Math 196 Lecture 1 10/2/19 ①

- Warmup problem: Solve for x and y:

$$\left| \begin{array}{l} 2x + 4y = 8 \\ 3x + 6y = 4 \end{array} \right|$$

("system of 2 equations in 2 variables")

- Introductions: pair up and present:
  - name (preferred), year
  - where from
  - major, interests
  - fun fact
- Syllabus — Piazza
- Questions?
- What do you think linear algebra is about?
- Back to warmup problem: Why are there no solutions?

$2x + 4y = 8$ is a line in the xy plane.

$$y = -\frac{1}{2}x + 2$$

and $3x + 6y = 4 \rightsquigarrow y = -\frac{1}{2}x + \frac{2}{3}$

These are parallel lines.

- Gauss-Jordan elimination: Example.

$$\left|\begin{array}{l} 3x+6y=3 \\ 2x+3y=1 \end{array}\right| \div 3$$

(divide row 1 by 3).

$$\left|\begin{array}{l} x+2y=1 \\ 2x+3y=1 \end{array}\right| -2(\mathrm{I})$$

(replace row 2 by row 2 − 2×(row 1))

$$\left|\begin{array}{l} x+2y=1 \\ \quad\ -y=-1 \end{array}\right| \div(-1)$$

(divide row 2 by (−1).)

$$\left|\begin{array}{l} x+2y=1 \\ \qquad\ y=1 \end{array}\right| -2(\mathrm{II})$$

$$\left|\begin{array}{l} x \qquad =-1 \\ \qquad\ y=1 \end{array}\right|$$

Done!

We can do all this without writing so much:

$$\begin{bmatrix} 3 & 6 & 3 \\ 2 & 3 & 1 \end{bmatrix} \div 3$$

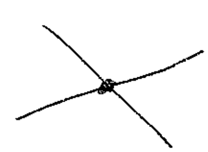
$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} -2(\mathrm{I})$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \end{bmatrix} \div(-1)$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} -2(\mathrm{II})$$

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

interpretation:
2 lines intersect at one point:



A rectangular grid of numbers is called a matrix. The example we just did used 2×3 matrices.

$$\underset{\#\text{ rows}}{2} \times \underset{\#\text{ columns}}{3}$$

- Another example: 3 eq 3 vars (from textbook)

$$\left|\begin{array}{l} 2x + 8y + 4z = 2 \\ 2x + 5y + z = 5 \\ 4x + 10y - z = 1 \end{array}\right|$$

① First, write the system as a 3×4 matrix:

$$\begin{bmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{bmatrix}$$

② Now, use 1st eq to clear a column.

"pivot"

$$\begin{bmatrix} \boxed{2} & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{bmatrix} \div 2$$

$$\begin{bmatrix} 1 & 4 & 2 & 1 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{bmatrix} \begin{array}{l} \\ -2(\mathrm{I}) \\ -4(\mathrm{I}) \end{array}$$

$$\begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -3 & 3 \\ 0 & -6 & -9 & -3 \end{bmatrix}$$

Done with first column!

(2 steps:

1. make the pivot 1.
2. make the remaining entries in the column 0.)

3 Now, use 2nd eq.

pivot

$$\begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & \boxed{-3} & -3 & 3 \\ 0 & -6 & -9 & -3 \end{bmatrix} \begin{matrix} \\ \div(-3) \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & -6 & -9 & -3 \end{bmatrix} \begin{matrix} -4(\mathrm{II}) \\ \\ +6(\mathrm{II}) \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -3 & -9 \end{bmatrix} \begin{matrix} \\ \\ \div(-3) \end{matrix}$$

4 3rd eq.

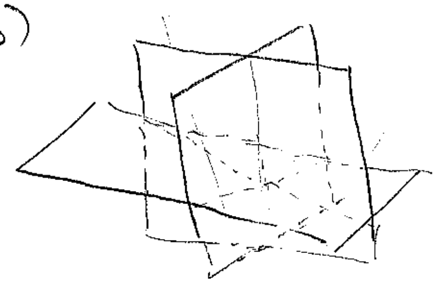
$$\begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{matrix} +2(\mathrm{III}) \\ -(\mathrm{III}). \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 11 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Done!

$x = 11$
$y = -4$
$z = 3.$

interpretation. these 3 planes intersect at $(11, -4, 3)$



At the end, the matrix we have is in "reduced row-echelon form." ie.

a. If a row has nonzero entries, the first nonzero entry is a 1 (called the <u>pivot</u> of this row)

b. If a column contains a pivot, then all other entries in that column are 0.

c. If a row contains a leading 1, then each row above it contains a leading 1 further to the left.

Example from warm-up problem:

$$\begin{bmatrix} 2 & 4 & 8 \\ 3 & 6 & 4 \end{bmatrix} \begin{matrix} \div 2 \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 3 & 6 & 4 \end{bmatrix} \begin{matrix} \\ -3(\mathrm{I}) \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & -8 \end{bmatrix} \begin{matrix} \\ \div(-8) \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{matrix} -4(\mathrm{II}) \\ \end{matrix}$$

$$\underbrace{\begin{bmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}}_{\text{RREF.}}$$

pivots.

The last row says $0 = 1$ so there are no solutions.

[[NOTE: I skipped this example in lecture]]

- Yet another example: (from textbook)

$$\begin{cases} 2x+4y+6z=0 \\ 4x+5y+6z=3 \\ 7x+8y+9z=6 \end{cases}$$

$$\begin{bmatrix} 2 & 4 & 6 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{bmatrix} \div 2$$

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{bmatrix} \begin{array}{l} \\ -4(\mathrm{I}) \\ -7(\mathrm{I}) \end{array}$$

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{bmatrix} \begin{array}{l} \\ \div(-3) \\ \\ \end{array}$$

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & -6 & -12 & 6 \end{bmatrix} \begin{array}{l} -2(\mathrm{II}) \\ \\ +6(\mathrm{II}) \end{array}$$

$$\begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

RREF. Done!

what happened? we ended up with

$$\begin{cases} x-z=2 \\ y+2z=-1 \end{cases}$$

$$\Rightarrow \begin{cases} x = z+2 \\ y=-2z-1 \end{cases}$$

We can choose z. "free var"

E.g.

$z=1$ gives $(x,y,z)=(3,-3,1)$.

$z=7$ gives $(x,y,z)=(9,-15,7)$.

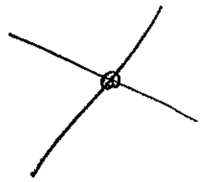
Choose $z=t$ gives

(x,y,z)
$=(t+2,-2t-1,t)$
$=(2,-1,0)$
$+t(1,-2,1)$.

This is a line in $\mathbb{R}^3$.

Geometric interpretations:

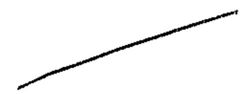
- 2 eq in 2 vars:

1. 

one solution (two lines intersect)

2. 

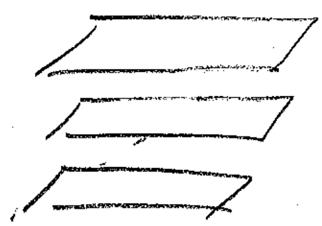
no solutions (two parallel lines)

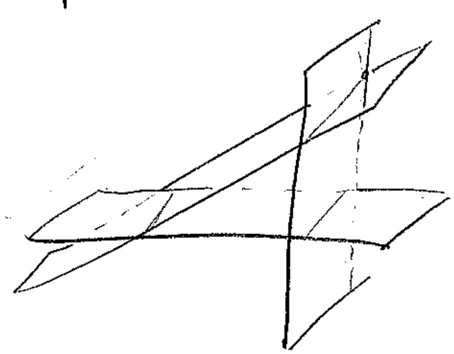
3. 

infinitely many solutions (the two equations are the same line)

- 3 eq in 3 vars.

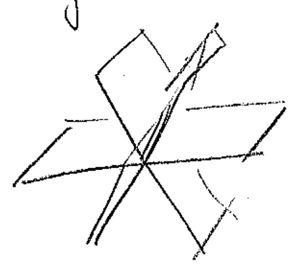
1. one sol'n: 3 planes intersect in one point
2. no sol'ns. For example.





3. infinitely many sol'ns. e.g.

all 3 planes are the same | they intersect in a line



You can have m eq's in n var's.

m = # of "hyperplanes" n = dimension

In higher dimensions, it is harder to picture, but situation is the same.

Any system of linear equations has either:

- 1 solution
- infinitely many solutions

→ system is "consistent"

- no solutions → "inconsistent"

→ this occurs if one of the rows says "zero = nonzero"

e.g. $\begin{bmatrix} 2 & 4 & 8 \\ 3 & 6 & 4 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & -8 \end{bmatrix}$

second row says $0 = -8$.

- one more example (in textbook)

$$\begin{bmatrix} \boxed{2} & 4 & -2 & 2 & 4 & 2 \\ 1 & 2 & -1 & 2 & 0 & 4 \\ 3 & 6 & -2 & 1 & 9 & 1 \\ 5 & 10 & -4 & 5 & 9 & 9 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 2 & -1 & 1 & 2 & 1 \\ 0 & 0 & 0 & \boxed{1} & -2 & 3 \\ 0 & 0 & 1 & -2 & 3 & -2 \\ 0 & 0 & 1 & 0 & -1 & 4 \end{bmatrix}$$

$$\rightsquigarrow \begin{bmatrix} 1 & 2 & -1 & 0 & 4 & -2 \\ 0 & 0 & 0 & 1 & -2 & 3 \\ 0 & 0 & \boxed{1} & 0 & -1 & 4 \\ 0 & 0 & 1 & 0 & -1 & 4 \end{bmatrix}$$

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

But this is not reduced row echelon form.

so at the end: reorder the rows.

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

$$\longrightarrow \quad \begin{aligned} x_1 + 2x_2 + 3x_5 &= 2 \\ x_3 - x_5 &= 4 \\ x_4 - 2x_5 &= 3 \end{aligned}$$

choose $x_5 \longrightarrow$ get x_3, x_4

Then choose $x_2 \longrightarrow$ get x_1

(x_2, x_5 are called "free variables")

The three types of "elementary row operations" used in Gauss-Jordan elimination

1. Divide a row by a nonzero scalar
2. Subtract a multiple row from another row.
3. Swap two rows.

(Note: you may have learned Gauss-Jordan elimination slightly differently. But as long as you stick to those elementary row operations, the end result will always be the same.)

Lecture 2:

Last time: Consider $x=5$.

Is this a line? a plane? It depends.

$\{(x,y) \mid x=5\}$ is a line. (in $\mathbb{R}^2$)

$\{(x,y,z) \mid x=5\}$ is a plane. (in $\mathbb{R}^3$)

$\{x \mid x=5\}$ is a point (in $\mathbb{R}^1$)

$\{(x,y,z,w) \mid x=5\}$ is a "3-dim plane" (in $\mathbb{R}^4$)

So back to

$$\left| \begin{array}{l} x_1 + 2x_2 + 3x_5 = 2 \\ x_3 - x_5 = 4 \\ x_4 - 2x_5 = 3 \end{array} \right|$$

each equation represents a "4D plane" in $\mathbb{R}^5$.

The solutions are everything of the form

$x_1 = 2-2t-3r$
$x_2 = t$
$x_3 = 4+r$
$x_4 = 3+2r$
$x_5 = r$

can write it in set notation

$\{(2-2t-3r,\ t,\ 4+r,\ 3+2r,\ r) \mid t\in\mathbb{R}, r\in\mathbb{R}\}$

Some ideas:

- $x_1 + 2x_2 + 3x_5 = 2$ is a 4-plane in $\mathbb{R}^5$
- If we consider this equation by itself there are 4 free variables: x_2, x_3, x_4, x_5.
- "usually" when you add one equation, you decrease the number of free variables

---

Draw some pictures of lines and planes in $\mathbb{R}^2$, $\mathbb{R}^3$.

Def: The number of nonzero rows (or number of pivots) in rref(A) is called the rank of A.

(we'll see this number again in the future)

Consider (from before):

$$\left|\begin{aligned} 2x+8y+4z&=2 \\ 2x+5y+z&=5 \\ 4x+10y-z&=1 \end{aligned}\right|$$

$\longrightarrow$ so far we've looked at the augmented matrix

$$A = \begin{bmatrix} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{bmatrix}$$

We can also just look at the coefficient matrix $\longrightarrow$ $B = \begin{bmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{bmatrix}$

Last time we calculated:

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 11 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

Q: What is rref(B)?

A: We can just remove the column on the right.

$$\text{rref}(B) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This is because of how Gauss-Jordan elimination works.

Theorem: Consider a linear system of n eq's n vars. Let A be the coefficient matrix. (A is $n \times n$.) Then the system has a unique solution if and only if

$$\text{rref}(A) = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

(Note: • "Identity matrix."

• we don't care about the values on the RHS of the equations.)

What does "if and only if" mean?

Now consider n eq's m var's.

Let A be coeff matrix. $(n \times m)$.

$$m = \begin{pmatrix} \text{total \# of} \\ \text{variables} \end{pmatrix} = \begin{pmatrix} \text{\# of free} \\ \text{variables} \end{pmatrix} + \underbrace{\begin{pmatrix} \text{\# of} \\ \text{pivots} \end{pmatrix}}_{\text{rank}(A)}.$$

This gives the relationship between (# of free variables) and rank(A) and (# of variables).

Another geometric interpretation of a system of equations.

A column vector is a matrix with 1 column.
A row vector - - - - - - row.

e.g.
$$\begin{bmatrix} 1 \\ 2 \\ 9 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 5 & 5 & 3 & 7 \end{bmatrix}$$

4×1 matrix column vec. $\qquad$ 1×5 matrix row vector.

We can use a $n \times 1$ column vector to represent a point of $\mathbb{R}^n$.

Basic operation on vectors:

① Vector addition: e.g.

$$\vec{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \qquad \vec{w} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\longrightarrow \quad \vec{v} + \vec{w} = \begin{bmatrix} 1+2 \\ 4-1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

(just add entry by entry).

② Scalar multiplication e.g.

$$3 \cdot \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \cdot 2 \\ 3 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}.$$

③ Dot product

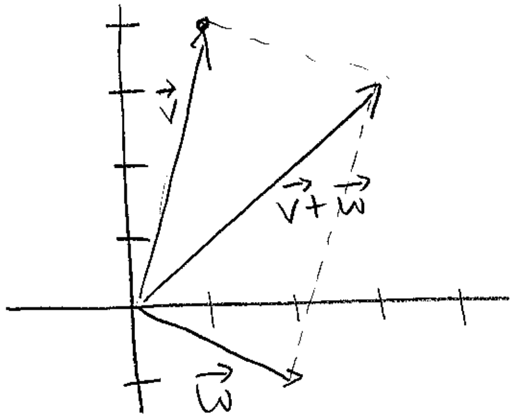
$\vec{v}, \vec{w}$ as above

$$\vec{v} \cdot \vec{w} = (1 \cdot 2) + (4 \cdot (-1)) = -2.$$

(Why don't we define vector multiplication entry by entry? It's not very useful.)

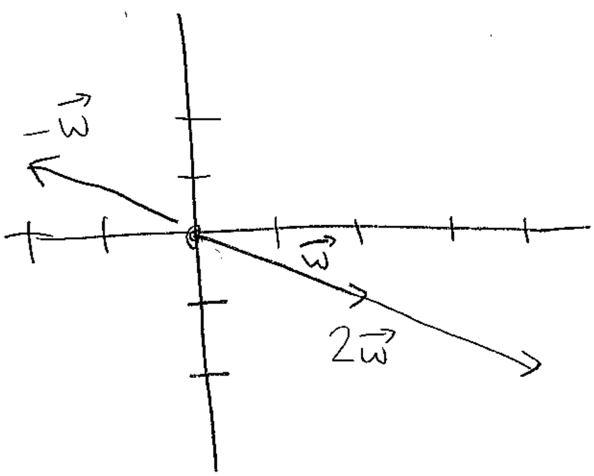
Geometric interpretations:

1. vector addition.



"go along $\vec{v}$, then go along $\vec{w}$"

2. scalar multiplication



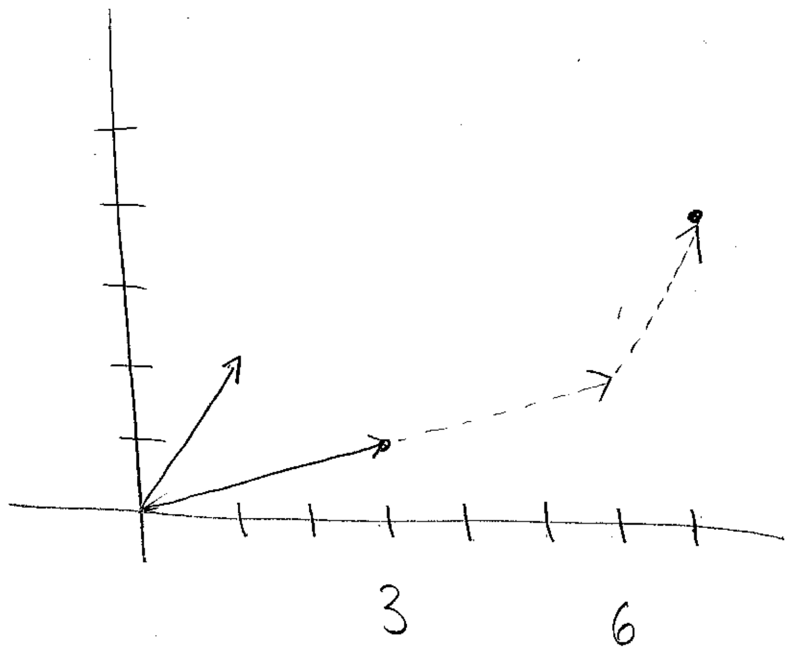
"just rescaling the vector"

3. dot product — we'll come back to this later...

Consider $\left| \begin{array}{l} 3x_1 + x_2 = 7 \\ x_1 + 2x_2 = 4 \end{array} \right|$

We can write this as

$$x_1\begin{bmatrix}3\\1\end{bmatrix} + x_2\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}7\\4\end{bmatrix}.$$



we want to find some multiples of $\begin{bmatrix}3\\1\end{bmatrix}$ and $\begin{bmatrix}1\\2\end{bmatrix}$ which combine to make $\begin{bmatrix}7\\4\end{bmatrix}$.

Solution is $x_1 = 2$, $x_2 = 1$ i.e.

$$\underbrace{2\cdot\begin{bmatrix}3\\1\end{bmatrix} + 1\cdot\begin{bmatrix}1\\2\end{bmatrix}} = \begin{bmatrix}7\\4\end{bmatrix}.$$

this is called a "linear combination" of $\begin{bmatrix}3\\1\end{bmatrix}$ and $\begin{bmatrix}1\\2\end{bmatrix}$.

Lec #3 starts here

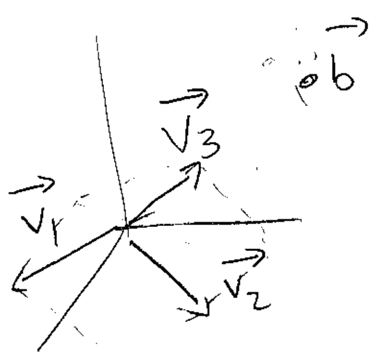
In general: n eqs in m vars:
can be written as:

$$x_1 \begin{bmatrix} | \\ \vec{v}_1 \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ \vec{v}_2 \\ | \end{bmatrix} + \dots + x_m \begin{bmatrix} | \\ \vec{v}_m \\ | \end{bmatrix} = \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix}$$

$\vec{v}_1 \in \mathbb{R}^n$ $\vec{v}_2 \in \mathbb{R}^n$ etc $\vec{b} \in \mathbb{R}^n$.

<u>Note</u>. In this interpretation, the dimension is the <u>number of eqs</u> NOT the number of vars. (# vars = # vectors)

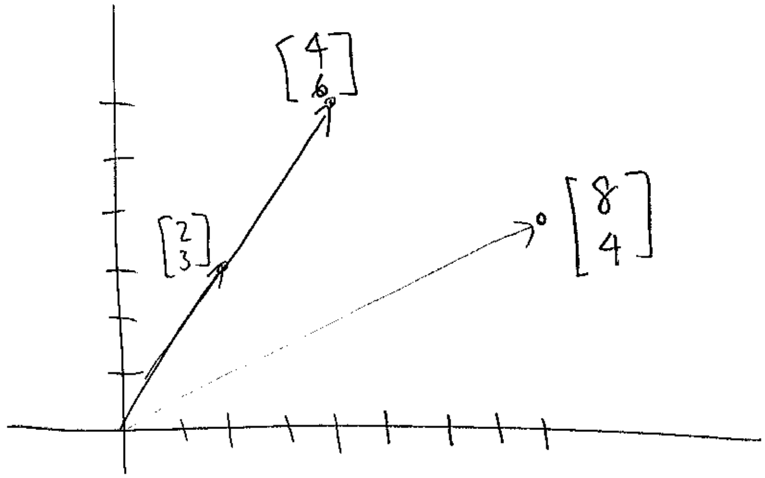
3 eqs in 3 vars.



