



# 无条件极值

---

徐海峰整理

March 21, 2025

数学科学学院, 扬州大学

此幻灯片的教学内容面向非数学专业的理工科本科生.

内容完全基于下面的教材:

梅加强编著《数学分析》，高等教育出版社.

---

其他参考文献.

## 极值的定义

---

### 定义

设  $f: A \rightarrow \mathbb{R}$  为多元函数, 其中  $A$  为  $\mathbb{R}^n$  的子集,  $x^0 = (x_1^0, \dots, x_n^0) \in A$ .

### 定义

设  $f: A \rightarrow \mathbb{R}$  为多元函数, 其中  $A$  为  $\mathbb{R}^n$  的子集,  $x^0 = (x_1^0, \dots, x_n^0) \in A$ . 如果存在  $\delta > 0$ , 使得

$$f(x^0) \geq f(x) \quad (\text{或 } f(x^0) \leq f(x)), \quad \forall x \in A \cap B_\delta(x^0) - \{x^0\},$$

### 定义

设  $f: A \rightarrow \mathbb{R}$  为多元函数, 其中  $A$  为  $\mathbb{R}^n$  的子集,  $x^0 = (x_1^0, \dots, x_n^0) \in A$ . 如果存在  $\delta > 0$ , 使得

$$f(x^0) \geq f(x) \quad (\text{或 } f(x^0) \leq f(x)), \quad \forall x \in A \cap B_\delta(x^0) - \{x^0\},$$

则称  $x^0$  为  $f$  的极大(小)值点,

## 定义

设  $f: A \rightarrow \mathbb{R}$  为多元函数, 其中  $A$  为  $\mathbb{R}^n$  的子集,  $x^0 = (x_1^0, \dots, x_n^0) \in A$ . 如果存在  $\delta > 0$ , 使得

$$f(x^0) \geq f(x) \quad (\text{或 } f(x^0) \leq f(x)), \quad \forall x \in A \cap B_\delta(x^0) - \{x^0\},$$

则称  $x^0$  为  $f$  的极大(小)值点,  $f(x^0)$  为  $f$  的极大(小)值.

## 定义

设  $f: A \rightarrow \mathbb{R}$  为多元函数, 其中  $A$  为  $\mathbb{R}^n$  的子集,  $x^0 = (x_1^0, \dots, x_n^0) \in A$ . 如果存在  $\delta > 0$ , 使得

$$f(x^0) \geq f(x) \quad (\text{或 } f(x^0) \leq f(x)), \quad \forall x \in A \cap B_\delta(x^0) - \{x^0\},$$

则称  $x^0$  为  $f$  的极大(小)值点,  $f(x^0)$  为  $f$  的极大(小)值. 当上式中“ $\geq$ ” (“ $\leq$ ”) 换成“ $>$ ” (“ $<$ ”) 时, 相应地把  $x^0$  称为严格极值点,  $f(x^0)$  为严格极值.



## 达到极值的必要条件

### 命题

设  $x^0$  为  $f$  的极值点, 如果  $x^0$  为  $A$  的内点, 且  $f$  在  $x^0$  处存在一阶偏导数, 则

$$f'_{x_1}(x^0) = f'_{x_2}(x^0) = \cdots = f'_{x_n}(x^0) = 0.$$

## 达到极值的必要条件

### 命题

设  $x^0$  为  $f$  的极值点, 如果  $x^0$  为  $A$  的内点, 且  $f$  在  $x^0$  处存在一阶偏导数, 则

$$f'_{x_1}(x^0) = f'_{x_2}(x^0) = \cdots = f'_{x_n}(x^0) = 0.$$

### Proof.

以  $f'_{x_1}$  为例, 考虑一元函数  $\varphi(x) = f(x, x_2^0, \dots, x_n^0)$ , 则  $\varphi$  可导, 且以  $x_1^0$  为极值点, 由 Fermat 定理知  $\varphi'(x_1^0) = 0$ , 即  $f'_{x_1}(x^0) = 0$ . □

