Stochastic Processes

Topic 2

Jointly Gaussian Random Variables

nctuee09f

Summary

This lecture reviews several important properties of Gaussian random variables. Specifically, I will discuss:

- Gaussian Random Variable
- Moment Generating Function
- Central Limit Theorem
- ***Jointly*** Gaussian Random Variables (***Multivariate Normal***)
- Joint Gaussian Density Function

Notation

We will use the following notation rules, unless otherwise noted, to represent symbols during this course.

- Boldface upper case letter to represent **MATRIX**
- Boldface lower case letter to represent **vector**
- Superscript $(\cdot)^T$ and $(\cdot)^H$ to denote transpose and hermitian (conjugate transpose), respectively
- Upper case italic letter to represent $RANDOM\ VARIABLE$

Motivation: Why a special attention to Gaussian RVs?

- They are analytically tractable

  $\rightarrow$ Preserved by linear systems

- Central limit theorem

  $\rightarrow$ Gaussian can approximate a large variety of distributions in large samples

- Useful as models of communication links

  — noise

  — channel fading effects

1 Gaussian Random Variables

Definition 1 *The probability density function (pdf) $f_X(x)$ for a Gaussian random variable X with mean μ and variance σ^2, denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$, is given by*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

for $-\infty < x < \infty$.

Remarks:

— Verify that X has mean μ and variance σ^2

— If $X \sim \mathcal{N}(\mu, \sigma^2)$, then the random variable $Z = \frac{X-\mu}{\sigma}$ has a $\mathcal{N}(0, 1)$ distribution, also known as the ***standard normal***

— In general, the random variable $Z_1 = aX + b$ for any real scalars a and b is also Gaussian with mean $a\mu + b$ and variance $a^2\sigma^2$

— A Gaussian random variable can be characterized by its first 2 moments; *i.e.* $E[X]$ and $E[X^2]$

2 Moment Generating Function

Definition 2 *The moment generating function (MGF) of a continuous random variable X with pdf $f_X(x)$ is defined by*

$$\theta(t) \triangleq E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx,$$

where t is a complex variable. For a discrete random variable X with probability mass function (pmf) $p_X(x) \triangleq P_X[X = x]$, the MGF is defined by

$$\theta(t) \triangleq E[e^{tX}] = \sum_i e^{tx_i} P_X[X = x_i].$$

Remarks:

(1) It's similar to the Laplace transform. Thus, in general, there is a one-to-one correspondence between $\theta(t)$ and $f_X(x)$.

(2) As the name suggests, MGF is commonly used for computing moments

The nth moment of X can be obtained by

$$E[X^n] = \frac{d^n}{dt^n}(\theta(t))\Big|_{t=0}.$$

(3) ***Solving problems involving sum of independent RV's***

(4) Analytical tool to demonstrate Central Limit Theorem

Example: Let $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$. Its MGF is given by

$$\theta(t) = \exp(\mu_X t + \sigma_X^2 t^2/2). \tag{2}$$

(1) We can compute $E[X] = \theta'(0) = \mu_X$ and $E[X^2] = \theta''(0) = \mu_X^2 + \sigma_X^2$.

(2) ***Sum of 2 indep. Gaussian***

Suppose that $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ is independent with X. Then, the MGF for $Z = X + Y$ is

$$E[e^{tZ}] = E[e^{tX}] \cdot E[e^{tY}] = \exp((\mu_X + \mu_Y)t + (\sigma_X^2 + \sigma_Y^2)t^2/2),$$

which bears the same form as (2). Therefore, from the 1-to-1 correspondence between distribution of a random variable and its MGF, we conclude that $Z \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

3 Central Limit Theorem

Theorem 3 (CLT in Special Case) *Let $X_1, X_2, ..., X_n$ be independent and identically distributed (i.i.d.) random variables with $E[X_i] = 0$ and $\mathrm{Var}[X_i] = 1$ for all i. Then,*

$$Z_n \triangleq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$$

tends to the standard normal in the sense that its MGF satisfies

$$\lim_{n\to\infty} \theta_{Z_n}(t) = e^{t^2/2},$$

which is the MGF of standard normal.

Remarks:

(1) For i.i.d. random variables $X_1 \cdots X_n$ with mean μ and variance σ^2, the CLT says that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right) = \sqrt{n} \left(\frac{\overline{X} - \mu}{\sigma} \right)$$

tends to standard normal where $\overline{X} = (1/n) \sum_{i=1}^{n} X_i$.