Q: When are there no solutions?

Q: When are there infinitely many solutions?

consider $\begin{vmatrix} 2x+4y=8 \\ 3x+6y=4 \end{vmatrix} \rightsquigarrow x\begin{bmatrix} 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$



Problem:

All linear combinations of $\begin{bmatrix}2\\3\end{bmatrix}$ and $\begin{bmatrix}4\\6\end{bmatrix}$ lie inside a line in $\mathbb{R}^2$. $\begin{bmatrix}8\\4\end{bmatrix}$ is not in this line.

$\Longrightarrow$ no solution.

Next

$$x\begin{bmatrix}2\\3\end{bmatrix} + y\begin{bmatrix}4\\6\end{bmatrix} = \begin{bmatrix}4\\6\end{bmatrix}.$$

There are many ways to get to $\begin{bmatrix}4\\6\end{bmatrix}$.

e.g. $x=0 \quad y=1$.

or $x=2 \quad y=0$

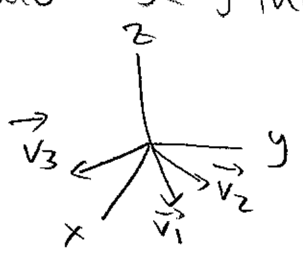
In fact, any choice of y works.

---

3 eqs in 3 vars. What could go wrong?

$$x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{b}$$

All linear combinations of $\vec{v}_1$ $\vec{v}_2$ $\vec{v}_3$ could be stuck inside a plane in $\mathbb{R}^3$

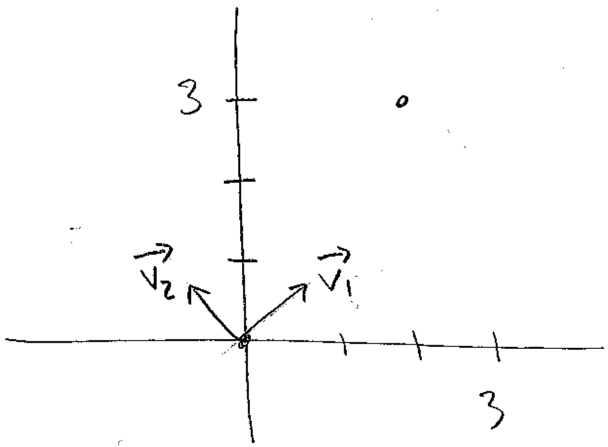


(e.g. $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are all inside the xy-plane).

An example of trying to solve

$$x_1 \vec{v_1} + x_2 \vec{v_2} = \vec{b}$$

Maybe you want to use $\vec{v_1}, \vec{v_2}$ as your coordinate axes instead of $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$.



$$\begin{bmatrix}2\\3\end{bmatrix} = \frac{5}{2}\sqrt{2}\,\vec{v_1} + \frac{1}{2}\sqrt{2}\,\vec{v_2}$$

$$\approx 3.54\,\vec{v_1} + 0.71\,\vec{v_2}$$

$$\vec{v_1} = \begin{bmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix} = \text{rotate } \begin{bmatrix}1\\0\end{bmatrix} \text{ by } 45^\circ \text{ clockwise}$$

$$\vec{v_2} = \begin{bmatrix}-\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix} = \text{rotate } \begin{bmatrix}0\\1\end{bmatrix} \text{ by } 45^\circ \text{ clockwise.}$$

e.g. if you want to rotate an image by 45°.

Lecture 3

Start on pg 18.

- Review (quickly) geometric interp. as intersection of planes.

Note: planes go on forever!

- Finish discussion of geom. interp with vectors.

Yet another interpretation: use matrix multiplication (n equations, m variables).

$$\left| \begin{array}{l} 2x_1 - 3x_2 + 5x_3 = 7 \\ 9x_1 + 4x_2 - 6x_3 = 8 \end{array} \right|$$

a column vector.

$$A = \begin{bmatrix} 2 & -3 & 5 \\ 9 & 4 & -6 \end{bmatrix} \qquad b = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

coefficient matrix for the system

The augmented matrix for the system is $[A \,\vdots\, b]$

The matrix form of this system is

$$\begin{bmatrix} 2 & -3 & 5 \\ 9 & 4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}.$$

or $\boxed{A\vec{x} = \vec{b}}$ where $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$

We define $A\vec{x}$ as follows:

if $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & & \cdots & a_{nm} \end{bmatrix}$ $\quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$.

$$A\vec{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m \end{bmatrix}$$

A is $n \times m$, $\vec{x}$ is $m \times 1$

Two ways to think about this:

① If $\vec{w_1}, \ldots, \vec{w_n}$ are the rows of A

(i.e. $A = \begin{bmatrix} \text{---} & \vec{w_1} & \text{---} \\ \text{---} & \vec{w_2} & \text{---} \\ & \vdots & \\ \text{---} & \vec{w_n} & \text{---} \end{bmatrix}$)

then $A\vec{x} = \begin{bmatrix} \vec{w}_1 \cdot \vec{x} \\ \vec{w}_2 \cdot \vec{x} \\ \vdots \\ \vec{w}_n \cdot \vec{x} \end{bmatrix}$ so you're taking dot products of $\vec{x}$ with each row of A.

(2) If $\vec{v}_1, \ldots, \vec{v}_m$ are the columns of A,

$\left(\text{i.e. } A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix} \right)$ then

$$A\vec{x} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_m\vec{v}_m.$$

So $A\vec{x}$ is a <u>linear combination</u> of $\vec{v}_1, \ldots, \vec{v}_m$.

The two ways to think about $A\vec{x}$ corresp. to the two geometric interpretations of systems of eq'ns. (In one way, you look at columns of A. In the other, you look at rows.)

If A is an $n \times m$ matrix, we can think of it also as a function.

Input: $\vec{x}$ (a vector in $\mathbb{R}^m$)

Output: $A\vec{x}$ (a vector in $\mathbb{R}^n$).

More generally, a <u>function T from X to Y</u> is:

input $\qquad\qquad\qquad\qquad$ output

$$x \rightsquigarrow \boxed{\ T\ } \rightsquigarrow y$$

$(x \in X) \qquad\qquad\qquad\qquad (y \in Y)$.

X is the <u>domain</u>. Y is the <u>target space</u>

If the input is x, the output is denoted $T(x)$.

e.g. ① $T: \mathbb{R} \longrightarrow \mathbb{R}$

$T(x) = x^2$

② $T: \mathbb{R}^2 \longrightarrow \mathbb{R}$.

$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \sqrt{x_1^2 + x_2^2}$ (distance from origin)

③ If A is a $n \times m$ matrix, then $T(\vec{x}) = A\vec{x}$ defines a function from $\mathbb{R}^m$ to $\mathbb{R}^n$.

So every $n \times m$ matrix gives you a function from $\mathbb{R}^m$ to $\mathbb{R}^n$.

Lecture #4 starts here!

Def: A fn T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is called a linear transformation is there exists an $n \times m$ matrix A such that

$$T(\vec{x}) = A\vec{x} \quad (\text{for all } \vec{x} \in \mathbb{R}^m).$$

$$\underbrace{\begin{bmatrix} 2 & -3 & 5 \\ 9 & 4 & -6 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}.$$

Interpretation #3: which inputs $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ will give the output of $\begin{bmatrix} 7 \\ 8 \end{bmatrix}$?

We need to "invert" this linear transformation.

Q: $T: \mathbb{R}^n \to \mathbb{R}^n$ $\quad T(\vec{x}) = \vec{x}$.

Is this a linear transformation?

Yes: Let $I_n = \underbrace{\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}_{n} \Big\} n$ ← "identity matrix"

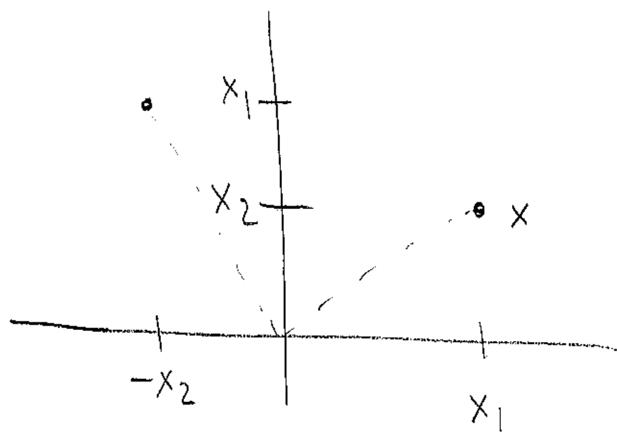
Then $T(\vec{x}) = I_n \vec{x}$.

$$\left(\text{e.g. } \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right).$$

Example: $T : \mathbb{R}^2 \to \mathbb{R}^2$

$$T(\vec{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{x}$$

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}.$$



T represents 90° rotation counterclockwise wrt the origin.

So: 90° counter-c.w. rot wrt origin is a linear transformation.

Q: $T : \mathbb{R}^2 \to \mathbb{R}^2$. what is something that is not a linear transformation? (In other words, it is not given by matrix mult.)

A: consider. e.g. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. (nonzero const fn).

or $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+1 \\ y \end{bmatrix}$ translation to left by 1.

How do we know these are not linear transfs?

One strategy: find some property that all lin. transf. satisfy. Show these do not satisfy them.

For example. If T is a lin transf, then $T\vec{0} = \vec{0}$ ($\vec{0}$ is zero-vector).

<u>Theorem</u>: Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a lin. transf. Then

a. $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w})$ for all $\vec{v}, \vec{w} \in \mathbb{R}^m$

b. $T(k\vec{v}) = k\,T(\vec{v})$ for all $\vec{v} \in \mathbb{R}^m$, $k \in \mathbb{R}$.

<u>Proof</u>: matrix mult satisfies

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

$$A(k\vec{v}) = k\,(A\vec{v})$$

(You can check this yourself) $\square$

Note: This is really powerful.

For example if $T: \mathbb{R}^2 \to \mathbb{R}^n$ is a lin transf (w/ matrix A, $n \times 2$).

we know $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

so $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + x_2 T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$.

Hence: If we know $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$, we know $T\vec{x}$ for all $\vec{x} \in \mathbb{R}^2$!

e.g. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ is the matrix for T.

Then $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$.

$T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$.

In general. $T(\vec{e_i})$ is just the ith column of A, where $\vec{e_i} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{th}$.

So: If someone tells us to find $T: \mathbb{R}^m \to \mathbb{R}^n$ and we know

1. that T is a lin transf
2. the values of $T(\vec{e_1}), \ldots, T(\vec{e_m})$

then we know what the function T is!

In fact $T(\vec{x}) = A\vec{x}$, where

$$A = \begin{bmatrix} | & & | \\ T(\vec{e_1}) & \cdots & T(\vec{e_m}) \\ | & & | \end{bmatrix}.$$

So, two ways of thinking about lin. transf.

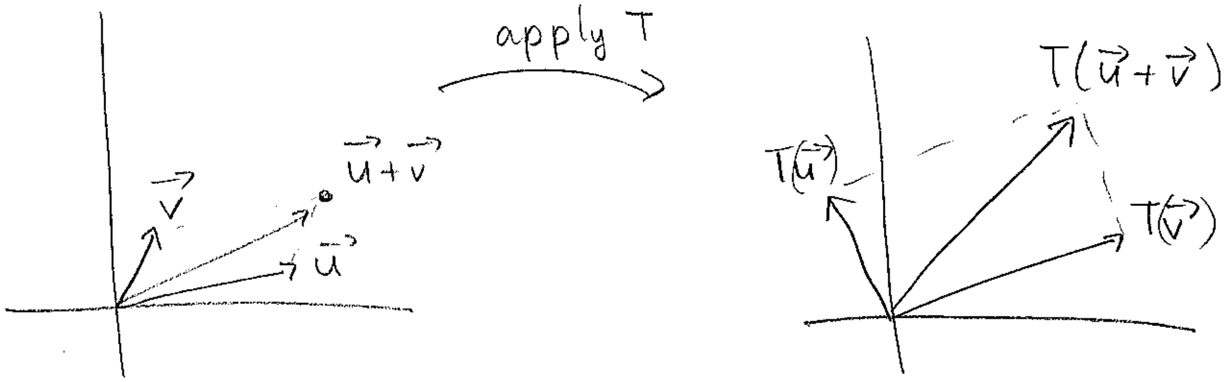
1. T is a lin transf iff there is a matrix A s.t. $T(\vec{x}) = A\vec{x}$.] concrete
2. T is a lin transf iff the following hold

$$\text{abstract}\begin{cases} ① \ T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w}) & \text{for all } \vec{v}, \vec{w} \in \mathbb{R}^m \\ ② \ T(k\vec{v}) = k\,T(\vec{v}) & \text{for all } \vec{v} \in \mathbb{R}^m,\ k \in \mathbb{R}. \end{cases}$$

~~Lecture 4~~: ← (Not anymore! Lecture 5 starts here!)

- Start on page 25.

- Draw a picture to illustrate $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$:



(In this example, $T: \mathbb{R}^2 \to \mathbb{R}^2$)

- Why are these called "linear transformations"?

- In grade school, you learned about linear functions $f: \mathbb{R} \to \mathbb{R}$.

$$f(x) = \underbrace{a}_{\text{"slope"}}x + \underbrace{b}_{\text{"y-intercept"}}$$

- Q: what are the linear transformations $T: \mathbb{R} \to \mathbb{R}$?

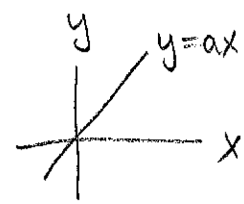
$T(\vec{x}) = A\vec{x}$ where $\vec{x} \in \mathbb{R}^1$ (1×1 col vec)
A is a 1×1 matrix

So $\vec{x}$ and A are really just scalars!

And $A\vec{x}$ is just normal multiplication of scalars.

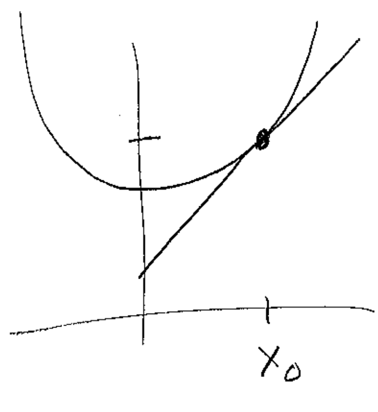
So really, $T(x)$ is a function of the form $\boxed{T(x) = ax.}$ These are all possible linear transformations from $\mathbb{R}$ to $\mathbb{R}$.

(Remark: In linear algebra terminology, $T(\vec{x}) = A\vec{x} + \vec{b}$ is called an "affine transformation".)



- $\left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; y = T(x) \right\}$ is a line in $\mathbb{R}^2$.
  ($T: \mathbb{R} \to \mathbb{R}$ a lin. transf.)
- $\left\{ \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \in \mathbb{R}^{m+n} \;\middle|\; \vec{y} = T(\vec{x}) \right\}$ is a m-dim plane in $\mathbb{R}^{m+n}$.
  ($T: \mathbb{R}^m \to \mathbb{R}^n$ a linear transf.)

Side remark: In calculus, the derivative



gives you a "linear approximation" to a function near a point.

$$f(x_0+h) - f(x_0) \approx f'(x_0)h$$

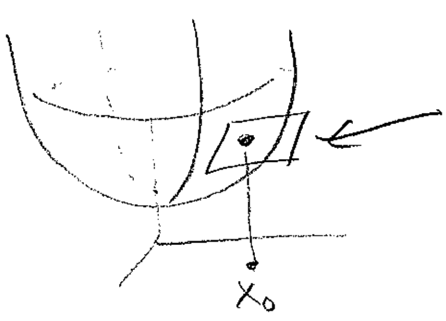
when h is small.

When f is $\mathbb{R} \to \mathbb{R}$, then $f'(x_0)$ is a scalar. The lin transf $T(\vec{v}) = f'(x_0)\vec{v}$ gives the best linear approx near x_0.

When f is $\mathbb{R}^m \to \mathbb{R}^n$, we can still define the derivative as a lin transf.

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) \approx f'(\vec{x}_0)\,\vec{h}.$$

The equation is <u>exactly the same</u>, but now $f'(\vec{x}_0)$ is a $n \times m$ matrix (ie. a lin transf $\mathbb{R}^m \to \mathbb{R}^n$).



the tangent plane is a translated copy of

$$\left\{ \begin{bmatrix} \vec{x} \\ \vec{y} \end{bmatrix} \;\middle|\; \vec{y} = f'(x_0)\,\vec{x} \right\}$$

Let's consider some functions $\mathbb{R}^2 \to \mathbb{R}$ and see if they are linear.

Given a function $T: \mathbb{R}^2 \to \mathbb{R}^2$ there are two ways to check if it's linear

① (concrete) Find a 2×2 matrix A s.t. $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^2$.

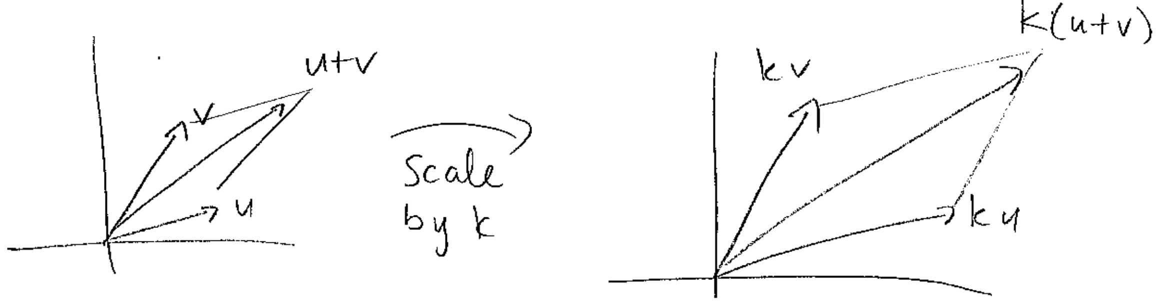
② (abstract) Check that

$$\begin{cases} T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) & \text{for all } \vec{u},\vec{v} \in \mathbb{R}^2 \\ T(l\vec{u}) = l\,T(\vec{u}) & \text{for all } \vec{u} \in \mathbb{R}^2,\ l \in \mathbb{R}. \end{cases}$$

1. Scaling by k: $T(\vec{x}) = k\vec{x}$.

(This was covered in lecture 4.)

Scaling does in fact satisfy the conditions in ②. e.g.



What is the matrix A?

The first col of A is $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}k\\0\end{bmatrix}$

.. second ... is $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\k\end{bmatrix}$.

So $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}k & 0\\0 & k\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

2. Constant function. $T(\vec{x}) = \vec{b}$.

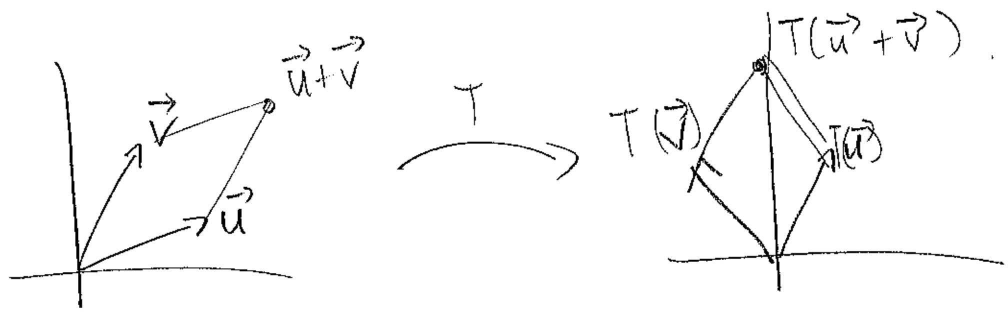
If $\vec{b} = \vec{0}$, it's linear. $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

If $\vec{b} \neq \vec{0}$, it's not since $T\left(\begin{bmatrix}0\\0\end{bmatrix}\right) \neq \begin{bmatrix}0\\0\end{bmatrix}$.

3. Rotation by angle θ. (counter clockwise).

$T(\vec{x})$ = (rotate $\vec{x}$ c.c.w about origin by angle θ).

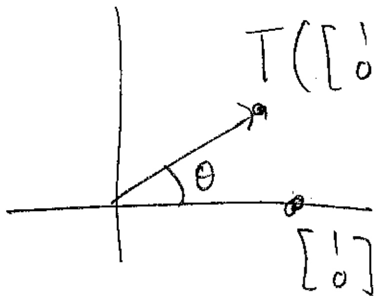
linear!

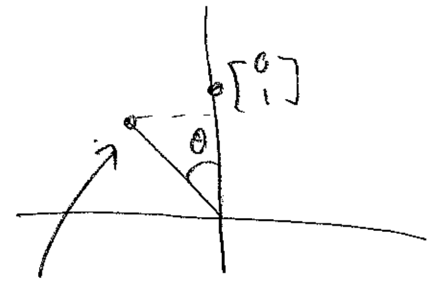


What's the matrix?

$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$

$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}$.

