

统计信号处理

2024 期末考试解答

整理：scpsyl

2026 年 1 月 12 日

注：本试题及解答为个人自行总结，仅供学习交流使用，版权归原作者所有。

1 各类判决准则综述

阐述各类判决准则的由来（文字）、判决规定、门限确定、ROC 曲线上的解（画图）。包括：Bayes 准则、极大极小准则、Neyman-Pearson 准则、最小错误概率准则、最大似然准则、最大后验概率准则。

1.1 Bayes 准则

由来：在先验概率 $(P(H_0), P(H_1))$ 和代价因子 C_{ij} 已知的情况下，选择使“期望损失”（平均风险）最小的判决规则。

判决规定：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_B,$$

门限

$$\lambda_B = \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})}.$$

ROC 曲线上的解：曲线上任意 $\forall \lambda_T = \lambda_B$ 的点，对应 P_F 为虚警概率， P_D 为检测概率。

1.2 极大极小准则

由来：在事先不知道先验概率的情况下，选择使得在最坏情况下（最大风险）风险最小的判决规则。

$$\min_{\delta} \max_{P(H_0)} R(\delta; P(H_0))$$

极大极小方程：

$$C_{00}P(D_0|H_0) + C_{10}P(D_1|H_0) = C_{01}P(D_0|H_1) + C_{11}P(D_1|H_1)$$

判决准则：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{MM}$$

门限：

$$\lambda_{MM} = \frac{\xi_{MM}(C_{10} - C_{00})}{(1 - \xi_{MM})(C_{01} - C_{11})}$$

其中 ξ_{MM} 由极大极小方程求出。

ROC 曲线上的解：

由极大极小方程可得出 P_D 与 P_F 的线性关系：

$$P_D = \frac{C_{01} - C_{00}}{C_{01} - C_{11}} P_F + \frac{C_{10} - C_{00}}{C_{01} - C_{11}}$$

曲线上的解为 ROC 曲线上与上述直线的交点。

1.3 Neyman-Pearson 准则

由来：在未知先验概率和代价因子情况下，希望虚警概率尽可能小，检测概率尽可能大，通过固定 P_F 为一个较小的值的情况下，求条件极值。

判决规定：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{NP}$$

门限：

$$\lambda_{NP} \text{ 由 } P_F = \int_{\lambda_{NP}}^{\infty} p(\gamma|H_0)d\gamma = \alpha \text{ 确定}$$

ROC 曲线上的解：

曲线上给定 $P_F = \alpha$ 时对应的 P_D 值。

1.4 最小错误概率准则

由来：在已知先验概率和代价因子情况下，希望平均错误概率 $P_e = P(H_0)P_F + P(H_1)P_M$ 最小。

判决规定：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_e$$

门限：

$$\lambda_e = \frac{P(H_0)}{P(H_1)}$$

ROC 曲线上的解：

曲线上对应斜率为 $\frac{P(H_0)}{P(H_1)}$ 的切线与曲线的切点的横纵坐标对应的 P_F 和 P_D 。

1.5 最大似然准则

由来：在已知先验概率和代价因子情况下，选择“使观测最可能出现”的假设，即似然函数最大的假设。

判决规定：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_{ML}$$

门限：

$$\lambda_{ML} = 1$$

ROC 曲线上的解：

曲线上对应斜率为 1 的切线与曲线的切点的横纵坐标对应的 P_F 和 P_D 。

1.6 最大后验概率准则

由来：在得到观测之后，取更大概率的假设为最终判决结果，即满足最大后验概率的假设。

$$\frac{P(H_1|\mathbf{z})}{P(H_0|\mathbf{z})} \underset{H_0}{\overset{H_1}{\geq}} 1$$

判决规定：

$$\lambda(\mathbf{z}) = \frac{p(\mathbf{z}|H_1)}{p(\mathbf{z}|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \lambda_{MAP}$$

门限：

$$\lambda_{MAP} = \frac{P(H_0)}{P(H_1)}$$

ROC 曲线上的解：

曲线上对应斜率为 $\frac{P(H_0)}{P(H_1)}$ 的切线与曲线的切点的横纵坐标对应的 P_F 和 P_D 。

2 简答题

2.1 确知信号在色高斯信道和白高斯信道中，最小平均错误概率准则下的最优接收机结构？他们的异同点？

解答：

(1) 白高斯信道 判决准则：

$$\int_0^T [s_1(t) - s_0(t)] z(t) dt \underset{H_0}{\overset{H_1}{\geq}} V_T$$

门限设计：

$$V_T = \frac{N_0}{2} \mathcal{L}_T + \frac{1}{2}(E_1 - E_0), \quad \mathcal{L}_T \text{ 由具体性能指标确定}$$

