MATH 1853: Introduction to Linear Algebra

Tutorial 1: Matrix

Wenyong Zhou

HKU EEE

September 2024

Who am I

- Wenyong Zhou (Richard)
- Contact me: wenyongz@connect.hku.hk
- Office: Room 806, Chow Yei Ching Building

Important Notes

Before coming, please send me an email about:

- when will you come
- what questions you have

Note: The tutorial and lecture contents **cannot be in sync** due to the tight scheduling. You should treat tutorials as extra materials for Math 1853.

Outline

What is Linearity (Linear)

Consider an equation of the form $ax + by = c$ (a, b, and c are constants), which, when solved, forms a straight line on a plane. In this case, x and y exhibit a linear relationship. The equation $ax + by = c$ is called a linear equation.

Extending this idea to n variables, we have the general form of a first-degree equation:

$$k_1x_1 + k_2x_2 + \cdots + k_nx_n = b$$

This is called a linear equation, where x_1, x_2, $\cdots$, x_n are variables, and k_1, k_2, $\cdots$, k_n, b are constants. The variables x_1, x_2, $\cdots$, x_n exhibit a linear relationship.

What is Linear Algebra

Linear Algebra is a branch of algebra that primarily deals with linear relationships between variables. The core content of linear algebra involves:

1. The study of the structure of finite-dimensional linear spaces.
2. Linear transformations between these spaces.

This course introduces the fundamental concepts of linear algebra, with a central focus on solving systems of linear equations.

Gaussian Elimination

In general, a system of linear equations with n unknowns $x_1, x_2, \ldots, x_n$ can be written as:

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

This system contains m equations. Here, a_{ij} are the coefficients, b_i are constant terms, with $i = 1, 2, \ldots, m$, and $j = 1, 2, \ldots, n$. The coefficient a_{ij} has two subscripts, where i denotes the row and j denotes the column. **Gaussian Elimination** is a classic method for solving systems of linear equations. It is simple, practical, and timeless.

Definition of a Matrix

An $m \times n$ matrix is a rectangular array of numbers with m rows and n columns:

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

This is called an $m \times n$ matrix, where a_{ij} represents the number located in the i-th row and j-th column of the matrix. These numbers are also called the **elements** of the matrix.

- **Real Matrix:** A matrix whose elements are real numbers.
- **Complex Matrix:** A matrix whose elements are complex numbers.

Types of Matrices

- **Square Matrix**: A matrix with the same number of rows and columns ($m = n$).
- **Row Matrix**: A matrix with a single row ($1 \times n$).
- **Column Matrix**: A matrix with a single column ($m \times 1$).
- **Zero Matrix**: A matrix in which all elements are zero.
- **Identity Matrix**: A square matrix with ones on the diagonal and zeros elsewhere.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrix Addition and Scalar Multiplication

1. Let $\mathbf{A}$ and $\mathbf{B}$ be matrices of **the same size**. The sum $\mathbf{A}+\mathbf{B}$ is defined as:

$$(\mathbf{A}+\mathbf{B})_{ij} = \mathbf{A}_{ij} + \mathbf{B}_{ij}.$$

2. Scalar multiplication is defined as $k\mathbf{A}$, where k is a scalar:

$$(k\mathbf{A})_{ij} = k \cdot \mathbf{A}_{ij}, \quad \text{where } k \text{ is a scalar. In particular, } -\mathbf{A} = (-1)\mathbf{A}.$$

3. The subtraction of matrices is defined as:

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$$

Matrix Multiplication (Key Concept)

Let $\mathbf{A} = [a_{ik}]_{m\times p}$ and $\mathbf{B} = [b_{kj}]_{p\times n}$. The product $\mathbf{C} = [c_{ij}]_{m\times n}$ of $\mathbf{A}$ and $\mathbf{B}$ is defined as:

$$\mathbf{C} = \mathbf{AB},$$

where

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

That is, the element in the i-th row and j-th column of the product matrix $\mathbf{AB}$ is the inner product of the i-th row of $\mathbf{A}$ and the j-th column of $\mathbf{B}$.

$$\begin{pmatrix} * & * & \dots & * \\ a_{i1} & a_{i2} & \dots & a_{ip} \\ * & * & \dots & * \end{pmatrix} \begin{pmatrix} * & b_{1j} & * \\ * & b_{2j} & * \\ \vdots & \vdots & \vdots \\ * & b_{pj} & * \end{pmatrix} = \begin{pmatrix} * & * & \dots & * \\ * & c_{ij} & * & \\ * & * & \dots & * \end{pmatrix}$$

Matrix Multiplication

Matrix multiplication does not satisfy the commutative property:

(a) Even if $\mathbf{A}$ and $\mathbf{B}$ can be multiplied, $\mathbf{BA}$ may not be defined.

(b) Even if both $\mathbf{AB}$ and $\mathbf{BA}$ are defined, they may not be of the same type.

(c) Even if $\mathbf{AB}$ and $\mathbf{BA}$ are of the same type, they are not necessarily equal.

Example:

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} -2 & -1 \\ 3 & 2 \end{pmatrix}$$

In this example, we see that $\mathbf{AB} \neq \mathbf{BA}$.

Matrix Multiplication

The product of two non-zero matrices can be a zero matrix.

Example: $$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \times \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Matrix Multiplication

The cancellation law does not hold for matrix multiplication:

$$\mathbf{AB} = \mathbf{CB} \quad \not\Rightarrow \quad \mathbf{A} = \mathbf{C}$$

Example:

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 2 & 0 \\ 4 & 5 \end{pmatrix}$$

$$\mathbf{BA} = \mathbf{CA} = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$$

Matrix Powers

Let $\mathbf{A}$ be an $n \times n$ matrix, and let m be a natural number. Define $\mathbf{A}^m = \mathbf{A} \times \mathbf{A} \times \cdots \times \mathbf{A}$ (where m copies of $\mathbf{A}$ are multiplied). For a polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0$, define

$$f(\mathbf{A}) = a_n \mathbf{A}^n + a_{n-1} \mathbf{A}^{n-1} + \cdots + a_2 \mathbf{A}^2 + a_1 \mathbf{A} + a_0 \mathbf{I}.$$

Operational Rules:

(1) $\mathbf{A}^k \mathbf{A}^l = \mathbf{A}^{k+l}$

(2) $(\mathbf{A}^k)^l = \mathbf{A}^{kl}$

Matrix Transpose

For an $m \times n$ matrix $\mathbf{A} = [a_{ij}]$, the transpose of $\mathbf{A}$ is an $n \times m$ matrix $\mathbf{B} = [b_{ij}]$, where:

$$b_{ij} = a_{ji},$$

denoted as $\mathbf{B}$ being the transpose of $\mathbf{A}$, written as $\mathbf{A}^T$. Therefore, by definition, we have:

$$(\mathbf{A}^T)_{ij} = \mathbf{A}_{ji}.$$

Rules of Operation:

1. $(\mathbf{A}^T)^T = \mathbf{A}$
2. $(\mathbf{B} + \mathbf{C})^T = \mathbf{B}^T + \mathbf{C}^T$
3. $(k\mathbf{A})^T = k\mathbf{A}^T$
4. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$, $(\mathbf{A}_1\mathbf{A}_2 \dots \mathbf{A}_m)^T = \mathbf{A}_m^T \dots \mathbf{A}_2^T\mathbf{A}_1^T$
5. $(\mathbf{A}^m)^T = (\mathbf{A}^T)^m$
6. $|\mathbf{A}| = |\mathbf{A}^T|$
7. $\mathbf{A}$ is symmetric if and only if $\mathbf{A} = \mathbf{A}^T$. $\mathbf{A}$ is skew-symmetric if and only if $\mathbf{A} = -\mathbf{A}^T$.

