Outline of MA265

Instructor: Xu Wang

This is an outline of MA265 Linear Algebra. All the definitions can be found in the textbook and are omitted here for brevity.

1 Chapter 1

1.1 Systems of linear equations

- **Def**: linear equation

  Ex: Are they linear equations?

  $\sqrt{3}x_1 + x_2 = 1, \qquad \sqrt{x_1} + x_2 = 2, \qquad x_1x_2 + x_3 = 1$

- **Def**: linear system

  Ex: Construct a linear system according to the following problem: An unknown amount of chickens and rabbits were locked in a cage. The total amount of them is 6, and there are 16 feet in total. What is the amount of chickens and rabbits, respectively? (Hint: assume that there are x_1 chickens and x_2 rabbits.)

$$\begin{cases} x_1 + x_2 = 6 \\ 2x_1 + 4x_2 = 16 \end{cases} \xLeftrightarrow{\text{Collect all coefficients}} \begin{bmatrix} 1 & 1 & 6 \\ 2 & 4 & 16 \end{bmatrix} \text{ (augmented matrix)} \tag{1}$$

  To get the solution

$$\begin{cases} x_1 = * \\ x_2 = ** \end{cases} \xLeftrightarrow{\text{corresponding matrix}} \begin{bmatrix} 1 & 0 & * \\ 0 & 1 & ** \end{bmatrix}, \tag{2}$$

  we only need to transform the matrix in (1) into the form in (2).

♣ **Elementary row operations**

1. Interchange two rows.
2. Multiply a row by a scalar.
3. Replace a row by the sum of itself and a multiple of another row.

Ex:

$$\begin{bmatrix} 1 & 1 & 6 \\ 2 & 4 & 16 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 6 \\ 0 & 2 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 6 \\ 0 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \Longleftrightarrow \begin{cases} x_1 = 4 \\ x_2 = 2 \end{cases} \quad \text{One solution}$$

Ex:

$$\begin{bmatrix} 1 & 1 & 6 \\ 2 & 2 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 6 \\ 0 & 0 & 0 \end{bmatrix} \iff \begin{cases} x_1 = 6 - x_2 \\ x_2 \text{ is free} \end{cases} \quad \text{Infinitely many solutions}$$

Ex:

$$\begin{bmatrix} 1 & 1 & 6 \\ 1 & 1 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 6 \\ 0 & 0 & 2 \end{bmatrix} \iff \begin{cases} x_1 + x_2 = 6 \\ 0 = 2 \end{cases} \quad \text{No solution}$$

- **Def**: solution/solution set

  1. only one solution 2. infinitely many solutions 3. no solution

- **Def**: row equivalent

  Properties: systems are equivalent $\iff$ corresponding matrices are row equivalent

  $\iff$ they have the same solution set

1.2 Row reduction and echelon forms

- **Def**: Nonzero row/column

  Def: leading entry

- **Def**: echelon form (3 conditions)/reduced echelon form (5 conditions)

  Ex: Find echelon forms and the reduced echelon form of the original matrix:

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

- **Thm**: Each matrix may be row equivalent to more than one matrix in echelon form, but is row equivalent to only one matrix in reduced echelon form.

- **Def**: pivot position/pivot column

♣ **Thm**: A linear system is consistent if and only if its rightmost column is not a pivot column.

Ex: Recall examples in Lesson 1.1:

$$\begin{bmatrix} \textcircled{1} & 1 & 6 \\ 2 & \textcircled{4} & 16 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 0 & 4 \\ 0 & \textcircled{1} & 2 \end{bmatrix} \quad \text{the rightmost column is NOT a pivot colum, so consistent}$$

$$\begin{bmatrix} ① & 1 & 6 \\ 2 & ② & 12 \end{bmatrix} \sim \begin{bmatrix} ① & 1 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$ the rightmost column is NOT a pivot column, co consistent

$$\begin{bmatrix} ① & 1 & 6 \\ 1 & 1 & ⑧ \end{bmatrix} \sim \begin{bmatrix} ① & 1 & 6 \\ 0 & 0 & ② \end{bmatrix}$$ the rightmost column is a pivot column, so inconsistent

- **Remark**: For a linear system:

  consistent + no free variable $\iff$ only one solution   e.g. the first matrix above

  consistent + free variable $\iff$ infinitely many solutions   e.g. the second one above

1.3 Vector equations

A linear system has the following equivalent expressions.

$$\begin{bmatrix} 1 & 1 & 6 \\ 2 & 4 & 16 \end{bmatrix} \xleftarrow{\text{row view}} \begin{cases} x_1 + x_2 = 6 \\ 2x_1 + 4x_2 = 16 \end{cases} \xrightarrow{\text{column view}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 4 \end{bmatrix} x_2 = \begin{bmatrix} 6 \\ 16 \end{bmatrix}$$

- **Def**: (column) vector

  1. Vectors in $\mathbb{R}^2$: $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

  (1) $\mathbf{u} = \mathbf{v}$ if and only if $u_1 = v_1$ and $u_2 = v_2$, e.g. $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

  (2) $\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$

  (3) $c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}$, c is a scalar

  2. Vectors in $\mathbb{R}^3$: $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$

  3. Vectors in $\mathbb{R}^n$: $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$

  Geometric description: Identify a geometric point (a, b) with a vector $\begin{bmatrix} a \\ b \end{bmatrix}$. Four vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$ and the origin could form a parallelogram.

- **Def**: linear combination

  Ex: For the vector equation $\begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 4 \end{bmatrix} x_2 = \begin{bmatrix} 6 \\ 16 \end{bmatrix}$, we have already known its solution $\begin{cases} x_1 = 4 \\ x_2 = 2. \end{cases}$ That is,

$$4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \end{bmatrix}, \quad \text{so } \begin{bmatrix} 6 \\ 16 \end{bmatrix} \text{ is a linear combination of } \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{and} \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

- **Thm**: Vector $\mathbf{y}$ is a linear combination of vectors $\mathbf{v}_1, \cdots, \mathbf{v}_p$

  $\iff$ The vector equation $\mathbf{v}_1 x_1 + \cdots \mathbf{v}_p x_p = \mathbf{y}$ has a solution

  $\iff$ The augmented matrix $\begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_p & \mathbf{y} \end{bmatrix}$ is consistent

- **Def**: Given vectors $v_1, \cdots, v_p$,

$$\begin{aligned} \text{Span}\{\mathbf{v}_1, \cdots, \mathbf{v}_p\} &= \{\text{all linear combinations of } \mathbf{v}_1, \cdots, \mathbf{v}_p\} \\ &= \{c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p : c_1, \cdots, c_p \text{ are scalars}\} \\ &= \text{subset spanned (generated) by vectors } \mathbf{v}_1, \cdots, \mathbf{v}_p \end{aligned}$$

  Geometric description:

  Span$\{\mathbf{u}\}$ denotes a straight line

  Span$\{\mathbf{u}, \mathbf{v}\}$ denotes a plane

1.4 Matrix equations $A\mathbf{x} = \mathbf{b}$

- **Def**: product between A and $\mathbf{x}$

  Ex: $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}_{2\times 3} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}_{3\times 1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \times 1 + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \times 2 + \begin{bmatrix} 3 \\ 4 \end{bmatrix} \times 3 = \begin{bmatrix} 14 \\ 20 \end{bmatrix}$

  Ex: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$: identity matrix

  Ex: For vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, we can rewrite $\mathbf{v}_1 + \mathbf{v}_2 - 2\mathbf{v}_3 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

- **Properties**: $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$, A is a matrix and $\mathbf{u}, \mathbf{v}$ are vectors

  $A(c\mathbf{u}) = cA\mathbf{u}$, c: scalar

  Ex: Let $A = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Then

  $$A(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

  $$A\mathbf{u} + A\mathbf{v} = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

- **Thm**: Let A be an $m \times n$ matrix with columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$ and $\mathbf{b} \in \mathbb{R}^m$.

  The solution set of $A\mathbf{x} = \mathbf{b} \Longleftrightarrow$ The solution set of $\mathbf{a}_1 x_1 + \cdots + \mathbf{a}_p x_p = \mathbf{b}$

  $\Longleftrightarrow$ The solution set of the system determined by the augmented matrix $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$

- **Question**: Determine if for each vector $\mathbf{b} \in \mathbb{R}^m$, $A\mathbf{x} = \mathbf{b}$ is consistent

  Ex: $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

  $\begin{bmatrix} 1 & 1 & b_1 \\ 2 & 2 & b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & b_1 \\ 0 & 0 & b_2 - 2b_1 \end{bmatrix}$ is consistent if and only if $b_2 - 2b_1 = 0$

  Ex: $A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

  $\begin{bmatrix} 1 & 1 & b_1 \\ 2 & 4 & b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & b_1 \\ 0 & 2 & b_2 - 2b_1 \end{bmatrix}$ is consistent for any $\mathbf{b}$

♣ **Thm**: The following statements are equivalent:

For each $\mathbf{b} \in \mathbb{R}^m$, $A\mathbf{x} = \mathbf{b}$ is consistent

$\Longleftrightarrow$ For each $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b}$ is a linear combination of $\mathbf{a}_1, \cdots, \mathbf{a}_n$

$\Longleftrightarrow$ $\mathbb{R}^m = \text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$

$\Longleftrightarrow$ A has a pivot position in every row

1.5 Solution sets of $A\mathbf{x} = \mathbf{b}$

- **Def**: A homogeneous linear system is in the form $A\mathbf{x} = \mathbf{0}$. It must be consistent with the trivial solution $\mathbf{x} = \mathbf{0}$.

  If $\mathbf{x} \neq \mathbf{0}$, it is called a nontrivial solution.

  Remark: $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions $\Longleftrightarrow$ $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions

$\iff$ $A\mathbf{x} = \mathbf{0}$ has free variables

Ex: $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix}$. Find all the solutions of $A\mathbf{x} = \mathbf{0}$.

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 0 \\ 1 & 2 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 0 & -1 & 0 \\ 0 & \textcircled{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \iff \begin{cases} x_1 = x_3 \\ x_2 = -2x_3 \\ x_3 = x_3 (\text{free}) \end{cases}$$

$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, x_3 can be chosen as any real numbers.

- **Def**: $\mathbf{x} = t\mathbf{v}$, $t \in \mathbb{R}$, is call the parametric vector form of the solution.

Ex: Find all solutions of $x_1 - x_2 - x_3 = 0$.

$$\begin{bmatrix} \textcircled{1} & -1 & -1 & 0 \end{bmatrix} \iff \begin{cases} x_1 = x_2 + x_3 \\ x_2 = x_2 \text{ (free)} \\ x_3 = x_3 \text{ (free)} \end{cases} \iff \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_3$$

Ex: Given $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Find matrix A such that $A\mathbf{x}_0 = \mathbf{0}$.

Suppose that x_3 is a free variable and all the solution can be written as $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x_3$.

Then $\begin{cases} x_1 = x_3 \\ x_2 = x_3 \\ x_3 = x_3 \text{ (free)} \end{cases} \iff \begin{cases} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \\ 0 = 0 \end{cases} \iff$ augmented matrix $\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

So we can choose $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

Ex: Find all the solutions of $A\mathbf{x} = \mathbf{b}$ with $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

$$\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 4 & 0 \\ 1 & 2 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & -1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 0 & -1 & -3 \\ 0 & \textcircled{1} & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \iff \begin{cases} x_1 = -3 + x_3 \\ x_2 = 2 - 2x_3 \\ x_3 = x_3 \text{ (free)} \end{cases}$$

All the solutions are in the form $\mathbf{x} = \begin{bmatrix} -3 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

Compare it with the first example on this page, we get the following Thm.

- **Thm**: Assume that $A\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{p}$. Then any solution of $A\mathbf{x} = \mathbf{b}$ has the form $\mathbf{x} = \mathbf{p} + \mathbf{v}$, where $\mathbf{v}$ is any solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$.

1.7 Linear independence

- **Def**: 1. linearly independent

  2. linearly dependent

  Ex: Determine if the columns of $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ are linearly dependent

  Augmented matrix $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 4 & 0 \\ 3 & 4 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -2 & -4 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 2 & 3 & 0 \\ 0 & \textcircled{1} & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

  There is infinitely many solutions for $A\mathbf{x} = \mathbf{0}$, so of course there is nontrivial ones, since there is one free variable. Thus, the columns of A are linear dependent.

  Ex: Determine if $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ are linear dependent.

  Method 1: consider the augmented matrix $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{0} \end{bmatrix}$ as above

  Method 2: note that $v_2 = 2v_1$, so they are linearly dependent. See also what follows.

- **Thm**: Vectors $\mathbf{v}_1, \cdots, \mathbf{v}_p$ are linearly dependent $\iff$ One of them is a linear combination of the others.

- **Thm**: Any set of vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

  Reason: Consider linear system $\mathbf{v}_1 x_1 + \cdots + \mathbf{v}_p x_p = \mathbf{0}$. There is p variables in total. There is at most n pivot variables since there is n equations. As a result, there is at least $p - n(> 0)$ free variables. So the system has nontrivial solutions, and thus the vectors are linearly dependent.

- **Thm**: Any set of vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ containing the zero vector is linearly dependent.

