Tutorial note for MATH 2111

Pan Liang

1. System of linear equation

1.1. Some concepts

- Linear equation: $a_1x_1 + a_2x_2 + ... + a_nx_n = b$
- System of linear equation: a collection of one or more linear equations, involving same variables $x_1, x_2, ...x_n$.
- Solution of system of linear equation: a collection of $x_1, x_2, ...x_n$ satisfies all of the equations.
- Three possibilities of solution:
  1. no solution;$\rightarrow$ the system is inconsistent.
  2. exactly one solution; $\rightarrow$ the system is consistent.
  3. infinitely many solutions; $\rightarrow$ the system is consistent.

1.2. Solving a linear system

Gaussian Elimination

Example 1.1.

$$2x_1 + 4x_2 = -4, \tag{1.1}$$
$$5x_1 + 7x_2 = 11. \tag{1.2}$$

Sol: Eq.1.2-$\frac{5}{2}$ Eq.1.1 $\Rightarrow -3x_2 = 21 \Rightarrow x_2 = -7$. (Elimination $5x_1$ in Eq.1.2).
Substituting $x_2 = -7$ in Eq.1.1, we have $x_1 = 12$.
Thus, we have the solution $x_1 = 12, x_2 = -7$.

1.3. Matrix form

Example 1.1 can be rewritten into the following matrix form

$$\begin{pmatrix} 2 & 4 \\ 5 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 11 \end{pmatrix},$$

where $\begin{pmatrix} 2 & 4 \\ 5 & 7 \end{pmatrix}$ is called the coefficient matrix, and $\begin{pmatrix} 2 & 4 & -4 \\ 5 & 7 & 11 \end{pmatrix}$ is called the augmented matrix.

The solution procedure can be written into the matrix notation

$$\begin{pmatrix} 2 & 4 & -4 \\ 5 & 7 & 11 \end{pmatrix} \xrightarrow{r_2 \to r_2 - 5/2r_1} \begin{pmatrix} 2 & 4 & -4 \\ 0 & -3 & 21 \end{pmatrix} \xrightarrow{r_2 \to -1/3r_2} \begin{pmatrix} 2 & 4 & -4 \\ 0 & 1 & -7 \end{pmatrix}$$
$$\xrightarrow{r_1 \to r_1 - 4r_2} \begin{pmatrix} 2 & 0 & 24 \\ 0 & 1 & -7 \end{pmatrix} \xrightarrow{r_1 \to 1/2r_1} \begin{pmatrix} 1 & 0 & 12 \\ 0 & 1 & -7 \end{pmatrix}$$

$\Rightarrow$ The solution of the linear system is

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 12 \\ -7 \end{pmatrix}.$$

1.4. Elementary row operation

There are three kinds of Elementary row operations (ERO)

1. replacement $\to$ replace one row by sum of itself and a multiple of another row;
2. interchange $\to$ interchange two rows;
3. scaling $\to$ multiply all the entries in a row by some non-zero number.

Definition 1.1. *Two matrices are called row equivalent if they differ by a sequence of EROs.*

Theorem 1.1. *Linear systems corresponding to row equivalent augmented matrices will hava the same solution set.*

Example 1.2. *consider the following linear system*

$$\begin{aligned} x_2 + 4x_3 &= -5, \\ x_1 + 3x_2 + 5x_3 &= -2, \\ 3x_1 + 7x_2 + 7x_3 &= 6. \end{aligned}$$

The matrix form of the this linear system

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

Sol: *The solution procedure can be written into the matrix notation*

$$\begin{pmatrix} 0 & 1 & 4 & -5 \\ 1 & 3 & 5 & -2 \\ 3 & 7 & 7 & 6 \end{pmatrix} \xrightarrow[\textit{interchange}]{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & 1 & 4 & -5 \\ 3 & 7 & 7 & 6 \end{pmatrix} \xrightarrow[\textit{replacement}]{r_3 \to r_3 - 3r_1} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & 1 & 4 & -5 \\ 0 & -2 & -8 & 22 \end{pmatrix}$$
$$\xrightarrow[\textit{replacement}]{r_3 \to r_3 + 2r_2} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$\Rightarrow$ *The third row implies* $0 = 2 \Rightarrow$ *The linear system is inconsistent.*

Example 1.3. *consider the following linear system*

$$\begin{aligned} x_1 + x_2 &= 1, \\ -3x_1 + 3x_2 &= -3. \end{aligned}$$

Sol:

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

$\Rightarrow$ *The linear system has infinitely many solutions.*

1.5. REF and RREF

- Row echelon form (REF)
  1. All non-zero rows are above any zero rows;
  2. Leading entry moves to right by at least one column when rows going down;
  3. All entries in a column below a leading entry are zeros.

  Based on the REF, we have the row reduction algorithm (Gaussian elimination), which contains two parts

  1. Phase 1: Forward elimination (obtain the REF);
  2. Phase 2: Backward elimination (obtain the solution).

- Reduced row echelon form (RREF)
  1. All non-zero rows are above any zero rows;
  2. Leading entry moves to right by at least one column when rows going down;
  3. All entries in a column below a leading entry are zeros.
  4. Leading entry in each row is 1;
  5. Leading entry 1 is the only non-zero entry in its column.

  Object: find a form that is easy to obtain the solution of the linear system.

Example 1.4.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix} (RREF), \qquad \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} (not),$$

$$\begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} (REF), \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} (RREF).$$

Example 1.5. *Find RREF of the following matrix*

$$\begin{pmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{pmatrix}$$

Sol: *By the REOs, we have*

$$\begin{pmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{pmatrix} \xrightarrow[r_3 \to r_3 - 5r_1]{r_2 \to r_2 - 3r_1} \begin{pmatrix} 1 & 3 & 5 & 7 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -34 \end{pmatrix} \xrightarrow[r_2 \to r_2/(-4)]{r_3 \to r_3 - 2r_2} \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -10 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 3 & 5 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} \boldsymbol{1} & 0 & -1 & 0 \\ 0 & \boldsymbol{1} & 2 & 0 \\ 0 & 0 & 0 & \boldsymbol{1} \end{pmatrix} (\boldsymbol{1} \textit{ is pivot position})$$

1.6. Rank of a matrix

Some concepts based on RREF

1. uniqueness of RREF;
2. A pivot position $\leftrightarrow$ leading 1 in RREF;
3. A pivot column $\leftrightarrow$ containing pivot position.
4. Basic variables $\to$ pivot position;
5. Free variables $\to$ non-pivot position.

Definition 1.2. *rank A= no. of pivot positions in A.*

Theorem 1.2. *Linear systems $[A|b]$ is consistent iff rank A= rank $[A|b]$.*

Theorem 1.3. *Linear systems $[A|b]$ is consistent with n variables, then the no. of free variables =n -rank A.*

Example 1.6. *Find solution of linear system*

$$\begin{aligned} 3x_1 - 4x_2 + 2x_3 &= 0, \\ -9x_1 + 12x_2 - 6x_3 &= 0, \\ -6x_1 + 8x_2 - 4x_3 &= 0. \end{aligned}$$

Sol:

$$\begin{pmatrix} 3 & -4 & 2 & 0 \\ -9 & 12 & -6 & 0 \\ -6 & 8 & -4 & 0 \end{pmatrix} \xrightarrow[r_2 \to r_2/(-4)]{r_3 \to r_3 - 2r_2} \begin{pmatrix} 3 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & -\frac{4}{3} & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\Rightarrow$ *Rank A= 1, 1 basic variable and 2 free variables* $\Rightarrow$

$$\begin{cases} x_1 = \frac{4}{3}s + \frac{2}{3}t \\ x_2 = s \\ x_3 = t \end{cases}.$$

2. Vector

2.1. Two descriptions of vector

- geometric:
  direction interval from **A** to **B** $\overrightarrow{AB}$;
  length and direction

- algebraic:
  a collection of number arranged in column form;
  size: number of entries in the vector;
  notation: **u**, **v**.