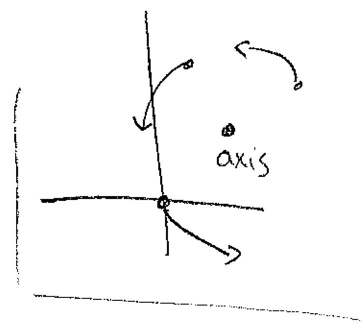




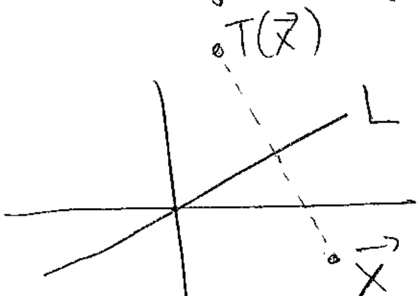
$T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$.

So $$T\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

If you rotate around a pt other than $\vec{0}$, the function is not linear.



4. reflection about a line through origin

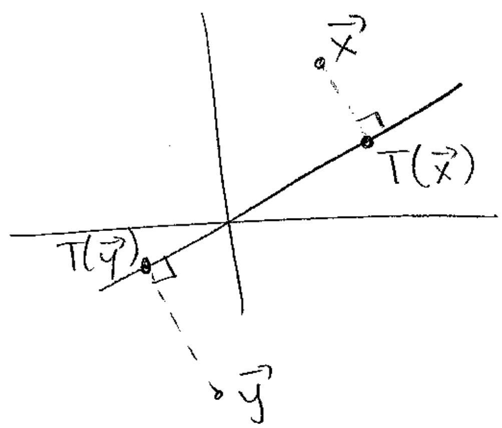


It's linear! (convince yourself geometrically that it is).

What is the matrix? We'll come back to that (a lot) later.

If L does not go through origin, then not linear $T(\vec{0}) \neq \vec{0}$.

5. orthogonal projection onto line L.



If L goes through origin, linear (think geometrically)

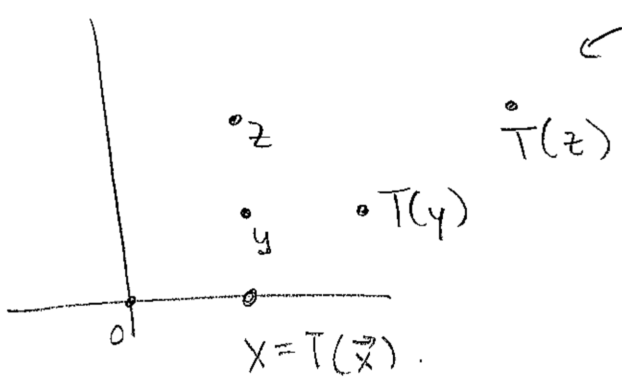
If not through origin, then not linear.

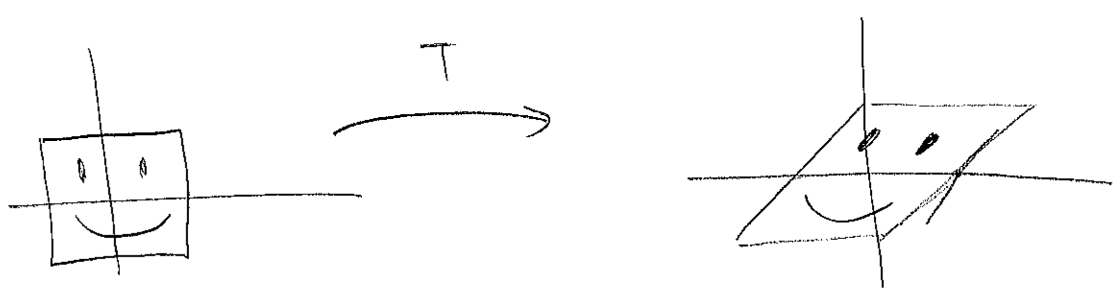
Another class of lin. transf.: shears.

A horizontal shear is

$$T(\vec{x}) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + kx_2 \\ x_2 \end{bmatrix}.$$

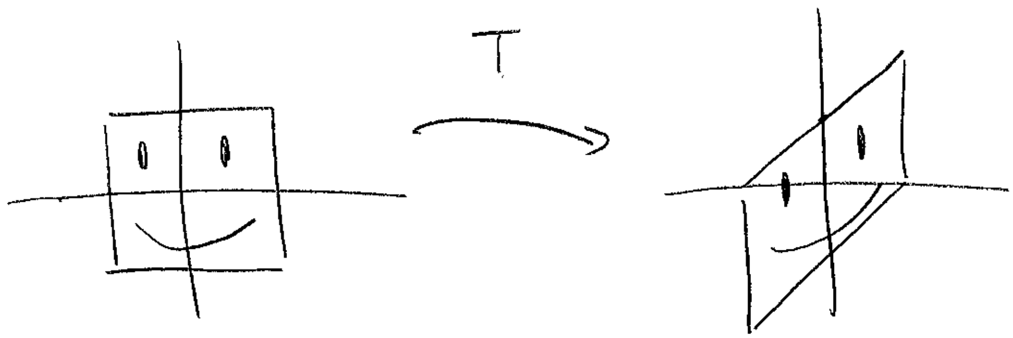
e.g. with $k=1$.





larger $k \longrightarrow$ stretched more.

A vertical shear:



$$T(\vec{x}) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 + kx_1 \end{bmatrix}.$$

Review:

1. A linear transformation is a function $T: \mathbb{R}^m \to \mathbb{R}^n$ such that there is an $n \times m$ matrix A such that

$$T(\vec{x}) = A\vec{x}.$$

2. Theorem 2.1.3.: $T: \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation if and only if
   a. $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$ for all $\vec{x}, \vec{y} \in \mathbb{R}^m$
   b. $T(k\vec{x}) = kT(\vec{x})$ for all $\vec{x} \in \mathbb{R}^m$ $k \in \mathbb{R}$.

3. Theorem 2.1.2: Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then the matrix A for T is

$$A = \begin{bmatrix} | & & | \\ T(\vec{e}_1) & \cdots & T(\vec{e}_m) \\ | & & | \end{bmatrix} \quad \text{where } \vec{e}_i = \begin{array}{c} \\ \\ \end{array} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \leftarrow i^{th} \\ \\ \end{array}$$

We can use these theorems to find formulas for some geometric transformations.

→ Go back to page 30.

Composition of functions

$f(x) = x+1$ "add 1"

$g(x) = x^2$ "square"

$$\begin{tikzcd} 3 \arrow[r] \arrow[rr, bend right, "g \circ f"'] & f(3)=4 \arrow[r] & g(4)=16 \end{tikzcd}$$

$(g \circ f)(3) = 16$

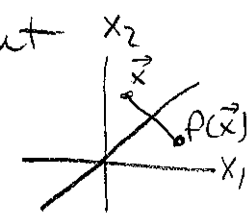
$$(g \circ f)(x) = g(f(x)) = g(x+1) = (x+1)^2$$

$$(f \circ g)(x) = f(g(x)) = f(x^2) = x^2+1$$

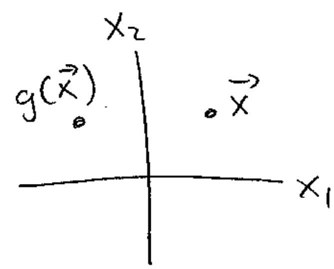
IMPORTANT!
Function composition is not commutative!

Another example: $f: \mathbb{R}^2 \to \mathbb{R}^2$ $\quad g: \mathbb{R}^2 \to \mathbb{R}^2$

$$f(\vec{x}) = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$ ← reflection about diagonal line $x_2 = x_1$

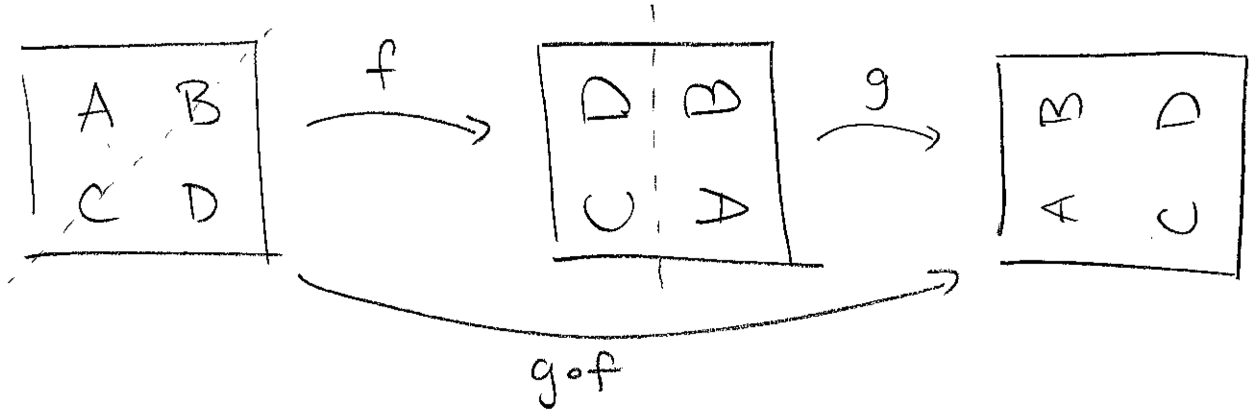


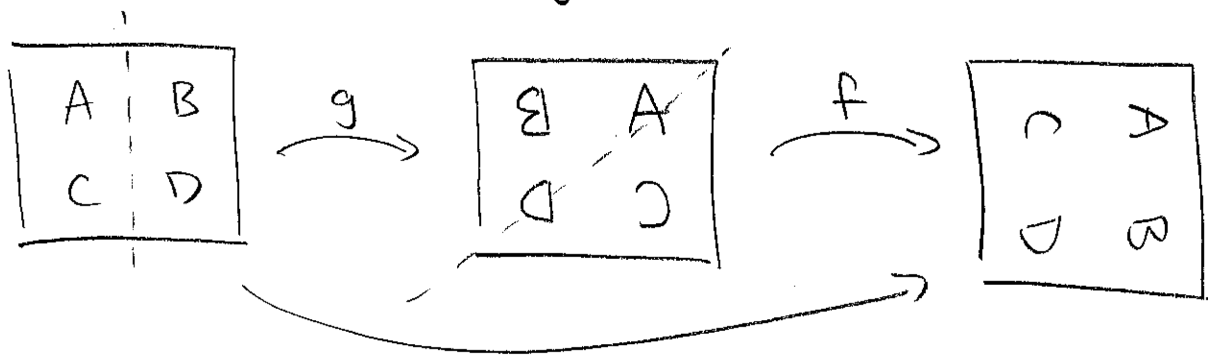
$$g(\vec{x}) = \underbrace{\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}}_{B} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$ ← reflection about vertical line $x_1 = 0$.



Q: What are $g \circ f$ and $f \circ g$?

Take a sheet of paper and try!





$f \circ g$.

90° ccw rotation

$$(g \circ f)(\vec{x}) = g(A\vec{x}) = B(A\vec{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{x}$$

90° cw rotation

$$(f \circ g)(\vec{x}) = f(g(\vec{x})) = A(B\vec{x}) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \vec{x}$$

We define

$$BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad AB = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

In general, $f: \mathbb{R}^m \to \mathbb{R}^p$, $g: \mathbb{R}^p \to \mathbb{R}^n$
linear transformations

$$f(\vec{x}) = A\vec{x} \qquad g(\vec{y}) = B\vec{y}.$$

Then $g \circ f$ is a linear transformation
$(\mathbb{R}^m \to \mathbb{R}^n)$ and BA is
<u>by definition</u> the $n \times m$ matrix for $g \circ f$.

Caution: If $f: \mathbb{R}^m \to \mathbb{R}^q$ $g: \mathbb{R}^p \to \mathbb{R}^n$
with $p \neq q$ then $g \circ f$ is <u>not defined</u>
(So BA is <u>not defined</u>.)

<u>Q1</u>: Why is $g \circ f$ a lin. transf.?

$$(g \circ f)(\vec{x} + \vec{y}) = g(f(\vec{x} + \vec{y})) = g(f(\vec{x}) + f(\vec{y}))$$
$$= g(f(\vec{x})) + g(f(\vec{y})).$$

$$(g \circ f)(k\vec{x}) = g(f(k\vec{x})) = g(k f(\vec{x}))$$
$$= k g(f(\vec{x})).$$

<u>Q2</u>: What exactly is the matrix BA?

Let $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$.

Lec. 6 starts here!

The i^{th} column of BA is

$$(g \circ f)(\vec{e}_i) = g(f(\vec{e}_i)) = B(A\vec{e}_i)$$
$$= B\vec{v}_i$$

<u>So:</u>
$$\boxed{BA = \begin{bmatrix} | & & | \\ B\vec{v}_1 & \cdots & B\vec{v}_m \\ | & & | \end{bmatrix}}$$

e.g.
$$\underbrace{\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}}_{B} \begin{array}{cc} \vec{v}_1 & \vec{v}_2 \\ \multicolumn{2}{c}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \end{array} = \begin{bmatrix} B\vec{v}_1 & B\vec{v}_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

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

Matrix multiplication is <u>not</u> commutative.

let $f: \mathbb{R}^m \to \mathbb{R}^q$ $\quad g: \mathbb{R}^p \to \mathbb{R}^n$

$f(\vec{x}) = A\vec{x}$ $\quad g(\vec{y}) = B\vec{y}$

(A is $q \times m$) $\quad$ (B is $n \times p$)

Composition : $(g \circ f)(\vec{x}) = g(f(\vec{x}))$

Addition : $(f+g)(\vec{x}) = f(\vec{x}) + g(\vec{x})$.

Q1: When is addition defined?

Both f, g need to have $\overbrace{\text{same domain}}$ and $\underbrace{\text{same target space}}$

or else one of $f(\vec{x})$, $g(\vec{x})$ doesn't make sense

or else you can't add the two outputs.

so we need $m = p$, $q = n$.

(A and B have the same dimensions).

$f+g$ is a lin. transf. $(\mathbb{R}^m \to \mathbb{R}^n)$.

$\Longrightarrow$ $(f+g)(\vec{x}) = C\vec{x}$. for some matrix C.

What is C?

$$\vec{e_i} = \begin{array}{c} \\ \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \end{array} \leftarrow i\text{th} \quad \in \mathbb{R}^m$$

The ith column of C is $C\vec{e_i}$.

$$C\vec{e_i} = (f+g)(\vec{e_i}) = f(\vec{e_i}) + g(\vec{e_i})$$

$$= \underbrace{A\vec{e_i}}_{i\text{th col of } A} + \underbrace{B\vec{e_i}}_{i\text{th col of } B}.$$

so we just add A and B entrywise

we define A+B this way.

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}.$$

<u>So:</u> A+B is defined as the entrywise sum because this corresponds to the sum of two functions $f+g$.

<u>Q2.</u> Composition $g \circ f$. $f: \mathbb{R}^m \to \mathbb{R}^q$

when is $g \circ f$ defined? $g: \mathbb{R}^p \to \mathbb{R}^n$

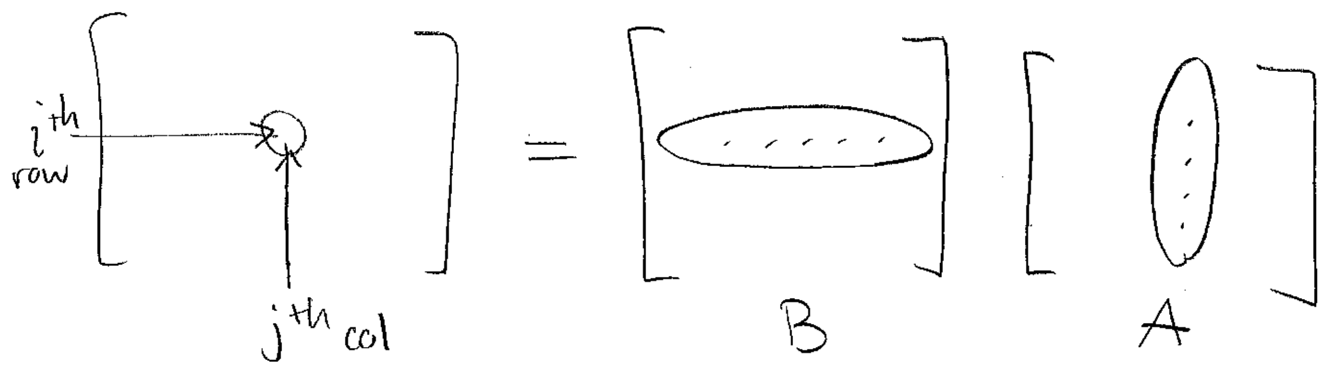
Need [target space of f] = [domain of g]

so need $p = q$.

$g \circ f(\vec{x}) = D\vec{x}$. What is D? See pg 42.

<u>So</u>: BA is defined the way it is because this corresponds to the composition $g \circ f$.

From examples, we can see that



Take the dot product of the ith row of B and jth column of A.

Matrix multiplication is <u>not</u> commutative.

But it is associative.

$f(\vec{x}) = A\vec{x}$

$g(\vec{y}) = B\vec{y}$

$h(\vec{z}) = C\vec{z}$

$$\underbrace{(AB)C}_{\downarrow} = \underbrace{A(BC)}_{\downarrow}$$

$$(f \circ g) \circ h \qquad f \circ (g \circ h)$$

Both mean: do h, then g, then f (in that order) we also have the distributive property.

$$\begin{cases} A(C+D) = AC + AD \\ (A+B)C = AC + BC \end{cases}$$

Identity function $\quad id_n : \mathbb{R}^n \to \mathbb{R}^n$

$$id_n(\vec{x}) = \vec{x}.$$

This is a linear transf.

with matrix $\quad I_n = \underbrace{\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}}_{n} \Big\} n$

called "identity matrix"

If $f(\vec{x}) = A\vec{x} \quad f : \mathbb{R}^m \to \mathbb{R}^n$

$(A \text{ is } n \times m).$

Then we can form the compositions

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, bend left, "f"] \arrow[rr, bend right, "f"'] & \mathbb{R}^n \arrow[r, bend left, "id_n"] & \mathbb{R}^n \end{tikzcd}$$

$id_n \circ f = f$

$$\boxed{I_n A = A}$$

$$\begin{tikzcd} \mathbb{R}^m \arrow[r, bend left, "id_m"] \arrow[rr, bend right, "f"'] & \mathbb{R}^m \arrow[r, bend left, "f"] & \mathbb{R}^n \end{tikzcd}$$

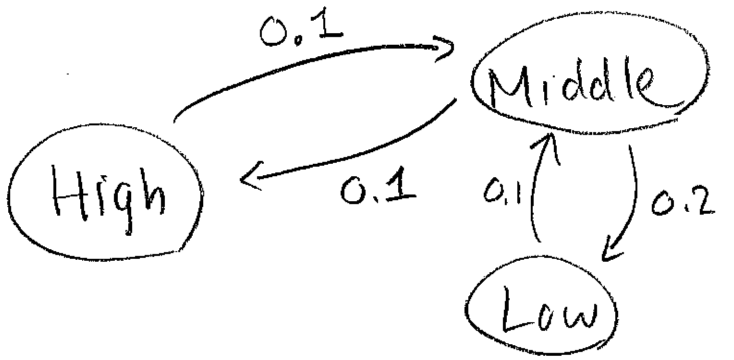
$f \circ id_m = f$

$$\boxed{A I_m = A}$$

Application: Markov Chains. (skip this for now...)

Example: Sort the countries of the world into 3 groups: High GDP / Middle / Low.

Each year, a country may move from one group to another.



suppose there are $H(n)$ countries in "High" in year n
$M(n)$ — — — "Middle" —
$L(n)$ — — — "Low"

Then next year:

$$H(n+1) = 0.9\,H(n) + 0.1\,M(n)$$
$$M(n+1) = 0.1\,H(n) + 0.7\,M(n) + 0.1\,L(n)$$
$$L(n+1) = 0.2\,M(n) + 0.8\,L(n)$$

i.e.
$$\begin{bmatrix} H(n+1) \\ M(n+1) \\ L(n+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.9 & 0.1 & \\ 0.1 & 0.7 & 0.1 \\ & 0.2 & 0.8 \end{bmatrix}}_{\text{call this matrix } A.} \begin{bmatrix} H(n) \\ M(n) \\ L(n) \end{bmatrix}$$

A is called a stochastic matrix or transition matrix or Markov matrix.
(i.e. it is square and all entries ≥ 0 and all columns sum to 1).

Then $\begin{bmatrix} H(n) \\ M(n) \\ L(n) \end{bmatrix} = A^n \begin{bmatrix} H(0) \\ M(0) \\ L(0) \end{bmatrix}$.

So if we're interested in what happens far into the future, we need to study A^n.

(Note: we can let H, M, L be numbers (actual counts), or we can let them be $\underbrace{\text{proportions}}_{\text{so } H+M+L=1}$.

In the latter case $\begin{bmatrix} H \\ M \\ L \end{bmatrix}$ is called a distribution vector.)

Q: Is there a limit as $n \to \infty$?

