第 7 章 估计理论

7.1　设有 N 次独立观测 $z_i = A + v_i (i = 1, 2, \cdots, N)$, 其中 $v_i \sim N\left(0, \sigma^2\right)$, A 在 $(-A_0, A_0)$ 上服从均匀分布, 证明 A 的最大后验概率估计为

$$\hat{A}_{\text{map}} = \begin{cases} -A_0 & \left(\bar{z} < -A_0\right) \\ \bar{z} & \left(-A_0 \leqslant \bar{z} \leqslant A_0\right) \\ A_0 & \left(\bar{z} > A_0\right) \end{cases},$$

其中 $\bar{z} = \dfrac{1}{N} \sum\limits_{i=1}^{N} z_i$.

7.2　单次观测 $z = s + n$ 是高斯分布信号与均匀分布噪声之和, 且两者相互独立, 其概率密度如图所示.

(1) 对 z 的各个范围 $z < -1, z > 1, |z| < 1$, 画出 $f(s|z)$;

(2) 求 $\hat{S}_{\text{map}}$;

(3) 当 $z > 1$ 时, 将最小均方估计表示为两积分之比.

7.3　从有噪声的观测中估计天线方位角. 在观测之前已知角度 s 在 $[-1, 1]$(单位为 mrad) 上均匀分布, 噪声 n_i 是各自独立的且与 s 无关, 噪声的分布密度为

$$f\left(n_i\right) = \begin{cases} 1 - \left|n_i\right| & \left(-1 < n_i < 1\right) \\ 0 & (\text{else}) \end{cases},$$

观测样本为 $z_i = s + n_i$.

(1) 求单次观测 $z_1 = 1.5$ 时的最小均方估计;

(2) 求单次观测 $z_1 = 1.5$ 时的最大后验概率估计.

7.4　设随机常量 θ 的后验概率密度为

$$f(\theta|z) = \frac{\varepsilon}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\theta - z)^2\right] + \frac{1-\varepsilon}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\theta + z)^2\right],$$

其中 ε 是任意常数, $0 < \varepsilon < 1$, 求 θ 的最小均方估计和最大后验概率估计.

7.5　给定 $z = \dfrac{s}{2} + n, n$ 是均值为零, 方差为 1 的高斯随机变量:

(1) 求 s 的最大似然估计 $\hat{s}_{\text{ml}}$;

(2) 对下列 $f(s)$ 求最大后验估计估计 $\hat{s}_{\text{map}}$.

$$f(s) = \begin{cases} \dfrac{1}{4} \exp\left(-\dfrac{s}{4}\right) & (s \geqslant 0) \\ 0 & (s < 0) \end{cases}$$

7.6　设观测信号 $z_i = \theta + v_i (i = 1, 2, \cdots, N)$, 已知 v_i 是相互独立, 具有相同分布的高斯噪声, 其均值为 0, 方差为 σ_v^2; 信号 θ 也是一零均值方差为 σ_θ^2 的高斯信号, 且与噪声统计独立. 通过 N 次观测对信号 θ 进行估计, 求信号 θ 的最小均方估计 $\hat{\theta}_{\text{ms}}$, 最大后验估计 $\hat{\theta}_{\text{map}}$ 和条件中位数估计 $\hat{\theta}_{\text{med}}$.

解 先求后验概率密度:

$$
\begin{aligned}
f(\theta \mid \mathbf{z}) &= \frac{f(\mathbf{z} \mid \theta) f(\theta)}{\int_{-\infty}^{\infty} f(\mathbf{z} \mid \theta) f(\theta) d\theta} \\
&= \frac{\frac{1}{\left(2\pi\sigma^2\right)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(z_i-\theta\right)^2\right] \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left[-\frac{1}{2\sigma_0^2}\theta^2\right]}{\int_{-\infty}^{\infty} \frac{1}{\left(2\pi\sigma^2\right)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(z_i-\theta\right)^2\right] \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{1}{2\sigma_0^2}\theta^2\right] d\theta} \\
&= \frac{\exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\left(N\theta^2-2N\theta\bar{z}\right)+\frac{1}{\sigma_\theta^2}\theta^2\right]\right\}}{\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\left(N\theta^2-2N\theta\bar{z}\right)+\frac{1}{\sigma_\theta^2}\theta^2\right]\right\} d\theta} \\
&= \frac{\exp\left(-\frac{1}{2}W(\theta)\right)}{\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}W(\theta)\right) d\theta},
\end{aligned} \tag{1}
$$

7.7 在题 7.6 中以 $\hat{\theta}(N)$ 表示 N 次观测对 θ 的估计, 以 $\sigma^2(N)$ 表示相应的方差. 现观测到另一取样值 z_{N+1}, 试用 $\hat{\theta}(N)$、z_{N+1} 和 $\sigma^2(N)$ 表示估计 $\hat{\theta}(N+1)$, 此为 $\hat{\theta}(N)$ 提供了一种序列估计算法.