## 达到极值的必要条件

### 命题

设  $x^0$  为  $f$  的极值点, 如果  $x^0$  为  $A$  的内点, 且  $f$  在  $x^0$  处存在一阶偏导数, 则

$$f'_{x_1}(x^0) = f'_{x_2}(x^0) = \cdots = f'_{x_n}(x^0) = 0.$$

### Proof.

以  $f'_{x_1}$  为例, 考虑一元函数  $\varphi(x) = f(x, x_2^0, \dots, x_n^0)$ , 则  $\varphi$  可导, 且以  $x_1^0$  为极值点, 由 Fermat 定理知  $\varphi'(x_1^0) = 0$ , 即  $f'_{x_1}(x^0) = 0$ . □

### 注

满足条件  $f'_{x_1}(x^0) = \cdots = f'_{x_n}(x^0) = 0$  的点  $x^0$  称为  $f$  的驻点或临界点.

## 定理

设  $U$  为  $\mathbb{R}^n$  中开集,  $f: U \rightarrow \mathbb{R}$  具有二阶连续偏导数,  $x^0$  为  $f$  的驻点, 则

- (1) 如果  $x^0$  为  $f$  的极小 (大) 值点, 则  $\text{Hess}(f)(x^0)$  为半正 (负) 定方阵;
- (2) 如果  $\text{Hess}(f)(x^0)$  为正 (负) 定方阵, 则  $x^0$  为  $f$  的严格极小 (大) 值点;
- (3) 如果  $\text{Hess}(f)(x^0)$  为不定方阵, 则  $x^0$  不是  $f$  的极值点.

## 最小二乘法

---

## 最小二乘法

例

设  $(x_1, y_1), \dots, (x_n, y_n)$  为平面  $\mathbb{R}^2$  上  $n$  个点, 求一条直线  $y = ax + b$ , 使得

$$F(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2 \text{ 最小.}$$

## 最小二乘法

例

设  $(x_1, y_1), \dots, (x_n, y_n)$  为平面  $\mathbb{R}^2$  上  $n$  个点, 求一条直线  $y = ax + b$ , 使得

$$F(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2 \text{ 最小.}$$

解. 函数  $F(a, b)$  是关于  $a, b$  的光滑函数.

## 最小二乘法

例

设  $(x_1, y_1), \dots, (x_n, y_n)$  为平面  $\mathbb{R}^2$  上  $n$  个点, 求一条直线  $y = ax + b$ , 使得

$$F(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2 \text{ 最小.}$$

解. 函数  $F(a, b)$  是关于  $a, b$  的光滑函数. 先求驻点:

$$0 = \frac{\partial F}{\partial a}(a, b) = 2 \sum_{i=1}^n (ax_i + b - y_i)x_i,$$

$$0 = \frac{\partial F}{\partial b}(a, b) = 2 \sum_{i=1}^n (ax_i + b - y_i),$$





解. 这推出

$$\left\{ \begin{array}{l} (\sum_{i=1}^n x_i^2)a + (\sum_{i=1}^n x_i)b = \sum_{i=1}^n x_i y_i, \\ (\sum_{i=1}^n x_i)a + nb = \sum_{i=1}^n y_i. \end{array} \right.$$

解. 这推出

$$\begin{cases} (\sum_{i=1}^n x_i^2)a + (\sum_{i=1}^n x_i)b = \sum_{i=1}^n x_i y_i, \\ (\sum_{i=1}^n x_i)a + nb = \sum_{i=1}^n y_i. \end{cases}$$

改写为矩阵相乘的形式:

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

解. 这推出

$$\begin{cases} (\sum_{i=1}^n x_i^2)a + (\sum_{i=1}^n x_i)b = \sum_{i=1}^n x_i y_i, \\ (\sum_{i=1}^n x_i)a + nb = \sum_{i=1}^n y_i. \end{cases}$$

改写为矩阵相乘的形式:

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

当系数矩阵的行列式

$$n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 = \frac{1}{2} \sum_{i \neq j} (x_i - x_j)^2 \neq 0$$

时, 上述二元一次方程组有惟一的解(驻点).



解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

故在此驻点处,  $F$  的 Hessian 矩阵为

$$2 \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix},$$

它是正定方阵, 故该驻点为极小值点.