We'll come back to these kinds of questions later, but in case you're interested:

<u>Theorem 2.3.11</u>: Let A be a transition matrix. Suppose there is an m such that all entries of A^m are positive. Then the limit $\lim_{n \to 0} A^n \vec{x}$ exists and does not depend on the distribution $\vec{x}$. !

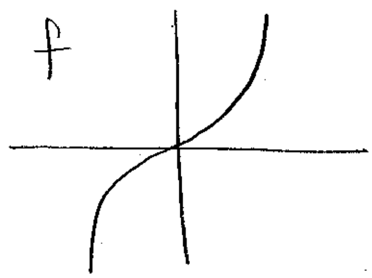
Lecture 7

10/23/18

Inverse functions

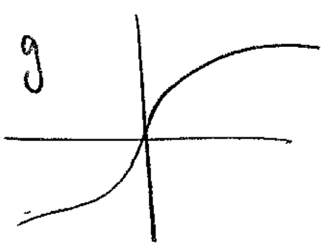
e.g. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ $\qquad$ $g : \mathbb{R} \to \mathbb{R}$, $g(y) = \sqrt[3]{y}$.

f



$$\begin{tikzcd} 3 \arrow[r, bend left, "f"] & 27 \arrow[l, bend left, "g"] \end{tikzcd}$$

g



$$(g \circ f)(x) = x \quad \underline{\underline{\text{and}}} \quad (f \circ g)(x) = x.$$

We say that <u>f and g are inverses.</u>

and we write
$f^{-1} = g$
$g^{-1} = f$

Another example: $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$

$$R_\theta(\vec{x}) = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, bend left, "R_\theta"] & \mathbb{R}^2 \arrow[l, bend left, "R_{-\theta}"] \end{tikzcd}$$

$$R_{-\theta}(\vec{x}) = \underbrace{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}_{B} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

so $R_\theta \circ R_{-\theta} = \text{id}$ $\qquad R_{-\theta} \circ R_\theta = \text{id}$

Then

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

or $\quad AB = I_2 \qquad$ we write $\quad A^{-1} = B$

$\quad\;\; BA = I_2 \qquad\qquad\qquad\quad B^{-1} = A.$

Definition (with lin transf): A square matrix A is invertible if the lin transf $T(\vec{x}) = A\vec{x}$ is invertible.

Alt. def (with matrix mult): A sq mx A is invertible if there is another sq mx B such that $AB = BA = I_n$. We say that B is the inverse of A and write $A^{-1} = B$.

How to find the inverse of an $n \times n$ matrix A? (example: $A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$).

Step 1: Start with the $n \times (2n)$ matrix $[A \mid I_n]$

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right].$$

Step 2: Find the rref of this matrix

$$\left[\begin{array}{cc|cc} 1 & 0 & -1/5 & 3/5 \\ 0 & 1 & 2/5 & -1/5 \end{array}\right]$$

Step 3: If the rref is of the form $[I_n \mid B]$ then $A^{-1} = B$.

If not, then A is not invertible.

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{bmatrix}$$

Why does this work?

We want to solve $A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ for x_1 and x_2. so:

$$\left| \begin{array}{l} x_1 + 3x_2 = y_1 \\ 2x_1 + x_2 = \quad y_2 \end{array} \right|$$

Gauss-Jordan elimination

$$\left| \begin{array}{l} x_1 \quad = -\frac{1}{5}y_1 + \frac{3}{5}y_2 \\ x_2 = \frac{2}{5}y_1 - \frac{1}{5}y_2 \end{array} \right|$$

Another example $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \begin{array}{c} \\ -(I) \end{array} \longrightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{array}\right] \quad (*)$$

$$\xrightarrow[\text{rref}]{} \left[\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right].$$

The rref is not of the form $\left[\begin{array}{cc|c} 1 & 0 & B \\ 0 & 1 & \end{array}\right]$ so A is not invertible. In fact we could have stopped at $(*)$.

Inverse of a 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad-bc \neq 0$.

$ad-bc$ is called the <u>determinant</u> of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

We'll talk about the determinant later in the course.

$$\left(\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ad-bc & 0 \\ 0 & -bc+ad \end{bmatrix} \right)$$

Q: If A, B are invertible $n \times n$ mxs, what is $(BA)^{-1}$?

Note $A^{-1}B^{-1}BA = I_n$ since LHS is

1. Do A
2. Do B
3. undo B
4. Undo A. $\longrightarrow$ gets you back to where you started.

So $(BA)^{-1} = A^{-1}B^{-1}$

warning! matrix mult is <u>not</u> commutative, so $(BA)^{-1}$ is <u>not</u> $B^{-1}A^{-1}$!

Matrix inverses give us a way to solve system of n eq and n vars in certain cases.

e.g. $\left| \begin{aligned} x_1 + 3x_2 &= 5 \\ 2x_1 + x_2 &= -5 \end{aligned} \right|$

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{bmatrix} \begin{bmatrix} 5 \\ -5 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

When A (an $n \times n$ mx) is invertible, things are very nice.

1. For every $\vec{y} \in \mathbb{R}^n$, we can find a $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{y}$.
2. The equation $A\vec{x} = \vec{0}$ has a unique solution: $\vec{x} = \vec{0}$.

This brings us to the topics of chapter 3.

Lecture 8 starts here!

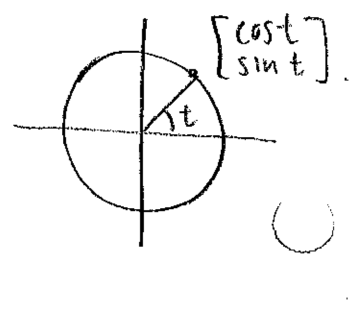
Image of a function: 10/25/18

<u>Def</u>: If $f: X \to Y$, then

$$\operatorname{im}(f) = \{ f(x) \mid x \in X \}.$$

Examples.

1. $f: \mathbb{R} \to \mathbb{R}$. $f(x) = x$.
   $\operatorname{im}(f) = \mathbb{R}$.
2. $f: \mathbb{R} \to \mathbb{R}$ $f(x) = x^2$
   $\operatorname{im}(f) = [0, \infty) = \mathbb{R}_{\geq 0}$.
3. $f: \mathbb{R} \to \mathbb{R}^2$ $f(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}$.



$\operatorname{Im}(f) =$ unit circle in $\mathbb{R}^2$.

4. $T: \mathbb{R}^2 \to \mathbb{R}^2$

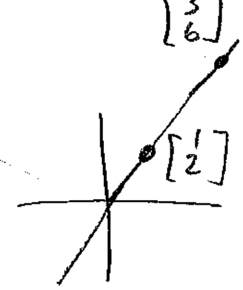
$T(\vec{x}) = A\vec{x} \qquad A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}.$

Note: A is invertible, so for all $\vec{y} \in \mathbb{R}^2$, there is a solution $\vec{x}$ to $A\vec{x} = \vec{y}$, namely $\vec{x} = A^{-1}\vec{y}$.

So $\mathrm{im}(T) = \mathrm{im}(A) = \mathbb{R}^2$.

5. $T: \mathbb{R}^2 \to \mathbb{R}^2$

$T(\vec{x}) = A\vec{x}. \qquad A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}.$



$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Observe that $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ are parallel. (Hence A is not invertible)

So $\mathrm{im}(T) = \mathrm{im}(A) = \left\{ \begin{bmatrix} t \\ 2t \end{bmatrix} \middle| \, t \in \mathbb{R} \right\}.$

Recall that in general, if

$$A = \underbrace{\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}}_{m.} \Big\} n$$

then $A\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \underbrace{x_1\vec{v}_1 + \cdots + x_m\vec{v}_m.}_{\text{a "linear combination" of } \vec{v}_1, \ldots, \vec{v}_m.}$

<u>Def</u>: The span of $\vec{v_1}, \dots, \vec{v_m}$ is the set of <u>all</u> linear comb.s of $\vec{v_1}, \dots, \vec{v_m}$.

i.e., $\text{span}(\vec{v_1}, \dots, \vec{v_m}) = \{ c_1\vec{v_1} + \dots + c_m\vec{v_m} \mid c_1 \in \mathbb{R}, \dots, c_m \in \mathbb{R} \}$.

Then if $A = \underbrace{\begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_m} \\ | & & | \end{bmatrix}}_{m} \Big\} n$.

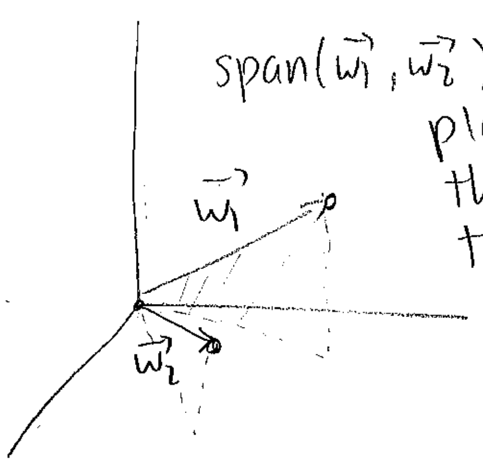
then $\boxed{\text{im}(A) = \text{span}(\vec{v_1}, \dots, \vec{v_m})}$

Some properties of the image of a lin. transf. $T: \mathbb{R}^m \to \mathbb{R}^n$.

a. $\vec{0} \in \text{im}\, T$

b. If $\vec{v_1}, \vec{v_2} \in \text{im}\, T$ then $\vec{v_1} + \vec{v_2} \in \text{im}\, T$

c. If $\vec{v_1} \in \text{im}\, T$ $k \in \mathbb{R}$ then $k\vec{v_1} \in \text{im}\, T$



span($\vec{w_1}, \vec{w_2}$) is the plane passing through these two pts and the origin.

Let's prove these properties.

a. $T\vec{0}=\vec{0}$.

b. Suppose $\vec{v_1}, \vec{v_2} \in \operatorname{im} T$. then there exist $\vec{w_1}, \vec{w_2} \in \mathbb{R}^m$ s.t. $\vec{v_1}=T(\vec{w_1})$ $\vec{v_2}=T(\vec{w_2})$. Then

$$T(\vec{w_1}+\vec{w_2})=T(\vec{w_1})+T(\vec{w_2})=\vec{v_1}+\vec{v_2}.$$

c. Suppose $\vec{v}\in \operatorname{im} T$. Then there exists $\vec{w}\in\mathbb{R}^m$ s.t. $\vec{v}=T(\vec{w})$. Then

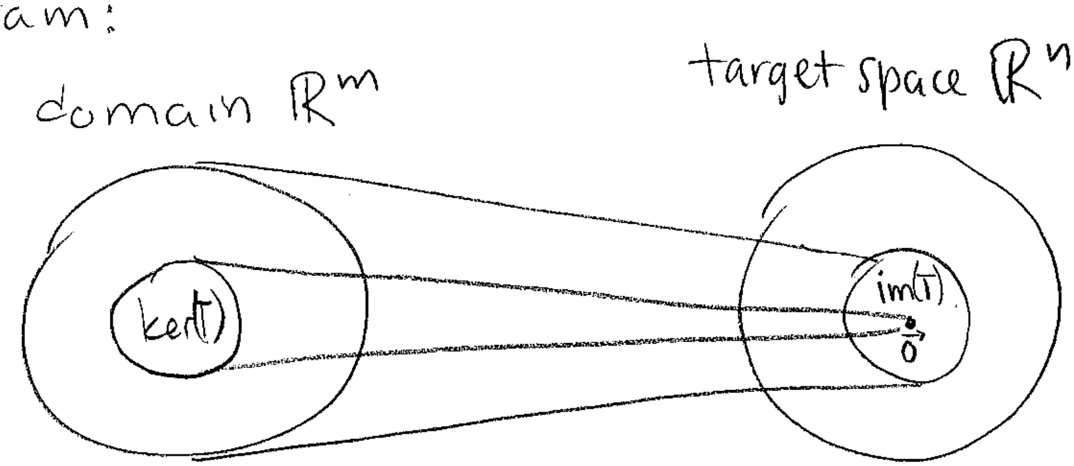
$$T(k\vec{w})=kT(\vec{w})=k\vec{v}.$$

Next topic: the kernel of a lin transf.

<u>Def</u>: $T:\mathbb{R}^m \to \mathbb{R}^n$ $T(\vec{x})=A\vec{x}$ (A $n\times m$)

$$\ker(T)=\ker(A)=\{\vec{x}\in\mathbb{R}^m \mid T(\vec{x})=0\}.$$

A diagram:

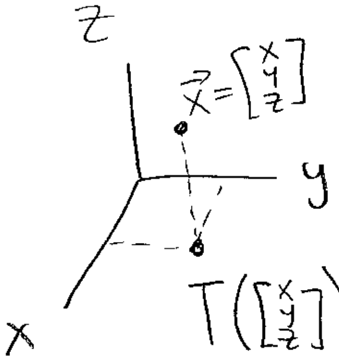


Examples:

1. $T: \mathbb{R}^2 \to \mathbb{R}^2$ is $90°$ counter-cw rotation.

   Then $\ker(T) = \{0\}$.

2. $T: \mathbb{R}^3 \to \mathbb{R}^3$ is orthog proj onto x-y plane.



$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

$$\ker(T) = \left\{ \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\} = z\text{-axis} = \operatorname{span}\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$$

Another way to see this: we're solving

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

The aug. matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Already in rref. We see z is a free var.

3. $T: \mathbb{R}^3 \to \mathbb{R}^2$

$$T(\vec{x}) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \vec{x}.$$

Aug matrix $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array}\right]$

$$\Big\downarrow \text{rref}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

free variable.

$$\left|\begin{array}{l} x - z = 0 \\ y + 2z = 0 \end{array}\right| \to \text{let } z = t \to \left|\begin{array}{l} x = t \\ y = -2t \\ z = t \end{array}\right|$$

so $\ker(T) = \left\{ \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}$

$$= \operatorname{span}\left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right).$$

4. $T: \mathbb{R}^5 \to \mathbb{R}^4 \quad T(\vec{x}) = A\vec{x}$.

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

$$\text{rref}\left[A \,\middle|\, \vec{0}\right] = \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & -4 & 0 \\ 0 & 0 & 1 & -4 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

3 free variables: x_2, x_4, x_5.

$$\left| \begin{array}{l} x_1 = -2x_2 - 3x_4 + 4x_5 \\ x_3 = 4x_4 - 5x_5 \end{array} \right|$$

so
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_4 + 4x_5 \\ x_2 \\ 4x_4 - 5x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix}$$

So
$$\ker(T) = \operatorname{span}\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix} \right)$$

Some properties of the kernel of $T: \mathbb{R}^m \to \mathbb{R}^n$.

a. $\vec{0} \in \ker T$

b. $\ker T$ is closed under addition

c. $\ker T$ is closed under scalar mult.

(You did this on a previous HW assignment.)

e.g. proof of b. Suppose $\vec{v_1}, \vec{v_2} \in \ker T$.

Then $T(\vec{v_1} + \vec{v_2}) = T(\vec{v_1}) + T(\vec{v_2}) = \vec{0} + \vec{0} = \vec{0}$

so $\vec{v_1} + \vec{v_2} \in \ker T$.

When is $\ker(A) = \{\vec{0}\}$?

Recall A $(n \times m)$

To solve $A\vec{x} = \vec{0}$, we look at $\mathrm{rref}(A)$.

e.g. if $\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & 5 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

free variable

$\longrightarrow$ kernel contains a nonzero vector.

e.g. if $\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

no free vars $\longrightarrow$ $\ker(A) = \{\vec{0}\}$.

So: $\ker(A) = \{\vec{0}\}$ if and only if there are no free variables.

But recall: no free variables is the same as

$$\underbrace{\mathrm{rank}(A)}_{\#\text{ of pivots in rref}(A)} = \underbrace{m}_{\#\text{ of columns in } A}$$

So: (Thm 3.1.7a) $\ker(A) = \{\vec{0}\}$ if and only if $\text{rank}(A) = m$.

Let A be $n \times n$ mx (square!).

Then: A is invertible $\Longrightarrow$ $\ker(A) = \{\vec{0}\}$.

A not invertible $\Longrightarrow$ $\ker(A) \neq \{\vec{0}\}$.

(since there are free variables).

Various characterizations of invertible mxs. (3.1.8).

Let A be an $n \times n$ matrix. The following statements are equivalent.

i. A is invertible

ii. For every $\vec{b} \in \mathbb{R}^n$ the system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$.

iii. $\text{rref}(A) = I_n$

iv. $\text{rank}(A) = n$

v. $\text{im}(A) = \mathbb{R}^n$

vi. $\ker(A) = \{\vec{0}\}$

Lecture 9

10/30/18

<u>Recall</u>: $T: \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation.

$$\text{im}(T) = \{ T(\vec{x}) \mid \vec{x} \in \mathbb{R}^m \}$$

$$\ker(T) = \{ \vec{x} \in \mathbb{R}^m \mid T(\vec{x}) = \vec{0} \}.$$

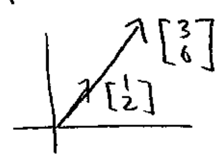
Examples from last time:

1. $T: \mathbb{R}^2 \to \mathbb{R}^2 \quad T(\vec{x}) = A\vec{x} \quad A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}.$

$$\text{im}(T) = \text{im}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}\right) = \mathbb{R}^2.$$

$$\ker(T) = \ker(A) = \{ \vec{x} \in \mathbb{R}^2 \mid A\vec{x} = \vec{0} \} = \{ \vec{0} \}$$

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



$$\text{im}(T) = \text{im}(A) = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}\right) = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$$

$$= \left\{ \begin{bmatrix} t \\ 2t \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}.$$

$\ker(T)$? what are the solutions to

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \;?$$

$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 6 & 0 \end{array}\right] \begin{array}{l} \\ -2(I) \end{array} \xrightarrow{\text{rref}} \left[\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

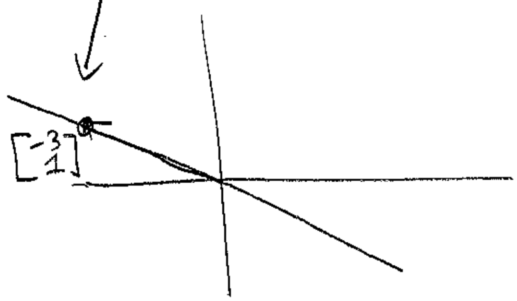
(aug. matrix)

this x_2 is a free variable

$\longrightarrow \quad x_1 + 3x_2 = 0. \qquad$ if $x_2 = t$

then $x_1 = -3t$.

So $\ker(T) = \left\{ \begin{bmatrix} -3t \\ t \end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\} = \operatorname{span}\left(\begin{bmatrix} -3 \\ 1 \end{bmatrix} \right)$.



3. Do example #4 on pg. 59. Find the kernel.

Continue w/ pg 60 etc...

$\operatorname{Im}(T)$ and $\ker(T)$ are examples of <u>subspaces</u>.

<u>Definition</u>: A subset W of $\mathbb{R}^n$ is called a linear subspace of $\mathbb{R}^n$ if it satisfies

a. $\vec{0} \in W$.

b. W is closed under addition

c. W is closed under scalar multiplication.

(Note: (b) and (c) together mean W is closed under linear combinations, ie. if $\vec{w_1}, \dots, \vec{w_m} \in W$, $k_1, \dots k_m \in \mathbb{R}$ then $k_1\vec{w_1} + \dots + k_m\vec{w_m} \in W$. Equivalently, if $\vec{w_1}, \dots, \vec{w_m} \in W$, then $\operatorname{span}(\vec{w_1}, \dots, \vec{w_m}) \subseteq W$.)

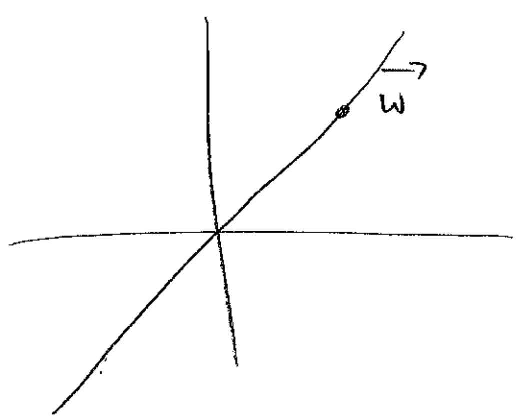
If $T(\vec{x}) = A\vec{x}$ is a lin transf $T: \mathbb{R}^m \to \mathbb{R}^n$, then

- $\ker(T)$ is a subspace of $\mathbb{R}^m$
- $\text{im}(T)$ is a subspace of $\mathbb{R}^n$.

<u>Q</u>: What are the subspaces of $\mathbb{R}^2$?

<u>A</u>: $\{\vec{0}\}$ is a subspace.

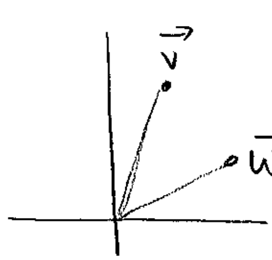
Suppose W is a subspace of $\mathbb{R}^2$ and it contains a nonzero vector $\vec{w}$.



