School of Mathematical Science
Jiangsu University

**Lecture Notes for the course**

# Advanced Algebra

**Version # 2**

Peter Tan

August 2023

School of Mathematical Science
Jiangsu University

**Important Questions in Linear Algebra**

1. Solve system of linear equations.
   **Key Words:** three elementary row operations, row reduced echelon form, solution set.

2. Compute determinant of a square matrix.
   **Key Words:** expansion, elementary row operations, adjoint matrix.

3. Basis of an abstract vector space.
   **Key Words:** linear combination, span, linear independent.

4. Inner Product space.
   **Key Words:** Inner product, Orthogonal, Gram-Schimidt Process, Projection.

5. Linear Transformation.
   **Key Words:** Kernel and range, Standard matrix of a linear transformation, matrix of a linear transformation.

6. Eigen-theory of a Linear transformation.
   **Key Words:** Eigenvalue and eigenvectors of a matrix, Diagonalization, Orthogonally diagonalization.

学而不思则罔, 思而不学则殆.
Learning without thinking is ambiguous;
thinking without learning is exhausted.

– Confucious

# Preface

江苏大学与美国阿卡迪亚大学的中外合作办学项目办的非常成功, 许多学生被世界百强名校录取为研究生。本专业的目标是培养高素质复合型国际化创新人才. 毫无疑问, 数学专业是培养高素质人才的抓手. 关于复合型和国际化, 我校学生可在美方大学选修数学、精算、数据科学和计算机科学等四个本科专业。

高等代数的教学可为人才培养贡献一份力量。首先, 我们教学内容为培养目标服务，对高等代数课程的教学内容进行了如下修改。首先，作为专业基础课，深挖数学思想和思维，强调高等代数是代数和几何的统一。在教学中强调教学内容的几何意义，帮助学生更好的理解教学内容。其次，加强计算机软件在教学中的应用。将学生从大量的计算中解放出来，引领学生用所得数据中发现规律，提高解决问题的能力, **加强对数学概念的理解**。最后，将编程引入教学过程，提高学生利用编程解决数学问题的能力。在这个人工智能飞速发展的年代，高等代数已经成为它的最重要的数学基础之一，在课上和课后发掘高等代数知识在人工智能里的应用是一个极好的教学方法，对培养高层次人才有帮助。

线性代数是代数和几何的统一, 希望在高等代数的学习中多多体会. 在学年结束时，你问你自己是否掌握了线性代数的核心要义，该问题等价于: 我是否学会了将问题转化为矩阵问题?

Yilan Tan at UJS

# Contents

# List of Figures



# List of Tables

# Part 1: First Semester

# System of Linear Equations 1

Variable[1] is an expression, usually denoted by a letter, that is defined for values within a given set. Variable can be used to represent elements of sets which are not numbers but frequently it relates to numerical quantities and functions defined in them together with the relationship between them.

Solving system of linear equation is an important topic in linear algebra. It serves as a tool of this course. In most case we use Gauss-Jordan elimination to solve system of linear equations. The general solutions give the idea of vector.



1: 中文意思: 变量.
练习: 复述第一段话.

## 1.1 System of linear equations

The following definitions are important.

- **Linear equation.**
  A linear equation in the variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

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

  where the **constant** $b$ and the **coefficients**, $a_1, a_2, \ldots, a_n$ are real numbers.

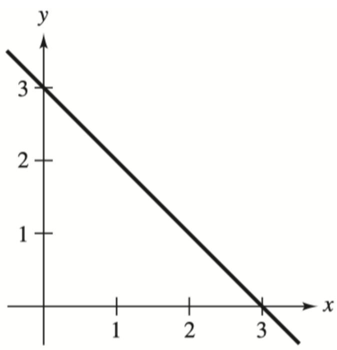



Figure 1.1: a linear equation $x + y = 3$.

- System of linear equations. A system of linear equations is a **collection** of one or more linear equations involving the same variables, $x_1, x_2, \ldots, x_n$.

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

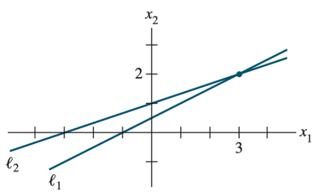



Figure 1.2: System of linear equations

- Solution set.
  A solution of the system is a list $(s_1, s_2, \ldots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, \ldots, s_n$ are substituted for $x_1, \ldots, x_n$, respectively. The set of

all possible solutions is called the solution set of the linear system.

- Consistent and inconsistent.
  A system of linear equations is said to be consistent if it has either one solution or infinitely many solutions; a system is inconsistent if it has no solution.

Hint: Think lines in the plane or planes in space!

Discussion: (1) Give two examples of linear systems which are consistent and inconsistent, respectively. For the consistent system, find the solution set.

(2) Give an example of a linear system which has infinitely many solutions.

## Motivations

**Example 1.1.1** Which linear system is easy to solve?

$$\begin{cases} x + \phantom{2}y = 25 \\ \phantom{x+} 2y = 30 \end{cases}, \qquad \begin{cases} \phantom{2}x + \phantom{4}y = 25 \\ 2x + 4y = 80 \end{cases}$$

The linear system on the left is called in row-echelon form, which means that it has a “stair-step”pattern with nonzero leading coefficients. For the first linear system, using back-substitution, we easily obtain the solution $y = 15$ and then $x = 10$.

**Definition 1.1.1** Two linear system are called equivalent when they have the same solution set.

It is a common sense that one can transform a linear system into an equivalent linear system using the following three operations.

**Operations which produce equivalent linear systems**

Each of these operations on a system of linear equations produces an equivalent system.

1. Add a multiple of an equation to another equation.
2. Multiply an equation by a nonzero constant.
3. Interchange two equations.

In what next, we show the second linear system in Example 1.1.1 is equivalent to the first one.

$$\begin{cases} x + y = 25 & (1) \\ 2x + 4y = 80 & (2). \end{cases} \tag{1.1}$$

Equation (2) subtract 2 times of Equation (1) to obtain Equation (3), that is,

$$\begin{cases} x + y = 25 & (1) \\ 2y = 30 & (3). \end{cases} \tag{1.2}$$

In this course, we go further more. $\frac{1}{2}$ times of Equation (3) yields that

$$\begin{cases} x + y = 25 & (1) \\ y = 15 & (4). \end{cases} \tag{1.3}$$

Then it follows from Equation (1) substracting Equation (4) that

$$\begin{cases} x = 10 & (5) \\ y = 15 & (4). \end{cases} \tag{1.4}$$

There is a great mathematical idea behind this example, that is, to solve a linear system, one find an equivalent linear system which is much easier to solve. The algorithm to transform a system of linear equations into a unique equivalence system is called the Gauss-Jordan elimination[2].

**Discussion:** the idea.

**This is the most important idea in this chapter.**

2: 高斯消元法

**Example 1.1.2** Using the ideas above, find the solution set of the linear system

$$\begin{cases} 3x_2 - 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases}$$

Try this Exercise.

## 1.2 Gauss-Jordan Elimination

In what next, we introduce a notion to simplify the expression of Example 1.1.2. The essential information of a linear system can be

3: Matrix is a way to record information!

recorded compactly in a rectangular array called a matrix.[3] Given the system

$$\begin{cases} \quad\quad 3x_2 - \ \ 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + \ \ 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases}$$

Try this Exercise.

The matrices

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 3 & -9 & 12 & -9 & 6 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix}$$

Critical Thinking:
Linear system $\Longleftrightarrow$ Matrix.

are called the coefficient matrix and augmented matrix of the linear system, respectively.

4: **数学思维：类比.** 三类矩阵的初等行变换无非是三类线性方程组等价变换的另一种表述。

Recall that, for a linear system, there are three operations to produce an equivalent linear system. [4] Analogue for the three operations, we also have three row operations.

Discussion: terminologies:

1. row
2. column
3. entry

**Three Operations for matrix**

1. Add a multiple of a row to another row.
2. Multiply a row by a nonzero constant.
3. Interchange two rows.

**Exercise 1.2.1** Answer the following three questions below.

Exercise A. Exercise B. Exercise C.

Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other. Row operations are reversible. Therefore, if the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.[5]

5: Do you understand this paragraph? Circle your answer below:
Yes. No.

A nonzero row or column in a matrix means a row or column that contains at least one nonzero entry; a leading entry of a row refers to the leftmost nonzero entry in a nonzero row.

**Definition 1.2.1** A rectangular matrix is in row echelon form (REF) if it has the following three properties:

1. All nonzero rows are above any row of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zero.
   If a matrix in row echelon form satisfies the following additional conditions, then it is in reduced row echelon form (RREF)
4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

Example of matrices who has REF and RREF, respectively:

$$\begin{pmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Example of matrices who are not in REF.

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

**Exercise:** Turn the matrices above into RREF.

Answer:

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

**Definition 1.2.2** A pivot position in a matrix $A$ is a location in $A$ that corresponds to a leading 1 in the reduced echelon form of $A$. A pivot column is a column of $A$ that contains a pivot position.

Discussion: How to decide the number of the pivot columns of a matrix $A$?

**Example 1.2.1** Supposed that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -3 & 3 & 0 \\ 0 & 1 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The pivot positions are $(1,1)$(row, column), $(2,2)$ and $(3,5)$; The pivot columns are columns $1$, $2$ and $5$.

**Example 1.2.2** Find the RREF of $\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix}$

**Solution:**

**Step 1** Begin with the leftmost nonzero column, select a nonzero entry in that column as a pivot, and then move the row to the first row.

Interchange Row 1 and Row 3.

$$\begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix}$$

**Step 2** Use the row operations to create zeros in all positions below the pivot.

Adding −1 times Row 1 to Row 2.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix} \xrightarrow{-R_1+R_2} \begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix}$$

**Step 3** Cover the row containing the pivot position and cover all rows above it. Apply steps 1 and 2 to submatrix that remains. Repeat the process until there are no more nonzero row to modify.

Adding $-\frac{3}{2}$ time the row 2 to row 3.

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{pmatrix} \xrightarrow{-\frac{3}{2}\cdot R_2+R_3} \begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

**Step 4** Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation. Adding -2 times row 3 to row 2,

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{-2R_3+R_2} \begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}.$$

- −6 times row 3 to Row 1,

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{-6R_3+R_1} \begin{pmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}.$$

- Multiply $\frac{1}{2}$ to Row 2,

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{pmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}.$$

- 9 times Row 2 to Row 1,

$$\begin{pmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{9R_2+R_1} \begin{pmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}.$$

- Multiply $\frac{1}{3}$ to Row 1,

$$\begin{pmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{\frac{1}{3}R_1} \begin{pmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix},$$

which is the reduced row echelon form of the original matrix.

Exercise: find the reduced echelon form of the matrices $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 2 & -2 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & 2 & 2 & 2 \end{pmatrix}$, respectively.

Answer: $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

**Chanllenge question:** Write a code to perform the Gauss-Jordan elimination automatically.

**Theorem 1.2.1** Each matrix is row equivalent to one and only one reduced echelon matrix.

- If a matrix $A$ is row equivalent to an echelon matrix $U$, we call $U$ an echelon form (or row echelon form) of $A$;
- if $U$ is in reduced row echelon form, we call $U$ the reduced (row) echelon form of $A$.

True or False:

1. An echelon form of a matrix $A$ is unique.

2. One can get an echelon form of a matrix $A$ by only replacement and interchange.

The following exercise is also important. You may write the answer immediately by otained result above.

**Exercise 1.2.2** Find the reduce echelon form of

$$\begin{pmatrix} 0 & 3 & -6 & 4 \\ 3 & -7 & 8 & 8 \\ 3 & -9 & 12 & 6 \end{pmatrix}.$$

Answer: $\begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

The next exercise looks similar to the above one. However, it is totally different.

**Exercise 1.2.3** Find the reduce echelon form of

$$\begin{pmatrix} 0 & -6 & 6 & 4 \\ 3 & 8 & -5 & 8 \\ 3 & 12 & -9 & 6 \end{pmatrix}$$

Answer: $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

头脑风暴: 为什么练习1.2.2的结果可从 Example 1.2.2得到, 但练习1.2.3的结果不可以?

## 1.3 Solution set

Try this Exercise.

Recall that an augmented matrix records the essential information of a linear system. Here is a question: what we can obtain from the reduced echelon form of the augmented matrix of a linear system?

The reduced echelon form tells many information regarding the linear system.

**Theorem 1.3.1** (Consistent) A linear system is consistent if and only if the rightmost column of the augmented matrix is not a pivot column - that is, if and only if an echelon form of the augmented matrix has no row of the form

$$\begin{pmatrix} 0 & \cdots & 0 & b \end{pmatrix} \text{ with } b \text{ nonzero}$$

Form now on in this section, we suppose that the linear system is consistent. We define two important concepts below.

