2. 最小二乘求解： $\min f_0(x)=\|Ax-b\|_2^2$

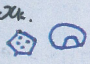
解：$\|x\|_2^2=x^Tx \Rightarrow \|Ax-b\|_2^2=(Ax-b)^T(Ax-b)$
$=x^TA^TAx-x^TA^Tb-b^TAx+b^Tb$
$=x^TA^TAx-2x^TA^Tb+b^Tb$. $\frac{\partial f_0}{\partial x}=2A^TAx-2A^Tb=0$
$\Rightarrow x^*=(A^TA)^{-1}A^Tb$. $\frac{\partial^2 f_0}{\partial x^2}=2A^TA$. $(A^TA\succeq 0)$
所以 x^* 即为其最小值.

3. 证明 $A^TA\succeq 0$： 利用 x^TA^TAx 来证明
$x^TA^TAx=\|Ax\|_2^2\geq 0$ 对于任意的 x，故 $A^TA\succeq 0$

4. Affine set 仿射集
① 通过点 x_1, x_2 的直线可以定义为下述形式：
$y=\theta x_1+(1-\theta)x_2$ $(\theta\in R)$.
过集合两点的直线也要在此集合中，即为仿射集.
② 仿射集的解：$\{x\mid Ax=b\}$
仿射集都可以用线性方程组的解来表示.

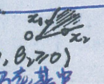
5. 线段：line segment
$y=\theta x_1+(1-\theta)x_2$ $0\leq\theta\leq 1$，x_1, x_2 为线段端点

6. 凸集：Convex set
集合内任意两点连接的线段仍在这个集合内.
$x_1, x_2\in C, 0\leq\theta\leq 1 \Rightarrow \theta x_1+(1-\theta)x_2\in C$.
① 凸组合：$x=\theta_1x_1+\theta_2x_2+\cdots+\theta_kx_k$.
$\theta_1+\theta_2+\cdots+\theta_k=1$. $\theta_i\geq 0$



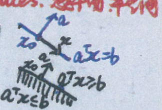
② 凸包：包含集合最小的凸集

7. 凸锥 Convex cone.



锥组合：$x=\theta_1x_1+\theta_2x_2$ $(\theta_1\geq 0, \theta_2\geq 0)$
凸锥即为所有点的锥组合仍在其中.

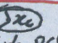
8. Hyperplanes and halfspaces. 超平面 半空间
① 超平面 $\{x\mid a^Tx=b\}$ $(a\neq 0)$
② 半空间 $\{x\mid a^Tx\leq b\}$ $(a\neq 0)$



关于超平面的解释说明：
$a^Tx=b$, $a^Tx_0=b$. $a^Tx-a^Tx_0=0 \Rightarrow a^T(x-x_0)=0$

9. Euclidean ball 欧式球. center x_c. radius r
$B(x_c, r)=\{x\mid \|x-x_c\|_2\leq r\}=\{x_c+ru\mid \|u\|_2\leq 1\}$
令 $x-x_c=w \Rightarrow \|w\|_2\leq r \Rightarrow \|\frac{w}{r}\|_2\leq 1 \Rightarrow \frac{w}{r}=u$

10. ellipsoid 椭圆球



$\{x\mid (x-x_c)^TP^{-1}(x-x_c)\leq 1\}$ with $P\in S_{++}^n$
示例：$\frac{x^2}{a^2}+\frac{y^2}{b^2}\leq 1 \Rightarrow \begin{bmatrix}x\\y\end{bmatrix}^T\begin{bmatrix}a^2&0\\0&b^2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}\leq 1$
另一种表示方法如下：$\{x_c+Au\mid \|u\|_2\leq 1\}$

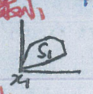
11. 范数球和范数锥. Norm ball and cones
① 范数球：$\{x\mid \|x-x_c\|\leq r\}$
② 范数锥：$\{(x, t)\mid \|x\|\leq t\}$
欧式范数锥 也叫二阶锥
范数球和范数锥都是凸的.

12. 多面体. Polyhedra $Ax\preceq b$, $Cx=d$.
结合半空间来理解：$Ax\preceq b \Rightarrow a_i^Tx\leq b_i$

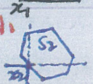
13. 半正定锥 Positive semidefinite cone
① symmetric matrices 对称矩阵
S^n: 对称矩阵. S_+^n: 对称半正定. S_{++}^n: 对称正定
② $S_+^n=\{X\in S^n\mid X\succeq 0\}$ $X\in S_+^n \Leftrightarrow z^TXz\geq 0$ $\forall z$
S_+^n is a convex cone 凸锥.
ex. $\begin{bmatrix}x&y\\y&z\end{bmatrix}\in S_+^2 \Leftrightarrow xz-y^2\geq 0$

14. 保持集合凸性的操作 ※
① 定义：$x_1, x_2\in C$, $0\leq\theta\leq 1 \Rightarrow \theta x_1+(1-\theta)x_2\in C$
② 操作：intersection | affine functions
perspective function | linear-fractional F
交集 仿射函数. 透视函数. 线性分段函数均是凸操作. (超平面, 半空间, 范数球…都是凸集)
① 仿射函数：$f(x)=Ax+b$ with $A\in R^{m\times n}$.
(1) $S\subseteq R^n$ convex $\Rightarrow f(S)=\{f(x)\mid x\in S\}$ convex
定义域是凸集，那么值域也是凸集.
(2) $S\subseteq R^m$ convex $\Rightarrow f^{-1}(C)=\{x\in R^n\mid f(x)\in C\}$
值域是凸集，定义域也是凸集
image 为值域、inverse image 为定义域.
例：hyperbolic cone. (双曲锥)
$\{x\mid x^TPx\leq (c^Tx)^2, c^Tx\geq 0\}$ (with $P\in S_+^n$)
② 透视函数：$P(x, t)=x/t$. 值域
定义域：$\mathrm{dom}\,P=\{(x, t)\mid t>0\}$
定义域和值域都是凸的.
③ 线性分段函数：$f(x)=\frac{Ax+b}{c^Tx+d}$.
$\mathrm{dom}\,f=\{x\mid c^Tx+d>0\}$
定义域和值域都是凸的.. (理解过程如下)
$P\circ f(x)=P(Ax+b, c^Tx+d)=\frac{Ax+b}{c^Tx+d}$
因为是仿射函数和透视函数的复合运算

15. minimum element 全局最小
$x\in S$ 关于 $\preceq_K$ 若是全局最小
$y\in S \Rightarrow x\preceq_K y$ 即 $y-x\in K$



minimal element 局部最小
$x\in S$ 关于 $\preceq_K$ 若是局部最小
$y\in S$. $y\preceq_K x \Rightarrow y=x$ 即 $x-y\in K$



16. Dual Cones 对偶锥
锥 K 的对偶锥为 K^*.
$K^*=\{y\mid y^Tx\geq 0$ for all $x\in K\}$.
例子：$K=R_+^n$; $K^*=R_+^n$. $K=S_+^n$. $K^*=S_+^n$
$K=\{(x, t)\mid \|x\|_2\leq t\}$: $K^*=\{(x, t)\mid \|x\|_2\leq t\}$
$K=\{(x, t)\mid \|x\|_1\leq t\}$: $K^*=\{(x, t)\mid \|x\|_\infty\leq t\}$

相关题目： ① 用归纳法证明凸组合也是凸集 (By induction 用归纳法)
illustrate the idea for $k=3$ (举例).
$y=\theta_1x_1+\theta_2x_2+\theta_3x_3\in C$. $(\theta_1+\theta_2+\theta_3=1$ and $\theta_1, \theta_2, \theta_3\geq 0$. So at least one of the θ_i is not equal to one. Assume $\theta_1\neq 1$
$\Rightarrow y=\theta_1x_1+(1-\theta_1)(\mu_2x_2+\mu_3x_3)$.
$\mu_2=\theta_2/(1-\theta_1)$ and $\mu_3=\theta_3/(1-\theta_1)$ $\mu_2, \mu_3\geq 0$
we know $\mu_1+\mu_2=\frac{\theta_2+\theta_3}{1-\theta_1}=1$. C is convex
$x_2, x_3\in C \Rightarrow \mu_2x_2+\mu_3x_3\in C$ 得证.

