② 假设 $\exists \vec{\alpha} \in W_2^{\perp}$ 但 $\vec{\alpha} \notin W_1$

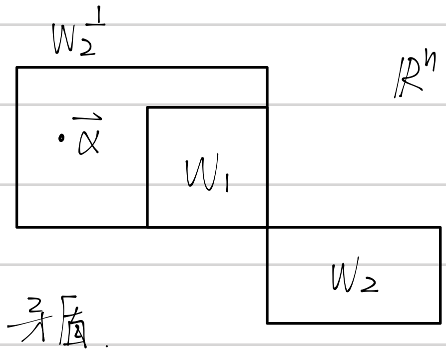
取 $W = \text{span}\{W_1, \vec{\alpha}\}$，则:

$W \perp W_2$，且 $\dim W = \dim W_1 + 1$

但 $\dim W + \dim W_2 = \dim W_1 + \dim W_2 + 1 = n+1 > n$. 矛盾.

$\therefore \forall \vec{\alpha} \in W_2^{\perp}$，有 $\vec{\alpha} \in W_1$. 即 $W_2^{\perp} \subseteq W_1$

综合①,②，$W_1 = W_2^{\perp}$. #




Thm. $\mathbb{R}^n$ 中，$N(A) \perp C(A^T)$，且 $\dim N(A) + \dim C(A^T) = n \Rightarrow N(A)$ 与 $C(A^T)$ 互为 $\mathbb{R}^n$ 中的正交补

$\mathbb{R}^m$ 中，$N(A^{\perp}) \perp C(A)$，且 $\dim N(A^T) + \dim C(A) = m \Rightarrow N(A^T)$ 与 $C(A)$ ··· $\mathbb{R}^m$ ·······

$\forall \vec{x} \in \mathbb{R}^n = N(A) + C(A^{\perp})$，有 $\vec{x} = \vec{x}_n + \vec{x}_r$，其中 $\vec{x}_n \in N(A)$，$\vec{x}_r \in C(A^{\perp})$  Figure 4.3

$A\vec{x} = A\vec{x}_n + \underbrace{A\vec{x}_r}_{\text{this part is invertible.}} = \vec{0} + A\vec{x}_r = A\vec{x}_r \in C(A)$

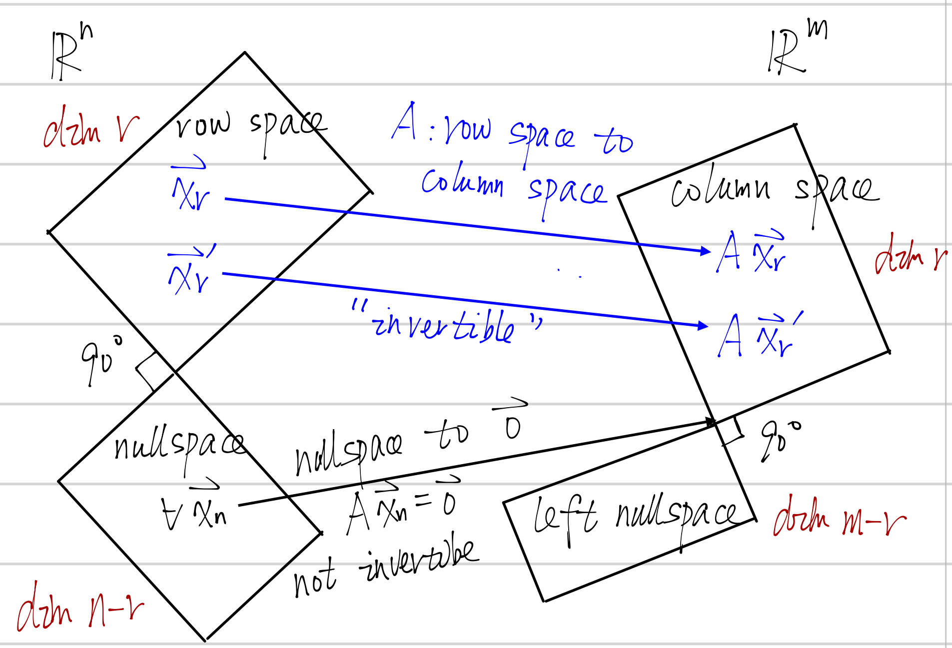
(1) $\forall \vec{x}_n \in$ nullspace $\xrightarrow{\text{映射} A} A\vec{x}_n = \vec{0}$ （多对一）

