Multivariate: Matrix algebra

Table of contents

1 Multivariate

1.1 Goals

1.1.1 Goals of this course

- Familiarize you with **classic** and **modern** multivariate statistics
- Make sure that you understand how to **conduct** these analyses and **interpret** the results
- Prepare you for **further study** in applied statistics
- Give you enough background to **understand** current applied statistics research

1.1.2 Goals of this lecture

- Introduce **matrix algebra**
- Review basic descriptive **statistics**
- Convert expressions for those statistics into **matrix format**

1.2 What is multivariate?

1.2.1 Multivariate?

Variate $\approx$ Variable

Univariate = one variable
Mean, variance

Bivariate = two variables
Correlation between two variables

Multivariate = multiple variables
How several outcomes are related to one another or to predictor(s)

1.2.2 Multivariate?

Multivariate data

- Dataset with many variables
- Typically across many participants
- Often across multiple time points
- "Data cube"

1.2.3 Multivariate?

Multivariate statistics

- Simultaneous analysis of many dependent variables or outcomes
- (There may also be many independent variables or predictors)
- Multivariate analysis is typically accompanied by univariate and bivariate analyses too: means, variances, correlations

1.3 Multivariate techniques

1.3.1 Classic multivariate techniques

Technique	Predictor (IV)	Outcome (DV)
t test	1 discrete, 2 levels	1
One-way ANOVA	1 discrete, >2 levels	1
Factorial ANOVA	≥ 2 discrete	1
Correlation	1 continuous	1
Regression	≥ 2 continuous	1
ANCOVA	Discrete, continuous	1

From Harris, R.N. (1985). A Primer on Multivariate Statistics

1.3.2 Classic multivariate techniques

Technique	Predictor (IV)	Outcome (DV)
Discriminant analysis	1 discrete, 2 levels	≥ 2
MANOVA	1 discrete, >2 levels	≥ 2
Canonical correlation	≥ 2 continuous	≥ 2
PCA	≥ 2 continuous	
FA	≥ 2 continuous	

From Harris, R.N. (1985). A Primer on Multivariate Statistics

1.3.3 Modern multivariate techniques

Models for one outcome variable

- Linear regression, logistic regression, Poisson regression

Models for multiple indicators of a construct

- Factor analysis (FA), principal components analysis (PCA), latent class / profile analysis (LCA / LPA)

Models for multiple outcomes

- Repeated measures ANOVA, MANOVA, mixed models, mediation, path analysis

2 Matrix algebra

2.1 Definitions

2.1.1 Matrix algebra?

Why do we start here?

- Statistics is applied math
- Matrices help us organize a lot of numbers
- Matrix algebra lets us manipulate a lot of numbers at once

Matrix algebra is the language of statistics

2.1.2 Scalar

A scalar is an "ordinary" number

The algebra of scalars is arithmetic

- 4
- $2 + 5 = 7$

2.1.3 Matrix

- **Doubly ordered** arrangement of scalars
  - **Rows** represent one aspect (e.g., subjects)
  - **Columns** represent another (e.g., variables)
- Denoted by capital letters (often **bold** capital letters)
- The **order** (size) of the matrix is (# rows × # columns)
  - In general, matrices are order $p \times q$
- The **elements** in the matrix are identified by subscripts
  - Element a_{ij} is in row i and column j

2.1.4 Matrix X with 4 rows and 3 columns

$$\underset{(4,3)}{\mathbf{X}} = \begin{bmatrix} 2 & 6 & 5 \\ 8 & 1 & 4 \\ 9 & 3 & 6 \\ 2 & 0 & 5 \end{bmatrix}$$

2.1.5 Matrix A with 2 rows and 5 columns

$$\underset{(2,5)}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \end{bmatrix}$$

2.1.6 Vector

- A **row vector** is a matrix of order $1 \times q$

- A **column vector** is a matrix of order $p \times 1$
- Denoted by lower case, underlined letters

2.1.7 Row vector

$$\underset{(1,q)}{\underline{x}} = \begin{bmatrix} x_1 & x_2 & \cdots & x_q \end{bmatrix}$$

2.1.8 Column vector

$$\underset{(p,1)}{\underline{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

2.2 Matrix algebra

2.2.1 Matrix algebra

Matrix algebra is the set of rules for performing mathematical operations on matrices (and vectors)

