07　估计理论

7.1估计的基本概念

估计理论：

从含有噪声的数据中估计信号的某些特征参量的理论和方法。

估计问题通常是以下三种情况：

- ■ **根据观测样本直接对观测样本的各类统计特性作出估计；**
- ■ **根据观测样本，对观测样本中的信号中的未知的待定参量作出估计，称为信号的参量估计问题，又分为点估计和区间估计；**
- ■ **根据观测样本对随时间变化的信号作出波形估计，又称为过程估计。**

7. 估计理论

估计理论：

从含有噪声的数据中估计信号的某些特征参量的理论和方法。

- 估计方法
  - 参数估计
    - 贝叶斯估计
      - 最小均方估计
      - 最大后验概率估计
      - 条件中位数估计
    - 最大似然估计
    - 线性最小均方估计
    - 最小二乘估计
  - 波形估计

7. 估计理论

7.1 估计的基本概念

7.2 贝叶斯估计： 已知代价函数及先验概率，使估计付出的平均代价最小

7.3 最大似然估计： 使似然函数最大

7.4 估计量的性能

7.5 线性最小均方估计： 已知估计量的一、二阶矩,使均方误差最小的线性估计

7.6 最小二乘估计： 观测与估计偏差的平方和最小

7.7 波形估计

7.1估计的基本概念

例7.1：有被测电压θ，θ的取值范围为$(-\theta_0,\theta_0)$，由于测量设备的不完善，测量总会有些误差，测量误差可归结为噪声，因此，实际得到的测量值为

$$z = \theta + v$$

其中v一般服从零均值的正太分布，方差为σ_v^2。

问题：如何根据测量值z来估计θ的值？

7.1估计的基本概念

例7.1解：

θ为随机变量： $f(\theta|z)|_{\theta=\hat{\theta}_{map}}$=max 最大后验概率估计

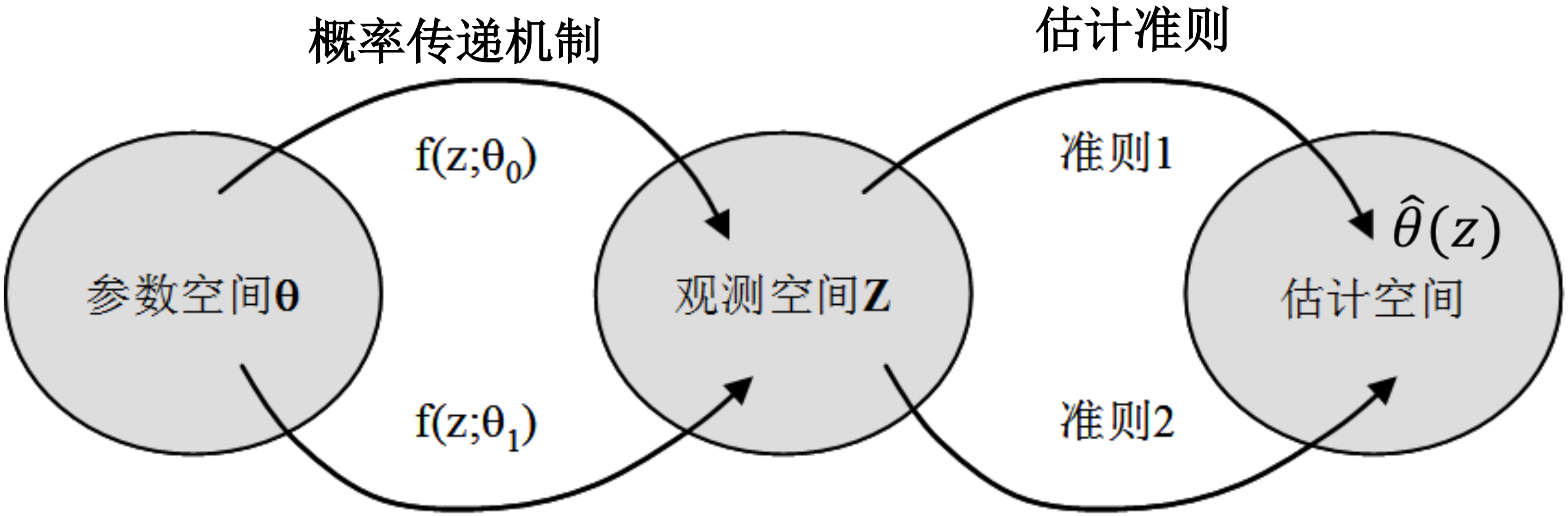
$$\hat{\theta}_{map} = \frac{\sigma_\theta^2 z}{\sigma_v^+ + \sigma_\theta^2} \qquad \theta \sim N(0，\sigma_\theta^2)$$

