统计信号分析与处理

第4章 信号估计理论

本章内容

4.1 引言

4.2 估计准则

4.3 多参量的常用估计准则

4.4 估计量评价的指标

4.5 克拉美-罗不等式

4.6 应用

4.1 引言

- 信号检测：判断是否存在信号或存在哪种信号
  信号估计：对信号的参量甚至波形进行定量的推断

- 从信号检测到信号估计，是对事物从定性的判断到定量的描述。

不同应用领域：

- 数理统计领域：估计总体的均值、方差、各阶矩、相关函数等；
- 信息与通信工程领域：估计信号的振幅、相位、频率、时延等；
- 控制工程领域：估计动态系统的参量和状态，如飞行体的质量、位置、速度、加速度等；
- 经济领域：估计、预测各种反应经济运行的指标，如人均国民生产总值、物价指数等。

分类

- 确定参量估计和随机参量估计
- 单维（标量）参量估计和多维（矢量）参量估计
- 时变参量估计和时不变参量估计
- 线性参量估计与非线性参量估计

4.2 估计准则

4.2.1 最大后验概率估计

4.2.2 最大似然估计

4.2.3 最小均方误差估计

4.2.4 线性最小均方误差估计

4.2.5 最小平均绝对误差估计

4.2.6 贝叶斯估计

4.2.7 最小二乘估计

观测信号

$$x(t) = s(t;\theta_1,\theta_2,\cdots,\theta_M) + n(t) \quad 0 \le t \le T$$

其中 $s(t;\theta_1,\theta_2,\cdots\theta_M)$ 为有用信号，$\boldsymbol{\theta} = [\theta_1,\theta_2,\cdots,\theta_M]^T$ 为待估计参量，$n(t)$ 为观测噪声。

利用N个观测样本 $\mathbf{x} = [x_1, x_2, \cdots, x_N]^T$ 进行估计, 估计量记为

$$\hat{\boldsymbol{\theta}} = g(\mathbf{x})$$

4.2.1 最大后验概率估计

使后验概率密度最大的一种估计，即

$$\hat{\theta}_{MAP} = \arg\max_{\theta} f\left(\theta \mid \mathbf{x}\right)$$

其中 $f\left(\theta \mid \mathbf{x}\right)$ 为单个待估计量 θ 的后验概率密度函数。

估计量 $\hat{\theta}_{MAP}$ 可以通过如下方程得到

$$\left[\frac{\partial}{\partial\theta} f\left(\theta \mid \mathbf{x}\right)\right]\Bigg|_{\theta=\hat{\theta}_{MAP}} = 0 \quad or \quad \left[\frac{\partial}{\partial\theta} \ln f\left(\theta \mid \mathbf{x}\right)\right]\Bigg|_{\theta=\hat{\theta}_{MAP}} = 0$$

利用

$$f(\theta\,|\,\mathbf{x})=\frac{f(\mathbf{x}\,|\,\theta)\,f(\theta)}{f(\mathbf{x})}$$

进一步可得

$$\left[\frac{\partial}{\partial\theta}\ln f(\mathbf{x}\,|\,\theta)+\frac{\partial}{\partial\theta}\ln f(\theta)\right]\Bigg|_{\theta=\hat{\theta}_{MAP}}=0$$

例 4. 1

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$, 噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{MAP}$ 。

4.2.2 最大似然估计

使观测样本的似然函数 $f(\mathbf{x}|\theta)$ 取得最大值的一种估计，即

$$\hat{\theta}_{ML} = \arg\max_{\theta} f(\mathbf{x}|\theta)$$

估计量 $\hat{\theta}_{ML}$ 可以通过如下方程得到

$$\left.\frac{\partial}{\partial\theta} f(\mathbf{x}|\theta)\right|_{\theta=\hat{\theta}_{ML}} = 0 \quad or \quad \left.\frac{\partial}{\partial\theta} \ln f(\mathbf{x}|\theta)\right|_{\theta=\hat{\theta}_{ML}} = 0$$

- 适用于确定参量估计和先验分布未知的随机参量估计。

例 4.2

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$, 噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{ML}$ 。

4.2.3 最小均方误差估计

使估计的均方误差最小的一种估计。定义估计误差

$$e\left(\hat{\theta}\right)=\theta-\hat{\theta}$$

估计的均方误差

$$\xi\left(\hat{\theta}\right)=E\left\{e^{2}\left(\hat{\theta}\right)\right\}=\int_{(\theta)}\int_{(\mathbf{x})}\left(\theta-\hat{\theta}\right)^{2}f\left(\theta,\mathbf{x}\right)d\mathbf{x}d\theta=\int_{(\mathbf{x})}\xi\left(\hat{\theta}\,|\,\mathbf{x}\right)f\left(\mathbf{x}\right)d\mathbf{x}$$

则

$$\hat{\theta}_{MS}=\arg\min_{\hat{\theta}}\xi\left(\hat{\theta}\right)=\arg\min_{\hat{\theta}}\xi\left(\hat{\theta}\,|\,\mathbf{x}\right)$$

