Lecture #5: Fundamental Operations with Matrices

i.e. permutations, rotations, projections

First let's talk about lengths and angles

<u>def:</u> The <u>length</u> of a vector $v = \begin{bmatrix} v_1 \\ \vdots \\ v_d \end{bmatrix}$ is

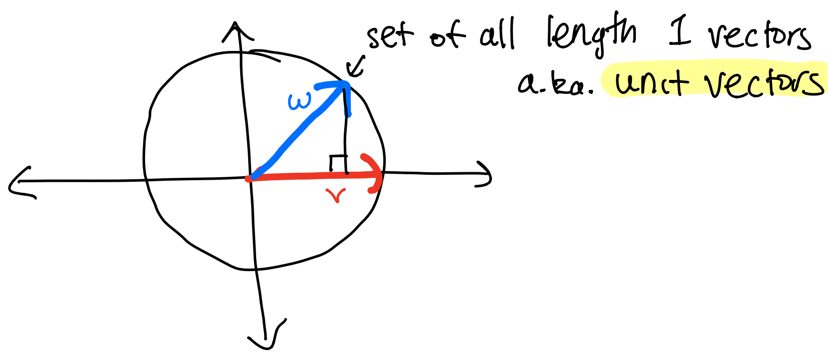
$$\|v\| = \sqrt{v_1^2 + v_2^2 \dots + v_d^2}$$

Now we can use the inner-product to talk about angles

<u>def.</u> The <u>angle</u> θ between two vectors v and w is determined by the expression

$$\cos\theta = \frac{\langle v, w \rangle}{\|v\| \, \|w\|}$$

The way to think about it geometrically is, assuming $\|v\| = \|w\| = 1$, then



If v is on the x-axis then the x-coordinate of w is the cosine of the angle btwn them

Lengths and angles are helpful things to keep track of when we apply various ops.

Now let's investigate an important op.

Q: How do I apply a permutation to the coordinates of a vector?

Formally, a permutation π is a function

$$\pi : \underbrace{[d]}_{\text{integers from 1 to } d} \longrightarrow [d]$$

that is one-to-one, i.e. each element of the image is mapped to exactly once

e.g. $d=3$, $\pi(1)=2$, $\pi(2)=3$, $\pi(3)=1$

Q2: How could I apply this permutation to a 3-d vector using a matrix-vector product?

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ x \\ y \end{bmatrix}$$

The first coordinate becomes the second, etc def: A permutation matrix is a square matrix with 0s and 1s where there is exactly one 1 in each row/column

In particular if I want to construct the permutation matrix corresponding to π I set

$$A_{i,j} = \mathbb{1}_{\{\pi(j) = i\}}$$

indicator function that is 1 if $\pi(j)=i$ and 0 else

Now let's test our intuition:

Q3: Does applying a permutation matrix change the length?

No, e.g. $\left\| \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right\| = \sqrt{x^2+y^2+z^2} = \left\| \begin{bmatrix} z \\ x \\ y \end{bmatrix} \right\|$

Q4: Does it change the angle between vectors?

Also no, can see from definition of angle via the inner-product

Something special about permutation matrices (that is also true ~~of~~ other families of matrices we will see later) is computing their inverse is easy

<u>def:</u> The transpose of an $m \times n$ matrix A, denote A^T is an $n \times m$ matrix with

$$(A^T)_{i,j} = A_{j,i}$$

e.g. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

Now let's take our permutation matrix from earlier:

$$\underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

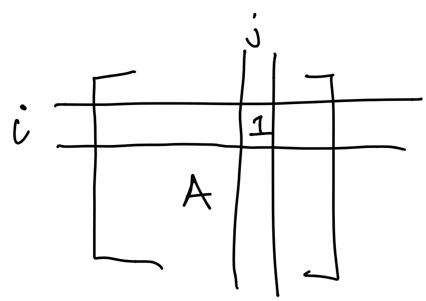
Aha! So $A^T = A^{-1}$ (for permutation matrices)

Actually, this is always true

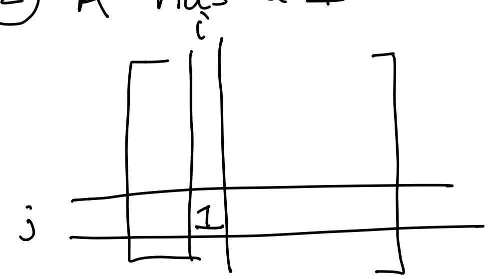
<u>Fact:</u> If A is a permutation matrix, then $A^T = A^{-1}$

Here's the intuition: Suppose A corresponds to π and $\pi(j) = i$ then

(1) multiplying by A sends the j^{th} coordinate to the i^{th} coordinate



(2) A^T has a 1 at the location (j,i):



thus it sends the i^{th} coordinate to the j^{th} coordinate

And so the two operations undo each other

Let's dive deeper into transposition

<u>Fact:</u> $(AB)^T = B^T A^T$

This should remind you of the prelecture check-up

In particular for permutation matrices

$$(AB)^{-1} = (AB)^T$$

$$B^{-1}A^{-1} = B^T A^T$$

Inverses and transpositions both reverse the order of matrix multiplication

Some more basic facts

<u>Fact:</u> $(A^T)^T = A$

<u>Fact:</u> $(A+B)^T = A^T + B^T$

<u>Fact:</u> For square and invertible A,

$$(A^{-1})^T = (A^T)^{-1}$$

Actually, transposes are a helpful way to express many kinds of operations

<u>Q5:</u> How can we express $\langle v, w \rangle$ (column vectors $d \times 1$) through matrix multiplication, using transposes?

