第二章 误差概论和最小二乘法

§ 误差的定义和分类

误差定义
- 绝对误差 $\Delta = x - a$ x：测量值
- 相对误差 $r = \Delta / a$ a：真值

误差的表现形式
- 系统误差 变化规律
- 随机误差 不可预测
- 过失误差 明显不符

随机误差的统计特征：(1) $\lim\limits_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\Delta_i = 0$

(2) $P(\Delta_i) > P(\Delta_j)$ $\quad \Delta_i < \Delta_j$

随机误差的分布密度

$$P(\Delta_i) = \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta_i^2} \qquad h = \frac{1}{\sqrt{2}\sigma}$$：精密度指数

§ 观测精度

一 精度标准

$x_i \quad i = 1,2,\cdots,n$

a 真值

$\Delta_i = x_i - a$ 真误差

$\sigma = \sqrt{\dfrac{\sum \Delta_i^2}{n}}$ 标准偏差

$$D(\xi) = E[(x_i - a)^2] = \frac{1}{n}\sum (x_i - a)^2$$

误差概论表 $\lvert\Delta_i\rvert \le \mu_\xi \sigma$			
置信水平	误差限	置信水平	误差限
50.0%	0.674σ	95.5%	2σ
68.3%	1σ	99.0%	2.58σ
95.0%	1.96σ	99.7%	3σ

误差概论和最小二乘法

通常，误差值大于3σ的观测值会被舍去，这种取舍观测值的原则称为拉依达准则（3σ准则）。

精度的好坏
- 标准误差
- 平均误差　残差绝对值的算术平均
- 概率误差　误差绝对值序列的中位数
- 半峰宽度　极大值半高度处的全宽

最靠近真值的观测值x_0称为**最或然值**。通常，$x_0 = \bar{x}$

观测值与最或然值之差称为**残差**　$v_i = x_i - \bar{x}$

$$\Delta_i = x_i - a = (x_i - \bar{x}) + (\bar{x} - a) = v_i + \bar{\Delta}$$

$$\sum \Delta_i^2 = \sum v_i^2 + \sum \bar{\Delta}^2 + 2\bar{\Delta}\sum v_i = \sum v_i^2 + n\bar{\Delta}^2$$

$$D(\bar{x}) = E[(\bar{x} - a)^2] = \sum (\bar{x} - a)^2 P(\bar{x}) = \frac{1}{n}\sum (\bar{x} - a)^2 = \frac{n\bar{\Delta}^2}{n} = \bar{\Delta}^2$$

$$\sigma_{\bar{x}}^2 = \sigma^2 / n \quad \Rightarrow \quad \sigma = \sqrt{\frac{\sum v_i^2}{n-1}}$$ 标准偏差

二　误差传递公式

1. $y = kx$　　k :常数　x :观测值

$$y + \Delta y = k(x + \Delta x)$$

$$\Delta y_i = k \cdot \Delta x_i$$

$$\sigma_y^2 = \frac{\sum \Delta y_i^{\ 2}}{n} = \frac{k^2 \sum \Delta x_i^{\ 2}}{n} = k^2 \sigma_x^2$$

$$\sigma_y = k \sigma_x$$

2. $y = x_1 + x_2$　　x_1, x_2 :观测值且互相独立

$$y + \Delta y = (x_1 + \Delta x_1) + (x_2 + \Delta x_2)$$

$$\Delta y_i = \Delta x_{1i} + \Delta x_{2i}$$

$$\sigma_y^2 = \frac{\sum \Delta y_i^{\ 2}}{n} = \frac{\sum \Delta x_{1i}^{\ 2} + 2\sum \Delta x_{1i} \Delta x_{2i} + \sum \Delta x_{2i}^{\ 2}}{n} = \sigma_{x_1}^2 + \sigma_{x_2}^2$$

3. $y = k_1 x_1 + k_2 x_2$　　k_1, k_2 :常数　x_1, x_2 :观测值且互相独立

$$\sigma_y^{\ 2} = k_1^{\ 2}\sigma_{x_1}^{\ 2} + k_2^{\ 2}\sigma_{x_2}^{\ 2}$$

4. 一般非线性函数形式

$$y = f(x_1, x_2, \cdots, x_m)$$

$$y + \Delta y = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \cdots, x_m + \Delta x_m)$$

泰勒级数展开　$y = f(x_{10}, x_{20}, \cdots, x_{m0}) + \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0 (x_k - x_{k0})$

$$\Delta y_i = \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0 (x_k - x_{k0})$$

$$\sigma_y^2 = \frac{\sum \Delta y_i^{\ 2}}{n} = \sum\left[\sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0 (x_k - x_{k0})\right]^2 / n$$

$$= \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0^2 \sigma_{x_k}^2 + 2\sum_{\substack{k<j \\ k\neq j}}^{m} \frac{\partial f}{\partial x_k}\frac{\partial f}{\partial x_j}\rho_{kj}\sigma_{x_k}\sigma_{x_j}$$

