教 案

无条件极值

教学内容

最大值和最小值问题大量地出现于理论研究和客观实际之中，例如，路程最短、用料最省、产量最多、收益最大等等。而最大值或最小值问题往往通过极值问题来解决。由于许多问题往往受到多个因素的影响和制约，因此有必要讨论多元函数的极值问题和最值问题。本节中主要讲解以下几方面的内容：

（1） 多元函数极值的概念与取极值的必要条件；

（2） 多元函数取极值的充分条件；

（3） 多元函数的最值问题；

（4） 最小二乘法与矛盾方程组。

教学思路和要求

（1）与一元函数类似，多元函数的最值与极值有着密切联系，且最值问题的讨论，往往先从极值问题入手，因此我们先从函数的极值问题展开。

（2）由于关于一元函数已经讨论过极值和最值问题，学生掌这些概念、方法和计算并不困难，但教师应通过实例指出多元函数的极值和最值问题仍有其特点与复杂性。当充分性条件失效时，对于极值问题如何处理；对最值问题如何讨论，都是要请清楚其复杂性的。同时，还要介绍一些计算技巧。

（3）由于最小二乘法与求矛盾方程组的最小二乘解的思想和方法在实际中应用比较广泛，因此有必要将问题的来龙去脉讲清楚，计算方法讲清楚，不能一带而过。

教学安排

一．多元函数的无条件极值

多元函数的极值刻画了多元函数的一个局部性质。

定义 7.7.1 设 n 元函数 f 定义于开集 $\Omega \subset \mathbf{R}^n$ 上，$\boldsymbol{x}_0 \in \Omega$。如果存在 $\delta > 0$，使得

$$f(\boldsymbol{x}) \ge f(\boldsymbol{x}_0) \quad （或\ f(\boldsymbol{x}) \le f(\boldsymbol{x}_0)），\quad \boldsymbol{x} \in O(\boldsymbol{x}_0, \delta)，$$

则称 $\boldsymbol{x}_0$ 为 f 的一个**极小值点**（或**极大值点**），称 $f(\boldsymbol{x}_0)$ 为相应的**极小值**（或**极大值**），极小值点和极大值点统称为**极值点**，极小值和极大值统称为**极值**。

例 函数 $z = x^2 + 2y^2$ 在 $(0, 0)$ 点取极小值 0，这是因为在 $(0, 0)$ 的任何邻域中异于 $(0, 0)$ 的点处，函数均取正值。

例 函数 $z = xy$ 在 $(0, 0)$ 点既不取到极大值，也不取到极小值。因为在点 $(0, 0)$ 处函数值为 0，而在该点的任一邻域内，总有使函数值为正的点，也有使函数值为负的点。

Fermat 定理指出，对于一元函数而言，如果 f 在点 x_0 处可导，那末 x_0 是 f 的极值点的必要条件是 $f'(x_0) = 0$。由这个结果可以直接导出多元函数极值点的一个必要条件。

定理 7.7.1（极值点的必要条件） 设 $\boldsymbol{x}_0 = (x_1^{(0)}, \cdots, x_n^{(0)})$ 是 n 元函数 f 的一个

极值点，且 f 在 $\boldsymbol{x}_0$ 处各个一阶偏导数均存在，则必有

$$f'_{x_i}(\boldsymbol{x}_0)=0, \quad i=1,2,\cdots,n。$$

证 固定除 x_i 外的其余 $n-1$ 个量，考察一元函数

$$\varphi_i(x_i)=f(x_1^{(0)},\cdots,x_{i-1}^{(0)},x_i,x_{i+1}^{(0)},\cdots,x_n^{(0)}),$$

则 φ_i 在点 $x_i^{(0)}$ 处可导，且 $x_i^{(0)}$ 是 φ_i 的极值点。由 Fermat 定理，即得 $\varphi_i'(x_i^{(0)})=0,$ 亦即

$$f'_{x_i}(\boldsymbol{x}_0)=0。$$

f 的各个一阶偏导数均为 0 的点称为它的**驻点**。定理 7.7.1 就是说：在偏导数存在的前提下，极值点必定是驻点。

驻点未必是极值点，例 7.7.2 就说明了这一点。

偏导数不存在的点也可能是极值点。例如 $f(x,y)=|x|$，函数 $z=f(x,y)$ 所对应的图象是一个柱面，可见在 Oxy 平面上整个 y 轴上每一个点 $(0,y)$ 都是 f 的极小值点，但是在这些点上 f 关于 x 的偏导数均不存在。