- Addition and subtraction are straight-forward extensions

- Multiplication and division are not

- Other matrix-specific operations, such as the transpose

2.2.2 Transpose

- Switch the rows and columns

- Rows become columns and columns become rows

$$\underset{(2,3)}{\mathbf{A}} = \begin{bmatrix} 6 & 2 & 4 \\ 8 & 1 & 0 \end{bmatrix}$$

$$\underset{(3,2)}{\mathbf{A'}} = \underset{(3,2)}{\mathbf{A}^T} = \begin{bmatrix} 6 & 8 \\ 2 & 1 \\ 4 & 0 \end{bmatrix}$$

2.2.3 Transpose

- The transpose of a column vector is a row vector
- The transpose of a row vector is a column vector

$$\underset{(4,1)}{\underline{x}} = \begin{bmatrix} 2 \\ 8 \\ 9 \\ 2 \end{bmatrix}$$

$$\underset{(1,4)}{\underline{x}'} = \underset{(1,4)}{\underline{x}^T} = \begin{bmatrix} 2 & 8 & 9 & 2 \end{bmatrix}$$

2.2.4 Addition

Matrices must be of the **same order** to be **conformable** for addition

- $\mathbf{A} + \mathbf{B} = \mathbf{C}$
- Add corresponding elements: $c_{ij} = a_{ij} + b_{ij}$

$$\begin{bmatrix} 3 & 1 & 5 \\ 2 & 4 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 7 \\ 8 & 11 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 5 & 12 \\ 10 & 15 & 10 \end{bmatrix}$$

- Commutative: $(\mathbf{A} + \mathbf{B}) = (\mathbf{B} + \mathbf{A})$
- Associative: $[\mathbf{A} + (\mathbf{B} + \mathbf{C})] = [(\mathbf{A} + \mathbf{B}) + \mathbf{C}]$

2.2.5 Subtraction

Matrices must be of the **same order** to be **conformable** for subtraction

- $\mathbf{A} - \mathbf{B} = \mathbf{D}$
- Add corresponding elements: $d_{ij} = a_{ij} - b_{ij}$

$$\begin{bmatrix} 3 & 1 & 5 \\ 2 & 4 & 6 \end{bmatrix} - \begin{bmatrix} 2 & 4 & 7 \\ 8 & 11 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -3 & -2 \\ -6 & -7 & 2 \end{bmatrix}$$

2.2.6 Multiplication

- Multiplication of matrices is more complicated
- This is "inner" or "dot" product matrix multiplication
  - There is also an "outer" product, which we won't use

Rules for matrix multiplication

- $\underset{(p,q)}{\mathbf{A}} \times \underset{(q,r)}{\mathbf{B}} = \underset{(p,r)}{\mathbf{C}}$
- $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum a_{ik}b_{kj}$

2.2.7 Multiplication example: matrix times matrix

$$\underset{(2,3)}{\mathbf{A}} \; \underset{(3,2)}{\mathbf{B}} = \underset{(2,2)}{\mathbf{C}}$$

$$\begin{bmatrix} 2 & 4 & 1 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 8 \\ 19 & 17 \end{bmatrix}$$

Row 1, Column 1: $(2 \times 1) + (4 \times 2) + (1 \times 4) = 14$

Row 1, Column 2: $(2 \times 3) + (4 \times 0) + (1 \times 2) = 8$

Row 2, Column 1: $(3 \times 1) + (0 \times 2) + (4 \times 4) = 19$

Row 2, Column 2: $(3 \times 3) + (0 \times 0) + (4 \times 2) = 17$

2.2.8 Multiplication example: matrix times column vector

$$\underset{(2,3)}{\mathbf{A}}\ \underset{(3,1)}{\underline{b}} = \underset{(2,1)}{\underline{c}}$$

$$\begin{bmatrix} 2 & 4 & 1 \\ 3 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 14 \\ 19 \end{bmatrix}$$

Row 1, Column 1: $(2 \times 1) + (4 \times 2) + (1 \times 4) = 14$

Row 2, Column 1: $(3 \times 1) + (0 \times 2) + (4 \times 4) = 19$

2.2.9 Multiplication example: data matrix times column vector of regression coefficients

$$\underset{(5,3)}{\mathbf{X}}\ \underset{(3,1)}{\underline{b}} = \underset{(5,1)}{\underline{c}}$$

