

October 7

October 2018

## Tutorial 11: Random Processes

*Lecturer: Utpal Mukherji/ Parimal Parag*

*Prepared by: Karthik*

**Note:** *LaTeX template courtesy of UC Berkeley EECS dept.*

### Contents

|             |  |              |
|-------------|--|--------------|
| <b>11.1</b> | <b>A Remark On The Previous Tutorial</b>                                       | <b>11-2</b>  |
| <b>11.2</b> | <b>Moment Generating Functions</b>   | <b>11-2</b>  |
| <b>11.3</b> | <b>Characteristic Functions</b>  | <b>11-4</b>  |
| 11.3.1      | Joint Characteristic Function and Independence . . . . .                       | 11-5         |
| 11.3.2      | Characteristic Function and Moments . . . . .                                  | 11-7         |
| 11.3.3      | Characteristic Function and Distribution . . . . .                             | 11-7         |
| <b>11.4</b> | <b>Jointly Gaussian Random Variables</b>                                       | <b>11-7</b>  |
| 11.4.1      | Jointly Gaussian Random Variables, Independence and Uncorrelatedness . . . . . | 11-9         |
| 11.4.2      | Non-Singularity and Singularity of Covariance Matrix . . . . .                 | 11-10        |
| <b>11.5</b> | <b>Exercises</b>   | <b>11-11</b> |

## 11.1 A Remark On The Previous Tutorial

Towards the end of the previous tutorial, we discussed about the geometric interpretation of the conditional expectation  $E[Y|X]$  for finite second moment random variables  $X$  and  $Y$  as the MMSE estimate of  $Y$ , given the observation  $X$ ; see <http://ece.iisc.ernet.in/~parimal/random/TASession10.pdf>, Section 10.4 for the details. Here, we shall note the following points:

1. Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let

$$S := \{Z : Z \text{ is } \mathcal{F}\text{-measurable, } E[Z^2] < \infty\}$$

denote the set of all  $\mathcal{F}$ -measurable random variables with finite second moment. Then,  $S$  is a vector space over the field  $\mathbb{F} = \mathbb{R}$ , with the usual addition and scalar multiplication operations on random variables.

2. For any two random variables (or vectors)  $U, V \in S$ ,

$$\langle U, V \rangle = E[UV]$$

is a well-defined notion of inner product between  $U$  and  $V$ .

We first prove that  $S$  is a vector space. Towards this, note that the zero random variable belongs to  $S$ . Further, if  $X, Y \in S$ , then  $E[X^2] < \infty$  and  $E[Y^2] < \infty$  by definition. Further, by the Cauchy-Schwartz inequality,

$$E[XY] \leq \sqrt{E[X^2]} \sqrt{E[Y^2]} < \infty.$$

Therefore,

$$E[(X + Y)^2] = E[X^2] + E[Y^2] + 2E[XY] < \infty,$$

thereby implying that  $X + Y \in S$ . Finally, for any  $\alpha \in \mathbb{R}$ ,  $\alpha X \in S$  if  $X \in S$ . Hence,  $S$  is a vector space.

To show that  $\langle U, V \rangle = E[UV]$  is a well-defined notion of inner product, we note the following:

1.  $\langle U, U \rangle = E[U^2] \geq 0$  for all  $U \in S$ , with equality if and only if  $E[U^2] = 0$ , which in turn is true if and only if  $U = 0$  almost surely.
2.  $\langle U, V \rangle = E[UV] = E[VU] = \langle V, U \rangle$ .
3. For any  $a, b \in \mathbb{R}$ ,

$$\langle aU + bV, W \rangle = E[(aU + bV)W] = aE[UW] + bE[VW] = a \langle U, W \rangle + b \langle V, W \rangle.$$

Since the above properties characterize an inner product, we conclude that  $\langle U, V \rangle = E[UV]$  is a well-defined notion of inner product between any two random variables  $U$  and  $V$  belonging to  $S$ .

## 11.2 Moment Generating Functions

In this section, we define the moment generating function of a random variable or a collection of random variables, and list down some properties.

