AAE 590 Applied Optimal Control and Estimation Problem Set 1

Baboo J. Cui, Yangang Cao

June 26, 2019

Problem 1

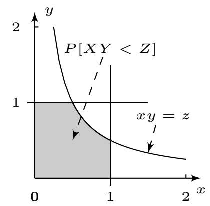
Let the random variables X and Y have the probability density function (pdf)

$$f(x,y) = \begin{cases} 1 & 0<x<1, 0<y<1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the probability distribution function (PDF) and pdf of the product $Z = XY$

Solution

PDF of random variable $Z, F_Z(z) = P[Z \leq z] = P[XY \leq z]$ is given by following figure.



$$\therefore F_Z(z) = \begin{cases} 1 & z \geq 1 \\ 0 & z \leq 0 \\ z + \int_z^1 \int_0^{\frac{2}{x}} f_{XY}(x,y)dydx & 0<z<1 \end{cases}$$

Since $f_{XY}(x,y) = 1 \quad 0<x<1, \quad 0<y<1,$

$$\begin{aligned} \Rightarrow F_Z(z) &= z + \int_z^1 \int_0^{\frac{2}{x}} dydx = z + \int_z^1 \frac{z}{x}dx \\ &= z - z\ln z & 0<z<1 \\ \Rightarrow f_Z(z) &= \frac{dF_Z(z)}{dz} = -\ln z & 0<z<1 \end{aligned}$$

Problem 2

Determine the pdf of $Y, f_Y(y)$, where $Y = \sin X$, and X is a uniform random variable with pdf

$$f_X(x) = \begin{cases} \frac{1}{2\pi} & -\pi < x \leq \pi \\ 0 & \text{otherwise} \end{cases}$$

Solution

PDF of random variable Y is given by:

$$F_Y(y) = P[Y \leq y] = \begin{cases} 1 & y \geq 1 \\ 0 & y < -1 \\ P[\sin X \leq y] & -1 \leq y < 1 \end{cases}$$

Then,

$$\begin{aligned} P[\sin X \leq y] &= 1 - P[\sin X \geq y] \\ &= 2P[X \leq \arcsin y] \\ &= 2\int_{-\frac{\pi}{2}}^{\arcsin y} f_X(x)dx = 2\int_{-\frac{\pi}{2}}^{\arcsin y} \frac{1}{2\pi}dx \\ &= \frac{1}{\pi}\left(\arcsin y + \frac{\pi}{2}\right) \end{aligned}$$

Therefore,

$$\therefore f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{\pi\sqrt{1-y^2}} \qquad -1 \leq y < 1$$

Problem 3

Let X_1, X_2, and X_3 be random variables with equal variances but with correlation coefficients $\rho_{12} = 0.3, \rho_{13} = 0.5$, and $\rho_{23} = 0.2$. Find the correlation coefficient of the linear functions $Y = X_1 + X_2$ and $Z = X_2 + X_3$

Solution

Let σ^2 be the variance of X_1, X_2, X_3 . Then,

$$\begin{cases} \text{Cov}\left[X_1, X_2\right] = \sigma^2\rho_{12} \\ \text{Cov}\left[X_1, X_3\right] = \sigma^2\rho_{13} \\ \text{Cov}\left[X_2, X_3\right] = \sigma^2\rho_{23} \end{cases}$$

Using this,

$$\begin{aligned} \text{Cov}[Y, Z] &= \mathbb{E}\left[\left(X_1 + X_2\right)\left(X_2 + X_3\right)\right] - \mathbb{E}\left[X_1 + X_2\right]\mathbb{E}\left[X_2 + X_3\right] \\ &= \mathbb{E}\left[X_1X_2\right] + \mathbb{E}\left[X_1X_3\right] + \mathbb{E}\left[X_1^2\right] + \mathbb{E}\left[X_2X_3\right] - \mathbb{E}\left[X_1\right]\mathbb{E}\left[X_2\right] \\ &\quad - \mathbb{E}\left[X_1\right]\mathbb{E}\left[X_3\right] - \mathbb{E}\left[X_1\right]^2 - \mathbb{E}\left[X_2\right]\mathbb{E}\left[X_3\right] \\ &= \text{Cov}\left[X_1, X_2\right] + \text{Cov}\left[X_1, X_3\right] + \sigma^2 + \text{Cov}\left[X_2, X_3\right] \\ &= \sigma^2\left(\rho_{12} + \rho_{13} + 1 + \rho_{23}\right) \end{aligned}$$

