AMD

Xiaolu Hou

FIIT, STU
xiaolu.hou @ stuba.sk

Course Outline

- Vectors and matrices
- System of linear equations
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices
- Vectors and matrices

Recommended reading

- Andrilli, Stephen, and David Hecker. Elementary linear algebra. Academic Press, 2022. Fifth edition
  - Sections 1.1, 1.2, 1.4, 1.5
  - Free copy online
- Anton, Howard, and Chris Rorres. Elementary linear algebra: applications version. John Wiley & Sons, 2013.
  - Sections 1.3
  - Free copy online

Lecture Outline

- Definitions
- Vectors
- Matrices

Vectors and matrices

- Definitions
- Vectors
- Matrices

Sets

- $\emptyset$: empty set
- $|S|$: cardinality of S
- A set S is *finite* if $|S| < \infty$
- $a \in S$: a is an element in set S
- $a \notin S$: a is not an element in set S
- $S \subseteq T$: if $s \in S$, then $s \in T$, S is a subset of T
- $S = T$: $S \subseteq T$ and $T \subseteq S$
- The *power set* of a set S, denoted by 2^S, is the set of all subsets of S.

Example

Let $T = \{ 0, 1, 2, 3 \}$ and $S = \{ 2, 3 \}$, then

- $S \subseteq T$ and $T \nsubseteq S$.
- $2 \in S$, $0 \notin S$.
- $|S| = 2$, $|T| = 4$.
- $2^S = \{ \emptyset, S, \{ 2 \}, \{ 3 \} \}$.

Sets

- Union: $A \cup B$
- Intersection: $A \cap B$
- Difference: $A - B = \{ a \in A, a \notin B \}$
- Complement of A in S: $A^c = S - A$
- Cartesian product $A \times B = \{ (a, b) \mid a \in A, b \in B \}$
  - ordered pairs

Example

- $A = \{ 0, 1, 2 \}$, $B = \{ 2, 3, 4 \}$
- $A \cup B = \{ 0, 1, 2, 3, 4 \}$, $A \cap B = \{ 2 \}$

Example

- $A = \{ 2, 4, 6 \}$, $B = \{ 1, 3, 5 \}$, $S = A \cup B$
- $A - B = A$. Complement of A in S is B

$$A \times B = \{ (2,1), (2,3), (2,5), (4,1), (4,3), (4,5), (6,1), (6,3), (6,5) \} .$$

Functions

Definition

A *function/map* $f : S \to T$ is a rule that assigns each element $s \in S$ a **unique** element $t \in T$.

- S – *domain* of f; T – *codomain* of f.
- If $f(s) = t$, then t is called the *image* of s, s is a *preimage* of t.
- For any $A \subseteq T$, *preimage of A under f* is

$$f^{-1}(A) := \{\, s \in S \mid f(s) \in A \,\}$$

Example

Define

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

where $\mathbb{R}$ is the set of real numbers. Then f has domain $\mathbb{R}$ and codomain $\mathbb{R}$.

Functions – example

Example

Define

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2 \end{aligned}$$

where $\mathbb{R}$ is the set of real numbers. Then f has domain $\mathbb{R}$ and codomain $\mathbb{R}$.
Let $A = \{ 1 \} \subseteq \mathbb{R}$, the preimage of A under f is given by

$$f^{-1}(A) = \{ -1, 1 \} .$$

1 is the image of -1 and -1 is a preimage of 1. 1 is another preimage of 1.
Let $B = \{ -1 \} \subseteq \mathbb{R}$, then $f^{-1}(B) = \emptyset$.

Composition of functions

Definition

For two functions $f : T \to U$, $g : S \to T$, the *composition* of f and g, denoted by $f \circ g$, is the function

$$\begin{aligned} f \circ g : S &\to U \\ s &\mapsto f(g(s)). \end{aligned}$$

Example

What is $f \circ g$?

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2, \end{aligned}$$

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3. \end{aligned}$$

Composition of functions

Example

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^2, \end{aligned}$$

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto x^3. \end{aligned}$$

$$\begin{aligned} f \circ g : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto (x^3)^2 = x^6. \end{aligned}$$

Matrices

- $\mathbb{R}$: the set of all real numbers

Definition

A *matrix with coefficients in* $\mathbb{R}$ is a rectangular array where each entry is an element of $\mathbb{R}$.

Matrix A is said to have m *rows*, n *columns* and is of size $m \times n$.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Example

The matrix

$$A = \begin{bmatrix} 1 & 7.5 & 6 \\ 2 & 3 & 4 \end{bmatrix}$$

has 2 rows, 3 columns and is of size 2×3.

Vectors

- A $1 \times n$ matrix is called a *row vector*.
- An $n \times 1$ matrix is called a *column vector*.

Example

- $\boldsymbol{a} = \begin{bmatrix} 1, & -1, & 3 \end{bmatrix}$ is a row vector
- $\boldsymbol{b} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$ is a column vector

Note

- By "vector," we refer specifically to a row vector.
- $\mathbb{R}^n$ represents the set of all vectors with n entries, also referred to as *coordinates*.
- When written by hand, $\vec{a}$ is used to denote a vector.

Example

$\boldsymbol{a} \in \mathbb{R}^3$

Vectors and scalars

- Two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ are equal, written $\boldsymbol{a} = \boldsymbol{b}$, if all corresponding coordinates are *equal*
- $\mathbf{0} = \begin{bmatrix} 0, & 0, & \cdots, & 0 \end{bmatrix}$ is the *zero vector*.
- An element $x \in \mathbb{R}$ is called a *scalar*

Example

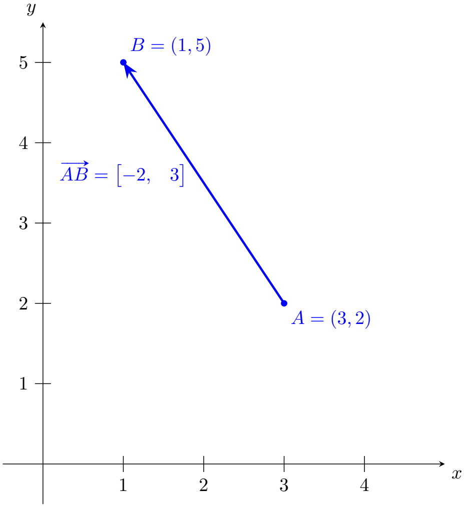
- $\begin{bmatrix} 1, & 0, & 4 \end{bmatrix} \neq \begin{bmatrix} 1, & 0, & -4 \end{bmatrix}$
- $5 \in \mathbb{R}$ is a scalar

Vectors and matrices

- Definitions
- Vectors
- Matrices

Geometric interpretation of vectors

- A vector with two coordinates, i.e. an element of $\mathbb{R}^2$, is frequently used to represent a movement from one point to another in a coordinate plane
- From an initial point $(3,2)$ to a terminal point $(1,5)$, there is a net decrease of 2 units along the $x-$axis and a net increase of 3 units along the $y-$axis. A vector representing this change would thus be $\begin{bmatrix}-2, & 3\end{bmatrix}$, as indicated by the arrow in the figure



