Linear Algebra Notes

January 2024

Chapter 1

Matrices and Gaussian Elimination

1.1 introduction

1.1.1 Linear Equations

This section begins with the central problem of linear algebra solving linear equations.The most important ease,and the simplest,is when the number of unknowns equals the number of equations.We have n equations in n unknowns

Example 1.1. *Solving equations*
We illustrate two methods to solve:

$$x + 2y = 7$$

$$3x + y = 6$$

1.Elimination
2.Determinants (Cramers Rule)

Both has its own benefit:

Elimination- better efficiency for large system
Determinants- explicit formula

Definition 1.1. *linear system*
A linear equation in n unknowns is an equation of the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

where a1,...a_n and b are real numbers. $x_1...x_n$are called the variables or unknowns.

A system of m equations in n unknowns (or mn system) is of the form

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1, \tag{1.1}$$

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2, \tag{1.2}$$

$$... \tag{1.3}$$

$$a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n, \tag{1.4}$$

1.1.2 Solution, Consistency, Solution Set

Definition 1.2. *Solution set*

By a solution of an mn system (2), we mean an ordered n-tuple of numbers $(x_1, x_2, , x_n)$ that satisfies all the equations of the system. The set of all solution is called the solution set

In our previous example (1), the solution set is (1,3)

Definition 1.3. *Consistency*

If a linear system has at least one solution, it is called consistent. Otherwise, it is called inconsistent.

Example 1.2.

$$x + 2y = 3$$

$$4x + 8y = 6$$

is inconsistent because there is no solution.

Remark 1. *There are three scenario:*

1.unique solution: consistent

2.infinitely many solutions: consistent (singular case)

3.no solution: inconsistent (singular case)

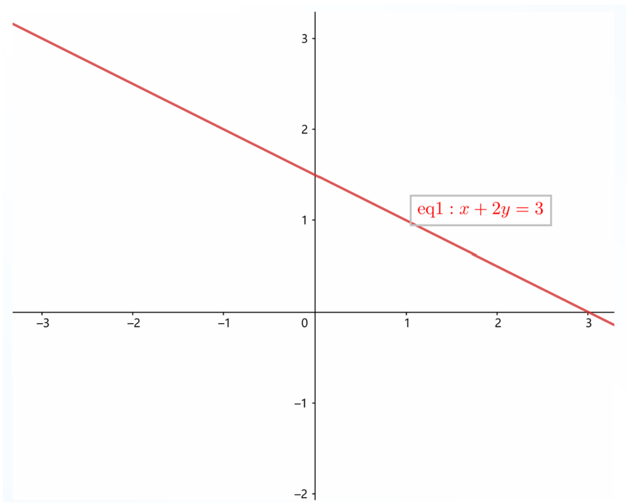
1.2 The Geometry of Linear Equations

1.2.1 Row Picture

Example 1.3. *A single linear equation*

$$ax + by = c$$

cut out a hyperplane (a line) in $\mathbb{R}^2$. e.g

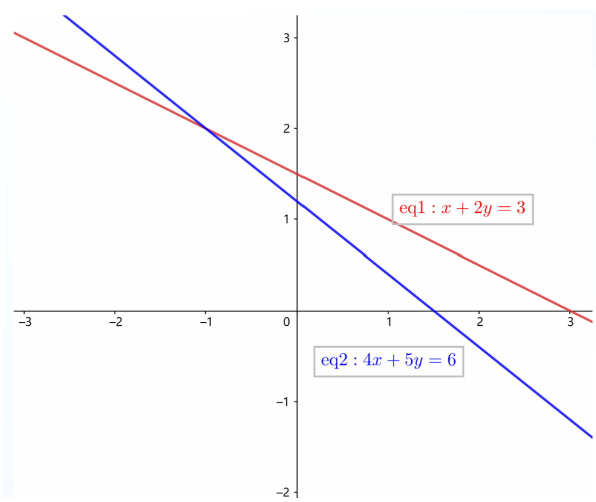


This is the case with infinitely many solutions.

Example 1.4. *A system of equations cut out the intersections of hyperplanes associate to each equation e.g*

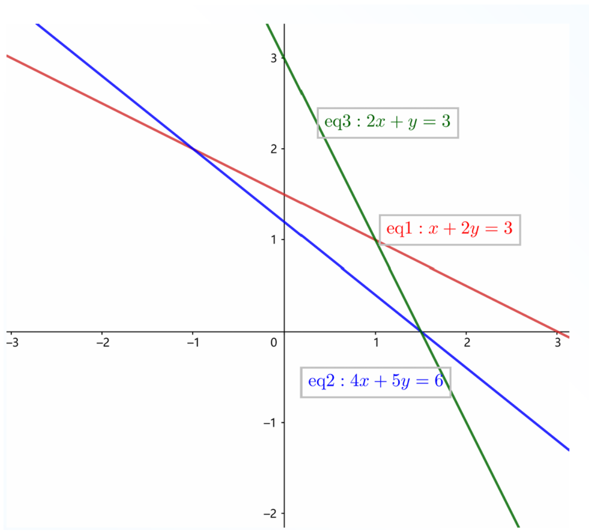
$$x + 2y = 3$$

$$4x + 5y = 6$$



In this example, we have unique intersection point at $(1, 2)$.

Example 1.5. *We may have infinitely many solutions e.g*



In this example, we have unique intersection point at $(1, 2)$*. In this example, we have no common intersection point. i.e. no solution*

Example 1.6. *Let's consider* 3×3 *system:*

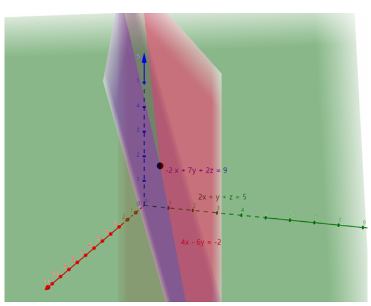
$$2x + y + z = 5 \tag{1.5}$$

$$4x - 6y \quad = -2 \tag{1.6}$$

$$-2x + 7y + 2z = 9 \tag{1.7}$$

where each equation describe a plane in $\mathbb{R}^3$

First two equation intersections at a line:



The solution (1,1,2) would be the unique intersection of the line with the plane defined by the last equation. This is the non-singular case.

Question: What is the general picture with m equations in n dimension (i.e. $n > 3$)?

1.Each non-trivial equation cut out a hyperplane (of dimension $n - 1$) in $\mathbb{R}^n$

2.the solution set is their common intersection, which may be empty;

3.whenever the solution set is non empty, it is at least max$(n - m,0)$ dimensional

4.to read off the dimension of the intersection, one have to transform the equation into row echelon form

1.2.2 Column Picture

We may treat a $n \times 1$ matrix as a vector in $\mathbb{R}^n$, e.g.

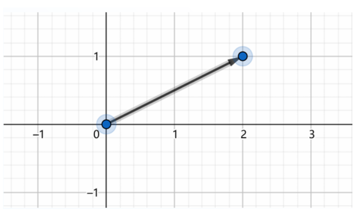
$$\vec{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

is a vector in $\mathbb{R}^2$. Two operations for vectors (similarly in $\mathbb{R}^n$):

$$\text{scalar multiplication : } c \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ cx_3 \end{bmatrix}$$

$$\text{vector addition : } \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$$

Pictorially, the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in $\mathbb{R}^2$. can be drawn as arrow:



1.3 An Example of Gaussian Elimination

1.3.1 Equivalent System

Definition 1.4. *Equivalent*

If two systems of linear equations have the same solution set, they are said to be equivalent.

There are three operations that can be used to obtain an equivalent system from an given one:

1. *The order of any two equations in a system can be interchanged;*
2. *Both sides of an equation can be multiplied by the same nonzero number;*
3. *A multiple of one equation can be added to another*

Theorem Gaussian Elimination Using the row operations 1, 2 and 3, one can always transform augmented matrix associated to a linear system in its reduced row echelon form.

This process is called Gaussian Elimination.

Once the equation is in row Echelon form, we can easily read off the solution

Example 1.7. *We solve:*

$$2u + v + w = 5 \tag{1.8}$$

$$4u - 6v + \quad = -2 \tag{1.9}$$

$$-2u + 7v + 2w = 9 \tag{1.10}$$

By using only operation 1 and 3, we can transform the equation into

$$2u + v + w = 5 \tag{1.11}$$

$$-8v - 2w = -2 \tag{1.12}$$

$$w = 2 \tag{1.13}$$

We end up with a strictly triangular system, with the diagonal coefficient 2, 8,1 being the pivots

one can use operation 1 and 3 to bring it into a diagonal system:

$$2u \qquad = 5 \tag{1.14}$$

$$-8v \quad = -2 \tag{1.15}$$

$$w = 2 \tag{1.16}$$

Therefore one can read off the solution (1,1,2) from a diagonal system.Equivalently, one can simply use the back-substitution method.

Remark 2. *For an nn non-singular system, we can always use operations 1 and 3 to obtain a strictly triangular system.*

1.3.2 The "Breakdown" of Elimination

Let consider the following system:

$$\begin{array}{rcrcrcl} u &+& v &+& w &=& \_ \\ 2u &+& 2v &+& 5w &=& \_ \\ 4u &+& 4v &+& 8w &=& \_ \end{array} \Longrightarrow \begin{array}{rcrcrcl} u &+& v &+& w &=& \_ \\ &&&& 3w &=& \_ \\ &&&& 4w &=& \_ \end{array}$$

So there is no way to avoid zero in the second diagonal position. This corresponds to the singular case. Two scenario in this example:

1.there is no solution to last two equations,inconsistent.

2.there is solution to last two equations, consistent but infinitely many solutions.

We can further eliminate $4w$ in last row by $3w$, and in this case we call 3 the second pivot. There is no third pivot.

1.4 Matrix Notation and Matrix Multiplication

1.4.1 Matrix Notation

A matrix is an arrangement of mn elements in n rows and n columns, given by

$$\begin{bmatrix} a_1 1 & a_1 2 & ... & a_1 n \\ a_2 1 & a_2 2 & ... & a_2 n \\ ... & ... & ... & ... \\ a_m 1 & a_m 2 & ... & a_m n \end{bmatrix}$$

Example 1.8.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$B = \begin{bmatrix} 5 & 4 & 1 \\ 3 & 2 & 2 \\ 9 & 7 & 9 \end{bmatrix}$$

$$C = \begin{bmatrix} 5 & 3 & 9 & 4 \end{bmatrix}$$

Suppose we have the equation

$$x_1 + 2x_2 + x_3 = 3 \tag{1.17}$$

$$3x_1 - x_2 - 3x_3 = -1 \tag{1.18}$$

$$2x_1 + 3x_2 + x_3 = 4 \tag{1.19}$$

We write it as$Ax = b$,where

Example 1.9.

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$b = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}$$

Here A is called the Coefficient matrix.The associated Augmented matrix is:

$$[A|b] = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{bmatrix}$$

1.5 Triangular Factors and Row Exchanges

1.6 Inverses and Transposes

1.7 Special Matrices and Applications

Chapter 2

Vector Spaces

For the linear system $Ax = b$, the basic questions are existence and uniqueness– Is there no solution, or one solution, or an infinity of solutions? They are much easy to answer after elimination. But we need to devote one more chapter to those questions, to find every solution for an m by n system. It goes to the heart of linear algebra.

2.1 Vector Spaces and Subspaces

2.1.1 Vector Spaces

Let's go back to $Ax = 0$, which may have infinitely solutions.

Example 2.1.

$$\begin{bmatrix} a & b \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

is a line l in $\mathbb{R}^2$ through origin, which is determined by one vector.

Example 2.2.

$$\begin{bmatrix} a & b & c \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

is a plane P in $\mathbb{R}^3$ through origin, which is determined by two vectors. The remarkable thing is that we can choose different sets of vectors $\vec{v_1}, \vec{v_2}$ to express the same plane.

We can get the common feature of $Ax = 0$
1.If $Ax = 0$ then $A(cx) = 0$.
2.If $Ax = 0, Ay = 0$ then we have $A(x + y) = 0$.
Than extract these properties into the notion of vector spaces.

Definition 2.1. *Once the vector in a space can have addition and scalar multiplication, and the result is still in this space, then the space can be called* ***vector space****.*

Remark 3. *Addition and scalar multiplication are required to satisfy these eight rules:*

1.$x + y = y + x$.
2.$x + (y + z) = (x + y) + z$.
3.There is a unique "zero vector" such that $x + 0 = x$ *for all* x.
4.For each x *there is a unique vector* $-x$ *such that* $x + (-x) = 0$.
5.$1x = x$.
6.$(c_1c_2)x = c_1(c_2x)$.
7.$c(x + y) = cx + cy$.
8.$(c_1 + c_2)x = c_1x + c_2x$.

Remark 4. *The formal definition allows other things to be "vectors"-provided that addition and scalar multiplication are right.*

1.The infinite-dimensional space $\mathbb{R}$.

2.The space of 3 by 2 matrices. Here the "vectors" are matrices, the space is almost the same as $\mathbb{R}^6$*(The six components are arranged in a rectangle instead of a column), indeed it is* ***isomorphic****(to* $\mathbb{R}^6$.

3.The space of even fuctions $f(x)$. *Here the "vectors" are fuctions, the dimension is somehow a larger infinity than for* $\mathbb{R}^n$.

Are these sets below vector spaces? You can give some examples.
1.$\mathbb{R}^3$ *without the origin.*
2.$\mathbb{R}^2$ *without the third quadrant.*
3.$\mathbb{C}^2$

2.1.2 Subspaces

Geometrically, think of the usual there-dimensional $\mathbb{R}^3$ *and choose any plane through the origin. Obviously, that plane is a vector space: it is a* ***subspace*** *of the original space* $\mathbb{R}^3$.

Definition 2.2. *If* S *in a nonempty subset of a vector space* V, *and* S *satisfies the conditions: linear combinations stay in the subspace.*
1.$\vec{u} + \vec{v} \in S$ *wnhenever* $\vec{u} \in S$ *and* $\vec{v} \in S$.
2.$c\vec{u} \in S$ *whenever* $\vec{u} \in S$ *for any scalar* c.
Then S *is said to be a* ***subspace*** *of* V.

Notice in particular that the zero vector will belong to every subspace. That comes from rule (ii): Choose the scalar to be $c = 0$.

Remark 5. *The subsets* $\{0\}$ *and* V *are subspaces of* V, *and are called* ***trivial subspaces****().*

Remark 6. *Vector vs Points*

It is common to represent each vector with just a point in space, the point at the tip of that vector with its tail on the origin. That way, if you want to think about every possible vector whose tip sits on a certain line, just think about the line itself. Likewise, to think about all possible two-dimensional vectors all at once. What you will be thinking about is the infinite, flat sheet of two-dimensional space itself, laeving the arrows out of it.

In linear algebra, lines, planes, spaces can all be treated as a collection of vectors. That is why they are called vector spaces.

Example 2.3. *Subspaces of* $\mathbb{R}^2$ *(three types):*

1.The origin. That is the zero vector $\begin{bmatrix}0\\0\end{bmatrix}$.

2.Straight line across the origin.

3.$\mathbb{R}^2$ *itself.*

Try to list the subspaces of $\mathbb{R}^5$.

Example 2.4. *Start from the vector space of 3 by 3 matrices. One possible subspace is the set of lower triangular matrices. Another is the set of symmetric matrices.*

Property

Property 1. *The intersection of two subspaces is still a subspace.*

Property 2. *Tf* W *is a subspace of* V, *then* $\vec{0_w} = \vec{0_v}$. *(Is* $\mathbb{R}^1$ *a subspace of* $\mathbb{R}^2$*?)*

Try to decide whether the following ones are subspaces of $\mathbb{R}^2$.

1.$S_1 = \{(x, 0)|x \in \mathbb{R}\}$

2.$S_2 = \{(x, y)|4x + y = 0\}(x, y \in \mathbb{R})$

3.$S_3 = \{(x, y)|x + y = 1\}(x, y \in \mathbb{R})$

4.$S_4 = \{(x, y)|xy = 0\}(x, y \in \mathbb{R})$

5.$S_5 = S_1 \cap S_2$

6.$S_6 = S_1 \cup S_2$

The Column Space of A

We can first associate column space with linear equation $Ax = b$.

$$\begin{bmatrix}1 & 1 & 2\\2 & 1 & 3\\4 & 1 & 5\\7 & 1 & 8\end{bmatrix} \times \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}b_1\\b_2\\b_3\\b_4\end{bmatrix}$$

With m>n we have more equations than unknownsand usually there will be no solution. An important question is which b allow this equation system to have solution?

Some right side b which make the equation system solvable are given:

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

How about we write the solution first, find which b lead to this solution?

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 3 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

The left side gives all posible linear combinations of the three column vectors.

Definition 2.3. *Given a m by n matrix* $A = \begin{bmatrix} \vec{a_1} & \vec{a_2} & & \vec{a_n} \end{bmatrix}$, ***the column space of A*** *is defined to be* $C(A) = \{c_1\vec{a_1} + c_2\vec{a_2} + \ + c_n\vec{a_n} | c_i \in \mathbb{R}\}$

You may already discover. The system $Ax = b$ is solvable if and only if the vector b can be expressed as a combination of the columns of A. Then b is in the column space.

For matrix A, do its three column vectors all contribute to expand the dimension of column space? No, because even if we delete a column, the column space will not change. We will call the ***linearly independent****(columns(contribute to column space) that come first pivot columns later. Hence,* $C(A) = span\{ \begin{bmatrix} 1 \\ 2 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \}$

Example 2.5. *Say* $R = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Any elements of the form

$$\begin{aligned} \vec{x} &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} \\ &= (x_1 + 3x_2 + 2x_4) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (x_3 + 3x_4) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

is in $C(R)$.

$$C(R) = span\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \} = span\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} | a, b \in \mathbb{R}\} = span\{\vec{a_1}, \vec{a_3}\}$$

C(A)=span{$\vec{a_i}$|i is the pivot column in the RREF}. This is a minimal spanning set for C(A).

For another matrix, the smallest possible column space (one vector only) comes from the zero matrix $A = 0$. The only combination of the columns is $b = 0$. At the other extreme, suppose A is the 5 by 5 identity matrix. Then $C(I)$ is the whole of $\mathbb{R}^5$; the five columns of I can combine to produce any five-dimensional vector b. This is not at all special to the identity matrix. Any 5 by 5 matrix that is nonsingular will have the whole of $\mathbb{R}^5$ as its column space. For such a matrix we can solve $Ax = b$ by Gaussian Elimination: there are five pivots. Therefore every b is in $C(A)$ for a nonsingular matrix.

The Nullspace of A

Definition 2.4. *The nullspace of A consists of all vectors x such that $Ax = 0$. It is denoted by $N(A)$. It is a subspace of $\mathbb{R}^n$, just as the column space was a subspace of $\mathbb{R}^m$.*

Definition means the nullspace contains all the solutions $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *to*

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

Some obvious solutions:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 4 & 1 & 5 \\ 7 & 1 & 8 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 4 & 1 & 5 \\ 7 & 1 & 8 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Then, we can write a complete solution to this equation system by observation.

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 4 & 1 & 5 \\ 7 & 1 & 8 \end{bmatrix} \times \begin{bmatrix} c \\ c \\ -c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where c is a constant. So the nullspace looks like a straight line across the origin which is a subspace of $\mathbb{R}^3$.

Is the solutions stil a subspace when the right-hand side b not equals to zero? Definitely no, because zero vector isn't in the solution space.

Remark 7. *Contrast between $N(A)$ and $C(A)$*

$\boldsymbol{N}(A)$	$\boldsymbol{C}(A)$
1. $\boldsymbol{N}(A)$ is a subspace of $\mathbb{R}^n$.	1. $\boldsymbol{C}(A)$ is a subspace of $\mathbb{R}^m$.
2. $\boldsymbol{N}(A)$ is implicitly defined; *i.e.*, you are given only a condition $A\mathbf{x} = 0$ that vectors in $\boldsymbol{N}(A)$ must satisfy.	2. $\boldsymbol{C}(A)$ is explicitly defined; *i.e.*, you are told how to build vectors in $\boldsymbol{C}(A)$.
3. It takes time to find vectors in $\boldsymbol{N}(A)$. Row operations on $[A\ 0]$ are required.	3. It is easy to find vectors in $\boldsymbol{C}(A)$. The columns of A are displayed; others are formed from them.

$\boldsymbol{N}(A)$	$\boldsymbol{C}(A)$
4. There is no obvious relation between $\boldsymbol{N}(A)$ and the entries in A.	4. There is an obvious relation between $\boldsymbol{C}(A)$ and the entries in A, since each column of A is in $\boldsymbol{C}(A)$.
5. A typical vector $\mathbf{v}$ in $\boldsymbol{N}(A)$ has the property that $A\mathbf{v} = 0$	5. A typical vector $\mathbf{v}$ in $\boldsymbol{C}(A)$ has the property that the linear system $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector $\mathbf{v}$, it is easy to tell if $\mathbf{v}$ is in $\boldsymbol{N}(A)$. Just compute $A\mathbf{v}$	6. Given a specific vector $\mathbf{v}$, it may take time to tell if $\mathbf{v}$ is in $\boldsymbol{C}(A)$. Row operations on $[A\ \mathbf{v}]$ are required.

Now, you have know two important subspaces defined by a matrix. At the end of this chapter, you will know 4 subspaces-the column space of A, the nullspace of A, and their two perpendicular spaces. You are now approaching the core of linear algebra.

2.2 Solving $Ax = 0$ and $Ax = b$

2.2.1 $Ax = 0$

This section introduces an algorithm. To understand the solving process, let's look at this matrix.

$$A = \begin{bmatrix} 1 & 3 & 3 & 3 \\ 2 & 6 & 8 & 10 \\ 3 & 9 & 11 & 13 \end{bmatrix}$$

How to find the solution of $Ax = 0$?

We have learnt an algorithm for solving linear equation systems-Gauss Elimination, which can help us simplify the questions and find the pivots.
Do G.E. for matrix A
The first step:

$$\begin{bmatrix} 1 & 3 & 3 & 3 \\ 2 & 6 & 8 & 10 \\ 3 & 9 & 11 & 13 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

The second column has no pivot. In chapter 1(square matrix cases), we call this situation permanent failure(or singular). But it doesn't mean $Ax = 0$ has no solution. We can go on to find the next pivot.

$$A = \begin{bmatrix} 1 & 3 & 3 & 3 \\ 2 & 6 & 8 & 10 \\ 3 & 9 & 11 & 13 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

*G.E. ends. the result is not a strict upper triangular matrix, we call it **Echelon Form** U.*
*The matrix A has two pivots, which is called the **rank**(of matrix A.*

Definition 2.5. *The rank of a $m \times n$ matrix A is the number of pivots in its row echelon from. It is written as rk(A) or rank(A).*

Property 3. *1.rk(A) $\leq min(m, n)$ for $m \times n$ matrix A.*
2.For $n \times n$ matrix A: A non-singular $\Longleftrightarrow$ A invertible $\Longleftrightarrow$ rk(A)=n.
3.rk(AB) $\leq min(rk(A), rk(B))$.
*4.rk(A)= **dimension**() of C(A)= minimal no. of columns needed for spanning C(A)*

During this elimination process the nullspace is unchanged.

proof: *$PA = LU$,where P and L are invertible $m \times m$ matrices.*

$$\begin{aligned} 0 &= Ax \\ &= (L^{-1}P)Ax \\ &= Ux \end{aligned}$$

Then $N(A) = N(U)$

The simplified equation system becomes

$$\begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Find the pivots and they determine the pivot columns, the rest are free columns. In this example, column 1 and column 3 are pivot columns, column 2 and 4 are free columns.
We can take any number for the coefficient of free columns and solve the coefficient for pivot columns by back-substitution. Commonly, we choose to set a free column coefficient 1 and the others 0.

$$\begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} x_1 \\ 1 \\ x_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} x_1 \\ 0 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Write their corresponding equation systems.

$$\begin{cases} x_1 + 3x_2 + 3x_3 + 3x_4 = 0 \\ \qquad\qquad\quad 2x_3 + 4x_4 = 0 \\ \qquad\qquad\qquad\qquad\ \ 0 = 0 \end{cases}$$

Solve the coefficients of pivot columns by back-substitution now.

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

Add them, the solution is $x = x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 0 \\ -2 \\ 1 \end{bmatrix}$. *They span a 2-dimensional space* $N(A) = span\{ \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \\ 1 \end{bmatrix} \}$.

Recall in Chapter 1, we can do row operations to let the matrix become an identity matirx, then the solution appears on the right side (that is how you find the inverse of matrix) . Similarly, we can have further simplication to this linear equation system.

$$U = \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 3 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 & -3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R$$

After this process, we have the ***Reduced Row Echelon Form(RREF)*** R. *During this elimination process the nullspace is unchanged too.*

proof: $U = D\tilde{U}$ *and* $E\tilde{U} = R$, *where* D *and* E *are invertible* $m \times m$ *matrices.*

$$\begin{aligned} 0 &= Ux \\ &= D\tilde{U}x \\ &= D^{-1}D\tilde{U}x \\ &= \tilde{U}x \\ &= E\tilde{U}x \\ &= Rx \end{aligned}$$

Then $N(A) = N(U) = N(R)$

2.2.2 $Ax = b$

If we slightly change the right-hand side b, *the augmented matrix is:*

$$\begin{bmatrix} 1 & 3 & 3 & 3 & 1 \\ 0 & 0 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Write its corresponding equation systems.

$$\begin{cases} x_1 + 3x_2 + 3x_3 + 3x_4 = 1 \\ 2x_3 + 4x_4 = 4 \\ 0 = 0 \end{cases}$$

An important question: What is the particular solution? It can be any solution to the linear equation system, no matter what method you use.
I'd like to introduce you one common method: Set all free variables 0.

$$\begin{cases} x_1 + 3x_3 = 1 \\ \qquad\; 2x_3 = 4 \end{cases}$$

So, a particular solution is $x_p = \begin{bmatrix} -5 \\ 0 \\ 2 \\ 0 \end{bmatrix}$

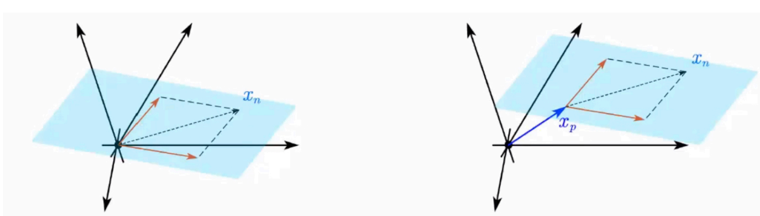
To find the complete solution: ***add the particular solution by nullspace vectors****.*

$$\begin{cases} Ax_p = b \\ Ax_n = 0 \\ A(x_p + x_n) = b \end{cases}$$

The complete solution is $x_{complete} = x_{particular} + x_{nullspace}$.
For the example above, the complete solution is

$$x_c = x_p + x_n = \begin{bmatrix} -5 \\ 0 \\ 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 0 \\ -2 \\ 1 \end{bmatrix}, x_2, x_4 \in \mathbb{R}$$

Can the whole solution set construct a subspace of $\mathbb{R}^4$*? Firstly, let's consider nullspace only. It is a 2-dimensional plane through the origin in* $\mathbb{R}^4$ *space. Then add a particular solution. It is now a 2-dimensional plane away form the origin in* $\mathbb{R}^4$ *space.*



Nullspace of A — $2-$*dimensional plane through the origin*

Solution set of $Ax = b$ — $2-$*dimensional plane away from the origin*

Now try to understand any particular solution is OK if the tip of $\vec{x_p}$ *is in the blue plane.*

Example 2.6. $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}, \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$*, when* $Ax = b$ *is solvable?*

The augmented matrix:

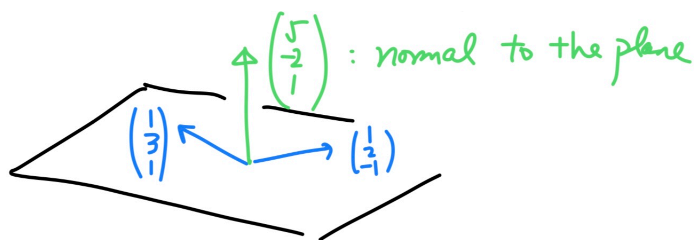
$$\begin{bmatrix} 1 & 3 & 3 & 2 & b_1 \\ 2 & 6 & 9 & 7 & b_2 \\ -1 & -3 & 3 & 4 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 3 & 3 & b_2 - 2b_1 \\ 0 & 0 & 6 & 6 & b_3 + b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 3 & 3 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - 2b_2 + 5b_1 \end{bmatrix}$$

Hence, when $b_3 - 2b_2 + 5b_1 = 0$*,* $Ax = 0$ *is solvable.*

It is worth mentioning that

$$C(A) = span\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}\} = span\{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} | b_3 - 2b_2 + 5b_1 = 0\}$$

Pictorially:



Example 2.7. *Consider* $\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{b}$

1.Find column space, null space of A.
2.Find condition on $\vec{b}$ *s.t. the equation is consistent.*
3.Find general solutions for $Ax = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix}$.

What we have to do is G.E.:

$$\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & -2 & -2 & b_3 - 3b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 1 & 1 & 1/2b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 2 & 4b_1 - 3/2b_2 \\ 0 & 0 & 1 & 1 & 1/2b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix}$$

If we stop at: $\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix}$

$$C(A) = span\{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 7 \end{bmatrix}\} = \{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} | b_3 + b_2 - 5b_1 = 0\}.$$

For null space: two free variables x_2 and x_4 with two equations:

$$\begin{cases} 2x_3 + 2x_4 = 0 \\ x_1 + 2x_2 + 3x_3 + 5x_4 = 0 \end{cases} \rightarrow \begin{cases} x_3 = -x_4 \\ x_1 = -2x_2 - 2x_4 \end{cases}$$

Hence, $N(A) = \{ \begin{bmatrix} -2x_2 - 2x_4 \\ x_2 \\ -x_4 \\ x_4 \end{bmatrix} | x_2, x_4 \in \mathbb{R} \} = span\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \}.$

For $\vec{b} = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix}$, *we find the particular solution.*

$$\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 5 & 0 \\ 0 & 0 & 2 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Set free variables $x_2 = x_4 = 0$:

$$\begin{cases} x_1 + 3x_3 = 0 \\ 2x_3 = 6 \end{cases} \rightarrow \begin{cases} x_3 = 3 \\ x_1 = -9 \end{cases}$$

Therefore, $x_p = \begin{bmatrix} -9 \\ 0 \\ 3 \\ 0 \end{bmatrix}$. $x = \begin{bmatrix} -9 \\ 0 \\ 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$

Remark 8. *1.Permuting rows may be needed in the Elimination, since our focus is not LU factorization, we don't need to fix the permutation P at the beginning.*

2.It is enough to reduce to U instead of R for solving the equation, finding $C(A)$ and $N(A)$.

3.We have the equality: no. pivot columns + no. free variables of $N(A) = n$.

2.3 Linear Independence, Basis, and Dimension

For a $m \times n$ matrix, the numbers m and n give an incomplete picture of the true size of a linear system. The important number is the rank r, which was introduced as the number of pivots in the elimination process. But it would be wrong to leave it there because the rank has a simple and intuitive meaning: The rank counts the number of genuinely independent rows in the matrix A.

The goal of this section is to explain and use four ideas:

1.Linear independence or dependence.
2.Spanning a subspace.
3.Basis for a subspace(a set of vectors).
4.Dimension of a subspace(a number).

2.3.1 Linear Independence

Definition 2.6. *A set of vectors $\vec{v_1},,v_k$ are said to be* ***linearly independent*** *if whenever we have an expression*

$$c_1\vec{v_1} + \quad + c_k\vec{v_k} = \vec{0}$$

then we must have $c_1 = \quad = c_k = 0$.
Otherwise, they are called ***linearly dependent****.*
In the other word: we cannot find a non-trivial linear combination to express zero vector.

Suppose $v_1 = \begin{bmatrix}1\\3\end{bmatrix}, v_2 = \begin{bmatrix}2\\6\end{bmatrix}$, they are not linearly independent because $2v_1 - v_2 = 0$.
Suppose $v_1 = \begin{bmatrix}1\\3\end{bmatrix}, v_2 = \begin{bmatrix}0\\0\end{bmatrix}$, they are not linearly independent because $0v_1 - 3v_2 = 0$.
A zero vector makes the whole set of vectors dependent because we can take zero other vectors and one zero vector to get a zero vector.

Can the flowing 3 vectors in $\mathbb{R}^2$ be linearly independent?

$$v_1 = \begin{bmatrix}1\\3\end{bmatrix}, v_2 = \begin{bmatrix}2\\1\end{bmatrix}, v_3 = \begin{bmatrix}3\\5\end{bmatrix}$$

We can represent every vector using 2 vectors that are not in the same line: $v_3 = 7/5v_1 + 4/5v_2$, which means $7/5v_1 + 4/5v_2 - v_3 = 0$. Hence, these 3 vectors are not linearly independent.
Another interpretation:

$$\begin{bmatrix}1 & 2 & 3\\3 & 1 & 5\end{bmatrix} \times \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$$

This system always have nonzero solutions. Free columns always exist.
To repeat, if a matrix A have nonzero solutions to $Ax = 0$, then the columns are linearly dependent. The solution shows us which combinations of columns of A gives the zero vector.

Remark 9. *If If we write the vectors in columns of A:*

They are independent if and only if the nullspace has only the zero vector, no free columns appear and $n = r$.

They are dependent if and only if the nullspace has not only the zero vector, has free columns appear and $n > r$.

Example 2.8. *Let $A = \begin{bmatrix}3 & 1 & 4 & 2\\0 & 0 & 1 & 5\\0 & 0 & 0 & 2\end{bmatrix}$, then the pivots columns $\vec{a_1}, \vec{a_3}, \vec{a_4}$ are linearly independent, and adding any non-pivot column become dependent.*

If $c_1 \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} = 0$
$\rightarrow c_3 = 0$ *form last row.*
$\rightarrow c_2 = 0$ *from second row.*
$\rightarrow c_1 = 0$ *from first row.*
$\rightarrow \vec{a_1}, \vec{a_3}, \vec{a_4}$ *are linearly independent.*
$\vec{a_1} - 3\vec{a_2} = 0$
$\rightarrow$ *linearly dependent.*

Remark 10. $\vec{a_1},, \vec{a_n}$ *are linearly independent* $\iff$ $Q\vec{a_1},, Q\vec{a_n}$ *are linearly independent*

Example 2.9. *If* $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$, *then* $\exists$ *invertible* 3×3 *matrix* E *s.t.*

$$A = E \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

After a series of row operations, the row space and the nullspace remain unchanged, but the column space changed.

Check if this relation of columns is still unchanged?

$$(Col2) + (Col3) = 3(Col4), 3(Col1) = (Col2)$$

The answer is yes. That is because nullspace never changes. The nullspace shows us which combination of columns can lead us to zero and the relation is ever-lasting.

Hence, $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ *and* $\begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}$ *as well as* $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ *and* $\begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix}$ *are linearly independent.*

Remark 11. *If we have* $\vec{a_1},, \vec{a_n}$ *in* $\mathbb{R}^m$ *and*

$$\vec{a_i}^T \vec{a_j} = \begin{cases} non-zero(i = j) \\ 0(i \neq j) \end{cases}$$

then $\vec{a_1},, \vec{a_n}$ *are linearly independent.*

i.e. ***orthogonal vectors****() are linearly independent.*

Proof: If $c_1\vec{a_1} + \; + c_n\vec{a_n} = 0$, then $\vec{a_j}^T(c_1\vec{a_1} + \; + c_n\vec{a_n}) = c_j(\vec{a_j}^T\vec{a_j})$
Because $\vec{a_j}^T\vec{a_j}$ is a non-zero number, $c_j = 0$.

Remark 12. *Supposed* $\vec{v}_1,, \vec{v}_k$ *are linearly independent,*
adding $\vec{v}_{k+1}$ *are still linearly independent* $\iff$ $\vec{v}_{k+1}$ *not in span*$\{\vec{v}_1,, \vec{v}_k\}$

Proof:

not in $\rightarrow$ *l.i.* $\iff$ *l.d.* $\rightarrow$ *in*

If $\vec{v_1},,\vec{v}_{k+1}$ are linearly dependent, then $c_1\vec{v_1}++c_{k+1}\vec{v}_{k+1}=0$ for some c_i not all zero.

$c_{k+1}\neq 0$ because $\vec{v_1},,\vec{v_k}$ are linearly independent, then $\vec{v}_{k+1}=-c_1/c_{k+1}\vec{v_1}--c_k/c_{k+1}\vec{v}_k$, which means $\vec{v}_{k+1}$ in $span\{\vec{v_1},,\vec{v_k}\}$.

The proof of **l.i.** $\rightarrow$ **not in** is similar.