On the other hand, the variances of Y and Z are:

$$\begin{aligned}
\sigma_Y^2 &= \mathbb{E}\left[\left(X_1+X_2\right)\left(X_1+X_2\right)\right]-\mathbb{E}\left[X_1+X_2\right]\mathbb{E}\left[X_1+X_2\right]\\
&= 2\sigma^2+2\operatorname{Cov}\left[X_1,X_2\right]=2\sigma^2\left(1+\rho_{12}\right)\\
\sigma_Z^2 &= \mathbb{E}\left[\left(X_2+X_3\right)\left(X_2+X_3\right)\right]-\mathbb{E}\left[X_2+X_3\right]\mathbb{E}\left[X_2+X_3\right]\\
&= 2\sigma^2+2\operatorname{Cov}\left[X_2,X_3\right]=2\sigma^2\left(1+\rho_{23}\right)\\
\therefore \rho_{YZ} &= \frac{\operatorname{Cov}[Y,Z]}{\sigma_Y\sigma_Z}=\frac{\rho_{12}+\rho_{13}+1+\rho_{23}}{\sqrt{2+2\rho_{12}}\sqrt{2+\rho_{23}}}=\frac{2}{\sqrt{2.6}\sqrt{2.4}}=0.8006
\end{aligned}$$

Problem 4

Let

$$f\left(x_1,x_2\right)=\begin{cases} 6x_1 & 0<x_1<x_2<1\\ 0 & \text{elsewhere}\end{cases}$$

be the joint pdf of the random variables X_1 and X_2 .

1. Find the conditional mean and variance of X_1, given $X_2=x_2, 0<x_2<1$
2. Find the pdf of the random variable $Y=\mathbb{E}\left[X_1|X_2\right]$
3. Determine $\mathbb{E}[Y]$ and $\operatorname{Var}[Y]$ and compare these to $\mathbb{E}\left[X_1\right]$ and $\operatorname{Var}\left[X_1\right]$, respectively. What can we know about the results?

Solution

1. The marginal pdf of X_1 is:

$$f_2\left(x_2\right)=\begin{cases} \int_0^{x_2} 6x_1 dx_1 = 3x_2^2 & 0<x_2<1\\ 0 & \text{elsewhere}\end{cases}$$

   Then, the conditional pdf of X_1, given $X_2=x_2$, is:

$$f_{1|2}\left(x_1|x_2\right)=\begin{cases} \frac{f(x_2,x_1)}{f_2(x_2)}=\frac{6x_1}{3x_2^2}=\frac{2x_1}{x_2^2} & 0<x_1<x_2\\ 0 & \text{elsewhere}\end{cases}$$

   Therefore,

$$\therefore \mathbb{E}\left[X_1|x_2\right]=\int_0^{x_1} x_2\left(\frac{2x_1}{x_2^2}\right)dx_1=\frac{2}{3}x_2 \qquad 0<x_2<1$$

2. From the above result, $Y=\mathbb{E}\left[X_1|X_2\right]=\frac{2X_2}{3}$. Then the PDF of Y is:

$$F_Y(y)=P[Y\le y]=P\left[X_2\le\frac{3y}{2}\right] \qquad 0\le y<\frac{2}{3}$$

   Using the marginal pdf $f_2\left(x_2\right)$,

$$F_Y(y)=\int_0^{\frac{3y}{2}} 3x_2^2 dx_1=\frac{27y^3}{8} \quad 0\le y<\frac{2}{3}$$

$$\therefore f_Y(y)=\frac{dF_Y(y)}{dy}=\frac{81y^2}{8} \quad 0\le y<\frac{2}{3}$$

