

Vectors and Matrix (Elementary)

HU-HTAKM

July 10, 2024

This is just all the knowledge about vectors and matrices that you may have learnt in high school mathematics.

Chapter 1

Introduction of vectors

In daily life, many quantities are described by **magnitude** alone, such as length, time and temperature. These are called **scalars**. However, some quantities are described by both **magnitude** and **direction**. For example, the acceleration of an object. For simplicity, we visualize vectors in two dimension.

Definition 1.1. A **vector** is represented by a directed line segment. Arrowhead represents the direction while the length represents the magnitude. Vector is denoted by \overrightarrow{AB} (A to B), \mathbf{a} , \vec{a} . Magnitude (length) of a vector \mathbf{a} is denoted by $|\vec{a}|$. We say that two vectors \vec{a} and \vec{b} are equal if they have the same magnitude and direction.

In 2D, it is represented as following:

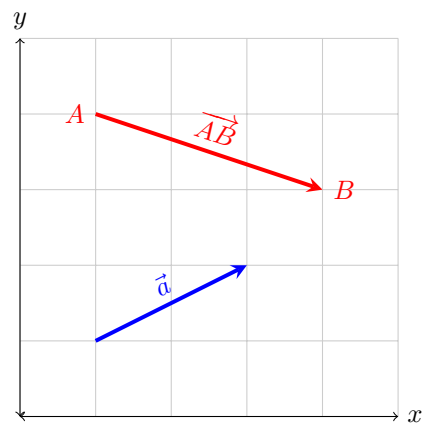


Figure 1.1: Vector

Definition 1.2. **Negative vector** of a vector \vec{u} is the vector with equal magnitude but opposite direction. It is denoted by $-\vec{u}$.

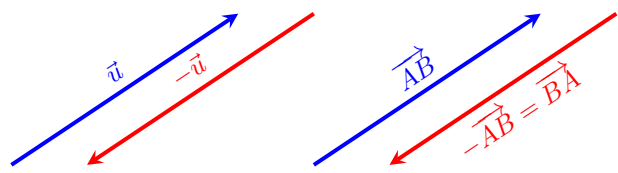


Figure 1.2: Vector and its corresponding negative vector

Definition 1.3. **Zero vector** is a vector with zero magnitude and no direction. It is denoted by $\vec{0}$.

Definition 1.4. **Unit vector** is a vector with magnitude 1. It is denoted by \hat{u} .

Lemma 1.5. $\hat{u} = \frac{\vec{u}}{|\vec{u}|}$ is unit vector with the same direction as \vec{u} .

Example 1.1. The most common unit vectors are the unit vectors along the axis.

\hat{i} corresponds to unit vector along the x -axis.

\hat{j} corresponds to unit vector along the y -axis.

\hat{k} corresponds to unit vector along the z -axis.

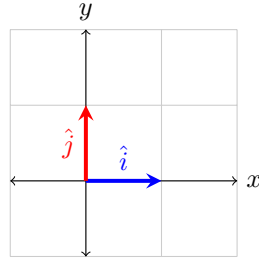
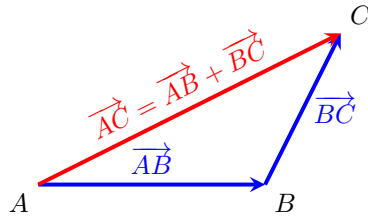
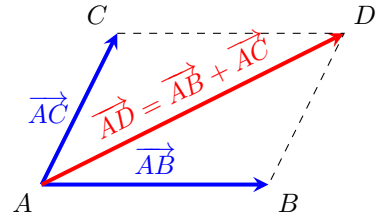


Figure 1.3: Unit vectors along the x -axis and y -axis

Definition 1.6. Addition of two vectors are also a vector and is defined as below.



(a) Triangle Law of Addition



(b) Parallelogram Law of Addition

Definition 1.7. Subtraction of two vectors \vec{u} and \vec{v} is $\vec{u} - \vec{v} = \vec{u} + (-\vec{v})$.

