§14 最小二乘逼近

14.1 复习

回忆一向量 $\alpha \in \mathbb{R}^m$ 在一子空间 $V \subset \mathbb{R}^m$ 上投影 α_p, 求法:

1.找 V 的一组基 $\beta_1, \cdots, \beta_k$. 令 $M = (\beta_1, \beta_2, \cdots, \beta_k)$,
则 M 列满秩, 且 $V = C(M)$(M 的列空间).

2. $\alpha_p \in C(M)$, 则 $\exists \mathbf{x} \in \mathbb{R}^k, \alpha_p = M\mathbf{x}$.
$\alpha - \alpha_p \perp C(M)$, 则 $(\alpha - \alpha_p)^T M = M^T(\alpha - \alpha_p) = \mathbf{0}$.
$\Longrightarrow M^T M\mathbf{x} = M^T\alpha$.

3.解方程组 $M^T M\mathbf{x} = M^T\alpha$, 即 $\mathbf{x} = (M^T M)^{-1} M^T\alpha$.

因此 $\alpha_p = M(M^T M)^{-1} M^T\alpha$.

14.1 复习

注：

1. M列满秩 $\Longleftrightarrow$ $M\mathbf{y}=\mathbf{0}$只有零解

$\Longleftrightarrow$ $M^TM\mathbf{y}=\mathbf{0}$只有零解

$\Longleftrightarrow$ M^TM 列满秩(即M^TM 可逆，因为它是方阵)

2. $P=M(M^TM)^{-1}M^T$ 称为投影矩阵. 一个矩阵 P 称为投影阵，如果 $P^2=P, P^T=P.$

第二个条件$P^T=P$是因为对$\forall\alpha,\beta\in\mathbb{R}^m$,我们需要保证

$P\alpha\perp(I-P)\beta\Longleftrightarrow P^T(I-P)=(I-P)^TP=0$

即$P^T=P.$

14.1 复习

命题：设 P 为 n 阶投影阵, 则 $C(P)=N(I-P), C(I-P)=N(P)$.

证明: $P^2=P$, 则 $P(I-P)=0 \Longrightarrow C(I-P)\subset N(P)$.

$\forall \alpha\in\mathbb{R}^n, \alpha=P\alpha+(I-P)\alpha$

$$\Longrightarrow \mathbb{R}^n=C(I-P)+C(P)=N(P)+C(P).$$

设 $\alpha\in N(P)$, 则 $\alpha=(I-P)\alpha \Longrightarrow N(P)\subset C(I-P)$.

14.1 复习

3.一般地，设 A 为 $m\times n$ 阶阵，则 $\mathbb{R}^m=C(A)+N(A^T)$.

我们求 $\forall\alpha\in\mathbb{R}^m$ 关于这个和的分解 $\alpha=\alpha_1+\alpha_2\in C(A)+N(A^T)$.

使用以上方法，即找 $C(A)$ 的一组基 $\beta_1,\cdots,\beta_r$.

令 $M=(\beta_1,\cdots,\beta_r), C(A)=C(M)$. 求 α 在 $C(M)$ 上的投影.

14.2 最小二乘法

回到解方程组 $A_{m\times n}\mathbf{x}=\mathbf{b}$.

$$A\mathbf{x}=\mathbf{b}\text{有解}\Longleftrightarrow \mathbf{b}\in C(A).$$

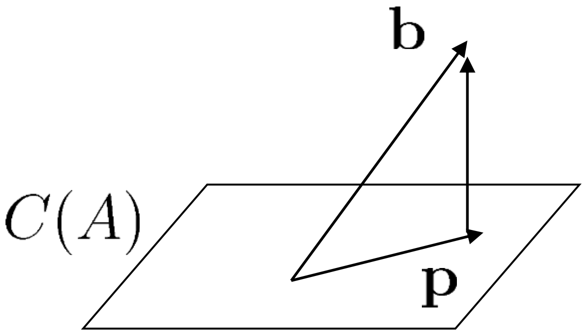
假设它无解，则 $\mathbf{b}\notin C(A)$.

此时转化问题为：

求 $\hat{\mathbf{x}}$ 使得$||A\hat{\mathbf{x}}-\mathbf{b}||$最小，即 $\min\limits_{\mathbf{x}\in\mathbb{R}^n}||A\hat{\mathbf{x}}-\mathbf{b}||$ 的最小值点.

