MA2001 Linear Algebra I Notes

Contents

Chapter 1

Linear Systems & Gaussian Elimination

1.1 Linear Systems & Their Solutions

1.1.1 Linear Equations

Definition. A **linear equation** in n **variables** (**unknowns**) $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are real constants.

1.1.1.1 Zero Equation

Definition. A linear equation is a **zero equation** if

$$a_1 = a_2 = \ldots = a_n = b = 0$$

1.1.1.2 Nonzero Equation

Definition. A linear equation is a **nonzero equation** if it is not a zero equation

1.1.1.3 Inconsistent Equation

Definition. A linear equation is **inconsistent** if

$$a_1 = a_2 = \ldots = a_n = 0 \text{ but } b \neq 0$$

1.1.2 Solutions of a Linear Equation

Let $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ be a linear equation in n variables $x_1, x_2, \ldots, x_n$.

Definition. For all $s_1, s_2, \ldots, s_n \in \mathbb{R}$, if

$$a_1s_1 + a_2s_2 + \ldots + a_ns_n = b$$

then $\boldsymbol{x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n}$ is a **solution** to the given linear equation.

1.1.2.1 Solution Set

Definition. The set of all solutions is called the **solution set**.

1.1.2.2 General Solution

Definition. An expression that gives the entire solution set is a **general solution**.

1.1.3 Solution Set of Zero Equation

The zero equation in n variables $x_1, x_2, \dots, x_n$ is satisfied by any values of $x_1, x_2, \dots, x_n$.

1.1.4 Solution Set of Inconsistent Linear Equation

An inconsistent linear equation in n variables $x_1, x_2, \dots, x_n$ is not satisfied by any values of $x_1, x_2, \dots, x_n$.

1.1.5 General Solution of Consistent Nonzero Equation

The general solution of a consistent nonzero equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$ in n variables $x_1, x_2, \dots, x_n$ has $n-1$ arbitrary parameters and is of the following form: For any integer $i \in [1, n]$,

$$\begin{cases} x_1 &= s_1, \\ x_2 &= s_2, \\ & \vdots \\ x_{i-1} &= s_{i-1}, \\ x_i &= \frac{1}{a_i}(b - a_1s_1 - a_2s_2 - \ldots - a_{i-1}s_{i-1} - a_{i+1}s_{i+1} - \ldots - a_ns_n), \\ x_{i+1} &= s_{i+1}, \\ & \vdots \\ x_n &= s_n \end{cases}$$

where $s_1, s_2, \dots, s_{i-1}, s_{i+1}, \dots, s_n$ are arbitrary parameters.

1.1.6 Geometrical Interpretation of Linear Equations

1.1.6.1 In Two Variables

The solution set of the linear equation

$$ax + by = c \text{ (in } x, y)$$

where a and b are not both zero represents a **straight line** in the xy-plane.

1.1.6.2 In Three Variables

The solution set of the linear equation

$$ax + by + cz = d \text{ (in } x, y, z)$$

where a, b, c are not all zero represents a **plane** in the xyz-space.

1.1.7 Linear Systems

Definition. A **linear system** (**system of linear equations**) of m linear equations in n variables $x_1, x_2, \ldots, x_n$ is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad\vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

where a_{ij} and b_i are real constants.

- a_{ij} is the **coefficient** of x_j in the ith equation,
- b_i is the **constant term** of the ith equation.

1.1.7.1 Zero System

Definition. If all a_{ij} and b_i are zero, the linear system is a **zero system**.

1.1.7.2 Nonzero System

Definition. If some a_{ij} or b_i is nonzero, the linear system is a **nonzero system**.

1.1.8 Solution of Linear Systems

Given a linear system in n variables $x_1, x_2, \ldots, x_n$, if $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ is a solution to **every equation** of the linear system, then it is a **solution** to the system.

1.1.8.1 Solution Set

Definition. The set of all solutions to the linear system is called the **solution set**.

1.1.8.2 General Solution

Definition. An expression that gives the entire solution set of the linear system is a **general solution**.

1.1.9 Consistency of Linear Systems

Definition. A linear system is

- **consistent** if it has at least one solution;
- **inconsistent** if it has no solution.

1.1.10 Number of Solutions to Linear Systems

A linear system has either

- no solution, or
- exactly one solution, or
- infinitely many solutions

1.1.11 Geometric Interpretation of Linear Systems

1.1.11.1 In Two Variables of Two Equations

Given a linear system in variables x, y of two equations

$$\begin{cases} a_1x + b_1y = c_1, & (L_1) \\ a_2x + b_2y = c_2, & (L_2) \end{cases}$$

where a_1, b_1 are not both zero and a_2, b_2 are not both zero. The system has

- no solution $\iff$ L_1 and L_2 are parallel but distinct;
- exactly one solution $\iff$ L_1 and L_2 are not parallel;
- infinitely many solutions $\iff$ L_1 and L_2 are the same line.

1.1.11.2 In Three Variables of Two Equations

Given a linear system in variables x, y, z of two equations

$$\begin{cases} a_1x + b_1y + c_1z = d_1, & (P_1) \\ a_2x + b_2y + c_2z = d_2, & (P_2) \end{cases}$$

where a_1, b_1, c_1 are not all zero and a_2, b_2, c_2 are not all zero. The system has

- no solution $\iff$ P_1 and P_2 are parallel but distinct

  $\iff \dfrac{a_1}{a_2} = \dfrac{b_1}{b_2} = \dfrac{c_1}{c_2}$;

- infinitely many solutions $\iff$ P_1 and P_2 intersect at a straight line

  $\iff \dfrac{a_1}{a_2}, \dfrac{b_1}{b_2}, \dfrac{c_1}{c_2}$ are not all the same;

- infinitely many solutions $\iff$ P_1 and P_2 are the same plane

$$\iff \frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2}.$$

1.2 Elementary Row Operations

1.2.1 Augmented Matrix Representation of a Linear System

Definition. Given a linear system in variables $x_1, x_2, \ldots, x_n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

The rectangular array of constants

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right)$$

is called the **augmented matrix** of the linear system.

1.2.2 Elementary Row Operations on Augmented Matrices

Definition. The **elementary row operations** are the following operations on rows of an augmented matrix:

Description of Operation	Notation
Multiply the ith row by a nonzero constant k	kR_i
Interchange the ith and jth rows	$R_i \leftrightarrow R_j$
Add k times the ith row to the jth row	$R_j + kR_i$

1.2.2.1 Correspondence to Operations on Equations in Linear System

Each elementary row operation corresponds to operations on the equations of the linear system as follows:

Elementary Row Operation	Equation Operation
kR_i	Multiply the ith equation by a nonzero constant k
$R_i \leftrightarrow R_j$	Interchange the ith and jth equations
$R_j + kR_i$	Add k times the ith equation to the jth equation

1.2.2.2 Interchanging two rows can be decomposed further

Interchanging two rows can be obtained by using the other two operations.

$$\begin{pmatrix} a \\ b \end{pmatrix} \xrightarrow{R_1 + R_2} \begin{pmatrix} a+b \\ b \end{pmatrix} \xrightarrow{R_2 - R_1} \begin{pmatrix} a+b \\ -a \end{pmatrix}$$
$$\xrightarrow{R_1 + R_2} \begin{pmatrix} b \\ -a \end{pmatrix} \xrightarrow{-R_2} \begin{pmatrix} b \\ a \end{pmatrix}$$

1.2.2.3 Inverse Elementary Row Operations

Each **elementary row operation** has an inverse operation, which undoes the said **elementary row operation** and is also an **elementary row operation**, as follows:

Elementary Row Operation	**Its Inverse**
$\boldsymbol{A} \xrightarrow{kR_i} \boldsymbol{B}$	$\boldsymbol{B} \xrightarrow{\frac{1}{k}R_i} \boldsymbol{A}$
$\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B}$	$\boldsymbol{B} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{A}$
$\boldsymbol{A} \xrightarrow{R_j + kR_i} \boldsymbol{B}$	$\boldsymbol{B} \xrightarrow{R_j - kR_i} \boldsymbol{A}$

1.2.3 Row Equivalent Matrices

Definition. Two **augmented matrices** are **row equivalent** if one can be obtained from the other by a **series** of **elementary row operations**.

1.2.4 Row Equivalence as an Equivalence Relation

Theorem. Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ be any **augmented matrices**.

- $\boldsymbol{A}$ is row equivalent to $\boldsymbol{A}$ (reflexive);
- $\boldsymbol{A}$ is row equivalent to $\boldsymbol{B} \iff \boldsymbol{B}$ is row equivalent to $\boldsymbol{A}$ (symmetric);
- $\boldsymbol{A}$ is row equivalent to $\boldsymbol{B}$ and $\boldsymbol{B}$ is row equivalent to $\boldsymbol{C} \implies \boldsymbol{A}$ is row equivalent to $\boldsymbol{C}$ (transitive).

Therefore **row equivalence** is an **equivalence relation**.

1.2.5 Row Equivalence Implies Same Solution Set

Theorem. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be **augmented matrices** of two **linear systems**. Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are **row equivalent**.

- Then the corresponding linear systems have the same solution set.

1.3 Row-Echelon Form

1.3.1 Leading Entry

Definition. The **leading entry** for any **nonzero row** of any **augmented matrix** is the first nonzero number of that row, from leftmost to rightmost column.

1.3.2 Row-Echelon Form (REF)

Definition. An **augmented matrix** is in **row-echelon form (REF)** if the following properties are satisfied:

- The **zero rows** are grouped together at the bottom;
- For any two successive **nonzero rows**, the **leading entry** in the lower row appears to the right of the **leading entry** in the higher row.

1.3.3 Pivot Points, Pivot Columns, and Non-pivot Columns

Definition. Suppose an **augmented matrix** is in **row-echelon form**.

1.3.3.1 Pivot Point

Then the **leading entry** of a **nonzero row** is a **pivot point**.

1.3.3.2 Pivot and Non-pivot Columns

A column of the **augmented matrix** is called a

- **pivot column** if it contains a **pivot point**;
- **non-pivot column** if it contains no **pivot point**.

1.3.3.3 Each Pivot Column Contains Exactly One Pivot Point

By the second property of the **row-echelon form**, every **pivot column** contains exactly one **pivot point**.

1.3.4 Reduced Row-Echelon Form (RREF)

Definition. Suppose an **augmented matrix** is in **row-echelon form**. It is in **reduced row-echelon form (RREF)** if

- The **leading entry**, or equivalently the **pivot point**, of every **nonzero row** is 1;
- In each **pivot column**, all entries except the **pivot point** are 0.

1.3.5 General Solution From Row-echelon Form

Suppose that the **augmented matrix** corresponding to a **linear system** is in **row-echelon form**. Then the **general solution** of the **linear system** can be obtained by the following algorithm:

1. **Set** the variables corresponding to **non-pivot columns** to be arbitrary parameters.
2. **Solve** the variables corresponding to **pivot columns** by **back substitution** (from the bottom equation to the top).

1.3.6 General Solution From Reduced Row-echelon Form

Suppose that the **augmented matrix** corresponding to a **linear system** is in **reduced row-echelon form**. Then the **general solution** of the **linear system** can be obtained by the following algorithm:

1. **Set** the variables corresponding to **non-pivot columns** to be arbitrary parameters.
2. **Solve** the variables corresponding to **pivot columns** in any order.

1.4 Gaussian Elimination

1.4.1 REF/RREF of Augmented Matrics

Definition. Let $\boldsymbol{A}$ and $\boldsymbol{R}$ be **augmented matrices**. Suppose that $\boldsymbol{A}$ is **row equivalent** to $\boldsymbol{R}$.

- If $\boldsymbol{R}$ is in **row-echelon form**,
    1. then $\boldsymbol{R}$ is a **row-echelon form** of $\boldsymbol{A}$;
- If $\boldsymbol{R}$ is in **reduced row-echelon form**,
    1. then $\boldsymbol{R}$ is a **reduced row-echelon form** of $\boldsymbol{A}$.

Therefore, by section 1.16 and 1.21, solving a **linear system** with **augmented matrix** $\boldsymbol{A}$ $\iff$ solving a **linear system** with **augmented matrix** that is the **REF/RREF** of $\boldsymbol{A}$.

1.4.2 Gaussian Elimination: Finding REF of Augmented Matrices

Given an **augmented matrix**, we can find its **row-echelon form** by an algorithm called **Gaussian Elimination**. The algorithm is as follows:

1. Find the **leftmost column** which is not entirely zero.
2. Check the **top entry** of such column. If it is 0,

- replace it by a nonzero number by interchanging the top row with another row below.

3. For **each row below** the top row,

   - add a suitable multiple of the **top row** to it so that its **leading entry** becomes 0.

4. If the entire matrix is not in **row-echelon form**,

   - then cover the top row and repeat steps 1-3 to the remaining matrix.

1.4.3 Gauss-Jordan Elimination: Finding RREF of Augmented Matrices

Given an **augmented matrix**, we can find its **reduced row-echelon form** by an algorithm called **Gauss-Jordan Elimination**. The algorithm is as follows:

1. Use **Gaussian Elimination** to get a **row-echelon form**.

2. For **each nonzero row**, multiply a suitable constant so that the **pivot point** becomes 1.

3. For each row, starting from the last nonzero row and working backwards,

   - Add a suitable multiple of the current row to each of the rows above to introduce 0 above the **pivot point** of the current row.

1.4.4 Infinitely Many REFs

Every nonzero matrix has infinitely many non-reduced **row-echelon forms**.

1.4.5 Uniqueness of RREF

Every matrix has a unique **reduced row-echelon form**.

1.4.6 Deducing Consistency of Linear Systems from its REF

Suppose that $\boldsymbol{A}$ is the **augmented matrix** of a **linear system**, and $\boldsymbol{R}$ is a **row-echelon form** of $\boldsymbol{A}$.

1.4.6.1 Inconsistent With No Solutions

The **linear system** is **inconsistent**, i.e. **no solution** if

- the last column is a **pivot column** $\iff$ the last nonzero row has a **pivot point** in the last column.

1.4.6.2 Consistent With One Solution

The **linear system** is **consistent** with exactly one solution if

- the last column is a **non-pivot column**, and
- all other columns are **pivot columns**.

1.4.6.3 Consistent With Infinitely Many Solutions & Arbitrary Parameters

The **linear system** is **consistent** with infinitely many solution if

- the last column is a **non-pivot column**, and
- some other column(s) is/are **non-pivot column(s)**.

The number of arbitrary parameters = the number of **non-pivot columns** excluding the last column.

1.4.7 Geometrical Interpretation of Linear Systems with RREF

1.4.7.1 Linear System in Three Variables of Three Equations

Given a linear system in variables x, y, z of three equations:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 & (P_1), \\ a_{21}x + a_{22}y + a_{23}z = b_2 & (P_2), \\ a_{31}x + a_{32}y + a_{33}z = b_3 & (P_3) \end{cases}$$

where a_{i1}, a_{i2}, a_{i3} are not all zero for $i = 1, 2, 3$. The **reduced row-echelon form $\boldsymbol{R}$** has three rows and four coloumns. The following table summarises the possible solutions sets:

Consistent	**Last Column**	**Other Pivot Columns**	**Other Non-pivot Columns / Arbitrary Parameters**	**Intersection**
No	Pivot	0-2	1-3	null set
Yes	Non-pivot	1	2	plane (three planes coincide)
Yes	Non-pivot	2	1	line
Yes	Non-pivot	3	0	point

1.5 Homogeneous Linear Systems

1.5.1 Homogeneous Linear Equation

Definition. A **linear equation** in variables $x_1, x_2, \ldots, x_n$ is **homogeneous** if it is of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$$

1.5.1.1 Remarks

A **linear equation** is homogeneous $\iff x_1 = 0, x_2 = 0, \ldots x_n = 0$ is a solution.

1.5.2 Geometrical Interpretation of Homogeneous Linear Equations

1.5.2.1 In Two Variables

The solution set of the linear equation

$$ax + by = 0 \text{ (in } x, y)$$

where a and b are not both zero represents a **straight line**, in the xy-plane, that passes through the origin $O(0,0)$.

1.5.2.2 In Three Variables

The solution set of the linear equation

$$ax + by + cz = 0 \text{ (in } x, y, z)$$

where a, b, c are not all zero represents a **plane**, in the xyz-space, that contains the origin $O(0,0,0)$.

1.5.3 Homogeneous Linear System

Definition. A **linear system** is **homogeneous** if every linear equation of the system is homogeneous. That is the linear system is of the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases}$$

1.5.4 Trivial Solution of Homogeneous Linear Systems

A **linear system** in $x_1, x_2, \ldots, x_n$ is **homogeneous** $\iff x_1 = 0, x_2 = 0, \ldots, x_n = 0$ is a solution.

- This is the **trivial solution** of a **homogeneous linear system**
- Other solutions are called **non-trivial solutions**.

1.5.5 Geometric Interpretation of Homogeneous Linear Systems

1.5.5.1 In Two Variables of Two Equations

Given a **homogeneous linear system** in variables x, y of two equations

$$\begin{cases} a_1x + b_1y = 0, & (L_1) \\ a_2x + b_2y = 0, & (L_2) \end{cases}$$

where a_1, b_1 are not both zero and a_2, b_2 are not both zero. The system has

- only the **trivial solution** $\iff$ L_1 and L_2 are different;
- **non-trivial solutions** $\iff$ L_1 and L_2 are the same.

1.5.5.2 In Three Variables of Two Equations

Given a linear system in variables x, y, z of two equations

$$\begin{cases} a_1x + b_1y + c_1z = 0, & (P_1) \\ a_2x + b_2y + c_2z = 0, & (P_2) \end{cases}$$

where a_1, b_1, c_1 are not all zero and a_2, b_2, c_2 are not all zero. The system has infinitely many solutions with only two cases

- The two planes are the same, or
- The two planes intersect at a straight line passing through $O(0, 0, 0)$.

Chapter 2

Matrices

2.1 Introduction to Matrices

2.1.1 Definition of Matrix

Definition. A **matrix** (plural **matrices**) is a rectangular array of numbers

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

where

- m is the number of **rows** in the matrix.
- n is the number of **columns** in the matrix.
- The **size** of the matrix is given by $\boldsymbol{m \times n}$.
- The $\boldsymbol{(i, j)}$**-entry** is the entry in the ith row and jth column.
  - In the given matrix, the (i, j)-entry is $\boldsymbol{a_{ij}}$.

2.1.1.1 Remarks

- Some books use $[\dots]$ instead of $(\dots)$.
- A 1×1 matrix is usually treated as a real number in computation.

2.1.2 Notation of Matrices

A **matrix** is usually denoted by capital letters $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \ldots$. Given a

$$\boldsymbol{m \times n} \text{ matrix } \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

- a_{ij} is the (i, j)-entry of $\boldsymbol{A}$.
- The matrix is denoted by $\boldsymbol{A = (a_{ij})_{m \times n}}$
- If the size of $\boldsymbol{A}$ is known (or not important)
  - The matrix can also be denoted by $\boldsymbol{A = (a_{ij})}$.

2.1.3 Row Matrix

Definition. A **row matrix** (**row vector**) is a **matrix** with only one **row**.

2.1.4 Column Matrix

Definition. A **column matrix** (**column vector**) is a **matrix** with only one **column**.

2.1.5 Square Matrix

Definition. A **square matrix** is a **matrix** with the same number of **rows** and **columns**.

2.1.5.1 Order of Square Matrix

Definition. An $n \times n$ matrix is a **square matrix** of **order** n.

2.1.5.2 Diagonal / Diagonal Entries of Square Matrix

Definition. Let $\boldsymbol{A} = (a_{ij})$ be a **square matrix** of order n.