2.2. Vector operations

- vector addition: $\mathbf{u} + \mathbf{v}$

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$$

- scalar multiplication: $k\mathbf{u}$

$$k \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} ku_1 \\ ku_2 \end{pmatrix}$$

Operations rules

2.3. Linear combination and span

Definition 2.1. *Let $S = \boldsymbol{v}_1, ..., \boldsymbol{v}_k$ be a collection of vectors in $\boldsymbol{R}^n$ and let $c_1, ..., c_k$ be numbers. The following $\boldsymbol{y}$ is a **linear combination** of the vectors in S*

$$\boldsymbol{y} = c_1\boldsymbol{v}_1 + ... + c_k\boldsymbol{v}_k,$$

or: a l.c. of $\boldsymbol{v}_1, ..., \boldsymbol{v}_k$.

Theorem 2.1. *Let $S = \boldsymbol{v}_1, ..., \boldsymbol{v}_k \in \boldsymbol{R}^n$, $\boldsymbol{y} \in \boldsymbol{span}S$ iff $[\boldsymbol{v}_1, ..., \boldsymbol{v}_k|\boldsymbol{y}]$ is consistent.*

Example 2.1. *Whether $\boldsymbol{b}$ is linear combination of the columns of $A = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$?*

$$[A|\boldsymbol{b}] = \begin{pmatrix} 1 & -4 & 2 \\ 0 & 3 & 5 \\ -2 & 8 & -4 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 3 \\ -7 \\ -3 \end{pmatrix}.$$

***Sol**: $\Leftrightarrow$ $\boldsymbol{b} \in (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ $\Leftrightarrow$ $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3|\boldsymbol{b})$ is consistent.*

$$[A|\boldsymbol{b}] = \begin{pmatrix} 1 & -4 & 2 & 3 \\ 0 & 3 & 5 & -7 \\ -2 & 8 & -4 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -4 & 2 & 3 \\ 0 & 3 & 5 & -7 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

$\Rightarrow$ *$[A|\boldsymbol{b}]$ is inconsistent $\Rightarrow$ $\boldsymbol{b}$ is not linear combination of the columns of A.*

Example 2.2.

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} -3 \\ 1 \\ 8 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} h \\ -5 \\ -3 \end{pmatrix}.$$

Whether $\boldsymbol{b}$ is linear combination of the columns of $(\boldsymbol{v}_1, \boldsymbol{v}_2)$?

Sol: *$\Leftrightarrow$ $(\boldsymbol{v}_1, \boldsymbol{v}_2|\boldsymbol{b})$ is consistent.*

$$[\boldsymbol{v}_1, \boldsymbol{v}_2|\boldsymbol{b}] = \begin{pmatrix} 1 & -3 & h \\ 0 & 1 & -5 \\ -2 & 8 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -3 & h \\ 0 & 1 & -5 \\ 0 & 2 & -3+2h \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -3 & h \\ 0 & 1 & -5 \\ 0 & 0 & 2h+7 \end{pmatrix}$$

$[\boldsymbol{v}_1, \boldsymbol{v}_2|\boldsymbol{b}]$ is consistent iff $2h + 7 = 0$ i.e. $h = -7/2$.

Remark 2.1. *Linear combination is a very important concept in this course!!!*

Theorem 2.2. *Let A be $m \times n$ matrix. The followings will be equivalent*

- *$A\boldsymbol{x} = \boldsymbol{b}$ is consistent for each $\boldsymbol{b} \in \boldsymbol{R}^m$;*
- *Each $\boldsymbol{b} \in \boldsymbol{R}^m$ is a l.c. of the columns of A;*
- *The column of A span $\boldsymbol{R}^m$;*
- *A has a pivot position in every row.*

Remark 2.2. *Try to find the relation between the linear combination and linear system.*

- *linear combination*

$$x_1\boldsymbol{v}_1 + ... + x_k\boldsymbol{v}_k = \boldsymbol{b}.$$

- *matrix equation $A = [\boldsymbol{v}_1, ..., \boldsymbol{v}_k]$*

$$A\boldsymbol{x} = \boldsymbol{b}.$$

Key point

1. *$A\boldsymbol{x} = \boldsymbol{b}$ is consistent $\Leftrightarrow$*
2. *$x_1\boldsymbol{a}_1 + ... + x_n\boldsymbol{a}_n = \boldsymbol{b}$ is consistent $\Leftrightarrow$*
3. *$\boldsymbol{b}$ is a l.c. of $\boldsymbol{a}_1, ..., \boldsymbol{a}_n$ $\Leftrightarrow$*
4. *$\boldsymbol{b}$ is contained in Span$\{\boldsymbol{a}_1, ..., \boldsymbol{a}_n\}$.*

Example 2.3. *matrix form $\Leftrightarrow$ vector equation*

- $A\boldsymbol{x} = b \Rightarrow x_1\boldsymbol{a}_1 + ... + x_n\boldsymbol{a}_n = \boldsymbol{b}$,

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \Leftrightarrow x_1\begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2\begin{pmatrix} 2 \\ 6 \end{pmatrix} + x_3\begin{pmatrix} 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

- $y_1\boldsymbol{a}_1 + ... + y_n\boldsymbol{a}_n = \boldsymbol{b} \Rightarrow A\boldsymbol{y} = b$,

$$y_1\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + y_2\begin{pmatrix} 2 \\ 6 \\ 3 \end{pmatrix} + y_3\begin{pmatrix} 3 \\ 7 \\ 8 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 2 & 6 & 7 \\ 1 & 3 & 8 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}.$$

Example 2.4. *Determine whether the columns of A spans $\boldsymbol{R}^3$, where*

$$A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3] = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ -2 & -4 & -1 \end{pmatrix}.$$

The columns of A spans $\boldsymbol{R}^3 \Leftrightarrow \forall\ \boldsymbol{b} \in \boldsymbol{R}^3$, $A\boldsymbol{x} = b$ is consistent i,e. $[A|\boldsymbol{b}]$ is consistent. By REO, we can obtain the REF

$$[A|\boldsymbol{b}] = \begin{pmatrix} 1 & 2 & 1 & b_1 \\ 0 & 1 & 3 & b_2 \\ -2 & -4 & -1 & b_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & b_1 \\ 0 & 1 & 3 & b_2 \\ 0 & 0 & 1 & b_3 + 2b_1 \end{pmatrix}$$

Obviously, this system has solution, so $A\boldsymbol{x} = b$ is consistent and A spans $\boldsymbol{R}^3$.

Actually, for the matrix A, $\boldsymbol{Rank}A = \boldsymbol{Rank}[A|\boldsymbol{b}]$, so $A\boldsymbol{x} = b$ is consistent and $[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3]$ span $\boldsymbol{R}^3$.

2.4. Homogeneous and Non-homogeneous system

2.4.1. Homogeneous system

For the linear system

$$A\mathbf{x} = \mathbf{0},$$

then it is called homogeneous system. This system is always consistent, because $\mathbf{x} = \mathbf{0}$ is solution of the system.

- Basic variables $\rightarrow$ rank A;
- Free variables $\rightarrow$ n-rank A.

Example 2.5. *Solve the homogeneous system $A\boldsymbol{x} = \boldsymbol{0}$, where*

$$A = \begin{pmatrix} 1 & 3 & -3 & 7 \\ 0 & 1 & -4 & 5 \\ 0 & 2 & -8 & 10 \end{pmatrix}$$

2.5. Linearly dependent and linearly independent set

Definition 2.2. *$\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n$ are a set of vectors, for the vector equation $a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + ... + a_n\boldsymbol{v}_n = 0$, if it has the unique zero solution, then they are called linearly independent; if it has non-zero solution, they are called linearly independent set.*

Example 2.6. *If the following vectors are linearly independent or not?*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 7 \\ 2 \\ -6 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 9 \\ 4 \\ -11 \end{pmatrix}$$

Based on the definition, we only need to check whether $Ax = b$ has the unique zero solution.

$$A = \begin{pmatrix} 1 & 7 & 9 \\ 0 & 2 & 4 \\ 0 & -6 & -11 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 7 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Obviously, this system has the unique solution $\boldsymbol{x} = 0$, so they are linearly independent.

