

1. Math 221 221.02 Fall 2021

Linear Algebra

Tue / Thu 12:00 - 13:15

Physics 119

Text: Shifrin & Adams ← all exercises taken from here

Office Hours: Mon 16:15 - 17:15 } outside if possible
Thu 13:15 - 14:30 right after class } Physics 209 if not

Safety: • masks, hand wash/sanitize politely point out noncompliance — even me!
• distance if possible
• know where the exits are from the room and the building

Policies • covered on Thu ⇒ fair game for HW or exam Tue
• collaboration/academic honesty
- Yes on HW
- No on exams I have • brought numerous cases to the Office of Student Conduct
• never lost

Index cards

1. Ezra Miller

2. he/him

3. 42nd grade

4. Major or potential major: Math, Music

5. What you hope to get out of this course

students who know linear algebra, especially the right way to think about rank

6. The most important thing you've learned about how you learn

not to take notes!

7. Hobbies: frisbee, gardening, photography, beer

8. Something unique about yourself

hold breath for 4 minutes
screws in right hand
told by doctor in hospital I was going to die of rabies
so radioactive I set off a Geiger counter from across room
remarkable bike accident without injury

X "I'm from MA"

✓ "I'm from HI —

but I'm allergic to pineapple!"

What this course covers

1. things that are straight or flat
2. systems of linear equations and their solutions
 $y=x+2$ not $y=x^2$
3. ways to manipulate these while preserving straightness, flatness, or linearity
4. how to write down
 • understand (e.g. decompose into simpler pieces)
 manipulations ("operations") of this sort
5. solving linear systems
6. applications: so many phenomena behave linearly or — as you've seen in Calculus — approximately so \Rightarrow linear algebra is the foundation of modern math & stat

"linear"

"algebra"

includes

This is a serious mathematical, abstract introduction.

Proofs are essential, both in class and in written assignments, including exams.

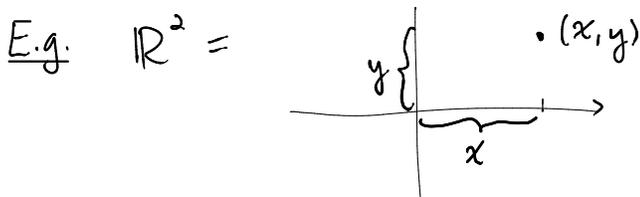
If you are taking this class purely to use linear algebra, then consider Math 212; we'll help you or 218 find a section. Course is geared toward (potential) math majors and minors.

Vectors § 1.1 - 1.2

Ask about any unfamiliar symbols!

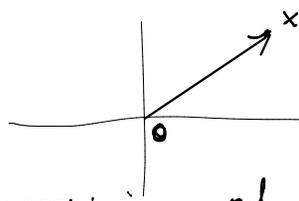
\mathbb{R} = real numbers $x \in \mathbb{R}$ is a partition of the rational #s into a left half and a right half

$\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}, \dots, x_n \in \mathbb{R}\}$ ← This is a sentence!



Def: A vector in \mathbb{R}^n is a point $X = (x_1, \dots, x_n) \in \mathbb{R}^n$.

to visualize a vector:
represent
draw

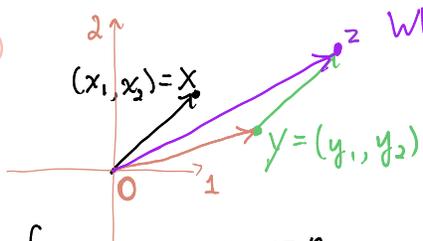


an arrow with tail at O
head at X

Convention: same vector if arrow is moved around.

Q. If you see two arrows, how do you tell if they represent the same vector?

A. (poll class)



What must the coordinates of this point be?