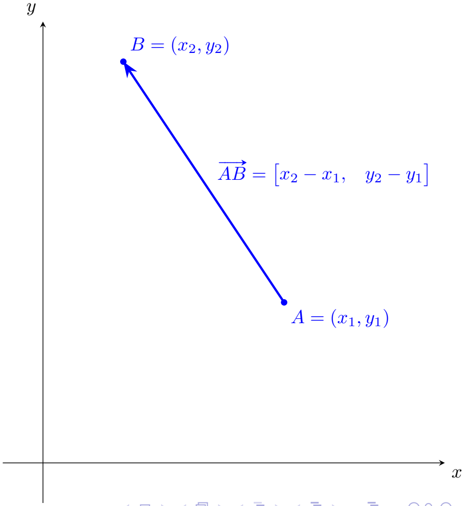
Remark

- Points of a coordinate system: parentheses
- Vectors: brackets

Geometric interpretation of vectors

- In general, a vector starting at point $A = (x_1, y_1)$ and ending at $B = (x_2, y_2)$, denoted $\overrightarrow{AB}$ is given by

$$\overrightarrow{AB} = \begin{bmatrix} x_2 - x_1, & y_2 - y_1 \end{bmatrix}$$



Norm of a vector

- The *distance* between two points (x_1, y_1) and (x_2, y_2) in the plane is given by

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

- The vector between the points is $\begin{bmatrix} x_2 - x_1, & y_2 - y_1 \end{bmatrix}$
- This motivates the following definition

Definition

The *norm* (also called *length*) of a vector $\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix}$, denoted $\|\boldsymbol{a}\|$, is given by

$$\|\boldsymbol{a}\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

A vector of norm 1 is called a *unit vector*

Norm of a vector – Example

Example

- The norm of $\boldsymbol{a} = \begin{bmatrix} 4, & -3, & 0, & 2 \end{bmatrix}$ is given by

$$\|\boldsymbol{a}\| = \sqrt{4^2 + (-3)^2 + 0^2 + 2^2} = \sqrt{16 + 9 + 0 + 4} = \sqrt{29}$$

- $\begin{bmatrix} \frac{3}{5}, & -\frac{4}{5} \end{bmatrix}$ is a unit vector in $\mathbb{R}^2$

$$\sqrt{\left(\frac{3}{5}\right)^2 + \left(-\frac{4}{5}\right)^2} = 1$$

- $\begin{bmatrix} -\frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & -\frac{1}{2} \end{bmatrix}$ is a unit vector in $\mathbb{R}^4$

$$\sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = 1$$

Scalar multiplication

- $\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \in \mathbb{R}^n$
- $\alpha \in \mathbb{R}$
- The scalar multiple of $\boldsymbol{a}$ by α is the vector

$$\alpha\boldsymbol{a} = \begin{bmatrix} \alpha a_1, & \alpha a_2, & \cdots, & \alpha a_n \end{bmatrix}$$

- It is easy to see that

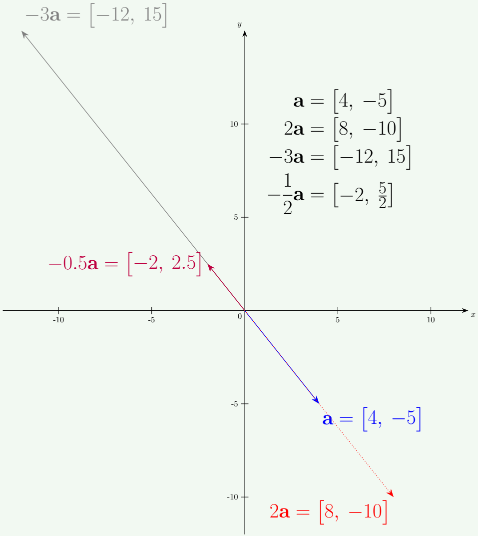
$$\|\alpha\boldsymbol{a}\| = |\alpha|\|\boldsymbol{a}\|$$

since

$$\|\alpha\boldsymbol{a}\| = \sqrt{(\alpha a_1)^2 + (\alpha a_2)^2 + \cdots + (\alpha a_n)^2} = \sqrt{\alpha^2(a_1^2 + a_2^2 + \cdots + a_n^2)} = |\alpha|\|\boldsymbol{a}\|.$$

Scalar multiplication

Example



$$\mathbf{a} = [4,\ -5]$$
$$2\mathbf{a} = [8,\ -10]$$
$$-3\mathbf{a} = [-12,\ 15]$$
$$-\frac{1}{2}\mathbf{a} = [-2,\ \tfrac{5}{2}]$$

Scalar multiplication

$$\|\alpha \boldsymbol{a}\| = |\alpha|\|\boldsymbol{a}\|$$

- Multiplication by α dilates (expands) the norm of the vector when $|\alpha| > 1$ and contracts (shrinks) the norm when $|\alpha| < 1$
- Scalar multiplication by 1 or -1 does not affect the norm
- Scalar multiplication by 0 always yields the zero vector.

Direction

Definition

$\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \mathbf{0}$, $\boldsymbol{b} \neq \mathbf{0}$, $\boldsymbol{a}$ and $\boldsymbol{b}$ are said to be

- *in the same direction* if $\exists \alpha \in \mathbb{R}_{>0}$ s.t. $\boldsymbol{b} = \alpha \boldsymbol{a}$
- *in the opposite direction* if $\exists \alpha \in \mathbb{R}_{<0}$ s.t. $\boldsymbol{b} = \alpha \boldsymbol{a}$
- *parallel* if they are in the same or in the opposite direction

Example

- $\begin{bmatrix} 1, & -3, & 2 \end{bmatrix}$ and $\begin{bmatrix} 3, & -9, & 6 \end{bmatrix}$ are in the same direction

$$\begin{bmatrix} 1, & -3, & 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 3, & -9, & 6 \end{bmatrix}.$$

- $\begin{bmatrix} -3, & 6, & 15 \end{bmatrix}$ and $\begin{bmatrix} 4, & -8, & 20 \end{bmatrix}$ are in the opposite direction

$$\begin{bmatrix} -3, & 6, & 15 \end{bmatrix} = -\frac{3}{4} \begin{bmatrix} 4, & -8, & 20 \end{bmatrix}.$$

Normalization of a vector

Lemma

For any $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \mathbf{0}$,

$$\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|}$$

is unit vector in the same direction as $\boldsymbol{a}$.

Proof

By the above observations

$$\left\| \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} \right\| = \frac{1}{\|\boldsymbol{a}\|} \|\boldsymbol{a}\| = 1.$$

This process of "dividing" a vector by its norm to obtain a unit vector in the same direction is called *normalizting* the vector.

Vector addition

- Take two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$

$$\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1, & b_2, & \cdots, & b_n \end{bmatrix}$$

- The sum of $\boldsymbol{a}$ and $\boldsymbol{b}$ is given by the vector

$$\begin{bmatrix} a_1 + b_1, & a_2 + b_2, & \cdots, a_n + b_n \end{bmatrix}$$

Example

If $\boldsymbol{a} = \begin{bmatrix} 2, & -3, & 5 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -6, & 4, & -2 \end{bmatrix}$, then

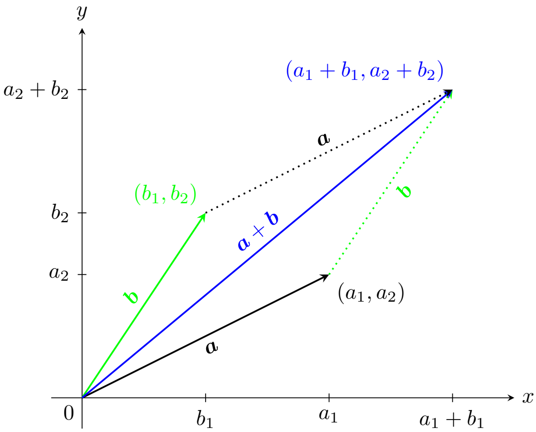
$$\boldsymbol{a} + \boldsymbol{b} = \begin{bmatrix} 2-6, & -3+4, & 5-2 \end{bmatrix} = \begin{bmatrix} -4, & 1, & 3 \end{bmatrix}$$

