STAT-445/645: Applied Multivariate Analysis

K. Ken Peng

Department of Statistics and Actuarial Science
Simon Fraser University

Spring 2026

What to do today?

Part I. Introduction and Preparation

Part I.2. Review on Matrix Algebra (Chp 2.1-4, Supplement 2A)

Part I.3. Introduction to R

Part I.4. Multivariate Random Variables and Distributions (Chp 1, 2.5-6, 3)

I.2.1. Why do we need matrix/vector algebra in STAT445/645?

Simple linear regression:

$$Y = \beta_0 + \beta X + \epsilon, \quad E(\epsilon) = 0, \;\; V(\epsilon) = \sigma^2$$

with observations $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ from n independent units. That is, for $i = 1, \ldots, n$,

$$y_i = \beta_0 + \beta x_i + \epsilon_i, \quad E(\epsilon_i) = 0, \;\; V(\epsilon_i) = \sigma^2$$

with $\epsilon_1, \ldots, \epsilon_n$ indpt.

- LSE: $\hat{\beta} = \frac{S_{XY}}{S_{XX}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$; $\hat{\beta}_0 = \bar{y} - \hat{\beta}\bar{x}$.
- $V(\hat{\beta}) = \sigma^2 / S_{XX}$ and $V(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{XX}}\right]$;
- $\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

How about to consider

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon, \quad E(\epsilon) = 0, \;\; V(\epsilon) = \sigma^2$$

with observations $\{(x_{11}, x_{21}, y_1), \ldots, (x_{1n}, x_{2n}, y_n)\}$ from n independent units?
That is, for $i = 1, \ldots, n$,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad E(\epsilon_i) = 0, \;\; V(\epsilon_i) = \sigma^2$$

with $\epsilon_1, \ldots, \epsilon_n$ indpt.

- LSE for $\beta_0, \beta_1, \beta_2$?
- How about the estimators' variances?
- $\hat{\sigma}^2 = \frac{1}{n-3} \sum_{i=1}^{n} e_i^2$ with $e_i = y_i - \hat{y}_i = y_i - [\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}]$.

What if

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \epsilon, \quad E(\epsilon) = 0, \ V(\epsilon) = \sigma^2$$

with observations $\{(x_{11}, \ldots, x_{p1}, y_1), \ldots, (x_{1n}, \ldots, x_{pn}, y_n)\}$ from n independent units?

What if to study how $(Y_1, \ldots, Y_k)$ depends on $(X_1, \ldots, X_p)$?

$\Longrightarrow$ **vector/matrix algebra** as a tool for communication in general, together with software packages such as R and SAS to conduct the required computing.

I.2.2. Notation and Basic Definitions

- a real number; a **scalar**; a physical quantity

  e.g. $a = 3.6, \quad b = -7.21, \quad x = 5, \quad y = \sqrt{8}$

- A **vector** is a group of p numbers/elements arranged in a *column*: a p-dim vector.

$$\mathbf{a} = \begin{pmatrix} 1.0 \\ 2.3 \\ 4.7 \end{pmatrix}, \quad \mathbf{b} = \begin{bmatrix} -.6 \\ 5.9 \end{bmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

- A $p \times q$ **matrix** is a group of pq numbers/elements arranged into a rectangular array with p *rows* and q *columns*.

$$\mathbf{A} = \begin{pmatrix} 1.0 & 8 \\ 2.3 & 6 \\ 4.7 & -9.5 \end{pmatrix}, \; \mathbf{B} = \begin{bmatrix} 5.9 & b_1 & b_2 \end{bmatrix}, \; \mathbf{X} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \; \mathbf{Y} = \begin{pmatrix} y_{11} & \cdots & y_{1q} \\ \vdots & \vdots & \vdots \\ y_{p1} & \cdots & y_{pq} \end{pmatrix}$$

- An $p \times q$ matrix is called a *square* matrix if $p = q$.

- An $p \times q$ matrix is a *row* vector if $p = 1$; a *column* vector if $q = 1$; a scalar if $p = q = 1$.

- Two matrices $\mathbf{A} = \left(a_{ij}\right)$ and $\mathbf{B} = \left(b_{ij}\right)$ are the same iff $a_{ij} = b_{ij}$: $\mathbf{A} = \mathbf{B}$.