解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

故在此驻点处,  $F$  的 Hessian 矩阵为

$$2 \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix},$$

它是正定方阵, 故该驻点为极小值点. 由于  $\{x_1, x_2, \dots, x_n\}$  不全相等, 故  $\{ax_i + b\}_{i=1}^n$  不是全部有界的.

解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

故在此驻点处,  $F$  的 Hessian 矩阵为

$$2 \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix},$$

它是正定方阵, 故该驻点为极小值点. 由于  $\{x_1, x_2, \dots, x_n\}$  不全相等, 故  $\{ax_i + b\}_{i=1}^n$  不是全部有界的. 事实上, 若存在  $M > 0$ , 使得对于某些  $i \in \Lambda \subsetneq \{1, 2, \dots, n\}$ , 有  $|ax_i + b| \leq M$ , 则必存在  $j \in \{1, 2, \dots, n\} - \Lambda$ ,

$$|ax_j + b| = |a(x_j - x_i) + ax_i + b| \geq |a(x_j - x_i)| - |ax_i + b|,$$

从而当  $(a, b) \rightarrow \infty$  (即  $\|(a, b)\| = \sqrt{a^2 + b^2} \rightarrow +\infty$ ) 时,  $|ax_j + b| \rightarrow +\infty$ .

解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

故在此驻点处,  $F$  的 Hessian 矩阵为

$$2 \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix},$$

它是正定方阵, 故该驻点为极小值点. 由于  $\{x_1, x_2, \dots, x_n\}$  不全相等, 故  $\{ax_i + b\}_{i=1}^n$  不是全部有界的. 事实上, 若存在  $M > 0$ , 使得对于某些  $i \in \Lambda \subsetneq \{1, 2, \dots, n\}$ , 有  $|ax_i + b| \leq M$ , 则必存在  $j \in \{1, 2, \dots, n\} - \Lambda$ ,

$$|ax_j + b| = |a(x_j - x_i) + ax_i + b| \geq |a(x_j - x_i)| - |ax_i + b|,$$

从而当  $(a, b) \rightarrow \infty$  (即  $\|(a, b)\| = \sqrt{a^2 + b^2} \rightarrow +\infty$ ) 时,  $|ax_j + b| \rightarrow +\infty$ . 因此当  $(a, b) \rightarrow \infty$  时  $F(a, b) \rightarrow +\infty$ ,



解.

$$F''_{aa} = 2 \sum_{i=1}^n x_i^2, \quad F''_{ab} = F''_{ba} = 2 \sum_{i=1}^n x_i, \quad F''_{bb} = 2n.$$

故在此驻点处,  $F$  的 Hessian 矩阵为

$$2 \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix},$$

它是正定方阵, 故该驻点为极小值点. 由于  $\{x_1, x_2, \dots, x_n\}$  不全相等, 故  $\{ax_i + b\}_{i=1}^n$  不是全部有界的. 事实上, 若存在  $M > 0$ , 使得对于某些  $i \in \Lambda \subsetneq \{1, 2, \dots, n\}$ , 有  $|ax_i + b| \leq M$ , 则必存在  $j \in \{1, 2, \dots, n\} - \Lambda$ ,

$$|ax_j + b| = |a(x_j - x_i) + ax_i + b| \geq |a(x_j - x_i)| - |ax_i + b|,$$

从而当  $(a, b) \rightarrow \infty$  (即  $\|(a, b)\| = \sqrt{a^2 + b^2} \rightarrow +\infty$ ) 时,  $|ax_j + b| \rightarrow +\infty$ . 因此当  $(a, b) \rightarrow \infty$  时  $F(a, b) \rightarrow +\infty$ , 故该驻点为惟一的最小值点(见后面的引理). ■

解. 当  $(x, y)$  在此直线上时, 自然有  $y = ax + b$ , 将其写为  $ax - y + b = 0$ . 连同之前的两个式子

$$\begin{cases} (\sum_{i=1}^n x_i^2)a + (\sum_{i=1}^n x_i)b = \sum_{i=1}^n x_i y_i, \\ (\sum_{i=1}^n x_i)a + nb = \sum_{i=1}^n y_i, \end{cases}$$

