COMP 761: Lecture 10 – Linear Algebra I

David Rolnick

September 25, 2020

Course Announcements

Course Announcements

- Vincent's optional list of practice problems available on Slack

Course Announcements

- Vincent's optional list of practice problems available on Slack
- Monday class attendance optional if observing Yom Kippur

Vectors

Vectors

- A (real) *vector* v is a tuple of real numbers $[v_1, \ldots, v_n]$.

Vectors

- A (real) *vector* v is a tuple of real numbers $[v_1, \ldots, v_n]$.
- We write $v \in \mathbb{R}^n$ and say that v has *dimension* n.

Vectors

- A (real) *vector* v is a tuple of real numbers $[v_1, \ldots, v_n]$.
- We write $v \in \mathbb{R}^n$ and say that v has *dimension* n.
- One can also work with vectors where the v_k are complex numbers, rational numbers, etc, but we won't need that

Vectors

- A (real) *vector* v is a tuple of real numbers $[v_1, \ldots, v_n]$.
- We write $v \in \mathbb{R}^n$ and say that v has *dimension* n.
- One can also work with vectors where the v_k are complex numbers, rational numbers, etc, but we won't need that
- Order matters, so e.g. $[2, 1, 1] \neq [1, 2, 1]$

Vectors

- A (real) *vector* v is a tuple of real numbers $[v_1, \ldots, v_n]$.
- We write $v \in \mathbb{R}^n$ and say that v has *dimension* n.
- One can also work with vectors where the v_k are complex numbers, rational numbers, etc, but we won't need that
- Order matters, so e.g. $[2, 1, 1] \neq [1, 2, 1]$
- When working with real vectors, we say real numbers are *scalars* to distinguish from vectors. So $[2, 2]$ is a vector, 2 is a scalar.

Basic properties

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$[v_1, \ldots, v_n] + [w_1, \ldots, w_n] = [v_1 + w_1, \ldots, v_n + w_n]$$
$$[v_1, \ldots, v_n] - [w_1, \ldots, w_n] = [v_1 - w_1, \ldots, v_n - w_n]$$

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$[v_1, \ldots, v_n] + [w_1, \ldots, w_n] = [v_1 + w_1, \ldots, v_n + w_n]$$
$$[v_1, \ldots, v_n] - [w_1, \ldots, w_n] = [v_1 - w_1, \ldots, v_n - w_n]$$

- Multiplication or division of a vector by a scalar also works component-wise:

$$c[v_1, \ldots, v_n] = [cv_1, \ldots, cv_n]$$
$$\frac{1}{c}[v_1, \ldots, v_n] = [\frac{v_1}{c}, \ldots, \frac{v_n}{c}]$$

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$[v_1, \ldots, v_n] + [w_1, \ldots, w_n] = [v_1 + w_1, \ldots, v_n + w_n]$$
$$[v_1, \ldots, v_n] - [w_1, \ldots, w_n] = [v_1 - w_1, \ldots, v_n - w_n]$$

- Multiplication or division of a vector by a scalar also works component-wise:

$$c[v_1, \ldots, v_n] = [cv_1, \ldots, cv_n]$$
$$\frac{1}{c}[v_1, \ldots, v_n] = [\frac{v_1}{c}, \ldots, \frac{v_n}{c}]$$

- If $v = [1, 0]$ and $w = [0, 1]$, what is $5v - 2w$?

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$[v_1, \dots, v_n] + [w_1, \dots, w_n] = [v_1 + w_1, \dots, v_n + w_n]$$
$$[v_1, \dots, v_n] - [w_1, \dots, w_n] = [v_1 - w_1, \dots, v_n - w_n]$$

- Multiplication or division of a vector by a scalar also works component-wise:

$$c[v_1, \dots, v_n] = [cv_1, \dots, cv_n]$$
$$\frac{1}{c}[v_1, \dots, v_n] = [\frac{v_1}{c}, \dots, \frac{v_n}{c}]$$

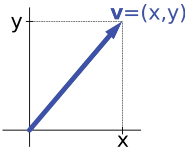
- If $v = [1, 0]$ and $w = [0, 1]$, what is $5v - 2w$?

