
误差理论与测量平差基础

主讲： 戴华阳 教授
中国矿业大学（北京）

第三章 平差数学模型与原则

§3-1 平差的任务和思路

§3-2 平差的函数模型

§3-3 平差的随机模型

§3-4 参数估计的性质和方法

§3-5 平差的最小二乘原则

误差理论与平差方法的关系

■ 误差理论error theory

认识误差来源、性质分类、随机特性；

描述误差作为随机变量的分布、数字特征；

揭示误差数字特征的传播规律。 (理论基础)

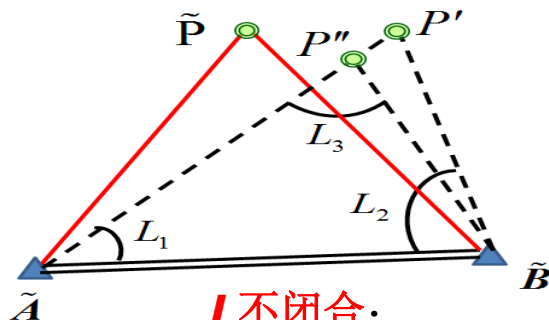
■ 测量平差survey adjustment

处理误差问题的理论和方法。

- 根据几何关系和观测特征，建立平差数学模型； (解决办法)
 - 按一定的平差原则，估算观测值变量分布参数的最优值；
 - 解决观测误差引起的不闭合问题和最优性问题，评价估值精度。
-

1、误差问题与平差的任务

误差问题：观测误差
→带来几何不闭合→
造成工程多解矛盾



L不闭合:
 $L_1 + L_2 + L_3 - 180^\circ = w \neq 0$

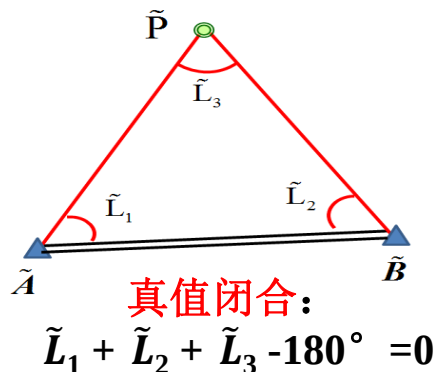
平差的任务：
调整观测值，分配
闭合差；估算 $\tilde{L}$ 的
最优值 $\hat{L}$ 。

$$L_i + v_i = \hat{L}_i \quad w = 3''$$

$v_i = -k_i * w$ 使 $\hat{L}$ 闭合:

↑ $\hat{L}_1 + \hat{L}_2 + \hat{L}_3 - 180^\circ = 0$

最优分配原则 $\Rightarrow$
接近于 $\tilde{L}$

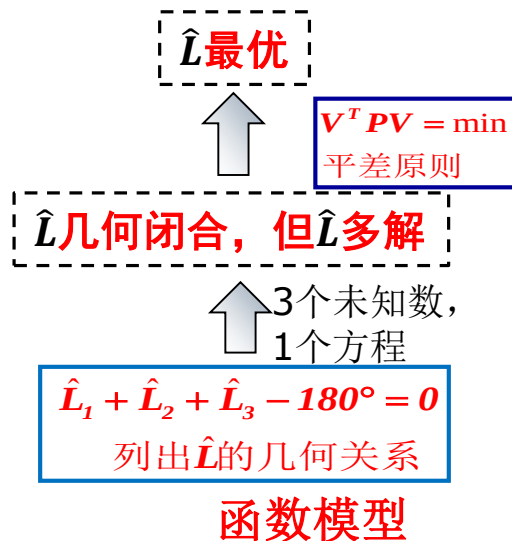
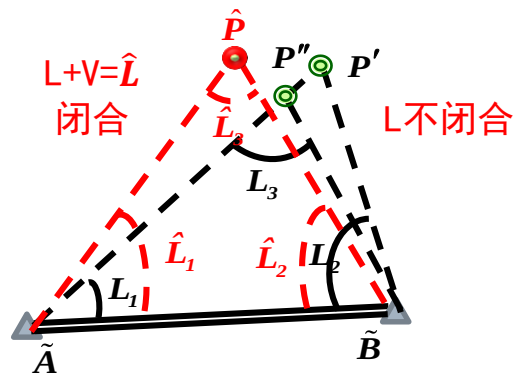


真值闭合:
 $\tilde{L}_1 + \tilde{L}_2 + \tilde{L}_3 - 180^\circ = 0$

就像分奖金
三组测3角，3万奖金
定分配原则 k_i ，对 k_i
的是原则性的粗略要
求，如平均分配1/3，
但奖金数必需细算

2、平差的思路

平差的任务：求解**闭合、最优**唯一的 $\hat{L}$ (平差值)