ρ_{kj} :相关系数

三 等精度和非等精度测量

等精度 $\sigma = \sqrt{\dfrac{\sum v_i^2}{n-1}}$

非等精度 $\sigma = \sqrt{\dfrac{\sum p_i v_i^2}{n-1}}$

p_i :权，对观测结果的重视程度 $p_i \propto 1/\sigma_i^2$

$\sigma_1 = \sqrt{2} \quad \sigma_2 = 2 \quad \sigma_3 = 2\sqrt{2}$

$p_1 = \sigma_1^2 / \sigma_1^2 = 1$

$p_2 = \sigma_1^2 / \sigma_2^2 = 0.5$

$p_3 = \sigma_1^2 / \sigma_3^2 = 0.25$

$\sigma_1 = \max \sigma_i \quad i = 1,2,\cdots,n$

$p_i = \sigma_1^2 / \sigma_i^2$ 权：相对概念

x_1 单位权观测值

§ 直接观测量的最或然值及标准偏差

1. 等精度情况

$$v_i = x_i - x_0 \quad Q = \sum v_i^{\ 2} = \min$$

最小二乘 → 最或然值

$$\frac{\partial Q}{\partial x_0} = 0 \quad \Rightarrow \quad x_0 = \frac{1}{n}\sum x_i = \bar{x}$$

$$\sigma_{x_0}^2 = \frac{\sigma^2}{n} = \frac{\sum v_i^{\ 2}}{n(n-1)}$$

2. 非等精度情况

$$Q = \sum p_i v_i^{\ 2} = \min$$

$$\frac{\partial Q}{\partial x_0} = 0 \quad \Rightarrow \quad x_0 = \frac{\sum p_i x_i}{\sum p_i}$$

加权平均

$$\sigma_{x_0}^2 = \frac{\sum p_i^{\ 2}\sigma_{x_i}^2}{\left(\sum p_i\right)^2} = \frac{1}{\left(\sum p_i\right)^2}\left(\sum p_i^{\ 2}\cdot\frac{\sigma^2}{p_i}\right) = \frac{\sigma^2}{\sum p_i} = \frac{\sum p_i v_i^{\ 2}}{(n-1)\sum p_i}$$

§ 间接观测量的最或然值及标准偏差

一 线性情况

1. 等精度观测列

观测站到月面的距离ρ_0 $\rightarrow$ 地球自转参数

误差方程 $l_i = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{im}x_m$

由于残差存在 $v_i = l_i - (b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{im}x_m)$

$$Q = \sum {v_i}^2 = \sum_{i=1}^{n}(l_i - \sum_{k=1}^{m} b_{ik}x_k)^2 = \min$$

$$\partial Q / \partial x_j = 0 \quad \Rightarrow \quad \sum_i (l_i - \sum_k b_{ik}x_k) b_{ij} = 0$$

引入高斯符号 $[b_j b_k] = \sum_i b_{ij} b_{ik}$

正规方程

$$\begin{cases} [b_1b_1]x_1 + [b_1b_2]x_2 + \cdots + [b_1b_m]x_m = [b_1 l] \\ [b_2b_1]x_1 + [b_2b_2]x_2 + \cdots + [b_2b_m]x_m = [b_2 l] \\ \cdots\cdots \\ [b_mb_1]x_1 + [b_mb_2]x_2 + \cdots + [b_mb_m]x_m = [b_m l] \end{cases}$$

2. 非等精度观测列

$$l_1, l_2, \cdots, l_n \xrightarrow{\text{加权}} \sqrt{p_1}l_1, \sqrt{p_2}l_2, \cdots, \sqrt{p_n}l_n$$

误差方程

$$\sqrt{p_i}v_i = \sqrt{p_i}l_i - (\sqrt{p_i}b_{i1}x_1 + \sqrt{p_i}b_{i2}x_2 + \cdots + \sqrt{p_i}b_{im}x_m)$$

$$\Downarrow \quad Q = \sum_i p_i v_i^{\,2} = \min$$

正规方程

$$\begin{cases} [pb_1b_1]x_1 + [pb_1b_2]x_2 + \cdots + [pb_1b_m]x_m = [pb_1l] \\ [pb_2b_1]x_1 + [pb_2b_2]x_2 + \cdots + [pb_2b_m]x_m = [pb_2l] \\ \cdots\cdots \\ [pb_mb_1]x_1 + [pb_mb_2]x_2 + \cdots + [pb_mb_m]x_m = [pb_ml] \end{cases}$$