Remark 13. *Let V be a vector space and $v_1,v_2,,v_k(k\geq 2)$ be a set of vectors. Then $v_1,v_2,,v_k$ are linearly dependent if and only if one of the vectors v_j can be written as a linear combination of the other vectors.*

Proof: (1)Assume $v_1,v_2,,v_k$ are linearly dependent, the equation $c_1v_1+c_2v_2++c_kv_k=0$ has nonzero solution.

This means at least one of the c_i is nonzero.

Let c_r be the last one, with $c_r\neq 0$.

$$c_1v_1+c_2v_2++c_rv_r=0$$

$$v_r=-c_1/c_rv_1-c_2/c_rv_2--c_{r-1}/c_rv_{r-1}$$

So v_r is a linear combination of other vectors.

(2)Assume V_r is linear combination of other vectors.

$$v_r=(c_1v_1+c_2v_2++c_{r-1}v_{r-1})+(c_{r+1}v_{r+1}++c_kv_k)$$

$$c_1v_1+c_2v_2++c_{r-1}v_{r-1}+(-1)v_r+c_{r+1}v_{r+1}++c_kv_k=0$$

So $v_1,v_2,,v_k$ are linearly dependent.

2.3.2 Spanning, Basis and Dimension

spanning

Definition 2.7. *Vectors $\vec{v_1},\vec{v_2},,\vec{v_n}$ span a space means: the space* ***consists of*** *all linear combinations of these vectors.*

Can I change the word "consists of" with "contains"? No.
Vectors $\vec{v}_1,\vec{v}_2,,\vec{v}_n$ span a space means: the space is the smallest vector space that contains all the combinations of these vectors. The spanning space of a set of vectors is exactly all the combinations of the vectors.
The columns of A span the column space $C(A)$.
There is no need for these vectors to be linearly independent.

Example 2.10. *The vectors $w_1=(1,0,0),w_2=(0,1,0)$ and $w_3=(-2,0,0)$ span a plane(the $x-y$ plane) in $\mathbb{R}^3$. The first two vectors also span this plane, where w_1 and w_3 span only a line.*

Example 2.11. *The coordinate vectors $e_1,,e_n$ coming from the identity matrix span $\mathbb{R}^n$. Every vector $b=(b_1,,b_n)$ is a combination of those columns. In this example the weights are the components b_i themselves:*
$b=b_1e_1++b_ne_n$.

Basis

To decide if the columns are independent, we solve $Ax = 0$. To decide if b is a combination of the columns, we try to solve $Ax = b$. Independence involves the nullspace, and spanning involves the column space. The coordinate vectors $e_1, , e_n$ span $\mathbb{R}^n$ and they are linearly independent. Roughly speaking, no vectors in that set are wasted.

This leads to the crucial idea of a basis.

Definition 2.8. *A basis for V is a sequence of vectors having two properties at once:*

1. The vectors are linearly independent (not too many vectors).
2. They span the space V (not too few vectors).

Example 2.12. $\vec{v_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ *is a basis for* $\mathbb{R}^2$ *because* $N(A) = 0, C(A) = \mathbb{R}^2$ *with* $A = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$

$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *is not a basis for* $\mathbb{R}^2$ *because* $span\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\} \neq \mathbb{R}^2$

$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *is not a basis for* $\mathbb{R}^2$ *because* $\vec{v}_1, \vec{v}_2, \vec{v}_3$ *are linearly dependent.*

Remark 14. *There is one and only one way to write v as a combination of the basis vectors*

Proof: *If $v = a_1v_1 + + a_kv_k$ and also $v = b_1v_1 + + b_kv_k$, then subtraction gives $0 = \sum(a_i - b_i)v_i$. because independence, every coefficient $a_i - b_i$ must be zero. Therefore $a_i = b_i$.*

Remark 15. *A basis is an optional spanning set of V, it is not unique:*

$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *is a basis for* $\mathbb{R}^2$, $\vec{v_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ *is also.*

*And in $\mathbb{R}^n$, the set $\{e_1, e_2, , e_n\}$ is called the **standard basis** of $\mathbb{R}^n$*

Remark 16. *Consider* $U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ *and* $C(U)$.

$$C(U) = span\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}\}$$

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ *are linearly independent as* $N(\begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}) = 0$

Hence, pivot columns is a set of basis for $C(U)$ in the case that U is in row echelon form. If we are asking the columns to be a basis for the whole space $\mathbb{R}^n$, then the matrix must be square and invertible.

Example 2.13. *1.Let $V = \{a_0 + a_1x + a_2x^2 + a_3x^3 | a_i \in \mathbb{R}\}$, then $1, x, x^2, x^3$ is a basis for V.*

2.Let $\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$, then $1, i$ is a basis for $\mathbb{C}$.

3.Let $M_{2\times2}(\mathbb{R}) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} | a, b, c, d \in \mathbb{R}\}$, then $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is a basis for $M_{2\times2}(\mathbb{R})$

4.Let $S_{2\times2}(\mathbb{R}) = \{A | A \in M_{2\times2}(\mathbb{R}), A^T = A\}$, then $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a basis for $S_{2\times2}(\mathbb{R})$

5.Let $B_{2\times2}(\mathbb{R}) = \{A \in M_{2\times2}(\mathbb{R}) | tr(A) = 0\}$, then $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is a basis for $B_{2\times2}(\mathbb{R})$

6.Let $D_{3\times3}(\mathbb{R}) = \{A | A \in M_{3\times3}(\mathbb{R}), A^T = -A\}$, then $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ is a basis for $D_{3\times3}(\mathbb{R})$

Dimension

A space has infinitely many different basis, but there is something common to all of these choices. The number of basis vectors is a property of the space itself.

Definition 2.9. *If V is a vector space with basis $\vec{v_1}, , \vec{v_n}$, we say that the dimension of V is n or V is n-dimensional.*

Dimension of $\mathbb{R}^n$ is n.
Dimension of $B_{2\times2}(\mathbb{R})$ is 3.
Dimension of $D_{3\times3}(\mathbb{R})$ is 3.

Property 4. *If $v_1, ..., v_m$ and $w_1, ..., w_n$ are both bases for the same vector space, then $m = n$. Dimension of V is well-defined.*

Proof: Suppose there are more ws than vs $(n > m)$. We will arrive at a contradiction.

Say $n = 3$, $m = 2$:

$$w_1 = a_{11}v_1 + a_{21}v_2$$

$$w_2 = a_{12}v_1 + a_{22}v_2$$

$$w_3 = a_{13}v_1 + a_{23}v_2$$

Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$.

Because $Ax = 0$ has a **non-zero** solution $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$.

We have:

$$\begin{bmatrix} \vec{w_1} & \vec{w_2} & \vec{w_3} \end{bmatrix} = \begin{bmatrix} \vec{v_1} & \vec{v_2} \end{bmatrix} \times \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Then we have:

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0 \to c_1w_1 + c_2w_2 + c_3w_3 = 0$$

In this case, the ws could not be a basisso we cannot have $n > m$.

Property 5. *Given two vector subspaces W_1 and W_2 of a finite-dimensional vector space V, then $dim(W_1) + dim(W_2) = dim(W_1 + W_2) + dim(W_1 \cap W_2)$.*

Proof: Let $x_1, , x_n$ be basis of W_1, $y_1, , y_m$ be basis of W_2.

1.We want to prove $dim(W_1 + W_2) = dimC(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$

(1)$\forall z \in W_1 + W_2$, $\exists x \in W_1, y \in W_2$ s.t. $z = x + y$
Notice $x = \sum a_ix_i$, $y = \sum b_iy_i$
$\to z = \sum a_ix_i + \sum b_iy_i$
$\to z \in C(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$
(2)$\forall z \in C(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$
$\to z = \sum a_ix_i + \sum b_iy_i$
$\to z = x + y$ with $x \in W_1, y \in W_2$
$\to z \in W_1 + W_2$

Hence,$dim(W_1 + W_2) = dimC(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$

2.We want to prove $dim(W_1 \cap W_2) = dimN(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$, the nullspace is equal to $\sum a_ix_i + \sum b_iy_i = 0$

(1)If $\sum a_ix_i = 0$, then $a_i = 0$ and $\sum b_iy_i = 0 \to b_i = 0$
If $\sum a_ix_i \neq 0$, then $\sum a_ix_i \in W_1$ and $\sum a_ix_i = -\sum b_iy_i \in W_2 \to \sum a_ix_i \in W_1 \cap W_2$
$\to dimN(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix}) \leq dim(W_1 \cap W_2)$
(2)$\forall u \in W_1 \cap W_2, \exists \{a_i\}, \{b_i\}$ s.t. $\sum a_ix_i = -\sum b_iy_i = u$
$\to \sum a_ix_i + \sum b_iy_i = 0$
$\to dimN(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix}) \geq dim(W_1 \cap W_2)$

Hence,$dim(W_1 \cap W_2) = dimN(\begin{bmatrix} x_1 & & x_n & y_1 & & y_m \end{bmatrix})$

In conclusion, $dim(W_1) + dim(W_2) = dim(W_1 + W_2) + dim(W_1 \cap W_2)$

Property 6. *Let V be a space, then*
1.Any linearly independent set of V can be extended to a basis, by adding more vectors if necessary.
2.Any spanning set for V can be reduced to a basis, by discarding vectors if necessary.

Example 2.14. *Consider the matrix A:*

$$A = \begin{bmatrix} 1 & 3 & 4 & 2 \\ 2 & 7 & 9 & 4 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

1.Can the 4 column vectors span $C(A)$?
Of course, it is the definition of column space.
2.Are they a basis for the column space $C(A)$?
No, they are not independent because $Ax = 0$ has nonzero solutions.
3.Can you find a basis for the column space $C(A)$?
Column 1 and column 2. They are pivot columns also. And $dim(C(A)) = 2$ because one set of basis have 2 vectors.

rank r = the number of pivot columns = $dim(C(A))$
n - r = the number of free columns = $dim(N(A))$
$dim(C(A)) + dim(N(A)) = n$ for $m \times n$ matrix A

Remark 17. *We must say $= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ being a four dimensional vector, meaning it is a vector in $\mathbb{R}^4$, but span$\{\vec{v}\}$ is 1-dimensional subspace of $\mathbb{R}^4$. So the word "four dimensional vector" here refer to its **ambient space**(.*

2.4 The Four Fundamental Subspace

2.4.1 Null space

Given any n*n matrix A we can have

$$P = P_1, P_2, ..., P_k(permutation\ matrix)$$
$$L = E_1, E_2, ..., E_s(type\ 3\ elementary\ matrices)$$

s.t.

$$PA = LU$$

we talked about the case that A is a $n \times n$ matrix the same method works here

U could be in the row echelon form

$$U = \begin{bmatrix} P_1 & * & * & * & * & * & * \\ 0 & 0 & P_2 & * & * & * & * \\ 0 & 0 & 0 & 0 & P_3 & * & * \\ 0 & 0 & 0 & 0 & 0 & P_4 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We can further write

$$U = \begin{bmatrix} P_1 & & & & \\ & P_2 & & & \\ & & P_3 & & \\ & & & P_4 & \\ & & & & 1 \end{bmatrix} \times \begin{bmatrix} 1 & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Let's consider = $\begin{bmatrix} 1 & * & * & * & * & * & * \\ 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ By further row operation, using $m \times m$ upper triangular matrices E

$$E\tilde{U} = \begin{bmatrix} 1 & * & 0 & * & 0 & 0 & * \\ 0 & 0 & 1 & * & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = R$$

R is in **Reduced Row Echelon form**

Property

Property 7. $NA) = N(U) = N(\tilde{U}) = N(R)$

Proof as PA=LU
(P,L are invertible matrices)
then

$$\begin{aligned} Ax &= 0 \\ (L^{-1}P)Ax &= 0 \\ Ux &= 0 \end{aligned}$$

similarity

$$\begin{aligned} U &= D\tilde{U} \\ E\tilde{U} &= R \end{aligned}$$

(E,U are invertible $m \times m$ matrices)

Example 2.15.

$$R = \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, solve\ R \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 0$$

We solve from the bottom rows:

$$x_3 + x_4 = 0 \rightarrow x_3 = -x_4$$

Then the rows above is

$$x_1 + 3x_2 - x_4 = 0 \rightarrow x_1 = -3x_2 + x_4$$

i.e. $\vec{x}=\begin{bmatrix} -3x_2+x_4 \\ x_2 \\ -x_4 \\ x_4 \end{bmatrix}$ *or N(R)=*$\{\begin{bmatrix} -3x_2+x_4 \\ x_2 \\ -x_4 \\ x_4 \end{bmatrix} /x_2, x_3 \in \mathbb{R}\}$

The variables NOT corresponding to pivots are free variables, variables correspond to pivots are expressed in terms of free variables

we can write $\vec{x} = \begin{bmatrix} -3x_2+x_4 \\ x_2 \\ -x_4 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$

or $N(R) = span\{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}\}$

minimal size of spanning set = the number of free variables=the number of number of non-zero rows.

Example 2.16. $A=\begin{bmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & 2 & 0 & 8 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ $,A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$ *=0*

we start from the bottom

$$x_4 + 3x_5 = 0 \to x_4 = -3x_5$$

then the next row

$$x_2 + 2x_3 + 8x_5 = 0 \to x_2 = -2x_3 - 8x_5$$

Finally:

$$x_1 + x_3 + x_5 = 0 \to x_1 = -x_3 - 3x_5$$

$$N(A) = \{\begin{bmatrix} -x_3 - 3x_5 \\ -2x_3 - 8x_5 \\ x_3 \\ -3x_5 \\ x_5 \end{bmatrix} |x_3, x_5 \in \mathbb{R}\} = span\{\begin{bmatrix} -3 \\ -8 \\ 0 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\}$$

(non-pivot columns $\to$ *free variables* x_3 *and* x_5*)*
maybe you realize a quick way to write down N(A)
In this case: the number of free variables=2=the number of non-pivot columns

2.4.2 Column space

Example 2.17. *If* $\begin{bmatrix} 1 & 3 & 3 & 2 & b_1 \\ 0 & 0 & 3 & 3 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - 2b_2 + 5b_1 \end{bmatrix}$
with $b_3 - 2b_2 + 5b_1$ *is non-zero,then it is incomsistent*
Therefore

$$C(A) = \{ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} | b_3 - 2b_2 + 5b_1 = 0\}$$

Say $R = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ *in rref(reduced row echelon form)*
Any elements of the form

$$\vec{x} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} = (x_1 + 3x_2 + 2x_4) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (x_3 + 3x_4) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

is in C(R)

$$C(R) = span\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \} = \{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} | a, b \in \mathbb{R}\}$$

Another view on C(A)

$$E^{-1}A = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R$$

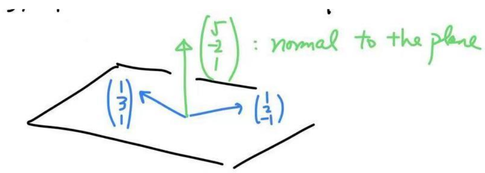
$$A = ER = \begin{bmatrix} E \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & E \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} & E \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} & E \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \end{bmatrix}$$

$$C(A) = \{x_1 E \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 E \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} + x_3 E \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_4 E \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}\} = \{aE \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + bE \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} | a, b, \in \mathbb{R}\} = span\{\begin{pmatrix} a_1 \end{pmatrix}, \begin{pmatrix} a_3 \end{pmatrix}\}$$

i.e. If $L^{-1}PA = U$ *U is one upper triangular matrix*
then C(A)=span{$\begin{pmatrix} a_i \end{pmatrix}$*/i is a pivot column in the rref }*
This is a minimal spanning set for C(A)
minimal size of spanning set = the number of pivot columns

Example 2.18. $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -2 & 3 & 4 \end{bmatrix}$

$$C(A) = span\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \\ 3 \end{pmatrix} \} = span\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} \} = \{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} | b_3 - 2b_2 + 5b_1 = 0\}$$



Picturially:

2.4.3 Column space and Null space

Example 2.19. *Consider*

$$\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 4 & 8 & 12 \\ 3 & 6 & 7 & 13 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \vec{b}$$

Want: *1.Find column space ,null space of A*
2.Find conditon on s.t. the equation is consistent
3.Find general solutions for $Ax = \begin{bmatrix} 0 \\ 6 \\ -6 \end{bmatrix}$
what we have to do is Gaussian Elimination

$$\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 2 & 4 & 8 & 12 & b_2 \\ 3 & 6 & 7 & 13 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & -2 & -2 & b_3 - 3b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Let say we stop at: $\begin{bmatrix} 1 & 2 & 3 & 5 & b_1 \\ 0 & 0 & 2 & 2 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix}$

$$C(A) = span\{(a_1), (a_3)\} = span\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \\ 7 \end{pmatrix}\} = \{\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} | b_3 + b_2 - 5b_1 = 0\}$$

For null space: two free variables x_2 *and* x_4
Hence

$$N(A) = \{\begin{pmatrix} -2x_2 - 2x_4 \\ x_2 \\ -x_4 \\ x_4 \end{pmatrix} | x_2, x_4 \in \mathbb{R}\} = span\{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}\}$$

or, looking at the rref:

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 4b_1 - 2/3b_2 \\ 0 & 0 & 1 & 1 & 1/2b_2 - b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_2 - 5b_1 \end{bmatrix}$$

For $b=\begin{pmatrix}0\\6\\-6\end{pmatrix}$*,we find the particular solution*

$$\begin{bmatrix}1&2&3&5&b_1\\0&0&2&2&b_2-2b_1\\0&0&0&0&b_3+b_2-5b_1\end{bmatrix}=\begin{bmatrix}1&2&3&5&0\\0&0&2&2&6\\0&0&0&0&0\end{bmatrix}$$

Set free variables $x_2=0=x_4$

$$\begin{aligned}x_1+3x_3&=0\\2x_3&=6\end{aligned}\to x_3=3,x_1=-9$$

Therefore

$$\vec{x_p}=\begin{bmatrix}-9\\0\\3\\0\end{bmatrix}$$

Summary:A:$m\times m$ matrix

A$\times\vec{x}=\vec{b}$

There exist **Gaussian elimination** s.t.

$$\mathbf{LP}\begin{bmatrix}A&b\end{bmatrix}=T(upper\ triangular\ matrix)$$

Column space :

1.span by column of A = pivot columns

2. given by $l_j(\vec{b})=o$,for those j which correspond to zero row in Row Echelon form of A

Summary:

1.variables correspond to non-pivot columns(free)

2.variables correspond to pivot columns(not free)

3.Solving $A\vec{x}=\vec{b}$,solution is $\vec{x_n}+\vec{x_p}$

4.Find$\vec{x_p}$: set free variables be any constant

Remark 18. *1. Permuting rows maybe needed in the Elimination, since our focus is not LU factorization, we don't need to fix the permutation P at the beginning.*

2. It's enough to reduce to U(upper triangular matrix), for solving the equation, finding C(A),N(A).

3.We have the equation

the number of pivot columns= the number of free variables of N(A)=n

2.4.4 Rank of matrix

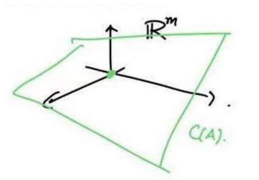
Definition 2.10. *1.The rank of a* $m\times n$ *matrix Ais the no. of pivots in its row echelon form.*

2. it is written as rk(A) or rank(A)

Property 8. *1.$rk(A)min(m, n)$ for $m \times n$ matrix A*
2.For $n \times n$ matrix A

$$Anon-singular \iff Ainvertible \iff rk(A) = n$$

3.$rk(AB) \leq min(rk(A), rk(B))$



Geometrically:
C(A) is a subspace of R^m($sayAis$m$\times n$ matrix
rk(A)="dimension" of C(A)=minimal no. of columns needed for spanning C(A)

2.4.5 Row space

Example 2.20. *suppose* $A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}$ *and we have*
$C(A^T) = span\{\begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ 9 \\ 7 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 3 \\ 4 \end{bmatrix}\}$ *by definition*
Row operation:

$$A \to U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Clain: $C(A^T) = span\{\begin{bmatrix} 1 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \\ 3 \end{bmatrix}\}$
$dim(C(A^T)) = the\ number\ of\ pivot\ rows = rk(A)$
Row operation means: $EA = U$ for invertible E, then we have $A^T E^T = U^T$ and $A^T = U^T (E^T)^{-1}$

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \\ 3 & 9 & 3 \\ 2 & 7 & 4 \end{bmatrix} \times \begin{bmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 0 & 0 \\ 3 & 3 & 0 \\ 2 & 3 & 0 \end{bmatrix}$$

$$\to \begin{cases} e_{11}\vec{a_1} + e_{12}\vec{a_2} + e_{13}\vec{a_3} = \vec{u_1} \\ e_{21}\vec{a_1} + e_{22}\vec{a_2} + e_{23}\vec{a_3} = \vec{u_2} \\ e_{31}\vec{a_1} + e_{32}\vec{a_2} + e_{33}\vec{a_3} = \vec{u_3} \end{cases} \to span\{\vec{u_1}, \vec{u_2}, \vec{u_3}\} \in span\{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$$

Similialy

$$A^T = U^T (E^T)^{-1} \rightarrow span\{\vec{u_1}, \vec{u_2}, \vec{u_3}\} \in span\{\vec{a_1}, \vec{a_2}, \vec{a_3}\}$$

Therefore:

$$C(A^T) = C(U^T)$$

Fact: In general

$$C(AB) \in C(A)$$

Proof: write $A = \begin{bmatrix}\vec{a_1}, ..., \vec{a_n}\end{bmatrix}$,then the i-th column of $AB = A\vec{b_i}$

$$A\vec{b_i} = b_{1i}\vec{a_1} + b_{2i}\vec{a_2} + ... + b_{ni}\vec{a_n}$$

s.t. every column in AB lies in C(A)

2.4.6 Left nullspace

Definition 2.11. *the lefr nullspace of A is defined by $N(A^T)$*

$$N(A^T) = \{\vec{y} | \vec{y} \in \mathbb{R}^m, A^T\vec{y} = 0\}$$

so we can think of $N(A^T) as space of row vectors s.t. A = \begin{bmatrix}0 & . & . & . & . & 0\end{bmatrix}$ *from the left*$dim(V(A^T)) + dim(C(A)) = m$
using the earlier discussion:

$$C(A) \perp N(A^T) \text{ and } C(A) + N(A^T) = \mathbb{R}^m$$

$N(A^T)$ is NOT preserved by row operation, but related as follows:

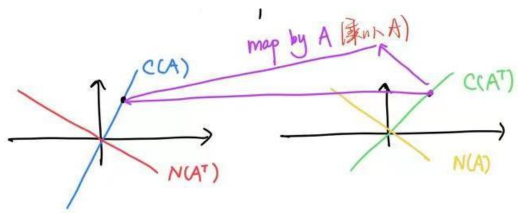
$$EA = U = \begin{bmatrix} * & & & \\ & * & & \\ & & * & \\ & & & \end{bmatrix}$$

$$\rightarrow A^T E^T = U^T = \begin{bmatrix} * & & \\ & * & \\ & & * \\ & & \end{bmatrix}$$

Example 2.21. $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ *with* $m = n = 2$

$$C(A) = span\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\} N(A^T) = span\{\begin{bmatrix} -3 \\ 1 \end{bmatrix}\}$$

$$C(A^T) = span\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\} N(A) = span\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\}$$



2.4.7 Left inverse and Right inverse

Recall that: If A is $n \times n$ matrix, the following are equivalent
1.A is invertible.
2.$A\vec{x} = \vec{0}$ has only as solution, or $N(A) = \{\vec{0}\}$
3.$rk(A) = 0$ or $C(A) = \mathbb{R}^N$
A is invertible: we have B s.t.

$$AB = BA = I_n$$

and we say $B = A^{-1}$ (because inverse is unique)
For $m \times n$ matrix A with $m \neq n$, at best we have only one-side inverse.(and not unique)
there are two case that these inverse exist: 1. $m > n = rk(A)$: left inverse exists i.e.$BA = I_n$ 2. $rk(A) = m < n$: right inverse exists i.e.$AB = I_m$

Example 2.22. $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$ *with* $rk(A) = 2$ *has right inverse*
$C = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{bmatrix}$ *it can't has left inverse.*

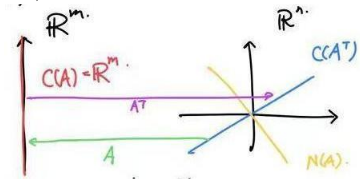
Mind: A way to find right inverse $C = A^T(AA^T)^{-1}$
Symbolically: $AC = (AA^T)(AA^T)^{-1} = I_m$

Example 2.23. $A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$, $AA^T = \begin{bmatrix} 16 & 0 \\ 0 & 25 \end{bmatrix}$

$$C = \begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1/16 & 0 \\ 0 & 1/25 \end{bmatrix} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ 0 & 0 \end{bmatrix}$$

is one of the right inverse.

Q: Why (AA^T) is invertible?

Picturially: 

at least we see that $rk(AA^T) = m$ in this case.
To write down the argument:$AA^T\vec{x} = 0$
then

$$\vec{x^T}AA\vec{x} = 0 \rightarrow < A^T\vec{x}, A^T\vec{x} >= 0 \rightarrow A^T\vec{x} = 0 \rightarrow \vec{x} = 0$$

$N(AA^T) = 0$ and it is a square matrix $\rightarrow$ *it is invertible*.

For the discussion on left inverse: $m \times n$ matrix A with $m > n = rk(A)$, once again we can use transpose A^T to get that $A^T(A^T)^T = A^TA$ is invertible. Taking transepose again of $A^TC = I_n$, we have $C^TA = I_n$, with $C^T = (A^TA)^{-1}A^T$

Summery:

1.$C(A) = \mathbb{R}^n$
2.$N(A) = \{0\}$, or columns are l.i.
3.$C(A^T) = \mathbb{R}^n$
4.$N(A^T) = \{0\}$, or rows are l.i.
5.It's row Echelon form has n pivots.
6.It has left inverse.
7.It has right inverse.

2.5 Linear transformation

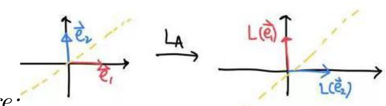
2.5.1 Introduction

Recall that we can think of a $m \times n$ matrix A as an assignment/map

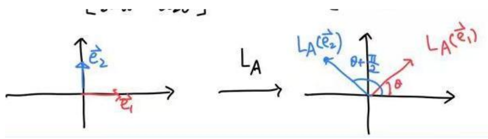
$$L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

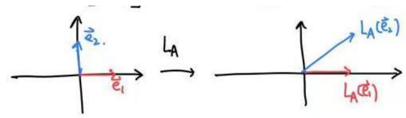
sending a vector $\vec{x} \in \mathbb{R}^n$ to the vector $A\vec{x}$ in $\mathbb{R}^m$
we write $L_A(\vec{x}) = A\vec{x}$

Example 2.24. *1.* $A = cI,\ L_I(\vec{x}) = c\vec{x}$
2.$A = 0_{m\times n},\ L_0(\vec{x}) = \vec{0}$
3.$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $L_A(\vec{e_2}) = \vec{e_1},\ L_A(\vec{e_1}) = \vec{e_2}\ L_A(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$

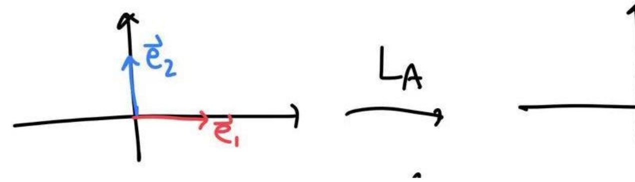
Picture: 

It gives a reflection along the 45^o line
4.$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, $L_A(\vec{e_1}) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$, $L_A(\vec{e_2}) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$





5. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ $L_A(\vec{e_1}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $L_A(\vec{e_2}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



6. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $L_A(\vec{e_1}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $L_A(\vec{e_2}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

observe that for $A = \begin{bmatrix} \vec{a_1}, \vec{a_2}, ..., \vec{a_n} \end{bmatrix}$, by definition 1.$L_A(\vec{e_i}) = \vec{a_i}$ is the i-th column of A

2.L_A is determined by its value at $\vec{e_i}$'s, more explicitly

$$L_A(\begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ . \\ c_n \end{bmatrix} = A \begin{bmatrix} c_1 \\ c_2 \\ . \\ . \\ . \\ . \\ c_n \end{bmatrix} = c_1\vec{a_1} + c_2\vec{a_2} + ... + c_n\vec{a_n}$$

$$= c_1 L_A(\vec{e_1}) + c_2 L_A(\vec{e_2}) + ... + c_n L_A(\vec{e_n}) = L_A(c_1\vec{e_1} + c_2\vec{e_2} + ... + c_n\vec{e_n})$$

3.$L_A(\vec{c}) \in C(A)$ for any $\vec{c} \in \mathbb{R}$, $C(A)$ is also called the image of A.

Definition 2.12. *A map f from a set S_1 to another set S_2 is an assignment of element x in S_1 to an element $f(x)$ in S_2*
we write

$$f : S_1 \longrightarrow S_2$$

for a map

Example 2.25. *let S_1 be the set of students in linear algebra 7 class, $S_2 = 0, 1, ..., 100$, let $f : S_1 \longrightarrow S_2$ sending student x to his/her score in HW1. f is a map.*

Definition 2.13. *Given two vector space V, W, a linear map/tranformation is a map,$L : V \longrightarrow W$*
satisfying:

$$L(a\vec{x} + b\vec{y}) = aL(\vec{x}) + bL(\vec{y})$$

Example 2.26. *we learn that* $P_3 = \{a_0 + a_1x + a_2x^2 + a_3x^3 | a_i \in \mathbb{R}\}$*, the space of degree 3 polynomials is a vector space. difine:*

$$D : P_3 \longrightarrow P_3$$

be

$$D(a_0+a_1x+a_2x^2+a_3x^3) = (d/dx)(a_0+a_1x+a_2x^2+a_3x^3) = a_1+2a_2x+3a_3x^2$$

It is a linear tranformation:

$$\begin{cases} D(cf(x)) = cD(f(x)) \\ D(g(x)f(x)) = D(g(x)) + D(f(x)) \end{cases}$$

we let $m : P_3 \longrightarrow P_4$ *given by multiplication* $m(f(x)) = xf(x)$*, then explicitly:*

$$m(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0x + a_1x^2 + a_2x^3 + a_3x^4$$

let $D = x(d^2/dx^2), D(f(x)) = x(d^2f(x)/d(x^2))$

$$D(a_0+a_1x+a_2x^2+a_3x^3) = x(d/dx)(a_1+2a_2x+3a_3x^2) = x(2a_2+6a_3x) = 2a_2x+6a_3x^2$$

Example 2.27. *Integration:* $I : P_3 \longrightarrow P_4$ *defined by* $I(f(x)) = \int_0^x f(x)dt$ *explicitly:*

$$I(a_0 + a_1x + a_2x^2 + a_3x^3) = a_0x + (a_1/2)x^2 + a_2/3x^3 + (a_3/4)x^4$$

$$I(cf(x)) = \int_0^x cf(x)dt = c\int_0^x f(x)dt = cI(f(x))$$

$$I(f(x)+g(x)) = \int_0^x (f(x)+g(x))dt = \int_0^x f(x)dt+\int_0^x (gx)dt = I(f(x))+I(g(x))$$

Composition

Property 9. *let* $f : V \longrightarrow W$,$g : W \longrightarrow U$ *be linear transformation, then* $g \circ f : V \longrightarrow U$ *is a linear transformation*
$g \circ f$ *is defined as* $g \circ f(\vec{v}) = g(f(\vec{v}))$
Proof:

$$\begin{aligned} g \circ f(a\vec{v_1} + b\vec{v_2}) &= g(f(a\vec{v_1} + b\vec{v_2})) \\ &= g(af(\vec{v_1}) + bf(\vec{v_2})) \\ &= g(a\vec{w_1} + b\vec{w_2}) = ag(\vec{w_1}) + bg(\vec{w_2}) \\ &= ag(f(\vec{v_1})) + bg(f(v_2)) \\ &= a(g \circ f)(\vec{v_1}) + b(g \circ f)(v_2) \end{aligned}$$

Example 2.28. *we consider $D = x(d^2/dx^2)$,we can think of $f = d^2/dx^2 : P_3 \longrightarrow P_1$, and $g : P_1 \longrightarrow P_2$, then we have*

$$\begin{aligned}(g \circ f)(a_0 + a_1x + a_2x^2 + a_3x^3) &= g(f(a_0 + a_1x + a_2x^2 + a_3x^3)) \\ &= g(2a_2 + 6a_3x) \\ &= 2a_2x + 6a_3x^2 \\ &= D(a_0 + a_1x + a_2x^2 + a_3x^3)\end{aligned}$$

Back to matrix:$A : m \times n$ matrix,$Bn \times r$ matrix.

$$L_B : \mathbb{R}^r \longrightarrow \mathbb{R}^n, L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

$$L_A \circ L_B : \mathbb{R}^r \longrightarrow \mathbb{R}^m$$

which is given by

$$L_A \circ L_B(\vec{X}) = L_A(L_B(\vec{X})) = L_A(B\vec{X})) = A(B\vec{X}) = (AB)\vec{X} = L_{AB}(\vec{X})$$

2.5.2 Matrix representation

If we look at $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ associated to a matrix, it is determined by $L_A(\vec{e_i})$ i.e.it action on a basis $\vec{e_1}, ..., \vec{e_n}$

Property 10. *Any linear transformation $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is given by $L = L_A$ for some $m \times n$ matrix A*

Proof:*(say $n = 3, m = 2$), write* $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_1\vec{e_1} + c_2\vec{e_2} + c_3\vec{e_3}$

$$L(c_1\vec{e_1} + c_2\vec{e_2} + c_3\vec{e_3}) = c_1L(\vec{e_1} + c_2L(\vec{e_2} + c_3L(\vec{e_3}$$

set $A = \left[L(\vec{e_1}), L(\vec{e_2}), L(\vec{e_3})\right]$ be the matrix.

for general$L : V \longrightarrow W$ between two vector space, let say,$\vec{v_1}, ..., \vec{v_n}$: a basis for V;$\vec{w_1}, ..., \vec{w_n}$: a basis for W.
(say $n = 3, m = 2$) we can write every vector $\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + a_3\vec{v_3}$

$$L(\vec{c} = L(a_1\vec{v_1} + a_2\vec{v_2} + a_3\vec{v_3}) = a_1L(\vec{v_1} + a_2L(\vec{v_2} + a_3L(\vec{v_3}$$

L is uniquely determined by its value on a basis $\vec{v_1}, \vec{v_2}, \vec{v_3}$
let's go back to the situation that $L : V \longrightarrow W$ with basis $\vec{v_1}, \vec{v_2}, \vec{v_3}$ for V, and $\vec{w_1}, \vec{w_2}$ for W.
we know L is determined by $L(\vec{v_1}), L(\vec{v_2}), L(\vec{v_3})$, let's further write:

$$L(\vec{v_1}) = a_{11}\vec{w_1} + a_{21}\vec{w_2}$$

$$L(\vec{v_2}) = a_{12}\vec{w_1} + a_{22}\vec{w_2}$$

$$L(\vec{v_3}) = a_{13}\vec{w_1} + a_{23}\vec{w_2}$$

or

$$\left[L(\vec{v_1}), L(\vec{v_2}), L(\vec{v_3})\right] = \left[\vec{w_1}, \vec{w_2}\right] \times \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

finding basis $\vec{v_i}$'s for V, and $\vec{w_i}$'s for W, the matrix $[L]$ encode all information about L.if $\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3}$

$$\begin{aligned} L(\vec{v}) &= c_1 L(\vec{v_1}) + c_2 L(\vec{v_2}) + c_3 L(\vec{v_3}) \\ &= \left[L(\vec{v_1}), L(\vec{v_2}), L(\vec{v_3})\right] \times \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \\ &= \left[\vec{w_1}, \vec{w_2}\right] \left[L\right] \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \end{aligned}$$

Example: Let's consider $D = d/dx : P_3 \Longrightarrow P_3$ be the linear transformation. We fixed a basis $1(\vec{v_1}\vec{w_1}), x(\vec{v_2}\vec{w_2}), x^2(\vec{v_3}\vec{w_3}), x^3(\vec{v_4}\vec{w_4})$ for P_3

$$D(\vec{v_1}) = 0 = 0 \cdot \vec{w_1} + 0 \cdot \vec{w_2} + 0 \cdot \vec{w_3} + 0 \cdot \vec{w_4} = \left[\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\right] \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$D(\vec{v_2}) = 1 = 1 \cdot \vec{w_1} + 0 \cdot \vec{w_2} + 0 \cdot \vec{w_3} + 0 \cdot \vec{w_4} = \left[\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\right] \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$D(\vec{v_3}) = 2x = 0 \cdot \vec{w_1} + 2 \cdot \vec{w_2} + 0 \cdot \vec{w_3} + 0 \cdot \vec{w_4} = \left[\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\right] \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

$$D(\vec{v_4}) = 3x^2 = 0 \cdot \vec{w_1} + 0 \cdot \vec{w_2} + 3 \cdot \vec{w_3} + 0 \cdot \vec{w_4} = \left[\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\right] \begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

we collect the column vector representing

$$L(v_i) = \left[\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\right] \left[a_i\right]$$

we put $\left[a_i\right]$ as column of $\left[L\right]$: $\left[L\right] = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix}$

In the example $\left[L\right] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Let say we want to find $D(1+x-3x^2+5x^3)$ we can write

$$1+x-3x^2+5x^3 = \begin{bmatrix}\vec{v_1}, \vec{v_2}, \vec{v_3}, \vec{v_4}\end{bmatrix}\begin{bmatrix}1\\1\\-3\\5\end{bmatrix}$$

we multiply $\begin{bmatrix}1\\1\\-3\\5\end{bmatrix}$ to $\begin{bmatrix}L\end{bmatrix}$ from R.H.S:

$$\begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{bmatrix}\begin{bmatrix}1\\1\\-3\\5\end{bmatrix} = \begin{bmatrix}1\\-6\\15\\0\end{bmatrix}$$

then

$$\begin{bmatrix}\vec{w_1}, \vec{w_2}, \vec{w_3}, \vec{w_4}\end{bmatrix}\begin{bmatrix}1\\-6\\15\\0\end{bmatrix} = 1-6x+15x^2$$

is the resulting vector.