平差的解题思路：

- 1) 由**函数模型**消除闭合差；
- 2) 依据**最小二乘原则**
估算最优值参数。
- 3) 通过**随机模型****调整**
观测值，求解平差值。

3、平差的数学模型和原则

平差的数学模型 = 函数模型+随机模型

求解的 $\hat{L}$ 的数学关系

- 平差的函数模型—待定的平差值 $\hat{L}=L+V$ 应满足的几何关系

如： $\hat{L}_1 + \hat{L}_2 + \hat{L}_3 - 180^\circ = 0$

- 平差的随机模型—观测值 L 具有的随机特征

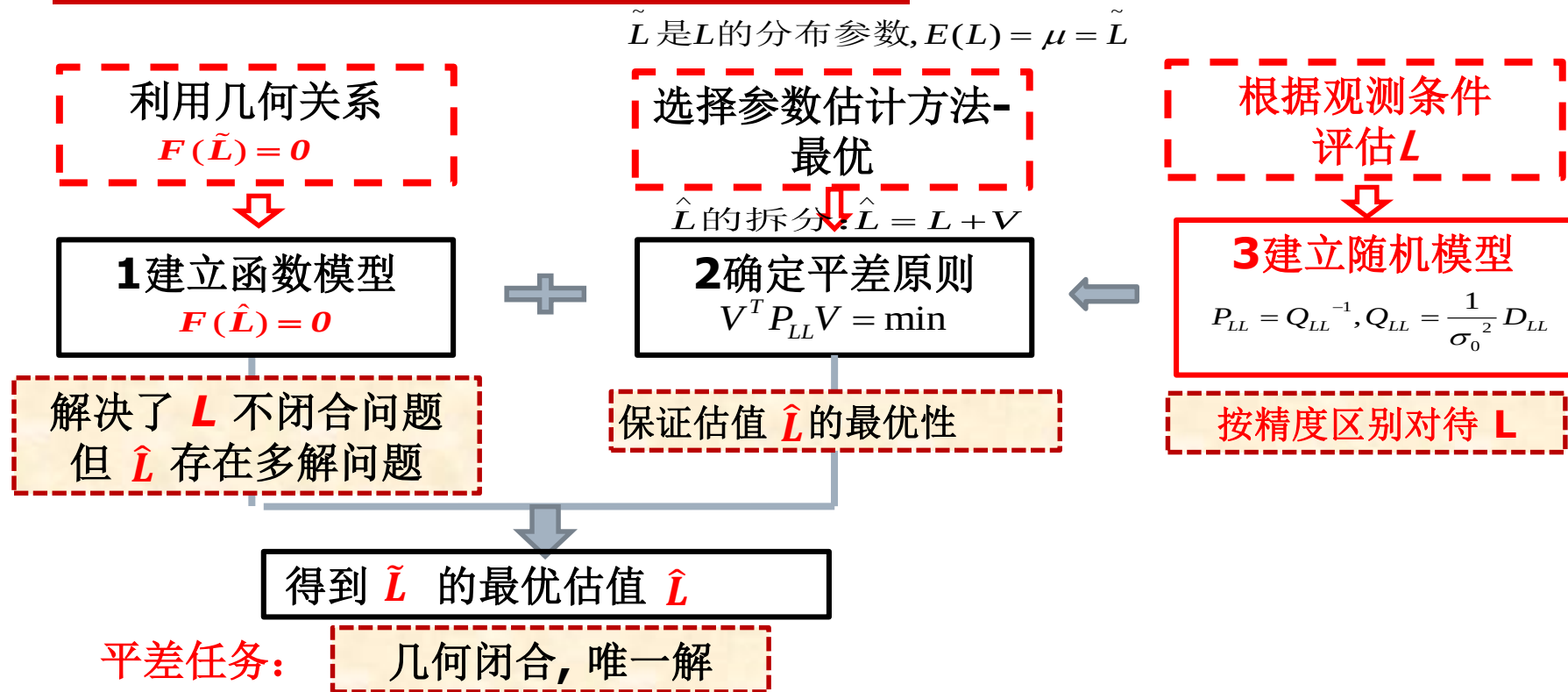
$$P_{LL} = Q_{LL}^{-1} = \sigma_0^2 D_{LL}^{-1}$$

- 平差的原则—最小二乘原则（参数最优估计）

$$V^T P V = \min$$

可得 $\hat{L}$

4、平差的实现过程

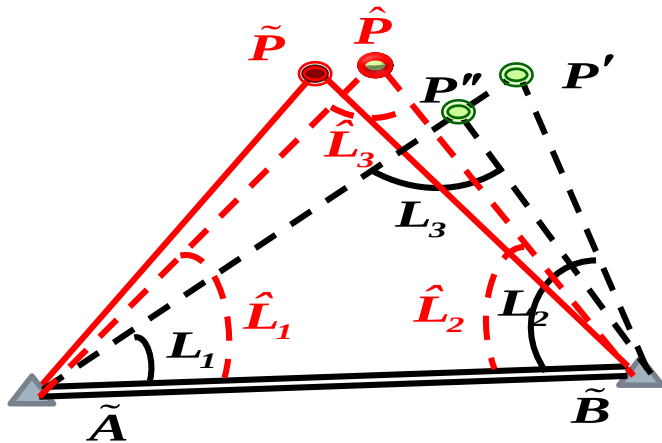


5、平差的效果

平差的效果如何？

$$L \xrightarrow{+V} \hat{L} \rightarrow \tilde{L}$$

$$P', P'' \longrightarrow \hat{P} \rightarrow \tilde{P}$$



以上是本节课的内容。

同学们，下节课再见！

第三章 平差数学模型与原则

§3-1 平差的任务和思路

§3-2 平差的函数模型

§3-3 平差的随机模型

§3-4 参数估计的性质和方法

§3-5 平差的最小二乘原则

§3-2 平差的函数模型

形式、个数、建立方法

平差问题——由观测值求平差值（估计被观测量真值）的问题。

平差的函数模型——待定的平差值 $\hat{L}=L+V$ 应满足的几何关系。

1、函数模型的变量

观测值 L

平差值 $\hat{L}$

改正数 V

变量关系： $\hat{L} = L + V$

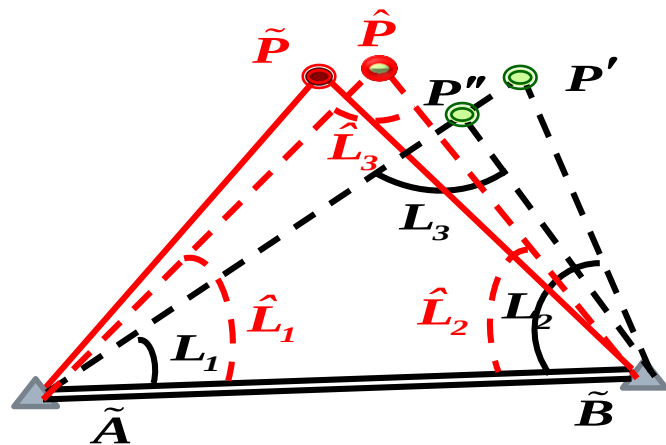
数值上看： L 已知， V 、 L 待定

常量真值 $\tilde{L}$

$$\tilde{L} = L + \Delta$$

L 只含 Δ

$$\hat{L}_{n1} = \begin{pmatrix} \hat{L}_1 \\ \hat{L}_2 \\ \vdots \\ \hat{L}_n \end{pmatrix} \quad L_{n1} = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix} \quad V_{n1} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$



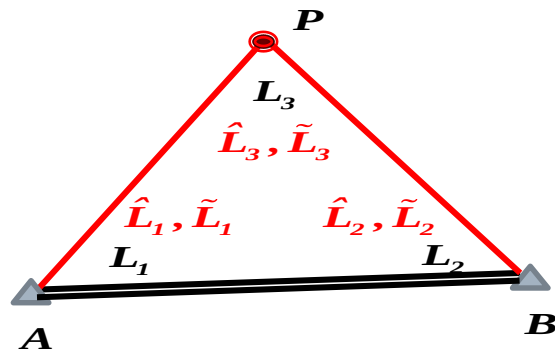
$$\text{如: } \hat{L}_1 + \hat{L}_2 + \hat{L}_3 - 180^\circ = 0$$

