Introduction to Linear Algebra

Haoran Li

2025-06-14

Contents

1. Lecture 1 - System of linear equations

1.1. Linear systems

Throughout this course, we adopt the following notations:

- Natural numbers: $\mathbb{N} = \{0, 1, 2, 3, ...\}$
- Integers: $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$
- Rational numbers: $\mathbb{Q} = \left\{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\right\}$ is the set of fractions. Here $\in$ means belong to.
- Real numbers: $\mathbb{R}$ is the set of numbers on the whole real number line. It includes:
  - irrational numbers (like $\sqrt{2}$, $\sqrt[3]{3}$)
  - transcendental numbers (like π, e)
- Complex numbers: $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$, $i = \sqrt{-1}$ is the imaginary number such that $i^2 = -1$.
- $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}$
- $\mathbb{R}^n = \{(r_1, r_2, r_3, ..., r_n) \mid r_1, r_2, ..., r_n \in \mathbb{R}\}$ is the set of all n-tuples of real numbers. Geometrically:
  - $\mathbb{R}^1 = \mathbb{R}$ is a line.
  - $\mathbb{R}^2$ is a plane.
  - $\mathbb{R}^3$ is our usual physical space.

Definition 1.1.1: A linear equation in the variables $x_1, x_2, x_3, ..., x_n$ is an equation that can be written in the form

$$a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + a_n x_n = b \tag{1.1}$$

where the coefficients $a_1, a_2, a_3, ..., a_n$ and b are real or complex numbers, usually known in advance.

Example 1.1.1:

- $x_1 + \frac{1}{2} x_2 = 2$, ✓
- $\pi(x_1 + x_2) - 9.9 x_3 = e$, ✓. Because if we expand it, we got $\pi x_1 + \pi x_2 - 9.9 x_3 = e$ in which case $a_1 = \pi, a_2 = \pi, a_3 = -9.9, b = e$ as in the form of (1.1)
- $|x_2| - 1 = 0$, ✘
- $x_1 + x_2^2 = 9$, ✘
- $\sqrt{x_1} + \sqrt{x_2} = 1$, ✘

Definition 1.1.2: A system of linear equations (or a linear system) is a collection of one or more linear equations involving the same variables, say $x_1, x_2, x_3, ..., x_n$.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.2}$$

Example 1.1.2: For $n = m = 2$, (1.2) is just

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{1.3}$$

Example 1.1.3: (The Nine Chapters on the Mathematical Art). In a cage full of chickens and rabbits. The total number of heads is 10 and the total number of legs is 26. Calculate the number of chickens and rabbits.

Solution: Let's assume the number of chickens and rabbits are x_1 and x_2, then we can write down a linear system

$$\begin{cases} x_1 + x_2 = 10 \\ 2x_1 + 4x_2 = 26 \end{cases} \tag{1.4}$$

Let's solve this linear system
step 1. Replace `Row2` by `Row2` $-$ `2Row1`, we get

$$\begin{cases} x_1 + 2x_2 = 20 \\ 2x_2 = 6 \end{cases} \tag{1.5}$$

step 2. Divide `Row2` by 2, we get

$$\begin{cases} x_1 + 2x_2 = 20 \\ x_2 = 3 \end{cases} \tag{1.6}$$

step 3. Replace `Row1` by `Row1` $-$ `2Row 2`, we have the solution

$$\begin{cases} x_1 = 14 \\ x_2 = 3 \end{cases} \tag{1.7}$$

Remark: This process is call the Gaussian elimination

Definition 1.1.3: A solution of the linear system (1.2) is

$$\begin{cases} x_1 = s_1 \\ x_2 = s_2 \\ x_3 = s_3 \\ \dots \\ x_n = s_n \end{cases} \tag{1.8}$$

where $s_1, s_2, s_3, \dots, s_n$ are numbers that make (1.2) true. The set of all possible solutions is called the solution set of the linear system. To solve a linear system is to find all its solutions.

1.2. Geometric interpretation

Example 1.2.1:

$$\begin{cases} x_1 + x_2 = 10 \\ 2x_1 + 4x_2 = 26 \end{cases} \tag{1.9}$$

describes two lines in $\mathbb{R}^2$, and the solution set is the intersection.

Question: How many solutions does the following linear system have?

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{1.10}$$

Answer: It may have

- a *unique* solution if these two lines *intersect*.
- (uncountably) *infinitely many* solutions if these two lines *overlap*.
- *no* solutions if these two lines are *parallel* but not overlapping.

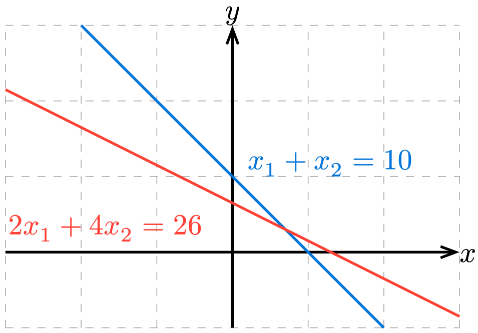
Example 1.2.2: Compare the following three linear systems

$$\begin{cases} x_1 + x_2 = 10 \\ 2x_1 + 4x_2 = 26 \end{cases} \quad (1.11) \qquad \begin{cases} x_1 + 2x_2 = 10 \\ 2x_1 + 4x_2 = 26 \end{cases} \quad (1.12) \qquad \begin{cases} x_1 + 2x_2 = 13 \\ 2x_1 + 4x_2 = 26 \end{cases} \quad (1.13)$$

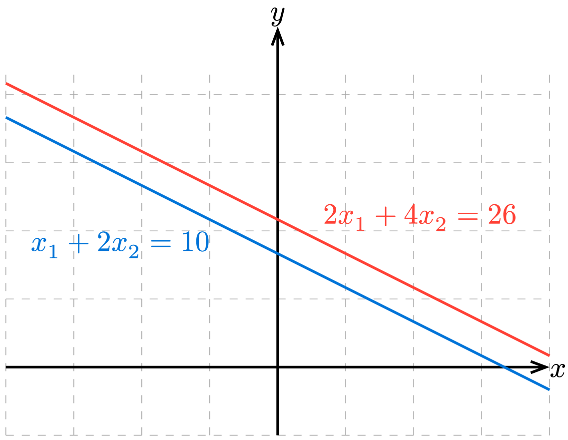
- (1.11) has a unique solution

$$\begin{cases} x_1 = 7 \\ x_2 = 3 \end{cases} \tag{1.14}$$

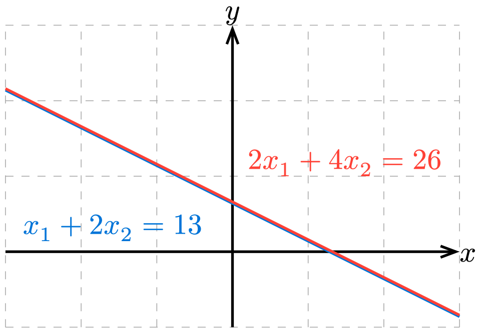
- (1.12) has no solutions since the 1st equation contradicts the 2nd.
- (1.13) has infinitely many solutions since the 2nd equation is twice of the 1st.



(a) Two lines intersect



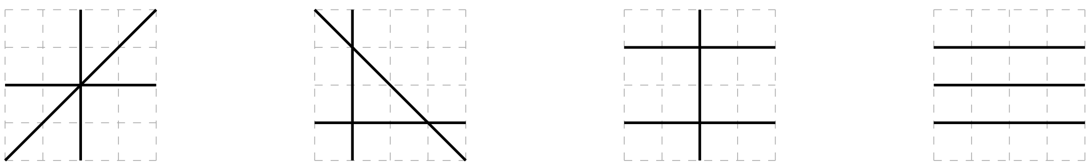
(b) Two lines intersect



(c) Two lines intersect

Figure 1: Two lines in a plane

If we increase the number of equations, we get more lines, it might looks like



If we increase the number of variables, we get
- $a_1x_1 + a_2x_2 + a_3x_3 = b$ describes a plane in $\mathbb{R}^3$.
- $a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b$ describes a *hyperplane* in $\mathbb{R}^n$.
- Therefore the solution set of (1.2) is the intersection of m hyperplanes.

Example 1.2.3: Geometric interpretation of

$$\begin{cases} x_1 - 3x_2 + 2x_3 = 0 \\ -5x_1 + 12x_2 - x_3 = 0 \end{cases} \tag{1.15}$$

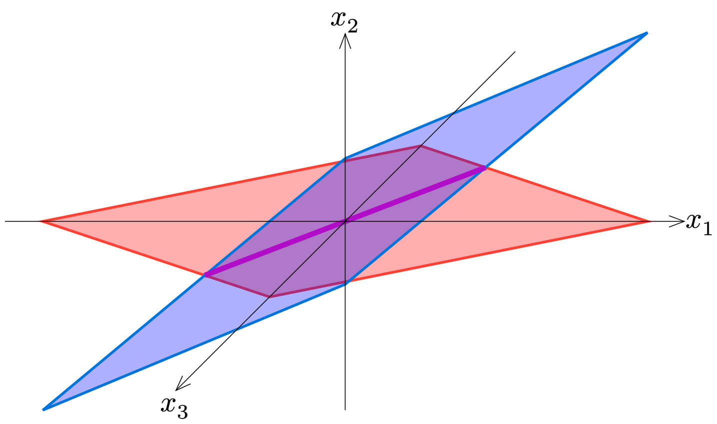


Figure 3: Two planes intersect

Remark: It is geometrically clear that for a system of 2 equations in 3 variables, there are either no solutions or infinitely many, since two planes either intersects at a line, or overlap, or simply parallel.

Definition 1.2.1: We say a linear system is consistent if it has solution(s), and inconsistent if it has none.

Example 1.2.4: In the previous Example Example 1.2.2, (1.11) and (1.13) are consistent, while (1.12) is inconsistent

Exercise 1.2.1:

1. Try Gaussian elimination on the following linear systems
   - $$\begin{cases} x_1 + 5x_2 = 7 \\ -2x_1 - 7x_2 = -5 \end{cases} \tag{1.16}$$
   - $$\begin{cases} 2x_1 + 4x_2 = -4 \\ 5x_1 + 7x_2 = 11 \end{cases} \tag{1.17}$$
   - $$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \end{cases} \tag{1.18}$$
2. Find the point of intersection of the lines $x_1 - 5x_2 = 1$ and $3x_1 - 7x_2 = 5$.
3. For what values of h and k is the following system consistent?

$$\begin{cases} 2x_1 - x_2 = h \\ -6x_1 + 3x_2 = k \end{cases} \tag{1.19}$$

2. Lecture 2 - Matrices and row echelon form

2.1. Matrices

Definition 2.1.1: A m by n (or $m \times n$) matrix is a rectangular array of numbers with m rows and n columns

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{2.20}$$

We use the (i, j)-th entry to mean the entry on the i-th row and j-column (i.e. a_{ij}).

Definition 2.1.2: A matrix is

- a zero matrix is a matrix with all entries zeros.

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{2.21}$$

- a square matrix is a matrix with the same number of rows and columns, i.e. $m = n$.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{2.22}$$

- a vector if it only has one column, i.e. $n = 1$.

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \tag{2.23}$$

- a row vector if it only has one row, i.e. $m = 1$.

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \tag{2.24}$$

- the identity matrix if it is a square matrix with diagonal elements $1'$s, and $0'$s otherwise. Here the diagonal are the (i, i)-th entries.

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.25}$$

Definition 2.1.3: Soon we will be getting tired of writing all these equations in the linear system (2.2), instead we write down its augmented matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & & \vdots & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & b_m \end{bmatrix} \tag{2.26}$$

Which is obtained by omitting x_i's, pluses, and equal signs. If we delete the last column, we will get the coefficient matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \tag{2.27}$$

Example 2.1.1:

- For (2.4), its augmented matrix and coefficient matrix are

$$\begin{bmatrix} 1 & 1 & 10 \\ 2 & 4 & 26 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \tag{2.28}$$

- For

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \end{cases} \tag{2.29}$$

  , its augmented matrix and coefficient matrix are

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \tag{2.30}$$

- In general, a linear system of m equations in n variables has a m by $(n+1)$ augmented matrix and a m by n coefficient matrix.

Definition 2.1.4: Inspired by Gaussian elimination, we define the following three elementary row operations

- Replacement: Replace one row by the sum of itself and a multiple of another row.
- Interchangement: Interchange two rows.
- Scaling: Multiply all entries in a row by a *nonzero* constant.

We say matrices A, B are row equivalent ($A \sim B$ or $A \to B$) if B can obtained by applying a sequence of elementary row operations to A (or vise versa).

Example 2.1.2: Let's rewrite the process in Example Example 1.1.3

$$\begin{bmatrix} 1 & 1 & 10 \\ 2 & 4 & 26 \end{bmatrix} \xrightarrow{\text{R2}\to\text{R2}-\text{2R1}} \begin{bmatrix} 1 & 1 & 10 \\ 0 & 2 & 6 \end{bmatrix} \xrightarrow{\text{R2}\to\frac{\text{R2}}{2}} \begin{bmatrix} 1 & 1 & 10 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{\text{R1}\to\text{R1}-\text{R2}} \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 3 \end{bmatrix} \tag{2.31}$$

Example 2.1.3: Solve

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_3 = 1 \\ x_1 + x_2 + x_3 = 3 \end{cases} \tag{2.32}$$

with augmented matrix.

$$\begin{aligned} &\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{\text{R2}\to\text{R2}-\text{2R1}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \xrightarrow{\text{R3}\to\text{R3}-\text{R1}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & 0 & 2 \end{bmatrix} \\ &\xrightarrow{\text{R3}\to\text{R3}-\text{R2}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix} \xrightarrow{\text{R3}\to\frac{\text{R3}}{3}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow[\text{R2}\to\text{R2}+\text{3R3}]{\text{R1}\to\text{R1}-\text{R3}} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ &\xrightarrow{\text{R2}\to\frac{\text{R2}}{2}} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\text{R1}\to\text{R1}+\text{R2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{aligned} \tag{2.33}$$

This gives the unique solution

$$\begin{cases} x_1 = 1 \\ x_2 = 1 \\ x_3 = 1 \end{cases} \tag{2.34}$$

2.2. Row echelon form

Definition 2.2.1:
- A leading entry of a row refers to the leftmost nonzero entry (in a nonzero row).
- A matrix is of row echelon form (REF) if it is of a "staircase shape".

$$\overset{\text{REF}}{\begin{bmatrix} \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}} \tag{2.35}$$

$\blacksquare$ are the leading entries, $*$ are some unknown numbers.
- The leading entries of an REF matrix are called pivots.
- The position of pivots are called pivot positions.
- The column pivots are in are called pivot columns.
- An REF of reduced row echelon form (RREF) if all its pivots are 1's and in each pivot column, every entry except the pivot are $0'$s.

$$\overset{\text{RREF}}{\begin{bmatrix} 1 & * & 0 & * & 0 & 0 & * & * \\ 0 & 0 & 1 & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}} \tag{2.36}$$

Example 2.2.1: In Example Example 2.1.3,

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & 0 & 2 \end{bmatrix} \tag{2.37}$$

is not an REF.

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix} \tag{2.38}$$

is an REF, but not an RREF.

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{2.39}$$

is an RREF.

Theorem 2.2.1: Every matrix is row equivalent to some REF matrix (which is not in general unique), but it is row equivalent to some unique RREF matrix.

Remark: This ensures that the pivot positions are well-defined, i.e. you won't get different pivot positions if you applied different row operations

Example 2.2.2: In Example Example 2.1.3,

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 3 & 3 \end{bmatrix} \tag{2.40}$$

is an REF that is row equivalent to the original matrix

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \tag{2.41}$$

and

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \tag{2.42}$$

is its unique row equivalent RREF.

Remark: A linear system has a unique solution if and only if its RREF deleting the last column gives the identity matrix.

Example 2.2.3: Solve

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_3 = 1 \\ x_1 + x_2 - 2x_3 = 1 \end{cases} \tag{2.43}$$

with augmented matrix.

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & -2 & 1 \end{bmatrix} \xrightarrow[]{\substack{\text{R2}\to\text{R2}-\text{2R1} \\ \text{R3}\to\text{R3}-\text{R1}}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & -3 & 0 \end{bmatrix} \xrightarrow{\text{R3}\to\text{R3}-\text{R2}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.44}$$

You might notice that the last row represents $0x_1 + 0x_2 + 0x_3 = 1$, this is a contradiction, therefore the linear system is inconsistent.

Remark: This only happens if and only if the last pivot column is the last column

Example 2.2.4:

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_3 = 1 \end{cases} \tag{2.45}$$

we write down its augmented matrix

$$\begin{aligned} &\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \end{bmatrix} \xrightarrow{\text{R2}\to\text{R2}-\text{2R1}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \end{bmatrix} \\ &\xrightarrow{\frac{\text{R2}}{2}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \end{bmatrix} \xrightarrow{\text{R1}\to\text{R1}+\text{R2}} \begin{bmatrix} 1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{3}{2} & -\frac{1}{2} \end{bmatrix} \end{aligned} \tag{2.46}$$

This gives the solution set

$$\begin{cases} x_1 - \frac{1}{2}x_3 = \frac{1}{2} \\ x_2 - \frac{3}{2}x_3 = \frac{1}{2} \end{cases} \Rightarrow \begin{cases} x_1 = \frac{1}{2}x_3 + \frac{1}{2} \\ x_2 = \frac{3}{2}x_3 - \frac{1}{2} \end{cases} \tag{2.47}$$

Let's formalize these as row reduction algorithm

step 1. Begin with the leftmost nonzero column. This is a pivot column. The pivot position should be at the top.

step 2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

step 3. Use row replacement operations to create zeros in all positions below the pivot.

step 4. Cover (or ignore) the rows containing the pivot positions. Apply Steps 1-3 to the rows that remains. Repeat the process until you are left with an REF.

step 5. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

Steps 1-4 are call forward phase, after which you get an REF. Step 5 is called backward phase, after which you get the RREF.

Example 2.2.5: Consider the augmented matrix

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \tag{2.48}$$

- Forward phase

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \xrightarrow[\text{Step 1,2}]{\text{R1}\leftrightarrow\text{R4}} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \tag{2.49}$$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \xrightarrow{\substack{\text{R2}\to\text{R2+R1} \\ \text{R3}\to\text{R3+2R1} \\ \text{Step 3}}} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \tag{2.50}$$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \xrightarrow{\substack{\text{R3}\to\text{R3}-\frac{5}{2}\text{R2} \\ \text{R4}\to\text{R4}+\frac{3}{2}\text{R2} \\ \text{Step 4,1,2,3}}} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \tag{2.51}$$

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \xrightarrow[\text{Step 4,1}]{\text{R3}\leftrightarrow\text{R4}} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.52}$$

- Backward phase

$$\begin{aligned} &\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[\text{Step 5}]{\text{R3}\to\frac{\text{R3}}{-5}} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\substack{\text{R1}\to\text{R1+9R3} \\ \text{R2}\to\text{R2+6R3} \\ \text{Step 5}}} \begin{bmatrix} 1 & 4 & 5 & 0 & -7 \\ 0 & 2 & 4 & 0 & -6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ &\xrightarrow[\text{Step 5}]{\text{R2}\to\frac{\text{R2}}{2}} \begin{bmatrix} 1 & 4 & 5 & 0 & -7 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[\text{Step 5}]{\text{R1}\to\text{R1}-4\text{R2}} \begin{bmatrix} 1 & 0 & -3 & 0 & 5 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{2.53}$$

Definition 2.2.2: The variables corresponding to pivot columns in a matrix are called basic variables, the other variables are called free variables. In a solution set, basic variables are expressed in terms of free variables, and a free variable can take any value.