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \\ X_{41} & X_{42} & X_{43} \\ X_{51} & X_{52} & X_{53} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1X_{11} + b_2X_{12} + b_3X_{13} \\ b_1X_{21} + b_2X_{22} + b_3X_{23} \\ b_1X_{31} + b_2X_{32} + b_3X_{33} \\ b_1X_{41} + b_2X_{42} + b_3X_{43} \\ b_1X_{51} + b_2X_{52} + b_3X_{53} \end{bmatrix}$$

2.2.10 Multiplication example: row vector times column vector

$$\underset{(1,3)}{\underline{a}'}\ \underset{(3,1)}{\underline{b}} = \underset{(1,1)}{\underline{c}}$$

$$\begin{bmatrix} 3 & 1 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 55 \end{bmatrix}$$

$(3 \times 2) + (1 \times 4) + (5 \times 9) = 55$

2.2.11 Multiplication example: column vector times row vector

$$\underset{(3,1)}{\underline{b}}\ \underset{(1,3)}{\underline{a}'} = \underset{(3,3)}{\mathbf{C}}$$

$$\begin{bmatrix} 2 \\ 4 \\ 9 \end{bmatrix} \begin{bmatrix} 3 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 10 \\ 12 & 4 & 20 \\ 27 & 9 & 45 \end{bmatrix}$$

Row 1, Column 1: $(2 \times 3) = 6$

Row 1, Column 2: $(2 \times 1) = 2$

Row 1, Column 3: $(2 \times 5) = 10$

Row 2, Column 1: $(4 \times 3) = 12$

Row 2, Column 2: $(4 \times 1) = 4$

Row 2, Column 3: $(4 \times 5) = 20$

Row 3, Column 1: $(9 \times 3) = 27$

Row 3, Column 2: $(9 \times 1) = 9$

Row 3, Column 3: $(9 \times 5) = 45$

2.2.12 Multiplication example: matrix times a scalar

Multiply every element in the matrix by the scalar

$$3 \times \begin{bmatrix} 3 & 2 & 4 \\ 1 & 2 & 3 \\ 8 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 9 & 6 & 12 \\ 3 & 6 & 9 \\ 24 & 3 & 3 \end{bmatrix}$$

2.2.13 Multiplication of matrices: Associative

$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

- When multiplying > 2 matrices, it doesn't matter which two you start with
- Must keep the same overall order

2.2.14 Multiplication of matrices: Distributive

With respect to addition and subtraction:

$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

- Distribute the **A** matrix as you would in arithmetic

2.2.15 Multiplication of matrices: NOT commutative

$\mathbf{AB} \neq \mathbf{BA}$

- **Order matters for matrix multiplication**
- See *row x column* and *column x row* examples

2.2.16 Multiplication of matrices: Transpose and multiplication

- If $\mathbf{D} = \mathbf{AB}$, then $\mathbf{D}' = \mathbf{B}'\mathbf{A}'$
  - The transpose is equal to the product of the transposes in **reverse order**

2.2.17 Division

Division is not defined for matrices

Instead of dividing, we **multiply by the inverse**

- The inverse of a number is the number raised to the power of -1
  - e.g., inverse of $5 = 5^{-1} = \frac{1}{5}$

We do the same in regular arithmetic:

- Divide: $30/5 = 6$
- Multiply by inverse: $30 \times 5^{-1} = 30 \times \frac{1}{5} = 6$

2.2.18 Division

Calculating the inverse of a matrix is complicated (more later)

Multiplying the matrix by its inverse results in the **identity matrix**:

$\mathbf{A}\ \mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

2.2.19 Identity matrix

The identity matrix ($\mathbf{I}$) is a special matrix that looks like:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

1s on the **main diagonal**, 0s elsewhere

Multiplying a matrix by $\mathbf{I}$ is like multiplying it by 1

- Why would you do this? To be able to perform other matrix operations

2.2.20 Matrix times inverse = identity matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 2 & 4 & 1 \end{bmatrix} \qquad \mathbf{A}^{-1} = \begin{bmatrix} 1 & -2 & 1 \\ -\frac{3}{5} & 1 & -\frac{1}{5} \\ \frac{2}{5} & 0 & -\frac{1}{5} \end{bmatrix}$$