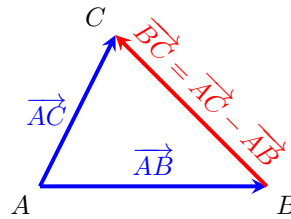


Figure 1.5: Subtraction

Definition 1.8. Scalar multiplication of a vector \vec{u} is defined by:

1. If $\lambda > 0$, then $\lambda\vec{u}$ is a vector with magnitude $\lambda|\vec{u}|$ and same direction with \vec{u} .
2. If $\lambda < 0$, then $\lambda\vec{u}$ is a vector with magnitude $-\lambda|\vec{u}|$ and opposite direction with \vec{u} .
3. If $\lambda = 0$, then $\lambda\vec{u}$ is a zero vector.

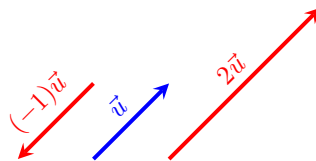


Figure 1.6: Scalar multiplication of vector

Lemma 1.9. For any vectors \vec{u} , \vec{v} , \vec{w} and any real number λ , μ ,

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
3. $\vec{u} + \vec{0} = \vec{u}$
4. $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$
5. $\lambda(\vec{u} \pm \vec{v}) = \lambda\vec{u} \pm \lambda\vec{v}$
6. $(\lambda \pm \mu)\vec{u} = \lambda\vec{u} \pm \mu\vec{u}$

Theorem 1.10. If \vec{u} and \vec{v} are non-zero and not parallel, then

1. If $\lambda\vec{u} + \mu\vec{v} = \vec{0}$, then $\lambda = \mu = 0$
2. If $\lambda_1\vec{u} + \mu_1\vec{v} = \lambda_2\vec{u} + \mu_2\vec{v}$, then $\lambda_1 = \lambda_2$ and $\mu_1 = \mu_2$.

Remark 1.10.1. Knowing the properties, we can find that we can represent any vectors in 2D space in terms of \hat{i} and \hat{j} , and any vectors in 3D space in terms of \hat{i} , \hat{j} and \hat{k} .

We have now finished introducing the basics of vectors. We can use the vectors to do all sort of stuff.

Definition 1.11. **Position vector** is a vector that starts at origin O .

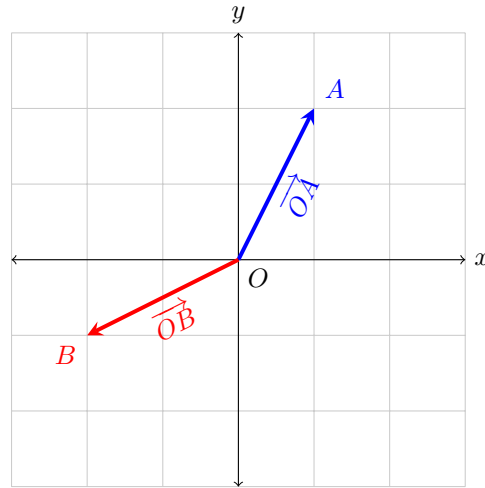


Figure 1.7: Position Vectors

Theorem 1.12. (Section formula) If P is a point of AB such that $AP : PB = r : s$, then $\vec{p} = \frac{s\vec{a} + r\vec{b}}{s+r}$ where \vec{a} , \vec{b} and \vec{p} are the position vector of A , B and P respectively.

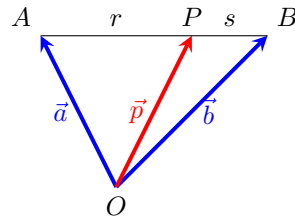


Figure 1.8: Section formula

Chapter 2

Introduction to matrix

In real life, we may process data using a rectangular array of real numbers.

Definition 2.1. **Matrix** is a rectangular array of real numbers arranged in form:

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

The **dimension** of this matrix is $m \times n$. Each number inside the matrix is the **elements**. It is denoted by **A**.

In some advanced mathematics, we use vectors to define matrix as in form

$$\begin{pmatrix} | & | & \dots & | \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \\ | & | & & | \end{pmatrix}$$

where $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ are m -dimensional vectors.

Definition 2.2. **Zero matrix** is a matrix where all the elements are zero.

$$\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

A $m \times n$ zero matrix is denoted as $\mathbf{0}_{m \times n}$ or $\mathbf{0}$.

Definition 2.3. **Row matrix** is a matrix with only 1 row. **Column matrix** is a matrix with only 1 column.