  Reason: Without loss of generality, we assume that $\mathbf{v}_1 = 0$. Then apparently

  $$\mathbf{v}_1 \cdot 1 + \mathbf{v}_2 \cdot 0 + \cdots + \mathbf{v}_2 \cdot 0 = \mathbf{0}$$

  is always true, that is, $\mathbf{v}_1 x_1 + \cdots + \mathbf{v}_p x_p = \mathbf{0}$ has a nontrivial solution $\begin{cases} x_1 = 1 \\ x_2 = 0 \\ \quad \vdots \\ x_p = 0 \end{cases}$

1.8 Linear transformations

- **Def**: transformation (mapping)

$$T: \mathbb{R}^n \to \mathbb{R}^m \qquad \mathbb{R}^n: \text{ domain}, \qquad \mathbb{R}^m: \text{ codomain}$$
$$\mathbf{x} \mapsto T(\mathbf{x}) \qquad T(\mathbf{x}): \text{ image of } \mathbf{x}, \qquad \text{range of } T\text{: all the images}$$

Ex: Define the following transformation

$$T: \mathbb{R}^2 \to \mathbb{R}^2$$
$$\mathbf{x} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

What is $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$, $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)$?

Answer: $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Ex: Define another transformation

$$T: \mathbb{R}^2 \to \mathbb{R}^2$$
$$\mathbf{x} \mapsto 2\mathbf{x}$$

What is $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$, $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)$?

Answer: $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- **Def**: matrix transformation ($T(\mathbf{x}) = A\mathbf{x}$)

- **Def**: linear transformation

♣ For a matrix transformation $T(\mathbf{x}) = A\mathbf{x}$, we have the following three kinds of problems.

1. Given A, $\mathbf{u} \Longrightarrow T(\mathbf{u})$

Ex: $T: \mathbb{R}^2 \to \mathbb{R}^2$ with $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. What is the image $T(\mathbf{u})$?

Answer: $T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$.

2. Given A, $T(\mathbf{u}) \Longrightarrow \mathbf{u}$

Ex: $T: \mathbb{R}^2 \to \mathbb{R}^2$ with $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $T(\mathbf{u}) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. What is $\mathbf{u}$?

Answer: Since $\mathbf{u}$ satisfies $T(\mathbf{u}) = A\mathbf{u}$, we have $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Then it suffices to consider the augmented matrix and do the row reduction:

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \text{ that is, } \mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

3. For each $\mathbf{x}$, the image $T(\mathbf{x})$ is given $\Longrightarrow A$

Ex: For each $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $T(\mathbf{x}) = \begin{bmatrix} x_1 - x_2 \\ 2x_2 \\ x_1 + x_3 \end{bmatrix}$. What is A?

Answer: Rewrite $T(\mathbf{x}) = \begin{bmatrix} x_1 - x_2 \\ 2x_2 \\ x_1 + x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \mathbf{x_1} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \mathbf{x_2} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mathbf{x_3} =$
$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, so $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

Ex: Consider $T : \mathbb{R}^2 \to \mathbb{R}^2$. Let $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ be the two columns of the identity matrix. If we know $T(\mathbf{e_1}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $T(\mathbf{e_2}) = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, what is A?

Answer: For each $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{e_1} x_1 + \mathbf{e_2} x_2$, we have

$$T(\mathbf{x}) = T(\mathbf{e_1} x_1 + \mathbf{e_2} x_2) = T(\mathbf{e_1}) x_1 + T(\mathbf{e_2}) x_2 = \begin{bmatrix} T(\mathbf{e_1}) & T(\mathbf{e_2}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

So $A = \begin{bmatrix} T(\mathbf{e_1}) & T(\mathbf{e_2}) \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

1.9 The matrix of a linear transformation

- **Thm**: Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that $T(\mathbf{x}) = A\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$. In fact,

  $$A = \begin{bmatrix} T(\mathbf{e_1}) & \cdots & T(\mathbf{e_n}) \end{bmatrix},$$

  where $\mathbf{e_1}, \cdots, \mathbf{e_n}$ are the columns of the identity matrix $I_{n \times n}$.

- **Geometric description in $\mathbb{R}^2$**: $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

1. Reflections: $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
2. Contractions and expansions: $A = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$
3. Shears: $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
4. Rotation: $A = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$
5. Projections: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

- **Def**: onto mapping

  Ex: The mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ is NOT onto.

  $$\mathbf{x} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

♣ **Thm**: Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

T is onto. $\iff$ For each $\mathbf{b} \in \mathbb{R}^m$, $A\mathbf{x} = \mathbf{b}$ is consistent.

$\iff$ A has a pivot position in every row.

$\iff$ $\mathbb{R}^m = Span\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ with $\mathbf{a}_1, \cdots, \mathbf{a}_n$ being the columns of A

- **Def**: one-to-one mapping

  Ex: The mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ is NOT one-to-one.

  $$\mathbf{x} \mapsto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

♣ **Thm**: Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

T is one-to-one. $\iff$ $A\mathbf{x} = 0$ has only the trivial solution.

$\iff$ The columns of A are linearly independent.

2 Chapter 2

2.1 Matrix operations

$$A_{m\times n} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \text{ with } \mathbf{a}_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} \Longrightarrow A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \Big(= [a_{ij}]_{m\times n} \Big)$$

Diagnal matrix: a square matrix with zero non-diagonal entries, for example, $I_n = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}_{n\times n}$

1. Sum and scalar multiple

   $A = B$: same size & same corresponding entries

   $A + B$: the sum has the same size as A and B & adding corresponding entries

   cA: same size as A & each entry in A is multiplied by c

   Properties: $A + B = B + A$, $c(A + B) = cA + cB$

2. Multiplication

   Def: Given $A_{m\times n}$ and $B_{n\times p} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix}$, the product is defined by

   $$AB = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

   Ex: Given $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}_{2\times 3}$ and $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}_{3\times 3}$. What is AB?

   Answer: $AB = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}_{2\times 3}$

   $\Longrightarrow$ The (i, j)-entry in AB can be calculated as $(AB)_{ij} = \text{row}_i(A) \cdot \text{column}_j(B)$

   Ex: Since any given matrix could define a linear transformation, we have

   $$A_{m\times n} \iff T_A : \mathbb{R}^n \to \mathbb{R}^m, \qquad B_{n\times p} \iff T_B : \mathbb{R}^p \to \mathbb{R}^n$$
   $$\mathbf{x} \mapsto A\mathbf{x}, \qquad\qquad \mathbf{x} \mapsto B\mathbf{x}$$

   That is, for any $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{x} \overset{T_B}{\longmapsto} B\mathbf{x} \overset{T_A}{\longmapsto} AB\mathbf{x}$, which define a new mapping

   $$(AB)_{m\times p} \iff T_{AB} : \mathbb{R}^p \to \mathbb{R}^n$$
   $$\mathbf{x} \mapsto AB\mathbf{x}$$

Properties: $A(BC) = (AB)C$, $A(B+C) = AB + AC$, $c(AB) = (cA)B = A(cB)$

♣ In general, $AB \neq BA$ e.g. $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

♣ In general, $AB = AC \not\Rightarrow B = C$ e.g. A, B as above, $C = \begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}$

♣ In general, $AB = 0 \not\Rightarrow A = 0$ or $B = 0$ A as above, $B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$

3. Transpose

   Def: Given $A_{m\times n}$. Its transpose, denoted by $A^\top$, is an $n \times m$ matrix whose columns are the corresponding rows of A

   Properties: $(A^\top)^\top = A$, $(A+B)^\top = A^\top + B^\top$, $(cA)^\top = cA^\top$, $(AB)^\top = B^\top A^\top$

2.2 & 2.3 Inverse of a matrix

- **Def**: invertible

  ♣ If $AB = AC$ and A is invertible $\Longrightarrow B = C$

  ♣ If $AB = 0$ and A is invertible (resp. B is invertible) $\Longrightarrow B = 0$ (resp. $A = 0$)

- **Properties**: $(A^{-1})^{-1} = A$, $(AB)^{-1} = B^{-1}A^{-1}$, $(A^\top)^{-1} = (A^{-1})^\top$

- **Thm**: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad-bc \neq 0$, then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
  If $ad - bc = 0$, then A is not invertible.

  Ex: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$. What is A^{-1}?

  Answer: $ad - bc = 1 \times 5 - 2 \times 3 = -1$, so A is invertible and

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

- **Thm**: If $A_{n\times n}$ is invertible, then for each vector $\mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

  $\Longrightarrow$ In this case, A has a pivot position in every row.

♣ **Thm**: $A_{n\times n}$ is invertible $\Longleftrightarrow$ A is row equivalent to I_n

- **Def**: elementary matrix

  Ex: $E_1 = \begin{bmatrix} 1 & 0 \\ r & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}$

  For any 2×2 matrix A, we have

  $$E_1 A = \begin{bmatrix} 1 & 0 \\ r & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ra + c & rb + d \end{bmatrix}$$

  $\Longrightarrow$ EA is obtained by performing the same row operation to A

♣ Calculation of A^{-1}: If $A_{n\times n}$ is invertible, then $A \sim I_n$ and there exists a matrix A^{-1} such that $A^{-1}A = I_n$. That is, A^{-1} is a kind of row operations that transform A to I_n. Moreover,

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

That is, under the operation A^{-1}, we have $\begin{bmatrix} A & I_n \end{bmatrix} \sim \begin{bmatrix} I_n & A^{-1} \end{bmatrix}$

Ex: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

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

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

2.8 Subspaces of $\mathbb{R}^n$

- **Def**: subspace

  Ex: For $\mathbf{u} \in \mathbb{R}^3$, Span$\{\mathbf{u}\}$ is a subspace of $\mathbb{R}^3$.

  For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, Span$\{\mathbf{u}, \mathbf{v}\}$ is a subspace of $\mathbb{R}^3$.

  Ex: $\mathbb{R}^n$, $\{\mathbf{0}\}$ are both subspaces of $\mathbb{R}^n$.

- **Def**: column space of A: Col$A$

  $\Longrightarrow$ For $A_{m\times n}$, Col$A$ is a subspace of $\mathbb{R}^m$

- **Def**: null space of A: Nul$A$

  $\Longrightarrow$ For $A_{m\times n}$, Nul$A$ is a subspace of $\mathbb{R}^n$

  Ex: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$. Is $\mathbf{u}$ in Col$A$ or Nul$A$?

① Consider $\begin{bmatrix} A & \mathbf{u} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 4 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \end{bmatrix}$. The rightmost column is not a pivot column, so the system is consistent. Equivalently, there is a solution $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{u}$, that is, $\mathbf{u}$ is a linear combination of the columns of A. Hence, $\mathbf{u}$ is in Col$A$.

② Consider $A\mathbf{u} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 19 \end{bmatrix}$. That is, $\mathbf{u}$ is not a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$, so $\mathbf{u}$ is not in Nul$A$.

Ex: Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Then

$$\text{Col}A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

Question: How to find the smallest amount of vectors that span a subspace?

- **Def**: basis

Ex: Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$. Find a basis for Col$A$\Nul$A$.

① Nul$A$: We need to find all the solutions of $A\mathbf{x} = \mathbf{0}$. Consider the augmented matrix

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

The solution is in the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} x_3, \quad x_3 \text{ is a free parameter.}$$

So Nul$A$=Span$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$, and the set $\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$ is a basis for Nul$A$.

② Col$A$: We need to find linearly independent columns of A. Based on the echelon of $\begin{bmatrix} A & \mathbf{0} \end{bmatrix}$ calculated above, we can get the echelon form of A directly

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

The third column can be written as a linear combination of the first two columns, and the first two columns are linear independent. So

$$\text{Col}A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\},$$

and the set $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ is a basis for Col$A$.

♣ **Thm**: The pivot columns of A form a basis for Col$A$.

2.9 Dimension and rank

- **Def**: coordinate vector

  Ex: $\mathbf{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} = 5\mathbf{e}_1 + 6\mathbf{e}_2$ where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ form a basis for $\mathbb{R}^2$.

  Hence, $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ is the coordinate vector of $\mathbf{x}$ relative to the standard basis $\{\mathbf{e}_1, \mathbf{e}_2\}$.

  Ex: $\mathbf{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

  ① $\{\mathbf{b}_1, \mathbf{b}_2\}$ is also a basis for $\mathbb{R}^2$: $\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 0 \\ 0 & \textcircled{1} \end{bmatrix}$

  ② Hence, we can find the coordinate vector of $\mathbf{x}$ relative to the new basis $\{\mathbf{b}_1, \mathbf{b}_2\}$, that is, find $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ such that $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2$:

  $$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{x} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \text{so} \quad \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

- **Def**: dimension

  Ex: $\mathbb{R}^n$ has the standard basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$, so dim $\mathbb{R}^n = n$.

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

  ① Col$A$={the set generated by the pivot columns}=Span$\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_4\}$, so dim Col$A$=3

  ② Nul$A$={all the solutions of $A\mathbf{x} = \mathbf{0}$}:

  $$\begin{bmatrix} A & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 2 & 0 & 0 & 0 \\ 0 & 0 & \textcircled{1} & 0 & 0 \\ 0 & 0 & 0 & \textcircled{1} & 0 \end{bmatrix}, \text{ so } \mathbf{x} = \begin{bmatrix} -2x_2 \\ x_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} x_2$$

  Hence, Nul$A$ = Span$\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$, and dim Nul$A$=1

  $\Longrightarrow$ dim Col$A_{m\times n}$(No. of basic variables)+dim Nul$A_{m\times n}$(No. of free variables)= n(No. of variables)

- **Def**: rank$A$=dim Col$A$

- **Thm** (The rank theorem): For $A_{m\times n}$, rank$A$+dim Nul$A = n$

- **Thm** (The basis theorem): Let H be a p-dimensional subspace of $\mathbb{R}^n$. Any linearly independent set of exactly p vectors in H is a basis for H.