14.2 最小二乘法

由右图，$\mathbf{b}$ 和它在 $C(A)$上投影点 $\mathbf{p}$ 距离最小. 设 $\mathbf{p} = A\hat{\mathbf{x}}$, 则 $\hat{\mathbf{x}}$ 即为所求.



$\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$ 称为误差向量(error),

$\hat{\mathbf{x}}$ 最小二乘解(the least square solution).

如何求 $\mathbf{p}$?

1. $\mathbf{p} = A\hat{\mathbf{x}}$.

2. $\mathbf{e} \perp C(A) \Longleftrightarrow A^T(\mathbf{b} - \mathbf{p}) = \mathbf{0} \Longrightarrow A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$.

因此，求 $\mathbf{p} \Longleftrightarrow$ 解方程组 $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$.

$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ 称为法方程组(normal equations).

14.2 最小二乘法

性质：

1.法方程组总有解(无论 A 是否列满秩).

这是因为$C(A^T)=C(A^TA), A^T\mathbf{b}\in C(A^T)=C(A^TA)$.

2. $A^TA\hat{\mathbf{x}}=A^T\mathbf{b}$ 的解可能有无穷个, 但 $A\hat{\mathbf{x}}$ (投影 $\mathbf{p}$)唯一.

即设 $\hat{\alpha},\hat{\beta}$ 均满足$A^TA\hat{\mathbf{x}}=A^T\mathbf{b}$, 则 $A^TA(\hat{\alpha}-\hat{\beta})=\mathbf{0}$

$$\Longrightarrow \hat{\alpha}-\hat{\beta}\in N(A^TA)=N(A)$$

因此 $A(\hat{\alpha}-\hat{\beta})=\mathbf{0}\Longrightarrow A\hat{\alpha}=A\hat{\beta}$.

14.2 最小二乘法

例：$A=\begin{pmatrix}2&0\\-1&1\\0&2\end{pmatrix},\mathbf{b}=\begin{pmatrix}1\\0\\-1\end{pmatrix}$. 则 $\begin{pmatrix}2&0\\-1&1\\0&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}1\\0\\-1\end{pmatrix}$ 无解.

考虑法方程组 $A^TA\hat{\mathbf{x}}=A^T\mathbf{b}$ 即 $\begin{pmatrix}5&-1\\-1&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2\\-2\end{pmatrix}$.

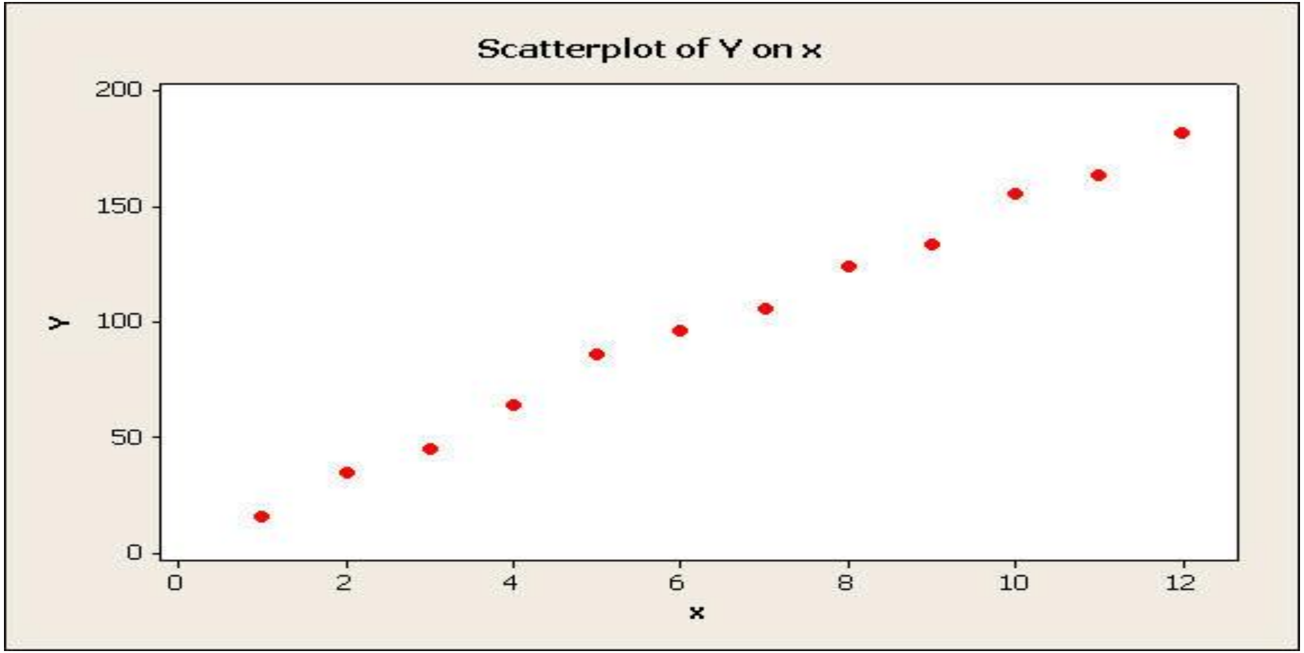
$$\Longrightarrow\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}\frac{1}{3}\\-\frac{1}{3}\end{pmatrix}.$$