利用

$$\left.\frac{\partial}{\partial\hat{\theta}}\xi(\hat{\theta}\,|\,\mathbf{x})\right|_{\hat{\theta}=\hat{\theta}_{MS}}=0$$

可得

$$\hat{\theta}_{MS}=\int_{(\theta)}\theta f\left(\theta\,|\,\mathbf{x}\right)\mathrm{d}\theta=E\left\{\theta\,|\,\mathbf{x}\right\}$$

例 4.3

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$，噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{MS}$ 。

4.2.4 线性最小均方误差估计

最小均方误差估计的一种特例，要求估计量与观测样本之间必须满足线性关系，即：

$$\hat{\theta} = g(\mathbf{x}) = a + \sum_{k=1}^{N} b_k x_k = a + \mathbf{b}^T \mathbf{x}$$

其中　和 $\mathbf{b} = [b_1, b_2, \cdots, b_N]^T$ 是待定系数，根据最小均方误差准则确定。估计的均方误差为

$$\xi(\hat{\theta}) = E\left\{\left[\theta - \left(a + \sum_{k=1}^{N} b_k x_k\right)\right]^2\right\}$$

则

$$\frac{\partial}{\partial a}E\left\{\xi\left(\hat{\theta}\right)\right\}=E\left\{-2\left(\theta-a_L-\sum_{k=1}^{N}b_k x_k\right)\right\}=0$$

$$\frac{\partial}{\partial b_k}E\left\{\xi\left(\hat{\theta}\right)\right\}=-2E\left\{\left(\theta-a_L-\sum_{k=1}^{N}b_k x_k\right)x_k\right\}=0$$

求解得

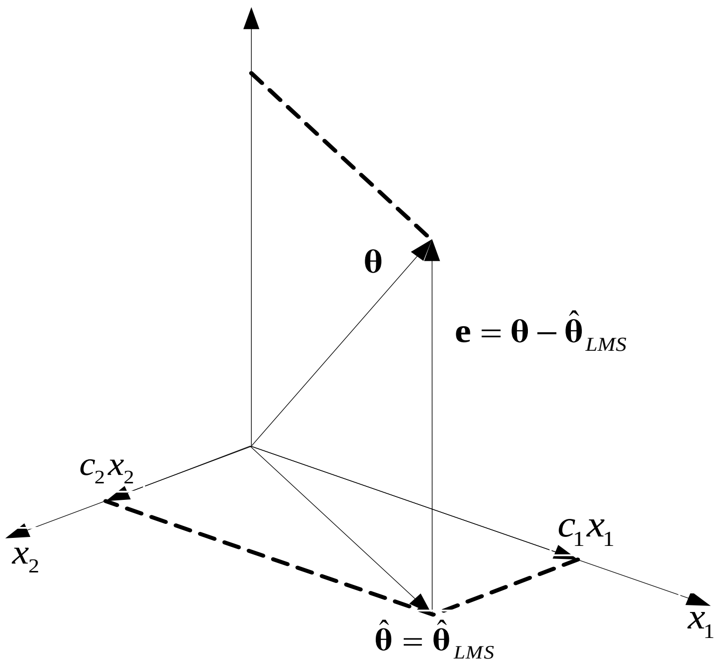
$$a_L=E\{\theta\}-Cov\{\theta,\mathbf{x}\}Cov^{-1}\{\mathbf{x},\mathbf{x}\}E\{\mathbf{x}\}$$

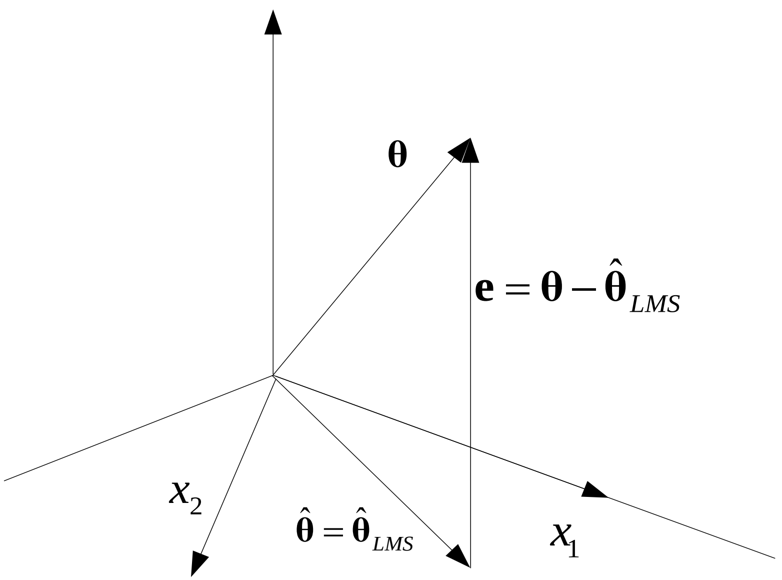
$$\mathbf{b}_L^T=Cov\{\theta,\mathbf{x}\}Cov^{-1}\{\mathbf{x},\mathbf{x}\}$$