Invertible Matrix (Key Concept)

Let $\mathbf{A}$ be an $n \times n$ matrix. If there exists an $n \times n$ matrix $\mathbf{B}$ such that:

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I},$$

then $\mathbf{B}$ is called the **inverse** of $\mathbf{A}$, denoted as $\mathbf{B} = \mathbf{A}^{-1}$. If no such $\mathbf{B}$ exists, then $\mathbf{A}$ is called **non-invertible** or **singular**.

Theorem
If a square matrix $\mathbf{A}$ is invertible, then its inverse is unique.

Proof: Assume both $\mathbf{B}$ and $\mathbf{C}$ are inverses of $\mathbf{A}$. Then:

$$\mathbf{B} = \mathbf{BI} = \mathbf{B}(\mathbf{AC}) = (\mathbf{BA})\mathbf{C} = \mathbf{IC} = \mathbf{C}.$$

Thus, the inverse of $\mathbf{A}$ is unique.

Linear Transformations

Suppose there are two linear transformations:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \end{cases} \qquad (1)$$

$$\begin{cases} x_1 = b_{11}t_1 + b_{12}t_2, \\ x_2 = b_{21}t_1 + b_{22}t_2, \\ x_3 = b_{31}t_1 + b_{32}t_2. \end{cases} \qquad (2)$$

How to represent the y_1 and y_2 in terms of t_1 and t_2?

Solution

Substituting (2) into (1), we can obtain the linear transformation from t_1, t_2 to y_1, y_2:

$$\begin{cases} y_1 = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})t_1 + (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32})t_2, \\ y_2 = (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31})t_1 + (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32})t_2. \end{cases}$$

Matrix Multiplication Example

Compute the following product:

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

Solution

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{I} \\ \mathbf{O} & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{B}_1 \\ \mathbf{O} & \mathbf{B}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1\mathbf{I} & \mathbf{A}_1\mathbf{B}_1 + \mathbf{B}_2 \\ \mathbf{O} & \mathbf{A}_2\mathbf{B}_2 \end{pmatrix}$$

which simplifies to:

$$\begin{pmatrix} 1 & 2 & 5 & 2 \\ 0 & 1 & 2 & -4 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & -9 \end{pmatrix}.$$

Matrix Operations Example

Given matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & 4 \\ 0 & 5 & 1 \end{pmatrix},$$

find $3\mathbf{AB} - 2\mathbf{A}$ and $\mathbf{A}^T\mathbf{B}$.

Solution

$$3\mathbf{AB} - 2\mathbf{A} = 3\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ -1 & -2 & 4\\ 0 & 5 & 1\end{pmatrix} - 2\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix}$$

$$= 3\begin{pmatrix}0 & 5 & 8\\ 0 & -5 & 6\\ 2 & 9 & 0\end{pmatrix} - 2\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix} = \begin{pmatrix}-2 & 13 & 22\\ -2 & -17 & 20\\ 4 & 29 & -2\end{pmatrix}$$

$$\mathbf{A}^T\mathbf{B} = \begin{pmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ -1 & -2 & 4\\ 0 & 5 & 1\end{pmatrix} = \begin{pmatrix}0 & 5 & 8\\ 0 & -5 & 6\\ 2 & 9 & 0\end{pmatrix}$$

Linear Transformation Example

Given the linear transformations

$$\begin{cases} x_1 = 2y_1 + y_3, \\ x_2 = -2y_1 + 3y_2 + 2y_3, \\ x_3 = 4y_1 + y_2 + 5y_3, \end{cases} \qquad \begin{cases} y_1 = -3z_1 + z_2, \\ y_2 = 2z_1 + z_3, \\ y_3 = -z_2 + 3z_3, \end{cases}$$

find the linear transformation from z_1, z_2, z_3 to x_1, x_2, x_3.

Solution

Using matrix knowledge:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & 1 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & 1 & 5 \end{pmatrix} \begin{pmatrix} -3 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

Therefore,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -6 & 1 & 3 \\ 12 & -4 & 9 \\ -10 & -1 & 16 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix},$$

leading to the system of equations:

$$\begin{cases} x_1 = -6z_1 + z_2 + 3z_3, \\ x_2 = 12z_1 - 4z_2 + 9z_3, \\ x_3 = -10z_1 - z_2 + 16z_3. \end{cases}$$

Matrix Operations Example

Given matrices $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, determine:

1. Does $\mathbf{AB} = \mathbf{BA}$?
2. Is $(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$?
3. Is $(\mathbf{A} + \mathbf{B})(\mathbf{A} - \mathbf{B}) = \mathbf{A}^2 - \mathbf{B}^2$?

Solution - Part 1

(1) Since

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix},$$

$$\mathbf{BA} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 8 \end{pmatrix},$$

it follows that $\mathbf{AB} \neq \mathbf{BA}$.

Solution - Part 2

(2) Since

$$(\mathbf{A}+\mathbf{B})^2 = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 8 & 14 \\ 14 & 29 \end{pmatrix},$$

but

$$\mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2 = \begin{pmatrix} 3 & 8 \\ 4 & 11 \end{pmatrix} + \begin{pmatrix} 6 & 8 \\ 8 & 12 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 16 \\ 15 & 27 \end{pmatrix},$$

it follows that $(\mathbf{A}+\mathbf{B})^2 \neq \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2$.

Solution - Part 3

(3) Since

$$(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 6 \\ 0 & 9 \end{pmatrix},$$

but

$$\mathbf{A}^2 - \mathbf{B}^2 = \begin{pmatrix} 3 & 8 \\ 4 & 11 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ 1 & 7 \end{pmatrix},$$

it follows that $(\mathbf{A}+\mathbf{B})(\mathbf{A}-\mathbf{B}) \neq \mathbf{A}^2 - \mathbf{B}^2$.

Thank you!

