Geometry of Vectors, Matrices

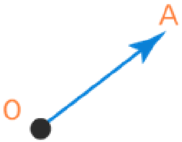
Hayk Aprikyan, Hayk Tarkhanyan

Armenian Code Academy

March 21, 2025

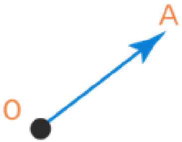
Norm

What if we want to measure the length of some vector?



Norm

What if we want to measure the length of some vector?



What we can say, is that

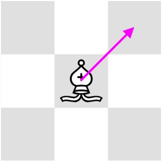
$$\text{the length of the vector}$$
$$=$$
$$\text{the distance between } O \text{ and } A.$$

Norm

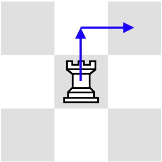
But how to measure distance?

Norm

But how to measure distance?

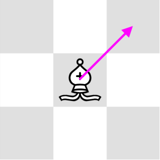


For a bishop, the distance to its upper-right neighbor is 1.

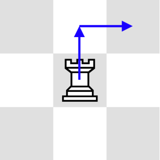


While for a rook, it is 2.

But how to measure distance?



For a bishop, the distance to its upper-right neighbor is 1.



While for a rook, it is 2.

So there are different ways to measure distance and length.

Norm

For a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, its **Euclidean norm** or **L2 norm** is:

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

or, equivalently,

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

Norm

For a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, its **Euclidean norm** or **L2 norm** is:

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

or, equivalently,

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

Example

Let $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. The Euclidean norm of $\mathbf{v}$ is:

$$\|\mathbf{v}\|_2 = \sqrt{3^2 + 4^2} = 5$$

Norm

For a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, its **Euclidean norm** or **L2 norm** is:

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

or, equivalently,

$$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

Example

Let $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. The Euclidean norm of $\mathbf{v}$ is:

$$\|\mathbf{v}\|_2 = \sqrt{3^2 + 4^2} = 5$$

Euclidean norm is the standard length we use in classic geometry.
Sometimes we omit the little "2" and just write $\|\mathbf{v}\|$ instead of $\|\mathbf{v}\|_2$.

Norm

For a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, its **Manhattan norm** or **L1 norm** is:

$$\|\mathbf{v}\|_1 = |v_1| + |v_2| + \cdots + |v_n|$$

Norm

For a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$, its **Manhattan norm** or **L1 norm** is:

$$\|\mathbf{v}\|_1 = |v_1| + |v_2| + \cdots + |v_n|$$

Example

Let $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. The Manhattan norm of $\mathbf{v}$ is:

$$\|\mathbf{v}\|_1 = |3| + |4| = 7$$

Norm

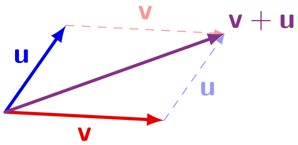
As we have seen, there are different types of norms (=many different ways to calculate the length of a vector), and one of them is chosen depending on the problem.

Norm

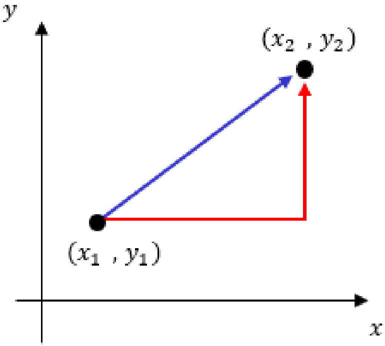
As we have seen, there are different types of norms (=many different ways to calculate the length of a vector), and one of them is chosen depending on the problem.

Notice, however, that independently of which one we take, all norms always satisfy the following three properties:

1. $\|\mathbf{v}\| \geq 0$, and equals 0 if only if $\mathbf{v} = \mathbf{0}$,
2. $\|c\mathbf{v}\| = |c| \cdot \|\mathbf{v}\|$,
3. $\|\mathbf{v} + \mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$.



Norm



Manhattan Distance L^1

Euclidean Distance L^2

$$L^1 = |x_2 - x_1| + |y_2 - y_1|$$

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

Angle between vectors

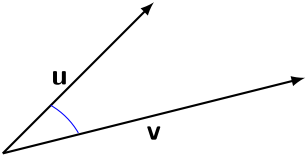
Geometrically, a vector is an arrow in space, that is, it has both a length and direction.

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?

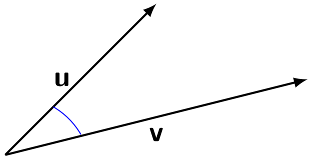
Angle between vectors

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 **u** and **v** is the angle $0 \leq \theta \leq \pi$ for which:

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

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 $\mathbf{v}$.

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

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

Angle between vectors

Corollary 1

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 1

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 $\mathbf{u}$.

