STA4002: Multivariate Techniques with Business Applications

Tutorial 1

Instructor: Dr.Lyu Teaching

Assistant: Zhou Yanan

1 Content Review

Definition of Multivariate Normal Distribution

$$X \sim MVN_p[\mu,\ \Sigma]$$

the density function:

$$f_X(X) = (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\}$$

Theorem 1. *(Theorem 8) Let X (with* p *components) be distributed according to $N(\mu,\ \Sigma)$ .Then*

$$Y = CX$$

is distributed according to $N(C\mu,\ C\Sigma C^T)$ for C *is nonsingular.*

Theorem 2. *(Theorem 10) If X is distributed according to $N(\mu,\ \Sigma)$. Then*

$$Y = DX$$

is distributed according to $N(D\mu,\ D\Sigma D^T)$, where D *is a $q \times p$ matrix of rand $q \leq p$.*

Equivalent Def 1 $X = (X_1, \ldots, X_p), X_p i.i.d.N(0,1)$. If $Y = \mu + AX$, where $\mu \in \mathbb{R}^p$ is a constant vector and $A \in \mathbb{R}^{p\times k}$ is constant matrix, then $Y \sim N_p(\mu,\ \Sigma)$ with $\Sigma = AA^T$.

Equivalent Def 2 If $Y \in \mathbb{R}^p$ and $E[e^{it^Ty}] = exp\{it^T\mu - t^T\Sigma t/2\}$ (the characteristic function), then $Y \sim N_p(\mu,\ \Sigma)$.

(i) The characteristic function of $N(0,1)$ is $e^{-t^2/2}$.

(ii) $X = (X_1, \ldots, X_p),\ X_p i.i.d.N(0,1)$.

$$E\mathrm{e}^{\mathrm{i}\boldsymbol{t}'\boldsymbol{X}} = \prod_{j=1}^{p} E\mathrm{e}^{\mathrm{i}t_j x_j} = \exp\left\{-\sum_{j=1}^{p}\frac{t_j^2}{2}\right\} = \exp\left\{-\frac{\boldsymbol{t}'\boldsymbol{t}}{2}\right\}.$$

(iii) $Y = \mu + AX$.

$$E\mathrm{e}^{\mathrm{i}\boldsymbol{t}'\boldsymbol{y}} = E\mathrm{e}^{\mathrm{i}\boldsymbol{t}'(\boldsymbol{\mu}+\boldsymbol{A}\boldsymbol{X})} = \exp\{\mathrm{i}\boldsymbol{t}'\boldsymbol{\mu}\}E\mathrm{e}^{\mathrm{i}(\boldsymbol{A}'\boldsymbol{t})'\boldsymbol{X}} = \exp\left\{\mathrm{i}\boldsymbol{t}'\boldsymbol{\mu} - \frac{(\boldsymbol{A}'\boldsymbol{t})'\boldsymbol{A}'\boldsymbol{t}}{2}\right\}$$
$$= \exp\left\{\mathrm{i}\boldsymbol{t}'\boldsymbol{\mu} - \frac{\boldsymbol{t}'\boldsymbol{A}\boldsymbol{A}'\boldsymbol{t}}{2}\right\} = \exp\left\{\mathrm{i}\boldsymbol{t}'\boldsymbol{\mu} - \frac{\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t}}{2}\right\},$$

2 Some Properties of MVN

$$D = \begin{pmatrix} d_{1,1} & d_{1,2} & \dots & d_{1,p} \\ d_{2,1} & d_{2,2} & \dots & d_{2,p} \\ \vdots \\ d_{q,1} & d_{q,2} & \dots & d_{q,p} \end{pmatrix}$$

$$Y = \begin{pmatrix} d_{1,1}X_1 + d_{1,2}X_2 + \ldots + d_{1,p}X_p \\ d_{2,1}X_1 + d_{2,2}X_2 + \ldots + d_{2,p}X_p \\ \vdots \\ d_{q,1}X_1 + d_{q,2}X_2 + \ldots + d_{q,p}X_p \end{pmatrix}$$

In other words, the distribution of q linear combinations of the components of a normal random vector is still multivariate normal distribution.