(2) It happens a lot that we'll be asked to find

$$P\left[a \le \sum_{i=1}^{n} X_i \le b \right]$$

with moderately large n for i.i.d. samples X_i. We don't really need to find the pdf for the sum $\sum_i X_i$. CLT tells us we can approximate the normalized sum to a standard normal.

(3) For a detailed proof of CLT, see [1]

4 Jointly Gaussian Random Variables

Definition 4 (Jointly Gaussian) *A collection of random variables $X_1, X_2, \cdots, X_n$ are **jointly Gaussian** if*

$$\sum_{i=1}^{n} a_i X_i$$

is a Gaussian random variable for real a_i for $i = 1 \cdots n$.

(1) In plain words, any linear combination of jointly Gaussian RV's is again ***a*** Gaussian random variable

(2) Let $X_1, X_2, \cdots, X_n$ be i.i.d. Gaussian. Then, they are jointly Gaussian. $\rightarrow$ Check the MGF of $\sum_{i=1}^{n} a_i X_i$.

(3) Let $\mathbf{x} = [X_1 \cdots X_n]^T$ be a vector of n i.i.d. $\mathcal{N}(0,1)$ random variables. Then,

$$\mathbf{Ax} + \mathbf{b}$$

is a vector of jointly Gaussian random variables for all real deterministic $\mathbf{A}$ and $\mathbf{b}$

(4) JG$\rightarrow$linear system$\rightarrow$JG

(5) Jointly Gaussian implies Marginal Gaussian. But the converse is not true

5 Basic Definitions

Definition 5 (Joint Density) *Let* $\mathbf{x} = [X_1 \cdots X_n]^T$ *be a random vector with probability distribution function* $F_{\mathbf{x}}(\boldsymbol{x})$*. Then, by definition*

$$F_{\mathbf{x}}(\boldsymbol{x}) \triangleq P\left[X_1 \leq x_1, \cdots, X_n \leq x_n\right] \triangleq P\left[\mathbf{x} \leq \mathsf{x}\right],$$

where $\boldsymbol{x} = [x_1, \cdots, x_n]^T$ *is a realization of the random vector* $\mathbf{x}$*. If the nth partial derivative of* $F_{\mathbf{x}}(\boldsymbol{x})$ *exists, the joint pdf is defined as*

$$f_{\mathbf{x}}(\boldsymbol{x}) \triangleq \frac{\partial^n F_{\mathbf{x}}(\boldsymbol{x})}{\partial x_1 \cdots \partial x_n}.$$

Definition 6 *The joint moment generating function for* N *random variables* $X_1 \cdots X_N$ *is defined by*

$$\theta(t_1, t_2, \cdots, t_N) = E\left[\exp\left(\sum_{i=1}^{N} t_i X_i\right)\right].$$

Remarks:

(1) Joint MGF uniquely determines the joint distribution of $X_1, \cdots, X_n$.

(2) $X_1, \cdots, X_n$ are independent if and only if $\theta(t_1, t_2, \cdots, t_n) = \theta_{X_1}(t_1) \cdots \theta_{X_n}(t_n)$.

Covariance Matrix

(a) The cross-covariance matrix $\mathbf{K}_{\mathbf{xy}}$ of two random vectors $\mathbf{x}$ and $\mathbf{y}$ consisting of random variables $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_n$, respectively, is defined by

$$\begin{aligned} \mathbf{K}_{\mathbf{xy}} &\triangleq \mathrm{Cov}\,(\mathbf{x}, \mathbf{y}) \\ &\triangleq E\left[(\mathbf{x} - E[\mathbf{x}])(\mathbf{y} - E[\mathbf{y}])^H\right], \end{aligned}$$

where $E[\mathbf{x}] = [E[X_1], \cdots, E[X_n]]^T$ is the vector of expected values of X_i, and likewise for $E[\mathbf{y}]$.

(b) The covariance matrix $\mathbf{K}_{\mathbf{xx}}$ of a random vector $\mathbf{x}$ is defined by $\mathbf{K}_{\mathbf{xx}} = E\left[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^H\right]$.

