{\alpha}=[1,2,3]^{\mathrm{T}},\boldsymbol{\beta}=\left[1,\frac{1}{2},\frac{1}{3}\right]^{\mathrm{T}}$，则 $(\boldsymbol{\alpha\beta}^{\mathrm{T}})^n=(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha})^{n-1}\boldsymbol{\alpha\beta}^{\mathrm{T}}=3^{n-1}\begin{bmatrix}1&\frac{1}{2}&\frac{1}{3}\\2&1&\frac{2}{3}\\3&\frac{3}{2}&1\end{bmatrix}$.

若 $\boldsymbol{\alpha}=\boldsymbol{\beta}=[a_1,a_2,a_3]^{\mathrm{T}}$，则 $(\boldsymbol{\alpha\beta}^{\mathrm{T}})^n=(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha})^{n-1}\boldsymbol{\alpha\alpha}^{\mathrm{T}}=\left(\sum\limits_{i=1}^{3}a_i^2\right)^{n-1}\boldsymbol{\alpha\alpha}^{\mathrm{T}}$.

(3)$r(\boldsymbol{A})=r\left(\begin{bmatrix}a_1b_1&a_1b_2&a_1b_3\\a_2b_1&a_2b_2&a_2b_3\\a_3b_1&a_3b_2&a_3b_3\end{bmatrix}\right)=1$，行向量（或列向量）成比例，列向量是 $\boldsymbol{\alpha}=[a_1,a_2,a_3]^{\mathrm{T}}$，比例系数为 b_1,b_2,b_3，提出比例系数，有

$$A=\begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix}=\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}[b_1,b_2,b_3].$$

上式从右到左是作乘法，从左到右是将 A 分解. 当 $r(A)=1$ 时，都可作这种分解. 以后可以看到，众多题目用到这种分解方法，读者需牢记.

例 2.2.2 设 $A=\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$，计算 $A^n(n\geqslant 2)$.

解 因

$$A^2=AA=\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}=\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}=4E,$$

故有

$$A^n=\begin{cases} 4^kE, & n=2k, \\ 4^kA, & n=2k+1, \end{cases} \quad k=1,2,\cdots.$$

例 2.2.3 已知 $A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$，求 A^n.

解

$$A=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}+\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}=E+B,$$

则

$$A^n=(E+B)^n=E^n+nE^{n-1}B+\frac{n(n-1)}{2!}E^{n-2}B^2+\frac{n(n-1)(n-2)}{3!}E^{n-3}B^3+\cdots+B^n,$$

因 E 和任何矩阵可交换，故展开式和初等代数中的展开式一样. 因

$$B=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\quad B^2=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B^3=B^2B=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

故 $B^k=O,k\geqslant 3$，所以

$$A^n=E+nB+\frac{n(n-1)}{2}B^2$$

$$=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}+\begin{bmatrix} 0 & n & 0 \\ 0 & 0 & n \\ 0 & 0 & 0 \end{bmatrix}+\begin{bmatrix} 0 & 0 & \frac{n(n-1)}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}.$$

【注】 因单位矩阵和任何同阶方阵可交换，故本题将 $\boldsymbol{A}$ 分解成 $\boldsymbol{A}=\boldsymbol{E}+\boldsymbol{B}$，再利用主对角元素为零的上三角矩阵 $\boldsymbol{B}$，有 $\boldsymbol{B}^n=\boldsymbol{O}\ (n\geqslant 3)$ 的特性，来求解 $\boldsymbol{A}^n$ 是非常方便的.

例 2.2.4 与向量 $\boldsymbol{\alpha}_1=[2,-1,-3]$，$\boldsymbol{\alpha}_2=[-3,1,5]$都正交的单位向量 $\boldsymbol{\beta}^{\circ}=$__________.

解 应填$\pm\dfrac{1}{\sqrt{6}}[2,1,1]$.

设向量 $\boldsymbol{\beta}=[x_1,x_2,x_3]$与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 都正交，则有

$$\begin{cases}(\boldsymbol{\alpha}_1,\boldsymbol{\beta})=2x_1-x_2-3x_3=0,\\(\boldsymbol{\alpha}_2,\boldsymbol{\beta})=-3x_1+x_2+5x_3=0.\end{cases}\qquad(*)$$

解方程组$(*)$，由

$$\begin{bmatrix}2&-1&-3\\-3&1&5\end{bmatrix}\longrightarrow\begin{bmatrix}-1&0&2\\-3&1&5\end{bmatrix}\longrightarrow\begin{bmatrix}-1&0&2\\0&1&-1\end{bmatrix},$$

解得 $\boldsymbol{\beta}=k[2,1,1]$，$k$ 是任意常数.

取 $\boldsymbol{\beta}^{\circ}=\dfrac{\pm\boldsymbol{\beta}}{\|\boldsymbol{\beta}\|}=\pm\dfrac{1}{\sqrt{6}}[2,1,1]$，则 $\boldsymbol{\beta}^{\circ}$为所求的与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 都正交的单位向量.

【注】 与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 都正交的单位向量有两个，即 $\boldsymbol{\beta}^{\circ}=\pm\dfrac{\boldsymbol{\beta}}{\|\boldsymbol{\beta}\|}$.

例 2.2.5 设向量组 $\boldsymbol{\alpha}_1=[1,1,1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[0,1,1]^{\mathrm{T}}$，用施密特正交化方法将向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 化成标准正交向量组.

解 取 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=[1,1,1]^{\mathrm{T}}$；

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\begin{bmatrix}0\\1\\1\end{bmatrix}-\frac{2}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}-\dfrac{2}{3}\\\dfrac{1}{3}\\\dfrac{1}{3}\end{bmatrix}.$$

将 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 单位化，得

$$\boldsymbol{\xi}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix},\quad\boldsymbol{\xi}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}=\frac{1}{\sqrt{6}}\begin{bmatrix}-2\\1\\1\end{bmatrix},$$

则 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 即为所求的标准正交向量组.

例 2.2.6 设 $\boldsymbol{A}=\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}$，其中 $\boldsymbol{\xi}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，且 $\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}=1$. 证明：

(1)$\boldsymbol{A}$ 是对称矩阵；

(2)$\boldsymbol{A}^2=\boldsymbol{E}$；

(3)$\boldsymbol{A}$ 是正交矩阵.

证明 (1)$\boldsymbol{A}^{\mathrm{T}}=(\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{E}^{\mathrm{T}}-2(\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}=\boldsymbol{A}$，故 $\boldsymbol{A}$ 是对称矩阵.

(2)
$$\begin{aligned}\boldsymbol{A}^2&=(\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}})(\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}})=\boldsymbol{E}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}-2\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}+4\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}\\&=\boldsymbol{E}-4\boldsymbol{\xi}\boldsymbol{\xi}^{\mathrm{T}}+4\boldsymbol{\xi}(\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi})\boldsymbol{\xi}^{\mathrm{T}}=\boldsymbol{E}.\end{aligned}$$

(3)由(1)，(2)知，$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，$\boldsymbol{A}^2=\boldsymbol{E}$，得 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E}$，故 $\boldsymbol{A}$ 是正交矩阵.

二、证明 $\boldsymbol{A}$ 可逆及求 $\boldsymbol{A}^{-1}$ 的方法

1. 求数值矩阵 $\boldsymbol{A}$ 的逆矩阵

方法一 利用下述公式求矩阵的逆矩阵：

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}=\frac{1}{|\boldsymbol{A}|}\begin{bmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{bmatrix},$$

注意 A_{ij} 的位置及正、负号.

例 2.2.7 已知 $\boldsymbol{A}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$，写出 $\boldsymbol{A}$ 可逆的一个充要条件，当 $\boldsymbol{A}$ 可逆时，求 $\boldsymbol{A}^{-1}$.

解 $\boldsymbol{A}$ 可逆 $\Leftrightarrow|\boldsymbol{A}|=\begin{vmatrix}a&b\\c&d\end{vmatrix}=ad-bc\neq0$，且当 $ad-bc\neq0$ 时，

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}=\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}.$$

【注】 利用公式 $\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}$ 求 $\boldsymbol{A}^{-1}$ 时，应注意：

(1)不要忘记 $\boldsymbol{A}^{*}$ 要除以 $|\boldsymbol{A}|$；

(2)$\boldsymbol{A}^{*}$ 的元素是 $|\boldsymbol{A}|$ 中元素的代数余子式，注意正、负号；

(3)A_{ij} 位于矩阵 $\boldsymbol{A}^{*}$ 中相对于矩阵 $\boldsymbol{A}$ 的元素 a_{ji} 的位置上.

对于 2 阶矩阵求 $\boldsymbol{A}^{*}$，只需将 a_{11}，a_{22} 互换，及 a_{12}，a_{21} 上添加负号.

例 2.2.8 求矩阵 $\boldsymbol{A}=\begin{bmatrix}3&2&1\\1&1&1\\1&0&1\end{bmatrix}$ 的逆矩阵.

分析 利用公式 $\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}$.

解 $|\boldsymbol{A}|=\begin{vmatrix}3&2&1\\1&1&1\\1&0&1\end{vmatrix}=2\neq0$，$\boldsymbol{A}$ 可逆. 记 $\boldsymbol{A}=(a_{ij})_{3\times3}$，各元素的代数余子式分别是

$$A_{11}=\begin{vmatrix}1&1\\0&1\end{vmatrix}=1,\quad A_{12}=-\begin{vmatrix}1&1\\1&1\end{vmatrix}=0,\quad A_{13}=\begin{vmatrix}1&1\\1&0\end{vmatrix}=-1,$$

$$A_{21}=-\begin{vmatrix}2&1\\0&1\end{vmatrix}=-2,\quad A_{22}=\begin{vmatrix}3&1\\1&1\end{vmatrix}=2,\quad A_{23}=-\begin{vmatrix}3&2\\1&0\end{vmatrix}=2,$$

$$A_{31}=\begin{vmatrix}2&1\\1&1\end{vmatrix}=1,\quad A_{32}=-\begin{vmatrix}3&1\\1&1\end{vmatrix}=-2,\quad A_{33}=\begin{vmatrix}3&2\\1&1\end{vmatrix}=1.$$

故

$$A^{-1}=\frac{1}{|A|}A^*=\frac{1}{2}\begin{bmatrix}1&-2&1\\0&2&-2\\-1&2&1\end{bmatrix}.$$

【注】 (1)$A_{ij}=(-1)^{i+j}M_{ij}$，注意正、负号.

(2)注意 A_{ij} 在 A^* 中的位置：A_{11}, A_{22}, A_{33} 是主对角线元素，A 的上三角元素的代数余子式 A_{12}，A_{13}, A_{23} 放在 A^* 的下三角对称的位置上，同样，A_{21}, A_{31}, A_{32} 放在 A^* 的上三角对称的位置上.

方法二 用初等行变换求矩阵的逆矩阵：

$$[A\,\vdots\,E]\xrightarrow{\text{初等行变换}}[E\,\vdots\,A^{-1}].$$

例 2.2.9 设 $A=\begin{bmatrix}0&2&-1\\1&1&2\\-1&-1&-1\end{bmatrix}$，求 A^{-1}.

解

$$\left[\begin{array}{ccc:ccc}0&2&-1&1&0&0\\1&1&2&0&1&0\\-1&-1&-1&0&0&1\end{array}\right]\xrightarrow{①\leftrightarrow②}\left[\begin{array}{ccc:ccc}1&1&2&0&1&0\\0&2&-1&1&0&0\\-1&-1&-1&0&0&1\end{array}\right]$$

$$\xrightarrow{③+①}\left[\begin{array}{ccc:ccc}1&1&2&0&1&0\\0&2&-1&1&0&0\\0&0&1&0&1&1\end{array}\right]\xrightarrow[②+③]{①-2③}\left[\begin{array}{ccc:ccc}1&1&0&0&-1&-2\\0&2&0&1&1&1\\0&0&1&0&1&1\end{array}\right]$$

$$\xrightarrow{\frac{1}{2}②}\left[\begin{array}{ccc:ccc}1&1&0&0&-1&-2\\0&1&0&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\0&0&1&0&1&1\end{array}\right]\xrightarrow{①-②}\left[\begin{array}{ccc:ccc}1&0&0&-\frac{1}{2}&-\frac{3}{2}&-\frac{5}{2}\\0&1&0&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\0&0&1&0&1&1\end{array}\right],$$

故

$$A^{-1}=\begin{bmatrix}-\frac{1}{2}&-\frac{3}{2}&-\frac{5}{2}\\\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\0&1&1\end{bmatrix}.$$

【注】 (1)关于用公式求逆矩阵，由于求解 A^* 要计算 n^2 个 $n-1$ 阶行列式，同时还要计算 $|A|$，计算量较大，一般只适用于一些特殊的或阶数较低的矩阵. 但公式本身及矩阵可逆的充要条件在理论上很重要，要充分重视，且无论 $|A|$ 是否为零，$AA^*=|A|E$ 总是成立的. 用初等行变换求逆矩阵则是一种更方便且实用的方法，应熟练掌握. 注意运算正确，每年的试卷中这类数值运算的出错率是很高的.

(2)为了保证计算正确，求出的逆矩阵应予以验证：$AA^{-1}=E$ 或 $A^{-1}A=E$.

2. 求抽象矩阵的逆矩阵

方法一　利用定义求逆矩阵.

根据题设条件，找出一个矩阵 $\boldsymbol{B}$，使 $\boldsymbol{AB}=\boldsymbol{E}$，则 $\boldsymbol{A}$ 可逆，且 $\boldsymbol{A}^{-1}=\boldsymbol{B}$.

例 2.2.10　$\boldsymbol{A},\boldsymbol{B}$ 均是 n 阶矩阵，且 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$. 证明 $\boldsymbol{A}-\boldsymbol{E}$ 可逆，并求 $(\boldsymbol{A}-\boldsymbol{E})^{-1}$.

证明　因 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$，即

$$\boldsymbol{AB}-\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{O},\quad \boldsymbol{AB}-\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{E}=\boldsymbol{E},$$
$$\boldsymbol{A}(\boldsymbol{B}-\boldsymbol{E})-(\boldsymbol{B}-\boldsymbol{E})=\boldsymbol{E},$$

即

$$(\boldsymbol{A}-\boldsymbol{E})(\boldsymbol{B}-\boldsymbol{E})=\boldsymbol{E},\tag{$*$}$$

故 $\boldsymbol{A}-\boldsymbol{E}$ 可逆，且

$$(\boldsymbol{A}-\boldsymbol{E})^{-1}=\boldsymbol{B}-\boldsymbol{E}.$$

【注】　(1)利用定义求逆矩阵，要求将矩阵满足的关系式 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$ 作恒等变形，化成“左端为待求逆矩阵与另一个矩阵的乘积，右端为单位矩阵”的形式.

(2)把矩阵的和、差关系化成乘积的一种方法是提取公因子，提公因子时，左边的矩阵(如 $\boldsymbol{A}$)往左边提，右边的矩阵(如 $\boldsymbol{B}-\boldsymbol{E}$)往右边提.

(3)本题中式($*$)成立已经说明 $\boldsymbol{A}-\boldsymbol{E}$ 可逆，当然也可在式($*$)两边取行列式，说明 $|\boldsymbol{A}-\boldsymbol{E}|\neq 0$，从而先证 $\boldsymbol{A}-\boldsymbol{E}$ 可逆，再求 $(\boldsymbol{A}-\boldsymbol{E})^{-1}$.

方法二　将 $\boldsymbol{A}$ 分解成若干个可逆矩阵的乘积.

若 $\boldsymbol{A}=\boldsymbol{BC}$，其中 $\boldsymbol{B},\boldsymbol{C}$ 均是可逆矩阵，则 $\boldsymbol{A}$ 可逆，且 $\boldsymbol{A}^{-1}=\boldsymbol{C}^{-1}\boldsymbol{B}^{-1}$.

例 2.2.11　设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶可逆方阵，且 $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 是可逆矩阵，证明 $\boldsymbol{A}+\boldsymbol{B}$ 是可逆矩阵，并求 $(\boldsymbol{A}+\boldsymbol{B})^{-1}$.

证明　因 $\boldsymbol{A},\boldsymbol{B}$ 是可逆矩阵，故

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B})=\boldsymbol{A}(\boldsymbol{B}^{-1}+\boldsymbol{A}^{-1})\boldsymbol{B}.$$

又因 $\boldsymbol{A},\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1},\boldsymbol{B}$ 可逆，故 $\boldsymbol{A}+\boldsymbol{B}$ 可逆，且

$$(\boldsymbol{A}+\boldsymbol{B})^{-1}=[\boldsymbol{A}(\boldsymbol{B}^{-1}+\boldsymbol{A}^{-1})\boldsymbol{B}]^{-1}=\boldsymbol{B}^{-1}(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}\boldsymbol{A}^{-1}.$$

【注】　(1)要证 $\boldsymbol{A}+\boldsymbol{B}$ 可逆，并求 $(\boldsymbol{A}+\boldsymbol{B})^{-1}$，应设法将 $\boldsymbol{A}+\boldsymbol{B}$ 化成已知可逆矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 的乘积. 这里表面上 $\boldsymbol{A}+\boldsymbol{B}$ 没有公因子，但 $\boldsymbol{A}$ 可逆，故有

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{A}+\boldsymbol{EB}=\boldsymbol{A}+\boldsymbol{AA}^{-1}\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B}).$$

同理，$\boldsymbol{B}$ 可逆，有

$$\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{A}^{-1}\boldsymbol{B})=\boldsymbol{A}(\boldsymbol{B}^{-1}\boldsymbol{B}+\boldsymbol{A}^{-1}\boldsymbol{B})=\boldsymbol{A}(\boldsymbol{B}^{-1}+\boldsymbol{A}^{-1})\boldsymbol{B},$$

从而将 $\boldsymbol{A}+\boldsymbol{B}$ 化成了已知可逆矩阵的乘积. 提公因子是和、差化积的一个好办法.

(2)$\boldsymbol{A},\boldsymbol{B}$ 均可逆，并不能得出 $\boldsymbol{A}+\boldsymbol{B}$ 可逆，在增加条件 $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$ 可逆时，才能保证 $\boldsymbol{A}+\boldsymbol{B}$ 可逆.

3. 分块矩阵的逆

例 2.2.12 设

$$A=\begin{bmatrix}0&0&1&-1&0\\0&0&0&1&-1\\0&0&0&0&1\\1&3&0&0&0\\-1&2&0&0&0\end{bmatrix},$$

求 A^{-1}.

解 将 A 分块如下，

$$A=\left[\begin{array}{cc|ccc}0&0&1&-1&0\\0&0&0&1&-1\\0&0&0&0&1\\\hline 1&3&0&0&0\\-1&2&0&0&0\end{array}\right]\overset{记}{=\!=}\begin{bmatrix}O&B\\C&O\end{bmatrix},$$

其中，

$$B=\begin{bmatrix}1&-1&0\\0&1&-1\\0&0&1\end{bmatrix},\quad C=\begin{bmatrix}1&3\\-1&2\end{bmatrix}.$$

因

$$B^{-1}=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix},\quad C^{-1}=\begin{bmatrix}\frac{2}{5}&-\frac{3}{5}\\\frac{1}{5}&\frac{1}{5}\end{bmatrix},$$

故

$$A^{-1}=\begin{bmatrix}O&C^{-1}\\B^{-1}&O\end{bmatrix}=\begin{bmatrix}0&0&0&\frac{2}{5}&-\frac{3}{5}\\0&0&0&\frac{1}{5}&\frac{1}{5}\\1&1&1&0&0\\0&1&1&0&0\\0&0&1&0&0\end{bmatrix}.$$

【注】 副对角线分块矩阵的逆，可推广如下，

$$A=\begin{bmatrix}&&&A_1\\&&A_2&\\&\mathinner{\cdot^{\cdot^{\cdot}}}&&\\A_s&&&\end{bmatrix},$$

其中 $A_i(i=1,2,\cdots,s)$ 可逆，则 A 可逆，且

$$A^{-1}=\begin{bmatrix}&&&A_s^{-1}\\&&\mathinner{\cdot^{\cdot^{\cdot}}}&\\&A_2^{-1}&&\\A_1^{-1}&&&\end{bmatrix}.$$

例 2.2.13 已知 $A=\begin{bmatrix}B & O\\ D & C\end{bmatrix}$，其中 B 是 $r\times r$ 可逆矩阵，C 是 $s\times s$ 可逆矩阵. 证明 A 可逆，并求 A^{-1}.

证明 因 $|A|=\begin{vmatrix}B & O\\ D & C\end{vmatrix}=|B|\,|C|\neq 0$，故 A 可逆. 设 $A^{-1}=\begin{bmatrix}X & Y\\ Z & W\end{bmatrix}$，由定义，有

$$AA^{-1}=\begin{bmatrix}B & O\\ D & C\end{bmatrix}\begin{bmatrix}X & Y\\ Z & W\end{bmatrix}=\begin{bmatrix}BX & BY\\ DX+CZ & DY+CW\end{bmatrix}=\begin{bmatrix}E & O\\ O & E\end{bmatrix},$$

得

$$BX=E\Rightarrow X=B^{-1},\quad BY=O\Rightarrow Y=O(B\text{ 可逆}),$$

$$DX+CZ=O\Rightarrow Z=-C^{-1}DB^{-1}(X=B^{-1}),\quad DY+CW=E\Rightarrow W=C^{-1}(Y=O),$$

故

$$A^{-1}=\begin{bmatrix}B^{-1} & O\\ -C^{-1}DB^{-1} & C^{-1}\end{bmatrix}.$$

【注】 若

$$A_1=\begin{bmatrix}B & D\\ O & C\end{bmatrix},\quad A_2=\begin{bmatrix}O & B\\ C & D\end{bmatrix},\quad A_3=\begin{bmatrix}D & B\\ C & O\end{bmatrix},$$

其中 B,C 可逆，则有

$$A_1^{-1}=\begin{bmatrix}B^{-1} & -B^{-1}DC^{-1}\\ O & C^{-1}\end{bmatrix},\quad A_2^{-1}=\begin{bmatrix}-C^{-1}DB^{-1} & C^{-1}\\ B^{-1} & O\end{bmatrix},\quad A_3^{-1}=\begin{bmatrix}O & C^{-1}\\ B^{-1} & -B^{-1}DC^{-1}\end{bmatrix}.$$

请读者用本题的方法推得此类求逆公式，不必记忆，但应理解推导方法，要用时，临时推导即可.

三、伴随矩阵及其运算

例 2.2.14 设 $A=\begin{bmatrix}1 & 2 & 0\\ 2 & 2 & 0\\ 3 & 4 & 5\end{bmatrix}$，$A^*$ 是 A 的伴随矩阵，则 $(A^*)^{-1}=$________.

分析 利用公式 $AA^*=|A|E$. 当 $|A|\neq 0$ 时，有 $\frac{A}{|A|}A^*=E$，故 $(A^*)^{-1}=\frac{A}{|A|}$，这样不需要先求出 A^*，再求 $(A^*)^{-1}$.

解 应填 $-\frac{1}{10}\begin{bmatrix}1 & 2 & 0\\ 2 & 2 & 0\\ 3 & 4 & 5\end{bmatrix}$.

$$|A|=\begin{vmatrix}1 & 2 & 0\\ 2 & 2 & 0\\ 3 & 4 & 5\end{vmatrix}=-10\neq 0,$$

由公式 $AA^*=|A|E$，得 $\frac{A}{|A|}A^*=E$，故 A^* 可逆，且 $(A^*)^{-1}=\frac{A}{|A|}=-\frac{1}{10}\begin{bmatrix}1 & 2 & 0\\ 2 & 2 & 0\\ 3 & 4 & 5\end{bmatrix}$.

例 2.2.15 设 A 是 3 阶矩阵，已知 $A^{-1}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}$，则 $|A^*|=$________.

解 应填$\frac{1}{16}$.

由 $AA^*=|A|E$，当 A 可逆时，有 $A^*=|A|A^{-1}$.

两边取行列式，得

$$|A^*|=||A|A^{-1}|=|A|^n\cdot|A^{-1}|=\frac{|A|^n}{|A|}=|A|^{n-1}=\frac{1}{|A^{-1}|^{n-1}},$$

其中 $n=3$，

$$|A^{-1}|=\begin{vmatrix}2&1&1\\1&2&1\\1&1&2\end{vmatrix}=\begin{vmatrix}0&-3&-1\\1&2&1\\0&-1&1\end{vmatrix}=4,$$

故 $|A^*|=\frac{1}{4^{3-1}}=\frac{1}{16}$.

例 2.2.16 设 A,B 是 n 阶方阵，$|A|=2$，$|B|=-4$，则 $|2B^*A^{-1}|=$________.

解 应填$(-8)^{n-1}$.

因 $|kA|=k^n|A|$，$|AB|=|A||B|$，$|A^{-1}|=\frac{1}{|A|}$，$|B^*|=|B|^{n-1}$，故

$$|2B^*A^{-1}|=2^n|B^*||A^{-1}|=2^n|B|^{n-1}\cdot\frac{1}{|A|}=(-8)^{n-1}.$$

四、初等变换与初等矩阵

例 2.2.17 设 A 是 3 阶可逆矩阵，将 A 的第 1 行和第 2 行互换后得到矩阵 B，其中 $A^{-1}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}$，则 B 可逆，且 $B^{-1}=$________.

解 应填$\begin{bmatrix}a_{12}&a_{11}&a_{13}\\a_{22}&a_{21}&a_{23}\\a_{32}&a_{31}&a_{33}\end{bmatrix}$.

因 $B=E_{12}A$，

其中 $E_{12}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$，故 $B^{-1}=(E_{12}A)^{-1}=A^{-1}E_{12}^{-1}=A^{-1}E_{12}=\begin{bmatrix}a_{12}&a_{11}&a_{13}\\a_{22}&a_{21}&a_{23}\\a_{32}&a_{31}&a_{33}\end{bmatrix}$.

例 2.2.18 设 A 是 3 阶矩阵，将 A 的第 1 列与第 2 列互换得到 B，再将 B 的第 2 列加到第 3 列得到 C，则满足 $AQ=C$ 的可逆矩阵 Q 为(　　).

(A) $\begin{bmatrix}0&1&0\\1&0&0\\1&0&1\end{bmatrix}$　(B) $\begin{bmatrix}0&1&0\\1&0&1\\0&0&1\end{bmatrix}$　(C) $\begin{bmatrix}0&1&0\\1&0&0\\0&1&1\end{bmatrix}$　(D) $\begin{bmatrix}0&1&1\\1&0&0\\0&0&1\end{bmatrix}$

解 应选(D).

将 $\boldsymbol{A}$ 的第 1 列与第 2 列互换得到 $\boldsymbol{B}$，即

$$\boldsymbol{A}\boldsymbol{E}_{12}=\boldsymbol{A}\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}=\boldsymbol{B}.$$

再将 $\boldsymbol{B}$ 的第 2 列加到第 3 列得到 $\boldsymbol{C}$，即

$$\boldsymbol{B}\boldsymbol{E}_{23}(1)=\boldsymbol{B}\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}=\boldsymbol{C},$$

从而有

$$(\boldsymbol{A}\boldsymbol{E}_{12})\boldsymbol{E}_{23}(1)=\boldsymbol{A}\boldsymbol{E}_{12}\boldsymbol{E}_{23}(1)=\boldsymbol{A}\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}0&1&1\\1&0&0\\0&0&1\end{bmatrix}=\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{C},$$

其中 $\boldsymbol{Q}=\begin{bmatrix}0&1&1\\1&0&0\\0&0&1\end{bmatrix}$，故应选(D).

五、矩阵方程

含有未知矩阵的方程称为**矩阵方程**.

解矩阵方程，应先根据题设条件和矩阵的运算规则，将方程进行恒等变形，使方程化成 $\boldsymbol{AX}=\boldsymbol{B}$，$\boldsymbol{XA}=\boldsymbol{B}$ 或 $\boldsymbol{AXB}=\boldsymbol{C}$ 的形式.

若 $\boldsymbol{A}$，或 $\boldsymbol{A}$ 且 $\boldsymbol{B}$ 可逆，则分别可得解为 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$，$\boldsymbol{X}=\boldsymbol{B}\boldsymbol{A}^{-1}$，$\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{B}^{-1}$.

例 2.2.19 设 3 阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}=6\boldsymbol{A}+\boldsymbol{B}\boldsymbol{A}$，且

$$\boldsymbol{A}=\begin{bmatrix}\frac{1}{3}&0&0\\0&\frac{1}{4}&0\\0&0&\frac{1}{5}\end{bmatrix},$$

求 $\boldsymbol{B}$.

解 显然 $\boldsymbol{A}$ 可逆，用 $\boldsymbol{A}^{-1}$ 右乘方程两边，得

$$\boldsymbol{A}^{-1}\boldsymbol{B}=6\boldsymbol{E}+\boldsymbol{B},$$

即

$$(\boldsymbol{A}^{-1}-\boldsymbol{E})\boldsymbol{B}=6\boldsymbol{E},\quad \boldsymbol{B}=6(\boldsymbol{A}^{-1}-\boldsymbol{E})^{-1}.$$

$$\boldsymbol{A}^{-1}=\begin{bmatrix}3&0&0\\0&4&0\\0&0&5\end{bmatrix},\quad \boldsymbol{A}^{-1}-\boldsymbol{E}=\begin{bmatrix}2&0&0\\0&3&0\\0&0&4\end{bmatrix},\quad (\boldsymbol{A}^{-1}-\boldsymbol{E})^{-1}=\begin{bmatrix}\frac{1}{2}&0&0\\0&\frac{1}{3}&0\\0&0&\frac{1}{4}\end{bmatrix},$$

得
$$B=6(A^{-1}-E)^{-1}=\begin{bmatrix}3&0&0\\0&2&0\\0&0&\frac{3}{2}\end{bmatrix}.$$

例 2.2.20 设矩阵 A 的伴随矩阵

$$A^*=\begin{bmatrix}1&0&0&0\\0&1&0&0\\1&0&1&0\\0&-3&0&8\end{bmatrix},$$

且 $ABA^{-1}=BA^{-1}+3E$，其中 E 是 4 阶单位矩阵，求矩阵 B.

解 由题设条件得
$$(A-E)BA^{-1}=3E,$$
$$B=3(A-E)^{-1}A \tag{①}$$
$$=3(A-E)^{-1}(A^{-1})^{-1}$$
$$=3[A^{-1}(A-E)]^{-1}=3(E-A^{-1})^{-1} \tag{②}$$
$$=3\left(E-\frac{A^*}{|A|}\right)^{-1},$$

其中由 $AA^*=|A|E$，$|A^*|=|A|^{n-1}$，得 $|A|^3=|A^*|=8$，即 $|A|=2$，故 $B=3\left(E-\dfrac{A^*}{|A|}\right)^{-1}=6(2E-A^*)^{-1}$.

由
$$2E-A^*=\begin{bmatrix}1&0&0&0\\0&1&0&0\\-1&0&1&0\\0&3&0&-6\end{bmatrix},\quad (2E-A^*)^{-1}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\1&0&1&0\\0&\frac{1}{2}&0&-\frac{1}{6}\end{bmatrix},$$

可得
$$B=\begin{bmatrix}6&0&0&0\\0&6&0&0\\6&0&6&0\\0&3&0&-1\end{bmatrix}.$$

【注】 化简到表达式①或②，即可直接计算 B，但要先求出 A 或 A^{-1}，此时还必须利用 $AA^*=|A|E$ 及 $|A|^3=|A^*|$ 求出 $|A|$，并由 A^* 求出 A 或 A^{-1}.

六、矩阵的秩和等价矩阵

例 2.2.21 设 $A=\begin{bmatrix}1&1&1&-1\\1&3&x&1\\-3&-5&-2&1\\3&5&y&-1\end{bmatrix}$，已知 $r(A)=2$，则参数 x,y 分别为__________.

解 应填 0，2.

对 A 作初等变换，有

$$
\boldsymbol{A}=\begin{bmatrix}1&1&1&-1\\1&3&x&1\\-3&-5&-2&1\\3&5&y&-1\end{bmatrix}\xrightarrow[\substack{[2]+[4]\\ [3]+[4]}]{[1]+[4]}\begin{bmatrix}0&0&0&-1\\2&4&x+1&1\\-2&-4&-1&1\\2&4&y-1&-1\end{bmatrix}\xrightarrow[\substack{③+①\\ ④-①}]{②+①}\begin{bmatrix}0&0&0&-1\\2&4&x+1&0\\-2&-4&-1&0\\2&4&y-1&0\end{bmatrix}
$$

$$
\xrightarrow[③+②]{④+③}\begin{bmatrix}0&0&0&-1\\2&4&x+1&0\\0&0&x&0\\0&0&y-2&0\end{bmatrix}=\boldsymbol{A}'.
$$

由 $r(\boldsymbol{A})=r(\boldsymbol{A}')=2$，得 $x=0,y=2$.

【注】 初等变换不改变矩阵的秩.这里可以只作初等行变换，也可以只作初等列变换，也可以既作初等行变换，又作初等列变换.

例 2.2.22 设 $n(n\geqslant 3)$ 阶矩阵

$$
\boldsymbol{A}=\begin{bmatrix}1&a&a&\cdots&a\\a&1&a&\cdots&a\\a&a&1&\cdots&a\\\vdots&\vdots&\vdots&&\vdots\\a&a&a&\cdots&1\end{bmatrix},
$$

若矩阵 $\boldsymbol{A}$ 的秩为 $n-1$，则 a 为（　　）.

(A) 1　　(B) $\dfrac{1}{1-n}$　　(C) -1　　(D) $\dfrac{1}{n-1}$

解　应选(B).

对 $\boldsymbol{A}$ 作初等变换，有

$$
\boldsymbol{A}=\begin{bmatrix}1&a&a&\cdots&a\\a&1&a&\cdots&a\\a&a&1&\cdots&a\\\vdots&\vdots&\vdots&&\vdots\\a&a&a&\cdots&1\end{bmatrix}\xrightarrow[i=2,3,\cdots,n]{ⓘ-①}\begin{bmatrix}1&a&a&\cdots&a\\a-1&1-a&0&\cdots&0\\a-1&0&1-a&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\a-1&0&0&\cdots&1-a\end{bmatrix}
$$

$$
\xrightarrow{[1]+\sum\limits_{i=2}^{n}[i]}\begin{bmatrix}1+(n-1)a&a&a&\cdots&a\\0&1-a&0&\cdots&0\\0&0&1-a&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&1-a\end{bmatrix}=\boldsymbol{A}'.
$$

因 $r(\boldsymbol{A})=r(\boldsymbol{A}')=n-1$，故应取 $1+(n-1)a=0$，即 $a=\dfrac{-1}{n-1}=\dfrac{1}{1-n}$（因 $n\geqslant 3,a=\dfrac{1}{1-n}\neq 1$），故应选(B).

【注】 也可以利用 $|\boldsymbol{A}|=0$，再作判别.

例 2.2.23 设矩阵

$$A=\begin{bmatrix}1&0&1\\-1&-1&1\\0&2&a\end{bmatrix},\quad B=\begin{bmatrix}1&0&1\\0&-1&2\\0&0&0\end{bmatrix}.$$

(1)当 a 为何值时，矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 等价；

(2)当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 等价时，求一个可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{PA}=\boldsymbol{B}$.

解 (1)$\boldsymbol{A},\boldsymbol{B}$ 同型，$\boldsymbol{A}$ 等价于 $\boldsymbol{B}\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{B})$.

对于 $\boldsymbol{B}$，显然有 $r(\boldsymbol{B})=2$.

对于 $\boldsymbol{A}$，因 $\begin{vmatrix}1&0\\-1&-1\end{vmatrix}\neq0$，要求 $r(\boldsymbol{A})=r(\boldsymbol{B})=2$，故应有

$$|A|=\begin{vmatrix}1&0&1\\-1&-1&1\\0&2&a\end{vmatrix}=\begin{vmatrix}1&0&1\\0&-1&2\\0&2&a\end{vmatrix}=\begin{vmatrix}1&0&1\\0&-1&2\\0&0&a+4\end{vmatrix}=-(a+4)=0,$$

解得 $a=-4$.

因此，当 $a=-4$ 时，有 $r(\boldsymbol{A})=r(\boldsymbol{B})=2$，矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 等价.

(2)当 $a=-4$ 时，

$$A=\begin{bmatrix}1&0&1\\-1&-1&1\\0&2&-4\end{bmatrix}.$$

对 $\boldsymbol{A}$ 作初等行变换，消 a_{21},a_{32} 为零. 因

$$E_{21}(1)A=\begin{bmatrix}1&0&0\\1&1&0\\0&0&1\end{bmatrix}A=\begin{bmatrix}1&0&0\\1&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&1\\-1&-1&1\\0&2&-4\end{bmatrix}$$

$$=\begin{bmatrix}1&0&1\\0&-1&2\\0&2&-4\end{bmatrix}=C,$$

$$E_{32}(2)C=\begin{bmatrix}1&0&0\\0&1&0\\0&2&1\end{bmatrix}\begin{bmatrix}1&0&1\\0&-1&2\\0&2&-4\end{bmatrix}=\begin{bmatrix}1&0&1\\0&-1&2\\0&0&0\end{bmatrix}=B,$$

故有

$$B=E_{32}(2)C=E_{32}(2)E_{21}(1)A.$$

取

$$P=E_{32}(2)E_{21}(1)=\begin{bmatrix}1&0&0\\0&1&0\\0&2&1\end{bmatrix}\begin{bmatrix}1&0&0\\1&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\1&1&0\\2&2&1\end{bmatrix},$$

使得 $\boldsymbol{PA}=\boldsymbol{B}$.

【注】 本题(2)若改为求满足条件的全部矩阵 $\boldsymbol{P}$，可用下述方法.

解 考虑 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}$，将 $\boldsymbol{P}^{\mathrm{T}}$ 按列分块为 $[\xi_1,\xi_2,\xi_3]$.

$$[\boldsymbol{A}^{\mathrm{T}} \vdots \boldsymbol{B}^{\mathrm{T}}]=\left[\begin{array}{ccc:ccc}1&-1&0&1&0&0\\0&-1&2&0&-1&0\\1&1&-4&1&2&0\end{array}\right]\longrightarrow\left[\begin{array}{ccc:ccc}1&-1&0&1&0&0\\0&1&-2&0&1&0\\0&0&0&0&0&0\end{array}\right],$$

解得

$$\xi_1=k_1\begin{bmatrix}2\\2\\1\end{bmatrix}+\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \xi_2=k_2\begin{bmatrix}2\\2\\1\end{bmatrix}+\begin{bmatrix}1\\1\\0\end{bmatrix},\quad \xi_3=k_3\begin{bmatrix}2\\2\\1\end{bmatrix},$$

则有

$$\boldsymbol{P}^{\mathrm{T}}=\begin{bmatrix}2k_1+1&2k_2+1&2k_3\\2k_1&2k_2+1&2k_3\\k_1&k_2&k_3\end{bmatrix},$$

此时，

$$\boldsymbol{P}=\begin{bmatrix}2k_1+1&2k_1&k_1\\2k_2+1&2k_2+1&k_2\\2k_3&2k_3&k_3\end{bmatrix},$$

其中，因为 $\boldsymbol{P}$ 可逆，则 $k_3\neq0$，即 k_1,k_2 为任意常数，k_3 为任意非零常数.

基础习题精练

习 题

2.2.1 设 $\boldsymbol{A}=\begin{bmatrix}1&2&-1\\-2&-4&2\\3&6&-3\end{bmatrix}$，则 $\boldsymbol{A}^n=$ ______.

2.2.2 设 $\boldsymbol{A}=\begin{bmatrix}0&1&0\\1&0&-1\\0&-1&0\end{bmatrix}$，计算 $\boldsymbol{A}^n(n\geqslant2)$.

2.2.3 设 $\boldsymbol{A}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}3&0&0\\0&2&0\\0&0&1\end{bmatrix}$，$\boldsymbol{C}=\begin{bmatrix}1&0&0\\0&1&0\\1&0&1\end{bmatrix}$，则 $(\boldsymbol{ABC})^{-1}=$ ______.

2.2.4 设 n 阶矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{E}=\boldsymbol{O}$.

(1)证明 $\boldsymbol{A},\boldsymbol{A}+2\boldsymbol{E}$ 可逆，并求 $\boldsymbol{A}^{-1}$ 及 $(\boldsymbol{A}+2\boldsymbol{E})^{-1}$；

(2)当 $\boldsymbol{A}\neq\boldsymbol{E}$ 时，判别 $\boldsymbol{A}-2\boldsymbol{E}$ 是否可逆，并说明理由.

2.2.5 设 n 阶矩阵 $\boldsymbol{A}$ 可逆($n\geqslant 2$)，$\boldsymbol{A}^*$ 是矩阵 $\boldsymbol{A}$ 的伴随矩阵，则(　　).

(A)$(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n-1}\boldsymbol{A}$　　(B)$(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n+1}\boldsymbol{A}$

(C)$(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n-2}\boldsymbol{A}$　　(D)$(\boldsymbol{A}^*)^*=|\boldsymbol{A}|^{n+2}\boldsymbol{A}$

2.2.6 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶可逆矩阵，证明：

(1)$(\boldsymbol{A}^{-1})^*=(\boldsymbol{A}^*)^{-1}$；

(2)$(\boldsymbol{AB})^*=\boldsymbol{B}^*\boldsymbol{A}^*$.

2.2.7 设 $\boldsymbol{A}=\begin{bmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{bmatrix}$，则$(\boldsymbol{A}^*)^{-1}=$__________.

2.2.8 设 $\boldsymbol{A}$ 为3阶矩阵，将 $\boldsymbol{A}$ 的第2行加到第1行得到 $\boldsymbol{B}$，再将 $\boldsymbol{B}$ 的第1列的-1倍加到第2列得到 $\boldsymbol{C}$，记

$$\boldsymbol{P}=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix},$$

则(　　).

(A)$\boldsymbol{C}=\boldsymbol{P}^{-1}\boldsymbol{AP}$　　(B)$\boldsymbol{C}=\boldsymbol{PAP}^{-1}$

(C)$\boldsymbol{C}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$　　(D)$\boldsymbol{C}=\boldsymbol{PAP}^{\mathrm{T}}$

2.2.9 设4阶矩阵

$$\boldsymbol{B}=\begin{bmatrix}0&-1&0&0\\0&0&-1&0\\0&0&0&-1\\0&0&0&0\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}2&1&3&4\\0&2&1&3\\0&0&2&1\\0&0&0&2\end{bmatrix},$$

且矩阵 $\boldsymbol{A}$ 满足关系式 $\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{C}^{-1}\boldsymbol{B})^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{E}+\boldsymbol{A}$，求矩阵 $\boldsymbol{A}$.

2.2.10 设3阶矩阵 $\boldsymbol{A}=\begin{bmatrix}a&b&b\\b&a&b\\b&b&a\end{bmatrix}$，其中 $b\neq 0$，已知 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 的秩为 $r(\boldsymbol{A}^*)=1$，则 a,b 应满足关系(　　).

(A)$a=b$　　(B)$a=-2b$　　(C)$a=0$　　(D)$a=2b$

2.2.11 设 $\boldsymbol{A}$ 是 4×3 矩阵，$\boldsymbol{B}$ 是 3×4 的非零矩阵，且满足 $\boldsymbol{AB}=\boldsymbol{O}$，其中

$$\boldsymbol{A}=\begin{bmatrix}1&2&t\\3&t&18\\2&4&2t\\1&8-t&4t-18\end{bmatrix},$$

则(　　).

(A)当 $t\neq 6$ 时，必有 $r(\boldsymbol{B})=1$　　(B)当 $t=6$ 时，必有 $r(\boldsymbol{B})=2$

(C)当 $t\neq 6$ 时，必有 $r(\boldsymbol{B})=2$　　(D)当 $t=6$ 时，必有 $r(\boldsymbol{B})=1$

解答

2.2.1 $(-6)^{n-1}\begin{bmatrix}1 & 2 & -1\\ -2 & -4 & 2\\ 3 & 6 & -3\end{bmatrix}$ 解

$$A=\begin{bmatrix}1 & 2 & -1\\ -2 & -4 & 2\\ 3 & 6 & -3\end{bmatrix}=\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}[1,2,-1]\xlongequal{记}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta},$$

则 $$A^n=(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta})^n=(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta})(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta})\cdots(\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{\beta}\boldsymbol{\alpha}^{\mathrm{T}})\cdots(\boldsymbol{\beta}\boldsymbol{\alpha}^{\mathrm{T}})\boldsymbol{\beta}=(\boldsymbol{\beta}\boldsymbol{\alpha}^{\mathrm{T}})^{n-1}\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta},$$

其中 $\boldsymbol{\beta}\boldsymbol{\alpha}^{\mathrm{T}}=[1,2,-1]\begin{bmatrix}1\\ -2\\ 3\end{bmatrix}=-6$，故 $A^n=(-6)^{n-1}\begin{bmatrix}1 & 2 & -1\\ -2 & -4 & 2\\ 3 & 6 & -3\end{bmatrix}$.

2.2.2 解 因为

$$A^2=AA=\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & -1\\ 0 & -1 & 0\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & -1\\ 0 & -1 & 0\end{bmatrix}=\begin{bmatrix}1 & 0 & -1\\ 0 & 2 & 0\\ -1 & 0 & 1\end{bmatrix},$$

$$A^3=A^2A=\begin{bmatrix}1 & 0 & -1\\ 0 & 2 & 0\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & -1\\ 0 & -1 & 0\end{bmatrix}=\begin{bmatrix}0 & 2 & 0\\ 2 & 0 & -2\\ 0 & -2 & 0\end{bmatrix}=2A,$$

$$A^4=A^3A=2AA=2A^2,$$

故 $$A^n=\begin{cases}2^{k-1}A^2, & n=2k,\\ 2^kA, & n=2k+1,\end{cases}\quad k=1,2,\cdots.$$

2.2.3 $\begin{bmatrix}\frac{1}{3} & 0 & 0\\ 0 & 0 & \frac{1}{2}\\ -\frac{1}{3} & 1 & 0\end{bmatrix}$ 解

$$(ABC)^{-1}=C^{-1}B^{-1}A^{-1}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}^{-1}\begin{bmatrix}3 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}^{-1}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}^{-1}$$

$$=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}\frac{1}{3} & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}=\begin{bmatrix}\frac{1}{3} & 0 & 0\\ 0 & 0 & \frac{1}{2}\\ -\frac{1}{3} & 1 & 0\end{bmatrix}.$$

2.2.4 (1)证明 因 $A^2-3A+2E=O$，故 $A(A-3E)=-2E$，即

$$A\left[-\frac{1}{2}(A-3E)\right]=E,$$

故 A 可逆，且 $A^{-1}=-\frac{1}{2}(A-3E)$.

又 $$A^2-3A+2E=(A+2E)(A-5E)+10E+2E=(A+2E)(A-5E)+12E=O,$$

故 $A+2E$ 可逆，且 $(A+2E)^{-1}=-\frac{1}{12}(A-5E)$.

(2)解 因 $A^2-3A+2E=(A-2E)(A-E)=O$，因 $A\neq E$，$A-E\neq O$，则 $(A-2E)x=0$ 有非零解，故 $A-2E$ 不可逆.

2.2.5 (C) 解 由 $A^*=|A|A^{-1}$，得 $(A^*)^*=|A^*|(A^*)^{-1}$，又 $|A^*|=|A|^{n-1}$，故

$$(A^*)^*=|A|^{n-1}(|A|A^{-1})^{-1}=|A|^{n-1}\frac{1}{|A|}A=|A|^{n-2}A.$$

2.2.6 证明 (1)A 可逆，$r(A)=n\Rightarrow r(A^*)=n$，$A^*$ 也可逆，且有

$$A^{-1}(A^{-1})^*=|A^{-1}|E,$$

$$(A^{-1})^*=|A^{-1}|A=|A^{-1}||A|(A^*)^{-1}=\frac{|A|}{|A|}(A^*)^{-1}=(A^*)^{-1}.$$

(2)由 $AA^*=|A|E$，$BB^*=|B|E$，得 $A^*=|A|A^{-1}$，$B^*=|B|B^{-1}$，则

$$B^*A^*=|B|B^{-1}|A|A^{-1}=|AB|(AB)^{-1}=(AB)^*.$$

【注】 (1)类似地，$(A^{\mathrm{T}})^*=(A^*)^{\mathrm{T}}$($A$ 可逆时).

(2)作为公式 $AA^*=|A|E$，这里的 A 是任意方阵，故有

$$(kA)(kA)^*=|kA|E,\quad A^{\mathrm{T}}(A^{\mathrm{T}})^*=|A^{\mathrm{T}}|E,\quad A^{-1}(A^{-1})^*=|A^{-1}|E,\quad A^*(A^*)^*=|A^*|E,$$

这些关系式只是 $AA^*=|A|E$ 的演变，作为考研的学生，应具备这种推理能力.

2.2.7 $-\frac{1}{16}A$ 解 由 $(A^*)^{-1}=\frac{A}{|A|}$，又

$$|A|=\begin{vmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{vmatrix}=\begin{vmatrix}1&1&1&1\\0&0&-2&-2\\0&-2&0&-2\\0&-2&-2&0\end{vmatrix}=-16,$$

所以 $$(A^*)^{-1}=-\frac{1}{16}A.$$

2.2.8 (B) 解 将 A 的第 2 行加到第 1 行得到 B，即

$$B=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}A=PA.$$

将 B 的第 1 列的 -1 倍加到第 2 列得到 C，即

$$C=B\begin{bmatrix}1&-1&0\\0&1&0\\0&0&1\end{bmatrix}\xlongequal{\text{记}}BQ.$$

因

$$PQ=\begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&-1&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=E,$$

故 $Q=P^{-1}$，从而有 $C=BQ=BP^{-1}=PAP^{-1}$，故应选(B).

2.2.9 解 $(E-C^{-1}B)^{\mathrm{T}}C^{\mathrm{T}}=[C(E-C^{-1}B)]^{\mathrm{T}}=(C-B)^{\mathrm{T}}$，故

$$A(E-C^{-1}B)^{\mathrm{T}}C^{\mathrm{T}}=A(C-B)^{\mathrm{T}},$$

从而

$$A(C-B)^{\mathrm{T}}-A=E,$$

则

$$A[(C-B)^{\mathrm{T}}-E]=A(C-B-E)^{\mathrm{T}}=E.$$

显然 $(C-B-E)^{\mathrm{T}}$ 可逆，则

$$A=[(C-B-E)^{\mathrm{T}}]^{-1}.$$

将 C,B,E 代入，得

$$A=\left[\left(\begin{bmatrix}2&1&3&4\\0&2&1&3\\0&0&2&1\\0&0&0&2\end{bmatrix}-\begin{bmatrix}0&-1&0&0\\0&0&-1&0\\0&0&0&-1\\0&0&0&0\end{bmatrix}-\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\right)^{\mathrm{T}}\right]^{-1}$$

$$=\left(\begin{bmatrix}1&2&3&4\\0&1&2&3\\0&0&1&2\\0&0&0&1\end{bmatrix}^{\mathrm{T}}\right)^{-1}=\begin{bmatrix}1&0&0&0\\2&1&0&0\\3&2&1&0\\4&3&2&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0&0&0\\-2&1&0&0\\1&-2&1&0\\0&1&-2&1\end{bmatrix}.$$

【注】 再次提醒读者：解这类矩阵方程，应先将方程变形，再化简，最后代入具体数值矩阵进行运算. 因为通过化简可使步骤清晰，并减少计算量. 评分标准也是化简有化简的分数，计算有计算的分数，按步骤给分.

2.2.10 (B) 解 $r(A^{*})=1\Rightarrow r(A)=2$. 对 A 作初等变换，有

$$A=\begin{bmatrix}a&b&b\\b&a&b\\b&b&a\end{bmatrix}\xrightarrow[③-①]{②-①}\begin{bmatrix}a&b&b\\b-a&a-b&0\\b-a&0&a-b\end{bmatrix}\xrightarrow{[1]+[2]+[3]}\begin{bmatrix}a+2b&b&b\\0&a-b&0\\0&0&a-b\end{bmatrix}.$$

$r(A)=2\Leftrightarrow a\neq b$ 且 $a=-2b$，因 $b\neq0$，故 $a=-2b(a\neq0)$.

2.2.11 (A) 解 $AB=O, r(A)+r(B)\leqslant3$，$B$ 是非零矩阵，故 $r(B)\geqslant1$.

因

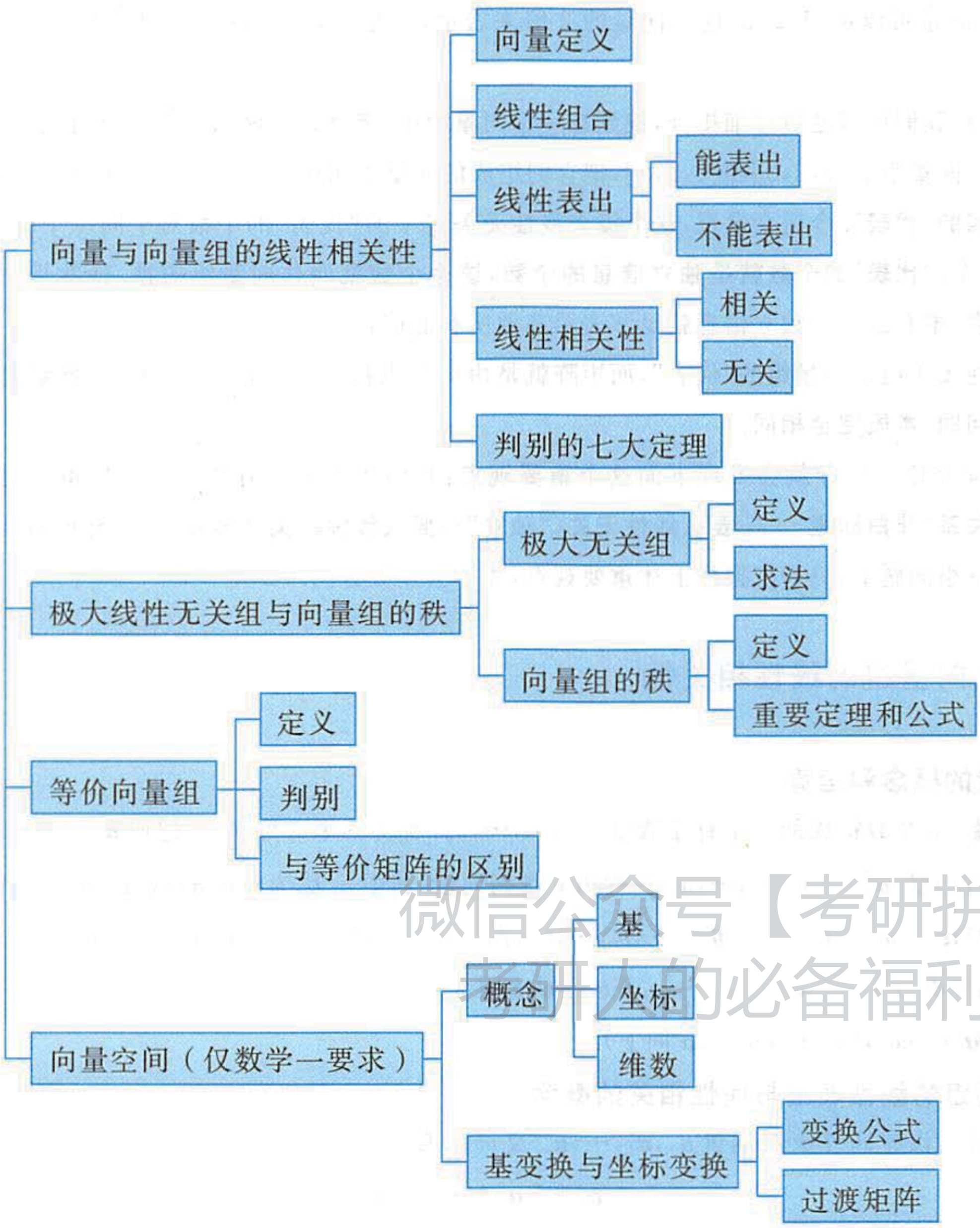
$$A=\begin{bmatrix}1&2&t\\3&t&18\\2&4&2t\\1&8-t&4t-18\end{bmatrix}\xrightarrow[\substack{③-2①\\④-①}]{②-3①}\begin{bmatrix}1&2&t\\0&t-6&18-3t\\0&0&0\\0&6-t&3t-18\end{bmatrix}\xrightarrow{④+②}\begin{bmatrix}1&2&t\\0&t-6&18-3t\\0&0&0\\0&0&0\end{bmatrix}.$$

当 $t\neq6$ 时，$r(A)=2\Rightarrow r(B)=1$，故(A)成立，(C)不成立；

当 $t=6$ 时，$r(A)=1\Rightarrow r(B)=1$ 或 $r(B)=2$，故(B)，(D)不成立.

第3讲 向量组

基础知识结构



基础内容精讲

一、线性代数中的一号人物——向量

有了前面行列式和矩阵两个话题的铺垫,可以让**线性代数的主人翁——向量**出场了.前面指出,矩阵中的若干个行(列)都是向量,它们之间存在着某种联系,这种联系说到底就是**线性无关的向量个数(独立信息的个数)的问题**,也就是这若干个向量组成的向量组中,有几个就足够代表这个向量组,其他的向量都可以由这几个向量线性表示出来.比如,向量组[1,2,3],[6,7,9]与[2,4,6],有三个成员,可是[2,4,6]这个向量可以由[1,2,3]这个向量的2倍来表示,于是[2,4,6]这个向量就是"多余"的,不是独立的信息.

进一步地,我们可以通过仔细排查,找到在一个向量组中,能够代表该向量组中所有成员的一组向量,比如上面这个向量组的[1,2,3]和[6,7,9],把它们组成的向量组叫作原向量组的**极大线性无关组,这个组就是原向量组的"代表"**.今后会发现,这个极大线性无关组中的"代表"的个数对于同一个向量组来说,是唯一的,事实上,**"代表"的个数就是独立信息的个数,这个个数就叫作向量组的秩**.秩就是独立信息的个数,就是"代表"中有这几个独立信息就足够表示其他所有的信息了.

极大线性无关组是向量组的"代表",而矩阵就是由向量组拼成的,所以**矩阵的秩、向量组的秩都反映了"代表"的问题,本质完全相同**.

读者研读完这一部分就会发现下面这个**重要观点**:我们致力于搞清楚的就是**向量与向量之间的关系——这种关系"非白即黑"——要么线性无关("独立"),要么线性相关("多余")**.在充斥着诸多抽象理论和方法的向量组问题上,只有把握住上述重要观点,才不会迷失方向.

二、向量及向量组的线性相关性

1. 向量的概念和运算

n维向量 n个数构成的一个有序数组$[a_1,a_2,\cdots,a_n]$称为一个n维向量,记成$\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n]$,并称$\boldsymbol{\alpha}$为$n$维行向量,$\boldsymbol{\alpha}^{\mathrm{T}}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$称为$n$维列向量,其中$a_i$称为向量$\boldsymbol{\alpha}$(或$\boldsymbol{\alpha}^{\mathrm{T}}$)的第$i$个分量.

相等 若$\boldsymbol{\alpha}=[a_1,a_2,\cdots,a_n]$,$\boldsymbol{\beta}=[b_1,b_2,\cdots,b_n]$都是$n$维向量,则当且仅当$a_i=b_i(i=1,2,\cdots,n)$,$\boldsymbol{\alpha}=\boldsymbol{\beta}$.

加法 $\boldsymbol{\alpha}+\boldsymbol{\beta}=[a_1+b_1,a_2+b_2,\cdots,a_n+b_n]$($\boldsymbol{\alpha},\boldsymbol{\beta}$同上).

数乘 $k\boldsymbol{\alpha}=[ka_1,ka_2,\cdots,ka_n]$($\boldsymbol{\alpha}$同上).

2. 向量组的线性表出与线性相关的概念

线性组合 设有m个n维向量$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$及$m$个数$k_1,k_2,\cdots,k_m$,则向量

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$$

称为向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的线性组合.

线性表出 若向量$\boldsymbol{\beta}$能表示成向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的线性组合,即存在$m$个数$k_1,k_2,\cdots,k_m$,使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m,$$

则称向量$\boldsymbol{\beta}$能被向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表出.

线性相关 对 m 个 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，若存在一组不全为零的数 $k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0},$$

则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关.

含有零向量或有成比例的向量的向量组必线性相关.

线性无关 若不存在不全为零的数 $k_1,k_2,\cdots,k_m$，使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$ 成立，就称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关，亦即只有当 $k_1=k_2=\cdots=k_m=0$ 时，才有 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$ 成立，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

单个非零向量，两个不成比例的向量均线性无关.

向量组或线性相关或线性无关，二者必居其一且仅居其一.

3. 判别线性相关性的七大定理

定理 1 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n(n\geqslant 2)$ 线性相关的充要条件是向量组中至少有一个向量可由其余的 $n-1$ 个向量线性表出.

【注】 **证明** 先证必要性. 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n(n\geqslant 2)$ 线性相关，则存在 n 个不全为零的数 $k_1,k_2,\cdots,k_n$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0}.$$

不妨设 $k_1\neq 0$，于是由向量的线性运算规则得

$$\boldsymbol{\alpha}_1=-\frac{k_2}{k_1}\boldsymbol{\alpha}_2+\frac{k_3}{k_1}\boldsymbol{\alpha}_3-\cdots-\frac{k_n}{k_1}\boldsymbol{\alpha}_n.$$

再证充分性. 不妨设 $\boldsymbol{\alpha}_1$ 可用 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_n$ 线性表示，即

$$\boldsymbol{\alpha}_1=l_2\boldsymbol{\alpha}_2+l_3\boldsymbol{\alpha}_3+\cdots+l_n\boldsymbol{\alpha}_n,$$

于是有

$$1\boldsymbol{\alpha}_1-l_2\boldsymbol{\alpha}_2-l_3\boldsymbol{\alpha}_3-\cdots-l_n\boldsymbol{\alpha}_n=\mathbf{0},$$

显然 $1,-l_2,-l_3,\cdots,-l_n$ 不全为零，故向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关.

其逆否命题：向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n(n\geqslant 2)$ 线性无关的充要条件是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 中任一向量都不能由其余的 $n-1$ 个向量线性表出.

定理 2 若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，而 $\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关，则 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，且表示法唯一.

【注】 **证明** 因为 $\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关，所以存在不全为零的数 $k,k_1,k_2,\cdots,k_n$，使得

$$k\boldsymbol{\beta}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=\mathbf{0},$$

其中 $k\neq 0$(如果 $k=0$，则由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，得 $k_1,k_2,\cdots,k_n$ 必须全为零，这与 $k,k_1,k_2,\cdots,k_n$ 不全为零矛盾)，于是 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示为

$$\boldsymbol{\beta}=-\frac{k_1}{k}\boldsymbol{\alpha}_1-\frac{k_2}{k}\boldsymbol{\alpha}_2-\cdots-\frac{k_n}{k}\boldsymbol{\alpha}_n.$$

再证表示法唯一，设有两种表示法，

$$\begin{aligned}\boldsymbol{\beta}&=l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_n\boldsymbol{\alpha}_n\\&=h_1\boldsymbol{\alpha}_1+h_2\boldsymbol{\alpha}_2+\cdots+h_n\boldsymbol{\alpha}_n,\end{aligned}$$

于是

$$(l_1-h_1)\boldsymbol{\alpha}_1+(l_2-h_2)\boldsymbol{\alpha}_2+\cdots+(l_n-h_n)\boldsymbol{\alpha}_n=\mathbf{0},$$

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，所以必有

$$l_i-h_i=0, \text{即 } l_i=h_i, i=1,2,\cdots,n,$$

故 $\boldsymbol{\beta}$ 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示的表示法唯一.

定理 3　如果向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表示，且 $t>s$，则 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性相关（以少表多，多的相关）.

【注】　以 $t=3,s=2$ 来证明，便于读者理解.

证明　设 $\begin{cases}\boldsymbol{\beta}_1=k_{11}\boldsymbol{\alpha}_1+k_{21}\boldsymbol{\alpha}_2,\\ \boldsymbol{\beta}_2=k_{12}\boldsymbol{\alpha}_1+k_{22}\boldsymbol{\alpha}_2,\\ \boldsymbol{\beta}_3=k_{13}\boldsymbol{\alpha}_1+k_{23}\boldsymbol{\alpha}_2,\end{cases}$ 欲证 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关，只需证存在不全为零的数 l_1,l_2,l_3，使得

$$l_1\boldsymbol{\beta}_1+l_2\boldsymbol{\beta}_2+l_3\boldsymbol{\beta}_3=\mathbf{0},$$

即

$$l_1(k_{11}\boldsymbol{\alpha}_1+k_{21}\boldsymbol{\alpha}_2)+l_2(k_{12}\boldsymbol{\alpha}_1+k_{22}\boldsymbol{\alpha}_2)+l_3(k_{13}\boldsymbol{\alpha}_1+k_{23}\boldsymbol{\alpha}_2)=\mathbf{0},$$

整理得

$$(k_{11}l_1+k_{12}l_2+k_{13}l_3)\boldsymbol{\alpha}_1+(k_{21}l_1+k_{22}l_2+k_{23}l_3)\boldsymbol{\alpha}_2=\mathbf{0}. \qquad (*)$$

当 $\begin{cases}k_{11}l_1+k_{12}l_2+k_{13}l_3=0,\\ k_{21}l_1+k_{22}l_2+k_{23}l_3=0\end{cases}$ 时，显然（*）式成立，而这个方程组是 3 个未知数 l_1,l_2,l_3，2 个方程的情形，未知数个数大于方程个数，必有非零解 l_1,l_2,l_3，故 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关.

其等价命题：如果向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表示，且 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 线性无关，则 $t\leqslant s$.

定理 4　设 m 个 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$，其中

$$\boldsymbol{\alpha}_1=[a_{11},a_{21},\cdots,a_{n1}]^{\mathrm{T}},$$

$$\boldsymbol{\alpha}_2=[a_{12},a_{22},\cdots,a_{n2}]^{\mathrm{T}},$$

$$\cdots\cdots$$

$$\boldsymbol{\alpha}_m=[a_{1m},a_{2m},\cdots,a_{nm}]^{\mathrm{T}}.$$

则向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关的充分必要条件是齐次线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\mathbf{0} \qquad (*)$$

有非零解，其中

$$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m]=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\a_{21}&a_{22}&\cdots&a_{2m}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_m\end{bmatrix}.$$

【注】 证明 设

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m=\mathbf{0},\tag{**}$$

即

$$x_1\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{bmatrix}+x_2\begin{bmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{bmatrix}+\cdots+x_m\begin{bmatrix}a_{1m}\\a_{2m}\\\vdots\\a_{nm}\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}.\tag{***}$$

将(＊＊＊)式左端写成矩阵形式，即得线性方程组(＊)．因此，如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关，就必有不全为零的数 $x_1,x_2,\cdots,x_m$ 使(＊＊)式成立，即齐次线性方程组(＊)有非零解；反之，如果线性方程组(＊)有非零解，也就是有不全为零的数 $x_1,x_2,\cdots,x_m$ 使(＊＊)式成立，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关．定理得证．

其等价命题：m 个 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关的充分必要条件是齐次线性方程组(＊)只有零解．

【注1】 如果 $n<m$，即方程个数小于未知数个数，线性方程组(＊)求解时必有自由未知量，即必有非零解．因此，任何 $n+1$ 个 n 维向量都是线性相关的．所以在 n 维空间中，任何一个线性无关的向量组最多只能含 n 个向量．

【注2】 n 个 n 维列向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关$\Leftrightarrow|\boldsymbol{A}|=|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n|=0\Leftrightarrow\boldsymbol{Ax}=\mathbf{0}$ 有非零解．反之则有，

$$\text{线性无关}\Leftrightarrow|\boldsymbol{A}|\neq0\Leftrightarrow\boldsymbol{Ax}=\mathbf{0}\text{ 仅有零解．}$$

定理5 仿定理4的研究方法，便有

向量 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出

$$\Leftrightarrow\text{非齐次线性方程组}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s]\begin{bmatrix}x_1\\x_2\\\vdots\\x_s\end{bmatrix}=\boldsymbol{\alpha}_1x_1+\boldsymbol{\alpha}_2x_2+\cdots+\boldsymbol{\alpha}_sx_s=\boldsymbol{\beta}\text{ 有解}$$

$$\Leftrightarrow r([\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s])=r([\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}]).$$

反之则有，不能线性表出$\Leftrightarrow[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s]\begin{bmatrix}x_1\\x_2\\\vdots\\x_s\end{bmatrix}=\boldsymbol{\alpha}_1x_1+\boldsymbol{\alpha}_2x_2+\cdots+\boldsymbol{\alpha}_sx_s=\boldsymbol{\beta}$ 无解$\Leftrightarrow r([\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s])\neq r([\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}])$．

定理6 如果向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 中有一部分向量线性相关，则整个向量组也线性相关．

证明 不妨设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_j(j<m)$ 线性相关，于是有不全为零的数 $k_1,k_2,\cdots,k_j$，使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_j\boldsymbol{\alpha}_j=\mathbf{0}.$$

从而有不全零的数为 $k_1,k_2,\cdots,k_j,0,\cdots,0$，使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_j\boldsymbol{\alpha}_j+0\boldsymbol{\alpha}_{j+1}+\cdots+0\boldsymbol{\alpha}_m=\mathbf{0},$$

故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 也线性相关．

其逆否命题：如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关，则其任一部分向量组也线性无关．

总之，向量组部分线性相关，则整体也线性相关；整体线性无关，则任一部分都线性无关．

定理 7 如果一组 n 维向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关，那么把这些向量各任意添加 m 个分量所得到的新向量（$n+m$ 维）组 $\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*,\cdots,\boldsymbol{\alpha}_s^*$ 也是线性无关的；如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，那么它们各去掉相同的若干个分量所得到的新向量组也是线性相关的.

【注】 事实上，对于

$$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s],\quad \boldsymbol{B}=[\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*,\cdots,\boldsymbol{\alpha}_s^*],$$

其中 $\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*,\cdots,\boldsymbol{\alpha}_s^*$ 是分别在 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 后面任加 m 个分量. 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关，即齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解，则 $\boldsymbol{Bx}=\boldsymbol{0}$ 显然也只有零解（即 $\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*,\cdots,\boldsymbol{\alpha}_s^*$ 也线性无关）；反之，如果 $\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*,\cdots,\boldsymbol{\alpha}_s^*$ 线性相关，即 $\boldsymbol{Bx}=\boldsymbol{0}$ 有非零解，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 也有非零解（即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 也线性相关）.

三、极大线性无关组、等价向量组、向量组的秩

1. 极大线性无关组

在向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中，若存在部分组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 满足：

①$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性无关；

②向量组中任一向量 $\boldsymbol{\alpha}_i(i=1,2,\cdots,s)$ 均可由 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 线性表出.

则称向量组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是原向量组的极大线性无关组.

向量组的极大线性无关组一般不唯一，只由一个零向量组成的向量组不存在极大线性无关组，一个线性无关向量组的极大线性无关组就是该向量组本身.

2. 等价向量组

设两个向量组：（Ⅰ）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$，（Ⅱ）$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$. 若（Ⅰ）中每个 $\boldsymbol{\alpha}_i(i=1,2,\cdots,s)$ 均可由（Ⅱ）线性表出，则称向量组（Ⅰ）可由向量组（Ⅱ）线性表出；若向量组（Ⅰ），（Ⅱ）可以相互线性表出，则称向量组（Ⅰ）与向量组（Ⅱ）是等价向量组，记作（Ⅰ）$\cong$（Ⅱ）.

等价向量组满足：

①（Ⅰ）$\cong$（Ⅰ）（反身性）；

②若（Ⅰ）$\cong$（Ⅱ），则（Ⅱ）$\cong$（Ⅰ）（对称性）；

③若（Ⅰ）$\cong$（Ⅱ），（Ⅱ）$\cong$（Ⅲ），则（Ⅰ）$\cong$（Ⅲ）（传递性）.

向量组和它的极大线性无关组是等价向量组.

【注】 读者应注意等价矩阵和等价向量组概念的区别.

3. 向量组的秩

向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的极大线性无关组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 中所含向量的个数 r 称为向量组的秩，记作 $\mathrm{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r$ 或 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r$.

等价向量组等秩，反之未必成立.

4. 有关向量组的秩的重要定理和公式

（1）三秩相等.

$r(\boldsymbol{A})$（矩阵的秩）$=\boldsymbol{A}$ 的行秩（$\boldsymbol{A}$ 的行向量组的秩）$=\boldsymbol{A}$ 的列秩（$\boldsymbol{A}$ 的列向量组的秩）.

（2）若 $\boldsymbol{A}\xrightarrow{\text{初等行变换}}\boldsymbol{B}$，则

①$\boldsymbol{A}$ 的行向量组和 $\boldsymbol{B}$ 的行向量组是等价向量组；

②$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的任何相应的部分列向量组具有相同的线性相关性.

(3)设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 及 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$. 若 $\boldsymbol{\beta}_i(i=1,2,\cdots,t)$ 均可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，则

$$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\leqslant r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s).$$

四、向量空间(仅数学一要求)

1. 基本概念

若 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 维向量空间 $\mathbf{R}^n$ 中的线性无关的有序向量组，则任一向量 $\boldsymbol{\alpha}\in\mathbf{R}^n$ 均可由 $\xi_1,\xi_2,\cdots,\xi_n$ 线性表出，记表出式为

$$\boldsymbol{\alpha}=a_1\xi_1+a_2\xi_2+\cdots+a_n\xi_n,$$

称有序向量组 $\xi_1,\xi_2,\cdots,\xi_n$ 是 $\mathbf{R}^n$ 的一个基，基向量的个数 n 称为向量空间的维数，而$[a_1,a_2,\cdots,a_n]$($[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$)称为向量 $\boldsymbol{\alpha}$ 在基 $\xi_1,\xi_2,\cdots,\xi_n$ 下的坐标，或称为 $\boldsymbol{\alpha}$ 的坐标行(列)向量.

2. 基变换、坐标变换

定理 8 若 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 和 $\xi_1,\xi_2,\cdots,\xi_n$ 是 $\mathbf{R}^n$ 中的两个基，且有关系

$$[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n]=[\xi_1,\xi_2,\cdots,\xi_n]\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}=[\xi_1,\xi_2,\cdots,\xi_n]\boldsymbol{C}, \tag{$*$}$$

则($*$)式称为由基 $\xi_1,\xi_2,\cdots,\xi_n$ 到基 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 的基变换公式，矩阵 $\boldsymbol{C}$ 称为由基 $\xi_1,\xi_2,\cdots,\xi_n$ 到基 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 的过渡矩阵，$\boldsymbol{C}$ 的第 i 列即是 $\boldsymbol{\eta}_i$ 在基 $\xi_1,\xi_2,\cdots,\xi_n$ 下的坐标列向量，且过渡矩阵 $\boldsymbol{C}$ 是可逆矩阵.

定理 9 设 $\boldsymbol{\alpha}$ 在基 $\xi_1,\xi_2,\cdots,\xi_n$ 和基 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 下的坐标分别是 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，$\boldsymbol{y}=[y_1,y_2,\cdots,y_n]^{\mathrm{T}}$，即

$$\boldsymbol{\alpha}=[\xi_1,\xi_2,\cdots,\xi_n]\boldsymbol{x}=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n]\boldsymbol{y}.$$

又基 $\xi_1,\xi_2,\cdots,\xi_n$ 到基 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$ 的过渡矩阵为 $\boldsymbol{C}$，即

$$[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n]=[\xi_1,\xi_2,\cdots,\xi_n]\boldsymbol{C},$$

则

$$\boldsymbol{\alpha}=[\xi_1,\xi_2,\cdots,\xi_n]\boldsymbol{x}=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n]\boldsymbol{y}=[\xi_1,\xi_2,\cdots,\xi_n]\boldsymbol{C}\boldsymbol{y},$$

得

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}\text{ 或 }\boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}. \tag{$**$}$$

($**$)式称为坐标变换公式.

基础例题精解

一、向量组的线性表出与线性相关

其理论依据是本讲“基础内容精讲”中的“二”的“2. 向量组的线性表出与线性相关的概念”与“3. 判别线性相关性的七大定理”.

例 2.3.1 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 s 个 n 维向量，下列论断正确的是（　　）.

(A)$\boldsymbol{\alpha}_s$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出，则向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关

(B)已知存在不全为零的数 $k_1,k_2,\cdots,k_{s-1}$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{s-1}\boldsymbol{\alpha}_{s-1}+0\boldsymbol{\alpha}_s=\mathbf{0},$$

则 $\boldsymbol{\alpha}_s$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出

(C)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，则任一向量均可由其余向量线性表出

(D)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，$\boldsymbol{\alpha}_s$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性相关

解　应选(D).

(A)不正确. 因 $\boldsymbol{\alpha}_s$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出，并不能排除别的向量能由其余向量线性表出，如 $\boldsymbol{\alpha}_1=[1,0]$，$\boldsymbol{\alpha}_2=[2,0]$，$\boldsymbol{\alpha}_3=[0,2]$，虽然 $\boldsymbol{\alpha}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性表出，但 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关.

(B)不正确. 因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关的线性组合系数不唯一（特别是 $\boldsymbol{\alpha}_s=\mathbf{0}$ 时，显然 $\boldsymbol{\alpha}_s$ 仍可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出）. 当另一组不全为零的数 $l_1,l_2,\cdots,l_s(l_s\neq0)$，使等式

$$l_1\boldsymbol{\alpha}_1+l_2\boldsymbol{\alpha}_2+\cdots+l_s\boldsymbol{\alpha}_s=\mathbf{0}$$

成立时，$\boldsymbol{\alpha}_s$ 仍可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出. 例如：$\boldsymbol{\alpha}_1=[1,1]$，$\boldsymbol{\alpha}_2=[2,2]$，$\boldsymbol{\alpha}_3=[3,3]$，有 $2\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3=\mathbf{0}$，但又有 $\boldsymbol{\alpha}_3=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$.

(C)不正确. 例如 $\boldsymbol{\alpha}_1=[0,0]$，$\boldsymbol{\alpha}_2=[1,0]$线性相关，$\boldsymbol{\alpha}_1$ 可由 $\boldsymbol{\alpha}_2$ 线性表出，但 $\boldsymbol{\alpha}_2$ 不能由 $\boldsymbol{\alpha}_1$ 线性表出.

(D)正确. 因若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性无关，而已知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，故由定理 2 知，$\boldsymbol{\alpha}_s$ 必可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出（且表出法唯一），这和已知矛盾，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性相关.

例 2.3.2 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_{s+1}$ 线性无关，问：

(1)$\boldsymbol{\alpha}_1$ 能否由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s$ 线性表出，证明你的结论；

(2)$\boldsymbol{\alpha}_{s+1}$ 能否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，证明你的结论.

解　(1)$\boldsymbol{\alpha}_1$ 能由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s$ 线性表出.

方法一　由已知 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_{s+1}$ 线性无关，得 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s$ 线性无关，而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，故 $\boldsymbol{\alpha}_1$ 可由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s$ 线性表出（且表出法唯一）.

方法二　因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关，故存在不全为零的数 $k_1,k_2,\cdots,k_s$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0},$$

其中，$k_1\neq0$. 若 $k_1=0$，则 $k_2,k_3,\cdots,k_s$ 不全为零，则 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s$ 线性相关，从而 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_{s+1}$ 线性相关，这和已知矛盾，故 $k_1\neq0$，故有

$$\boldsymbol{\alpha}_1=-\frac{1}{k_1}(k_2\boldsymbol{\alpha}_2+k_3\boldsymbol{\alpha}_3+\cdots+k_s\boldsymbol{\alpha}_s)\overset{\text{记}}{=\!=}l_2\boldsymbol{\alpha}_2+l_3\boldsymbol{\alpha}_3+\cdots+l_s\boldsymbol{\alpha}_s.\quad(*)$$

(2)$\boldsymbol{\alpha}_{s+1}$不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出. 用反证法. 设 $\boldsymbol{\alpha}_{s+1}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，

$$\boldsymbol{\alpha}_{s+1}=\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\alpha}_2+\cdots+\lambda_s\boldsymbol{\alpha}_s,$$

由(1)，将(*)式 $\boldsymbol{\alpha}_1=l_2\boldsymbol{\alpha}_2+l_3\boldsymbol{\alpha}_3+\cdots+l_s\boldsymbol{\alpha}_s$ 代入上式，得

$$\boldsymbol{\alpha}_{s+1}=(\lambda_1l_2+\lambda_2)\boldsymbol{\alpha}_2+(\lambda_1l_3+\lambda_3)\boldsymbol{\alpha}_3+\cdots+(\lambda_1l_s+\lambda_s)\boldsymbol{\alpha}_s,$$

则 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_{s+1}$ 线性相关，这和已知矛盾，故 $\boldsymbol{\alpha}_{s+1}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出.

【注】 考研真题“设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关，向量组 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关. 问：(1)$\boldsymbol{\alpha}_1$ 能否由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出？证明你的结论；(2)$\boldsymbol{\alpha}_4$ 能否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出？证明你的结论.”是由本题改编而得，由本书作者之一在当年命题时编制，而考生做得并不理想.

例 2.3.3 已知 4 维列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，则下列向量组中线性无关的是（　　）.

(A)$\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1$

(B)$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_1$

(C)$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1$

(D)$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1$

解 应选(D).

方法一 排除法.

(A)$(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)+(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3)+(\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4)+(\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1)=\mathbf{0}$;

(B)$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)-(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+(\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4)-(\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_1)=\mathbf{0}$;

(C)$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)-(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+(\boldsymbol{\alpha}_3-\boldsymbol{\alpha}_4)+(\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1)=\mathbf{0}$.

故(A),(B),(C)均是线性相关的向量组，由排除法，应选(D).

方法二

$$[\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4]\begin{bmatrix}1&0&0&-1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{bmatrix},$$

两边取行列式，因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，则 $|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4|\neq 0$. 又

$$\begin{vmatrix}1&0&0&-1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{vmatrix}=2\neq 0,$$

故 $|\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1|\neq 0$，向量组 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_4-\boldsymbol{\alpha}_1$ 线性无关，应选(D).

例 2.3.4 设 $a_1,a_2,\cdots,a_s$ 是 s 个互不相同的数，试讨论 s 个 n 维列向量 $\boldsymbol{\alpha}_i=[1,a_i,a_i^2,\cdots,a_i^{n-1}]^{\mathrm{T}}$ $(i=1,2,\cdots,s)$ 的线性相关性.

解 若 $s>n$，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关（向量个数超过向量维数时，向量组必线性相关）.

