

Multivariate Analysis Lecture 5: Normal and Multivariate Normal

Zhaoxia Yu
Professor, Department of Statistics

2023-04-19

Section 1

The Big Picture

The Big Picture: Univariate vs Multivariate

- **Review:** A random sample, denoted by X_1, \dots, X_n , from a (univariate) normal distribution $N(\mu, \sigma^2)$
 - What are the distributions of \bar{X}, s^2 ? What useful statistics can be constructed?
- **New material:** A random sample, denoted by $\mathbf{X}_1, \dots, \mathbf{X}_n$, from a multivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
 - What are the distributions of $\bar{\mathbf{X}}, \mathbf{S}$? What useful statistics can be constructed?

The Big Picture: Univariate

- A random sample, denoted by X_1, \dots, X_n , from a (univariate) normal distribution $N(\mu, \sigma^2)$
- Let $\mathbf{X}_{n \times 1} = (X_1, \dots, X_n)^T$. It is random vector with a multivariate normal distribution, i.e.,

$$\mathbf{X}_{n \times 1} = (X_1, \dots, X_n)^T \sim \mathbf{N}(\mu \mathbf{1}, \sigma^2 \mathbf{I})$$

- 1 $\bar{X} \sim N(\mu, \sigma^2/n)$
- 2 $\frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2$
- 3 Independence between \bar{X} and s^2 .
- 4 a t-statistic is

$$\frac{\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}}{\sqrt{\frac{(n-1)s^2/\sigma^2}{n-1}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{s} \sim t_{n-1}$$

The Big Picture: Multivariate

- A random sample $\mathbf{X}_1, \dots, \mathbf{X}_n$ from a multivariate normal distribution $\mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- Let

$$\mathbf{X}_{n \times p} = \begin{pmatrix} \mathbf{X}_1^T \\ \vdots \\ \mathbf{X}_n^T \end{pmatrix}$$

\mathbf{X} follows a matrix normal distribution.

- 1 Sample mean vector follows a multivariate normal, i.e., $\bar{\mathbf{X}} \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$
- 2 Sample covariance matrix $(n-1)\mathbf{S}$ follows a Wishart distribution, i.e., $(n-1)\mathbf{S} \sim \text{Wishart}_p(n-1, \boldsymbol{\Sigma})$
- 3 Independence between $\bar{\mathbf{X}}$ and S .
- 4 Hoetelling's T^2 : $T^2 = (\bar{\mathbf{X}} - \boldsymbol{\mu})^T \left(\frac{\mathbf{S}}{n}\right)^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu})$

Outline

- Multivariate normal distribution (MVN)
- Moment generating function (MGF)
 - Apply MGF to univariate normal
 - Apply MGF to multivariate normal
- Zero-Cov vs Independence
- MVN: $\bar{\mathbf{X}}$ and \mathbf{S}

Section 2

MVN

PDF of Normal of Distributions

- Univariate normal distribution:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Bivariate normal distribution:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} \right)}$$

The formula for a $p \geq 3$ -dimensional multivariate normal distribution is much messier, so we use a compact way:

- Multivariate normal distribution:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Section 3

MGF

Tools to Characterize a Distribution

- Probability density function (PDF) or probability mass function (PMF)
- Cumulative distribution (CDF)
- Characteristic function (CF)
- Moment generating function (MGF)
-

Moment Generating Function (MGF)

- The moment generating function of random variable X is defined

$$M_X(t) = \mathbb{E}[e^{tX}]$$

- Like a PDF/PMF or CDF, a MGF uniquely determines/identifies a distribution
- The definition can be extended to random vectors and random matrices
 - Consider a random vector $\mathbf{X}_{p \times 1}$. Let t be a $p \times 1$ vector.

$$M_{\mathbf{X}} = \mathbb{E}[e^{t^T \mathbf{X}}]$$

- Consider a random matrix $\mathbf{X}_{n \times p}$. Let t be a $n \times p$ matrix.

$$M_{\mathbf{X}} = \mathbb{E}[e^{\text{trace}(t^T \mathbf{X})}]$$

Moment Generating Function: Univariate

- Where does the name of MGF come from?

