Angles, Vector Spaces, Matrices

Hayk Aprikyan, Hayk Tarkhanyan

Angle between vectors

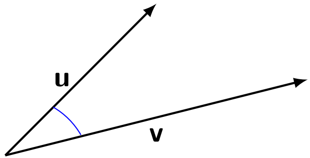
Geometrically, a vector is an arrow in space, that is, it has both a length and direction.

Angle between vectors

Geometrically, a vector is an arrow in space, that is, it has both a length and direction. How do we describe a direction in mathematics?

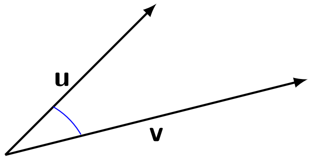
Angle between vectors

Geometrically, a vector is an arrow in space, that is, it has both a length and direction. How do we describe a direction in mathematics? By angles!



Angle between vectors

Geometrically, a vector is an arrow in space, that is, it has both a length and direction. How do we describe a direction in mathematics? By angles!



Remember the formula from high school geometry:
$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \alpha$, where α is the angle between $\mathbf{a}$ and $\mathbf{b}$.

Angle between vectors

Definition

The angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ is the angle $0 \leq \theta \leq \pi$ for which:

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$$

Angle between vectors

Definition

The angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ is the angle $0 \leq \theta \leq \pi$ for which:

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$$

Example

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Find the angle θ between $\mathbf{u}$ and $\mathbf{v}$.

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|} = \frac{(3 \cdot 7) + (4 \cdot 1)}{\sqrt{3^2 + 4^2} \cdot \sqrt{7^2 + 1^2}} = \frac{25}{\sqrt{25} \cdot \sqrt{50}} = \frac{1}{\sqrt{2}}$$

$$\frac{1}{\sqrt{2}} = \cos\frac{\pi}{4} \quad \Rightarrow \quad \theta = \arccos\frac{1}{\sqrt{2}} = \frac{\pi}{4} = 45^\circ$$

Angle between vectors

Corollary

For any vectors $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^n$,

$$\mathbf{v} \cdot \mathbf{u} = \|\mathbf{v}\| \cdot \|\mathbf{u}\| \cdot \cos\theta,$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{u}$.

Angle between vectors

Corollary

For any vectors $\mathbf{u}$, $\mathbf{v} \in \mathbb{R}^n$,

$$\mathbf{v} \cdot \mathbf{u} = \|\mathbf{v}\| \cdot \|\mathbf{u}\| \cdot \cos\theta,$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{u}$.

In particular, this also means that:

$$|\mathbf{v} \cdot \mathbf{u}| \leq \|\mathbf{v}\| \cdot \|\mathbf{u}\|$$

Angle between vectors

Corollary

For any vectors **u**, **v** $\in \mathbb{R}^n$,

$$\mathbf{v} \cdot \mathbf{u} = \|\mathbf{v}\| \cdot \|\mathbf{u}\| \cdot \cos\theta,$$

where θ is the angle between **v** and **u**.

In particular, this also means that:

$$|\mathbf{v} \cdot \mathbf{u}| \leq \|\mathbf{v}\| \cdot \|\mathbf{u}\|$$

i.e. for any numbers $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ we have

$$|a_1 b_1 + a_2 b_2 + \cdots + a_n b_n| \leq \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \cdot \sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}$$

a fact known as the **Cauchy-Schwarz inequality**.

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,
- Two vectors form an angle of 180° if they point in the opposite directions,

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,
- Two vectors form an angle of 180° if they point in the opposite directions,

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,
- Two vectors form an angle of 180° if they point in the opposite directions,

and most importantly:

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,
- Two vectors form an angle of 180° if they point in the opposite directions,

and most importantly:

- Two vectors are **perpendicular** (or **orthogonal**) to each other (i.e. form a 90° angle), if and only if their dot product equals 0.

Angle between vectors

The way we defined the angle between two vectors, it is clear that:

- Two vectors form an angle of 0° if they point in the same direction,
- Two vectors form an angle of 180° if they point in the opposite directions,

and most importantly:

- Two vectors are **perpendicular** (or **orthogonal**) to each other (i.e. form a 90° angle), if and only if their dot product equals 0.