解. 当  $(x, y)$  在此直线上时, 自然有  $y = ax + b$ , 将其写为  $ax - y + b = 0$ . 连同之前的两个式子

$$\begin{cases} (\sum_{i=1}^n x_i^2)a + (\sum_{i=1}^n x_i)b = \sum_{i=1}^n x_i y_i, \\ (\sum_{i=1}^n x_i)a + nb = \sum_{i=1}^n y_i, \end{cases}$$

将它们改写为矩阵的形式

$$\begin{pmatrix} x & y & 1 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{pmatrix} \begin{pmatrix} a \\ -1 \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$



解. 这说明  $(a, -1, b)^T$  是相应的齐次线性方程组的非零解, 故

$$\begin{vmatrix} x & y & 1 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix} = 0,$$

解. 这说明  $(a, -1, b)^T$  是相应的齐次线性方程组的非零解, 故

$$\begin{vmatrix} x & y & 1 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix} = 0,$$

这就是所求直线方程.

解. 这说明  $(a, -1, b)^T$  是相应的齐次线性方程组的非零解, 故

$$\begin{vmatrix} x & y & 1 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix} = 0,$$

这就是所求直线方程. 将此行列式按第一行展开,

$$\begin{vmatrix} \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix} x - \begin{vmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{vmatrix} y + \begin{vmatrix} \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \end{vmatrix} = 0.$$



解. 因此

$$y = \frac{\begin{vmatrix} \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{vmatrix}} x + \frac{\begin{vmatrix} \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{vmatrix}}$$

解. 因此

$$y = \frac{\begin{vmatrix} \sum_{i=1}^n y_i & n \\ \sum_{i=1}^n x_i y_i & \sum_{i=1}^n x_i \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{vmatrix}} x + \frac{\begin{vmatrix} \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \end{vmatrix}}{\begin{vmatrix} \sum_{i=1}^n x_i & n \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \end{vmatrix}}$$

即

$$y = \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i) - n \sum_{i=1}^n x_i y_i}{(\sum_{i=1}^n x_i)^2 - n \sum_{i=1}^n x_i^2} x + \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i) - (\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i^2)}{(\sum_{i=1}^n x_i)^2 - n \sum_{i=1}^n x_i^2}.$$





改写为 解.

$$\mathbf{y} = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n y_i\right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \mathbf{x} + \frac{\left(\sum_{i=1}^n y_i\right) \left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n x_i y_i\right)}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}.$$

改写为 解.

$$\mathbf{y} = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \mathbf{x} + \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}.$$

注意到分母  $n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2 = \frac{1}{2} \sum_{i \neq j} (x_i - x_j)^2$ . 若记

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \overline{x \cdot y} = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \overline{x \cdot x} = \frac{1}{n} \sum_{i=1}^n x_i^2,$$

改写为解.

$$y = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} x + \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i^2) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}.$$

注意到分母  $n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2 = \frac{1}{2} \sum_{i \neq j} (x_i - x_j)^2$ . 若记

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \overline{x \cdot y} = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \overline{x \cdot x} = \frac{1}{n} \sum_{i=1}^n x_i^2,$$

则直线的斜率为

$$a = \frac{n \sum_{i=1}^n x_i y_i - n \bar{x} \cdot n \bar{y}}{n \sum_{i=1}^n x_i^2 - (n \bar{x})^2} = \frac{\frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x \cdot x} - \bar{x}^2}.$$

解. 直线在  $y$  轴上的截距为

$$b = \frac{n\bar{y} \cdot \left(\sum_{i=1}^n x_i^2\right) - (n\bar{x}) \cdot \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (n\bar{x})^2} = \frac{\bar{y} \cdot \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x} \cdot \frac{1}{n} \sum_{i=1}^n x_i y_i}{\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2} = \frac{\bar{y} \cdot \overline{x \cdot x} - \bar{x} \cdot \overline{x \cdot y}}{\overline{x \cdot x} - \bar{x}^2}.$$

■

欢迎访问 [atzjg.net](http://atzjg.net)

## Backup slides

Sometimes, it is useful to add slides at the end of your presentation to refer to during audience questions.

The best way to do this is to include the `appendixnumberbeamer` package in your preamble and call `\appendix` before your backup slides.

The theme will automatically turn off slide numbering and progress bars for slides in the appendix.

欢迎访问 [atzjg.net](http://atzjg.net)