2、函数模型的个数

观测的**目的**，是为了**确定几何模型**。

但**测多少个量**合适呢？

由于模型的**几何量之间**存在**函数关系**，因此，要确定一个几何模型不需测定全部、**只需要测定部分几何量**。



1) **必要观测个数** t ：确定一个几何模型所需的最少的观测个数。

必要元素个数 t ：确定一个**几何关系**所需的最少的**元素**个数。

测定三角形形状只需测定2个内角， $t=2$ 。
第3个内角有内角和关系确定。

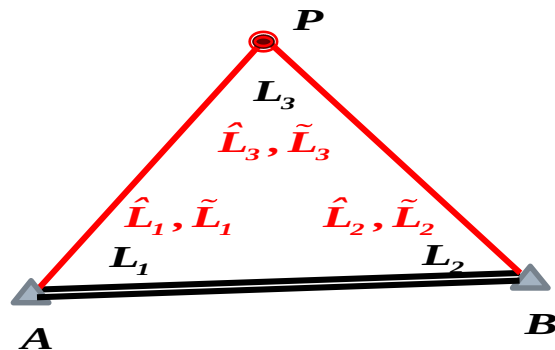
2、函数模型的个数

1) **必要观测个数 t** ：确定一个几何模型所需的最少的观测个数。

观测个数 n 与必要观测个数 t 的**关系**

- 当 $n < t$ 时（少测），模型不能确定。
- 当 $n = t$ 时，模型唯一确定，但不能发现粗差。
- **当 $n > t$ 时**（多测），模型多解，可以发现粗差。

显然少测不行，多测为好！



测定三角形形状只需测定2个内角， $t=2$ 。
第3个内角由内角和关系确定。

2、函数模型的产生（个数）

2) 多余观测个数 r ：观测个数与必要观测个数之差。

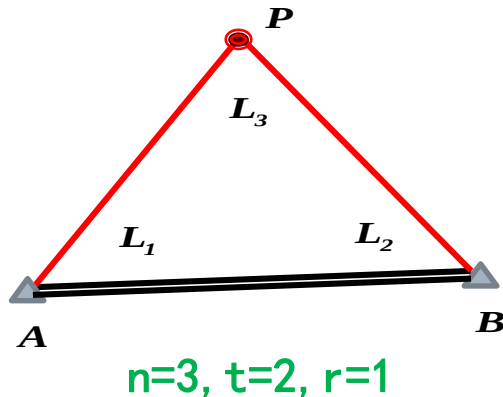
$$r=n-t$$

平差的任务是由 n 个观测值 L 求 n 个平差值 $\hat{L}$ 。需要建立关于 $\hat{L}$ 的函数模型
多余观测个数 r 与函数模型个数的关系：

图例中 $t=2$,

- 当 $n=1, r=0$ 时, $\hat{L}_1$ 不能形成函数模型。
- 当 $n=2, r=0$ 时, $\hat{L}_1, \hat{L}_2$ 不能形成函数模型
- 当 $n=3, r=1$ 时, $\hat{L}_1, \hat{L}_2, \hat{L}_3$ 形成1个函数模型。

$$\hat{L}_1 + \hat{L}_2 + \hat{L}_3 - 180^\circ = 0$$



可见，一个多余观测，产生一个函数模型。 r 个多余观测，产生 r 个独立的函数模型。

3、必要观测的性质与其个数 t 确定方法

◆ 必要观测的性质

- ①必要观测个数 t 取决于观测模型本身；
 - ②必要元素彼此函数独立(不能列出函数关系)，模型中量都可由必要元素来表示；
 - ③一个模型中函数独立的量最多只有 t 个。
-

◆ 必要观测个数 t 的确定

(一) 水准网 t 的确定

1) 有已知水准点时, $t=p$

P --未知点个数

$$p_0=1, p=3, t=3$$

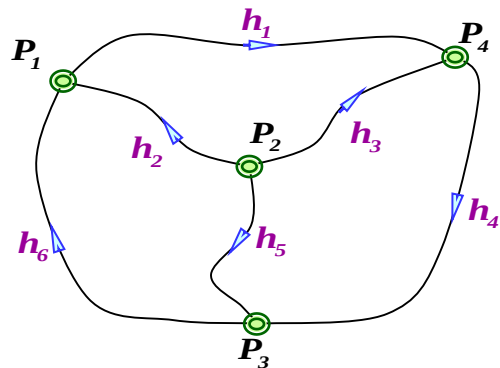
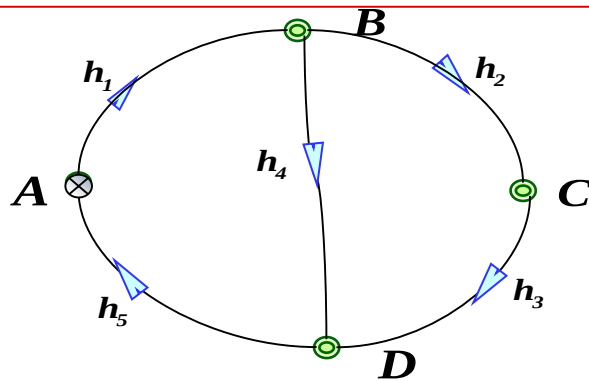
$$n=5, r=2$$