Example 2.1. $\begin{pmatrix} 1 & 2 \end{pmatrix}$ is a row matrix.

Example 2.2. $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ is a column matrix.

Definition 2.4. **Square matrix** with order n is a matrix with dimension $n \times n$. The elements $a_{11}, a_{22}, \dots, a_{nn}$ are the principal diagonal elements.

Definition 2.5. **Diagonal matrix** is a square matrix where elements that are not principal diagonal elements are zero.

Example 2.3. $\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$ is a diagonal matrix.

Definition 2.6. **Identity matrix** is a diagonal matrix where principal diagonal elements are all 1.

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

A $n \times n$ identity matrix is denoted as \mathbf{I}_n or \mathbf{I} .

Definition 2.7. For two $m \times n$ matrices \mathbf{A} with elements a_{ij} and \mathbf{B} with elements b_{ij} , addition and subtraction of \mathbf{A} and \mathbf{B} is defined by

$$\mathbf{A} \pm \mathbf{B} = \begin{pmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & \dots & a_{1n} \pm b_{1n} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & \dots & a_{2n} \pm b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} \pm b_{m1} & a_{m2} \pm b_{m2} & \dots & a_{mn} \pm b_{mn} \end{pmatrix}$$

Definition 2.8. For a $m \times n$ matrix \mathbf{A} with elements a_{ij} and a scalar k , scalar multiplication is defined by:

$$k\mathbf{A} = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}$$

Lemma 2.9. For $m \times n$ matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and scalars λ, μ ,

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
3. $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$
4. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$
5. $(\lambda \pm \mu)\mathbf{A} = \lambda\mathbf{A} \pm \mu\mathbf{A}$
6. $\lambda(\mathbf{A} \pm \mathbf{B}) = \lambda\mathbf{A} \pm \lambda\mathbf{B}$
7. $\lambda(\mu\mathbf{A}) = (\lambda\mu)\mathbf{A}$
8. $0\mathbf{A} = \mathbf{0}$
9. $\lambda\mathbf{0} = \mathbf{0}$

There is an important scalar that can be computed from square matrices. It is called determinant.

Definition 2.10. For any square matrix \mathbf{A} , **determinant** of \mathbf{A} is a real number that is denoted by $|\mathbf{A}|$ or $\det \mathbf{A}$.

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

How do we calculate determinant?

Definition 2.11. For a 2×2 matrix,

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

Definition 2.12. For a 3×3 matrix,

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

Definition 2.13. Given a matrix \mathbf{A} with elements a_{ij} . **Minor** of the element a_{ij} , denoted by M_{ij} , is a determinant of order $(n-1)$ which is obtained by removing the i -th row and j -th column of $|\mathbf{A}|$.

Cofactor of the element a_{ij} , denoted by A_{ij} , is defined by $A_{ij} = (-1)^{i+j}M_{ij}$.

Example 2.4. Given a matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

$$M_{12} = \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix}$$

$$M_{33} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}$$

$$A_{12} = - \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix}$$

$$A_{33} = + \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}$$

Lemma 2.14. For any identity matrix \mathbf{I} , $|\mathbf{I}| = 1$.

Theorem 2.15. (Cofactor expansion) We can expand the determinant using cofactors along i -th row or j -th column.

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj}$$

It is sometimes very tedious to calculate determinant directly. Luckily, we can perform row and column operations to determinant so that we can simplify the process.

Definition 2.16. By definition, we can perform addition and scalar multiplication to determinant.

$$\begin{aligned} & \begin{vmatrix} a_{11} & a_{12} + b_1 & \dots & a_{1n} \\ a_{21} & a_{22} + b_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} + b_n & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & b_1 & \dots & a_{1n} \\ a_{21} & b_2 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{vmatrix} \\ & \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} + b_1 & a_{22} + b_2 & \dots & a_{2n} + b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ b_1 & b_2 & \dots & b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\ & \begin{vmatrix} a_{11} & \lambda a_{12} & \dots & a_{1n} \\ a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \lambda a_{n2} & \dots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \\ & \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \mu a_{21} & \mu a_{22} & \dots & \mu a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \mu \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \end{aligned}$$

We can also change the order.