Then since W is closed under scalar mult, W must contain the entire line through $\vec{0}$ and $\vec{w}$.

→ Any line through $\vec{0}$ is a subspace of $\mathbb{R}^2$

What if we want something more?



Since W is closed under lin comb.

$\vec{v}, \vec{w} \in W \Rightarrow$ every lin comb of $\vec{v}, \vec{w}$ is in W

$\Rightarrow W = \mathbb{R}^2$

So the subspaces of $\mathbb{R}^2$ are.

1. $\{\vec{0}\}$ (0-dim)
2. A line through $\vec{0}$ (1-dim)
3. All of $\mathbb{R}^2$ itself (2-dim).

The subspaces of $\mathbb{R}^3$ are:

1. $\{\vec{0}\}$ (0-dim)
2. A line through $\vec{0}$. (1-dim)
3. A plane through $\vec{0}$ (2-dim)
4. All of $\mathbb{R}^3$ (3-dim).

The subspaces of $\mathbb{R}^n$ are:

any k-dim plane through $\vec{0}$.
$(0 \leq k \leq n)$. → use your imagination ☺.

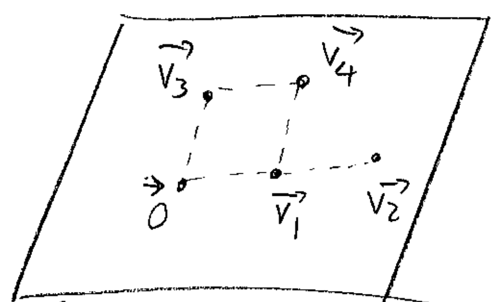
How can we describe a particular subspace? One way is to write a list of vectors whose span is that space. Often, it is better if our list of vectors has no "redundancy".

Example: $\vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $\vec{v_2} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$ $\vec{v_3} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\vec{v_4} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$.

Consider $\text{span}(\vec{v_1}, \vec{v_2}, \vec{v_3}, \vec{v_4})$. We can check for redundancy by adding one vector at a time.

1. $\text{span}(\vec{v_1})$ = a line. no redundancy.
2. $\vec{v_2}$ is a scalar multiple of $\vec{v_1}$ (i.e. $\vec{v_2} \in \text{span}(\vec{v_1})$). so
   $\text{span}(\vec{v_1}, \vec{v_2}) = \text{span}(\vec{v_1}) \rightsquigarrow \vec{v_2}$ is redundant.
3. $\vec{v_3}$ is not parallel to $\vec{v_1}$.
   $\rightarrow \text{span}(\vec{v_1}, \vec{v_3})$ = a plane. $\vec{v_3}$ is not redund
4. $\text{span}(\vec{v_1}, \vec{v_3}, \vec{v_4}) = \text{span}(\vec{v_1}, \vec{v_3})$ since $\vec{v_4} = \vec{v_1} + \vec{v_3}$.

From the picture,



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

$$\text{im}\begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

$= \text{span}(\vec{v_1}, \vec{v_3})$.

$= \text{span}(\vec{v_1}, \vec{v_4})$

$= \text{span}(\vec{v_2}, \vec{v_3}) = \text{span}(\vec{v_2}, \vec{v_4}) = \text{span}(\vec{v_3}, \vec{v_4})$.

Def: Consider a list of vectors $\vec{v_1}, \dots, \vec{v_m}$ in $\mathbb{R}^n$

a. We say $\vec{v_i}$ is redundant in this list if $\vec{v_i} \in \text{span}(\vec{v_1}, \dots, \vec{v_{i-1}})$

b. If none of the vecs in the list are redund., then we say $\vec{v_1}, \dots, \vec{v_m}$ are linearly independent. Otherwise, they are linearly dependent.

c. Let $V \subset \mathbb{R}^n$ be a subspace. We say $\vec{v_1}, \dots, \vec{v_m}$ is a basis of V if

① $V = \text{span}(\vec{v_1}, \dots \vec{v_m})$

and ② $\vec{v_1}, \dots, \vec{v_m}$ are lin. indep.

So from the prev example:

- $\vec{v_1}, \vec{v_2}, \vec{v_3}, \vec{v_4}$ are lin. dep.
  ($\vec{v_2}$ is redund. $\vec{v_4}$ is redund)
- $\vec{v_1}, \vec{v_3}$ are lin. indep.
  ($\vec{v_1}, \vec{v_4}$ is also lin indep. etc).
- $\vec{v_1}, \vec{v_3}$ is a basis of $\text{im} A$.

Another definition of linearly independent:

The vectors $\vec{v_1}, \dots, \vec{v_m} \in \mathbb{R}^n$ are <u>linearly dependent</u> if we can find scalars $c_1, c_2, \dots, c_m$ not all zero such that.

$$c_1\vec{v_1} + \dots + c_m\vec{v_m} = \vec{0}.$$

(If we can't find such scalars, then the vecs are lin. indep.)

(this is called a "nontrivial relation")

From the previous example.

$$2\vec{v_1} + (-1)\vec{v_2} + 0\vec{v_3} + 0\vec{v_4} = \vec{0}$$

so $\vec{v_1}, \dots, \vec{v_4}$ are lin. dep.

Note: This means: $A\begin{bmatrix}2\\-1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$.

So $\begin{bmatrix}2\\-1\\0\\0\end{bmatrix} \in \ker A$.

So in general,

$\vec{v_1}, \dots, \vec{v_m}$ are lin indep if and only if

$$\ker\begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_m} \\ | & & | \end{bmatrix} = \{\vec{0}\}.$$

(Lecture 10 = midterm) 11/6/18

Lecture 11

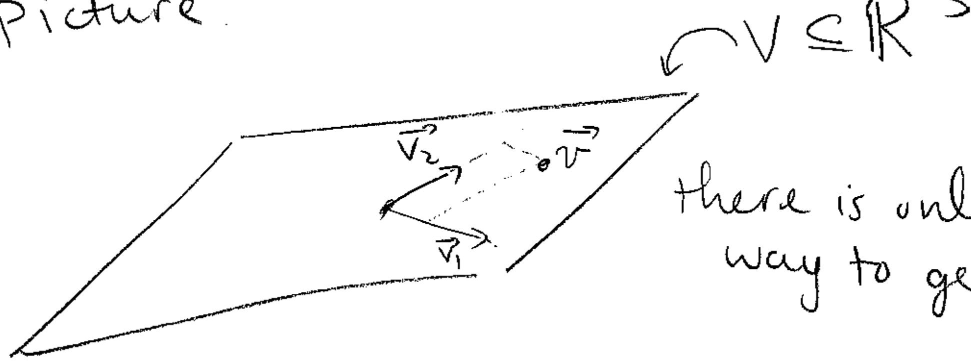
- Start on pg 68.

More on bases.

- $\vec{e_1}=\begin{bmatrix}1\\0\end{bmatrix}$ $\vec{e_2}=\begin{bmatrix}0\\1\end{bmatrix}$ is a basis for $\mathbb{R}^2$
- $\vec{e_1}=\begin{bmatrix}1\\0\\0\end{bmatrix}$ $\vec{e_2}=\begin{bmatrix}0\\1\\0\end{bmatrix}$ $\vec{e_3}=\begin{bmatrix}0\\0\\1\end{bmatrix}$ is a basis for $\mathbb{R}^3$.
- In general: suppose $\vec{v_1},\dots,\vec{v_m}$ are vectors in a subspace V of $\mathbb{R}^m$. Then $\vec{v_1},\dots,\vec{v_m}$ form a basis of V if and only if every $\vec{v}\in V$ can be expressed uniquely as a linear combination

$$\vec{v}=c_1\vec{v_1}+\dots+c_m\vec{v_m}.$$

Picture.



there is only one way to get $\vec{v}$

This is because: if $\vec{v_1}, \dots, \vec{v_m}$ is a basis for V, then by def:

1. $\text{span}(\vec{v_1}, \dots, \vec{v_m}) = V$ $\longrightarrow$ you can get every vector in V in at least one way.

2. $\vec{v_1}, \dots, \vec{v_m}$ are lin indep
$\longrightarrow$ no redundancy, so you can get every vector in V in at most one way.

Consider V = a plane in $\mathbb{R}^3$ (through $\vec{0}$)

(Think geometrically)

a. We can find at most <u>2</u> lin. indep vecs in V.

b. We need at least <u>2</u> vectors to span V.

c. Every basis of V has exactly <u>2</u> vectors.

In general, if V is a subspace of $\mathbb{R}^n$, the <u>dimension of V</u> is defined to be the number of vectors in a basis of V.
(If $V = \{\vec{0}\}$ then $\dim V = 0$).

If $V \subset \mathbb{R}^n$ is a subspace and $\dim V = m$, then.

a. We can find at most m lin. indep. vectors in V.

b. We need at least m vectors to span V.

c. If m vectors in V are lin indep, then they form a basis of V.

d. If m vectors in V span V, then they form a basis of V.

(Just think about a plane in $\mathbb{R}^3$!)

Algorithms for finding bases of kernel and image

Ex: $A = \begin{bmatrix} 1 & 2 & 2 & -5 & 6 \\ -1 & -2 & -1 & 1 & -1 \\ 4 & 8 & 5 & -8 & 9 \\ 3 & 6 & 1 & 5 & -7 \end{bmatrix}$.

$B = \mathrm{rref}(A) = \begin{bmatrix} 1 & 2 & 0 & 3 & -4 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

We already know how to write

$\ker(B) = \mathrm{span}(\text{something})$.

Use rref(A) : x_2, x_4, x_5 are free vars

and $x_1 = -2x_2 - 3x_4 + 4x_5$

$x_3 = 4x_4 - 5x_5$

so, any element of ker(A) is of the form

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_4 + 4x_5 \\ x_2 \\ 4x_4 - 5x_5 \\ x_4 \\ x_5 \end{bmatrix}$$

$$= x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix}.$$

$$\text{so } \ker(A) = \operatorname{span}\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ -5 \\ 0 \\ 1 \end{bmatrix} \right)$$

From the 2nd, 4th, and 5th rows, we see these 3 vectors cannot have any nontrivial relations. so they are lin indep. (This works in general.) Hence they are a basis for ker(A).

How to use rref (A) to find a basis for im(A)?

Algorithm:

1. Find rref(A) and identify the columns with pivots.
2. Take the columns of A that correspond to those columns in rref(A) with pivots.

In the example, it would be 1st and 3rd cols.

so a basis for im(A) is $\begin{bmatrix} 1 \\ -1 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 5 \\ 1 \end{bmatrix}$

Why does this work?

1. The non-redundant cols of B are those with pivots
2. Elementary row operations do not change relations between columns.

e.g. $\vec{b_5} = -4\vec{b_1} + 5\vec{b_3}$

$\Downarrow$

$\vec{a_5} = -4\vec{a_1} + 5\vec{a_3}$.

Lecture 12 11/8/18

We saw:

$\dim(\ker(A)) = \#$ of free variables.

$\dim(\mathrm{im}(A)) = \#$ of pivots $= \mathrm{rank}(A)$.

So $\dim(\ker A) + \dim(\mathrm{im}\, A)$

$= \#$ of free vars $+ \#$ of pivots

$= \#$ of cols of A.

Rank-nullity theorem: A is an $n \times m$ matrix.

Then $\dim(\ker A) + \dim(\mathrm{im}\, A) = m$.

Example

1. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ A is invertible.

$\Rightarrow \begin{cases} \mathrm{im}\, A = \mathbb{R}^2 \\ \ker A = \{\vec{0}\}. \end{cases}$

$\dim \mathrm{im}\, A = 2$

$\dim \ker A = 0$.

2. $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. $\mathrm{im}\, A = \mathrm{span}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$.

$\ker A = \mathrm{span}\left(\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right)$.

$\dim \mathrm{im}\, A = 1$

$\dim \ker A = 1$.

Fact: $\vec{v_1}, \ldots, \vec{v_n}$ form a basis of $\mathbb{R}^n$ if and only if

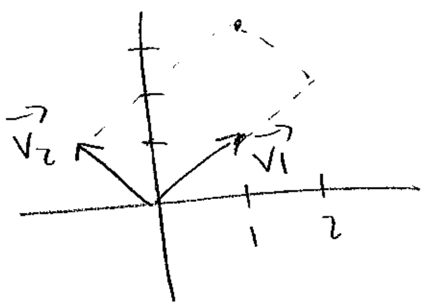
$$\begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_n} \\ | & | & & | \end{bmatrix}$$

is invertible.

Next topic: coordinates. (very quickly)

$\vec{v_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\vec{v_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. $\vec{v_1}, \vec{v_2}$ is a basis for $\mathbb{R}^2$

Let $\mathcal{B} = (\vec{v_1}, \vec{v_2})$.



$$\vec{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 2\vec{v_1} + 1\vec{v_2}$$

we can think of (2, 1) as "coordinates" of this point with respect to $\mathcal{B}$.

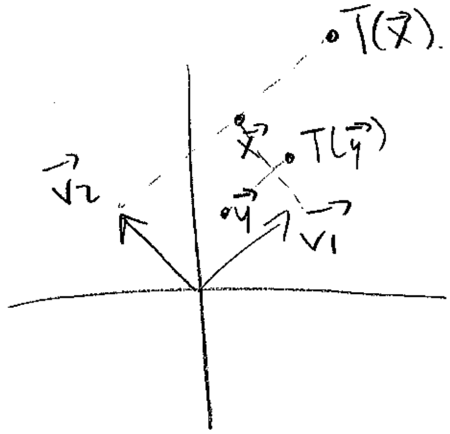
So we write $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\begin{matrix} \leftarrow \text{coeff of } \vec{v_1} \\ \leftarrow \text{coeff of } \vec{v_2}. \end{matrix}$

Something it is easier to describe a linear transformation w.r.t. another coordinate system.

Note: $\vec{v} = 2\vec{v_1} + \vec{v_2} = \begin{bmatrix} | & | \\ \vec{v_1} & \vec{v_2} \\ | & | \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Suppose $T: \mathbb{R}^2 \to \mathbb{R}^2$.

$$\left.\begin{aligned} T\vec{v_1} &= 2\vec{v_1} \\ T\vec{v_2} &= \vec{v_2} \end{aligned}\right] \text{ stretching in } \vec{v_1} \text{ direction.}$$



$$[T(\vec{x})]_{\mathcal{B}} = \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}_{\substack{\uparrow \\ \text{called the} \\ \mathcal{B}\text{-matrix of } T.}} [\vec{x}]_{\mathcal{B}}$$

But what is the regular matrix of T?

Note $[\vec{y}]_{\mathcal{B}} = \underbrace{\begin{bmatrix} | & | \\ \vec{v_1} & \vec{v_2} \\ | & | \end{bmatrix}^{-1}}_{S^{-1}} \vec{y}$ let $S = \begin{bmatrix} | & | \\ \vec{v_1} & \vec{v_2} \\ | & | \end{bmatrix}$.

Then $[T(\vec{x})]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} [\vec{x}]_{\mathcal{B}}$ becomes

$$S^{-1} T(\vec{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} S^{-1} \vec{x}.$$

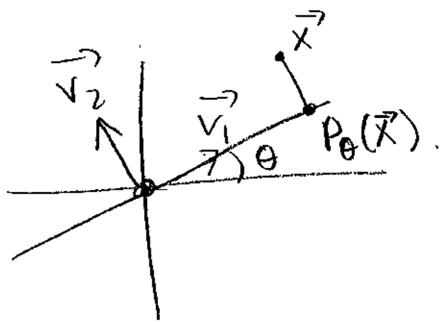
$$T(\vec{x}) = \underbrace{S \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} S^{-1}}_{\text{the matrix for } T.} \vec{x}$$

So: in general. if $\beta = (\vec{v}_1, \dots, \vec{v}_n)$.

$$[T(\vec{x})]_\beta = B\,[\vec{x}]_\beta . \qquad S = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix}.$$

Then the matrix for T is $\boxed{SBS^{-1}}$.

Example: projection onto a line.



P_θ = projection onto line with angle θ wrt. x-axis.

let $\vec{v}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ $\quad \vec{v}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ $\quad \beta = (\vec{v}_1, \vec{v}_2)$.

Then $[P_\theta(\vec{x})]_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} [\vec{x}]_\beta$. so the matrix for P_θ is.

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix}.$$

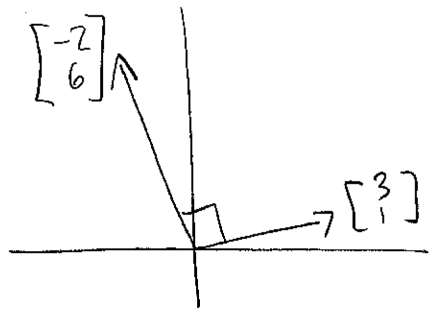
There's an easier way to think about orthogonal projections.

Recall the dot product of two vectors:

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

In $\mathbb{R}^2$:



note: $\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 6 \end{bmatrix} = 3(-2) + 1(6)$
$= 0.$

Also, the length of $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is

$$\sqrt{3^2 + 1^2} = \sqrt{\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix}}.$$

In general:

a. $\vec{v}, \vec{w} \in \mathbb{R}^n$ are orthogonal if $\vec{v} \cdot \vec{w} = 0$

b. The length (or magnitude) of $\vec{v}$ is
$\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$

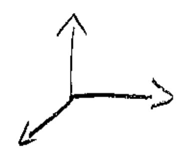
c. $\vec{v}$ is a unit vector if $\|\vec{v}\| = 1$

Def: $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m} \in \mathbb{R}^n$ are <u>orthonormal</u> if they are all unit vectors and orthogonal to one another. i.e.

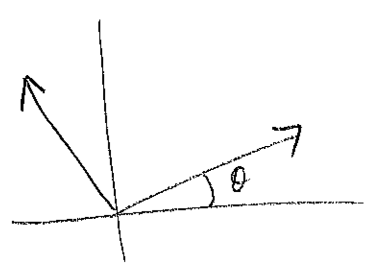
$$\vec{u_i} \cdot \vec{u_j} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$$

For <u>example</u>: ① In $\mathbb{R}^2$, $\vec{e_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $\vec{e_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are orthonormal.

② In $\mathbb{R}^n$, $\vec{e_1}, \ldots, \vec{e_n}$ are orthonormal.



③ In $\mathbb{R}^2$, $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$, $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ are orthonormal.



<u>Fact</u>: Orthonormal vectors are linearly independent.
How to show this? Suppose $\vec{u_1}, \ldots, \vec{u_m} \in \mathbb{R}^n$ are orthonormal, and consider

$$c_1\vec{u_1} + \cdots + c_m\vec{u_m} = \vec{0}.$$

We want to find solutions $c_1, \ldots, c_m$.

Take the dot product of both sides with $\vec{u_1}$.

$$(c_1\vec{u_1} + c_2\vec{u_2} + \dots + c_m\vec{u_m}) \cdot \vec{u_1} = 0 \cdot \vec{u_1}$$

$$c_1\underbrace{(\vec{u_1}\cdot\vec{u_1})}_{1} + c_2\underbrace{(\vec{u_2}\cdot\vec{u_1})}_{0} + \dots + c_m\underbrace{(\vec{u_m}\cdot\vec{u_1})}_{0} = 0.$$

$$\boxed{c_1 = 0}$$

Similarly $c_2 = 0, \dots, c_m = 0$.

So the only solution to $c_1\vec{u_1} + \dots + c_m\vec{u_m} = 0$ is when $c_1 = c_2 = \dots = c_m = 0$.

$\Longrightarrow$ $\vec{u_1}, \dots, \vec{u_m}$ are lin indep.

This proof also tells us something else.

Suppose
- $\vec{u_1}, \dots, \vec{u_m} \in \mathbb{R}^n$ are orthonormal.
- $\vec{v} \in \text{span}(\vec{u_1}, \dots, \vec{u_m})$.

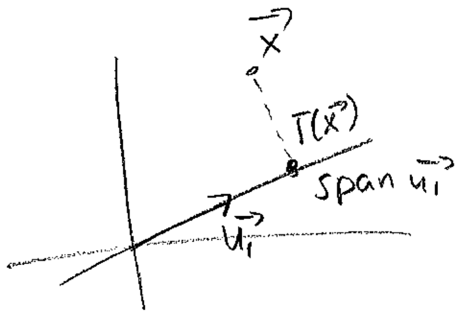
How do we find $c_1, \dots, c_m$ such that $\vec{v} = c_1\vec{u_1} + \dots + c_m\vec{u_m}$?

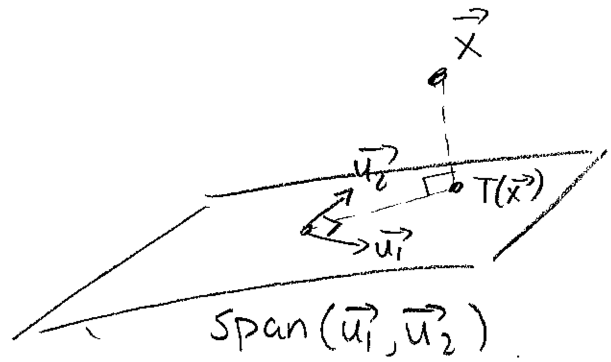
Dot both sides with $\vec{u_1}$ to get $\boxed{\vec{v}\cdot\vec{u_1} = c_1}$.

