S1 Sains Data
FMIPA UNESA

Analisis Multivariat – Materi 02

Review Matrix Algebra

Prodi S1 Sains Data

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Surabaya

2025

Outline

- Scalar, Vectors and Matrix
- **Variance and Covariance Matrix**
- **Correlation Matrix**

Scalar, Vectors and Matrix

Scalar and Vectors

- Scalars: A scalar is just a single number, in contrast to most of the other objects studied in linear algebra, which are usually arrays of multiple numbers.
- Vectors: A vector is an array of numbers. The numbers are arranged in order.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Matrix and Vectors

A matrix $\mathcal{A}$ is a system of numbers with n rows and p columns:

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & a_{1p} \\ \vdots & a_{22} & & & & \vdots \\ \vdots & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & \dots & a_{np} \end{pmatrix}.$$

We also write (a_{ij}) for $\mathcal{A}$ and $\mathcal{A}(n \times p)$ to indicate the numbers of rows and columns.

Vectors are matrices with one column and are denoted as x or $x(p \times 1)$.

Special Matices and Vectors

Special matrices and vectors are defined in Table. Note that we use small letters for scalars as well as for vectors.

Name	Definition	Notation	Example
Scalar	$p = n = 1$	a	3
Column vector	$p = 1$	a	$\begin{pmatrix} 1 \\ 3 \end{pmatrix}$
Row vector	$n = 1$	$a^\top$	$\begin{pmatrix} 1 & 3 \end{pmatrix}$
Vector of ones	$(\underbrace{1, \dots, 1}_{n})^\top$	1_n	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Vector of zeros	$(\underbrace{0, \dots, 0}_{n})^\top$	0_n	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
Square matrix	$n = p$	$\mathcal{A}(p \times p)$	$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$
Diagonal matrix	$a_{ij} = 0, i \neq j, n = p$	$\mathrm{diag}(a_{ii})$	$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$
Identity matrix	$\mathrm{diag}(\underbrace{1, \dots, 1}_{p})$	$\mathcal{I}_p$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
Identity matrix	$\mathrm{diag}(\underbrace{1, \dots, 1}_{p})$	$\mathcal{I}_p$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
Unit matrix	$a_{ij} = 1, n = p$	$1_n 1_n^\top$	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
Symmetric matrix	$a_{ij} = a_{ji}$		$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$
Null matrix	$a_{ij} = 0$	0	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$
Upper triangular matrix	$a_{ij} = 0, i < j$		$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$
Idempotent matrix	$\mathcal{A}\mathcal{A} = \mathcal{A}$		$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$
Orthogonal matrix	$\mathcal{A}^\top \mathcal{A} = \mathcal{I} = \mathcal{A}\mathcal{A}^\top$		$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$

Matrix Operation

Elementary operations are summarized below:

$$\begin{aligned} \mathcal{A}^\top &= (a_{ji}) \\ \mathcal{A} + \mathcal{B} &= (a_{ij} + b_{ij}) \\ \mathcal{A} - \mathcal{B} &= (a_{ij} - b_{ij}) \\ c \cdot \mathcal{A} &= (c \cdot a_{ij}) \end{aligned}$$

$$\mathcal{A} \cdot \mathcal{B} = \mathcal{A}(n \times p)\, \mathcal{B}(p \times m) = \mathcal{C}(n \times m) = (c_{ij}) = \left(\sum_{j=1}^{p} a_{ij} b_{jk} \right).$$

Properties of Matrix Operation

$$\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$$ Comutative of addition

$$\mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C}$$ distributive

$$\mathcal{A}(\mathcal{B}\mathcal{C}) = (\mathcal{A}\mathcal{B})\mathcal{C}$$ asosiative

$$(\mathcal{A}^\top)^\top = \mathcal{A}$$

$$(\mathcal{A}\mathcal{B})^\top = \mathcal{B}^\top \mathcal{A}^\top$$

Not comutative of multiplication except multiply with scalar, but dot product of two vectors is commutative.