Remarks:

(1) For any matrices $\mathbf{A}$ and $\mathbf{B}$, and vectors $\mathbf{c}$ and $\mathbf{d}$ with proper dimension, one has $\mathrm{Cov}(\mathbf{Ax} + \mathbf{c}, \mathbf{By} + \mathbf{d}) = \mathbf{A} \cdot \mathrm{Cov}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{B}^T$

(2) The diagonal terms of $\mathbf{K}_{\mathbf{xx}}$, simply denoted by $\mathbf{K}_{\mathbf{x}}$, are the variances $\sigma_1^2, \cdots, \sigma_n^2$ of $X_1, X_2, \cdots, X_n$.

(3) $\mathbf{K}_{\mathbf{x}}$ is a Hermitian (or real symmetric) matrix. ($\mathbf{K}_{\mathbf{x}} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^H$)

Uncorrelated, orthogonal, and independent

Consider two real $n \times 1$ random vectors $\mathbf{x}$ and $\mathbf{y}$. Then, we say

(a) $\mathbf{x}$ and $\mathbf{y}$ are uncorrelated if $E[\mathbf{x}\mathbf{y}^T] = E[\mathbf{x}] \cdot E[\mathbf{y}^T]$.

(b) $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if $E[\mathbf{x}\mathbf{y}^T] = \mathbf{0}$.

(c) $\mathbf{x}$ and $\mathbf{y}$ are independent if $f_{\mathbf{x}\mathbf{y}}(\mathbf{x}, \mathbf{y}) = f_{\mathbf{x}}(\mathbf{x}) f_{\mathbf{y}}(\mathbf{y})$.

Remarks:

(1) Independence implies uncorrelatedness, but the converse is **NOT** true.

(2) Covariance matrix of uncorrelated random variables $X_1, \cdots, X_n$ is a diagonal matrix. (Off-diagonal terms are zero.)

Example:(Decorrelation of random vectors) (See 5.4-2 in textbook!)

6 Jointly Gaussian Density Function

We will determine in this section the ***joint density function*** for a jointly Gaussian random vector. Let's see some useful facts first.

(1) We denote $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$ to indicate that $\mathbf{x}$ is a jointly Gaussian ***random vector*** with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{K}$.

(2) Assume that $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$, then $\mathbf{Ax} + \mathbf{b} \sim \mathcal{N}(\mathbf{Am} + \mathbf{b}, \mathbf{AKA}^T)$. (We have already seen that any linear transformation of JG is also JG.)

(3) The joint MGF of a jointly Gaussian random vector $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$ is given by

$$\theta_{\mathbf{x}}(\mathbf{t}) = E[e^{\mathbf{t}^T\mathbf{x}}] = \exp\left(\mathbf{t}^T\mathbf{m} + \frac{1}{2}\mathbf{t}^T\mathbf{K}\mathbf{t}\right).$$

This can be verified by noting that $Y \triangleq \mathbf{t}^T\mathbf{x}$ is a Gaussian random variable from Definition 4. And, the mean and variance of Y are $\mathbf{t}^T\mathbf{m}$ and $\mathbf{t}^T\mathbf{K}\mathbf{t}$, respectively. Then, $\theta_{\mathbf{x}}(\mathbf{t}) = E[e^Y] = \theta_Y(t)\big|_{t=1}$.

(4) Jointly Gaussian random variables are independent if and only if they are uncorrelated.

($\Rightarrow$) Much easier to check $E[XY] = E[X]E[Y]$ if independence.

($\Leftarrow$) Use the concept of MGF.

Suppose that $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$ is a jointly Gaussian random vector of uncorrelated Gaussian random variables $X_1 \cdots X_n$, implying that $\mathbf{K}$ is a diagonal matrix $\mathbf{K} = \mathrm{diag}(\sigma_1^2 \cdots \sigma_n^2)$. Therefore, the joint MGF is

$$\begin{aligned}
\theta_{\mathbf{x}}(\mathbf{t}) &= \exp\left(\mathbf{t}^T\mathbf{m} + \frac{1}{2}\mathbf{t}^T\mathbf{K}\mathbf{t}\right) = \exp\left(\sum_{i=1}^{n}(t_i m_i + \frac{1}{2}t_i^2\sigma_i^2)\right) \\
&= \prod_{i=1}^{n}\exp\left(t_i m_i + \frac{1}{2}t_i^2\sigma_i^2\right) = \theta_{X_1}(t_1)\cdots\theta_{X_n}(t_n).
\end{aligned}$$

■

Important:
This is an important result, since we can determine whether jointly Gaussian random variables are ***independent*** by simply ***checking its correlation***. With independence, we can easily calculate, e.g., the joint pdf and the conditional expectation $E[X|Y] = E[X]$.

