Real-Valued Gaussian Random Vectors

Prof. Joon Ho Cho
Department of Electrical Engineering, POSTECH

- Review. The received signal contains a noise $N(t)$ often modeled by an **additive white Gaussian noise (AWGN)**. To understand it, we started to review a pre-requisite of this course, *Probability, Random Variables, and Random Processes*.
  - statistical independence
    - of multiple events
    - of multiple random variables
  - correlation, covariance, and correlation coefficient
    - of two real-valued random variables
    - of two complex-valued random variables.

- Preview. We will examine the notion of **Gaussianity**.
  - characteristic function of a real-valued random variable
  - real-valued Gaussian random vectors

- Characteristic Function (CF)
  - Q. What is the characteristic function of a real-valued random variable X?
  - A. It is denoted by $\Phi_X(\omega)$ and defined as

$$\Phi_X(\omega) \triangleq \mathbb{E}\left\{e^{\mathrm{j}\omega X}\right\},$$

    i.e., it is related to the Fourier transform of the PDF $f_X(x)$.
    * The CF fully characterizes the random variable.
    * In general, the collection of all the nth order moments for $n \in \mathbb{N}$ characterizes the random variable.

    * There are some random variables where a finite number of moments fully characterizes the random variables.

– Q. What is the joint CF of a real-valued random vector $\underline{X}$?

– A. It is denoted by $\Phi_{\underline{X}}(\underline{\omega})$ and defined as

$$\Phi_{\underline{X}}(\underline{\omega}) \triangleq \mathbb{E}\left\{e^{\mathrm{j}\underline{\omega}^T \underline{X}}\right\},$$

i.e., it is related to the multi-dimensional Fourier transform of the joint PDF $f_{\underline{X}}(\underline{x})$.

* The CF fully characterizes the random vector.
* There are some random variables where a finite number of moments fully characterizes the random variables.
* Random vectors are independent iff the joint CF factors.

- Real-valued Gaussian random variable
- In what follows, "real-valued" is omitted.
  - Def. A random variable X is Gaussian if its PDF $f_X(x)$ is parameterized by only two constants as

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (1)$$

    * It turns out that $\mu = \mathbb{E}\{X\}$ and $\sigma^2 = \mathrm{Var}\{X\} > 0$.
    * Often denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$. ('Normal' means Gaussian.)
    * Fig. PDF, a bell curve

– Def. A random variable X is Gaussian if its CF $\Phi_X(\omega)$ is parameterized by only two constants as

$$\Phi_X(\omega) = e^{\mathrm{j}\omega\mu - \frac{\omega^2\sigma^2}{2}} = \exp\left(\mathrm{j}\omega\mu - \frac{\omega^2\sigma^2}{2}\right).$$

* Fig. CF when $\mu = 0$, bell curve, uncertainty principle

* It turns out that $\mu = \mathbb{E}\{X\}$ and $\sigma^2 = \mathrm{Var}\{X\}$.
* If $\sigma^2 > 0$, then the PDF is given by (1).
* The CF allows $\sigma^2 = 0$, which implies $f_X(x) = \delta(x - \mu)$.
* Thus, the CF is more useful than the PDF for Gaussian random variables.

– Properties of a Gaussian random variable

* The PDF and the CF of a Gaussian random variable are parameterized only by the two moments: the mean μ and the variance σ^2.

* $$\mathbb{E}[(X-\mu)^n] = \begin{cases} 0, & \text{for odd } n \\ \dfrac{n!}{\left(\frac{n}{2}\right)!}\left(\dfrac{\sigma}{\sqrt{2}}\right)^n, & \text{for even } n \end{cases}$$

/ex/
$$\begin{aligned} \mathbb{E}[(X-\mu)^2] &= 1\cdot\sigma^2 \\ \mathbb{E}[(X-\mu)^4] &= 1\cdot 3\cdot\sigma^4 \\ \mathbb{E}[(X-\mu)^6] &= 1\cdot 3\cdot 5\cdot\sigma^6 \\ &\vdots \end{aligned}$$

- Central Limit Theorem (CLT) and a Gaussian random variable
  * If X_i's are i.i.d. random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2$, then

$$Y_n \triangleq \frac{(X_1 - \mu) + (X_2 - \mu) + \cdots + (X_n - \mu)}{\sigma\sqrt{n}} \xrightarrow{D} Y \sim \mathcal{N}(0, 1).$$

  * The sequence of random variables Y_n converges to a zero-mean unit-variance Gaussian random variable in distribution.
  * In other words, the CDF or the CF of Y_n is very close to the CDF or the CF of a zero-mean unit-variance Gaussian if n is sufficient large.
- Fig. (An LNA is equivalent to a system that multiplies the gain g and adds an additive noise $N(t)$.)
  * $N(t)$ is a noise generated by electrons that bounce back and forth while they flow.
  * Since there are a lot of electrons, we can apply the CLT to claim that we observe a Gaussian noise at each time instant.