- Basic Variables.
  The variables corresponding to pivot columns in the augmented matrix are called basic(leading) variables.

Fill the blank:
1. The number of basic variables equals to the number of ____________, or ____________.

2. The number of free variables equals to ________________.

Feel free to add correct answers in this list.

▶ Free Variables.
The variables which are not basic variables are called free variables.

The solution set of a linear system can be described explicitly by solving the reduced system for the basic variables in terms of the free variables.

**Example 1.3.1** Solve the linear system

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

**Solution.** We divides the solution into steps.

**Step 1** The augmented coefficient matrix of the system of linear equations is

$$A = \begin{pmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{pmatrix}$$

**Step 2** The row reduced echelon form

$$\text{RREF(A)}=\begin{pmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

**Step 3** How to interpret the RREF(A)?

$$\begin{aligned} x_1 \quad - 2x_3 + 3x_4 &= -24 \\ x_2 - 2x_3 + 2x_4 &= -7 \\ x_5 &= 4 \end{aligned}$$

**Step 4** The variables $x_1$, $x_2$ and $x_5$ are the basic (leading) variables. The other variables, $x_3$ and $x_4$, are the free variable.

6: It is a convention(惯例).

**Step 5** express the basic variables by the free variables [6]

$$\begin{aligned} x_1 &= -24 + 2x_3 - 3x_4 \\ x_2 &= -7 + 2x_3 - 2x_4 \\ x_5 &= 4 \end{aligned}$$

**Step 6** The General solution is

$$\begin{cases} x_1 = -24 + 2x_3 - 3x_4 \\ x_2 = -7 + 2x_3 - 2x_4 \\ x_3 \text{ is free} \\ x_4 \text{ is free} \\ x_5 = 4 \end{cases} \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -24 + 3x_3 - 3x_4 \\ -7 + 2x_3 - 2x_4 \\ x_3 \\ x_4 \\ 4 \end{pmatrix},$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -24 + 2x_3 - 3x_4 \\ -7 + 2x_3 - 2x_4 \\ x_3 \\ x_4 \\ 4 \end{pmatrix} = \begin{pmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, x_3, x_4 \in \mathbb{R}$$

Repeat the answer below: how to express the general solution?

**Step 7** The statement "$x_3$ is free" means that you are free to choose any value for $x_3$; The same is for $x_4$. For example, let $x_3 = 1$ and $x_4 = -2$, then

$$\begin{pmatrix} -13 \\ -1 \\ 1 \\ -2 \\ 4 \end{pmatrix} \text{ is a particular solution.}$$

The next exercise is a very tough one, but trying it well definitely does benefit your computational ability.

Try this Exercise.

The rref of the augmented matrix is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{2}a + \frac{1}{2}b - \frac{1}{2}c - \frac{1}{2}d \\ 0 & 1 & 0 & 0 & \frac{1}{2}a - \frac{1}{2}b + \frac{1}{2}c + \frac{1}{2}d \\ 0 & 0 & 1 & 0 & \frac{1}{2}b - \frac{1}{4}d \\ 0 & 0 & 0 & 1 & -\frac{1}{2}a - \frac{1}{2}c + \frac{1}{4}d \end{pmatrix}.$$

**Exercise 1.3.1** Let $a, b, c, d \in \mathbb{R}$. Solve the linear system

$$\begin{cases} x_1 + x_2 = a \\ x_2 + 3x_3 + x_4 = b \\ -x_1 + x_3 - x_4 = c \\ 2x_2 - 2x_3 + 2x_4 = d \end{cases}$$

The next example is different with the examples above!

**Example 1.3.2** Solve the system

$$\begin{cases} x_1 - x_2 + 2x_3 = 4 \\ x_1 + x_3 = 6 \\ 2x_1 - 3x_2 + 5x_3 = 4 \\ 3x_1 + 2x_2 - x_3 = 1 \end{cases}$$

Answer: The augmented matrix for this system is $\begin{pmatrix} 1 & -1 & 2 & 4 \\ 1 & 0 & 1 & 6 \\ 2 & -3 & 5 & 4 \\ 3 & 2 & -1 & 1 \end{pmatrix}$, and the RREF of this matrix is $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Since there is a row $[0 \quad 0 \quad 0 \quad 1]$ in the reduced echelon form of the augmented matrix, the system is inconsistent.

Scan for the solution!

**Exercise 1.3.2** Find a relation of $a, b, c, d \in \mathbb{R}$ such that the linear system below is consistent.

$$\begin{cases} x_1 + x_2 = a \\ x_2 + 3x_3 + x_4 = b \\ -x_1 + x_3 - x_4 = c \\ 2x_2 + 2x_3 - 2x_4 = d \end{cases}$$

Discussion: Suppose that the linear system $A\mathbf{x} = \mathbf{b}$ is always consistent for any $\mathbf{b} \in \mathbb{R}^n$. What we can say about the reduced echelon form of a the coefficient matrix?

**Theorem 1.3.2** Suppose that $A$ is an $n \times n$ matrix and the linear system $A\mathbf{x} = \mathbf{b}$ is always consistent. Then

$$\text{rref}(A) = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}.$$

## Homogeneous Linear System

Homogeneous linear system

Systems of linear equations in which each of the constant terms is zero are called **homogeneous**. A homogeneous system of $m$ equations in $n$ variables has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Homogeneous linear system is always consistent.

A homogeneous system must have at least one solution. Specifically, if all variables in a homogeneous system have the value zero, then each of the equations is satisfied. Such a solution is called **trivial**[7].

7: 注意这个词!

**Example 1.3.3** Solve the linear system

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= 0 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 0 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 0 \end{aligned}$$

The next example is called two birds with one stone. Keep in mind that some questions can be solved **simultaneous**, which saves time and energy.

**Example 1.3.4** Solve the systems of linear equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + 2x_2 + 4x_3 &= 0 \\ x_1 + 3x_2 + 9x_3 &= 0 \end{aligned} \qquad \begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_2 + 4x_3 &= 1 \\ x_1 + 3x_2 + 9x_3 &= 0 \end{aligned}$$

## 1.4 Applications

**Example 1.4.1** (polynomial curve fitting) Find a polynomial function of degree $n-1$

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n-1} x^{n-1}$$

whose graph passes through the $n$-points in the $xy$-plane

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

where $x_i \neq x_j$ for $i \neq j$.

**Solution:** To solve for the $n$ coefficients of $p(x)$, substitute each of the $n$ points into the polynomial function and obtain $n$ linear equations in $n$ variables $a_0, a_1, a_2, \ldots, a_{n-1}$.

The coefficient matrix of the system is:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

This kind of matrix is called Vandermont matrix.

$$\begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_{n-1} x_1^{n-1} = y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_{n-1} x_2^{n-1} = y_2 \\ \qquad \vdots \\ a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_{n-1} x_n^{n-1} = y_n \end{cases}$$

We claim the linear system has a unique solution. It is too tough to prove it using the method we learn in this chapter. It will become clear once we finish Chapter 3.

Hint: $\begin{pmatrix} 1 & 1 & 1 & 4 \\ 1 & 2 & 4 & 0 \\ 1 & 3 & 9 & 12 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & 24 \\ 0 & 1 & 0 & -28 \\ 0 & 0 & 1 & 8 \end{pmatrix}.$

**Exercise 1.4.1** Determine the polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ whose graph passes through the points $(1, 4)$, $(2, 0)$, and $(3, 12)$.

**Exercise 1.4.2** Find a polynomial of degree four that fits the points $(-2, 3)$, $(-1, 5)$, $(0, 1)$, $(1, 4)$, and $(2, 10)$.

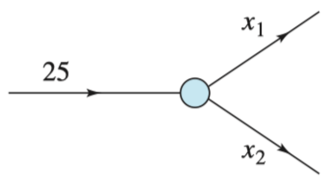
For the next question, note that $2012^4 =$ ?? is a very large number. We should think of a smart way to solve this question.

**Exercise 1.4.3** Find a polynomial of degree 4 that fits the points

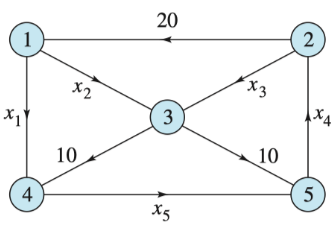
$$(2011, 3), \quad (2012, 5), \quad (2013, 1), \quad (2014, 4), \quad (2015, 10)$$

**Network Analysis**

Networks composed of branches and junctions are used as models in such fields as economics, traffic analysis, and electrical engineering. In a network model, you assume that the total flow into a junction is equal to the total flow out of the junction. For example, the junction shown below has 25 units flowing into it, so there must be 25 units flowing out of it. You can represent this with the linear equation




**Example 1.4.2** Set up a system of linear equations to represent the network shown below. Then solve the system.



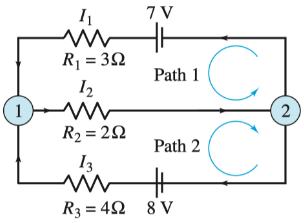
Discussion: How to interpret the result?

**Kirchhoff's Laws.**

In an electrical network,

1. All the current flowing into a junction must flow out of it.
2. The sum of the products $IR$ ( $I$ is current and $R$ is resistance) around a closed path is equal to the total voltage in the path.

**Example 1.4.3** Determine the currents $I_1$, $I_2$, and $I_3$ for the electrical network shown below.




Try this Exercise.

Interpret the solution set of linear system from geometric viewpoint?

## 1.5 SageMath

SageMath is a free open-source mathematics software system with Python-based language. its mission is to create a viable free open source alternative to (the expensive) Magma, Maple, Mathematica and Matlab (famous software).

**Figure 1.3:** The logo of SageMath.

The official website is: https://www.sagemath.org/.

In this course, I will show you how to solve exercise 1.3.1 in the accompanied lab manual, and much more.

# 2 Matrix

A football stadium has three concession areas, located in the south, north, and west stands. The top-selling items are peanuts, hot dogs, and soda. Sales for one day are given in the first matrix below, and the prices (in dollars) of the three items are given in the second matrix.



**Numbers of Items Sold**

| | Peanuts | Hot Dogs | Sodas | Selling Price | |
|---|---|---|---|---|---|
| South Stand | 120 | 250 | 305 | 2.00 | Peanuts |
| North Stand | 207 | 140 | 419 | 3.00 | Hot Dogs |
| West Stand | 29 | 120 | 190 | 2.75 | Soda |

Discussion: let's discuss the means of rows and columns of the matrix.

One can see that the way to record information is very effective. The information is stored in matrix.

## 2.1 The equality of matrices

It is standard **mathematical convention** to represent matrices in any one of the three ways listed below.

1. An uppercase letter such as $A$, $B$, or $C$.
2. A representative element enclosed in brackets, such as $[a_{ij}]$ , $[b_{ij}]$.
3. A rectangular array of numbers

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

The entry $a_{ij}$ of a matrix $A$ is usually denoted as $A_{ij}$.

**Definition 2.1.1** Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are equal when they have the same size $(m \times n)$ and $a_{ij} = b_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Summerize here how to check two matrices equal.

1.

2.

Consider the four matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, C = \begin{bmatrix} 1 & 3 \end{bmatrix}, \text{ and } D = \begin{bmatrix} 1 & 2 \\ x & 4 \end{bmatrix}$$

Matrices $A$ and $B$ are **not** equal because they are of different sizes. Similarly, $B$ and $C$ are not equal. Matrices $A$ and $D$ are equal if and only if $x = 3$.

## 2.2 The operations of matrices

**Definition 2.2.1** (Addition and Scalar Multiplication)

**Addition** If $A = [a_{ij}]$ and $B = [b_{ij}]$ are matrices of size $m \times n$, then their sum is the $m \times n$ matrix $A + B = [a_{ij} + b_{ij}]$. [1]

**Scalar Multiplication** If $A = [a_{ij}]$ is an $m \times n$ matrix and $c$ is a scalar, then the scalar multiple of $A$ by $c$ is the $m \times n$ matrix $cA = [ca_{ij}]$.

1: Remark: $A - B = A + (-1)B$.
问题: 我们有没有定义矩阵减法?

Try this Exercise.