Example:
Let X and Y be jointly Gaussian random variables with zero mean, $\mathrm{Var}(X) = \sigma_X^2$ and $\mathrm{Var}(Y) = \sigma_Y^2$. We can find a scalar α such that $X - \alpha Y$ and Y are independent Gaussian random variables by letting

$$E\left[(X - \alpha Y)Y\right] = E[X - \alpha Y]E[Y] = 0.$$

From which, we have

$$\alpha = \frac{E[XY]}{E[Y^2]} = \rho\frac{\sigma_X}{\sigma_Y},$$

where $\rho \triangleq \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$ is the correlation coefficient between X and Y.

Joint Density of Two JG RVs

The joint pdf for two real jointly Gaussian random variables X and Y is given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot \exp\left(\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}\right),$$

where $\mu_X = E[X]$, $\sigma_X^2 = \mathrm{Var}[X]$, $\mu_Y = E[Y]$, $\sigma_Y^2 = \mathrm{Var}[Y]$, and

$$\rho = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y}$$

is the correlation coefficient.

Proof

Let's assume $\mu_X = \mu_Y = 0$ for simplicity. Since X and Y are jointly Gaussian, we know that

$$U = X - \alpha Y \quad \text{and} \quad V = Y$$

are also jointly Gaussian random variables. From the example in the last page, we know that U and V are independent if $\alpha = \rho\frac{\sigma_X}{\sigma_Y}$. From Section 3.4 of the textbook, the joint pdf $f_{X,Y}(x,y)$ can be determined from $f_{U,V}(u,v)$ by

$$f_{X,Y}(x,y) = \frac{1}{|\mathbf{J}|} f_{U,V}(u,v),$$

where $|\mathbf{J}| = \det(\mathbf{J})$ and the matrix $\mathbf{J}$ is given by

$$\mathrm{J} = \begin{bmatrix} \partial X/\partial U & \partial X/\partial V \\ \partial Y/\partial U & \partial Y/\partial V \end{bmatrix} = \begin{bmatrix} 1 & \rho\sigma_X/\sigma_Y \\ 0 & 1 \end{bmatrix}.$$

So, actually, we have

$$\begin{aligned} f_{X,Y}(x,y) &= f_{U,V}(x-\alpha y, y) = f_U(x-\alpha y) f_V(y) \\ &= \frac{1}{2\pi\sigma_U\sigma_V}\exp\left(-\frac{(x-\alpha y)^2}{2\sigma_U^2}\right)\exp\left(-\frac{y^2}{2\sigma_V^2}\right), \end{aligned}$$

where $\sigma_U^2 = E[U^2] = E[(X-\alpha Y)^2] = (1-\rho^2)\sigma_X^2$ and $\sigma_V = \sigma_Y$. Plugging the results and performing some manipulations, we can show that $f_{X,Y}(x,y)$ takes the form mentioned in the above. ■

Remarks:

(1) This joint pdf is commonly used to define two jointly Gaussian random variables. (See p. 201 in textbook.)

(2) If $\rho = 0$, we have $f_{X,Y}(x,y) = f_X(x) f_Y(y)$, showing that uncorrelatedness implies independence for jointly Gaussian random variables

(3) Recall that the joint MGF for JG only depends on the mean vector and covariance matrix. We can deduce that the joint pdf for JG is also the case.

Let $\mathbf{z} \triangleq [X\ Y]^T$. The mean vector $\mathbf{m}_z = [E[X]\ E[Y]]^T$, and covariance matrix $\mathbf{K}_z = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$ determines the joint pdf in the form

$$f_{\mathbf{z}}(\boldsymbol{z}) = \frac{1}{2\pi \det(\mathbf{K}_z)^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{z} - \mathbf{m}_z)^T \mathbf{K}_z^{-1} (\boldsymbol{z} - \mathbf{m}_z) \right).$$

(4) The ***contour*** and the surface of the joint pdf for two zero mean jointly Gaussian X_1 and X_2 with variance 2 and correlation coefficients $\rho = 0.5$ are plotted respectively in Fig. 1.