Note

Vectors cannot be added unless they have the same number of coordinates

Vector addition – geometric interpretation

- Draw a vector $\boldsymbol{a}$. Then draw a vector $\boldsymbol{b}$ whose initial point is the terminal point of $\boldsymbol{a}$.
- The sum of $\boldsymbol{a}$ and $\boldsymbol{b}$ is the vector whose initial point is the same as that of $\boldsymbol{a}$ and whose terminal point is the same as that of $\boldsymbol{b}$.
- The total movement $\boldsymbol{a}+\boldsymbol{b}$ is equivalent to first moving along $\boldsymbol{a}$ and then along $\boldsymbol{b}$



Subtraction of vectors

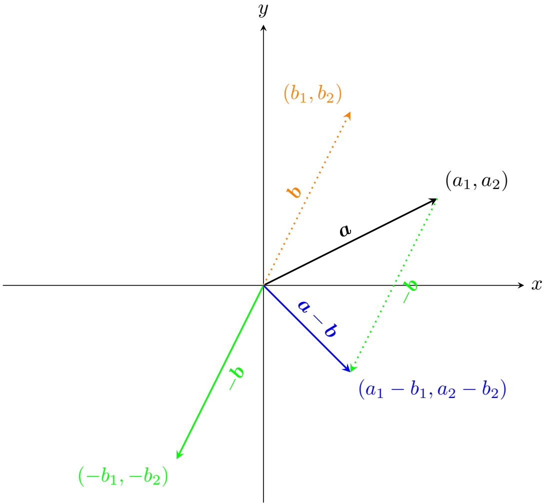
- Let $-\boldsymbol{b}$ denote the scalar product between -1 and $\boldsymbol{b}$
- Define

$$\boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a} + (-\boldsymbol{b})$$

Example

$\boldsymbol{a} = \begin{bmatrix} 2, & 1 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 1, & 2 \end{bmatrix}$

$$\begin{aligned} \boldsymbol{a} - \boldsymbol{b} &= \boldsymbol{a} + (-\boldsymbol{b}) \\ &= \begin{bmatrix} 2, & 1 \end{bmatrix} + \begin{bmatrix} -1, & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1, & -1 \end{bmatrix}. \end{aligned}$$



Fundamental properties of vector addition and scalar multiplication

Theorem

Take any $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$, *any* $\alpha, \beta \in \mathbb{R}$, *we have*

1. $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$	*Commutative law of addition*
2. $\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$	*Associative law of addition*
3. $\mathbf{0} + \boldsymbol{a} = \boldsymbol{a} + \mathbf{0} = \boldsymbol{a}$	*Existence of identity element for addition*
4. $\boldsymbol{a} + (-\boldsymbol{a}) = (-\boldsymbol{a}) + \boldsymbol{a} = \mathbf{0}$	*Existence of inverse elements for addition*
5. $\alpha(\boldsymbol{a} + \boldsymbol{b}) = \alpha\boldsymbol{a} + \alpha\boldsymbol{b}$	*Distributive laws of scalar multiplication*
6. $(\alpha + \beta)\boldsymbol{a} = \alpha\boldsymbol{a} + \beta\boldsymbol{a}$	*over vector addition*
7. $(\alpha\beta)\boldsymbol{a} = \alpha(\beta\boldsymbol{a})$	*Associativity of scalar multiplication*
8. $1\boldsymbol{a} = \boldsymbol{a}$	*Identity property for scalar multiplication*

- $\mathbf{0}$ is called an *identity element for vector addition* because $\mathbf{0}$ does not change the identity of any vector to which it is added
- $-\boldsymbol{a}$ is called the *Additive inverse of* $\boldsymbol{a}$ because it "cancels out $\boldsymbol{a}$" to produce the additive identity element (i.e. the zero vector)

Proof of property 6

$$
\begin{aligned}
(\alpha+\beta)\boldsymbol{a} &= (\alpha+\beta)\begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \\
&= \begin{bmatrix} (\alpha+\beta)a_1, & (\alpha+\beta)a_2, & \cdots, & (\alpha+\beta)a_n \end{bmatrix} \\
&\quad \text{definition of scalar multiplication} \\
&= \begin{bmatrix} \alpha a_1+\beta a_1, & \alpha a_2+\beta a_2, & \cdots, \alpha a_n+\beta a_n \end{bmatrix} \\
&\quad \text{coordinate-wise use of distributive law in } \mathbb{R} \\
&= \begin{bmatrix} \alpha a_1, & \alpha a_2, & \cdots, & \alpha a_n \end{bmatrix} + \begin{bmatrix} \beta a_1, & \beta a_2, & \cdots, & \beta a_n \end{bmatrix} \\
&\quad \text{definition of vector addition} \\
&= \alpha\begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} + \beta\begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \\
&\quad \text{definition of scalar multiplication} \\
&= \alpha\boldsymbol{a}+\beta\boldsymbol{a}
\end{aligned}
$$

Dot product

Definition

Let

$$\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1, & b_2, & \cdots, & b_n \end{bmatrix} \in \mathbb{R}^n$$

be two vectors. The *dot product (inner product)* of $\boldsymbol{a}$ and $\boldsymbol{b}$ is given by

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i.$$

Example

$\boldsymbol{a} = \begin{bmatrix} 2, & -4, & 3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1, & 5, & -2 \end{bmatrix}, \boldsymbol{a} \cdot \boldsymbol{b} = 2 \times 1 + (-4) \times 5 + 3 \times (-2) = -24.$

Note

- Dot product is not defined for vectors having different numbers of coordinates.
- Dot product involves two vectors and the result is a scalar, whereas scalar multiplication involves a scalar and a vector and the result is a vector.

Properties of dot product

Theorem

Take any $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, then

1. $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$	*Commutativity of dot product*
2. $\boldsymbol{a} \cdot \boldsymbol{a} = \|\boldsymbol{a}\| \geq 0$	*Relationship between dot product and norm*
3. $\boldsymbol{a} \cdot \boldsymbol{a} = 0$ *iff* $\boldsymbol{a} = 0$	
4. $\alpha(\boldsymbol{a} \cdot \boldsymbol{b}) = (\alpha \boldsymbol{a}) \cdot \boldsymbol{b} = \boldsymbol{a} \cdot (\alpha \boldsymbol{b})$	*Relationship between scalar multiplication and dot product*
5. $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$	*Distributive laws of dot product*
6. $(\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{c} = (\boldsymbol{a} \cdot \boldsymbol{c}) + (\boldsymbol{b} \cdot \boldsymbol{c})$	*over addition*

Proof

We provide the proof for a few properties.