7.8 假定概率密度 $f(z;\theta)$ 满足正则条件:

$$
E\left\{\frac{\partial \ln f(z;\theta)}{\partial\theta}\right\} = 0.
$$

证明任何无偏估计量 $\hat{\theta}$ 的方差满足

$$
\operatorname{Var}(\hat{\theta}) = E\left[(\theta-\hat{\theta})^2\right] \geqslant \frac{1}{E\left\{\left[\frac{\partial \ln f(z;\theta)}{\partial\theta}\right]^2\right\}} = -\frac{1}{E\left\{\frac{\partial^2 \ln f(z;\theta)}{\partial\theta^2}\right\}},
$$

当且仅当

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

时, 上式的等号成立.

7.9 给定一独立观测序列 $x_1, z_2, \ldots, z_N$, 其均值为 m, 方差为 σ^2, 试问:

(1) 样本均值 $u = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} x_i$ 是否是 m 的无偏估计? 并求 u 的方差;

(2) 若方差的估计是 $v = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\left(z_i-u\right)^2$, 试问 v 是否是 σ^2 的无偏估计.

7.10 设 N 次观测为 $z_i = A + n_i (i=1,2,\ldots,N)$, 其中 A 为未知的确定信号, 噪声 $n_i (i=1,2,\ldots,N)$ 相互独立并服从同样的分布 $N\left(0,\sigma^2\right)$, 噪声的方差 σ^2 已知.

(1) 求 A 的极大似然估计 $\hat{A}_{\mathrm{m\,1}}$;

(2) $\hat{A}_{\mathrm{m\,1}}$ 是否为无偏估计?

(3) $\hat{A}_{\mathrm{m\,1}}$ 是否为有效估计? 估计的方差等于多少?

7.11 在题 7.10 中, 如果 A 和 σ^2 都是未知的, 试分别求其最大似然估计, 并讨论它们的无偏性.

7.12 要传输两个确定参数 A_1 和 A_2, 为了保证传输可靠, 现构造两个信号 s_1 和 s_2 分别在两个信道上传输, 则有

$$s_1 = x_{11}A_1 + x_{12}A_2, \quad s_2 = x_{21}A_1 + x_{22}A_2, \quad \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} \neq 0,$$

其中 $x_{ij}(i, j = 1, 2)$ 均为已知常数. 接收端获得的观测为

$$r_1 = s_1 + n_1, \quad r_2 = s_2 + n_2,$$

其中噪声 n_1 和 n_2 统计独立并且服从同样的分布 $N\left(0, \sigma_n^2\right)$.

(1) 求 A_1 和 A_2 的最大似然估计;

(2) A_1 和 A_2 的最大似然估计是否为无偏估计?

(3) A_1 和 A_2 的最大似然估计是否为有效估计?

7.13 证明: 矢量形式的线性最小均方估计是无偏估计, 均方误差阵满足

$$\begin{aligned}\boldsymbol{P}_{\theta} = \boldsymbol{E}\left[\tilde{\theta}\tilde{\theta}^{\mathrm{T}}\right] &= \boldsymbol{P}_{\theta} - \boldsymbol{P}_{\theta z}\boldsymbol{P}_z^{-1}\boldsymbol{P}_{z\theta} \\ &= \operatorname{Var}(\boldsymbol{\theta}) - \operatorname{Cov}(\boldsymbol{\theta}, \boldsymbol{z}) \operatorname{Var}^{-1}(\boldsymbol{z}) \operatorname{Cov}(\boldsymbol{z}, \boldsymbol{\theta}).\end{aligned}$$

7.14 设观察模型为 $z_i = a + v_i(i = 1, 2, \cdots, N)$, 其中随机参量 a 为信号幅值, $E(a) = 0, E(a^2) = A(A\text{为已知常数}), v_i$ 是零均值白噪声, $E\left(v_i v_j\right) = \sigma^2\delta_{ij}, E\left(v_i a\right) = 0$, 求 a 的线性最小均方估计.