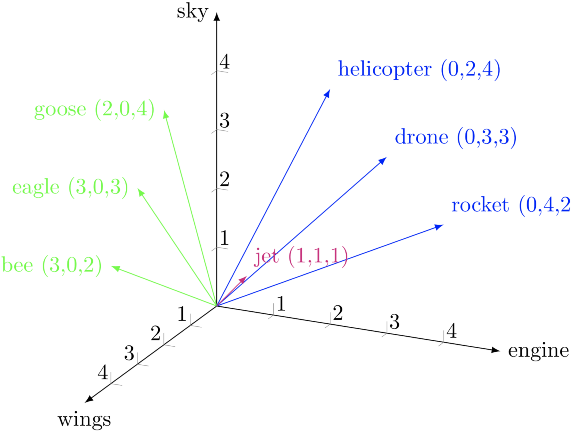
This is important because being orthogonal is the mathematical way of saying "pointing in completely different directions".

Cosine similarity (optional)

Suppose that we somehow represent a couple of words as vectors in $\mathbb{R}^3$.

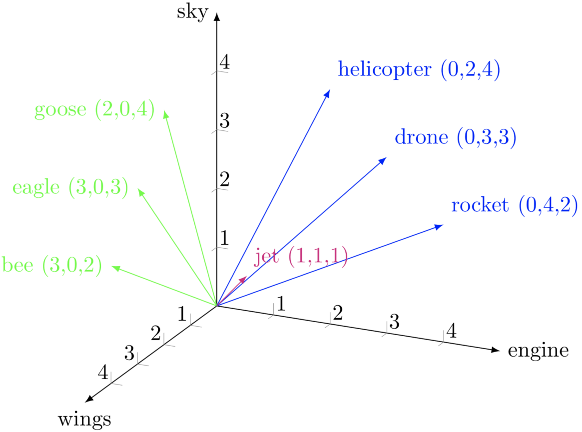
Cosine similarity (optional)

Suppose that we somehow represent a couple of words as vectors in $\mathbb{R}^3$.



Cosine similarity (optional)

Suppose that we somehow represent a couple of words as vectors in $\mathbb{R}^3$.



Which vectors are more similar, **drone** and **jet**, or **drone** and **rocket**?

Cosine similarity (optional)

A useful way to measure how similar two vectors are, is to measure the *cosine* of the angle between them:

$$\cos\theta \qquad \text{(where } \theta \text{ is the angle between } \mathbf{u} \text{ and } \mathbf{v}\text{)}$$

Cosine similarity (optional)

A useful way to measure how similar two vectors are, is to measure the *cosine* of the angle between them:

$$\cos\theta \qquad \text{(where } \theta \text{ is the angle between } \mathbf{u} \text{ and } \mathbf{v}\text{)}$$

While the L1 or L2 distances between two vectors measure how far or how *close* their endpoints are, the cosine similarity measures how *aligned* they are – independently of their lengths.

In the example of **drone**, **jet**, and **rocket**, the cosine between **drone** and **jet** is ≈ 0.95, while the cosine between **drone** and **rocket** is ≈ 0.7.

Cosine similarity (optional)

A useful way to measure how similar two vectors are, is to measure the *cosine* of the angle between them:

$$\cos\theta \qquad \text{(where } \theta \text{ is the angle between } \mathbf{u} \text{ and } \mathbf{v}\text{)}$$

While the L1 or L2 distances between two vectors measure how far or how *close* their endpoints are, the cosine similarity measures how *aligned* they are – independently of their lengths.

In the example of **drone**, **jet**, and **rocket**, the cosine between **drone** and **jet** is ≈ 0.95, while the cosine between **drone** and **rocket** is ≈ 0.7.

(check some actual word embeddings here)

Vector Space

Finally, we are left to notice two things. Take, for example,

- the set $D = \{0, 1, 2, \ldots, 9\}$ of digits, and
- the set $P = \{x \in \mathbb{R} : x > 0\}$ of positive numbers.

Vector Space

Finally, we are left to notice two things. Take, for example,

- the set $D = \{0, 1, 2, \ldots, 9\}$ of digits, and
- the set $P = \{x \in \mathbb{R} : x > 0\}$ of positive numbers.