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx \\&= \int_{-\infty}^{\infty} \left[1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots\right] f(x) dx \\&= 1 + t\mathbb{E}[X] + \frac{t^2}{2!}\mathbb{E}[X^2] + \cdots\end{aligned}$$

- $M_X^{(k)}(0) = E[X^k]$, where $M_X^{(k)}(t)$ is the k th derivative of $M_X(t)$.

Subsection 1

MGF: Univariate Normal

MGF of Univariate Normal

- Recall that the MGF of a random variable X is defined as:
 $M_X(t) = \mathbb{E}[e^{tX}]$.
- For the normal distribution with mean μ and variance σ^2 , the MGF is given by:

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

- The mean is $\mathbb{E}[X] = M'_X(0) = \mu$.
- The variance is

$$\begin{aligned}\text{Var}(X) &= E[(X - \mu)^2] = \dots = E[X^2] - (E[X])^2 \\ &= M''_X(0) - M'_X(0)^2 = \sigma^2\end{aligned}$$

MGF of Univariate Normal: Examples

- Recall that $M_X(t) = \exp\{\mu t + \frac{1}{2}\sigma^2 t^2\}$ for $X \sim N(\mu, \sigma^2)$.

What is the distribution corresponding to each of the following MGFs?

1

$$M_X(t) = \exp\left(\frac{1}{2}t^2\right)$$

2

$$M_X(t) = \exp\left(2t + \frac{9}{2}t^2\right)$$

3

$$M_X(t) = \exp\left(-t + \frac{1}{8}t^2\right)$$

MGF of Univariate Normal: Examples (continued)

- ① Standard normal distribution, i.e., $\mu = 0, \sigma^2 = 1$:

$$M_X(t) = \exp\left(\frac{1}{2}t^2\right)$$

- ② Normal distribution with mean $\mu = 2$ and standard deviation $\sigma = 3$:

$$M_X(t) = \exp\left(2t + \frac{9}{2}t^2\right)$$

- ③ Normal distribution with mean $\mu = -1$ and standard deviation $\sigma = 0.5$:

$$M_X(t) = \exp\left(-t + \frac{1}{8}t^2\right)$$

MGF of Univariate Normal: A Linear Function

- Let $X \sim N(\mu, \sigma^2)$. We know that $M_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$
- Let $Y = aX + b$, where a and b are constants.
- We now find $M_Y(t)$:

$$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(aX+b)}] = e^{bt} \mathbb{E}[e^{(at)X}]$$

Since at is just another constant, we can treat it as a new variable, say $s = at$. Then:

$$\begin{aligned} M_Y(t) &= e^{bt} M_X(s) = e^{bt} \exp\left\{\mu s + \frac{1}{2}\sigma^2 s^2\right\} \\ &= \exp\{bt + a\mu t + \frac{1}{2}\sigma^2 a^2 t^2\} = \exp\{(a\mu + b)t + \frac{1}{2}(a\sigma)^2 t^2\} \end{aligned}$$

- $M_Y(t)$ has the form of the MGF of a normal distribution:
 $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.

MGF of Univariate Normal: Sum of Two Independent Normal

- Let X and Y be two independent and $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$.

$$M_X(t) = \exp\{\mu_X t + \frac{1}{2}\sigma_X^2 t^2\}, M_Y(t) = \exp\{\mu_Y t + \frac{1}{2}\sigma_Y^2 t^2\}$$

- Let $Z = X + Y$.

$$\begin{aligned} M_Z(t) &\stackrel{X \perp Y}{=} M_X(t)M_Y(t) = \exp\{\mu_X t + \frac{1}{2}\sigma_X^2 t^2\} \exp\{\mu_Y t + \frac{1}{2}\sigma_Y^2 t^2\} \\ &= \exp\{(\mu_X + \mu_Y)t + \frac{1}{2}(\sigma_X^2 + \sigma_Y^2)t^2\} \end{aligned}$$