$$5v - 2w = [5, 0] - [0, 2] = [5, -2].$$

Graphical representations

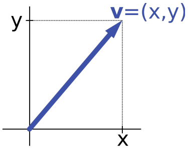
Graphical representations

- Often represented by points in Cartesian coordinates:

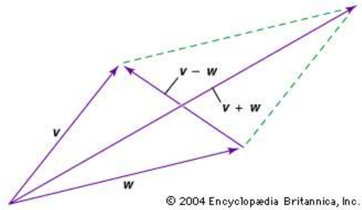


Graphical representations

- Often represented by points in Cartesian coordinates:



- Or by arrows with a fixed length and direction:



Norms of a vector

Norms of a vector

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

Norms of a vector

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

- What is the ℓ^1-norm of x?

Norms of a vector

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

- What is the ℓ^1-norm of x? $||x||_1 = |x_1| + |x_2| + \cdots + |x_n|$ (sometimes called the *taxi-cab distance*)

Norms of a vector

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- What is the ℓ^1-norm of x? $||x||_1 = |x_1| + |x_2| + \cdots + |x_n|$ (sometimes called the *taxi-cab distance*)
- Most common norm is the ℓ^2-norm:

$$||x||_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Norms of a vector

- For $p \geq 1$, the ℓ^p-*norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

- What is the ℓ^1-norm of x? $||x||_1 = |x_1| + |x_2| + \cdots + |x_n|$ (sometimes called the *taxi-cab distance*)
- Most common norm is the ℓ^2-norm:

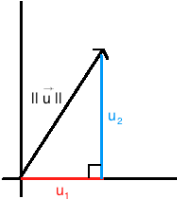
$$||x||_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

- Often, $||x||_2$ is just written $||x||$.

Geometric interpretation

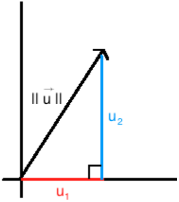
Geometric interpretation

- Geometrically, the ℓ^2-norm of a vector is equal to its length:



Geometric interpretation

- Geometrically, the ℓ^2-norm of a vector is equal to its length:



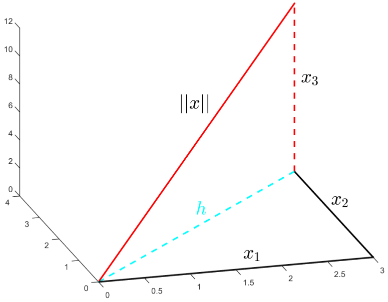
- By the Pythagorean Theorem:

$$\text{Length}^2 = u_1^2 + u_2^2 = ||u||_2^2.$$

Geometric interpretation

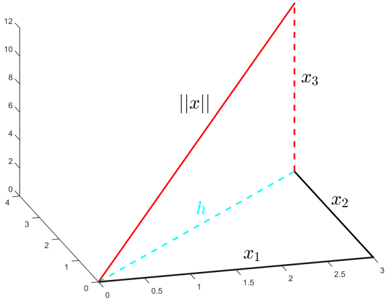
Geometric interpretation

- Why is this true in higher dimensions?



Geometric interpretation

- Why is this true in higher dimensions?



- In 3D, we can apply the Pythagorean Theorem twice:

$$x_1^2 + x_2^2 = h^2$$
$$h^2 + x_3^2 = \text{Length}^2$$
$$\text{so Length}^2 = x_1^2 + x_2^2 + x_3^2 = ||x||_2^2.$$

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- Let's look at some different norms of the vector $x = [1, -10, 8, 0]$:

$$||x||_1 = 19, \quad ||x||_2 \approx 12.85, \quad ||x||_5 \approx 10.58, \ \ldots, \quad ||x||_{20} \approx 10.01.$$

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p} .$$

- Let's look at some different norms of the vector $x = [1, -10, 8, 0]$:

$$||x||_1 = 19, \quad ||x||_2 \approx 12.85, \quad ||x||_5 \approx 10.58, \ \ldots, \quad ||x||_{20} \approx 10.01.$$

- What does it look like?

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- Let's look at some different norms of the vector $x = [1, -10, 8, 0]$:

$$||x||_1 = 19, \quad ||x||_2 \approx 12.85, \quad ||x||_5 \approx 10.58, \ \ldots, \quad ||x||_{20} \approx 10.01.$$

- What does it look like?
- They are getting close to $\max_k |x_k| = 10$.