Notice that,

- while the sum of two digits may not be a digit (e.g. $5 + 7 = 12$), the sum of two vectors is *always* a vector,

Vector Space

Finally, we are left to notice two things. Take, for example,

- the set $D = \{0, 1, 2, \ldots, 9\}$ of digits, and
- the set $P = \{x \in \mathbb{R} : x > 0\}$ of positive numbers.

Notice that,

- while the sum of two digits may not be a digit (e.g. $5 + 7 = 12$), the sum of two vectors is *always* a vector,

and

- while the product of a positive number with an arbitrary scalar c may not be positive (e.g. $4 \cdot (-1) = -4$), the product of a vector with a scalar is *always* a vector.

Vector Space

Finally, we are left to notice two things. Take, for example,

- the set $D = \{0, 1, 2, \ldots, 9\}$ of digits, and
- the set $P = \{x \in \mathbb{R} : x > 0\}$ of positive numbers.

Notice that,

- while the sum of two digits may not be a digit (e.g. $5 + 7 = 12$), the sum of two vectors is *always* a vector,

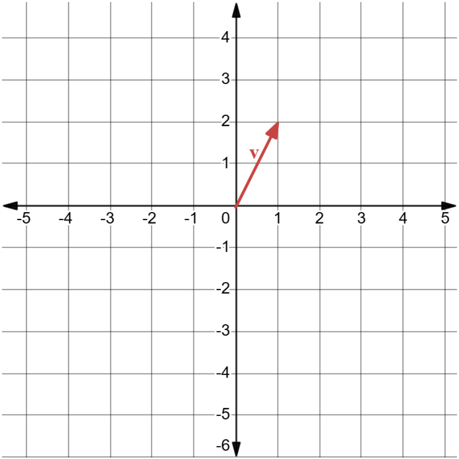
and

- while the product of a positive number with an arbitrary scalar c may not be positive (e.g. $4 \cdot (-1) = -4$), the product of a vector with a scalar is *always* a vector.

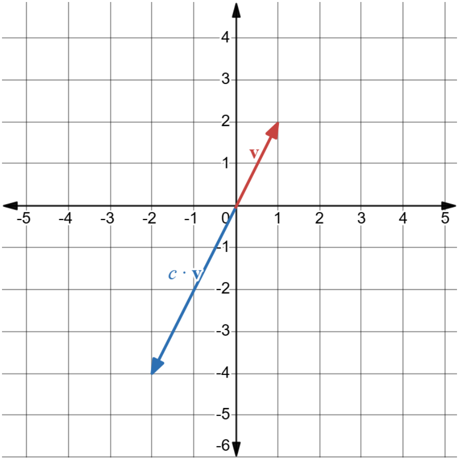
In this case we say that the set of vectors is **closed under addition and scalar multiplication**, while D or P are not (P is closed under addition only).

Vector Space

Furthermore, take the line $y = 2x$ and choose any vector on it:



After multiplying it with any number c, it will still stay on the line $y = 2x$:



Vector Space

Similarly, if we add two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which both lie on the line $y = 2x$, their sum would again be on the same line.

Vector Space

Similarly, if we add two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which both lie on the line $y = 2x$, their sum would again be on the same line.

In other words, the line $y = 2x$ is **closed under addition and scalar multiplication**, just like the whole set of vectors $\mathbb{R}^2$. This motivates us to give a special name to the good sets like the line $y = 2x$ and $\mathbb{R}^2$.

Similarly, if we add two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which both lie on the line $y = 2x$, their sum would again be on the same line.

In other words, the line $y = 2x$ is **closed under addition and scalar multiplication**, just like the whole set of vectors $\mathbb{R}^2$. This motivates us to give a special name to the good sets like the line $y = 2x$ and $\mathbb{R}^2$.

We say that $\mathbb{R}^2$ is a **vector space**, and the set of vectors lying on the line $y = 2x$ are a **vector subspace** of $\mathbb{R}^2$.

