6.1.2 Gauss-Markov 定理：

特定条件下 LS 最好，使得 LS 变得伟大

在参数估计理论中，称参数向量 θ 的估计 $\hat{\theta}$ 为无偏估计，若 $E(\hat{\theta})=\theta$（真实的未知参数向量）。如果一个无偏估计还具有最小方差则称这一无偏估计为最优无偏估计

* 最优无偏解（统计信号处理）

须满足的条件：

$\underline{A}\,\underline{x} = \underline{b} + \underline{e}$ → 随机误差向量

误差均值为 0 $E\{\underline{e}\}=0$

协方差 $\mathrm{cov}\{\underline{e}\} = E\{\underline{e}\,\underline{e}^H\} = \sigma^2 \underline{I}$ 高斯白噪声

↓ 最优无偏解由最小二乘解给出

* 最大似然解 6.1.3

（概率）

3. LS 解与最大似然解的等价性（了解）

LS 视广泛应用

复高斯随机向量（加性误差向量 $\underline{e}=[e_1, e_2 \dots e_m]^T$ 为高斯分布）

$$f(\underline{e}) = \frac{1}{\pi^m |\underline{\Gamma}_e|} \exp[-(\underline{e}-\underline{u}_e)^H \underline{\Gamma}_e^{-1} (\underline{e}-\underline{u}_e)]$$

概率密度函数

协方差矩阵：

误差向量的各分量均具有零均值和相同方差 σ^2，

$\underline{\Gamma}_e = \sigma^2 \underline{I}$

$$= \frac{1}{(\pi\sigma^2)^m} \exp[-\frac{1}{\sigma^2} \underline{e}^H \underline{e}]$$

$\|\underline{A}\underline{x} - \underline{b}\|^2$

当误差向量 $\underline{e}$ 为零均值的高斯随机向量，但其元素具有不同方差时，这种情况下的最大似然解将不可能等于最小二乘解公式，最小二乘解不再是最优的

估计 $\underline{x}$ 使得 $f(\underline{e})$ 最大 $\Leftrightarrow \min \|\underline{A}\underline{x} - \underline{b}\|_2^2$

4. 数据最小二乘

$\underline{A} = \underline{A}_0 + \underline{E}$ $\underline{A}$：数据矩阵：观测误差或噪声

对 $\underline{A}$ 校准 $(\underline{A} + \Delta\underline{A})\underline{x} = \underline{b}$

$\min_{\underline{x}} \|\Delta \underline{\underline{A}}\|_F^2$, s.t. $(\underline{\underline{A}}+\Delta\underline{\underline{A}})\underline{x}=\underline{b}$

F范数

$\min L(\underline{x})$

$=tr(\Delta\underline{\underline{A}}(\Delta\underline{\underline{A}})^H)$ 迹 $+\underbrace{\underline{\lambda}^H}_{1\times m}(\underbrace{\underline{\underline{A}}}_{m\times n}\underbrace{\underline{x}}_{n\times 1}+\Delta\underline{\underline{A}}\underline{x}-\underbrace{\underline{b}}_{m\times 1})$

拉格朗日乘子法

min[ ] 无约束优化问题

$\frac{\partial L(\underline{x})}{\partial \Delta\underline{\underline{A}}^T}=0 \Rightarrow \Delta\underline{\underline{A}}=-\underline{\lambda}\,\underline{x}^H$ $\quad \Delta\underline{\underline{A}}=-\underline{\lambda}\,\underline{x}^H$

$(\Delta\underline{\underline{A}})^H+\underline{x}\,\underline{\lambda}^H=0$

应该是对$\Delta\underline{\underline{A}}^T$求导

浙江大学
ZHEJIANG UNIVERSITY

第八讲

$Ax=b$. 与普通最小二乘不同，这里假定数据向量 b 无观测误差或噪声，只有数据矩阵 $A=A_0+E$ 有观测误差或噪声，且误差矩阵 E 的每个元素符合零均值、等方差的独立高斯分布。

*. 考虑用校正矩阵 ΔA 干扰有误差的数据矩阵 A，使得

$$A+\Delta A=A_0+E+\Delta A\to A_0.$$

---

* 将 $\Delta A=-\lambda x^H$ 代入约束条件 $(A+\Delta A)x=b$，即有 $\lambda=\dfrac{Ax-b}{x^Hx}$

从而有 $\Delta A=-\dfrac{(Ax-b)x^H}{x^Hx}$，原目标函数

$$J(x)=\|\Delta A\|_F^2=tr\left[\Delta A(\Delta A)^H\right]=tr\left(\frac{(Ax-b)x^H}{x^Hx}\,\frac{x(Ax-b)^H}{x^Hx}\right)$$

$$tr(BC)=tr(CB)$$

$$J(x)=\frac{(Ax-b)^H(Ax-b)}{x^Hx}$$

由此得：$\hat{x}_{DLS}=\underset{x}{\arg\min}\dfrac{(Ax-b)^H(Ax-b)}{x^Hx}$