即 $\hat{\theta}_{LMS}=a_L+\mathbf{b}_L^T\mathbf{x}=E\{\theta\}+Cov\{\theta,\mathbf{x}\}Cov^{-1}\{\mathbf{x},\mathbf{x}\}\left[\mathbf{x}-E\{\mathbf{x}\}\right]$

正交条件：线性最小均方误差估计的估计误差与观测样本是正交的。

$$E\left\{\left(\theta-\hat{\theta}_{LMS}\right)\mathbf{x}^T\right\}=0$$





例 4.4

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$，噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{LMS}$ 。

4.2.5 最小平均绝对误差估计

使绝对估计误差的统计平均值最小的一种估计。

定义绝对估计误差

$$e_{ABS}\left(\hat{\theta}\right)=\left|\theta-\hat{\theta}\right|$$

其统计平均绝对误差

$$\xi_{ABS}\left(\hat{\theta}\right)=\int_{(\theta)}\int_{(\mathbf{x})}\left|\theta-\hat{\theta}\right|f\left(\theta,\mathbf{x}\right)d\mathbf{x}d\theta=\int_{(\mathbf{x})}\left[\int_{(\theta)}\left|\theta-\hat{\theta}\right|f\left(\theta|\mathbf{x}\right)d\theta\right]f\left(\mathbf{x}\right)d\mathbf{x}$$

$$=\int_{(\mathbf{x})}\xi_{ABS}\left(\hat{\theta}\middle|\mathbf{x}\right)f\left(\mathbf{x}\right)d\mathbf{x}$$

则 $$\hat{\theta}_{ABS} = \arg\min_{\hat{\theta}} \xi_{ABS}\left(\hat{\theta}\right) = \arg\min_{\hat{\theta}} \xi_{ABS}\left(\hat{\theta} \mid \mathbf{x}\right)$$

即 $$\frac{\partial}{\partial \hat{\theta}} \xi_{ABS}\left(\hat{\theta} \middle| \mathbf{x}\right)\Big|_{\hat{\theta}=\hat{\theta}_{ABS}} = 0$$

解得

$$\int_{-\infty}^{\hat{\theta}_{ABS}} f\left(\theta \middle| \mathbf{x}\right) d\theta = \int_{\hat{\theta}_{ABS}}^{+\infty} f\left(\theta \middle| \mathbf{x}\right) d\theta$$

可见，$\hat{\theta}_{ABS}$ 是条件概率密度的中位数，故又称作条件中位数估计。

例 4.5

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$，噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{ABS}$ 。

4.2.6 贝叶斯估计

使估计所承担的平均风险最小的一种估计。定义 θ 估计为 $\hat{\theta}$ 所承担的风险（代价函数）$c\left(\theta,\hat{\theta}\right)$, 则估计的平均风险为

$$C\left(\hat{\theta}\right)=\int_{(\mathbf{x})}\int_{(\theta)}c\left(\theta,\hat{\theta}\right)f\left(\theta,\mathbf{x}\right)d\theta d\mathbf{x}=\int_{(\mathbf{x})}C\left(\hat{\theta}\,|\,\mathbf{x}\right)f\left(\mathbf{x}\right)d\mathbf{x}$$

即

$$\hat{\theta}_{BAY}=\arg\min_{\hat{\theta}}C\left(\hat{\theta}\right)=\arg\min_{\hat{\theta}}C\left(\hat{\theta}\,|\,\mathbf{x}\right)$$

（1）最小均方误差估计与贝叶斯估计

若定义

$$c\left(\theta,\hat{\theta}\right)=\left(\theta-\hat{\theta}\right)^2$$

平均风险

$$C\left(\hat{\theta}\right)=\int\limits_{(\mathbf{x})}\int\limits_{(\theta)}\left(\theta-\hat{\theta}\right)^2 f\left(\theta,\mathbf{x}\right)d\theta d\mathbf{x}=E\left\{\left(\theta-\hat{\theta}\right)^2\right\}$$

则

$$\hat{\theta}_{BAY}=\hat{\theta}_{MS}$$

（2）最小平均绝对误差估计与贝叶斯估计

若定义

$$c\left(\theta,\hat{\theta}\right)=\left|\theta-\hat{\theta}\right|$$

则平均风险

$$C\left(\hat{\theta}\right)=\int_{(\mathbf{x})}\int_{(\theta)}\left|\theta-\hat{\theta}\right| f\left(\theta,\mathbf{x}\right)d\theta d\mathbf{x}=E\left\{\left|\theta-\hat{\theta}\right|\right\}$$

则 $\hat{\theta}_{BAY}=\hat{\theta}_{ABS}$

（3）最大后验概率估计与贝叶斯估计

若定义

$$c\left(\theta,\hat{\theta}\right)=\begin{cases}1, & \left|\theta-\hat{\theta}\right|\geq\dfrac{\Delta}{2}\\[2ex] 0, & \left|\theta-\hat{\theta}\right|<\dfrac{\Delta}{2}\end{cases}$$