② 到a点的距离比到b点的距离近. (半空间.
$\{x\mid \|x-a\|_2\leq\|x-b\|_2\} \Rightarrow \|x-a\|_2^2\leq\|x-b\|_2^2$
$\Rightarrow (x-a)^T(x-a)\leq (x-b)^T(x-b)$
$\Rightarrow x^Tx-2a^Tx+a^Ta\leq x^Tx-2b^Tx+b^Tb$
$\Rightarrow 2(b-a)^Tx\leq b^Tb-a^Ta$ The set is indeed a halfspace. and when $\|x-a\|_2=\|x-b\|_2$ the point that are equidistant to a and b are given by a hyperplane whose normal is in the direction $b-a$

③ 证明凸集的另一种方法.
$C=\{x\in R^n\mid x^TAx+b^Tx+c\leq 0\}$ $A\in S^n$.
证明如果 $A\succeq 0$. 那么 C 是凸集
解：A set is convex if and only if its intersection with an arbitrary line $\{\hat{x}+tv\mid t\in R\}$ is convex.
故 $(\hat{x}+tv)^TA(\hat{x}+tv)+b^T(\hat{x}+tv)+c=\alpha t^2+\beta t+\gamma$
where. $\alpha=v^TAv$. $\beta=b^Tv+2\hat{x}^TAv$.
$\gamma=c+b^T\hat{x}+\hat{x}^TA\hat{x}$. then the intersection of C with the line defined by $\hat{x}$ and v is the set $\{\hat{x}+tv\mid \alpha t^2+\beta t+\gamma\leq 0\}$.
which is convex if $\alpha\geq 0$. This is true for any v, if $v^TAv\geq 0$ for all v, $A\succeq 0$
第二问：$H=\{x\mid g^Tx+h=0\}$ $x=\hat{x}+tv$
$\delta=g^Tv$. $\varepsilon=g^T\hat{x}+h$. Assume $\hat{x}\in H$. i.e. $\varepsilon=0$
The intersection of $C\cap H$ with the line defined by $\hat{x}$ and v is
$\{\hat{x}+tv\mid \alpha t^2+\beta t+\gamma\leq 0, \delta t=0\}$
$\delta=g^Tv\neq 0$. the intersection is the singleton $\{\hat{x}\}$. so it is a convex set. If $\delta=g^Tv=0$ the set reduces to $\{\hat{x}+tv\mid \alpha t^2+\beta t+\gamma\leq 0\}$ which is convex if $\alpha\geq 0$. therefore $C\cap H$ is convex if $g^Tv=0 \Rightarrow v^TAv\geq 0$
This is true if there exists λ such that $A+\lambda gg^T\succeq 0$. then $v^TAv=v^T(A+\lambda gg^T)v\geq 0$ for all v satisfying $g^Tv=0$.

④ Show set $\{x\in R_+^2\mid x_1x_2\geq 1\}$ is convex
consider a convex combination z of two points (x_1, x_2) and (y_1, y_2) in the set. If $x\succeq y$, then $z=\theta x+(1-\theta)y\succeq y$ and obviously $z_1z_2\geq y_1y_2\geq 1$. Similar proof if $y\succeq x$.

⑤ Show that $\{x\in R_+^n\mid \prod_{i=1}^n x_i\geq 1\}$ is convex
Hint. If $a, b\geq 0$ and $0\leq\theta\leq 1$, then $a^\theta b^{1-\theta}\leq\theta a+(1-\theta)b$.
Solution: Assume that $\prod_i x_i\geq 1$ and $\prod_i y_i\geq 1$. Using the inequality in the hint, we have
$\prod_i(\theta x_i+(1-\theta)y_i)\geq\prod x_i^\theta y_i^{1-\theta}=(\prod_i x_i)^\theta(\prod_i y_i)^{1-\theta}\geq 1$

⑥ $\{x\in R^n\mid \alpha\leq a^Tx\leq\beta\}$
is an intersection of two halfspaces. hence it is a convex set (and a polyhedron)
$\{x\in R^n\mid \alpha_i\leq x_i\leq\beta_i, i=1, \ldots, n\}$
is a convex set and a polyhedron because it is a finite intersection of halfspaces
$\{x\in R^n\mid a_1^Tx\leq b_1, a_2^Tx\leq b_2\}$
is an intersection of two halfspaces, so it is convex set. It is also a polyhedron. It is a cone if $b_1=0$ and $b_2=0$
$\{x\mid \|x-x_0\|_2\leq\|x-y\|_2$ for all $y\in S\}$, $S\subseteq R^n$
it is convex because it can be expressed as $\bigcap_{y\in S}\{x\mid \|x-x_0\|_2\leq\|x-y\|_2\}$ an intersection of halfspaces.
The set $\{x\mid x+S_2\subseteq S_1\}$. where $S_1, S_2\subseteq R^n$ with S_1 convex.
$x+y\in S_1$ for all $y\in S_2$. Therefore
$\{x\mid x+S_2\subseteq S_1\}=\bigcap_{y\in S_2}\{x\mid x+y\in S_1\}=\bigcap_{y\in S_2}(S_1-y)$
the intersection of convex sets S_1-y.
$\{x\mid \|x-a\|_2\leq\theta\|x-b\|_2\}$ Assume $a\neq b$ $0\leq\theta\leq 1$ The set is convex, in fact a ball. $\{x\mid \|x-a\|_2\leq\theta\|x-b\|_2\}$
$=\{x\mid \|x-a\|_2^2\leq\theta^2\|x-b\|_2^2\}$
$=\{x\mid (1-\theta^2)x^Tx-2(a-\theta^2b)^Tx+(a^Ta-\theta^2b^Tb)\leq 0\}$
If $\theta=1$, this is a halfspace. If $\theta<1$, it is a ball $\{x\mid (x-x_0)^T(x-x_0)\leq R^2\}$
with center x_0 and radius R given by
$x_0=\frac{a-\theta^2b}{1-\theta^2}$, $R=\left(\frac{\theta^2\|b\|_2^2-\|a\|_2^2}{1-\theta^2}-\|x_0\|_2^2\right)^{\frac{1}{2}}$

1. 求导 $f(x)=a^Tx$, $\frac{\partial f}{\partial x}=a$ | $f(x)=x^Ta$. $\frac{\partial f}{\partial x}=a$

1. 凸函数 $\mathrm{dom}\,f$ is convex set
定义域是凸集 $x, y\in\mathrm{dom}\,f$, $0\leq\theta\leq 1$
$f(\theta x+(1-\theta)y)\leq\theta f(x)+(1-\theta)f(y)$
strictly convex $x\neq y$, $0<\theta<1$
$f(\theta x+(1-\theta)y)<\theta f(x)+(1-\theta)f(y)$

2. 降维 Restriction of a Convex function to a line.
$f: R^n\to R$ is convex if and only if the function $g: R\to R$. $g(t)=f(x+tv)$
$\mathrm{dom}\,g=\{t\mid x+tv\in\mathrm{dom}\,f\}$ is convex in t for any $x\in\mathrm{dom}\,f$, $v\in R^n$

3. First-order condition 一阶条件.
f is differentiable. if $\mathrm{dom}\,f$ is open and the gradient $\nabla f(x)$ exists.
当 $\boxed{f(y)\geq f(x)+\nabla f(x)^T(y-x)}$ for all $x, y\in\mathrm{dom}\,f$
那么在凸集下. 函数 f 是凸函数

4. Second-order conditions 二阶条件
f 是凸函数, 当且仅当 $\boxed{\nabla^2 f(x)\succeq 0}$
for all $x\in\mathrm{dom}\,f$ (定义域是凸集)
If $\nabla^2 f(x)\succ 0$ for all $x\in\mathrm{dom}\,f$, then f is strictly convex.
例子：quadratic function:
$f(x)=(1/2)x^TPx+q^Tx+r$ (with $P\in S^n$)
$\nabla f(x)=Px+q$, $\nabla^2 f(x)=P$ convex if $P\succeq 0$
least-squares objective: $f(x)=\|Ax-b\|_2^2$
$\nabla f(x)=2A^T(Ax-b)$. $\nabla^2 f(x)=2A^TA$
convex for any A.
quadratic-over-linear: $f(x, y)=x^2/y$
$\nabla^2 f(x, y)=\frac{2}{y^3}\begin{bmatrix}y\\-x\end{bmatrix}\begin{bmatrix}y\\-x\end{bmatrix}^T\succeq 0$.
convex for $y>0$