Example 2.7. *For what values of h, $\boldsymbol{v}_1, \boldsymbol{v}_2\boldsymbol{v}_3$ are linearly dependent?*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} -3 \\ 3 \\ -9 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 5 \\ -7 \\ h \end{pmatrix}$$

***sol**: For $A = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3]$ we have*

$$A = \begin{pmatrix} 1 & -3 & 5 \\ -1 & 3 & 3 \\ 3 & -9 & h \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -3 & 5 \\ 0 & 0 & -2 \\ 0 & 0 & h-15 \end{pmatrix}$$

This system is always has a solution $x_1 = 3, x_2 = 1, x_3 = 0$ for any h, so they are always linearly dependent.

Example 2.8. *Are $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ linearly dependent?*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

Based on the definition, we need to verify that $\boldsymbol{v}_1x_1 + \boldsymbol{v}_2x_2 + \boldsymbol{v}_3x_3 = 0$ has unique zero solution.

$$A = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3] = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 7 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

$\Rightarrow$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \Rightarrow \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} x_3$$

Remark 2.3. *n vectors in $\mathbb{R}^m$, where $n > m$ are always linear dependent.*

2.6. Linaer transformation

For the linear system $A\mathbf{x} = \mathbf{b}$, two points of view

- old one: when $\mathbf{b}$ is given, find $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$, ***linear system***.
- new one: study $\mathbf{x} \mapsto A\mathbf{x}$, see which $\mathbf{x}$ lands on $\mathbf{b}$, ***transformation***.

Notations for the following transformation

$$T : R^n \to R^m,$$
$$T : \mathbf{x} \mapsto A\mathbf{x}.$$

where

1. domain R^n,
2. codomain R^m,
3. $T(\mathbf{x})$: image of $\mathbf{x}$, actually $A\mathbf{x}$,
4. range: the collection of all possible $T(\mathbf{x})$.

Definition 2.3. *A transformation $T : R^n \to R^m$ is called linear, if*

- $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$,
- $T(k\boldsymbol{u}) = kT(\boldsymbol{u})$,
- *combine these two points $T(k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2) = k_1T(\boldsymbol{u}_1) + k_2T(\boldsymbol{u}_2)$ (usually use this one).*

For A linear transformation T, two properties

- $T(0) = 0$,
- $T(c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n) = c_1T(\mathbf{u}_1) + ... + c_nT(\mathbf{u}_n)$.

Example 2.9. *Linear transformation?*

$$T_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} (\times), T_2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1^3 \\ x_2 \\ x_3 \end{pmatrix} (\times), T_3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} (\surd)$$

Example 2.10. *T($\boldsymbol{x}$)=A$\boldsymbol{x}$, find $\boldsymbol{x}$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$.*
Case 1.

$$A = \begin{pmatrix} 1 & 0 & -2 \\ -2 & 1 & 6 \\ 3 & -2 & -5 \end{pmatrix}, \boldsymbol{b} \begin{pmatrix} -1 \\ 7 \\ -3 \end{pmatrix}$$

sol:

$$[A|b] = \begin{pmatrix} 1 & 0 & -2 & -1 \\ -2 & 1 & 6 & 7 \\ 3 & -2 & -5 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 5 \\ 0 & -2 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$\Rightarrow$ *this system has a unique solution*

$$\boldsymbol{x} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$$

Case 2.

$$A = \begin{pmatrix} 1 & -5 & -7 \\ -3 & 7 & 5 \end{pmatrix}, \boldsymbol{b} \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

sol:

$$[A|b] = \begin{pmatrix} 1 & -5 & -7 & -2 \\ -3 & 7 & 5 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -5 & -7 & -2 \\ 0 & -8 & -18 & -8 \end{pmatrix} \to \begin{pmatrix} 1 & -5 & -7 & -2 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & 3 & 3 \\ 0 & 1 & 2 & 1 \end{pmatrix}$$

$\Rightarrow$ *this system has infinitely many solutions*

$$\boldsymbol{x} = \begin{pmatrix} -3 \\ -2 \\ 1 \end{pmatrix} x_3 + \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$

Based on the definition, for these two cases, we need to solve the linear system $\boldsymbol{A}x = \boldsymbol{b}$.

Actually, for matrix A, with matrix-vector multiplication, we have

- $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$,
- $A(k\mathbf{u}) = kA\mathbf{u}$.

So, based on the definition, if we have a matrix, with with matrix-vector multiplication, we can define a linear transformation. This point implies that when we have a matrix A, we can define a linear transformation $T : \mathbf{x} \mapsto A\mathbf{x}$. i.e. $A \to T$. In the following, we will also know that if we have a linear transformation T, we will find its corresponding matrix A, i.e. $T \to A$.

Theorem 2.3. *$T : R^n \to R^m$ be a linear transformation, and set $A = [T\boldsymbol{e}_1, ..., T\boldsymbol{e}_n]$, then $T(\boldsymbol{x}) = A\boldsymbol{x}$, and A is the matrix corresponding to the linear transformation T.*

Let $x = x_1\boldsymbol{e}_1 + ... + x_n\boldsymbol{e}_n \Rightarrow$

$$T(x) = T(x_1\boldsymbol{e}_1 + ... + x_n\boldsymbol{e}_n) = x_1T(\boldsymbol{e}_1) + ... + x_nT(\boldsymbol{e}_n)$$
$$= [T\boldsymbol{e}_1, ..., T\boldsymbol{e}_n] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A\boldsymbol{x}.$$

Example 2.11. *Find the matrix corresponding to the linear transformation.*
case 1.

$$T\boldsymbol{x} = (0, x_1 + x_2, x_2 + x_3, x_3 + x_4)^T$$

For this transformation, it can be verified that it is a linear transformation, and

$$\begin{aligned} T\boldsymbol{e}_1 &= (0,1,0,0)^T, \\ T\boldsymbol{e}_2 &= (0,1,1,0)^T, \\ T\boldsymbol{e}_3 &= (0,0,1,1)^T, \\ T\boldsymbol{e}_4 &= (0,0,0,1)^T, \end{aligned}$$

so

$$A = [T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), T(\boldsymbol{e}_3), T(\boldsymbol{e}_4)] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

case 2.

$$T(\boldsymbol{x}) = (x_1 - 5x_2 + 4x_3, x_2 - 6x_3)^T$$

For this transformation, it can be verified that it is a linear transformation, and

$$\begin{aligned} T\boldsymbol{e}_1 &= (1,0)^T, \\ T\boldsymbol{e}_2 &= (-5,1)^T, \\ T\boldsymbol{e}_3 &= (4,-6)^T, \end{aligned}$$

so

$$A = [T\boldsymbol{e}_1, T\boldsymbol{e}_2, T\boldsymbol{e}_3] = \begin{pmatrix} 1 & -5 & 4 \\ 0 & 1 & -6 \end{pmatrix}.$$

In summary, we can draw the conclusion **Matrix** $A \Leftrightarrow$ **linear transformation** T.

Definition 2.4. $T : R^n \to R^m$ *is a linear transformation, the* ***kernel*** *of* T *is defined as* $kerT = \{x \in R^n | T(x) = 0\}$.

Example 2.12.

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

We need to solve $A\boldsymbol{x} = 0$, so

$$A = \rightarrow \begin{pmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 0 & 2 & -8 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -4 & 7 & -5 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -9 & 7 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} 9 \\ 4 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} -7 \\ -3 \\ 0 \\ 1 \end{pmatrix} x_4 = \boldsymbol{u}x_3 + \boldsymbol{v}x_4,$$

thus $\ker T = Span\{\boldsymbol{u}, \boldsymbol{v}\}$.

2.6.1. Two properties of linear transformation

- One to one property: If $T(\mathbf{x}_1) = b = T(\mathbf{x}_2) \Rightarrow \mathbf{x}_1 = \mathbf{x}_2$.
  T is one to one$\Leftrightarrow T(\mathbf{x}) = 0$ has trivial solution $\Leftrightarrow a_1x_1 + ... + a_nx_n = \mathbf{0}$ has unique zero solution $\Leftrightarrow$ the columns of A are linearly dependent.

- Onto property: $\forall \mathbf{b} \Rightarrow \exists \mathbf{x}$ s.t. $T(\mathbf{x}) = \mathbf{b}$.
  $A\mathbf{x} = \mathbf{b}$ is consistent $\Leftrightarrow$ T is onto, so T is onto $\Leftrightarrow$ the columns of A span $\mathbb{R}^m$.