0 to X = y to z (do "rise and run")

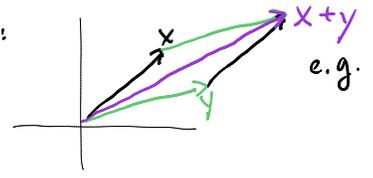
$$\Leftrightarrow z_1 - y_1 = x_1 - 0 \quad z_2 - y_2 = x_2 - 0$$

$$z_1 = x_1 + y_1 \quad z_2 = x_2 + y_2$$

Def: The sum of x and y in \mathbb{R}^n is

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \Rightarrow z = x + y.$$

Better picture:

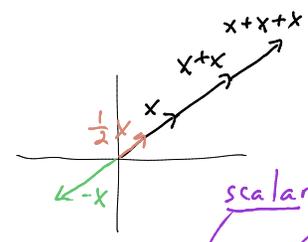


e.g. $x = (5, 4) \Rightarrow x + y = (5 + 7, 4 + 2) = (12, 6)$
 $y = (7, 2)$

Q. What does $x + x$ look like?
 $x + x + x$?

What would $\frac{1}{2}x$ be?

A.



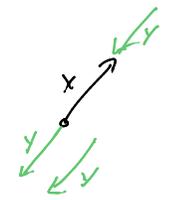
Def: For $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$, the scalar multiple of x by c is $cx = (cx_1, \dots, cx_n) \in \mathbb{R}^n$.

scalar set of all scalars "change of scale"

really "set" space of vectors ("vector space")

Def: Nonzero x and y are parallel if $y = cx$ for some $c \in \mathbb{R}$.

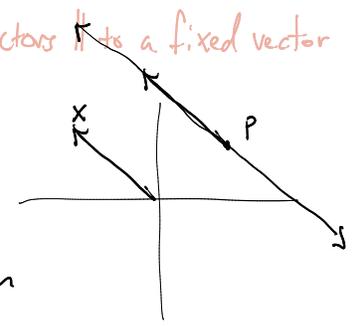
looks like x, y collinear, but now draw various reps: $x \nearrow y$



Q. What is a line in \mathbb{R}^n ?

A. through 0: $\{cx \mid c \in \mathbb{R}\}$ for some $x \in \mathbb{R}^n$ the set of all vectors \uparrow is a fixed vector
 $\stackrel{\text{def}}{=} \text{span}(x)$ parameter

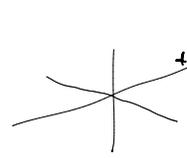
through $p = (p_1, \dots, p_n)$: $\{p + cx \mid c \in \mathbb{R}\}$ for some $x \in \mathbb{R}^n$
 point vector



in \mathbb{R}^2 , implicit description: $z_2 = mz_1 + b$ parametric description of a line; $c = \text{parameter}$

Q. What is a plane in \mathbb{R}^n ?

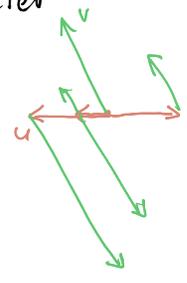
through 0:



3 chop sticks + 2 pencils forget the axes \rightsquigarrow

$$\{su + tv \mid s \in \mathbb{R}, t \in \mathbb{R}\} = \text{span}(u, v)$$

through p : $p + \text{span}(u, v)$



2.

Def: The span of $v_1, \dots, v_k \in \mathbb{R}^n$ is

$$\text{span}(v_1, \dots, v_k) = \{c_1 v_1 + \dots + c_k v_k \mid c_1, \dots, c_k \in \mathbb{R}\}$$

"the set of all linear combinations of v_1, \dots, v_k "

a linear combination of v_1, \dots, v_k

adding k vectors is like adding two because $+$ is associative!
 $(\dots((v_1+v_2)+v_3)+\dots)+v_k$
↑ why?

The dimension of this span is the minimum number of these vectors needed to span. e.g. $\dim(\text{line}) = ?$ $\dim(\text{plane}) = ?$ $\dim(\text{point}) = ?$

E.g. Is $x = (1, 3, -1, -2)$ a linear combination of $u = (1, 1, 0, -1)$ and $v = (2, 0, 1, 1)$? Does x lie in the plane spanned by u and v ? \Leftrightarrow

Can we find $\frac{s}{c_1}, \frac{t}{c_2}$ so that

$$\frac{s}{c_1} u + \frac{t}{c_2} v = x?$$

$$\parallel (s, s, 0, -s) + (2t, 0, t, t) = (s+2t, s, t, -s+t) \stackrel{?}{=} (1, 3, -1, -2)$$

$$\Rightarrow \begin{cases} s+2t = 1 \\ s = 3 \\ t = -1 \\ -s+t = -2 \end{cases}$$

