

估值概述

1. 估值的意义

随机变量的分布函数完全确定了它的统计特性，但是，基于这种统计特性的描述，我们依然无法预测某一次试验的观察结果确定地是什么。

如果想用某一个确定的值去作为观察结果的估计值，可以[利用概率分布函数](#)去选择一个合理的值。

数学期望：统计平均值，无条件估值，统计误差为方差

估值：利用已有的观测量的信息，估计估计量的取值，减小关于估计量的不确定性。

2. 估值的应用

最基本的问题是利用对一个随机变量的观测对另一个随机变量进行估计

根据当前取值来预测下一个时刻取值的预测编码；

根据接收的信号序列来估计发送序列某个时刻发送值的均衡问题；

根据信号和噪声之和的序列来估计信号的噪声抑制问题；

3. 估值问题的研究内容与方法

估值研究的是统计意义上的估计问题。由于是对一个不确定量的估计，一定存在估计误差，关键在于[利用某种原则、方法选择估计值](#)，使[估计误差在某种意义下是最小的](#)。

估值准则的选取：有实际意义，有用；并且是能够得到简单解的。

“估计的均方误差最小”的估计准则是应用最广的一种估计思想，简称 MS 估计，相应的估计误差简称 MMSE。

“输出信噪比最大”是雷达系统中最佳接收的工作原理。

在一定的估计准则下，有不同的[解决具体估值问题的方法](#)：

- ✓ 非线性估值
- ✓ 线性估值
- ✓ 维纳滤波
- ✓ 匹配滤波
- ✓ 递归线性均方估值

其中一个重要的问题就是[确定估值参数](#)

估值的评价：分析估值误差

均方误差最小的最佳线性估计

估值问题

设 ξ 和 η 是两个随机矢量，两者存在联合分布，其中 η 是观察矢量，通过 η 对 ξ 进行估值，得到符合某种准则的最佳估值 ξ 。

1 均方误差最小的估值问题

均方误差最小的估值问题

设 ξ 和 η 是两个随机矢量，两者存在联合分布，设 η 是观察矢量，通过 η 对 ξ 进行估值，求均方误差最小的估值 ξ 。

$$E\left\{\left\|\xi - \hat{\xi}(\eta)\right\|^2 / \eta = Y\right\} = \min E\left\{\left\|\xi - K\right\|^2 / \eta = Y\right\}, \text{ 其中 } K = (k_1, k_2, \dots, k_n)^T$$

为任意矢量。

均方误差最小估值的解：均方估值和条件均值的关系

$$\begin{aligned} & E\left\{\left\|\xi - K\right\|^2 / \eta = Y\right\} \\ &= E\left\{\left\|\xi - E\{\xi/\eta\} + E\{\xi/\eta\} - K\right\|^2 / \eta = Y\right\} \\ &= E\left\{\left\|E\{\xi/\eta\} - K\right\|^2 / \eta = Y\right\} \\ &\quad + E\left\{\left\|\xi - E\{\xi/\eta\}\right\|^2 / \eta = Y\right\} \\ &\quad + 2E\left\{\left[\xi - E\{\xi/\eta\}\right] \cdot \left[E\{\xi/\eta\} - K\right] / \eta = Y\right\} \\ &\quad + 2E\left\{\left[E\{\xi/\eta\} - K\right] \cdot \left[\xi - E\{\xi/\eta\}\right] / \eta = Y\right\} \\ &= E\left\{\left\|E\{\xi/\eta\} - K\right\|^2 / \eta = Y\right\} + E\left\{\left\|\xi - E\{\xi/\eta\}\right\|^2 / \eta = Y\right\} \end{aligned}$$

为了使均方误差最小，

$$\text{应使 } K = E\{\xi/\eta\}, \text{ 即: } \hat{\xi}(\eta = Y) = E\{\xi/\eta = Y\}$$

定理

设随机矢量 $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ 和 $\eta = (\eta_1, \eta_2, \dots, \eta_m)$ 的联合概率密度函数是，