7.15 设观测模型为 $z_k = ks + v_k(k = 1, 2, \cdots, N)$, 其中 ks 代表随时间增长的信号幅度, s 为增长速度, 若信号 ks 代表测量运动物体的距离, 那么 s 代表运动的速度, 假定 $E(s) = 0, E\left(s^2\right) = \boldsymbol{S}, v_k$ 是零均值白噪声, $E\left(v_k v_j\right) = \sigma^2\delta_{kj}, E\left(v_k a\right) = 0$, 求 s 的线性最小均方估计.

7.16 对于 (7.6.2) 式的线性观测模型, 假定测量噪声的均值为零, 证明最小二乘和加权最小二乘估计误差的方差阵分别为

$$\begin{aligned}\operatorname{Var}\left(\tilde{\theta}_{\mathrm{ls}}\right) &= \boldsymbol{E}\left\{\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{ls}}\right]\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{ls}}\right]^{\mathrm{T}}\right\} = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}, \\ \operatorname{Var}\left(\tilde{\theta}_{\mathrm{ls}}\right) &= E\left\{\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{law}}\right]\left[\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{lsw}}\right]^{\mathrm{T}}\right\} = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{R}\boldsymbol{W}\boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{H}\right)^{-1}.\end{aligned}$$

7.17 证明 $\operatorname{Var}\left(\tilde{\theta}_{\mathrm{lsw}}\right) \geqslant \operatorname{Var}\left(\tilde{\theta}_{\mathrm{ls}\,R^{-1}}\right) = \left(\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}\right)^{-1}$.(提示: 利用矩阵的 Schwartz 不等式, 设 $\boldsymbol{A}$ 为 $n \times m$ 的矩阵, $\boldsymbol{B}$ 为 $m \times l$ 的矩阵且 $\boldsymbol{A}\boldsymbol{A}^T$ 可逆.)

7.18 按题 7.3 给定的条件, 若新的观测技术容许在第二次观测中减少噪声, 即信道观测样本为 $y_1 = s + n_1, y_2 = s + \dfrac{n_1}{2}$, 其中 s 和 n_i 的统计特性不变, 求线性最小均方估计 $\hat{s}_{\mathrm{lms}} = h_1 y_1 + h_2 y_2$.

7.19 通过位移的测量估计车辆的加速度 a, 测量是具有噪声的, 以致实际数据样本有下列形式: $z_j = aj^2 + n_j(j = 1, 2, \cdots)$. 已知 $E(a) = 0, E\left(a^2\right) = \sigma_a^2, E\left(n_j\right) = 0, E\left(n_j^2\right) = \sigma_n^2, E\left(an_j\right) = 0, E\left(n_i n_j\right) = 0(i \neq j)$.

(1) 设噪声样本是正态的

$$f\left(n_j\right) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{n_j^2}{2\sigma_n^2}\right),$$

两个数据样本分别为 $z_1=a+n_1, z_2=4a+n_2$, 求最大似然估计;

(2) 两个数据样本同上, 如设 a, n_i 均是正态的, 且 $E(a)=E\left(n_j\right)=0, E\left(a^2\right)=\sigma_a^2=\sigma_n^2$, 求最大后验估计.

解

7.20 按题 7.19 给出的条件, 两个数据样本 $z_1=a+n_1, z_2=4a+n_2$, 并设 $\sigma_a^2=\sigma_n^2$, 求线性最小均方估计.

7.21 已知电压 s, 其概率密度为

$$f(s)=\frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{s^2}{8}\right),$$

现用几次电压测量的线性组合对 a 进行估计, 两只高级仪表的读书为 z_1, z_2, 读数为真实电压上零均值的正态误差, 误差是各自独立的, 且误差的方差 $\sigma^2=4$, 求最佳线性均方估计.

7.22 在某星球上, 有物体自由下落, 在 t 秒内下降距离 $s=g_\mathrm{e}t^2/2$, 现在用一台有噪声的仪器进行观测来估计重力加速度 g_e, 取样值为

$$z_j=\frac{j^2}{2}g_\mathrm{e}+n_j \quad (j=1,2,\cdots),$$

其中 g_e 的均值已扣除, 故四级估计 g_e 的均值为 0, 方差为 $1\mathrm{m}^2/\mathrm{s}$, 噪声样本也是零均值, 其 $R_n(k)=E\left[n_in_{i+k}\right]=1/2^k$, 且与信号不相关.

(1) 一次取样 $z_1=\dfrac{1}{2}g_\mathrm{e}+n_1$, 求最佳线性估计 $\hat{g}_\mathrm{e}$;

(2) 用两次取样求 $\hat{g}_\mathrm{e}$.