若 $s=n$，则由范德蒙德行列式，

$$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n|=\begin{vmatrix}1&1&\cdots&1\\a_1&a_2&\cdots&a_n\\\vdots&\vdots&&\vdots\\a_1^{n-1}&a_2^{n-1}&\cdots&a_n^{n-1}\end{vmatrix}$$

$$=\prod_{n\geqslant i>j\geqslant 1}(a_i-a_j)\neq 0$$

知，方程组 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\mathbf{0}$ 有唯一零解，故向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关.

若 $s<n$，则由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关（整体）知，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s(s<n)$ 线性无关（部分）.

综上，当 $s>n$ 时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性相关；当 $s\leqslant n$ 时，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关.

例 2.3.5 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 都是 n 维向量$(n\geqslant 3)$，且 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2,\boldsymbol{\beta}_3=3\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2$，证明：向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关.

分析 只要找到不全为零的数 k_1,k_2,k_3，使

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+k_3\boldsymbol{\beta}_3=\mathbf{0}$$

即可.

证明　设有数 k_1,k_2,k_3，使得

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+k_3\boldsymbol{\beta}_3=\mathbf{0},$$

代入题设条件，得

$$k_1(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+k_2(\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2)+k_3(3\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2)=\mathbf{0},$$

整理得

$$(k_1+k_2+3k_3)\boldsymbol{\alpha}_1+(k_1-2k_2+2k_3)\boldsymbol{\alpha}_2=\mathbf{0}.$$

要使上式成立，不论 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是否线性相关，只需

$$\begin{cases}k_1+k_2+3k_3=0,\\k_1-2k_2+2k_3=0\end{cases} \qquad ①$$

即可. 由方程组①求得一个解为 $k_1=8,k_2=1,k_3=-3$，故

$$8\boldsymbol{\beta}_1+\boldsymbol{\beta}_2-3\boldsymbol{\beta}_3=\mathbf{0},$$

从而得证：$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关.

【注】　本题中 3 个向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 由 2 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性表出，由定理 3 知“以少表多，多的相关”，故 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性相关，本题的证明思路即是定理 3 的证明思路.

例 2.3.6　设 $\boldsymbol{A}$ 是 n 阶矩阵，若存在正整数 k，使线性方程组 $\boldsymbol{A}^k\boldsymbol{x}=\mathbf{0}$ 有解向量 $\boldsymbol{\alpha}$，且 $\boldsymbol{A}^{k-1}\boldsymbol{\alpha}\neq\mathbf{0}$，证明：向量组 $\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha},\cdots,\boldsymbol{A}^{k-1}\boldsymbol{\alpha}$ 线性无关.

证明　设有一组常数 $\lambda_1,\lambda_2,\cdots,\lambda_k$，使得

$$\lambda_1\boldsymbol{\alpha}+\lambda_2\boldsymbol{A}\boldsymbol{\alpha}+\cdots+\lambda_k\boldsymbol{A}^{k-1}\boldsymbol{\alpha}=\mathbf{0}, \qquad ①$$

式①两端左乘 $\boldsymbol{A}^{k-1}$，得

$$\lambda_1\boldsymbol{A}^{k-1}\boldsymbol{\alpha}+\lambda_2\boldsymbol{A}^k\boldsymbol{\alpha}+\cdots+\lambda_k\boldsymbol{A}^{2k-2}\boldsymbol{\alpha}=\mathbf{0}.$$

由题设条件 $\boldsymbol{A}^k\boldsymbol{\alpha}=\mathbf{0}$，知 $\boldsymbol{A}^{k+1}\boldsymbol{\alpha}=\boldsymbol{A}^{k+2}\boldsymbol{\alpha}=\cdots=\boldsymbol{A}^{2k-2}\boldsymbol{\alpha}=\mathbf{0}$，从而得 $\lambda_1\boldsymbol{A}^{k-1}\boldsymbol{\alpha}=\mathbf{0}$.

由于 $\boldsymbol{A}^{k-1}\boldsymbol{\alpha}\neq\mathbf{0}$，故 $\lambda_1=0$，代入式①得

$$\lambda_2\boldsymbol{A}\boldsymbol{\alpha}+\lambda_3\boldsymbol{A}^2\boldsymbol{\alpha}+\cdots+\lambda_k\boldsymbol{A}^{k-1}\boldsymbol{\alpha}=\mathbf{0}. \qquad ②$$

将式②两端左乘 $\boldsymbol{A}^{k-2}$，同上可证 $\lambda_2=0$. 同理可证 $\lambda_3=\lambda_4=\cdots=\lambda_k=0$，从而得证 $\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha},\boldsymbol{A}^2\boldsymbol{\alpha},\cdots,\boldsymbol{A}^{k-1}\boldsymbol{\alpha}$ 线性无关.

二、极大线性无关组及向量组秩的求法

在掌握了向量组线性相关与线性无关的基本概念后，需要进一步确定某向量组的极大线性无关组，并确定该向量组的秩.

例 2.3.7　设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的秩为 $r(r<s)$，则下列说法错误的是（　　）.

(A) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中至少有一个由 r 个向量组成的部分组线性无关

(B) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中任何 r 个线性无关向量组成的部分组与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是等价向量组

(C) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中任何 r 个向量的部分组都线性无关

(D)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中任何 $r+1$ 个向量的部分组都线性相关

解 应选(C).

$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r$，由定义，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中最大的线性无关的部分向量组中的向量个数是 r，故至少有一个包含 r 个向量的部分组线性无关，(A)成立. 向量组的极大无关组不一定是唯一的，任何包含 r 个向量的线性无关部分组均是极大线性无关组，均与原向量组等价，故(B)成立. 因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中最大的线性无关部分组的向量个数为 r，所以任何 $r+1$ 个向量的部分组都线性相关，因此(D)成立. 故由排除法知应选(C).

事实上，$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r$，并非 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 中任何 r 个向量都线性无关，如 $\boldsymbol{\alpha}_1=[1,0]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[2,0]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=[0,1]^{\mathrm{T}}$，$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=2$，但 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是线性相关的. 故选项(C)不成立，应选(C).

例 2.3.8 设向量组

$\boldsymbol{\alpha}_1=[1,-1,2,4]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[0,3,1,2]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=[3,0,7,14]^{\mathrm{T}}$，$\boldsymbol{\alpha}_4=[1,-2,2,0]^{\mathrm{T}}$，$\boldsymbol{\alpha}_5=[2,1,5,10]^{\mathrm{T}}$，

则该向量组的极大线性无关组是（　　）.

(A)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$　　(B)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$　　(C)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_5$　　(D)$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5$

解 应选(B).

将向量组合并成矩阵 $\boldsymbol{A}$，并作初等行变换，化成阶梯形矩阵.

$$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5]=\begin{bmatrix}1&0&3&1&2\\-1&3&0&-2&1\\2&1&7&2&5\\4&2&14&0&10\end{bmatrix}$$

$$\rightarrow\begin{bmatrix}1&0&3&1&2\\0&3&3&-1&3\\0&1&1&0&1\\0&2&2&-4&2\end{bmatrix}\rightarrow\begin{bmatrix}1&0&3&1&2\\0&1&1&0&1\\0&0&0&-1&0\\0&0&0&0&0\end{bmatrix}$$

$$=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4,\boldsymbol{\beta}_5]=\boldsymbol{B}.$$

$\boldsymbol{B}$ 中的极大线性无关组为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_4$ 或 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 或 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_4,\boldsymbol{\beta}_5$，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5$ 对应的极大线性无关组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 或 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 或 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5$. 对照选项可知，应选(B).

【注】 列向量组经初等行变换，化成新的列向量组，即

$$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s]\xrightarrow{\text{初等行变换}}[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s],$$

因 $[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s]\boldsymbol{x}=\boldsymbol{0}$ 和 $[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s]\boldsymbol{x}=\boldsymbol{0}$ 是同解方程组，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 有相同的线性相关性. 同样，任何对应的部分向量组

$$[\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}]\xrightarrow{\text{初等行变换}}[\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}],$$

因 $[\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}]\boldsymbol{x}=\boldsymbol{0}$ 和 $[\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}]\boldsymbol{x}=\boldsymbol{0}$ 同解，故 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 和 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 有相同的线性相关性.

例 2.3.9 求向量组

$$\boldsymbol{\beta}_1=[1,2,1,3]^{\mathrm{T}},\quad \boldsymbol{\beta}_2=[1,1,-1,1]^{\mathrm{T}},\quad \boldsymbol{\beta}_3=[1,3,3,5]^{\mathrm{T}},$$
$$\boldsymbol{\beta}_4=[4,5,-3,6]^{\mathrm{T}},\quad \boldsymbol{\beta}_5=[-3,-5,-2,-7]^{\mathrm{T}}$$

的秩、极大线性无关组，并将其余的向量用极大线性无关组线性表出.

解 用列向量组作初等行变换，因初等行变换将方程组变成同解方程组，故变换前后的任何相应的部分(或全部)列向量组成的方程组仍同解，即它们具有相同的线性相关性.

$$\begin{bmatrix}1&1&1&4&-3\\2&1&3&5&-5\\1&-1&3&-3&-2\\3&1&5&6&-7\end{bmatrix}\xrightarrow[\substack{③-① \\ ④-3①}]{②-2①}\begin{bmatrix}1&1&1&4&-3\\0&-1&1&-3&1\\0&-2&2&-7&1\\0&-2&2&-6&2\end{bmatrix}$$

$$\xrightarrow[④-2②]{③-2②}\begin{bmatrix}1&1&1&4&-3\\0&-1&1&-3&1\\0&0&0&-1&-1\\0&0&0&0&0\end{bmatrix}$$

$$\xrightarrow[\substack{(-1)② \\ (-1)③}]{①+②}\begin{bmatrix}1&0&2&1&-2\\0&1&-1&3&-1\\0&0&0&1&1\\0&0&0&0&0\end{bmatrix}=[\boldsymbol{\beta}_1',\boldsymbol{\beta}_2',\boldsymbol{\beta}_3',\boldsymbol{\beta}_4',\boldsymbol{\beta}_5'].$$

由右端阶梯形矩阵知 $r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4,\boldsymbol{\beta}_5)=3$，由 $\boldsymbol{\beta}_1',\boldsymbol{\beta}_2',\boldsymbol{\beta}_4'$ 线性无关(由第 1，2，4 列组成的齐次线性方程组有唯一零解推得)知，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_4$ 线性无关，是向量组的极大线性无关组，$\boldsymbol{\beta}_3,\boldsymbol{\beta}_5$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_4$ 线性表出，由方程

$$\begin{cases}\boldsymbol{\beta}_1'x_1+\boldsymbol{\beta}_2'x_2+\boldsymbol{\beta}_4'x_3=\boldsymbol{\beta}_3',\\ \boldsymbol{\beta}_1'y_1+\boldsymbol{\beta}_2'y_2+\boldsymbol{\beta}_4'y_3=\boldsymbol{\beta}_5',\end{cases}$$

可解得

$$\begin{cases}\boldsymbol{\beta}_3'=2\boldsymbol{\beta}_1'-\boldsymbol{\beta}_2'+0\boldsymbol{\beta}_4',\\ \boldsymbol{\beta}_5'=-3\boldsymbol{\beta}_1'-4\boldsymbol{\beta}_2'+\boldsymbol{\beta}_4',\end{cases}$$

即

$$\begin{cases}\boldsymbol{\beta}_3=2\boldsymbol{\beta}_1-\boldsymbol{\beta}_2,\\ \boldsymbol{\beta}_5=-3\boldsymbol{\beta}_1-4\boldsymbol{\beta}_2+\boldsymbol{\beta}_4.\end{cases}$$

【注】 (1)将向量组处理成行向量并作初等列变换，其效果和将向量组处理成列向量并作初等行变换是一样的.

(2)向量组(或矩阵)的秩是唯一的，其极大线性无关组可以是不唯一的. 本题中除 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_4$ 外，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$；$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_5$；$\boldsymbol{\beta}_1,\boldsymbol{\beta}_3,\boldsymbol{\beta}_5$ 均是向量组的极大线性无关组，其余向量由极大线性无关组表出时，表出法唯一.

(3)求向量组的极大线性无关组时，只能都作初等行变换(或都作初等列变换)，不能混合着既作初等行变换又作初等列变换.

三、有关秩的证明题

例 2.3.10 证明：$r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}$.

分析 对于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 及 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$，若任一 $\boldsymbol{\beta}_i(i=1,2,\cdots,t)$ 均可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，则 $r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)\leqslant r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$. 这是本题证明的关键依据.

证明 不妨设 $A_{m\times n}$，$B_{n\times s}$. 将 B，AB 按行分块为 $B=\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{bmatrix}$，$AB=C=\begin{bmatrix}\gamma_1\\\gamma_2\\\vdots\\\gamma_m\end{bmatrix}$，于是

$$AB=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{bmatrix}=\begin{bmatrix}a_{11}\beta_1+a_{12}\beta_2+\cdots+a_{1n}\beta_n\\a_{21}\beta_1+a_{22}\beta_2+\cdots+a_{2n}\beta_n\\\vdots\\a_{m1}\beta_1+a_{m2}\beta_2+\cdots+a_{mn}\beta_n\end{bmatrix}=\begin{bmatrix}\gamma_1\\\gamma_2\\\vdots\\\gamma_m\end{bmatrix},$$

所以 AB 的行向量 $\gamma_i(i=1,2,\cdots,m)$ 均可由 B 的行向量线性表出，故

$$r(AB)\leqslant r(B).$$

同理可证 $r(AB)\leqslant r(A)$，故有 $r(AB)\leqslant\min\{r(A),r(B)\}$.

例 2.3.11 证明：$r(A+B)\leqslant r([A,B])\leqslant r(A)+r(B)$.

证明 设 $r(A)=p$，$r(B)=q$，将 A，B 按列分块为

$$A=[\alpha_1,\alpha_2,\cdots,\alpha_s],\quad B=[\beta_1,\beta_2,\cdots,\beta_s],$$

于是

$$[A,B]=[\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_s],$$

$$A+B=[\alpha_1+\beta_1,\alpha_2+\beta_2,\cdots,\alpha_s+\beta_s].$$

因 $\alpha_i+\beta_i(i=1,2,\cdots,s)$ 均可由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_s$ 线性表出，故

$$r(A+B)\leqslant r([A,B]).$$

又设 A，B 的列向量组的极大线性无关组分别为 $\alpha_1,\alpha_2,\cdots,\alpha_p$ 和 $\beta_1,\beta_2,\cdots,\beta_q$，将 A 的极大线性无关组 $\alpha_1,\alpha_2,\cdots,\alpha_p$ 扩充成 $[A,B]$ 的极大线性无关组，设为 $\alpha_1,\alpha_2,\cdots,\alpha_p,\beta_1,\beta_2,\cdots,\beta_w$，显然 $w\leqslant q$，故有

$$r([A,B])=p+w\leqslant p+q=r(A)+r(B),$$

故 $r(A+B)\leqslant r([A,B])\leqslant r(A)+r(B)$.

例 2.3.12 设 A 是 n 阶方阵，A^* 是 A 的伴随矩阵. 证明：

$$r(A^*)=\begin{cases}n, & r(A)=n,\\1, & r(A)=n-1,\\0, & r(A)<n-1.\end{cases}$$

分析 证明前，①应先研究已知条件及其等价条件是什么. 例如，已知 $r(A)=n$（应联想到）$\Leftrightarrow|A|\neq 0\Leftrightarrow A$ 可逆 $\Leftrightarrow A$ 的列向量组、行向量组线性无关 $\Leftrightarrow Ax=0$ 只有零解等.

$r(A)=n-1$，应想到 $|A|=0$，但存在 $n-1$ 阶子式不为零，故至少有一个 $|A|$ 中元素的余子式或代数余子式不为零.

$r(A)<n-1$，则 $|A|$ 中全部元素的代数余子式都为零.

②再研究要证什么. 要证 $r(A^*)=n$，可证 $|A^*|\neq 0$，证 A^* 可逆，证 $A^*x=0$ 只有零解，且应想到 A^* 是什么样的矩阵.

③已知条件和要证结论之间有何种联系：$AA^*=A^*A=|A|E$.

证明 当 $r(A)=n$ 时，A 可逆，$|A|\neq 0$，由 $AA^*=|A|E$ 知，A 和 A^* 均是可逆矩阵，故 $r(A^*)=n$.（或两边取行列式，得 $|AA^*|=|A||A^*|=||A|E|=|A|^n$，$|A^*|=|A|^{n-1}\neq 0$，$r(A^*)=n$.）

当 $r(A)=n-1$ 时，由矩阵的秩的定义知，$|A|$ 中存在 $n-1$ 阶子式不等于零，而 A^* 由 $|A|$ 的元素 a_{ij} 的代数余子式 A_{ij} 组成，故 $r(A^*)\geqslant 1$. 又 $r(A)=n-1$，$|A|=0$，$AA^*=|A|E=O$，得 $r(A)+r(A^*)\leqslant n$，而

$r(\boldsymbol{A})=n-1$，故 $r(\boldsymbol{A}^*)\leqslant 1$（或 $\boldsymbol{A}^*$ 的每一列均是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解向量，因 $r(\boldsymbol{A})=n-1$，故 $r(\boldsymbol{A}^*)\leqslant 1$），所以 $r(\boldsymbol{A}^*)=1$.

当 $r(\boldsymbol{A})<n-1$ 时，由 $\boldsymbol{A}$ 的秩的定义知，$|\boldsymbol{A}|$ 的代数余子式全部为零，即 $\boldsymbol{A}^*$ 的全部元素为零，$\boldsymbol{A}^*=\boldsymbol{O}$，故 $r(\boldsymbol{A}^*)=0$.

【注】 (1)上述过程是可逆的，即 $r(\boldsymbol{A})=n\Leftrightarrow r(\boldsymbol{A}^*)=n$，$r(\boldsymbol{A})=n-1\Leftrightarrow r(\boldsymbol{A}^*)=1$，$r(\boldsymbol{A})<n-1\Leftrightarrow r(\boldsymbol{A}^*)=0$. 考题也考过这些.

(2)由本题可知，$\boldsymbol{A}^*$ 的秩只有三种情况：$n,1,0$. 对行(列)向量组而言，$\boldsymbol{A}^*$ 只有三种情况：n 行(列)线性无关；n 行(列)两两成比例；行(列)向量全部是零向量.

四、等价矩阵和等价向量组

(1)向量组等价和矩阵等价是两个不同的概念. 矩阵等价要同型，当然行数、列数都要相等；向量组等价要同维，但向量个数可以不等.

(2)$\boldsymbol{A},\boldsymbol{B}$ 同型时，$\boldsymbol{A}\cong\boldsymbol{B}\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{B})\Leftrightarrow \boldsymbol{PAQ}=\boldsymbol{B}$（$\boldsymbol{P},\boldsymbol{Q}$ 是可逆矩阵）.

(3)$\boldsymbol{\alpha}_i,\boldsymbol{\beta}_j\,(i=1,2,\cdots,s;j=1,2,\cdots,t)$ 同维，则

$$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}\cong\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t\}$$

$\Leftrightarrow\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s\}$ 与 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t\}$ 可以相互表出

$\Leftrightarrow r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)$，且可单方向表出，即只知 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 这两个向量组中的某一个向量组可由另一个向量组线性表出

$\Leftrightarrow r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)=r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t)$（三秩相同）.

例 2.3.13 设 n 维向量组(Ⅰ)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$，(Ⅱ)$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$，记

$$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s],\quad \boldsymbol{B}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t].$$

则下列结论正确的是（　　）.

(A)若 $r(\text{Ⅰ})=r(\text{Ⅱ})$，则 $\boldsymbol{A}\cong\boldsymbol{B}$

(B)若(Ⅰ)可由(Ⅱ)线性表出，则(Ⅰ)≅(Ⅱ)

(C)若 $r(\boldsymbol{A})=r(\boldsymbol{B})$，且(Ⅱ)可由(Ⅰ)线性表出，则(Ⅰ)≅(Ⅱ)

(D)若 $r(\boldsymbol{A})=r(\boldsymbol{B})$，则(Ⅰ)≅(Ⅱ)

解 应选(C).

$\boldsymbol{A}\cong\boldsymbol{B}$ 有一个大前提，要求同型. 尽管 $r(\boldsymbol{A})=r(\text{Ⅰ})$，$r(\boldsymbol{B})=r(\text{Ⅱ})$，由 $r(\text{Ⅰ})=r(\text{Ⅱ})$ 可知 $r(\boldsymbol{A})=r(\boldsymbol{B})$，但 $\boldsymbol{A},\boldsymbol{B}$ 不同型，所以 $\boldsymbol{A},\boldsymbol{B}$ 无等价可言，故(A)不成立.

(Ⅰ)≅(Ⅱ)⇔(Ⅰ)，(Ⅱ)可以相互线性表出. 仅由(Ⅰ)可由(Ⅱ)线性表出不能说明(Ⅰ)≅(Ⅱ)，故(B)不成立.

例 $\boldsymbol{A}=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2]$，$\boldsymbol{B}=\begin{bmatrix}0&1&-1\\0&0&0\\1&1&1\end{bmatrix}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]$，$r(\boldsymbol{A})=r(\boldsymbol{B})=2$，但 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 和 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 之间

不能相互表出，故(D)不成立.

由排除法可知，应选(C).

对于(C)，因由 $r(\boldsymbol{A})=r(\boldsymbol{B})$，知 $r(\text{Ⅰ})=r(\text{Ⅱ})$，并设其秩为 r，且设(Ⅰ)，(Ⅱ)的极大线性无关组分别为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r;\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$.

记向量组(Ⅲ)为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$. 由于(Ⅱ)可由(Ⅰ)线性表出，故(Ⅰ)的极大线性无关组也是(Ⅲ)的极大线性无关组，又极大线性无关组不唯一，则(Ⅱ)的极大线性无关组也是(Ⅲ)的极大线性无关组，故(Ⅰ)中任一向量均可由(Ⅱ)的极大线性无关组线性表出，故(Ⅰ)，(Ⅱ)可相互表出，(Ⅰ)$\cong$(Ⅱ)得证.

例 2.3.14 设向量组(Ⅰ)$\boldsymbol{\alpha}_1=[1,0,2]^{\mathrm{T}},\boldsymbol{\alpha}_2=[0,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[2,-1,a+4]^{\mathrm{T}}$，向量组(Ⅱ)$\boldsymbol{\beta}_1=[1,2,4]^{\mathrm{T}},\boldsymbol{\beta}_2=[1,-1,a+2]^{\mathrm{T}},\boldsymbol{\beta}_3=[3,3,10]^{\mathrm{T}}$，矩阵 $\boldsymbol{A}=\begin{bmatrix}1&0&2\\0&1&-1\\2&1&a+4\end{bmatrix}$，$\boldsymbol{B}=\begin{bmatrix}1&1&3\\2&-1&3\\4&a+2&10\end{bmatrix}$.

问：(1)$\boldsymbol{A},\boldsymbol{B}$ 是否等价，说明理由；

(2)向量组(Ⅰ)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 与(Ⅱ)$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 是否等价，说明理由.

解 将 $\boldsymbol{A},\boldsymbol{B}$ 合并，一起作初等行变换.

$$[\boldsymbol{A}\,\vdots\,\boldsymbol{B}]=\left[\begin{array}{ccc:ccc}1&0&2&1&1&3\\0&1&-1&2&-1&3\\2&1&a+4&4&a+2&10\end{array}\right]\rightarrow\left[\begin{array}{ccc:ccc}1&0&2&1&1&3\\0&1&-1&2&-1&3\\0&1&a&2&a&4\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc:ccc}1&0&2&1&1&3\\0&1&-1&2&-1&3\\0&0&a+1&0&a+1&1\end{array}\right],$$

$$[\boldsymbol{B}\,\vdots\,\boldsymbol{A}]=\left[\begin{array}{ccc:ccc}1&1&3&1&0&2\\2&-1&3&0&1&-1\\4&a+2&10&2&1&a+4\end{array}\right]\rightarrow\left[\begin{array}{ccc:ccc}1&1&3&1&0&2\\0&-3&-3&-2&1&-5\\0&a-2&-2&-2&1&a-4\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc:ccc}1&1&3&1&0&2\\0&1&1&\frac{2}{3}&-\frac{1}{3}&\frac{5}{3}\\0&a&0&-\frac{2}{3}&\frac{1}{3}&a-\frac{2}{3}\end{array}\right].$$

(1)当 $a\neq-1$ 且 $a\neq0$ 时，$r(\boldsymbol{A})=r(\boldsymbol{B})=3\Leftrightarrow\boldsymbol{A}\cong\boldsymbol{B}$；

当 $a=-1$ 时，$r(\boldsymbol{A})=2\neq r(\boldsymbol{B})=3$，$\boldsymbol{A}$ 不等价于 $\boldsymbol{B}$；

当 $a=0$ 时，$r(\boldsymbol{A})=3\neq r(\boldsymbol{B})=2$，$\boldsymbol{A}$ 不等价于 $\boldsymbol{B}$.

(2)当 $a\neq-1$ 且 $a\neq0$ 时，$r(\boldsymbol{A})=r(\boldsymbol{B})=3$，故由克拉默法则知$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\boldsymbol{x}=\boldsymbol{\beta}_i(i=1,2,3)$，以及$[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]\boldsymbol{y}=\boldsymbol{\alpha}_j(j=1,2,3)$均有解，故(Ⅰ)，(Ⅱ)可相互表出，即(Ⅰ)$\cong$(Ⅱ)；

当 $a=-1$ 时，$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=2$，$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}_3)=3$，$\boldsymbol{\beta}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出，(Ⅰ)与(Ⅱ)不等价；

当 $a=0$ 时，$r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=3$，$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=2$，$r(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\alpha}_j)=3(j=1,2,3)$，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 均不能由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表出，故(Ⅰ)与(Ⅱ)不等价.

五、向量空间（仅数学一要求）

考研数学对向量空间这一内容要求不高，但所考知识都很重要，读者仍需认真研究并掌握.

例 2.3.15 设 $\mathbf{R}^3$ 中两个基（Ⅰ）$\boldsymbol{\alpha}_1=[1,1,0]^{\mathrm{T}},\boldsymbol{\alpha}_2=[0,1,1]^{\mathrm{T}},\boldsymbol{\alpha}_3=[1,0,1]^{\mathrm{T}}$，（Ⅱ）$\boldsymbol{\beta}_1=[1,0,0]^{\mathrm{T}}$，$\boldsymbol{\beta}_2=[1,1,0]^{\mathrm{T}},\boldsymbol{\beta}_3=[1,1,1]^{\mathrm{T}}$.

（1）求基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 到基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的过渡矩阵；

（2）已知 ξ 在基（Ⅱ）$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下的坐标为 $[1,0,2]^{\mathrm{T}}$，求 ξ 在基（Ⅰ）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的坐标.

解 （1）设 $[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]\boldsymbol{C}$，其中 $\boldsymbol{C}$ 是过渡矩阵，则

$$\boldsymbol{C}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]^{-1}[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}^{-1}\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}=\begin{bmatrix}0&-1&1\\1&0&-1\\0&1&1\end{bmatrix}.$$

（2）设 $\xi=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]\begin{bmatrix}1\\0\\2\end{bmatrix}$，则

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]^{-1}[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]\begin{bmatrix}1\\0\\2\end{bmatrix}$$

$$=\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}^{-1}\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}\begin{bmatrix}1\\0\\2\end{bmatrix}$$

$$=\frac{1}{2}\begin{bmatrix}1&1&-1\\-1&1&1\\1&-1&1\end{bmatrix}\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}\begin{bmatrix}1\\0\\2\end{bmatrix}=\frac{1}{2}\begin{bmatrix}3\\1\\3\end{bmatrix},$$

得 ξ 在基（Ⅰ）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的坐标为 $\left[\frac{3}{2},\frac{1}{2},\frac{3}{2}\right]^{\mathrm{T}}$.

基础习题精练

习 题

2.3.1 设向量 $\boldsymbol{\beta}$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出，但不能由向量组（Ⅰ）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出，记向量组（Ⅱ）$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1},\boldsymbol{\beta}$，则 $\boldsymbol{\alpha}_s$（　　）.

（A）不能由（Ⅰ）线性表出，也不能由（Ⅱ）线性表出

(B)不能由(Ⅰ)线性表出,但可由(Ⅱ)线性表出

(C)可由(Ⅰ)线性表出,也可由(Ⅱ)线性表出

(D)可由(Ⅰ)线性表出,但不能由(Ⅱ)线性表出

2.3.2 设向量组 $\boldsymbol{\alpha}_1=[1,0,-1,2]^{\mathrm{T}}$,$\boldsymbol{\alpha}_2=[2,-1,-2,6]^{\mathrm{T}}$,$\boldsymbol{\alpha}_3=[3,1,t,4]^{\mathrm{T}}$ 线性无关,则参数 t 满足__________.

2.3.3 设 $n(n\geqslant 3)$ 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关,若向量组 $l\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1,m\boldsymbol{\alpha}_3-2\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_3$ 线性相关,则 m,l 应满足条件__________.

2.3.4 求向量组

$$\boldsymbol{\alpha}_1=[1,2,-5]^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=[2,-1,2]^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=[4,3,-8]^{\mathrm{T}},\quad \boldsymbol{\alpha}_4=[7,-1,1]^{\mathrm{T}}$$

的秩、极大线性无关组,并将其余向量由极大线性无关组线性表出.

2.3.5 设向量组(Ⅰ)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的秩为 r_1,向量组(Ⅱ)$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 的秩为 r_2,向量组(Ⅲ)$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$ 的秩为 r_3,则下列结论不正确的是( ).

(A)若(Ⅰ)可由(Ⅱ)线性表示,则 $r_2=r_3$

(B)若(Ⅱ)可由(Ⅰ)线性表示,则 $r_1=r_3$

(C)若 $r_1=r_3$,则 $r_2>r_1$

(D)若 $r_2=r_3$,则 $r_1\leqslant r_2$

2.3.6(仅数学一) 已知 $\mathbf{R}^3$ 的两组基 $\boldsymbol{\alpha}_1=[1,0,-1]^{\mathrm{T}}$,$\boldsymbol{\alpha}_2=[2,1,1]^{\mathrm{T}}$,$\boldsymbol{\alpha}_3=[1,1,1]^{\mathrm{T}}$ 与 $\boldsymbol{\beta}_1=[0,1,1]^{\mathrm{T}}$,$\boldsymbol{\beta}_2=[-1,1,0]^{\mathrm{T}}$,$\boldsymbol{\beta}_3=[1,2,1]^{\mathrm{T}}$.

(1)求由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵;

(2)求 $\boldsymbol{\gamma}=[9,6,5]^{\mathrm{T}}$ 在这两组基下的坐标;

(3)求向量 $\boldsymbol{\delta}$,使它在这两组基下有相同的坐标.

解答

2.3.1 (B) 解 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性表出,不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出,即存在数 $k_1,k_2,\cdots,k_s$ 使得 $\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s$,且 $k_s\neq 0$,故 $k_s\boldsymbol{\alpha}_s=\boldsymbol{\beta}-k_1\boldsymbol{\alpha}_1-k_2\boldsymbol{\alpha}_2-\cdots-k_{s-1}\boldsymbol{\alpha}_{s-1}$,表出式中不能缺少 $\boldsymbol{\beta}$,即 $\boldsymbol{\alpha}_s$ 可由 $\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出,但不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{s-1}$ 线性表出.

2.3.2 $t\neq -3$ 解 $$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=\begin{bmatrix}1&2&3\\0&-1&1\\-1&-2&t\\2&6&4\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&3\\0&1&-1\\0&0&t+3\\0&0&0\end{bmatrix},$$ $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关 $\Leftrightarrow t\neq -3$.

2.3.3 $lm=6$ 解 $$[l\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1,m\boldsymbol{\alpha}_3-2\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_3]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}-1&0&1\\l&-2&0\\0&m&-3\end{bmatrix}.$$

因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关,则 $r([\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3])=3$,则

$$r([l\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_1,m\boldsymbol{\alpha}_3-2\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1-3\boldsymbol{\alpha}_3])=r\left(\begin{bmatrix}-1&0&1\\l&-2&0\\0&m&-3\end{bmatrix}\right)<3(\text{相关}),$$

所以$\begin{vmatrix}-1&0&1\\l&-2&0\\0&m&-3\end{vmatrix}=0$，则 $lm=6$.

【注】 本题与例 2.3.3 的方法二很相似，但这里 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 $n(n\geqslant 3)$ 维列向量，故使用$|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|$是不妥的，这里用到如下一个更一般的秩的结论：

若 $r(\boldsymbol{A}_{m\times n})=n$(列满秩)，则 $r(\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s})=r(\boldsymbol{B}_{n\times s})$.

2.3.4 解 $$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4]=\begin{bmatrix}1&2&4&7\\2&-1&3&-1\\-5&2&-8&1\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&4&7\\0&-5&-5&-15\\0&12&12&36\end{bmatrix}$$

$$\longrightarrow\begin{bmatrix}1&2&4&7\\0&1&1&3\\0&0&0&0\end{bmatrix}=[\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\boldsymbol{\alpha}_3',\boldsymbol{\alpha}_4'].$$

由 $\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\boldsymbol{\alpha}_3',\boldsymbol{\alpha}_4'$ 知 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=2$，极大线性无关组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$(不唯一，$\boldsymbol{\alpha}_1$ 和其他任意一个向量的组合均是)．且 $\boldsymbol{\alpha}_3=2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4=\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2$.

2.3.5 (C) **解** 这里(Ⅰ)显然天生可由(Ⅲ)线性表示，若再有 $r_1=r_3$，则(Ⅰ)和(Ⅲ)等价，此时(Ⅲ)便可由(Ⅰ)线性表示，即此时(Ⅱ)便可由(Ⅰ)线性表示，则 $r_2\leqslant r_1$，故(C)错误，同理可知(D)正确，而(A)和(B)显然是正确的.

2.3.6 解 (1)设从基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 的过渡矩阵是 $\boldsymbol{C}$，则

$$[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\boldsymbol{C},$$

故

$$\begin{aligned}\boldsymbol{C}&=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]^{-1}[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]\\&=\begin{bmatrix}1&2&1\\0&1&1\\-1&1&1\end{bmatrix}^{-1}\begin{bmatrix}0&-1&1\\1&1&2\\1&0&1\end{bmatrix}=\begin{bmatrix}0&1&1\\-1&-3&-2\\2&4&4\end{bmatrix}.\end{aligned}$$

(2)设 $\boldsymbol{\gamma}$ 在基 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 下坐标是$[y_1,y_2,y_3]^{\mathrm{T}}$，即 $y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2+y_3\boldsymbol{\beta}_3=\boldsymbol{\gamma}$，亦即

$$\begin{cases}-y_2+y_3=9,\\y_1+y_2+2y_3=6,\Rightarrow y_1=0,y_2=-4,y_3=5.\\y_1+y_3=5\end{cases}$$

设 $\boldsymbol{\gamma}$ 在基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 下的坐标是$[x_1,x_2,x_3]^{\mathrm{T}}$，按坐标变换公式 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$，有

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0&1&1\\-1&-3&-2\\2&4&4\end{bmatrix}\begin{bmatrix}0\\-4\\5\end{bmatrix}=\begin{bmatrix}1\\2\\4\end{bmatrix},$$

可见 $\boldsymbol{\gamma}$ 在这两组基下的坐标分别是$[1,2,4]^{\mathrm{T}}$和$[0,-4,5]^{\mathrm{T}}$.

(3)设 $\boldsymbol{\delta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=x_1\boldsymbol{\beta}_1+x_2\boldsymbol{\beta}_2+x_3\boldsymbol{\beta}_3$，即

$$x_1(\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1)+x_2(\boldsymbol{\alpha}_2-\boldsymbol{\beta}_2)+x_3(\boldsymbol{\alpha}_3-\boldsymbol{\beta}_3)=\mathbf{0},$$

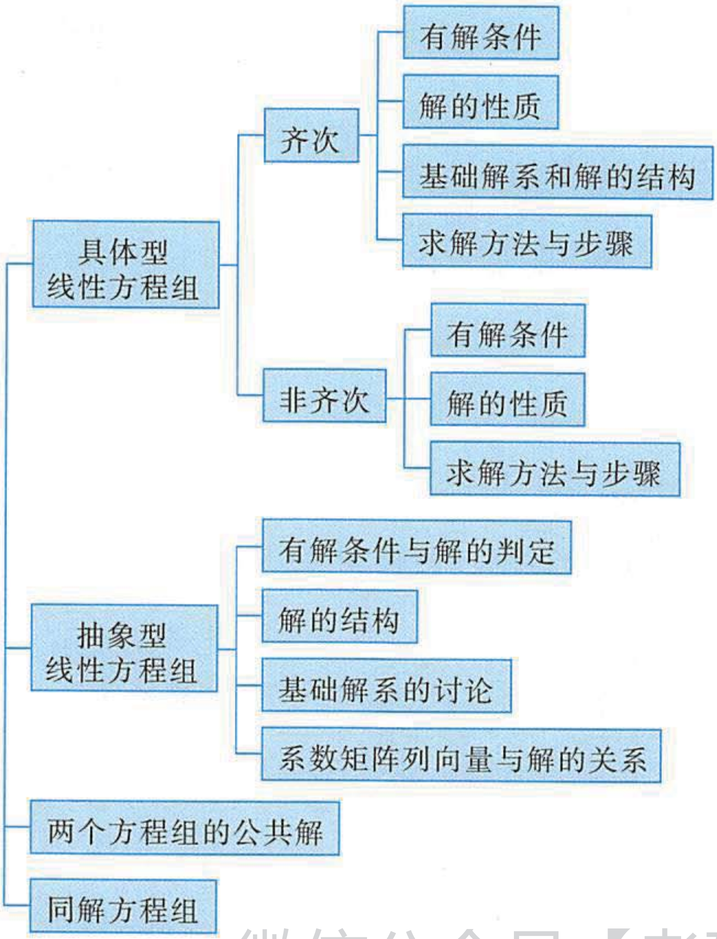
亦即

$$\begin{cases}x_1+3x_2=0,\\-x_1-x_3=0,\\-2x_1+x_2=0\end{cases}\Rightarrow x_1=x_2=x_3=0.$$

所以，仅零向量在这两组基下有相同的坐标.

第4讲 线性方程组

基础知识结构



基础内容精讲

一、线性方程组与向量组其实是一回事

我们来看一般的非齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\qquad\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m, \end{cases}$$

该方程组的系数矩阵

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

就是若干个列向量拼成的，且其增广矩阵

$$\left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

就是系数矩阵再添加一个列向量拼成的.

仔细观察，把上述方程组写成向量的形式便不难看出，**该方程组的未知数就是向量组中各成员的系数**：

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta},$$

其中

$$\boldsymbol{\alpha}_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, j=1,2,\cdots,n,\quad \boldsymbol{\beta}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

所以从本质上说，**方程组问题就是向量组问题，方程组和向量组是同一个问题的两种表现形式，其本质一样，所以解决方法也一样.**

求解线性方程组，就是对增广矩阵作初等行变换，化成行阶梯形矩阵，然后求解. 这个基本方法贯穿这门课程始终.

接下来，该方程组有无穷多解时，如何表示出所有的解？这又是某个(无穷)向量组用什么"代表"来表示的问题. **这个"代表"就是基础解系.** 于是，解线性方程组便成了这一部分的关键. 只是希望读者发现：解方程组所得到的解 $x_1,x_2,\cdots,x_n$，不就是描述向量与向量之间关系的表示系数吗？

二、齐次线性方程组

方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0, \\ \cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases} \quad (\text{I})$$

称为 m 个方程 n 个未知量的齐次线性方程组，其向量形式为

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{0},$$

其中

$$\boldsymbol{\alpha}_j=\begin{bmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{bmatrix},j=1,2,\cdots,n.$$

其矩阵形式为

$$\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0},$$

其中

$$\boldsymbol{A}_{m\times n}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

1. 有解的条件

当 $r(\boldsymbol{A})=n$ 时($\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关)，方程组(Ⅰ)有唯一零解；

当 $r(\boldsymbol{A})=r<n$ 时($\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性相关)，方程组(Ⅰ)有非零解，且有 $n-r$ 个线性无关解.

2. 解的性质

若 $\boldsymbol{A}\boldsymbol{\xi}_1=\boldsymbol{0},\boldsymbol{A}\boldsymbol{\xi}_2=\boldsymbol{0}$，则 $\boldsymbol{A}(k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2)=\boldsymbol{0}$，其中 k_1,k_2 是任意常数.

3. 基础解系和解的结构

(1)基础解系：设 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$ 满足

①是方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解；

②线性无关；

③方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的任一解均可由 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$ 线性表出，

则称 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$ 为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系.

(2)通解：设 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系，则 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}$ 是方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的通解，其中 $k_1,k_2,\cdots,k_{n-r}$ 是任意常数.

4. 求解方法与步骤

①将系数矩阵 $\boldsymbol{A}$ 作初等行变换化成阶梯形矩阵 $\boldsymbol{B}$(或最简阶梯形矩阵 $\boldsymbol{B}$)，初等行变换将方程组化为同解方程组，故 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 和 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 同解，只需解 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 即可. 设 $r(\boldsymbol{A})=r$，

$$\boldsymbol{A}\xrightarrow{\text{初等行变换}}\boldsymbol{B}=\begin{bmatrix}c_{11}&c_{12}&\cdots&c_{1r}&\cdots&c_{1n}\\0&c_{22}&\cdots&c_{2r}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots&&\vdots\\0&0&\cdots&c_{rr}&\cdots&c_{rn}\\0&0&\cdots&0&\cdots&0\\\vdots&\vdots&&\vdots&&\vdots\\0&0&\cdots&0&\cdots&0\end{bmatrix}_{m\times n},$$

其中，m 是原方程组中方程个数，n 是未知量个数.

②按列找出一个秩为 r 的子矩阵，则剩余列位置的未知数即设为自由变量.

③按基础解系定义求出 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$，并写出通解.

三、非齐次线性方程组

方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \tag{Ⅱ}$$

称为 m 个方程 n 个未知量的非齐次线性方程组，其向量形式为 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{b}$，其中

$$\boldsymbol{\alpha}_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, j=1,2,\cdots,n, \quad \boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

其矩阵形式为 $\boldsymbol{Ax}=\boldsymbol{b}$，其中

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

矩阵 $\left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$ 称为矩阵 $\boldsymbol{A}$ 的增广矩阵，简记成 $[\boldsymbol{A} \mid \boldsymbol{b}]$ 或 $[\boldsymbol{A},\boldsymbol{b}]$.

1. 有解的条件

若 $r(\boldsymbol{A})\neq r([\boldsymbol{A},\boldsymbol{b}])$（$\boldsymbol{b}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表出），则方程组（Ⅱ）无解；

若 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])=n$（即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{b}$ 线性相关），则方程组（Ⅱ）有唯一解；

若 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])=r<n$，则方程组（Ⅱ）有无穷多解.

2. 解的性质

设 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}$ 是非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，$\boldsymbol{\xi}$ 是对应齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，则：①$\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解；②$k\boldsymbol{\xi}+\boldsymbol{\eta}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解.

3. 求解方法与步骤

将增广矩阵作初等行变换化成阶梯形（或最简阶梯形）矩阵，求出对应齐次线性方程组的通解，再加上一个非齐次线性方程组的特解即是非齐次线性方程组的通解.

①写出 $\boldsymbol{Ax}=\boldsymbol{b}$ 的导出方程组 $\boldsymbol{Ax}=\boldsymbol{0}$，并求 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}$.

②求出 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个特解 $\boldsymbol{\eta}$.

③则 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解为 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}+\boldsymbol{\eta}$，其中 $k_1,k_2,\cdots,k_{n-r}$ 为任意常数.

基础例题精解

一、具体型线性方程组

线性方程组是线性代数课程的核心内容之一，几乎每年的考研数学大题，均对线性方程组的求解提出了较高要求，读者要高度重视.

本部分例题均为"系数矩阵为具体数字"的方程组.只不过有些系数矩阵全是已知数，而有些系数矩阵含有未知参数.读者应认真研究这些例题，掌握求解方法，规范解题步骤.

例 2.4.1 求齐次线性方程组

$$\begin{cases} x_1+x_2-3x_4-x_5=0, \\ x_1-x_2+2x_3-x_4=0, \\ 4x_1-2x_2+6x_3+3x_4-4x_5=0, \\ 2x_1+4x_2-2x_3+4x_4-7x_5=0 \end{cases}$$

的通解.

解 将系数矩阵作初等行变换，化成阶梯形矩阵.

$$\boldsymbol{A}=\begin{bmatrix} 1 & 1 & 0 & -3 & -1 \\ 1 & -1 & 2 & -1 & 0 \\ 4 & -2 & 6 & 3 & -4 \\ 2 & 4 & -2 & 4 & -7 \end{bmatrix}$$

$$\xrightarrow[\substack{③-4① \\ ④-2①}]{②-①}\begin{bmatrix} 1 & 1 & 0 & -3 & -1 \\ 0 & -2 & 2 & 2 & 1 \\ 0 & -6 & 6 & 15 & 0 \\ 0 & 2 & -2 & 10 & -5 \end{bmatrix}$$

$$\xrightarrow[④+②]{③-3②}\begin{bmatrix} 1 & 1 & 0 & -3 & -1 \\ 0 & -2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 9 & -3 \\ 0 & 0 & 0 & 12 & -4 \end{bmatrix}$$

$$\xrightarrow[\frac{1}{3}③]{④-\frac{4}{3}③}\begin{bmatrix} 1 & 1 & 0 & -3 & -1 \\ 0 & -2 & 2 & 2 & 1 \\ 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}=\boldsymbol{B},$$

则 $\boldsymbol{Ax}=\boldsymbol{0}$ 和 $\boldsymbol{Bx}=\boldsymbol{0}$ 是同解方程组，且 $r(\boldsymbol{A})=r(\boldsymbol{B})=3$.

按列找出一个秩为 3 的子矩阵，可取第一、二、四列，则剩余第三、五列位置的元素 x_3，x_5 即设为自由未知量.

取自由未知量 $x_3=k_1$，$x_5=3k_2$，代入方程得

$$x_4=k_2,$$

$$x_2=x_3+x_4+\frac{1}{2}x_5=k_1+\frac{5}{2}k_2,$$

$$x_1=-x_2+3x_4+x_5=-k_1+\frac{7}{2}k_2.$$

由此得通解

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=\begin{bmatrix}-k_1+\frac{7}{2}k_2\\k_1+\frac{5}{2}k_2\\k_1+0\\0+k_2\\0+3k_2\end{bmatrix}=k_1\begin{bmatrix}-1\\1\\1\\0\\0\end{bmatrix}+k_2\begin{bmatrix}\frac{7}{2}\\\frac{5}{2}\\0\\1\\3\end{bmatrix},\text{其中 } k_1,k_2 \text{ 是任意常数}.$$

【注】 (1)由阶梯形矩阵 $\boldsymbol{B}$ 知，未知量 x_2,x_3 地位等同，同样 x_4,x_5 的地位也等同，自由未知量也可以取为 x_2,x_4 或 x_2,x_5 或 x_3,x_4.

(2)方程组求解的回代过程也可以用初等行变换来实现，即对阶梯形矩阵 $\boldsymbol{B}$ 继续作初等行变换，化成行最简阶梯形矩阵.

$$\boldsymbol{B}=\begin{bmatrix}1&1&0&-3&-1\\0&-2&2&2&1\\0&0&0&3&-1\\0&0&0&0&0\end{bmatrix}\xrightarrow[\frac{1}{3}③]{-\frac{1}{2}②}\begin{bmatrix}1&1&0&-3&-1\\0&1&-1&-1&-\frac{1}{2}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}$$

$$\xrightarrow{①-②}\begin{bmatrix}1&0&1&-2&-\frac{1}{2}\\0&1&-1&-1&-\frac{1}{2}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}\xrightarrow[①+2③]{②+③}\begin{bmatrix}1&0&1&0&-\frac{7}{6}\\0&1&-1&0&-\frac{5}{6}\\0&0&0&1&-\frac{1}{3}\\0&0&0&0&0\end{bmatrix}=\boldsymbol{C}.$$

取自由未知量 $x_3=k_1,x_5=3k_2$，则

$$x_4=k_2,\quad x_2=\frac{5}{2}k_2+k_1,\quad x_1=\frac{7}{2}k_2-k_1,$$

故通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}=\begin{bmatrix}\frac{7}{2}k_2-k_1\\\frac{5}{2}k_2+k_1\\k_1\\k_2\\3k_2\end{bmatrix}=k_2\begin{bmatrix}\frac{7}{2}\\\frac{5}{2}\\0\\1\\3\end{bmatrix}+k_1\begin{bmatrix}-1\\1\\1\\0\\0\end{bmatrix},\text{其中 } k_1,k_2 \text{ 是任意常数}.$$

或者取自由未知量 $x_3=1, x_5=0$ 及 $x_3=0, x_5=3$ 分别代入，求得基础解系

$$\xi_1=[-1,1,1,0,0]^{\mathrm{T}},\quad \xi_2=\left[\frac{7}{2},\frac{5}{2},0,1,3\right]^{\mathrm{T}},$$

故通解为 $k_1\xi_1+k_2\xi_2$，其中 k_1,k_2 是任意常数.

例 2.4.2 求解非齐次线性方程组

$$\begin{cases}x_1+5x_2-x_3-x_4=-1,\\x_1-2x_2+x_3+3x_4=3,\\3x_1+8x_2-x_3+x_4=1,\\x_1-9x_2+3x_3+7x_4=7,\end{cases}$$

并用对应的齐次线性方程组的基础解系表示通解.

解 对增广矩阵作初等行变换化成阶梯形矩阵.

$$[\boldsymbol{A}\,\vdots\,\boldsymbol{b}]=\left[\begin{array}{cccc|c}1&5&-1&-1&-1\\1&-2&1&3&3\\3&8&-1&1&1\\1&-9&3&7&7\end{array}\right]\xrightarrow[④-①]{\substack{②-①\\③-3①}}\left[\begin{array}{cccc|c}1&5&-1&-1&-1\\0&-7&2&4&4\\0&-7&2&4&4\\0&-14&4&8&8\end{array}\right]$$

$$\xrightarrow[④-2②]{③-②}\left[\begin{array}{cccc|c}1&5&-1&-1&-1\\0&-7&2&4&4\\0&0&0&0&0\\0&0&0&0&0\end{array}\right].$$

令自由未知量 $x_3=k_1, x_4=k_2$，代入得

$$x_2=-\frac{1}{7}(4-2k_1-4k_2)=-\frac{4}{7}+\frac{2}{7}k_1+\frac{4}{7}k_2,$$

$$x_1=-1+k_1+k_2-5\left(-\frac{4}{7}+\frac{2}{7}k_1+\frac{4}{7}k_2\right)=\frac{13}{7}-\frac{3}{7}k_1-\frac{13}{7}k_2.$$

得通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}\frac{13}{7}-\frac{3}{7}k_1-\frac{13}{7}k_2\\-\frac{4}{7}+\frac{2}{7}k_1+\frac{4}{7}k_2\\k_1\\k_2\end{bmatrix}=\begin{bmatrix}\frac{13}{7}\\-\frac{4}{7}\\0\\0\end{bmatrix}+k_1\begin{bmatrix}-\frac{3}{7}\\\frac{2}{7}\\1\\0\end{bmatrix}+k_2\begin{bmatrix}-\frac{13}{7}\\\frac{4}{7}\\0\\1\end{bmatrix},$$

其中 k_1,k_2 是任意常数，$\left[-\frac{3}{7},\frac{2}{7},1,0\right]^{\mathrm{T}},\left[-\frac{13}{7},\frac{4}{7},0,1\right]^{\mathrm{T}}$ 为对应的齐次线性方程组的基础解系.

例 2.4.3 已知线性方程组

$$\begin{cases}x_1+x_2+x_3+x_4=1,\\3x_1+2x_2+x_3+x_4=a,\\x_2+2x_3+2x_4=3,\\5x_1+4x_2+3x_3+4x_4=b.\end{cases}$$

则 a,b 为何值时，方程组无解？a,b 为何值时，方程组有解？方程组有解时，求其全部解.

解 对方程组的增广矩阵$[\boldsymbol{A} \vdots \boldsymbol{b}]$作初等行变换.

$$[\boldsymbol{A} \vdots \boldsymbol{b}]=\left[\begin{array}{cccc|c}1&1&1&1&1\\3&2&1&1&a\\0&1&2&2&3\\5&4&3&4&b\end{array}\right]\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&1\\0&-1&-2&-2&a-3\\0&1&2&2&3\\0&-1&-2&-1&b-5\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&1\\0&1&2&2&3\\0&0&0&0&a\\0&0&0&1&b-2\end{array}\right]\longrightarrow\left[\begin{array}{cccc|c}1&1&1&1&1\\0&1&2&2&3\\0&0&0&1&b-2\\0&0&0&0&a\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{cccc|c}1&0&-1&-1&-2\\0&1&2&2&3\\0&0&0&1&b-2\\0&0&0&0&a\end{array}\right].$$

当$a\neq0$,b任意时,$r(\boldsymbol{A})=3\neq r([\boldsymbol{A},\boldsymbol{b}])=4$,方程组无解;

当$a=0$,b任意时,$r(\boldsymbol{A})=3=r([\boldsymbol{A},\boldsymbol{b}])<4$,方程组有无穷多解.

取自由未知量$x_3=k$,代入方程,解得

$$x_4=b-2,\quad x_2=3-2(b-2)-2k=-2k+7-2b,\quad x_1=-2+k+(b-2)=k+b-4,$$

故通解为

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}k+b-4\\-2k+7-2b\\k\\b-2\end{bmatrix}=k\begin{bmatrix}1\\-2\\1\\0\end{bmatrix}+\begin{bmatrix}b-4\\7-2b\\0\\b-2\end{bmatrix},$$

其中k是任意常数.

二、抽象型线性方程组

"系数矩阵为抽象矩阵"的方程组,要把握以下四种基本问题.

1. 方程组有解的条件及解的判别

这里的理论主要有以下三条.

(1)$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$: 总有解,至少有零解.

(2)$\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$: $r(\boldsymbol{A})=n$,只有零解;

$r(\boldsymbol{A})<n$,有无穷多解.

(3)$\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{b}$: $r(\boldsymbol{A})\neq r([\boldsymbol{A},\boldsymbol{b}])$,无解;

$r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])=n$,有唯一解;

$r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])=r<n$,有无穷多解.

例 2.4.4 设非齐次线性方程组为$\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{b}$,则( ).

(A)当$r(\boldsymbol{A})=m$时,方程组有解 (B)当$r(\boldsymbol{A})=n$时,方程组有唯一解

(C)当 $m=n$ 时，方程组有唯一解　　　　(D)当 $r(\boldsymbol{A})<n$ 时，方程组有无穷多解

分析　$\boldsymbol{Ax}=\boldsymbol{b}$ 有解（无解）$\Leftrightarrow r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])$（$r(\boldsymbol{A})\neq r([\boldsymbol{A},\boldsymbol{b}])$）.

题设条件未提及 $r([\boldsymbol{A},\boldsymbol{b}])$，这正是要求读者考虑的问题.

解　应选(A).

因 $r(\boldsymbol{A})=m$ 时，$\boldsymbol{A}$ 的行向量组线性无关，增广矩阵在 $\boldsymbol{A}$ 的基础上增加一列，即行向量增加一个分量，仍线性无关，故有 $r(\boldsymbol{A})=m=r([\boldsymbol{A},\boldsymbol{b}])$，方程组有解，应选(A).

(B)$r(\boldsymbol{A})=n$ 时可能无解；(C)可能无解或有无穷多解；(D)可能无解，均可排除.

例 2.4.5　设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{Ax}=\boldsymbol{0}$ 是非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 所对应的齐次线性方程组，则下列结论正确的是(　　).

(A)若 $\boldsymbol{Ax}=\boldsymbol{0}$ 仅有零解，则 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解

(B)若 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解，则 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解

(C)若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 仅有零解

(D)若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解

解　应选(D).

$\boldsymbol{Ax}=\boldsymbol{0}$ 仅有零解，则 $r(\boldsymbol{A})=n$，此时 $r([\boldsymbol{A},\boldsymbol{b}])$ 可能是 n，也可能是 $n+1$，故(A)不成立；

$\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解，则 $r(\boldsymbol{A})<n$，此时 $r([\boldsymbol{A},\boldsymbol{b}])$ 可能等于 $r(\boldsymbol{A})$，也可能不等于 $r(\boldsymbol{A})$，故(B)不成立；

$\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解，则 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])<n$，故此时 $\boldsymbol{Ax}=\boldsymbol{0}$ 有无穷多解，即有非零解，故(C)不成立. 应选(D).

【注】　(1)若 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解，则 $r(\boldsymbol{A})=n$(列满秩)$\nRightarrow r([\boldsymbol{A},\boldsymbol{b}])=n$，故 $\boldsymbol{Ax}=\boldsymbol{b}$ 可能有解，可能无解.

(2)若 $\boldsymbol{Ax}=\boldsymbol{0}$ 有无穷多解(有非零解)，则 $r(\boldsymbol{A})<n$(列不满秩)$\nRightarrow r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])$，故 $\boldsymbol{Ax}=\boldsymbol{b}$ 可能有解，可能无解.

(3)若 $\boldsymbol{A}$ 行满秩，则 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])$，故 $\boldsymbol{Ax}=\boldsymbol{b}$ 必有解.

(4)若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解，则 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])=\boldsymbol{A}$ 的列数，故 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解.

(5)若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解，则 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])<\boldsymbol{A}$ 的列数，故 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解.

由(1),(2)可知,(4),(5)不能倒推.

2. 解的结构

这里的理论依据：

(1)齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有基础解系 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$，则通解为

$$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r},$$

其中 $k_1,k_2,\cdots,k_{n-r}$ 是任意常数；

(2)非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有特解 $\boldsymbol{\eta}$，对应的齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有基础解系 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_{n-r}$，则 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解为 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r}+\boldsymbol{\eta}$，其中 $k_1,k_2,\cdots,k_{n-r}$ 是任意常数.

例 2.4.6　设 $r(\boldsymbol{A}_{4\times4})=2$，$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的 3 个解向量，其中

$$\begin{cases}\boldsymbol{\eta}_1-\boldsymbol{\eta}_2=[-1,0,3,-4]^{\mathrm{T}},\\ \boldsymbol{\eta}_1+\boldsymbol{\eta}_2=[3,2,1,-2]^{\mathrm{T}},\\ \boldsymbol{\eta}_3+2\boldsymbol{\eta}_2=[5,1,0,3]^{\mathrm{T}},\end{cases}$$

则 $Ax=b$ 的通解是__________.

解 应填 $k_1[-1,0,3,-4]^{\mathrm{T}}+k_2[-1,4,3,-12]^{\mathrm{T}}+\left[\frac{3}{2},1,\frac{1}{2},-1\right]^{\mathrm{T}}$，其中 k_1,k_2 是任意常数.

由题设 $A_{4\times4}x=b$，$r(A)=2$，$n=4$，知 $Ax=b$ 的通解结构为

$$k_1\xi_1+k_2\xi_2+\boldsymbol{\eta},$$

其中 ξ_1,ξ_2 是 $Ax=0$ 的基础解系，$\boldsymbol{\eta}$ 是 $Ax=b$ 的一个特解.

因 $A(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2)=b-b=0$，$A[3(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)-2(\boldsymbol{\eta}_3+2\boldsymbol{\eta}_2)]=6b-6b=0$，故 $\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$，$3(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)-2(\boldsymbol{\eta}_3+2\boldsymbol{\eta}_2)$ 为 $Ax=0$ 的解向量，因此可取

$$\xi_1=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2=[-1,0,3,-4]^{\mathrm{T}},$$

$$\xi_2=3(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)-2(\boldsymbol{\eta}_3+2\boldsymbol{\eta}_2)=[9,6,3,-6]^{\mathrm{T}}-[10,2,0,6]^{\mathrm{T}}=[-1,4,3,-12]^{\mathrm{T}},$$

显然 ξ_1,ξ_2 线性无关.

又因 $A\left[\frac{1}{2}(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)\right]=\frac{1}{2}(b+b)=b$，故 $\frac{1}{2}(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)$ 为 $Ax=b$ 的一个特解，因此可取

$$\boldsymbol{\eta}=\frac{1}{2}(\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)=\left[\frac{3}{2},1,\frac{1}{2},-1\right]^{\mathrm{T}},$$

故 $Ax=b$ 的通解为

$$k_1\xi_1+k_2\xi_2+\boldsymbol{\eta}=k_1[-1,0,3,-4]^{\mathrm{T}}+k_2[-1,4,3,-12]^{\mathrm{T}}+\left[\frac{3}{2},1,\frac{1}{2},-1\right]^{\mathrm{T}},$$

其中 k_1,k_2 是任意常数.

【注】 读者还可以尝试使用如下方法解答：

因 $[\boldsymbol{\eta}_1-\boldsymbol{\eta}_2,\boldsymbol{\eta}_1+\boldsymbol{\eta}_2,\boldsymbol{\eta}_3+2\boldsymbol{\eta}_2]=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3]\begin{bmatrix}1&1&0\\-1&1&2\\0&0&1\end{bmatrix}$，故

$$[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3]=[\boldsymbol{\eta}_1-\boldsymbol{\eta}_2,\boldsymbol{\eta}_1+\boldsymbol{\eta}_2,\boldsymbol{\eta}_3+2\boldsymbol{\eta}_2]\begin{bmatrix}1&1&0\\-1&1&2\\0&0&1\end{bmatrix}^{-1},$$

由此可以直接求出 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$，从而得到通解为 $k_1(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2)+k_2(\boldsymbol{\eta}_2-\boldsymbol{\eta}_3)+\boldsymbol{\eta}_3$，其中 k_1,k_2 是任意常数.

3. 方程组 $Ax=0$ 的基础解系的讨论

基础解系是一个十分重要的知识点，给出

$$A_{m\times n}x=0,r(A)=r.$$

若向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 满足：①$A\boldsymbol{\alpha}_i=0,i=1,2,\cdots,s$；②$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关；③$s=n-r$，则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 为 $Ax=0$ 的基础解系. 读者要研究上述①②③是否都成立，缺一不可.

例 2.4.7 设 $\xi_1=[1,-2,3,1]^{\mathrm{T}}$，$\xi_2=[2,0,5,-2]^{\mathrm{T}}$ 是齐次线性方程组 $A_{3\times4}x=0$ 的解，且 $r(A)=2$，则下列向量中是 $Ax=0$ 的解向量的是(　　).

(A)$\boldsymbol{\alpha}_1=[1,-2,3,2]^{\mathrm{T}}$　　(B)$\boldsymbol{\alpha}_2=[0,0,5,-2]^{\mathrm{T}}$

(C)$\boldsymbol{\alpha}_3=[-1,-6,-1,7]^{\mathrm{T}}$　　(D)$\boldsymbol{\alpha}_4=[1,6,1,6]^{\mathrm{T}}$

分析 $\boldsymbol{\alpha}_i(i=1,2,3,4)$ 是齐次线性方程组的解$\Leftrightarrow\boldsymbol{\alpha}_i$ 可由基础解系线性表出.

解 应选(C).

$r(\boldsymbol{A}_{3\times 4})=2,n=4,\xi_1,\xi_2$ 线性无关，故 ξ_1,ξ_2 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系.

若向量 $\boldsymbol{\alpha}_i(i=1,2,3,4)$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，则 $\boldsymbol{\alpha}_i$ 可由其基础解系线性表出. 将 $\xi_1,\xi_2,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 合并成矩阵

$$\boldsymbol{B}=[\xi_1,\xi_2,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4],$$

并作初等行变换，有

$$\boldsymbol{B}=[\xi_1,\xi_2,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4]=\left[\begin{array}{cc|cccc}1&2&1&0&-1&1\\-2&0&-2&0&-6&6\\3&5&3&5&-1&1\\1&-2&2&-2&7&6\end{array}\right]$$

$$\to\left[\begin{array}{cc|cccc}1&2&1&0&-1&1\\0&4&0&0&-8&8\\0&-1&0&5&2&-2\\0&-4&1&-2&8&5\end{array}\right]\to\left[\begin{array}{cc|cccc}1&2&1&0&-1&1\\0&-1&0&5&2&-2\\0&0&0&20&0&0\\0&0&1&-22&0&13\end{array}\right]$$

$$=[\xi_1',\xi_2',\boldsymbol{\alpha}_1',\boldsymbol{\alpha}_2',\boldsymbol{\alpha}_3',\boldsymbol{\alpha}_4']\xlongequal{\text{记}}\boldsymbol{C}.$$

由矩阵 $\boldsymbol{C}$ 可知，$\boldsymbol{\alpha}_3'$ 可由 ξ_1',ξ_2' 线性表出，故 $\boldsymbol{\alpha}_3$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解向量，其余向量均不是，故应选(C).

【注】 作为练习，请读者解下题.

设非齐次线性方程组 $\boldsymbol{A}_{3\times 4}\boldsymbol{x}=\boldsymbol{b}$ 有通解

$$k_1\xi_1+k_2\xi_2+\boldsymbol{\eta}=k_1[1,2,0,-2]^{\mathrm{T}}+k_2[4,-1,-1,-1]^{\mathrm{T}}+[1,0,-1,1]^{\mathrm{T}},$$

则下列向量中是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解的是( ).

(A) $\boldsymbol{\alpha}_1=[1,2,0,-2]^{\mathrm{T}}$ (B) $\boldsymbol{\alpha}_2=[6,1,-2,-2]^{\mathrm{T}}$

(C) $\boldsymbol{\alpha}_3=[3,1,-2,4]^{\mathrm{T}}$ (D) $\boldsymbol{\alpha}_4=[5,1,-1,-3]^{\mathrm{T}}$

解 应选(B).

$\boldsymbol{\alpha}_i(i=1,2,3,4)$ 是否是方程组的解，可以考查是否属于通解，即是否存在数 k_1,k_2，使得

$$\boldsymbol{\alpha}_i=k_1\xi_1+k_2\xi_2+\boldsymbol{\eta},$$

即 $\boldsymbol{\alpha}_i-\boldsymbol{\eta}=k_1\xi_1+k_2\xi_2$，接下来的方法可仿例 2.4.7，不再赘述.

例 2.4.8 设 ξ_1,ξ_2,ξ_3 是方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系，则下列向量组中也是方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系的是( ).

(A) $\xi_1-\xi_2,\xi_2-\xi_3,\xi_3-\xi_1$ (B) $\xi_1+\xi_2,\xi_2-\xi_3,\xi_3+\xi_1$

(C) $\xi_1+\xi_2-\xi_3,\xi_1+2\xi_2+\xi_3,2\xi_1+3\xi_2$ (D) $\xi_1+\xi_2,\xi_2+\xi_3,\xi_3+\xi_1$

分析 本题基础解系应由三个线性无关的解向量组成，题设的四个选项，均是三个向量，且由解的性质知，三个向量均是解向量，故关键是看哪个选项是线性无关向量组.

解 应选(D).

$(\xi_1-\xi_2)+(\xi_2-\xi_3)+(\xi_3-\xi_1)=\boldsymbol{0}$;

$-(\xi_1+\xi_2)+(\xi_2-\xi_3)+(\xi_3+\xi_1)=\boldsymbol{0}$;

$2\xi_1+3\xi_2=(\xi_1+\xi_2-\xi_3)+(\xi_1+2\xi_2+\xi_3)$.

故(A),(B),(C)中向量组都是线性相关的，由排除法，应选(D).

对于(D)，$[\boldsymbol{\xi}_1+\boldsymbol{\xi}_2,\boldsymbol{\xi}_2+\boldsymbol{\xi}_3,\boldsymbol{\xi}_3+\boldsymbol{\xi}_1]=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3]\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}$，两边取行列式，因 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 线性无关，则 $|\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3|\neq 0$. 又

$$\begin{vmatrix}1&0&1\\1&1&0\\0&1&1\end{vmatrix}=2\neq 0,$$

故 $|\boldsymbol{\xi}_1+\boldsymbol{\xi}_2,\boldsymbol{\xi}_2+\boldsymbol{\xi}_3,\boldsymbol{\xi}_3+\boldsymbol{\xi}_1|\neq 0$，因此 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2,\boldsymbol{\xi}_2+\boldsymbol{\xi}_3,\boldsymbol{\xi}_3+\boldsymbol{\xi}_1$ 是线性无关的，故应选(D).

4. 线性方程组系数矩阵列向量和解的关系

(1)齐次线性方程组

$$\boldsymbol{\alpha}_1x_1+\boldsymbol{\alpha}_2x_2+\cdots+\boldsymbol{\alpha}_nx_n=\boldsymbol{0}$$

的解是使 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的线性组合为零向量时线性组合的系数.

(2)非齐次线性方程组

$$\boldsymbol{\alpha}_1x_1+\boldsymbol{\alpha}_2x_2+\cdots+\boldsymbol{\alpha}_nx_n=\boldsymbol{\beta}$$

的解是 $\boldsymbol{\beta}$ 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表出的表出系数.

简而言之，“方程组的解就是描述列向量组中各向量之间数量关系的系数.”这个观点对于解题也很有用处.

例 2.4.9 已知 $\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4]$，其中 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 是 4 维列向量，且 $\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，$r(\boldsymbol{A})=3$. 若 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3+4\boldsymbol{\alpha}_4$，则线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的通解是__________.

解 应填 $k[1,-2,-1,0]^{\mathrm{T}}+[1,2,3,4]^{\mathrm{T}}$，其中 k 是任意常数.

因 $\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3=\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3+0\boldsymbol{\alpha}_4=\boldsymbol{0}$，故 $\boldsymbol{\xi}=[1,-2,-1,0]^{\mathrm{T}}$ 是对应齐次线性方程组的非零解，又 $\boldsymbol{\eta}=[1,2,3,4]^{\mathrm{T}}$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的特解，$r(\boldsymbol{A})=3$. 故线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的通解为 $k\boldsymbol{\xi}+\boldsymbol{\eta}$，其中 k 是任意常数.

例 2.4.10 已知线性方程组

$$\boldsymbol{\alpha}_1x_1+\boldsymbol{\alpha}_2x_2+\boldsymbol{\alpha}_3x_3+\boldsymbol{\alpha}_4x_4=\boldsymbol{\alpha}_5$$

有通解 $[2,0,0,1]^{\mathrm{T}}+k[1,-1,2,0]^{\mathrm{T}}$，则下列说法正确的是(　　).

(A) $\boldsymbol{\alpha}_5$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出　　(B) $\boldsymbol{\alpha}_4$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出

(C) $\boldsymbol{\alpha}_5$ 不能由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性表出　　(D) $\boldsymbol{\alpha}_4$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_5$ 线性表出

解 应选(B).

由题设条件知

$$\boldsymbol{\alpha}_5=(k+2)\boldsymbol{\alpha}_1-k\boldsymbol{\alpha}_2+2k\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4.$$

$\boldsymbol{\alpha}_5$ 的表出式中必定有 $\boldsymbol{\alpha}_4$，故 $\boldsymbol{\alpha}_5$ 不可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 表出，排除(A)；

$\boldsymbol{\alpha}_5$ 可由 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 表出，取 $k=-2$，即有 $\boldsymbol{\alpha}_5=2\boldsymbol{\alpha}_2-4\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，排除(C)；

$\boldsymbol{\alpha}_4$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_5$ 表出，取 $k=0$，即有 $\boldsymbol{\alpha}_4=\boldsymbol{\alpha}_5-2\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2$，排除(D).

由排除法，应选(B).

对于(B)，由题设条件知 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=3$. 若 $\boldsymbol{\alpha}_4$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出，因 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3=\boldsymbol{0}$，则 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\leqslant 2$，这和 $r(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=3$ 矛盾，故(B)成立.

三、两个方程组的公共解

(1)齐次线性方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 和 $\boldsymbol{B}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 的公共解是满足方程组 $\begin{bmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$ 的解，即联立求解. 同理，可求 $\boldsymbol{Ax}=\boldsymbol{\alpha}$ 与 $\boldsymbol{Bx}=\boldsymbol{\beta}$ 的公共解. 这里对读者的计算能力提出较高要求，理论上没有什么难点.

(2)求出 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 的通解 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_s\boldsymbol{\xi}_s$，代入 $\boldsymbol{B}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$，求出 $k_i(i=1,2,\cdots,s)$之间的关系，代回 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 的通解，即得公共解.

(3)若给出 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_s$ 与 $\boldsymbol{B}_{m\times n}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$，则公共解 $\boldsymbol{\gamma}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_s\boldsymbol{\xi}_s=l_1\boldsymbol{\eta}_1+l_2\boldsymbol{\eta}_2+\cdots+l_t\boldsymbol{\eta}_t$，即

$$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_s\boldsymbol{\xi}_s-l_1\boldsymbol{\eta}_1-l_2\boldsymbol{\eta}_2-\cdots-l_t\boldsymbol{\eta}_t=\boldsymbol{0},$$

解此式，求出 k_i 或 l_j，$i=1,2,\cdots,s$，$j=1,2,\cdots,t$，即可写出 $\boldsymbol{\gamma}$.

例 2.4.11 已知线性方程组

$$(\text{Ⅰ})\begin{cases}x_1+x_2=0,\\x_2-x_4=0,\end{cases}\qquad(\text{Ⅱ})\begin{cases}x_1-x_2+x_3=0,\\x_2-x_3+x_4=0.\end{cases}$$

(1)分别求方程组(Ⅰ),(Ⅱ)的基础解系；

(2)求方程组(Ⅰ),(Ⅱ)的公共解.

解 (1)由方程组(Ⅰ)得其系数矩阵为

$$\boldsymbol{A}=\begin{bmatrix}1&1&0&0\\0&1&0&-1\end{bmatrix},$$

解得其基础解系为

$$\boldsymbol{\xi}_1=[0,0,1,0]^{\mathrm{T}},\quad\boldsymbol{\xi}_2=[-1,1,0,1]^{\mathrm{T}}.$$

同理，方程组(Ⅱ)的系数矩阵为

$$\boldsymbol{B}=\begin{bmatrix}1&-1&1&0\\0&1&-1&1\end{bmatrix},$$

其基础解系为 $\boldsymbol{\eta}_1=[0,1,1,0]^{\mathrm{T}},\quad\boldsymbol{\eta}_2=[-1,-1,0,1]^{\mathrm{T}}.$

(2)方法一 直接解(Ⅰ),(Ⅱ)的联立方程组，即求解 $\begin{bmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$. 因

$$\begin{bmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{bmatrix}=\begin{bmatrix}1&1&0&0\\0&1&0&-1\\1&-1&1&0\\0&1&-1&1\end{bmatrix}\to\begin{bmatrix}1&1&0&0\\0&1&0&-1\\0&-2&1&0\\0&0&-1&2\end{bmatrix}$$

$$\to\begin{bmatrix}1&1&0&0\\0&1&0&-1\\0&0&1&-2\\0&0&-1&2\end{bmatrix}\to\begin{bmatrix}1&1&0&0\\0&1&0&-1\\0&0&1&-2\\0&0&0&0\end{bmatrix},$$

故得方程组(Ⅰ),(Ⅱ)的公共解为 $k[-1,1,2,1]^{\mathrm{T}}$，其中 k 是任意常数.

方法二 在方程组（Ⅰ）的通解中找出满足方程组（Ⅱ）的解（或在（Ⅱ）的通解中找出满足（Ⅰ）的解），即是（Ⅰ），（Ⅱ）的公共解.

方程组（Ⅰ）的通解为 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=[-k_2,k_2,k_1,k_2]^{\mathrm{T}}$，代入方程组（Ⅱ），得

$$\begin{cases}-k_2-k_2+k_1=0,\\ k_2-k_1+k_2=0.\end{cases}$$

解得 $k_1=2k_2$，代入方程组（Ⅰ）的通解，得方程组（Ⅰ），（Ⅱ）的公共解是

$$[-k_2,k_2,2k_2,k_2]^{\mathrm{T}}=k_2[-1,1,2,1]^{\mathrm{T}},$$

其中 k_2 是任意常数.

方法三 从方程组（Ⅰ），（Ⅱ）的通解中找出公共解.

（Ⅰ）的通解为 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$，（Ⅱ）的通解为 $l_1\boldsymbol{\eta}_1+l_2\boldsymbol{\eta}_2$，则公共解应满足 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=l_1\boldsymbol{\eta}_1+l_2\boldsymbol{\eta}_2$，即

$$k_1\begin{bmatrix}0\\0\\1\\0\end{bmatrix}+k_2\begin{bmatrix}-1\\1\\0\\1\end{bmatrix}=l_1\begin{bmatrix}0\\1\\1\\0\end{bmatrix}+l_2\begin{bmatrix}-1\\-1\\0\\1\end{bmatrix}.$$

由上式可得 $k_2=l_2,k_2=l_1-l_2,k_1=l_1$. 故

$$k_2=l_1-l_2=k_1-k_2,\text{得 } k_1=2k_2,$$

或

$$k_2=l_2=l_1-l_2,\text{得 } l_1=2l_2.$$

因此，公共解为

$$2k_2\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=k_2(2\boldsymbol{\xi}_1+\boldsymbol{\xi}_2)=k_2[-1,1,2,1]^{\mathrm{T}},$$

其中 k_2 是任意常数；

或

$$2l_2\boldsymbol{\eta}_1+l_2\boldsymbol{\eta}_2=l_2(2\boldsymbol{\eta}_1+\boldsymbol{\eta}_2)=l_2[-1,1,2,1]^{\mathrm{T}},$$

其中 l_2 是任意常数.

【注】 本题除了给出两个方程组求其公共解之外，还可以转化成给出一个方程组和另一个方程组的通解（或基础解系），然后求公共解，或者给出两个方程组的通解（或基础解系），然后求公共解. 上述三种解法，均已给出了求解方法.

当然，也可将一个方程组改成满足某个（或某些）条件（满足另一个方程组就是满足某些条件）的方程组，再求解.

四、同解方程组

若两个方程组 $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\mathbf{0}$ 和 $\boldsymbol{B}_{s\times n}\boldsymbol{x}=\mathbf{0}$ 有完全相同的解，则称为同解方程组.

于是， $\boldsymbol{Ax}=\mathbf{0},\boldsymbol{Bx}=\mathbf{0}$ 是同解方程组

$\Leftrightarrow\boldsymbol{Ax}=\mathbf{0}$ 的解满足 $\boldsymbol{Bx}=\mathbf{0}$，且 $\boldsymbol{Bx}=\mathbf{0}$ 的解满足 $\boldsymbol{Ax}=\mathbf{0}$（互相把解代入求出结果即可）

$\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{B})$，且 $\boldsymbol{Ax}=\mathbf{0}$ 的解满足 $\boldsymbol{Bx}=\mathbf{0}$（或 $\boldsymbol{Bx}=\mathbf{0}$ 的解满足 $\boldsymbol{Ax}=\mathbf{0}$）

$\Leftrightarrow r(\boldsymbol{A})=r(\boldsymbol{B})=r\left(\begin{bmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{bmatrix}\right)$（三秩相同较方便）.

例 2.4.12 设线性方程组

$$(\mathrm{I})\begin{cases}x_1+3x_3+5x_4=0,\\x_1-x_2-2x_3+2x_4=0,\\2x_1-x_2+x_3+3x_4=0.\end{cases}$$

在（Ⅰ）的基础上，添加一个方程 $4x_1+ax_2+bx_3+13x_4=0$，得

$$(\mathrm{II})\begin{cases}x_1+3x_3+5x_4=0,\\x_1-x_2-2x_3+2x_4=0,\\2x_1-x_2+x_3+3x_4=0,\\4x_1+ax_2+bx_3+13x_4=0.\end{cases}$$

（1）求方程组（Ⅰ）的通解；

（2）a,b 满足什么条件时，方程组（Ⅰ），（Ⅱ）是同解方程组？

解 （1）对方程组（Ⅰ）的系数矩阵作初等行变换，即

$$\begin{bmatrix}1&0&3&5\\1&-1&-2&2\\2&-1&1&3\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&3&5\\0&-1&-5&-3\\0&-1&-5&-7\end{bmatrix}\longrightarrow\begin{bmatrix}1&0&3&5\\0&-1&-5&-3\\0&0&0&-4\end{bmatrix}.$$

自由未知量 x_3 赋值为 1，回代方程组，得基础解系 $\boldsymbol{\xi}=[-3,-5,1,0]^{\mathrm{T}}$，故方程组（Ⅰ）的通解为 $k[-3,-5,1,0]^{\mathrm{T}}$，$k$ 是任意常数.

（2）由方程组（Ⅰ），（Ⅱ）同解，知（Ⅰ）的解应满足（Ⅱ）. 因（Ⅰ），（Ⅱ）前 3 个方程相同，故只需满足第 4 个方程即可，将（Ⅰ）的通解代入（Ⅱ）的第 4 个方程，得

$$4(-3k)+a(-5k)+bk+13(0k)=0,$$

即 $(-12-5a+b)k=0$，k 是任意常数. 故 a,b 满足

$$b-5a=12.$$

当 $b-5a=12$ 时，（Ⅰ）的解满足（Ⅱ），又（Ⅱ）的前 3 个方程即是（Ⅰ），所以（Ⅱ）的解也满足（Ⅰ），方程组（Ⅰ），（Ⅱ）是同解方程组.

例 2.4.13 设方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解，证明：$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 和 $\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$ 是同解方程组.

证明 ①显然 $\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$ 的解必满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$.

②因 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解，故 $r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{b}])$，得 $r\left(\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\right)=r(\boldsymbol{A}^{\mathrm{T}})$. 故 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 和 $\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系中所含解向量个数相等.

由①，②知，方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 和 $\begin{bmatrix}\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{b}^{\mathrm{T}}\end{bmatrix}\boldsymbol{x}=\boldsymbol{0}$ 是同解方程组.

基础习题精练

习 题

2.4.1 确定常数 a，使向量组 $\boldsymbol{\alpha}_1=[1,1,a]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[1,a,1]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=[a,1,1]^{\mathrm{T}}$ 可由向量组 $\boldsymbol{\beta}_1=[1,1,a]^{\mathrm{T}}$，$\boldsymbol{\beta}_2=[-2,a,4]^{\mathrm{T}}$，$\boldsymbol{\beta}_3=[-2,a,a]^{\mathrm{T}}$ 线性表示，但向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不能由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示.

2.4.2 已知 $\boldsymbol{\alpha}_1=[1,0,1,0]^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=[2,2,a,2]^{\mathrm{T}}$，$\boldsymbol{\alpha}_3=[3,1,1,1]^{\mathrm{T}}$，$\boldsymbol{\beta}=[4,-1,6,b]^{\mathrm{T}}$，问 a,b 取何值时，$\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出？取何值时能够线性表出？当 $\boldsymbol{\beta}$ 能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出时，写出其表出式.

