CST 3526 Stochastic Process Lecture 13 - 12/11/2025

Lecture 13: Gaussian Processes I

Lecturer: Ziqiao Wang

In this lecture, we begin our study of Gaussian processes. In fact, a solid understanding of multiple (jointly) Gaussian random variables is the key to understanding Gaussian processes. Therefore, we will mainly focus on how to characterize joint Gaussian random variables. Specifically, we will see how a Gaussian distribution is completely described by its first two moments, and how the characteristic function can be used to reveal and prove several important properties of Gaussian random variables.

1 Jointly Gaussian Random Variables

Recall that the PDF for the Gaussian random variable X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \qquad -\infty < x < \infty.$$

The CDF of the Gaussian random variable is given by

$$P[X \le x] = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{(x'-m)^2}{2\sigma^2}}\, dx'.$$

Theorem 1.1 (Central Limit Theorem (CLT))**.** *Suppose $X_1, X_2, X_3, \ldots$ is a sequence of i.i.d. random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$. Let $\bar{X}_n = \sum_{i=1}^{n} X_i$. Then, as $n \to \infty$, the random variables*

$$\sqrt{n}\,(\bar{X}_n - \mu)$$

converge in distribution to a normal random variable $\mathcal{N}(0, \sigma^2)$:

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

Remark 1.1. *The amazing part about the central limit theorem is that the summands X_i can have* **any distribution** *as long as they have a finite mean and finite variance. This gives the result its wide applicability.*

The random variables $X_1, X_2, \ldots, X_n$ are said to be *jointly Gaussian* if their joint PDF is given by

$$f_{\mathbf{X}}(\mathbf{x}) \triangleq f_{X_1, X_2, \ldots, X_n}(x_1, \ldots, x_n) = \frac{\exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^{\mathrm{T}}\mathbf{K}^{-1}(\mathbf{x}-\mathbf{m})\right\}}{(2\pi)^{n/2}|\mathbf{K}|^{1/2}}, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{m}$ are column vectors defined by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_n] \end{bmatrix},$$

and $\mathbf{K}$ is the covariance matrix defined by

$$\mathbf{K} = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_n) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_n, X_1) & \mathrm{Cov}(X_n, X_2) & \cdots & \mathrm{Var}(X_n) \end{bmatrix}. \tag{2}$$

The $(\cdot)^{\mathrm{T}}$ in Eq. (2) denotes the transpose of a matrix or vector. Note that the covariance matrix is symmetric since $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$.

Eq. (1) shows that **the PDF of jointly Gaussian random variables is completely specified by the individual means and variances and the pairwise covariances**. It can be shown using the joint characteristic function that all the marginal PDF's associated with Eq. (1) are also Gaussian and that these too are completely specified by the same set of means, variances and covariances.

Remark 1.2. *In general, the independence and uncorrelatedness of two random variables are different concepts. When second-order moments exist, independence implies uncorrelatedness, but the converse is not necessarily true. A similar conclusion holds for random vectors. Since the statistical properties of jointly Gaussian random vectors are completely determined by their second-order moments, they exhibit a special relationship between independence and uncorrelatedness.*

Example 1. Derive the PDF of two-dimensional Gaussian (X, Y). The covariance matrix for the two-dimensional case is given by

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 & \rho_{X,Y}\sigma_1\sigma_2 \\ \rho_{X,Y}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

where we have used the fact that $\mathrm{Cov}(X, Y) = \rho_{X,Y}\sigma_1\sigma_2$. The determinant of $\mathbf{K}$ is $\sigma_1^2\sigma_2^2(1 - \rho_{X,Y}^2)$. The inverse of the covariance matrix is also a real symmetric matrix:

$$\mathbf{K}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{X,Y}^2)} \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y}\sigma_1\sigma_2 \\ -\rho_{X,Y}\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}.$$

The term in the exponent is therefore

$$\begin{aligned}
&\frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{X,Y}^2)} [x - m_1,\ y - m_2] \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y}\sigma_1\sigma_2 \\ -\rho_{X,Y}\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x - m_1 \\ y - m_2 \end{bmatrix} \\
=&\frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{X,Y}^2)} [x - m_1,\ y - m_2] \begin{bmatrix} \sigma_2^2(x-m_1) - \rho_{X,Y}\sigma_1\sigma_2(y-m_2) \\ -\rho_{X,Y}\sigma_1\sigma_2(x-m_1) + \sigma_1^2(y-m_2) \end{bmatrix} \\
=&\frac{\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2}{1-\rho_{X,Y}^2}.
\end{aligned}$$

