Class Exercise 4 – Second Order Statistics of Random Vectors, Gaussian Random Vector

Reminder

We say that $X \sim N(\mu,\sigma^2)$ when it has a PDF of $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Equivalently, its characteristic function is $\phi_X(\omega) = E\left[e^{i\omega X}\right] = e^{i\omega\mu - \frac{1}{2}\sigma^2\omega^2}$.

For $\sigma^2 = 0$, X is deterministic and $f_X(x) = \delta(x-\mu)$.

The set of the random variables $X_1, X_2, \cdots X_n$ is jointly Gaussian iff (if and only if) every linear combination of them is a Gaussian variable. Equivalently, $X_1, X_2, \cdots X_n$ are jointly Gaussian iff (if and only if) they can be represented as an affine transformation of n independent Gaussian variables:

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = A \begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix} + b, \qquad Z_i \sim N(0,1), \qquad \{Z_i\} \text{ are i.i.d}$$

Problem 1:

Given are three random variables X_1, X_2, X_3, who satisfy:

$$Var(X_1) = Var(X_2) = Var(X_3) = 1$$
$$Var(X_1 + X_2 + X_3) = 0$$

Find the covariance matrix of $\underline{X} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}$.

Solution:

The variance of a random variable is zero iff it is deterministic, i.e. equal to a **constant** with probability 1.

This means that exists a constant a such that, with probability 1:

$$X_1 + X_2 + X_3 = a$$

or put differently:

$$X_1 = a - (X_2 + X_3)$$

By finding the variance on both sides we get:

$$Var(X_1) = Var(X_2 + X_3) = Var(X_2) + Var(X_3) + 2Cov(X_2, X_3)$$

It is given that all the variances are equal to 1, thus:

$$1 = 2 + 2Cov(X_2, X_3)$$
$$\Rightarrow Cov(X_2, X_3) = -0.5$$

And, similarly, both of the other covariances are equal to -0.5.

All in all, we got:

$$C_X = \begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{pmatrix}$$

Let us check the result we got: we'll define a new random variable:

$$Y = X_1 + X_2 + X_3 = \underbrace{\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

As required, the variance of Y is:

$$Var(Y) = Var(AX) = AC_X A^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0$$

Problem 2:

Consider $X \sim N(0, \sigma_X^2)$ and $\Theta \sim U[0, 2\pi]$, independent of each other.
Let us define the random variable $W = X\cos(\theta)$. Is W Gaussian? Prove or refute.

Solution:

W is not Gaussian.
Recall that for a Gaussian random variable $X \sim N(0, \sigma_X^2)$:

$$\mathrm{E}\left[X^n\right] = \begin{cases} 0, & n \quad odd \\ \sigma_X^n \cdot 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (n-1), & n \quad even \end{cases}$$

Let us check whether W satisfies the same characteristic as X.
For this cause, we will first check whether the expectation of W is zero:

$$\mathrm{E}\left[W\right] = \mathrm{E}\left[X\cos(\Theta)\right] = \mathrm{E}\left[X\right] \cdot \mathrm{E}\left[\cos(\Theta)\right] = 0$$

If so, let us check whether W satisfies the characteristic of a Gaussian random variable with expected value zero and the same variance as W, namely, with distribution $\sim N(0, \sigma_W^2)$:

$$\mathrm{E}\left[W^4\right] \overset{?}{=} 3 \cdot \sigma_W^4$$

We will calculate each side of the equation separately and then compare:

$$\mathrm{E}\left[W^4\right] = \mathrm{E}\left[X^4\right] \cdot \mathrm{E}\left[\cos^4(\Theta)\right] = 3\sigma^4 \cdot \frac{3}{8}$$

On the other hand:

$$\sigma_W^2 = \mathrm{E}\left[W^2\right] = \mathrm{E}\left[X^2\right] \cdot \mathrm{E}\left[\cos^2(\Theta)\right] = \sigma^2 \cdot \frac{1}{2}$$

Therefore:

$$3 \cdot \sigma_W^4 = 3\left(\mathrm{E}\left[W^2\right]\right)^2 = 3\sigma^4 \cdot \frac{1}{4}$$

We got that:

$$\mathrm{E}\left[W^4\right] = 3\sigma^4 \cdot \frac{3}{8} \quad \neq \quad 3\sigma^4 \cdot \frac{1}{4} = 3\sigma_W^4$$

$\Rightarrow W$ is not Gaussian.

