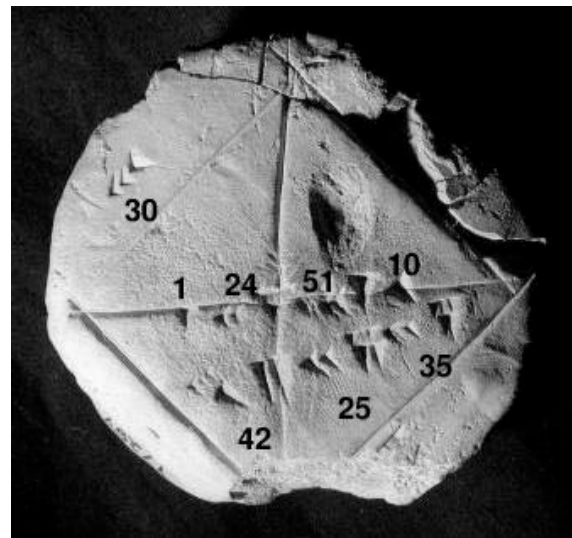


计算方法 (A)



童伟华 管理科研楼1205室

E-mail: tongwh@ustc.edu.cn

中国科学技术大学 数学科学学院

<http://math.ustc.edu.cn/>



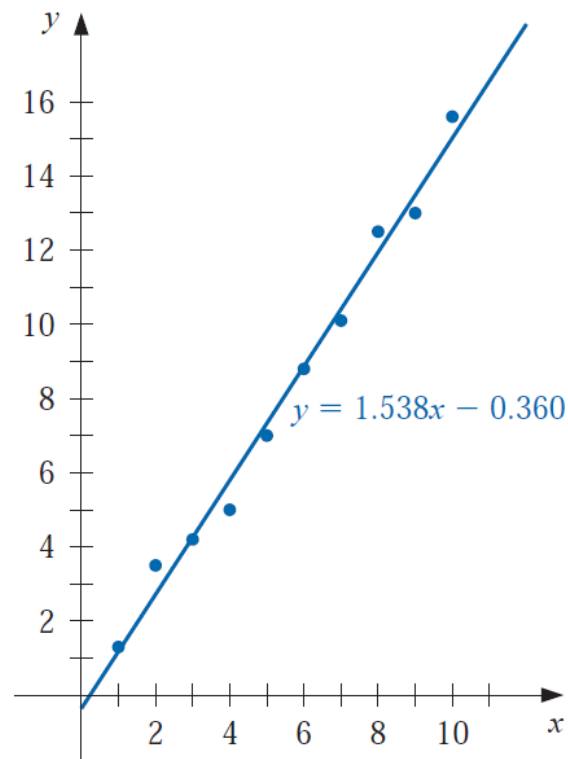


第二章 最小二乘拟合

拟合函数



- 在实际问题中，数据不可避免的会有误差，插值函数会将这些误差也包括在内
- 若误差较大时，用插值方法构造函数的效果不好
- 用逼近的方法构造函数
 - 不要求过所有的点（可以减少误差影响）
 - 尽可能表现数据的趋势，靠近这些点



- 如何刻画数据点和逼近函数之间的靠近程度?
- 给定一组观察或测量得到的一组离散数据序列 $\{(x_i, y_i)\}_{i=1}^m$ ，选定函数空间 Φ ，构造函数 $\varphi(x) \in \Phi$ ，使得函数 $\varphi(x)$ 最优的靠近离散数据序列 $\{(x_i, y_i)\}_{i=1}^m$ ，即向量

$Q = (\varphi(x_1), \varphi(x_2), \dots, \varphi(x_m))^T$ 与 $Y = (y_1, y_2, \dots, y_m)^T$ 的误差或距离最小

- 如何定义向量之间的距离：向量范数

$$R_1 = \|Q - Y\|_1 = \sum_{i=1}^m |\varphi(x_i) - y_i| \quad \Longrightarrow \quad \text{最小平均逼近}$$

$$R_2 = \|Q - Y\|_2 = \left(\sum_{i=1}^m |\varphi(x_i) - y_i|^2 \right)^{1/2} \quad \Longrightarrow \quad \text{最小平方逼近}$$

$$R_\infty = \|Q - Y\|_\infty = \max_{1 \leq i \leq m} \{|\varphi(x_i) - y_i|\} \quad \Longrightarrow \quad \text{最小一致逼近}$$

■ Gauss和Legendre分别发明

■ 最小二乘法: $R = R_2^2 = \left(\sum_{i=1}^m |\varphi(x_i) - y_i|^2 \right)$

- 均方误差
- 按均方误差达到最小
- 优点: 易于求解

■ 最小二乘法是最基本的函数逼近方法, 被广泛应用于运筹学、统计学、逼近论等各个领域

■ 问题: 如何选定函数空间 Φ ?