Row 1, Column 1: $1 = (1 \times 1) + (2 \times -\frac{3}{5}) + (3 \times \frac{2}{5})$

Row 1, Column 2: $0 = (1 \times -2) + (2 \times 1) + (3 \times 0)$

Row 1, Column 3: $0 = (1 \times 1) + (2 \times -\frac{1}{5}) + (3 \times -\frac{1}{5})$

Row 2, Column 1: $0 = (1 \times 1) + (3 \times -\frac{3}{5}) + (2 \times \frac{2}{5})$

Row 2, Column 2: $1 = (1 \times -2) + (3 \times 1) + (2 \times 0)$

Row 2, Column 3: $0 = (1 \times 1) + (3 \times -\frac{1}{5}) + (2 \times -\frac{1}{5})$

Row 3, Column 1: $0 = (2 \times 1) + (4 \times -\frac{3}{5}) + (1 \times \frac{2}{5})$

Row 3, Column 2: $0 = (2 \times -2) + (4 \times 1) + (1 \times 0)$

Row 3, Column 3: $1 = (2 \times 1) + (4 \times -\frac{1}{5}) + (1 \times -\frac{1}{5})$

3 Review: Basic statistics

3.1 Statistics

This section **reviews** material you should already be familiar with from previous courses

You don't need to memorize these equations, but you should be *comfortable* with using them

All of the material in this course (and in all of statistics!) builds on these basic concepts of *central tendency*, *variability*, and *covariability*

3.2 Central tendency

3.2.1 Arithmetic mean

- **Population**:

$\mu_X = \frac{\sum X}{N}$

where N is the size of the *population*

- **Sample**:

$\overline{X} = \frac{\sum X}{n}$

where n is the size of the *sample*

3.3 Variability of one variable

3.3.1 Variation

Sum of squared deviations from mean: *sum of squares (SS)*

- **Population**:

$SS_X = \sum(X_i - \mu_X)^2 = \sum X^2 - \frac{(\sum X)^2}{N}$

where μ_X is the *population* mean

- **Sample**:

$SS_X = \sum(X_i - \overline{X})^2 = \sum X^2 - \frac{(\sum X)^2}{n}$

where $\overline{X}$ is the *sample* mean

3.3.2 Variance

Average squared deviation of scores around the mean

- **Population**:

${\sigma^2}_X = \frac{\sum(X_i - \mu_X)^2}{N} = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}$

where μ_X is the *population* mean and N is the *population* size

- **Sample**:

${s^2}_X = \frac{\sum(X_i - \overline{X})^2}{n-1} = \frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}$

where $\overline{X}$ is the *sample* mean and n is the *sample* size

3.3.3 Standard deviation

Square root of variance: in the same units as the original variable

- **Population**:

$\sigma_X = \sqrt{\frac{\sum(X_i - \mu_X)^2}{N}} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$

- **Sample**:

$s_X = \sqrt{\frac{\sum(X_i - \overline{X})^2}{n-1}} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}}$

3.4 Relationship between 2 variables

3.4.1 Covariation

Analogous to **variation**: *sum of cross-products* of the deviations or sum of products (SP)

- **Population**:

$SP_{XY} = \sum(X_i - \mu_X)(Y_i - \mu_Y) = \sum XY - \frac{(\sum X)(\sum Y)}{N}$
where μ_X and μ_Y are the *population* means

- **Sample**:

$SP_{XY} = \sum(X_i - \overline{X})(Y_i - \overline{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{n}$
where $\overline{X}$ and $\overline{Y}$ are the *sample* means

3.4.2 Covariance

Analogous to **variance**: *average* sum of cross-products of the deviations around the mean

- **Population**:

$\sigma_{XY} = \frac{\sum(X_i - \mu_X)(Y_i - \mu_Y)}{N} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{N}$

- **Sample**:

$s_{XY} = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{n-1} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{n-1}$

3.4.3 Correlation

Standardized measure of how two variables are related

- **Population**:

$\rho_{XY} = \frac{\sum z_X z_Y}{N} = \frac{SP_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$

- **Sample**:

$r_{XY} = \frac{\sum z_X z_Y}{n} = \frac{SP_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}} = \frac{s_{XY}}{s_X s_Y}$

where z_X and z_Y are standard scores (z-scores):