<u>Problem 3:</u>

Consider a Gaussian random variable $\underline{X} = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}^T$ with expectation vector $\eta_{\underline{X}} = \underline{0}$, and covariance matrix:

$$C_{\underline{X}\,\underline{X}} = \begin{bmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}$$

1. Given is the discrete random variable K that takes the values:

$$P\{K = 1\} = P\{K = 2\} = P\{K = 3\} = P\{K = 4\} = \frac{1}{4}$$

Calculate the expectation $E\left\{\prod_{i=1}^{K} X_i\right\}$. Assume that K is independent of $\underline{X}$.

2. The random variables $\begin{cases} Z = X_1 + X_2 \\ W = X_3 - X_4 \end{cases}$ are defined. Are W, Z independent?

<u>Solution:</u>

1. We will use the smoothing theorem:

$$E\left\{\prod_{i=1}^{K} X_i\right\} = E\left\{E\left\{\prod_{i=1}^{K} X_i \mid K\right\}\right\}$$

We will start by calculating the conditional expectation for any value of K:

$$E\left\{\prod_{i=1}^{K} X_i \mid K=j\right\} = E\left\{\prod_{i=1}^{j} X_i \mid K=j\right\} = E\left\{\prod_{i=1}^{j} X_i\right\} = \begin{cases} E\{X_1\} & j=1 \\ E\{X_1X_2\} & j=2 \\ E\{X_1X_2X_3\} & j=3 \\ E\{X_1X_2X_3X_4\} & j=4 \end{cases}$$

For joint Gaussian random variables with expected value zero, it holds that:

$$E\{X_1\} = E\{X_1X_2X_3\} = 0$$

and, thus:

$$E\{X_1\} = 0$$
$$E\{X_1X_2\} = 0.5$$
$$E\{X_1X_2X_3\} = 0$$
$$E\{X_1X_2X_3X_4\} = E\{X_1X_2\}E\{X_3X_4\} + E\{X_1X_3\}E\{X_2X_4\} + E\{X_1X_4\}E\{X_2X_3\} = 0.75$$

Therefore, we get:

$$E\left\{\prod_{i=1}^{K} X_i\right\} = E\left\{E\left\{\prod_{i=1}^{K} X_i \mid K\right\}\right\} = \sum_{j=1}^{4} P(K=j)\cdot E\left\{\prod_{i=1}^{K} X_i \mid K=j\right\} = \frac{1}{4}\left[0+\frac{1}{2}+0+\frac{3}{4}\right] = \frac{5}{16}$$

2. Since W, Z were created by a linear transformation on a Gaussian random vector, they are joint Gaussian random variables. Consequently, **it is enough to check correlation between them** and it is not necessary to find the distribution explicitly. In this case, it can be shown that:

$$E\{Z\} = E\{X_1\} + E\{X_2\} = 0$$
$$E\{W\} = E\{X_3\} - E\{X_4\} = 0$$

Thus, it is enough to check orthogonality, namely check whether $E\{ZW\} = 0$:

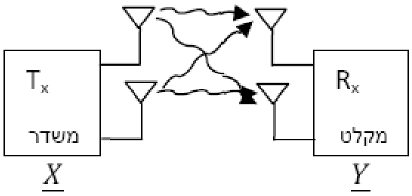
$$E\{ZW\} = E\{(X_1+X_2)(X_3-X_4)\} = E\{X_1X_3 + X_2X_3 - X_1X_4 - X_2X_4\} =$$
$$= E\{X_1X_3\} + E\{X_2X_3\} - E\{X_1X_4\} - E\{X_2X_4\} = 0.5+0.5-0.5-0.5 = 0$$

W, Z are in fact orthogonal. Since the expected value is equal to 0, the variables are uncorrelated. To conclude, recall that uncorrelated and joint Gaussian random variables are independent.

Extra Questions

Problem 4 (Estimation teaser):

Given is a digital communication system with a transmitter with two antennas and a receiver with two antennas (MIMO).