3. Matrix

3.1. Matrix operations

1. Matrix addition and scalar multiplication

- $A + B$, A and B have the same size.
- $k \cdot A$, k is a real number.
- $A - B$.

Example 3.1.

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

$$3 \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

2. Matrix multiplication
How to define AB,

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

where

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

Example 3.2.

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

what k, such that $AB = BA$?

$$AB \begin{pmatrix} 23 & 5k-10 \\ -9 & k+15 \end{pmatrix}, BA = \begin{pmatrix} 23 & 15 \\ 6-3k & k+15 \end{pmatrix}$$

$\Rightarrow$

$$\begin{aligned} 5k - 10 &= 15 \\ 6 - 3k &= -9. \end{aligned}$$

Thus $k = 5$.

Generally, $AB \neq BA$**.**
4. Power
$A^k = \underbrace{A \cdot A \cdot A \cdot \dots \cdot A}_{k}$.
5. Transpose $A \in R^{m \times n}$ and $A^T \in R^{n \times m}$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{m2} \\ \dots & \dots & \dots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

- If $A^T = A$, A is called symmetric.
- If $A^T = -A$, A is called skew-symmetric.

3.2. Invertible matrix

Definition 3.1. *For a matrix A, if $\exists C$, s.t. $AC = I_n = CA$, then A is called invertible.*

Example 3.3.

$$I_n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & 1 \end{pmatrix}, I_n^{-1} = I_n.$$

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, AA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow A^{-1} = A.$$

$$A^5 - 5A - I_n = 0 \Rightarrow (A^4 - 5I_n) \cdot A = A \cdot (A^4 - 5I_n) = I_n$$
$$\Rightarrow A^{-1} = A^4 - 5I_n.$$

Theorem 3.1. *For 2×2 matrix*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

when $ad - bc \neq 0$, A is invertible and

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

Based on the theorem above, it is easy to obtain the inverse for 2×2 matrix.

How to find the inverse of matrix for $n \times n$ matrix? For the matrix product, we have

$$AB = A[\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n] = [A\mathbf{b}_1, A\mathbf{b}_2, ..., A\mathbf{b}_n] = [\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n] = I_n,$$

so $A^{-1} = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n]$, and $\mathbf{b}_i$ is the unique solution of $A\mathbf{x} = \mathbf{e}_i$. we have know that we can solve this equation by EROs, i.e., $[A|\mathbf{e}_i] \Rightarrow [I_n|\mathbf{b}_i]$.

Theorem 3.2. *A is invertible matrix, if by EROs, $[A|I_n] \Rightarrow [I_n|B]$, then $B = A^{-1}$.*

Example 3.4. *Find the inverse of the following matrix*

$$A \begin{pmatrix} 1 & 1 & 0 & -3 \\ -1 & -2 & 1 & 8 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

We performed EROs for $[A|I_4]$

$$\left(\begin{array}{cccc|cccc} 1 & 1 & 0 & -3 & 1 & 0 & 0 & 0 \\ -1 & -2 & 1 & 8 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[r_4 \to -r_3 + r_4]{r_2 \to r_2 + r_1} \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & -3 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 5 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow[r_3 \to 3r_4 + r_3]{r_1 \to 3r_4 + r_1, r_2 \to -5r_4 + r_2} \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & -3 & 3 \\ 0 & -1 & 1 & 0 & 1 & 1 & 5 & -5 \\ 0 & 0 & 1 & 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow[r_3 \to -r_3]{r_2 \to -r_3 + r_2} \left(\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & -3 & 3 \\ 0 & 1 & 0 & 0 & -1 & -1 & -7 & 8 \\ 0 & 0 & 1 & 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow{r_1 \to -r_2 + r_1} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 2 & 1 & 4 & -5 \\ 0 & 1 & 0 & 0 & -1 & -1 & -7 & 8 \\ 0 & 0 & 1 & 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 1 \end{array}\right)$$

$$\Rightarrow A^{-1} = \begin{pmatrix} 2 & 1 & 4 & -5 \\ -1 & -1 & -7 & 8 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

3.2.1. Elementary matrix

Definition 3.2. *An elementary matrix is obtained by performing a single ERO on I_n.*

When $A \to B$ by a single ERO, we have $[A|I] \to [B|E]$. The matrix E is the elementary matrix corresponding to the single ERO.

The procedure of obtaining the inverse matrix is a sequence of EROs, so we have a sequence of elementary matrices $E_n, E_{n-1}, ..., E_1$, and

$$[A|I] \to [E_n E_{n-1} ... E_1 A | E_n E_{n-1} ... E_1 I] = [I|B]$$

and $A^{-1} = B = E_n E_{n-1} ... E_1$.

As we know, there are three kinds of Elementary row operations (ERO)

1. replacement,
2. interchange,
3. scaling.

Example 3.5. *A is 3×4 matrix.*

(1) Write down the corresponding to the following EROs: (1) $r_1 \leftrightarrow r_2$, (2) $r_3 \to 2r_1 + r_3$, (3) $-r_2$.

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(2) B is obtained by (1), (2), (3), write P, s.t $B = PA$
We have know that $B = E_3E_2E_1A$, so $P = E_3E_2E_1$ and

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

For Example. 3.4, we can also obtain the sequence of elementary matrices $E_n, E_{n-1}, ..., E_1$, s.t $A^{-1} = E_nE_{n-1}...E_1$

4. Determinant

4.1. Definition of determinant

- for 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the determinant of A is defined as $\det A = ad - bc$.

- for 3×3 matrix

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

the determinant of A is defined

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

- for $n \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix},$$

the determinant of A is defined

$$\begin{aligned}\det A &= \sum_{i=1}^{n} a_{1i}\cdot(-1)^{i+1}\det A_{1i} = \sum_{j=1}^{n} a_{j1}\cdot(-1)^{1+j}\det A_{j1}\\ &= \sum_{i=1}^{n} a_{ki}\cdot(-1)^{i+k}\det A_{ki} = \sum_{j=1}^{n} a_{jk}\cdot(-1)^{k+j}\det A_{jk},\end{aligned}$$

where $C_{ki} = (-1)^{i+k}\det A_{ki}$ and $C_{jk} = (-1)^{k+j}\det A_{jk}$ are cofactor of A.

- for $n\times n$ triangle matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ 0 & a_{22} & \dots & a_{1n}\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & a_{nn}\end{pmatrix},$$

$\det A = a_{11}\cdot a_{22}\cdot \dots \cdot a_{nn}$.

Theorem 4.1. *Let A be a $n\times n$ matrix,*

- *If $A \to B$ by $kr_j + r_i$, then $\det B = \det A$,*
- *If $A \to B$ by $r_j \leftrightarrow r_i$, then $\det B = -\det A$,*
- *If $A \to B$ by lr_i, then $\det B = l\det A$.*

Example 4.1.

1. *Calculate the determinant of $A+B$, $A-2B$, where*

$$A = \begin{pmatrix} 1 & 2\\ 2 & 1\end{pmatrix}, B = \begin{pmatrix} 1 & -2\\ 1 & 3\end{pmatrix}.$$

2. *Calculate the determinant of A, where*

$$A = \begin{pmatrix} 9 & 0 & 0 & 2\\ 7 & 3 & 2 & 8\\ 3 & 0 & 0 & 0\\ 5 & -3 & 1 & 11\end{pmatrix}$$

$$\det A = \begin{vmatrix} 9 & 0 & 0 & 2\\ 7 & 3 & 2 & 8\\ 3 & 0 & 0 & 0\\ 5 & -3 & 1 & 11\end{vmatrix} = 3\begin{vmatrix} 0 & 0 & 2\\ 3 & 2 & 8\\ -3 & 1 & 11\end{vmatrix} = 3\cdot 2\begin{vmatrix} 3 & 2\\ -3 & 1\end{vmatrix} = 6\cdot 9 = 54.$$

3. *Calculate the determinant of A, where*

$$A = \begin{pmatrix} 1 & x & x^2 & x^3 \\ x^3 & 1 & x & x^2 \\ x^2 & x^3 & 1 & x \\ x & x^2 & x^3 & 1 \end{pmatrix}$$