- The **diagonal/principle diagonal/major diagonal** of $\boldsymbol{A}$ is the **sequence** of entries
  - $\boldsymbol{a_{11}, a_{22}, \ldots, a_{nn}}$
- The entries a_{ii}, for $i = 1, \ldots, n$ are the **diagonal entries**
- The entries a_{ij}, $i \neq j$ are the **non-diagonal entries**
- The **anti-diagonal/minor diagonal** of $\boldsymbol{A}$ is the **sequence** of entries from the right top to the left bottom
  - $\boldsymbol{a_{1n}, a_{2,n-1}, \ldots, a_{n1}}$

2.1.6 Diagonal Matrix

Definition. A **square matrix** is a **diagonal matrix** if all its **non-diagonal entries** are zero.

- $\boldsymbol{A} = (a_{ij})_{n \times n}$ is **diagonal** $\iff a_{ij} = 0$ for all $i \neq j$

2.1.7 Scalar Matrix

Definition. A **diagonal matrix** is a **scalar matrix** if all its **diagonal entries** are the same.

- $\boldsymbol{A} = (a_{ij})_{n \times n}$ is **scalar** $\iff a_{ij} = \begin{cases} c & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

2.1.8 Identity Matrix

Definition. A **scalar matrix** is an **identity matrix** if all its **diagonal entries** are 1.

- $\boldsymbol{A} = (a_{ij})_{n \times n}$ is **identity** $\iff a_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$

Note: There is exactly one identity matrix of order n.

2.1.8.1 Notation for Identity Matrix

- The **identity matrix** of **order** n is denoted by $\boldsymbol{I}_n$.
- If no confusion in order, we may write $\boldsymbol{I}$ instead of $\boldsymbol{I}_n$.

2.1.9 Zero Matrix

Definition. A **matrix** with all entries equal to zero is a **zero matrix**.

- $\boldsymbol{A} = (a_{ij})_{m \times n}$ is **zero** $\iff a_{ij} = 0$ for all i, j

Note: There is exactly one zero matrix of size $m \times n$.

2.1.9.1 Notation for Zero Matrix

- The **zero matrix** of **size** $m \times n$ is denoted by $\mathbf{0}_{m \times n}$.
- If no confusion in size, we may write $\mathbf{0}$ instead of $\mathbf{0}_{m \times n}$.

2.1.10 Symmetric Square Matrix

Definition. A **square matrix** is **symmetric** if it is symmetric with respect to the diagonal.

- $\boldsymbol{A} = (a_{ij})_{n \times n}$ is **symmetric** $\iff a_{ij} = a_{ji}$ for all i, j
  - There is no restriction to the **diagonal entries**.

2.1.11 Upper Triangular Square Matrix

Definition. A **square matrix** is **upper triangular** if all the entries **below** the diagonal are zero.

- $\boldsymbol{A} = (a_{ij})_{n\times n}$ is **upper triangular** $\Longleftrightarrow$ $a_{ij} = 0$ if $i > j$
  - There is no restriction to the **diagonal entries**.

2.1.12 Lower Triangular Square Matrix

Definition. A **square matrix** is **lower triangular** if all the entries **above** the diagonal are zero.

- $\boldsymbol{A} = (a_{ij})_{n\times n}$ is **lower triangular** $\Longleftrightarrow$ $a_{ij} = 0$ if $i < j$
  - There is no restriction to the **diagonal entries**.

2.1.13 Triangular Square Matrix

Definition. Both **upper triangular matrices** and **lower triangular matrices** are **triangular matrices**.

2.1.14 Diagonal and Triangular Square Matrices

A **square matrix** is both upper and lower triangular $\Longleftrightarrow$ it is diagonal.

2.2 Matrix Operations

2.2.1 Identical Matrices

- A matrix is completely determined by its **size** and **entries**.
- **Definition.** Two matrices are **equal** if
  - they have the same **size** (same number of rows and same number of columns), and
  - all the corresponding entries are the same.

  Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{p\times q}$. Then

  - $\boldsymbol{A} = \boldsymbol{B} \Longleftrightarrow m = p \;\&\; n = q \;\&\; a_{ij} = b_{ij}$ for all i, j

2.2.2 Matrix Addition, Subtraction & Scalar Multiplication

Definition. Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ and $\boldsymbol{B} = (b_{ij})_{m\times n}$ be matrices, and c a constant. The following operations are defined:

- **Addition**: $\boldsymbol{A} + \boldsymbol{B} = (a_{ij} + b_{ij})_{m\times n}$
- **Substraction**: $\boldsymbol{A} - \boldsymbol{B} = (a_{ij} - b_{ij})_{m\times n}$
- **Scalar Multiplication**: $c\boldsymbol{A} = (ca_{ij})_{m\times n}$

Remarks.

- $(-1)\boldsymbol{A}$ is usually denoted by $-\boldsymbol{A}$.
- It can be proved that $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B})$.
  - In the discussion we usually only consider addition and scalar multiplication.

2.2.3 Properties of Matrix Addition, Subtraction & Scalar Multiplication

Theorem. Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ be matrices of the same size, and c, d be constants. Then

- $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B})$
- Commutative Law for Matrix Addition:
  - $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$
- Associative Law for Matrix Addition:
  - $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$
- Let $\mathbf{0}$ be the zero matrix of the same size as $\boldsymbol{A}$. Then
  - $\mathbf{0} + \boldsymbol{A} = \boldsymbol{A}; \quad \boldsymbol{A} - \boldsymbol{A} = \mathbf{0}; \quad 0\boldsymbol{A} = \mathbf{0}; \quad c\mathbf{0} = \mathbf{0}.$
- Distributive Law for Scalar Multiplication over Addition:
  - $c(\boldsymbol{A} + \boldsymbol{B}) = c\boldsymbol{A} + c\boldsymbol{B}$
  - $(c + d)\boldsymbol{A} = c\boldsymbol{A} + d\boldsymbol{A}$
- $c(d\boldsymbol{A}) = (cd)\boldsymbol{A}, \quad 1\boldsymbol{A} = \boldsymbol{A}$

2.2.4 Matrix Multiplication

Definition. Let $\boldsymbol{A} = (a_{ij})_{m\times p}$ and $\boldsymbol{B} = (b_{ij})_{p\times n}$. Then $\boldsymbol{AB}$ is the $m \times n$ matrix such that its (i,j)-entry is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

Note: No. of columns of $\boldsymbol{A}$ = the no. of rows of $\boldsymbol{B}$.

2.2.4.1 Matrix Multiplication Explained in Words

In order to get the (i,j)-entry of the **product** matrix:

1. Find the ith **row** of the first matrix;
2. Find the jth **column** of the second matrix;
3. Multiply the corresponding entries;
4. Add the products together.

2.2.5 Noncommutativity of Matrix Multiplication

Matrix Multiplication is **not commutative** in general. For example, let $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. $\boldsymbol{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{BA}$.

- $\boldsymbol{AB}$ is the **pre-multiplication** of $\boldsymbol{A}$ to $\boldsymbol{B}$ (to $\boldsymbol{B}$ by $\boldsymbol{A}$).
- $\boldsymbol{BA}$ is the **post-multiplication** of $\boldsymbol{A}$ to $\boldsymbol{B}$ (to $\boldsymbol{B}$ by $\boldsymbol{A}$).

2.2.6 Properties of Matrix Multiplication

Theorem.

- Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ be $m \times p, p \times q, q \times n$ matrices, resp. Then
  - Associative Law: $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$.
- Let $\boldsymbol{A}$ be a $m \times p$ matrix, $\boldsymbol{B}_1, \boldsymbol{B}_2$ be $p \times n$ matrices. Then
  - Distributive Law: $\boldsymbol{A}(\boldsymbol{B}_1 + \boldsymbol{B}_2) = \boldsymbol{AB}_1 + \boldsymbol{AB}_2$.
- Let $\boldsymbol{A}_1, \boldsymbol{A}_2$ be $m \times p$ matrices, $\boldsymbol{B}$ be a $p \times n$ matrix. Then
  - Distributive Law: $(\boldsymbol{A}_1 + \boldsymbol{A}_2)\boldsymbol{B} = \boldsymbol{A}_1\boldsymbol{B} + \boldsymbol{A}_2\boldsymbol{B}$.
- Let $\boldsymbol{A}, \boldsymbol{B}$ be $m \times p, p \times n$ matrices resp. and c be a constant. Then

  - $c(\boldsymbol{AB}) = (c\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(c\boldsymbol{B})$.

- Let $\boldsymbol{A}$ be a $m \times n$ matrix. Then
  - $\boldsymbol{A}\mathbf{0}_{n\times p} = \mathbf{0}_{m\times p}$; $\quad \mathbf{0}_{p\times m}\boldsymbol{A} = \mathbf{0}_{p\times n}$.
  - $\boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$; $\quad \boldsymbol{I}_m\boldsymbol{A} = \boldsymbol{A}$.

2.2.7 Nonnegative Integer Powers of Square Matrices

2.2.7.1 Multiplying a Matrix With Itself

Let $\boldsymbol{A}$ be a $m \times n$ matrix. Then

- $\boldsymbol{AA}$ is well-defined $\iff m = n \iff \boldsymbol{A}$ is a **square matrix**.

2.2.7.2 Nonnegative Integer Powers of Square Matrices

Definition. Let $\boldsymbol{A}$ be a **square matrix** of order n. For nonnegative integers k, the **powers** of $\boldsymbol{A}$ are defined as

$$\boldsymbol{A}^k = \begin{cases} \boldsymbol{I}_n & \text{if } k = 0 \\ \underbrace{\boldsymbol{AA}\ldots\boldsymbol{A}}_{k \text{ times}} & \text{if } k \geq 1 \end{cases}$$

2.2.8 Properties of Nonnegative Integer Powers of Square Matrices

Let $\boldsymbol{A}, \boldsymbol{B}$ be square matrices of the same size, and m, n nonnegative integers. Then

- $\boldsymbol{A}^m\boldsymbol{A}^n = \boldsymbol{A}^{m+n}$.
- In general, $(\boldsymbol{AB})^n \neq \boldsymbol{A}^n\boldsymbol{B}^n$ for $n = 2, 3, \ldots$
  - Since matrix multiplication is not commutative,

$$\begin{aligned}(\boldsymbol{AB})^n &= \underbrace{(\boldsymbol{AB})(\boldsymbol{AB})\ldots(\boldsymbol{AB})}_{n \text{ times}} \\ &\neq \underbrace{\boldsymbol{AA}\ldots\boldsymbol{A}}_{n \text{ times}}\underbrace{\boldsymbol{BB}\ldots\boldsymbol{B}}_{n \text{ times}} \\ &= \boldsymbol{A}^n\boldsymbol{B}^n\end{aligned}$$

- However, suppose that $\boldsymbol{AB} = \boldsymbol{BA}$. Then
  - $(\boldsymbol{AB})^n = \boldsymbol{A}^n\boldsymbol{B}^n$ for all nonnegative integers n.

2.2.9 Matrix Representation

Let $\boldsymbol{A} = (a_{ij})_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$

2.2.9.1 Matrix as a Column of Row Vectors

- Let $\boldsymbol{a}_i$ denote the ith row of $\boldsymbol{A}$, for $i = 1, 2, \ldots, m$.

$$\boldsymbol{a}_1 = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \end{pmatrix}$$
$$\boldsymbol{a}_2 = \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2n} \end{pmatrix}$$
$$\vdots$$
$$\boldsymbol{a}_m = \begin{pmatrix} a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

- Then each $\boldsymbol{a}_i$ is a $1 \times n$ matrix (row vector).
- Then $\boldsymbol{A}$ can be represented as

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{pmatrix}$$

2.2.9.2 Matrix as a Row of Column Vectors

- Let $\boldsymbol{b}_j$ denote the jth column of $\boldsymbol{A}$, for $j = 1, 2, \ldots, n$.

$$\boldsymbol{b}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \boldsymbol{b}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \boldsymbol{b}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

- Then each $\boldsymbol{b}_j$ is a $m \times 1$ matrix (column vector).
- Then $\boldsymbol{A}$ can be represented as

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \ldots & \boldsymbol{b}_n \end{pmatrix}$$

2.2.10 Decomposing Matrix Multiplication

Suppose $\boldsymbol{A} = (a_{ij})_{m \times p}$.

- Let $\boldsymbol{a}_i = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{ip} \end{pmatrix}$ be the ith row of $\boldsymbol{A}$.

Suppose $\boldsymbol{B} = (b_{ij})_{p \times n}$.

- Let $\boldsymbol{b}_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix}$ be the jth column of $\boldsymbol{B}$.

Then

$$\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{b}_1 & \boldsymbol{a}_1\boldsymbol{b}_2 & \dots & \boldsymbol{a}_1\boldsymbol{b}_n \\ \boldsymbol{a}_2\boldsymbol{b}_1 & \boldsymbol{a}_2\boldsymbol{b}_2 & \dots & \boldsymbol{a}_2\boldsymbol{b}_n \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}_m\boldsymbol{b}_1 & \boldsymbol{a}_m\boldsymbol{b}_2 & \dots & \boldsymbol{a}_m\boldsymbol{b}_n \end{pmatrix}$$

2.2.10.1 Decomposing Into Entries

$$\begin{aligned}(i,j)\text{-entry of } \boldsymbol{AB} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj} \\ &= \boldsymbol{a}_i\boldsymbol{b}_j\end{aligned}$$

2.2.10.2 Decomposing Into Rows

$$\begin{aligned} i\text{th row of } \boldsymbol{AB} &= \begin{pmatrix} \boldsymbol{a}_i\boldsymbol{b}_1 & \boldsymbol{a}_i\boldsymbol{b}_2 & \dots & \boldsymbol{a}_i\boldsymbol{b}_n \end{pmatrix} \\ &= \boldsymbol{a}_i \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \dots & \boldsymbol{b}_n \end{pmatrix} \\ &= \boldsymbol{a}_i\boldsymbol{B} \end{aligned}$$

Thus, $\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{pmatrix} \boldsymbol{B} = \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{B} \\ \boldsymbol{a}_2\boldsymbol{B} \\ \vdots \\ \boldsymbol{a}_m\boldsymbol{B} \end{pmatrix}$

2.2.10.3 Decomposing Into Columns

$$\begin{aligned} j\text{th column of } \boldsymbol{AB} &= \begin{pmatrix} \boldsymbol{a}_1\boldsymbol{b}_j \\ \boldsymbol{a}_2\boldsymbol{b}_j \\ \vdots \\ \boldsymbol{a}_m\boldsymbol{b}_j \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_m \end{pmatrix} \boldsymbol{b}_j \\ &= \boldsymbol{A}\boldsymbol{b}_j \end{aligned}$$

Thus, $\boldsymbol{AB} = \boldsymbol{A} \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \dots & \boldsymbol{b}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}\boldsymbol{b}_1 & \boldsymbol{A}\boldsymbol{b}_2 & \dots & \boldsymbol{A}\boldsymbol{b}_n \end{pmatrix}$

2.2.10.4 Decomposing Into Blocks

Matrices can be multiplied in blocks (provided that the sizes are matched).

2.2.11 Matrix Representation of Linear Equation

Given a **linear equation** in n variables $x_1, \ldots, x_n$

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

The corresponding matrix representation is

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = b$$

2.2.12 Matrix Representation of Linear System

Given a **linear system** of m equations in n variables $x_1, \dots, x_n$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Then the corresponding matrix representation is as follows:

- The **coefficient matrix** is $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$
- The **variable matrix** is $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- The **constant matrix** is $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$
- Then $\boldsymbol{Ax} = \boldsymbol{b}$

2.2.12.1 Solution Vector to Linear System

Let $\boldsymbol{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$. Then

$$\begin{aligned} & x_1 = u_1, \dots, x_n = u_n \text{ is a solution to the system} \\ \iff & \boldsymbol{Au} = \boldsymbol{b} \\ \iff & \boldsymbol{u} \text{ is a solution to } \boldsymbol{Ax} = \boldsymbol{b} \end{aligned}$$

2.2.12.2 Alternative Matrix Representation of Linear System

Let $\boldsymbol{a}_j$ denote the jth column of $\boldsymbol{A}$. Then

$$\begin{aligned}
\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \\
&= x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n \\
&= \sum_{j=1}^{n} x_j \boldsymbol{a}_j
\end{aligned}$$

2.2.13 Transpose of Matrix

Definition. Let $\boldsymbol{A} = (a_{ij})_{m\times n}$ be a **matrix**. The **transpose** of $\boldsymbol{A}$ is the $n \times m$ matrix $\boldsymbol{A}^T$ (or $\boldsymbol{A}^t$)

- whose (i, j)-entry is a_{ji}.

Remarks.

- The ith row of $\boldsymbol{A}^T$ is the ith column of $\boldsymbol{A}$.
- The jth column of $\boldsymbol{A}^T$ is the jth row of $\boldsymbol{A}$.

2.2.14 Properties of Transpose of Matrix

Theorem. Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then

- $(\boldsymbol{A}^T)^T = \boldsymbol{A}$.
- $\boldsymbol{A}$ is symmetric $\iff$ $\boldsymbol{A} = \boldsymbol{A}^T$.
- Let c be a scalar. Then $(c\boldsymbol{A})^T = c\boldsymbol{A}^T$.
- Let $\boldsymbol{B}$ be $m \times n$. Then $(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$.
- Let $\boldsymbol{B}$ be $n \times p$. Then $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$.
  - **Note:** In general, $(\boldsymbol{A}\boldsymbol{B})^T \neq \boldsymbol{A}^T\boldsymbol{B}^T$.

2.3 Inverse of Square Matrices

2.3.1 Additive Inverse of Matrices

Let $\boldsymbol{A}$ be a **matrix**. Then $-\boldsymbol{A}$ is the **additive inverse** of $\boldsymbol{A}$.

2.3.2 Inverse of Square Matrices

Definition. Let $\boldsymbol{A}$ be a **square matrix** of order n.

- If there exists a **square matrix** $\boldsymbol{B}$ of order n so that
  - $\boldsymbol{AB} = \boldsymbol{I}_n$ and $\boldsymbol{BA} = \boldsymbol{I}_n$,

  then $\boldsymbol{A}$ is **invertible**, and $\boldsymbol{B}$ is an **inverse** of $\boldsymbol{A}$.
- If $\boldsymbol{A}$ is not **invertible**, then $\boldsymbol{A}$ is **singular**.

Note: Non-square matrix is neither **invertible** nor **singular**.

2.3.3 Uniqueness of Inverse of Invertible Matrices

Theorem. Let $\boldsymbol{A}$ be a **square matrix**. If $\boldsymbol{A}$ is **invertible**

- then its **inverse** is unique.
- **Notation.** The unique inverse of $\boldsymbol{A}$, is denoted by $\boldsymbol{A}^{-1}$
  - $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

2.3.4 Cancellation Law for Invertible Matrices

Theorem. Let $\boldsymbol{A}$ be an **invertible matrix**. Then

- $\boldsymbol{AB}_1 = \boldsymbol{AB}_2 \implies \boldsymbol{B}_1 = \boldsymbol{B}_2$.
- $\boldsymbol{C}_1\boldsymbol{A} = \boldsymbol{C}_2\boldsymbol{A} \implies \boldsymbol{C}_1 = \boldsymbol{C}_2$.

Remark. The cancellation law fails if $\boldsymbol{A}$ is **singular**.

2.3.5 Inverse of Square Matrix of Order Two

Theorem. Let $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

- $\boldsymbol{A}$ is **invertible** $\iff ad - bc \neq 0$.
- If $\boldsymbol{A}$ is **invertible**, then $\boldsymbol{A}^{-1} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

2.3.6 Properties of Invertible Matrices

Theorem. Let $\boldsymbol{A}, \boldsymbol{B}$ be **invertible matrices** of same size. Then

- Let $c \neq 0$. $c\boldsymbol{A}$ is **invertible**, and $(c\boldsymbol{A})^{-1} = \frac{1}{c}\boldsymbol{A}^{-1}$.
- $\boldsymbol{A}^T$ is **invertible**, and $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$.
- $\boldsymbol{A}^{-1}$ is **invertible**, and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
- $\boldsymbol{AB}$ is **invertible**, and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ be **invertible matrices** of the same size. Then

- $(\boldsymbol{A}_1\boldsymbol{A}_2 \ldots \boldsymbol{A}_k)^{-1} = \boldsymbol{A}_k^{-1} \ldots \boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}$.