相关接收机结构：

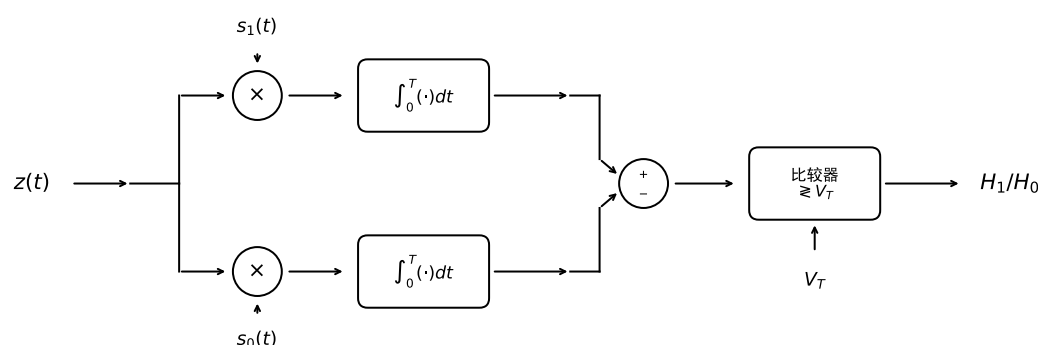


图 1: 白高斯信道 - 相关接收机

匹配滤波接收机结构：

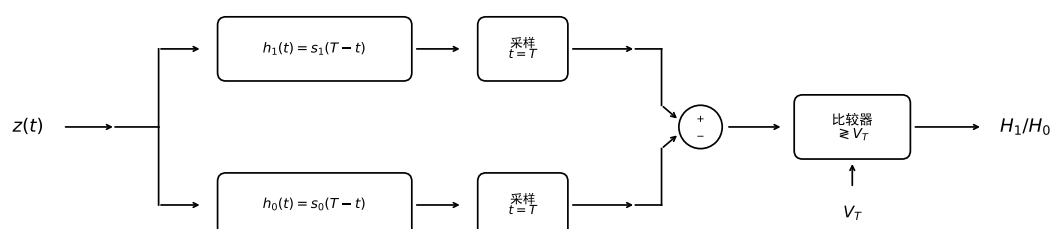


图 2: 白高斯信道 - 匹配滤波接收机

(2) 色高斯信道 需要设计“白化滤波器” $h_w(t, \tau)$ ，使 $z(t)$ 输入后输出噪声“白化”。需使满足：

$$R_{n_w}(t, \tau) = \mathbb{E}[n_w(t) n_w(\tau)] = \delta(t - \tau)$$

对白化滤波器核的展开：

$$h_w(t, \tau) = \sum_{k=1}^{\infty} \frac{1}{\sqrt{\lambda_k}} g_k(t) g_k(\tau), \quad g_k(t) \text{ 为基函数}$$

相关接收机结构（含白化滤波器）：

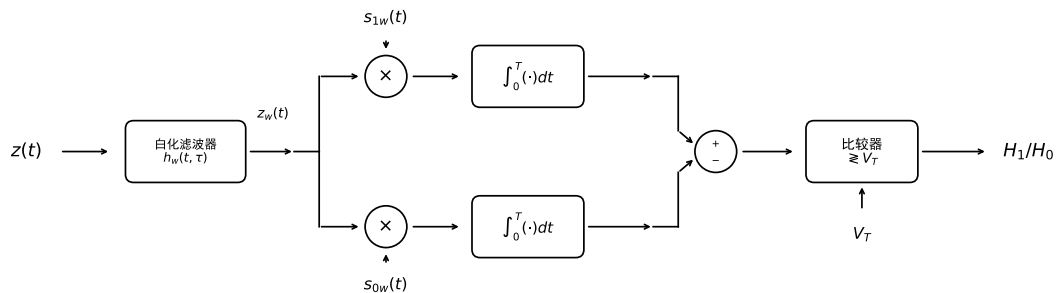


图 3: 色高斯信道 - 相关接收机（含白化滤波器）

匹配滤波接收机结构（含白化滤波器）：

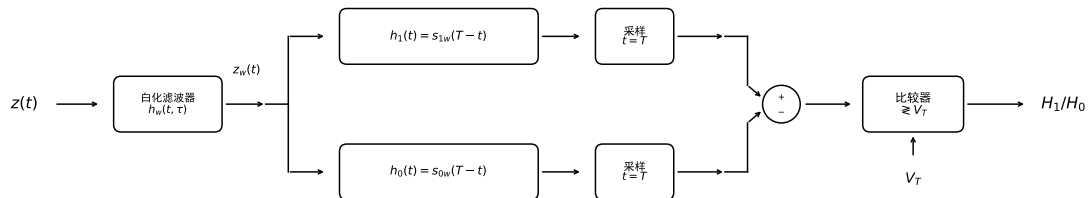


图 4: 色高斯信道 - 匹配滤波接收机（含白化滤波器）

(3) 异同点

- 相同点：都源于似然比/假设检验的思想，本质都是将观测投影到“最能区分假设”的方向上。（等先验下，门限常取中点。）
- 不同点：白噪声可直接做（概率意义下的）匹配滤波/相关器；色噪声由于噪声频谱不均、存在相关性，需要先做归一化/加权（“白化”），之后再匹配滤波/相关器。

2.2 线性最小均方估计、最小二乘估计、加权最小二乘估计，他们的异同点？

(1) 线性最小均方估计 (LMS) 已知随机参数 $\boldsymbol{\theta}$ 与观测 \mathbf{z} 的一、二阶统计量 (如 $\mathbb{E}[\boldsymbol{\theta}]$ 、 $\mathbb{E}[\mathbf{z}]$ 、 $\text{Var}(\boldsymbol{\theta})$ 、 $\text{Var}(\mathbf{z})$ 、 $\text{Cov}(\boldsymbol{\theta}, \mathbf{z})$ 等)，考虑线性估计器

$$\hat{\boldsymbol{\theta}} = \mathbf{A}\mathbf{z} + \mathbf{B}.$$

令均方误差最小：

$$\min_{\mathbf{A}, \mathbf{B}} \mathbb{E}[(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^\top (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})] \iff \min_{\mathbf{A}, \mathbf{B}} \mathbb{E}[(\boldsymbol{\theta} - \mathbf{A}\mathbf{z} - \mathbf{B})^\top (\boldsymbol{\theta} - \mathbf{A}\mathbf{z} - \mathbf{B})].$$