- A square matrix $\mathbf{A} = \left(a_{ij}\right)$ is *diagonal* if all its off-diagonal elements are zero: $a_{ij} = 0$ if $i \neq j$.
  e.g. $\mathbf{A} = \begin{pmatrix} 1.3 & 0 \\ 0 & 6 \end{pmatrix}$, denoted by $\mathbf{A} = diag(1.3, 6)$.
  $\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3.5 \end{pmatrix}$, denoted by $\mathbf{B} = diag(1, 6, 3.5)$.

- Two important matrices: the *identity* matrix $\mathbf{I} = diag(1, \ldots, 1)$; the *zero* matrix $\mathbf{0} = diag(0, \ldots, 0)$.

I.2.3. Vector Operations

- **addition**. The sum of two p-dim vectors $\mathbf{a}$ and $\mathbf{b}$ is a new p-dim vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$:

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_p + b_p \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix} = \mathbf{c}$$

  e.g.

$$\mathbf{x} = \begin{pmatrix} 1 \\ 3.3 \\ -.5 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 6.7 \\ -1.2 \\ 9 \end{pmatrix}, \quad \mathbf{x} + \mathbf{y} = \begin{pmatrix} 7.7 \\ 2.1 \\ 8.5 \end{pmatrix}.$$

- **scalar multiplication**. If a is a scalar and $\mathbf{x}$ is a p-dim vector with components (entries) x_i, the *product* $a\mathbf{x}$ is a new p-dim vector with components ax_i.
  e.g.

$$a = 10, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3.3 \\ -.5 \end{pmatrix}, \quad a\mathbf{x} = \begin{pmatrix} 10 \\ 33 \\ -5 \end{pmatrix}$$

- **subtraction**. The subtraction of two p-dim vectors $\mathbf{a}$ and $\mathbf{b}$ is a new p-dim vector $\mathbf{c} = \mathbf{a} - \mathbf{b}$, which is $\mathbf{a} + (-1)\mathbf{b}$.
  e.g.
  $$\mathbf{x} = \begin{pmatrix} 1 \\ 3.3 \\ -.5 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 6.7 \\ -1.2 \\ 9 \end{pmatrix}, \quad \mathbf{x} - \mathbf{y} = \begin{pmatrix} -5.7 \\ 4.5 \\ -9.5 \end{pmatrix}.$$
- The vector $\mathbf{y} = a_1\mathbf{x}_1 + \ldots + a_k\mathbf{x}_k$ is a *linear combination* of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$.
  - If there exist k numbers $c_1, \ldots, c_k$, not all zero, such that $c_1\mathbf{x}_1 + \ldots + c_k\mathbf{x}_k = \mathbf{0}$, the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are *linearly dependent*.
    Otherwise the set of vectors are *linearly independent* (Iff No vector in the set can be built linearly using the others).
  - Every p-dim vector can be expressed as
    $$\mathbf{a} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots a_p \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = a_1\mathbf{e}_1 + \ldots + a_p\mathbf{e}_p.$$
    The set linearly independent vectors $\mathbf{e}_1, \ldots, \mathbf{e}_p$ is a *basis* for the p-dim vector space, and $a_1, \ldots, a_p$ are the coordinates of $\mathbf{a}$.

- The **inner product** of two p-dim vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$x_1y_1 + x_2y_2 + \ldots + x_py_p,$$

denoted by $\mathbf{x}^{'}\mathbf{y} = \mathbf{y}^{'}\mathbf{x}$.

- The **length** of a p-dim vector $\mathbf{x}$ is

$$||\mathbf{x}|| = \sqrt{x_1^2 + \ldots + x_p^2}.$$

It is $L_{\mathbf{x}} = \sqrt{\mathbf{x}'\mathbf{x}}$.

- The **angle** θ between two p-dim vectors $\mathbf{x}$ and $\mathbf{y}$ is determined from

$$cos(\theta) = \frac{\mathbf{x}^{'}\mathbf{y}}{||\mathbf{x}||||\mathbf{y}||} = \frac{\mathbf{x}^{'}\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}.$$

Geometric interpretation

- a 2-dim vector $\mathbf{x}$; $||\mathbf{x}||$; a scalar multiplication $a\mathbf{x}$.
- addition of two 2-dim vectors $\mathbf{x}$ and $\mathbf{y}$; inner product of them

I.2.4. Matrix Operations

- The **transpose** of $\mathbf{A} = \left(a_{ij}\right)$ is $\mathbf{A}' = \mathbf{B} = \left(b_{ij}\right)$ with $b_{ij} = a_{ji}$.
- **scalar multiplication**. Let c be a scalar and $\mathbf{A} = \left(a_{ij}\right)$. Then $c\mathbf{A} = \left(b_{ij}\right)$ with $b_{ij} = ca_{ij}$.
- The **addition** of $\mathbf{A} = \left(a_{ij}\right)$ and $\mathbf{B} = \left(b_{ij}\right)$ is $\mathbf{A} + \mathbf{B} = \mathbf{C} = \left(c_{ij}\right)$ with $c_{ij} = a_{ij} + b_{ij}$.
  - **subtraction**. $\mathbf{A} - \mathbf{B} = \mathbf{C} = \left(c_{ij}\right)$ with $c_{ij} = a_{ij} - b_{ij}$.