3 Chapter 3

3.1 Determinants of $A_{n\times n}$

- **Def**: submatrix A_{ij}

  Ex: Consider the 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. $A_{11} = [a_{22}]$, $A_{12} = [a_{21}]$, $A_{22} = [a_{11}]$

  Ex: For $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}_{3\times 3}$, $A_{11} = \begin{bmatrix} 2 & 4 \\ 3 & 5 \end{bmatrix}_{2\times 2}$, $A_{12} = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}_{2\times 2}$

- **Def**: determinant of A: $\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + \cdots + a_{1n}(-1)^{1+n}\det A_{1n}$

  In particular, $\det[a_{11}] = a_{11}$.

  Ex: For $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\det A = a_{11}\det A_{11} - a_{12}\det A_{12} = a_{11}a_{22} - a_{12}a_{21}$

- **Thm**: $A_{n\times n}$ is invertible $\iff$ $\det A \neq 0$

- **Def**: the (i,j)-cofactor of A is denoted by $C_{ij} = (-1)^{i+j}\det A_{ij}$

  $\implies$ Then the definition of $\det A$ above can be rewritten as

  $$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n},$$

  which is called the cofactor expansion across the first row.

- **Thm**: $\det A$ can be calculated by the cofactor expansion across any row of down any column

  $$\begin{aligned} \det A &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \end{aligned}$$

  Ex: Calculate the following determinant

  $$\begin{vmatrix} 1 & 0 & 2 & 3 & 1 \\ 2 & 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 0 & 0 \\ 1 & 2 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 0 \end{vmatrix} \xlongequal{3rd\ row} 3(-1)^{3+3} \begin{vmatrix} 1 & 0 & 3 & 1 \\ 2 & 0 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 0 \end{vmatrix} \xlongequal{4th\ row} 3\cdot 2(-1)^{4+3} \begin{vmatrix} 1 & 0 & 1 \\ 2 & 0 & 3 \\ 1 & 2 & 4 \end{vmatrix}$$

  $$\xlongequal{2nd\ column} (-6)2(-1)^{3+2} \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 12$$

  Ex: $\begin{vmatrix} 2 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} \xlongequal{1st\ column} 2(-1)^{1+1} \begin{vmatrix} 4 & 5 \\ 0 & 6 \end{vmatrix} = 2\cdot 4\cdot 6$

- **Thm**: If $A_{n\times n}$ is a triangular matrix, then its determinant is the product of the main diagonals, that is, $\det A = \Pi_{i=1}^{n} a_{ii}$.

- **Thm** (Row operations): Let A be a square matrix.

  ① If a scalar multiple of one row of A is added to another row to produce B, then $\det B = \det A$.

  ② If two rows of A are interchanged to produce B, then $\det B = -\det A$.

  ③ If a scalar k is multiplied to one row of A to produce B, then $\det B = k\det A$.

  Ex:

$$\begin{vmatrix} 5 & 6 & 7 \\ 5 & 6 & 8 \\ 50 & 260 & 150 \end{vmatrix} \overset{use\ ③}{=\!=\!=} 10\begin{vmatrix} 5 & 6 & 7 \\ 5 & 6 & 8 \\ 5 & 26 & 15 \end{vmatrix} \overset{use\ ①}{=\!=\!=} 10\begin{vmatrix} 5 & 6 & 7 \\ 0 & 0 & 1 \\ 0 & 20 & 8 \end{vmatrix}$$

$$\overset{use\ ②}{=\!=\!=} -10\begin{vmatrix} 5 & 6 & 7 \\ 0 & 20 & 8 \\ 0 & 0 & 1 \end{vmatrix} = -1000$$

3.2 Properties of determinants

- **Thm**: Let A be a square matrix, then $\det A^\top = \det A$.

$$\begin{aligned} \Longrightarrow \det A^\top &= \text{cofactor expansion across the } i\text{th row of } A^\top \\ &= \text{cofactor expansion down the } i\text{th column of } A \\ &= \det A \end{aligned}$$

- **Thm** (Multiplicative property): Let A and B be $n \times n$ square matrices. Then $\det(AB) = \det A \cdot \det B$

  $\Longrightarrow$ If A is invertible, then $1 = |I| = |AA^{-1}| = |A||A^{-1}|$. Hence, $|A^{-1}| = \frac{1}{|A|}$.

  $\Longrightarrow$ In general, $\det(A + B) \neq \det A + \det B$

- **Thm** (Linearity property): Assume that the jth column of $A_{n\times n}$ is allowed to vary $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{j-1} & \mathbf{x} & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{bmatrix}$. Define the mapping $T : \mathbb{R}^n \to \mathbb{R}$ by $T(\mathbf{x}) = \det A$. Then T is linear: $T(c\mathbf{x}) = cT(\mathbf{x})$ and $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$.

$$\Longrightarrow \begin{vmatrix} a_{11} & cx_1 \\ a_{21} & cx_2 \end{vmatrix} = c\begin{vmatrix} a_{11} & x_1 \\ a_{21} & x_2 \end{vmatrix} \text{ and } \begin{vmatrix} a_{11} & x_1 + y_1 \\ a_{21} & x_2 + y_2 \end{vmatrix} = \begin{vmatrix} a_{11} & x_1 \\ a_{21} & x_2 \end{vmatrix} + \begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix}$$

  Ex:

$$\begin{vmatrix} 17 & 17 & 17 \\ 25 & 26 & 25 \\ 55 & 88 & 56 \end{vmatrix} = \begin{vmatrix} 17 & 17+0 & 17 \\ 25 & 25+1 & 25 \\ 55 & 55+33 & 56 \end{vmatrix} = \begin{vmatrix} 17 & 17 & 17 \\ 25 & 25 & 25 \\ 55 & 55 & 56 \end{vmatrix} + \begin{vmatrix} 17 & 0 & 17 \\ 25 & 1 & 25 \\ 55 & 33 & 56 \end{vmatrix}$$

$$= \begin{vmatrix} 17 & 0 & 17+0 \\ 25 & 1 & 25+0 \\ 55 & 33 & 55+1 \end{vmatrix} = \begin{vmatrix} 17 & 0 & 17 \\ 25 & 1 & 25 \\ 55 & 33 & 55 \end{vmatrix} + \begin{vmatrix} 17 & 0 & 0 \\ 25 & 1 & 0 \\ 55 & 33 & 1 \end{vmatrix} = \begin{vmatrix} 17 & 0 & 0 \\ 25 & 1 & 0 \\ 55 & 33 & 1 \end{vmatrix} = 17.$$

- **Def**: Let A be an $n \times n$ matrix and $\mathbf{b}$ is vector in $\mathbb{R}^n$. Denote

$$A_i(\mathbf{b}) = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \end{bmatrix}$$

♣ **Thm** (Cramer's rule): If $A_{n\times n}$ is invertible, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the system $A\mathbf{x} = \mathbf{b}$ has a unique solution $\mathbf{x}$ with entries

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

Ex: Consider $\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \end{bmatrix}$. We have got $x_1 = 4$ and $x = 2$ in Chapter 1. Next we use Cramer's rule to check these results.

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{\begin{vmatrix} 6 & 1 \\ 16 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix}} = \frac{8}{2} = 4$$

$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{\begin{vmatrix} 1 & 6 \\ 2 & 16 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix}} = \frac{4}{2} = 2$$

3.3 Volume and linear transformation

Recall: For $A_{2\times 2} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, if A is invertible, then $A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

- **Def**: The adjugate (adjoint) of $A_{n\times n}$ is

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

where $C_{ij} = (-1)^{i+j}\det A_{ij}$ is the (i,j)-cofactor of A.

Ex: Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, calculate adj$A$.

Answer: $C_{11} = (-1)^{1+1}\det[d] = d$, $C_{12} = (-1)^{1+2}\det[c] = -c$

$C_{21} = (-1)^{2+1}\det[b] = -b$, $C_{22} = (-1)^{2+2}\det[a] = a$

Hence, $\text{adj}A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

$\Longrightarrow (A_{2\times 2})^{-1} = \frac{1}{\det A}\text{adj}A$

- **Thm** (An inverse formula): Let A be an $n \times n$ invertible matrix. Then

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

$\Longrightarrow$ The (i, j) entry of A^{-1} is $\frac{C_{ji}}{\det A}$.

Ex: For $A = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$, the area determined by the columns $\begin{bmatrix} k \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is $|k|$.

$\Longrightarrow$ Moreover, the parallelogram determined by two vectors $\begin{bmatrix} k \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is the same as the parallelogram determined by four points $(0, 0)$, $(k, 0)$, $(0, 1)$ and $(k, 1)$.

- **Thm**: For $A_{n\times n}$, the volume determined by its columns is $|\det A|$.

- **Thm**: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping with $T(\mathbf{x}) = A\mathbf{x}$. Then for any region S in $\mathbb{R}^n$,

$$\{\text{The volume of } T(S)\} = |\det A| \cdot \{\text{The volume of } S\}.$$

Review of Chapter 3

1. Determinant of $A_{n\times n}$:

$$\begin{aligned}\det A &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \qquad \text{(the cofactor expansion across the } i\text{th row)} \\ &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \qquad \text{(the cofactor expansion down the } j\text{th column)}\end{aligned}$$

2. Properties of determinants:

   ① row operations: three kinds of elementary row operations

   ② transpose: $|A^\top| = |A|$

   ③ multiplication: $|AB| = |A| \cdot |B|$

   ④ linearity: $\left|\begin{bmatrix} \mathbf{a}_1 & \mathbf{x}+\mathbf{y} \end{bmatrix}\right| = \left|\begin{bmatrix} \mathbf{a}_1 & \mathbf{x} \end{bmatrix}\right| + \left|\begin{bmatrix} \mathbf{a}_1 & \mathbf{y} \end{bmatrix}\right|$

3. Solve $A\mathbf{x} = \mathbf{b}$:

   ① $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$

   ② If A is invertible ($\det A \neq 0$), then $\mathbf{x} = A^{-1}\mathbf{b}$

   ③ If A is invertible ($\det A \neq 0$), then the ith entry in $\mathbf{x}$ is $x_i = \frac{\det A_i(\mathbf{b})}{\det A}$

4. Calculate A^{-1}:

   ① $\begin{bmatrix} A & I \end{bmatrix} \sim \begin{bmatrix} I & A^{-1} \end{bmatrix}$

   ② $A^{-1} = \frac{1}{\det A}\text{adj}A$ (this can be used to calculate the (i, j) entry of A^{-1})

5. Matlab code (for the ones who are interested):

   Define a vector: $>> \mathbf{b} = [1; 2]$

   Define a matrix: $>> A = [1, 2; 3, 4]$

   Determinant of A: $>> det(A)$

   Inverse of A: $>> inv(A)$

   Adjoint of A: $>> adjoint(A)$

   Solution of $A\mathbf{x} = \mathbf{b}$ if: $>> A\backslash\mathbf{b}$

4 Chapter 4

4.1 Vector spaces and subspaces

- **Def**: vector spaces

  Ex: $\mathbb{R}^n$ is a vector space with zero object $\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}$

  Ex: The polynomial space $\mathbb{P}_n = \{$all polynomials of the form $p(t) = a_0 + a_1 t + \cdots + a_n t^n\}$ is a vector space with zero object 0 (constant).

  Ex: The matrix space $\mathbb{M}_{m\times n} = \{$ all $m \times n$ matrices $A\}$ is a vector space with zero object $\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{m\times n}$

- **Def**: For general vector spaces V and W, a linear transformation $T : V \to W$ satisfies

  (i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for $\mathbf{u}, \mathbf{v} \in V$;

  (ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for $\mathbf{u} \in V$.

- **Def**: subspace H of general vector space V

  Ex: $\{\mathbf{0}\}$ and V are subspaces of V

  Ex: For $\mathbf{v}_1, \mathbf{v}_2 \in V$, the spanning set $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ is a subspace of V.

  Ex: Determine if $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ is in the subspace spanned by $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$.

  $\iff$ Determine if $w \in \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

  $\iff$ Consider the augmented matrix $\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{w} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

  The system above is not consistent, so $\mathbf{w}$ is not in the spanning set.

4.2 Column/Null spaces and linear transformation

- **Def**: $\text{Col}A_{m\times n} = \text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$

  $= \{\mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x} \text{ for some } x \in \mathbb{R}^n\}$

Ex: Given a set $S = \left\{ \begin{bmatrix} 2s+3t \\ r+s-2t \\ 4r+s \\ 3r-s-t \end{bmatrix} : r, s, t \text{ real} \right\}$. Find A such that $S = \text{Col}A$.

Answer: Note that

$$S = \left\{ \begin{bmatrix} 0 \\ 1 \\ 4 \\ 3 \end{bmatrix} r + \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} s + \begin{bmatrix} 3 \\ -2 \\ 0 \\ -1 \end{bmatrix} t : r, s, t \text{ real} \right\}$$

$$= \text{Span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 0 \\ -1 \end{bmatrix} \right\}$$

As a result, $A = \begin{bmatrix} 0 & 2 & 3 \\ 1 & 1 & -2 \\ 4 & 1 & 0 \\ 3 & -1 & -1 \end{bmatrix}$

Ex: Given $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$. Is $\mathbf{b}$ in Col$A$?

$\iff$ Determine if $\mathbf{b} \in \text{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$

$\iff$ Consider the augmented matrix $\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 3 & 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

The system is not consistent, so $\mathbf{b}$ is not in Col$A$.

Ex: Given A as above. Find k such that Col$A$ is a subspace of $\mathbb{R}^k$.

Answer: $k = 3$

- **Def**: $\text{Nul}A_{m\times n} = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$

Ex: Given $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix}$. Find Nul$A$.

