Homework of Optimal Estimation

Junda Wu 255065010

Homework 1

1. Proof Since X and Y are uncorrelated, we have $Cov(X,Y) = E[(X-EX)(Y-EY)] = 0$. Hence

$$C_Z = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$$

Then $|\det C_Z|^{1/2} = \sigma_x \sigma_y$, $C_Z^{-1} = \begin{bmatrix} 1/\sigma_x^2 & \\ & 1/\sigma_y^2 \end{bmatrix}$

Since X, Y are jointly normal, we have

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{(x-EX)^2}{2\sigma_x^2} - \frac{(y-EY)^2}{2\sigma_y^2} \right]$$

Then we can calculate marginal distribution:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left[-\frac{(x-EX)^2}{2\sigma_x^2} \right] \frac{1}{\sqrt{2\pi}\,\sigma_y} \int_{-\infty}^{+\infty} \exp \frac{(y-Ey)^2}{2\sigma_y^2}\, dy$$

recall that $\frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{+\infty} \exp \frac{(x-\mu)^2}{2\sigma^2}\, dx = 1$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left[-\frac{(x-EX)^2}{2\sigma_x^2} \right]$$

Similarly, $f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma_y} \exp\left[-\frac{(y-EY)^2}{2\sigma_y^2} \right]$

Note that $f(x,y) = f_X(x) f_Y(y)$, we can conclude that X and Y are independent. $\square$

2. Solution

(1) $$R_Z(i,j) = E\Big(z(i)\,z(j)\Big) = E\Big(s^2 + s\big(v(i)+v(j)\big) + v(i)\,v(j) \Big)$$

$$= \underbrace{E(s^2)}_{=2} + \underbrace{E\big(sv(i)\big)}_{=1} + \underbrace{E\big(sv(j)\big)}_{=1} + \underbrace{E\big(v(i)\,v(j)\big)}_{=\begin{cases} 0, & i\neq j \\ Var(v(i)) = 1, & i=j \end{cases}} = \begin{cases} 5, & i=j \\ 4, & i\neq j \end{cases}$$

(2) $E\big(z(n)\big) = E(s) + E\big(v(n)\big) = 1 \Rightarrow E\big(z(n)\big)$ is constant

$R_Z(k) = \begin{cases} 4, & k\neq 0 \\ 5, & k=0 \end{cases} \Rightarrow R_Z(k)$ is independent of time

$\Rightarrow z(n)$ is WSS. $\square$

Homework 2

1. Solution Maximum Likelihood:

$$P(Y=\{6,2,6\} \mid X=\{1,1,1\}) = \overset{\frac{1}{6}}{P(Y=6 \mid X=1)}\ \overset{\frac{1}{6}}{P(Y=2 \mid X=1)}\ \overset{\frac{1}{6}}{P(Y=6 \mid X=1)} = \frac{1}{216}$$

$$P(Y=\{6,2,6\} \mid X=\{1,0,1\}) = \frac{1}{6} \times \frac{1}{10} \times \frac{1}{6} = \frac{1}{360}$$

$$P(Y=\{6,2,6\} \mid X=\{0,1,1\}) = \frac{1}{2} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{72}$$

$$P(Y=\{6,2,6\} \mid X=\{1,1,0\}) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{2} = \frac{1}{72}$$

$$P(Y=\{6,2,6\} \mid X=\{1,0,0\}) = \frac{1}{6} \times \frac{1}{10} \times \frac{1}{2} = \frac{1}{120} \qquad \Rightarrow X_{ML} = \{0,1,0\}$$

$$P(Y=\{6,2,6\} \mid X=\{0,0,1\}) = \frac{1}{2} \times \frac{1}{10} \times \frac{1}{6} = \frac{1}{120}$$

$$P(Y=\{6,2,6\} \mid X=\{0,1,0\}) = \frac{1}{2} \times \frac{1}{6} \times \frac{1}{2} = \frac{1}{24}$$

$$P(Y=\{6,2,6\} \mid X=\{0,0,0\}) = \frac{1}{2} \times \frac{1}{10} \times \frac{1}{2} = \frac{1}{40}$$