**Definition 11.2.1.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X$  be an  $\mathcal{F}$ -measurable random variable. Then, the moment generating function of  $X$  is denoted by  $M_X(t)$  and is defined for all  $t \in \mathbb{R}$  as

$$M_X(t) = E[e^{tX}], \quad t \in \mathbb{R}.$$

*Remark 1.* Since for any  $t \in \mathbb{R}$ ,  $e^{tX}$  is a strictly positive random variable, it follows that  $E[e^{tX}]$  is well-defined (i.e., the problem of  $\infty - \infty$  does not arise). However,  $E[e^{tX}]$  could be equal to  $+\infty$  for some values of  $t \in \mathbb{R}$ .

*Remark 2.*  $M_X(0) = 1$ .

*Example:*

Let  $X \sim \text{Exp}(1)$ . Then,

$$\begin{aligned} E[e^{tX}] &= \int_0^{\infty} e^{tx} f_X(x) dx \\ &= \int_0^{\infty} e^{tx} e^{-x} dx \\ &= \int_0^{\infty} e^{-(1-t)x} dx \\ &= \begin{cases} \frac{1}{1-t}, & t < 1, \\ +\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

*Remark 3.* For a random variable  $X$  with pdf  $f_X$ , the moment generating function of  $X$  is analogous to the Laplace transform of  $f_X$ .

Since  $M_X(t)$  can be equal to  $+\infty$  for some values of  $t \in \mathbb{R}$ , such as in the above example, the usefulness of moment generating functions is limited only to the case when they are finite. However, for most of the commonly used distributions such as Bernoulli, Poisson, Gaussian, Geometric, Exponential, Uniform, Gamma, etc, the corresponding moment generating functions are finite on an open interval containing the origin. For instance, in the example presented above, we note that  $M_X(t) < \infty$  for all  $t \in (-\infty, 1)$ . In particular,  $M_X(t) < \infty$  for all  $t \in (-1, 1)$ . Such a condition of finiteness of moment generating functions on an open interval containing the origin usually suffices to speak about some nice properties of moment generating functions.

*Exercises:*

1. Find the moment generating function for the following distributions:

Ber( $p$ ), Poisson( $\lambda$ ), Exp( $\lambda$ ), Geo( $p$ ), uniform( $-a, a$ ),  $a > 0$ .

2. Show that for any  $a, b \in \mathbb{R}$ , we have

$$M_{aX+b}(t) = e^{tb} M_X(at) = e^{tb} M_{aX}(t), \quad t \in \mathbb{R}.$$

Due to the restrictive setting (of requiring finiteness on an open interval containing the origin) for working with moment generating functions, we shall not discuss moment generating functions in detail here. We shall move to discussing about an alternative class of well-behaved functions, namely characteristic functions.

### 11.3 Characteristic Functions

We first state the definition of the characteristic function of a random variable  $X$ .

**Definition 11.3.1.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X$  be an  $\mathcal{F}$ -measurable random variable. Then, the characteristic function of  $X$  is denoted by  $\Phi_X(\omega)$  and is defined for all  $\omega \in \mathbb{R}$  as

$$\Phi_X(\omega) = E[e^{j\omega X}], \quad \omega \in \mathbb{R},$$

where  $j = \sqrt{-1}$  is the imaginary number.

**Note that the value of the characteristic function could be a complex number.** Unlike moment generating functions that may be equal to  $+\infty$ , we note the following important property of characteristic functions:

**Proposition 11.3.2.** For any random variable  $X$ ,  $|\Phi_X(\omega)| \leq 1$  for all values of  $\omega \in \mathbb{R}$ , where for any complex number  $z = a + jb$ ,  $|z| = \sqrt{a^2 + b^2}$ .