(2) $\forall \vec{x}_r \in$ row space $\xrightarrow{\text{映射} A} A\vec{x}_r$ in column space

若 $\exists \vec{x}_r, \vec{x}_r' \in C(A^T)$. s.t. $A\vec{x}_r = A\vec{x}_r'$，则

$\begin{cases} A(\vec{x}_r - \vec{x}_r') = \vec{0} \Rightarrow \vec{x}_r - \vec{x}_r' \in N(A) \\ \text{又 } \vec{x}_r - \vec{x}_r' \in C(A^T) \end{cases}$

而 $N(A) \perp C(A^T)$. $\therefore \vec{x}_r = \vec{x}_r'$ （一一映射，可逆）




注. 每个秩r矩阵包含一个 $r \times r$ 可逆子矩阵

(Example 4).

例. $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. $\vec{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \vec{x}_r + \vec{x}_n$ where $\vec{x}_r = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\vec{x}_n = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

$\begin{cases} \vec{x}_r \text{ is in the row space.} \\ \vec{x}_n \text{ is in the null space. } A\vec{x}_n = \vec{0}. \end{cases}$

定理. 设 $\dim V = n$   (HW)

(1) 若 $\alpha_1, \alpha_2, \cdots, \alpha_n \in V$ 线性无关，则 $V = \text{span}\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$. 从而 $\alpha_1, \cdots, \alpha_n$ 为V的一组基.

(2) 若 $V = \text{span}\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$. 则 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关. 从而 $\alpha_1, \cdots, \alpha_n$ 为V的一组基.

注. When the count is right, one property of a basis implies the other.

定理. $A\in M_{n\times n}(\mathbb{R})$. $A=[\vec{\alpha}_1\ \vec{\alpha}_2 \cdots \vec{\alpha}_n]$, $\vec{\alpha}_j\in\mathbb{R}^n$, $\dim\mathbb{R}^n=n$,

(1) 若 $\vec{\alpha}_1,\vec{\alpha}_2,\cdots,\vec{\alpha}_n$ 线性无关，则 $\mathbb{R}^n=\operatorname{span}\{\vec{\alpha}_1,\vec{\alpha}_2,\cdots,\vec{\alpha}_n\}=C(A)$ ($\forall\vec{b}\in\mathbb{R}^n$. $A\vec{x}=\vec{b}$ 有解)

(2) 若 $\mathbb{R}^n=\operatorname{span}\{\vec{\alpha}_1,\vec{\alpha}_2,\cdots,\vec{\alpha}_n\}$，则 $\vec{\alpha}_1,\vec{\alpha}_2,\cdots,\vec{\alpha}_n$ 线性无关. ($A\vec{x}=\vec{b}$ 若有解，则有唯一解)

existence　　　　　　uniqueness

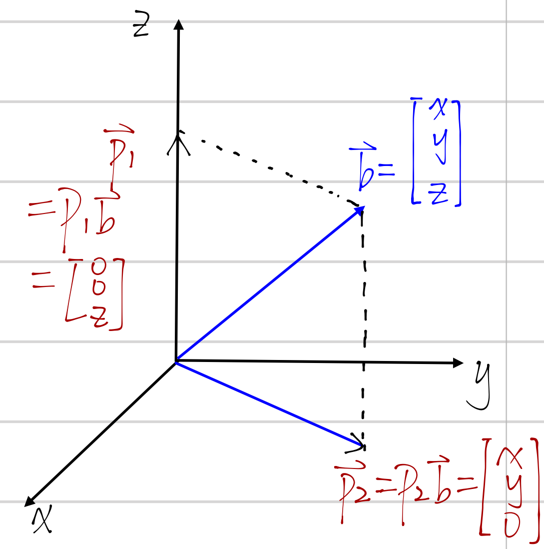
定理. $A\in M_n(\mathbb{R})$. 对 $\forall\vec{b}\in\mathbb{R}^n$，$A\vec{x}=\vec{b}$ 有解 $\Leftrightarrow$ $A\vec{x}=\vec{b}$ 若有解，则有唯一解