$\begin{bmatrix}H+A^TQA&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x\\w\end{bmatrix}=\begin{bmatrix}-g-2A^TQAx+2A^TQb\\0\end{bmatrix}$
$\begin{bmatrix}H&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x\\w\end{bmatrix}=\begin{bmatrix}-g-2A^TQAx+2A^TQb\\0\end{bmatrix}$
$\begin{bmatrix}H&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x\\\hat{w}\end{bmatrix}=\begin{bmatrix}-g\\0\end{bmatrix}$, $\hat{w}=w+2QAx-2Qb$
$\begin{bmatrix}H&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x\\w\end{bmatrix}=\begin{bmatrix}-g\\0\end{bmatrix}$ $\Delta x^TH\Delta x=-g^T\Delta x$
$\hat{f}(x+\Delta x)=f(x)+g^Tv+\frac{1}{2}v^THv$
$=f(x)+\frac{1}{2}g^Tv$
$=f(x)-\frac{1}{2}\lambda(x)^2$
a

6. Jensen's inequality
basic inequality: if f is convex. then for $0\leq\theta\leq 1$. $f(\theta x+(1-\theta)y)\leq\theta f(x)+(1-\theta)f(y)$
例题：$f: R\to R$ is convex, and $a, b\in\mathrm{dom}\,f$ with $a<b$. 证明 $f(x)\leq\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b)$ for all $x\in[a, b]$
Solution: This is Jensen's inequality with $\lambda=(b-x)/(b-a)$
$f(x)=f(\lambda a+(1-\lambda)b)\leq\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b)$

穿插一个小例子：$\{x\mid \mathrm{dist}(x, S)\leq\mathrm{dist}(x, T)\}$. where $S, T\subseteq R^n$. and
$\mathrm{dist}(x, S)=\inf\{\|x-z\|_2\mid z\in S\}$
解：In general this set is not convex, as the following example in R shows. With $S=\{-1, 1\}$ and $T=\{0\}$. we have
$\{x\mid \mathrm{dist}(x, S)\leq\mathrm{dist}(x, T)\}=\{x\in R\mid x\leq -1/2$ or $x\geq\frac{1}{2}\}$
which clearly is not convex.

7. Operations that preserve convexity.
简单的凸函数经过保持凸性的操作后，仍然还是凸函数
1. nonnegative weighted sum
2. composition with affine function
3. pointwise maximum and supremum
4. composition 5. minimization
6. perspective. 非负加权和、仿射函数的组合、逐点取最大和取上确界、合成、最小和投影都可以保持凸性.

8. composition with affine function
$f(Ax+b)$ is convex if f is convex
(any) norm of affine function
$f(x)=\|Ax+b\|$ is convex
If $f_1, \ldots, f_m$ are convex, then $f(x)=\max\{f_1(x), \ldots, f_m(x)\}$ is convex. 逐点取最大
例子：piecewise-linear function:
$f(x)=\max_{i=1,\ldots,m}(a_i^Tx+b_i)$ is convex

9. Pointwise supremum 上确界 (最小上界)
If $f(x, y)$ is convex in x for each $y\in A$, then $g(x)=\sup_{y\in A}f(x, y)$ is convex.

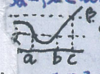
10. Composition with scaler functions
标量复合 (纯量函数的复合) $f(x)=h(g(x))$
composition of $g: R^n\to R$ and $h: R\to R$:
f is convex if $\begin{cases}g\text{ 凸}, h\text{ 凸}, \tilde{h}\text{ 非减函数},\\ g\text{ 凹}, h\text{ 凸}, \tilde{h}\text{ 非增函数}\end{cases}$
Proof (for $n=1$, differentiable g, h)
$f''(x)=h''(g(x))g'(x)^2+h'(g(x))g''(x)\geq 0$
Ex: $\exp g(x)$ is convex if g is convex
$1/g(x)$ is convex if g is concave and positive
Vector composition:
$f''(x)=g'(x)^T\nabla^2h(g(x))g'(x)+\nabla h(g(x))^Tg''(x)$

11. Minimization 下确界 (最大下界)
If $f(x, y)$ is convex in (x, y) and C is a convex set, then $g(x)=\inf_{y\in C}f(x, y)$ is convex
应用：Schur complement 舒尔补
$f(x, y)=x^TAx+2x^TBy+y^TCy$ with $\begin{bmatrix}A&B\\B^T&C\end{bmatrix}\succeq 0, C\succ 0$
minimizing over y gives $g(x)=\inf_y f(x, y)$
$=x^T(A-BC^{-1}B^T)x$, g is convex, hence Schur complement $A-BC^{-1}B^T\succeq 0$.
证明：$\nabla f_y=2B^Tx+2Cy=0 \Rightarrow y=-C^{-1}B^Tx$
$\Rightarrow g(x)=x^T(A-BC^{-1}B^T)x$

12. Perspective 透视函数
The perspective of a function $f: R^n\to R$ is the function $g: R^n\times R\to R$. $g(x, t)=tf(x/t)$.
$\mathrm{dom}\,g=\{(x, t)\mid x/t\in\mathrm{dom}\,f, t>0\}$
g is convex if f is convex

13. The conjugate function (共轭函数).
the conjugate of a function f is.
$f^*(y)=\sup_{x\in\mathrm{dom}\,f}(y^Tx-f(x))$ | f^* is convex (even if f is not
例题：$f(x)=-\log x$. $f^*(y)=\sup_{x>0}(xy+\log x)$
$=\begin{cases}-1-\log(-y), & y<0\\ \infty, & \text{otherwise}\end{cases}$ | $(xy+\log x)'=y+\frac{1}{x}=0 \Rightarrow x=-1/y>0$
$(xy+\log x)''=-1/x^2<0$. 故 $x=-1/y$ 是极大值点.
例题：$f(x)=(1/2)x^TQx$ with $Q\in S_{++}^n$
$f^*(y)=\sup_x(y^Tx-(1/2)x^TQx)=\frac{1}{2}y^TQ^{-1}y$
$x=Q^{-1}y$ 是其极大值点.

14. Quasiconvex functions (准凸函数).
$f: R^n\to R$ is quasiconvex if $\mathrm{dom}\,f$ is convex and the sublevel sets
$S_\alpha=\{x\in\mathrm{dom}\,f\mid f(x)\leq\alpha\}$ are convex for all α.



① f is quasiconcave if $-f$ is quasiconvex
② f is quasilinear if it is quasiconvex and quasiconcave
两个准凸函数的和不一定是准凸函数

15. 准凸函数的性质 Properties.
① modified Jensen inequality: for 准凸 f
$0\leq\theta\leq 1 \Rightarrow f(\theta x+(1-\theta)y)\leq\max\{f(x), f(y)\}$
② 一阶条件：可微的 f 在定义域是凸集的情况下，如果是准凸函数当且仅当：
$f(y)\leq f(x) \Rightarrow \nabla f(x)^T(y-x)\leq 0$

16. Log-concave and log-convex functions
a positive function f is log-concave if $\log f$ is concave: $f(x)=e^{\log f(x)}$
即：$f(\theta x+(1-\theta)y)\geq f(x)^\theta\cdot f(y)^{1-\theta}$ for $0\leq\theta\leq 1$
f is log-convex if $\log f$ is convex.
$\log f(\theta x+(1-\theta)y)\geq\theta\log f(x)+(1-\theta)\log f(y)$
性质：twice differentiable f with convex domain is log-concave if and only if
$f(x)\nabla^2 f(x)\preceq\nabla f(x)\nabla f(x)^T$. for all $x\in\mathrm{dom}\,f$

$f(\alpha x+\beta y)=\alpha f(x)+\beta f(y)$ $f(X)=Xz$
$f(\alpha X+\beta Y)=(\alpha X+\beta Y)z=\alpha Xz+\beta Yz=\alpha f(X)+\beta f(Y)$

$z^TX^2z=\|Xz\|_2^2$ is convex (因为是范数的平方)
$z^T(\theta X+(1-\theta)Y)^2z\leq\theta z^TX^2z+(1-\theta)z^TY^2z$
for $X, Y\in S^m$, $0\leq\theta\leq 1$. $\Rightarrow$
$(\theta X+(1-\theta)Y)^2\preceq\theta X^2+(1-\theta)Y^2$