*Proof.* We have

$$E[e^{j\omega X}] = E[\cos(\omega X) + j \sin(\omega X)] = E[\cos(\omega X)] + jE[\sin(\omega X)], \quad (11.1)$$

from which it follows that for all  $\omega \in \mathbb{R}$ , we have

$$\begin{aligned} |\Phi_X(\omega)| &= \sqrt{(E[\cos(\omega X)])^2 + (E[\sin(\omega X)])^2} \\ &\stackrel{(a)}{\leq} \sqrt{E[\cos^2(\omega X)] + E[\sin^2(\omega X)]} \\ &= 1, \end{aligned}$$

where (a) above follows from the fact that  $(E[Z])^2 \leq E[Z^2]$ , and the last line above follows from the fact that  $\cos^2(\omega X) + \sin^2(\omega X) = 1$  for all  $\omega \in \mathbb{R}$ .  $\square$

**Thus, although the value of the characteristic function could be a complex number, it is always bounded in magnitude between 0 and 1.**

*Example:*

Let  $X \sim \mathcal{N}(0, 1)$ . Then,

$$\Phi_X(\omega) = E[\cos(\omega X) + j \sin(\omega X)] = E[\cos(\omega X)] + jE[\sin(\omega X)].$$

We note that

$$E[\sin(\omega X)] = \int_{-\infty}^{\infty} \sin(\omega x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

Therefore, we have

$$\begin{aligned} \Phi_X(\omega) &= E[\cos(\omega X)] \\ &= \int_{-\infty}^{\infty} \cos(\omega x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= 2 \int_0^{\infty} \cos(\omega x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \end{aligned}$$

We now differentiate both sides of the above equation with respect to  $\omega$ . Noting that  $|\cos(\omega x)| \leq 1$  for all  $\omega \in \mathbb{R}$  and using the “bounded convergence theorem” to pass  $\frac{d}{d\omega}$  through the integral, we get

$$\begin{aligned}
 \frac{d}{d\omega} \Phi_X(\omega) &= 2 \int_0^\infty -x \sin(\omega x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
 &= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^\infty \sin(\omega x) \left( -x e^{-\frac{x^2}{2}} \right) dx \\
 &\stackrel{(a)}{=} \frac{2}{\sqrt{2\pi}} \left( \sin(\omega x) e^{-\frac{x^2}{2}} \Big|_0^\infty - \int_0^\infty e^{-\frac{x^2}{2}} \omega \cos(\omega x) dx \right) \\
 &= -\omega \int_0^\infty \cos(\omega x) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
 &= -\omega \Phi_X(\omega),
 \end{aligned}$$

where (a) above follows from integration by parts. Using the condition  $\Phi_X(0) = 1$ , the unique solution to the ODE

$$\begin{aligned}
 \frac{d}{d\omega} \Phi_X(\omega) &= -\omega \Phi_X(\omega), \quad \omega \in \mathbb{R}, \\
 \Phi_X(0) &= 1,
 \end{aligned}$$

is given by  $\Phi_X(\omega) = e^{-\frac{\omega^2}{2}}$ ,  $\omega \in \mathbb{R}$ .

*Exercises:*

1. Find the characteristic function for the following distributions:

Ber( $p$ ), Poisson( $\lambda$ ), Exp( $\lambda$ ), Geo( $p$ ), uniform( $-a, a$ ),  $a > 0$ .

2. Show that for any  $a, b \in \mathbb{R}$ , we have

$$\Phi_{aX+b}(\omega) = e^{j\omega b} \Phi_X(a\omega) = e^{j\omega b} \Phi_{aX}(\omega), \quad \omega \in \mathbb{R}.$$

Using this result and the expression for the characteristic function of  $Y \sim \mathcal{N}(0, 1)$  distribution, show that the characteristic function of  $\mathcal{N}(\mu, \sigma^2)$ , where  $\mu \in \mathbb{R}$  and  $\sigma > 0$ , is given by

$$\Phi_Y(\omega) = e^{j\omega\mu - \frac{1}{2}\omega^2\sigma^2}, \quad \omega \in \mathbb{R}.$$

### 11.3.1 Joint Characteristic Function and Independence

We now discuss the characteristic function of a random vector, and then its implications on independence of the elements of the random vector. We have the following definition for the joint characteristic function of  $n$  random variables  $X_1, \dots, X_n$ .

