§13 投影

13.1 引言

上一讲中，我们得到如下结果：

设 A 为 $m \times n$ 阶阵，$\mathbf{b} \in \mathbb{R}^m$.

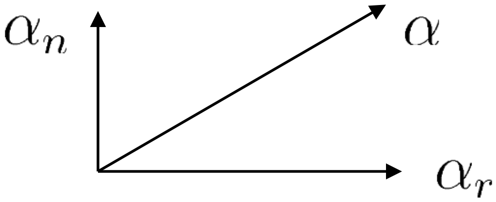
1. $\mathbb{R}^n = C(A^T) + N(A), C(A^T) = N(A)^{\perp}$.

2. $\mathbb{R}^m = C(A) + N(A^T), C(A) = N(A^T)^{\perp}$.

3. 设 $A\mathbf{x} = \mathbf{b}$ 有解，则 $A\mathbf{x} = \mathbf{b}$ 在 $C(A^T)$ 中有唯一解.

设 $\alpha \in \mathbb{R}^n$ 是 $A\mathbf{x} = \mathbf{b}$ 的解. 则 $\alpha = \alpha_r + \alpha_n, \alpha_r \in C(A^T), \alpha_n \in N(A)$.

直观上，



13.1 引言

例：$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. $A\mathbf{x} = \mathbf{b}$ 有解 $\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 4 \\ -2 \end{pmatrix}, \alpha_r = \frac{1}{5}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \alpha_n = \frac{1}{5}\begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

直观上，α_r 是 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 在 $C(A^T) = \left\{ c\begin{pmatrix} 1 \\ 2 \end{pmatrix} | c \in \mathbb{R} \right\}$ 这条直线上投影.

另一方面，若 $A\mathbf{x} = \mathbf{b}$ 无解，此时我们可以考虑问题：

求 $\hat{\mathbf{x}} \in \mathbb{R}^n$, 使得 $||A\hat{\mathbf{x}} - \mathbf{b}||$ 极小(或最小)?

直观上，$A\mathbf{x} = \mathbf{b}$ 无解 $\Longleftrightarrow$ $\mathbf{b} \notin C(A)$. 上述问题意味着求 $C(A)$ 上距离 $\mathbf{b}$ 最近的点 $A\hat{\mathbf{x}}$, 它是 $\mathbf{b}$ 在 $C(A)$ 上的投影点.

13.1 引言

例： $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$ 则 $A\mathbf{x} = \mathbf{b}$ 无解

即 $\mathbf{b} \notin C(A)$(平面 $x + y - z = 0$).

$\mathbf{b}$ 在平面上投影点为 $\mathbf{p} = (p_x, p_y, p_z)^T$.

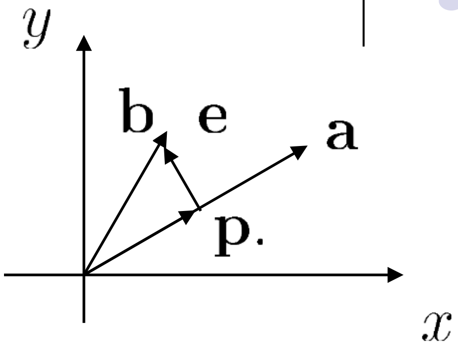
则 $$\left.\begin{array}{l} p_x + p_y = p_z \\ (p_x, p_y, p_z - 1) = \lambda(1, 1, -1) \end{array}\right\} \Longrightarrow \mathbf{p} = (\frac{1}{3}, \frac{1}{3}, \frac{2}{3})^T \in C(A).$$

$$A\hat{\mathbf{x}} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \Longrightarrow \hat{\mathbf{x}} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

这一讲，我们讨论点(或向量)在空间投影问题.

13.2 点在直线和平面上的投影

如右图，我们求 $\mathbf{b}$ 在 $\mathbf{a}$ 上的投影向量 $\mathbf{p}$.



$$\begin{cases} \mathbf{p}+\mathbf{e}=\mathbf{b},\ \mathbf{e}\perp\mathbf{a} \\ \mathbf{p}=t\mathbf{a}(t\in\mathbb{R}) \end{cases}$$

$$\mathbf{e}\perp\mathbf{a}\Longrightarrow\mathbf{a}^T(\mathbf{b}-t\mathbf{a})=0 \quad\Longrightarrow t=\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}(\mathbf{a}\neq\mathbf{0})$$

即 $\mathbf{b}$ 在 $\mathbf{a}$ 上投影向量为 $\left(\dfrac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{a}=\mathbf{p}$.