二．取极值的充分条件

如何判别一个驻点是否为极值点？下面的定理提供了一个充分条件。

定理 7.7.2（极值点的充分条件） 设 n 元函数 f 在 $\boldsymbol{x}_0$ 的某邻域上具有各个二阶连续偏导数，$\boldsymbol{x}_0$ 是 f 的一个驻点，f 的 Hessian 矩阵为

$$\boldsymbol{H}=\begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots & f''_{x_1x_n} \\ f''_{x_2x_1} & f''_{x_2x_2} & \cdots & f''_{x_2x_n} \\ \cdots & & & \\ f''_{x_nx_1} & f''_{x_nx_2} & \cdots & f''_{x_nx_n} \end{pmatrix},$$

则当在点 $\boldsymbol{x}_0$ 处 $\boldsymbol{H}$ 正定时，$\boldsymbol{x}_0$ 为 f 的极小值点；在点 $\boldsymbol{x}_0$ 处 $\boldsymbol{H}$ 负定时，$\boldsymbol{x}_0$ 是 f 的极大值点。

证 由于 $\boldsymbol{x}_0$ 是 f 的驻点，所以

$$f'_{x_i}(\boldsymbol{x}_0)=0, \quad i=1,2,\cdots,n。$$

由上一节末给出的 f 在 $\boldsymbol{x}_0$ 处的一阶 Taylor 公式得到

$$f(\boldsymbol{x})-f(\boldsymbol{x}_0)=\frac{1}{2}(\Delta\boldsymbol{x})^T\boldsymbol{H}(\boldsymbol{x}_0+\theta\Delta\boldsymbol{x})\Delta\boldsymbol{x},$$

其中 $\Delta\boldsymbol{x}=\boldsymbol{x}-\boldsymbol{x}_0$，$0<\theta<1$。

根据正定矩阵特征可知 $\boldsymbol{H}(\boldsymbol{x}_0)$ 正定等价于其各阶顺序主子式均取正值，而 Hessian 矩阵中所有元素 $f''_{x_ix_j}$ 在 $\boldsymbol{x}_0$ 某邻域中连续，从而各阶顺序主子式也在该邻域中连续，因此可取到 $\delta>0$，使得 $\boldsymbol{x}\in O(\boldsymbol{x}_0,\delta)$ 时，$\boldsymbol{H}(\boldsymbol{x})$ 的各阶顺序主子式均取正值，也就是说，此时 $\boldsymbol{H}(\boldsymbol{x})$ 是正定矩阵。于是当 $0<\|\Delta\boldsymbol{x}\|<\delta$ 时，

$$(\Delta\boldsymbol{x})^T\boldsymbol{H}(\boldsymbol{x}_0+\theta\Delta\boldsymbol{x})\Delta\boldsymbol{x}>0,$$

所以

$$f(\boldsymbol{x})-f(\boldsymbol{x}_0)>0,$$

即 $\boldsymbol{x}_0$ 是 f 的极小值点。

当 H 在 $\boldsymbol{x}_0$ 处负定时，$-\boldsymbol{H}(\boldsymbol{x}_0)$ 正定，从而 $\boldsymbol{x}_0$ 是 $-f$ 的极小值点，即是 f 的极大值点。

证毕

例 求$u=f(x,y,z)=x^3+y^2+z^2+6xy+2z$的极值点。

解 先求出函数 f 的驻点，为此，解方程组

$$\begin{cases} u'_x=3x^2+6y=0, \\ u'_y=2y+6x=0, \\ u'_z=2z+2=0, \end{cases}$$

即得两个驻点$P(6,-18,-1)$， $Q(0,0,-1)$。函数 f 在点(x,y,z)处的 Hessian 矩阵为

$$\boldsymbol{H}=\begin{pmatrix} 6x & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