Property 11. *Given V and W as above, with $\vec{v_1}, ..., \vec{v_n}$ basis for V. Fix $\vec{l_1}, ..., \vec{l_n}$ be n vectors in W.*
there exist unique $L : V \Longrightarrow W$ linear transformations, such that $L(\vec{v_i}) = \vec{l_i}$)
Proof:*(say $n = 3$): Given $\vec{v} \in V$, we can write $\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + a_3\vec{v_3}$ uniquely(because of basis).*
Defined $L(\vec{v}) = a_1\vec{l_1} + a_2\vec{l_2} + a_3\vec{l_3}$, we check for linearity.

Example 2.29. *Let $M_{2\times 2}(\mathbb{R})$ be the space of 2×2 matrices, let $A = \begin{bmatrix}1&1\\0&1\end{bmatrix}$ and we let $L : M_{2\times 2}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$, given by $L(B) = B \cdot A$*
It is a linear transformation

$$\begin{aligned}L(c_1B_1 + c_2B_2) &= (c_1B_1 + c_2B_2)A\\ &= c_1B_1A + c_2B_2A\\ &= c_1(B_1A) + c_2(B_2A)\\ &= c_1L(B_1) + c_2L(B_2)\end{aligned}$$

we take the basis
$\begin{bmatrix}1&0\\0&0\end{bmatrix}(\vec{v_1}, \vec{w_1})$, $\begin{bmatrix}0&1\\0&0\end{bmatrix}(\vec{v_2}, \vec{w_2})$, $\begin{bmatrix}0&0\\1&0\end{bmatrix}(\vec{v_3}, \vec{w_3})$, $\begin{bmatrix}0&0\\0&1\end{bmatrix}(\vec{v_4}, \vec{w_4})$ *for $M_{2\times 2}(\mathbb{R})$*

$$L(\vec{v_1}) = \begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix} = \begin{bmatrix}1&1\\0&0\end{bmatrix} = \vec{w_1} + \vec{w_2} = \begin{bmatrix}, \vec{w_1}\vec{w_2}, \vec{w_3}, \vec{w_4}\end{bmatrix}\begin{bmatrix}1\\1\\0\\0\end{bmatrix}$$

$$L(\vec{v_2}) = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = \vec{w_2} = \left[,\vec{w_1}\vec{w_2},\vec{w_3},\vec{w_4}\right]\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$L(\vec{v_3}) = \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\1 & 1\end{bmatrix} = \vec{w_3} + \vec{w_4} = \left[,\vec{w_1}\vec{w_2},\vec{w_3},\vec{w_4}\right]\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

$$L(\vec{v_4}) = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = \vec{w_4} = \left[,\vec{w_1}\vec{w_2},\vec{w_3},\vec{w_4}\right]\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

$\left[L\right] = \begin{bmatrix}1 & 0 & 0 & 0\\1 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 1 & @\end{bmatrix}$ *is the matrix representation.*

Q: What if we chose a basis in $\mathrm{R}^m differfromthestandardbasis,.....,$?

Example 2.30. *we recall that the map $L_P : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ with $\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$*

$$L_P(\vec{v_1}) = \vec{e_2} = \vec{w_1} = \left[\vec{w_1},\vec{w_2}\right]\begin{bmatrix}1\\0\end{bmatrix}$$

$$L_P(\vec{v_2}) = \vec{e_1} = \vec{w_2} = \left[\vec{w_1},\vec{w_2}\right]\begin{bmatrix}0\\1\end{bmatrix}$$

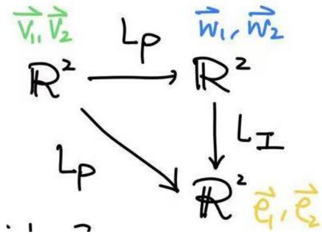
therefore the matrix representation is $\left[L_P\right] = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
What happened here is the change of basis: $\vec{w_1} = \vec{e_2}, \vec{w_2} = \vec{e_1}$, consider $L_I : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be the identity map.

$$L_I(\vec{w_1}) = \vec{e_2} = \left[\vec{e_1}, vece_2\right]\begin{bmatrix}0\\1\end{bmatrix}$$

$$L_I(\vec{w_2}) = \vec{e_1} = \left[\vec{e_1}, vece_2\right]\begin{bmatrix}1\\0\end{bmatrix}$$

$\left[L_I\right] = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$, *so we consider the composition:*

$$L_I \circ L_P = L_P \longrightarrow \left[L_I\right] \circ \left[L_P\right] = \left[L_P\right]$$



let $V \xrightarrow{L} W \xrightarrow{Z}$ be a composition of linear map, with a corresponding choice of basis, then we have the identity

$$_{\vec{z_i}}\left[M\right]_{\vec{w_i}} \cdot_{\vec{w_i}} \left[L\right]_{\vec{v_i}} =_{\vec{z_i}} \left[M \circ L\right]_{\vec{v_i}}$$

In working, after the basis are fixed according the composition of linear map correspond to matrix multiplication.

We may look at the change of basis in this way: say

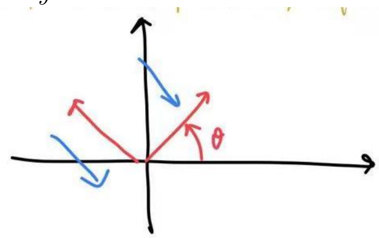
$$L : V \longrightarrow W$$

then in particular we have

$$V \xrightarrow{id} V \xrightarrow{L} W$$

and we get $_{\vec{w_i}}\left[L\right]_{\vec{u_i}} \cdot_{\vec{u_i}} \left[id\right]_{\vec{v_i}} =_{\vec{w_i}} \left[L\right]_{\vec{v_i}}$

Example 2.31. *we may use this relation to write down some linear transfor-*



mation:want to find

the projection onto the $\theta rotation of e_i$*, take a second basis* $\vec{w_1} = \vec{v_1} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \vec{w_2} = \vec{v_2} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$*, and then under this basis the Projection* $P,_{\vec{w_i}} \left[P\right]_{\vec{v_i}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \iff P\vec{v_i} = \vec{w_i}, P\vec{v_0} = 0$

In our earlier discussion:

$$\mathbb{R}^2(\vec{e_1}, \vec{e_2}) \xrightarrow{id} \mathbb{R}^2(\vec{v_1}, \vec{v_2}) \xrightarrow{P} \mathbb{R}^2(\vec{v_1}, \vec{v_2}) \xrightarrow{id} \mathbb{R}^2(\vec{e_1}, \vec{e_2})$$

$$_{\vec{e_i}}\left[P\right]^{\vec{e_a}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

That is we $R \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R^{-1} = \left[P\right]$, give us the matrix representation of the projection P.

$$\left[P\right]^2 = (R \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R^{-1})(R \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R^{-1}) = R \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R^{-1} = \left[P\right]$$

Indeed we can read it in the other way $R{-}1\left[P\right]R = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$

suppose we want to write down reflection along the line, we can take a basis $\vec{v_1} = \begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}, \vec{v_2} = \begin{bmatrix}-\sin\theta\\\cos\theta\end{bmatrix}$

$$_{\vec{e_i}}\left[id\right]_{\vec{v_i}} \cdot_{\vec{v_i}} \left[H\right]_{\vec{v_i}} \cdot_{\vec{v_i}} \left[id\right]_{\vec{e_i}} =_{\vec{e_i}} \left[H\right]_{\vec{e_i}}$$

$$_{\vec{e_i}}\left[F\right]_{\vec{e_i}} = \begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\theta & \sin\theta\\\sin\theta & -\cos\theta\end{bmatrix} = \begin{bmatrix}\cos^2\theta-\sin^2\theta & 2\cos\theta\sin\theta\\2\cos\theta\sin\theta & \sin^2\theta-\cos^2\theta\end{bmatrix}$$

2.5.3 Image and Kernel

Definition 2.14. *Giveb a linear transformation $L : V \longrightarrow W$, we define*
1.$ker(L) = \{\vec{v}|\vec{v} \in V, L(\vec{v}) = 0\}$
2.$Im(L) = \{\vec{w}|\vec{w} \in W, there\ exist\ \vec{v} \in V s.t.L(\vec{v}) = \vec{w}\}$

Example 2.32.
$L{=}L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m, then ker(L_A) = \{\vec{v} \in \mathbb{R}^n|L_A(\vec{v}) = 0\} = N(A)$

Im$(L_A) = \{\vec{w} \in \mathbb{R}^m|\vec{w} = L_A(\vec{v}) for some \vec{v} \in \mathbb{R}^n\} = C(A)$

Now with $L : V \longrightarrow W$ equipped with basis $\vec{v_1}, ..., \vec{v_n}$ for V, $\vec{w_1}, ..., \vec{w_m}$ for W, matrix representation:(say $m = 2, n = 3$) write $\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3}$, then

$L(\vec{v}) = \left[\vec{w_1}, \vec{w_2}\right]\left[L\right]\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}$

$N([L]) \iff Ker(L)$ given by the 1-1 correspondence

$\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} \to \left[\vec{v_1}, \vec{v_2}, \vec{v_3}\right]\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3}$

Summary: 1.$C([L])$

$\iff$ Im(L)$isa1 - 1correspondence\begin{bmatrix}a_1\\a_2\end{bmatrix} \to \left[\vec{w_1}, \vec{w_2}\right]\begin{bmatrix}a_1\\a_2\end{bmatrix} = a_1\vec{w_1} + a_2\vec{w_2}$

2.$Ker(L) \in$V$isasubspace$

$Im(L) \in$W$isasubspace$

3. $rk([L]) = dim(C([L])) = dim(Im(L))$

4.$dim(Ker(L)) + dim(Im(L)) = dim(V)$

2.5.4 Isomorphism

Definition 2.15. *$f : V \longrightarrow W$ linear map is an isomorphism(, if we can find another $g : W \longrightarrow V$ be a linear map s.t.$g \circ f = id_V, f \circ g = id_W$*

Example 2.33. *$f: L_A: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ for invertible A is isomorphism.*
$g: \mathbb{R}^n \longrightarrow \mathbb{R}^n$, then we have $g \circ f(B) = g(A \cdot B) = A^{-1}AB = B$,
$f \circ g(B) = A(A^{-1}B) = B$

Example 2.34. *Recall that we have the vector space $W = \mathbb{R}_+$ with*
$x + y = x \cdot y$, $c \cdot x = x^2$, in fact it is isomorphism to $(\mathbb{R}, +, \cdot)$
we take $f: \mathbb{R} \longrightarrow \mathbb{R}$ with $f(a) = e^a$; $g: \mathbb{R}_+ \longrightarrow \mathbb{R}$ with $g(x) = log(x)$
Claim:f, g are linear and $f \circ g = id_{\mathbb{R}_+}$, $g \circ f = id_{\mathbb{R}}$
$L: V \longrightarrow W$, with basis $\vec{v_1}, ...\vec{v_n}$ for V, $\vec{w_1}, ..., \vec{w_n}$ for W
Claim:L is an isomorphism if and only if $[L]$ is invertible.
Proof: *Assume L is an isomorphism, then we have $T: W \longrightarrow V$ s.t.*
$T \circ L = id_V$, $L \circ T - id_W$, then we take their matrix representation:

$$_{\vec{v_i}}[T]_{\vec{w_i}} \cdot_{\vec{w_i}} [L]_{\vec{v_i}} =_{\vec{v_i}} [id]_{\vec{v_i}} = I_{n \times n}$$

and

$$[L] \cdot [T] =_{\vec{w_i}} [id]_{\vec{w_i}} = Im \times m$$

$\rightarrow$m=nand$[L]^{-1} = [T]$
if $[L]$ is invertible, then $m = n$ and we have $[L]^{-1}$, we difine $T: W \longrightarrow V$ for
$\vec{w} = a_1\vec{w_1} + a_2\vec{w_2} + ... + a_n\vec{w_n}$ by

$$T(\vec{w}) = \begin{bmatrix}\vec{v_1}, \vec{v_2}, ..., \vec{v_n}\end{bmatrix} \begin{bmatrix}L\end{bmatrix}^{-1} \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{bmatrix}, i.e.[T] = [L]^{-1}$$

Then $_{\vec{v_i}}[T \circ L]_{\vec{v_i}} = I_{n \times n},_{\vec{w_i}} [L \circ T]_{\vec{w_i}} = I_{n \times n} \rightarrow Tistheinverseof L$

2.5.5 Summary

1.Linear transformation $L: V \longrightarrow W$ (with a choice of basis) $\iff$ m$\times n$ $[L]$
2.Composition $\iff$ $matrixmultiplication$
3.Isomorphism $\iff$ $Invertible$
4.$Im(L)$ $\iff$ C([L])
5.$Ker(L)$ $\iff$ N([L])
6.$dim(Im(L))$ $\iff$ rk([L])

Chapter 3

Orthogonality

In high school, we learned about vertical in geometry. In linear algebra, we also find algebraic properties similar to vertical, which we call **orthogonality**().

3.1 Orthogonal Vectors and Subspaces

The basis of linear algebra is vectors ,matrices and Subspaces, and we'll start with vectors to understand the properties of orthogonality.

3.1.1 Orthogonal Vectors()

Definition 3.1. *Let $u = (a_1, a_2 \cdots, a_n)^T$ and $v = (b_1, b_2 \cdots, b_n)^T$ be vectors of a vector space $\mathbb{R}^n$.Then the **inner product()**of u and v is defined as*

$$\langle u, v\rangle = u^T v = a_1 b_1 + a_2 b_2 + \cdots a_n b_n$$

For A and B are $n \times n$ matrices ,we define:

$$\langle A, B\rangle = Trace(A^T B)$$

Property 1. $\langle u, v\rangle = \langle v, u\rangle$.

Property 2. $\langle cu, v\rangle = c\langle v, u\rangle$.

Property 3. $\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle$.

Definition 3.2. *For a vector x, we define the length of the vector:*

$$\|x\| = \langle x, x\rangle^{1/2}$$

Property 1. $\|x\|$ *= 0 if and only if* $x = 0$

Property 2. $\|x\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y}\rangle = \|\vec{x} - \vec{y}\|^2$.

Proof:$\|x-y\|^2 = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots (x_n - y_n)^2$
$= x_1^2 + x_2^2 + \cdots x_n^2 + y_1^2 + y_2^2 + \cdots y_n^2 - 2(x_1y_1 + x_2y_2 + \cdots x_ny_n)$
$= \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle$
Now we can define what is orthogonality for vectors and matrices.

Definition 3.3. *We define v and u are* ***orthogonal*** *if:*

$$\langle v, u\rangle = 0$$

We can see if v and u are **orthogonal**, from property 2.2,
$\|u-v\|^2 = \|u\|^2 + \|v\|^2$.So they are vertical in geometry.

Lemma. *If* ***nonzero*** *vectors* $v_1, v_2 \cdots, v_n$ *are* ***mutually orthogonal****(i.c.every vector is perpendicular to every other)(), then they are linearly independent.()*

Proof:If we want to prove they are **mutually orthogonal**,we need to prove for any $c_i \in \mathbb{R}$, if $c_1v_1 + c_2v_2 + \cdots c_nv_n = 0$,$c_i = 0$

This kind of thought is very important: we multiply both sides of this equation by v_i for $i \in [1, n]$.

$$v_i \times 0 = 0 = v_i \times (c_1v_1 + c_2v_2 + \cdots c_nv_n)$$

And for v_iv_j,if $i \neq j$, $v_iv_j = 0$.The equation will be:

$$0 = c_i \|v_i\|^2$$

Because v_i is a **nonzero** vector, so $\|v_i\|^2 \neq 0$, so $c_i = 0$ for $i \in [1, n]$, then we can prove $v_1, v_2 \cdots, v_n$ are **mutually orthogonal**

Tips:If we want to find a orthogonal vector of other vectors $v_1, v_2, \ldots v_m$, we need to make a matrix $A = [v_1, v_2, \ldots v_m]$, and find it's N(A), the basis of it is want we want.

Example 3.1. *Find all vectors perpendicular to* $v = [1, 1]^T$*, if v is small, we can assume*

$$w = [a, b]^T, w^Tv = 0$$

a + b = 0 ,b = -a , $w = a[1, -1]^T$*, for* $a \in \boldsymbol{R}^T$

Example 3.2. *Find all vectors in* $\boldsymbol{R}^3$ *that are orthogonal to* $v = [1, 1, 1]^T$ *and* $v = [1, -1, 0]^T$*.If vector is more than two, we need to form a matrix.*

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}, Ax = 0$$

x is what we want, the row simplest form matrix is

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \end{bmatrix}, N(A) = span \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

Finally : $w = a[1, 1, -2]^T$*, for* $a \in \boldsymbol{R}^T$

Example 3.3. *Find all vectors that are perpendicular to* $v_1 = [1, 4, 4, 1]^T$ *and* $v_2 = [2, 9, 8, 2]^T$*.*

$$A = \begin{bmatrix} 1 & 4 & 4 & 1 \\ 2 & 9 & 8 & 2 \end{bmatrix}, Ax = 0$$

x is what we want, the simplest form matrix is

$$A = \begin{bmatrix} 1 & 0 & 4 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

the null space of the matrix is

$$N(A) = span \begin{bmatrix} -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Finally : $w = a[-4, 0, 1, 0]^T + b[-1, 0, 0, 1]^T$*, for* $a, b \in \boldsymbol{R}^T$

Example 3.4. *Show that xy is orthogonal to x+y if and only if* $\|x\| = \|y\|$*.*

$$\langle x + y, x - y \rangle = \|x\|^2 - \|y\|^2 = 0$$

then $\|x\| = \|y\|$

Example 3.5. *Prove* $A^T A$ *has the same nullspace as* A*.*
As we all know, $\langle Ax, Ax \rangle = 0$ *if and only if Ax = 0, assume* $x \in N(A)$*, Ax = 0,then* $A^T Ax = 0$*,* $x^T A^T Ax = 0 = \langle Ax, Ax \rangle$
As we all know, $\langle Ax, Ax \rangle = 0$ *if and only if Ax = 0, so* $N(A) = N(A^T A)$

3.1.2 Orthogonal Subspaces

In the previous study, we found that for the four basic subspaces of a matrix, they will have some special properties

Recall 1. *We take a matrix A $(m \times n)$ for L_A :$\mathbb{R}^n$ —>$\mathbb{R}^m$,We find*

$$x \in C(A) and y \in N(A^T)\langle x, y\rangle = 0, C(A) + N(A^T) = \mathbb{R}^m$$

$$x \in C(A^T) and y \in N(A)\langle x, y\rangle = 0, C(A^T) + N(A) = \mathbb{R}^n$$

Proof:take $x \in C(A^T), y \in N(A), z \in \mathbb{R}^m$.

$$x = A^T, Ay = 0$$

$$\langle x, y\rangle = \langle A^T z, y\rangle = (A^T z)^T y = z^T Ay = 0$$

Definition 3.4. *For any $x \in W_1$ and $y \in W_2$(W_1, W_2 are subspaces and $W_1, W_2 \in \mathbb{R}^n$),if $\langle x, y\rangle = 0$,we call they are orthogonal().*

$$W_1 \perp W_2$$

*If $W_1 + W_2 = \mathbb{R}^n$, we call them **orthogonality complement**()*

$$W_1 = W_2^{\perp}$$

Sofor a matrix $C(A^T)$ and $N(A)$ are orthogonality complement, $C(A)$ and $N(A^T)$ are orthogonality complement,too.

fact 1: if $\langle x, y_i\rangle = 0$ for a basis $y_1 \cdots y_n$ of $\mathbf{V}$, then $\langle x, y\rangle = 0$ for all $v \in \mathbf{V}$

fact 2:if $W \subseteq V$ be subspaces, dimW = dimV —-> W = V

Theorem 3.1. *As a **linear transformation**(),each matrix A transforms its row space $C(A^T)$ to its column space $C(A)$.*

proof: assume A is $m \times n$ matrix, $C(A^T), N(A) \in \mathbb{R}^n$, for any $z \in \mathbb{R}^n, z = v_1 + v_2, (v_1 \in C(A^T), v_2 \in N(A))$, so Az can be written as:

$$A(v_1 + v_2) = Av_1 + Av_2 = Av_1$$

Av_1 is in $C(A)$, so A transforms its row space $C(A^T)$ to its column space $C(A)$.

Note:Every vector $b \in C(A)$comes from exactly one vector $x \in C(A^T)$.

proof:There are many problems of falsification(), and we generally adopt the idea of proof by contradiction.

assume x_1, x_2 can transform to b by A.

$$A(x_1 - x_2) = b - b = 0$$

However $x_1 - x_2 \in C(A^T)$,if $A(x_1 - x_2) = 0$, then $(x_1 - x_2) \in N(A)$. $C(A^T) \bigcap N(A) = 0 = x_1 - x_2$, so $x_1 = x_2$

Example 3.1. *Put bases for the orthogonal subspaces V and W into the columns of matrices V and W. Why does V^TW = zero matrix?*

if V^TW = zero matrix, $V^TWx = 0, x \in \boldsymbol{R}^n$.

$$V^T = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_i \end{bmatrix}, W = \begin{bmatrix} w_1, w_2, \dots, w_j \end{bmatrix}$$

for any $x \in \boldsymbol{R}^n$,

$$\begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_i \end{bmatrix} w_k = 0, k \in \boldsymbol{R}$$

*So $Wx \in$ **C(W)**, so $V^TWx = 0$ for $x \in \boldsymbol{R}^n$, then V^TW = zero matrix.*

Example 3.2. *Suppose S is spanned by the vectors [1,2,2,3] and [1,3,3,2]. Find two vectors that span $S^\perp$.*

we know vectors in $S^\perp$ are orthogonal with vectors in S. So we can assume v_1, v_2 can span $S^\perp$.So the question is actually solving Ax = 0.

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix}, Av_1 = Av_2 = 0$$

As we learn before, the row simplest form matrix is :

$$A --- > \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 1 & -1 \end{bmatrix}, v_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} v_2 = \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Example 3.3. *Suppose an n by n matrix is invertible: $AA1 = I$. Then the first column of $A1$ is orthogonal to the space spanned by which rows of A?*

assume:

$$A = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}, A^-1 = \begin{bmatrix} w_1, w_2, \dots, w_n \end{bmatrix}$$

From the question, we can know :

$$A = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} \times w_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

So we can know:the first column of $A1$ is orthogonal to the space spanned by$v_2, v_3, \dots, v_n$

Example 3.4. *If P is the plane of vectors in $\boldsymbol{R}^4$ satisfying $x_1+x_2+x_3+x_4=0$, write a basis for $P^{\perp}$.*

$v=[x_1,x_2,x_3,x_4]^T, v\in P, x_i\in \boldsymbol{R}$, so we can find v can be written as:

$$v=\begin{bmatrix}-x_2-x_3-x_4\\x_2\\x_3\\x_4\end{bmatrix}=x_2\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}+x_3\begin{bmatrix}-1\\0\\1\\0\end{bmatrix}+x_4\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$$

to find $P^{\perp}$, we need to find N(A), for:

$$A=\begin{bmatrix}-1&1&0&0\\-1&0&1&0\\-1&0&0&1\end{bmatrix}, N(A)=span\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=P^{\perp}=N(A)$$

3.2 Cosines and Projections onto Lines

3.2.1 Cosines

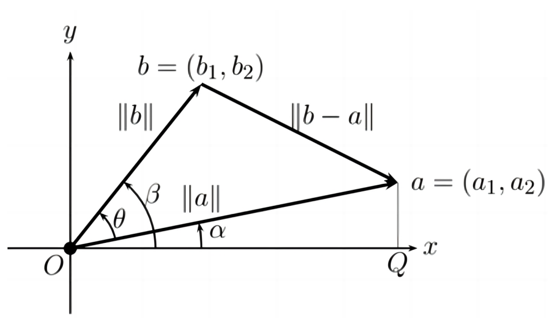
If $x^Ty=0$,then x,y are orthogonal,also called perpendicular().

The orthogonal case is the most important.

Now we allow inner products that are not zero,and angles that are not right angles.

$$\cos\theta=\frac{\langle\vec{x},\vec{y}\rangle}{\|x\|\,\|y\|}, if\vec{x},\vec{y}\neq\vec{0}$$

.



proof:$\cos\theta=\cos\beta\cos\alpha-\sin\beta\sin\alpha=\frac{a_1b_1}{\|a\|\|b\|}+\frac{a_2b_2}{\|a\|\|b\|}=\frac{\vec{a}^T\vec{b}}{\|a\|\|b\|}$

since: $|\cos\theta|\leq 1$, $\|a\|\,\|b\|\geq\left|a^Tb\right|$

Example 3.5. *In n dimensions, what angle does the vector $a=[1,1,...,1]^T$ make with the coordinate axes()?*

assume an unit vector on coordinate axes is e_i.

$$\langle\vec{a},\vec{e_i}\rangle=1, \|a\|\,\|e_i\|=\sqrt{n}$$

$$\cos\theta = \frac{\langle \vec{a}, \vec{e_i} \rangle}{\|a\| \, \|e_i\|} = \frac{1}{\sqrt{n}}$$

Example 3.6. *By choosing the correct vector b , prove that*

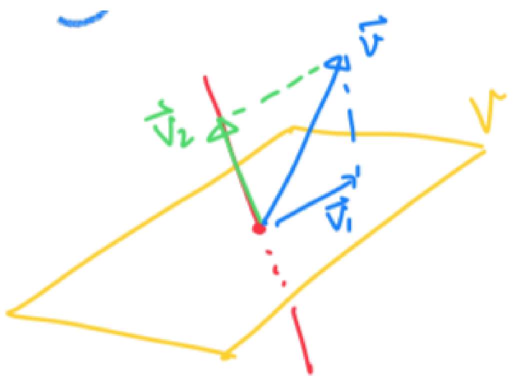
$$(a_1 + \cdots + a_n)^2 \leq n(a_1^2 + \cdots + a_n^2)$$

we choose $b = [1, \ldots, 1]^T$, $a = [a_1 + \cdots + a_n]^T$,*we can know* $\|a\| \, \|b\| \geq \left|a^T b\right|$, *So*

$$\|a\|^2 \, \|b\|^2 = n(a_1^2 + \cdots + a_n^2) \geq \left|a^T b\right|^2 = (a_1 + \cdots + a_n)^2$$

3.2.2 Projection one a line

Now, if we want to find the distance from a point b to the line in the direction of the vector a, what we need to do?



fact:Any vector $\in \mathbf{R}^n$ can be uniquely written as $v = v_1 + v_2 (v_1 \in \mathbf{V}, v_1 \in \mathbf{V}^T)$

Definition 3.5. *We define the orthogonal projection()* $\boldsymbol{P}$:$\boldsymbol{V}$ *—->* $\boldsymbol{V}$, *by* $\boldsymbol{P}(v)$ $= v_1$

Now we need to find c in:$\mathbf{P}(v) = pv$

1. step1: Find unit vector(), of v_1, $v_1' = \frac{v_1}{\|v_1\|}$. For any vector v $\mathbf{P}$(v) = cv_1'
2. step2: Find c, $v_2 = v - \mathbf{P}(v) = v - cv_1'$
   $\langle v_2, v_1' \rangle = 0 = \langle v, v_1' \rangle - \langle cv_1', v_1' \rangle$, and $\langle v_1', v_1' \rangle = 1$
   $$\langle v, v_1' \rangle = c^T = c$$
3. step3: $\mathbf{P}$(v) = c$v_1' = v_1'^T \times v \times v_1'$, and $v_1'^T \times v$ is a number, so $v_1'^T \times v \times v_1' = v_1'^T v_1' v$

4. step4: Finally, we can find the matrix [P]: $v_1'^T v_1' = \frac{v_1^T v_1}{\|v_1\|^2}$
   the matrix can transform any vector on the line

we find that [P] have some interesting properties.

Property 1 (P)**.** $= [P]^T$, *because* $(v_1^T v_1)^T = v_1^T v_1$.

Property 2. $[P]^2 = [P]$
proof:*(1)[P][P]*$v = [P]v_1 = v_1$
or we can see it in other way.
(2)

$$[P]^2 = \frac{v_1^T v_1 v_1^T v_1}{\|v_1\|^4}$$

$$v_1 v_1^T = \|v_1\|^2, [P]^2 = \frac{v_1^T v_1}{\|v_1\|^4} v_1 v_1^T = \frac{v_1^T v_1}{\|v_1\|^2} = [P]$$

Property 3. *For a matrix [P] transforms vectors to* $\boldsymbol{V}$, *matrix I-[P] can transforms vectors to* $\boldsymbol{V}^T$
proof:$v - v_1 = v_2 = v(I - [P])$

Example 3.7. *Is the projection matrix P invertible? Why or why not?*
we know that $P^2 = P$, *if P is invertible,* $P = I$, *so obviously P isn't invertible.*

Example 3.8. *Find the projection matrix P_1 onto the line through $a = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and also the matrix P_2 that projects onto the line perpendicular to a.*

we all know $P_1 = \frac{1}{10}\begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$

if we want to find P_2 that projects onto the line perpendicular to a, we didn't need to find vector perpendicular to a.

$P_2 = I - P_1 = \frac{1}{10}\begin{bmatrix} 9 & -3 \\ -3 & 1 \end{bmatrix}$

3.2.3 Projection one a subspace

Now we need to learn how to find a projection onto V?

Definition 3.6. *$q_1, \ldots, q_k$ are in $\boldsymbol{R}^n$, they are* ***mutually orthogonal****() and $\langle q_i, q_i \rangle = 1$.*

If they are basis for $\boldsymbol{V} \in \boldsymbol{R}^n$, we call them are a ***orthonormal basis****() of* $\boldsymbol{V}$

At that time, P(v) can be written as:

$$P(v) = \langle v, q_1 \rangle q_1 + \langle v, q_2 \rangle q_2 + \ldots \langle v, q_k \rangle q_k$$

check:$v_2 = v - \langle v, q_1 \rangle q_1 + \langle v, q_2 \rangle q_2 + \ldots \langle v, q_k \rangle q_k, \in \mathbf{V}^T$
we need to check $\langle v_2, q_i \rangle = 0$ for all $i = 1, \ldots, k$

$$\langle v_2, q_i \rangle = \langle (v - \langle v, q_1 \rangle q_1 + \langle v, q_2 \rangle q_2 + \ldots \langle v, q_k \rangle q_k), q_i \rangle$$

$$= \langle v, q_i \rangle - \langle v, q_i \rangle \langle q_i, q_i \rangle = 0$$

So $[P] = q_1 q_1^T + q_2 q_2^T + \ldots q_k, q_k^T$

Tips:After we define inner product, we can define **Transpose**() in a new way.

Definition 3.7. *we find that*

$$\langle Ax, y \rangle = \langle x, A^T y \rangle$$

if any matrix satisfy $\langle Ax, y \rangle = \langle x, By \rangle$ we say:$B = A^T$

Property 1. *there is only one transpose matrix of a matrix.*

proof*:$\langle Ax, y \rangle = \langle x, By \rangle$ for not only one B,$x^T A^T y = x^T By$, then $x^T(A^T - B)y = 0 for x, y \in \boldsymbol{R}^n$, so $A^T - B = 0, A^T = B$*

Property 2. $(AB)^T = B^T A^T$

proof*:$\langle ABx, y \rangle = \langle Bx, A^T y \rangle = \langle x, A^T B^T y \rangle$, so $(AB)^T = B^T A^T$*

Example 3.9. *Let A be a 9 Œ 5 matrix whose rank is 4.*

(a) Find the rank of A^TA.

If we want to find rank of A^TA, we need to find dim(N(A^TA)), we know N(A^TA) = N(A), and dim(N(A)) = n - dim(c(A)).

So N(A^TA) = N(A) = 1, rank of A^TA = dim(c(A^TA)) = n - dim(N(A)) = 4

(b) Show that the equation Ax = b has solution if and only if the orthogonal projection of b on the the left nullspace $N(A^T)$ is 0.

Any b can be written as b = v + w (v $\in N(A^T)$,w $\in C(A)$). w is orthogonal projection of b on the the left nullspace $N(A^T)$, so w can not be written as Ax = w, if w $\neq 0$, $Ax = b$ can not be written.

Example 3.10. *Project the vector b onto the line through a. Check that e = b-$\vec{p}$ is perpendicular to a or not.*

$$b = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

we need to find matrix $P = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, *the vector* $\vec{p} = Pb = \frac{5}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

so $e = \begin{bmatrix} -\frac{2}{3} \\ \frac{4}{3} \\ -\frac{2}{3} \end{bmatrix}$, *we can find* $e \times a = 0$

Example 3.11. *Find a 3Œ3 matrix P whose associated linear transformation $L(v) = Pv$ is the orthogonal projection onto V, for $V = spanv_1, v_2$.*

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

if we have two vectors are mutually vertical, we can find P easily.

$$P_1 = \frac{v_1^T v_1}{\|v_1\|^2} = \frac{1}{6} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

then $v_{2p} = v_2 - v_2 \times P_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}$

$$P = P_1 + P_2 = \frac{v_1^T v_1}{\|v_1\|^2} + \frac{v_{2p}^T v_{2p}}{\|v_{2p}\|^2} = \frac{1}{14} \begin{bmatrix} 5 & 6 & -3 \\ 6 & 10 & 2 \\ -3 & 2 & 13 \end{bmatrix}$$

3.3 Projection and Least Squares

3.3.1 Review

Recall 2. *In the section before, we have learnt the inner product of the vectors. For $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have*

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

, which we call it the inner product of $\vec{x}, \vec{y}$.

The inner product has some properties.

Property 1. $\|x\|^2 + \|y\|^2 - 2\langle \vec{x}, \vec{y} \rangle = \|\vec{x} - \vec{y}\|^2$.

Property 2. $\cos\theta = \frac{\langle \vec{x}, \vec{y} \rangle}{\|x\| \|y\|}$, *if* $\vec{x}, \vec{y} \neq \vec{0}$.

Then, we define the concept of orthogonality.

Definition 3.8. $\vec{x} \perp \vec{y}$, *if* $\langle \vec{x}, \vec{y} \rangle = 0$.

3.3.2 Projection as Linear Transformation

After the definition of orthogonality, we wonder the projection.

Question. *$V \subseteq$ subspace, How we write down the orthogonal projection onto V ?*

We need to define what is the projection.

Abstractly, if $V \subseteq \mathbb{R}^n$, then V satisfies that $V + V^\perp = \mathbb{R}^n$ and $V \cap V^\perp = \{0\}$.

Any vector $\vec{v} \in \mathbb{R}^n$ can be **uniquely** decomposed as $\vec{v} = \vec{v_1} + \vec{v_2}$, where $\vec{v_1} \in V$, $\vec{v_2} \in V^\perp$.

Definition 3.9. *We define the orthogonal projection $P : V \longrightarrow V$ by $P(\vec{v}) = \vec{v_1}$.*

check. *(linearity) Let $\vec{v}, \vec{w} \in \mathbb{R}^n, \vec{v} = \vec{v_1} + \vec{v_2}, \vec{w} = \vec{w_1} + \vec{w_2}$. Then $c\vec{v} + d\vec{w} = (c\vec{v_1} + d\vec{w_1}) + (c\vec{v_2} + d\vec{w_2})$, where $c\vec{v_1} + d\vec{w_1} \in V$ and$c\vec{v_2} + d\vec{w_2} \in V^\perp$. Then $P(c\vec{v} + d\vec{w}) = c\vec{v_1} + d\vec{w_1}$.*

As a linear transformation, P has some properties.