($\mathbf{a}$, $\mathbf{b}$ 表示相应列向量.)

13.2 点在直线和平面上的投影

$$(\mathbf{a}^T\mathbf{b})\mathbf{a} = \mathbf{a}(\mathbf{a}^T\mathbf{b}) = (\mathbf{a}\mathbf{a}^T)\mathbf{b}.$$

因此，$\mathbf{p} = \left(\dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{b}.$ (注意：$\mathbf{a}\mathbf{a}^T$是一个 2×2 矩阵.)

$S = \dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$称为投影矩阵.

$\forall \mathbf{b} \in \mathbb{R}^2, S\mathbf{b}$ 是$\mathbf{b}$ 在 $\mathbf{a}$ 上投影向量.

设 l 是 $\mathbf{a}$ 所在直线，我们得到一个映射(向量空间之间的映射):

$$\begin{aligned} \mathbb{R}^2 &\longrightarrow l \subset \mathbb{R}^2 \\ \mathbf{b} &\longmapsto S\mathbf{b} \end{aligned}$$

13.2 点在直线和平面上的投影

例： $\mathbf{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\qquad S = \dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$\forall \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, S\mathbf{b} = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}$ 即 $\mathbf{b}$ 在 x 轴上的投影.

13.2 点在直线和平面上的投影

三维空间情形是类似的.

求 $\mathbf{b}=\begin{pmatrix}1\\1\\1\end{pmatrix}$ 在直线 $l=\left\{c\begin{pmatrix}1\\2\\2\end{pmatrix}|c\in\mathbb{R}\right\}$ 上投影 $\mathbf{p}$.

令 $\mathbf{p}=t\mathbf{a}=\begin{pmatrix}t\\2t\\2t\end{pmatrix}$ $\quad \mathbf{p}\perp(\mathbf{b}-\mathbf{p})=\mathbf{e} \implies \begin{pmatrix}t & 2t & 2t\end{pmatrix}\begin{pmatrix}1-t\\1-2t\\1-2t\end{pmatrix}=0$

$\implies t=\dfrac{5}{9}$ $\qquad S=\dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}=\dfrac{1}{9}\begin{pmatrix}1&2&2\\2&4&4\\2&4&4\end{pmatrix}$

S 满足：$S^2=S, S^T=S.$

我们得到一个映射：$\mathbb{R}^3\longrightarrow l\subset\mathbb{R}^3$

$$\mathbf{b}\longmapsto S\mathbf{b}$$

13.2 点在直线和平面上的投影

下面我们考虑点在平面上的投影.

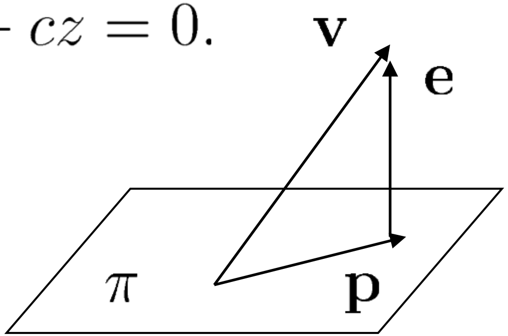
给定 $\mathbf{v}=(v_1,v_2,v_3)^T\in\mathbb{R}^3$, 平面 $\pi: ax+by+cz=0.$

设 $\mathbf{p}$ 是 $\mathbf{v}$ 在 π 上的投影. 求 $\mathbf{p}$.

令 α_1,α_2 是平面 π 上两无关向量, 即

$ax+by+cz=0$ 的基础解系或 $N((a,b,c))$

的一组基.



令 $A=(\alpha_1,\alpha_2)$, 则平面 $\pi=C(A)$.

求投影 $\mathbf{p}\Longleftrightarrow$ 求 $\mathbf{v}$ 关于 $\mathbb{R}^3=C(A)+N(A^T)$ 的分解 $\mathbf{v}_l+\mathbf{v}_{ln}$.

其中，$\mathbf{v}_l=\mathbf{p},\mathbf{v}_{ln}=\mathbf{e}\in N(A^T)$.

13.2 点在直线和平面上的投影

$$\mathbf{p}=\mathbf{v}_l\in C(A)\iff \exists\hat{\mathbf{x}}, A\hat{\mathbf{x}}=\mathbf{p}.$$

$$\mathbf{e}=\mathbf{v}-\mathbf{p}\perp\pi\implies A^T(A\hat{\mathbf{x}}-\mathbf{v})=\mathbf{0}.$$

即 $\hat{\mathbf{x}}$ 是 $A^TA\mathbf{x}=A^T\mathbf{v}$ 的解.