$$\underset{1\times d}{v^T}\,\underset{d\times 1}{w} = [v_1\ v_2 \ldots v_d] \begin{bmatrix} w_1 \\ \vdots \\ w_d \end{bmatrix} = w^T v = (v^T w)^T$$

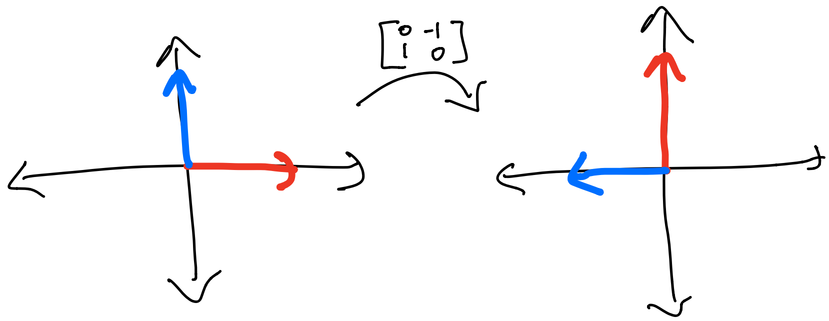
Now let's talk about rotations and projections just in 2-d for now

We'll get to higher-d in a few lectures

Q6: How do I rotate a 2-d vector by 90° counter-clockwise in the plane?

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

First column describes what happens to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
second column describes what happens to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$



More generally, if I want to rotate counter-clockwise by θ I get

$$\underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{A}$$

Q7: Does a rotation preserve lengths?

Yes, but how do we see it?

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$$

$$\text{length}^2 = (x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2$$
$$= x^2\cos^2\theta + y^2\sin^2\theta + x^2\sin^2\theta + y^2\cos^2\theta =$$
$$x^2 + y^2 = \text{old length}^2$$

<u>Application Domain</u>: Robotics, describe camera angle, etc

Oftentimes useful to act on just two coordinates at a time

$$\underbrace{\begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{yaw}} \underbrace{\begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}}_{\text{pitch}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}}_{\text{roll}}$$

<u>Q8</u>: How do we invert a rotation in 2-d?

In 2-d, can rotate by $-\theta$

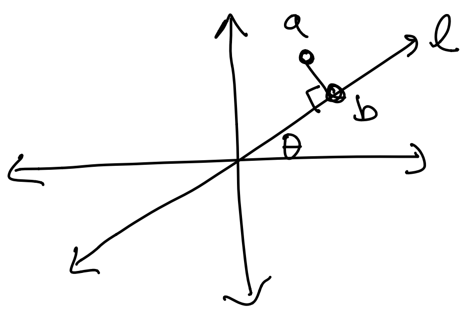
$$\underbrace{\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}}_{A^{-1}} = \underbrace{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}_{A^T}$$

Does this look familiar?

<u>Fact:</u> For any rotation matrix R, its inverse is R^T

We'll need to wait for more definitions to understand what this means in high-d, but it's still true

Finally, let's talk about projections e.g.



$\|a-b\|$

we want to map a to b, the closest point to a on line ℓ

Actually some projections are easier than others

<u>Q9:</u> What is the projection of v onto the x-axis?

Let $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, then its projection is $\begin{bmatrix} v_1 \\ 0 \end{bmatrix} = v'$

It is easy to see that v' is the closest to v along the x-axis (here distance $\triangleq \|v - v'\|$)

Actually this operation can also be expressed as a matrix-vector product:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$$

<u>Q10</u>: What does a general projection look like, like the one onto line ℓ?

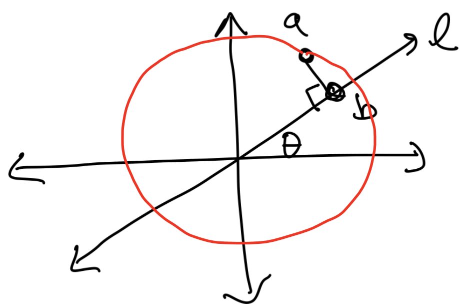
$$R_\theta \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R_{-\theta}$$

It's easy to read off geometric properties from this expression

<u>Q11</u>: Does a projection preserve length?

$$P = \underbrace{R_\theta}_{\text{preserves length}} \overbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}^{\text{decreases length sometimes}} \underbrace{R_{-\theta}}_{\text{preserves length}}$$

Can visualisize this geometrically too:



b is strictly inside the circle we get by rotating a around the origin

<u>Fact:</u> Let A and B be square matrices and suppose A is invertible. Then

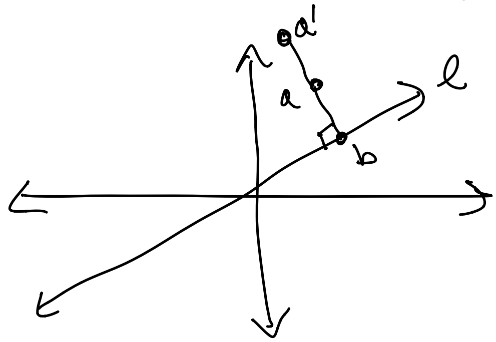
B is invertible iff AB is invertible

<u>Q12:</u> Is a projection onto ℓ invertible?

Again, we can go back to the expression

$$\underset{\substack{\uparrow \\ \text{has an inverse}}}{R_\theta} \underset{\substack{\uparrow \\ \text{doesn't have an inverse}}}{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}} \underset{\substack{\uparrow \\ \text{has an inverse}}}{R_{-\theta}}$$

Let's go back to the geometric picture one last time to check our algebraic reasoning



a and a' both map to b, hence we can't undo the operation uniquely

<u>Application Domain</u>: Projections are essential in data science, and give you a way to map a high-d dataset into low-d so you can visualsize it.

But how do you find a good projection?