于是

$$\boldsymbol{H}(P)=\begin{pmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \qquad \boldsymbol{H}(Q)=\begin{pmatrix} 0 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

因为

$$|36|>0，\begin{vmatrix} 36 & 6 \\ 6 & 2 \end{vmatrix}=36>0，\begin{vmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix}=72>0。$$

所以$\boldsymbol{H}(P)$是正定阵，由定理 7.7.2 可知，P 是函数 f 的极小值点。

容易验证$\boldsymbol{H}(Q)$并非正定阵，从而不能用定理 7.7.2 判断 Q 是否为极值点。由直接计算可得$\boldsymbol{H}(Q)$的三个特征值为

$$\lambda_1=2，\lambda_2=1+\sqrt{37}，\lambda_3=1-\sqrt{37}。$$

当我们取$\Delta\boldsymbol{x}$为$\boldsymbol{H}(Q)$相应于特征值λ_i的特征向量时，可得

$$(\Delta\boldsymbol{x})^T\boldsymbol{H}(Q)\Delta\boldsymbol{x}=\lambda_i(\Delta\boldsymbol{x})^T\Delta\boldsymbol{x},$$

即$(\Delta\boldsymbol{x})^T\boldsymbol{H}(Q)\Delta\boldsymbol{x}$与$\lambda_i$同号，再利用 Hessian 矩阵中各元素的连续性可知，当特征向量的范数$\|\Delta\boldsymbol{x}\|$充分小时

$$f(\boldsymbol{x})-f(\boldsymbol{x}_0)=\frac{1}{2}(\Delta\boldsymbol{x})^T\boldsymbol{H}(\boldsymbol{x}_0+\theta\Delta\boldsymbol{x})\Delta\boldsymbol{x}$$

与λ_i同号。注意到$\lambda_1,\lambda_2>0$，$\lambda_3<0$，这样，在$\boldsymbol{x}_0$的任何邻域内，既能取到$\boldsymbol{x}$，使得$f(\boldsymbol{x})>f(\boldsymbol{x}_0)$，又能取得$\boldsymbol{x}$，使得$f(\boldsymbol{x})<f(\boldsymbol{x}_0)$，所以$\boldsymbol{x}_0$不是 f 的极值点。

实际应用中经常遇到二元函数的情况，我们把这个特例作为一个推论。

推论 7.7.1 设函数$z=f(x,y)$在点(x_0,y_0)的某邻域上各二阶偏导数均连续，(x_0,y_0)是 f 的一个驻点，记

$$\Delta=f''_{xx}(x_0,y_0)f''_{yy}(x_0,y_0)-[f''_{xy}(x_0,y_0)]^2。$$

则

（1）当$\Delta>0$时，(x_0,y_0)是 f 的极值点。且当$f''_{xx}(x_0,y_0)>0$时，(x_0,y_0)是极小值点；当$f''_{xx}(x_0,y_0)<0$时，(x_0,y_0)是极大值点；

（2）当$\Delta<0$时，(x_0,y_0)不是 f 的极值点。

注 当$\Delta = 0$时，(x_0, y_0)是否为极值点需另行讨论。

证 因为$\Delta > 0$且$f''_{xx}(x_0, y_0) > 0$等价于f在(x_0, y_0)处的Hessian阵正定；同样，$\Delta > 0$且$f''_{xx}(x_0, y_0) < 0$等价于f在(x_0, y_0)处的Hessian阵负定，由此即得（1）的结论。

又因为f在(x_0, y_0)处Hessian阵两特征值之积即Δ，因而当$\Delta < 0$时两特征值异号，此时易知(x_0, y_0)一定不是f的极值点。

证毕

例 设$a \neq 0$，求函数$f(x, y) = xy(a - x - y)$的极值。

解 先求出f的驻点，即解方程组

$$\begin{cases} f'_x = y(a - x - y) - xy = 0, \\ f'_y = x(a - x - y) - xy = 0 \end{cases}$$

