EE 511 Solutions to Problem Set 4

1. (i) $\phi_X(s) = e^{s^2/2}$. (ii) We have

$$P[X \geq a] = e^{-as}\phi_X(s) \quad \text{for all } s > 0.$$

This upper bound should be minimized with respect to s to obtain the Chernoff bound.

$$e^{-as}\phi_X(s) = e^{-as}e^{s^2/2}$$

Setting the derivative with respect to s to 0, we get

$$se^{-as}e^{s^2/2} + (-a)e^{-as}e^{s^2/2} = 0$$

i.e., $s = a$. The second derivative at $s = a$ can be shown to be positive. Therefore, the Chernoff bound is given by

$$P[X \geq a] \leq e^{-a^2/2}$$

(iii) From the Chebyshev inequality, we get

$$P[|X| \geq a] \leq \frac{1}{a^2}.$$

Since $f_X(x)$ is symmetric, we get

$$P[X \geq a] \leq \frac{1}{2a^2}.$$

2. $E[Z] = E[X] + aE[Y] = 0$.

$$\begin{aligned} E[X|Y = y] &= E[Z|Y = y] - aE[Y|Y = y] \\ &= E[Z] - ay \\ &= -ay. \end{aligned}$$

3.

$$E[X] = \int_0^{100} x f_X(x) dx = \int_0^{100} \frac{x}{100} dx = 50.$$

Given that $X \geq 65$, X is uniformly distributed in $[65, 100]$. Therefore, we have

$$E[X|X \geq 65] = \int_{65}^{100} x f_X(x|X \geq 65) dx = \int_{65}^{100} \frac{x}{35} dx = 82.5.$$

4. $E[X] = \sum_{k=0}^{\infty} \frac{ke^{-a}a^k}{k!}$. We know

$$\sum_{k=0}^{\infty} \frac{e^{-a}a^k}{k!} = 1. \tag{1}$$

Differentiating with respect to a, we get

$$\sum_{k=0}^{\infty} \frac{ke^{-a}a^{k-1}}{k!} - e^{-a}\sum_{k=0}^{\infty} \frac{a^k}{k!} = 0.$$

$$\frac{1}{a}E[X] = e^{-a}\sum_{k=0}^{\infty}\frac{a^k}{k!} = 1.$$

Therefore, we have $E[X] = a$.

Differentiating (1) twice with respect to a, we get

$$\sum_{k=0}^{\infty}\frac{k(k-1)e^{-a}a^{k-2}}{k!} - \sum_{k=0}^{\infty}\frac{ke^{-a}a^{k-1}}{k!} - \sum_{k=0}^{\infty}\frac{ke^{-a}a^{k-1}}{k!} + \sum_{k=0}^{\infty}\frac{e^{-a}a^{k}}{k!} = 0$$

$$\frac{1}{a^2}\sum_{k=0}^{\infty}\frac{k^2e^{-a}a^{k}}{k!} - \frac{1}{a^2}\sum_{k=0}^{\infty}\frac{ke^{-a}a^{k}}{k!} - \frac{2}{a}\sum_{k=0}^{\infty}\frac{ke^{-a}a^{k}}{k!} + 1 = 0$$

$$\frac{1}{a^2}E[X^2] - \frac{1}{a^2}a - \frac{2}{a}a + 1 = 0.$$

Therefore, we have $E[X^2] = a^2 + a \Rightarrow Var(X) = a$.

5.

$$E[X] = \int_0^{\infty} x\lambda e^{-\lambda x}dx = \int_0^{\infty} xd(-e^{-\lambda x}) = -xe^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}dx = 0 + \frac{e^{-\lambda x}}{-\lambda}\Big|_0^{\infty} = \frac{1}{\lambda}.$$

$$f_X(x|X \geq 2) = \begin{cases} 0 & x < 2 \\ \frac{f_X(x)}{P[X\geq 2]} & x \geq 2 \end{cases}$$

$$P[X \geq 2] = \int_2^{\infty} \lambda e^{-\lambda x}dx = \frac{e^{-\lambda x}}{-\lambda}\Big|_2^{\infty} = e^{-2\lambda}.$$