Example 2.2.6: In Example Example 2.2.4, x_1, x_2 are basic variables and x_3 is a free variable. And we formally write our solution set as

$$\begin{cases} x_1 = \frac{1}{2}x_3 + \frac{1}{2} \\ x_2 = \frac{3}{2}x_3 - \frac{1}{2} \\ x_3 \text{ is free} \end{cases} \tag{2.54}$$

Exercise 2.2.1: Find the general solution of the system

$$\begin{cases} x_1 - 2x_2 - x_3 + 3x_4 = 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 = 3 \\ 3x_1 - 6x_2 - 4x_3 + 8x_4 = 2 \end{cases} \tag{2.55}$$

Solution:

$$\begin{gathered} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -4 & 8 & 2 \end{bmatrix} \xrightarrow[\text{R3}\to\text{R3}-3\text{R1}]{\text{R2}\to\text{R2}+2\text{R1}} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -1 & -1 & 2 \end{bmatrix} \xrightarrow{\text{R3}\to(-1)\cdot\text{R3}} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 1 & 1 & -2 \end{bmatrix} \\ \xrightarrow{\text{R2}\to\text{R2}-3\text{R3}} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 0 & -2 & 9 \\ 0 & 0 & 1 & 1 & -2 \end{bmatrix} \xrightarrow{\text{R2}\leftrightarrow\text{R3}} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & -2 & 9 \end{bmatrix} \\ \xrightarrow{\text{R3}\to\frac{\text{R3}}{-2}} \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} \xrightarrow[\text{R1}\to\text{R1}-3\text{R3}]{\text{R2}\to\text{R2}-\text{R3}} \begin{bmatrix} 1 & -2 & -1 & 0 & \frac{27}{2} \\ 0 & 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} \xrightarrow{\text{R1}\to\text{R1}+\text{R2}} \begin{bmatrix} 1 & -2 & 0 & 0 & 16 \\ 0 & 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} \end{gathered} \tag{2.56}$$

Write this as solution set, we get

$$\begin{cases} x_1 - 2x_2 = 16 \\ x_3 = \frac{5}{2} \\ x_4 = -\frac{9}{2} \end{cases} \Rightarrow \begin{cases} x_1 = 2x_2 + 16 \\ x_2 \text{ is free} \\ x_3 = \frac{5}{2} \\ x_4 = -\frac{9}{2} \end{cases} \tag{2.57}$$

Theorem 2.2.2: Suppose the augmented matrix of a linear system is $[A \;\; \boldsymbol{b}]$, and its RREF is $[U \;\; \boldsymbol{d}]$, then the linear system has

1. no solutions $\Leftrightarrow$ $\boldsymbol{d}$ is a pivot column, i.e. contains a pivot.
2. has solutions $\Leftrightarrow$ $\boldsymbol{d}$ is not a pivot column
   - a unique solution $\Leftrightarrow$ every column of U is a pivot column.
   - infinitely many solutions $\Leftrightarrow$ some columns of U is not a pivot column.

Example 2.2.7:

- In Example Example 2.2.3, the linear system has no solutions since

$$[A \;\; \boldsymbol{b}] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = [U \;\; \boldsymbol{d}] \tag{2.58}$$

- In Example Example 2.1.3, the linear system has a unique solution since

$$[A \;\; \boldsymbol{b}] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = [U \;\; \boldsymbol{d}] \tag{2.59}$$

- In Example Exercise 2.2.1, the linear system has infinitely many solutions since

$$[A \;\; \boldsymbol{b}] = \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -4 & 8 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & 0 & 0 & 16 \\ 0 & 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 0 & 1 & -\frac{9}{2} \end{bmatrix} = [U \;\; \boldsymbol{d}] \tag{2.60}$$

Exercise 2.2.2: Find the general solutions of the system with given augmented matrix, name the pivot columns, pivot positions, basic and free variables.

- $$\begin{bmatrix} 0 & 1 & -6 & 5 \\ 1 & -2 & 7 & -4 \end{bmatrix} \tag{2.61}$$

- $$\begin{bmatrix} 1 & -7 & 0 & 6 & 5 \\ 0 & 0 & 1 & -2 & -3 \\ -1 & 7 & -4 & 2 & 7 \end{bmatrix} \tag{2.62}$$

Question: How does the size of the augmented matrix affect the solution set?

3. Lecture 3 - Matrix algebra

3.1. Matrix addition and scalar multiplication

Definition 3.1.1: Let's use $M_{m \times n}(\mathbb{R})$ to denote the set of all (real-valued) m by n matrices.

Definition 3.1.2: Suppose A, B are $m \times n$ matrices, c is a scalar (i.e. a number), then we can define

- Addition

$$\begin{aligned} &\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \\ = &\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \end{aligned} \tag{3.63}$$

- Scalar multiplication

$$c\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix} \tag{3.64}$$

Example 3.1.1:

1.
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \tag{3.65}$$

2.
$$2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} \tag{3.66}$$

3.2. Matrix multiplication

Definition 3.2.1: Suppose A is a $m \times n$ matrix, and B is a $n \times p$ matrix, we can define matrix multiplication AB to be the $m \times p$ matrix, computed via the row-column rule: The (i, j)-entry is to multiply the i-row and j-th column

$$\begin{bmatrix} & & & \\ & & & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & & \\ & & & \end{bmatrix} \begin{bmatrix} & b_{1j} & \\ & b_{2j} & \\ & \vdots & \\ & b_{nj} & \end{bmatrix} = \begin{bmatrix} & & \\ & \blacksquare & \\ & & \end{bmatrix} \tag{3.67}$$

Where the (i, j)-entry $\blacksquare = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$.

If A is a square matrix, then we could define matrix power A^k to be simply $\overbrace{A \cdots A}^{k \text{ times}}$

Example 3.2.1:

- $$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix} \tag{3.68}$$

- $$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1\cdot 3 + 2\cdot 1 + 2\cdot 2 & 1\cdot 1 + 2\cdot 2 + 2\cdot 3 \\ 2\cdot 3 + 1\cdot 1 + 1\cdot 2 & 2\cdot 1 + 1\cdot 2 + 1\cdot 3 \end{bmatrix} = \begin{bmatrix} 9 & 11 \\ 9 & 7 \end{bmatrix} \tag{3.69}$$

- $$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_1 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + x_1 \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \tag{3.70}$$

- $$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0\cdot 1 + 1\cdot 0 & 0\cdot 0 + 1\cdot 0 \\ 0\cdot 1 + 0\cdot 0 & 0\cdot 0 + 0\cdot 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{3.71}$$

- $$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1\cdot 0 + 0\cdot 0 & 1\cdot 1 + 0\cdot 0 \\ 0\cdot 0 + 0\cdot 0 & 0\cdot 1 + 0\cdot 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{3.72}$$

item

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3.73}$$

item

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3.74}$$

Example 3.2.2:

1.
$$\begin{bmatrix}1 & 0 & -3\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} = \begin{bmatrix}a_{11}-3a_{31} & a_{12}-3a_{32} & a_{13}-3a_{33}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} \tag{3.75}$$

2.
$$\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\2 & 0 & 1\end{bmatrix} = \begin{bmatrix}a_{11}+2a_{13} & a_{12} & a_{13}\\a_{21}+2a_{23} & a_{22} & a_{23}\\a_{31}+2a_{33} & a_{32} & a_{33}\end{bmatrix} \tag{3.76}$$

3.
$$\begin{bmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} = \begin{bmatrix}a_{11} & a_{12} & a_{13}\\2a_{21} & 2a_{22} & 2a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} \tag{3.77}$$

4.
$$\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 3 & 0\\0 & 0 & 1\end{bmatrix} = \begin{bmatrix}a_{11} & 3a_{12} & a_{13}\\a_{21} & 3a_{22} & a_{23}\\a_{31} & 3a_{32} & a_{33}\end{bmatrix} \tag{3.78}$$

5.
$$\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} = \begin{bmatrix}a_{31} & a_{32} & a_{33}\\a_{21} & a_{22} & a_{23}\\a_{11} & a_{12} & a_{13}\end{bmatrix} \tag{3.79}$$

6.
$$\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix} = \begin{bmatrix}a_{12} & a_{11} & a_{13}\\a_{22} & a_{21} & a_{23}\\a_{32} & a_{31} & a_{33}\end{bmatrix} \tag{3.80}$$

Exercise 3.2.1: Suppose $A = \begin{bmatrix}1 & 1 & 1\\2 & 2 & 1\\2 & 1 & 1\end{bmatrix}, B = \begin{bmatrix}1 & 1 & 1\\2 & 2 & 1\\2 & 1 & 1\end{bmatrix}$, computes matrix multiplication AB

Fact 3.2.1: Suppose A, B, C, D are matrices, c is a scalar, 0 is the zero matrix, I is the identity matrix. we have the following facts

- Matrix multiplication is generally *NOT commutative*, i.e. $AB \neq BA$
- Matrix multiplication is *associative*, i.e. the order of multiplication doesn't matter, in other words $(AB)C = A(BC)$, so it makes sense to write successive multiplication $A_1A_2A_3\cdots A_n$
- Scalar multiplication and matrix multiplication commutes, $A(cB) = c(AB) = (cA)B$. so it makes sense to write $cA_1A_2A_3\cdots A_n$
- Matrix multiplication is *distributive* over addition, i.e. $A(B+C) = AB+AC$, $(A+B)C = AC+BC$
- Zero matrix and identity matrix acts as 0 and 1, i.e. $A+0 = 0+A = A$, $A0 = 0A = 0$, $IA = AI = A$
- Even if $A \neq 0$, $B \neq 0$, AB could still be 0, take eqref{ eq:equation1} for an example
- $AB = AC$ does NOT imply $B = C$

Remark: Some of the properties of matrices are really similar to that of numbers, so we dub this the name of *matrix algebra*

3.3. Partitioned matrix

Definition 3.3.1: A is a partitioned (or block) matrix if is divided into smaller submatrix by some horizontal and vertical lines. And the submatrices are the blocks

$$\left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right] = \left[\begin{array}{cc|cccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} & a_{47} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} & a_{57} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} & a_{67} \end{array}\right] \tag{3.81}$$

Here the blocks are $A_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A_{12} = \begin{bmatrix} a_{13} & a_{14} & a_{15} & a_{16} \\ a_{23} & a_{24} & a_{25} & a_{26} \end{bmatrix}$, $A_{13} = \begin{bmatrix} a_{17} \\ a_{27} \end{bmatrix}$, $A_{21} = \begin{bmatrix} a_{31} & a_{32} \end{bmatrix}$,

$A_{22} = \begin{bmatrix} a_{33} & a_{34} & a_{35} & a_{36} \end{bmatrix}$, $A_{23} = \begin{bmatrix} a_{37} \end{bmatrix}$, $A_{31} = \begin{bmatrix} a_{41} & a_{42} \\ a_{51} & a_{52} \\ a_{61} & a_{62} \end{bmatrix}$, $A_{32} = \begin{bmatrix} a_{43} & a_{44} & a_{45} & a_{46} \\ a_{53} & a_{54} & a_{55} & a_{56} \\ a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix}$,

$A_{33} = \begin{bmatrix} a_{47} \\ a_{57} \\ a_{67} \end{bmatrix}$.

Fact 3.3.1: Suppose $A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix}$, $B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \cdots & B_{2r} \\ \vdots & \vdots & & \vdots \\ B_{q1} & B_{q2} & \cdots & B_{qr} \end{bmatrix}$ are partitioned matrices, and the number of columns of A_{1k} is equal to the number of rows of B_{k1} (so that all submatrices multiplications make sense). Then the usual row-column rule still WORKS!!! By treating submatrices as if they are numbers.

$$AB = C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1r} \\ C_{21} & C_{22} & \cdots & C_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ C_{p1} & C_{p2} & \cdots & C_{pr} \end{bmatrix}, C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{iq}B_{qj} \tag{3.82}$$

Example 3.3.1: Consider $\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 2 & 1 \\ \hline 2 & 1 & 1 \end{array}\right]$, $\left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] = \left[\begin{array}{c|cc} 1 & 1 & 1 \\ 2 & 2 & 1 \\ \hline 2 & 1 & 1 \end{array}\right]$, then

$$A_{11}B_{11} + A_{12}B_{21} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}[2] = \begin{bmatrix} 5 \\ 8 \end{bmatrix} \tag{3.83}$$

$$A_{11}B_{12} + A_{12}B_{22} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}[1 \ \ 1] = \begin{bmatrix} 4 & 3 \\ 7 & 5 \end{bmatrix} \tag{3.84}$$

$$A_{21}B_{11} + A_{22}B_{21} = [2 \ \ 1]\begin{bmatrix} 1 \\ 2 \end{bmatrix} + [1][2] = [6] \tag{3.85}$$

$$A_{21}B_{12} + A_{22}B_{22} = [2 \ \ 1]\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} + [1][1 \ \ 1] = [5 \ \ 4] \tag{3.86}$$

$$\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right]\left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right] = \left[\begin{array}{cc|c} 5 & 4 & 3 \\ 8 & 7 & 5 \\ \hline 6 & 5 & 4 \end{array}\right] \tag{3.87}$$

Example 3.3.2: Suppose 3×3 matrix A can be partitioned into $[\mathrm{R1R2R3}]$ or $[\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]$, then Example Example 3.2.2 reads

- $E = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $EA = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}[\mathrm{R1R2R3}] = [\mathrm{R1{-}3R3R2R3}]$, EA acts as subtracting 3 times row 3 from row 1.
- $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, $AE = [\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} = [\boldsymbol{a}_1{+}2\boldsymbol{a}_3 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]$, AE acts as adding 2 times column 3 to column 1.
- $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $EA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}[\mathrm{R1R2R3}] = [\mathrm{R12R2R3}]$, EA acts as scaling the third row by 2.
- $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $AE = [\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\boldsymbol{a}_1 \ \ 3\boldsymbol{a}_2 \ \ \boldsymbol{a}_3]$, AE acts as scaling the third column by 3.
- $E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $EA = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}[\mathrm{R1R2R3}] = [\mathrm{R3R2R1}]$, EA acts as interchanging row 1 and row 3.
- $E = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $AE = [\boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \boldsymbol{a}_3]\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [\boldsymbol{a}_2 \ \ \boldsymbol{a}_1 \ \ \boldsymbol{a}_3]$, AE acts as interchanging column 1 and column 2.

Definition 3.3.2: Matrices E in the previous example is called elementary matrices. They describe row and column elementary operations.

Exercise 3.3.1: If A is a 4 by 5 matrix, what is the elementary matrix E that acts as replacing the fourth row by adding twice of the second row.

Exercise 3.3.2: Suppose $B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{bmatrix}$, show that $AB = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{bmatrix}$

Exercise 3.3.3: Verify that $A^2 = I_2$ where $A = \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}$, and use partitioned matrices to show that $M^2 = I_4$, where $M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -3 & 1 \end{bmatrix}$

Exercise 3.3.4: Suppose $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} \sim \begin{bmatrix} U & \boldsymbol{d} \end{bmatrix}$ is the REF/RREF, then U will also be the REF/RREF of A.

Solution: Suppose the row elementary operations applied are $E_1, E_2, \cdots, E_k$, then

$$\begin{array}{ccccccc} [A\ \ \boldsymbol{b}] \sim & E_1[A\ \ \boldsymbol{b}] & \sim & E_2E_1[A\ \ \boldsymbol{b}] & \sim \cdots \sim & E_k\cdots E_2E_1[A\ \ \boldsymbol{b}] & = [U\ \ \boldsymbol{d}] \\ & \| & & \| & & \| & \\ & [E_1A\ \ E_1\boldsymbol{b}] & & [E_2E_1A\ \ E_2E_1\boldsymbol{b}] & & [E_k\cdots E_2E_1A\ \ E_k\cdots E_2E_1\boldsymbol{b}] & \end{array}$$

The same sequence of row elementary operations would reduce A to U.

4. Lecture 4 - Matrix equations and linear independence

4.1. Vector and matrix equations

Recall a vector is a matrix with one column, the zero vector is a vector with all entries zero. For scalar (i.e. a number) c, and vectors $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, we have

- Addition $\boldsymbol{a} + \boldsymbol{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{bmatrix}$
- Scalar multiplication $c\boldsymbol{a} = c\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{bmatrix}$
- Subtraction $\boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a} + (-1)\boldsymbol{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1-b_1 \\ a_2-b_2 \\ \vdots \\ a_n-b_n \end{bmatrix}$

Remark: In handwritings, we use $\vec{v}$ or $\vec{v}$ to denote a vector, while in printing materials we often use the math bold font $\boldsymbol{v}$.

Definition 4.1.1: A vector equation is of the form

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b} \tag{4.88}$$

We can also write the vector equation eqref{eq:veceq} as a matrix equation with partitioned matrix

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} = A\boldsymbol{x} = \boldsymbol{b} \tag{4.89}$$

Here $A = \begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n\end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$

Example 4.1.1: (4.4) can be written as a vector equation

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 = x_1\begin{bmatrix}1\\2\end{bmatrix} + x_2\begin{bmatrix}1\\4\end{bmatrix} = \begin{bmatrix}x_1\\2x_1\end{bmatrix} + \begin{bmatrix}x_2\\4x_2\end{bmatrix} = \begin{bmatrix}x_1+x_2\\2x_1+4x_2\end{bmatrix} = \begin{bmatrix}10\\26\end{bmatrix} = \boldsymbol{b} \tag{4.90}$$

Or a matrix equation

$$A\boldsymbol{x} = \begin{bmatrix}1 & 1\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}10\\26\end{bmatrix} = \boldsymbol{b} \tag{4.91}$$

Example 4.1.2: The corresponding vector equation of

$$\begin{cases}x_1 + x_3 = 1\\2x_2 + x_3 = 2\end{cases} \tag{4.92}$$

is

$$x_1\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\2\end{bmatrix} + x_3\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix} \tag{4.93}$$

And the corresponding matrix equation is

$$\begin{bmatrix}1 & 0 & 1\\0 & 2 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix} \tag{4.94}$$

%

Question: % Suppose A is a $m \times n$ matrix, when does the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ always has a solution for any $\boldsymbol{b}$ in $\mathbb{R}^n$ %

4.2. Span

Definition 4.2.1:

- A linear combination of $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ is a sum $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n$ for some scalars $c_1, \cdots, c_n$.
- The span of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ is the set of all its linear combinations, which we denote $\mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$.

Theorem 4.2.1: $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$ has solution(s) $\Leftrightarrow$ $\boldsymbol{b}$ is a linear combination of $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ $\Leftrightarrow$ $\boldsymbol{b} \in \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n\}$

Exercise 4.2.1: Let $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Is $\boldsymbol{b}$ in $\mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$?

