

# STAT 3100 Lecture Note

Week Two (Sep 13 & 15, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Sep/09 17:42:18

---

## Multivariate Transformation

### Multivariate distribution

- Random vector  $\mathbf{X} = (X_1, \dots, X_n)$  with realization  $\mathbf{x} = (x_1, \dots, x_n)$ 
  - cdf  $F_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 \leq x_1, \dots, X_n \leq x_n)$
- Discrete
  - Joint pmf

$$p_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 = x_1, \dots, X_n = x_n)$$

- $\text{supp}(\mathbf{X}) = \text{supp}(p_{\mathbf{X}}) = \{\mathbf{x} \in \mathbb{R}^n : p_{\mathbf{X}}(\mathbf{x}) > 0\}$
- Marginal pmf of  $(X_1, \dots, X_k)$

$$p_{X_1, \dots, X_k}(x_1, \dots, x_k) = \sum_{(x_{k+1}, \dots, x_n) \in \mathbb{R}^{n-k}} p_{\mathbf{X}}(\mathbf{x})$$

- 
- Continuous
    - Joint pdf

$$f_{\mathbf{X}}(\mathbf{x}) = (\partial^n / \partial x_1 \cdots \partial x_n) F_{\mathbf{X}}(\mathbf{x})$$

- \*  $\Pr(\mathbf{X} \in B) = \int_B f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$  for each Borel set  $B \subset \mathbb{R}^n$
- $\text{supp}(\mathbf{X}) = \text{supp}(f_{\mathbf{X}}) = \{\mathbf{x} \in \mathbb{R}^n : f_{\mathbf{X}}(\mathbf{x}) > 0\}$
- Marginal pdf of  $(X_1, \dots, X_k)$ 
  - \*  $f_{X_1, \dots, X_k}(x_1, \dots, x_k) = \int_{\mathbb{R}^{n-k}} f_{\mathbf{X}}(\mathbf{x}) dx_{k+1} \cdots dx_n$

**Find the joint pdf of random vector  $\mathbf{Y} = g(\mathbf{X})$  by multivariate transformation (CB Sec. 4.3 & 4.6)**

- Conditions
  - $\mathbf{X}$  and  $\mathbf{Y}$  both of  $n$  dimensions
  - $g(\cdot) = (g_1(\cdot), \dots, g_n(\cdot)) : \text{supp}(\mathbf{X}) \rightarrow \text{supp}(\mathbf{Y})$  is one-to-one, i.e.,
    - \*  $\mathbf{y} = (y_1, \dots, y_n) = (g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) = g(\mathbf{x})$
    - \*  $\mathbf{x} = (x_1, \dots, x_n) = g^{-1}(\mathbf{y}) = (g_1^{-1}(\mathbf{y}), \dots, g_n^{-1}(\mathbf{y}))$
- Jacobian matrices
  - Jacobian matrix of transformation  $g^{-1}$

$$\mathbf{J}_{g^{-1}} = \mathbf{J}_{g^{-1}}(\mathbf{y}) = \left[ \frac{\partial g_i^{-1}(\mathbf{y})}{\partial y_j} \right]_{n \times n} = \begin{bmatrix} \frac{\partial g_1^{-1}(\mathbf{y})}{\partial y_1} & \cdots & \frac{\partial g_1^{-1}(\mathbf{y})}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n^{-1}(\mathbf{y})}{\partial y_1} & \cdots & \frac{\partial g_n^{-1}(\mathbf{y})}{\partial y_n} \end{bmatrix}$$

- Jacobian matrix of transformation  $\mathbf{g}$

$$\mathbf{J}_{\mathbf{g}} = \mathbf{J}_{\mathbf{g}}(\mathbf{x}) = \left[ \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right]_{n \times n} = \begin{bmatrix} \frac{\partial g_1(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial g_1(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial g_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