Therefore, we have

$$f_X(x|X \geq 2) = \begin{cases} 0 & x < 2 \\ \lambda e^{-\lambda(x-2)} & x \geq 2 \end{cases}$$

$$E[X|X \geq 2] = \int_2^{\infty} x\lambda e^{-\lambda(x-2)}dx = -xe^{-\lambda(x-2)}\Big|_2^{\infty} + \int_2^{\infty} e^{-\lambda(x-2)}dx = 2 + \frac{e^{-\lambda(x-2)}}{-\lambda}\Big|_2^{\infty} = 2 + \frac{1}{\lambda}.$$

6. (i) $MSE(c) = E[(Y-c)^2] = \int_{-\infty}^{\infty}(y-c)^2 f_Y(y)dy$. Setting the derivative of $MSE(c)$ with respect to c to be 0, we get

$$\frac{dMSE(c)}{dc} = -\int_{-\infty}^{\infty} 2(y-c)f_Y(y)dy = 0$$

i.e.,

$$c = \int_{-\infty}^{\infty} yf_Y(y)dy = E[Y].$$

Also,

$$\frac{d^2MSE(c)}{dc^2} = \int_{-\infty}^{\infty} 2f_Y(y)dy = 2 > 0.$$

(ii) $E[(Y-g(X))^2] = E[E[(Y-g(X))^2|X]]$, i.e.,

$$E[(Y-g(X))^2] = \int_{\infty}^{\infty} E[(Y-g(X))^2|X=x] f_X(x)dx.$$

Since $f_X(x) \geq 0$, we minimize $E[(Y-g(X))^2]$ by minimizing $E[(Y-g(X))^2|X=x]$ for each x.

$$E[(Y-g(X))^2|X=x] = \int_{\infty}^{\infty} (y-g(x))^2 f_Y(y|X=x)dy.$$

As in part (i), the best choice for $g(x)$ is

$$g(x) = \int_{-\infty}^{\infty} y f_Y(y|X=x)dy = E[Y|X=x].$$

7. Since X is a zero-mean Gaussian with variance σ^2

$$\phi_x(s) = e^{\frac{s^2\sigma^2}{2}} = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{s^2\sigma^2}{2}\right)^k.$$

$$E[X^n] = \left.\frac{\partial^n \phi_X(s)}{\partial s^n}\right|_{s=0}.$$

Therefore, $E[X^n] = 0$ when n is odd. When n is even and $n = 2m$, we have

$$\left.\frac{\partial^n \phi_X(s)}{\partial s^n}\right|_{s=0} = \frac{(2m)!}{m!2^m}\sigma^{2m} = (1.3.\cdots(2m-3).(2m-1))\,\sigma^{2m}.$$

8. a)

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{3/2}|C|^{1/2}} \exp\left\{-\frac{1}{2}\underline{x}^T C^{-1}\underline{x}\right\}$$

where $|C| = 36$ and

$$C^{-1} = \frac{1}{36}\begin{bmatrix} 30 & -18 & 0 \\ -18 & 18 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

b) $E[Y] = 0$ and

$$\sigma_Y^2 = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = 41.$$

Y is Gaussian with zero-mean and variance 41.

c) $E[\underline{Z}] = \underline{0}$ and

$$C_{\underline{Z}} = \begin{bmatrix} 5 & -3 & -1 \\ -1 & 3 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 & 0 \\ 3 & 5 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -1 \\ -3 & 3 & 0 \\ -1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 36 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$\underline{Z}$ is a 3-dimensional zero-mean Gaussian random vector with covariance matrix $C_{\underline{Z}}$.