Property 3. *$P^2 = P$ as linear transformation.*

Proof. $P(\vec{v}) = \vec{v_1} \Longrightarrow P(P(\vec{v})) = P(\vec{v_1}) = \vec{v_1}$. ☐

Property 4. *$\tilde{P}(\vec{v}) = \vec{v} - P(\vec{v})$ gives projection onto $V^\perp$.*

3.3.3 Projection Matrix obtained by Transition

We wonder how to express a exact linear transformation of projection in the form of matrix.

Question. *How to write down matrix w.r.t. basis $\vec{e_i}'s$?*

Step 3.3.3.1. *Find a basis of V, say $\vec{v_1}, \dots, \vec{v_k}$.*

Step 3.3.3.2. *Let $A = \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_k} \\ | & & | \end{bmatrix}$, $N(A) = V^\perp$, find a basis of $V^\perp$, say $\vec{v_{k+1}}, \dots, \vec{v_n}$.*

Step 3.3.3.3. ${}_{\vec{v_i}}\left[P\right]_{\vec{v_i}} =$

$$\begin{bmatrix} 1_1 & & & & & & & \\ & 1_2 & & & & & O & \\ & & \ddots & & & & & \\ & & & 1_k & & & & \\ & & & & 0_1 & & & \\ & & & & & 0_2 & & \\ & & O & & & & \ddots & \\ & & & & & & & 0_{n-k} \end{bmatrix}$$

Step 3.3.3.4. *Transition matrix* ${}_{\vec{e_i}}\left[id\right]_{\vec{v_i}} = \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_n} \\ | & & | \end{bmatrix} = S$

Step 3.3.3.5. *At last, ${}_{\vec{e_i}}\left[P\right]_{\vec{e_i}} = S_{\vec{v_i}}\left[P\right]_{\vec{v_i}} S^{-1}$ is the matrix expression of the projection linear transformation.*

3.3.4 Projection Matrix onto 1-dimensional Subspace

We want more geometric way of writing it. First, we consider the easy case. When the dimension of V equals 1, it implies that we are projecting onto a single straight line.

In this case, $dimV = 1$, i.e. $V = Span\{\vec{v}\}$ for $\vec{v} \neq \vec{0}$.

First, we need to normalize $\vec{v}$, write $\hat{v} = \frac{1}{\|\vec{v}\|}\vec{v}$. Then, we have $\|\hat{v}\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| = 1$.

For any other vector $\vec{x}$, the projection of $\vec{x}$ onto V is parallel to $\hat{v}$. i.e. $P(\vec{x}) = c \cdot \hat{v}$ for a c needed to find out.

In high school, we say in $\mathbb{R}^2$, and we have the formula $P(\vec{x}) = \langle \vec{x}, \hat{v} \rangle \hat{v}$.

We observe that this formula holds for any dimension.

Claim. *$P(\vec{x}) = \langle \vec{x}, \hat{v} \rangle \hat{v} = \frac{\langle \vec{x}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v}$ is the required projection.*

Given any $\vec{x}$, write $\vec{x} = (\vec{x} - \langle \vec{x}, \hat{v} \rangle \hat{v}) + \langle \vec{x}, \hat{v} \rangle \hat{v}$. Then $\vec{x} - \langle \vec{x}, \hat{v} \rangle \hat{v} \in V^{\perp}$, and $\vec{x} - \langle \vec{x}, \hat{v} \rangle \hat{v} \in V$.

check. *Whether $\vec{x} - \langle \vec{x}, \hat{v} \rangle \hat{v}$ is in $V^{\perp}$?*

We have $\langle \vec{x} - \langle \vec{x}, \hat{v} \rangle \hat{v}, \hat{v} \rangle = \langle \vec{x}, \hat{v} \rangle - \langle \vec{x}, \hat{v} \rangle \langle \hat{v}, \hat{v} \rangle = \langle \vec{x}, \hat{v} \rangle - \langle \vec{x}, \hat{v} \rangle = 0$

Example 3.12. *Let $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, compute the projection of $\vec{x}$ onto $Span\{\vec{v}\}$.*

$\|\vec{v}\|^2 = 1^2 + 1^2 + 1^2 = 3$.

$\langle \vec{x}, \vec{v} \rangle = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 = 6$.

$P(\vec{x}) = \frac{\langle \vec{x}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v} = 2 \cdot \vec{v} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$, *which is independent on the choice of $\vec{v}$.*

Question. *How above the matrix correspond to projection onto $Span\{\vec{v}\}$?*

$$P(\vec{x}) = \frac{\langle \vec{x}, \vec{v} \rangle}{\|\vec{v}\|^2}\vec{v} = \frac{(\vec{v}^T \cdot \vec{x})}{\|\vec{v}\|^2}\vec{v} = \vec{v}\frac{(\vec{v}^T \cdot \vec{x})}{\|\vec{v}\|^2} = \frac{\vec{v} \cdot \vec{v}^T \cdot \vec{x}}{\|\vec{v}\|^2} = (\frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2})\vec{x}$$

Thus, the projection matrix is given by $\frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2}$, which is a $n \times n$ matrix of $rk = 1$.

Example 3.13. $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\frac{1}{\|\vec{v}\|^2} = \frac{1}{3}$, $\vec{v} \cdot \vec{v}^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, *then*

$$\left[P\right] = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

.

We can observe that for the projection P onto a line, $\begin{bmatrix}P\end{bmatrix}$ is a symmetric matrix.

Proof. $\begin{bmatrix}P\end{bmatrix}^T = (\frac{\vec{v}\cdot\vec{v}^T}{\|\vec{v}\|^2})^T = \frac{\vec{v}\cdot\vec{v}^T}{\|\vec{v}\|^2} = \begin{bmatrix}P\end{bmatrix}$ □

Indeed, any projection P onto V has its matrix representation under $e_i's$ being symmetric.(Prove later)

Example 3.14. *We may check the formula for projection onto the θ-line in $\mathbb{R}^2$.*

Take $v \neq 0$ parallel to the line with $v = \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}$.

$$\|\vec{v}\|^2 = \cos^2\theta + \sin^2\theta = 1$$

$$\vec{v}\cdot\vec{v}^T = \begin{bmatrix}\cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta\end{bmatrix}$$

As a result, $\begin{bmatrix}P\end{bmatrix} = \begin{bmatrix}\cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta\end{bmatrix}$, which agrees with the one we get using transition matrix.

3.3.5 General Projection

How to write down a general projection onto V?

Definition 3.10. *$\vec{q_1},\ldots,\vec{q_k}$ be a sequence of vectors in $\mathbb{R}^n$, it is said to be orthonormal() if*

$$\langle\vec{q_i},\vec{q_j}\rangle = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

If further it is a basis for $V \subseteq \mathbb{R}^n$, we call it a orthonormal basis for V.

Example 3.15. $\vec{e_1} = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\vec{e_2} = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $\vec{e_3} = \begin{bmatrix}0\\0\\1\end{bmatrix}$ *is an orthonormal basis for* $\mathbb{R}^3$.

Example 3.16. $\vec{v_1} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$, $\vec{v_2} = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$ *is an orthonormal basis for* $\mathbb{R}^2$.

Let say given V, and orthonormal basis $\vec{q_1},\vec{q_2},\ldots,\vec{q_k}$.(we will learn how to find it later)

We can write the orthogonal projection onto V:

$$P(v) \coloneqq \langle v,\vec{q_1}\rangle\,\vec{q_1} + \langle v,\vec{q_2}\rangle\,\vec{q_2} + \cdots + \langle v,\vec{q_k}\rangle\,\vec{q_k} \in V$$

check.

$$v_2 = v - \langle v,\vec{q_1}\rangle\,\vec{q_1} - \cdots - \langle v,\vec{q_k}\rangle\,\vec{q_k} \in V^{\perp}?$$

We only need to check $\langle v_2,\vec{q_i}\rangle = 0$ for all $i = 1,\ldots,k$.

$$\langle v_2,\vec{q_i}\rangle = \langle v - \langle v,\vec{q_1}\rangle\,\vec{q_1} - \cdots - \langle v,\vec{q_k}\rangle\,\vec{q_k},\vec{q_i}\rangle = \langle v,\vec{q_i}\rangle - \langle v,\vec{q_i}\rangle\,\langle\vec{q_i},\vec{q_i}\rangle = 0$$

Suppose we want to write down matrix representation of P w.r.t. $e_i's$.

$$\big[P\big] = \vec{q}_1\vec{q}_1^T + \vec{q}_2\vec{q}_2^T + \cdots + \vec{q}_k\vec{q}_k^T$$

We can observe that $\big[P\big]^T = \big[P\big]$, i.e. it is a symmetric matrix.

Example 3.17. $\vec{v}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}$, *$Span\{\vec{v}_1, \vec{v}_2\} = V$, please find the projection matrix onto V.*

Because $\langle\vec{v}_1, \vec{v}_2\rangle = 0$, $q_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}$, $q_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \begin{bmatrix}0\\0\\1\end{bmatrix}$

$\big[P\big] = \vec{q}_1\vec{q}_1^T + \vec{q}_2\vec{q}_2^T = \begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & 1\end{bmatrix}$

3.3.6 Least Square and another Projection Matrix

Once we can measure length, we can talk about approximation.

Given an inconsistent equation: $A\vec{x} = \vec{b}$ has no solution. i.e. $\vec{b}$ is outside $C(A)$.

We want best approximation: $\left\|A\vec{x} - \vec{b}\right\|^2$ attain minimum.

We want to find a vector $A\hat{x}$ s.t. it is closest to $\vec{b}$.

Let $A\hat{x} = \vec{b}_1$, which is the orthogonal projection onto $C(A)$.

Then $\vec{b} - A\hat{x}$ is in $C(A)^\perp = N(A^T)$.

Let $e = \vec{c} - A\hat{x}$, where e for error vector.

That means we have:

$$A^T(\vec{b} - A\hat{x}) = 0$$

i.e. $A^TA\hat{x} = A^T\vec{b}$, which we called it the **Normal Equation** and it is always solvable.

The solution $\vec{x} = \hat{x}$ would be a minimizer of $\left\|A\vec{x} - \vec{b}\right\|^2$.

Here A^TA is **cross-product** matrix $n \times n$ matrix.

It has some properties.

Property 5. *A^TA is symmetric.*

Property 6. $N(A^TA) = N(A)$

Proof. For product of matrices, we generally have $N(A) \subseteq N(BA)$.

Let $B = A^T$, we have $N(A) \subseteq N(A^TA)$

Now, we only need to show $N(A^TA) \subseteq N(A)$.

Take $\vec{x} \in N(A^TA)$, by definition $A^TA\vec{x} = 0$.

$$\Longrightarrow \vec{x}^TA^TA\vec{x} = 0 \Longrightarrow (A\vec{x})^TA\vec{x} = 0$$

$$\Longrightarrow \|A\vec{x}\|^2 = 0 \Longrightarrow A\vec{x} = 0 \Longrightarrow \vec{x} \in N(A)$$

□

Tip: This is a useful trick. Please memorize it. Similarly, we have $N(AA^T) = N(A^T)$.

Property 7. *If $N(A) = 0 \Longrightarrow A^TA$ invertible. Then $\hat{x}$ has unique solution $\hat{x} = (A^TA)^{-1}A^Tb$.*

Example 3.18. $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, *the equation $A\vec{x} = \vec{b}$ has no solution.*

$C(A) = Span\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}\}$ *not containing $\vec{b}$.*

Then, we consider the approximation. i.e. Solving $A^TA\vec{x} = A^T\vec{b}$.

$A^T\vec{b} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ 23 \end{bmatrix}$

$N(A) = \{0\}$, $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$

$(A^TA)^{-1} = \frac{1}{\det(A^TA)} \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$

$\Longrightarrow \hat{x} = (A^TA)^{-1}A^Tb = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 23 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Come back to the Normal Equation $A^TA\hat{x} = A^T\vec{b}$.

Recall that $A\hat{x}$ is the projection of $\vec{b}$ onto $C(A)$.

$P(b) = A\hat{x} = A(A^TA)^{-1}A^T\vec{b}$, where $A(A^TA)^{-1}A^T$ is the matrix for projection onto $C(A)$.

Let $\begin{bmatrix} P \end{bmatrix} = A(A^TA)^{-1}A^T$, it has following properties.

Property 8. $\begin{bmatrix} P \end{bmatrix}^2 = A(A^TA)^{-1}A^TA(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = \begin{bmatrix} P \end{bmatrix}$

Property 9. $\begin{bmatrix} P \end{bmatrix}^T = (A(A^TA)^{-1}A^T)^T = A((A^TA)^{-1})^TA^T = A((A^TA)^T)^{-1}A^T = A(A^TA)^{-1}A^T = \begin{bmatrix} P \end{bmatrix}$

Claim. *P is a projection iff $P^2 = P$ and $P^T = P$.*

Proof. (for the converse) Assuming $P^2 = P$ and $P^T = P$.

Consider $\vec{v} = P\vec{v} + (\vec{v} - P\vec{v})$.

Take $P(\vec{v} - P\vec{v}) = P\vec{v} - P^2\vec{v} = 0$.

Hence $\vec{v} - P\vec{v} \in N(P) = C(P)^T$ and $P\vec{v} \in C(P)$.

Therefore, P is a projection onto $C(P)$. □

Given $V \subseteq \mathbb{R}^n$, there are two ways to obtain projection matrix.

1. Find an orthogonal basis $\vec{q}_1, \dots, \vec{q}_k$ of V. Then $P = \vec{q}_1\vec{q}_1^T + \vec{q}_2\vec{q}_2^T + \cdots + \vec{q}_k\vec{q}_k^T$.
2. Find a basis $\vec{v}_1, \dots, \vec{v}_k$ of V, and let $A = [\vec{v}_1, \dots, \vec{v}_k] : \mathbb{R}^k \longrightarrow \mathbb{R}^n$ with $C(A) = V$. Then $P = A(A^TA)^{-1}A^T$

Both satisfy the two properties of projection matrix.

3.3.7 The Application of Least Square

Question. *What is **Data fitting**?*

Say we have a input quantity t and measured quantity b.
We have guessed the formula relating them without knowing the coefficient().

Example 3.19. $b = C + Dt$

Example 3.20. $b = Ce^{-\lambda t} + De^{-\mu t}$

Example 3.21. $b = Const + Dt^2$

If we are exactly right about their relation and exactly accurate when we do measurement $\Longrightarrow$ two measurement would determine C, D.
However, in reality, it is hoping too much and there must be error.
Want: Find the best fitted coefficient to minimize error.

Example 3.22. *Assuming you are a money manager:*
Say $b =$ price of gasoline, and $t =$ price of crude oil.
Our intuition() may tell us they are related by a linear formula: $b = C + Dt$
You try to collect a lot of historical data.
***Q**: How to find the best fit line to the data set?*
Say we have 3 data:

- $b = 1$ *at* $t = -1$
- $b = 1$ *at* $t = 1$
- $b = 3$ *at* $t = 2$

Then we have the equation:
$C + D(-1) = 1$
$C + D(1) = 1$
$C + D(2) = 3.$
Or in the matrix way: $\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$
***Rk**: More data $\Longrightarrow$ bigger matrix*
During the data fitting, we will minimize the length of $\vec{e} = \vec{b} - A\vec{x}$.
$\|\vec{e}\|^2 = [1 - (C + D(-1))]^2 + [1 - (C + D)]^2 + [3 - (C + 2D)]^2$
From our earlier discussion: we need to solve $A^TA\hat{x} = A^T\vec{b}$.

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}; \; A^T b = \begin{bmatrix} 5 \\ 6 \end{bmatrix};$$

$$\hat{x} = (A^T A)^{-1} A^T \vec{b} = \begin{bmatrix} 9/7 \\ 4/7 \end{bmatrix}.$$

Summary: (For fitting a straight line), say $b = C + tD$.

If we are given $(t_1, b_1), \ldots, (t_m, b_m)$ s.t. $t_1, \ldots, t_m$ are distinct.($\Longrightarrow N(A) = \{0\}$)

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum_i t_i \\ \sum_i t_i & \sum_i t_i^2 \end{bmatrix}$$

$$A^T \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum_i b_i \\ \sum_i t_i b_i \end{bmatrix}$$

Equation to Solve: $\begin{bmatrix} m & \sum_i t_i \\ \sum_i t_i & \sum_i t_i^2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum_i b_i \\ \sum_i t_i b_i \end{bmatrix}$

If not fitting a straight line: Say $b = Ce^{-\lambda t} + De^{-\mu t}$, maybe for radioactive material.

$$\left.\begin{aligned} b_1 &= Ce^{-\lambda t_1} + De^{-\mu t_1} \\ b_2 &= Ce^{-\lambda t_2} + De^{-\mu t_2} \\ b_3 &= Ce^{-\lambda t_3} + De^{-\mu t_3} \end{aligned}\right\} \Longrightarrow \begin{bmatrix} e^{-\lambda t_1} & e^{-\mu t_1} \\ e^{-\lambda t_2} & e^{-\mu t_2} \\ e^{-\lambda t_3} & e^{-\mu t_3} \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Rk: We need to linearity in the coefficient, not linearity in t or b.

Rk2: The most difficult part in reality is to guess what is the formula relating b and t.

3.3.8 *Weighted Least Square

Back to the least square example: $\begin{cases} C + D(-1) = 1 \\ C + D(1) = 1 \\ C + D(2) = 3 \end{cases}$

What if we weight each data differently:

$\|\vec{e}\|_w^2 = w_1^2[1 - (C + D(-1))]^2 + w_2^2[1 - (C + D)]^2 + w_3^2[3 - (C + 2D)]^2$

How do we think about it?

$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \vec{e} \in \mathbb{R}^3$: we introduce $W : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the invertible matrix

$$W = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix}$$

$\|\vec{e}\|_w^2 = \langle We, We \rangle_{\mathbb{R}}^3$, $\mathbb{R}^3$ for usual inner product in $\mathbb{R}^3$.

Formulation: If we have W invertible $m \times m$ matrix.

We can look at $\vec{e} = A\vec{x} - \vec{b}$, but consider $\|\vec{e}\|_w^2 = \langle W\vec{e}, W\vec{e}\rangle$.

$\hat{x}$ is a minimizer iff $W\vec{e} = WA\vec{x} - W\vec{b} = \tilde{A}\vec{x} - \tilde{b}$ attains minimum for usual inner product.

That means we are solving for the equation: $\tilde{A}^T\tilde{A}\hat{x} = \tilde{A}^T\vec{b}$ or $A^TW^TWA\hat{x} = A^TW^T\vec{b}$.

We introduce an extended inner product on $\mathbb{R}^m$: $\langle x, y\rangle_w := \langle Wx, Wy\rangle = x^TW^TWy$.

It also satisfies:

1. $\langle x, y\rangle_w = \langle y, x\rangle_w$
2. $\langle cx, y\rangle_w = c\,\langle x, y\rangle_w$
3. $\langle x_1 + x_2, y\rangle_w = \langle x_1, y\rangle_w + \langle x_2, y\rangle_w$
4. $\langle x, x\rangle_w \geq 0$ and equal zero iff $x = 0$

We can repeat the abstract theory: $V + V^w = \mathbb{R}^m$ and define the projection $P : \mathbb{R}^m \longrightarrow V$ using this decomposition.

The relation with matrix is different.

Example 3.23. $\langle Ax, y\rangle_w \neq \langle x, A^Ty\rangle_w$

3.4 Orthogonal Bases and Gram-Schmidt

3.4.1 Review

Recall: What we learn about orthogonality.

Definition 3.11. *$q_1, \ldots, q_k$ are said to be **orthogonal** if $\langle q_i, q_j\rangle = 0$ for $i \neq j$.*

*$q_1, \ldots, q_k$ are said to be **orthonormal** if we further have $\|q_i\| = 1$ for all i.*

Definition 3.12. *If a basis $q_1, \ldots, q_k$ for V is **orthogonal (resp. orthonormal)**, we call it a **orthogonal basis(resp. orthonormal basis)**.*

3.4.2 Orthogonal Matrix

Definition 3.13. *An $n \times n$ matrix Q is said to be **orthogonal** if $Q^TQ = I_n$.*

Attention: The matrix is called orthogonal not orthonormal by tradition.

Example 3.24. $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$*, a rotation matrix, or* $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$*, a reflection matrix.*

The orthogonal matrix has the following properties. Let Q be an $n \times n$ orthogonal matrix.

Property 10. *$Q^{-1} = Q^T$ and hence $QQ^T = I_n$.*

Proof. $Q^TQ = I \implies n = rk(I_n) \leq rk(Q) \implies rk(Q) = n \implies Q$ invertible.
Then $Q^TQ = I \implies Q^TQQ^{-1} = Q^{-1} \implies Q^T = Q^{-1}$.
And then, $I_n = Q^TQ = Q^{-1}Q = QQ^{-1} = QQ^T$. □

Property 11. *Its columns $q_1, \dots, q_n$ are* ***orthonormal basis*** *for $\mathbb{R}^n$*

Proof. (cont'd(continued)) Write $Q = [q_1, \dots, q_n]$:
$rk(Q) = n \iff q_1, \dots, q_n$ a basis for $\mathbb{R}^n$
(say $n = 3$): $Q^T = \begin{bmatrix} q_1^T \\ q_2^T \\ q_3^T \end{bmatrix}$,
$Q^TQ = \begin{bmatrix} q_1^T \\ q_2^T \\ q_3^T \end{bmatrix} \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \langle q_1, q_1 \rangle & \langle q_1, q_2 \rangle & \langle q_1, q_3 \rangle \\ \langle q_2, q_1 \rangle & \langle q_2, q_2 \rangle & \langle q_2, q_3 \rangle \\ \langle q_3, q_1 \rangle & \langle q_3, q_2 \rangle & \langle q_3, q_3 \rangle \end{bmatrix}$, which holds for any $n \times n$ matrix.
$Q^TQ = I \longleftrightarrow q_1, \dots, q_n$ orthonormal. □

Rk: Q^T is also orthogonal if Q is. Then we also have the transpose of rows of Q forms a orthonormal basis for $\mathbb{R}^n$, whichi is not obvioous from definition.

Example 3.25. $Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$ *is an orthogonal matrix.*

Tip: This matrix and its variants are very common when you work on previous semesters final exam papers.

3.4.3 NOT Square Q

What if Q is NOT square:
If Q is $m \times n$ matrix with orthonormal columns $\iff Q^TQ = I_n$.
$n = rk(I_n) \leq rk(Q) \leq m \implies n \leq m$.
$Q = [q_1, \dots, q_n$, with orthonormal column.
$q_1, \dots, q_n$ is an orthonormal basis for $C(Q)$.
In the Normal Equation $\hat{x} = Q^TQ\hat{x} = Q^Tb$
Projection onto $C(Q)$: $P = Q(Q^TQ)^{-1}Q^T = QQ^T \neq I$
(say $n = 3$): $P\vec{x} = QQ^T\vec{x} = [q_1, q_2, q_3] \begin{bmatrix} q_1^T \\ q_2^T \\ q_3^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [q_1, q_2, q_3] \begin{bmatrix} \langle q_1, \vec{x} \rangle \\ \langle q_2, \vec{x} \rangle \\ \langle q_3, \vec{x} \rangle \end{bmatrix} =$
$\langle q_1, \vec{x} \rangle q_1 + \langle q_2, \vec{x} \rangle q_2 + \langle q_3, \vec{x} \rangle q_3$
We have $L_Q : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be the associated linear transformation:

Property 12. $Q^TQ = I_n$ *iff* $\|Q\vec{x}\| = \|\vec{x}\|$ *for all* $\vec{x} \in \mathbb{R}^n$

Proof. $\implies$): $\|Q\vec{x}\|^2 = \vec{x}^TQ^TQ\vec{x} = \vec{x}^T\vec{x} = \|\vec{x}\|^2 \implies \|Q\vec{x}\| = \|\vec{x}\|$
$\impliedby$): if $\|Q\vec{x}\|^2 = \|\vec{x}\|^2$ for all $\vec{x} \implies \vec{x}^TQ^TQ\vec{x} = \vec{x}^T\vec{x} \implies \vec{x}^T(Q^TQ - I)\vec{x} = 0 \implies$ Let $C = Q^TQ - I$, then $\vec{x}^TC\vec{x} = 0$ for all $\vec{x} \in \mathbb{R}^n$ with some symmetric $n \times n$ matrix C.

Input $\vec{x} = \vec{w} + \vec{y}$ for any $\vec{w}, \vec{y} \in \mathbb{R}^n$, we have $(\vec{w} + \vec{y})^T C(\vec{w} + \vec{y} = 0$
$\vec{w}^T C\vec{w} + \vec{y}^T C\vec{y} + \vec{w}^T C\vec{y} + \vec{y}^T C\vec{w} = 0$
$\because$ $\vec{y}^T C\vec{w}$ is a 1×1 matrix and C is symmetric, $\vec{y}^T C\vec{w} = (\vec{y}^T C\vec{w})^T = \vec{w}^T C^T\vec{y} = \vec{w}^T C\vec{y}$
We also have $\vec{w}^T C\vec{w} = \vec{y}^T C\vec{y} = 0$.
Thus, $2\vec{w}^T C\vec{y} = 0 \Longrightarrow \vec{w}^T C\vec{y} = 0$ for all $\vec{w}, \vec{y} \in \mathbb{R}^n$
Trick: choose $\vec{w} = e_i$, $\vec{y} = e_j$, we will have $0 = e_i^T C e_j = C_{ij}$
We have $C = 0 \Longrightarrow Q^T Q = I$ □

3.4.4 Gram-Schmidt Process

Q: Given a subspace $V \subseteq \mathbb{R}^m$, how to find an orthonormal basis?

Example 3.26. *$dim(V) = 1$: take $v \in V$ s.t. $V = Span\{v\}$ then $q_1 = \frac{v}{\|v\|}$ is the basis.*

Example 3.27. *$dim(V) = 2$: take v_1, v_2 a basis for V*

1. *Let $q_1 = \frac{v_1}{\|v_1\|}$*
2. *Let $\tilde{q}_2 = v_2 - \langle v_2, q_1\rangle\, q_1$, and we have $\tilde{q}_2 \neq \vec{0}$*
   Let $q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|}$

Example 3.28. *$dim(V) = 3$: take v_1, v_2, v_3 a basis for V*

1. *Let $q_1 = \frac{v_1}{\|v_1\|}$*
2. *Let $\tilde{q}_2 = v_2 - \langle v_2, q_1\rangle\, q_1$, $q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|}$*
3. *Let $\tilde{q}_3 = v_3 - (\langle v_3, q_1\rangle\, q_1 + \langle v_3, q_2\rangle\, q_2)$, where $\langle v_3, q_1\rangle\, q_1 + \langle v_3, q_2\rangle\, q_2$ is the projection onto $Span\{q_1, q_2\}$, $q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|}$*

Example 3.29. *The Example of Gram-Schmidt Process $v_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}$, $v_2 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $v_3 = \begin{bmatrix}2\\1\\0\end{bmatrix}$*

$$q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix};\ \tilde{q}_2 = v_2 - \langle v_2, q_1\rangle\, q_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{\sqrt{2}}\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{bmatrix} = \begin{bmatrix}1/2\\0\\-1/2\end{bmatrix}$$

$$\|\tilde{q}_2\| = \frac{1}{\sqrt{2}},\ q_2 = \sqrt{2}\begin{bmatrix}1/2\\0\\-1/2\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\0\\-1/\sqrt{2}\end{bmatrix}$$

$q_3 = v_3 - \langle v_3, q_1 \rangle q_1 - \langle v_3, q_2 \rangle q_2$

$$= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right\rangle \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} \right\rangle \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \; q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$Q = [q_1, q_2, q_3] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}, \textit{ satisfying } Q^T Q = I$$

In general: Say we have $v_1, \ldots, v_n$ basis for $V \subseteq \mathbb{R}^n$

Suppose we have already found the first j orthogonal vectors $q_1, \ldots, q_j$ s.t. $Span\{q_1, \ldots, q_j\} = Span\{v_1, \ldots, v_j\}$.

Take $\tilde{q}_{j+1} \coloneqq v_{j+1} - (\langle v_{j+1}, q_1 \rangle q_1 + \cdots + \langle v_{j+1}, q_j \rangle q_j)$, where $\langle v_{j+1}, q_1 \rangle q_1 + \cdots + \langle v_{j+1}, q_j \rangle q_j$ is the projection onto $Span\{q_1, \ldots, q_j\}$.

Take $q_{j+1} \coloneqq \frac{1}{\|\tilde{q}_{j+1}\|} \tilde{q}_{j+1}$

Then $Span\{q_1, \ldots, q_{j+1}\} = Span\{v_1, \ldots, v_{j+1}\}$, with $q_1, \ldots, q_{j+1}$ orthogonal.

3.4.5 The QR Decomposition

What if we compute the transition matrix from the basis $v_i's$ to the orthonormal basis $q_i's$?

Say $n = 3 (m \geq 3)$: $v_1 = \langle v_1, q_1 \rangle q_1$

$v_2 = \langle v_2, q_1 \rangle q_1 + \langle v_2, q_2 \rangle q_2$

$v_3 = \langle v_3, q_1 \rangle q_1 + \langle v_3, q_2 \rangle q_2 + \langle v_3, q_3 \rangle q_3$

Fact: With an orthonormal basis $q_1, \ldots, q_n$, any vector v is written as $v = \langle v, q_1 \rangle q_1 + \cdots + \langle v, q_n \rangle q_n$

($\because P(v) = v$ for $v \in Span\{q_1, \ldots, q_n\}$)

Using the theory of transition matrix:

$[v_1, v_2, v_3] = [q_1, q_2, q_3]_{qi} [id]_{vi}$

$$[v_1, v_2, v_3] = [q_1, q_2, q_3] \begin{bmatrix} \langle v_1, q_1 \rangle & \langle v_2, q_1 \rangle & \langle v_3, q_1 \rangle \\ 0 & \langle v_2, q_2 \rangle & \langle v_3, q_2 \rangle \\ 0 & 0 & \langle v_3, q_3 \rangle \end{bmatrix}$$

$$A = [v_1, v_2, v_3],\ Q = [q_1, q_2, q_3],\ R = \begin{bmatrix} \langle v_1, q_1 \rangle & \langle v_2, q_1 \rangle & \langle v_3, q_1 \rangle \\ 0 & \langle v_2, q_2 \rangle & \langle v_3, q_2 \rangle \\ 0 & 0 & \langle v_3, q_3 \rangle \end{bmatrix}$$

Q, an $m \times n$ matrix, satisfies that $Q^T Q = I$

R is upper triangular invertible matrix.

This is called the **QR decomposition** for the matrix A.

Theorem 3.2. *Any $m \times n$ matrix $A = [v_1, \ldots, v_n]$ with linear independent columns can be written as $A = QR$,with $Q = [q_1, \ldots, q_n]$ $m \times n$ matrix with $Q^T Q = I$, R upper triangular $n \times n$ invertible matrix.*

Example 3.30. *In The Example of Gram-Schmidt Processour previous example:*

$$[v_1, v_2, v_3] = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \ [q_1, q_2, q_3] = \begin{bmatrix} {}^1\!/\!\sqrt{2} & {}^1\!/\!\sqrt{2} & 0 \\ 0 & 0 & 1 \\ {}^1\!/\!\sqrt{2} & -{}^1\!/\!\sqrt{2} & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} \langle v_1, q_1 \rangle & \langle v_2, q_1 \rangle & \langle v_3, q_1 \rangle \\ 0 & \langle v_2, q_2 \rangle & \langle v_3, q_2 \rangle \\ 0 & 0 & \langle v_3, q_3 \rangle \end{bmatrix} = \begin{bmatrix} \sqrt{2} & {}^1\!/\!\sqrt{2} & \sqrt{2} \\ 0 & {}^1\!/\!\sqrt{2} & \sqrt{2} \\ 0 & 0 & 1 \end{bmatrix}$$

Chapter 4

Determinants

4.1 Introduction

Determinant is a very important conception in the linear algebra. In the future, it can help you with several things listed as below.

1. Observe whether a matrix is invertible.
2. Find the inverse of the matrix with the help of cofactor matrix and the determinant.
3. Find the solution of $Ax = b$ by the determinant and the cofactor matrix.
4. Find the eigenvalue of a matrix.
5. Examine a matrix to be positive definite or not.

The application of the determinant is so important that we have to have a deeper understanding on it.
In this chapter, we will give you the algebraic definition of determinant first, after that, we can tell you more about what is the meaning of the determinant especially when it is zero. There is one very important thing which is the geometric view of determinant, from my opinion, it can help you understanding the whole conception more easily.

4.2 Three Basic Algebraic Definition of Determinant

To learn determinant, I hope you can watch it as a mathematical operation first, which will follow certain rules, and the target of the operation is matrices when the output of the operation are numbers.
Now we are going to introduce the three basic rule of the determinant as a algebraic operation.

4.2.1 Normalization

$$\det(I) = 1 \tag{4.1}$$

Example 4.1. $\det \begin{bmatrix} 1 \end{bmatrix} = 1$

Example 4.2. $\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1$

Example 4.3. $\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1$

4.2.2 Alternating

$$\det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = -\det \begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix} \tag{4.2}$$

Example 4.4. $\det \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix} = -\det \begin{bmatrix} 3 & 6 & 9 \\ 1 & 6 & 7 \\ 3 & 7 & 8 \end{bmatrix}$

Example 4.5. $\det \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} = -\det \begin{bmatrix} 3 & 6 & 9 & 2 \\ 1 & 6 & 7 & 9 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}$

Example 4.6. $\det \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} = -\det \begin{bmatrix} 3 & 6 & 9 & 2 & 9 \\ 1 & 6 & 7 & 9 & 7 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix}$

4.2.3 Multi-Linearity

The First Part

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} \tag{4.3}$$

Example 4.7. $\det \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix} = \det \begin{bmatrix} 2 & 4 & 1 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix} + \det \begin{bmatrix} -1 & 2 & 6 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}$

Example 4.8. $\det \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} = \det \begin{bmatrix} 2 & 4 & 1 & 3 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} + \det \begin{bmatrix} -1 & 2 & 7 & 6 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}$

Example 4.9. $\det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}=\det\begin{bmatrix}2&4&1&3&5\\1&6&7&9&7\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}+\det\begin{bmatrix}-1&2&7&6&2\\1&6&7&9&7\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

The Second Part

$$\det\begin{bmatrix}-cv_1-\\-v_2-\\-v_3-\\...\end{bmatrix}=c\times\det\begin{bmatrix}-v_1-\\-v_2-\\-v_3-\\...\end{bmatrix}\tag{4.4}$$

Example 4.10. $\det\begin{bmatrix}2&12&14\\3&6&9\\3&7&8\end{bmatrix}=2\times\det\begin{bmatrix}1&6&7\\3&6&9\\3&7&8\end{bmatrix}$

Example 4.11. $\det\begin{bmatrix}3&18&21&27\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}=3\times\det\begin{bmatrix}1&6&7&9\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}$

Example 4.12. $\det\begin{bmatrix}4&24&28&36&28\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}=4\times\det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

4.3 Property Derived from Three Basic Definition

From the Previous three basic definition of the operation called the determinant, we can derive more useful properties from it, we will give the prove of them based on the four equations in section two.

4.3.1 When Having Two Same Rows

$$\det\begin{bmatrix}-v_1-\\-v_1-\\-v_3-\\...\end{bmatrix}=0\tag{4.5}$$

Proof. We have the formula 4.2.2 ,

So we find that

$$\det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = -\det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix}$$

Which you can think that you change the first two rows.

However if we just look at the structure of both sides of the equation, we will find that

$$\det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix}$$

So what we have now is

$$\begin{cases} \det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = -\det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} \\ \det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} \end{cases}$$

If you add two equations up, then you will have

$$2 \times \det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = 0$$

Which is

$$\det \begin{bmatrix} -v_1- \\ -v_1- \\ -v_3- \\ \dots \end{bmatrix} = 0$$

□

Example 4.13. $\det \begin{bmatrix} 1 & 6 & 7 \\ 1 & 6 & 7 \\ 3 & 7 & 8 \end{bmatrix} = 0$

Example 4.14. $\det \begin{bmatrix} 1 & 6 & 7 & 9 \\ 1 & 6 & 7 & 9 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} = 0$

Example 4.15. $\det \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 1 & 6 & 7 & 9 & 7 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} = 0$

4.3.2 When Having One Row to be Zero

$$\det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = 0 \tag{4.6}$$

Actually, based on the three fundamental definition, we can prove this conclusion in several ways, here I will simply offer two of them.