MATH 1853: Introduction to Linear Algebra

Tutorial 2: Solving Linear Systems

Wenyong Zhou

HKU EEE

September 2024

Who am I

- Wenyong Zhou (Richard)
- Contact me: wenyongz@connect.hku.hk
- Office: Room 806, Chow Yei Ching Building

Important Notes

Before coming, please send me an email about:

- when will you come
- what questions you have

Note: The tutorial and lecture contents **cannot be in sync** due to the tight scheduling. You should treat tutorials as extra materials for Math 1853.

Outline

n-Dimensional Vector

An ordered tuple $a_1, a_2, \ldots, a_n$ of n elements from the field $\mathbb{R}$ is called an n-dimensional vector over field $\mathbb{R}$, abbreviated as n-dimensional vector, denoted as

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

where a_i is called the i-th component of $\mathbf{a}$.

Linear Combination

Let $\mathbf{a}_i \in \mathbb{R}^n$, $k_i \in \mathbb{R}$, $i = 1, 2, \ldots, m$. Then the vector

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m$$

is called a linear combination of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ over the field $\mathbb{R}$.

Linear Combination

Given the vector set $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ and a vector β, if there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that

$$\beta = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \cdots + \lambda_m \mathbf{a}_m,$$

then vector β is said to be a linear combination of the vector set $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$.

Linear System Representation

Given the linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an $m \times n$ matrix, let

$$\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n),$$

then

$$(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{b},$$

which means that the linear system can be equivalently expressed as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}.$$

Linear Combination

The vector $\mathbf{b}$ can be expressed as a linear combination of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, which is equivalent to the system of equations

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}.$$

has a solution.

Linear Combination

For example, let

$$\mathbf{a}_1 = (a_{11}, a_{21}, a_{31}, a_{41})^\top, \quad \mathbf{a}_2 = (a_{12}, a_{22}, a_{32}, a_{42})^\top,$$

$$\mathbf{a}_3 = (a_{13}, a_{23}, a_{33}, a_{43})^\top,$$

then

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3$$

is equivalent to

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \\ a_{43} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 \end{pmatrix}.$$

Linear Combination

Thus, equivalently, the expression

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$$

is equivalent to the system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 = b_4. \end{cases}$$

Linear Independent

Given a set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, if there exist non-zero scalars $k_1, k_2, \ldots, k_m \in \mathbb{R}$ such that

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_m\mathbf{a}_m = 0, \tag{1}$$

then the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is said to be *linearly dependent*. Otherwise, they are called *linearly independent*.

Linear Independent

The set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is linearly dependent, which is equivalent to the homogeneous system of equations

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_m\mathbf{a}_m = 0$$

having a non-trivial solution.

Linear Independent

The set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is linearly independent, which is equivalent to the homogeneous system of equations

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_m\mathbf{a}_m = 0$$

having only the trivial solution.

Example

Let n-dimensional vectors

$$\mathbf{e}_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0),$$

where the i-th component is 1, and the remaining components are 0. Then, the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are linearly independent.

Example

Proof: Let

$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n = 0.$$

This implies

$$(x_1, x_2, \ldots, x_n) = 0,$$

which means that $x_1 = x_2 = \cdots = x_n = 0$. Thus, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are linearly independent.

Example

A set of vectors that contains the zero vector is linearly dependent.

Example

Proof: Let the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ contain $\mathbf{a}_1 = 0$. Then, there exist non-zero scalars $1, 0, 0, \ldots, 0$ such that

$$1\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 0\mathbf{a}_m = 0.$$

Thus, the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is linearly dependent.

Example

If a subset of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is linearly dependent, then the entire set of vectors is also linearly dependent.

Example

Proof: Suppose the first j vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_j$ are linearly dependent, where $j \leq m$. Then, there exist non-zero scalars $k_1, k_2, \ldots, k_j$ such that

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_j\mathbf{a}_j = 0,$$

and thus there exist non-zero scalars $k_1, k_2, \ldots, k_j, 0, \ldots, 0$ such that

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_j\mathbf{a}_j + 0\mathbf{a}_{j+1} + \cdots + 0\mathbf{a}_m = 0,$$

proving that the entire set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is linearly dependent.

Example

Any $n+1$ vectors in n-dimensional space must be linearly dependent.

Example

Proof: For the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{a}_{n+1} \in \mathbb{R}^n$, assume

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_{n+1}\mathbf{a}_{n+1} = 0.$$

Notice that in this homogeneous system of linear equations, the number of unknowns is $n+1$, while the number of equations is n. Therefore, there must exist non-trivial solutions.

Thus, $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \mathbf{a}_{n+1}$ are linearly dependent.

Example

If the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ is linearly independent, and the set $\beta, \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ is linearly dependent, then β can be expressed as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$, and the expression is unique.

Example

Proof: There exist non-zero scalars $k, k_1, \ldots, k_r$ such that

$$k\beta + k_1\mathbf{a}_1 + \cdots + k_r\mathbf{a}_r = 0, \tag{1}$$

which implies that $k \neq 0$. In fact, assuming $k = 0$, then equation (1) becomes

$$k_1\mathbf{a}_1 + \cdots + k_r\mathbf{a}_r = 0,$$

and since $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ are linearly independent, we must have

$$k_1 = k_2 = \cdots = k_r = 0.$$

This contradicts the assumption that $k, k_1, \ldots, k_r$ are not all zero. Thus, we have

$$\beta = -\frac{k_1}{k}\mathbf{a}_1 - \cdots - \frac{k_r}{k}\mathbf{a}_r.$$

Hence, β can be expressed as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$.

Uniqueness of Representation

We now prove the uniqueness of the representation. Suppose β has two representations:

$$\beta = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + \cdots + l_r\mathbf{a}_r, \quad \beta = h_1\mathbf{a}_1 + h_2\mathbf{a}_2 + \cdots + h_r\mathbf{a}_r.$$

Subtracting the two equations gives

$$(l_1 - h_1)\mathbf{a}_1 + (l_2 - h_2)\mathbf{a}_2 + \cdots + (l_r - h_r)\mathbf{a}_r = 0.$$

Since $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ are linearly independent, we must have

$$l_1 - h_1 = l_2 - h_2 = \cdots = l_r - h_r = 0,$$

which implies $l_i = h_i$, for $i = 1, 2, \ldots, r$. Therefore, the representation is unique.

Example 1

Let $\mathbf{a}_1 = (1, -1, 1)$, $\mathbf{a}_2 = (1, 2, 0)$, and $\mathbf{a}_3 = (1, 0, 3)$. Are $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ linearly dependent?

Solution

Suppose

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = 0. \tag{3}$$

The system of equations can be written in matrix form as

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 3 \end{vmatrix} = 7 \neq 0.$$

Thus, the system of equations (3) has only the trivial solution. Therefore, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are linearly independent.