**Example 2.2.1** For the matrices $A$ and $B$, find (a) $3A$, (b) $-B$, and (c) $3A - B$.

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix}$$

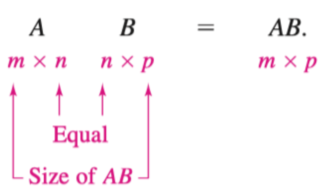


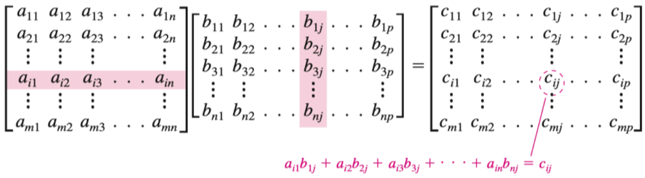

**Figure 2.1:** The size of $AB$

**Definition 2.2.2** (Matrix Multiplication) If $A = [a_{ij}]$ is an $m \times \mathbf{n}$ matrix and $B = [b_{ij}]$ is an $\mathbf{n} \times p$ matrix, then the product $AB$ is an $m \times p$ matrix

$$AB = [c_{ij}]$$

where

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{n} a_{ik} b_{kj} \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} \end{aligned}$$




**Example 2.2.2** Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

1. Find $A + B$ and $B + A$. Is matrix addition commutative? that is, $A + B = B + A$.

2. Find $AB$ and $BA$. Is matrix multiplication commutative? that is, $AB = BA$.

Put your answer here!

Try this Exercise.

**Definition 2.2.3** (Vector)

**Column vector** A matrix has only one column.
**Row vector** A matrix has only one row.

Let us fix the notations for vectors.

1. Boldface lowercase letters often designate column matrices and row matrices. For instance, **u**, **v**, $\alpha_1, \alpha_2$.
2. A matrix can be partitioned into column vectors.

- Vector can be considered as a record of information.
- Matrix can be considered as a record of a multiple-dimensional information.

Discussion: Come back to the example at the beginning of this chapter. How to interpret the result of matrix multiplication below?

$$\begin{bmatrix} 120 & 250 & 305 \\ 207 & 140 & 419 \\ 29 & 120 & 190 \end{bmatrix} \begin{bmatrix} 2.00 \\ 3.00 \\ 2.75 \end{bmatrix} = \begin{bmatrix} 1828.75 \\ 1986.25 \\ 940.50 \end{bmatrix}$$

Question: What is the total sales of all the three stands?

Let us try the next exercise as a motivation.

Write down here what you found from this exercise.

**Exercise 2.2.1** Compute

1. $\begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ -7 \\ -9 \end{bmatrix} - \begin{bmatrix} -6 \\ 8 \\ 12 \end{bmatrix} + 3 \begin{bmatrix} 6 \\ -5 \\ -9 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix}.$

2. $\begin{pmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 3 & -9 & 12 & -9 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -1 \\ 3 \\ -2 \end{pmatrix}$

One application of matrix multiplication is that the linear system

Critical thinking: Linear system ⇔ Matrix equation. Someone claims that linear algebra is a generalization of $ax = b$.

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

can be written as the **matrix equation** $A\mathbf{x} = \mathbf{b}$,

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

2: 动词; 词义: 分块.

Let $A = [a_{ij}]$ be a matrix of size $m \times n$ and let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$. Partition[2] matrix $A$ as $A = [\alpha_1 \ \alpha_2 \ \ldots \ \alpha_n]$. Show that

Critical Thinking: The idea of partition. Usually we see a matrix as a collection of column vectors!

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = A\mathbf{x}$$

▶ The linear system

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ 4x_1 + 5x_2 + 6x_3 &= 3 \\ 7x_1 + 8x_2 + 9x_3 &= 6 \end{aligned}$$

- The matrix equation.

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

- The vector equation.

$$x_1 \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

Critical thinking The nature of linear system is to express a vector as a linear combination of the other vectors.

**Theorem 2.2.1** Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}.$$

Partition the matrix $B$ as $B = [\beta_1 \quad \beta_2 \quad \dots \quad \beta_p]$. Then

$$AB = [A\beta_1 \quad A\beta_2 \quad \dots \quad A\beta_p]$$

思维训练：学习矩阵乘法的深层次原因是什么？将一个矩阵分解成几个矩阵的乘积。这些矩阵比较特殊，具有明显的特征。体现的数学思想：将一个复杂问题 (类比矩阵) 分解成几个简单的阶段性问题（类比矩阵分解)。通过复合（矩阵乘法)，较好的理解问题的**实质**。
通俗的讲，你学习数学的意义是：先研究一个特殊问题。然后学习将一般问题转换为特殊问题的方式方法，最终解决问题。该思想在常微分方程中由重要应用!

Proof. It is easy to see that both $AB$ and $[A\beta_1 \quad A\beta_2 \quad \dots \quad A\beta_p]$ have size $m \times p$. We next show that

$$(AB)_{ij} = [A\beta_1 \quad A\beta_2 \quad \dots \quad A\beta_p]_{ij}.$$

RHS $= [A\beta_1 \quad A\beta_2 \quad \dots \quad A\beta_p]_{ij} = [A\beta_j]_i = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} =$ LHS. Therefore, $AB = [A\beta_1 \quad A\beta_2 \quad \dots \quad A\beta_p]$. □

The theorem have a great consequence.

**Brainstorm:**

Let $A = \begin{pmatrix} -2 & 4 & -4 \\ 1 & -1 & -1 \\ -2 & 2 & -1 \end{pmatrix}$. Find a matrix $B$ such that

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

What you learned from this example? Share with us.

Moreover, verify $BA$.

## 2.3 Properties of matrix operations

**Theorem 2.3.1** If $A$, $B$, and $C$ are $m \times n$ matrices, and $c$ and $d$ are scalars, then the properties below are true.

1. $A + B = B + A$ Commutative property of addition
2. $A+(B+C) = (A+B)+C$ Associative property of addition.
3. $(cd)A = c(dA)$ Associative property of multiplication
4. $1A = A$ Multiplicative identity
5. $c(A + B) = cA + cB$ Distributive property
6. $(c + d)A = cA + dA$ Distributive property

类比: 数的加法中的数字零与矩阵加法中的零矩阵.

$$0_{mn} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

One important property of the addition of real numbers is that the number 0 is the additive identity. That is, $c + 0 = c$ for any real number $c$. For matrices, a similar property holds. Specifically, if $A$ is an $m \times n$ matrix and $O_{mn}$ is the $m \times n$ matrix consisting entirely of zeros, then $A + O_{mn} = A$. The matrix $O_{mn}$ is a zero matrix, and it is the additive identity for the set of all $m \times n$ matrices.

When the size of the matrix is understood, you may denote a zero matrix simply by $O$ or $\mathbf{0}$.

**Theorem 2.3.2** (Properties of Zero Matrices) If $A$ is an $m \times n$ matrix and $c$ is a scalar, then the properties below are true.

1. $A + O_{mn} = A$
2. $A + (-A) = O_{mn}$
3. If $cA = O_{mn}$, then $c = 0$ or $A = O_{mn}$.

**Example 2.3.1** Solve for $X$ in the equation $3X + A = B$, where

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix}.$$

请体会数学证明的严谨性. 我们没有证明"移项"性质, 就不可以使用.

Try this Exercise.

**Solution:**

$$\begin{aligned}
3X + A &= B \\
(3X + A) + (-A) &= B + (-A) && (\text{两边同时加上}(-A)) \\
3X + (A + (-A)) &= B - A && ((\text{由定理2.3.1, 第 2 条})) \\
3X + 0_{mn} &= B - A && (\text{由定理2.3.2, 第 2 条}) \\
3X &= B - A && (\text{由定理2.3.2, 第 1 条}) \\
\frac{1}{3}(3X) &= \frac{1}{3}(B - A) && (\text{等式两边同时乘以}\frac{1}{3}) \\
(\frac{1}{3}3)X &= \frac{1}{3}(B - A) && (\text{由定理2.3.1, 第 3 条}) \\
1X &= \frac{1}{3}(B - A) && (\text{由定理2.3.1, 第 4 条}) \\
X &= \frac{1}{3}(B - A) \\
&= \begin{bmatrix} \frac{-4}{3} & 2 \\ \frac{2}{3} & \frac{-2}{3} \end{bmatrix}
\end{aligned}$$

## Properties of matrix multiplications

**Theorem 2.3.3** If $A$, $B$, and $C$ are matrices (with sizes such that the matrix products are defined), and $c$ is a scalar, then the properties below are true.

1. $A(BC) = (AB)C$ Associative property of multiplication.
2. $A(B + C) = AB + AC$ Distributive property
3. $(A + B)C = AC + BC$ Distributive property
4. $c(AB) = (cA)B = A(cB)$

Proof.

1. Let $A = [a_{ij}]_{mn}, B = [b_{ij}]_{np}$ and $C = [c_{ij}]_{pq}$. To show this statement, it is enough to show that $(A(BC))_{ij} = ((AB)C)_{ij}$.

$$
\begin{aligned}
\text{LHS} &= \sum_{k=1}^{n} a_{ik}(BC)_{kj} \\
&= \sum_{k=1}^{n} \left( a_{ik} \left( \sum_{l=1}^{p} b_{kl}c_{lj} \right) \right) \\
&= \sum_{k=1}^{n} \left( \sum_{l=1}^{p} a_{ik}b_{kl}c_{lj} \right) \\
&= a_{i1}b_{11}c_{1j} + a_{i1}b_{12}c_{2j} + \ldots + a_{i1}b_{1p}c_{pj} \\
&+ a_{i2}b_{21}c_{1j} + a_{i1}b_{12}c_{2j} + \ldots + a_{i2}b_{2p}c_{pj} \\
&+ \ldots \\
&+ a_{in}b_{n1}c_{1j} + a_{in}b_{n2}c_{2j} + \ldots + a_{in}b_{np}c_{pj} \\
&= a_{i1}b_{11}c_{1j} + a_{i2}b_{21}c_{1j} + \ldots + a_{in}b_{n1}c_{1j} \\
&+ a_{i1}b_{12}c_{2j} + a_{i1}b_{12}c_{2j} + \ldots + a_{in}b_{n2}c_{2j} \\
&+ a_{i1}b_{1p}c_{pj} + a_{i2}b_{2p}c_{pj} + \ldots + a_{in}b_{np}c_{pj} \\
&= (a_{i1}b_{11} + a_{i2}b_{21} + \ldots + a_{in}b_{n1})c_{1j} \\
&+ (a_{i1}b_{12} + a_{i1}b_{12} + \ldots + a_{in}b_{n2})c_{2j} \\
&+ (a_{i1}b_{1p} + a_{i2}b_{2p} + \ldots + a_{in}b_{np})c_{pj} \\
&= (AB)_{i1}c_{1j} + (AB)_{i2}c_{2j} + \ldots + (AB)_{ip}c_{pj} \\
&= ((AB)C)_{ij}.
\end{aligned}
$$

2. Practice.
3. Practice.
4. Let $A = [a_{ij}]_{mn}, B = [b_{ij}]_{np}$. It is easy to see that the the matrices $c(AB)$, $(cA)B$ and $A(cB)$ have the same size. We are going to show that

We use this property very often in this course.

$$(c(AB))_{ij} = ((cA)B)_{ij} = (A(cB))_{ij}.$$

$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, thus $(c(AB))_{ij} = c(AB)_{ij} = c\sum_{k=1}^{n} a_{ik}b_{kj}$.
$((cA)B)_{ij} = \sum_{k=1}^{n} (ca_{ik})b_{kj} = c\sum_{k=1}^{n} a_{ik}b_{kj}$.
$(A(cB))_{ij} = \sum_{k=1}^{n} a_{ik}(cb_{kj}) = c\sum_{k=1}^{n} a_{ik}b_{kj}$. The statement is proved.

□

Commutativity is a strong property for matrix multiplication. However, it does not hold in general[3].

3: To show a statement is false, you only need to give an example.

**Example 2.3.2** Show $AB$ and $BA$ are not equal for the matrices

$$A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$$

Solution

$$AB = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 4 & -4 \end{bmatrix},$$

$$BA = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 7 \\ 4 & -2 \end{bmatrix}.$$

Thus $AB \neq BA$.

We next show that cancellation property is not valid in general.

**Example 2.3.3** Show that $AC = BC$, where

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}.$$

What is your conclusion of Example 2.3.3?

## Identity matrix

A special type of square matrix that has 1's on the main diagonal and 0 's elsewhere is called an **identity matrix**.

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

When the order of the matrix is understood to be $n$, you may denote $I_n$ simply as $I$.

**Theorem 2.3.4** (Properties of the Identity Matrix) If $A$ is a matrix of size $m \times n$, then the properties below are true.

1. $AI_n = A$.
2. $I_m A = A$.

There is an unexpected results from Theorem 2.3.4. Partition identity matrix $I_n = [e_1\ e_2\ \dots\ e_n]$, where

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

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

and $F = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix}$,

$M = a_{n1}F^{n-1} + a_{n-11}F^{n-2} + \dots + a_{11}E$.