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
3. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$
7. $1 \cdot \mathbf{v} = \mathbf{v}$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$
7. $1 \cdot \mathbf{v} = \mathbf{v}$
8. $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$
7. $1 \cdot \mathbf{v} = \mathbf{v}$
8. $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$
9. $(c + d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$
7. $1 \cdot \mathbf{v} = \mathbf{v}$
8. $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$
9. $(c + d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

then it is called a **vector space** (or a **linear space**).

Vector Space

Definition

If V is a set such that:

1. it is closed under addition and scalar multiplication,
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
4. There exists a vector $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$
5. For every $\mathbf{v} \in V$, there exists a vector $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
6. $(cd) \cdot \mathbf{v} = c \cdot (d \cdot \mathbf{v})$
7. $1 \cdot \mathbf{v} = \mathbf{v}$
8. $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$
9. $(c + d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

then it is called a **vector space** (or a **linear space**).

No need to memorize–just the natural laws of addition and multiplication.

Vector Space

Definition

A subset U of a vector space V is called a **subspace** of V if U itself is a vector space.

Definition

A subset U of a vector space V is called a **subspace** of V if U itself is a vector space.

How to check if a subset U of V is a sub*space*? Just make sure that it is closed under addition and scalar multiplication!

Vector Space

Definition

A subset U of a vector space V is called a **subspace** of V if U itself is a vector space.

How to check if a subset U of V is a sub*space*? Just make sure that it is closed under addition and scalar multiplication!

Theorem

Assume V is a vector space, and U is a subset of V. Then U is a subspace of V if and only if

1. $\mathbf{x} + \mathbf{y} \in U$, for all $\mathbf{x}, \mathbf{y} \in U$,
2. $c\mathbf{x} \in U$, for all $\mathbf{x} \in U$ and $c \in \mathbb{R}$.

Vector Space

Definition

A subset U of a vector space V is called a **subspace** of V if U itself is a vector space.

How to check if a subset U of V is a sub*space*? Just make sure that it is closed under addition and scalar multiplication!

Theorem

Assume V is a vector space, and U is a subset of V. Then U is a subspace of V if and only if

1. $\mathbf{x} + \mathbf{y} \in U$, for all $\mathbf{x}, \mathbf{y} \in U$,
2. $c\mathbf{x} \in U$, for all $\mathbf{x} \in U$ and $c \in \mathbb{R}$.

- So $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, ... are all vector spaces.
- The set of all vectors that lie on the same line (e.g. $y = kx$) form a subspace (on the condition that the line also contains the $\mathbf{0}$ vector).

Matrices

Definition

An $m \times n$ table A of elements a_{ij} is called an (m, n) **matrix**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad a_{ij} \in \mathbb{R}$$

(here $i = 1, \dots, m$ and $j = 1, \dots, n$).

Matrices

Definition

An $m \times n$ table A of elements a_{ij} is called an (m, n) **matrix**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad a_{ij} \in \mathbb{R}$$

(here $i = 1, \dots, m$ and $j = 1, \dots, n$).

The set of all (m, n) matrices is denoted by $\mathbb{R}^{m \times n}$.

Matrices

Definition

An $m \times n$ table A of elements a_{ij} is called an (m, n) **matrix**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad a_{ij} \in \mathbb{R}$$

(here $i = 1, \dots, m$ and $j = 1, \dots, n$).

The set of all (m, n) matrices is denoted by $\mathbb{R}^{m \times n}$.

Example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3} \qquad B = \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

Matrices

Definition

An $m \times n$ table A of elements a_{ij} is called an (m, n) **matrix**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad a_{ij} \in \mathbb{R}$$

(here $i = 1, \dots, m$ and $j = 1, \dots, n$).

The set of all (m, n) matrices is denoted by $\mathbb{R}^{m \times n}$.

Example

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{2\times 3} \qquad B = \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

Note that the first number in (m, n) **always** shows rows, second – columns.