Similarly $\boxed{c_i = \vec{v}\cdot\vec{u_i}}$ (when

Technique: When we have orthonormal vectors, we can often take dot products to help us find coefficients.

Example: orthogonal projections onto subspaces

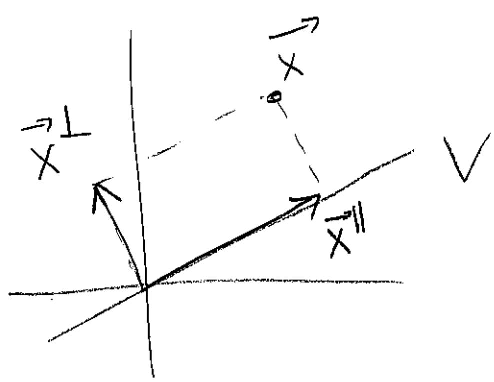


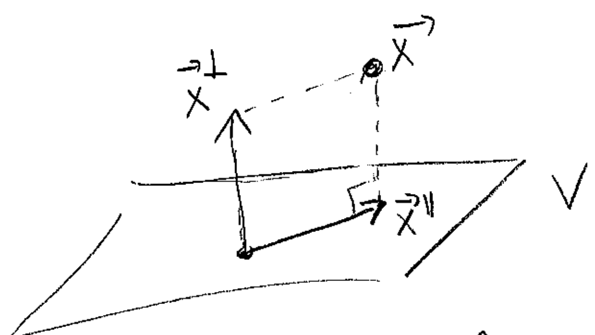


let $\vec{x} \in \mathbb{R}^n$ and V be a subspace of $\mathbb{R}^n$.

we can split $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$

where $\vec{x}^{\parallel} \in V$, $\vec{x}^{\perp}$ is perp to V.





If $\vec{u_1}, \dots, \vec{u_m}$ is an orthonormal basis of V,

then $\vec{x}^{\parallel} = c_1\vec{u_1} + \dots + c_m\vec{u_m}$.

Observe: $\vec{x}^{\parallel}$ is the orthog. proj. of $\vec{x}$ onto V.

$$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$$

$$\vec{x} \cdot \vec{u}_1 = \vec{x}^{\parallel} \cdot \vec{u}_1 + \underbrace{\vec{x}^{\perp} \cdot \vec{u}_1}_{\text{orthogonal}} = c_1 + 0.$$

So $\quad c_1 = \vec{x} \cdot \vec{u}_1 \quad c_2 = \vec{x} \cdot \vec{u}_2$ etc.

So $\quad \vec{x}^{\parallel} = \text{proj}_V \vec{x} = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + \dots + (\vec{x} \cdot \vec{u}_m)\vec{u}_m$.

This gives us a way to find the projection of $\vec{x}$ onto V.

If $\vec{x} \in \mathbb{R}^n$, $V \subset \mathbb{R}^n$ is a subspace, to find $\text{proj}_V \vec{x}$:

1. Find an orthonormal basis for V: $\vec{u}_1, \dots, \vec{u}_m$.
2. Compute $\vec{u}_1 \cdot \vec{x}, \dots, \vec{u}_m \cdot \vec{x}$.
3. $\text{proj}_V \vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \dots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$.

We can also use matrices to describe this process. For step 2, observe:

$$\begin{bmatrix} \text{---} & \vec{u}_1 & \text{---} \\ \text{---} & \vec{u}_2 & \text{---} \\ & \vdots & \\ \text{---} & \vec{u}_m & \text{---} \end{bmatrix} \begin{bmatrix} \\ \vec{x} \\ \\ \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \cdot \vec{x} \\ \vdots \\ \vec{u}_m \cdot \vec{x} \end{bmatrix}.$$

For step 3, observe:

$$\begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_m \\ | & & | \end{bmatrix} \begin{bmatrix} \vec{u}_1 \cdot \vec{x} \\ \vdots \\ \vec{u}_m \cdot \vec{x} \end{bmatrix} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$$
$$= \mathrm{proj}_V \vec{x}.$$

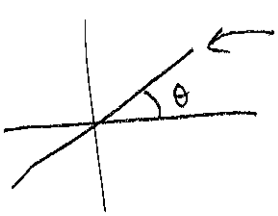
so let $Q = \begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_m \\ | & & | \end{bmatrix}$ $\quad Q^T = \begin{bmatrix} — & \vec{u}_1 & — \\ & \vdots & \\ — & \vec{u}_m & — \end{bmatrix}$

T = transpose = flip the matrix
(interchange rows and cols)

Then $\boxed{\mathrm{proj}_V \vec{x} = QQ^T\vec{x}}$

The matrix for orthogonal projection onto V is QQ^T. (Theorem 5.3.10.)

Example: $T: \mathbb{R}^2 \to \mathbb{R}^2$ orthog. proj. onto line w/ angle θ wrt x-axis.



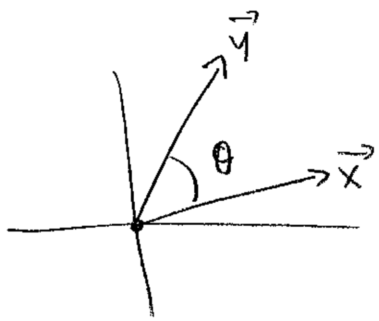
← let this be V.

An orthonormal basis for V: $\vec{u}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$.

so $Q = \begin{bmatrix} | \\ \vec{u}_1 \\ | \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ $\quad Q^T = [\cos\theta \;\; \sin\theta]$.

$$QQ^T = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} [\cos\theta \;\; \sin\theta] = \underbrace{\begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}}_{\text{this is the matrix for the projection.}}$$

Angle between two vectors. [skip this section]



Identity: $\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \cos\theta$.
This gives you the angle θ between $\vec{x}$ and $\vec{y}$.

The identity works in any dimensions.

Observe: If $\vec{x} \cdot \vec{y} = 0$ then $\cos\theta = 0$. so $\theta = 90°$ as expected.

Correlation. (See section 5.1).

If $\vec{x}, \vec{y} \in \mathbb{R}^n$ have mean 0 (ie $x_1 + \dots + x_n = 0$ $\;y_1 + \dots + y_n = 0$)

then the correlation coefficient is

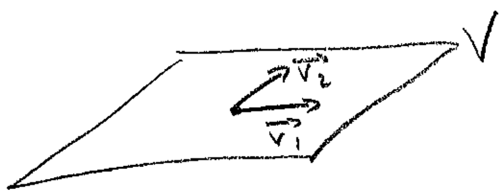
$$r = \cos\theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}$$

Lecture 14 11/15/18

Q: Given a subspace V of $\mathbb{R}^n$, how to find an orthonormal basis for it?

Example: $V \subset \mathbb{R}^4$ is a 2-dim plane

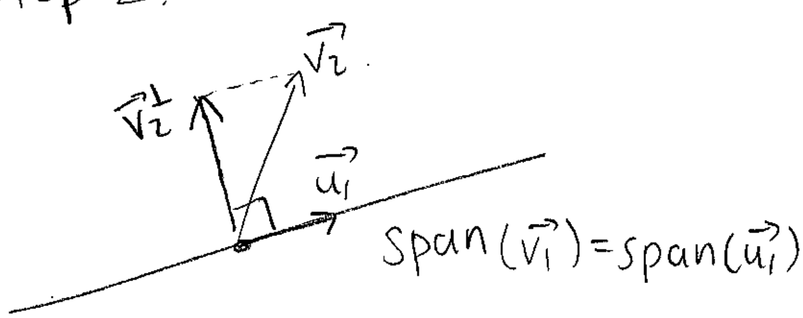
$$V = \operatorname{span}(\vec{v_1}, \vec{v_2}) \qquad \vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{v_2} = \begin{bmatrix} 1 \\ 9 \\ 9 \\ 1 \end{bmatrix}.$$



Step 1: Normalize $\vec{v_1}$: $\vec{u_1} = \dfrac{\vec{v_1}}{\|\vec{v_1}\|} = \dfrac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

(so $\vec{u_1}$ is a unit vector)

Step 2.



consider $\vec{v_2}^{\perp} = \vec{v_2} - (\vec{v_2} \cdot \vec{u_1})\vec{u_1} = \begin{bmatrix} -4 \\ 4 \\ 4 \\ -4 \end{bmatrix}$

let $\vec{u_2} = \dfrac{\vec{v_2}^{\perp}}{\|\vec{v_2}^{\perp}\|} = \dfrac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$

Done! $\vec{u_1}, \vec{u_2}$ is an orthonormal basis of V.

Another example:

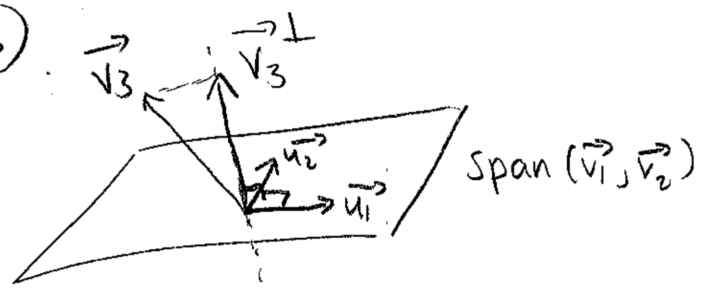
Suppose $V = \text{span}(\vec{v_1}, \vec{v_2}, \vec{v_3})$ where $\vec{v_1}, \vec{v_2}, \vec{v_3}$ is a basis for V.

1. $\vec{u_1} = \dfrac{\vec{v_1}}{\|\vec{v_1}\|}$ $\qquad$ $\text{span}(\vec{v_1}) = \text{span}(\vec{u_1})$

2. $\vec{v_2}^{\perp} = \vec{v_2} - (\text{proj of } \vec{v_2} \text{ onto span}(\vec{u_1}))$

$= \vec{v_2} - (\vec{v_2} \cdot \vec{u_1})\vec{u_1}$ $\longleftarrow$ this is orthogonal to $\vec{u_1}$.

$\vec{u_2} = \dfrac{\vec{v_2}^{\perp}}{\|\vec{v_2}^{\perp}\|}$ $\qquad$ $\text{span}(\vec{v_1}, \vec{v_2}) = \text{span}(\vec{u_1}, \vec{u_2})$.

3. 

$\vec{v_3}^{\perp} = \vec{v_3} - (\text{proj of } \vec{v_3} \text{ onto span}(\vec{u_1}, \vec{u_2}))$

$= \vec{v_3} - \left((\vec{v_3} \cdot \vec{u_1})\vec{u_1} + (\vec{v_3} \cdot \vec{u_2})\vec{u_2} \right)$.

$\vec{u_3} = \dfrac{\vec{v_3}^{\perp}}{\|\vec{v_3}^{\perp}\|}$ $\qquad$ $V = \text{span}(\vec{v_1}, \vec{v_2}, \vec{v_3}) = \text{span}(\vec{u_1}, \vec{u_2}, \vec{u_3})$.

$\vec{u_1}, \vec{u_2}, \vec{u_3}$ is an orthonormal basis for V!

(If you have more vectors, keep going.) This is called the Gram-Schmidt process.

So: To find the orthog. projection onto V,

1. Start with a basis for V: $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$
2. Gram-Schmidt process gives an orthonormal basis $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_m}$.
3. Then $\text{proj}_V(\vec{x}) = (\vec{u_1} \cdot \vec{x})\vec{u_1} + \cdots + (\vec{u_m} \cdot \vec{x})\vec{u_m}$.

$$= Q Q^T \vec{x}.$$

$$Q = \begin{bmatrix} | & | & & | \\ \vec{v_1} & \vec{v_2} & \cdots & \vec{v_m} \\ | & | & & | \end{bmatrix}$$

$$Q^T = \begin{bmatrix} — & \vec{v_1} & — \\ — & \vec{v_2} & — \\ & \vdots & \\ — & \vec{v_m} & — \end{bmatrix}$$

If Gram-Schmidt seems like too much trouble, there's another way to find the projection matrix. Before we get to it, let's talk about <u>least-squares solutions</u>.

Setup: We have a system of equations:

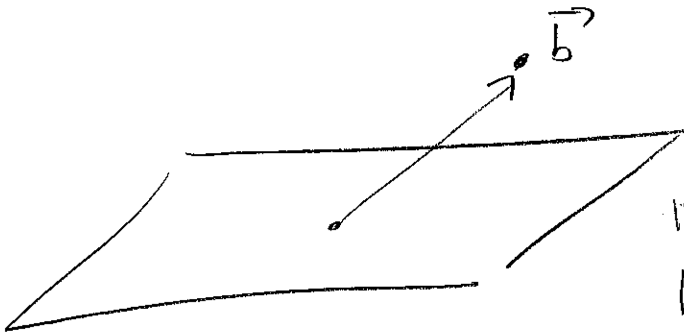
$A\vec{x} = \vec{b}$. (A, $\vec{b}$ are given and we want to find $\vec{x}$).

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}$$

A is $n \times m$
(n equations, m variables).

$A\vec{x} = \vec{b}$ is the same as

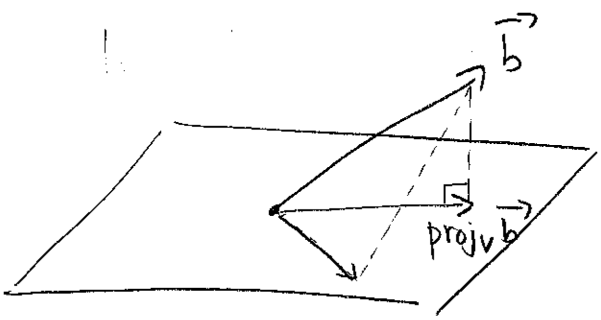
$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_m\vec{v}_m = \vec{b}.$$



$\text{im}(A) = \text{span}(\vec{v}_1, \ldots, \vec{v}_m)$.
Let $V = \text{im}(A)$.

If $\vec{b} \notin \text{im}(A)$ then the system has no solutions

The best we can do is find $\vec{x}^*$ which minimizes $\|A\vec{x} - \vec{b}\|$.



This is the same as solving $A\vec{x}^* = \text{proj}_V\vec{b}$ for $\vec{x}^*$.

Any $\vec{x}^*$ which solves this is called a least-squares solution of $A\vec{x} = \vec{b}$.

(since we're minimizing $\|A\vec{x} - \vec{b}\| = \sqrt{(\cdot)^2 + (\cdot)^2 + \cdots + (\cdot)^2}$)

How to find $\vec{x}^*$? Key observation:

$\vec{b}^\perp = \vec{b} - \operatorname{proj}_V \vec{b}$ is orthogonal to

$$V = \operatorname{im}(A) = \operatorname{span}(\vec{v}_1, \dots, \vec{v}_m).$$

so $\vec{v}_1 \cdot \vec{b}^\perp = 0, \dots, \vec{v}_m \cdot \vec{b}^\perp = 0$.

so
$$\begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix} \begin{bmatrix} | \\ \vec{b}^\perp \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

i.e. $$A^T \vec{b}^\perp = \vec{0}.$$

But $A^T \vec{b}^\perp = A^T(\vec{b} - \operatorname{proj}_V \vec{b}) = A^T(\vec{b} - A\vec{x}^*)$.
$$= A^T\vec{b} - A^T A \vec{x}^*$$

so: $$\boxed{A^T A \vec{x}^* = A^T \vec{b}}$$

Therefore, the least-squares solutions to $A\vec{x} = \vec{b}$ are the exact solutions to $\underbrace{A^T A \vec{x} = A^T \vec{b}}$.

"the normal equation of $A\vec{x} = \vec{b}$"

$(A^TA)\vec{x} = A^T\vec{b}$ ← properties of the normal equation:

1. A is $n \times m$ $\quad A^T$ is $m \times n$

$\Rightarrow A^TA$ is $m \times m$.

This is a system of m equations in m variables.

2. This system is consistent (i.e. always has at least one solution).

3. If the columns of A are lin. indep. (i.e. if $\ker(A) = \{\vec{0}\}$) then A^TA is invertible, so there is a <u>unique</u> least-squares solution $\boxed{\vec{x}^* = (A^TA)^{-1}A^T\vec{b}}$

Recall that $A\vec{x}^* = \text{proj}_V \vec{b}$ $\quad (V = \text{im} A)$

So $\text{proj}_V \vec{b} = A(A^TA)^{-1}A^T\vec{b}$.

So: If V is a subspace of $\mathbb{R}^n$ and $\vec{v_1}, \ldots, \vec{v_m}$ is a basis of V, the matrix for the orthogonal projection onto V is

$$A(A^TA)^{-1}A^T, \quad \text{where } A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}.$$

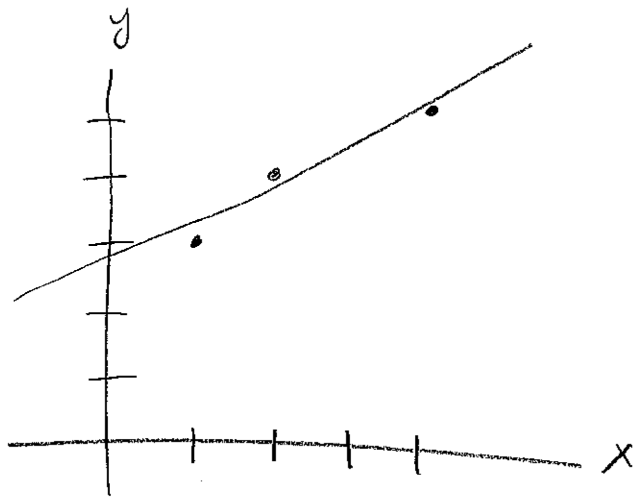
Lecture 15

11/20/18

Application to data fitting.

Example:

inches of snow (x)	cups of coffee sold (y)
1	3
2	4
4	5



Q: Can we find a line which approximates our data well? $y = c_0 + c_1 x$ is the general form for a line.

The 3 data points give us:

$$\left|\begin{array}{l} c_0 + c_1 = 3 \\ c_0 + 2c_1 = 4 \\ c_0 + 4c_1 = 5 \end{array}\right| \longrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}.$$

3 equations in 2 variables, $\longrightarrow$ system may not be consistent.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}.$$

The least squares solution of $A\vec{x} = \vec{b}$ is.

$$\vec{x}^* = (A^T A)^{-1} A^T \vec{b} = \begin{bmatrix} 5/2 \\ 9/14 \end{bmatrix} \approx \begin{bmatrix} 2.5 \\ 0.643 \end{bmatrix}$$

so the line is $y = 2.5 + 0.643x = f(x)$

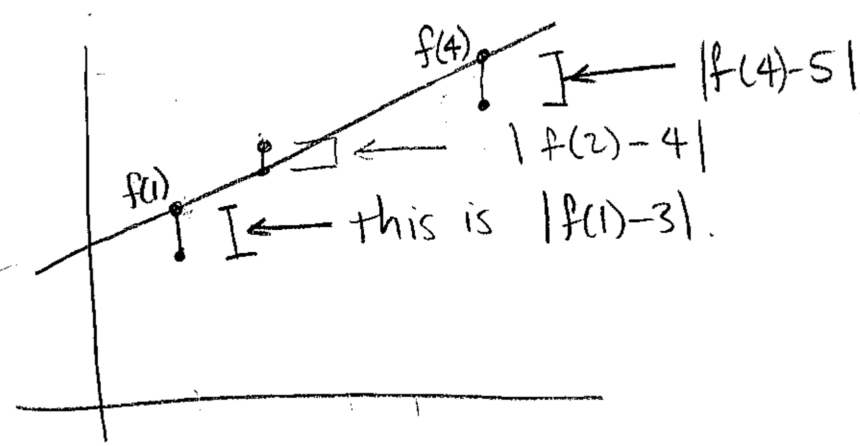
Recall $\vec{x}^*$ is the minimizer of $\|A\vec{x} - \vec{b}\|^2$.

$$A\vec{x}^* = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} c_0 + c_1 \\ c_0 + 2c_1 \\ c_0 + 4c_1 \end{bmatrix}.$$

"predicted values"

$$\|A\vec{x}^* - \vec{b}\|^2 = \left\| \begin{bmatrix} f(1) - 3 \\ f(2) - 4 \\ f(4) - 5 \end{bmatrix} \right\|^2$$

$$= [f(1)-3]^2 + [f(2)-4]^2 + [f(4)-5]^2.$$



You can also use this method to fit data to quadratics. ($y = c_0 + c_1 x + c_2 x^2$), etc.

Now, we move to Chapter 6: <u>Determinants</u>

- Def of determinant for 1×1, 2×2, 3×3 matrices.

1×1: $$\det\begin{bmatrix} a \end{bmatrix} = a.$$

2×2: $$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

3×3: $$\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$ ← to calculate this, use Sarrus's rule:

$$-gec - hfa - idb.$$

$$\left[\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \end{array}\right]\begin{array}{cc} a & b \\ d & e \\ g & h \end{array}$$

$$+aei + bfg + cdh$$

$$\det\begin{bmatrix} \quad & \quad & \quad \end{bmatrix} = aei + bfg + cdh - gec - hfa - dbi$$

4×4 and above: too complicated.