- $\mathbf{J}_{\mathbf{g}^{-1}}(\mathbf{y}) = \{\mathbf{J}_{\mathbf{g}}(\mathbf{g}^{-1}(\mathbf{y}))\}^{-1}$ 
  - \* Alternative way to reach  $\mathbf{J}_{\mathbf{g}^{-1}}(\mathbf{y})$
- Then
 
$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}\{\mathbf{g}^{-1}(\mathbf{y})\} |\det\{\mathbf{J}_{\mathbf{g}^{-1}}(\mathbf{y})\}| \mathbf{1}_{\text{supp}(\mathbf{Y})}(\mathbf{y}).$$
  - Never miss  $\mathbf{1}_{\text{supp}(\mathbf{Y})}(\mathbf{y})$
- If  $\mathbf{g}$  is NOT one-to-one, one may figure out the cdf of  $\mathbf{Y}$  and then differentiate it.

### Example Lec3.1

$X_1$  and  $X_2$  are iid from  $\mathcal{N}(0, 1)$ . Find the joint pdf of  $Y_1 = (X_1 + X_2)/\sqrt{2}$  and  $Y_2 = (X_1 - X_2)/\sqrt{2}$  and show their independence.

Note: the sample mean and standard deviation are respectively  $\bar{X} = (X_1 + X_2)/2 = Y_1/\sqrt{2}$  and  $S = \sqrt{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2} = |Y_2|$ .

### Find the marginal pdf

1. Figure out the joint pdf first
2. Taking the Integral

### Example Lec3.2

$X_1$  and  $X_2$  are iid from  $\mathcal{N}(0, 1)$ . Find the pdf of  $U = \sqrt{X_1^2 + X_2^2}$ .

## Basics on square matrices

### Eigen-decomposition

- $\mathbf{A}$  is a real  $n \times n$  matrix
- Eigenvalues of  $\mathbf{A}$ , say  $\lambda_1 \geq \dots \geq \lambda_n$ :  $n$  roots of characteristic equation  $\det(\lambda \mathbf{I}_n - \mathbf{A}) = 0$
- The  $i$ th (Right) eigenvector  $\mathbf{v}_i$ :  $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$
- Eigen-decomposition:  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ 
  - $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$  and  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$  both  $n \times n$  matrices
  - Specifically  $\mathbf{V}^{-1} = \mathbf{V}^T$  for symmetric  $\mathbf{A}$
- Numerical implementation in  $R$ : `eigen()`
- Connection to determinant and trace
  - Determinant
    - \*  $\det \mathbf{A} = \prod_{i=1}^n \lambda_i$
    - \*  $\det(\mathbf{A}^T) = \det \mathbf{A}$
    - \*  $\det(\mathbf{A}^{-1}) = (\det \mathbf{A})^{-1}$
    - \*  $\det(c\mathbf{A}) = c^n \det \mathbf{A}$  for  $n \times n$  matrix  $\mathbf{A}$  and scalar  $c$
    - \*  $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$  for squared  $\mathbf{A}$  and  $\mathbf{B}$
  - Trace

- \*  $\text{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$
- \*  $\text{tr}(c\mathbf{A}) = c\text{tr}(\mathbf{A})$  for scalar  $c$
- \*  $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$  for squared  $\mathbf{A}$  and  $\mathbf{B}$
- \*  $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$

## Square root of matrices

- $\mathbf{A}^{1/2} = \mathbf{V}\mathbf{\Lambda}^{1/2}\mathbf{V}^\top$  if for semi-positive definite  $\mathbf{A}$ 
  - Semi-positive/non-negative definite: symmetric  $\mathbf{A}$  with eigenvalues all non-negative
  - $\mathbf{\Lambda}^{1/2} = \text{diag}(\lambda_1^{1/2}, \dots, \lambda_n^{1/2})$
  - $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$
- $\mathbf{A}^{-1/2} = \mathbf{V}\mathbf{\Lambda}^{-1/2}\mathbf{V}^\top$  for positive definite  $\mathbf{A}$ 
  - Positive definite: symmetric  $\mathbf{A}$  with eigenvalues all positive
  - $\mathbf{\Lambda}^{-1/2} = \text{diag}(\lambda_1^{-1/2}, \dots, \lambda_n^{-1/2})$
  - $\mathbf{A}^{-1/2}\mathbf{A}^{-1/2} = \mathbf{A}^{-1}$  and  $\mathbf{A}^{1/2}\mathbf{A}^{-1/2} = \mathbf{I}$