Answer: Consider the augmented matrix of the homogeneous system

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

Its solutions are in the form $\begin{cases} x_1 = x_3 + 2x_4 \\ x_2 = -2x_3 - 3x_4 \\ x_3 = x_3 \text{ (free)} \\ x_4 = x_4 \text{ (free)} \end{cases} \iff \mathbf{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} x_4.$

Hence, Nul$A$ = Span$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Ex: Given A as above and $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$. Is $\mathbf{u}$ in Nul$A$?

Answer:

① One way is to find Nul$A$ first, and then check if $\mathbf{u}$ is in the spanning set. It will need a lot of calculations.

② The simplest way is to check if $A\mathbf{u} = \mathbf{0}$: $A\mathbf{u} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, so $\mathbf{u}$ is in Nul$A$.

4.3 Linearly independent sets and bases

- **Def**: The set of vectors $\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ in V is linearly independent if $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = 0$ has only the trivial solution $c_1 = \cdots = c_p = 0$.

  Ex: Is the set $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \right\}$ in $\mathbb{R}^3$ linearly independent?

  Answer: Consider the augmented matrix $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. There is only the trivial solution, so the set above is a linearly independent set.

  Ex: It the set $\{1, t, t^2\}$ in $\mathbb{P}_2$ linearly independent?

  Answer: Consider the homogeneous equation $c_1 \cdot 1 + c_2 t + c_3 t^2 = 0$. It has only the trivial solution $c_1 = c_2 = c_3 = 0$. So the set is a linear independent set.

- **Def**: Let H be a subspace of V. Then the set $\mathcal{B} = \{v_1, \cdots, v_p\}$ is a basis for H if

  ① $\mathcal{B}$ is a linearly independent set,

  ② H =Span$\{v_1, \cdots, v_p\}$.

  Ex: $\mathbb{R}^n$ =Span$\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$. The set $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is called the standard basis for $\mathbb{R}^n$.

  Ex: $\mathbb{P}_n = \{c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n : c_0, c_1, \cdots, c_n \text{ real}\}$ =Span$\{1, t, t^2, \cdots, t^n\}$.

  The set $\{1, t, t^2, \cdots, t^n\}$ is called the standard basis for $\mathbb{P}_n$.

Ex (8 in the textbook): Given the set $\left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix} \right\}$ in $\mathbb{R}^3$.

(1) Is it a basis for $\mathbb{R}^3$?

No, because any basis for $\mathbb{R}^3$ should contain exactly 3 vectors.

(2) Find a basis for the set spanned by above vectors.

It suffices to find the linearly independent vectors in above set:

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

So $\left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix} \right\}$ is a basis for $\text{Span}\left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix} \right\}$.

Since there is exactly three vectors in the set $\left\{ \begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 4 \end{bmatrix} \right\}$, it is also a basis for $\mathbb{R}^3$.

- **Thm** (The spanning set thm): For $\{v_1, \cdot, v_p\}$ in V, if v_k is a linear combination of the other vectors, then

$$\text{Span}\{v_1, \cdots, v_p\} = \text{Span}\{v_1, \cdots, v_{k-1}, v_{k+1}, \cdots, v_p\}.$$

Ex: According to theorem above, $\text{Span}\{u, 2u\} = \text{Span}\{u\} = \text{Span}\{2u\}$

Ex: $\text{Col}A = \text{Span}\{$ all the columns $\} = \text{Span}\{$ pivot columns $\}$

Ex: Find a basis for the set of vectors in the plane $x + 2y + z = 0$.

Answer: Denote the set above by

$$S = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + 2y + z = 0 \right\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \right\} = \text{Nul} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}.$$

We only need to find a basis for $\text{Nul}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$:

$$\begin{bmatrix} \textcircled{1} & 2 & 1 & 0 \end{bmatrix} \Longrightarrow \begin{cases} x = -2y - z \\ y = y \text{ (free)} \\ z = z \text{ (free)} \end{cases} \Longrightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} y + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} z.$$

So $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for S.

4.5 Dimension of vector spaces

- **Def**: dim$V$ = number of vectors in a basis

  Ex: dim$\mathbb{R}^n = n$ with a standard basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$

  Ex: dim$\mathbb{P}_n = n + 1$ with a standard basis $\{1, t, \cdots, t^n\}$

- **Thm**: If V is a vector space with a basis $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$, then

  (1) any basis for V has exactly p vectors;

  (2) any set of more than p vectors in V is linearly dependent.

  Ex: $\mathbb{R}^2$ has a standard basis $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$.

  Is $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\}$ a linearly independent set? No

  Is the set above a basis for $\mathbb{R}^2$? No

  Ex: $\mathbb{P}_1$ has a standard basis $\{1, t\}$.

  Are the following sets bases for $\mathbb{P}_1$?

  $\{1, 1+t\}$ Yes

  $\{2, t\}$ Yes

  $\{t, 2+t\}$ Yes

  $\{t, 2t\}$ No, cause one is a scalar multiple of the other one

  $\{1, t, 1+t\}$ No, cause there is more than 2 vectors

  Ex: Define a set $S = \left\{ \begin{bmatrix} a+2b \\ 2a+4b \\ -a-2b \end{bmatrix} : a, b \text{ real} \right\}$. What is dim$S$?

  Answer: $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} a + \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} b : a, b \text{ real} \right\}$

  $$= \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right\}$$

  So dim$S$=1.

  Ex: Define a set $T = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 0 \right\}$. What is dim$T$?

  Answer: $T = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0 \right\} = \text{Nul} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$

$$= \left\{ \begin{bmatrix} -b-c \\ b \\ c \end{bmatrix} : b, c \text{ real} \right\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} b + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} c : b, c \text{ real} \right\}$$

$$= \text{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

So dim$T$=2.

Ex: $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & -2 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 5}$. Then

dim Col$A = 2$, and Col$A$ is a subspace of $\mathbb{R}^4$

dim Nul$A = 3$, and Nul$A$ is a subspace of $\mathbb{R}^5$

- **Thm**: If H is a subspace of a finite-dimensional vector space V, then

  (1) dim$H \leq$ dim$V$;

  (2) H is also a finite-dimensional vector space;

  (3) any basis for H can be extended to a basis for V.

  Ex: Given A as above. Then Col$A$ is a subspace of $\mathbb{R}^4$. We now check the above three results:

  (1) dim Col$A \leq$ dim$\mathbb{R}^4$ holds;

  (2) holds apparently;

  (3) The pivot columns form a basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 0 \\ 0 \end{bmatrix} \right\}$ for Col$A$.

  Now we extend it to a basis for $\mathbb{R}^4$: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

4.6 Rank

For $A_{m\times n} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$, Col$A$ = Span$\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$.

- **Def**: For $A_{m\times n} = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}$, the row space is $\text{Row}A = \text{Span}\{\mathbf{r}_1, \cdots, \mathbf{r}_m\}$, which is a subspace of $\mathbb{R}^n$.

  $\Longrightarrow \text{Row}A = \text{Col}A^\top$

  Ex: $A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}$, then $\text{Row}A = \text{Span}\{\mathbf{r}_1, \mathbf{r}_2\}$.

  If we use the three kinds of elementary row operations:

  $A \overset{\text{row interchange}}{\sim} A_1 = \begin{bmatrix} \mathbf{r}_2 \\ \mathbf{r}_1 \end{bmatrix}$, then $\text{Row}A_1 = \text{Span}\{\mathbf{r}_2, \mathbf{r}_1\}$;

  $A \overset{\text{scalar multiple}}{\sim} A_2 = \begin{bmatrix} c\mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}$, then $\text{Row}A_2 = \text{Span}\{c\mathbf{r}_1, \mathbf{r}_2\}$;

  $A \overset{\text{row replacement}}{\sim} A_3 = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 + c\mathbf{r}_1 \end{bmatrix}$, then $\text{Row}A_3 = \text{Span}\{\mathbf{r}_1, \mathbf{r}_2 + c\mathbf{r}_1\}$.

  The above row spaces are the same: $\text{Row}A = \text{Row}A_1 = \text{Row}A_2 = \text{Row}A_3$.

  That is, elementary row operations won't change the row space.

- **Thm**: If matrices A and B are row equivalent, then they have the same row space.

  If B is in echelon form, then its non-zero rows form a basis for $\text{Row}A = \text{Row}B$.

  Ex: Given $A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 1 & 2 & 1 \end{bmatrix}$. Find bases for Col$A$, Row$A$ and Nul$A$.

  ① echelon form:

$$A \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & -2 & -4 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 1 & 1 & 1 & 1 \\ 0 & \textcircled{1} & 2 & 3 & 4 \\ 0 & 0 & \textcircled{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  $\text{Col}A = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 4 \\ 1 \end{bmatrix} \right\}$

  $\text{Row}A = \text{Span}\{(1,1,1,1,1), (0,1,2,3,4), (0,0,1,1,2)\}$

  ② reduced echelon form:

$$A \sim \begin{bmatrix} \textcircled{1} & 1 & 0 & 0 & -1 \\ 0 & \textcircled{1} & 0 & 1 & 0 \\ 0 & 0 & \textcircled{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 0 & 0 & -1 & -1 \\ 0 & \textcircled{1} & 0 & 1 & 0 \\ 0 & 0 & \textcircled{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} x_1 = x_4 + x_5 \\ x_2 = -\, x_4 \\ x_3 = -\, x_4 - 2x_5 \\ x_4 = x_4 \text{ (free)} \\ x_5 = x_5 \text{ (free)} \end{cases} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} x_4 + \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} x_5$$

$$\text{Nul}A = \text{Span}\left\{ \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- **Def**: $\text{rank}A = \text{dimCol}A$

  Ex: Given A as above. We have

  $\text{dimCol}A = \text{dimRow}A =$ number of pivot positions $= 3$.

♣ **Thm** (The rank thm): For $A_{m\times n}$, it holds

  $\text{dimCol}A = \text{dimRow}A = \text{rank}A$ and $\text{rank}A + \text{dimNul}A = n$.

  $\Longrightarrow$ For $(A^\top)_{n\times m}$, $\text{rank}A^\top + \text{dimNul}A^\top = m$, where

  $\text{rank}A^\top = \text{dimCol}A^\top = \text{dimRow}A = \text{dimCol}A = \text{rank}A$.

  Ex: If the null space of a 7×6 matrix A is 5-dimensional, what are $\text{dimCol}A$ and $\text{dimRow}A$?

  Answer: $\text{dimCol}A = \text{dimRow}A = 6 - \text{dimNul}A = 1$.

- **Thm**: Let A be an $n \times n$ matrix. Then

$$\begin{aligned} A \text{ is invertible } &\iff \det(A) \neq 0 \\ &\iff A \sim I_n \\ &\iff \text{dimCol}A = \text{dimRow}A = \text{rank}A = n \\ &\iff \text{dimNul}A = 0 \\ &\iff \text{Nul}A = \{\mathbf{0}\} \\ &\iff \text{Col}A = \mathbb{R}^n \end{aligned}$$

5 Chapter 5

5.1 Eigenvalues and eigenvectors

- **Def**: eigenvalues and eigenvectors

  Ex: Is $\mathbf{x} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ an eigenvector of $A = \begin{bmatrix} 3 & 6 & 7 \\ 3 & 3 & 7 \\ 5 & 6 & 5 \end{bmatrix}$?

  Answer: Calculate $A\mathbf{x} = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix} = -2\mathbf{x}$.

  So $\mathbf{x}$ is an eigenvector of A corresponding to the eigenvalue -2.

  Ex: Is $\lambda = 2$ is an eigenvalue of $A = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$?

  Answer: If λ is an eigenvalue, then $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.

  Consider the augmented matrix $\begin{bmatrix} A - \lambda I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

  The system has a free variable, so has nontrivial solutions.

  Hence, $\lambda = 2$ is an eigenvalue of A.

♣ **Calculation**:

① Eigenvalues: $|A - \lambda I| = 0 \quad \big((A - \lambda I)\mathbf{x} = \mathbf{0}$ has nontrivial solutions$\big)$

Ex: Given $A = \begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$. Consider $|A - \lambda I| = \begin{vmatrix} 3-\lambda & 2 \\ 3 & 8-\lambda \end{vmatrix} = (\lambda - 2)(\lambda - 9) = 0.$

So its eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 9$.

② Eigenvectors: nontrivial solutions of $(A - \lambda I)\mathbf{x} = \mathbf{0}$

$\Longrightarrow$ The eigenspace for λ is actually $\text{Nul}(A - \lambda I)\backslash\{\mathbf{0}\}$

Ex: For $\lambda_1 = 2$, consider $\begin{bmatrix} A - \lambda_1 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

All the nontrivial solutions are of the form $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} x_2$ except $\mathbf{0}$.

$\left\{\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} x_2 : \mathbf{x} \neq \mathbf{0}\right\}$ is called the eigenspace corresponding to $\lambda_1 = 2$.

For $\lambda_2 = 9$, similarly, $\begin{bmatrix} A - \lambda_2 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} -6 & 2 & 0 \\ 3 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} \textcircled{1} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

All the nontrivial solutions are of the form $\mathbf{x} = \begin{bmatrix} \frac{1}{3} \\ 1 \end{bmatrix} x_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} t$ except $\mathbf{0}$.

$\left\{\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} t : \mathbf{x} \neq \mathbf{0}\right\}$ is called the eigenspace corresponding to $\lambda_2 = 9$.

- **Thm**: The eigenvectors corresponding to distinct eigenvalues are linearly independent.

