Multivariate Analysis
Chapter 2: Sample
Chapter 3: Random Matrix

Shaobo Jin

Department of Mathematics

Intended Learning Outcome

Through this chapter, you should be able to

1. apply properties of random vector and random matrix,
2. understand random sample,
3. understand sample mean and sample covariance matrix,
4. apply properties of sample mean and sample covariance matrix.

Expectation

The expected value of a random vector/matrix is the vector/matrix consisting of the expected values of each of its elements:

$$\mathbb{E}\left(\boldsymbol{X}\right) = \begin{bmatrix} \mathbb{E}\left(X_{11}\right) & \mathbb{E}\left(X_{12}\right) & \cdots & \mathbb{E}\left(X_{1p}\right) \\ \mathbb{E}\left(X_{21}\right) & \mathbb{E}\left(X_{22}\right) & \cdots & \mathbb{E}\left(X_{2p}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}\left(X_{n1}\right) & \mathbb{E}\left(X_{n2}\right) & \cdots & \mathbb{E}\left(X_{np}\right) \end{bmatrix}.$$

Covariance Matrix

For a $p \times 1$ random vector $\boldsymbol{X}$ with mean $\boldsymbol{\mu}_X$ and a $q \times 1$ random vector $\boldsymbol{Y}$ with mean $\boldsymbol{\mu}_Y$, its covariance matrix is

$$\begin{aligned} \operatorname{cov}(\boldsymbol{X}, \boldsymbol{Y}) &= \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu}_X)(\boldsymbol{Y} - \boldsymbol{\mu}_Y)^T\right] \\ &= \mathbb{E}\left(\boldsymbol{X}\boldsymbol{Y}^T\right) - \boldsymbol{\mu}_X \boldsymbol{\mu}_Y^T, \end{aligned}$$

where its (i,k)th element is $\mathbb{E}\left[(X_i - \mu_{X,i})(Y_k - \mu_{Y,k})\right]$.

Covariance Matrix

For a $p \times 1$ random vector $\boldsymbol{X}$ with mean $\boldsymbol{\mu}$, its (variance-) covariance matrix is

$$\begin{aligned} \operatorname{var}(\boldsymbol{X}) = \operatorname{cov}(\boldsymbol{X}, \boldsymbol{X}) = \boldsymbol{\Sigma} &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \\ &= \mathbb{E}\left[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T\right] \\ &= \mathbb{E}\left(\boldsymbol{X}\boldsymbol{X}^T\right) - \boldsymbol{\mu}\boldsymbol{\mu}^T, \end{aligned}$$

where its (i,k)th element is

$$\sigma_{ik} = \mathbb{E}\left[(X_i - \mu_i)(X_k - \mu_k)\right].$$

- $\boldsymbol{\Sigma}$ is symmetric, i.e., $\sigma_{ik} = \sigma_{ki}$.
- $\boldsymbol{\Sigma}$ is positive semi-definite.

Linear Combination

A linear combination of p variables is

$$\boldsymbol{c}^T \boldsymbol{X} = c_1 X_1 + c_2 X_2 + \cdots + c_p X_p,$$

where $\boldsymbol{c}$ is a vector of fixed (not random) values and $\boldsymbol{X}$ is a $p \times 1$ random vector.

A linear combination of can also be

$$\boldsymbol{C}\boldsymbol{X} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^p c_{1j} X_j \\ \sum_{j=1}^p c_{2j} X_j \\ \vdots \\ \sum_{j=1}^p c_{qj} X_j \end{bmatrix},$$

where $\boldsymbol{C}$ is a $q \times p$ matrix of fixed (not random) values and $\boldsymbol{X}$ is a $p \times 1$ random vector.