Through the first transmitting antenna, the random variable X_1 is transmitted and through the second transmitting antenna, the random variable X_2 is transmitted, when it is given that:

$$X_{1,2} \sim \begin{cases} 1 & w.p.\ 0.5 \\ -1 & w.p.\ 0.5 \end{cases}$$

The received signal in both of the receiving antennas is given by:

$$Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1$$
$$Y_2 = h_{21}X_1 + h_{22}X_2 + Z_2$$

Or, in short:

$$\underline{Y} = H\,\underline{X} + \underline{Z}$$

(the matrix H represents the communication channel from transmitter to receiver)
Assume that $Z_{1,2} \sim N(0,0.1)$, H is invertible and X_1, X_2, Z_1, Z_2 are independent.
The purpose of the receiver is to discover what are X_1 and X_2 given the received information Y_1, Y_2. In order to guess the transmitted vector, it is proposed to multiply first the received vector $\underline{Y}$ by the inverse matrix of H. This results in:

$$\underline{Y'} = H^{-1}\underline{Y} = H^{-1}H\,\underline{X} + H^{-1}\underline{Z} = \underline{X} + H^{-1}\underline{Z} = \underline{X} + \underline{Z'}$$

We can write this as:

$$Y_1' = X_1 + Z_1'$$
$$Y_2' = X_2 + Z_2'$$

1. Find the distribution of $\underline{Z'} = H^{-1}\underline{Z}$, the "new" noise vector, when it is given that:

$$H = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$$

Solution:

Notice that $\underline{Z}$ is a Gaussian random vector and, thus, so is $\underline{Z'}$, since it is a linear transformation of $\underline{Z}$.

Let us find the expectation vector of $\underline{Z}'$ and its covariance matrix:

$$\mathrm{E}\left[\underline{Z}'\right] = H^{-1}\mathrm{E}\left[\underline{Z}\right] = H^{-1}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

$$\operatorname{cov}\left(\underline{Z}'\right) = H^{-1}\operatorname{cov}\left(\underline{Z}\right)\left(H^{-1}\right)^{T} = \begin{bmatrix}4/3 & 2/3\\2/3 & 4/3\end{bmatrix}\begin{bmatrix}0.1 & 0\\0 & 0.1\end{bmatrix}\begin{bmatrix}4/3 & 2/3\\2/3 & 4/3\end{bmatrix} = \begin{bmatrix}2/9 & 8/45\\8/45 & 2/9\end{bmatrix}$$

Recall that the distribution of a Gaussian random vector is defined exclusively by its expectation vector and its covariance matrix therefore it is sufficient to find them.

2. Now it is suggested to decide on X_1 and X_2 on the following way:

$$\hat{X}_1 = sign\left(Y_1'\right)$$
$$\hat{X}_2 = sign\left(Y_2'\right)$$

Find the probability of error for the decision of X_1 and for the decision of X_2.

Solution:

The probability of error of the estimator $\hat{X}$ of the random variable X is defined as the probability that the event $\hat{X} \neq X$ occurs, i.e., $P_{error} = \Pr\{\hat{X} \neq X\}$.

Let us find the probability of error when estimating X_1:

$$P_{error1} = \Pr\{\hat{X}_1 \neq X_1\} = \Pr\{X_1 = 1\}\cdot\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = 1\} + \Pr\{X_1 = -1\}\cdot\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = -1\} =$$
$$= \frac{1}{2}\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = 1\} + \frac{1}{2}\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = -1\}$$

$$\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = 1\} = \Pr\{Sign(Y_1') \neq 1 \mid X_1 = 1\} = \Pr\{Sign(Y_1') = -1 \mid X_1 = 1\} =$$
$$= \Pr\{Sign(X_1 + Z_1') = -1 \mid X_1 = 1\} = \Pr\{Sign(Z_1' + 1) = -1 \mid X_1 = 1\} = \Pr\{Sign(Z_1' + 1) = -1\}$$
$$= \Pr\{Z_1' + 1 \leq 0\} = \Pr\{Z_1' \leq -1\} = \Pr\{Z_1' \geq 1\} = Q\left(\frac{1}{\sqrt{2/9}}\right) = Q\left(\sqrt{4.5}\right) \approx 0.0169$$