- Real-valued Gaussian random vector
  - Def. A random vector $\underline{X} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$ is Gaussian if its joint PDF $f_{\underline{X}}(\underline{x})$ is parameterized by a length-N vector and a size N symmetric positive definite matrix as

$$\begin{aligned} f_{\underline{X}}(\underline{x}) &= \frac{1}{\sqrt{2\pi}^N \sqrt{\det C}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T C^{-1}(\underline{x}-\underline{\mu})} \\ &= \frac{1}{\sqrt{2\pi}^N \sqrt{\det C}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^T C^{-1}(\underline{x}-\underline{\mu})\right). \quad (2) \end{aligned}$$

    * It turns out that $\underline{\mu} = \mathbb{E}\{\underline{X}\}$ is the mean vector and $C = \mathbb{E}\{(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T\} = C_{\underline{X}\underline{X}} = \mathrm{Cov}\{\underline{X}\} = C^T \succ 0$ is the covariance matrix of $\underline{X}$.
    * Often denoted by $\underline{X} \sim \mathcal{N}(\underline{\mu}, C)$. (Again, 'N' means Normal or Gaussian.)
    * Often called a multivariate ($\longleftrightarrow$ uni-variate, bi-variate) Gaussian random variable.

* Simplify! $N = 1$

* Simplify! $N = 2$

  · When $N = 2$, $\underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$, $C = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$, $\rho \neq \pm 1 \Longrightarrow$

$$
\begin{aligned}
f_{\underline{X}}(\underline{x}) &= f_{X_1,X_2}(x_1, x_2) \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right.\right. \\
&\qquad \left.\left. -2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}\right]
\end{aligned}
$$

  · For $N = 2$ (bi-variate Gaussian), $\underline{X} \sim N(\mu_1, \mu_2; \sigma_1, \sigma_2, \rho)$ or $\underline{X} \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2, \rho)$.

- Fig. $\rho = 0$, $0 < \rho < 1$, $-1 < \rho < 0$

- Any N-variate Gaussian distribution can be obtained by rotating, reflecting, dilating, translating the Gaussian distribution $\mathcal{N}(\underline{0}, I_N)$.

– Def. A random vector $\underline{X}$ is Gaussian if its joint CF $\Phi_{\underline{X}}(\underline{\omega})$ is parameterized by a length-N vector and a size N symmetric positive **semi-definite** matrix as

$$\Phi_{\underline{X}}(\underline{\omega}) = e^{\mathrm{j}\underline{\omega}^T \underline{\mu} - \frac{\underline{\omega}^T C \underline{\omega}}{2}} = \exp\left(\mathrm{j}\underline{\omega}^T \underline{\mu} - \frac{\underline{\omega}^T C \underline{\omega}}{2}\right).$$

* It turns out that $\underline{\mu} = \mathbb{E}\{\underline{X}\}$ and $C = C_{\underline{X}\underline{X}} = C^T \succeq \mathbf{0}$.
* If $C \succ 0$, then the joint PDF is given by (2).
* The joint CF allows un-invertible C, i.e., C^{-1} does not exist, which leads to the use of a multi-dimensional Dirac delta function for the PDF. We encounter an impulse fence.
* Thus, the joint CF is more useful than the joint PDF for Gaussian random vectors.
* If $\underline{X}$ is Gaussian, then $X_1, X_2, \cdots, X_N$ are called jointly Gaussian (random variables).

- Def. Random vectors $\underline{X}_1, \underline{X}_2, ..., \underline{X}_N$ are jointly Gaussian if

$$\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_N \end{bmatrix}$$

is Gaussian.

- Properties of a Gaussian random vector
  * The joint PDF and the joint CF of a Gaussian random vector are parameterized only by the two moments: the mean vector $\underline{\mu}$ and the covariance matrix C.
  * Theorem. An affine transform $\underline{Y} = A\underline{X} + \underline{b}$ of $\underline{X} \sim N(\underline{\mu}, C)$ leads to $\underline{X} \sim N(A\underline{\mu} + \underline{b}, ACA^T)$.

- Specialize! Each element of a Gaussian random vector is a Gaussian random variable.
- Specialize! A weighted sum of some elements of a Gaussian random vector is a Gaussian random variable.
- Specialize! Any collection of the elements of a Gaussian random vector is a Gaussian random vector.

∗ Proof.

$$
\begin{aligned}
\Phi_{\underline{Y}}(\underline{\omega}) &\triangleq \mathbb{E}\left[e^{\mathrm{j}\underline{\omega}^T \underline{Y}}\right] = \mathbb{E}\left[e^{\mathrm{j}\underline{\omega}^T (A\underline{X}+\underline{b})}\right] \\
&= e^{\mathrm{j}\underline{\omega}^T \underline{b}}\, \mathbb{E}\left[e^{\mathrm{j}\left(A^T \underline{\omega}\right)^T \underline{X}}\right] \\
&= e^{\mathrm{j}\underline{\omega}^T \underline{b}}\, \Phi_{\underline{X}}\left(A^T \underline{\omega}\right) \\
&= e^{\mathrm{j}\underline{\omega}^T \underline{b}} \exp\left(\mathrm{j}\left(A^T \underline{\omega}\right)^T \underline{\mu} - \frac{1}{2}\left(A^T \underline{\omega}\right)^T C \left(A^T \underline{\omega}\right)\right) \\
&= \exp\left(\mathrm{j}\underline{\omega}^T \left(A\underline{\mu} + \underline{b}\right) - \frac{1}{2}\underline{\omega}^T \left(ACA^T\right) \underline{\omega}\right)
\end{aligned}
$$

which is the joint CF of a Gaussian random vector with mean $A\underline{\mu} + \underline{b}$ and covariance ACA^T.