# 线性代数
# 投影矩阵

曾吉文
哈尔滨工业大学数学科学学院

2022年 9 月

主要内容：对于一个 $n$ 阶矩阵 $A$, 证明有一个零化多项式 $f(x)$, 满足 $f(A)=0$，且这个零化多项式就是矩阵 $A$ 的特征多项式。

1. 空间中向量到向量的投影
2. $n$维向量的内积，距离，正交
3. 向量 $Y$ 到一维向量空间的投影
4. 几何向量空间中应用问题
   - 点到直线的投影

# 1 空间中向量到向量的投影

给定空间中两个点$A=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B=\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$,

**定义 1.1**

向量 $\overrightarrow{OB}$ 到向量 $\overrightarrow{OA}$ 的投影向量，是指向量 $\overrightarrow{OP}$，其中点 $P$ 在直线 $OA$ 上，且向量 $\overrightarrow{PB}$ 垂直于向量 $\overrightarrow{OA}$.

注意：向量$\overrightarrow{PB}$ 的长度 $|\overrightarrow{PB}|$ 定义为点 $B$ 到直线 $OA$ 的距离。

设

$$
\begin{aligned}
\overrightarrow{OP} &= k\overrightarrow{OA} \\
\overrightarrow{PB} &= \overrightarrow{OB} - \overrightarrow{OP} \\
\overrightarrow{OA} \bullet \overrightarrow{PB} &= 0 \\
\overrightarrow{OA} \bullet (\overrightarrow{OB} - \overrightarrow{OP}) &= 0 \\
\overrightarrow{OA} \bullet (\overrightarrow{OB} - k\overrightarrow{OA}) &= 0 \\
k &= \frac{\overrightarrow{OA} \bullet \overrightarrow{OB}}{\overrightarrow{OA} \bullet \overrightarrow{OA}}
\end{aligned}
$$

现在用列向量分别代替上述表示：

$$\overrightarrow{OA} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = X, \overrightarrow{OB} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = Y$$

向量的内积可以用矩阵的乘积表示：

$$k = \frac{X'Y}{X'X}, \overrightarrow{OP} = k\overrightarrow{OA} = \frac{X'Y}{X'X}X = X\frac{X'Y}{X'X}$$

特别，记

$$\overrightarrow{OP} = \frac{XX'}{X'X}Y$$

对于列向量 $X$,称矩阵 $\frac{XX'}{X'X}$ 为向量 $X$ 的投影矩阵，记为：

$$P_X = \frac{XX'}{X'X}$$

由此，任意向量 $Y$ 在向量 $X$ 的投影向量记为：

$$Y_X = \overrightarrow{OP} = P_X Y$$

**Theorem 1.2**

给定空间向量 $X, Y$, 则向量 $Y$ 在向量 $X$ 的投影向量为：

$$Y_X = P_X Y = \frac{XX'}{X'X} Y, P_X = \frac{XX'}{X'X}$$

称为向量 $X$ 的投影矩阵。

# 2 $n$维向量的内积，距离，正交

设有两个 $n$ 维向量 $X, Y$,

**定义 2.1**

对于向量 $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 与向量 $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$，两个向量的内积定义为：

$$(X, Y) = X'Y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

容易证明：

$$(X, X) \geq 0$$

等号成立，当且仅当 $X = 0$

内积具有双线性性质和对称性质：

- $(X, Y) = (Y, X)$
- $(kX, Y) = k(X, Y), (X, kY) = k(X, Y)$
- $(X + Y, Z) = (X, Z) + (Y, Z), (X, Y + Z) = (X, Y) + (X, Z)$

## 定义 2.2

对于向量 $X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 向量$X$的长度为：

$$|X|=\sqrt{X'X}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$$

## Theorem 2.3

$$|(X,Y)| \leq |X||Y|$$

等号成立，当且仅当 $X$ 与 $Y$ 线性相关。

## Proof.

当 $X$ 与 $Y$ 线性相关，可以设 $Y = kX$,

$$|(X,Y)| = |(X,kX)| = |k||(X,X)| = \\ |k||X|^2 = |k||X||X| = |X||kX| = |X||Y|$$

当 $X$ 与 $Y$ 线性相关，$\lambda X - Y \neq 0, \forall \lambda$.

$$\begin{aligned} 0 < (\lambda X - Y, \lambda X - Y) = \lambda^2(X,X) - 2\lambda(X,Y) + (Y,Y) \\ \Rightarrow 4(X,Y)^2 - 4(X,X)(Y,Y) < 0 \\ \Rightarrow (X,Y)^2 < (X,X)(Y,Y) \Rightarrow |(X,Y)| < |X||Y| \end{aligned}$$