其解为

$$\mathbf{A} = \text{Cov}(\boldsymbol{\theta}, \mathbf{z}) [\text{Var}(\mathbf{z})]^{-1}, \quad \mathbf{B} = \mathbb{E}[\boldsymbol{\theta}] - \text{Cov}(\boldsymbol{\theta}, \mathbf{z}) [\text{Var}(\mathbf{z})]^{-1} \mathbb{E}[\mathbf{z}].$$

因此

$$\hat{\boldsymbol{\theta}}_{\text{LMS}} = \mathbb{E}[\boldsymbol{\theta}] + \text{Cov}(\boldsymbol{\theta}, \mathbf{z}) [\text{Var}(\mathbf{z})]^{-1} (\mathbf{z} - \mathbb{E}[\mathbf{z}]).$$

性质

1. 无偏性：

$$\mathbb{E}[\hat{\boldsymbol{\theta}}_{\text{LMS}}] = \mathbb{E}[\boldsymbol{\theta}].$$

2. 误差协方差矩阵 (令 $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$):

$$\begin{aligned} \text{Var}(\tilde{\boldsymbol{\theta}}) &= \mathbb{E}[\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^\top] \\ &= \text{Var}(\boldsymbol{\theta}) - \text{Cov}(\boldsymbol{\theta}, \mathbf{z}) [\text{Var}(\mathbf{z})]^{-1} \text{Cov}(\mathbf{z}, \boldsymbol{\theta}). \end{aligned}$$

3. 线性最小均方估计误差与观测在二阶意义下 “不相关” ($\text{Cov}(\tilde{\boldsymbol{\theta}}, \mathbf{z}) = \mathbf{0}$)。

(2) 最小二乘估计 (LS) 对参数 $\boldsymbol{\theta}$ 的线性观测模型：

$$\mathbf{z} = \mathbf{C}\boldsymbol{\theta} + \mathbf{n}.$$

最小二乘准则：

$$\min_{\hat{\boldsymbol{\theta}}} (\mathbf{z} - \mathbf{C}\hat{\boldsymbol{\theta}})^\top (\mathbf{z} - \mathbf{C}\hat{\boldsymbol{\theta}}),$$

其解 (\mathbf{C} 满列秩)：

$$\hat{\boldsymbol{\theta}}_{\text{LS}} = (\mathbf{C}^\top \mathbf{C})^{-1} \mathbf{C}^\top \mathbf{z}.$$

性质

1. $\hat{\boldsymbol{\theta}}_{\text{LS}}$ 是 \mathbf{z} 的线性函数（线性估计器）。
2. 若 $\mathbb{E}[\mathbf{n}] = \mathbf{0}$ ，则 $\hat{\boldsymbol{\theta}}_{\text{LS}}$ 为无偏估计。
3. 估计误差协方差（记 $\mathbf{R}_n = \text{Var}(\mathbf{n})$ ）：

$$\mathbb{E}[\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^T] = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{R}_n\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}.$$

（若 $\mathbb{E}[\mathbf{n}] = \mathbf{0}$ ，则 $\mathbf{R}_n = \text{Var}(\mathbf{n})$ 。）

(3) 加权最小二乘估计（LSW） 当观测噪声相关/异方差时，引入权矩阵 \mathbf{W} （对称正定）：

$$\min_{\hat{\boldsymbol{\theta}}} (\mathbf{z} - \mathbf{C}\hat{\boldsymbol{\theta}})^T \mathbf{W} (\mathbf{z} - \mathbf{C}\hat{\boldsymbol{\theta}}).$$

其解：

$$\hat{\boldsymbol{\theta}}_{\text{WLS}} = (\mathbf{C}^T \mathbf{W} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{W} \mathbf{z}.$$

性质

1. $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ 是 \mathbf{z} 的线性估计。
2. 若 $\mathbb{E}[\mathbf{n}] = \mathbf{0}$ ，则 $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ 为无偏估计。
3. 估计误差协方差：

$$\mathbb{E}[\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^T] = (\mathbf{C}^T \mathbf{W} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{W} \mathbf{R}_n \mathbf{W} \mathbf{C} (\mathbf{C}^T \mathbf{W} \mathbf{C})^{-1}.$$

其有下界

$$\mathbb{E}[\tilde{\boldsymbol{\theta}}\tilde{\boldsymbol{\theta}}^T] \succeq (\mathbf{C}^T \mathbf{R}_n^{-1} \mathbf{C})^{-1},$$

当且仅当取 $\mathbf{W} = \mathbf{R}_n^{-1}$ 时取等号；此时 \mathbf{W} 为最优加权矩阵，对应估计方差（协方差）最小。

(4) 异同点比较

• 相同点：

1. 都是基于“平方误差/二次型”构造的最优问题，得到闭式线性解
2. 都可视为“相关/投影”的思想：把观测投影到模型张成的子空间

• 不同点：

1. LMS 基于随机参数与观测的统计特性（均值、方差、协方差等）进行估计，适用于参数与观测均为随机变量的情形。
2. LS 基于确定性参数的线性观测模型，通过最小化观测误差平方和进行估计，适用于参数为确定性未知量的情形。
3. LSW 是 LS 的推广，适用于观测噪声相关或异方差的情形，通过引入权矩阵对不同观测赋予不同权重，以提高估计性能。

2.3 Kalman 滤波和 Wiener 滤波的异同点？

解答：共同点：两者都是根据线性最小均方误差（LMS）准则来进行估计的，即都致力于使估值的均方误差达到最小。

不同点：

- Wiener 滤波要求观测过程是平稳的；Kalman 滤波不要求平稳性，可适用于时变系统。
- Wiener 滤波基于传输函数或冲激响应，通过对过去观测数据的线性加权（卷积积分）进行估计；Kalman 滤波建立状态空间模型，采用状态变量法在时域内直接求解。
- Wiener 滤波通常适用于“离线”处理。每获取新数据需重新计算相关矩阵及其逆矩阵，运算量大且需存储所有历史数据，难以实时实现；Kalman 滤波采用递推算法。根据新的观测数据不断修正估计量，存储量小，运算效率高，具有很好的实时性。
- Wiener 滤波分析单个信号尚可，处理多信号（多维信号）非常困难；Kalman 滤波能够很好地解决多信号（多维）的估计问题。