7.23 取两个数据样本 $z_1=s+n_1, z_2=2s+n_2$, 已知 $E(s)=E(n)=0, E\left(s^2\right)=3, E\left(n_1^2\right)=E\left(n_2^2\right)=2$, 信号和噪声不相关, 但 $E\left(n_1n_2\right)=1$, 试求其线性最小均方估计.

7.24 设 $x(t)=s\cos\omega_0t+n(t)$, 通过取样对幅度 s 做线性估计. 设 $x(t)$ 在 $\omega_0t=0, \omega_0t=\pi/4$ 处取样, 并设 $E(s)=E\left(n_i\right)=0, E\left(n_1n_2\right)=0, E\left(s^2\right)=\sigma_i^2, E\left(sn_i\right)=0, E\left(n_1^2\right)=E\left(n_2^2\right)=\sigma_n^2$, 求 s 的最佳线性估计 $\hat{s}=h_1x_1+h_2x_2$.

7.25 在一段时间 $[0,T]$ 的两个端点对随机过程 $X(t)$ 进行观测, 得 $X(0)$ 和 $X(T)$, 若用这两个观测数据求 $X(t)$ 在 $[0,T]$ 上的积分值 $I=\int_0^T X(t)\mathrm{d}t$ 的线性最小均方估计, 即求

$$\hat{I}=aX(0)+bX(T).$$

(1) 求 a, b;

(2) 讨论当 T 很小时会出现什么结果? 这个结果合理吗?

7.26 观测某个点目标的等速直线运动. 设观测数据 x 为 $x_k=\theta_0+\theta_1t_k+n_k(k=1,2,\ldots,N)$, 其中 x_k 代表点目标在 t_k 时刻的距离, 参量 θ_0 代表 $t=0$ 时刻的初始距离, 参量 θ_1 代表目标的平均速度, 要求根据对距离 x 的测量做出对 θ_0 和 θ_1 的最小二乘估计.

7.27 测量某地每分钟穿越一段公路的车辆, 用自动计数器作了四次测量, 得到四个值 9505,5610,6050,6220.

(1) 若从前面测量知道每分钟平均车辆数为 5800, 标准偏差在该值上下为 300(每分钟车辆数), 求线性最小均方估计, 仪器误差可以假设平均为零、各样之间相互独立、标准差为 200(每分钟车辆数);

(2) 如果作更多测量, 第五次测得每分钟车辆数为 5880, 比较这两种情形的均方误差 (四个样本和五个样本);

(3) 如第五次测得每分钟车辆数为 6280 重复上面 (2).

7.28　设有随机变量 X 和 Y, 已知 $E\{Y^n\} = m_n (n \geqslant 2), E\{Y\} = 0$ 以及 $X = Y^2$. 若以 Y 的观察值对 X 作线性估计, 求其最佳估计.

7.29　设有零均值实平稳随机过程 $s(t), t$ 是 $[0, T]$ 内的一点. 若已知 $s(0)$ 和 $s(T)$, 利用 $s(0)$ 和 $s(T)$ 求 $s(t)$ 的最佳估计.

7.30　设随机信号 $s(t)$ 加白噪声 $n(t)$ 通过线性滤波器, 信号与噪声的自相关函数分别为 $R_s(\tau) = \dfrac{1}{2}\mathrm{e}^{-|\tau|}$ 和 $R_n(\tau) = \delta(\tau)$, 要求滤波器输出的信号波形均方误差最小. 求滤波器的特性及均方误差.

7.31　设 $z(t) = s(t) + n(t), s(t)$ 和 $n(t)$ 皆为零均值互不相关的实平稳随机过程, 且

$$G_s(\omega) = \frac{1}{1+\omega^2}, \quad G_n(\omega) = 1,$$

若不考虑网络的可实现性, 求最佳滤波器.

7.32　设观测为 $r(u) = a(u) + n(u) (-\infty < u \leqslant t)$, 其中 $a(u)$ 和噪声 $n(u)$ 互不相关, 且均值均为零, 功率谱密度分别为

$$G_a(\omega) = \frac{2K\sigma^2}{\omega^2 + K^2}, \quad G_n(\omega) = N_0\omega^2,$$

要求根据观测 $r(u)$ 估计信号 $a(t)$, 求可实现的维纳滤波器.

补充习题

1　假定要测量某个电压值 θ, 电压 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 一般服从零均值正态分布, 方差为 σ_v^2. 如何根据测量值 z 来估计 θ 的值.