$z_X = \frac{X_i - \mu_X}{\sigma_X}$ (population) or $z_X = \frac{X_i - \overline{X}}{s_X}$ (sample)

4 Basic statistics: Matrix!

4.1 Data matrix

4.1.1 Data matrix

Data is usually presented as

- Rows for subjects
- Columns for variables

The *data matrix* is an $n \times p$ matrix

- n subjects (rows: $1, 2, \dots, i, \dots, n$)
- p variables (columns: $X_1, X_2, \dots, X_j, \dots, X_p$)

4.1.2 Data matrix

$$\mathbf{X} = \begin{bmatrix} X_{11} & \cdots & X_{1j} & \cdots & X_{1p} \\ \vdots & & \vdots & & \vdots \\ X_{i1} & \cdots & X_{ij} & \cdots & X_{ip} \\ \vdots & & \vdots & & \vdots \\ X_{n1} & \cdots & X_{nj} & \cdots & X_{np} \end{bmatrix}$$

4.2 Unit vector and matrix

4.2.1 Unit vector

The unit vector is a vector filled with 1s

$$\underset{(n,1)}{\underline{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

- Used to **add numbers in a matrix together**
- Same function as $\sum$ in arithmetic: $\sum_{i=1}^{n} X = \underline{1}'\underline{x}$
- The unit vector $\underline{1}$ is typically a *column vector* but we can also us its transpose $\underline{1}'$ when we need a *row vector*

4.2.2 Unit vector adds up elements

$$\underset{(1,4)}{\underline{1}'} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \quad \underset{(4,1)}{\underline{x}} = \begin{bmatrix} 4 \\ 3 \\ 8 \\ 2 \end{bmatrix}$$

$\underline{1}'\underline{x} = (1 \times 4) + (1 \times 3) + (1 \times 8) + (1 \times 2) =$

$4 + 3 + 8 + 2 = 17$

4.2.3 Unit matrix

The unit matrix is a matrix filled with 1s

$$\underset{(n,n)}{\mathbf{E}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

- Used to **add numbers** and **create products of numbers**
- For most of our purposes, we'll use an $n \times n$ version, but it can be any size / order

4.3 Central tendency

4.3.1 Mean of a single variable

Pre-multiply the vector of values by the unit vector and multiply by inverse of n

$\overline{X} = \frac{1}{n}\, \underline{1}'\, \underline{x}$

4.3.2 Mean of a single variable

Example: Variable X is observed for $n = 4$ subjects

$$\underset{(1,4)}{\underline{1}'} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \quad \underset{(4,1)}{\underline{x}} = \begin{bmatrix} 4 \\ 3 \\ 8 \\ 2 \end{bmatrix}$$

$\overline{X} = \frac{1}{n}\, \underline{1}'\, \underline{x} =$

$\frac{1}{4}[(1 \times 4) + (1 \times 3) + (1 \times 8) + (1 \times 2)] =$

$\frac{1}{4}(4 + 3 + 8 + 2) = \frac{17}{4} = 4.25$

4.3.3 Mean of several variables

Example: Variables X_1, X_2, and X_3 for $n = 4$ subjects

$$\underset{(1,4)}{\underline{1}'} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \quad \underset{(4,3)}{\mathbf{X}} = \begin{bmatrix} 4 & 2 & 4 \\ 3 & 1 & 1 \\ 8 & 3 & 2 \\ 2 & 5 & 5 \end{bmatrix}$$

$$\underline{\overline{x}} = \frac{1}{n}\, \underline{1}'\, \mathbf{X} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 4 \\ 3 & 1 & 1 \\ 8 & 3 & 2 \\ 2 & 5 & 5 \end{bmatrix} =$$

$\frac{1}{4}\begin{bmatrix} 4+3+8+2 & 2+1+3+5 & 4+1+2+5 \end{bmatrix} =$

$\frac{1}{4}\begin{bmatrix} 17 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 4.25 & 2.75 & 3 \end{bmatrix}$

4.4 (Co)variation

4.4.1 Some matrix algebra properties

Sum a variable across all subjects:

$\sum X = \sum_{i=1}^{n} X = \underline{1}' \, \mathbf{X} = \mathbf{X}' \, \underline{1}$

Sum THEN square:

$(\sum X)^2 = (\sum_{i=1}^{n} X)^2 = \mathbf{X}' \, \underline{1} \, \underline{1}' \, \mathbf{X} = \mathbf{X}' \, \mathbf{E} \, \mathbf{X}$

Square THEN sum:

$\sum(X^2) = \sum_{i=1}^{n}(X^2) = \mathbf{X}' \, \mathbf{X}$

4.4.2 Variation

Recall that the *sample variation* is:

$SS_X = \sum(X_i - \overline{X})^2 = \sum(X^2) - \frac{(\sum X)^2}{n}$

4.4.3 Variation

Recall that the *sample variation* is:

$SS_X = \sum(X_i - \overline{X})^2 = \sum(X^2) - \frac{(\sum X)^2}{n}$

And that:

$\sum(X^2) = \mathbf{X}' \, \mathbf{X}$

$(\sum X)^2 = \mathbf{X}' \, \underline{1} \, \underline{1}' \, \mathbf{X} = \mathbf{X}' \, \mathbf{E} \, \mathbf{X}$

4.4.4 Variation

Recall that the *sample variation* is:

$SS_X = \sum(X_i - \overline{X})^2 = \sum(X^2) - \frac{(\sum X)^2}{n}$

And that:

$\sum(X^2) = \mathbf{X}'\,\mathbf{X}$

$(\sum X)^2 = \mathbf{X}'\,\underline{1}\,\underline{1}'\,\mathbf{X} = \mathbf{X}'\,\mathbf{E}\,\mathbf{X}$

Substitute matrix expressions:

$SS_X = \mathbf{X}'\,\mathbf{X} - \frac{1}{n}\,(\mathbf{X}'\,\mathbf{E}\,\mathbf{X})$

4.4.5 Covariation

Recall that the *sample covariation* is:

$SP_{XY} = \sum(X_i - \overline{X})(Y_i - \overline{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{n}$

4.4.6 Covariation

Recall that the *sample covariation* is:

$SP_{XY} = \sum(X_i - \overline{X})(Y_i - \overline{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{n}$

And that (extending to the X and Y situation):

$\sum(XY) = \mathbf{X}'\,\mathbf{Y}$

$(\sum X)(\sum Y) = \mathbf{X}'\,\underline{1}\,\underline{1}'\,\mathbf{Y} = \mathbf{X}'\,\mathbf{E}\,\mathbf{Y}$

4.4.7 Covariation

Recall that the *sample covariation* is:

$SP_{XY} = \sum(X_i - \overline{X})(Y_i - \overline{Y}) = \sum XY - \frac{(\sum X)(\sum Y)}{n}$

And that (extending to the X and Y situation):

$\sum(XY) = \mathbf{X}'\,\mathbf{Y}$

$(\sum X)(\sum Y) = \mathbf{X}'\,\underline{1}\,\underline{1}'\,\mathbf{Y} = \mathbf{X}'\,\mathbf{E}\,\mathbf{Y}$

Substitute matrix expressions:

$SP_{XY} = \mathbf{X}'\,\mathbf{Y} - \frac{1}{n}\,(\mathbf{X}'\,\mathbf{E}\,\mathbf{Y})$

4.4.8 Variation-covariation matrix (P)

- Involves many variables
- Subscripts indicate which variables are involved: $\mathbf{P}_{XX}$, $\mathbf{P}_{XY}$
- **Variation** along the diagonal, **covariation** elsewhere

$$\mathbf{P}_{XX} = \mathbf{X'X} - \frac{1}{n}\mathbf{X'EX} = \begin{bmatrix} SS_{X_1} & SP_{X_1X_2} & \cdots & SP_{X_1X_p} \\ SP_{X_2X_1} & SS_{X_2} & \cdots & SP_{X_2X_p} \\ \vdots & \vdots & \ddots & \vdots \\ SP_{X_pX_1} & SP_{X_pX_2} & \cdots & SS_{X_p} \end{bmatrix}$$

4.5 (Co)variance

4.5.1 Variance

Recall that the *sample variance* is:

$s_X^2 = \frac{variation}{n-1} = \frac{SS_X}{n-1}$

Multiply the matrix expression for variation by $\frac{1}{n-1}$:

$s_X^2 = \frac{1}{n-1}\Big(\mathbf{X}'\,\mathbf{X} - \frac{1}{n}\,(\mathbf{X}'\,\mathbf{E}\,\mathbf{X})\Big)$

