

Lecture #5: Gaussian Vectors

Definition - A random vector $\vec{X} = (X_1, \dots, X_n)$ is Gaussian if and only if any linear combination of $X_1, \dots, X_n$ is a Gaussian random variable. That is,

$$Z = a_1 X_1 + \dots + a_n X_n$$

is Gaussian for all $a_1, \dots, a_n \in \mathbb{R}$.

Characterization of Gaussian Random Variables: If X is Gaussian then:

1. $X \sim \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $E[X] = \mu$, $E[(X-\mu)^2] = \sigma^2$

2. $\phi(t) = E[e^{tX}] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$

Is there a similar way to characterize Gaussian random variables?

Notation - For a random vector $\vec{X} = (X_1, \dots, X_n)$ we define

1. $\vec{\mu} = (E[X_1], \dots, E[X_n]) = (\mu_1, \dots, \mu_n)$

2. The covariance matrix is defined by

$$C_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$$

Consequences - If $\vec{X} = (X_1, \dots, X_n)$ is a Gaussian random vector then

1. Each X_i is a Gaussian random variable (set $a_i = 1$, $a_j = 0$ for $i \neq j$).

2. If $Z = a_1 X_1 + \dots + a_n X_n$ then

$$E[Z] = E[a_1 X_1 + \dots + a_n X_n]$$

$$= a_1 E[X_1] + \dots + a_n E[X_n]$$

$$= a_1 \mu_1 + \dots + a_n \mu_n$$

$$= \vec{a} \cdot \vec{\mu} \text{ (Dot Product)}$$

3. If $Z = a_1 X_1 + \dots + a_n X_n$ then

$$\begin{aligned} \text{Var}(Z) &= E[(a_1 X_1 + \dots + a_n X_n - a_1 \mu_1 - \dots - a_n \mu_n)^2] \\ &= E[(a_1 (X_1 - \mu_1) + \dots + a_n (X_n - \mu_n))^2] \\ &= E[(\sum_{i=1}^n a_i (X_i - \mu_i))^2] \\ &= E[\sum_{i=1}^n \sum_{j=1}^n a_i a_j (X_i - \mu_i)(X_j - \mu_j)] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j E[(X_i - \mu_i)(X_j - \mu_j)] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j C_{ij} = \vec{a}^T C \vec{a} \end{aligned}$$

$$\Rightarrow \text{Var}(Z) = \vec{a}^T C \vec{a} = \vec{a} \cdot C \vec{a}$$

* The mean vector $\vec{\mu}$ and covariance matrix C characterize the Gaussian random variable.

Example:

Consider the Gaussian random vector defined by

$$X_1 = Z_1 + Z_2$$

$$X_2 = Z_1 - Z_2$$

where Z_1, Z_2 are IID standard Gaussian random variables. In this case,

$$\vec{X} = A \vec{Z} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix},$$

$$\Rightarrow \vec{Z} = A^{-1} \vec{X} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Rightarrow \begin{aligned} Z_1 &= \frac{1}{2} X_1 + \frac{1}{2} X_2 \\ Z_2 &= \frac{1}{2} X_1 - \frac{1}{2} X_2 \end{aligned}$$

We also have that

$$E[X_1] = E[Z_1] + E[Z_2] = 0$$

$$E[X_2] = E[Z_1] - E[Z_2] = 0$$

$$\begin{aligned} E[X_1^2] &= E[Z_1^2 + 2Z_1 Z_2 + Z_2^2] = E[Z_1^2] + 2E[Z_1 Z_2] + E[Z_2^2] \\ &= 1 + 2E[Z_1]E[Z_2] + 1 = 2 \end{aligned}$$

$$E[X_1 X_2] = E[(Z_1 + Z_2)(Z_1 - Z_2)] = E[Z_1^2] - E[Z_2^2] = 0$$

$$E[X_2^2] = E[(Z_1 - Z_2)^2] = E[Z_1^2 - 2Z_1 Z_2 + Z_2^2] = 2$$

Therefore,

$$C = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \text{ and } C^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

