Prof. Dr. Rudolf Mathar, Dr. Gholamreza Alirezaei, Dr. Arash Behboodi

Exercise 2
- Proposed Solution -

Friday, November 11, 2016

Solution of Problem 1

Note that for any random variable $\boldsymbol{Y} = g(\boldsymbol{X})$ the expectation $\mathrm{E}(\boldsymbol{Y}) = \mathrm{E}(g(\boldsymbol{X}))$ is defined by

$$\mathrm{E}(\boldsymbol{Y}) = \begin{cases} \sum_i g(\boldsymbol{x}_i)\, p_{\boldsymbol{X}}(\boldsymbol{x}_i)\,, & \text{if } \boldsymbol{X} \text{ is discrete,} \\ \int_{\mathrm{supp}\{\boldsymbol{X}\}} g(\boldsymbol{x})\, f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}\,, & \text{if } \boldsymbol{X} \text{ is continuous}\,. \end{cases} \tag{1}$$

Because of the linearity of both operators (sum and integral), it follows that:

a)

$$\begin{aligned} \mathrm{E}(\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}) &= \sum_i (\boldsymbol{A}\boldsymbol{x}_i + \boldsymbol{b})\, p_{\boldsymbol{X}}(\boldsymbol{x}_i) \\ &\underset{\text{linearity}}{=} \boldsymbol{A} \sum_i \boldsymbol{x}_i\, p_{\boldsymbol{X}}(\boldsymbol{x}_i) + \boldsymbol{b} \sum_i p_{\boldsymbol{X}}(\boldsymbol{x}_i) \\ &\underset{\text{definition}}{=} \boldsymbol{A}\,\mathrm{E}(\boldsymbol{X}) + \boldsymbol{b}\,, \end{aligned}$$

b)

$$\begin{aligned} \mathrm{E}(c_X \boldsymbol{X} + c_Y \boldsymbol{Y}) &= \sum_{i,j} (c_X \boldsymbol{x}_i + c_Y \boldsymbol{y}_j)\, p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x}_i, \boldsymbol{y}_j) \\ &\underset{\text{linearity}}{=} c_X \sum_{i,j} \boldsymbol{x}_i\, p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x}_i, \boldsymbol{y}_j) + c_Y \sum_{i,j} \boldsymbol{y}_j\, p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x}_i, \boldsymbol{y}_j) \\ &\underset{\text{independency}}{=} c_X \sum_{i,j} \boldsymbol{x}_i\, p_{\boldsymbol{X}}(\boldsymbol{x}_i)\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) + c_Y \sum_{i,j} \boldsymbol{y}_j\, p_{\boldsymbol{X}}(\boldsymbol{x}_i)\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) \\ &\underset{\text{separation}}{=} c_X \sum_i \boldsymbol{x}_i\, p_{\boldsymbol{X}}(\boldsymbol{x}_i) \sum_j p_{\boldsymbol{Y}}(\boldsymbol{y}_j) + c_Y \sum_i p_{\boldsymbol{X}}(\boldsymbol{x}_i) \sum_j \boldsymbol{y}_j\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) \\ &\underset{\text{unitary}}{=} c_X \sum_i \boldsymbol{x}_i\, p_{\boldsymbol{X}}(\boldsymbol{x}_i) + c_Y \sum_j \boldsymbol{y}_j\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) \\ &\underset{\text{definition}}{=} c_X\,\mathrm{E}(\boldsymbol{X}) + c_Y\,\mathrm{E}(\boldsymbol{Y})\,, \end{aligned}$$

c)

$$\begin{aligned} \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) &= \sum_{i,j} \boldsymbol{x}_i\, \boldsymbol{y}_j\, p_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{x}_i, \boldsymbol{y}_j) \\ &\underset{\text{independency}}{=} \sum_i \sum_j \boldsymbol{x}_i\, \boldsymbol{y}_j\, p_{\boldsymbol{X}}(\boldsymbol{x}_i)\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) \\ &\underset{\text{separation}}{=} \sum_i \boldsymbol{x}_i\, p_{\boldsymbol{X}}(\boldsymbol{x}_i) \sum_j \boldsymbol{y}_j\, p_{\boldsymbol{Y}}(\boldsymbol{y}_j) \\ &\underset{\text{definition}}{=} \mathrm{E}(\boldsymbol{X})\,\mathrm{E}(\boldsymbol{Y})\,. \end{aligned}$$