2.4 参量检测、非参量检测、鲁棒检测的异同点？

解答：

相同点：都是在给定观测 z 的情况下，对假设 H_0 和 H_1 进行判决：选取判决域或判决规则；评价指标一致：错误概率 (P_F, P_M, P_e) 或风险/代价，并在某个约束下优化。

不同点：

- 参量检测：假设检验中， H_0, H_1 下的观测分布形式已知（属于某参量族），参数已知或仅含少量未知参量；典型准则为 Neyman-Pearson 准则、Bayes 准则（常导出似然比检验）。

- 非参量检测：不对 H_0, H_1 下的分布形式作具体假设（或仅作很弱的结构性假设），主要基于观测样本的秩、符号、经验分布等构造统计量进行判决；典型方法如秩检验、符号检验。
- 鲁棒检测： H_0, H_1 下的分布不精确已知，而属于某个不确定集合（参数区间/分布扰动/邻域模型等），按极大极小准则或最不利分布设计判决规则，使最坏情况下的风险/性能受控（代价是规则更保守）。

3 简答题

Bayes 风险 R_B 与先验概率 $P(H_0)$ 之间，一定是单凸曲线的形式吗？为什么？

解答：

令 $\pi = P(H_0)$ ，则 $P(H_1) = 1 - \pi$ 。对任意固定判决规则 δ ，其平均风险可写为

$$R(\delta; \pi) = \pi R_0(\delta) + (1 - \pi) R_1(\delta),$$

其中

$$R_j(\delta) = \sum_i C_{ij} P(\delta = i | H_j), \quad j \in \{0, 1\}.$$

由于 C_{ij} 与先验无关，且对固定 δ ， $P(\delta = i | H_j)$ 与 π 无关，故 $R_0(\delta), R_1(\delta)$ 均为常数，从而 $R(\delta; \pi)$ 关于 π 为仿射函数（直线）。

Bayes 最小风险定义为

$$R_B(\pi) = \min_{\delta} R(\delta; \pi).$$

取任意 $\pi_1, \pi_2 \in [0, 1]$ 与 $\lambda \in [0, 1]$ ，令

$$\pi = \lambda \pi_1 + (1 - \lambda) \pi_2.$$

则对任意 δ ，

$$R(\delta; \pi) = \lambda R(\delta; \pi_1) + (1 - \lambda) R(\delta; \pi_2).$$

对 δ 取最小并利用 \min 的基本不等式 $\min_x (\lambda a_x + (1 - \lambda) b_x) \geq \lambda \min_x a_x + (1 - \lambda) \min_x b_x$ ，得

$$R_B(\pi) = \min_{\delta} R(\delta; \pi) \geq \lambda \min_{\delta} R(\delta; \pi_1) + (1 - \lambda) \min_{\delta} R(\delta; \pi_2) = \lambda R_B(\pi_1) + (1 - \lambda) R_B(\pi_2).$$

因此 $R_B(\pi)$ 为凹函数（上凸）。

任取区间内部点 $\pi_0 \in (0, 1)$ ，取足够小的 $\varepsilon > 0$ 使得 $\pi_0 \pm \varepsilon \in (0, 1)$ 。由凹性，取 $\lambda = \frac{1}{2}$ 、 $\pi_1 = \pi_0 - \varepsilon$ 、 $\pi_2 = \pi_0 + \varepsilon$ ，则

$$R_B(\pi_0) = R_B\left(\frac{1}{2}(\pi_0 - \varepsilon) + \frac{1}{2}(\pi_0 + \varepsilon)\right) \geq \frac{R_B(\pi_0 - \varepsilon) + R_B(\pi_0 + \varepsilon)}{2}.$$

上式蕴含至少有一侧满足

$$R_B(\pi_0) \geq R_B(\pi_0 - \varepsilon) \quad \text{或} \quad R_B(\pi_0) \geq R_B(\pi_0 + \varepsilon).$$

因此 π_0 不可能同时严格小于其左右邻域的函数值，即 π_0 不可能是严格局部极小点。

因此， $R_B(\pi)$ 随 $\pi = P(H_0)$ 的变化并不呈现包含局部极小值的“单谷/多谷单凸曲线”；相反，它是凹（上凸）的下包络形状，区间内部不产生局部极小值（最多出现平台）。

4 观测模型的参数估计

给定观测模型

$$z_i = a + n_i, \quad i = 1, \dots, N,$$

其中 $a \sim \mathcal{N}(0, \sigma_a^2)$, $n_i \sim \mathcal{N}(0, \sigma_n^2)$, 统计独立。

1. 求最大后验估计 \hat{a}_{MAP}
2. 求最小均方估计 \hat{a}_{MS}
3. 计算相应的估计均方误差 (MSE)
4. 判断估计量是否无偏、是否有效, 并与 C-R 下界比较

解答：

由独立性，

$$p(\mathbf{z} | a) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(z_i - a)^2}{2\sigma_n^2}\right), \quad p(a) = \frac{1}{\sqrt{2\pi\sigma_a^2}} \exp\left(-\frac{a^2}{2\sigma_a^2}\right),$$