θ为未知常数： $f(z;\theta)|_{\theta=\hat{\theta}_{ml}}$=max 最大似然估计

$$\hat{\theta}_{ml} = z$$

7.1估计的基本概念

估计问题基本要素



7.2贝叶斯估计

1、贝叶斯估计

在已知代价函数及先验概率基础上，使估计付出的平均代价最小。

设观测值为z，待估参量为θ。

估计误差： $\tilde{\theta} = \theta - \hat{\theta}(z)$

设代价函数： $C(\tilde{\theta})$

贝叶斯估计准则： $\hat{\theta}(z) \Leftarrow \min_{\hat{\theta}} E[C(\tilde{\theta})]$

7.2贝叶斯估计

统计平均代价：

$$E[C(\tilde{\theta})]=E[C(\theta,\hat{\theta}(z))]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}C(\theta,\hat{\theta}(z))f(\theta,z)d\theta dz$$

$$=\int_{-\infty}^{\infty}\underbrace{\left[\int_{-\infty}^{\infty}C(\theta,\hat{\theta}(z))f(\theta\mid z)d\theta\right]}_{\text{条件平均代价}}f(z)dz$$

$$=\int_{-\infty}^{\infty}\bar{C}(\theta\mid z)f(z)dz$$

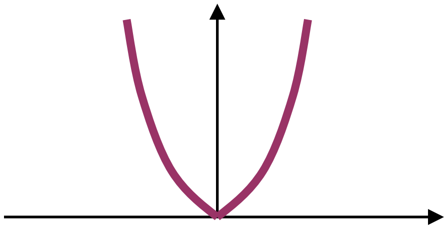
等价于使下式最小：

$$\int_{-\infty}^{\infty}C(\theta,\hat{\theta}(z))f(\theta\mid z)d\theta=\text{最小}$$

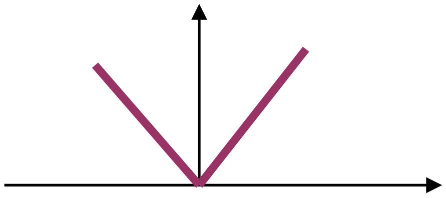
7.2贝叶斯估计

2、典型代价函数及贝叶斯估计

平方代价：$C(\theta,\hat{\theta})=(\theta-\hat{\theta})^2$

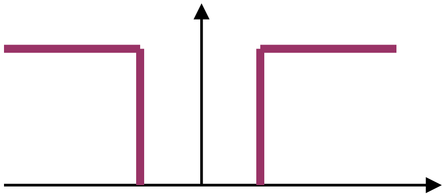


绝对值代价：$C(\theta,\hat{\theta})=|\theta-\hat{\theta}|$



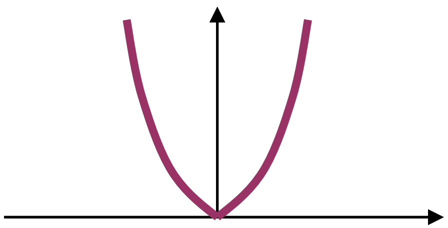
均匀代价：

$$C(\theta,\ \hat{\theta})=\begin{cases}1, & |\theta-\hat{\theta}|\geq\dfrac{\Delta}{2}\\ 0, & |\theta-\hat{\theta}|<\dfrac{\Delta}{2}\end{cases}$$



7.2贝叶斯估计

平方代价： $C(\theta,\hat{\theta})=(\theta-\hat{\theta})^2$



☞ **最小均方估计(Minimal Square)**

$$\bar{C}(\theta\,|\,z)=\int_{-\infty}^{\infty}(\theta-\hat{\theta})^2 f(\theta\,|\,z)d\theta = \text{最小}$$

对 $\hat{\theta}$ 求导数，并使其等于零：

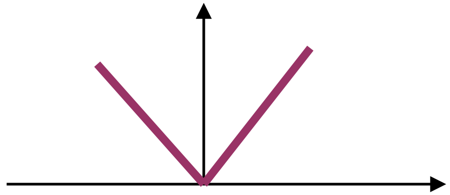
$$\frac{d\bar{C}(\theta\,|\,z)}{d\hat{\theta}}=-2\int_{-\infty}^{\infty}\theta f(\theta\,|\,z)d\theta+2\hat{\theta}\int_{-\infty}^{\infty}f(\theta\,|\,z)d\theta$$

得： $\hat{\theta}=\int_{-\infty}^{\infty}\theta f(\theta\,|\,z)d\theta$

即 $\hat{\theta}=E[\theta\,|\,z]$，也称为条件均值估计。

7.2贝叶斯估计

绝对值代价： $C(\theta,\hat{\theta}) = |\theta-\hat{\theta}|$



☞条件中位数估计(Median)

$$\bar{C}(\theta\,|\,z) = \int_{-\infty}^{\infty} |\theta-\hat{\theta}|\, f(\theta\,|\,z)d\theta$$

$$= \int_{-\infty}^{\hat{\theta}} (\hat{\theta}-\theta) f(\theta\,|\,z)d\theta + \int_{\hat{\theta}}^{\infty} (\theta-\hat{\theta}) f(\theta\,|\,z)d\theta$$