Properties. Associative property; distributive property; commutative property

- **matrix multiplication**. The **product** of $p \times q$ matrix $\mathbf{A} = \left(a_{ij}\right)$ and $q \times k$ matrix $\mathbf{B} = \left(b_{ij}\right)$ is $\mathbf{AB} = \mathbf{C} = \left(c_{ij}\right)$, a $p \times k$ matrix with $c_{ij} = \sum_{l=1}^{q} a_{il} b_{lj}$.

Properties. Associative; distributive over addition; not commutative (!); $(\mathbf{AB})^{'} = \mathbf{B}^{'}\mathbf{A}^{'}$

- The **inverse** of a square matrix $\mathbf{A} = \left(a_{ij}\right)$ is $\mathbf{B} = \mathbf{A}^{-1}$ such that $\mathbf{BA} = \mathbf{AB} = \mathbf{I}$, the identity matrix.
  - If $\mathbf{A}^{-1}$ exists, $\mathbf{A}$ is *invertible (nonsingular, full rank)*.
  - $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ if well-defined.
  - $(\mathbf{A}^{'})^{-1} = \left(\mathbf{A}^{-1}\right)^{'}$ if well-defined.
  - If $\mathbf{A}^{'}\mathbf{A} = \mathbf{I}$, $\mathbf{A}$ is **orthogonal**.
    Iff $\mathbf{A}^{'} = \mathbf{A}^{-1}$, $\mathbf{A}$ is **orthogonal**.

- The **determinant** of a square matrix $\mathbf{A} = \left(a_{ij}\right)_{k\times k}$ is denoted by $|\mathbf{A}|$ or $det(\mathbf{A})$:

$$\begin{array}{ll} |\mathbf{A}| = a_{11}, & k = 1 \\ |\mathbf{A}| = \sum_{j=1}^{k}(-1)^{1+j} a_{1j}|\mathbf{A}_{1j}|, & k > 1 \end{array}$$

  where $\mathbf{A}_{1j}$ is the $(k-1)\times(k-1)$ matrix obtained from $\mathbf{A}$ after deleting its first row and jth column.

e.g.

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 * |4|(-1)^2 + 2 * |3|(-1)^3 = -2$$

$$\begin{vmatrix} 3 & 1 & 6 \\ 7 & 4 & 5 \\ 2 & -7 & 1 \end{vmatrix} = 3*\begin{vmatrix} 4 & 5 \\ -7 & 1 \end{vmatrix}(-1)^2 + 1*\begin{vmatrix} 7 & 5 \\ 2 & 1 \end{vmatrix}(-1)^3 + 6*\begin{vmatrix} 7 & 4 \\ 2 & -7 \end{vmatrix}(-1)^4 = -222.$$

Property. $|\mathbf{A}| = |\mathbf{A}'|$; $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$; $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$; $|c\mathbf{A}| = c^k\,|\mathbf{A}|$.

- The **trace** of a square matrix $\mathbf{A} = \left(a_{ij}\right)_{k \times k}$ is $tr(\mathbf{A}) = \sum_{i=1}^{k} a_{ii}$.

Property.

- $tr(c\mathbf{A}) = ctr(\mathbf{A})$.
- $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$.
- $tr(\mathbf{AB}) = tr(\mathbf{BA})$.
- $tr(\mathbf{AA}') = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij}^2$.

Part I.3. Introduction to R

Part I.3. Introduction to R

Examples of using R:

- as a calculator ...
- use its functions and create functions ...
- conduct data analyses ...

What will we study in the next class?

- **Part I. Introduction and Preparation**
  - *I.1. General Introduction*
  - *I.2. Review on Matrix Algebra*
  - *I.3. Introduction to R*
  - I.4. Multivariate Random Variables and Distributions
- *Part II. Inference under Multivariate Normal Distribution (Textbook Chp 4-7)*
- *Part III. Commonly-Used Multivariate Analysis Methods (Textbook Chp 8-11)*
- *Part IV. Other Topics (Textbook Chp 12)*