$f_{\xi\eta}(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m)$ 且 $f_{\eta}(y_1, y_2, \dots, y_m) \neq 0$ ，则 ξ 关于 η 的最小均方误差估计为 $\hat{\xi}(\eta) = E\{\xi/\eta\}$ ，它满足 $E\left\{\left\|\xi - \hat{\xi}(\eta)\right\|^2\right\} = \min E\left\{\left\|\xi - g(\eta)\right\|^2\right\}$ 。

例：

设 ξ 和 η 服从二元高斯分布，求最佳估值 $\hat{\xi}(\eta)$

解：

$$f_{\xi/\eta}(x/\eta=y) = \frac{1}{\sqrt{2\pi(1-r^2)}\sigma_{\xi}} \exp\left(-\frac{1}{2(1-r^2)\sigma_{\xi}^2}\left(x-\mu_{\xi}-\rho\frac{\sigma_{\xi}}{\sigma_{\eta}}(y-\mu_{\eta})\right)^2\right)$$

$$\hat{\xi}(\eta) = E\{\xi/\eta=Y\} = \mu_{\xi} + \rho\frac{\sigma_{\xi}}{\sigma_{\eta}}(y-\mu_{\eta})$$

$$E\{\hat{\xi}(\eta)\} = E\{E\{\xi/\eta=Y\}\} = \mu_{\xi}$$

$$\begin{aligned} E\left[\left[\hat{\xi}(\eta)-\mu_{\xi}\right]^2\right] &= E\left\{\left[\rho(\sigma_{\xi}/\sigma_{\eta})\cdot(y-\mu_{\eta})\right]^2\right\} \\ &= (\rho\sigma_{\xi})^2 \end{aligned}$$

2 均方误差最小的线性估计

对于正态分布的随机变量， ξ 关于的 η 最小均方误差估计 $\hat{\xi}(\eta) = E(\xi/\eta)$ 是一个线性估计，对于非正态分布的随机变量，估计 $\hat{\xi}(\eta) = E(\xi/\eta)$ 一般不是线性的。这时需要找到一个关于 η 的线性函数作为最佳估计。

定义：

ξ 关于的 η 估计 $\hat{\xi}(\eta) = \mathbf{A}\eta + \mathbf{b}$ ，使得 $E\left\{\left\|\xi - \hat{\xi}\right\|^2\right\} = E\left\{\left\|\xi - (\mathbf{A}\eta + \mathbf{b})\right\|^2\right\}$ 达到最小

值，这时 $\hat{\xi}(\eta) = \mathbf{A}\eta + \mathbf{b}$ 称作为最佳线性估值。其中 ξ 是 $n \times 1$ 的矢量， η 是 $m \times 1$ 的矢量， \mathbf{A} 是 $n \times m$ 的矩阵， \mathbf{b} 是 $n \times 1$ 的矢量。

问题：

确定 \mathbf{A} 、 \mathbf{b}

结论：

$$b_j = E\xi_j - A_j E\eta$$

$$A_j = \left(C_{[\xi_j \eta]}\right) \left(C_{[\eta \eta]}\right)^{-1}$$

即，

$$\hat{\xi}_j = E\xi_j - C_{[\xi_j \eta]} C_{[\eta \eta]}^{-1} E\eta + C_{[\xi_j \eta]} C_{[\eta \eta]}^{-1} \eta$$

$$\hat{\xi} = E\xi - C_{\xi\eta} C_{\eta\eta}^{-1} E\eta + C_{\xi\eta} C_{\eta\eta}^{-1} \eta$$

如果 ξ , $\eta = (\eta_1 \ \eta_2 \ \cdots \ \eta_m)^T$ 均值为零, 则有

$$\hat{\xi}_j = C_{[\xi_j \eta]} C_{[\eta \eta]}^{-1} \eta$$

方法:

1) 极值分析法

求估值的均方误差

$$E \left\{ \left[\xi_j - \hat{\xi}_j \right]^2 / \eta \right\} = D\xi_j + A_j C_{[\eta \eta]} A_j^H - A_j \left(C_{[\xi_j \eta]} \right)^H - \overline{A_j} \left(C_{[\xi_j \eta]} \right)^T + \left[E\xi_j - A_j E\eta - b_j \right]^2$$