To find the joint density of $\vec{X}$ we compute

$$\begin{aligned} P(\vec{X} \in B) &= P(A\vec{Z} \in B) = P(\vec{Z} \in A^{-1}(B)) \\ &= \iint_{A^{-1}(B)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_2^2}{2}} dz_1 dz_2 \end{aligned}$$

In the x_1, x_2 variables we have

$$dz_1 = \frac{1}{2} dx_1 + \frac{1}{2} dx_2$$

$$dz_2 = \frac{1}{2} dx_1 - \frac{1}{2} dx_2$$

$$\Rightarrow |dz_1 dz_2| = \left(\frac{1}{2} dx_1 + \frac{1}{2} dx_2 \right) \left(\frac{1}{2} dx_1 - \frac{1}{2} dx_2 \right) = \frac{1}{4} dx_1^2 - \frac{1}{4} dx_1 dx_2 + \frac{1}{4} dx_2 dx_1 + \frac{1}{4} dx_2^2$$

$$\Rightarrow |dz_1 dz_2| = \left| -\frac{1}{4} dx_1 dx_2 + \frac{1}{4} dx_2 dx_1 \right| = \left| -\frac{1}{4} dx_1 dx_2 - \frac{1}{4} dx_1 dx_2 \right| = \frac{1}{2} dx_1 dx_2$$

Therefore, changing variables we have:

$$\begin{aligned} P(\vec{X} \in B) &= \iint_B \frac{1}{2\pi} e^{-\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right)^2/2} e^{-\left(\frac{1}{2}x_1 - \frac{1}{2}x_2\right)^2/2} \frac{1}{2} dx_1 dx_2 \\ &= \iint_B \frac{1}{2\pi \cdot 2} \exp\left(-\frac{x_1^2}{4} - \frac{x_2^2}{4}\right) dx_1 dx_2 \end{aligned}$$

The joint density is therefore,

$$f(x_1, x_2) = \frac{1}{2\pi \cdot 2} \exp\left(-\frac{x_1^2}{4} - \frac{x_2^2}{4}\right) = \frac{1}{2\pi (\det(C))^{1/2}} \exp\left(-\frac{1}{2} \vec{x}^T C^{-1} \vec{x}\right)$$

Note: From the matrix A we can generate the random variables X_1, X_2 by matrix multiplication with Z_1, Z_2 .

Note: The matrix A is the square root of C in the sense

$$\text{that } C = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^T = AA^T$$

For this reason, A is often denoted Σ .

Fact: If $\vec{X} = (X_1, \dots, X_n)$ is a Gaussian random vector with mean $\vec{\mu}$ and invertible covariance matrix C then the joint density is given by

$$f(\vec{X}) = \frac{1}{(2\pi)^{n/2} (\det(C))^{1/2}} \exp\left(-\frac{(\vec{X}-\vec{\mu})^T C^{-1} (\vec{X}-\vec{\mu})}{2}\right)$$

Fact: If $\vec{X} = (X_1, \dots, X_n)$ is a Gaussian random vector and $i \neq j$, then X_i and X_j are independent if and only if $C_{ij} = \text{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = 0$.

Linear Algebra Emerges:

If we define a scalar random variable, i.e., $X = a$, to be a Gaussian with $\mu = \mathbb{E}[X] = a$ and $\text{Var}(X) = 0$ then the set of all Gaussian random variables with mean 0 is a vector space V_g .

Recall, an inner product on V_g , denoted $\langle \cdot, \cdot \rangle$, must satisfy

1. For all $X \in V_g$, $\langle X, X \rangle \geq 0$ and $\langle X, X \rangle = 0$ if and only if $X = 0$.
2. For all $X, Y \in V_g$, $\langle X, Y \rangle = \langle Y, X \rangle$
3. For all $a \in \mathbb{R}$, $X, Y, Z \in V_g$, $\langle aX + Y, Z \rangle = a\langle X, Z \rangle + \langle Y, Z \rangle$

The expected value defines an inner product on V_g :

$$\langle X, Y \rangle = \mathbb{E}[XY]$$

*Note $\langle X, Y \rangle = 0$, i.e., X and Y are orthogonal, if and only if X, Y are independent.