  Ex: Given $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{bmatrix}$. Find its eigenvalues.

  Answer: $|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & -\lambda & 4 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (1-\lambda)(-\lambda)(5-\lambda) = 0.$

  Its eigenvalues are $\lambda = 1, 0, 5$.

- **Thm**: The eigenvalues of a triangular matrix are its diagonals.

- **Thm**: Let A be an $n \times n$ matrix. Then

  A is invertible $\iff |A| \neq 0 \ \big(\text{i.e. } |A - 0I| \neq 0\big) \iff 0$ is not an eigenvalue of A

  A is not invertible $\iff |A| = 0 \ \big(\text{i.e. } |A - 0I| = 0\big) \iff 0$ is an eigenvalue of A

  Ex: Without calculation, we know that the matrix $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ has eigenvalue 0 cause it is not invertible.

5.2 The characteristic equation

- **Thm** (Properties of determinants): Let A and B be $n \times n$ matrices. Then

  ① A is invertible $\iff |A| \neq 0 \iff 0$ is not an eigenvalue of A

  ② $|AB| = |A| \cdot |B|, \quad |A^\top| = |A|, \quad |A^{-1}| = \frac{1}{|A|}$

  ③ If A is triangular, then $|A| = a_{11}a_{22}\cdots a_{nn}$ (product of the diagonals)

  ④ $A = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \overset{\text{row replacement}}{\sim} B = \begin{bmatrix} r_1 \\ r_2 + cr_1 \end{bmatrix}$, then $|B| = |A|$

  $A = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \overset{\text{row interchange}}{\sim} B = \begin{bmatrix} r_2 \\ r_1 \end{bmatrix}$, then $|B| = -|A|$

  $A = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \overset{\text{row scaling}}{\sim} B = \begin{bmatrix} cr_1 \\ r_2 \end{bmatrix}$, then $|B| = c|A|$

  ⑤ linearity property (see below)

  Ex: $\begin{vmatrix} 18 & 56 \\ 17 & 56 \end{vmatrix} = \begin{vmatrix} 17+1 & 56 \\ 17+0 & 56 \end{vmatrix} \overset{\text{linearity}}{=} \begin{vmatrix} 17 & 56 \\ 17 & 56 \end{vmatrix} + \begin{vmatrix} 1 & 56 \\ 0 & 56 \end{vmatrix} = 56$

  Ex: If A is of size $n \times n$, then $|cA| = c^n|A|$.

- **Def**: $|A - \lambda I| = 0$: Characteristic equation

  $|A - \lambda I|$: Characteristic polynomial (CP)

  Ex: Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$. Then its characteristic polynomial is

  $$\text{CP} = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & 4-\lambda & 5 \\ 0 & 0 & 6-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda)(6-\lambda),$$

  and its eigenvalues are $\lambda = 1, 4, 6$.

  Ex: Let $A = \begin{bmatrix} 4 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$. Then its characteristic polynomial is

  $$\text{CP} = |A - \lambda I| = \begin{vmatrix} 4-\lambda & 2 & 3 \\ 0 & 4-\lambda & 5 \\ 0 & 0 & 6-\lambda \end{vmatrix} = (4-\lambda)^2(6-\lambda),$$

  and its eigenvalues are $\lambda = 4, 4, 6$.

- **Def**: The multiplicity of $\lambda = 4$ in the above example is 2.

  Ex: For $A_{4\times 4}$, it has eigenvalues 1,2,2,6. What's its CP?

  Answer: CP$= (1-\lambda)(2-\lambda)^2(6-\lambda)$

  Ex: Let $A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & h & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Find h such that the eigenspace for $\lambda = 5$ is two.

  Answer: The eigenspace for $\lambda = 5$ is $\text{Nul}(A - 5I)\backslash\{\mathbf{0}\}$. It suffices to consider the null space $\text{Nul}(A - 5I)$:

  $$\begin{bmatrix} 0 & -2 & 6 & -1 & 0 \\ 0 & -2 & h & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & -4 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & -2 & 6 & -1 & 0 \\ 0 & 0 & h-6 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  The eigenspace is of dimension two if there is two free variables, that is, $h = 6$.

5.3 Diagonalization

- **Def**: similar ($A = PBP^{-1}$)

- **Thm**: If A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues.

  Ex: If $A = PBP^{-1}$ with $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ and

$$A^k = PB^kP^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0 & 2^k \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

- **Def**: diagonalizable ($A = PDP^{-1}$ with D a diagonal matrix)

♣ **Thm** (The diagonalization thm): An $n \times n$ matrix is diagonalizable $\iff$ A has n linearly independent eigenvectors.

Reason: Let $\mathbf{p}_1, \cdots, \mathbf{p}_n$ be the n linearly indepedent eigenvectors. Then there must be corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$ such that

$$\begin{cases} A\mathbf{p}_1 = \lambda_1 \mathbf{p}_1 \\ \quad\vdots \\ A\mathbf{p}_n = \lambda_1 \mathbf{p}_n \end{cases} \Longrightarrow \begin{bmatrix} A\mathbf{p}_1 & \cdots & A\mathbf{p}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{p}_1 & \cdots & \lambda_n\mathbf{p}_n \end{bmatrix}$$

$$\Longrightarrow A \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

$$\Longrightarrow AP = PD$$

$$\Longrightarrow A = PDP^{-1}$$

with $P = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$.

Ex: Is $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ diagonalizable?

Answer: Its eigenvalues are $\lambda = 1, 3$. Next we calculate the corresponding eigenvectors.

For $\lambda_1 = 1$: $\begin{bmatrix} 0 & 2 & 0 \\ 0 & 2 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & \textcircled{1} & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1$. We can choose $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

For $\lambda_1 = 3$: $\begin{bmatrix} \textcircled{1} & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_2$. We can choose $\mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Now we get $D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ and $P = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ such that $A + PDP^{-1}$.

So A is diagonalizable.

- **Thm**: An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Ex: Is $A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ diagonalizable?

Its CP= $\begin{vmatrix} 2-\lambda & 0 & 1 \\ 1 & 3-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2(3-\lambda)$. So it has eigenvalues $\lambda = 2, 2, 3$.

For $\lambda = 2$: $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & 1 & 0 & 0 \\ 0 & 0 & ① & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} x_2$. We can choose $\mathbf{p}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

For $\lambda = 3$: $\begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & 0 & 0 & 0 \\ 0 & 0 & ① & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x_2$. We can choose $\mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

We can not find $\mathbf{p}_3$ to get an invertible matrix P. So A is NOT diagonalizable.

Ex: Is $A = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ diagonalizable?

Its CP= $\begin{vmatrix} 3-\lambda & 0 & 1 \\ 1 & 2-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2(3-\lambda)$. So it has eigenvalues $\lambda = 2, 2, 3$.

For $\lambda = 2$: $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} x_3$. We can choose $\mathbf{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{p}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, which are linearly independent.

For $\lambda = 3$: $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & -1 & 0 & 0 \\ 0 & 0 & ① & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} x_2$. We can choose $\mathbf{p}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

Now we get the invertible matrix $P = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

So A is diagonalizable.

- **Thm**: Let A be an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \cdots, \lambda_p$ $(p \leq n)$.

  ① The dimension of the eigenspace for λ_k $(1 \leq k \leq p)$ is less than or equal to the multiplicity of λ_k.

  ②: A is diagonalizable $\iff$ the dimension of the eigenspace for λ_k is equal to the multiplicity of λ_k (i.e., the sum of the dimensions of the eigenspaces is n)

5.4 Eigenvectors and linear transformations

Recall that $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear $\iff$ $T(\mathbf{x}) = A\mathbf{x}$ with $A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}_{m \times n}$.

- **Def**: If V has a basis $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ (that is, $\dim V = n$), then any $\mathbf{x} \in V$ is $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$. Define the coordinate vector

$$\big[\mathbf{x}\big]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

  Ex: Let $V = \mathbb{P}_2$ which has the standard basis $\mathcal{B} = \{1, t, t^2\}$. For the polynomial $p(t) = 3 - t^2$, what is its coordinate vector $\big[p(t)\big]_{\mathcal{B}}$?

  Answer: $\big[p(t)\big]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$

- **Def**: Assume that V is a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ (i.e., dim $V = n$), and W is a vector space with basis $\mathcal{C} = \{\mathbf{c}_1, \cdots, \mathbf{c}_m\}$ (i.e., dim $W = m$). Then

$$\begin{array}{ccc} T : \mathbf{x} & \longrightarrow & T(\mathbf{x}) \\ \downarrow & & \downarrow \\ \big[\mathbf{x}\big]_{\mathcal{B}} & \xrightarrow{A} & \big[T(\mathbf{x})\big]_{\mathcal{C}} \end{array}$$

  with $A = \begin{bmatrix} \big[T(\mathbf{b}_1)\big]_{\mathcal{C}} & \cdots & \big[T(\mathbf{b}_n)\big]_{\mathcal{C}} \end{bmatrix}$ the matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$.

  In particular, if $V = W$ and $\mathcal{B} = \mathcal{C}$, we denote the standard matrix A by $\big[T\big]_{\mathcal{B}}$.

♣ **Ex**: Let $T : \mathbb{P}_2 \to \mathbb{P}_1$ be a linear transformation defined by $T(a_0 + a_1 t + a_2 t^2) = a_0 + (a_2 - a_1)t$ for any real numbers a_0, a_1 and a_2. What is the standard matrix relative to the standard bases for $\mathbb{P}_2$ and $\mathbb{P}_1$?

Answer:

① Find $\mathcal{B}$ and $\mathcal{C}$:

The standard basis for $\mathbb{P}_2$ is $\mathcal{B} = \{1, t, t^2\}$ and the standard basis for $\mathbb{P}_1$ is $\mathcal{C} = \{1, t\}$.

② Find $A = \big[\ [T(\mathbf{b}_1)]_{\mathcal{C}} \ \cdots \ [T(\mathbf{b}_n)]_{\mathcal{C}}\ \big]$:

Note that in this example $\mathbf{b}_1 = 1$, $\mathbf{b}_2 = t$ and $\mathbf{b}_3 = t^2$. According to the map T defined above, we have

$$\begin{aligned} T(\mathbf{b}_1) &= T(1) = 1 && \text{(in this case } a_0 = 1, a_1 = a_2 = 0),\\ T(\mathbf{b}_2) &= T(t) = -t && \text{(in this case } a_1 = 1, a_0 = a_2 = 0),\\ T(\mathbf{b}_3) &= T(t^2) = t && \text{(in this case } a_2 = 1, a_0 = a_1 = 0), \end{aligned}$$

and hence

$$\begin{aligned} \big[T(\mathbf{b}_1)\big]_{\mathcal{C}} &= \big[1\big]_{\{1,t\}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\\ \big[T(\mathbf{b}_2)\big]_{\mathcal{C}} &= \big[-t\big]_{\{1,t\}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix},\\ \big[T(\mathbf{b}_2)\big]_{\mathcal{C}} &= \big[t\big]_{\{1,t\}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{aligned}$$

Finally, we get the standard matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Ex: Let $T : \mathbb{P}_2 \to \mathbb{R}^3$ be a linear transformation defined by $T(p(t)) = \begin{bmatrix} p(-1) \\ p(0) \\ p(1) \end{bmatrix}$. What is the standard matrix relative to the standard bases for $\mathbb{P}_2$ and $\mathbb{R}^3$?

Answer:

① Find $\mathcal{B}$ and $\mathcal{C}$:

The standard basis for $\mathbb{P}_2$ is $\mathcal{B} = \{1, t, t^2\}$ and the standard basis for $\mathbb{R}^3$ is $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

② Find $A = \big[\ [T(\mathbf{b}_1)]_{\mathcal{C}} \ \cdots \ [T(\mathbf{b}_n)]_{\mathcal{C}}\ \big]$:

$$T(\mathbf{b}_1) = T(1) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \big[T(\mathbf{b}_1)\big]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$T(\mathbf{b}_2) = T(t) = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \big[T(\mathbf{b}_1)\big]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$T(\mathbf{b}_3) = T(t^2) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \big[T(\mathbf{b}_1)\big]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

So $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

Ex: If $A_{n\times n} = PDP^{-1}$ is diagonalizable with an invertible matrix $P = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix}$ and a diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$, it defines a linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^n \quad \text{with} \quad T(\mathbf{x}) = A\mathbf{x}.$$

Define a new basis $\mathcal{B} = \{\mathbf{p}_1, \cdots, \mathbf{p}_n\}$ for $\mathbb{R}^n$. What is the standard matrix $[T]_{\mathcal{B}}$?

Answer:

①: Find $\mathcal{B}$ and $\mathcal{C}$:

In this example, the domain and codomain are the same, so their bases are the same: $\mathcal{B} = \mathcal{C} = \{\mathbf{p}_1, \cdots, \mathbf{p}_n\}$ as is given above.

② Find $[T]_{\mathcal{B}} = \begin{bmatrix} \big[T(\mathbf{p}_1)\big]_{\mathcal{B}} & \cdots & \big[T(\mathbf{p}_n)\big]_{\mathcal{B}} \end{bmatrix}$:

$$T(\mathbf{p}_1) = A\mathbf{p}_2 = \lambda_1\mathbf{p}_1 \quad \Longrightarrow \quad \big[T(\mathbf{p}_1)\big]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$T(\mathbf{p}_2) = A\mathbf{p}_2 = \lambda_2\mathbf{p}_2 \quad \Longrightarrow \quad \big[T(\mathbf{p}_1)\big]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{bmatrix}$$

$$\vdots$$

$$T(\mathbf{p}_3) = A\mathbf{p}_3 = \lambda_3\mathbf{p}_3 \quad \Longrightarrow \quad \big[T(\mathbf{p}_1)\big]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{bmatrix}$$

So $[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} = D.$

- **Thm** (Diagonal representation thm): Suppose $A = PDP^{-1}$ with a diagonal matrix D. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from columns of P, then D is the $\mathcal{B}$-matrix for the mapping $T : \mathbf{x} \mapsto A\mathbf{x}$.