得到四个驻点

$$(0, 0)，(a, 0)，(0, a)，\left(\frac{a}{3}, \frac{a}{3}\right)。$$

由计算得$f''_{xx}(x, y) = -2y$，又

$$\begin{aligned} \Delta &= f''_{xx} f''_{yy} - f''^2_{xy} = (-2x)(-2y) - (a - 2x - 2y)^2 \\ &= -a^2 - 4x^2 - 4y^2 + 4ax + 4ay。 \end{aligned}$$

以驻点代入，得

$$\Delta|_{(0,0)} = \Delta|_{(a,0)} = \Delta|_{(0,a)} = -a^2 < 0，$$

因而$(0, 0)$，$(a, 0)$和$(0, a)$均非f的极值点。

由于

$$\Delta|_{\left(\frac{a}{3},\frac{a}{3}\right)} = \frac{1}{3}a^2 > 0，$$

且

$$f''_{xx}\left(\frac{a}{3}, \frac{a}{3}\right) = -\frac{2}{3}a，$$

因而当$a > 0$时，$f\left(\frac{a}{3}, \frac{a}{3}\right) = \frac{a^3}{27}$为极大值；当$a < 0$时，$f\left(\frac{a}{3}, \frac{a}{3}\right) = \frac{a^3}{27}$为极小值。

例 讨论函数$f(x, y) = (x - 2)^4 + (x - y)^4$的极值。

解 解方程组

$$\begin{cases} f'_x(x, y) = 4(x - 2)^3 + 4(x - y)^3 = 0, \\ f'_y(x, y) = -4(x - y)^3 = 0 \end{cases}$$

得驻点$(2, 2)$。在该点处$f''_{xx} = f''_{yy} = f''_{xy} = 0$，所以

$$\Delta|_{(2,2)} = 0。$$

这样，不能用定理7.7.2来判定f在$(2, 2)$是否取得极值。

但是，由于$(x, y) \neq (2, 2)$时，必有

$$f(x, y) = (x - 2)^4 + (x - y)^4 > 0 = f(2,2)，$$

所以$(2, 2)$是函数f的极小值点。

例 讨论$f(x, y) = x^2 - 2xy^2 + y^4 - y^5$的极值。

解 解方程组

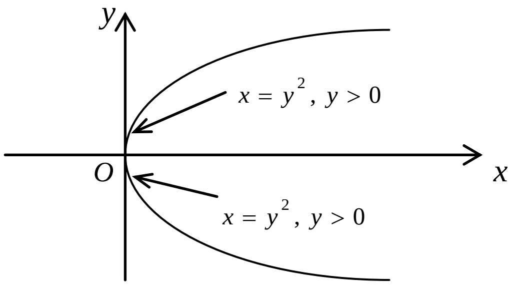
$$\begin{cases} \dfrac{\partial f}{\partial x} = 2x - 2y^2 = 0, \\ \dfrac{\partial f}{\partial y} = -4xy + 4y^3 - 5y^4 = 0. \end{cases}$$

求得驻点 $(0,0)$。再计算二阶偏导数，

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = -4y, \quad \frac{\partial^2 f}{\partial y^2} = -4x + 12y^2 - 20y^3,$$

在 $(0,0)$ 处有 $AC - B^2 = 0$，这时候无法用定理判定。

注意到 $f(0,0) = 0$，以及 $f(x,y) = (x - y^2)^2 - y^5$，那么，在曲线 $x = y^2,\ y > 0$ 上 $f(x,y) < 0$；在曲线 $x = y^2,\ y < 0$ 上 $f(x,y) > 0$，因此 $f(0,0) = 0$ 不是极值（见下图）。



三．函数的最值

函数的最值是指函数在某区域上的最大值或最小值。如果说函数的极值是一个局部性的概念，那么函数的最值却是一个涉及整体性质的概念。

当我们考虑函数的最值时，应当把区域内部所有极值点上的函数值和在区域边界上的函数值作比较来确定。在通常遇到的实际问题中，如果所讨论的函数在区域内部偏导数处处存在，而根据问题性质，又能确定其最值一定在区域内部取得，当这个函数只有一个驻点时，可以肯定这个驻点就是函数的最值点。