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

Lemma 2.17. We can add multiples of rows and columns into another rows and columns respectively.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} - \lambda a_{1j} & a_{12} & \cdots & a_{1n} \\ a_{21} - \lambda a_{2j} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} - \lambda a_{nj} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} - \mu a_{i1} & a_{12} - \mu a_{i2} & \cdots & a_{1n} - \mu a_{in} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Example 2.5.

$$\begin{vmatrix} 2 & 5 & 10 \\ 10 & 3 & 4 \\ 2 & 5 & 10 \end{vmatrix} = \begin{vmatrix} 2 & 5 & 10 \\ 10 & 3 & 4 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

Chapter 3

Product of vectors and matrices

When we calculate the products of vectors and matrices, it is quite different with our scalar multiplication. In this chapter, we will purely discuss them and their applications. Let's start with vector.

Definition 3.1. Given two vectors \vec{u} and \vec{v} . **Dot product** (Scalar product) of \vec{u} and \vec{v} is defined by:

$$\vec{u} \cdot \vec{v} = |\vec{u}| |\vec{v}| \cos \theta$$

where θ is the angle between two vectors.

Lemma 3.2. For any vectors $\vec{u}, \vec{v}, \vec{w}$ and any scalar λ ,

1. $\vec{u} \cdot \vec{0} = \vec{0} \cdot \vec{u} = 0$
2. $\vec{u} \cdot \vec{u} = |\vec{u}|^2 \geq 0$
3. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
4. $\vec{u} \cdot \vec{u} = 0$ if and only if $\vec{u} = \vec{0}$
5. $\lambda(\vec{u} \cdot \vec{v}) = (\lambda\vec{u}) \cdot \vec{v} = \vec{u} \cdot (\lambda\vec{v})$
6. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
7. $|\vec{u} \cdot \vec{v}| \leq |\vec{u}| |\vec{v}|$
8. $|\vec{u} - \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 - 2(\vec{u} \cdot \vec{v})$

Theorem 3.3. For any two non-zero vectors \vec{u} and \vec{v} , $\vec{u} \cdot \vec{v} = 0$ if and only if $\vec{u} \perp \vec{v}$.

Lemma 3.4. In the 3D coordinate system,

$$\hat{i} \cdot \hat{i} = \hat{j} \cdot \hat{j} = \hat{k} \cdot \hat{k} = 1 \qquad \hat{i} \cdot \hat{j} = \hat{j} \cdot \hat{k} = \hat{k} \cdot \hat{i} = 0$$

Given two vectors $\vec{u} = u_1\hat{i} + u_2\hat{j} + u_3\hat{k}$ and $\vec{v} = v_1\hat{i} + v_2\hat{j} + v_3\hat{k}$. We can define dot product of \vec{u} and \vec{v} by:

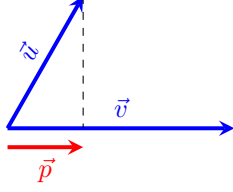
$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + u_3v_3$$

Interestingly, from the definition, we can project a vector into another vector.

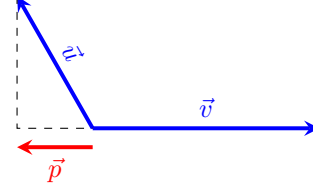
Definition 3.5. Given two non-zero vectors \vec{u} and \vec{v} . **Projection** of \vec{u} onto \vec{v} is a vector \vec{p} defined by:

$$\vec{p} = \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v}$$

Magnitude of \vec{p} is $|\vec{p}| = \frac{|\vec{u} \cdot \vec{v}|}{|\vec{v}|}$.



(a) Projection with equal direction



(b) Projection with opposite direction

We have another type of product for vectors.

Definition 3.6. Given two vectors \vec{u} and \vec{v} . **Cross product** (Vector product) of \vec{u} and \vec{v} is defined by:

$$\vec{u} \times \vec{v} = (|\vec{u}| |\vec{v}| \sin \theta) \hat{n}$$

where θ is the angle between \vec{u} and \vec{v} , and \hat{n} is the unit vector perpendicular to both \vec{u} and \vec{v} and its direction is determined by right-hand rule.