Δ>0是很小的常数

则 $\hat{\theta}_{BAY}=\hat{\theta}_{MAP}$

4.2.7 最小二乘估计

基于参量的线性观测模型，把估计作为确定的最优化问题来处理。

线性观测方程为

$$x_i = h_i\theta + n_i, \quad i = 1, 2, \cdots, N$$

估计的误差平方和为

$$\xi\left(\hat{\theta}\right) = \sum_{i=1}^{N}\left(x_i - h_i\hat{\theta}\right)^2$$

则 $$\hat{\theta}_{LS} = \arg\min_{\hat{\theta}} \xi\left(\hat{\theta}\right) = \arg\min_{\hat{\theta}} \sum_{i=1}^{N}\left(x_i - h_i\hat{\theta}\right)^2$$

即 $$\frac{\partial}{\partial \hat{\theta}} \xi\left(\hat{\theta}\right)\bigg|_{\hat{\theta}=\hat{\theta}_{LS}} = -2\sum_{i=1}^{N}\left(x_i - h_i\hat{\theta}\right)h_i = 0$$

解得

$$\hat{\theta}_{LS} = \frac{\sum_{i=1}^{N} h_i x_i}{\sum_{i=1}^{N} h_i^2}$$

线性观测方程的矢量形式

$$\mathbf{x}=\mathbf{h}\theta+\mathbf{n}$$

$$\mathbf{x}=\left[x_1,x_2,\cdots,x_N\right]^T,\ \mathbf{h}=\left[h_1,h_2,\cdots,h_N\right]^T,\ \mathbf{n}=\left[n_1,n_2,\cdots,n_N\right]^T。$$

误差平方和为

$$\xi\left(\hat{\theta}\right)=\left(\mathbf{x}-\mathbf{h}\hat{\theta}\right)^T\left(\mathbf{x}-\mathbf{h}\hat{\theta}\right)$$

解得

$$\hat{\theta}_{LS}=\left(\mathbf{h}^T\mathbf{h}\right)^{-1}\mathbf{h}^T\mathbf{x}$$

例 4.6

已知如下观测样本

$$x_i = s + n_i, \qquad i = 1, 2, \cdots, N$$

其中信号 $s \sim N\left(0, \sigma_s^2\right)$，噪声 $n_i \sim N\left(0, \sigma_n^2\right)$ 独立同分布，并且信号与噪声不相关。

求 $\hat{s}_{LS}$ 。

4.3 多参量的常用估计准则

多参量矢量

$$\boldsymbol{\theta} = \left[\theta_1, \theta_2, \cdots, \theta_M\right]^T$$

估计量

$$\hat{\boldsymbol{\theta}} = \left[\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_M\right]^T$$

估计误差矢量

$$\mathbf{e}\left(\hat{\boldsymbol{\theta}}\right) = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} = \left[\left(\theta_1 - \hat{\theta}_1\right), \left(\theta_2 - \hat{\theta}_2\right), \cdots, \left(\theta_M - \hat{\theta}_M\right)\right]^T$$

（1）多参量最大后验概率估计

$$\hat{\boldsymbol{\theta}}_{MAP}=\arg\max_{\boldsymbol{\theta}} f\left(\boldsymbol{\theta}\,|\,\mathbf{x}\right)$$

即

$$\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\boldsymbol{\theta}\,|\,\mathbf{x}\right)\right]\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{MAP}}=0$$

其中第 i 个方程为

$$\left[\frac{\partial}{\partial\theta_i}\ln f\left(\boldsymbol{\theta}\,|\,\mathbf{x}\right)\right]\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{MAP}}=0\ ,\quad i=1,2,\cdots,M$$

（2）多参量最大似然估计

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} f\left(\mathbf{x} \mid \boldsymbol{\theta}\right)$$

即

$$\left.\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f\left(\mathbf{x} \mid \boldsymbol{\theta}\right)\right]\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}} = 0$$

其中第 i 个方程为

$$\left.\left[\frac{\partial}{\partial \theta_i} \ln f\left(\mathbf{x} \mid \boldsymbol{\theta}\right)\right]\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}} = 0,\ i = 1,2,\cdots,M$$

（3）多参量最小均方误差估计

$$\hat{\boldsymbol{\theta}}_{MS} = \arg\min_{\hat{\boldsymbol{\theta}}} \boldsymbol{\xi}\left(\hat{\boldsymbol{\theta}}\right) = \arg\min_{\hat{\boldsymbol{\theta}}} \boldsymbol{\xi}\left(\hat{\boldsymbol{\theta}} \mid \mathbf{x}\right)$$

得

$$\hat{\boldsymbol{\theta}}_{MS}\left(\mathbf{x}\right) = \int_{(\boldsymbol{\theta})} \boldsymbol{\theta} f\left(\boldsymbol{\theta} \mid \mathbf{x}\right) d\boldsymbol{\theta}$$