In particular, $(\underbrace{\boldsymbol{AA}\ldots\boldsymbol{A}}_{k \text{ times}})^{-1} = \underbrace{\boldsymbol{A}^{-1}\ldots\boldsymbol{A}^{-1}\boldsymbol{A}^{-1}}_{k \text{ times}}$. Thus,

- $(\boldsymbol{A}^k)^{-1} = (\boldsymbol{A}^{-1})^k$.

2.3.7 Negative Integer Powers of Square Matrices

Definition. Let $\boldsymbol{A}$ be an **invertible matrix**. For any positive integer k, $\boldsymbol{A}^{-k} = (\boldsymbol{A}^{-1})^k$.

2.3.8 Properties of Integer Powers of Square Matrices

Theorem. Let $\boldsymbol{A}$ be an **invertible matrix**. For any integers m and n,

- $\boldsymbol{A}^{m+n} = \boldsymbol{A}^m\boldsymbol{A}^n$ and $(\boldsymbol{A}^m)^n = \boldsymbol{A}^{mn}$.

Note: If $\boldsymbol{A}$ is **singular**, then $\boldsymbol{A}^{-1}$ is undefined.

2.4 Elementary Matrices

2.4.1 Definition of Elementary Matrices

Definition. A **square matrix** is called an **elementary matrix** if it can be obtained from the **identity matrix** by performing a single **elementary row operation**.

2.4.2 Connection of Elementary Matrices to Pre-Matrix Multiplication

Theorem. Let $\boldsymbol{E}$ be an **elementary matrix** obtained by performing an **elementary row operation** on $\boldsymbol{I}_m$. Then for any $m \times n$ matrix $\boldsymbol{A}$, $\boldsymbol{EA}$ can be obtained by performing the same **elementary row operation** to $\boldsymbol{A}$.

- Let $\boldsymbol{A}$ be an $m \times n$ matrix.

- $\boldsymbol{I}_m \xrightarrow{cR_i} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{EA}$.
- $\boldsymbol{I}_m \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{EA}$.
- $\boldsymbol{I}_m \xrightarrow{R_i + cR_j} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{EA}$.

2.4.3 Elementary Matrices are Invertible

Theorem.

- Every **elementary matrix** is **invertible**.
- The **inverse** of an **elementary matrix** is **elementary**.

Let $\boldsymbol{E}$ be an **elementary matrix**.

- $\boldsymbol{I} \xrightarrow{cR_i} \boldsymbol{E} \implies \boldsymbol{I} \xrightarrow{\frac{1}{c}R_i} \boldsymbol{E}^{-1}$.
- $\boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \implies \boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E}^{-1}$. (So $\boldsymbol{E} = \boldsymbol{E}^{-1}$).
- $\boldsymbol{I} \xrightarrow{R_i + cR_j} \boldsymbol{E} \implies \boldsymbol{I} \xrightarrow{R_i - cR_j} \boldsymbol{E}^{-1}$.

2.4.4 Row Equivalent Matrices are Linked by Elementary Matrices

Theorem. Two **matrices** $\boldsymbol{A}$ and $\boldsymbol{B}$ are **row equivalent** $\Longleftrightarrow$ there exist **elementary matrices** $\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_k$ such that $\boldsymbol{B} = \boldsymbol{E}_k\boldsymbol{E}_{k-1}\ldots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}$.
Remarks. Suppose that $\boldsymbol{B} = \boldsymbol{E}_k\boldsymbol{E}_{k-1}\ldots\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}$. Then $\boldsymbol{A} = \boldsymbol{E}_1^{-1}\boldsymbol{E}_2^{-1}\ldots\boldsymbol{E}_{k-1}^{-1}\boldsymbol{E}_k^{-1}\boldsymbol{B}$

2.4.5 Main Theorem for Invertible Matrices

Theorem. Let $\boldsymbol{A}$ be a **square matrix**. Then the following are equivalent:

1. $\boldsymbol{A}$ is an **invertible matrix**.
2. Linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution.
3. Linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution.
4. The **reduced row-echelon form** of $\boldsymbol{A}$ is $\boldsymbol{I}$.
5. $\boldsymbol{A}$ is the product of **elementary matrices**.

2.4.6 Checking Whether a Square Matrix is Invertible

- A **square matrix** is **invertible**

$$\begin{aligned}&\iff \text{Its reduced row-echelon form is } \boldsymbol{I}\\&\iff \text{All the columns in its row-echelon form are pivot.}\\&\iff \text{All the rows in its row-echelon form are nonzero.}\end{aligned}$$

- A **square matrix** is **singular**

$$\begin{aligned}&\iff \text{Its reduced row-echelon form is not } \boldsymbol{I}\\&\iff \text{Some columns in its row-echelon form are non-pivot.}\\&\iff \text{Some rows in its row-echelon form are zero.}\end{aligned}$$

2.4.7 Finding Inverse of an Invertible Matrix

Theorem. Let $\boldsymbol{A}$ be an **invertible matrix**. The **reduced row-echelon form** of $\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I} \end{array}\right)$ is $\left(\begin{array}{c|c} \boldsymbol{I} & \boldsymbol{A}^{-1} \end{array}\right)$.

2.4.8 Weaker Requirement for Invertible Matrices

Theorem. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be **square matrices** of the same size. If $\boldsymbol{AB} = \boldsymbol{I}$, then $\boldsymbol{A}$ and $\boldsymbol{B}$ are **invertible**, and $\boldsymbol{A}^{-1} = \boldsymbol{B}$, $\boldsymbol{B}^{-1} = \boldsymbol{A}$.

2.4.9 Invertibility of Product of Matrices

Corollary. Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$ be **square matrices** of the same size. Then

- $\boldsymbol{A}_1\boldsymbol{A}_2 \ldots \boldsymbol{A}_k$ is **invertible** $\iff$ all $\boldsymbol{A}_i$ are **invertible**.
- $\boldsymbol{A}_1\boldsymbol{A}_2 \ldots \boldsymbol{A}_k$ is **singular** $\iff$ some $\boldsymbol{A}_i$ are **singular**.

2.4.10 Elementary Column Operations

Definition. The **elementary column operations** are the following operations on columns of a matrix:

Description of Operation	Notation
Multiply the ith column by a nonzero constant k	kC_i
Interchange the ith and jth columns	$C_i \leftrightarrow C_j$
Add k times the jth column to the ith column	$C_i + kC_j$

2.4.11 Performing a Single Elementary Column Operation on Identity Matrix Gives an Elementary Matrix

Let $\boldsymbol{E}$ be the **matrix** obtained from $\boldsymbol{I}$ by a single **elementary column operation**. Then $\boldsymbol{E}$ is an **elementary matrix** (i.e. $\boldsymbol{E}$ can be obtained from $\boldsymbol{I}$ by a single **elementary row operation**).

- $\boldsymbol{I} \xrightarrow{kC_i} \boldsymbol{E} \iff \boldsymbol{I} \xrightarrow{kR_i} \boldsymbol{E}$.
- $\boldsymbol{I} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{E} \iff \boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E}$.
- $\boldsymbol{I} \xrightarrow{C_i + kC_j} \boldsymbol{E} \iff \boldsymbol{I} \xrightarrow{R_j + kR_i} \boldsymbol{E}$.

2.4.12 Connection of Elementary Matrices to Matrix Post-Multiplication

Theorem. Let $\boldsymbol{E}$ be an **elementary matrix** obtained by performing an **elementary column operation** on $\boldsymbol{I}_n$. Then for any $m \times n$ matrix $\boldsymbol{A}$, $\boldsymbol{AE}$ can be obtained by performing the same **elementary column operation** to $\boldsymbol{A}$.

- $\boldsymbol{I} \xrightarrow{kC_i} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{kC_i} \boldsymbol{AE}$.
- $\boldsymbol{I} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{C_i \leftrightarrow C_j} \boldsymbol{AE}$.
- $\boldsymbol{I} \xrightarrow{C_i + kC_j} \boldsymbol{E} \implies \boldsymbol{A} \xrightarrow{C_i + kC_j} \boldsymbol{AE}$.

2.5 Determinant

2.5.1 (i, j)-cofactor

Definition. Let $\boldsymbol{A} = (a_{ij})_{n\times n}$. Let $\boldsymbol{M}_{ij}$ be the **submatrix** obtained from $\boldsymbol{A}$ by deleting the ith row and jth column. Then for $1 \le i, j \le n$, the **(i, j)-cofactor** of $\boldsymbol{A}$ is

$$A_{ij} = (-1)^{i+j} \det(\boldsymbol{M}_{ij})$$

2.5.2 Definition of Determinant of Square Matrix

Definition. Let $\boldsymbol{A} = (a_{ij})_{n\times n}$. Its **determinant** is:

- If $n = 1$, define $\det(\boldsymbol{A}) = a_{11}$.
- If $n > 1$, let A_{ij} be its (i, j)-cofactor, define
  - $\det(\boldsymbol{A}) = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}$

2.5.3 Determinant of 2×2 Matrix

Let $\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\det(A) = ad - bc$.

2.5.4 Broken Diagonal Formula for Determinant of 3×3 Matrix

Let $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Then

$$\det(\boldsymbol{A}) = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{22}a_{31})$$

- The positive terms come from the
  - 3 (broken) diagonals from the top left to the bottom right.
- The negative terms come from the
  - 3 (broken) diagonals from the top right to the bottom left.

Warning: The "diagonal expansion" of $\det(A)$ for 2×2 or 3×3 matrices is not valid if the order ≥ 4.

2.5.5 Properties of Determinant

2.5.5.1 Properties of Determinant

Theorem. For any **square matrices** $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same size and any scalar c,

- For $n \in \mathbb{N}$, $\det(\boldsymbol{I}_n) = 1$.
- $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$.
- $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.
- $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$, where $\boldsymbol{A}$ is $n \times n$.
- $\det(\boldsymbol{A}^{-1}) = \det(\boldsymbol{A})^{-1}$ if $\boldsymbol{A}$ is **invertible**.

2.5.5.2 Change in Determinant by Elementary Row Operations

Theorem. For any square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$,

- $\boldsymbol{A} \xrightarrow{cR_i} \boldsymbol{B} \implies \det(\boldsymbol{B}) = c\det(\boldsymbol{A})$.
  - In particular, $\boldsymbol{I} \xrightarrow{cR_i} \boldsymbol{E} \implies \det(\boldsymbol{E}) = c$.
- $\boldsymbol{A} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{B} \implies \det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.

- In particular, $\boldsymbol{I} \xrightarrow{R_i \leftrightarrow R_j} \boldsymbol{E} \implies \det(\boldsymbol{E}) = -1$.

- $\boldsymbol{A} \xrightarrow{R_i + cR_j} \boldsymbol{B} \implies \det(\boldsymbol{B}) = \det(\boldsymbol{A})$.
  - In particular, $\boldsymbol{I} \xrightarrow{R_i + cR_j} \boldsymbol{E} \implies \det(\boldsymbol{E}) = 1$.

2.5.5.3 Change in Determinant by Elementary Matrices

Theorem. Let $\boldsymbol{A}$ be a **square matrix**. For any **elementary matrix** of the same order,

$$\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$$

2.5.5.4 Change in Determinant Between Row Equivalent Matrices

Theorem. Suppose **square matrices** $\boldsymbol{A}$ and $\boldsymbol{B}$ are **row equivalent**. Then there exist **elementary matrices** $\boldsymbol{E}_1, \boldsymbol{E}_2, \dots, \boldsymbol{E}_k$ such that

$$\boldsymbol{B} = \boldsymbol{E}_k \dots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}$$

Then

$$\det(\boldsymbol{B}) = \det(\boldsymbol{E}_k) \dots \det(\boldsymbol{E}_2)\det(\boldsymbol{E}_1)\det(\boldsymbol{A})$$

Since $\det(\boldsymbol{E}) \neq 0$ for every **elementary matrix** $\boldsymbol{E}$,

- $\det(\boldsymbol{A}) = 0 \iff \det(\boldsymbol{B}) = 0$;
- Equivalently, $\det(\boldsymbol{A}) \neq 0 \iff \det(\boldsymbol{B}) \neq 0$.

2.5.5.5 Determinant of Square Matrix With Zero Row

Theorem. Suppose a **square matrix** $\boldsymbol{A}$ has a **zero row**. Then $\det(\boldsymbol{A}) = 0$.

2.5.5.6 Invertibility and Determinant

Theorem. For any **square matrix** $\boldsymbol{A}$

- $\det(\boldsymbol{A}) = 0 \iff \boldsymbol{A}$ is **singular**;
- Equivalently, $\det(\boldsymbol{A}) \neq 0 \iff \boldsymbol{A}$ is **invertible**,

2.5.5.7 Determinant of Triangular Matrix

Theorem. Suppose $\boldsymbol{A} = (a_{ij})_{n\times n}$ is **triangular**. Then

$$\det(\boldsymbol{A}) = a_{11}a_{22}\dots a_{nn}$$

Remarks.

- Note that a **row-echelon form** of a **square matrix** is always **upper triangular**.
  - To find the **determinant** using **elementary row operation**, it suffices to use **Gaussian elimination** to get a **row-echelon form**.

2.5.6 Cofactor Expansion

Theorem. Let $\boldsymbol{A} = (a_{ij})_{n\times n}$ and A_{ij} denote the (i,j)-cofactor of $\boldsymbol{A}$. Then for $1 \le i \le n$,

- $\det(\boldsymbol{A}) = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$
- This is called the **cofactor expansion along the ith row**.

and for $1 \le j \le n$,

- $\det(\boldsymbol{A}) = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}$
- This is called the **cofactor expansion along the jth column**.

Remarks.

- In evaluating the **determinant** using **cofactor expansion**,
  - expand along the row or column with the most zeros.

2.5.7 Finding Determinant Efficiently

Given $\boldsymbol{A} = (a_{ij})_{n\times n}$, $\det(\boldsymbol{A})$ may be found as follows:

- $\boldsymbol{A}$ has a zero row/column $\implies \det(\boldsymbol{A}) = 0$.
- $\boldsymbol{A}$ is **triangular** $\implies \det(\boldsymbol{A}) = a_{11}a_{22}\ldots a_{nn}$.
- Suppose that $\boldsymbol{A}$ is not **triangular**.
  - $n = 2 \implies \det(\boldsymbol{A}) = a_{11}a_{22} - a_{12}a_{21}$
  - If a row/column has many 0, use **cofactor expansion**.
  - Otherwise, use **elementary row operations** to get **row-echelon form**.
    - $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$.

2.5.8 Adjoint Matrix

2.5.8.1 Definition of Adjoint Matrix

Definition. Let $\boldsymbol{A}$ be a **square matrix** of order n. The **(classical) adjoint** (or **adjugate**, or **adjunct**) of $\boldsymbol{A}$ is

$$\mathbf{adj}(\boldsymbol{A}) = (A_{ji})_{n\times n}$$

where A_{ij} is the (i,j)-cofactor of $\boldsymbol{A}$.

2.5.8.2 Properties of Adjoint Matrix

Theorem. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be **square matrices** of order n. Then

- $\boldsymbol{A}[\mathbf{adj}(\boldsymbol{A})] = \det(\boldsymbol{A})\boldsymbol{I}$.
- $[\mathbf{adj}(\boldsymbol{A})]\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}$.
- If $\boldsymbol{A}$ is **invertible**, then
  - $\boldsymbol{A}^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})$
  - $[\mathbf{adj}(\boldsymbol{A})]^{-1} = \dfrac{1}{\det(\boldsymbol{A})}\boldsymbol{A}$
  - $\mathbf{adj}(\boldsymbol{A}^{-1}) = [\mathbf{adj}(\boldsymbol{A})]^{-1}$
  - $\det(\mathbf{adj}(\boldsymbol{A})) = [\det(\boldsymbol{A})]^{n-1}$
  - $\mathbf{adj}(\boldsymbol{AB}) = \mathbf{adj}(\boldsymbol{B})\,\mathbf{adj}(\boldsymbol{A})$

2.5.9 Cramer's Rule

Let $\boldsymbol{A}$ be an **invertible matrix** of order n. Then for every **column matrix** $\boldsymbol{b}$ of size $n \times 1$, the **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution:

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})}\begin{pmatrix} \det(\boldsymbol{A}_1) \\ \det(\boldsymbol{A}_2) \\ \vdots \\ \det(\boldsymbol{A}_n) \end{pmatrix}$$

where $\boldsymbol{A}_j$ is obtained from $\boldsymbol{A}$ by replacing its jth column by $\boldsymbol{b}$. Therefore, for $j = 1, 2, \ldots, n$,

$$x_j = \frac{\det(\boldsymbol{A}_j)}{\det(\boldsymbol{A})}$$

Chapter 3

Vector Spaces

3.1 Euclidean n-Spaces

3.1.1 Introduction to n-vectors

3.1.1.1 Definition of n-vector

Definition. An ***n*-vector** or **ordered *n*-tuple** of real numbers is

$$\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$$

where $v_i \in \mathbb{R}$ is the ***i*th component** or ***i*th coordinate** of $\boldsymbol{v}$.

3.1.1.2 Zero Vector

Definition. The n-vector $\mathbf{0} = (0, 0, \ldots, 0)$ is the **zero vector**.

3.1.2 n-vector Operations

3.1.2.1 Equal n-vectors

Definition. Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$. Then $\boldsymbol{u}$ and $\boldsymbol{v}$ are **equal** if $u_i = v_i$ for all $i = 1, \ldots, n$.

3.1.2.2 n-vector Addition, Subtraction & Scalar Multiplication

Definition. Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$, and $c \in \mathbb{R}$. The following operations are defined:

- **Scalar Multiplication:** $c\boldsymbol{v} = (cv_1, cv_2, \ldots, cv_n)$
- **Negative:** $-\boldsymbol{v} = (-1)\boldsymbol{v}$
- **Addition:** $\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$
- **Subtraction:** $\boldsymbol{u} - \boldsymbol{v} = (u_1 - v_1, u_2 - v_2, \ldots, u_n - v_n)$

3.1.3 Properties of n-vectors

Theorem. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ be n-vectors and $c, d \in \mathbb{R}$. Then

- $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{u} + (-\boldsymbol{v})$
- Commutative Law for Vector Addition:
  - $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
- Associative Law for Vector Addition:
  - $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$
- Additive Identity and Additive Inverse:
  - $\boldsymbol{v} + \boldsymbol{0} = \boldsymbol{v}$ and $\boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}$
- Distributive Law for Scalar Multiplication over Addition:
  - $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$
  - $(c + d)\boldsymbol{v} = c\boldsymbol{v} + d\boldsymbol{v}$
- $c(d\boldsymbol{v}) = (cd)\boldsymbol{v} = d(c\boldsymbol{v})$
  - In particular, $-c\boldsymbol{v} = (-c)\boldsymbol{v} = c(-\boldsymbol{v})$ and $-(-\boldsymbol{v}) = \boldsymbol{v}$
- $1\boldsymbol{v} = \boldsymbol{v}$

3.1.4 n-vectors as Matrices

An n-vector $(v_1, v_2, \ldots, v_n)$ can be viewed as

- a **row matrix** (**row vector**) $\begin{pmatrix} v_1 & v_2 & \ldots & v_n \end{pmatrix}$, or
- a **column matrix** (**column vector**) $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

3.1.5 Vectors in xy-Plane

3.1.5.1 Vectors as Points on the xy-Plane

Every point P on the xy-plane is represented by a vector $\boldsymbol{v} = (a, b)$ where

- a is the x-coordinate and b is the y-coordinate; and
- $\boldsymbol{v}$ is the arrow from the origin O to the point P, denoted by $\overrightarrow{OP}$

3.1.5.2 Vectors as Change from Initial Point to End Point

A vector (a, b) represents the change from the **initial point** (x_1, y_1) to the **end point** (x_2, y_2) where $a = x_2 - x_1, b = y_2 - y_1$.

3.1.5.3 Length of Vectors

Definition. Let $\boldsymbol{v} = (v_1, v_2)$ be a vector in xy-plane. Its **length** is

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2}$$

3.1.5.4 Properties of Length of Vectors

Theorem.