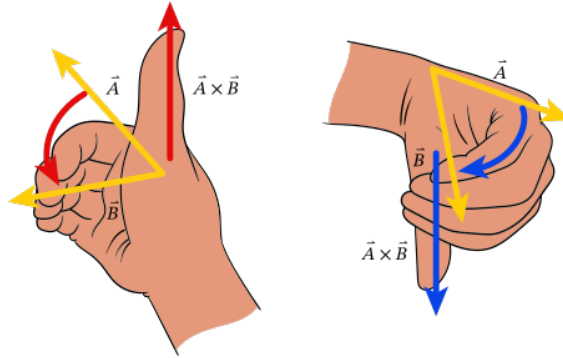


Figure 3.2: Right Hand Rule

Lemma 3.7. For any vectors $\vec{u}, \vec{v}, \vec{w}$ and any scalar λ ,

1. $\vec{u} \times \vec{0} = \vec{0} \times \vec{u} = \vec{0}$
2. $\vec{u} \times \vec{u} = \vec{0}$
3. $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$
4. $(\lambda \vec{u}) \times \vec{v} = \vec{u} \times (\lambda \vec{v}) = \lambda(\vec{u} \times \vec{v})$
5. $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
6. $(\vec{u} + \vec{v}) \times \vec{w} = \vec{u} \times \vec{w} + \vec{v} \times \vec{w}$
7. $|\vec{u} \times \vec{v}|^2 = |\vec{u}|^2 |\vec{v}|^2 - (\vec{u} \cdot \vec{v})^2$

We can make use of the determinant learnt to calculate the cross product.

Lemma 3.8. In the 3D coordinate system,

$$\hat{i} \times \hat{i} = \hat{j} \times \hat{j} = \hat{k} \times \hat{k} = 0$$

$$\hat{i} \times \hat{j} = \hat{k}$$

$$\hat{j} \times \hat{k} = \hat{i}$$

$$\hat{k} \times \hat{i} = \hat{j}$$

Given two vectors $\vec{u} = u_1\hat{i} + u_2\hat{j} + u_3\hat{k}$ and $\vec{v} = v_1\hat{i} + v_2\hat{j} + v_3\hat{k}$. We can define cross product of \vec{u} and \vec{v} by:

$$\vec{u} \times \vec{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Now that we have introduced both products about vectors. What are the applications? The most simple one is to calculate the area and volume.

Theorem 3.9. Area of the parallelogram formed by two vectors \vec{u} and \vec{v} is $|\vec{u} \times \vec{v}|$.

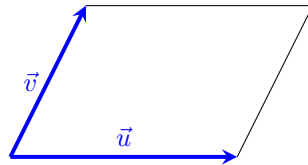


Figure 3.3: Parallelogram formed by two vectors

Lemma 3.10. Area of the triangle formed by two vectors \vec{u} and \vec{v} is $\frac{1}{2} |\vec{u} \times \vec{v}|$.

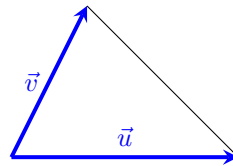


Figure 3.4: Triangle formed by two vectors

Theorem 3.11. Volume of the parallelepiped formed by three vectors \vec{u} , \vec{v} and \vec{w} is $|\vec{u} \cdot (\vec{v} \times \vec{w})|$.

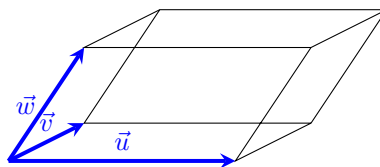


Figure 3.5: Parallelepiped formed by three vectors

Lemma 3.12. Volume of the tetrahedron formed by three vectors \vec{u} , \vec{v} and \vec{w} is $\frac{1}{6} |\vec{u} \cdot (\vec{v} \times \vec{w})|$.

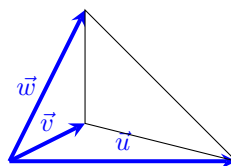


Figure 3.6: Tetrahedron formed by three vectors

We can also use cross product to determine if 3 points are collinear.

Theorem 3.13. Given three points A , B and C . They are collinear if and only if $\overrightarrow{AB} \times \overrightarrow{AC} = \vec{0}$.



What about multiplying matrices?