By the row operations, we have

$$\det A = \begin{vmatrix} 1 & x & x^2 & x^3 \\ 0 & 1-x^4 & x-x^5 & x^2-x^6 \\ 0 & 0 & 1-x^4 & x-x^5 \\ 0 & 0 & 0 & 1-x^4 \end{vmatrix} = (1-x^4)^3$$

4. *Calculate the determinant of A, where*

$$A = \begin{pmatrix} 1 & x & x & x \\ x & x & x & x \\ x & x & x^2 & x \\ x & x & x & x^3 \end{pmatrix}$$

By the row operations, we have

$$\det A = \begin{vmatrix} 1-x & 0 & 0 & 0 \\ x & x & x & x \\ 0 & 0 & x^2-x & 0 \\ 0 & 0 & 0 & x^3-x \end{vmatrix} = (1-x)\begin{vmatrix} x & x & x \\ 0 & x^2-x & 0 \\ 0 & 0 & x^3-x \end{vmatrix} = -(1-x)^2x^2(x^3-x).$$

5. *Calculate the determinant of A, where*

$$A_5 = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}.$$

For the matrix A_5, we have

$$\det A_5 = \begin{vmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & 2 & 0 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{vmatrix} - 4 \cdot \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{vmatrix} = \det A_4 - 4\det A_3,$$

so$\det A_5 = \det A_4 - 4\det A_3$*, and* $\det A_i$ *can be considered as a sequence with* $\det A_1 = 1$ *and* $\det A_2 = -3$*.*

6. *Let A denote the following matrix*

$$A = \begin{pmatrix} a & d & c & b \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{pmatrix}.$$

(i) Assume $a = 1, b = 2, c = 3, d = 4$, $\det A$?

$$\det A = \begin{vmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 4 & 3 \\ 3 & 2 & 1 & 4 \\ 4 & 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 3 & 2 \\ 0 & -7 & -2 & -1 \\ 0 & -10 & -8 & -2 \\ 0 & -13 & -10 & -7 \end{vmatrix} = \begin{vmatrix} -7 & -2 & -1 \\ -10 & -8 & -2 \\ -13 & -10 & -7 \end{vmatrix} = \begin{vmatrix} -7 & -2 & -1 \\ 4 & -4 & 0 \\ 36 & 4 & 0 \end{vmatrix}$$

so $\det A = -160$

(ii) Show that $\det A$ *contains the* $(a+b+c+d)$ *factor*
By row operations

$$\det A = \begin{vmatrix} a & d & c & b \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix} = \begin{vmatrix} a+b+c+d & a+b+c+d & a+b+c+d & a+b+c+d \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix}$$

$$= (a+b+c+d) \begin{vmatrix} 1 & 1 & 1 & 1 \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix}.$$

(ii) Show that $\det A$ *contains the* $(a-b+c+d)$ *factor*
By row operations

$$\det A = \begin{vmatrix} a & d & c & b \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix} = \begin{vmatrix} a-b+c-d & d-a+b-c & c-d+a-b & b-c+d-a \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix}$$

$$= (a-b+c-d) \begin{vmatrix} 1 & -1 & 1 & -1 \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{vmatrix}.$$

4.2. Cramer's rule

Theorem 4.2. *When* $\det A \neq 0$, *A is invertible, the unique solution of $A\boldsymbol{x} = \boldsymbol{b}$ is given by*

$$x_i = \frac{\det A_i(b)}{\det A},$$

where $A_i(\boldsymbol{b}) = [\boldsymbol{a}_1, ..., \boldsymbol{a}_{i-1}, \boldsymbol{b}, \boldsymbol{a}_{i+1}, ..., \boldsymbol{a}_n]$.

Example 4.2. *Use Cramer's rule to solve the system*

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8 \end{aligned}$$

For this system,

$$A = \begin{pmatrix} 3 & -2 \\ -5 & 4 \end{pmatrix}, A_1(\boldsymbol{b}) = \begin{pmatrix} 6 & -2 \\ 8 & 4 \end{pmatrix}, A_2(\boldsymbol{b}) = \begin{pmatrix} 3 & 6 \\ -5 & 8 \end{pmatrix},$$

so $x_1 = \dfrac{\det A_1(b)}{\det A} = 20$ *and* $x_2 = \dfrac{\det A_2(b)}{\det A} = 27$.

5. Vector space

This section contains some abstract concepts, which are not easy for new learners, so you should firstly review these concepts. In this course, calculation is not difficult, but the understanding of these concepts is the most important thing. The concepts in this chapter is similar with that in the chapters before, you should find and construct the relations between these concepts and it will be helpful to understand them well.

5.1. Vector space and subspace

Definition 5.1. *A collection of vectors V, with vector addition and scalar multiplication, is called vector space, if the following points are satisfied.*

1. $\boldsymbol{u} + \boldsymbol{v}$ *is always in* V,
2. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$,
3. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$,
4. *has* $\boldsymbol{0}$, *s.t.* $\boldsymbol{u}+\boldsymbol{0}=\boldsymbol{0}$,
5. *for each* $\boldsymbol{u} \in U$, *s.t.* $\boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{0}$,
6. $c\boldsymbol{u}$ *is always in* V,
7. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$,
8. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$,
9. $1\boldsymbol{u}=\boldsymbol{u}$.

Example 5.1. *Some case for vector space*

- $\{\boldsymbol{0}\}$,
- R^n,
- $P_n = \{p(t)|p(t) = p + 0 + p_1 t + ... + p_n t^n, t \in R\}$, *the polynomials with degree at most equal to n,*

- *P, all the polynomials,*
- *$M_{m\times n}$, collection of all $m \times n$ matrices.*

Definition 5.2. *A subspace $H \subseteq V$ is a non-empty subset if V, s.t. H itself forms a vector space under the same vector addition and scalar multiplication.*

An alternative definition

1. *$\boldsymbol{0} \in V$ in H,*
2. *$\boldsymbol{u}, \boldsymbol{v} \in H$, then $\boldsymbol{u} + \boldsymbol{v} \in H$,*
3. *$\boldsymbol{u} \in H$ and $a \in R$, then $a\boldsymbol{u} \in H$.*

Example 5.2. *Some case for vector subspace*

1. *$\{\boldsymbol{0}\}$,*
2. *R^n,*
3. *$P_n = \{p(t)|p(t) = p + 0 + p_1 t + ... + p_n t^n, t \in R\}$, the polynomials with degree at most equal to n,*
4. *P, all the polynomials,*
5. *$M_{m\times n}$, collection of all $m \times n$ matrices.*

Definition 5.3. *$S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$ is a set of vectors in V, $\boldsymbol{y} = c_1\boldsymbol{v}_1 + ... + c_n\boldsymbol{v}_n$ is called the linear combination of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$. Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$ is the collection of all the possible linear combinations of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_n\}$.*

Based on the definition, $H = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ is a subspace of V.

Definition 5.4. *A is $m \times n$ matrix, the null space of A is defined as $\boldsymbol{Null}A = \{\boldsymbol{x}|A\boldsymbol{x} = 0, \boldsymbol{x} \in R^n\}$.*

Based on the definition, we can also verify that $\mathbf{Null}A$ is a subspace of R^n.

Example 5.3. *Describe the null space of A, where*

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 5 & 10 & 3 \\ 1 & 3 & 8 & 2 \end{pmatrix}.$$

Based on the definition, we need to solve the linear system $A\boldsymbol{x} = 0$. By the row operations, we have

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 5 & 10 & 3 \\ 1 & 3 & 8 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -10 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

the solution can be written as

$$\boldsymbol{x} = \begin{pmatrix} 10 \\ -6 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} x_4 = \boldsymbol{v}_3 x_3 + \boldsymbol{v}_4 x_4.$$

Consequently, $NullA = Span\{\boldsymbol{v}_3, \boldsymbol{v}_4\}$.

