E2:202 Random Processes **Nov. 13, 2020**

Tutorial 6: Characteristic Functions and Jointly Gaussian RVs

Lecturer: Parimal Parag *TA: Arvind* *Scribes: Krishna Chaythanya KV*

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

6.1 Characteristic Functions

Definition (Characteristic Function)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space, and let X be an $\mathcal{F}$-measurable RV. Then, the characteristic function of X denoted as $\phi_X(\omega)$, for $\omega \in \mathbb{R}$, is defined as

$$\phi_X(\omega) = \mathbb{E}\left[\exp(j\omega X)\right], \quad \omega \in \mathbb{R},$$

where $j = \sqrt{-1}$.

Remark. $|\phi_X(\omega)| \leq 1, \forall \omega \in \mathbb{R}$. Hence, the characteristic function is bounded in magnitude.

We now compute the characteristic function of the standard normal RV.

Example (Characteristic function of the standard normal)**.** Let $X \sim \mathcal{N}(0,1)$. For a fixed $\omega \in \mathbb{R}$,

$$\begin{aligned}\phi_X(\omega) &= \mathbb{E}\left[\exp(j\omega X)\right] \\ &= \underbrace{\mathbb{E}\left[\cos(\omega X)\right]}_{L_1(\omega)} + j\underbrace{\mathbb{E}\left[(\sin(\omega X))\right]}_{L_2(\omega)}.\end{aligned}$$

$$\text{Now, } L_2(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin(\omega x)\, e^{-\frac{x^2}{2}}\, dx = 0,$$

since $\sin(\omega x) : \mathbb{R} \to [-1,1]$ is an odd function.

$$\begin{aligned}\text{Further, } L_1(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos(\omega X)\, e^{-\frac{x^2}{2}}\, dx \\ &= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} \cos(\omega X)\, e^{-\frac{x^2}{2}}\, dx.\end{aligned}$$

The second equality is valid because $\cos(\omega x) : \mathbb{R} \to [-1,1]$ is an even function. We now take the derivative of $\phi_X(\omega)$ and use the *bounded convergence theorem* (because $|\cos(\omega X)| \le 1$) to get

$$\begin{aligned}
\frac{d\phi_X(\omega)}{d\omega} &= \frac{dL_1(\omega)}{d\omega} \\
&= 2\int_0^\infty -x\sin(\omega x)\frac{1}{\sqrt{2\pi}}e^{-\frac{-x^2}{2}}\,dx \\
&= \frac{2}{\sqrt{2\pi}}\int_0^\infty \sin(\omega x)\left(-xe^{-\frac{x^2}{2}}\right)dx \\
&= -\omega\int_0^\infty \cos(\omega x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx \quad \text{(integration by parts)} \\
&= -\omega L_1(\omega) = -\omega\phi_X(\omega).
\end{aligned}$$

The unique solution to the above ODE is

$$\phi_X(\omega) = e^{-\frac{\omega^2}{2}}, \quad \omega \in \mathbb{R}.$$

Exercise 6.1. Show that for $a, b \in \mathbb{R}$

$$\phi_{aX+b}(\omega) = e^{j\omega b}\phi_{aX}(\omega) = e^{j\omega b}\phi_X(a\omega).$$

Use this to show that for $Y \sim \mathcal{N}(\mu, \sigma^2)$, for $\mu \in \mathbb{R}, \sigma^2 > 0$,

$$\phi_Y(\omega) = e^{j\omega\mu - \omega^2\sigma^2/2}$$

6.1.1 Joint Characteristic Functions

Definition (Joint Characteristic Function)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X_1, X_2, \ldots, X_n$ be $\mathcal{F}$-measurable RVs. Then, the joint characteristic function of $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is denoted by $\phi_{\boldsymbol{X}}(\boldsymbol{\omega})$ and is defined for all $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n$ as

$$\phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \mathbb{E}\left[e^{j\boldsymbol{\omega}^\intercal \mathbf{X}}\right] = \mathbb{E}\left[\exp\left(j\sum_{i=1}^n \omega_i X_i\right)\right], \ \omega_i \in \mathbb{R}, \ \forall i \in [n].$$

Remark. Note that $\boldsymbol{\omega}$ is an n-length vector of real numbers, which are not necessarily identical.

We now state two important results about characteristic functions.