MAP: $\lim_{n\to\infty} A^n = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \Rightarrow p(x=0) = \frac{1}{3} \qquad p(x=1) = \frac{2}{3}$

$$p(x \mid Y) = \frac{p(Y \mid x)\, p(x)}{p(Y)} \propto p(Y \mid x)\, p(x)$$

So we can compute

$$P(x=\{1,1,1\})\ P(Y=\{6,2,6\} \mid X=\{1,1,1\}) = \frac{1}{216} \times \frac{2}{3} \times 0.95 \times 0.95 = \frac{361}{129600}$$

$$P(x=\{1,0,1\})\ P(Y=\{6,2,6\} \mid X=\{1,0,1\}) = \frac{1}{360} \times \frac{2}{3} \times 0.05 \times 0.1 = \frac{1}{108000}$$

$$P(x=\{0,1,1\})\ P(Y=\{6,2,6\} \mid X=\{0,1,1\}) = \frac{1}{72} \times \frac{1}{3} \times 0.1 \times 0.95 = \frac{19}{43200} = \frac{57}{129600}$$

$$P(x=\{1,1,0\})\ P(Y=\{6,2,6\} \mid X=\{1,1,0\}) = \frac{1}{72} \times \frac{2}{3} \times 0.95 \times 0.05 = \frac{19}{43200}$$

$$P(x=\{1,0,0\})\ P(Y=\{6,2,6\} \mid X=\{1,0,0\}) = \frac{1}{120} \times \frac{2}{3} \times 0.05 \times 0.9 = \frac{9}{36000} = \frac{1}{4000}$$

$$P(x=\{0,0,1\})\ P(Y=\{6,2,6\} \mid X=\{0,0,1\}) = \frac{1}{120} \times \frac{1}{3} \times 0.9 \times 0.1 = \frac{1}{4000} \qquad \text{So, } X_{MAP} = \{0,0,0\}.$$

$$P(x=\{0,1,0\})\ P(Y=\{6,2,6\} \mid X=\{0,1,0\}) = \frac{1}{24} \times \frac{1}{3} \times 0.1 \times 0.05 = \frac{1}{14400}$$

$$P(x=\{0,0,0\})\ P(Y=\{6,2,6\} \mid X=\{0,0,0\}) = \frac{1}{40} \times \frac{1}{3} \times 0.9 \times 0.9 = \frac{81}{12000} = \frac{27}{4000} \rightarrow \text{maximum}$$

2、 Solution

(1) (a) $f(z|s) = f_V(v)\big|_{v=z-s} = \frac{1}{\sqrt{2\pi}V} e^{-(z-s)^2/2V^2}$

$$f_S(s) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(s-\eta)^2}{2\sigma^2}\right] \qquad f_Z(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+V^2}}\exp\left[\frac{-(z-\eta)^2}{2(\sigma^2+V^2)}\right]$$

$$\Rightarrow f_S(s|z=z) = \frac{f(z|s)f_S(s)}{f_Z(z)}$$

$$= \frac{1}{2\pi\sigma V \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+V^2}}} \exp\left\{-\frac{(z-s)^2}{2V^2} - \frac{(s-\eta)^2}{2\sigma^2} + \frac{(z-\eta)^2}{2(\sigma^2+V^2)}\right\}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2V^2}{\sigma^2+V^2}}}\exp\frac{-\sigma^2(\sigma^2+V^2)(z^2-2zs+s^2) - V^2(\sigma^2+V^2)(s^2-\eta s\times 2+\eta^2)+\sigma^2V^2(z^2-\eta z\times 2+\eta^2)}{2(V^2)(\sigma^2)(\sigma^2+V^2)}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2V^2}{\sigma^2+V^2}}}\exp \frac{1}{2V^2\sigma^2(\sigma^2+V^2)}\left\{-(\sigma^2+V^2)^2 s^2 + s\left(\sigma^2(\sigma^2+V^2)2z + V^2(\sigma^2+V^2)2\eta\right) - \underbrace{\left[\sigma^4 z^2 + V^4\eta^2 + 2\eta z\sigma^2V^2\right]}_{=(\sigma^2 z+V^2\eta)^2}\right\}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2V^2}{\sigma^2+V^2}}}\exp\left[-\frac{\left[(\sigma^2+V^2)s-(\sigma^2z+V^2\eta)\right]^2}{2V^2\sigma^2(\sigma^2+V^2)}\right]$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{\sigma^2V^2}{\sigma^2+V^2}}}\exp\left[-\frac{1}{2\frac{V^2\sigma^2}{\sigma^2+V^2}}\left[s-\left(\frac{\sigma^2}{\sigma^2+V^2}z+\frac{V^2}{\sigma^2+V^2}\eta\right)\right]^2\right]$$

$$\Rightarrow E[s|z=z] = \frac{\sigma^2}{\sigma^2+V^2}z+\frac{V^2}{\sigma^2+V^2}\eta = \eta+\frac{\sigma^2}{\sigma^2+V^2}(z-\eta)$$