2. $\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix}$

$$\boldsymbol{a} \cdot \boldsymbol{a} = a_1 a_1 + a_2 a_2 + \cdots + a_n a_n = a_1^2 + a_2^2 + \cdots a_n^2 = \|\boldsymbol{a}\|^2 \geq 0$$

Properties of dot product

5. $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$ Distributive laws of dot product over addition

Proof

5. $\boldsymbol{a} = \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1, & b_2, & \cdots, & b_n \end{bmatrix}, \boldsymbol{c} = \begin{bmatrix} c_1, & c_2, & \cdots, & c_n \end{bmatrix}$

$$\begin{aligned}
\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) &= \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \cdot \left(\begin{bmatrix} b_1, & b_2, & \cdots, & b_n \end{bmatrix} + \begin{bmatrix} c_1, & c_2, & \cdots, & c_n \end{bmatrix} \right) \\
&= \begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 + c_1, & b_2 + c_2, & \cdots, b_n + c_n \end{bmatrix} \\
&= a_1(b_1 + c_1) + a_2(b_2 + c_2) + \cdots + a_n(b_n + c_n) \\
&= a_1 b_1 + a_1 c_1 + a_2 b_2 + a_2 c_2, + \cdots + a_n b_n + a_n c_n \\
&= (a_1 b_1 + a_2 b_2 + \cdots + a_n b_n) + (a_1 c_1 + a_2 c_2 + \cdots + a_n c_n)
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c} &= \left(\begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \cdot \begin{bmatrix} b_1, & b_2, & \cdots, & b_n \end{bmatrix} \right) \\
&\quad + \left(\begin{bmatrix} a_1, & a_2, & \cdots, & a_n \end{bmatrix} \cdot \begin{bmatrix} c_1, & c_2, & \cdots, & c_n \end{bmatrix} \right) \\
&= (a_1 b_1 + a_2 b_2 + \cdots + a_n b_n) + (a_1 c_1 + a_2 c_2 + \cdots + a_n c_n)
\end{aligned}$$

Dot product – example

Example

$$\boldsymbol{a} = \begin{bmatrix} 1, & 2, & 3 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 4, & 5, & 6 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} -1, & -2, & -3 \end{bmatrix}$$

$$\begin{aligned}
\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) &= \begin{bmatrix} 1, & 2, & 3 \end{bmatrix} \cdot \left(\begin{bmatrix} 4-1, & 5-2, & 6-3 \end{bmatrix} \right) = \begin{bmatrix} 1, & 2, & 3 \end{bmatrix} \cdot \begin{bmatrix} 3, & 3, & 3 \end{bmatrix} \\
&= 1 \times 3 + 2 \times 3 + 3 \times 3 = 3 + 6 + 9 = 18, \\
\boldsymbol{a} \cdot \boldsymbol{b} &= \begin{bmatrix} 1, & 2, & 3 \end{bmatrix} \cdot \begin{bmatrix} 4, & 5, & 6 \end{bmatrix} = 1 \times 4 + 2 \times 5 + 3 \times 6 = 4 + 10 + 18 = 32, \\
\boldsymbol{a} \cdot \boldsymbol{c} &= \begin{bmatrix} 1, & 2, & 3 \end{bmatrix} \cdot \begin{bmatrix} -1, & -2, & -3 \end{bmatrix} = 1 \times (-1) + 2 \times (-2) + 3 \times (-3) \\
&= -1 - 4 - 9 = -14, \\
\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c} &= 32 - 14 = 18.
\end{aligned}$$

Orthogonal vectors

Definition

$\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ are *orthogonal* if $\boldsymbol{a} \cdot \boldsymbol{b} = 0$

Example

$\boldsymbol{a} = \begin{bmatrix} 2, & -5 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} -10, & -4 \end{bmatrix}$ are orthogonal in $\mathbb{R}^2$

$$\boldsymbol{a} \cdot \boldsymbol{b} = 2 \times (-10) + (-5) \times (-4) = -20 + 20 = 0.$$

Dot product of unit vectors

Recall

A vector of norm 1 is called a *unit vector*

Lemma

If $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ *are unit vectors, then*

$$-1 \leq \boldsymbol{a} \cdot \boldsymbol{b} \leq 1.$$

Proof

We make use the results from different part of the previous theorem.

Dot product of unit vectors

Lemma

If $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ *are unit vectors, then*

$$-1 \leq \boldsymbol{a} \cdot \boldsymbol{b} \leq 1.$$

Proof

$$\begin{aligned}
&(\boldsymbol{a}+\boldsymbol{b}) \cdot (\boldsymbol{a}+\boldsymbol{b}) = \|\boldsymbol{a}+\boldsymbol{b}\|^2 \geq 0 && \text{part 2} \\
&\implies \boldsymbol{a} \cdot \boldsymbol{a} + \boldsymbol{b} \cdot \boldsymbol{a} + \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b} \geq 0 && \text{parts 5, 6} \\
&\implies \|\boldsymbol{a}\|^2 + 2(\boldsymbol{a} \cdot \boldsymbol{b}) + \|\boldsymbol{b}\|^2 \geq 0 && \text{parts 1, 2} \\
&\implies 1 + 2\boldsymbol{a} \cdot \boldsymbol{b} + 1 \geq 0 && \boldsymbol{a}, \boldsymbol{b} \text{ are unit vectors} \\
&\implies \boldsymbol{a} \cdot \boldsymbol{b} \geq -1
\end{aligned}$$

A similar argument beginning with $(\boldsymbol{a}-\boldsymbol{b}) \cdot (\boldsymbol{a}-\boldsymbol{b}) = \|\boldsymbol{a}-\boldsymbol{b}\|^2 \geq 0$ shows $\boldsymbol{a} \cdot \boldsymbol{b} \leq 1$

Cauchy-Schwarz Inequality

Theorem

Take any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$*, we have*

$$|\boldsymbol{a} \cdot \boldsymbol{b}| \leq \|\boldsymbol{a}\| \|\boldsymbol{b}\|.$$

Proof

- If $\boldsymbol{a} = \mathbf{0}$ or $\boldsymbol{b} = \mathbf{0}$, the theorem holds
- Otherwise, the theorem is equivalent to

$$-1 \leq \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\| \|\boldsymbol{b}\|} \leq 1$$

We have discussed that

$$\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} \quad \frac{\boldsymbol{b}}{\|\boldsymbol{b}\|}$$

are unit vectors. The result follows from the previous lemma.

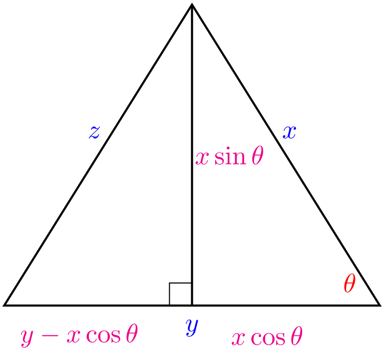
Cauchy-Schwarz Inequality – example

Example

$\boldsymbol{a} = \begin{bmatrix} -1, & 4, & 2, & 0, & -3 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 2, & 1, & -4, & -1, & 0 \end{bmatrix}.$

$$
\begin{aligned}
\boldsymbol{a} \cdot \boldsymbol{b} &= -2 + 4 - 8 + 0 + 0 = -6 \\
\|\boldsymbol{a}\| &= \sqrt{1 + 16 + 4 + 0 + 9} = \sqrt{30} \\
\|\boldsymbol{b}\| &= \sqrt{4 + 1 + 16 + 1 + 0} = \sqrt{22} \\
\|\boldsymbol{a}\|\|\boldsymbol{b}\| &= \sqrt{30 \times 22} = 2\sqrt{165} \approx 25.7 \\
|\boldsymbol{a} \cdot \boldsymbol{b}| = 6 &\leq \|\boldsymbol{a}\|\|\boldsymbol{b}\|
\end{aligned}
$$

Law of Cosines

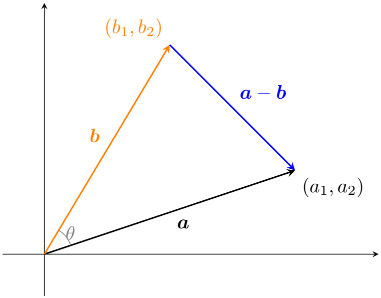


$$\begin{aligned}
(y - x\cos\theta)^2 + (x\sin\theta)^2 &= z^2 \\
y^2 + x^2\cos^2\theta - 2yx\cos\theta + x^2\sin^2\theta &= z^2 \\
x^2 + y^2 - 2yx\cos\theta &= z^2
\end{aligned}$$