Theorem (Joint Characteristic Functions and Independence)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for $n \in \mathbb{N}$, let $X_1, X_2, \ldots, X_n$ be $\mathcal{F}$-measurable RVs. Then $X_1, X_2, \ldots, X_n$ are mutually independent iff

$$\phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \prod_{i=1}^n \phi_{X_i}(\omega_i) \quad \forall \boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n.$$

Remark. Note that the condition stated above must hold for all $\boldsymbol{\omega}$ and for not just $\boldsymbol{\omega} = (\hat{\omega}, \hat{\omega}, \ldots, \hat{\omega}) \in \mathbb{R}$, for independence to hold true.

Theorem (Characteristic Functions and Distributions)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X and Y be two $\mathcal{F}$-measurable RVs such that

$$\phi_X(\omega) = \phi_Y(\omega), \quad \forall \omega \in \mathbb{R}.$$

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

Remark. Thus, in order to show that two RVs have the same distribution, it suffices to show that their characteristic functions are identical.

6.2 Jointly Gaussian Random Variables

Before we proceed to study joint Gaussianity in general, here is a small exercise on independent Gaussian RVs.

Exercise 6.2. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and $X_1, X_2, \ldots, X_n$ are independent $\mathcal{F}$-measurable RVs. Then, any linear combination $Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n$, for $a_i \in \mathbb{R}, i \in [n]$, is a Gaussian RV.
Hint: **We know the form of the characteristic function of a RV $Z = \mathcal{N}\left(\mu, \sigma^2\right)$. Use the independence property to compute $\phi_Y(\omega)$, for $\omega \in \mathbb{R}$, and use the theorem about characteristic functions and distributions stated above.**

Exercise 6.3 (Supplementary exercise on Normal RVs)**.** Let X be a RV with the $\mathcal{N}(0,1)$ distribution and let $a > 0$. Show that the RV Y given by

$$Y = \begin{cases} X, & \text{if } |X| < a, \\ -X, & \text{if } |X| \geq a, \end{cases}$$

has the $\mathcal{N}(0,1)$ distribution.

Let us now define jointly Gaussian RVs.

Definition (Jointly Gaussian RV)**.** A collection $(X_i : i \in I)$, for some index set $I \subseteq \mathbb{R}$, of RVs has a joint Gaussian distribution if every *finite* linear combination of $(X_i : i \in I)$ is a Gaussian RV.

Remark. A random vector X is called a Gaussian random vector if its co-ordinate random variables are jointly Gaussian.

Let $X_1, X_2, \ldots, X_n$ be $\mathcal{F}$-measurable jointly Gaussian RVs with finite means, and let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$. Further let

$$\boldsymbol{\mu} \triangleq \left(\mathbb{E}[X_1], \mathbb{E}[X_2], \ldots, \mathbb{E}[X_n]\right),$$

and

$$\boldsymbol{K} \triangleq \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^\intercal\right].$$

Remark. $\boldsymbol{K}$ is called the covariance matrix of the random vector $\boldsymbol{X}$. Note that $\boldsymbol{K}$, by definition, is *positive semi-definite*, i.e.,

$$\boldsymbol{x}^\intercal \boldsymbol{K} \boldsymbol{x} \geq 0, \quad \boldsymbol{x} \in \mathbb{R}^n.$$

Then, the following lemma holds.

Lemma (Characteristic Function of Jointly Gaussian RVs)**.** If $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is a Gaussian random vector, then its characteristic function is given by

$$\phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \mathbb{E}\left[e^{j\boldsymbol{\omega}^\intercal \mathbf{X}}\right] = \exp\left(j\boldsymbol{\omega}^\intercal \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\omega}^\intercal \boldsymbol{K} \boldsymbol{\omega}\right), \quad \forall \omega \in \mathbb{R}^n.$$

Proof. Fix $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n$. Then, the RV $\boldsymbol{\omega}^\intercal \mathbf{X}$ is Gaussian (**why?**)

$$\text{with mean } \mathbb{E}\left[\boldsymbol{\omega}^\intercal \boldsymbol{X}\right] = \boldsymbol{\omega}^\intercal \boldsymbol{\mu},$$

$$\text{and variance } \mathbb{E}\left[(\boldsymbol{\omega}^\intercal \boldsymbol{X} - \boldsymbol{\omega}^\intercal \boldsymbol{\mu})^\intercal (\boldsymbol{\omega}^\intercal \boldsymbol{X} - \boldsymbol{\omega}^\intercal \boldsymbol{\mu})\right] = \boldsymbol{\omega}^\intercal \boldsymbol{K} \boldsymbol{\omega}.$$