2) 没有已知水准点时, $t=p-1$

假设1个已知点

$$p_0=0, p=4, t=3$$

$$n=6, r=3$$



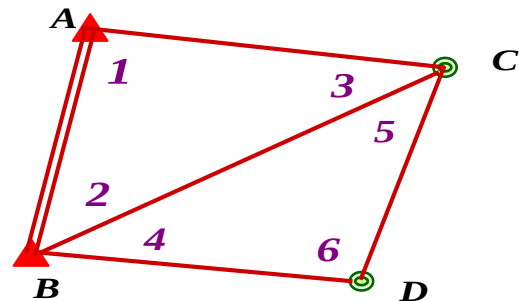
(二) 测角网

1) 有两个或以上已知三角点的三角网

$t=2$ p 未知点个数 p

$p_0=2, p=2, t=4$

$n=6, r=2$

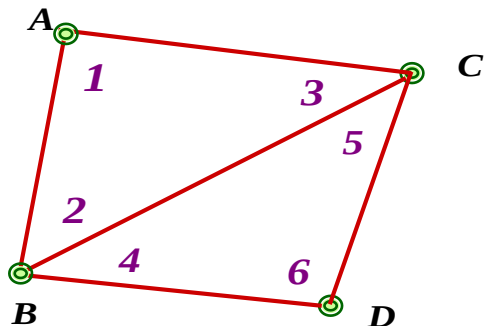


2) 无已知点的三角网

$t=2 * (p-2)$

$p_0=0, p=4, t=4$

$n=6, r=2$



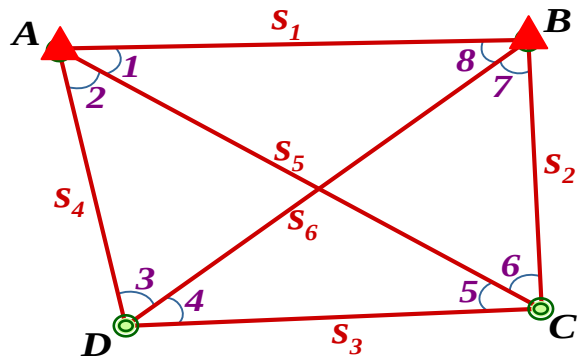
3) 有一个已知点的三角网

$t=2 * (p-1)$

(三) 测边或边角同测网

1) 有两个或以上已知点时

$$t = 2 * p$$

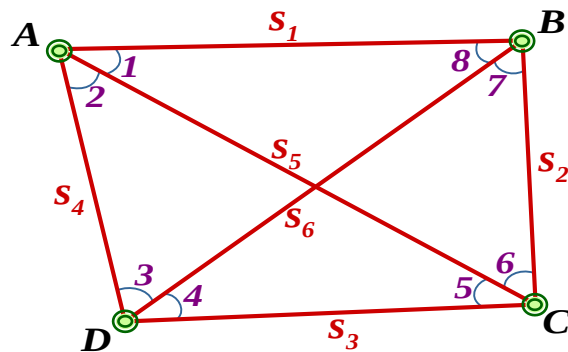


$$p_0 = 2, p = 2, t = 4 \\ n = 8 + 6 = 14, r = 10$$

2) 无已知点时

$$t = 2 * (p - 2) + 1$$

(假定一已知点和一方位角)



$$p_0 = 0, p = 4, t = 2 * 2 + 1 = 5 \\ n = 8 + 6 = 14, r = 9$$

以上是本节课的内容。

同学们，下节课再见！

4、函数模型的形式

函数模型来源于两方面：

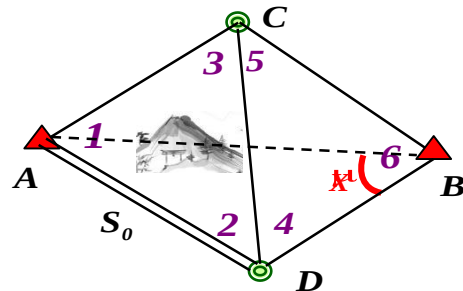
- 1) r 个多余观测，产生 r 个独立的函数模型；
- 2) 引入 u 个独立的未知参数 $\hat{X}$ ，相当于增加了多余观测，增加 u 个独立的函数模型。

函数模型个数： $c=r+u$

函数模型—平差值 $\hat{L}$ 、 $\hat{X}$ 应满足的几何关系

平差（求 $\hat{L}$ ）时，为什么要引入参数 $\hat{X}$ ？ 为了解题方便！

由于引入参数的情况不同，产生4种基本形式的函数模型，产生4种基本平差方法。



4、函数模型的形式

函数模型（关于 $\hat{L}$ 、 $\hat{X}$ 的关系式）来源于多余观测和引入参数，函数模型个数 $c=r+u$ 。引入参数情况不同，产生**4种**形式的**函数模型**和**4种基本平差方法**。

函数模型与引入参数个数的关系

引入参数 $\hat{X}$ 个数 u	函数模型			平差方法名称
	个数	形式	名称	
$u=0$	$c=r$	$F_{r1}(\hat{L}_{n1})=0$	条件方程	条件平差(4)
$u < t$ 独立	$c=r+u$	$F_{c1}(\hat{L}_{n1}, \hat{X}_{u1})=0$	附有参数的条件方程	附有参数的条件平差(5)
$u=t$ 独立	$c=r+t=n$	$\hat{L}_{n1} = F_{n1}(\hat{X}_{t1})$	误差方程	间接平差(6)
$u > t$ 其中 $s=u-t$ 不独立	$c=r+u$ $=n+s$	$\hat{L}_{n1} = F_{n1}(\hat{X}_{u1})$ $\Phi_{s1}(\hat{X}_{u1})=0$	误差方程+限制 条件方程	附有限制条件的间接平差(7)

函数模型变换 $(\hat{L}, \hat{X}) \rightarrow (V, \hat{x})$

作变量拆分: $\hat{L}_{n1} = L_{n1} + V_{n1}$ $\hat{X}_{u1} = \hat{X}_{u1}^0 + \hat{x}_{u1}$