## Multivariate normal (MVN) distribution

MVN( $\mathbf{0}, \mathbf{I}_p$ )

- Random  $p$ -vector  $\mathbf{Z} = (Z_1, \dots, Z_p)^\top \sim \text{MVN}(\mathbf{0}, \mathbf{I}_p) \Leftrightarrow \text{iid } Z_1, \dots, Z_p \sim \mathcal{N}(0, 1)$ .
- pdf of MVN( $\mathbf{0}, \mathbf{I}_p$ ):

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= \prod_{i=1}^p (2\pi)^{-1/2} \exp(-z_i^2/2) \\ &= (2\pi)^{-p/2} \exp(-\mathbf{z}^\top \mathbf{z}/2), \quad \mathbf{z} \in \mathbb{R}^p \end{aligned}$$

MVN( $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ )

- pdf of MVN( $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ ),  $\boldsymbol{\Sigma} > 0$ :

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-p/2} (\det \boldsymbol{\Sigma})^{-1/2} \exp\{-(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})/2\}, \quad \mathbf{x} \in \mathbb{R}^p$$

- $\mathbf{X} \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Rightarrow \mathbf{AX} + \mathbf{a} \sim \text{MVN}(\mathbf{A}\boldsymbol{\mu} + \mathbf{a}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$  for arbitrary  $\mathbf{A} \in \mathbb{R}^{q \times p}$  and  $\mathbf{a} \in \mathbb{R}^q$
- $\mathbf{X} \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \mathbf{X} = \boldsymbol{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}$  with  $\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim \text{MVN}(\mathbf{0}, \mathbf{I}_p)$

## Marginals of MVN

- Suppose  $p$ -vector  $\mathbf{X} = (X_1, \dots, X_p)^\top$  and  $q$ -vector  $\mathbf{Y} = (Y_1, \dots, Y_q)^\top$  are jointly normally distributed. Then,  $\mathbf{X}$  and  $\mathbf{Y}$  are independent  $\Leftrightarrow \text{cov}(\mathbf{X}, \mathbf{Y}) = \mathbf{0}_{p \times q}$ .
- If  $\mathbf{X}$  is of MVN, then its all margins are still of MVN. The inverse proposition does NOT hold.
  - Cautionary example: Let  $Y = XZ$ , where  $X \sim \mathcal{N}(0, 1)$ ;  $Z$  is independent of  $X$  with  $\Pr(Z = 1) = \Pr(Z = -1) = .5$ .  $X$  and  $Y$  both turn out to be of standard normal, but they are not jointly normal.

## Normal sampling theory (CB Sec. 5.3)

### Stochastic representations for $\chi^2$ -, $t$ -, and $F$ -r.v. (HMC Chp. 3)

- If iid  $X_1, \dots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ , then
  - $\sum_{i=1}^n X_i^2 \sim \chi^2(n)$  if iid  $X_1, \dots, X_n \sim \mathcal{N}(0, 1)$ ;
  - $X/\sqrt{Y/n} \sim t(n)$  if  $X \sim \mathcal{N}(0, 1)$  and  $Y \sim \chi^2(n)$  are independent;

–  $(X/m)/(Y/n) \sim F(m, n)$  if  $X \sim \chi^2(m)$  and  $Y \sim \chi^2(n)$  are independent.

### Important identities for normal samples

- $\bar{X} = n^{-1} \sum_{i=1}^n X_i$  and  $S^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X})^2$  are independent
- $n^{1/2}(\bar{X} - \mu)/\sigma \sim \mathcal{N}(0, 1)$
- $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
- $n^{1/2}(\bar{X} - \mu)/S \sim t(n-1)$