第七节　向量到子空间的距离

最小二乘法

主要内容

- 定义
- 向量到子空间各向量间的最短距离
- 最小二乘法

一、定义

在解析几何中，两个点 α 和 β 间的距离等于向量 $\alpha-\beta$ 的长度. 在欧氏空间中同样可引入

定义 13 长度 $|\alpha-\beta|$ 称为向量 α 和 β 的距离记为 $d(\alpha,\beta)$.

容易证明距离的三条基本性质：

1) $d(\alpha,\beta)=d(\beta,\alpha)$；

2) $d(\alpha,\beta)\geq 0$，并且仅当 $\alpha=\beta$ 时等号才成立;

3) $d(\alpha,\beta)\leq d(\alpha,\gamma)+d(\gamma,\beta)$ (三角不等式).

二、向量到子空间各向量间的最短距离

在中学所学几何中知道一个点到一个平面(或一条直线)上所有点的距离以垂线最短. 下面可证明一个固定向量和一个子空间中各向量的距离也是以“垂线最短”

先设一个子空间 W，它是由向量 $\alpha_1, \alpha_2, \ldots, \alpha_k$ 所生成，即 $W = L(\alpha_1, \alpha_2, \ldots, \alpha_k)$．说一个向量 α 垂直于子空间 W，就是指向量 α 垂直于 W 中任何一个向量．容易验证 α 垂直于 W 的充分必要条件是 α 垂直于每个α_i $(i = 1, 2, \ldots, k)$．

下证向量到子空间各向量间的距离以垂线最短.

设 β 是给定的一向量，γ 是 W 中的向量，且满足 $\beta-\gamma$ 垂直于 W. 要证明 β 到 W 中各向量的距离以垂线最短，就是要证明，对 W 中任一向量 δ，有

$$|\beta-\gamma|\leq|\beta-\delta|.$$

可以画出下面的示意图:

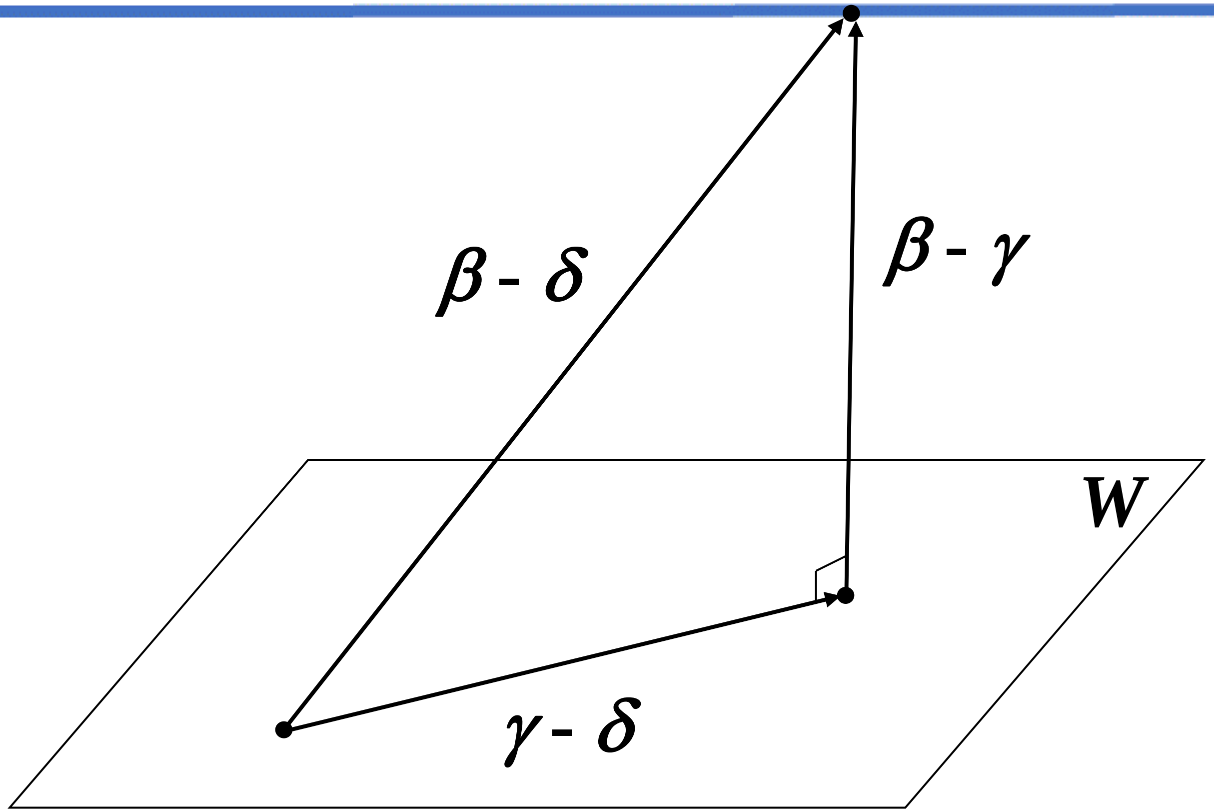

图 9-2

证明　$\beta-\delta=(\beta-\gamma)+(\gamma-\delta)$．因 W 是子空间，$\gamma\in W$，$\delta\in W$，则 $\gamma-\delta\in W$．故 $\beta-\gamma$ 垂直于 $\gamma-\delta$．由勾股定理，有

$$|\beta-\gamma|^2+|\gamma-\delta|^2=|\beta-\delta|^2,$$

故

$$|\beta-\gamma|\le|\beta-\delta|.$$

证毕

三、最小二乘法

1. 引例

引例 已知某种材料在生产过程中的废品率 y 与某种化学成分 x 有关. 下列表中记载了某工厂生产中 y 与相应的 x 的几次数值:

y (%)	1.00	0.9	0.9	0.81	0.60	0.56	0.35
x (%)	3.6	3.7	3.8	3.9	4.0	4.1	4.2

我们想找出 y 对 x 的一个近似公式.

解 把表中数值画出图来看，发现它的变化趋势近于一条直线. 因此我们决定选取 x 的一次式 $ax+b$ 来表达. 当然最好能选到适当的 a,b 使得下面的等式

$$3.6a+b-1.00=0\,,$$

$$3.7a+b-0.9=0\,,$$

$$3.8a+b-0.9=0,$$

$$3.9a+b-0.81=0,$$

$$4.0a+b-0.60=0,$$

$$4.1a+b-0.56=0,$$

$$4.2a+b-0.35=0$$

都成立．实际上是不可能的．任何 a,b 代入上面各式都会发生些误差．于是想找 a,b 使得上面各式的误差的平方和最小，即找 a,b 使

$$(3.6a + b - 1.00)^2 + (3.7a + b - 0.9)^2$$

$$+ (3.8a + b - 0.9)^2 + (3.9a + b - 0.81)^2$$

$$+ (4.0a + b - 0.60)^2 + (4.1a + b - 0.56)^2$$

$$+ (4.2a + b - 0.35)^2$$

最小. 这里讨论的是误差的平方即二乘方, 故称为最小二乘法. 现在转向一般的最小二乘法问题.

2. 定义

定义14 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s - b_1 = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s - b_2 = 0, \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ns}x_s - b_n = 0 \end{cases}$$

可能无解. 即任何一组数 $x_1, x_2, \ldots, x_s$ 都可能使

$$\sum_{i=1}^{n}(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{is}x_s - b_i)^2 \qquad (1)$$

不等于零. 法找 $x_1^0, x_2^0, \ldots, x_s^0$ 使(1)最小,这样的 $x_1^0, x_2^0, \ldots, x_s^0$ 称为方程组的**最小二乘解**.

这种问题就叫做**最小二乘法问题**.