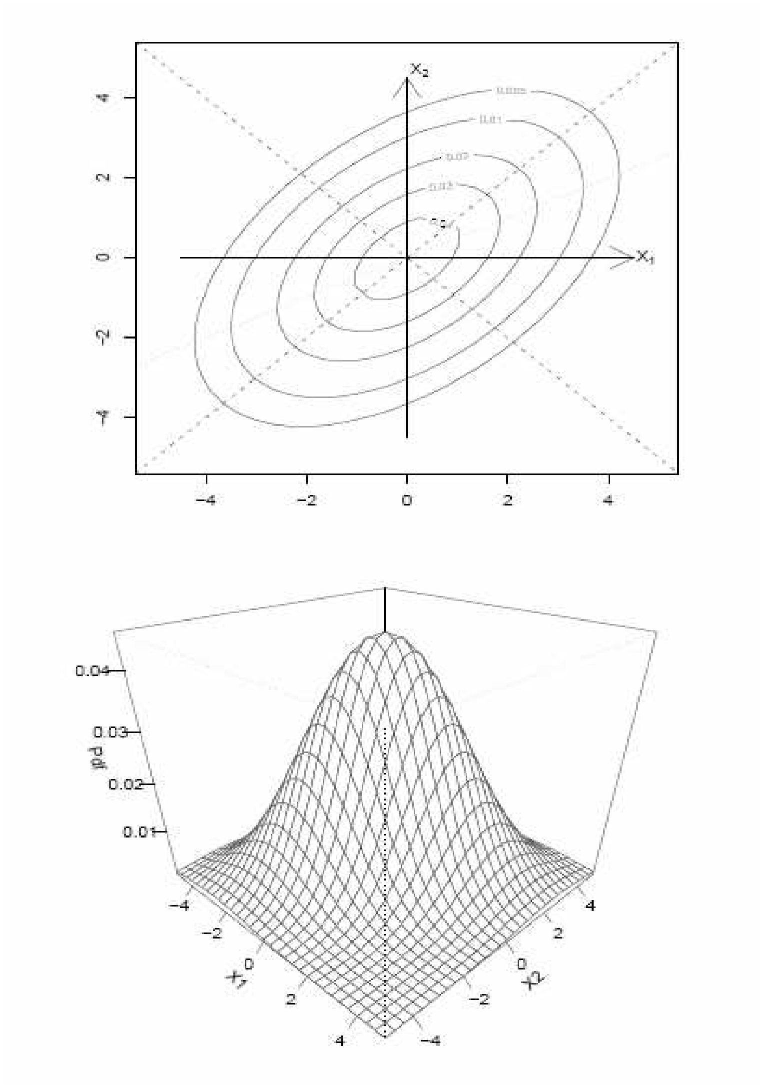

Figure 1: The contour and the surface of the pdf for jointly Gaussian X_1 and X_2 with variance 2 and $\rho = 0.5$.

Joint Density of n JG RVs

1. (Recall) Any jointly Gaussian random vector $\mathbf{x}$ can be represented by a linear combination of the vector of i.i.d. standard normal random variables $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$.

   That is, if $\mathbf{x} \sim \mathcal{N}(\mathbf{m_x}, \mathbf{K_x})$, we can write

   $$\mathbf{x} = \mathbf{K_x}^{1/2}\mathbf{z} + \mathbf{m_x},$$

   where $\mathbf{K_x}^{1/2} = \mathbf{E}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{E}^H$ with $\mathbf{E}$ being the matrix of orthonormal eigenvectors and $\mathbf{\Lambda}$ the diagonal matrix of eigenvalues of $\mathbf{K_x}$.

2. Let $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ and $\mathbf{U}$ be a unitary matrix. Then, $\mathbf{Uz}$ has an identical distribution as $\mathbf{z}$, denoted by

   $$\mathbf{z} \stackrel{d}{=} \mathbf{Uz}.$$

   (Justify)

   a. $\mathbf{Uz}$ is jointly Gaussian.

   b. Mean vector of $\mathbf{Uz}$ is a zero vector.

   c. $\text{Cov}(\mathbf{Uz}, \mathbf{Uz}) = \mathbf{U}\text{Cov}(\mathbf{z}, \mathbf{z})\mathbf{U}^T = \mathbf{U}\mathbf{U}^T = \mathbf{I}$

3. (General Expression) Let $\mathbf{x} = [X_1, X_2, \cdots, X_n]$ be a real jointly Gaussian random vector (Normal random vector) with mean vector $\mathbf{m}_{\mathbf{x}}$ and covariance matrix $\mathbf{K}_{\mathbf{x}}$. Then, the joint pdf $f_{\mathbf{x}}(\boldsymbol{x})$ is given by

$$f_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\det(\mathbf{K}_x)^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \mathbf{m}_x)^T \mathbf{K}_x^{-1}(\boldsymbol{x} - \mathbf{m}_x)\right).$$

Remarks:

(1) Please note the difference between random vector $\mathbf{x}$ and deterministic vector $\boldsymbol{x}$.