1. Show that $Ae_1 = Me_1$.
2. Suppose that $AF = FA$ from now on. Show that $Ae_2 = Me_2$. (**Hint:** $e_2 = Fe_1$)
3. Show that $A = M$.

**Example 2.3.4** Let $A = [\alpha_1\ \alpha_2\ \dots\ \alpha_n]$. Show that $Ae_i = \alpha_i$.

**Definition 2.3.1** If $A$ is an $n \times n$ matrix and if $k$ is a positive integer, then $A^k$ denotes the product of $k$ copies of $A$:

$$A^k = \underbrace{A \dots A}_{k}$$

**Example 2.3.5** Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Use mathematical induction to show that

$$A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

In mathematics, the **Fibonacci sequence** is a sequence in which each number is the sum of the two preceding ones, that is,

$$F_{n+1} = F_{n+1} + F_n.$$

Numbers that are part of the Fibonacci sequence are known as Fibonacci numbers, commonly denoted $F_n$. the first few values in the sequence are:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144$$

The Fibonacci sequence $\{F_n\}$ can be described by this matrix. Indeed, for any $n \in \mathbb{N}$,

$$\begin{bmatrix} F_{n+2} \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = A^{n+1} \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$$

**Exercise 2.3.1** Find $A^k$ if $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

### Transpose

**Definition 2.3.2** Let $A$ be a $m \times n$ matrix. The **transpose**[4] of $A$ is the $n \times m$ matrix,denoted by $A^T$, whose columns are formed from the corresponding rows of $A$.

4: 转置

**Theorem 2.3.5** Let $A$ and $B$ denote matrices whose sizes are appropriate for the following sums and products.

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. For any scalar, $(rA)^T = r(A^T)$
4. $(AB)^T = B^T A^T$

In Europe, people study the linear system like

$$[x_1 \quad x_2] \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = [5 \quad 6].$$

With the notion of transpose, we are able to solve it in our way $A\mathbf{x} = \mathbf{b}$, that is,

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

In mathematics, we solve one case, and then try to convert the others into this case. **It is a great idea!**

Proof. □

## 2.4 The inverse of a matrix

**Definition 2.4.1** (Inverse of a Matrix) An $n \times n$ matrix $A$ is **invertible** (or nonsingular) when there exists an $n \times n$ matrix $B$ such that

$$AB = BA = I_n$$

where $I_n$ is the identity matrix of order $n$. The matrix $B$ is the (multiplicative) **inverse** of $A$. A matrix that does not have an inverse is **noninvertible** (or **singular**).

**Theorem 2.4.1** (Uniqueness) If $A$ is an invertible matrix, then its inverse is **unique**.

**Questions:** how to show the inverse is unique?

Since the inverse is unique, the inverse of $A$ is denoted by $A^{-1}$.

**Theorem 2.4.2** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

- $A$ is invertible if and only if $ad - bc \neq 0$.
- if $ad - bc \neq 0$, then the inverse is

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

**Comment:** As a math major student, you should have the mathematical thinking. For instance, from the definition of invertible, you can show or prove the statement, not just memorize it.

The method used next is totally different with the one for $n = 2$.

Discussion: Suppose $A$ is invertible. Are you able to see that $AB = I_n$ as $n$-linear systems?

Let $B = [\beta_1, \beta_2, \dots, \beta_n]$ and $I_n = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n]$. $AB = I_n$ can be read as

$$AB = A[\beta_1, \beta_2, \dots, \beta_n] = [A\beta_1, A\beta_2, \dots, A\beta_n] = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n].$$

Thus we have **consistent** linear systems

$$A\beta_1 = \mathbf{e}_1, A\beta_2 = \mathbf{e}_2, \dots, A\beta_n = \mathbf{e}_n.$$

**Lemma 2.4.3** Let $A$ be a matrix of size $n \times n$ such that the $n$ Linear system $A\mathbf{x} = \mathbf{e}_1, A\mathbf{x} = \mathbf{e}_2, \dots, A\mathbf{x} = \mathbf{e}_n$ are consistent. Then for every $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ is always consistent.

Proof. For any $i = 1, 2, \dots, n-1$ or $n$, since the linear system $A\mathbf{x} = \mathbf{e}_i$ is consistent, we may assume the solution is $\beta_i$, that is, $A\beta_i = e_i$. Let $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$. Thus

$$\mathbf{b} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \dots + b_n\mathbf{e}_n.$$

$$\begin{aligned} A(b_1\beta_1 + b_2\beta_2 + \dots + b_n\beta_n) &= A(b_1\beta_1) + A(b_2\beta_2) + \dots + A(b_n\beta_n) \\ &= b_1(A\beta_1) + b_2(A\beta_2) + \dots + b_n(A\beta_n) \\ &= b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \dots + b_n\mathbf{e}_n \\ &= \mathbf{b} \end{aligned}$$

□

Recall from Theorem 1.3.2 that the linear system $A\mathbf{x} = \mathbf{b}$ is always consistent, then the coefficient matrix $A$ can be reduced to $I_n$.

## Algorithm for finding $A^{-1}$

The key point here is: if $A$ is row equivalent to $I$, then $[A \quad I]$ is row equivalent to $[I \quad A^{-1}]$. Otherwise, $A$ does not have an inverse.

**Example:** Find the inverse of the matrix $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -6 & 2 & 3 \end{bmatrix}$.

**Example:** Show the matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \\ -2 & 3 & -2 \end{bmatrix}$ has no inverse.

**Theorem 2.4.4** If $A$ is an invertible matrix, $k$ is a positive integer, and $c$ is a nonzero scalar, then $A^{-1}$, $A^k$, $cA$, and $A^T$ are invertible and the statements below are true.

1. $\left(A^{-1}\right)^{-1} = A$
2. $\left(A^k\right)^{-1} = A^{-1}A^{-1}\cdots A^{-1} = \left(A^{-1}\right)^k$
3. $(cA)^{-1} = \frac{1}{c}A^{-1}$
4. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

**Theorem 2.4.5** If $A$ and $B$ are invertible matrices of order $n$, then $AB$ is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}$$

**Corollary 2.4.6** If $A_i$, $i = 1, 2, \ldots, n$ are invertible square matrices of same size, then

$$\left(A_1A_2A_3 \cdot\cdot A_n\right)^{-1} = A_n^{-1} \cdot\cdot A_3^{-1}A_2^{-1}A_1^{-1}$$

**Theorem 2.4.7** If $C$ is an invertible matrix, then the properties below are true.

1. If $AC = BC$, then $A = B$.
2. If $CA = CB$, then $A = B$.

**Theorem 2.4.8** If $A$ is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

**Discussion:** Let $A$ be an invertible $n \times n$ matrix. Sometimes, the inverse question of Theorem 2.4.8 will be asked. That is, if the matrix equation $A\mathbb{x} = \mathbf{b}$ is always consistent for any $\mathbf{b} \in \mathbb{R}^n$, is $A$ invertible? hint: $A$ is invertible if and only if $A$ is row reduced to the identity matrix.

## 2.5 Elementary matrices

**Elementary Row Operations**

There are three elementary row operations for matrices listed below.

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.
2. (Interchange) Interchange two row.
3. (Scaling) Multiply all entries in a row by a nonzero constant.

Similarly one can define elementary column operations.

**Elementary Column Operations**

There are three elementary column operations for matrices listed below.

1. (Replacement) Replace one column by the sum of itself and a multiple of another column.
2. (Interchange) Interchange two columns.
3. (Scaling) Multiply all entries in a column by a nonzero constant.

**Definition 2.5.1** (Elementary Matrix) An $n \times n$ matrix is an elementary matrix when it can be obtained from the identity matrix $I_n$ by a single elementary row operation.

**Question:** Which of the matrices below are elementary matrix? For those that are, describe the corresponding elementary row operation.

a. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ e. $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ f. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

There are three types of elementary matrix. We denote

$$E(i,j) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & \cdots & 1 & & \\ & & & \vdots & \ddots & \vdots & & \\ & & & 1 & \cdots & 0 & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i-\text{th row} \\ \\ j-\text{th row} \\ \\ \\ \\ \end{matrix},$$

$$E(i;k) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i-\text{th row} \\ \\ \\ \\ \end{matrix},$$

$$E(i,j;k) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & k & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ i-\text{th row} \\ \\ j-\text{th row} \\ \\ \\ \end{matrix}.$$

**Theorem 2.5.1** (Elementary Matrices Are Invertible) If $E$ is an elementary matrix, then $E^{-1}$ exists and is an elementary matrix.

**Proof:** (Direct computation)

1. $E^{-1}(i,j) = E(i,j)$.
2. $E^{-1}(i;k) = E(i;\frac{1}{k})$.
3. $E^{-1}(i,j;k) = E(i,j;-k)$

**Example 2.5.1** Verify

$$(a)\quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 1 & -3 & 6 \\ 3 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 6 \\ 0 & 2 & 1 \\ 3 & 2 & -1 \end{bmatrix}$$

$$(b)\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 2 & 6 & -4 \\ 0 & 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & 3 & -2 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

$$(c)\quad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -2 & -2 & 3 \\ 0 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 1 \\ 0 & 4 & 5 \end{bmatrix}$$

Exercise: Fill the blank below.

$$(1).\quad \begin{bmatrix} 0 & 2 & 1 \\ 1 & -3 & 6 \\ 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{\hspace{4cm}}.$$

$$(2).\quad \begin{bmatrix} 1 & 0 & -4 \\ 0 & 2 & 6 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{\hspace{4cm}}.$$

$$(3).\quad \begin{bmatrix} 1 & 0 & -1 \\ -2 & -2 & 3 \\ 0 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{\hspace{4cm}}.$$

Why this theorem is so important? We can write the elementary matrix operations into matrix multiplication. The latter can be understood by the computer!

**Theorem 2.5.2** (Representing Elementary Row(Column) Operations) Let $A$ be an $m \times n$ matrix. Let $E$ be the elementary matrix obtained by performing an elementary row(column) operation on $I_m(I_n)$. If that same elementary row(column) operation is performed on $A$, then the resulting matrix is the product $EA(AE)$.

**Example 2.5.2** (Using Elementary Matrices) Find a sequence of elementary matrices that can be used to write the matrix $A$ in row-echelon form.

$$A = \begin{bmatrix} 0 & 1 & 3 & 5 \\ 1 & -3 & 0 & 2 \\ 2 & -6 & 2 & 0 \end{bmatrix}$$

**SOLUTION**

| Matrix | Elementary Row Operation | Elementary Matrix |
|---|---|---|
| $\begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 2 & -6 & 2 & 0 \end{bmatrix}$ | $R_1 \leftrightarrow R_2$ | $E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ |
| $\begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & -4 \end{bmatrix}$ | $R_3 + (-2)R_1 \rightarrow R_3$ | $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ |
| $\begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & -2 \end{bmatrix}$ | $\left(\frac{1}{2}\right)R_3 \rightarrow R_3$ | $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$ |

The three elementary matrices $E_1$, $E_2$, and $E_3$ can be used to perform the same elimination.

$$B = E_3E_2E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 3 & 5 \\ 1 & -3 & 0 & 2 \\ 2 & -6 & 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 1 & 3 & 5 \\ 2 & -6 & 2 & 0 \end{bmatrix}$$

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

**Definition 2.5.2** (Row Equivalence) Let $A$ and $B$ be $m \times n$ matrices. Matrix $B$ is row-equivalent to $A$ when there exists a finite number of elementary matrices $E_1, E_2, \dots, E_k$ such that

$$B = E_kE_{k-1} \cdots E_2E_1A$$

**Theorem 2.5.3** (A Property of Invertible Matrices) A square matrix $A$ is invertible if and only if it can be written as the product of elementary matrices, that is,

$$A = E_kE_{k-1} \cdots E_2E_1,$$

where $E_i$ are elementary matrices.

**Corollary 2.5.4** A square matrix $A$ is invertible if and only if it row equivalent to identity matrix.

**Example:** Find a sequence of elementary matrices whose product is the nonsingular matrix

$$A = \begin{bmatrix} -1 & -2 \\ 3 & 8 \end{bmatrix}$$

**Theorem 2.5.5** (Equivalent Conditions) If $A$ is an $n \times n$ matrix, then the statements below are equivalent.

1. $A$ is invertible.
2. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $n \times 1$ column matrix $\mathbf{b}$.
3. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
4. $A$ is row-equivalent to $I_n$.
5. A can be written as the product of elementary matrices.