Since the characteristic function of a Gaussian RV Y with mean $\tilde{\mu} \in \mathbb{R}$ and variance $\theta^2 > 0$ is given by $\phi_Y(\omega) = e^{j\omega\tilde{\mu} - \omega^2\theta^2/2}$, it follows that

$$\phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \phi_{\boldsymbol{\omega}^\intercal \mathbf{X}}(1) = \exp\left(j\boldsymbol{\omega}^\intercal \boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\omega}^\intercal \boldsymbol{K} \boldsymbol{\omega}\right).$$

□

Remark. In fact, it can be show that a random vector $\boldsymbol{X}$ with characteristic function $\phi_X(\boldsymbol{\omega}) = \exp\left(j\boldsymbol{\omega}^\intercal \boldsymbol{\delta} - \frac{1}{2}\boldsymbol{\omega}^\intercal \boldsymbol{Q} \boldsymbol{\omega}\right)$ for some $\boldsymbol{\delta} \in \mathbb{R}^n$ and a positive semi-definite real matrix $\boldsymbol{Q} \in \mathcal{M}_{n\times n}$. The proof of this statement is left as an exercise.
Hint: **We need to show that any finite linear combination** $Y = a_1X_1 + \ldots + a_nX_n$ **for** $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^b$ **is Gaussian. To show this, it suffices to show that the characteristic function** $\phi_Y(\omega), \omega \in \mathbb{R}$, **is in the form of the characteristic function of a Gaussian RV** $\forall (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$**.**

From the above observation, it follows that the vector $\boldsymbol{\mu} = (\mathbb{E}[X_1], \mathbb{E}[X_2], \ldots, \mathbb{E}[X_n])$, and the covariance matrix $\boldsymbol{K}$ are sufficient to fully characterize a Gaussian random vector.

Exercise 6.4. Let $\boldsymbol{X} = (X_1, X_1, \ldots, X_n)$ be a Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{K}$. Show that $(X_i : i \in [n])$ are independent iff $\boldsymbol{K}$ is diagonal.
Hint: **Use the theorem on joint characteristic functions and independence.**

Proposition. A Gaussian random vector $\boldsymbol{X}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{K}$, such that $\boldsymbol{K}$ is non-singular, has a pdf given by

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}(\det(\boldsymbol{K}))^{\frac{1}{2}}} \exp\left(-\frac{(\boldsymbol{x} - \boldsymbol{\mu})^\intercal \boldsymbol{K}(\boldsymbol{x} - \boldsymbol{\mu})}{2}\right).$$

Proof. Let $\boldsymbol{X}$ be a Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{K}$. Read the following fact about non-singular positive semi-definite (or positive definite) matrices.

Fact 6.5. Since $\boldsymbol{K}$ is positive semi-definite and non-singular, it can be written as $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\intercal$, where $\boldsymbol{U}$ is an *orthonormal matrix*, i.e., $\boldsymbol{U}\boldsymbol{U}^\intercal = \boldsymbol{U}^\intercal\boldsymbol{U} = \boldsymbol{I}$, and $\boldsymbol{\Lambda}$ is a diagonal matrix with positive eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ along the diagonal.

Let $\boldsymbol{Y} = \boldsymbol{U}^\intercal(\boldsymbol{X} - \boldsymbol{\mu})$. Then $\boldsymbol{Y}$ is a Gaussian vector (**why?**) of mean $\mathbf{0}$ and covariance matrix

$$\boldsymbol{Q} = \mathbb{E}\left[\boldsymbol{Y}\boldsymbol{Y}^\intercal\right] = \boldsymbol{U}^\intercal\boldsymbol{K}\boldsymbol{U} = \boldsymbol{\Lambda}.$$

Since $\boldsymbol{\Lambda}$ is diagonal, $\boldsymbol{Y}$ is a vector of *independent* Gaussian RVs (follows from the exercise), and $Y_i \sim \mathcal{N}(0, \lambda_i)$. Hence, $\boldsymbol{Y}$ has the joint pdf

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\lambda_i}} \exp\left(-\frac{y_i^2}{2\lambda_i}\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\lambda_1\lambda_2\ldots\lambda_n}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{y_i^2}{\lambda_i}\right). \end{aligned}$$