Which indicates that $Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

MGF of Univariate Normal: Sample Mean

- If $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$.
- We have showed that $E[\bar{X}] = \mu$ and $Var[\bar{X}] = \sigma^2/n$.
- How to prove \bar{X} follows a normal distribution?
- A compact proof:

$$M_{\bar{X}}(t) = \prod_{i=1}^n M_{X_i}\left(\frac{t}{n}\right) = \left(\exp\left\{ \mu \frac{t}{n} + \frac{1}{2} \sigma^2 \frac{t^2}{n^2} \right\} \right)^n = \exp\left\{ \mu t + \frac{1}{2} \frac{\sigma^2}{n} t^2 \right\}$$

Based on the $M_{\bar{X}}(t)$, $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

MGF of Univariate Normal: Sample Mean

- A proof with more details explained

$$\begin{aligned}M_{\bar{X}}(t) &= E[e^{t\bar{X}}] = E[e^{\frac{t}{n} \sum_{i=1}^n X_i}] = E[e^{\frac{t}{n} X_1 + \frac{t}{n} X_2 + \dots + \frac{t}{n} X_n}] \\&\stackrel{iid}{=} E[e^{\frac{t}{n} X_1}] \dots E[e^{\frac{t}{n} X_n}] = M_{X_1}\left(\frac{t}{n}\right) \dots M_{X_n}\left(\frac{t}{n}\right) \\&= \exp\left\{\mu \frac{t}{n} + \frac{1}{2} \sigma^2 \left(\frac{t}{n}\right)^2\right\} \dots \exp\left\{\mu \frac{t}{n} + \frac{1}{2} \sigma^2 \left(\frac{t}{n}\right)^2\right\} \\&= \exp\left\{\mu t + \frac{1}{2} \frac{\sigma^2}{n} t^2\right\}\end{aligned}$$

Based on the $M_{\bar{X}}(t)$, $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

Subsection 2

MGF of MVN

MGF of Multivariate Normal

- The moment generating function (MGF) of a random vector $\mathbf{X}_{p \times 1}$ is defined as: $M_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[e^{\mathbf{t}^T \mathbf{X}}]$.
- Here \mathbf{t} is a $p \times 1$ vector.
- For the multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, the MGF is given by:

$$M_{\mathbf{X}}(\mathbf{t}) = \exp \left(\boldsymbol{\mu}^T \mathbf{t} + \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right)$$

MGF of MVN: Examples

- ① Bivariate standard normal distribution:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, M_{\mathbf{X}}(\mathbf{t}) = \exp\left(\frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)$$

- ② Bivariate normal distribution with specific mean vector and covariance matrix:

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}, M_{\mathbf{X}}(\mathbf{t}) = \exp\left(\boldsymbol{\mu}^T \mathbf{t} + \frac{1}{2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)$$

MGF of MVN: A Linear Combination

- Let $\mathbf{X}_{p \times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- We want to show that the linear combinations $\mathbf{Y} = \mathbf{A}_{q \times p} \mathbf{X}$ also follows a multivariate normal distribution.
- The MGF of \mathbf{X} is

$$M_{\mathbf{X}}(\mathbf{t}) = E[e^{\mathbf{t}^T \mathbf{X}}] = \exp \left(\boldsymbol{\mu}^T \mathbf{t} + \frac{1}{2} \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t} \right)$$

MGF of MVN: A Linear Combination

- To find the distribution of \mathbf{Y} , we derive the MGF of \mathbf{Y} .

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= E[e^{\mathbf{t}^T \mathbf{A} \mathbf{X}}] = M_{\mathbf{X}}(\mathbf{A}^T \mathbf{t}) = \exp \left(\boldsymbol{\mu}^T \mathbf{A}^T \mathbf{t} + \frac{1}{2} \mathbf{t}^T \mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T \mathbf{t} \right) \\ &= \exp \left((\mathbf{A} \boldsymbol{\mu})^T \mathbf{t} + \frac{1}{2} \mathbf{t}^T (\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T) \mathbf{t} \right), \end{aligned}$$

- $M_{\mathbf{Y}}(\mathbf{t})$ has the form of the MGF of a multivariate normal distribution with mean vector $\mathbf{A} \boldsymbol{\mu}$ and covariance matrix $\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T$.
- As a result, the linear combination

$$\mathbf{Y} = \mathbf{A} \mathbf{X} \sim N(\mathbf{A} \boldsymbol{\mu}, \mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T)$$

MGF of MVN: The Sample Mean Vector

- Let $\mathbf{X}_1, \dots, \mathbf{X}_n$ be a random sample from $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- We have defined then the sample mean vector $\bar{\mathbf{X}}$
- We have shown that
 - $\mathbb{E}[\bar{\mathbf{X}}] = \boldsymbol{\mu}$
 - $\text{Cov}[\bar{\mathbf{X}}] = \frac{\boldsymbol{\Sigma}}{n}$
- Next, we will show that it follows a multivariate normal distribution.