其中 $\mathbf{z} = [z_1, \dots, z_N]^T$ 。

则

$$\begin{aligned} p(a|\mathbf{z}) &= \frac{p(\mathbf{z} | a)p(a)}{p(\mathbf{z})} = A(z) \exp\left(-\frac{1}{2} \left[\frac{1}{\sigma_n^2} \sum_{i=1}^N (z_i - a)^2 + \frac{a^2}{\sigma_a^2} \right]\right), \\ &= B(z) \exp\left(-\frac{1}{2} \left[\left(\frac{N}{\sigma_n^2} + \frac{1}{\sigma_a^2} \right) a^2 - \frac{2}{\sigma_n^2} \sum_{i=1}^N z_i a \right]\right) = C(z) \exp\left(-\frac{1}{2\sigma_{a|z}^2} (a - \mu_{a|z})^2\right) \end{aligned}$$

其中

$$\sigma_{a|z}^2 = \frac{\sigma_a^2 \sigma_n^2}{N\sigma_a^2 + \sigma_n^2}, \quad \mu_{a|z} = \frac{N\sigma_a^2}{N\sigma_a^2 + \sigma_n^2} \bar{z}, \quad \bar{z} = \frac{1}{N} \sum_{i=1}^N z_i.$$

(1) \hat{a}_{MAP}

$$\hat{a}_{\text{MAP}} = \arg \max_a p(a | \mathbf{z}) = \arg \max_a \ln p(\mathbf{z} | a) + \ln p(a).$$

略去与 a 无关常数，记 $\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i$ ，则

$$\ln p(\mathbf{z} | a) + \ln p(a) = -\frac{1}{2\sigma_n^2} \sum_{i=1}^N (z_i - a)^2 - \frac{1}{2\sigma_a^2} a^2 + \text{const.}$$

对 a 求导并令零：

$$\begin{aligned} \frac{\partial}{\partial a} \left[-\frac{1}{2\sigma_n^2} \sum_{i=1}^N (z_i - a)^2 - \frac{1}{2\sigma_a^2} a^2 \right] &= 0 \\ \Rightarrow \frac{1}{\sigma_n^2} \sum_{i=1}^N (z_i - a) - \frac{a}{\sigma_a^2} &= 0 \Rightarrow \frac{N}{\sigma_n^2} \bar{z} - \left(\frac{N}{\sigma_n^2} + \frac{1}{\sigma_a^2} \right) a = 0. \end{aligned}$$

因此

$$\hat{a}_{\text{MAP}} = \frac{\frac{N}{\sigma_n^2}}{\frac{N}{\sigma_n^2} + \frac{1}{\sigma_a^2}} \bar{z} = \frac{N\sigma_a^2}{N\sigma_a^2 + \sigma_n^2} \bar{z}.$$

(2) \hat{a}_{MS}

$$\hat{a}_{\text{MS}} = \mathbb{E}[a | \mathbf{z}].$$

由于先验 a 与似然均为高斯，后验 $p(a | \mathbf{z})$ 仍为高斯分布，其均值等于 MAP 解（高斯分布的均值 = 众数）：

$$\hat{a}_{\text{MS}} = \hat{a}_{\text{MAP}} = \frac{N\sigma_a^2}{N\sigma_a^2 + \sigma_n^2} \bar{z}.$$

(3)MSE 由于 $\hat{a}_{\text{MS}} = \mathbb{E}[a | \mathbf{z}]$, 其 MSE 等于后验方差的期望, 而此处后验方差与 \mathbf{z} 无关, 为常数:

$$\text{MSE}(\hat{a}_{\text{MS}}) = \mathbb{E}[(a - \hat{a}_{\text{MS}})^2] = \mathbb{E}[\text{Var}(a | \mathbf{z})] = \sigma_{a|\mathbf{z}}^2 = \frac{\sigma_a^2 \sigma_n^2}{N\sigma_a^2 + \sigma_n^2}.$$

同理因为 $\hat{a}_{\text{MAP}} = \hat{a}_{\text{MS}}$, 故

$$\text{MSE}(\hat{a}_{\text{MAP}}) = \text{MSE}(\hat{a}_{\text{MS}}) = \frac{\sigma_a^2 \sigma_n^2}{N\sigma_a^2 + \sigma_n^2} \triangleq \sigma^2$$

(4) 无偏性与有效性 (a) 无偏性

参数 a 的无偏性定义为

$$\mathbb{E}[\hat{a}] = \mathbb{E}[a].$$

$\mathbb{E}[a] = 0$ 。又

$$\mathbb{E}[\bar{z}] = \mathbb{E}[a + \bar{n}] = \mathbb{E}[a] + \mathbb{E}[\bar{n}] = 0,$$

因此

$$\mathbb{E}[\hat{a}_{\text{MAP}}] = \frac{N\sigma_a^2}{N\sigma_a^2 + \sigma_n^2} \mathbb{E}[\bar{z}] = 0 = \mathbb{E}[a].$$

故

$$\hat{a}_{\text{MAP}} (\text{亦即} \hat{a}_{\text{MS}}) \text{ 是无偏的。}$$

(b) 有效性

由结论: 具有高斯型分布的后验概率密度的估值是有效估值, 其估值方差达到 C-R 界限。证明如下:

C-R 界限为:

$$\mathbb{E}[(\hat{a} - a)^2] \geq \frac{1}{\mathbb{E}\left[\left(\frac{\partial}{\partial a} \ln p(\mathbf{z}, a)\right)^2\right]} = -\frac{1}{\mathbb{E}\left[\frac{\partial^2}{\partial a^2} \ln p(\mathbf{z}, a)\right]},$$

