Jared Yu 914640019

HW 4 STA 135

1. Given a pair of jointly distributed random vectors $\vec{X}$ and $\vec{Y}$, it will be shown that,

$$Cov(\boldsymbol{C}\vec{X}, \boldsymbol{D}\vec{Y}) = \boldsymbol{C}Cov(\vec{X}, \vec{Y})\boldsymbol{D}^T.$$

$$\begin{aligned} Cov(\boldsymbol{C}\vec{X}, \boldsymbol{D}\vec{Y}) &= E\left[\left(\boldsymbol{C}\vec{X} - E(\boldsymbol{C}\vec{X})\right)\left(\boldsymbol{D}\vec{Y} - E(\boldsymbol{D}\vec{Y})\right)^T\right] \\ &= \boldsymbol{C}E\left[\left(\vec{X} - E(\vec{X})\right)\left(\vec{Y} - E(\vec{Y})\right)\right]\boldsymbol{D}^T \\ &= \boldsymbol{C}Cov(\vec{X}, \vec{Y})\boldsymbol{D}^T \;\blacksquare \end{aligned}$$

2. Given mutually independent random vectors $\vec{X}_1, \cdots, \vec{X}_n \in \mathbb{R}^p$, it will be shown that,

$$Cov(a_1\vec{X}_1 + \cdots + a_n\vec{X}_n + \vec{c}) = a_1^2 Cov(\vec{X}_1) + \cdots + a_n^2 Cov(\vec{X}_n).$$

$$Cov(a_1\vec{X}_1 + \cdots + a_n\vec{X}_n + \vec{c}) = Cov(\Sigma_{i=1}^n a_i\vec{X}_i + \vec{c})$$

$$= E\left\{\left[(\Sigma_{i=1}^n a_i\vec{X}_i + \vec{c}) - E(\Sigma_{i=1}^n a_i\vec{X}_i + \vec{c})\right]\left[(\Sigma_{j=1}^n a_j\vec{X}_j + \vec{c}) - E(\Sigma_{j=1}^n a_j\vec{X}_j + \vec{c})\right]^T\right\}$$

$$= E\left\{\left[\Sigma_{i=1}^n a_i\vec{X}_i - E(\Sigma_{i=1}^n a_i\vec{X}_i)\right]\left[\Sigma_{j=1}^n a_j\vec{X}_j - E(\Sigma_{j=1}^n a_j\vec{X}_j)\right]^T\right\}$$

$$= E\left\{\left[\Sigma_{i=1}^n a_i\vec{X}_i - \Sigma_{i=1}^n a_i E(\vec{X}_i)\right]\left[\Sigma_{j=1}^n a_j\vec{X}_j - \Sigma_{j=1}^n a_j E(\vec{X}_j)\right]^T\right\}$$

$$= E\left\{\left[\Sigma_{i=1}^n a_i\left(\vec{X}_i - E(\vec{X}_i)\right)\right]\left[\Sigma_{j=1}^n a_j\left(\vec{X}_j - E(\vec{X}_j)\right)\right]^T\right\}$$

$$= E\left\{\Sigma_{i=1}^n \Sigma_{j=1}^n a_i a_j \left(\vec{X}_i - E(\vec{X}_i)\right)\left(\vec{X}_j - E(\vec{X}_j)\right)^T\right\}$$

$$= E\left\{\sum_{\substack{i,j=1 \\ i=j}}^n a_i a_j \left(\vec{X}_i - E(\vec{X}_i)\right)\left(\vec{X}_j - E(\vec{X}_j)\right)^T + \sum_{\substack{i,j=1 \\ i\neq j}}^n a_i a_j \left(\vec{X}_i - E(\vec{X}_i)\right)\left(\vec{X}_j - E(\vec{X}_j)\right)^T\right\}$$