Definition 3.14. Given an $m \times n$ matrix \mathbf{A} with elements a_{ij} and an $n \times p$ matrix \mathbf{B} with elements b_{ij} . The product of \mathbf{A} and \mathbf{B} is an $m \times p$ matrix $\mathbf{C} = \mathbf{AB}$ with each elements c_{ij} obtained by:

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

Example 3.1. Given that $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}$. Then

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix} \quad \mathbf{BA} = \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 39 & 54 & 69 \\ 49 & 68 & 87 \\ 59 & 82 & 105 \end{pmatrix}$$

Lemma 3.15. For any scalars λ and μ and matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ that allows for matrix multiplications between them,

1. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
2. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
3. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
4. $(\lambda\mathbf{A})(\mu\mathbf{B}) = \lambda\mu(\mathbf{AB})$
5. $\mathbf{0A} = \mathbf{A0} = \mathbf{0}$
6. $\mathbf{I}_m\mathbf{A} = \mathbf{AI}_n = \mathbf{A}$ if \mathbf{A} is an $m \times n$ matrix
7. $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$

Definition 3.16. Let \mathbf{A} be a square matrix and n be a positive integer. Then

$$\mathbf{A}^n = \prod_{i=1}^n \mathbf{A}$$

Some of the matrices have its multiplicative inverse, just like scalar.

Definition 3.17. Given an $n \times n$ matrix \mathbf{A} . It is **invertible** (non-singular) if there exists another $n \times n$ matrix \mathbf{B} such that

$$\mathbf{AB} = \mathbf{I}_n$$

\mathbf{B} is the **inverse matrix** of \mathbf{A} , and is denoted by \mathbf{A}^{-1} .

Theorem 3.18. For any square matrices \mathbf{A} and \mathbf{B} , if $\mathbf{AB} = \mathbf{I}$, then $\mathbf{BA} = \mathbf{I}$ and \mathbf{B} is the inverse of \mathbf{A} .

Lemma 3.19. For any invertible matrix \mathbf{A} , $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$.

Theorem 3.20. (Uniqueness) Let \mathbf{A} be an invertible matrix. If both \mathbf{B} and \mathbf{C} are inverses of \mathbf{A} , then $\mathbf{B} = \mathbf{C}$.

Theorem 3.21. Let \mathbf{A} and \mathbf{B} be two invertible $n \times n$ matrices, and λ be a non-zero scalar.

1. \mathbf{A}^{-1} is invertible and $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
2. $\lambda\mathbf{A}$ is invertible and $(\lambda\mathbf{A})^{-1} = \frac{1}{\lambda}\mathbf{A}^{-1}$
3. \mathbf{AB} is invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Lemma 3.22. If \mathbf{A} is invertible and k is a positive integer, then \mathbf{A}^k is invertible and $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$.

Based on definition, it seems we may need to find the inverse by brute force. Luckily, there is a way to easily find the inverse. Before that, we should introduce the transpose.

Definition 3.23. Given a $m \times n$ matrix \mathbf{A} . **Transpose** of matrix \mathbf{A} , denoted by \mathbf{A}^T , is a $n \times m$ matrix obtained by interchanging the rows and columns of \mathbf{A} .

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

Lemma 3.24. Let λ be a scalar, and matrices \mathbf{A}, \mathbf{B} that allows for matrix multiplications.

1. $\mathbf{I}^T = \mathbf{I}$
2. $(\mathbf{A}^T)^T = \mathbf{A}$
3. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
4. $(\lambda\mathbf{A})^T = \lambda\mathbf{A}^T$
5. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
6. If \mathbf{A} is invertible, then \mathbf{A}^T is invertible and $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

We can now introduce the adjoint matrix.

Definition 3.25. For any $n \times n$ square matrix \mathbf{A} , **adjoint matrix** is defined by:

$$\text{adj } \mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T$$

where A_{ij} is the cofactor of a_{ij} .

Using this matrix, we can get a matrix that is a multiple of identity matrix using cofactor expansion.

Theorem 3.26. Let \mathbf{A} be a square matrix. Then

$$\mathbf{A}(\text{adj } \mathbf{A}) = (\text{adj } \mathbf{A})\mathbf{A} = \begin{pmatrix} |\mathbf{A}| & 0 & \dots & 0 \\ 0 & |\mathbf{A}| & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & |\mathbf{A}| \end{pmatrix} = |\mathbf{A}|\mathbf{I}$$

As we can see, we have obtained a formula for inverse matrix.

Theorem 3.27. Let \mathbf{A} be a square matrix.