2.4.3 对于 n 元方程组，下列结论成立的是（　　）.

(A)若 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解，则 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解

(B)$\boldsymbol{Ax}=\boldsymbol{0}$有非零解的充要条件是$|\boldsymbol{A}|=0$

(C)$\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解的充要条件是 $r(\boldsymbol{A})=n$

(D)若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有两个不同的解，则 $\boldsymbol{Ax}=\boldsymbol{0}$有无穷多解

2.4.4 设 ξ_1,ξ_2,ξ_3 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系，则该方程组的基础解系还可以表示成（　　）.

(A)ξ_1,ξ_2,ξ_3 的一个等价向量组

(B)ξ_1,ξ_2,ξ_3 的一个等秩向量组

(C)$\xi_1+\xi_2,\xi_2+\xi_3,\xi_3+\xi_1$

(D)$\xi_1-\xi_2,\xi_2-\xi_3,\xi_3-\xi_1$

2.4.5 已知线性方程组 $\boldsymbol{A}_{4\times4}\boldsymbol{x}=\boldsymbol{0}$ 有基础解系

$$\xi_1=[1,2,-3,0]^{\mathrm{T}},\quad \xi_2=[2,-1,0,1]^{\mathrm{T}},\quad \xi_3=[1,0,-2,1]^{\mathrm{T}},$$

则下列向量是该方程组的解的是（　　）.

(A)$[-1,5,3,7]^{\mathrm{T}}$

(B)$[0,4,-9,1]^{\mathrm{T}}$

(C)$[1,2,-3,1]^{\mathrm{T}}$

(D)$[2,-1,-2,0]^{\mathrm{T}}$

2.4.6 已知 $\xi_1=[-9,1,2,11]^{\mathrm{T}}$，$\xi_2=[1,-5,13,0]^{\mathrm{T}}$，$\xi_3=[-7,-9,24,11]^{\mathrm{T}}$是方程组

$$\begin{cases}2x_1+a_2x_2+3x_3+a_4x_4=d_1,\\3x_1+b_2x_2+2x_3+b_4x_4=4,\\9x_1+4x_2+x_3+c_4x_4=d_3\end{cases}$$

的三个解，求此方程组的通解.

2.4.7 设 $\boldsymbol{A}$ 是 n 阶实矩阵，$\boldsymbol{A}^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的转置矩阵，证明：方程组(Ⅰ)$\boldsymbol{Ax}=\boldsymbol{0}$ 和(Ⅱ)$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{0}$是同解方程组.

解 答

2.4.1 解 记 $\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]$，$\boldsymbol{B}=[\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3]$. 由于 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示，所以 $r(\boldsymbol{A})<3$，从而行列式

$$|\boldsymbol{A}|=\begin{vmatrix}1&1&a\\1&a&1\\a&1&1\end{vmatrix}=-(a-1)^2(a+2)=0,$$

得 $a=1$ 或 $a=-2$.

当 $a=1$ 时,$\boldsymbol{\alpha}_1=\boldsymbol{\alpha}_2=\boldsymbol{\alpha}_3=\boldsymbol{\beta}_1=[1,1,1]^{\mathrm{T}}$,显然 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,而 $\boldsymbol{\beta}_2=[-2,1,4]^{\mathrm{T}}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示,即 $a=1$ 符合题意.

当 $a=-2$ 时,则有

$$[\boldsymbol{B}\,\vdots\,\boldsymbol{A}]=\left[\begin{array}{ccc:ccc}1&-2&-2&1&1&-2\\1&-2&-2&1&-2&1\\-2&4&-2&-2&1&1\end{array}\right]\longrightarrow\left[\begin{array}{ccc:ccc}1&-2&-2&1&1&-2\\0&0&-6&0&3&-3\\0&0&0&0&-3&3\end{array}\right].$$

考虑非齐次线性方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$,由上述可知 $r(\boldsymbol{B})=2$,而 $r([\boldsymbol{B},\boldsymbol{\alpha}_2])=3$,则方程组 $\boldsymbol{Bx}=\boldsymbol{\alpha}_2$ 无解,即 $\boldsymbol{\alpha}_2$ 不能由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,所以 $a=-2$ 不符合题意,应舍去. 综上,$a=1$.

【注】 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 可由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 线性表示,则三个方程组 $x_{i1}\boldsymbol{\beta}_1+x_{i2}\boldsymbol{\beta}_2+x_{i3}\boldsymbol{\beta}_3=\boldsymbol{\alpha}_i(i=1,2,3)$ 均有解;向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 不可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示,则三个方程组 $x_{i1}\boldsymbol{\alpha}_1+x_{i2}\boldsymbol{\alpha}_2+x_{i3}\boldsymbol{\alpha}_3=\boldsymbol{\beta}_i(i=1,2,3)$ 至少有一个无解.

2.4.2 **解** 设 $\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3$,$\boldsymbol{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]$,对该非齐次方程组的增广矩阵作初等行变换,有

$$[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\,\vdots\,\boldsymbol{\beta}]=\left[\begin{array}{ccc:c}1&2&3&4\\0&2&1&-1\\1&a&1&6\\0&2&1&b\end{array}\right]\longrightarrow\left[\begin{array}{ccc:c}1&2&3&4\\0&2&1&-1\\0&a-2&-2&2\\0&2&1&b\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{ccc:c}1&2&3&4\\0&2&1&-1\\0&a+2&0&0\\0&0&0&b+1\end{array}\right].$$

当 $b\neq-1$,a 任意时,$r(\boldsymbol{A})<r([\boldsymbol{A},\boldsymbol{\beta}])$,方程组无解,$\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出.

当 $b=-1$,$a\neq-2$ 时,$r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{\beta}])=3$,方程组有唯一解,解得 $x_1=7$,$x_2=0$,$x_3=-1$,故 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出,其表出式为 $\boldsymbol{\beta}=7\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3=7\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_3$.

当 $b=-1$,$a=-2$ 时,$r(\boldsymbol{A})=r([\boldsymbol{A},\boldsymbol{\beta}])=2$,方程组有无穷多解.取自由未知量 $x_2=k$(也可取 x_3 为自由未知量),解得 $x_1=4k+7$,$x_3=-2k-1$,此时,$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出,且表出法有无穷多种,其表出式为

$$\boldsymbol{\beta}=(4k+7)\boldsymbol{\alpha}_1+k\boldsymbol{\alpha}_2-(2k+1)\boldsymbol{\alpha}_3,$$

其中 k 是任意常数.

2.4.3 (D) **解** 注意审题,"对于 n 元方程组"这里只指出方程组未知量个数为 n,未指出方程的个数,故 $\boldsymbol{A}$ 不一定是方阵,所以,若"$\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解$\Rightarrow\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解"不一定成立,方程组可能无解,(A)不正确;"$\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充要条件是 $|\boldsymbol{A}|=0$"不一定成立,因 $\boldsymbol{A}$ 不一定是方阵,(B)不正确;"$\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解的充要条件是 $r(\boldsymbol{A})=n$",同样 $\boldsymbol{A}$ 不一定是方阵,(C)不正确. 对于(D),若 $\boldsymbol{A}\xi_1=\boldsymbol{b}$,$\boldsymbol{A}\xi_2=\boldsymbol{b}$,则必有 $\boldsymbol{A}k(\xi_1-$

$\xi_2)=0$，k 是任意常数，故选(D).

2.4.4　(C)　**解**　和 ξ_1,ξ_2,ξ_3 等价的向量组不一定线性无关（此时向量个数大于秩），和 ξ_1,ξ_2,ξ_3 等秩的向量组中的向量不一定是解向量，由 $(\xi_1-\xi_2)+(\xi_2-\xi_3)+(\xi_3-\xi_1)=\mathbf{0}$，知 $\xi_1-\xi_2,\xi_2-\xi_3,\xi_3-\xi_1$ 线性相关，故(A)，(B)，(D)均应排除. 而

$$[\xi_1+\xi_2,\xi_2+\xi_3,\xi_3+\xi_1]=[\xi_1,\xi_2,\xi_3]\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix},\ \begin{vmatrix}1&0&1\\1&1&0\\0&1&1\end{vmatrix}=2\neq0,$$

又 ξ_1,ξ_2,ξ_3 线性无关，则 $\xi_1+\xi_2,\xi_2+\xi_3,\xi_3+\xi_1$ 线性无关，且均是解向量，故选(C).

2.4.5　(B)　**解**　$\boldsymbol{Ax}=\mathbf{0}$ 的解向量可由基础解系线性表出，选项中的向量分别记为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$. 对于 $\boldsymbol{\beta}_i(i=1,2,3,4)$，若存在 x_1,x_2,x_3，使得

$$x_1\xi_1+x_2\xi_2+x_3\xi_3=\boldsymbol{\beta}_i,$$

则 $\boldsymbol{\beta}_i$ 可由 ξ_1,ξ_2,ξ_3 线性表出，即是解向量. 因

$$[\xi_1,\xi_2,\xi_3\ \vdots\ \boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4]=\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\2&-1&0&5&4&2&-1\\-3&0&-2&3&-9&-3&-2\\0&1&1&7&1&1&0\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\0&-5&-2&7&4&0&-5\\0&6&1&0&-9&0&4\\0&1&1&7&1&1&0\end{array}\right]\longrightarrow\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\0&-5&-2&7&4&0&-5\\0&1&-1&7&-5&0&-1\\0&0&2&0&6&1&1\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\0&1&-1&7&-5&0&-1\\0&0&-7&42&-21&0&-10\\0&0&2&0&6&1&1\end{array}\right]\longrightarrow\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\0&1&-1&7&-5&0&-1\\0&0&2&0&6&1&1\\0&0&-1&42&-3&3&-7\end{array}\right]$$

$$\longrightarrow\left[\begin{array}{ccc:cccc}1&2&1&-1&0&1&2\\0&1&-1&7&-5&0&-1\\0&0&-1&42&-3&3&-7\\0&0&0&84&0&7&-13\end{array}\right],$$

故 $[\xi_1,\xi_2,\xi_3]\boldsymbol{x}=\boldsymbol{\beta}_2$ 有解，其他无解，应选(B).

2.4.6　**分析**　求 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解关键是求 $\boldsymbol{Ax}=\mathbf{0}$ 的基础解系，$\xi_1-\xi_2,\xi_2-\xi_3$ 都是 $\boldsymbol{Ax}=\mathbf{0}$ 的解，现在就要判断秩 $r(\boldsymbol{A})$，以确定基础解系中解向量的个数.

解　系数矩阵 $\boldsymbol{A}$ 是 3×4 矩阵，$r(\boldsymbol{A})\leqslant3$，由于 $\boldsymbol{A}$ 中第 2，3 行不成比例，故 $r(\boldsymbol{A})\geqslant2$，又因

$$\boldsymbol{\eta}_1=\xi_1-\xi_2=[-10,6,-11,11]^{\mathrm{T}},\quad \boldsymbol{\eta}_2=\xi_2-\xi_3=[8,4,-11,-11]^{\mathrm{T}}$$

是 $\boldsymbol{Ax}=\mathbf{0}$ 的两个线性无关的解，于是 $4-r(\boldsymbol{A})\geqslant2$，因此 $r(\boldsymbol{A})=2$，所以 $\xi_1+k_1\boldsymbol{\eta}_1+k_2\boldsymbol{\eta}_2$ 是此方程组的通解，其中 k_1,k_2 为任意常数.

2.4.7　**证明**　①若存在 $\boldsymbol{x}$，使 $\boldsymbol{Ax}=\mathbf{0}$，则两端左乘 $\boldsymbol{A}^{\mathrm{T}}$，得 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\mathbf{0}$，故方程组（Ⅰ）的解必是方程组（Ⅱ）的解.

② 若存在 $\boldsymbol{x}$，使 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\mathbf{0}$，则两端左乘 $\boldsymbol{x}^{\mathrm{T}}$，得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=(\boldsymbol{Ax})^{\mathrm{T}}\boldsymbol{Ax}=0$.

因 $\boldsymbol{A}$ 是实矩阵，故 $\boldsymbol{Ax}$ 是实向量. 设 $\boldsymbol{Ax}=[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$，则 $(\boldsymbol{Ax})^{\mathrm{T}}\boldsymbol{Ax}=\sum_{i=1}^{n}a_i^2=0$，得 $a_i=0(i=1,2,\cdots,n)$，故 $\boldsymbol{Ax}=\boldsymbol{0}$，从而知方程组(Ⅱ)的解必是方程组(Ⅰ)的解.

由①，②知，方程组(Ⅰ)，(Ⅱ)是同解方程组.

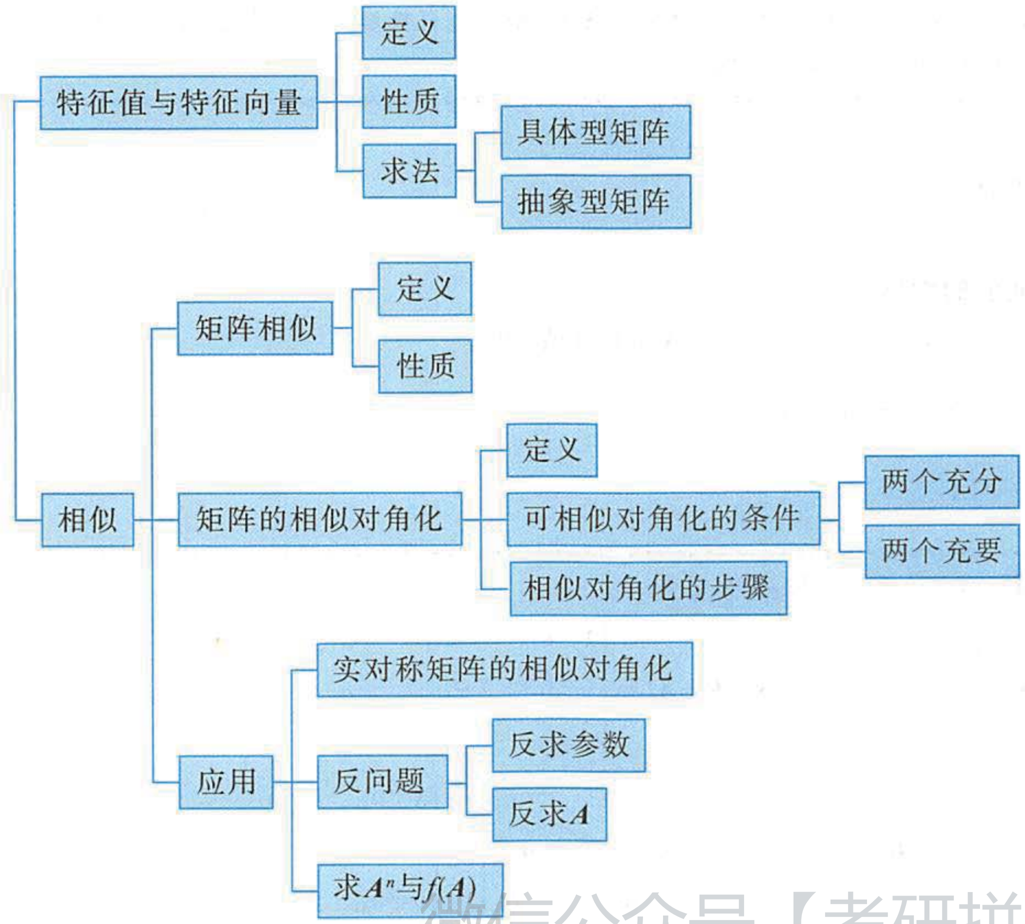
【注】 (1)若 $\boldsymbol{A}$ 是 $m\times n$ 实矩阵，则 $\boldsymbol{Ax}=\boldsymbol{0}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{0}$ 也是同解方程组.

(2)方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{0}$ 是同解方程组，从而有 $r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.

(3)因 $r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})$，故 $r(\boldsymbol{A})=r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r(\boldsymbol{AA}^{\mathrm{T}})$.

第5讲 特征值与特征向量

基础知识结构



基础内容精讲

一、基本概念

设 A 是 n 阶矩阵，λ 是一个数，若存在 n 维非零列向量 $\xi\neq 0$，使得

$$A\xi=\lambda\xi, \qquad ①$$

则称 λ 是 A 的特征值，ξ 是 A 的对应于特征值 λ 的特征向量.

【注】 由①式，得

$$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{\xi}=\boldsymbol{0},$$

因 $\boldsymbol{\xi}\neq\boldsymbol{0}$，故

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{vmatrix}=0. \quad ②$$

②式称为 $\boldsymbol{A}$ 的特征方程，是未知量 λ 的 n 次方程，有 n 个根（重根按照重数计），$\lambda \boldsymbol{E}-\boldsymbol{A}$ 称为特征矩阵，$|\lambda \boldsymbol{E}-\boldsymbol{A}|$ 称为特征多项式.

具体型矩阵的特征值与特征向量的求法见“基础例题精解”的“一”. 抽象型矩阵的特征值与特征向量的求法见“基础例题精解”的“二”.

二、基本性质

1. 特征值的性质

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，$\lambda_i(i=1,2,\cdots,n)$是 $\boldsymbol{A}$ 的特征值，则

① $\sum\limits_{i=1}^{n}\lambda_i=\sum\limits_{i=1}^{n}a_{ii}=\operatorname{tr}(\boldsymbol{A})$；

② $\prod\limits_{i=1}^{n}\lambda_i=|\boldsymbol{A}|$.

【注】 以 3 阶矩阵为例，做个证明.

设 $\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$，则 $|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda-a_{33} \end{vmatrix}$ 是 λ 的一元三次多项式，则

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda-a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda-a_{33} \end{vmatrix}=\begin{vmatrix} \lambda-a_{11} & 0-a_{12} & 0-a_{13} \\ 0-a_{21} & \lambda-a_{22} & 0-a_{23} \\ 0-a_{31} & 0-a_{32} & \lambda-a_{33} \end{vmatrix},$$

将上述行列式全部拆开，得

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix}+\begin{vmatrix} \lambda & 0 & -a_{13} \\ 0 & \lambda & -a_{23} \\ 0 & 0 & -a_{33} \end{vmatrix}+\begin{vmatrix} \lambda & -a_{12} & 0 \\ 0 & -a_{22} & 0 \\ 0 & -a_{32} & \lambda \end{vmatrix}+\begin{vmatrix} -a_{11} & 0 & 0 \\ -a_{21} & \lambda & 0 \\ -a_{31} & 0 & \lambda \end{vmatrix}+$$

$$\begin{vmatrix} \lambda & -a_{12} & -a_{13} \\ 0 & -a_{22} & -a_{23} \\ 0 & -a_{32} & -a_{33} \end{vmatrix}+\begin{vmatrix} -a_{11} & 0 & -a_{13} \\ -a_{21} & \lambda & -a_{23} \\ -a_{31} & 0 & -a_{33} \end{vmatrix}+\begin{vmatrix} -a_{11} & -a_{12} & 0 \\ -a_{21} & -a_{22} & 0 \\ -a_{31} & -a_{32} & \lambda \end{vmatrix}+\begin{vmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{vmatrix}$$

$$=\lambda^3-(a_{11}+a_{22}+a_{33})\lambda^2+(A_{11}+A_{22}+A_{33})\lambda-|\boldsymbol{A}|. \quad ①$$

设
$$|\lambda E-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$$
$$=\lambda^3-(\lambda_1+\lambda_2+\lambda_3)\lambda^2+(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3)\lambda-\lambda_1\lambda_2\lambda_3. \quad ②$$

比较①,②两式,得

$$a_{11}+a_{22}+a_{33}=\sum_{i=1}^{3}a_{ii}=\lambda_1+\lambda_2+\lambda_3=\sum_{i=1}^{3}\lambda_i,$$
$$|A|=\lambda_1\lambda_2\lambda_3.$$

在求$|A|$及A的特征值时会经常用到这两个性质.

2. 特征向量的性质

(1)k重特征值λ至多只有k个线性无关的特征向量(直接使用,不用证明).

(2)若ξ_1,ξ_2是A的属于不同特征值λ_1,λ_2的特征向量,则ξ_1,ξ_2线性无关(见例2.5.5).

(3)若ξ_1,ξ_2是A的属于同一特征值λ的特征向量,则$k_1\xi_1+k_2\xi_2$(k_1,k_2不同时为零的任意常数)仍是A的属于特征值λ的特征向量(见例2.5.7).

三、矩阵的相似

1. 定义

设A,B是两个n阶方阵,若存在n阶可逆矩阵P,使得$P^{-1}AP=B$,则称A相似于B,记成$A\sim B$.

【注】 (1)$A\sim A$;(反身性)

(2)若$A\sim B$,则$B\sim A$;(对称性)

(3)若$A\sim B,B\sim C$,则$A\sim C$.(传递性)这个性质以后常用.

2. 相似矩阵的性质

(1)若$A\sim B$,则有:①$r(A)=r(B)$;②$|A|=|B|$;③$|\lambda E-A|=|\lambda E-B|$;④$A,B$有相同的特征值.

【注】 以上结论,反之不成立.例如:若取$A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$,$B=\begin{bmatrix}1&1\\0&1\end{bmatrix}$,则①$r(A)=r(B)=2$;②$|A|=|B|=1$;③$|\lambda E-A|=|\lambda E-B|=(\lambda-1)^2$;④$\lambda_A=\lambda_B=1$(二重根),但对任意可逆矩阵$P$,$P^{-1}AP=P^{-1}EP=E\neq B$,故$A,B$不相似.

(2)若$A\sim B$,则$A^m\sim B^m$,$f(A)\sim f(B)$(其中$f(x)$是多项式)(见例2.5.13).

(3)若$A\sim B$,且A可逆,则$A^{-1}\sim B^{-1}$,$f(A^{-1})\sim f(B^{-1})$(其中$f(x)$是多项式).

【注】 证明 $A\sim B$,则存在可逆矩阵P,使得$P^{-1}AP=B$,两边取逆,有$P^{-1}A^{-1}P=B^{-1}$,故$A^{-1}\sim B^{-1}$,由(2)知$f(A^{-1})\sim f(B^{-1})$.

(4)若$A\sim B$,则$A^T\sim B^T$.

【注】 证明 $A\sim B$,则存在可逆矩阵P,使得$P^{-1}AP=B$,两边取转置,有$P^T A^T (P^{-1})^T=B^T$,即$P^T A^T (P^T)^{-1}=B^T$,故$A^T\sim B^T$.

(5)若 $A \sim B$，且 A 可逆，则 $A^* \sim B^*$.

【注】 证明 $A \sim B$，则存在可逆矩阵 P，使得 $P^{-1}AP=B$，两边取伴随运算，有 $P^* A^* (P^{-1})^* = B^*$，即 $P^* A^* (P^*)^{-1} = B^*$，故 $A^* \sim B^*$.

四、矩阵的相似对角化

1. 定义

设 n 阶矩阵 A，若存在 n 阶可逆矩阵 P，使得 $P^{-1}AP=\Lambda$，其中 Λ 是对角矩阵，则称 A 可相似对角化，记 $A \sim \Lambda$，称 Λ 是 A 的相似标准形.

2. 矩阵可相似对角化的条件

由定义可知，若 A 可相似对角化，即 $P^{-1}AP=\Lambda$，其中 P 可逆，等式两边同时左乘 P，有 $AP=P\Lambda$，记

$$P=[\xi_1,\xi_2,\cdots,\xi_n],\quad \Lambda=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

则

$$A[\xi_1,\xi_2,\cdots,\xi_n]=[\xi_1,\xi_2,\cdots,\xi_n]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

即

$$[A\xi_1,A\xi_2,\cdots,A\xi_n]=[\lambda_1\xi_1,\lambda_2\xi_2,\cdots,\lambda_n\xi_n],$$

也即

$$A\xi_i=\lambda_i\xi_i,\ i=1,2,\cdots,n.$$

由 P 可逆，则 $\xi_1,\xi_2,\cdots,\xi_n$ 线性无关. 上述过程可逆，于是

①n 阶矩阵 A 可相似对角化 $\Leftrightarrow$ A 有 n 个线性无关的特征向量.

由第 5 讲“基础内容精讲”“二”中“2”的“(1)”，k 重特征值 λ 至多只有 k 个线性无关的特征向量，得

②n 阶矩阵 A 可相似对角化 $\Leftrightarrow$ A 对应于每个 k_i 重特征值都有 k_i 个线性无关的特征向量.

由第 5 讲“基础内容精讲”“二”中“2”的“(2)”，对应于不同特征值的特征向量线性无关，得

③n 阶矩阵 A 有 n 个不同特征值 $\Rightarrow$ A 可相似对角化.

由下面的“五”可知：

④n 阶矩阵 A 为实对称矩阵 $\Rightarrow$ A 可相似对角化.

以上①，②为 A 可相似对角化的充要条件；③，④为 A 可相似对角化的充分条件.

五、实对称矩阵必可相似于对角矩阵

在第 2 讲中已经学过，若 $A^{\mathrm{T}}=A$，则 A 为对称矩阵，进一步，若组成 A 的元素都是实数，则 A 为实对称矩阵.

(1) A 是实对称矩阵，则 A 的特征值是实数，特征向量是实向量(不用证).

(2)实对称矩阵 **A** 的属于不同特征值的特征向量相互正交(证明见例 2.5.16).

(3)实对称矩阵 **A** 必相似于对角矩阵,即必有 n 个线性无关的特征向量 $\xi_1,\xi_2,\cdots,\xi_n$,即必有可逆矩阵

$$P=[\xi_1,\xi_2,\cdots,\xi_n],使得 P^{-1}AP=\Lambda,其中 \Lambda=\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix},且存在正交矩阵 Q,使得 Q^{-1}AQ=Q^{\mathrm{T}}AQ=$$

Λ,故 **A** 正交相似于 Λ(不用证,熟练掌握方法与步骤,见“基础例题精解”的“六”).

基础例题精解

一、具体型矩阵的特征值与特征向量

求解具体型矩阵的特征值与特征向量,一般用“特征方程法”,即先用特征方程 $|\lambda E-A|=0$ 求出 λ,再解齐次线性方程组 $(\lambda E-A)x=0$,求出特征向量.

例 2.5.1　求下列矩阵的特征值和特征向量.

(1)$A=\begin{bmatrix}1&1&1\\0&2&2\\0&0&3\end{bmatrix}$;　　(2)$B=\begin{bmatrix}1&0&1\\0&1&0\\0&0&1\end{bmatrix}$.

解　(1)

$$|\lambda E-A|=\begin{vmatrix}\lambda-1&-1&-1\\0&\lambda-2&-2\\0&0&\lambda-3\end{vmatrix}=(\lambda-1)(\lambda-2)(\lambda-3)=0,$$

解得 **A** 的特征值为 $\lambda_1=1,\lambda_2=2,\lambda_3=3$.

当 $\lambda_1=1$ 时,由

$$(\lambda_1E-A)x=\begin{bmatrix}0&-1&-1\\0&-1&-2\\0&0&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\mathbf{0},$$

解得基础解系为 $\xi_1=[1,0,0]^{\mathrm{T}}$,$k_1\xi_1$($k_1$ 不为零的任意常数)为对应于 $\lambda_1=1$ 的全部特征向量.

当 $\lambda_2=2$ 时,由

$$(\lambda_2E-A)x=\begin{bmatrix}1&-1&-1\\0&0&-2\\0&0&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\mathbf{0},$$

解得基础解系为 $\xi_2=[1,1,0]^{\mathrm{T}}$,$k_2\xi_2$($k_2$ 不为零的任意常数)为对应于 $\lambda_2=2$ 的全部特征向量.

当 $\lambda_3=3$ 时,由

$$(\lambda_3\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\begin{bmatrix}2&-1&-1\\0&1&-2\\0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0},$$

解得基础解系为 $\xi_3=[3,4,2]^{\mathrm{T}}$，$k_3\xi_3$（$k_3$ 不为零的任意常数）为对应于 $\lambda_3=3$ 的全部特征向量.

（2）由

$$|\lambda\boldsymbol{E}-\boldsymbol{B}|=\begin{vmatrix}\lambda-1&0&-1\\0&\lambda-1&0\\0&0&\lambda-1\end{vmatrix}=(\lambda-1)^3=0,$$

解得 $\boldsymbol{B}$ 的特征值为 $\lambda_1=\lambda_2=\lambda_3=1$（三重根）.

当 $\lambda=1$ 时，由

$$(\lambda\boldsymbol{E}-\boldsymbol{B})\boldsymbol{x}=\begin{bmatrix}0&0&-1\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0},$$

解得线性无关的基础解系 $\xi_1=[1,0,0]^{\mathrm{T}}$，$\xi_2=[0,1,0]^{\mathrm{T}}$，$k_1\xi_1+k_2\xi_2$（$k_1,k_2$ 不同时为零的任意常数）为 $\boldsymbol{B}$ 的对应于特征值 $\lambda=1$（三重根）的全部特征向量.

【注 1】 （1）上、下三角矩阵与对角矩阵的特征值就是对角线元素.

（2）题中 $\boldsymbol{B}$ 的特征值 $\lambda=1$ 是三重根，但对应的线性无关的特征向量只有两个.

【注 2】 用 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 求出全部的 λ 是起步，也是关键. 上面的例子都是用行列式的性质写成了 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，但如果有些技巧想不到，而直接展开成了 λ 的 n 次多项式 $f(\lambda)$，那又该如何求特征根呢？

例如，设 $\boldsymbol{A}=\begin{bmatrix}2&-2&0\\-2&1&-2\\0&-2&0\end{bmatrix}$，求矩阵 $\boldsymbol{A}$ 的特征多项式 $|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 及 $\boldsymbol{A}$ 的全部特征值.

解

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&2&0\\2&\lambda-1&2\\0&2&\lambda\end{vmatrix}=\begin{vmatrix}\lambda+2&2&0\\2(\lambda+2)&\lambda-1&2\\2(\lambda+2)&2&\lambda\end{vmatrix}$$

$$=(\lambda+2)\begin{vmatrix}1&2&0\\2&\lambda-1&2\\2&2&\lambda\end{vmatrix}=(\lambda+2)\begin{vmatrix}1&0&0\\2&\lambda-5&2\\2&-2&\lambda\end{vmatrix}$$

$$=(\lambda+2)\begin{vmatrix}\lambda-5&2\\-2&\lambda\end{vmatrix}=(\lambda+2)(\lambda^2-5\lambda+4)$$

$$=(\lambda+2)(\lambda-1)(\lambda-4),$$

所以 $\boldsymbol{A}$ 的特征值为 $-2,1,4$.

但以上过程可能有的读者一时想不到，而直接用展开做：

$$\begin{vmatrix}\lambda-2&2&0\\2&\lambda-1&2\\0&2&\lambda\end{vmatrix}=\lambda(\lambda-1)(\lambda-2)-4(\lambda-2)-4\lambda$$

$$=\lambda^3-3\lambda^2-6\lambda+8.$$

矩阵 **A** 的特征多项式 $f(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 求出来了，但 $f(\lambda)$ 是 λ 的 3 次多项式，求特征根有些困难. 为此介绍三个简单而有效的 k 次多项式方程

$$f(\lambda)=a_k\lambda^k+\cdots+a_1\lambda+a_0=0$$

的求根方法：

①若 $a_0=0$，则 $f(0)=0$，所以此时 0 是 $f(\lambda)=0$ 的根.

②若 $a_k+a_{k-1}+\cdots+a_1+a_0=0$，则 $f(1)=0$，所以此时 1 是 $f(\lambda)=0$ 的根.

③若 $f(\lambda)$ 的偶次项（包括 a_0）系数之和等于奇次项系数之和，则 $f(-1)=0$，所以此时 -1 是 $f(\lambda)=0$ 的根.

据此，在本例中 $f(\lambda)=\lambda^3-3\lambda^2-6\lambda+8$，因 $1+(-3)+(-6)+8=0$，所以 1 是 $f(\lambda)=0$ 的根. 从而$\lambda-1$是 $f(\lambda)$的一个因式，进一步分解 $f(\lambda)$，可用多项式的带余除法

$$\begin{array}{r@{\,}r}
 & \lambda^2-2\lambda-8 \\ \cline{2-2}
\lambda-1 \big) & \lambda^3-3\lambda^2-6\lambda+8 \\
 & \lambda^3-\lambda^2 \\ \cline{2-2}
 & -2\lambda^2-6\lambda \\
 & -2\lambda^2+2\lambda \\ \cline{2-2}
 & -8\lambda+8 \\
 & -8\lambda+8 \\ \cline{2-2}
 & 0
\end{array}$$

则 $f(\lambda)=(\lambda-1)(\lambda^2-2\lambda-8)$. 而 $\lambda^2-2\lambda-8$ 作为二次三项式的因式分解读者是熟悉的，故有

$$f(\lambda)=(\lambda-1)(\lambda+2)(\lambda-4),$$

所以 **A** 的 3 个特征值为 $-2,1,4$.

这里所说的三个求根方法常称为**试根法**. 多项式的带余除法和常数的带余除法形式上基本相同. 上述这些原理浅显易懂，很容易掌握，且很实用，再看一题.

例如，设

$$\boldsymbol{A}=\begin{bmatrix}5 & 6 & -10\\ -6 & 1 & 6\\ -2 & 3 & 0\end{bmatrix},$$

求矩阵 **A** 的特征多项式和全部特征值.

解

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-5 & -6 & 10\\ 6 & \lambda-1 & -6\\ 2 & -3 & \lambda\end{vmatrix}=\begin{vmatrix}\lambda-5 & -6 & \lambda+5\\ 6 & \lambda-1 & 0\\ 2 & -3 & \lambda+2\end{vmatrix}$$

$$=\begin{vmatrix}\lambda-9 & 0 & -\lambda+1\\ 6 & \lambda-1 & 0\\ 2 & -3 & \lambda+2\end{vmatrix}=\begin{vmatrix}\lambda-9 & 0 & -\lambda+1\\ 6 & \lambda-1 & 0\\ 0 & -\dfrac{\lambda}{3}-\dfrac{8}{3} & \lambda+2\end{vmatrix}$$

$$=(\lambda-9)(\lambda-1)(\lambda+2)+6(-\lambda+1)\left(-\frac{\lambda}{3}-\frac{8}{3}\right)$$

$$\overset{(*)}{=\!=}(\lambda-1)[(\lambda-9)(\lambda+2)+(2\lambda+16)]$$

$$=(\lambda-1)(\lambda^2-5\lambda-2)$$

$$=(\lambda-1)\left(\lambda-\frac{5+\sqrt{33}}{2}\right)\left(\lambda-\frac{5-\sqrt{33}}{2}\right).$$

所以 $\boldsymbol{A}$ 的 3 个特征值为 $1,\frac{5+\sqrt{33}}{2},\frac{5-\sqrt{33}}{2}$.

(＊)处提出公因子 $\lambda-1$ 是重要的，有了这一步下面就能将特征多项式完全分解，从而求出 $\boldsymbol{A}$ 的全部特征值. 如有读者忽略了提出公因子，或在计算 $|\lambda\boldsymbol{E}-\boldsymbol{A}|$ 时用其他方法得出

$$f(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda^3-6\lambda^2+3\lambda+2,$$

可利用上例介绍的试根法求根. 因

$$1+(-6)+3+2=0,$$

故 1 是 $f(\lambda)=0$ 的一个根，从而 $\lambda-1$ 是 $f(\lambda)$ 的一个因式. 再用多项式的带余除法

$$\begin{array}{r} \lambda^2-5\lambda-2 \\ \lambda-1\sqrt{\lambda^3-6\lambda^2+3\lambda+2} \\ \underline{\lambda^3-\lambda^2} \\ -5\lambda^2+3\lambda \\ \underline{-5\lambda^2+5\lambda} \\ -2\lambda+2 \\ \underline{-2\lambda+2} \\ 0 \end{array}$$

可得

$$f(\lambda)=(\lambda-1)(\lambda^2-5\lambda-2),$$

同样可求得 $\boldsymbol{A}$ 的 3 个特征值为 $1,\frac{5+\sqrt{33}}{2},\frac{5-\sqrt{33}}{2}$.

其实，还有一个定理可供读者试根：

定理 设

$$f(x)=1\cdot x^k+a_{k-1}x^{k-1}+\cdots+a_1x+a_0$$

是系数 $a_i(i=0,1,2,\cdots,k-1)$ 都是整数的多项式，则 $f(x)=0$ 的有理根都是整数，且均是 a_0 的因子.

例如，设 3 阶方阵 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda\boldsymbol{E}-\boldsymbol{A}|=\lambda^3-4\lambda^2+3\lambda+2,$$

求矩阵 $\boldsymbol{A}$ 的特征值.

解 据上述定理，$f(\lambda)=0$ 的有理根只可能是 $\pm1,\pm2$. 计算得

$$f(1)=2\neq0,\quad f(-1)=-6\neq0,$$
$$f(2)=0,\quad f(-2)=-28\neq0,$$

故 $f(\lambda)=0$ 的有理根只有一个，即 $\lambda_1=2$. 从而 $\lambda-2$ 是 $f(\lambda)$ 的一个因式. 用多项式的带余除法可得

$$f(\lambda)=(\lambda-2)(\lambda^2-2\lambda-1).$$

对二次多项式方程 $\lambda^2-2\lambda-1=0$ 用求根公式得出两个根为

$$\lambda_2=1+\sqrt{2},\quad \lambda_3=1-\sqrt{2}.$$

综上可知 $\boldsymbol{A}$ 的 3 个特征值为 $2,1+\sqrt{2},1-\sqrt{2}$.

二、抽象型矩阵的特征值与特征向量

抽象型矩阵 $\boldsymbol{A}$ 的特征值和特征向量的问题主要有四点.

(1)利用定义，即若满足关系式

$$\boldsymbol{A}\boldsymbol{\xi}=\lambda\boldsymbol{\xi},\boldsymbol{\xi}\neq\boldsymbol{0},$$

则 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{\xi}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量.

(2)由定义推导得出 $\boldsymbol{A}$ 的特征方程，即若 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 成立，则 λ 是 $\boldsymbol{A}$ 的特征值，且若有 $(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{\xi}=\boldsymbol{0}$ $(\boldsymbol{\xi}\neq\boldsymbol{0})$，则 $\boldsymbol{\xi}$ 是 $\boldsymbol{A}$ 对应于 λ 的特征向量. 这与上一部分例题的理论一样.

(3) $|\boldsymbol{A}|=\lambda_1\lambda_2\cdots\lambda_n$.

(4)特征向量的性质.

例 2.5.2 设 n 阶矩阵 $\boldsymbol{A}$ 有特征值 λ，对应的特征向量为 $\boldsymbol{\xi}$. 求 $k\boldsymbol{A},\boldsymbol{A}^2,\boldsymbol{A}^k,f(\boldsymbol{A})$ 的特征值和特征向量，其中 $f(x)=a_0+a_1x+\cdots+a_nx^n$.

解 利用定义，由题设条件，有 $\boldsymbol{\xi}\neq\boldsymbol{0}$，

$$\boldsymbol{A}\boldsymbol{\xi}=\lambda\boldsymbol{\xi}. \qquad ①$$

式①两端乘 k，得

$$k\boldsymbol{A}\boldsymbol{\xi}=k\lambda\boldsymbol{\xi},\quad (k\boldsymbol{A})\boldsymbol{\xi}=(k\lambda)\boldsymbol{\xi}, \qquad ②$$

故由式②知，$k\boldsymbol{A}$ 有特征值 $k\lambda$，特征向量仍是 $\boldsymbol{\xi}$.

式①两端左乘 $\boldsymbol{A}$，并代入式①，得

$$\boldsymbol{A}^2\boldsymbol{\xi}=\lambda\boldsymbol{A}\boldsymbol{\xi}=\lambda^2\boldsymbol{\xi}, \qquad ③$$

由式③知，$\boldsymbol{A}^2$ 有特征值 λ^2，特征向量仍是 $\boldsymbol{\xi}$.

式①两端连着左乘 $k-1$ 个 $\boldsymbol{A}$，并逐个代入式①，得

$$\boldsymbol{A}^k\boldsymbol{\xi}=\lambda\boldsymbol{A}^{k-1}\boldsymbol{\xi}=\lambda^2\boldsymbol{A}^{k-2}\boldsymbol{\xi}=\cdots=\lambda^k\boldsymbol{\xi}, \qquad ④$$

由式④知，$\boldsymbol{A}^k$ 有特征值 λ^k，特征向量仍是 $\boldsymbol{\xi}$.

由式①，②，③，④得

$$\begin{aligned}f(\boldsymbol{A})\boldsymbol{\xi}&=(a_0\boldsymbol{E}+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+\cdots+a_n\boldsymbol{A}^n)\boldsymbol{\xi}\\&=a_0\boldsymbol{\xi}+a_1\boldsymbol{A}\boldsymbol{\xi}+a_2\boldsymbol{A}^2\boldsymbol{\xi}+\cdots+a_n\boldsymbol{A}^n\boldsymbol{\xi}\\&=a_0\boldsymbol{\xi}+a_1\lambda\boldsymbol{\xi}+a_2\lambda^2\boldsymbol{\xi}+\cdots+a_n\lambda^n\boldsymbol{\xi}=f(\lambda)\boldsymbol{\xi},\end{aligned} \qquad ⑤$$

由式⑤知，$f(\boldsymbol{A})$ 有特征值 $f(\lambda)$，特征向量仍是 $\boldsymbol{\xi}$.

例 2.5.3 设 n 阶可逆矩阵 $\boldsymbol{A}$ 有特征值 λ，对应的特征向量为 $\boldsymbol{\xi}$.

(1)证明 $\lambda\neq0$；

(2)求 $\boldsymbol{A}^{-1},\boldsymbol{A}^*,\boldsymbol{E}-\boldsymbol{A}^{-1}$ 的特征值和特征向量.

(1)**证明** 若 $\lambda=0$，则 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=|-\boldsymbol{A}|=0$，得 $|\boldsymbol{A}|=0$，这与矩阵 $\boldsymbol{A}$ 可逆矛盾，故 $\lambda\neq0$.

(2)**解** 由题设条件知，

$$\boldsymbol{A}\boldsymbol{\xi}=\lambda\boldsymbol{\xi},\boldsymbol{\xi}\neq\boldsymbol{0}. \qquad ①$$

式①两端左乘 $\boldsymbol{A}^{-1}$，且知 $\lambda\neq0$，得

$$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{\xi}=\boldsymbol{\xi}=\lambda\boldsymbol{A}^{-1}\boldsymbol{\xi},$$

$$A^{-1}\xi=\frac{1}{\lambda}\xi,\qquad ②$$

故 A^{-1} 有特征值 $\frac{1}{\lambda}$，特征向量仍是 ξ.

式①两端左乘 A^*，得

$$A^*A\xi=|A|E\xi=\lambda A^*\xi,$$

$$A^*\xi=\frac{|A|}{\lambda}\xi,\qquad ③$$

故 A^* 有特征值 $\frac{|A|}{\lambda}$，特征向量仍是 ξ.（若 A 有特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，因 $|A|=\prod_{i=1}^{n}\lambda_i$，故 A^* 的特征值为 $\lambda_2\lambda_3\cdots\lambda_n,\lambda_1\lambda_3\cdots\lambda_n,\cdots,\lambda_1\lambda_2\cdots\lambda_{n-1}$）

同理，

$$(E-A^{-1})\xi=\xi-A^{-1}\xi=\xi-\frac{1}{\lambda}\xi=\left(1-\frac{1}{\lambda}\right)\xi,\qquad ④$$

故 $E-A^{-1}$ 有特征值 $1-\frac{1}{\lambda}$，特征向量仍是 ξ.

【注】 (1)A 以及与 A 有关的常用矩阵的特征值和特征向量，总结如下：

矩阵	A	kA	A^k	$f(A)$	A^{-1}	A^*	$P^{-1}AP$
特征值	λ	$k\lambda$	λ^k	$f(\lambda)$	$\frac{1}{\lambda}$	$\frac{\lvert A\rvert}{\lambda}$	λ
对应的特征向量	ξ	ξ	ξ	ξ	ξ	ξ	$P^{-1}\xi$

表中 λ 在分母上，设 $\lambda\neq0$.

(2)$f(x)$ 为多项式，若矩阵 A 满足 $f(A)=O$，λ 是 A 的任一特征值，则 λ 满足 $f(\lambda)=0$.

(3)A^{T} 的特征值与 A 相同，但特征向量不再是 ξ，要单独计算才能得出.

例 2.5.4 设 A 是 n 阶矩阵，且满足 $A^2=A$（此时 A 称为幂等矩阵）.

(1)求 A 的特征值可能的取值；

(2)证明：$E+A$ 是可逆矩阵.

(1)**解** **方法一** 用定义法. 设 A 有特征值 λ，其对应的特征向量是 $\xi(\xi\neq0)$，则

$$A\xi=\lambda\xi,$$

上式两端左乘 A，得 $A^2\xi=\lambda A\xi=\lambda^2\xi$. 又 $A^2\xi=A\xi=\lambda\xi$，有 $\lambda^2\xi=\lambda\xi$，即 $(\lambda^2-\lambda)\xi=0$，而 $\xi\neq0$，故 $\lambda^2-\lambda=\lambda(\lambda-1)=0$，得 A 的特征值可能的取值是 $\lambda=0$ 或 $\lambda=1$.

方法二 由题设条件 $A^2=A$，有 $A^2-A=O$. 设 A 有特征值为 λ，则由例 2.5.3 的“注”知 A^2-A 有特征值 $\lambda^2-\lambda$. 但 $A^2-A=O$ 是零矩阵，其特征值为零，故有

$$\lambda^2-\lambda=0，得 \lambda=0 或 \lambda=1.$$

(2)**证明** 由(1)知，满足 $A^2=A$ 的矩阵 A 的特征值可能的取值是 $\lambda=0$ 或 $\lambda=1$，则 $E+A$ 的特征值可能的取值是 $\lambda=1$ 或 $\lambda=2$，均不为 0，故 $|E+A|\neq0$（或假设 $|E+A|=0$，则 -1 是 A 的特征值，这与(1)中结论矛盾），故 $E+A$ 是可逆矩阵.

【注】 满足条件 $A^2=A$ 的矩阵有很多，例如，$A=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ 或 $\begin{bmatrix}1&0\\0&0\end{bmatrix}$ 或 $\begin{bmatrix}0&0\\0&0\end{bmatrix}$，均有 $A^2=A$，而它们的特征值分别是 1,1;1,0;0,0.

例 2.5.5 证明：n 阶方阵 $\boldsymbol{A}$ 的任意两个不同特征值 λ_1,λ_2 对应的两个特征向量线性无关.

证明 设 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2(\lambda_1\neq\lambda_2)$ 对应的特征向量分别是 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2(\boldsymbol{\xi}_i\neq\boldsymbol{0},i=1,2)$，即

$$\boldsymbol{A\xi}_1=\lambda_1\boldsymbol{\xi}_1,\quad \boldsymbol{A\xi}_2=\lambda_2\boldsymbol{\xi}_2.$$

考查

$$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=\boldsymbol{0}. \qquad ①$$

式①两端左乘 $\boldsymbol{A}$，得

$$\boldsymbol{A}(k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2)=k_1\boldsymbol{A\xi}_1+k_2\boldsymbol{A\xi}_2=k_1\lambda_1\boldsymbol{\xi}_1+k_2\lambda_2\boldsymbol{\xi}_2=\boldsymbol{0}. \qquad ②$$

式①两边乘 λ_1，得

$$k_1\lambda_1\boldsymbol{\xi}_1+k_2\lambda_1\boldsymbol{\xi}_2=\boldsymbol{0}. \qquad ③$$

式③减式②，得 $k_2(\lambda_1-\lambda_2)\boldsymbol{\xi}_2=\boldsymbol{0}$. 因 $\lambda_1\neq\lambda_2,\boldsymbol{\xi}_2\neq\boldsymbol{0}$，故得 $k_2=0$.

将 $k_2=0$ 代入式①，因 $\boldsymbol{\xi}_1\neq\boldsymbol{0}$，得 $k_1=0$，故不同特征值 λ_1,λ_2 对应的特征向量线性无关.

例 2.5.6 设 λ_1,λ_2 是 $\boldsymbol{A}$ 的两个不同的特征值，$\boldsymbol{\xi}$ 是对应于 λ_1 的特征向量，证明：$\boldsymbol{\xi}$ 不是 λ_2 的特征向量(即一个特征向量不能属于两个不同的特征值).

证明 用反证法. 由题设知 $\boldsymbol{A\xi}=\lambda_1\boldsymbol{\xi}$. 若 $\boldsymbol{\xi}$ 也是 $\boldsymbol{A}$ 的对应于 λ_2 的特征向量，则有

$$\boldsymbol{A\xi}=\lambda_2\boldsymbol{\xi},$$

即得 $\lambda_1\boldsymbol{\xi}=\lambda_2\boldsymbol{\xi}$，即 $(\lambda_1-\lambda_2)\boldsymbol{\xi}=\boldsymbol{0}$. 由于 $\boldsymbol{\xi}\neq\boldsymbol{0}$，故 $\lambda_1=\lambda_2$，这与已知 $\lambda_1\neq\lambda_2$ 矛盾，故一个特征向量不能属于两个不同的特征值.

例 2.5.7 已知 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 是 $\boldsymbol{A}$ 的对应于 λ 的特征向量，问 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2(k_1,k_2$ 是任意常数)是否属于 $\boldsymbol{A}$ 的对应于 λ 的特征向量?

解 当 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2=\boldsymbol{0}$ 时，$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ 不是 $\boldsymbol{A}$ 的特征向量(特征向量是非零向量).

当 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2\neq\boldsymbol{0}$ 时，$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ 是 $\boldsymbol{A}$ 的对应于 λ 的特征向量. 由题设知，

$$\boldsymbol{A\xi}_1=\lambda\boldsymbol{\xi}_1,\quad \boldsymbol{A\xi}_2=\lambda\boldsymbol{\xi}_2,$$

则

$$\boldsymbol{A}k_1\boldsymbol{\xi}_1=\lambda k_1\boldsymbol{\xi}_1,\quad \boldsymbol{A}k_2\boldsymbol{\xi}_2=\lambda k_2\boldsymbol{\xi}_2,$$

$$\boldsymbol{A}(k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2)=\lambda(k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2),$$

故 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ 仍是 $\boldsymbol{A}$ 的对应于 λ 的特征向量.

或由题设知，$\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 均是齐次线性方程组 $(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解，故 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ 仍是上述方程组的解，即当 $k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2\neq\boldsymbol{0}$ 时，$k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2$ 仍是 $\boldsymbol{A}$ 的对应于 λ 的特征向量.

例 2.5.8 已知 $\boldsymbol{A}$ 是 3 阶方阵，特征值为 1,2,3，则 $|\boldsymbol{A}|$ 的元素 a_{11},a_{22},a_{33} 的代数余子式 A_{11},A_{22},A_{33} 的和 $\sum\limits_{i=1}^{3}A_{ii}=$ ________.

解 应填 11.

3 阶矩阵 $\boldsymbol{A}$ 的特征值 $\lambda=1,2,3$，故 $|\boldsymbol{A}|=6$. 由于代数余子式 $A_{ii}(i=1,2,3)$ 与 $\boldsymbol{A}^*$ 有关，是 $\boldsymbol{A}^*$ 的对角元素，即 $\sum\limits_{i=1}^{3}A_{ii}=\operatorname{tr}(\boldsymbol{A}^*)$，又 $\boldsymbol{A}^*$ 有特征值 $\lambda^*=\dfrac{|\boldsymbol{A}|}{\lambda}$，即 6,3,2，故 $\sum\limits_{i=1}^{3}A_{ii}=\operatorname{tr}(\boldsymbol{A}^*)=\sum\limits_{i=1}^{3}\lambda_i^*=6+3+2=11$.

例 2.5.9 设 $\boldsymbol{A}$ 是 5 阶方阵，满足 $\boldsymbol{A}^5=\boldsymbol{O}$. 则 $|\boldsymbol{A}+3\boldsymbol{E}|=$ ________.

解 应填 3^5.

设 $\boldsymbol{A}$ 的特征值为 λ，$\boldsymbol{A}^5$ 的特征值为 λ^5，由已知 $\boldsymbol{A}^5=\boldsymbol{O}$，得 $\boldsymbol{A}$ 的特征值 $\lambda=0$(5 重根). $\boldsymbol{A}+3\boldsymbol{E}$ 的特征值为 $\lambda=3$(5 重根)，故 $|\boldsymbol{A}+3\boldsymbol{E}|=3^5$.

三、矩阵能否相似对角化的判别与证明

用好上面讲的矩阵可相似对角化的两个充要条件与两个充分条件.

例 2.5.10 下列矩阵中，不能相似于对角矩阵的是(　　).

(A)$\boldsymbol{A}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$　(B)$\boldsymbol{B}=\begin{bmatrix}1&1&1\\0&2&2\\0&0&3\end{bmatrix}$　(C)$\boldsymbol{C}=\begin{bmatrix}1&-2&1\\2&-4&2\\1&-2&1\end{bmatrix}$　(D)$\boldsymbol{D}=\begin{bmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{bmatrix}$

解　应选(D).

$\boldsymbol{A}$ 是实对称矩阵，必相似于对角矩阵. $\boldsymbol{B}$ 有三个不同的特征值 $\lambda_1=1,\lambda_2=2,\lambda_3=3$，能相似于对角矩阵. $r(\boldsymbol{C})=1$，$\lambda=0$ 是 2 重特征值，有两个线性无关的特征向量，另一个特征值 $\lambda_3=-2$，故 $\boldsymbol{C}$ 能相似于对角矩阵. 由排除法，应选(D).

对(D)，
$$|\lambda\boldsymbol{E}-\boldsymbol{D}|=\begin{vmatrix}\lambda-2&1&-2\\-5&\lambda+3&-3\\1&0&\lambda+2\end{vmatrix}\xlongequal{[1]+[2]-[3]}\begin{vmatrix}\lambda+1&1&-2\\\lambda+1&\lambda+3&-3\\-\lambda-1&0&\lambda+2\end{vmatrix}$$
$$=(\lambda+1)\begin{vmatrix}1&1&-2\\0&\lambda+2&-1\\0&1&\lambda\end{vmatrix}=(\lambda+1)^3=0.$$

知 $\lambda=-1$ 是 3 重特征值，但 $-\boldsymbol{E}-\boldsymbol{D}\neq\boldsymbol{O}$，线性无关的特征向量最多有 2 个(实际上只有一个)，故 $\boldsymbol{D}$ 不能相似于对角矩阵.

【注】 判别一个矩阵是否相似于对角矩阵的步骤如下.

(1)是否是实对称矩阵，实对称矩阵必相似于对角矩阵(如(A)).

(2)特征值是否都是实单根，若是，则相似于对角矩阵(如(B)).

(3)特征值是 r 重根，若对应有 r 个线性无关的特征向量，则相似于对角矩阵(如(C))；若对应的线性无关的特征向量少于 r 个(不可能多于 r 个)，则不能相似于对角矩阵(如(D)).

例 2.5.11 设 $\boldsymbol{A}$ 是 3 阶矩阵，已知 $|\boldsymbol{E}+\boldsymbol{A}|=0$，$(3\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 有非零解，$\boldsymbol{E}-3\boldsymbol{A}$ 不可逆，问 $\boldsymbol{A}$ 是否相似于对角矩阵，说明理由.

解　$\boldsymbol{A}$ 相似于对角矩阵.

由题意，知 $\boldsymbol{E}+\boldsymbol{A},3\boldsymbol{E}-\boldsymbol{A},\boldsymbol{E}-3\boldsymbol{A}$ 均不可逆，即
$$|\boldsymbol{E}+\boldsymbol{A}|=0,\quad |3\boldsymbol{E}-\boldsymbol{A}|=0,\quad |\boldsymbol{E}-3\boldsymbol{A}|=0,$$
改写成 $(-1)^3|-\boldsymbol{E}-\boldsymbol{A}|=0$，$|3\boldsymbol{E}-\boldsymbol{A}|=0$，$3^3\left|\frac{1}{3}\boldsymbol{E}-\boldsymbol{A}\right|=0$，因此可知，$\boldsymbol{A}$ 有 3 个不同的特征值：$-1,3,\frac{1}{3}$. 故 $\boldsymbol{A}$ 有 3 个线性无关的特征向量，故 $\boldsymbol{A}$ 相似于对角矩阵，即
$$\boldsymbol{A}\sim\begin{bmatrix}-1&&\\&3&\\&&\frac{1}{3}\end{bmatrix}.$$

【注】 $|E+A|=0,(3E-A)x=0$ 有非零解，$E-3A$ 不可逆，都在告诉你 A 的特征值.

四、两个矩阵是否相似的判别与证明

（1）**定义法** 若存在可逆矩阵 P，使得 $P^{-1}AP=B$，则 $A\sim B$.

（2）**性质** 若 $A\sim B$，则 $r(A)=r(B)$，$|A|=|B|$，$\operatorname{tr}(A)=\operatorname{tr}(B)$，$\lambda_A=\lambda_B$.

（3）**传递性** 若 $A\sim\Lambda$，$\Lambda\sim B$，则 $A\sim B$.

例 2.5.12 设 $W=\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}$，则与 W 相似的矩阵是（ ）.

(A) $A=\begin{bmatrix}1&0&1\\0&2&0\\0&0&0\end{bmatrix}$ (B) $B=\begin{bmatrix}1&0&0\\0&0&2\\0&3&0\end{bmatrix}$ (C) $C=\begin{bmatrix}1&1&1\\a&a&a\\-a&-a&-a\end{bmatrix}$ (D) $D=\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix}$

解 应选（D）.

$r(W)=1\neq r(A)=2$，W 与 A 不相似，排除（A）；

$|W|=0\neq|B|=-6$，W 与 B 不相似，排除（B）；

$\operatorname{tr}(W)=6\neq\operatorname{tr}(C)=1$，$W$ 与 C 不相似，排除（C）.

由排除法，应选（D）.

对于（D）：$r(W)=r(D)=1$，$\lambda=0$ 均为 2 重特征值，且均有两个线性无关的特征向量，且 $\lambda_3=\sum\limits_{i=1}^{3}w_{ii}=\sum\limits_{i=1}^{3}d_{ii}=6$，故 $W\sim D$，即

$$\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}\sim\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix},$$

应选（D）.

【注】 选项（A），（B），（C）均不满足与 W 相似的必要条件，应予排除，利用必要条件来排除不成立的结论是选择题常用的办法.

例 2.5.13 设 n 阶方阵 A 与 B 相似，证明：

（1）对任意的正整数 k，都有 A^k 与 B^k 相似；

（2）对任意一个多项式 $f(x)=a_kx^k+a_{k-1}x^{k-1}+\cdots+a_1x+a_0$，矩阵多项式 $f(A)$ 和 $f(B)$ 相似；

（3）当 A,B 都是可逆矩阵时，A^* 和 B^* 相似.

证明 （1）因 A,B 相似，故有可逆矩阵 P，使得 $P^{-1}AP=B$.

由
$$B^k=\overbrace{BB\cdots B}^{k个}=\overbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)}^{k个}$$

$$=P^{-1}A(PP^{-1})A(PP^{-1})\cdots AP=P^{-1}A^kP,$$

知 $A^k\sim B^k$.

（2）由于
$$P^{-1}(a_0E)P=a_0E,$$
并且由（1）可知
$$P^{-1}(a_1A)P=a_1B,\quad \cdots,\quad P^{-1}(a_kA^k)P=a_kB^k.$$

等式两边全部相加，得
$$P^{-1}(a_kA^k+\cdots+a_1A+a_0E)P=a_kB^k+\cdots+a_1B+a_0E,$$
即 $P^{-1}f(A)P=f(B)$，故 $f(A)\sim f(B)$.

（3）因 A,B 相似，故有可逆矩阵 P 使 $P^{-1}AP=B$，且 $|A|=|B|$，两边求逆，得
$$P^{-1}A^{-1}P=B^{-1},$$
上式两边乘 $|A|$，得
$$P^{-1}(|A|A^{-1})P=|A|B^{-1}=|B|B^{-1}.$$

因 $|A|A^{-1}=A^*$，$|B|B^{-1}=B^*$，故 $P^{-1}A^*P=B^*$，于是 $A^*\sim B^*$.

【注】 若 $A\sim B$，则 $kA\sim kB$，$A^k\sim B^k$，$f(A)\sim f(B)$，又 A,B 可逆时，$A^{-1}\sim B^{-1}$，$A^*\sim B^*$，且 $f(A^{-1})\sim f(B^{-1})$，$f(A^*)\sim f(B^*)$.

例 2.5.14 （1）A,B 是 n 阶方阵，且 A 是实对称矩阵. 证明 A 相似于 B 的充分必要条件是 A,B 相似于同一个对角矩阵 Λ；

（2）设 $A=\begin{bmatrix}0&-1&4\\-1&3&-1\\4&-1&0\end{bmatrix}$，$B=\begin{bmatrix}5&0&0\\-5&-1&3\\4&3&-1\end{bmatrix}$，问 A,B 是否相似，说明理由.

（1）**证明** 必要性. 因 A 是实对称矩阵，所以存在可逆矩阵 P，使 $P^{-1}AP=\Lambda$，Λ 是对角矩阵，即 $A\sim\Lambda$. 若 A 相似于 B，则由相似的传递性知，$B\sim\Lambda$. 故 A,B 相似于同一个对角矩阵 Λ.

充分性. 若 A,B 相似于同一个对角矩阵，设为 $A\sim\Lambda$，$B\sim\Lambda$，即 $A\sim\Lambda\sim B$. 故 $A\sim B$.

（2）**解** 因
$$A=\begin{bmatrix}0&-1&4\\-1&3&-1\\4&-1&0\end{bmatrix},$$
有
$$|\lambda E-A|=\begin{vmatrix}\lambda&1&-4\\1&\lambda-3&1\\-4&1&\lambda\end{vmatrix}=\begin{vmatrix}\lambda+4&1&-4\\0&\lambda-3&1\\-\lambda-4&1&\lambda\end{vmatrix}=\begin{vmatrix}\lambda+4&1&-4\\0&\lambda-3&1\\0&2&\lambda-4\end{vmatrix}$$
$$=(\lambda+4)(\lambda^2-7\lambda+10)=(\lambda+4)(\lambda-2)(\lambda-5)=0.$$

故 A 有 3 个不同的特征值 $\lambda_1=-4$，$\lambda_2=2$，$\lambda_3=5$，故 $A\sim\Lambda=\begin{bmatrix}-4&0&0\\0&2&0\\0&0&5\end{bmatrix}$.

又
$$B=\begin{bmatrix}5&0&0\\-5&-1&3\\4&3&-1\end{bmatrix},$$
有
$$|\lambda E-B|=\begin{vmatrix}\lambda-5&0&0\\5&\lambda+1&-3\\-4&-3&\lambda+1\end{vmatrix}=(\lambda-5)(\lambda^2+2\lambda-8)=(\lambda-5)(\lambda+4)(\lambda-2)=0.$$

所以 B 有 3 个不同的特征值 $\lambda_1=-4,\lambda_2=2,\lambda_3=5$，故

$$B\sim\Lambda=\begin{bmatrix}-4&0&0\\0&2&0\\0&0&5\end{bmatrix}.$$

综上知 $A\sim B$.

【注】 特征值相同是两个矩阵相似的必要条件，如 $\begin{bmatrix}1&0\\0&1\end{bmatrix}$ 与 $\begin{bmatrix}1&0\\0&2\end{bmatrix}$ 的特征值不同肯定不相似，但特征值相同的矩阵也不一定相似，如 $\begin{bmatrix}1&1\\0&1\end{bmatrix}$ 与 $\begin{bmatrix}1&0\\0&1\end{bmatrix}$ 特征值相同，但并不相似；当 A,B 均是实对称矩阵时，A,B 相似的充分必要条件是 A,B 有相同的特征值.

五、求可逆矩阵 P，使得 $P^{-1}AP=\Lambda$

在已知 A 可相似对角化的条件下，其基本步骤如下.

(1)求 A 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$.

(2)求 A 的对应于特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的线性无关的特征向量 $\xi_1,\xi_2,\cdots,\xi_n$.

(3)令 $P=[\xi_1,\xi_2,\cdots,\xi_n]$，则 $P^{-1}AP=\Lambda=\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix}$.

例 2.5.15　设 A 为 3 阶矩阵，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是线性无关的 3 维列向量，且满足

$$A\boldsymbol{\alpha}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\quad A\boldsymbol{\alpha}_2=2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\quad A\boldsymbol{\alpha}_3=2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3.$$

(1)求矩阵 B，使得 $A[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]B$；

(2)求矩阵 A 的特征值；

(3)求可逆矩阵 P，使得 $P^{-1}AP$ 为对角矩阵.

解　(1)将题设条件合并成矩阵形式有

$$A[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]=[A\boldsymbol{\alpha}_1,A\boldsymbol{\alpha}_2,A\boldsymbol{\alpha}_3]=[\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3]$$

$$=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}1&0&0\\1&2&2\\1&1&3\end{bmatrix},$$

故

$$B=\begin{bmatrix}1&0&0\\1&2&2\\1&1&3\end{bmatrix}.$$

(2)因 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，故 $C=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]$ 是可逆矩阵，且有 $C^{-1}AC=B$，即 A 和 B 相似，相似矩阵有

相同的特征值. 因

$$|\lambda\boldsymbol{E}-\boldsymbol{B}|=\begin{vmatrix}\lambda-1 & 0 & 0\\ -1 & \lambda-2 & -2\\ -1 & -1 & \lambda-3\end{vmatrix}=(\lambda-1)(\lambda^2-5\lambda+4)=(\lambda-1)^2(\lambda-4)=0,$$

故 $\boldsymbol{B}$ 的特征值为 1,1,4,所以 $\boldsymbol{A}$ 的特征值也为 1,1,4.

(3)对于矩阵 $\boldsymbol{B}$,当 $\lambda_1=\lambda_2=1$ 时,由 $(\boldsymbol{E}-\boldsymbol{B})\boldsymbol{x}=\boldsymbol{0}$,即

$$\begin{bmatrix}0 & 0 & 0\\ -1 & -1 & -2\\ -1 & -1 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

得基础解系 $\boldsymbol{\xi}_1=[1,-1,0]^{\mathrm{T}}$,$\boldsymbol{\xi}_2=[-2,0,1]^{\mathrm{T}}$;

当 $\lambda_3=4$ 时,由 $(4\boldsymbol{E}-\boldsymbol{B})\boldsymbol{x}=\boldsymbol{0}$,即

$$\begin{bmatrix}3 & 0 & 0\\ -1 & 2 & -2\\ -1 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

得基础解系 $\boldsymbol{\xi}_3=[0,1,1]^{\mathrm{T}}$.

令 $\boldsymbol{Q}=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3]=\begin{bmatrix}1 & -2 & 0\\ -1 & 0 & 1\\ 0 & 1 & 1\end{bmatrix}$,有 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{\Lambda}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{bmatrix}$,故有

$$\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}\boldsymbol{Q}=\boldsymbol{\Lambda}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{bmatrix}.$$

取 $\boldsymbol{P}=\boldsymbol{C}\boldsymbol{Q}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]\begin{bmatrix}1 & -2 & 0\\ -1 & 0 & 1\\ 0 & 1 & 1\end{bmatrix}=[\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2,-2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3]$,则有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{bmatrix}.$$

六、实对称矩阵的相似对角化

若 $\boldsymbol{A}$ 为实对称矩阵,则其用正交矩阵 $\boldsymbol{Q}$ 相似对角化的基本步骤如下.

(1)求 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$.

(2)求 $\boldsymbol{A}$ 的对应于特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的特征向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$.

(3)将 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n$ 正交化(若需要的话)、单位化为 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$.

(4)令 $\boldsymbol{Q}=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n]$,则 $\boldsymbol{Q}$ 为正交矩阵,且 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$.

例 2.5.16 证明：实对称矩阵 $\boldsymbol{A}$ 对应于不同特征值的特征向量是正交的.

证明 设 $\boldsymbol{A}\boldsymbol{x}_i=\lambda_i\boldsymbol{x}_i(\boldsymbol{x}_i\neq\boldsymbol{0},i=1,2)$，$\lambda_1\neq\lambda_2$，$\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，则

$$\begin{aligned}\lambda_1\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1&=\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_1=(\boldsymbol{A}\boldsymbol{x}_2)^{\mathrm{T}}\boldsymbol{x}_1\\&=(\lambda_2\boldsymbol{x}_2)^{\mathrm{T}}\boldsymbol{x}_1=\lambda_2\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1.\end{aligned}$$

由于 $\lambda_1\neq\lambda_2$，所以 $\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1=0$，即 $(\boldsymbol{x}_1,\boldsymbol{x}_2)=0$，故当 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 为实对称矩阵 $\boldsymbol{A}$ 对应于不同特征值的特征向量时，$\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 正交.

例 2.5.17 设

$$\boldsymbol{A}=\begin{bmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{bmatrix},$$

求正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ 为对角矩阵.

解

$$\begin{aligned}|\lambda\boldsymbol{E}-\boldsymbol{A}|&=\begin{vmatrix}\lambda-1&2&-2\\2&\lambda+2&-4\\-2&-4&\lambda+2\end{vmatrix}=\begin{vmatrix}\lambda-1&2&-4\\0&\lambda-2&0\\-2&-4&\lambda+6\end{vmatrix}\\&=(\lambda-2)(\lambda^2+5\lambda-14)\\&=(\lambda-2)^2(\lambda+7),\end{aligned}$$

令 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$，得 $\lambda_1=\lambda_2=2,\lambda_3=-7$.

对于 $\lambda_1=\lambda_2=2$，由 $(\lambda_1\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},$$

得线性无关的特征向量 $\boldsymbol{\xi}_1=[2,-1,0]^{\mathrm{T}}$，$\boldsymbol{\xi}_2=[2,0,1]^{\mathrm{T}}$. 用施密特正交化方法，先正交化，得

$$\boldsymbol{\beta}_1=\boldsymbol{\xi}_1,$$

$$\begin{aligned}\boldsymbol{\beta}_2&=\boldsymbol{\xi}_2-\frac{(\boldsymbol{\xi}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1\\&=\begin{bmatrix}2\\0\\1\end{bmatrix}-\frac{4}{5}\begin{bmatrix}2\\-1\\0\end{bmatrix}=\frac{1}{5}\begin{bmatrix}2\\4\\5\end{bmatrix},\end{aligned}$$

再将 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 单位化得

$$\boldsymbol{\eta}_1=\left[\frac{2\sqrt{5}}{5},-\frac{\sqrt{5}}{5},0\right]^{\mathrm{T}},$$

$$\boldsymbol{\eta}_2=\left[\frac{2\sqrt{5}}{15},\frac{4\sqrt{5}}{15},\frac{\sqrt{5}}{3}\right]^{\mathrm{T}}.$$

对于 $\lambda_3=-7$，由 $(\lambda_3\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix}-8&2&-2\\2&-5&-4\\-2&-4&-5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},$$

得特征向量 $\boldsymbol{\xi}_3=[1,2,-2]^{\mathrm{T}}$，单位化得 $\boldsymbol{\eta}_3=\left[\frac{1}{3},\frac{2}{3},-\frac{2}{3}\right]^{\mathrm{T}}$，取正交矩阵

$$Q=[\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3]=\begin{bmatrix}\frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3}\\ -\frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3}\\ 0 & \frac{\sqrt{5}}{3} & -\frac{2}{3}\end{bmatrix},$$

则$Q^{-1}AQ=\begin{bmatrix}2 & & \\ & 2 & \\ & & -7\end{bmatrix}$.

七、反求参数

常用方法如下.

(1)若 $A\sim B$，则 $|A|=|B|$，$r(A)=r(B)$，$\mathrm{tr}(A)=\mathrm{tr}(B)$，$\lambda_A=\lambda_B$. 这些等式可求参数.

(2)若 ξ_0 是 A 的属于特征值 λ_0 的特征向量，则有 $A\xi_0=\lambda\xi_0$，这便建立了若干等式(方程组)，可求参数.

(3)若 λ_0 是 A 的特征值，则有 $|\lambda_0 E-A|=0$，此等式可求参数.

例 2.5.18 已知

$$A=\begin{bmatrix}2 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & x\end{bmatrix},\quad B=\begin{bmatrix}2 & & \\ & y & \\ & & -1\end{bmatrix},$$

且 $A\sim B$，则参数 $x=$________，$y=$________.

解 应填 0；1.

因为 $A\sim B$，则有 $|A|=|B|$，$\mathrm{tr}(A)=\mathrm{tr}(B)$，得

$$|A|=-2=|B|=-2y,\quad y=1.$$

$$\mathrm{tr}(A)=2+x=\mathrm{tr}(B)=1+y,\quad x=0.$$

例 2.5.19 已知 $A=\begin{bmatrix}0 & 0 & 1\\ x & 1 & y\\ 1 & 0 & 0\end{bmatrix}$ 相似于对角矩阵，则 x,y 应满足________.

解 应填 $x+y=0$.

A 相似于对角矩阵，对角矩阵的对角元素是 A 的特征值，且应存在 3 个线性无关的特征向量. 因为

$$|\lambda E-A|=\begin{vmatrix}\lambda & 0 & -1\\ -x & \lambda-1 & -y\\ -1 & 0 & \lambda\end{vmatrix}=(\lambda-1)(\lambda^2-1)=(\lambda-1)^2(\lambda+1)=0,$$

故无论 x,y 为何值，均有 $\lambda_1=\lambda_2=1$，$\lambda_3=-1$.

对于二重特征值 $\lambda_1=\lambda_2=1$，应有两个线性无关的特征向量，即方程组

$$(E-A)x=0$$

应有两个线性无关的解向量. 因

$$\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}1 & 0 & -1\\ -x & 0 & -y\\ -1 & 0 & 1\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & -1\\ 0 & 0 & -x-y\\ 0 & 0 & 0\end{bmatrix},$$

故当 $x+y=0$ 时，有 $r(\boldsymbol{E}-\boldsymbol{A})=1$，从而 $(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 有两个线性无关的解向量，此时 $\boldsymbol{A}$ 相似于对角矩阵.

例 2.5.20　已知 $\boldsymbol{\alpha}=[1,k,1]^{\mathrm{T}}$ 是 $\boldsymbol{A}^{-1}$ 的特征向量，其中 $\boldsymbol{A}=\begin{bmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}$，求 k 及 $\boldsymbol{\alpha}$ 所对应的 $\boldsymbol{A}^{-1}$ 的特征值.

解　由题设 $\boldsymbol{A}^{-1}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}$，$\lambda$ 是 $\boldsymbol{A}^{-1}$ 的对应于 $\boldsymbol{\alpha}$ 的特征值，两端左乘 $\boldsymbol{A}$，得 $\boldsymbol{\alpha}=\lambda\boldsymbol{A}\boldsymbol{\alpha}$.

因 $\boldsymbol{\alpha}\neq\boldsymbol{0}$，故 $\lambda\neq 0$，$\boldsymbol{A}\boldsymbol{\alpha}=\dfrac{1}{\lambda}\boldsymbol{\alpha}\overset{\text{记}}{=\!=}\mu\boldsymbol{\alpha}$，即

$$\begin{bmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}\begin{bmatrix}1\\ k\\ 1\end{bmatrix}=\mu\begin{bmatrix}1\\ k\\ 1\end{bmatrix}.$$

由对应分量相等，得
$$\begin{cases}3+k=\mu,\\ 2+2k=k\mu,\\ 3+k=\mu,\end{cases}$$

得 $2+2k=k(3+k)$，$k^2+k-2=0$，故 $k=1$ 或 $k=-2$.

当 $k=1$ 时，$\boldsymbol{\alpha}=[1,1,1]^{\mathrm{T}}$，$\mu=4$，$\lambda=\dfrac{1}{\mu}=\dfrac{1}{4}$；

当 $k=-2$ 时，$\boldsymbol{\alpha}=[1,-2,1]^{\mathrm{T}}$，$\mu=1$，$\lambda=\dfrac{1}{\mu}=1$.

【注】　已知 $\boldsymbol{A}^{-1}$ 的特征向量，因 $\boldsymbol{A}$，$\boldsymbol{A}^{-1}$ 有相同的特征向量，故不必求出 $\boldsymbol{A}^{-1}$.

八、由特征值、特征向量反求 A

若有可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$，则 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，这便是求解的基本思路.

例 2.5.21　设 $\boldsymbol{A}$ 是 3 阶矩阵，已知 $\boldsymbol{A}\boldsymbol{\xi}_i=i\boldsymbol{\xi}_i\,(i=1,2,3)$，其中 $\boldsymbol{\xi}_1=[1,0,0]^{\mathrm{T}}$，$\boldsymbol{\xi}_2=[1,1,0]^{\mathrm{T}}$，$\boldsymbol{\xi}_3=[1,1,1]^{\mathrm{T}}$，求矩阵 $\boldsymbol{A}$.

解　**方法一**　由题设 $\boldsymbol{A}\boldsymbol{\xi}_i=i\boldsymbol{\xi}_i\,(i=1,2,3)$ 知，3 阶矩阵 $\boldsymbol{A}$ 有 3 个互不相同的特征值，所以 $\boldsymbol{A}$ 可以相似于对角矩阵，且 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 是 3 个线性无关的特征向量，故存在可逆矩阵 $\boldsymbol{P}=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3]$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}1 & & \\ & 2 & \\ & & 3\end{bmatrix},$$

故
$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}1 & & \\ & 2 & \\ & & 3\end{bmatrix}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & & \\ & 2 & \\ & & 3\end{bmatrix}\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\ 0 & 2 & 1\\ 0 & 0 & 3\end{bmatrix}.$$

方法二　由题设条件 $A\xi_i=i\xi_i(i=1,2,3)$，合并成矩阵形式，得

$$[A\xi_1,A\xi_2,A\xi_3]=A[\xi_1,\xi_2,\xi_3]=[\xi_1,2\xi_2,3\xi_3],$$

两端右乘$[\xi_1,\xi_2,\xi_3]^{-1}$，得

$$A=[\xi_1,2\xi_2,3\xi_3][\xi_1,\xi_2,\xi_3]^{-1}$$

$$=\begin{bmatrix}1&2&3\\0&2&3\\0&0&3\end{bmatrix}\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}^{-1}=\begin{bmatrix}1&2&3\\0&2&3\\0&0&3\end{bmatrix}\begin{bmatrix}1&-1&0\\0&1&-1\\0&0&1\end{bmatrix}=\begin{bmatrix}1&1&1\\0&2&1\\0&0&3\end{bmatrix}.$$

九、求 A^k 及 $f(A)$

主要理论：当 $A\sim B=\Lambda=\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix}$ 时，有

$$P^{-1}A^kP=\Lambda^k,\quad A^k=P\Lambda^kP^{-1}=P\begin{bmatrix}\lambda_1^k&&&\\&\lambda_2^k&&\\&&\ddots&\\&&&\lambda_n^k\end{bmatrix}P^{-1},$$

$$P^{-1}f(A)P=f(\Lambda),\quad f(A)=Pf(\Lambda)P^{-1}=P\begin{bmatrix}f(\lambda_1)&&&\\&f(\lambda_2)&&\\&&\ddots&\\&&&f(\lambda_n)\end{bmatrix}P^{-1},$$

从而提供了计算 A^k 及 $f(A)$ 的重要方法.

例 2.5.22　已知 A 是 3 阶矩阵，且 $A\sim\Lambda=\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}$. 设 $B=A^3-6A^2+11A-8E$，则 $B=$________.

解　应填 $-2E$.

由 $A\sim\Lambda=\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}$ 知，存在可逆矩阵 P，使 $P^{-1}AP=\Lambda$，$A=P\Lambda P^{-1}$，则

$$\begin{aligned}B&=A^3-6A^2+11A-8E=(P\Lambda P^{-1})^3-6(P\Lambda P^{-1})^2+11P\Lambda P^{-1}-8E\\&=P(\Lambda^3-6\Lambda^2+11\Lambda-8E)P^{-1}\\&=P\left(\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}^3-6\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}^2+11\begin{bmatrix}1&&\\&2&\\&&3\end{bmatrix}-8\begin{bmatrix}1&&\\&1&\\&&1\end{bmatrix}\right)P^{-1}\end{aligned}$$

$$=P\begin{bmatrix}1-6+11-8 & & \\ & 8-24+22-8 & \\ & & 27-54+33-8\end{bmatrix}P^{-1}$$

$$=P\begin{bmatrix}-2 & & \\ & -2 & \\ & & -2\end{bmatrix}P^{-1}=-2E.$$

例 2.5.23 设 $A=\begin{bmatrix}1 & -2 & 0 & 0\\ -1 & 0 & 0 & 0\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 2\end{bmatrix}$，求 A^n.

解 A 是分块对角矩阵，记 $A=\begin{bmatrix}B & O\\ O & C\end{bmatrix}$，其中 $B=\begin{bmatrix}1 & -2\\ -1 & 0\end{bmatrix}$，$C=\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}$，则

$$A^n=\begin{bmatrix}B & O\\ O & C\end{bmatrix}^n=\begin{bmatrix}B^n & O\\ O & C^n\end{bmatrix}.$$

对于 $B=\begin{bmatrix}1 & -2\\ -1 & 0\end{bmatrix}$，$|\lambda E-B|=\begin{vmatrix}\lambda-1 & 2\\ 1 & \lambda\end{vmatrix}=(\lambda+1)(\lambda-2)=0$，对应的特征值为 $\lambda_1=-1$，$\lambda_2=2$.

当 $\lambda_1=-1$ 时，由 $(\lambda_1 E-B)x=0$，得 $\xi_1=[1,1]^T$；

当 $\lambda_2=2$ 时，由 $(\lambda_2 E-B)x=0$，得 $\xi_2=[-2,1]^T$.