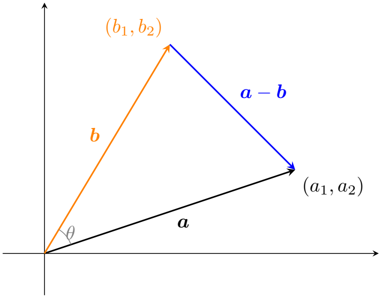
The angle between two vectors

- There are two angles formed by the two vectors, but we always choose the angle θ between two vectors to be the one measuring between 0 and π radians, inclusive.
- By the Law of Cosines

$$\|\boldsymbol{a}-\boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2-2\|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos\theta$$



Angle between two vectors



$$\|\boldsymbol{a} - \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 - 2\|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos\theta$$

$$\|\boldsymbol{a} - \boldsymbol{b}\|^2 = (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b}) = \boldsymbol{a} \cdot \boldsymbol{a} - 2\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b} = \|\boldsymbol{a}\|^2 - 2\boldsymbol{a} \cdot \boldsymbol{b} + \|\boldsymbol{b}\|^2$$
$$\implies \|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos\theta = \boldsymbol{a} \cdot \boldsymbol{b}$$

$$\cos\theta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$$

Angle between two vectors – example

Example

$\boldsymbol{a} = \begin{bmatrix} 6, & -4 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -2, & 3 \end{bmatrix}$

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} = \frac{6\times(-2)+(-4)\times 3}{\sqrt{36+9}\sqrt{4+9}} = -\frac{24}{\sqrt{52}\sqrt{13}} = -\frac{12}{13} \approx -0.9231,$$

which gives $\theta \approx 2.74$ radians (using calculator).

Angle between two vectors

- For higher dimensions we are outside the geometry of everyday experience
- We give the following definition

Definition

For any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ with $n \geq 2$, the *angle between $\boldsymbol{a}$ and $\boldsymbol{b}$* is the unique angle θ such that $0 \leq \theta \leq \pi$ and

$$\cos\theta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}.$$

Note that according to Cauchy-Schwarz Inequality,

$$-1 \leq \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} \leq 1.$$

Thus this value equals $\cos\theta$ for a unique θ from 0 to π radians.

Angle between two vectors

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}.$$

By the properties of the cosine function, we have

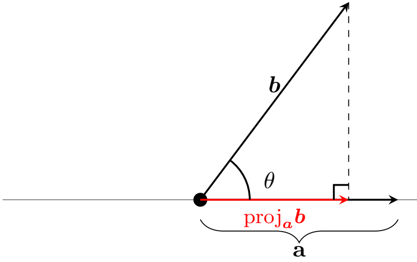
$$\begin{aligned}
\boldsymbol{a}\cdot\boldsymbol{b} > 0 \quad &\iff \quad 0 \le \theta < \frac{\pi}{2} \\
\boldsymbol{a}\cdot\boldsymbol{b} = 0 \quad &\iff \quad \theta = \frac{\pi}{2} \\
\boldsymbol{a}\cdot\boldsymbol{b} < 0 \quad &\iff \quad \frac{\pi}{2} < \theta \le \pi
\end{aligned}$$

Note

By definition of orthogonal vectors, two *nonzero* vectors are orthogonal if and only if they are perpendicular to each other (i.e. $\theta = \frac{\pi}{2}$)

Projection vectors

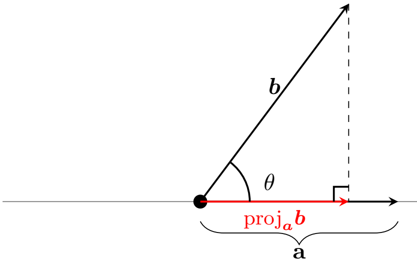
- The projection of one vector onto another is useful in physics, engineering, computer graphics, and statistics.
- $\boldsymbol{a}, \boldsymbol{b}$ both in $\mathbb{R}^2$ or $\mathbb{R}^3$, drawn at the same initial point
- Let θ represent the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$
- Drop a perpendicular line segment from the terminal point of $\boldsymbol{b}$ to the straight line containing the vector $\boldsymbol{a}$
- The project of $\boldsymbol{b}$ onto $\boldsymbol{a}$, denoted $\text{proj}_{\boldsymbol{a}}\boldsymbol{b}$, is the vector from the initial point of $\boldsymbol{a}$ to the point where the dropped perpendicular meets the straight line



Projection vectors

- Using trigonometry, for $0 \leq \theta \leq \frac{\pi}{2}$, $\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}$ is in the direction of the unit vector $\boldsymbol{a}/\|\boldsymbol{a}\|$, and

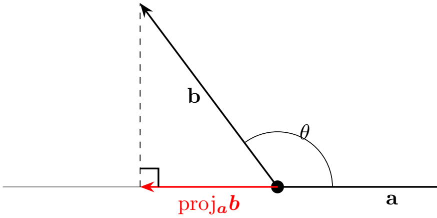
$$\|\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}\| = \|\boldsymbol{b}\| \cos\theta$$



Projection vectors

- Using trigonometry, when $\frac{\pi}{2} < \theta \leq \pi$, $\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}$ is in the direction of the unit vector $-\boldsymbol{a}/\|\boldsymbol{a}\|$, and

$$\|\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}\| = -\|\boldsymbol{b}\| \cos\theta$$



Projection vectors

- When $0 \leq \theta \leq \frac{\pi}{2}$, $\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}$ is in the direction of the unit vector $\boldsymbol{a}/\|\boldsymbol{a}\|$, and

$$\|\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}\| = \|\boldsymbol{b}\|\cos\theta$$

- When $\frac{\pi}{2} < \theta \leq \pi$, $\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}$ is in the direction of the unit vector $-\boldsymbol{a}/\|\boldsymbol{a}\|$, and

$$\|\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b}\| = -\|\boldsymbol{b}\|\cos\theta$$

- We know that

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}.$$

- We have

$$\mathrm{proj}_{\boldsymbol{a}}\boldsymbol{b} = \|\boldsymbol{b}\|\cos\theta\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} = \|\boldsymbol{b}\|\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|^2}\boldsymbol{a}.$$

Projection vectors – example

Example

$$\boldsymbol{a} = \begin{bmatrix} 4, & 0, & -3 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 3, & 1, & -7 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\boldsymbol{a}}\boldsymbol{b} &= \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|^2}\boldsymbol{a} = \frac{4 \times 3 + 0 \times 1 + (-3) \times (-7)}{4^2 + 0^2 + (-3)^2}\boldsymbol{a} \\ &= \frac{33}{25}\begin{bmatrix} 4, & 0, & -3 \end{bmatrix} = \begin{bmatrix} \frac{132}{25}, & 0, & -\frac{99}{25} \end{bmatrix}. \end{aligned}$$

Vectors and matrices

- Definitions
- Vectors
- Matrices

Matrices

- $\mathbb{R}$: the set of all real numbers

Definition

A *matrix with coefficients in* $\mathbb{R}$ is a rectangular array where each *entry* is an element of $\mathbb{R}$.

Matrix A is said to have m *rows*, n *columns* and is of size $m \times n$.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

We also write $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$. When the size is clear from the context, we write $A = (a_{ij})$.