其中期望 $\mathbb{E}[\cdot]$ 是对联合分布 $p(\mathbf{z}, a) = p(\mathbf{z} | a)p(a)$ 的期望。

本题中,

$$\begin{aligned} \frac{\partial}{\partial a} \ln p(\mathbf{z} | a) &= -\frac{1}{\sigma^2} \left[a - \frac{N\sigma^2}{\sigma_n^2} \bar{z} \right], \\ \frac{\partial^2}{\partial a^2} \ln p(\mathbf{z} | a) &= -\frac{1}{\sigma^2}, \end{aligned}$$

则 C-R 界限为

$$\mathbb{E}[(\hat{a} - a)^2] \geq \sigma^2.$$

由此, 二者都达到了 C-R 界限, 故为有效估计量。

5 二维成像/模糊观测模型

给定二维成像/模糊观测模型

$$v(k, \ell) = \sum_i \sum_j h(k, \ell; i, j) u(i, j) + n(k, \ell), \quad 1 \leq k, \ell \leq N.$$

将其写成矩阵形式： $\mathbf{v} = \mathbf{H}\mathbf{u} + \mathbf{n}$ 。已知 \mathbf{u} , \mathbf{n} 的二阶统计量（如协方差/相关矩阵）分别为 \mathbf{R}_u , \mathbf{R}_n ，且 \mathbf{u} 与 \mathbf{n} 不相关。

1. 在线性最小均方意义下，设 $\hat{\mathbf{u}} = \mathbf{C}\mathbf{v}$ ，求最优滤波矩阵 \mathbf{C}
2. 给出对应的估值均方误差

解答：

1) 最优滤波矩阵 \mathbf{C}^* (LMMSE/Wiener 解) 线性最小均方估计 (LMMSE) 的标准形式为

$$\mathbf{C}^* = \mathbf{R}_{uv} \mathbf{R}_v^{-1},$$

其中

$$\mathbf{R}_{uv} = \mathbb{E}[\mathbf{u}\mathbf{v}^T], \quad \mathbf{R}_v = \mathbb{E}[\mathbf{v}\mathbf{v}^T].$$

(a) 计算 \mathbf{R}_{uv} 由

$$\mathbf{v} = \mathbf{H}\mathbf{u} + \mathbf{n},$$

可得

$$\mathbf{R}_{uv} = \mathbb{E}[\mathbf{u}(\mathbf{H}\mathbf{u} + \mathbf{n})^T] = \mathbb{E}[\mathbf{u}\mathbf{u}^T]\mathbf{H}^T + \mathbb{E}[\mathbf{u}\mathbf{n}^T].$$

由于 $\mathbb{E}[\mathbf{u}\mathbf{n}^T] = \mathbf{0}$, 因此

$$\mathbf{R}_{uv} = \mathbf{R}_u \mathbf{H}^T.$$

(b) 计算 \mathbf{R}_v 同理

$$\mathbf{R}_v = \mathbb{E}[(\mathbf{H}\mathbf{u} + \mathbf{n})(\mathbf{H}\mathbf{u} + \mathbf{n})^T] = \mathbf{H}\mathbb{E}[\mathbf{u}\mathbf{u}^T]\mathbf{H}^T + \mathbb{E}[\mathbf{n}\mathbf{n}^T],$$

且交叉项因不相关为零, 于是

$$\mathbf{R}_v = \mathbf{H}\mathbf{R}_u\mathbf{H}^T + \mathbf{R}_n.$$

(c) 代入得到 \mathbf{C}^* 将上述结果代入 $\mathbf{C}^* = \mathbf{R}_{uv}\mathbf{R}_v^{-1}$, 得

$$\boxed{\mathbf{C}^* = \mathbf{R}_u \mathbf{H}^T (\mathbf{H}\mathbf{R}_u \mathbf{H}^T + \mathbf{R}_n)^{-1}}.$$

2) 误差协方差与 MSE 定义估计误差

$$\mathbf{e} = \mathbf{u} - \hat{\mathbf{u}} = \mathbf{u} - \mathbf{C}\mathbf{v},$$

误差协方差矩阵

$$\mathbf{R}_e = \mathbb{E}[\mathbf{e}\mathbf{e}^T].$$

LMMSE 最优解对应的误差协方差为

$$\mathbf{R}_e^* = \mathbf{R}_u - \mathbf{R}_{uv} \mathbf{R}_v^{-1} \mathbf{R}_{vu},$$

其中 $\mathbf{R}_{vu} = \mathbf{R}_{uv}^T$ 。代入

$$\mathbf{R}_{uv} = \mathbf{R}_u \mathbf{H}^T, \quad \mathbf{R}_v = \mathbf{H}\mathbf{R}_u \mathbf{H}^T + \mathbf{R}_n, \quad \mathbf{R}_{vu} = \mathbf{H}\mathbf{R}_u,$$

得到

$$\mathbf{R}_e^* = \mathbf{R}_u - \mathbf{R}_u \mathbf{H}^T (\mathbf{H} \mathbf{R}_u \mathbf{H}^T + \mathbf{R}_n)^{-1} \mathbf{H} \mathbf{R}_u.$$

若题目所指的估计均方误差（MSE）定义为

$$\text{MSE} = \mathbb{E}[\|\mathbf{e}\|^2],$$

则有

$$\text{MSE} = \text{tr}(\mathbf{R}_e^*) = \text{tr}\left(\mathbf{R}_u - \mathbf{R}_u \mathbf{H}^T (\mathbf{H} \mathbf{R}_u \mathbf{H}^T + \mathbf{R}_n)^{-1} \mathbf{H} \mathbf{R}_u\right).$$

6 Kalman 滤波器与 Wiener 滤波器设计

给定状态空间模型

$$\begin{aligned}\dot{x}(t) &= -x(t) + w(t), \\ z(t) &= x(t) + v(t),\end{aligned}$$