解　这是一个参数估计问题, 解决这一问题有许多办法. 如果 θ 为**随机变量**, 那么, 可以计算后验概率密度 $f(\theta|z)$, 然后求出使 $f(\theta|z)$ 最大的 θ 作为对 θ 的估计值, 即

$$f(\theta|z)|_{\theta=\hat{\theta}_{\text{map}}}=\max.$$

$\hat{\theta}_{\text{map}}$ 称为最大后验概率估计. 这一估计的合理性可以这样来解释: 得到观测 z 后, 计算后验概率密度 $f(\theta|z)$, 如图所示, 很显然, θ 落在以 $\hat{\theta}_{\text{map}}$ 为中心, 以 δ 为半径的领域内的概率要大于落在其他值为中心相同大小领域的概率, 因此有理由认为, 之所以得到观测 z, 是因为 θ 的取值为 $\hat{\theta}_{\text{map}}$, 从后验概率最大这个角度讲是合理的选择. 对于本题描述的估计问题, 假定 $\theta\sim N\left(0,\sigma_\theta^2\right)$, 那么

$$\hat{\theta}_{\text{map}}=\frac{\sigma_\theta^2 z}{\sigma_v^2+\sigma_\theta^2}.$$

如果 θ 为**未知常数**, 这时可以求出似然函数 $f(z;\theta)$, 求出使 $f(z;\theta)$ 最大的 θ 作为对 θ 的估计, 记为 $\hat{\theta}_{\text{ml}}$, 称 $\hat{\theta}_{\text{ml}}$ 为 θ 的最大似然估计, 即

$$f(z;\theta)|_{\theta=\hat{\theta}_{\text{ml}}}=\max.$$

可求得 $\hat{\theta}_{\text{ml}}=z$.

注　*将参数视为随机变量是贝叶斯学派的价值观; 将参数视为未知参数是经典频率学派的价值观.*

2　(贝叶斯估计) 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 $v\sim N\left(0,\sigma_v^2\right)$, 求 A 的最大后验概率和最小均方估计.

解　先求最大后验概率, 因为

$$f(z|A)=\frac{1}{\sqrt{2\pi}\sigma_v}\exp\left\{-\frac{(z-A)^2}{2\sigma_v^2}\right\},$$
$$f(A)=\begin{cases}\dfrac{1}{2A_0} & \left(-A_0\leqslant A\leqslant A_0\right)\\ 0 & (\text{else}),\end{cases}$$
$$f(A|z)=\frac{f(z|A)f(A)}{f(z)},$$

由于 $f(z)$ 与 A 无关, 所以 $f(A|z)$ 的最大值对应的 A 值只取决于 $f(z|A)$ 与 $f(A)$ 的乘积, 当 $-A_0\leqslant z\leqslant A_0$, $f(A|z)$ 的最大值出现在 $A=z$ 处, 所以, $\hat{A}_{\text{map}}=z$; 当 $z>A_0$ 时, $f(A|z)$ 的最大值出现在 $A=A_0$ 处, $\hat{A}_{\text{map}}=A_0$; 当 $z<-A_0$ 时, $f(A|z)$ 的最大值出现在 $A=-A_0$ 处, $\hat{A}_{\text{map}}=-A_0$, 即

$$\hat{A}_{\text{map}}=\begin{cases}-A_0 & \left(z<-A_0\right)\\ z & \left(-A_0\leqslant z\leqslant A_0\right)\\ A_0 & \left(z>A_0\right).\end{cases}$$

再求最小均方估计, 由 $\hat{\theta}_{\text{ms}} = \int_{-\infty}^{+\infty} \theta f(\theta|z)\mathrm{d}\theta = E[\theta|z]$ 得