例 有一块宽 12cm 的薄金属片，把它的两边折起，做成一个截面为等腰梯形的水槽（图 7.7.1），问怎样析方能使梯形截面的面积最大？

解 设截面梯形的腰长为 x，腰与底边夹角为 α，于是截面积 F 为

$$F(x,\alpha) = \frac{1}{2}[(12 - 2x) + (12 - 2x + 2x\cos\alpha)]x\sin\alpha$$

$$= 12x\sin\alpha - 2x^2\sin\alpha + x^2\sin\alpha\cos\alpha。$$

根据问题的实际背景

$$D(F) = \{(x,\alpha) \mid 0 \le x \le 6,\ 0 \le \alpha \le \pi\}。$$

先求 F 在区域 $D(F)$ 内部的驻点。解方程组

$$\begin{cases} F'_x(x,\alpha) = 12\sin\alpha - 4x\sin\alpha + 2x\sin\alpha\cos\alpha = 0, \\ F'_\alpha(x,\alpha) = 12x\cos\alpha - 2x^2\cos\alpha + x^2(\cos^2\alpha - \sin^2\alpha) = 0. \end{cases}$$

当 $x > 0$，$0 < \alpha < \pi$ 时，原方程可简化为

$$\begin{cases} 6 - 2x + x\cos\alpha = 0, \\ 12\cos\alpha - 2x\cos\alpha + x(2\cos^2\alpha - 1) = 0. \end{cases}$$

由此可得 $\cos\alpha=\dfrac{1}{2}$， $x=4$，即

$$\alpha=\frac{\pi}{3},\quad x=4$$

$$F\left(4,\frac{\pi}{3}\right)=12\sqrt{3}\approx 20.8\ (\mathrm{cm}^2)\text{。}$$

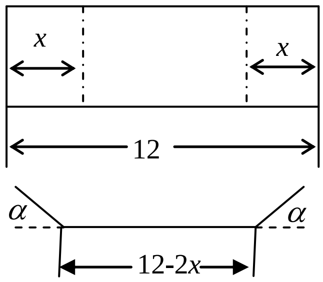


图 7.7.1

由于在边界 $x=0$， $\alpha=0$ 及 $\alpha=\pi$ 上 $F(x,\alpha)$ 均为0，且 $\max F(6,\alpha)=\max(18\sin 2\alpha)=18$，因此 F 的最值必定在 $D(F)$ 的内部达到，而 $x=4$， $\alpha=\dfrac{\pi}{3}$ 是唯一可能达到最值的解，因此这就是所求的。

四．最小二乘法

在实践活动中，经常需要从一组统计数据 (x_i,y_i)（$i=1,2,\cdots,n$）中寻求变量 x 与 y 间函数关系的近似表达式，即经验公式。最小二乘法就是建立经验公式的一种常用方法。

先从简单的情况谈起。在平面直角坐标系中，即坐标为 (x_i,y_i) 的点为 A_i（$i=1,2,\cdots,n$）。假设这些点分布在某条直线附近，我们就认为 x,y 间存在着线性关系

$$y=ax+b\text{。}$$

如果所有的 A_i 恰好都在直线 $y=ax+b$ 上，自然令人满意，但这类理想的情况往往并不可能。当 $x=x_i$ 时，经实测获得的数据是 y_i，而直线上相应的纵坐标却是 ax_i+b，两者有误差 $\delta_i=y_i-(ax_i+b)$。显然，合理的挑选 a,b 的标准应满足“总体偏差不很大”的原则。而总体的偏差通常又以诸 δ_i 的平方和作标准来衡量。

以偏差 δ_i 的平方和为标准来选取 a,b 的方法就是**最小二乘法**。

记

$$\Delta(a,b)=\sum_{i=1}^{n}[y_i-(ax_i+b)]^2\text{。}$$

为求使 $\Delta(a,b)$ 最小的 (a,b)，应用多元函数求极值的方法可知，a,b 应满足方程组

$$\begin{cases}\dfrac{\partial\Delta}{\partial a}=\displaystyle\sum_{i=1}^{n}2(y_i-ax_i-b)(-x_i)=0,\\[2ex]\dfrac{\partial\Delta}{\partial b}=\displaystyle\sum_{i=1}^{n}2(y_i-ax_i-b)(-1)=0,\end{cases}$$