其中第 i 个方程估计量为

$$\hat{\theta}_{iMS}\left(\mathbf{x}\right) = \int_{(\boldsymbol{\theta})} \theta_i f\left(\boldsymbol{\theta} \mid \mathbf{x}\right) d\boldsymbol{\theta} \qquad i = 1, 2, \cdots, M$$

（4）多参量线性最小均方误差估计

线性关系 $\hat{\boldsymbol{\theta}} = \mathbf{a} + \mathbf{B}\mathbf{x}$

估计的均方误差为

$$\xi\left(\hat{\boldsymbol{\theta}}\right) = E\left\{\left[\boldsymbol{\theta} - \mathbf{a} - \mathbf{B}\mathbf{x}\right]^T \left[\boldsymbol{\theta} - \mathbf{a} - \mathbf{B}\mathbf{x}\right]\right\}$$

则

$$\frac{\partial}{\partial \mathbf{a}} E\left\{\xi\left(\hat{\boldsymbol{\theta}}\right)\right\}\Bigg|_{\substack{\mathbf{a}=\mathbf{a}_L \\ \mathbf{B}=\mathbf{B}_L}} = 0 \qquad \frac{\partial}{\partial \mathbf{B}} E\left\{\xi\left(\hat{\boldsymbol{\theta}}\right)\right\}\Bigg|_{\substack{\mathbf{a}=\mathbf{a}_L \\ \mathbf{B}=\mathbf{B}_L}} = 0$$

解得

$$\hat{\boldsymbol{\theta}}_{LMS} = E\left\{\boldsymbol{\theta}\right\} + Cov\left\{\boldsymbol{\theta}, \mathbf{x}\right\} Cov^{-1}\left\{\mathbf{x}, \mathbf{x}\right\}\left[\mathbf{x} - E\left\{\mathbf{x}\right\}\right]$$

正交条件 $E\left\{\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LMS}\right)\mathbf{x}^T\right\} = 0$

（5）多参量最小平均绝对误差估计

$$\hat{\boldsymbol{\theta}}_{ABS}=\arg\min_{\hat{\boldsymbol{\theta}}}\xi\left(\hat{\boldsymbol{\theta}}\right)=\arg\min_{\hat{\boldsymbol{\theta}}}\xi\left(\hat{\boldsymbol{\theta}}\,|\,\mathbf{x}\right)$$

解得

$$\int_{-\infty}^{\hat{\theta}_{iABS}} f\left(\boldsymbol{\theta}\,|\,\mathbf{x}\right)d\boldsymbol{\theta}=\int_{\hat{\theta}_{iABS}}^{+\infty} f\left(\boldsymbol{\theta}\,|\,\mathbf{x}\right)d\boldsymbol{\theta}\quad i=1,2,\cdots,M$$

（6）多参量贝叶斯估计

平均风险为

$$C\left(\hat{\boldsymbol{\theta}}\right)=\int_{(\mathbf{x})}\int_{(\boldsymbol{\theta})}c(\mathbf{e})\,f\left(\mathbf{x},\boldsymbol{\theta}\right)d\boldsymbol{\theta}d\mathbf{x}$$

则

$$\hat{\boldsymbol{\theta}}_{BAY}=\arg\min_{\hat{\boldsymbol{\theta}}}C\left(\hat{\boldsymbol{\theta}}\right)=\arg\min_{\hat{\boldsymbol{\theta}}}C\left(\hat{\boldsymbol{\theta}}\mid\mathbf{x}\right)$$

当 $c(\mathbf{e})=\mathbf{e}\left(\hat{\boldsymbol{\theta}}\right)^T\mathbf{e}\left(\hat{\boldsymbol{\theta}}\right)$ $\qquad\hat{\boldsymbol{\theta}}_{BAY}=\hat{\boldsymbol{\theta}}_{MS}$

当 $c(\mathbf{e})=\sum_{i=1}^{M}\left|\theta_i-\hat{\theta}_i\right|$ $\qquad\hat{\boldsymbol{\theta}}_{BAY}=\hat{\boldsymbol{\theta}}_{ABS}$

（7）多参量最小二乘估计

多参量线性观测方程为 $\mathbf{x}=\mathbf{H}\boldsymbol{\theta}+\mathbf{n}$

误差的平方和为 $\xi\left(\hat{\boldsymbol{\theta}}\right)=\left[\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right]^{T}\left[\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right]$

则 $$\hat{\boldsymbol{\theta}}_{LS}=\arg\min_{\hat{\boldsymbol{\theta}}}\xi\left(\hat{\boldsymbol{\theta}}\right)$$

即 $$\left.\frac{\partial}{\partial\hat{\boldsymbol{\theta}}}\xi\left(\hat{\boldsymbol{\theta}}\right)\right|_{\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}_{LS}}=-2\mathbf{H}^{T}\left(\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right)=0$$

解得 $$\hat{\boldsymbol{\theta}}_{LS}=\left(\mathbf{H}^{T}\mathbf{H}\right)^{-1}\mathbf{H}^{T}\mathbf{x}$$