于是有

$$P^{-1}BP=\begin{bmatrix}-1 & 0\\ 0 & 2\end{bmatrix}，其中\ P=\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix}，P^{-1}=\frac{1}{3}\begin{bmatrix}1 & 2\\ -1 & 1\end{bmatrix},$$

故

$$B^n=\left(P\begin{bmatrix}-1 & 0\\ 0 & 2\end{bmatrix}P^{-1}\right)^n=P\begin{bmatrix}-1 & 0\\ 0 & 2\end{bmatrix}^nP^{-1}$$

$$=\frac{1}{3}\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix}\begin{bmatrix}(-1)^n & 0\\ 0 & 2^n\end{bmatrix}\begin{bmatrix}1 & 2\\ -1 & 1\end{bmatrix}$$

$$=\frac{1}{3}\begin{bmatrix}(-1)^n+2^{n+1} & (-1)^n\cdot 2-2^{n+1}\\ (-1)^n-2^n & (-1)^n\cdot 2+2^n\end{bmatrix}.$$

对于 $C=\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}$，因当 $n\geqslant 2$ 时，$\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}^n=O$，故

$$C^n=\left(2E+\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\right)^n=2^nE^n+n\cdot 2^{n-1}E^{n-1}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}=\begin{bmatrix}2^n & n\cdot 2^{n-1}\\ 0 & 2^n\end{bmatrix},$$

因此，

$$A^n=\begin{bmatrix}B^n & O\\ O & C^n\end{bmatrix}=\begin{bmatrix}\frac{1}{3}(-1)^n+\frac{1}{3}\cdot 2^{n+1} & \frac{2}{3}(-1)^n-\frac{1}{3}\cdot 2^{n+1} & 0 & 0\\ \frac{1}{3}(-1)^n-\frac{1}{3}\cdot 2^n & \frac{2}{3}(-1)^n+\frac{1}{3}\cdot 2^n & 0 & 0\\ 0 & 0 & 2^n & n\cdot 2^{n-1}\\ 0 & 0 & 0 & 2^n\end{bmatrix}.$$

基础习题精练

习 题

2.5.1 求下列矩阵的特征值和特征向量：

(1)$\boldsymbol{A}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$；　　(2)$\boldsymbol{B}=\begin{bmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{bmatrix}$.

2.5.2 设 $\boldsymbol{A}$ 是 3 阶矩阵，已知 $-3\boldsymbol{E}+\boldsymbol{A}$ 不可逆，$|2\boldsymbol{E}+\boldsymbol{A}|=0$. $(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 有非零解，则 $|\boldsymbol{A}^*-\boldsymbol{E}|=$ ________.

2.5.3 设 $\boldsymbol{A}$ 是 4 阶实矩阵，$\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵，已知 $\boldsymbol{A}^*$ 有特征值 $1,-1,2,-4$，求 $|\boldsymbol{A}^3+2\boldsymbol{A}^2-\boldsymbol{A}-3\boldsymbol{E}|$.

2.5.4 设 $\boldsymbol{A}=\begin{bmatrix}a&-1&c\\5&b&3\\1-c&0&-a\end{bmatrix}$，且 $|\boldsymbol{A}|=-1$，$\boldsymbol{A}^*$ 是 $\boldsymbol{A}$ 的伴随矩阵，$\boldsymbol{A}^*$ 有特征值 λ_0，对应于 λ_0 的特征向量为 $\boldsymbol{\xi}=[-1,-1,1]^{\mathrm{T}}$，求 a,b,c 及 λ_0.

2.5.5 设 $\boldsymbol{A}\sim\begin{bmatrix}1&&\\&-1&\\&&2\end{bmatrix}$，则 $r(\boldsymbol{A}-\boldsymbol{E})+r(\boldsymbol{A}+\boldsymbol{E})=$ ________.

2.5.6 设 $\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}$ 是可逆矩阵，$\boldsymbol{BA}=\begin{bmatrix}2a_{11}&-a_{12}&3a_{13}\\2a_{21}&-a_{22}&3a_{23}\\2a_{31}&-a_{32}&3a_{33}\end{bmatrix}$，则 $\boldsymbol{B}\sim$(　　).

(A) $\begin{bmatrix}-1&&\\&2&\\&&-3\end{bmatrix}$　　(B) $\begin{bmatrix}2&&\\&-1&\\&&3\end{bmatrix}$　　(C) $\begin{bmatrix}3&&\\&-2&\\&&-1\end{bmatrix}$　　(D) $\begin{bmatrix}1&&\\&3&\\&&2\end{bmatrix}$

2.5.7 证明：n 阶矩阵 $\begin{bmatrix}1&1&\cdots&1\\1&1&\cdots&1\\\vdots&\vdots&&\vdots\\1&1&\cdots&1\end{bmatrix}$ 与 $\begin{bmatrix}0&\cdots&0&1\\0&\cdots&0&2\\\vdots&&\vdots&\vdots\\0&\cdots&0&n\end{bmatrix}$ 相似.

2.5.8 已知 $\boldsymbol{A}=\begin{bmatrix}-4&-10&0\\1&y&0\\3&x&1\end{bmatrix}\sim\begin{bmatrix}-2&&\\&1&\\&&1\end{bmatrix}$，则 x,y 分别为 ________.

2.5.9 已知 $\boldsymbol{\xi}=[1,1,-1]^{\mathrm{T}}$ 是矩阵 $\boldsymbol{A}=\begin{bmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{bmatrix}$ 的一个特征向量.

(1)确定参数 a,b 及 $\boldsymbol{\xi}$ 对应的特征值 λ；

(2)$\boldsymbol{A}$ 是否相似于对角矩阵？说明理由.

2.5.10 已知1，1，−1是3阶实对称矩阵$\boldsymbol{A}$的特征值，向量$\xi_1=[1,1,1]^{\mathrm{T}}$，$\xi_2=[2,2,1]^{\mathrm{T}}$是$\boldsymbol{A}$的对应于特征值$\lambda_1=\lambda_2=1$的特征向量，求矩阵$\boldsymbol{A}$.

2.5.11 已知$\boldsymbol{A}=\begin{bmatrix}2 & x & 1\\ 0 & 3 & 0\\ 3 & -6 & 0\end{bmatrix}$能相似于对角矩阵，求$\boldsymbol{A}^{100}$.

解答

2.5.1 **解** (1)对矩阵$\boldsymbol{A}=\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}$，由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda & 0 & -1\\ 0 & \lambda-1 & 0\\ -1 & 0 & \lambda\end{vmatrix}=(\lambda-1)(\lambda^2-1)=(\lambda-1)^2(\lambda+1)=0,$$

得特征值$\lambda_1=\lambda_2=1$，$\lambda_3=-1$.

当$\lambda_1=\lambda_2=1$时，由$(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix}1 & 0 & -1\\ 0 & 0 & 0\\ -1 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

解得基础解系$\xi_1=[1,0,1]^{\mathrm{T}}$，$\xi_2=[0,1,0]^{\mathrm{T}}$，$k_1\xi_1+k_2\xi_2$为对应于$\lambda_1=\lambda_2=1$的全部特征向量，其中$k_1$，$k_2$是不全为零的任意常数.

当$\lambda_3=-1$时，由$(-\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix}-1 & 0 & -1\\ 0 & -2 & 0\\ -1 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

解得基础解系为$\xi_3=[1,0,-1]^{\mathrm{T}}$，$k\xi_3$为对应于$\lambda_3=-1$的全部特征向量，其中$k$为任意非零常数.

(2)对矩阵$\boldsymbol{B}=\begin{bmatrix}2 & -1 & 2\\ 5 & -3 & 3\\ -1 & 0 & -2\end{bmatrix}$，由

$$|\lambda\boldsymbol{E}-\boldsymbol{B}|=\begin{vmatrix}\lambda-2 & 1 & -2\\ -5 & \lambda+3 & -3\\ 1 & 0 & \lambda+2\end{vmatrix}$$

$$\xlongequal{[1]+[2]-[3]}\begin{vmatrix}\lambda+1 & 1 & -2\\ \lambda+1 & \lambda+3 & -3\\ -\lambda-1 & 0 & \lambda+2\end{vmatrix}$$

$$=(\lambda+1)\begin{vmatrix}1&1&-2\\1&\lambda+3&-3\\-1&0&\lambda+2\end{vmatrix}=(\lambda+1)\begin{vmatrix}1&1&-2\\0&\lambda+2&-1\\0&1&\lambda\end{vmatrix}$$

$$=(\lambda+1)(\lambda^2+2\lambda+1)=(\lambda+1)^3=0,$$

得特征值 $\lambda_1=\lambda_2=\lambda_3=-1$.

当 $\lambda_1=\lambda_2=\lambda_3=-1$ 时，由 $(-\boldsymbol{E}-\boldsymbol{B})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix}-3&1&-2\\-5&2&-3\\1&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}, \tag{*}$$

解得全部特征向量为 $k[1,1,-1]^{\mathrm{T}}$，其中 k 为任意非零常数.

【注】 在第(2)问中，因 $|\lambda\boldsymbol{E}-\boldsymbol{B}|=0$，特征矩阵 $\lambda\boldsymbol{E}-\boldsymbol{B}$ 的行向量组线性相关，任意两个向量线性无关，方程组中总有一个方程是多余方程，故在求解方程组(*)时，去掉第二个方程是方便的.

2.5.2 42 **解** 由题知，$-3\boldsymbol{E}+\boldsymbol{A}$ 不可逆，故 $|3\boldsymbol{E}-\boldsymbol{A}|=0$，则 $\boldsymbol{A}$ 有特征值 $\lambda_1=3$；

$|2\boldsymbol{E}+\boldsymbol{A}|=(-1)^3|-2\boldsymbol{E}-\boldsymbol{A}|=0$，则 $\boldsymbol{A}$ 有特征值 $\lambda_2=-2$；

$(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 有非零解，故 $|\boldsymbol{E}-\boldsymbol{A}|=0$，则 $\boldsymbol{A}$ 有特征值 $\lambda_3=1$.

综上，3 阶矩阵 $\boldsymbol{A}$ 有特征值 $\lambda_1=3,\lambda_2=-2,\lambda_3=1$，得 $|\boldsymbol{A}|=\prod_{i=1}^{3}\lambda_i=-6$. 则 $\boldsymbol{A}^*$ 的特征值为

$$\mu_1=\frac{-6}{3}=-2,\quad \mu_2=\frac{-6}{-2}=3,\quad \mu_3=-6.$$

则 $\boldsymbol{A}^*-\boldsymbol{E}$ 的特征值为 $m_1=-2-1=-3,m_2=3-1=2,m_3=-6-1=-7$，故

$$|\boldsymbol{A}^*-\boldsymbol{E}|=(-3)\times2\times(-7)=42.$$

2.5.3 解 $\boldsymbol{A}^*$ 有特征值 $1,-1,2,-4$，则 $|\boldsymbol{A}^*|=\prod_{i=1}^{4}\lambda_i=8\neq0$，故 $\boldsymbol{A}^*$ 可逆，$\boldsymbol{A}$ 可逆. 又 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}=|\boldsymbol{A}|^3=8$，得 $|\boldsymbol{A}|=2$. 故 $\boldsymbol{A}$ 的特征值 $\lambda_A=\dfrac{|\boldsymbol{A}|}{\lambda_{A^*}}$，即为 $2,-2,1,-\dfrac{1}{2}$.

设 $f(\boldsymbol{A})=\boldsymbol{A}^3+2\boldsymbol{A}^2-\boldsymbol{A}-3\boldsymbol{E}$，则 $f(\boldsymbol{A})$ 的特征值为

$$f(2)=2^3+2\times2^2-2-3=11,$$

$$f(-2)=(-2)^3+2\times(-2)^2+2-3=-1,$$

$$f(1)=1+2-1-3=-1,$$

$$f\left(-\frac{1}{2}\right)=\left(-\frac{1}{2}\right)^3+2\times\left(-\frac{1}{2}\right)^2+\frac{1}{2}-3=-\frac{17}{8},$$

故 $|\boldsymbol{A}^3+2\boldsymbol{A}^2-\boldsymbol{A}-3\boldsymbol{E}|=f(2)\cdot f(-2)\cdot f(1)\cdot f\left(-\dfrac{1}{2}\right)=-\dfrac{187}{8}$.

2.5.4 解 由题知 $\boldsymbol{A}^*\boldsymbol{\xi}=\lambda_0\boldsymbol{\xi}$，两端左乘 $\boldsymbol{A}$，$\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{\xi}=\lambda_0\boldsymbol{A}\boldsymbol{\xi}=|\boldsymbol{A}|\boldsymbol{\xi}=-\boldsymbol{\xi}$，即

$$\lambda_0\begin{bmatrix}a&-1&c\\5&b&3\\1-c&0&-a\end{bmatrix}\begin{bmatrix}-1\\-1\\1\end{bmatrix}=\begin{bmatrix}1\\1\\-1\end{bmatrix},$$

$$\begin{cases}\lambda_0(-a+1+c)=1,\\ \lambda_0(-5-b+3)=1,\\ \lambda_0(-1+c-a)=-1,\end{cases}$$

得 $\lambda_0=1,a=c,b=-3$,又 $|\boldsymbol{A}|=\begin{vmatrix} a & -1 & a \\ 5 & -3 & 3 \\ 1-a & 0 & -a \end{vmatrix}=a-3=-1$,得 $a=c=2$.

2.5.5 4 **解** 因 $\boldsymbol{A}\sim\boldsymbol{\Lambda}=\begin{bmatrix} 1 & & \\ & -1 & \\ & & 2 \end{bmatrix}$,即存在可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$,即 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$,故

$$\begin{aligned} r(\boldsymbol{A}-\boldsymbol{E})+r(\boldsymbol{A}+\boldsymbol{E}) &= r(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}-\boldsymbol{E})+r(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}+\boldsymbol{E}) \\ &= r(\boldsymbol{P}(\boldsymbol{\Lambda}-\boldsymbol{E})\boldsymbol{P}^{-1})+r(\boldsymbol{P}(\boldsymbol{\Lambda}+\boldsymbol{E})\boldsymbol{P}^{-1})=r(\boldsymbol{\Lambda}-\boldsymbol{E})+r(\boldsymbol{\Lambda}+\boldsymbol{E}) \\ &= r\left(\begin{bmatrix} 0 & & \\ & -2 & \\ & & 1 \end{bmatrix}\right)+r\left(\begin{bmatrix} 2 & & \\ & 0 & \\ & & 3 \end{bmatrix}\right) \\ &= 2+2=4. \end{aligned}$$

2.5.6 (B) **解** 因

$$\boldsymbol{B}\boldsymbol{A}=\begin{bmatrix} 2a_{11} & -a_{12} & 3a_{13} \\ 2a_{21} & -a_{22} & 3a_{23} \\ 2a_{31} & -a_{32} & 3a_{33} \end{bmatrix}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} 2 & & \\ & -1 & \\ & & 3 \end{bmatrix}=\boldsymbol{A}\begin{bmatrix} 2 & & \\ & -1 & \\ & & 3 \end{bmatrix},$$

又 $\boldsymbol{A}$ 可逆,故 $\boldsymbol{B}=\boldsymbol{A}\begin{bmatrix} 2 & & \\ & -1 & \\ & & 3 \end{bmatrix}\boldsymbol{A}^{-1}=(\boldsymbol{A}^{-1})^{-1}\begin{bmatrix} 2 & & \\ & -1 & \\ & & 3 \end{bmatrix}\boldsymbol{A}^{-1}$,故 $\boldsymbol{B}\sim\begin{bmatrix} 2 & & \\ & -1 & \\ & & 3 \end{bmatrix}$.

2.5.7 **证明** 设 $\boldsymbol{A}=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix},\boldsymbol{B}=\begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & n \end{bmatrix}$.

因为

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-1 & -1 & \cdots & -1 \\ -1 & \lambda-1 & \cdots & -1 \\ \vdots & \vdots & & \vdots \\ -1 & -1 & \cdots & \lambda-1 \end{vmatrix}=(\lambda-n)\lambda^{n-1}=0,$$

$$|\lambda\boldsymbol{E}-\boldsymbol{B}|=\begin{vmatrix} \lambda & 0 & \cdots & -1 \\ 0 & \lambda & \cdots & -2 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda-n \end{vmatrix}=(\lambda-n)\lambda^{n-1}=0,$$

所以 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值 $\lambda_1=n,\lambda_2=0(n-1$ 重).

由于 $\boldsymbol{A}$ 为实对称矩阵,所以 $\boldsymbol{A}$ 相似于对角矩阵

$$\boldsymbol{\Lambda}=\begin{bmatrix} n & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}.$$

因为 $r(\lambda_2 \boldsymbol{E}-\boldsymbol{B})=r(\boldsymbol{B})=1$，所以 $\boldsymbol{B}$ 对应于特征值 $\lambda_2=0$ 有 $n-1$ 个线性无关的特征向量，于是 $\boldsymbol{B}$ 也相似于 $\boldsymbol{A}$.

故 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似.

【注】 (1)本题是 2014 年考研实考题，阅卷的结果表明，很多考生没有考虑用相似的传递性来证明，或者没有说明 $\boldsymbol{A},\boldsymbol{B}$ 可以相似对角化，就直接由特征值相同推得 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$，这是不完整的.

(2)本题也道出了一个判别矩阵相似的重要定理：若 $\boldsymbol{A},\boldsymbol{B}$ 均可相似对角化，且特征值相同，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似.

(3)读者应进一步明白，若存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$，只能得出 $\boldsymbol{A}\sim\boldsymbol{B}$，是推不出 $\boldsymbol{A}\sim\boldsymbol{B}\sim\boldsymbol{\Lambda}$ 的. 如 $\boldsymbol{A}=\begin{bmatrix}1&1\\0&1\end{bmatrix}$，其特征值 $\lambda_1=\lambda_2=1$，但 $2-r(\boldsymbol{E}-\boldsymbol{A})=2-1=1$，这个二重特征值 $\lambda_1=\lambda_2=1$ 没有两个线性无关的特征向量，于是 $\boldsymbol{A}$ 不可相似对角化；再取可逆矩阵 $\boldsymbol{P}=\begin{bmatrix}2&0\\0&1\end{bmatrix}$，则 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{bmatrix}1&\frac{1}{2}\\0&1\end{bmatrix}=\boldsymbol{B}$，故 $\boldsymbol{A}\sim\boldsymbol{B}$，$\boldsymbol{B}$ 也不可相似对角化.

2.5.8 6，3 **解** 因 $\boldsymbol{A}\sim\boldsymbol{\Lambda}$，则 $\sum\limits_{i=1}^{3}a_{ii}=-4+y+1=\sum\limits_{i=1}^{3}\lambda_i=0$，故 $y=3$.

又 $\lambda_2=\lambda_3=1$ 是二重特征值，已知 $\boldsymbol{A}\sim\boldsymbol{\Lambda}$，故 $r(\boldsymbol{E}-\boldsymbol{A})=1$. 由

$$\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}5&10&0\\-1&-2&0\\-3&-x&0\end{bmatrix}\longrightarrow\begin{bmatrix}1&2&0\\0&-x+6&0\\0&0&0\end{bmatrix},$$

知 $r(\boldsymbol{E}-\boldsymbol{A})=1\Leftrightarrow x=6$.

2.5.9 **解** (1)设 $\boldsymbol{A}$ 的特征向量 $\boldsymbol{\xi}$ 所对应的特征值为 λ，则有 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$，即

$$\begin{bmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{bmatrix}\begin{bmatrix}1\\1\\-1\end{bmatrix}=\lambda\begin{bmatrix}1\\1\\-1\end{bmatrix},$$

即

$$\begin{cases}2-1-2=\lambda,\\5+a-3=\lambda,\\-1+b+2=-\lambda,\end{cases}$$

解得 $\lambda=-1,a=-3,b=0$.

(2)当 $a=-3,b=0$ 时，令

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&1&-2\\-5&\lambda+3&-3\\1&0&\lambda+2\end{vmatrix}=(\lambda+1)^3=0,$$

可知 $\lambda=-1$ 是 $\boldsymbol{A}$ 的 3 重特征值，但

$$-\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}-3&1&-2\\-5&2&-3\\1&0&1\end{bmatrix},r(-\boldsymbol{E}-\boldsymbol{A})=2.$$

当 $\lambda=-1$ 时，对应的线性无关特征向量只有一个，因此 $\boldsymbol{A}$ 不能相似于对角矩阵.

2.5.10　分析　因 $\boldsymbol{A}$ 是实对称矩阵，故 $\boldsymbol{A}$ 一定相似于对角矩阵，且不同的特征值对应的特征向量相互正交，故由此可求出 $\lambda_3=-1$ 对应的特征向量（和例 2.5.21 不同，在例 2.5.21 中，$\boldsymbol{A}$ 不是对称矩阵，求 $\boldsymbol{A}$ 时，$\boldsymbol{A}$ 的 3 个特征向量必须给出）.

解　设 $\lambda_3=-1$ 对应的特征向量为 $\boldsymbol{\xi}_3=[x_1,x_2,x_3]^{\mathrm{T}}$，则由 $\boldsymbol{A}$ 是实对称矩阵，有

$$\begin{cases}(\boldsymbol{\xi}_1,\boldsymbol{\xi}_3)=x_1+x_2+x_3=0,\\(\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=2x_1+2x_2+x_3=0,\end{cases}$$

取 $\boldsymbol{\xi}_3=[-1,1,0]^{\mathrm{T}}$，从而得可逆矩阵 $\boldsymbol{P}=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3]$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda},$$

则有

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}=\begin{bmatrix}1&2&-1\\1&2&1\\1&1&0\end{bmatrix}\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}\begin{bmatrix}1&2&-1\\1&2&1\\1&1&0\end{bmatrix}^{-1}$$

$$=\frac{1}{2}\begin{bmatrix}1&2&-1\\1&2&1\\1&1&0\end{bmatrix}\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}\begin{bmatrix}-1&-1&4\\1&1&-2\\-1&1&0\end{bmatrix}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}.$$

【注】　若将 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 标准正交化，还可求得正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1}=\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$，用转置来替代求逆运算，结果当然是一样的.

2.5.11　解　因 $\boldsymbol{A}=\begin{bmatrix}2&x&1\\0&3&0\\3&-6&0\end{bmatrix}$ 的特征方程为

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&-x&-1\\0&\lambda-3&0\\-3&6&\lambda\end{vmatrix}=(\lambda-3)(\lambda^2-2\lambda-3)=(\lambda-3)^2(\lambda+1)=0,$$

故有特征值 $\lambda_1=\lambda_2=3,\lambda_3=-1$.

已知 $\boldsymbol{A}$ 能相似于对角矩阵，故当 $\lambda=3$（二重根）时，应有 $r(3\boldsymbol{E}-\boldsymbol{A})=1$. 由

$$3\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}1&-x&-1\\0&0&0\\-3&6&3\end{bmatrix}\longrightarrow\begin{bmatrix}1&-x&-1\\0&0&0\\0&6-3x&0\end{bmatrix},$$

知 $r(3\boldsymbol{E}-\boldsymbol{A})=1\Leftrightarrow x=2$.

故当 $\lambda_1=\lambda_2=3$ 时，$(3\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的同解方程为 $x_1-2x_2-x_3=0$，解得线性无关的特征向量

$$\boldsymbol{\xi}_1=[1,0,1]^{\mathrm{T}},\quad \boldsymbol{\xi}_2=[2,1,0]^{\mathrm{T}};$$

当 $\lambda_3=-1$ 时，

$$(-\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\begin{bmatrix}-3&-2&-1\\0&-4&0\\-3&6&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0},$$

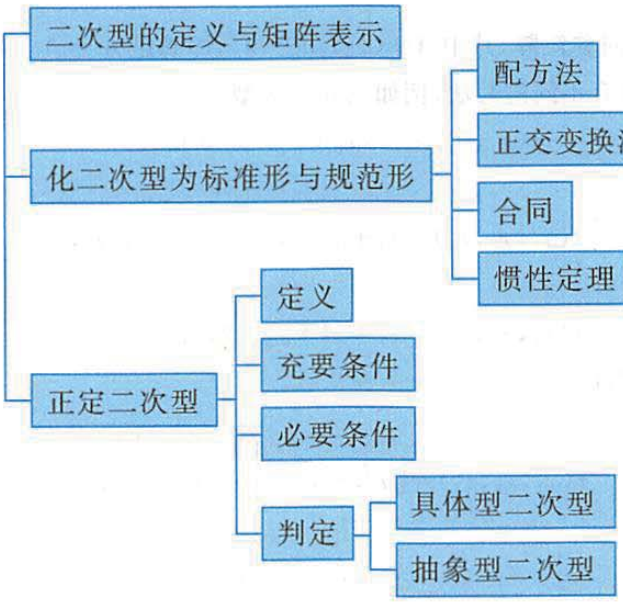
解得线性无关的特征向量 $\boldsymbol{\xi}_3=[1,0,-3]^{\mathrm{T}}$. 于是有可逆矩阵 $\boldsymbol{P}=[\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3]$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\begin{bmatrix}3&&\\&3&\\&&-1\end{bmatrix}$，则 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，故

$$A^{100}=P\Lambda^{100}P^{-1}=\begin{bmatrix}1&2&1\\0&1&0\\1&0&-3\end{bmatrix}\begin{bmatrix}3^{100}&&\\&3^{100}&\\&&(-1)^{100}\end{bmatrix}\frac{1}{4}\begin{bmatrix}3&-6&1\\0&4&0\\1&-2&-1\end{bmatrix}$$

$$=\frac{1}{4}\begin{bmatrix}3^{101}+1&2\times3^{100}-2&3^{100}-1\\0&4\times3^{100}&0\\3^{101}-3&-6\times3^{100}+6&3^{100}+3\end{bmatrix}.$$

第6讲 二次型

基础知识结构



基础内容精讲

一、二次型的定义及其矩阵表达式

n 元变量 $x_1,x_2,\cdots,x_n$ 的二次齐次多项式

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2&+\cdots+2a_{1n}x_1x_n\\&+a_{22}x_2^2+\cdots+2a_{2n}x_2x_n\\&+\cdots\\&+a_{nn}x_n^2\end{aligned}$$

称为 **n 元二次型**，简称**二次型**.

考研只研究系数 $a_{ij}\in\mathbf{R}(i\leqslant j,i,j=1,2,\cdots,n)$ 的情况，故称此二次型 f 为**实二次型**.

因为 $x_ix_j=x_jx_i$，若令 $a_{ij}=a_{ji}(i,j=1,2,\cdots,n)$，则 $2a_{ij}x_ix_j=a_{ij}x_ix_j+a_{ji}x_jx_i$，于是

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n+a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n\\+\cdots+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2\qquad(*)\end{aligned}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j, \tag{* *}$$

其中（ * ）式称为完全展开式，（ * * ）式称为和式. 令

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},$$

则二次型可表示为

$$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}. \tag{* * *}$$

（ * * * ）式称为二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵表达式，实对称矩阵 $\boldsymbol{A}$ 称为**二次型 $f(\boldsymbol{x})$ 的矩阵**. 这里需要着重多说几句.

二次型的矩阵 $\boldsymbol{A}$ 是一个**对称矩阵**，其中 $\boldsymbol{A}=(a_{ij})$，$a_{ij}=a_{ji}$，即满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$. 为什么要强调这一点？

事实上，一个二次型可以有不同的写法，例如三元二次型

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+4x_1x_2,$$

可以写成

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+x_1x_2+3x_2x_1,$$

也可以写成

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+2x_1x_2+2x_2x_1$$

等. 对应的用矩阵表示也有多种形式：

$$f(x_1,x_2,x_3)=[x_1,x_2,x_3]\begin{bmatrix}1&4&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \tag{①}$$

$$=[x_1,x_2,x_3]\begin{bmatrix}1&1&0\\3&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \tag{②}$$

$$=[x_1,x_2,x_3]\begin{bmatrix}1&2&0\\2&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}. \tag{③}$$

这样，代表二次型的矩阵就不唯一了，不利于研究二次型问题. 现在我们立了"规矩"，**规定二次型的矩阵必须是对称矩阵，代表二次型的矩阵就是唯一的.** 所以只有对称矩阵才是二次型的矩阵，也只有用对称矩阵表达的式子才是二次型的矩阵表达式，上面第③种形式是按"规矩"写的. 再看下面这个例子.

注例 写出三元二次型 $f(x_1,x_2,x_3)=2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_2x_3+2x_1x_3$ 的二次型矩阵 $\boldsymbol{A}$.

解 将二次型表示成矩阵形式是基本要求，方法：$\boldsymbol{A}$ 的主对角线元素 a_{ii} 是平方项 x_i^2 的系数，a_{ij} 是混合项 x_ix_j 的系数的一半，或利用矩阵乘法（见本题解答）.

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_1^2+2x_2^2+2x_3^2-2x_1x_2-2x_2x_3+2x_1x_3\\&=2x_1^2-x_1x_2+x_1x_3-x_2x_1+2x_2^2-x_2x_3+x_3x_1-x_3x_2+2x_3^2\\&=x_1(2x_1-x_2+x_3)+x_2(-x_1+2x_2-x_3)+x_3(x_1-x_2+2x_3)\\&=[x_1,x_2,x_3]\begin{bmatrix}2x_1-x_2+x_3\\-x_1+2x_2-x_3\\x_1-x_2+2x_3\end{bmatrix}=[x_1,x_2,x_3]\begin{bmatrix}2&-1&1\\-1&2&-1\\1&-1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\end{aligned}$$

得
$$A=\begin{bmatrix}2&-1&1\\-1&2&-1\\1&-1&2\end{bmatrix}.$$

系数矩阵 A 的秩称为二次型 $f(x)$的秩. 比如注例中,$A=\begin{bmatrix}2&-1&1\\-1&2&-1\\1&-1&2\end{bmatrix}$,其秩 $r(A)=3$,故其对应的二次型的秩也是 3.

二、合同变换,二次型的合同标准形、规范形

1. 线性变换的定义

对于 n 元二次型 $f(x_1,x_2,\cdots,x_n)$,若令

$$\begin{cases}x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n,\\x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n,\\\quad\cdots\cdots\\x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n,\end{cases}$$

记 $x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$,$C=\begin{bmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{m}\end{bmatrix}$,$y=\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}$,则上式可写为

$$x=Cy,$$

称为从 $y_1,y_2,\cdots,y_n$ 到 $x_1,x_2,\cdots,x_n$ 的**线性变换**. 若线性变换的系数矩阵 C 可逆,即$|C|\neq0$,则称为**可逆线性变换**. 现给出 $f(x)=x^{\mathrm{T}}Ax$,令 $x=Cy$,则

$$f(x)=(Cy)^{\mathrm{T}}A(Cy)=y^{\mathrm{T}}(C^{\mathrm{T}}AC)y.$$

记 $B=C^{\mathrm{T}}AC$,则

$$f(x)=y^{\mathrm{T}}By=g(y).$$

此时,二次型 $f(x)=x^{\mathrm{T}}Ax$ 通过线性变换 $x=Cy$ 得到一个新二次型 $g(y)=y^{\mathrm{T}}By$.

2. 矩阵合同的定义与性质

读者不难发现,上述二次型 $f(x)$与 $g(y)$的系数矩阵 A 与 B 有下述关系

$$B=C^{\mathrm{T}}AC,$$

这种关系叫什么呢? 下面就给出数学上的描述.

定义 设 A,B 为 n 阶实对称矩阵,若存在可逆矩阵 C,使得

$$C^{\mathrm{T}}AC=B,$$

则称 A 与 B **合同**. 记作 $A\simeq B$. 此时称 $f(x)$与 $g(y)$为**合同二次型**.

可以看出,在二次型背景下,A 表征的是 $f(x)=x^{\mathrm{T}}Ax$ 的"形态",B 表征的是 $g(y)=y^{\mathrm{T}}By$ 的"形态",但是毕竟 $f(x)=g(y)$,是同一个东西,所谓 A,B 分别表征了不同的"形态",无非是因为在 $x=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$ 的参照系下与在 $y=[y_1,y_2,\cdots,y_n]^{\mathrm{T}}$ 的参照系下看到的是同一个事物的不同"形态"罢了.

若在 $x=[x_1,x_2,x_3]^{\mathrm{T}}$ 的参照系下,"$f(x)=$某常数"的图形是"正"着放的,如图 2-6-1 所示;在 $y=$

$[y_1,y_2,y_3]^{\mathrm{T}}$ 的参照系下，“$g(y)=$某常数”的图形是“歪”着放的，如图 2-6-2 所示（读者只需将头稍稍歪一下，使得 y_3 轴平行于你的鼻梁，即可看到图形在 $\boldsymbol{y}=[y_1,y_2,y_3]^{\mathrm{T}}$ 的参照系下的“歪”的形态）.

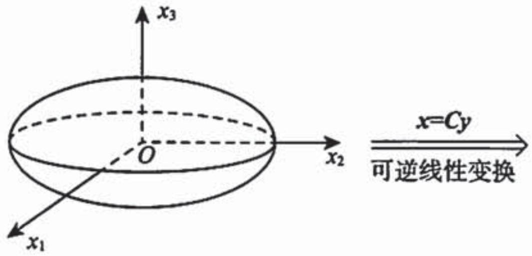


图 2-6-1　$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=$某常数

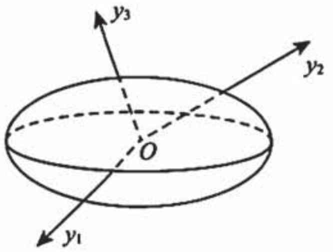


图 2-6-2　$g(\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y}=$某常数

上面这段分析让我们理解到：在二次型中，$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的合同，**就是指同一个二次型在可逆线性变换下的两个不同状态的联系.**

合同有以下三个性质：

(1)$\boldsymbol{A}\simeq\boldsymbol{A}$（反身性）；

(2)若 $\boldsymbol{A}\simeq\boldsymbol{B}$，则 $\boldsymbol{B}\simeq\boldsymbol{A}$（对称性）；

(3)若 $\boldsymbol{A}\simeq\boldsymbol{B}$，且 $\boldsymbol{B}\simeq\boldsymbol{C}$，则 $\boldsymbol{A}\simeq\boldsymbol{C}$（传递性）.

特别需要指出的是，由矩阵合同的定义知，若 $\boldsymbol{A}\simeq\boldsymbol{B}$，则有 $r(\boldsymbol{A})=r(\boldsymbol{B})$. 因此有重要结论：**可逆线性变换不会改变二次型的秩.** 顺便再说一句，由于在二次型中，二次型的矩阵都是对称矩阵，所以**和对称矩阵合同的矩阵也必是对称矩阵.** 这是因为若 $\boldsymbol{A}\simeq\boldsymbol{B}$，即存在可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$，其中 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，则

$$\boldsymbol{B}^{\mathrm{T}}=(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}.$$

3. 二次型的标准形、规范形

若二次型中只含有平方项，没有交叉项（即所有交叉项的系数全为零），即形如

$$d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$$

的二次型称为**标准形.**

若标准形中，系数 $d_i(i=1,2,\cdots,n)$ 仅为 $1,-1,0$，即形如 $x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2$ 的二次型称为**规范形.**

若二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 合同于标准形 $d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$（或合同于规范形 $x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2$），则称 $d_1x_1^2+d_2x_2^2+\cdots+d_nx_n^2$ 为 $f(\boldsymbol{x})$ 的**合同标准形**（则称 $x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2$ 为 $f(\boldsymbol{x})$ 的**合同规范形**）.

任何二次型均可通过配方法（作可逆线性变换）化成标准形及规范形，用矩阵语言表述：任何实对称矩阵 $\boldsymbol{A}$，必存在可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{\Lambda}$，其中

$$\underset{\text{（标准形）}}{\boldsymbol{\Lambda}=\begin{bmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{bmatrix}}\text{或}\underset{\text{（规范形）}}{\boldsymbol{\Lambda}=\begin{bmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&-1&&&&&\\&&&&\ddots&&&&\\&&&&&-1&&&\\&&&&&&0&&\\&&&&&&&\ddots&\\&&&&&&&&0\end{bmatrix}}.$$

任何二次型也可以通过正交变换化成标准形，用矩阵语言表达即是任何实对称矩阵 $\boldsymbol{A}$，一定存在正交矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda},$$

其中

$$\boldsymbol{\Lambda}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}.$$

在众多实际问题中，人们都会将二次型化成标准形、规范形，因为它们是“最佳形态”.

三、惯性定理

无论选取什么样的可逆线性变换，将二次型化成标准形或规范形，其正项个数 p，负项个数 q 都是不变的，p 称为**正惯性指数**，q 称为**负惯性指数**.

【注】 (1)若二次型的秩为 r，则 $r=p+q$，可逆线性变换不改变正、负惯性指数；

(2)两个二次型(或实对称矩阵)合同的充要条件是有相同的正、负惯性指数，或有相同的秩及正(或负)惯性指数.

四、正定二次型及其判别

1. 定义

n 元二次型 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. 若对任意的 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}\neq\boldsymbol{0}$，均有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$，则称 f 为**正定二次型**，称二次型的对应矩阵 $\boldsymbol{A}$ 为**正定矩阵**.

2. 二次型正定的充要条件

n 元二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定$\Leftrightarrow$对任意 $\boldsymbol{x}\neq\boldsymbol{0}$，有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$(定义)

$\Leftrightarrow f$ 的正惯性指数 $p=n$

$\Leftrightarrow$存在可逆矩阵 $\boldsymbol{D}$，使 $\boldsymbol{A}=\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}$

$\Leftrightarrow\boldsymbol{A}\simeq\boldsymbol{E}$

$\Leftrightarrow\boldsymbol{A}$ 的特征值 $\lambda_i>0(i=1,2,\cdots,n)$

$\Leftrightarrow\boldsymbol{A}$ 的全部顺序主子式均大于 0.

【注】 设 $\boldsymbol{A}=(a_{ij})_{n\times n}$，则

$$|\boldsymbol{A}_k|=\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & & \vdots\\ a_{k1} & a_{k2} & \cdots & a_{kk}\end{vmatrix}$$

称为 n 阶矩阵 $\boldsymbol{A}$ 的 k 阶顺序(或左上角)主子式. 当 k 取 $1,2,\cdots,n$ 时，就得到 $\boldsymbol{A}$ 的 n 个顺序主子式.

3. 二次型正定的必要条件

(1)$a_{ii}>0(i=1,2,\cdots,n)$；

(2)$|\boldsymbol{A}|>0$.

基础例题精解

一、用配方法化二次型为标准形

本节的理论依据与注意事项请看例 2.6.1 与例 2.6.2 后的注.

例 2.6.1 化二次型

$$f(x_1,x_2,x_3)=x_1^2+2x_1x_2+2x_1x_3-x_2^2-2x_2x_3-x_3^2$$

为标准形,并写出所作的可逆线性变换.

解 先对 x_1^2 及所有含有 x_1 的混合项 $2x_1x_2,2x_1x_3$ 配完全平方,有

$$\begin{aligned}f(x_1,x_2,x_3)&=(x_1+x_2+x_3)^2-x_2^2-x_3^2-2x_2x_3-x_2^2-2x_2x_3-x_3^2\\&=(x_1+x_2+x_3)^2-2x_2^2-4x_2x_3-2x_3^2.\end{aligned}$$

再对 $2x_2^2$ 及所有含 x_2 的混合项 $-4x_2x_3$ 配完全平方,有

$$f(x_1,x_2,x_3)=(x_1+x_2+x_3)^2-2(x_2+x_3)^2.$$

作线性变换

$$\begin{cases}y_1=x_1+x_2+x_3,\\y_2=x_2+x_3,\\y_3=x_3,\end{cases}\text{即}\begin{cases}x_1=y_1-y_2,\\x_2=y_2-y_3,\\x_3=y_3.\end{cases}$$

把二次型化为标准形 $f(x_1,x_2,x_3)\xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}}y_1^2-2y_2^2$.

将线性变换表示成矩阵形式,即

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y},$$

其中 $\boldsymbol{C}=\begin{bmatrix}1&-1&0\\0&1&-1\\0&0&1\end{bmatrix}$. 因 $|\boldsymbol{C}|=\begin{vmatrix}1&-1&0\\0&1&-1\\0&0&1\end{vmatrix}=1\neq0$,故所作变换是可逆线性变换.

【注】 (1)用配方法化二次型为标准形时,要遵循"将某个变量的平方项及与其有关的混合项一次配完,配成一个完全平方,减少一个未配完全平方的变量,使得总的平方项的项数小于等于变量个数". 目的是保证所用变换是可逆的.

(2)当总的完全平方项的项数小于变量个数时,例如本题是三元二次型,完全平方个数是 2 时,应视作

$$f(x_1,x_2,x_3)=(x_1+x_2+x_3)^2-2(x_2+x_3)^2+0x_3^2,$$

变换为

$$\begin{cases} y_1=x_1+x_2+x_3, \\ y_2=x_2+x_3, \\ y_3=x_3. \end{cases}$$

如变换为

$$\begin{cases} y_1=x_1+x_2+x_3, \\ y_2=x_2+x_3, \end{cases}$$

显然是不可逆的.

(3)配方法中配方次序可以不同,故所作可逆线性变换不唯一,标准形不唯一,但标准形中不为零的项数($r(f)$)、正项个数(正惯性指数)、负项个数(负惯性指数)是不变的.

(4)本题用矩阵表述:设 $\boldsymbol{A}=\begin{bmatrix}1 & 1 & 1\\ 1 & -1 & -1\\ 1 & -1 & -1\end{bmatrix}$,求可逆矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{\Lambda}$,并写出 $\boldsymbol{\Lambda}$.

例 2.6.2 将二次型 $f(x_1,x_2,x_3)=x_1x_2+x_1x_3-x_2x_3$ 化为规范形,并求所用的可逆线性变换.

解 令

$$\begin{cases} x_1=y_1+y_2, \\ x_2=y_1-y_2, \\ x_3=y_3, \end{cases}$$

则

$$\begin{aligned} f(x_1,x_2,x_3)&=y_1^2-y_2^2+y_1y_3+y_2y_3-y_1y_3+y_2y_3=y_1^2-y_2^2+2y_2y_3 \\ &=y_1^2-(y_2-y_3)^2+y_3^2. \end{aligned}$$

令

$$\begin{cases} z_1=y_1, \\ z_2=y_2-y_3, \\ z_3=y_3, \end{cases} \text{即} \begin{cases} y_1=z_1, \\ y_2=z_2+z_3, \\ y_3=z_3, \end{cases}$$

得二次型的规范形为

$$f(x_1,x_2,x_3)\xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{z}}z_1^2-z_2^2+z_3^2.$$

所用线性变换为

$$\begin{cases} x_1=y_1+y_2, \\ x_2=y_1-y_2, \\ x_3=y_3 \end{cases} \text{与} \begin{cases} y_1=z_1, \\ y_2=z_2+z_3, \\ y_3=z_3, \end{cases} \text{得} \begin{cases} x_1=z_1+z_2+z_3, \\ x_2=z_1-z_2-z_3, \\ x_3=z_3. \end{cases}$$

用矩阵表示为

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\ 1 & -1 & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}z_1\\ z_2\\ z_3\end{bmatrix}\xlongequal{\text{记}}\boldsymbol{C}\boldsymbol{z},$$

其中 $|C|=\begin{vmatrix}1&1&1\\1&-1&-1\\0&0&1\end{vmatrix}=-2\neq0$，故 $x=Cz$ 是可逆线性变换.

【注】 (1)用配方法化二次型为标准形、规范形，若有平方项，应将平方项及其有关混合项配成完全平方(如例 2.6.1). 没有平方项时，作可逆线性变换 $\begin{cases}x_1=y_1+y_2,\\x_2=y_1-y_2,\\x_3=y_3,\end{cases}$ 使其出现平方项，然后再配完全平方(如本例). 这样，任何二次型总可以用配方法化成标准形、规范形，且所作变换均是可逆线性变换. 用矩阵的语言表述：任何实对称矩阵 A，必存在可逆矩阵 C，使得 $C^TAC=\Lambda$，其中 Λ 是对角矩阵，或是对角元素只取 $1,-1,0$ 的对角矩阵.

(2)配方是中学已经掌握的内容，并不难，但为我们提供了①所作的可逆线性变换；②与 A 合同的对角矩阵；③二次型(或 A)的秩；④正、负惯性指数；⑤是否正定等二次型的重要信息，故在求上述问题时，别忘了有一招是配方法.

二、用正交变换化二次型为标准形

这是考研重点，几乎每年都考大题，事实上，当写出二次型的矩阵 A 时，便成为“将 A 用正交矩阵相似对角化”的问题了，也就是第 5 讲“基础例题精解”的“六”.

例 2.6.3 用正交变换化二次型

$$f(x_1,x_2,x_3)=2x_1^2+5x_2^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$$

为标准形，并求所作的正交变换.

解 二次型的对应矩阵为

$$A=\begin{bmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{bmatrix},$$

由特征方程

$$|\lambda E-A|=\begin{vmatrix}\lambda-2&-2&2\\-2&\lambda-5&4\\2&4&\lambda-5\end{vmatrix}=\begin{vmatrix}\lambda-2&-2&0\\-2&\lambda-5&\lambda-1\\2&4&\lambda-1\end{vmatrix}=\begin{vmatrix}\lambda-2&-2&0\\-4&\lambda-9&0\\2&4&\lambda-1\end{vmatrix}$$

$$=(\lambda-1)(\lambda^2-11\lambda+10)=(\lambda-1)^2(\lambda-10)=0,$$

知 A 有特征值 $\lambda_1=\lambda_2=1,\lambda_3=10$.

当 $\lambda_1=\lambda_2=1$ 时，有

$$(E-A)x=\begin{bmatrix}-1&-2&2\\-2&-4&4\\2&4&-4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\mathbf{0}, \qquad (*)$$

其同解方程组为 $-x_1-2x_2+2x_3=0$，解得基础解系为 $\boldsymbol{\xi}_1=[-2,1,0]^{\mathrm{T}}$，$\boldsymbol{\xi}_2=[2,0,1]^{\mathrm{T}}$.

当 $\lambda_3=10$ 时，有

$$(10\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\begin{bmatrix}8&-2&2\\-2&5&4\\2&4&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\boldsymbol{0}, \qquad (**)$$

解得基础解系为 $\boldsymbol{\xi}_3=[1,2,-2]^{\mathrm{T}}$.

将 $\lambda_1=\lambda_2=1$ 对应的特征向量 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 标准正交化.

取

$$\boldsymbol{\eta}_1=\boldsymbol{\xi}_1=[-2,1,0]^{\mathrm{T}},$$

$$\boldsymbol{\eta}_2=\boldsymbol{\xi}_2-\frac{(\boldsymbol{\xi}_2,\boldsymbol{\eta}_1)}{(\boldsymbol{\eta}_1,\boldsymbol{\eta}_1)}\boldsymbol{\eta}_1=[2,0,1]^{\mathrm{T}}-\frac{-4}{5}[-2,1,0]^{\mathrm{T}}=\left[\frac{2}{5},\frac{4}{5},1\right]^{\mathrm{T}},$$

不妨取 $\boldsymbol{\eta}_2=[2,4,5]^{\mathrm{T}}$.

再将 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\xi}_3$ 单位化

$$\boldsymbol{\eta}_1^{\circ}=\begin{bmatrix}\frac{-2}{\sqrt{5}}\\\frac{1}{\sqrt{5}}\\0\end{bmatrix},\quad \boldsymbol{\eta}_2^{\circ}=\begin{bmatrix}\frac{2}{3\sqrt{5}}\\\frac{4}{3\sqrt{5}}\\\frac{5}{3\sqrt{5}}\end{bmatrix},\quad \boldsymbol{\eta}_3^{\circ}=\begin{bmatrix}\frac{1}{3}\\\frac{2}{3}\\\frac{-2}{3}\end{bmatrix},$$

得正交矩阵 $\boldsymbol{Q}=[\boldsymbol{\eta}_1^{\circ},\boldsymbol{\eta}_2^{\circ},\boldsymbol{\eta}_3^{\circ}]$，令 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$，则原二次型化为标准形

$$f(x_1,x_2,x_3)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\xlongequal{\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{y}=y_1^2+y_2^2+10y_3^2,$$

其中正交变换为

$$\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}=\begin{bmatrix}\frac{-2}{\sqrt{5}}&\frac{2}{3\sqrt{5}}&\frac{1}{3}\\\frac{1}{\sqrt{5}}&\frac{4}{3\sqrt{5}}&\frac{2}{3}\\0&\frac{5}{3\sqrt{5}}&\frac{-2}{3}\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}.$$

【注】 (1)正交变换只能化二次型为标准形，不能化为规范形(除非特征值都属于{1，-1，0}).

(2)正交变换不唯一，但经正交变换所得标准形是唯一的，求得特征值后即可得标准形为

$$\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.$$

(3)解方程组(* *)的技巧：

$\lambda=10$ 是单根，$\boldsymbol{A}$ 是实对称矩阵，必可相似对角化，故 $r(10\boldsymbol{E}-\boldsymbol{A})=2$，必有一个方程是多余方程. 因 $10\boldsymbol{E}-\boldsymbol{A}$ 中三行均不成比例，任一方程均是另两方程的线性组合，故可去除任一方程，例如去掉第一个方程，剩下的仍是同解方程组，

$$\begin{bmatrix}-2&5&4\\2&4&5\\0&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}-2&5&4\\0&9&9\\0&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}-2&5&4\\0&1&1\\0&0&0\end{bmatrix}\longrightarrow\begin{bmatrix}-2&0&-1\\0&1&1\\0&0&0\end{bmatrix},$$

解得 $\boldsymbol{\xi}_3=[1,2,-2]^{\mathrm{T}}$.

(4) $\lambda=1$ 是重根时，方程组(*)有两个线性无关解，为避免后面的正交化计算，可在解方程组(*)时，同时考虑正交化. $\lambda_1=\lambda_2=1$，对应特征向量应满足

$$-x_1-2x_2+2x_3=0.$$

取 $\xi_1=[-2,1,0]^{\mathrm{T}}$，再求 ξ_2 时，可先令 ξ_2 和 ξ_1 正交，取 $\xi_2=[1,2,a]^{\mathrm{T}}$（$(\xi_1,\xi_2)=0$，$a$ 待定），再代入方程以确定 a，即 $-1-4+2a=0$，得 $a=\dfrac{5}{2}$，则 $\xi_2=\left[1,2,\dfrac{5}{2}\right]^{\mathrm{T}}$，取整数得 $\xi_2'=[2,4,5]^{\mathrm{T}}$.

(5) 计算过程还可作如下验算：① $\sum\limits_{i=1}^{3}\lambda_i=\sum\limits_{i=1}^{3}a_{ii}$，$|\boldsymbol{A}|=\prod\limits_{i=1}^{3}\lambda_i$；② 当 $\lambda_1\neq\lambda_3$ 时，$(\xi_1,\xi_3)=0$.

(6) 本题用矩阵语言表述即：设 $\boldsymbol{A}=\begin{bmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{bmatrix}$，求正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$，并写出对角矩阵 $\boldsymbol{\Lambda}$.

三、合同矩阵与合同二次型

合同矩阵与合同二次型本质上一样，关键是掌握判别矩阵合同的方法.

例 2.6.4 设 $\boldsymbol{A}=\begin{bmatrix}1&2&0\\2&1&0\\0&0&1\end{bmatrix}$，则下列矩阵中与 $\boldsymbol{A}$ 合同的是(　　).

(A) $\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$　(B) $\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}$　(C) $\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}$　(D) $\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}$

解　应选(B).

方法一　$\boldsymbol{A}$ 对应的二次型为

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+4x_1x_2,$$

经配方法化为

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+4x_1x_2=(x_1+2x_2)^2-3x_2^2+x_3^2,$$

知 $r(f)=3$，正惯性指数 $p=2$，与选项(B)相同，故应选(B).

方法二　求 $\boldsymbol{A}$ 的特征值. 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-2&0\\-2&\lambda-1&0\\0&0&\lambda-1\end{vmatrix}=(\lambda-1)(\lambda^2-2\lambda-3)=(\lambda-1)(\lambda-3)(\lambda+1)=0,$$

知 $\boldsymbol{A}$ 的特征值为 $\lambda_1=1,\lambda_2=3,\lambda_3=-1$，且 $\boldsymbol{A}$ 对应的二次型在正交变换下化标准形为 $y_1^2+3y_2^2-y_3^2$，与选项(B)的正、负惯性指数相同，即

$$\boldsymbol{A}\simeq\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix},$$

故应选(B).

四、具体型二次型的正定问题

本部分的理论依据请参看“基础内容精讲”中“四”，对于例 2.6.5 的五种方法，请认真研读并加以总结.

例 2.6.5 判别二次型

$$f(x_1,x_2,x_3)=2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$$

的正定性.

解
$$f(x_1,x_2,x_3)=[x_1,x_2,x_3]\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},$$

故 $f(x_1,x_2,x_3)$ 的对应矩阵是

$$\boldsymbol{A}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}.$$

方法一 判别对应矩阵 $\boldsymbol{A}$ 的各阶顺序主子式是否大于零.

由
$$2>0,\quad \begin{vmatrix}2&1\\1&2\end{vmatrix}=3>0,$$

$$|\boldsymbol{A}|=\begin{vmatrix}2&1&1\\1&2&1\\1&1&2\end{vmatrix}=4\begin{vmatrix}1&1&1\\1&2&1\\1&1&2\end{vmatrix}=4\begin{vmatrix}1&1&1\\0&1&0\\0&0&1\end{vmatrix}=4>0,$$

故 $\boldsymbol{A}$ 正定，f 是正定二次型.

方法二 判别对应矩阵 $\boldsymbol{A}$ 的特征值是否全部大于零.

由
$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&-1&-1\\-1&\lambda-2&-1\\-1&-1&\lambda-2\end{vmatrix}=(\lambda-4)\begin{vmatrix}1&-1&-1\\1&\lambda-2&-1\\1&-1&\lambda-2\end{vmatrix}$$

$$=(\lambda-4)\begin{vmatrix}1&-1&-1\\0&\lambda-1&0\\0&0&\lambda-1\end{vmatrix}=(\lambda-4)(\lambda-1)^2=0,$$

知 $\boldsymbol{A}$ 的特征值 $\lambda_1=4,\lambda_2=\lambda_3=1$ 全部大于零，故 $\boldsymbol{A}$ 正定，f 是正定二次型.

方法三 利用配方法化为标准形，判别 f 的正惯性指数 p 是否等于 n(未知量个数).

$$\begin{aligned}f&=2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3\\&=2\left(x_1+\frac{x_2}{2}+\frac{x_3}{2}\right)^2+\frac{3}{2}x_2^2+\frac{3}{2}x_3^2+x_2x_3\\&=2\left(x_1+\frac{x_2}{2}+\frac{x_3}{2}\right)^2+\frac{3}{2}\left(x_2+\frac{1}{3}x_3\right)^2+\frac{4}{3}x_3^2,\end{aligned}$$

由二次型的标准形知，f 的正惯性指数 $p=3=n$，故 f 是正定二次型.

方法四 用定义，验证是否对任意的 $\boldsymbol{x}=[x_1,x_2,x_3]^{\mathrm{T}}\neq\boldsymbol{0}$，有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$. 将二次型配成完全平方和（注意：这里配完全平方和的做法与方法三中的配方法不同，这里没有要求将某平方项及与其有关的混合项一次配完），即

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3\\&=(x_1+x_2)^2+(x_2+x_3)^2+(x_3+x_1)^2,\end{aligned}$$

故有 $f\geqslant 0$，且

$$f=0\Leftrightarrow\begin{cases}x_1+x_2=0,\\x_2+x_3=0,\\x_3+x_1=0.\end{cases}\tag{$*$}$$

因方程组（∗）的系数行列式

$$\begin{vmatrix}1&1&0\\0&1&1\\1&0&1\end{vmatrix}=2\neq 0,$$

故方程组（∗）只有零解，故 $f=0\Leftrightarrow x_1=x_2=x_3=0$.

因此，当 $\boldsymbol{x}=[x_1,x_2,x_3]^{\mathrm{T}}\neq\boldsymbol{0}$ 时，恒有 $f(x_1,x_2,x_3)>0$，从而知 f 是正定二次型.

方法五 由方法四知，$f(x_1,x_2,x_3)$ 可配完全平方和

$$f(x_1,x_2,x_3)=(x_1+x_2)^2+(x_2+x_3)^2+(x_3+x_1)^2,$$

且利用内积可表示成如下矩阵形式

$$\begin{aligned}f(x_1,x_2,x_3)&=[x_1+x_2,x_2+x_3,x_3+x_1]\begin{bmatrix}x_1+x_2\\x_2+x_3\\x_3+x_1\end{bmatrix}\\&=[x_1,x_2,x_3]\begin{bmatrix}1&0&1\\1&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\\&\xlongequal{\text{记}}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x}\xlongequal{\text{记}}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x},\end{aligned}$$

其中 $\boldsymbol{A}=\boldsymbol{D}^{\mathrm{T}}\boldsymbol{D}$，$|\boldsymbol{D}|=\begin{vmatrix}1&1&0\\0&1&1\\1&0&1\end{vmatrix}=2\neq 0$，$\boldsymbol{D}$ 可逆，故 $\boldsymbol{A}$ 合同于单位矩阵，是正定矩阵，$f(x_1,x_2,x_3)$ 是正定二次型.

【注】 判别二次型是否正定，以上五种方法是利用定义及 4 个充要条件. 对具体的数值二次型（或实对称矩阵）主要用方法一，即判别顺序主子式是否全部大于零，其他问题要具体分析并根据你对这些充要条件的熟练程度作出选择.

例 2.6.6 下列矩阵中，为正定矩阵的是（　　）.

(A) $\begin{bmatrix}4&2&3\\2&2&1\\3&1&-1\end{bmatrix}$　(B) $\begin{bmatrix}1&2&2\\2&1&2\\2&2&4\end{bmatrix}$　(C) $\begin{bmatrix}1&2&2\\2&1&0\\2&0&2\end{bmatrix}$　(D) $\begin{bmatrix}3&0&1\\0&2&1\\1&1&2\end{bmatrix}$

解 应选(D).

利用正定矩阵的必要条件及顺序主子式判别.

(A)中矩阵 $a_{33}=-1<0$,不正定.

(B),(C)中矩阵二阶顺序主子式 $\begin{vmatrix}1&2\\2&1\end{vmatrix}=-3<0$,不正定.

由排除法,应选(D).

对 $\begin{bmatrix}3&0&1\\0&2&1\\1&1&2\end{bmatrix}$,各阶顺序主子式为

$$\Delta_1=3>0,\quad \Delta_2=6>0,\quad \Delta_3=\begin{vmatrix}3&0&1\\0&2&1\\1&1&2\end{vmatrix}=\begin{vmatrix}3&-2&1\\0&0&1\\1&-3&2\end{vmatrix}=-\begin{vmatrix}3&-2\\1&-3\end{vmatrix}=7>0,$$

故其是正定矩阵,应选(D).

例 2.6.7 $f(x_1,x_2,x_3)=x_1^2+2tx_1x_2-2x_1x_3+x_2^2+4x_2x_3+5x_3^2$ 正定,则 t 的取值范围是__________.

解 应填 $-\dfrac{4}{5}<t<0$.

二次型的对应矩阵为

$$\boldsymbol{A}=\begin{bmatrix}1&t&-1\\t&1&2\\-1&2&5\end{bmatrix}.$$

$\boldsymbol{A}$ 的各阶顺序主子式为

$$\Delta_1=1,$$

$$\Delta_2=\begin{vmatrix}1&t\\t&1\end{vmatrix}=1-t^2,$$

$$\Delta_3=\begin{vmatrix}1&t&-1\\t&1&2\\-1&2&5\end{vmatrix}=\begin{vmatrix}1&0&0\\t&1-t^2&2+t\\-1&2+t&4\end{vmatrix}=4(1-t^2)-(2+t)^2=-5t^2-4t.$$

于是,f 正定$\Leftrightarrow\boldsymbol{A}$ 正定$\Leftrightarrow\begin{cases}1-t^2>0,\\-t(5t+4)>0\end{cases}\Leftrightarrow\begin{cases}-1<t<1,\\-\dfrac{4}{5}<t<0\end{cases}\Leftrightarrow-\dfrac{4}{5}<t<0.$

五、抽象型二次型的正定问题

例 2.6.8 设实二次型 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,其中 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 是实对称矩阵,则 $f(x_1,x_2,\cdots,x_n)$ 为正定二次型的充要条件是(　　).

(A)$\boldsymbol{A}^*$ 是正定矩阵　　(B)$\boldsymbol{A}^{-1}$ 是正定矩阵

(C)$f(x_1,x_2,\cdots,x_n)$ 的负惯性指数为零　　(D)存在 n 阶实矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$

解 应选(B).

因 $\boldsymbol{A}$ 正定,且 $\boldsymbol{A}$ 为实对称矩阵,有 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,且 $\boldsymbol{A}$ 的特征值 $\lambda>0$. 两边求逆,得 $(\boldsymbol{A}^{\mathrm{T}})^{-1}=\boldsymbol{A}^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$,所

以 $\boldsymbol{A}^{-1}$ 是实对称矩阵，且其特征值 $\frac{1}{\lambda}>0$，故 $\boldsymbol{A}^{-1}$ 是正定矩阵.

反之，因 $\boldsymbol{A}^{-1}$ 正定，故 $(\boldsymbol{A}^{-1})^{\mathrm{T}}=\boldsymbol{A}^{-1}$，且 $\boldsymbol{A}^{-1}$ 的特征值 $\mu>0$. 两边求逆，得 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，所以 $\boldsymbol{A}$ 为实对称矩阵，且 $\boldsymbol{A}$ 的特征值 $\frac{1}{\mu}>0$，故 $\boldsymbol{A}$ 是正定矩阵.

综上，$\boldsymbol{A}$ 正定 $\Leftrightarrow \boldsymbol{A}^{-1}$ 正定，应选(B).

(A)$\boldsymbol{A}^*$ 正定是 $\boldsymbol{A}$ 正定的必要条件，但不充分，如

$$\boldsymbol{A}=\begin{bmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}=-\boldsymbol{E},\quad \boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}=-1(-\boldsymbol{E})^{-1}=\boldsymbol{E}^{-1}=\boldsymbol{E},$$

$\boldsymbol{A}^*$ 是正定矩阵，但 $\boldsymbol{A}$ 不是正定矩阵，故(A)不成立；

(C)f 的负惯性指数为零，但正惯性指数 p 不一定为 n，即可能 $p=r(f)<n$，故(C)不成立；

(D)$\boldsymbol{A}$ 正定 $\Leftrightarrow$ 存在 n 阶实可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$，但(D)中 $\boldsymbol{C}$ 没有要求可逆，故(D)不成立. 例如取

$$\boldsymbol{C}=\begin{bmatrix}1&1\\0&0\end{bmatrix},$$

此时，

$$\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\begin{bmatrix}1&0\\1&0\end{bmatrix}\begin{bmatrix}1&1\\0&0\end{bmatrix}=\begin{bmatrix}1&1\\1&1\end{bmatrix},$$

$|\boldsymbol{A}|=0$，$\boldsymbol{A}$ 不正定.

例 2.6.9 已知 $\boldsymbol{A}$ 是 n 阶方阵，有特征值 $\lambda_1<\lambda_2<\cdots<\lambda_n$，问：

(1)a 满足什么条件时，$a\boldsymbol{E}+\boldsymbol{A}$ 正定；

(2)正数 ε 满足什么条件时，$\boldsymbol{E}+\varepsilon\boldsymbol{A}$ 正定.

解 (1)$\boldsymbol{A}$ 有特征值 $\lambda_i(i=1,2,\cdots,n)$，则 $a\boldsymbol{E}+\boldsymbol{A}$ 有特征值 $a+\lambda_i(i=1,2,\cdots,n)$，由题设条件，知

$$a+\lambda_1<a+\lambda_2<\cdots<a+\lambda_n.$$

当 $a+\lambda_1>0$，即 $a>-\lambda_1$ 时，有 $a+\lambda_i>0(i=1,2,\cdots,n)$，$a\boldsymbol{E}+\boldsymbol{A}$ 是正定矩阵.

(2)$\boldsymbol{A}$ 有特征值 λ_i，则 $\boldsymbol{E}+\varepsilon\boldsymbol{A}$ 有特征值 $1+\varepsilon\lambda_i(i=1,2,\cdots,n)$. 因 $\varepsilon>0$，由题设条件，知

$$1+\varepsilon\lambda_1<1+\varepsilon\lambda_2<\cdots<1+\varepsilon\lambda_n.$$

当 $\lambda_1\geqslant 0$ 时，对任意的 $\varepsilon>0$，$\boldsymbol{E}+\varepsilon\boldsymbol{A}$ 正定；

当 $\lambda_1<0$ 时，只需 $1+\varepsilon\lambda_1>0$，即 $\varepsilon\lambda_1>-1$，$0<\varepsilon<\frac{-1}{\lambda_1}$，就有 $1+\varepsilon\lambda_i>0(i=1,2,\cdots,n)$，从而 $\boldsymbol{E}+\varepsilon\boldsymbol{A}$ 正定.

基础习题精练

习 题

2.6.1 设二次型 $f(x_1,x_2,x_3)=2x_1^2+4x_2^2+ax_3^2+4x_1x_2+4x_1x_3+4x_2x_3$ 的秩为 2，则参数 $a=$ ________.

2.6.2 二次型 $f(x_1,x_2,x_3)=2x_2^2+2x_3^2+4x_1x_2-4x_1x_3+8x_2x_3$ 的规范形为（　　）.

(A) $y_1^2+y_2^2+y_3^2$　　(B) $y_1^2+y_2^2-y_3^2$　　(C) $y_1^2-y_2^2-y_3^2$　　(D) $y_1^2-y_2^2$

2.6.3 设二次型

$$f(x_1,x_2,x_3)=kx_1^2+2x_2^2-2x_3^2+2mx_1x_3\ (m>0),$$

其中二次型的矩阵的特征值之和为 1，特征值之积为 -12.

(1)求 k,m；

(2)用正交变换化二次型为标准形，并求所作的正交变换及对应的正交矩阵.

2.6.4 设 $f(x_1,x_2)=x_1^2+3x_2^2+4x_1x_2$，则 $f(x_1,x_2)$ 对应的矩阵与下列矩阵不合同的是（　　）.

(A) $\begin{bmatrix}-1&1\\1&2\end{bmatrix}$　　(B) $\begin{bmatrix}1&2\\2&1\end{bmatrix}$　　(C) $\begin{bmatrix}0&-1\\-1&1\end{bmatrix}$　　(D) $\begin{bmatrix}1&1\\1&2\end{bmatrix}$

2.6.5 已知 $\boldsymbol{A}=\begin{bmatrix}1&0&0\\0&-4&0\\0&0&\frac{1}{9}\end{bmatrix}$ 合同于 $\boldsymbol{\Lambda}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}$，即存在可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{\Lambda}$，其中 $\boldsymbol{C}=$________.

2.6.6 已知 $\boldsymbol{A}=\begin{bmatrix}k&1&0\\1&1&0\\0&0&k+2\end{bmatrix}$ 正定，则 k 应满足条件________.

2.6.7 已知二次型 $f(x_1,x_2,x_3)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 在正交变换 $\boldsymbol{x}=\boldsymbol{Qy}$ 下的标准形为 $y_1^2+y_2^2$，且 $\boldsymbol{Q}$ 的第 3 列为 $\left[\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right]^{\mathrm{T}}$.

(1)求矩阵 $\boldsymbol{A}$；

(2)证明 $\boldsymbol{A}+\boldsymbol{E}$ 为正定矩阵，其中 $\boldsymbol{E}$ 为 3 阶单位矩阵.

解答

2.6.1 2　解　f 对应的矩阵 $\boldsymbol{A}=\begin{bmatrix}2&2&2\\2&4&2\\2&2&a\end{bmatrix}$，对 $\boldsymbol{A}$ 作初等行变换，

$$\boldsymbol{A}=\begin{bmatrix}2&2&2\\2&4&2\\2&2&a\end{bmatrix}\longrightarrow\begin{bmatrix}2&2&2\\0&2&0\\0&0&a-2\end{bmatrix},$$

于是 $r(\boldsymbol{A})=r(f)=2\Leftrightarrow a=2$.

2.6.2 (B)　解

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_2^2+2x_3^2+4x_1x_2-4x_1x_3+8x_2x_3=2(x_1+x_2+2x_3)^2-2x_1^2-6x_3^2-12x_1x_3\\&=2(x_1+x_2+2x_3)^2-2(x_1+3x_3)^2+12x_3^2,\end{aligned}$$

故知 f 的规范形为 $y_1^2+y_2^2-y_3^2$.

或
$$|\lambda E-A|=\begin{vmatrix}\lambda & -2 & 2\\ -2 & \lambda-2 & -4\\ 2 & -4 & \lambda-2\end{vmatrix}=(\lambda-6)(\lambda+4)(\lambda-2)=0,$$

因 A 的特征值为 $\lambda_1=6,\lambda_2=2,\lambda_3=-4$，故 f 的规范形为 $y_1^2+y_2^2-y_3^2$. 故应选(B).

2.6.3 解 (1)二次型 f 的对应矩阵为
$$A=\begin{bmatrix}k & 0 & m\\ 0 & 2 & 0\\ m & 0 & -2\end{bmatrix}.$$

设 A 的特征值为 $\lambda_1,\lambda_2,\lambda_3$，由题设条件，得
$$\lambda_1+\lambda_2+\lambda_3=k+2-2=1,$$
$$\lambda_1\lambda_2\lambda_3=|A|=-4k-2m^2=-12,$$

解得 $k=1,m=\pm2$，已知 $m>0$，故 $m=2$，即 $k=1,m=2$.

(2)由矩阵 A 的特征方程
$$|\lambda E-A|=\begin{vmatrix}\lambda-1 & 0 & -2\\ 0 & \lambda-2 & 0\\ -2 & 0 & \lambda+2\end{vmatrix}=(\lambda-2)(\lambda^2+\lambda-6)=(\lambda-2)^2(\lambda+3)=0,$$

知 A 的特征值为 $\lambda_1=\lambda_2=2,\lambda_3=-3$，故二次型的标准形为 $2y_1^2+2y_2^2-3y_3^2$.

对 $\lambda_1=\lambda_2=2$，由
$$(2E-A)x=\begin{bmatrix}1 & 0 & -2\\ 0 & 0 & 0\\ -2 & 0 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\mathbf{0},$$

解得基础解系为 $\xi_1=[0,1,0]^T,\xi_2=[2,0,1]^T$.

对 $\lambda_3=-3$，由
$$(-3E-A)x=\begin{bmatrix}-4 & 0 & -2\\ 0 & -5 & 0\\ -2 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\mathbf{0},$$

解得基础解系为 $\xi_3=[1,0,-2]^T$.

因 ξ_1,ξ_2,ξ_3 已两两正交，故只需单位化，得
$$\xi_1^\circ=[0,1,0]^T,\quad \xi_2^\circ=\left[\frac{2}{\sqrt5},0,\frac{1}{\sqrt5}\right]^T,\quad \xi_3^\circ=\left[\frac{1}{\sqrt5},0,\frac{-2}{\sqrt5}\right]^T,$$

则所求正交矩阵
$$Q=[\xi_1^\circ,\xi_2^\circ,\xi_3^\circ]=\begin{bmatrix}0 & \frac{2}{\sqrt5} & \frac{1}{\sqrt5}\\ 1 & 0 & 0\\ 0 & \frac{1}{\sqrt5} & \frac{-2}{\sqrt5}\end{bmatrix},$$

故 $x=Qy$ 即为所求正交变换，且有
$$f=x^TAx\xlongequal{x=Qy}y^TQ^TAQy=y^T\begin{bmatrix}2 & & \\ & 2 & \\ & & -3\end{bmatrix}y=2y_1^2+2y_2^2-3y_3^2.$$

2.6.4 (D) **解** 由题设,二次型为 $f(x_1,x_2)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=x_1^2+3x_2^2+4x_1x_2=(x_1+2x_2)^2-x_2^2$.

(A)对应二次型为 $f=\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}-1&1\\1&2\end{bmatrix}\boldsymbol{x}=-x_1^2+2x_2^2+2x_1x_2=-(x_1-x_2)^2+3x_2^2$;

(B)对应二次型为 $f=\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}1&2\\2&1\end{bmatrix}\boldsymbol{x}=x_1^2+x_2^2+4x_1x_2=(x_1+2x_2)^2-3x_2^2$;

(C)对应二次型为 $f=\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0&-1\\-1&1\end{bmatrix}\boldsymbol{x}=x_2^2-2x_1x_2=(x_2-x_1)^2-x_1^2$;

(D)对应二次型为 $f=\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}1&1\\1&2\end{bmatrix}\boldsymbol{x}=x_1^2+2x_2^2+2x_1x_2=(x_1+x_2)^2+x_2^2$.

故应选(D).

2.6.5 $\begin{bmatrix}1&0&0\\0&0&\frac{1}{2}\\0&3&0\end{bmatrix}$ **解** 设 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=[x_1,x_2,x_3]\begin{bmatrix}1&0&0\\0&-4&0\\0&0&\frac{1}{9}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_1^2-4x_2^2+\frac{1}{9}x_3^2$.

令 $\begin{cases}x_1=y_1,\\x_2=\frac{1}{2}y_3,\\x_3=3y_2,\end{cases}$ 即 $\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&\frac{1}{2}\\0&3&0\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\xlongequal{\text{记}}\boldsymbol{C}\boldsymbol{y}$,得

$$f=[x_1,x_2,x_3]\begin{bmatrix}1&&\\&-4&\\&&\frac{1}{9}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

$$\xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}}[y_1,y_2,y_3]\begin{bmatrix}1&0&0\\0&0&3\\0&\frac{1}{2}&0\end{bmatrix}\begin{bmatrix}1&&\\&-4&\\&&\frac{1}{9}\end{bmatrix}\begin{bmatrix}1&0&0\\0&0&\frac{1}{2}\\0&3&0\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

$$=[y_1,y_2,y_3]\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix},$$

故 $\boldsymbol{C}=\begin{bmatrix}1&0&0\\0&0&\frac{1}{2}\\0&3&0\end{bmatrix}$,使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{\Lambda}=\begin{bmatrix}1&&\\&1&\\&&-1\end{bmatrix}$.

【注】 有读者不理解为何要使用变换 $\begin{cases}x_1=y_1,\\x_2=\frac{1}{2}y_3,\\x_3=3y_2,\end{cases}$ 为何是 $x_2=\frac{1}{2}y_3$ 而不是 $x_2=\frac{1}{2}y_2$? 这里大家要注意变换的目的是将 $\boldsymbol{A}$ 合同于对角矩阵 $\boldsymbol{\Lambda}$. 而 $\boldsymbol{\Lambda}$ 有两个特点:一是对角元素的绝对值为1;二是对角元素的符号分别是正、正、负,先正后负的顺序决定了必须是 $x_2=\frac{1}{2}y_3$,$x_3=3y_2$.

2.6.6 $k>1$ **解**

$$\Delta_1=k>0;\quad \Delta_2=\begin{vmatrix} k & 1 \\ 1 & 1 \end{vmatrix}=k-1>0,\quad k>1;$$

$$\Delta_3=\begin{vmatrix} k & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & k+2 \end{vmatrix}=(k+2)(k-1)>0,\quad k<-2 \text{ 或 } k>1.$$

取公共部分 $k>1$，此时有 $\Delta_1>0,\Delta_2>0,\Delta_3>0$，故 $\boldsymbol{A}$ 正定.

2.6.7 (1)**解** 因为二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 在正交变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$ 下的标准形为 $y_1^2+y_2^2$，所以其系数 1，1，0 就是矩阵 $\boldsymbol{A}$ 的特征值，即

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix},$$

且矩阵 $\boldsymbol{Q}$ 的第 3 列就是属于特征值 0 的特征向量.

设$[x_1,x_2,x_3]^{\mathrm{T}}$ 为 $\boldsymbol{A}$ 的属于特征值 1 的特征向量. 由于实对称矩阵属于不同特征值的特征向量是正交的，故有

$$[1,0,1]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=0,$$

即 $x_1+x_3=0$，解得 $\boldsymbol{\xi}_2=\left[\frac{\sqrt{2}}{2},0,-\frac{\sqrt{2}}{2}\right]^{\mathrm{T}}$，$\boldsymbol{\xi}_3=[0,1,0]^{\mathrm{T}}$，即为属于特征值 1 的两个正交单位特征向量. 以 $\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 分别为 $\boldsymbol{Q}$ 的第 1，2 列(或第 2，1 列)得到

$$\boldsymbol{Q}=\begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \text{ 或 } \boldsymbol{Q}=\begin{bmatrix} 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix},$$

并有 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}$. 从而得

$$\boldsymbol{A}=\boldsymbol{Q}\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}\boldsymbol{Q}^{\mathrm{T}}=\frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

(2)**证明** 因 $\boldsymbol{A}$ 的特征值为 1，1，0，所以矩阵 $\boldsymbol{A}+\boldsymbol{E}$ 的特征值为 2，2，1. 又 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，则 $\boldsymbol{A}+\boldsymbol{E}$ 为实对称矩阵，故 $\boldsymbol{A}+\boldsymbol{E}$ 是正定矩阵(实对称矩阵正定的一个充要条件是其所有特征值均为正数).

【注】 第(2)问也可计算 $\boldsymbol{A}+\boldsymbol{E}$ 的顺序主子式 $\Delta_1=\frac{3}{2}>0,\Delta_2=3>0,\Delta_3=4>0$，故 $\boldsymbol{A}+\boldsymbol{E}$ 正定.

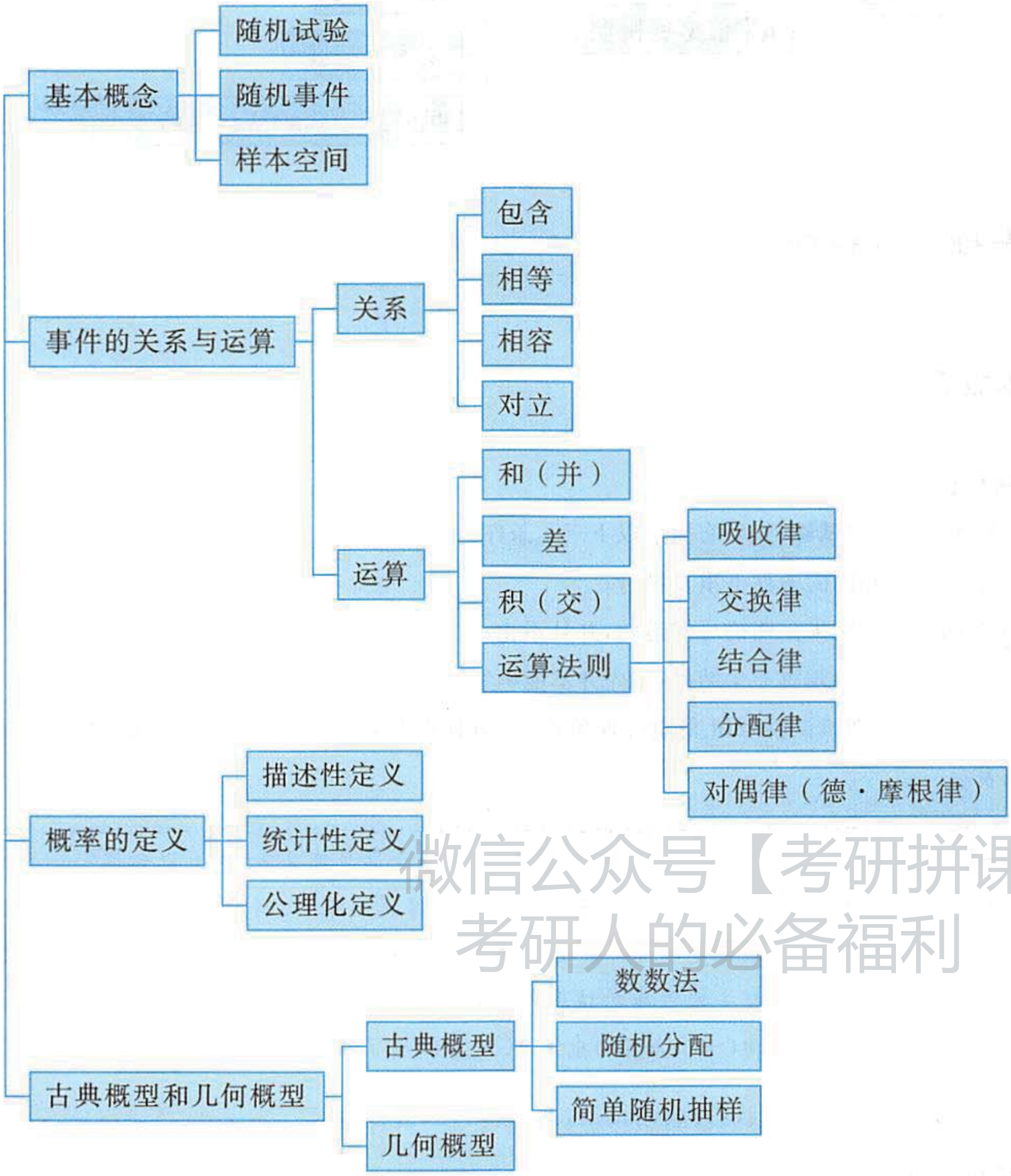
第三部分

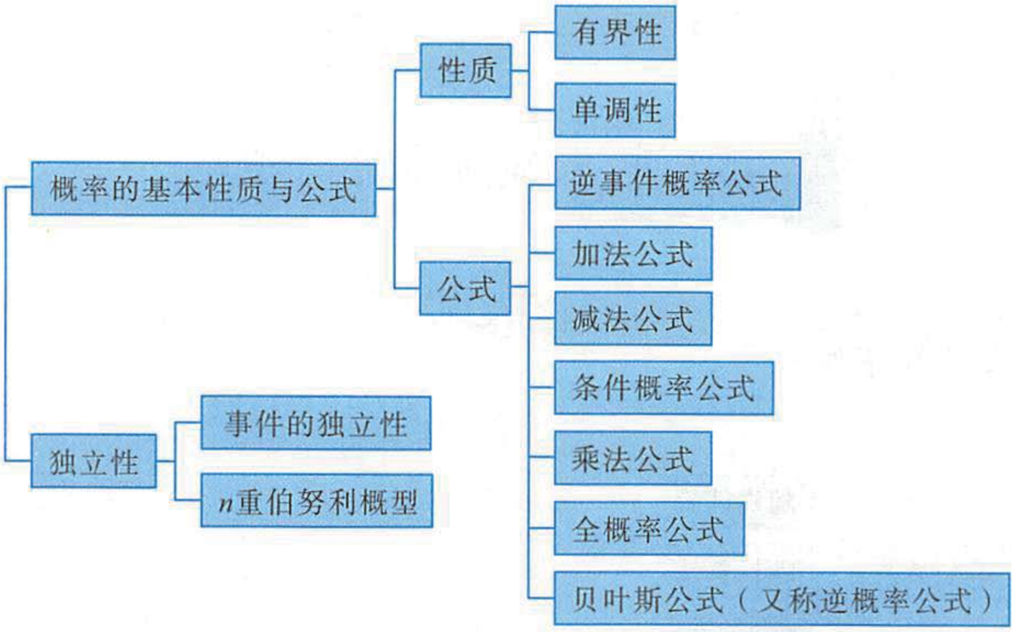
概率论与数理统计

（仅数学一、数学三要求）

第1讲 随机事件与概率

基础知识结构





基础内容精讲

一、基本概念

1. 随机试验

称一个试验为**随机试验**，如果它满足以下三个条件：

(1)试验可以在相同的条件下重复进行；

(2)试验所有可能结果是明确可知道的，并且不止一个；

(3)每一次试验会出现哪一个结果，事先并不能确定.