- $\|c\boldsymbol{v}\| = |c|\|\boldsymbol{v}\|$
- $\boldsymbol{v} = \mathbf{0} \iff \|v\| = 0$

3.1.5.5 Geometrical Interpretation of Scalar Multiplication

Let $\boldsymbol{v} = (v_1, v_2)$ and $c \in \mathbb{R}$. Then

- $c\boldsymbol{v}$ is a vector **parallel** to $\boldsymbol{v}$ such that
    - its **length** is $|c|$ times the **length** of $\boldsymbol{v}$

    1. $c = 0 \implies c\boldsymbol{v} = 0\boldsymbol{v} = \mathbf{0}$ is the **zero vector**.
    2. $c > 0 \implies c\boldsymbol{v}$ has the same direction as $\boldsymbol{v}$
    3. $c < 0 \implies c\boldsymbol{v}$ has the opposite direction of $\boldsymbol{v}$

3.1.5.6 Geometrical Interpretation of Vector Addition

Let $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$. Then $\boldsymbol{u} + \boldsymbol{v}$ is the vector obtained as follows:

1. Parallel shift $\boldsymbol{v}$ so that its initial point is the same as the end of $\boldsymbol{u}$.
2. Then $\boldsymbol{u} + \boldsymbol{v}$ is the vector from the intial point of $\boldsymbol{u}$ to the end point of $\boldsymbol{v}$.

3.1.5.7 Geometrical Interpretation of Vector Subtraction

Let $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$. Then $\boldsymbol{u} - \boldsymbol{v}$ is the vector obtained as follows:

1. Parallel shift $\boldsymbol{v}$ so that $\boldsymbol{u}$ and $\boldsymbol{v}$ have the same initial point.
2. Then $\boldsymbol{u} - \boldsymbol{v}$ is the vector from the end point of $\boldsymbol{v}$ to the end point of $\boldsymbol{u}$.

3.1.6 Vectors in xyz-Space

3.1.6.1 Vectors as Points on the xyz-Space

Every point P on the xyz-space is represented by a vector $\boldsymbol{v} = (a, b, c)$ where

- a is the x-coordinate, b is the y-coordinate, and c is the z-coordinate; and
- $\boldsymbol{v}$ is the arrow from the origin O to the point P, denoted by $\overrightarrow{OP}$

3.1.7 Euclidean Spaces

Definition. The **Euclidean $\boldsymbol{n}$-space** (or simply **$\boldsymbol{n}$-space**) is the set of all n-vectors of real numbers, denoted by $\mathbb{R}^n$ where

$$\mathbb{R}^n = \{(v_1, v_2, \ldots, v_n) \mid v_1, v_2, \ldots, v_n \in \mathbb{R}\}$$

$\boldsymbol{v} \in \mathbb{R}^n \iff \boldsymbol{v}$ is of the form $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ for real numbers $v_1, v_2, \ldots, v_n$

- In particular,
  - $n = 1 \implies \mathbb{R} = \mathbb{R}^1$ is the real line.
  - $n = 2 \implies \mathbb{R}^2$ is the xy-plane.
  - $n = 3 \implies \mathbb{R}^3$ is the xyz-space.
- Given a **linear system $\boldsymbol{Ax} = \boldsymbol{b}$** in m equations and n variables
  - $\boldsymbol{x}$ as be viewed as an n-vector, i.e. $\boldsymbol{x} \in \mathbb{R}^n$
  - Then the **solution set** of $\boldsymbol{Ax} = \boldsymbol{b}$ is a subset of $\mathbb{R}^n$

3.1.8 Implicit and Explicit Forms of Linear Systems

3.1.8.1 Definition of Implicit and Explicit Forms

Definition.

- A **linear system** is given in the **implicit form** as follows:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \quad \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

- Its general solution is in the **explicit form**.

3.1.8.2 Lines in $\mathbb{R}^2$

3.1.8.2.1 Vector Equation for line in $\mathbb{R}^2$

A **straight line** in $\mathbb{R}^2$ is determined by a point (x_0, y_0) on the line, and its direction vector $(a, b) \neq \mathbf{0}$.

- A point on the line is of the form $(x_0, y_0) + t(a, b)$, where $t \in \mathbb{R}$

3.1.8.2.2 Implicit to Explicit Form

Refer to Chapter 1.1.5 on how to convert implicit form to explicit form.

3.1.8.2.3 Explicit to Implicit Form

Given an explicit form of a line in $\mathbb{R}^2$:

$$\{(x_0 + at, y_0 + bt) \mid t \in \mathbb{R}\}$$

where x_0, y_0, a, b are real constants. An implicit form may be obtained as follows:

1. Let $x = x_0 + at$, and $y = y_0 + bt$
2. Then $t = \dfrac{x - x_0}{a}$ and $t = \dfrac{y - y_0}{b} \implies \dfrac{x - x_0}{a} = \dfrac{y - y_0}{b}$
3. Cross multiply to get $bx - bx_0 = ay - ay_0 \implies bx - ay = bx_0 - ay_0$
4. Hence, the implicit form is $\{(x, y) \mid bx - ay = bx_0 - ay_0\}$

3.1.8.3 Planes in $\mathbb{R}^3$

3.1.8.3.1 Vector Equation for plane in $\mathbb{R}^3$

A **plane** in $\mathbb{R}^3$ is determined by a point (x_0, y_0, z_0) on the plane, and two non-parallel vectors parallel to the plane (a_1, b_1, c_1) and (a_2, b_2, c_2).

- A point on the plane is of the form $(x_0, y_0, z_0) + s(a_1, b_1, c_1) + t(a_2, b_2, c_2)$, where $s, t \in \mathbb{R}$

3.1.8.3.2 Implicit to Explicit Form

Refer to Chapter 1.1.5 on how to convert implicit form to explicit form.

3.1.8.3.3 Explicit to Implicit Form

Given an explicit form of a plane in $\mathbb{R}^3$:

$$\{(x_0 + a_1 s + a_2 t, y_0 + b_1 s + b_2 t, z_0 + c_1 s + c_2 t) \mid t \in \mathbb{R}\}$$

where $x_0, y_0, z_0, a_i, b_i, c_i$ are real constants. An implicit form may be obtained as follows:

1. Let $x = x_0 + a_1 s + a_2 t$, $y = y_0 + b_1 s + b_2 t$, and $z = z_0 + c_1 s + c_2 t$

2. Then we obtain the following **linear system** in s, t:

$$\begin{cases} a_1 s + a_2 t = x - x_0 \\ b_1 s + b_2 t = y - y_0 \\ c_1 s + c_2 t = z - z_0 \end{cases}$$

3. Perform **Gaussian Elimination** to its corresponding **augmented matrix**.

4. The $(3,3)$-entry of its **row-echelon form** is a function f in variables x, y, z

5. Since the system is **consistent**, $f(x, y, z) = 0$.

6. Hence, the implicit form is $\{(x, y, z) \mid f(x, y, z) = 0\}$

3.1.8.4 Lines in $\mathbb{R}^3$

3.1.8.4.1 Vector Equation for line in $\mathbb{R}^3$

A **straight line** in $\mathbb{R}^3$ is determined by a point (x_0, y_0, z_0) on the line, and its direction vector $(a, b, c) \neq \mathbf{0}$.

- A point on the line is of the form $(x_0, y_0, z_0) + t(a, b, c)$, where $t \in \mathbb{R}$

3.1.8.4.2 Implicit to Explicit Form

Solve the **linear system** of two equations (each representing a plane) in three variables to convert implicit form to explicit form.

3.1.8.4.3 Explicit to Implicit Form

Given an explicit form of a line in $\mathbb{R}^3$:

$$\{(x_0 + at, y_0 + bt, z_0 + ct) \mid t \in \mathbb{R}\}$$

where x_0, y_0, z_0, a, b, c are real constants. An implicit form may be obtained as follows:

1. Let $x = x_0 + at$, $y = y_0 + bt$, and $z = z_0 + ct$

2. Find the relation between x and y, say $f(x, y) = 0$

3. Find the relation between x and z, say $g(x, z) = 0$

4. Then, the implicit form is $\{(x, y, z) \mid f(x, y) = 0 \;\&\; g(x, z) = 0\}$

3.2 Linear Combinations and Linear Spans

3.2.1 Definition of Linear Combination

Definition. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ be vectors in $\mathbb{R}^n$. A **linear combination** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ has the form

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k$$

where $c_1, c_2, \ldots, c_k \in \mathbb{R}$

Remarks. In particular, $\mathbf{0}$ is a **linear combination** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$:

$$\mathbf{0} = 0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \ldots + 0\boldsymbol{v}_k$$

3.2.2 Definition of Linear Span

Definition. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be a subset of $\mathbb{R}^n$. The **set of all linear combinations** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ is

$$\{c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k \mid c_1, c_2, \ldots, c_k \in \mathbb{R}\}$$

and is called the **linear span** of S (or $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$), denoted by $\operatorname{span}(S)$ or $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ respectively.

Remarks. $\boldsymbol{v}$ is a linear combination of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \iff \boldsymbol{v} \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$

3.2.3 Criterion for $\operatorname{span}(S) = \mathbb{R}^n$

3.2.3.1 When $k \geq n$

Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. The following steps allow us to check whether $\operatorname{span}(S) = \mathbb{R}^n$:

1. View each $\boldsymbol{v}_j$ as a column vector
2. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_k \end{pmatrix}$
3. Find a **row-echelon form** $\boldsymbol{R}$ of $\boldsymbol{A}$
   - If $\boldsymbol{R}$ has a zero row, then $\operatorname{span}(S) \neq \mathbb{R}^n$
   - If $\boldsymbol{R}$ has no zero row, then $\operatorname{span}(S) = \mathbb{R}^n$

3.2.3.2 When $k < n$

Theorem. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. $k < n \implies \operatorname{span}(S) \neq \mathbb{R}^n$

3.2.4 Properties of Linear Spans

3.2.4.1 Linear Spans Always Contains Zero Vector

Theorem. Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subseteq \mathbb{R}^n$. Then

- $\mathbf{0} \in \operatorname{span}(S)$, where $\mathbf{0}$ is the zero vector in $\mathbb{R}^n$
  - Hence, $\operatorname{span}(S) \neq \varnothing$

3.2.4.2 Linear Spans Are Closed Under Linear Combination

Theorem. Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subseteq \mathbb{R}^n$, $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r \in \text{span}(S)$, and $c_1, c_2, \ldots, c_r \in \mathbb{R}$. Then

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_r\boldsymbol{v}_r \in \text{span}(S)$$

In particular,

- $\text{span}(S)$ is **closed** under scalar multiplication.
  - $\boldsymbol{v} \in \text{span}(S)$ and $c \in \mathbb{R} \implies c\boldsymbol{v} \in \text{span}(S)$
- $\text{span}(S)$ is **closed** under addition.
  - $\boldsymbol{u} \in \text{span}(S)$ and $\boldsymbol{v} \in \text{span}(S) \implies \boldsymbol{u} + \boldsymbol{v} \in \text{span}(S)$

3.2.4.3 When a Span is a Subset of Another Span

Theorem. Given two subsets of $\mathbb{R}^n$:

$$S_1 = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}, S_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$$

Then

$$\text{span}(S_1) \subseteq \text{span}(S_2) \iff \text{Every } \boldsymbol{u}_i \text{ is a linear combination of } \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$$

3.2.4.4 Linear Spans with Redundant Vectors

Theorem. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k \in \mathbb{R}^n$. Then

$$\boldsymbol{v}_k \text{ is a linear combination of } \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1} \implies \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}\} = \text{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k\}$$

3.2.5 Criterion for Vector Belongs to Span

Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. The following steps allow us to check whether a vector $\boldsymbol{v} \in \text{span}(S)$:

1. View each $\boldsymbol{v}_j$ as a **column vector**.
2. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_k \end{pmatrix}$
3. Check if the **linear system** $\boldsymbol{Ax} = \boldsymbol{v}$ is consistent.
   - If $\boldsymbol{Ax} = \boldsymbol{v}$ is consistent, then $\boldsymbol{v} \in \text{span}(S)$
   - If $\boldsymbol{Ax} = \boldsymbol{v}$ is inconsistent, then $\boldsymbol{v} \notin \text{span}(S)$

3.3 Subspaces

3.3.1 Definition of Subspaces

Definition. Let V be a subset of $\mathbb{R}^n$. Then V is called a **subspace** of $\mathbb{R}^n$ if there exist $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ such that:

$$V = \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$$

More precisely,

- V is the **subspace spanned** by $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$;
- $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ **spans** the subspace V.

3.3.2 Zero Space

Let $\mathbf{0} \in \mathbb{R}^n$ be the zero vector. Then

$\{\mathbf{0}\} = \operatorname{span}\{\mathbf{0}\}$ is a subspace of $\mathbb{R}^n$ called the **zero space**

3.3.3 Euclidean n-Space is a Subspace

Let $\boldsymbol{e}_i$ denote the n-vector whose ith coordinate is 1 and elsewhere 0. Then

$\mathbb{R}^n = \operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ is a subspace of $\mathbb{R}^n$

3.3.4 Showing That a Subset is a Subspace

To show that a subset V of $\mathbb{R}^n$ is a subspace:

1. Find $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ such that

   - V is the set containing all vectors of the form:

     $$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k$$

     where $c_1, c_2, \ldots, c_k \in \mathbb{R}$

3.3.5 Showing That a Subset is Not a Subspace

A subset V of $\mathbb{R}^n$ is not a subspace if any of the following fails:

- $\mathbf{0} \in V$
- $c \in \mathbb{R}$ & $\boldsymbol{v} \in V \implies c\boldsymbol{v} \in V$
- $\boldsymbol{u} \in V$ & $\boldsymbol{v} \in V \implies \boldsymbol{u} + \boldsymbol{v} \in V$

3.3.6 Subspaces of $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3$

3.3.6.1 Subspaces of $\mathbb{R}^1 = \mathbb{R}$

Let the nonzero vector $\boldsymbol{v} = v \in \mathbb{R}$. The following are the **subspaces** of $\mathbb{R}^1$:

- $\{0\}$,
- $\mathbb{R} = \text{span}\{\boldsymbol{v}\}$.

3.3.6.2 Subspaces of $\mathbb{R}^2$

Let the nonzero vectors $\boldsymbol{u} = (u_1, u_2) \in \mathbb{R}^2, \boldsymbol{v} = (v_1, v_2) \in \mathbb{R}^2$. The following are the **subspaces** of $\mathbb{R}^2$:

- $\{\mathbf{0}\} = \{(0,0)\}$,
- $\text{span}\{\boldsymbol{v}\}$ (straight line passing through the origin),
- $\mathbb{R}^2 = \text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ if $\boldsymbol{u}$ and $\boldsymbol{v}$ are not parallel.

3.3.6.3 Subspaces of $\mathbb{R}^3$

Let the nonzero vectors $\boldsymbol{u} = (u_1, u_2, u_3) \in \mathbb{R}^3, \boldsymbol{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$. The following are the **subspaces** of $\mathbb{R}^3$:

- $\{\mathbf{0}\} = \{(0,0,0)\}$,
- $\text{span}\{\boldsymbol{v}\}$ (straight line passing through the origin or intersection of two planes containing the origin),
- $\text{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ if $\boldsymbol{u}$ and $\boldsymbol{v}$ are not parallel (a plane containing the origin),
- $\mathbb{R}^3$

3.3.7 Solution Set of Homogeneous Linear System is a Subspace of $\mathbb{R}^n$

Theorem. The **solution set** of a **homogeneous linear system** of n variables is a **subspace** of $\mathbb{R}^n$.

3.3.8 Solution Space

Definition. The **solution set** of a **homogeneous linear system** is called the **solution space** of the system.

3.3.9 A Subspace of $\mathbb{R}^n$ is Always The Solution Space of a Homogeneous Linear System

Theorem. The **subspace** of $\mathbb{R}^n$ is always the **solution space** of a **homogeneous linear system**.

3.4 Vector Spaces

3.4.1 Definition of Vector Space

Definition. A set V is a **vector space** if V is a **subspace** of $\mathbb{R}^n$ for some $n \in \mathbb{N}$.

3.4.2 Expanded Definition of Subspace

Definition. If W and V are **vector spaces** such that $W \subseteq V$, then W is a **subspace** of V.

3.5 Linear Independence

3.5.1 Definition of Linear Independence

Definition. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. If the equation $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k = \mathbf{0}$ has a **non-trivial solution**, then

- S is a **linearly dependent set**,
- $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are **linearly dependent**.

If the equation has **only the trivial solution**, then

- S is a **linearly independent set**,
- $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are **linearly independent**.

3.5.2 Linear Independence of Subset/Superset

Theorem. Let S_1, S_2 be finite subsets of $\mathbb{R}^n$ such that $S_1 \subseteq S_2$. Then

3.5.2.1 Superset Preserves Linear Dependence

S_1 linearly dependent $\Longrightarrow S_2$ linearly dependent.

3.5.2.2 Subset Preserves Linear Independence

S_2 linearly independent $\Longrightarrow S_1$ linearly independent.

3.5.3 Linear Independence of Sets Containing Zero Vector

Theorem. $\mathbf{0} \in S \subseteq \mathbb{R}^n \implies S$ is **linearly dependent**.

3.5.4 Linear Independence of Sets Containing Only One Vector

Theorem. Let $\boldsymbol{v} \in \mathbb{R}^n$. Then $\{\boldsymbol{v}\}$ is **linearly independent** $\iff \boldsymbol{v} \neq \mathbf{0}$.

3.5.5 Linear Independence and Linear Combinations

Theorem. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n, k \geq 2$. Then S is **linearly dependent** $\iff$ there exists $\boldsymbol{v}_i$ such that $\boldsymbol{v}_i \in \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_k\}$.

3.5.6 Cardinality Upper Bound for Linear Independence

Theorem. Let $S \subseteq \mathbb{R}^n$. $|S| > n \implies$ S is **linearly dependent**.

3.5.7 Adding Vectors to a Linearly Independent Set

Theorem. Suppose $S \subseteq \mathbb{R}^n$ is **linearly independent**, $\boldsymbol{v} \in \mathbb{R}^n$, and $\boldsymbol{v} \notin \operatorname{span}(S)$. Then $S \cup \{\boldsymbol{v}\}$ is **linearly independent**.

3.6 Bases

3.6.1 Definition of Bases

Definition. Let S be a finite subset of a **vector space** V. Then S is a **basis** (plural **bases**) for V if

- S is **linearly independent**, and
- $\operatorname{span}(S) = V$.

3.6.1.1 Remarks

- A **basis** for a **vector space** V contains
  - the smallest possible number of vectors that spans V, and
  - the largest possible number of vectors that are linearly independent
- For convenience, the **basis** for $\{\mathbf{0}\}$ is defined to be $\varnothing$.
- Other than $\{\mathbf{0}\}$, any **vector space** has infinitely many different **bases**.

3.6.2 Basis and Unique Linear Combination

Theorem. Let S be a finite subset of a **vector space** V. Then the following are equivalent

- S is a **basis** for V.
- Every $\boldsymbol{v} \in V$ can be **uniquely** expressed (there is exactly one $c_1, \dots, c_k \in \mathbb{R}$) as
  - $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$

3.6.3 Coordinate Vector

Definition. Let S be a finite subset of a **vector space** V.

- For every $\boldsymbol{v} \in V$, there exist unique $c_1, \dots, c_k \in \mathbb{R}$ such that
  - $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$
- $c_1, \dots, c_k$ are the **coordinates** of $\boldsymbol{v}$ **relative** to S.
- $(c_1, \dots, c_k)$ is the **coordinate vector** of $\boldsymbol{v}$ **relative** to the basis S, denoted by $(\boldsymbol{v})_S$.

3.6.3.1 Remark

The order of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ is fixed.