其中 $[pb_jb_k] = \sum_i p_i b_{ij} b_{ik}$

二　非线性情况

黑体辐射　$B_\lambda(T,A,\lambda)=\dfrac{2Ahc^2}{\lambda^5}\cdot\dfrac{1}{e^{hc/kT\lambda-1}}$

观测量　$(\lambda, B_\lambda) \to (T, A)$

非线性关系　$y_i = f_i(x_1, x_2, \cdots, x_m)$

$$\Delta y_i = y_i - y_0 = \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0 \Delta x_k \quad \text{观测方程}$$

$$\Delta y_i - \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}\right)_0 \Delta x_k = -v_i \quad \text{误差方程}$$

高斯—牛顿迭代法　$\Delta x_k \to$

$$x_k^{(1)} = x_{0k} + \Delta x_k$$

$$x_k^{(2)} = x_k^{(1)} + \Delta x_k^{(1)}$$

$$\cdots\cdots$$

$$x_k = x_k^{(l)} + \Delta x_k^{(l)}$$

$$\begin{cases} \left|\Delta x_k^{(s+1)}\right| < \varepsilon_1 \\ \left|\dfrac{Q^{(s)} - Q^{(s+1)}}{Q^{(s+1)}}\right| < \varepsilon_2 \end{cases}$$ 迭代收敛，否则发散

三　最或然值的标准偏差

$$[\varDelta^2] > [v^2]$$

$$\sigma^2 = \frac{[\varDelta^2]}{n} = \frac{[v^2]}{n-i}$$ i:待定常数

无偏估计 $$\hat{\sigma}^2 = \frac{[v^2]}{n-m}$$

$$\sigma_{x_k} = \hat{\sigma} / \sqrt{p_{x_k}}$$

$$[b_1 l],[b_2 l],\cdots,[b_m l] \quad \rightarrow \quad 1,0,\cdots,0$$

$$\Downarrow$$

解出x_1 $\rightarrow$ $p_{x_1} = 1/x_1$

§ 最小二乘曲线拟合

一 目标函数

理论公式 $y = f(x, c_1, c_2, \cdots, c_m)$

根据n对观测值(x_i, y_i)，确定参数c_k使得目标函数

$$d = d(x_1, x_2, \cdots, x_n; y_1, y_2, \cdots, y_n; c_1, c_2, \cdots, c_m)$$

取极值。

通常，目标函数可选为：

$$1)\quad d = \max_{1<i<n} \left| y_i - f(x_i, \mathbf{c}) \right|$$

$$2)\quad d = \sum_{i=1}^{n} \left| y_i - f(x_i, \mathbf{c}) \right|$$

$$3)\quad d = \sum_{i=1}^{n} [y_i - f(x_i, \mathbf{c})]^2$$

二 最小二乘曲线拟合

这种选取各观测点的残差平方和作为目标函数的拟合称为**最小二乘曲线拟合**。拟合量记为χ^2，为目标函数

$$\chi^2 = \sum_{i=1}^{n} p_i [y_i - f(x_i, \mathbf{c})]^2 = \sum p_i \delta_i^{\,2}$$

$$\frac{\partial}{\partial c_j} \sum_{i=1}^{n} p_i \delta_i^{\,2} = -2 \sum_{i=1}^{n} p_i [y_i - f(x_i, \mathbf{c})] \frac{\partial f(x_i, \mathbf{c})}{\partial c_j} = 0 \quad k = 1 \sim m$$

1. 线性情况

$$y = y_0(x) + \sum_{k=1}^{m} c_k f_k(x)$$

得到未知参数$\boldsymbol{c}$的线性方程组

$$\sum_{k=1}^{m} c_k \sum_{i=1}^{n} p_i f_k(x_i) f_j(x_i) = \sum_{i=1}^{n} p_i [y_i - y_0(x_i)] f_j(x_i) \quad j = 1 \sim m$$

写成矩阵形式 $(\mathbf{F}^T \mathbf{P}_y \mathbf{F})\mathbf{c} = \mathbf{F}^T \mathbf{P}_y (\mathbf{y} - \mathbf{y}_0)$