Corollary 2

The dot product of two vectors equals 0 if and only if they are perpendicular to each other (form a 90° angle).

Angle between vectors

Corollary 1

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 $\mathbf{u}$.

Corollary 2

The dot product of two vectors equals 0 if and only if they are perpendicular to each other (form a 90° angle).

Corollary 3

Any vector $\mathbf{v} \in \mathbb{R}^n$ forms an angle of 0° with itself and 180° with its negative.

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 $D = \{0, 1, 2, \dots, 9\}$ 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.

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,

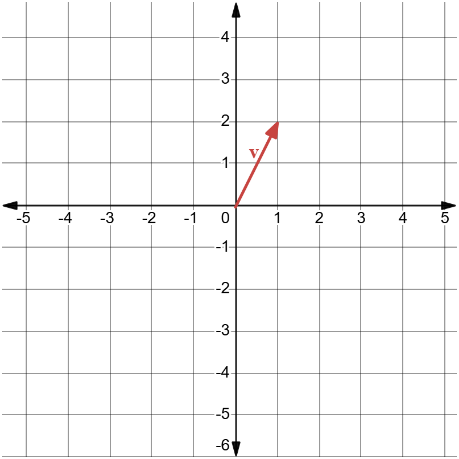
and

- 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.

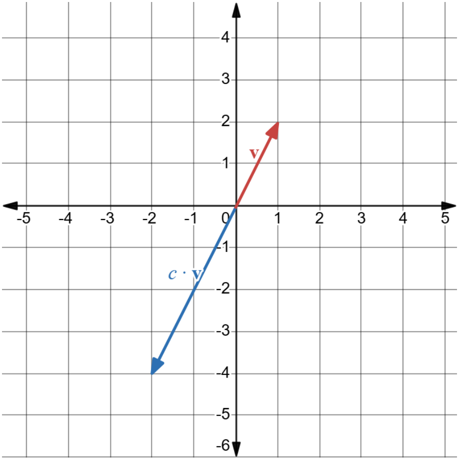
In this case we say that the set of vectors is **closed under addition and scalar multiplication**, while D or P are not (P is closed under addition only).

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 $\mathbb{R}^2$.

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 $\mathbb{R}^2$.

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

A set V is called a **vector space** if

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. $(c + d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

No need to memorize the properties–just the natural laws of 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.

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 $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\ldots$ 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 $m \times n$ tuple A of elements a_{ij} ($i = 1, \ldots, m$ and $j = 1, \ldots, n$), is called a real-valued (m, n) **matrix**:

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

The set of all real-valued (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 $m \times n$ tuple A of elements a_{ij} $(i = 1, \ldots, m$ and $j = 1, \ldots, n)$, is called a real-valued (m, n) **matrix**:

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

The set of all real-valued (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

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

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

Remark

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

Scalar Multiplication of a Matrix

Definition

The product of a scalar c and a matrix A, denoted as cA, is obtained by multiplying each element of the matrix by the scalar.

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

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

Scalar Multiplication of a Matrix

Definition

The product of a scalar c and a matrix A, denoted as cA, is obtained by multiplying each element of the matrix by the scalar.

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

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

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

Negative of a Matrix

Definition

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

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

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

Negative of a Matrix

Definition

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

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

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

Remark

The negative of a matrix equals (-1) times the matrix.

Matrix Subtraction

Definition

The difference of two matrices A and B, denoted as $A - B$, is obtained by subtracting corresponding elements, or by adding A and $-B$. If A and B are both of size $m \times n$, then $A - B$ is also of size $m \times n$.

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

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

Matrix Subtraction

Definition

The difference of two matrices A and B, denoted as $A - B$, is obtained by subtracting corresponding elements, or by adding A and $-B$. If A and B are both of size $m \times n$, then $A - B$ is also of size $m \times n$.

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

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

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

Zero Matrix

Definition

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

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

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

Remark

$A + O = O + A = A$ for any matrix A.

Transpose of a Matrix

Definition

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

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}$$

Transpose of a Matrix

Definition

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

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}$$

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.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}$$

Example

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

Remark

The transpose of an (m, n) matrix 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 $m, n \in \mathbb{N}$ the set of real-valued matrices $\mathbb{R}^{m \times n}$ forms a vector space.

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.