Matrix Addition

The vectors are practically 1-column matrices: $\mathbb{R}^n = \mathbb{R}^{n \times 1}$. Similar to vectors, we define the following operations with the matrices:

Matrix Addition

The vectors are practically 1-column matrices: $\mathbb{R}^n = \mathbb{R}^{n\times 1}$. Similar to vectors, we define the following operations with the matrices:

Definition

The sum of two matrices A and B, denoted as $A+B$, is obtained by adding corresponding elements. If A is of size $m\times n$ and B is of the same size, then $A+B$ is also of size $m\times n$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}$$

Matrix Addition

Example

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

$$A + B = \begin{bmatrix} -1 & 2 \\ 4 & 7 \end{bmatrix}$$

Matrix Addition

Example

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

$$A + B = \begin{bmatrix} -1 & 2 \\ 4 & 7 \end{bmatrix}$$

Matrix addition is only defined for matrices of the same size.

Matrix Addition

Example

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

$$A + B = \begin{bmatrix} -1 & 2 \\ 4 & 7 \end{bmatrix}$$

Matrix addition is only defined for matrices of the same size.

Question

What do you think happens if we multiply a matrix by a scalar?

Scalar Multiplication of a Matrix

Definition

If A is a matrix and c is any scalar, we define the product cA as the matrix:

$$c \cdot A = c \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} c \cdot a_{11} & c \cdot a_{12} & \dots & c \cdot a_{1n} \\ c \cdot a_{21} & c \cdot a_{22} & \dots & c \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c \cdot a_{m1} & c \cdot a_{m2} & \dots & c \cdot a_{mn} \end{bmatrix}$$

i.e. each element of the matrix is multiplied by the scalar c.

Scalar Multiplication of a Matrix

Definition

If A is a matrix and c is any scalar, we define the product cA as the matrix:

$$c \cdot A = c \cdot \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} c \cdot a_{11} & c \cdot a_{12} & \dots & c \cdot a_{1n} \\ c \cdot a_{21} & c \cdot a_{22} & \dots & c \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c \cdot a_{m1} & c \cdot a_{m2} & \dots & c \cdot a_{mn} \end{bmatrix}$$

i.e. each element of the matrix is multiplied by the scalar c.

Scalar multiplication can be performed for any scalar c and any matrix A.

Negative of a Matrix

Similarly, we define the negative of a matrix to be the matrix with all elements multiplied by -1:

Definition

The negative of a matrix A, denoted as $-A$, is obtained by changing the sign of each element in the matrix.

$$-A = -\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & -a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{m1} & -a_{m2} & \dots & -a_{mn} \end{bmatrix}$$

Matrix Subtraction

Definition

The difference of two matrices A and B (both of the same size), denoted as $A - B$, is obtained by subtracting corresponding elements:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2n} - b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \dots & a_{mn} - b_{mn} \end{bmatrix}$$

Matrix Subtraction

Definition

The difference of two matrices A and B (both of the same size), denoted as $A - B$, is obtained by subtracting corresponding elements:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2n} - b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \dots & a_{mn} - b_{mn} \end{bmatrix}$$

What happens if we add a matrix to its negative?

Zero Matrix

Definition

The **zero matrix**, denoted as O or $O_{m\times n}$, is a matrix where all elements are zero.

Example

$$O_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad O_{3\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Zero Matrix

Definition

The **zero matrix**, denoted as O or $O_{m\times n}$, is a matrix where all elements are zero.

Example

$$O_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad O_{3\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Of course, adding O has no effect on a matrix:

Remark

For any zero matrix O of the same size as matrix A, we have

$$A + O = O + A = A$$

Transpose of a Matrix

Definition

The **transpose** of a matrix A, denoted as A^T, is obtained by swapping its rows and columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

Transpose of a Matrix

Definition

The **transpose** of a matrix A, denoted as A^T, is obtained by swapping its rows and columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

Example

$$A = \begin{bmatrix} 7 & 4 & 2 \\ 0 & 1 & -3 \end{bmatrix} \qquad A^T = \begin{bmatrix} 7 & 0 \\ 4 & 1 \\ 2 & -3 \end{bmatrix}$$

Transpose of a Matrix

Definition

The **transpose** of a matrix A, denoted as A^T, is obtained by swapping its rows and columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

Example

$$A = \begin{bmatrix} 7 & 4 & 2 \\ 0 & 1 & -3 \end{bmatrix} \qquad A^T = \begin{bmatrix} 7 & 0 \\ 4 & 1 \\ 2 & -3 \end{bmatrix}$$

If A is an (m, n) matrix, then A^T is an

Transpose of a Matrix

Definition

The **transpose** of a matrix A, denoted as A^T, is obtained by swapping its rows and columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}$$

Example

$$A = \begin{bmatrix} 7 & 4 & 2 \\ 0 & 1 & -3 \end{bmatrix} \qquad A^T = \begin{bmatrix} 7 & 0 \\ 4 & 1 \\ 2 & -3 \end{bmatrix}$$

If A is an (m, n) matrix, then A^T is an (n, m) matrix.