Matrix Characteristics

- Rank
- Trace
- Determinant
- Transpose
- Inverse
- G-Inverse
- Eigen Value and Eigen Vector

Invers

If $|\mathcal{A}| \neq 0$ and $\mathcal{A}(p \times p)$, then the inverse $\mathcal{A}^{-1}$ exists:

$$\mathcal{A}\,\mathcal{A}^{-1} = \mathcal{A}^{-1}\,\mathcal{A} = \mathcal{I}_p.$$

For small matrices, the inverse of $\mathcal{A} = (a_{ij})$ can be calculated as

$$\mathcal{A}^{-1} = \frac{\mathcal{C}}{|\mathcal{A}|},$$

where $\mathcal{C} = (c_{ij})$ is the adjoint matrix of $\mathcal{A}$. The elements c_{ji} of $\mathcal{C}^\top$ are the co-facto of $\mathcal{A}$:

$$c_{ji} = (-1)^{i+j} \begin{vmatrix} a_{11} & \dots & a_{1(j-1)} & a_{1(j+1)} & \dots & a_{1p} \\ \vdots & & & & & \\ a_{(i-1)1} & \dots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \dots & a_{(i-1)p} \\ a_{(i+1)1} & \dots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \dots & a_{(i+1)p} \\ \vdots & & & & & \\ a_{p1} & \dots & a_{p(j-1)} & a_{p(j+1)} & \dots & a_{pp} \end{vmatrix}.$$

The relationship between determinant and inverse of matrix $\mathcal{A}$ is $|\mathcal{A}^{-1}| = |\mathcal{A}|^{-1}$.

Invers using OBE (Operasi Baris Elementer)

Diberikan

$$E_3 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$R_1 \leftarrow R_1 + 2R_2$$

Invers matriks elementer diperoleh dari *I* dengan menerapkan operasi baris elementer kebalikannya.

Jika E_3 diperoleh dari *I* dengan **menjumlahkan baris ke-*i* dengan hasil kali baris *ke-j*** dengan konstanta tak nol *k*, maka $(E_3)^{-1}$ diperoleh dari *I* dengan **mengurangkan baris ke-*i* dengan *k* kali baris *ke-j*.**

$$(E_3)^{-1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Longrightarrow \underbrace{\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_3} \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{(E_3)^{-1}} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{I}$$

G-Invers

A more general concept is the *G-inverse* (Generalized Inverse) $\mathcal{A}^-$ which satisfies the following:

$$\mathcal{A}\mathcal{A}^-\mathcal{A} = \mathcal{A}.$$

Later we will see that there may be more than one G-inverse.

Example 2.2 The generalized inverse can also be calculated for singular matrices. We have

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

which means that the generalized inverse of $\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is $\mathcal{A}^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ even though the inverse matrix of $\mathcal{A}$ does not exist in this case.

Eigen Value and Eigen Vector

Consider a $(p \times p)$ matrix $\mathcal{A}$. If there a scalar λ and a vector γ exists such as

$$\mathcal{A}\gamma = \lambda\gamma, \tag{2.1}$$

then we call

λ	an eigenvalue
γ	an eigenvector.

It can be proven that an eigenvalue λ is a root of the p-th order polynomial $|\mathcal{A} - \lambda\mathcal{I}_p| = 0$. Therefore, there are up to p eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ of $\mathcal{A}$. For each eigenvalue λ_j, a corresponding eigenvector γ_j exists given by Eq. (2.1). Suppose the matrix $\mathcal{A}$ has the eigenvalues $\lambda_1, \ldots, \lambda_p$. Let $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$.

The determinant $|\mathcal{A}|$ and the trace $\operatorname{tr}(\mathcal{A})$ can be rewritten in terms of the eigenvalues:

$$|\mathcal{A}| = |\Lambda| = \prod_{j=1}^{p} \lambda_j \tag{2.2}$$

$$\operatorname{tr}(\mathcal{A}) = \operatorname{tr}(\Lambda) = \sum_{j=1}^{p} \lambda_j. \tag{2.3}$$

An idempotent matrix $\mathcal{A}$ (see the definition in Table 2.1) can only have eigenvalues in $\{0, 1\}$; therefore, $\operatorname{tr}(\mathcal{A}) = \operatorname{rank}(\mathcal{A}) = $ number of eigenvalues $\neq 0$.