MGF of MVN: The Sample Mean Vector

- We first calculate its MGF:

$$\begin{aligned}M_{\bar{\mathbf{X}}}(\mathbf{t}) &\stackrel{iid}{=} \prod_{i=1}^n M_{\mathbf{X}_i}\left(\frac{1}{n}\mathbf{t}\right) = \left(\exp\left(\boldsymbol{\mu}^T \frac{1}{n}\mathbf{t} + \frac{1}{2}\left(\frac{1}{n}\mathbf{t}\right)^T \boldsymbol{\Sigma} \left(\frac{1}{n}\mathbf{t}\right)\right)\right)^n \\&= \exp\left(n\left(\boldsymbol{\mu}^T \frac{1}{n}\mathbf{t} + \frac{1}{2n^2}\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}\right)\right) \\&= \exp\left(\boldsymbol{\mu}^T \mathbf{t} + \frac{1}{2}\mathbf{t}^T \frac{1}{n}\boldsymbol{\Sigma} \mathbf{t}\right)\end{aligned}$$

- $M_{\bar{\mathbf{X}}}(\mathbf{t})$ has the form of the MGF of a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\frac{1}{n}\boldsymbol{\Sigma}$, i.e., $\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \frac{\boldsymbol{\Sigma}}{n})$

Section 4

Zero-Covariance

Independence of Normals Under Jointly Normal

- In general, zero-correlation does not guarantee independence
- Independence of normals under jointly normal: If the joint distribution of \mathbf{X}_1 (a $p \times 1$ random vector) and \mathbf{X}_2 is jointly/multivariate normal, i.e.,

$$\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right),$$

then $\mathbf{X}_1 \perp \mathbf{X}_2 \Leftrightarrow \Sigma_{12} = 0$

- Proof: omitted. A result about MGF can be used to prove independence.

The Joint Distribution of Two Linear Functions

- Let $\mathbf{X}_{p \times 1} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
- Let $\mathbf{Y} = \mathbf{A}\mathbf{X}$ and $\mathbf{Z} = \mathbf{B}\mathbf{X}$.
- What is the joint distribution of \mathbf{Y} and \mathbf{Z} ?
- Note that

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \mathbf{X}$$

- A linear combination of MVN random vector also follows a MVN

$$\begin{aligned} \begin{pmatrix} \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} &\sim N\left(\begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \boldsymbol{\mu}, \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \boldsymbol{\Sigma} \begin{pmatrix} \mathbf{A}^T & \mathbf{B}^T \end{pmatrix}\right) \\ &\sim N\left(\begin{pmatrix} \mathbf{A}\boldsymbol{\mu} \\ \mathbf{B}\boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T & \mathbf{A}\boldsymbol{\Sigma}\mathbf{B}^T \\ \mathbf{B}\boldsymbol{\Sigma}\mathbf{A}^T & \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T \end{pmatrix}\right) \end{aligned}$$

We have

$$\mathbf{Y} \perp \mathbf{Z} \Leftrightarrow \mathbf{A}\boldsymbol{\Sigma}\mathbf{B}^T = 0$$

A Simulation Question

- Suppose we have a random sample from a normal distribution.
- How to use a simulation to show that sample mean and sample variance are uncorrelated (in fact they are also independent)?

Subsection 1

Sample Mean and Sample Variance

The Independence Between Sample Mean and Sample Variance

- For a random sample from a normal distribution, the sample mean and sample variance are independent.
- Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a normal distribution with mean μ and variance σ^2 .
- The sample mean and sample variance are defined as:
 - Sample mean: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
 - Sample variance: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
- We want to show that \bar{X} and s^2 are independent.