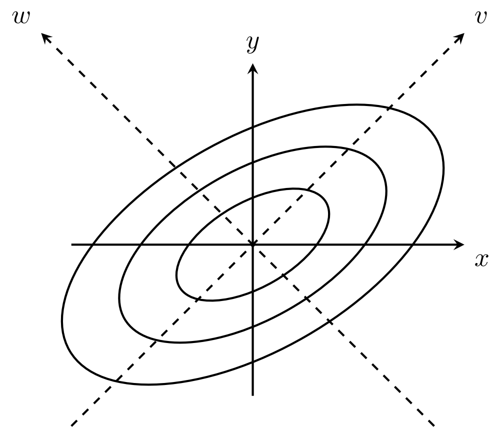
Thus, the random variables X and Y are said to be *jointly Gaussian* if their joint PDF has the form

$$f_{X,Y}(x,y) = \frac{\exp\left\{-\frac{1}{2(1-\rho_{X,Y}^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{X,Y}^2}}, \tag{3}$$

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

The PDF is centered at the point (m_1, m_2), and it has a bell shape that depends on the values of σ_1, σ_2, and $\rho_{X,Y}$. As shown in the figure below, the PDF is constant for values x and y for which the argument of the exponent is constant:

$$\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right] = \text{constant}. \tag{4}$$

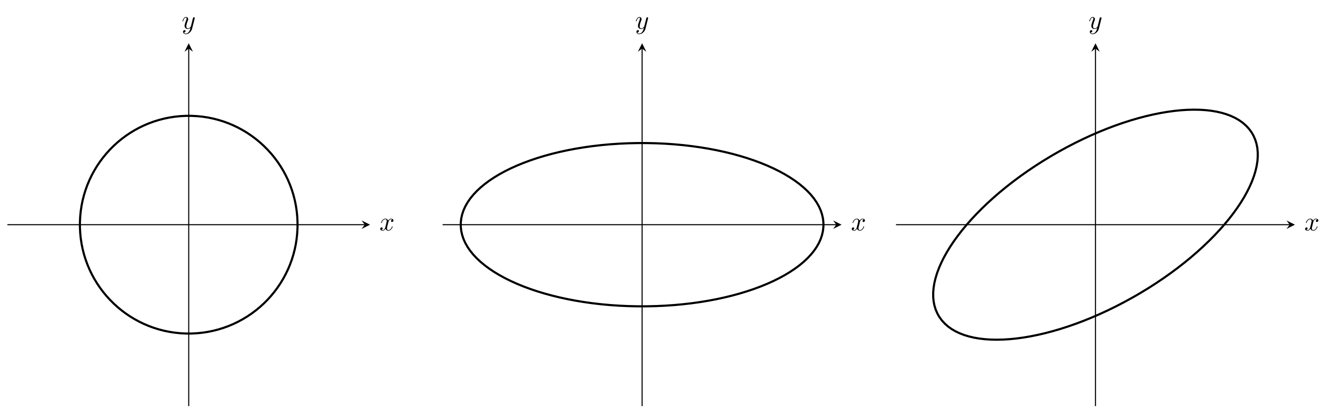


Regarding the orientation of the elliptical contours for various values of σ_1, σ_2, and $\rho_{X,Y}$. When $\rho_{X,Y} = 0$, that is, when X and Y are independent, the equal-pdf contour is an ellipse with principal axes aligned with the x- and y-axes. When $\rho_{X,Y} \neq 0$, the major axis of the ellipse is oriented along the angle

$$\theta = \frac{1}{2} \arctan\left(\frac{2\rho_{X,Y}\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \right). \tag{5}$$

Note that the angle is 45° when the variances are equal.