Example 2

Let the set of vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ be linearly independent. Additionally, $\beta_1 = \mathbf{a}_1 + \mathbf{a}_2 + 2\mathbf{a}_3$, $\beta_2 = \mathbf{a}_1 - \mathbf{a}_2$, and $\beta_3 = \mathbf{a}_1 + \mathbf{a}_3$. Prove that $\beta_1, \beta_2, \beta_3$ are linearly dependent.

Solution

Proof: Since

$$\beta_1 = -\beta_2 + 2\beta_3,$$

it follows that $\beta_1, \beta_2, \beta_3$ are linearly dependent.

Exercise 3

Express the vector $\mathbf{a}$ as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$:

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

Solution

Suppose $\mathbf{a} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 + x_4\mathbf{a}_4$, then the system of equations is:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1, \\ x_1 + x_2 - x_3 - x_4 = 2, \\ x_1 - x_2 + x_3 - x_4 = 1, \\ x_1 - x_2 - x_3 + x_4 = 1. \end{cases}$$

Solution

By performing elementary row operations on the augmented matrix:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 \end{array}\right) \xrightarrow{\text{elementary row operations}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{5}{4} \\ 0 & 1 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{4} \end{array}\right)$$

We get: $x_1 = \frac{5}{4}, x_2 = \frac{1}{4}, x_3 = -\frac{1}{4}, x_4 = -\frac{1}{4}$.
Thus:

$$\mathbf{a} = \frac{5}{4}\mathbf{a}_1 + \frac{1}{4}\mathbf{a}_2 - \frac{1}{4}\mathbf{a}_3 - \frac{1}{4}\mathbf{a}_4.$$

Exercise 4

Please proof that: If $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent, then $\mathbf{a}_1 + \mathbf{a}_2$ and $\mathbf{a}_1 - \mathbf{a}_2$ are also linearly independent.

Solution

Suppose $k_1(\mathbf{a}_1 + \mathbf{a}_2) + k_2(\mathbf{a}_1 - \mathbf{a}_2) = 0$, then:

$$(k_1 + k_2)\mathbf{a}_1 + (k_1 - k_2)\mathbf{a}_2 = 0.$$

Since $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent, we must have:

$$\begin{cases} k_1 + k_2 = 0, \\ k_1 - k_2 = 0. \end{cases}$$

The system of equations has the unique solution: $k_1 = 0, k_2 = 0$.
Thus, $\mathbf{a}_1 + \mathbf{a}_2$ and $\mathbf{a}_1 - \mathbf{a}_2$ are linearly independent.

Exercise 5

If $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ are linearly dependent, but any three of them are linearly independent, prove that there must exist a set of nonzero numbers k_1, k_2, k_3, k_4 such that:

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + k_3\mathbf{a}_3 + k_4\mathbf{a}_4 = 0.$$

Solution

By contradiction. Since $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ are linearly dependent, there exist nonzero numbers k_1, k_2, k_3, k_4 such that:

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + k_3\mathbf{a}_3 + k_4\mathbf{a}_4 = 0.$$

Assume that at least one of k_1, k_2, k_3, k_4 is zero. Without loss of generality, let $k_1 = 0$, then:

$$k_2\mathbf{a}_2 + k_3\mathbf{a}_3 + k_4\mathbf{a}_4 = 0.$$

where k_2, k_3, k_4 are nonzero, it follows that $\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ are linearly dependent, which contradicts the assumption that any three of the vectors are linearly independent.
Therefore, k_1, k_2, k_3, k_4 must all be nonzero.

Thank you!

MATH 1853: Introduction to Linear Algebra

Tutorial 3: Vector Spaces and Linear Spaces

Wenyong Zhou

HKU EEE

September 2024

Who am I

- Wenyong Zhou (Richard)
- Contact me: wenyongz@connect.hku.hk
- Office: Room 806, Chow Yei Ching Building

Important Notes

Before coming, please send me an email about:

- when will you come
- what questions you have

Note: The tutorial and lecture contents **cannot be in sync** due to the tight scheduling. You should treat tutorials as extra materials for Math 1853.

Outline

Vector Spaces

Let $B = \{\beta_1, \beta_2, \ldots, \beta_n\} \subseteq \mathbb{R}^n$. If B is linearly independent, then any vector $\alpha \in \mathbb{R}^n$ can be expressed as a linear combination of B, that is,

$$\alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n,$$

The ordered tuple $(a_1, a_2, \ldots, a_n)$ is called the coordinates of vector α with respect to the basis B (or in the basis B). We denote it as

$$\alpha_B = (a_1, a_2, \ldots, a_n), \quad \text{or} \quad \alpha_B = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

and call α_B the coordinate vector of α.

Vector Spaces

Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of $\mathbb{R}^n$, and let

$$\begin{aligned} \eta_1 &= a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n, \\ \eta_2 &= a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n, \\ &\vdots \\ \eta_n &= a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n, \end{aligned}$$

Then a necessary and sufficient condition for $\eta_1, \eta_2, \ldots, \eta_n$ to be linearly independent is that

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0.$$

Vector Spaces

Let $B_1 = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $B_2 = \{\eta_1, \eta_2, \dots, \eta_n\}$ be two bases of $\mathbb{R}^n$, and suppose the following relationship holds:

$$(\eta_1, \eta_2, \dots, \eta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

Then the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

is called the transition matrix from the old basis B_1 to the new basis B_2.

Vector Spaces

Let the coordinates of vector α under two bases $B_1 = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $B_2 = \{\eta_1, \eta_2, \ldots, \eta_n\}$ be

$$x = (x_1, x_2, \ldots, x_n)^\top \quad \text{and} \quad y = (y_1, y_2, \ldots, y_n)^\top.$$

If A is the transition matrix from basis B_1 to basis B_2, then

$$Ay = x \quad \text{or} \quad y = A^{-1}x.$$

Linear Spaces

- Vectors in addition and scalar multiplication follow certain arithmetic rules, such as:
  - **Commutative law of addition**: $x + y = y + x$,
  - **Associative law of addition**: $(x + y) + z = x + (y + z)$,
  - **Distributive law of scalar multiplication**: $(\lambda + \mu)x = \lambda x + \mu x$;
- Matrices, polynomials, continuous functions, etc., also satisfy exactly the same arithmetic rules.
- By studying the common properties of different objects in terms of linear operations, we arrive at the axiomatic definition of a vector space.

Linear Spaces

A vector space (also called a linear space) consists of four components: two sets V and F, and two operations — one called vector addition, and the other called scalar multiplication.

- V: A set of objects called vectors. For example, n-dimensional vectors, matrices.
- F: A field — typically the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$.
- **Vector addition** (denoted as $x + y$): An operation between two elements in set V. The operation must satisfy closure: $x + y \in V$, $\forall x, y \in V$.
- **Scalar multiplication** (denoted as λx): An operation between an element in set F and an element in set V. The operation must satisfy closure: $\lambda x \in V$, $\forall \lambda \in F, \forall x \in V$.