我们是通过研究随机试验来研究随机现象的，为方便起见，将随机试验简称为**试验**，并用字母 E 或 $E_1,E_2,\cdots$表示.

> **【注】** 在不少情况下，不能确切知道某一随机试验的全部可能结果，但可以知道它不超出某个范围. 这时，也可以用这个范围来作为该试验的全部可能结果的集合. 例如，需要记录某个城市一天的交通事故数量，则试验结果将是非负整数 x. 无法确定 x 的可能取值的确切范围，但可以把这个范围取为$[0,+\infty)$，它总能包含一切可能的试验结果，尽管明知某些结果，如 $x>10\ 000$是不会出现的. 甚至可以把这个范围取为$(-\infty,+\infty)$也无妨. 这里就有了一定的数学抽象，它可以带来很大的方便.

2. 随机事件

在一次试验中可能出现，也可能不出现的结果称为**随机事件**，简称**事件**，并用大写字母 A,B,C 等表示. 为讨论需要，将每次试验中一定发生的事件称为**必然事件**，记为 Ω. 每次试验中一定不发生的事件称为**不可能事件**，记为$\varnothing$.

【注】 随机事件在一次试验中是否发生虽然不能事先确定，但是在大量重复试验的情况下，它的发生呈现出一定的规律性，这门课程正是要研究这种规律性，读者应在学习这门课程后，对此有较为深刻的认识.

3. 样本空间

随机试验的每一个可能结果称为**样本点**，记为 ω. 样本点的全体组成的集合称为**样本空间（或基本事件空间）**，记为 Ω，即 $\Omega=\{\omega\}$. 由一个样本点构成的事件称为**基本事件**. 随机事件 A 总是由若干个基本事件组成，即 A 是 Ω 的子集.

二、事件的关系与运算

1. 定义（关系：包含、相等、相容、互斥、对立；运算：和（并）、差、积（交））

(1)如果事件 A 发生必导致事件 B 发生，则称事件 B **包含**事件 A（或 A 被 B 包含），记为 $A\subset B$.

(2)如果 $A\subset B$ 且 $B\subset A$，则称事件 A 与 B **相等**，记为 $A=B$. A 与 B 相等，事实上也就是说，A 与 B 由完全相同的一些试验结果构成，它不过是同一事件表面上看来两个不同的说法而已.

(3)称“事件 A 与 B 同时发生”的事件为事件 A 与 B 的**积（或交）**，记为 $A\cap B$ 或 AB.

【注】 称“有限个（或可列个）事件 $A_1,A_2,\cdots,A_n(\cdots)$ 同时发生”的事件为事件 $A_1,A_2,\cdots,A_n(\cdots)$ 的积（或交），记为 $\bigcap_{i=1}^{n} A_i$（或 $\bigcap_{i=1}^{\infty} A_i$）.

(4)若 $AB\neq\varnothing$，则称事件 A 和 B **相容**；若 $AB=\varnothing$，则称事件 A 与 B **互不相容**，也叫**互斥**. 如果一些事件中任意两个事件都互斥，则称这些事件是**两两互斥的**，或简称**互斥的**.

(5)称“事件 A 与 B 至少有一个发生”的事件为事件 A 与 B 的**和（或并）**，记为 $A\cup B$.

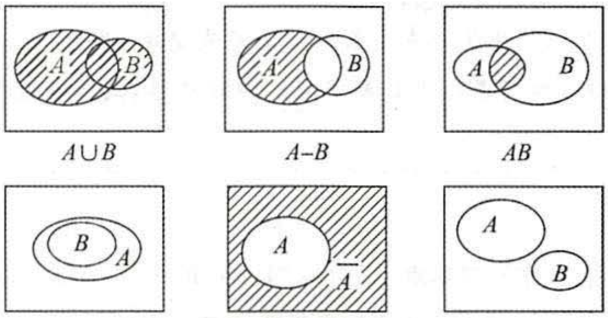
【注】 称“有限个（或可列个）事件 $A_1,A_2,\cdots,A_n(\cdots)$ 至少有一个发生”的事件为事件 $A_1,A_2,\cdots,A_n(\cdots)$ 的和（或并），记为 $\bigcup_{i=1}^{n} A_i$（或 $\bigcup_{i=1}^{\infty} A_i$）.

(6)称“事件 A 发生而事件 B 不发生”的事件为事件 A 与 B 的**差**，记为 $A-B$；称“事件 A 不发生”的事件为事件 A 的**逆事件**或**对立事件**，记为 $\overline{A}$.

由定义易知 $\quad A-B=A-AB=A\overline{B}$，$\quad B=\overline{A}\Leftrightarrow AB=\varnothing$ 且 $A\cup B=\Omega$.

(7)称有限个（或可列个）事件 $A_1,A_2,\cdots,A_n(\cdots)$ 构成**一个完备事件组**，如果 $\bigcup_{i=1}^{n} A_i$（或 $\bigcup_{i=1}^{\infty} A_i$）$=\Omega$，$A_iA_j=\varnothing$（对一切 $i\neq j;i,j=1,2,\cdots,n(\cdots)$）.

(8)事件的关系与运算可以用**文氏图**形象地表示出来（如图 3-1-1），图中的矩形表示必然事件 Ω.



$A\cup B$　　$A-B$　　AB

$B\subset A$　　$\overline{A}$　　$AB=\varnothing$

图 3-1-1

2. 运算法则

(1)吸收律　若 $A\subset B$,则 $A\cup B=B,A\cap B=A$;

(2)交换律　$A\cup B=B\cup A,A\cap B=B\cap A$;

(3)结合律　$(A\cup B)\cup C=A\cup(B\cup C),(A\cap B)\cap C=A\cap(B\cap C)$;

(4)分配律　$A\cap(B\cup C)=(A\cap B)\cup(A\cap C),A\cup(B\cap C)=(A\cup B)\cap(A\cup C),A\cap(B-C)=(A\cap B)-(A\cap C)$;

(5)对偶律(德·摩根律)　$\overline{A\cup B}=\overline{A}\cap\overline{B},\overline{A\cap B}=\overline{A}\cup\overline{B}$.

【注】 (1)事件运算顺序约定为先进行逆运算,而后进行交运算,最后进行并或差运算.

(2)事件的关系、运算与集合的关系、运算相当,且具有相同的运算法则,所以我们可以对比着理解记忆,并要学会用集合关系去考虑事件关系.

三、概率的定义

1. 描述性定义

通常将随机事件 A 发生的可能性大小的度量(非负值),称为事件 A 发生的**概率**,记为 $P(A)$.

2. 统计性定义

在相同条件下做重复试验,事件 A 出现的次数 k 和总的试验次数 n 之比 $\frac{k}{n}$,称为事件 A 在这 n 次试验中出现的**频率**.当试验次数 n 充分大时,频率将"稳定"于某常数 p 的"附近".n 越大,频率偏离这个常数 p 的可能性越小.这个常数 p 就称为事件 A 的**概率**.

【注】 (1)概率的统计性定义实质上是说,用频率 $\frac{k}{n}$ 作为事件 A 的概率 $P(A)$ 的估计.其直观背景为某事件出现的可能性大小,可由其在多次重复试验中出现的频率去刻画.

(2)从上述(1)可以看出,频率只是概率的估计,而非概率本身.也就是说,概率的统计性定义是无法准确给出某事件的概率的.其重要性主要基于以下两点.

①它提供了估计概率的方法.比如在一批产品中抽取样品,来估计该批产品的合格率(合格率是客观的数据,抽取样品计算出来的合格率,只是一种估计).

②它提供了一种检验某结论是否正确的准则.比如,你说某批产品的合格率是 95%,我们做试验,抽取样品进行计算,得出的结果是合格率为 20%,远远低于你所说的 95%,于是毫不犹豫地拒绝你的结论.

3. 公理化定义

设随机试验的样本空间为Ω，如果对每一个事件A都有一个确定的实数$P(A)$，且事件函数$P(\cdot)$满足：

(1)非负性：$P(A)\geqslant 0$；

(2)规范性：$P(\Omega)=1$；

(3)可列可加性：对任意可列个互不相容事件$A_1,A_2,\cdots,A_n,\cdots$(即$A_iA_j=\varnothing,i\neq j;i,j=1,2,\cdots$)，有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i),$$

则称$P(\cdot)$为**概率**，$P(A)$为事件A的**概率**.

【注】 (1)数学上所说的"公理"，就是一些不加证明而承认的前提，上述公理化定义只是界定了概率这个概念所必须满足的一些一般性质，它不解决具体场合下的概率计算.

(2)概率$P(\cdot)$是事件的函数.

(3)虽然它不解决具体场合下的概率计算，但是我们却常常用它来判断某事件函数$P(\cdot)$是否是概率，这种题型在考研试题中也是经常遇到的.

四、古典概型和几何概型

下面研究两种非常重要的概率类型：古典概型和几何概型.

(1)称随机试验(随机现象)的概率模型为**古典概型**，如果其样本空间(基本事件空间)满足：

①只有有限个样本点(基本事件)；

②每个样本点(基本事件)发生的可能性都一样.

如果古典概型的基本事件总数为n，事件A包含k个基本事件，也叫作有利于A的基本事件为k个，则A的概率为

$$P(A)=\frac{k}{n}=\frac{\text{事件 }A\text{ 所含基本事件的个数}}{\text{基本事件总数}}.$$

由上式计算得出的概率称为A的**古典概率**.

(2)称随机试验(随机现象)的概率模型为**几何概型**，如果：

①样本空间(基本事件空间)Ω是一个可度量的有界区域；

②每个样本点(基本事件)发生的可能性都一样，即样本点落入Ω的某一可度量的子区域S的可能性大小与S的几何度量成正比，而与S的位置及形状无关.

在几何概型随机试验中，如果S_A是样本空间Ω的一个可度量的子区域，则事件$A=\{$样本点落入区域$S_A\}$的概率为

$$P(A)=\frac{S_A\text{ 的几何度量}}{\Omega\text{ 的几何度量}}.$$

由上式计算得出的概率称为A的**几何概率**.

【注】 古典概型与几何概型的区别：基本事件有限、等可能的随机试验为**古典概型**；基本事件无限且具有几何度量、等可能的随机试验为**几何概型**.

五、概率的基本性质与公式

1. 性质

(1)有界性:对于任一事件 A,有 $0\leqslant P(A)\leqslant 1$,且 $P(\varnothing)=0, P(\Omega)=1$.

【注】 $P(A)=0$,不能断言 $A=\varnothing$;$P(A)=1$,不能断言 $A=\Omega$,其中道理可见例 3.1.12.

(2)单调性:设 A,B 是两个事件,若 $A\subset B$,则有

$$P(B-A)=P(B)-P(A),\quad P(B)\geqslant P(A).$$

2. 公式

(1)逆事件概率公式:对于任一事件 A,有 $P(\overline{A})=1-P(A)$.

(2)加法公式:对于任意两个事件 A,B,有 $P(A\cup B)=P(A)+P(B)-P(AB)$.

【注】 (1)设 A_1,A_2,A_3 为任意三个事件,则有

$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3).$$

(2)若 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件,则有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n).$$

(3)减法公式:$P(A-B)=P(A)-P(AB)=P(A\overline{B})$.

(4)条件概率公式:设 A,B 为任意两个事件,若 $P(A)>0$,我们称在已知事件 A 发生的条件下,事件 B 发生的概率为**条件概率**,记为 $P(B|A)$,且

$$P(B|A)=\frac{P(AB)}{P(A)}.$$

【注】 (1)条件概率 $P(\cdot|A)$是概率,概率的一切性质和重要结论对条件概率都适用.

例如:

$$P(\overline{B}|A)=1-P(B|A),$$
$$P(B-C|A)=P(B|A)-P(BC|A),$$

等等.

(2)条件概率就是在附加了一定的条件之下所计算的概率.当说到"条件概率"时,总是指另外附加的条件,其形式可归结为"已知某事件发生了".

(5)乘法公式:如果 $P(A)>0$,则 $P(AB)=P(A)P(B|A)$.

一般地,对于 $n>2$,如果 $P(A_1A_2\cdots A_{n-1})>0$,则

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

(6)全概率公式:如果 $\bigcup\limits_{i=1}^{n} A_i=\Omega, A_iA_j=\varnothing(i\neq j;i,j=1,2,\cdots,n), P(A_i)>0$,则对任一事件 B,有

$$B=\bigcup_{i=1}^{n} A_iB,\quad P(B)=\sum_{i=1}^{n} P(A_i)P(B\mid A_i).$$

(7)贝叶斯公式(又称逆概率公式):如果 $\bigcup\limits_{i=1}^{n} A_i=\Omega, A_iA_j=\varnothing(i\neq j;i,j=1,2,\cdots,n), P(A_i)>0$,则对任一事件 B,只要 $P(B)>0$,就有

$$P(A_j|B)=\frac{P(A_j)P(B\mid A_j)}{\sum\limits_{i=1}^{n}P(A_i)P(B\mid A_i)}(j=1,2,\cdots,n).$$

【注】 (1)要注意 $P(B)$ 与 $P(B|A)$ 的区别与联系，虽然二者都是计算事件 B 的概率，但前者 $P(B)$ 实际上是指在样本空间 Ω 下计算的，后者 $P(B|A)$ 则是在事件 A 已经发生的条件下（即样本空间现在缩减至 A）计算的.

(2)全概率公式是用于计算某个“结果”B 发生的可能性大小. 如果一个“结果”B 的发生总是与某些前提条件（或原因、因素或前一阶段结果）A_i 相联系，那么在计算 $P(B)$ 时，我们总是用 A_i 对 B 作分解：

$$B=\bigcup_{i=1}^{n}A_iB,$$

应用全概率公式计算 $P(B)$，我们常称这种方法为**全集分解法**. 如果在 B 发生的条件下，探求导致这一“结果”的各种“原因”，即 A_i 发生的可能性大小 $P(A_i|B)$，则要应用贝叶斯公式.

六、事件的独立性和独立重复试验

1. 事件的独立性

(1)描述性定义(直观性定义) 设 A,B 为两个事件，如果其中任何一个事件发生的概率不受另外一个事件发生与否的影响，则称事件 A 与 B **相互独立**. 设 $A_1,A_2,\cdots,A_n$ 是 $n(n\geqslant 2)$ 个事件，如果其中任何一个或几个事件发生的概率都不受其余的某一个或几个事件发生与否的影响，则称事件 $A_1,A_2,\cdots,A_n$ 相互独立.

(2)数学定义 设 A,B 为两个事件，如果 $P(AB)=P(A)P(B)$，则称事件 A 与 B **相互独立**，简称为 A 与 B **独立**.

【注】 设 $A_1,A_2,\cdots,A_n$ 为 $n(n\geqslant 2)$ 个事件，如果对其中任意有限个事件 $A_{i_1},A_{i_2},\cdots,A_{i_k}(2\leqslant k\leqslant n)$，有

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}),$$

则称 n 个事件 $A_1,A_2,\cdots,A_n$ 相互独立.

考研中常考的是 $n=3$ 时的情形. 细致说来，设 A_1,A_2,A_3 为三个事件，若

$$P(A_1A_2)=P(A_1)P(A_2),\tag{3-1-1}$$

$$P(A_1A_3)=P(A_1)P(A_3),\tag{3-1-2}$$

$$P(A_2A_3)=P(A_2)P(A_3),\tag{3-1-3}$$

$$P(A_1A_2A_3)=P(A_1)P(A_2)P(A_3),\tag{3-1-4}$$

则称事件 A_1,A_2,A_3 **相互独立**. 当去掉上述(3-1-4)式后，称只满足(3-1-1),(3-1-2),(3-1-3)式的事件 A_1,A_2,A_3 **两两独立**. 见例 3.1.21.

2. 试验的独立性

如果各个试验结果是相互独立的，则称这些试验是相互独立的. 例如，对试验 E_1 的任一结果 A_1，试验

E_2 的任一结果 A_2，事件 A_1 与 A_2 相互独立，即 $P(A_1A_2)=P(A_1)P(A_2)$，称随机试验 E_1 和 E_2 是相互独立的. 对试验 E_i 中的任一结果 $A_i(i=1,2,\cdots,n)$，事件 $A_1,A_2,\cdots,A_n$ 相互独立，即对其中任意 $k(2\leqslant k\leqslant n)$ 个事件有 $P\left(\bigcap\limits_{j=1}^{k}A_{i_j}\right)=\prod\limits_{j=1}^{k}P(A_{i_j})$，称 n 个试验 $E_1,E_2,\cdots,E_n$ 是相互独立的.

3. 独立试验序列概型与 n 重伯努利概型

在同样条件下独立重复地进行一系列完全相同的试验，即每次试验结果及其发生的概率都不变，各次试验是相互独立的，称这种重复试验序列的数学模型为**独立试验序列概型**. 如果每次试验只有两个结果 A 与 $\overline{A}$，且在每次试验中 A 发生的概率都相等(即 $P(A)=p$)，将这种试验独立重复 n 次，则称这种试验为 **n 重伯努利概型**.

在第 2 讲中会看到，在 n 重伯努利概型中，事件 A 发生 k 次(只管次数，不论位置)的概率为 $C_n^kp^k(1-p)^{n-k}(k=0,1,\cdots,n)$，且如果用 X 表示 n 重伯努利概型中事件 A 发生的次数，则 X 服从二项分布$B(n,p)$.

【注】 (1)事件相互独立是概率论中一个重要的概念，它是定义随机试验独立性、随机变量独立性的基础. 我们总是由试验的方式来判定事件的独立性，进而判定事件的相互独立性，再应用事件独立性定义中所揭示的概率关系计算与之有关的事件的概率.

(2)要善于判定独立试验序列概型，只要题目中出现“将……重复进行 n 次”“对……重复观察 n 次”等字样(注意要求每次试验只有两个结果 A 与 $\overline{A}$)，或可以转换为 n 次独立重复试验概型的问题，这些都是要考虑应用二项分布计算与之有关的事件的概率.

基础例题精解

一、事件的关系与运算

随机事件的描述和运算是概率计算的基础，也是考研数学中的一个考点.

(1)求解概率试题，通常要做的是首先**将实际问题符号化**，即用符号及其运算来描述事件，这种描述可以体现出考生求解问题的思路. 用简明的符号和运算准确反映题意、说明问题，是考生数学能力的一种体现.

(2)需要指出，在讨论和推断随机事件之间的关系时，**采用直观的文氏图**常常是一种有效的方法.

(3)此部分虽然比较基础，但很重要，将实际问题符号化的能力将伴随着概率论的整个课程.

例 3.1.1 以 A 表示事件“甲产品畅销，乙产品滞销”，则其对立事件 $\overline{A}$ 为(　　).

(A)“甲产品滞销，乙产品畅销”　　(B)“甲、乙产品均畅销”

(C)“甲产品滞销或乙产品畅销”　　(D)“甲产品滞销”

解 应选(C).

先将事件符号化，再利用事件运算解答，即以 A_1 表示事件“甲产品畅销”，A_2 表示事件“乙产品滞销”，则 $A=A_1A_2$，其对立事件为 $\overline{A}=\overline{A_1A_2}=\overline{A_1}\cup\overline{A_2}$，表示事件“甲产品滞销或乙产品畅销”，故选择(C).

例 3.1.2 设 A,B 和 C 是任意三个事件，则下列选项中正确的是(　　).

(A) 若 $A\cup C=B\cup C$，则 $A=B$

(B) 若 $A-C=B-C$，则 $A=B$

(C) 若 $AC=BC$，则 $A=B$

(D) 若 $AB=\varnothing$ 且 $\overline{A}\,\overline{B}=\varnothing$，则 $\overline{A}=B$

解 应选(D).

方法一(直选法) 由事件运算的对偶律，有 $\overline{\overline{A}\,\overline{B}}=A\cup B=\overline{\varnothing}=\Omega$. 而由 $A\cup B=\Omega$ 且 $AB=\varnothing$，可见 A 和 B 互为对立事件，即 $\overline{A}=B$，因此(D)正确.

方法二(排除法) 前三个选项都不成立，只需分别举出反例. 例如，由于 A,B,C 是任意三个事件，若取 $A\neq B$，而 $C=\Omega$ 是必然事件，则 $A\cup C=B\cup C$ 且 $A-C=B-C$，但 $A\neq B$，从而(A)和(B)不成立.

设 $A\neq B$，$C=\varnothing$，则 $AC=BC$，但 $A\neq B$，因此(C)不成立. 从而选(D).

【注】 本题的结果反映了事件的运算与数的运算的不同之处.

例 3.1.3 试问下列命题是否成立，并说明理由.

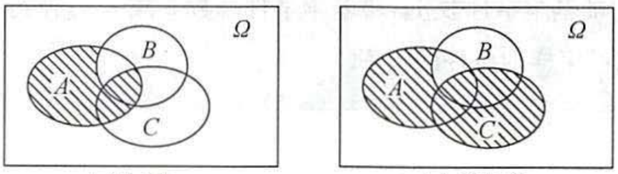
(1)$A-(B-C)=(A-B)\cup C$；

(2)若 $AB=\varnothing$ 且 $C\subset A$，则 $BC=\varnothing$；

(3)$(A\cup B)-B=A$；

(4)$(A-B)\cup B=A$.

解 (1)不成立. $A-(B-C)=A-B\overline{C}=A\overline{B\overline{C}}=A(\overline{B}\cup C)=A\overline{B}\cup AC=(A-B)\cup AC\neq(A-B)\cup C$；也可由图 3-1-2 快速判断.



$A-(B-C)$
(a)

$(A-B)\cup C$
(b)

图 3-1-2

(2)成立. 因 $C\subset A$，有 $BC\subset AB=\varnothing$，故 $BC=\varnothing$.

(3)不成立. 因 $(A\cup B)-B=(A\cup B)\overline{B}=A\overline{B}\cup B\overline{B}=A\overline{B}=A-B\neq A$.

(4)不成立. 因 $(A-B)\cup B=A\overline{B}\cup B=(A\cup B)(\overline{B}\cup B)=A\cup B\neq A$.

【注】 读者可再做练习，证明下列事件的运算公式：

(1)$A=AB\cup A\overline{B}$；

(2)$A\cup B=A\cup\overline{A}B$.

证明 (1)$AB\cup A\overline{B}=A(B\cup\overline{B})=A\Omega=A$.

(2)$A\cup\overline{A}B=(A\cup\overline{A})(A\cup B)=\Omega(A\cup B)=A\cup B$.

二、古典概型和几何概型

1. 古典概型

计算的关键是基本事件、样本空间的选定以及基本事件数的计算.

计数方法常用的有三种.

(1)列举法(直接查数法)：基本事件数不多时常用这种方法.

(2)集合对应法：①加法原理——完成一件事有 n 类办法，第一类办法中有 m_1 种方法，第二类办法中有 m_2 种方法，……，第 n 类办法中有 m_n 种方法，则完成此事共有 $\sum\limits_{i=1}^{n} m_i$ 种方法.

②乘法原理——完成一件事有 n 个步骤. 第一步有 m_1 种方法，第二步有 m_2 种方法，……，第 n 步有 m_n 种方法，则完成此事共有 $\prod\limits_{i=1}^{n} m_i$ 种方法.

③排列——从 n 个不同的元素中取出 $m(\leqslant n)$ 个元素，并按照一定顺序排成一列，叫作排列. 所有排列的个数叫作排列数，记作

$$\mathrm{P}_n^m=n(n-1)(n-2)\cdots(n-m+1)=\frac{n!}{(n-m)!}.$$

当 $m=n$ 时，$\mathrm{P}_n^n=\frac{n!}{0!}=n!$，叫作全排列.

④组合——从 n 个不同的元素中取出 $m(\leqslant n)$ 个元素，并成一组，叫作组合. 所有组合的个数叫组合数，记作

$$\mathrm{C}_n^m=\frac{\mathrm{P}_n^m}{m!}.$$

(3)逆数法：先求 $\overline{A}$ 中的基本事件数 $n_{\overline{A}}$，将基本事件总数 n 减去 $n_{\overline{A}}$ 便得 A 中的基本事件数，这种方法常用于计算含有“至少”字样的事件的概率.

古典概型的常见类型如下.

(1)直接用定义求概率.

古典概型的概率计算最基本的做法就是“数数”，数出总样本数和事件所含样本点数，在以往考研试题中也出现过这种直接数数的题型.

例 3.1.4 考虑一元二次方程 $x^2+Bx+C=0$，其中 B,C 分别是将一枚骰子(色子)接连抛两次前后出现的点数，求该方程有实根的概率 p 和有重根的概率 q.

分析 本题很难套用现有概型模式求解，一种最简单的也是最直接的做法：由定义，数出总样本数和事件所含样本点数给出答案. 后者可借助列表显示所有可能的结果.

解 方程有实根，即事件“$B^2-4C\geqslant 0$”；方程有重根，即事件“$B^2-4C=0$”. 每枚骰子抛一次共有 6 种可能结果，抛两次共有 6^2 种结果，列表如下：

B^2-4C符号 C / B	1	2	3	4	5	6
1	−	−	−	−	−	−
2	0	−	−	−	−	−
3	+	+	−	−	−	−
4	+	+	+	0	−	−
5	+	+	+	+	+	+
6	+	+	+	+	+	+

事件“$B^2-4C\geqslant 0$”所含样本点数为19，事件“$B^2-4C=0$”所含样本点数为2，故

$$p=\frac{19}{36},\quad q=\frac{2}{36}=\frac{1}{18}.$$

(2)随机分配.

随机分配也叫随机占位，突出一个“放”字，即将 n 个可辨质点随机地分配到 N 个盒中，区分每盒最多可以容纳一个和可以容纳任意多个质点，不同分法的总数列表如下.

将 n 个质点随机地分配到 N 个盒中	分配方式	不同分法的总数
	每盒可容纳任意多个质点	N^n（见注1）
	每盒容纳至多一个质点	$\mathrm{P}_N^n=N(N-1)\cdot\cdots\cdot(N-n+1)$（见注2）

【注1】 每个质点均可放到 N 个盒子中的任何一个，即有 N 种放法，于是 n 个可辨质点放到 N 个盒子中共有 N^n 种不同放法.

【注2】 质点可辨，且一个盒子容纳至多一个质点，故 n 个质点放到 N 个盒子中的所有不同放法即为从 N 个元素中选取 n 个元素的排列数 P_N^n.

例 3.1.5　5人共钓到3条鱼，每条鱼被各人钓到的可能性相同，求：

(1)3条鱼由不同的人钓到的概率；

(2)有1人钓到2条鱼的概率；

(3)3条鱼由同一人钓到的概率.

分析　看作随机分配问题，即把鱼视作“质点(可辨)”，把人视作“盒子(可容纳任意多个质点)”.

解　样本空间基本事件总数为 5^3.

(1)先从5个人中选3个人，共有 C_5^3 种选法，再将3条鱼给3个人，每人一条，共有3！种方法，由乘法原理，基本事件数为 $\mathrm{C}_5^3 3!$，于是

$$P\{3\text{ 条鱼由不同的人钓到}\}=\frac{\mathrm{C}_5^3 3!}{5^3}=\frac{12}{25}=0.48.$$

(2)先在5个人中任选1人，为 C_5^1，再从3条鱼中任取2条，为 C_3^2，分给这个人，剩下1条鱼随机给剩下

的 4 个人，为$(5-1)^{3-2}=4^1=4$，基本事件数为$C_5^1C_3^2(5-1)^{3-2}$，于是

$$P\{有1人钓到2条鱼\}=\frac{C_5^1C_3^2(5-1)^{3-2}}{5^3}=\frac{12}{25}=0.48.$$

(3)5 个人中任选 1 人，为C_5^1，3 条鱼全给一个人，即C_3^3，基本事件数为C_5^1，于是

$$P\{3条鱼由同一人钓到\}=\frac{C_5^1}{5^3}=\frac{1}{25}=0.04.$$

(3)简单随机抽样.

设$\Omega=\{\omega_1,\omega_2,\cdots,\omega_N\}$含$N$个元素，称$\Omega$为总体. 如果各元素被抽到的可能性相同，自总体$\Omega$的抽样称作简单随机抽样，突出一个“取”字.

分为先后有放回、先后无放回及任取这三种不同的简单随机抽样方式. 在每种抽样方式下各种不同抽法（基本事件）的总数列表如下.

自含N个元素的总体Ω中n次简单随机抽样	抽样方式	抽法总数
	先后有放回取n次	N^n（见注 1）
	先后无放回取n次	$P_N^n=N(N-1)\cdot\cdots\cdot(N-n+1)$（见注 2）
	任取n个	C_N^n（见注 3）

【注 1】 既考虑抽到何元素，又考虑各元素出现的顺序，每次从Ω中随意抽取一个元素，并在抽取下一元素前将其放回Ω. 于是每次都有N个元素可被抽取，即有N种抽法，抽n次，即N^n.

【注 2】 既考虑抽到何元素，又考虑各元素出现的顺序，凡是抽出的元素均不再放回Ω，于是每次抽取时都比上一次少了一个元素，抽n次，即

$$P_N^n=N(N-1)\cdot\cdots\cdot(N-n+1).$$

【注 3】 任取n个是指一次性取n个元素，相当于将n个元素无序且无放回取走，其抽法总数为$C_N^n=\frac{P_N^n}{n!}$.

例 3.1.6 从 1,2,…,15 这 15 个数中随机取出 3 个，试求下列事件的概率：$A_1=\{3个数中最大的是10\}$；$A_2=\{3个数中大于、等于和小于7的各1个\}$；$A_3=\{3个数中2个大于7,1个小于7\}$.

解 这是“无放回且无序”的问题. 从 15 个数中随机取出 3 个，总共有$C_{15}^3=455$(种)不同取法，即总共有C_{15}^3个基本事件，其中有利于A_1的取法有$C_9^2=36$(种)(3 个数中最大的是 10，在小于 10 的 9 个数中随意取 2 个有C_9^2种不同取法)；有利于A_2的取法有$6\times8=48$(种)(在小于 7 的 6 个数中随意取 1 个，在大于 7 的 8 个数中随意取 1 个，有6×8种不同取法)；有利于A_3的取法有$6\times C_8^2=168$(种)(在小于 7 的 6 个数中随意取 1 个，在大于 7 的 8 个数中随意取 2 个). 于是

$$P(A_1)=\frac{36}{455},\quad P(A_2)=\frac{48}{455},\quad P(A_3)=\frac{168}{455}.$$

2. 几何概型

几何概型的特点是总样本数和事件所含样本点数不能如古典概型那样可以数数，但具有几何特性，如长度、面积或体积等．解题时应将数量关系用几何直观图形表现出来，其概率通常是事件所含样本数和总样本数对应区域大小之比．

例 3.1.7　在区间(0,1)中随机地取两个数，则这两个数之差的绝对值小于 0.5 的概率为________．

解　应填 0.75．

设两个数分别为 x,y，则依题意 (x,y) 的取值范围为正方形区域 $D=\{(x,y)\mid 0<x<1,0<y<1\}$．又 $|x-y|<0.5$，则所在区域如图 3-1-3 中阴影部分所示，于是所求概率为

$$p=\frac{1^2-0.5^2}{1^2}=0.75.$$

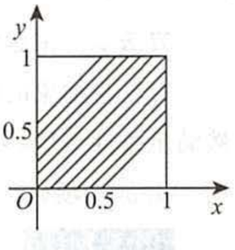


图 3-1-3

例 3.1.8　随机地向半圆 $0<y<\sqrt{2ax-x^2}(a>0)$ 内投掷一点，点均匀落在半圆内任何一个区域，求该点和原点连线同 x 轴夹角 $\theta\leqslant\frac{\pi}{4}$ 的概率．

解　由题设，半圆与直线 $y=x$ 的交点坐标为 (a,a)，如图 3-1-4 所示，阴影部分区域 D 内任意点与原点连线同 x 轴的夹角 $\theta\leqslant\frac{\pi}{4}$，因此所求概率为

$$P\{(x,y)\in D\}=\left(\frac{\pi a^2}{4}+\frac{a^2}{2}\right)\Big/\frac{\pi a^2}{2}=\frac{1}{2}+\frac{1}{\pi}.$$

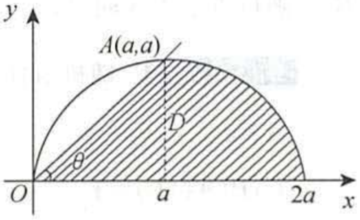


图 3-1-4

三、概率的基本性质与公式

1. 加法公式、减法公式、逆事件概率公式

首先，要记住概率性质公式(等量关系有：加法公式，减法公式与逆事件概率公式．不等式关系有：有界性，单调性，$\sqrt{P(A)P(B)}\leqslant\frac{P(A)+P(B)}{2}$，$P(B|A)=\frac{P(AB)}{P(A)}\geqslant P(AB)$，等等)．

其次，要将题目中的已知条件，即所求事件的概率及其数量关系用数学符号明确表示出来．

最后，通过解方程或等量代换即可求得所要结果．

例 3.1.9　设 A,B 为两个随机事件，且 $P(A)=0.4,P(A\cup B)=0.7$．若 A,B 互不相容，则 $P(B)=$________；若 A,B 相互独立，则 $P(B)=$________；若 A 发生 B 必发生，则 $P(B)=$________．

解　应填 0.3；0.5；0.7．

由加法公式，有 $P(A\cup B)=P(A)+P(B)-P(AB)$，于是：

若 A,B 互不相容，则 $P(AB)=0,P(B)=P(A\cup B)-P(A)=0.3$；

若 A,B 相互独立，则 $P(AB)=P(A)P(B),P(A\cup B)=P(A)+P(B)-P(A)P(B)$，解得

$$P(B)=\frac{P(A\cup B)-P(A)}{1-P(A)}=0.5;$$

若 A 发生 B 必发生，则 $P(B)=P(A\cup B)=0.7$．

例 3.1.10 若随机事件 A,B 同时发生，C 也必然发生，则下列选项必然成立的是(　　).

(A)$P(C)<P(A)+P(B)-1$　　(B)$P(C)\geqslant P(A)+P(B)-1$

(C)$P(C)=P(AB)$　　(D)$P(C)=P(A\cup B)$

解　应选(B).

方法一　由题设知，$AB\subset C$，则 $P(AB)\leqslant P(C)$. 又由加法公式 $P(A\cup B)=P(A)+P(B)-P(AB)$，得 $P(AB)=P(A)+P(B)-P(A\cup B)\geqslant P(A)+P(B)-1$，选(B).

方法二　由题设知，$AB\subset C$，但未必有 $C=AB$ 或 $C=A\cup B$，应排除选项(C)，(D).

对较为抽象的问题，可在条件允许的范围内选择较为特殊的情况进行验证. 如本题若取 $A=B=C$，仍然满足条件 $AB=A\subset C$，代入选项(A)，有 $P(A)<P(A)+P(A)-1$，即 $P(A)>1$，与性质不符，应排除选项(A). 故选项(B)正确，选之.

例 3.1.11 设 A 和 B 是任意两事件，讨论下列命题的正确性.

(1)若 $P(A)=P(B)$，则 $A=B$；

(2)若 $P(AB)=0$，则 $AB=\varnothing$.

解　两个命题都不成立. 考虑向区间$[0,2]$上掷随机点试验的事件. $A=\{$随机点落入区间$[0,1]$上$\}$，$B=\{$随机点落入区间$[1,2]$上$\}$. 显然 $P(A)=P(B)$，$P(AB)=0$，然而 $A\neq B$，$AB\neq\varnothing$.

例 3.1.12 随机事件 A,B，满足 $P(A)=P(B)=\frac{1}{2}$ 和 $P(A\cup B)=1$，则有(　　).

(A)$A\cup B=\Omega$　　(B)$AB=\varnothing$

(C)$P(\overline{A}\cup\overline{B})=1$　　(D)$P(A-B)=0$

解　应选(C).

根据加法公式 $P(A\cup B)=P(A)+P(B)-P(AB)$，将已知条件代入得 $1=\frac{1}{2}+\frac{1}{2}-P(AB)$，即 $P(AB)=0$，从而 $P(\overline{A}\cup\overline{B})=P(\overline{AB})=1-P(AB)=1$，选(C).

> **【注】** $P(A)=1$ 不能推出 $A=\Omega$；$P(B)=0$ 不能推出 $B=\varnothing$.

2. 条件概率公式

例 3.1.13 已知 $P(\overline{A})=0.3$，$P(B)=0.4$，$P(A\overline{B})=0.5$，求 $P(B|A\cup\overline{B})$.

解　由条件概率的定义知 $P(B|A\cup\overline{B})=\dfrac{P(AB)}{P(A\cup\overline{B})}$，其中

$$P(A\cup\overline{B})=P(A)+P(\overline{B})-P(A\overline{B})=0.7+0.6-0.5=0.8.$$

又由 $P(A\overline{B})=P(A)-P(AB)$，可得

$$P(AB)=P(A)-P(A\overline{B})=0.7-0.5=0.2.$$

代回原式，可得

$$P(B|A\cup\overline{B})=\frac{P(AB)}{P(A\cup\overline{B})}=\frac{0.2}{0.8}=0.25.$$

例 3.1.14 设 A,B,C 是三个随机事件，A 与 C 互斥，$P(AB)=\frac{1}{2}$，$P(C)=\frac{1}{3}$，则 $P(AB|\overline{C})=$ ______.

解　应填 $\frac{3}{4}$.

因为 A 与 C 互斥，所以 $P(AC)=0$. 又 $ABC\subset AC$，所以 $P(ABC)=0$，故

$$P(AB|\overline{C})=\frac{P(AB\overline{C})}{P(\overline{C})}=\frac{P(AB)-P(ABC)}{1-P(C)}=\frac{\frac{1}{2}}{1-\frac{1}{3}}=\frac{3}{4}.$$

【注】 $AC=\varnothing$，则 $ABC=\varnothing B=\varnothing$，同样得到 $P(ABC)=0$.

3. 乘法公式

例 3.1.15　已知 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{3}$，$P(A|B)=\frac{1}{2}$，求 $P(A\cup B)$.

解　由乘法公式知

$$P(AB)=P(A)P(B|A)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12},\quad P(B)=\frac{P(AB)}{P(A|B)}=\frac{1/12}{1/2}=\frac{1}{6},$$

所以

$$P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{4}+\frac{1}{6}-\frac{1}{12}=\frac{1}{3}.$$

例 3.1.16　某人给四位亲友各写一封信，然后随机地装入 4 个写好地址的信封中，且每个信封装一封信．问：

(1)4 封信都装对了的概率；

(2)4 封信都装错了的概率.

解　设事件 $A_i(i=1,2,3,4)$ 为“第 i 封信装对了”，事件 A 为“4 封信都装对了”，事件 B 为“4 封信都装错了”．于是

$$(1)P(A)=P(A_1A_2A_3A_4)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)P(A_4|A_1A_2A_3)$$

$$=\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}\times\frac{1}{1}=\frac{1}{24}.$$

$$(2)P(B)=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3}\,\overline{A_4})$$

$$=1-P\left(\overline{\overline{A_1}\,\overline{A_2}\,\overline{A_3}\,\overline{A_4}}\right)=1-P(A_1\cup A_2\cup A_3\cup A_4)$$

$$=1-\{P(A_1)+P(A_2)+P(A_3)+P(A_4)-$$

$$[P(A_1A_2)+P(A_1A_3)+P(A_1A_4)+P(A_2A_3)+P(A_2A_4)+P(A_3A_4)]+$$

$$[P(A_1A_2A_3)+P(A_1A_2A_4)+P(A_1A_3A_4)+P(A_2A_3A_4)]-P(A_1A_2A_3A_4)\},$$

其中 $P(A_i)=\frac{1}{4}$，$P(A_iA_j)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$，$P(A_iA_jA_k)=\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{24}(1\leqslant i,j,k\leqslant 4,i\neq j\neq k)$，于是

$$P(B)=1-4\times\frac{1}{4}+6\times\frac{1}{12}-4\times\frac{1}{24}+\frac{1}{24}=\frac{3}{8}.$$

【注】　本题装信过程可看作无放回抽取信件，依次装入信封．在计算第(2)问时，直接计算是较为困难的，这种情况下，采用逆向思维方法求解问题更加简捷.

4. 全概率公式

全概型的特点是事件的发生是在多个因素或条件下发生的，且这多个因素或条件均构成一个完备事件组，抓住并设定完备事件组是解决这类问题的切入点和关键.

例 3.1.17　每箱产品有 10 件，其次品数从 0 到 2 是等可能的，开箱检验时，从中任取 1 件，如检验出是次品，则认为该箱产品不合格而拒收．假设由于检验有误，将 1 件正品误认为次品的概率为 2%，1 件

次品被漏查而判为正品的概率为 5%，求该箱产品通过验收的概率.

解 题中有两个完备事件组：一是箱中次品的个数为 0，1，2；二是通过验收是在正确判断和错误判断的两种情况下发生. 设 $A_i(i=0,1,2)$为“箱中有 i 个次品”，B 为“通过验收”，B_1 为“抽取正品”，于是有

$$P(A_i)=\frac{1}{3}(i=0,1,2),$$

$$P(B_1|A_i)=\frac{10-i}{10},$$

从而由全概率公式得

$$P(B_1)=\sum_{i=0}^{2}P(A_i)P(B_1|A_i)=\frac{1}{3}\sum_{i=0}^{2}\frac{10-i}{10}=0.9,\quad P(\overline{B_1})=0.1,$$

因此由全概率公式得

$$\begin{aligned}P(B)&=P(B_1)P(B|B_1)+P(\overline{B_1})P(B|\overline{B_1})\\&=0.9\times0.98+0.1\times0.05=0.887.\end{aligned}$$

【注】 本题中，通过验收是在完备事件组 $B_1,\overline{B_1}$ 背景下发生的，抽取正品，则是在完备事件组 A_0,A_1,A_2 背景下实现的，因此，具备全概型的特征.

5. 贝叶斯公式（又称逆概率公式）

贝叶斯概型的特点是事件在多种因素或条件下发生，在事件已经发生后反过来讨论该事件是在哪个因素或条件下发生的概率.

例 3.1.18 设有两批数量相同的零件，已知有一批产品全部合格，另一批产品有 25%不合格. 从两批产品中任取 1 只，经检验是正品，放回原处，并在原所在批次再取 1 只，试求这只产品是次品的概率.

分析 两次抽取情况有不同，第一次是在完全不知情的情况下等可能地从两批产品中抽取，第二次是第一次抽取产品并检验合格后，这时对所抽取是第一批还是第二批产品的可能性已有推断，因此计算要分两步走. 第一步在已知抽取产品合格的条件下，推断抽取的是第一批还是第二批产品的概率，是属贝叶斯概型，第二步是以分别从两个批次抽取为完备事件组计算第二次抽到次品的概率，属全概型，两个概型的复合，常常是该类题型的特点.

解 设 $H_i(i=1,2)$为“第一次从第 i 批产品中抽取”，A 为“取正品”，则有

$$P(H_1)=P(H_2)=\frac{1}{2},\ P(A|H_1)=1,\ P(A|H_2)=\frac{3}{4},$$

即有
$$P(A)=P(H_1)P(A|H_1)+P(H_2)P(A|H_2)=\frac{7}{8}.$$

从而有

$$P(H_1|A)=\frac{P(H_1)P(A|H_1)}{P(A)}=\frac{4}{7},$$

$$P(H_2|A)=1-P(H_1|A)=\frac{3}{7}.$$

又设 $C_i(i=1,2)$为“第二次从第 i 批产品中抽取”，则有

$$\begin{aligned}P(\overline{A})&=P(C_1)P(\overline{A}|C_1)+P(C_2)P(\overline{A}|C_2)\\&=\frac{4}{7}\times0+\frac{3}{7}\times\frac{1}{4}=\frac{3}{28}.\end{aligned}$$

四、事件的独立性和 n 重伯努利试验

1. 事件的独立性

例 3.1.19 三人独立地破译一个密码，他们能单独译出的概率分别为$\frac{1}{5}$，$\frac{1}{3}$，$\frac{1}{4}$，求此密码被译出的概率.

解 记事件 A_i 为“第 i 个人译出密码”，$i=1,2,3$，B 为“密码被译出”. 则

$$P(B)=P(A_1\cup A_2\cup A_3)=1-P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=1-\frac{4}{5}\times\frac{2}{3}\times\frac{3}{4}=\frac{3}{5}.$$

【注】 互不相容可简化事件并的概率计算，相互独立可简化事件交的概率计算. 这里为了利用相互独立性，把事件的并在对偶律下转化为事件的交，这一方法会经常用到.

例 3.1.20 设 A,B 是任意两个事件，其中 $0<P(A)<1$，证明：$P(B|A)=P(B|\overline{A})$是事件 A,B 相互独立的充分必要条件.

证明 由 $0<P(A)<1$，知条件概率 $P(B|A)$，$P(B|\overline{A})$均存在.

必要性. 若 A,B 相互独立，则有 $P(B|A)=P(B)$，$P(B|\overline{A})=P(B)$，从而有 $P(B|A)=P(B|\overline{A})$.

充分性. 若 $P(B|A)=P(B|\overline{A})$，则有

$$\frac{P(AB)}{P(A)}=\frac{P(\overline{A}B)}{P(\overline{A})}=\frac{P(B)-P(AB)}{1-P(A)},$$

$$P(AB)[1-P(A)]=P(A)[P(B)-P(AB)],$$

从而有 $P(AB)=P(A)P(B)$，即 A,B 相互独立.

例 3.1.21 将一枚硬币独立地掷两次，引进事件：$A_1=\{$掷第一次出现正面$\}$，$A_2=\{$掷第二次出现正面$\}$，$A_3=\{$正反面各出现一次$\}$，$A_4=\{$正面出现两次$\}$，则事件(　　).

(A)A_1,A_2,A_3 相互独立　　(B)A_2,A_3,A_4 相互独立

(C)A_1,A_2,A_3 两两独立　　(D)A_2,A_3,A_4 两两独立

解 应选(C).

由题设，根据古典概型，有

$$P(A_1)=P(A_2)=P(A_3)=\frac{1}{2},$$

$$P(A_1A_2)=\frac{1}{4},\quad P(A_1A_3)=\frac{1}{4},\quad P(A_2A_3)=\frac{1}{4},$$

但

$$P(A_1A_2A_3)=0\neq P(A_1)P(A_2)P(A_3),$$

知 A_1,A_2,A_3 两两独立，但 A_1,A_2,A_3 不相互独立.

又

$$A_4A_3=\varnothing,\quad P(A_4)=\frac{1}{4},\quad P(A_3)=\frac{1}{2},\quad P(A_3A_4)=0\neq P(A_3)P(A_4),$$

知 A_2,A_3,A_4 不两两独立，也不相互独立，故选择(C).

2. n 重伯努利试验

(1)**独立试验** 称试验 $E_1,E_2,\cdots,E_n$ 为相互独立的，如果分别与各个试验相联系的任意 n 个事件之间相互独立.

(2)**独立重复试验** “独立”表示“与各试验相联系的事件之间相互独立”，其中“重复”表示“每个事件在各次试验中出现的概率不变”.

(3)**伯努利试验** 只计“成功”和“失败”两种对立结果的试验，称作**伯努利试验**. 将伯努利试验独立地重复进行 n 次，称作 **n 重伯努利试验，亦简称伯努利试验.**

于是，伯努利试验的特点：

①只有两种对立的结果；

②各次试验相互独立；

③各次试验成功的概率相同.

相应地，n 重伯努利概型是满足下述条件的随机试验：

①每次试验只有 A 与 $\overline{A}$ 两个结果；

②每次试验 A 发生的概率 $p=P(A)$ 不变；

③试验独立重复进行 n 次.

如果用 X 表示 n 次独立试验事件 A 发生的次数，则 $X\sim B(n,p)$，即

$$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}(k=0,1,\cdots,n).$$

例 3.1.22 对同一目标接连进行 3 次独立重复射击，假设至少命中目标一次的概率为 $\frac{7}{8}$，则每次射击命中目标的概率 $p=$________.

解 应填 $\frac{1}{2}$.

记事件 $A_i=\{$第 i 次命中目标$\}(i=1,2,3)$. 由条件知，事件 A_1,A_2,A_3 相互独立，且其概率均为 p. 已知 3 次独立重复射击至少命中目标一次的概率为

$$\begin{aligned}P(A_1\cup A_2\cup A_3)&=1-P(\overline{A_1}\;\overline{A_2}\;\overline{A_3})\\&=1-P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})\\&=1-(1-p)^3=\frac{7}{8},\end{aligned}$$

由此得 $p=\frac{1}{2}$.

例 3.1.23 某人向同一目标独立重复射击，每次命中目标的概率为 $p(0<p<1)$，则此人第 4 次射击恰好第二次命中目标的概率为( ).

(A) $3p(1-p)^2$ (B) $6p(1-p)^2$ (C) $3p^2(1-p)^2$ (D) $6p^2(1-p)^2$

解 应选(C).

依题设，4 次射击，最后一次命中. 前 3 次有一次命中，若设前 3 次射击有 ξ 次命中，则 $\xi\sim B(3,p)$，且有

$$P\{\xi=1\}=\mathrm{C}_3^1 p(1-p)^2.$$

于是第 4 次射击恰好第二次命中目标的概率为 $3p^2(1-p)^2$，故选择(C).

基础习题精练

习　题

3.1.1　用事件 A,B,C 的运算关系表示事件：①A,B,C 都不发生；②A,B,C 不都发生；③A,B,C 不多于一个发生.

3.1.2　甲口袋有 5 个白球、3 个黑球，乙口袋有 4 个白球、6 个黑球. 从两个口袋中各任取一球，求取到的两个球颜色相同的概率.

3.1.3　从$[0,1]$中随机地取两个数，求其积大于$\frac{1}{4}$，其和小于$\frac{5}{4}$的概率.

3.1.4　两个人相约 7 点至 8 点到某地点会面，先到者等另一人 20 分钟，过时就可以离去，试求两个人能会面的概率.

3.1.5　已知事件 A,B 满足 $P(AB)=P(\overline{A}\cap\overline{B})$，记 $P(A)=p$，试求 $P(B)$.

3.1.6　已知 $P(A)=0.7$，$P(A-B)=0.3$，试求 $P(\overline{AB})$.

3.1.7　甲、乙两人独立地对同一目标射击一次，其命中率分别为 0.6 和 0.7. 现已知目标被击中，求它是甲射中的概率.

3.1.8　已知装有同种零件的产品两箱，第一箱内装 50 件产品，其中一等品 10 件；第二箱内装 30 件产品，其中一等品 18 件，现从两箱中任意挑选一箱. 试求：从中先后取出两件产品（取后不放回），先取出的产品是一等品的概率 p；已知先取出的产品是一等品，那么第二次取出的产品仍然是一等品的概率 q.

3.1.9　口袋中有 1 个球，不知它的颜色是黑的还是白的，现再往口袋中放入 1 个白球，然后从口袋中任意取出 1 个，发现取出的是白球，试求口袋中原来那个球是白球的概率.

3.1.10　设 A,B,C 为三个随机事件，且 A 与 C 相互独立，B 与 C 相互独立，则 $A\cup B$ 与 C 相互独立的充分必要条件是(　　).

(A)A 与 B 相互独立　　(B)A 与 B 互不相容

(C)AB 与 C 相互独立　　(D)AB 与 C 互不相容

3.1.11　设事件 A 在每次试验中出现的概率为 p，则在 n 次独立重复试验中事件 A 最多出现一次的概率 $P=$________.

解　答

3.1.1　**解**　这是一道要求将概率论语言叙述的事件用事件关系来表示的题目，只要知道各种运算所描述的事件关系，此题是不难解答的.

①$\{A,B,C$ 都不发生$\}=\{A$ 不发生，且 B 不发生，且 C 不发生$\}=\overline{A}\,\overline{B}\,\overline{C}=\overline{A\cup B\cup C}$.

②$\{A,B,C$ 不都发生$\}=\{A,B,C$ 至少有一个不发生$\}=\{A,B,C$ 都发生$\}$的逆事件

$=\overline{A}\cup\overline{B}\cup\overline{C}=\overline{ABC}$.

③$\{A,B,C$ 不多于一个发生$\}=\{A,B,C$ 至多有一个发生$\}=\{A,B,C$ 至少有两个不发生$\}$

$=\overline{A}\,\overline{B}\cup\overline{A}\,\overline{C}\cup\overline{B}\,\overline{C}=A\,\overline{B}\,\overline{C}\cup\overline{A}B\,\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}\,\overline{B}\,\overline{C}$.

3.1.2 解 从两个口袋中各取一球，共有 $C_8^1C_{10}^1$ 种等可能取法，而两个球颜色相同有两种情况：第一种是从甲口袋取出白球，从乙口袋也取出白球；第二种是从甲口袋取出黑球，从乙口袋也取出黑球. 共有 $C_5^1C_4^1+C_3^1C_6^1$ 种取法，于是 $P\{$取到的两个球颜色相同$\}=\dfrac{5\times4+3\times6}{8\times10}=\dfrac{19}{40}$.

3.1.3 解 从$[0,1]$中随机地取两个数 x 与 y，可以看成从图 3-1-5 正方形区域中任取一点，满足几何概型的两个特点，$\Omega=\{(x,y)\mid 0\leqslant x,y\leqslant 1\}$，记 $A=\left\{(x,y)\,\middle|\,x+y<\dfrac{5}{4},xy>\dfrac{1}{4}\right\}$对应图 3-1-5 中阴影部分.

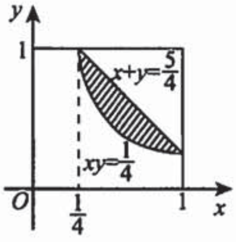


图 3-1-5

$$P(A)=\frac{S_A}{S_\Omega}=\frac{\int_{\frac{1}{4}}^{1}\left(\frac{5}{4}-x-\frac{1}{4x}\right)\mathrm{d}x}{1}=\left(\frac{5}{4}x-\frac{1}{2}x^2-\frac{1}{4}\ln x\right)\Bigg|_{\frac{1}{4}}^{1}$$

$$=\frac{15}{32}-\frac{1}{2}\ln 2.$$

3.1.4 解 设两个人到达的时间分别为 x,y，则(x,y)的取值范围为边长为 60 的正方形区域，设两个人会面的事件为 A.

则 A 成立的范围是$\{|x-y|\leqslant 20\}$，即有当 $x\geqslant y$ 时，$x-y\leqslant 20$；当 $x<y$ 时，$x-y>-20$. 于是，如图 3-1-6 所示，有

$$P(A)=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

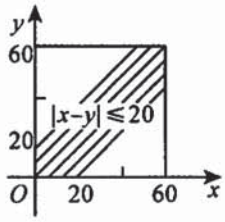


图 3-1-6

3.1.5 解 因为

$$P(AB)=P(\overline{A}\cap\overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-P(A)-P(B)+P(AB),$$

由此得
$$1-P(A)-P(B)=0,$$
所以
$$P(B)=1-P(A)=1-p.$$

3.1.6 解 因为 $0.3=P(A-B)=P(A)-P(AB)=0.7-P(AB)$，由此得 $P(AB)=0.7-0.3=0.4$，故

$$P(\overline{AB})=1-P(AB)=0.6.$$

3.1.7 解 记事件 A 为“目标被击中”，事件 B_1 为“甲射中目标”，事件 B_2 为“乙射中目标”，因为 $A=B_1\cup B_2$，所以

$$P(A)=P(B_1\cup B_2)=P(B_1)+P(B_2)-P(B_1B_2)$$
$$=0.6+0.7-0.6\times0.7=0.88.$$

考虑到 $B_1\subset A$，故有
$$P(B_1|A)=\frac{P(AB_1)}{P(A)}=\frac{P(B_1)}{P(A)}=\frac{0.6}{0.88}\approx0.682.$$

3.1.8 解 记 $A=\{$第一次取出的产品是一等品$\}$，$B=\{$第二次取出的产品是一等品$\}$，则

$$p=P(A),\quad q=P(B|A)=\frac{P(AB)}{P(A)}.$$

要计算 p,q，仅需求出 $P(A)$，$P(AB)$. 而事件 A，AB 都是试验的结果，它与其前提条件产品取自哪一箱有关. 因此我们自然想到：写出前提条件，应用全概率公式计算 $P(A)$，$P(AB)$.

记 $C_i=\{$产品取自第 i 箱$\}(i=1,2)$，则

$$C_1C_2=\varnothing,\quad C_1\cup C_2=\Omega,\quad A=AC_1+AC_2,\quad AB=ABC_1+ABC_2,$$

故
$$P(A)=P(AC_1)+P(AC_2)=P(C_1)P(A|C_1)+P(C_2)P(A|C_2)$$
$$=\frac{1}{2}\times\frac{10}{50}+\frac{1}{2}\times\frac{18}{30}=\frac{2}{5},$$

$$
\begin{aligned}
P(AB)&=P(ABC_1)+P(ABC_2)=P(C_1)P(AB|C_1)+P(C_2)P(AB|C_2)\\
&=\frac{1}{2}\times\frac{10}{50}\times\frac{9}{49}+\frac{1}{2}\times\frac{18}{30}\times\frac{17}{29}=\frac{276}{1\,421},
\end{aligned}
$$

$$p=P(A)=\frac{2}{5},\quad q=\frac{P(AB)}{P(A)}=\frac{276}{1\,421}\times\frac{5}{2}=\frac{690}{1\,421}.$$

3.1.9　解　记事件 A 为“取出的是白球”，事件 B 为“原来那个球是白球”. 容易看出：

$$P(A|B)=1,\quad P(A|\overline{B})=\frac{1}{2}.$$

另外，由于不知道袋中原来那个球的颜色，故 $P(B)=P(\overline{B})=\frac{1}{2}$，由贝叶斯公式得

$$P(B|A)=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\overline{B})P(A|\overline{B})}=\frac{\frac{1}{2}\times 1}{\frac{1}{2}\times 1+\frac{1}{2}\times\frac{1}{2}}=\frac{2}{3}.$$

3.1.10　(C)　**解**　$A\cup B$ 与 C 相互独立，即 $P[(A\cup B)\cap C]=P(A\cup B)P(C)$.

由
$$
\begin{aligned}
P[(A\cup B)\cap C]&=P(AC\cup BC)\\
&=P(AC)+P(BC)-P(AC\cap BC)\\
&=P(A)P(C)+P(B)P(C)-P(ABC),
\end{aligned}
$$

而
$$
\begin{aligned}
P(A\cup B)P(C)&=[P(A)+P(B)-P(AB)]P(C)\\
&=P(A)P(C)+P(B)P(C)-P(AB)P(C),
\end{aligned}
$$

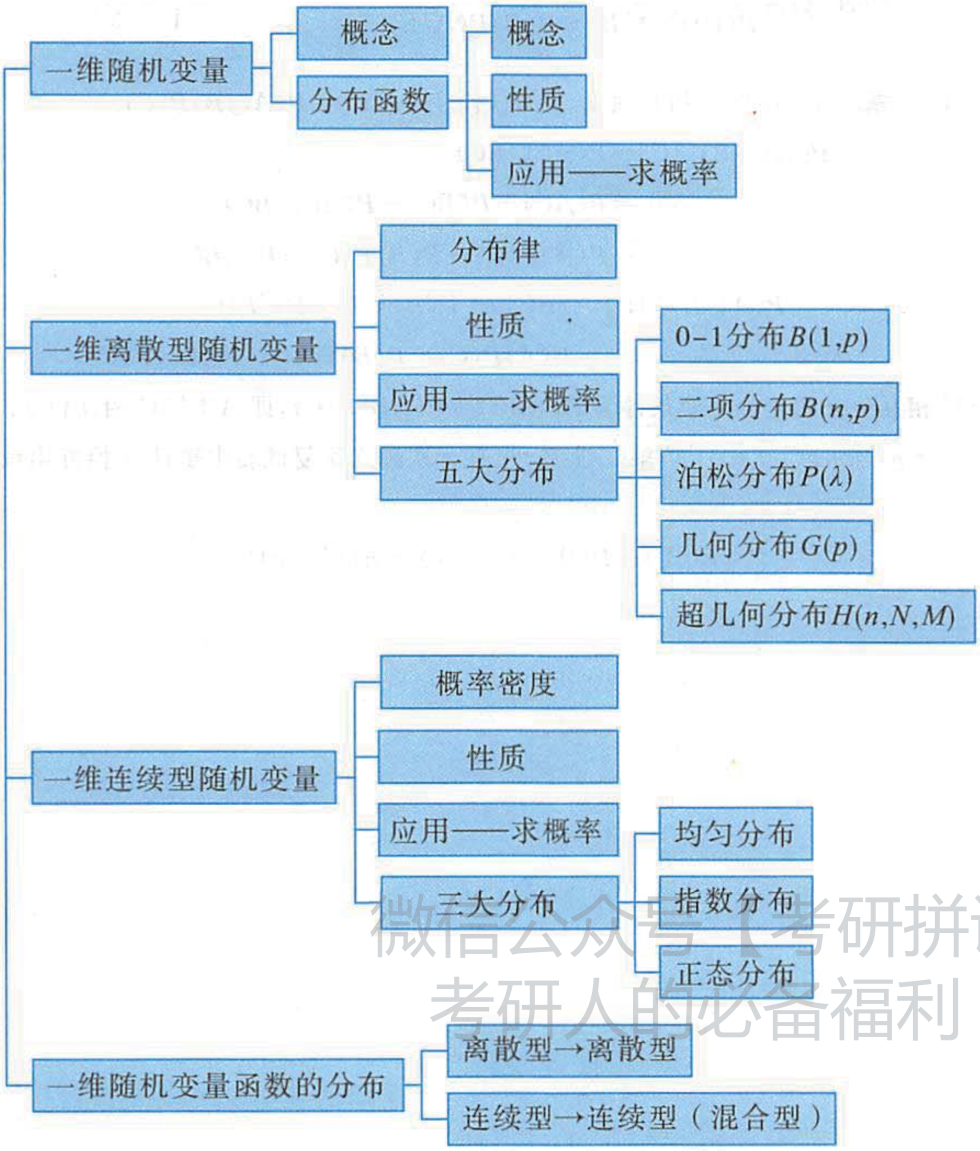
所以 $A\cup B$ 与 C 相互独立的充分必要条件为 $P(ABC)=P(AB)P(C)$，即 AB 与 C 相互独立. 应选(C).

3.1.11　$(1-p)^n+np(1-p)^{n-1}$　**解**　设 $B_k=\{$在 n 次独立重复试验中事件 A 恰好出现 k 次$\}(k=0,1)$，则

$$P=P(B_0)+P(B_1)=(1-p)^n+np(1-p)^{n-1}.$$

第2讲 一维随机变量及其分布

基础知识结构



基础内容精讲

一、随机变量及其分布函数的概念、性质及应用

1. 随机变量的概念

随机变量就是“其值会随机而定”的变量. 设随机试验 E 的样本空间为 $\Omega=\{\omega\}$，如果对每一个 $\omega\in\Omega$，都有唯一的实数 $X(\omega)$ 与之对应，并且对任意实数 x，$\{\omega|X(\omega)\leqslant x,\omega\in\Omega\}$ 是随机事件，则称定义在 Ω 上的实值单值函数 $X(\omega)$ 为随机变量. 简记为随机变量 X. 一般用大写字母 $X,Y,Z,\cdots$ 或希腊字母 $\xi,\eta,\zeta,\cdots$ 来表示随机变量.

【注】 (1)随机事件是从静态的观点来研究随机现象，而随机变量则是一种动态的观点，如高等数学中常量与变量的区别与联系.

(2)随机变量的实质是“实值单值函数”，这个定义不同于高等数学中函数的定义(其“定义域”一定是实数集)，它的“定义域”不一定是实数集，这点需要读者注意.

2. 分布函数的概念及性质

(1)概念.

设 X 是随机变量，x 是任意实数，称函数 $F(x)=P\{X\leqslant x\}(x\in\mathbf{R})$ 为随机变量 X 的分布函数，或称 X 服从分布 $F(x)$，记为 $X\sim F(x)$.

【注】 分布函数完整地描述了随机变量的概率规律性.

(2)性质(也是充要条件).

①$F(x)$ 是 x 的单调不减函数，即对任意实数 $x_1<x_2$，有 $F(x_1)\leqslant F(x_2)$；

②$F(x)$ 是 x 的右连续函数，即对任意 $x_0\in\mathbf{R}$，有 $\lim\limits_{x\to x_0^+}F(x)=F(x_0+0)=F(x_0)$；

③$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0,F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$.

【注】 (1)务必记住分布函数是事件的概率，由此知 $0\leqslant F(x)\leqslant 1$，即 $F(x)$ 是有界函数.

(2)满足以上三条性质的函数 $F(x)$ 必是某个随机变量的分布函数，所以，这三条性质也是判断函数 $F(x)$ 是否为某一随机变量 X 的分布函数的充要条件.

3. 分布函数的应用——求概率

$P\{X\leqslant a\}=F(a)$；

$P\{X<a\}=F(a-0)$；

$P\{X=a\}=F(a)-F(a-0)$.

二、常见的两类随机变量——离散型随机变量和连续型随机变量

1. 离散型随机变量及其概率分布

如果随机变量 X 只可能取有限个或可列个值 $x_1,x_2,\cdots$，则称 X 为**离散型随机变量**，称

$$P\{X=x_i\}=p_i,i=1,2,\cdots$$

为 X 的**分布列**、**分布律**或**概率分布**，记为 $X\sim p_i$，概率分布常常用表格形式或矩阵形式表示，即

X	x_1	x_2	$\cdots$
P	p_1	p_2	$\cdots$

或 $X\sim\begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$.

数列 $\{p_i\}$ 是离散型随机变量的概率分布的充要条件：$p_i\geqslant 0(i=1,2,\cdots)$，且 $\sum\limits_i p_i=1$.

设离散型随机变量 X 的概率分布为 $P\{X=x_i\}=p_i$，则 X 的分布函数

$$F(x)=P\{X\leqslant x\}=\sum_{x_i\leqslant x}P\{X=x_i\},$$

$$P\{X=x_i\}=P\{X\leqslant x_i\}-P\{X<x_i\}=F(x_i)-F(x_i-0),$$

并且对实数轴上的任一集合 B 有

$$P\{X\in B\}=\sum_{x_i\in B}P\{X=x_i\},$$

特别地，
$$P\{a<X\leqslant b\}=P\{X\leqslant b\}-P\{X\leqslant a\}=F(b)-F(a).$$

2. 连续型随机变量及其概率密度

如果随机变量 X 的分布函数可以表示为

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t\ (x\in\mathbf{R}),$$

其中 $f(x)$ 是非负可积函数，则称 X 为**连续型随机变量**，称 $f(x)$ 为 X 的**概率密度函数**，简称**概率密度**，记为 $X\sim f(x)$.

$f(x)$ 为某一随机变量 X 的概率密度的充分必要条件：$f(x)\geqslant 0$，且 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$（由此可知，改变 $f(x)$ 有限个点的值，$f(x)$ 仍然是概率密度）.

设 X 为连续型随机变量，$X\sim f(x)$，则对任意实数 c 有 $P\{X=c\}=0$；对实数轴上任一集合 B 有

$$P\{X\in B\}=\int_B f(x)\mathrm{d}x,$$

特别地，

$$P\{a<X<b\}=P\{a\leqslant X<b\}=P\{a<X\leqslant b\}=P\{a\leqslant X\leqslant b\}=\int_a^b f(x)\mathrm{d}x=F(b)-F(a).$$

【注】 (1)"概率密度"这个名词的由来可解释如下：取定一个点 x，则按分布函数的定义，事件 $\{x<X\leqslant x+h\}$ 的概率 $(h>0$，为常数）应为 $F(x+h)-F(x)$. 所以，比值 $[F(x+h)-F(x)]/h$ 可以解释为在点 x 附近长为 h 的区间 $(x,x+h]$ 内，单位长所占的概率. 令 $h\to 0$，则这个比值的极限，即 $F'(x)=f(x)$，也就是在点 x 处（无穷小区间内）单位长的概率. 或者说，它反映了概率在点 x 处的"密集程度". 设想一条极细的无穷长的金属杆，总质量为 1，概率密度相当于杆上各点的质量密度.

(2) $P\{a<X<b\}=\int_a^b f(x)\mathrm{d}x$ 意味着 X 落入某一区间的概率等于该区间之上、概率密度曲线之下曲边梯形的面积. 应用概率的这种几何意义，常常有助于问题的分析与求解.

(3) 设 $X\sim f(x)$，则 X 的分布函数 $F(x)$ 是 x 的连续函数. 在 $f(x)$ 的连续点 x_0 处有 $F'(x_0)=f(x_0)$. 如果 $F(x)$ 是连续函数，除有限个点外，$F'(x)$ 存在且连续，则 X 为连续型随机变量，且 $f(x)=F'(x)$（在 $F'(x)$ 不存在的地方可以令 $f(x)=0$ 或取其他值）.

三、常见的随机变量分布类型

1. 离散型

(1)0—1分布 $B(1,p)$.

如果 X 的概率分布为 $X\sim\begin{pmatrix}1 & 0\\ p & 1-p\end{pmatrix}$，即 $P\{X=1\}=p,P\{X=0\}=1-p$，则称 X 服从参数为 p 的**0—1分布**，记为 $X\sim B(1,p)(0<p<1)$.

(2)二项分布 $B(n,p)$.

如果 X 的概率分布为 $P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}(k=0,1,\cdots,n;0<p<1)$，则称 X 服从参数为 (n,p) 的**二项分布**，记为 $X\sim B(n,p)$.

【注】 如果 X 是 n 重伯努利试验中事件 A 发生的次数，则 $X\sim B(n,p)$，其中 $p=P(A)$. 这个结论在解题中会经常用到.

(3)泊松分布 $P(\lambda)$.

如果 X 的概率分布为 $P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}(k=0,1,\cdots;\lambda>0)$，则称 X 服从参数为 λ 的**泊松分布**，记为 $X\sim P(\lambda)$.

(4)几何分布 $G(p)$.

如果 X 的概率分布为 $P\{X=k\}=(1-p)^{k-1}p\ (k=1,2,\cdots;0<p<1)$，则称 X 服从参数为 p 的**几何分布**，记为 $X\sim G(p)$.

【注】 设 X 表示伯努利试验中事件 A 首次发生所需做的试验次数，则 $X\sim G(p)$，其中 $p=P(A)$.

(5)超几何分布 $H(n,N,M)$.

如果 X 的概率分布为 $P\{X=k\}=\dfrac{\mathrm{C}_M^k\mathrm{C}_{N-M}^{n-k}}{\mathrm{C}_N^n}$ $(\max\{0,n-N+M\}\leqslant k\leqslant\min\{M,n\};M,N,n$ 为正整数且 $M\leqslant N,n\leqslant N,k$ 为整数)，则称 X 服从参数为 (n,N,M) 的**超几何分布**，记为 $X\sim H(n,N,M)$.

2. 连续型

(1)均匀分布 $U(a,b)$.

如果随机变量 X 的概率密度或分布函数分别为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其他},\end{cases}\qquad F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b,\end{cases}$$

则称 X 在区间 (a,b) 上服从**均匀分布**，记为 $X\sim U(a,b)$.(如图3-2-1,3-2-2)

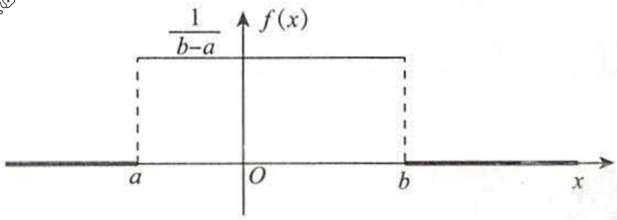


图 3-2-1

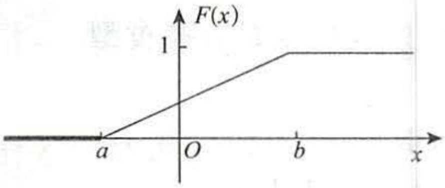


图 3-2-2

> **【注】** 区间(a,b)可以是闭区间$[a,b]$；几何概型是均匀分布的实际背景，用几何概率计算事件概率时，已假设点在区域内服从均匀分布；几何概率可以用均匀分布计算.

(2)指数分布 $E(\lambda)$.

如果 X 的概率密度或分布函数分别为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & 其他,\end{cases}\quad F(x)=\begin{cases}1-e^{-\lambda x}, & x\geqslant 0,\\ 0, & x<0\end{cases}\quad(\lambda>0),$$

则称 X 服从参数为 λ 的**指数分布**，记为 $X\sim E(\lambda)$.（如图 3-2-3，3-2-4）

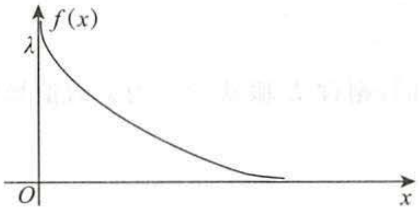


图 3-2-3

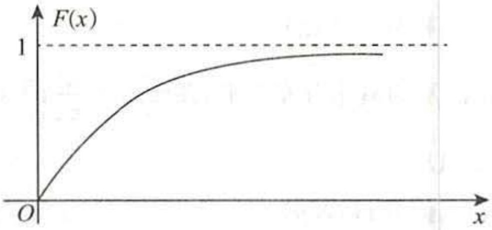


图 3-2-4

(3)正态分布 $N(\mu,\sigma^2)$.

如果 X 的概率密度

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\quad(-\infty<x<+\infty),$$

其中 $-\infty<\mu<+\infty,\sigma>0$，则称 X 服从参数为(μ,σ^2)的**正态分布**或称 X 为**正态变量**，记为 $X\sim N(\mu,\sigma^2)$. 此时 $f(x)$的图形关于直线 $x=\mu$ 对称，即 $f(\mu-x)=f(\mu+x)$，并在 $x=\mu$ 处有唯一最大值 $f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$. 如图 3-2-5 所示.

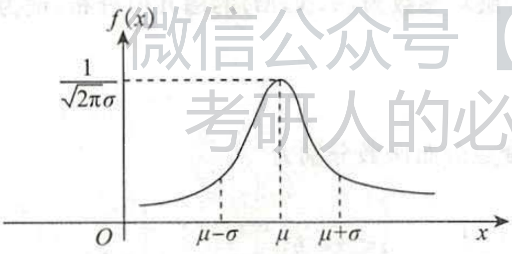


图 3-2-5

称 $\mu=0,\sigma=1$ 时的正态分布 $N(0,1)$ 为**标准正态分布**，通常记标准正态分布的概率密度为 $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$，分布函数为 $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$. 显然 $\varphi(x)$ 为偶函数，则

$$\Phi(0)=\frac{1}{2},\quad \Phi(-x)=1-\Phi(x).$$

如图 3-2-6 所示.

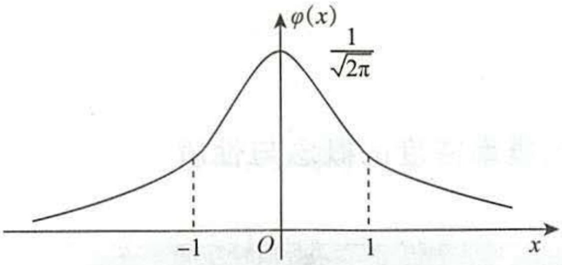


图 3-2-6

若 $X\sim N(0,1)$，$P\{X>\mu_\alpha\}=\alpha$，则称 μ_α 为标准正态分布的上侧 α 分位数（上 α 分位点）.

如果 $X\sim N(\mu,\sigma^2)$，则

$$F(x)=P\{X\leqslant x\}=\Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$F(\mu-x)+F(\mu+x)=1,$$

$$P\{a<X<b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right),$$

$$aX+b\sim N(a\mu+b,a^2\sigma^2)(a\neq0).$$

四、一维随机变量函数的分布

1. 概念

设 X 为随机变量，函数 $y=g(x)$，则以随机变量 X 作为自变量的函数 $Y=g(X)$ 也是随机变量，称为随机变量 X 的函数. 例如：$Y=aX^2+bX+c$，$Y=|X-a|$，$Y=\begin{cases}X, & X\leqslant1,\\ 1, & X>1,\end{cases}$ 等等.

2. 随机变量函数的分布

（1）离散型.

设 X 为离散型随机变量，其概率分布为 $p_i=P\{X=x_i\}(i=1,2,\cdots)$，则 X 的函数 $Y=g(X)$ 也是离散型随机变量，其概率分布为 $P\{Y=g(x_i)\}=p_i(i=1,2,\cdots)$，即

$$Y\sim\begin{bmatrix}g(x_1) & g(x_2) & \cdots\\ p_1 & p_2 & \cdots\end{bmatrix}.$$

如果有若干个 $g(x_k)$ 相同，则合并诸项为一项 $g(x_k)$，并将相应概率相加作为 Y 取 $g(x_k)$ 值的概率.

（2）连续型.

设 X 为连续型随机变量，其分布函数、概率密度分别为 $F_X(x)$ 与 $f_X(x)$，随机变量 $Y=g(X)$ 是 X 的函数，则 Y 的分布函数或概率密度可用分布函数法求得.

$$F_Y(y)=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}=\int_{g(X)\leqslant y}f_X(x)\mathrm{d}x.$$

如果 $F_Y(y)$ 连续，且除有限个点外，$F'_Y(y)$ 存在且连续，则 Y 的概率密度 $f_Y(y)=F'_Y(y)$.

基础例题精解

一、分布函数、概率分布、概率密度的概念与性质

(1)$F(x)$是分布函数$\Leftrightarrow F(x)$是x的单调不减、右连续函数，且$F(-\infty)=0,F(+\infty)=1$.

【注】 若$F(x)$为分段函数，必须在分段点处考虑右连续.

(2)$\{p_i\}$是概率分布$\Leftrightarrow p_i\geqslant 0$，且$\sum\limits_i p_i=1$.

(3)$f(x)$是概率密度$\Leftrightarrow f(x)\geqslant 0$，且$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$.

例 3.2.1 已知离散型随机变量X的正概率点为$-1,0,2$，它们各自的概率互不相等且成等差数列，则X的分布列为________，分布函数为________，$P\{|X|\leqslant 1|X\geqslant 0\}=$________.

解 应填 $X\sim\begin{pmatrix}-1 & 0 & 2\\ \frac{1}{3}+d & \frac{1}{3} & \frac{1}{3}-d\end{pmatrix}$，$-\frac{1}{3}<d<\frac{1}{3}$，$d\neq 0$；$F(x)=\begin{cases}0, & x<-1,\\ \frac{1}{3}+d, & -1\leqslant x<0,\\ \frac{2}{3}+d, & 0\leqslant x<2,\\ 1, & x\geqslant 2;\end{cases}$ $\frac{1}{2-3d}$.

设$P\{X=-1\}=a+d,P\{X=0\}=a,P\{X=2\}=a-d$，其中$0<a+d<1,0<a<1,0<a-d<1$，由离散型随机变量$X$分布列的性质，有

$$a+d+a+a-d=3a=1,$$

解得$a=\frac{1}{3}$，且$-\frac{1}{3}<d<\frac{1}{3}$，$d\neq 0$，因此$X\sim\begin{pmatrix}-1 & 0 & 2\\ \frac{1}{3}+d & \frac{1}{3} & \frac{1}{3}-d\end{pmatrix}$，其中$-\frac{1}{3}<d<\frac{1}{3}$，$d\neq 0$.

所以

$$F(x)=\begin{cases}0, & x<-1,\\ \frac{1}{3}+d, & -1\leqslant x<0,\\ \frac{2}{3}+d, & 0\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

$$P\{|X|\leqslant 1|X\geqslant 0\}=\frac{P\{0\leqslant X\leqslant 1\}}{P\{X\geqslant 0\}}=\frac{1}{2-3d}.$$

例 3.2.2 设 $F_1(x)$，$F_2(x)$是随机变量的分布函数，$f_1(x)$，$f_2(x)$是相应的概率密度，则( ).

(A)$F_1(x)+F_2(x)$是分布函数 (B)$F_1(x)\cdot F_2(x)$是分布函数

(C)$f_1(x)+f_2(x)$是概率密度 (D)$f_1(x)\cdot f_2(x)$是概率密度

解 应选(B).

根据连续型随机变量分布函数和概率密度的性质，有

$$F_1(+\infty)+F_2(+\infty)=2,\quad \int_{-\infty}^{+\infty}[f_1(x)+f_2(x)]\mathrm{d}x=2,$$

$$\int_{-\infty}^{+\infty}f_1(x)f_2(x)\mathrm{d}x\neq\int_{-\infty}^{+\infty}f_1(x)\mathrm{d}x\int_{-\infty}^{+\infty}f_2(x)\mathrm{d}x=1,$$

应排除选项(A)，(C)，(D)，故选择(B). 事实上，由于 $F_1(x)\cdot F_2(x)$ 是单调不减的右连续函数，且

$$F_1(-\infty)\cdot F_2(-\infty)=0,\quad F_1(+\infty)\cdot F_2(+\infty)=1,$$

可直接判断选项(B)正确.

二、分布函数、概率分布、概率密度之间的关系与转换

(1) 已知$\{p_i\}$为概率分布，则分布函数 $F(x)=P\{X\leqslant x\}=\sum\limits_{x_i\leqslant x}P\{X=x_i\}=\sum\limits_{x_i\leqslant x}p_i$；已知概率密度为 $f(x)$，则分布函数 $F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t$.

(2)反之，已知 X 的分布函数为 $F(x)$，那么首先要判断 X 是什么类型的随机变量：如果 $F(x)$是阶梯函数，则 X 为离散型的，X 可能取得的值 x_i 必为 $F(x)$的间断点，且 $p_i=P\{X=x_i\}=F(x_i)-F(x_i-0)$；如果 $F(x)$是连续函数，且除有限个点外，$F'(x)$存在且连续，则 X 是连续型随机变量，其概率密度 $f(x)=F'(x)$.

(3)除上述两种情况，X 常常是混合型的随机变量. 见例 3.2.19.

例 3.2.3 已知随机变量 X 的概率分布为

X	1	2	3
$P\{X=k\}$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

且 $P\{X\geqslant 2\}=\frac{3}{4}$，求未知参数 θ 及 X 的分布函数 $F(x)$.

解 由 $P\{X\geqslant2\}=P\{X=2\}+P\{X=3\}=2\theta(1-\theta)+(1-\theta)^2=1-\theta^2=\frac{3}{4}$，解得 $\theta=\pm\frac{1}{2}$.