The main property of determinants that we need: A square matrix A is invertible if and only if $\det A \neq 0$.

This is all you need to know about determinants for this course

Some other properties/uses of determinants.

(1). Cramer's rule: formula for solution to a system of n eqs in n vars.

$\longrightarrow$ not very useful in practice (Gauss-Jordan elimination is better)

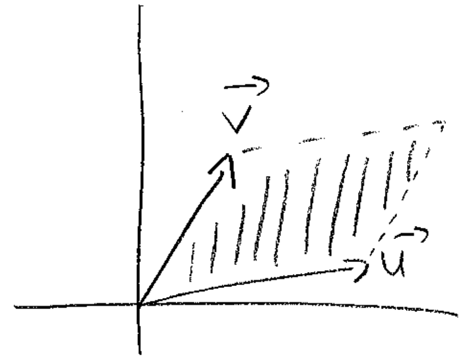
(2). Determinants show up in the formula for the inverse of a matrix

$\longrightarrow$ not very useful in practice (Gauss-Jordan elimination is better)

3 Area and volume

If $\vec{u}, \vec{v} \in \mathbb{R}^2$,

$$A = \begin{bmatrix} | & | \\ \vec{u} & \vec{v} \\ | & | \end{bmatrix}.$$



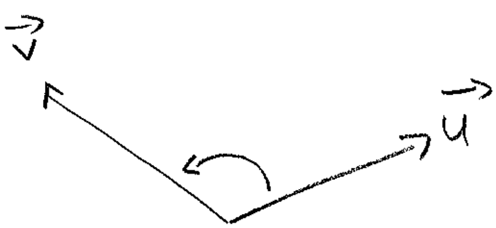
then • $|\det A| =$ area of parallelogram spanned by $\vec{u}, \vec{v}$.

• $\det A > 0$ if $\vec{u}$ moves to $\vec{v}$ in counterclockwise direction

$\det A < 0$ if $\vec{u}$ - - - - - - clockwise direction

Examples:

①



$\det\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} > 0$

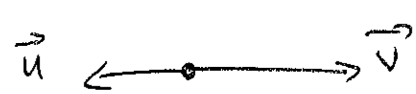
②

$\det\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} < 0$.

③

$\det\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} = 0$

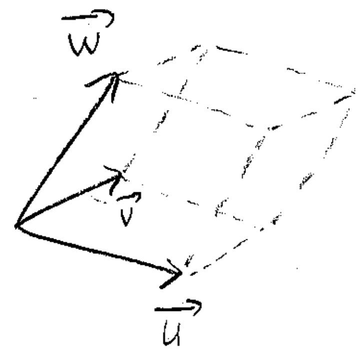
④



$\det\begin{bmatrix} \vec{u} & \vec{v} \end{bmatrix} = 0$

In $\mathbb{R}^3$: if $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, then

$$\left| \det \begin{bmatrix} | & | & | \\ \vec{u} & \vec{v} & \vec{w} \\ | & | & | \end{bmatrix} \right| = \text{area of parallelepiped spanned by } \vec{u}, \vec{v}, \vec{w}$$



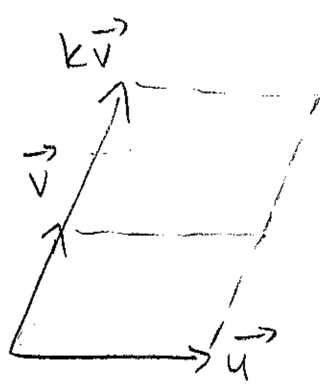
The sign of $\det[\vec{u}\ \vec{v}\ \vec{w}]$ is given by the "right-hand rule"

Back to $\mathbb{R}^2$

How to find the signed area of the parallelogram spanned by $\vec{u}, \vec{v}$?

Define $f(\vec{u}, \vec{v})$ = signed area.

Let's study properties of f.

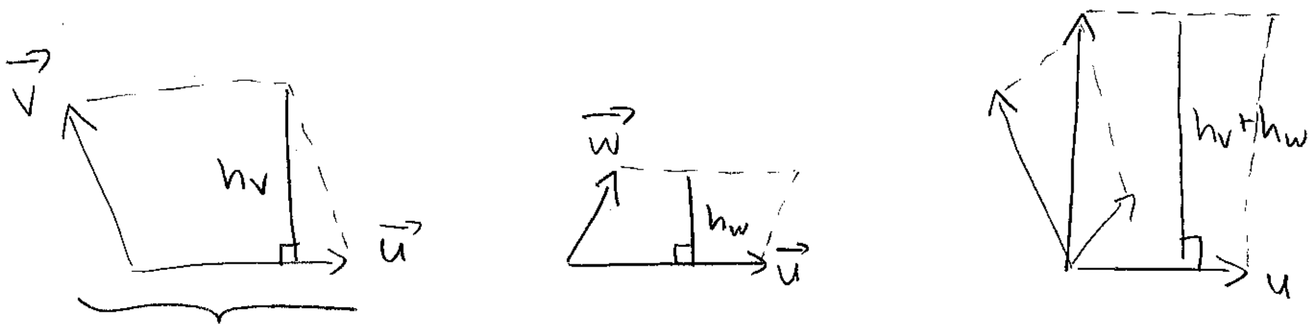


use the formula

Area = base · height

the base ($\vec{u}$) is unchanged. The height is scaled by k.

$$f(\vec{u}, k\vec{v}) = k f(\vec{u}, \vec{v})$$



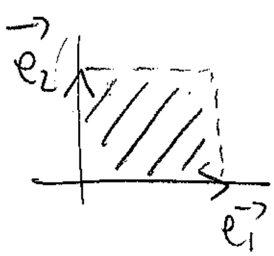
$$f(\vec{u},\vec{v}) \quad + \quad f(\vec{u},\vec{w}) \quad = \quad f(\vec{u}, \vec{v}+\vec{w}).$$

So: ①

$$\left.\begin{aligned} &f(\vec{u}, k\vec{v}) = k f(\vec{u},\vec{v}) \\ &f(\vec{u}, \vec{v}+\vec{w}) = f(\vec{u},\vec{v}) + f(\vec{u},\vec{w}). \\ &f(k\vec{u},\vec{v}) = k f(\vec{u},\vec{v}). \\ &f(\vec{u}+\vec{w},\vec{v}) = f(\vec{u},\vec{v}) + f(\vec{w},\vec{v}). \end{aligned}\right\} \text{ f is \underline{bilinear}.}$$

② $f(\vec{u},\vec{u}) = 0$. $\longleftarrow$ f is <u>alternating</u>.

③. $f(\vec{e_1}, \vec{e_2}) = 1$
(normalization).



These properties uniquely determine the function f.

First note that for any $\vec{u}, \vec{v}$,

$$\begin{aligned} 0 &= f(\vec{u}+\vec{v}, \vec{u}+\vec{v}) \\ &= f(\vec{u}, \vec{u}+\vec{v}) + f(\vec{v}, \vec{u}+\vec{v}) \\ &= \left(f(\vec{u}, \vec{u}) + f(\vec{u}, \vec{v})\right) + \left(f(\vec{v}, \vec{u}) + f(\vec{v}, \vec{v})\right) \\ &= 0 + f(\vec{u}, \vec{v}) + f(\vec{v}, \vec{u}) + 0. \end{aligned}$$

So $f(\vec{u}, \vec{v}) = -f(\vec{v}, \vec{u})$

which is why 2 is called "alternating".

Then $f\left(\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}\right)$.

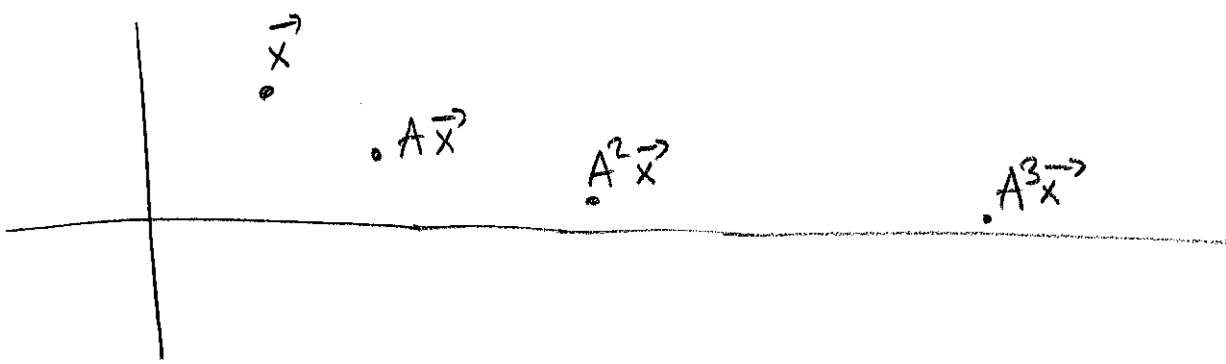
$$\begin{aligned} &= f(a\vec{e_1} + c\vec{e_2}, b\vec{e_1} + d\vec{e_2}) \\ &= f(a\vec{e_1}, b\vec{e_1}) + f(a\vec{e_1}, d\vec{e_2}) \\ &\qquad + f(c\vec{e_2}, b\vec{e_1}) + f(c\vec{e_2}, d\vec{e_2}). \\ &= ab\, f(\vec{e_1}, \vec{e_1}) + ad\, f(\vec{e_1}, \vec{e_2}) \\ &\qquad + cb\, f(\vec{e_2}, \vec{e_1}) + cd\, f(\vec{e_2}, \vec{e_2}). \\ &= 0 + ad + bc(-1) + 0. = ad - bc. \end{aligned}$$

Chapter 7: We'll apply linear algebra to study dynamical systems and Markov chains.

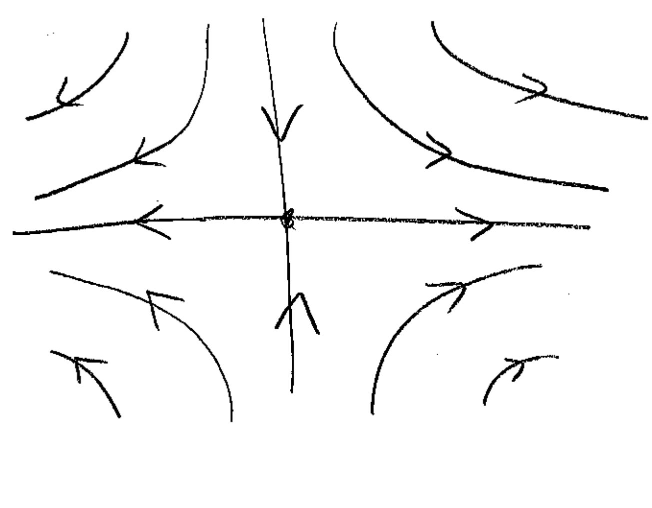
First, let's look at diagonal matrices

e.g. $A = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \longrightarrow \begin{array}{l} A\vec{e_1} = 2\vec{e_1} \\ A\vec{e_2} = 0.5\vec{e_2} \end{array}$

A stretches by 2 in $\vec{e_1}$ direction and by 0.5 in $\vec{e_2}$ direction



We can use a "phase portrait" to illustrate how points move as we apply A over and over again



$$\begin{cases} A\vec{e_1} = 2\vec{e_1} \\ A\vec{e_2} = 0.5\vec{e_2} \end{cases} \Longrightarrow \begin{cases} A^n\vec{e_1} = 2^n\vec{e_1} \\ A^n\vec{e_2} = 0.5^n e_2 \end{cases} \Longrightarrow A^n = \begin{bmatrix} 2^n & 0 \\ 0 & 0.5^n \end{bmatrix}$$

(It is easy to raise a diagonal matrix to some power.)

<u>Example</u>:

$c(t)$ = population of coyotes t years from now

$r(t)$ = - - - - - - - roadrunners - - - - - - - -

Suppose that

$$\begin{cases} c(t+1) = 0.86c(t) + 0.08r(t) \\ r(t+1) = -0.12c(t) + 1.14r(t). \end{cases}$$

(The textbook asks: significance of these coeffs?)

Define: $\vec{x}(t) = \begin{bmatrix} c(t) \\ r(t) \end{bmatrix}$ the "state vector"

$$A = \begin{bmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{bmatrix}$$

Then $\boxed{\vec{x}(t+1) = A\vec{x}(t)}$

We can write $\vec{x}(t)$ in terms of $\vec{x}(0)$:

$$\boxed{\vec{x}(t) = A^t \vec{x}(0)}$$

$$\left(\vec{x}(0) \xrightarrow{A} \vec{x}(1) \xrightarrow{A} \vec{x}(2) \xrightarrow{A} \vec{x}(3) \xrightarrow{A} \cdots \right).$$

A is not a diagonal matrix so finding A^t is not as easy before.

<u>Case 1</u>: $\vec{x}(0) = \vec{x}_0 = \begin{bmatrix} 100 \\ 300 \end{bmatrix}$. $\left(\begin{array}{l} \text{100 coyotes} \\ \text{300 roadrunners} \end{array} \right)$.

$$\Rightarrow \vec{x}(1) = A\vec{x}_0 = \begin{bmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{bmatrix} \begin{bmatrix} 100 \\ 300 \end{bmatrix}$$

$$= \begin{bmatrix} 110 \\ 330 \end{bmatrix} = 1.1\,\vec{x}_0$$

Cool! A stretches $\vec{x}_0$ by a factor of 1.1.

So $\vec{x}(t) = A^t \vec{x}_0 = (1.1)^t \vec{x}_0$.

$$\Rightarrow \begin{array}{l} c(t) = 100\,(1.1)^t \\ r(t) = 300\,(1.1)^t \end{array}$$

Done!

Case 2: $\vec{x}_0 = \begin{bmatrix} 200 \\ 100 \end{bmatrix}$

$$A\vec{x}_0 = \begin{bmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{bmatrix} \begin{bmatrix} 200 \\ 100 \end{bmatrix} = \begin{bmatrix} 180 \\ 90 \end{bmatrix} = 0.9\vec{x}_0.$$

$$\Rightarrow \quad \vec{x}(t) = A^t \vec{x}_0 = (0.9)^t \vec{x}_0$$

$$\Rightarrow \quad c(t) = 200(0.9)^t$$
$$r(t) = 100(0.9)^t$$

Case 3: $\vec{x}_0 = \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}$.

$A\vec{x}_0 = \begin{bmatrix} 940 \\ 1020 \end{bmatrix} \leftarrow$ Not a scalar multiple of $\vec{x}_0$. ☹.

What do we do here?

Key idea: let $\vec{v}_1 = \begin{bmatrix} 100 \\ 300 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 200 \\ 100 \end{bmatrix}$.

We know what A does to $\vec{v}_1$ and $\vec{v}_2$.

So let's write $\vec{x}_0 = \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}$ as a linear combination of $\vec{v}_1, \vec{v}_2$.

$$\vec{x_0} = 2\vec{v_1} + 4\vec{v_2}$$

So $\vec{x}(t) = A^t\vec{x_0} = A^t(2\vec{v_1} + 4\vec{v_2}) = 2A^t\vec{v_1} + 4A^t\vec{v_2}$

$$= 2(1.1)^t\vec{v_1} + 4(0.9)^t\vec{v_2}.$$

$$= 2(1.1)^t\begin{bmatrix}100\\300\end{bmatrix} + 4(0.9)\begin{bmatrix}200\\100\end{bmatrix}.$$

$$\Rightarrow \begin{cases} c(t) = 200(1.1)^t + 800(0.9)^t \\ r(t) = 600(1.1) + 400(0.9)^t. \end{cases}$$

Done!

Another way of getting this answer. (Well, it's really the same thing...): Let $\mathfrak{B} = (\vec{v_1}, \vec{v_2})$.

$S = \begin{bmatrix} | & | \\ \vec{v_1} & \vec{v_2} \\ | & | \end{bmatrix}$

$$\begin{cases} A\vec{v_1} = 1.1\vec{v_1} \\ A\vec{v_2} = 0.9\vec{v_2}. \end{cases} \Longrightarrow [A\vec{x}]_{\mathfrak{B}} = \begin{bmatrix}1.1 & 0\\0 & 0.9\end{bmatrix}[\vec{x}]_{\mathfrak{B}}$$

Let $B = \begin{bmatrix}1.1 & 0\\0 & 0.9\end{bmatrix}$. Then $A = SBS^{-1}$.

$$A^t = (SBS^{-1})^t = \underbrace{(SBS^{-1})(SBS^{-1})\cdots(SBS^{-1})}_{t \text{ times}}.$$

$$= SB^tS^{-1}$$

since $S^{-1}S = I$

Then $\vec{X}(t) = A^t \vec{X_0} = S B^t S^{-1} \vec{X_0}$

$$= \begin{bmatrix} 100 & 200 \\ 300 & 100 \end{bmatrix} \begin{bmatrix} (1.1)^t & 0 \\ 0 & (0.9)^t \end{bmatrix} \underbrace{\begin{bmatrix} 100 & 200 \\ 300 & 100 \end{bmatrix}^{-1} \begin{bmatrix} 1000 \\ 1000 \end{bmatrix}}_{\begin{bmatrix} 2 \\ 4 \end{bmatrix}}$$

(same as saying $\vec{X_0} = 2\vec{v_1} + 4\vec{v_2}$) $\rightarrow \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

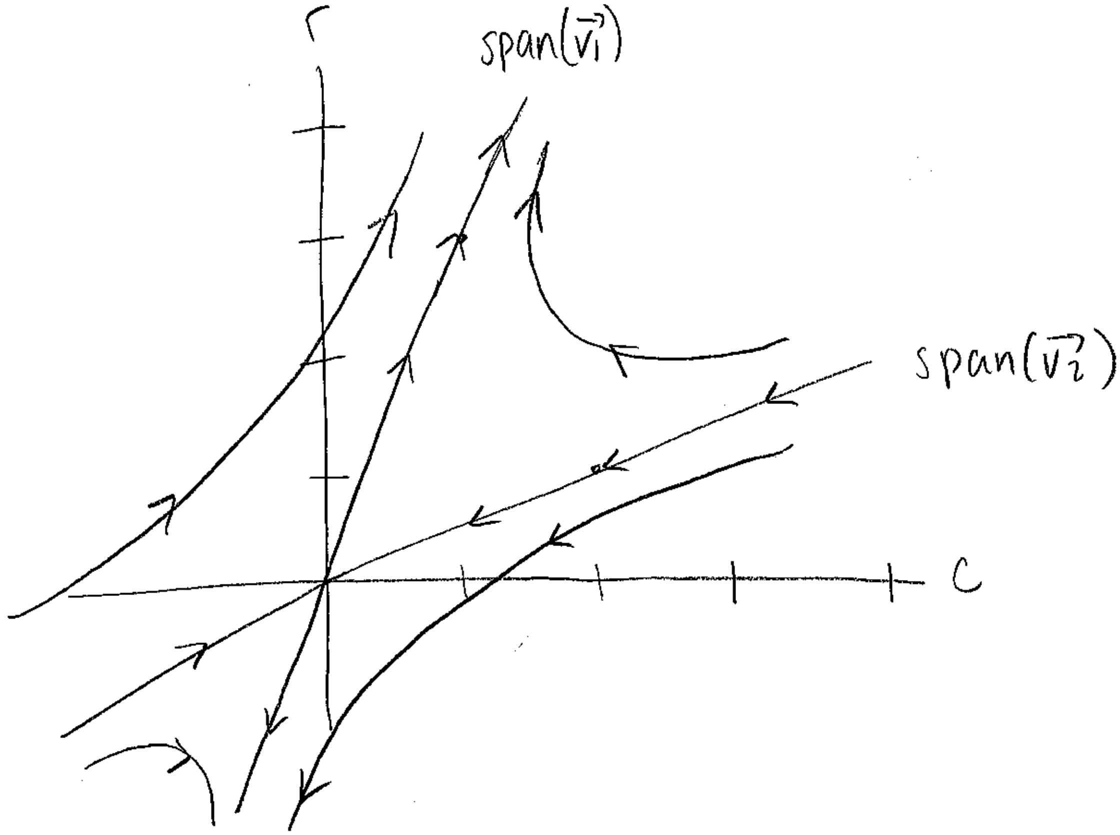
$$= \begin{bmatrix} 100 & 200 \\ 300 & 100 \end{bmatrix} \begin{bmatrix} (1.1)^t & 0 \\ 0 & (0.9)^t \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 100 & 200 \\ 300 & 100 \end{bmatrix} \begin{bmatrix} 2(1.1)^t \\ 4(0.9)^t \end{bmatrix}.$$

$$= 2(1.1)^t \begin{bmatrix} 100 \\ 300 \end{bmatrix} + 4(0.9)^t \begin{bmatrix} 200 \\ 100 \end{bmatrix}.$$

Same answer as before!

Now let's draw the phase portrait of this system.