Solution: This is equivalent of asking if whether the vector equation $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 = \boldsymbol{b}$ has solution(s), we find an REF of its augmented matrix

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ 2 & 0 & -1 & 1 \\ 1 & 1 & -2 & 1 \end{bmatrix} \xrightarrow[\text{R3}\to\text{R3}-\text{R1}]{\text{R2}\to\text{R2}-2\text{R1}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 2 & -3 & 0 \end{bmatrix} \xrightarrow{\text{R3}\to\text{R3}-\text{R2}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.95}$$

Since there is a pivot in the last column, by Theorem Theorem 2.2.2, the linear system is inconsistent, hence $\boldsymbol{b} \notin \mathrm{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$

4.3. Linear independence

Definition 4.3.1: $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is linearly dependent if some $\boldsymbol{v}_i$ is in the span of the others (so it is somewhat redundant), or equivalently, if there is a non-trivial solution $c_1, \cdots, c_n$ (i.e. not all c_i's are 0) to the vector equation

$$c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0} \tag{4.96}$$

(4.96) is referred to as a linear dependence between $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$. If (4.96) has only the trivial solution (i.e. $c_1, \cdots, c_n$ are all 0, which is of course always a solution), $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is said to be linearly independent

Remark: Equivalence between two different definitions of linear dependence

- If $\boldsymbol{v}_i = c_1\boldsymbol{v}_1 + \cdots + c_{\{i-1\}}\boldsymbol{v}_{\{i-1\}} + c_{\{i+1\}}\boldsymbol{v}_{\{i+1\}} + \cdots + c_n\boldsymbol{v}_n$, then $c_1\boldsymbol{v}_1 + \cdots + (-1)\boldsymbol{v}_i + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$
- If $c_1\boldsymbol{v}_1 + \cdots + c_i\boldsymbol{v}_i + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$ and $c_i \neq 0$ (since not all c_i's are zero, we may assume some c_i is nonzero), then $\boldsymbol{v}_i = -\frac{c_1}{c_i}\boldsymbol{v}_1 - \cdots - \frac{c_{i-1}}{c_i}\boldsymbol{v}_{\{i-1\}} - \frac{c_{i+1}}{c_i}\boldsymbol{v}_{\{i+1\}} - \cdots - \frac{c_n}{c_i}\boldsymbol{v}_n$

Question: How do we determine and find linear dependence of $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$?

Answer: Let $A = [\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_n]$, then non-trivial solutions to $A\boldsymbol{x} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n = \mathbf{0}$ would be the linear dependences of $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$. Therefore it is linearly independent if it has only the trivial(zero) solution.

Theorem 4.3.1: $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is linearly independent $\Leftrightarrow$ $A\boldsymbol{x} = \mathbf{0}$ has only the trivial(zero) solution $\Leftrightarrow$ each column of the RREF of A is a pivot column.

Proof: Consider the RREF of the augmented matrix $[A \ \mathbf{0}]$, it is necessarily $[U \ \mathbf{0}]$ for $A\boldsymbol{x} = \mathbf{0}$ to have only the trivial solution. ■

Example 4.3.1: Suppose $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_1\}$ is linearly dependent since

$$[\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \boldsymbol{e}_1] = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{bmatrix} \tag{4.97}$$

The solution to this augmented matrix would be $\begin{cases} x_1 = -\frac{1}{2}x_3 \\ x_2 = -\frac{1}{2}x_3 \\ x_3 \text{ is free} \end{cases}$, by choosing any value nonzero value of x_3 (say 1) we get a linear dependence $-\frac{1}{2}\boldsymbol{v}_1 - \frac{1}{2}\boldsymbol{v}_2 + \boldsymbol{e}_1 = \mathbf{0}$. On the other hand, $\boldsymbol{v}_1, \boldsymbol{v}_2$ are linearly independent since

$$[\boldsymbol{v}_1 \ \ \boldsymbol{v}_2] = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.98}$$

Where each column is a pivot column.

Exercise 4.3.1: Write the system

$$\begin{cases} 8x_1 - x_2 = 4 \\ 5x_1 + 4x_2 = 1 \\ x_1 - 3x_2 = 2 \end{cases} \tag{4.99}$$

first as a vector equation and then as a matrix equation.

Exercise 4.3.2: Let $\boldsymbol{v}_1 = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 8 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} 4 \\ -1 \\ -5 \end{bmatrix}$.

- Does $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ span $\mathbb{R}^3$? Why or why not?
- Is $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ linearly independent? Why or why not?

5. Lecture 5 - Geometric interpretation of solutions of linear systems

5.1. Geometric interpretation of vectors

We like to identify vectors in $M_{\{n\times 1\}}(\mathbb{R})$ with points in $\mathbb{R}^n$. And there are very nice geometric interpretation of vector additions and scalar multiplications.

Example 5.1.1: Let $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, then $\boldsymbol{a}+\boldsymbol{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $2\boldsymbol{a} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $-\boldsymbol{b} = (-1)\boldsymbol{b} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$, $\boldsymbol{a}-\boldsymbol{b} = \boldsymbol{a}+(-\boldsymbol{b}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

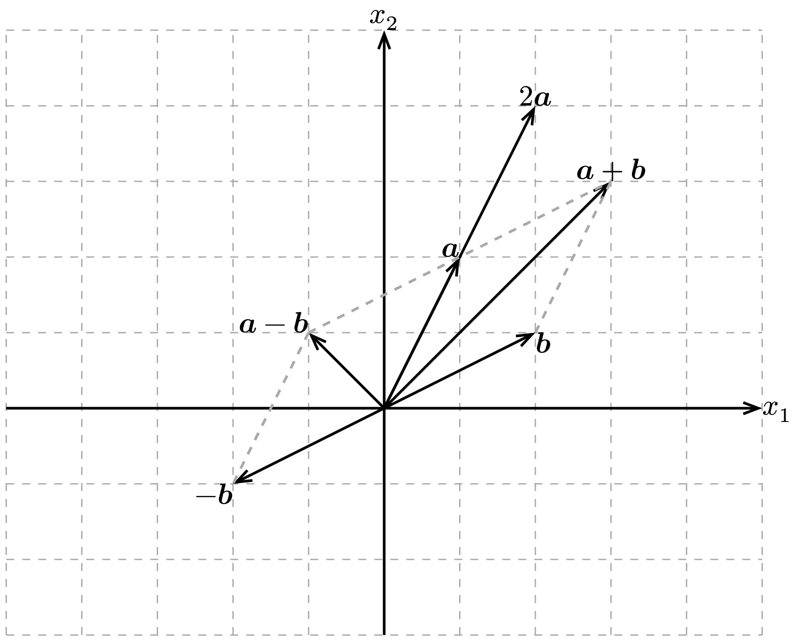


Figure 5: Vector-point correspondence when $n = 2$

5.2. Basis

Theorem 5.2.1: Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

- For each $\boldsymbol{b}$ in $\mathbb{R}^m$, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution.
- Each $\boldsymbol{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
- The columns of A span $\mathbb{R}^m$.
- A has a pivot position in every row. (Equivalently, in the last row)

Definition 5.2.1: $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ is said to be a basis for $\mathbb{R}^n$ if it is linearly independent and spans all of $\mathbb{R}^n$

Theorem 5.2.2: Let $A = [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \cdots \ \boldsymbol{v}_n]$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n\}$ forms basis of $\mathbb{R}^n \Leftrightarrow A \sim I_n$, in other words, each row and each column of A has a pivot.

Definition 5.2.2: The standard basis for $\mathbb{R}^n$ is the set of vectors $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$, where

$$\boldsymbol{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \longleftarrow j\text{-th entry}$$

Example 5.2.1: $\left\{ \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is the standard basis for $\mathbb{R}^3$, and

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3 \end{aligned} \tag{5.100}$$

5.3. Geometric meaning of spans

Example 5.3.1: Consider Example Example 4.3.1 where $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_1\}$ and Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ are both the plane $\mathbb{R}^2$, $\boldsymbol{e}_1$ is in the span of $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ because $\boldsymbol{e}_1 = \frac{1}{2}\boldsymbol{v}_1 + \frac{1}{2}\boldsymbol{v}_2$. The gray grids illustrate the span of $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ and the purple grids illustrate the span of $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$.

Exercise 5.3.1: Suppose $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

- Determine whether $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ forms a basis for $\mathbb{R}^3$.
- Without performing elementary row operations, how many solutions does $A\boldsymbol{x} = \boldsymbol{b}$ have? Where $\boldsymbol{b} = \begin{bmatrix} 1 \\ 45 \\ -9 \end{bmatrix}$, what about $\boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

5.4. Parametric vector form

Example 5.4.1: In Example 2.2.4, the solution set can be written as parametric vector form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}x_3 + \frac{1}{2} \\ \frac{3}{2}x_3 - \frac{1}{2} \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}x_3 \\ \frac{3}{2}x_3 \\ x_3 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} = x_3 \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} \tag{5.101}$$

In Example Exercise 2.2.1, the solution set can be written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2x_2 + 16 \\ x_2 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 16 \\ 0 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 16 \\ 0 \\ \frac{5}{2} \\ -\frac{9}{2} \end{bmatrix} \tag{5.102}$$

Exercise 5.4.1: Suppose the augmented matrix of a linear system is equivalent to the following matrix

$$\begin{bmatrix} 1 & 1 & 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \tag{5.103}$$

Write down the solution set in parametric vector form

Solution:

$$\begin{cases} x_1 + x_2 + 2x_4 = 3 \\ x_3 - 2x_4 = 2 \\ x_5 = 1 \end{cases} \Rightarrow \begin{cases} x_1 = 3 - x_2 - 2x_4 \\ x_2 \text{ is free} \\ x_3 = 2 + 2x_4 \\ x_4 \text{ is free} \\ x_5 = 1 \end{cases} \tag{5.104}$$

So the solution in parametric vector form would be

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3 - x_2 - 2x_4 \\ x_2 \\ 2 + 2x_4 \\ x_4 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -x_2 \\ x_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_4 \\ 0 \\ 2x_4 \\ x_4 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} \tag{5.105}$$

5.5. Geometric interpretation of solution set to linear system

Definition 5.5.1: A linear system is homogeneous if it has matrix equation $A\boldsymbol{x} = \mathbf{0}$ (note that this always have the zero solution, called the trivial solution).

Theorem 5.5.1: Suppose $[A \;\; b] \sim [U \;\; d]$ is the RREF, then U will be the RREF of A. The solutions of $A\boldsymbol{x} = \mathbf{0}$ and $A\boldsymbol{x} = \boldsymbol{b}$ differs by $\tilde{\boldsymbol{d}}$ ($\tilde{\boldsymbol{d}}$ is $\boldsymbol{d}$ with 0's inserted at the free variable positions), i.e.

$$\tilde{\boldsymbol{d}} + \{\text{solutions of } A\boldsymbol{x} = \mathbf{0}\} = \{\text{solutions of } A\boldsymbol{x} = \boldsymbol{b}\} \tag{5.106}$$

Geometrically speaking, the solution set of $A\boldsymbol{x} = \boldsymbol{b}$ is the hyperplane translated from the hyperplane through the origin (solution set of $A\boldsymbol{x} = \mathbf{0}$) by $\boldsymbol{d}$.

Example 5.5.1: $x_1 - 2x_2 = 2$ has augmented matrix $[1 \;\; -2 \;\; 2]$ which is already an RREF, which has solution

$$\begin{cases} x_1 = 2 + 2x_2 \\ x_2 \text{ is free} \end{cases} \Rightarrow \begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{5.107}$$

Its corresponding homogeneous equation $x_1 - 2x_2 = 0$ has solution

$$\begin{cases} x_1 = 2x_2 \\ x_2 \text{ is free} \end{cases} \Rightarrow x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{5.108}$$

Example 5.5.2: Consider homogeneous linear system $3x_1 + x_2 - x_3 = 0$, and non-homogeneous linear system $3x_1 + x_2 - x_3 = 3$ .The parametric vector form of the solution sets to both are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{1}{3} \\ 0 \\ 1 \end{bmatrix} \tag{5.109}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{1}{3} \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{5.110}$$

6. Lecture 6 - Linear transformations

6.1. Linear transformations and matrix transformations

Definition 6.1.1: A Linear transformation (or linear mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a mapping satisfying

- $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ for any $\boldsymbol{u}, \boldsymbol{v}$ in $\mathbb{R}^n$
- $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$ for any scalar c and any $\boldsymbol{u}$ in $\mathbb{R}^n$

Example 6.1.1: Reflection, rotation and scaling are all linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Exercise 6.1.1: Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y \end{bmatrix}$. Is T is a linear mapping?

Solution: Suppose $\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, then

item

$$T(\boldsymbol{u}+\boldsymbol{v}), = T\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right) = \begin{bmatrix} (u_1+v_1)+(u_2+v_2) \\ u_2+v_2 \end{bmatrix};, = \begin{bmatrix} (u_1+u_2)+(v_1+v_2) \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} u_1+u_2 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1+v_2 \\ v_2 \end{bmatrix} = T(\boldsymbol{u}) + T(\boldsymbol{v}) \tag{6.111}$$

item

$$T(c\boldsymbol{u}) = T\left(\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}\right) = \begin{bmatrix} cu_1+cu_2 \\ cu_2 \end{bmatrix} = c\begin{bmatrix} u_1+u_2 \\ u_2 \end{bmatrix} = cT(\boldsymbol{u}) \tag{6.112}$$

Definition 6.1.2: A matrix transformation is the mapping defined via matrix multiplication, i.e. $T(\boldsymbol{x}) = A\boldsymbol{x}$ for some $m \times n$ matrix A. It is a linear transformation thanks to Fact Fact 3.2.1 c),d) since

- $T(\boldsymbol{u}+\boldsymbol{v}) = A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u}+A\boldsymbol{v} = T(\boldsymbol{u})+T(\boldsymbol{v})$
- $T(c\boldsymbol{u}) = A(c\boldsymbol{u}) = c(A\boldsymbol{u}) = cT(\boldsymbol{u})$

Example 6.1.2: In fact, T in Exercise Exercise 6.1.1 is a matrix transformation

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1+x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\boldsymbol{x} \tag{6.113}$$

Definition 6.1.3: In general, if $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ is the standard basis, then any $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$, and

$$T(\boldsymbol{x}) = x_1T(\boldsymbol{e}_1) + \cdots + x_nT(\boldsymbol{e}_n) = [T(\boldsymbol{e}_1) \ \cdots \ T(\boldsymbol{e}_n)]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{6.114}$$

Denote $A = [T(\boldsymbol{e}_1) \ \cdots \ T(\boldsymbol{e}_n)]$ (which is called the standard matrix for the linear transformation T), then $T(\boldsymbol{x}) = A\boldsymbol{x}$, so every linear transformation T from $\mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation.

Example 6.1.3: In Exercise Exercise 6.1.1

$$T(\boldsymbol{e}_1) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1+0\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}, \quad T(\boldsymbol{e}_2) = T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0+1\\1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} \tag{6.115}$$

Therefore the standard matrix for T is $\begin{bmatrix}1&1\\0&1\end{bmatrix}$.

Exercise 6.1.2: Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}x_1+x_2+1\\x_2\end{bmatrix}$. Is T is a linear mapping?

Solution: T is not a linear transformation since $T(\boldsymbol{x}) = \begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{p}$ is not a matrix transformation (6.114) ($\boldsymbol{p} \neq 0$)

Example 6.1.4: Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation

- Suppose T is the reflection over x_2-axis, then the standard matrix for T is

$$A = [T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)] = \begin{bmatrix}-1&0\\0&1\end{bmatrix} \tag{6.116}$$

- Suppose T is the rotation by 60° counter-clockwise, then the standard matrix for T is

$$A = [T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)] = \begin{bmatrix}\frac{1}{2}&-\frac{\sqrt{3}}{2}\\\frac{\sqrt{3}}{2}&\frac{1}{2}\end{bmatrix} \tag{6.117}$$

Exercise 6.1.3: $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation that rotate 60° counter-clockwise and then reflects over x_2-axis, what is its standard matrix?

Solution: The standard matrix for T is

$$A = [T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)] = \begin{bmatrix}-\frac{1}{2}&\frac{\sqrt{3}}{2}\\\frac{\sqrt{3}}{2}&\frac{1}{2}\end{bmatrix} \tag{6.118}$$

6.2. Properties of linear transformations

Definition 6.2.1: Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a mapping. We call

- $\mathbb{R}^n$ the domain of T
- $\mathbb{R}^m$ the codomain of T
- $T(\boldsymbol{x})$ the image of $\boldsymbol{x}$ under T
- $T^{-1}(\boldsymbol{b}) = \boldsymbol{x} | T(\boldsymbol{x}) = \boldsymbol{b}$ the set of preimages of $\boldsymbol{b}$ under T
- the set of images $\{T\boldsymbol{x} | \boldsymbol{x} \in \mathbb{R}^n\}$ the range of T

Exercise 6.2.1: Suppose the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$, what is the standard matrix of T? What is the image of $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$, what is the set of vectors with image $\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$, what is the range?

Solution: The standard matrix is $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, the image $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ under T is

$$T\left(\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \tag{6.119}$$

The set of vectors with image $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ under T is the solution set to $T(\boldsymbol{x}) = A\boldsymbol{x} = \boldsymbol{b}$ (this is Example Example 2.1.3), which is $\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$. And since there is a pivot in each row, by Theorem Theorem 5.2.1, the range of T is $\mathbb{R}^3$

Definition 6.2.2: A mapping T is said to be onto $\mathbb{R}^m$ if each $b \in \mathbb{R}^m$ is the image of at least one $x \in \mathbb{R}^n$.

Codomain is larger than the range if T is not onto

Definition 6.2.3: A mapping T is said to be one-to-one if each $b \in \mathbb{R}^m$ is the image of at most one $x \in \mathbb{R}^n$.

Theorem 6.2.1: Suppose A is the standard matrix for linear transformation T (i.e. $T(\boldsymbol{x}) = A\boldsymbol{x}$), then

- T is one-to-one $\Leftrightarrow$ $A\boldsymbol{x} = \boldsymbol{b}$ has at most one solution $\Leftrightarrow$ $A\boldsymbol{x} = \boldsymbol{0}$ has the unique trivial solution $\Leftrightarrow$ RREF of A has a pivot in each column $\Leftrightarrow$ columns of A are linearly independent.
- T is onto $\Leftrightarrow$ $A\boldsymbol{x} = \boldsymbol{b}$ always has solution $\Leftrightarrow$ the columns of A span $\mathbb{R}^m$ $\Leftrightarrow$ RREF of A has a pivot in each row.

Exercise 6.2.2: Suppose the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ 2x_1 - x_3 \end{bmatrix}$, Is T onto? Is T one-to-one?

Solution: This is the Example Example 2.2.4. The standard matrix for T is $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \end{bmatrix} \xrightarrow{\text{R2}\to\text{R2}-2\text{R1}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \end{bmatrix}$, since there is a pivot in each row but not in each column, it is onto but not one-to-one

Exercise 6.2.3:
- If $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation, could it be one-to-one?
- If $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation, could it be onto?

Solution: Both no! Due to Theorem Theorem 6.2.1

- Since A is a 2×3 matrix, there will be at most 2 pivots (only 2 rows), so there won't be enough pivots to fill all columns.
- Since A is a 3×2 matrix, there will be at most 2 pivots (only 2 columns), so there won't be enough pivots to fill all rows.

Exercise 6.2.4: Suppose $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 + x_3 \\ 2x_1 - x_3 \\ x_1 + x_2 - 2x_3 \end{bmatrix}$.

- What is the domain of T?
- What is the codomain of T?
- What is the standard matrix of T?
- What is the image of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$?
- What is the set of vectors with image being $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$?
- What is the set of vectors with image being $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$?
- Is T onto?
- Is T one-to-one?