对 $\hat{\theta}$ 求导数，并使其等于零，得：

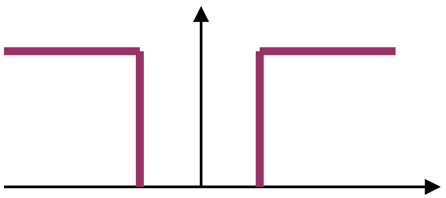
$$\int_{-\infty}^{\hat{\theta}_{abs}} f(\theta\,|\,z)d\theta = \int_{\hat{\theta}_{abs}}^{\infty} f(\theta\,|\,z)d\theta$$

可见，估计为条件概率密度 $f(\theta\,|\,z)$ 的中位数。

7.2贝叶斯估计

均匀代价：

$$C(\theta,\ \hat{\theta}) = \begin{cases} 1, & |\theta - \hat{\theta}| \geq \dfrac{\Delta}{2} \\ 0, & |\theta - \hat{\theta}| < \dfrac{\Delta}{2} \end{cases}$$



☞**最大后验概率估计（maximal posterior probability)**

$$\bar{C}(\theta \mid z) = 1 - \int_{\hat{\theta}_{map}-\frac{\Delta}{2}}^{\hat{\theta}_{map}+\frac{\Delta}{2}} f(\theta \mid z) d\theta$$

应当选择 $\hat{\theta}$ ，使它处在后验概率 $f(\theta \mid z)$的最大处。

最大后验概率方程：

$$\left.\frac{\partial f(\theta \mid z)}{\partial \theta}\right|_{\theta=\hat{\theta}_{map}} = 0 \text{ 或 } \left.\frac{\partial \ln f(\theta \mid z)}{\partial \theta}\right|_{\theta=\hat{\theta}_{map}} = 0$$

7.2贝叶斯估计

由关系式：

$$f(\theta \mid z)=\frac{f(z \mid \theta)f(\theta)}{f(z)}$$

两边取对数并对θ求导，得最大后验概率方程的另一形式：

$$\left[\frac{\partial \ln f(z \mid \theta)}{\partial \theta}+\frac{\partial \ln f(\theta)}{\partial \theta}\right]\Bigg|_{\theta=\hat{\theta}_{map}}=0$$

7.2贝叶斯估计

例1设观测为$z = A + v$，其中被估计量A在[-A_0,A_0]上均匀分布，测量噪声v~N(0, σ_v^2)，求A的最大后验概率估计和最小均方估计。

$$f(z|A) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left\{-\frac{(z-A)^2}{2\sigma_v^2}\right\}$$

$$f(A|z) = \frac{f(z|A)f(A)}{f(z)}$$

$$\hat{A}_{ms} = \int_{-\infty}^{\infty} Af(A|z)dA = \int_{-\infty}^{\infty} A\frac{f(z|A)f(A)}{f(z)}dA = \frac{\int_{-\infty}^{\infty} Af(z|A)f(A)dA}{\int_{-\infty}^{\infty} f(z|A)f(A)dA}$$

7.2贝叶斯估计

例2 高斯白噪声中的直流电平估计-高斯先验分布。设有N次独立观测$z_i=A+v_i$, i=1, 2, …. N，其中v~N(0, σ^2)，A~ $N(\mu_A, \sigma_A^2)$，求A的估计。

$$f(A|\mathbf{z}) = \frac{f(\mathbf{z}|A)f(A)}{\int_{-\infty}^{\infty} f(\mathbf{z}|A)f(A)dA}$$

$$= \frac{1}{\sqrt{2\pi\sigma_{A|z}^2}} \exp\left[-\frac{1}{2\sigma_{A|z}^2}(A-\mu_{A|z})^2\right]$$

7.3最大似然估计

1、最大似然估计 (Maximum Likelihood Estimate)

设观测矢量z=[z1,z2,…,zN]T,把$f(z;\theta)$称为似然函数，使似然函数最大所对应的参数θ作为对θ的估计，称为最大似然估计，记为$\hat{\boldsymbol{\theta}}_{\mathrm{m}l}$，即

$$\boldsymbol{f}(\boldsymbol{z};\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathbf{ml}}}=\boldsymbol{max} \text{ 或 } \ln\boldsymbol{f}(\boldsymbol{z};\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathbf{ml}}}=\boldsymbol{max}\text{。}$$