**Definition 11.3.3.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and for  $n \in \{1, 2, \dots\}$ , let  $X_1, \dots, X_n$  be  $\mathcal{F}$ -measurable random variables. Then, the joint characteristic function of  $X = (X_1, \dots, X_n)$  is denoted by  $\Phi_X(\omega)$  and defined for all  $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$  as

$$\Phi_X(\omega) = E[e^{j\omega^T X}] = E \left[ \exp \left( j \sum_{i=1}^n \omega_i X_i \right) \right], \quad \omega \in \mathbb{R}^n.$$

*Note:* We shall write  $\Phi_X(\omega)$  to denote both the characteristic function of a single random variable and the joint characteristic function of a collection  $(X_1, \dots, X_n)$  of random variables. The notation will be clear from the context without causing confusion.

*Remark 4.* Note that  $\Phi_X(\omega) = \Phi_{\omega^T X}(1)$  for all  $\omega \in \mathbb{R}^n$ .

*Remark 5.* The joint moment generating function of  $X = (X_1, \dots, X_n)$  is similarly defined as

$$M_X(t) = E[e^{t^T X}] = E \left[ \exp \left( \sum_{i=1}^n t_i X_i \right) \right], \quad t \in \mathbb{R}^n.$$

We now have the following important result concerning independence, which we state without proof.

**Theorem 11.3.4.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and for  $n \in \{1, 2, \dots\}$ , let  $X_1, \dots, X_n$  be  $\mathcal{F}$ -measurable random variables. Then,  $X_1, \dots, X_n$  are mutually independent if and only if

$$\Phi_X(\omega) = \prod_{i=1}^n \Phi_{X_i}(\omega_i)$$

holds for all  $(\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ .

*Remark 6.* In order to check independence of  $n = 2$  random variables  $X_1$  and  $X_2$ , it DOES NOT SUFFICE to check that

$$E[e^{j(\omega X_1 + \omega X_2)}] = \Phi_{X_1}(\omega) \Phi_{X_2}(\omega)$$

holds for all  $\omega \in \mathbb{R}$ . This is weaker than checking that

$$E[e^{j(\omega_1 X_1 + \omega_2 X_2)}] = \Phi_{X_1}(\omega_1) \Phi_{X_2}(\omega_2)$$

holds for all  $\omega_1, \omega_2 \in \mathbb{R}$ , which is what the statement of the theorem demands. See the exercise below.

*Exercise:*

Let  $X$  and  $Y$  be two random variables whose joint density is given by

$$f(x, y) = \frac{1}{4} (1 + xy(x^2 - y^2)), \quad |x| < 1, \quad |y| < 1.$$

Then, show that for all  $\omega \in \mathbb{R}$ ,

$$E[e^{j(\omega X + \omega Y)}] = \Phi_X(\omega) \cdot \Phi_Y(\omega),$$

but  $X$  and  $Y$  are not independent.

*Remark 7.* An analogous if and only if result regarding independence holds for the joint moment generating function, provided it is finite on a open ball in  $\mathbb{R}^n$  containing the origin.

### 11.3.2 Characteristic Function and Moments

The following result gives a method to obtain the moments from the characteristic function. Note that this result is applicable only knowing beforehand that such a moment exists and is finite. We now state the result without proof.

**Theorem 11.3.5.** *Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X$  be an  $\mathcal{F}$ -measurable random variable such that  $E[|X|^n] < \infty$  for some  $n \in \{1, 2, \dots\}$ . Then,*

$$E[X^n] = (-j)^n \frac{d^n}{d\omega^n} \Phi_X(\omega) \Big|_{\omega=0}.$$

*Remark 8. An analogous result is true for moment generating functions, which gives it the name “moment” generating function.*

### 11.3.3 Characteristic Function and Distribution

The term “characteristic” in characteristic function refers to the fact that it characterizes a distribution (CDF) completely and uniquely. In other words, given two random variables that have identical characteristic functions, it follows that both random variables have identical distribution, and vice-versa. This is stated formally below.