Matrices

Matrices can be added together and multiplied by numbers, and these operations share the same "good" properties (e.g. $A + B = B + A$) with vectors.

Matrices

Matrices can be added together and multiplied by numbers, and these operations share the same "good" properties (e.g. $A + B = B + A$) with vectors.

In that sense, it is not difficult to prove that:

Theorem

For each fixed (m, n), the set $\mathbb{R}^{m \times n}$ of all $m \times n$ matrices forms a vector space.

Matrices

Matrices can be added together and multiplied by numbers, and these operations share the same "good" properties (e.g. $A + B = B + A$) with vectors.

In that sense, it is not difficult to prove that:

Theorem

For each fixed (m, n), the set $\mathbb{R}^{m\times n}$ of all $m \times n$ matrices forms a vector space.

Okay, but why do we need matrices at all?

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \\ \\ \\ \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \mathbf{A}_2 \cdot \mathbf{v} \\ \\ \\ \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \mathbf{A}_2 \cdot \mathbf{v} \\ \vdots \\ \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \mathbf{A}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{A}_m \cdot \mathbf{v} \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \mathbf{A}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{A}_m \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \ldots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \ldots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \ldots + a_{mn}v_n \end{bmatrix}$$

Matrix-Vector Multiplication

Let's define the multiplication of a matrix by a vector.

Say, we have a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \dots & \mathbf{A}_1 & \dots \\ \dots & \mathbf{A}_2 & \dots \\ & \vdots & \\ \dots & \mathbf{A}_m & \dots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

To multiply A by $\mathbf{v}$, we take each row of A and compute its dot product with the vector $\mathbf{v}$:

$$\begin{bmatrix} \mathbf{A}_1 \cdot \mathbf{v} \\ \mathbf{A}_2 \cdot \mathbf{v} \\ \vdots \\ \mathbf{A}_m \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \ldots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \ldots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \ldots + a_{mn}v_n \end{bmatrix}$$

As a result we get a new vector of size $m \times 1$.

Matrix-Vector Multiplication

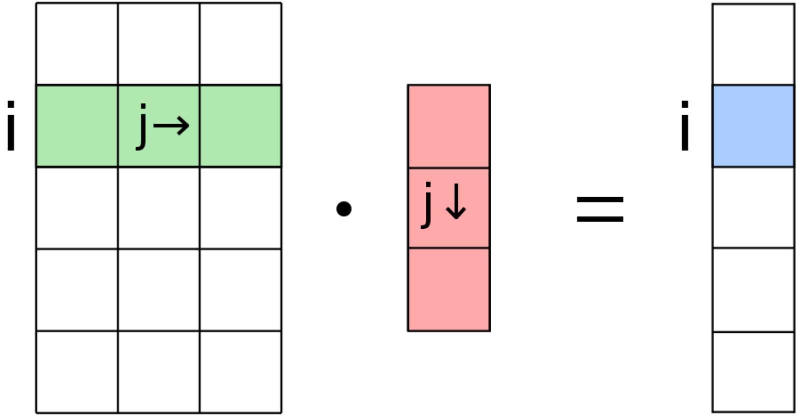
Definition

Let A be an $m \times n$ matrix and $\mathbf{v}$ be a column vector of size $n \times 1$. The product $A\mathbf{v}$ is a column vector of size $m \times 1$ obtained by multiplying each row of A by the corresponding element of $\mathbf{v}$ and summing the results.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$A\mathbf{v} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n \end{bmatrix}$$

Matrix-Vector Multiplication



Matrix-Vector Multiplication

Example

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

$$A\mathbf{v} = \begin{bmatrix} 1 \cdot 2 + 2 \cdot (-1) + 3 \cdot 3 \\ 4 \cdot 2 + 5 \cdot (-1) + 6 \cdot 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 21 \end{bmatrix}$$