Proof

- We first rewrite the sample mean:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \mathbf{1}^T \mathbf{X}$$

- We have shown that $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \mathbf{X}^T \mathbb{C} \mathbf{X}$, where $\mathbb{C} = \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T$. In addition, it is easy to verify that $\mathbb{C} = \mathbb{C}^T$, $\mathbb{C}^2 = \mathbb{C}$. Thus, the distribution of s^2 can be rewritten to

$$s^2 = \frac{1}{n-1} (\mathbb{C} \mathbf{X})^T (\mathbb{C} \mathbf{X}),$$

which indicates that the distribution s^2 is determined by the distribution of $\mathbb{C} \mathbf{X}$.

Proof (continued)

- Clearly $\bar{\mathbf{X}}$ and $\mathbb{C}\mathbf{X}$ are linear combinations of \mathbf{X} , which follows a multivariate normal with covariance $\Sigma = \mathbf{I}$. Thus,

$$\text{cov}(\bar{\mathbf{X}}, \mathbb{C}) = \frac{1}{n} \mathbf{1}^T \Sigma \mathbb{C} = \frac{1}{n} \mathbf{1}^T \mathbb{C} = 0$$

Please verify that last step.

By Theorem on “Independence of Normals Under Jointly Normal”, we can conclude the $\bar{\mathbf{X}}$ and s^2 are independent.

Section 5

MVN: $\bar{\mathbf{X}}$ and S

The Independence Between Sample Mean Vector and Sample Covariance Matrix

- How to prove that the sample mean vector and the sample covariance matrix are independent
- Messier way: vectorize the $n \times p$ matrix \mathbf{X} to a $(np) \times 1$ vector and then apply the condition for independent linear combinations under MVN
- Neater way: use properties of Matrix Normal Distribution

Sample Mean Vector and Sample Covariance Matrix

- If we have a random sample from MVN, we will show that $\bar{\mathbf{X}}$ and S are independent
- Proof outline
 - 1 Vectorize $\mathbf{X}_{n \times p}$ to a vector $\tilde{\mathbf{X}}_{(np) \times 1}$, which follows a MVN
 - 2 Show that the distribution of $\tilde{\mathbf{X}}$ is determined by a linear function of $\tilde{\mathbf{X}}_{(np) \times 1}$
 - 3 Show that the distribution of S is determined by a linear function of $\tilde{\mathbf{X}}_{(np) \times 1}$
 - 4 Find the covariance of the two linear functions
 - 5 Conclude that the two linear functions are independent, which indicates that the sample mean vector and the sample covariance matrix are independent

Step 1a: Vectorize

- We vectorize $\mathbf{X}_{n \times p}$ such that
 - the first n random variables are for the first feature
 -
 - the last n random variables are for the last feature

$$\tilde{\mathbf{X}}_{(np) \times 1} = \begin{pmatrix} \mathbf{X}_{(1)} \\ \vdots \\ \mathbf{X}_{(p)} \end{pmatrix}$$

- What is the distribution of $\mathbf{X}_{(1)}$?

- The distribution of $\mathbf{X}_{(1)}$?

$$\mathbf{X}_{(1)} \sim N\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \sigma_{11}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{11}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{11}^2 \end{pmatrix}\right) \sim N(\mu_1 \mathbf{1}_n, \sigma_{11}^2 \mathbf{I})$$

Step 1b: The distribution of $\mathbf{x}_{n \times p}$

$$\begin{aligned}\tilde{\mathbf{X}}_{(np) \times 1} &= \begin{pmatrix} \mathbf{X}_{(1)} \\ \vdots \\ \mathbf{X}_{(p)} \end{pmatrix} \sim N(\boldsymbol{\mu} \otimes \mathbf{1}_n, \boldsymbol{\Sigma} \otimes \mathbf{I}_n) \\ &\sim N\left(\begin{pmatrix} \mu_1 \mathbf{1}_n \\ \vdots \\ \mu_p \mathbf{1}_n \end{pmatrix}, \begin{pmatrix} \sigma_{11} \mathbf{I}_n & \cdots & \sigma_{1p} \mathbf{I}_n \\ \cdots & \cdots & \cdots \\ \sigma_{p1} \mathbf{I}_n & \cdots & \sigma_{pp} \mathbf{I}_n \end{pmatrix}\right)\end{aligned}$$