3.6.4 Standard Basis

Definition. Let $E = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} \subseteq \mathbb{R}^n$ where

- $\boldsymbol{e}_1 = (1, 0, \dots, 0), \boldsymbol{e}_2 = (0, 1, \dots, 0), \dots, \boldsymbol{e}_n = (0, 0, \dots, 1)$

E is the **standard basis** for $\mathbb{R}^n$.

- For any $\boldsymbol{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$,
  - $(\boldsymbol{v})_E = (v_1, \dots, v_n) = \boldsymbol{v}$

3.6.5 Properties of Coordinate Vector

Theorem. Let S be a **basis** for a **vector space** V. Suppose $|S| = k$. Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r \in V$

3.6.5.1 Zero Coordinate Vector

$(\boldsymbol{v})_S = \mathbf{0} \Leftrightarrow \boldsymbol{v} = \mathbf{0}$

3.6.5.2 Homogeneity of Degree One / Scalar Multiplication Preserving

For any $c \in \mathbb{R}$ and $\boldsymbol{v} \in V$, $(c\boldsymbol{v})_S = c(\boldsymbol{v})_S$.

3.6.5.3 Additivity / Addition Preserving

For any $\boldsymbol{u}, \boldsymbol{v} \in V$, $(\boldsymbol{u}+\boldsymbol{v})_S = (\boldsymbol{u})_S + (\boldsymbol{v})_S$.

3.6.5.4 Equal Coordinate Vectors

For any $\boldsymbol{u}, \boldsymbol{v} \in V$, $\boldsymbol{u} = \boldsymbol{v} \Leftrightarrow (\boldsymbol{u})_S = (\boldsymbol{v})_S$.

3.6.5.5 Linear Combination Preserving

For any $c_1, \ldots, c_r \in \mathbb{R}$, $(c_1\boldsymbol{v}_1 + \cdots + c_r\boldsymbol{v}_r)_S = c_1(\boldsymbol{v}_1)_S + \cdots + c_r(\boldsymbol{v}_r)_S$.

3.6.5.6 Linear Independence Preserving

$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ are **linearly independent** $\Leftrightarrow (\boldsymbol{v}_1)_S, \ldots, (\boldsymbol{v}_r)_S$ are **linearly independent**.

3.6.5.7 Span Preserving

$\mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\} = V \Leftrightarrow \mathrm{span}\{(\boldsymbol{v}_1)_S, \ldots, (\boldsymbol{v}_r)_S\} = \mathbb{R}^k$.

3.7 Dimensions

3.7.1 Dimension Theorem for Vector Spaces

Theorem. Let V be a **vector space** having a **basis** with k vectors. Then

- Any subset of V of $> k$ vectors is **linearly dependent**.
- Any subset of V of $< k$ vectors cannot span V.

3.7.1.1 Corollary

All **bases** of a **vector space** have the same cardinality.

- To be more precise, if S_1 and S_2 are two **bases** of a **vector space** V,
  - then $|S_1| = |S_2|$

3.7.2 Definition of Dimension

Definition. Let V be a **vector space** and S a **basis** for V. Then

- the **dimension** of V, denoted by $\dim(V)$, is $|S|$

3.7.2.1 Examples

- $\dim(\{\mathbf{0}\}) = 0$
- $\dim(\mathbb{R}^n) = n$
- **Straight Line Through the Origin in $\mathbb{R}^2$ and $\mathbb{R}^3$**
    - $\dim(\operatorname{span}(\boldsymbol{v})) = 1$ with $\boldsymbol{v} \neq \mathbf{0}$.
- **Plane Containing Origin in $\mathbb{R}^3$**
    - $\dim(\operatorname{span}(\boldsymbol{u}, \boldsymbol{v})) = 2$ where $\boldsymbol{u}$ and $\boldsymbol{v}$ are **linearly independent**.

3.7.3 Dimension of Solution Space

Let $\boldsymbol{Ax} = \mathbf{0}$ be a **homogeneous linear system** and V be its **solution space**. Let $\boldsymbol{R}$ be a **row-echelon form** of $\boldsymbol{A}$. Then

$$\begin{aligned}\text{no. of non-pivot colns of } \boldsymbol{R} &= \text{no. of aribitrary parameters in soln}\\ &= \dim(V)\end{aligned}$$

3.7.4 Easier Criterion for Basis

Theorem. Let S be a subset of a **vector space** V. The following are equivalent:

1. S is a **basis** for V.
2. S is **linearly independent**, and $|S| = \dim(V)$.
3. $\operatorname{span}(S) = V$ and $|S| = \dim(V)$.

3.7.5 Dimension for Subspaces

Let U be a **subspace** of a **vector space** V. Then

- $U = V \Leftrightarrow \dim(U) = \dim(V)$.

3.7.5.1 Corollary

Let U be a **subspace** of a **vector space** V. Then

- $U \neq V \Leftrightarrow \dim(U) < \dim(V)$.

3.7.6 Invertibility and Rows and Columns as Basis

Theorem. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. Then the following are equivalent:

1. $\boldsymbol{A}$ is **invertible**.
2. The rows of $\boldsymbol{A}$ form a **basis** for $\mathbb{R}^n$.
3. The columns of $\boldsymbol{A}$ form a **basis** for $\mathbb{R}^n$.

3.8 Transition Matrices

3.8.1 Coordinate Vector As a Column Vector

Definition. Let S be a **basis** for a **vector space** V and $\boldsymbol{v} \in V$ such that $(\boldsymbol{v})_S = (c_1, \dots, c_k)$. Then the **column vector** $[\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$ is also called the **coordinate vector** of $\boldsymbol{v}$ relative to S.

3.8.2 Definition of Transition Matrix

Definition. Let V be a **vector space**, and $S = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ and T be **bases** for V. Then the **transition matrix** from S to T is

$$\boldsymbol{P} = \begin{pmatrix} [\boldsymbol{u}_1]_T & \cdots & [\boldsymbol{u}_k]_T \end{pmatrix}$$

3.8.3 Pre-Multiplication of Transition Matrix

Theorem. Let S and T be **bases** for a **vector space** V, and $\boldsymbol{P}$ be the **transition matrix** from S to T. Then $\forall \boldsymbol{w} \in V \quad \boldsymbol{P}[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$.

3.8.4 Inverse of Transition Matrices

Theorem. Let S and T be **bases** for a **vector space** V, and $\boldsymbol{P}$ be the **transition matrix** from S to T. Then

3.8.4.1 Transition Matrices Are Invertible

P is **invertible**.

3.8.4.2 Inverse is a Transition Matrix

P^{-1} is the **transition matrix** from T to S.

Chapter 4

Vector Spaces Associated With Matrices

4.1 Row Spaces and Column Spaces

4.1.1 Definition of Row Spaces and Column Spaces

Definition. Let $\boldsymbol{A} = (a_{ij})_{m\times n}$.

4.1.1.1 Row Space

For $1 \le i \le m$, let

$$\begin{aligned}\boldsymbol{r}_i &= i\text{th row of } \boldsymbol{A}\\ &= \begin{pmatrix} a_{i1} & \cdots & a_{in}\end{pmatrix} \in \mathbb{R}^n\end{aligned}$$

Then the **row space** of $\boldsymbol{A} = \operatorname{span}\{\boldsymbol{r}_1, \dots, \boldsymbol{r}_m\}$ is a **subspace** of $\mathbb{R}^n$.

4.1.1.2 Column Space

For $1 \le j \le n$, let

$$\begin{aligned}\boldsymbol{c}_j &= j\text{th column of } \boldsymbol{A}\\ &= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj}\end{pmatrix} \in \mathbb{R}^m\end{aligned}$$

Then the **column space** of $\boldsymbol{A} = \operatorname{span}\{\boldsymbol{c}_1, \dots, \boldsymbol{c}_n\}$ is a **subspace** of $\mathbb{R}^m$.

4.1.2 Row Equivalent Matrices And Their Rows

4.1.2.1 Row Equivalence Implies Same Row Space

Theorem. Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are **row equivalent matrices**. Then **row space** of $\boldsymbol{A}$ = **row space** of $\boldsymbol{B}$.

4.1.2.1.1 Remark

Let $\boldsymbol{R}$ be a **row-echelon form** of $\boldsymbol{A}$. Then the **row space** of $\boldsymbol{A}$ = **row space** of $\boldsymbol{R}$.

4.1.2.2 Nonzero Rows of REF are Linearly Independent

Theorem. Let $\boldsymbol{R}$ be a **row-echelon form** of a matrix $\boldsymbol{A}$. Then the nonzero rows of R are **linearly independent**.

4.1.2.3 Finding Basis/Dimension For Row Space

Theorem. Let $\boldsymbol{R}$ be a **row-echelon form** of a matrix $\boldsymbol{A}$. Then

- the nonzero rows of $\boldsymbol{R}$ form a **basis** for the **row space** of $\boldsymbol{A}$, and
- dim(row space of $\boldsymbol{A}$) = no. of nonzero rows of $\boldsymbol{R}$.

4.1.3 Row Equivalent Matrices And Their Columns

4.1.3.1 Row Equivalence Preserve Linear Relations Between Columns

Theorem. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be **row equivalent matrices**. Then

- There is a linear relation among a given set of columns of $\boldsymbol{A}$ $\Leftrightarrow$ the same linear relation exists among the corresponding set of columns of $\boldsymbol{B}$.
- A given set of columns of $\boldsymbol{A}$ is **linearly independent** $\Leftrightarrow$ the corresponding set of columns of $\boldsymbol{B}$ is **linearly independent**.
- A given set of columns of $\boldsymbol{A}$ is a **basis** for the column space of $\boldsymbol{A}$ $\Leftrightarrow$ the corresponding set of columns of $\boldsymbol{B}$ is a **basis** for the column space of $\boldsymbol{B}$

4.1.3.2 Pivot Columns of REF is a Basis For Column Space

Theorem. Let $\boldsymbol{R}$ be a **row-echelon form** of a matrix $\boldsymbol{A}$. Then the **pivot columns** of R is a **basis** for the **column space** of $\boldsymbol{R}$.

4.1.3.3 Finding Basis/Dimension For Column Space

Theorem. Let $\boldsymbol{R}$ be a **row-echelon form** of a matrix $\boldsymbol{A}$. The **pivot columns** of $\boldsymbol{R}$ is a **basis** for the **column space** of $\boldsymbol{R}$. Then

- the corresponding columns of $\boldsymbol{A}$ is a **basis** for the column space of $\boldsymbol{A}$, and
- dim(column space of $\boldsymbol{A}$) = no. of **pivot columns** of $\boldsymbol{R}$.

4.1.4 Finding Basis for Vector Space Spanned By a Set of Vectors

To find a **basis** for a **vector space** $\boldsymbol{V} = \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$, choose either one of the methods below:

4.1.4.1 Method 1: From New Vectors Using Row Space

1. Let each $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ be a **row vector**
2. Let matrix $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_m \end{pmatrix}$
3. The problem is now equivalent to finding the **basis** for the **row space** of $\boldsymbol{A}$ (refer to section 1.1.2.3).

4.1.4.2 Method 2: Select From Original Vectors Using Column Space

1. Let each $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ be a **column vector**
2. Let matrix $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_m \end{pmatrix}$
3. The problem is now equivalent to finding the **basis** for the **column space** of $\boldsymbol{A}$ (refer to section 1.1.3.3).

4.1.5 Consistency of Linear Systems And Column Space of Coefficient Matrix

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix**. The **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ is **consistent** $\Leftrightarrow \boldsymbol{b} \in \{\boldsymbol{Av} \mid \boldsymbol{v} \in \mathbb{R}^n\}$ = column space of $\boldsymbol{A}$.

4.2 Rank

4.2.1 The Dimension of Row Space And Column Space Are The Same

Theorem. Let $\boldsymbol{A}$ be a **matrix**. Then

$$\dim(\text{row space of } \boldsymbol{A}) = \dim(\text{column space of } \boldsymbol{A})$$

4.2.2 Definition of Rank

Definition. Let $\boldsymbol{A}$ be a **matrix**. The **rank** of $\boldsymbol{A}$, denoted by $\operatorname{rank}(\boldsymbol{A})$ is

$$\operatorname{rank}(\boldsymbol{A}) = \dim(\text{row space of } \boldsymbol{A}) = \dim(\text{column space of } \boldsymbol{A})$$

4.2.3 Properties of Rank

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix**. Then

- $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^T)$

- $\text{rank}(\boldsymbol{A}) = 0 \Leftrightarrow \boldsymbol{A} = \boldsymbol{0}$
- $\text{rank}(\boldsymbol{A}) \leq m$ and $\text{rank}(\boldsymbol{A}) \leq n$
  - $\text{rank}(\boldsymbol{A}) \leq \min\{m, n\}$
  - $\boldsymbol{A}$ is **full rank** if $\text{rank}(\boldsymbol{A}) = \min\{m, n\}$
- A **square matrix** $\boldsymbol{A}$ is **full rank** $\Leftrightarrow$ $\boldsymbol{A}$ is **invertible**.

4.2.4 Rank & Consistency of Linear System

Theorem. Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a **linear system**. Let $\left(\begin{array}{c|c} \boldsymbol{R} & \boldsymbol{b}' \end{array}\right)$ be the **row-echelon form** of $\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{b} \end{array}\right)$. Then

$$\begin{aligned} \boldsymbol{Ax} = \boldsymbol{b} \text{ is } \textbf{consistent} &\Leftrightarrow \text{rank}(\boldsymbol{A}) = \text{rank}\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{b} \end{array}\right) \\ &\Leftrightarrow \text{rank}(\boldsymbol{R}) = \text{rank}\left(\begin{array}{c|c} \boldsymbol{R} & \boldsymbol{b}' \end{array}\right) \end{aligned}$$

4.2.4.1 Remark

In general, $\text{rank}(\boldsymbol{A}) \leq \text{rank}\left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{b} \end{array}\right) \leq \text{rank}(\boldsymbol{A}) + 1$

4.2.5 Row/Column Spaces and Matrix Multiplication

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix** and $\boldsymbol{B}$ be a $n \times p$ **matrix**. Then

- column space of $\boldsymbol{AB} \subseteq$ column space of $\boldsymbol{A}$;
- row space of $\boldsymbol{AB} \subseteq$ row space of $\boldsymbol{B}$.

4.2.5.1 Ranks and Matrix Multiplication

In particular,

- $\text{rank}(\boldsymbol{AB}) \leq \text{rank}(\boldsymbol{A})$;
- $\text{rank}(\boldsymbol{AB}) \leq \text{rank}(\boldsymbol{B})$.

That is, $\text{rank}(\boldsymbol{AB}) \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$.

- $\boldsymbol{A}$ is **invertible** $\implies \text{rank}(\boldsymbol{AB}) = \text{rank}(\boldsymbol{B})$
- $\boldsymbol{B}$ is **invertible** $\implies \text{rank}(\boldsymbol{AB}) = \text{rank}(\boldsymbol{A})$

4.3 Nullspaces and Nullities

4.3.1 Definition of Nullspace

Definition. Let $\boldsymbol{A}$ be a $m \times n$ **matrix**. The **nullspace** of $\boldsymbol{A}$ is the **solution space** of $\boldsymbol{Ax} = \boldsymbol{0}$, which is

$$\{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{Av} = \boldsymbol{0}\}.$$

4.3.2 Definition of Nullity

Definition. The **dimension** of the **nullspace** of a **matrix** $\boldsymbol{A}$ is the **nullity** of $\boldsymbol{A}$, denoted by $\text{nullity}(\boldsymbol{A})$.

4.3.3 Finding Nullity of a Matrix

Let $\boldsymbol{R}$ be a **row-echelon form** of $\boldsymbol{A}$. Then

- $\boldsymbol{Ax} = \boldsymbol{0} \Leftrightarrow \boldsymbol{Rx} = \boldsymbol{0}$
- **nullspace** of $\boldsymbol{A}$ = **nullspace** of $\boldsymbol{R}$
- $\text{nullity}(\boldsymbol{A}) = \text{nullity}(\boldsymbol{R})$ = no. of **non-pivot columns** of $\boldsymbol{R}$.

4.3.4 Dimension Theorem

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix**. Then

$$\text{rank}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{A}) = n$$

4.3.5 Solution to Inhomogeneous Linear System and Nullspace

Theorem. Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ has a **solution** $\boldsymbol{v}$. Let W = **nullspace** of $\boldsymbol{A}$. Then the **solution set** of $\boldsymbol{Ax} = \boldsymbol{b}$ is

$$\boldsymbol{v} + W = \{\boldsymbol{v} + \boldsymbol{w} \mid \boldsymbol{w} \in W\}$$

Hence

(A general solution of $\boldsymbol{Ax} = \boldsymbol{b}$) = (A particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$)+(A general solution of $\boldsymbol{Ax} = \boldsymbol{0}$)

4.3.5.1 Condition for Linear System to have Unique Solution

Theorem. Suppose that a **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ is **consistent**. Then

$$\begin{aligned} \boldsymbol{Ax} = \boldsymbol{b} \text{ has a unique } \textbf{solution} &\Leftrightarrow \boldsymbol{Ax} = \boldsymbol{0} \text{ has only the } \textbf{trivial solution} \\ &\Leftrightarrow \textbf{nullspace} \text{ of } \boldsymbol{A} \text{ is } \{\boldsymbol{0}\} \\ &\Leftrightarrow \text{nullity}(\boldsymbol{A}) = 0 \\ &\Leftrightarrow \text{rank}(\boldsymbol{A}) = \text{no. of columns of } \boldsymbol{A} \end{aligned}$$

4.3.5.2 Condition for the Solution Set of a Linear System to be a Vector Space

Theorem. The **solution set** of a **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ is a **vector space** $\Leftrightarrow$ $\boldsymbol{b} = \mathbf{0}$.

Chapter 5

Orthogonality

5.1 The Dot Product

5.1.1 Pythagoras' Theorem

Theorem. In a right-angled triangle, let c be the length of the **hypotenuse**, and a and b be the lengths of the other two sides. Then

$$a^2 + b^2 = c^2$$

5.1.2 Cosine Rule

Theorem. Given any triangle with sides of length a, b, and c. Let θ be the angle contained between the sides of lengths a and b. Then

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

5.1.3 Definitions

Definition. Let $\boldsymbol{u} = (u_1, \dots, u_n), \boldsymbol{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$.

5.1.3.1 Dot Product (Inner Product)

The **dot product** (**inner product**) of $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + \cdots + u_nv_n$$

5.1.3.2 Norm (Length)

The **norm** (**length**) of $\boldsymbol{v}$ is

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

5.1.3.3 Unit Vector

Definition. $\boldsymbol{v}$ is a **unit vector** if $\|v\| = 1$.

5.1.3.4 Distance

The **distance** between $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2}$$

5.1.3.5 Angle

The **angle** between $\boldsymbol{u}$ and $\boldsymbol{v}$ ($\boldsymbol{u} \neq \boldsymbol{0}$ and $\boldsymbol{v} \neq \boldsymbol{0}$) is

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right), \quad 0 \leq \theta \leq \pi$$

5.1.4 Dot Product and Matrix Multiplication

5.1.4.1 Dot Product As Matrix Multiplication

Theorem. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$.

- If $\boldsymbol{u}$ and $\boldsymbol{v}$ are viewed as **row vectors**, then $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^T$.
- If $\boldsymbol{u}$ and $\boldsymbol{v}$ are viewed as **column vectors**, then $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v}$.

