Stochastic Processes

Midterm 2

Solutions

Total: 100 points

1. (10+10=20 points)

   (a) Suppose the state of nature is $\Omega = \{x_1, \cdots, x_N\}$. The MAP decision rule based on the observation $Y = y$ is given by

$$\begin{aligned}\hat{X}_{\text{MAP}} &= \arg\max_{x_j} P[X = x_j \mid Y = y] \\ &= \arg\max_{x_j} f_Y(y \mid X = x_j) P[X = x_j].\end{aligned}$$

   When $c_{i,j} = 1 - \delta_{i,j}$, the average cost is

$$\begin{aligned}C(D) &\triangleq E[C] \\ &= \sum_i \sum_j c_{i,j} P[\text{decide } H_i \text{ and } H_j \text{ is true}] \\ &= \sum_i \sum_j (1 - \delta_{i,j}) P[\text{decide } H_i \text{ and } H_j \text{ is true}] \\ &= \sum_{i \neq j} P[\text{decide } H_i \text{ and } H_j \text{ is true}],\end{aligned}$$

   which is exactly the probability of decision error. Bayes detection rule guarantees that this decision error probability is minimized.

   (b) The ML rule states the following decision

$$f_{\mathbf{y}}(\mathsf{y}|\alpha = 1) \underset{\hat{\alpha}=-1}{\overset{\hat{\alpha}=1}{\gtrless}} f_{\mathbf{y}}(\mathsf{y}|\alpha = -1).$$

   After certain manipulations, the maximum likelihood decision rule is

$$\mathsf{y}^T \mathbf{K}_n^{-1} \mathbf{h} \underset{\hat{\alpha}=-1}{\overset{\hat{\alpha}=1}{\gtrless}} 0.$$

   And, the probability of error is

$$P_e = P(\text{decide } 1|\alpha = -1)P(\alpha = -1) + P(\text{decide -}1|\alpha = 1)P(\alpha = 1).$$

   Further algebraic effort gives

$$\begin{aligned}P(\text{decide } 1|\alpha = -1) &= P(\mathsf{y}^T \mathbf{K}_n^{-1} \mathbf{h} > 0|\alpha = -1) \\ &= P((-\mathbf{h} + \mathbf{n})^T \mathbf{K}_n^{-1} \mathbf{h} > 0) \\ &= Q\left(\sqrt{\mathbf{h}^T \mathbf{K}_n^{-1} \mathbf{h}}\right).\end{aligned}$$

   It can be found that $P(\text{decide -}1|\alpha = 1) = P(\text{decide } 1|\alpha = -1)$. Thus, we have

$$P_e = Q\left(\sqrt{\mathbf{h}^T \mathbf{K}_n^{-1} \mathbf{h}}\right).$$

2. ($10 \times 5 = 50$ points)

(a) We first find the moment generating function $M_{Z_i^2}$ of Z_i^2

$$\begin{aligned} M_{Z_i^2} &= E[e^{tZ_i^2}] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}} e^{tz^2} dz \\ &= (1-2t)^{\frac{-1}{2}}. \end{aligned}$$

Then, the moment generating function of Y is obtained as

$$\begin{aligned} M_Y(t) &= E[e^{tY}] \\ &= E[e^{t\sum_{i=1}^n Z_i^2}] \\ &= \prod_{i=1}^n E[e^{tZ_i^2}] \quad \text{since } Z_i\text{'s for } i = 1\ldots n \text{ are indep.} \\ &= (1-2t)^{\frac{-n}{2}}. \end{aligned}$$

(b) The mean of Y is

$$\begin{aligned} E[Y] &= E\left[\sum_{i=1}^n Z_i^2\right] \\ &= \sum_{i=1}^n E[Z_i^2] \\ &= n \quad (\text{since } Z_i \sim \mathcal{N}(0,1), \text{ and therefore } E[Z_i^2] = 1). \end{aligned}$$

We need the 2nd moment of Y in order to find the variance of it. The 2nd moment is given by

$$\begin{aligned} E[Y^2] &= E\left[\left(\sum_{i=1}^n Z_i^2\right)^2\right] \\ &= E\left[\sum_{i=1}^n Z_i^4\right] + E\left[\sum_{i\neq j} Z_i^2 Z_j^2\right] \\ &= \sum_{i=1}^n E\left[Z_i^4\right] + \sum_{i\neq j} E\left[Z_i^2\right]\left[Z_j^2\right] \\ &= 3n + n(n-1) = n^2 + 2n. \end{aligned}$$

Therefore, the variance of Y is obtained as

$$\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 = 2n.$$

(c) For each deviation $X_i - \bar{X}_n$, we have

$$\begin{aligned} E\left[\bar{X}_n(X_i - \bar{X}_n)\right] &= E[\bar{X}_n X_i] - E[\bar{X}_n^2] \\ &= E\left[\frac{1}{n}\sum_{j=1}^n X_j X_i\right] - \frac{1}{n^2}\sum_{j=1}^n E[X_j^2] \\ &= \frac{1}{n}E[X_j^2] - \frac{1}{n}E[X_j^2] \\ &= 0, \end{aligned}$$

which equals to $E\left[\bar{X}_n\right] E\left[(X_i - \bar{X}_n)\right]$. Therefore, we conclude that $\bar{X}_n$ and $X_i - \bar{X}_n$ are uncorrelated, thus independent for Gaussian sample, for all i. We therefore can conclude that the sample mean is independent with the sample variance.