Matrices – examples

Example

- $A = \begin{bmatrix} 1 & 7.5 & 6 \\ 2 & 3 & 4 \end{bmatrix}$ has 2 rows, 3 columns and is of size 2×3. $a_{11} = 1,\ a_{22} = 3$.
- $B = \begin{bmatrix} 4 & -2 \\ 1 & 7 \\ -5 & 3 \end{bmatrix}$ is of size 3×2. $b_{12} = -2,\ b_{31} = -5$
- $C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ is of size 3×3
- $D = \begin{bmatrix} 7 \\ 1 \\ -2 \end{bmatrix}$ is a 3×1 matrix
- $E = \begin{bmatrix} 4, & -3, & 0 \end{bmatrix}$ is a 1×3 matrix
- $F = \begin{bmatrix} 4 \end{bmatrix}$ is a 1×1 matrix

Matrices

- The size of a matrix is always specified by stating the number of rows first. For example, a 3×4 matrix always has three rows and four columns, never four rows and three columns
- An $m \times n$ matrix can be thought of either as a collection of m row vectors, each having n coordinates, or as a collection of n column vectors, each having m coordinates.

Definition

Let $\mathcal{M}_{m \times n}$ denote the set of all $m \times n$ matrices

Rows and columns of matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

- a_{ij} denotes the entry in the ith row and jth column.
- The ith row of A is
$$\begin{bmatrix} a_{i1}, & a_{i2}, & \dots & a_{in} \end{bmatrix}.$$
- The jth column of A is
$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Rows and columns of matrices – Examples

Example

$$A = \begin{bmatrix} 1 & 7.5 & 6 \\ 2 & 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 3.5 & 7 \end{bmatrix}$$

- The 1st row of A is
$$\begin{bmatrix} 1, & 7.5, & 6 \end{bmatrix}.$$
- The 2nd column of B is
$$\begin{bmatrix} 0 \\ 7 \end{bmatrix}.$$
- $A \in \mathcal{M}_{2\times 3}$, $B \in \mathcal{M}_{2\times 2}$

Main diagonal of a matrix

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

The *main diagonal* entries of A are $a_{11}, a_{22}, a_{33}, \dots$, those that lie on a diagonal line drawn down to the right, beginning from the upper-left corner of the matrix.

Example

$$A = \begin{bmatrix} 1 & 7.5 & 6 \\ 2 & 3 & 4 \end{bmatrix}$$

has main diagonal entries $1, 3$.

Equal matrices

- Two matrices A and B are *equal* iff they have the same size and all corresponding entries are equal
- $A, B \in \mathcal{M}_{m\times n}$, then
$$A = B \iff a_{ij} = b_{ij},\ \forall i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n$$

Zero matrices

A *zero matrix* is any matrix with all entries equal to 0, denoted O.

Example

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Square matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

An $n \times n$ matrix is called a *square matrix* (i.e. a matrix with the same number of rows and columns).

Example

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

Diagonal matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

A square matrix of size $n \times n$ is a *diagonal matrix* if $a_{ij} = 0$ for $i \neq j$

Example

$$A = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} -4 & 0 \\ 0 & 5 \end{bmatrix}$$

The following matrix is not a diagonal matrix

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

Identity matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

An $n \times n$ *identity matrix*, normally denoted I_n, is a diagonal matrix whose diagonal entries are 1 and all other entries are 0, i.e. $a_{ii} = 1$ for $i = 1, 2 \dots, n$ and $a_{ij} = 0$ for $i \neq j$.

Example

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Upper triangular matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

An *upper triangular matrix* is a square matrix with all entries below the main diagonal equal to zero.
In other words, $A \in \mathcal{M}_{n \times n}$ is an *upper triangular matrix* if $a_{ij} = 0$ for $i > j$.

Example

$$P = \begin{bmatrix} 6 & 9 & 11 \\ 0 & -2 & 3 \\ 0 & 0 & 5 \end{bmatrix}, \quad U = \begin{bmatrix} 7 & -2 & 2 & 0 \\ 0 & -4 & 9 & 5 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Lower triangular matrices

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

An *lower triangular matrix* is a square matrix with all entries above the main diagonal equal to zero.
In other words, $A \in \mathcal{M}_{n\times n}$ is an *lower triangular matrix* if $a_{ij} = 0$ for $i < j$.

Example

$$L = \begin{bmatrix} 3 & 0 & 0 \\ 9 & -2 & 0 \\ 14 & -6 & 1 \end{bmatrix}$$

Transpose of a matrix

The *transpose* of $A \in \mathcal{M}_{n\times n}$, denoted $A^\top$, is the $n \times m$ matrix obtained by interchanging the rows and columns of A.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ & \vdots & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad A^\top = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{m2} \\ & \vdots & \\ a_{1n} & \dots & a_{mn} \end{bmatrix}.$$

Example

$$A = \begin{bmatrix} 1 & 7.5 & 6 \\ 2 & 3 & 4 \end{bmatrix}, \quad A^\top = \begin{bmatrix} 1 & 2 \\ 7.5 & 3 \\ 6 & 4 \end{bmatrix}.$$

A is of size 2×3, $A^\top$ is of size 3×2.

Note

$(A^\top)^\top = A$

Symmetric matrices

Definition

$A \in \mathcal{M}_{m\times n}$ is *symmetric* if $A = A^\top$. It is *skew-symmetric* if $A = -A^\top$

Since $A^\top \in \mathcal{M}_{n\times m}$, it is easy to see that any symmetric or skew-symmetric matrix is a square matrix:

Example

$$A = \begin{bmatrix} 2 & 6 & 4 \\ 6 & -1 & 0 \\ 4 & 0 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -1 & 3 & 6 \\ 1 & 0 & 2 & -5 \\ -3 & -2 & 0 & 4 \\ -6 & 5 & -4 & 0 \end{bmatrix}$$

A is symmetric and B is skew-symmetric

Matrix addition

Definition

Take $A = (a_{ij}), B = (b_{ij}) \in \mathcal{M}_{m\times n}$, the *sum* of A and B, $A + B$, is the $m \times n$ matrix whose $(i, j)-$entry is equal to $a_{ij} + b_{ij}$

Example

$$\begin{bmatrix} 6 & -3 & 2 \\ -7 & 0 & 4 \end{bmatrix} + \begin{bmatrix} 5 & -6 & -3 \\ -4 & -2 & -4 \end{bmatrix} = \begin{bmatrix} 11 & -9 & -1 \\ -11 & -2 & 0 \end{bmatrix}$$

Matrix addition – example

Example

Notice that the definition does not allow addition of matrices with different sizes.

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

We cannot add those two matrices together. But

$$A + B^\top = \begin{bmatrix} 1+4 \\ 2+5 \\ 3+6 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$$

$$A^\top + B = \begin{bmatrix} 1, & 2, & 3 \end{bmatrix} + \begin{bmatrix} 4, & 5, & 6 \end{bmatrix} = \begin{bmatrix} 1+4, & 2+5, & 3+6 \end{bmatrix} = \begin{bmatrix} 5, & 7, & 9 \end{bmatrix}.$$

Multiply a matrix by a scalar

Definition

Let $A = (a_{ij}) \in \mathcal{M}_{m\times n}$ and $\alpha \in \mathbb{R}$. The scalar multiple of A by α is $\alpha A \in \mathcal{M}_{m\times n}$ whose $(i,j)-$entry is equal to αa_{ij}.