Matrix-Vector Multiplication

Example

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

$$A\mathbf{v} = \begin{bmatrix} 1 \cdot 2 + 2 \cdot (-1) + 3 \cdot 3 \\ 4 \cdot 2 + 5 \cdot (-1) + 6 \cdot 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 21 \end{bmatrix}$$

Example

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

$$A\mathbf{v} = \begin{bmatrix} (-2) \cdot 4 + 1 \cdot 2 \\ 0 \cdot 4 + 3 \cdot 2 \\ 1 \cdot 4 + (-1) \cdot 2 \end{bmatrix} = \begin{bmatrix} -6 \\ 6 \\ 2 \end{bmatrix}$$

Matrix-Vector Multiplication

Matrix-vector multiplication shares properties with scalar multiplication and addition of vectors.

- **Distributive Property:**
  For a matrix A and vectors $\mathbf{v}$ and $\mathbf{w}$ of appropriate sizes:

$$A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w}$$

- **Scalar Multiplication:**
  For a matrix A and a scalar c:

$$A(c\mathbf{v}) = c(A\mathbf{v})$$

Matrix-Vector Multiplication

Matrix-vector multiplication shares properties with scalar multiplication and addition of vectors.

- **Distributive Property:**
  For a matrix A and vectors $\mathbf{v}$ and $\mathbf{w}$ of appropriate sizes:

$$A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w}$$

- **Scalar Multiplication:**
  For a matrix A and a scalar c:

$$A(c\mathbf{v}) = c(A\mathbf{v})$$

Note that we can only multiply a matrix by a vector if the number of columns of the matrix equals the length of the vector
(i.e. **not all** matrices and vectors can be multiplied)!

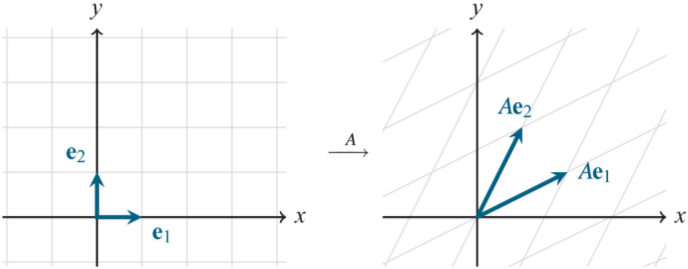
Geometric Interpretation

Why do we define the matrix-vector multiplication this way? Turns out, it has a beautiful geometrical interpretation.

Geometric Interpretation

Why do we define the matrix-vector multiplication this way? Turns out, it has a beautiful geometrical interpretation.

Think this way: when you multiply, say, a 2×2 matrix A by a vector $\mathbf{v} \in \mathbb{R}^2$, what you get is another vector $\mathbf{u} = A\mathbf{v} \in \mathbb{R}^2$. We call this $\mathbf{u}$ the **transformed version** of $\mathbf{v}$ (and we say that A is a linear transformation).



Geometric Interpretation

As we will see later, the resulting "transformed version" $\mathbf{u}$ is just the same old $\mathbf{v}$ except it is **rotated** and **scaled** to become longer or shorter (and possibly, flipped).

Geometric Interpretation

As we will see later, the resulting "transformed version" **u** is just the same old **v** except it is **rotated** and **scaled** to become longer or shorter (and possibly, flipped).

In this sense, all matrices are either just rotating vectors by some degree, or flipping them horizontally/vertically, or scale them, or do all three.

The key thing is: whatever a matrix "does" to one vector, it does the same to all other vectors too (when being multiplied with them).