很显然，如果函数可导函数，那么最大似然估计的必要条件是

$$\frac{\partial\boldsymbol{f}(\boldsymbol{z};\boldsymbol{\theta})}{\partial x\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathbf{ml}}}=\mathbf{0} \text{ 或 } \frac{\partial\ln\boldsymbol{f}(\boldsymbol{z};\boldsymbol{\theta})}{\partial x\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathbf{ml}}}=\mathbf{0}$$

最大似然方程

7.3最大似然估计

例1、高斯白噪声中的直流电平估计-未知参数。设有N次独立观测$z_i=A+v_i$，$i=1,2,\dots.N$，其中$v_i\sim N(0,\sigma^2)$，A为未知参数，σ^2已知，求A的最大似然估计。

$$f(\mathbf{z}/A)=\left(\frac{1}{2\pi\sigma^2}\right)^{N/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(z_i-A)^2\right]$$

$$\hat{A}_{ml}=\bar{z}=\frac{1}{N}\sum_{i=1}^{N}z_i$$

7.3最大似然估计

例2、设有N次独立观测$z_i=v_i$，$i=1,2,\ldots.N$，其中$v_i \sim N(0,\sigma^2)$，求σ^2 的最大似然估计。

$$f(\mathbf{z}/\sigma^2)=\left(\frac{1}{2\pi\sigma^2}\right)^{N/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}z_i^2\right\}$$

$$\hat{\sigma}_{ml}^2=\frac{1}{N}\sum_{i=1}^{N}z_i^2$$

7.3最大似然估计

例3、高斯白噪声中的直流电平估计-未知参数与未知方差。设有N次独立观测$z_i=A+v_i$，$i=1,2,\ldots.N$，其中$v\sim N(0,\sigma^2)$，σ^2、A均为未知参数，求A和σ^2的最大似然估计。

$$f(\mathbf{z}/\boldsymbol{\theta})=\left(\frac{1}{2\pi\sigma^2}\right)^{N/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(z_i-A)^2\right] \qquad \boldsymbol{\theta}=[A\ \sigma^2]^T$$

$$\hat{\boldsymbol{\theta}}_{ml}=\begin{bmatrix}\hat{A}_{ml}\\ \hat{\sigma}^2\end{bmatrix}=\begin{bmatrix}\overline{z}\\ \dfrac{1}{N}\sum_{i=1}^{N}(z_i-\overline{z})^2\end{bmatrix}$$

