Probability and Statistics: MA6.101
Surprise Quiz Solutions

1 5-Marks Questions

Question 1

Question: Suppose $X \sim N(\mu, \Sigma)$. Let $Y = W^{-1}(X - \mu)$, where W is a matrix that satisfies $W^2 = \Sigma$. Obtain the distribution for the random vector Y.
Answer:

Let the length of $\mathbf{X}$ be n and $\mathbf{X} = H(\mathbf{Y}) = W\mathbf{Y} + \mu$.
We denote the jacobian of H with J and let $W = [w_{ij}]_{n\times n}$. Then

$$J = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$
$$J = W$$

(Refer to Question 4 solution for explanation to calculate the Jacobian)
The pdf of $\mathbf{Y}$ is given by

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{X}}(H(\mathbf{y}))|\det(\mathbf{J})| \\ &= f_{\mathbf{X}}(W\mathbf{y} + \mu)|W| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}(W\mathbf{y} + \mu - \mu)^T \Sigma^{-1}(W\mathbf{y} + \mu - \mu)\right\}|W| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}|W|} \exp\left\{-\frac{1}{2}(W\mathbf{y})^T W^{-1}W^{-1}(W\mathbf{y})\right\}|W| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}\mathbf{y}^T\mathbf{y}\right\} \end{aligned}$$

Which is the same pdf as the n length standard normal vector. $Y \sim \mathcal{N}(\mathbf{0}, I_n)$. Note that W is a symmetric matrix, because it is decomposition of a given covariance matrix Σ.

Note: Other solutions will be given marks appropriately, if they follow correct logic and are coherent with the information given in the question.

Question 2

Question:

1. State the Central Limit Theorem.
2. Write the three equivalent definitions for a multivariate Gaussian.

Answer:

1. **Central Limit Theorem:**
   The Central Limit Theorem states that, given a sufficiently large sample size from a population with a finite level of variance, the mean of all samples from the same population will be approximately normally distributed. This holds regardless of the original distribution of the population, as long as the sample size is large enough.

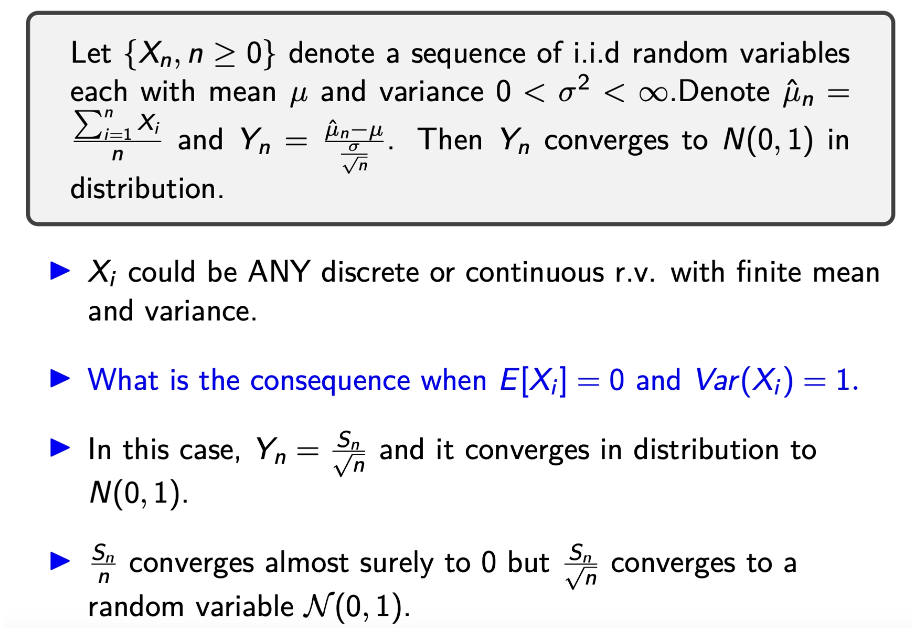


Figure 1: CLT Definition (Refer to L19 Slide 3)

2. **Equivalent Definitions for a Multivariate Gaussian:**
   (a) A random vector $\mathbf{X} \in \mathbb{R}^n$ is said to follow a multivariate Gaussian (normal) distribution if its probability density function is given by:

$$f_{\mathbf{X}}(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$$

where $\mu \in \mathbb{R}^n$ is the mean vector and $\Sigma \in \mathbb{R}^{n\times n}$ is the covariance matrix.