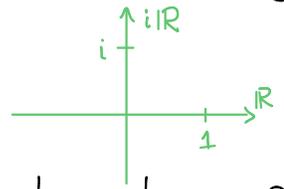
but $3u - v = (1, 3, -1, -4) \neq x$, so: No. This linear system is inconsistent.

systematic (algorithmic) way to solve: next week

Note about \mathbb{R} :

Nothing so far used \mathbb{R} ! Could have used

- complex numbers $\mathbb{C} = \{a+bi \mid a, b \in \mathbb{R}\}$, where $i^2 = -1$, so $i = \sqrt{-1}$



= the set of linear combinations of 1 and i
= the plane spanned by 1 and i

- rational numbers $\mathbb{Q} = \{\frac{a}{b} \mid a, b \text{ integers and } b \neq 0\}$ where $\frac{a}{b} = \frac{a'}{b'}$ if $ab' = a'b$

- binary field $\mathbb{F}_2 = \{0, 1\}$ where $0+0=0$ $0 \cdot 0 = 0$
 $0+1=1$ $0 \cdot 1 = 0$
 $1+1=0$ $1 \cdot 1 = 1$

but not integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.

↑ why? Because you might have to divide by 3 to solve a linear system

Q. What is special about \mathbb{R} ?

A. For $x \in \mathbb{R}$, $x^2 \geq 0$. \leftarrow that's the special thing: \mathbb{R} is ordered.

\mathbb{C} ?
 \mathbb{F}_2 ?
 \mathbb{Q} ?

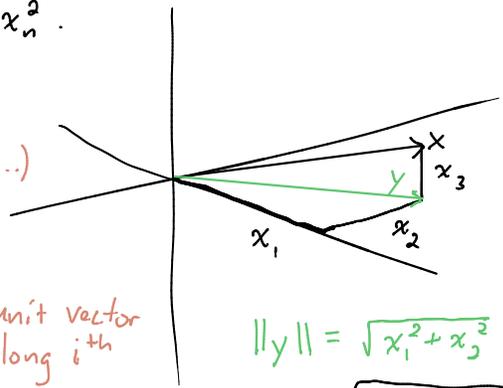
Def: length of $x \in \mathbb{R}^n$ is $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$.

(or magnitude)

Why? $x = x_1 e_1 + \dots + x_n e_n$ $e_i = (\dots, 0, 1, 0, \dots)$

$$= \underbrace{x_1 e_1 + \dots + x_{n-1} e_{n-1}}_{x_{n-1}} + x_n e_n$$

i^{th} slot
 $\Rightarrow e_i = \text{unit vector along } i^{\text{th}} \text{ axis}$



$$\|y\| = \sqrt{x_1^2 + x_2^2}$$

$$\|x\| = \sqrt{\underbrace{\|y\|^2}_{x_1^2 + x_2^2} + x_3^2}$$

$$\Rightarrow \|x\|^2 = \|x_{n-1}\|^2 + x_n^2 \text{ by Pythagoras}$$

$x_1^2 + \dots + x_{n-1}^2$ by induction on n

$$x = 0 \Leftrightarrow \|x\| = 0$$

$x \neq 0 \Rightarrow \frac{x}{\|x\|}$ is a unit vector (length = 1) in the "same direction" as x .

Another way to express it:

Def: For $x, y \in \mathbb{R}^n$ (or $\mathbb{C}^n, \mathbb{Q}^n, \mathbb{F}_2^n, \dots$)

their dot product is $x \cdot y = x_1 y_1 + \dots + x_n y_n$.

Thus $\|x\|^2 = x \cdot x$ in \mathbb{R}^n .

Note: $x \cdot x < 0$ is possible in \mathbb{C}^n : $(1, 2i) \cdot (1, 2i) = 1 - 4 = -3$.