6.3. Composition of linear transformations

Definition 6.3.1: Suppose

- $T_1 : \mathbb{R}^p \to \mathbb{R}^n$ is a linear transformation with standard matrix A_1 (which should be $n \times p$)
- $T_2 : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix A_2 (which should be $m \times n$)

Define the composition $T_2 \circ T_1$ of T_1 and T_2 as $(T_2 \circ T_1)(\boldsymbol{x}) = T_2(T_1(\boldsymbol{x}))$ begin{center} begin{tikzcd} RR^p arrow[r, “T_1”] arrow[rr, “T_2circle.small T_1”‘, bend right] , RR^n arrow[r, “T_2”] , RR^m end{tikzcd} end{center} Then $T_2 \circ T_1 : \mathbb{R}^p \to \mathbb{R}^m$ is also a linear transformation (Why? Verify this). For $\boldsymbol{x} \in \mathbb{R}^p$,

$$(T_2 \circ T_1)(\boldsymbol{x}) = T_2(T_1(\boldsymbol{x})) = A_2(T_1(\boldsymbol{x})) = A_2(A_1 \boldsymbol{x}) = (A_2 A_1)\boldsymbol{x} \tag{6.120}$$

So we have concluded that the standard matrix for $T_2 \circ T_1$ is the $m \times p$ matrix $A_2 A_1$.

Note: You should compose maps to the left.

Example 6.3.1: Consider Example Example 6.1.4., If we let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ to denote the rotation by $60°$ counter-clockwise, $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ to denote reflection over x_2-axis, and their standard matrices are

$$\begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6.121}$$

Then look at Exercise Exercise 6.1.3, this is the composition $T_2 \circ T_1$, which has the standard matrix

$$A_2 A_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \tag{6.122}$$

Question: Suppose $A = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$ is the standard matrix for the linear transformation of rotating $60°$ counter-clockwise (Example Example 6.1.4). What is A^7?

Answer: $A^7 = AAAAAAA$ is the standard matrix for composition of linear transformations $T \circ T \circ T \circ T \circ T \circ T \circ T$ which is rotate $7 \times 60° = 420°$, but that is the same as rotating $420° - 360° = 60°$ which is the same linear transformation as T, so $A^7 = A$

Exercise 6.3.1: Let T be the linear transformation that rotate $\mathbb{R}^2$ counter-clockwise of angle θ, find the standard matrix for T. What about the standard matrix for $T^{\circ 100}$

7. Lecture 7 - Matrix transpose and matrix inverse

7.1. Matrix transpose

Definition 7.1.1: Suppose $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ is a $m \times n$ matrix, we define its transpose by flipping it over the diagonal, and this is the $n \times m$ matrix

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{mn} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \tag{7.123}$$

Example 7.1.1: Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

Theorem 7.1.1: Here are some properties of matrix transpose

- $\left(A^T\right)^T = A$
- $(A + B)^T = A^T + B^T$
- $(cA)^T = cA^T$
- $(AB)^T = B^T A^T$

Definition 7.1.2: For any $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \in \mathbb{R}^n$, we can define the dot product to be $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^T \boldsymbol{w} = v_1 w_1 + \cdots + v_n w_n$. $|\boldsymbol{v}| = \sqrt{\boldsymbol{v}^T \boldsymbol{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$ is the length $\boldsymbol{v}$

Remark: There is a nice geometric interpretation of dot product. Suppose the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$ is θ, then $\boldsymbol{v} \cdot \boldsymbol{w} = |\boldsymbol{v}| \|\boldsymbol{w}| \cos \theta$. % With a bit of trigonometry, we see it is the product of one vector and the projection of the other onto the first one. % Here $\boldsymbol{u}$ is the projection of $\boldsymbol{v}$ onto $\boldsymbol{w}$

Exercise 7.1.1: Let $\boldsymbol{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

- What is the length of $\boldsymbol{v}$?
- What is $\boldsymbol{v} \cdot \boldsymbol{w}$?
- what is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$?

Question: Have you wondered about how linear algebra would look like if we make the identification of row vectors $M_{\{1\times n\}}(\mathbb{R}^n)$ and $\mathbb{R}^n$ instead of column vectors.

Answer: The standard basis for $\mathbb{R}^n$ would be $\{\boldsymbol{e}_1^T, \boldsymbol{e}_2^T, \cdots, \boldsymbol{e}_n^T\}$, where $\boldsymbol{e}_j^T = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}$, where every entry is 0 except the j-th entry being 1. For any $\boldsymbol{x}^T \in \mathbb{R}^n$ and any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, we have

$$= T(\boldsymbol{x})^T = (x_1 T(\boldsymbol{e}_1) + x_2 T(\boldsymbol{e}_2) + \cdots + x_n T(\boldsymbol{e}_n))^T ;, = x_1 T(\boldsymbol{e}_1)^T + x_2 T(\boldsymbol{e}_2)^T + \cdots + x_n T(\boldsymbol{e}_n)^T ;, = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} T(\boldsymbol{e}_1)^T \\ T(\boldsymbol{e}_2)^T \\ \vdots \\ T(\boldsymbol{e}_n)^T \end{bmatrix} = \tag{7.124}$$

Here $A = \begin{bmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_n) \end{bmatrix}$.

If you consider composition $T_2 \circ T_1$ where T_1, T_2 are linear transformations with standard matrices A, B, you would get

$$(T_2 \circ T_1)(\boldsymbol{x}^T) = T_2(T_1(\boldsymbol{x}^T)) = T_1(\boldsymbol{x}^T) B^T = (\boldsymbol{x}^T A^T) B^T = \boldsymbol{x}^{T(A^T B^T)} \tag{7.125}$$

On the other hand, we should know that $(T_2 \circ T_1)(\boldsymbol{x}^T) = \boldsymbol{x}^{T(A^T B^T)}$, so this implies $A^T B^T = (BA)^T$.

7.2. Matrix inverse

Definition 7.2.1: Suppose linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is both onto and one-to-one (i.e. for each vector $\boldsymbol{b}$ in the codomain $\mathbb{R}^m$ there is a unique preimage, which we denote as $T^{-1}(\boldsymbol{b})$). Suppose A is the standard matrix for T, then m necessarily equal n as shown in Exercise Exercise 6.2.3, so A must be a square matrix. We know $T(\boldsymbol{x}) = \boldsymbol{b}$ always has a unique solution which should be $T^{-1}(\boldsymbol{b})$, it can be shown that $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ as mapping is actually also a linear transformation (Why? See if you can figure this out). Then the standard matrix of T^{-1} is defined to be A^{-1} (the inverse matrix of A). Note that

$$, (T \circ T^{-1})(\boldsymbol{b}) = T(T^{-1}(\boldsymbol{b})) = T(\boldsymbol{x}) = \boldsymbol{b};, (T^{-1} \circ T)(\boldsymbol{x}) = T^{-1}(T(\boldsymbol{x})) = T^{-1}(\boldsymbol{b}) = \boldsymbol{x} \tag{7.126}$$

Since $T \circ T^{-1}$, $T^{-1} \circ T$ work like the identity mappings, so $AA^{-1} = A^{-1}A = I$. In this case, we see that A is equivalent to the identity matrix (because of Theorem Example 4.3.1, A has a pivot in each row and column).

Remark: Because we can write elementary row operations as left elementary matrix multiplications, so we know there are elementary matrices $E_1, E_2, \cdots, E_k$ such that

$$A, \sim E_1 A \sim E_2 E_1 A \sim E_3 E_2 E_1 A \sim \cdots;, \sim E_k E_{k-1} \cdots E_2 E_1 A = I \tag{7.127}$$

If we multiply A^{-1} on the right on both sides, we get $E_k E_{k-1} \cdots E_2 E_1 = A^{-1}$. Thanks to this observation, we introduce an algorithm for computing matrix inverses. Let's consider the RREF of the following partitioned matrix

$$\begin{aligned} [A \,|\, I] &\sim [E_1 A \,|\, E_1] \sim [E_2 E_1 A \,|\, E_2 E_1] \sim \cdots \\ &\sim [E_k E_{k-1} \cdots E_2 E_1 A \,|\, E_k E_{k-1} \cdots E_2 E_1] = [I \,|\, A^{-1}] \end{aligned} \tag{7.128}$$

Exercise 7.2.1: Find the inverse of the following matrices. begin{tasks}(3) task $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ task $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ task $\begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ end{tasks}

Solution:

- $$\left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{\text{R1}\to(-1)\text{R1}} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \tag{7.129}$$

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

- $$,\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right] \xrightarrow{\text{R2}\to\text{R2}-3\text{R1}} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -1 & -3 & 1 \end{array}\right];, \xrightarrow{\text{R1}\to\text{R1}+2\text{R2}} \left[\begin{array}{ccc|c} 1 & 0 & -5 & 2 \\ 0 & -1 & -3 & 1 \end{array}\right] \xrightarrow{\text{R2}\to(-1)\text{R2}} \left[\begin{array}{ccc|c} 1 & 0 & -5 & 2 \\ 0 & 1 & 3 & -1 \end{array}\right] \tag{7.130}$$

Hence $A^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$

- $$\begin{aligned} &\left[\begin{array}{cccccc|c} 1 & -1 & 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\text{R3}\to\text{R3}-\text{R1}]{\text{R2}\to\text{R2}-2\text{R1}} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 2 & 0 & -1 & 0 & 1 \end{array}\right] \\ &\xrightarrow{\text{R3}\to\text{R3}-\text{R2}} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -3 & -2 & 1 & 0 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array}\right] \xrightarrow{\text{R2}\to\text{R2}+\text{R3}} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 1 \\ 0 & 0 & 3 & 1 & -1 & 1 \end{array}\right] \\ &\xrightarrow[\text{R3}\to\frac{\text{R3}}{3}]{\text{R2}\to\frac{\text{R2}}{2}} \left[\begin{array}{ccccc|c} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}\right] \xrightarrow{\text{R1}\to\text{R1}+\text{R2}-\text{R3}} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{array}\right] \end{aligned} \tag{7.131}$$

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

7.3. Properties of matrix transposes and inverses

Definition 7.3.1: A square matrix A is invertible (or non-singular) if it has an inverse A^{-1} such that $AA^{-1} = A^{-1}A = I$. A is called singular if A is not invertible.

Theorem 7.3.1: Suppose T is a linear transformation with standard matrix A, then

$$T \text{ is invertible with inverse } T^{-1} \Leftrightarrow A \text{ is invertible with inverse } A^{-1} \Leftrightarrow A \sim I \tag{7.132}$$

Theorem 7.3.2: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, here $\det A = ad - bc$

Example 7.3.1: If $A = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$, then

$$A^{-1} = \texttt{dfrac\{1\}}\left\{\frac{1}{2}\frac{1}{2} - \frac{\sqrt{3}}{2}\left(-\frac{\sqrt{3}}{2}\right)\right\} \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \tag{7.133}$$

Theorem 7.3.3: If A is invertible, then the linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$

Proof: Multiply A^{-1} on the left on both sides ■

Example 7.3.2: Let's consider (7.4), in which case $A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 10 \\ 26 \end{bmatrix}$, then $A^{-1} = \frac{1}{1 \cdot 4 - 1 \cdot 2}\begin{bmatrix} 4 & -1 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix}$, and

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 10 \\ 26 \end{bmatrix} = \begin{bmatrix} 2 \cdot 10 - \frac{1}{2} \cdot 26 \\ -10 + \frac{1}{2} \cdot 26 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix} \tag{7.134}$$

Theorem 7.3.4: Here are some properties of matrix inverse

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

Exercise 7.3.1: What is $(A^T)^{-1}$ in Exercise Exercise 7.2.1, c)?

Solution: Use Theorem Theorem 7.3.4, we know

$$(A^T)^{-1} = (A^{-1})^T = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}^T = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & 0 & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \end{bmatrix} \tag{7.135}$$

Exercise 7.3.2: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with standard matrix A.

- If A is invertible, then A has n pivots. ✓
- If T is one-to-one, then A is invertible. ✓
- If columns of A span $\mathbb{R}^n$, then A is invertible. ✓
- If A is invertible, $A\boldsymbol{x} = \mathbf{0}$ only has the trivial solution. ✓
- If T is onto, then T is one-to-one. ✓
- If T is one-to-one, then T is onto. ✓

Exercise 7.3.3: Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation with standard matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. Is A^{-1} invertible? Is T one-to-one? Is A^T invertible? If so, what is $\left(A^T\right)^{-1}$? If so, what is $\left(A^{-1}\right)^{-1}$. Is T invertible (i.e. does T^{-1} exist)? What is the standard matrix of T^{-1}? Is T onto?

Solution:

$$\begin{aligned} [A\,|\,I] = &\left[\begin{array}{ccc|cc|c} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\text{R2}\to\text{R2}+2\text{R3}]{\text{R1}\to\text{R1}-3\text{R3}} \left[\begin{array}{ccc|cc|c} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & -1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \\ &\xrightarrow{\text{R2}\to(-1)\text{R2}} \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{R1}\to\text{R1}-2\text{R2}} \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] = [I\,|\,A^{-1}] \end{aligned} \tag{7.136}$$

So $A^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. By Theorem Theorem 7.3.4, we know

$$\left(A^{-1}\right)^{-1} = A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.137}$$

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \tag{7.138}$$

Therefore we know A^{-1} and A^T are invertible.
In general, if T is invertible, then A is invertible, so A^{-1} will be the standard matrix for T^{-1} as $T^{-1}(\boldsymbol{x}) = A^{-1}\boldsymbol{x}$, in more explict terms, we have

$$T^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + x_3 \\ -x_2 - 2x_2 \\ x_3 \end{bmatrix} \tag{7.139}$$

8. Lecture 8 - Determinant

Definition 8.1: We say a square matrix A is upper triangular if it only has zeros to the left of the diagonal

$$\begin{bmatrix} * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \tag{8.140}$$

A is lower triangular if it only has zeros to the right of the diagonal

$$\begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 \\ * & * & * & * & * & 0 \\ * & * & * & * & * & * \end{bmatrix} \tag{8.141}$$

A is diagonal if A only has nonzero entries on the diagonal

$$\begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \tag{8.142}$$

A diagonal matrix is both upper triangular and lower triangular

8.1. Geometric definition of determninants

Now let's introduce determinants (ONLY for square matrices!!!). Consider the parallelepiped P with edges $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ in $\mathbb{R}^n$. We would like the following geometric definition of determinants.

Definition 8.1.1: The determinant of $A = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_n]$ (Usually denoted $\det A$ or $|A| = |\,[\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_n]\,|$, replacing brackets with vertical lines) as *signed* volumes of P . Therefore we have $\mathrm{Vol}(P) = |\det A|$, i.e. actual volume is the absolute value of the determinant.

Example 8.1.1: $[n = 1]$ Suppose $A = [a]$ is a 1 by 1 matrix, then $\det A$ is the signed length of $a \in \mathbb{R}^1$, which is a itself! Namely $\det A = a$.

Example 8.1.2: $[n = 2]$ Suppose $A = [\boldsymbol{a}_1 \ \boldsymbol{a}_2]$ is a 2 by 2 matrix. det A is the actual positive area of the parallelogram if $\boldsymbol{a}_1$ turns counter-clockwise to $\boldsymbol{a}_2$, otherwise the negative area.

Example 8.1.3: $[n = 3]$ Suppose $A = [\boldsymbol{a}_1 \ \boldsymbol{a}_2 \ \boldsymbol{a}_3]$ is a 3 by 3 matrix. To decide the sign of the volume of the parallelepiped, we follow the right-hand rule.

Determinant has following three properties:

- Interchanging $\boldsymbol{a}_1, \boldsymbol{a}_2$ changes the sign of determinant. Namely $\det [\boldsymbol{a}_2 \ \boldsymbol{a}_1] = -\det [\boldsymbol{a}_1 \ \boldsymbol{a}_2]$

Remark: If $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$ is linearly dependent, then A is singular, i.e. not invertible, then the determinant will be zero, since the parallelepiped will be constraint in a hyperplane which has zero volume. Take $n = 2$ and 3 for examples

8.2. Properties of determinants

Theorem 8.2.1: Suppose $A = \begin{bmatrix} h & 0 \\ * & B \end{bmatrix}$ or $\begin{bmatrix} h & * \\ \mathbf{0} & B \end{bmatrix}$ where h is a scalar, B is a $(n-1) \times (n-1)$ submatrix, then $\det A = h \cdot \det B$.

begin{corollary} Determinant of a triangular matrix is the product of the diagonal elements. end{corollary}

Proof: Apply Theorem Theorem 8.2.1 repeatedly. ■

Theorem 8.2.2: $\det(A) = \det(A^T)$

Thanks to Theorem Theorem 8.2.2, we can compute determinants via elementary row and column operations.

Example 8.2.1: Use elementary row operations to evaluate the following (Note that we omit the backward phase (which are replacements) since it doesn't change the determinants)

1.

$$\begin{vmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 1 \end{vmatrix} \xlongequal[\text{R3}\to\text{R3}-\text{R1}]{\text{R2}\to\text{R2}-2\text{R1}} \begin{vmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \\ 0 & 3 & 0 \end{vmatrix} \xlongequal{\text{factor R3}} 3\begin{vmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \\ 0 & 1 & 0 \end{vmatrix} \xlongequal{\text{R2}\to\text{R2}-2\text{R3}} 3\begin{vmatrix} 1 & -1 & 1 \\ 0 & 0 & -3 \\ 0 & 1 & 0 \end{vmatrix} \xlongequal{\text{R2}\leftrightarrow\text{R3}} (-1)\cdot 3\begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{vmatrix} = (-1)\cdot 3\cdot 1\cdot 1\cdot(-3) = 9 \tag{8.143}$$

1.

$$\left|\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & -1 \\ -1 & 2 & 2 \end{bmatrix}\right| \xlongequal[\text{R3}\to\text{R3}+\text{R1}]{\text{R2}\to\text{R2}-2\text{R1}} \begin{vmatrix} 1 & 2 & 1 \\ 0 & -4 & -3 \\ 0 & 4 & 3 \end{vmatrix} \xlongequal{\text{R3}\to\text{R3}+\text{R2}} \begin{vmatrix} 1 & 2 & 1 \\ 0 & -4 & -3 \\ 0 & 0 & 0 \end{vmatrix} = 1\cdot(-4)\cdot 0 = 0 \tag{8.144}$$

1.

$$\begin{vmatrix} 1 & 2 & 0 & 0 \\ 2 & 8 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ -4 & 5 & 7 & 1 \end{vmatrix} \xlongequal{\text{C2}\to\text{C2}-2\text{C1}} \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 1 & 2 & 0 \\ -4 & 13 & 7 & 1 \end{vmatrix} \xlongequal{\text{transpose}} \begin{vmatrix} 1 & 2 & 1 & -4 \\ 0 & 4 & 1 & 13 \\ 0 & 0 & 2 & 7 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 1\cdot 4\cdot 2\cdot 1 = 8 \tag{8.145}$$

Exercise 8.2.1: Suppose I is the $n \times n$ identity matrix, what is $\det I$, $\det(-I)$, $\det(2I)$ and $\det(aI)$?

Solution: Note that I is a diagonal matrix. $\det I = 1$, $\det(-I) = (-1)^n$, $\det(2I) = 2^n$, and in general $\det\,(aI) = a^n$ by factoring each row.

Example 8.2.2: Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \xlongequal{\text{R2}\to\text{R2}-\frac{c}{a}\text{R1}} \begin{vmatrix} a & b \\ 0 & d - b\frac{c}{a} \end{vmatrix} = a\left(d - b\frac{c}{a}\right) = ad - bc \tag{8.146}$$

Definition 8.2.1: Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with standard matrix A, the determinant of T is defined to be $\det T = \det A$.