相关题目： ① 证明 $F(x)=\frac{1}{x}\int_0^x f(t)dt$ 是凸的.
$F'(x)=-(1/x^2)\int_0^x f(t)dt+f(x)/x$
$F''(x)=(2/x^3)\int_0^x f(t)dt-2f(x)/x^2+f'(x)/x$
$=(2/x^3)\int_0^x(f(t)-f(x)-f'(x)(t-x))dt$
因为 f 是凸函数 Convexity now follows from the fact that: $f(t)\geq f(x)+f'(x)(t-x)$ for all $x, t\in\mathrm{dom}\,f$. $\Rightarrow F''(x)\geq 0$

② 利用一阶条件来证明二阶条件. 即 $\nabla^2 f(x)\geq 0$
解：一阶条件：$f'(x)(y-x)\leq f(y)-f(x)\leq f'(y)(y-x)$
假设 $y>x \Rightarrow \frac{f'(y)-f'(x)}{y-x}\geq 0 \Rightarrow f''(x)\geq 0$
若假设 $f''(z)\geq 0$，并且 $x<y$.
$\Rightarrow 0\leq\int_x^y f''(z)(y-z)dz=(f'(z)(y-z))\Big|_{z=x}^{z=y}+\int_x^y f'(z)dz$
$=-f'(x)(y-x)+f(y)-f(x) \Rightarrow$
$f(y)\geq f(x)+f'(x)(y-x)$. this shows f 是凸的
To generalize to $n>1$, $g(t)=f(x_0+tv)$ 是凸的当且仅当 $g''(t)=v^T\nabla^2f(x_0+tv)v\geq 0 \Rightarrow \nabla^2 f(x)\geq 0$
③ For $0<\alpha\leq 1$, $u_\alpha(x)=(x^\alpha-1)/\alpha$, $\mathrm{dom}\,u_\alpha=R_+$
$u_0(x)=\log x$ (with $\mathrm{dom}\,u_0=R_{++}$).

解: $\lim_{\alpha\to 0} u_\alpha(x) = \lim_{\alpha\to 0} \frac{(d/d\alpha)(x^\alpha-1)}{(d/d\alpha)\alpha} = \lim_{\alpha\to 0}\frac{x^\alpha \log x}{1}$
$=\log x$. 证明 u_α 是凹函数, 单调递增且 $u_\alpha(1)=0$

解: $u_\alpha(1)=\frac{1^\alpha-1}{\alpha}=0$, 且 $u'_\alpha(x)=x^{\alpha-1}$.
当 $0<\alpha<1$ 时, 其值是正的. $u''_\alpha(x)=(\alpha-1)x^{\alpha-2}$
$\Rightarrow u''_\alpha(x)<0 \Rightarrow u_\alpha$ is strictly concave

⑭ (a) $f(X)=\mathrm{tr}(X^{-1})$ is convex on $\mathrm{dom} f=S^n_{++}$

解: Define $g(t)=f(Z+tV)$, where $Z\succ 0$ and $V\in S^n$
$g(t)=\mathrm{tr}((Z+tV)^{-1})=\mathrm{tr}(Z^{-1}(I+tZ^{-1/2}VZ^{-1/2})^{-1})$
$=\mathrm{tr}(Z^{-1}Q(I+t\Lambda)^{-1}Q^T)=\mathrm{tr}(Q^TZ^{-1}Q(I+t\Lambda)^{-1})$
$=\sum_{i=1}^n (Q^TZ^{-1}Q)_{ii}(1+t\lambda_i)^{-1}$, we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2}=Q\Lambda Q^T$. In the last equality we express g as a positive weighted sum of convex functions $1/(1+t\lambda_i)$. 因此是凸

(b) $f(X)=(\det X)^{1/n}$ is concave on $\mathrm{dom} f=S^n_{++}$

解: Define $g(t)=f(Z+tV)$, where $Z\succ 0$ and $V\in S^n$
$g(t)=(\det(Z+tV))^{1/n}$
$=(\det Z^{1/2}\det(I+tZ^{-1/2}VZ^{-1/2})\det Z^{1/2})^{1/n}$
$=(\det Z)^{1/n}\left(\prod_{i=1}^n(1+t\lambda_i)\right)^{1/n}$ where $\lambda_i, i=1,\cdots,n$, are the eigenvalues of $Z^{-1/2}VZ^{-1/2}$. From the last equality we see that g is a concave function of t on $\{t \mid Z+tV\succ 0\}$, since $\det Z>0$ and the geometric mean $\left(\prod_{i=1}^n x_i\right)^{1/n}$ is concave on R^n_{++}.

1. 优化问题的标准形式

minimize $f_0(x)$
subject to $f_i(x)\le 0$, $i=1,\cdots,m$
$h_i(x)=0$, $i=1,\cdots,p$

最优值:
$p^*=\inf\{f_0(x) \mid f_i(x)\le 0, i=1,\cdots,m, h_i(x)=0, i=1,\cdots,p\}$

2. 凸优化问题 (标准形式)

minimize $f_0(x)$ subject to $f_i(x)\le 0$, $a_i^Tx=b_i$ $\Leftrightarrow$ minimize $f_0(x)$ subject to $f_i(x)\le 0$, $Ax=b$

f_0 和 f_i 都是凸的, 若 f_0 为准凸则变为准凸优化
凸优化问题的局部最优解均是全局最优解

3. Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and $\nabla f_0(x)^T(y-x)\ge 0$ for all feasible y
If nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x.

4. 1.无约束问题 若 x 是最优值, 那么 $x\in\mathrm{dom} f_0$, $\nabla f_0(x)=0$
2.等式约束问题: minimize $f_0(x)$ s.t. $Ax=b$.
$\Leftrightarrow x\in\mathrm{dom} f_0$, $Ax=b$, $\nabla f_0(x)+A^T\nu=0$
3.非负象限最小化. $\min f_0(x)$. s.t. $x\ge 0$.
x 是最优解, $x\in\mathrm{dom} f_0$, $x\ge 0$, 可以得出
$\nabla f_0(x)_i\ge 0$, $x_i=0$ | $\nabla f_0(x)_i=0$, $x_i>0$

5. 消除等式约束. $\min f_0(x)$. s.t. $f_i(x)\le 0$, $Ax=b$
$Ax=b \Rightarrow x=x_0+Fz$. 故原优化问题转化成.
$\min f_0(Fz+x_0)$ s.t. $f_i(Fz+x_0)\le 0$.
引入等式约束: $\min f_0(A_0x+b_0)$ s.t. $f_i(A_ix+b_i)\le 0$
$\Leftrightarrow \min f_0(y_0)$ s.t. $f_i(y_i)\le 0$, $y_i=A_ix+b_i$
对于线性不等式引入松弛变量: $\min f_0(x)$ s.t. $a_i^Tx\le b_i$
$\Leftrightarrow \min f_0(x)$ s.t. $a_i^Tx+s_i=b_i$, $s_i\ge 0$
epigraph form: $\min t$ s.t. $f_0(x)-t\le 0$, $f_i(x)\le 0$, $Ax=b$

最小化一些变量: $\min f_0(x_1,x_2)$ s.t. $f_i(x_1)\le 0$
$\Leftrightarrow \min \tilde f_0(x_1)$. s.t. $f_i(x_1)\le 0$. $\tilde f_0(x_1)=\inf f_0(x_1,x_2)$

$(\theta x_1+(1-\theta)y_1)(\theta x_2+(1-\theta)y_2)$
$=\theta^2x_1x_2+(1-\theta^2)y_1y_2+\theta(1-\theta)x_1y_2+\theta(1-\theta)x_2y_1$
$=\theta x_1x_2+(1-\theta)y_1y_2-\theta(1-\theta)(y_1-x_1)(y_2-x_2)$
≥ 1

$\delta J(x(t),\delta x)=\int_{t_0}^{t_f}\left[g_x-\frac{d}{dt}g_{\dot x}\right]\delta x(t)dt$
$+g_{\dot x}(x(t_f),\dot x(t_f),t_f)\,\delta x(t_f)$
the $+[g_{\dot x}]\Big|_{t_0}^{t_f}$

$\Leftrightarrow \max r$ s.t. $a_i^Tx_c+r\|a_i\|_2\le b_i$
利用公式: $|a^Tb|\le\|a\|_2\cdot\|b\|_2$

Linear-fractional program 线性分式规划
$\min f_0(x)$ s.t. $Gx\le h$, $Ax=b$
$f_0(x)=\frac{c^Tx+d}{e^Tx+f}$, $\mathrm{dom} f_0(x)=\{x \mid e^Tx+f>0\}$.
令 $y=x/e^Tx+f$, $z=1/e^Tx+f$ $\Leftrightarrow$
$\min c^Ty+dz$ s.t. $Gy\le hz$, $Ay=bz$.
$e^Ty+fz=1$, $z\ge 0$.