Linear Spaces

Let V be a non-empty set, and F be a field.

- A binary operation, called **vector addition**, is defined between elements of V. This means that for any two elements x and y in V, there exists a unique element $z \in V$ such that z is the sum of x and y, denoted as $z = x + y$.
- A binary operation, called **scalar multiplication**, is defined between elements of F and V. This means that for any scalar $\lambda \in F$ and any element $x \in V$, there exists a unique element $y \in V$ such that y is the scalar product of λ and x, denoted as $y = \lambda x$.

If the addition and scalar multiplication satisfy the following rules, then V is called a **vector space** over the field F.

Linear Spaces

Vector addition must satisfy the following four rules:

1. $x + y = y + x$ (Commutativity of addition)
2. $(x + y) + z = x + (y + z)$ (Associativity of addition)
3. There exists an element $\theta \in V$ such that for all $x \in V$,

$$x + \theta = x$$

   (The element θ with this property is called the zero element of V, and is denoted 0).
4. For every $x \in V$, there exists an element $y \in V$ such that

$$x + y = 0$$

   (The element y is called the additive inverse of x, and is denoted $-x$).

Linear Spaces

Scalar multiplication must satisfy the following two rules:

1. $1x = x$;
2. $\lambda(\mu x) = (\lambda\mu)x$.

Scalar multiplication and vector addition must satisfy the following two rules:

3. $(\lambda + \mu)x = \lambda x + \mu x$;
4. $\lambda(x + y) = \lambda x + \lambda y$.

In the above rules, λ and μ represent arbitrary elements from the field F, and x, y, and z represent arbitrary elements from the set V.

Linear Spaces

- $\mathbb{R}^n$: n-dimensional real vector space. Operations: (1) vector addition, (2) scalar multiplication with vectors.

$$V = \left\{ x \mid x = (\xi_1, \xi_2, \ldots, \xi_n)^T, \, \xi_i \in \mathbb{R}, \, i = 1, 2, \ldots, n \right\}, \quad F = \mathbb{R}.$$

- $\mathbb{C}^n$: n-dimensional complex vector space. Operations: (1) vector addition, (2) scalar multiplication with vectors.

$$V = \left\{ x \mid x = (\xi_1, \xi_2, \ldots, \xi_n)^T, \, \xi_i \in \mathbb{C}, \, i = 1, 2, \ldots, n \right\}, \quad F = \mathbb{C}.$$

Linear Spaces

- If $V = \left\{x \mid x = (\xi_1, \xi_2, \ldots, \xi_n)^T,\ \xi_i \in \mathbb{R},\ i = 1, 2, \ldots, n\right\}$ and $F = \mathbb{C}$, then V is **not** a vector space, because the scalar multiplication does not satisfy closure under these conditions.
- However, if $V = \left\{x \mid x = (\xi_1, \xi_2, \ldots, \xi_n)^T,\ \xi_i \in \mathbb{C},\ i = 1, 2, \ldots, n\right\}$ and $F = \mathbb{R}$, then V is a vector space, denoted as $\mathbb{C}^n_{\mathbb{R}}$.

Linear Spaces

Matrices of size $m \times n$ with elements from the field F, under:

- (1) matrix addition, and
- (2) scalar multiplication with matrices,

form a vector space over the field F, denoted as $F^{m \times n}$.

- $\mathbb{R}^{m \times n}$: $m \times n$ real matrix space.
- $\mathbb{C}^{m \times n}$: $m \times n$ complex matrix space.

Linear Spaces

The polynomial ring $F[x]$ over the field F, under:

- (1) the usual polynomial addition, and
- (2) scalar multiplication with polynomials,

forms a vector space over the field F.

If we only consider polynomials of degree less than n, and allow the addition of zero polynomials, this also forms a vector space over the field F, denoted as $F[x]_n$.

1. $\mathbb{R}[x]_n$: The space of polynomials with degree less than n over the real field $\mathbb{R}$,
$$\mathbb{R}[x]_n = \left\{p(x) \mid p(x) = \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0,\, \alpha_i \in \mathbb{R}\right\}.$$
2. $\mathbb{C}[x]_n$: The space of polynomials with degree less than n over the complex field $\mathbb{C}$,
$$\mathbb{C}[x]_n = \left\{p(x) \mid p(x) = \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0,\, \alpha_i \in \mathbb{C}\right\}.$$

Linear Spaces

The set of all real-valued functions, under:

- (1) function addition, and
- (2) scalar multiplication with functions,

forms a vector space over the real field.

Let $C[a, b]$ denote the set of all continuous functions on the interval $[a, b]$. Then $C[a, b]$, under:

- (1) function addition, and
- (2) scalar multiplication with functions,

forms a vector space over the real field.

A field P, under its own addition and multiplication, forms a vector space over itself.

Example 1

Let V be the set of positive real numbers, and $\mathbb{R}$ be the set of real numbers. Define addition $\oplus$ and scalar multiplication $\odot$ as follows:

$$\alpha \oplus \beta = \alpha\beta \quad \text{(i.e., the product of } \alpha \text{ and } \beta\text{)},$$

$$k \odot \alpha = \alpha^k \quad \text{(i.e., } \alpha \text{ raised to the power of } k\text{)},$$

where $\alpha, \beta \in V$ and $k \in \mathbb{R}$.

Question: Does V, under the operations $\oplus$ and $\odot$, form a vector space over $\mathbb{R}$?

Solution

For addition $\oplus$, closure holds: $\forall \alpha, \beta \in V = \mathbb{R}^+$, $\alpha \oplus \beta = \alpha\beta \in V$. Moreover, the following properties are satisfied:

1. **Commutativity**: $\alpha \oplus \beta = \alpha\beta = \beta\alpha = \beta \oplus \alpha$.
2. **Associativity**: $(\alpha \oplus \beta) \oplus \gamma = (\alpha\beta)\gamma = \alpha(\beta\gamma) = \alpha \oplus (\beta \oplus \gamma)$.
3. **Existence of a zero element**: The constant 1 is the zero element: $1 \oplus \alpha = 1\alpha = \alpha$.
4. **Existence of an additive inverse**: The inverse of any $\alpha \in V$ is $\frac{1}{\alpha}$: $\frac{1}{\alpha} \oplus \alpha = 1$.