Theorem 5.1. *For a matrix A, $NullA = \boldsymbol{0}$ equivalent to $A\boldsymbol{x} = 0$ has unique solution. When $A\boldsymbol{x} = 0$ has $\{\boldsymbol{v}_1, ..., \boldsymbol{v}_n\}$ as a set of basic solution, then $NullA = \{\boldsymbol{v}_1, ..., \boldsymbol{v}_n\}$.*

Definition 5.5. *$A = (\boldsymbol{a}_1, ..., \boldsymbol{a}_n)$ is $m \times n$ matrix, the column space of A is defined as $\boldsymbol{Col}A = Span\{\boldsymbol{a}_1, ..., \boldsymbol{a}_n\}$.*

Based on the definition, $\mathbf{a} \in \mathbf{Col}A$ equivalent to the linear systen $[A|\mathbf{v}]$ is consistent.

Example 5.4. *whether $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $\boldsymbol{Col}A$, where*

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 0 & 6 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

Based on the definition, we check the linear system $A\boldsymbol{x} = \boldsymbol{v}_i$ is consistent or not.

By the row operations, we have

$$[A|\boldsymbol{v}_1] = \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 2 & 4 & 0 & 6 & 2 \\ 1 & 2 & 2 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & -2 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix},$$

so $[A|\boldsymbol{v}_1]$ is consistent and $\boldsymbol{v}_1 \in \boldsymbol{Col}A$.

$$[A|\boldsymbol{v}_2] = \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 2 & 4 & 0 & 6 & 4 \\ 1 & 2 & 2 & 1 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & -2 & 2 & 2 \\ 0 & 0 & 1 & -1 & 6 \end{pmatrix},$$

so $[A|\boldsymbol{v}_2]$ is inconsistent and $\boldsymbol{v}_2$ is not in $\boldsymbol{Col}A$.

Remark 5.1. *T is the linear transformation corresponding to A, $\boldsymbol{v} \in$ the range of $T \leftrightarrow \exists \boldsymbol{x} \in R^n$, s.t. $T(\boldsymbol{x}) = \boldsymbol{v} \leftrightarrow [A|\boldsymbol{v}]$ is consistent, thus $\boldsymbol{Col}A$ is the same as the range of T. $\boldsymbol{Col}A = R^m$ iff every rows of A has a pivot position.*

With the definition of column space, the row space can be defined as $\mathbf{Row}A = \mathbf{Col}A^T$.

5.2. Linear transformation in general

Definition 5.6. *V and W are two vector spaces, a transformation $T : V \to W$ is called linear, if*

- $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$,
- $T(k\boldsymbol{u}) = kT(\boldsymbol{u})$,
- *combine these two points* $T(k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2) = k_1T(\boldsymbol{u}_1) + k_2T(\boldsymbol{u}_2)$ *(usually use this one).*

If $V = W$, T is called linear operator.

It can be observed that this definition is similar with that of a linear transformation $T : R^n \to R^m$. However, here $T : V \to W$ and V and W are two vector spaces. If we take $V = R^n$ and $W = R^m$, two definitions will be the same. So it is called linear transformation in general.

Definition 5.7. *V and W are two vector spaces, Let $T : V \to W$ is linear transformation*

- *the kernel of T,* $\ker T = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = 0\}$,
- *the image of T,* $ImT = \{\boldsymbol{w} \in W : \exists \boldsymbol{v} s.t. T(\boldsymbol{v}) = \boldsymbol{w}\}$.

For the linear transformation $T : V \to W$, we also define the one-to-one and onto properties.

- T is one-to-one, iff $\ker T = \mathbf{0}$,
- T is onto, iff $ImT = W$.

Example 5.5. *V=W=$P(R)$ the vector composed of all the polynomials.*

- $D(p) = \frac{d}{dt}p(t)$ *(Differential operator)* $\Rightarrow$ *onto not one-to-one,*
- $I_a(p) = \int_a^t p(s)ds$ *(Integral operator)* $\Rightarrow$ *one-to-one not onto.*

It can be verified that $\ker T$ and ImT are subspace of V and W respectively. If $V = R^n$ and $W = R^m$, we can find $A \leftrightarrow T$, then $\ker T = \mathbf{Null}A$ and $ImT = \mathbf{Col}A$. (prove)

5.3. Basis for subspace

Definition 5.8. *H is a subspace of V, B is called a basis for H, if*

- *B is a linearly independent set; (moving any vector from it will make the span smaller),*
- ***Span**B=H; (every vector in H can be written as a l.c. of the vector in B).*

Example 5.6. *Some examples about basis.*

- *$\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ is a basis for R^n.*
- *$\{1, t, ..., t^n\}$ is a basis for P^n.*
- *$\{[0, 1]^T, [1, 1]^T\}$ and $\{[1, -1]^T, [1, 0]^T\}$ are two basis for R^2.*

It can be observed that for the vector space R^n, the basis must contains n vectors, otherwise if $m < n$, the Rank of the matrix composed by the m vectors must be less than m i.e. the span of these vector can not be R^n. If we have set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$, they are the basis of R^n iff $\det A \neq 0$ where $A = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n]$.

5.4. Coordinate vector relative to a basis

$B = \{\mathbf{b}_1, ..., \mathbf{b}_p\}$ is a basis for H, the coordinate vector of $\mathbf{x}$ relative to B is the vector in R^p formed by numbers $c_1, ..., c_p$, such that $\mathbf{x} = c_1\mathbf{b}_1 + ... + c_p\mathbf{b}_p$, then the coordinate vector is denoted as $[\mathbf{x}]_b = [c_1, ..., c_p]^T$.

Example 5.7. *$B = \{[3, 1]^T, [1, 3]^T\}$, find the coordinate vector for any $x = [x_1, x_2]^T \in R^2$.*

Firstly, we need to check these to vectors are basis for R^2. Based on the definition of the coordinate vector, we need to find c_1, c_2, s.t.

$$\begin{aligned} x_1 &= 3c_1 + c_3, \\ x_2 &= c_1 + 3c_2, \end{aligned}$$

$\Rightarrow$

$$\begin{aligned} c_1 &= \frac{3x_1 - x_2}{8}, \\ c_2 &= \frac{3x_2 - x_1}{8}, \end{aligned}$$

so, for any $\boldsymbol{x}$, the coordinate vector is $[\boldsymbol{x}]_b = [\frac{3x_1 - x_2}{8}, \frac{3x_2 - x_1}{8}]^T$

Example 5.8. $B = \{\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\}$, *find the coordinate vector for $\boldsymbol{e}_1$.*

Firstly, we need to check these to vectors are basis for R^3. To find c_1, c_2, c_3, we need to solve the following equation

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

The solution is $c_1 = \frac{1}{2}, c_2 = \frac{1}{2}, c_3 = 0$, so the coordinate vector is $[\boldsymbol{e}_1]_b = [\frac{1}{2}, \frac{1}{2}, 0]^T$.

5.5. Dimension of vector space

Let $B = \{\mathbf{b}_1, ..., \mathbf{b}_p\}$ be a basis for H, then any subset of H containing $q > p$ vectors must be linearly independent. With this observation, we can prove that if B and C are two basis of H containing n and m vectors, then $n = m$. Thus, for a vector space, any basis must containing the same of vectors, we have the following definition

Definition 5.9. *When V has a basis B with n vectors, V will be called finite dimensional space, and $\dim V = n$.*

- $\dim R^n = n$, $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ is a basis of R^n.
- $\dim P_n = n$, $\{1, t, ..., t^n\}$ is a basis of P_n.
- $\dim P = \infty$, $\{1, t, t^2, ...\}$ is a basis of P and it contains infinity number of elements.