$\mathbf{b}$ 在 $C(A)$ 上投影 $\mathbf{p}=A\hat{\mathbf{x}}=\begin{pmatrix}\frac{2}{3}\\-\frac{2}{3}\\-\frac{2}{3}\end{pmatrix}$.

3.若 A 列满秩，则 A^TA 可逆，$\hat{\mathbf{x}}=(A^TA)^{-1}A^T\mathbf{b}$.

14.3 最小二乘法的应用：曲线拟合

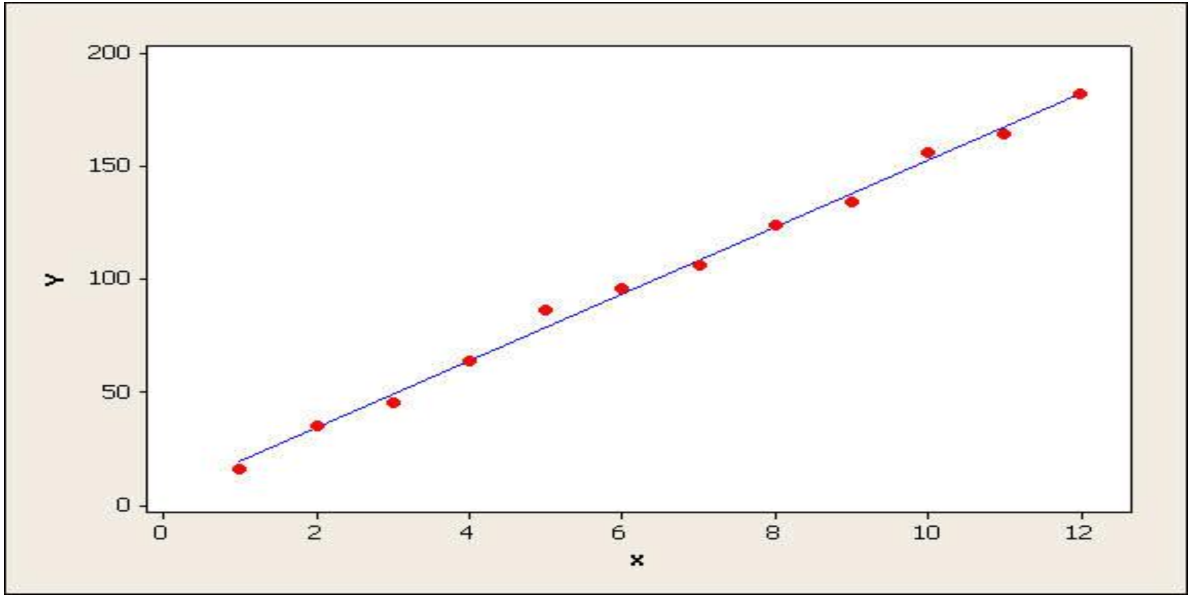
x	1	2	3	4	5	6	7	8	9	10	11	12
y	16	35	45	64	86	96	106	124	134	156	164	182



14.3 最小二乘法的应用：曲线拟合

$$\hat{y} = a + bx$$

$$e_i = y_i - \hat{y}_i$$



14.3 最小二乘法的应用：曲线拟合

例：设给定数据 $\{(x_1, y_1), \cdots, (x_N, y_N)\}$.

寻找直线 $y = C + Dx$, 使得误差

$$E(C, D) = [y_1 - (C + Dx_1)]^2 + \cdots + [y_N - (C + Dx_N)]^2$$

最小.

即向量 $\begin{pmatrix} y_1 - (C + Dx_1) \\ \vdots \\ y_N - (C + Dx_N) \end{pmatrix}$ 的长度最小.