$$
\begin{aligned}
\hat{A}_{\text{ms}} &= \int_{-\infty}^{+\infty} A f(A|z)\mathrm{d}A \\
&= \int_{-\infty}^{+\infty} A \frac{f(z|A)f(A)}{f(z)}\mathrm{d}A \\
&= \frac{\int_{-\infty}^{+\infty} A f(z|A)f(A)\mathrm{d}A}{\int_{-\infty}^{+\infty} f(z|A)f(A)\mathrm{d}A} \\
&= \frac{\int_{-A_0}^{A_0} \frac{A}{\sqrt{2\pi}\sigma_v} \exp\left[-\frac{(z-A)^2}{2\sigma_v^2}\right] \cdot \frac{1}{2A_0}\mathrm{d}A}{\int_{-A_0}^{A_0} \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left[-\frac{(z-A)^2}{2\sigma_v^2}\right] \cdot \frac{1}{2A_0}\mathrm{d}A} \\
&= \frac{\int_{z-A_0}^{z+A_0} (z-u) \cdot \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u}{\int_{z-A_0}^{z+A_0} \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u} \quad (\text{令} u = z - A) \\
&= \frac{z \int_{z-A_0}^{z+A_0} \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u}{\int_{z-A_0}^{z+A_0} \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u} - \frac{\int_{z-A_0}^{z+A_0} u \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u}{\int_{z-A_0}^{z+A_0} \exp\left[-\frac{u^2}{2\sigma_v^2}\right] \mathrm{d}u} \\
&= z - \frac{2\sigma_v^2 \int_{(x-a)/\sqrt{2}}^{(x+a)/\sqrt{2}} u \exp\left[-u^2\right] \mathrm{d}u}{\sigma_v \int_{x-a}^{x+a} \exp\left[-\frac{u^2}{2}\right] \mathrm{d}u} \quad (\text{令} a = A_0/\sigma_v, x = z/\sigma_v) \\
&= z - \frac{\sigma_v \left\{\exp\left[-(x-a)^2/2\right] - \exp\left[-(x+a)^2/2\right]\right\}}{\sqrt{2\pi}[Q(x-a) - Q(x+a)]},
\end{aligned}
$$

式中 $a = A_0/\sigma_v$ 代表信噪比, $x = z/\sigma_v$ 表示归一化观测值, $Q(\cdot)$ 为标准正态概率密度函数的概率右尾函数:

$$
Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp\left(-u^2/2\right) \mathrm{d}u.
$$

下图给出了 $\hat{A}_{\text{map}}$ 和 $\hat{A}_{\text{ms}}$ 对 z 的关系曲线, 并且由图可以看出, $\hat{A}_{\text{map}}$ 和 $\hat{A}_{\text{ms}}$ 并不相等, 而两个估计量都是非线性的, 所以是非线性估计.

---

3 (贝叶斯估计) 高斯白噪声中的恒定电平估计——高斯先验分布. 设有 N 次独立观测 $z_i = A + v_i, i = 1, 2, \ldots, N$, 其中 $v_i \sim N\left(0, \sigma^2\right), A \sim N\left(\mu_A, \sigma_A^2\right)$, 求 A 的估计.

解 先求后验概率密度:

$$
\begin{aligned}
f(A|\boldsymbol{z}) &= \frac{f(\boldsymbol{z}|A)f(A)}{f(\boldsymbol{z})} \quad (f(z|A)\text{是先验概率, 就是似然, 由于独立同分布}, z_i\text{连乘.}) \\
&= \frac{f(\boldsymbol{z}|A)f(A)}{\int_{-\infty}^{+\infty} f(\boldsymbol{z}|A)f(A)\mathrm{d}A} \\
&= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N (z_i - A)^2\right] \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]}{\int_{-\infty}^{+\infty} \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N (z_i - A)^2\right] \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]\mathrm{d}A} \\
&= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N (z_i^2 - 2z_iA + A^2)\right] \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]}{\frac{1}{(2\pi\sigma^2)^{N/2}} \frac{1}{\sqrt{2\pi\sigma_A^2}} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^N (z_i^2 - 2z_iA + A^2)\right] \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]\mathrm{d}A} \qquad (2) \\
&= \frac{\exp\left(\sum_{i=1}^N z_i^2\right) \exp\left[-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{z})\right] \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]}{\exp\left(\sum_{i=1}^N z_i^2\right) \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{z})\right] \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]\mathrm{d}A} \\
&= \frac{\exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}(NA^2 - 2NA\bar{z}) + \frac{1}{\sigma_A^2}(A-\mu_A)^2\right]\right\}}{\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}(NA^2 - 2NA\bar{z}) + \frac{1}{\sigma_A^2}(A-\mu_A)^2\right]\right\}\mathrm{d}A} \\
&= \frac{\exp\left\{-\frac{1}{2}W(A)\right\}}{\int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2}W(A)\right\}\mathrm{d}A},
\end{aligned}
$$

上式中 $\bar{z} = \frac{1}{N}\sum_{i=1}^N z_i$ 为样本均值, $W(A)$ 为

$$
W(A) = \frac{1}{\sigma^2}(NA^2 - 2NA\bar{z}) + \frac{1}{\sigma_A^2}(A-\mu_A)^2.
$$

注意到 $f(A|z)$ 的分母与 A 无关, $W(A)$ 是 A 的二次型, 经过配方可以把上式写成

$$
W(A) = \frac{1}{\sigma_{A|z}^2}(A-\mu_{A|z})^2 - \frac{\mu_{A|z}^2}{\sigma_{A|z}^2} + \frac{\mu_A^2}{\sigma_A^2}, \qquad (3)
$$