9. (a) X_2 is a zero-mean Gaussian with variance 2.

$$f_{X_2}(x_2) = \frac{1}{\sqrt{4\pi}} \exp\left\{-\frac{x_2^2}{4}\right\}.$$

(b) $f_{X_1}(x_1|X_2 = x_2) = \dfrac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}$.

$$\det(C) = 2 - r^2 \quad \text{and} \quad C^{-1} = \frac{1}{2-r^2}\begin{bmatrix} 2 & -r \\ -r & 1 \end{bmatrix}.$$

Therefore, we have

$$\begin{aligned} f_{X_1}(x_1|X_2 = x_2) &= \frac{\frac{1}{2\pi\sqrt{2-r^2}} \exp\left\{-\frac{1}{2(2-r^2)}(2x_1^2 - 2rx_1x_2 + x_2^2)\right\}}{\frac{1}{\sqrt{4\pi}} \exp\left\{-\frac{x_2^2}{4}\right\}} \\ &= \frac{\sqrt{2}}{\sqrt{2\pi}\sqrt{2-r^2}} \exp\left\{-\frac{x_1^2 - rx_1x_2 + \frac{r^2x_2^2}{4}}{2-r^2}\right\} \\ &= \frac{\sqrt{2}}{\sqrt{2\pi}\sqrt{2-r^2}} \exp\left\{-\frac{(x_1 - \frac{rx_2}{2})^2}{2-r^2}\right\} \end{aligned}$$

Given $X_2 = x_2$, X_1 is Gaussian with mean $\frac{rx_2}{2}$ and variance $1 - \frac{r^2}{2}$.

10. (a)

$$\phi_{\underline{X}}(\underline{s}) = \exp\left\{\underline{s}^T\underline{m} + \frac{1}{2}\underline{s}^T C\underline{s}\right\}.$$

$$\phi_{X_1}(s_1) = \phi_{\underline{X}}(\underline{s})\big|_{s_2=0} = \exp\left\{m_1 s_1 + \frac{\sigma_2^2 s_1^2}{2}\right\}$$

Therefore, X_1 is Gaussian. Similarly, X_2 can be shown to be Gaussian.

(b)

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1,x_2)dx_2 \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_1^2}{2}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_2^2}{2}\right\} dx_2 \\ &\quad + \frac{x_1}{\sqrt{2\pi}} \exp\left\{-\frac{x_1^2 - 2}{2}\right\} \int_{-\infty}^{\infty} \frac{x_2}{\sqrt{2\pi}} \exp\left\{-\frac{x_2^2}{2}\right\} dx_2 \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_1^2}{2}\right\} + 0 \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_1^2}{2}\right\} \end{aligned}$$

Similarly, we can show that

$$f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_2^2}{2}\right\}.$$

11. $E[\underline{Y}] = AE[\underline{X}] + \underline{b}$. Let C_X and C_Y denote the covariance matrices of $\underline{X}$ and $\underline{Y}$ respectively.

$$\begin{aligned} C_Y &= E[(\underline{Y} - E[\underline{Y}])(\underline{Y} - E[\underline{Y}])^T] \\ &= E[(A(\underline{X} - E[\underline{X}]))(A(\underline{X} - E[\underline{X}]))^T] \\ &= AC_X A^T \end{aligned}$$

12. $\underline{X}$ is proper if all the elements of $E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T]$ are zero.

$$\underline{X} - E[\underline{X}] = (\underline{X}_r - E[\underline{X}_r]) + j(\underline{X}_i - E[\underline{X}_i])$$

$$\begin{aligned} E[(\underline{X} - E[\underline{X}])(\underline{X} - E[\underline{X}])^T] &= E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_r - E[\underline{X}_r])^T] \\ &\quad + jE[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T] \\ &\quad + jE[(\underline{X}_i - E[\underline{X}_i])(\underline{X}_r - E[\underline{X}_r])^T] \\ &\quad - E[(\underline{X}_i - E[\underline{X}_i])(\underline{X}_i - E[\underline{X}_i])^T] \end{aligned}$$

Therefore, we need

$$E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_r - E[\underline{X}_r])^T] = E[(\underline{X}_i - E[\underline{X}_i])(\underline{X}_i - E[\underline{X}_i])^T],$$

and

$$E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T] = -E[(\underline{X}_i - E[\underline{X}_i])(\underline{X}_r - E[\underline{X}_r])^T].$$

Since $E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T] = E[(\underline{X}_i - E[\underline{X}_i])(\underline{X}_r - E[\underline{X}_r])^T]$, we have

$$E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T] = -E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T]^T.$$

This means that the diagonal elements of $E[(\underline{X}_r - E[\underline{X}_r])(\underline{X}_i - E[\underline{X}_i])^T]$ are zero, i.e., the real and imaginary part of each component in $\underline{X}$ are uncorrelated. Thus, the required conditions are:

(i) The vectors $\underline{X}_r$ and $\underline{X}_i$ should have the same covariance matrix.