（8）多参量加权最小二乘估计

对$\xi\left(\hat{\boldsymbol{\theta}}\right)$加权后做最小二乘估计，获得更好的估计结果。

性能指标

$$\xi_W\left(\hat{\boldsymbol{\theta}}\right)=\left[\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right]^T\mathbf{W}\left[\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right]$$

是$N\times N$ 维的对称正定加权矩阵。

则

$$\hat{\boldsymbol{\theta}}_{LSW}=\arg\min_{\hat{\boldsymbol{\theta}}}\xi_W\left(\hat{\boldsymbol{\theta}}\right)$$

即 $$\left.\frac{\partial}{\partial\hat{\boldsymbol{\theta}}}\xi_W\left(\hat{\boldsymbol{\theta}}\right)\right|_{\hat{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}_{LSW}}=-2\mathbf{H}^T\mathbf{W}\left[\mathbf{x}-\mathbf{H}\hat{\boldsymbol{\theta}}\right]=0$$

解得

$$\hat{\boldsymbol{\theta}}_{LSW}=\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{x}$$

估计误差矩阵为

$$E\left\{\left[\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{LSW}\right]\left[\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{LSW}\right]^T\right\}=\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{R}_\mathbf{n}\mathbf{W}\mathbf{H}\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}$$

其中$\mathbf{R}_\mathbf{n}=E\left\{\mathbf{n}\mathbf{n}^T\right\}$是对称正定矩阵，可分解为

$$\mathbf{R}_\mathbf{n}=\mathbf{D}^T\mathbf{D}$$

利用矩阵施瓦兹不等式

$$\mathbf{B}^T\mathbf{B} \geq (\mathbf{AB})^T (\mathbf{AA}^T)^{-1} (\mathbf{AB})$$

令 $\mathbf{A} = \mathbf{H}^T\mathbf{D}^{-1},\ \mathbf{B} = \mathbf{DWH}(\mathbf{H}^T\mathbf{WH})^{-1}$

可得 $$E\left\{\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LSW}\right]\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LSW}\right]^T\right\} \geq \left(\mathbf{H}^T\mathbf{R}_\mathbf{n}^{-1}\mathbf{H}\right)^{-1}$$

当 $\mathbf{W} = \mathbf{R}_\mathbf{n}^{-1}$ 时上式中等式成立，即估计误差达到最小。

此时 $$\hat{\boldsymbol{\theta}}_{LSW} = \left(\mathbf{H}^T\mathbf{R}_\mathbf{n}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}_\mathbf{n}^{-1}\mathbf{x}$$

例 4. 7

已知线性观测方程为

$$\mathbf{x} = \mathbf{Cs} + \mathbf{n}$$

其中

$\mathbf{n}$ 是N 维高斯噪声列矢量，$E\{\mathbf{n}\} = 0, Cov\{\mathbf{n}, \mathbf{n}\} = \mathbf{V_n}$;

$\mathbf{s}$ 是M维高斯信号列矢量，$E\{\mathbf{s}\} = 0, Cov\{\mathbf{s}, \mathbf{s}\} = \mathbf{V_s}$;

信号与噪声相互独立，$\mathbf{C}$是 $N \times M$ 维观测矩阵

试求 $\hat{\mathbf{s}}_{MAP}, \hat{\mathbf{s}}_{MS}, \hat{\mathbf{s}}_{LMS}$ 。

例 4.8

对某二维矢量 $\boldsymbol{\theta}$ 做了两次观测，观测方程如下

$$\mathbf{x}_1 = \mathbf{H}_1\boldsymbol{\theta} + \mathbf{n}_1$$

$$x_2 = \mathbf{h}_2\boldsymbol{\theta} + n_2$$

其中

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{H}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix},$$

$$x_2 = 4, \qquad \mathbf{h}_2 = \begin{bmatrix} 2 & 2 \end{bmatrix},$$

试求 $\hat{\boldsymbol{\theta}}_{LS}$ 。

4.4 估计量评价的指标

（1）无偏性

对于确定参量 θ，若估计量 $\hat{\theta}$ 满足

$$E\left\{\hat{\theta}\right\}=\theta$$

或对于随机参量 θ，若估计量　满足

$$E\left\{\hat{\theta}\right\}=E\left\{\theta\right\}$$

则称所求的估计量 $\hat{\theta}$ 具有无偏性，是无偏估计，否则就是有偏估计。

确定参量有偏估计的偏差量为 $E\left\{\hat{\boldsymbol{\theta}}\right\}-\boldsymbol{\theta}$

随机参量有偏估计的偏差量为 $E\left\{\hat{\boldsymbol{\theta}}\right\}-E\left\{\boldsymbol{\theta}\right\}$

当观测样本数趋于无穷时，若

$$\lim_{N\to+\infty} E\left\{\hat{\boldsymbol{\theta}}\right\}=\boldsymbol{\theta}$$