  More generally, if $A = PCP^{-1}$ where C may not be a diagonal matrix, and $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from columns of P, then C is the $\mathcal{B}$-matrix for the mapping $T : \mathbf{x} \mapsto A\mathbf{x}$.

  $\Longrightarrow$ The standard matrix C can be calculated by $C = P^{-1}AP$.

  Ex: Let $T : \mathbf{x} \mapsto A\mathbf{x}$ with $A = \begin{bmatrix} 3 & 4 \\ -1 & -1 \end{bmatrix}$. Define a basis $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2\}$ with $\mathbf{p}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and $\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. What is the standard matrix $[T]_{\mathcal{B}}$?

  Answer: According to the thm above, $[T]_{\mathcal{B}} = C = P^{-1}AP$ with

  $$P = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \quad \text{and thus} \quad P^{-1} = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}.$$

  So $[T]_{\mathcal{B}} = P^{-1}AP = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 3 & 4 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}.$

Appendix B Complex numbers

Question: What is the eigenvalues of the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

Consider the characteristic polynomial: $|A - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$

What are the roots of $\lambda^2 + 1 = 0$?

- **Def**: Denote by $\mathbf{i}$ the imaginary unit such that $\mathbf{i}^2 = -1$. A complex number is in the form $z = a + b\mathbf{i}$ with $a = \text{Re}z$ being the real part and $b = \text{Im}z$ being the imaginary part.

  Ex: For the complex number $z = 3 + 2\mathbf{i}$, its real part is $\text{Re}z = 3$, and its imaginary part is $\text{Im}z = 2$.

- **Properties**:

  ① $z_1 = z_2 \iff \text{Re}z_1 = \text{Re}z_2$ and $\text{Im}z_1 = \text{Im}z_2$

  ② summation: $(a + b\mathbf{i}) + (c + d\mathbf{i}) = (a + c) + (b + d)\mathbf{i}$

  ③ multiplication: $(a + b\mathbf{i}) \cdot (c + d\mathbf{i}) = (ac - bd) + (bc + ad)\mathbf{i}$

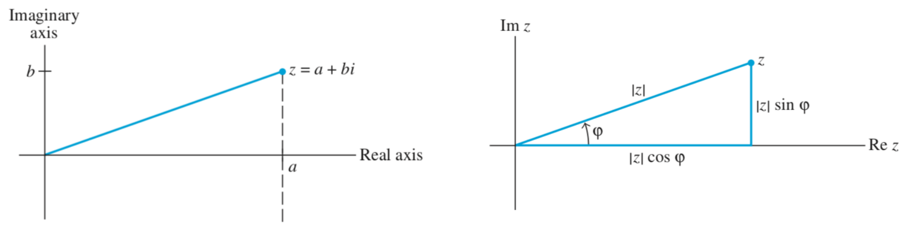
④ the conjugate of $z = a + b\mathbf{i}$ is $\bar{z} = a - b\mathbf{i}$

⑤ the absolute value of $z = a + b\mathbf{i}$ is $|z| = \sqrt{z \cdot \bar{z}} = \sqrt{a^2 + b^2}$

⑥ the inverse of $z = a + b\mathbf{i}$ is $z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z \cdot \bar{z}} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}\mathbf{i}$

Ex: For $z = 3 + 4\mathbf{i}$, we have $\bar{z} = 3 - 4\mathbf{i}$, $|z| = 5$, $z^{-1} = \frac{3}{25} - \frac{4}{25}\mathbf{i}$.

- **Geometric discription**:



Based on these figures, we get $a = |z| \cos \varphi$ and $b = |z| \sin \varphi$.

Hence, there are two ways to determine a complex number:

(1) $z = a + b\mathbf{i}$

(2) $z = |z| \cos \varphi + (|z| \sin \varphi)\, \mathbf{i} = |z|e^{\mathbf{i}\varphi}$

Ex: If $z = |z|e^{\varphi}$, then $z^k = |z|^k e^{\mathbf{i}k\varphi} = |z|^k \cos(k\varphi) + |z|^k \sin(k\varphi)\mathbf{i}$

Ex: Find all real and complex roots of the equation $z^8 = 2^8$.

Answer: Assume that $z = |z|e^{\mathbf{i}\varphi}$. It then suffices to determine $|z|$ and φ.

Note that $z^8 = |z|^8 \cos(8\varphi) + |z|^8 \sin(8\varphi)\mathbf{i} = 2^8$. Their real (resp. imaginary) parts should be the same, that is

Firstly, $|z|^8 \sin(8\varphi) = 0 \Longrightarrow 8\varphi = k\pi$ for any integer k.

Secondly, $|z|^8 \cos(8\varphi) = 2^8$. If $8\varphi = k\pi$, $\cos(8\varphi) = \pm 1$. However, $\cos(8\varphi)$ can not be -1, otherwise we will get a contradiction $-|z|^8 = 2^8$. So we finally get $8\varphi = 2k\pi$, that is, $\varphi = \frac{k\pi}{4}$ such that $\cos(8\varphi) = 1$. Hence, $|z| = 2$.

So $z = 2e^{\mathbf{i}\frac{k\pi}{4}}$, k can be any integer.

5.5 Complex eigenvalues

- **Ex**: Let $A = \begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$. What are its eigenvalues and corresponding eigenvectors?

① Find all the eigenvalues: $|A - \lambda I| = \begin{vmatrix} 1-\lambda & -2 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 5 = (\lambda - 2)^2 + 1$.

$\Longrightarrow A$ has eigenvalues $\lambda = 2 \pm \mathbf{i}$

② Find corresponding eigenvectors:

For $\lambda_1 = 2+\mathbf{i}$, $\begin{bmatrix} A-\lambda_1 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} -1-\mathbf{i} & -2 & 0 \\ 1 & 1-\mathbf{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1-\mathbf{i} & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\Longrightarrow$ Solutions $\mathbf{x} = \begin{bmatrix} -1+\mathbf{i} \\ 1 \end{bmatrix} x_2$. Choose $\mathbf{p}_1 = \begin{bmatrix} -1+\mathbf{i} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}\mathbf{i}$.

For $\lambda_2 = 2-\mathbf{i}$, $\begin{bmatrix} A-\lambda_2 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} -1+\mathbf{i} & -2 & 0 \\ 1 & 1+\mathbf{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1+\mathbf{i} & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\Longrightarrow$ Solutions $\mathbf{x} = \begin{bmatrix} -1-\mathbf{i} \\ 1 \end{bmatrix} x_2$. Choose $\mathbf{p}_1 = \begin{bmatrix} -1-\mathbf{i} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\mathbf{i}$.

$\Longrightarrow$ In this example, we have $\lambda_2 = \overline{\lambda_1}$ and $\mathbf{p}_2 = \overline{\mathbf{p}_1}$.

$\Longrightarrow$ If $A\mathbf{p} = \lambda\mathbf{p}$, then $A\overline{\mathbf{p}} = \overline{\lambda}\overline{\mathbf{p}}$. (If λ is an eigenvalue of A, then $\overline{\lambda}$ is also an eigenvalue)

For a real matrix A, its complex eigenvalues occur in conjugate pairs.

- **Ex**: For $A_{2\times 2}$ given above, consider one of the eigenvalues $\lambda = 2-\mathbf{i}$ and its corresponding eigenvector $\mathbf{p} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\mathbf{i}$.

  Denote $P = \begin{bmatrix} \mathrm{Re}\mathbf{p} & \mathrm{Im}\mathbf{p} \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$. Is there a matrix C such that $A = PCP^{-1}$?

  Answer: $C = P^{-1}AP = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \left(= \begin{bmatrix} \mathrm{Re}\lambda & \mathrm{Im}\lambda \\ -\mathrm{Im}\lambda & \mathrm{Re}\lambda \end{bmatrix} \right)$

- **Thm**: Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - b\mathbf{i}$ $(b \neq 0)$ and an associated eigenvector $\mathbf{p}$. Then $A = PCP^{-1}$ with $P = \begin{bmatrix} \mathrm{Re}\mathbf{p} & \mathrm{Im}\mathbf{p} \end{bmatrix}$ and $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

  Ex: For $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with $|C-\lambda I| = (a-\lambda)^2 + b^2$, its eigenvalues are $\lambda = a \pm b\mathbf{i}$ with $|\lambda| = \sqrt{a^2+b^2}$. Then

  $$C = |\lambda| \begin{bmatrix} \frac{a}{|\lambda|} & -\frac{b}{|\lambda|} \\ \frac{b}{|\lambda|} & \frac{a}{|\lambda|} \end{bmatrix} = |\lambda| \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

  which is a composition of a rotation through the angle θ and a scaling by $|\lambda|$.

  Ex: Let $C = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$. What are the rotation angle θ and the scaling constant $|\lambda|$?

  Answer: $|\lambda| = \sqrt{(\sqrt{3})^2 + 1^2} = 2$.

  The angle θ satisfies $\cos\theta = \frac{a}{|\lambda|} = \frac{\sqrt{3}}{2}$ and $\sin\theta = \frac{1}{2}$. Hence, $\theta = \frac{\pi}{6}$.

5.7 Applications to differential equations

- For $y'(t) = \lambda y(t)$, $t \geq 0$, all its solutions are in the form $y(t) = ce^{\lambda t}$ with a free parameter c. No matter what c is, $y(t)$ above is a solution of the differential equation.

  If, in addition, the initial value is given $y(0) = y_0$, then the constant c is determined and the solution is unique: $y(t) = y_0 e^{\lambda t}$.

  If $\lambda < 0$, the solution $y(t)$ will go to 0 as $t \to +\infty$.

  If $\lambda > 0$, the solution $y(t)$ will go to positive or negative infinity as $t \to +\infty$.

- For a system of linear differential equations

$$\begin{cases} y_1'(t) = \lambda_1 y_1(t) \\ y_2'(t) = \lambda_2 y_2(t) \\ \qquad \vdots \\ y_n'(t) = \lambda_n y_n(t) \end{cases} \iff \begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix} \iff Y'(t) = DY(t),$$

  it has solutions $\begin{cases} y_1(t) = c_1 e^{\lambda_1 t} \\ \qquad \vdots \\ y_n(t) = c_n e^{\lambda_n t} \end{cases}$

- What are the solutions of $X'(t) = AX(t)$ if A is not a diagonal matrix as above?

  If $A = PDP^{-1}$, then $X'(t) = PDP^{-1}X(t) \iff [P^{-1}X(t)]' = D\,[P^{-1}X(t)]$.

  Denote $Y(t) = P^{-1}X(t)$, we get $Y'(t) = DY(t)$. Solve this auxiliary equation to get $Y(t)$ and then get $X(t) = PY(t)$.

♣ **Ex**: Solve $X'(t) = AX(t)$ with $A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$ and $X(0) = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

Answer:

① Find D and P:

$|A - \lambda I| = (\lambda + 1)(\lambda + 2) \Longrightarrow \lambda = -1, -2 \Longrightarrow D = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$.

For $\lambda_1 = -1$, $\begin{bmatrix} 2 & -2 & 0 \\ 3 & -3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_2 \Longrightarrow \mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For $\lambda_1 = -2$, $\begin{bmatrix} 3 & -2 & 0 \\ 3 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix} x_2 \Longrightarrow \mathbf{p}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

So $P = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$.

② Solve $Y'(t) = DY(t)$ and get $X(t) = PY(t)$:

Based on D, $\begin{cases} y_1(t) = c_1 e^{-t} \\ y_2(t) = c_2 e^{-2t} \end{cases} \Longrightarrow Y(t) = \begin{bmatrix} c_1 \\ 0 \end{bmatrix} e^{-t} + \begin{bmatrix} 0 \\ c_2 \end{bmatrix} e^{-2t}$. Hence,

$$X(t) = PY(t) = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \left(\begin{bmatrix} c_1 \\ 0 \end{bmatrix} e^{-t} + \begin{bmatrix} 0 \\ c_2 \end{bmatrix} e^{-2t} \right) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} e^{-2t}$$

$\Longrightarrow X(t) = c_1 \mathbf{p}_1 e^{\lambda_1 t} + c_2 \mathbf{p}_2 e^{\lambda_2 t}$

③ Use $X(0)$ to determine c_1 and c_2:

Based on the formula above and the initial condition,

$$X(0) = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Solve $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & -1 \end{bmatrix}$, and get $c_1 = 5$ and $c_2 = -1$.

- **Def**: For $X'(t) = AX(t)$, denote by λ the eigenvalues of A.

  1. If $\lambda < 0$, the origin is an attractor/sink.

  The direction of greatest attraction is corresponding to the most negative eigenvalue.

  2. If $\lambda > 0$, the origin is a repeller/source.

  The direction of greatest repulsion is corresponding to the largest positive eigenvalue.

  3. If λ has both positive and negative values, the origin is a saddle point.

- If $A_{2\times 2}$ has a pair of complex eigenvalues λ and $\overline{\lambda}$ with $\mathbf{p}$ and $\overline{\mathbf{p}}$, then

  $X(t) = c_1 \mathbf{p} e^{\lambda t} + c_2 \overline{\mathbf{p}} e^{\overline{\lambda} t}$ are complex solutions!