Circular (isotropic) | Axis-aligned Ellipse | Rotated ellipse



i) $\sigma_x = \sigma_y,\ \rho = 0$ | ii) $\sigma_x \neq \sigma_y,\ \rho = 0$ | iii) $\sigma_x \neq \sigma_y,\ \rho \neq 0$

Example 2 (Rotation of Jointly Gaussian Random Variables)**.** The ellipse corresponding to an arbitrary two-dimensional Gaussian vector forms an angle

$$\theta = \frac{1}{2} \arctan\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \right)$$

relative to the x-axis. Suppose we define a new coordinate system by using the following rotation matrix:

$$\begin{bmatrix} V \\ W \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}.$$

To show that the new random variables are independent, it suffices to show that they have covariance zero:

$$
\begin{aligned}
\operatorname{Cov}(V,W) =& \mathbb{E}[(V-\mathbb{E}[V])(W-\mathbb{E}[W])] \\
=& \mathbb{E}\Big[\big((X-m_1)\cos\theta+(Y-m_2)\sin\theta\big)\big(-(X-m_1)\sin\theta+(Y-m_2)\cos\theta\big)\Big] \\
=& -\sigma_1^2\sin\theta\cos\theta+\operatorname{Cov}(X,Y)\cos^2\theta-\operatorname{Cov}(X,Y)\sin^2\theta+\sigma_2^2\sin\theta\cos\theta \\
=& \frac{(\sigma_2^2-\sigma_1^2)\sin 2\theta+2\operatorname{Cov}(X,Y)\cos 2\theta}{2} \\
=& \frac{\cos 2\theta\left[(\sigma_2^2-\sigma_1^2)\tan 2\theta+2\operatorname{Cov}(X,Y)\right]}{2}.
\end{aligned}
$$

If we let the angle of rotation θ be such that

$$\tan 2\theta = \frac{2\operatorname{Cov}(X,Y)}{\sigma_1^2-\sigma_2^2},$$

then the covariance of V and W is zero as required.

Theorem 1.2. *Multiple jointly Gaussian real random variables are independent if and only if the covariance matrix of the random variables is diagonal.*

Proof. Suppose $X_1, X_2, \ldots, X_n$ are jointly Gaussian random variables with

$$\operatorname{Cov}(X_i, X_j) = 0 \quad \text{for } i \neq j.$$

Show that $X_1, X_2, \ldots, X_n$ are independent random variables.

From Eq. (2), we see that the covariance matrix is a diagonal matrix:

$$\mathbf{K} = \operatorname{diag}[\operatorname{Var}(X_i)] = \operatorname{diag}[\sigma_i^2].$$

Therefore, $\mathbf{K}^{-1} = \operatorname{diag}\left[\frac{1}{\sigma_i^2}\right]$, and $(\mathbf{x}-\mathbf{m})^T\mathbf{K}^{-1}(\mathbf{x}-\mathbf{m}) = \sum_{i=1}^n \left(\frac{x_i-m_i}{\sigma_i}\right)^2$.

From Eq. (1), we have

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\exp\left\{-\frac{1}{2}\sum_{i=1}^n \left[(x_i-m_i)/\sigma_i\right]^2\right\}}{(2\pi)^{n/2}\,|\mathbf{K}|^{1/2}} = \prod_{i=1}^n \frac{\exp\left\{-\frac{1}{2}\left[(x_i-m_i)/\sigma_i\right]^2\right\}}{\sqrt{2\pi\sigma_i^2}} = \prod_{i=1}^n f_{X_i}(x_i).$$

Thus, $X_1, X_2, \ldots, X_n$ are independent Gaussian random variables. □

Example 3 (Decorrelation)**.** Let $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ be a jointly Gaussian random vector. Define the linear transformation $\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} I & A \\ 0 & I \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$.

We want Y_1 and Y_2 to be uncorrelated, i.e.,

$$\mathbb{E}[(Y_1-\mathbb{E}Y_1)(Y_2-\mathbb{E}Y_2)^T] = 0.$$

This leads to

$$0 = \Sigma_{12} + A\Sigma_{22} \quad \Rightarrow \quad A = -\Sigma_{12}\Sigma_{22}^{-1}.$$

Then the covariance matrix of $\mathbf{Y}$ becomes

$$\mathbb{E}[(\mathbf{Y}-\mathbb{E}\mathbf{Y})(\mathbf{Y}-\mathbb{E}\mathbf{Y})^T] = \begin{pmatrix} \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}.$$