即

$$\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i & n\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^{n}x_i\,y_i\\ \sum_{i=1}^{n}y_i\end{pmatrix}\text{。}$$

由这个方程组解得 a,b，再代入 $y=ax+b$，即得到所求的经验公式。

例 经实际测定，某种型号空调的制冷功率（单位：千瓦）与噪声（单位：分贝）有如下关系：

制冷功率	3.34	3.4	3.36	3.32	3.34	3.36	3.38	3.4	3.4	3.42
噪声	49.2	50	49.3	49	49	49.5	49.8	49.9	50.2	50.2

当它以 4 千瓦功率能力工作时，噪声能否小于 60 分贝？

解　记制冷功率为 x ，噪声为 y。将这些数据画在坐标纸上，发现它们呈直线关系，所以有理由猜测，y 是 x 的一次函数，即

$$y = ax + b。$$

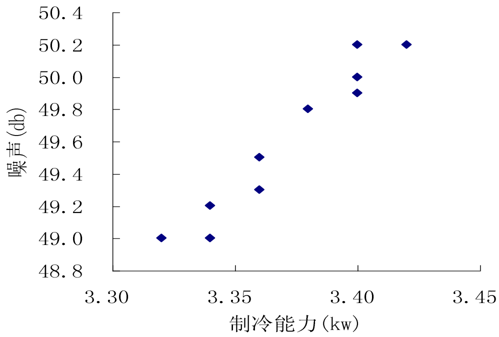


图 7.7.2

分别以 x_i， y_i 记实测所得的制冷功率与噪声，按前述方法便得方程组

$$\begin{pmatrix} 113.7136 & 33.72 \\ 33.72 & 10 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1672.984 \\ 496.1 \end{pmatrix}$$

由此解得

$$\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} 13.81 \\ 3.04 \end{pmatrix}。$$

于是，制冷功率 x 与噪声 y 近似满足

$$y = 13.81\,x + 3.04。$$

制冷功率	3.34	3.40	3.36	3.32	3.34	3.36	3.38	3.40	3.40	3.42
实测噪声	49.20	50.00	49.30	49.00	49.00	49.50	49.80	49.90	50.20	50.20
计算噪声	49.17	49.99	49.44	48.89	49.17	49.44	49.72	49.99	49.99	50.27
相对误差(%)	0.07	0.01	0.29	0.23	0.34	0.12	0.17	0.19	0.41	0.14

当制冷功率 $x = 4$ 千瓦，噪声 y 近似值为 58.28 分贝，根据以上数据的相对误差，有把握不超过 60 分贝。

例　通过观察知道，红铃虫的产卵数与温度有关，下面是一组实验观察值：

温度	21	23	25	27	29	32	35
产卵数	7	11	21	24	66	105	325

试确定产卵数与温度的近似函数关系。

解　记温度为 x，产卵数为y。将数据画在坐标纸上，看起来两者呈指数关系。设对应的近似关系是

$$y = \beta \mathrm{e}^{ax} = \mathrm{e}^{b}\cdot \mathrm{e}^{ax}，$$

其中 $\beta = e^b$ 。两边取对数，再作代换 $\tilde{y} = \ln y$，就得到

$$\tilde{y}=ax+b。$$

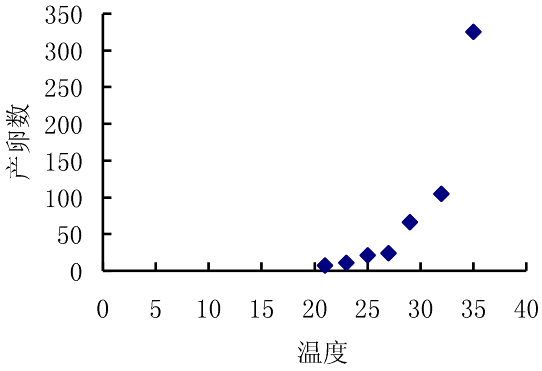


图 7.7.3

将观察数据转化为下表

x	21	23	25	27	29	32	35
$\tilde{y}=\ln y$	1.9459	2.3979	3.0445	3.1781	4.1897	4.6540	5.7839