4.5.2 Covariance

Recall that the *sample covariance* is:

$cov_{XY} = s_{XY} = \frac{covariation}{n-1} = \frac{SP_{XY}}{n-1}$

Mulitply the matrix expression for covariation by $\frac{1}{n-1}$:

$cov_{XY} = s_{XY} = \frac{1}{n-1}\Big(\mathbf{X}'\,\mathbf{Y} - \frac{1}{n}\,(\mathbf{X}'\,\mathbf{E}\,\mathbf{Y})\Big)$

4.5.3 Variance-covariance matrix (S)

- Involves many variables
- Subscripts indicate which variables are involved: $\mathbf{S}_{XX}$, $\mathbf{S}_{XY}$
- **Variance** along the diagonal, **covariance** elsewhere
- **One of THE most important matrices in statistics**

$$\mathbf{S}_{XX} = \frac{1}{n-1}(\mathbf{X'X} - \frac{1}{n}\mathbf{X'EX}) = \begin{bmatrix} s_{X_1}^2 & s_{X_1X_2} & \cdots & s_{X_1X_p} \\ s_{X_2X_1} & s_{X_2}^2 & \cdots & s_{X_2X_p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{X_pX_1} & s_{X_pX_2} & \cdots & s_{X_p}^2 \end{bmatrix}$$

4.6 Correlation

4.6.1 Correlation

The correlation between X and Y is:

$r_{XY} = \frac{SP_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}}$

Since division for matrices means **multiplication by the inverse**:

- We need the inverse of $\sqrt{SS_X}$ and $\sqrt{SS_Y}$
- i.e., $\sqrt{SS_X}^{-1}$ and $\sqrt{SS_Y}^{-1}$

4.6.2 Reciprocals of square root of variation

$\mathbf{D}_P$ is a matrix with the square root of **variation** on the diagonal:

$$\mathbf{D}_P = \begin{bmatrix} \sqrt{SS_{X_1}} & 0 & \cdots & 0 \\ 0 & \sqrt{SS_{X_2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{SS_{X_p}} \end{bmatrix}$$

The inverse of $\mathbf{D}_P$:

$$\mathbf{D}_P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{SS_{X_1}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{SS_{X_2}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{SS_{X_p}}} \end{bmatrix}$$

4.6.3 Reciprocals of square root of variance

$\mathbf{D}_S$ is a matrix with the square root of **variance** on the diagonal:

$$\mathbf{D}_S = \begin{bmatrix} \sqrt{s^2_{X_1}} & 0 & \cdots & 0 \\ 0 & \sqrt{s_{X_2^2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{s^2_{X_p}} \end{bmatrix}$$

The inverse of $\mathbf{D}_S$:

$$\mathbf{D}_S^{-1} = \begin{bmatrix} \frac{1}{\sqrt{s^2_{X_1}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{s^2_{X_2}}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{s^2_{X_p}}} \end{bmatrix}$$

4.6.4 Correlation matrix (R)

- Involves many variables
- Subscripts indicate which variables are involved: $\mathbf{R}_{XX}$, $\mathbf{R}_{XY}$
- 1**s** along the diagonal, **correlations** elsewhere
- **One of THE most important matrices in statistics**

4.6.5 Correlation matrix (R)

In terms of *variation and covariation*:

$$\mathbf{R}_{XX} = \mathbf{D}_P^{-1}\,\mathbf{P}\,\mathbf{D}_P^{-1} = \begin{bmatrix} 1 & r_{X_1X_2} & \cdots & r_{X_1X_p} \\ r_{X_2X_1} & 1 & \cdots & r_{X_2X_p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{X_pX_1} & r_{X_pX_2} & \cdots & 1 \end{bmatrix}$$

4.6.6 Correlation matrix (R)

In terms of *variance and covariance*:

$$\mathbf{R}_{XX} = \mathbf{D}_S^{-1}\,\mathbf{S}\,\mathbf{D}_S^{-1} = \begin{bmatrix} 1 & r_{X_1X_2} & \cdots & r_{X_1X_p} \\ r_{X_2X_1} & 1 & \cdots & r_{X_2X_p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{X_pX_1} & r_{X_pX_2} & \cdots & 1 \end{bmatrix}$$