5.1.4.2 Matrix Multiplication As Dot Products

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix** and $\boldsymbol{B}$ be a $n \times p$ **matrix**. Let the **row vector** $\boldsymbol{a}_i \in \mathbb{R}^n$ be the ith row of $\boldsymbol{A}$ and the **column vector** $\boldsymbol{b}_j \in \mathbb{R}^n$ be the jth column of $\boldsymbol{B}$. Then

$$(i, j)\text{-entry of } \boldsymbol{AB} = \boldsymbol{a}_i\boldsymbol{b}_j = \boldsymbol{a}_i \cdot \boldsymbol{b}_j$$

5.1.5 Properties of Dot Product and Norm

Theorem. Let $\boldsymbol{u}, \boldsymbol{v} = (v_1, \ldots, v_n), \boldsymbol{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

5.1.5.1 Dot Product is Commutative

- $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$

5.1.5.2 Dot Product is Distributive Over Vector Addition

- $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$
- $\boldsymbol{w} \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{w} \cdot \boldsymbol{u} + \boldsymbol{w} \cdot \boldsymbol{v}$

5.1.5.3 Dot Product and Scalar Multiplication

- $(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v})$

5.1.5.4 Absolute Homogeneity of Norm

- $\|c\boldsymbol{v}\| = |c|\|\boldsymbol{v}\|$

5.1.5.5 Dot Product of a Vector with Itself

- $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$
- $\boldsymbol{v} \cdot \boldsymbol{v} = 0 \Leftrightarrow \boldsymbol{v} = \mathbf{0}$

5.1.5.6 Norm and Dot Product

- $\|\boldsymbol{v}\| = \sqrt{v_1^2 + \cdots + v_n^2} = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$

5.1.5.7 Nonnegativity of Norm and When Norm is Zero

- $\|\boldsymbol{v}\| \geq 0$
- $\|\boldsymbol{v}\| = 0 \Leftrightarrow \boldsymbol{v} = \mathbf{0}$

5.1.6 Inequalities of Norm and Distance

Theorem. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$

5.1.6.1 Cauchy-Schwarz Inequality

- $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$

5.1.6.2 Triangle Inequality (Norm Version)

- $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

5.1.6.3 Triangle Inequality (Distance Version)

- $d(\boldsymbol{u}, \boldsymbol{w}) \leq d(\boldsymbol{u}, \boldsymbol{v}) + d(\boldsymbol{v}, \boldsymbol{w})$

5.1.7 Results for Multiplying a Matrix with its Transpose

5.1.7.1 When Multiplying a Matrix with its Transpose Gives the Zero Matrix

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ matrix.

$$\boldsymbol{A}\boldsymbol{A}^T = \mathbf{0}_{m\times m} \implies \boldsymbol{A} = \mathbf{0}_{m\times n}$$

5.1.7.2 When the Trace of Multiplying a Matrix with its Transpose is Zero

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ matrix. Then

$$\operatorname{tr}(\boldsymbol{A}\boldsymbol{A}^T) = 0 \Leftrightarrow \boldsymbol{A} = \mathbf{0}_{m\times n}$$

5.2 Orthogonality, Orthogonal and Orthonormal Sets

5.2.1 Definition of Orthogonal for Two Vectors

Definition. Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$. They are **orthogonal**, denoted by $\boldsymbol{u} \perp \boldsymbol{v}$, if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

5.2.1.1 Zero Vector is Orthogonal to Every Vector in $\mathbb{R}^n$

Theorem. Let $\mathbf{0} \in \mathbb{R}^n$. Then $\forall \boldsymbol{v} \in \mathbb{R}^n \quad \mathbf{0} \perp \boldsymbol{v}$

5.2.2 Definition of Orthogonal / Orthonormal Sets

5.2.2.1 Definition of Orthogonal for a Subset of $\mathbb{R}^n$

Definition. Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. S is **orthogonal** if $\boldsymbol{v}_i \perp \boldsymbol{v}_j$ for all $i \neq j$, that is

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0 \quad \text{for all } i \neq j$$

5.2.2.2 Definition of Orthonormal for a Subset of $\mathbb{R}^n$

Definition. Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$. S is **orthonormal** if S is **orthogonal** and $\forall \boldsymbol{v}_i \in S \quad \|\boldsymbol{v}_i\| = 1$, that is

$$\boldsymbol{v}_i \cdot \boldsymbol{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

5.2.3 Properties of Orthogonal / Orthonormal Sets

- S is **orthonormal** $\implies$ S is **orthogonal**.
- S is **orthogonal** $\implies$ a subset of S is **orthogonal**.
- S is **orthonormal** $\implies$ a subset of S is **orthonormal**.
- S is **orthogonal** $\implies$ $S \cup \{\mathbf{0}\}$ is **orthogonal**.
- S is **orthonormal** $\implies$ $\mathbf{0} \notin S$.

5.2.4 Normalizing: Converting an Orthogonal Set to an Orthonormal Set

The following process of converting an **orthogonal** set of nonzero vectors to an **orthonormal** set of vectors, is called **normalizing**:

1. Let $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\} \subseteq \mathbb{R}^n$ be an **orthogonal** set of **nonzero vectors**
2. For all $\boldsymbol{u}_i \in S$, set $\boldsymbol{v}_i = \dfrac{\boldsymbol{u}_i}{\|\boldsymbol{u}_i\|}$.
3. Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an **orthonormal** set.

5.2.5 Using Matrix Multiplication to Check for Orthogonal / Orthonormal Set

Theorem. To check whether a set $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \subseteq \mathbb{R}^n$ is **orthogonal** / **orthonormal**, let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$. Then

- S is **orthogonal** $\Leftrightarrow \boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ for all $i \neq j \Leftrightarrow \boldsymbol{A}^T\boldsymbol{A}$ is **diagonal**

- S is **orthonormal** $\Leftrightarrow \boldsymbol{v}_i \cdot \boldsymbol{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \Leftrightarrow \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_k$.

5.2.5.1 Standard Basis is An Orthonormal Set

Theorem. The **standard basis** $E = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} \subseteq \mathbb{R}^n$ is an **orthonormal** set.

5.2.5.2 An Orthonormal Subset of $\mathbb{R}^n$ with n Vectors is a Basis for $\mathbb{R}^n$

Theorem. Let $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\} \subseteq \mathbb{R}^n$. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix}$. Then

$$\begin{aligned} \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}_n &\implies \boldsymbol{A} \text{ is invertible} \\ &\implies \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\} \text{ is a } \textbf{basis} \text{ for } \mathbb{R}^n \quad \text{(refer to section 3.7.6)} \end{aligned}$$

5.2.6 Orthogonal Nonzero Sets are Linearly Independent

Theorem. Let $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ be an **orthogonal set** of **nonzero vectors** in $\mathbb{R}^n$. Then S is **linearly independent**.

5.2.6.1 Orthonormal Sets are Linearly Independent

Corollary. Let $S = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ be an **orthonormal set** in $\mathbb{R}^n$. Then S is **linearly independent**.

5.3 Orthogonal and Orthonormal Bases

5.3.1 Orthogonal / Orthonormal Bases

Definition. Let S be a **basis** for a **vector space**.

- S is an **orthogonal basis** if it is **orthogonal**.

- S is an **orthonormal basis** if it is **orthonormal**.

5.3.2 Criterion for Orthogonal / Orthonormal Set to be Basis

Theorem. Suppose S is a subset of **vector space** V, where S is an **orthogonal/orthonormal set** and $\mathbf{0} \notin S$, then

$|S| = \dim(V)$ or $\operatorname{span}(S) = V \implies S$ is an **orthogonal/orthonormal** (respectively) **basis** for V

5.3.3 Simpler Formula for Coordinate Vectors Relative to Orthogonal / Orthonormal Bases

5.3.3.1 Orthogonal Bases

Theorem. Let $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be an **orthogonal basis** for a **vector space** V. Then

$$\forall \boldsymbol{w} \in V \quad (\boldsymbol{w})_S = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}, \ldots, \frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\boldsymbol{u}_k \cdot \boldsymbol{u}_k}\right)$$

5.3.3.2 Orthonormal Bases

Theorem. Let $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be an **orthonormal basis** for a **vector space** V. Then

$$\forall \boldsymbol{w} \in V \quad (\boldsymbol{w})_S = (\boldsymbol{w} \cdot \boldsymbol{u}_1, \ldots, \boldsymbol{w} \cdot \boldsymbol{u}_k)$$

5.3.4 Orthogonality between a Vector and a Subspace

Definition. Let V be a **subspace** of $\mathbb{R}^n$. $\boldsymbol{u} \in \mathbb{R}^n$ is **orthogonal** (**perpendicular**) to V if $\forall \boldsymbol{v} \in V \quad \boldsymbol{u} \perp \boldsymbol{v}$.

5.3.4.1 Example: Normal Vector of a Plane in $\mathbb{R}^3$

Let $V = \{(x, y, z) \mid ax + by + cz = 0\}$ be a plane in $\mathbb{R}^3$ containing the origin. Then $\boldsymbol{n} = (a, b, c)$ is **orthogonal** to V and is a **normal vector** of the plane V.

5.3.5 Easier Criterion for Vector to be Orthogonal to a Vector Space

Theorem. Let $V = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be a **vector space**. Then

$\boldsymbol{w}$ is **orthogonal** to $V \Leftrightarrow \boldsymbol{w} \perp \boldsymbol{v}_i$ for all $\boldsymbol{v}_i \in V$

5.3.6 The Set of All Vectors Orthogonal to a Subspace is a Subspace

Theorem. Let W be a **subspace** of $\mathbb{R}^n$. Then $W^\perp = \{\boldsymbol{v} \in \mathbb{R}^n \mid \boldsymbol{v}$ is **orthogonal** to $W\}$ is a **subspace** of $\mathbb{R}^n$.

5.3.7 Projection of a Vector Onto a Vector Space

5.3.7.1 With Orthonormal Basis

Theorem. Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be an **orthonormal basis** for a **vector space** V. Then the **projection** of $\boldsymbol{w}$ onto V is

$$(\boldsymbol{w} \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + \cdots + (\boldsymbol{w} \cdot \boldsymbol{v}_k)\boldsymbol{v}_k$$

5.3.7.2 With Orthogonal Basis

Theorem. Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be an **orthogonal basis** for a **vector space** V. Then the **projection** of $\boldsymbol{w}$ onto V is

$$\left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_1}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\right)\boldsymbol{u}_1 + \cdots + \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u}_k}{\boldsymbol{u}_k \cdot \boldsymbol{u}_k}\right)\boldsymbol{u}_k$$

Remarks. It is the sum of projections of $\boldsymbol{w}$ onto $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$.

5.3.8 Gram-Schmidt Process: Generating Orthogonal / Orthonormal Basis

Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be a **basis** for a **vector space** V. Define

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{u}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_3 \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 \\
\vdots \quad & \quad \vdots \\
\boldsymbol{v}_k &= \boldsymbol{u}_k - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_2}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{u}_k \cdot \boldsymbol{v}_{k-1}}{\boldsymbol{v}_{k-1} \cdot \boldsymbol{v}_{k-1}}\boldsymbol{v}_{k-1}
\end{aligned}$$

- Then $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an **orthogonal basis** for V.
- $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ can be **normalized** to obtain an **orthonormal basis** for V (refer to section 5.2.4).

5.3.9 Solving a Linear System whose Coefficient Matrix's Columns are Linearly Independent

5.3.9.1 Decomposition

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ matrix whose columns are **linearly independent**. Then there exists

- A $m \times n$ **matrix $\boldsymbol{Q}$** whose columns form an **orthonormal set**, and

- An **invertible** $n \times n$ **uppper triangular matrix** $\boldsymbol{R}$

such that $\boldsymbol{A} = \boldsymbol{QR}$.

5.3.9.2 Algorithm

To solve a **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ where the columns of $\boldsymbol{A}$ are **linearly independent**.

1. $(\boldsymbol{QR})\boldsymbol{x} = \boldsymbol{b}$
2. $\boldsymbol{Rx} = \boldsymbol{IRx} = \boldsymbol{Q}^T\boldsymbol{QRx} = \boldsymbol{Q}^T\boldsymbol{b}$
3. Solve $\boldsymbol{x}$ by **back-substitution**.

5.3.9.3 Remark

There exists a $\boldsymbol{R}$ such that the **diagonal entries** are all positive.

5.4 Best Approximations

5.4.1 The Projection of a Vector on a Subspace is its Best Approximation in that Subspace

Theorem. Let V be a **subspace** of $\mathbb{R}^n$. $\forall \boldsymbol{u} \in \mathbb{R}^n$, let $\boldsymbol{p}$ be the **projection** of $\boldsymbol{u}$ onto V. Then

- $\boldsymbol{p}$ is the **best approximation** of $\boldsymbol{u}$ in V, i.e.
  - $\forall \boldsymbol{v} \in V \quad d(\boldsymbol{u}, \boldsymbol{p}) \leq d(\boldsymbol{u}, \boldsymbol{v})$
- Moreover, $\boldsymbol{p}$ is the **only best approximation** of $\boldsymbol{u}$ in V.
  - $d(\boldsymbol{u}, \boldsymbol{p}) = d(\boldsymbol{u}, \boldsymbol{v}) \Leftrightarrow \boldsymbol{v} = \boldsymbol{p}$

5.4.2 Least Squares Solution

5.4.2.1 Definition of Least Squares Solution

Definition. Let $\boldsymbol{A}$ be a $m \times n$ **matrix** and $\boldsymbol{b} \in \mathbb{R}^m$. $\boldsymbol{u} \in \mathbb{R}^n$ is a **least squares solution** to the **linear system** $\boldsymbol{Ax} = \boldsymbol{b}$ if

$$\forall \boldsymbol{v} \in \mathbb{R}^n \quad \|\boldsymbol{b} - \boldsymbol{Au}\| \leq \|\boldsymbol{b} - \boldsymbol{Av}\|$$

5.4.2.2 Least Squares Solution and Projection

Theorem. Let $\boldsymbol{A}$ be a $m \times n$ **matrix** and $\boldsymbol{b} \in \mathbb{R}^m$. Let $\boldsymbol{p}$ be the **projection** of $\boldsymbol{b}$ onto the **column space** of $\boldsymbol{A}$. Then

- $\forall \boldsymbol{v} \in \mathbb{R}^n \quad \|\boldsymbol{b} - \boldsymbol{p}\| \leq \|\boldsymbol{b} - \boldsymbol{Av}\|$
- $\boldsymbol{u}$ is a **least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{u}$ is a **solution** to $\boldsymbol{Ax} = \boldsymbol{p}$.

5.4.3 Methodology: Finding Least Squares Solution

5.4.3.1 Tedious Method

To find a **least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b}$, proceed as follows:

1. Find an **orthogonal** (**orthonormal**) **basis** for $V =$ **column space** of $\boldsymbol{A}$.

2. Find the **projection** $\boldsymbol{p}$ of $\boldsymbol{b}$ onto V.

3. Solve the **linear system** $\boldsymbol{Ax} = \boldsymbol{p}$.

    - A **solution** to $\boldsymbol{Ax} = \boldsymbol{p}$ is a **least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b}$.

Remarks.

- The **linear system** $\boldsymbol{Ax} = \boldsymbol{p}$ is always **consistent** since $\boldsymbol{p}$ lies in the **column space** of $\boldsymbol{A}$.

- If $\boldsymbol{Ax} = \boldsymbol{b}$ is already **consistent**, then

    - $\boldsymbol{b} = \boldsymbol{p} \in V$
    - (**solution** to $\boldsymbol{Ax} = \boldsymbol{b}$) = (**least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b}$).

5.4.3.2 Easy Method

Theorem. $\boldsymbol{u}$ is a **least squares solution** to $\boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \boldsymbol{u}$ is a **solution** to $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.

5.5 Orthogonal Matrices

5.5.1 Definition of Orthogonal Matrices

Definition. Let $\boldsymbol{A}$ be a **square matrix**. Then

$$\begin{aligned}\boldsymbol{A} \text{ is an } \textbf{orthogonal matrix} &\Leftrightarrow \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}\\ &\Leftrightarrow \boldsymbol{A}^{-1} = \boldsymbol{A}^T\\ &\Leftrightarrow \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}\end{aligned}$$

5.5.1.1 Example: Identity Matrix is an Orthogonal Matrix

Theorem. The **identity matrix** $\boldsymbol{I}_n$ is an **orthogonal matrix**.

5.5.2 Properties of Orthogonal Matrices

5.5.2.1 Rows and Columns of Orthogonal Matrix Form an Orthonormal Basis for $\mathbb{R}^n$

Theorem. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. Then

$$\boldsymbol{A} \text{ is an } \textbf{orthogonal matrix} \Leftrightarrow \text{columns of } \boldsymbol{A} \text{ form an } \textbf{orthonormal basis} \text{ for } \mathbb{R}^n$$

$\Leftrightarrow$ rows of $\boldsymbol{A}$ form an **orthonormal basis** for $\mathbb{R}^n$

5.5.2.2 The Transpose or Inverse of an Orthogonal Matrix is Orthogonal

Theorem. If $\boldsymbol{A}$ is an **orthogonal matrix**, then $\boldsymbol{A}^T = \boldsymbol{A}^{-1}$ is an **orthogonal matrix**.

5.5.2.3 The Product of Two Orthogonal Matrices is Orthogonal

Theorem. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are **orthogonal matrices** of the same size, then $\boldsymbol{AB}$ is an **orthogonal matrix**.

5.5.2.4 Condition for Rows/Columns of a Matrix to Form an Orthonormal Set

5.5.2.4.1 Condition for Columns of a Matrix to Form an Orthonormal Set

Refer to section 5.2.5.

5.5.2.4.2 Condition for Rows of a Matrix to Form an Orthonormal Set

Theorem. For any $m \times n$ **matrix** $\boldsymbol{A}$, $\boldsymbol{AA}^T = \boldsymbol{I}_m \Leftrightarrow$ the **rows** of $\boldsymbol{A}$ form an **orthonormal set**.

5.5.2.5 Pre-Multiplication of an Orthogonal Matrix to an Orthonormal Set

Theorem. Let $S = \{\boldsymbol{u}_1 \ldots, \boldsymbol{u}_k\}$ be an **orthonormal** subset of $\mathbb{R}^n$ and $\boldsymbol{P}$ be a $n \times n$ **orthogonal matrix**. Then $\{\boldsymbol{Pu}_1, \ldots, \boldsymbol{Pu}_k\}$ is an **orthonormal set**.

5.5.3 Transition Matrices Between Orthonormal Bases

5.5.3.1 Formula for Transition Matrices Between Orthonormal Bases

Theorem. Let $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ be **orthonormal bases** for a **vector space** V. Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_k \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_k \end{pmatrix}$. Then

- $\boldsymbol{P} = \boldsymbol{B}^T\boldsymbol{A}$ is the **transition matrix** from S to T.
- $\boldsymbol{Q} = \boldsymbol{A}^T\boldsymbol{B}$ is the **transition matrix** from T to S.

5.5.3.2 Transition Matrices Between Orthonormal Bases are Orthogonal

Theorem. Let S and T be **orthonormal bases** for a **vector space** V and $\boldsymbol{P}$ be the **transition matrix** from S to T. Then $\boldsymbol{P}$ is an **orthogonal matrix**.

5.5.4 Classification of Orthogonal Matrices

5.5.4.1 Table of Orthogonal Matrices By Order

Theorem. The following table lists all the **orthogonal matrices** in increasing **order**:

Order	Determinant	Formulae
1	1	$\begin{pmatrix} 1 \end{pmatrix}$
	-1	$\begin{pmatrix} -1 \end{pmatrix}$
2	1	$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
	-1	$\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$

5.5.5 Geometric Representation of Orthogonal Matrix

5.5.5.1 Rotation Matrix

- Let $\boldsymbol{u}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$.
  - Then $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is an **orthonormal basis** for $\mathbb{R}^2$.
- Let $\boldsymbol{P}_\theta = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Then
  - $\boldsymbol{P}_\theta$ is **orthogonal**, and
  - $\boldsymbol{P}_\theta$ is the **transition matrix** from $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ to $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$.
- $\forall \boldsymbol{u} \in \mathbb{R}^2 \quad \boldsymbol{P}_\theta \boldsymbol{u}$ is the **rotation** of $\boldsymbol{u}$ about the **origin** by θ anticlockwise.

5.5.5.2 Composition of Rotation Matrices

- Moreover, $\boldsymbol{P}_\beta \boldsymbol{P}_\alpha = \boldsymbol{P}_{\alpha+\beta}$

5.5.5.3 Deriving Sum Laws for Sine and Cosine

- $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$
- $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$

Chapter 6

Diagonalization

6.1 Eigenvalues and Eigenvectors

6.1.1 Definition of Eigenvalues and Eigenvectors

Definition. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. Suppose that for some $\lambda \in \mathbb{R}$ and **nonzero** $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$. Then

- λ is an **eigenvalue** of $\boldsymbol{A}$.
- $\boldsymbol{v}$ is an **eigenvector** of $\boldsymbol{A}$ associated with λ.