其中

$$
\begin{aligned}
\sigma_{A|z}^2 &= \left(\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}\right)^{-1}, \\
\mu_{A|z} &= \left(\frac{N}{\sigma^2}\bar{z} + \frac{\mu_A}{\sigma_A^2}\right)\sigma_{A|z}^2.
\end{aligned}
$$

将 (3) 式代入 (2) 式中, 得到

$$
\begin{aligned}
f(A|\boldsymbol{z}) &= \frac{\exp\left[-\frac{1}{2\sigma_{A|z}^2}(A-\mu_{A|z})^2\right] \exp\left[-\frac{1}{2}\left(\frac{\mu_A^2}{\sigma_A^2} - \frac{\mu_{A|z}^2}{\sigma_{A|z}^2}\right)\right]}{\int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2\sigma_{A|z}^2}(A-\mu_{A|z})^2\right] \exp\left[-\frac{1}{2}\left(\frac{\mu_A^2}{\sigma_A^2} - \frac{\mu_{A|z}^2}{\sigma_{A|z}^2}\right)\right]\mathrm{d}A} \qquad (4) \\
&= \frac{1}{\sqrt{2\pi\sigma_{A|z}^2}} \exp\left[-\frac{1}{2\sigma_{A|z}^2}(A-\mu_{A|z})^2\right].
\end{aligned}
$$

由 (4) 式可以看出, 后验概率密度是高斯的. 由于**最小均方估计为被估计量的条件均值**, 所以

$$\hat{A}_{\mathrm{ms}} = \mu_{A|z} = \left(\frac{N}{\sigma^2}\bar{z} + \frac{\mu_A}{\sigma_A^2}\right)\sigma_{A|z}^2 = \frac{\frac{N}{\sigma^2}\bar{z} + \frac{\mu_A}{\sigma_A^2}}{\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}} \quad (\text{加权平均}),$$

令 $k = \dfrac{\sigma_A^2}{\sigma_A^2 + \sigma^2/N}$, 则

$$\hat{A}_{\mathrm{ms}} = k\bar{z} + (1-k)\mu_A.$$

另外, 由于最大后验概率估计是使后验概率密度最大所对应的 A 值, 因此由 (4) 式可得

$$\hat{A}_{\mathrm{map}} = \mu_{A|z} = \hat{A}_{\mathrm{ms}},$$

即最大后验概率与最小均方估计相等.

注 *此题对应习题 7.6.*

回顾 贝叶斯估计

---

4 (最大似然估计) 高斯白噪声中的恒定电平估计——未知参数. 设有 N 次独立观测 $z_i = A + v_i (i = 1, 2, \ldots, N)$, 其中 $v \sim N\left(0, \sigma^2\right)$, A 为未知参数, σ^2 已知, 求 A 的最大似然估计.

解 先求似然函数:

$$f(\boldsymbol{z}; A) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(z_i - A\right)^2\right] \quad (\text{视}A\text{为未知参数}, z_i\text{由于独立同分布连乘.}),$$

$$\ln f(\boldsymbol{z}; A) = -\frac{N}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(z_i - A\right)^2,$$

$$\frac{\partial \ln f(\boldsymbol{z}; A)}{\partial A} = \frac{1}{\sigma^2}\sum_{i=1}^{N}\left(z_i - A\right) = \frac{N}{\sigma^2}\left(\frac{1}{N}\sum_{i=1}^{N} z_i - A\right),$$

根据最大似然方程, 得

$$\hat{A}_{\mathrm{ml}} = \bar{z} = \frac{1}{N}\sum_{i=1}^{N} z_i, \tag{5}$$

$\bar{z}$ 为观测的样本均值, 由于

$$\frac{\partial^2 \ln f(z; A)}{\partial A^2} = -\frac{N}{\sigma^2} < 0,$$

所以 (5) 式求得的是极大值, 也就是 A 的最大似然估计.

---

5 (最大似然估计) 设有 N 次独立观测 $z_i = v_i (i = 1, 2, \ldots, N)$, 其中 $v_i \sim N\left(0, \sigma^2\right)$, 求噪声方差 σ^2 的最大似然估计.