3. 最小二乘法的代数表示

下面利用欧氏空间的概念来表达最小二乘法，并给出最小二乘解所满足的条件. 令

$$
\left.
\begin{aligned}
&A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ns} \end{pmatrix},\ B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \\
&X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix},\ Y=\begin{pmatrix} \sum_{j=1}^{s} a_{1j}x_j \\ \sum_{j=1}^{s} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{s} a_{nj}x_j \end{pmatrix}=AX.
\end{aligned}
\right\} \qquad \textbf{(2)}
$$

用距离的概念，(1) 就是

$$|Y-B|^2.$$

最小二乘法就是找 $x_1^0, x_2^0, \ldots, x_s^0$ 使 Y 与 B 的距离最短. 但从(2) 知道向量 Y 就是

$$Y = x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \cdots + x_s\begin{pmatrix} a_{1s} \\ a_{2s} \\ \vdots \\ a_{ns} \end{pmatrix}.$$

文山学院 WENSHAN UNIVERSITY

把 A 的各列向量分别记成 $\alpha_1, \alpha_2, \ldots, \alpha_s$. 由它们生成的子空间为 $L(\alpha_1, \alpha_2, \ldots, \alpha_s)$. Y 就是 $L(\alpha_1, \alpha_2, \ldots, \alpha_s)$ 中的向量. 于是最小二乘法问题可叙述成:

找 X 使(1)最小，就是在 $L(\alpha_1, \alpha_2, \ldots, \alpha_s)$ 中找一向量 Y，使得 B 到它的距离比到子空间 $L(\alpha_1, \alpha_2, \ldots, \alpha_s)$ 中其他向量的距离都短.

4. 最小二乘解的求法

应用前面所讲的结论，设

$$Y = AX = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_s\alpha_s$$

是所要求的向量，则

$$C = B - Y = B - AX$$

必须垂直于子空间 $L(\alpha_1, \alpha_2, \ldots, \alpha_s)$. 为此只须

而且必须

$$(C, \alpha_1) = (C, \alpha_2) = \ldots = (C, \alpha_s) = 0.$$

回忆矩阵乘法规则，上述一串等式可以写成矩阵相乘的式子，即

$$\alpha_1^{\mathrm{T}} C = 0\,,\ \alpha_2^{\mathrm{T}} C = 0\,,\ \ldots\,,\ \alpha_s^{\mathrm{T}} C = 0\,.$$

而 $\alpha_1^{\mathrm{T}}, \alpha_2^{\mathrm{T}}, \ldots, \alpha_s^{\mathrm{T}}$ 按行正好排成矩阵 A^{T}，上述一串等式合起来就是

$$A^{\mathrm{T}}(B - AX) = 0\,,$$

或

$$A^{\mathrm{T}} A X = A^{\mathrm{T}} B\,.$$

这就是最小二乘解所满足的代数方程，它是一个线性方程组，系数矩阵是$A^{\mathrm{T}}A$，常数项是$A^{\mathrm{T}}B$. 这种线性方程组总是有解的.（参见第5章习题17）

现在回到之前的例子，得

$$A=\begin{pmatrix} 3.6 & 1 \\ 3.7 & 1 \\ 3.8 & 1 \\ 3.9 & 1 \\ 4.0 & 1 \\ 4.1 & 1 \\ 4.2 & 1 \end{pmatrix}, \qquad B=\begin{pmatrix} 1.00 \\ 0.90 \\ 0.90 \\ 0.81 \\ 0.60 \\ 0.56 \\ 0.35 \end{pmatrix}.$$

最小二乘解 a,b 所满足的方程是

$$A^{\mathrm{T}}A\begin{pmatrix} a \\ b \end{pmatrix} - A^{\mathrm{T}}B = 0,$$

即为

$$\begin{cases} 106.75a + 27.3b - 19.675 = 0, \\ 27.3a + 7b - 5.12 = 0. \end{cases}$$

解得

a = -1.05 , b = 4.81 (取三位有效数字).