Note that the covariance $\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})$ between two random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined by $\mathrm{E}\left([\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{Y} - \mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}\right)$ while the covariance matrix of the random variable $\boldsymbol{Z}$ is given by $\mathrm{Cov}(\boldsymbol{Z}, \boldsymbol{Z})$ or in simple notation $\mathrm{Cov}(\boldsymbol{Z})$. Hence, this leads to

d)

$$\begin{aligned}
\mathrm{Cov}(\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}) &= \mathrm{Cov}(\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}) \\
&\underset{\text{definition}}{=} \mathrm{E}\left([\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} - \mathrm{E}(\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b})][\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} - \mathrm{E}(\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b})]^{\mathrm{H}}\right) \\
&\underset{\text{a)}}{=} \mathrm{E}\left([\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} - \boldsymbol{A}\,\mathrm{E}(\boldsymbol{X}) - \boldsymbol{b}][\boldsymbol{A}\boldsymbol{X} + \boldsymbol{b} - \boldsymbol{A}\,\mathrm{E}(\boldsymbol{X}) - \boldsymbol{b}]^{\mathrm{H}}\right) \\
&\underset{\text{apply brackets}}{=} \mathrm{E}\left(\boldsymbol{A}[\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\right) \\
&\underset{\text{a)}}{=} \boldsymbol{A}\,\mathrm{E}\left([\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^{\mathrm{H}}\right)\boldsymbol{A}^{\mathrm{H}} \\
&\underset{\text{definition}}{=} \boldsymbol{A}\,\mathrm{Cov}(\boldsymbol{X})\,\boldsymbol{A}^{\mathrm{H}},
\end{aligned}$$

e) Similarly to the proof in d)

$$\begin{aligned}
\mathrm{Cov}(c_X\boldsymbol{X}+c_Y\boldsymbol{Y}) &= \mathrm{E}\left([c_X\boldsymbol{X}+c_Y\boldsymbol{Y}-\mathrm{E}(c_X\boldsymbol{X}+c_Y\boldsymbol{Y})][c_X\boldsymbol{X}+c_Y\boldsymbol{Y}-\mathrm{E}(c_X\boldsymbol{X}+c_Y\boldsymbol{Y})]^{\mathrm{H}}\right) \\
&= \mathrm{E}\left([c_X\boldsymbol{X}+c_Y\boldsymbol{Y}-c_X\,\mathrm{E}(\boldsymbol{X})-c_Y\,\mathrm{E}(\boldsymbol{Y})][c_X\boldsymbol{X}+c_Y\boldsymbol{Y}-c_X\,\mathrm{E}(\boldsymbol{X})-c_Y\,\mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}\right) \\
&= \mathrm{E}\left([c_X(\boldsymbol{X}-\mathrm{E}(\boldsymbol{X}))+c_Y(\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y}))][c_X(\boldsymbol{X}-\mathrm{E}(\boldsymbol{X}))+c_Y(\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y}))]^{\mathrm{H}}\right) \\
&= \mathrm{E}\left(|c_X|^2[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})][\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathrm{H}}+|c_Y|^2[\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})][\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}\right. \\
&\left.+c_X c_Y^{\mathrm{H}}[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})][\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}+c_Y c_X^{\mathrm{H}}[\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})][\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathrm{H}}\right) \\
&= |c_X|^2\,\mathrm{Cov}(\boldsymbol{X}) + |c_Y|^2\,\mathrm{Cov}(\boldsymbol{Y}) \\
&+\mathrm{E}\left(c_X c_Y^{\mathrm{H}}[\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})][\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}+c_Y c_X^{\mathrm{H}}[\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})][\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathrm{H}}\right) \\
&= |c_X|^2\,\mathrm{Cov}(\boldsymbol{X}) + |c_Y|^2\,\mathrm{Cov}(\boldsymbol{Y}) \\
&+c_X c_Y^{\mathrm{H}}\,\mathrm{E}\left(\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})\right)\mathrm{E}\left(\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})\right)^{\mathrm{H}}+c_Y c_X^{\mathrm{H}}\,\mathrm{E}\left(\boldsymbol{Y}-\mathrm{E}(\boldsymbol{Y})\right)\mathrm{E}\left(\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})\right)^{\mathrm{H}} \\
&= |c_X|^2\,\mathrm{Cov}(\boldsymbol{X}) + |c_Y|^2\,\mathrm{Cov}(\boldsymbol{Y}) \\
&+c_X c_Y^{\mathrm{H}}\,[\mathrm{E}(\boldsymbol{X})-\mathrm{E}(\boldsymbol{X})]\,[\mathrm{E}(\boldsymbol{Y})-\mathrm{E}(\boldsymbol{Y})]^{\mathrm{H}}+c_Y c_X^{\mathrm{H}}\,[\mathrm{E}(\boldsymbol{Y})-\mathrm{E}(\boldsymbol{Y})]\,[\mathrm{E}(\boldsymbol{X}-\mathrm{E}(\boldsymbol{X})]^{\mathrm{H}} \\
&= |c_X|^2\,\mathrm{Cov}(\boldsymbol{X}) + |c_Y|^2\,\mathrm{Cov}(\boldsymbol{Y})\,.
\end{aligned}$$