$$= \sum_{\substack{i,j=1 \\ i=j}}^n a_i a_j\, E\left[\left(\vec{X}_i - E(\vec{X}_i)\right)\left(\vec{X}_j - E(\vec{X}_j)\right)^T\right] + \sum_{\substack{i,j=1 \\ i\neq j}}^n a_i a_j\, E\left[\left(\vec{X}_i - E(\vec{X}_i)\right)\left(\vec{X}_j - E(\vec{X}_j)\right)^T\right]$$

$$= \sum_{\substack{i,j=1 \\ i=j}}^n a_i a_j\, Cov(\vec{X}_i, \vec{X}_j) + \sum_{\substack{i,j=1 \\ i\neq j}}^n a_i a_j\, Cov(\vec{X}_i, \vec{X}_j)$$

$\because \vec{X}_1, \cdots, \vec{X}_n$ are mutually independent, $Cov(\vec{X}_i, \vec{X}_j) = 0$ when $i \neq j$

$$= \sum_{\substack{i,j=1 \\ i=j}}^n a_i a_j\, Cov(\vec{X}_i, \vec{X}_j)$$

$$= \sum_{i=1}^n a_i^2\, Cov(\vec{X}_i, \vec{X}_i) = \sum_{i=1}^n a_i^2\, Cov(\vec{X}_i) = a_1^2 Cov(\vec{X}_1) + \cdots + a_n^2 Cov(\vec{X}_n) \;\blacksquare$$

3. Given $\vec{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix} \sim \mathcal{N}(\vec{0}, \boldsymbol{I}_p)$, it will be shown that:

$$\vec{Z}^T\vec{Z} = Z_1^2 + \cdots + Z_p^2 \sim \chi_p^2.$$

First, it will be shown that $\vec{Z}^T\vec{Z} = Z_1^2 + \cdots + Z_p^2$:

$$\vec{Z}^T\vec{Z} = [Z_1 \quad \cdots \quad Z_p] \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix}$$

$$= Z_1^2 + \cdots + Z_p^2.$$

It first can be mentioned that Z_i are uncorrelated. The reason is that the variance-covariance matrix for $\vec{Z}$ is $\boldsymbol{I}_p$. The variance-covariance matrix shows that the covariance is zero for Z_i and Z_j where $i \neq j$ and one when $i = j$. The variance-covariance matrix $\boldsymbol{\Sigma}$ would appear as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,p} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,p} & \sigma_{2,p} & \cdots & \sigma_{p,p} \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix},$$

where $\sigma_{i,j}$ is $Cov(Z_i, Z_j)$ for $i, j = 1, \cdots, p$. This shows that the Z_i and Z_j have zero correlation for $i \neq j$, or in other words they are uncorrelated in such a case. Typically, this is not enough to show independence. It is the case the independent random variables have zero correlation, but it is not always the case vice-versa. However, the exception applies here where the random variables are normally distributed. Therefore, it can be said that since $Z_i's$ are uncorrelated and normally distributed, they are independent.

Next, given that $Z_i \sim \mathcal{N}(0,1)$, then let $Y_i = Z_i^2$. The goal is to show that $Y_i \sim \chi_1^2$. Below is a calculation of the C.D.F. of Y_i:

$$F_{Y_i}(y) = P(Y_i \leq y) = P(Z_i^2 \leq y) = P(-\sqrt{y} \leq Z_i \leq \sqrt{y}) = P(Z_i \leq \sqrt{y}) - P(Z_i \leq -\sqrt{y})$$
$$= \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

The p.d.f. then can be shown as follows:

$$f_{Y_i}(y) = \frac{d}{dy}F_{Y_i}(y) = \frac{d}{dy}\Phi(\sqrt{y}) - \frac{d}{dy}\Phi(-\sqrt{y}) = \frac{d}{d\sqrt{y}}\Phi(\sqrt{y})\frac{d\sqrt{y}}{dy} - \frac{d}{d(-\sqrt{y})}\Phi(-\sqrt{y})\frac{d(-\sqrt{y})}{dy}$$
$$= \phi(\sqrt{y})\left(\frac{1}{2\sqrt{y}}\right) - \phi(-\sqrt{y})\left(\frac{1}{-2\sqrt{y}}\right) = \phi(\sqrt{y})\left(\frac{1}{2\sqrt{y}}\right) + \phi(\sqrt{y})\left(\frac{1}{2\sqrt{y}}\right)$$
$$= \frac{1}{\sqrt{y}}\phi(\sqrt{y}) = \frac{1}{\sqrt{2\pi y}}\exp\left(-\frac{1}{2}y\right)$$