1. Identity matrix: $I_n$
2. Scalar matrix: $\lambda I_n$
3. Diagonal matrix $A$ such that $a_{ij} = 0$ for all $i \neq j$.
4. Elementary matrices: $E(i, j)$, $E(i; k)$ and $E(i, j; k)$.
5. Zero matrix.
6. Upper triangular matrix: $a_{ij} = 0$ for all $i > j$.
7. Lower triangular matrix: $a_{ij} = 0$ for all $i < j$.
8. Strictly Upper triangular matrix: $a_{ij} = 0$ for all $i \geq j$.
9. Strictly lower triangular matrix: $a_{ij} = 0$ for all $i \leq j$.

**Exercise:** Let $A$ be a $n \times n$ strictly upper triangular matrix. Show that $A^n = 0$.

# 3 Determinant

## Motivation



The purpose of this section is to define the determinant of the square matrices of$2 \times 2$ and $3 \times 3$.

**linear system in two variables**

Solve the linear system

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases}$$

1. We may assume that $a_{11} \neq 0$ and $a_{21} \neq 0$. Otherwise, the linear system is equivalent to an linear system in echelon form which is easy to solve. Then we obtain an equivalent linear system

$$\begin{cases} a_{21}a_{11}x + a_{21}a_{12}y = a_{21}b_1 & (1) \\ a_{11}a_{21}x + a_{11}a_{22}y = a_{11}b_2 & (2), \end{cases}$$

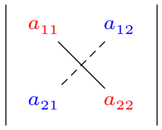
2. equation (2) subtracts equation (1) gives (an equivalent linear system)

$$\begin{cases} a_{21}a_{11}x + a_{21}a_{12}y & = a_{21}b_1 \\ (a_{11}a_{22} - a_{12}a_{21})\, y & = a_{11}b_2 - a_{21}b_1, \end{cases}$$

3. If $a_{11}a_{22} - a_{12}a_{21} \neq 0$, then $y = \frac{a_{11}b_2 - b_1 a_{21}}{a_{11}a_{22} - a_{12}a_{21}}$.
4. (Question:) Figure out the value of $x$.

▶ It seems that $a_{11}a_{22} - a_{12}a_{21}$ is critical for the linear system.

How to memorize it?




▶ Define the determinant of the $2 \times 2$ matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ by

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

Solve

$$\begin{cases} 3x_1 - 2x_2 = 12 \\ 2x_1 + x_2 = 1 \end{cases}$$

▶ The solution can be expressed as

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{D_1}{D}, \qquad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{D_2}{D}.$$

**linear system of three variables**

For the linear system in three variables,

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

If $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \neq 0$, then the linear system has a unique solution:

$$x = \frac{b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-b_1a_{23}a_{32}-a_{12}b_2a_{33}-a_{13}a_{22}b_3}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$

$$y = \frac{a_{11}b_2a_{33}+b_1a_{23}a_{31}+a_{13}a_{21}b_3-a_{11}a_{23}b_3-b_1a_{21}a_{33}-a_{13}b_2a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$

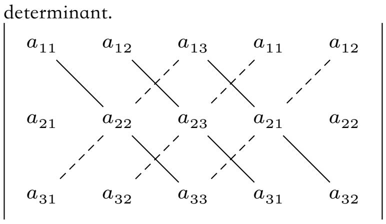
$$z = \frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-a_{11}b_2a_{32}-a_{12}a_{21}b_3-b_1a_{22}a_{31}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}}$$

In 2022, I thought it is not easy to see how to get this formula! With the help of SageMath, it seems a easy question.

Define the determinant of $3 \times 3$ matrix

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

A way to memorize the definition of determinant.




An observation

$$\begin{aligned}&a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}\\=&a_{11}(a_{22}a_{33}-a_{23}a_{32})+a_{12}(a_{23}a_{31}-a_{21}a_{33})+a_{13}(a_{21}a_{32}-a_{22}a_{31})\\=&a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix}+a_{13}\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}.\end{aligned}$$

Therefore

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}=a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix}+a_{13}\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}$$

Critical thinking: We can define determinants of $3\times 3$ matrices in terms of determinants of $2\times 2$ matrices. This encourages us to define determinants of $4\times 4$ matrices in terms of determinants of $3\times 3$ matrices. Inductively we can define the determinants of $n\times n$ matrices.

### Area and Volumes

Compute the area of a parallelgram determined by vectors $\langle 1,3\rangle$ and $\langle 4,1\rangle$.

One way to get the area is to compute the absolute value of the determinant of the matrix

$$A=\begin{pmatrix}1&3\\4&1\end{pmatrix}.$$

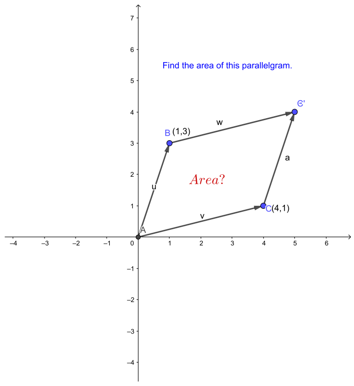
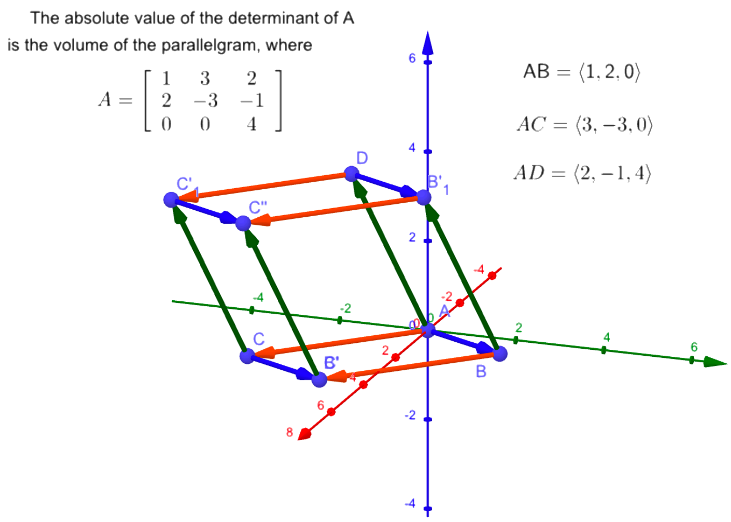



**Figure 3.1:** Geometry interpretation of determinants

Here is another example.




## 3.1 Definition of determinant

**Definition 3.1.1** (Minors and Cofactors of a square matrix) If $A$ is a square matrix, then the minor $M_{ij}$ of the entry $a_{ij}$ is the determinant of the matrix obtained by deleting the $i$-th row and $j$-th column of $A$. The cofactor $C_{ij}$ of the entry $a_{ij}$ is $C_{ij} = (-1)^{i+j} M_{ij}$.

**Example 3.1.1** Let $A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}$. Find the minor $M_{11}$, $M_{12}$, $M_{13}$, $C_{11}$, $C_{12}$ and $C_{13}$, respectively.

**Definition 3.1.2** For $n \geq 2$, the determinant of an $n \times n$ square matrix $A = [a_{ij}]$ is defined as

$$\begin{aligned} \det(A) &= a_{11}M_{11} - a_{12}M_{12} + \ldots + (-1)^{n+1} a_{1n} M_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} M_{1j} \end{aligned}$$

**Example 3.1.2** Using definition compute the determinants of the matrices below, respectively.

Observation. $\det(B)$ equals to the product of diagonal entries.

$$A = \begin{bmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -7 & 8 & 9 \\ 0 & 2 & -5 & 7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

There is a better way to compute the determinant of a square matrix.

**Theorem 3.1.1** The determinant of an $n \times n$ matrix $A$ can be computed by a cofactor expansion across any row or down any column. The expansion across the $i$ th row using the cofactors is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

The cofactor expansion down the $j$ th column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

Toughest Question so far: How to prove the theorem?
Vote: Should we prove this theorem?

**Example 3.1.3** Compute $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

**Corollary 3.1.2** Let $A$ be an $n \times n$ matrix. If there is a zero row or zero column of $A$, then $\det(A) = 0$.

Next, we give the definitions of upper and lower triangular matrices.

1. Upper triangular matrices.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,n-1} & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

2. Lower triangular matrices $A = [a_{ij}]$ with $a_{ij} = 0$ for all $i < j$.

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \cdots & a_{n-1,n-1} & 0 \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & 0 & a_{n,n} \end{bmatrix}$$

**Example 3.1.4** Compute the determinants of the upper and lower triangular matrices, respectively. Briefly explain your method.

**Example 3.1.5** Compute the determinants of the elementary matrices $E(i,j)$, $E(i;k)$ and $E(i,j;k)$, respectively.

Critical thinking: a square matrix is row equivalent to an upper triangular matrix! that is, there exist elementary matrices $E_1, E_2, \dots, E_n$ such that $E_n E_{n-1} \dots E_1 A$ is an upper triangular matrix.

**Exercise 3.1.1** Compute the determinant of the matrix

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & a_{1n} \\ 0 & 0 & \cdots & a_{2,n-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n-1,2} & \cdots & 0 & 0 \\ a_{n1} & 0 & 0 & 0 & 0 \end{bmatrix}$$

## 3.2 Row operations and Determinant.

**Lemma 3.2.1** If two rows of a matrix $A$ are equal, then $\det(A) = 0$.

This lemma can be proved by mathematical induction.

**Theorem 3.2.2** (Elementary Row Operations and determinants) Let $A$ be a square matrix.

1. If one row of $A$ is multiplied by $k$ to produce $B$, then
$$\det(B) = k \det(A).$$
2. If a multiple of one row of $A$ is added to another row to produce a matrix $B$, then
$$\det(B) = \det(A).$$
3. If two rows of $A$ are interchanged to produce $B$, then
$$\det(B) = -\det(A).$$

**Interpretation of Theorem 3.2.2**

Let $A$ be a square matrix and $E$ is an elementary matrix. Then
$$\det(EA) = \det(E)\det(A).$$

Critical thinking: It is more efficient to compute the determinant of a matrix through it is echelon form.

**Corollary 3.2.3** Let $A$ be a square matrix and $E_1, E_2, \dots, E_m$ are elementary matrices. Then
$$\det(E_m E_{m-1} \dots E_1 A) = \det(E_m)\det(E_{m-1}) \dots \det(E_1)\det(A).$$

**Corollary 3.2.4** (Determinant of a Scalar Multiple of a Matrix) If $A$ is a square matrix of order $n$ and $c$ is a scalar, then the determinant of $cA$ is $\det(cA) = c^n \det(A)$

Compute the determinant of the matrix $A$, where $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$

**Example 3.2.1** Compute $\det(A)$, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

## 3.3 Properties of Determinants

Recall that A square matrix $A$ is invertible if and only if it can be written as the product of elementary matrices. Thus it follows from Corollary 3.2.3 that

**Theorem 3.3.1** A square matrix $A$ is invertible iff $\det(A) \neq 0$.

**Theorem 3.3.2** (Determinant of a Matrix Product) If $A$ and $B$ are square matrices of order $n$, then $\det(AB) = \det(A)\det(B)$.

The next corollary follows from Theorem 3.3.2.

**Corollary 3.3.3** (Determinant of an Inverse Matrix) If $A$ is an $n \times n$ invertible matrix, then $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$.

**Example 3.3.1** Use two ways to compute $|A^{-1}|$, where

$$A = \begin{bmatrix} 3 & 1 & -2 \\ 2 & 0 & 0 \\ -4 & -1 & 5 \end{bmatrix}.$$

**Example 3.3.2** Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 2 \\ 2 & 1 & 0 \end{bmatrix}$, and $B$ be a $3 \times 3$ matrix whose determinant is $-2$.

1. Find $\det(A)$.
2. Find $\det(2A)$.
3. Find $|A^{-1}|$.
4. Find $\det(A^3)$.
5. Find $\det(A^2B^{-3})$.

### Determinant of a Transpose

**Theorem 3.3.4** If $A$ is a square matrix, then

$$\det(A) = \det\left(A^T\right)$$

**Conditions That Yield a Zero Determinant**

If $A$ is a square matrix and any one of the conditions below is true, then $\det(A) = 0$.

1. An entire row (or an entire column) consists of zeros.
2. Two rows (or columns) are equal.
3. One row (or column) is a multiple of another row (or column).

Observation: the second and third conditions yield a zero row (column) after the elementary row(column) operation: replacement.

## 3.4 Cramer's Rule

Let $A = [a_{ij}]$ be an $n \times n$ matrix and $|A| = D$.

Observation: Compute the determinant of the matrix

$$B = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ & & \vdots & & \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ & & \vdots & & \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ & & \vdots & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{matrix} \\ \\ i-\text{th row} \\ \\ j-\text{th row} \\ \\ \\ \end{matrix}$$

Note that you should consider the cases $i = j$ and $i \neq j$.