或

$$\lim_{N\to+\infty} E\left\{\hat{\boldsymbol{\theta}}\right\}=E\left\{\boldsymbol{\theta}\right\}$$

则称估计量 $\hat{\boldsymbol{\theta}}$ 具有渐进无偏性，是渐近无偏估计。

（2）有效性

- 均方误差衡量了估计量在真值附近的密集程度。
- 如果某个无偏估计量的均方误差是所有估计量均方误差的最小值，称该估计量是有效估计量。
- 均方误差的最小值由克拉美-罗不等式给出。
- 无偏估计量的有效率

$$\eta = \frac{\xi_{\min}\left\{\hat{\boldsymbol{\theta}}\right\}}{\xi\left\{\hat{\boldsymbol{\theta}}_1\right\}} \qquad \xi\left\{\hat{\boldsymbol{\theta}}_1\right\} = E\left\{\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}\right)^T\left(\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}\right)\right\}$$

η 越大估计质量越好，对于有效估计 $\eta = 1$ 。

当观测样本数 N 趋于无穷时，若

$$\lim_{N\to+\infty}\eta=1$$

则称该估计量为渐近有效估计量。

对某一估计量，若

$$\lim_{N\to+\infty}\eta=\eta_0\neq 1$$

则称 η_0 为渐进有效率。

（3）一致性

若观测样本数N 趋于无穷时，估计量越来越接近其真值，则此时的估计量称为一致估计量。

一致估计的两种度量方法：

A 当 $N \to +\infty$时，估计量 $\hat{\boldsymbol{\theta}}$ 在概率意义上收敛于$\boldsymbol{\theta}$

$$\lim_{N \to +\infty} P\left(\left\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right\| < \varepsilon\right) = 1$$

B 当 $N \to +\infty$时，估计量 $\hat{\boldsymbol{\theta}}$ 在均方意义上趋近于$\boldsymbol{\theta}$

$$\lim_{N \to +\infty} E\left\{\left\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right\|^2\right\} = 0$$

（4）充分性

对待定参量 $\boldsymbol{\theta}$ 及其估计量$\hat{\boldsymbol{\theta}}(\mathbf{x})$, 如果似然函数满足

$$f(\mathbf{x}|\boldsymbol{\theta}) = g\left[\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}\right]h(\mathbf{x})$$

其中，$h(\mathbf{x}) \geq 0$且与$\boldsymbol{\theta}$无关，$g\left[\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}\right]$ 是 $\hat{\boldsymbol{\theta}},\boldsymbol{\theta}$ 的函数，则称 $\hat{\boldsymbol{\theta}}(\mathbf{x})$为充分估计量。

表明$g(\cdot)$中的 $\hat{\boldsymbol{\theta}}$包含了观测样本中关于待定参量的全部信息。即：从充分估计量中可以获取待定参量的全部信息，而其它估计量中关于待定参量的信息总是小于充分估计量的。

4.5 克拉美-罗不等式

4.5.1 确定单参量估计的克拉美-罗不等式

4.5.2 随机单参量估计的克拉美-罗不等式

4.5.3 确定矢量估计的克拉美-罗不等式

4.5.4 随机矢量估计的克拉美-罗不等式

4.5.1 确定单参量估计的克拉美-罗不等式

对于确定单参量θ 的无偏估计为$\hat{\theta}$ ，有

$$E\left\{\left(\hat{\theta}-\theta\right)^2\right\}\geq\frac{1}{E\left\{\left[\frac{\partial}{\partial\theta}\ln f\left(\mathbf{x}\mid\theta\right)\right]^2\right\}}=-\frac{1}{E\left\{\frac{\partial^2}{\partial\theta^2}\ln f\left(\mathbf{x}\mid\theta\right)\right\}}$$

当 $\frac{\partial}{\partial\theta}\ln f\left(\mathbf{x}\mid\theta\right)=K\left(\theta\right)\left(\hat{\theta}-\theta\right)$ 时等号成立，此时的估计量 $\hat{\theta}$ 为有效估计量。

确定单参量的有效估计是它的最大似然估计。

4.5.2 随机单参量估计的克拉美-罗不等式

对随机单参量 θ 的无偏估计 $\hat{\theta}$，定义 θ 给定时估计误差的条件数学期望为

$$g(\theta)=\int_{(\mathbf{x})}\left(\hat{\theta}-\theta\right)f\left(\mathbf{x}\,|\,\theta\right)d\mathbf{x}$$

若 $f\left(\mathbf{x}\,|\,\theta\right)$ 对 θ 的一二阶导数都存在且绝对可积，并满足

$$\lim_{\theta\to+\infty} f\left(\theta\right)g\left(\theta\right)=0,\quad \lim_{\theta\to-\infty} f\left(\theta\right)g\left(\theta\right)=0$$

有

$$E\left\{\left(\hat{\theta}-\theta\right)^2\right\} \geq \frac{1}{E\left\{\left[\frac{\partial}{\partial\theta}\ln f\left(\mathbf{x},\theta\right)\right]^2\right\}} = -\frac{1}{E\left\{\frac{\partial^2}{\partial\theta^2}\ln f\left(\mathbf{x},\theta\right)\right\}}$$