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- Let's look at some different norms of the vector $x = [1, -10, 8, 0]$:

$$||x||_1 = 19, \quad ||x||_2 \approx 12.85, \quad ||x||_5 \approx 10.58, \ \ldots, \quad ||x||_{20} \approx 10.01.$$

- What does it look like?
- They are getting close to $\max_k |x_k| = 10$.
- That is because for p large,

$$(|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p} \approx \left(\max_k |x_k|^p\right)^{1/p} = \max_k |x_k|.$$

The ℓ^∞ norm

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \dots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- Let's look at some different norms of the vector $x = [1, -10, 8, 0]$:

$$||x||_1 = 19, \quad ||x||_2 \approx 12.85, \quad ||x||_5 \approx 10.58, \; \dots, \quad ||x||_{20} \approx 10.01.$$

- What does it look like?
- They are getting close to $\max_k |x_k| = 10$.
- That is because for p large,

$$(|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p} \approx \left(\max_k |x_k|^p\right)^{1/p} = \max_k |x_k|.$$

- Formally, we define the ℓ^∞-norm of x to be $\max_k |x_k|$.

The ℓ^0-"norm"

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

The ℓ^0-"norm"

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}.$$

- Sometimes, people will talk about the ℓ^0*-norm* $||x||_0$ of a vector x, meaning the number of non-zero entries. Also called the *sparsity*.

The ℓ^0-"norm"

- For $p \geq 1$, the ℓ^p*-norm* (or just *p-norm*) of a vector $x = (x_1, \ldots, x_n)$ is the scalar value:

$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

- Sometimes, people will talk about the ℓ^0*-norm* $||x||_0$ of a vector x, meaning the number of non-zero entries. Also called the *sparsity*.
- Technically, this is not a mathematical "norm" (because $||cx||_0 \neq |c| \cdot ||x||_0$), but still very useful.

Dot products

Dot products

- The *dot product* (or *scalar product*) of vectors $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_n]$ is given by:

$$x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Dot products

- The *dot product* (or *scalar product*) of vectors $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_n]$ is given by:

$$x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

- Example: if $x = [3, -1, 0]$, $y = [2, 1, 2]$, then

$$x \cdot y = (3)(2) + (-1)(1) + (0)(2) = 5.$$

Dot products

- The *dot product* (or *scalar product*) of vectors $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_n]$ is given by:

$$x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

- Example: if $x = [3, -1, 0]$, $y = [2, 1, 2]$, then

$$x \cdot y = (3)(2) + (-1)(1) + (0)(2) = 5.$$

- What is $x \cdot x$?

Dot products

- The *dot product* (or *scalar product*) of vectors $x = [x_1, \ldots, x_n]$ and $y = [y_1, \ldots, y_n]$ is given by:

$$x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

- Example: if $x = [3, -1, 0]$, $y = [2, 1, 2]$, then

$$x \cdot y = (3)(2) + (-1)(1) + (0)(2) = 5.$$

- What is $x \cdot x$?

$$x \cdot x = x_1^2 + x_2^2 + \cdots + x_n^2 = ||x||_2^2.$$

Problem

Prove that the dot product between vectors *a* and *b* is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

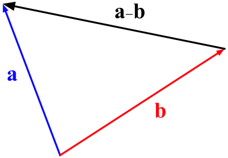
Problem

Prove that the dot product between vectors *a* and *b* is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

- Let's look at this geometrically:



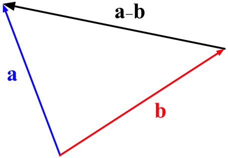
Problem

Prove that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

- Let's look at this geometrically:



- $a, b, a - b$ form a triangle.

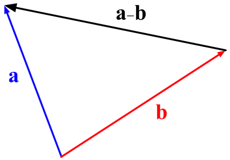
Problem

Prove that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

- Let's look at this geometrically:



- $a, b, a-b$ form a triangle.
- We know the lengths of the sides: $||a||_2, ||b||_2, ||a-b||_2$.