(ii) The vectors $\underline{X}_r$ and $\underline{X}_i$ should have a cross-covariance matrix that is skew-symmetric.

13.

$$\underline{s}^T \underline{m} = \sum_{i=1}^{n} s_i m_i \quad \text{and} \quad \underline{s}^T C \underline{s} = \sum_{i=1}^{n} s_i \left(\sum_{j=1}^{n} C_{ij} s_j \right)$$

$$E[X_k] = \left. \frac{\partial \phi_{\underline{X}}(\underline{s})}{\partial s_k} \right|_{\underline{s}=\underline{0}}$$

$$\frac{\partial \phi_{\underline{X}}(\underline{s})}{\partial s_k} = \phi_{\underline{X}}(\underline{s}) \left[m_k + C_{kk} s_k + \frac{1}{2} \sum_{j=1, j\neq k}^{n} (C_{kj} + C_{jk}) s_j \right]$$

Therefore, we have

$$E[X_k] = m_k \quad \text{and} \quad E[\underline{X}] = \underline{m}.$$

$$R = E[\underline{X}\underline{X}^T] = C + \underline{m}\underline{m}^T.$$

$$R_{kl} = \frac{\partial}{\partial s_k}\frac{\partial}{\partial s_l}\phi_{\underline{X}}(\underline{s})\Big|_{\underline{s}=\underline{0}}$$

$$\begin{aligned}
R_{kk} &= \left\{\phi_{\underline{X}}(\underline{s})\left[C_{kk} + \left(m_k + C_{kk}s_k + \frac{1}{2}\sum_{j=1,j\neq k}^{n}(C_{kj}+C_{jk})s_j\right)^2\right]\right\}_{\underline{s}=\underline{0}} \\
&= C_{kk} + m_k^2 \\
R_{kl} &= \left\{\phi_{\underline{X}}(\underline{s})\left[\frac{1}{2}(C_{kl}+C_{lk}) + \left(m_k + C_{kk}s_k + \frac{1}{2}\sum_{j=1,j\neq k}^{n}(C_{kj}+C_{jk})s_j\right)\right.\right. \\
&\quad \left.\left.\left(m_l + C_{ll}s_l + \frac{1}{2}\sum_{j=1,j\neq l}^{n}(C_{lj}+C_{jl})s_j\right)\right]\right\}_{\underline{s}=\underline{0}} \\
&= \frac{1}{2}(C_{kl}+C_{lk}) + m_k m_l \\
&= C_{kl} + m_k m_l
\end{aligned}$$

Therefore, the covariance matrix of $\underline{X}$ is C.

14. The covariance matrix C_Y of $[Y_1 \quad Y_2]^T$ is

$$C_Y = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

$$C_Y = \begin{bmatrix} 2(1+\rho) & 0 \\ 0 & 2(1-\rho) \end{bmatrix}.$$

Therefore, Y_1 and Y_2 are uncorrelated. Since they are also jointly Gaussian, they are independent.

15. The covariance matrix C_Y of $[Y_1 \quad Y_2]^T$ is

$$C_Y = \begin{bmatrix} \frac{1}{\sigma_1} & \frac{1}{\sigma_2} \\ \frac{1}{\sigma_1} & -\frac{1}{\sigma_2} \end{bmatrix}\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\begin{bmatrix} \frac{1}{\sigma_1} & \frac{1}{\sigma_2} \\ \frac{1}{\sigma_1} & -\frac{1}{\sigma_2} \end{bmatrix}.$$

$$C_Y = \begin{bmatrix} 2(1+\rho) & 0 \\ 0 & 2(1-\rho) \end{bmatrix}.$$

Therefore, Y_1 and Y_2 are uncorrelated. Since they are also jointly Gaussian, they are independent.