$\square$

if $|(X,Y)| = |X||Y|$, it means that:

$$(\lambda X - Y, \lambda X - Y) = \lambda^2 (X,X) - 2\lambda (X,Y) + (Y,Y) = 0$$

有解 $\lambda_0$, 因此有：

$$\lambda_0 X - Y = 0, Y = \lambda_0 X$$

得到 $X$ 与 $Y$ 线性相关。

**定义 2.4**

对向量 $X$ 和向量 $Y$，定义它们的夹角 $\theta$ 满足：

$$\cos\theta = \frac{(X,Y)}{|X||Y|}, \theta = \arccos\frac{(X,Y)}{|X||Y|}$$

$0 \le \theta \le \pi$

如果两个向量的夹角为 $\frac{\pi}{2}$，称两个向量正交。

根据夹角定义，向量 $X$ 与向量 $Y$ 正交，当且仅当 $(X,Y) = X'Y = 0$

设有两个 $n$ 维向量 $X, Y$，考虑向量 $Y$ 到向量 $X$ 生成的向量空间 $L(X)=\{aX|a\in F\}$的距离。

**定义 2.5**

对于向量 $X=\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{pmatrix}$ 与向量 $Y=\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n \end{pmatrix}$，两个向量的距离定义为：

$$|X-Y|=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}$$

**定义 2.6**

给定向量 $n$ 维向量 $Y$ 和一个$n$ 维 向量空间 $L$, 向量$Y$ 到向量空间 $L$ 的距离 定义为：

$$d = \min\{|Y - X||X \in L\}$$

即向量 $X$ 跑遍向量空间 $L$ 时，最小的值 $|Y - X|$

## 3. 向量 $Y$ 到$n$ 维向量空间的投影

假设有向量空间 $L(X)=\{aX|a\in F\}$, 定义向量 $Y$ 到向量空间 $L(X)$ 的投影为向量 $Y_X\in L(X)$,满足条件：

$$X'(Y-Y_X)=0$$

注意，如果 $Y\in L(X)$, 则有$Y-Y_X=aX\in L(X)$

$$(Y-Y_X)'(Y-Y_X)=aX'(Y-Y_X)=0\Rightarrow Y=Y_X$$

既然 $Y_X \in L(X)$,设 $Y_X = aX$

$$X'(Y - Y_X) = 0 \Rightarrow X'(Y - aX) = 0$$

$$a = \frac{X'Y}{X'X} \Rightarrow Y_X = \frac{X'Y}{X'X}X = X\frac{X'Y}{X'X} = \frac{XX'}{X'X}Y$$

**定义 3.1**

对于 $n$维列向量 $X$， 称矩阵 $\frac{XX'}{X'X}$ 为向量 $X$ 的投影矩阵。

**Theorem** 3.2

对于任意向量 $Y$ 和向量空间 $L(X)=\{aX|a\in F\}$, 向量 $Y$ 在空间 $L$ 的投影为：

$$Y_X=\frac{XX'}{X'X}Y$$

且$|Y-Y_X|$ 为向量 $Y$ 到空间 $L(X)$ 的距离。

Proof.

结论的第一部分已经证明。现在证明第二部分。 对任意向量 $aX\in L(X)$，有：$|Y-Y_X|\le|Y-aX|$ □

$$
\begin{aligned}
|Y-aX| = |Y-Y_X+Y_X-aX|, Y_X-aX \in L(X) \\
\Rightarrow Y-Y_X \text{正交于} Y_X-aX \\
\Rightarrow |Y-Y_X|^2+|Y_X-aX|^2 = |Y-aX|^2 \\
\Rightarrow |Y-Y_X| \leq |Y-aX|
\end{aligned}
$$

证毕。注意：一般，当向量 $X$ 与向量 $Y$ 正交时，有：

$$|X|^2+|Y|^2=|X+Y|^2$$

这是因为：

$$
\begin{aligned}
|X+Y|^2 &= (X+Y, X+Y) \\
&= (X,X) + 2(X,Y) + (Y,Y) \\
&= |X|^2 + |Y^2|
\end{aligned}
$$

继续以上思想，考虑更一般的向量空间 $L(X_1, X_2, ..., X_r)$，由线性无关的向量：$X_1, X_2, ..., X_r$ 生成。研究向量 $Y$ 到向量空间 $L(X_1, X_2, ..., X_r)$ 的距离和投影。