又 $P\{X=2\}=2\theta(1-\theta)\geqslant0$，故取 $\theta=\frac{1}{2}$，从而得 X 的概率分布

X	1	2	3
$P\{X=k\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

X 的分布函数

$$F(x)=P\{X\leqslant x\}=\sum_{i\leqslant x}P\{X=i\}=\begin{cases}0, & x<1,\\ \frac{1}{4}, & 1\leqslant x<2,\\ \frac{3}{4}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

例 3.2.4 已知 X 的分布函数为 $F(x)$，概率密度为 $f(x)$. 当 $x\leqslant 0$ 时，$f(x)$ 连续，且 $f(x)=F(x)$，若 $F(0)=1$，求 $F(x)$，$f(x)$.

解 由于 $F(x)$ 是单调不减的函数，故对任意的 $x>0$，有 $F(x)=1$. 又 $x\leqslant 0$ 时，$f(x)$ 连续，故 $f(x)=F'(x)$，已知 $f(x)=F(x)$，由微分方程 $F'(x)=F(x)$，$F(0)=1$，解得 $F(x)=\mathrm{e}^x(x\leqslant 0)$，从而得

$$F(x)=\begin{cases}\mathrm{e}^x, & x<0,\\ 1, & x\geqslant 0,\end{cases}\quad f(x)=\begin{cases}\mathrm{e}^x, & x<0,\\ 0, & \text{其他}.\end{cases}$$

【注】 当分布函数是不连续的分段函数时，为保证其右连续性，自变量 x 的分段小区间除第一个区间外，其余都应是左闭右开的.

例 3.2.5 随机变量 X 的概率密度为

$$f(x)=\begin{cases}2\left(1-\frac{1}{x^2}\right), & 1\leqslant x\leqslant 2,\\ 0, & \text{其他}.\end{cases}$$

求 X 的分布函数 $F(x)$.

解 当 $1\leqslant x<2$ 时，

$$F(x)=\int_1^x 2\left(1-\frac{1}{t^2}\right)\mathrm{d}t=\left(2t+\frac{2}{t}\right)\Big|_1^x$$

$$=2\left(x+\frac{1}{x}\right)-4=2\left(x+\frac{1}{x}-2\right);$$

当 $x<1$ 时，

$$F(x)=\int_{-\infty}^x 0\mathrm{d}t=0;$$

当 $x\geqslant 2$ 时，

$$F(x)=\int_1^2 2\left(1-\frac{1}{t^2}\right)\mathrm{d}t=2\left(t+\frac{1}{t}\right)\Big|_1^2=2\times\frac{1}{2}=1.$$

故

$$F(x)=\begin{cases}0, & x<1,\\ 2\left(x+\frac{1}{x}-2\right), & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

三、利用分布计算概率及其逆问题

已知 X 的分布，则可以求出用 X 的取值范围表示的事件的概率.

如果已知 X 的分布函数 $F(x)$，则 $P\{X\leqslant a\}=F(a)$，$P\{X<a\}=F(a-0)$，由此应用概率性质可以求得 X 在任一区间内的概率，例如，

$$P\{X=a\}=P\{X\leqslant a\}-P\{X<a\}=F(a)-F(a-0),$$
$$P\{a<X<b\}=P\{X<b\}-P\{X\leqslant a\}=F(b-0)-F(a),$$

等等.

（1）如果已知 X 的概率分布 $\{p_i\}$，则对实数轴上任意集合 B 有

$$P\{X\in B\}=\sum_{x_i\in B}P\{X=x_i\}.$$

（2）如果已知 X 的概率密度 $f(x)$，则对实数轴上任意集合 B 有

$$P\{X\in B\}=\int_B f(x)\,\mathrm{d}x.$$

特别地，

$$P\{X=c\}=0(\forall c\in\mathbf{R}),$$
$$P\{a<X<b\}=\int_a^b f(x)\mathrm{d}x.$$

（3）反问题：已知某个事件的概率值求与其有关的未知参数，则只要写出与此事件有关的相应公式及概率，反解等式或不等式，即可求得未知参数或其取值范围.

例 3.2.6 设随机变量 X 的概率分布为 $P\{X=k\}=Ak(k=1,2,3,4,5)$，求常数 A 及概率 $P\left\{\frac{1}{2}<X<\frac{5}{2}\right\}$.

解 由于 $\sum\limits_{k=1}^{5}P\{X=k\}=1$，即 $A\sum\limits_{k=1}^{5}k=15A=1$，解得 $A=\frac{1}{15}$，且

$$P\left\{\frac{1}{2}<X<\frac{5}{2}\right\}=P\{X=1\}+P\{X=2\}=\frac{1}{15}+\frac{2}{15}=\frac{1}{5}.$$

例 3.2.7 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}Ax, & 1<x<2,\\ B, & 2\leqslant x<3,\\ 0, & \text{其他},\end{cases}$$

且 $P\{1<X<2\}=P\{2<X<3\}$，求常数 A,B，分布函数 $F(x)$ 及概率 $P\{2<X<4\}$.

解 由于 $$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_1^2 Ax\,\mathrm{d}x+\int_2^3 B\mathrm{d}x=\frac{3}{2}A+B.$$

又 $P\{1<X<2\}=P\{2<X<3\}$，即 $\int_1^2 Ax\,\mathrm{d}x=\int_2^3 B\mathrm{d}x$，$\frac{3}{2}A=B$，解得 $A=\frac{1}{3}$，$B=\frac{1}{2}$，且

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\begin{cases}0, & x<1,\\ \int_1^x \frac{1}{3}t\mathrm{d}t, & 1\leqslant x<2,\\ \int_1^2\frac{1}{3}t\mathrm{d}t+\int_2^x\frac{1}{2}\mathrm{d}t, & 2\leqslant x<3,\\ 1, & x\geqslant 3\end{cases}$$

$$=\begin{cases}0, & x<1,\\ \frac{1}{6}(x^2-1), & 1\leqslant x<2,\\ \frac{1}{2}(x-1), & 2\leqslant x<3,\\ 1, & x\geqslant 3,\end{cases}$$

$$P\{2<X<4\}=F(4-0)-F(2)=1-\frac{1}{2}=\frac{1}{2},$$

或

$$P\{2<X<4\}=\int_2^4 f(x)\mathrm{d}x=\int_2^3\frac{1}{2}\mathrm{d}x=\frac{1}{2}.$$

例 3.2.8 设 $X\sim N(2,\sigma^2)$，且 $P\{2<X<4\}=0.3$，求 $P\{X<0\}$.

解 已知 $X\sim N(2,\sigma^2)$，故

$$P\{2<X<4\}=\Phi\left(\frac{4-2}{\sigma}\right)-\Phi\left(\frac{2-2}{\sigma}\right)=\Phi\left(\frac{2}{\sigma}\right)-0.5=0.3,\quad \Phi\left(\frac{2}{\sigma}\right)=0.8,$$

所以 $P\{X<0\}=\Phi\left(\frac{0-2}{\sigma}\right)=1-\Phi\left(\frac{2}{\sigma}\right)=0.2.$

【注】 结合图 3-2-7，本题应用面积求解更为直观、简便.

已知 $P\{2<X<4\}=0.3$，故

$$P\{X<0\}=0.5-0.3=0.2.$$

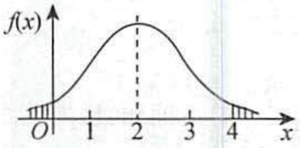


图 3-2-7

例 3.2.9 设随机变量 X 的概率密度为 $f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & 其他,\end{cases}$ Y 表示对 X 的 3 次独立重复试验中事件 $\left\{X\leqslant\frac{1}{2}\right\}$ 出现的次数，求 $P\{Y=2\}$.

解 依题意 $Y\sim B(3,p)$，其中 $p=P\left\{X\leqslant\frac{1}{2}\right\}=\int_0^{\frac{1}{2}}2x\mathrm{d}x=\frac{1}{4}$，故

$$P\{Y=2\}=\mathrm{C}_3^2p^2(1-p)=\frac{9}{64}.$$

四、八个常见随机变量分布类型及其应用

1. 离散型随机变量的分布

(1)0－1 分布.

0－1 分布就是 $n=1$ 情况下的二项分布，即只进行一次试验，该事件发生的概率为 p，不发生的概率为 $q=1-p$. 这是一个最简单的分布，只有两种结果的随机现象，比如抛硬币观察正、反面，新生儿是男还是女，检查产品是否合格等，都可用它来描述.

例 3.2.10 设 X 服从 0−1 分布，其分布律为 $P\{X=k\}=p^k(1-p)^{1-k}$，$k=0,1$，求 X 的分布函数，并作出其图形.

解 当 $x<0$ 时，$F(x)=P\{X\leqslant x\}=0$；

当 $0\leqslant x<1$ 时，

$$F(x)=P\{X\leqslant x\}=P\{X=0\}=1-p;$$

当 $x\geqslant 1$ 时，

$$F(x)=P\{X\leqslant x\}=P\{X=0\}+P\{X=1\}=(1-p)+p=1.$$

所以

$$F(x)=\begin{cases}0, & x<0,\\ 1-p, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

作图，如图 3-2-8 所示.

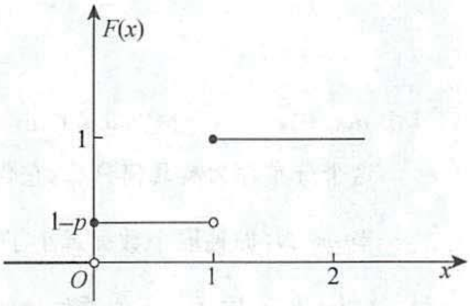


图 3-2-8

(2)二项分布 $B(n,p)$.

例 3.2.11 设 $X\sim B(2,p)$，$Y\sim B(4,p)$，且 $P\{X\geqslant 1\}=\frac{5}{9}$，则 $P\{Y\geqslant 1\}=$(　　).

(A)$\frac{65}{81}$　　(B)$\frac{16}{81}$　　(C)1　　(D)$\frac{4}{7}$

解 应选(A).

由题设 $P\{X\geqslant 1\}=\frac{5}{9}=1-P\{X=0\}=1-C_2^0p^0(1-p)^2$，即 $(1-p)^2=\frac{4}{9}$，解得 $p=\frac{1}{3}$，从而有

$$P\{Y\geqslant 1\}=1-P\{Y=0\}=1-C_4^0p^0(1-p)^4=1-\left(\frac{2}{3}\right)^4=\frac{65}{81},$$

故选择(A).

(3)几何分布及其推广.

例 3.2.12 设随机变量 X 的概率密度为 $f(x)=\begin{cases}2^{-x}\ln 2, & x>0,\\ 0, & \text{其他}.\end{cases}$ 对 X 进行独立重复的观测，直到第 2 个大于 3 的观测值出现时停止，记 Y 为观测次数. 求 Y 的概率分布.

解 记 p 为“观测值大于 3”的概率，则 $p=P\{X>3\}=\int_3^{+\infty}2^{-x}\ln 2\,\mathrm{d}x=\frac{1}{8}$.

依题意 Y 为离散型随机变量，而且取值为 2，3，…，则 Y 的概率分布为

$$P\{Y=k\}=C_{k-1}^1p(1-p)^{k-2}\cdot p=C_{k-1}^1p^2(1-p)^{k-2}=(k-1)\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^{k-2},k=2,3,\cdots.$$

(4)超几何分布.

从一个有限总体中进行不放回抽样常会遇到超几何分布.

设由 N 个产品组成的总体，其中含有 M 个不合格品. 若从中随机不放回地抽取 n 个，则其中含有不

合格品的个数 X 是一个离散型随机变量．假如 $n\leqslant M$，则 X 可能取 $0,1,\cdots,n$；若 $n>M$，则 X 可能取 $0,1,\cdots,M$．由古典概率方法容易算得

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\qquad (*)$$

其中 $\max\{0,n-N+M\}\leqslant k\leqslant\min\{M,n\}$；$M,N,n$ 为正整数且 $M\leqslant N,n\leqslant N,k$ 为整数．

这个分布称为**超几何分布**，它含有三个参数 N,M 和 n，记为 $H(n,N,M)$．

当 $n\ll N$（即抽取个数 n 远小于产品总数 N）时，每次抽取后，总体中的不合格品率 $p=\frac{M}{N}$ 改变甚微，这时不放回抽样可近似看作放回抽样，这时超几何分布可用二项分布近似．

例 3.2.13 20 个产品中有 5 个不合格品，若从中随机取出 8 个，试求其中不合格品数 X 的概率分布．

解 按题意有 $N=20,M=5,n=8$．由（ * ）式可算得

$$P\{X=0\}=\frac{C_{15}^8}{C_{20}^8}=\frac{6\ 435}{125\ 970}=0.0511,$$

$$P\{X=1\}=\frac{C_5^1C_{15}^7}{C_{20}^8}=\frac{32\ 175}{125\ 970}=0.2554,$$

$$P\{X=2\}=\frac{C_5^2C_{15}^6}{C_{20}^8}=\frac{50\ 050}{125\ 970}=0.3973.$$

类似可得 $X=3,4,5$ 的概率，则 X 的概率分布为

X	0	1	2	3	4	5
P	0.0511	0.2554	0.3973	0.2384	0.0542	0.0036

【注】 由此分布可算得各种事件的概率．譬如，不合格品不多于 3 个的概率为

$$\begin{aligned}P\{X\leqslant 3\}&=P\{X=0\}+P\{X=1\}+P\{X=2\}+P\{X=3\}\\&=0.0511+0.2554+0.3973+0.2384=0.9422.\end{aligned}$$

（5）泊松分布．

例 3.2.14 设一本书的各页印刷错误的个数 X 服从泊松分布．已知只有一个和只有两个印刷错误的页数相同，则随意抽查的 4 页中无印刷错误的概率 $p=$______．

解 应填 e^{-8}．

以 X 表示“随意抽取的一页上印刷错误的个数”，以 $X_k(k=1,2,3,4)$ 表示“随意抽取的第 k 页上印刷错误的个数”，由条件知 X 和 $X_k(k=1,2,3,4)$ 服从同一泊松分布 $P\{X=j\}=\frac{\lambda^j}{j!}e^{-\lambda}$，未知参数 λ 取决于

$$P\{X=1\}=P\{X=2\},\quad \frac{\lambda}{1!}e^{-\lambda}=\frac{\lambda^2}{2!}e^{-\lambda},$$

于是 $\lambda=2$．由于随机变量 $X_k(k=1,2,3,4)$ 显然相互独立，因此

$$P\{X_1=0,X_2=0,X_3=0,X_4=0\}=P\{X_1=0\}P\{X_2=0\}P\{X_3=0\}P\{X_4=0\}=(e^{-2})^4=e^{-8}.$$

2. 连续型随机变量的分布

（1）均匀分布 $U(a,b)$．

例 3.2.15 已知 $X\sim U(a,b)(a>0)$，且 $P\{0<X<3\}=\frac{1}{4}, P\{X>4\}=\frac{1}{2}$，求 X 的概率密度及 $P\{1<X<5\}$.

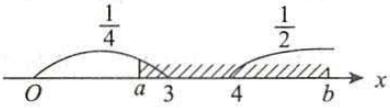


图 3-2-9

解 如图 3-2-9 所示，$X\sim U(a,b)$.

由 $P\{X>4\}=\frac{1}{2}$，且 $P\left\{X>\frac{a+b}{2}\right\}=\frac{1}{2}$，得 $\frac{a+b}{2}=4$，即 $a+b=8$.

由于 $P\{3<X<4\}=1-P\{X\geqslant 4\}-P\{X\leqslant 3\}=1-P\{X>4\}-P\{0<X<3\}=1-\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$，

又 $P\{3<X<4\}=\frac{4-3}{b-a}$，故 $\frac{4-3}{b-a}=\frac{1}{4}$，解得 $b-a=4$.

联立得 $\begin{cases}a+b=8,\\ b-a=4,\end{cases}$ 解得 $\begin{cases}a=2,\\ b=6.\end{cases}$

所以 X 的概率密度为

$$f(x)=\begin{cases}\frac{1}{4}, & 2<x<6,\\ 0, & \text{其他}.\end{cases}$$

$$P\{1<X<5\}=P\{2<X<5\}=\frac{5-2}{6-2}=\frac{3}{4}.$$

(2)指数分布.

参数为 $\lambda(\lambda>0)$ 的指数分布 $E(\lambda)$ 的概率密度为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

例 3.2.16 设随机变量 X 服从参数为 1 的指数分布，a 为常数且大于零，则 $P\{X\leqslant a+1\mid X>a\}=$ ________.

解 应填 $1-\frac{1}{e}$.

方法一 $$P\{X\leqslant a+1\mid X>a\}=\frac{P\{a<X\leqslant a+1\}}{P\{X>a\}}=\frac{\int_a^{a+1}e^{-x}dx}{\int_a^{+\infty}e^{-x}dx}=1-\frac{1}{e}.$$

方法二 由指数分布的无记忆性知

$$P\{X\leqslant a+1\mid X>a\}=1-P\{X>a+1\mid X>a\}=1-P\{X>1\}=P\{X\leqslant 1\}=1-\frac{1}{e}.$$

(3)正态分布.

有关正态分布的问题：

①所有有关正态分布不等式的判别都应先标准化；

②充分利用标准正态分布图形的对称性，有时对于解题可以起到事半功倍的效果；

③要注意利用正态分布的参数与其数字特征的关系求解问题.

例 3.2.17 设随机变量 $X\sim N(2,2^2)$，且 $aX+b\sim N(0,1)$，则 a,b 取值为________.

解 应填 $\frac{1}{2},-1$ 或 $-\frac{1}{2},1$.

方法一 利用正态分布标准化公式，即由 $X\sim N(2,2^2)$，有

$$\frac{X-2}{2}\sim N(0,1),$$

得 $a=\frac{1}{2},b=-1$. 同时有 $-\frac{X-2}{2}\sim N(0,1)$，故也有 $a=-\frac{1}{2},b=1$.

方法二 利用正态分布的参数与其数字特征的关系，有

$$E(aX+b)=aEX+b=2a+b=0,$$
$$D(aX+b)=a^2DX=4a^2=1,$$

解得 $a=\frac{1}{2},b=-1$ 或 $a=-\frac{1}{2},b=1$.

【注】 对于正态分布，若 $X\sim N(0,1)$，则也必有 $-X\sim N(0,1)$. 因此，本题若用正态分布标准化公式求解，可能丢掉另一个解. 若用正态分布的参数与其数字特征的关系求解，就容易得出两个解，做法更为保险.

例 3.2.18 设随机变量 X 服从正态分布 $N(0,1)$，对给定的 $\alpha(0<\alpha<1)$，数 u_α 满足 $P\{X>u_\alpha\}=\alpha$，若 $P\{|X|<x\}=\alpha$，则 x 等于(　　).

(A) $u_{\frac{\alpha}{2}}$　　(B) $u_{1-\frac{\alpha}{2}}$　　(C) $u_{\frac{1-\alpha}{2}}$　　(D) $u_{1-\alpha}$

解 应选(C).

如图 3-2-10 所示，从几何直观看，条件给出 u_α 的定值原则，即为上 α 分位数. 类比可得 $P\{|X|<u_{\frac{1-\alpha}{2}}\}=\alpha$，故选择(C).

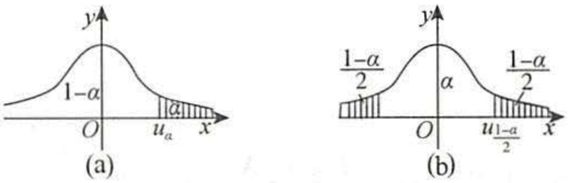


图 3-2-10

3. 一般类型随机变量的概率分布

同时兼有离散型与连续型随机变量特征的随机变量，或未明确分布类型的随机变量都可称为一般类型的随机变量也叫混合型随机变量. 需要强调的是，该类型的分布问题只能以分布函数为工具.

例 3.2.19 一水渠出口闸门挡板是边长为 1(单位)的正方形. 已知初始水面高为 $\frac{3}{4}$(单位)，现发现挡板某一部位出现一个小孔(小孔等可能出现在挡板的任一位置)，水经过小孔流出，求剩余液面高度 X 的分布函数 $F(x)$.

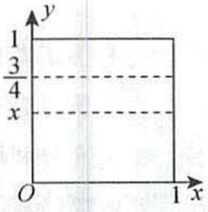


图 3-2-11

解 如图 3-2-11 所示，显然 X 是取非负值的非离散型随机变量，且满足条件 $0\leqslant X\leqslant\frac{3}{4}$，所以当 $x<0$ 时，X 的分布函数 $F(x)=P\{X\leqslant x\}=0$；

当 $0\leqslant x<\frac{3}{4}$ 时，$F(x)=P\{X\leqslant x\}=P\{X<0\}+P\{0\leqslant X\leqslant x\}=P\{0\leqslant X\leqslant x\}$，其中事件 $\{0\leqslant X\leqslant x\}$ 意味着小孔出现在初始水面下部且位于边长为 x 和 1 的长方形内. 由于小孔是等可能地出现在挡板任一位置，根据几何概率知 $P\{0\leqslant X\leqslant x\}=\frac{x\cdot 1}{1\cdot 1}=x$，所以当 $0\leqslant x<\frac{3}{4}$ 时，$F(x)=x$；

当 $x\geqslant\frac{3}{4}$ 时，$F(x)=P\{X\leqslant x\}=1$.

综上可得

$$F(x)=\begin{cases}0, & x<0,\\ x, & 0\leqslant x<\frac{3}{4},\\ 1, & x\geqslant\frac{3}{4}.\end{cases}$$

五、一维随机变量函数的分布

例 3.2.20　设 X 是仅可能取 6 个值的离散型随机变量，其分布为

X	-2	-1	0	1	2	3
P	0.05	0.15	0.20	0.25	0.20	0.15

求 $Y=2X+1,Z=X^2$ 的概率分布.

解　$Y=2X+1$，则 Y 仍是离散型随机变量，它可取 $-3,-1,1,3,5,7$ 这 6 个值. 由于它们没有相同的值，故 Y 取这些值的概率仍如上述，即 Y 的概率分布为

$Y=2X+1$	-3	-1	1	3	5	7
P	0.05	0.15	0.20	0.25	0.20	0.15

$Z=X^2$，虽然 Z 仍是离散型随机变量，但它可能取的 6 个值 4,1,0,1,4,9 中出现了相同的值，Z 取相同值的概率应加起来，如

$$P\{Z=1\}=P\{X=1\}+P\{X=-1\}=0.25+0.15=0.40,$$

则 Z 的概率分布为

$Z=X^2$	0	1	4	9
P	0.20	0.40	0.25	0.15

例 3.2.21　假设随机变量 X 服从参数为 λ 的指数分布，$F(x)$ 是其分布函数，证明：随机变量 $Y=F(X)$ 在区间 $(0,1)$ 上服从均匀分布.

证明　指数分布函数为 $F(x)=1-\mathrm{e}^{-\lambda x}(x\geqslant 0)$. 设 $G(y)=P\{Y\leqslant y\}$ 为 $Y=F(X)$ 的分布函数. 由于分布函数 $F(x)$ 的值域为 $[0,1]$，可见当 $y\leqslant 0$ 时，$G(y)=0$；当 $y\geqslant 1$ 时，$G(y)=1$；当 $0<y<1$ 时，

$$G(y)=P\{Y\leqslant y\}=P\{1-\mathrm{e}^{-\lambda X}\leqslant y\}=P\left\{X\leqslant-\frac{1}{\lambda}\ln(1-y)\right\}$$

$$=F\left[-\frac{1}{\lambda}\ln(1-y)\right]=y,$$

于是 $G(y)$ 是区间 $(0,1)$ 上的均匀分布函数，从而 $Y=F(X)$ 在区间 $(0,1)$ 上服从均匀分布.

【注】 本题更一般的结论见习题 3.2.12.

例 3.2.22 设随机变量 X 的概率密度为 $f_X(x)=\frac{k}{1+x^2}\ (x\in\mathbf{R})$. 求：

(1)常数 k；

(2)随机变量 $Y=1-\sqrt[3]{X}$ 的概率密度 $f_Y(y)$.

解 (1)由于
$$\int_{-\infty}^{+\infty}f_X(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{k}{1+x^2}\mathrm{d}x=k\arctan x\Big|_{-\infty}^{+\infty}=k\pi=1,$$
因此 $k=\frac{1}{\pi}$. X 的概率密度为
$$f_X(x)=\frac{1}{\pi(1+x^2)}\ (x\in\mathbf{R}).$$

(2)随机变量 $Y=1-\sqrt[3]{X}$ 的分布函数为
$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{1-\sqrt[3]{X}\leqslant y\}=P\{X\geqslant(1-y)^3\}\\&=\int_{(1-y)^3}^{+\infty}f_X(x)\mathrm{d}x=\int_{(1-y)^3}^{+\infty}\frac{1}{\pi(1+x^2)}\mathrm{d}x\\&=\frac{1}{\pi}\arctan x\Big|_{(1-y)^3}^{+\infty}=\frac{1}{\pi}\left[\frac{\pi}{2}-\arctan(1-y)^3\right]\ (y\in\mathbf{R}),\end{aligned}$$
由此可得 $Y=1-\sqrt[3]{X}$ 的概率密度为 $f_Y(y)=F_Y'(y)=\frac{3(1-y)^2}{\pi[1+(1-y)^6]}\ (y\in\mathbf{R})$.

例 3.2.23 设随机变量 X 的概率密度为
$$f_X(x)=\begin{cases}1+x, & -1\leqslant x<0,\\1-x, & 0\leqslant x\leqslant 1,\\0, & \text{其他},\end{cases}$$
求随机变量 $Y=X^2+1$ 的分布函数.

解 随机变量 $Y=X^2+1$ 的分布函数为
$$F_Y(y)=P\{Y\leqslant y\}=P\{X^2+1\leqslant y\}(y\in\mathbf{R}).$$

当 $y<1$ 时，$F_Y(y)=0$.

当 $y\geqslant 1$ 时，$F_Y(y)=P\{X^2+1\leqslant y\}=P\{-\sqrt{y-1}\leqslant X\leqslant\sqrt{y-1}\}$.

如果 $1\leqslant y<2$，则 $0\leqslant\sqrt{y-1}<1$，因此有
$$\begin{aligned}F_Y(y)&=P\{-\sqrt{y-1}\leqslant X\leqslant\sqrt{y-1}\}=\int_{-\sqrt{y-1}}^{\sqrt{y-1}}f_X(x)\mathrm{d}x\\&=\int_{-\sqrt{y-1}}^{0}(1+x)\mathrm{d}x+\int_{0}^{\sqrt{y-1}}(1-x)\mathrm{d}x\\&=2\sqrt{y-1}-y+1.\end{aligned}$$

如果 $y\geqslant 2$，则 $\sqrt{y-1}\geqslant 1$，因此有
$$F_Y(y)=\int_{-\sqrt{y-1}}^{\sqrt{y-1}}f_X(x)\mathrm{d}x=\int_{-1}^{0}(1+x)\mathrm{d}x+\int_{0}^{1}(1-x)\mathrm{d}x=1.$$

综上所述，随机变量 $Y=X^2+1$ 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<1,\\ 2\sqrt{y-1}-y+1, & 1\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$$

基础习题精练

习 题

3.2.1 离散型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ 0.4, & -1\leqslant x<2,\\ 0.8, & 2\leqslant x<3,\\ 1, & x\geqslant 3,\end{cases}$$

则 X 的分布列为 ______，$P\{X\leqslant 3.2\}=$ ______；方程 $x^2+Xx+1=0$ 有实根的概率为 ______.

3.2.2 已知 $F(x)$ 是分布函数，下列函数：

①$aF(x)(a>0,a\neq 1)$；②$F(x)+F(-x)$；③$F(x)-F(-x)$；④$F(x)\cdot F(-x)$.

其中不能作为分布函数的个数是(　　).

(A)1　　(B)2　　(C)3　　(D)4

3.2.3 随机变量 X 的概率密度为

$$f(x)=\begin{cases}x, & 0\leqslant x<1,\\ 2-x, & 1\leqslant x<2,\\ 0, & \text{其他}.\end{cases}$$

求 X 的分布函数 $F(x)$.

3.2.4 设连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{A}{x^2}, & x>100,\\ 0, & x\leqslant 100,\end{cases}$$

则 $A=$ ______，$P\{100<X\leqslant 150\}=$ ______，$P\{X\geqslant 1\,000\}=$ ______，$P\{X=1\,000\}=$ ______，X 的分布函数 $F(x)=$ ______.

3.2.5 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-a,\\ A+B\arcsin\dfrac{x}{a}, & -a\leqslant x<a,\\ 1, & x\geqslant a,\end{cases}$$

则 $A=$ ______，$B=$ ______，X 的概率密度 $f(x)=$ ______，$P\left\{-\dfrac{a}{2}<X<\dfrac{a}{2}\right\}=$ ______.

3.2.6 设随机变量 X 服从区间$(2,5)$上的均匀分布，求对 X 进行 3 次独立观测中，至少有 2 次的观测值大于 3 的概率.

3.2.7 设随机变量 X 服从正态分布 $N(\mu_1,\sigma_1^2)$，随机变量 Y 服从正态分布 $N(\mu_2,\sigma_2^2)$，$\sigma_1,\sigma_2>0$，且 $P\{|X-\mu_1|<1\}>P\{|Y-\mu_2|<1\}$，则必有(　　).

(A)$\sigma_1<\sigma_2$　　(B)$\sigma_1>\sigma_2$　　(C)$\mu_1<\mu_2$　　(D)$\mu_1>\mu_2$

3.2.8 设随机变量 X 的分布函数 $F(x)=\begin{cases}0, & x<0,\\ \dfrac{1}{2}, & 0\leqslant x<1,\\ 1-\mathrm{e}^{-x}, & x\geqslant 1,\end{cases}$ 则 $P\{0\leqslant X\leqslant 1\}=$________，$P\{0<X<1\}=$________.

3.2.9 设随机变量 X 和 Y 同分布，X 的概率密度为

$$f(x)=\begin{cases}\dfrac{3}{8}x^2, & 0<x<2,\\ 0, & \text{其他}.\end{cases}$$

已知事件 $A=\{X>a\}$和 $B=\{Y>a\}$独立，且 $P(A\cup B)=\dfrac{3}{4}$，求常数 a.

3.2.10 设连续型随机变量 X 的概率密度 $f(x)$是一个偶函数，$F(x)$为 X 的分布函数，求证：对任意实数$a>0$，有

(1)$F(-a)=1-F(a)=0.5-\int_0^a f(x)\mathrm{d}x$；

(2)$P\{|X|<a\}=2F(a)-1$；

(3)$P\{|X|>a\}=2[1-F(a)]$.

3.2.11 已知 X 为随机变量，$Y=X^2+X+1$，已知 X 的概率分布为 $P\{X=-1\}=P\{X=0\}=P\{X=1\}=\dfrac{1}{3}$，试求 Y 的分布函数 $F_Y(y)$.

3.2.12 设随机变量 X 的分布函数 $F_X(x)$是严格增函数，其反函数 $F_X^{-1}(y)$存在，$Y=F_X(X)$. 证明：Y 服从区间$(0,1)$上的均匀分布.

3.2.13 设随机变量 $X\sim U(-1,2)$，求 $Y=|X|$的概率密度.

3.2.14 设随机变量 X 具有概率密度 $f_X(x)$，$-\infty<x<+\infty$，求 $Y=X^2$ 的概率密度.

3.2.15 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{9}x^2, & 0<x<3,\\ 0, & \text{其他}.\end{cases}$$

令随机变量 $Y=\begin{cases}2, & X\leqslant 1,\\ X, & 1<X<2,\\ 1, & X\geqslant 2.\end{cases}$

(1)求 Y 的分布函数；

(2)求概率 $P\{X\leqslant Y\}$.

解答

3.2.1 $X\sim\begin{pmatrix}-1 & 2 & 3\\ 0.4 & 0.4 & 0.2\end{pmatrix}$；1；0.6 **解** 分布函数的分段点即为离散型随机变量的正概率点，正概率点的概率即为对应点跃度，即$X=-1,2,3$，分布列为

$$X\sim\begin{pmatrix}-1 & 2 & 3\\ 0.4 & 0.4 & 0.2\end{pmatrix},\quad P\{X\leqslant 3.2\}=F(3.2)=1.$$

方程$x^2+Xx+1=0$有实根，即$X^2-4\geqslant 0$，从而有

$$P\{X^2-4\geqslant 0\}=P\{X\leqslant -2\}+P\{X\geqslant 2\}=F(-2)+1-F(2-0)=1-0.4=0.6.$$

3.2.2 (D) **解** 应用分布函数的充要条件求解. 由于

$$aF(+\infty)=a\neq 1,$$
$$F(-\infty)+F(+\infty)=1\neq 0,$$
$$F(-\infty)-F(+\infty)=-1\neq 0,$$
$$F(+\infty)\cdot F(-\infty)=0\neq 1,$$

故①，②，③，④都不能作为分布函数，故答案选(D).

3.2.3 **解** 由题设知，当$0\leqslant x<1$时，

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{0}^{x}t\mathrm{d}t=\frac{1}{2}t^2\Big|_0^x=\frac{x^2}{2};$$

当$1\leqslant x<2$时，

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{1}^{x}(2-t)\mathrm{d}t+\int_{0}^{1}t\mathrm{d}t$$
$$=\left(2t-\frac{t^2}{2}\right)\Big|_1^x+\frac{t^2}{2}\Big|_0^1=2x-\frac{x^2}{2}-1;$$

当$x<0$时，$F(x)=0$；

当$x\geqslant 2$时，

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{0}^{1}t\mathrm{d}t+\int_{1}^{2}(2-t)\mathrm{d}t$$
$$=\frac{t^2}{2}\Big|_0^1+\left(2t-\frac{t^2}{2}\right)\Big|_1^2=1.$$

所以

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{2}, & 0\leqslant x<1,\\ 2x-\dfrac{x^2}{2}-1, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

3.2.4 100；$\frac{1}{3}$；$\frac{1}{10}$；0；$\begin{cases}0, & x<100,\\ 1-\dfrac{100}{x}, & x\geqslant 100\end{cases}$ **解** 根据连续型随机变量概率密度的性质，有

$$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{100}^{+\infty}\frac{A}{x^2}\mathrm{d}x=-\frac{A}{x}\Big|_{100}^{+\infty}=\frac{A}{100}=1,\quad A=100,$$

于是

$$P\{100<X\leqslant 150\}=\int_{100}^{150}\frac{100}{x^2}\mathrm{d}x=-\frac{100}{x}\bigg|_{100}^{150}=\frac{1}{3},$$

$$P\{X\geqslant 1\,000\}=\int_{1\,000}^{+\infty}\frac{100}{x^2}\mathrm{d}x=-\frac{100}{x}\bigg|_{1\,000}^{+\infty}=\frac{1}{10},$$

$$P\{X=1\,000\}=0,$$

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\begin{cases}0, & x<100,\\ \displaystyle\int_{100}^{x}\frac{100}{t^2}\mathrm{d}t, & x\geqslant 100\end{cases}=\begin{cases}0, & x<100,\\ 1-\dfrac{100}{x}, & x\geqslant 100.\end{cases}$$

3.2.5 $\frac{1}{2}$；$\frac{1}{\pi}$；$\begin{cases}\dfrac{1}{\pi\sqrt{a^2-x^2}}, & -a<x<a,\\ 0, & 其他；\end{cases}$ $\frac{1}{3}$ **解** 根据连续型随机变量分布函数的连续性，有

$$F(-a)=F(-a-0),\quad F(a)=F(a-0),$$

$$\begin{cases}A+B\cdot\left(-\dfrac{\pi}{2}\right)=0,\\ A+B\cdot\dfrac{\pi}{2}=1,\end{cases}$$

解得 $A=\frac{1}{2}$，$B=\frac{1}{\pi}$.

于是

$$f(x)=\begin{cases}F'(x), & -a<x<a,\\ 0, & 其他\end{cases}=\begin{cases}\dfrac{1}{\pi\sqrt{a^2-x^2}}, & -a<x<a,\\ 0, & 其他.\end{cases}$$

$$P\left\{-\frac{a}{2}<X<\frac{a}{2}\right\}=\int_{-\frac{a}{2}}^{\frac{a}{2}}f(x)\mathrm{d}x=\int_{-\frac{a}{2}}^{\frac{a}{2}}\frac{1}{\pi\sqrt{a^2-x^2}}\mathrm{d}x=\frac{1}{3},$$

或

$$P\left\{-\frac{a}{2}<X<\frac{a}{2}\right\}=F\left(\frac{a}{2}\right)-F\left(-\frac{a}{2}\right)=\frac{1}{\pi}\left[\arcsin\frac{1}{2}-\arcsin\left(-\frac{1}{2}\right)\right]=\frac{1}{3}.$$

3.2.6 **解** 在一次观测中，观测值大于3的概率为

$$p=P\{X>3\}=\int_{3}^{5}\frac{1}{3}\mathrm{d}x=\frac{2}{3}.$$

记 $Y=\{3$ 次独立观测中 X 的观测值大于3的次数$\}$，由 $Y\sim B\left(3,\frac{2}{3}\right)$ 得

$$P\{Y\geqslant 2\}=1-P\{Y=0\}-P\{Y=1\}=1-\left(\frac{1}{3}\right)^3-3\times\frac{2}{3}\times\left(\frac{1}{3}\right)^2=\frac{20}{27}.$$

3.2.7 (A) **解** 所有有关正态分布不等式的判别都应先标准化，即由

$$P\{|X-\mu_1|<1\}>P\{|Y-\mu_2|<1\},$$

有

$$P\left\{\frac{|X-\mu_1|}{\sigma_1}<\frac{1}{\sigma_1}\right\}>P\left\{\frac{|Y-\mu_2|}{\sigma_2}<\frac{1}{\sigma_2}\right\},$$

从而有

$$\Phi\left(\frac{1}{\sigma_1}\right)>\Phi\left(\frac{1}{\sigma_2}\right).$$

又由函数 $\Phi(x)$ 的单调性，有 $\frac{1}{\sigma_1}>\frac{1}{\sigma_2}>0$，即 $\sigma_1<\sigma_2$，故选择(A).

3.2.8 $1-\mathrm{e}^{-1}$；0 **解**
$$P\{0\leqslant X\leqslant 1\}=P\{X\leqslant 1\}-P\{X<0\}=F(1)-F(0-0)$$
$$=1-\mathrm{e}^{-1}-0=1-\mathrm{e}^{-1};$$
$$P\{0<X<1\}=P\{X<1\}-P\{X\leqslant 0\}=F(1-0)-F(0)=\frac{1}{2}-\frac{1}{2}=0.$$

3.2.9 **解** 由 X 和 Y 同分布可得 $P(A)=P(B)$，从而
$$\frac{3}{4}=P(A\cup B)=P(A)+P(B)-P(AB)=2P(A)-[P(A)]^2,$$
由此解得 $P(A)=\frac{1}{2}$. 又
$$\frac{1}{2}=P(A)=P\{X>a\}=\int_a^2\frac{3}{8}x^2\mathrm{d}x=\frac{1}{8}x^3\bigg|_a^2=1-\frac{1}{8}a^3,$$
解得 $a=\sqrt[3]{4}$.

【注】 随机变量 X 与 Y 同分布，并不意味着 $X=Y$，反之成立，即 $X=Y$，则 X 与 Y 同分布.

3.2.10 **证明** 因为 $f(x)$ 是一个偶函数，所以 $f(-x)=f(x)$，且 $1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=2\int_0^{+\infty}f(x)\mathrm{d}x$，可得
$$\int_0^{+\infty}f(x)\mathrm{d}x=0.5.$$
又 $F(-a)=\int_{-\infty}^{-a}f(x)\mathrm{d}x$，令 $x=-t$，则 $F(-a)=\int_{+\infty}^{a}f(t)(-\mathrm{d}t)=\int_a^{+\infty}f(t)\mathrm{d}t=1-F(a)$.

(1)
$$F(-a)=\int_a^{+\infty}f(t)\mathrm{d}t=\int_0^{+\infty}f(t)\mathrm{d}t-\int_0^a f(t)\mathrm{d}t$$
$$=0.5-\int_0^a f(x)\mathrm{d}x.$$

(2)
$$P\{|X|<a\}=P\{-a<X<a\}=F(a-0)-F(-a)$$
$$=F(a)-[1-F(a)]$$
$$=2F(a)-1.$$

(3)
$$P\{|X|>a\}=P\{X<-a\}+P\{X>a\}=F(-a-0)+1-F(a)$$
$$=1-F(a)+1-F(a)$$
$$=2[1-F(a)].$$

3.2.11 **解** 由题设知 Y 为离散型随机变量，当 $X=-1$ 时，Y 取 1；当 $X=0$ 时，Y 取 1；当 $X=1$ 时，Y 取 3. 则 Y 的概率分布为

$$P\{Y=1\}=P\{X=-1\}+P\{X=0\}=\frac{2}{3},\quad P\{Y=3\}=P\{X=1\}=\frac{1}{3}.$$

故 Y 的分布函数为
$$F_Y(y)=P\{Y\leqslant y\}=\begin{cases}0, & y<1,\\ \frac{2}{3}, & 1\leqslant y<3,\\ 1, & y\geqslant 3.\end{cases}$$

3.2.12 **证明** 为了证明这个结论，首先要看到 $Y=F_X(X)$ 是在区间 $(0,1)$ 上取值的随机变量，所以：

当 $y<0$ 时，$F_Y(y)=0$；

当 $y\geqslant 1$ 时，$F_Y(y)=1$；

当 $0\leqslant y<1$ 时，

$$F_Y(y)=P\{Y\leqslant y\}=P\{F_X(X)\leqslant y\}=P\{X\leqslant F_X^{-1}(y)\}=F_X[F_X^{-1}(y)]=y.$$

综上所述，$Y=F_X(X)$ 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0,\\ y, & 0\leqslant y<1,\\ 1, & y\geqslant 1,\end{cases}$$

这就是在区间 $(0,1)$ 上的均匀分布函数，所以 $Y\sim U(0,1)$.

3.2.13 解 X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{3}, & -1<x<2,\\ 0, & \text{其他}.\end{cases}$$

记 X 的分布函数为 $F_X(x)$，先求 Y 的分布函数 $F_Y(y)$.

当 $y<0$ 时，$F_Y(y)=P\{Y\leqslant y\}=0$；

当 $y\geqslant 0$ 时，$F_Y(y)=P\{|X|\leqslant y\}=P\{-y\leqslant X\leqslant y\}=F_X(y)-F_X(-y)$.

将 $F_Y(y)$ 关于 y 求导可得 Y 的概率密度为

$$f_Y(y)=\begin{cases}f_X(y)+f_X(-y), & y>0,\\ 0, & \text{其他}.\end{cases}$$

当 $0\leqslant y<1$ 时，$-1<-y\leqslant 0$，因而 $f_X(y)=\dfrac{1}{3}$，$f_X(-y)=\dfrac{1}{3}$，此时

$$f_Y(y)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3};$$

当 $1\leqslant y<2$ 时，$-2<-y\leqslant -1$，因而 $f_X(y)=\dfrac{1}{3}$，$f_X(-y)=0$，此时

$$f_Y(y)=\frac{1}{3};$$

当 $y\geqslant 2$ 时，$f_X(y)=0$，$f_X(-y)=0$，因而 $f_Y(y)=0$. 故 Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{2}{3}, & 0\leqslant y<1,\\ \dfrac{1}{3}, & 1\leqslant y<2,\\ 0, & \text{其他}.\end{cases}$$

3.2.14 解 分别记 X,Y 的分布函数为 $F_X(x)$，$F_Y(y)$. 先求 Y 的分布函数 $F_Y(y)$. 由于 $Y=X^2\geqslant 0$，故当 $y<0$ 时，$F_Y(y)=0$；当 $y\geqslant 0$ 时，

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{X^2\leqslant y\}=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}\\&=F_X(\sqrt{y})-F_X(-\sqrt{y}),\end{aligned}$$

对 $F_Y(y)$ 关于 y 求导，得 Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})], & y>0,\\ 0, & \text{其他}.\end{cases}$$

3.2.15 解 (1)由题设知，$P\{1\leqslant Y\leqslant 2\}=1$. 记 Y 的分布函数为 $F_Y(y)$，则

当 $y<1$ 时，$F_Y(y)=0$；

当 $y\geqslant 2$ 时，$F_Y(y)=1$；

当 $1\leqslant y<2$ 时，

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{Y=1\}+P\{1<Y\leqslant y\}\\&=P\{X\geqslant 2\}+P\{1<X\leqslant y\}=\int_2^3\frac{x^2}{9}\mathrm{d}x+\int_1^y\frac{x^2}{9}\mathrm{d}x=\frac{y^3+18}{27}.\end{aligned}$$

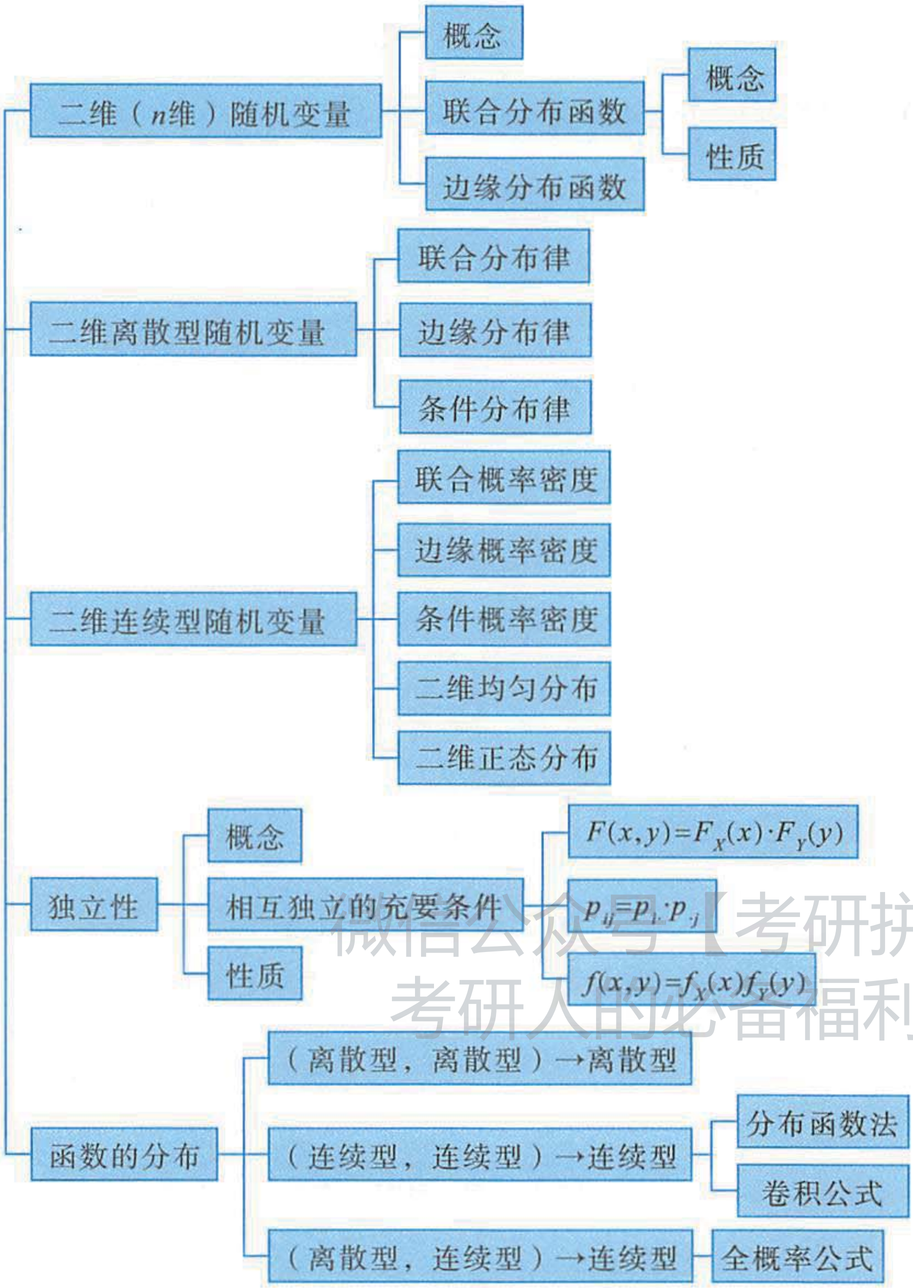
所以 Y 的分布函数为

$$F_Y(y)=\begin{cases}0, & y<1,\\ \dfrac{y^3+18}{27}, & 1\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$$

(2)$P\{X\leqslant Y\}=P\{X<2\}=\int_0^2\frac{x^2}{9}\mathrm{d}x=\frac{8}{27}$.

第3讲 多维随机变量及其分布

基础知识结构



基础内容精讲

一、二维(n 维)随机变量及其分布函数

1. 多维随机变量的概念

如果 $X_1, X_2, \cdots, X_n$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则称 $(X_1, X_2, \cdots, X_n)$ 为 n 维随机变量或 n 维随机向量，$X_i(i=1,2,\cdots,n)$ 称为第 i 个分量.

当 $n=2$ 时，记 (X,Y) 为**二维随机变量**(或**二维随机向量**).

2. 多维随机变量的分布函数的概念和性质

(1)概念.

对任意的 n 个实数 $x_1, x_2, \cdots, x_n$，称 n 元函数

$$F(x_1, x_2, \cdots, x_n)=P\{X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n\}$$

为 n 维随机变量 $(X_1, X_2, \cdots, X_n)$ 的**联合分布函数**.

当 $n=2$ 时，则对任意的实数 x, y，称二元函数

$$F(x,y)=P\{X \leqslant x, Y \leqslant y\}$$

为二维随机变量 (X,Y) 的**联合分布函数**，简称**分布函数**，记为 $(X,Y) \sim F(x,y)$.

【注】 注意 $F(x,y)$ 是事件 $A=\{X \leqslant x\}$ 与 $B=\{Y \leqslant y\}$ 同时发生的概率，因此 $0 \leqslant F(x,y) \leqslant 1$. 更为重要的是，可以利用概率及其有关的性质、公式求分布函数及讨论分布函数的性质，等等.

(2)性质.

①**单调性** $F(x,y)$ 是 x, y 的单调不减函数：

对任意固定的 y，当 $x_1<x_2$ 时，$F(x_1,y) \leqslant F(x_2,y)$；

对任意固定的 x，当 $y_1<y_2$ 时，$F(x,y_1) \leqslant F(x,y_2)$.

②**右连续性** $F(x,y)$ 是 x, y 的右连续函数：

$$\lim_{x \to x_0^+} F(x,y)=F(x_0+0,y)=F(x_0,y);$$

$$\lim_{y \to y_0^+} F(x,y)=F(x,y_0+0)=F(x,y_0).$$

③**有界性** $F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0, F(+\infty,+\infty)=1$.

④**非负性** 对任意 $x_1<x_2, y_1<y_2$ 有

$$P\{x_1<X \leqslant x_2, y_1<Y \leqslant y_2\}=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1) \geqslant 0.$$

3. 边缘分布函数

设二维随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，随机变量 X 与 Y 的分布函数 $F_X(x)$ 与 $F_Y(y)$ 分别称为 (X,Y) 关于 X 和关于 Y 的**边缘分布函数**. 由概率性质得

$$\begin{aligned} F_X(x) &= P\{X \leqslant x\}=P\{X \leqslant x, Y<+\infty\} \\ &= \lim_{y \to +\infty} P\{X \leqslant x, Y \leqslant y\} \\ &= \lim_{y \to +\infty} F(x,y)=F(x,+\infty). \end{aligned}$$

同理，有 $F_Y(y)=F(+\infty,y)$.

二、常见的两类二维随机变量——离散型随机变量与连续型随机变量

（一）二维离散型随机变量的概率分布、边缘分布和条件分布

1. 概率分布

(1)如果二维随机变量 (X,Y) 只能取有限对值或可列对值 $(x_1,y_1),(x_1,y_2),\cdots,(x_n,y_n),\cdots$，则称 (X,Y) 为二维离散型随机变量.

称
$$p_{ij}=P\{X=x_i,Y=y_j\},i,j=1,2,\cdots$$
为 (X,Y) 的**分布律**或称为随机变量 X 和 Y 的**联合分布律**，记为 $(X,Y)\sim p_{ij}$. 联合分布律常用表格形式表示.

Y X	y_1	$\cdots$	y_j	$\cdots$	$P\{X=x_i\}$
x_1	p_{11}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	p_{i1}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$P\{Y=y_j\}$	$p_{\cdot 1}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

(2)数列 $\{p_{ij}\},i,j=1,2,\cdots$ 是某一二维离散型随机变量的概率分布的充分必要条件为
$$p_{ij}\geqslant 0,\quad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}p_{ij}=1.$$

2. 联合分布函数、边缘分布、条件分布

(1)联合分布函数.

设 (X,Y) 的概率分布为 $p_{ij},i,j=1,2,\cdots$，则 (X,Y) 的**联合分布函数**为
$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}p_{ij}.$$

【注】 它是以 (x,y) 为顶点的左下角平面上 (X,Y) 取所有可能值的概率的和.

设 G 是平面上的某个区域，则
$$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)\in G}p_{ij}.$$

【注】 (X,Y) 落入 G 的概率等于 (X,Y) 在 G 内取所有可能值的概率的和. 这是计算概率、求随机变量函数分布的重要公式.

(2)边缘分布.

X,Y 的**边缘分布**分别为
$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}P\{X=x_i,Y=y_j\}=\sum_{j=1}^{\infty}p_{ij}\ (i=1,2,\cdots);$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{\infty}P\{X=x_i,Y=y_j\}=\sum_{i=1}^{\infty}p_{ij}\,(j=1,2,\cdots).$$

(3)条件分布.

如果$(X,Y)\sim p_{ij}(i,j=1,2,\cdots)$,对固定的$j$,如果$p_{\cdot j}=P\{Y=y_j\}>0$,则称

$$P\{X=x_i\mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}}\ (i=1,2,\cdots)$$

为X在“$Y=y_j$”条件下的条件分布.

同理,可定义Y在“$X=x_i$”条件下的条件分布

$$P\{Y=y_j\mid X=x_i\}=\frac{p_{ij}}{p_{i\cdot}}\ (j=1,2,\cdots).$$

(二)二维连续型随机变量的概率密度、边缘概率密度和条件概率密度

1. 概率密度

(1)概念.

如果二维随机变量(X,Y)的联合分布函数$F(x,y)$可以表示为

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v,(x,y)\in\mathbf{R}^2,$$

其中$f(x,y)$是非负可积函数,则称(X,Y)为**二维连续型随机变量**,称$f(x,y)$为(X,Y)的**概率密度**,记为$(X,Y)\sim f(x,y)$.

(2)二元函数$f(x,y)$是概率密度的充分必要条件为

$$f(x,y)\geqslant 0,\quad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=1.$$

【注】 改变$f(x,y)$的有限个点值(仍取非负值),$f(x,y)$仍然是概率密度.

2. 联合分布函数与概率密度、边缘概率密度、条件概率密度

(1)联合分布函数与概率密度.

设(X,Y)的联合分布函数为$F(x,y)$,概率密度为$f(x,y)$,则

①$F(x,y)$为(x,y)的二元连续函数,且

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v;$$

②设G为平面上某个区域,则

$$P\{(X,Y)\in G\}=\iint_{G}f(x,y)\mathrm{d}x\mathrm{d}y;$$

【注】 这是计算概率、求随机变量分布函数的依据.

③若$f(x,y)$在点(x,y)处连续,则$\frac{\partial^2F(x,y)}{\partial x\partial y}=f(x,y)$;

④若$F(x,y)$连续且可导,则(X,Y)是连续型随机变量,则$\frac{\partial^2F(x,y)}{\partial x\partial y}$是它的概率密度.

(2)边缘概率密度.

设$(X,Y)\sim f(x,y)$,则X的边缘分布函数为

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty}f(u,v)\mathrm{d}v\right]\mathrm{d}u,$$

所以 X 是连续型随机变量，其概率密度

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,$$

称 $f_X(x)$ 为 (X,Y) 关于 X 的**边缘概率密度**.

同理，Y 也是连续型随机变量，其概率密度为 $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$.

(3)条件概率密度.

设 $(X,Y)\sim f(x,y)$，边缘概率密度 $f_X(x)>0$，则称

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$

为 Y 在“$X=x$”条件下的**条件概率密度**.

同理可定义 X 在“$Y=y$”条件下的条件概率密度

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}(f_Y(y)>0).$$

由上讨论知，若 $f_X(x)>0,f_Y(y)>0$，则有概率密度乘法公式

$$f(x,y)=f_X(x)f_{Y|X}(y|x)=f_Y(y)f_{X|Y}(x|y).$$

【注】 称 $$F_{Y|X}(y|x)=\int_{-\infty}^{y}f_{Y|X}(y|x)\mathrm{d}y=\int_{-\infty}^{y}\frac{f(x,y)}{f_X(x)}\mathrm{d}y$$

为 Y 在“$X=x$”条件下的**条件分布函数**.

同理可定义 X 在“$Y=y$”条件下的条件分布函数

$$F_{X|Y}(x|y)=\int_{-\infty}^{x}f_{X|Y}(x|y)\mathrm{d}x=\int_{-\infty}^{x}\frac{f(x,y)}{f_Y(y)}\mathrm{d}x.$$

3. 常见的二维分布

(1)二维均匀分布.

称 (X,Y) 在平面有界区域 D 上服从**均匀分布**，如果 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$$

其中 S_D 为区域 D 的面积.

(2)二维正态分布.

如果 (X,Y) 的概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\},$$

其中 $\mu_1\in\mathbf{R},\mu_2\in\mathbf{R},\sigma_1>0,\sigma_2>0,-1<\rho<1$，则称 (X,Y) 服从参数为 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 的**二维正态分布**，记为 $(X,Y)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$. 此时有

①$X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$，$\rho$ 为 X 与 Y 的相关系数，即

$$\rho=\frac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}=\frac{\mathrm{Cov}(X,Y)}{\sigma_1\sigma_2}.$$

②X,Y 的条件分布都是正态分布.

③$aX+bY(a\neq0$ 或 $b\neq0)$服从正态分布.

④X 与 Y 相互独立的充分必要条件是 X 与 Y 不相关，即 $\rho=0$.

三、随机变量的相互独立性

1. 概念

(1)设二维随机变量(X,Y)的联合分布函数为 $F(x,y)$，边缘分布函数分别为 $F_X(x),F_Y(y)$，如果对任意的实数 x,y 都有

$$F(x,y)=F_X(x)\cdot F_Y(y)(x\in\mathbf{R},y\in\mathbf{R})(\text{即事件}\{X\leqslant x\}\text{与}\{Y\leqslant y\}\text{相互独立}),$$

则称 X 与 Y **相互独立**，否则称 X 与 Y **不相互独立**.

(2)如果 n 维随机变量$(X_1,X_2,\cdots,X_n)$的联合分布函数等于边缘分布函数的乘积，即

$$F(x_1,x_2,\cdots,x_n)=F_1(x_1)\cdot F_2(x_2)\cdot\cdots\cdot F_n(x_n),$$

其中 $F_i(x_i)(i=1,2,\cdots,n)$为 X_i 的边缘分布函数，x_i 为任意实数，则称 $X_1,X_2,\cdots,X_n$ **相互独立**.

(3)随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$相互独立，如果对任意 $k(k\geqslant2)$，$X_1,X_2,\cdots,X_k$ 相互独立.

(4)两个多维随机变量$(X_1,X_2,\cdots,X_n)$与$(Y_1,Y_2,\cdots,Y_m)$相互独立，如果对任意实数 $x_i(i=1,2,\cdots,n)$与$y_j(j=1,2,\cdots,m)$有

$$\begin{aligned}&P\{X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n;Y_1\leqslant y_1,Y_2\leqslant y_2,\cdots,Y_m\leqslant y_m\}\\=&P\{X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n\}\cdot P\{Y_1\leqslant y_1,Y_2\leqslant y_2,\cdots,Y_m\leqslant y_m\},\end{aligned}$$

即联合分布函数等于各自的分布函数相乘：

$$F(x_1,x_2,\cdots,x_n,y_1,y_2,\cdots,y_m)=F_1(x_1,x_2,\cdots,x_n)\cdot F_2(y_1,y_2,\cdots,y_m).$$

2. 相互独立的充要条件

(1) n 个随机变量 $X_1,X_2,\cdots,X_n$ 相互独立

$\Leftrightarrow$对任意的 n 个实数 $x_i(i=1,2,\cdots,n)$，n 个事件$\{X_1\leqslant x_1\},\{X_2\leqslant x_2\},\cdots,\{X_n\leqslant x_n\}$相互独立.

(2)①设(X,Y)为二维离散型随机变量，则 X 与 Y 相互独立

$\Leftrightarrow$联合分布等于边缘分布相乘，即

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}(i,j=1,2,\cdots).$$

②n 个离散型随机变量 $X_1,X_2,\cdots,X_n$ 相互独立

$\Leftrightarrow$对任意的 $x_i\in D_i=\{X_i$ 的一切可能值$\}(i=1,2,\cdots,n)$，有

$$P\{X_1=x_1,\cdots,X_n=x_n\}=\prod_{i=1}^{n}P\{X_i=x_i\}.$$

(3)①(X,Y)为二维连续型随机变量，则 X 与 Y 相互独立

$\Leftrightarrow$概率密度等于边缘密度相乘，即

$$f(x,y)=f_X(x)\cdot f_Y(y).$$

②设$(X_1,X_2,\cdots,X_n)$为 n 维连续型随机变量，则 $X_1,X_2,\cdots,X_n$ 相互独立

$\Leftrightarrow$概率密度等于边缘密度相乘，即

$$f(x_1,x_2,\cdots,x_n)=f_1(x_1)\cdot f_2(x_2)\cdot\cdots\cdot f_n(x_n),$$

其中 $f_i(x)$ 为 X_i 的边缘概率密度.

3. 相互独立的性质

(1)设 $X_1,X_2,\cdots,X_n$ 相互独立，则其中任意 $k(2\leqslant k\leqslant n)$ 个随机变量也相互独立.

(2)①设 (X,Y) 为二维离散型随机变量，X 与 Y 独立，则条件分布等于边缘分布：

$$P\{X=x_i|Y=y_j\}=P\{X=x_i\}(P\{Y=y_j\}>0),$$

$$P\{Y=y_j|X=x_i\}=P\{Y=y_j\}(P\{X=x_i\}>0).$$

②设 (X,Y) 为二维连续型随机变量，X 与 Y 独立，则条件概率密度等于边缘概率密度：

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=f_X(x)(f_Y(y)>0),$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=f_Y(y)(f_X(x)>0).$$

(3)若 $X_1,X_2,\cdots,X_n$ 相互独立，$g_1(x),g_2(x),\cdots,g_n(x)$ 为一元连续函数，则 $g_1(X_1),g_2(X_2),\cdots,g_n(X_n)$ 相互独立.

一般地，若 $X_{11},\cdots,X_{1t_1},X_{21},\cdots,X_{2t_2},\cdots,X_{n1},\cdots,X_{nt_n}$ 相互独立，g_i 是 $t_i(i=1,2,\cdots,n)$ 元连续函数，则 $g_1(X_{11},\cdots,X_{1t_1}),g_2(X_{21},\cdots,X_{2t_2}),\cdots,g_n(X_{n1},\cdots,X_{nt_n})$ 也相互独立.

四、多维随机变量函数的分布

1. 概念

设 X,Y 为随机变量，$g(x,y)$ 是二元函数，则以随机变量 X,Y 作为变量的函数 $U=g(X,Y)$ 也是随机变量，称之为随机变量 X,Y 的函数. 例如：$U=X+Y,U=\begin{cases}1, & X<Y,\\ 0, & X\geqslant Y\end{cases}$ 等等.

问题：已知 (X,Y) 的联合分布，求 $U=g(X,Y)$ 的分布；又 $V=h(X,Y)$，求 (U,V) 的联合分布.

2. 求法

已知 (X,Y) 的联合分布，求函数 $Z=g(X,Y)$ 的分布. 首先要确定 X,Y 的类型，而后采用相应的公式计算.

(1)如果 (X,Y) 是二维离散型随机变量，则 $Z=g(X,Y)$ 也是离散型的，先确定 Z 的值，而后求其相应的概率，用一般解题模式(矩阵法)即可求得 Z 的分布，具体见下面的注例及例 3.3.10.

(2)如果 X,Y 其中一个是离散型的，另一个是非离散型的，我们总是将事件对离散型的一切可能进行全集分解，而后应用全概率公式求得 Z 的分布，具体见例 3.3.11.

(3)如果 (X,Y) 是二维连续型随机变量，即 $(X,Y)\sim f(x,y)$，则 $Z=g(X,Y)$ 的分布函数

$$F(z)=P\{g(X,Y)\leqslant z\}=\iint\limits_{g(x,y)\leqslant z}f(x,y)\mathrm{d}x\mathrm{d}y,$$

要求 $f(z)$，亦可直接用下面的“3. 相互独立随机变量函数的分布及卷积公式”中所讲公式，具体见例 3.3.12，例 3.3.13，例 3.3.14.

注例　已知(X,Y)的联合概率分布为

Y X	-1	1
-1	$\frac{1}{6}$	$\frac{1}{6}$
0	$\frac{2}{6}$	$\frac{1}{6}$
1	0	$\frac{1}{6}$

(1)求$Z=X-Y$的概率分布；

(2)设$U_1=XY,V_1=\frac{X}{Y}$，求(U_1,V_1)的概率分布；

(3)设$U_2=\max\{X,Y\},V_2=\min\{X,Y\}$，求$(U_2,V_2)$的概率分布，$U_2V_2$的概率分布.

解　由题设得

p_{ij}	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$
(X,Y)	$(-1,-1)$	$(-1,1)$	$(0,-1)$	$(0,1)$	$(1,-1)$	$(1,1)$
$Z=X-Y$	0	-2	1	-1	2	0
$U_1=XY$	1	-1	0	0	-1	1
$V_1=X/Y$	1	-1	0	0	-1	1
$U_2=\max\{X,Y\}$	-1	1	0	1	1	1
$V_2=\min\{X,Y\}$	-1	-1	-1	0	-1	1
U_2V_2	1	-1	0	0	-1	1

(1)由此即得$Z=X-Y$的概率分布为

$Z=X-Y$	-2	-1	0	1	2
$P\{Z=k\}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	0

(2)(U_1,V_1)的概率分布为

V_1 U_1	-1	0	1
-1	$\frac{1}{6}$	0	0
0	0	$\frac{3}{6}$	0
1	0	0	$\frac{2}{6}$

(3)(U_2,V_2)的概率分布为

U_2 \ V_2	-1	0	1
-1	$\frac{1}{6}$	0	0
0	$\frac{2}{6}$	0	0
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

U_2V_2 的概率分布为

U_2V_2	-1	0	1
$P\{U_2V_2=k\}$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

3. 相互独立随机变量函数的分布及卷积公式

(1)和的分布.

设$(X,Y)\sim f(x,y)$，则 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x=\int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y.$$

【注 1】 **证明** 按照定义，如图 3-3-1 所示，

$$F_Z(z)=P\{Z\leqslant z\}=P\{X+Y\leqslant z\}=\iint\limits_{D:x+y\leqslant z} f(x,y)\mathrm{d}\sigma=\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{z-y} f(x,y)\mathrm{d}x,$$

于是 $f_Z(z)=F'_Z(z)=\left\{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y} f(x,y)\mathrm{d}x\right]\mathrm{d}y\right\}'_z \xlongequal{(*)} \int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-y} f(x,y)\mathrm{d}x\right]'_z \mathrm{d}y$

$$=\int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y.$$

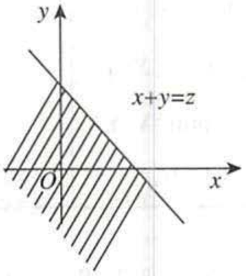


图 3-3-1

(*)处将先积分再求导的运算顺序交换，成了先求导再积分，可以证明(这里不证)，这是成立的.

同理，若 $F_Z(z)=\iint\limits_{D:x+y\leqslant z} f(x,y)\mathrm{d}\sigma=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{-\infty}^{z-x} f(x,y)\mathrm{d}y,$

于是 $f_Z(z)=F'_Z(z)=\left\{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-x} f(x,y)\mathrm{d}y\right]\mathrm{d}x\right\}'_z=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z-x} f(x,y)\mathrm{d}y\right]'_z \mathrm{d}x$

$$=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x.$$

【注 2】 当 X 与 Y 相互独立时，有卷积公式

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x=\int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)\mathrm{d}y.$$

(2)差的分布.

设$(X,Y)\sim f(x,y)$，则$Z=X-Y$的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,x-z)\mathrm{d}x=\int_{-\infty}^{+\infty}f(y+z,y)\mathrm{d}y\overset{\text{独立}}{=\!=\!=}\int_{-\infty}^{+\infty}f_X(y+z)f_Y(y)\mathrm{d}y.$$

【注1】 **证明** 按照定义，如图3-3-2所示，

$$F_Z(z)=P\{Z\leqslant z\}=P\{X-Y\leqslant z\}=\iint\limits_{D:x-y\leqslant z}f(x,y)\mathrm{d}\sigma=\int_{-\infty}^{+\infty}\mathrm{d}y\int_{-\infty}^{z+y}f(x,y)\mathrm{d}x,$$

于是

$$\begin{aligned}f_Z(z)=F'_Z(z)&=\left\{\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z+y}f(x,y)\mathrm{d}x\right]\mathrm{d}y\right\}'_z\\&=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{z+y}f(x,y)\mathrm{d}x\right]'_z\mathrm{d}y\\&=\int_{-\infty}^{+\infty}f(z+y,y)\mathrm{d}y.\end{aligned}$$

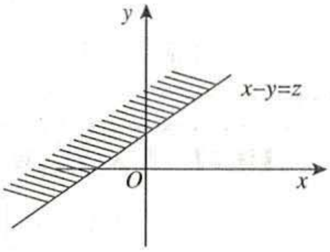


图 3-3-2

同理，若 $F_Z(z)=\iint\limits_{D:x-y\leqslant z}f(x,y)\mathrm{d}\sigma=\int_{-\infty}^{+\infty}\mathrm{d}x\int_{x-z}^{+\infty}f(x,y)\mathrm{d}y,$

于是

$$\begin{aligned}f_Z(z)=F'_Z(z)&=\left\{\int_{-\infty}^{+\infty}\left[\int_{x-z}^{+\infty}f(x,y)\mathrm{d}y\right]\mathrm{d}x\right\}'_z\\&=\int_{-\infty}^{+\infty}\left[\int_{x-z}^{+\infty}f(x,y)\mathrm{d}y\right]'_z\mathrm{d}x\\&=\int_{-\infty}^{+\infty}[-f(x,x-z)\cdot(-1)]\mathrm{d}x\\&=\int_{-\infty}^{+\infty}f(x,x-z)\mathrm{d}x.\end{aligned}$$

【注2】 $Z=Y-X$的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,x+z)\mathrm{d}x\overset{\text{独立}}{=\!=\!=}\int_{-\infty}^{+\infty}f_X(x)f_Y(x+z)\mathrm{d}x.$$

(3)积的分布.

设$(X,Y)\sim f(x,y)$，则$Z=XY$的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{1}{|y|}f\left(\frac{z}{y},y\right)\mathrm{d}y.$$

【注1】 **证明** 按照定义，如图3-3-3所示，

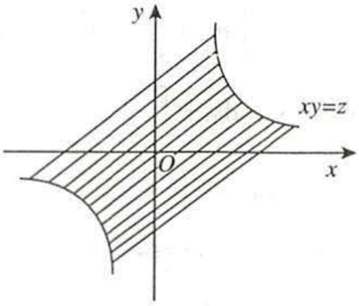


图 3-3-3

$$F_Z(z)=P\{Z\leqslant z\}=P\{XY\leqslant z\}=\iint\limits_{D:xy\leqslant z} f(x,y)\mathrm{d}\sigma=\int_{-\infty}^{0}\mathrm{d}x\int_{\frac{z}{x}}^{+\infty}f(x,y)\mathrm{d}y+\int_{0}^{+\infty}\mathrm{d}x\int_{-\infty}^{\frac{z}{x}}f(x,y)\mathrm{d}y,$$

于是 $f_Z(z)=F_Z'(z)=\int_{-\infty}^{0}\left[\int_{\frac{z}{x}}^{+\infty}f(x,y)\mathrm{d}y\right]_z'\mathrm{d}x+\int_{0}^{+\infty}\left[\int_{-\infty}^{\frac{z}{x}}f(x,y)\mathrm{d}y\right]_z'\mathrm{d}x$

$$=\int_{-\infty}^{0}\left[-f\left(x,\frac{z}{x}\right)\cdot\frac{1}{x}\right]\mathrm{d}x+\int_{0}^{+\infty}\left[f\left(x,\frac{z}{x}\right)\cdot\frac{1}{x}\right]\mathrm{d}x$$

$$=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x.$$

同理可证后一等号.

【注 2】 当 X 与 Y 独立时,有

$$f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f_X(x)f_Y\left(\frac{z}{x}\right)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{1}{|y|}f_X\left(\frac{z}{y}\right)f_Y(y)\mathrm{d}y.$$

(4)商的分布.

设$(X,Y)\sim f(x,y)$,则 $Z=\dfrac{X}{Y}$的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}|y|f(yz,y)\mathrm{d}y.$$

【注 1】 **证明** 按照定义,如图 3-3-4 所示,

$$F_Z(z)=P\{Z\leqslant z\}=P\left\{\frac{X}{Y}\leqslant z\right\}=\iint\limits_{D:\frac{x}{y}\leqslant z}f(x,y)\mathrm{d}\sigma$$

$$=\int_{-\infty}^{0}\mathrm{d}y\int_{zy}^{+\infty}f(x,y)\mathrm{d}x+\int_{0}^{+\infty}\mathrm{d}y\int_{-\infty}^{zy}f(x,y)\mathrm{d}x,$$

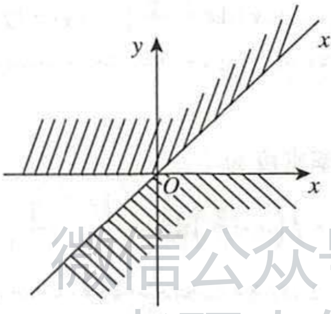


图 3-3-4

于是 $f_Z(z)=F_Z'(z)=\int_{-\infty}^{0}\left[\int_{zy}^{+\infty}f(x,y)\mathrm{d}x\right]_z'\mathrm{d}y+\int_{0}^{+\infty}\left[\int_{-\infty}^{zy}f(x,y)\mathrm{d}x\right]_z'\mathrm{d}y$

$$=\int_{-\infty}^{0}[-f(zy,y)\cdot y]\mathrm{d}y+\int_{0}^{+\infty}[f(zy,y)\cdot y]\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}|y|f(zy,y)\mathrm{d}y.$$

【注 2】 当 X 与 Y 独立时，有 $f_Z(z)=\int_{-\infty}^{+\infty}|y|f_X(yz)f_Y(y)\mathrm{d}y$.

(5)$\max\{X,Y\}$分布.

设$(X,Y)\sim F(x,y)$，则 $Z=\max\{X,Y\}$的分布函数为

$$F_{\max}(z)=P\{\max\{X,Y\}\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=F(z,z).$$

当 X 与 Y 独立时，

$$F_{\max}(z)=F_X(z)\cdot F_Y(z).$$

(6)$\min\{X,Y\}$分布.

设$(X,Y)\sim F(x,y)$，则 $Z=\min\{X,Y\}$的分布函数为

$$\begin{aligned}F_{\min}(z)&=P\{\min\{X,Y\}\leqslant z\}=P\{\{X\leqslant z\}\cup\{Y\leqslant z\}\}\\&=P\{X\leqslant z\}+P\{Y\leqslant z\}-P\{X\leqslant z,Y\leqslant z\}\\&=F_X(z)+F_Y(z)-F(z,z).\end{aligned}$$

当 X 与 Y 独立时，

$$\begin{aligned}F_{\min}(z)&=F_X(z)+F_Y(z)-F_X(z)F_Y(z)\\&=1-[1-F_X(z)][1-F_Y(z)].\end{aligned}$$

【注】 上述结果容易推广到 n 个相互独立的随机变量 $X_1,X_2,\cdots,X_n$ 的情况，即

$$F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z),$$

$$F_{\min}(z)=1-[1-F_{X_1}(z)][1-F_{X_2}(z)]\cdots[1-F_{X_n}(z)].$$

特别地，当 $X_i(i=1,2,\cdots,n)$相互独立且有相同的分布函数 $F(x)$与概率密度 $f(x)$时，

$$F_{\max}(x)=[F(x)]^n,$$

$$f_{\max}(x)=n[F(x)]^{n-1}f(x).$$

$$F_{\min}(x)=1-[1-F(x)]^n,$$

$$f_{\min}(x)=n[1-F(x)]^{n-1}f(x).$$

这些结果在数理统计部分是要用到的.

(7)常见分布的可加性.

有些相互独立且服从同类型分布的随机变量，其和的分布也是同类型的，它们是二项分布、泊松分布、正态分布与 χ^2 分布. 设随机变量 X 与 Y 相互独立，则：

若 $X\sim B(n,p),Y\sim B(m,p)$，则 $X+Y\sim B(n+m,p)$(注意 p 相同)；

若 $X\sim P(\lambda_1),Y\sim P(\lambda_2)$，则 $X+Y\sim P(\lambda_1+\lambda_2)$；

若 $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$，则 $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$；

若 $X\sim\chi^2(n),Y\sim\chi^2(m)$，则 $X+Y\sim\chi^2(n+m)$.

【注】 上述结果对 n 个相互独立的随机变量也成立.

基础例题精解

一、联合分布、边缘分布、条件分布和独立性的概念、性质及其相互转化

1. 联合分布、边缘分布、条件分布之间的相互转化

已知(X,Y)的联合分布（联合分布函数$F(x,y)$、联合分布p_{ij}、概率密度$f(x,y)$），可以确定边缘分布（边缘分布函数、边缘分布、边缘概率密度）与条件分布（条件分布、条件概率密度），其转换公式为

(1)已知$(X,Y)\sim F(x,y)$，则$F_X(x)=F(x,+\infty)$，$F_Y(y)=F(+\infty,y)$.

(2)已知$(X,Y)\sim p_{ij}=P\{X=x_i,Y=y_j\}$，则

$$F(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}p_{ij},$$

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij},$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{\infty}p_{ij}.$$

条件分布：

$$P\{X=x_i|Y=y_j\}=\frac{p_{ij}}{p_{\cdot j}}(p_{\cdot j}>0),$$

$$P\{Y=y_j|X=x_i\}=\frac{p_{ij}}{p_{i\cdot}}(p_{i\cdot}>0).$$

(3)已知$(X,Y)\sim f(x,y)$，则$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$,

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y,$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x.$$

若$f(x,y)$在点(x,y)处连续，则

$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y).$$

条件概率密度：

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}(f_Y(y)>0),$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}(f_X(x)>0).$$

反之，仅知边缘分布，不可确定联合分布.若再附加某些其他条件，如概率、独立性、某些数字特征等，则有可能确定联合分布.

2. X与Y相互独立的充要条件

(1)对任意实数x,y,

$$\text{事件}A=\{X\leqslant x\},B=\{Y\leqslant y\}\text{相互独立}\Leftrightarrow F(x,y)=F_X(x)F_Y(y).$$

(2)若(X,Y)为二维离散型随机变量，则

$$X\text{与}Y\text{独立}\Leftrightarrow P\{X=x_i,Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}$$

$$\Leftrightarrow 若 P\{Y=y_j\}>0, P\{X=x_i \mid Y=y_j\}=P\{X=x_i\}.$$

(3)若(X,Y)为二维连续型随机变量，则

$$X 与 Y 独立 \Leftrightarrow f(x,y)=f_X(x)f_Y(y) \Leftrightarrow f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=f_X(x)(f_Y(y)>0).$$

(4)(X,Y)服从二维正态分布，则 X 与 Y 独立$\Leftrightarrow X$ 与 Y 不相关.

3. X 与 Y 不独立的判断与证明

X 与 Y 不独立$\Leftrightarrow$存在 x_0, y_0 使 $A=\{X\leqslant x_0\}$与 $B=\{Y\leqslant y_0\}$不独立，即

$$F(x_0,y_0)\neq F_X(x_0)F_Y(y_0).$$

因此证明不独立的常用方法：求 x_0, y_0，使 $0<P\{X\leqslant x_0\}, P\{Y\leqslant y_0\}<1$，

$$\{X\leqslant x_0\}\subset\{Y\leqslant y_0\} 或 \{Y\leqslant y_0\}\subset\{X\leqslant x_0\} 或 \{X\leqslant x_0, Y\leqslant y_0\}\neq\varnothing.$$

(1)当(X,Y)为离散型时，

$$X 与 Y 不独立 \Leftrightarrow 存在 x_0, y_0，使 P\{X=x_0, Y=y_0\}\neq P\{X=x_0\}P\{Y=y_0\}.$$

【注】 二维离散型随机变量的独立性的讨论主要是在掌握两随机变量的联合分布和边缘分布的基础上进行的，主要验证其中只要有一项不成立，则独立性不成立，同时要注意两随机变量不相关未必独立，独立必不相关.

(2)二维连续型随机变量的独立性讨论主要看概率密度或分布函数是否满足

$$f(x,y)=f_X(x)f_Y(y) 或 F(x,y)=F_X(x)F_Y(y).$$

【注】 作为特殊类型，若(X,Y)服从二维正态分布 $N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$，则 X 与 Y 相互独立的充分必要条件是 $\rho=0$，但不具一般性，即一般情况下，不相关未必独立.

例 3.3.1 设 X 与 Y 相互独立，下表是随机变量 X 与 Y 的联合分布律以及两个边缘分布律中的部分数值，试将其余数值填入表中空白处.