**Theorem 11.3.6.** *Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X$  and  $Y$  be two  $\mathcal{F}$ -measurable random variables such that*

$$\Phi_X(\omega) = \Phi_Y(\omega) \text{ for all } \omega \in \mathbb{R}.$$

*Then,  $F_X(x) = F_Y(x)$  for all  $x \in \mathbb{R}$ .*

*Remark 9. A similar property holds for moment generating functions if they are finite on an open interval containing the origin.*

## 11.4 Jointly Gaussian Random Variables

In this section, we discuss some properties of jointly Gaussian random variables, also known by other names such as Gaussian random vectors,  $\mathbb{R}^n$ -valued Gaussian random variables, or multivariate Gaussian random variable. We first give a definition of this below.

**Definition 11.4.1.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and for  $n \in \{1, 2, \dots\}$ , let  $X_1, \dots, X_n$  be  $\mathcal{F}$ -measurable random variables with finite means. Then,  $X_1, \dots, X_n$  are said to be jointly Gaussian if the joint characteristic function of  $X = (X_1, \dots, X_n)$  is of the form

$$\Phi_X(\omega) = e^{j\omega^T \mu - \frac{1}{2} \omega^T K \omega}, \quad \omega \in \mathbb{R}^n,$$

for some  $\mu \in \mathbb{R}^n$  and some positive semi-definite real matrix  $K \in \mathbb{R}^{n \times n}$ .

Indeed, it can be shown that  $\mu = (E[X_1], \dots, E[X_n])^T$  is the vector of means, and  $K = E[(X - \mu)(X - \mu)^T]$  is the covariance matrix.

*Remark 10. Some textbooks use the above as the definition for jointly Gaussian random variables, while some others use the following alternative definition:  $X_1, \dots, X_n$  are jointly Gaussian if every linear combination of  $X_1, \dots, X_n$  is Gaussian distributed. We show below that both these definitions are equivalent (i.e., one implies the other).*

**Proposition 11.4.2.** *The following statements are equivalent.*

1. *The joint characteristic function of  $X_1, \dots, X_n$  is of the form*

$$\Phi_X(\omega) = e^{j\omega^T \mu - \frac{1}{2}\omega^T K \omega}, \quad \omega \in \mathbb{R}^n,$$

*for some  $\mu \in \mathbb{R}^n$  and some positive semi-definite real matrix  $K \in \mathbb{R}^{n \times n}$ .*

2. *Every linear combination of  $X_1, \dots, X_n$  is Gaussian distributed.*

*Proof.* We first show that the first statement implies the second. Let

$$Y = a_1 X_1 + \dots + a_n X_n = a^T X$$

denote any linear combination of  $X_1, \dots, X_n$ , where  $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ . In order to show that  $Y$  is Gaussian distributed, we shall show that the characteristic function of  $Y$  resembles that of a Gaussian distribution. Towards this, for any  $t \in \mathbb{R}$ , we have

$$\begin{aligned} \Phi_Y(t) &= E[e^{jtY}] = E \left[ \exp \left( j \sum_{i=1}^n t a_i X_i \right) \right] \\ &= E \left[ \exp \left( j \sum_{i=1}^n \omega_i X_i \right) \right] \\ &= \Phi_X(\omega), \end{aligned}$$

where  $\omega = (\omega_1, \dots, \omega_n)$ , with  $\omega_i = t a_i$  for all  $i = 1, \dots, n$ . By the first statement of the proposition, we then have

$$\begin{aligned} \Phi_Y(t) &= \Phi_X(\omega) \\ &= e^{j\omega^T \mu - \frac{1}{2}\omega^T K \omega} \\ &= e^{jt(\omega^T \mu) - \frac{1}{2}t^2(a^T K a)} \\ &= e^{j t \nu - \frac{1}{2}t^2 \sigma^2}, \end{aligned}$$

where  $\nu = a^T \mu$  and  $\sigma^2 = a^T K a$ . Thus, we notice that the characteristic function of  $Y$  resembles that of a Gaussian distribution with mean  $\nu$  and variance  $\sigma^2$ . Since a characteristic function uniquely characterizes a distribution, we conclude that  $Y \sim \mathcal{N}(\nu, \sigma^2)$ .