方法平差	引入参数 $\hat{X}$ 个数 u	$\hat{L}, \hat{X}$ 的函数模型			$V, \hat{x}$ 的函数模型
		名称	非线性	线性	
条件平差	$u=0$	条件方程	$F_{r1}(\hat{L}_{n1}) = 0$	$A_{rn\ n1} \hat{L}_{n1} + A_{r1} = 0$	$A_{rn\ n1} V_{n1} + W_{r1} = 0$ $W_{r1} = F(L) = AL + A_0$
附参条件平差	$u < t$ 独立 $\hat{X}_{u1}$	附参条件方程	$F_{c1}(\hat{L}_{n1}, \hat{X}_{u1}) = 0$	$A_{cn\ n1} \hat{L}_{n1} + B_{cu\ u1} \hat{X}_{u1} + A_{c1} = 0$	$A_{cn\ n1} V_{n1} + B_{cu\ u1} \hat{x}_{u1} + W_{c1} = 0$ $W_{c1} = F(L, X^0) = AL + BX^0 + A_0$
间接平差	$u=t$ 独立 $\hat{X}_{t1}$	误差方程	$\hat{L}_{n1} = F_{n1}(\hat{X}_{t1})$	$\hat{L}_{n1} = B_{nt\ t1} \hat{X}_{t1} + d_{n1}$	$V_{n1} = B_{nt\ t1} \hat{x}_{t1} - l_{n1}$ $l = L - F(X)$ $= L - (BX + d)$
附限间接平差	$u > t$ 其中 $s = u - t$ 不独立 $\hat{X}_{u1}$	误差方程+限制条件方程	$\hat{L}_{n1} = F_{n1}(\hat{X}_{u1})$ $\Phi_{s1}(\hat{X}_{u1}) = 0$	$\hat{L}_{n1} = B_{nu\ u1} \hat{X}_{u1} + d_{n1}$ $C_{su\ u1} \hat{X}_{u1} - C_{s1} X^0 + W_{s1}^x = 0$	$V_{n1} = B_{nu\ u1} \hat{x}_{u1} - l_{n1}$ $C_{su\ u1} \hat{x}_{u1} + W_{s1}^x = 0$ $W_{s1}^x = \Phi_s(X^0)$

$$A = \left(\frac{\partial F}{\partial \hat{L}} \right)_0 \quad B = \left(\frac{\partial F}{\partial \hat{X}} \right)_0 \quad C = \left(\frac{\partial \Phi_s}{\partial \hat{X}} \right)_0$$

函数模型线性化

4种函数模型中涉及多个关于 $\hat{L}$, $\hat{X}$ 的非线性函数, 为了简便运算和精度评价, 对其在 L 、 X^0 处进行线性化

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \frac{f'(\mathbf{x}_0)}{1!}(\mathbf{x} - \mathbf{x}_0)$$

$$\hat{L}_{n1} = L_{n1} + V_{n1}$$

$$\hat{X}_{u1} = X_{u1}^0 + \hat{x}_{u1}$$

$$1) F_{r1}(\hat{L}_{n1}) \approx F_{r1}(L_{n1}) + \left(\frac{\partial F}{\partial \hat{L}}\right)_o(\hat{L} - L) = A_{rn} V_{n1} + W_{r1}$$

$$A = \left(\frac{\partial F}{\partial \hat{L}}\right)_o$$

$$B = \left(\frac{\partial F}{\partial \hat{X}}\right)_o$$

$$C = \left(\frac{\partial \Phi_s}{\partial \hat{X}}\right)_o$$

$$2) F_{c1}(\hat{L}_{n1}, \hat{X}_{u1}) \approx F_{c1}(L_{n1}, X_{u1}^0) + \left(\frac{\partial F}{\partial \hat{L}}\right)_o(\hat{L} - L) + \left(\frac{\partial F}{\partial \hat{X}}\right)_o(\hat{X} - X^0)$$

$$= A_{cn} V_{n1} + B_{cu} \hat{x}_{u1} + W_{c1}$$

函数模型线性化

4种函数模型中涉及多个关于 $\hat{L}$, $\hat{X}$ 的非线性函数, 为了简便运算和精度评价, 对其在 L 、 X^0 处进行线性化

$$\hat{L}_{n1} = L_{n1} + V_{n1} \quad \hat{X}_{u1} = X_{u1}^0 + \hat{x}_{u1}$$

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \frac{f'(\mathbf{x}_0)}{1!}(\mathbf{x} - \mathbf{x}_0)$$

$$\begin{aligned} 3) \quad F_{n1}(\hat{X}_{t1}) &\approx F_{n1}(X_{t1}^0) + \left(\frac{\partial F}{\partial \hat{X}_{nt}}\right)_o(\hat{X}_{t1} - X_{t1}^0) \\ &= B_{nt \ t1} \hat{x}_{t1} + F_{n1}(X_{t1}^0) \end{aligned}$$

$$A = \left(\frac{\partial F}{\partial \hat{L}}\right)_o$$

$$B = \left(\frac{\partial F}{\partial \hat{X}}\right)_o$$

$$4) \quad F_{n1}(\hat{X}_{u1}) \approx B_{nu \ u1} \hat{x}_{u1} + F_{n1}(X_{u1}^0)$$

$$C = \left(\frac{\partial \Phi_s}{\partial \hat{X}}\right)_o$$

$$\begin{aligned} \Phi_{s1}(\hat{X}_{u1}) &\approx \Phi_{s1}(X_{u1}^0) + \left(\frac{\partial \Phi_s}{\partial \hat{X}}\right)_o(\hat{X}_{u1} - X_{u1}^0) \\ &= C_{su \ u1} \hat{x}_{u1} + W_{s1} \end{aligned}$$

以上是本节课的内容。

同学们，下节课再见！

第三章 平差数学模型与原则

§3-1 平差的任务和思路

§3-2 平差的函数模型

§3-3 平差的随机模型

§3-4 参数估计的性质和方法

§3-5 平差的最小二乘原则

§3-3 平差的随机模型

平差的任务—由观测值 L 求平差值 $\hat{L}$ （观测量真值 $\tilde{L}$ 的估值）

$$L_i \sim N(\mu_i, \sigma_i^2)$$

1、平差的随机模型

观测值
$$\underset{n1}{L} = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$$

假设：观测值只含偶然误差，观测成果精度符合预设规律，服从正态分布，可对观测值进行先验评价。平差后还有验后评价。

$$\underset{n1}{\mu} = \underset{n1}{E(L)} = E(\tilde{L} - \Delta) = \tilde{L}_{n1} \quad \hat{\mu} = \hat{L} \text{ 是对 } \mu = \tilde{L} \text{ 的估计}$$

平差的随机模型—是描述观测值先验精度和相关性的指标

$$\underset{nn}{D}_{LL} = \sigma_0^2 \underset{nn}{Q}_{LL} = \sigma_0^2 \underset{nn}{P}_{LL}^{-1} \quad \text{取 } \sigma_0^2$$

简写
$$\underset{nn}{D} = \sigma_0^2 \underset{nn}{Q} = \sigma_0^2 \underset{nn}{P}^{-1}$$