Solution of Problem 2

The density of a normal distribution is given by

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\exp\Big(-\tfrac{1}{2}\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big)$$

The mean can be calculated by

$$
\begin{aligned}
\mathrm{E}(\boldsymbol{X}) &= \int\limits_{\mathrm{supp}\{\boldsymbol{X}\}} \boldsymbol{x}\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\underset{\mathrm{supp}\{\boldsymbol{X}\}=\mathbb{R}^n}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{x}\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n \\
&\underset{\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} (\boldsymbol{y}+\boldsymbol{\mu})\, f_{\boldsymbol{X}}(\boldsymbol{y}+\boldsymbol{\mu})\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n \\
&\underset{\text{linearity}}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{y}\, f_{\boldsymbol{X}}(\boldsymbol{y}+\boldsymbol{\mu})\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n \\
&\qquad + \boldsymbol{\mu}\underbrace{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} f_{\boldsymbol{X}}(\boldsymbol{y}+\boldsymbol{\mu})\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n}_{=1} \\
&\underset{\text{unitary}}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{y}\, f_{\boldsymbol{X}}(\boldsymbol{y}+\boldsymbol{\mu})\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n + \boldsymbol{\mu} \\
&= \boldsymbol{\mu} + \underbrace{\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{y}\,\exp\!\Big(-\tfrac{1}{2}\,\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{y}\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n}_{=\mathrm{E}(\boldsymbol{Y})\ \text{with}\ \boldsymbol{Y}\sim\mathcal{N}_n(\boldsymbol{0},\boldsymbol{\Sigma})} \\
&= \boldsymbol{\mu}\,.
\end{aligned}
$$

The covariance can similarly be calculated by

$$
\begin{aligned}
\mathrm{Cov}(\boldsymbol{X}) &\underset{\text{definition}}{=} \int\limits_{\mathrm{supp}\{\boldsymbol{X}\}} [\boldsymbol{x}-\boldsymbol{\mu}][\boldsymbol{x}-\boldsymbol{\mu}]^{\mathrm{T}}\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\underset{\mathrm{supp}\{\boldsymbol{X}\}=\mathbb{R}^n}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} [\boldsymbol{x}-\boldsymbol{\mu}][\boldsymbol{x}-\boldsymbol{\mu}]^{\mathrm{T}}\, f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n \\
&\underset{\boldsymbol{y}=\boldsymbol{x}-\boldsymbol{\mu}}{=} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{y}\,\boldsymbol{y}^{\mathrm{T}}\, f_{\boldsymbol{X}}(\boldsymbol{y}+\boldsymbol{\mu})\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n \\
&= \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{y}\,\boldsymbol{y}^{\mathrm{T}}\,\exp\!\Big(-\tfrac{1}{2}\,\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{y}\Big)\,\mathrm{d}y_1\,\mathrm{d}y_2\cdots\mathrm{d}y_n
\end{aligned}
$$