Quadratic program (QP)
$\min (1/2)x^TPx+q^Tx+r$.
s.t. $Gx\le h$, $Ax=b$.
$P\in S^n_+$, so objective is convex quadratic

二次约束二次规划. (QCQP)
$\min (1/2)x^TP_0x+q_0^Tx+r_0$
s.t. $(1/2)x^TP_ix+q_i^Tx+r_i\le 0$, $Ax=b$. quadratic
$P_i\in S^n_+$, objective and constraints are convex.

二阶锥规划 (SOCP)
$\min f^Tx$, s.t. $\|A_ix+b_i\|_2\le c_i^Tx+d_i$, $Fx=g$

Geometric programming 几何规划.
monomial function 单项式
$f(x)=cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$. $\mathrm{dom} f=R^n_{++}$, $c>0$.
poynomial function 多项式 (即单项式相加)
$f(x)=\sum_{k=1}^K c_kx_1^{a_{1k}}x_2^{a_{2k}}\cdots x_n^{a_{nk}}$, $\mathrm{dom} f=R^n_{++}$
GP: $\min f_0(x)$. s.t. $f_i(x)\le 1$. $h_i(x)=1$
with f_i 多项式. h_i 是单项式

Semidefinite program (SDP)
$\min c^Tx$. s.t. $x_1F_1+x_2F_2+\cdots+x_nF_n+G\le 0$, $Ax=b$

舒尔补: $\begin{bmatrix}A&B\\C&D\end{bmatrix}>0 \Rightarrow A-B^TC^{-1}B\ge 0$
$\Rightarrow C-B^TA^{-1}B\ge 0$.

证明: $x>0$, $u_0(x)=\lim_{\alpha\to 0}u_\alpha(x)$

LP: $\min c^Tx$ s.t. $Ax\le b$ $\Leftrightarrow$ SDP: s.t. $\mathrm{diag}(Ax-b)\le 0$

Eigenvalue minimization: $\min \lambda_{\max}(A(x))$
$\Leftrightarrow$ SDP: $\min t$. s.t. $A(x)\le tI$
$\lambda_{\max}(A)\le t \Leftrightarrow A\le tI$.

Matrix norm minimization:
$\min \|A(x)\|_2=(\lambda_{\max}(A(x)^TA(x)))^{1/2}$
$\Leftrightarrow \min t$. s.t. $\begin{bmatrix}tI&A(x)\\A(x)^T&tI\end{bmatrix}\ge 0$. 即通过 $\|A\|_2\le t$
$\Leftrightarrow A^TA\le t^2I$, $t\ge 0$. $\Leftrightarrow \begin{bmatrix}tI&A(x)\\A(x)^T&tI\end{bmatrix}\ge 0$

① Lagrangian 拉格朗日

标准形式: $\min f_0(x)$. s.t. $f_i(x)\le 0$, $h_i(x)=0$
$L(x,\lambda,\nu)=f_0(x)+\sum_{i=1}^m\lambda_if_i(x)+\sum_{i=1}^p\nu_ih_i(x)$

② Lagrange dual function 拉格朗日对偶

$g(\lambda,\nu)=\inf_{x\in D}L(x,\lambda,\nu)=\inf_{x\in D}(f_0(x)+\sum\lambda_if_i(x)+\sum\nu_ih_i(x))$
g is concave, can be $-\infty$ for some λ,ν.
Lower bound property: if $\lambda\ge 0$, then $g(\lambda,\nu)\le p^*$

例: $\min x^Tx$ s.t. $Ax=b$.
解: $L(x,\nu)=x^Tx+\nu^T(Ax-b)$ 令梯度为零 $\Rightarrow$
$\nabla_xL(x,\nu)=2x+A^T\nu=0 \Rightarrow x=-(1/2)A^T\nu$
故 $g(\nu)=L((-1/2)A^T\nu,\nu)=-\frac14\nu^TAA^T\nu-b^T\nu$
Lower bound property: $p^*\ge -(1/4)\nu^TAA^T\nu-b^T\nu$.

例: $\min c^Tx$ s.t. $Ax=b$, $x\ge 0$
解: $L(x,\lambda,\nu)=c^Tx+\nu^T(Ax-b)-\lambda^Tx=-b^T\nu+(c+A^T\nu-\lambda)^Tx$
$g(\lambda,\nu)=\inf_xL(x,\lambda,\nu)=\begin{cases}-b^T\nu & A^T\nu-\lambda+c=0\\-\infty & \text{otherwise}.\end{cases}$
下界性质: $p^*\ge -b^T\nu$ 若 $A^T\nu+c\ge 0$

例: $\min \|x\|$ s.t. $Ax=b$
解: $g(\nu)=\inf_x(\|x\|-\nu^TAx+b^T\nu)=\begin{cases}b^T\nu. & \|A^T\nu\|_*\le 1\\-\infty & \text{otherwise}.\end{cases}$
$\|\nu\|_*=\sup_{\|u\|\le 1}u^T\nu$ is dual norm of $\|\cdot\|$.
证明: $\inf(\|x\|-y^Tx)=0$ 若 $\|y\|_*\le 1$, $-\infty$ otherwise.
$\|y\|_*\le 1 \Rightarrow \|x\|-y^Tx\ge 0$ for all x. equality if $x=0$
$\|y\|_*>1$, chose $x=tu$. $\|u\|\le 1$. $u^Ty=\|y\|_*>1$
$\|x\|-y^Tx=t(\|u\|-\|y\|_*)\to-\infty$. as $t\to\infty$
下界性质: $p^*\ge b^T\nu$ 若 $\|A^T\nu\|_*\le 1$

③ 拉格朗日对偶和共轭函数

$\min f_0(x)$. s.t. $Ax\le b$, $Cx=d$.
解: $g(\lambda,\nu)=\inf_{x\in\mathrm{dom} f_0}(f_0(x)+(A^T\lambda+C^T\nu)^Tx-b^T\lambda-d^T\nu)$
$=-f_0^*(-A^T\lambda-C^T\nu)-b^T\lambda-d^T\nu$.
其中共轭函数: $f^*(y)=\sup_{x\in\mathrm{dom} f}(y^Tx-f(x))$ *

④ 对偶问题. The dual problem.

原问题: $\min c^Tx$ s.t. $Ax=b$. $x\ge 0$
对偶问题: $\max -b^T\nu$ s.t. $A^T\nu+c\ge 0$
在拉格朗日对偶函数中取最好的下界 p^*
假设最优的最优值为 d^*. 弱对偶: $d^*\le p^*$.
强对偶: $d^*=p^*$. 通过 Slater's condition 可以弱对偶变成强对偶

例. 原问题 $\min c^Tx$. s.t. $Ax\le b$. 则对偶函数.
$g(\lambda)=\inf_x((c+A^T\lambda)^Tx-b^T\lambda)=\begin{cases}-b^T\lambda, & A^T\lambda+c=0\\-\infty & \text{otherwise}.\end{cases}$
对偶问题: $\max -b^T\lambda$. s.t. $A^T\lambda+c=0$, $\lambda\ge 0$
From Slater's 条件: $p^*=d^*$ 若 $A\tilde x<b$ for some $\tilde x$.

例: 原问题 $\min x^TPx$. s.t. $Ax\le b$
对偶函数: $g(\lambda)=\inf_x(x^TPx+\lambda^T(Ax-b))=-\frac14\lambda^TAP^{-1}A^T\lambda-b^T\lambda$
对偶问题: $\max -(1/4)\lambda^TAP^{-1}A^T\lambda-b^T\lambda$ s.t. $\lambda\ge 0$

⑤ Complementary slackness 互补松弛条件.

假设强对偶: x^* 原问题最优. (λ^*,ν^*) 对偶最优.
那么: $\lambda_i^*>0 \Rightarrow f_i(x^*)=0$. $f_i(x^*)<0 \Rightarrow \lambda_i^*=0$

⑥ KKT条件: 下面4个条件称为KKT条件.