Proposition:

- commutative 1. $x \cdot y = y \cdot x$ for all $x, y \in \mathbb{R}^n$
 - associative 2. $(cx) \cdot y = c(x \cdot y)$ " and $c \in \mathbb{R}$
 - distributive 3. $x \cdot (y+z) = x \cdot y + x \cdot z$ " and $z \in \mathbb{R}^n$
 - positive... 4. $x \cdot x = \|x\|^2 \geq 0$ for all $x \in \mathbb{R}^n$
 - ...definite 5. $x \cdot x = 0 \Leftrightarrow x = 0$.
- } works over any field
} needs \mathbb{R}

Pf: 1. $x_i y_i = y_i x_i$. \leftarrow still a sentence!

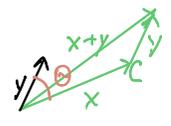
2. $(c x_i) y_i = c(x_i y_i)$.

3. $x_i (y_i + z_i) = x_i y_i + x_i z_i$.

4. $\|x\|^2$ is a sum of squares...

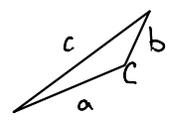
5. ... that is nonzero if $x \neq 0$. \square

E.g. $\|x+y\|^2 = (x+y) \cdot (x+y)$
 $= (x+y) \cdot x + (x+y) \cdot y$
 $= \|x\|^2 + y \cdot x + x \cdot y + \|y\|^2$
 $= \|x\|^2 + 2x \cdot y + \|y\|^2 \Rightarrow \|x\|^2 + \|y\|^2 = \|x+y\|^2 - 2x \cdot y$



Over \mathbb{R} : "recall" Law of cosines: $a^2 + b^2 = c^2 + 2ab \cos C$.

$2ab \cos C = -2\|x\|\|y\|\cos \theta$
 since $C = \pi - \theta$



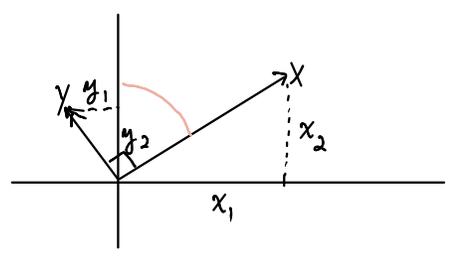
Thus $-2x \cdot y$ "should be" $-2\|x\|\|y\|\cos \theta$.

Def: The angle θ between x and y is defined by

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

x and y are orthogonal ($x \perp y$) if $x \cdot y = 0$.

E.g. in \mathbb{R}^2 :



similar Δs why? Both have an angle $\pi/2 - \theta$
 $\Rightarrow \frac{x_1}{y_2} = \frac{x_2}{-y_1}$
 since $y_1 < 0$
 $\Rightarrow x_1 y_1 + x_2 y_2 = 0$

For higher dim check by induction, or use invariance of $x \cdot y$ under rotation later in the course: put x, y in the plane. Actually, put x, y on axes!

Def of θ needs

Prop (Cauchy - Schwarz inequality):

$|x \cdot y| \leq \|x\| \|y\|$. "=" holds \Leftrightarrow one is a scalar times the other.

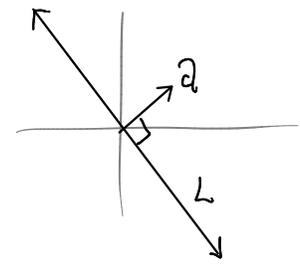
Pf: Easy if $x=0$ or $y=0$ so assume not.

First do case of unit vectors. Need $-1 \leq x \cdot y \leq 1$.

$\|x+y\|^2 = \|x\|^2 + 2x \cdot y + \|y\|^2 = 2(x \cdot y + 1) \geq 0 \Rightarrow x \cdot y \geq -1$
 $\|x-y\|^2 = \dots \Rightarrow x \cdot y \leq 1$
 } "=" $\Leftrightarrow \|x \pm y\| = 0$
 $\Leftrightarrow x = \mp y$

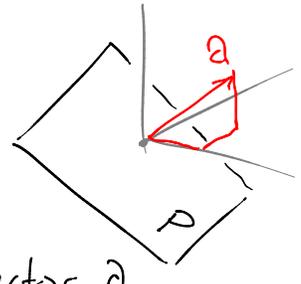
General: $\left| \frac{x}{\|x\|} \cdot \frac{y}{\|y\|} \right| \leq 1 \Rightarrow |x \cdot y| \leq \|x\| \|y\|$. \square