We now show that the second statement implies the first. Towards this, for any  $\omega \in \mathbb{R}^n$ , we have

$$\Phi_X(\omega) = \Phi_{\omega^T X}(1).$$

Since  $\omega^T X$  is a linear combination of  $X_1, \dots, X_n$ , by the second statement, we have that this is Gaussian distributed with mean

$$\nu = E[\omega^T X] = \sum_{i=1}^n \omega_i E[X_i] = \omega^T \mu,$$

where  $\mu = (E[X_1], \dots, E[X_n])^T$ , and variance

$$\begin{aligned} \sigma^2 &= E[(Y - E[Y])^2] \\ &= E[(\omega^T \mu - \omega^T X)^2] \\ &= E[\omega^T (X - \mu)(X - \mu)^T \omega] \\ &= \omega^T K \omega, \end{aligned}$$



where  $K = E[(X - \mu)(X - \mu)^T]$ . Therefore, we get

$$\begin{aligned}\Phi_X(\omega) &= \Phi_{\omega^T X}(1) \\ &= e^{j\nu - \frac{1}{2}\sigma^2} \\ &= e^{j\omega^T \mu - \frac{1}{2}\omega^T K \omega},\end{aligned}$$

hence completing the proof.  $\square$

*Remark 11.* Note that  $X_1, \dots, X_n$  jointly Gaussian implies that EVERY linear combination is Gaussian distributed. In particular, this implies that for each  $i \in \{1, 2, \dots, n\}$ ,  $X_i$  is Gaussian distributed, since  $X_i$  may be expressed as  $\sum_{k=1}^n a_k X_k$ , with  $a_i = 1$  and  $a_k = 0$  for all  $k \neq i$ .

*Remark 12.* The all-zero linear combination is treated, by convention, as a Gaussian random variable that puts a mass of 1 at the origin  $0 \in \mathbb{R}^n$ .

### 11.4.1 Jointly Gaussian Random Variables, Independence and Uncorrelatedness

One of the most important properties of jointly Gaussian random variables is the relationship between independence. It is clear that if two random variables  $X$  and  $Y$  are independent, then they are uncorrelated. The following result shows that **the converse is true if  $X$  and  $Y$  are jointly Gaussian**.

**Theorem 11.4.3.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X_1, \dots, X_n$  be  $\mathcal{F}$ -measurable random variables that are jointly Gaussian with mean  $\mu \in \mathbb{R}^n$  and covariance matrix  $K$ . Then,  $X_1, \dots, X_n$  are mutually independent if and only if  $K$  is diagonal.

*Proof.* Exercise.

Hint: For showing independence, it suffices to show that

$$\Phi_X(\omega) = \prod_{i=1}^n \Phi_{X_i}(\omega_i)$$

holds for all  $\omega = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ .  $\square$

The requirement of **joint Gaussianity is very crucial** for the above result to hold. The following exercise shows that if  $X$  and  $Y$  are individually Gaussian, but not jointly Gaussian, then  $X$  and  $Y$  being uncorrelated need not imply that  $X$  and  $Y$  are independent.

*Exercise:*

Let  $X \sim \mathcal{N}(0, 1)$ . Further, let  $W$  be independent of  $X$ , with the distribution<sup>1</sup>

$$P(W = w) = \begin{cases} \frac{1}{2}, & w = \pm 1, \\ 0, & \text{otherwise.} \end{cases}$$

Define a new random variable  $Y$  as  $Y = WX$ .

1. Show that  $Y \sim \mathcal{N}(0, 1)$ .
2. Show that  $X$  and  $Y$  are uncorrelated.

---

<sup>1</sup>The distribution of  $W$  is known as Rademacher distribution.

3. Demonstrate that

$$X + Y = \begin{cases} 0, & \text{with probability } \frac{1}{2}, \\ 2X, & \text{with probability } \frac{1}{2}, \end{cases}$$

hence proving that the linear combination  $X + Y$  is not Gaussian, and hence that  $X$  and  $Y$  are not jointly Gaussian.