3. Using the pdf $f_Y(y)$,

$$\mathbb{E}[Y] = \int_0^{\frac{2}{3}} y \frac{81y^2}{8} dy = \frac{1}{2}$$

$$\operatorname{Var}[Y] = \int_0^{\frac{2}{3}} y^2 \frac{81y^2}{8} dy - \left(\frac{1}{2}\right)^2 = \frac{1}{60}$$

On the other hand, the marginal pdf of X_1 is:

$$f_1(x_1) = \int_{x_1}^{1} 6x_1 dx_2 = 6x_1(1 - x_1) \qquad 0 < x_1 < 1$$

zero elsewhere. Thus, it is easy to show that $\mathbb{E}[X_1] = \frac{1}{2}$ and $\operatorname{Var}[X_1] = \frac{1}{20}$. That is, here

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[X_1|X_2]] = \mathbb{E}[X_1]$$

and

$$\operatorname{Var}[Y] = \operatorname{Var}[\mathbb{E}[X_1|X_2]] \leq \operatorname{Var}[X_1]$$

$\Rightarrow$ This result has the useful interpretation. Both the random variables X_1 and $\mathbb{E}[X_1|X_2]$ have the same mean. If we did not know the mean, we could use either of the two random variables to guess at the mean. Since, however, Var $[\mathbb{E}[X_1|X_2]) \leq \operatorname{Var}[X_1]$, we would put more reliance in $\mathbb{E}[X_1|X_2]$ as a guess. That is we observe x_2, we could prefer to use $\mathbb{E}[X_1|x_2]$ as a guess at the unknown mean of X_1. Indeed, the general estimation approaches based on the incoming measurements are motivated by this idea.

Problem 5

Consider a Gaussian random vector $X = [X_1, X_2]^T$ with expectation and covariance matrix given by:

$$E(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, K := \operatorname{Cov}(X) = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

1. Find the eigenvalues and eigenvectors of K

2. The contours of equal probability density (likelihood ellipse) are given by an equation of the form
$$x^T K^{-1} x = c^2$$

3. Plot the likelihood ellipses for $c = 0.25, 1, 1.5$

4. What is the probability of finding x inside each of these ellipses?

Solution

1. Eigenvalues λ are given by

$$\det(\lambda I - K) = \lambda^2 - 6\lambda + 7 = 0$$

$$\therefore \lambda_1 = 1.5858, \lambda_2 = 4.4141$$

and the corresponding eigenvectors are

$$\det\left(\lambda_1 I - K\right) v_1 = 0 \rightarrow v_1 = [-0.9239 \quad 0.3827]^{\mathrm{T}}$$

$$\det\left(\lambda_2 I - K\right) v_2 = 0 \rightarrow v_2 = [-0.3827 \quad 0.9239]^{\mathrm{T}}$$

2. Let $Q = [v_1 v_2]$ and $x := Qy$. Then, since,

$$Q^{\mathrm{T}} K Q = \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \rightarrow Q^{\mathrm{T}} K^{-1} Q = \Lambda^{-1}$$

we have

$$x^{\mathrm{T}} K^{-1} x = y^{\mathrm{T}} \Lambda^{-1} y = \sum_{i=1}^{2} \frac{y_i^2}{\lambda_i} = c^2$$

which is an ellipse equation. Thus, the directions of eigenvectors are principal axes. Specifically, v_2 is major axis and v_1 in minor axis.

3. The distances from the origin to the ellipsoid in the principal axes directions are $c\sqrt{\lambda_2}$ and $c\sqrt{\lambda_1}$. Therefore, the ellipses for different c are given by

4. The probability of finding x inside each of these ellipses is given by:

$$P\left[x^T K^{-1} x \leq c^2\right] = P\left[y^T \Lambda^{-1} y \leq c^2\right] = 1 - \exp\left(-\frac{c^2}{2}\right)$$

and thus for the differnt c cases,

$$\begin{cases} c = 0.25 \rightarrow P = 3.08\% \\ c = 1 \rightarrow P = 39.35\% \\ c = 1.5 \rightarrow P = 67.53\% \end{cases}$$