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}$
x_1		$\frac{1}{8}$		
x_2	$\frac{1}{8}$			
$P\{Y=y_j\}$	$\frac{1}{6}$			1

解　由于 X 与 Y 相互独立，有

$$p_{ij}=P\{X=x_i, Y=y_j\}=P\{X=x_i\}P\{Y=y_j\}=p_{i\cdot}p_{\cdot j}(i=1,2;j=1,2,3).$$

由边缘分布的性质，有

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{3}p_{ij},\quad p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{2}p_{ij},$$

于是

$$p_{11}=p_{\cdot 1}-p_{21}=\frac{1}{6}-\frac{1}{8}=\frac{1}{24},$$

$$p_{1\cdot}=\frac{p_{11}}{p_{\cdot 1}}=\frac{1}{24}\div\frac{1}{6}=\frac{1}{4},\quad p_{2\cdot}=1-p_{1\cdot}=\frac{3}{4},$$

$$p_{\cdot 2}=\frac{p_{12}}{p_{1\cdot}}=\frac{1}{8}\div\frac{1}{4}=\frac{1}{2},\quad p_{\cdot 3}=1-p_{\cdot 1}-p_{\cdot 2}=\frac{1}{3},$$

$$p_{22}=p_{\cdot 2}p_{2\cdot}=\frac{1}{2}\times\frac{3}{4}=\frac{3}{8},\quad p_{13}=p_{\cdot 3}p_{1\cdot}=\frac{1}{3}\times\frac{1}{4}=\frac{1}{12},$$

$$p_{23}=p_{\cdot 3}p_{2\cdot}=\frac{1}{3}\times\frac{3}{4}=\frac{1}{4}.$$

因此所得数值填空如下表：

X \ Y	y_1	y_2	y_3	$P\{X=x_i\}$
x_1	$\mathbf{\frac{1}{24}}$	$\frac{1}{8}$	$\mathbf{\frac{1}{12}}$	$\mathbf{\frac{1}{4}}$
x_2	$\frac{1}{8}$	$\mathbf{\frac{3}{8}}$	$\mathbf{\frac{1}{4}}$	$\mathbf{\frac{3}{4}}$
$P\{Y=y_j\}$	$\frac{1}{6}$	$\mathbf{\frac{1}{2}}$	$\mathbf{\frac{1}{3}}$	1

例 3.3.2 二维随机变量(X,Y)的概率分布为

X \ Y	0	1
0	0.4	a
1	b	0.1

若随机事件$\{X=0\}$与$\{X+Y=1\}$相互独立，则(　　).

(A)$a=0.2,b=0.3$　　(B)$a=0.1,b=0.4$

(C)$a=0.3,b=0.2$　　(D)$a=0.4,b=0.1$

解 应选(D).

由联合分布、边缘分布及条件分布的性质，有

$$P\{X=0\}=p_{11}+p_{12}=0.4+a,$$

$$P\{X+Y=1\}=p_{12}+p_{21}=a+b,$$

$$p_{11}+p_{12}+p_{21}+p_{22}=a+b+0.5=1\Rightarrow a+b=0.5,$$

$$P\{X=0,X+Y=1\}=P\{X=0\}P\{Y=1|X=0\}=(0.4+a)\frac{a}{0.4+a}=a.$$

又$\{X=0\}$与$\{X+Y=1\}$相互独立，有

$$P\{X=0,X+Y=1\}=P\{X=0\}\cdot P\{X+Y=1\}$$
$$=(0.4+a)(a+b)=0.5(0.4+a).$$

因此$0.5(0.4+a)=a$，解得$a=0.4,b=0.5-0.4=0.1$，故选择(D).

例 3.3.3 设二维随机变量(X_1,Y_1)和(X_2,Y_2)的概率密度分别为$f_1(x,y)$与$f_2(x,y)$，令

$$f(x,y)=af_1(x,y)+bf_2(x,y),$$

则 $f(x,y)$是某个二维连续型随机变量的概率密度，当且仅当 a,b 满足条件(　　).

(A)$a+b=1$　　　　(B)$a>0$ 且 $b>0$

(C)$0\leqslant a\leqslant 1,0\leqslant b\leqslant 1$　　　　(D)$a\geqslant 0,b\geqslant 0$ 且 $a+b=1$

解　应选(D).

函数 $f(x,y)$为某个二维连续型随机变量的概率密度的充分必要条件是

$$f(x,y)\geqslant 0,\quad \int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty}f(x,y)dy=1,$$

即同时有

$$f(x,y)=af_1(x,y)+bf_2(x,y)\geqslant 0,$$

$$\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty}f(x,y)dy=a\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty}f_1(x,y)dy+b\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty}f_2(x,y)dy=a+b=1,$$

故选择(D).

例 3.3.4　设二维随机变量 X,Y 相互独立，且 $X\sim N(0,4),Y\sim N(1,9)$，则$(X,Y)$的概率密度为________，随机变量 $Z=X+Y$ 的概率密度为________.

解　应填$\frac{1}{12\pi}e^{-\frac{x^2}{8}-\frac{(y-1)^2}{18}}$；$\frac{1}{\sqrt{26\pi}}e^{-\frac{(z-1)^2}{26}}$.

由题设，X,Y 的概率密度分别为

$$f_1(x)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{x^2}{2\cdot 2^2}},\quad f_2(y)=\frac{1}{3\sqrt{2\pi}}e^{-\frac{(y-1)^2}{2\cdot 3^2}},$$

则在 X,Y 相互独立的条件下，(X,Y)的概率密度为

$$f(x,y)=f_1(x)f_2(y)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{x^2}{2\cdot 2^2}}\cdot\frac{1}{3\sqrt{2\pi}}e^{-\frac{(y-1)^2}{2\cdot 3^2}}=\frac{1}{12\pi}e^{-\frac{x^2}{8}-\frac{(y-1)^2}{18}}.$$

又 $\xi=X\pm Y\sim N(0\pm 1,4+9)$，因此 $Z=X+Y$ 的概率密度为

$$f(z)=\frac{1}{\sqrt{26\pi}}e^{-\frac{(z-1)^2}{26}}.$$

例 3.3.5　已知二维随机变量(X,Y)在 G 上服从均匀分布，G 由直线$x-y=0,x+y=2$ 与 $y=0$ 围成，求：

(1)边缘概率密度 $f_X(x)$和 $f_Y(y)$；

(2)条件概率密度 $f_{X|Y}(x|y)$.

解　(X,Y)的正概率密度区域如图 3-3-5 所示.

由题设，(X,Y)的概率密度为

$$f(x,y)=\begin{cases}1, & 0\leqslant y\leqslant 1,y\leqslant x\leqslant 2-y,\\ 0, & 其他.\end{cases}$$

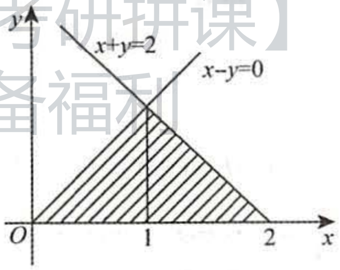


图 3-3-5

(1)

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\begin{cases}\int_0^x 1dy, & 0<x<1,\\ \int_0^{2-x}1dy, & 1<x<2,\\ 0, & 其他\end{cases}$$

$$=\begin{cases}x, & 0<x<1,\\ 2-x, & 1<x<2,\\ 0, & 其他,\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x=\begin{cases}\int_y^{2-y} 1\mathrm{d}x, & 0<y<1,\\ 0, & \text{其他}\end{cases}=\begin{cases}2-2y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(2) 在 $0<y<1$ 的条件下，条件概率密度

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{2-2y},$$

所以

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{1}{2-2y}, & 0<y<1, y\leqslant x\leqslant 2-y,\\ 0, & \text{其他}.\end{cases}$$

例 3.3.6 已知随机变量 X 在区间$[0,1]$上服从均匀分布，在 $X=x(0<x<1)$的条件下，随机变量 Y 在区间$[0,x]$上服从均匀分布，求：

(1)(X,Y)的概率密度；

(2)Y 的概率密度；

(3)概率 $P\{X+Y>1\}$.

解 (1)由题设，X 的边缘概率密度为

$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

在 $X=x(0<x<1)$的条件下，随机变量 Y 在区间$[0,x]$上的条件概率密度为

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{x}, & 0<y<x<1,\\ 0, & \text{其他},\end{cases}$$

因此，当 $0<y<x<1$ 时，$f(x,y)=f_X(x)f_{Y|X}(y|x)=\dfrac{1}{x}$，故$(X,Y)$的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{x}, & 0<y<x<1,\\ 0, & \text{其他}.\end{cases}$$

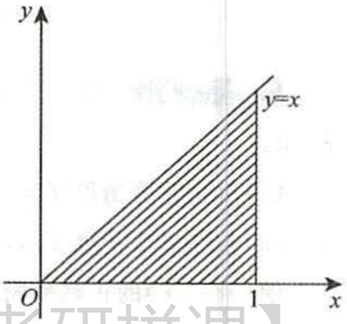


图 3-3-6

(2)(X,Y)的概率密度的正概率密度区域如图 3-3-6 所示.

当 $0<y<1$ 时，$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x=\int_y^1 \frac{1}{x}\mathrm{d}x=-\ln y$；

当 $y\leqslant 0$ 或 $y\geqslant 1$ 时，$f_Y(y)=0$.

从而有

$$f_Y(y)=\begin{cases}-\ln y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(3)$P\{X+Y>1\}=\iint\limits_{x+y>1} f(x,y)\mathrm{d}x\mathrm{d}y=\int_{\frac{1}{2}}^1 \mathrm{d}x\int_{1-x}^x \frac{1}{x}\mathrm{d}y=1-\ln 2.$

【注】 第(1)问中，当 $f_X(x)>0$ 时，才有 $f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}$，所以 $f(x,y)=f_X(x)\cdot f_{Y|X}(y|x)$ 是在 $f_X(x)>0$ 时的表达式，故最终结果要加上 $f_X(x)=0$ 时的定义，一般说来，在分段表达式的“其他”中也就定义全面了.

例 3.3.7 设随机变量 X 与 Y 分别表示两个电子元件的寿命（千小时），已知 X,Y 的联合分布函数为

$$F(x,y)=\begin{cases}1-e^{-0.5x}-e^{-0.5y}+e^{-0.5(x+y)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(1)判断 X 与 Y 的独立性；

(2)求二维随机变量 (X,Y) 的概率密度；

(3)求 2 个元件都可以使用 100 个小时的概率.

解 (1)**方法一** 先计算 X 与 Y 的边缘分布函数.

$$F_X(x)=F(x,+\infty)=\begin{cases}1-e^{-0.5x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

$$F_Y(y)=F(+\infty,y)=\begin{cases}1-e^{-0.5y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

由于 $F(x,y)=F_X(x)F_Y(y)$，所以 X 与 Y 相互独立.

方法二 先计算 X 与 Y 的联合概率密度和边缘概率密度.

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}=\begin{cases}0.25e^{-0.5(x+y)}, & x>0,y>0,\\ 0, & \text{其他},\end{cases}$$

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}0.5e^{-0.5x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}0.5e^{-0.5y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

由于 $f(x,y)=f_X(x)f_Y(y)$，所以 X 与 Y 相互独立.

(2)由(1)方法二已求得二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y}=\begin{cases}0.25e^{-0.5(x+y)}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

(3)2 个元件都可以使用 100 个小时的概率为

$$P\{X\geqslant 0.1,Y\geqslant 0.1\}=P\{X\geqslant 0.1\}P\{Y\geqslant 0.1\}=[1-F_X(0.1)][1-F_Y(0.1)]=e^{-0.1}.$$

二、利用分布计算概率（求与二维随机变量相关事件的概率）

(1)计算公式.

设 D 为平面上的某个区域，则

$$P\{(X,Y)\in D\}=\begin{cases}\displaystyle\sum_{(x_i,y_j)\in D}P\{X=x_i,Y=y_j\}, & (X,Y)\sim p_{ij},\\ \displaystyle\iint_D f(x,y)\mathrm{d}x\mathrm{d}y, & (X,Y)\sim f(x,y).\end{cases}$$

特别地，

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}=\begin{cases}\sum\limits_{x_i\leqslant x,y_j\leqslant y}P\{X=x_i,Y=y_j\}, & (X,Y)\sim p_{ij},\\ \int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v, & (X,Y)\sim f(x,y).\end{cases}$$

(2) 计算步骤.

① 确定随机变量的类型，写出 $p_{ij}>0$ 或画出 $f(x,y)\neq 0$ 的区域 G.

② 画出求和或求积分的区域 D.

③ 在 $D\cap G$ 上计算二重和值或二重积分值.

【注】 对不同 x,y 要将二重积分(二重求和)化为累次积分(累次求和)进行计算.

例 3.3.8 设随机变量 X,Y 相互独立，且均服从标准正态分布 $N(0,1)$，则(　　).

(A) $P\{X+Y\geqslant 1\}=\dfrac{1}{2}$　　(B) $P\{X-Y\geqslant 1\}=\dfrac{1}{2}$

(C) $P\{\max\{X,Y\}\geqslant 0\}=\dfrac{1}{4}$　　(D) $P\{\min\{X,Y\}\geqslant 0\}=\dfrac{1}{4}$

解 应选(D).

因为 X,Y 相互独立，且均服从标准正态分布 $N(0,1)$，所以 $U=X+Y\sim N(0,2)$，$V=X-Y\sim N(0,2)$，排除(A)，(B). 又

$$P\{\max\{X,Y\}\geqslant 0\}=1-P\{\max\{X,Y\}<0\}=1-P\{X<0,Y<0\}=\frac{3}{4},$$

$$P\{\min\{X,Y\}\geqslant 0\}=P\{X\geqslant 0,Y\geqslant 0\}=\frac{1}{4},$$

故应选择(D).

例 3.3.9 已知随机变量 (X,Y) 的概率密度

$$f(x,y)=\begin{cases}A\mathrm{e}^{-y}, & 0<x<y,\\ 0, & \text{其他},\end{cases}$$

试求常数 A，并计算概率 $P\{X+Y\geqslant 1\}$，$P\left\{\dfrac{X}{Y}\leqslant\dfrac{1}{2}\right\}$.

解 由于

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x\mathrm{d}y=\int_{0}^{+\infty}\mathrm{d}x\int_{x}^{+\infty}A\mathrm{e}^{-y}\mathrm{d}y=A\int_{0}^{+\infty}\mathrm{e}^{-x}\mathrm{d}x=A,$$

所以 $A=1$.

$$\begin{aligned}P\{X+Y\geqslant 1\}&=1-P\{X+Y<1\}=1-\iint\limits_{x+y<1}f(x,y)\mathrm{d}x\mathrm{d}y\\&=1-\int_{0}^{\frac{1}{2}}\mathrm{d}x\int_{x}^{1-x}\mathrm{e}^{-y}\mathrm{d}y=1-\int_{0}^{\frac{1}{2}}(\mathrm{e}^{-x}-\mathrm{e}^{x-1})\mathrm{d}x\\&=1-(1+\mathrm{e}^{-1}-2\mathrm{e}^{-\frac{1}{2}})=2\mathrm{e}^{-\frac{1}{2}}-\mathrm{e}^{-1},\end{aligned}$$

积分区域如图 3-3-7 中阴影部分所示.

$$P\left\{\frac{X}{Y}\leqslant\frac{1}{2}\right\}=\iint\limits_{\frac{x}{y}\leqslant\frac{1}{2}}f(x,y)\mathrm{d}x\mathrm{d}y=\int_{0}^{+\infty}\mathrm{d}x\int_{2x}^{+\infty}\mathrm{e}^{-y}\mathrm{d}y=\frac{1}{2},$$

积分区域如图 3-3-8 中阴影部分所示.

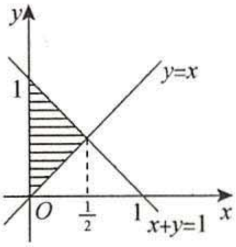


图 3-3-7

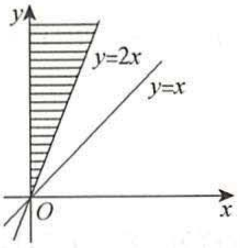


图 3-3-8

三、求函数的分布

例 3.3.10 将两封信投入 3 个信箱，设 X_1,X_2 分别表示第一个和第二个信箱投进的信的数量，求：

(1)(X_1,X_2)的联合分布、边缘分布，并判断 X_1 与 X_2 的独立性；

(2)在条件 $X_2=1$ 下，X_1 的条件分布；

(3)随机变量 $Y_1=X_1+X_2$ 和 $Y_2=X_1-X_2$ 的分布.

解 (1)X_1 与 X_2 可能取值为 0,1,2，样本点总数为 $3^2=9$，则

$$p_{00}=P\{X_1=0,X_2=0\}=\frac{1}{9},\quad p_{01}=P\{X_1=0,X_2=1\}=\frac{2}{9},$$

$$p_{02}=P\{X_1=0,X_2=2\}=\frac{1}{9},\quad p_{10}=P\{X_1=1,X_2=0\}=\frac{2}{9},$$

$$p_{11}=P\{X_1=1,X_2=1\}=\frac{2}{9},\quad p_{12}=P\{X_1=1,X_2=2\}=0,$$

$$p_{20}=P\{X_1=2,X_2=0\}=\frac{1}{9},\quad p_{21}=p_{22}=0,$$

于是(X_1,X_2)的联合分布为

X_2 / X_1	0	1	2	$p_{i\cdot}$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{4}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0	$\frac{4}{9}$
2	$\frac{1}{9}$	0	0	$\frac{1}{9}$
$p_{\cdot j}$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$	1

因为 $p_{00}\neq p_{\cdot 0}\cdot p_{0\cdot}$，所以 X_1 与 X_2 不相互独立.

(2)
$$P\{X_1=0|X_2=1\}=\frac{p_{01}}{p_{\cdot 1}}=\frac{\frac{2}{9}}{\frac{4}{9}}=\frac{1}{2},\quad P\{X_1=1|X_2=1\}=\frac{p_{11}}{p_{\cdot 1}}=\frac{\frac{2}{9}}{\frac{4}{9}}=\frac{1}{2},$$

$$P\{X_1=2|X_2=1\}=0,$$

所以在条件 $X_2=1$ 下，X_1 的条件分布为 $\begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$.

(3)$Y_1=X_1+X_2$ 取值为 0,1,2,于是

$$P\{Y_1=0\}=p_{00}=\frac{1}{9},\quad P\{Y_1=1\}=p_{01}+p_{10}=\frac{4}{9},\quad P\{Y_1=2\}=p_{02}+p_{11}+p_{20}=\frac{4}{9},$$

从而有
$$Y_1\sim\begin{bmatrix}0&1&2\\ \frac{1}{9}&\frac{4}{9}&\frac{4}{9}\end{bmatrix}.$$

又 $Y_2=X_1-X_2$ 取值为 $-2,-1,0,1,2$,于是

$$P\{Y_2=-2\}=p_{02}=\frac{1}{9},\quad P\{Y_2=-1\}=p_{01}=\frac{2}{9},\quad P\{Y_2=0\}=p_{00}+p_{11}=\frac{1}{3},$$

$$P\{Y_2=1\}=p_{10}=\frac{2}{9},\quad P\{Y_2=2\}=p_{20}=\frac{1}{9},$$

从而有
$$Y_2\sim\begin{bmatrix}-2&-1&0&1&2\\ \frac{1}{9}&\frac{2}{9}&\frac{1}{3}&\frac{2}{9}&\frac{1}{9}\end{bmatrix}.$$

【注】 随机变量 $Y_1=X_1+X_2$ 的概率取值恰好为(X_1,X_2)的联合分布中平行于副对角线连线上的元素之和(见分布表),随机变量 $Y_2=X_1-X_2$ 的概率取值恰好为(X_1,X_2)的联合分布中平行于主对角线连线上的元素之和(见分布表),很有规律.

例 3.3.11 设随机变量 X 与 Y 相互独立,其中 X 的概率分布为$\begin{pmatrix}1&2\\0.3&0.7\end{pmatrix}$,而 Y 的概率密度为 $f(y)$,求随机变量 $U=X+Y$ 的概率密度 $g(u)$.

解 本题 X 是离散型随机变量,Y 是连续型随机变量,在独立条件下,$U=X+Y$ 仍然是连续型随机变量,求其概率密度,应先从求分布函数入手. 由全概率公式有

$$\begin{aligned}G(u)&=P\{U\leqslant u\}=P\{X+Y\leqslant u\}\\&=P\{X=1\}P\{X+Y\leqslant u\mid X=1\}+P\{X=2\}P\{X+Y\leqslant u\mid X=2\}\\&=0.3P\{Y\leqslant u-1\}+0.7P\{Y\leqslant u-2\}=0.3F(u-1)+0.7F(u-2).\end{aligned}$$

故 U 的概率密度为 $\quad g(u)=G'(u)=0.3f(u-1)+0.7f(u-2).$

【注】 本题给出应用全概率公式讨论一个离散型、一个连续型随机变量和的分布问题的方法,不论是计算概率还是求随机变量的分布我们常常都是这样处理的,它具有一般性,请读者务必注意.

例 3.3.12 设随机变量 X,Y 相互独立,它们的概率密度分别为

$$f_X(x)=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & 其他,\end{cases}\quad f_Y(y)=\begin{cases}2y, & 0<y<1,\\ 0, & 其他.\end{cases}$$

(1)求(X,Y)的概率密度;

(2)求 $Z=X+Y$ 的概率密度.

解 (1)由于 X,Y 相互独立,(X,Y)的概率密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}2y\mathrm{e}^{-x}, & x>0,0<y<1,\\ 0, & 其他.\end{cases}$$

(2)**方法一** 分布函数法. 正概率区域 D 与所求概率 $F_Z(z)=P\{X+Y\leqslant z\}$ 的积分区域的公共部分有三种不同的组合形式(如图 3-3-9)，于是：

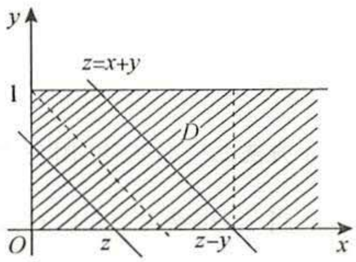


图 3-3-9

当 $z<0$ 时，$F_Z(z)=0$；

当 $0\leqslant z<1$ 时，$F_Z(z)=\int_0^z \mathrm{d}x\int_0^{z-x} 2y\mathrm{e}^{-x}\mathrm{d}y=z^2-2z-2\mathrm{e}^{-z}+2$；

当 $z\geqslant 1$ 时，$F_Z(z)=1-\int_0^1 \mathrm{d}y\int_{z-y}^{+\infty} 2y\mathrm{e}^{-x}\mathrm{d}x=1-2\mathrm{e}^{-z}$.

因此 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\begin{cases}2z+2\mathrm{e}^{-z}-2, & 0<z<1,\\ 2\mathrm{e}^{-z}, & z>1,\\ 0, & \text{其他}.\end{cases}$$

方法二 公式法. 在 X,Y 相互独立的条件下，用公式 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x$ 计算，有

$$f_X(x)f_Y(z-x)=\begin{cases}2(z-x)\mathrm{e}^{-x}, & x>0,0<z-x<1,\\ 0, & \text{其他},\end{cases}$$

其中 $f_X(x)f_Y(z-x)$ 正值区域如图 3-3-10 所示，于是：

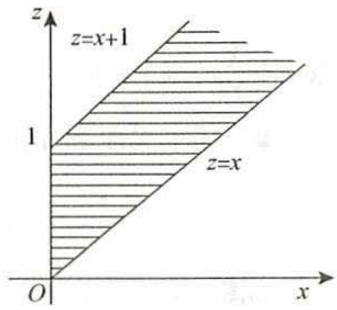


图 3-3-10

当 $z<0$ 时，$f_Z(z)=0$；

当 $0<z<1$ 时，$f_Z(z)=\int_0^z 2(z-x)\mathrm{e}^{-x}\mathrm{d}x=2z+2\mathrm{e}^{-z}-2$；

当 $z>1$ 时，$f_Z(z)=\int_{z-1}^z 2(z-x)\mathrm{e}^{-x}\mathrm{d}x=2\mathrm{e}^{-z}$.

因此 $Z=X+Y$ 的概率密度为

$$f_Z(z)=\begin{cases}2z+2\mathrm{e}^{-z}-2, & 0<z<1,\\ 2\mathrm{e}^{-z}, & z>1,\\ 0, & \text{其他}.\end{cases}$$

例 3.3.13 设二维随机变量 (X,Y) 在矩形区域 $D=\{(x,y)\mid 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$ 上服从均匀分布，求边长为 X 和 Y 的矩形面积 Z 的概率密度.

解 由题设 $Z=XY(Z>0)$，(X,Y) 的概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{2}, & 0\leqslant x\leqslant 2,0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

方法一 分布函数法. 正概率区域 D 与所求概率 $F_Z(z)=P\{XY\leqslant z\}$ 的积分区域的公共部分有三种不同的组合形式(如图 3-3-11)，于是：

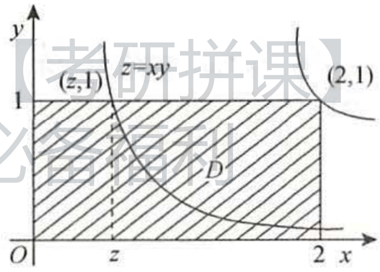


图 3-3-11

当 $z<0$ 时，$F_Z(z)=0$；

当 $0\leqslant z<2$ 时，$F_Z(z)=\int_0^z \mathrm{d}x\int_0^1 \frac{1}{2}\mathrm{d}y+\int_z^2 \mathrm{d}x\int_0^{\frac{z}{x}} \frac{1}{2}\mathrm{d}y$

$$=\frac{1}{2}z(1-\ln z+\ln 2);$$

当 $z\geqslant 2$ 时，$F_Z(z)=1$.

因此 $Z=XY$ 的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{1}{2}(\ln 2-\ln z), & 0<z<2,\\ 0, & 其他.\end{cases}$$

方法二　公式法. 用公式 $f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)dx$, $0<z<x\leqslant 2$.

当 $z\leqslant 0$ 或 $z\geqslant 2$ 时, $f_Z(z)=0$;

当 $0<z<2$ 时, $f_Z(z)=\frac{1}{2}\int_z^2\frac{1}{x}dx=\frac{1}{2}(\ln 2-\ln z)$.

因此 $Z=XY$ 的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{1}{2}(\ln 2-\ln z), & 0<z<2,\\ 0, & 其他.\end{cases}$$

例 3.3.14　设随机变量 X,Y 相互独立, 它们都服从指数分布, 概率密度分别为

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0,\\ 0, & 其他,\end{cases}\quad f_Y(y)=\begin{cases}\lambda e^{-\lambda y}, & y\geqslant 0,\\ 0, & 其他,\end{cases}$$

求 $Z=\frac{X}{Y}$ 的概率密度.

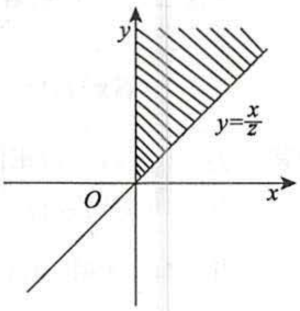


图 3-3-12

解　由于 X,Y 相互独立, (X,Y) 的概率密度为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\lambda^2 e^{-\lambda(x+y)}, & x\geqslant 0, y\geqslant 0,\\ 0, & 其他.\end{cases}$$

如图 3-3-12 所示: 当 $z<0$ 时, $F_Z(z)=0$;

当 $z\geqslant 0$ 时,

$$F_Z(z)=P\left\{\frac{X}{Y}\leqslant z\right\}=\int_0^{+\infty}dx\int_{\frac{x}{z}}^{+\infty}f(x,y)dy=\frac{z}{z+1}.$$

因此 $Z=\frac{X}{Y}$ 的概率密度为

$$f_Z(z)=\begin{cases}\dfrac{1}{(1+z)^2}, & z>0,\\ 0, & 其他.\end{cases}$$

基础习题精练

习题

3.3.1　设随机变量 X 和 Y 的联合分布函数为

$$F(x,y)=\begin{cases}0, & \min\{x,y\}<0,\\ \min\{x,y\}, & 0\leqslant\min\{x,y\}<1,\\ 1, & \min\{x,y\}\geqslant 1,\end{cases}$$

则随机变量 X 的分布函数为________.

3.3.2　如果二维随机变量 (X,Y) 的分布函数为

$$F(x,y)=\begin{cases}1-e^{-\lambda_1 x}-e^{-\lambda_2 y}+e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}}, & \lambda_1,\lambda_2,\lambda_{12}>0,x>0,y>0,\\ 0, & 其他,\end{cases}$$

试求 X 和 Y 各自的边缘分布函数.

3.3.3　随机变量(X,Y)的概率密度

$$f(x,y)=\frac{1+\sin x\sin y}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}\ (-\infty<x,y<+\infty)$$

的两个边缘概率密度 $f_1(x),f_2(y)$分别为________.

3.3.4　求以下给出的(X,Y)的概率密度的边缘概率密度 $f_X(x)$和 $f_Y(y)$：

(1)$f_1(x,y)=\begin{cases}e^{-y}, & 0<x<y,\\ 0, & 其他;\end{cases}$

(2)$f_2(x,y)=\begin{cases}\frac{5}{4}(x^2+y), & 0<y<1-x^2,\\ 0, & 其他.\end{cases}$

3.3.5　设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}k, & 0<x^2<y<x<1,\\ 0, & 其他.\end{cases}$$

(1)试求常数 k；

(2)求 $P\left\{X>\frac{1}{2}\right\}$和 $P\left\{Y<\frac{1}{2}\right\}$.

3.3.6　已知随机变量 X 与 Y 相互独立，X 服从参数为 λ 的指数分布，$P\{Y=-1\}=\frac{1}{4}$，$P\{Y=1\}=\frac{3}{4}$，试计算概率 $P\{X-Y\leqslant 1\}$，$P\{XY\leqslant 2\}$.

3.3.7　设随机变量 X_1,X_2,X_3,X_4 相互独立且同分布，$P\{X_i=0\}=0.6$，$P\{X_i=1\}=0.4$，$i=1,2,3,4$，求行列式 $X=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}$ 的概率分布.

3.3.8　袋中有编号为1,1,2,3的四个球，现从中无放回地任取两次，每次任取一个，设 X_1,X_2 分别为第一次、第二次取到的球的号码，试求：

(1)(X_1,X_2)的联合分布，并判断 X_1 与 X_2 的独立性；

(2)在 $X_2=2$ 的条件下，X_1 的条件分布；

(3)随机变量 $Y=X_1X_2$ 的分布.

3.3.9　设随机变量 X 和 Y 独立同分布，且 X 的分布函数为 $F(x)$，则 $Z=\max\{X,Y\}$ 的分布函数为(　　).

(A)$F^2(z)$　　(B)$F(x)F(y)$　　(C)$1-[1-F(z)]^2$　　(D)$[1-F(x)][1-F(y)]$

3.3.10　设二维随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}2-x-y, & 0<x<1,0<y<1,\\ 0, & 其他.\end{cases}$$

(1)求 $P\{X>2Y\}$；

(2)求 $Z=X+Y$ 的概率密度 $f_Z(z)$.

3.3.11　设随机变量 X 的概率分布为 $P\{X=1\}=P\{X=2\}=\frac{1}{2}$. 在给定 $X=i$ 的条件下，随机变量 Y

服从均匀分布$U(0,i)(i=1,2)$. 求 Y 的分布函数 $F_Y(y)$和概率密度 $f_Y(y)$.

解答

3.3.1 $F(x)=\begin{cases}0, & x<0,\\ x, & 0\leqslant x<1,\\ 1, & x\geqslant 1\end{cases}$ **解** $F(x)$是 $F(x,y)$的边缘分布函数, $F(x)=F(x,+\infty)$.

$\lim\limits_{y\to+\infty}\min\{x,y\}=x$,因此 $F(x)=\begin{cases}0, & x<0,\\ x, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$

3.3.2 **解** 因为

$$\lim_{y\to+\infty}(1-\mathrm{e}^{-\lambda_1 x}-\mathrm{e}^{-\lambda_2 y}+\mathrm{e}^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}})=1-\mathrm{e}^{-\lambda_1 x},$$
$$\lim_{x\to+\infty}(1-\mathrm{e}^{-\lambda_1 x}-\mathrm{e}^{-\lambda_2 y}+\mathrm{e}^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}})=1-\mathrm{e}^{-\lambda_2 y},$$

所以 X 和 Y 各自的边缘分布函数为

$$F_X(x)=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)=\begin{cases}1-\mathrm{e}^{-\lambda_1 x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

$$F_Y(y)=F(+\infty,y)=\lim_{x\to+\infty}F(x,y)=\begin{cases}1-\mathrm{e}^{-\lambda_2 y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

3.3.3 $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$; $\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$ **解** 由边缘概率密度的公式,有

$$f_1(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2}{2}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y+\frac{\sin x}{2\pi}\mathrm{e}^{-\frac{x^2}{2}}\int_{-\infty}^{+\infty}\sin y\mathrm{e}^{-\frac{y^2}{2}}\mathrm{d}y=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},$$

即 $f_1(x)$是标准正态概率密度. 由对称性知 $f_2(y)$也是标准正态概率密度,所以 $f_2(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}$.

3.3.4 **解** (1)当 $x>0$ 时,有 $f_X(x)=\int_x^{+\infty}\mathrm{e}^{-y}\mathrm{d}y=\mathrm{e}^{-x}$, 所以 X 的边缘概率密度为

$$f_X(x)=\begin{cases}\mathrm{e}^{-x}, & x>0,\\ 0, & \text{其他},\end{cases}$$

这是指数分布 $E(1)$.

而当 $y>0$ 时,有 $f_Y(y)=\int_0^y\mathrm{e}^{-y}\mathrm{d}x=y\mathrm{e}^{-y}$,所以 Y 的边缘概率密度为 $f_Y(y)=\begin{cases}y\mathrm{e}^{-y}, & y>0,\\ 0, & \text{其他}.\end{cases}$

(2)因为 $f_2(x,y)$的非零区域如图 3-3-13 阴影部分所示,所以当$-1<x<1$时,有

$$f_X(x)=\int_0^{1-x^2}\frac{5}{4}(x^2+y)\mathrm{d}y=\frac{5}{4}x^2(1-x^2)+\frac{5}{8}(1-x^2)^2=\frac{5}{8}(1-x^4),$$

所以 X 的边缘概率密度为

$$f_X(x)=\begin{cases}\frac{5}{8}(1-x^4), & -1<x<1,\\ 0, & \text{其他}.\end{cases}$$

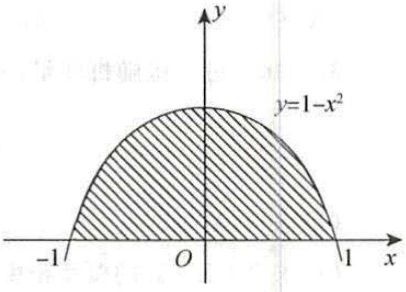


图 3-3-13

又因为当 $0<y<1$ 时,有

$$f_Y(y)=\int_{-\sqrt{1-y}}^{\sqrt{1-y}}\frac{5}{4}(x^2+y)\mathrm{d}x=\frac{5}{12}x^3\Big|_{-\sqrt{1-y}}^{\sqrt{1-y}}+\frac{5}{2}y\sqrt{1-y}$$

$$=\frac{5}{6}\sqrt{1-y}(1+2y),$$

所以 Y 的边缘概率密度为 $f_Y(y)=\begin{cases}\frac{5}{6}\sqrt{1-y}(1+2y), & 0<y<1,\\ 0, & 其他.\end{cases}$

3.3.5 解 (1) $f(x,y)$ 的非零区域如图 3-3-14(a)阴影部分所示.

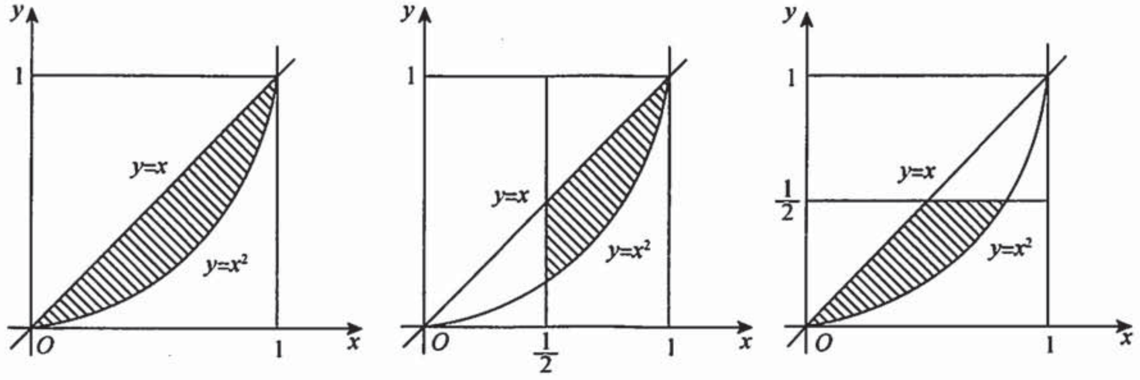
由 $k\int_0^1 \mathrm{d}x\int_{x^2}^{x}\mathrm{d}y = k\int_0^1 (x-x^2)\mathrm{d}x = \frac{k}{6}=1$，解得 $k=6$.

(2) $f(x,y)$ 的非零区域与 $x>\frac{1}{2}$ 的交集如图 3-3-14(b)阴影部分所示，所以

$$P\left\{X>\frac{1}{2}\right\}=6\int_{\frac{1}{2}}^{1}\mathrm{d}x\int_{x^2}^{x}\mathrm{d}y=6\int_{\frac{1}{2}}^{1}(x-x^2)\mathrm{d}x=6\left(\frac{1}{2}x^2-\frac{1}{3}x^3\right)\Bigg|_{\frac{1}{2}}^{1}=\frac{1}{2}.$$

又因为 $f(x,y)$ 的非零区域与 $y<\frac{1}{2}$ 的交集如图 3-3-14(c)阴影部分所示，所以

$$P\left\{Y<\frac{1}{2}\right\}=6\int_0^{\frac{1}{2}}\mathrm{d}y\int_y^{\sqrt{y}}\mathrm{d}x=6\int_0^{\frac{1}{2}}(\sqrt{y}-y)\mathrm{d}y=\sqrt{2}-\frac{3}{4}.$$



(a) (b) (c)

图 3-3-14

3.3.6 解 由于 Y 是离散型随机变量，X 与 Y 相互独立，因此计算与 X,Y 有关的事件 A 的概率时，自然想到将 A 对 Y 的可能值作全集分解，应用全概率公式计算 $P(A)$.

已知 $X\sim f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$ $P\{Y=-1\}=\frac{1}{4}$，$P\{Y=1\}=\frac{3}{4}$，X 与 Y 独立，所以由全概率公式得

$$\begin{aligned}P\{X-Y\leqslant 1\}&=P\{X-Y\leqslant 1,Y=1\}+P\{X-Y\leqslant 1,Y=-1\}\\&=P\{X-1\leqslant 1,Y=1\}+P\{X+1\leqslant 1,Y=-1\}\\&=P\{Y=1\}P\{X\leqslant 2\}+P\{Y=-1\}P\{X\leqslant 0\}\\&=\frac{3}{4}\int_0^2\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x=\frac{3}{4}(1-\mathrm{e}^{-2\lambda}),\end{aligned}$$

$$\begin{aligned}P\{XY\leqslant 2\}&=P\{XY\leqslant 2,Y=1\}+P\{XY\leqslant 2,Y=-1\}\\&=P\{X\leqslant 2,Y=1\}+P\{-X\leqslant 2,Y=-1\}\\&=P\{Y=1\}P\{X\leqslant 2\}+P\{Y=-1\}P\{X\geqslant -2\}\\&=\frac{3}{4}\int_0^2\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x+\frac{1}{4}\int_0^{+\infty}\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x=1-\frac{3}{4}\mathrm{e}^{-2\lambda}.\end{aligned}$$

3.3.7 解 因为 $X=X_1X_4-X_2X_3$，所以 X 的可能取值为 $-1,0,1$. 又因为 $Z_1=X_1X_4$ 与 $Z_2=X_2X_3$ 独立同分布，且 $P\{Z_1=1\}=P\{X_1=1,X_4=1\}=0.4^2=0.16$,

$$P\{Z_1=0\}=1-P\{Z_1=1\}=0.84,$$

同理，得 $P\{Z_2=1\}=0.16, P\{Z_2=0\}=0.84$. 由此得

$$P\{X=-1\}=P\{Z_1=0,Z_2=1\}=P\{Z_1=0\}P\{Z_2=1\}=0.84\times 0.16=0.1344,$$

$$P\{X=0\}=P\{Z_1=0,Z_2=0\}+P\{Z_1=1,Z_2=1\}=0.84^2+0.16^2=0.7312,$$

$$P\{X=1\}=P\{Z_1=1,Z_2=0\}=P\{Z_1=1\}P\{Z_2=0\}=0.16\times 0.84=0.1344.$$

所以 X 的概率分布为

X	-1	0	1
P	0.1344	0.7312	0.1344

3.3.8 解 (1) X_1 与 X_2 可能的取值为 1,2,3，则

$$p_{11}=P\{X_1=1,X_2=1\}=\frac{2}{4}\times\frac{1}{3}=\frac{1}{6},\quad p_{12}=P\{X_1=1,X_2=2\}=\frac{2}{4}\times\frac{1}{3}=\frac{1}{6},$$

$$p_{13}=P\{X_1=1,X_2=3\}=\frac{2}{4}\times\frac{1}{3}=\frac{1}{6},\quad p_{21}=P\{X_1=2,X_2=1\}=\frac{1}{4}\times\frac{2}{3}=\frac{1}{6},$$

$$p_{22}=0,\quad p_{23}=P\{X_1=2,X_2=3\}=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12},\quad p_{31}=P\{X_1=3,X_2=1\}=\frac{1}{4}\times\frac{2}{3}=\frac{1}{6},$$

$$p_{32}=P\{X_1=3,X_2=2\}=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12},\quad p_{33}=0,$$

于是 (X_1,X_2) 的联合分布为

X_2 X_1	1	2	3	$p_{i\cdot}$
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
2	$\frac{1}{6}$	0	$\frac{1}{12}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{12}$	0	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

因为 $p_{11}\neq p_{\cdot 1}\cdot p_{1\cdot}$，所以 X_1 与 X_2 不相互独立.

(2)由于

$$P\{X_1=1\mid X_2=2\}=\frac{p_{12}}{p_{\cdot 2}}=\frac{\frac{1}{6}}{\frac{1}{4}}=\frac{2}{3},$$

$$P\{X_1=2\mid X_2=2\}=0,\quad P\{X_1=3\mid X_2=2\}=\frac{p_{32}}{p_{\cdot 2}}=\frac{\frac{1}{12}}{\frac{1}{4}}=\frac{1}{3},$$

所以在 $X_2=2$ 的条件下，X_1 的条件分布为 $\begin{bmatrix} 1 & 2 & 3 \\ \frac{2}{3} & 0 & \frac{1}{3} \end{bmatrix}$.

(3) $Y=X_1X_2$ 取 1,2,3,6，于是

$$P\{Y=1\}=p_{11}=\frac{1}{6},\quad P\{Y=2\}=p_{12}+p_{21}=\frac{1}{3},$$

$$P\{Y=3\}=p_{13}+p_{31}=\frac{1}{3},\quad P\{Y=6\}=p_{23}+p_{32}=\frac{1}{6},$$

从而有

$$Y\sim\begin{pmatrix}1 & 2 & 3 & 6\\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}\end{pmatrix}.$$

3.3.9 (A) **解** 由于 X 和 Y 独立同分布，则 Y 的分布函数为 $F(y)$.

$$G(z)=P\{Z\leqslant z\}=P\{\max\{X,Y\}\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=P\{X\leqslant z\}P\{Y\leqslant z\}=F^2(z),$$

故选择(A).

3.3.10 **解** (1)$P\{X>2Y\}=\iint\limits_{x>2y}f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\mathrm{d}x\int_0^{\frac{x}{2}}(2-x-y)\mathrm{d}y=\int_0^1\left(x-\frac{5}{8}x^2\right)\mathrm{d}x=\frac{7}{24}$.

(2)$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\mathrm{d}x$，其中

$$f(x,z-x)=\begin{cases}2-x-(z-x), & 0<x<1,0<z-x<1,\\ 0, & \text{其他}\end{cases}=\begin{cases}2-z, & 0<x<1,0<z-x<1,\\ 0, & \text{其他}.\end{cases}$$

当 $z\leqslant 0$ 或 $z\geqslant 2$ 时，$f_Z(z)=0$；当 $0<z<1$ 时，$f_Z(z)=\int_0^z(2-z)\mathrm{d}x=z(2-z)$；

当 $1<z<2$ 时，$f_Z(z)=\int_{z-1}^1(2-z)\mathrm{d}x=(2-z)^2$.

则 Z 的概率密度为 $f_Z(z)=\begin{cases}z(2-z), & 0<z<1,\\ (2-z)^2, & 1<z<2,\\ 0, & \text{其他}.\end{cases}$

3.3.11 **解** $F_Y(y)=P\{Y\leqslant y\}=P\{X=1\}P\{Y\leqslant y|X=1\}+P\{X=2\}P\{Y\leqslant y|X=2\}$

$$=\frac{1}{2}P\{Y\leqslant y|X=1\}+\frac{1}{2}P\{Y\leqslant y|X=2\}.$$

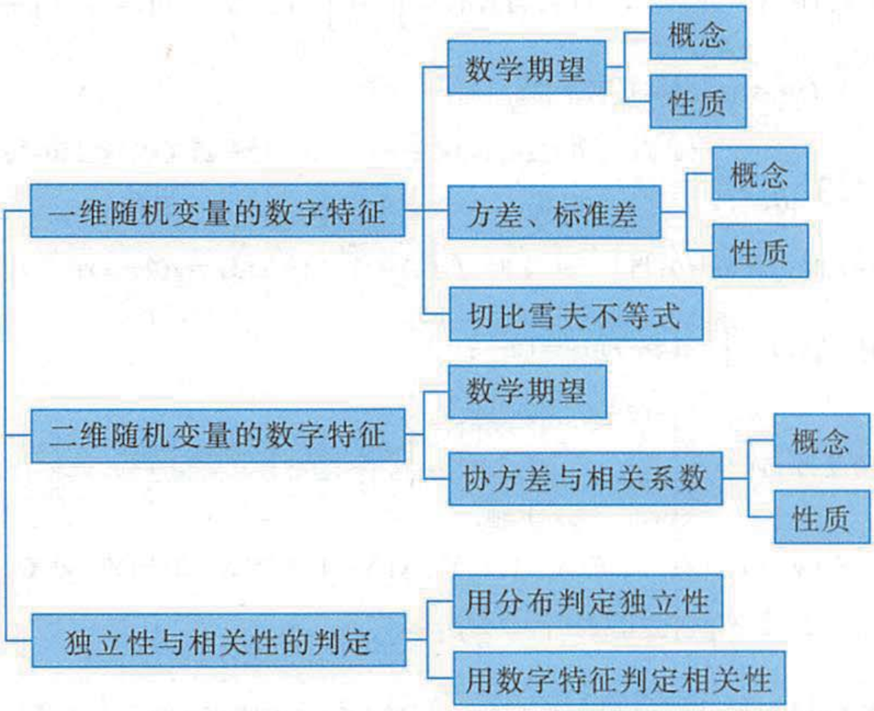
当 $y<0$ 时，$F_Y(y)=0$；当 $0\leqslant y<1$ 时，$F_Y(y)=\frac{3y}{4}$；当 $1\leqslant y<2$ 时，$F_Y(y)=\frac{1}{2}+\frac{y}{4}$；

当 $y\geqslant 2$ 时，$F_Y(y)=1$.

所以随机变量 Y 的分布函数为 $F_Y(y)=\begin{cases}0, & y<0,\\ \frac{3y}{4}, & 0\leqslant y<1,\\ \frac{1}{2}+\frac{y}{4}, & 1\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$ 概率密度为 $f_Y(y)=\begin{cases}\frac{3}{4}, & 0<y<1,\\ \frac{1}{4}, & 1<y<2,\\ 0, & \text{其他}.\end{cases}$

第4讲 随机变量的数字特征

基础知识结构



基础内容精讲

一、一维随机变量的数字特征

（一）随机变量的数学期望

1. 概念

设 X 是随机变量，Y 是 X 的函数，$Y=g(X)$.

(1)如果 X 是离散型随机变量，其分布列为 $p_i=P\{X=x_i\}(i=1,2,\cdots)$. 若级数 $\sum\limits_{i=1}^{\infty}x_ip_i$ 绝对收敛，则称随机变量 X 的数学期望存在，并将级数和 $\sum\limits_{i=1}^{\infty}x_ip_i$ 称为随机变量 X 的**数学期望**，记为$E(X)$或 EX，即

$$EX=\sum_{i=1}^{\infty}x_ip_i,$$

否则称 X 的数学期望不存在.

若级数 $\sum\limits_{i=1}^{\infty} g(x_i)p_i$ 绝对收敛,则称 $Y=g(X)$的数学期望 $E[g(X)]$存在,且 $E[g(X)]=\sum\limits_{i=1}^{\infty} g(x_i)p_i$,否则称 $g(X)$的数学期望不存在.

(2)如果 X 是连续型随机变量,其概率密度为 $f(x)$.若积分 $\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$ 绝对收敛,则称 X 的数学期望存在,且 $EX=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$,否则称 X 的数学期望不存在.

若积分 $\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$ 绝对收敛,则称 $g(X)$的数学期望存在,且 $E[g(X)]=\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$,否则称 $g(X)$的数学期望不存在.

【注】 (1)数学期望又称为概率平均值,常常简称期望或均值.数学期望是描述随机变量平均取值状况特征的指标,它刻画随机变量的一切可能值的集中位置.

(2)在数学期望的定义中要求级数(或积分)绝对收敛,否则称期望不存在.这是因为 X 的期望存在要求与 X 的取值顺序无关,即要求任意改变 x_i 的次序不应改变EX 的存在性,这在数学上就要求级数(或积分)绝对收敛,况且绝对收敛又有很多性质也便于数学上的处理.只不过在考题中,命题人基本上会避开对“绝对收敛”的考查,而仅要求考生会计算即可.

2. 性质

(1)对任意常数 a_i 和随机变量 $X_i(i=1,2,\cdots,n)$有

$$E\left(\sum_{i=1}^{n} a_iX_i\right)=\sum_{i=1}^{n} a_iEX_i.$$

特别地,

$$Ec=c,\quad E(aX+c)=aEX+c,\quad E(X\pm Y)=EX\pm EY.$$

(2)设 X 与 Y 相互独立,则

$$E(XY)=EX\cdot EY,\quad E[g_1(X)\cdot g_2(Y)]=E[g_1(X)]\cdot E[g_2(Y)].$$

一般地,设 $X_1,X_2,\cdots,X_n$ 相互独立,则

$$E\left(\prod_{i=1}^{n} X_i\right)=\prod_{i=1}^{n} EX_i,\quad E\left[\prod_{i=1}^{n} g_i(X_i)\right]=\prod_{i=1}^{n} E[g_i(X_i)].$$

(二)随机变量的方差、标准差、切比雪夫不等式

1. 概念

设 X 是随机变量,如果 $E[(X-EX)^2]$存在,则称 $E[(X-EX)^2]$为 X 的方差,记为 DX,即

$$DX=E[(X-EX)^2]=E(X^2)-(EX)^2.$$

称 $\sqrt{DX}$为 X 的标准差或均方差,记为 $\sigma(X)$,称随机变量 $X^*=\dfrac{X-EX}{\sqrt{DX}}$为 X 的标准化随机变量,此时 $EX^*=0,DX^*=1$.

2. 性质

(1)$DX\geqslant 0,E(X^2)=DX+(EX)^2\geqslant(EX)^2$.

(2)$Dc=0$(c 为常数).

(3)$D(aX+b)=a^2DX$.

(4)$D(X\pm Y)=DX+DY\pm 2\mathrm{Cov}(X,Y)$.

(5)如果 X 与 Y 相互独立，则

$$D(aX+bY)=a^2DX+b^2DY.$$

一般地，如果 $X_1,X_2,\cdots,X_n$ 相互独立，$g_i(x)$ 为 x 的连续函数，则

$$D\left(\sum_{i=1}^{n}a_iX_i\right)=\sum_{i=1}^{n}a_i^2DX_i,$$

$$D\left[\sum_{i=1}^{n}g_i(X_i)\right]=\sum_{i=1}^{n}D[g_i(X_i)].$$

3. 切比雪夫不等式

如果随机变量 X 的方差 DX 存在，则对任意 $\varepsilon>0$，有

$$P\{|X-EX|\geqslant\varepsilon\}\leqslant\frac{DX}{\varepsilon^2}\text{或}P\{|X-EX|<\varepsilon\}\geqslant 1-\frac{DX}{\varepsilon^2}.$$

【注】 由切比雪夫不等式知，当 DX 愈小时，概率 $P\{|X-EX|<\varepsilon\}$ 愈大，这表明方差是刻画随机变量与其期望值偏离程度的量，是描述随机变量 X“分散程度”特征的指标.

我们将常用分布的期望和方差列表如下.

分布	分布列 p_k 或概率密度 $f(x)$	期望	方差
0-1 分布	$P\{X=k\}=p^k(1-p)^{1-k},k=0,1$	p	$p(1-p)$
二项分布 $B(n,p)$	$P\{X=k\}=\mathrm{C}_n^kp^k(1-p)^{n-k},k=0,1,\cdots,n$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$P\{X=k\}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},k=0,1,\cdots$	λ	λ
几何分布 $G(p)$	$P\{X=k\}=(1-p)^{k-1}p,k=1,2,\cdots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},-\infty<x<+\infty$	μ	σ^2
均匀分布 $U(a,b)$	$f(x)=\frac{1}{b-a},a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x)=\lambda\mathrm{e}^{-\lambda x},x>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

注：表中仅列出各分布概率密度的非零区域.

二、二维随机变量的数字特征

(一)二维随机变量函数的数学期望

设 X,Y 为随机变量，$g(X,Y)$ 为 X,Y 的函数.

(1)如果(X,Y)为离散型随机变量，其联合分布为

$$p_{ij}=P\{X=x_i,Y=y_j\}(i,j=1,2,\cdots),$$

若级数 $\sum_i\sum_j g(x_i,y_j)p_{ij}$ 绝对收敛，则定义

$$E[g(X,Y)]=\sum_i\sum_j g(x_i,y_j)p_{ij}.$$

（2）如果(X,Y)为连续型随机变量，其概率密度为 $f(x,y)$，若积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$ 绝对收敛，则定义

$$E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y.$$

（二）两个随机变量的协方差与相关系数

1. 概念

如果随机变量 X 与 Y 的方差存在且 $DX>0,DY>0$，则称 $E[(X-EX)(Y-EY)]$为随机变量 X 与 Y 的**协方差**，并记为 $\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)=E[(X-EX)(Y-EY)]=E(XY)-EX\cdot EY,$$

其中

$$E(XY)=\begin{cases}\sum_i\sum_j x_iy_jP\{X=x_i,Y=y_j\}(\text{离散型}),\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)\mathrm{d}x\mathrm{d}y(\text{连续型}).\end{cases}$$

称 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}$为随机变量 X 与 Y 的**相关系数**. 如果 $\rho_{XY}=0$，则称 X 与 Y **不相关**；如果 $\rho_{XY}\neq0$，则称 X 与 Y **相关**.

【注】 协方差 $\mathrm{Cov}(X,Y)$是描述随机变量 X 与 Y 之间偏差的关联程度的，比如研究父亲身高 X 与孩子身高 Y 之间的偏差程度，便可用 $\mathrm{Cov}(X,Y)$来刻画. 相关系数 ρ_{XY}是描述随机变量 X 与 Y 之间的线性相依性，$|\rho_{XY}|$的大小是刻画 X 与 Y 之间线性相关程度的一种度量. $\rho_{XY}=0$ 表示 X 与 Y 之间不存在线性关系，故称 X 与 Y 不相关，但这并不意味着 X 与 Y 之间不存在相依关系，它们之间还可存在某种非线性关系.

2. 性质

（1）对称性 $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$， $\rho_{XY}=\rho_{YX}$，

$\mathrm{Cov}(X,X)=DX$， $\rho_{XX}=1$.

（2）线性性 $\mathrm{Cov}(X,c)=0$， $\mathrm{Cov}(aX+b,Y)=a\mathrm{Cov}(X,Y)$，

$\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$.

一般地，

$$\mathrm{Cov}\left(\sum_{i=1}^{n}a_iX_i,Y\right)=\sum_{i=1}^{n}a_i\mathrm{Cov}(X_i,Y).$$

（3）相关系数有界性 $|\rho_{XY}|\leqslant1$.

（4）线性关系下的相关系数

如果 $Y=aX+b$，则 $\rho_{XY}=\begin{cases}1, & a>0,\\ -1, & a<0.\end{cases}$

基础例题精解

一、一维随机变量的数字特征

1. 离散型随机变量的数学期望与方差

例 3.4.1 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ 0.2, & -1\leqslant x<0,\\ 0.8, & 0\leqslant x<1,\\ 1, & x\geqslant 1,\end{cases}$$

则 $EX=$________，$E(2X+5)=$________，$E(X^2)=$________，$D(X^2)=$________.

解 应填 0；5；0.4；0.24.

求 X 的期望与方差，先求 X 的分布列，即有

$$X\sim\begin{pmatrix}-1 & 0 & 1\\ 0.2 & 0.6 & 0.2\end{pmatrix}.$$

因此

$$EX=-1\times 0.2+0\times 0.6+1\times 0.2=0,$$

$$E(2X+5)=2EX+5=5,$$

$$E(X^2)=(-1)^2\times 0.2+0^2\times 0.6+1^2\times 0.2=0.4,$$

$$D(X^2)=E(X^4)-[E(X^2)]^2=(-1)^4\times 0.2+0^4\times 0.6+1^4\times 0.2-0.4^2=0.24.$$

例 3.4.2 随机变量 X 服从参数为 λ 的泊松分布，且已知 $E[(X-1)(X-2)]=1$，则 $\lambda=$________.

解 应填 1.

由于随机变量 X 服从参数为 λ 的泊松分布，所以 $EX=DX=\lambda$，从而有

$$E[(X-1)(X-2)]=E(X^2)-3EX+2=DX+(EX)^2-3EX+2=\lambda^2-2\lambda+2=1,$$

解方程 $\lambda^2-2\lambda+1=0$，得 $\lambda=1$.

例 3.4.3 设试验成功一次的概率为 p，进行 100 次独立重复试验，当 $p=$________时成功次数的标准差最大，且其最大值为________.

解 应填 $\frac{1}{2}$；5.

设成功次数为 X，则 $X\sim B(100,p)$，X 的标准差为 $\sqrt{DX}=\sqrt{np(1-p)}$.

由于 $p(1-p)\leqslant\frac{1}{4}$，当且仅当 $p=\frac{1}{2}$ 时等号成立，故当 $p=\frac{1}{2}$ 时成功次数的标准差最大，且其最大值为 $\sqrt{100\times\frac{1}{4}}=5$.

2. 离散型随机变量函数的数学期望与方差

主要用两种方法计算：一是直接利用离散型随机变量函数数学期望的计算公式；二是先计算离散型随机变量函数的概率分布，再用定义计算.

例 3.4.4 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2}\cos\dfrac{x}{2}, & 0\leqslant x<\pi,\\ 0, & \text{其他},\end{cases}$$

对 X 独立观察 4 次，用 Y 表示观察值大于$\frac{\pi}{3}$的次数，求 Y^2 的数学期望.

解 由题易知 $P\left\{X>\frac{\pi}{3}\right\}=\int_{\frac{\pi}{3}}^{+\infty}f(x)\mathrm{d}x=\int_{\frac{\pi}{3}}^{\pi}\frac{1}{2}\cos\frac{x}{2}\mathrm{d}x=\frac{1}{2}$，所以 $Y\sim B\left(4,\frac{1}{2}\right)$，于是

$$EY=4\times\frac{1}{2}=2,\quad DY=4\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=1,$$

所以

$$E(Y^2)=DY+(EY)^2=1+4=5.$$

例 3.4.5 假设一部机器一天内发生故障的概率为 0.2，机器发生故障时，停业工作一天，若一周 5 个工作日里机器无故障可获利润 10 万元，发生一次故障仍可获利 5 万元，发生两次故障所获利润为零，发生三次或三次以上故障亏损 2 万元，求一周内的期望利润值.

解 设一周内发生故障的次数为 X，则 $X\sim B(5,0.2)$，有

$$P\{X=0\}=C_5^0\times0.2^0\times0.8^5=0.32768,$$

$$P\{X=1\}=C_5^1\times0.2\times0.8^4=0.4096,$$

$$P\{X=2\}=C_5^2\times0.2^2\times0.8^3=0.2048,$$

$$P\{X\geqslant3\}=1-\sum_{k=0}^{2}P\{X=k\}=0.05792.$$

设一周内可获利润为 Y 万元，则

$$Y=f(X)=\begin{cases}10, & X=0,\\ 5, & X=1,\\ 0, & X=2,\\ -2, & X\geqslant3,\end{cases}$$

所以

$$EY=10\times0.32768+5\times0.4096+0\times0.2048+(-2)\times0.05792=5.20896(\text{万元}).$$

3. 连续型随机变量的数学期望与方差

例 3.4.6 连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{\pi(1+x^2)}(-\infty<x<+\infty),$$

则 EX(　　).

(A)等于 0　　(B)等于 1　　(C)等于 π　　(D)不存在

解 应选(D).

由连续型随机变量数学期望的定义，有

$$EX=\int_{-\infty}^{+\infty}\frac{x}{\pi(1+x^2)}\mathrm{d}x=\frac{1}{2\pi}\ln(1+x^2)\Big|_{-\infty}^{+\infty},$$

由于该积分发散，所以 EX 不存在，故选择(D).

例 3.4.7 连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{\pi}}\mathrm{e}^{-x^2+2x-1}\ (-\infty<x<+\infty),$$

则 $EX=$________，$\sqrt{DX}=$________.

解 应填 $1;\dfrac{1}{\sqrt{2}}$.

在讨论涉及形如 e^{-x^2+2x-1} 的概率密度的数字特征时，应尽可能地利用正态分布概率密度的性质，于是由

$$f(x)=\frac{1}{\sqrt{\pi}}\mathrm{e}^{-x^2+2x-1}=\frac{1}{\sqrt{2\pi}\cdot\frac{1}{\sqrt{2}}}\exp\left\{-\frac{(x-1)^2}{2\left(\frac{1}{\sqrt{2}}\right)^2}\right\},$$

知 $EX=1$，$\sqrt{DX}=\dfrac{1}{\sqrt{2}}$.

例 3.4.8 连续型随机变量 X 服从参数为 1 的指数分布，则 $E(X+\mathrm{e}^{-2X})=$________，$D(\mathrm{e}^{-2X})=$________.

解 应填 $\dfrac{4}{3};\dfrac{4}{45}$.

$EX=\dfrac{1}{1}=1$，$E(\mathrm{e}^{-2X})=\displaystyle\int_0^{+\infty}\mathrm{e}^{-2x}\mathrm{e}^{-x}\mathrm{d}x=\int_0^{+\infty}\mathrm{e}^{-3x}\mathrm{d}x=\frac{1}{3}$，所以

$$E(X+\mathrm{e}^{-2X})=1+\frac{1}{3}=\frac{4}{3}.$$

又由

$$E[(\mathrm{e}^{-2X})^2]=E(\mathrm{e}^{-4X})=\int_0^{+\infty}\mathrm{e}^{-4x}\mathrm{e}^{-x}\mathrm{d}x=\int_0^{+\infty}\mathrm{e}^{-5x}\mathrm{d}x=\frac{1}{5},$$

得 $D(\mathrm{e}^{-2X})=E[(\mathrm{e}^{-2X})^2]-[E(\mathrm{e}^{-2X})]^2=\dfrac{1}{5}-\left(\dfrac{1}{3}\right)^2=\dfrac{4}{45}$.

4. 连续型随机变量函数的数学期望与方差

例 3.4.9 连续型随机变量 X 的概率密度为

$$f(x)=\frac{1}{\pi(1+x^2)}\ (-\infty<x<+\infty),$$

求 $E(\min\{|X|,1\})$.

解 由对称性知，

$$\begin{aligned}E(\min\{|X|,1\})&=\int_{-\infty}^{+\infty}\min\{|x|,1\}f(x)\mathrm{d}x=2\int_0^{+\infty}\min\{x,1\}f(x)\mathrm{d}x\\&=2\left[\int_0^1 xf(x)\mathrm{d}x+\int_1^{+\infty}f(x)\mathrm{d}x\right]\\&=\frac{2}{\pi}\left(\int_0^1\frac{x}{1+x^2}\mathrm{d}x+\int_1^{+\infty}\frac{1}{1+x^2}\mathrm{d}x\right)\\&=\frac{1}{\pi}\ln(1+x^2)\Big|_0^1+\frac{2}{\pi}\arctan x\Big|_1^{+\infty}\\&=\frac{1}{\pi}\ln 2+\frac{1}{2}.\end{aligned}$$

二、二维随机变量的数字特征

1. 二维离散型随机变量的数字特征的计算

(1)给出 X,Y 的联合分布，相当于给出了一个充分条件，在此基础上求数字特征，完全可以通过准确套用公式解决. 由于较简单，考研在此命题频率不高.

(2)给出 X,Y 的边缘分布，相当于给出了一个必要条件. 需再附加一些条件，如某些数字特征或事件概率，才能确定 X,Y 的联合分布. 这种问题是考研常考的.

例 3.4.10 随机变量

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix},\quad Y \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

且 $\mathrm{Cov}(X,Y)=\frac{1}{8}$，求 X 与 Y 的联合分布.

解 依题设，$EX=\frac{3}{4}$，$EY=\frac{1}{2}$，于是

$$\mathrm{Cov}(X,Y)=E(XY)-EXEY=E(XY)-\frac{3}{8}=\frac{1}{8},$$

得 $E(XY)=\frac{1}{2}$，即有

$$E(XY)=\frac{1}{2}=P\{X=1,Y=1\},$$

$$P\{X=0,Y=1\}=P\{Y=1\}-P\{X=1,Y=1\}=0,$$

$$P\{X=0,Y=0\}=P\{X=0\}-P\{X=0,Y=1\}=\frac{1}{4},$$

$$P\{X=1,Y=0\}=1-\frac{1}{4}-\frac{1}{2}-0=\frac{1}{4},$$

因此，X 与 Y 的联合分布为

Y X	0	1	$p_{i\cdot}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{2}$	1

例 3.4.11 设随机变量 X 与 Y 的概率分布分别为

X	0	1
P	$\frac{1}{3}$	$\frac{2}{3}$

Y	-1	0	1
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

且 $P\{X^2=Y^2\}=1$.

(1)求随机变量(X,Y)的概率分布；

(2)求 $Z=XY$ 的概率分布；

(3)求 X 与 Y 的相关系数 ρ_{XY}.

解　(1)由于 $P\{X^2=Y^2\}=1$,知 $P\{X^2\neq Y^2\}=0$,即

$$P\{X=0,Y=-1\}=P\{X=0,Y=1\}=P\{X=1,Y=0\}=0,$$

从而有

$$P\{X=0,Y=0\}=P\{X=0\}=\frac{1}{3},$$

$$P\{X=1,Y=-1\}=P\{Y=-1\}=\frac{1}{3},$$

$$P\{X=1,Y=1\}=P\{Y=1\}=\frac{1}{3},$$

因此,(X,Y)的概率分布为

X \ Y	-1	0	1	$p_{i\cdot}$
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

(2)$Z=XY$ 的可能取值为$-1,0,1$.

$$P\{Z=-1\}=P\{X=1,Y=-1\}=\frac{1}{3},$$

$$P\{Z=1\}=P\{X=1,Y=1\}=\frac{1}{3},$$

$$P\{Z=0\}=1-\frac{1}{3}-\frac{1}{3}=\frac{1}{3},$$

所以

$$Z\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}.$$

(3)由于

$$EX=0\times\frac{1}{3}+1\times\frac{2}{3}=\frac{2}{3},\quad E(X^2)=\frac{2}{3},\quad DX=\frac{2}{3}-\left(\frac{2}{3}\right)^2=\frac{2}{9},$$

$$EY=-1\times\frac{1}{3}+0\times\frac{1}{3}+1\times\frac{1}{3}=0,\quad E(Y^2)=\frac{2}{3},\quad DY=\frac{2}{3},$$

$$E(XY)=-1\times1\times\frac{1}{3}+1\times1\times\frac{1}{3}=0,$$

因此

$$\rho_{XY}=\frac{E(XY)-EXEY}{\sqrt{DX}\sqrt{DY}}=0.$$

2. 二维连续型随机变量的数字特征的计算

除了上一部分所述，还要注意这里增加了概率密度 $f(x,y)$ 等，可能涉及积分计算.

例 3.4.12 已知随机变量 $(X,Y)\sim N(0,0;1,1;0.5)$，$U=\max\{X,Y\}$，$V=\min\{X,Y\}$，则 $E(U+V)=$ ________，$E(UV)=$ ________.

解 应填 0；0.5.

$$U=\max\{X,Y\}=\frac{1}{2}[(X+Y)+|X-Y|],$$

$$V=\min\{X,Y\}=\frac{1}{2}[(X+Y)-|X-Y|],$$

则

$$U+V=X+Y,\quad U-V=|X-Y|,\quad UV=XY.$$

所以

$$E(U+V)=EX+EY=0.$$

$$E(UV)=E(XY)=\operatorname{Cov}(X,Y)+EXEY=0.5+0=0.5.$$

三、独立性与相关性的判定、切比雪夫不等式

1. 独立性与相关性的判定

(1)随机变量 X 与 Y 相互独立，意指对任意实数 x,y，事件 $\{X\leqslant x\}$ 与 $\{Y\leqslant y\}$ 相互独立，即 (X,Y) 的联合分布等于边缘分布相乘：$F(x,y)=F_X(x)\cdot F_Y(y)$.

若 (X,Y) 是连续型的，则 X 与 Y 独立的充要条件是

$$f(x,y)=f_X(x)\cdot f_Y(y);$$

若 (X,Y) 是离散型的，则 X 与 Y 独立的充要条件是

$$P\{X=x_i,Y=y_j\}=P\{X=x_i\}\cdot P\{Y=y_j\}.$$

记住，我们是通过分布来判定独立性的.

(2)随机变量 X 与 Y 不相关，意指 X 与 Y 之间不存在线性相依性，即 $\rho_{XY}=0$，其充要条件是

$$\rho_{XY}=0\Leftrightarrow\operatorname{Cov}(X,Y)=0\Leftrightarrow E(XY)=EX\cdot EY\Leftrightarrow D(X\pm Y)=DX+DY.$$

记住，我们一般是通过数字特征来判定相关性的.

(3)几个重要结论：

① 如果 X 与 Y 独立，则 X,Y 不相关，反之不然；

② 如果 (X,Y) 服从二维正态分布，则 X,Y 独立 $\Leftrightarrow X,Y$ 不相关；

③ 由①知，如果 X 与 Y 相关，则 X,Y 不独立.

综上所述，我们在讨论随机变量 X 与 Y 相关性、独立性时，总是先计算 $\operatorname{Cov}(X,Y)$，而后按下列程序进行判断或再计算：

$$\operatorname{Cov}(X,Y)=E(XY)-EXEY\begin{cases}\neq0\Leftrightarrow X\text{ 与 }Y\text{ 相关}\Rightarrow X\text{ 与 }Y\text{ 不独立},\\ =0\Leftrightarrow X\text{ 与 }Y\text{ 不相关，通过分布推断}\begin{cases}X,Y\text{ 独立},\\ X,Y\text{ 不独立}.\end{cases}\end{cases}$$

【注】 上述讨论均假设方差存在并且不为零.

例 3.4.13 设随机变量(X,Y)服从二维正态分布，则随机变量$\xi=X+Y,\eta=X-Y$不相关的充分必要条件为(　　).

(A)$EX=EY$　　　　(B)$E(X^2)-(EX)^2=E(Y^2)-(EY)^2$

(C)$E(X^2)=E(Y^2)$　　　　(D)$E(X^2)+(EX)^2=E(Y^2)+(EY)^2$

解 应选(B).

随机变量ξ,η不相关的充分必要条件是协方差$\mathrm{Cov}(\xi,\eta)$为零，由

$$E(\xi\eta)=E(X^2)-E(Y^2),\quad E\xi=EX+EY,\quad E\eta=EX-EY,$$

得

$$\begin{aligned}\mathrm{Cov}(\xi,\eta)&=E(\xi\eta)-E\xi E\eta=E(X^2)-E(Y^2)-(EX+EY)(EX-EY)\\&=E(X^2)-(EX)^2-[E(Y^2)-(EY)^2],\end{aligned}$$

因此，随机变量ξ,η不相关的充分必要条件为$E(X^2)-(EX)^2-[E(Y^2)-(EY)^2]=0$，即$E(X^2)-(EX)^2=E(Y^2)-(EY)^2$，故选择(B).

例 3.4.14 设随机变量X的概率密度为

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},x\in(-\infty,+\infty),$$

证明：(1)X与$|X|$不相关；(2)X与$|X|$不独立.

证明 (1)由对称性

$$EX=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=0,\quad E(X|X|)=\int_{-\infty}^{+\infty}x|x|f(x)\mathrm{d}x=0,$$

因此，$\mathrm{Cov}(X,|X|)=E(X|X|)-EXE(|X|)=0$，知$X$与$|X|$不相关.

(2)对任意常数a，不妨设$a>0$.

$$P\{X\leqslant a\}=\frac{1}{2}\int_{-\infty}^{a}\mathrm{e}^{-|x|}\mathrm{d}x=\frac{1}{2}\int_{-\infty}^{0}\mathrm{e}^{x}\mathrm{d}x+\frac{1}{2}\int_{0}^{a}\mathrm{e}^{-x}\mathrm{d}x=1-\frac{1}{2}\mathrm{e}^{-a},$$

$$P\{|X|\leqslant a\}=\frac{1}{2}\int_{-a}^{a}\mathrm{e}^{-|x|}\mathrm{d}x=\int_{0}^{a}\mathrm{e}^{-x}\mathrm{d}x=1-\mathrm{e}^{-a}.$$

又事件$\{|X|\leqslant a\}\subset\{X\leqslant a\}$，即$\{|X|\leqslant a\}\cap\{X\leqslant a\}=\{|X|\leqslant a\}$，故有

$$P\{X\leqslant a,|X|\leqslant a\}=P\{|X|\leqslant a\}=1-\mathrm{e}^{-a},$$

从而有

$$P\{X\leqslant a,|X|\leqslant a\}\neq P\{X\leqslant a\}\cdot P\{|X|\leqslant a\},$$

因此，X与$|X|$不独立.

2. 切比雪夫不等式

应用切比雪夫不等式：(1)可以估算随机变量在某范围取值的概率；(2)可以证明某些收敛性问题(见例 3.6.11).

例 3.4.15 设X,Y为随机变量，数学期望都是 2，方差分别为 1 和 4，相关系数为 0.5，试用切比雪夫不等式估计概率$P\{|X-Y|\geqslant 6\}$.

解 应用不等式$P\{|\xi-E\xi|\geqslant\varepsilon\}\leqslant\frac{D\xi}{\varepsilon^2}$估计.

已知$EX=EY=2,DX=1,DY=4,\rho_{XY}=0.5=\frac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\cdot\sqrt{DY}}=\frac{\mathrm{Cov}(X,Y)}{2}$. 记$\xi=X-Y$，则

$$E\xi=0,\quad D\xi=D(X-Y)=DX+DY-2\mathrm{Cov}(X,Y)=1+4-2\times0.5\times2=3.$$

取 $\varepsilon=6$. 由切比雪夫不等式，得

$$P\{|X-Y|\geqslant 6\}\leqslant\frac{D(X-Y)}{6^2}=\frac{3}{36}=\frac{1}{12}.$$

基础习题精练

习题

3.4.1 已知连续型随机变量 X 与 Y 有相同的概率密度，且

$$X\sim f(x)=\begin{cases}2x\theta^2, & 0<x<\dfrac{1}{\theta},\\ 0, & \text{其他}\end{cases}\quad(\theta>0),\quad E[a(X+2Y)]=\frac{1}{\theta},$$

则 $a=($　　$)$.

(A) $\frac{2}{3}$　　(B) $\frac{1}{2}$　　(C) $\frac{1}{3}$　　(D) $\frac{1}{6}$

3.4.2 甲、乙两人相约于某地在 12:00～13:00 会面，设 X,Y 分别是甲、乙到达的时间，且假设 X 和 Y 相互独立，已知 X,Y 的概率密度分别为

$$f_X(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他},\end{cases}\qquad f_Y(y)=\begin{cases}2y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

求先到达者需要等待的时间的数学期望.

3.4.3 设 A,B 为两个随机事件，且 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{3}$，$P(A|B)=\frac{1}{2}$. 令

$$X=\begin{cases}1, & A\text{ 发生},\\ 0, & A\text{ 不发生},\end{cases}\qquad Y=\begin{cases}1, & B\text{ 发生},\\ 0, & B\text{ 不发生},\end{cases}$$

求：(1)二维随机变量 (X,Y) 的概率分布；

(2) X,Y 的相关系数 ρ_{XY}；

(3) $Z=X^2+Y^2$ 的概率分布.

3.4.4 设随机变量 (X,Y) 服从二维正态分布，X 与 Y 的期望值为 μ，方差为 σ^2，X,Y 的相关系数为 $\rho_{XY}=0$，记 $Z_1=2X+Y$，$Z_2=2X-Y$，则 Z_1,Z_2 的相关系数为(　　).

(A) 0　　(B) $\frac{3}{5}$　　(C) $\frac{3}{\sqrt{15}}$　　(D) $\frac{3}{\sqrt{10}}$

3.4.5 设 $X\sim N(0,\sigma^2)$，$Y=X^3$.

(1)求 EY；

(2)判断 X 与 Y 是否相关？是否独立？说明理由.

3.4.6 设随机变量 X 在 $[-1,1]$ 上服从均匀分布，$Y=X^3$，则 X 与 Y(　　).

(A)不相关且相互独立　　(B)不相关且相互不独立

(C)相关且相互独立　　(D)相关且相互不独立

3.4.7 设 X 为随机变量，$P\{|X-EX|<\varepsilon\}\geqslant 0.9$，$DX=0.009$，试用切比雪夫不等式估计 ε 的取值范围.

解答

3.4.1 (B) 解 由于 X 与 Y 有相同的概率密度，因此 $EX=EY$，于是

$$E[a(X+2Y)]=a(EX+2EY)=3aEX=3a\int_{-\infty}^{+\infty}xf(x)\,\mathrm{d}x=3a\int_{0}^{\frac{1}{\theta}}2x^2\theta^2\,\mathrm{d}x$$

$$=3a\cdot\frac{2}{3}x^3\theta^2\Big|_0^{\frac{1}{\theta}}=\frac{2a}{\theta}=\frac{1}{\theta},$$

解得 $a=\frac{1}{2}$，故选择(B).

3.4.2 解 由题意可得 $f(x,y)=\begin{cases}6x^2y, & 0<x<1,0<y<1,\\ 0, & \text{其他}.\end{cases}$

根据题意可知，所求即为 $|X-Y|$ 的数学期望，如图 3-4-1 所示，有

$$E(|X-Y|)=\int_0^1\int_0^1|x-y|\,6x^2y\,\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{D_1}[-(x-y)6x^2y]\,\mathrm{d}x\mathrm{d}y+\iint\limits_{D_2}(x-y)6x^2y\,\mathrm{d}x\mathrm{d}y$$

$$=\frac{1}{12}+\frac{1}{6}=\frac{1}{4}(\text{小时}).$$

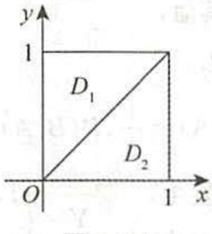


图 3-4-1

3.4.3 解 (1)(X,Y)的取值为$(0,0),(0,1),(1,0),(1,1)$，并有

$$P\{X=0,Y=0\}=P(\overline{A}\,\overline{B})=P(\overline{A\cup B})=1-P(A)-P(B)+P(AB),$$

其中 $P(AB)=P(A)P(B|A)=\frac{1}{12}$，$P(B)=\frac{P(AB)}{P(A|B)}=\frac{\frac{1}{12}}{\frac{1}{2}}=\frac{1}{6}$，因此

$$P\{X=0,Y=0\}=1-\frac{1}{4}-\frac{1}{6}+\frac{1}{12}=\frac{2}{3},$$

$$P\{X=0,Y=1\}=P(\overline{A}B)=P(B)-P(AB)=\frac{1}{6}-\frac{1}{12}=\frac{1}{12},$$

$$P\{X=1,Y=0\}=P(A\overline{B})=P(A)-P(AB)=\frac{1}{4}-\frac{1}{12}=\frac{1}{6},$$

$$P\{X=1,Y=1\}=P(AB)=\frac{1}{12},$$

从而得(X,Y)的概率分布为

X \ Y	0	1	$p_{i\cdot}$
0	$\frac{2}{3}$	$\frac{1}{12}$	$\frac{3}{4}$
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{5}{6}$	$\frac{1}{6}$	1

(2)由 X 与 Y 的联合分布,得 X 与 Y 的边缘分布分别为

$$X\sim\begin{pmatrix}0 & 1\\ \frac{3}{4} & \frac{1}{4}\end{pmatrix},\quad Y\sim\begin{pmatrix}0 & 1\\ \frac{5}{6} & \frac{1}{6}\end{pmatrix},$$

从而

$$EX=\frac{1}{4},\quad E(X^2)=\frac{1}{4},\quad DX=\frac{1}{4}-\left(\frac{1}{4}\right)^2=\frac{3}{16},$$

$$EY=\frac{1}{6},\quad E(Y^2)=\frac{1}{6},\quad DY=\frac{1}{6}-\left(\frac{1}{6}\right)^2=\frac{5}{36},$$

$$E(XY)=\frac{2}{3}\times0\times0+\frac{1}{12}\times0\times1+\frac{1}{6}\times1\times0+\frac{1}{12}\times1\times1=\frac{1}{12},$$

$$\mathrm{Cov}(X,Y)=E(XY)-EXEY=\frac{1}{24},$$

因此

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}=\frac{1}{\sqrt{15}}=\frac{\sqrt{15}}{15}.$$

(3)$Z=X^2+Y^2$ 可能取值为 0,1,2,因此

$$P\{Z=0\}=P\{X=0,Y=0\}=\frac{2}{3},$$

$$P\{Z=1\}=P\{X=0,Y=1\}+P\{X=1,Y=0\}=\frac{1}{4},$$

$$P\{Z=2\}=P\{X=1,Y=1\}=\frac{1}{12},$$

从而

$$Z\sim\begin{pmatrix}0 & 1 & 2\\ \frac{2}{3} & \frac{1}{4} & \frac{1}{12}\end{pmatrix}.$$

3.4.4 (B) **解** 利用已有的关系计算,不必从 (Z_1,Z_2) 的分布去考虑.