Proof. 1 we have the formula 4.2.3

$$\begin{aligned} & \det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} \\ = & \det \begin{bmatrix} 0+0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} \\ = & \det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} \end{aligned}$$

So we proved that

$$\det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = 0$$

□

Proof. 2

We have the formula 4.2.3

$$\begin{aligned}
&\det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} \\
=&\det \begin{bmatrix} c \times 0 \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} \\
=&c \times \det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}
\end{aligned}$$

So we proved that

$$\det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} = 0$$

□

Example 4.16. $\det \begin{bmatrix} 0 & 0 & 0 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix} = 0$

Example 4.17. $\det \begin{bmatrix} 0 & 0 & 0 & 0 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} = 0$

Example 4.18. $\det \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} = 0$

4.3.3 When Two Rows Have Common Parts

$$\det \begin{bmatrix} -v_1 + c \times v_2- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} \tag{4.7}$$

Proof. We will make a short cut here by using the extra property which we have proved in the final step.

We have the formula 4.2.3 and the formula 4.2.3

$$\begin{aligned}
&\det\begin{bmatrix} -v_1 + c \times v_2- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} \\
= &\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} + \det\begin{bmatrix} -c \times v_2- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} \\
= &\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} + c \times \det\begin{bmatrix} -v_2- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix}
\end{aligned}$$

We also have the formula 4.3.1
So we know that

$$c \times \det\begin{bmatrix} -v_2- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} = 0$$

Then we proved that

$$\det\begin{bmatrix} -v_1 + c \times v_2- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix} = \det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix}$$

□

Example 4.19. $\det\begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix} = \det\begin{bmatrix} -2 & 0 & -2 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}$

Example 4.20. $\det\begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix} = \det\begin{bmatrix} -2 & 0 & -2 & 7 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}$

Example 4.21. $\det\begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} = -\det\begin{bmatrix} -2 & 0 & -2 & 7 & -2 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix}$

4.4 Row Operation and Determinant

4.4.1 Three Types of Elementary Matrix and Determinant

The most important thing in this subsection is to identify that all the elementary matrix has one characteristic in common that left-product of such matrix obeys the associative law.

Type I

The Type one in the row operation we learnt in the Gaussian elimination is known as the permutation matrix.

Conclusion 4.4.1.0.1. $\det \begin{bmatrix} p \end{bmatrix} = -1$

Proof. From the Alternating principle of the determinant which is the formula 4.2.2, we will have

$$\det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = -\det \begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix}$$

And according to the Normalization principle of the determinant which is the formula 4.2.1, we will have

$$\det(I) = 1$$

So

$$\begin{aligned} &\det \begin{bmatrix} p \end{bmatrix} \\ = &-\det \begin{bmatrix} I \end{bmatrix} \\ = &-1 \end{aligned}$$

□

Conclusion 4.4.1.0.2. $\det \left(\begin{bmatrix} p \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix} \right) = \det \begin{bmatrix} p \end{bmatrix} \times \det \begin{bmatrix} A \end{bmatrix}$

Proof. We have the formula 4.2.2, which is

$$\det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = -\det \begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix}$$

Let's think about the row operation in the Gaussian elimination, which is

$$\begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & ... \\ 1 & 0 & 0 & ... \\ 0 & 0 & 1 & ... \\ ... & ... & ... & ... \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} 0 & 1 & 0 & ... \\ 1 & 0 & 0 & ... \\ 0 & 0 & 1 & ... \\ ... & ... & ... & ... \end{bmatrix}$ to be the permutation matrix $\begin{bmatrix} p \end{bmatrix}$

Then we can take a look at the formula 4.2.2, the formula can be changed into

$$\det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = -\det(\begin{bmatrix} p \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix})$$

If we set the matrix $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ to be the matrix $\begin{bmatrix} A \end{bmatrix}$, Then the formula will continue to change to be

$$\det \begin{bmatrix} A \end{bmatrix} = -\det(\begin{bmatrix} p \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix})$$

If we take a look at the determinant of the permutation matrix $\begin{bmatrix} p \end{bmatrix}$.

Luckily we find that

$$\det \begin{bmatrix} A \end{bmatrix} = -(\det \begin{bmatrix} p \end{bmatrix} \times \det \begin{bmatrix} A \end{bmatrix})$$

We already have $\det \begin{bmatrix} A \end{bmatrix} = -\det(\begin{bmatrix} p \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix})$.

So we proved that for the permutation matrix $\begin{bmatrix} p \end{bmatrix}$ and the random square matrix $\begin{bmatrix} A \end{bmatrix}$, we will have $\det(\begin{bmatrix} p \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix}) = \det \begin{bmatrix} p \end{bmatrix} \times \det \begin{bmatrix} A \end{bmatrix}$ □

Example 4.22. $\det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = -1$ *and* $\det(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}) = \det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}$

Example 4.23. $\det \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = -1$ *and* $\det(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}) = \det \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}$

Example 4.24. $\det \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = -1$ *and* $\det \left(\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} \right) =$
$\det \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix}$

Type II

The Type II elementary matrix is a diagonal matrix with one diagonal entry not equals to zero and one when the rest of which are all one, which we will call it E_2 for convenient in this part.

Conclusion 4.4.1.0.3.

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

Proof. We have the Normalization principle of the determinant, which is formula 4.2.1, So we have $\det \begin{bmatrix} I \end{bmatrix} = 1$.

Also we have the second part of the Multi-Linearity principle of the determinant, which is the formula 4.2.3, so we have

$$\det \begin{bmatrix} -cv_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Then, we have

$$\begin{aligned} &\det \begin{bmatrix} E_2 \end{bmatrix} \\ =&c \times \det \begin{bmatrix} I \end{bmatrix} \\ =&c \times 1 \\ =&c \end{aligned}$$

□

Conclusion 4.4.1.0.4. $\det \left(\begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix} \right) = \det \begin{bmatrix} E_2 \end{bmatrix} \times \det \begin{bmatrix} A \end{bmatrix}$

Proof. We have the formula 4.2.3, which is

$$\det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -c \times v_i- \\ ... \end{bmatrix} = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix}$$

Let's think about the row operation in the Gaussian elimination, which is

$$\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -c \times v_i- \\ ... \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & c & ... \\ ... & ... & ... & ... & ... & ... \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & c & ... \\ ... & ... & ... & ... & ... & ... \end{bmatrix}$ to be the type II elementary matrix $\begin{bmatrix} E_2 \end{bmatrix}$

Then we can take a look at the formula 4.2.3, the formula can be changed into

$$\det \left(\begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix} \right) = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix}$ to be the matrix $\begin{bmatrix} A \end{bmatrix}$, Then the formula will continue to change to be

$$\det \left(\begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix} \right) = c \times \det \begin{bmatrix} A \end{bmatrix}$$

If we take a look at the determinant of the type II matrix $\begin{bmatrix} E_2 \end{bmatrix}$.

Luckily we find that

$$\det \begin{bmatrix} E_2 \end{bmatrix} \times \det \begin{bmatrix} A \end{bmatrix} = c \times \det \begin{bmatrix} A \end{bmatrix}$$

We already have $\det\left(\left[E_2\right] \times \left[A\right]\right) = c \times \det\left[A\right]$.

So we proved that for the type II elementary matrix $\left[E_2\right]$ and the random square matrix $\left[A\right]$, we will have $\det\left(\left[E_2\right] \times \left[A\right]\right) = \det\left[E_2\right] \times \det\left[A\right]$ □

Example 4.25. $\det\begin{bmatrix}2 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} = 2$ *and* $\det\left(\begin{bmatrix}2 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \times \begin{bmatrix}1 & 6 & 7\\3 & 6 & 9\\3 & 7 & 8\end{bmatrix}\right) = \det\begin{bmatrix}2 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \times \det\begin{bmatrix}1 & 6 & 7\\3 & 6 & 9\\3 & 7 & 8\end{bmatrix}$

Example 4.26. $\det\begin{bmatrix}1 & 0 & 0 & 0\\0 & -3 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix} = -3$ *and* $\det\left(\begin{bmatrix}1 & 0 & 0 & 0\\0 & -3 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix} \times \begin{bmatrix}1 & 6 & 7 & 9\\3 & 6 & 9 & 2\\3 & 7 & 8 & 1\\2 & 6 & 8 & 9\end{bmatrix}\right) = \det\begin{bmatrix}1 & 0 & 0 & 0\\0 & -3 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix} \times \det\begin{bmatrix}1 & 6 & 7 & 9\\3 & 6 & 9 & 2\\3 & 7 & 8 & 1\\2 & 6 & 8 & 9\end{bmatrix}$

Example 4.27. $\det\begin{bmatrix}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 5 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{bmatrix} = -3$ *and* $\det\left(\begin{bmatrix}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 5 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{bmatrix} \times \begin{bmatrix}1 & 6 & 7 & 9 & 7\\3 & 6 & 9 & 2 & 9\\3 & 7 & 8 & 1 & 6\\2 & 6 & 8 & 9 & 5\\1 & 4 & 8 & 7 & 5\end{bmatrix}\right) = \det\begin{bmatrix}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 5 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{bmatrix} \times \det\begin{bmatrix}1 & 6 & 7 & 9 & 7\\3 & 6 & 9 & 2 & 9\\3 & 7 & 8 & 1 & 6\\2 & 6 & 8 & 9 & 5\\1 & 4 & 8 & 7 & 5\end{bmatrix}$

Type III

The Type III elementary matrix is very similar to the identity matrix when one off diagonal entry is not zero, which we will call it E_3 for convenience in this part.

Conclusion 4.4.1.0.5. $\det\begin{bmatrix}1 & 0 & 0 & ... & 0 & ... & 0 & ...\\0 & 1 & 0 & ... & 0 & ... & 0 & ...\\0 & 0 & 1 & ... & 0 & ... & 0 & ...\\... & ... & ... & ... & ... & ... & ... & ...\\0 & 0 & 0 & ... & c_{ij} & ... & 1 & ...\\... & ... & ... & ... & ... & ... & ... & ...\end{bmatrix} = 1$

Proof. We have the Normalization principle of the determinant, which is formula 4.2.1, So we have $\det\left[I\right] = 1$.

Also we have the second part of the Multi-Linearity principle of the determinant, which is the formula 4.2.3, so we have

$$\det \begin{bmatrix} -cv_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

What's more we also have the first part of Multi-Linearity principle of the determinant, which is the formula 4.2.3, so we have

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Then, we have

$$\begin{aligned} &\det \left[E_3\right] \\ =&\det \left[I\right] \\ =&1 \end{aligned}$$

□

Conclusion 4.4.1.0.6. $\det(\left[E_3\right] \times \left[A\right]) = \det \left[E_3\right] \times \det \left[A\right]$

Proof. We have the formula 4.2.3, which is

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Let's think about the row operation in the Gaussian elimination, which is

$$\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_j- \\ ... \\ -v_i + v_j- \\ ... \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & c_{ij} & ... & 1 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_j- \\ ... \\ -v_i- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & c_{ij} & ... & 1 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix}$ to be the type III elementary matrix $\left[E_3\right]$

Then we can take a look at the formula 4.2.3, the formula can be changed into

$$\det\left(\begin{bmatrix}E_3\end{bmatrix} \times \begin{bmatrix}-v_1-\\-v_2-\\-v_3-\\\cdots\\-v_j-\\\cdots\\-v_i-\\\cdots\end{bmatrix}\right) = \det \begin{bmatrix}-v_1-\\-v_2-\\-v_3-\\\cdots\\-v_j-\\\cdots\\-v_i-\\\cdots\end{bmatrix}$$

If we set the matrix $\begin{bmatrix}-v_1-\\-v_2-\\-v_3-\\\cdots\\-v_j-\\\cdots\\-v_i-\\\cdots\end{bmatrix}$ to be the matrix $\begin{bmatrix}A\end{bmatrix}$, Then the formula will continue to change to be

$$\det\left(\begin{bmatrix}E_3\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}A\end{bmatrix}$$

If we take a look at the determinant of the type III matrix $\begin{bmatrix}E_3\end{bmatrix}$.

Luckily we find that

$$\det\begin{bmatrix}E_3\end{bmatrix} \times \det\begin{bmatrix}A\end{bmatrix} = \det\begin{bmatrix}A\end{bmatrix}$$

We already have $\det\left(\begin{bmatrix}E_3\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}A\end{bmatrix}$.

So we proved that for the type III elementary matrix $\begin{bmatrix}E_3\end{bmatrix}$ and the random square matrix $\begin{bmatrix}A\end{bmatrix}$, we will have $\det\left(\begin{bmatrix}E_3\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}E_3\end{bmatrix} \times \det\begin{bmatrix}A\end{bmatrix}$ ☐

Example 4.28. $\det\begin{bmatrix}1&0&0\\2&1&0\\0&0&1\end{bmatrix} = 1$ *and* $\det\left(\begin{bmatrix}1&0&0\\2&1&0\\0&0&1\end{bmatrix} \times \begin{bmatrix}1&6&7\\3&6&9\\3&7&8\end{bmatrix}\right) = \det\begin{bmatrix}1&0&0\\2&1&0\\0&0&1\end{bmatrix} \times \det\begin{bmatrix}1&6&7\\3&6&9\\3&7&8\end{bmatrix}$

Example 4.29. $\det\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&-1&1&0\\0&0&0&1\end{bmatrix} = 1$ *and* $\det\left(\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&-1&1&0\\0&0&0&1\end{bmatrix} \times \begin{bmatrix}1&6&7&9\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}\right) = \det\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&-1&1&0\\0&0&0&1\end{bmatrix} \times \det\begin{bmatrix}1&6&7&9\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}$

Example 4.30. $\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&9&1&0\\0&0&0&0&1\end{bmatrix} = 1 \textit{ and} \det\left(\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&9&1&0\\0&0&0&0&1\end{bmatrix} \times \begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}\right) =$
$\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&9&1&0\\0&0&0&0&1\end{bmatrix} \times \det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

4.4.2 Elementary Matrix

This subsection is mainly a conclusion of the previous subsection, because we want to see what all the elementary matrix has in common.

For all three types of elementary matrix, we all have

$$\det\left(\begin{bmatrix}E\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}E\end{bmatrix} \times \det\begin{bmatrix}A\end{bmatrix}$$

When $\begin{bmatrix}E\end{bmatrix}$ is elementary matrix and $\begin{bmatrix}A\end{bmatrix}$ is the random square matrix That means all the elementary matrix has one characteristic in common that left-product of such matrix obeys the associative law.

Example 4.31. $\det\begin{bmatrix}0&0&1&0&0\\0&1&0&0&0\\1&0&0&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix} = -1 \textit{ and} \det\left(\begin{bmatrix}0&0&1&0&0\\0&1&0&0&0\\1&0&0&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix} \times \begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}\right) =$
$\det\begin{bmatrix}0&0&1&0&0\\0&1&0&0&0\\1&0&0&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix} \times \det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

Example 4.32. $\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&-2&0\\0&0&0&0&1\end{bmatrix} = -3 \textit{ and} \det\left(\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&-2&0\\0&0&0&0&1\end{bmatrix} \times \begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}\right) =$
$\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&-2&0\\0&0&0&0&1\end{bmatrix} \times \det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

Example 4.33. $\det \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -4 & 0 & 0 & 1 \end{bmatrix} = 1$ *and* $\det \left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -4 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix} \right) =$
$\det \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -4 & 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 & 9 & 7 \\ 3 & 6 & 9 & 2 & 9 \\ 3 & 7 & 8 & 1 & 6 \\ 2 & 6 & 8 & 9 & 5 \\ 1 & 4 & 8 & 7 & 5 \end{bmatrix}$

4.4.3 Back Substitution Matrix

When we look back at the Gaussian elimination, we can actually split it into two sub steps listed as below.

1. Left multiplied by elementary matrix to turn it into reduced row echelon form.

2. Left multiplied by upper triangular matrix, to turn the reduced row echelon form into diagonal matrix, through the step called the back substitution matrix.

From the previous discovery, we now know the determinant of all the elementary matrix, but we didn't spend our time on the back substitution matrix.

In this subsection we will find the secret of the back substitution matrix, and you will find it is very similar to type III elementary matrix.

Definition 4.1. *The back substitution matrix is matrix in form of* $\begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & ... & c_{ij} & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix}$ *and I will take the sign of S wo represent the back substitution matrix.*

Conclusion 4.4.3.0.1. $\det \begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & ... & c_{ij} & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix} = 1$

Proof. We have the Normalization principle of the determinant, which is formula 4.2.1, So we have $\det \begin{bmatrix} I \end{bmatrix} = 1$.

Also we have the second part of the Multi-Linearity principle of the determinant, which is the formula 4.2.3, so we have

$$\det \begin{bmatrix} -cv_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

What's more we also have the first part of Multi-Linearity principle of the determinant, which is the formula 4.2.3, so we have

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Then, we have

$$\begin{aligned} &\det \left[S\right] \\ =&\det \left[I\right] \\ =&1 \end{aligned}$$

□

Conclusion 4.4.3.0.2. $\det\left(\left[S\right] \times \left[A\right]\right) = \det\left[S\right] \times \det\left[A\right]$

Proof. We have the formula 4.2.3, which is

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Let's think about the row operation in the Gaussian elimination, which is

$$\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i + v_j- \\ ... \\ -v_j- \\ ... \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & ... & c_{ij} & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \\ -v_j- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} 1 & 0 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 1 & 0 & ... & 0 & ... & 0 & ... \\ 0 & 0 & 1 & ... & 0 & ... & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & ... & c_{ij} & ... \\ ... & ... & ... & ... & ... & ... & ... & ... \end{bmatrix}$ to be the back substitution matrix $\left[S\right]$

Then we can take a look at the formula 4.2.3, the formula can be changed into

$$\det\left(\begin{bmatrix}S\end{bmatrix} \times \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \\ -v_i- \\ \cdots \\ -v_j- \\ \cdots \end{bmatrix}\right) = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \\ -v_i- \\ \cdots \\ -v_j- \\ \cdots \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \\ -v_i- \\ \cdots \\ -v_j- \\ \cdots \end{bmatrix}$ to be the matrix $\begin{bmatrix}A\end{bmatrix}$, Then the formula will continue to change to be

$$\det\left(\begin{bmatrix}S\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det \begin{bmatrix}A\end{bmatrix}$$

If we take a look at the determinant of the back substitution matrix $\begin{bmatrix}S\end{bmatrix}$.

Luckily we find that

$$\det \begin{bmatrix}S\end{bmatrix} \times \det \begin{bmatrix}A\end{bmatrix} = \det \begin{bmatrix}A\end{bmatrix}$$

We already have $\det\left(\begin{bmatrix}S\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det \begin{bmatrix}A\end{bmatrix}$.

So we proved that for the back substitution matrix $\begin{bmatrix}S\end{bmatrix}$ and the random square matrix $\begin{bmatrix}A\end{bmatrix}$, we will have $\det\left(\begin{bmatrix}S\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det \begin{bmatrix}S\end{bmatrix} \times \det \begin{bmatrix}A\end{bmatrix}$ $\square$

Example 4.34. $\det \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1$ *and* $\det\left(\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}\right) = \det \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 \\ 3 & 6 & 9 \\ 3 & 7 & 8 \end{bmatrix}$

Example 4.35. $\det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 1$ *and* $\det\left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}\right) = \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \det \begin{bmatrix} 1 & 6 & 7 & 9 \\ 3 & 6 & 9 & 2 \\ 3 & 7 & 8 & 1 \\ 2 & 6 & 8 & 9 \end{bmatrix}$

Example 4.36. $\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&9&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix} = 1$ *and* $\det\left(\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&9&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}\times\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}\right) =$
$\det\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&9&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}\times\det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

4.4.4 All the Matrix Representing Row Operation

We know, all the greater row operation can be split into row operations that simply happening between two rows or the simple change inside one row itself. And the elementary matrix with the back substitution matrix have already contained every situation of the slight row operation.
The bigger row operation can be seen as constructed by the simpler ones. Algebraically speaking, we can represent the more complex row operation by the product of the elementary matrix with the back substitution matrix.
Let's see whether the left-product of such matrix obeys the associative law.
The answer is Yes.

Conclusion 4.4.4.0.1. *If the matrix $\begin{bmatrix}B\end{bmatrix}$ is the product of elementary matrix and the back substitution matrix, and $\begin{bmatrix}A\end{bmatrix}$ is a random square matrix.*

$$\det\left(\begin{bmatrix}B\end{bmatrix}\times\begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}B\end{bmatrix}\times\det\begin{bmatrix}A\end{bmatrix}$$

Proof. We know the matrix $\begin{bmatrix}B\end{bmatrix}$ is the product of elementary matrices and back substitution matrix.

If we generate the elementary matrix to one particular kind of matrix called the fundamental matrix, with sign $\begin{bmatrix}F\end{bmatrix}$

Then we can split the matrix $\begin{bmatrix}B\end{bmatrix}$, which is

$$\begin{bmatrix}B\end{bmatrix} = \begin{bmatrix}F\end{bmatrix}_k\times\begin{bmatrix}F\end{bmatrix}_{k-1}\times ...\times\begin{bmatrix}F\end{bmatrix}_2\times\begin{bmatrix}F\end{bmatrix}_1$$

Then we need to use the property of associative law for the fundamental matrices multiplied at the left side.

Then we can find that

$$\begin{aligned}&\det\left(\begin{bmatrix}B\end{bmatrix}\right)\times\begin{bmatrix}A\end{bmatrix})\\=&\det\left(\begin{bmatrix}F\end{bmatrix}_k\times\begin{bmatrix}F\end{bmatrix}_{k-1}\times ...\times\begin{bmatrix}F\end{bmatrix}_2\times\begin{bmatrix}F\end{bmatrix}_1\times\begin{bmatrix}A\end{bmatrix}\right)\\=&\det\begin{bmatrix}F\end{bmatrix}_k\times\det\left(\begin{bmatrix}F\end{bmatrix}_{k-1}\times ...\times\begin{bmatrix}F\end{bmatrix}_2\times\begin{bmatrix}F\end{bmatrix}_1\times\begin{bmatrix}A\end{bmatrix}\right)\\=&\det\left(\begin{bmatrix}F\end{bmatrix}_k\times\begin{bmatrix}F\end{bmatrix}_{k-1}\times ...\times\begin{bmatrix}F\end{bmatrix}_2\times\begin{bmatrix}F\end{bmatrix}_1\right)\times\det\begin{bmatrix}A\end{bmatrix}\\=&\det\left(\begin{bmatrix}B\end{bmatrix}\right)\times\det\left(\begin{bmatrix}A\end{bmatrix}\right)\end{aligned}$$

□

What we now know is for any matrix that is the product of the elementary matrices and back substitution matrix obeys the associative law when was multiplied left.

Example 4.37. $\det\left(\begin{bmatrix}5&1&7\\1&2&4\\3&2&1\end{bmatrix}\times\begin{bmatrix}1&6&7\\3&6&9\\3&7&8\end{bmatrix}\right)=\det\begin{bmatrix}5&1&7\\1&2&4\\3&2&1\end{bmatrix}\times\det\begin{bmatrix}1&6&7\\3&6&9\\3&7&8\end{bmatrix}$

Example 4.38. $\det\left(\begin{bmatrix}5&1&7&8\\1&2&4&7\\3&2&1&6\\1&5&4&3\end{bmatrix}\times\begin{bmatrix}1&6&7&9\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}\right)=\det\begin{bmatrix}5&1&7&8\\1&2&4&7\\3&2&1&6\\1&5&4&3\end{bmatrix}\times$
$\det\begin{bmatrix}1&6&7&9\\3&6&9&2\\3&7&8&1\\2&6&8&9\end{bmatrix}$

Example 4.39. $\det\left(\begin{bmatrix}5&1&7&8&6\\1&2&4&7&7\\3&2&1&6&5\\1&5&4&3&4\\1&1&5&6&4\end{bmatrix}\times\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}\right)=\det\begin{bmatrix}5&1&7&8&6\\1&2&4&7&7\\3&2&1&6&5\\1&5&4&3&4\\1&1&5&6&4\end{bmatrix}\times$
$\det\begin{bmatrix}1&6&7&9&7\\3&6&9&2&9\\3&7&8&1&6\\2&6&8&9&5\\1&4&8&7&5\end{bmatrix}$

4.5 The Associative Law and Determinant I

4.5.1 When Multiplied at the Left Side

Fundamental Matrix

Definition 4.2. *The Fundamental matrix is a conception that created by myself for convenience in this part, which is the product of the elementary matrix and the back substitution matrix.*

This part is the brief review of the previous knowledge.
We now know that for any fundamental matrix $\left[B\right]$ multiplied by the left of a random square matrix $\left[A\right]$,

$$\det\left(\left[B\right]\times\left[A\right]\right)=\det\left[B\right]\times\det\left[A\right]$$

We will find the left side of the product follow the associative law.

The conclusion here is that the determinant of the fundamental matrix follows the associative law when multiplied by the left of a random square matrix.

Square Matrix

Conclusion 4.5.1.0.1. *For any square matrix* $\left[S\right]$ *and random square matrix* $\left[A\right]$ *we all have*

$$\det\left(\left[S\right] \times \left[A\right]\right) = \det\left[S\right] \times \det\left[A\right]$$

Proof. **Hint: LDU decomposition**

The basic calculation of the determinant have close relationship with the Gaussian elimination, so we can always take our eyes on it. But we have to set two cases here by whether the matrix $\left[S\right]$ is invertible.

Condition 4.5.1.0.1 (When the matrix $\left[S\right]$ is invertible)**.** *From the previous study in the chapter one, we will find that there are totally Two steps of Gaussian elimination to complete the LDU decomposition, when the invertible matrix must can be decomposited.*

*1. Using the elementary matrix to do the row operation until we got the **Upper Triangular Matrix**.*

*2. Using the diagonal matrix to change the upper triangular matrix to the **Reduced Row Echelon Form**.*

*3.using the back substitution matrix to do the back substitution step until we get the **Diagonal Matrix***

So when we change the matrix $\left[S\right]$ *to the upper triangular matrix* $\left[\hat{U}\right]$

We have to multiply several elementary matrix by the left side of the matrix $\left[S\right]$

That is

$$\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right]\left[E_1\right] \times \left[S\right] = \left[\hat{U}\right]$$

We know that the elementary matrices are all invertible.

So

$$\left[S\right] = \left(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right]\left[E_1\right]\right)^{-1} \times \left[\hat{U}\right]$$

When we change the upper triangular matrix $\left[\hat{U}\right]$ *to the reduced row echelon form* $\left[R\right]$*.*

We will have to left multiply the upper triangular matrix

That is

$$\left[D_1\right] \times \left[\hat{U}\right] = \left[R\right]$$

We know that the diagonal matrix must be invertible, so we have

$$\left[\hat{U}\right] = \left(\left[D_1\right]\right)^{-1} \times \left[R\right]$$

Now we come to the final step, left multiply the back substitution matrices $\left[BS\right]$ *on the reduced row echelon form, changing it to an identity matrix.*

Caution 1. *In the former sections, I used the symbol $\left[S\right]$ to represent the back substitution matrix, but considering we have used the same symbol here to represent the square matrix here, so we change the symbol of back substitution matrix to be $\left[BS\right]$ in this subsection.*

That is

$$\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right] \times \left[R\right] = \left[I\right]$$

And similarly we know the back substitution matrices are all invertible, so we have

$$\left[R\right] = (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right])^{-1} \times \left[I\right]$$

To have a short conclusion here, what now we have are

$$\begin{cases} \left[S\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right])^{-1} \times \left[\hat{U}\right] \\ \left[\hat{U}\right] = (\left[D_1\right])^{-1} \times \left[R\right] \\ \left[R\right] = (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right])^{-1} \times \left[I\right] \end{cases}$$

So we have

$$\left[S\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right])^{-1} \times (\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right])^{-1}$$

So we found that ***when the square matrix $\left[S\right]$ is invertible, it can must be decomposited to be the product of several fundemantal matrix, we know that the determinant of the fundamental matrix follows the associative law when multiplied by the left of a random square matrix, so here we proved that when the square matrix $\left[S\right]$ is invertible, the determinant of the matrix follows the associative law when multiplied by the left of a random square matrix.*** *In the algebraic way, it is just For any square matrix $\left[S\right]$ (but it should be invertible) and random square matrix $\left[A\right]$ we all have*

$$\det(\left[S\right] \times \left[A\right]) = \det\left[S\right] \times \det\left[A\right]$$

Condition 4.5.1.0.2 (When the matrix $\left[S\right]$ is not invertible)**.** *This time we have to use the knowledge on the not invertible matrices. Now let's take our look back on the equation we want to prove:*

$$\det(\left[S\right] \times \left[A\right]) = \det\left[S\right] \times \det\left[A\right]$$

When the matrix $\left[S\right]$ is not invertible, according to the former knowledge which I will not going to prove here, we have the matrix $\left[S\right] \times \left[A\right]$ is also not invertible.
Here we try to find what will the determinant be if the matrix itself is not invertible.
We set a not invertible matrix $\left[V\right]$, now we will do the Gaussian elimination on it.

We know that when the matrix is not invertible, it can not be LDU decomposited. But we will use the same steps as the first condition did to decomposite the matrix $\left[V\right]$.
The first step is to change the matrix $\left[V\right]$ *to the upper triangular matrix* $\left[\hat{U}\right]$
We have to multiply several elementary matrix by the left side of the matrix $\left[V\right]$
That is

$$\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right] \times \left[V\right] = \left[\hat{U}\right]$$

We know that the elementary matrices are all invertible.
So

$$\left[V\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right])^{-1} \times \left[\hat{U}\right]$$

But here, because The matrix $\left[\hat{U}\right]$ *is not invertible, so we can't use a diagonal matrix to change it into the reduced row echelon form with full pivots.*
So I consider making some change here.
What we already have is

$$\left[V\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right])^{-1} \times \left[\hat{U}\right]$$

we know that when the elementary matrices are multiplied by the left side of the square matrix, the determinant follows the associative law, in the conclusion 4.4.4,
So we have

$$\begin{aligned} &\det \left[V\right] \\ =&\det ((\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right])^{-1} \times \left[\hat{U}\right]) \\ =&\det (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \left[E_1\right]^{-1}) \times \det \left[\hat{U}\right] \end{aligned}$$

Now look at the $\det \left[\hat{U}\right]$
We know there is one way to describe a square matrix when it is not invertible, which is saying that the rows of the matrix is not linear independent.
Suppose for the matrix $\left[\hat{U}\right]$, *the ith row is the construction of the rest of the rows.*
We have the Multi-Linearity property for the determinant as the formula 4.2.3, which is

$$\det \begin{bmatrix} -v_1 + v_1^{'} - \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

We also have the corollary 4.3.2, which is

$$\det \begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = 0$$

After this method for several times. we will find

$$\begin{aligned}
&\det\left[\hat{U}\right]\\
=&\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix}\\
=&\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ 0 \\ ... \end{bmatrix}\\
=&0
\end{aligned}$$

Then we will got $\det \hat{U} = 0$.
So

$$\begin{aligned}
&\det\left[V\right]\\
=&\det\left(\left(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right]\left[E_1\right]\right)^{-1} \times \left[\hat{U}\right]\right)\\
=&\det\left(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right]\left[E_1\right]^{-1}\right) \times \det\left[\hat{U}\right]\\
=&0
\end{aligned}$$

Then we will got ***the determinant for any invertible square matrix is zero****.*
Now we again look at the thing that we are going to prove, which is

$$\det\left(\left[S\right] \times \left[A\right]\right) = \det\left[S\right] \times \det\left[A\right]$$

When the matrix $\left[S\right]$ *is not invertible, we will find the matrix* $\left[S\right] \times \left[A\right]$ *is not invertible, too.*
So now we will have

$$\det\left(\left[S\right] \times \left[A\right]\right) = 0 = \det\left[S\right] \times \det\left[A\right]$$

□

The most important thing we proved in this chapter is that the determinant of any square matrix follow the associative law when it is multiplied on the left of other square matrices

4.5.2 When Multiplied at the right side

This part will be discussed in the next section.

4.6 Column Operation and Determinant

In the previous sections the case we are analyzing are always multiply the target matrix on the left when finding for any square matrix the determinant follow the associative law.
We start that section from the idea of Gaussian elimination, which analyzing the matrix by multiple times of row operations.
We understand every idea from perspective of row operation, which limit the "left multiplication" because of the product rule for the matrix.
We have successfully understand how everything goes in the situation of $\det\left(\begin{bmatrix}S\end{bmatrix} \times \begin{bmatrix}A\end{bmatrix}\right) = \det\begin{bmatrix}S\end{bmatrix} \times \det\begin{bmatrix}A\end{bmatrix}$, but something is still missing.
How about finding the relationship between
$\det\left(\begin{bmatrix}A\end{bmatrix} \times \begin{bmatrix}S\end{bmatrix}\right) and \det\begin{bmatrix}A\end{bmatrix} \times \det\begin{bmatrix}S\end{bmatrix}$
Whether the associative law will still be there when the target matrix is multiplied on the right side?
Taking a matrix up from the right is just doing the column change to the original matrix, this characteristic comes from the basic definition of matrix multiplication.
So in this section, **algebraically speaking**, we will starting to analyze the situation like $\det\left(\begin{bmatrix}A\end{bmatrix} \times \begin{bmatrix}S\end{bmatrix}\right) = \det\begin{bmatrix}A\end{bmatrix} \times \det\begin{bmatrix}S\end{bmatrix}$, when the matrix $\begin{bmatrix}S\end{bmatrix}$ is chosen from elementary matrix, to fundamental matrix, then to any square matrix. **Abstractly speaking**, we will start consider the column change instead of row change.

4.6.1 determinant of the Transposed and the Original

we do can try to estimate the steps we did on the row operation, **but please remember the original three definition on the determinant are fixed, which has great characteristic of row operation rather than the column operation.**

So here we need a really powerful equation to get some information about the column operation.

But how to find one?

You must have an idea of doing something on the basic three definition on the determinant, so we'd better take a good look on it.

Let's take the Alternating rule as an example

According to the formula 4.2.2, which is

$$\det\begin{bmatrix}-v_1- \\ -v_2- \\ -v_3- \\ \dots\end{bmatrix} = -\det\begin{bmatrix}-v_2- \\ -v_1- \\ -v_3- \\ \dots\end{bmatrix}$$

,

If we set the matrix $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ to be the matrix $\begin{bmatrix} A \end{bmatrix}$, when the matrix $\begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix}$ can be gained from the matrix $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ after the row exchange. If we set the symbol of the permutation matrix here to be $\begin{bmatrix} P \end{bmatrix}$,

Then the matrix $\begin{bmatrix} -v_2- \\ -v_1- \\ -v_3- \\ ... \end{bmatrix}$ can be treat as $\begin{bmatrix} P \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix}$.

Back substitute them into the formula 4.2.2, we will have

$$\det \begin{bmatrix} A \end{bmatrix} = -\det \begin{bmatrix} P \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix}$$

This formula shows the direct relationship between the determinant of one particular matrix and the determinant of which after row exchange

Now the problem to us is as expected, no clue for column exchange.

Hint: Algebraically speaking, column exchange asks the original matrix to show up at the left side of the multiplication, so try the transpose

So my conclusion is the determinant of the transpose is very important.

Conclusion 4.6.1.0.1. *The matrix $\begin{bmatrix} S \end{bmatrix}$ is any square matrix.*

$$\det \begin{bmatrix} S \end{bmatrix} = \det \left(\begin{bmatrix} S \end{bmatrix}\right)^T$$

Proof. **Hint: The step here is very similar to the one in the proof of conclusion 4.5.1**

Condition 4.6.1.0.1 (When the matrix $\begin{bmatrix} S \end{bmatrix}$ is invertible)**.** *We know in the previous section when we are proving the square matrices follow the associative law when taken up from the left, we split the complete version of Gaussian elimination into three sub-steps.*

1. Using the elementary matrix to do the row operation until we got the ***Upper Triangular Matrix****.*

2. Using the diagonal matrix to change the upper triangular matrix to the ***Reduced Row Echelon Form****.*

3.using the back substitution matrix to do the back substitution step until we get the ***Diagonal Matrix***

So when we change the matrix $\begin{bmatrix} S \end{bmatrix}$ to the upper triangular matrix $\begin{bmatrix} \hat{U} \end{bmatrix}$

We have to multiply several elementary matrix by the left side of the matrix $\begin{bmatrix} S \end{bmatrix}$

That is

$$\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix} \times \begin{bmatrix} S \end{bmatrix} = \begin{bmatrix} \hat{U} \end{bmatrix}$$

We know that the elementary matrices are all invertible.
So

$$\left[S\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right])^{-1} \times \left[\hat{U}\right]$$

When we change the upper triangular matrix $\left[\hat{U}\right]$ *to the reduced row echelon form* $\left[R\right]$.
We will have to left multiply the upper triangular matrix
That is

$$\left[D_1\right] \times \left[\hat{U}\right] = \left[R\right]$$

We know that the diagonal matrix must be invertible, so we have

$$\left[\hat{U}\right] = (\left[D_1\right])^{-1} \times \left[R\right]$$

Now we come to the final step, left multiply the back substitution matrices $\left[BS\right]$ *on the reduced row echelon form, changing it to an identity matrix.*

Caution 2. *In the former sections, I used the symbol* $\left[S\right]$ *to represent the back substitution matrix, but considering we have used the same symbol here to represent the square matrix here, so we change the symbol of back substitution matrix to be* $\left[BS\right]$ *in this subsection.*

That is

$$\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \times \left[BS_1\right] \times \left[R\right] = \left[I\right]$$

And similarly we know the back substitution matrices are all invertible, so we have

$$\left[R\right] = (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \times \left[BS_1\right])^{-1} \times \left[I\right]$$

To have a short conclusion here, what now we have are

$$\begin{cases} \left[S\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right])^{-1} \times \left[\hat{U}\right] \\ \left[\hat{U}\right] = (\left[D_1\right])^{-1} \times \left[R\right] \\ \left[R\right] = (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \times \left[BS_1\right])^{-1} \times \left[I\right] \end{cases}$$

So we have

$$\left[S\right] = (\left[E_k\right]\times\left[E_{k-1}\right]\times...\times\left[E_2\right]\times\left[E_1\right])^{-1}\times(\left[D_1\right])^{-1}\times(\left[BS_k\right]\times\left[BS_{k-1}\right]\times...\times\left[BS_2\right]\times\left[BS_1\right])^{-1}$$

We don't have any useful information about what will happen after we take the transpose of the matrix, but now we can find the result is really clear.
We know that $\left[S\right] = (\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right])^{-1} \times (\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \times \left[BS_1\right])^{-1}$,
After we take the transpose on the both sides of the equation.