2、随机模型分类

$$\underset{nn}{D} = \underset{nn}{D}_{LL} = E(L - E(L))(L - E(L))^T$$

□ 相关观测值的函数模型

$$D = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

$$Q = \begin{bmatrix} Q_{11} & Q_{12} & \cdots & Q_{1n} \\ Q_{21} & Q_{22} & \cdots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \cdots & Q_{nn} \end{bmatrix}$$

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}$$

□ 独立观测值的函数模型

取 σ_0^2

$$D = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}$$

$$Q = \begin{pmatrix} Q_{11} & & & \\ & Q_{22} & & \\ & & \ddots & \\ & & & Q_{nn} \end{pmatrix}$$

$$P = \begin{pmatrix} P_{11} & & & \\ & P_{22} & & \\ & & \ddots & \\ & & & P_{nn} \end{pmatrix} = \begin{pmatrix} 1/Q_{11} & & & \\ & 1/Q_{22} & & \\ & & \ddots & \\ & & & 1/Q_{nn} \end{pmatrix} = \begin{pmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_n \end{pmatrix}$$

第三章 平差数学模型与原则

§3-1 平差的任务和思路

§3-2 平差的函数模型

§3-3 平差的随机模型

§3-4 参数估计的性质和方法

§3-5 平差的最小二乘原则

§3-4 参数估计的性质和方法

1、参数估计的性质

参数估计：数理统计中，由随机变量的样本值来估计其分布参数。

观测值
正态随机
变量

$$L: (\mu, \sigma^2) \quad f(L) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(L-\mu)^2}{2\sigma^2}}$$
$$\mu = E(L) \stackrel{\text{只含}\Delta}{=} E(\tilde{L} - \Delta) = \tilde{L}$$
$$\bar{L} = \frac{1}{2}(L_1 + L_2) \text{ 是一个参数 } \mu \text{ 的一个估计量}$$

平差的任务—由观测值 L 求平差值 $\hat{L}$ （观测量真值 $\tilde{L}$ 的估值）

$\hat{\mu} = \hat{L}$ 是对分布参数 $\mu = \tilde{L}$ 的估计，而且需要最优估计

参数估计的效果如何评价？—参数估计性质（无偏性、一致性、有效性）

如何找到最优估计？—参数估计方法（点估计：极大似然法、最小二乘法
区间估计）

1、参数估计的性质

1) 无偏性

$\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ 为随机变量 $\mathbf{X}$ 分布参数 θ 的估计量，是样本 $x_1, x_2, \dots, x_n$ 的函数式

若 $E(\hat{\theta}) = \theta$ 则称 $\hat{\theta}$ 为 θ 的无偏估计量，否则为有偏估计量。

$\bar{L} = \frac{1}{2}(L_1 + L_2)$ 是参数 μ 的无偏估计量

$x = L_1 + L_2$ 是参数 μ 的有偏估计

2) 一致性

若 $\lim_{n \rightarrow \infty} D(\hat{\theta}) = 0$

则称 $\hat{\theta}$ 为 θ 的一致估计。

一致 + 无偏，则称严格一致估计。

1、参数估计的性质

当 θ 有多个估计量时，哪个更好？

3) 有效性（最有效）

若 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$ 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。

若 $D(\hat{\theta}_m) = \min \{ D(\hat{\theta}_1), D(\hat{\theta}_2) \}$ 则称 $\hat{\theta}_m$ 最有效估计量。

单参数Cramer-Rao不等式，提供了一种求估计量最小方差的方法

$$D(\hat{\theta}) \geq \frac{-1}{E\left(\frac{\partial^2}{\partial \theta^2} \ln G\right)} = D(\hat{\theta}_m)$$

G为似然函数，即观测样本联合密度函数

1、参数估计的性质

当 θ 有多个估计量时，哪个更好？

若 $\hat{\theta}$ 具有“无偏+一致+最有效”，则称之为最优无偏估计。

若 $\hat{\theta}$ 无偏+最有效，则 $\hat{\theta}$ 具有一致性，进而 $\hat{\theta}$ 为最优无偏估计量。

在测量平差中，参数估计函数模型常为线性，最优估计量称为最优线性无偏估计量。

例1：某总体观测值 $x \sim N(\mu, \sigma^2)$ ，其子样观测值为 $x_1, x_2, \dots, x_n$ 。试证，子样的算术平均值为母体均值的**无偏估计**。

证：因母体 $x \sim N(\mu, \sigma^2)$ ，则 母体均值 $E(x) = \mu$

子样分布同母体，

即 子样 $x_i \sim N(\mu, \sigma^2)$ ， $E(x_i) = \mu \quad i=1, 2, \dots, n$

估计量-子样观测值的算术平均值 $\bar{x}$

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

$\bar{x}$ 的数学期望

$$\begin{aligned} E(\bar{x}) &= \frac{1}{n}(E(x_1) + E(x_2) + \dots + E(x_n)) \\ &= \frac{1}{n} n\mu = \mu \end{aligned}$$

所以， $\bar{x}$ 是 μ 的无偏估计。

特别地

$$\bar{x}_2 = \frac{1}{2}(x_1 + x_2) \quad \bar{x}_3 = \frac{1}{3}(x_1 + x_2 + x_3)$$

$$E(\bar{x}_2) = \frac{1}{2}[E(x_1) + E(x_2)] = \mu$$

$$E(\bar{x}_3) = \frac{1}{3}[E(x_1) + E(x_2) + E(x_3)] = \mu$$

$\bar{x}_2$ 是 μ 的无偏估计， $\bar{x}_3$ 也是有偏

实际中无偏估计的要求**较易满足**

例2: 某总体观测值 $x \sim N(\mu, \sigma^2)$, 其子样观测值为 $x_1, x_2, \dots, x_n$ **独立、不含系统误差**。试证, 子样的算术平均值为母体均值的**严格一致估计**。

证: 因母体 $x \sim N(\mu, \sigma^2)$, 则 母体均值 $E(x) = \mu = \tilde{L}$, $D(x) = \sigma^2$
子样分布同母体, 子样 $x_i \sim N(\mu, \sigma^2)$, $E(x_i) = \mu$, $D(x_i) = \sigma^2$

估计量 $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$

因为 $E(\bar{x}) = \mu$ $\bar{x}$ 是 μ 的**无偏**估计量

$$D(\bar{x}) = D\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n D(x_i) = \frac{1}{n} \sigma^2 \quad x_i \text{ 独立}$$

$$\lim_{n \rightarrow \infty} D(\bar{x}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sigma^2 = 0 \quad \bar{x} \text{ 是 } \mu \text{ 的 } \textbf{一致估计量}$$