7.4估计量的性能

1、估计量的性能标准

- 无偏性

如果估计量的均值等于非随机参量或等于随机参量的均值，则称估计量具有无偏性。即满足：

对于确定量，有： $E[\hat{\theta}] = \theta$

对于随机量，有： $E[\hat{\theta}] = E[\theta]$

7.4估计量的性能

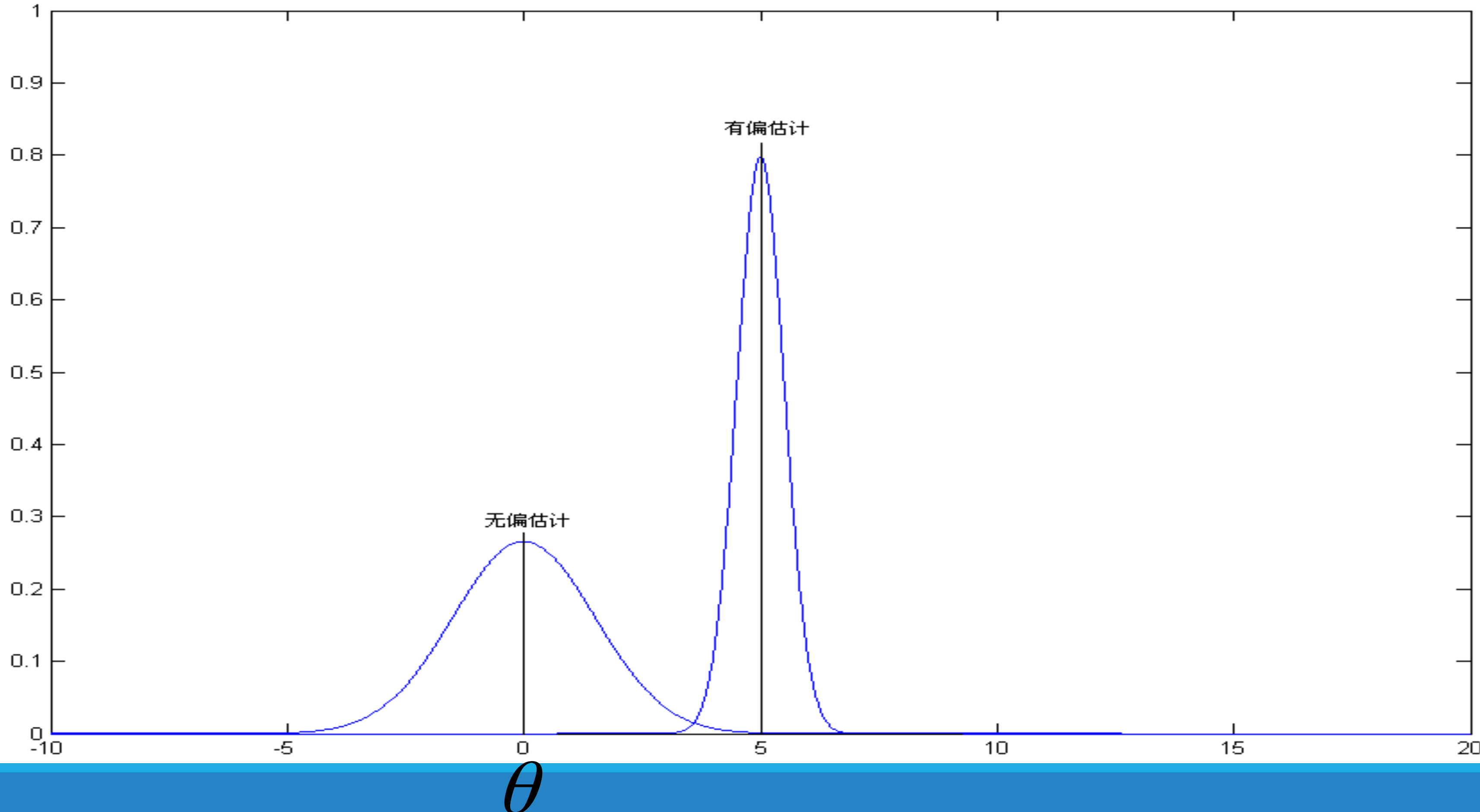
- **有效性**

对于无偏估计，如果估计的方差越小，表明估计量的取值越集中于真值附近，估计的性能越好。

$$Var(\hat{\theta}) = E\{[\hat{\theta} - E(\hat{\theta})]^2\}$$

7.4估计量的性能

对于有偏估计，尽管估计的方差很小，但估计的误差可能仍然很大。



7.4估计量的性能

- **有效性**

对于无偏估计，如果估计的方差越小，表明估计量的取值越集中于真值附近，估计的性能越好。

$$Var(\hat{\theta}) = E\{[\hat{\theta} - E(\hat{\theta})]^2\}$$

用估计的方差还不能准确地描述估计的性能，所以我们可以用均方误差作为评价估计量性能的一个指标。

$$Mse(\hat{\theta}) = E\{[\theta - \hat{\theta}]^2\}$$

7.4估计量的性能

- **一致性**

即对于任意小数ε，若有： $\lim_{N\to\infty} P\left(\left|\hat{\theta}(z_1, z_2, \cdots, z_N) - \theta\right| > \varepsilon\right) = 0$

则估计量 $\hat{\theta}$ 为一致估计量。

若满足 $\lim_{N\to\infty} E\left\{\left[\theta - \hat{\theta}(\mathbf{z}_N)\right]^2\right\} = 0$

则称 $\hat{\theta}$ 为均方一致估计量。

7.4估计量的性能

例1、高斯白噪声中的直流电平估计-未知参数。设有N次独立观测$z_i=A+v_i$，$i=1,2,\ldots.N$，其中$v_i \sim N(0,\sigma^2)$，σ^2已知。

$$\hat{A}_{map} = \begin{cases} -A_0 & z < -A_0 \\ z & -A_0 \le z \le A_0 \\ A_0 & z > A_0 \end{cases} \qquad \hat{A}_{ml} = \bar{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$$

$$E(\hat{A}_{map}) \ne A \qquad E(\hat{A}_{ml}) = E\left(\frac{1}{N}\sum_{i=1}^{N} z_i\right) = A$$

$$Mse(\hat{A}_{ml}) > Mse(\hat{A}_{map})$$

7. 4估计量的性能

2、克拉美－罗限(Cramer-Rao Low bound)

无偏估计量的估计方差的最小值

- 非随机参量

任何无偏估计量的方差满足

$$Var(\hat{\theta}) = E\{[\theta - \hat{\theta}]^2\} \geq J^{-1}$$

克拉美–罗限

$$J = E\left\{\left[\frac{\partial \ln f(z|\theta)}{\partial \theta}\right]^2\right\} = -E\left\{\frac{\partial^2 \ln f(z|\theta)}{\partial \theta^2}\right\}$$

等号成立的条件： $\dfrac{\partial \ln f(\mathbf{z}|\theta)}{\partial \theta} = (\hat{\theta} - \theta)k(\theta)$

7.4估计量的性能

2、克拉美－罗限(Cramer-Rao bound)