Linear Combination

1. Random variable X, constants c and d:
$$\begin{aligned} \mathbb{E}\left(cX+d\right) &= c\mathbb{E}\left(X\right)+d, \\ \text{var}\left(cX+d\right) &= c^2\text{var}\left(X\right). \end{aligned}$$
2. Random vector $\boldsymbol{X}$, constant vector $\boldsymbol{c}$, constant d:
$$\begin{aligned} \mathbb{E}\left(\boldsymbol{c}^T\boldsymbol{X}+d\right) &= \boldsymbol{c}^T\mathbb{E}\left(\boldsymbol{X}\right)+d, \\ \text{var}\left(\boldsymbol{c}^T\boldsymbol{X}+d\right) &= \boldsymbol{c}^T\text{var}\left(\boldsymbol{X}\right)\boldsymbol{c}. \end{aligned}$$
3. Random vector $\boldsymbol{X}$, constant matrix $\boldsymbol{C}$, constant vector $\boldsymbol{d}$:
$$\begin{aligned} \mathbb{E}\left(\boldsymbol{C}\boldsymbol{X}+\boldsymbol{d}\right) &= \boldsymbol{C}\mathbb{E}\left(\boldsymbol{X}\right)+\boldsymbol{d}, \\ \text{var}\left(\boldsymbol{C}\boldsymbol{X}+\boldsymbol{d}\right) &= \boldsymbol{C}\text{var}\left(\boldsymbol{X}\right)\boldsymbol{C}^T. \end{aligned}$$
4. Random matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, and constant matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{D}$
$$\begin{aligned} \mathbb{E}\left(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}+\boldsymbol{D}\right) &= \boldsymbol{A}\mathbb{E}\left(\boldsymbol{X}\right)\boldsymbol{B}+\boldsymbol{D}, \\ \text{cov}\left(\boldsymbol{A}\boldsymbol{X},\boldsymbol{B}\boldsymbol{Y}\right) &= \boldsymbol{A}\text{cov}\left(\boldsymbol{X},\boldsymbol{Y}\right)\boldsymbol{B}^T. \end{aligned}$$

Additive and Scaling Property

1. Random variables X, Y, and Z, constant a:

$$\begin{aligned} \mathbb{E}(X+Y) &= \mathbb{E}(X)+\mathbb{E}(Y), \\ \operatorname{cov}(X+Y,Z) &= \operatorname{cov}(X,Z)+\operatorname{cov}(Y,Z), \\ \operatorname{cov}(aX,Z) &= a\operatorname{cov}(X,Z). \end{aligned}$$

2. Random vectors $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$, constant matrix $\boldsymbol{A}$:

$$\begin{aligned} \mathbb{E}(\boldsymbol{X}+\boldsymbol{Y}) &= \mathbb{E}(\boldsymbol{X})+\mathbb{E}(\boldsymbol{Y}), \\ \operatorname{cov}(\boldsymbol{X}+\boldsymbol{Y},\boldsymbol{Z}) &= \operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})+\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z}), \\ \operatorname{cov}(\boldsymbol{A}\boldsymbol{X},\boldsymbol{Z}) &= \boldsymbol{A}\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z}), \\ \operatorname{cov}(\boldsymbol{X},\boldsymbol{A}\boldsymbol{Z}) &= \operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})\boldsymbol{A}^T. \end{aligned}$$

3. Random matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$\mathbb{E}(\boldsymbol{X}+\boldsymbol{Y}) = \mathbb{E}(\boldsymbol{X})+\mathbb{E}(\boldsymbol{Y}).$$

Independence

1. Independent random variables X and Y:

$$\begin{aligned} \mathbb{E}(XY) &= \mathbb{E}(X)\,\mathbb{E}(Y), \\ \operatorname{var}(X+Y) &= \operatorname{var}(X) + \operatorname{var}(Y), \\ \operatorname{cov}(X,Y) &= \mathbf{0}. \end{aligned}$$

2. Independent random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$\begin{aligned} \operatorname{var}(\boldsymbol{X}+\boldsymbol{Y}) &= \operatorname{var}(\boldsymbol{X}) + \operatorname{var}(\boldsymbol{Y}), \\ \operatorname{cov}(\boldsymbol{X},\boldsymbol{Y}) &= \mathbf{0}. \end{aligned}$$

3. Independent random matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$:

$$\mathbb{E}(\boldsymbol{X}\boldsymbol{Y}) = \mathbb{E}(\boldsymbol{X})\,\mathbb{E}(\boldsymbol{Y}).$$

Random Sample

$$\boldsymbol{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_1^T \\ \boldsymbol{X}_2^T \\ \vdots \\ \boldsymbol{X}_n^T \end{bmatrix}.$$

If the row vectors represent independent observations from a common joint distribution $f(\boldsymbol{x})$, then $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$ form a random sample from $f(\boldsymbol{x})$.