## 定义 3.3

假设有向量空间 $L(X_1, X_2, ..., X_r)$，定义向量 $Y$ 到向量空间 $L(X_1, X_2, ..., X_r)$ 的投影为向量 $Y_L \in L(X_1, X_2, ..., X_r)$,满足条件：

$$X'(Y - Y_L) = 0, \forall X \in L(X_1, X_2, ..., X_r)$$

注意，当 $Y \in L(X_1, X_2, ..., X_r)$，因为 $Y - Y_L \in L$，有：

$$(Y - Y_L)'(Y - Y_L) = 0 \Rightarrow Y - Y_L = 0, Y = Y_L$$

现在设 $n$ 维向量 $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$，向量空间 $L$ 由线性无关向量

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, , \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \ldots, \alpha_r = \begin{pmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{nr} \end{pmatrix}$$

生成。

$$X'(Y - Y_L) = 0 \Leftrightarrow \alpha_j'(Y - Y_L) = 0, j = 1, 2, \ldots, r$$

假设

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_r \end{pmatrix}$$

$$Y_L = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_r\alpha_r$$

$$Y_L = A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix}$$

因此有：

$$0 = A'(Y - Y_L) = A'(Y - AX) = 0, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{pmatrix}$$

$$\begin{aligned} A'AX = A'Y, X &= (A'A)^{-1}A'Y \\ Y_L = AX &= A(A'A)^{-1}A'Y \\ A(A'A)^{-1}A' &= P_A \end{aligned}$$

称为矩阵 $A$ 的列空间的投影矩阵。总结以上讨论，得到结论：

**Theorem 3.4**

对于$n$ 维列向量 $Y$ 和线性无关的向量 $\alpha_1, \alpha_2, ..., \alpha_r$ 生成的子空间 $L(\alpha_1, \alpha_2, ..., \alpha_r)$

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_r \end{pmatrix}$$
$$Y_L = P_A Y = A(A'A)^{-1}A'Y$$

为向量 $Y$ 在向量空间 $L$ 的投影。而且 $|Y - Y_L|$ 是向量 $Y$ 到空间 $L$ 的距离。

**Proof.**

定理的第一部分已经证明。现在证明第二部分：任取向量空间 $L(\alpha_1, \alpha_2, ..., \alpha_r)$ 中的点 $P$, 希望证明

$$|Y - Y_L| \leq |Y - P|$$

因为$Y_L - P \in L$

$$|Y - P| = |Y - Y_L + Y_L - P|, (Y_L - P)'(Y - Y_L) = 0$$

$$|Y - Y_L|^2 \leq |Y - Y_L|^2 + |Y_L - P|^2 = |Y - P|^2$$

所以有：

$$|Y - Y_L| \leq |Y - P|$$

□

关于投影的一个判别定理：

**Theorem** 3.5

对于$n$ 维列向量 $Y$ 和线性无关的向量 $\alpha_1, \alpha_2, ..., \alpha_r$ 生成的子空间 $L(\alpha_1, \alpha_2, ..., \alpha_r)$。 $\alpha \in L$为向量 $Y$ 在向量空间 $L$ 的投影,当且仅当

$$|Y - \alpha| \leq |Y - \beta|, \forall \beta \in L$$

Proof.

定理的必要性已经证明。现在证明充分性：只需证明 $Y_L - \alpha = 0$. 因为$\alpha - Y_L \in L$

$$|Y - \alpha| = |Y - Y_L + Y_L - \alpha|, (Y_L - \alpha)'(Y - Y_L) = 0$$

$$|Y - Y_L|^2 \leq |Y - Y_L|^2 + |Y_L - \alpha|^2 = |Y - \alpha|^2 \leq |Y - Y_L|^2$$

□

所以有：

$$|Y - Y_L|^2 + |Y_L - \alpha|^2 = |Y - \alpha|^2 = |Y - Y_L|^2$$

$$|Y_L - \alpha|^2 = 0, Y_L = \alpha$$

# 4. 几何向量空间中应用问题

# 4.1 点到直线的投影

## 4.1 点到直线的投影

设有点： $P=(x_1,y_1,z_1)$ 和直线：

$$l:\frac{x-x_0}{m}=\frac{y-y_0}{n}=\frac{z-z_0}{p}$$

$M=(x_0,y_0,z_0)$ 为直线上的点， $\vec{s}=(m,n,p)$ 为直线的方向向量。

**定义 4.1**

给定直线 $l$ 和直线外一点 $P=(x_1,y_1,z_1)$，经过点 $P$ 和直线 $l$ 垂直的平面为 $\pi$, 则平面 $\pi$ 与直线 $l$ 的交点 $P_l$ 称为点 $P$ 在直线 $l$ 的投影点。

