信号统计分析

第五章 信号估计理论

本章内容

5.1 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

5.2 估计准则

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

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

5.2.1 最大后验概率估计

使后验概率密度最大的一种估计，即

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

其中 $f(\theta\,|\,\mathbf{x})$ 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

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

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 5.1

已知如下观测样本

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

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

求 $\hat{s}_{MAP}$ 。

5.2.2 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

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

- 适用于确定参量估计和先验分布未知的随机参量估计。

例 5.2

已知如下观测样本

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

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

求 $\hat{s}_{ML}$ 。

5.2.3 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

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

估计的均方误差

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

则

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

利用

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

可得

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

例 5.3

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 $\hat{s}_{MS}$ 。

5.2.4 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 a 和 $\mathbf{b} = \left[b_1, b_2, \cdots, b_N\right]^T$ 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

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

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

求解得

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

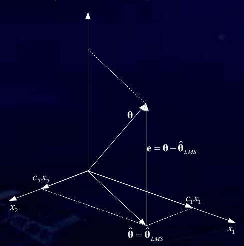
$$\mathbf{b}_L^T = Cov\{\theta,\mathbf{x}\}Cov^{-1}\{\mathbf{x},\mathbf{x}\}$$

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 5.4

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 $\hat{s}_{LMS}$ 。

5.2.5 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

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

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

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

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 5.5

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 $\hat{s}_{ABS}$ 。

5.2.6 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

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

即

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

（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

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

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

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

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

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

Δ>0是很小的常数

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

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

解得

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

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 5.6

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 $\hat{s}_{LS}$ 。

5.3 多参量的常用估计准则

多参量矢量

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

估计量

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

估计误差矢量

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

（1）多参量最大后验概率估计

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

即

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

其中第 i 个方程为

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

（2）多参量最大似然估计

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

即

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

其中第 i 个方程为

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

（3）多参量最小均方误差估计

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

得

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

其中第 i 个方程估计量为

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

（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

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

则

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

解得

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

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

（5）多参量最小平均绝对误差估计

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

解得

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

（6）多参量贝叶斯估计

平均风险为

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

则

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

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

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

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

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

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

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

性能指标

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

$\mathbf{W}$ 是 $N \times N$ 维的对称正定加权矩阵。

则

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

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

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

$$\mathbf{B}^T\mathbf{B} \geq (\mathbf{A}\mathbf{B})^T \left(\mathbf{A}\mathbf{A}^T\right)^{-1} (\mathbf{A}\mathbf{B})$$

令　$\mathbf{A} = \mathbf{H}^T\mathbf{D}^{-1},\ \mathbf{B} = \mathbf{D}\mathbf{W}\mathbf{H}\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}$

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 5.7

已知线性观测方程为

$$\mathbf{x} = \mathbf{C}\mathbf{s} + \mathbf{n}$$

其中

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

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

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 5.8

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

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{H}_1 = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix},$$
$$x_2 = 4, \quad \mathbf{h}_2 = \begin{bmatrix} 2 & 2 \end{bmatrix},$$

试求 $\hat{\boldsymbol{\theta}}_{LS}$ 。

5.4 估计量评价的指标

（1）无偏性

对于确定参量 $\boldsymbol{\theta}$ ，若估计量 $\hat{\boldsymbol{\theta}}$ 满足

$$E\{\hat{\boldsymbol{\theta}}\} = \boldsymbol{\theta}$$

或对于随机参量 $\boldsymbol{\theta}$ ，若估计量 满足

$$E\{\hat{\boldsymbol{\theta}}\} = E\{\boldsymbol{\theta}\}$$

则称所求的估计量 $\hat{\boldsymbol{\theta}}$ 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 $E\{\hat{\boldsymbol{\theta}}\} - \boldsymbol{\theta}$

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

当观测样本数趋于无穷时，若

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

或

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

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

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

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。

5.5 克拉美-罗不等式

5.5.1 确定单参量估计的克拉美-罗不等式

5.5.2 确定矢量估计的克拉美-罗不等式

5.5.3 随机单参量估计的克拉美-罗不等式

5.5.4 随机矢量估计的克拉美-罗不等式

5.5.1 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

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

当且仅当 $\frac{\partial}{\partial\theta}\ln f(\mathbf{x}|\theta)=K(\theta)\left(\hat{\theta}-\theta\right)$ 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。

5.5.2 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

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

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

5.5.3 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 $\hat{\theta}$，为定义 θ给定时估计误差的条件数学期望为

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

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

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

有

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

当且仅当 $\frac{\partial}{\partial\theta}\ln f(\mathbf{x},\theta)=K\left(\hat{\theta}-\theta\right)$ 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。

5.5.4 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

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

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

例 5.9

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

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

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 5.10

当待估计的未知参量为信号的初相位时，接收信号波形可表示成

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

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初相位 θ 。

例 5.11

以雷达系统为例，假设观测信号为

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

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。

本章内容小结

1 单参量估计准则
2 多参量估计准则
3 估计量评价的指标
4 克拉美-罗不等式
5 应用举例