A^TA 是可逆阵(A 列满秩) $\implies \hat{\mathbf{x}}=(A^TA)^{-1}A^T\mathbf{v}.$

则 $\mathbf{p}=A(A^TA)^{-1}A^T\mathbf{v}.$

此时 $A(A^TA)^{-1}A^T$ 称为投影矩阵.

13.2 点在直线和平面上的投影

例：$A=\begin{pmatrix}1&0\\1&1\\1&2\end{pmatrix}, \mathbf{b}=\begin{pmatrix}6\\0\\0\end{pmatrix}$. 求 $\hat{\mathbf{x}}$ 使得 $\mathbf{p}=A\hat{\mathbf{x}}$ 是 $\mathbf{b}$ 在 $C(A)$ 上的投影向量.

解：

$$A^TA=\begin{pmatrix}1&1&1\\0&1&2\end{pmatrix}\begin{pmatrix}1&0\\1&1\\1&2\end{pmatrix}=\begin{pmatrix}3&3\\3&5\end{pmatrix} \qquad A^T\mathbf{b}=\begin{pmatrix}6\\0\end{pmatrix}$$

$$A^TA\hat{\mathbf{x}}=A^T\mathbf{b} \Longrightarrow \begin{pmatrix}3&3\\3&5\end{pmatrix}\begin{pmatrix}\hat{x}_1\\\hat{x}_2\end{pmatrix}=\begin{pmatrix}6\\0\end{pmatrix} \Longrightarrow \hat{\mathbf{x}}=\begin{pmatrix}5\\-3\end{pmatrix} \qquad \mathbf{p}=A\hat{\mathbf{x}}=\begin{pmatrix}5\\2\\-1\end{pmatrix}.$$

注：A^TA 可逆，因为 A 的列线性无关.

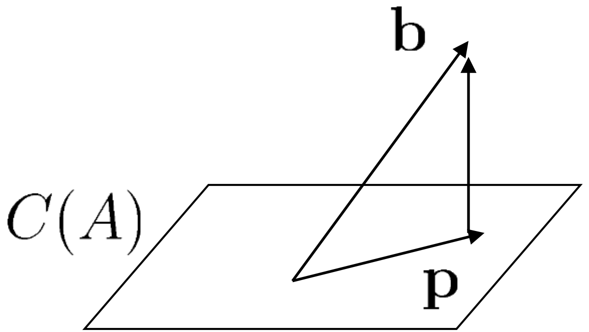
13.3 一般情形

问题：A 为 $m \times n$ 阶阵. 设 $\mathbf{b} \in \mathbb{R}^m$, 求 $\mathbf{b}$ 在 $C(A)$ 上的投影 $\mathbf{p}$?

$\mathbf{p} \in C(A) \Longleftrightarrow \exists \hat{\mathbf{x}} \in \mathbb{R}^n, A\hat{\mathbf{x}} = \mathbf{p}.$

$\mathbf{e} = \mathbf{b} - \mathbf{p} \perp C(A)$ 即 $\mathbf{e} \in N(A^T)$.

$\Longrightarrow A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}.$



即 $\hat{\mathbf{x}}$ 是 $A^TA\mathbf{x} = A^T\mathbf{b}$ 的解. $\mathbf{p} = A\hat{\mathbf{x}}$.

1. $A^TA\mathbf{x} = A^T\mathbf{b}$ 总有解.

2. 设 $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$ 是 $A^TA\mathbf{x} = A^T\mathbf{b}$ 的两个解, 则 $A\hat{\mathbf{x}}_1 = A\hat{\mathbf{x}}_2$.

$\Longrightarrow \mathbf{p}$ 是唯一的. $(\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2 \in N(A^TA) = N(A))$

注：一般情形中，A^TA 未必是可逆阵，除非 A 列满秩.

13.3 一般情形

若 A^TA 可逆，投影阵 $P=A(A^TA)^{-1}A^T$ 满足

$$P^2=P, P^T=P.$$

一般地，一个矩阵 P 满足 $P^2=P, P^T=P,$ 则称 P 为投影矩阵.

自然问题：关于哪个空间的投影矩阵？

检查投影的例子. $\mathbf{b}=\mathbf{p}+\mathbf{e}.$ 设 P 为投影阵，则

$$P\mathbf{p}=\mathbf{p}, P\mathbf{e}=\mathbf{0}.$$

定理：设 P 是一个投影矩阵，则

$$C(P)=N(I-P), N(P)=C(I-P).$$