应用最小二乘法即得

$$\tilde{y}=0.26921x-3.78495，$$

转换回去，就得到了红铃虫的产卵数与温度的关系为

$$y=0.02271\,\mathrm{e}^{0.26921x}。$$

也可以设想，本题的数据分布是双曲线的一部分，这时可设对应的近似关系为

$$\frac{1}{y}=\frac{a}{x}+b，$$

作代换 $\tilde{y}=\frac{1}{y}$， $\tilde{x}=\frac{1}{x}$， 同样变成了线性问题

$$\tilde{y}=a\tilde{x}+b，$$

求出 a，b 后，代回上面的式子就可以了。

五．矛盾方程组

用最小二乘法寻求经验公式的思路又表现为寻求“矛盾方程组”的解。

我们知道，若线性方程组

$$A\boldsymbol{x}=\boldsymbol{b}$$

的系数矩阵的秩与其增广矩阵的秩不相等，则方程组无解。此时，称原方程组为**矛盾方程组**。

从理论上判断，许多实际问题一定是有解的。但由于许多数据是通过测量或统计得到的，因此方程组的系数难免有这样那样的误差；同时，将实际问题归结为数学问题（即**数学建模**）的思想、方法、工具常常会受到种种限制，未必能恰如其分地反映问题固有的内在规律，这些都会造成本来可解的问题变得不可解。然而，科学研究、生产过程、经济活动等领域又要求我们必须找到这些问题的解，所以，研究如何处理这一类方程具有非常重要的意义。

既然数据本身就带有误差，也就是说，即使求得了它的精确解，对问题来说，也仅仅是一个近似解而已。那么一个合理的思路是，是否可以就以矛盾方程组为对象，求出它的一个尽可能精确的近似解，作为问题的近似解。

所谓“尽可能精确”是这么定义的：

定义 7.7.2 设 $\boldsymbol{A}\in\mathbf{R}^{n\times m}$，$\boldsymbol{x}\in\mathbf{R}^{m}$，$\boldsymbol{b}\in\mathbf{R}^{n}$，线性方程组

$$A\boldsymbol{x}=\boldsymbol{b}$$

的系数矩阵是列满秩的，且系数矩阵的秩不等于其增广矩阵的秩。若存在向量 $\tilde{\boldsymbol{x}}$，使得

$$\sum_{i=1}^{n}[b_i-(A\tilde{\boldsymbol{x}})_i]^2=\min_{\boldsymbol{x}\in\mathbf{R}^n}\sum_{i=1}^{n}[b_i-(A\boldsymbol{x})_i]^2,$$

则称 $\tilde{\boldsymbol{x}}$ 为这个方程组的**最小二乘解**。

当 $\boldsymbol{A}$ 列满秩时，$\boldsymbol{A}^T\boldsymbol{A}$ 是非奇异矩阵。这是因为如果 $\det \boldsymbol{A}^T\boldsymbol{A}=0$，则存在 $\boldsymbol{x}\in\mathbf{R}^n$（$\boldsymbol{x}\neq 0$），使得 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$，于是 $\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$，进而得 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$，与 $\boldsymbol{A}$ 列满秩矛盾。因此以 $\boldsymbol{A}^T\boldsymbol{A}$ 为系数矩阵的线性方程组的解存在唯一。

定理 7.7.3 $\tilde{\boldsymbol{x}}$ 是线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$$

的最小二乘解的充分必要条件是，$\tilde{\boldsymbol{x}}$ 是线性方程组

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^T\boldsymbol{b}$$

的解。

证 根据定义，如果 $\tilde{\boldsymbol{x}}=(x_1,\cdots,x_m)^T$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，即 $\tilde{\boldsymbol{x}}$ 是 m 元函数

$$f(x_1,\cdots,x_m)=\sum_{i=1}^{n}(b_i-\sum_{j=1}^{m}a_{ij}x_j)^2$$

的最小值点。由极值点的必要条件，应有 $\dfrac{\partial f}{\partial x_k}=0$（$k=1,2,\cdots,m$）。而

$$\frac{\partial f}{\partial x_k}=2\sum_{i=1}^{n}(b_i-\sum_{j=1}^{m}a_{ij}x_j)(-a_{ik}),$$