Let $A$ be an $n \times n$ matrix and $|A| = D$. Compute

$$a_{1i}C_{1j} + a_{2i}C_{2j} + \cdots + a_{ni}C_{nj}.$$

**Lemma 3.4.1** Let $A$ be an $n \times n$ matrix and $|A| = D$.

$$a_{i1}C_{r1} + a_{i2}C_{r2} + \cdots + a_{in}C_{rn} = \begin{cases} D, & \text{if } i = r \\ 0, & \text{if } i \neq r \end{cases}$$

**Example 3.4.1** Let $A = \begin{vmatrix} 3 & 2 & 1 & -2 \\ 0 & 1 & 3 & 0 \\ 4 & -6 & 0 & 5 \\ -1 & 3 & -2 & 1 \end{vmatrix}$.

Find

1. $3C_{41} + 4C_{42} - 5C_{43} - 2C_{44}$.
2. $M_{23} + 2M_{33} + 3M_{43}$.
3. $3C_{31} + 2C_{32} + C_{33} - 2C_{34}$.

This lemma has two unexpected consequences: the second way to find the inverse of an invertible matrix and solving linear system.

### Inverse of a matrix

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$. Define the **adjoint matrix** of $A$, denoted by $\text{adj}(A)$,

$$\text{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$

**Theorem 3.4.2** If $A$ is an $n \times n$ invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A).$$

Critical thinking: Compute $A\,\text{adj}(A)$ and $\text{adj}(A)A$, respectively.

### Cramer's rule

**Theorem 3.4.3** Let $A$ be a square matrix of size $n$ and $\det(A) \neq 0$. The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)} \text{adj}(A)\mathbf{b}$$

Objective: Find another way to compute $x_i$.

$$A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} = \underline{\hspace{6cm}}$$

Then

$$x_i = [\frac{1}{\det(A)} \operatorname{adj}(A)\mathbf{b}]_i = \frac{1}{\det(A)}(b_1C_{1i} + b_2C_{2i} + \ldots + b_nC_{ni})$$

$$= \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

For any $n \times n$ matrix $A$ and any $\mathbf{b} \in \mathbb{R}^n$, let $A_i(\mathbf{b})$ be the matrix obtained from $A$ by replacing column $i$ by the vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{col } i}}{\mathbf{b}} \quad \cdots \quad \mathbf{a}_n]$$

**Theorem 3.4.4** (Cramer's Rule) Let $A$ be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \quad i = 1, 2, \ldots, n$$

# Vector Spaces 4

Vector space is a set of objects which satisfy 10 axioms[1]. The objects are called vectors. Vector space plays the role as the real numbers in Calculus. The set of all column (row) vectors with vector addition and scalar multiplication is a vector space, denoted by $\mathbb{R}^n$. That is,



$$\mathbb{R}^n = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \middle| a_i \in \mathbb{R} \right\}, \text{ with the operations}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \text{ and } c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}.$$

1: 公理. 公理无需证明

## Motivation: Vectors in the plane

**Observations**

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in the plane, and let $c$ and $d$ be scalars. Then the following ten equations hold.

1. $\mathbf{u} + \mathbf{v}$ is a vector in the plane.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
4. $\mathbf{u} + \mathbf{0} = \mathbf{u}$
5. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
6. $c\mathbf{u}$ is a vector in the plane.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1(\mathbf{u}) = \mathbf{u}$

**Theorem 4.0.1** Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$, and let $c$ be a scalar. Then the properties below are true.

1. The additive identity is unique. That is, if $\mathbf{v} + \mathbf{u} = \mathbf{v}$, then $\mathbf{u} = \mathbf{0}$.

2. The additive inverse of $\mathbf{v}$ is unique. That is, if $\mathbf{v}+\mathbf{u}=\mathbf{0}$, then $\mathbf{u}=-\mathbf{v}$.
3. $0\mathbf{v}=\mathbf{0}$
4. $c0=0$
5. If $c\mathbf{v}=\mathbf{0}$, then $c=0$ or $\mathbf{v}=\mathbf{0}$.
6. $-(-\mathbf{v})=\mathbf{v}$.

## 4.1 Vector Space $\mathbb{R}^n$

In this section, all vectors are in $\mathbb{R}^n$.

### Contents



### Linear Combinations

An important type of problem in linear algebra involves writing one vector $\mathbf{x}$ as the sum of scalar multiples of other vectors $\mathbf{v}_1, \mathbf{v}_2$, ..., and $\mathbf{v}_n$. That is, for scalars $c_1, c_2, \ldots, c_n$

$$\mathbf{x}=c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n$$

The vector $\mathbf{x}$ is called a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2$, ..., and $\mathbf{v}_n$ with weight $c_1, c_2, \ldots, c_n$.

**Example 4.1.1** Let vectors $\mathbf{x}=(-1,-2,-2), \mathbf{u}=(0,1,4), \mathbf{v}=(-1,1,2)$, and $\mathbf{w}=(3,1,2)$ in $\mathbb{R}^3$. Find scalars $a$, $b$, and $c$ such that

$$\mathbf{x}=a\mathbf{u}+b\mathbf{v}+c\mathbf{w}$$

**Example 4.1.2**

Given that $\begin{bmatrix} 2 & 2 & -3 & -16 \\ 1 & 2 & -2 & -14 \\ 1 & 1 & -1 & -6 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 4 \end{bmatrix}$ Write $\begin{bmatrix} -16 \\ -14 \\ -6 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \\ -1 \end{bmatrix}$.

$$\begin{bmatrix} -16 \\ -14 \\ -6 \end{bmatrix} = \_\_\_ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + \_\_\_ \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + \_\_\_ \begin{bmatrix} -3 \\ -2 \\ -1 \end{bmatrix}$$

### Linear independent

**Definition 4.1.1** An indexed set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be linearly independent if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be linearly dependent if there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

**Example 4.1.3** Show that $\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}3\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\6\\1\end{bmatrix}\right\}$ linear independent.

## Span

**Definition 4.1.2** If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ is denoted by $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. That is, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p | c_1, c_2, \ldots, c_p \in \mathbb{R}\}.$$

**Example 4.1.4** show $\mathbb{R}^3 = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}3\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\6\\1\end{bmatrix}\right\}$.

## Basis

**Definition 4.1.3** A basis for $\mathbb{R}^n$ is a linearly independent set that spans $\mathbb{R}^n$.

**Example 4.1.5** Is $\left\{\begin{bmatrix}1\\-4\\-3\end{bmatrix}, \begin{bmatrix}3\\2\\-2\end{bmatrix}, \begin{bmatrix}4\\-6\\-7\end{bmatrix}\right\}$ a basis of $\mathbb{R}^3$?

**Example 4.1.6** Is $\left\{\begin{bmatrix}1\\-4\\-3\\0\end{bmatrix}, \begin{bmatrix}3\\2\\-2\\1\end{bmatrix}, \begin{bmatrix}4\\-6\\-7\\2\end{bmatrix}\right\}$ a basis of $\mathbb{R}^4$?

## 4.2 Abstract Vector Space

Contents



**Definition 4.2.1** A **vector space** is a **nonempty set** $V$ of objects, called **vectors**, on which are defined **two operations**, called **addition and multiplication by scalars** (real number), subject to the **ten axioms** list below. The axioms must hold for all vector $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in $V$ and for all scalars $c$ and $d$.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u}+\mathbf{v}$, is in $V$.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.
4. There is a zero vector $\mathbf{0}$ in $V$ such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$.
5. For each $\mathbf{u}$ in $V$, there is a vector $-\mathbf{u}$ in $V$ such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by $c$, denoted by $c\mathbf{u}$, is in $V$.
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$.
9. $c(d\mathbf{u})=(cd)\mathbf{u}$
10. $1\mathbf{u}=\mathbf{u}$.

**Five Examples of Vector Spaces**

▶ The space $\mathbb{R}^n=\left\{\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\middle| x_i\in\mathbb{R}\right\}$ is a vector space with operations

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}+\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}=\begin{bmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{bmatrix} \text{ and } c\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}cx_1\\cx_2\\\vdots\\cx_n\end{bmatrix}.$$

▶ Let $\mathrm{P}_n(x)=\{a_0+a_1x+...+a_nx^n|a_i\in\mathbb{R}\}$ with two operations

$$\begin{aligned}&(a_0+a_1x+...+a_nx^n)+(b_0+b_1x+...+b_nx^n)\\=&(a_0+b_0)+(a_1+b_1)x+...+(a_n+b_n)x^n.\end{aligned}$$

$$k(a_0+a_1x+...+a_nx^n)=(ka_0)+(ka_1)x+...+(ka_n)x^n$$

▶ Let $M_{mn}(\mathbb{R})=\left\{\begin{bmatrix}a_{11}&\dots&a_{1n}\\\vdots&&\vdots\\a_{m1}&\dots&a_{mn}\end{bmatrix}\middle| a_{ij}\in\mathbb{R}\right\}$ is a vector space with the usual matrix addition and scalar multiplication.

**Zero vectors**

1. The zero vector of $\mathbb{R}^n$ is $\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$.
2. The zero vector of $\mathrm{P}_n(x)$ is $0=0+0x+0x^2+...+0x^n$.
3. The zero vector of $M_{mn}(\mathbb{R})$ is $\begin{bmatrix}0&\dots&0\\\vdots&\ddots&\vdots\\0&\dots&0\end{bmatrix}$
4. The zero vector of $C(-\infty,\infty)$ is $0$, the zero function.

- Let $C(-\infty, \infty)$ be the set of all real-valued continuous functions defined on the entire real line. $C(-\infty, \infty)$ is a vector space with operations

$$(f+g)(x) = f(x) + g(x) \text{ and } (cf)(x) = c[f(x)]$$

- Define $C[a, b]$ be the set of all real-valued continuous functions defined on the interval $[a, b]$. $C[a, b]$ is a vector space in the above sense.

**Theorem 4.2.1** Let $\mathbf{v}$ be any element of a vector space $V$, and let $c$ be any scalar. Then the properties below are true.

1. $0\mathbf{v} = \mathbf{0}$
2. $c\mathbf{0} = \mathbf{0}$
3. If $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.
4. $(-1)\mathbf{v} = -\mathbf{v}$.

**Corollary 4.2.2** Let $V$ be a vector space. Then the zero element of $V$ is unique.

**Sets that are not Vector Spaces**

1. The set of all integers (with the standard operations) does not form a vector space.
2. The set of all second-degree polynomials is not a vector space.
3. $P = \{1 + a_1x + \ldots a_nx^n | a_i \in \mathbb{R}\}$ is not a vector space.

**Example 4.2.1** Let $V = \mathbb{R}^2$, the set of all ordered pairs of real numbers, with the standard operation of addition and the nonstandard definition of scalar multiplication listed below.

$$c(x_1, x_2) = (cx_1, 0)$$

Show that $V$ is not a vector space.

## Subspace

**Definition 4.2.2** A nonempty subset $W$ of a vector space $V$ is a subspace of $V$ when $W$ is a vector space under the operations of addition and scalar multiplication defined in $V$.

Observation: If $W$ is a subset of a vector space $V$, the axioms 2, 3, 7, 8, 9, and 10 are satisfied automatically.

2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$.

In many applications in linear algebra, vector spaces occur as subspaces of larger spaces.

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} + \mathbf{v}$, is in $V$.

4. There is a zero vector $\mathbf{0}$ in $V$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.

5. For each $\mathbf{u}$ in $V$, there is a vector $-\mathbf{u}$ in $V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

6. The scalar multiple of $\mathbf{u}$ by $c$, denoted by $c\mathbf{u}$, is in $V$.

**Theorem 4.2.3** (Test for a Subspace) If $W$ is a nonempty subset of a vector space $V$, then $W$ is a subspace of $V$ if and only if the two closure conditions listed below hold.

1. If $\mathbf{u}$ and $\mathbf{v}$ are in $W$, then $\mathbf{u} + \mathbf{v}$ is in $W$.
2. If $\mathbf{u}$ is in $W$ and $c$ is any scalar, then $c\mathbf{u}$ is in $W$.

**Trivial subspace**

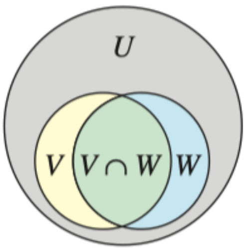
There are two obvious subspace of a vector space $V$.

1. $\{0\}$ 2. $V$.

We call them trivial.

**Example 4.2.2** Let $W$ be the set of all $2 \times 2$ matrices such that $A^T = A$. Show that $W$ is a subspace of the vector space $M_{2,2}$, with the standard operations of matrix addition and scalar multiplication.