Solution

For scalar multiplication $\odot$, closure holds: $\forall \alpha \in V = \mathbb{R}^+, k \in \mathbb{R}$, $k \odot \alpha = \alpha^k \in V = \mathbb{R}^+$. Moreover, the following properties are satisfied:

1. **Identity element of scalar multiplication**: $1 \odot \alpha = \alpha^1 = \alpha$.
2. **Compatibility of scalar multiplication**: $(k_1 k_2) \odot \alpha = \alpha^{k_1 k_2} = (\alpha^{k_2})^{k_1} = k_1 \odot (k_2 \odot \alpha)$.
3. **Distributive property over scalar addition**: $(k_1 + k_2) \odot \alpha = \alpha^{k_1 + k_2} = \alpha^{k_1} \alpha^{k_2} = k_1 \odot \alpha \oplus k_2 \odot \alpha$.
4. **Distributive property over vector addition**: $k \odot (\alpha \oplus \beta) = k \odot (\alpha\beta) = (\alpha\beta)^k = \alpha^k \beta^k = (k \odot \alpha) \oplus (k \odot \beta)$.

Therefore, V forms a vector space over $\mathbb{R}$ under the operations $\oplus$ and $\odot$.

Example 2

Given a basis $B_2 = \{\beta_1, \beta_2, \beta_3\}$ of $\mathbb{R}^3$, where

$$\beta_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \beta_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

find the transition matrix A from the standard basis $B_1 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to the basis B_2.

Solution

$$(\beta_1, \beta_2, \beta_3) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix},$$

we obtain the transition matrix:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

Exercise 3

Given two bases of $\mathbb{R}^3$, $B_1 = \{\alpha_1, \alpha_2, \alpha_3\}$ and $B_2 = \{\beta_1, \beta_2, \beta_3\}$, where:

$$\alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$\beta_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \beta_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

(i) Find the transition matrix A from the basis B_1 to the basis B_2.

(ii) Given that the coordinates of α in the basis B_1 are $(1, -2, -1)^\top$, find the coordinates of α in the basis B_2.

Solution

(i) Let

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)A,$$

which gives the equation:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} A.$$

Thus,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}.$$

Solution

By performing row operations:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & -1 & 0 \end{array}\right) \xrightarrow{r_3 - r_2,\, r_2 - r_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 & -2 & -2 \end{array}\right),$$

we obtain the transition matrix:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & -2 & -2 \end{pmatrix}.$$

Solution

(ii) Let the coordinates to be found be $(y_1, y_2, y_3)^\top$. Then:

$$\alpha = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} = (\beta_1, \beta_2, \beta_3)\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Substituting $(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)A$, we get:

$$(\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} = (\alpha_1, \alpha_2, \alpha_3)A\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Thus:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = A^{-1}\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ -4 \end{pmatrix}.$$

Thank you!

MATH 1853: Introduction to Linear Algebra

Tutorial 4: Determinant

Wenyong Zhou

HKU EEE

Sep 2024

Who am I

- Wenyong Zhou (Richard)
- Contact me: wenyongz@connect.hku.hk
- Office: Room 806, Chow Yei Ching Building

Important Notes

Before coming, please send me an email about:

- when will you come
- what questions you have

Note: The tutorial and lecture contents **cannot be in sync** due to the tight scheduling. You should treat tutorials as extra materials for Math 1853.

Outline

Why Discuss Determinants?

For the two-variable linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases}$$

it is easy to solve using the elimination method: When $a_{11}a_{22} - a_{12}a_{21} \neq 0$, the system has a unique solution:

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11} b_2 - a_{12} b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

We introduce a notation

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \triangleq ad - bc, \tag{8}$$

and this notation is called a 2-by-2 determinant.

Why Discuss Determinants?

Thus, the solution to the linear system can be summarized as follows: When the 2-by-2 determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$$

is non-zero, the system of equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases}$$

has a unique solution:

$$x_1 = \frac{\begin{vmatrix} \mathbf{b_1} & a_{12} \\ \mathbf{b_2} & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & \mathbf{b_1} \\ a_{21} & \mathbf{b_2} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Cramer's rule

Let's consider Cramer's rule: Given the linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{1}$$

If the determinant of the coefficient matrix

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \neq 0, \tag{2}$$

then the linear system (1) has a solution, and the solution is unique.

Cramer's rule

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \cdots, \quad x_n = \frac{D_n}{D}, \tag{3}$$

where D_j is the determinant formed by replacing the j-th column of D with the constant terms $b_1, b_2, \cdots, b_n$. The determinant D_j is given by:

$$D_j = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & \mathbf{b_1} & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & \mathbf{b_2} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & \mathbf{b_n} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}. \tag{4}$$

The reason for the appearance of determinant formulas can be understood as follows: They perfectly express the regularity of a portion of the solution to the linear system.

- Using Cramer's Rules solve

$$\begin{cases} x_1 - 2x_2 = 4 \\ -x_1 + x_2 = 2 \end{cases}$$

$$A = \begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

- Using Cramer's Rules solve

$$\begin{cases} x_1 - 2x_2 = 4 \\ -x_1 + x_2 = 2 \end{cases}$$

$$A = \begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$$x_1 = \frac{\det(A_1(b))}{\det(A)} = \frac{\det \begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix}}{1 \times 1 - (-2 \times -1)} = \frac{4+4}{-1} = -8$$

$$x_2 = \frac{\det(A_2(b))}{\det(A)} = \frac{\det \begin{bmatrix} 1 & 4 \\ -1 & 2 \end{bmatrix}}{1 \times 1 - (-2 \times -1)} = \frac{2+4}{-1} = -6$$

Definition of an $n \times n$ Determinant

Given an $n \times n$ array of numbers, the n-order determinant D is defined as:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

This number is expressed as follows:

1. When $n = 1$, $D = \begin{vmatrix} a_{11} \end{vmatrix} = a_{11}$;
2. When $n \geq 2$,

$$D = (-1)^{1+1}a_{11}M_{11} + (-1)^{1+2}a_{12}M_{12} + \cdots + (-1)^{1+n}a_{1n}M_{1n}, \tag{16}$$

where M_{1j} (for $j = 1, 2, \ldots, n$) is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the first row and the j-th column from D.