Example 5.9. *Describe all possible subspace H of R^3*

- $\dim H = 0$, *H=$\{\boldsymbol{0}\}$*
- $\dim H = 1$, *H=Span$\{\boldsymbol{v}\}$, where $\boldsymbol{v} \in R^3$*
- $\dim H = 2$, *H=Span$\{\boldsymbol{u}, \boldsymbol{v}\}$, where $\boldsymbol{u}, \boldsymbol{v} \in R^3$ and $\boldsymbol{u}, \boldsymbol{v}$ are linearly independent.*

6. Eigenvalue and eigenvector

6.1. Eigenvalue and eigenvector

Definition 6.1. *A is a square matrix, when $A\boldsymbol{x} = \lambda\boldsymbol{x}$ has a non-trivial solution $\boldsymbol{x}$, λ is called an eigenvalue of A and $\boldsymbol{x}$ is an eigenvector of A corresponding to λ.*

Example 6.1. *Find the eigenvalue and eigenvector of A*

- *$A = I_n$, for any $\boldsymbol{x} \in R^n$, we have $A\boldsymbol{x} = I_n\boldsymbol{x} = \boldsymbol{x} = 1 \cdot \boldsymbol{x}$, so 1 is the eigenvalue of A and any $\boldsymbol{x} \in R^n$ are eigenvectors corresponding to 1.*
- *$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$, when $\boldsymbol{x} = (a, -a)^T$, we have $A\boldsymbol{x} = \boldsymbol{0} = 0 \cdot \boldsymbol{x}$, so 0 is the eigenvalue of A and any $\boldsymbol{x} = (a, -a)^T$ are eigenvectors corresponding to 0.*

How to find eigenvalue and eigenvector of any matrix?

Based on the definition, if λ is an eigenvalue of A, the linear system $(A - \lambda I)\mathbf{x} = 0$ has non-zero solution $\Leftrightarrow \det(A - \lambda I) = 0$, which implies that eigenvalue λ is the root of polynomial $p(\lambda) = \det(A - \lambda I)$, and $p(\lambda)$ is called the characteristic equation of A.

Theorem 6.1. *λ is an eigenvalue of A, iff λ is the root* $\det(A - \lambda I_n) = 0$.

Example 6.2. *Find the eigenvalue and eigenvector of A, where*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 5 & 2 & 0 \\ 6 & 7 & 4 \end{pmatrix}.$$

For matrix A, $\det(A - \lambda I) = (1 - \lambda)(2 - \lambda)(4 - \lambda)$, so the eigenvalues are $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 4$. Solving the linear system $(A - \lambda_i I)\boldsymbol{x} = 0$, we can obtain the eigenvectors corresponding to λ_i.

Example 6.3. *Find the eigenvalue and eigenvector of A, where*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \xrightarrow{ERO} \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} = B$$

For matrix A, $\det(A - \lambda I) = (1 - \lambda)(2 - \lambda)$, so the eigenvalues are $\lambda_1 = 1, \lambda_2 = 2$. However, for B, $\lambda_{1,2} = \pm\sqrt{2}$. This example implies that by ERO, the eigenvalues are different.

Example 6.4. *For a matrix A,*

- *If A is not invertible, $\lambda = 0$ is an eigenvalue of A ($\det A = 0$).*
- *If λ is an eigenvalue of A, for any n, $A^n\boldsymbol{x} = A^{n-1} \cdot A\boldsymbol{x} = \lambda A^{n-1}\boldsymbol{x} = \lambda^n\boldsymbol{x}$, So λ^n is an eigenvalues of A^n.*

6.2. Diagonalization

Aim: To find a diagonal matrix D and an invertible matrix P, such that $A = PDP^{-1}$.

Suppose that $D = diag(\lambda_1, \lambda_2, ..., \lambda_n)$ and $P = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n]$, $A = PDP^{-1} \Rightarrow AP = PD \Rightarrow [Av_1, Av_2, ..., Av_n] = [\lambda\mathbf{v}_1, \lambda\mathbf{v}_2, ..., \lambda\mathbf{v}_n]$, i.e. the elements of D is the eigenvalues of A and the columns of P are eigenvector corresponding to the eigenvector.

Definition 6.2. *For a matrix A, if there exist eigenvectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, .., \boldsymbol{v}_n\}$ of A such that $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, .., \boldsymbol{v}_n]$ is invertible, A is called diagonalizable and $A = PDP^{-1}$ is a diagonalization of A.*

Example 6.5. *Whether A is diagnalizable, where*

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

$|A - \lambda I| = (\lambda - 1)^2 - 4 = 0$, $\lambda_1 = 3, \lambda_2 = -1$.

For $\lambda_1 = 3$, $(A - \lambda_1 I)\boldsymbol{x} = 0 \Rightarrow x = (\frac{1}{2}, 1)^T$. *For* $\lambda_2 = -1$, $(A - \lambda_2 I)\boldsymbol{x} = 0 \Rightarrow x = (-\frac{1}{2}, 1)^T$.

Thus A is diagnalizable, $P = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & 1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 1 & \frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix}$ *and* $D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

Example 6.6. *Whether A is diagnalizable, where*

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

For A, $\lambda_1 = \lambda_2 = 1$, *the eigenvectors corresponding to* λ_i *is* $\boldsymbol{v}_i = (s, 0)^T$ *and* $p = [\boldsymbol{v}_1, \boldsymbol{v}_2]$ *is not invertible. Thus, A is not diagnalizable.*

In general, how to determine the diagonalizability of A?
Since $AP = PD$, P is invertible $\Leftrightarrow$ A has n linear independent eigenvectors.

Theorem 6.2. *A is a* $n \times n$ *matrix*

- *A is diagonaliable iff it has n linear independent eigenvectors;*
- *The eigenvectors corresponding to distinct eigenvalues are automatically linear independent;*
- *If A has n distinct eigenvalues, A is always diagonaliable. (When A has some repeated eigenvalues, whether it is diagonaliable dependent on whether the eigenvectors are linear independent or not.)*

The step of diagonalization of a matrix

1. Find the eigenvalues of A;
2. For each eigenvalues, find a basis for the eigenspace $E_\lambda = Null(A - \lambda I)$;
3. Construct P by putting these vectors as its columns;
4. Construct D using the corresponding eigenvalues following the same order.

Example 6.7. *Exercise Diagonalize A, where*

$$A = \begin{pmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix}.$$

7. Inner product

7.1. Inner product

Inner product is a generalization of dot product. Let $\mathbf{u}, \mathbf{v}$ are vectors in R^n, the inner product is defined as

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + ... + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

Theorem 7.1. *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in R^n$ and $c \in R$ then*

1. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$;
2. $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$;
3. $(c\boldsymbol{u}) \cdot \boldsymbol{v} = c\boldsymbol{u} \cdot \boldsymbol{v}$;
4. $(\boldsymbol{u}) \cdot \boldsymbol{u} \geq 0$, *and* $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ *iff* $\boldsymbol{u} = 0$;

Some related definitions

1. Norm: $||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$;
2. Distance: $dist(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$;
3. Angle: $\angle(\mathbf{u}, \mathbf{v}) = \cos^{-1} \dfrac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$.

Based on these definitions, we have the following properties

1. Cauchy-Schwarz inequality: $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$;
2. Triangle inequality: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$;
3. Pythagoras' theorem in vector form: $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ iff $\mathbf{u} \cdot \mathbf{v} = 0$ i.e. they are orthogonal with each other.

Definition 7.1. *Let $\boldsymbol{u}, \boldsymbol{v} \in R^n$, when $\boldsymbol{u} \cdot \boldsymbol{v} = 0$, they are orthogonal, denoted as $\boldsymbol{u} \perp \boldsymbol{v}$.*

Remark

1. This is a general form of perpendicularity;
2. $\mathbf{0}$ is orthogonal with any other vectors.

Definition 7.2. *Let $S \subset R^n$, if $\boldsymbol{u}$ is orthogonal to any vectors in S, $\boldsymbol{u}$ is orthogonal to S denoted as $\boldsymbol{u} \perp S$. The orthogonal complement of S is defined as $S^\perp = \{u \in R^n | u \perp S\}$.*

1. In R^3 let S_1=x-axis, then $\mathbf{e}_2 \perp S_1$ and $\mathbf{e}_3 \perp S_1$, $S_1^\perp = Span\{\mathbf{e}_2, \mathbf{e}_3\}$;
2. In R^n let $S_2 = Span\{\mathbf{e}_1, .., \mathbf{e}_k\}$ then $\mathbf{e}_j \perp S_2$, $j = k+1, ..., n$, $S_2^\perp = Span\{\mathbf{e}_{k+1}, \mathbf{e}_n\}$.

Let A be a $m \times n$ matrix, then $(RowA)^\perp = NullA$ and $(ColA)^\perp = NullA^T$.