Definisi

- Jika A adalah matriks n x n maka vektor tidak-nol $\mathbf{x}$ di R^n disebut **vektor eigen** dari A jika $A\mathbf{x}$ sama dengan perkalian suatu skalar λ dengan $\mathbf{x}$, yaitu

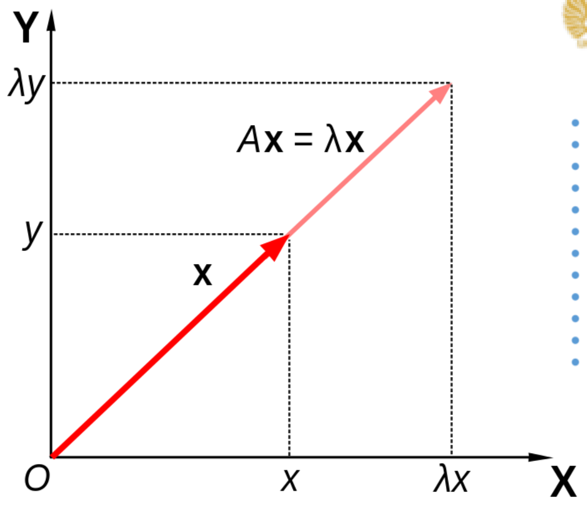
$$\boxed{A\mathbf{x} = \lambda\mathbf{x}}$$

Skalar λ disebut **nilai eigen** dari A, dan $\mathbf{x}$ dinamakan vektor eigen yang berkoresponden dengan λ.

- Kata "eigen" berasal dari Bahasa Jerman yang artinya "asli" atau "karakteristik".

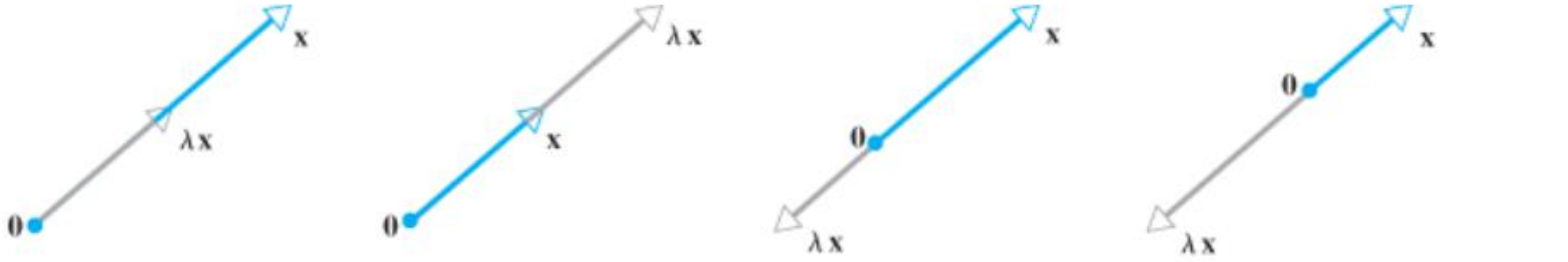
- Dengan kata lain, nilai eigen menyatakan nilai karakteristik dari sebuah matriks yang berukuran n x n.

- Vektor eigen **x** menyatakan matriks kolom yang apabila dikalikan dengan sebuah matriks n x n menghasilkan vektor lain yang merupakan kelipatan vektor itu sendiri.



Sumber gambar: Wikipedia

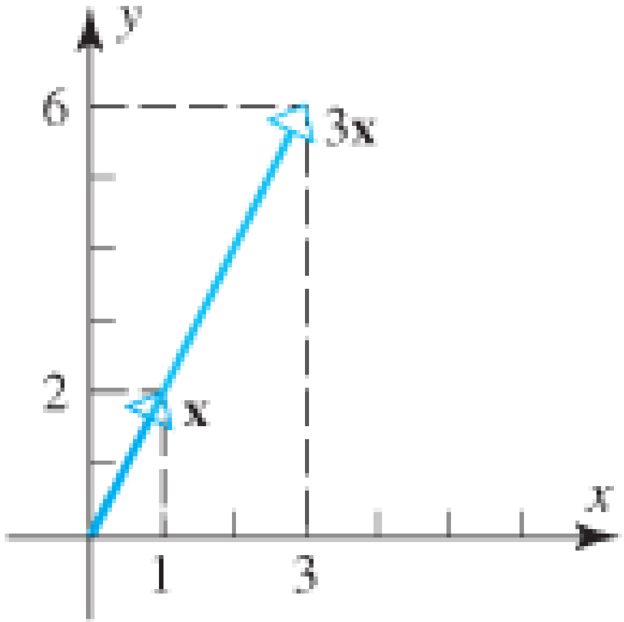
- Dengan kata lain, operasi $A\mathbf{x} = \lambda\mathbf{x}$ menyebabkan vektor **x** menyusut atau memanjang dengan faktor λ dengan arah yang sama jika λ positif dan arah berkebalikan jika λ negatif.