The p.d.f. for a Chi-squared distribution is as follows:

$$\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}\exp\left(-\frac{x}{2}\right),$$

where k is the degrees of freedom. In the case of $k = 1$, the above formula translates to:

$$\frac{1}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)}x^{-\frac{1}{2}}\exp\left(-\frac{x}{2}\right).$$

The value $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, where $\Gamma(\cdot)$ denotes the gamma function. So, the above can be further rewritten as:

$$\frac{1}{\sqrt{2\pi x}}\exp\left(-\frac{x}{2}\right).$$

This is identical to the above result for the p.d.f. of Y_i:

$$\frac{1}{\sqrt{2\pi y}}\exp\left(-\frac{y}{2}\right),$$

therefore, we can conclude that $Y_i \sim \chi_1^2$, where $Y_i = Z_i^2$.

The last part to be shown is that $W = \sum_{i=1}^{p} Z_i^2 \sim \chi_p^2$. This will be done using the m.g.f. of W.

$$M_W(t) = E(e^{tW}) = E\left(e^{t\sum_{i=1}^{p} Y_i}\right) = E\left(\Pi_{i=1}^{p} e^{tY_i}\right)$$

$\because$ Y_i are independent

$$= \Pi_{i=1}^{p} E(e^{tY_i})$$

$$= \Pi_{i=1}^{p} M_{Y_i}(t)$$

(also, above it has been shown that $Y_i \sim \chi_1^2$, so $M_{Y_i}(t)$ can be rewritten as the m.g.f. of a Chi-squared distribution with 1 degree of freedom)

$$= \Pi_{i=1}^{p}\left(\frac{1}{1-2t}\right)^{\frac{1}{2}}$$

$$= \left(\frac{1}{1-2t}\right)^{\frac{p}{2}}$$

The above is the m.g.f. of $\chi_{(p)}^2$, therefore it can be said that $W = \sum_{i=1}^{p} Z_i^2 \sim \chi_p^2$. The conclusion then is that, $\vec{Z}^T\vec{Z} = Z_1^2 + \cdots + Z_p^2 \sim \chi_p^2$. ■

4. Given that the joint p.d.f. of X and Y is

$$f(x,y)=\frac{1}{\sqrt{3}\pi}\exp\left(-\frac{2}{3}[(x-3)^2+(y+2)^2+(x-3)(y+2)]\right),$$

the goal is to find $Z=h(X,Y)$, such that

$$Z\sim\chi_2^2.$$

In the case of a bivariate normal distribution:

$$\begin{pmatrix}X_1\\X_2\end{pmatrix}\sim N\left(\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix},\begin{pmatrix}\sigma_{1,1}&\sigma_{1,2}\\\sigma_{2,1}&\sigma_{2,2}\end{pmatrix}\right)$$

$$\vec{X}\sim N(\vec{\mu},\mathbf{\Sigma})$$

The p.d.f. of the bivariate normal random variables is:

$$f(X_1,X_2)=\frac{1}{\sqrt{(2\pi)^2|\mathbf{\Sigma}|}}\exp\left\{-\frac{1}{2}\left(\vec{X}-\vec{\mu}\right)^T\mathbf{\Sigma}^{-1}\left(\vec{X}-\vec{\mu}\right)\right\}$$

$$=\frac{1}{2\pi}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\vec{X}-\vec{\mu}\right)^T\mathbf{\Sigma}^{-1}\left(\vec{X}-\vec{\mu}\right)\right\}$$