(b) A random vector $\mathbf{X}$ is multivariate Gaussian if every linear combination of its components is univariate Gaussian. That is, for any vector $\mathbf{a} \in \mathbb{R}^n$, the scalar $\mathbf{a}^\top \mathbf{X}$ is normally distributed.

(c) A random vector $\mathbf{X}$ is multivariate Gaussian if it can be written as $\mathbf{X} = \mu + A\mathbf{Z}$, where $\mathbf{Z}$ is a vector of independent standard normal variables and A is a matrix such that $\Sigma = AA^\top$.

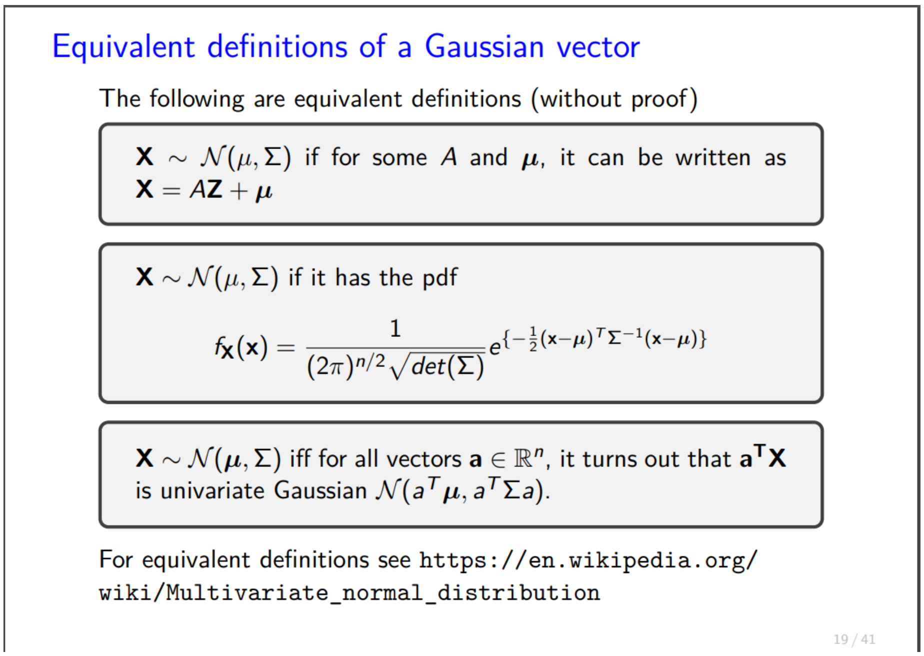


Figure 2: Gaussian Vector Definitions (Refer to L20 Slide 5)

Question 3

Question: Find the stationary distribution π for Markov chain with the following transition probability matrix. (3 marks) State if π is unique. (1 mark) Is the chain irreducible? Give reasons (1 mark)

$$P = \begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.9 & 0.1 \\ 0.1 & 0.9 & 0 \end{bmatrix}$$

Answer: To find the stationary distribution π, we need to solve $\pi = \pi P$ and $\pi_1 + \pi_2 + \pi_3 = 1$

This gives us the following equations:

$$\pi_1 = 0.2\pi_1 + 0\pi_2 + 0.1\pi_3$$
$$\pi_2 = 0.8\pi_1 + 0.9\pi_2 + 0.9\pi_3$$
$$\pi_3 = 0\pi_1 + 0.1\pi_2 + 0\pi_3$$
$$\pi_1 + \pi_2 + \pi_3 = 1$$

Solving these equations, we get: $\pi = (\frac{1}{89}, \frac{80}{89}, \frac{8}{89})$

We can independently verify the uniqueness of this stationary distribution in the following way:

$$\pi P = \pi,$$

can be written as:

$$P^T \pi^T = \pi^T.$$

which implies that π belongs to the null space of $P^T - I$, where I is the identity matrix. Specifically, we have