1. If $|\mathbf{A}| \neq 0$, then $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj } \mathbf{A}$
2. \mathbf{A} is invertible if and only if $|\mathbf{A}| \neq 0$.

Example 3.2. Given $\mathbf{A} = \begin{pmatrix} 1 & 1 & 4 \\ 2 & 0 & 3 \\ -1 & 2 & 0 \end{pmatrix}$. We can find that

$$|\mathbf{A}| = \begin{vmatrix} 1 & 1 & 4 \\ 2 & 0 & 3 \\ -1 & 2 & 0 \end{vmatrix} = 7$$

Therefore, we can see that \mathbf{A} is invertible. Then

$$\text{adj } \mathbf{A} = \begin{pmatrix} \begin{vmatrix} 0 & 3 \\ 2 & 0 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ -1 & 0 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ -1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 1 & 4 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ -1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 2 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} \end{pmatrix}^T = \begin{pmatrix} -6 & 8 & 3 \\ -3 & 4 & 5 \\ 4 & -3 & -2 \end{pmatrix} \quad \mathbf{A}^{-1} = \frac{1}{7} \begin{pmatrix} -6 & 8 & 3 \\ -3 & 4 & 5 \\ 4 & -3 & -2 \end{pmatrix} = \begin{pmatrix} -\frac{6}{7} & \frac{8}{7} & \frac{3}{7} \\ -\frac{3}{7} & \frac{4}{7} & \frac{5}{7} \\ \frac{4}{7} & -\frac{3}{7} & -\frac{2}{7} \end{pmatrix}$$

Chapter 4

System of linear equations

One of the very important applications of matrices is solving a system of multiple linear equations.

Definition 4.1. A system of m linear equations in n unknowns is in form:

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

It is **consistent** if it has solutions. It is **inconsistent** if it has no solutions.

For simplicity, we focus on system of n linear equations in n unknown. One of the obvious way is to use invertible matrix.

Theorem 4.2. We can turn the system of linear equations into form:

$$\mathbf{AX} = \mathbf{B} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

If \mathbf{A} is a square matrix, \mathbf{A} is invertible if and only if $\mathbf{AX} = \mathbf{B}$ has a unique solution $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$.

Lemma 4.3. Given a system of n linear equation in n unknowns $\mathbf{AX} = \mathbf{B}$.

1. $\mathbf{AX} = \mathbf{B}$ has a unique solution if and only if $|\mathbf{A}| \neq 0$.
2. $\mathbf{AX} = \mathbf{B}$ has either no solutions or infinitely many solutions if and only if $|\mathbf{A}| = 0$.

It can be tedious to find invertible matrix in order to solve the system of linear equations. We introduce Cramer's Rule, which make uses of determinant.

Theorem 4.4. (Cramer's Rule) For a system of 2 linear equations in 2 unknowns, we let

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases} \quad \Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \Delta_x = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} \quad \Delta_y = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

If $\Delta \neq 0$, then $x = \frac{\Delta_x}{\Delta}$ and $y = \frac{\Delta_y}{\Delta}$ is the unique solution of the system.

For a system of 2 linear equations in 2 unknowns, we let

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases} \quad \Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \Delta_x = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} \quad \Delta_y = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} \quad \Delta_z = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}$$

If $\Delta \neq 0$, then $x = \frac{\Delta_x}{\Delta}$, $y = \frac{\Delta_y}{\Delta}$ and $z = \frac{\Delta_z}{\Delta}$ is the unique solution of the system.

Interestingly, Cremer's Rule can help us determine whether a system of linear equations has no solutions or infinitely many solutions.

Theorem 4.5. For a system of 2 linear equations in 2 unknowns:

1. If $\Delta \neq 0$, then the system has an unique solution.
2. If $\Delta = 0$ and either $\Delta_x \neq 0$ or $\Delta_y \neq 0$, then the system has no solutions.
3. If $\Delta = 0$ and $\Delta_x = \Delta_y = 0$, then the system has infinite many solutions.

For a system of 3 linear equations in 3 unknowns:

1. If $\Delta \neq 0$, then the system has an unique solution.
2. If $\Delta = 0$ and either $\Delta_x \neq 0$, $\Delta_y \neq 0$ or $\Delta_z \neq 0$, then the system has no solutions.
3. If $\Delta = 0$ and $\Delta_x = \Delta_y = \Delta_z = 0$, then the system has infinite many solutions.