如果一个无偏估计，它的方差达到CRLB，那么，这个估计必定是最大似然估计。这时最大似然估计是最好的。但如果不存在达到CRLB的估计，最大似然估计就不一定是最好的估计。

$$\left.\frac{\partial \ln f(\mathbf{z}\,|\,\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_{ml}} = 0$$

$$\left.\frac{\partial \ln f(\mathbf{z}\,|\,\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}_{ml}} = (\hat{\theta}-\hat{\theta}_{ml})k(\hat{\theta}_{ml}) = 0$$

7.4估计量的性能

例2、高斯白噪声中的DC电平。DC电平的最大似然估计的方差是否达到CRLB？它的估计方差是多少？

$$\hat{A}_{ml} = \bar{z} = \frac{1}{N}\sum_{i=1}^{N} z_i$$

$$E[\hat{A}_{ml}] = E\left[\frac{1}{N}\sum_{i=1}^{N} z_i\right] = \frac{1}{N}\sum_{i=1}^{N} E[z_i] = A$$

$$\frac{\partial \ln f(\mathbf{z}/A)}{\partial A} = \frac{N}{\sigma^2}\left(\frac{1}{N}\sum_{i=1}^{N} z_i - A\right) = \frac{N}{\sigma^2}(\hat{A}_{ml} - A)$$

$$Var(\hat{A}_{ml}) = -\frac{1}{E\left\{\dfrac{\partial^2 \ln f(z/A)}{\partial A^2}\right\}} = \frac{\sigma^2}{N}$$

7.4估计量的性能

2、克拉美－罗限(Cramer-Rao Low bound)

无偏估计量的估计方差的最小值

- **随机参量**

任何无偏估计量的均方误差满足

克拉美−罗限

$$Mse(\hat{\theta}) = E\{[\theta - \hat{\theta}]^2\} \geq J^{-1}$$

$$J = E\left\{\left[\frac{\partial \ln f(z,\theta)}{\partial \theta}\right]^2\right\} = -E\left\{\frac{\partial^2 \ln f(z,\theta)}{\partial \theta^2}\right\}$$

等号成立的条件：

$$\frac{\partial \ln f(\mathbf{z},\theta)}{\partial \theta} = (\hat{\theta} - \theta)k$$

7.4估计量的性能

2、克拉美－罗限(Cramer-Rao Low bound)

无偏估计量的估计方差的最小值

- **随机参量**

如果有某个无偏估计达到CRLB,那么该估计必定是最大后验概率估计.而最小均方估计的均方误差也是最小的,所以这时最小均方估计与最大后验概率估计等价.

$$\frac{\partial \ln f(\mathbf{z},\theta)}{\partial \theta} = (\hat{\theta}-\theta)k$$

7.4估计量的性能

例2 高斯白噪声中的直流电平估计-高斯先验分布。设有N次独立观测$z_i=A+v_i$, i=1, 2, …. N，其中v~N(0, σ^2)，A~ $N(\mu_A,\sigma_A^2)$，求A的估计的CRLB。

$$E(\hat{A}_{map}) = E(A)$$

$$f(\mathbf{z},A) = f(\mathbf{z}\,|\,A)f(A)$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(z_i-A)^2\right]\frac{1}{\sqrt{2\pi\sigma_A^2}}\exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]$$

$$\frac{\partial \ln f(\mathbf{z},A)}{\partial A} = \left(\frac{N}{\sigma^2}+\frac{1}{\sigma_A^2}\right)\left(\hat{A}_{map}-A\right) \qquad \frac{\partial^2 \ln f(\mathbf{z},A)}{\partial A^2} = -\frac{N}{\sigma^2}-\frac{1}{\sigma_A^2}$$

$$Mse(\hat{\theta}) = E[(\hat{\theta}-\theta)^2] = \left(\frac{N}{\sigma^2}+\frac{1}{\sigma_A^2}\right)^{-1} = \frac{\sigma_A^2\sigma^2}{N\sigma_A^2+\sigma^2}$$

7.5线性最小均方估计

1、线性最小均方估计(linear minimum mean square error estimation)

前提：不知道 $f(\theta)$，知道 θ 的一、二阶矩特性

准则：使均方误差最小的线性估计

实现：

$$\hat{\theta}_{lms} = \sum_{i=1}^{N} a_i z_i + b$$

选择适当的系数a_i及b，使估计均方误差最小。

$$Mse(\hat{\theta}) = E[(\theta - \hat{\theta})^2] = E\left[\left(\theta - \sum_{i=1}^{N} a_i z_i - b\right)^2\right] = \min$$