所以

$$\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}a_{ik}x_j=\sum_{i=1}^{n}a_{ik}b_i,\quad k=1,2,\cdots,m,$$

即

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^T\boldsymbol{b}。$$

现在只要证明 $\boldsymbol{A}^T\boldsymbol{A}\tilde{\boldsymbol{x}}=\boldsymbol{A}^T\boldsymbol{b}$ 时，$\tilde{\boldsymbol{x}}$ 一定是函数 f 的最小值点。因为 $\boldsymbol{A}^T\boldsymbol{A}\tilde{\boldsymbol{x}}=\boldsymbol{A}^T\boldsymbol{b}$ 等价于 $\tilde{\boldsymbol{x}}$ 是 f 的驻点，所以

$$f(\boldsymbol{x})-f(\tilde{\boldsymbol{x}})=\frac{1}{2}(\boldsymbol{x}-\tilde{\boldsymbol{x}})^T\boldsymbol{H}(\tilde{\boldsymbol{x}}+\theta(\boldsymbol{x}-\tilde{\boldsymbol{x}}))(\boldsymbol{x}-\tilde{\boldsymbol{x}})。$$

但是，因为 f 是二次函数，故 $\boldsymbol{H}$ 是常值矩阵，由直接计算可知

$$\frac{\partial^2 f}{\partial x_k\partial x_l}=2\sum_{i=1}^{n}a_{ik}a_{il},$$

所以 $\boldsymbol{H}=2\boldsymbol{A}^T\boldsymbol{A}$。于是

$$f(\boldsymbol{x})-f(\tilde{\boldsymbol{x}})=(\boldsymbol{x}-\tilde{\boldsymbol{x}})^T\boldsymbol{A}^T\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})。$$

因为 $\boldsymbol{A}$ 是列满秩的，故当 $\boldsymbol{x}\neq\tilde{\boldsymbol{x}}$ 时 $\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})\neq 0$，从而

$$f(\boldsymbol{x})-f(\tilde{\boldsymbol{x}})>0,$$

即 $\tilde{\boldsymbol{x}}$ 是函数 f 的最小值点。

证毕

例　有些商品的销售量是有季节性的。某商店一年中各个月份出售某种型号的单冷空调的数量如下（单位：台）：

月份	1	2	3	4	5	6	7	8	9	10	11	12
销售量	132	199	241	339	323	360	383	375	297	244	178	98

试确定销售量与月份的近似函数关系。

解　记月份为 x ，销售量为 y。将这些数据画在坐标纸上，发现它们呈抛物线。于是设

$$y = ax^2 + bx + c\text{。}$$

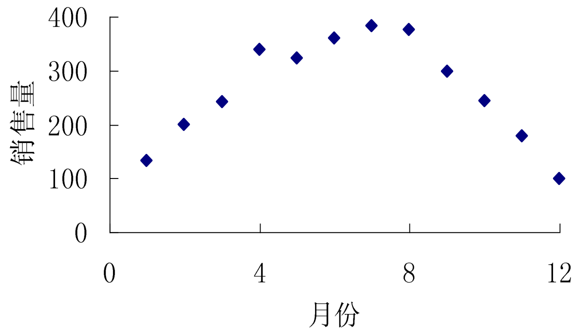


图 7.7.4

将其表成方程组

$$\begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ \vdots & \vdots & \vdots \\ 144 & 12 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 132 \\ 199 \\ 241 \\ \vdots \\ 98 \end{pmatrix}\text{。}$$

这是一个矛盾方程组，为求其最小二乘解，可在两边左乘系数矩阵的转置，得到方程组

$$\begin{pmatrix} 60710 & 6084 & 650 \\ 6084 & 650 & 78 \\ 650 & 78 & 12 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 156430 \\ 20312 \\ 3169 \end{pmatrix}\text{。}$$

解出 a、b、c 后，就得到近似关系式

$$y = -8.62\,x^2 + 110.01\,x + 15.75\text{。}$$

六．进一步的问题

作为本节讨论的继续，下一节将讨论条件极值问题。

七．习题

1；2；4；6．（1）；7；8；10；11；13。