  Denote $X_1 = \mathbf{p} e^{\lambda t}$ and $X_2 = \overline{\mathbf{p}} e^{\overline{\lambda} t}$. It holds $X_2 = \overline{X_1}$.

  $$\Longrightarrow \begin{cases} \dfrac{X_1 + X_2}{2} = \text{Re}\left[\mathbf{p} e^{\lambda t}\right] \\ \dfrac{X_1 - X_2}{2\mathbf{i}} = \text{Im}\left[\mathbf{p} e^{\lambda t}\right] \end{cases}$$

  $\Longrightarrow X(t) = \tilde{c}_1 \text{Re}\left[\mathbf{p} e^{\lambda t}\right] + \tilde{c}_2 \text{Im}\left[\mathbf{p} e^{\lambda t}\right]$ are the real solutions!

  Ex: Find all the real solutions of $X'(t) = AX(t)$ with $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$.

  ① Find all the eigenvalues: $|A - \lambda I| = (\lambda + 2)^2 + 1 \Longrightarrow \lambda = -2 \pm \mathbf{i}$

  Since the eigenvalues are complex and form a conjugate pair, we only need to use one of them.

② Choose λ and calculate $\mathbf{p}$: Choose $\lambda = -2 + \mathbf{i}$, and solve

$$\begin{bmatrix} A - \lambda I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} -1-\mathbf{i} & 2 & 0 \\ -1 & 1-\mathbf{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1+\mathbf{i} & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \mathbf{x} = \begin{bmatrix} 1-\mathbf{i} \\ 1 \end{bmatrix} x_2$$

to get $\mathbf{p} = \begin{bmatrix} 1-\mathbf{i} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \mathbf{i}$.

③ Calculate $\mathrm{Re}\left[\mathbf{p}e^{\lambda t}\right]$ and $\mathrm{Im}\left[\mathbf{p}e^{\lambda t}\right]$:

$$\begin{aligned} \mathbf{p}e^{\lambda t} &= \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \mathbf{i}\right) e^{-2t+\mathbf{i}t} \\ &= e^{-2t}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \mathbf{i}\right)(\cos t + \sin t\ \mathbf{i}) \\ &= e^{-2t}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\cos t - \begin{bmatrix} -1 \\ 0 \end{bmatrix}\sin t\right) + e^{-2t}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\sin t + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\cos t\right)\mathbf{i} \end{aligned}$$

$$\begin{aligned} \Longrightarrow \mathrm{Re}\left[\mathbf{p}e^{\lambda t}\right] &= e^{-2t}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\cos t - \begin{bmatrix} -1 \\ 0 \end{bmatrix}\sin t\right) \\ \mathrm{Im}\left[\mathbf{p}e^{\lambda t}\right] &= e^{-2t}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\sin t + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\cos t\right) \end{aligned}$$

- In this case, the origin is a spiral point.

$$\begin{cases} \text{the trajectories of the solution spiral inward if } \mathrm{Re}\lambda < 0 \\ \text{the trajectories of the solution spiral outward if } \mathrm{Re}\lambda > 0 \end{cases}$$

6 Chapter 6

6.1 Inner product, length, and orthogonality

- **Def**: For two vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, their inner product is

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\top \mathbf{v} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n$$

$\Longrightarrow$ Properties:

① $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, $\quad (\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$, $\quad (c\mathbf{u}) \cdot v = \mathbf{u} \cdot (c\mathbf{v}) = c\mathbf{u} \cdot \mathbf{v}$

② $\mathbf{u} \cdot \mathbf{u} \geq 0$ for any $\mathbf{u}$ in $\mathbb{R}^n$; $\quad \mathbf{u} \cdot \mathbf{u} = 0 \iff \mathbf{u} = \mathbf{0}$

- **Def**: For $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ in $\mathbb{R}^n$, the length (norm) of $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + \cdots + u_n^2}$$

$\Longrightarrow$ Properties:

① If $\|\mathbf{u}\| = 1$, then $\mathbf{u}$ is called a unit vector.

② If $\|\mathbf{u}\| \neq 1$, then it can be normalized as $\widehat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|}\mathbf{u}$.

- **Def**: For $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, the distance between $\mathbf{u}$ and $\mathbf{v}$ is

$$dist(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Ex: Given $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Calculate the following quantities.

$$\mathbf{u} \cdot \mathbf{v} = 1, \quad \|\mathbf{u}\| = \sqrt{3^2 + 4^2} = 5, \quad dist(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \left\| \begin{bmatrix} 4 \\ 3 \end{bmatrix} \right\| = 5$$

- **Def**: For $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, they are orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

$\Longrightarrow$ Properties:

① $\mathbf{0}$ is orthogonal to any vectors in $\mathbb{R}^n$.

② $\mathbf{u}$ and $\mathbf{v}$ are orthogonal $\iff \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

Ex: Given $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$. Then $\mathbf{u} \cdot \mathbf{v} = 0$, and

$$\|\mathbf{u}+\mathbf{v}\|^2 = \left\| \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix} \right\|^2 = 1 + 3^2 + 3^2 = 19,$$

$$\|\mathbf{u}\|^2 = 1 + 2^2 + 3^2 = 14, \quad \|\mathbf{v}\|^2 = (-2)^2 + 1^2 + 0^2 = 5.$$

Hence, it holds $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

- **Def**: Let W be a subspace of $\mathbb{R}^n$. A vector $\mathbf{z}$ in $\mathbb{R}^n$ is called orthogonal to W if $\mathbf{z}$ is orthogonal to each vector in W. Denote the set

  $$W^{\perp} = \{\mathbf{z} : \mathbf{z} \text{ is orthogonal to } W\}$$

  $\Longrightarrow$ Properties:

  ① $W^{\perp}$ is also a subspace of $\mathbb{R}^n$, which is orthogonal to W.

  ② $(\text{Row}A)^{\perp} = \text{Nul}A = \left(\text{Col}A^{\top}\right)^{\perp}$

6.2 Orthogonal sets

- **Def**: A set of vectors $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is an orthogonal set if any two vectors inside are orthogonal.

- **Thm**: An orthogonal set of nonzero vectors is also a linearly independent set.

  Ex: The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \right\}$ is linearly independent, but is not orthogonal.

  The set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\}$ is both linearly independent and orthogonal.

- **Def**: An orthogonal basis for a subspace W is a basis that is also an orthogonal set.

  An orthonormal basis for W is a basis that is also an orthogonal set containing only unit vectors.

  Ex: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \right\}$ is a basis.

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\} \text{ is an orthogonal basis.}$$

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ is an orthonormal basis.}$$

- **Thm**: Let $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ be an orthogonal basis for W. Then for each $\mathbf{y}$ in W,

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p \qquad \text{with} \quad c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}, \quad j = 1, 2, \cdots, p.$$

- **Def**: Given two vectors $\mathbf{y}$ and $\mathbf{u}$. Rewrite $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$ such that $\widehat{\mathbf{y}} = c\mathbf{u}$ is a scalar multiple of $\mathbf{u}$, and $\mathbf{z}$ is orthogonal to $\mathbf{u}$.

  Then $\widehat{\mathbf{y}} = c\mathbf{u} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$ is the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$.

  The distance from $\mathbf{y}$ to the line through $\mathbf{u}$ is $\|\mathbf{z}\| = \|\mathbf{y} - \widehat{\mathbf{y}}\|$.

  Ex: Let $\mathbf{y} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. What is the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$?

  Answer: The projection

$$\widehat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} = \frac{\mathbf{10}}{\mathbf{20}} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix},$$

  and

$$\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}} = \begin{bmatrix} 1 \\ 7 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

  such that $\mathbf{y} \cdot \mathbf{z} = 0$. That is $\widehat{\mathbf{y}}$ and $\mathbf{z}$ are orthogonal.

- **Thm**: The matrix $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_p \end{bmatrix}_{m \times p}$ has orthonormal columns $\iff U^\top U = I$.

  Reason:

$$U^\top U = \begin{bmatrix} \mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_p^\top \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^\top\mathbf{u}_1 & \mathbf{u}_1^\top\mathbf{u}_2 & \cdots & \mathbf{u}_1^\top\mathbf{u}_p \\ \mathbf{u}_1^\top\mathbf{u}_2 & \mathbf{u}_2^\top\mathbf{u}_2 & \cdots & \mathbf{u}_2^\top\mathbf{u}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_p^\top\mathbf{u}_1 & \mathbf{u}_p^\top\mathbf{u}_2 & \cdots & \mathbf{u}_p^\top\mathbf{u}_p \end{bmatrix} = I$$

6.3 Orthogonal projections

- **Thm** (The orthogonal decomposition thm): Let W be a subspace of $\mathbb{R}^n$. Then any vector $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$ with $\widehat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$.

  If W has an orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$, then the orthogonal projection of $\mathbf{y}$ onto W, which is also denoted by $\widehat{\mathbf{y}} = \text{proj}_W\mathbf{y}$, is

$$\widehat{\mathbf{y}} = \text{proj}_W\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p.$$

  $\Longrightarrow$ **Remark**: If $\mathbf{y}$ is in W, then $\text{proj}_W\mathbf{y} = \mathbf{y}$ and $\mathbf{z} = \mathbf{0}$.

  Ex: Given $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Find the orthogonal projection of $\mathbf{y}$ onto $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

  Answer: Noting that $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W. Hence, the orthogonal decomposition thm can be used directly:

$$\widehat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}.$$

- **Thm** (The best approximation thm): Let W be a subspace of $\mathbb{R}^n$. Then the orthogonal projection $\widehat{\mathbf{y}}$ of $\mathbf{y}$ onto W is the closest point(best approximation) in W to $\mathbf{y}$. That is,

$$\|\mathbf{y} - \widehat{\mathbf{y}}\| \le \|\mathbf{y} - \mathbf{v}\| \quad \text{for any } \mathbf{v} \in W.$$

  $\Longrightarrow$ $\|\mathbf{z}\| = \|\mathbf{y} - \widehat{\mathbf{y}}\|$ denotes the distance from $\mathbf{y}$ to W.

- **Ex**: Given $\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$.

  ① Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ an orthogonal basis? $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$ Yes.

  ② Find the orthogonal projection of $\mathbf{y}$ onto $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$:

$$\widehat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{35}{35}\begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix} + \frac{-28}{14}\begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \\ -1 \end{bmatrix}$$

  ③ Find the closest point to $\mathbf{y}$ in $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$: same as above $\begin{bmatrix} 3 \\ -9 \\ -1 \end{bmatrix}$

④ Find the best approximation of $\mathbf{y}$ in W =Span$\{\mathbf{u}_1, \mathbf{u}_2\}$: same as above $\begin{bmatrix} 3 \\ -9 \\ -1 \end{bmatrix}$

⑤ What is the distance from $\mathbf{y}$ to W? $\|\mathbf{z}\| = \|\mathbf{y} - \widehat{\mathbf{y}}\| = \left\| \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} \right\| = \sqrt{40}$

- **Thm**: If $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W in $\mathbb{R}^n$, then

$$\widehat{\mathbf{y}} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p.$$

If $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_p \end{bmatrix}$, then $U^\top U = I$.

6.4 The Gram–Schmidt process

- **Ex**: Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$ with $\{\mathbf{x}_1, \mathbf{x}_2\}$ being a basis. To obtain an orthogonal basis for W, define

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{x}_1 \\ \mathbf{u}_2 &= \mathbf{x}_2 - \text{proj}_{\mathbf{u}_1}\mathbf{x}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \end{aligned}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W.

For example, $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Then

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ \mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \end{aligned}$$

and apparently $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$.

♣ **Thm** (The Gram–Schmidt process): Given a basis $\{\mathbf{x}_1, \cdots, \mathbf{x}_p\}$ for W. Then

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{x}_1 \\ \mathbf{u}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\ \mathbf{u}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \end{aligned}$$

$$\vdots$$

$$\mathbf{u}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{u}_{p-1}}{\mathbf{u}_{p-1} \cdot \mathbf{u}_{p-1}}\mathbf{u}_{p-1}$$

form an orthogonal basis for W. In addition,

$$\text{Span}\{\mathbf{x}_1, \cdots, \mathbf{x}_k\} = \text{Span}\{\mathbf{u}_1, \cdots, \mathbf{u}_k\} \quad \text{for any} \quad k = 1, 2, \cdots, p.$$

- **Ex**: Given $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}$. Then the column space Col$A$ = Span$\{\mathbf{a}_1, \mathbf{a}_2\}$ has a basis $\{\mathbf{a}_1, \mathbf{a}_2\}$ since the columns $\mathbf{a}_1, \mathbf{a}_2$ of A are linearly independent.

① Find an orthogonal basis for Col$A$.

$$\mathbf{u}_1 = \mathbf{a}_1 = \begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 = \begin{bmatrix} 9 \\ 7 \\ -5 \\ 5 \end{bmatrix} - \frac{72}{36}\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix}$$

② Find an orthonormal basis for Col$A$.

$$\mathbf{v}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{6}\begin{bmatrix} 5 \\ 1 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ \frac{1}{6} \\ -\frac{1}{2} \\ \frac{1}{6} \end{bmatrix}$$

$$\mathbf{v}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{6}\begin{bmatrix} -1 \\ 5 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{6} \\ \frac{5}{6} \\ \frac{1}{6} \\ \frac{1}{2} \end{bmatrix}$$

③ Denote a matrix $Q = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}$, which satisfies $Q^\top Q = I$. If $A = QR$, then

$$R = Q^\top A = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & -\frac{1}{2} & \frac{1}{6} \\ -\frac{1}{6} & \frac{5}{6} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix},$$

which is a triangular matrix with positive diagonals.