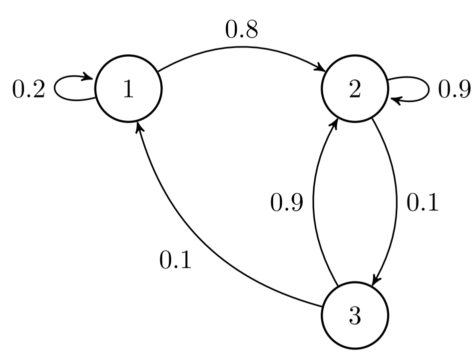
$$P^T - I = \begin{pmatrix} -0.8 & 0 & 0.1 \\ 0.8 & -0.1 & 0.9 \\ 0 & 0.1 & -1 \end{pmatrix},$$

which has the following row-reduced echelon form:

$$\begin{pmatrix} 1 & 0 & -0.125 \\ 0 & 1 & -10 \\ 0 & 0 & 0 \end{pmatrix}.$$

The row-reduced form indicates that $P^T - I$ has rank 2. So by the rank-nullity theorem, its null space (π) is one-dimensional (3-2 = 1). Since π must also be non-negative and sum to 1 (as it's a probability distribution), these additional constraints uniquely determine the stationary distribution.

Irreducibility: Markov chain:



$p_{ij}^{(n)}$ denotes the n-step transition probability from state i to state j. It is defined as:

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$$

The 2-step transition matrix $\mathbf{P^2}$ is:

$$P^2 = \begin{pmatrix} 0.04 & 0.88 & 0.08 \\ 0.01 & 0.90 & 0.09 \\ 0.02 & 0.89 & 0.09 \end{pmatrix}$$

State 1 and State 2:
$p_{12}^1 = 0.8 > 0, \text{so } 1 \to 2$
$p_{21}^2 = 0.1 \times 0.1 = 0.01 > 0, \text{so } 2 \to 1$
$\Rightarrow 1 \leftrightarrow 2.$

State 1 and State 3:
$p_{13}^2 = 0.8 \times 0.1 = 0.08 > 0, \text{so } 1 \to 3$
$p_{31}^1 = 0.1 > 0, \text{so } 3 \to 1$
$\Rightarrow 1 \leftrightarrow 3.$

State 2 and State 3:
$p_{23}^1 = 0.1 > 0, \text{so } 2 \to 3$
$p_{32}^1 = 0.9 > 0, \text{so } 3 \to 2$
$\Rightarrow 2 \leftrightarrow 3.$

Since we have demonstrated that any pair of states i and j communicate with each other ($i \leftrightarrow j$), the Markov chain is irreducible. In other words, the chain is irreducible since we can go from any state to any other state in a finite number of steps.

Question 4

Suppose $\mathbf{X}$ is a standard normal vector of size n. Let $\mathbf{Y} = A\mathbf{X} + b$, where A is a square symmetric invertible matrix. Derive the expression for $f_Y(y)$.

Answer:

Let H denote the inverse of $G(X) = AX + b$
Since A is an invertible matrix $H(y) = A^{-1}(y - b)$

where $\mathbf{J}$ is the Jacobian of $\mathbf{H}$ defined by

$$\mathbf{J} = \begin{bmatrix} \frac{\partial H_1}{\partial y_1} & \frac{\partial H_1}{\partial y_2} & \cdots & \frac{\partial H_1}{\partial y_n} \\ \frac{\partial H_2}{\partial y_1} & \frac{\partial H_2}{\partial y_2} & \cdots & \frac{\partial H_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial H_n}{\partial y_1} & \frac{\partial H_n}{\partial y_2} & \cdots & \frac{\partial H_n}{\partial y_n} \end{bmatrix}$$

We can see that, $\frac{\partial H_i}{\partial y_j}$ is nothing but the (i, j) element of the inverse matrix of $\mathbf{A}$. (Since differentiation drops the constant term $\mathbf{b}$)

$$\mathbf{J} = \begin{bmatrix} A_{11}^{-1} & A_{12}^{-1} & \cdots & A_{1n}^{-1} \\ A_{21}^{-1} & A_{22}^{-1} & \cdots & A_{2n}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}^{-1} & A_{n2}^{-1} & \cdots & A_{nn}^{-1} \end{bmatrix}$$

$$|\det(\mathbf{J})| = \left|\det(\mathbf{A}^{-1})\right| = \frac{1}{|\det(\mathbf{A})|}$$