(2) Assume the elements $X_1, X_2, \cdots, X_n$ of the random vector $\mathbf{x}$ are uncorrelated, each with variances $\mathrm{Var}(X_i) = \sigma_i^2$, then the joint pdf is reduced to

$$f_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\prod_{i=1}^{n}\sigma_i} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu_i}{\sigma_i}\right)^2\right).$$

Proof Recall that any jointly Gaussian random vector can be represented by a linear combination of a standard Gaussian random vector, we can write

$$\begin{aligned}\mathbf{x} &= \mathbf{K_x}^{\frac{1}{2}}\mathbf{z} + \mathbf{m_x} \\ &= \mathbf{E}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{E}^T\mathbf{z} + \mathbf{m_x}\end{aligned}$$

where $\mathbf{K_x}^{\frac{1}{2}} \triangleq \mathbf{E}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{E}^T$. Since $\mathbf{E}$ is a unitary matrix, $\mathbf{E}^T\mathbf{z}$ is also a standard Gaussian random vector. It follows that $\mathbf{x}$ has an identical distribution with $\mathbf{E}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{z} + \mathbf{m_x}$.

Let $\mathbf{y} = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{z}$. It is clear that $\mathbf{y}$ is also jointly Gaussian distributed with $N(0, \Lambda)$, and also a vector of independent Gaussian RVs. Then, the joint pdf for $\mathbf{y}$ is given by

$$\begin{aligned}f_{\mathbf{y}}(\boldsymbol{y}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\lambda_i^{1/2}} \exp\left(-\frac{y_i^2}{2\lambda_i}\right) \\ &= \frac{1}{(2\pi)^{n/2}\det(\mathbf{\Lambda})^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{y}^T\mathbf{\Lambda}^{-1}\boldsymbol{y}\right).\end{aligned}$$

Once we have the joint pdf of $\mathbf{y}$, we can use the concept of linear transformation to determine the joint pdf of $\mathbf{x}$ from

$$\mathbf{x} = \mathbf{E}\mathbf{y} + \mathbf{m_x}.$$

That is,

$$\begin{aligned}f_{\mathbf{x}}(\boldsymbol{x}) &= \frac{1}{|\mathbf{J}|} f_{\mathbf{y}}\left(\mathbf{E}^T(\boldsymbol{x} - \mathbf{m_x})\right) \\ &= f_{\mathbf{y}}\left(\mathbf{E}^T(\boldsymbol{x} - \mathbf{m_x})\right) \quad \text{since} \quad |\mathbf{J}| = 1 \\ &= \frac{1}{(2\pi)^{n/2}\det(\mathbf{\Lambda})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \mathbf{m_x})^T\mathbf{E}\mathbf{\Lambda}^{-1}\mathbf{E}^T(\boldsymbol{x} - \mathbf{m_x})\right) \\ &= \frac{1}{(2\pi)^{n/2}\det(\mathbf{K_x})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \mathbf{m_x})^T\mathbf{K}_{\mathbf{x}}^{-1}(\boldsymbol{x} - \mathbf{m_x})\right).\end{aligned}$$

■

4. To conclude, the following 3 statements are equivalent:
    - Random variables $X_1, X_2, \cdots, X_n$ are jointly Gaussian.
    - The random variable $Y = \sum_{i=1}^{n} a_i X_i$ is a Gaussian random variable for any real a_i.
    - The joint pdf for $X_1, X_2, \cdots, X_n$ is given by

$$f_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\det(\mathbf{K}_{\mathbf{x}})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \mathbf{m}_{\mathbf{x}})^T \mathbf{K}_{\mathbf{x}}^{-1} (\boldsymbol{x} - \mathbf{m}_{\mathbf{x}})\right).$$

References

[1] W. Feller, *An Introduction to Probability Theory and Its Applications*, New York: John Wily, 2nd edition, 1971.