A_j, b_j 的取值应使估值的均方误差对 A_j, b_j 的偏微商为零, 从而确定 A_j, b_j ;

2) 正交性原理

定理: 满足正交性的线性估计 (从正交性原理进行推导)

如果用复随机变量 $\eta_1, \eta_2, \dots, \eta_m$ 的线性组合 $\hat{\xi} = a_1 \eta_1 + a_2 \eta_2 + \dots + a_m \eta_m$ 来估计复随机变量 ξ , 其估计误差 $e = \xi - \hat{\xi}$ 。若 e 和 $\eta_1, \eta_2, \dots, \eta_m$ 正交, 则此估计是最小均方误差线性估计。

$$\text{利用 } E \left\{ e \cdot \overline{\eta_j} \right\} = E \left\{ [\xi - (a_1 \eta_1 + a_2 \eta_2 + a_3 \eta_3 + \dots + a_m \eta_m)] \cdot \overline{\eta_j} \right\} = 0$$

注意: 两种方法得到的结论一致, 正交性原理的推导证明中, 先假设 $E\{\xi\} = E\{\eta\} = 0$, 进一步的论述中说明了 $E\{\xi\} \neq 0, E\{\eta\} \neq 0$ 时, 满足正交性原理的线性估计仍然是最小均方误差估计。

最佳线性估计的均方误差

根据正交性原理, 均方误差最小的最佳线性估计的误差与观察量正交, 从而与估计值正交:

$$E \left\{ e \cdot \overline{\eta_j} \right\} = E \left\{ (\xi - \hat{\xi}) \overline{\eta_j} \right\} = 0$$

则:

$$E \left\{ (\xi - \hat{\xi}) \hat{\xi} \right\} = 0$$

因此:

$$E \left\{ \left| \xi - \hat{\xi} \right|^2 \right\} = E \left\{ \left| \xi \right|^2 \right\} - E \left\{ \left| \hat{\xi} \right|^2 \right\},$$

$$E \left\{ \left| \xi - \hat{\xi} \right|^2 \right\} = C_{\xi\xi} - [C_{\xi\eta}] [C_{\eta\eta}]^{-1} [C_{\eta\xi}]。$$

推论 1

如果复随机变量 ξ 和复随机变量 $\eta_1, \eta_2, \eta_3, \dots, \eta_m$ 不相关, 即相互正交, 此时无法用复随机变量 $\eta_1, \eta_2, \eta_3, \dots, \eta_m$ 对复随机变量 ξ 进行估计。此时的最佳估计就是 $\hat{\xi}(\eta) = 0$ 。

推论 2

如果复随机变量 ξ 和复随机变量 $\eta_1, \eta_2, \eta_3, \dots, \eta_m$ 有相关性, 进行估计, 相应的均方误差就减少了,

$$E\{e^2\} = E\left\{\left|\xi - \hat{\xi}\right|^2\right\} = E\left\{(\xi - \hat{\xi})(\xi - \hat{\xi})^*\right\} \\ = C_{\xi\xi} - [C_{\xi\eta}] [C_{\eta\eta}]^{-1} [C_{\xi\eta}]^T$$

减少量是 $[C_{\xi\eta}] [C_{\eta\eta}]^{-1} [C_{\xi\eta}]^T$ 。

推论 3

如果复随机变量 $\eta_1, \eta_2, \eta_3, \dots, \eta_m$ 彼此之间相互正交, 有

$$a_j = (C_{\xi\eta} \cdot C_{\eta\eta}^{-1})_{jj} = C_{\xi\eta_j} C_{\eta_j\eta_j}^{-1}$$

推论 4

增加新的数据可以改善原有的估值, 但是如果增加用来估值的数据 $\eta_{1+m}, \eta_{2+m}, \dots$ 与原来的估计误差 $e = \xi - (a_1\xi_1 + a_2\xi_2 + \dots + a_m\xi_m)$ 正交, 这时增加数据不能改善原来的估计。