Problem 6

Given the three estimates of the scalar x

$$y_i := \hat{x}_i = x + \tilde{x}_i, \quad i = 1, 2, 3$$

with the estimation error $\tilde{x}_i$ jointly Gaussian, zero-mean, with

$$E\left[\tilde{x}_i \tilde{x}_i\right] = P_{ij}, \quad i, j = 1, 2, 3$$

Find

1. The maximum likelihood estimator
2. The variance of the maximum likelihood estimator

Solution

1. The maximum likelihood estimator of x given $y_i, i = 1, 2, 3$ is represented as

$$\begin{aligned}\hat{x}_{ML} &= \arg\max_x P\left[y_1, y_2, y_3 | x\right] \\ &= \arg\max_x \Lambda_Y(x)\end{aligned}$$

where

$$\Lambda_Y(x) = \frac{1}{2\pi^{\frac{3}{2}}\sqrt{|P|}} \exp\left(-\frac{1}{2}\left[y_1 - x \quad y_2 - x y_3 - x\right] P^{-1} \left[y_1 - x \quad y_2 - x \quad y_3 - x\right]^{\mathrm{T}}\right)$$

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

Then, it is equivalnet that

$$\max \Lambda_Y(x) = \min\left[y_1 - x \quad y_2 - x y_3 - x\right] P^{-1} \left[y_1 - x \quad y_2 - x \quad y_3 - x\right]^{\mathrm{T}}$$

Therefore, the given ML estimator can be interpreted as the least-square (LS) estimator as follows

$$\hat{x}_{ML} = \hat{x}_{LS} = \left(H^{\mathrm{TP}^{-1}} H\right)^{-1} H^{\mathrm{T}} P^{-1} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

where $H = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$.

2. The estimation error of the ML estimator, $e := \hat{x}_{ML} - x$, is given by

$$e = \left(H^{\mathrm{T}} P^{-1} H\right)^{-1} H^{\mathrm{T}} P^{-1} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix}$$

Since $\mathbb{E}[e] = 0$,

$$\begin{aligned}\mathrm{Var}(e) = \mathbb{E}\left[e^2\right] &= \mathbb{E}\left[\left(H^{\mathrm{T}} P^{-1} H\right)^{-1} H^{\mathrm{T}} P^{-1} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix}\right. \\ &= \left(H^{\mathrm{T}} P^{-1} H\right)^{-1} H^{\mathrm{T}} P^{-1} \mathbb{E}\left[\left[\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_2 \end{bmatrix} \left[\begin{array}{ccc} \tilde{x}_1 & \tilde{x}_2 & \tilde{x}_3 \end{array}\right]\right] P^{-1} H \left(H^{\mathrm{T}} P^{-1} H\right)^{-1}\right] \\ &= \left(H^{\mathrm{T}} P^{-1} H\right)^{-1} H^{\mathrm{T}} P^{-1} P P^{-1} H \left(H^{\mathrm{T}} P^{-1} H\right)^{-1} \\ &= \left(H^{\mathrm{T}} P^{-1} H\right)^{-1}\end{aligned}$$

Problem 7

Consider a random vector $Y = [y(1) y(2) \cdots y(k)]^{\mathrm{T}}$ where the elements $y(j)$ are made

$$y(j) = x + w(j), \quad j = 1, \cdots, k$$

where $w(j)$ are independent, identically distributed, Gaussian, zero-mean, and with the variance σ^2 , i.e., $\mathcal{N}\left(0, \sigma^2\right)$.