Pf. 左 $\Leftrightarrow C(A)=\mathbb{R}^n \Leftrightarrow \operatorname{rank}(A)=n \Leftrightarrow N(A)=\{\vec{0}\}\Leftrightarrow$ 右.　　$A\vec{x}=\vec{b}$ 有唯一解, two parts

$\Leftrightarrow A\to R$ 没零行 $\Leftrightarrow$ $n$ 个主元 $\Leftrightarrow$ 无自由变量 $\Leftrightarrow$

注. 以上都是 $A$ 可逆的充要条件，其它充要条件?

# §4.2 正交投影




## 1. $\mathbb{R}^3$中的投影.

z轴一组基

$W_1$ = z轴: $\mathbb{R}^3$中的1维子空间; $W_1 = C(A_1)$, where $A_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

xy平面一组基

$W_2$ = xy平面: $\mathbb{R}^3$中的2维子空间; $W_2 = C(A_2)$, where $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

z轴与xy平面正交, 且维数之和为3 $\Rightarrow$ 互为$\mathbb{R}^3$中的正交补

$\forall \vec{b} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$

(1) z轴上对$\vec{b}$的最佳逼近? $(x,y,z)$点到z轴上何点最近?

find $\vec{p_1} \in W_1 (= C(A_1))$, s.t. $(\vec{b} - \vec{p_1}) \perp W_1$ (正交投影)

$\vec{b}$在z轴上的投影: $\vec{p_1} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$, 误差 $\vec{b} - \vec{p_1} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \perp$ z轴. $\Rightarrow \vec{b} - \vec{p_1} \in W_1^{\perp} = W_2$

用矩阵$P_1$描述投影? $P_1 : 3\times 3$. $\vec{b} \in \mathbb{R}^3 \to \vec{p_1} = P_1\vec{b} \in \mathbb{R}^3$

$$P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow P_1 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \qquad \mathrm{rank}(P_1) = 1, \quad P_1^2 = P_1$$

(2) $\vec{b}$到xy平面的投影? find $\vec{p_2} \in W_2 (= C(A_2))$, s.t. $(\vec{b} - \vec{p_2}) \perp W_2$

$\vec{p_2} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in W_2$, 误差 $\vec{b} - \vec{p_2} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \perp W_2 \Rightarrow \vec{b} - \vec{p_2} \in W_2^{\perp} = W_1$

矩阵$P_2 : 3\times 3$. $\vec{b} \in \mathbb{R}^3 \to \vec{p_2} = P_2\vec{b} \in \mathbb{R}^3$

$$P_2 = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} \Rightarrow P_2 \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}. \qquad \mathrm{rank}(P_2) = 2. \quad P_2^2 = P_2.$$

(3) $W_1$与$W_2$互为$\mathbb{R}^3$中正交补. $W_1 + W_2 = \mathbb{R}^3$.

$\forall \vec{b} \in \mathbb{R}^3 = W_1 + W_2$, $\exists \vec{p_1} \in W_1, \vec{p_2} \in W_2$, s.t. $\vec{b} = \vec{p_1} + \vec{p_2}$.

$\vec{p_1}, \vec{p_2}$ are just projections of $\vec{b}$ in (1), (2).

$P_1$的投影误差

Notice that: $P_1 + P_2 = I_3$. $\vec{b} = P_1\vec{b} + P_2\vec{b} = \vec{p_1} + \vec{p_2}$

当$P_1$将$\vec{b}$投影到$W_1$中的$\vec{p_1}$时, $I_3 - P_1$恰好将$\vec{b}$投影到$W_1^{\perp}$中的向量 $\vec{b} - \vec{p_1} = \vec{p_2}$.