Ql: What is the distribution of $X_1 + 2X_2 - 3X_3$,when $X = (X_1, X_2, X_3)$ have the distribution defined as below.

The density function:

$$f_x(x) = (2\pi)^{-p/2}|A|^{1/2} \exp\{-\frac{1}{2}(x-b)^T A(x-b)\}.$$

Let b $= 0$, and $A = \begin{pmatrix} 7 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & 2 \end{pmatrix}$.

Solution Find $\Sigma = A^{-1} = \frac{A^*}{|A|}$, where A^* is the adjoint matrix. We have $|A| = 7\times 7 - 3\times 4 + 2\times(-5) = 49 - 12 - 10 = 27$.

$$A^{-1} = \frac{1}{27}\begin{pmatrix} 7 & -4 & -5 \\ -4 & 10 & -1 \\ -5 & -1 & 19 \end{pmatrix}, Y = X_1 + 2X_2 - 3X_3 = (1, 2, -3)\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

According to Theorem 10, $Y \sim N(D\mu,\ D\Sigma D^T)$, where $D = (1, 2, -3)$. Thus, we have $D\mu = 0$ and

$$D\Sigma D^T = \frac{1}{27}(1, 2, -3)\begin{pmatrix} 7 & -4 & -5 \\ -4 & 10 & -1 \\ -5 & -1 & 19 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \frac{244}{27},$$

that is $Y \sim N(0,\ \frac{244}{27})$.

Corollary 1. *Let $X = (X_1, X_2)^T$, and $X \sim MVN_p[\mu,\ \Sigma]$ with $\mu = (\mu_1, \mu_2)^T$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$, where $|\Sigma_{22}| > 0$. Then the conditional distribution of X_1, given that $X_2 = x_2$, is normal and has Mean $= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$, and Covariance $= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.*

Remark of Corollary (1)

(i) The conditional distributions derived from joint normal distribution are normal.

(ii) Note that the covariance does not depend on the value x_2 of the conditioning variable.

Proof. We shall give an indirect proof. (See Exercise 4.13, which uses the densities directly.) Take

$$\underset{(p\times p)}{\mathbf{A}} = \left[\begin{array}{c:c} \underset{(q\times q)}{\mathbf{I}} & \underset{q\times(p-q)}{-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}} \\ \hdashline \underset{(p-q)\times q}{\mathbf{0}} & \underset{(p-q)\times(p-q)}{\mathbf{I}} \end{array}\right]$$

so

$$\mathbf{A}(\mathbf{X}-\boldsymbol{\mu}) = \mathbf{A}\left[\begin{array}{c} \mathbf{X}_1-\boldsymbol{\mu}_1 \\ \hdashline \mathbf{X}_2-\boldsymbol{\mu}_2 \end{array}\right] = \boxed{\left[\begin{array}{c} \mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2) \\ \hdashline \mathbf{X}_2-\boldsymbol{\mu}_2 \end{array}\right]}$$

is jointly normal with covariance matrix $\boxed{\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'}$ given by

$$\left[\begin{array}{c:c} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \hdashline \mathbf{0} & \mathbf{I} \end{array}\right]\left[\begin{array}{c:c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hdashline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]\left[\begin{array}{c:c} \mathbf{I} & \mathbf{0}' \\ \hdashline (-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})' & \mathbf{I} \end{array}\right] = \boxed{\left[\begin{array}{c:c} \boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{0}' \\ \hdashline \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{array}\right]}$$

Since $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$ and $\mathbf{X}_2-\boldsymbol{\mu}_2$ have zero covariance, they are independent. Moreover, the quantity $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$ has distribution $N_q(\mathbf{0},\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$. Given that $\mathbf{X}_2=\mathbf{x}_2$, $\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)$ is a constant. Because $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$ and $\mathbf{X}_2-\boldsymbol{\mu}_2$ are independent, the conditional distribution of $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)$ is the same as the unconditional distribution of $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$. Since $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2-\boldsymbol{\mu}_2)$ is $N_q(\mathbf{0},\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$, so is the random vector $\mathbf{X}_1-\boldsymbol{\mu}_1-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2)$ when $\mathbf{X}_2$ has the particular value $\mathbf{x}_2$. Equivalently, given that $\mathbf{X}_2=\mathbf{x}_2$, $\mathbf{X}_1$ is distributed as $N_q(\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2-\boldsymbol{\mu}_2),\ \boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$. ■

Corollary 2. *If X is distributed according to $N(\mu,\ \Sigma)$, the marginal distribution of any set of components of X is multivariate normal with means, variances, and covariances obtained by taking the corresponding components of μ and Σ, respectively.*

Lemma 1.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{2|1} \end{pmatrix},$$

where $A_{2|1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$.