If $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$ form a random sample, then their joint density is given by

$$f(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n) = f(\boldsymbol{x}_1) f(\boldsymbol{x}_2) \cdots f(\boldsymbol{x}_n).$$

Sample Mean and Covariance Matrix

Result 2.1

Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, ..., $\boldsymbol{X}_n$ be a random sample from a joint distribution that has mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then $\bar{\boldsymbol{X}}$ is an unbiased estimator of $\boldsymbol{\mu}$ and its covariance matrix is $n^{-1}\boldsymbol{\Sigma}$. That is,

$$\mathbb{E}\left(\bar{\boldsymbol{X}}\right) = \boldsymbol{\mu} \quad \text{and} \quad \operatorname{var}\left(\bar{\boldsymbol{X}}\right) = n^{-1}\boldsymbol{\Sigma}.$$

But for the sample covariance matrix

$$\mathbb{E}\left(\boldsymbol{S}_n\right) = \frac{n-1}{n}\boldsymbol{\Sigma}.$$

Thus,

$$\boldsymbol{S} = \frac{n}{n-1}\boldsymbol{S}_n$$

is an unbiased estimator of $\boldsymbol{\Sigma}$.

Some Notes on Sample Covariance Matrix

- The biased estimator $\boldsymbol{S}_n$ uses n^{-1}, and the unbiased estimator $\boldsymbol{S}$ uses $(n-1)^{-1}$.
- Even though $\boldsymbol{S}$ is an unbiased estimator of $\boldsymbol{\Sigma}$, $\sqrt{s_{ii}}$ is a biased estimator of $\sqrt{\sigma_{ii}}$.
- Sometimes, it has to be $\boldsymbol{S}$. But for a large enough n, the difference between $\boldsymbol{S}_n$ and $\boldsymbol{S}$ can often be ignored.

Descriptive Statistics: Sample Mean

The sample mean of variable X_k, for $k = 1, 2, ..., p$, is

$$\bar{x}_k = \frac{1}{n}\sum_{j=1}^{n} x_{jk}.$$

The sample mean is a $p \times 1$ vector

$$\bar{\boldsymbol{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{x}_j$$

Sample Mean

For the sample mean,

$$
\begin{aligned}
\bar{\boldsymbol{x}} &= \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} n^{-1}\sum_{j=1}^{n} x_{j1} \\ n^{-1}\sum_{j=1}^{n} x_{j2} \\ \vdots \\ n^{-1}\sum_{j=1}^{n} x_{jp} \end{bmatrix} = \frac{1}{n} \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \frac{1}{n} \boldsymbol{X}^T \mathbf{1},
\end{aligned}
$$

where $\mathbf{1}$ is a $n \times 1$ vector of ones.

Descriptive Statistics: Sample Covariance Matrix

The sample variance of variable X_k is

$$s_{kk} = \frac{1}{n}\sum_{j=1}^{n}\left(x_{jk}-\bar{x}_k\right)^2.$$

The sample covariance between X_i and X_k is

$$s_{ik} = \frac{1}{n}\sum_{j=1}^{n}\left(x_{ji}-\bar{x}_i\right)\left(x_{jk}-\bar{x}_k\right).$$

The sample covariance matrix is

$$\boldsymbol{S}_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix} = \frac{1}{n}\sum_{j=1}^{n}\left(\boldsymbol{x}_j-\bar{\boldsymbol{x}}\right)\left(\boldsymbol{x}_j-\bar{\boldsymbol{x}}\right)^T,$$

which must be symmetric and positive (semi-) definite.