$\vec{p_1}$为$W_1$中对$\vec{b}$的最佳逼近, 误差 $\vec{b} - \vec{p_1}$ 为正交补 $W_1^{\perp}$对$\vec{b}$的最佳逼近.

(1), (2)计算其一即可

## 2. $\mathbb{R}^m$ 中的正交投影.

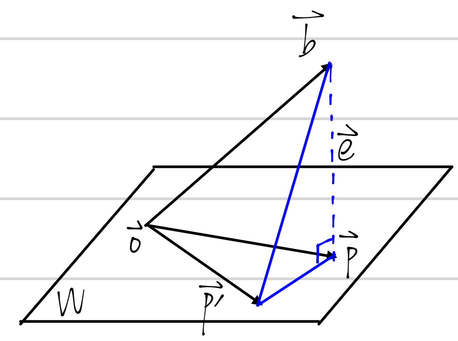
W 为 $\mathbb{R}^m$ 的 n 维子空间 ($n \le m$), 取 W 一组基 $\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_n \in \mathbb{R}^m$, 令 $A = [\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_n]$, 则

$W = \operatorname{span}\{\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_n\} = C(A)$.

定义. $\forall \vec{b} \in \mathbb{R}^m$, 则 $\vec{b}$ 到 W 上的投影 $\vec{p}$ 满足: $\vec{p} \in W$, 且 $\vec{e} \perp W$, 其中 $\vec{e} := \vec{b} - \vec{p}$ 为投影误差.

若 $\vec{p} = P\vec{b}$, 称 $P \in M_{m\times m}(\mathbb{R})$ 为投影矩阵.

定理. $\vec{b}$ 到 W 上的投影 $\vec{p}$ 是 W 中距离 $\vec{b}$ 最近的点, 即误差的长度最小.




Pf. $\forall \vec{p}' \in W$, 有 $\vec{b} - \vec{p}' = \vec{b} - \vec{p} + \vec{p} - \vec{p}' = \vec{e} + (\vec{p} - \vec{p}')$

又 $\vec{e} \perp (\vec{p} - \vec{p}')$

$\Rightarrow \|\vec{b} - \vec{p}'\|^2 = \|\vec{e}\|^2 + \|\vec{p} - \vec{p}'\|^2 \geq \|\vec{e}\|^2$.

当 $\vec{p}' = \vec{p}$ 时, $\|\vec{b} - \vec{p}'\|^2$ 达到最小. #.

定理. 当投影矩阵 P 将 $\vec{b} \in \mathbb{R}^m$ 投影到子空间 $W = C(A)$ 中的 $\vec{p} = P\vec{b}$ 时, 其投影误差 $\vec{e} := \vec{b} - \vec{p}$ 即为 $\vec{b}$ 在子空间 $W^{\perp} = N(A^T)$ 上的投影, 且对应的投影矩阵为 $I_m - P$.

Pf. 由题设, 知: $\vec{p} = P\vec{b} \in W$ 且 $\vec{e} \perp W$.

$\therefore \vec{e} \in W^{\perp}$, 且 $\vec{b} - \vec{e} = \vec{p} \perp W^{\perp}$ $\Rightarrow$ $\vec{e}$ 是 $\vec{b}$ 在子空间 $W^{\perp}$ 上的投影.

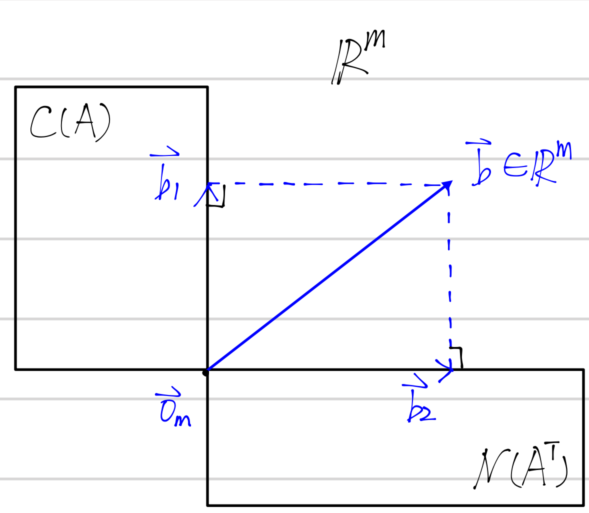
易看出. $\vec{e} = \vec{b} - \vec{p} = (I_m - P)\vec{b}$, $\therefore$ 对应的投影矩阵为 $I_m - P$. #.