- 专业知识或工作经验
- 观察数据

线性拟合和二次拟合函数

- 最小二乘问题：给定数据序列 $\{(x_i, y_i)\}_{i=1}^m$ 和函数空间

$$\Phi = \text{span}\{\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)\},$$

求函数 $\varphi(x) \in \Phi$ ，使得

$$\min_{\varphi \in \Phi} \sum_{i=1}^m (\varphi(x_i) - y_i)^2$$

其中 $\varphi(x) = a_1\varphi_1(x) + a_2\varphi_2(x) + \dots + a_n\varphi_n(x)$, $a_i \in \mathbb{R}$, $i = 1, 2, \dots, n$

- 若定义

$$Q(a_1, a_2, \dots, a_n) = \sum_{i=1}^m (\varphi(x_i) - y_i)^2 = \sum_{i=1}^m (a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \dots + a_n\varphi_n(x_i) - y_i)^2$$

则最小二乘问题可写成

$$\min_{a_1, a_2, \dots, a_n \in \mathbb{R}} Q(a_1, a_2, \dots, a_n)$$

线性拟合和二次拟合函数

- 将 $Q(a_1, a_2, \dots, a_n) = \sum_{i=1}^m (a_1 \varphi_1(x_i) + a_2 \varphi_2(x_i) + \dots + a_n \varphi_n(x_i) - y_i)^2$ 写成矩阵的表示形式：记

$$\mathbf{A} = \begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{pmatrix}_{m \times n}, \quad \mathbf{x} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}_{n \times 1}, \quad \mathbf{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}_{m \times 1},$$

则有

$$Q(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b}$$

- 二次型取最小值的条件

$$\frac{\partial Q(\mathbf{x})}{\partial \mathbf{x}} = 0 \Rightarrow 2\mathbf{A}^T \mathbf{Ax} - 2\mathbf{A}^T \mathbf{b} = 0 \Rightarrow \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b}$$

线性拟合和二次拟合函数



- 线性拟合：取函数空间 $\Phi = \text{span}\{1, x\}$ ， $\varphi(x) = a + bx$ ，则

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b} \quad \longrightarrow \quad \begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{pmatrix}$$

线性拟合和二次拟合函数

- 二次拟合：取函数空间 $\Phi = \text{span}\{1, x, x^2\}$, $\varphi(x) = a + bx + cx^2$, 则

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b} \quad \longrightarrow \quad \begin{pmatrix} m & \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 \\ \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 & \sum_{i=1}^m x_i^4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m x_i^2 y_i \end{pmatrix}$$



线性拟合和二次拟合函数

■ 形如 ae^{bx} 曲线拟合：取函数空间 $\Phi = \{ae^{bx} \mid a, b \in \mathbb{R}\}$

- 非线性空间
- 线性化：作变换 $z = \ln y$ ，有

$$z_i = \ln y_i$$

$$\ln \varphi(x) = \ln(ae^{bx}) = \ln a + bx = A + Bx$$

- 利用线性拟合求 A, B

■ 类似地，可求形如 $\frac{1}{a+bx}$ 曲线拟合

■ 注意：变换后的所得拟合曲线，已经不在是平方误差极小意义下的拟合曲线

解矛盾方程组



■ 矛盾方程组:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m \times n}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{n \times 1}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m \times 1}$$

■ 当 $m \gg n$ 时, 方程组 $\mathbf{Ax} = \mathbf{b}$ 通常无解, 故称为矛盾方程组

■ 修改: 要求均方误差最小

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2$$

- 最小二乘意义下的解

解矛盾方程组



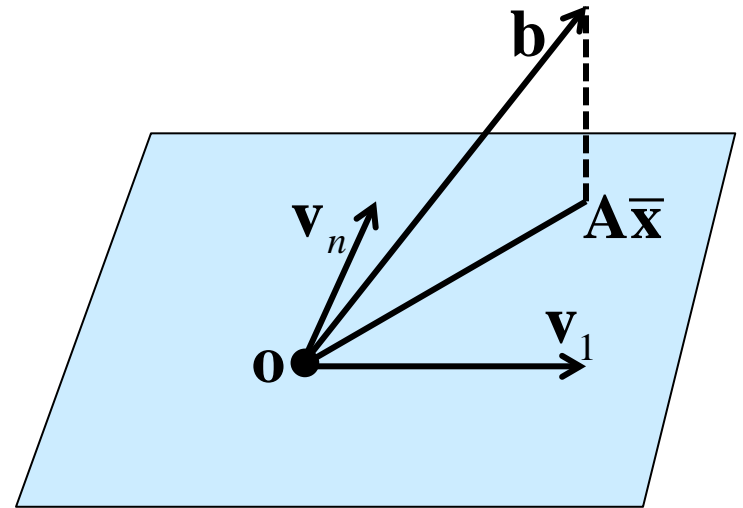
■ 最小二乘法的几何解释:

$$\mathbf{Ax} = \mathbf{b} \Rightarrow x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n = \mathbf{b}$$

$$(\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) \perp \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{R}^n\}$$

$$\mathbf{x}^T \mathbf{A}^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = 0, \quad \forall \mathbf{x} \in \mathbb{R}^n$$

$$\mathbf{A}^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = 0 \Rightarrow \mathbf{A}^T \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$$



解矛盾方程组



■ 定理：设 A 为 $m \times n$ 阶的矩阵， b 为 m 维列向量，则

(1) $A^T A x = A^T b$ 称为矛盾方程 $Ax = b$ 的法方程，恒有解

(2) x 是 $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$ 的解 $\Leftrightarrow x$ 是法方程 $A^T A x = A^T b$ 的解