Step 2: The sample mean vector

- The sample mean vector can be written as linear functions of $\tilde{\mathbf{X}}$:

$$\bar{\mathbf{X}} = \frac{1}{n} \begin{pmatrix} \mathbf{1}_n^T & \mathbf{0}_n^T & \cdots & \mathbf{0}_n^T \\ \mathbf{0}_n^T & \mathbf{1}_n^T & \cdots & \mathbf{0}_n^T \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_n^T & \mathbf{0}_n^T & \cdots & \mathbf{1}_n^T \end{pmatrix} \tilde{\mathbf{X}}$$

Step 3: The sample covariance matrix

- Recall that we have shown the following result

$$\mathbf{S} = \frac{1}{n-1} \mathbf{X}^T \mathbb{C} \mathbf{X} = \frac{1}{n-1} (\mathbb{C} \mathbf{X})^T \mathbb{C} \mathbf{X}$$

- So we only need to focus on $\mathbb{C} \mathbf{X}$, the centered random matrix.
- The vectorized version of the centered random matrix is

$$\text{vec}(\mathbb{C} \mathbf{X}) = \begin{pmatrix} \mathbb{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbb{C} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbb{C} \end{pmatrix} \tilde{\mathbf{X}}_{(np) \times (np)}$$

Step 4: The covariance of the two linear functions

So we have the following results -

$$\tilde{\mathbf{X}}_{(np) \times 1} \sim N(\boldsymbol{\mu} \otimes \mathbf{1}_n, \boldsymbol{\Sigma} \otimes \mathbf{I}_n) \sim N\left(\begin{pmatrix} \mu_1 \mathbf{1}_n \\ \vdots \\ \mu_p \mathbf{1}_n \end{pmatrix}, \begin{pmatrix} \sigma_{11} \mathbf{I}_n & \cdots & \sigma_{1p} \mathbf{I}_n \\ \cdots & \cdots & \cdots \\ \sigma_{p1} \mathbf{I}_n & \cdots & \sigma_{pp} \mathbf{I}_n \end{pmatrix}\right)$$

- The distribution of $\bar{\mathbf{X}}$ and S depend on

$$\bar{\mathbf{X}} = \frac{1}{n} \begin{pmatrix} \mathbf{1}_n^T & \mathbf{0}_n^T & \cdots & \mathbf{0}_n^T \\ \mathbf{0}_n^T & \mathbf{1}_n^T & \cdots & \mathbf{0}_n^T \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0}_n^T & \mathbf{0}_n^T & \cdots & \mathbf{1}_n^T \end{pmatrix} \tilde{\mathbf{X}}, \text{vec}(\mathbb{C}\mathbf{X}) = \begin{pmatrix} \mathbb{C} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbb{C} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbb{C} \end{pmatrix} \tilde{\mathbf{X}}$$

Step 4: The covariance of the two linear functions

- Let $\tilde{\Sigma}$ denote the covariance matrix of $\tilde{\mathbf{x}}$
- Let \mathbb{A} denote the matrix such that $\bar{\mathbf{X}} = \mathbb{A}\tilde{\mathbf{x}}$
- Let \mathbb{B} denote the matrix such that $\text{vec}(\mathbf{C}\mathbf{X}) = \mathbb{B}\tilde{\mathbf{X}}$
- It can be verified that $\mathbb{A}\tilde{\Sigma}\mathbb{B}^T = \mathbf{0}$.

Step 5: The independence of the two linear functions

- Both $\bar{\mathbf{X}}$ and $\text{vec}(\mathbb{C}\mathbf{X})$ are linear function of the same MVN-distributed random vector $\tilde{\mathbf{X}}$
- Their covariance matrix is zero, which indicates that they are independent by Theorem on “Independence of Normals Under Jointly Normal”.
- The sample covariance matrix only depends on the centered data, $\text{vec}(\mathbb{C}\mathbf{X})$ (the vector form) up to a constant
- Therefore, if we have a random sample from a **MVN**, the sample mean vector and the sample covariance matrix are independent
- The proof is lengthy. It can be more compact if we introduce matrix normal distribution.