(It's like the $\begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$ matrix from the beginning, but now, the axes that are being stretched are not the x and y axes)

The thing that allowed us to find A^t was that we found vectors $\vec{v}$ which satisfy.

$$\underbrace{A\vec{v} = \lambda\vec{v}}_{\text{A stretches } \vec{v} \text{ by a factor of } \lambda.} \quad \text{for some scalar } \vec{v}$$

<u>Definition</u>: Let $T(\vec{x}) = A\vec{x}$ be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$.

A nonzero vector $\vec{v}$ is called an <u>eigenvector</u> of A if $A\vec{v} = \lambda\vec{v}$ for some scalar λ.
This λ is called the <u>eigenvalue</u> associated with eigenvector $\vec{v}$.

A basis $\vec{v_1}, \dots, \vec{v_n}$ of $\mathbb{R}^n$ is called an <u>eigenbasis</u> of A if the vectors $\vec{v_1}, \dots, \vec{v_n}$ are eigenvectors of A.

<u>Example</u>: $A = \begin{bmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{bmatrix}$.

$$\vec{v_1} = \begin{bmatrix} 100 \\ 300 \end{bmatrix} \qquad \vec{v_2} = \begin{bmatrix} 200 \\ 100 \end{bmatrix}.$$

Then: $\vec{v_1}$ is an evec of A with eval 1.1

$\vec{v_2}$ - - - - - - - - - - - - 0.9

$\vec{v_1}, \vec{v_2}$ is an ebasis of A.

Suppose
- A is $n\times n$.
- $A\vec{v_1} = \lambda_1 \vec{v_1}$ $\quad \dots \quad$ $A\vec{v_n} = \lambda_n \vec{v_n}$
- $\vec{v_1}, \dots, \vec{v_n}$ is a basis of $\mathbb{R}^n$.

Let $\mathcal{B} = (\vec{v_1}, \dots, \vec{v_n})$ $\qquad S = \begin{bmatrix} | & & | \\ \vec{v_1} & \cdots & \vec{v_n} \\ | & & | \end{bmatrix}$.

Then $$[A\vec{x}]_{\mathcal{B}} = \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{B} [\vec{x}]_{\mathcal{B}}.$$

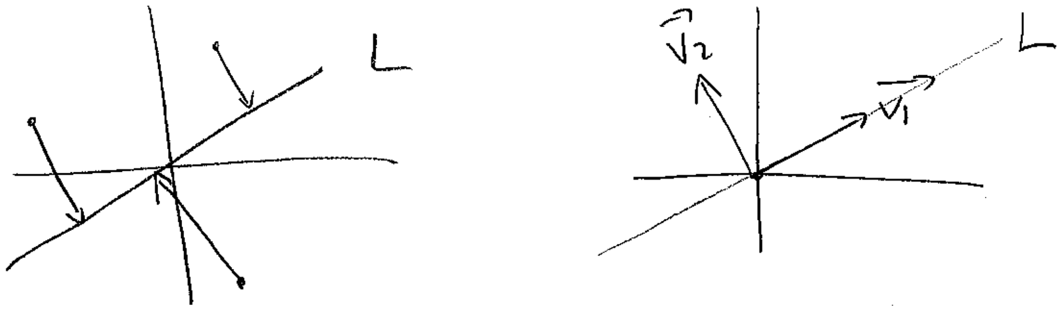
$$\Rightarrow \begin{cases} A = SBS^{-1} \\ S^{-1}AS = B. \end{cases}$$

(We say $\overset{n\times n}{A}$ is <u>diagonalizable</u> is we can find $n\times n$ matrices B, S such that.

- B is a diagonal matrix
- S is invertible
- $A = SBS^{-1}$.)

Lecture 17: $T: \mathbb{R}^2 \to \mathbb{R}^2$ 11/29/18

<u>Example</u>: $T(\vec{x})$ = orthog. proj. onto line L.



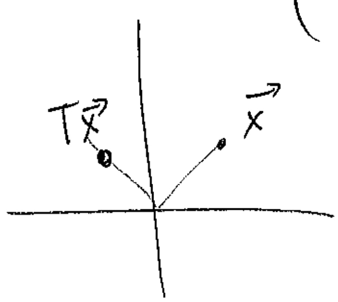
$\vec{v}_1$ is an evec of T with eval 1.

$\vec{v}_2$ - - - - - - - - - - - - 0.

($\vec{v}_1, \vec{v}_2$ is an eigenbasis).

<u>Example</u>: $T: \mathbb{R}^2 \to \mathbb{R}^2$. $90°$ counterclockwise rotation.

$$\left(A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right)$$



$T(\vec{x})$ is never a scalar multiple of $\vec{x}$.

So: T has <u>no</u> $\boxed{\text{real}}$ evecs or evals.

(But! $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix}$.

$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -i \end{bmatrix} = -i\begin{bmatrix} -i \\ 1 \end{bmatrix}$. Not need for this class.)

Example:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$$

$\vec{e_1}$ is evec w/ eval 2
$\vec{e_2}$ is evec w/ eval 0.5
($\vec{e_1}, \vec{e_2}$ is an eigenbasis).

Example:

$$A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

For any $\vec{v} \in R^2$,
$A\vec{v} = 1\vec{v}$

so all nonzero vecs in $\mathbb{R}^2$ are evecs with eval 1.

Any basis for $\mathbb{R}^2$ is an eigenbasis of A.

How to find evals and evecs of a matrix

First: consider a square matrix A ($n \times n$) and a nonzero vec $\vec{v} \in \mathbb{R}^n$.

The sentence "$\vec{v}$ is an evec of A w/ eval 0"

is the same as "$A\vec{v} = 0\vec{v}$."

i.e. "$A\vec{v} = \vec{0}$".

i.e. "$\vec{v} \in \ker A$."

So "$\vec{v}$ is an evec of A w/ eval 0"
is the same as "$\vec{v} \neq \vec{0}$ and $\vec{v} \in \ker A$"

So 0 is an eval of A if and only if $\ker A \neq \{\vec{0}\}$. (ie. ker A has something other than $\vec{0}$.)

But recall: ① $\ker A \neq \{\vec{0}\}$ if and only if A is not invertible.

② A is not invertible if and only if $\det A = 0$.

Put all this together to get:

$$\boxed{\text{0 is an eval of A if and only if } \det A = 0.}$$

Next question.

λ is an eval of A if and only if ???

$A\vec{v} = \lambda\vec{v}$

$A\vec{v} - \lambda\vec{v} = 0$

$A\vec{v} - \lambda I_n\vec{v} = 0$

$(A - \lambda I_n)\vec{v} = 0$

$\vec{v}$ is an evec of A w/ eval λ if and only if $\vec{v} \neq \vec{0}$ and $\vec{v} \in \ker(A - \lambda I_n)$.

$$\left(\lambda I_n = \begin{bmatrix} \lambda & & 0 \\ & \lambda & \\ 0 & & \ddots \; \lambda \end{bmatrix} \right)$$

Def: $E_\lambda = \ker(A - \lambda I_n)$ is called the <u>eigenspace</u> associated with λ.

As before, λ is an eval of A if and only if $\det(A - \lambda I_n) = 0$.

This gives a way of computing evals and evecs:

<u>Example</u>: $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

The eigenvalues are the solutions to

$\det(A - \lambda I_2) = 0$.

$$A-\lambda I_2 = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{bmatrix}.$$

$$\begin{aligned} \det(A-\lambda I_2) &= (1-\lambda)(3-\lambda) - 2\cdot 4 \\ &= 3 - 4\lambda + \lambda^2 - 8 \\ &= \lambda^2 - 4\lambda - 5. \\ &= (\lambda-5)(\lambda+1). \end{aligned}$$

so the evals of A are 5 and -1.

$$A-5I_2 = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix}.$$

$$E_5 = \ker(A-5I_2) = \operatorname{span}\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

$\Longrightarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an evec of A with eval 5.

$$A-(-1)I_2 = \begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}.$$

$$E_{(-1)} = \ker(A+I_2) = \operatorname{span}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$\Longrightarrow \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an evec of A with eval -1.

$\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}$ is an eigenbasis of A.

Let $S = \begin{bmatrix}1 & 1\\2 & -1\end{bmatrix}$ $B = \begin{bmatrix}5 & 0\\0 & -1\end{bmatrix}$.

Then $A = SBS^{-1}$.

Let's check this to make sure.

$$AS = \begin{bmatrix}1 & 2\\4 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\2 & -1\end{bmatrix} = \begin{bmatrix}5 & -1\\10 & 1\end{bmatrix}$$

equal

$$SB = \begin{bmatrix}1 & 1\\2 & -1\end{bmatrix}\begin{bmatrix}5 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}5 & -1\\10 & 1\end{bmatrix}$$

In general, to find the evals and evecs of A:

Step 1: Solve $\det(A - \lambda I_n) = 0$ to get the eigenvalues.

Step 2: The eigenvectors w/ eval λ are the nonzero vectors in $E_\lambda = \ker(A - \lambda I_n)$.

If we want to diagonalize A, then:

<u>Step 3</u>: For each eval λ, find a basis of E_λ

<u>Step 4</u>: After doing this for each λ:

- if we have n vectors total, then they are an eigenbasis
- if we have $< n$ vectors, then A does not have an eigenbasis.

<u>Example</u>

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$A - \lambda I_3 = \begin{bmatrix} 1-\lambda & 1 & 1 \\ 0 & -\lambda & 1 \\ 0 & 0 & 1-\lambda \end{bmatrix}.$$

$$\det(A - \lambda I_3) = (1-\lambda)(-\lambda)(1-\lambda).$$

$$= -\lambda(\lambda - 1)^2$$

$\Longrightarrow$ evals are 1 and 0.

$$E_1 = \ker(A - I_3) = \ker \begin{bmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \operatorname{span} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

$$E_0 = \ker A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \operatorname{span} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an evec w/ eval 1

$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ is an evec w/ eval 0.

A has no eigenbasis and is not diagonalizable.

<u>Example</u>: $A = I_2$.

$$\det(A - \lambda I_2) = \det \begin{bmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2.$$

$\Rightarrow$ 1 is the only eval of A.

$$E_1 = \ker(A - I_2) = \ker \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbb{R}^2$$

So (as we've already seen) every nonzero vec of $\mathbb{R}^2$ is an evec.

Example

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

$$\det(A - \lambda I_2) = \det\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

$\Longrightarrow$ A has no real eigenvalues or eigenvectors.

Fact (Theorem 7.3.4).

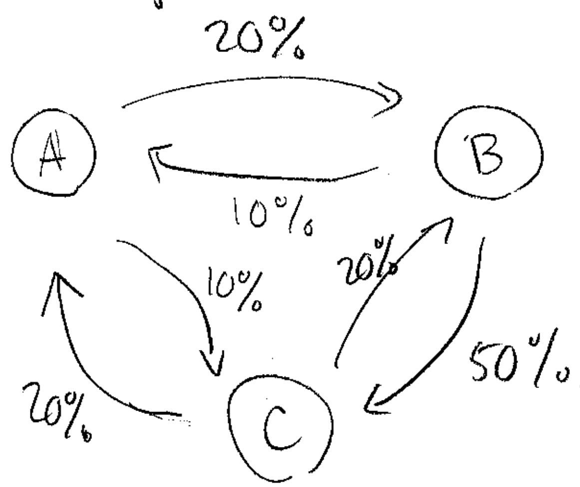
If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

This is because each eigenspace has at least one nonzero vector, so we are guaranteed to find n eigenvectors.

(This is a useful fact, but you don't need it for this course.)

Application of linear algebra to Markov chains:

<u>Example</u>: Suppose there are 3 cities and each year, people move between the cities according to the following "state diagram":



e.g. at the end of each year,

- 20% of people in city A move to city B
- 10% of people in city A move to city C.
- The remaining (70%) people in city A stay there.

let $a(t)$ = proportion of people in city A in year t
$b(t)$ = — — — — — B - - -
$c(t)$ = — — — — — C - - -

(By proportion I mean $a(t)+b(t)+c(t)=1$)

Then
$$a(t+1) = 0.7\,a(t) + 0.1\,b(t) + 0.2\,c(t)$$
$$b(t+1) = 0.2\,a(t) + 0.4\,b(t) + 0.2\,c(t)$$
$$c(t+1) = 0.1\,a(t) + 0.5\,b(t) + 0.6\,c(t)$$

let $\vec{x}(t) = \begin{bmatrix} a(t) \\ b(t) \\ c(t) \end{bmatrix}$ $\quad A = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.4 & 0.2 \\ 0.1 & 0.5 & 0.6 \end{bmatrix}$.

Then $\vec{x}(t+1) = A\vec{x}(t)$. and $\vec{x}(t) = A^t\vec{x}(0)$.

<u>Q</u>: If $\vec{x}(0) = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$ then what will happen many years from now?

e.g. $\vec{x}(1) = A\vec{x}(0) = \begin{bmatrix} 5/15 \\ 4/15 \\ 6/15 \end{bmatrix}$.

Note that the 3 components of $\vec{x}(1)$ add up ⤳

to 1, as they should (since they represent proportions)

The same is also true for $\vec{x}(t)$.

Anyways, we can do what we did before. find evecs and evals:

evecs: $\vec{v}_1 = \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ $\vec{v}_3 = \begin{bmatrix} -1 \\ -3 \\ 4 \end{bmatrix}$

w/ evals: $\lambda_1 = 1$ $\lambda_2 = 0.5$ $\lambda_3 = 0.2$.

Let $S = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$ $B = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}$

Then $$A = SBS^{-1}$$

$$A^t \vec{x}(0) = SB^t S^{-1} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \qquad S^{-1} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/20 \\ -2/45 \\ 1/36 \end{bmatrix}.$$

$$= \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 1 & & \\ & 0.5^t & \\ & & 0.2^t \end{bmatrix} \begin{bmatrix} 1/20 \\ -2/45 \\ 1/36 \end{bmatrix}.$$

$$= \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 1/20 \\ -2/45 \cdot (0.5)^t \\ 1/36 \cdot (0.2)^t \end{bmatrix}$$

$$= \frac{1}{20}\vec{v}_1 - \frac{2}{45}(0.5)^t \vec{v}_2 + \frac{1}{36}(0.2)^t \vec{v}_3 .$$

Observe that $\lim\limits_{t \to \infty} A^t \vec{x}(0) = \frac{1}{20}\vec{v}_1 = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}.$

In fact, no matter what your initial distribution vector $\vec{x}(0)$ is,

$$\lim_{t \to \infty} A^t \vec{x}(0) = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}.$$

so we call this the <u>equilibrium distribution</u> of A

$$\vec{X}_{equ} = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}.$$

Why is this true?

$$\lim_{t \to \infty} B^t = \lim_{t \to \infty} \begin{bmatrix} 1 & & \\ & 0.5^t & \\ & & 0.2^t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

so $\lim_{t\to\infty} S B^t S^{-1} \vec{x}(0) = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix} S^{-1} \vec{x}(0)$

$$= \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{0} & \vec{0} \\ | & | & | \end{bmatrix} S^{-1} \vec{x}(0).$$

$$= \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{0} & \vec{0} \\ | & | & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = c_1 \vec{v}_1 = c_1 \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix}$$

But the components need to add up to 1, so $c_1 = 1/20$. (we don't even need to calculate $S^{-1}\vec{x}(0)$ to know this must be true! Also, this must be true for every $\vec{x}(0)$!)

__Def__: A vector $\vec{x} \in \mathbb{R}^n$ is a __distribution vector__ if

- its components add up to 1
- all components are ≥ 0.

Def: A square matrix A is a **transition matrix** (or stochastic matrix, or Markov matrix) if

- all entries are ≥ 0.
- the entries in each column add up to 1.

Another example:

$$(A) \xrightarrow{50\%} (B)$$

Does this have an equilibrium?

$$a(t+1) = 0.5a(t)$$
$$b(t+1) = 0.5a(t) + b(t)$$

$$\vec{x}(t+1) = \underbrace{\begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix}}_{A} \vec{x}(t)$$

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

$$\lambda_1 = 1 \qquad \lambda_2 = 0.5$$

Using the same reasoning as before,

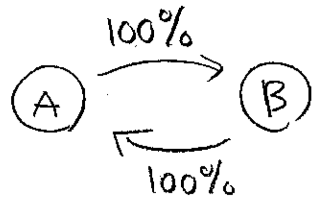
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is the equilibrium distribution

for A, i.e.

$$\lim_{t \to \infty} A^t \vec{x}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$ no matter what

the initial distribution is.

<u>Example</u>:



$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Suppose $\vec{x}(0) = \begin{bmatrix} 1/4 \\ 3/4 \end{bmatrix}$.

Then $\vec{x}(1) = \begin{bmatrix} 3/4 \\ 1/4 \end{bmatrix}$, $\vec{x}(2) = \begin{bmatrix} 1/4 \\ 3/4 \end{bmatrix}$, ...

$$\lim_{t \to \infty} A^t \begin{bmatrix} 1/4 \\ 3/4 \end{bmatrix}$$ <u>does not exist</u>.

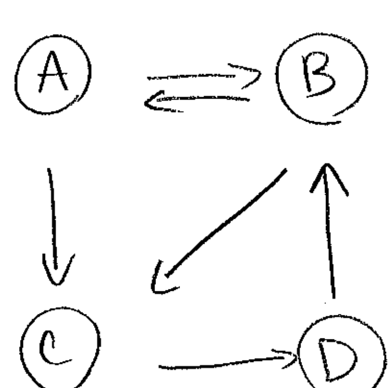
This matrix does not have an equilibrium distribution.

Theorem: Let A be an $n \times n$ transition matrix and suppose all the entries of A are positive. Then.

1. there exists exactly one distribution vector $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{x}$. Call this vector $\vec{x}_{equ}$.
2. For any initial distribution $\vec{x}_0$,

$$\lim_{t \to \infty} A^t \vec{x}_0 = \vec{x}_{equ}$$

Application of Markov chains to PageRank (alg. used by Google to rank websites. their first alg.):



Idea: suppose you start at some page. After each minute, you click on a random link on that page.

After a long time, what is the probability you are at page A? B? C? D?

$$A = \begin{array}{c} \text{(start)} \\ \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ \\ A \\ B \\ C \\ D \end{array} \text{(end)}$$

(Higher prob = many pages linking to it = page is important)

We would like to apply the theorem to find an equilibrium distribution, but we can't. The matrix has zeros.

To fix this, consider.

$$\tilde{A} = 0.9A + 0.1 \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

"random teleporting"

$\tilde{A}$ has an equilibrium distribution!

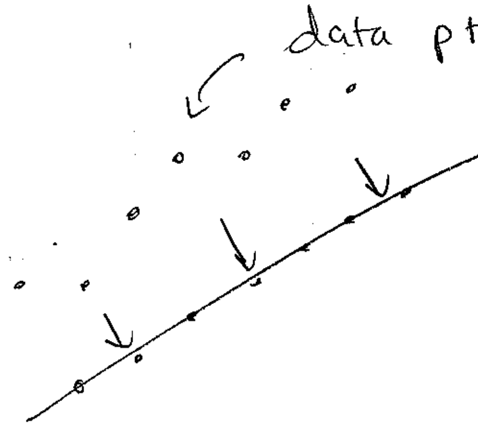
Just compute $\ker(\tilde{A} - I)$ to get the equilibrium distribution.

This is the idea behind PageRank!

That's it for this class!

Some other topics/applications of linear algebra (just for fun).

- singular value decomposition.



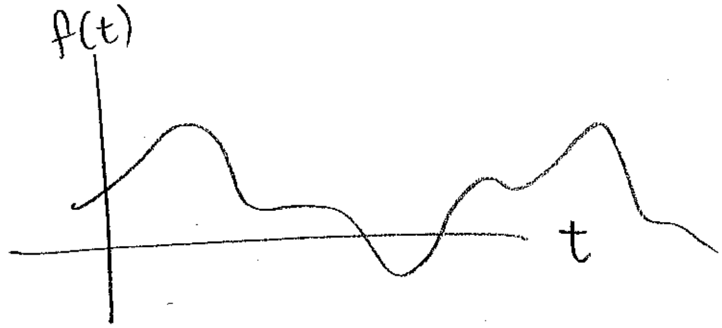
project the data pts onto this line. you wont lose much information.

principal component analysis (PCA)

eg. If you have data pts in $\mathbb{R}^{1000}$ you can project onto a 10-dim subspace to save lots of space

For SVD, you need Gram-Schmidt to find orthonormal eigenbases for "symmetric matrices".

- Fourier transform.



Data is transmitted via waves. How to break up into different freqs?

$$f(t) = a_0 + (b_1 \cos t + b_2 \cos 2t + b_3 \cos 3t + \dots)$$
$$+ (c_1 \sin t + c_2 \sin 2t + \dots).$$

Given $f(t)$, how to find $a_0, b_1, b_2, \dots$
$c_1, c_2, \dots$?

Fact: $\left\{ \begin{matrix} 1, \cos t, & \cos 2t, & \cos 3t, & \dots \\ \sin t, & \sin 2t, & \sin 3t, & \dots \end{matrix} \right\}$.

is an "orthonormal basis" for "the space of functions" $f: [0, 2\pi] \to \mathbb{R}$.

So to find these coefficients, we use "dot products" to calculate "projections"!