(a) $0 \le \lambda \le 1$ (b) $\lambda \ge 1$ (c) $-1 \le \lambda \le 0$ (d) $\lambda \le -1$

Contoh 1: Misalkan $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$. Vektor $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ merupakan vektor eigen dari A dengan nilai eigen yang berkoresponden $\lambda = 3$, karena

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$



Properties of Matrix Characteristics

$\mathcal{A}(n \times n),\ \mathcal{B}(n \times n),\ c \in \mathbb{R}$

$$\operatorname{tr}(\mathcal{A} + \mathcal{B}) = \operatorname{tr} \mathcal{A} + \operatorname{tr} \mathcal{B} \tag{2.4}$$

$$\operatorname{tr}(c\mathcal{A}) = c \operatorname{tr} \mathcal{A} \tag{2.5}$$

$$|c\mathcal{A}| = c^n |\mathcal{A}| \tag{2.6}$$

$$|\mathcal{A}\mathcal{B}| = |\mathcal{B}\mathcal{A}| = |\mathcal{A}||\mathcal{B}| \tag{2.7}$$

$\mathcal{A}(n \times p),\ \mathcal{B}(p \times n)$

$$\operatorname{tr}(\mathcal{A} \cdot \mathcal{B}) = \operatorname{tr}(\mathcal{B} \cdot \mathcal{A}) \tag{2.8}$$

$$\operatorname{rank}(\mathcal{A}) \leq \min(n, p)$$

$$\operatorname{rank}(\mathcal{A}) \geq 0 \tag{2.9}$$

$$\operatorname{rank}(\mathcal{A}) = \operatorname{rank}(\mathcal{A}^\top) \tag{2.10}$$

$$\operatorname{rank}(\mathcal{A}^\top A) = \operatorname{rank}(\mathcal{A}) \tag{2.11}$$

$$\operatorname{rank}(\mathcal{A} + \mathcal{B}) \leq \operatorname{rank}(\mathcal{A}) + \operatorname{rank}(\mathcal{B}) \tag{2.12}$$

$$\operatorname{rank}(\mathcal{A}\mathcal{B}) \leq \min\{\operatorname{rank}(\mathcal{A}), \operatorname{rank}(\mathcal{B})\} \tag{2.13}$$

Properties of Matrix Characteristics

$\mathcal{A}(n \times p),\ \mathcal{B}(p \times q),\ \mathcal{C}(q \times n)$

$$\begin{aligned} \operatorname{tr}(\mathcal{ABC}) &= \operatorname{tr}(\mathcal{BCA}) \\ &= \operatorname{tr}(\mathcal{CAB}) \qquad (2.14) \end{aligned}$$

$$\operatorname{rank}(\mathcal{ABC}) = \operatorname{rank}(\mathcal{B}) \quad \text{for nonsingular } \mathcal{A}, \mathcal{C} \qquad (2.15)$$

$\mathcal{A}(p \times p)$

$$|\mathcal{A}^{-1}| = |\mathcal{A}|^{-1} \qquad (2.16)$$

$$\operatorname{rank}(\mathcal{A}) = p \quad \text{if and only if } \mathcal{A} \text{ is nonsingular.} \qquad (2.17)$$

Derivative of Matrix

$f : \mathbb{R}^p \to \mathbb{R}$ and a $(p \times 1)$ vector x, then $\frac{\partial f(x)}{\partial x}$ is the column vector of partial derivatives $\left\{\frac{\partial f(x)}{\partial x_j}\right\}$, $j = 1, \ldots, p$ and $\frac{\partial f(x)}{\partial x^\top}$ is the row vector of the same derivative ($\frac{\partial f(x)}{\partial x}$ is called the *gradient* of f).