定理. 已知 $\mathbb{R}^m = W + W^{\perp} = C(A) + N(A^T)$, $\forall \vec{b} \in \mathbb{R}^m$, 其可"唯一"地分解成 $\vec{b} = \vec{b}_1 + \vec{b}_2$, $\vec{b}_1 \in C(A)$, $\vec{b}_2 \in N(A^T)$

更进一步地, $\vec{b}_1$ 和 $\vec{b}_2$ 分别为 $\vec{b}$ 在 $C(A)$ 和 $N(A^T)$ 上的投影.

Pf. 取 $\vec{b}_1$ 为 $\vec{b}$ 在 $C(A)$ 上的投影, 令 $\vec{b}_2 = \vec{b} - \vec{b}_1$, 则 $\vec{b}_2 := \vec{b} - \vec{b}_1$ 为 $\vec{b}$ 在 $N(A^T)$ 上的投影.

即 $\vec{b} = \vec{b}_1 + \vec{b}_2$ 且 $\vec{b}_1 \in C(A)$, $\vec{b}_2 \in N(A^T)$.




"唯一性": 若 $\vec{b} = \vec{b}_1' + \vec{b}_2'$, 其中 $\vec{b}_1' \in C(A)$, $\vec{b}_2' \in N(A^T)$

则 $\vec{0} = (\vec{b}_1 - \vec{b}_1') + (\vec{b}_2 - \vec{b}_2')$

$\Rightarrow \vec{b}_1 - \vec{b}_1' = -(\vec{b}_2 - \vec{b}_2') \in N(A^T)$

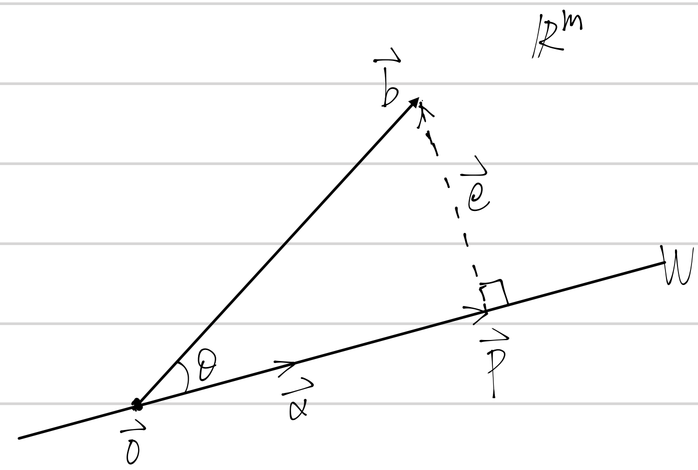
而 $\vec{b}_1 - \vec{b}_1' \in C(A)$

$\Rightarrow \vec{b}_1 - \vec{b}_1' \in C(A) \cap N(A^T)$

而 $C(A) \perp N(A^T)$, $C(A) \cap N(A^T) = \{\vec{0}\}$

$\Rightarrow \vec{b}_1 = \vec{b}_1'$, 从而 $\vec{b}_2 = \vec{b}_2'$. #.

## 3. 求解 $\hat{x}$ 及投影矩阵 P.




(1) 一维子空间: $W=\text{span}\{\vec\alpha\}$ is a line in $\mathbb{R}^m$.

① 求 $\vec b\in\mathbb{R}^m$ 在 W 中的投影 $\vec p\in W$.