6.1.2 Characteristic Polynomial

6.1.2.1 Definition of Characteristic Polynomial

Definition. Let $\boldsymbol{A}$ be a **square matrix**. $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ is the **characteristic polynomial** of $\boldsymbol{A}$.

6.1.2.2 Characteristic Polynomial is a Monic Polynomial of Degree n

Theorem. Let $\boldsymbol{A}$ be a **square matrix**. Then the **characteristic polynomial** of $\boldsymbol{A}$ is a **monic polynomial** in λ of **degree** n:

$$\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

6.1.3 Definition of Characteristic Equation

Definition. Let $\boldsymbol{A}$ be a **square matrix**. $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ is the **characteristic equation** of $\boldsymbol{A}$.

6.1.4 Eigenvalues are the Roots to the Characteristic Equation

Theorem. Let $\boldsymbol{A}$ be a **square matrix**. Then the **eigenvalues** of $\boldsymbol{A}$ are all the **roots** to the **characteristic equation** of $\boldsymbol{A}$.

6.1.5 Eigenvalue and Invertibility

Theorem. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. Then

0 is not an **eigenvalue** of $\boldsymbol{A} \Leftrightarrow \det(\boldsymbol{A}) \neq 0 \Leftrightarrow \boldsymbol{A}$ is **invertible**

6.1.6 Eigenvalues of Triangular Matrices are its Diagonal Entries

Theorem. Let $\boldsymbol{A}$ be a **triangular matrix**. Then its **eigenvalues** are all the **diagonal entries** of $\boldsymbol{A}$.

6.1.7 Eigenspace

6.1.7.1 Definition of Eigenspace

Definition. Let $\boldsymbol{A}$ be a **square matrix** and λ an **eigenvalue** of $\boldsymbol{A}$. Then the **eigenspace** of $\boldsymbol{A}$ associated to λ, denoted by E_λ (or $E_{\boldsymbol{A},\lambda}$), is the **nullspace** of $\lambda\boldsymbol{I} - \boldsymbol{A}$.

6.1.7.2 Properties of Eigenspace

6.1.7.2.1 Vectors in Eigenspace

Theorem. $E_{\boldsymbol{A},\lambda}$ consists of all the **eigenvectors** of $\boldsymbol{A}$ associated to λ and the **zero vector** $\boldsymbol{0}$.

6.1.7.2.2 Dimension of Eigenspace

Theorem. Since $\lambda\boldsymbol{I} - \boldsymbol{A}$ is **singular**, $\dim(E_\lambda) \geq 1$.

6.1.7.2.3 Eigenspace Associated To Eigenvalue Zero

Theorem. If $\boldsymbol{A}$ is **singular**, then $E_0 =$ **nullspace** of $\boldsymbol{A}$.

6.2 Diagonalization

6.2.1 Definition of Diagonalizable Matrix

Definition. Let $\boldsymbol{A}$ be a **square matrix**. $\boldsymbol{A}$ is **diagonalizable** if $\exists$ an **invertible matrix** $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a **diagonal matrix**.

6.2.2 Criterion for Diagonalizability

Theorem. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. Then

$\boldsymbol{A}$ is **diagonalizable** $\Leftrightarrow$ $\boldsymbol{A}$ has n **linearly independent eigenvectors**

6.2.3 Eigenvectors associated to Distinct Eigenvalues are Linearly Independent

Theorem. Let $\lambda_1, \ldots, \lambda_k$ be distinct **eigenvalues** of $\boldsymbol{A}$ and $\boldsymbol{v}_i$ be an **eigenvector** of $\boldsymbol{A}$ associated to λ_i. Then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are **linearly independent**.

6.2.4 Algorithm of Diagonalization

Let $\boldsymbol{A}$ be a **square matrix** of **order** n.

1. Find the **characteristic polynomial** $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$.

2. **Factorise** $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ over $\mathbb{R}$ (not $\mathbb{C}$) to find **eigenvalues** of $\boldsymbol{A}$.

   - If $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ cannot be completely factorised, then $\boldsymbol{A}$ is not **diagonalizable**.
   - If $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ can be completely factorised, say $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \prod\limits_{i=1}^{k} (\lambda - \lambda_i)^{r_i}$ where $\lambda_1, \ldots, \lambda_k$ are all distinct. Then
     - **algebraic multiplicity** of λ_i, denoted by $a(\lambda_i)$, is r_i
     - **geometric multiplicity** of λ_i, denoted by $g(\lambda_i)$, is $\dim(E_{\lambda_i})$
     - Moreover, $1 \leq g(\lambda_i) \leq a(\lambda_i)$, and $\sum\limits_{i=1}^{k} a(\lambda_i) = n$.

3. For each **eigenvalue** λ_i of $\boldsymbol{A}$, find a **basis** S_i for the **eigenspace** E_{λ_i}.

   - $\exists i \quad g(\lambda_i) < a(\lambda_i) \Rightarrow \sum\limits_{i=1}^{k} \dim(E_{\lambda_i}) < n \Rightarrow \boldsymbol{A}$ is not **diagonalizable**.

4. $\forall i \quad g(\lambda_i) = a(\lambda_i) \Rightarrow \sum\limits_{i=1}^{k} \dim(E_{\lambda_i}) = n \Rightarrow \boldsymbol{A}$ is **diagonalizable**, then

   - $\bigcup\limits_{i=1}^{k} S_i = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a **basis** for $\mathbb{R}^n$.
   - $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ **diagonalizes** $\boldsymbol{A}$, i.e. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$ is **diagonal**.
   - $\boldsymbol{v}_i$ is an **eigenvector** of $\boldsymbol{A}$ associated to the ith **diagonal entry** (eigenvalue) of $\boldsymbol{D}$.

6.2.5 Matrix with n Distinct Eigenvalues is Diagonalizable

Theorem. Let $\boldsymbol{A}$ be a **square matrix** of **order** n. $\boldsymbol{A}$ has n distinct **eigenvalues** $\Rightarrow \boldsymbol{A}$ is **diagonalizable**.

6.2.6 Application of Diagonalization

6.2.6.1 Matrix Exponentiation

Theorem. Suppose that $\boldsymbol{P}$ **diagonalizes** $\boldsymbol{A}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$. Then for

$$\begin{cases} m \in \mathbb{Z}, m \geq 0, & \boldsymbol{A} \text{ is } \textbf{singular} \\ m \in \mathbb{Z}, & \boldsymbol{A} \text{ is } \textbf{invertible} \end{cases} \qquad \boldsymbol{A}^m = \boldsymbol{P} \begin{pmatrix} \lambda_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^m \end{pmatrix} \boldsymbol{P}^{-1}$$

6.2.6.2 Example: Fibonacci Numbers

- The **fibonacci numbers** are defined as $a_n = \begin{cases} n, & n = 0 \text{ or } n = 1 \\ a_{n-1} + a_{n-2}, & n \geq 2 \end{cases}$
- Note that $a_{n+1} = a_{n-1} + a_n$ for $n \geq 1$
- Let $\boldsymbol{x}_n = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} a_n \\ a_{n-1} + a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}$ and $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Then
- $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \boldsymbol{A}^2\boldsymbol{x}_{n-2} = \cdots = \boldsymbol{A}^n\boldsymbol{x}_0$, where $\boldsymbol{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- To find $\boldsymbol{A}^n$ for large n, it will be easier to **diagonalize** $\boldsymbol{A}$.

6.3 Orthogonal Diagonalization

6.3.1 Definition of Orthogonally Diagonalizable Matrix

Definition. A **square matrix** $\boldsymbol{A}$ is **orthogonally diagonalizable** if it can be **diagonalized** by an **orthogonal matrix**, i.e.,

- $\exists$ an **orthogonal matrix** $\boldsymbol{P}$ such that $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a **diagonal matrix**.
- $\boldsymbol{P}$ is said to **orthogonally diagonalize** $\boldsymbol{A}$.

6.3.1.1 Remarks

- For any **eigenvalue** λ of $\boldsymbol{A}$, we can always choose an **orthonormal basis** for E_λ.
- Suppose further that $\boldsymbol{A}$ is **orthogonally diagonalizable**. Then
  - for distinct **eigenvalues** $\lambda \neq \mu$, every **eigenvector** of λ is **orthogonal** to that of μ.

6.3.2 Only Symmetric Matrices are Orthogonally Diagonalizable

Theorem. A **square matrix** is **orthogonally diagonalizable** $\Leftrightarrow$ it is a **symmetric matrix**.

6.3.3 Algorithm of Orthogonal Diagonalization

Let $\boldsymbol{A}$ be a **symmetric matrix** of order n.

1. Find the **characteristic polynomial** $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$.
2. **Factorise** $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ over $\mathbb{R}$ to find **eigenvalues** of $\boldsymbol{A}$.
   - $\det(\lambda \boldsymbol{I} - \boldsymbol{A})$ can definitely be completely factorised, say $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{r_i}$ where $\lambda_1, \dots, \lambda_k$ are all distinct.
     - $\sum_{i=1}^{k} a(\lambda_i) = n$.
3. For each **eigenvalue** λ_i of $\boldsymbol{A}$, find an **orthonormal basis** T_{λ_i} for E_{λ_i}.
   (a) Find a basis S_{λ_i} for E_{λ_i}.
   - $\forall i \quad 1 \leq g(\lambda_i) = a(\lambda_i) \Rightarrow \sum_{i=1}^{k} \dim(E_{\lambda_i}) = n$

   (b) Use Gram-Schmidt process to transfer S_{λ_i} to an **orthonormal basis** T_{λ_i} for E_{λ_i}.
4. $\forall i \quad g(\lambda_i) = a(\lambda_i) \Rightarrow \sum_{i=1}^{k} \dim(E_{\lambda_i}) = n \Rightarrow \boldsymbol{A}$ is **diagonalizable**, then
5. $\bigcup_{i=1}^{k} T_{\lambda_i} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ is an **orthonormal basis** for $\mathbb{R}^n$.
   - $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ **orthogonally diagonalizes** $\boldsymbol{A}$

6.4 Quadratic Forms and Conic Sections

6.4.1 Quadratic Forms

6.4.1.1 Definition of Quadratic Form

Definition. A **quadratic form/homogeneous** polynomial in degree 2 in n variables $x_1, \dots, x_n$ is

$$Q(x_1, \dots, x_n) = \sum_{i=1}^{n} q_{ii} x_i^2 + \sum_{i<j} q_{ij} x_i x_j$$

6.4.1.2 Matrix Representation of Quadratic Forms

Theorem. Let $Q(x_1, \dots, x_n) = \sum_{i=1}^{n} q_{ii} x_i^2 + \sum_{i<j} q_{ij} x_i x_j$ be a **quadratic form**. Let $\boldsymbol{x} = (x_1, \dots, x_n)^T$ and $\boldsymbol{A} = (a_{ij})_{n \times n}$ where $a_{ii} = q_{ii}$ and $a_{ij} = a_{ji} = \dfrac{1}{2} q_{ij}$ for $i < j$. Then $Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n$.

6.4.1.3 Simplification of Quadratic Forms

Theorem. Suppose $Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ is a **quadratic form** where $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ and $\boldsymbol{A}$ is a **symmetrix matrix** of order n. Then

1. $\exists$ an **orthogonal matrix** $\boldsymbol{P}$ such that $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$

2. Let $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$. Then

3. $Q(\boldsymbol{x}) = \sum\limits_{i=1}^{n} \lambda_i y_i^2$.

6.4.2 Quadratic Equation

6.4.2.1 Quadratic Equation in One Variable

Definition. A **quadratic equation** in variable x is of the form $ax^2 + bx = c$.

6.4.2.2 Quadratic Equation in Two Variables

Definition. A **quadratic equation** in variables x and y is

$$ax^2 + bxy + cy^2 + dx + ey = f$$

6.4.2.2.1 Matrix Representation

Let $\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\boldsymbol{A} = \begin{pmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{pmatrix}$, and $\boldsymbol{b} = \begin{pmatrix} d \\ e \end{pmatrix}$. Then $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} = f$.

6.4.2.2.2 Quadratic Form Associated With Quadratic Equation

Definition. $ax^2 + bxy + cy^2 = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ is the quadratic form **associated** with the quadratic equation.

6.4.2.3 Graph of a Quadratic Equation

Theorem. The graph of a **quadratic equation** is a **conic section**.

6.4.3 Classification of Conic Sections

6.4.3.1 Table of Conics

Degeneracy	Name	Equation / Standard Form
Degenerated	The whole plane $\mathbb{R}^2$	$0 = 0$
	Empty Set	$x^2 + y^2 = -1$
	A point	$x^2 + y^2 = 0$
	A line	$x = 0$ or $x^2 = 0$
	A pair of distinct lines	$x^2 - y^2 = 0$
Non-degenerated	Circle	$\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\alpha^2} = 1, \quad \alpha > 0$
	Ellipse	$\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0$
	Hyperbola	$\pm\dfrac{x^2}{\alpha^2} \mp \dfrac{y^2}{\beta^2} = 1, \quad \alpha > 0, \beta > 0$
	Parabola	$x^2 = \alpha y$ or $y^2 = \alpha x$

6.4.3.2 Algorithm to Classify Conic Sections

Given a **quadratic equation** $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} = f, \boldsymbol{x} \in \mathbb{R}^2$.

1. **Orthogonally diagonalize $\boldsymbol{A}$.**
   - $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$, $\boldsymbol{P}$ an **orthogonal matrix**.
2. Let $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x}$. Then
   - $\boldsymbol{y}^T \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \boldsymbol{y} + \boldsymbol{b}^T\boldsymbol{P}\boldsymbol{y} = f$
3. Complete the squares.

6.4.3.2.1 Determinant of Symmetric Matrix and Type of Conic Sections

Suppose the conic section is **non-degenerate**. Since $\lambda\mu = \det(\boldsymbol{A})$, then

- $\det(\boldsymbol{A}) > 0 \Leftrightarrow$ ellipse (or circle).
- $\det(\boldsymbol{A}) < 0 \Leftrightarrow$ hyperbola.
- $\det(\boldsymbol{A}) = 0 \Leftrightarrow$ parabola.

6.4.3.2.2 Rotating and Reflecting Conic Sections

Let $\boldsymbol{P}$ be **orthogonal** of order 2. Then $\det(\boldsymbol{P}) = \pm 1$.

- $\det(\boldsymbol{P}) = 1 \Rightarrow \boldsymbol{P} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Let $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x}$. Then

  - the new axes are obtained by rotating the original axes about origin anticlockwise by θ.

- $\det(\boldsymbol{P}) = -1 \Rightarrow \boldsymbol{P} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Let $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x}$. Then

  - the new axes are obtained by first rotating the original axes about origin anticlockwise by θ, then reflecting w.r.t. the x'-axis.

6.4.3.2.3 Can Always Orthogonally Diagonalize by Matrix with Determinant One

By multiplying the 2nd column of $\boldsymbol{P}$ by -1 if necessary, we can always **diagonalize** a **symmetric matrix** $\boldsymbol{A}$ by an **orthogonal matrix** with **determinant** 1.

Chapter 7

Linear Transformation

7.1 Linear Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

7.1.1 Definition of Linear Transformation

Definition. The mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

a **linear transformation** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

7.1.1.1 Linear Operator

T is a **linear operator** on $\mathbb{R}^n$ if $m = n$.

7.1.2 Linear Transformation As Matrix Form

A **linear transformation** is viewed as the **matrix form**:

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$T : \mathbb{R}^n \to \mathbb{R}^m$ such that $\forall \boldsymbol{x} \in \mathbb{R}^n \quad T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$

- $\boldsymbol{A} = (a_{ij})_{m\times n}$ is the **standard matrix** for T.

7.1.3 Examples of Linear Transformation

7.1.3.1 Identity Transformation

Definition. Let $I : \mathbb{R}^n \to \mathbb{R}^n$ be the **linear transformation** $\forall \boldsymbol{x} \in \mathbb{R}^n \quad I(\boldsymbol{x}) = \boldsymbol{x}$.

- It is the
  - **identity transformation**; or
  - **identity operator** on $\mathbb{R}^n$.
- $I(\boldsymbol{x}) = \boldsymbol{x} = \boldsymbol{I}_n\boldsymbol{x} \Rightarrow \boldsymbol{I}_n$ is the **standard matrix** for I.

7.1.3.2 Zero Transformation

Definition. Let $O : \mathbb{R}^n \to \mathbb{R}^m$ be the **linear transformation** $\forall \boldsymbol{x} \in \mathbb{R}^n \quad O(\boldsymbol{x}) = \boldsymbol{0}$.

- It is the **zero transformation**.
- $O(\boldsymbol{x}) = \boldsymbol{0} = \boldsymbol{0}_{m\times n}\boldsymbol{x} \Rightarrow \boldsymbol{0}_{m\times n}$ is the **standard matrix** for O.

7.1.4 Standard Matrix of a Linear Transformation is Unique

Theorem. The **standard matrix** of a **linear transformation** is unique.

7.1.5 To Prove that a Mapping is a Linear Transformation

To show that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation**, it suffices to find a **matrix $\boldsymbol{A}$** so that $\forall \boldsymbol{x} \in \mathbb{R}^n \quad T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$.

7.1.6 Linearity of Linear Transformations

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation**. Then

- $T(\boldsymbol{0}) = \boldsymbol{0}$
- $\forall c \in \mathbb{R}, \forall \boldsymbol{v} \in \mathbb{R}^n, \quad T(c\boldsymbol{v}) = cT(\boldsymbol{v})$
- $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \quad T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$
- $\forall c_1, \ldots, c_k \in \mathbb{R}, \forall \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n, \quad T(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) = c_1T(\boldsymbol{v}_1) + \cdots + c_kT(\boldsymbol{v}_k)$

7.1.7 To Prove that a Mapping is Not a Linear Transformation

To show that a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is **not** a **linear transformation**,

- Show that $T(\boldsymbol{0}) \neq \boldsymbol{0}$; or
- Find $c \in \mathbb{R}, \boldsymbol{v} \in \mathbb{R}^n$ such that $T(c\boldsymbol{v}) \neq cT(\boldsymbol{v})$; or
- Find $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ such that $T(\boldsymbol{u} + \boldsymbol{v}) \neq T(\boldsymbol{u}) + T(\boldsymbol{v})$.

7.1.8 Representation of Linear Transformations

7.1.8.1 Linear Transformations are Completely Determined by a Basis of $\mathbb{R}^n$ and its Images

Suppose that $T : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation** and $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a **basis** for $\mathbb{R}^n$. Then $\forall \boldsymbol{v} \in \mathbb{R}^n \quad (\boldsymbol{v})_S = (c_1, \ldots, c_n)$, and

$$\begin{aligned} T(\boldsymbol{v}) &= T(c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n) \\ &= c_1T(\boldsymbol{v}_1) + \cdots + c_nT(\boldsymbol{v}_n) \end{aligned}$$

$T(\boldsymbol{v})$ is completely determined by $T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)$. Let

- $\boldsymbol{A}$ be the **standard matrix** for T,
- $\boldsymbol{B} = \begin{pmatrix} T(\boldsymbol{v}_1) & \cdots & T(\boldsymbol{v}_n) \end{pmatrix}$, and
- $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$

Then $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{P}^{-1}$.

7.1.8.2 Columns of Standard Matrix are Images of Standard Basis of $\mathbb{R}^n$

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation** and $\boldsymbol{A}$ be the **standard matrix** for T. Then $\boldsymbol{A} = \begin{pmatrix} T(\boldsymbol{e}_1) & \cdots & T(\boldsymbol{e}_n) \end{pmatrix}$.