It will be like

$$
\begin{aligned}
&\left[S^T\right]\\
=&(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right])^{-1} \times (\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \times \left[BS_1\right])^{-1}\\
=&[(\left[E_1\right])^{-1} \times (\left[E_2\right])^{-1} \times ... \times (\left[E_{k-1}\right])^{-1} \times (\left[E_k\right])^{-1} \times (\left[D_1\right])^{-1}\\
&\times (\left[BS_1\right])^{-1} \times (\left[BS_2\right])^{-1} \times ... \times (\left[BS_{k-1}\right])^{-1} \times (\left[BS_k\right])^{-1}]^T\\
=&(\left[BS_k^T\right])^{-1} \times (\left[BS_{k-1}^T\right])^{-1} \times ... \times (\left[BS_2^T\right])^{-1} \times (\left[BS_1^T\right])^{-1} \times (\left[D_1^T\right])^{-1}\\
&\times (\left[E_k^T\right])^{-1} \times (\left[E_{k-1}^T\right])^{-1} \times ... \times (\left[E_2^T\right])^{-1} \times (\left[E_1^T\right])^{-1}
\end{aligned}
$$

We also have the conclusion that the fundamental matrix follows the associative law in the determinant when taken up from the left.
Also we know that the determinant of the type III elementary matrix and the back substitution matrix are one from the conclusion 4.4.1 and the conclusion 4.4.3
So now we have

$$
\begin{aligned}
&\det\left[S^T\right]\\
=&\det((\left[BS_k^T\right])^{-1} \times (\left[BS_{k-1}^T\right])^{-1} \times ... \times (\left[BS_2^T\right])^{-1} \times (\left[BS_1^T\right])^{-1} \times (\left[D_1^T\right])^{-1}\\
&\times (\left[E_k^T\right])^{-1} \times (\left[E_{k-1}^T\right])^{-1} \times ... \times (\left[E_2^T\right])^{-1} \times (\left[E_1^T\right])^{-1})\\
=&\det(\left[BS_k^T\right]^{-1}) \times \det(\left[BS_{k-1}^T\right]^{-1}) \times ... \times \det(\left[BS_2^T\right]^{-1}) \times \det(\left[BS_1^T\right]^{-1}) \times \det(\left[D_1^T\right]^{-1})\\
&\times \det(\left[E_k^T\right]^{-1}) \times \det(\left[E_{k-1}^T\right]^{-1}) \times ... \times \det(\left[E_2^T\right]^{-1}) \times \det(\left[E_1^T\right]^{-1})\\
=&\det(\left[D_1^T\right]^{-1})
\end{aligned}
$$

Similarly, we also have

$$
\begin{aligned}
&\det\left[S\right]\\
=&\det((\left[E_1\right])^{-1} \times (\left[E_2\right])^{-1} \times ... \times (\left[E_{k-1}\right])^{-1} \times (\left[E_k\right])^{-1} \times (\left[D_1\right])^{-1}\\
&\times (\left[BS_1\right])^{-1} \times (\left[BS_2\right])^{-1} \times ... \times (\left[BS_{k-1}\right])^{-1} \times (\left[BS_k\right])^{-1})\\
=&\det(\left[E_1\right]^{-1}) \times \det(\left[E_2\right]^{-1}) \times ... \times \det(\left[E_{k-1}\right]^{-1}) \times \det(\left[E_k\right]^{-1}) \times \det(\left[D_1\right]^{-1})\\
&\times \det(\left[BS_1\right]^{-1}) \times \det(\left[BS_2\right]^{-1}) \times ... \times \det(\left[BS_{k-1}\right]^{-1}) \times \det(\left[BS_k\right]^{-1})\\
=&\det(\left[D_1\right]^{-1})
\end{aligned}
$$

We know the matrix $\left[D_1\right]$ *is a diagonal matrix, so we have* $\left[D_1\right] = \left[D_1^T\right]$ *So now we have* $\det\left[S\right] = \det(\left[D_1\right]^{-1}) = \det(\left[D_1^T\right]^{-1}) = \det\left[S^T\right]$
So we proved that, when the matrix $\left[S\right]$ ***is invertible we have*** $\det\left[S\right] = \det\left[S^T\right]$

Condition 4.6.1.0.2 (When the matrix $\left[S\right]$ is not invertible)**.** *This time we have to use the knowledge on the not invertible matrices. Now let's take our look back on the equation we want to prove:*

$$\det(\left[S\right] \times \left[A\right]) = \det\left[S\right] \times \det\left[A\right]$$

When the matrix $\begin{bmatrix} S \end{bmatrix}$ is not invertible, according to the former knowledge which I will not going to prove here, we have the matrix $\begin{bmatrix} S \end{bmatrix} \times \begin{bmatrix} A \end{bmatrix}$ is also not invertible.

Here we try to find what will the determinant be if the matrix itself is not invertible.

We set a not invertible matrix $\begin{bmatrix} V \end{bmatrix}$, now we will do the Gaussian elimination on it.

We know that when the matrix is not invertible, it can not be LDU decomposited. But we will use the same steps as the first condition did to decomposite the matrix $\begin{bmatrix} V \end{bmatrix}$.

The first step is to change the matrix $\begin{bmatrix} V \end{bmatrix}$ to the upper triangular matrix $\begin{bmatrix} \hat{U} \end{bmatrix}$ We have to multiply several elementary matrix by the left side of the matrix $\begin{bmatrix} V \end{bmatrix}$

That is

$$\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix} \times \begin{bmatrix} V \end{bmatrix} = \begin{bmatrix} \hat{U} \end{bmatrix}$$

We know that the elementary matrices are all invertible.

So

$$\begin{bmatrix} V \end{bmatrix} = (\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix})^{-1} \times \begin{bmatrix} \hat{U} \end{bmatrix}$$

But here, because The matrix $\begin{bmatrix} \hat{U} \end{bmatrix}$ is not invertible, so we can't use a diagonal matrix to change it into the reduced row echelon form with full pivots.

So I consider making some change here.

What we already have is

$$\begin{bmatrix} V \end{bmatrix} = (\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix})^{-1} \times \begin{bmatrix} \hat{U} \end{bmatrix}$$

we know that when the elementary matrices are multiplied by the left side of the square matrix, the determinant follows the associative law, in the conclusion 4.4.4,

So we have

$$\begin{aligned} & \det \begin{bmatrix} V \end{bmatrix} \\ = & \det ((\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix})^{-1} \times \begin{bmatrix} \hat{U} \end{bmatrix}) \\ = & \det (\begin{bmatrix} E_k \end{bmatrix} \times \begin{bmatrix} E_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} E_2 \end{bmatrix} \times \begin{bmatrix} E_1 \end{bmatrix}^{-1}) \times \det \begin{bmatrix} \hat{U} \end{bmatrix} \end{aligned}$$

Now look at the $\det \begin{bmatrix} \hat{U} \end{bmatrix}$

We know there is one way to describe a square matrix when it is not invertible, which is saying that the rows of the matrix is not linear independent.

Suppose for the matrix $\begin{bmatrix} \hat{U} \end{bmatrix}$, the ith row is the construction of the rest of the rows.

We have the Multi-Linearity property for the determinant as the formula 4.2.3, which is

$$\det \begin{bmatrix} -v_1 + v_1^{'} - \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

We also have the corollary 4.3.2, which is

$$\det\begin{bmatrix} 0 \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = 0$$

After this method for several times. we will find

$$\begin{aligned} &\det\left[\hat{U}\right] \\ =&\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ -v_i- \\ ... \end{bmatrix} \\ =&\det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \\ 0 \\ ... \end{bmatrix} \\ =&0 \end{aligned}$$

Then we will got $\det \hat{U} = 0$.
So

$$\begin{aligned} &\det\left[V\right] \\ =&\det\left(\left(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right]\right)^{-1} \times \left[\hat{U}\right]\right) \\ =&\det\left(\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right] \times \left[E_1\right]^{-1}\right) \times \det\left[\hat{U}\right] \\ =&0 \end{aligned}$$

Then we will got ***the determinant for any invertible square matrix is zero****.*
So we know if the matrix $\left[S\right]$ *is not invertible, then the matrix* $\left[S^T\right]$ *is also not invertible.*
So we know that $\det\left[S\right] = 0 = \det\left[S^T\right]$.
So we proved that, when the matrix $\left[S\right]$ ***is not invertible we have*** $\det\left[S\right] = \det\left[S^T\right]$

□

4.6.2 Square Matrix Taken Up from the Right

Conclusion 4.6.2.0.1. *When the target square matrix is taken up from the right we will have the determinant follow the associative law, which is* $\det\left(\left[A\right] \times \left[S\right]\right) and \det\left[A\right] \times$

$\det \left[S\right]$ *when the matrix* $\left[S\right]$ *is the target matrix, and the matrix* $\left[A\right]$ *is a random matrix.*

Proof. We have known the conclusion 4.5.1, which is

$$\det\left(\left[S\right] \times \left[A\right]\right) = \det\left[S\right] \times \det\left[A\right]$$

So now we also have to use the conclusion in the last subsection conclusion 4.6.1, which is $\det\left[S\right] = \det\left(\left[S\right]\right)^T$
We will have

$$\begin{aligned}
&\det\left(\left[A\right] \times \left[S\right]\right) \\
=&\det[(\left[S\right]^T \times \left[A\right]^T)^T] \\
=&\det\left(\left[S\right]^T \times \left[A\right]^T\right) \\
=&\det\left[S\right]^T \times \det\left[A\right]^T \\
=&\det\left[S\right] \times \det\left[A\right] \\
=&\det\left[A\right] \times \det\left[S\right]
\end{aligned}$$

□

4.7 The Associative Law and Determinant II

We have already finish the section The Associative Law and Determinant I before, but it only focus on the the situation when the target matrix is taken up from the left, so in this section I will make a conclusion on the associative law in a complete version.

4.7.1 Theorem

Conclusion 4.7.1.0.1.

$$\det\left(\left[A\right] \times \left[B\right]\right) = \det\left[A\right] \times \det\left[B\right]$$

when the matrix $\left[A\right]$ *and the matrix* $\left[B\right]$ *are random square matrix.*

Proof. The full set of proof will not be show here, it is just a conclusion for the previous sections. □

4.7.2 A Brief Introduction on the Application of the Associative Law

The most important thing for the application of the associative law here is it can help you to calculate teh determinant for all the square matrix

Caution 3. *But for a lot of cases, this kind of algorithm is not convenient because what you have you to do is to split the given matrix up to the version of* $(\det[\left[E_k\right] \times \left[E_{k-1}\right] \times ... \times \left[E_2\right]\left[E_1\right])^{-1} \times (\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right]\left[BS_1\right])^{-1}]$

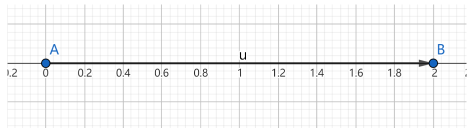


Figure 4.1: the Box of $\begin{bmatrix} 2 \end{bmatrix}$

So there does exist some better ways to calculate the value for the determinant of a square matrix, I will talk about it later, **but I think what I want to tell that there will not be a better one or worse one on all the methods to calculate the determinant**, when you need to calculate the determinant, you also need to think about the simplest way that just introduced because it is the most direct one.

4.8 The Geometric View on Determinant

We have already introduce the basic algebraic rules of the determinant, and I think it is time now to introduce the most important thing to you.

Personally, I love the geometric view of determinant.

4.9 Some Preparation in Geometry

4.9.1 Volume of the Box

Square Matrix and Box

There is a very important conception here called box
In this subsection, we are going to understand the relation ship between the box and the matrix.
I will show you several examples here first before I use words to define the meaning of the box.
And my examples will come from matrix beneath dimension 2 because it will be more easier ot see that

Example 4.40. *The box of the matrix* $\begin{bmatrix} 2 \end{bmatrix}$ *is a line in the one dimensional coordinate system.*

Example 4.41. *The box of the matrix* $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *is a square in the graph which is ABCD in the two dimensional coordinate system.*

Example 4.42. *The box of the matrix* $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$ *is a parallelogram in the graph which is ABCD in the two dimensional coordinate system.*

So know, I think it is time to give out the algebraic definition of the box.

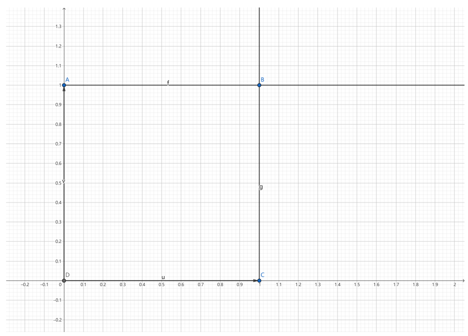


Figure 4.2: the Box of $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

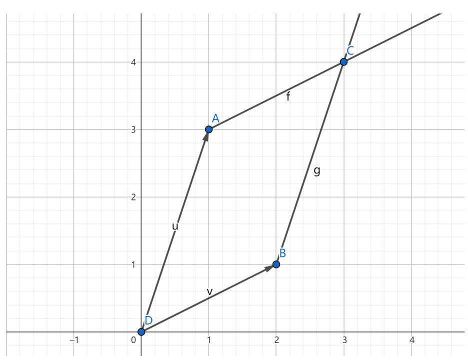


Figure 4.3: the Box of $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

Definition 4.3. *For the square matrix* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$*, we will have half the frame for the box are* v_1*, to* v_n*, when we define the box to be a parallelogram, so you can find out what the box is by adjusting the whole graph into a parallelogram*

The definition is too abstract, so I will make an example here.
I will show all the step to find the box for the matrix $\begin{bmatrix} 2,4 \\ 1,0 \end{bmatrix}$

Step 4.9.1.1. *We find the matrix is* $\begin{bmatrix} 1 & 1 \\ 1 & 0.5 \end{bmatrix}$*, so we can find that two frames are (1,1) and (1,0.5). here is the graph*

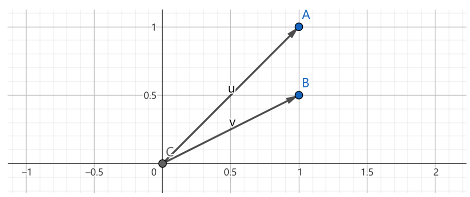


Figure 4.4: Step One

Step 4.9.1.2. *Then you have to make the graph changed into a parallelogram, here is the graph*

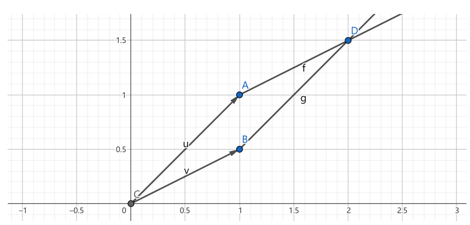


Figure 4.5: Step Two

So I think now you have already had the ability to find the box of the the square matrix, you can see it by eyes in dimension one which is a line; On dimension two, which is a parallelogram; On dimension three, which is a Parallelepiped.
In the higher dimensional space we can't see what the box is actually looks like, but we know that with several perpendicular axis, there is somehow a box, the most thing here is to feel it.

Volume

In the last subsection we have accept the conception of box.

In the real life box is in the three dimensional space, so we define the space in it to be the volume of it.

But in the other spaces like two dimensional space, when the box is a parallelogram, so as you guess, now the volume of the box is just the area.

Just like that, the volume of the box in the one dimensional space is the length.

But these things are not enough, we have to admit that the value of th e volume of the box can be a negative number.

And here I will help you to understand.

Negative volume of Box

If I define this box in the two dimensional space is positive.

Figure 4.6: Positive

Then I think we can define the volume of the box of the following is negetive.

Figure 4.7: Negative

We now define the volume of the box can be negative, every time you flip the box you will add an extra negative sign to the volume of the box.

Conclusion 4.9.1.0.1. *The volume of the box can be a negative number.*

For the absolute value,

in one dimensional space it is length, in two dimensional space it is area,

in three dimensional space it is volume,

and for higher dimensional space, you don't need to see the volume, but to accept and feel that there is something like volume but we now can not understand.

Volume of the Box to be Zero

There do exist situation when the volume of the box is zero, I would love to give you some practical examples that you might understand it better before I gave the definition by words.

Example 4.43. *The volume of a dot in the one dimensional space is zero, because* ***we define the volume of the box in the one dimensional space is the length, and of course we have the length of the box is zero.***

Example 4.44. *The volume of a line or dot in the two dimensional space is zero,*
because ***we define the volume of the box in the two dimensional space is the area, and of course we have the length of the dot or line is zero.***

Definition 4.4. *When the box is a n-1 or less dimensional object in the n dimensional space, then the volume of which is zero.*

Caution 4. *1. The volume of the box depend on the object itself, but also the space.*
2.The most important thing here is dimension, if the volume of the box of a particular matrix is zero, then it is a dimension killer.

4.9.2 Determinant and Volume of the Box

Here, we need to start from the definition of determinant.
I will show every thing in three basic definitions as below.

Normalization

In the normalization part of the determinant we defined $\det(I) = 1$.
Then I think it is best to see why this equation links the volume of the box with the determinant in different conditions.

Condition 4.9.2.0.1 (When $\begin{bmatrix} I \end{bmatrix}$ is $\begin{bmatrix} 1 \end{bmatrix}$)**.** *We will have the graph for the volume of the box here.*

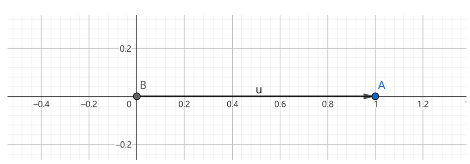


Figure 4.8: Volume of the Box for $\begin{bmatrix} 1 \end{bmatrix}$

Geometrically speaking, the volume of the box now is the length of the line which is 1.
So here we have

When $[I]$ ***is*** $[1]$***,***
det $[I]$ $=$ ***1*** $=$ ***Volume of the Box***

Condition 4.9.2.0.2 (When $[I]$ is $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$)**.** *We will have the graph for the volume of the box here.*

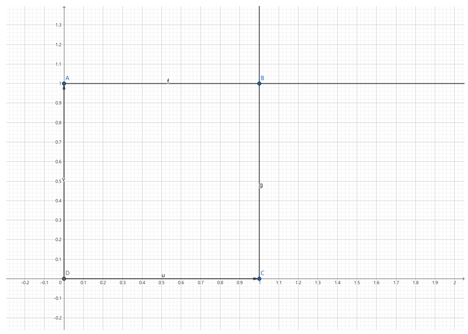


Figure 4.9: Volume of the Box for $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$

Geometrically speaking, the volume of the box now is the area of the square which is 1.
So here we have
When $[I]$ ***is*** $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$***,***
det $[I]$ $=$ ***1*** $=$ ***Volume of the Box***

So for the higher dimensional space it is the same, we can permanently prove that we have
det $[I]$ $=$ **1** $=$ **Volume of the Box**

Conclusion 4.9.2.0.1. det $[I]$ $=$ *1* $=$ *Volume of the Box*

Alternating

From the alternating property of the determinant, we have the function 4.2.2, which is

$$\det\begin{bmatrix}-v_1-\\-v_2-\\-v_3-\\\ldots\end{bmatrix} = -\det\begin{bmatrix}-v_2-\\-v_1-\\-v_3-\\\ldots\end{bmatrix}$$

We know this shows some kinds of linear transformation, then I think it is best to see why this equation links the volume of the box with the determinant in different conditions.

Condition 4.9.2.0.3 (When in One Dimensional Space)**.** *We can't do row exchange for lack of row.*

Condition 4.9.2.0.4 (When in Two Dimensional Space)**.** *We set* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \cdots \end{bmatrix}$ *to be* $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$,

So we will find that in the determinant of the matrix after the row exchange is opposite to the one before exchange.

And we now look at the change in the volume of the box.

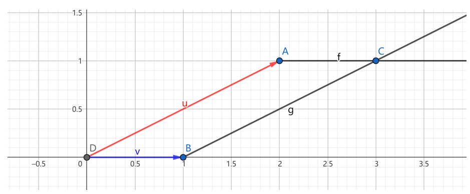


Figure 4.10: Before Row Exchange

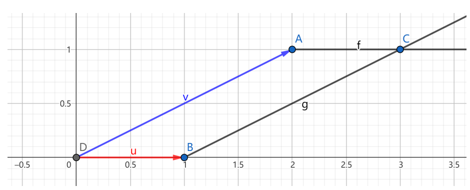


Figure 4.11: After Row Exchange

So it's obviously the same as the situation of flipping as we have mentioned before.

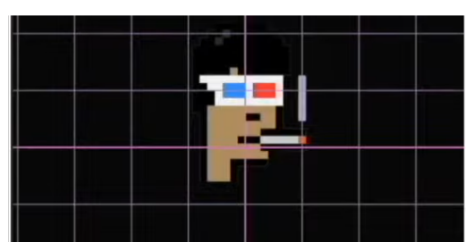


Figure 4.12: Before Row Exchange

Figure 4.13: After Row Exchange

So we will find that after the row exchange, the volume of the box is opposite to which before row exchange.

So now what we find is if you make a row exchange, the corresponding change in the determinant and the volume of the box is the same.
I can say that for any matrices come from identity matrix after row exchange, their determinant and the volume of the box are the same.

Conclusion 4.9.2.0.2. det $\begin{bmatrix} A \end{bmatrix}$ = *Volume of the Box*
When the $\begin{bmatrix} A \end{bmatrix}$ *is in form of*

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} P_k \end{bmatrix} \times \begin{bmatrix} P_{k-1} \end{bmatrix} \times ... \times \begin{bmatrix} P_2 \end{bmatrix} \times \begin{bmatrix} P_1 \end{bmatrix} \times \begin{bmatrix} P_I \end{bmatrix}$$

Caution 5. *What the row exchange is doing is just flipping the box.*

Multi-Linearity I

From the first part of property of Multi-Linearity we have the function 4.2.3, which is

$$\det \begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det \begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

We know this shows some kinds of linear transformation, then I think it is best to see why this equation links the volume of the box with the determinant in different conditions.

Condition 4.9.2.0.5 (When in the One Dimensional Space)**.** *If we set the matrix* $\begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ *to be* $\begin{bmatrix} 3 \end{bmatrix}$, *and the matrix* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ *and matrix* $\begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$ *to be* $\begin{bmatrix} 2 \end{bmatrix}$ *and* $\begin{bmatrix} 1 \end{bmatrix}$ *respectively.*
we have known the determinant satisfies this equation, but what about the volume of the box?
Here we will show you the graphs.

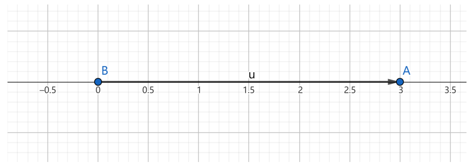


Figure 4.14: Before Split

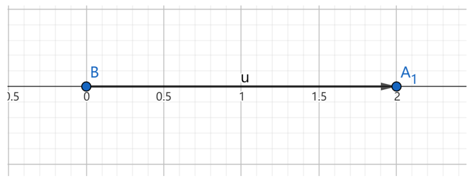


Figure 4.15: Part I After Split

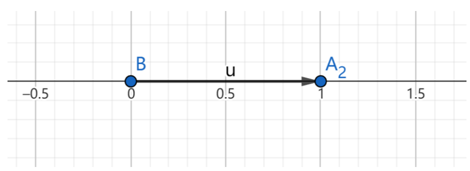


Figure 4.16: Part II After Split

we will see that the volume of the box of the matrix before split is 3, the part I and part II are 2 and 1 respectively. Then we find

$$Volumn\ of\ Box\ (Before\ Split)\ =\ Volumn\ of\ Box\ (Part\ I) + Volumn\ of\ Box\ (Part\ II)$$

Condition 4.9.2.0.6 (When in the Two Dimensional Space)**.** *If we set the matrix* $\begin{bmatrix} -v_1 + v_1^{'} - \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 3 \end{bmatrix}$*, and the matrix* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *and matrix* $\begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 1 \end{bmatrix}$ *and* $\begin{bmatrix} 2 \end{bmatrix}$ *respectively.*

we have known the determinant satisfies this equation, but what about the volume of the box?

Here we will show you the graphs.

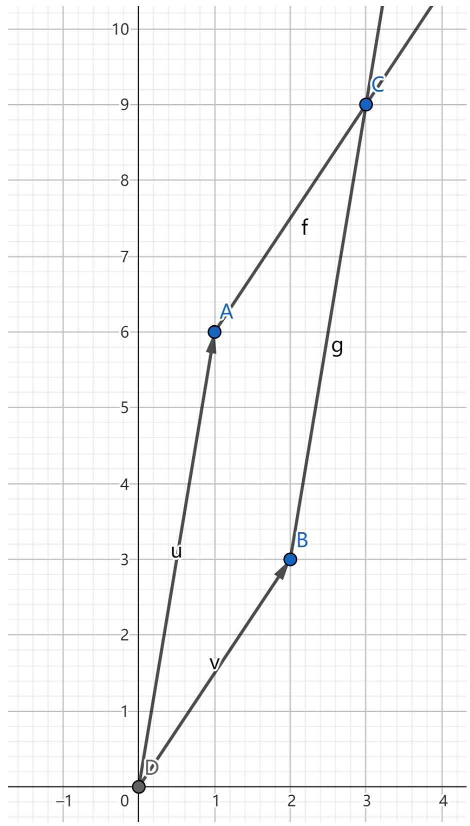


Figure 4.17: Before Split

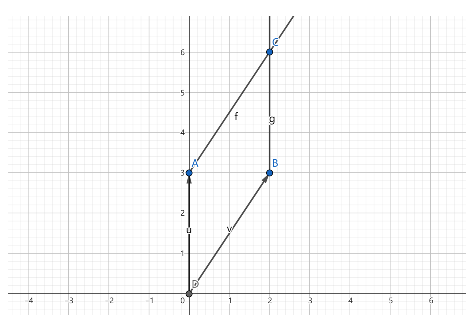


Figure 4.18: Part I After Split

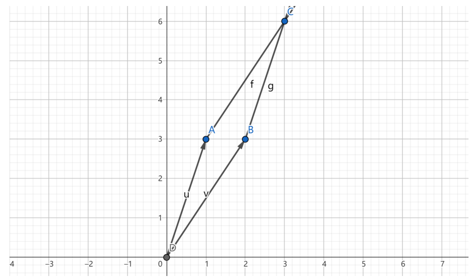


Figure 4.19: Part II After Split

We have the area of original parallelogram is -9, the area of the parallelogram part I after split is -6, the area of the parallelogram part II after split is -3. we find that in this situation the sum of the area of the part I and part II parallelogram is just equal to the area of the original parallelogram.
Geometrically which is

$$Volumn\ of\ Box\ (Before\ Split)\ =\ Volumn\ of\ Box\ (Part\ I) + Volumn\ of\ Box\ (Part\ II)$$

So we found that if you make a row exchange, the corresponding change in the determinant and the volume of the box is the same.
And we know that **any matrix can be the combination of a lot of other matrix, and those matrix can finally come to the identity matrix through specific method.**

Conclusion 4.9.2.0.3. *For any matrix* $\begin{bmatrix} A \end{bmatrix}$
we have $\det \begin{bmatrix} A \end{bmatrix}$ $=$ *Volume of the Box*

Multi-Linearity II

Actually I think there is no need to tell story about the second rule on Multi-Linearity, because actually the part II of Multi-Linearity is just a block with multiple times of part I Multi-Linearity.
But I think it is better to take a look at it again.
From the rule of part II of Multi-Linearity, we have the equation 4.2.3, which is

$$\det \begin{bmatrix} -cv_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = c \times \det \begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

We know this shows some kinds of linear transformation, then I think it is best to see why this equation links the volume of the box with the determinant in different conditions.

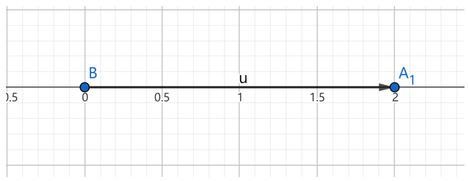


Figure 4.20: Before Degenerated

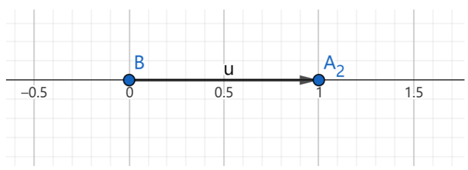


Figure 4.21: After Degenerated

Condition 4.9.2.0.7 (When in the One Dimensional Space)**.** *If we set the matrix* $\begin{bmatrix} -cv_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 2 \end{bmatrix}$*, and the matrix* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 1 \end{bmatrix}$
we have known the determinant satisfies this equation, but what about the volume of the box?
Here we will show you the graphs.
we will see that the volume of the box of the matrix before degeneration is 2, after degeneration is 1.
Then we find

$$Volumn\ of\ Box\ (Before\ Degeneration)\ =\ c \times Volumn\ of\ Box\ (AfterDegeneration)$$

Condition 4.9.2.0.8 (When in the Two Dimensional Space)**.** *If we set the matrix* $\begin{bmatrix} -c \times v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 4, 2 \\ 3, 1 \end{bmatrix}$*, and the matrix* $\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$ *to be* $\begin{bmatrix} 2, 1 \\ 3, 1 \end{bmatrix}$.
we have known the determinant satisfies this equation, but what about the volume of the box?
Here we will show you the graphs.
We have the area of original parallelogram is -2, the area of the parallelogram after degeneration is -1, and the value of c is 2.
Geometrically which is

$$Volumn\ of\ Box\ (Before\ Degeneration)\ =\ c \times Volumn\ of\ Box\ (AfterDegeneration)$$

So we found that if you make a row exchange, the corresponding change in the determinant and the volume of the box is the same.

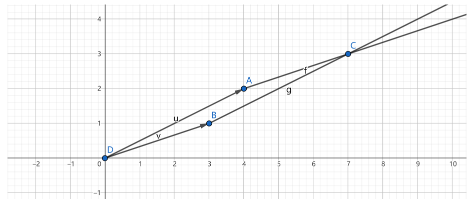


Figure 4.22: Before Degeneration

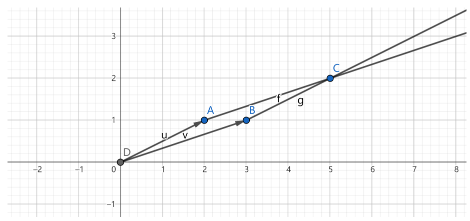


Figure 4.23: After Degeneration

And we know that **any matrix can be the combination of a lot of other matrix, and those matrix can finally come to the identity matrix through specific method.**

Conclusion 4.9.2.0.4. *For any matrix* $\begin{bmatrix} A \end{bmatrix}$
we have $\det \begin{bmatrix} A \end{bmatrix} =$ *Volume of the Box*

Square Matrix

In this subsection I just want to make a conclusion of this section.
We have known that for the identity matrix we have the volume of the box is just the determinant, and for the determinant and the volume of the box, they has the same computation rule when doing Multi-Linearity, Alternating.
Every matrix comes from the identity matrix after the computation of Multi-Linearity and Alternating.

Conclusion 4.9.2.0.5. *For any square matrix* $\begin{bmatrix} A \end{bmatrix}$
we have $\det \begin{bmatrix} A \end{bmatrix} =$ *Volume of the Box*

4.10 Determinant Calculation & Formula for Determinant

This part is very skillful, because there are many ways to calculate the determinant of a matrix.
All the methods are correct when only some of them are convenient.
In this section I will try to introduce all the general method to you.

4.10.1 Incomplete Split

Conclusion 4.10.1.0.1. *When the size of matrix* $\left[A\right]$ *is n by n*

$$\det\left[A\right] = \Sigma_{k=1}^{n} a_{kj} \times c_{kj}$$

or

$$\det\left[A\right] = \Sigma_{k=1}^{n} a_{ik} \times c_{ik}$$

When $c_{mn} = (-1)^{m+n} \times \det\left[A_{mn}\right]$, *and* $\left[A_{mn}\right]$ *is the matrix delete the* m^{th} *row and* n^{th} *column.*

Algebraic Proof

we have the conclusion 4.10.1

Proof. I am going to prove $\det\left[A\right] = \Sigma_{k=1}^{n} a_{ik} \times c_{ik}$ first.

We know the property of the determinant called the first part of Multi-Linearity, as the function 4.2.3, which is

$$\det\begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} = \det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix} + \det\begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ \dots \end{bmatrix}$$

If we set the matrix $\left[A\right]$ with the size of $n \times n$ to be

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} & \dots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Then we will use the function 4.2.3 multiple times.

If we take the i^{th} row to split, that is

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} & \cdots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$= \det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ a_{i1} & 0 & 0 & 0 & \cdots & 0 \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$+ \det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ 0 & a_{i2} & a_{i3} & a_{i4} & \cdots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$= \det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ a_{i1} & 0 & 0 & 0 & \cdots & 0 \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$+\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ 0 & a_{i2} & 0 & 0 & \cdots & 0 \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$+\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ 0 & 0 & a_{i3} & a_{i4} & \cdots & a_{in} \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$=\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ a_{i1} & 0 & 0 & 0 & \cdots & 0 \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$+\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \cdots & a_{(i-1)n} \\ 0 & a_{i2} & 0 & 0 & \cdots & 0 \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \cdots & a_{(i+1)n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

$$
+\det\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\
a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\
a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\
0 & 0 & a_{i3} & 0 & \dots & 0 \\
a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\
\dots & \dots & \dots & \dots & \dots & \dots
\end{bmatrix}
$$

$$
+\det\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\
a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\
a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\
0 & 0 & 0 & a_{i4} & \dots & 0 \\
a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\
\dots & \dots & \dots & \dots & \dots & \dots
\end{bmatrix}
$$

$$
+\dots
$$

$$
+\det\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\
a_{21} & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\
a_{31} & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\
0 & 0 & 0 & 0 & \dots & a_{in} \\
a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\
\dots & \dots & \dots & \dots & \dots & \dots
\end{bmatrix}
$$

And we have to use the Alternating property of determinant to transport the entries of the a_{ik} to the Seattle corner.