Now consider the Choleski decomposition of $\boldsymbol{\Sigma}$ by $\boldsymbol{\Sigma} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}$. Notice that $\boldsymbol{M}$ is invertible and it holds that $|\boldsymbol{\Sigma}| = |\boldsymbol{M}^{\mathrm{T}}|\,|\boldsymbol{M}| = |\boldsymbol{M}|^2$. So the Jacobian of the transformation $\boldsymbol{z} = \boldsymbol{M}\boldsymbol{y}$ is equal to $|\boldsymbol{M}|$ and is positive. Note the identity $\boldsymbol{M}^{\mathrm{T}} = \boldsymbol{M}^{-1}$. Hence, the above equation becomes

$$
\begin{aligned}
\mathrm{Cov}(\boldsymbol{X}) &\underset{\boldsymbol{y}=\boldsymbol{M}^{-1}\boldsymbol{z}}{=} \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{M}^{-1}\,\boldsymbol{z}\,\boldsymbol{z}^{\mathrm{T}}\,\boldsymbol{M}\,\exp\!\Big(-\tfrac{1}{2}\,\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}\Big)\,|\boldsymbol{M}|\,\mathrm{d}z_1\,\mathrm{d}z_2\cdots\mathrm{d}z_n \\
&= \boldsymbol{M}^{-1}\underbrace{\frac{1}{\sqrt{(2\pi)^n}}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\cdots\int\limits_{-\infty}^{\infty} \boldsymbol{z}\,\boldsymbol{z}^{\mathrm{T}}\,\exp\!\Big(-\tfrac{1}{2}\,\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}\Big)\,|\boldsymbol{M}|\,\mathrm{d}z_1\,\mathrm{d}z_2\cdots\mathrm{d}z_n\,\boldsymbol{M}}_{=\mathrm{Cov}(\boldsymbol{Z})\ \text{with}\ \boldsymbol{Z}\sim\mathcal{N}_n(\boldsymbol{0},\boldsymbol{I_n})} \\
&= \boldsymbol{M}^{-1}\,\boldsymbol{I_n}\,\boldsymbol{M} = \boldsymbol{\Sigma}\,.
\end{aligned}
$$

Solution of Problem 3

Since the number of variables in $\mathrm{E}(X_1X_2X_3X_4X_5)$ is odd, its solution is zero, i.e., $\mathrm{E}(X_1X_2X_3X_4X_5) = 0$.

For $\mathrm{E}(X_1X_2X_3X_4)$ we need to sum up over all permutations $\boldsymbol{\pi}$ of two random variables. Hence, it leads to

$$\begin{aligned}\mathrm{E}(X_1X_2X_3X_4) &= \mathrm{E}(X_1X_2)\,\mathrm{E}(X_3X_4) + \mathrm{E}(X_1X_3)\,\mathrm{E}(X_2X_4) + \mathrm{E}(X_1X_4)\,\mathrm{E}(X_2X_3)\\ &= \Sigma_{1,2}\Sigma_{3,4} + \Sigma_{1,3}\Sigma_{2,4} + \Sigma_{1,4}\Sigma_{2,3}\end{aligned}$$

For $\mathrm{E}(X_1^2X_5^4)$ we rewrite it as $\mathrm{E}(X_1X_1X_5X_5X_5X_5)$ and use the Isserlis' Theorem to obtain

$$\begin{aligned}\mathrm{E}(X_1X_1X_5X_5X_5X_5) &= 3\,\mathrm{E}(X_1X_1)\,\mathrm{E}(X_5X_5)\,\mathrm{E}(X_5X_5) + 12\,\mathrm{E}(X_1X_5)\,\mathrm{E}(X_1X_5)\,\mathrm{E}(X_5X_5)\\ &= 3\Sigma_{1,1}\Sigma_{5,5}^2 + 12\Sigma_{1,5}^2\Sigma_{5,5}\end{aligned}$$

By Isserlis' Theorem, all higher moments of a sequence $X_1, X_2, \ldots, X_n$ can be calculated by the aid of the covariance entries. This shows that all higher moments of the normal distribution are dependent to the second order moments. Thus, higher order moments cannot contain more information compared to the information included in the second order moments.