所以, $\bar{x}$ 是 μ 的**严格一致估计量**

例3: 某总体观测值 $x \sim N(\mu, \sigma^2)$, 其子样观测值为 $x_1, x_2, \dots, x_n$ **独立、不含系统误差**。试证, 子样的算术平均值为母体均值的最**优无偏估计**。

证明: 前已证明估计量 $\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$ 的特征

$$E(\bar{x}) = \mu \quad D(\bar{x}) = \frac{1}{n} \sigma^2$$

$$\lim_{n \rightarrow \infty} D(\bar{x}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sigma^2 = 0$$

(无偏性+一致性+最有效性)



最优无偏

$\bar{x}$ 是母体均值 μ 的无偏、一致估值, 现只需证明 $\bar{x}$ 的**最有效性**。

一组独立同精度子样观测值, 其似然函数为:

$$G = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}$$

$$\ln G = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

例3: 某总体观测值 $x \sim N(\mu, \sigma^2)$, 其子样观测值为 $x_1, x_2, \dots, x_n$ 独立、不含系统误差。试证, 子样的算术平均值为母体均值的最优无偏估计。

似然函数 $G = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}$

$$\ln G = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

对参数 μ 求导: $\frac{\partial}{\partial \mu} \ln G = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = \frac{1}{\sigma^2} \left(\sum_{i=1}^n x_i - n\mu \right)$

$$\frac{\partial^2}{\partial \mu^2} \ln G = -\frac{1}{\sigma^2} n$$

代入罗-克拉美不等式得

$$\sigma_{\min}^2 = \frac{-1}{E\left(\frac{\partial^2}{\partial \mu^2} \ln G\right)} = \frac{\sigma^2}{n}$$

所以 $D(\bar{x}) = \sigma_{\min}^2$

算术平均值 $\bar{x}$ 是观测值母体均值 μ 的无偏、最有效估计量, 因而是最优无偏估计量

小结

1、参数估计的性质

(无偏性、一致性、有效性)



(严格一致+最有效性)



最优无偏

无偏+最有效



无偏+一致



严格一致+最有效性



最优无偏

当存在多种估计量的不确定问题时，如何找最优估计？下节再讲

以上是本节课的内容。

同学们，下节课再见！

第三章 平差数学模型与原则

§3-1 平差的任务和思路

§3-2 平差的函数模型

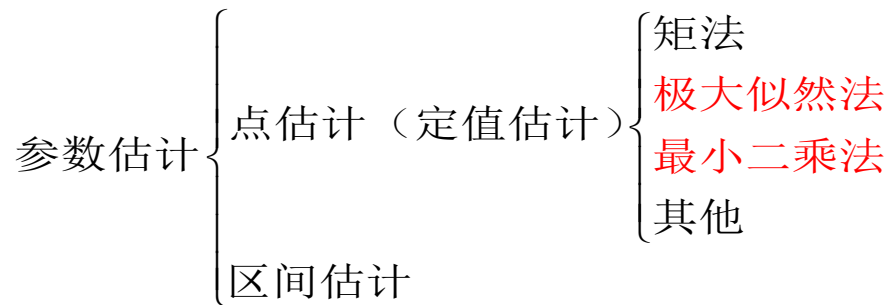
§3-3 平差的随机模型

§3-4 参数估计的性质和方法

§3-5 平差的最小二乘原则

1、参数估计的方法

参数估计方法有**两类**：



区间估计法，是通过设置置信度，求参数的**取值区间**

下面介绍两种**常用的**点参数估计方法

1) 极大似然法

Maximum likelihood method

2) 最小二乘法

Least square method

2、极大似然法

参数估计：数理统计中，利用子样估计母体的分布参数。**极大似然法**是一种常见的估计方法。

问题：设母体 x 的分布密度函数 $f(x;\theta)$ 已知，参数 θ 待估计。由一组子样值有多个参数估计量 $\hat{\theta}$ ，哪种估计量好？

(1) 极大似然法的**定参原则**：估计量 $\hat{\theta}$ 应使子样共同出现的概率最大。

(2) 极大似然估计方法-求估值的过程

总体 x 的子样观测值为 x_i ， x_i 在邻域 dx 上的概率为 $f(x_i;\theta)dx$ ，子样相互独立，则 n 个子样同时出现的概率为

$$P = f(x_1; \theta)dx f(x_2; \theta)dx \cdots f(x_n; \theta)dx = \prod_{i=1}^n f(x_i; \theta)(dx)^n$$

取似然函数

$$G = \prod_{i=1}^n f(x_i; \theta)$$

2、极大似然法

取似然函数 $G = \prod_{i=1}^n f(x_i; \theta)$

按照极大似然法的要求，当 $P = \max$ ，即 $G = \max$ 时，所求 $\hat{\theta}$ 最佳。

似然函数 G 取极值的必要条件为：

$$\frac{\partial G}{\partial \hat{\theta}} = \frac{\partial}{\partial \hat{\theta}} \prod_{i=1}^n f(x_i; \hat{\theta}) = 0$$

$$\text{或 } \frac{\partial \ln G}{\partial \hat{\theta}} = \frac{\partial}{\partial \hat{\theta}} \prod_{i=1}^n \ln f(x_i; \hat{\theta}) = 0$$

由此得到的参数估计量 $\hat{\theta}$ 称为极大似然估计量。

极大似然估计，是渐进的有效、无偏、最优估计。但其缺点是 x 的 $f(x; \theta)$ 需已知。

3、最小二乘法

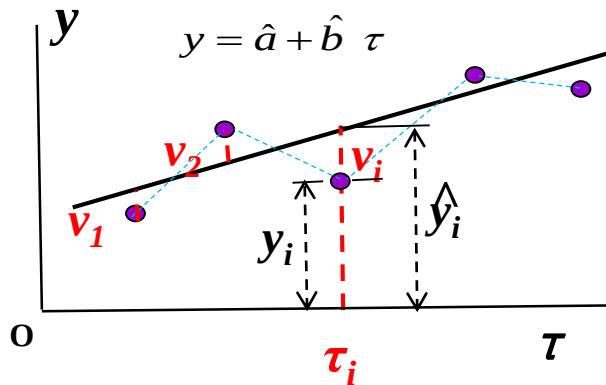
实际中，**分布函数不知道**，需要利用一组观测数据来估计未知参数。

例如，匀速运动物体的位置与时刻的物理（理论）关系为 $y = a + b\tau$ 。 y_i 是不同时刻 τ_i 的一组观测值。**如何求参数估值 $\hat{a}$ 、 $\hat{b}$ ，得到 y - τ 最佳的拟合关系？**