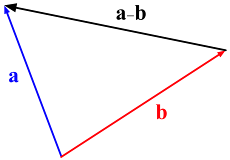
Problem

Prove that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

- Let's look at this geometrically:



- $a, b, a-b$ form a triangle.
- We know the lengths of the sides: $||a||_2, ||b||_2, ||a-b||_2$.
- There is a geometric result called the Law of Cosines (generalizing the Pythagorean Theorem) between the sides a, b, c of any triangle:

$$\text{len}(a)^2 + \text{len}(b)^2 = \text{len}(c)^2 + 2\text{len}(a)\text{len}(b)\cos\theta,$$

where θ is the angle between a and b.

Problem

Prove that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta,$$

where θ is the geometric angle between the vectors.

- $a, b, a-b$ form a triangle.
- We know the lengths of the sides: $||a||_2, ||b||_2, ||a-b||_2$.
- There is a geometric result called the Law of Cosines (generalizing the Pythagorean Theorem) between the sides a, b, c of any triangle:

$$\text{len}(a)^2 + \text{len}(b)^2 = \text{len}(c)^2 + 2\text{len}(a)\text{len}(b)\cos\theta,$$

  where θ is the angle between a and b.
- We can substitute in our lengths $||a||_2, ||b||_2, ||a-b||_2$:

$$||a||_2^2 + ||b||_2^2 = ||a-b||_2^2 + 2||a||_2||b||_2\cos\theta$$

Problem

- We can substitute in our lengths $||a||_2, ||b||_2, ||a-b||_2$:

$$||a||_2^2 + ||b||_2^2 = ||a-b||_2^2 + 2||a||_2||b||_2 \cos\theta$$

Problem

- We can substitute in our lengths $||a||_2, ||b||_2, ||a-b||_2$:

$$||a||_2^2 + ||b||_2^2 = ||a-b||_2^2 + 2||a||_2||b||_2 \cos\theta$$

- We can expand this:

$$\left(a_1^2 + \cdots + a_n^2\right) + \left(b_1^2 + \cdots + b_n^2\right)$$
$$= \left((a_1 - b_1)^2 + \cdots + (a_n - b_n)^2\right) - 2||a||_2||b||_2 \cos\theta$$
$$= \left(a_1^2 - 2a_1b_1 + b_1^2\right) + \cdots + \left(a_n^2 - 2a_nb_n + b_n^2\right) + 2||a||_2||b||_2 \cos\theta$$

Problem

- We can substitute in our lengths $||a||_2, ||b||_2, ||a-b||_2$:

$$||a||_2^2 + ||b||_2^2 = ||a-b||_2^2 + 2||a||_2||b||_2 \cos\theta$$

- We can expand this:

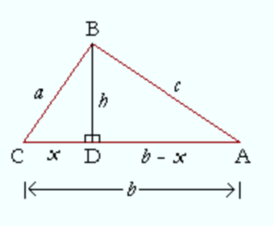
$$\begin{aligned}
&\left(a_1^2 + \cdots + a_n^2\right) + \left(b_1^2 + \cdots + b_n^2\right) \\
&= \left((a_1 - b_1)^2 + \cdots + (a_n - b_n)^2\right) - 2||a||_2||b||_2 \cos\theta \\
&= \left(a_1^2 - 2a_1b_1 + b_1^2\right) + \cdots + \left(a_n^2 - 2a_nb_n + b_n^2\right) + 2||a||_2||b||_2 \cos\theta
\end{aligned}$$

- Therefore:

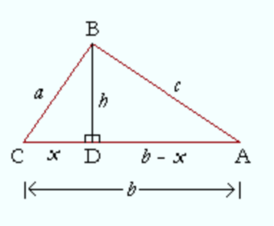
$$2a_1b_1 + \cdots + 2a_nb_n = 2||a||_2||b||_2 \cos\theta,$$

and therefore $a \cdot b = ||a||_2||b||_2 \cos\theta$.

So why is the Law of Cosines true?



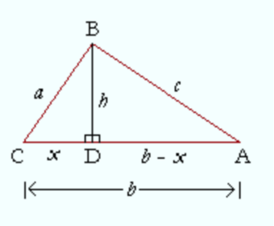
So why is the Law of Cosines true?