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

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

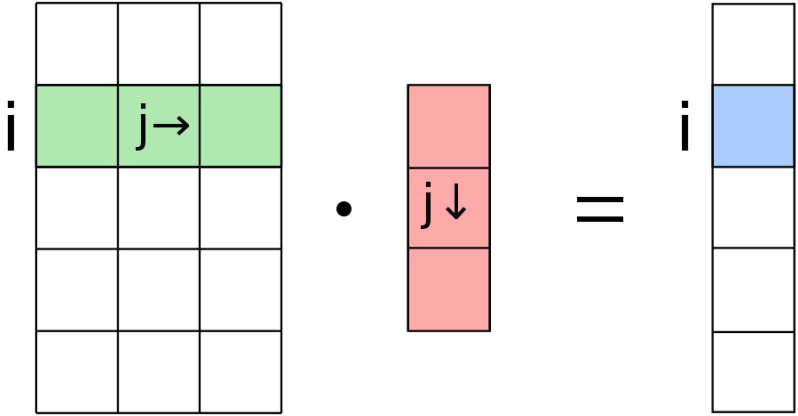
Matrix-Vector Multiplication

Or, in other words, if we denote the rows of A by $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m$, the product $A\mathbf{v}$ will be a column vector of size $m \times 1$ obtained by taking the dot product of each row of A with the vector $\mathbf{v}$:

$$A = \begin{bmatrix} \ldots & \mathbf{A}_1 & \ldots \\ \ldots & \mathbf{A}_2 & \ldots \\ & \vdots & \\ \ldots & \mathbf{A}_m & \ldots \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

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

Matrix-Vector Multiplication



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

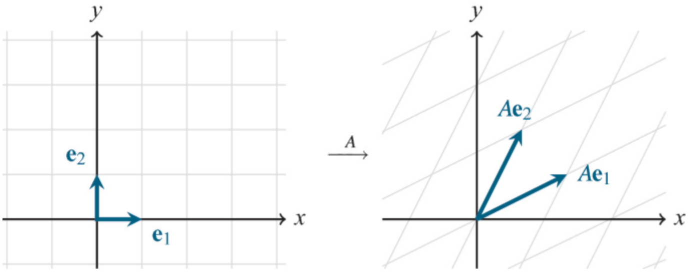
Note that we can only multiply a matrix by a vector if the number of columns of the matrix equals the length of the vector.

Geometric Interpretation

Why do we define the matrix-vector multiplication this way? Turns out, it has a beautiful geometrical 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" $\mathbf{u}$ is just the same old $\mathbf{v}$ 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

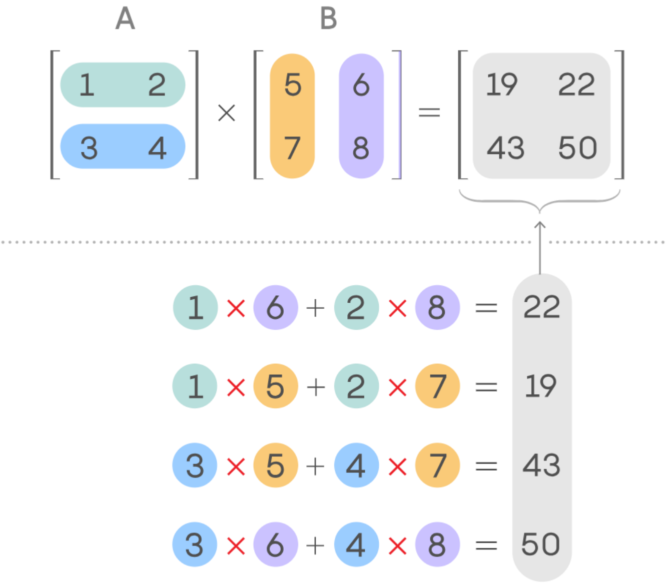
Definition

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 obtained by taking the dot product of the i-th row of A and the j-th column of B:

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

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

$$\text{where } c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{p=1}^{n} a_{ip}b_{pj}$$


A
B
1 2
3 4
×
5 6
7 8
=
19 22
43 50
1 × 6 + 2 × 8 = 22
1 × 5 + 2 × 7 = 19
3 × 5 + 4 × 7 = 43
3 × 6 + 4 × 8 = 50

Matrix Multiplication

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

- **Distributive Property:**
  For matrices A, B, and C of appropriate sizes:
  $$A(B + C) = AB + AC \quad \text{and} \quad (A + B)C = AC + BC$$
- **Associativity Property:**
  For matrices A, B, and C of appropriate sizes:
  $$A(BC) = (AB)C$$
- **Scalar Multiplication:**
  For matrices A, B of appropriate sizes and a scalar c:
  $$A(cB) = c(AB) = (cA)B$$

Matrix Multiplication

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

- **Distributive Property:**
  For matrices A, B, and C of appropriate sizes:
  $$A(B+C) = AB + AC \quad \text{and} \quad (A+B)C = AC + BC$$
- **Associativity Property:**
  For matrices A, B, and C of appropriate sizes:
  $$A(BC) = (AB)C$$
- **Scalar Multiplication:**
  For matrices A, B of appropriate sizes and a scalar c:
  $$A(cB) = c(AB) = (cA)B$$

Note that we can only multiply two matrices if the number of columns of the first matrix equals the number of rows of the second matrix: $(m \times n)$ with $(n \times k)$.

Matrix Multiplication

Example

Let

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

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

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