(b) $E[s^2|z=z] = D[s|z=z] + E^2[s|z=z] = \frac{\sigma^2V^2}{\sigma^2+V^2} + \left(\eta+\frac{\sigma^2}{\sigma^2+V^2}(z-\eta)\right)^2$

(2) (a) Maximum Likelihood Estimate:

$$f_Z(z=z(i)|s) = \frac{1}{\sqrt{2\pi}\sigma_V}\exp\left[-\frac{(z(i)-s)^2}{2\sigma_V^2}\right]$$

So we have the likelihood function

$$f_Z(z=z(1),\cdots,z(n)|s) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_V}\exp\left[-\frac{(z(i)-s)^2}{2\sigma_V^2}\right]$$

To maximize the above function, consider logarithm of it:

$$\ln f_Z = -n\ln\left(\sqrt{2\pi}\,\sigma_V\right) - \sum_{i=1}^{n}\frac{(z(i)-s)^2}{2\sigma_V^2} \Rightarrow \frac{\partial f_Z}{\partial s} = \sum_{i=1}^{n}\frac{z(i)-s}{\sigma_V^2} = 0 \Rightarrow \hat{S}_{ML} = \frac{1}{n}\sum_{i=1}^{n} z(i)$$

(b) Maximum A Posteriori estimate (MAP):

$$f_S(s \mid Z = z(1),\cdots,z(n)) \propto f_S(s)\, f_Z(Z = z(1),\cdots,z(n) \mid s)$$

$$= \underbrace{\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\,\sigma_V}\exp\left[-\frac{(z(i)-s)^2}{2\sigma_V^2}\right]}_{\text{likelihood}} \times \underbrace{\frac{1}{\sqrt{2\pi}\,\sigma_S}\exp\left(-\frac{(s-\eta_s)^2}{2\sigma_S^2}\right)}_{\text{prior}}$$

Again, consider the logarithm to maximize it:

$$\ln f_Z = -n\ln\left(\sqrt{2\pi}\,\sigma_V\right) - \ln\left(\sqrt{2\pi}\,\sigma_S\right) - \sum_{i=1}^{n}\frac{(z(i)-s)^2}{2\sigma_V^2} - \frac{(s-\eta_s)^2}{2\sigma_S^2}$$

$$\frac{\partial \ln f_Z}{\partial s} = \sum_{i=1}^{n}\frac{(z(i)-s)}{\sigma_V^2} - \frac{s-\eta_s}{\sigma_S^2} = 0 \Rightarrow \hat{S}_{MAP} = \frac{\sigma_S^2}{n\sigma_S^2+\sigma_V^2}\sum_{i=1}^{n} z(i) + \frac{\sigma_V^2}{n\sigma_S^2+\sigma_V^2}\eta_s$$

$$\Downarrow$$

$$-\frac{n}{\sigma_V^2}s + \frac{1}{\sigma_V^2}\sum_{i=1}^{n} z(i) - \frac{s}{\sigma_S^2} + \frac{\eta_s}{\sigma_S^2} = 0 \Rightarrow -n\sigma_S^2 s + \sigma_S^2\sum_{i=1}^{n} z(i) - \sigma_V^2 s + \sigma_V^2\eta_s = 0$$

(c) $f(S=s, z_1 = z(1), \cdots, z_n = z(n)) = f(S=s, v_1 = z(1)-s, \cdots, v_n = z(n)-s)$

Since s and v are independent, define $\vec{v} = [v_1, \cdots, v_n]$, we have $\begin{bmatrix} S \\ \vec{v} \end{bmatrix} \sim N\left(\begin{bmatrix} \eta_s \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_S^2 & \\ & \sigma_V^2 I \end{bmatrix}\right)$

Hence $f(S=s, v_1 = z(1)-s, \cdots, v_n = z(n)-s)$

$$= \frac{1}{(2\pi)^{\frac{n+1}{2}}\sigma_S\sigma_V^n}\exp\left(-\frac{1}{2}\left[(s-\eta_s) \quad (\vec{z}-s\vec{1})^T\right]\begin{bmatrix} \sigma_S^2 & \\ & \sigma_V^2 I \end{bmatrix}^{-1}\begin{bmatrix} s-\eta_s \\ \vec{z}-s\vec{1} \end{bmatrix}\right)$$

$$= \frac{1}{\sigma_S^2}(s-\eta_s)^2 + \frac{1}{\sigma_V^2}\sum_{i=1}^{n}(z_i - s)^2 = \frac{1}{\sigma_S^2}(s-\eta_s)^2 + \frac{1}{\sigma_V^2}\sum_{i=1}^{n}(z_i^2 + s^2 - 2sz_i)$$