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & 0 \\ 0 & A_{2|1} \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}.$$

Thus, we have

$$|A| = |A_{11}| \times |A_{2|1}|, \tag{1}$$

$$A^{-1} = \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{2|1}^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}. \tag{2}$$

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Now we proof Corollary (2).

Proof. According to Equation (1) and (2), $\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

$$|\Sigma| = |\Sigma_{11}| \times |\Sigma_{2|1}|,$$

$$\Sigma^{-1} = \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & \Sigma_{2|1}^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\Sigma_{21}A_{11}^{-1} & I \end{pmatrix},$$

$$\begin{aligned} f_X(X) &= (2\pi)^{-p/2}|\Sigma|^{-1/2}\exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\} \\ &= f_1(x_1)f_{2|1}(x_2|x_1), \\ f_1(x_1) &= (2\pi)^{-q/2}|\Sigma_{11}|^{-1/2}\exp\{-\frac{1}{2}(x_1-\mu_1)^T\Sigma_{11}^{-1}(x_1-\mu_1)\}, \\ f_{2|1}(x_2|x_1) &= (2\pi)^{(p-q)/2}|\Sigma_{2|1}|^{-1/2}\exp\{-\frac{1}{2}(x_2-\mu_{2|1})^T\Sigma_{2|1}^{-1}(x_2-\mu_{2|1})\}. \end{aligned}$$

where $\mu_{2|1} = \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(y_1-\mu_1)$.
Thus, we have $X_1 \sim N_q(\mu_1, \Sigma_{11})$. □

Remark of Proof If we regard y_1 as a constant, then $f_{2|1}(x_2|x_1)$ is the pdf of $N_{p-q}(\mu_{2|1},\ \Sigma_{2|1})$.

Another description of Corollary (2) :

Corollary 3. *If $X \sim MVN_p[\mu,\ \Sigma]$, then all subsets of X are normally distributed. Specifically, if we respectively partition X, μ and Σ as* $\underset{(p\times 1)}{X}{}^T = \left(\underset{(q\times 1)}{X_1}{}^T, \underset{((p-q)\times 1)}{X_2}{}^T\right)$, $\underset{(p\times 1)}{\mu}{}^T = \left(\underset{(q\times 1)}{\mu_1}{}^T, \underset{((p-q)\times 1)}{\mu_2}{}^T\right)$, *and*

$$\underset{(p\times p)}{\Sigma} = \begin{pmatrix} \underset{(q\times q)}{\Sigma_{11}} & \underset{(q\times(p-q))}{\Sigma_{12}} \\ \underset{((p-q)\times q)}{\Sigma_{21}} & \underset{((p-q)\times(p-q))}{\Sigma_{22}} \end{pmatrix},$$

then $X_1 \sim MVN_q[\mu_1,\ \Sigma_{11}]$.

Proof. Set $D = (I_{q\times q}, 0)$ in Theorem 10. $\square$

To apply this result to an arbitrary subset of the components of X, we simply relabel the subset of the interest as X_1 and select the corresponding component means and covariances as μ_1 and Σ_{11}, respectively. For example, $X = (x_1, x_2, x_3)^T \sim MVN_p[\mu,\ \Sigma]$, we want to find the distribution of $(x_1, x_3)^T$. We set $X = (x_1, x_3, x_2)^T$, $\mu = (\mu_1, \mu_3, \mu_2)^T$, and $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{13} & \sigma_{21} \\ \sigma_{31} & \sigma_{33} & \sigma_{23} \\ \sigma_{21} & \sigma_{23} & \sigma_{22} \end{pmatrix}$. In the end, $X_1 = (x_1, x_3)^T$