解 $\hat{\mathbf{c}} = (\mathbf{F}^T \mathbf{P}_y \mathbf{F})^{-1} \mathbf{F}^T \mathbf{P}_y (\mathbf{y} - \mathbf{y}_0)$

$$E(\hat{\mathbf{c}}) = (\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}\mathbf{F}^T\mathbf{P}_y E(\mathbf{y}-\mathbf{y}_0)$$

$$= (\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}\mathbf{F}^T\mathbf{P}_y\mathbf{F}\mathbf{c}$$

$$= \mathbf{c}$$

广义误差传递公式

$$\hat{\mathbf{y}} = \mathbf{y}_0 + \mathbf{F}\hat{\mathbf{c}}$$

$$\mathbf{V}_{\hat{y}} = E\{[\hat{\mathbf{y}} - E(\hat{\mathbf{y}})][\hat{\mathbf{y}} - E(\hat{\mathbf{y}})]^T\}$$

$$= E\{[\mathbf{y}_0 + \mathbf{F}\hat{\mathbf{c}} - \mathbf{y}_0 - E(\mathbf{F}\hat{\mathbf{c}})][\mathbf{y}_0 + \mathbf{F}\hat{\mathbf{c}} - \mathbf{y}_0 - E(\mathbf{F}\hat{\mathbf{c}})]^T\}$$

$$= E\{[\mathbf{F}\hat{\mathbf{c}} - E(\mathbf{F}\hat{\mathbf{c}})][\mathbf{F}\hat{\mathbf{c}} - E(\mathbf{F}\hat{\mathbf{c}})]^T\}$$

$$= E\{\mathbf{F}[\hat{\mathbf{c}} - E(\hat{\mathbf{c}})][\hat{\mathbf{c}} - E(\hat{\mathbf{c}})]^T\mathbf{F}^T\}$$

$$= \mathbf{F}\cdot E\{[\hat{\mathbf{c}} - E(\hat{\mathbf{c}})][\hat{\mathbf{c}} - E(\hat{\mathbf{c}})]^T\}\cdot\mathbf{F}^T$$

$$= \mathbf{F}\cdot\mathbf{V}_{\hat{c}}\cdot\mathbf{F}^T$$

$\hat{\mathbf{c}}$的方差 $\mathbf{V}_{\hat{c}} = [(\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}\mathbf{F}^T\mathbf{P}_y]\mathbf{V}_y[(\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}\mathbf{F}^T\mathbf{P}_y]^T$

$$= (\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}$$

这里 $\mathbf{V}_y = \mathbf{P}_y^{-1}$

2. 非线性情况

$$y = f(x,\mathbf{c}^{(0)}) + \sum_{k=1}^{m}\left(\frac{\partial f(x,\mathbf{c})}{\partial c_k}\right)_0 (c_k - c_k^{(0)})$$

$$\sum_{k=1}^{m}\Delta c_k \sum_{i=1}^{n} p_i \left(\frac{\partial f}{\partial c_k}\right)_0 \left(\frac{\partial f}{\partial c_j}\right)_0 = \sum_{i=1}^{n} p_i [y_i - f(x_i,\mathbf{c}^{(0)})] \left(\frac{\partial f}{\partial c_j}\right)_0$$

写成矩阵形式 $(\mathbf{F}^T\mathbf{P}_y\mathbf{F})\Delta\mathbf{c} = \mathbf{F}^T\mathbf{P}_y(\mathbf{y}-\mathbf{y}_0)$

迭代求解 $\Delta\mathbf{c}^{(k)} = (\mathbf{F}^T\mathbf{P}_y\mathbf{F})^{-1}\mathbf{F}^T\mathbf{P}_y(\mathbf{y}-\mathbf{y}_0)$

$\mathbf{c}^{(k)} = \mathbf{c}^{(k-1)} + \Delta\mathbf{c}^{(k)}$

3. 最优化求解

目标函数d达到最小

1) 网格搜索法

划分参数c_j的搜索区间$[a_j,b_j]$，构成$(n\text{-}1)^m$个网格，计算这n^m个网格节点上的目标函数值，取对应最小值的参数点为最优解

两个参数 $d_{lk} = \sum [y_i - f(x_i, c_{1l}, c_{2k})]^2$

$$d_{l_0 k_0} = \min(d_{lk})$$

2) 随机搜索法

参数c_j的搜索区间$[a_j,b_j]$，定义中点$u_j=(a_j+b_j)/2$及半宽度$\delta_j=(b_j-a_j)/2$，产生T组均匀分布的随机数$\xi_j \in (a_j,b_j)$，计算目标函数，其值最小对应的随机数为最优解。

$$[a_j, b_j] = u_j \pm \delta_j$$

$$\bar{u}_j = \sum_{t=1}^{T_0} W_t \xi_{jt} \bigg/ \sum_{t=1}^{T_0} W_t \qquad \text{其中} W_t = \mathrm{d}t_0 / \mathrm{d}t$$

$$\bar{\delta}_j = \sqrt{\sum_{t=1}^{T_0} (\xi_{jt} - \bar{u}_j)^2 W_t \Big/ \sum_{t=1}^{T_0} W_t}$$

新的搜索区间 $\bar{u}_j \pm \bar{\delta}_j$

重复下去直至 $\dfrac{d^{(k-1)} - d^{(k)}}{d^{(k)}} \le \varepsilon$