Question: Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation with standard matrix $A = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. What is the area of the image of the unit circle under T?

Remark: In multi-variable calculus, this is known as the Jacobian.

Theorem 8.2.3: Suppose A, B are $n \times n$ matrices, then $\det(AB) = (\det\ A)(\det\ B)$

Proof: Suppose T_1, T_2 are linear transformations with standard matrices A, B, respectively. Consider $T_2 \circ T_1$ which should scale area by $\det(BA)$. On the other hand, this should scale area by $\det\ A$ and then $\det\ B$ ■

Theorem 8.2.4: A is invertible $\Leftrightarrow \det\ A \neq 0$. In addition, $\det(A^{-1}) = d\texttt{frac}\{1\}\{\det\ A\}$.

Proof: If A is invertible, then A^{-1} is well-defined, then $1 = \det\ I = \det(AA^{-1}) = (\det\ A)(\det(A^{-1})) \Rightarrow \det(A^{-1}) = d\texttt{frac}\{1\}\{\det\ A\}$, so $\det\ A \neq 0$. Conversely, if $\det\ A \neq 0$, A would have n pivots, so a pivot in each row and column, thus A will be invertible.■

Exercise 8.2.2: Compute the determinants of the following matrices

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

8.3. Cofactor expansion

There is one more property of determinants. $\det\ [\boldsymbol{w}\ \ \boldsymbol{u}+\boldsymbol{v}] = \det\ [\boldsymbol{w}\ \ \boldsymbol{u}] + \det\ [\boldsymbol{w}\ \ \boldsymbol{v}]$

Definition 8.3.1: We use a_{ij} to be denote the (i, j)-th entry of the matrix A, and A_{ij} to denote the submatrix of A by deleting the i-th row and the j-th column We define the (i, j)-cofactor to be $C_{ij} = (-1)^{\{i+j\}} \det\ A_{ij}$ (we also call $\det\ A_{ij}$ a minor).

For $n \geq 2$, with the help of Lemma “lemma:det-A-lemma”, we derive cofactor expansion formula.

The cofactor expansion across the i-th row is

$$\det\ A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{\in} \tag{8.147}$$

The cofactor expansion across the j-th column is

$$\det\ A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{8.148}$$

Proof: [Proof of Theorem Theorem 8.2.2] Note that row cofactor expansion of A is the same as column cofactor expansion of A^T. Hence we can prove this inductively on the size of the matrix. ■

Example 8.3.1:

$$\begin{vmatrix} 1 & 2 & 3 & 0 \\ 0 & 3 & -1 & 0 \\ -1 & 2 & 1 & 2 \\ 2 & -3 & 1 & 0 \end{vmatrix} \xlongequal{\text{cofactor expansion across last column}} 2(-1)^{\{3+4\}} \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & -1 \\ 2 & -3 & 1 \end{vmatrix}$$
$$\xlongequal{\text{R3}\to\text{R3}-2\text{R1}} (-2) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & -1 \\ 0 & -7 & -5 \end{vmatrix} \xlongequal{\text{cofactor expansion across first column}} (-2)\cdot 1(-1)^{\{1+1\}} \begin{vmatrix} 3 & -1 \\ -7 & -5 \end{vmatrix}$$
$$= (-2)(3(-5) - (-1)(-7)) = 44 \tag{8.149}$$

Remark: When use the cofactor expansion, we want to apply it to rows/columns with more $0's$

Exercise 8.3.1: Suppose $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$. Please find the cofactor expansion of A across the begin{tasks}(2) task 1st row task 2nd column end{tasks} And evaluate determinant of A.

Solution:

$$\begin{aligned} \det A &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}(-1)^{\{1+1\}}\det A_{11} + a_{12}(-1)^{\{1+2\}}\det A_{12} + a_{13}(-1)^{\{1+3\}}\det A_{13} \\ &= 1\cdot(-1)^{\{1+1\}}| \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} | + (-1)\cdot(-1)^{\{1+2\}}| \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} | + 1\cdot(-1)^{\{1+3\}}| \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} | \\ &= 1\cdot(0\cdot 1 - (-1)\cdot 1) + (-1)\cdot(2\cdot 1 - (-1)\cdot 1) + 1\cdot(-1)(2\cdot 1 - (-1)\cdot 1) \\ &= 6 \end{aligned} \tag{8.150}$$

$$\begin{aligned} \det A &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= a_{11}(-1)^{\{1+1\}}\det A_{11} + a_{12}(-1)^{\{1+2\}}\det A_{12} + a_{13}(-1)^{\{1+3\}}\det A_{13} \\ &= (-1)\cdot(-1)^{\{1+2\}}| \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} | + 0\cdot(-1)^{\{2+2\}}| \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} | + 1\cdot(-1)^{\{3+2\}}| \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} | \\ &= (-1)\cdot(-1)(2\cdot 1 - (-1)\cdot 1) + 0\cdot(1\cdot 1 - 1\cdot 1) + 1\cdot(-1)(1\cdot(-1) - 1\cdot 2) \\ &= 6 \end{aligned} \tag{8.151}$$

Exercise 8.3.2: Suppose $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -1 \\ -1 & 2 & 1 \end{bmatrix}$. Write out the cofactor expansion of A across the second row, and evaluate the determinant $\det A$.

Solution:

$$
\begin{aligned}
\det A &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} \\
&= a_{21}(-1)^{\{2+1\}}\det A_{21} + a_{22}(-1)^{\{2+2\}}\det A_{22} + a_{23}(-1)^{\{2+3\}}\det A_{23}; = 1\cdot(-1)^{\{2+1\}}| \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} | + 0 \\
&(-1)\cdot(-1)(1\cdot 2 - 2\cdot(-1)) \\
&= 4
\end{aligned}
\tag{8.152}
$$

Remark: The REF of a square matrix A is upper triangular, and $\det A = 0$ if A has less then n pivots.

Exercise 8.3.3: Suppose $A = \begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Please find use cofactor expansion to find the $\det A$

Solution: Note that A is upper triangular, so we could do cofactor expansions across first columns multiple times

$$
| \begin{bmatrix} 2 & -1 & 3 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} | = 2(-1)^{\{1+1\}}| \begin{bmatrix} -2 & 0 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix} | = 2\cdot(-2)(-1)^{\{1+1\}}| \begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix} | = 2\cdot(-2)\cdot 3\cdot(-1)^{\{1+1\}}(1) = -12 \tag{8.153}
$$

Exercise 8.3.4: Consider $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. What is $A_{11}, A_{12}, A_{21}, A_{22}$? What is $C_{11}, C_{12}, C_{21}, C_{22}$. Write down the cofactor expansion of A across the

- 1st row
- 2nd row
- 1st column
- 2nd column

Solution: $A_{11} = [d], A_{12} = [c], A_{21} = [b], A_{22} = [a]$ are all 1 by 1 matrices. $C_{11} = (-1)^{\{1+1\}}\det A_{11} = d, C_{12} = (-1)^{\{1+2\}}\det A_{21} = -c, C_{21} = (-1)^{\{2+1\}}\det A_{21} = -b, C_{22} = (-1)^{\{2+2\}}\det A_{22} = a$. So the cofactor expansions are

- $\det A = aC_{11} + bC_{12} = ad - bc$
- $\det A = cC_{21} + dC_{22} = -bc + ad$
- $\det A = aC_{11} + cC_{21} = ad - bc$
- $\det A = bC_{12} + dC_{22} = -bc + ad$

Note that all of the above calculations show that $\det A = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

9. Lecture 9 - Vector spaces and subspaces

9.1. Vector space

To motivate the definition of a vector space, let's consider the following example

Example 9.1.1: Let $\mathbb{P}_n$ denote the set of (real) polynomials of degree less or equal to n. For example $\mathbb{P}_0 = \mathbb{R}$ is just the set of real numbers, and

$$\begin{aligned} \mathbb{P}_1, &= \{a_0 + a_1 t | a_0, a_1 \in \mathbb{R}\} \\ \mathbb{P}_2, &= \{a_0 + a_1 t + a_2 t^2 | a_0, a_1, a_2 \in \mathbb{R}\} \\ \mathbb{P}_3, &= \{a_0 + a_1 t + a_2 t^2 + a_3 t^3 | a_0, a_1, a_2, a_3 \in \mathbb{R}\} \\ ,\vdots; \mathbb{P}_n, &= \{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n | a_0, a_1, a_2, \cdots, a_n \in \mathbb{R}\} \end{aligned} \tag{9.154}$$

You may soon realize that $\mathbb{P}_n$ can be identified with $\mathbb{R}^{n+1}$. More concrete examples could be

- For $\mathbb{P}_1 \cong \mathbb{R}^2$, $1 + 2t \rightsquigarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
- For $\mathbb{P}_2 \cong \mathbb{R}^3$, $3t^2 - 1 \rightsquigarrow \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix}$

If we consider addition and scalar multiplication, we have So we may conclude that addition and scalar multiplication in $\mathbb{P}_n$ can be identically translated to addition and scalar multiplication in $\mathbb{R}^{n+1}$

Remark: We call $\{1, t, t^2, \cdots, t^n\}$ the *standard basis* of $\mathbb{P}_n$, corresponding to the standard basis for $\mathbb{R}^{n+1}$

Example 9.1.2: $\{1, t, t^2\}$ is the standard basis for $\mathbb{P}_2$, and

$$p(t), = a_0 + a_1 t + a_2 t^2 = a_0 \cdot 1 + a_1 \cdot t + a_2 \cdot t^2 \tag{9.155}$$

Example 9.1.3: Let's denote $M_{m\times n}(\mathbb{R})$ the set of $m \times n$ matrices. For example

$$\begin{aligned} M_{\{2\times 2\}}(\mathbb{R}), &= \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \middle| a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\} \\ M_{\{3\times 2\}}(\mathbb{R}), &= \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \middle| a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32} \in \mathbb{R} \right\} \\ M_{\{2\times 3\}}(\mathbb{R}), &= \left\{ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \middle| a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23} \in \mathbb{R} \right\} \\ ,\vdots; M_{m\times n}(\mathbb{R}), &= \left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \middle| a_{11}, \cdots, a_{1n}, \cdots, a_{m1}, \cdots, a_{mn} \in \mathbb{R} \right\}. \end{aligned} \tag{9.156}$$

You may realize that $M_{m\times n}(\mathbb{R})$ can be identified with $\mathbb{R}^{\{mn\}}$ So we may conclude that addition and scalar multiplication in $M_{m\times n}(\mathbb{R})$ can be identically translated to addition and scalar multiplication in $\mathbb{R}^{\{mn\}}$

Remark: We call $\{E_{ij}\}$ the *standard basis* of $M_{m\times n}(\mathbb{R})$, corresponding to the standard basis for $\mathbb{R}^{\{mn\}}$. Here E_{ij} is the $m\times n$ matrix that only has a single 1 in the (i,j)-th spot, but $0'$s elsewhere.

Example 9.1.4:

$$\left\{E_{11}=\begin{bmatrix}1&0\\0&0\end{bmatrix},E_{12}=\begin{bmatrix}0&1\\0&0\end{bmatrix},E_{21}=\begin{bmatrix}0&0\\1&0\end{bmatrix},E_{22}=\begin{bmatrix}0&0\\0&1\end{bmatrix}\right\} \tag{9.157}$$

is the standard basis for $M_{\{2\times 2\}}(\mathbb{R})$, and

$$\begin{aligned}\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}&=\begin{bmatrix}a_{11}&0\\0&0\end{bmatrix}+\begin{bmatrix}0&a_{12}\\0&0\end{bmatrix}+\begin{bmatrix}0&0\\a_{21}&0\end{bmatrix}+\begin{bmatrix}0&0\\0&a_{22}\end{bmatrix}\\&=a_{11}\begin{bmatrix}1&0\\0&0\end{bmatrix}+a_{12}\begin{bmatrix}0&1\\0&0\end{bmatrix}+a_{21}\begin{bmatrix}0&0\\1&0\end{bmatrix}+a_{22}\begin{bmatrix}0&0\\0&1\end{bmatrix}\\&=a_{11}E_{11}+a_{12}E_{12}+a_{21}E_{21}+a_{22}E_{22}\end{aligned} \tag{9.158}$$

Definition 9.1.1: A (real) vector space is a set V of objects, called *vectors*, on which are defined two operations, called *addition* fatplus and *(left) scalar multiplication* fatdot, subject to axioms begin{enumerate}setcounter{enumi}{-1}

- bmufatplusbmv and $c$fatdotbmv are still in V
- bmufatplus$bmv = bmv$fatplusbmu
- $(bmu$fatplus$bmv)$fatplus$bmw = bmu$fatplus$(bmv$fatplus$bmw)$
- There is a *zero vector* $bm0$ such that bmufatplus$bm0 = bmu$
- For each bmu in V, there is a vector fatminusbmu in V such that bmufatplus(fatminus$bmu) = bm0$
- $c$fatdot$(bmu$fatplus$bmv) = c$fatdotbmufatplus $c$fatdotbmv
- $(c+d)$fatdot$bmu = c$fatdotbmufatplus $d$fatdotbmu
- $c$fatdot$(d$fatdot$bmu) = (cd)$fatdotbmu
- 1fatdot$bmu = bmu$

Example 9.1.5: Set V to be $\mathbb{R}^n$, fatplus to be addition $+$ for vectors, fatdot to be scalar multiplication

$$\cdot \tag{9.159}$$

for vectors, then this is a vector space

Example 9.1.6: [non-example] Suppose $V = \mathbb{R}$, $a$fatplus $b = a + b + 1$, $c$fatdot $a = c \cdot a = ca$, we can check begin{enumerate}setcounter{enumi}{-1}

- $a$fatplus $b = a + b + 1 \in \mathbb{R}$, $c$fatdot $a = ca \in \mathbb{R}$
- $a$fatplus $b = a + b + 1 = b + a + 1 = b$fatplus a
- $(a$fatplus $b)$fatplus $c = (a + b + 1) + c + 1 = a + (b + c + 1) + 1 = a$fatplus$(b$fatplus $c)$
- There is a *zero vector* $bm0 = -1$ such that $a$fatplus$bm0 = a + (-1) + 1 = a$
- For each a, we have fatminus $a = -a - 2$ such that $a$fatplus(fatminus $a) = a + (-a - 2) + 1 = -1 = bm0$

However 2fatdot$(a$fatplus $b) = 2(a + b + 1) \neq 2a + 2b + 1 =$ 2fatdot $a$fatplus 2fatdot b. Therefore, this is not a vector space

9.2. Subspace

Definition 9.2.1: Suppose V is a vector space with addition fatplus and scalar multiplication fatdot. A subspace H is a non-empty subset which closed under addition and scalar multiplication, i.e. for any $bmu, bmv \in H$, $c \in \mathbb{R}$, bmufatplusbmv, $c$fatdot$bmu \in H$

Remark: It is easy to check that a subspace H is again a vector space.

Exercise 9.2.1: Recall that $M_{\{2\times 2\}}(\mathbb{R})$ is the set of 2 by 2 matrices, and that a square matrix A is *symmetric* if $A^T = A$. Consider a subset V consists of 2 by 2 symmetric matrices, i.e. $V = \left\{A \in M_{\{2\times 2\}}(\mathbb{R}) \;\middle|\; A^T = A\right\}$

- Show that V is a vector space.
- Find a basis for V.

Solution:

- For any $A, B \in V$, $c \in \mathbb{R}$, by definition we know that $A^T = A$, $B^T = B$, we want to show that $A + B \in V$, $cA \in V$ (condition for subspace), i.e. $(A + B)^T = A^T + B^T$, $(cA)^T = cA$. This is true because

$$(A + B)^T = A^T + B^T = A + B, \quad (cA)^T = cA^T = cA \tag{9.160}$$

Therefore V is a subspace of $M_{\{2\times 2\}}(\mathbb{R})$, and thus a vector space

- Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_{\{2\times 2\}}(\mathbb{R})$, then $a_{12} = a_{21}$, so we may conclude that

$$V = \left\{ \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in M_{\{2\times 2\}}(\mathbb{R}) \;\middle|\; a, b, c \in \mathbb{R} \right\} \tag{9.161}$$

Note that

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \\ b & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{9.162}$$

And that linear combination (9.162) is the zero matrix $\Leftrightarrow a = b = c = 0$, thus $\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for V

Exercise 9.2.2: Suppose $H = \{p(t) = a_0 + a_1 t + a_2 t^2 \in \mathbb{P}_2 | a_0 + a_1 + a_2 = 0\}$ is the set of polynomials of degree $l = 2$ and sum of coefficients zero. Show that H is a vector space.

Example 9.2.1: Consider the vector space $V = \mathbb{R}^2$, and H is the set of solutions to the linear equation $2x_1 - x_2 + 2 = 0$, then H is not a subspace. For example, if we choose $\boldsymbol{u} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, then $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ is not in H, nor is $2\boldsymbol{u} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$ The reason is that H is not homogeneous. If we consider H_1 to be solution set of the homogeneous equation $2x_1 - x_2 = 0$, we see that H_1 is a subspace as it is the span of a single vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

10. Lecture 10 - Null spaces, Column spaces, Row spaces, Rank and Nullity

10.1. Null space, Column Space and Row space

Definition 10.1.1: Suppose V is a vector space, $\{bmv_1, bmv_2, \cdots, bmv_n\} \subseteq V$ is a set of linearly independent vectors that span V, we define the dimension of V to be $\dim V = n$.

Definition 10.1.2: Suppose A is a $m \times n$ matrix, we define the null space of A to be

$$\mathrm{Nul} A = \{\boldsymbol{x} \in \mathbb{R}^n | A\boldsymbol{x} = \boldsymbol{0}\} = \text{solution set of \$A bold(x)=bold(0)\$} \qquad (10.163)$$

Note that the solution set to linear system $A\boldsymbol{x} = \boldsymbol{0}$ is the intersection of m hyperplanes (one for each homogeneous equation) that pass through the origin.

Example 10.1.1: $A = \begin{bmatrix} 1 & -3 & 2 \\ -5 & 12 & -1 \end{bmatrix}$, the to find the Nul$A$ is equivalent to solve $A\boldsymbol{x} = \boldsymbol{0}$

$$[A\ \ \boldsymbol{0}] \xrightarrow{\texttt{R2}\to\texttt{R2+5R1}} \begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & -3 & 9 & 0 \end{bmatrix} \xrightarrow{\frac{\texttt{R2}}{-3}} \begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & 1 & -3 & 0 \end{bmatrix} \xrightarrow{\texttt{R1}\to\texttt{R1+3R2}} \begin{bmatrix} 1 & 0 & -7 & 0 \\ 0 & 1 & -3 & 0 \end{bmatrix} \tag{10.164}$$

Hence the solution set is

$$\begin{cases} x_1 = 7x_3 \\ x_2 = 3x_3 \\ x_3 \text{ is free} \end{cases} \tag{10.165}$$

, in parametric form, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$, which describes a line in $\mathbb{R}^3$ of the direction $\begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$ that passes through the origin, and this line is the intersection of planes $x_1 - 3x_2 + 2x_3 = 0$ and $-5x_1 + 12x_2 - x_3 = 0$

Remark: As discussed in Example Example 9.2.1, in general, the solution set of $A\boldsymbol{x} = \boldsymbol{b}$ is not a subspace of $\mathbb{R}^n$ unless $\boldsymbol{b} = \boldsymbol{0}$. And in fact, any subspace of $\mathbb{R}^n$ is the null space for some $m \times n$ matrix A, i.e. the intersection of hyperplanes passing through the origin

Definition 10.1.3: Suppose $A = [\boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \cdots\ \ \boldsymbol{a}_n]$ is an $m \times n$ matrix, then the column space (denote as Col$A$) is the subspace Span$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n\}$ in $\mathbb{R}^m$. Suppose $A = \begin{bmatrix} \texttt{R1} \\ \texttt{R2} \\ \vdots \\ \texttt{Rm} \end{bmatrix}$, then the row space (denote as Row$A$) is the subspace spaned by row vectors Span$\{\texttt{R1}, \texttt{R2}, \cdots, \texttt{Rm}\}$ in $\mathbb{R}^n$ written horizontally.