Example;

Given $C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, how do we find A such that:

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = A \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$$

where Z_1, Z_2 are IID standard Gaussian random variables.

1. $\text{Var}(X_1) = 2$, so define $Z_1 = \frac{1}{\sqrt{2}} X_1$.

2. Set $Z'_2 = X_2 - \langle X_2, Z_1 \rangle Z_1 = X_2 - \mathbb{E}[X_2, Z_1] Z_1$,
projection onto Z_1

Note,

$$\begin{aligned} \langle Z'_2, Z_1 \rangle &= \langle X_2 - \langle X_2, Z_1 \rangle Z_1, Z_1 \rangle \\ &= \langle X_2, Z_1 \rangle - \langle X_2, Z_1 \rangle \langle Z_1, Z_1 \rangle \\ &= 0 \end{aligned}$$

Therefore, Z'_2 and Z_1 are independent. Now,

$$\langle X_1, Z'_2 \rangle = \langle X_1, \frac{1}{\sqrt{2}} X_2 \rangle = \frac{1}{\sqrt{2}} \langle X_1, X_2 \rangle = \frac{1}{\sqrt{2}}$$

Therefore,

$$Z'_2 = X_2 - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} X_1 = X_2 - \frac{1}{2} X_1$$

$$\begin{aligned} \Rightarrow \langle Z'_2, Z'_2 \rangle &= \langle X_2 - \frac{1}{2} X_1, X_2 - \frac{1}{2} X_1 \rangle \\ &= \langle X_2, X_2 \rangle - \langle X_2, X_1 \rangle + \frac{1}{4} \langle X_1, X_1 \rangle \\ &= 2 - 1 + \frac{1}{2} \\ &= \frac{3}{2} \end{aligned}$$

Consequently, we define

$$Z_2 = \sqrt{\frac{2}{3}} Z'_2 = \sqrt{\frac{2}{3}} (X_2 - \frac{1}{2} X_1) = \sqrt{\frac{2}{3}} X_2 - \frac{1}{\sqrt{6}} X_1$$

Therefore,

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = A^{-1} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\text{Consequently, } A = \sqrt{3} \begin{bmatrix} -\sqrt{\frac{2}{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \end{bmatrix}$$

From 1 and 2 we have:

$$X_1 = \sqrt{2} Z_1$$

$$X_2 = \frac{1}{\sqrt{2}} Z_1 + \sqrt{\frac{3}{2}} Z_2$$

Notice

$$A \cdot A^T = \begin{bmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{\frac{3}{2}} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & \frac{1}{2} + \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Cholesky
Decomposition

Example:

What if $C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$?

1. $\text{Var}(X_1) = E[X_1^2] = C_{11} = 1$. Therefore,

$$Z_1 = X_1$$

2. $Z'_2 = X_2 - E[X_2 Z_1] Z_1 = X_2 - E[X_2 X_1] X_1 = X_2 - X_1$

$$\text{Var}(Z'_2) = E[(X_2 - X_1)(X_2 - X_1)] = C_{22} - 2C_{12} + C_{11} = 2 - 2 + 1 = 1$$

Therefore,

$$Z_2 = X_2 - X_1$$

3. $Z'_3 = X_3 - \langle X_3, Z_1 \rangle Z_1 - \langle X_3, Z_2 \rangle Z_2$
 $= X_3 - \langle X_3, X_2 - X_1 \rangle (X_2 - X_1) - \langle X_3, X_1 \rangle X_1$
 $= X_3 - (C_{32} - C_{31})(X_2 - X_1) - C_{31} X_1$
 $= X_3 - (X_2 - X_1) - X_1$
 $= X_3 - X_2$

$$\text{Var}(Z'_3) = E[(X_3 - X_2)(X_3 - X_2)] = C_{33} - 2C_{32} + C_{22} = 3 - 4 + 2 = 1$$

Therefore,

$$Z_3 = X_3 - X_2$$

Consequently

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \Rightarrow A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$