4. Finally, show that  $X$  and  $Y$  are not independent.

### 11.4.2 Non-Singularity and Singularity of Covariance Matrix

In this subsection, we show that for a collection  $X_1, \dots, X_n$  of jointly Gaussian random variables, their joint density is well-defined if and only if the covariance matrix is non-singular (i.e., has non-zero determinant). Further, we also demonstrate the implications of the covariance matrix being singular.

We have the following definition.

**Definition 11.4.4.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and let  $X_1, \dots, X_n$  be jointly Gaussian  $\mathcal{F}$ -measurable random variables, with mean  $\mu \in \mathbb{R}^n$  and covariance matrix  $K$ . If  $K$  is non-singular, then the joint density of  $X_1, \dots, X_n$  is denoted by  $f_X(x)$  for all  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  and is given by

$$f_X(x) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(K)}} e^{-\frac{1}{2}x^T K^{-1}x}, \quad x \in \mathbb{R}^n, \quad (11.2)$$

where  $\det(K)$  denotes the determinant of matrix  $K$ .

Clearly, we note that the above quantity is not well-defined when  $\det(K) = 0$  (in which case  $K^{-1}$  does not exist). We now examine this case in detail. We note that

$$K = \begin{bmatrix} E[(X_1 - \mu_1)^2] & E[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & E[(X_1 - \mu_1)(X_n - \mu_n)] \\ E[(X_2 - \mu_2)(X_1 - \mu_1)] & E[(X_2 - \mu_2)^2] & \cdots & E[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \cdots & \vdots \\ E[(X_n - \mu_n)(X_1 - \mu_1)] & E[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & E[(X_n - \mu_n)^2] \end{bmatrix}_{n \times n},$$

where  $\mu_i = E[X_i]$  for each  $i \in \{1, 2, \dots\}$ . If  $K$  is singular, then there exists a row of  $K$  that is linearly dependent on the remaining  $(n - 1)$  rows. Without loss of generality, we assume that the first row of  $K$  is linearly dependent on the remaining rows. This then implies that there exists  $(a_2, \dots, a_n) \in \mathbb{R}^{n-1}$ ,  $a \neq 0$ , such that

$$(\text{Row 1 of } K) = a_2 \cdot (\text{Row 2 of } K) + \cdots + a_n \cdot (\text{Row } n \text{ of } K).$$

This translates to the following set of equations:

$$E[(X_1 - \mu_1)^2] = a_2 E[(X_2 - \mu_2)(X_1 - \mu_1)] + \cdots + a_n E[(X_n - \mu_n)(X_1 - \mu_1)] \quad (11.3)$$

$$E[(X_1 - \mu_1)(X_2 - \mu_2)] = a_2 E[(X_2 - \mu_2)^2] + \cdots + a_n E[(X_n - \mu_n)(X_2 - \mu_2)] \quad (11.4)$$

$\vdots$

$$E[(X_1 - \mu_1)(X_n - \mu_n)] = a_2 E[(X_2 - \mu_2)(X_n - \mu_n)] + \cdots + a_n E[(X_n - \mu_n)^2]. \quad (11.5)$$

Multiplying (11.4) by  $(-a_2)$ ,  $\dots$ , (11.5) by  $(-a_n)$ , and adding with (11.3), we get

$$E[(X_1 - (a_2 X_2 + \cdots + a_n X_n))^2] = 0,$$

from which it follows that  $X_1 = a_2X_2 + \dots + a_nX_n$  almost surely.

Thus, if the covariance matrix is singular, it represents a redundancy in the information contained in the vector  $X = (X_1, \dots, X_n)$  in the sense that one of the random variables is a linear combination of the remaining variables. We then have the following protocol to determine the joint density of a collection  $X_1, \dots, X_n$  of jointly Gaussian random variables:

1. Check if the covariance matrix is non-singular. If so, the joint density exists and may be written in the form given in (11.2).
2. If the covariance matrix is singular, then there exists an element  $X_i$  which may be written as a linear combination of the remaining variables. Delete the row and column corresponding to  $X_i$  from the covariance matrix, and consider the new  $(n-1) \times (n-1)$  sub-matrix formed by the remaining variables.
3. If this sub-matrix is non-singular, a joint density may be written, but now with this sub-matrix as the covariance matrix, only for the remaining  $(n-1)$  variables. If the sub-matrix is singular, then follow the above procedure of deletion to obtain a lower dimensional sub-matrix, until a sub-matrix that is non-singular is obtained. A joint density may then be written for the variables participating in the sub-matrix.