Theorem 10.1.1: Row reductions preserve Nul$A$, Row$A$, but not Col$A$. However, row reductions preserve linear dependences of the column vectors.

Remark: Suppose $A\boldsymbol{x} = \boldsymbol{0}$ is some linear dependence of the column vectors of A, after elementary row reduction $A \sim EA$, $(EA)\boldsymbol{x} = E(A\boldsymbol{x}) = E\boldsymbol{0} = \boldsymbol{0}$ is again the same linear dependence of columns of EA. In other words, the linear dependence of columns of a matrix is preserved by row equivalence.

Theorem 10.1.2: Suppose A is a $m \times n$ matrix, $A \sim U$ is of RREF form

- The vectors in the parametric vector form of the solution set of $A\boldsymbol{x} = \boldsymbol{0}$ gives a basis for Nul$A$. Note that dimNul$A$ = the number of free variables.
- A basis for Col$A$ could be the set of pivot columns in A. Note that dim Col$A$ = the number of pivots.
- A basis for Row$A$ could be the set of non-zero row vectors in U (Or any REF of A). Note that dim Row$A$ = the number of pivots.

10.2. Rank and Nullity

Definition 10.2.1: $\dim \mathrm{Nul} A$ is also name the nullity of A. The number of pivots of A (which is equal to both $\dim \mathrm{Col} A$ and $\dim \mathrm{Row} A$) is called the rank of A

Theorem 10.2.1: [Rank-Nullity theorem] Notice that the number of columns in A (say a $m \times n$ matrix) is equal to the number of free variables and the number of pivot columns, thus we have

$$n = \text{nullity} + \text{rank} \tag{10.166}$$

Example 10.2.1: $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}$, which is an REF and an RREF respectively. There is only one free variable x_3, so the nullity is 1, and the 1st, 2nd columns are pivot columns, so the rank is 2. We see that Theorem Theorem 10.2.1 holds as $3 = 1 + 2$, and

$$\mathrm{Nul} A = \mathrm{Span}\left\{ \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} \right\} \tag{10.167}$$

$$\mathrm{Col} A = \mathrm{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{10.168}$$

$$\mathrm{Row} A = \mathrm{Span}\left\{ \begin{bmatrix} 1 & 0 & -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 & 1 & -\frac{3}{2} \end{bmatrix} \right\} \text{ or } \mathrm{Span}\{[1 \ \ -1 \ \ 1], [1 \ \ 2 \ \ -3]\} \tag{10.169}$$

Exercise 10.2.1: $A = \begin{bmatrix} 1 & 2 & 0 & 4 & 5 \\ 0 & 0 & 5 & -7 & 8 \\ 0 & 0 & 0 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -\frac{7}{5} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ Note that here we have 2 free variables x_2, x_4, so the nullity is 2, and the 1st, 3rd, 5th columns are pivot columns, so the rank is 3. We see that Theorem Theorem 10.2.1 holds as $5 = 2 + 3$, and

$$\text{Nul}A = \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ \frac{7}{5} \\ 1 \end{bmatrix} \right\} \tag{10.170}$$

$$\text{Col}A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 8 \\ -9 \\ 0 \end{bmatrix} \right\} \tag{10.171}$$

$$\text{Row}A = \text{Span}\{[1\ \ 2\ \ 0\ \ 4\ \ 5], [0\ \ 0\ \ 5\ \ -7\ \ 8], [0\ \ 0\ \ 0\ \ 0\ \ -9]\} \tag{10.172}$$

Question: If you have a set $\mathcal{S}$ of vectors in $\mathbb{R}^m$, how do you find a subset of $\mathcal{S}$ that is a basis for Span$\{\mathcal{S}\}$ (i.e. remove linear dependences)?

Answer: Collect these vectors as the column vectors of a matrix, and then find its columns space.

Exercise 10.2.2: Suppose $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$. Find a basis for Span$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$.

Exercise 10.2.3: Suppose $A = \begin{bmatrix} 1 & -7 & 0 & 6 & 5 \\ 0 & 0 & 1 & -2 & -3 \\ -1 & 7 & -4 & 2 & 7 \end{bmatrix}$. Find Nul$A$, Row$A$, Col$A$, and the nullity, rank of A.

11. Lecture 11 - Linear transformations in general

11.1. Linear transformation

Definition 11.1.1: Suppose V, W are vector spaces, a *linear transformation* is a mapping $T : V \to W$ such that

- $T(bmu fatplus bmv) = T(bmu) fatplus\ T(bmv)$
- $T(cfatdot\ u) = cfatdot\ T(bmu)$

Just as before, we call V the *domain* of T, W the *codomain* of T, the *image* of bmx under T is $T(bmx)$, the set of images $\{T(bmx) | bmx \in V\}$ the *range* (denoted as Range T), and the set $\{bmx | T(bmx) = bmb\}$ the preimage of bmb under T. We still say that T is *one-to-one* if any $bmb \in W$, there is at most one $bmx \in V$ such that $T(bmx) = bmb$. T is *onto* the range is the codomain. T is said to be *invertible* if T has a inverse (this happens if and only if T is both one-to-one and onto, and we can easily check that T^{-1} is also an linear transformation), in this case we also call T an isomorphism.

Definition 11.1.2: We call $\{bmx | T(bmx) = \mathbb{0}\}$ the kernel (or null space) of T, and denote it as $\ker T$.

Remark: label{14:24/06/30/2022} Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(\boldsymbol{x}) = A\boldsymbol{x}$ is a matrix transformation, then $\ker T = \mathrm{Nul} A$, Range $T = \mathrm{Col} A$

Example 11.1.1: The identification

$$T : \mathbb{P}_2 \to \mathbb{R}^3, \quad T(a_0 + a_1 t + a_2 t^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{11.173}$$

in Eaxmple Example 9.1.1 is an invertible linear transformation with inverse linear transformation

$$T^{-1} : \mathbb{R}^3 \to \mathbb{P}_2, \quad T^{-1}\left(\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \right) = a_0 + a_1 t + a_2 t^2 \tag{11.174}$$

Example 11.1.2: The identification

$$T : M_{\{2\times 2\}}(\mathbb{R}) \to \mathbb{R}^4, \quad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \tag{11.175}$$

in Eaxmple Example 9.1.3 is an invertible linear transformation with inverse linear transformation

$$T^{-1} : \mathbb{R}^4 \to M_{\{2\times 2\}}(\mathbb{R}), \quad T^{-1}\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{11.176}$$

Theorem 11.1.1: Suppose $T : V \to W$ is a linear transformation between vector spaces, then

- $\ker T$ is a subspace of V.
- Range T is a subspace of W.

Proof:

- Suppose $bmu, bmv \in \ker T$, then by definition $T(bmu) = \mathbb{0}$, $T(bmv) = \mathbb{0}$, so $T(bmu\text{fatplus}bmv) = T(bmu)\text{fatplus}\ T(bmv) = \mathbb{0}$. And for any $c \in \mathbb{R}$, $T(c\text{fatdot}bmu) = c\text{fatdot}\ T(bmu) = \mathbb{0}$. In other words, we have shown that $bmu\text{fatplus}bmv, c\text{fatdot}bmu \in \ker T$, so $\ker T$ is a subspace.
- For any $T(bmu), T(bmv) \in \text{Range } T$, $T(bmu)\text{fatplus}\ T(bmv) = T(bmu\text{fatplus}bmv) \in \text{Range } T$, and for any $c \in \mathbb{R}$, $c\text{fatdot}\ T(bmu) = T(c\text{fatdot}bmu) \in \text{Range } T$, so Range T is a subspace

■

Exercise 11.1.1: Suppose $T : \mathbb{P}_2 \to \mathbb{R}$ takes the sum of coefficients, i.e. $T(a_0 + a_1 t + a_2 t^2) = a_0 + a_1 + a_2$. Show that T is a linear transformation, and H in Exercise Exercise 9.2.2 is a vector space.

Solution: since for any $p(t) = a_0 + a_1 + a_2, q(t) = b_0 + b_1 t + b_2 t^2 \in \mathbb{P}_2$, $c \in \mathbb{R}$, we have

$$q), = T((a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2) = (a_0 + b_0) + (a_1 + b_1) + (a_2 + b_2);, = (a_0 + a_1 + a_2) + (b_0 + b_1 + b_2) = T(p) + \tag{11.177}$$

$$T(cp) = T((ca_0) + (ca_1)t + (ca_2)t^2) = (ca_0) + (ca_1) + (ca_2) = c(a_0 + a_1 + a_2) = cT(p) \tag{11.178}$$

11.2. Matrices of linear transformations

Question: Can we realize a general linear transformation T as a matrix transformation?

Answer: Just need to know its effect on any basis! Suppose $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ is a basis of $\mathbb{R}^n$, then any vector $\boldsymbol{v}$ can be written as $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$, and by linearality, we have

Theorem 11.2.1: [Unique representation theorem]label{12:53-06/28/2022} Suppose $\mathcal{B} = \{bmb_1, \cdots, bmb_n\}$ is a basis of a vector sapce V, then any vector $bmv \in V$ can be uniquely represented as a linear combination $x_1$fatdotbmb_1fatplus$\cdots$fatplus $x_n$fatdotbmb_n

Example 11.2.1: Consider linear transformation

$$T : \mathbb{P}_2 \to \mathbb{R}^3, \quad T(a_0 + a_1 t + a_2 t^2) = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{11.179}$$

and $\mathcal{B} = \{1 + t, t + t^2, 1 + t^2\}$ is a basis for $\mathbb{P}_2$, $\mathcal{E}$ is the standard basis for $\mathbb{R}^3$, then matrix of T relative to bases $\mathcal{B}$ and $\mathcal{E}$ can be read from eqref{12:55-06/29/2022} as

$$A = [T(1+t) \;\; T(t+t^2) \;\; T(1+t^2)] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tag{11.180}$$

Exercise 11.2.1: Suppose $T : \mathbb{P}_2 \to \mathbb{R}_3$ is evaluation at 2, i.e. $T(p(t)) = p(2)$ for $p(t) \in \mathbb{P}_2$. Show that T is a linear transformation. Is T onto? Is T one-to-one? Is T invertible? Find a basis for $\ker T$. Find a basis for Range T.

Exercise 11.2.2: Suppose H is the set of 2 by 2 matrices that are symmetric with the sum of diagonal being zero. Show H is a vector space. Find a basis of H.

11.3. Change of basis

12. Lecture 12 - Eigendecomposition

12.1. Eigenvalues, eigenvectors and eigenspaces

12.2. Characteristic polynomials

12.3. Similarity

12.4. Eigendecomposition and diagonalization

13. Lecture 13 - Orthogonalization

13.1. Orthogonal and orthonormal basis

Definition 13.1.1: We say a vector $\boldsymbol{u}$ is of unit length (or a unit vector, or a normalized vector) if $|\boldsymbol{u}| = 1$

Definition 13.1.2: $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ is called an orthogonal set if $\boldsymbol{v}_i, \boldsymbol{v}_j$ $(i \neq j)$ are orthogonal. $\mathcal{B}$ is called an orthogonal basis if $\mathcal{B}$ is in addition a basis. $\mathcal{B}$ is called an orthonormal basis if the basis vectors are in addition normalized.

Remark: Suppose $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$, to test if $\mathcal{B}$ is an orthogonal (or orthonormal) set, we just need to write $A = [\boldsymbol{v}_1 \ \cdots \ \boldsymbol{v}_n]$, and test if

$$A^T A = \begin{bmatrix} \boldsymbol{v}_1^T \\ \boldsymbol{v}_2^T \\ \vdots \\ \boldsymbol{v}_n^T \end{bmatrix} [\boldsymbol{v}_1 \ \ \boldsymbol{v}_2 \ \ \cdots \ \ \boldsymbol{v}_n] = \begin{bmatrix} \boldsymbol{v}_1^T \boldsymbol{v}_1 & \boldsymbol{v}_1^T \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_1^T \boldsymbol{v}_n \\ \boldsymbol{v}_2^T \boldsymbol{v}_1 & \boldsymbol{v}_2^T \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_2^T \boldsymbol{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{v}_n^T \boldsymbol{v}_1 & \boldsymbol{v}_n^T \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n^T \boldsymbol{v}_n \end{bmatrix} \tag{13.181}$$

is diagonal (or if $A^T A = I$)

13.2. Gram-Schmidt process

Question: Suppose you have two vectors $\boldsymbol{u}, \boldsymbol{v}$ (here $\boldsymbol{v} \neq \boldsymbol{0}$), what is the orthogonal projection of $\boldsymbol{u}$ onto $\boldsymbol{v}$ (Which we denote as $\mathrm{Proj}_{\boldsymbol{v}} \boldsymbol{u}$)?

Answer: First you realize that $\mathrm{Proj}_{\boldsymbol{v}} \boldsymbol{u}$ is parallel to $\boldsymbol{v}$, so we write it as $\lambda \boldsymbol{v}$, and we know $|\mathrm{Proj}_{\boldsymbol{v}} \boldsymbol{u}| = \lambda \ |\boldsymbol{v}| = |\boldsymbol{u}| \cos\theta$, so we may conclude that

$$\lambda = |\boldsymbol{u}| \frac{\cos\theta}{|\boldsymbol{v}|} = |\boldsymbol{u}\|\boldsymbol{v}| \frac{\cos\theta}{|\boldsymbol{v}|^2} = \boldsymbol{u} \cdot \frac{\boldsymbol{v}}{\boldsymbol{v}} \cdot \boldsymbol{v} \tag{13.182}$$

Therefore we have derived the equation

$$\text{Proj}_{\boldsymbol{v}}\boldsymbol{u} = \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\boldsymbol{v} \tag{13.183}$$

Example 13.2.1: Consider $\mathcal{B} = \left\{ \boldsymbol{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \right\}$, $\boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $W = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$. Let $U = [\boldsymbol{u}_1 \ \ \boldsymbol{u}_2] = \begin{bmatrix} 2 & -1 \\ 1 & 2 \\ 0 & 0 \end{bmatrix}$, then $A^T A = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$ is diagonal, hence $\mathcal{B}$ is an orthogonal set, the orthogonal projection of $\boldsymbol{y}$ onto W is

$$= \boldsymbol{y} \cdot \frac{\boldsymbol{u}_1}{\boldsymbol{u}_1} \cdot \boldsymbol{u}_1 \boldsymbol{u}_1 + \boldsymbol{y} \cdot \frac{\boldsymbol{u}_2}{\boldsymbol{u}_2} \cdot \boldsymbol{u}_2 \boldsymbol{u}_2;, = 1\cdot 2 + 1 \cdot 1 + 1 \cdot \frac{0}{2^2} + 1^2 + 0^2 \boldsymbol{u}_1 + 1 \cdot (-1) + 1 \cdot 2 + 1 \cdot \frac{0}{(-1)^2} + 2^2 + 0^2 \boldsymbol{u}_2;, = \frac{3}{5}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{5}\begin{bmatrix} - \\ 2 \\ 0 \end{bmatrix} \tag{13.184}$$

Question: Suppose we are given arbitrary basis $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ for a subspace W of $\mathbb{R}^n$, how could we get a orthogonal (or orthonormal) basis $\mathcal{U} = \{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ from it

Answer: We apply the Gram-Schmidt process

Then $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ would be an orthogonal basis. To get an orthonormal basis, just normalize these vectors.

Example 13.2.2: Consider $\left\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \right\}$. Let's use Gram-Schmidt process to find an orthogonal (and an orthonormal) basis from it.

- First we choose $\boldsymbol{u}_1 = \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- $\boldsymbol{v}_2 - d\mathtt{frac}\{\boldsymbol{v}_2 \cdot \boldsymbol{u}_1\}\{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1\}\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{5}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$. Let's instead take a multiple of this to be our $\boldsymbol{u}_2$, namely we set $\boldsymbol{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$
- $\boldsymbol{u}_3 = \boldsymbol{v}_3 - d\mathtt{frac}\{\boldsymbol{v}_3 \cdot \boldsymbol{u}_1\}\{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1\}\boldsymbol{u}_1 - d\mathtt{frac}\{\boldsymbol{v}_3 \cdot \boldsymbol{u}_2\}\{\boldsymbol{u}_2 \cdot \boldsymbol{u}_2\}\boldsymbol{u}_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

Thus $\left\{ \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ is an orthogonal basis for $\mathbb{R}^3$. If we further normalize it, we have

$$\begin{aligned} \boldsymbol{w}_1 &= \frac{\boldsymbol{u}_1}{|\boldsymbol{u}_1|} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \\ \boldsymbol{w}_2 &= \frac{\boldsymbol{u}_2}{|\boldsymbol{u}_2|} = \frac{1}{\sqrt{6}}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \\ \boldsymbol{w}_3 &= \frac{\boldsymbol{u}_3}{|\boldsymbol{u}_3|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned} \tag{13.185}$$

is an orthonormal basis.

Exercise 13.2.1: Use Gram-Schmidt process to find an orthogonal and orthonormal basis of the following basis

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

13.3. Orthogonal matrix and QR factorization

Definition 13.3.1: A square matrix U is called an orthogonal matrix if the columns of U form an orthonormal basis $\Leftrightarrow U^T U = UU^T = I$. Note that in this particular case, $U^T = U^{-1}$

Remark: The rows of U also form an orthonormal basis.

Note: The name *orthogonal matrix* is unfortunate since it is made out of orthonormal basis. It should really be called the orthonormal matrix. But since mathematicians has been using this for a long time. So we will stick to this usage.

Definition 13.3.2: Suppose W is a subspace of $\mathbb{R}^n$, the orthogonal complement $W^\perp = \{\boldsymbol{v} | \boldsymbol{v} \perp W\}$ is the set of vectors that are orthogonal to W

Remark: It is easy to show that W is a subspace of $\mathbb{R}^n$ and $\left(W^\perp\right)^\perp = W$. This is kind of a *duality*, take $\mathbb{R}^3$ for an example, the orthogonal complement of a line would be the perpendicular plane, the orthogonal complement of a plane would be the perpendicular line, the orthogonal complement of the origin (a point) would be the whole space ($\mathbb{R}^3$) and the orthogonal complement of $\mathbb{R}^3$ would be the origin $\{\mathbf{0}\}$. From this we also see that $\dim W + \dim W^\perp = n$.

We make the following crucial observation: $\boldsymbol{x}$ is in Nul$A \Leftrightarrow \boldsymbol{x}$ is orthogonal all row vectors of $A \Leftrightarrow \boldsymbol{x} \perp \text{Row}A$, so we may conclude that $(\text{Row}A)^\perp = \text{Nul}A$, and similarly we know that $(\text{Col}A)^\perp = \left(\text{Row }\left(A^T\right)\right)^\perp = \text{Nul }\left(A^T\right)$.