其中 $w(t)$, $v(t)$ 为零均值白噪声, 且

$$\begin{aligned}\text{cov}(w(t), w(\tau)) &= 2\alpha\delta(t - \tau), \\ \text{cov}(v(t), v(\tau)) &= \alpha\delta(t - \tau),\end{aligned}$$

并假设二者互不相关。

1. 设计对应的稳态 Kalman 滤波器
2. 设计物理可实现的 Wiener 滤波器
3. 二者对比分析

解答:

(1) 稳态 Kalman 滤波器

$$\dot{x}(t) = Ax(t) + Bw(t), \quad z(t) = Hx(t) + v(t),$$

其中

$$A = -1, \quad B = 1, \quad H = 1, \quad Q = 2\alpha, \quad R = \alpha.$$

由

$$\dot{P}(t) = AP(t) + P(t)A^\top - P(t)H^\top R^{-1}HP(t) + BQB^\top.$$

稳态时 $\dot{P}(t) = 0$, 令 $P(t) \rightarrow P$, 得到

$$-2P + 2\alpha - \frac{P^2}{\alpha} = 0,$$

由于 $P \geq 0$,

$$P = \alpha(\sqrt{3} - 1).$$

故

$$K(t) = P(t)H^\top R^{-1} = \sqrt{3} - 1.$$

则

$$\dot{\hat{x}}(t) = A\hat{x}(t) + K(t)(z(t) - H\hat{x}(t)) = -\hat{x}(t) + (\sqrt{3} - 1)(z(t) - \hat{x}(t)).$$

$$\dot{\hat{x}}(t) = -\sqrt{3}\hat{x}(t) + (\sqrt{3} - 1)z(t).$$

$$H_K(s) = \frac{\hat{X}(s)}{Z(s)} = \frac{\sqrt{3} - 1}{s + \sqrt{3}}.$$

(2) 物理可实现 (因果) Wiener 滤波器 由状态方程

$$\dot{x}(t) = -x(t) + w(t),$$

可视为稳定系统

$$G(s) = \frac{1}{s + 1}$$

对输入白噪声 $w(t)$ 的输出。由于

$$\text{Cov}(w(t), w(\tau)) = 2\alpha\delta(t - \tau),$$

故其双边功率谱密度为

$$\Phi_w(s) = 2\alpha.$$

于是

$$\Phi_x(s) = |G(j\omega)|^2 \Big|_{s=j\omega} \Phi_w(s) = \frac{1}{1 - s^2} \cdot 2\alpha = \frac{2\alpha}{1 - s^2}.$$

观测为 $z = x + v$, 且

$$\text{Cov}(v(t), v(\tau)) = \alpha \delta(t - \tau) \Rightarrow \Phi_v(s) = \alpha,$$

且 x 与 v 不相关, 因此

$$\Phi_z(s) = \Phi_x(s) + \Phi_v(s) = \frac{2\alpha}{1-s^2} + \alpha = \alpha \frac{3-s^2}{1-s^2}.$$

同时互谱为

$$\Phi_{xz}(s) = \Phi_x(s) = \frac{2\alpha}{1-s^2}.$$

$$\Phi_z^+(s) = \sqrt{\alpha} \frac{\sqrt{3}+s}{1+s}, \quad \Phi_z^-(s) = \frac{1-s}{\sqrt{3}-s} \sqrt{\alpha}.$$

Wiener-Hopf 公式可写为

$$H_W(s) = \frac{1}{\Phi_z^+(s)} \left[\frac{\Phi_{xz}(s)}{\Phi_z^-(s)} \right]^{t+},$$

$$\frac{\Phi_{xz}(s)}{\Phi_z^-(s)} = \frac{2\sqrt{\alpha}}{(1+s)(\sqrt{3}-s)} = \frac{\sqrt{\alpha}(\sqrt{3}-1)}{1+s} + \frac{\sqrt{\alpha}(\sqrt{3}-1)}{\sqrt{3}-s}.$$

得

$$\left[\frac{\Phi_{xz}(s)}{\Phi_z^-(s)} \right]^{t+} = \frac{\sqrt{\alpha}(\sqrt{3}-1)}{1+s}.$$

则,

$$H_W(s) = \frac{1}{\sqrt{\alpha}} \frac{1+s}{\sqrt{3}+s} \cdot \frac{\sqrt{\alpha}(\sqrt{3}-1)}{1+s} = \frac{\sqrt{3}-1}{s+\sqrt{3}}.$$

物理可实现 (因果) Wiener 滤波器的冲激响应为

$$h_W(t) = (\sqrt{3}-1)e^{-\sqrt{3}t}u(t),$$

等价的微分方程实现为

$$\dot{\hat{x}}(t) = -\sqrt{3}\hat{x}(t) + (\sqrt{3}-1)z(t).$$

(3) 结果分析与比较 本题中

稳态 Kalman 滤波器 \equiv 物理可实现 (因果) Wiener 滤波器.