2 Characteristic Function

The **characteristic function** of a random variable X is defined by

$$\Phi_X(\omega) = \mathbb{E}\big[e^{j\omega X}\big] = \int_{-\infty}^{\infty} f_X(x)\, e^{j\omega x}\, dx, \tag{6}$$

where $j = \sqrt{-1}$ is the imaginary unit number. The two expressions on the RHS motivate two interpretations of the characteristic function. In the first expression, $\Phi_X(\omega)$ can be viewed as the expected value of a function of X, $e^{j\omega X}$, in which the parameter ω is left unspecified. In the second expression, $\Phi_X(\omega)$ is simply the Fourier transform of the PDF $f_X(x)$ (with a reversal in the sign of the exponent). Both of these interpretations prove useful in different contexts.

If we view $\Phi_X(\omega)$ as a Fourier transform, then we have from the Fourier transform inversion formula that the PDF of X is given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(\omega)\, e^{-j\omega x}\, d\omega. \tag{7}$$

It then follows that **every PDF and its characteristic function form a unique Fourier transform pair**.

Since $f_X(x)$ and $\Phi_X(\omega)$ form a transform pair, we would expect to be able to obtain the moments of X from $\Phi_X(\omega)$. The **moment theorem** states that the moments of X are given by

$$\mathbb{E}[X^n] = \frac{1}{j^n} \frac{d^n}{d\omega^n} \Phi_X(\omega)\bigg|_{\omega=0}. \tag{8}$$

To show this, first expand $e^{j\omega x}$ in a power series in the definition of $\Phi_X(\omega)$:

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} f_X(x) \left\{1 + j\omega x + \frac{(j\omega x)^2}{2!} + \cdots \right\} dx. \tag{9}$$

Assuming that all the moments of X are finite and that the series can be integrated term by term, we obtain

$$\Phi_X(\omega) = 1 + j\omega\mathbb{E}[X] + \frac{(j\omega)^2\mathbb{E}[X^2]}{2!} + \cdots + \frac{(j\omega)^n\mathbb{E}[X^n]}{n!} + \cdots. \tag{10}$$

If we differentiate the above expression once and evaluate the result at $\omega = 0$ we obtain

$$\frac{d}{d\omega}\Phi_X(\omega)\bigg|_{\omega=0} = j\,\mathbb{E}[X]. \tag{11}$$

If we differentiate n times and evaluate at $\omega = 0$, we finally obtain

$$\frac{d^n}{d\omega^n}\Phi_X(\omega)\bigg|_{\omega=0} = j^n\mathbb{E}[X^n], \tag{12}$$

which yields Eq. (8).

Characteristic Function of Gaussian Random Variables Let $X \sim \mathcal{N}(m, \sigma^2)$. By definition, the characteristic function is

$$\Phi_X(\omega) = \mathbb{E}\big[e^{j\omega X}\big] = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)\, dx,$$

where

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

Substituting and letting $y = x - m$, we obtain

$$\Phi_X(\omega) = e^{j\omega m} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2} + j\omega y\right) dy.$$

Completing the square,

$$-\frac{y^2}{2\sigma^2} + j\omega y = -\frac{(y - j\omega\sigma^2)^2}{2\sigma^2} - \frac{1}{2}\sigma^2\omega^2.$$

Thus,

$$\Phi_X(\omega) = e^{j\omega m} e^{-\frac{1}{2}\sigma^2\omega^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - j\omega\sigma^2)^2}{2\sigma^2}\right) dy.$$

The integral is the integral of a (complex-shifted) Gaussian density. Its value is 1: it is the same as the standard Gaussian integral, just translated by a constant (here a complex shift, but the integral over $\mathbb{R}$ is unchanged). We obtain

$$\boxed{\Phi_X(\omega) = \exp\left(jm\omega - \frac{1}{2}\sigma^2\omega^2\right).}$$