7.2. Orthogonal sets and orthonormal sets

Definition 7.3. *Let $S \subset R^n$ is called* ***orthogonal*** *if any two vectors in S are always orthogonal to each other. Let $S \subset R^n$ is called* ***orthonormal*** *if (1) $S \subset$ is called orthogonal; (2) each vector in S is of unit length.*

Definition 7.4. *A basis for a subspace W is called an* ***orthogonal basis*** *if it is an orthogonal set; a basis for a subspace W is called an* ***orthonormal basis*** *if it is an orthonormal set.*

Definition 7.5. *Let $S \subset R^n$ is called* ***orthogonal*** *if any two vectors in S are always orthogonal to each other. Let $S \subset R^n$ is called* ***orthonormal*** *if (1) $S \subset$ is called orthogonal; (2) each vector in S is of unit length.*

Example 7.1. *$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is an orthonormal basis for R^3. $S = \{(1,2)^T, (2,-1)^T\}$ is an orthogonal basis for R^2.*

The vectors in an orthogonal set are always linear independent. Let $S = \{\mathbf{v}_1, ..., \mathbf{v}_p\}$ be an orthogonal set, then for any $\mathbf{v} \in SpanS$, $\mathbf{v}$ can be written as

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|}\mathbf{v}_1 + ... + \frac{\mathbf{v} \cdot \mathbf{v}_p}{\|\mathbf{v}_p\|}\mathbf{v}_p. \tag{7.1}$$

Let $S' = \{\mathbf{v}'_1, ..., \mathbf{v}'_p\}$ be an orthonormal set, then for any $\mathbf{v} \in SpanS'$, $\mathbf{v}$ can be written as

$$\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{v}'_1}{\|\mathbf{v}'_1\|}\mathbf{v}'_1 + ... + \frac{\mathbf{v} \cdot \mathbf{v}'_p}{\|\mathbf{v}'_p\|}\mathbf{v}'_p.$$

Let $S = \{\mathbf{u}_1, ..., \mathbf{u}_p\}$ be an orthogonal basis for W, for each $\mathbf{v} \in R^n$, the following vector in W

$$proj_W\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|}\mathbf{u}_1 + ... + \frac{\mathbf{v} \cdot \mathbf{u}_p}{\|\mathbf{u}_p\|}\mathbf{u}_p \tag{7.2}$$

is called the orthogonal projection of $\mathbf{v}$ onto W.

Based on Eq.(7.1) and Eq.(7.2), we have $\mathbf{v} = proj_W\mathbf{v} \Leftrightarrow v \in W$.

Example 7.2. *$W = Span\{(2,-2,1)^T, (2,1,-2)^T\}$, find $proj_W\boldsymbol{v}$, where $\boldsymbol{v} = (1,2,3)^T$.*

Exercise.

Definition 7.6. *The decomposition $\boldsymbol{v} = proj_W\boldsymbol{v} + (\boldsymbol{v} - proj_W\boldsymbol{v})$ is called orthogonal decomposition of $\boldsymbol{v}$. It can be proved that the orthogonal decomposition is the unique way to write $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{z}$ with $\boldsymbol{w} \in W$ and $\boldsymbol{z} \in W^{\perp}$.*

Let $\mathbf{v} \in R^n$ and $\mathbf{w} \in W$, we have $\|\mathbf{v} - proj_W\mathbf{v}\| < \|\mathbf{v} - \mathbf{w}\|$, the equality holds when $\mathbf{w} = proj_W\mathbf{v}$, so, in some sense, $proj_W\mathbf{v}$ is the best approximation of $\mathbf{v}$ by the vectors in W.

Definition 7.7. *The distance of $\boldsymbol{v}$ to W is defined as* $dist(\boldsymbol{v}, W) = \|\boldsymbol{v} - proj_W \boldsymbol{v}\|$

Let $S = \{\mathbf{u}_1, ..., \mathbf{u}_p\}$ be an orthogonal basis for $W = SpanS$, when $W \neq R^n$, then $W^{\perp} \neq \Phi$ and there exists $\mathbf{w}$ in $W^{\perp}$, s.t. $\mathbf{w} \perp S$, hence $S \cup \{\mathbf{w}\}$ is still orthogonal. In other words, we can extend an orthogonal set by adding a vector $\mathbf{w}$. It can be summarized as follows

1. $S_1 = \{\mathbf{u}_1\}$, and $\mathbf{u}_2 \notin S_1$ then compute $\mathbf{z}_2$ s.t. $S_2 = \{\mathbf{u}_1, \mathbf{z}_2\}$ is orthogonal and $Span\{\mathbf{u}_1, \mathbf{z}_2\} = Span\{\mathbf{u}_1, \mathbf{u}_2\}$;
2. $S_2 = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, and $\mathbf{u}_3 \notin S_2$ then compute $\mathbf{z}_3$ s.t. $S_3 = \{\mathbf{u}_1, \mathbf{v}_2, \mathbf{z}_3\}$ is orthogonal, and $Span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{z}_3\} = Span\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.
3. ...

The process is celled **Gram-Schmidt orthogonalization process**.

Let $\{\mathbf{x}_1, ..., \mathbf{x}_p\}$ be a set of linear independent vectors, the process can be summarized as: choose $\mathbf{u}_1 = \mathbf{x}_1$, then

$$\begin{aligned}
\mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|}\mathbf{u}_1, \\
\mathbf{u}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|}\mathbf{u}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|}\mathbf{u}_2, \\
&\cdots \\
\mathbf{u}_p &= \mathbf{x}_p - \sum_{i=1}^{p-1} \frac{\mathbf{x}_p \cdot \mathbf{u}_i}{\|\mathbf{u}_i\|}\mathbf{u}_i.
\end{aligned}$$

Then $Span\{\mathbf{x}_1, ..., \mathbf{x}_p\} = Span\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ and $\{\mathbf{u}_1, ..., \mathbf{u}_p\}$ is an orthogonal set.

8. Symmetric matrix

8.1. Eigenvectors of a symmetric matrix

Let A be symmetric, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal to each other. With this conclusion, the collection of these eigenvalues is orthogonal (orthonormal).

If $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is eigenvectors of A, then A diagonalizable and $A = PDP^{-1}$, where $P = (\mathbf{u}_1, ..., \mathbf{u}_n)$. When $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is orthonormal set, $P^T = P^{-1}$ and we have $A = PDP^T$.

Definition 8.1. *A matrix is orthogonal diagonalizable if there exist orthogonal matrix P s.t $A = PDP^T$.*

For the symmetric matrix, it must be orthogonal diagonalizable (direct conclusion).

The step of diagonalization of a symmetric matrix

1. Find the eigenvalues of A;
2. For each eigenvalues, find a basis for the eigenspace $E_\lambda = Null(A - \lambda I)$, then apply Gram-Schmidt process to convert it into an orthonormal basis;

3. Construct P by putting these vectors as its columns;
4. Construct D using the corresponding eigenvalues following the same order.

Example 8.1. *Exercise Orthogonal diagonalize A, where*

$$A = \begin{pmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix}.$$

Theorem 8.1. *(**Spectral Theorem for Symmetric Matrices**) Let A be an $n \times n$ symmetric matrix. Then:*

1. *A has n real eigenvalues, counting multiplicities.*
2. *For each λ, $\dim Null(A - \lambda I)$ = multiplicity of λ.*
3. *Eigenvectors corresponding to distinct eigenvalues are orthogonal.*
4. *A is orthogonally diagonalizable, namely, there is an orthogonal matrix P and a diagonal matrix D s.t. $A = PDP^T$.*

Theorem 8.2. *(**Principal Axes Theorem**) Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $\boldsymbol{x} = P\boldsymbol{y}$, that transforms the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ into a quadratic form $\boldsymbol{y}^T D\boldsymbol{y}$ with no cross-product terms*

$$\boldsymbol{x}^T A\boldsymbol{x} = \sum_{i;j=1}^{n} a_{ij}x_i x_j \underset{\longleftrightarrow}{\boldsymbol{x} = P\boldsymbol{y}} \sum_{i=1}^{n} \lambda_i y_i^2 = \boldsymbol{y}^T D\boldsymbol{y}.$$

This theorem can describe behavior of quadratic forms (depending on the sign of λ_i)

1. positive definite;
2. negative definite;
3. indefinite.