On the other hand, $\vec{z} \sim N(\eta_s\vec{1}, (\sigma_S^2+\sigma_V^2)I)$. Therefore,

$$\frac{1}{\sigma_V^2+\sigma_S^2}\sum_{i=1}^{n}(z_i^2 - 2\eta_s z_i + \eta_s^2)$$

$$\parallel$$

$$f(z = z(1), \cdots, z(n)) = \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma_S+\sigma_V)^n}\exp\left(-\frac{1}{2}\left[\vec{z}-\eta_s\vec{1}\right]^T\left[(\sigma_V^2+\sigma_S^2)I\right]^{-1}\left[\vec{z}-\eta_s\vec{1}\right]\right)$$

Then, by Bayes equation,

$$f(s \mid z_1 = z(1), \cdots, z_n = z(n)) = \frac{f(S=s, z_1 = z(1), \cdots, z_n = z(n))}{f(z_1 = z(1), \cdots, z_n = z(n))}$$

$$= \frac{(\sigma_S+\sigma_V)^n}{\sqrt{2\pi}\,\sigma_S\sigma_V^n}\exp\left[\left(\frac{1}{\sigma_S^2}+\frac{n}{\sigma_V^2}\right)s^2 - \frac{2\eta_s}{\sigma_S^2}s - \frac{2}{\sigma_V^2}\sum_{i=1}^{n} z(i)\, s + X\right]$$

Note that it's a Gaussian distribution, thus

$$\hat{S}_{MMSE} = E(s \mid z_1 = z(1), \cdots, z_n = z(n)) = \frac{1}{\frac{1}{\sigma_S^2}+\frac{n}{\sigma_V^2}}\left(\frac{\eta_s}{\sigma_S^2} + \frac{1}{\sigma_V^2}\sum_{i=1}^{n} z(i)\right) = \frac{1}{\sigma_V^2+n\sigma_S^2}\left(\eta_s\sigma_V^2 + \sigma_S^2\sum_{i=1}^{n} z(i)\right)$$

Note that in this case, $\hat{s}_{MAP} = \hat{s}_{MMSE}$!

(d) $E[s z(i)] = E[s(s+v(i))] = E[s^2] + E(s)\underbrace{E(v(i))}_{=0} = Var(s) + E^2(s) = \eta_s^2 + \sigma_s^2$

$$E[z(i)z(j)] = \underbrace{E[s^2]}_{=\eta_s^2+\sigma_s^2} + \underbrace{E(s)}_{=0} E[v(i)+v(j)] + \underbrace{E(v(i)+v(j))}_{=\begin{cases} 0 & i\neq j \\ \sigma_v^2 & i=j\end{cases}} = \begin{cases} \sigma_s^2+\eta_s^2 & i\neq j \\ \sigma_s^2+\eta_s^2+\sigma_v^2 & i=j \end{cases}$$

$$E(zz^T) = \sigma_v^2 I + (\eta_s^2+\sigma_s^2) M \qquad M = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

$$= \sigma_v^2 I + (\eta_s^2+\sigma_s^2)\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}[1 \cdots 1]$$

Lemma. For $A = kI + lxx^T$, we have its inverse as $A^{-1} = \frac{1}{k}I - \frac{l}{k(k+lx^Tx)}xx^T$

Proof of Lemma.

$$AA^{-1} = (kI + lxx^T)\left(\frac{1}{k}I - \frac{l}{k(k+lx^Tx)}xx^T\right)$$

$$= I - \frac{l^2x^Tx}{k(k+lx^Tx)}xx^T + \frac{l}{k}xx^T - \frac{kl}{k(k+lx^Tx)}xx^T$$