3 典型估值问题举例

估值理论的应用

随机过程的估值

随机过程的预测

信号的统计插值和滤波

典型的例子

例 1, 设有零均值实平稳随机过程 $s(t)$, 已知 t 时刻的值, 试估计 $t+s$ 时刻的值。

例 2, 设有零均值实平稳随机过程的信号 $s(t)$ 和零均值实平稳随机过程的噪声 $n(t)$, 它们是统计独立的随机过程, 实际测到的是 $x(t)=s(t)+n(t)$, 试根据 $x(t)$ 对 $s(t)$ 进行估值。

例 3, 设有零均值实平稳随机过程的信号 $s(t)$, 试根据 $s(0)$ 及 $s(T)$ 对 $s(t)$ 进行估值。

例 4, 利用 t 时刻随机过程 $s(t)$ 的观察值和它的导数 $s'(t)$ 观察值, 对未来的 $s(t+\lambda)$, $\lambda>0$, 进行估值。

例 5, 已知 t_1, t_2 两个时刻 $t_1>t_2$, 利用这两个随机过程 $s(t)$ 的观察值对未来的 $s(t)$, 进行估值。

例 6, 已知零均值实随机过程 $s(t)$, 利用 $s(0)$ 和 $s(T)$ 的观察值, 对随机过程在 $[0, T]$ 区间的积分进行估计。求最佳估值。

假定研究的随机过程是零均值实平稳的随机过程

例 1

设有零均值实平稳随机过程 $s(t)$ ，已知 t 时刻的值，试估计 $t+s$ 时刻的值。

解：

利用最佳线性估值原理， $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\hat{s}(t+s) = as(t), \text{ 且}$$

$$a = E\{s(t+\lambda)s(t)\} / E\{s(t)s(t)\} = R_{ss}(\lambda) / R_{ss}(0)$$

相应的最小均方误差是， $E\left\{\left|\xi - \hat{\xi}\right|^2\right\} = E\left\{\xi^2\right\} - E\left\{\xi \hat{\xi}^*\right\}$

$$\begin{aligned} E\left\{[s(t+\lambda) - \hat{s}(t+\lambda)]^2\right\} &= E\{[s(t+\lambda) - as(t)]s(t+\lambda)\} \\ &= E\{s(t+\lambda)s(t+\lambda)\} - aE\{s(t)s(t+\lambda)\} \\ &= R_{ss}(0) - aR_{ss}(\lambda) \\ &= R_{ss}(0) - [R_{ss}(\lambda)]^2 / R_{ss}(0) \end{aligned}$$

讨论 1：如果随机过程的相关函数是 $R_{ss}(\tau) = \sigma^2 e^{-\alpha|\tau|}$

$$\begin{aligned} a &= R_{ss}(\lambda) / R_{ss}(0) = e^{-\alpha\lambda} \\ E\left\{[s(t+\lambda) - \hat{s}(t+\lambda)]^2\right\} &= \sigma^2(1 - e^{-2\alpha\lambda}) \end{aligned}$$

讨论 2：在上述条件下，估计误差正交于 $s(u)$ ， $u < t$

$$\begin{aligned} E\{[s(t+\lambda) - as(t)] \cdot s(u)\} &= R_{ss}(t+\lambda-u) - aR_{ss}(t-u) \\ &= \sigma^2 e^{-\alpha(t+\lambda-u)} - a\sigma^2 e^{-\alpha(t-u)} = \sigma^2 e^{-\alpha(t+\lambda-u)} - e^{-\alpha\lambda} \sigma^2 e^{-\alpha(t-u)} \\ &= 0 \end{aligned}$$

例 2

设有零均值实平稳随机过程的信号 $s(t)$ 和零均值实平稳随机过程的噪声 $n(t)$ ，它们是统计独立的随机过程，实际测到的是 $x(t)=s(t)+n(t)$ ，试根据 $x(t)$ 对 $s(t)$ 进行估值。