解 因为

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

$$\ln f\left(\boldsymbol{z}; \sigma^2\right) = -\frac{N}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{N} z_i^2,$$

$$\frac{\partial \ln f\left(\boldsymbol{z}; \sigma^2\right)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{N} z_i^2 = -\frac{N}{2\sigma^4}\left(\sigma^2 - \frac{1}{N}\sum_{i=1}^{N} z_i^2\right),$$

令上式等于零, 得

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

很容易验证

$$\left.\frac{\partial^2 \ln f\left(\boldsymbol{z};\sigma^2\right)}{\partial\left(\sigma^2\right)^2}\right|_{\sigma^2=\frac{1}{N}\sum_{i=1}^{N}z_i^2}<0,$$

所以 $\hat{\sigma}_{\mathrm{ml}}^2=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}z_i^2$ 是最大似然估计.

---

6 (最大似然估计) 高斯白噪声中的恒定电平估计——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 $v\sim N\left(0,\sigma^2\right)$, σ^2、A 为未知参数, 求 A 和 σ^2 的最大似然估计.

解 这是一个多参量的同时估计, 最大似然估计方式不变, 只是标量变成了矢量 $\boldsymbol{\theta}$. 在本题中, $\boldsymbol{\theta}=\left[A\ \sigma^2\right]^{\mathrm{T}}$.

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

$$\ln f(\boldsymbol{z};\boldsymbol{\theta})=-\frac{N}{2}\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(z_i-A\right)^2,$$

$$\frac{\partial\ln f(\boldsymbol{z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=\begin{bmatrix}\dfrac{N}{\sigma^2}\left(\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}z_i-A\right)\\ -\dfrac{N}{2\sigma^4}\left[\sigma^2-\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\left(z_i-A\right)^2\right]\end{bmatrix},$$

令 $\left.\dfrac{\partial\ln f(\boldsymbol{z};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right|_{\theta=\hat{\theta}_{\mathrm{ml}}}=0$ 可求得最大似然估计为

$$\hat{\boldsymbol{\theta}}_{\mathrm{ml}}=\begin{bmatrix}\hat{A}_{\mathrm{ml}}\\ \hat{\sigma}^2\end{bmatrix}=\begin{bmatrix}\bar{z}\\ \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\left(z_i-\bar{z}\right)^2\end{bmatrix}.$$

注意 这里参数 A 的估计与题 4 一致, 在噪声方差已知和未知的情况下是相同的; 而参数 σ^2 的估计与题 5 不同.

---

7 (最大似然估计) 正弦信号相位的估计. 希望估计高斯白噪声中正弦信号的相位, 即 $z(n)=A\cos\left(2\pi f_0 n+\phi\right)+v(n)(n=0,1,\ldots,N-1)$, 假定幅度 A 和频率 f_0 是已知的, $v(n)$ 是高斯白噪声序列, 均值为零, 方差为 σ^2, 求相位 ϕ 的最大似然估计.

解 似然函数为

$$f(\boldsymbol{z};\phi)=\frac{1}{\left(2\pi\sigma^2\right)^{N/2}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left[z(n)-A\cos\left(2\pi f_0 n+\phi\right)\right]^2\right\},$$

对数似然函数为

$$\ln f(\boldsymbol{z};\phi)=-\frac{N}{2}\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left[z(n)-A\cos\left(2\pi f_0 n+\phi\right)\right]^2,$$

$$\frac{\partial\ln f(\boldsymbol{z};\phi)}{\partial\phi}=-\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left[z(n)-A\cos\left(2\pi f_0 n+\phi\right)\right]A\sin\left(2\pi f_0 n+\phi\right),$$

令上式等于零, 得

$$\sum_{n=0}^{N-1} z(n)\sin\left(2\pi f_0 n+\hat{\phi}_{\mathrm{ml}}\right)=A\sum_{n=0}^{N-1}\cos\left(2\pi f_0 n+\hat{\phi}_{\mathrm{ml}}\right)\sin\left(2\pi f_0 n+\hat{\phi}_{\mathrm{ml}}\right).$$

当 f_0 不在 0 或 1/2 附近时, 上式右边近似为零. 因此似然估计近似满足

$$\sum_{n=0}^{N-1} z(n)\sin\left(2\pi f_0 n+\hat{\phi}_{m1}\right)=0,$$

展开上式, 得

$$\sum_{n=0}^{N-1} z(n)\sin 2\pi f_0 n\cos\hat{\phi}_{\mathrm{ml}}=-\sum_{n=0}^{N-1} z(n)\cos 2\pi f_0 n\sin\hat{\phi}_{\mathrm{m1}},$$

$$\hat{\phi}_{\mathrm{ml}}=-\arctan\frac{\sum_{n=0}^{N-1} z(n)\sin 2\pi f_0 n}{\sum_{n=0}^{N-1} z(n)\cos 2\pi f_0 n}.$$