$$=\frac{1}{2\pi}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}X_1-\mu_1\\X_2-\mu_2\end{pmatrix}^T\begin{pmatrix}\sigma_{1,1}&\sigma_{1,2}\\\sigma_{2,1}&\sigma_{2,2}\end{pmatrix}^{-1}\begin{pmatrix}X_1-\mu_1\\X_2-\mu_2\end{pmatrix}\right\}$$

$$=\frac{1}{2\pi}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}X_1-\mu_1&X_2-\mu_2\end{pmatrix}\begin{pmatrix}\sigma_{1,1}&\sigma_{1,2}\\\sigma_{2,1}&\sigma_{2,2}\end{pmatrix}^{-1}\begin{pmatrix}X_1-\mu_1\\X_2-\mu_2\end{pmatrix}\right\}$$

($\begin{pmatrix}\sigma_{1,1}&\sigma_{1,2}\\\sigma_{2,1}&\sigma_{2,2}\end{pmatrix}^{-1}$ will be rewritten as $\begin{pmatrix}a&b\\b&c\end{pmatrix}$ and the actual values will be calculated later)

$$=\frac{1}{2\pi}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}X_1-\mu_1&X_2-\mu_2\end{pmatrix}\begin{pmatrix}a&b\\b&c\end{pmatrix}\begin{pmatrix}X_1-\mu_1\\X_2-\mu_2\end{pmatrix}\right\}$$

$$=\frac{1}{2\pi}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}[a(X_1-\mu_1)^2+c(X_2-\mu_2)^2+2b(X_1-\mu_1)(X_2-\mu_2)]\right\}$$

Next, the given formula for the problem will be matched with the above to determine the appropriate values for a,b,c,μ_1, and μ_2. Looking at the values in the exponent, it can be seen that $a=\frac{4}{3},b=\frac{2}{3},c=\frac{4}{3},\mu_1=3$, and $\mu_2=-2$. Then going back to $\begin{pmatrix}a&b\\b&c\end{pmatrix}$, it follows that,

$$\mathbf{\Sigma}^{-1}=\begin{pmatrix}\frac{4}{3}&\frac{2}{3}\\\frac{2}{3}&\frac{4}{3}\end{pmatrix}.$$

To find $\mathbf{\Sigma}$, the inverse can be taken again to show that:

$$\mathbf{\Sigma}=\frac{1}{\left(\frac{4}{3}\right)^2-\left(\frac{2}{3}\right)^2}\begin{pmatrix}\frac{4}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{4}{3}\end{pmatrix}=\frac{3}{4}\begin{pmatrix}\frac{4}{3}&-\frac{2}{3}\\-\frac{2}{3}&\frac{4}{3}\end{pmatrix}=\begin{pmatrix}1&-\frac{1}{2}\\-\frac{1}{2}&1\end{pmatrix}.$$

Also, it will be noted first that the determinant $|\mathbf{\Sigma}|$ is as follows:

$$|\mathbf{\Sigma}| = \frac{1}{1-\frac{1}{4}}\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} \frac{4}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{4}{3} \end{pmatrix} > 0$$

It then follows from the above results that X and Y follow the bivariate normal distribution,

$$\binom{X}{Y} \sim N\left(\binom{3}{-2}, \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}\right).$$

Let $\vec{X} = \begin{bmatrix} X \\ Y \end{bmatrix}$, then $\vec{X} \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, where $\boldsymbol{\mu} = \binom{3}{-2}$ and $\mathbf{\Sigma} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$ with $|\mathbf{\Sigma}| > 0$. Then $(\vec{X} - \boldsymbol{\mu})^T \mathbf{\Sigma}(\vec{X} - \boldsymbol{\mu})$ is distributed as χ_2^2. So, in the case of this problem, the $h(X, Y)$ function is $(\vec{X} - \boldsymbol{\mu})^T \mathbf{\Sigma}(\vec{X} - \boldsymbol{\mu})$ by Result 4.7 from the textbook.■