由于 $\rho_{XY}=0$,知 X 与 Y 不相关且相互独立.于是

$$DZ_1=4DX+DY=5\sigma^2,\quad DZ_2=4DX+DY=5\sigma^2,$$

$$\begin{aligned}\mathrm{Cov}(Z_1,Z_2)&=E[(2X+Y)(2X-Y)]-E(2X+Y)E(2X-Y)\\&=4E(X^2)-4(EX)^2-E(Y^2)+(EY)^2\\&=4DX-DY=3\sigma^2,\end{aligned}$$

因此,Z_1,Z_2 的相关系数 $\rho=\frac{\mathrm{Cov}(Z_1,Z_2)}{\sqrt{DZ_1}\sqrt{DZ_2}}=\frac{3}{5}$,故选择(B).

3.4.5 **解** (1)已知 X 的概率密度 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$,$f(x)$ 为 x 的偶函数,故

$$EY=E(X^3)=\int_{-\infty}^{+\infty}x^3f(x)\mathrm{d}x=0.$$

$$(2)E(XY)=E(X^4)=\int_{-\infty}^{+\infty} x^4 \cdot \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \mathrm{d}x=\frac{2}{\sqrt{2\pi\sigma}}\int_{0}^{+\infty} x^4 e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} \mathrm{d}x \neq 0,$$

故 $EX \cdot EY \neq E(XY)$，故 X 与 Y 相关⇒X 与 Y 不独立.

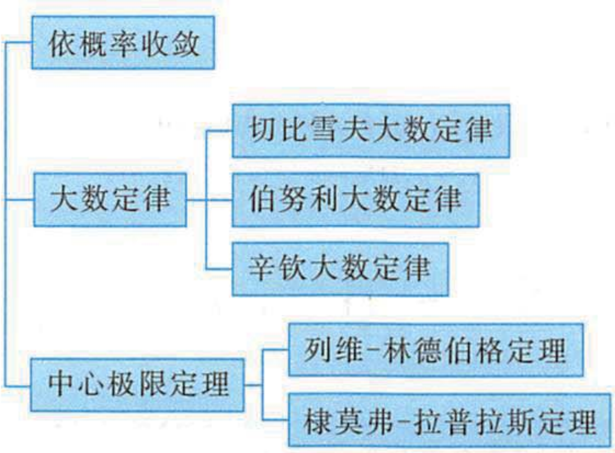
3.4.6 (D) **解** 显然，结论(C)对任意两个随机变量都不正确. 又 X 与 Y 之间存在函数关系，因而不可能独立，结论(A)不正确. 由题设，$EX=0$，$E(XY)=\int_{-1}^{1} x^4 \cdot \frac{1}{2} \mathrm{d}x \neq 0$，从而知 $\mathrm{Cov}(X,Y)\neq 0$，则 X 与 Y 必相关，也必相互不独立，选择(D).

3.4.7 **解** 应用切比雪夫不等式 $P\{|X-EX|<\varepsilon\}\geqslant 1-\frac{DX}{\varepsilon^2}$ 求解. 由题设得

$$P\{|X-EX|<\varepsilon\}\geqslant 1-\frac{0.009}{\varepsilon^2}\geqslant 0.9,\quad \varepsilon^2\geqslant 0.09,\quad \varepsilon\geqslant 0.3.$$

第5讲 大数定律与中心极限定理

基础知识结构



基础内容精讲

一、依概率收敛

设随机变量 X 与随机变量序列 $\{X_n\}(n=1,2,3,\cdots)$，如果对任意的 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\{|X_n-X|\geqslant\varepsilon\}=0 \text{ 或 } \lim_{n\to\infty}P\{|X_n-X|<\varepsilon\}=1,$$

则称随机变量序列 $\{X_n\}$ **依概率收敛于随机变量 X**，记为

$$\lim_{n\to\infty}X_n=X(P)\text{ 或 }X_n\xrightarrow{P}X\ (n\to\infty).$$

【注】 以上定义中将随机变量 X 写成数 a 也成立.

二、大数定律

定理 1（切比雪夫大数定律） 假设 $\{X_n\}(n=1,2,\cdots)$ 是相互独立的随机变量序列，如果方差 DX_i $(i\geqslant1)$ 存在且一致有上界，即存在常数 C，使 $DX_i\leqslant C$ 对一切 $i\geqslant1$ 均成立，则 $\{X_n\}$ 服从大数定律：

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}EX_i.$$

定理 2(伯努利大数定律) 假设 μ_n 是 n 重伯努利试验中事件 A 发生的次数，在每次试验中事件 A 发生的概率为 $p(0<p<1)$，则 $\frac{\mu_n}{n}\xrightarrow{P}p$，即对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{\mu_n}{n}-p\right|<\varepsilon\right\}=1.$$

定理 3(辛钦大数定律) 假设 $\{X_n\}$ 是独立同分布的随机变量序列，如果 $EX_i=\mu(i=1,2,\cdots)$ 存在，则 $\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}\mu$，即对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|<\varepsilon\right\}=1.$$

三、中心极限定理

定理 4(列维-林德伯格定理) 假设 $\{X_n\}$ 是独立同分布的随机变量序列，如果

$$EX_i=\mu,DX_i=\sigma^2>0(i=1,2,\cdots)$$

存在，则对任意的实数 x，有

$$\lim_{n\to\infty}P\left\{\frac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\Phi(x).$$

【注】 (1)定理的三个条件"独立、同分布、期望方差存在"，缺一不可.

(2)只要 X_n 满足定理条件，那么当 n 很大时，独立同分布随机变量的和 $\sum_{i=1}^{n}X_i$ 近似服从正态分布 $N(n\mu,n\sigma^2)$，由此可知，当 n 很大时，有

$$P\left\{a<\sum_{i=1}^{n}X_i<b\right\}\approx\Phi\left(\frac{b-n\mu}{\sqrt{n}\sigma}\right)-\Phi\left(\frac{a-n\mu}{\sqrt{n}\sigma}\right),$$

这常常是解题的依据. 只要题目涉及独立同分布随机变量的和 $\sum_{i=1}^{n}X_i$，我们就要考虑独立同分布中心极限定理.

定理 5(棣莫弗-拉普拉斯定理) 假设随机变量 $Y_n\sim B(n,p)(0<p<1,n\geqslant1)$，则对任意实数 x，有

$$\lim_{n\to\infty}P\left\{\frac{Y_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=\Phi(x).$$

【注】 如果记 $X_i\sim B(1,p)(0<p<1)$，即 $X_i\sim\begin{pmatrix}1 & 0\\ p & 1-p\end{pmatrix}$ 且相互独立，则

$$Y_n=\sum_{i=1}^{n}X_i\sim B(n,p),$$

由定理 4 推出定理 5.

基础例题精解

一、大数定律

例 3.5.1　设 $X_1,X_2,\cdots,X_n,\cdots$是相互独立的随机变量序列，$X_n$ 服从参数为 n 的指数分布$(n\geqslant1)$，则下列随机变量序列中不服从切比雪夫大数定律的是(　　).

(A)$X_1,\frac{1}{2}X_2,\cdots,\frac{1}{n}X_n,\cdots$　　(B)$X_1,X_2,\cdots,X_n,\cdots$

(C)$X_1,2X_2,\cdots,nX_n,\cdots$　　(D)$X_1,2^2X_2,\cdots,n^2X_n,\cdots$

解　应选(D).

切比雪夫大数定律要求$\{X_n\}$相互独立，方差存在且一致有界，即 $DX_n\leqslant C$. 逐一验证各选项是否满足这一条件，从而确定正确选项.

由题设知$\{X_n\}$相互独立，且 $DX_n=\frac{1}{n^2}\leqslant1$，所以(B)满足切比雪夫大数定律的条件.

又

$$D\left(\frac{1}{n}X_n\right)=\frac{1}{n^2}DX_n=\frac{1}{n^4}\leqslant1,\quad D(nX_n)=n^2DX_n=1\leqslant2,$$

由此可知，(A)，(B)，(C)均满足切比雪夫大数定律的条件，然而 $D(n^2X_n)=n^4DX_n=n^2$，选项(D)不满足切比雪夫大数定律的条件，故选择(D).

例 3.5.2　设随机变量序列 $X_1,X_2,\cdots,X_n,\cdots$相互独立，根据辛钦大数定律，当 $n\to\infty$时，$\frac{1}{n}\sum_{i=1}^{n}X_i$依概率收敛于数学期望，只要$\{X_n\}$(　　).

(A)有相同的数学期望　　(B)服从同一离散型分布

(C)服从同一泊松分布　　(D)服从同一连续型分布

解　应选(C).

辛钦大数定律要求$\{X_n\}$独立同分布且数学期望存在. 选项(A)缺少“同分布”，(B)，(D)缺少“数学期望存在”这一条件，因而正确选项是(C).

二、中心极限定理

例 3.5.3　设随机变量 $X_1,X_2,\cdots,X_n$相互独立，$S_n=X_1+X_2+\cdots+X_n$，则根据列维-林德伯格定理，当 n 充分大时，S_n近似服从正态分布，只要 $X_1,X_2,\cdots,X_n$(　　).

(A)有相同期望和方差　　(B)服从同一离散型分布

(C)服从同一指数分布　　(D)服从同一连续型分布

解　应选(C).

列维-林德伯格定理的条件：随机变量 $X_1,X_2,\cdots,X_n$相互独立同分布，并且其数学期望和方差均存在. 由于有相同的数学期望和方差未必服从相同的分布，可见选项(A)不满足定理条件. 满足选项(B)和(D)的随机变量 X_i的数学期望或方差未必存在，故选项(B)和(D)也不满足定理条件. 于是，只有选项(C)成立(指

数分布的数学期望和方差都存在).

例 3.5.4 设 X_n 表示将一枚硬币随意投掷 n 次出现"正面"的次数，则（　　）.

(A) $\lim\limits_{n\to\infty}P\left\{\dfrac{X_n-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$　　(B) $\lim\limits_{n\to\infty}P\left\{\dfrac{2X_n-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$

(C) $\lim\limits_{n\to\infty}P\left\{\dfrac{X_n-2n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$　　(D) $\lim\limits_{n\to\infty}P\left\{\dfrac{2X_n-2n}{\sqrt{n}}\leqslant x\right\}=\Phi(x)$

解　应选(B).

由题设知 $X_n\sim B\left(n,\dfrac{1}{2}\right)$，根据棣莫弗-拉普拉斯定理，得

$$\lim_{n\to\infty}P\left\{\frac{X_n-\frac{1}{2}n}{\sqrt{\frac{1}{4}n}}\leqslant x\right\}=\lim_{n\to\infty}P\left\{\frac{2X_n-n}{\sqrt{n}}\leqslant x\right\}=\Phi(x),$$

故选择(B).

例 3.5.5 生产线生产的产品成箱包装，每箱质量是随机的．假设每箱平均重 50 千克，标准差为 5，若用载重为 5 吨的汽车承运，试用中心极限定理说明每辆汽车最多可以装多少箱，才能保证不超载的概率大于 0.977.($\Phi(2)=0.977$)

解　设 $X_i(i=1,2,\cdots,n)$ 为"第 i 箱产品的质量(千克)"，由题意，X_i 独立同分布，且 $EX_i=50$，$\sqrt{DX_i}=5$，n 箱总质量为 $T_n=\sum\limits_{i=1}^{n}X_i$，从而有

$$ET_n=50n,\quad \sqrt{DT_n}=5\sqrt{n}.$$

由列维-林德伯格定理，$T_n\overset{\text{近似}}{\sim}N(50n,25n)$，因此有

$$P\{T_n\leqslant 5\,000\}=P\left\{\frac{T_n-50n}{5\sqrt{n}}\leqslant\frac{5\,000-50n}{5\sqrt{n}}\right\}\approx\Phi\left(\frac{1\,000-10n}{\sqrt{n}}\right)>0.977,$$

即 $\dfrac{1\,000-10n}{\sqrt{n}}>2$，$n<98.0199$，所以每辆汽车最多可以装 98 箱，才能保证不超载的概率大于 0.977.

基础习题精练

习 题

3.5.1　将一枚骰子重复掷 n 次，则当 $n\to\infty$ 时，n 次掷出点数的算术平均值 $\overline{X}_n$ 依概率收敛于________.

3.5.2　设 $X_1,X_2,\cdots,X_n,\cdots$ 为独立同分布随机变量序列，且均服从参数为 $\lambda(\lambda>1)$ 的指数分布，记 $\Phi(x)$ 为标准正态分布函数，则（　　）.

(A) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\lambda\sqrt{n}}\leqslant x\right\}=\Phi(x)$　　(B) $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum\limits_{i=1}^{n}X_i-n\lambda}{\sqrt{\lambda n}}\leqslant x\right\}=\Phi(x)$

(C) $\lim\limits_{n\to\infty} P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n} X_i - n}{\sqrt{n}} \leqslant x\right\} = \Phi(x)$ (D) $\lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - \lambda}{\sqrt{n\lambda}} \leqslant x\right\} = \Phi(x)$

3.5.3 设某种电气元件不能承受超负荷试验的概率为 0.05. 现在对 100 个这样的元件进行超负荷试验，以 X 表示“不能承受试验而烧毁的元件数”，则根据中心极限定理，$P\{5\leqslant X\leqslant 10\}\approx$________.

解答

3.5.1 $\dfrac{7}{2}$ **解** 设 $X_1, X_2, \cdots, X_n$ 是各次掷出的点数，显然它们独立同分布，每次掷出点数的数学期望等于$\dfrac{7}{2}$. 因此，根据辛钦大数定律，$\overline{X}_n$ 依概率收敛于$\dfrac{7}{2}$.

3.5.2 (C) **解** 由于 $E\left(\sum\limits_{i=1}^{n} X_i\right) = \dfrac{n}{\lambda}$，$D\left(\sum\limits_{i=1}^{n} X_i\right) = \dfrac{n}{\lambda^2}$，所以 $\lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - \dfrac{n}{\lambda}}{\dfrac{\sqrt{n}}{\lambda}} \leqslant x\right\} = \Phi(x)$，即 $\lim\limits_{n\to\infty} P\left\{\dfrac{\lambda\sum\limits_{i=1}^{n} X_i - n}{\sqrt{n}} \leqslant x\right\} = \Phi(x)$，故选择(C).

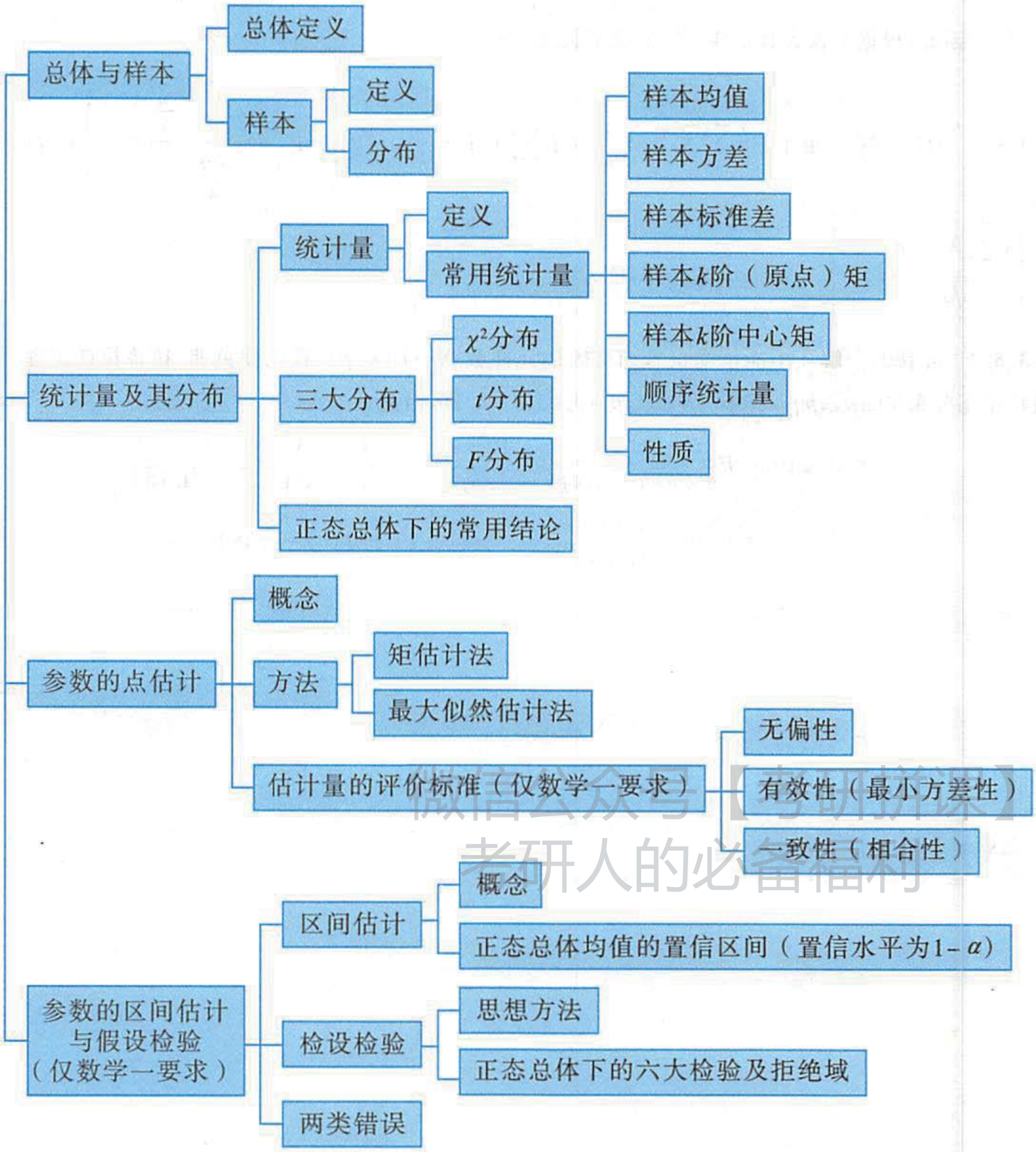
3.5.3 0.4890 **解** 不能承受试验而烧毁的元件数 $X\sim B(n,p)$. 根据棣莫弗-拉普拉斯定理，X 近似服从正态分布$N(np, npq)$，其中 $n=100$，$p=0.05$，$q=0.95$. 因此

$$P\{5\leqslant X\leqslant 10\} = P\left\{\frac{5-np}{\sqrt{npq}} \leqslant \frac{X-np}{\sqrt{npq}} \leqslant \frac{10-np}{\sqrt{npq}}\right\} = P\left\{0 \leqslant \frac{X-5}{\sqrt{4.75}} \leqslant \frac{10-5}{\sqrt{4.75}}\right\}$$

$$= P\left\{0 \leqslant \frac{X-5}{\sqrt{4.75}} \leqslant 2.29\right\} \approx \Phi(2.29) - \Phi(0) = 0.4890.$$

第6讲 数理统计

基础知识结构



- 总体与样本
  - 总体定义
  - 样本
    - 定义
    - 分布
- 统计量及其分布
  - 统计量
    - 定义
    - 常用统计量
      - 样本均值
      - 样本方差
      - 样本标准差
      - 样本k阶（原点）矩
      - 样本k阶中心矩
      - 顺序统计量
      - 性质
  - 三大分布
    - χ^2分布
    - t分布
    - F分布
  - 正态总体下的常用结论
- 参数的点估计
  - 概念
  - 方法
    - 矩估计法
    - 最大似然估计法
  - 估计量的评价标准（仅数学一要求）
    - 无偏性
    - 有效性（最小方差性）
    - 一致性（相合性）
- 参数的区间估计与假设检验（仅数学一要求）
  - 区间估计
    - 概念
    - 正态总体均值的置信区间（置信水平为$1-\alpha$）
  - 检设检验
    - 思想方法
    - 正态总体下的六大检验及拒绝域
  - 两类错误

基础内容精讲

一、总体与样本

1. 总体

研究对象的全体称为**总体**，组成总体的每一个元素称为**个体**. 在对总体进行统计研究时，我们所关心的是表征总体状况的某个(或某几个)数量指标 X(可以是向量)和该指标在总体中的分布情况. 我们把总体与随机变量 X 等同起来，说“总体 X”. 所谓总体的分布就是指随机变量 X 的分布.

2. 样本

n 个相互独立且与总体 X 具有相同概率分布的随机变量 $X_1,X_2,\cdots,X_n$ 所组成的整体$(X_1,X_2,\cdots,X_n)$称为来自总体 X，容量为 n 的一个**简单随机样本**，简称**样本**. 一次抽样结果的 n 个具体数值$(x_1,x_2,\cdots,x_n)$称为样本 $X_1,X_2,\cdots,X_n$ 的一个**观测值(或样本值)**.

3. 样本的分布

对于容量为 n 的样本 $X_1,X_2,\cdots,X_n$，有如下定理：

假设总体 X 的分布函数为 $F(x)$(概率密度为 $f(x)$，或概率分布为 $p_i=P\{X=x_i\}$)，则$(X_1,X_2,\cdots,X_n)$的分布函数为

$$F(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}F(x_i).$$

相应地：

(1)对于离散型随机变量的样本 $X_1,X_2,\cdots,X_n$，联合分布为

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}=\prod_{i=1}^{n}P\{X_i=x_i\};$$

(2)对于连续型随机变量的样本 $X_1,X_2,\cdots,X_n$，联合概率密度为

$$f(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i).$$

二、统计量及其分布

(一)统计量

设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个样本，$g(x_1,x_2,\cdots,x_n)$为 n 元函数，如果 g 中不含任何未知参数，则称$g(X_1,X_2,\cdots,X_n)$为样本 $X_1,X_2,\cdots,X_n$ 的一个**统计量**. 若$(x_1,x_2,\cdots,x_n)$为样本值，则称$g(x_1,x_2,\cdots,x_n)$为$g(X_1,X_2,\cdots,X_n)$的**观测值**.

【注】 (1)直观上，统计量就是由统计数据计算得来的量. 数学上，统计量是样本 $X_1,X_2,\cdots,X_n$ 的函数：$T=f(X_1,X_2,\cdots,X_n)$. 统计量不依赖于任何未知参数.

(2)作为随机样本的函数，统计量也是随机变量.

(二)常用统计量

样本数字特征和顺序统计量都是最常用的统计量. 统计量是统计分析和统计推断的重要工具.

1. 样本数字特征

数学期望(均值)、方差和标准差、矩等都是总体 X 最重要的数字特征. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本，则相应的样本数字特征定义：

(1)样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$；

(2)样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$；

样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$；

(3)样本 k 阶(原点)矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k(k=1,2,\cdots)$；

(4)样本 k 阶中心矩 $B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k(k=2,3,\cdots)$.

2. 顺序统计量

将样本 $X_1,X_2,\cdots,X_n$ 的 n 个观测量按其取值从小到大的顺序排列，得

$$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}.$$

随机变量 $X_{(k)}(k=1,2,\cdots,n)$ 称作第 k 顺序统计量，其中 $X_{(1)}$ 是最小顺序统计量，而 $X_{(n)}$ 是最大顺序统计量：

$$X_{(1)}=\min\{X_1,X_2,\cdots,X_n\},\quad X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}.$$

【注】 (1) $X_{(n)}=\max\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_{(n)}(x)=[F(x)]^n$$

(概率密度为 $f_{(n)}(x)=n[F(x)]^{n-1}f(x)$).

(2) $X_{(1)}=\min\{X_1,X_2,\cdots,X_n\}$ 的分布函数为

$$F_{(1)}(x)=1-[1-F(x)]^n$$

(概率密度为 $f_{(1)}(x)=n[1-F(x)]^{n-1}f(x)$).

3. 常用统计量的性质

设总体 X 的期望 $EX=\mu$，方差 $DX=\sigma^2$，$X_1,X_2,\cdots,X_n$ 是取自总体 X，容量为 n 的一个样本，$\overline{X},S^2$ 分别为样本的均值和方差，则

$$EX_i=\mu,\quad DX_i=\sigma^2(i=1,2,\cdots,n),\quad E\overline{X}=EX=\mu,\quad D\overline{X}=\frac{1}{n}DX=\frac{\sigma^2}{n},\quad E(S^2)=DX=\sigma^2.$$

(三)三大分布——χ^2 分布、t 分布和 F 分布

χ^2 分布、t 分布和 F 分布是统计推断中最常用的抽样分布. 读者不必记忆 χ^2 分布、t 分布和 F 分布的概率密度，只需了解相应变量的典型模式，以及它们的分布曲线的示意图和分位数，会查相应分位数的数值表.

1. χ^2 分布

(1)典型模式.

若随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且都服从标准正态分布，则随机变量 $X=\sum_{i=1}^{n}X_i^2$ 服从自由度为 n

的 χ^2 分布，记为 $X\sim\chi^2(n)$. 其概率密度 $f(x)$的图形如图3-6-1所示. 特别地，$X_i^2\sim\chi^2(1)$.

对给定的 $\alpha(0<\alpha<1)$，称满足

$$P\{\chi^2>\chi_\alpha^2(n)\}=\int_{\chi_\alpha^2(n)}^{+\infty} f(x)\mathrm{d}x=\alpha$$

的 $\chi_\alpha^2(n)$为 $\chi^2(n)$分布的**上 α 分位点**(如图 3-6-2). 对于不同的 $\alpha,n,\chi^2(n)$分布上 α 分位点可通过查表求得.

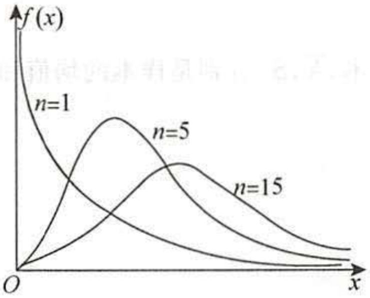


图 3-6-1

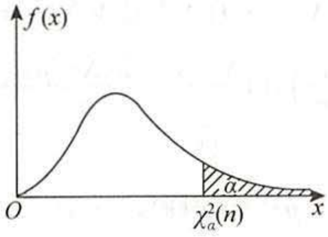


图 3-6-2

【注】 (1)自由度是指和式中独立变量的个数.

(2)某分布上 α 分位点(亦称上侧 α 分位数)为 μ_α 意指：点 μ_α 上侧(即右侧)，该概率密度曲线下方，x 轴上方图形面积为 α. 考研大纲中规定的便是上侧 α 分位数.

(2)χ^2 分布的性质.

①若 $X_1\sim\chi^2(n_1),X_2\sim\chi^2(n_2),X_1$ 与 X_2 相互独立，则 $X_1+X_2\sim\chi^2(n_1+n_2)$.

一般地，若 $X_i\sim\chi^2(n_i)(i=1,2,\cdots,m),X_1,X_2,\cdots,X_m$ 相互独立，则 $\sum\limits_{i=1}^{m}X_i\sim\chi^2\left(\sum\limits_{i=1}^{m}n_i\right)$.

②若 $X\sim\chi^2(n)$，则 $EX=n,DX=2n$.

2. t 分布

(1)典型模式.

设随机变量 $X\sim N(0,1),Y\sim\chi^2(n),X$ 与 Y 相互独立，则随机变量 $t=\dfrac{X}{\sqrt{Y/n}}$服从自由度为 n 的 t 分布，记为 $t\sim t(n)$.

t 分布的概率密度 $f(x)$的图形关于 $x=0$ 对称(如图 3-6-3)，因此 $Et=0(n\geqslant2)$.

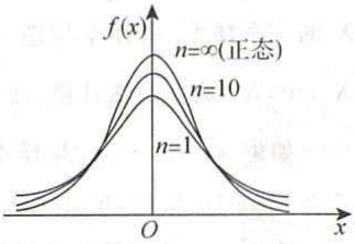


图 3-6-3

(2)t 分布的性质.

由 t 分布的概率密度 $f(x)$图形的对称性知

$$P\{t>-t_\alpha(n)\}=P\{t>t_{1-\alpha}(n)\},$$

故 $t_{1-\alpha}(n)=-t_\alpha(n)$.

3. F 分布

(1)典型模式.

设随机变量 $X\sim\chi^2(n_1),Y\sim\chi^2(n_2)$，且 X 与 Y 相互独立，则 $F=\dfrac{X/n_1}{Y/n_2}$服从自由度为(n_1,n_2)的 F 分布，记为 $F\sim F(n_1,n_2)$，其中 n_1 称为第一自由度，n_2 称为第二自由度. F 分布的概率密度 $f(x)$的图形如图 3-6-4 所示.

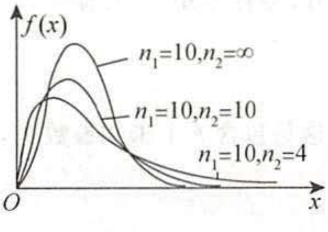


图 3-6-4

(2) F 分布的性质.

①若 $F\sim F(n_1,n_2)$，则 $\frac{1}{F}\sim F(n_2,n_1)$.

② $F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$.

（四）正态总体条件下的常用结论

设 $X_1,X_2,\cdots,X_n$ 是来自正态总体 $N(\mu,\sigma^2)$ 的一个样本，$\overline{X},S^2$ 分别是样本的均值和方差，则

(1) $\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$，即 $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\sim N(0,1)$；

(2) $\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$；

(3) $\frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)$（$\mu$ 未知时，在(2)中用 $\overline{X}$ 替代 μ）；

(4) $\overline{X}$ 与 S^2 相互独立，$\frac{\sqrt{n}(\overline{X}-\mu)}{S}\sim t(n-1)$（$\sigma$ 未知时，在(1)中用 S 替代 σ）. 进一步有

$$\frac{n(\overline{X}-\mu)^2}{S^2}\sim F(1,n-1).$$

三、参数的点估计

1. 概念

设总体 X 的分布函数为 $F(x;\theta)$（可以是多维的），其中 θ 是一个未知参数，$X_1,X_2,\cdots,X_n$ 是取自总体 X 的一个样本. 由样本构造一个适当的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 作为参数 θ 的估计，则称统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为 θ 的**估计量**，通常记为 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$.

如果 $x_1,x_2,\cdots,x_n$ 是样本的一个观察值，将其代入估计量 $\hat{\theta}$ 中得值 $\hat{\theta}(x_1,x_2,\cdots,x_n)$，并且此值作为未知参数 θ 的近似值，统计中称这个值为未知参数 θ 的**估计值**.

建立一个适当的统计量作为未知参数 θ 的估计量，并以相应的观察值作为未知参数估计值的问题，称为参数 θ 的**点估计问题**.

2. 方法

(1) 矩估计法.

设总体 X 分布中有 k 个未知参数 $\theta_1,\theta_2,\cdots,\theta_k$，$X_1,X_2,\cdots,X_n$ 是来自总体 X 的样本，如果 X 的原点矩 $E(X^l)(l=1,2,\cdots,k)$ 存在，即 $E(X^l)=\int_{-\infty}^{+\infty}x^l f(x;\theta_1,\cdots,\theta_k)\mathrm{d}x$ 或 $E(X^l)=\sum_i x_i^l P\{X=x_i;\theta_1,\cdots,\theta_k\}$ 存在. 令样本矩 = 总体矩，即

$$\frac{1}{n}\sum_{i=1}^{n}X_i^l=E(X^l)(l=1,2,\cdots,k),$$

这是包含 k 个未知参数 $\theta_1,\theta_2,\cdots,\theta_k$ 的 k 个联立方程组（称为矩法方程），由此解得

$$\hat{\theta}_l=\hat{\theta}_l(X_1,X_2,\cdots,X_n)(l=1,2,\cdots,k),$$

则 $\hat{\theta}_l$ 为 θ_l 的矩估计量，$\hat{\theta}_l(x_1,x_2,\cdots,x_n)$ 为 θ_l 的矩估计值.

【注】 矩估计法不要求总体服从什么分布，只要总体矩 $E(X^l)$ 存在即可. 计算 $E(X^l)$ 及解矩法方程 $\frac{1}{n}\sum_{i=1}^{n}X_i^l=E(X^l)$ 是求 θ 的矩估计量 $\hat{\theta}$ 的关键. 若这里 l 取多少没有明确规定的话，基于矩法方程求得的 θ 的估计量就可能不唯一了. 例如总体 X 服从参数为 λ 的泊松分布，λ 未知. 令 $\overline{X}=EX=\lambda$，得 λ 的矩估计量 $\hat{\lambda}_1=\overline{X}$；如果令 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=DX=\lambda$，得 λ 的另一个矩估计量 $\hat{\lambda}_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$. 为了使矩估计量有一个选定的标准，就需要约定或补充其他准则. **约定：用矩法方程求总体未知参数的估计量时，从低阶开始.**

(2)最大似然估计法.

①基本思想(最大似然原理).

对未知参数 θ 进行估计时，在该参数可能的取值范围 I 内选取，使“样本获此观测值 $x_1,x_2,\cdots,x_n$”的概率最大的参数值 $\hat{\theta}$ 作为 θ 的估计，这样选定的 $\hat{\theta}$ 最有利于 $x_1,x_2,\cdots,x_n$ 的出现.

(i)设总体 X 是离散型，其概率分布为 $P\{X=x\}=p(x;\theta),\theta\in I,\theta$ 为未知参数，$X_1,X_2,\cdots,X_n$ 为 X 的一个样本，则 $X_1,X_2,\cdots,X_n$ 取值为 $x_1,x_2,\cdots,x_n$ 的概率是

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}=\prod_{i=1}^{n}P\{X_i=x_i\}=\prod_{i=1}^{n}p(x_i;\theta),$$

显然这个概率值是 θ 的函数，将其记为

$$L(\theta)=L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}p(x_i;\theta),$$

称 $L(\theta)$ 为样本的**似然函数**. 若存在 $\hat{\theta}\in I$，使

$$L(x_1,x_2,\cdots,x_n;\hat{\theta})=\max_{\theta\in I}L(x_1,x_2,\cdots,x_n;\theta),$$

则称 $\hat{\theta}=\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为参数 θ 的**最大似然估计值**，而相应的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 称为参数 θ 的**最大似然估计量**.

(ii)同理，如果总体 X 是连续型随机变量，其概率密度为 $f(x;\theta),\theta\in I$，则样本的似然函数为

$$L(\theta)=L(x_1,x_2,\cdots,x_n;\theta)=\prod_{i=1}^{n}f(x_i;\theta).$$

若存在 $\hat{\theta}=\hat{\theta}(x_1,x_2,\cdots,x_n)\in I$，使

$$L(\hat{\theta})=\max_{\theta\in I}\prod_{i=1}^{n}f(x_i;\theta),$$

则称 $\hat{\theta}(x_1,x_2,\cdots,x_n)$ 为参数 θ 的**最大似然估计值**，而相应的统计量 $\hat{\theta}(X_1,X_2,\cdots,X_n)$ 称为参数 θ 的**最大似然估计量**.

②求最大似然估计量的步骤.

(i)写出样本的似然函数

$$L(\theta)=L(x_1,x_2,\cdots,x_n;\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n}p(x_i;\theta_1,\theta_2,\cdots,\theta_k)\text{ 或}\prod_{i=1}^{n}f(x_i;\theta_1,\theta_2,\cdots,\theta_k).$$

(ii)如果 $p(x;\theta_1,\theta_2,\cdots,\theta_k)$ 或 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$ 关于 $\theta_i(i=1,2,\cdots,k)$ 可微，则令

$$\frac{\partial L(\theta)}{\partial\theta_i}=0\text{ 或}\frac{\partial\ln L(\theta)}{\partial\theta_i}=0.$$

由于 $L(\theta)$ 是乘积形式，又 $\ln x$ 是 x 的单调增函数，因此 $L(\theta)$ 与 $\ln L(\theta)$ 在同一 θ 处取极值，所以更多的是采用解对数似然方程组的方法：$\frac{\partial \ln L(\theta)}{\partial \theta_i}=0$，求得 θ_i 的最大似然估计量

$$\hat{\theta}_i=\hat{\theta}_i(X_1,X_2,\cdots,X_n)(i=1,2,\cdots,k).$$

(iii)如果 $p(x;\theta_1,\theta_2,\cdots,\theta_k)$ 或 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$ 不可微，或似然方程组无解，则应由定义用其他方法求得 $\hat{\theta}_i$，例如当 $L(\theta)$ 为 θ 的单调增(或减)函数时，$\hat{\theta}$ 为 θ 取值上限(或下限).

【注】 求总体分布中未知参数 θ 的最大似然估计量必须知道总体的概率分布或密度. 写出样本的似然函数(或对数似然函数)，并求其最大值点是解题的关键.

③最大似然估计量的不变性原则.

设 $\hat{\theta}$ 是总体分布中未知参数 θ 的最大似然估计，函数 $u=u(\theta)$ 具有单值的反函数 $\theta=\theta(u)$，则 $\hat{u}=u(\hat{\theta})$ 是 $u(\theta)$ 的最大似然估计.

【注】 对于多个未知参数，不变性原理仍然成立.

3. 估计量的评价标准(仅数学一要求)

(1)无偏性.

若参数 θ 的估计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 对一切 n 及 $\theta\in I$，有 $E\hat{\theta}=\theta$，则称 $\hat{\theta}$ 为 θ 的**无偏估计量**.

(2)有效性(最小方差性).

设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 与 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$ 都是 θ 的无偏估计量，如果 $D(\hat{\theta}_1)<D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ **有效**.

(3)一致性(相合性).

设 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)$ 为未知参数 θ 的估计量，如果对任意 $\varepsilon>0$，有 $\lim\limits_{n\to\infty}P\{|\hat{\theta}-\theta|<\varepsilon\}=1$，即 $\hat{\theta}\xrightarrow{P}\theta$ $(n\to\infty)$，则称 $\hat{\theta}$ 为 θ 的**一致估计量(或相合估计量)**.

【注】 无偏性、有效性、一致性(相合性)是评价点估计量的一些基本标准，它们都是在某种意义下用于衡量估计量 $\hat{\theta}$ 与未知参数 θ 的"接近"程度，是从某一特定方面来评价其优良性的.

四、参数的区间估计(仅数学一要求)

1. 基本概念

设 θ 是总体 X 的一个未知参数，对于给定 $\alpha(0<\alpha<1)$，如果由样本 $X_1,X_2,\cdots,X_n$ 确定的两个统计量 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$，$\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)(\hat{\theta}_1<\hat{\theta}_2)$，使

$$P\{\hat{\theta}_1(X_1,X_2,\cdots,X_n)<\theta<\hat{\theta}_2(X_1,X_2,\cdots,X_n)\}=1-\alpha,$$

则称随机区间 $(\hat{\theta}_1,\hat{\theta}_2)$ 是 θ 置信度为 $1-\alpha$ 的**置信区间**，$\hat{\theta}_1$ 和 $\hat{\theta}_2$ 分别称为 θ 的置信度为 $1-\alpha$ 的双侧置信区间的**置信下限和置信上限**，$1-\alpha$ 称为**置信度或置信水平**，α 称为**显著性水平**.

给定置信度求未知参数置信区间的问题，称为**参数区间估计问题**.

【注】 置信区间的长度表示估计的精度，置信区间越短表示估计的精度越高.

2. 正态总体均值的置信区间（置信水平为 $1-\alpha$）

待估参数	其他参数	枢轴量的分布	置信区间
μ	σ^2 已知	$Z=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$	$\left(\overline{X}\pm\dfrac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$
μ	σ^2 未知	$t=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$	$\left(\overline{X}\pm\dfrac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$

五、假设检验（仅数学一要求）

1. 思想方法

关于总体（分布中的未知参数，分布的类型、特征、相关性、独立性…）的每一种论断（“看法”）称为**统计假设**. 然后根据样本观察数据或试验结果所提供的信息去推断（检验）这个“看法”（即假设）是否成立，这类统计推断问题称为**统计假设检验问题**，简称为**假设检验**. 如果总体分布函数 $F(x;\theta)$ 形式已知，但其中的参数 θ 未知，只涉及参数 θ 的各种统计假设称为**参数假设**. 如果一个统计假设完全确定总体的分布，则称这种假设为**简单假设**. 常常把着重考查、没有充分理由不能轻易否定的假设取为**原假设（基本假设或零假设）**，记为 H_0，将其否定的陈述（假设）称为**对立假设或备择假设**，记为 H_1. 对原假设 H_0 作出否定或不否定的推断，通常称为对 H_0 作**显著性检验**.

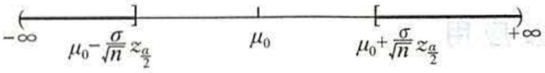
对这些假设进行检验的**基本思想是采用带有概率性质的反证法**，即“小概率原理”，也即“概率很接近于 0 的事件在一次试验或观察中认为它不会发生”，若发生了，则拒绝原假设 H_0. 小概率事件中“小概率”的值没有统一规定，通常是根据实际问题的要求，规定一个界限 $\alpha(0<\alpha<1)$，当一个事件的概率不大于 α 时，即认为它是小概率事件，在假设检验问题中，α 称为**显著性水平**，通常取 $\alpha=0.1,0.05,0.01$ 等.

在假设检验中，由拒绝原假设 H_0 的全体样本点所组成的集合 C 称为**否定域（或拒绝域）**，C 的补集 C^* 称为 H_0 的**接受域**.

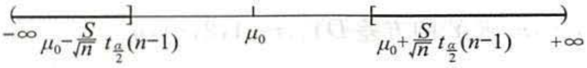
如果 H_0 的否定域形式为 $C=\{(x_1,x_2,\cdots,x_n)\mid T>\lambda_2$ 或 $T<\lambda_1\}$，即否定域位于接受域的两侧，则称这种检验为**双边检验**. 如果 H_0 的否定域形式为 $C=\{(x_1,x_2,\cdots,x_n)\mid T>\lambda\}$ 或者 $C=\{(x_1,x_2,\cdots,x_n)\mid T<\lambda\}$，即否定域位于接受域的一侧，右侧或左侧，称这种检验为**右边检验或左边检验**，统称为**单边检验**.

2. 正态总体下的六大检验及拒绝域

(1)σ^2 已知，μ 未知. $H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$.

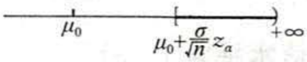


(2)σ^2 未知，μ 未知. $H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$.



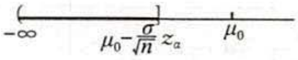
(3)σ^2 已知，μ 未知．$H_0:\mu\leqslant\mu_0$，$H_1:\mu>\mu_0$．

（或写 $\mu=\mu_0$）



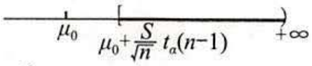
(4)σ^2 已知，μ 未知．$H_0:\mu\geqslant\mu_0$，$H_1:\mu<\mu_0$．

（或写 $\mu=\mu_0$）



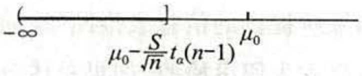
(5)σ^2 未知，μ 未知．$H_0:\mu\leqslant\mu_0$，$H_1:\mu>\mu_0$．

（或写 $\mu=\mu_0$）



(6)σ^2 未知，μ 未知．$H_0:\mu\geqslant\mu_0$，$H_1:\mu<\mu_0$．

（或写 $\mu=\mu_0$）



六、两类错误（仅数学一要求）

第一类错误（“弃真”）：若 H_0 为真，按检验法则，否定 H_0，此时犯了“弃真”的错误，这种错误称为第一类错误，犯第一类错误的概率为 $\alpha=P\{拒绝\ H_0\mid H_0\ 为真\}$．

第二类错误（“取伪”）：若 H_0 不真，按检验法则，接受 H_0，此时犯了“取伪”的错误，这种错误称为第二类错误，犯第二类错误的概率为 $\beta=P\{接受\ H_0\mid H_0\ 为假\}=P\{接受\ H_0\mid H_1\ 为真\}$．

【注】 犯两类错误的概率 α 与 β，并不满足 $\beta=1-\alpha$，在固定样本容量 n 的条件下，α 小，β 就大；β 小，α 就大．在实际应用中，我们总是在控制 α 的条件下，尽量使 β 小，这是因为人们常常把拒绝 H_0 比错误地接受 H_0 看得更重要些．

基础例题精解

一、统计量的概念、性质及应用

例 3.6.1 设 $X_1,X_2,\cdots,X_n(n>2)$ 为独立同分布的随机变量，且均服从正态分布 $N(0,1)$，记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$Y_i=X_i-\overline{X}$，$i=1,2,\cdots,n$．求 Y_i 的方差 DY_i，$i=1,2,\cdots,n$．

解 $DY_i=D(X_i-\overline{X})$

$$= D\left[\left(1-\frac{1}{n}\right)X_i - \frac{1}{n}\sum_{k\neq i}^{n} X_k\right] = \left(1-\frac{1}{n}\right)^2 + \frac{n-1}{n^2}$$

$$= \frac{n-1}{n}, i=1,2,\cdots,n.$$

例 3.6.2 (1)设总体 X 服从正态分布 $N(\mu,\sigma^2)$，其中 μ 已知，σ^2 未知. $X_1,X_2,\cdots,X_n$ 是取自总体 X 的简单随机样本，则下列样本函数中不是统计量的是(　　).

(A) $\frac{1}{n}\sum_{i=1}^{n} X_i$　　(B) $\max\limits_{1\leqslant i\leqslant n}\{X_i\}$

(C) $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$　　(D) $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

解 应选(C).

由统计量的定义“不含任何未知参数的样本函数”，即知不是统计量的选项应该是(C)，因为(C)中含有未知参数 σ，故选择(C).

(2)已知总体 X 的期望 $EX=0$，方差 $DX=\sigma^2$，从总体 X 中抽取容量为 n 的简单随机样本，其均值、方差分别为 $\overline{X}, S^2$. 记 $S_k^2=\frac{n}{k}\overline{X}^2+\frac{1}{k}S^2(k=1,2,3,4)$，则(　　).

(A) $E(S_1^2)=\sigma^2$　　(B) $E(S_2^2)=\sigma^2$

(C) $E(S_3^2)=\sigma^2$　　(D) $E(S_4^2)=\sigma^2$

解 应选(B).

这是一道计算性选择题，应用 $E\overline{X}=EX=0, D\overline{X}=\frac{\sigma^2}{n}, E(S^2)=\sigma^2$，通过计算 $E(S_k^2)$ 来确定正确选项.

由于 $E(\overline{X}^2)=D\overline{X}+(E\overline{X})^2=D\overline{X}=\frac{\sigma^2}{n}, E(S^2)=\sigma^2$，故

$$E(S_k^2)=\frac{n}{k}E(\overline{X}^2)+\frac{1}{k}E(S^2)=\frac{n}{k}\cdot\frac{\sigma^2}{n}+\frac{1}{k}\sigma^2=\frac{2}{k}\sigma^2.$$

当 $k=2$ 时，$E(S_2^2)=\sigma^2$，选择(B).

二、三大分布——χ^2 分布、t 分布和 F 分布及其应用

1. χ^2 分布

例 3.6.3 设 X_1,X_2,X_3,X_4 是来自正态总体 $N(0,2^2)$ 的简单随机样本，记

$$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2,$$

则当 $a=$________，$b=$________时，统计量 X 服从 χ^2 分布，其自由度为________. $(ab\neq 0)$

解 应填 $\frac{1}{20}$；$\frac{1}{100}$；2.

由条件知 X_1,X_2,X_3,X_4 相互独立且同服从正态分布 $N(0,2^2)$. 因此 X_1-2X_2 服从正态分布 $N(0,20)$，而 $3X_3-4X_4$ 服从正态分布 $N(0,100)$，并且相互独立. 由 χ^2 分布的典型模式知

$$T=\frac{(X_1-2X_2)^2}{20}+\frac{(3X_3-4X_4)^2}{100}$$

服从自由度为 2 的 χ^2 分布，从而 $a=\frac{1}{20}, b=\frac{1}{100}$ 时统计量 X 服从自由度为 2 的 χ^2 分布.

2. t 分布

例 3.6.4 假设总体 $X\sim N(0,3^2)$，$X_1,X_2,\cdots,X_8$ 是来自总体 X 的简单随机样本，则统计量

$$Y=\frac{X_1+X_2+X_3+X_4}{\sqrt{X_5^2+X_6^2+X_7^2+X_8^2}}$$

服从参数为________的________分布.

解 应填 4；t.

由于相互独立且服从正态分布的随机变量的线性组合仍然服从正态分布，易见

$$U=\frac{X_1+X_2+X_3+X_4}{\sqrt{D(X_1+X_2+X_3+X_4)}}$$
$$=\frac{X_1+X_2+X_3+X_4}{6}\sim N(0,1).$$

作为独立标准正态随机变量的平方和，

$$V=\frac{X_5^2}{9}+\frac{X_6^2}{9}+\frac{X_7^2}{9}+\frac{X_8^2}{9}$$

服从 χ^2 分布，自由度为 4. 随机变量 U 和 V 显然相互独立. 则随机变量 Y 可以表示为

$$Y=\frac{(X_1+X_2+X_3+X_4)\Big/6}{\sqrt{\frac{X_5^2+X_6^2+X_7^2+X_8^2}{9}\Big/4}}=\frac{U}{\sqrt{V/4}}.$$

由 t 分布的典型模式，可见随机变量 Y 服从自由度为 4 的 t 分布.

3. F 分布

例 3.6.5 设 $X_1,X_2,\cdots,X_{15}$ 是来自正态总体 $N(0,9)$ 的简单随机样本，则统计量

$$Y=\frac{1}{2}\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{X_{11}^2+X_{12}^2+\cdots+X_{15}^2}$$

服从参数为________的________分布.

解 应填(10,5)；F.

由 χ^2 分布的典型模式，知

$$\chi_1^2=\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{9}\text{和 }\chi_2^2=\frac{X_{11}^2+X_{12}^2+\cdots+X_{15}^2}{9}$$

服从自由度相应为 10 和 5 的 χ^2 分布，且相互独立. 从而，由 F 分布的典型模式，知

$$Y=\frac{1}{2}\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{X_{11}^2+X_{12}^2+\cdots+X_{15}^2}=\frac{\frac{\chi_1^2}{10}}{\frac{\chi_2^2}{5}}$$

服从自由度为(10,5)的 F 分布.

三、参数的点估计

1. 矩估计法

例 3.6.6 设来自总体 X 的简单随机样本 $X_1,X_2,\cdots,X_n$，总体 X 的概率分布为

$$X\sim\begin{pmatrix}-1 & 0 & 2\\ 2\theta & \theta & 1-3\theta\end{pmatrix},$$

其中 $0<\theta<\frac{1}{3}$. 试求未知参数 θ 的矩估计量.

解　总体 X 的数学期望为

$$EX=-2\theta+2(1-3\theta)=2-8\theta.$$

用样本均值 $\overline{X}$ 估计数学期望 EX，$\overline{X}=EX=2-8\theta$，得 θ 的矩估计量

$$\hat{\theta}=\frac{1}{8}(2-\overline{X}).$$

2. 最大似然估计法

例 3.6.7　设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta\left(0<\theta<\frac{1}{2}\right)$ 是未知参数，从总体 X 中抽取容量为 8 的一组样本，其样本值为 3，1，3，0，3，1，2，3. 求 θ 的矩估计值和最大似然估计值.

解　由于总体分布仅含一个未知参数，因此矩法方程为 $EX=\overline{X}$，其中

$$\overline{X}=\frac{1}{8}\sum_{i=1}^{8}X_i,$$

$$EX=0\times\theta^2+1\times2\theta(1-\theta)+2\theta^2+3(1-2\theta)=3-4\theta,$$

令 $3-4\theta=\overline{X}$，解得 θ 的矩估计量为 $\hat{\theta}=\frac{3-\overline{X}}{4}$，由样本值算得

$$\bar{x}=\frac{1}{8}\times(3+1+3+0+3+1+2+3)=2,$$

所以 θ 的矩估计值为 $\hat{\theta}=\frac{3-2}{4}=\frac{1}{4}$.

已知总体 X 的概率分布为 $p_i=P\{X=x_i\}=p(x_i;\theta)$，故样本值 $x_i(1\leqslant i\leqslant8)$ 的似然函数为

$$\begin{aligned}L(x_1,\cdots,x_8;\theta)&=\prod_{i=1}^{8}p(x_i;\theta)=\theta^2[2\theta(1-\theta)]^2\theta^2(1-2\theta)^4\\&=4\theta^6(1-\theta)^2(1-2\theta)^4.\end{aligned}$$

取对数，$\ln L(\theta)=\ln 4+6\ln\theta+2\ln(1-\theta)+4\ln(1-2\theta)$，

令

$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\frac{6}{\theta}-\frac{2}{1-\theta}-\frac{8}{1-2\theta}=\frac{24\theta^2-28\theta+6}{\theta(1-\theta)(1-2\theta)}=0,$$

即 $12\theta^2-14\theta+3=0$，解得 $\theta=\frac{7\pm\sqrt{13}}{12}$，因为 $\frac{7+\sqrt{13}}{12}>\frac{1}{2}$ 不合题意，故 θ 最大似然估计值为 $\hat{\theta}=\frac{7-\sqrt{13}}{12}$.

【注】　似然方程满足条件的解 $\hat{\theta}$ 一般就认为是最大似然估计而不加以验证. 原则上是需要证明的，然而有些问题证明是不容易的，甚至是不可能的. 如果能断言 $L(\theta)$ 有最大值点，而且似然方程只有唯一解 $\hat{\theta}$，则 $\hat{\theta}$ 为最大似然估计，此外有些问题可用微积分方法来验证，如本例. 事实上，由于

$$\frac{\mathrm{d}^2[\ln L(\theta)]}{\mathrm{d}\theta^2}=\frac{-6}{\theta^2}-\frac{2}{(1-\theta)^2}-\frac{16}{(1-2\theta)^2}<0,$$

所以 $\hat{\theta}=\frac{7-\sqrt{13}}{12}$ 使 $L(\theta)$ 达到最大，因而满足条件的解为 $\hat{\theta}=\frac{7-\sqrt{13}}{12}$.

3. 估计量的评价标准(仅数学一要求)

(1)无偏性.

例 3.6.8 设 $X_1,X_2,\cdots,X_n$是来自正态总体 $X\sim N(\mu,\sigma^2)$的简单随机样本,为使

$$D=k\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$$

成为总体方差 σ^2的无偏估计量,则 $k=$(　　).

(A)$\frac{1}{n-1}$　　(B)$\frac{1}{n}$　　(C)$\frac{1}{2(n-1)}$　　(D)$\frac{1}{2n}$

解　应选(C).

由条件知:$E(X^2)=\sigma^2+\mu^2$.假如统计量 D 是总体方差 σ^2的无偏估计量,则

$$\begin{aligned}ED&=k\sum_{i=1}^{n-1}E(X_{i+1}-X_i)^2=k\sum_{i=1}^{n-1}E(X_{i+1}^2+X_i^2-2X_iX_{i+1})\\&=k\sum_{i=1}^{n-1}(2\sigma^2+2\mu^2-2\mu^2)\\&=2k(n-1)\sigma^2=\sigma^2,\end{aligned}$$

于是,$k=\frac{1}{2(n-1)}$,故选(C).

例 3.6.9 设 $X_1,X_2,\cdots,X_n$是来自正态总体 $N(\mu,\sigma^2)$的一个简单随机样本,记

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,T=\overline{X}^2-\frac{1}{n}S^2,$$

证明:T 是 μ^2 的无偏估计量.

证明　由于

$$\begin{aligned}ET&=E\left(\overline{X}^2-\frac{1}{n}S^2\right)=E(\overline{X}^2)-\frac{1}{n}E(S^2)\\&=D\overline{X}+(E\overline{X})^2-\frac{1}{n}E(S^2)=\frac{\sigma^2}{n}+\mu^2-\frac{\sigma^2}{n}\\&=\mu^2,\end{aligned}$$

从而知,T 是 μ^2的无偏估计量.

(2)有效性(最小方差性).

例 3.6.10 设总体 $X\sim N(\mu,\sigma^2)$,X_1,X_2,X_3是来自 X 的简单随机样本,试证估计量

$$\hat{\mu}_1=\frac{1}{5}X_1+\frac{3}{10}X_2+\frac{1}{2}X_3,$$

$$\hat{\mu}_2=\frac{1}{3}X_1+\frac{1}{4}X_2+\frac{5}{12}X_3,$$

$$\hat{\mu}_3=\frac{1}{3}X_1+\frac{1}{6}X_2+\frac{1}{2}X_3$$

都是 μ 的无偏估计,并指出它们中哪一个最有效.

证明　因

$$E\hat{\mu}_1=\frac{1}{5}EX_1+\frac{3}{10}EX_2+\frac{1}{2}EX_3=\left(\frac{1}{5}+\frac{3}{10}+\frac{1}{2}\right)\mu=\mu,$$

$$E\hat{\mu}_2=\left(\frac{1}{3}+\frac{1}{4}+\frac{5}{12}\right)\mu=\mu,$$

$$E\hat{\mu}_3=\left(\frac{1}{3}+\frac{1}{6}+\frac{1}{2}\right)\mu=\mu,$$

故$\hat{\mu}_1,\hat{\mu}_2,\hat{\mu}_3$都是$\mu$的无偏估计.

又

$$D\hat{\mu}_1=\frac{1}{25}DX_1+\frac{9}{100}DX_2+\frac{1}{4}DX_3=\frac{38}{100}\sigma^2=0.38\sigma^2,$$

$$D\hat{\mu}_2=\left(\frac{1}{9}+\frac{1}{16}+\frac{25}{144}\right)\sigma^2=\frac{50}{144}\sigma^2=0.347\sigma^2,$$

$$D\hat{\mu}_3=\left(\frac{1}{9}+\frac{1}{36}+\frac{1}{4}\right)\sigma^2=\frac{14}{36}\sigma^2=0.389\sigma^2,$$

所以$\hat{\mu}_2$最有效.

(3)一致性(相合性).

例 3.6.11 设$X_1,X_2,\cdots,X_n$是来自总体X的一个简单随机样本,$EX=\mu,DX=\sigma^2$,证明统计量

$$Y=\frac{2}{n(n+1)}\sum_{i=1}^{n}iX_i$$

是μ的无偏相合估计量.

证明 由于

$$\begin{aligned}EY&=\frac{2}{n(n+1)}\sum_{i=1}^{n}E(iX_i)=\frac{2}{n(n+1)}\sum_{i=1}^{n}iEX_i\\&=\frac{2}{n(n+1)}(\mu+2\mu+\cdots+n\mu)\\&=\mu,\end{aligned}$$

从而知,Y是μ的无偏估计量.

又因

$$\begin{aligned}DY&=\left[\frac{2}{n(n+1)}\right]^2\sum_{i=1}^{n}i^2DX_i=\frac{4\sigma^2}{n^2(n+1)^2}(1^2+2^2+\cdots+n^2)\\&=\frac{4\sigma^2}{n^2(n+1)^2}\cdot\frac{1}{6}n(n+1)(2n+1)\\&=\frac{2}{3}\cdot\frac{2n+1}{n(n+1)}\sigma^2.\end{aligned}$$

于是由切比雪夫不等式

$$1\geqslant P\{|Y-\mu|<\varepsilon\}\geqslant1-\frac{DY}{\varepsilon^2}=1-\frac{2}{3}\cdot\frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2},$$

两边取极限,由$\lim\limits_{n\to\infty}1=\lim\limits_{n\to\infty}\left[1-\frac{2}{3}\cdot\frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2}\right]=1$,有

$$\lim_{n\to\infty}P\{|Y-\mu|<\varepsilon\}=1,$$

所以Y是μ的无偏相合估计量.

四、参数的区间估计(仅数学一要求)

例 3.6.12 设总体$X\sim N(\mu,\sigma^2)$(σ^2已知),则在给定样本容量n及置信度$1-\alpha$的情况下,未知参数μ的置信区间长度随着样本均值$\overline{X}$的增加而(　　).

(A)增加　　(B)减少

(C)不变　　(D)不能确定增或减

解 应选(C).

在给定条件下，未知参数的置信区间为$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$，其中$\Phi\left(z_{\frac{\alpha}{2}}\right)=1-\frac{\alpha}{2}$，可见该区间长度$2\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}$与$\overline{X}$无关，故选择(C).

例 3.6.13 设总体 X 的方差为 1，根据来自 X 的容量为 100 的简单随机样本，测得样本的均值为 5，则 X 的数学期望的置信度近似等于 0.95 的置信区间为________.

解 应填(4.804,5.196).

尽管 X 不是正态总体，但是由于样本容量 $n=100$ 属大样本，因此由中心极限定理易得

$$Z=\frac{\overline{X}-\mu}{1/\sqrt{100}}\overset{\text{近似}}{\sim}N(0,1),$$

查表易知，

$$P\{-1.96<Z<1.96\}=0.95,$$

故 μ 的置信度近似为 0.95 的置信区间为

$$\left(\overline{X}-\frac{1}{10}\times1.96,\overline{X}+\frac{1}{10}\times1.96\right),$$

将$\overline{X}=5$代入上式得置信区间为(4.804,5.196).

五、假设检验的两类错误(仅数学一要求)

例 3.6.14 在假设检验时，对于 $H_0:\mu=\mu_0$，$H_1:\mu\neq\mu_0$，则称(　　)为犯第一类错误.

(A)H_1真，接受 H_1　　(B)H_1不真，接受 H_1

(C)H_1真，拒绝 H_1　　(D)H_1不真，拒绝 H_1

解 应选(B).

犯第一类错误，即弃真错误，即 H_0真，但拒绝 H_0，也即 H_1不真，接受 H_1，故选择(B).

例 3.6.15 在假设检验中，显著性水平 α 的意义是(　　).

(A)原假设 H_0 成立，经检验被拒绝的概率

(B)原假设 H_0 成立，经检验被接受的概率

(C)原假设 H_0 不成立，经检验被拒绝的概率

(D)原假设 H_0 不成立，经检验被接受的概率

解 应选(A).

显著性水平 α 是确定小概率事件的一个界限，由检验准则知，$\alpha=P\{拒绝 H_0|H_0 为真\}$，所以正确选项是(A). 选项(B)所说的概率是 $P\{接受 H_0|H_0 为真\}=1-\alpha$；(D)是 $P\{接受 H_0|H_0 不成立\}=P\{犯第二类错误\}=\beta$；(C)是 $P\{拒绝 H_0|H_0 不成立\}=1-\beta$.

例 3.6.16 假定 X 是连续型随机变量，U 是对 X 的(一次)观测值；关于其概率密度 $f(x)$ 有如下假设：

$$H_0:f(x)=\begin{cases}\frac{1}{2}, & 0\leqslant x\leqslant2,\\ 0, & 其他,\end{cases}\qquad H_1:f(x)=\begin{cases}\frac{x}{2}, & 0\leqslant x\leqslant2,\\ 0, & 其他.\end{cases}$$

检验规则：当事件 $V=\left\{U>\frac{3}{2}\right\}$ 出现时，否定假设 H_0，接受 H_1. 则犯第一类错误的概率 $\alpha=$

________；犯第二类错误的概率 $\beta=$________.

解 应填$\frac{1}{4}$；$\frac{9}{16}$.

由 α 和 β 的意义，知

$$\alpha=P\left\{U>\frac{3}{2}\middle|H_0\right\}=\int_{\frac{3}{2}}^{2}\frac{1}{2}\mathrm{d}x=\frac{1}{4},$$

$$\beta=P\left\{U\leqslant\frac{3}{2}\middle|H_1\right\}=\int_{0}^{\frac{3}{2}}\frac{x}{2}\mathrm{d}x=\frac{9}{16}.$$

六、正态总体的假设检验（仅数学一要求）

例 3.6.17 已知某机器生产出的零件长度 X（单位：cm）服从正态分布 $N(\mu,\sigma^2)$，现从中随意抽取容量为 16 的一个样本，测得样本均值 $\overline{x}=10$，样本方差 $s^2=0.16$，$t_{0.025}(15)=2.132$.

(1)求总体均值 μ 置信水平为 0.95 的置信区间；

(2)在显著性水平为 0.05 下检验假设 $H_0:\mu=9.7$，$H_1:\mu\neq9.7$.

解 (1)在总体方差未知条件下，求均值 μ 的置信区间，即求置信水平为 $1-\alpha$ 的置信区间为$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$.

由样本值 $\overline{x}=10$，$s^2=0.16=0.4^2$，求得

$$\left(10-\frac{2.132\times0.4}{4},10+\frac{2.132\times0.4}{4}\right)=(9.7868,10.2132).$$

(2)在方差 σ^2 未知的条件下对总体均值进行检验.

①$H_0:\mu=9.7$，$H_1:\mu\neq9.7$.

②拒绝域为$\left(-\infty,\mu_0-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right]\cup\left[\mu_0+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),+\infty\right)$，代入数值，计算得$(-\infty,9.4868]\cup[9.9132,+\infty)$，由 $\overline{x}=10$，落入拒绝域，故否定 H_0，即认为 $\mu\neq9.7$.

基础习题精练

习题

3.6.1 设总体 X 与 Y 相互独立且都服从正态分布 $N(\mu,\sigma^2)$，$\overline{X}$，$\overline{Y}$ 是分别来自总体 X，Y，容量都为 n 的样本的均值，则当 n 固定时，概率 $P\{|\overline{X}-\overline{Y}|>\sigma\}$的值随 σ 的增大而(　　).

(A)单调增大　(B)单调减小　(C)保持不变　(D)增减不定

3.6.2 设 X_1,X_2,X_3,X_4 为来自总体 $N\left(0,\frac{1}{2}\right)$的简单随机样本，则统计量 $Y=\frac{X_1-X_2}{\sqrt{X_3^2+X_4^2}}$服从自由度为________的________分布.

3.6.3 设随机变量 X 服从自由度为(n_1,n_2)的 F 分布，则随机变量 $Y=\frac{1}{X}$服从参数为________的

________分布.

3.6.4 设总体 $X\sim N(\mu_1,4)$,$Y\sim N(\mu_2,5)$,X 与 Y 相互独立,$X_1,\cdots,X_8$ 和 $Y_1,\cdots,Y_{10}$ 是分别来自总体 X 和 Y 的两个样本,S_X^2 与 S_Y^2 分别为两个样本的方差,则(　　).

(A)$\frac{2S_X^2}{5S_Y^2}\sim F(7,9)$　(B)$\frac{5S_X^2}{2S_Y^2}\sim F(7,9)$　(C)$\frac{4S_X^2}{5S_Y^2}\sim F(7,9)$　(D)$\frac{5S_X^2}{4S_Y^2}\sim F(7,9)$

3.6.5 已知总体 X 的概率密度为

$$f(x)=\begin{cases}(1+\theta)x^{\theta}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

其中 $\theta>-1$ 是未知参数,设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本容量为 n 的简单随机样本,求 θ 的矩估计量和最大似然估计量.

3.6.6 设总体 X 的概率密度为 $f(x)=\begin{cases}\frac{2x}{\alpha^2}, & 0\leqslant x\leqslant\alpha,\\ 0, & \text{其他},\end{cases}$ 其中 $\alpha>1$ 未知,$X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本.

(1)求 $p=P\{0<X<\sqrt{\alpha}\}$;

(2)求 p 的最大似然估计量 $\hat{p}$.

3.6.7 设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}\frac{6x}{\theta^3}(\theta-x), & 0<x<\theta,\\ 0, & \text{其他},\end{cases}$$

$X_1,X_2,\cdots,X_n$ 是来自总体 X 的简单随机样本. 求 θ 的矩估计量$\hat{\theta}$,计算$\hat{\theta}$的方差 $D\hat{\theta}$,并讨论$\hat{\theta}$的无偏性、一致性(无偏性、一致性仅对数学一).

3.6.8 已知总体 X 的分布函数为

$$F(x;\alpha,\beta)=\begin{cases}1-\left(\frac{\alpha}{x}\right)^{\beta}, & \alpha\leqslant x,\\ 0, & \alpha>x,\end{cases}$$

其中参数 $\alpha\geqslant1,\beta>1,X_1,X_2,\cdots,X_n$ 为来自总体 X 的一个简单随机样本,求

(1)当 $\alpha=1$ 时,参数 β 的矩估计量;

(2)当 $\alpha=1$ 时,参数 β 的最大似然估计量;

(3)当 $\beta=2$ 时,参数 α 的最大似然估计量.

3.6.9(仅数学一) 一批零件的长度服从正态分布 $N(\mu,\sigma^2)$,其中 μ,σ^2 均未知. 现从中随机抽取 16 个零件,测得样本均值 $\overline{x}=20$ cm,样本标准差 $s=1$ cm,则 μ 的置信水平为 0.90 的置信区间为(　　).

(A)$\left(20-\frac{1}{4}t_{0.05}(16),20+\frac{1}{4}t_{0.05}(16)\right)$　(B)$\left(20-\frac{1}{4}t_{0.1}(16),20+\frac{1}{4}t_{0.1}(16)\right)$

(C)$\left(20-\frac{1}{4}t_{0.05}(15),20+\frac{1}{4}t_{0.05}(15)\right)$　(D)$\left(20-\frac{1}{4}t_{0.1}(15),20+\frac{1}{4}t_{0.1}(15)\right)$

解答

3.6.1 (C) 解　要计算概率需要知道 $\overline{X}-\overline{Y}$ 的分布. 依题意 $\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$,$\overline{Y}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$且相互

独立，故$\overline{X}-\overline{Y}\sim N\left(0,\frac{2\sigma^2}{n}\right)$，由此可知当 n 固定时，$P\{|\overline{X}-\overline{Y}|>\sigma\}=P\left\{\left|\frac{\overline{X}-\overline{Y}}{\sqrt{2}\sigma/\sqrt{n}}\right|>\sqrt{\frac{n}{2}}\right\}$与 σ 无关，故所求概率值保持不变，选择(C). 事实上，

$$P\{|\overline{X}-\overline{Y}|>\sigma\}=1-P\left\{\left|\frac{\overline{X}-\overline{Y}}{\sqrt{2}\sigma/\sqrt{n}}\right|\leqslant\sqrt{\frac{n}{2}}\right\}=1-\left[\Phi\left(\sqrt{\frac{n}{2}}\right)-\Phi\left(-\sqrt{\frac{n}{2}}\right)\right]$$
$$=2\left[1-\Phi\left(\sqrt{\frac{n}{2}}\right)\right].$$

3.6.2 2；t **解** 取 $X=X_1-X_2$，则 $EX=0,DX=DX_1+DX_2=1,X\sim N(0,1)$.

又 $E(\sqrt{2}X_i)=0,D(\sqrt{2}X_i)=2DX_i=1$，知$\sqrt{2}X_i\sim N(0,1)$，从而

$$Y_1=(\sqrt{2}X_3)^2+(\sqrt{2}X_4)^2\sim\chi^2(2),$$

因此 $Y=\dfrac{X}{\sqrt{\dfrac{Y_1}{2}}}\sim t(2)$，即 Y 服从自由度为 2 的 t 分布.

3.6.3 (n_2,n_1)；F **解** 因为服从自由度为(n_1,n_2)的 F 分布的随机变量 X 可以表示为

$$X=\frac{\dfrac{\chi_1^2}{n_1}}{\dfrac{\chi_2^2}{n_2}},\quad Y=\frac{1}{X}=\frac{\dfrac{\chi_2^2}{n_2}}{\dfrac{\chi_1^2}{n_1}},$$

其中 χ_1^2 和 χ_2^2 分别服从自由度为 n_1 和 n_2 的 χ^2 分布且相互独立. 由 F 分布的典型模式，知 Y 服从自由度为(n_2,n_1)的 F 分布.

3.6.4 (D) **解** 显然我们要用 $F(7,9)$典型模式来确定选项. 如果 $aQ\sim\chi^2(m)$，$bT\sim\chi^2(n)$，Q 与 T 独立，则$\frac{aQ/m}{bT/n}=\frac{naQ}{mbT}\sim F(m,n)$. 依题意

$$\frac{7S_X^2}{4}=\sum_{i=1}^{8}\left(\frac{X_i-\overline{X}}{2}\right)^2\sim\chi^2(7),\quad \frac{9S_Y^2}{5}=\sum_{i=1}^{10}\left(\frac{Y_i-\overline{Y}}{\sqrt{5}}\right)^2\sim\chi^2(9).$$

S_X^2 与 S_Y^2 相互独立，因此服从 $F(7,9)$分布的统计量 $A\frac{S_X^2}{S_Y^2}$，其系数应是 $A=\dfrac{\frac{7}{4}\Big/7}{\frac{9}{5}\Big/9}=\dfrac{5}{4}$，选择(D).

【注】 通过 F 分布典型模式的系数与自由度的关系确定选项.

3.6.5 **解** $\mu_1=EX=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_0^1(1+\theta)x^{\theta+1}\mathrm{d}x=\frac{1+\theta}{2+\theta}$，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$.

令$\mu_1=\overline{X}$，即$\frac{1+\theta}{2+\theta}=\overline{X}$，解得 θ 的矩估计量为$\hat{\theta}=\frac{2\overline{X}-1}{1-\overline{X}}$.

设 $x_1,x_2,\cdots,x_n$为样本值，则似然函数为

$$L(\theta)=(1+\theta)^n\prod_{i=1}^{n}x_i^{\theta}(0<x_i<1),$$

取对数得
$$\ln L(\theta)=n\ln(1+\theta)+\theta\sum_{i=1}^{n}\ln x_i,$$

令
$$\frac{\mathrm{d}\ln L(\theta)}{\mathrm{d}\theta}=\frac{n}{1+\theta}+\sum_{i=1}^{n}\ln x_i=0,$$

解得唯一解 $\hat{\theta}=-\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}-1$，则 θ 的最大似然估计量为 $\hat{\theta}=-\dfrac{n}{\sum\limits_{i=1}^{n}\ln X_i}-1$.

3.6.6 解 (1) $p=P\{0<X<\sqrt{\alpha}\}=\int_0^{\sqrt{\alpha}}f(x)\mathrm{d}x=\dfrac{1}{\alpha}$.

(2) 当 $0\leqslant x_1\leqslant\alpha,0\leqslant x_2\leqslant\alpha,\cdots,0\leqslant x_n\leqslant\alpha$ 时，似然函数为

$$L(\alpha)=f(x_1)f(x_2)\cdots f(x_n)=\frac{2^n}{\alpha^{2n}}(x_1x_2\cdots x_n),$$

显然 $L(\alpha)$ 关于 α 单调减少，且 $\alpha\geqslant\max\{x_1,x_2,\cdots,x_n\}$，则 α 的最大似然估计量为 $\hat{\alpha}=\max\{X_1,X_2,\cdots,X_n\}$.

又由(1)知 $p=\dfrac{1}{\alpha}$ 关于 α 是单调函数，根据最大似然估计的不变性，有 p 的最大似然估计量为 $\hat{p}=\dfrac{1}{\max\{X_1,X_2,\cdots,X_n\}}$.

3.6.7 解 由矩法方程 $EX=\overline{X}$，其中 $EX=\int_0^{\theta}\dfrac{6x^2(\theta-x)}{\theta^3}\mathrm{d}x=\dfrac{\theta}{2}$，即由 $\dfrac{\theta}{2}=\overline{X}$，解得 θ 的矩估计量 $\hat{\theta}=2\overline{X}$，且 $D\hat{\theta}=4D\overline{X}=\dfrac{4DX}{n}$，其中 $DX=E(X^2)-(EX)^2$.

$$E(X^2)=\int_0^{\theta}\frac{6x^3(\theta-x)}{\theta^3}\mathrm{d}x=\frac{6}{\theta^3}\left(\theta\frac{x^4}{4}-\frac{x^5}{5}\right)\bigg|_0^{\theta}=\frac{3}{10}\theta^2,$$

$$DX=\frac{3}{10}\theta^2-\frac{\theta^2}{4}=\frac{\theta^2}{20},$$

故 $D\hat{\theta}=\dfrac{4\theta^2}{20n}=\dfrac{\theta^2}{5n}$.

由于 $E\hat{\theta}=2E\overline{X}=2\cdot EX=2\cdot\dfrac{\theta}{2}=\theta$. 又由辛钦大数定律 $\overline{X}\xrightarrow{P}EX=\dfrac{\theta}{2}$，故 $\hat{\theta}=2\overline{X}\xrightarrow{P}2\cdot\dfrac{\theta}{2}=\theta$，所以 $\hat{\theta}=2\overline{X}$ 是 θ 的无偏估计、一致估计.

【注】 由于 $E\hat{\theta}=\theta,D\hat{\theta}=\dfrac{\theta^2}{5n}$，因此由切比雪夫不等式知：对任意 $\varepsilon>0$，有

$$0\leqslant P\{|\hat{\theta}-\theta|\geqslant\varepsilon\}=P\{|\hat{\theta}-E\hat{\theta}|\geqslant\varepsilon\}\leqslant\frac{D\hat{\theta}}{\varepsilon^2}=\frac{\theta^2}{5n\varepsilon^2}\to0(n\to\infty),$$

即 $\lim\limits_{n\to\infty}P\{|\hat{\theta}-\theta|\geqslant\varepsilon\}=0$，所以 $\hat{\theta}\xrightarrow{P}\theta$，$\hat{\theta}$ 为 θ 的一致估计.

3.6.8 解 (1) 当 $\alpha=1$ 时，X 的分布函数为

$$F(x;\beta)=\begin{cases}1-\dfrac{1}{x^{\beta}}, & 1\leqslant x,\\ 0, & 1>x,\end{cases}$$

则 X 的概率密度为

$$f(x;\beta)=\begin{cases}\dfrac{\beta}{x^{\beta+1}}, & 1<x,\\ 0, & \text{其他}.\end{cases}$$

由 $$\mu_1=EX=\int_{-\infty}^{+\infty}xf(x;\beta)\mathrm{d}x=\int_1^{+\infty}\beta x^{-\beta}\mathrm{d}x=\frac{\beta}{1-\beta}x^{1-\beta}\bigg|_1^{+\infty}=\frac{\beta}{\beta-1},$$

令 $\mu_1=\overline{X}$，由 $\dfrac{\beta}{\beta-1}=\overline{X}$，解得 β 的矩估计量为 $\hat{\beta}_1=\dfrac{\overline{X}}{\overline{X}-1}$.

(2)设 $x_1,x_2,\cdots,x_n$ 为样本值，则似然函数为

$$L(\beta)=\begin{cases}\dfrac{\beta^n}{(x_1x_2\cdots x_n)^{\beta+1}}, & 1<x_1,x_2,\cdots,x_n,\\ 0, & \text{其他},\end{cases}$$

当 $x_i>1(i=1,2,\cdots,n)$ 时，取对数得

$$\ln L(\beta)=n\ln\beta-(\beta+1)\sum_{i=1}^{n}\ln x_i,$$

令

$$\frac{\mathrm{d}\ln L(\beta)}{\mathrm{d}\beta}=\frac{n}{\beta}-\sum_{i=1}^{n}\ln x_i=0,$$

解得 $\hat{\beta}_2=\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}$，则 β 的最大似然估计量为 $\hat{\beta}_2=\dfrac{n}{\sum\limits_{i=1}^{n}\ln X_i}$.

(3)当 $\beta=2$ 时，X 的分布函数为 $F(x;\alpha)=\begin{cases}1-\dfrac{\alpha^2}{x^2}, & \alpha\leqslant x,\\ 0, & \alpha>x,\end{cases}$

则 X 的概率密度为

$$f(x;\alpha)=\begin{cases}\dfrac{2\alpha^2}{x^3}, & \alpha<x,\\ 0, & \text{其他}.\end{cases}$$

设 $x_1,x_2,\cdots,x_n$ 为样本值，则似然函数为

$$L(\alpha)=\begin{cases}\dfrac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}, & \alpha<x_1,x_2,\cdots,x_n,\\ 0, & \text{其他},\end{cases}$$

当 $x_i>\alpha(i=1,2,\cdots,n)$ 时，取对数得

$$\ln L(\alpha)=n\ln 2+2n\ln\alpha-3\sum_{i=1}^{n}\ln x_i,$$

则有

$$\frac{\mathrm{d}\ln L(\alpha)}{\mathrm{d}\alpha}=\frac{2n}{\alpha}>0,$$

故不能通过求导求出其最大似然估计量，可以直接从最大似然原理出发确定，即当 α 越大，则 $L(\alpha)$ 越大，总有

$$L(\alpha)=\frac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}\leqslant L(\min\{x_1,x_2,\cdots,x_n\})=\frac{2^n\min^{2n}\{x_1,x_2,\cdots,x_n\}}{(x_1x_2\cdots x_n)^3},$$

因此 $\hat{\alpha}=\min\{X_1,X_2,\cdots,X_n\}$ 为 α 的最大似然估计量.

【注】 在求最大似然估计量时可能出现似然方程无解的情况，但不表示题目无解，而是要根据最大似然原理，寻求其他定值方法，如利用单调性在参数取值区域的边界点取到最值.

3.6.9 (C) **解** σ^2 未知，均值 μ 的置信水平为 $1-\alpha$ 的置信区间为

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right),$$

其中，在给定条件下，$\overline{x}=20,s=1,t_{\frac{\alpha}{2}}(n-1)=t_{0.05}(15)$，则所求参数 μ 的置信水平为 0.90 的置信区间为

$$\left(20-\frac{1}{4}t_{0.05}(15),20+\frac{1}{4}t_{0.05}(15)\right),$$

故选择(C).

张宇

博士，全国著名考研数学辅导专家，教育部“国家精品课程建设骨干教师”，全国畅销书《张宇高等数学18讲》《张宇线性代数9讲》《张宇概率论与数理统计9讲》《张宇考研数学题源探析经典1000题》《张宇考研数学真题大全解》《张宇考研数学闭关修炼》《考研数学命题人终极预测8套卷》《张宇考研数学最后4套卷》作者，高等教育出版社原《全国硕士研究生入学统一考试数学考试大纲解析》一书编者之一，2007年斯洛文尼亚全球可持续发展大会受邀专家（发表15分钟主旨演讲），北京、上海、广州、西安等地全国著名考研数学辅导班首席主讲。

张宇数学教育系列丛书

○教材类

张宇考研数学基础30讲
张宇高等数学18讲
张宇线性代数9讲
张宇概率论与数理统计9讲

○题集类

张宇考研数学闭关修炼
张宇考研数学题源探析经典1000题（分数学一、数学二、数学三）
张宇考研数学真题大全解（分上、下册）（分数学一、数学二、数学三）
考研数学命题人终极预测8套卷（分数学一、数学二、数学三）
张宇考研数学最后4套卷（分数学一、数学二、数学三）

○教辅类

张宇带你学高等数学·同济七版（分上、下册）
张宇带你学线性代数·同济六版
张宇带你学概率论与数理统计·浙大四版

理工社网址：http://www.bitpress.com.cn

宇哥考研
新浪微博二维码

张宇考研数学
微信公众号

张宇考研数学
微信交流群

北京时代云图
官方微博

ISBN 978-7-5682-7466-1
9 787568 274661 >

定价：85.80元