Demean Variables

If we demean each variable, then we have

$$
\begin{aligned}
\begin{bmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1p}-\bar{x}_p \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2p}-\bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{np}-\bar{x}_p \end{bmatrix} &= \boldsymbol{X} - \begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_1 & \bar{x}_2 & \cdots & \bar{x}_p \end{bmatrix} \\
&= \boldsymbol{X} - \mathbf{1}\bar{\boldsymbol{x}}^T \\
&= \boldsymbol{X} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\boldsymbol{X}.
\end{aligned}
$$

Sample Covariance Matrix

Then,

$$
\begin{aligned}
&(n-1)\,\boldsymbol{S} = n\boldsymbol{S}_n \\
=&\begin{bmatrix} x_{11}-\bar{x}_1 & x_{21}-\bar{x}_1 & \cdots & x_{n1}-\bar{x}_1 \\ x_{12}-\bar{x}_2 & x_{22}-\bar{x}_2 & \cdots & x_{n2}-\bar{x}_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p}-\bar{x}_p & x_{2p}-\bar{x}_p & \cdots & x_{np}-\bar{x}_p \end{bmatrix}^T \begin{bmatrix} x_{11}-\bar{x}_1 & x_{12}-\bar{x}_2 & \cdots & x_{1p}-\bar{x}_p \\ x_{21}-\bar{x}_1 & x_{22}-\bar{x}_2 & \cdots & x_{2p}-\bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}-\bar{x}_1 & x_{n2}-\bar{x}_2 & \cdots & x_{np}-\bar{x}_p \end{bmatrix} \\
=&\left(\boldsymbol{X}-\frac{1}{n}\mathbf{1}\mathbf{1}^T\boldsymbol{X}\right)^T\left(\boldsymbol{X}-\frac{1}{n}\mathbf{1}\mathbf{1}^T\boldsymbol{X}\right) \\
=&\boldsymbol{X}^T\left(\boldsymbol{I}-\frac{1}{n}\mathbf{1}\mathbf{1}^T\right)\boldsymbol{X},
\end{aligned}
$$

where $\boldsymbol{I}$ is a $p \times p$ identity matrix.

Sample (Pearson) Correlation

The sample (Pearson) correlation coefficient between X_i and X_k is

$$r_{ik} = \frac{s_{ik}}{\sqrt{s_{ii}}\sqrt{s_{kk}}}.$$

Let

$$\boldsymbol{D}^{1/2} = \begin{bmatrix} \sqrt{s_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{s_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{s_{pp}} \end{bmatrix}.$$

Then,

$$\boldsymbol{R} = \boldsymbol{D}^{-1/2}\boldsymbol{S}\boldsymbol{D}^{-1/2}$$

is the sample (Pearson) correlation matrix.

Linear Combination(s)

Result 2.5

Let $\boldsymbol{X}$ be a random vector. Consider the linear combinations $\boldsymbol{b}^T\boldsymbol{X}$ and $\boldsymbol{c}^T\boldsymbol{X}$. Then,

$$\begin{aligned} \text{sample mean of } \boldsymbol{c}^T\boldsymbol{X} &= \boldsymbol{c}^T\bar{\boldsymbol{x}} \\ \text{sample variance of } \boldsymbol{c}^T\boldsymbol{X} &= \boldsymbol{c}^T\boldsymbol{S}\boldsymbol{c} \\ \text{sample covariance between } \boldsymbol{b}^T\boldsymbol{X} \text{ and } \boldsymbol{c}^T\boldsymbol{X} &= \boldsymbol{b}^T\boldsymbol{S}\boldsymbol{c}. \end{aligned}$$

Let $\boldsymbol{X}$ be a random vector. The linear combinations $\boldsymbol{A}\boldsymbol{X}$ have sample mean vector $\boldsymbol{A}\bar{\boldsymbol{x}}$ and sample covariance matrix is $\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^T$.