即由 n 个观测值 y_i ，拟合 y 与 τ 的“最佳”物理关系。

(1) 拟合指标:位置估计值 $\hat{y}_i$ 与其观测值 y_i 的**拟合偏差** v_i 来衡量。 n 个 v 如何要求？

采用何种拟合原则？



3、最小二乘法

(2) **拟合原则**-待定的参数应达到的拟合目标

拟合偏差: $v_i = \hat{y}_i - y_i = \hat{a} + \hat{b}\tau_i - y_i$

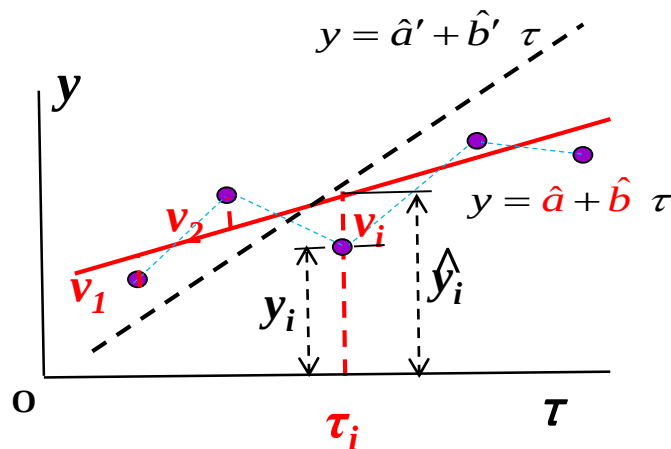
原则一: $|v_1| + |v_2| + \dots + |v_n| = \min$

原则二: $v_1^2 + v_2^2 + \dots + v_n^2 = \min$

$$V^T V = \min$$

原则二—拟合偏差的**平方和**最小。称为**最小二乘原则**, 1794年高斯提出。

最小二乘估计是**最优**的线性无偏估计。



3、最小二乘法

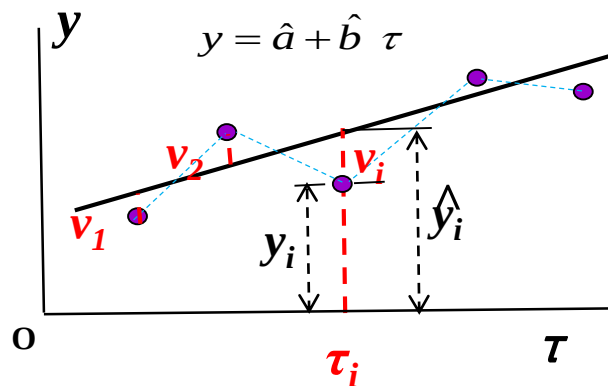
(3) 求参 $\hat{a}$ 、 $\hat{b}$ 模型—使 $V^T V = \sum v_i^2 = \min$ 的必要条件

$$v_i = \hat{y}_i - y_i = \hat{a} + \hat{b}\tau_i - y_i$$
$$\frac{\partial(V^T V)}{\partial \hat{a}} = 2 \sum_{i=1}^n v_i = 2 \sum_{i=1}^n (\hat{a} + \hat{b}\tau_i - y_i)$$

$$= 2(n\hat{a} + \sum_{i=1}^n \tau_i \hat{b} - \sum_{i=1}^n y_i) = 0$$

$$\frac{\partial(V^T V)}{\partial \hat{b}} = 2 \sum_{i=1}^n \tau_i v_i = 2 \sum_{i=1}^n (\tau_i \hat{a} + \tau_i^2 \hat{b} - \tau_i y_i)$$

$$= 2(\sum_{i=1}^n \tau_i \hat{a} + \sum_{i=1}^n \tau_i^2 \hat{b} - \sum_{i=1}^n \tau_i y_i) = 0$$



由此可求得拟合参数 $\hat{a}$ 、 $\hat{b}$ 。此种参数估计方法即最小二乘法，估计量具有最优线性无偏性。

若考虑观测值的精度和相关性，最小二乘原则可表示为： $V^T P V = \min$

4、极大似然法与最小二乘法的关系

极大似然法求解正态分布参数

若总体观测值为正态 $L \sim N(\mu, \sigma^2)$,

独立子样观测值 $L_i (i = 1, 2 \dots n)$ 的联合概率密度函数 $f(L_1, L_2 \dots L_n; \mu_L)$,

$\mu_L = (\mu_1, \mu_2 \dots \mu_n)^T$ 为 n 个待求参数。

似然函数为:

$$\begin{aligned} G &= f_L(L_1, L_2 \dots L_n; \mu_L) = \prod_{i=1}^n f(L_i; \mu_i) \\ &= \frac{1}{(2\pi)^{n/2} |D|^{1/2}} \exp\left[-\frac{1}{2} (L - \mu_L)^T D^{-1} (L - \mu_L)\right] \end{aligned}$$

$$\ln G = -\ln[(2\pi)^{n/2} |D|^{1/2}] - \frac{1}{2} (L - \mu_L)^T D^{-1} (L - \mu_L)$$

4、极大似然法与最小二乘法的关系

现按极大似然法要求 $G=\max$ ，求参数 $\mu_L=\tilde{L}$ 的估值 $\hat{\mu}_L=\hat{L}$ ，使

$$\ln G = -\ln[(2\pi)^{n/2} |D|^{1/2}] - \frac{1}{2} (\underline{L} - \hat{L})^T D^{-1} (\underline{L} - \hat{L}) = \max$$

$$= -\ln[(2\pi)^{n/2} |D|^{1/2}] - \frac{1}{2} (-V)^T \frac{1}{\sigma_0^2} P(-V) = \max$$



$$V^T P V = \min$$

即，对于正态分布观测值，由极大似然估计法导出了最小二乘原则。
两者有相同的求参要求。

这就是我们寻求和采纳的平差原则：最小二乘原则

以上是本节课的内容。

同学们，下节课再见！