Diagonal Matrix

$$\begin{vmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{vmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n.$$

$$\begin{vmatrix} & & & \lambda_1 \\ & & \lambda_2 & \\ & \cdots & & \\ \lambda_n & & & \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \lambda_1 \lambda_2 \cdots \lambda_n.$$

Triangular Matrix

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{n\geq i>j\geq 1} (x_i - x_j).$$

Example

- Using Cramer's Rules solve

$$\begin{cases} 2x_1 - 4x_2 = 9 \\ 3x_1 + 6x_2 = 7 \end{cases}$$

Example

- Using Cramer's Rules solve

$$\begin{cases} 2x_1 - 4x_2 = 9 \\ 3x_1 + 6x_2 = 7 \end{cases}$$

$$A = \begin{bmatrix} 2 & -4 \\ 3 & 6 \end{bmatrix}, \quad b = \begin{bmatrix} 9 \\ 7 \end{bmatrix}$$

$$x_1 = \frac{\det(A_1(b))}{\det(A)} = \frac{\det \begin{bmatrix} 9 & -4 \\ 7 & 6 \end{bmatrix}}{2 \times 6 - (-4 \times 3)} = \frac{41}{12}$$

$$x_2 = \frac{\det(A_2(b))}{\det(A)} = \frac{\det \begin{bmatrix} 2 & 9 \\ 3 & 7 \end{bmatrix}}{2 \times 6 - (-4 \times 3)} = \frac{-13}{24}$$

Example

Calculate the determinant:

$$D = \begin{vmatrix} 0 & -1 & -1 & 2 \\ 1 & -1 & 0 & 2 \\ -1 & 2 & -1 & 0 \\ 2 & 1 & 1 & 0 \end{vmatrix}.$$

Solution

$$D = (-1)^{1+1} \cdot 0 \cdot \begin{vmatrix} -1 & 0 & 2 \\ 2 & -1 & 0 \\ 1 & 1 & 0 \end{vmatrix} + (-1)^{1+2} \cdot (-1) \cdot \begin{vmatrix} 1 & 0 & 2 \\ -1 & -1 & 0 \\ 2 & 1 & 0 \end{vmatrix} +$$

$$(-1)^{1+3} \cdot (-1) \cdot \begin{vmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \\ 2 & 1 & 0 \end{vmatrix} + (-1)^{1+4} \cdot 2 \cdot \begin{vmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 2 & 1 & 1 \end{vmatrix}$$

$$= 0 + 2 + 10 - 8 = 4.$$

Example

Calculate the determinant D of the 4x4 matrix:

$$D = \begin{vmatrix} 1 & 1 & -1 & 2 \\ -1 & -1 & -4 & 1 \\ 2 & 4 & -6 & 1 \\ 1 & 2 & 4 & 2 \end{vmatrix}.$$

Solution

$$\begin{vmatrix} 1 & 1 & -1 & 2 \\ -1 & -1 & -4 & 1 \\ 2 & 4 & -6 & 1 \\ 1 & 2 & 4 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 & 2 \\ 0 & 0 & -5 & 3 \\ 0 & 0 & -14 & -3 \\ 1 & 2 & 4 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 & 2 \\ 0 & 0 & -5 & 3 \\ 0 & 0 & -14 & -3 \\ 0 & 1 & 5 & 0 \end{vmatrix}$$

$$= (-1)^3 \begin{vmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & -14 & -3 \\ 0 & 0 & -5 & 3 \end{vmatrix} = - \begin{vmatrix} -14 & -3 \\ -5 & 3 \end{vmatrix} = 57.$$

Example

Calculate the determinant D_n of the following $n \times n$ matrix:

$$D_n = \begin{vmatrix} x & a & a & \cdots & a \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \cdots & x \end{vmatrix}.$$

Solution

Step 1: Multiply the first row by (-1) and add it successively to the remaining rows to obtain:

$$D_n = \begin{vmatrix} x & a & a & \cdots & a \\ a-x & x-a & 0 & \cdots & 0 \\ a-x & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a-x & 0 & 0 & \cdots & x-a \end{vmatrix},$$

Solution

Next Step: Add all columns to the first column to obtain:

$$D_n = \begin{vmatrix} x+(n-1)a & a & a & \cdots & a \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x-a \end{vmatrix},$$

Solution

Last Step: It is a triangular matrix:

$$D_n = \begin{vmatrix} x+(n-1)a & a & a & \cdots & a \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x-a \end{vmatrix}$$

$$= [x+(n-1)a]\,(x-a)^{n-1}.$$

Thank you!

MATH 1853: Introduction to Linear Algebra

Tutorial 5: Eigenvalue and Eigenvector

Wenyong Zhou

HKU EEE

October 2022

Who am I

- Wenyong Zhou (Richard)
- Contact me: wenyongz@connect.hku.hk
- Office: Room 806, Chow Yei Ching Building

Important Notes

Before coming, please send me an email about:

- when will you come
- what questions you have

Note: The tutorial and lecture contents **cannot be in sync** due to the tight scheduling. You should treat tutorials as extra materials for Math 1853.

Outline

What is Eigenvalue

Let $\mathbf{A}$ be an $n \times n$ matrix. If there exists a scalar λ and a nonzero n-dimensional vector $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}, \tag{1}$$

then:

1. λ is called an **eigenvalue** of matrix $\mathbf{A}$;
2. $\mathbf{x}$ is called an **eigenvector** corresponding to the eigenvalue λ of matrix $\mathbf{A}$.

Finding Eigenvalues and Eigenvectors

Problem:

Given an $n \times n$ matrix $\mathbf{A}$, find the λ and $\mathbf{x}$ that satisfy the equation $\mathbf{Ax} = \lambda\mathbf{x}$ where $\mathbf{x} \neq 0$.

Solution Process:

1. Start with the equation $\mathbf{Ax} = \lambda\mathbf{x}$.
2. Rearrange to form $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0$.
3. Solve the determinant equation $|\lambda\mathbf{I} - \mathbf{A}| = 0$ to find the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.
4. Substitute each λ back into $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = 0$ to find the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots$.

Note: The equation $|\lambda\mathbf{I} - \mathbf{A}| = 0$ is called the **characteristic equation**.

Finding Eigenvalues and Eigenvectors

$$|\lambda \mathbf{I} - \mathbf{A}| = 0, \tag{2}$$

that is,

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{31} & -a_{32} & \lambda - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & -a_{n3} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0, \tag{3}$$

1. $|\lambda \mathbf{I} - \mathbf{A}|$ is called the **characteristic polynomial**.
2. $\lambda \mathbf{I} - \mathbf{A}$ is called the **characteristic matrix**.
3. A matrix of order n has n eigenvalues (including repeated roots) in the complex domain.

Eigen Subspace

Among all the eigenvectors corresponding to the eigenvalue λ_0, let $\xi_1, \xi_2, \dots, \xi_l$ be a maximal linearly independent set. The space spanned by this maximal linearly independent set is called the **eigen subspace** of matrix $\mathbf{A}$ corresponding to the eigenvalue λ_0, denoted as V_{λ_0}. The dimension of this eigen subspace is called the **geometric multiplicity** of the eigenvalue λ_0.

Note:

- The eigen subspace V_{λ_0} is the solution space of the equation $(\lambda_0 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$.
- "All eigenvectors corresponding to λ_0" + "Zero vector" = "Eigen subspace V_{λ_0}".

Equivalent Definitions of Geometric Multiplicity

The geometric multiplicity l of λ_0 is equivalently defined as:

1. l is the number of linearly independent solutions included in the fundamental solution system of the equation $(\lambda_0 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$, i.e., $l = n - \mathrm{r}(\lambda_0 \mathbf{I} - \mathbf{A})$.
2. l is the dimension of the eigen subspace corresponding to λ_0, i.e., $l = \dim V_{\lambda_0}$.