Note that $\det(\boldsymbol{K}) = \prod_{i=1}^{n} \lambda_i$ and $\sum_{i=1}^{n} \frac{y_i^2}{\lambda_i} = \boldsymbol{y}^\intercal \boldsymbol{\Lambda}^{-1} \boldsymbol{y}$ since $\boldsymbol{\Lambda}^{-1}$ has terms $(1/\lambda_i)_{i\in[n]}$ along its diagonals and zeros everywhere else. Thus,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{1}{\sqrt{2\pi}^{\frac{n}{2}} \sqrt{\det(\boldsymbol{K})}} \exp\left(-\frac{\boldsymbol{y}^\intercal \boldsymbol{\Lambda}^{-1} \boldsymbol{y}}{2}\right).$$

Observe that $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{Y} + \boldsymbol{\mu}$ and hence, by the transformation of random vectors, since $|\det(\boldsymbol{U})| = 1$ (and the determinant of the Jacobian inverse is 1),

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) &= f_{\boldsymbol{Y}}\left(\boldsymbol{U}^\intercal(\boldsymbol{x}-\boldsymbol{\mu})\right) \\ &= \frac{1}{\sqrt{2\pi}^{\frac{n}{2}} \sqrt{\det(\boldsymbol{K})}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^\intercal(\boldsymbol{x}-\boldsymbol{\mu})\right) \\ &= \frac{1}{\sqrt{2\pi}^{\frac{n}{2}} \sqrt{\det(\boldsymbol{K})}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\intercal \boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right). \quad (\text{since } \boldsymbol{K}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^\intercal) \end{aligned}$$

□

This simple exercise uses the Fact 6.5 presented in the proof above.

Exercise 6.6. Show that a positive semi-definite (symmetric) matrix $\boldsymbol{V}$ has a square root i.e., there exists a symmetric matrix $\boldsymbol{W}$ such that $\boldsymbol{W}^2 = \boldsymbol{V}$.

We have seen in the proposition above that, if the covariance matrix $\boldsymbol{K}$ is non-singular, the pdf of a Gaussian random vector is determined entirely by $\boldsymbol{\mu}$ and $\boldsymbol{K}$. But what if $\boldsymbol{K}$ is singular, i.e., $\det(\boldsymbol{K}) = 0$?

In such a situation, the random vector $\boldsymbol{X}$ *does not* have a pdf. We shall now analyze this situation.

Illustration. Let $\boldsymbol{X}$ be jointly Gaussian with covariance matrix $\boldsymbol{K}$ such that $\det(\boldsymbol{K}) = 0$. Hence, $\lambda_i = 0$ for some $i \in [n]$, where the λ_is are obtained from the decomposition in Fact 6.5.

In other words, there is a vector $\boldsymbol{\alpha}$ (which is an eigenvector of eigenvalue 0) such that $\boldsymbol{\alpha}^\intercal \boldsymbol{K}\boldsymbol{\alpha} = 0$. However,

$$\begin{aligned} \boldsymbol{\alpha}^\intercal \boldsymbol{K}\boldsymbol{\alpha} &= \operatorname{Var}(\boldsymbol{\alpha}^\intercal \boldsymbol{X}) \quad \text{Show this!} \\ &= 0. \end{aligned}$$

This implies, from the property of non-negative RVs with mean zero (we've seen this property in a previous tutorial), that

$$\boldsymbol{\alpha}^\intercal(\boldsymbol{X}-\boldsymbol{\mu}) = 0 \text{ w.p. } 1, \text{ or,}$$

that if $\boldsymbol{\mu} = 0$, there exists some linear combination of $\boldsymbol{X}$ that equals 0, or that if $\boldsymbol{\mu} = 0$, the X_is are *not* linearly independent. □

Exercise 6.7. Let X and Y be RVs such that the vector $\boldsymbol{Z} = (X, Y)$ is a Gaussian random vector with zero mean and covariance matrix

$$\boldsymbol{K} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

ρ is the correlation co-efficient between X and Y. Find the joint density of $X + Y$ and $X - Y$.

Exercise 6.8. let $\boldsymbol{X} = (X_1, X_2, X_3)$ be a zero mean Gaussian random vector with covariance matrix

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

1. Write down the marginal distributions of X_1, X_2, and X_3.
2. Does a joint density exist for $\boldsymbol{X}$?