1) 原问题约束: $f_i(x)\le 0$, $h_i(x)=0$
2) 对偶约束: $\lambda\ge 0$. 3) 互补松弛: $\lambda_if_i(x)=0$.
4) 拉格朗日函数关于 x 的梯度: $\nabla_xL(x,\lambda,\nu)=\nabla f_0(x)+\sum\lambda_i\nabla f_i(x)+\sum\nu_i\nabla h_i(x)=0$
如果强对偶成立. x,λ,ν 是最优的, 则满足KKT条件

⑦ 凸优化问题的KKT条件. $\tilde x,\tilde\lambda,\tilde\nu$ 对于一个凸优化问题满足KKT条件. 那么 $f_0(\tilde x)=L(\tilde x,\tilde\lambda,\tilde\nu)$
$g(\tilde\lambda,\tilde\nu)=L(\tilde x,\tilde\lambda,\tilde\nu) \Rightarrow f_0(\tilde x)=g(\tilde\lambda,\tilde\nu)$.
对于无约束问题. 取最优的条件为 $\nabla f_0(x)=0$
若为凸优化问题. 则KKT条件为充分必要条件.

⑧ 引入新的变量和等式约束 $\min f_0(Ax+b)$
对偶函数是常值: $g=\inf_xL(x)=\inf_xf_0(Ax+b)=p^*$
原问题: $\min f_0(y)$ s.t. $Ax+b-y=0$.
对偶问题: $\max b^T\nu-f_0^*(\nu)$ s.t. $A^T\nu=0$
对偶函数: $g(\nu)=\inf_{x,y}(f_0(y)-\nu^Ty+\nu^TAx+b^T\nu)$
$=\begin{cases}-f_0^*(\nu)+b^T\nu, & A^T\nu=0\\-\infty, & \text{otherwise}\end{cases}$ $f^*(y)=\sup_{x\in\mathrm{dom} f}(y^Tx-f(x))$

⑨ 范数近似问题. $\min \|Ax-b\|$.
$\min \|y\|$. s.t. $y=Ax-b$. 其拉格朗日对偶函数
$g(\nu)=\inf_{x,y}(\|y\|+\nu^Ty-\nu^TAx+b^T\nu)$
$=\begin{cases}b^T\nu+\inf_y(\|y\|+\nu^Ty), & A^T\nu=0\\-\infty, & \text{otherwise}.\end{cases}=\begin{cases}b^T\nu, & A^T\nu=0,\ \|\nu\|_*\le 1\\-\infty, & \text{otherwise}.\end{cases}$

相关题目: 找对偶问题 $\min c^Tx$. s.t. $Gx\le h$. $Ax=b$
解: The Lagrangian is $L(x,\lambda,\nu)=c^Tx+\lambda^T(Gx-h)+\nu^T(Ax-b)=(c^T+\lambda^TG+\nu^TA)x-h^T\lambda-\nu^Tb$
对偶函数: $g(\lambda,\nu)=\inf_xL(x,\lambda,\nu)=\begin{cases}-\lambda^Th-\nu^Tb, & c+G^T\lambda+A^T\nu=0\\-\infty, & \text{otherwise}.\end{cases}$
对偶问题: $\max g(\lambda,\nu)$. s.t. $\lambda\ge 0$. $\Rightarrow$
$\max -\lambda^Th-\nu^Tb$ s.t. $c+G^T\lambda+A^T\nu=0$. $\lambda\ge 0$.

相关题目: minimize $\|Ax-b\|_2^2$. s.t. $Gx=h$.
解: $L(x,\nu)=\|Ax-b\|_2^2+\nu^T(Gx-h)=x^TA^TAx+(G^T\nu-2A^Tb)^Tx-\nu^Th$. 最小 $x=-(1/2)(A^TA)^{-1}(G^T\nu-2A^Tb)$. 对偶函数
$g(\nu)=-(1/4)(G^T\nu-2A^Tb)^T(A^TA)^{-1}(G^T\nu-2A^Tb)-\nu^Th$
KKT条件: $2A^T(Ax^*-b)+G^T\nu^*=0$. $Gx^*=h$.
然后根据KKT条件. 去解出 x^* 和 ν^*

① 梯度下降法: $\Delta x=-\nabla f(x)$. $\|\nabla f(x)\|_2\le\epsilon$
② 最速下降法: $\Delta x_{sd}=\|\nabla f(x)\|_*\Delta x_{nsd}$.
满足 $\nabla f(x)^T\Delta_{sd}=-\|\nabla f(x)\|_*^2$
③ 牛顿方向: $\Delta x_{nt}=-\nabla^2f(x)^{-1}\nabla f(x)$
二阶近似: $\hat f(x+v)=f(x)+\nabla f(x)^Tv+\frac12v^T\nabla^2f(x)\cdot v$
一阶近似: $\nabla f(x+v)\approx\nabla\hat f(x+v)=\nabla f(x)+\nabla^2f(x)\cdot v=0$
(使得梯度为零). 二种解释方法.

LP and equivalent SDP

1. $\Delta x_{nt}=-\nabla^2f(x)^{-1}\nabla f(x)$.
$\lambda^2=\nabla f(x)^T\nabla^2f(x)^{-1}\nabla f(x)$. 判断 $\lambda^2/2\le\epsilon$
确定步长 t. $x=x+t\Delta x_{nt}$.

⑤ 牛顿方向: $\min f(x)$. s.t. $Ax=b$ $\xLeftrightarrow{KKT}$
$\nabla f(x^*)+A^Tw^*=0$. $Ax^*=b \Rightarrow \nabla f(x^k+\Delta x)+A^Tw=0$
$A(x^k+\Delta x)=b$. 利用 $\nabla f(x^k+\Delta x)=\nabla f(x^k)+\nabla^2f(x^k)\Delta x \Rightarrow$
$\nabla f(x^k)+\nabla^2f(x^k)\Delta x+A^Tw=0$. $A\Delta x=0 \Rightarrow$
当 x 是可行的. $\begin{bmatrix}\nabla^2f(x)&A^T\\A&0\end{bmatrix}\cdot\begin{bmatrix}\Delta x_{nt}\\w\end{bmatrix}=\begin{bmatrix}-\nabla f(x)\\0\end{bmatrix}$
当 x 不可行时. $\begin{bmatrix}\nabla^2f(x)&A^T\\A&0\end{bmatrix}\cdot\begin{bmatrix}\Delta x_{nt}\\w\end{bmatrix}=\begin{bmatrix}-\nabla f(x)\\b-Ax\end{bmatrix}$
二阶近似: $\min\hat f(x+v)=f(x)+\nabla f(x)^Tv+(1/2)v^T\nabla^2f(x)v$
s.t. $A(x+v)=b \xrightarrow{KKT} \nabla f(x+\Delta x_{nt})+A^Tw=0$.
$A(x+\Delta x_{nt})=b$. 其中步骤和普通的③中的步骤没有任何区别. 只是 $\lambda(x)=(-\nabla f(x)^T\Delta x_{nt})^{1/2}$

扩展: 不可行的时候, 优化条件为
$\begin{bmatrix}\nabla^2f(x)&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x_{nt}\\w\end{bmatrix}=-\begin{bmatrix}\nabla f(x)\\Ax-b\end{bmatrix} \xrightarrow{KKT条件} \begin{cases}\nabla f(x)+A^T\nu\\Ax-b\end{cases}$
$\begin{bmatrix}\nabla^2f(x)&A^T\\A&0\end{bmatrix}\begin{bmatrix}\Delta x_{nt}\\\Delta\nu_{nt}\end{bmatrix}=-\begin{bmatrix}\nabla f(x)+A^T\nu\\Ax-b\end{bmatrix}$ 其中 $w=\nu+\Delta\nu_{nt}$

Calculus of variations. 变分法 ☆

性能指标: $J=h(x(t_f),t_f)+\int_{t_0}^{t_f}g(x(t),u(t),t)dt$
$\dot x(t)=a(x(t),u(t),t)$.
Hamiltonian $H(x,u,p,t)=g(x,u,t)+p^T(t)a(x,u,t)$
即哈密顿函数: $H=g+p^Ta$
$\delta J_a=0$ 的必要条件是: 状态方程: $\dot x=a(x,u,t)$
协态方程: $\dot p=-H_x^T$. 控制方程: $H_u=0$.
边界条件: $h_t(t_f)+H(t_f)=0$ 计算 t_f 的值.
还要满足: $\dot x=a(x,u,t)=\frac{\partial H}{\partial p}$, $\dot p=-\left(\frac{\partial H}{\partial x}\right)^T$
如果 $x_i(t_f)$ 是自由的, 还要满足横截条件:
$p_i(t_f)=\frac{\partial h}{\partial x_i}(t_f)$ 下面给出例题:

相关问题: $\dot x=-x(t)+u(t)$. $x(t)$ 的边界条件为 $x(0)=1$, $x(t_f)=0$. 求最优控制 $u(t)$, 使下列性能指标 $J=\frac12\int_0^{t_f}(x^2+u^2)dt$ 为最小.
解: $x(0)$, $x(t_f)$ 均给定, 所以不需要横截条件
$H=\frac12(x^2+u^2)+\lambda(-x+u)$ 利用协态方程和控制方程: $\dot\lambda=-\frac{\partial H}{\partial x}=-x+\lambda$, $\frac{\partial H}{\partial u}=u+\lambda=0 \Rightarrow u=-\lambda$
$\Rightarrow \dot x(t)=-x(t)-\lambda(t)$. $\dot\lambda(t)=-x(t)+\lambda(t)$.
进行拉氏变换: $SX(S)-x(0)=-X(S)-\lambda(S)$
$S\lambda(S)-\lambda(0)=-X(S)+\lambda(S) \Rightarrow X(S)=\frac{x(0)-\lambda(S)}{S+1}$
于是可以解出 $\lambda(S)$ 和 $X(S)$, 然后通过拉氏反变换可以得出 $\lambda(0)$, 最后求出 $u^*(t)=-\lambda(t)$

相关例题: 状态方程 $\dot x_1(t)=x_2(t)$. $\dot x_2(t)=u(t)$
初始条件为: $x_1(0)=1$, $x_2(0)=1$
终端条件为: $x_1(1)=0$, $x_2(1)$ 自由, 求最优控制 $u^*(t)$. 使指标泛函 $J(u)=\frac12\int_0^1u^2(t)dt$ 最小.
解: 因为 $x_2(1)$ 是自由的, 所以用到横截条件. 但是 $h(x(t_f),t_f)=0$. 故 $P_2(1)=\frac{\partial h}{\partial x_2(1)}=0$.
$H=\frac12u^2+P_1x_2+P_2u$. 根据必要条件可以得出
$\dot P_1=-\frac{\partial H}{\partial x_1}=0$, $\dot P_2=-\frac{\partial H}{\partial x_2}=-P_1$, $\frac{\partial H}{\partial u}=0 \Rightarrow$
$u+P_2=0. \Rightarrow u^*(t)=-P_2(t)$.
所以正则方程: $\dot x_1=x_2(t)$. $\dot x_2=-P_2(t)$
$\dot P_1=0$. $\dot P_2=-P_1(t)$
边界条件: $x_1(0)=1$. $x_2(0)=1$. $x_1(1)=0$. $P_2(1)=0$
进行拉氏变换: $SX_1(S)-x_1(0)=X_2(S)$
$SX_2(S)-x_2(0)=-P_2(S)$, $SP_1(S)-P_1(0)=0$
$SP_2(S)-P_2(0)=-P_1(S)$. $\Rightarrow S^4X_1(S)=S^3x_1(0)+S^2x_2(0)-SP_2(0)+P_1(0)$ 代入初始条件条件: $x_1(0)=1$, $x_2(0)=1$
$X_1(S)=\frac1S+\frac1{S^2}-\frac1{S^3}P_2(0)+\frac1{S^4}P_1(0) \Rightarrow$
$x_1(t)=1+t-\frac12P_2(0)t^2+\frac16P_1(0)t^3$. 同理 $P_2(S)=\frac1SP_2(0)-\frac1{S^2}P_1(0)$, $P_2(t)=P_2(0)-P_1(0)t$ 利用条件
$x_1(1)=0$, $P_2(1)=0$, 可得 $2-\frac12P_2(0)+\frac16P_1(0)=0$.
可以解出 $P_1(0)=6$. $P_2(0)=6$. $x_1^*(t)=1+t-3t^2+t^3$
$P_2^*(t)=6(1-t)$. $u^*(t)=6(t-1)$. $x_2^*(t)=1-6t+3t^2$

相关例题: 双积分系统: $\ddot y=u$, $y(0)=10$.
$\dot y(0)=0$. $y(t_f)=\dot y(t_f)=0$. $J=\frac12\alpha t_f^2+\frac12\int_0^{t_f}bu^2(t)dt$
定义: $x_1=y$, $x_2=\dot y$. $x(t)=\begin{bmatrix}y(t)\\\dot y(t)\end{bmatrix}$, $u(t)=\ddot y(t)$
$\dot x(t)=Ax(t)+Bu(t)$ $x(0)=\begin{bmatrix}10\\0\end{bmatrix}$, $x(t_f)=\begin{bmatrix}0\\0\end{bmatrix}$
解: $H=g+p^Ta=\frac12bu^2+p_1x_2+p_2u$
因为 $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$, $B=\begin{bmatrix}0\\1\end{bmatrix}$ 根据: $\dot p=-H_x^T \Rightarrow$
$\begin{cases}\dot p_1=-\frac{\partial H}{\partial x_1}=0\to p_1(t)=c_1\\\dot p_2=-\frac{\partial H}{\partial x_2}=-p_1\to p_2(t)=-c_1t+c_2\end{cases}$
$H_u=bu+p_2=0\to u=-\frac{p_2}{b}=-\frac{c_2}{b}+\frac{c_1}{b}t$
$H(t_f)+h_t(t_f)=\frac12bu^2(t_f)+p_1(t_f)x_2(t_f)+p_2(t_f)u(t_f)+\alpha t_f=0=\frac12bu^2(t_f)+(-bu(t_f))\cdot u(t_f)+\alpha t_f$
$=-\frac12bu^2(t_f)+\alpha t_f=0\to t_f=\frac{1}{2b\alpha}(-c_2+c_1t_f)^2$
$x_2(t)=-\frac{c_2}{b}+\frac{c_1}{b}t\to x_2(t)=c_3-\frac{c_2}{b}t+\frac{c_1}{2b}t^2$
$x_2(0)=0$. $c_3=0$. $\dot x_1(t)=x_2(t)\to x_1(t)=c_4-\frac{c_2}{2b}t^2+\frac{c_1}{6b}t^3$, $x_1(0)=10$. $c_4=10$.
根据 $x_2(t_f)=0$. $x_1(t_f)=0 \Rightarrow$
$c_2=\frac{60b}{t_f^2}$. $c_1=\frac{120b}{t_f^3}$

拉氏变换: $1\to\frac1s$, $\delta(t)\to 1$, $K\to\frac Ks$
$t\to\frac1{s^2}$, $t^n\to\frac{n!}{s^{n+1}}$, $e^{-at}\to\frac{1}{s+a}$, $te^{-at}\to\frac{1}{(s+a)^2}$
$1-e^{-at}\to\frac{a}{s(s+a)}$, $\frac1Te^{-t/T}\to\frac{1}{Ts+1}$

第四章相关题目: 证明 $x^*=(1,1/2,-1)$ 是下列优化问题的最优解. $\min (1/2)x^TPx+q^Tx+r$
s.t. $-1\le x_i\le 1$, $i=1,2,3$. P, q, r 是已知量.
解: $\nabla f_0(x^*)=(-1,0,2)$ 且.
$\nabla f_0(x^*)\cdot(y-x)=-1(y_1-1)+2(y_3+1)\ge 0$ for all y satisfying $-1\le y_i\le 1$, which is clearly true.

④ 牛顿下降步骤: 终止条件 $\varepsilon>0$

$\phi(u)=\begin{cases}u^2, & |u|\le M\\M(2|u|-M), & |u|\ge M\end{cases}$
(b) $\min \sum_{i=1}^m(a_i^Tx-b_i)^2/(w_i+1)+M^21^Tw$. s.t. $w\ge 0$
(c) $\min \sum_{i=1}^m(u_i^2+2Mv_i)$. s.t. $-u-v\le Ax-b\le u+v$
$0\le u\le MI$. $v\ge 0$.
解: (a)⇔(b). fix u. $\min u^2/(w+1)+M^2w$ is given by $w=\begin{cases}|u|/M-1, & |u|\ge M\\0, & \text{otherwise}\end{cases}$ the optimal value.
$\inf_{w\ge 0}(u^2/(w+1)+M^2w)=\begin{cases}M(2|u|-M), & |u|\ge M\\u^2, & \text{otherwise}.\end{cases}$
the optimal $w_i=\begin{cases}|a_i^Tx-b_i|/M-1, & |a_i^Tx-b_i|\ge M\\0, & \text{otherwise}\end{cases}$
(a)⇔(c). fix c. $u_i+v_i=|a_i^Tx-b_i|$
$v_i=|a_i^Tx-b_i|-u_i \Rightarrow \min\sum_{i=1}^m(u_i^2-2Mu_i+2M|a_i^Tx-b_i|)$
s.t. $0\le u_i\le\min\{M,|a_i^Tx-b_i|\}$. If $|a_i^Tx-b_i|\le M$ then function reduces to $|a_i^Tx-b_i|$
If $|a_i^Tx-b_i|>M$. then function reduces to $2M|a_i^Tx-b_i|-M^2$. Finally. (c) is given by $\sum_{i=1}^m\phi(a_i^Tx-b_i)$.

