COL774: Machine Learning
Fall 2024-2025
More on Joint Distribution of Gaussains

Harshit Goyal

September 2024

1 Definition

Let random vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix}^T$. Let $\mathbf{Z} \in \mathcal{R}^l$ be the standard normal random vector (i.e. $Z_i \sim \mathcal{N}(0,1)$ for $i = 1, \dots, \ell$ are i.i.d.). Then $X_1, \dots, X_n$ are *jointly Gaussian* if there exist $\boldsymbol{\mu} \in \mathcal{R}^n$, $\mathbf{A} \in \mathcal{R}^{n \times l}$ such that $\mathbf{X} = \mathbf{AZ} + \boldsymbol{\mu}$.

From the above definition, note that the covariance matrix of $\mathbf{X}$ does not need to be invertible. Hence, every multivariate Gaussian does not have the p.d.f. of the form we saw in class. The constant random variable $\mathbf{X} = \boldsymbol{\mu}$ is also a Gaussian.

2 Other statements

The following statements are true

1. If $X_1, X_2, ..., X_n$ are marginally Gaussian then $X_1, X_2, ..., X_n$ need not be jointly Gaussian. See the example below.

2. If $X_1, X_2, ..., X_n$ are jointly Gaussian, that is, $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}^T$ follows multivariate Gaussian, then any projection $\mathbf{WX}$ follows multivariate Gaussian, $\mathbf{W} \in \mathcal{R}^{m \times n}$. Prove it using the definition.

3. The above implies that if X and Y are jointly Gaussian, then X, Y and $X + Y$ are Gaussian as

$$\begin{aligned} X &\sim \mathcal{N}(\mu_X, \sigma_X^2) \\ Y &\sim \mathcal{N}(\mu_Y, \sigma_Y^2) \\ X + Y &\sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2 + 2Cov(X, Y)) \end{aligned}$$

   Prove these by taking $\mathbf{W} = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$, $\mathbf{W} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, $\mathbf{W} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, respectively.

4. Given that $X_1, X_2, \dots, X_n \in \mathcal{R}$ are jointly Gaussian, $X_1, X_2, \dots, X_n$ are independent if and only if their covariance matrix is diagonal. The forward implication is easy to show. For the backwards, write the joint distribution of $X_1, X_2, \dots, X_n$. Then use:

$$\exp\left(\sum_{i=1}^{n} \frac{-(x_i - \mu_i)^2}{2\sigma_i^2}\right) = \prod_{i=1}^{n} \exp\left(\frac{-(x_i - \mu_i)^2}{2\sigma_i^2}\right)$$

3 Problem Setup

Consider independent random variables X and W distributed as:

$$X \sim \mathcal{N}(0, 1)$$

$$P(W = w) = \begin{cases} \frac{1}{2} & w = -1 \\ \frac{1}{2} & w = 1 \\ 0 & \text{otherwise} \end{cases}$$

Denote WX by Y. We'll show that

1) $Y \sim \mathcal{N}(0, 1)$
2) $X + Y$ is not Gaussian.
3) X and Y are uncorrelated, but not jointly Gaussian.

4 Proof

4.1 Y is marginally Gaussian

Let $\Phi(\cdot)$ denote the CDF of the Standard Normal. Let's prove:

$$P(Y \leq y) = \Phi(y)$$

Recall that the PDF of $\mathcal{N}(0, 1)$ is symmetric about $x = 0$, so

$$\Phi(x) = 1 - \Phi(-x) \quad \forall x \in \mathcal{R}$$

We compute

$$\begin{aligned} P(Y \leq y) &= P(Y \leq y, W = -1) + P(Y \leq y, W = 1) \quad \text{(since } W \text{ is either 1 or -1)} \\ &= P(XW \leq y, W = -1) + P(XW \leq y, W = 1) \\ &= P(-X \leq y, W = -1) + P(X \leq y, W = 1) \\ &= P(X \geq -y, W = -1) + P(X \leq y, W = 1) \\ &= (1 - \Phi(-y))\frac{1}{2} + \Phi(y)\frac{1}{2} \quad \text{(since } X \text{ and } W \text{ are independent)} \\ &= \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(y) \\ &= \Phi(y) \end{aligned}$$

This proves $WX \sim \mathcal{N}(0, 1)$.

4.2 CDF of $X + Y$

$$\begin{aligned} P(X + Y \leq \alpha) &= P(X + XW \leq \alpha) \\ &= P(X + XW \leq \alpha, W = -1) + P(X + XW \leq \alpha, W = 1) \\ &= P(0 \leq \alpha, W = -1) + P(2X \leq \alpha, W = 1) \\ &= P(0 \leq \alpha)P(W = -1) + P(2X \leq \alpha)P(W = 1) \\ &= P(0 \leq \alpha)\frac{1}{2} + P(X \leq \frac{\alpha}{2})\frac{1}{2} \\ &= \frac{1}{2}(\mathbb{1}_{(\alpha \geq 0)} + \Phi(\frac{\alpha}{2})) \end{aligned}$$

Which isn't the CDF of a Gaussian. Any univariate Gaussian will have a CDF of forrm $F(x) = \Phi(ax + b), a > 0$ or $F(x) = \mathbb{1}_{x \geq a}$.

4.3 Covariance of X and Y

Recall that if random variables X and Y are independent, then:

$$E[XY] = E[X]E[Y]$$

Since $E[W] = 0$ and X and W are independent:

$$\begin{aligned}\mathrm{Cov}(X, WX) &= E[X(WX)] - E[X]E[WX] \\ &= E[WX^2] - E[X]E[W]E[X] \\ &= E[W](E[X^2] - E[X]^2) = 0\end{aligned}$$

4.4 X and Y are not jointy Gaussian

Let $\alpha > 0$, then

$$(WX \leq -\alpha, X \leq -\alpha) = \frac{\Phi(-\alpha)}{2} \tag{1}$$

The above holds as $X \leq -\alpha$ with probability $\Phi(-\alpha)$ and

$$\begin{aligned}&(W = -1) \cap (X \leq -\alpha) \implies WX = -X \geq \alpha > -\alpha \\ &(W = 1) \cap (X \leq -\alpha) \implies WX = X \leq -\alpha\end{aligned}$$

The first case is not allowed. W takes value 1 with probability $\frac{1}{2}$, and X and W are independent. This justifies Equation 1. Let us assume (for contradiction) that X and Y are jointly Gaussian, noting that $\mathrm{Cov}(X, Y) = 0$, so the covariance matrix $\Sigma_{X,Y}$ will be I_2. Thus,

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

This shows that X and Y are independent and $P(X \leq -\alpha, WX \leq -\alpha) = \Phi(-\alpha)^2$, which contradicts Equation 1 as $\Phi(-\alpha)^2 = \frac{\Phi(-\alpha)}{2} \implies \Phi(-\alpha) = \frac{1}{2}$ which is not true as $\alpha > 0$.