Example

- $\alpha = -2$
$$A = \begin{bmatrix} 4 & -1 & 6 & 7 \\ 2 & 4 & 9 & -5 \end{bmatrix}, -2A = \begin{bmatrix} -8 & 2 & 12 & -14 \\ -4 & -8 & -18 & 10 \end{bmatrix}$$
- $0A = O$ for any matrix A

Subtraction of matrices

- Let $-A$ denote the matrix $-1A$, the scalar multiple of A by -1
- We define subtraction of matrices as

$$A - B = A + (-B)$$

Example

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

$$A - B = A + (-B) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1-5, & 2-6 \\ 3-7 & 4-8 \end{bmatrix} = \begin{bmatrix} -4, & -4 \\ -4, & -4 \end{bmatrix}$$

Fundamental properties of addition and scalar multiplication

Theorem

For any matrices $A, B, C \in \mathcal{M}_{m\times n}$ *and any scalars* $\alpha, \beta \in \mathbb{R}$, *we have*

1. $A + B = B + A$	*Commutative law of addition*
2. $A+(B+C) = (A+B)+C$	*Associative law of addition*
3. $O + A = A + O = A$	*Existence of identity element for addition*
4. $A+(-A) = (-A)+A = O$	*Existence of inverse elements for addition*
5. $\alpha(A + B) = \alpha A + \alpha B$	*Distributive laws of scalar multiplication*
6. $(\alpha + \beta)\boldsymbol{a} = \alpha A + \beta A$	*over vector addition*
7. $(\alpha\beta)A = \alpha(\beta A)$	*Associativity of scalar multiplication*
8. $1A = A$	*Identity property for scalar multiplication*

To prove each property, calculate corresponding entries on both sides and show they agree by applying an appropriate law of real numbers.

Fundamental properties of addition and scalar multiplication

Theorem

1. $A + B = B + A$ *Commutative law of addition*

To prove each property, calculate corresponding entries on both sides and show they agree by applying an appropriate law of real numbers.

Proof

Part 1. Suppose $A = (a_{ij})$, $B = (b_{ij})$, then the $(i,j)-$entry of $A + B$ is $a_{ij} + b_{ij}$. And the $(i,j)-$entry of $B + A$ is $b_{ij} + a_{ij}$. By the commutativity property of addition for real numbers, we have

$$a_{ij} + b_{ij} = b_{ij} + a_{ij},$$

which implies $A + B = B + A$.

Transpose of matrices

Theorem

For any $A, B \in \mathcal{M}_{m\times n},\ \alpha \in \mathbb{R}$, *we have*

1. $(A+B)^\top = A^\top + B^\top$
2. $(A-B)^\top = A^\top - B^\top$
3. $(\alpha A)^\top = \alpha A^\top$

Proof

Proof of part 2. Suppose $A = (a_{ij}), B = (b_{ij})$ First we note that $(A+B)^\top, A^\top + B^\top \in \mathcal{M}_{n\times m}$. Next, we should that each of their $(i,j)-$entries are equal for $i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m$:

$$(i,j) - \text{entry of } (A+B)^\top = (j,i) - \text{entry of } A+B = a_{ji} + b_{ji}$$

$$(i,j) - \text{entry of } A^\top + B^\top = (j,i) - \text{entry of } A + (j,i) - \text{entry of } A = a_{ji} + b_{ji}$$

Matrix multiplication

Definition

The product of $A \in \mathcal{M}_{m\times n}$ and $B \in \mathcal{M}_{n\times r}$ is the matrix $C = AB \in \mathcal{M}_{m\times r}$ whose $(i,j)-$entry is the dot product of the ith row of A with the jth column of B

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2r} \\ b_{31} & b_{32} & \cdots & b_{3j} & \cdots & b_{3r} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{nr} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1r} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2r} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{ir} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mj} & \cdots & c_{mr} \end{bmatrix},$$

where

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

Matrix multiplication

- Two matrices A, B can be multiplied (in that order) only if the number of columns of A is equal to the number of rows of B
- This ensures that each row of A contains the same number of entries as each column of B. Thus it is possible to perform the dot products needed to calculate C

Note

The dot product of two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ is the same as the product of $\boldsymbol{a} \in \mathcal{M}_{1\times n}$ and $\boldsymbol{b}^\top \in \mathcal{M}_{n\times 1}$

Example

$$[1, \quad 2, \quad 3] \cdot [4, \quad 5, \quad -6] = [1, \quad 2, \quad 3] \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = 1\times4+2\times5+3\times(-6) = 1+10-18 = -7.$$

Matrix multiplication – example

Example

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

$A \in \mathcal{M}_{2\times 3}$, $B \in \mathcal{M}_{3\times 4}$, A and B can be multiplied and the product $C \in \mathcal{M}_{2\times 4}$.
To calculate c_{11}, we compute the dot product of the 1st row of A and the 1st column of B:

$$c_{11} = \begin{bmatrix} 5, & -1, & 4 \end{bmatrix} \cdot \begin{bmatrix} 9, & 7, & -1 \end{bmatrix} = 5 \times 9 + (-1) \times 7 + 4 \times (-2) = 45 - 7 - 8 = 30.$$

$$c_{23} = \begin{bmatrix} -3, & 6, & 0 \end{bmatrix} \cdot \begin{bmatrix} -8, & -1, & 3 \end{bmatrix} = (-3) \times (-8) + 6 \times (-1) + 0 \times 3 = 24 - 6 = 18$$

The other entries are computed similarly, we have

$$C = AB = \begin{bmatrix} 30 & 34 & -27 & -6 \\ 15 & 24 & 18 & -6 \end{bmatrix}$$

Identity matrix

Let $A \in \mathcal{M}_{m\times n}$ be any matrix, $I_n \in \mathcal{M}_{n\times n}$ and $I_m \in \mathcal{M}_{m\times m}$ be identity matrices. We have

$$AI_n = I_m A = A$$

Proof

Suppose $I_n = (c_{ij})$, then for $i = 1, 2, \ldots, n$.

$$c_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Let $B = AI_n = (b_{ij})$, then $B \in \mathcal{M}_{m\times n}$. And for $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$

$$b_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj} = a_{ij} \times 1 = a_{ij}.$$

Transpose of matrix product

Theorem

Let $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{n\times r}$, *then*

$$(AB)^\top = B^\top A^\top$$

Proof

First we note that both matrices are of size $r \times m$.

$$\begin{aligned}
(i,j) - \text{entry of } (AB)^\top &= (j,i) - \text{entry of } AB \\
&= [j\text{th row of } A] \cdot [i\text{th column of } B] \\
(i,j) - \text{entry of } B^\top A^\top &= [i\text{th row of } B^\top] \cdot [j\text{th column of } A^\top] \\
&= [i\text{th column of } B] \cdot [j\text{th row of } A] \\
&= [j\text{th row of } A] \cdot [i\text{th column of } B]
\end{aligned}$$

Matrix multiplication – example

Example

$$D = \begin{bmatrix} -2 & 1 \\ 0 & 5 \\ 4 & -3 \end{bmatrix}, \; E = \begin{bmatrix} 1 & -6 \\ 0 & 2 \end{bmatrix}, \; F = \begin{bmatrix} -4, & 2, & 1 \end{bmatrix}, \; G = \begin{bmatrix} 7 \\ -1 \\ 5 \end{bmatrix}, \; H = \begin{bmatrix} 5 & 0 \\ 1 & -3 \end{bmatrix}$$

The order in which multiplication is performed is important. Given two matrices

- Neither product may be defined (e.g. DG, GD).
- One product may be defined but not the other (e.g. ED is not defined)
- Both products may be defined, but the resulting sizes may differ (e.g. F, G)
- Both products may be defined, and the resulting sizes may agree, but the entries may differ (e.g. E and H)

$$DE = \begin{bmatrix} -2 & 14 \\ 0 & 10 \\ 4 & -30 \end{bmatrix}, \; GF = \begin{bmatrix} -28 & 14 & 7 \\ 4 & -2 & -1 \\ -20 & 10 & 5 \end{bmatrix}, \; FG = [-25], \; EH = \begin{bmatrix} -1 & 18 \\ 2 & -6 \end{bmatrix}, \; HE = \begin{bmatrix} 5 & -30 \\ 1 & -12 \end{bmatrix}$$

Fundamental properties of matrix multiplication

Theorem

For any matrices A, B, C where the following operations are well-defined, and for any scalars $\alpha \in \mathbb{R}$, we have

1. $A(BC) = (AB)C$ — *Associative law of multiplication*
2. $A(B + C) = AB + AC$ — *Distributive law of matrix multiplication*
3. $(A + B)C = AC + BC$ — *over addition*
4. $\alpha(AB) = A(\alpha B)$ — *Associative law of scalar and matrix multiplication*

Proofs for 2, 3, 4 are easy - compute both sides, show they are equal. Proof for 1 can be found in the book Appendix A.

