《统计信号处理》讨论题二：贝叶斯估计

杨鼎

贝叶斯估计和非贝叶斯估计的区别是什么？贝叶斯估计的先验信息如何获得？对于高斯白噪声中指数信号参数估计问题

$$z_n = Ar^n + w_n, \quad n = 0, 1, \cdots, N-1$$

其中 $A \sim N(0, \sigma^2), w_n \overset{i,i,d}{\sim}$，A 与 w_n 相互独立；$0 < r < 1, \sigma^2 > 0, \sigma_A^2 > 0$ 均为已知参数。试求参数 A 的最小均方估计和最大后验概率估计，并分析二者的关系。若对参数 A 没有更多额先验信息时可如何处理？

1. 贝叶斯估计和非贝叶斯估计的区别是什么？贝叶斯估计的先验信息如何获得？

贝叶斯估计和非贝叶斯估计的区别是有无先验信息。

贝叶斯估计的先验信息通常是建立在物理约束的基础上；也可以从数据估计先验，称为经验贝叶斯。

2. 对于高斯白噪声中指数信号参数估计问题

$$z_n = Ar^n + w_n, \quad n = 0, 1, \cdots, N-1$$

其中 $A \sim \mathcal{N}(0,\sigma^2), w_n \overset{i,i,d}{\sim} \mathcal{N}(0,\sigma^2)$，$\mathbf{A}$ 与 w_n 相互独立；$0<r<1, \sigma^2>0, \sigma_A^2>0$ 均为已知参数。试求参数 $\mathbf{A}$ 的最小均方估计和最大后验概率估计，并分析二者的关系。若对参数 $\mathbf{A}$ 没有更多额先验信息时可如何处理？

(1) 参数 $\mathbf{A}$ 的最小均方估计

观测模型用矢量形式表示

$$\mathbf{Z} = A\mathbf{H} + \mathbf{w}$$

其中

$$\begin{aligned}
\mathbf{Z} &= \begin{bmatrix} z_0 & z_1 & \cdots & z_{N-1} \end{bmatrix}^T \\
\mathbf{w} &= \begin{bmatrix} w_0 & w_1 & \cdots & w_{N-1} \end{bmatrix}^T \\
\mathbf{H} &= \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r^{N-1} \end{bmatrix}^T
\end{aligned}$$

由于 $E(A)=0, E(\mathbf{w})=\mathbf{0}, E(\mathbf{Z})=\mathbf{0}$, 则

$$\begin{aligned}
\mathbf{C}_{A\mathbf{Z}} &= E[A(\mathbf{Z}-E(\mathbf{Z}))^T] \\
&= E[A(A\mathbf{H}+\mathbf{w})^T] \\
&= E[A^2\mathbf{H}^T] \\
&= \sigma_A^2\mathbf{H}^T
\end{aligned}$$

$$\begin{aligned}
\mathbf{C}_{\mathbf{Z}} &= E[(\mathbf{Z}-E(\mathbf{Z}))(\mathbf{Z}-E(\mathbf{Z}))^T] \\
&= E[(A\mathbf{H}+\mathbf{w})(A\mathbf{H}+\mathbf{w})^T] \\
&= E[A^2\mathbf{H}\mathbf{H}^T] + E[\mathbf{w}\mathbf{w}^T] \\
&= \sigma_A^2\mathbf{H}\mathbf{H}^T + \sigma^2\mathbf{I}
\end{aligned}$$

最小 MSE 估计量

$$
\begin{aligned}
\hat{A} = E(A|\mathbf{Z}) &= E(\mathbf{A}) + \mathbf{C}_{A\mathbf{Z}}\mathbf{C}_{\mathbf{Z}}^{-1}(\mathbf{Z} - E(\mathbf{Z})) \\
&= \sigma_A^2\mathbf{H}^T(\sigma_A^2\mathbf{H}\mathbf{H}^T + \sigma^2\mathbf{I})^{-1}\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}\mathbf{H}^T(\frac{\sigma_A^2}{\sigma^2}\mathbf{H}\mathbf{H}^T + \mathbf{I})^{-1}\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}\mathbf{H}^T(\mathbf{I} - \frac{\sigma_A^2}{\sigma^2}\mathbf{H}(\mathbf{I} + \frac{\sigma_A^2}{\sigma^2}\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}\mathbf{H}^T(\mathbf{I} - \frac{\sigma_A^2/\sigma^2\mathbf{H}\mathbf{H}^T}{\mathbf{I} + \sigma_A^2/\sigma^2\mathbf{H}^T\mathbf{H}})\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}(\mathbf{H}^T - \frac{\sigma_A^2/\sigma^2\mathbf{H}^T\mathbf{H}\mathbf{H}^T}{\mathbf{I} + \sigma_A^2/\sigma^2\mathbf{H}^T\mathbf{H}})\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}(1 - \frac{\sigma_A^2/\sigma^2\mathbf{H}^T\mathbf{H}}{1 + \sigma_A^2/\sigma^2\mathbf{H}^T\mathbf{H}})\mathbf{H}^T\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma^2}\frac{\sigma^2}{\sigma^2 + \sigma_A^2 + \mathbf{H}^T\mathbf{H}}\mathbf{H}^T\mathbf{Z} \\
&= \frac{\sigma_A^2}{\sigma_A^2\frac{1-r^{2N}}{1-r^2} + \sigma^2}\sum_{n=0}^{N-1} r_n z_n
\end{aligned}
$$

其中用到了矩阵求逆引理 $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1}$。

(2) 参数 A 的最大后验概率估计

最大后验概率估计为 $\hat{A}_{map} = arg\max_{A} p(A|\mathbf{Z})$，由贝叶斯公式 $p(A|\mathbf{Z}) = \frac{p(\mathbf{Z}|A)p(A)}{p(\mathbf{Z})}$，最大后验概率估计等价为 $\hat{A}_{map} = arg\max_{A} p(\mathbf{Z}|A)p(A)$，已知

$$
\begin{aligned}
p(\mathbf{Z}|A) &= (\frac{1}{2\pi\sigma^2})^{\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(z_n - Ar^n)^2\right\} \\
p(A) &= \frac{1}{\sqrt{2\pi\sigma_A^2}}\exp\{-\frac{A^2}{2\sigma_A^2}\}
\end{aligned}
$$

所以

$$
\ln[p(\mathbf{Z}|A)p(A)] = -\frac{N\ln(2\pi\sigma^2) + \ln(2\pi\sigma_A^2)}{2} - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(z_n - Ar^n)^2 - \frac{1}{2\sigma_A^2}A^2
$$

令 $\ln[p(\mathbf{Z}|A)p(A)] = 0$, 得到

$$
\frac{1}{\sigma^2}\sum_{n=0}^{N-1} r^n(z_n - Ar^n) = \frac{A}{\sigma_A^2}
$$

解得

$$\hat{A}_{map} = \frac{\sigma_A^2}{\sigma_A^2 \frac{1-r^{2N}}{1-r^2} + \sigma^2} \sum_{n=0}^{N-1} z_n r^n$$

与最小均方估计结果相同。

若对参数 A 没有更多的先验信息，可以考虑用经典估计方法。