Then we continue with our workout just now, we have

$$
\begin{aligned}
&\det\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & ... & a_{1n}\\
a_{21} & a_{22} & a_{23} & a_{24} & ... & a_{2n}\\
a_{31} & a_{32} & a_{33} & a_{34} & ... & a_{3n}\\
... & ... & ... & ... & ... & ...\\
a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n}\\
a_{i1} & a_{i2} & a_{i3} & a_{i4} & ... & a_{in}\\
a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n}\\
... & ... & ... & ... & ... & ...
\end{bmatrix}\\
&=(-1)^{i-1}\times\det\begin{bmatrix}
a_{i1} & 0 & 0 & 0 & ... & 0\\
a_{11} & a_{12} & a_{13} & a_{14} & ... & a_{1n}\\
a_{21} & a_{22} & a_{23} & a_{24} & ... & a_{2n}\\
a_{31} & a_{32} & a_{33} & a_{34} & ... & a_{3n}\\
... & ... & ... & ... & ... & ...\\
a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n}\\
a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n}\\
... & ... & ... & ... & ... & ...
\end{bmatrix}\\
&+(-1)^{i}\times\det\begin{bmatrix}
a_{i2} & 0 & 0 & 0 & ... & 0\\
a_{12} & a_{11} & a_{13} & a_{14} & ... & a_{1n}\\
a_{22} & a_{21} & a_{23} & a_{24} & ... & a_{2n}\\
a_{32} & a_{31} & a_{33} & a_{34} & ... & a_{3n}\\
... & ... & ... & ... & ... & ...\\
a_{(i-1)2} & a_{(i-1)1} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n}\\
a_{(i+1)2} & a_{(i+1)1} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n}\\
... & ... & ... & ... & ... & ...
\end{bmatrix}\\
&+...
\end{aligned}
$$

$$
\begin{aligned}
&=(-1)^{i+1} \times \det \begin{bmatrix} a_{i1} & 0 & 0 & 0 & \dots & 0 \\ 0 & a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\ 0 & a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\ 0 & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \\
&+(-1)^{i+2} \times \det \begin{bmatrix} a_{i2} & 0 & 0 & 0 & \dots & 0 \\ 0 & a_{11} & a_{13} & a_{14} & \dots & a_{1n} \\ 0 & a_{21} & a_{23} & a_{24} & \dots & a_{2n} \\ 0 & a_{31} & a_{33} & a_{34} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{(i-1)1} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\ 0 & a_{(i+1)1} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \\
&+\dots \\
&=(-1)^{i+1} \times a_{i1} \times \det \begin{bmatrix} a_{12} & a_{13} & a_{14} & \dots & a_{1n} \\ a_{22} & a_{23} & a_{24} & \dots & a_{2n} \\ a_{32} & a_{33} & a_{34} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\ a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \\
&+(-1)^{i+2} \times a_{i2} \times \det \begin{bmatrix} a_{11} & a_{13} & a_{14} & \dots & a_{1n} \\ a_{21} & a_{23} & a_{24} & \dots & a_{2n} \\ a_{31} & a_{33} & a_{34} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{(i-1)1} & a_{(i-1)3} & a_{(i-1)4} & \dots & a_{(i-1)n} \\ a_{(i+1)1} & a_{(i+1)3} & a_{(i+1)4} & \dots & a_{(i+1)n} \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \\
&+\dots
\end{aligned}
$$

If now we set $c_{mn} = (-1)^{m+n} \times \det \left[A_{mn}\right]$, and $\left[A_{mn}\right]$ is the matrix delete the m^{th} row and n^{th} column.

We will have

$$
\begin{aligned}
&\det A \\
&= a_{i1} \times ci1 + a_{i2} \times ci2 + \dots \\
&= \Sigma_{k=1}^{n} a_{ik} \times c_{ik}
\end{aligned}
$$

So we proved that $\det \left[A\right] = \Sigma_{k=1}^{n} a_{ik} \times c_{ik}$

And for the situation of $\det \left[A\right] = \Sigma_{k=1}^{n} a_{kj} \times c_{kj}$, I will use one of the conclusion we have before, the conclusion 4.6.1, which is $\det \left[S\right] = \det\left(\left[S\right]\right)^T$.

So that we have

$$\begin{aligned}&\det\left[A\right]\\&=\det\left[A^T\right]\\&=\Sigma_{k=1}^{n}a_{ki}\times c_{ki}\\&=\Sigma_{k=1}^{n}a_{kj}\times c_{kj}\end{aligned}$$

Then we have proved that When the size of matrix $\left[A\right]$ is n by n

$$\det\left[A\right]=\Sigma_{k=1}^{n}a_{kj}\times c_{kj}$$

or

$$\det\left[A\right]=\Sigma_{k=1}^{n}a_{ik}\times c_{ik}$$

When $c_{mn}=(-1)^{m+n}\times\det\left[A_{mn}\right]$, and $\left[A_{mn}\right]$ is the matrix delete the m^{th} row and n^{th} column. □

Geometric Explanation

This is the application of the method of split and degenerate as we have mentioned in the proof of conclusion 4.9.2.

There is a very good explanation for it, and I will do it with pictures.

Step 4.10.1.1. ***Here is a graph of box which I want to get the value of, the corresponding matrix is*** $\begin{bmatrix}3&1\\1&3\end{bmatrix}$

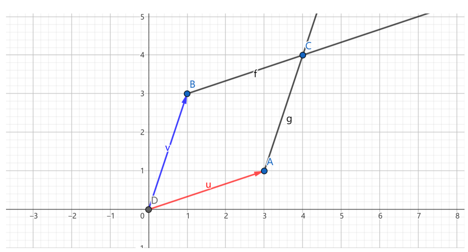


Figure 4.24: Box of Matrix $\begin{bmatrix}3&1\\1&3\end{bmatrix}$

Let's suppose we can't calculate the area out directly.

Step 4.10.1.2. *So we have to split this matrix into two matrices.*

If we split the matrix by the first row as an example.

We will have $\det\begin{bmatrix}3&1\\1&3\end{bmatrix}=\det\begin{bmatrix}3&0\\1&3\end{bmatrix}+\det\begin{bmatrix}0&1\\1&3\end{bmatrix}$

I will sign $\begin{bmatrix}3&0\\1&3\end{bmatrix}$ *as part I,* $\begin{bmatrix}0&1\\1&3\end{bmatrix}$ *as part II*

Then I will show what we do geometrically.

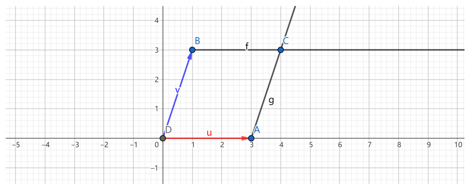


Figure 4.25: Box of Part I

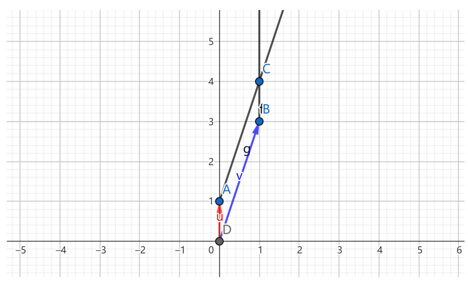


Figure 4.26: Box of Part II

Caution 6. *You have to remember the volume of the box here we have chance to be negative, such as the one of the volume of the box for the part II*

But let's suppose we still can't calculate the area of each box

Step 4.10.1.3. *We continue to make it easier to calculate the volume of the box.*

I want to mention a step we did when proving the conclusion 4.10.1, which is

$$(-1)^{i-1} \times \det \begin{bmatrix} a_{i1} & 0 & 0 & 0 & ... & 0 \\ a_{11} & a_{12} & a_{13} & a_{14} & ... & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & ... & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & ... & a_{3n} \\ ... & ... & ... & ... & ... & ... \\ a_{(i-1)1} & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n} \\ a_{(i+1)1} & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n} \\ ... & ... & ... & ... & ... & ... \end{bmatrix}$$

$$+(-1)^{i} \times \det \begin{bmatrix} a_{i2} & 0 & 0 & 0 & ... & 0 \\ a_{12} & a_{11} & a_{13} & a_{14} & ... & a_{1n} \\ a_{22} & a_{21} & a_{23} & a_{24} & ... & a_{2n} \\ a_{32} & a_{31} & a_{33} & a_{34} & ... & a_{3n} \\ ... & ... & ... & ... & ... & ... \\ a_{(i-1)2} & a_{(i-1)1} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n} \\ a_{(i+1)2} & a_{(i+1)1} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n} \\ ... & ... & ... & ... & ... & ... \end{bmatrix}$$

$$+...$$

$$=(-1)^{i+1} \times \det \begin{bmatrix} a_{i1} & 0 & 0 & 0 & ... & 0 \\ 0 & a_{12} & a_{13} & a_{14} & ... & a_{1n} \\ 0 & a_{22} & a_{23} & a_{24} & ... & a_{2n} \\ 0 & a_{32} & a_{33} & a_{34} & ... & a_{3n} \\ ... & ... & ... & ... & ... & ... \\ 0 & a_{(i-1)2} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n} \\ 0 & a_{(i+1)2} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n} \\ ... & ... & ... & ... & ... & ... \end{bmatrix}$$

$$+(-1)^{i+2} \times \det \begin{bmatrix} a_{i2} & 0 & 0 & 0 & ... & 0 \\ 0 & a_{11} & a_{13} & a_{14} & ... & a_{1n} \\ 0 & a_{21} & a_{23} & a_{24} & ... & a_{2n} \\ 0 & a_{31} & a_{33} & a_{34} & ... & a_{3n} \\ ... & ... & ... & ... & ... & ... \\ 0 & a_{(i-1)1} & a_{(i-1)3} & a_{(i-1)4} & ... & a_{(i-1)n} \\ 0 & a_{(i+1)1} & a_{(i+1)3} & a_{(i+1)4} & ... & a_{(i+1)n} \\ ... & ... & ... & ... & ... & ... \end{bmatrix}$$

$$+...$$

This is just changing the shape of the box

Then I will call $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ *to be the part I after shape change and* $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *to be the part II after shape change.*

Here are the graphs.

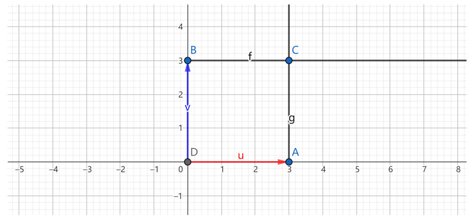


Figure 4.27: Part I After Shape Change

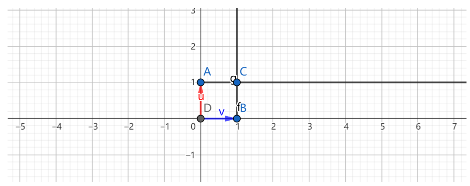


Figure 4.28: Part II After Shape Change

We will see that this step change the parallelogram into rectangle. And now it is very easy to calculate the determinant.

So what is the Geometric Meaning of each part of function

$$\det \left[A\right] = \Sigma_{k=1}^{n} a_{ik} \times c_{ik}$$

a_{ik} is like the height of a sub part.
c_{ik} is like the area of the base perpendicular to the height.

The basic skill here is still split and change the shape.

4.10.2 Complete Split & Big Formula

Conclusion 4.10.2.0.1.

$$\det \left[A\right] = \Sigma_{all\ P's}(a_{1\alpha_1} a_{2\alpha_2} \times ... \times a_{n\alpha_n}) \times \det \left[P\right]$$

when α_i takes any integer value between 1 and n when all of them will not be repeated, and $\left[P\right]$ is all the permutation matrix they corresponding to.

This formula is the most important one in the linear algebra, because it gives the full picture of determinant split.

Algebraic Proof

This is just a more complete split then the formula
$\det \left[A\right] = \Sigma_{k=1}^{n} a_{ik} \times c_{ik}$
I don't need to differ the situation when splitting by the column or the row because this time all the column and rows will be fully split up.

Proof. We know the property of the determinant called the first part of Multi-Linearity, as the function 4.2.3, which is

$$\det\begin{bmatrix} -v_1 + v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} + \det\begin{bmatrix} -v_1^{'}- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

Also we have the corollary 4.3.3, which is

$$\det\begin{bmatrix} -v_1 + c \times v_2- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix} = \det\begin{bmatrix} -v_1- \\ -v_2- \\ -v_3- \\ ... \end{bmatrix}$$

If we set the matrix $\begin{bmatrix} A \end{bmatrix}$ with the size of $n \times n$ to be

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix}$$

Then we will use the function 4.2.3 multiple times.
If we take the first row to split, that is

$$\begin{aligned}
&\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
=&\det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
=&\det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & a_{23} & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & a_{23} & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & a_{22} & 0 & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix}
\end{aligned}$$

And now we need to continue to split, I will try with the second row

$$
\begin{aligned}
&\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&= \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & a_{23} & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & a_{23} & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & a_{22} & 0 & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&+ \det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&= \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & 0 & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&+ \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & 0 & 0 & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&+ \det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & 0 & 0 & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&+ \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&= \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
&+ \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix}
\end{aligned}
$$

$$+\det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix}$$

$$+\det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix}$$

And we need to split the third row

$$
\begin{aligned}
&\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & ... \\ a_{21} & a_{22} & a_{23} & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
=&\det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & 0 & ... \\ a_{31} & a_{32} & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
=&\det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} a_{11} & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & a_{12} & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & a_{13} & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ a_{21} & 0 & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} \\
+&\det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & a_{22} & 0 & ... \\ 0 & 0 & a_{33} & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ a_{31} & 0 & 0 & .. \\ ... & ... & ... & ... \end{bmatrix} + \det\begin{bmatrix} 0 & 0 & 0 & ... \\ 0 & 0 & a_{23} & ... \\ 0 & a_{32} & 0 & .. \\ ... & ... & ... & ... \end{bmatrix}
\end{aligned}
$$

$$+\det\begin{bmatrix}0&0&0&...\\0&0&0&...\\a_{31}&0&0&..\\...&...&...&...\end{bmatrix}+\det\begin{bmatrix}0&0&0&...\\0&0&0&...\\0&a_{32}&0&..\\...&...&...&...\end{bmatrix}+\det\begin{bmatrix}0&0&0&...\\0&0&0&...\\0&0&a_{33}&..\\...&...&...&...\end{bmatrix}$$
$$+\det\begin{bmatrix}0&0&0&...\\0&0&0&...\\0&0&0&..\\...&...&...&...\end{bmatrix}$$
$$=\Sigma_{all\ P's}(a_{1\alpha_1}a_{2\alpha_2}\times...\times a_{n\alpha_n})\times\det\begin{bmatrix}P\end{bmatrix}$$

□

Some Confusion

I use this place to classify some conception in the big formula which is

$$\det\begin{bmatrix}A\end{bmatrix}=\Sigma_{all\ P's}(a_{1\alpha_1}a_{2\alpha_2}\times...\times a_{n\alpha_n})\times\det\begin{bmatrix}P\end{bmatrix}$$

I think there are basically two things need further explanation.

1. What is exactly the matrix $\begin{bmatrix}P\end{bmatrix}$ in the equation.
2. How can we calculate the determinant of matrix $\begin{bmatrix}P\end{bmatrix}$ fast.

So I will **start with the first question**.
I think it is good to give example to help you understand.
I will just say the permutation matrix for the matrix $\begin{bmatrix}1,0\\0,-9\end{bmatrix}$ is just the matrix $\begin{bmatrix}1,0\\0,1\end{bmatrix}$.

And the permutation matrix $\begin{bmatrix}0,4,0\\0,0,7\\-9,0,0\end{bmatrix}$ is just the matrix $\begin{bmatrix}0,1,0\\0,0,1\\1,0,0\end{bmatrix}$
Now is I will **give answer to the second problem.**
I will also give you answer by examples.

Example 4.45. *Find the determinant of the matrix* $\begin{bmatrix}0,1,0,0,0\\0,0,0,1,0\\0,0,1,0,0\\0,0,0,0,1\\1,0,0,0,0\end{bmatrix}$

Solution:
You have several options.
For example, you can calculate it by basic rules or big formulas.
But I will give you a convenient short cut here.
We now the matrix is $\begin{bmatrix}0,1,0,0,0\\0,0,0,1,0\\0,0,1,0,0\\0,0,0,0,1\\1,0,0,0,0\end{bmatrix}$ *, then we can just record it as (2,4,3,5,1) according to the position of one in each row.*

Now we will have

$$(2,4,3,5,1) \xrightarrow{I} (2,3,4,5,1) \xrightarrow{II} (2,3,4,1,5) \xrightarrow{III} (2,3,1,4,5) \xrightarrow{IV} (2,1,3,4,5) \xrightarrow{V} (1,2,3,4,5)$$

We will see that it take 5 steps to change the permutation matrix into the identity matrix.
So the linear transformation it represent is the combination of 5 simple permutation matrix.
We have proved that associative law for the determinant operation, then we will have $\det \begin{bmatrix} 0,1,0,0,0 \\ 0,0,0,1,0 \\ 0,0,1,0,0 \\ 0,0,0,0,1 \\ 1,0,0,0,0 \end{bmatrix} = (-1)^5 = -1$

Caution 7. *The basic logic here is to* ***find how many row operations are this permutation matrix represent****, because we know each single simple row exchange means to flip the box for one time, which will add a negative sign to the volume of the box.*

Geometric Explanation

We have known that the Geometric meaning of cofacter method is a slight split on the box, so here it is a complete split of box.
Let's make a conclusion on the things geometrically we do on the box when we are calculating the determinant.

1. **Split the box.**

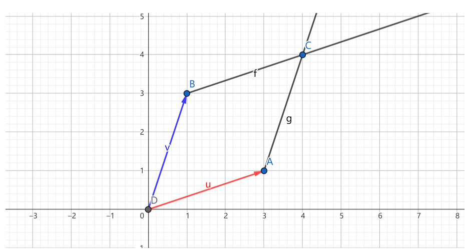


Figure 4.29: Before Explit

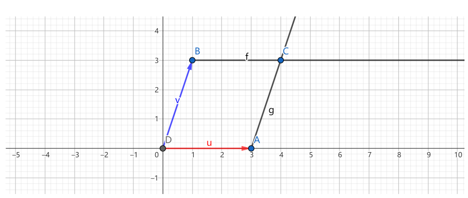


Figure 4.30: After Split Part I

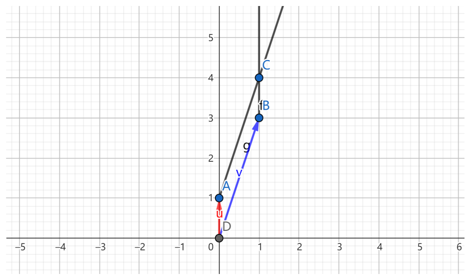


Figure 4.31: After Split Part II

2. Degenerate the Box

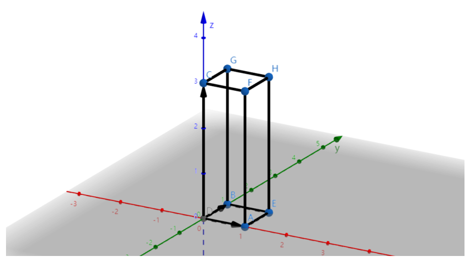


Figure 4.32: Before Degeneration

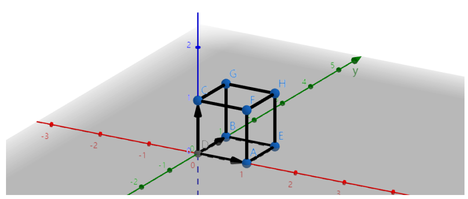


Figure 4.33: After Degeneration by 3

This step can always help you to degenerate the parallelepiped to the standard unit cubic that the identity matrix represent

3. Evaluate Everything by $\det\left[I\right]=1$

4.10.3 Upper Triangular Matrix

Conclusion 4.10.3.0.1. *For a upper triangular matrix* $\left[U\right]$ *with size* $n\times n$*, we have*

$$\det\left[U\right]=\prod_{i=1}^{n}u_{ii}$$

Proof. We have the knowledge on Gaussian Elimination, so $\left[U\right] = (\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right])^{-1}$ Then we will have

$$\begin{aligned} &\det \left[U\right] \\ =&\det [(\left[D_1\right])^{-1} \times (\left[BS_k\right] \times \left[BS_{k-1}\right] \times ... \times \left[BS_2\right] \left[BS_1\right])^{-1}] \\ =&\det (\left[D_1\right])^{-1} \\ =&\prod_{i=1}^{n} u_{ii} \end{aligned}$$

□

4.10.4 Formula for Specific Dimensional Space

I don't think this part is important, I think it is more like an application of the big formula
$\det \left[A\right] = \Sigma_{all\ P's}(a_{1\alpha_1} a_{2\alpha_2} \times ... \times a_{n\alpha_n}) \times \det \left[P\right]$
So I will not give proof for the conclusion in this subsection.

Two Dimensional Space

Conclusion 4.10.4.0.1.

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - cb$$

Three Dimensional Space

Conclusion 4.10.4.0.2.

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

4.10.5 Sample Problem

In the former section we learned several way to calculate the determinant, it is time to use those tricks now.

Problem 4.10.5.0.1. $A = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & ... & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & ... & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & ... & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & 0 & ... & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & ... & -1 & 2 \end{bmatrix}$, *find* $\det A$

Solution:

$$\begin{aligned}
&\det\left[A\right]\\
=&\det\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 & 0\\ -1 & 2 & -1 & 0 & 0 & ... & 0 & 0\\ 0 & -1 & 2 & -1 & 0 & ... & 0 & 0\\ 0 & 0 & -1 & 2 & -1 & ... & 0 & 0\\ 0 & 0 & 0 & -1 & 2 & ... & 0 & 0\\ ... & ... & ... & ... & ... & ... & ... & ...\\ 0 & 0 & 0 & 0 & 0 & ... & 2 & -1\\ 0 & 0 & 0 & 0 & 0 & ... & -1 & 2 \end{bmatrix}\\
=&\det\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 & 0\\ 0 & \frac{3}{2} & -1 & 0 & 0 & ... & 0 & 0\\ 0 & -1 & 2 & -1 & 0 & ... & 0 & 0\\ 0 & 0 & -1 & 2 & -1 & ... & 0 & 0\\ 0 & 0 & 0 & -1 & 2 & ... & 0 & 0\\ ... & ... & ... & ... & ... & ... & ... & ...\\ 0 & 0 & 0 & 0 & 0 & ... & 2 & -1\\ 0 & 0 & 0 & 0 & 0 & ... & -1 & 2 \end{bmatrix}\\
=&\det\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 & 0\\ 0 & \frac{3}{2} & -1 & 0 & 0 & ... & 0 & 0\\ 0 & 0 & \frac{4}{3} & -1 & 0 & ... & 0 & 0\\ 0 & 0 & -1 & 2 & -1 & ... & 0 & 0\\ 0 & 0 & 0 & -1 & 2 & ... & 0 & 0\\ ... & ... & ... & ... & ... & ... & ... & ...\\ 0 & 0 & 0 & 0 & 0 & ... & 2 & -1\\ 0 & 0 & 0 & 0 & 0 & ... & -1 & 2 \end{bmatrix}\\
=&...\\
=&\det\begin{bmatrix} 2 & -1 & 0 & 0 & 0 & ... & 0 & 0\\ 0 & \frac{3}{2} & -1 & 0 & 0 & ... & 0 & 0\\ 0 & 0 & \frac{4}{3} & -1 & 0 & ... & 0 & 0\\ 0 & 0 & 0 & \frac{5}{4} & -1 & ... & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{6}{5} & ... & 0 & 0\\ ... & ... & ... & ... & ... & ... & ... & ...\\ 0 & 0 & 0 & 0 & 0 & ... & \frac{n}{n-1} & -1\\ 0 & 0 & 0 & 0 & 0 & ... & 0 & \frac{n+1}{n} \end{bmatrix}\\
=&2\times\frac{3}{2}\times\frac{4}{3}\times...\times\frac{n+1}{n}\\
=&n+1
\end{aligned}$$

Problem 4.10.5.0.2.

$$\left[A\right]=\begin{bmatrix} a_1+b_1 & b_1+c_1 & c_1+a_1\\ a_2+b_2 & b_2+c_2 & c_2+a_2\\ a_3+b_3 & b_3+c_3 & c_3+a_3 \end{bmatrix}, find\det\left[A\right]$$

Solution:

$$
\begin{aligned}
&\det\left[A\right]\\
=&\det\left(\left[A\right]\begin{bmatrix}1&0&0\\-1&1&0\\1&0&1\end{bmatrix}\right)\\
=&\det\begin{bmatrix}2\times a_1 & b_1+c_1 & c_1+a_1\\2\times a_2 & b_2+c_2 & c_2+a_2\\2\times a_3 & b_3+c_3 & c_3+a_3\end{bmatrix}\\
=&2\times\det\begin{bmatrix}a_1 & b_1+c_1 & c_1+a_1\\a_2 & b_2+c_2 & c_2+a_2\\a_3 & b_3+c_3 & c_3+a_3\end{bmatrix}\\
=&2\times\det\begin{bmatrix}a_1 & b_1+c_1 & c_1\\a_2 & b_2+c_2 & c_2\\a_3 & b_3+c_3 & c_3\end{bmatrix}\\
=&2\times\det\begin{bmatrix}a_1 & b_1 & c_1\\a_2 & b_2 & c_2\\a_3 & b_3 & c_3\end{bmatrix}
\end{aligned}
$$

Problem 4.10.5.0.3. $\left[A\right]=\begin{bmatrix}a&b&b&...&b\\b&a&b&...&b\\b&b&a&...&b\\...&...&...&...&...\\b&b&b&...&a\end{bmatrix}$, *find* $\det\left[A\right]$

Solution: (**Say** 4×4): $\left[A\right]=\begin{bmatrix}a&b&b&b\\b&a&b&b\\b&b&a&b\\b&b&b&a\end{bmatrix}$

$$\begin{aligned}
&\det\left[A\right]\\
=&\det\left(\begin{bmatrix}1&1&1&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\left[A\right]\right)\\
=&\det\begin{bmatrix}a+3b&a+3b&a+3b&a+3b\\b&a&b&b\\b&b&a&b\\b&b&b&a\end{bmatrix}\\
=&(a+3b)\times\det\begin{bmatrix}1&1&1&1\\b&a&b&b\\b&b&a&b\\b&b&b&a\end{bmatrix}\\
=&(a+3b)\times\det\begin{bmatrix}1&1&1&1\\0&a-b&0&0\\0&0&a-b&0\\0&0&0&a-b\end{bmatrix}\\
=&(a+3b)(a-b)^3
\end{aligned}$$

general $(n\times n)$:

$$(a+(n-1)b)(a-b)^{n-1}$$

Problem 4.10.5.0.4. $\left[A_n\right]=\begin{bmatrix}a&b&0&...&0&0\\c&a&b&...&0&0\\0&c&a&...&0&0\\...&...&...&...&...&...\\0&0&0&...&a&b\\0&0&0&...&c&a\end{bmatrix}$, *find* $\det\left[A\right]$

Solution: If we try to split by the first column.
We will find

$$\det\left[A_n\right]=a\times\det\left[A_{n-1}\right]-cb\times\det\left[A_{n-2}\right]$$

So what we have is

$$\begin{cases}\det\left[A_n\right]=a\times\det\left[A_{n-1}\right]-bc\times\det\left[A_{n-2}\right]\\ \det\left[A_1\right]=a\\ \det\left[A_2\right]=a^2-bc\end{cases}$$

4.11 Formula for inverse

4.11.1 Recall

For 2×2 matrix,we have the formula:

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

We can find that $\frac{1}{ad-bc}$ is the determination of the matrix, and $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ is the transposed matrix of the cofactor matrix of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

4.11.2 Formula for inverse

We let $C_{ij} = (-1)^{i+j}det(A_{ij})$ be the (i,j) cofactor

We can group C_{ij} into the cofactor matrix: $C = \begin{bmatrix} C_{11} & ... & C_{1n} \\ ... & ... & ... \\ C_{n1} & ... & C_{nn} \end{bmatrix}$

Example: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, C = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$

According to this, we can infer the formula:

$$AC^T = det(A) \times I_n$$

Proof:

1.$(AC^T)_{ii} = a_{i1}(C^T)_{1i} + a_{i2}(C^T)_{2i} + ... + a_{in}(C^T)_{ni} = a_{i1}c_{i1} + a_{i2}c_{i2} + ... + a_{in}c_{in}$ It is exactly the i-th row expansion of $det(A)$.

So

$$(AC^T)_{ii} = det(A)$$

2.$(AC^T)_{ij} = a_{i1}c_{j1} + a_{i2}c_{j2} + ... + a_{in}c_{jn}$

Say n=3,i=1,j=2

$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$

$(AC^T)_{12} = a_{11}c_{21} + a_{12}c_{22} + a_{13}c_{23} = a_{11}(-1)^{2+1}det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} + a_{12}(-1)^{2+2}det\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} + a_{13}(-1)^{2+3}det\begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} = det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = 0$

In general: $(AC^T)_{ij} = a_{i1}(-1)^{j+1}c_{j1} + a_{i2}(-1)^{j+2}c_{j2} + ... + a_{in}(-1)^{j+n}c_{jn} = det\begin{bmatrix} ... & ... & ... \\ a_{i1} & ... & a_{in} \\ ... & ... & ... \\ a_{i1} & ... & a_{in} \\ ... & ... & ... \end{bmatrix} = 0$, the repeated row are i-th row and j-th row

It is exactly an operation to insert $a_{i1}a_{i2}...a_{in}$ to the j-th row

Above all

$$AC^T = det(A) \times I_n$$

If $det(A) \neq 0$,

$$A^{-1} = \frac{1}{det(A)} C^T$$

4.11.3 Application

Now we have two methods to find inverse matrix(Gauss-Jordan elimination and Formula for inverse)

Example:$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $det(A) = 1$, find A^{-1}

We can initially get $C = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

So $A^{-1} = \frac{1}{det(A)} C^T = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

4.12 Cramer's rule

4.12.1 Solving equation

Ax=b with $n \times n$matrix A,assuming A is invertible: $C^T Ax = C^T b$

Then $x = \frac{1}{det(A)} C^T b$

Let write down the entries$(C^T b)_j$: $(C^T b)_j = (C^T)_{j1} b_1 + (C^T)_{j2} b_2 + ... (C^T)_{jn} b_n = b_1 C_{1j} + b_2 C_{2j} + ... + b_n C_{nj} = b_1(-1)^{1+j} det A_{1j} + ... + b_n(-1)^{n+j} det A_{nj} = det \begin{bmatrix} a_{11} & ... & a_{1(j-1)} & b_1 & ... & a_{1n} \\ ... & ... & ... & ... & ... & ... \\ a_{n1} & ... & a_{n(j-1)} & b_n & ... & a_{nn} \end{bmatrix}$, it is exactly replacing the j-th column of A by $\vec{b}$

So

$$x_j = \frac{det(B_j)}{det(A)}$$

4.12.2 Application

Example: $x_1 + 3x_2 = 0, 2x_1 + 4x_2 = 6$ or $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$

We can initially get $det(A) = -2$

So $x_1 = \frac{1}{det(A)} det \begin{bmatrix} 0 & 3 \\ 6 & 4 \end{bmatrix} = 9, x_2 = \frac{1}{det(A)} det \begin{bmatrix} 1 & 0 \\ 2 & 6 \end{bmatrix} = -3$

4.13 Formula for pivots

4.13.1 Formula for pivots

Let A= $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, say $a \neq 0$

$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{d}{a} & 1 & 0 \\ -\frac{g}{a} & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 0 & e - \frac{bd}{a} & f - \frac{cd}{a} \\ 0 & * & * \end{bmatrix}$

Let $L_1 = \begin{bmatrix} 1 & 0 \\ -\frac{d}{a} & 1 \end{bmatrix}$, $A_1 = \begin{bmatrix} a & b \\ d & e \end{bmatrix}$

$L_1 \cdot A_1 = \begin{bmatrix} a & b \\ 0 & e - \frac{bd}{a} \end{bmatrix}$, we can find that a is the first pivot and $e - \frac{bd}{a}$ is the second pivot, and $det(L_1) \cdot det(A_1) =$ product of the first two pivots $= ae - bd$

In general, say if invertible $A = LU$, let $\begin{bmatrix} A_k & * \\ * & * \end{bmatrix} = \begin{bmatrix} L_k & 0 \\ * & * \end{bmatrix} \cdot \begin{bmatrix} U_k & * \\ * & * \end{bmatrix}$

so $A_k = L_k \cdot U_k$, so $det(A_k) = det(U_k)$

U_k is upper triangular, $det(U_k) =$ product of first k pivots

so $det(A_k) = d_1 \cdot d_2 \cdot ... \cdot d_k$

Therefore.

$$d_k = \frac{det(A_k)}{det(A_{k-1})}$$

4.13.2 Property

Property: For invertible A , iff $det(A_1) \cdot det(A_2) \cdot ... \cdot det(A_n) \neq 0$, A can be done $A = LDU$ without row operation

Proof:

1.$n = 1$, nothing to prove

2.$n = 2$, let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $det(A_1) = a \neq 0$, $det(A_2) = ad - bc$

if $det(A_2) = ad - bc \neq 0$, then A can be factored directly to $\begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix}$

3.Induction step, if $A_{n=1} = L_{n-1} \cdot \tilde{U}_{n-1}$ be the decomposition

$\begin{bmatrix} L_{n-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_{n-1} & * \\ * & a_{nn} \end{bmatrix} = \begin{bmatrix} \tilde{U}_{n-1} & * \\ * & * \end{bmatrix}$, where D_{n-1} has full pivots on diagonal

4.Further type III row operation can turn $\begin{bmatrix} \tilde{U}_{n-1} & * \\ * & * \end{bmatrix}$ into $\tilde{U}_n = \begin{bmatrix} \tilde{U}_{n-1} & * \\ 0 & d_{nn} \end{bmatrix}$

Chapter 5

Eigenvalues and eigenvectors

5.1 Eigenvalues and eigenvectors

Can we find a basis of R^n such that the linear transform has "simplest" representation.

i.e. We want a basis $v_1 ... v_n$ s.t.$S^{-1}AS = D$
$S = [v_1 ... v_n]$
In the other word,we are interested in

$$Ax = tx$$

(x is non-zero)

To diagonalize a matrix,we need a basis of R^nconsisting of eigenvectors.

How to find eigenvalues?

ů If we do row operation to get to r.r.e.f. U,does it help?

$$EA = U, EAx = tEx, Ux = tEx$$

ů $Ux = tEx$ is not very useful unless $t = 0$

Method:

ů We determine the eigenvalue first:

$Ax = tx$, (A-tI)x=0,

N(A-tI) is not equal to 0, $det(A - tI) = 0$

ů $P_A(t) = det(A - tI)$ is called the characteristic polynomial of A

ů $P_A(t)$ is a degree n polynomial.

ů It is become

$$P_A(t) = det(A - tI)$$

ů It is exactly of degree n because there is one degree n term:$(a_{11} - t)...(a_{nn} - t)det(I)$

Prop:

$P_A(t) = (-t)^n + tr(A)(-t)^{n-1} + ... + det(A)$

Thm:

ů An eigenvalue t is a solution to $P_A(t) = 0$

ů An eigenvector corresponding to t is a non-zero element in N(A-tI).

Benefit:

$A^2 = (SDS^{-1})(SDS^{-1}) = SD^2S^{-1}$

$A^3 = (SD^2S^{-1})(SDS^{-1}) = SD^3S^{-1}$

$A^{2021} = SD^{2021}S^{-1}$

Def:

Given an eigenvalue t,we let $E_t = \{v|Av = tv\}$ be the set of all eigenvectors.

E_t is a subspace called the eigensubspace w.r.t. t.

ů There are various cases that we cannot find a basis of eigenvectors:

Case 1

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

$P_A(t) = t^2 + 1$ has no solution.

Case 2

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

$P_A(t) = (1 - t)^2$

ů Take eigensubspace E,with basis $v_1...v_k$.

Extend it to basis $v_1...v_k, w_{k+1}...w_n$ of R^n

Lemma:

ů If $E_1...E_k$ are eigensubspace corresponding to distinct eigenvalues $u_1...u_k$.

ů Suppose we have l.i. vectors in E_i,

then putting them together are still l.i..

Thm:

An $n * n$ matrix A is diagonalizable iff

1.$P_A(t) = 0$ has n roots counting alg. multiplicity.

2.For each eigenvalue u:
alg. muli.=geo. multi.

Cor:

In particular,if $P_A(t)$ has n distinct root ,then A must be diagonalizable.

e.g.:

A=$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$,is A diagonalizable?

ans:

A-tI=$\begin{bmatrix} 1-t & 2 & -1 \\ 0 & 1-t & 1 \\ 1 & 0 & 1-t \end{bmatrix}$,

$$P_A(t) = det \begin{bmatrix} 1-t & 2 & -1 \\ 0 & 1-t & 1 \\ 1 & 0 & 1-t \end{bmatrix}$$

=-det$\begin{bmatrix} 1 & 0 & 1-t \\ 0 & 1-t & 1 \\ 1-t & 2 & -1 \end{bmatrix}$

=-det$\begin{bmatrix} 1 & 0 & 1-t \\ 0 & 1-t & 1 \\ 0 & 2 & -1-(1-t)^2 \end{bmatrix}$

=-det$\begin{bmatrix} 1-t & 1 \\ 2 & -1-(1-t)^2 \end{bmatrix}$

=$(1-t)^3+1-t+2$

What can be done:

A=$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$; $t_1 = 2$.

t_1: A-2I=$\begin{bmatrix} -1 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$

N(A-2I)=span$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

ů Take a basis v_1, v_2, v_3

S=$[v_1, v_2, v_3]$

The first column of $S^{-1}AS$=$\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$

ů If the sum of geo. mult.=k,then $S^{-1}AS$ has size $k * k$ diagonal block at the top-left corner.