We can also introduce second-order derivatives: $\frac{\partial^2 f(x)}{\partial x \partial x^\top}$ is the $(p \times p)$ matrix of elements $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$, $i = 1, \ldots, p$ and $j = 1, \ldots, p$. ($\frac{\partial^2 f(x)}{\partial x \partial x^\top}$ is called the *Hessian* of f).

Suppose that a is a $(p \times 1)$ vector and that $\mathcal{A} = \mathcal{A}^\top$ is a $(p \times p)$ matrix. Then

$$\frac{\partial a^\top x}{\partial x} = \frac{\partial x^\top a}{\partial x} = a, \tag{2.23}$$

$$\frac{\partial x^\top \mathcal{A} x}{\partial x} = 2\mathcal{A}x. \tag{2.24}$$

The Hessian of the quadratic form $Q(x) = x^\top \mathcal{A} x$ is

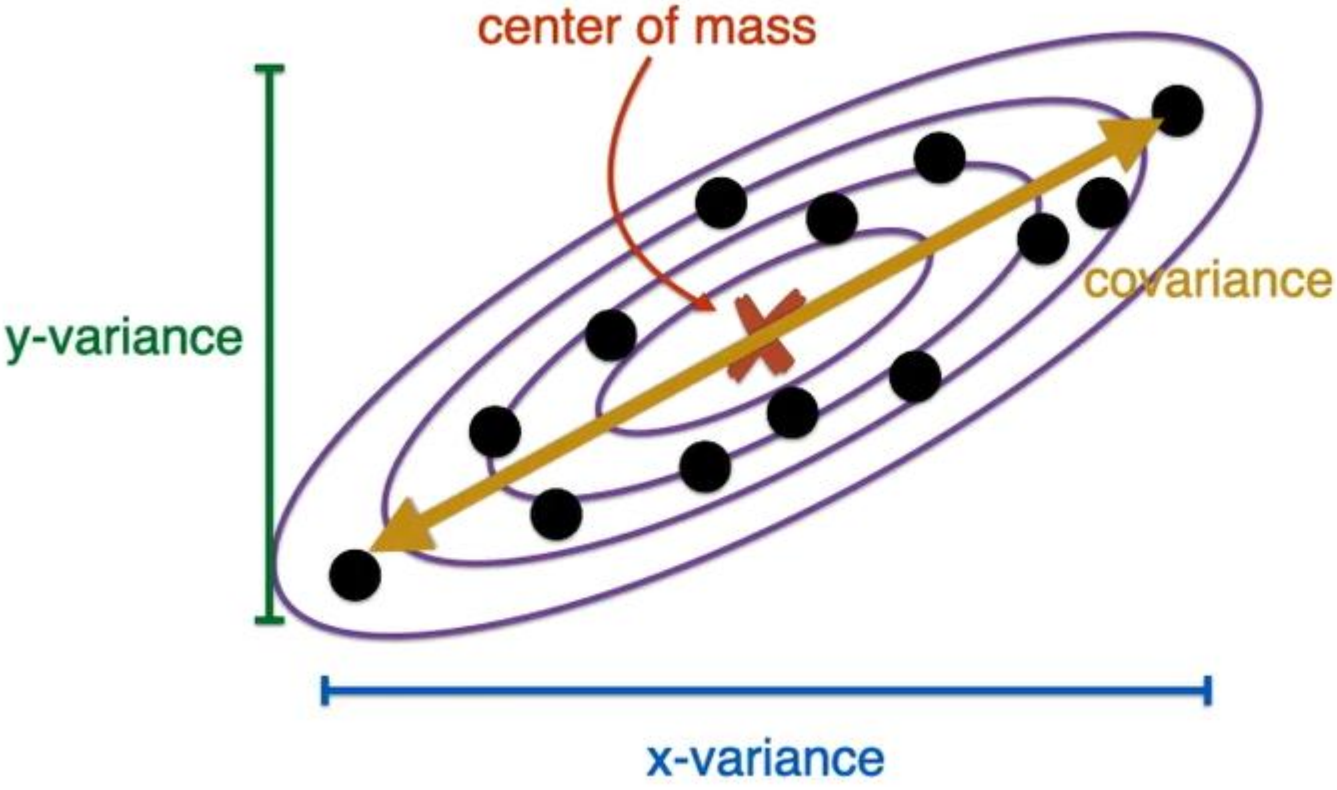
$$\frac{\partial^2 x^\top \mathcal{A} x}{\partial x \partial x^\top} = 2\mathcal{A}. \tag{2.25}$$