解：

利用最佳线性估值原理， $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\hat{s}(t) = ax(t), \text{ 且}$$

$$a = E\{s(t)x(t)\} / E\{x(t)x(t)\} = R_{ss}(0) / [R_{ss}(0) + R_{nn}(0)]$$

相应的最小均方误差是, $E\left\{\left|\xi - \hat{\xi}\right|^2\right\} = E\left\{\xi^2\right\} - E\left\{\xi \hat{\xi}^*\right\}$

$$\begin{aligned} E\left\{|s(t) - \hat{s}(t)|^2\right\} &= E\{[s(t) - ax(t)]s(t)\} \\ &= E\{s(t)s(t)\} - aE\{x(t)s(t)\} \\ &= R_{ss}(0) - aR_{ss}(0) \\ &= R_{ss}(0) - \frac{R_{ss}(0)}{R_{ss}(0) + R_{nn}(0)} R_{ss}(0) \\ &= R_{ss}(0) \frac{R_{nn}(0)}{R_{ss}(0) + R_{nn}(0)} \end{aligned}$$

例 3

设有零均值实平稳随机过程的信号 $s(t)$, 试根据 $s(0)$ 及 $s(T)$ 对 $s(t)$ 进行估值。

解:

利用最佳线性估值原理, $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\begin{aligned} s(t) &= (R_{ss}(t) \ R_{ss}(T-t)) \begin{pmatrix} R_{ss}(0) & R_{ss}(T) \\ R_{ss}(T) & R_{ss}(0) \end{pmatrix}^{-1} \begin{pmatrix} s(0) \\ s(T) \end{pmatrix} \\ &= as(0) + bs(T) \end{aligned}$$

其中,

$$\begin{aligned} a &= \frac{R_{ss}(t)R_{ss}(0) - R_{ss}(T-t)R_{ss}(T)}{[R_{ss}(0)]^2 - [R_{ss}(T)]^2} \\ b &= \frac{-R_{ss}(t)R_{ss}(T) + R_{ss}(T-t)R_{ss}(0)}{[R_{ss}(0)]^2 - [R_{ss}(T)]^2} \end{aligned}$$

相应的最小均方误差是, $E\left\{\left|\xi - \hat{\xi}\right|^2\right\} = E\left\{\xi^2\right\} - E\left\{\xi \hat{\xi}^*\right\}$

$$\begin{aligned} E\left\{|s(t) - \hat{s}(t)|^2\right\} &= E\{[s(t) - as(0) - bs(T)]s(t)\} \\ &= R_{ss}(0) - aR_{ss}(t) - bR_{ss}(T-t) \end{aligned}$$

例 4

利用 t 时刻随机过程 $s(t)$ 的观察值和它的导数 $s'(t)$ 观察值, 对未来的 $s(t+\lambda)$, $\lambda > 0$, 进行估值。

解:

利用最佳线性估值原理, $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\begin{aligned}
\hat{s}(t+\lambda) &= (R_{ss}(\lambda) - R'_{ss}(\lambda)) \begin{pmatrix} R_{ss}(0) & -R'_{ss}(0) \\ -R'_{ss}(0) & R_{ss}(0) \end{pmatrix}^{-1} \begin{pmatrix} s(t) \\ s'(t) \end{pmatrix} \\
&= (R_{ss}(\lambda) - R'_{ss}(\lambda)) \begin{pmatrix} R_{ss}(0) & 0 \\ 0 & R_{ss}(0) \end{pmatrix}^{-1} \begin{pmatrix} s(t) \\ s'(t) \end{pmatrix} \\
&= as(t) + bs'(t) \\
a &= \frac{R_{ss}(\lambda)}{R_{ss}(0)}, \quad b = \frac{R'_{ss}(\lambda)}{R_{ss}(0)}
\end{aligned}$$

相应的最小均方误差是, $E\left\{\left|\xi - \hat{\xi}\right|^2\right\} = E\left\{\xi^2\right\} - E\left\{\xi \hat{\xi}^*\right\}$

$$\begin{aligned}
E\left\{[s(t+\lambda) - \hat{s}(t+\lambda)]^2\right\} &= E\left\{[s(t+\lambda) - \hat{s}(t+\lambda)]s(t+\lambda)\right\} \\
&= E\left\{[s(t+\lambda) - as(t) - bs'(t)]s(t+\lambda)\right\} \\
&= R_{ss}(0) - aR_{ss}(\lambda) + bR'_{ss}(\lambda)
\end{aligned}$$