Exercise 13.3.1: Suppose $W = \text{Span}\left\{\begin{bmatrix}1\\1\\0\\2\\0\\3\end{bmatrix}, \begin{bmatrix}0\\0\\1\\-2\\0\\2\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\1\end{bmatrix}\right\}$, find $(\text{Col}A)^\perp$.

14. Lecture 14 - Least-square problem

14.1. Least-square solutions

Question: linear system $A\boldsymbol{x} = \boldsymbol{b}$ may not always be solvable, nonetheless, we still want to find the *best possible* solution $h@\{\boldsymbol{x}\}$

Answer: According to Gauss, by best possible we mean the least-square solutions. In concrete terms, such $h@\{\boldsymbol{x}\}$ stisfies that $|Ah@\{\boldsymbol{x}\} - \boldsymbol{b}|l = |A\boldsymbol{x} - \boldsymbol{b}|$ for any other choice of $\boldsymbol{x}$.

By a simple argument we can show that $Ah@\{\boldsymbol{x}\}$ must be the projection of $\boldsymbol{b}$ onto Col$A$, so we know that

$$(Ah@\{\boldsymbol{x}\} - \boldsymbol{b}) \perp \text{Col}A \Leftrightarrow (Ah@\{\boldsymbol{x}\} - \boldsymbol{b}) \in (\text{Col}A)^\perp = \text{Nul }\left(A^T\right) \Leftrightarrow A^{T(Ah@\{\boldsymbol{x}\} - \boldsymbol{b})} = \mathbf{0} \Leftrightarrow A^T Ah@\{\boldsymbol{x}\} = A^T\boldsymbol{b}$$

14.2. Machine Learning and approxi mation

Question: Suppose you know a curve has the form

$$y = \beta_1 f_1(x_1, \cdots, x_k) + \beta_2 f_2(x_1, \cdots, x_k) + \cdots + \beta_n f_{n(x_1, \cdots, x_k)} \quad (14.187)$$

And you know what functions $f_1, f_2, \cdots, f_n$ should be (they could be linear functions, quadratic functions, polynomial functions, exponential functions, logarithmic function, trigonometry functions, etc.). You also have collected a bunch of data points $\left(x_1^{(1)}, \cdots, x_k^{(1)}, y^{(1)}\right), \left(x_1^{(2)}, \cdots, x_k^{(2)}, y^{(2)}\right), \cdots, \left(x_1^{(m)}, \cdots, x_k^{(m)}, y^{(m)}\right)$. How do you find best fitting coefficients $\beta_1, \beta_2, \cdots, \beta_n$? First let's define the residual

$$\varepsilon^{(i)} = y^{(i)} - \beta_1 f_1\left(x_1^{(i)}, \cdots, x_k^{(i)}\right) - \beta_2 f_2\left(x_1^{(i)}, \cdots, x_k^{(i)}\right) - \cdots - \beta_n f_{n\left(x_1^{(i)}, \cdots, x_k^{(i)}\right)} \quad (14.188)$$

Our goal to minimize the residual. One might want to minimize the quantity $\sum_{i=1}^{m} |\varepsilon^{(i)}|$, however, this is cumbersome, difficult to work with, and mathematically unsatisfactory. So Gauss instead considered the square sum $\sum_{\{i=1\}}^{m} |\varepsilon^{(i)}|^2$, posed and solved the least square problem!!!

Answer: We use normal equations.

Let's write

$$\boldsymbol{y} = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon^{(1)} \\ \vdots \\ \varepsilon^{(m)} \end{bmatrix}, X = \begin{bmatrix} f_1\left(x_1^{(1)}, \cdots, x_k^{(1)}\right) & \cdots & f_{n\left(x_1^{(1)}, \cdots, x_k^{(1)}\right)} \\ \vdots & \ddots & \vdots \\ f_1\left(x_1^{(m)}, \cdots, x_k^{(m)}\right) & \cdots & f_{n\left(x_1^{(m)}, \cdots, x_k^{(m)}\right)} \end{bmatrix} \quad (14.189)$$

Then have $\boldsymbol{y} = X\boldsymbol{\beta} = \boldsymbol{\varepsilon}$. To minimize $| \boldsymbol{\varepsilon}|^2$ will be the same as solving the normal equation $X^T X\boldsymbol{\beta} = X^T\boldsymbol{y}$ of $X\boldsymbol{\beta} = \boldsymbol{y}$.

Example 14.2.1: Find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the data points $(2, 1), (5, 2), (7, 3)$ and $(8, 3)$.

15. Lecture 15 - Complex eigenvalues

15.1. Complex numbers

Theorem 15.1.1: Consider quadratic equation $@^2 + bt + c = 0$, the quadratic formula reads $t = \frac{-b \pm \sqrt{\Delta}}{2a}$, where $\Delta = b^2 - 4ac$ is the discriminant.

Let's review some general stuff about complex numbers.

Definition 15.1.1: The set of complex numbers $\mathbb{C}$ is the defined to be $\{z = a + bi | a, b \in \mathbb{R}\}$, with $i = \sqrt{-1}$ so that $i^2 = -1$. We call $\mathrm{Re}z = a$ the real part of z and $\mathrm{Im}z = b$ to be the imaginary part of z. We call $r = |z| = \sqrt{a^2 + b^2}$ the modulus (or absolute value) of z, and the angle

$$\varphi \tag{15.190}$$

between z and the real axis the argument of z.

Definition 15.1.2: There is a natural identification between the complex numbers and the plane $\mathbb{R}^2$ via $z = a + bi \rightsquigarrow (a, b)$ (this is why the set of complex numbers is often called the complex plane). We can define

- Addition via $(a + bi) + (c + di) = (a + c) + (b + d)i$
- Multiplication via $(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$.
- Conjugate as $\overline{\mathrm{z}} = a - bi$.

Remark: Through this identification it is easy to see that $a = r\cos\varphi$, $b = r\sin\varphi$, and we have Euler's formula $z = re^{\{i\varphi\}} = r\cos\varphi = i\sin\varphi$.

Remark: If we have a complex-valued matrix A, then the conjugation is defined entrywise (Note that column vectors are matrices). If we write a complex valued matrix $A = \mathrm{Re}A + i\mathrm{Im}A$, then the conjugation would be $\overline{\mathrm{A}} = \mathrm{Re}A - i\mathrm{Im}A$. For example, if $A = \begin{bmatrix} 1+2i & 3 \\ -1-i & i \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} + i\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$, then $\overline{\mathrm{A}} = \begin{bmatrix} 1-2i & 3 \\ -1+i & -i \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} - i\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$. It is easy to check that $\overline{\mathrm{A+B}} = \overline{\mathrm{A}} + \overline{\mathrm{B}}$, $\overline{\mathrm{c}}\ \overline{\mathrm{A}} = \overline{\mathrm{cA}}$, $\overline{\mathrm{A}}\ \overline{\mathrm{B}} = \overline{\mathrm{AB}}$.

15.2. Eigen-decomposition over the complex numbers

Question: $A = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$, can we diagonalize it?

Answer: First we compute its characteristic polynomial $t^2 - 4t + 5$, an then we can solve the quadratic equation as follows

$$\Delta = (-4)^2 - 4 \cdot 1 \cdot 5 = -4, t = \frac{-(-4) \pm \sqrt{-4}}{2} = -(-4) \pm \frac{2i}{2} = 2 \pm i \tag{15.191}$$

Hence the eigenvalues are $\lambda_1 = 2 + i, \lambda_2 = 2 - i$. The eigenvector for λ_1 can be computed via So we get $\boldsymbol{v}_1 = \begin{bmatrix} -1+i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} i \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} i = \mathrm{Re}\boldsymbol{v}_1 + i\mathrm{Im}\boldsymbol{v}_1$. The eigenvector for λ_2 can be computed via So we get $\boldsymbol{v}_2 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -i \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} i = \mathrm{Re}\boldsymbol{v}_2 + i\mathrm{Im}\boldsymbol{v}_2$. Now we may realize that we have λ_1 and λ_2 are conjugate of each other. $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are conjugate of each other.

Theorem 15.2.1: In general, if A is a 2 by 2 real-valued matrix with complex eigenvalues (characteristic polynomial have complex roots, no real roots), then they are conjugate of each other, we can write them as $\lambda = a - bi$ and $\overline{\text{lambda}} = a + bi$, suppose the eigenvector for

$$\lambda \tag{15.192}$$

is $\boldsymbol{v} = \text{Re}\boldsymbol{v} + i\text{Im}\boldsymbol{v}$, then the eigenvector for $\overline{\text{lambda}}$ will be $\overline{\text{bold(v)}} = \text{Re}\boldsymbol{v} - i\text{Im}\boldsymbol{v}$. If we write $P = \begin{bmatrix} \text{Re}\boldsymbol{v} & \text{Im}\boldsymbol{v} \end{bmatrix}$, and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then $AP = PC$, therefore we get decomposition $A = PCP^{-1}$.

Proof: We have $A\boldsymbol{v} = \lambda\boldsymbol{v}$ (and hence by conjugation we have $A\overline{\text{bold(v)}} = A\overline{\text{bold(v)}} = \overline{\text{Abold(v)}} = \overline{\text{A bold(v)}} = \overline{\text{lambdabold(v)}}$, note that here A is real-valued, so $\overline{\text{A}} = A$), we can rewrite this as

$$(A\text{Re}\boldsymbol{v}) + i(A\text{Im}\boldsymbol{v}) = A(\text{Re}\boldsymbol{v} + i\text{Im}\boldsymbol{v}) = A\boldsymbol{v} = \lambda\boldsymbol{v} = (a - bi)(\text{Re}\boldsymbol{v} + i\text{Im}\boldsymbol{v}) = (a\text{Re}\boldsymbol{v} + b\text{Im}\boldsymbol{v}) + i(a\text{Im}\boldsymbol{v} - b\text{Re}\boldsymbol{v}) \tag{15.193}$$

Looking at its real and imaginary parts we conclude that

$$A\text{Re}\boldsymbol{v} = a\text{Re}\boldsymbol{v} + b\text{Im}\boldsymbol{v}, \quad A\text{Im}\boldsymbol{v} = a\text{Im}\boldsymbol{v} - b\text{Re}\boldsymbol{v} \tag{15.194}$$

This is precisely $AP = PC$ ■

Remark: Matrix C is special in the following sense (it can be decomposed as a composition of rotation and scaling)

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r\cos\varphi & -r\sin\varphi \\ r\sin\varphi & r\cos\varphi \end{bmatrix} = r\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \tag{15.195}$$

Here we suppose $\overline{\text{lambda}} = a + bi = re^{\{i\varphi\}} = r\cos\varphi + ir\sin\varphi$ so that $a = r\cos\varphi$, $b = r\sin\varphi$.

Example 15.2.1: In the previous example, we could take $\lambda = 2 - i$ so that $a = 2, b = 1$, $\boldsymbol{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + i\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ so that $\text{Re}\boldsymbol{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\text{Im}\boldsymbol{v} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, and we have the decomposition

$$A = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} = PCP^{-1} \tag{15.196}$$

Exercise 15.2.1: Suppose $A = \begin{bmatrix} -1 & -1 \\ 5 & -5 \end{bmatrix}$. Find an invertible matrix P and a matrix C of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ such that $A = PCP^{-1}$

16. Lecture 16 - Orthogonal diagonalization

16.1. Orthogonal diagonalization

Theorem 16.1.1: Suppose A is a symmetric ($A^T = A$) real-valued matrix, and $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ are λ_1-eigenvector, λ_2-eigenvectors respectively. Then $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$

Proof: Consider

$$\begin{aligned} \lambda_1(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) &= (\lambda_1 \boldsymbol{v}_1)^T \boldsymbol{v}_2 = (A\boldsymbol{v}_1)^T \boldsymbol{v}_2 \\ &= \boldsymbol{v}_1^T A^T \boldsymbol{v}_2 = \boldsymbol{v}_1^T A \boldsymbol{v}_2 = \boldsymbol{v}_1 \cdot (\lambda_2 \boldsymbol{v}_2) = \lambda_2(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) \end{aligned} \tag{16.197}$$

We get $(\lambda_1 - \lambda_2)(\boldsymbol{v}_1 \cdot \boldsymbol{v}_2) = 0$, since $\lambda_1 - \lambda_2 \neq 0$, $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$ ■

Theorem 16.1.2: Suppose A is a symmetric real-valued matrix, then the eigenvalues are also real.

Proof: Suppose

$$\lambda \tag{16.198}$$

is an eigenvalue of A, then there exists some

$$\lambda \tag{16.199}$$

-eigenvector such that $A\boldsymbol{v} = \lambda \boldsymbol{v}$, and that the length of the complex vector $\boldsymbol{v}$ is $|\boldsymbol{v}|^2 = \overline{\text{bold(v)}} \cdot \boldsymbol{v}$ is a positive real number (Note that for a complex number $z = a + bi$, $|z|^2 = a^2 + b^2 = (a + bi)(a - bi) = z\overline{\text{z}}$). Since A is symmetric and real-valued, $\overline{\text{A}}^T = A$. We have

$$\overline{\text{lambda}}\ |\boldsymbol{v}|^2 = \overline{\text{A bold(v)}}^T \boldsymbol{v} = \overline{\text{bold(v)}}^T \overline{\text{A}}^T \boldsymbol{v} = \overline{\text{bold(v)}}^T A \boldsymbol{v} = \lambda\ |\boldsymbol{v}|^2 \tag{16.200}$$

Which implies that $(\lambda - \overline{\text{lambda}})|\boldsymbol{v}|^2 = 0$, so

$$\lambda = \overline{\text{lambda}} \tag{16.201}$$

, i.e.

$$\lambda \tag{16.202}$$

is real-valued. ■

Fact 16.1.1: A symmetric real-valued matrix is diagonalizable.

Theorem 16.1.3: Suppose A is a symmetric real-valued matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$ (maybe repeated). A can be orthogonal diagonalized as $A = PDP^T$, where D is the diagonal matrix $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, and P is an orthogonal matrix.

Proof: To get an orthonormal basis for each eigenspace Nul $(\lambda I - A)$, you just need to find an arbitrary basis, and then apply Gram-Schmidt process to get an orthonormal basis, then by Theorem Theorem 16.1.1, the set of eigenvectors form an orthonormal basis for $\mathbb{R}^n$, assume they are $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_n$, in corresponds to eigenvalues $\lambda_1, \cdots, \lambda_n$, then we get orthogonal diagonalization

$$A = PDP^{-1} = PDP^T \tag{16.203}$$

Here $P = [\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n]$ is an orthogonal basis and thus $P^{-1} = P^T$. ■

Example 16.1.1: Suppose $A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$, the characteristic polynomial is $(t-3)(t-8)$, so we have eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 8$, we can then find the eigenvectors $\boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, we may realized that $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$, i.e. $\boldsymbol{v}_1$ is orthogonal to $\boldsymbol{v}_2$, we can further normalize them into $\boldsymbol{u}_1 = \frac{\boldsymbol{v}_1}{|}\boldsymbol{v}_1| = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$ and $\boldsymbol{u}_2 = \frac{\boldsymbol{v}_2}{|}\boldsymbol{v}_2| = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$. So we have orthogonal diagonalization

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} = PDP^T \tag{16.204}$$

16.2. Spectral decomposition

Theorem 16.2.1: Suppose A is a symmetric real-valued matrix, and $A = PDP^T$ is its orthogonal diagonalization, then we have the so-called spectral decomposition

$$A = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T \tag{16.205}$$

Proof:

$$A = PDP^T = [\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} = \begin{bmatrix} \lambda_1 \boldsymbol{u}_1 & & \\ & \ddots & \\ & & \lambda_n \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \\ \vdots \\ \boldsymbol{u}_n^T \end{bmatrix} = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T \tag{16.206}$$

■

Remark: Recall that if $\boldsymbol{u}$ is of unit length, then $\text{Proj}_{\boldsymbol{u}} \boldsymbol{x} = \boldsymbol{u}\boldsymbol{u}^T \boldsymbol{x}$, so

$$A\boldsymbol{x} = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^T \boldsymbol{x} + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^T \boldsymbol{x} \tag{16.207}$$

You can think of this as decompose the matrix transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ into the sum of scaled orthogonal projections.

17. Lecture 21 - Quadratic forms

Let's briefly talk about quadratic forms: $ax_1 + 2bx_1x_2 + cx_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boldsymbol{x}^T A \boldsymbol{x}$. In Example Example 16.1.1, $\boldsymbol{x}^T A \boldsymbol{x}$ is the quadratic form $7x_1^2 + 4x_1x_2 + 4x_2^2$, note that $\boldsymbol{y} = P^T \boldsymbol{x}$ gives change of (orthonormal) coordinates (actually differ by a rotation) $\begin{cases} y_1 = -\frac{1}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_2 \\ y_2 = \frac{2}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_2 \end{cases}$, then the quadratic form becomes $\boldsymbol{x}^T A \boldsymbol{x} = \boldsymbol{x}^T P D P^T \boldsymbol{x} = (P^T \boldsymbol{x})^T D (P^T \boldsymbol{x}) = \boldsymbol{y}^T D \boldsymbol{y} = 3y_1^2 + 8y_2^2$, note this is without cross term y_1y_2.

18. Lecture 22 - Singular value decomposition

19. Appendix

Theorem 19.1: The RREF of a matrix A is unique.

Proof: The linear dependences of columns of A is presevered under row elementary operations, thus the pivot columns are unique. And for the same reason, we know how all other entries look like. ■

20. Online Assignments

20.1. Online Assignment 1

Problem 20.1.1: Rewrite the following linear systems as augmented matrices and then solve them, show all your work

- $$\begin{cases} 5x_1 + x_2 = 2 \\ 3x_1 - x_2 = 6 \end{cases} \tag{20.208}$$
- $$\begin{cases} x_1 + x_2 + x_3 = 6 \\ x_1 - x_2 + x_3 = 2 \\ -x_1 + x_2 + x_3 = 4 \end{cases} \tag{20.209}$$

Problem 20.1.2: How many solutions does the following linear systems of equations have

- $$\begin{cases} 5x_1 + 7x_2 = 3 \\ -10x_1 - 14x_2 = -3 \end{cases} \tag{20.210}$$

- $$\begin{cases} 2x_1 - x_2 = 4 \\ x_1 - \frac{1}{2}x_2 = 2 \end{cases} \tag{20.211}$$

Problem 20.1.3: Consider the following matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & -1 & 2 & 1 \\ 0 & 0 & 0 & 2 & 4 \end{bmatrix} \tag{20.212}$$

- Which columns are the pivot columns of A?
- Write down the RREF of the this matrix

Problem 20.1.4: Determine which of the following statements are true

- The following matrix is of row reduced echelon form

$$\begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix} \tag{20.213}$$

- The following two matrices are equivalent

$$\begin{bmatrix} 1 & -1 & 1 & 3 & 2 \\ 2 & 4 & 1 & 2 & 4 \\ 1 & 1 & -3 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & -1 & 2 & 1 \end{bmatrix} \tag{20.214}$$

Problem 20.1.5: Determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system

$$\begin{bmatrix} 1 & h & 1 \\ 2 & 4 & 4 \end{bmatrix} \tag{20.215}$$

Problem 20.1.6: Do the three lines $x_1 - 4x_2 = 1$, $2x_1 - x_2 = -3$, and $-x_1 - 3x_2 = 4$ have a common point of intersection? Explain.