Since Z_1' is a Gaussian random vector with expected value 0, it follows that it distributes symmetrically around the origin, and, thus:

$$\Pr\{\hat{X}_1 \neq X_1 \mid X_1 = -1\} = \Pr\{sign(Z_1' - 1) = 1\} = \Pr\{sign(Z_1' + 1) = -1\} =$$
$$= \Pr\{\hat{X}_1 \neq X_1 \mid X_1 = 1\}$$

All in all, we got:

$$P_{error1} = \frac{1}{2}\cdot Q\left(\sqrt{4.5}\right) + \frac{1}{2}\cdot Q\left(\sqrt{4.5}\right) = Q\left(\sqrt{4.5}\right) \approx 0.0169$$

For reasons of symmetry, the probability of error in the estimation of X_2 is exactly the same as the probability of error in the estimation of X_1.

Note: in section 1 we got that the noise elements are **not** uncorrelated. In other words, after multiplying by the inverse matrix, we indeed got that $Y_1^{'}$ is independent of X_2 and that $Y_2^{'}$ is independent of X_1, but the noises between the two measurements are dependent of one another. This hints that we did not use all the information that was received in the receiver, for $Y_1^{'}$ still includes some information on $Y_2^{'}$ and vice versa (due to the correlation between them).

Later on we will see how one can design a better receiver, which reaches smaller probabilities of error in the estimation of X_1 and X_2 - namely, the estimator proposed in this question is not optimal.

Appendix 1: Random Vectors – Second Order Statistics

- The expectation of a random vector is a deterministic vector (expectation is a number). They are marked as follows:

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad E\{\underline{X}\} = \underline{\eta}_X = \begin{bmatrix} E\{X_1\} \\ E\{X_2\} \\ \vdots \\ E\{X_n\} \end{bmatrix}$$

- Correlation matrix (deterministic matrix):

$$R_{XX} = E\{\underline{X}\,\underline{X}^T\} = \begin{bmatrix} E\{X_1X_1\} & E\{X_1X_2\} & \cdots & E\{X_1X_n\} \\ E\{X_2X_1\} & E\{X_2X_2\} & & E\{X_2X_n\} \\ \vdots & & \ddots & \vdots \\ E\{X_nX_1\} & E\{X_nX_2\} & \cdots & E\{X_nX_n\} \end{bmatrix}$$

- Covariance matrix (deterministic matrix):

$$C_{XX} = E\{(\underline{X} - \underline{\eta}_X)(\underline{X} - \underline{\eta}_X)^T\} = E\{\underline{X}\,\underline{X}^T\} - \underline{\eta}_X \underline{\eta}_X^T$$

- Cross-correlation matrix:

$$R_{XY} = E\{\underline{X}\,\underline{Y}^T\} = R_{YX}^T$$

- Cross-covariance matrix:

$$C_{XY} = E\{(\underline{X} - \underline{\eta}_X)(\underline{Y} - \underline{\eta}_Y)^T\} = E\{\underline{X}\,\underline{Y}^T\} - \underline{\eta}_X \underline{\eta}_Y^T = C_{YX}^T$$

Notice that the auto-covariance and auto-correlation matrices are necessarily positive semi-definite and symmetrical thus diagonalizable, whereas the cross-correlation and cross-covariance matrices are not necessarily positive semi-definite or symmetrical.

Linear transformation of a random vector:

Consider the random vector $\underline{X}$ with covariance matrix C_{XX} and expectation vector $\underline{\eta}_X$. Let us define $\underline{Y} = \mathbf{A}\underline{X} + \underline{b}$, where A is a deterministic matrix and $\underline{b}$ is a deterministic vector. Then:

$$\begin{aligned} C_{YY} &= \mathbf{A} C_{XX} \mathbf{A}^T \\ \underline{\eta}_Y &= \mathbf{A}\underline{\eta}_X + \underline{b} \end{aligned}$$

Appendix 2 – Q-Function

Seeing as how an analytical expression of the CDF of a Gaussian random variable does not exist, the Q-function is often used to deal with these cases.