**Example 4.2.3** Show that $W = \{(x_1, x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}$, with the standard operations, is not a subspace of $\mathbb{R}^2$.




**Theorem 4.2.4** If $V$ and $W$ are both subspaces of a vector space $U$, then the intersection of $V$ and $W$ (denoted by $V \cap W$) is also a subspace of $U$.

**Example 4.2.4** Let $W_5$ be the vector space of all functions defined on $[0,1]$, and let $W_1, W_2, W_3$, and $W_4$ be defined as shown below.

$$\begin{aligned} W_1 &= \text{ set of all polynomial functions defined on}[0,1] \\ W_2 &= \text{ set of all functions that are differentiable on } [0,1] \\ W_3 &= \text{ set of all functions that are continuous on } [0,1] \\ W_4 &= \text{ set of all functions that are integrable on } [0,1] \end{aligned}$$

Show that $W_1 \subset W_2 \subset W_3 \subset W_4$ and that $W_i$ is a subspace of $W_j$ for $i \leq j$.

### Subspace of $\mathbb{R}^2$

**Example 4.2.5** Determine whether each subset is a subspace of $\mathbb{R}^2$.

1. The set of points on the line $x + 2y = 0$
2. The set of points on the line $x + 2y = 1$

Remark: A subset $W$ is a subspace of $\mathbb{R}^n$ is a subspace only if $\mathbf{0}$ is in $W$ which is a very effective observation.

**subspaces of $\mathbb{R}^2$**

If $W$ is a subset of $\mathbb{R}^2$, then it is a subspace if and only if it has one of the forms listed below.

1. $W$ consists of the single point $(0,0)$.
2. $W$ consists of all points on a line that passes through the origin.
3. $W$ consists of all of $\mathbb{R}^2$.

### Subspace of $\mathbb{R}^3$

**subspace of $\mathbb{R}^3$**

A subset $W$ of $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$ if and only if it has one of the forms listed below.

1. $W$ consists of the single point $(0,0,0)$.
2. $W$ consists of points on a line that passes through the origin.
3. $W$ consists of points in a plane that passes through the origin.
4. $W$ consists of all of $\mathbb{R}^3$.

### Linear combination

**Definition 4.2.3** A vector $\mathbf{v}$ in a vector space $V$ is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ in $V$ when $\mathbf{v}$ can be written in the form

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$$

where $c_1, c_2, \ldots, c_k$ are scalars, and are called weights.

1. $(1,3,1) = 3(0,1,2) + (1,0,-5)$.
2. $\begin{bmatrix} 0 & 8 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} + 2\begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix}$.

**Example 4.2.6**

1. Write the vector $\mathbf{w} = (1,1,1)$ as a linear combination of vectors in the set

$$S = \{(1,2,3), (0,1,2), (-1,0,1)\}.$$

2. If possible, write the vector $\mathbf{w} = (1,-2,2)$ as a linear combination of vectors in the set $S$ above.

**Example 4.2.7**

1. Write the vector $\mathbf{w} = 1 + x + x^2$ as a linear combination of vectors in the set

$$S = \{1 + 2x + 3x^2, x + 2x^2, -1 + x^2\}.$$

2. If possible, write the vector $\mathbf{w} = 1 - 2x + 2x^2$ as a linear combination of vectors in the set $S$ above.

**Exercise 4.2.1** Write $\mathbf{w} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ as a linear combination of vectors $\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$.

## Spanning sets

**Definition 4.2.4** (Spanning Set) Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a subset of a vector space $V$. The set $S$ is a spanning set of $V$ when every vector in $V$ can be written as a linear combination of vectors in $S$. In such cases it is said that $S$ spans $V$.

**Example 4.2.8**

- The set $S = \{(1,0,0),(0,1,0),(0,0,1)\}$ spans $\mathbb{R}^3$ because any vector $\mathbf{u} = (u_1, u_2, u_3)$ in $\mathbb{R}^3$ can be written as

$$\mathbf{u} = u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1) = (u_1, u_2, u_3)\,.$$

- The set $S = \{1, x, x^2\}$ spans $P_2$ because any polynomial function $p(x) = a + bx + cx^2$ in $P_2$ can be written as

$$p(x) = a(1) + b(x) + c\left(x^2\right) = a + bx + cx^2$$

**Example 4.2.9** Show that $S = \{(1,2,3),(0,1,2),(-2,0,1)\}$ spans $\mathbb{R}^3$.

**Exercise 4.2.2**

- Show that $S = \{1 + 2x + 3x^2, x + x^2, -2 + x^2\}$ spans $P_2(x)$.
- $S = \{(1,2,3),(0,1,2),(-1,0,1)\}$ **does not span** $\mathbb{R}^3$.

## The span of a set

**Definition 4.2.5** If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a set of vectors in a vector space $V$, then **the span of** $S$ is the set of all linear combinations of the vectors in $S$,

$$\text{span}(S) = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k | c_1, c_2, \dots, c_k \in \mathbb{R}\}\,.$$

The span of $S$ is denoted by

$$\text{span}(S) \quad \text{or} \quad \text{span}\,\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}\,.$$

When $\text{span}(S) = V$, it is said that $V$ is **spanned** by $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, or that $S$ **spans** $V$.

**Example 4.2.10** Determine $\text{span}(S)$, where $S = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$.

**Example 4.2.11** Find the value of $a$ for which $v$ is in the set $H$, where

$$v = \begin{bmatrix} 2 \\ 6 \\ -9 \\ a \end{bmatrix}, H = \text{span}\left\{ \begin{bmatrix} -2 \\ -4 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -3 \\ 2 \end{bmatrix} \right\}.$$

This is an important theorem of this chapter.

**Theorem 4.2.5** If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a set of vectors in a vector space $V$, then span$(S)$ is a subspace of $V$. Moreover, span$(S)$ is the smallest subspace of $V$ that contains $S$, in the sense that every other subspace of $V$ that contains $S$ must contain span$(S)$.

## Linear independence

**Definition 4.2.6** A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ in a vector space $V$ is linearly independent when the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

has **only** the trivial solution

$$c_1 = 0, c_2 = 0, \dots, c_k = 0.$$

If there are also nontrivial solutions, then $S$ is linearly dependent.

**Example 4.2.12** Determine whether the set $S$ below of vectors in $\mathbb{R}^3$ is linearly independent or linearly dependent.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}.$$

**Example 4.2.13** Determine whether the set $S$ below of vectors in $P_2$ is linearly independent or linearly dependent.

$$S = \{1 + x - 2x^2, 2 + 5x - x^2, x + x^2\}$$

**Example 4.2.14** Determine whether the set of vectors in $M_{2,2}$ is linearly independent or linearly dependent.

$$S = \left\{ \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \right\}$$

**Example 4.2.15** Determine whether the set of vectors in $\mathbb{R}^4$ is linearly independent or linearly dependent.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \end{bmatrix} \right\}$$

**Exercise 4.2.3** (a). Determine whether the set of vectors in $\mathbb{R}^4$ is linearly dependent.

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

(b). show that $\text{span}(S) = \mathbb{R}^4$.

**Theorem 4.2.6** Suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ is linear independent, but $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is linear dependent. Then $\mathbf{v}_{k+1}$ must be able to written as linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$.

Proof. Suppose that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k + x_{k+1}\mathbf{v}_{k+1} = \mathbf{0} \qquad (*)$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is linear dependent, there exists a nontrivial solution $(a_1, a_2, \dots, a_{k+1}) \neq \mathbf{0}$ to the vector equation $(*)$.

We claim that $a_{k+1} \neq 0$. Suppose on the contrary that $a_{k+1} = 0$. Then $a_1v_1 + a_2v_2 + \cdots + a_kv_k = \mathbf{0}$. Since $\{v_1, \cdots, v_k\}$ is linearly independent, we know $a_1 = a_2 = \cdots = a_k = 0$, which contradicts to the assumption that $(a_1, a_2, \dots, a_{k+1}) \neq \mathbf{0}$. Thus $a_{k+1} \neq 0$. It follows from the vector equation (*) that $a_{k+1}\mathbf{v}_{k+1} = -a_1v_1 - a_2v_2 - \cdots - a_kv_k$, which implies that

$$\mathbf{v}_{k+1} = -\frac{a_1}{a_{k+1}}v_1 - \frac{a_2}{a_{k+1}}v_2 - \cdots - \frac{a_k}{a_{k+1}}v_k$$

□

Geometric view of Linear independent

1. $B = \{v_1\}$ is linear independent iff $v_1 \neq 0$.
2. $B = \{v_1, v_2\}$ is linear dependent iff they are co-line(On a same line).
3. $B = \{v_1, v_2, v_3\}$ is linear dependent iff they are co-plane(On a same plane ).

**Corollary 4.2.7** Two vectors $\mathbf{u}$ and $\mathbf{v}$ in a vector space $V$ are linearly dependent if and only if one is a scalar multiple of the other.

## Basis

**basis**

A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ in a vector space $V$ is a basis for $V$ when the conditions below are true.

1. $S$ spans $V$.
2. $S$ is linearly independent.

If a vector space V has a basis with a finite number of vectors, then V is **finite dimensional.** Otherwise, V is **infinite dimensional.**

### Standard Basis

1. The set $\left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$ is the standard basis of $\mathbb{R}^n$, i.e., $\mathbb{R}^n = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$.

2. The set $\{1, x, \dots, x^n\}$ is the standard basis of $\mathbf{P}_n(x)$.

3. Let $E_{ij} \in M_{mn}(\mathbb{R})$ be the matrix such that $(i, j)$-entry is 1 and all the other entries are zero. The set

$$\{E_{11}, \dots, E_{1n}, E_{21}, \dots, E_{2n}, \dots, E_{m1}, \dots, E_{mn}\}$$

is the standard basis of $M_{mn}(\mathbb{R})$.

**Example 4.2.16** Show that $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \\ 1 \end{bmatrix} \right\}$ is a basis of $\mathbb{R}^3$.

**Example 4.2.17** Show that $\{1 + x + x^2, 3 - x^2, 6x + x^2\}$ is a basis of $\mathbf{P}_2(x)$.

**Example 4.2.18** Let $\mathbf{u} = \{u_1, u_2, u_3\}$ be any vector in $\mathbb{R}^3$. Show that the equation $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ has a **unique solution** for the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$.

**Theorem 4.2.8** (Uniqueness of Basis Representation)
If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for a vector space $V$, then every vector in $V$ can be written in one and only one way as a linear combination of vectors in $S$.

**Theorem 4.2.9** (Bases and Linear Dependence)
If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for a vector space $V$, then every set containing more than $n$ vectors in $V$ is linearly dependent. That is, for any $m > n$, $W = \{u_1, u_2, \dots, u_m\}$ is linear dependent.

The theorem will lead to an important concepts in linear algebra.

Proof. Suppose that

$$x_1 u_1 + x_2 u_2 + \dots + x_m u_m = \mathbf{0}. \tag{4.1}$$

We are going to show that the vector equation (4.1) has a nontrivial solution. Since $S$ is a basis of $V$ and $W \subseteq V$, we have

$$\begin{aligned} u_1 &= a_{11} v_1 + a_{21} v_2 + \dots + a_{n1} v_n, \\ u_2 &= a_{12} v_1 + a_{22} v_2 + \dots + a_{n2} v_n, \\ &\vdots \\ u_m &= a_{1m} v_1 + a_{2m} v_2 + \dots + a_{nm} v_n. \end{aligned}$$

Then sub them into the vector equation (4.1), and we have

$$\begin{aligned} &x_1(a_{11} v_1 + a_{21} v_2 + \dots + a_{n1} v_n) \\ &+x_2(a_{12} v_1 + a_{22} v_2 + \dots + a_{n2} v_n) \\ &+\dots \\ &+x_m(a_{1m} v_1 + a_{2m} v_2 + \dots + a_{nm} v_n) = \mathbf{0}, \end{aligned}$$

which implies that

$$\begin{aligned} &(a_{11} x_1 + a_{12} x_2 + \dots + a_{1m} x_m) v_1 \\ &+(a_{21} x_1 + a_{22} x_2 + \dots + a_{2m} x_m) v_2 \\ &+\dots \\ &+(a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nm} x_m) v_n = \mathbf{0}. \end{aligned}$$

Since $S$ is a basis, $S$ is linear independent, thus we have

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \dots + a_{1m} x_m &= 0 \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2m} x_m &= 0 \\ &\vdots \\ a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nm} x_m &= 0 \end{aligned} \tag{4.2}$$

This homogeneous linear system (4.2) has nontrivial solution, which implies that the vector equation (4.1) has nontrivial solution. □

**Theorem 4.2.10** (Number of Vectors in a Basis) If a vector space $V$ has one basis with $n$ vectors, then every basis for $V$ has $n$ vectors.