Q:

Solve the initial value problem:

$dv/dt = 4v - 5w; dw/dt = 2v - 3w$

v(0)=8;w(0)=5

A:

Writing Z(t)=$\begin{bmatrix} v(t) \\ w(t) \end{bmatrix}$,we have $dZ/dt = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$Z,$Z(0) = \begin{bmatrix} 8 \\ 5 \end{bmatrix}$

We may treat Z(t) as a curve in R^2.

Let A=$\begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$

We try to diagonalize it.

det(A-tI)=det$\begin{bmatrix} 4-t & -5 \\ 2 & -3-t \end{bmatrix}$

=$t^2 - t - 2$

=(t-2)(t+1)

So : eigenvalues:$t_1 = 2, t_2 = -1$

$t_1 = 2$: A-2I=$\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix}$

with v_1=$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$

t_2=-1: A+I=$\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix}$

with v_2=$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$Q^{-1}AQ = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

with Q=$\begin{bmatrix} 5 & 1 \\ 2 & 1 \end{bmatrix}$

ů Then we have $dz/dt = (QDQ^{-1})Z$
or dy/dt=DY,where Y(t)=$Q^{-1}Z(t)$
ů Because it is $dy_1/dt = 2y_1(t)$,$dy_2/dt = -y_2(t)$
$y_1(t) = c_1e^{2t}, y_2(t) = c_2e^{-t}$
ů Y(t)=$\begin{bmatrix} c_1e^{2t} \\ c_2e^{-t} \end{bmatrix}$

ů $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$=Y(0)
=$Q^{-1}Z(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$
ů Z(t)=QY(t)=$\begin{bmatrix} 5e^{2t} + 3e^{-t} \\ 2e^{2t} + 3e^{-t} \end{bmatrix}$

$v(t) = 5e^{2t} + 3e^{-t}$
$w(t) = 2e^{2t} + 3e^{-t}$

5.2 Diagonalization

Def:

ů If there is non-zero v s.t. L(v)=tv,

we call t an eigenvalue of L, v eigenvector w.r.t. t.

Def:

ů We say that L is diagonalizable if it has a basis consisting of eigenvectors.

Def:

ů We define the characteristic polynomial of L by taking an arbitrary basis $v_1 ... v_n$ of V, and

$$P_L(t) = P_{[L]}(t)$$

prop:

If AB=BA, given eigensubspace E_t for A, we have B(E_t) contained in E_t

Thm:

If both AB=BA,and both diagonalizable,

then there is Q invertible s.t. $Q^{-1}AQ = D_1, Q^{-1}BQ = D_2$

Difference equation

$F_{k+2} = F_{k+1} + F_k, F_0 = 0, F_1 = 1.$

How to write down formula for F_n?

Idea: $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$ be a vector.Then

$u_k = \begin{bmatrix} F_k + F_{k-1} \\ F_k \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$=A$u_{k-1}$

$u_k = A^k u_0 = S \begin{bmatrix} t_1^k & 0 \\ 0 & t_2^k \end{bmatrix} S^{-1}$

ů If we consider a general difference equation
$F_k = a_1 F_{k-1} + ... + a_l F_{k-l}$

We let $u_k = \begin{bmatrix} F_{k+3} \\ F_{k+2} \\ F_{k+1} \\ F_k \end{bmatrix}$

$= \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+2} \\ F_{k+1} \\ F_k \\ F_{k-1} \end{bmatrix}$

$=Au_{k-1} = A^k u_0$

ů If A is diagonalizable:$A = SDS^{-1}$,then $u_k = SD^k(S^{-1}u_0)$=$c_1 t_1^k v_1 + ... + c_l t_l^k v_l$

u_0 is given, $\begin{bmatrix} c_1 \\ \vdots \\ c_l \end{bmatrix}$ is determined by initial data.

Markov matrices

Thm:

1.1 is an eigenvalue of A

2.The absolute value of the eigenvalue of A is less than or equal to one.

Proof:

If A,B are markov matrix,then AB is again a markov matrix.

[1,1,1...]AB=I , and A , B has non-negative entries.

So A^2 is a markov matrix,

A^3 is a markov matrix,

......

A^k is a markov matrix

So A^k has entries a_{ij} which is greater than or equal to 0 and less than or equal to 1.

But if there is an eigenvalue t_i with $|t_i|$>1,

then $A^k v_i = t_i^k v_i$.

$t_i^k v_i$ goes to infinity when k goes to positive infinity.

This is contradictory.

So the absolute value of the eigenvalue of A is less than or equal to one.

3.The sum of each column of A is one.

4.The elements of matrix A are non-zero.

Thm:

ů Let A=(a_{ij}) with a_{ij} is non-zero.Assume for some l s.t. A^l has all positive entries.

ů Then there is a largest eigenvalue t_1 greater than or equal to $|t_i|$,its eigenvector has positive coefficients.

Def:

A difference equation $U_{k+1} = AU_k$, with A diagonalizable

1.stable if $|t_i| < 1$ for all t_i.

2.neutrally stable if $|t_i| < 1$ or =1 and some $|t_i|$=1

3.unstable if $|t_i| > 1$ for some t_i.

Proof:

ů If A is stable: A=SDS^{-1},

with D^k going to 0 as k going to infinity.

Therefore A^k goes to 0 as k going to infinity.

i.e. $u_k = A^k u_0$ goes to 0 as k going to infinity.

ů If A is neutrally stable:

u_k=$A^k u$=$[v_1 ... v_l, ... v_n] D^k \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$

$= c_1(t_1)^k v_1 + ... + c_l(t_l)^k v_l + ...$

If t_1 =...=t_l =1, then u_k goes to $c_1 v_1 + ... + c_l v_l$

If A is unstable,limit of u_k may not always exists

(depending on initial condition u_k)

Complex numbers

Def:

A complex number is a number of the form $z = a + ib$ with the operations:

1.(a+ib)+(c+id)=(a+c)+i(b+d)

2.(a+ib)(c+id)=(ac-bd)+i(ad+bc)

The set of complex number is denoted by C.

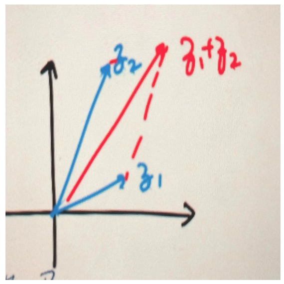
Given z=a+ib

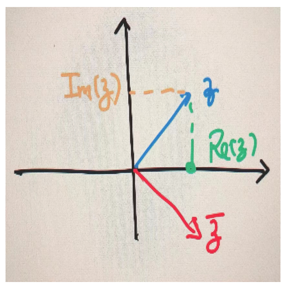
1.its complex conjugate = a-ib

2.its modulus= $\sqrt{a^2+b^2}$

3.its real part=a

4.its imaginary part Im(z)=b





Prop:

1.$(z_1+z_2)+z_3=z_1+(z_2+z_3)$

2.$z_1+z_2=z_2+z_1$

3.$z_1(z_2z_3)=(z_1z_2)z_3$

4.$z_1(z_2+z_3)=z_1z_2+z_1z_3$

5.There is 1=1+i0 such that 1*z=z for any z.

6.If z is non-zero,then $z^{-1}z=1$.

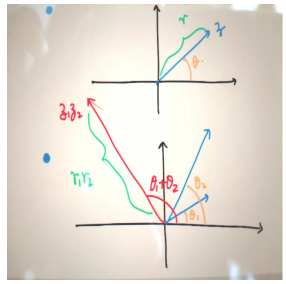
ů Given z is non-zero,we can write $|z|=r$,and $z/|z|=cos\alpha+isin\alpha$ for some α: $z=r(cos\alpha+isin\alpha)=re^{i\alpha}$

is called the polar form of z.

$z_1=r_1(cos\alpha_1+isin\alpha_1)=re^{i\alpha_1}$

$z_2=r_1(cos\alpha_2+isin\alpha_2)=re^{i\alpha_2}$

$z_1 z_2 = r_1 r_2 e^{i(\alpha_1+\alpha_2)}$



Warning:
$z^2 > 0$ is no longer true,indeed it does not even make sense.

Thm:

Any degree n polynomial $p(x)$ has exactly n roots counting multiplicity.

ů That means we don't need to worry about $P_A(t) = 0$ has not enough solutions.

Def:
A complex vector space is a set V together with two operation $+, *$, satisfying (1)—(8) as for real vector space.

ů A $m * n$ matrix A= $(a_{ij})_{mn}$ is the one with a_{ij} belonging to C.

ů We write $M_{m,n}(C)$ to stand for the set of complex matrices.

Rk:
ů Theory in Chapter 1,2,4,5.1—5.2 still holds true by replacing the role of R by C.

ů The properties for inner product:

1.$< x, cz >= c < x, z >$

2.$< x, y + z >=< x, y > + < x, z >$

3.$< x, x >$ is a non-negative real number with $< x, x >$=0 only if $x = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

Orthogonality:

ů x,y are said to be orthogonal if $< x, y >= 0$.

ů w_1 is perpendicular to w_2 if $< x, y >= 0$ for x belonging to w_1,y belonging to w_2.

ů $x_1 ... x_n$ orthogonal if $< x_i, x_j >= 0$ if i is not equal to j. $x_1 ... x_n$ orthogonal if $< x_i, x_j >= 1$ if i equal to j.

Formula:

ů Take an o.n. basis $q_1 ... q_k$ of v.

P(v)=$< q_1, v > q_1 + < q_2, v > q_2 + ... + < q_k, v > q_k$.

=$(q_1 q_1^H + ... + q_k q_k^H)$v

5.3 Introduction

In the previous chapters, the complex numbers could have been permitted, but would contributed to nothing new. Now we cannot avoid them. We now introduce the space $\mathbb{C}^n$ of vectors with n complex components where the length is computed differently. For example, the length squared of the vector in $\mathbb{C}^2$ with components $(1, i)$ is $1^2 + |i|^2 = 2$.

Before we get started, we need some preparations.

Proposition 5.1. $(AB)^H = B^H A^H$

Proof. We have $(AB)^T = B^T A^T$ and its conjugate gives $(AB)^H = B^H A^H$.

$$(AB)^H = \overline{(AB)}^T = \overline{(AB)^T} = \overline{(B^T A^T)} = \overline{B^T}\ \overline{A^T} = \overline{B}^T\ \overline{A}^T = B^H A^H$$

□

Definition 5.1.

- *A matrix A is said to be* Hermitian *if $A^H = A$.*
- *A matrix A is said to be* skew-Hermitian *if $A^H = -A$.*
- *A matrix A is said to be* unitary *if $A^H A = A A^H = I$.*

Remark 1. *The new definitions coincide with the old when the vectors and matrices are real. Thus we may view Hermitian matrices as the complex analogue of real symmetric matrices. Similarly, a real skew-Hermitian matrix is a skew-symmetric matrix. A real unitary matrix is an orthogonal matrix.*

5.4 Complex Matrices

In this section, we study matrices with complex entries and look at the complex analogues of symmetric and orthogonal matrices.

5.4.1 Unitary Matrices

Proposition 5.2. *Multiplication by a unitary matrix U has no effect on inner products, angles, or lengths:*

$$\|Ux\|^2 = x^H U^H U x = \|x\|^2 .$$

Proposition 5.3. *If U is unitary, then every eigenvalue of U has absolute value $|\lambda| = 1$. That is, $\lambda = \cos\theta + i\sin\theta$. If Q is a real orthogonal matrix, then there also exists $|\lambda| = 1$, which gives $\lambda = \pm 1$.*

Example 5.1. $U = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$ *has eigenvalues e^{it} and e^{-it}.*

The orthogonal eigenvectors are $x = (1, -i)$ and $y = (1, i)$. (Remember to take conjugates in $x^H y = 1 + i^2 = 0$.) After division by $\sqrt{2}$ they are orthonormal.

Proposition 5.4. *An $n \times n$ matrix U is unitary if and only if its column vectors form an orthonormal set in $\mathbb{C}^n$.*

Example 5.2. $U = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \cdot & 1 \\ 1 & w & \cdot & w^{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & w^{n-1} & \cdot & w^{(n-1)^2} \end{bmatrix} = \frac{\textit{Fourier matrix}}{\sqrt{n}}.$

The complex number w is on the unit circle at the angle $\theta = 2\pi/n$. It equals $e^{2\pi i/n} = \cos(2\pi/n) + i\sin(2\pi/n)$. Its powers are spaced evenly around the unit circle. That spacing assures that the sum of all n powers of w-all the nth roots of 1-is zero. Algebraically, the sum$1 + w + \cdots + w^{n-1}$ is $(w^n - 1)/(w - 1)$. And $(w^n - 1)/(w - 1)$ is zero.

$$\begin{aligned} \textit{row } i \textit{ of } U^H \textit{ times column } j \textit{ of } U \textit{ is } & \frac{1}{n}\left(1 + W + W^2 + \cdots + W^{n-1}\right) \\ &= \frac{W^{n-1} - 1}{W - 1} = 0 \end{aligned}$$

In the above case, $W = w^{j-i}$. Every entry of the original F has absolute value 1. The factor $\sqrt{n}$ shrinks the columns of U into unit vectors. The fundamental identity of the finite Fourier transform is $U^H U = I$. Thus U is unitary.

5.4.2 Hermitian Matrices

Hermitian matrices have many nice properties, as we shall see in the next proposition.

Proposition 5.5. *If $A = A^H$, then for all complex vectors x, the number $x^H A x$ is real.*

Proposition 5.6. *If $A = A^H$, every eigenvalue is real.*

Remark 2. *If $A = -A^H$ is skew-Hermitian, then the number $x^H A x$ is purely imaginary and all λ_i are purely imaginary.*

5.5 Similarity Transformations

5.5.1 Change of Basis = Similarity Transformation

Definition 5.2. *Let A and B be $n \times n$ matrices. B is said to be similar to A if there exists a nonsingular matrix S such that $B = S^{-1}AS$.*

Proposition 5.7. *Let A and B be $n \times n$ matrices. If B is similar to A, then the two matrices have the same characteristic polynomial and, consequently, the same eigenvalues.*

The similar matrix $B = M^{-1}AM$ is closely connected to A, if we go back to linear transformations. Remember the key idea: Every linear transformation is represented by a matrix. The matrix depends on the choice of basis! If we change the basis by M we change the matrix A to a similar matrix B.

Theorem 5.1. *Let $E = \{v_1, \ldots, v_n\}$ and $F = \{w_1, \ldots, w_n\}$ be two ordered bases for a vector space V, and let L be a linear operator on V. Let S be the transition matrix representing the change from F to E. If A is the matrix representing L with respect to E, and B is the matrix representing L with respect to F, then $B = S^{-1}AS$.*

$$
\begin{array}{ccccc}
{}_F[L]_F & = & {}_F[I]_E & {}_E[L]_E & {}_E[I]_F \\
B & = & S^{-1} & A & S
\end{array}
$$

5.5.2 The Jordan Form

Theorem 5.2. *If A has s independent eigenvectors, it is similar to a matrix with s blocks:*

$$
\textbf{\textit{Jordan form}} \qquad J = M^{-1}AM = \begin{bmatrix} J_1 & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & J_s \end{bmatrix}.
$$

Each Jordan block J_i is a triangular matrix that has only a single eigenvalue λ_i and only one eigenvector:

$$
J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}.
$$

The same λ_i will appear in several blocks, if it has several independent eigenvectors. Two matrices are similar if and only if they share the same Jordan form J.

Example 5.3. $T = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ *and* $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ *all lead to* $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

These four matrices have eigenvalues 1 *and* 1 *with only one eigenvector-so* J *consists of one block.*

$$M^{-1}TM = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

$$P^{-1}BP = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

$$U^{-1}AU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = T$$

$$\textit{and then } M^{-1}TM = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = J.$$

Example 5.4. $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Zero is a triple eigenvalue for A *and* B*, so it will appear in all their Jordan blocks. Then* A *and* B *have three possible Jordan forms:*

$$J_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The only eigenvector of A *is* $(1, 0, 0)$*. Its Jordan form has only one block, and* A *must be similar to* J_1*. The matrix* B *has the additional eigenvector* $(0, 1, 0)$*, and its Jordan form is* J_2 *with two blocks. As for* $J_3 =$ *zero matrix, it is in a family by itself; the only matrix similar to* J_3 *is* $M^{-1}0M = 0$*. A count of the eigenvectors will determine* J *when there is nothing more complicated than a triple eigenvalue.*

Example 5.5. Applications. *In every case we have Jordans similarity:* $A = MJM^{-1}$*, so now we need the powers of* J*:*

$$A^k = MJ^kM^{-1}.$$

J *is block diagonal, and the powers of each block can be taken separately:*

$$(J_i)^k = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \frac{1}{2}k(k-1)\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{bmatrix}.$$

This block J_i will enter when λ is a triple eigenvalue with a single eigenvector. Its exponential is in the solution to the corresponding differential equation:

$$e^{J_i t} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{1}{2}t^2 e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} \\ 0 & 0 & e^{\lambda t} \end{bmatrix}.$$

Here $I + J_i t + (J_i t)^2/2! + \cdots$ produces $1 + \lambda t + \lambda^2 t^2/2! + \cdots = e^{\lambda t}$ on the diagonal.

5.5.3 Diagonalizing Symmetric and Hermitian Matrices

Lemma. *There is a unitary matrix $M = U$ such that $U^{-1}AU = T$ is triangular. The eigenvalues of A appear along the diagonal of this similar matrix T.*

Proof. Every matrix, say 4 by 4, has at least one eigenvalue λ_1. In the worst case, it could be repeated four times. Therefore A has at least one unit eigenvector x_1, which we place in the first column of U. At this stage the other three columns are impossible to determine, so we complete the matrix in any way that leaves it unitary, and call it U_1. (The Gram-Schmidt process guarantees that this can be done.) $Ax_1 = \lambda_1 x_1$ column 1 means that the product $U_1^{-1}AU_1$ starts in the right form:

$$AU_1 = U_1 \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \text{ leads to } U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}.$$

Now work with the 3 by 3 submatrix in the lower right-hand corner. It has a unit eigenvector x_2, which becomes the first column of a unitary matrix M_2:

$$\text{If } U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix} \text{ then } U_2^{-1}(U_1^{-1}AU_1)U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

At the last step, an eigenvector of the 2 by 2 matrix in the lower right-hand corner goes into a unitary M_3, which is put into the corner of U_3:

$$U_3^{-1}\left(U_2^{-1}U_1^{-1}AU_1U_2\right)U_3 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & * \end{bmatrix} = T.$$

The product $U = U_1U_2U_3$ is still a unitary matrix, and $U^{-1}AU = T$. $\square$

This lemma applies to all matrices, with no assumption that A is diagonalizable.

This triangular form will show that any symmetric or Hermitian matrix-whether its eigenvalues are distinct or not-has a complete set of orthonormal eigenvectors. We need a unitary matrix such that $U^{-1}AU$ is diagonal. Schurs lemma has just found it. This triangular T must be diagonal, because it is also Hermitian when $A = A^H$:

$$T^H = (U^{-1}AU)^H = U^H A (U^{-1})^H = U^{-1}AU = T.$$

Theorem 5.3. Spectral Theorem. *Every real symmetric A can be diagonalized by an orthogonal matrix Q. Every Hermitian matrix can be diagonalized by a unitary U:*

$$\begin{array}{rlcl} \textit{(real)} & Q^{-1}AQ = \Lambda & \textit{or} & A = Q\Lambda Q^T \\ \textit{(complex)} & U^{-1}AU = \Lambda & \textit{or} & A = U\Lambda U^H \end{array}$$

The columns of Q (or U) contain orthonormal eigenvectors of A.

Example 5.6. *The spectral theorem says that this $A = A^T$ can be diagonalized:*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ with repeated eigenvalues } \lambda_1 = \lambda_2 = 1 \text{ and } \lambda_3 = -1.$$

$\lambda = 1$ has a plane of eigenvectors, and we pick an orthonormal pair x_1 and x_2:

$$x_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \text{ and } x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ and } x_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \text{ for } \lambda_3 = -1.$$

These are the columns of Q. Splitting $A = Q\Lambda Q^T$ into 3 columns times 3 rows gives

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \lambda_1 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $\lambda_1 = \lambda_2$, those first two projections $x_1x_1^T$ and $x_2x_2^T$ (each of rank 1) combine to give a projection P_1 of rank 2 (onto the plane of eigenvectors). then A is

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 P_1 + \lambda_3 P_3 = (+1) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1) \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Every Hermitian matrix with k different eigenvalues has a spectral decomposition into $A = \lambda_1 P_1 + \cdots + \lambda_k P_k$, where P_i is the projection onto the eigenspace for λ_i. Since there is a full set of eigenvectors, the projections add up to the identity. And since the eigenspace are orthogonal, two projections produce zero: $P_j P_i = 0$.

5.5.4 Normal matrices

We are very close to answering an important question, so we keep going: For which matrices is $T = \Lambda$? Symmetric, skew-symmetric, and orthogonal T's are all diagonal! Hermitian, skew-Hermitian, and unitary matrices are also in this class. They correspond to numbers on the real axis, the imaginary axis, and the unit circle. Now we want the whole class, corresponding to all complex numbers. The matrices are called "normal".

Definition 5.3. *The matrix N is normal if it commutes with N^H:*

$$NN^H = N^H N.$$

Normal matrices are exactly those that have a complete set of orthonormal eigenvectors.

Proposition 5.8. *If A is normal and $Ax = \lambda x$, then $A^H x = \overline{\lambda} x$.*

Proof.

$$\begin{aligned}
(A - \lambda I)\, x &= 0 \\
\|(A - \lambda I)\, x\|^2 &= 0 \\
((A - \lambda I)\, x)^H\, ((A - \lambda I)\, x) &= 0 \\
x^H\, (A - \lambda I)^H\, (A - \lambda I)\, x &= 0 \\
x^H\, (A - \lambda I)\, (A - \lambda I)^H\, x &= 0 \\
\left\|(A - \lambda I)^H\, x\right\|^2 &= 0 \\
\left\|\left(A^H - \overline{\lambda} I\right) x\right\|^2 &= 0 \\
\left(A^H - \overline{\lambda} I\right) x &= 0
\end{aligned}$$

□

Proposition 5.9. *If A is normal, $Ax = \lambda x$, $Ay = \mu y$, and $\lambda \neq \mu$, then x is orthogonal to y.*

Proof. The proof starts with $A^H y = \overline{\mu} y$:

$$(\lambda x)^H\, y = (Ax)\, y = x^H A^H y = x^H\, (\overline{\mu} y)\, .$$

The outside numbers are $\overline{\lambda} x^H y = \overline{\mu} x^H y$. Now we use the assumption that $\lambda \neq \mu$, which forces the conclusion that $x^H y = 0$. □

Theorem 5.4. *For such normal matrices, and no others, the triangular $T = U^{-1} N U$ is the diagonal Λ.*

Proof.
1. If N is normal, then so is the triangular $T = U^{-1}NU$:

$$TT^H = U^{-1}NUU^HN^HU = U^{-1}NN^HU,$$
$$T^HT = U^HN^HUU^{-1}NU = U^{-1}N^HNU.$$

2. Also a triangular T that is normal must be diagonal. (Compare the entries of TT^H and T^HT, and show that if they are equal, then T must be diagonal. □

Similarity Transformations

1. *A is diagonalizable: The columns of S are eigenvectors and $S^{-1}AS = \Lambda$.*
2. *A is arbitrary: The columns of M include "generalized eigenvectors" of A, and the Jordan form $M^{-1}AM = J$ is block diagonal.*
3. *A is arbitrary: The unitary U can be chosen so that $U^{-1}AU = T$ is triangular.*
4. *A is normal, $AA^H = A^HA$: then U can be chosen so that $U^{-1}AU = \Lambda$.*

   Special cases of normal matrices, all with orthonormal eigenvectors:

   (a) *If $A = A^H$ is Hermitian, then all λ_i are real.*
   (b) *If $A = A^T$ is real symmetric, then Λ is real and $U = Q$ is orthogonal.*
   (c) *If $A = -A^H$ is skew-Hermitian, then all λ_i are purely imaginary.*
   (d) *If A is orthogonal or unitary, then all $|\lambda_i| = 1$ are on the unit circle.*

article graphicx derivative amsmath physics amsthm ragged2e blindtext enumitem stepsenumerate1 [steps, 1]label = Step 0:
Chapter 6
January 2024

Introduction

This section is connected to the signs of the eigenvalues, and the signs are crucial. With the signs, we can figure out the minima or maxima of the function, thus solve the problem of optimization, which is of practical use in engineering and science.

6.1 Minima,Maxima,and Saddle Points

1. ***Basic forms***

   The first example in the textbook includes some partial differential equations, which may seem a bit unfamiliar to those who haven't learned it in *advance. But once you grasp the method of calculating them, you'll find that the form:*

$$\begin{bmatrix} ^2fx^2(0,0) & fxy(0,0) \\ fxy(0,0) & ^2fy^2(0,0) \end{bmatrix}$$

is actually the same as the quadratic form:

$$f(x,y) = ax^2 + 2bxy + cy^2$$

which is associated to

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

but note that the first one is 2 times of the second one.

2. *Figures of the quadratic form*

   After knowing this, we can then start to define the shape of the function.

   - ***A bowl:*** *(Also know as maximum for $a < 0$ or minimum for $a > 0$)*

     $$b^2 < ac$$

   - ***A saddle:***

     $$b^2 > ac$$

   - ***A valley:*** *(Also known as the singular case or degenerated case, which includes positive semidefinite and negative semidefinite)*

     $$b^2 = ac$$

   Proof:

   $$f(x,y) = ax^2 + 2bxy + cy^2 = a\left(x^2 + \frac{2b}{a}xy + \frac{b^2}{a^2}\right) + \left(c - \frac{b^2}{a}\right)y^2$$

3. ***Turning quadratic form into matrix***

Example

Example 5.1.

$$f(x,y) = x^2 - 10xy + y^2 = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} 1 & -5 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

6.2 Tests for Positive Definiteness

1. ***Positive Definiteness***

   they are listed on the textbook very clearly at 6B, and the proofs are evident, there is also a fifth one which is

   symmetric matrix A is positive definite if and only if there is

   (V) a matrix R with independent columns such that $A = R^T R$ the form can be written in many ways:

   $A = (L\sqrt{D})(L\sqrt{D})^T$

   $A = (Q\sqrt{\Lambda})(Q\sqrt{\Lambda})^T$

   $A = (Q\sqrt{\Lambda}Q^T)(Q\sqrt{\Lambda}Q^T)^T$

   it depends on how you want to split the A up

   now that we have learned the fifth test which could help us write the matrix A in another form, there is another way for us to complete the square, which is simply eliminating the matrix:

   Example

   Example 5.2.

   try to complete $A = 2x^2 + 2y^2 + 2z^2 - 2xy - 2yz$ into a square

   Solution1:*the algebraic way*

   $$A = 2x^2 + 2y^2 + 2z^2 - 2xy - 2yz$$
   $$= 2\left(x - \tfrac{1}{2}y\right)^2 + \tfrac{3}{2}y^2 + 2z^2$$
   $$= 2\left(x - \tfrac{1}{2}y\right)^2 + \tfrac{3}{2}\left(y - \tfrac{2}{3}\right)^2 + 2z^2$$

   we complete one square at a time,there are also many other ways ,you can search them up on some Chinese websites for the postgraduate entrance exams, but the second way is much easier. ***Solution 2:*** *the matrix way*

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

   so we get $x^T A x = x^T L D L^T x$

   and when we do matrix multiplication for

   $x' = L^T x = \begin{bmatrix} x - \frac{1}{2}y \\ y - \frac{2}{3}z \\ z \end{bmatrix}$ *we find that the new entries in x' correlate to the completed square, and the coefficients correlate to the entries in the diagonal matrix D*

$$A = 2x^2 + 2y^2 + 2z^2 - 2xy - 2yz$$
$$= x^T A x$$
$$= x'^T D x'$$
$$= 2\left(x - \tfrac{1}{2}y\right)^2 + \tfrac{3}{2}\left(y - \tfrac{2}{3}\right)^2 + 2z^2$$

2. ***Positive Semidefiniteness***

6D Each of the following tests is a necessary and sufficient condition for a symmetric matrix A to be ***positive semidefinite***:

(I′) $x^{\mathrm{T}}Ax \geq 0$ for all vectors x (this defines positive semidefinite).

(II′) All the eigenvalues of A satisfy $\lambda_i \geq 0$.

(III′) No principal submatrices have negative determinants.

(IV′) No pivots are negative.

(V′) There is a matrix R, possibly with dependent columns, such that $A = R^{\mathrm{T}}R$.

note that a principal submatrix is a square submatrix obtained by removing certain rows and columns

3. ***Negative Definiteness***

II. Tests for Negative Definiteness

定理　设A为n阶实对称矩阵，则下列命题等价：

(1) $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\,\boldsymbol{x}$ 是负定二次型（或$\boldsymbol{A}$是负定矩阵）；

(2) $\boldsymbol{A}$ 的 n个特征值$\lambda_1,\lambda_2,\cdots,\lambda_n$都小于零;

(3) $\boldsymbol{A}$ 的 n个主元 $d_1, d_2,\cdots, d_n$都小于零(无行交换);

(4) $\boldsymbol{A}$ 的奇数阶顺序主子式都小于零, 偶数阶顺序主子式都大于零;

(5) 存在可逆矩阵$\boldsymbol{R}$, 使得$\boldsymbol{A}= -\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$.

it is quite similar to positive definite, and the forth test is to ensure that upper left submatrices A_k has negative determinants

4. ***Law of inertia***

We have already elementary operations and their invariants on matrices.For instance, the row space,the nullspace,the rank,the determinant wont change when a multiple of one row is subtracted from another row. And as for eigenvalues, the similar transformation is the basic operation, which *leaves the eigenvalues unchanged. So what about the elementary operation and its invariant for $x^T Ax$?*

Definition (1)**.**

We say two symmetric matrices A and B are congruent if $C^T AC = B$ for some invertible C.

Theorem (1)**.**

Suppose A and B are two congruent symmetric matrices,then they have the same number of positive, negative and zero eigenvalues. Remark: the same number doesn't mean the same eigenvalue

Application (1)**.**

Finding out the right quadratic surface through elimination (in the exam paper)

First let's have a look at the six basic quadratic surfaces:

Ellipsoids: $a^2x^2 + b^2y^2 + c^2z^2 = 1$(a,b,c>0) (tips:no negative sign)

Hyperboloid of one sheet: $a^2x^2 + b^2y^2 - c^2z^2 = 1$ (a,b,c>0)

(tips:one negative sign)

Hyperboloid of two sheets: $-a^2x^2 - b^2y^2 + c^2z^2 = 1$(a,b,c>0)

(tips:two negative signs)

Elliptic paraboloid:$px^2 + qy^2 = z$ (p,q the same sign)

Hyperbolic paraboloid:$px^2 - qy^2 = z$ (p,q the same sign)

Cones:$a^2x^2 + b^2y^2 - c^2z^2 = 0$ (a,b,c>0)

So when we meet a quadratic form, we can turn it into its matrix form, eliminate it to its diagonal form, and finally find its quadratic surface easily.

Example 5.3.

$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + +x_3^2 + 4x_1x_2 + 4x_1x_3 + 4x_2x_3$ determine the type of the quadratic surface $f(x_1, x_2, x_3) = 2$

Let

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

and

$$f(\vec{x}) = \vec{x}^T A\vec{x}$$

then,we eliminate A into

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

we get 1 positive pivot, 2 negative pivots, so it is a Hyperboloid of two sheets

What if the diagonals are all zero?

Here's a trick: we multiply the matrix A with $\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *on both sides, which will give A non-zero diagonal entries,then we continue the two-sided elimination process.*

Example (1)**.**

$f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$ *determine the type of the quadratic surface* $f(x_1, x_2, x_3) = 1$

First we let

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

Then we use the trick of getting diagonal entries,

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

after that we use the two-sided elimination and get the standard form

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

it is a Hyperboloid of one sheet

6.3 Singular Value Decomposition

Definition:

The singular value decomposition of a matrix A is the factorization of A into the product of three matrices $A = U\Sigma V^T$ *where the columns of U and V are orthonormal and the matrix D is diagonal with positive real entries.*

Here is a direct image of what SVD should be like:

$$\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m} \left[\begin{array}{cccc:c} \sigma_1 & & & & \\ & \sigma_2 & & & \mathbf{0}_{r\times(n-r)} \\ & & \ddots & & \\ & & & \sigma_r & \\ \hdashline & \mathbf{0}_{(m-r)\times r} & & & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right]_{m\times n} (\boldsymbol{V}_{n\times n})^T$$

(orthogonal) (rectangular diagonal) (orthogonal)

where $\sigma_1, \sigma_2, \cdots, \sigma_r$ are the square roots of the nonzero eigenvalues of both $\boldsymbol{AA}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

- $\boldsymbol{U}$: the columns are orthonormal eigenvectors for $\boldsymbol{AA}^{\mathrm{T}}$.
- $\boldsymbol{V}$: the columns are orthonormal eigenvectors for $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$.

Remark

Remark 3. *For complex matrix, Σ remains real, but U and V becomes unitary*

Remark 4. *U and V give orthonormal bases for all four fundamental subspaces:*
first r columns of U: column space of A
last m-r columns of U: left nullspace of A
first r columns of U: row space of A
last n-r columns of U: nullspace of A

Proof. For U:

$$AA^T = U\Sigma\Sigma^T U^T$$

since $\Sigma\Sigma^T$ is an m by m matrix with $\sigma_1^2\ddot{\sigma}_r^2$ on its diagonal, also the eigenvalues are in descending order, then the first r entries are real numbers, and the the rest,which is m-r are zero.
Relating to $A = Q^T\Lambda Q$, we get:
first r columns of U: column space of A
last m-r columns of U: left nullspace of A
and the same goes for matrix V

□

Remark 5. *The SVD chooses those basis in a extremely special way, which is more than just orthonormal. This can be shown in the way we factorize the SVD as follows:*

How to do the SVD?

Using A^tA to get the the square of Σ as well as the eigenvectors of V

We use the formula $\frac{Av_j}{\sigma_j} = u_j$

Example 3 (SVD) Let $\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Then $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$.

The eigenvalues of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ are 3 and 1, which are also the non-zero eigenvalues of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. So $\sigma_1 = \sqrt{3}$, $\sigma_2 = 1$.

Finding three eigenvectors of $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ gives rise to

$$\boldsymbol{V} = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \boldsymbol{v}_3].$$

Two unit eigenvectors of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ are

$$\boldsymbol{u}_1 = \frac{\boldsymbol{A}\boldsymbol{v}_1}{\sigma_1} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \boldsymbol{u}_2 = \frac{\boldsymbol{A}\boldsymbol{v}_2}{\sigma_2} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Thus $\boldsymbol{U} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$.

Now $\boldsymbol{\Sigma}$ is a (2×3)-matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, it is easy to get that

$$\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

notice that you should not change the sequence of the steps, this is because of the special way SVD is formed, otherwise, the signs may be wrong.

Application of the SVD

1. *Image processing : Throwing away significant or extremely small eigenvalues.*
2. *The effective rank: Using A^TA or AA^T to get the squared singular values, the eigenvalues wouldn't be misleading. And then we decide a tolerance like 10^{-6}, and then count the singular values above it–that is the effective rank.*
3. *Polar decomposition*

$$A = U\Sigma V^T = (UV^T)(V\Sigma V^T)$$
$$A = QS$$

or the Reverse polar decomposition:

$$A = S'Q$$

So how can we use the polar decomposition to solve problems? Here is an example:

Example 5.4. *Use the QS decomposition to factorize the following matrix* $\begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix}$

Solution:

The normal way of getting the QS decomposition is to first solve the SVD to get the U Σ *and* V;

However, if we try in this way, we'll find that $\lambda = \frac{15+5\sqrt{5}}{2}$ *or* $\frac{15-5\sqrt{5}}{2}$ *and the rest of the computation will be quite disturbing; but there is another way, which is to factorize by QS itself:*

$$A^T A = S^T Q^T Q S = S^T S$$

Let $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *then*

$$A^T A = S^T S = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = \begin{bmatrix} 10 & -5 \\ -5 & 5 \end{bmatrix}$$

which gives us

$$\begin{cases} a^2 + b^2 = 10 \\ ac + bd = -5 \\ c^2 + d^2 = 5 \end{cases}$$

it is obvious that there are four sets of answers:

$$\begin{cases} a = 3 \\ b = 1 \\ c = -2 \\ d = 1 \end{cases}$$

$$\begin{cases} a = 3 \\ b = -1 \\ c = -2 \\ d = -1 \end{cases}$$

$$\begin{cases} a = -3 \\ b = 1 \\ c = 2 \\ d = 1 \end{cases}$$

$$\begin{cases} a = -3 \\ b = -1 \\ c = 2 \\ d = -1 \end{cases}$$

and then we get $Q = AS^{-1}$

and so the factorization is done