Let us denote $Z \sim N(0,1)$. It holds that:

$$Q(x) = \Pr\{Z \ge x\} = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{z^2}{2}} dz$$

For any Gaussian random variable $Y \sim N(\mu, \sigma^2)$, such that $\sigma^2 > 0$:

$$\Pr\{Y \ge y\} = \Pr\left\{\frac{Y-\mu}{\sigma} \ge \frac{y-\mu}{\sigma}\right\} = \Pr\left\{Z \ge \frac{y-\mu}{\sigma}\right\} = Q\left(\frac{y-\mu}{\sigma}\right)$$

Appendix 3 –Gaussian Random Vectors

A random vector is Gaussian iff every linear combination of its components is a Gaussian variable.

The characteristic function of a Gaussian vector is

$$\phi_{\underline{X}}(\underline{\omega}) = \exp\left\{ j\underline{\eta}_X^T \underline{\omega} - \frac{1}{2}\underline{\omega}^T C_{XX} \underline{\omega} \right\}$$

where η_X is the expectation vector of $\underline{X}$ and C_{XX} the covariance matrix of $\underline{X}$.

If the covariance matrix is invertible, then $\underline{X}$ is a continuous Gaussian vector which has the PDF of:

$$f_{\underline{X}}(\underline{x}) = \frac{1}{\sqrt{(2\pi)^n (\det C_{XX})}} \exp\left\{ -\frac{1}{2}\left(\underline{x} - \underline{\eta}_X\right)^T C_{XX}^{-1} \left(\underline{x} - \underline{\eta}_X\right) \right\}$$

The vector can be represented as an affine transformation of a Gaussian vector with any desired (but self-consistent) expectations, variances and correlations.

Some reminders:

- A random vector whose components are independent Gaussian variables is Gaussian. The opposite direction is not necessarily true.
- A linear transformation of a random Gaussian vector is a Gaussian vector.
- If $\underline{Y} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$ is Gaussian, then $\underline{X}_2 \mid \underline{X}_1$ is Gaussian.
- If $\underline{Y} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$ is Gaussian and the cross-covariance between $\underline{X}_1$ and $\underline{X}_2$ is zero, then $\underline{X}_1$ and $\underline{X}_2$ are independent Gaussian vectors.
- *Fourth moment of Gaussian random vector*: assume that the random vector $X = (X_1\ X_2\ X_3\ X_4)$ has expectation $\underline{\eta}_x = 0$. Thus

$$E\{X_1X_2X_3X_4\} = E\{X_1X_2\}E\{X_3X_4\} + E\{X_1X_3\}E\{X_2X_4\} + E\{X_1X_4\}E\{X_2X_3\}$$

Proof:

$$\Phi_{\underline{X}}(\underline{\omega}) = E\left[e^{j(\omega_1 X_1 + \omega_2 X_2 + \omega_3 X_3 + \omega_4 X_4)}\right] \overset{Taylor}{=} E[1] + \ldots + \frac{1}{4!} E\left[(\underline{\omega}^T \underline{X})^4\right] + \ldots$$

The fourth element includes, among others: $\dfrac{24E\{X_1X_2X_3X_4\}\omega_1\omega_2\omega_3\omega_4}{4!}$

$$\Phi_{\underline{X}}(\underline{\omega}) = e^{-\frac{1}{2}\underline{\omega}^T C \underline{\omega}} = e^{-\frac{1}{2}\sum_i \sum_j \omega_i \omega_j c_{ij}} \underset{Taylor}{=} 1 + \ldots + \frac{1}{2}\left(\frac{1}{2}\sum_{i,j} \omega_i \omega_j c_{ij}\right)^2 + \ldots =$$

$$= 1 + \ldots + \frac{8}{8}\left(c_{12}c_{34} + c_{13}c_{24} + c_{14}c_{23}\right)\omega_1\omega_2\omega_3\omega_4 + \ldots$$

By comparing the factor of $\omega_1\omega_2\omega_3\omega_4$ we get:

$$E[X_1X_2X_3X_4] = c_{12}c_{34} + c_{13}c_{24} + c_{14}c_{23} \underset{\underline{\eta}=0}{=} E[X_1X_2]E[X_3X_4] + E[X_1X_3]E[X_2X_4] + E[X_1X_4]E[X_2X_3]$$