7.5线性最小均方估计

$$b = E[\theta] - \sum_{i=1}^{N} a_i E[z_i]$$

$$E\left[\left(\theta - \sum_{i=1}^{N} a_i z_i - b\right) z_j\right] = 0$$

$$E\left[\tilde{\theta} z_j\right] = E[(\theta - \hat{\theta}_{lms}) z_j] = 0$$

正交条件

正交条件是信号最佳线性滤波和估计算法的基础，在随机信号处理中占有十分重要的地位。

7.5线性最小均方估计

性能分析：

☞线性最小均方估计为无偏估计，即有：

$$E\left[\hat{\theta}_{lms}\right] = E[\theta]$$

☞线性最小均方估计的均方误差等于误差与被估计量乘积的统计均值，即：

$$E\left[\tilde{\theta}^2\right] = E\left[\tilde{\theta}\theta\right]$$

其中： $\tilde{\theta} = \theta - \hat{\theta}_{lms}$

7.5线性最小均方估计

例1、设观测模型为$z_i=s+v_i$，$i=1,2,..$，其中随机参量s以等概率取{-2,-1,0,1,2}诸值，噪声干扰v_i以等概率取{-1,0,1}诸值，且$E[sv_i]=0$，$E[v_i v_j]=\sigma_v^2\delta_{ij}$，试根据一次、二次、三次观测数据求参量s的线性最小均方估计。

7.6最小二乘估计

1、最小二乘估计（Least square estimation）

前提：适用于线性观测模型；

不规定估计的概率或统计描述；

需要关于被估计量的观测信号模型；

准则：使观测与估计偏差的平方和最小。

假定观测模型为线性，即观测数据z_k与参量θ_1，θ_2，…，θ_M之间服从：

$$z_k = h_{k1}\theta_1 + h_{k2}\theta_2 + \cdots + h_{kM}\theta_M + n_k \quad k = 1,2,\cdots,N$$

其中h_{k1}，h_{k2}，…，h_{kM}为已知常系数。

7.6最小二乘估计

将观测方程用矢量及矩阵表示：

$$\mathbf{z} = \mathbf{H}\boldsymbol{\theta} + \mathbf{n}$$

最小二乘估计是使观测与估计偏差的平方和最小，即：

$$J\left(\hat{\boldsymbol{\theta}}\right) = \left(\mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}\right)^T \left(\mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}\right) = \min$$

$$J_W(\hat{\boldsymbol{\theta}}) = [\mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}]^T \mathbf{W}[\mathbf{z} - \mathbf{H}\hat{\boldsymbol{\theta}}] = \min$$

最小二乘估计为： $\hat{\boldsymbol{\theta}}_{ls} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1} \mathbf{H}^T\mathbf{z}$

加权最小二乘估计为： $\hat{\boldsymbol{\theta}}_{lsw} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z}$

7.6最小二乘估计

性能分析：

☞对于线性观测模型，最小二乘估计是线性估计，对测量噪声的统计特性无任何假设，应用十分广泛；

☞若噪声均值为零，最小二乘估计为无偏估计，即有：

$$E\left[\hat{\boldsymbol{\theta}}_{ls}\right] = \theta$$

$$E[\hat{\boldsymbol{\theta}}_{lsw}] = \boldsymbol{\theta}$$

7.6最小二乘估计

性能分析：

☞最小二乘估计的均方误差为： $V_n = E\left[\mathbf{n}\mathbf{n}^T\right]$

$$E\left[\tilde{\boldsymbol{\theta}}_{ls}^2\right] = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{V}_n\mathbf{H}\left(\mathbf{H}^T\mathbf{H}\right)^{-1}$$

$$E\left[\tilde{\boldsymbol{\theta}}_{lsw}^2\right] = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{R}\mathbf{W}\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}$$

☞对于加权最小二乘估计，如果有一些模型的知识，如 **E(v)=0**， $E\left[\mathbf{n}\mathbf{n}^T\right] = R$， 当 $W = R^{-1}$ 时，估计误差的方差阵达到最小，这个最小的方差阵为：

$$Var(\tilde{\boldsymbol{\theta}}_{lsR^{-1}}) = E\{[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{lsR^{-1}}][\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{lsR^{-1}}]^T\} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}$$

7.6最小二乘估计

例1、观测数据为：

$$z_k = a + n_k \quad k = 1, 2, \cdots, N$$

其中a为待估参量，n_k为观测噪声，求a的最小二乘估计。

7.6最小二乘估计

例2、根据以下对二维矢量 $\boldsymbol{\theta}$ 的两次观测，

$$z_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{\theta} + n_1$$

$$z_2 = 4 = \begin{bmatrix} 1 & 2 \end{bmatrix} \boldsymbol{\theta} + n_2$$

求 $\boldsymbol{\theta}$ 的线性最小二乘估计。

$$\hat{\boldsymbol{\theta}}_{ls} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{z} = \left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}^T \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 11 \end{bmatrix} = \begin{bmatrix} 1 \\ \dfrac{4}{3} \end{bmatrix}$$