设 $\vec p=\hat{x}\vec\alpha$, 则投影误差 $\vec e:=\vec b-\vec p=\vec b-\hat x\vec\alpha$

$\vec e\perp\vec\alpha \Rightarrow (\vec b-\hat x\vec\alpha)\cdot\vec\alpha=0 \Rightarrow \hat x=\frac{\vec\alpha\cdot\vec b}{\vec\alpha\cdot\vec\alpha}=\frac{\vec\alpha^T\vec b}{\vec\alpha^T\vec\alpha}$ (实数)

$\therefore \vec p=\frac{\vec\alpha^T\vec b}{\vec\alpha^T\vec\alpha}\vec\alpha$

Remark: $\hat x=\frac{\|\vec\alpha\|\cdot\|\vec b\|\cos\theta}{\|\vec\alpha\|^2}=\|\vec b\|\cos\theta\cdot\frac{1}{\|\vec\alpha\|}$

$\therefore \vec p=\underbrace{\|\vec b\|\cos\theta}_{\vec p\text{的长度}}\cdot\frac{\vec\alpha}{\|\vec\alpha\|}$ ← W中单位向量.

② 求投影矩阵 P s.t. $\vec p=P\vec b$

$\vec p=\vec\alpha\hat x=\vec\alpha\frac{\vec\alpha^T\vec b}{\vec\alpha^T\vec\alpha}=\frac{\vec\alpha\vec\alpha^T}{\vec\alpha^T\vec\alpha}\vec b$

$\therefore P=\frac{\vec\alpha\vec\alpha^T}{\vec\alpha^T\vec\alpha}$ ← a rank one matrix (numerator); ← a real numble (denominator). $P\in M_{m\times m}(\mathbb{R})$, $\text{rank}(P)=1$, $P^2=P$.

We are projecting onto a one-dimensional subspace. The line is the column space of P.

(2) n维子空间 $(n\le m)$ $W=\text{span}\{\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n\}=C(A)$, $\vec\alpha_i\in\mathbb{R}^m$, $i=1,2,\cdots,n$.

① 求 $\vec b\in\mathbb{R}^m$ 在 W 中的投影 $\vec p\in W$.

设 $\vec p=\hat x_1\vec\alpha_1+\hat x_2\vec\alpha_2+\cdots+\hat x_n\vec\alpha_n=A\hat{\vec x}\in C(A)$, 则投影误差为 $\vec e=\vec b-\vec p=\vec b-A\hat{\vec x}$.

$\vec e\perp C(A)=\text{span}\{\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n\}$:

(i) 几何角度: $\vec e\perp\vec\alpha_i$, $i=1,\cdots,n$. 于是有 $\begin{bmatrix}\vec\alpha_1^T\\ \vec\alpha_2^T\\ \vdots\\ \vec\alpha_n^T\end{bmatrix}(\vec b-A\hat{\vec x})=\vec 0$

即 $A^T(\vec b-A\hat{\vec x})=\vec 0 \Rightarrow (A^TA)\hat{\vec x}=A^T\vec b$  normal equation

↑ $A\vec x=\vec b$ 基础上左乘 $A^T$ ⇒ solvable.

(ii) 代数角度: $\vec e\in C(A)^\perp=N(A^T)$. $\therefore A^T\vec e=\vec 0$.

即 $(A^TA)\hat{\vec x}=A^T\vec b$.

$\vec\alpha_1,\vec\alpha_2,\cdots,\vec\alpha_n$ 线性无关 $\Rightarrow A^TA\in M_{n\times n}$ 可逆. (后证)

$\Rightarrow \hat{\vec x}=(A^TA)^{-1}A^T\vec b$.

$\Rightarrow \vec p=A\hat{\vec x}=A(A^TA)^{-1}A^T\vec b$

② 投影矩阵 $P=A(A^TA)^{-1}A^T$. $(m\times m)$  $P^2=P$, $P^T=P$.

注. · 当 $n=1$ 时, 与(1)中结果相符.

· 上述问题也可理解为: 当 $\vec b\notin C(A)$ 时, $A\vec x=\vec b$ 无解. 求 $\vec x(=\hat{\vec x})$, s.t. $\|\vec b-A\vec x\|$ 最小