- We can use the Pythagorean Theorem again:

$$
\begin{aligned}
a^2 &= x^2 + h^2 \\
&= x^2 + (c^2 - (b - x)^2) \\
&= x^2 + (c^2 - b^2 + 2bx - x^2) \\
&= c^2 - b^2 + 2bx \\
&= c^2 - b^2 + 2ab\cos C.
\end{aligned}
$$

So why is the Law of Cosines true?



- We can use the Pythagorean Theorem again:

$$\begin{aligned} a^2 &= x^2 + h^2 \\ &= x^2 + (c^2 - (b-x)^2) \\ &= x^2 + (c^2 - b^2 + 2bx - x^2) \\ &= c^2 - b^2 + 2bx \\ &= c^2 - b^2 + 2ab\cos C. \end{aligned}$$

- (Note: A slightly different proof is needed for obtuse triangles.)

Problem

So what is the dot product of two perpendicular vectors?

Problem

So what is the dot product of two perpendicular vectors?

- We just saw that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta.$$

Problem

So what is the dot product of two perpendicular vectors?

- We just saw that the dot product between vectors a and b is equal to:

$$||a||_2||b||_2 \cos\theta.$$

- If a and b are perpendicular, then $\theta = 90^\circ$, so $\cos\theta = 0$ and therefore

$$a \cdot b = 0.$$

Matrices

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \leq i \leq m$ and $1 \leq j \leq n$.

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \leq i \leq m$ and $1 \leq j \leq n$.
- This is often represented:

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

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \le i \le m$ and $1 \le j \le n$.
- This is often represented:

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

- We write $A \in \mathbb{R}^{m \times n}$.

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \leq i \leq m$ and $1 \leq j \leq n$.
- This is often represented:

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

- We write $A \in \mathbb{R}^{m \times n}$.
- If $m = n$, we say that A is a *square matrix*.

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \leq i \leq m$ and $1 \leq j \leq n$.
- This is often represented:

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

- We write $A \in \mathbb{R}^{m \times n}$.
- If $m = n$, we say that A is a *square matrix*.
- If $m = 1$, then A is a *row vector*, e.g. $\begin{bmatrix} 1 & 0 \end{bmatrix}$.

Matrices

- A matrix A is a set of real numbers a_{ij} with indices $1 \leq i \leq m$ and $1 \leq j \leq n$.
- This is often represented:

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

- We write $A \in \mathbb{R}^{m \times n}$.
- If $m = n$, we say that A is a *square matrix*.
- If $m = 1$, then A is a *row vector*, e.g. $\begin{bmatrix} 1 & 0 \end{bmatrix}$.
- If $n = 1$, then A is a *column vector*, e.g. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Basic properties

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \pm \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11} \pm b_{11} & \cdots & a_{1n} \pm b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} \pm b_{m1} & \cdots & a_{mn} \pm b_{mn} \end{bmatrix}$$

Basic properties

- Vector addition and subtraction are defined as performing those operations component-wise:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \pm \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} \pm b_{11} & \cdots & a_{1n} \pm b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} \pm b_{m1} & \cdots & a_{mn} \pm b_{mn} \end{bmatrix}$$

- Multiplication (or division) of a vector by a scalar also works component-wise:

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

Matrix-vector multiplication

Matrix-vector multiplication

- For a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $x \in \mathbb{R}^n$, the product Ax is defined by

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

Matrix-vector multiplication

- For a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $x \in \mathbb{R}^n$, the product Ax is defined by

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

- This is just the column vector formed by the dot products of x with each of the rows of A.

Matrix-vector multiplication

- For a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $x \in \mathbb{R}^n$, the product Ax is defined by

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

- This is just the column vector formed by the dot products of x with each of the rows of A.
- Note: the dimensions have to match up, and Ax is a vector in $\mathbb{R}^m$.

Matrix-vector multiplication

- For a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $x \in \mathbb{R}^n$, the product Ax is defined by

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

- This is just the column vector formed by the dot products of x with each of the rows of A.
- Note: the dimensions have to match up, and Ax is a vector in $\mathbb{R}^m$.
- Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -3 \\ 4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

The identity matrix

The identity matrix

- The $n \times n$ identity matrix is the matrix $A = [a_{ij}]$ where a_{ij} equals 1 if $i = j$ and otherwise equals 0.