例 5

已知 t_1 、 t_2 两个时刻 $t_1 > t_2$, 利用这两个随机过程 $s(t)$ 的观察值对未来的 $s(t)$, 进行估值。
解:

利用最佳线性估值原理, $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\begin{aligned}
\hat{s}(t) &= (R_{ss}(t-t_1) \quad R_{ss}(t-t_2)) \begin{pmatrix} R_{ss}(0) & R_{ss}(t_1-t_2) \\ R_{ss}(t_1-t_2) & R_{ss}(0) \end{pmatrix}^{-1} \begin{pmatrix} s(t_1) \\ s(t_2) \end{pmatrix} \\
&= as(t-t_1) + bs(t-t_2) \\
a &= \frac{R_{ss}(t-t_1)R_{ss}(0) - R_{ss}(t-t_2)R_{ss}(t_1-t_2)}{R_{ss}^2(0) - R_{ss}^2(t_1-t_2)} \\
b &= \frac{-R_{ss}(t-t_1)R_{ss}(t_1-t_2) + R_{ss}(t-t_2)R_{ss}(0)}{R_{ss}^2(0) - R_{ss}^2(t_1-t_2)}
\end{aligned}$$

相应的最小均方误差是,

$$\begin{aligned}
E\left\{[s(t) - \hat{s}(t)]^2\right\} &= E\left\{[s(t) - \hat{s}(t)]s(t)\right\} \\
&= E\left\{[s(t)s(t) - as(t_1)s(t) - bs(t_2)s(t)]\right\} \\
&= R_{ss}(0) - aR_{ss}(t-t_1) - bR_{ss}(t-t_2)
\end{aligned}$$

讨论 1, 若 t_2 时刻的信息不能改进对 $s(t)$ 的估计, 即 $b=0$,

$$b = 0,$$

$$R_{ss}(t-t_1)R_{ss}(t_1-t_2) = R_{ss}(t-t_2)R_{ss}(0)$$

$$R_{ss}(t-t_2) = \frac{R_{ss}(t-t_1)R_{ss}(t_1-t_2)}{R_{ss}(0)}$$

相应的相关函数是负指数形式。

例 6

已知零均值实随机过程 $s(t)$, 利用 $s(0)$ 和 $s(T)$ 的观察值, 对随机过程在 $[0, T]$ 区间的积分进行估计。求最佳估值。

解:

利用最佳线性估值原理, $\hat{\xi} = a^T \eta = [\vec{C}_{\xi\eta}]^T [C_{\eta\eta}]^{-1} \cdot \eta$

$$\text{定义 } \zeta = \int_0^T s(u) du$$

$$E[\zeta \cdot s(0)] = E\left(\int_0^T s(u) du \cdot s(0)\right) = \int_0^T R_{ss}(u) du$$

$$E[\zeta \cdot s(T)] = E\left(\int_0^T s(u) du \cdot s(T)\right) = \int_0^T R_{ss}(u-T) du$$

$$= \int_0^T R_{ss}(T-u) du = \int_0^T R_{ss}(u) du$$

记作

$$E[\zeta \cdot s(0)] = E[\zeta \cdot s(T)] = \int_0^T R_{ss}(u) du = \alpha$$

最佳估值是,

$$\begin{aligned} \hat{\zeta} &= (E[\zeta \cdot s(0)] \quad E[\zeta \cdot s(T)]) \begin{pmatrix} R_{ss}(0) & R_{ss}(-T) \\ R_{ss}(T) & R_{ss}(0) \end{pmatrix}^{-1} (s(0) \quad s(T))^T \\ &= (\alpha \quad \alpha) \begin{pmatrix} R_{ss}(0)s(0) - R_{ss}(T)s(T) \\ -R_{ss}(T)s(0) + R_{ss}(0)s(T) \end{pmatrix} / (R_{ss}^2(0) - R_{ss}^2(T)) \\ &= \alpha (R_{ss}(0) - R_{ss}(T)) (s(0) + s(T)) / (R_{ss}^2(0) - R_{ss}^2(T)) \\ &= \frac{\int_0^T R_{ss}(u) du}{(R_{ss}(0) + R_{ss}(T))} \cdot (s(0) + s(T)) \end{aligned}$$