Now, we know that

$$f_Y(\mathbf{y}) = f_X(H(\mathbf{y}) \cdot |\det(\mathbf{J})|$$

Thus,

$$f_Y(\mathbf{y}) = f_X(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})) \cdot \left|\det(\mathbf{A}^{-1})\right|$$

Substitute $f_X(\mathbf{x})$ [Note: X is a standard normal vector of size n] Thus,:

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{x}\right)$$

Thus:

$$f_Y(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}H(y)^T H(y)\right) \cdot \frac{1}{|\det(\mathbf{A})|}$$

Now,

$$H(y)^T H(y) = (\mathbf{y} - \mathbf{b})^T (\mathbf{A}^{-1})^T \mathbf{A}^{-1} (\mathbf{y} - \mathbf{b})$$

Since $(A^T)^{-1} = (A^{-1})^T$

$$H(y)^T H(y) = (\mathbf{y} - \mathbf{b})^T (\mathbf{A}^T)^{-1} \mathbf{A}^{-1} (\mathbf{y} - \mathbf{b})$$

Since $\mathbf{A}$ is symmetric:

$$H(y)^T H(y) = (\mathbf{y} - \mathbf{b})^T (\mathbf{A}^2)^{-1} (\mathbf{y} - \mathbf{b})$$

Thus,

$$f_Y(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}|\det(\mathbf{A})|} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{b})^T (\mathbf{A}^2)^{-1} (\mathbf{y} - \mathbf{b})\right)$$

2 10-Marks Questions

2.1 Question 1

Question:

Let (X, Y) be a pair of random variables with joint PDF $f_{X,Y}(x, y)$. Define a transformation from (X, Y) to (U, V) given by

$$U = X + Y, \quad V = X - Y$$

(a) Find the joint PDF $f_{U,V}(u,v)$ of the transformed random vector (U,V). (5 mks)

(b) Assume $f_{X,Y}(x,y) = ce^{-(x^2+y^2)}$ for all $x, y \in \mathbb{R}$. Find c. Find $f_{U,V}(u,v)$ for this specific case and elaborate on what kind of random variables are U and V.

Answer:

(a)

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} G_1(X,Y) \\ G_2(X,Y) \end{bmatrix} = \begin{bmatrix} X+Y \\ X-Y \end{bmatrix}$$

Solving for X and Y in terms of U and V we get

$$X = \frac{U+V}{2}, \quad Y = \frac{U-V}{2}$$

Let $(x,y) = H(u,v) = \left(\frac{u+v}{2}, \frac{u-v}{2}\right)$
We now find the determinant of the Jacobian.

$$J = \begin{vmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

We can now use the formula for $F_{U,V}(u,v)$.

$$\begin{aligned} f_{U,V}(u,v) &= |J|\, f_{X,Y}(H(u,v)) \\ &= \frac{1}{2} f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \end{aligned}$$

(b) To find c we do the following integration.

$$\begin{aligned} 1 &= c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\, dxdy \\ &= c\int_{-\infty}^{\infty} e^{-x^2}\, dx \int_{-\infty}^{\infty} e^{-y}\, dy \\ &= c\sqrt{\pi}\cdot\sqrt{\pi} \\ &\implies c = \frac{1}{\pi} \end{aligned}$$

Substituting into the formula from the previous question,

$$\begin{aligned} f_{U,V}(u,v) &= \frac{1}{2}\frac{1}{\pi}\exp\left\{-\left(\left(\frac{u+v}{2}\right)^2 + \left(\frac{u-v}{2}\right)^2\right)\right\} \\ &= \frac{1}{2\pi}\exp\left\{-\frac{1}{4}\left((u+v)^2 + (u-v)^2\right)\right\} \\ &= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(u^2+v^2)\right\} \end{aligned}$$

U, V have the standard bivariate normal distribution.