$$= I + \frac{-l^2x^Tx + l(k+lx^Tx) - kl}{k(k+lx^Tx)}xx^T = I \qquad \square$$

$x^Tx = n$, $k = \sigma_v^2$, $l = \sigma_s^2+\eta_s^2$

$$\Rightarrow E^{-1}(zz^T) = \frac{1}{\sigma_v^2}I - \frac{(\eta_s^2+\sigma_s^2)}{\sigma_v^2[\sigma_v^2+(\eta_s^2+\sigma_s^2)n]}M = \frac{1}{\sigma_v^2[\sigma_v^2+(\eta_s^2+\sigma_s^2)n]}\begin{bmatrix} \sigma_v^2+(n-1)(\sigma_s^2+\eta_s^2) & \cdots & -(\sigma_s^2+\eta_s)^2 \\ \vdots & & \vdots \\ -(\sigma_s^2+\eta_s^2) & \cdots & \sigma_v^2+(n-1)(\sigma_s^2+\eta_s^2) \end{bmatrix}$$

$$E(sz^T)[E(zz^T)]^{-1} = \frac{\eta_s^2+\sigma_s^2}{\sigma_v^2[\sigma_v^2+(\eta_s^2+\sigma_s^2)n]}[\sigma_v^2, \cdots, \sigma_v^2] = \frac{\eta_s^2+\sigma_s^2}{[\sigma_v^2+(\eta_s^2+\sigma_s^2)n]}\vec{1}$$

$$\Rightarrow \hat{s}_{LMMSE} = E(sz^T)[E(zz^T)]^{-1}[z(1), \cdots, z(n)]^T = \frac{\eta_s^2+\sigma_s^2}{\sigma_v^2+(\eta_s^2+\sigma_s^2)n}\sum_{i=1}^{n} z(i)$$

(e) Since $\begin{bmatrix} z(1) \\ \vdots \\ z(n) \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} s + v$, $\hat{s}_{LS} = (H^TH)^{-1}H^T\begin{bmatrix} z(1) \\ \vdots \\ z(n) \end{bmatrix}^T = \frac{1}{n}\sum_{i=1}^{n} z(i)$

Note that $\hat{s}_{LS} = \hat{s}_{ML}$!

(f) $z_k = s + v_k$, $\hat{s}_0 = E(s) = \eta_s$, $H_k = 1$, $R_k = \sigma_v^2$ $P_0 = \sigma_s^2$

Iteration:

$$\begin{cases} k_k = P_{k-1}H_k^T(H_kP_{k-1}H_k^T + R_k)^{-1} = P_{k-1}(P_{k-1}+R_k)^{-1} \\ P_k = [I - k_kH_k]P_{k-1} = (1-k_k)P_{k-1} \\ \hat{s}_k = \hat{s}_{k-1} + k_k(z_k - H_k\hat{s}_{k-1}) = (1-k_k)\hat{s}_{k-1} + k_kz_k \end{cases}$$

By iteration, we obtain

$$\begin{cases} k_k = \dfrac{\sigma_s^2}{k\sigma_s^2 + \sigma_v^2} \\ \hat{s}_k = \dfrac{\sigma_s^2 \sum_{i=1}^{k} z(i) + \sigma_v^2 \eta_s}{k\sigma_s^2 + \sigma_v^2} \\ P_k = \dfrac{\sigma_s^2 \sigma_v^2}{k\sigma_s^2 + \sigma_v^2} \end{cases}$$

$$\Rightarrow \quad \hat{S}_{RLS} = \hat{S}_n = \frac{\sigma_s^2 \sum_{i=1}^{n} z(i) + \sigma_v^2 \eta_s}{n\sigma_s^2 + \sigma_v^2}$$

Note that $\hat{S}_{RLS} = \hat{S}_{MAP} = \hat{S}_{MMSE}$!

(g) ① Maximum Likelihood Estimate:

It maximizes the likelihood function based on data only without prior knowledge.

Advantage: Simple, easy to compute.

Disadvantage: Lacks consideration of prior information, sensitive to noise

② Maximum A Posteriori (MAP) Estimate:

Combines likelihood and prior distribution.

Advantage: Utilizes prior knowledge

Disadvantage: Depends on the choice of prior distribution.

③ Minimum Mean Square Estimate (MMSE):

Minimizes mean square error using posteriori distribution.

Advantage: Optimal for mean squared error.

Disadvantage: Computationally complex.

④ Linear Minimum Mean Square Estimate (LMMSE)

A linear form of MMSE.

Advantage: Easier to compute.

Disadvantage: It's biased, may be less accurate for nonlinear relations.

⑤ Least Squares Estimate:

Minimizes sum of squared errors.

Advantage: Easy to understand and implement, low computational cost.

Disadvantage: Sensitive to outliers.

⑥ Recursive Least Squares Estimate:

Updates the estimate as new data comes in, based on principle of least squares.

Advantage: Saves computational resources.

Disadvantage: Depends on initial estimate.