3. E.g. $S = \{\text{vectors orthogonal to } a\} \subseteq \mathbb{R}^n$
 $= \{x \in \mathbb{R}^n \mid a \cdot x = 0\}$ with $a \neq 0$



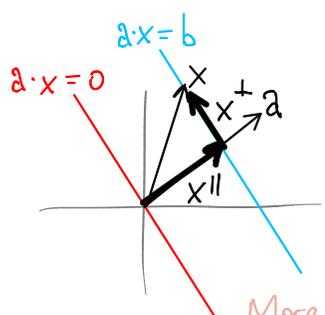
$n=2$: $S = \text{line } L \text{ through } 0 \text{ with normal vector } a$
 $n=3$: $S = \text{plane } P$ " " " "

$$3x_1 + 2x_2 + x_3 = 0 \Leftrightarrow (3, 2, 1) \cdot x = 0.$$



n arbitrary: $S = \text{hyperplane } H \text{ through } 0 \text{ with normal vector } a.$
omit?

Def: A hyperplane in \mathbb{R}^n is $H = \{x \in \mathbb{R}^n \mid a \cdot x = b\}$ for some $b \in \mathbb{R}$
 and nonzero $a \in \mathbb{R}^n$. dim = n-1



" $a \cdot x$ measures projection of x onto a "

$a \cdot x$ constant \Rightarrow all x 's on the line have same projection to a
 (= b)

More precisely: $x \in \mathbb{R}^n \Rightarrow x = x^|| + x^\perp$

$$\|a\| = 1 \Rightarrow x^|| = (x \cdot a) a$$

scalar \uparrow *vector* \nwarrow projection of x onto a .

general a : $(x \cdot \frac{a}{\|a\|}) \frac{a}{\|a\|} = \boxed{\frac{x \cdot a}{\|a\|^2} a = x^||}$ and $x^\perp = x - x^||$

Q: $x^|| \cdot x^\perp = ?$

A: $x^\perp \cdot a = (x - x^||) \cdot a = x \cdot a - (\frac{x \cdot a}{\|a\|^2}) a \cdot a$
 $= 0 \Rightarrow x^\perp \cdot x^|| = 0.$

Important notation

$$x = (x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ column vector}$$

not x^\perp

$$x^T = [x_1 \dots x_n] \text{ row vector}$$

$x^{(T)}$ = "transpose"

book often gets col vs. row wrong

Matrix multiplication

For column vectors a and x in \mathbb{R}^n ,

$$a^T x = a \cdot x$$

$$\left[\text{---} a^T \text{---} \right] \begin{bmatrix} | \\ x \\ | \end{bmatrix} = a_1 x_1 + \dots + a_n x_n.$$

objects of same type
different types

hyperplane $H_1 =$ solution to linear eqn.
 H_2
 \vdots
 H_m

$$\begin{aligned} A_1 x &= b_1 && \text{row } A_1 \in \mathbb{R}^n \\ &&& \text{scalar } b_1 \in \mathbb{R} \\ A_2 x &= b_2 \\ &\vdots \\ A_m x &= b_m && \text{row } A_m \in \mathbb{R}^n \\ &&& \text{scalar } b_m \in \mathbb{R} \end{aligned}$$

System of linear equations

solution to (*) = point in $H_1 \cap \dots \cap H_m$
 x_0

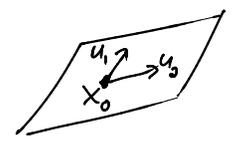
intersect

implicit description of solutions

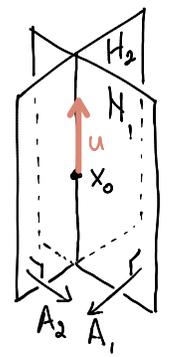
want: parametric description of all solutions

← e.g. $S = \{x_0 + t u \mid t \in \mathbb{R}\}$

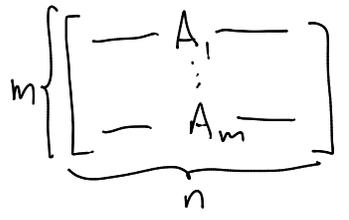
in general: $S = \{x_0 + t_1 u_1 + \dots + t_k u_k \mid t_1, \dots, t_k \in \mathbb{R}\}$
for fixed $u_1, \dots, u_k \in \mathbb{R}^n$
 $= x_0 + \text{span}(u_1, \dots, u_k).$



S could be empty!