We now introduce a third method of solving the system of linear equations, which is Gaussian elimination. Before that, we introduce the row echelon form.

Definition 4.6. An **augmented matrix** is a matrix that is obtained by appending an original matrix.

Example 4.1. Assume that we want to represent

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

We can use the following augmented matrix to represent this system of linear equations.

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array} \right)$$

Remark 4.6.1. What's special about system of linear equations is that we can perform row operations to augmented matrix.

Definition 4.7. An augmented matrix is in **row echelon form** if

1. The first non-zero element in each row is 1.
2. The first non-zero element in each row occurs in a column to the right of the first non-zero element in the preceding row.
3. Any zero rows are placed at the bottom of the matrix.

Example 4.2. Following are some augmented matrices that are in row echelon form.

$$\left(\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 1 & 3 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 6 & 1 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 7 & 6 & 5 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Once we turn an augmented matrix into row echelon form, we can obtain the solution by back substitutions.

Example 4.3. We want to find the solution of this system of linear equations:

$$\begin{cases} x + 3y + 4z = 19 \\ 2x + 5y + 3z = 21 \\ 3x + 7y + 3z = 26 \end{cases}$$

We turn this system into an augmented matrix, then perform row operations.

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 2 & 5 & 3 & 21 \\ 3 & 7 & 3 & 26 \end{array} \right) &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & -1 & -5 & -17 \\ 0 & -2 & -9 & -31 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & 1 & 5 & 17 \\ 0 & -2 & -9 & -31 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & 1 & 5 & 17 \\ 0 & 0 & 1 & 3 \end{array} \right) \end{aligned}$$

From this, we can turn the augmented matrix into linear equations again and get the solution.

$$\begin{cases} x + 3y + 4z = 19 \\ y + 5z = 17 \\ z = 3 \end{cases} \qquad \begin{cases} x = 1 \\ y = 2 \\ z = 3 \end{cases}$$

Example 4.4. We want to find the solution of this system of linear equations:

$$\begin{cases} x + 3y + 4z = 19 \\ 2x + 5y + 3z = 21 \\ 3x + 7y + 2z = 23 \end{cases}$$

We turn this system into an augmented matrix, then perform row operations.

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 2 & 5 & 3 & 21 \\ 3 & 7 & 2 & 23 \end{array} \right) &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & -1 & -5 & -17 \\ 0 & -2 & -10 & -34 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & 1 & 5 & 17 \\ 0 & -2 & -10 & -34 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|c} 1 & 3 & 4 & 19 \\ 0 & 1 & 5 & 17 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

From this, we can turn the augmented matrix into linear equations again and get the solution. Let t be any real number.

$$\begin{cases} x + 3y + 4z = 19 \\ y + 5z = 17 \end{cases} \qquad \begin{cases} x = 11t - 32 \\ y = -5t + 17 \\ z = t \end{cases}$$

There are a special kind of system of linear equations called the system of homogeneous linear equations. In this case, it is guarantee to have solutions.

Definition 4.8. A system of m homogeneous linear equations in n unknowns is in form:

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

Theorem 4.9. For a system of n homogeneous linear equations with n unknowns $\mathbf{AX} = \mathbf{0}$:

1. $\mathbf{AX} = \mathbf{0}$ has trivial solution (all 0) as its only solution if and only if $|\mathbf{A}| \neq 0$.
2. $\mathbf{AX} = \mathbf{0}$ has a trivial solution and infinitely many non-trivial solutions if and only if $|\mathbf{A}| = 0$.

There is a special use for augmented matrix. Note that when we represent system of linear equations in $\mathbf{AX} = \mathbf{B}$,

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

If $|\mathbf{A}| \neq 0$, by performing row operations, we get that

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

We know that $\mathbf{AX} = \mathbf{IB}$ can become $\mathbf{IX} = \mathbf{A}^{-1}\mathbf{B}$. It just so happen that we just obtained the inverse of \mathbf{A} ! Therefore,

Theorem 4.10. Given an invertible matrix \mathbf{A} . By performing row operations, we can get that

$$(\mathbf{A} \mid \mathbf{I}) \sim (\mathbf{I} \mid \mathbf{A}^{-1})$$