## 11.5 Exercises

1. Show that

$$\overline{\Phi_X(\omega)} = \Phi_X(-\omega) \quad \text{for all } \omega \in \mathbb{R}.$$

*Note: For a complex number  $z = x + jy$ ,  $\bar{z}$  denotes its complex conjugate, i.e.,  $\bar{z} = x - jy$ .*

2. Let  $X$  be a continuous random variable with pdf  $f_X$ . Show that  $\Phi_X(\omega)$  is real-valued for all  $\omega \in \mathbb{R}$  if and only if  $f_X(x) = f_X(-x)$  for all  $x \in \mathbb{R}$ . Argue similarly for a discrete random variable with pmf  $p_X$ .
3. Let  $X = (X_1, X_2, X_3)$  be a zero mean Gaussian random vector with covariance matrix

$$K = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

- (a) Write down the distributions of  $X_1$ ,  $X_2$  and  $X_3$ .
  - (b) Does a joint density exist for  $X$ ? If not, what is the relationship between  $X_1$ ,  $X_2$  and  $X_3$ ?
  - (c) Does a joint density exist for  $X' = (X_1, X_2)$ ? If not, what is the relationship between  $X_1$  and  $X_2$ ?
  - (d) Does a joint density exist for  $Y' = (X_2, X_3)$ ? If not, what is the relationship between  $X_2$  and  $X_3$ ?
  - (e) Does a joint density exist for  $Z' = (X_1, X_3)$ ? If not, what is the relationship between  $X_1$  and  $X_3$ ?
4. (*Rotation of a jointly Gaussian distribution yielding independence*)  
Let  $X$  be a Gaussian random vector with

$$E[X] = \begin{bmatrix} 10 \\ 5 \end{bmatrix}, \quad \text{Cov}(X) = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

- (a) Compute the pdf of  $X$  explicitly.

- (b) Find a vector  $b$  and an orthogonal matrix  $U$  such that the vector  $Y = [Y_1 \ Y_2]^T$  defined by  $Y = U(X - b)$  is a mean zero Gaussian vector such that  $Y_1$  and  $Y_2$  are independent. (Hint: Use SVD.)
5. Let  $X \sim \mathcal{N}(0, 1)$  and  $Y \sim \mathcal{N}(0, 1)$  be two independent random variables. Define  $S = X + Y$  and  $D = X - Y$ .
- (a) Are  $X$  and  $Y$  jointly Gaussian? Justify your answer.
- (b) Are  $S$  and  $D$  jointly Gaussian? Justify your answer.
- (c) Let  $U$  be a random variable independent of  $X$  and  $Y$  having the distribution

$$P(U = u) = \begin{cases} \frac{1}{2}, & u = \pm 1, \\ 0, & \text{otherwise.} \end{cases}$$

Are  $S$  and  $UD$  jointly Gaussian? Justify your answer.

6. Let  $Y \sim \mathcal{N}(0, 1)$  and  $a > 0$ . Define

$$Z = \begin{cases} Y, & \text{if } |Y| \leq a, \\ -Y, & \text{if } |Y| > a. \end{cases}$$

Show that  $Z \sim \mathcal{N}(0, 1)$  by showing that the characteristic function of  $Y$  and  $Z$  are identical.

7. Let  $(X, Y)$  be jointly Gaussian (or bivariate normal) with mean 0 and covariance matrix  $K$ , where

$$K = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}.$$

Suppose that  $|\rho| < 1$  (and hence  $K$  is invertible). Then  $X$  and  $Y$  have a joint density  $f$ . Show that the conditional density  $f_{Y|X}(y|x)$  is the density of a univariate normal with mean  $x \frac{\sigma_Y}{\sigma_X} \rho$  and variance  $\sigma_Y^2(1 - \rho^2)$ .