5. Given that $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ be a random vector with the population variance $\mathbf{\Sigma}$. In the case of

$$\mathbf{\Sigma} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 5 \end{bmatrix},$$

the goal is to find (α, β), such that

$$Cov(X_1, X_3 - (\alpha X_1 + \beta X_2)) = Cov(X_2, X_3 - (\alpha X_1 + \beta X_2)) = 0.$$

$$\begin{aligned} Cov(X_1, X_3 - (\alpha X_1 + \beta X_2)) &= Cov(X_1, X_3 - \alpha X_1 - \beta X_2) \\ &= Cov(X_1, X_3) - \alpha Cov(X_1, X_1) - \beta Cov(X_1, X_2) \\ &= Cov(X_1, X_3) - \alpha Var(X_1) - \beta Cov(X_1, X_2) = 0 \end{aligned}$$

$$\alpha Var(X_1) = Cov(X_1, X_3) - \beta Cov(X_1, X_2)$$

$$\alpha = \frac{Cov(X_1, X_3) - \beta Cov(X_1, X_2)}{Var(X_1)} = \frac{1 - \beta(1)}{4} = \frac{1-\beta}{4}$$

$$\begin{aligned} Cov(X_2, X_3 - (\alpha X_1 + \beta X_2)) &= Cov(X_2, X_3 - \alpha X_1 - \beta X_2) \\ &= Cov(X_2, X_3) - \alpha Cov(X_2, X_1) - \beta Cov(X_2, X_2) \\ &= Cov(X_2, X_3) - \alpha Cov(X_2, X_1) - \beta Var(X_2) \end{aligned}$$

$$\alpha Cov(X_2, X_1) = Cov(X_2, X_3) - \beta Var(X_2)$$

$$\alpha = \frac{Cov(X_2, X_3) - \beta Var(X_2)}{Cov(X_2, X_1)} = \frac{1 - \beta(3)}{1} = 1 - 3\beta$$

Then setting the two equations equal:

$$\frac{1-\beta}{4} = 1 - 3\beta$$

$$1 - \beta = 4 - 12\beta$$

$$11\beta = 3$$

$$\boxed{\beta = \frac{3}{11}}$$

Then plugging the result back into the original equation:

$$\alpha = 1 - 3\left(\frac{3}{11}\right) = 1 - \frac{9}{11} = \frac{2}{11}$$

$$\boxed{\alpha = \frac{2}{11}}$$

6. Let $\vec{X} \sim \mathcal{N}_p(\vec{\mu}, \mathbf{\Sigma})$. Let

$$\mathbf{\Sigma} = \sum_{j=1}^{p} \lambda_j \vec{v}_j \vec{v}_j^T$$

be the spectral decomposition. Let $Y_j = \vec{v}_j^T \vec{X}$ for all $j = 1, \cdots, p$. It will be shown that $Y_1, \cdots, Y_p$ are mutually independent.

It is given that $\vec{X}$ follows a normal distribution and $\vec{v}_j^T$ is a scalar and so the product of the two are a linear combination of a normal random variable. Therefore, Y_j is itself a normal random variable. To show that $Y_1, \cdots, Y_p$ are mutually independent, it must first be shown that that they are uncorrelated. Similar to previously problems, uncorrelated normal random variables are mutually independent.

$$Cov(Y_i, Y_j) = 0 \text{ for } i \neq j$$

$$Cov(\vec{v}_i^T \vec{X}, \vec{v}_j^T \vec{X}) = \vec{v}_i^T Cov(\vec{X}, \vec{X}) \vec{v}_j = \vec{v}_i^T \mathbf{\Sigma} \vec{v}_j = \vec{v}_i^T \lambda_j \vec{v}_j = \lambda_j \vec{v}_i^T \vec{v}_j = 0$$

The last part follows since $\vec{v}_i$ and $\vec{v}_j$ are orthogonal for $i \neq j$. We can then conclude that $Y_1, \cdots, Y_p$ are mutually independent. ■