**Definition 4.2.7** (Dimension of a Vector Space) If a vector space $V$ has a basis consisting of $n$ vectors, then the number $n$ is the dimension of $V$, denoted by $\dim(V) = n$. When $V$ consists of the zero vector alone, the dimension of $V$ is defined as zero.

1. The dimension of $\mathbb{R}^n$ with the standard operations is $n$.
2. The dimension of $P_n$ with the standard operations is $n + 1$.
3. The dimension of $M_{m,n}$ with the standard operations is $mn$.

**Example 4.2.19** Find the dimension of each subspace of $\mathbb{R}^3$.

1. $W = \{(d, c - d, c) : c$ and $d$ are real numbers $\}$
2. $W = \{(2b, b, 0) : b$ is a real number $\}$.

**Exercise 4.2.4** Find the dimension of the subspace $W$ of $\mathbb{R}^4$ spanned by

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(-1, 2, 5, 0), (3, 0, 1, -2), (-5, 4, 9, 2)\}.$$

**Example 4.2.20** Let $W$ be the subspace of all symmetric matrices in $M_{2,2}$. What is the dimension of $W$?

**Theorem 4.2.11** (Basis Tests in an $n$-Dimensional Space) Let $V$ be a vector space of dimension $n$.

1. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a linearly independent set of vectors in $V$, then $S$ is a basis for $V$.
2. If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ spans $V$, then $S$ is a basis for $V$.

Proof. 1. To show $S$ is a basis of $V$, we only need to show that $S$ spans $V$. $\forall v \in V$ and let $S' = \{v_1, v_2, \dots, v_n, v\}$. It follows from Theorem ?? that $S'$ is linear dependent. Suppose that

$$x_1 v_1 + x_2 v_2 + \dots + x_n v_n + x_{n+1} v_{n+1} = 0. \tag{4.3}$$

We claim that $x_{n+1} \neq 0$. Suppose on the contrary that $x_{n+1} = 0$. The vector equation above becomes

$$x_1 v_1 + x_2 v_2 + \ldots + x_n v_n = 0.$$

Since $\{v_1, v_2, \ldots, v_n\}$ is linear independent, we have

$$x_1 = x_2 = \ldots = x_n = 0.$$

Thus the vector equation 4.3 has only trivial solution. Thus $S'$ is linear independent, which contradicting to the established fact that $S'$ is linear dependent. So $x_{n+1} \neq 0$. Then

$$v = -\frac{x_1}{x_{n+1}} v_1 - \frac{x_2}{x_{n+1}} v_2 - \ldots - \frac{x_n}{x_{n+1}} v_n$$

That is, $S$ spans $V$. Therefore $S$ is a basis of $V$.

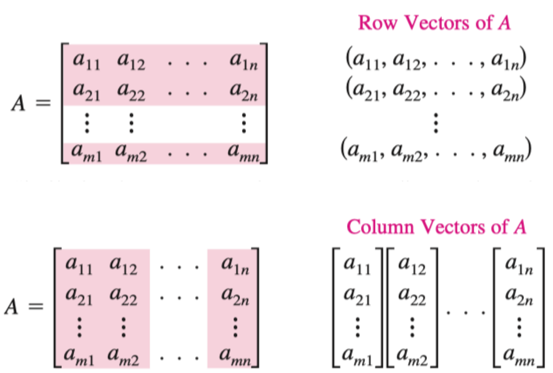
2. Let $S'$ be obtained in the following way.

□

### Project

Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. A subset $B'$ is called a maximal linear independent set in the sense that $\forall u \in B \backslash B'$, $B'' = B' \bigcup \{u\}$ is a linear dependent set. Write an algorithm to find a maximal linear independent subset $B'$ of $B$.

## 4.3 Row and Column spaces of a matrix $A$




**Example:**

For the matrix $A = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 3 & 4 \end{bmatrix}$, the row vectors are $(0, 1, -1)$ and $(-2, 3, 4)$, and the column vectors are $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, and $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$.

**Definition 4.3.1** (Row Space and Column Space of a Matrix)
Let $A$ be an $m \times n$ matrix.

1. The row space of $A$ is the subspace of $\mathbb{R}^n$ spanned by the row vectors of $A$.
2. The column space of $A$ is the subspace of $\mathbb{R}^m$ spanned by the column vectors of $A$.

**Example:** Let $A = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 3 & 4 \end{bmatrix}$.

- The row space is the set $\text{span}\{(0, 1, -1), (-2, 3, 4)\}$.
- The column space is the set $\text{span}\left\{\begin{bmatrix} 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \end{bmatrix}\right\}$.

Recall that span(S) is a subspace. The row or column space is a subspace of $\mathbb{R}^n$, thus it is important to find a basis for the subspaces.

**Theorem 4.3.1** (Basis for the Row Space of a Matrix)
If a matrix $A$ is row-equivalent to a matrix $B$ in row-echelon form, then the nonzero row vectors of $B$ form a basis for the row space of $A$.

**Theorem 4.3.2** (Basis for the Column Space of a Matrix)
the pivot columns of a matrix $A$ form a basis for the column space of $A$.

**Example:** Find a basis for the row space and column space of the matrix $A$ below, respectively.

$$A = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & -2 \end{bmatrix}, \quad \text{RREF}(A) = \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Theorem 4.3.3** The row space and column space of an $m \times n$ matrix $A$ have the same dimension.

**Definition 4.3.2** (Rank of a Matrix)
The dimension of the row (or column) space of a matrix $A$ is the rank of $A$ and is denoted by $\text{rank}(A)$.

### Null space

**Definition: Nullspace and nullity**

If $A$ is an $m \times n$ matrix, then the set of all solutions of the homogeneous system of linear equations $A\mathbf{x} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$ called the **nullspace** of $A$ and is denoted by $N(A)$. So,

$$N(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

The dimension of the nullspace of $A$ is the **nullity** of $A$.

**Example:** Find the nullspace of the matrix.

$$A = \begin{bmatrix} 1 & 2 & -2 & 1 \\ 3 & 6 & -5 & 4 \\ 1 & 2 & 0 & 3 \end{bmatrix}, \quad \text{RREF}(A) = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Theorem 4.3.4** Let $A$ be an $m \times n$ matrix. Then

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

**Theorem 4.3.5** (Solutions of a onhomogeneous linear system)
If $\mathbf{x}_p$ is a particular solution of the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$, then every solution of this system can be written in the form

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h,$$

where $\mathbf{x}_h$ is a solution of the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$.

Prove that

- The set of all solution vectors of the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} \neq \mathbf{0}$, is not a subspace
- Suppose that both $\mathbf{x}_p$ and $\mathbf{y}_p$ are two particular solution of the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$, then $\mathbf{x}_p - \mathbf{y}_p$ is a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$.
- Show that $\mathbf{y}_p = \mathbf{x}_p + \mathbf{x}_h$, where $\mathbf{x}_h$ is a solution of the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$.

## 4.4 From Abstract to concrete

### Coordinate vector

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \Longleftrightarrow \mathbf{x} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$$

**Definition 4.4.1** (coordinate vector)
Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an **ordered** basis for a vector space $V$ and let $\mathbf{x}$ be a vector in $V$ such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

The scalars $c_1, c_2, \dots, c_n$ are the coordinates of $\mathbf{x}$ relative to the basis $B$. The **coordinate vector** of $\mathbf{x}$ relative to $B$ is the column vector in $\mathbb{R}^n$ whose components are the coordinates of $\mathbf{x}$, denoted by

$$[\mathbf{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

**Example 4.4.1** Find the coordinate vector of $\mathbf{x} = (1, 2, -1)$ in $\mathbb{R}^3$ relative to the standard basis $B$ and a (nonstandard) basis $B' = \{(1, 0, 1), (0, -1, 2), (2, 3, -5)\}$, respectively.

**Example 4.4.2**
Let $B = \{(9, -3, 15, 4), (3, 0, 0, 1), (0, -5, 6, 8), (3, -4, 2, -3)\}$ be a basis of $\mathbb{R}^4$, $\mathbf{x} = (0, -20, 7, 15)$ and $\mathbf{y} = (15, -19, 23, -10)$. Find $[\mathbf{x}]_B$, $[\mathbf{y}]_B$ and $[\mathbf{x} + \mathbf{y}]_B$.

Hint: You may use the following fact.

$$\text{RREF}\left(\begin{bmatrix} 9 & 3 & 0 & 3 & 15 & 0 \\ -3 & 0 & -5 & -4 & -19 & -20 \\ 15 & 0 & 6 & 2 & 23 & 7 \\ 4 & 1 & 8 & -3 & -10 & 15 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 & 2 \end{bmatrix}$$

**Theorem 4.4.1** Let $B$ be a basis of a vector space $V$, $\mathbf{x}, \mathbf{y} \in V$ and $k \in \mathbb{R}$. Then

1. $[\mathbf{x} + \mathbf{y}]_B = [\mathbf{x}]_B + [\mathbf{y}]_B$.
2. $[k\mathbf{x}]_B = k([\mathbf{x}]_B)$

**Theorem 4.4.2** Let $S = \{v_1, v_2, \dots, v_m\}$ be a subset of $V$ and $B$ be a basis of a vector space $V$. Show that
$S$ is linear independent iff $\{[v_1]_B, \dots, [v_m]_B\}$ is linear independent.

**Example 4.4.3** Let $\mathbf{x} = -20x + 7x^2 + 15x^3$, $\mathbf{y} = 15 - 19x + 23x^2 -$

$10x^3$ and $B = \{9 - 3x + 15x^2 + 4x^3, 3 + x^3, -5x + 6x^2 + 8x^3, 3 - 4x + 2x^2 - 3x^3\}$.

1. Show that $B$ is a basis of $P_3(x)$.
2. Compute $[\mathbf{x}]_B$ and $[\mathbf{y}]_B$.

**Example 4.4.4** Let $\mathbf{x} = \begin{bmatrix} 0 & -20 \\ 7 & 15 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 15 & -19 \\ 23 & -10 \end{bmatrix}$ and $B = \left\{ \begin{bmatrix} 9 & -3 \\ 15 & 4 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -5 \\ 6 & 8 \end{bmatrix}, \begin{bmatrix} 3 & -4 \\ 2 & -3 \end{bmatrix} \right\}$.

1. Show that $B$ is a basis of $M_{22}(\mathbb{R})$.
2. Compute $[\mathbf{x}]_B$ and $[\mathbf{y}]_B$.

It follows from the previous three examples, we know that vector can disguise in different ways. However, One can understand any abstract vector with respect to a basis in the sense of $\mathbb{R}^n$.

## Transition Matrix

**Lemma 4.4.3** Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $B' = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ be two bases for a vector space $V$. If

$$\begin{aligned} \mathbf{v}_1 &= c_{11}\mathbf{u}_1 + c_{21}\mathbf{u}_2 + \cdots + c_{n1}\mathbf{u}_n \\ \mathbf{v}_2 &= c_{12}\mathbf{u}_1 + c_{22}\mathbf{u}_2 + \cdots + c_{n2}\mathbf{u}_n \\ &\vdots \\ \mathbf{v}_n &= c_{1n}\mathbf{u}_1 + c_{2n}\mathbf{u}_2 + \cdots + c_{nn}\mathbf{u}_n \end{aligned}$$

then the **transition matrix** from $B$ to $B'$ is defined as

$$Q = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} = [[\mathbf{v}_1]_{B'} \ [\mathbf{v}_2]_{B'} \ \cdots \ [\mathbf{v}_n]_{B'}]$$

How to find the the transition matrix from $B$ to $B'$?

1. How to find $[\mathbf{v}_1]_{B'}$?

2. How to find $[\mathbf{v}_1]_{B'}, [\mathbf{v}_2]_{B'}, \cdots, [\mathbf{v}_n]_{B'}$ together?

3. Let $S = [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \dots \ \ \mathbf{u}_n]$ and $T = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \dots \ \ \mathbf{v}_n]$. Express $Q$ by $S$ and $T$.

Exercise: Find the the transition matrix from $B'$ to $B$?

**Example 4.4.5** Find the transition matrix from $B$ to $B'$ for the bases for $\mathbb{R}^2$ below.

$$B = \{(-3, 2), (4, -2)\} \quad \text{and} \quad B' = \{(-1, 2), (2, -2)\}$$

**Theorem 4.4.4** (Change of basis)
Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be two bases for a vector space $V$, and $Q$ is the transition matrix from $B$ to $B'$. For any vector $\mathbf{x} \in V$,

$$[\mathbf{x}]_{B'} = Q[\mathbf{x}]_B$$

# Part 2: Second Semester

# 5 Inner product space