Therefore,

$$l = \dim V_{\lambda_0} = n - \mathrm{r}(\lambda_0 \mathbf{I} - \mathbf{A}).$$

Let λ_0 be an eigenvalue of the $n \times n$ matrix $\mathbf{A}$ with algebraic multiplicity k. Then k is called the **algebraic multiplicity** of the eigenvalue λ_0.

Property of Eigenvalue and Eigenvector

If $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue λ_0, then $k_1\mathbf{x}_1 + k_2\mathbf{x}_2$ is also an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_0 (where k_1 and k_2 are arbitrary constants, but $k_1\mathbf{x}_1 + k_2\mathbf{x}_2 \neq 0$).

Proof:
Since $\mathbf{x}_1$ and $\mathbf{x}_2$ are solutions of the homogeneous linear system

$$(\lambda_0\mathbf{I} - \mathbf{A})\mathbf{x} = 0,$$

it follows that $k_1\mathbf{x}_1 + k_2\mathbf{x}_2$ is also a solution of the above equation. Hence, when $k_1\mathbf{x}_1 + k_2\mathbf{x}_2 \neq 0$, it is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_0.

Property of Eigenvalue and Eigenvector

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be all the eigenvalues of an $n \times n$ matrix $\mathbf{A}$. The following important properties hold:

- $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$;
- $\lambda_1 \lambda_2 \cdots \lambda_n = |\mathbf{A}|$.

Here, $a_{11} + a_{22} + \cdots + a_{nn}$ is called the trace of the matrix and is denoted as $\text{tr}(\mathbf{A})$.

Property of Eigenvalue and Eigenvector

If λ is an eigenvalue of matrix $\mathbf{A}$, and $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ corresponding to λ, then:

(i) $k\lambda$ is an eigenvalue of $k\mathbf{A}$ (where k is any constant);

(ii) λ^m is an eigenvalue of $\mathbf{A}^m$ (where m is a positive integer);

(iii) If $\mathbf{A}$ is invertible, λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$;

(iv) If $\mathbf{A}$ is invertible, $\frac{|\mathbf{A}|}{\lambda}$ is an eigenvalue of $\mathbf{A}^*$ (where $\mathbf{A}^*$ denotes the adjugate of $\mathbf{A}$).

Property of Eigenvalue and Eigenvector

Let λ be an eigenvalue of the square matrix $\mathbf{A}$. Then

$$a_0 + a_1\lambda + \cdots + a_m\lambda^m$$

is an eigenvalue of

$$a_0\mathbf{I} + a_1\mathbf{A} + \cdots + a_m\mathbf{A}^m.$$

Property of Eigenvalue and Eigenvector

The eigenvalues of matrix $\mathbf{A}$ and matrix $\mathbf{A}^T$ are identical.

Proof:
Since

$$(\lambda\mathbf{I}-\mathbf{A})^T=(\lambda\mathbf{I})^T-\mathbf{A}^T=\lambda\mathbf{I}-\mathbf{A}^T,$$

we have

$$|\lambda\mathbf{I}-\mathbf{A}|=|\lambda\mathbf{I}-\mathbf{A}^T|,$$

hence, the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^T$ are identical.

Example

Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution

Step 1 Solve the characteristic equation $|\lambda \mathbf{I} - \mathbf{A}| = 0$.

$$|\lambda \mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2 - 4 = (\lambda - 3)(\lambda + 1),$$

By solving $|\lambda \mathbf{I} - \mathbf{A}| = 0$, the eigenvalues are obtained as

$$\lambda_1 = 3, \quad \lambda_2 = -1.$$

Solution

Step 2 Find all non-zero solutions to the equation $(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{x} = 0$.

- For $\lambda_1 = 3$, solve the linear system $(\lambda_1 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$. We have

$$3\mathbf{I} - \mathbf{A} = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix},$$

The fundamental solution is

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

Thus, all eigenvectors corresponding to $\lambda_1 = 3$ are

$$k_1 \mathbf{p}_1 = k_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad k_1 \neq 0.$$

Solution

- For $\lambda_2 = -1$, solve the linear system $(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{x} = 0$. We have

$$(-1)\mathbf{I} - \mathbf{A} = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

The fundamental solution is

$$\mathbf{p}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

Thus, all eigenvectors corresponding to $\lambda_2 = -1$ are

$$k_2\mathbf{p}_2 = k_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad k_2 \neq 0.$$

Examples

Given

$$\mathbf{A} = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}.$$

1. Find the eigenvalues and eigenvectors of $\mathbf{A}$;
2. Find an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a diagonal matrix.

Solution

By calculating the determinant,

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} 2-\lambda & -2 & 0 \\ -2 & 1-\lambda & -2 \\ 0 & -2 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda-4)(\lambda+2),$$

we find that the eigenvalues of matrix $\mathbf{A}$ are $\lambda_1 = -2$, $\lambda_2 = 1$, and $\lambda_3 = 4$.

Solution

When $\lambda_1 = -2$, from $(\lambda_1 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$, we have

$$\begin{pmatrix} 4 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

The fundamental solution is

$$\mathbf{x}_1 = (1, 2, 2)^T,$$

Thus, all eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue $\lambda_1 = -2$ are

$$k_1(1, 2, 2)^T,$$

where k_1 is any non-zero constant.

Solution

When $\lambda_2 = 1$, from $(\lambda_2 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$, we have

$$\begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

The fundamental solution is

$$\mathbf{x}_2 = (2, 1, -2)^T,$$

Thus, all eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue $\lambda_2 = 1$ are

$$k_2(2, 1, -2)^T,$$

where k_2 is any non-zero constant.

Solution

When $\lambda_3 = 4$, from $(\lambda_3 \mathbf{I} - \mathbf{A})\mathbf{x} = 0$, we have

$$\begin{pmatrix} -2 & -2 & 0 \\ -2 & -3 & -2 \\ 0 & -2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

The fundamental solution is

$$\mathbf{x}_3 = (2, -2, 1)^T,$$

Thus, all eigenvectors of $\mathbf{A}$ corresponding to the eigenvalue $\lambda_3 = 4$ are

$$k_3(2, -2, 1)^T,$$

where k_3 is any non-zero constant.

Solution

Given

$$\mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (\mathbf{A}\mathbf{x}_1, \mathbf{A}\mathbf{x}_2, \mathbf{A}\mathbf{x}_3) = (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \lambda_3\mathbf{x}_3)$$

$$= (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

Let

$$\mathbf{P} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

then

$$\mathbf{A}\mathbf{P} = \mathbf{P}\Lambda.$$

Solution

Since $|\mathbf{P}| = -27 \neq 0$, the matrix $\mathbf{P}$ is invertible, where

$$\mathbf{P} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix},$$

Thus,

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Thank you!