同时, 稳态误差协方差为

$$P = \alpha(\sqrt{3}-1),$$

稳态时的 Kalman 滤波和 Wiener 滤波是完全相同的。这也说明, 他们都是在最小均方误差准则下的最佳滤波器。

7 渐近相对效率计算

计算以下值及渐近相对效率的取值范围的大小关系。

分布 $f(x)$	$\text{ARE}_{\text{Wilcoxon}, t\text{-test}}$	$\text{ARE}_{\text{sign}, t\text{-test}}$
均匀分布	1	(填空/计算)
正态 (Gaussian, 记 N)	(填空/计算)	(填空/计算)
拉氏 (Laplace, 记 D)	(填空/计算)	(填空/计算)

表 1: 不同分布下的渐近相对效率

解答:

1) Pitman 渐近相对效率 (ARE) 定义 考虑位置参量微小偏移 (shift) 模型

$$X_i \sim f(x - \theta), \quad \theta = 0 \text{ vs } \theta > 0,$$

在局部备择 $\theta = h/\sqrt{n}$ 下, 许多检验统计量经标准化后满足

$$T_n \Rightarrow \mathcal{N}(h\Delta, 1),$$

其中 Δ 为该检验的 *efficacy* 参数。定义

$$\text{ARE}_{A,B} = \frac{\Delta_A^2}{\Delta_B^2}.$$

2) 通用公式 设 $\sigma^2 = \text{Var}(X)$ 。

• t 检验: $\Delta_t = \frac{1}{\sigma}$, 故 $\Delta_t^2 = \frac{1}{\sigma^2}$ 。

• 符号检验 (sign test):

$$\Delta_{\text{sign}} = 2f(0), \quad \Delta_{\text{sign}}^2 = 4f(0)^2,$$

因而

$$\boxed{\text{ARE}_{\text{sign}, t} = 4\sigma^2 f(0)^2.}$$

• Wilcoxon (秩检验):

$$\Delta_W = \sqrt{12} \int_{-\infty}^{\infty} f(x)^2 dx, \quad \Delta_W^2 = 12 \left(\int_{-\infty}^{\infty} f(x)^2 dx \right)^2,$$

因而

$$\boxed{\text{ARE}_{\text{Wilcoxon}, t} = 12\sigma^2 \left(\int_{-\infty}^{\infty} f(x)^2 dx \right)^2.}$$

3) 各分布下的计算结果

1. 均匀分布 $U(-a, a)$

$$f(x) = \frac{1}{2a} \mathbf{1}_{\{|x| \leq a\}}, \quad \sigma^2 = \frac{a^2}{3}, \quad f(0) = \frac{1}{2a}, \quad \int f^2 = \frac{1}{2a}.$$

$$\text{ARE}_{\text{Wilcoxon},t} = 12 \cdot \frac{a^2}{3} \left(\frac{1}{2a} \right)^2 = 1, \quad \text{ARE}_{\text{sign},t} = 4 \cdot \frac{a^2}{3} \left(\frac{1}{2a} \right)^2 = \frac{1}{3}.$$

2. 正态分布 $N(0, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}, \quad f(0) = \frac{1}{\sqrt{2\pi}\sigma}, \quad \int f^2 = \frac{1}{2\sqrt{\pi}\sigma}.$$

$$\text{ARE}_{\text{Wilcoxon},t} = 12\sigma^2 \left(\frac{1}{2\sqrt{\pi}\sigma} \right)^2 = \frac{3}{\pi} \approx 0.9549, \quad \text{ARE}_{\text{sign},t} = 4\sigma^2 \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^2 = \frac{2}{\pi} \approx 0.6366.$$

3. 拉普拉斯分布 $\text{Laplace}(0, b)$

$$f(x) = \frac{1}{2b} e^{-|x|/b}, \quad \sigma^2 = 2b^2, \quad f(0) = \frac{1}{2b}, \quad \int f^2 = \frac{1}{4b}.$$

$$\text{ARE}_{\text{Wilcoxon},t} = 12 \cdot 2b^2 \left(\frac{1}{4b} \right)^2 = \frac{3}{2} = 1.5, \quad \text{ARE}_{\text{sign},t} = 4 \cdot 2b^2 \left(\frac{1}{2b} \right)^2 = 2.$$

4. Logistic 分布 (位置 0、尺度 s)

$$f(x) = \frac{e^{-x/s}}{s(1 + e^{-x/s})^2} = \frac{1}{4s} \text{sech}^2\left(\frac{x}{2s}\right), \quad \sigma^2 = \frac{\pi^2 s^2}{3}, \quad f(0) = \frac{1}{4s}.$$

并且

$$\int f(x)^2 dx = \int \frac{1}{16s^2} \text{sech}^4\left(\frac{x}{2s}\right) dx = \frac{1}{16s^2} \cdot 2s \int_{-\infty}^{\infty} \text{sech}^4(u) du.$$

利用

$$\int_{-\infty}^{\infty} \text{sech}^4(u) du = \frac{4}{3},$$

得

$$\int f^2 = \frac{1}{16s^2} \cdot 2s \cdot \frac{4}{3} = \frac{1}{6s}.$$

于是

$$\text{ARE}_{\text{sign},t} = 4 \cdot \frac{\pi^2 s^2}{3} \left(\frac{1}{4s} \right)^2 = \frac{\pi^2}{12} \approx 0.822, \quad \text{ARE}_{\text{Wilcoxon},t} = 12 \cdot \frac{\pi^2 s^2}{3} \left(\frac{1}{6s} \right)^2 = \frac{\pi^2}{9} \approx 1.097.$$

4) 汇总表

分布 $f(x)$	$\text{ARE}_{\text{Wilcoxon},t}$	$\text{ARE}_{\text{sign},t}$	$\text{ARE}_{\text{Wilcoxon},\text{sign}}$
$U(-a, a)$	1	$\frac{1}{3}$	3
$N(0, \sigma^2)$	$\frac{3}{\pi} \approx 0.955$	$\frac{2}{\pi} \approx 0.637$	$\frac{3}{2} = 1.5$
$\text{Laplace}(0, b)$	$\frac{3}{2} = 1.5$	2	$\frac{3}{4} = 0.75$
$\text{Logistic}(0, s)$	$\frac{\pi^2}{9} \approx 1.097$	$\frac{\pi^2}{12} \approx 0.822$	$\frac{4}{3} \approx 1.333$