20.2. Online Assignment 2

Problem 20.2.1: Consider the following statements

- For any four distinct vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ in $\mathbb{R}^3$, $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\} = \mathbb{R}^3$
- Suppose we know that the augmented matrix of the linear system

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = d_1 \\ b_1x_1 + b_2x_2 + b_3x_3 = d_2 \\ c_1x_1 + c_2x_2 + c_3x_3 = d_3 \end{cases} \tag{20.216}$$

has two pivot columns, then how many solutions could the linear system have?

- Consider matrix equation $A\boldsymbol{x} = \boldsymbol{0}$ where A is a 3 by 4 matrix, then it always has more than one solution (obviously $\boldsymbol{x} = \boldsymbol{0}$ will be a solution)

Problem 20.2.2: Answer the following questions

- Determine whether $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ is in the span of $\left\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -1 \\ -2 \\ -4 \end{bmatrix} \right\}$, why?
- Assume $\boldsymbol{w} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 4 \\ -4 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$. Find constants c_1, c_2 such that $\boldsymbol{w} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + 2\boldsymbol{v}_3$. Show your work on how you found c_1, c_2.
- Suppose

$$A = \begin{bmatrix} 1 & 3 & 0 & 3 \\ -1 & -1 & -1 & 1 \\ 0 & -4 & 2 & -8 \\ 2 & 0 & 3 & -1 \end{bmatrix} \tag{20.217}$$

Is it true that for any vector $\boldsymbol{b}$ in $\mathbb{R}^4$, matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ always has solution(s)? If it is, please give your reason. If it is not, please find one such $\boldsymbol{b}$ and justify your answer.

20.3. Online Assignment 3

Problem 20.3.1: Suppose $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are vectors in $\mathbb{R}^4$, if $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is linearly independent, $\{\boldsymbol{v}_2, \boldsymbol{v}_3\}$ is linearly independent, and $\{\boldsymbol{v}_1, \boldsymbol{v}_3\}$ is linearly independent, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is linearly independent

Solution: False. For example $\left\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

Problem 20.3.2: Suppose A is a m by n matrix, and the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ always has solution for any vector $\boldsymbol{b}$ in $\mathbb{R}^m$, the columns of A are linearly independent

Solution: False. There could be free variables

Problem 20.3.3: Solve the linear system $\begin{cases} x_1+x_2+x_3+x_4=1 \\ 2x_1-x_2+x_3-x_4=-1 \end{cases}$ and express its solution set in parametric vector form

Problem 20.3.4: Suppose $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_3 = \begin{bmatrix} -2 \\ 3 \\ -11 \end{bmatrix}$, $\boldsymbol{v}_4 = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}$. Is $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ linearly independent? If so, please give your reason, if not, please find a linear dependence (i.e. some linear combination $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_2\boldsymbol{v}_3 + c_4\boldsymbol{v}_4 = 0$, c_1, c_2, c_3, c_4 not all zero)

Problem 20.3.5: Suppose linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ rotates the plane $\mathbb{R}^2$ counter-clockwise by 120°, what is the standard matrix for the this linear transformation?

Problem 20.3.6: Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $T(\boldsymbol{x}) = A\boldsymbol{x}$, where $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix}$. What is n and m? Find the preimage of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ under T. Is T one-to-one? Is T onto?

20.4. Online Assignment 4

Problem 20.4.1: Suppose $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$.

- Determine if A is invertible, if not, explain why, if it is, find the inverse matrix A^{-1}
- Find A^T, the transpose of A. Is A^T invertible? If yes, please evaluate $\left(A^T\right)^{-1}$

Please show all your work.

Problem 20.4.2: Suppose $A = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$ is the standard matrix for the linear transformation of rotating 120° counter-clockwise. Evaluate $A^{\{24\}}$ and explain why.

Problem 20.4.3: We say n is the order of a square matrix A if n is the smallest positive integer such that $A^n = I$, where I is the identity matrix. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation of reflecting over x_1-axis, and A is the standard matrix of T, the order of A is

Solution: Note that $T \circ T$ is reflecting over x_2-axis twice, which amounts to doing nothing, so A^2 as the standard matrix of $T \circ T$ is the identity matrix, i.e. the order of A is 2.

Problem 20.4.4: If A, B are both invertible, then $A + B$ is also invertible

Solution: False. For example we can take $B = -A = -I$, then $A + B = 0$ which is not invertible.

Problem 20.4.5: We call $q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$ a quadratic form, a, b, c are constants. Suppose $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, try to rewrite $q(x_1, x_2)$ as the form of a matrix multiplication $\boldsymbol{x}^T A \boldsymbol{x}$, where A is a symmetric matrix (i.e. $A^T = A$). Please find A.

Solution: We should choose $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, then

$$\boldsymbol{x}^T A \boldsymbol{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 + bx_2 & bx_1 + cx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = ax_1^2 + 2bx_1x_2 + cx_2^2 = q(x_1, x_2) \tag{20.218}$$

Problem 20.4.6: Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation defined by $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 - 2x_2 + x_3 \\ 2x_2 + 3x_3 \\ x_1 - x_3 \end{bmatrix}$. Please find T^{-1}, then standard matrix for T^{-1}. Is T^{-1} onto? Is T^{-1} is one-to-one? Show all your work.

Problem 20.4.7: An affine transformation is a mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$, Note that $A\boldsymbol{x} + \boldsymbol{b} = \begin{bmatrix} A & \boldsymbol{b} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix}$. If we use partitioned matrix and deploying the trick of adding“ 1 at the end of the coordinates, then T may be interpreted as a matrix transformation $\begin{bmatrix} T(\boldsymbol{x}) \\ 1 \end{bmatrix} = \begin{bmatrix} A\boldsymbol{x} + \boldsymbol{b} \\ 1 \end{bmatrix} = \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix}$

item Suppose $T_1(\boldsymbol{x}) = A_1\boldsymbol{x} + \boldsymbol{b}_1$ and $T_2(\boldsymbol{x}) = A_2\boldsymbol{x} + \boldsymbol{b}_2$ are both affine transformations. What is the composition $T_2 \circ T_1$? How should you deploy the trick?

- Suppose $T(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ is an affine transformation and A is invertible. What is T^{-1}? How should you deploy the trick?

Solution:

- $(T_2 \circ T_1)(\boldsymbol{x}) = T_2(T_1(\boldsymbol{x})) = A_2(T_1(\boldsymbol{x})) + \boldsymbol{b}_2 = A_2(A_1\boldsymbol{x} + \boldsymbol{b}_1) + \boldsymbol{b}_2 = A_2A_1\boldsymbol{x} + (A_2\boldsymbol{b}_1 + \boldsymbol{b}_2)$, this can be explained use the trick as follows

$$
\begin{bmatrix} A_2 & \boldsymbol{b}_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1 & \boldsymbol{b}_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A_2 A_1 & A_2 \boldsymbol{b}_1 + \boldsymbol{b}_2 \\ 0 & 1 \end{bmatrix} \tag{20.219}
$$

- The inverse should be $T^{-1}(\boldsymbol{y}) = A^{-1}\boldsymbol{y} - A^{-1}\boldsymbol{b}$, via the trick we can interpreted it as

$$
\begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}\boldsymbol{b} \\ 0 & 1 \end{bmatrix} \tag{20.220}
$$

Since

$$
\begin{bmatrix} A^{-1} & -A^{-1}\boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} I & \boldsymbol{0} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}\boldsymbol{b} \\ 0 & 1 \end{bmatrix} \tag{20.221}
$$

20.5. Online Assignment 5

Problem 20.5.1: Suppose $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 0 \end{bmatrix}$. Use cofactor expansion of A across the last row to evaluate the determinant of A

Solution:

$$
,| \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 0 \end{bmatrix} | = (-1) \cdot (-1)^{\{3+1\}} | \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} | + 2 \cdot (-1)^{\{3+2\}} | \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} | + 0;, = (-1)(2 \cdot 3 - 2 \cdot 1) + (-2)(1 \cdot 3 - 2 \cdot 2) = -2 \tag{20.222}
$$

Problem 20.5.2: Compute the determinant of $A = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 0 & 2 & 0 & 0 \\ 3 & 5 & 2 & 3 \\ 2 & 1 & 0 & -1 \end{bmatrix}$. Is A invertible?

Problem 20.5.3: Consider the parallelgram P with vertices $(-2,-3), (-1,1), (2,4), (3,8)$, use determinants to evaluate the area of P

Solution: Name these four points p_1, p_2, p_3, p_4, and the vectors $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\boldsymbol{a}_2 = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$. And we consider $A = [\boldsymbol{a}_1 \ \ \boldsymbol{a}_2] = \begin{bmatrix} 1 & 4 \\ 4 & 7 \end{bmatrix}$, then area of P would be

$$
|\det A| = |(1 \cdot 7 - 4 \cdot 4))| = 9 \tag{20.223}
$$

Problem 20.5.4: $\det(A - B) = \det A - \det B$

Solution: This is false, for example, we could just take $A = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = -I$, then $4 = \det(2I) = \det(A - B) \neq \det A - \det B = 1 - 1 = 0$

Problem 20.5.5: If A is a 3 by 3 matrix, then $\det(2A) = 8(\det\ A)$

Solution: This is true.

Problem 20.5.6: Suppose A, B are both 3 by 3 matrices, and $\det\ A = 2$, $\det\ B = \frac{1}{3}$, then the determinant of $A^T B^{-1}$ is

Solution: $\det(A^T B^{-1}) = \det(A^T)\det(B^{-1}) = (\det\ A)(\det\ B)^{-1} = 2 \cdot 3 = 6.$

Problem 20.5.7: Suppose A is a 3 by 3 matrix with entries integers, and $A^3 = I$ is the identity matrix. Then the determinant of A has to be

Solution: $\det\ A$ must be some real number. $1 = \det\ I = \det(A^3) = (\det\ A)^3 \Rightarrow \det\ A =$ `sqrt`$[3]\{1\} = 1.$

20.6. Online Assignment 6

Problem 20.6.1: Suppose $A = \begin{bmatrix} 1 & 2 & 1 & 2 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 3 & 1 & -2 \end{bmatrix}$

- Find a basis for the null space of A
- Find a basis for the column space of A
- Find a basis for the row space of A

Solution: First realize

$$A \xrightarrow{\text{R2}\leftrightarrow\text{R3}} \begin{bmatrix} 1 & 2 & 1 & 2 & -1 & 3 \\ 0 & 0 & 1 & 3 & 1 & -2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -1 & 0 & 3 \\ 0 & 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{20.224}$$

Hence the solution in parametric form is $x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. And the pivot columns are the 1st, 3rd and 5th columns

- A basis for Nul$A$ could be $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$
- A basis for Col$A$ could be $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$
- A basis for Row$A$ could be

$$\{[1\ \ 2\ \ 0\ \ -1\ \ 0\ \ 3], [0\ \ 0\ \ 1\ \ 3\ \ 0\ \ -1], [0\ \ 0\ \ 0\ \ 0\ \ 1\ \ -1]\} \tag{20.225}$$

Problem 20.6.2: Suppose $A = \begin{bmatrix} 3 & -1 & -5 \\ 1 & 1 & -1 \\ -2 & 2 & 4 \end{bmatrix}$

- Determine whether $\boldsymbol{u} = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}$ is in the null space of A. Explain your reasoning.
- Determine whether $\boldsymbol{b} = \begin{bmatrix} 1 \\ -3 \\ -4 \end{bmatrix}$ is in the column space of A. Explain your reasoning.

Solution:

- $\boldsymbol{u} \in \mathrm{Nul}A$ since $A\boldsymbol{u} = \mathbf{0}$
- $\boldsymbol{b} \in \mathrm{Col}A$ since linear system $A\boldsymbol{x} = \boldsymbol{b}$ is consistent

Problem 20.6.3: We say a square matrix A is anti-symmetric if $A^T = -A$. Denote the set of 3×3 anti-symmetric matrices V.

- Show that V is a vector space.
- What is the dimension of V?
- Find a basis of V.
- Show that

$$\mathcal{B} = \left\{ B_1 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, B_3 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix} \right\} \tag{20.226}$$

form a basis for V.

20.7. Online Assignment 7

Problem 20.7.1: Suppose $\boldsymbol{v}$ is an eigenvector for matrices A and B, then $\boldsymbol{v}$ is an eigenvector for $A + B$ and AB.

Solution: This is true. Since $\boldsymbol{v}$ is an eigenvector for both A and B, there exist eigenvalues λ_1, λ_2 such that $A\boldsymbol{v} = \lambda_1\boldsymbol{v}$, $B\boldsymbol{v} = \lambda_2\boldsymbol{v}$, so we have $(A+B)\boldsymbol{v} = A\boldsymbol{v} + B\boldsymbol{v} = \lambda_1\boldsymbol{v} + \lambda_2\boldsymbol{v} = (\lambda_1 + \lambda_2)\boldsymbol{v}$. In other words, $\boldsymbol{v}$ is an eigenvector for $A + B$ with eigenvalue $\lambda_1 + \lambda_2$. Similarly, we also have $AB\boldsymbol{v} = A(\lambda_1\boldsymbol{v}_1) = \lambda_2 A\boldsymbol{v} = \lambda_1\lambda_2\boldsymbol{v}$. In other words, $\boldsymbol{v}$ is an eigenvector for AB with eigenvector $\lambda_1\lambda_2$

Problem 20.7.2: Suppose $t + 3t^2 - 2t^3$ is the characteristic polynomial of a 3 by 3 matrix, then A is not invertible.

Solution: This is true. Note that $t = 0$ is a root of the characteristic polynomial, so the null space of A is not trivial, A is not invertible.

Problem 20.7.3:

Problem 20.7.4: If Nul$A$ is 2-dimensional, then 0 is an eigenvalue of A.

Solution: This is true. Since if Nul$A$ is 2-dimensional, then Nul$A$ is non-trivial, thus 0 is an eigenvalue of A

Problem 20.7.5: Is $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 4 & 3 \end{bmatrix}$? If so, find the corresponding eigenvalue, if not, please explain why.

Solution: Note that $\begin{bmatrix} 2 & 0 & 0 \\ 3 & 2 & 0 \\ 0 & 4 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$. so $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 3

Problem 20.7.6: Find all eigenvalues of $A = \begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$. Please show all your work.

Solution: Therefore eigenvalues for A are 3,5,10.

Problem 20.7.7: Assume that A is similar to an upper triangular matrix U, then det A is the product of all its eigenvalues (counting multiplicity). Please explain why.

20.8. Online Assignment 8

Problem 20.8.1: Suppose $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, determine whether A is diagonalizable. If it is, please find a diagonalization. If not, please explain why.

Problem 20.8.2: Suppose $A = \begin{bmatrix} -6 & 8 \\ -4 & 6 \end{bmatrix}$, Please evaluate $A^{\{101\}}$, show all your work.

Problem 20.8.3: Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation, $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$, $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$ are two different bases for $\mathbb{R}^3$. Determine whether the following is possible.

- $[T]_{\{\mathcal{B}\}} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & -1 & -2 \\ 2 & -1 & 3 \end{bmatrix}$ and $[T]_{\{\mathcal{C}\}} = \begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & 6 \\ 0 & 3 & 8 \end{bmatrix}$
- $[T]_{\{\mathcal{B}\}} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 3 & 4 \end{bmatrix}$ and $[T]_{\{\mathcal{C}\}} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 4 & 0 \\ 3 & -2 & 4 \end{bmatrix}$

Problem 20.8.4: Suppose A is similar to B.

- Could you conclude that $3A$ is similar to $3B$. If you can, please give your reasons, if not, please find a counter-example.
- Could you conclude that A^{-1} is similar to B^{-1}. If you can, please give your reasons, if not, please find a counter-example.

Solution: Since A is similar to B, we may assume $A = PBP^{-1}$.

- $3A = 3PBP^{-1} = P(3B)P^{-1}$ is similar.
- $A^{-1} = \left(PBP^{-1}\right)^{-1} = PB^{-1}P^{-1}$ is similar.

20.9. Online Assignment 9

Problem 20.9.1: Determine whether the following statements are true

- $|\boldsymbol{u}|^2 + |\boldsymbol{v}|^2 = |\boldsymbol{u} + \boldsymbol{v}|^2$ if and only if $\boldsymbol{u}, \boldsymbol{v}$ are orthogonal.
- If W is a subspace of $\mathbb{R}^n$, and vector $\boldsymbol{v}$ is orthogonal to both W and $W^\perp$, then $\boldsymbol{v} = \boldsymbol{0}$.
- If $W = \text{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ with $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$ linearly independent, and if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthogonal set in W, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for W.

Problem 20.9.2: Suppose $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & -1 \\ 0 & 2 & 4 \\ 3 & 5 & -1 \end{bmatrix}$, please find a basis for $(\text{Col}A)^\perp$.

Problem 20.9.3: Suppose we have $\mathcal{B} = \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, please justify that $\mathcal{B}$ is an orthogonal set, suppose $\boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, L is the subspace spanned by $\mathcal{B}$, compute the projection $\text{Proj}_L \boldsymbol{y}$ of $\boldsymbol{y}$ onto L.

Problem 20.9.4: Suppose $A = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$, please find an orthogonal basis for Col$A$.

20.10. Online Assignment 10

Problem 20.10.1: Determine whether the following statements are correct

- If A is symmetric and if vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ such that $A\boldsymbol{u} = \boldsymbol{u}$ and $\boldsymbol{v}$ is in Nul$A$, then $\boldsymbol{u} \cdot \boldsymbol{v} = \mathbf{0}$.
- Nul$A$ = Nul$A^T A$.

Solution:

- The eigenvalues for eigenvectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are 1 and 0, and A is symmetric, by Theorem Theorem 16.1.1, $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
- If $\boldsymbol{x} \in \text{Nul}A$, then $A\boldsymbol{x} = \mathbf{0}$, so $A^T A\boldsymbol{x} = \mathbf{0}$. If $\boldsymbol{x} \in \text{Nul}\left(A^T A\right)$, then $A^T A\boldsymbol{x} = \mathbf{0}$, so $0 = \boldsymbol{x}^T A^T A\boldsymbol{x} = |A\boldsymbol{x}|^2$, so $A\boldsymbol{x} = 0$.

Problem 20.10.2: Find the least-squares solution(s) to $\begin{bmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -3 \end{bmatrix}$.

Solution: Let's denote $A = \begin{bmatrix} 1 & 5 \\ 3 & 1 \\ -2 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 4 \\ -2 \\ -3 \end{bmatrix}$, then $A^T A = \begin{bmatrix} 14 & 0 \\ 0 & 42 \end{bmatrix}$, and $A^T \boldsymbol{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$, then we get the least -square solution $h@\{\boldsymbol{x}\} = \left(A^T A\right)^{-1} A^T \boldsymbol{b} = \begin{bmatrix} \frac{2}{7} \\ \frac{1}{7} \end{bmatrix}$

Problem 20.10.3: Orthogonal diagonalize the matrix $\begin{bmatrix} 3 & 4 \\ 4 & 9 \end{bmatrix}$.

Problem 20.10.4: Suppose $A = \begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix}$.

- Please find the eigenvalues of A.
- Please find the eigenvectors of A.
- Please write A as matrix multiplication PCP^{-1}, where C is of the form $\begin{bmatrix} a & -b \\ b & c \end{bmatrix}$.

21. Exams

21.1. Exam 1

Problem 21.1.1:

21.2. Exam 2

21.3. Final

Index