The identity matrix

- The $n \times n$ identity matrix is the matrix $A = [a_{ij}]$ where a_{ij} equals 1 if $i = j$ and otherwise equals 0.
- Example: the 4×4 identity matrix is:

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

- If $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is the identity matrix, then $Ax = x$.

The identity matrix

- The $n \times n$ identity matrix is the matrix $A = [a_{ij}]$ where a_{ij} equals 1 if $i = j$ and otherwise equals 0.
- Example: the 4×4 identity matrix is:

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

- If $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is the identity matrix, then $Ax = x$.
- Example:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Permutation matrices

Permutation matrices

- If a matrix has all entries 0 or 1, with exactly one 1 in each row and exactly one 1 in each column, then it is a *permutation matrix*.

Permutation matrices

- If a matrix has all entries 0 or 1, with exactly one 1 in each row and exactly one 1 in each column, then it is a *permutation matrix*.
- Example:

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

Permutation matrices

- If a matrix has all entries 0 or 1, with exactly one 1 in each row and exactly one 1 in each column, then it is a *permutation matrix*.
- Example:

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

- More generally, a permutation matrix A has the property that Ax is a permutation of the entries of x.

Matrix-matrix multiplication

Matrix-matrix multiplication

- Multiplying matrix $A \in \mathbb{R}^{m \times n}$ by matrix $B \in \mathbb{R}^{n \times p}$ is like multiplying A by each column vector of B and combining the resulting columns.

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}
\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} =
$$

$$
\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{bmatrix}
$$

Matrix-matrix multiplication

- Multiplying matrix $A \in \mathbb{R}^{m \times n}$ by matrix $B \in \mathbb{R}^{n \times p}$ is like multiplying A by each column vector of B and combining the resulting columns.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + a_{12}b_{2p} + \cdots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1} & \cdots & a_{21}b_{1p} + a_{22}b_{2p} + \cdots + a_{2n}b_{np} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1p} + a_{m2}b_{2p} + \cdots + a_{mn}b_{np} \end{bmatrix}$$

- The resulting product AB is a matrix in $\mathbb{R}^{m \times p}$.

Associativity

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- (We take this property for granted with multiplication of reals.)

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- (We take this property for granted with multiplication of reals.)
- Proof for $A \in \mathbb{R}^{1\times m}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times 1}$.

$$AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} =$$

$$\begin{bmatrix} a_1 b_{11} + a_2 b_{21} + \cdots + a_m b_{m1} & \cdots & a_1 b_{1n} + a_2 b_{2n} + \cdots + a_m b_{mn} \end{bmatrix},$$

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- (We take this property for granted with multiplication of reals.)
- Proof for $A \in \mathbb{R}^{1\times m}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times 1}$.

$$AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} =$$

$$\begin{bmatrix} a_1 b_{11} + a_2 b_{21} + \cdots + a_m b_{m1} & \cdots & a_1 b_{1n} + a_2 b_{2n} + \cdots + a_m b_{mn} \end{bmatrix},$$

- So

$$\begin{aligned} (AB)C &= (a_1 b_{11} + a_2 b_{21} + \cdots + a_m b_{m1})\, c_1 + \cdots \\ &\quad (a_1 b_{1n} + a_2 b_{2n} + \cdots + a_m b_{mn})\, c_n \\ &= \sum_{1\leq i\leq m, 1\leq j\leq n} a_i b_{ij} c_j. \end{aligned}$$

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- (We take this property for granted with multiplication of reals.)
- Proof for $A \in \mathbb{R}^{1\times m}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times 1}$.

$$BC = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} b_{11}c_1 + \cdots + b_{1n}c_n \\ b_{21}c_1 + \cdots + b_{2n}c_n \\ \vdots \\ b_{m1}c_1 + \cdots + b_{mn}c_n \end{bmatrix},$$

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- (We take this property for granted with multiplication of reals.)
- Proof for $A \in \mathbb{R}^{1\times m}$, $B \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{n\times 1}$.

$$BC = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} b_{11}c_1 + \cdots + b_{1n}c_n \\ b_{21}c_1 + \cdots + b_{2n}c_n \\ \vdots \\ b_{m1}c_1 + \cdots + b_{mn}c_n \end{bmatrix},$$

- So

$$\begin{aligned} A(BC) &= a_1\left(b_{11}c_1 + b_{12}c_2 + \cdots + b_{1n}c_n\right) + \cdots \\ &\quad a_m\left(b_{m1}c_1 + b_{m2}c_2 + \cdots + b_{mn}c_n\right) \\ &= \sum_{1\le i\le m, 1\le j\le n} a_i b_{ij} c_j = (AB)C. \end{aligned}$$

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- Example: If A and B are permutation matrices, then Bx is a permutation of x, so $A(Bx)$ is too.