1. Find the Maximum Likelihood (ML) estimator for x, i.e., $\hat{x}_{ML}$
2. Find the Mean Square Error (MSE) of ML estimator, i.e., $\mathrm{MSE}\left(\hat{\mathbf{x}}_{ML}\right) \equiv \mathrm{Var}\left(\hat{x}_{ML}\right)$
3. Is this estimator consistent? Prove your answer
4. Is this estimator efficient? Prove your answer

Solution

1. The ML estimator is given by

$$\hat{x}_{ML} = \arg\max_x P(Y|x) = \arg\max_x \Lambda_Y(x)$$

where the likelihood function $\Lambda_Y(x)$ is

$$\Lambda_Y(x) = \frac{1}{2\pi^{k/2}\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}[Y - \overline{1}x]^{\mathrm{T}}\Sigma^{-1}[Y - \overline{1}x]\right\}$$

where $\Sigma \in \mathbb{R}^{k\times k}$ and $\overline{1} \in \mathbb{R}^k$ are respectively given as

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 I, \quad \overline{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Then,

$$\begin{aligned} \arg\max_x \Lambda_Y(x) &= \arg\min_x \left(\frac{1}{2}[Y - \overline{1}x]^{\mathrm{T}}\Sigma^{-1}[Y - \overline{1}x]\right) \\ &= \arg\min_x \left(\frac{1}{2\sigma^2}[Y - \overline{1}x]^{\mathrm{T}}[Y - \overline{1}x]\right) \end{aligned}$$

From the right-hand-side term of the above equation, ML estimator can be interpreted as the exactly same formulation as the Least Square (LS) estimator whose minimum solution is

$$\hat{x}_{ML} = \hat{x}_{LS} = \left(\overline{1}^{\mathrm{T}}\overline{1}\right)^{-1}\overline{1}^{\mathrm{T}}Y = \frac{1}{k}\sum_{j=1}^{k} y(j)$$

2. MSE of the $\hat{x}_{ML}$ is defined as $\mathrm{MSE}\left(\hat{\mathrm{x}}_{\mathrm{ML}}\right) := \mathbb{E}\left[\left(\hat{x}_{ML} - x\right)^2\right]$. From the previous result,

$$\begin{aligned} \mathbb{E}\left[\left(\hat{x}_{ML} - x\right)^2\right] &= \mathbb{E}\left[\left(\frac{1}{k}\sum_{j=1}^{k} y(j) - x\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{1}{k}\sum_{j=1}^{k}(x + w(j)) - x\right)^2\right] = \mathbb{E}\left[\left(\frac{1}{k}\sum_{j=1}^{k} w(j)\right)^2\right] \\ &= \frac{\sigma^2}{k} \end{aligned}$$

3. Check the consistency:

$$\lim_{k\to\infty} \mathrm{MSE}\left(\hat{\mathrm{x}}_{\mathrm{ML}}\right) = \lim_{k\to\infty} \frac{\sigma^2}{k} = 0$$

$\therefore$ The given ML estimator is consistent so that the solution is getting more accurate as we take more measurements.

4. Check the efficiency: Cramer-Rao Lower Bound (CRLB) for the MSE of the given ML estimator is

$$\mathrm{MSE}\left(\hat{\mathbf{x}}_{ML}\right) \geq J^{-1}$$

where

$$\begin{aligned} J &:= \mathbb{E}\left[\left(\frac{\partial \ln \Lambda_Y(x)}{\partial x}\right)^2\right] = \mathbb{E}\left[\left(-\frac{1}{2\sigma^2}\frac{\partial[Y-\bar{1}x]^{\mathrm{T}}[Y-\bar{1}x]}{\partial x}\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{2\bar{1}^{\mathrm{T}}\bar{1}x - 2\bar{1}^{\mathrm{T}}Y}{2\sigma^2}\right)^2\right] = \mathbb{E}\left[\left(\frac{2\bar{1}^{\mathrm{T}}\bar{1}x - 2\bar{1}^{\mathrm{T}}\bar{1}x - 2\sum_{j=1}^{k} w(j)}{2\sigma^2}\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{-\sum_{j=1}^{k} w(j)}{\sigma^2}\right)^2\right] = \frac{k\sigma^2}{\sigma^4} = \frac{k}{\sigma^2} \end{aligned}$$

Therefore,

$$\mathrm{MSE}\left(\hat{\mathrm{x}}_{\mathrm{ML}}\right) = \frac{\sigma^2}{k} = J^{-1}$$

$\therefore$ Since the estimator's MSE(variance) is equal to its CRLB, the given ML estimator is efficient.