The joint characteristic function is very useful in developing the properties of jointly Gaussian random variables. We now show that **the joint characteristic function of n jointly Gaussian random variables $X_1, X_2, \ldots, X_n$ is given by**

$$\Phi_{X_1,X_2,\ldots,X_n}(\omega_1, \omega_2, \ldots, \omega_n) = \exp\left(j\sum_{i=1}^{n} \omega_i m_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} \omega_i\omega_k \operatorname{Cov}(X_i, X_k)\right), \tag{13}$$

which can be written more compactly as follows:

$$\Phi_{\mathbf{X}}(\boldsymbol{\omega}) \triangleq \Phi_{X_1,X_2,\ldots,X_n}(\omega_1, \omega_2, \ldots, \omega_n) = \exp\left(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{m} - \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\mathbf{K}\boldsymbol{\omega}\right), \tag{14}$$

where $\mathbf{m}$ is the vector of means and $\mathbf{K}$ is the covariance matrix defined in Eq. (2).

3 Properties of Multiple Gaussian Random Variables

Multivariate Gaussian distributions enjoy many nice properties that other distributions do not possess. Understanding these properties is very important for the study of Gaussian processes.

3.1 Marginal Distributions

Theorem 3.1. *If $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ follows an n-dimensional Gaussian distribution, then any subvector*

$$\tilde{\mathbf{X}} = (X_{k_1}, X_{k_2}, \ldots, X_{k_m})^T, \quad m < n$$

also follows an m-dimensional Gaussian distribution.

In other words, **the marginal distributions of a Gaussian distribution are still Gaussian**. This property can be verified using the characteristic function.

Proof. In fact, the characteristic function of $(X_{k_1}, X_{k_2}, \ldots, X_{k_m})$ satisfies

$$\Phi_{\tilde{\mathbf{X}}}(\boldsymbol{\omega}_{(k)}) = \Phi_{\tilde{\mathbf{X}}}(\omega_{k_1}, \omega_{k_2}, \ldots, \omega_{k_m}) = \mathbb{E}\exp\left(j(\omega_{k_1}X_{k_1} + \cdots + \omega_{k_m}X_{k_m})\right) = \Phi_{\mathbf{X}}(\tilde{\boldsymbol{\omega}}),$$

where $\Phi_{\mathbf{X}}(\boldsymbol{\omega})$ is the characteristic function of $\mathbf{X}$,

$$\boldsymbol{\omega}_{(k)} = (\omega_{k_1}, \omega_{k_2}, \ldots, \omega_{k_m})^T,$$

and

$$\tilde{\boldsymbol{\omega}} = (\dots, 0, \omega_{k_1}, 0, \dots, 0, \omega_{k_2}, 0, \dots, 0, \omega_{k_m}, 0, \dots)^T.$$

Note that

$$\tilde{\boldsymbol{\omega}}^T = \boldsymbol{\omega}_{(k)}^T A,$$

where $A \in \mathbb{R}^{m \times n}$ with elements

$$A_{pq} = \begin{cases} 1, & p = i,\ q = k_i, \\ 0, & p = i,\ q \neq k_i, \\ 0, & p \neq i. \end{cases}$$

Since $\mathbf{X}$ follows a multivariate Gaussian distribution, its characteristic function is

$$\Phi_{\mathbf{X}}(\boldsymbol{\omega}) = \exp\bigl(j\boldsymbol{\omega}^T\boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\Sigma}_{\mathbf{X}}\boldsymbol{\omega}\bigr).$$

Therefore,

$$\Phi_{\tilde{\mathbf{X}}}(\boldsymbol{\omega}_{(k)}) = \Phi_{\mathbf{X}}(\tilde{\boldsymbol{\omega}}) = \Phi_{\mathbf{X}}(\boldsymbol{\omega}_{(k)}A) = \exp\bigl(j\boldsymbol{\omega}_{(k)}^T A\boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{\omega}_{(k)}^T A\boldsymbol{\Sigma}_{\mathbf{X}}A^T\boldsymbol{\omega}_{(k)}\bigr).$$

Thus, $\tilde{\mathbf{X}}$ follows a Gaussian distribution with mean $A\boldsymbol{\mu}$ and covariance matrix $A\boldsymbol{\Sigma}_{\mathbf{X}}A^T$. □

Remark 3.1. *It should be pointed out that the converse is not true: even if each component of a random vector follows a univariate Gaussian distribution, this does not guarantee that the vector follows a joint multivariate Gaussian distribution. That is, Gaussian marginals do not imply a joint Gaussian distribution.*