Associativity

- The matrix product is *associative*, that is $(AB)C = A(BC)$.
- Example: If A and B are permutation matrices, then Bx is a permutation of x, so $A(Bx)$ is too.
- Since $A(Bx) = (AB)x$, the matrix AB is also a permutation matrix.

Commutativity?

Commutativity?

- Is it true that $AB = BA$? (This property is called *commutativity*.)

Commutativity?

- Is it true that $AB = BA$? (This property is called *commutativity*.)
- No! Matrix multiplication is not commutative.

Commutativity?

- Is it true that $AB = BA$? (This property is called *commutativity*.)
- No! Matrix multiplication is not commutative.
- Example:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Commutativity?

- Is it true that $AB = BA$? (This property is called *commutativity*.)
- No! Matrix multiplication is not commutative.
- Example:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, then AB is defined, but BA isn't even defined if $m \neq p$!

Commutativity?

- Is it true that $AB = BA$? (This property is called *commutativity*.)
- No! Matrix multiplication is not commutative.
- Example:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, then AB is defined, but BA isn't even defined if $m \neq p$!
- If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times m}$, then AB and BA are both defined, but the first is $m \times m$ and the second is $n \times n$.

The transpose of a matrix

The transpose of a matrix

- If $A = [a_{ij}]$, then the *transpose* A^T is the matrix $[a'_{ij}]$ where $a'_{ij} = a_{ji}$.

The transpose of a matrix

- If $A = [a_{ij}]$, then the *transpose* A^T is the matrix $[a'_{ij}]$ where $a'_{ij} = a_{ji}$.
- Example:

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

The transpose of a matrix

- If $A = [a_{ij}]$, then the *transpose* A^T is the matrix $[a'_{ij}]$ where $a'_{ij} = a_{ji}$.
- Example:

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

- Transposes and multiplication: $(AB)^T = B^T A^T$.

The transpose of a matrix

- If $A = [a_{ij}]$, then the *transpose* A^T is the matrix $[a'_{ij}]$ where $a'_{ij} = a_{ji}$.
- Example:

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

- Simple property: $(A^T)^T = A$.
- Transposes and multiplication: $(AB)^T = B^T A^T$.
- Example:

$$\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -3 \\ 4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Matrix products with the identity

Matrix products with the identity

- If $A \in \mathbb{R}^{n \times n}$ is the identity matrix and $x \in \mathbb{R}^n$ is a vector, then we know $Ax = x$.

Matrix products with the identity

- If $A \in \mathbb{R}^{n \times n}$ is the identity matrix and $x \in \mathbb{R}^n$ is a vector, then we know $Ax = x$.
- Therefore, $AB = B$ for any matrix $B \in \mathbb{R}^{n \times p}$.

Matrix products with the identity

- If $A \in \mathbb{R}^{n \times n}$ is the identity matrix and $x \in \mathbb{R}^n$ is a vector, then we know $Ax = x$.
- Therefore, $AB = B$ for any matrix $B \in \mathbb{R}^{n \times p}$.
- If $C \in \mathbb{R}^{m \times n}$, then why is $CA = C$?

Matrix products with the identity

- If $A \in \mathbb{R}^{n \times n}$ is the identity matrix and $x \in \mathbb{R}^n$ is a vector, then we know $Ax = x$.
- Therefore, $AB = B$ for any matrix $B \in \mathbb{R}^{n \times p}$.
- If $C \in \mathbb{R}^{m \times n}$, then why is $CA = C$?
- $(CA)^T = A^T C^T = AC^T = C^T$, so $CA = C$.

Next time!

Linear Algebra II