Check different matrices yourself:
- `visualize-it.github.io/linear_transformations/simulation.html`
- `www.shad.io/MatVis`

We will learn more about this later–now back to matrices∼

Matrix Multiplication

Definition

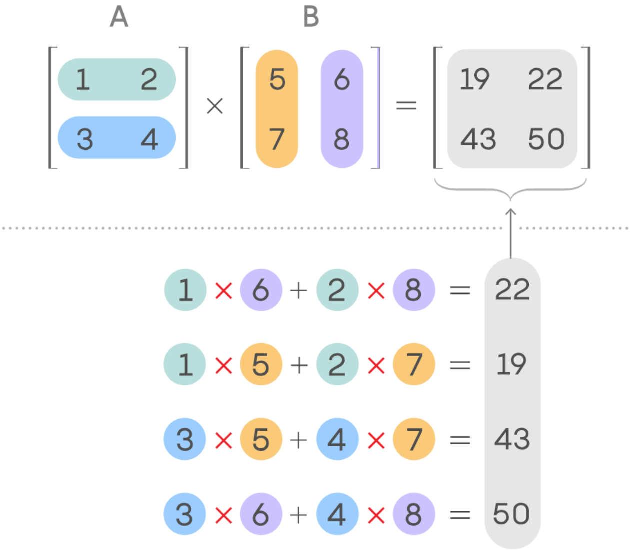
Let A be an $m \times n$ matrix, and let B be an $n \times k$ matrix. The product $C = AB$ is an $m \times k$ matrix, where each element c_{ij} is the dot product of the i^{th} row of A and the j^{th} column of B:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{bmatrix}$$

$$C = AB = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1k} \\ c_{21} & c_{22} & \dots & c_{2k} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mk} \end{bmatrix}$$

where $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$.

Matrix Multiplication



Matrix Multiplication

Matrix multiplication shares properties with scalar multiplication and addition of vectors, as well as matrix-vector multiplication.

- **Distributive Property:**

$$A(B+C) = AB + AC \quad \text{and} \quad (A+B)C = AC + BC$$

- **Associative Property:**

$$A(BC) = (AB)C$$

- **Scalar Multiplication:**

$$A(cB) = c(AB) = (cA)B$$

Matrix Multiplication

Matrix multiplication shares properties with scalar multiplication and addition of vectors, as well as matrix-vector multiplication.

- **Distributive Property:**

$$A(B+C) = AB + AC \quad \text{and} \quad (A+B)C = AC + BC$$

- **Associative Property:**

$$A(BC) = (AB)C$$

- **Scalar Multiplication:**

$$A(cB) = c(AB) = (cA)B$$

Note that we can only multiply A and B if the (number of columns in A) equals the (number of rows in B): $(m \times n)$ with $(n \times k)$.

In particular, this means that matrix multiplication is **not commutative**, i.e. in general, $AB \neq BA$!

Matrix Multiplication

Example

Let

$$C = \begin{bmatrix} -1 & 0 \\ 2 & -3 \\ 4 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2} \qquad D = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 0 & 7 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

$$CD = \begin{bmatrix} -1\cdot 5 + 0\cdot 3 & -1\cdot(-2) + 0\cdot 0 & -1\cdot 1 + 0\cdot 7 \\ 2\cdot 5 + (-3)\cdot 3 & 2\cdot(-2) + (-3)\cdot 0 & 2\cdot 1 + (-3)\cdot 7 \\ 4\cdot 5 + 1\cdot 3 & 4\cdot(-2) + 1\cdot 0 & 4\cdot 1 + 1\cdot 7 \end{bmatrix}$$

$$= \begin{bmatrix} -5 & 2 & -1 \\ 1 & -4 & -19 \\ 23 & -8 & 11 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

Matrix Multiplication

Example

Let

$$C = \begin{bmatrix} -1 & 0 \\ 2 & -3 \\ 4 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2} \qquad D = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 0 & 7 \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

$$CD = \begin{bmatrix} -1\cdot 5 + 0\cdot 3 & -1\cdot(-2) + 0\cdot 0 & -1\cdot 1 + 0\cdot 7 \\ 2\cdot 5 + (-3)\cdot 3 & 2\cdot(-2) + (-3)\cdot 0 & 2\cdot 1 + (-3)\cdot 7 \\ 4\cdot 5 + 1\cdot 3 & 4\cdot(-2) + 1\cdot 0 & 4\cdot 1 + 1\cdot 7 \end{bmatrix}$$

$$= \begin{bmatrix} -5 & 2 & -1 \\ 1 & -4 & -19 \\ 23 & -8 & 11 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

What about DC?