Distributive law of matrix multiplication – example

Example

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 0 \\ 4 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$$

$$B + C = \begin{bmatrix} -2+3 & 0+(-2) \\ 4+1 & -1+0 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 5 & -1 \end{bmatrix}$$

$$A(B+C) = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 \\ 5 & -1 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 + 2\cdot 5 & 1\cdot(-2) + 2\cdot(-1) \\ -1\cdot 1 + 3\cdot 5 & -1\cdot(-2) + 3\cdot(-1) \end{bmatrix} = \begin{bmatrix} 11 & -4 \\ 14 & -1 \end{bmatrix}$$

$$AB = \begin{bmatrix} -2+8 & 0+(-2) \\ 2+12 & 6-3 \end{bmatrix} = \begin{bmatrix} 6 & -2 \\ 14 & -3 \end{bmatrix}, \quad AC = \begin{bmatrix} 5 & -2 \\ 0 & 2 \end{bmatrix}$$

We have

$$AB + AC = \begin{bmatrix} 11 & -4 \\ 14 & -1 \end{bmatrix} = A(B+C)$$

Remark

Example

Continue from the previous example

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 0 \\ 4 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$$

$$B + C = \begin{bmatrix} -2+3 & 0+(-2) \\ 4+1 & -1+0 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 5 & -1 \end{bmatrix}$$

$$A(B+C) = AB + AC = \begin{bmatrix} 11 & -4 \\ 14 & -1 \end{bmatrix}$$

$$(B+C)A = \begin{bmatrix} 1 & -2 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1+2 & 2-6 \\ 5+1 & 10-3 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ 6 & 7 \end{bmatrix}$$

Cancellation laws do not hold

We note that if $AB = AC$ and $A \neq O$, it does not necessarily follow that $B = C$. For example

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 5 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 1 \\ -3 & 0 \end{bmatrix}$$

$$AB = \begin{bmatrix} 3 & 2 \\ 9 & 6 \end{bmatrix}, \quad AC = \begin{bmatrix} 3 & 2 \\ 9 & 6 \end{bmatrix}$$

Similarly, if $BA = CA$, $A \neq O$, it does not necessarily follow that $B = C$

Linear combination of matrices

Definition

Given $A_1, A_2, \ldots, A_r \in \mathcal{M}_{m\times n}$, $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$, an expression of the form

$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_r A_r$$

is called a *linear combination* of $A_1, A_2, \ldots, A_r$ with coefficients $\alpha_1, \alpha_2, \ldots, \alpha_r$.

Linear combination of matrices

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then

$$A\boldsymbol{x} = \begin{bmatrix} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ \vdots & & \vdots & & \ddots & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

Linear combination of matrices

Observation

Given $A \in \mathcal{M}_{m\times n}$, $\boldsymbol{x} \in \mathbb{R}^n$ (or equivalently $\boldsymbol{x} \in \mathcal{M}_{n\times 1}$), the product $A\boldsymbol{x}$ can be expressed as a linear combination of the columns of A in which the coefficients are the entries of $\boldsymbol{x}$.

Example

The matrix product

$$\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

can be written as

$$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ -3 \end{bmatrix}$$

Linear combination of matrices and matrix product

Consider

$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

It follows from the previous theorem that the jth column of AB can be expressed as a linear combination of the columns of A in which the coefficients in the linear combination are the entries from the jth column of B

$$\begin{aligned}
\begin{bmatrix} 12 \\ 8 \end{bmatrix} &= 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 27 \\ -4 \end{bmatrix} &= \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 7\begin{bmatrix} 4 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 30 \\ 26 \end{bmatrix} &= 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 5\begin{bmatrix} 4 \\ 0 \end{bmatrix} \\
\begin{bmatrix} 13 \\ 12 \end{bmatrix} &= 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix}
\end{aligned}$$

Linear combination of matrices and matrix product

$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

Similarly, the ith row of AB can be considered as linear combinations of the rows of B with coefficients given by the entries from the ith row of A

$$\begin{aligned} \begin{bmatrix} 12, & 27, & 30, & 13 \end{bmatrix} &= 1\begin{bmatrix} 4, & 1, & 4, & 3 \end{bmatrix} + 2\begin{bmatrix} 0, & -1, & 3, & 1 \end{bmatrix} + 4\begin{bmatrix} 2, & 7, & 5, & 2 \end{bmatrix} \\ \begin{bmatrix} 8, & -4, & 26, & 12 \end{bmatrix} &= 2\begin{bmatrix} 4, & 1, & 4, & 3 \end{bmatrix} + 6\begin{bmatrix} 0, & -1, & 3, & 1 \end{bmatrix} + 0\begin{bmatrix} 2, & 7, & 5, & 2 \end{bmatrix} \end{aligned}$$

The zero matrix

- $AO = OA = O$
- If $AB = O$, it is not necessarily true that $A = O$ or $B = O$. For example

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}, \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Powers of a square matrix

- Square matrices are the only matrices that can be multiplied by themselves
- $A \in \mathcal{M}_{m \times n}$, AA can be computed iff $m = n$

Definition

For $A \in \mathcal{M}_{n \times n}$, the *(nonnegative) powers* of A are given by

$$A^0 = I_n, \quad A^1 = A, \quad A^k = A^{k-1}A \text{ for } k \geq 2.$$

Example

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

$$A^2 = AA = \begin{bmatrix} 0 & 5 \\ -20 & 5 \end{bmatrix}, \quad A^3 = A^2A = \begin{bmatrix} -20 & 15 \\ -60 & -5 \end{bmatrix}.$$

Special cases when $AB = BA$

- Take $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{p\times r}$
- For AB to be defined, we need $p = n$
- For BA to be defined, we need $r = m$
- Then
$$AB \in \mathcal{M}_{m\times n}, \quad BA \in \mathcal{M}_{n\times m}$$
- For $AB = BA$, we need $m = n$, i.e.
$$A, B \in \mathcal{M}_{n\times n}$$
are square matrices
- In case $AB = BA$, we say A and B *commute* or A *commutes* with B.

Special cases when $AB = BA$

Take $A, B \in \mathcal{M}_{n\times n}$

- $n = 1$, $AB = BA$
- If $B = A$, then $AB = BA = A^2$
- If $A = O$ or $B = O$, then $AB = BA = O$
- If $\exists \alpha \in \mathbb{R}$ s.t. $A = \alpha I_n$, then according to the associative law of scalar and matrix multiplication, and the property of the identity matrix

$$BA = B(\alpha I_n) = \alpha(BI_n) = \alpha(I_n B) = (\alpha I_n)B = AB$$

- Similarly, if $\exists \alpha \in \mathbb{R}$ s.t. $B = \alpha I_n$, we have $AB = BA$

Example

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad AB = BA = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

Trace of a square matrix

Definition

Let $A = (a_{ij}) \in \mathcal{M}_{n \times n}$ be a square matrix. The *trace* of A, denoted $\operatorname{tr}(A)$ is given by the sum of the main diagonal entries of A, i.e.

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 5 & -3 & 4 \\ 1 & 1 & -2 \end{bmatrix}, \quad \operatorname{tr}(A) = 1 + (-3) + (-2) = -4.$$