Assignment

Aryan Jain - EE22BTECH11011*

Question: Suppose that $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}, \mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_n}$ are independent and identically distributed random vectors each having $N_p(\mu, \Sigma)$ distributions, where Σ is non-singular, $p > 1$ and $n > 1$. If $\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X_i}$ and $\overline{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y_i}$, then which one of the following statements is true?

(a) There exists $c > 0$ such that $c\left(\overline{\mathbf{X}} - \mu\right)^T \Sigma^{-1}\left(\overline{\mathbf{X}} - \mu\right)$ has χ^2-distribution with p degrees of freedom.

(b) There exists $c > 0$ such that $c\left(\overline{\mathbf{X}} - \overline{\mathbf{Y}}\right)^T \Sigma^{-1}\left(\overline{\mathbf{X}} - \overline{\mathbf{Y}}\right)$ has χ^2-distribution with $(p-1)$ degrees of freedom.

(c) There exists $c > 0$ such that $c\sum_{i=1}^{n}\left(\mathbf{X_i} - \overline{\mathbf{X}}\right)^T \Sigma^{-1}\left(\mathbf{X_i} - \overline{\mathbf{X}}\right)$ has χ^2-distribution with p degrees of freedom.

(d) There exists $c > 0$ such that $c\sum_{i=1}^{n}\left(\mathbf{X_i} - \mathbf{Y_i} - \overline{\mathbf{X}} + \overline{\mathbf{Y}}\right)^T \Sigma^{-1}\left(\mathbf{X_i} - \mathbf{Y_i} - \overline{\mathbf{X}} + \overline{\mathbf{Y}}\right)$ has χ^2-distribution with p degrees of freedom.

GATE ST Paper 2023

Solution:

We are given that,

$$\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n}, \mathbf{Y_1}, \mathbf{Y_2}, \ldots, \mathbf{Y_n} \sim N_p(\mu, \Sigma) \tag{1}$$

Also,

$$\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X_i} \tag{2}$$

$$\overline{\mathbf{Y}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{Y_i} \tag{3}$$

The mean of $\overline{\mathbf{X}}$ is given by:

$$\mu_{\overline{\mathbf{X}}} = E\left(\overline{\mathbf{X}}\right) \tag{4}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left(\mathbf{X_i}\right) \tag{5}$$

$$= \mu \tag{6}$$

Similarly,

$$\mu_{\overline{\mathbf{Y}}} = \mu \tag{7}$$

The covariance of $\overline{\mathbf{X}}$ is given by:

$$\begin{aligned}
\Sigma_{\overline{\mathbf{X}}} &= E\left[\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^T\right] && (8)\\
&= E\left[\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X_i}-\mu\right)\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X_i}-\mu\right)^T\right] && (9)\\
&= \frac{1}{n^2}E\left[\sum_{i=1}^{n}(\mathbf{X_i}-\mu_i)\sum_{j=1}^{n}\left(\mathbf{X_j}-\mu_j\right)\right] && (10)\\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[(\mathbf{X_i}-\mu)\left(\mathbf{X_j}-\mu\right)\right] && (11)\\
&= \Sigma && (12)
\end{aligned}$$

Similarly,

$$\Sigma_{\overline{\mathbf{Y}}} = \Sigma \tag{13}$$

To check option (A):
let us say,

$$\mathbf{A} = \left(\overline{\mathbf{X}}-\mu\right)^T\Sigma^{-1}\left(\overline{\mathbf{X}}-\mu\right) \tag{14}$$

$$\tag{15}$$

And,

$$\begin{aligned}
\Sigma^{-1} &= \mathbf{F}^T\mathbf{F} && (16)\\
\overline{\mathbf{y}} &= \mathbf{F}\left(\overline{\mathbf{X}}-\mu\right) && (17)\\
\implies \mathbf{A} &= \overline{\mathbf{y}}^T\overline{\mathbf{y}} && (18)\\
&= \left\|\overline{\mathbf{y}}\right\|^2 && (19)
\end{aligned}$$

Equation (19) shows that $\mathbf{A}$ can have χ^2-distribution.
To confirm that we will find the mean and covariance-matrix of $\overline{\mathbf{y}}$.

$$\begin{aligned}
\mu_{\overline{\mathbf{y}}} &= E\left(\overline{\mathbf{y}}\right) && (20)\\
&= E\left(\mathbf{F}\left(\overline{\mathbf{X}}-\mu\right)\right) && (21)\\
&= \mathbf{F}\left[E\left(\overline{\mathbf{X}}\right)-E\left(\mu\right)\right] && (22)\\
&= \mathbf{F}\left[\mu-\mu\right] \quad from(6) && (23)\\
&= 0 && (24)
\end{aligned}$$

And,

$$\begin{aligned}
\Sigma_{\overline{\mathbf{y}}} &= E\left[\left(\overline{\mathbf{y}}-\mu_{\overline{\mathbf{y}}}\right)\left(\overline{\mathbf{y}}-\mu_{\overline{\mathbf{y}}}\right)^T\right] && (25)\\
&= E\left[\left(\mathbf{F}\left(\overline{\mathbf{X}}-\mu\right)\right)\left(\mathbf{F}\left(\overline{\mathbf{X}}-\mu\right)\right)^T\right] && (26)\\
&= E\left[\mathbf{F}\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^T\mathbf{F}^T\right] && (27)\\
&= \mathbf{F}\left[E\left[\left(\overline{\mathbf{X}}-\mu\right)\left(\overline{\mathbf{X}}-\mu\right)^T\right]\right]\mathbf{F}^T && (28)\\
&= \mathbf{F}\Sigma\mathbf{F}^T && (29)\\
&= \mathbf{I} \quad from(16) && (30)
\end{aligned}$$

Hence **A** has χ^2-distribution.
So option (A) is correct.