7.6最小二乘估计

➢ 最小二乘估计在目标跟踪中的应用

匀速直线运动的观测模型：

$$z(i) = x_0 + \dot{x}_0 t_i + w(i)$$

$$\mathbf{X}_0 = \begin{bmatrix} x_0 & \dot{x}_0 \end{bmatrix}^T \qquad \mathbf{H}(i) = \begin{bmatrix} 1 & t_i \end{bmatrix}$$

$$z(i) = \mathbf{H}(i)\mathbf{X}_0 + w(i)$$

习题：7.26,7.31

7. 7波形估计

1、波形估计

参量估计适用于非时变参量，无法解决时变参量估计问题。

关于时变参量甚至时变信号本身的估计称为时变信号估计或波形估计，因此波形估计又称过程估计。

波形估计其实质就是给定有用信号和加性噪声的混合波形，寻求一种线性运算作用于此混合波形，使信号与噪声实现最佳分离，最佳的含义是使估计的均方误差最小，故又称为最佳线性滤波理论。

7.7波形估计

波形估计通常分为滤波、平滑、预测三种基本估计。

• 滤波： 根据当前和过去的观测值$\{z(k), k = n_0, n_0+1, \ldots, n\}$对信号$s(n)$进行估计(**Filtering**)；

•预测：根据当前和过去的观测值$\{z(k), k = n_0, n_0+1, \ldots, n_f\}$对未来时刻$n(n>n_f)$的信号$s(n)$进行估计，预测也称为外推；(**Prediction**)

•根据某一区间的观测数据$\{z(k), k = n_0, n_0+1, \ldots, n_f\}$对区间内的某一个时刻$n(n_0<n<n_f)$的信号进行估计，内插也称为平滑。(**Smoothing**)。

7.7波形估计

2、维纳滤波

线性最小均方估计是观测的线性函数，它可以看作为观测序列通过离散时间线性系统，即

$$\hat{s}(n)=\sum_{k=n_0}^{n}h(n,k)z(k)$$

滤波器系数的选择：

$$E\left\{\left[s(n)-\sum_{k=n_0}^{n}h(n,k)z(k)\right]z(i)\right\}=0$$ 正交原理

$$R_{sz}(n,i)=\sum_{k=n_0}^{n}h(n,k)R_z(k,i)$$ Wiener-Hopf方程

7.7波形估计

假定信号和观测过程是平稳随机序列，并且是联合平稳随机序列，

系统为因果的线性时不变离散时间线性系统

$$\hat{s}(n)=\sum_{k=n_0}^{n}h(n,k)z(k)$$

$$\Rightarrow \quad \hat{s}(n)=\sum_{k=-\infty}^{n}h(n-k)z(k)=h(n)\otimes z(n)$$

$$R_{sz}(n,i)=\sum_{k=n_0}^{n}h(n,k)R_z(k,i)$$

$$\Rightarrow \quad R_{sz}(n-i)=\sum_{k=-\infty}^{n}h(n-k)R_z(k-i)$$

$$\Rightarrow \quad R_{sz}(n)=\sum_{l=0}^{\infty}h(l)R_z(n-l)=h(n)\otimes R_z(n)$$

7.7波形估计

$$R_{sz}(n)=\sum_{l=0}^{\infty}h(l)R_z(n-l)=h(n)\otimes R_z(n)$$

$$\Longrightarrow\quad G_{sz}(z)=H(z)G_z(z)$$

$$\Longrightarrow\quad H(z)=\frac{G_{sz}(z)}{G_z(z)}$$

维纳滤波器

信号**s(n)**与观测噪声统计独立时，维纳滤波器为：

$$\Longrightarrow\quad H(z)=\frac{G_s(z)}{G_s(z)+G_v(z)}$$

观测为白噪声时，维纳滤波器为：　$n\geq 0$

$$\Longrightarrow\quad h(n)=R_{sz}(n)\qquad H(z)=G_{sz}^{+}(z)$$

维纳滤波和卡尔曼滤波是实现从噪声中提取信号，完成信号波形估计的两种线性最佳估计方法。

维纳滤波需要设计维纳滤波器，它的求解要求知道随机信号的统计特性，即相关函数或功率谱密度。当信号的功率谱为有理谱时，采用谱分解的方法求解滤波器的系统函数，简单易行，物理概念清楚，具有一定的工程实用价值，但当功率谱变化时，却不能进行实时处理。维纳滤波的限制是：它仅实用于一维平稳随机信号。

20世纪50年代，为了解决多输入、多输出非平稳随机信号的估计问题，卡尔曼于60年采用状态方程和观测方程描述系统的信号模型，提出离散状态估计的一组递推公式，即卡尔曼滤波器公式。由于卡尔曼滤波采用的递推算法非常适合计算机处理，广泛应用于许多领域。