相关题目: $\min c^Tx$, s.t. $Ax=b$ ①
解: The problem is infeasible ($b\notin R(A)$) the optimal value is ∞ | The problem is feasible and c is orthogonal to the nullspace of A. make $c=A^T\lambda+\hat c$. $A\hat c=0$. $\Rightarrow c^Tx=\lambda^TAx+\hat c^Tx=\lambda^Tb$ the optimal value is λ^Tb. | The problem is feasible, and c is not in the range of A^T ($\hat c\ne 0$). the problem is unbounded ($p^*=-\infty$) make $x=x_0-t\hat c$, $t\to\infty$.
故答案. $p^*=\begin{cases}+\infty, & b\notin R(A)\\\lambda^Tb, & c=A^T\lambda \text{ for some }\lambda\\-\infty, & \text{otherwise}.\end{cases}$

$\min c^Tx$. s.t. $a^Tx\le b$ where $a\ne 0$. ②
解: This problem is always feasible. make $c=a\lambda+\hat c$ with $a^T\hat c=0$. If $\lambda>0$. choose $x=-ta$ let $t\to\infty$. $c^Tx=-tc^Ta=-t\lambda a^Ta\to-\infty$ and $a^Tx-b=-ta^Ta-b\le 0$ for large t, so x is feasible for large t | If $\hat c\ne 0$, the problem is unbounded below. choose $x=ba-t\hat c$ and let t go to ∞ | If $c=a\lambda$ for some $\lambda\le 0$, the optimal value is $c^Tab=\lambda b$. In summary, the optimal value is $p^*=\begin{cases}\lambda b, & c=a\lambda \text{ for some }\lambda\le 0\\-\infty, & \text{otherwise}\end{cases}$

$\min c^Tx$. s.t. $l\le x\le u$, where $l\le u$ ③
解: 目标和约束是可以分离的. the objective is a sum of terms c_ix_i. $x_i\to l_i\le x_i\le u_i$
If $c_i>0$, then $x_i^*=l_i$, if $c_i<0$, then $x_i^*=u_i$
if $c_i=0$, $x_i\in[l_i,u_i]$. 故 $p^*=l^Tc^++u^Tc^-$
where $c_i^+=\max\{c_i,0\}$ and $c_i^-=\max\{-c_i,0\}$

$\min c^Tx$ s.t. $1^Tx=1$, $x\ge 0$ ④
解: Suppose $c_1=c_2=\ldots=c_k<c_{k+1}\le\cdots\le c_n$
$\Rightarrow c^Tx\ge c_1(1^Tx)=c_{\min}$ for all feasible x.
$x_1+\cdots+x_k=1$, $x_1\ge 0,\cdots,x_k\ge 0$, $x_{k+1}=\cdots=x_n=$
$p^*=c_1=c_{\min} \Rightarrow p^*=\min\{0,c_{\min}\}$

相关题目: $\min c^Tx$, s.t. $Ax\le b$. 求最优值
$p^*=\begin{cases}c^TA^{-1}b, & A^{-T}c\le 0\\-\infty, & \text{otherwise}.\end{cases}$ 解: make $y=Ax$
$\Rightarrow \min c^TA^{-1}y$, s.t. $y\le b$. If $A^{-T}c\le 0$, the optimal solution is $y=b$, with $p^*=c^TA^{-1}b$. Otherwise, the LP is unbounded below

相关题目: 1范数: $\|x\|_1=|x_1|+|x_2|+\cdots+|x_n|$
∞范数: $\|x\|_\infty=\max(|x_1|,|x_2|,\cdots,|x_n|)$.
Minimize $\|Ax-b\|_\infty$ ① Equivalent to the LP
$\min t$. s.t. $Ax-b\le tI$, $Ax-b\ge -tI$
$\Rightarrow -t\le a_k^Tx-b_k\le t. \Rightarrow t\ge|a_k^Tx-b_k| \Rightarrow$
$t\ge\max_k|a_k^Tx-b_k|=\|Ax-b\|_\infty$. so $p^*=\|Ax-b\|_\infty$

Minimize $\|Ax-b\|_1$ ② Equivalent to the LP
$\min 1^Ts$ s.t. $Ax-b\le s$, $Ax-b\ge -s$.
$\Rightarrow -s_k\le a_k^Tx-b_k\le s_k. \Rightarrow s_k\ge|a_k^Tx-b_k|$
so choose $s_k=|a_k^Tx-b_k| \Rightarrow p^*(x)=\|Ax-b\|_1$

Minimize $\|Ax-b\|_1$ subject to $\|x\|_\infty\le 1$ ③
Equivalent to the LP: $\min 1^Ty$
s.t. $-y\le Ax-b\le y$, $-1\le x\le 1$. , $x\in R^n$ and $y\in R^m$

Minimize $\|x\|_1$ subject to $\|Ax-b\|_\infty\le 1$ ④
Equivalent to the LP: $\min 1^Ty$
s.t. $-y\le x\le y$, $-1\le Ax-b\le 1$ (x,y 均为变量).

Minimize $\|Ax-b\|_1+\|x\|_\infty$ ⑤
Equivalent to: $\min 1^Ty+t$
s.t. $-y\le Ax-b\le y$, $-tI\le x\le tI$. (x,y,t 变量)

网络问题: $\min C=\sum_{i,j=1}^n c_{ij}x_{ij}$.
s.t. $b_i+\sum_{j=1}^n x_{ij}-\sum_{j=1}^n x_{ji}=0$, $i=1,2,\ldots,n$.
$l_{ij}\le x_{ij}\le u_{ij}$

Euler Equation: $\frac{\partial g(x(t),\dot x(t),t)}{\partial x}-\frac{d}{dt}\left(\frac{\partial g}{\partial\dot x}\right)=0$
$g_x(x^*(t),\dot x^*(t),t)-\frac{d}{dt}g_{\dot x}(x^*(t),\dot x^*(t),t)=0$ ①
$\begin{cases}g_x(x^*(t),\dot x^*(t),t)-\frac{d}{dt}g_{\dot x}(x^*(t),\dot x^*(t),t)=0\\g(x^*(t_f),\dot x^*(t_f),t_f)-g_{\dot x}(x^*(t_f),\dot x^*(t_f),t_f)\dot x^*(t_f)=0\\g_{\dot x}(x^*(t_f),\dot x^*(t_f),t_f)=0.\end{cases}$ ②
$\begin{cases}g_x(x^*(t),\dot x^*(t),t)-\frac{d}{dt}g_{\dot x}(x^*(t),\dot x^*(t),t)=0\\g_{\dot x}(x^*(t_f),\dot x^*(t_f),t_f)\left[\frac{d\theta}{dt}(t_f)-\dot x^*(t_f)\right]+g(x^*(t_f),\dot x^*(t_f),t_f)=0\end{cases}$ ③
$\begin{cases}g_x-\frac{d}{dt}g_{\dot x}=0\\w_x(t_f)+g_{\dot x}(t_f)=0\\w_{t_f}+g(t_f)-g_{\dot x}(t_f)\dot x(t_f)=0\end{cases}$ ④

$\min c^Tx$ s.t. $1^Tx=\alpha$, $0\le x\le 1$. 解:
$c_1\le\cdots\le c_{i_\alpha}<c_{i_\alpha+1}=\cdots=c_\alpha=\cdots=c_k<c_{k+1}\le\cdots\le c_n$
最优值 $c_1+c_2+\cdots+c_\alpha$ $x_1=\cdots=x_{i_\alpha}=1$ $x_{i_\alpha+1}+\cdots+x_k=\alpha-i_\alpha$
$x_{k+1}=\cdots=x_n=0$, $p^*=c_1+c_2+\cdots+c_{i_\alpha}+c_{i_\alpha+1}(\alpha-i_\alpha)$