7.1.9 Linear Combination Preserving Implies Linear Transformation

7.1.9.1 With k vectors

Theorem. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation**, i.e., $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \Leftrightarrow \forall c_1, \ldots, c_k \in \mathbb{R}, \forall \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n, \quad T(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) = c_1T(\boldsymbol{v}_1) + \cdots + c_kT(\boldsymbol{v}_k)$

7.1.9.2 With Two Vectors

Theorem. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation**, i.e., $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} \Leftrightarrow \forall c, d \in \mathbb{R}, \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \quad T(c\boldsymbol{u} + d\boldsymbol{v}) = cT(\boldsymbol{u}) + dT(\boldsymbol{v})$.

7.1.10 General Definition of Linear Transformations

Definition. Let V and W be **vector spaces**. A mapping $T : V \to W$ is a **linear transformation** if

$$\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \forall c, d, \in \mathbb{R}, \quad T(c\boldsymbol{u} + d\boldsymbol{v}) = cT(\boldsymbol{u}) + dT(\boldsymbol{v})$$

7.1.11 Change of Bases

7.1.11.1 Change Between the Images of Two Bases

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation**. Let

- $S = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ and $R = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be **bases** for $\mathbb{R}^n$,
- $\boldsymbol{B} = \begin{pmatrix} T(\boldsymbol{u}_1) & \cdots & T(\boldsymbol{u}_n) \end{pmatrix}$ and $\boldsymbol{C} = \begin{pmatrix} T(\boldsymbol{v}_1) & \cdots & T(\boldsymbol{v}_n) \end{pmatrix}$, and
- $\boldsymbol{P}$ be the **transition matrix** from S to R and $\boldsymbol{Q} = \boldsymbol{P}^{-1}$

Then $\boldsymbol{B} = \boldsymbol{C}\boldsymbol{P}$ and $\boldsymbol{C} = \boldsymbol{B}\boldsymbol{Q}$.

7.1.11.2 Change Basis for Linear Operator

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a **linear operation** on $\mathbb{R}^n$ and $\boldsymbol{A}$ be the **standard matrix**. Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be a **basis** for $\mathbb{R}^n$ and $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix}$. Then

- $[T(\boldsymbol{v})]_S = \boldsymbol{B}[\boldsymbol{v}]_S$, where $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$
- T can be represented by $[\boldsymbol{v}]_S \mapsto \boldsymbol{B}[\boldsymbol{v}]_S$
- $\boldsymbol{A}$ and $\boldsymbol{B}$ are **similar**.

7.1.12 Composition

7.1.12.1 Definition of Function Composition

Definition. Let $f : X \to Y$ and $g : Y \to Z$ be two functions. The **composition** of g with f, denoted by $g \circ f$, is the function $X \to Z$ such that $\forall x \in X \quad (g \circ f)(x) = g(f(x))$

7.1.12.2 Function Composition is not Commutative in General

Note: In general, $g \circ f \neq f \circ g$.

7.1.12.3 Linear Transformation Composition

Definition. Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be **linear transformations**. The **composition** of T with S, denoted by $T \circ S$, is the mapping $\mathbb{R}^n \to \mathbb{R}^k$ such that $\forall \boldsymbol{u} \in \mathbb{R}^n, \quad (T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u}))$

7.1.12.4 Linear Transformation Composition is not Commutative in General

Note: In general, $T \circ S \neq S \circ T$.

7.1.12.5 Composition of Linear Transformations is a Linear Transformation and Formula for its Standard Matrix

Theorem. Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be **linear transformations**. Then

- $T \circ S : \mathbb{R}^n \to \mathbb{R}^k$ is a **linear transformation**

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be the **standard matrices** for S and T respectively. Then

- the **standard matrix** for $T \circ S$ is $\boldsymbol{BA}$.

7.1.12.6 Properties of Linear Transformation Composition

Theorem. Let $S, T, T_1, T_2, S_1, S_2, U$ be **linear transformations** and $c \in \mathbb{R}$. Then

- **Identity & Zero Composition:** $I \circ S = S \circ I = S$; $O \circ S = S \circ O = O$
- **Constant Multiple:** $c(T \circ S) = (cT) \circ S = T \circ (cS)$
- **Associativity:** $U \circ (T \circ S) = (U \circ T) \circ S$
- **Distributive Over Addition:**
  - $(T_1 + T_2) \circ S = T_1 \circ S + T_2 \circ S$
  - $T \circ (S_1 + S_2) = T \circ S_1 + T \circ S_2$

7.2 Ranges and Kernels

7.2.1 Range

7.2.1.1 Range of Function

Definition. Let $f : X \to Y$ be a **function**. The **range** of f, denoted by $\mathrm{R}(f)$, is the set of all **images** of f:

$$\mathrm{R}(f) = \{f(x) \mid x \in X\} \subseteq Y$$

7.2.1.2 Range of Linear Transformation

Definition. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation**. The **range** of T, denoted by $\mathrm{R}(T)$, is the set of all **images** of T:

$$\mathrm{R}(T) = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

7.2.1.3 Representation of Range

7.2.1.3.1 Range as Subspace Spanned by the Images of Any Basis for $\mathbb{R}^n$

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation** and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be any basis for $\mathbb{R}^n$. Then

- $\mathrm{R}(T) = \mathrm{span}\{T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)\}$.
- $\mathrm{R}(T)$ is a **subspace** of $\mathbb{R}^m$.

7.2.1.3.2 Range as Column Space of Standard Matrix

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation** and $\boldsymbol{A}$ be the **standard matrix** for T. Then $\mathrm{R}(T) =$ **column space** of $\boldsymbol{A}$.

7.2.2 Rank of Linear Transformation

Definition. Let T be a **linear transformation**. The **rank** of T, denoted by $\mathrm{rank}(T)$, is defined as

$$\mathrm{rank}(T) = \dim(\mathrm{R}(T))$$

Let $\boldsymbol{A}$ be the **standard matrix** for T. Then

$$\mathrm{rank}(T) = \dim(\mathrm{R}(T)) = \dim(\textbf{column space} \text{ of } \boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A})$$

7.2.3 Kernel

7.2.3.1 Definition of Kernel of Linear Transformation

Definition. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation**. The **kernel** of T, denoted by $\mathrm{Ker}(T)$, is

$$\mathrm{Ker}(T) = \{\boldsymbol{v} \in \mathbb{R}^n \mid T(\boldsymbol{v}) = \boldsymbol{0}\} \subseteq \mathbb{R}^n$$

Recall that $T(\boldsymbol{0}) = \boldsymbol{0} \Rightarrow \boldsymbol{0} \in \mathrm{Ker}(T) \subseteq \mathbb{R}^n$.

7.2.3.2 Representation of Kernel

7.2.3.2.1 Kernel as Nullspace of Standard Matrix and Subspace of $\mathbb{R}^n$

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation** and $\boldsymbol{A}$ be the **standard matrix** for T. Then $\mathrm{Ker}(T) =$ (**nullspace** of $\boldsymbol{A}$) which is a **subspace** of $\mathbb{R}^n$.

7.2.3.2.2 Finding General Solution for Kernel

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation** and $\boldsymbol{A}$ be the **standard matrix** for T. To find a general solution for $\mathrm{Ker}(T)$, solve the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$.

7.2.4 Nullity of Linear Transformation

Definition. Let T be a **linear transformation**. The **nullity** of T, denoted by nullity(T), is defined as

$$\text{nullity}(T) = \dim(\text{Ker}(T))$$

Let $\boldsymbol{A}$ be the **standard matrix** for T. Then

$$\text{nullity}(T) = \dim(\text{Ker}(T)) = \dim(\textbf{nullspace} \text{ of } \boldsymbol{A}) = \text{nullity}(\boldsymbol{A})$$

7.2.5 Dimension Theorem for Linear Transformations

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a **linear transformation**. Then $\text{rank}(T) + \text{nullity}(T) = n$.

7.2.6 Linear Transformation Between Vector Spaces

Let $T : V \to W$ be a **linear transformation** between **vector spaces**.

7.2.6.1 Range

Definition. $\text{R}(T) = \{T(\boldsymbol{v} \mid \boldsymbol{v} \in V\}$ is a **subspace** of W.

7.2.6.2 Rank

Definition. $\text{rank}(T) = \dim(\text{R}(T))$.

7.2.6.3 Kernel

Definition. $\text{Ker}(T) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$ is a **subspace** of V.

7.2.6.4 Nullity

Definition. $\text{nullity}(T) = \dim(\text{Ker}(T))$.

7.2.6.5 Dimension Theorem

Definition. $\text{rank}(T) + \text{nullity}(T) = \dim(V)$.

7.3 Geometric Linear Transformations

7.3.1 Geometric Interpretation Completely Determined by Effect on a Basis for its Domain

Since the images of any **basis** for $\mathbb{R}^n$ completely determines a **linear transformation** $T : \mathbb{R}^n \to \mathbb{R}^m$, to study the geometric interpretation of a **linear transformation**, it suffices to check the effect of the **linear transformation** on a **basis** for its domain.

7.3.2 Scaling

7.3.2.1 Scaling in $\mathbb{R}^2$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** and $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ be the **standard matrix** for T, where $\lambda_1, \lambda_2 > 0$.

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda_1 x \\ \lambda_2 y \end{pmatrix}$$

Then T is a **scaling** in $\mathbb{R}^2$,

- along the x-axis by a factor of λ_1, and
- along the y-axis by a factor of λ_2.

7.3.2.1.1 Dilation & Contraction

Suppose further that $\lambda_1 = \lambda_2$. Let $\lambda = \lambda_1 = \lambda_2$. Then

- T is a **dilation** if $\lambda > 1$.
- T is a **contraction** if $0 < \lambda < 1$.

7.3.2.2 Scaling in other Axes in $\mathbb{R}^2$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** and $\boldsymbol{A}$ be the **standard matrix**.

- Suppose $\boldsymbol{A}$ is **diagonalizable**, say $\exists \boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, $\lambda_1, \lambda_2 > 0$
- Suppose $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix}$. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. Then S is a **basis** for $\mathbb{R}^2$.
- $\forall \boldsymbol{v} \in \mathbb{R}^2, \quad [T(\boldsymbol{v})]_S = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} [\boldsymbol{v}]_S$ (from section 7.1.11.2)
- Then T can be viewed as a scaling
  - along the direction of $\boldsymbol{v}_1$ by factor λ_1,
  - along the direction of $\boldsymbol{v}_2$ by factor λ_2,

7.3.2.3 Scaling in $\mathbb{R}^3$

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a **linear transformation** with **standard matrix** $\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, where $\lambda_1, \lambda_2, \lambda_3 > 0$.

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda_1 x \\ \lambda_2 y \\ \lambda_3 z \end{pmatrix}$$

Then T is a **scaling**,

- along the x-axis by a factor of λ_1,
- along the y-axis by a factor of λ_2,
- along the z-axis by a factor of λ_3.

7.3.2.3.1 Dilation & Contraction

Suppose further that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$. Then

- T is a **dilation** if $\lambda > 1$.
- T is a **contraction** if $0 < \lambda < 1$.

7.3.2.4 Scaling in other Axes in $\mathbb{R}^3$

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a **linear transformation** with **standard matrix** $\boldsymbol{A}$.

- Suppose $\boldsymbol{A}$ is **diagonalizable**, say $\exists \boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \lambda_1, \lambda_2, \lambda_3 > 0$
- Suppose $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix}$. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. Then S is a **basis** for $\mathbb{R}^3$.
- $\forall \boldsymbol{v} \in \mathbb{R}^3, \quad [T(\boldsymbol{v})]_S = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} [\boldsymbol{v}]_S$ (from section 7.1.11.2)
- Then T can be viewed as a scaling
  - along the direction of $\boldsymbol{v}_1$ by factor λ_1,
  - along the direction of $\boldsymbol{v}_2$ by factor λ_2,
  - along the direction of $\boldsymbol{v}_3$ by factor λ_3.

7.3.3 Reflection

7.3.3.1 Reflection in $\mathbb{R}^2$

7.3.3.1.1 Reflection w.r.t. x-axis

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** with **standard matrix** $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ -y \end{pmatrix}$$

T is the **reflection** w.r.t. the x-axis.

7.3.3.1.2 Reflection w.r.t. y-axis

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** with **standard matrix** $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} -x \\ y \end{pmatrix}$$

T is the **reflection** w.r.t. the y-axis.

7.3.3.1.3 Reflection w.r.t. the line $y = x$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** with **standard matrix** $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix}$$

T is the **reflection** w.r.t. the line $y = x$.

7.3.3.1.4 Reflection w.r.t. a Line Through the Origin

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a **linear transformation** with **standard matrix** $\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$
T is the **reflection** w.r.t. the line ℓ passing through the origin where θ is the angle between ℓ and the x-axis.

Orthogonal Matrices of Determinant -1
Every **orthogonal matrix** of $\det = -1$ is in this form.

Representation with Unit Vector Parallel to Line

- Let $\boldsymbol{n} = (\cos\theta, \sin\theta)^T$ be a unit vector on ℓ.
- $\boldsymbol{p}$ is the projection of $\boldsymbol{v}$ onto $\operatorname{span}\{\boldsymbol{n}\}$
  - $\boldsymbol{p} = (\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n}$
- $\boldsymbol{p}$ is the midpoint of $\boldsymbol{v}$ and $T(\boldsymbol{v})$
  - $T(\boldsymbol{v}) = 2\boldsymbol{p} - \boldsymbol{v} = 2(\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n} - \boldsymbol{v}$

Representation with Unit Vector Orthogonal to Line

- Let $\boldsymbol{n} = (\sin\theta, -\cos\theta)^T$ be a unit vector orthogonal to ℓ.
- $\boldsymbol{p}$ is the projection of $\boldsymbol{v}$ onto $\operatorname{span}\{\boldsymbol{n}\}$
  - $\boldsymbol{p} = (\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n}$
- Note that $T(\boldsymbol{v}) + 2\boldsymbol{p} = \boldsymbol{v}$
  - $T(\boldsymbol{v}) = \boldsymbol{v} - 2\boldsymbol{p} = \boldsymbol{v} - 2(\boldsymbol{v} \cdot \boldsymbol{n})\boldsymbol{n}$

7.3.3.2 Reflections in $\mathbb{R}^3$

7.3.3.2.1 Reflections w.r.t. Planes Formed by Coordinate Axes

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a **linear transformation**.

- If the **standard matrix** is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$,
  - then T is the reflection w.r.t. the xy-plane.
- If the **standard matrix** is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,
  - then T is the reflection w.r.t. the xz-plane.
- If the **standard matrix** is $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,
  - then T is the reflection w.r.t. the yz-plane.

7.3.3.2.2 Reflections w.r.t. Any Plane

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the reflection w.r.t. the plane $ax + by + cz = 0$, where a, b, c not all zero. Then $\boldsymbol{n} = (a, b, c)^T$ and $\forall \boldsymbol{v} \in \mathbb{R}^3, \quad T(\boldsymbol{v}) = \boldsymbol{v} - \left(2\frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\|\boldsymbol{n}\|^2}\right) \boldsymbol{n}$.

7.3.4 Rotation

7.3.4.1 Rotation in $\mathbb{R}^2$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the **rotation** about the origin by θ.

- Then T is a **linear transformation**.
- The **standard matrix** for T is $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

7.3.4.1.1 Orthogonal Matrices of Determinant 1

Every **orthogonal matrix** of det $= 1$ is in this form.

7.3.4.2 Rotation in $\mathbb{R}^3$

- Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the **rotation** about the z-axis anticlockwise by angle θ.
  - The z-coordinate does not change
  - It is the rotation about the origin on the plane $z = z_0$ anticlockwise by θ.

- **Standard Matrix** $\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$

- Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the **rotation** about the x-axis anticlockwise by angle θ.
  - The x-coordinate does not change
  - It is the rotation about the origin on the plane $x = x_0$ anticlockwise by θ.
  - **Standard Matrix** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$

- Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the **rotation** about the y-axis anticlockwise by angle θ.
  - The y-coordinate does not change
  - It is the rotation about the origin on the plane $y = y_0$ anticlockwise by θ.
  - **Standard Matrix** $\begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$

7.3.5 Reflections and Rotations in $\mathbb{R}^2$

7.3.5.1 Determinant Determines Rotation vs Reflection

Suppose **standard matrix $\boldsymbol{A}$** for $T : \mathbb{R}^2 \to \mathbb{R}^2$ is **orthogonal**.

- If $\det(\boldsymbol{A}) = 1$, T represents a **rotation** about the origin.
- If $\det(\boldsymbol{A}) = -1$, T represents the **reflection** w.r.t. a line passing through the origin.

7.3.5.2 Reflection about a Line can be Decomposed into Reflection and Rotation

Since the **standard matrix** for **reflection** about a line ℓ passing through the origin where θ is the angle between ℓ and the x-axis is

$$\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) \\ \sin(2\theta) & \cos(2\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Reflection w.r.t. ℓ can be decomposed into

1. Reflection w.r.t. x-axis
2. Rotation about the origin anticlockwise by 2θ.

7.3.6 Shears

7.3.6.1 Shears in $\mathbb{R}^2$

7.3.6.1.1 Shear in x-direction

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + ky \\ y \end{pmatrix}$$

Then T is a **shear** in the x-direction by a factor k.

7.3.6.1.2 Shear in y-direction

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ kx + y \end{pmatrix}$$

Then T is a **shear** in the y-direction by a factor k.

7.3.6.2 Shears in $\mathbb{R}^3$

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + k_1 z \\ y + k_2 z \\ z \end{pmatrix}$$

Then T is a **shear** in the x-direction by a factor k_1, and in the y-direction by factor k_2.

- On yz-plane $x = 0$, it is a shear in y-direction by k_2.
- On xz-plane $y = 0$, it is a shear in x-direction by k_1.
- On the plane $z = 1$,
  - $T\left(\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}\right) = \begin{pmatrix} x + k_1 \\ y + k_2 \\ 1 \end{pmatrix}$.

7.3.7 Translations

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + a \\ y + b \end{pmatrix}, a, b \in \mathbb{R}$$

- T is a **translation** by $(a, b)^T$.
- T is **not** a **linear transformation** unless $a = b = 0$.

7.3.8 2D Computer Graphic System

7.3.8.1 Representation of 2D Figures

- In 2D computer graphic, a figure is drawn by connecting points $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$.
- It can be written as an $2 \times n$ matrix:
  - $\boldsymbol{M} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}$.

7.3.8.2 Scaling, Reflection, Rotation, Shearing of 2D Figures

- Let T be a scaling/reflection/rotation/shearing on $\mathbb{R}^2$
  - Then T is a **linear transformation** with **standard matrix $\boldsymbol{A}$**.
- Let $\boldsymbol{M} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{pmatrix}$ be a 2D figure.
- Then the resulting figure by T is
  - $\begin{pmatrix} T(\boldsymbol{v}_1) & \cdots & T(\boldsymbol{v}_n) \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \cdots & \boldsymbol{A}\boldsymbol{v}_n \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix} = \boldsymbol{AM}$.

7.3.8.3 Translating 2D Figure with Homogeneous Coordinate System

- **Homogeneous coordinate system** is formed by identifying $\mathbb{R}^2$ with plane $z = 1$ in $\mathbb{R}^3$:
  $\begin{pmatrix} a \\ b \end{pmatrix} \leftrightarrow \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$
- A figure $(a_1, b_1), \ldots, (a_n, b_n)$ is identified by $(a_1, b_1, 1), \ldots, (a_n, b_n, 1)$.
- The associated matrix $\boldsymbol{M} = \begin{pmatrix} a_1 & \cdots & a_n \\ b_1 & \cdots & b_n \\ 1 & \cdots & 1 \end{pmatrix}$
- Suppose we want to do a translation by $(a, b)^T$.
- Define shear $T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + az \\ y + bz \\ z \end{pmatrix}$ with **standard matrix $\boldsymbol{A}$** $= \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$
- $\boldsymbol{AM} = \begin{pmatrix} a_1 + a & \cdots & a_n + a \\ b_1 + b & \cdots & b_n + b \\ 1 & \cdots & 1 \end{pmatrix}$ represent the translation by $(a, b)^T$.