- **Thm** (The QR factorization): If $A_{m\times n}$ has linearly independent columns, then $A = QR$ with columns of $Q_{m\times n}$ forming an orthonormal basis for Col$A$ and $R_{n\times n}$ being an upper triangular matrix with positive diagonals.

  $\Longrightarrow$ It implies that R is invertible.

6.5 Least-squares problems

If $A\mathbf{x} = \mathbf{b}$ has no solution but A has linearly independent columns, then $A = QR$ and

$$Q^\top QR\mathbf{x} = Q^\top \mathbf{b} \iff R\mathbf{x} = Q^\top \mathbf{b} \iff \mathbf{x} = R^{-1}Q^\top \mathbf{b}$$

Apparently, $\mathbf{x}$ above can not be a solution of $A\mathbf{x} = \mathbf{b}$. What is the meaning of $\mathbf{x}$?

- **Def**: A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a vector $\widehat{\mathbf{x}} \in \mathbb{R}^n$ such that

$$\|\mathbf{b} - A\widehat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\| \quad \text{for any} \quad \mathbf{x} \in \mathbb{R}^n.$$

$\Longrightarrow$ For any $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$, $A\mathbf{x} = \mathbf{a}_1 x_1 + \cdots + \mathbf{a}_n x_n \in \text{Col}A$. Then

$$A\widehat{\mathbf{x}} = \text{proj}_{\text{Col}A}\mathbf{b} \quad \text{is the orthogonal projection of } \mathbf{b} \text{ onto Col}A$$
$$\mathbf{b} - A\widehat{\mathbf{x}} \quad \text{is orthogonal to Col}A$$

That is, $\mathbf{b} - A\widehat{\mathbf{x}}$ is orthogonal to $\mathbf{a}_1, \cdots, \mathbf{a}_n$:

$$\left.\begin{array}{c} \mathbf{a}_1 \cdot (\mathbf{b} - A\widehat{\mathbf{x}}) = \mathbf{a}_1^\top (\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \\ \vdots \\ \mathbf{a}_n \cdot (\mathbf{b} - A\widehat{\mathbf{x}}) = \mathbf{a}_n^\top (\mathbf{b} - A\widehat{\mathbf{x}}) = 0 \end{array}\right\} \iff A^\top (\mathbf{b} - A\widehat{\mathbf{x}}) = \mathbf{0}$$

$$\iff A^\top A\widehat{\mathbf{x}} = A^\top \mathbf{b} \quad \text{(normal equation)}$$

- **Thm**: The least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincide with the solutions of the normal equation $A^\top A\widehat{\mathbf{x}} = A^\top \mathbf{b}$.

  Ex: Given $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

  ① Does $A\mathbf{x} = \mathbf{b}$ have solutions? $\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ No solution!

  ② Find the least-squares solutions of $A\mathbf{x} = \mathbf{b}$: Consider $A^\top A\widehat{\mathbf{x}} = A^\top \mathbf{b}$.

$$A^\top A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix}, \quad A^\top \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

The augmented matrix is $\begin{bmatrix} A^\top A & A^\top \mathbf{b} \end{bmatrix} = \begin{bmatrix} 3 & 6 & 6 \\ 6 & 12 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, and the solutions are in the form $\widehat{\mathbf{x}} = \begin{bmatrix} 2 - 2x_2 \\ x_2 \end{bmatrix}$ with x_2 being a free parameter.

$\Longrightarrow$ There are infinitely many least-squares solutions since $A^\top A$ is not invertible.

♣ **Ex**: Given $A = \begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$. Find the least-squares solution of $A\mathbf{x} = \mathbf{b}$.

Answer: Consider the normal equation $A^\top A \widehat{\mathbf{x}} = A^\top \mathbf{b}$.

$$A^\top A = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix}$$

$$A^\top \mathbf{b} = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \end{bmatrix}$$

The augmented matrix is $\begin{bmatrix} A^\top A & A^\top \mathbf{b} \end{bmatrix} = \begin{bmatrix} 6 & -11 & -4 \\ -11 & 22 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$, and hence $\widehat{\mathbf{x}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

$\Longrightarrow$ There is a unique least-squares solution of $A\mathbf{x} = \mathbf{b}$ since $A^\top A$ is invertible.

- **Thm**: $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution

  $\Longleftrightarrow$ $A^\top A$ is invertible

  $\Longleftrightarrow$ A has linearly independent columns

  Remark: In this case, A has linearly independent columns, then $A = QR$ and

  $$A^\top A = (QR)^\top (QR) = R^\top Q^\top QR = R^\top R$$

  is also invertible since R is invertible. Then the unique least-squares solution of $A\mathbf{x} = \mathbf{b}$ is

  $$\widehat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b} = \left(R^\top R\right)^{-1} R^\top Q^\top \mathbf{b} = R^{-1} Q^\top \mathbf{b},$$

  which answers the question proposed at the beginning of this lesson.

6.7 Inner product spaces

- **Def**: An inner product on a general vector space V is a function $\langle u, v\rangle$ such that

  1. $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$, $\quad \langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$, $\quad \langle c\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, c\mathbf{v}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$
  2. $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u}\rangle = 0$ iff $\mathbf{u} = \mathbf{0}$

  A vector space equipped with an inner product is called an inner product space.

  Ex: $\mathbb{R}^n$ with $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^\top\mathbf{v}$

  Ex: $\mathbb{P}_2$: Define an inner product by evaluation at $-1, 0, 1$

  $$\langle p(t), q(t)\rangle = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$$

  For example, let $x_1(t) = 1+t$ and $x_2(t) = 1-t$. Then

  $$\langle x_1(t), x_2(t)\rangle = x_1(-1)x_2(-1) + x_1(0)x_2(0) + x_1(1)x_2(1) = 1$$

  $$\langle x_1(t), x_1(t)\rangle = 0 + 1 + 4 = 5$$

  $\Longrightarrow$ norm(length): $\|x_1(t)\| = \sqrt{\langle x_1(t), x_1(t)\rangle} = \sqrt{5}$

  $\Longrightarrow$ distance between $x_1(t)$ and $x_2(t)$: $\|x_1(t) - x_2(t)\| = \sqrt{\langle 2t, 2t\rangle} = \sqrt{4+0+4} = \sqrt{8}$

- Gram–Schmidt process: basis $\{\mathbf{x}_1, \cdots, \mathbf{x}_p\} \longrightarrow$ orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$

  $$\begin{aligned}
  \mathbf{x}_1 &= \mathbf{u}_1 \\
  \mathbf{x}_2 &= \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 \\
  &\vdots \\
  \mathbf{x}_p &= \mathbf{x}_p - \frac{\langle \mathbf{x}_p, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 - \cdots - \frac{\langle \mathbf{x}_p, \mathbf{u}_{p-1}\rangle}{\langle \mathbf{u}_{p-1}, \mathbf{u}_{p-1}\rangle}\mathbf{u}_{p-1}
  \end{aligned}$$

  Ex: As above, transform $\{x_1(t), x_2(t)\}$ into an orthogonal basis $\{u_1(t), u_2(t)\}$.

  Answer: $u_1(t) = x_1(t) = 1+t$

  $$u_2(t) = x_2(t) - \tfrac{\langle x_2(t), u_1(t)\rangle}{\langle u_1(t), u_1(t)\rangle}u_1(t) = (1-t) - \tfrac{1}{5}(1+t) = \tfrac{4}{5} - \tfrac{6}{5}t$$

- Best approximation: W has an orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$, then for any vector $\mathbf{y}$, $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$ with

  $$\widehat{\mathbf{y}} = \frac{\langle \mathbf{y}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \cdots + \frac{\langle \mathbf{y}, \mathbf{u}_p\rangle}{\langle \mathbf{u}_p, \mathbf{u}_p\rangle}\mathbf{u}_p$$

  Ex: As above, find the best approximation of $y(t) = t^2$ in $W = \{x_1(t), x_2(t)\}$.

  Answer: ① Find an orthogonal basis: $\{x_1(t), x_2(t)\} \to \{u_1(t), u_2(t)\}$

② Find the best approximation (orthogonal projection)

$$\widehat{y}(t) = \frac{\langle y(t), u_1(t)\rangle}{\langle u_1(t), u_1(t)\rangle} u_1(t) + \frac{\langle y(t), u_2(t)\rangle}{\langle u_2(t), u_2(t)\rangle} u_2(t) = \frac{2}{5}(1+t) + \frac{8/5}{24/5}\left(\frac{4}{5} - \frac{6}{5}t\right) = \frac{2}{3}.$$

- **Thm** (The Cauchy–Schwarz inequality): $|\langle \mathbf{u}, \mathbf{v}\rangle| \le \|\mathbf{u}\|\|\mathbf{v}\|$

  Reason: $|\langle \mathbf{u}, \mathbf{v}\rangle| = |\langle c\mathbf{v} + \mathbf{z}, \mathbf{v}\rangle| = |c\langle \mathbf{v}, \mathbf{v}\rangle| = \|c\mathbf{v}\|\|\mathbf{v}\| \le \|\mathbf{u}\|\|\mathbf{v}\|$

- **Thm** (The triangle inequality): $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$

♣ **Ex**: Let $V = C[-1, 1]$ be the space of all continuous functions on $[-1, 1]$. Define an inner product

$$\langle p(t), q(t)\rangle = \int_{-1}^{1} p(t)q(t)dt.$$

Let $x_1(t) = 1$ and $x_2(t) = 2t - 1$. Then $\langle x_1(t), x_2(t)\rangle = \int_{-1}^{1}(2t-1)dt = -2 \neq 0$.

That is, $\{x_1, x_2\}$ are linearly independent but not orthogonal.

Find an orthogonal basis for $W = \text{Span}\{x_1, x_2\}$:

$$p_1(t) = x_1(t) = 1$$
$$p_2(t) = x_2(t) - \frac{\langle x_2(t), p_1(t)\rangle}{\langle p_1(t), p_1(t)\rangle} p_1(t) = (2t-1) - \frac{-2}{2}1 = 2t$$

Then $\{1, 2t\}$ is an orthogonal basis for W.

7 Chapter 7

7.1 Diagonalization of symmetric matrices

- **Def**: A symmetric matrix is a square matrix such that $A^\top = A$.

  Ex: Are the following matrices symmetric?

  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ Yes $\qquad \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$ No $\qquad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$ Yes

- **Def**: P is an orthogonal matrix if $P^{-1} = P^\top$, that is, columns of P are orthonormal.

  Ex: Are the following matrices orthogonal matrices?

  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ Yes $\qquad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ No $\qquad P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ Yes $\Longrightarrow P^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

- **Def**: A is called orthogonally diagonalizable if $A = PDP^\top$ with an orthogonal matrix P and a diagonal matrix D.

- **Thm** A is orthogonally diagonalizable $\iff$ A is symmetric ($A^\top = A$)

  Ex: Let $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ with distinct eigenvalues $-2, 7$.

  Decompose A such that $A = PDP^\top$:

  ① Find linearly independent eigenvectors:

  For $\lambda_1 = -2$, $\begin{bmatrix} A - \lambda_1 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 5 & -2 & 4 & 0 \\ -2 & 8 & 2 & 0 \\ 4 & 2 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & 0 & 1 & 0 \\ 0 & ① & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

  It has solutions $\mathbf{x} = \begin{bmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{bmatrix} x_3$. We can choose the first eigenvector $\mathbf{v}_1 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$

  For $\lambda_2 = 7$, $\begin{bmatrix} A - \lambda_2 I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} -4 & -2 & 4 & 0 \\ -2 & -1 & 2 & 0 \\ 4 & 2 & -4 & 0 \end{bmatrix} \sim \begin{bmatrix} ① & \frac{1}{2} & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

  It has solutions $\mathbf{x} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_3$.

  We can choose another two eigenvectors $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

② Find orthogonal eigenvectors:

Note that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = 0$ and $\mathbf{v}_2 \cdot \mathbf{v}_3 = -1$. Based on the Gram–Schmidt process:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{-1}{5}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \\ 1 \end{bmatrix}$$

③ Find orthonormal eigenvectors:

$$\mathbf{p}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

$$\mathbf{p}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{bmatrix}$$

$$\mathbf{p}_3 = \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = \frac{5}{3\sqrt{5}}\begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{3\sqrt{5}} \\ \frac{2}{3\sqrt{5}} \\ \frac{\sqrt{5}}{3} \end{bmatrix}$$

Then $P = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}$ and $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$ such that $A = PDP^\top$.

- **Spectral decomposition** of $A = PDP^\top$ with $P = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix}$:

$$A = \begin{bmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}\begin{bmatrix} \mathbf{p}_1^\top \\ \vdots \\ \mathbf{p}_n^\top \end{bmatrix} = \lambda_1 \mathbf{p}_1 \mathbf{p}_1^\top + \cdots + \lambda_n \mathbf{p}_n \mathbf{p}_n^\top$$

$\Longrightarrow$ Matrices $\mathbf{p}_i\mathbf{p}_i^\top$ above are called projection matrices:

$$\left(\mathbf{p}_i\mathbf{p}_i^\top\right)\mathbf{x} = \mathbf{p}_i\left(\mathbf{p}_i^\top\mathbf{x}\right) = \mathbf{p}_i\left(\mathbf{p}_i \cdot \mathbf{x}\right) = \frac{\mathbf{x} \cdot \mathbf{p}_i}{\mathbf{p}_i \cdot \mathbf{p}_i}\mathbf{p}_i$$