设点 $P$ 在直线 $l$ 上的投影点为 $P_l=(x,y,z)$. 向量 $MP_l$ 为向量 $MP$ 在向量 $\overrightarrow{s}=(m,n,p)$ 的投影。

$$\overrightarrow{MP_l} = \frac{\overrightarrow{s}'\overrightarrow{s}}{\overrightarrow{s}\,\overrightarrow{s}'}\overrightarrow{MP}$$

$$= \frac{1}{m^2+n^2+p^2}\begin{pmatrix} m^2 & mn & mp \\ nm & n^2 & np \\ pm & pn & p^2 \end{pmatrix}\begin{pmatrix} x_1-x_0 \\ y_1-y_0 \\ z_1-z_0 \end{pmatrix}$$

设 $P_l = (x_0+tm, y_0+tn, z_0+tp), \overrightarrow{MP_l} = t\begin{pmatrix} m \\ n \\ p \end{pmatrix}$

$$t\begin{pmatrix} m \\ n \\ p \end{pmatrix} = \frac{\vec{s}\,'\vec{s}}{\vec{s}\,\vec{s}\,'}\overrightarrow{MP}'$$

$$= \frac{1}{m^2+n^2+p^2}\begin{pmatrix} m \\ n \\ p \end{pmatrix}\begin{pmatrix} m & n & p \end{pmatrix}\begin{pmatrix} x_1-x_0 \\ y_1-y_0 \\ z_1-z_0 \end{pmatrix}$$

$$= \frac{m(x_1-x_0)+n(y_1-y_0)+p(z_1-z_0)}{m^2+n^2+p^2}\begin{pmatrix} m \\ n \\ p \end{pmatrix}$$

$$t = \frac{\vec{s}}{|\vec{s}|^2}\bullet\overrightarrow{MP} = \frac{1}{|\vec{s}|}\vec{s}_0\bullet\overrightarrow{MP}$$

$$P_l=(x,y,z),x=x_0+tm=x_0+\frac{m}{|\overrightarrow{s}|}\overrightarrow{s_0}\bullet\overrightarrow{MP}$$
$$y=y_0+tn=y_0+\frac{n}{|\overrightarrow{s}|}\overrightarrow{s_0}\bullet\overrightarrow{MP}$$
$$z=y_0+tp=z_0+\frac{p}{|\overrightarrow{s}|}\overrightarrow{s_0}\bullet\overrightarrow{MP}$$

- 直线$l$ 经过原点 $O(0,0,0)=M,\overrightarrow{MP}=(x_1,y_1,z_1)$,
- $t=\frac{1}{|\overrightarrow{s}|}\overrightarrow{s_0}\bullet\overrightarrow{MP}=\frac{mx_1+ny_1+pz_1}{m^2+n^2+p^2}$
- $$P_l=(x,y,z),x=tm=\frac{m^2x_1+mny_1+mpz_1}{m^2+n^2+p^2}$$

$$P_l = (x, y, z), y = tn = \frac{nmx_1 + n^2y_1 + npz_1}{m^2 + n^2 + p^2}$$

$$P_l = (x, y, z), z = tp = \frac{pmx_1 + pny_1 + p^2z_1}{m^2 + n^2 + p^2}$$

## Example 1

求点 $A(2,4,3)$ 在直线：$L : x = y = z$ 上投影点坐标，并求出点$A$ 到该直线的距离。

## 解.

分析：首先求出经过点$A$，并且与直线 $L$ 垂直的平面 $\pi$,然后求出平面 $\pi$ 与直线 $L$ 的交点，即为投影点。根据点法式：

$$\begin{aligned} \pi : x-2+y-4+z-3 &= 0 \\ x+y+z-9 &= 0 \end{aligned}$$

将直线参数方程 $x = y = z = t$ 代入平面 $\pi$ 的方程，得到$t = 3$. 所以投影点为 $M(3,3,3)$.

$$|\overrightarrow{AM}| = \sqrt{2}$$

为距离。

解法2， 应用前面公式：

$$\overrightarrow{s} = (1,1,1) = (m,n,p)$$
$$P = (2,4,3) = (x_1,y_1,z_1)$$
$$t = \frac{2+4+3}{3} = 3 \quad x = tm = 3, y = tn = 3, z = tp = 3$$
$$P_l = (3,3,3) \qquad |\overrightarrow{PP_l}| = \sqrt{2}$$

谢 谢！