(d) First, it is clear to see $E[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \mu$.
Second, we can use Chebyshev inequality to examine consistency of any unbiased estimators. Here, we have

$$P\left[\,\left|\bar{X}_n - \mu\right| > k\right] \leq \frac{\mathrm{Var}(\bar{X}_n)}{k^2}, \quad \text{for any} \quad k > 0. \tag{1}$$

The variance of the sample mean is

$$\begin{aligned}\mathrm{Var}(\bar{X}_n) &= \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\mathrm{Var}(\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{1}{n}\sigma^2.\end{aligned}$$

Therefore, by plugging the above result into (1), we can easily show the consistency of sample mean, i.e. for any $k > 0$,

$$\lim_{n\to\infty} P\left[\,\left|\bar{X}_n - \mu\right| > k\right] = 0.$$

(e) This is Example 4.8-1 in the textbook. We need to find the corresponding z in the following to specify a confidence interval with 90 percent confidence.

$$P\left[-z \leq \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \leq z\right] = 0.9,$$

where in this example $n = 21$, $\bar{X}_{21} = 3.5$ and $S_{21}/\sqrt{21} = 0.45$. The probability can be evaluated as

$$[-z \leq T_{20} \leq z] = F_T(z, 20) - F_T(-z, 20) = 2F_T(z, 20) - 1 = 0.9.$$

We therefore have $F_T(z, 20) = 0.95$, in which the value z can be looked up in the table given in the cover page. Thus, $z_{.95} = 1.725$.
And the confidence interval is given by

$$\begin{aligned}\left[\bar{X}_n - \frac{S_n z}{\sqrt{n}} \leq \mu \leq \bar{X}_n + \frac{S_n z}{\sqrt{n}}\right] &= [3.5 - 1.725 \cdot 0.45 \leq \mu \leq 3.5 + 1.725 \cdot 0.45] \\ &= [2.72, 4.28].\end{aligned}$$

3. ($10 \times 3 = 30$ points)

(a) We can show that $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian. Thus, the MMSE estimator is given by the conditional expectation

$$\begin{aligned}\hat{\mathbf{x}}_{mmse} &= E\Big[\mathbf{x}|\mathbf{y} = \boldsymbol{y}\Big] \\ &= \mathbf{m}_x + \mathbf{K}_{\mathbf{xy}}\mathbf{K}_y^{-1}(\boldsymbol{y} - \mathbf{m}_y) \\ &= \mathbf{K}_{xy}\mathbf{K}_y^{-1}\boldsymbol{y} \qquad (\text{since } \mathbf{m}_x = \mathbf{m}_y = \mathbf{0}) \\ &= \mathbf{K}_x\mathbf{H}^T\big(\mathbf{H}\mathbf{K}_x\mathbf{H}^T + \mathbf{K}_z\big)^{-1}\boldsymbol{y},\end{aligned}$$

where the cross-covariance matrix and the covariance are respectively

$$\begin{aligned}\mathbf{K}_{xy} &= E[\mathbf{x}\mathbf{y}^T] \\ &= E\Big[\mathbf{x}(\mathbf{H}\mathbf{x} + \mathbf{z})^T\Big] \\ &= \mathbf{K}_x\mathbf{H}^T \\ \mathbf{K}_y &= E[\mathbf{y}\mathbf{y}^T] \\ &= E\Big[(\mathbf{H}\mathbf{x} + \mathbf{z})(\mathbf{H}\mathbf{x} + \mathbf{z})^T\Big] \\ &= \mathbf{H}\mathbf{K}_x\mathbf{H}^T + \mathbf{K}_z.\end{aligned}$$

(b) The mean squared error for the MMSE estimator is

$$\begin{aligned}
\text{MSE} &= E\Big[||\mathbf{x} - \hat{\mathbf{x}}_{mmse}||^2\Big] \\
&= E\Big[\text{tr}\left\{\left(\mathbf{x} - \mathbf{K}_{xy}\mathbf{K}_y^{-1}\mathbf{y}\right)\left(\mathbf{x} - \mathbf{K}_{xy}\mathbf{K}_y^{-1}\mathbf{y}\right)^T\right\}\Big] \\
&= \text{tr}\left\{\mathbf{K}_x - \mathbf{K}_{xy}\mathbf{K}_y^{-1}\mathbf{K}_{yx}\right\} \\
&= \text{tr}\left\{\mathbf{K}_x - \mathbf{K}_x\mathbf{H}^T\left(\mathbf{H}\mathbf{K}_x\mathbf{H}^T + \mathbf{K}_z\right)^{-1}\mathbf{H}\mathbf{K}_x\right\}.
\end{aligned}$$

(c) This is the orthogonality principle we've discussed in class, and can be shown as follows.

$$\begin{aligned}
E\Big[\Big(\mathbf{x} - E[\mathbf{x}|\mathbf{y}]\Big) \cdot k^T(\mathbf{y})\Big] &= E\left[\mathbf{x}k^T(\mathbf{y})\right] - E\Big[E\left[\mathbf{x}|\mathbf{y}\right]k^T(\mathbf{y})\Big] \\
&= E\left[\mathbf{x}k^T(\mathbf{y})\right] - E\Big[E\left[\mathbf{x}k^T(\mathbf{y})|\mathbf{y}\right]\Big] \\
&= E\left[\mathbf{x}k^T(\mathbf{y})\right] - E\left[\mathbf{x}k^T(\mathbf{y})\right] \\
&= \mathbf{0}.
\end{aligned}$$

■