14.3 最小二乘法的应用：曲线拟合

令 $A=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_N\end{pmatrix}, \mathbf{b}=\begin{pmatrix}y_1\\ \vdots\\ y_N\end{pmatrix}, \hat{\mathbf{x}}=\begin{pmatrix}\hat{C}\\ \hat{D}\end{pmatrix}.$

即求 $\hat{\mathbf{x}}$ 使得$||\mathbf{b}-A\hat{\mathbf{x}}||$最小.

解法方程组 $A^TA\hat{\mathbf{x}}=A^T\mathbf{b}$ 即 $\begin{pmatrix}N & \sum_{i=1}^N x_i\\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2\end{pmatrix}\begin{pmatrix}\hat{C}\\ \hat{D}\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^N y_i\\ \sum_{i=1}^N x_iy_i\end{pmatrix}.$

令 $\bar{x}=\frac{1}{N}\sum_{i=1}^N x_i, \bar{y}=\frac{1}{N}\sum_{i=1}^N y_i$, 求得 $\hat{C}=\bar{y}-\hat{D}\bar{x}, \hat{D}=\frac{x_1y_1+\cdots+x_Ny_N-N\bar{x}\bar{y}}{x_1^2+\cdots+x_N^2-N\bar{x}^2}.$

直线 $y=\hat{C}+\hat{D}x$ 称为最小二乘直线.

特别地，若 $x_1+\cdots+x_N=0$, 则 $\hat{\mathbf{x}}$ 很容易计算，此时 A 两列正交, A^TA 是对角阵.

14.3 最小二乘法的应用：曲线拟合

例：求二次曲线拟合如下数据

x	19	25	31	38	44
y	19.0	32.3	49.0	73.3	97.8

解：求 $y = a + bx + cx^2$ 使得

$$E(a,b,c) = \sum_{k=1}^{5} [y_k - (a + bx_k + cx_k^2)]^2$$

最小.

14.3 最小二乘法的应用：曲线拟合

令 $A=\begin{pmatrix}1 & x_1 & x_1^2\\ \vdots & \vdots & \vdots\\ 1 & x_5 & x_5^2\end{pmatrix}, \mathbf{b}=\begin{pmatrix}y_1\\ \vdots\\ y_5\end{pmatrix}, \hat{\mathbf{x}}=\begin{pmatrix}a\\ b\\ c\end{pmatrix}.$

即求 $\hat{\mathbf{x}}$ 使得 $||\mathbf{b}-A\hat{\mathbf{x}}||$最小.

求解方程组 $A^TA\hat{\mathbf{x}}=A^T\mathbf{b} \implies \begin{pmatrix}a\\ b\\ c\end{pmatrix}=\begin{pmatrix}0.6882\\ 0.0193\\ 0.0497\end{pmatrix}.$

拟合曲线为 $y=0.6882+0.0193x+0.0497x^2$.

14.3 最小二乘法的应用：曲线拟合

最后，我们说明法方程组也来自于微积分.

求 $\min\limits_{\mathbf{x}\in\mathbb{R}^n} ||A\mathbf{x}-\mathbf{b}||$.　　　令 $f(x_1,\cdots,x_n)=||A\mathbf{x}-\mathbf{b}||^2$

$$
\begin{aligned}
&= (A\mathbf{x}-\mathbf{b})^T(A\mathbf{x}-\mathbf{b}) \\
&= \mathbf{x}^TA^TA\mathbf{x}-\mathbf{x}^TA^T\mathbf{b}-\mathbf{b}^TA\mathbf{x}+\mathbf{b}^T\mathbf{b}
\end{aligned}
$$

则 $\frac{\partial f}{\partial \mathbf{x}} := \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = 2A^TA\mathbf{x}-2A^T\mathbf{b}.$

若 $\hat{\mathbf{x}}$ 满足 $||A\hat{\mathbf{x}}-\mathbf{b}|| = min||A\mathbf{x}-\mathbf{b}||$, 则 $\frac{\partial f}{\partial \mathbf{x}}|_{\mathbf{x}=\hat{\mathbf{x}}}=\mathbf{0}.$

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

14.3 最小二乘法的应用：曲线拟合

例：$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \mathbf{b} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}.$

$$f(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \cdots$$

可以验证，$\frac{\partial f}{\partial \mathbf{x}} = 2 \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2 \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}.$