High Dimentional Data

Variance-Covariance Matrix

Correlation Matrix

Sum Square Matrix



$$\mu = Average$$

$$\Sigma = \begin{pmatrix} Var(x) & Cov(x,y) \\ Cov(x,y) & Var(y) \end{pmatrix}$$

Variance-Covariance Matrix

Covariance is a measure of dependency between random variables. Given two (random) variables X and Y the (theoretical) covariance is defined by

$$\sigma_{XY} = \mathsf{Cov}(X, Y) = \mathsf{E}(XY) - (\mathsf{E}X)(\mathsf{E}Y). \tag{3.1}$$

The precise definition of expected values is given in Chap. 4. If X and Y are independent of each other, the covariance $\mathsf{Cov}(X, Y)$ is necessarily equal to zero, see Theorem 3.1. The converse is not true. The covariance of X with itself is the variance:

$$\sigma_{XX} = \mathsf{Var}(X) = \mathsf{Cov}(X, X).$$

If the variable X is p-dimensional multivariate, e.g., $X = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$, then the theoretical covariances among all the elements are put into matrix form, i.e., the covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_{X_1X_1} & \dots & \sigma_{X_1X_p} \\ \vdots & \ddots & \vdots \\ \sigma_{X_pX_1} & \dots & \sigma_{X_pX_p} \end{pmatrix}.$$

Correlation Matrix

The correlation between two variables X and Y is defined from the covariance as the following:

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}. \tag{3.7}$$

The advantage of the correlation is that it is independent of the scale, i.e., changing the variables' scale of measurement does not change the value of the correlation. Therefore, the correlation is more useful as a measure of association between two random variables than the covariance. The empirical version of ρ_{XY} is as follows:

$$r_{XY} = \frac{s_{XY}}{\sqrt{s_{XX}s_{YY}}}. \tag{3.8}$$

The correlation is in absolute value always less or equal to 1. It is zero if the covariance is zero and vice-versa. For p-dimensional vectors $(X_1, \dots, X_p)^\top$ we have the theoretical correlation matrix

$$\mathcal{P} = \begin{pmatrix} \rho_{X_1X_1} & \dots & \rho_{X_1X_p} \\ \vdots & \ddots & \vdots \\ \rho_{X_pX_1} & \dots & \rho_{X_pX_p} \end{pmatrix},$$

Example 2

Table 1.1. For example, X_1 might be type of teaching technique, X_2 score on the GRE, X_3 GPA, and X_4 gender, with women coded 1 and men coded 2.

TABLE 1.1 A Data Matrix of Hypothetical Scores

Student	X_1	X_2	X_3	X_4
1	1	500	3.20	1
2	1	420	2.50	2
3	2	650	3.90	1
4	2	550	3.50	2
5	3	480	3.30	1
6	3	600	3.25	2

Find:

1. Variance-Covariance Matrix
2. Correlation Matrix

Assignment

Working with R

S1 Sains Data
FMIPA UNESA

Credit Card Approvals Dataset

	Gender	Age	Debt	Married	BankCustomer	Industry	YearsEmployed	PriorDefault	Employed	CreditScore	DriversLicense	Citizen	ZipCode	Income	Approved
0	1	30.83	0.000	1	1	Industrials	1.25	1	1	1	0	ByBirth	202	0	1
1	0	58.67	4.460	1	1	Materials	3.04	1	1	6	0	ByBirth	43	560	1
2	0	24.50	0.500	1	1	Materials	1.50	1	0	0	0	ByBirth	280	824	1
3	1	27.83	1.540	1	1	Industrials	3.75	1	1	5	1	ByBirth	100	3	1
4	1	20.17	5.625	1	1	Industrials	1.71	1	0	0	0	ByOtherMeans	120	0	1

1. Masukkan data numerik pada tabel tersebut ke R dan tulislah code sehingga menghasilkan

a) Eigen value dan eigen vector

b) Variance-Covariance Matrix

c) Correlation Matrix

2. Pelajari tentang Rpubs dan buatlah langkah-langkah cara publikasi tugas Anda tersebut pada Rpubs

Thank you