当 $\frac{\partial}{\partial\theta}\ln f\left(\mathbf{x},\theta\right) = K\left(\hat{\theta}-\theta\right)$ 时，等号成立。

此时 $\hat{\theta}$ 为有效估计量。

随机单参量的有效估计是它的最大后验概率估计。

4.5.3 确定矢量估计的克拉美-罗不等式

对确定矢量 $\boldsymbol{\theta}$ 的无偏估计为 $\hat{\boldsymbol{\theta}}$，有

$$E\left\{\left(\hat{\theta}_i-\theta_i\right)^2\right\}\geq\left(\mathbf{J}^{-1}\right)_{ii}\quad i=1,2,\cdots,M$$

其中

$$\mathbf{J}=E\left\{\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x}\,|\,\boldsymbol{\theta}\right)\right]\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x}\,|\,\boldsymbol{\theta}\right)\right]^T\right\}=-E\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x}\,|\,\boldsymbol{\theta}\right)\right]^T\right\}$$

当 $\hat{\theta}_i-\theta_i=\sum_{j=1}^{M}K_{ij}\frac{\partial}{\partial\theta_j}\ln f\left(\mathbf{x}\,|\,\boldsymbol{\theta}\right)\quad i=1,2,\cdots,M$时，等号成立。

4.5.4 随机矢量估计的克拉美-罗不等式

对随机矢量 $\boldsymbol{\theta}$ 的无偏估计为$\hat{\boldsymbol{\theta}}$ ，有

$$E\left\{\left(\hat{\theta}_i-\theta_i\right)^2\right\}\geq\left(\mathbf{J}_T^{-1}\right)_{ii}\quad i=1,2,\cdots,M$$

其中

$$\mathbf{J}_T=E\left\{\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x},\boldsymbol{\theta}\right)\right]\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x},\boldsymbol{\theta}\right)\right]^T\right\}=-E\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}}\ln f\left(\mathbf{x},\boldsymbol{\theta}\right)\right]^T\right\}$$

当 $\frac{\partial}{\partial\theta_i}\ln f\left(\mathbf{x},\boldsymbol{\theta}\right)=\sum_{j=1}^{M}\left(J_T\right)_{ij}\left(\hat{\theta}_j-\theta_j\right)\ i=1,2,\ldots,M$ 时，等号成立。

4.6 应用

例 4.9

探测器连续跟踪观测某个目标时，线性观测方程为

$$x_k = Aq^k + n_k \qquad k = 1, 2, \cdots, N$$

式中，目标信号的幅度A是待估计的参量，q是信号随时间的衰减因子，且 $0 < q \le 1$ ，n_k 是独立同分布的观测噪声，其均值为零、方差为 σ_n^2 。求A的估计值。

例 4.10

以雷达系统为例，假设观测信号为

$$x_k = A\cos\left(\omega_0 k + \theta\right) + n_k \qquad k = 1, 2, \cdots N$$

式中，振幅 A 和相位 $\theta \in [-\pi, \pi]$ 为待估计参量，频率 ω_0 已知，n_k 为观测噪声。

利用非线性最小二乘估计来估计 A 和 θ 。

例 4.11

当待估计的未知参量为信号的幅度时，接收信号波形可表示成

$$x(t)=s(t,A)+n(t)=As(t)+n(t) \qquad 0\le t\le T$$

式中，$s(t)$ 是已知的信号，$n(t)$ 是均值为零、功率谱密度为 $N_0/2$ 的高斯白噪声。

利用最大似然估计来估计信号幅度 A。

例 4.12

当待估计的未知参量为信号的相位时，接收信号波形可表示成

$$x(t)=s(t,\theta)+n(t)=A\sin(\omega_0 t+\theta)+n(t)\ \ 0\le t\le T$$

式中，幅度 A 和频率 ω_0 是已知的，$n(t)$ 是均值为零、功率谱密度为 $N_0/2$ 的高斯白噪声。

利用最大似然估计来估计信号的相位 θ。

例 4.13

当待估计的未知参量为信号的时延时，接收信号波形可表示成

$$x(t)=s(t,\tau)+n(t)=s(t-\tau)+n(t) \quad 0\le t\le T$$

式中，$s(t)$ 是已知的信号，$n(t)$ 是均值为零、功率谱密度为 $N_0/2$ 的高斯白噪声。

利用最大似然估计来估计信号的时延 τ 。

例 4.14

考虑高斯色噪声中信号未知参量的估计。接收信号波形为

$$x(t)=s(t,\boldsymbol{\theta})+n(t) \qquad 0\le t\le T$$

式中，$s(t,\boldsymbol{\theta})$是已知的信号，$\boldsymbol{\theta}$为待估计的参量，$n(t)$是均值为零、自相关函数为$R_n(\tau)$的高斯色噪声。

利用最大似然估计来估计信号的参量 $\boldsymbol{\theta}$ 。