Def: 1. An $m \times n$ matrix is an array A with \bullet m rows A_1, \dots, A_m
 \bullet n columns (so A_i has n entries for each i)



2. For $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$, $Ax = \begin{bmatrix} A_1 x \\ \vdots \\ A_m x \end{bmatrix} \in \mathbb{R}^m$

Thus the matrix eqn $Ax = b$ says precisely (*)

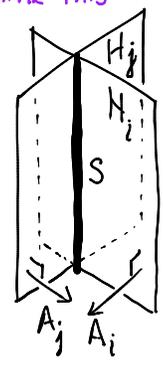
works over $\mathbb{R}, \mathbb{Q}, \mathbb{C}, \mathbb{F}_2, \dots$

3. augmented matrix $[A|b] = \left[\begin{array}{c|c} \text{---} A_1 \text{---} & b_1 \\ \vdots & \vdots \\ \text{---} A_m \text{---} & b_m \end{array} \right]$ shorthand for $Ax = b.$

equations: (*) or $Ax = b$ or $[A|b]$

solutions: a set of column vectors in \mathbb{R}^n : $S = \{x \in \mathbb{R}^n \mid Ax = b\}.$

4. $A \in \mathbb{R}^{m \times n}$ $b \in \mathbb{R}^m$ write this... \mathbb{R} could be any field ... think this



Q. What operations on $[A|b]$ preserve sol set S ?

A. (i) Swap any pair of rows.

geometric meanings

(ii) Multiply a row by a scalar $\neq 0$

(iii) Replace any row by its sum with a multiple of another row.

Def (i), (ii), and (iii) are elementary row operations.

Theorem 4.1 Applying any sequence of elementary row ops to $[A|b]$ results in a system with the same solution set.

Pf: (i) merely lists the equations $A_i x = b_i$ in a different order.

set complement

For (ii), note that $A_i x = b_i \Leftrightarrow c(A_i x) = c b_i$ if $c \in \mathbb{R} \setminus \{0\}$

$$\Leftrightarrow (c A_i) x = c b_i.$$

For (iii), let $[A'|b']$ be the system obtained from $[A|b]$ by replacing the row $[A_i|b_i]$ with $[A_i + c A_j | b_i + c b_j]$.

Let $S' = \text{sols of } A'x = b'$. Then

"is contained in" $S \subseteq S'$

$$\begin{aligned} \text{because } (A_i + c A_j)x &= A_i x + c A_j x \\ &= b_i + c b_j \text{ when } x \in S. \end{aligned}$$

$[A b]$	$[A' b']$
$A_1 x = b_1$	$A_1 x = b_1$
\vdots	\vdots
$A_{i-1} x = b_{i-1}$	$A_{i-1} x = b_{i-1}$
$A_i x = b_i$	$(A_i + c A_j)x = b_i + c b_j$
$A_{i+1} x = b_{i+1}$	$A_{i+1} x = b_{i+1}$
\vdots	\vdots
$A_m x = b_m$	$A_m x = b_m$

Aside: need $S = S'$, not just $S \subseteq S'$.

$$\begin{aligned} &\Updownarrow \\ &S \subseteq S' \text{ and } S' \subseteq S. \end{aligned}$$

But $A_i = A'_i - c A'_j$ (!) so also $S' \subseteq S$.

Finally, since each of (i), (ii), (iii) preserves S , any sequence of them does, as well. \square

E.g. $Ax = b$ for $A = \begin{bmatrix} 3 & -2 & 2 & 9 \\ 2 & 2 & -2 & -4 \end{bmatrix}$ and $b = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$:

$$[A|b] = \begin{bmatrix} 3 & -2 & 2 & 9 & | & 4 \\ 2 & 2 & -2 & -4 & | & 6 \end{bmatrix} \xrightarrow{(iii) A_1 \leftrightarrow A_1 - A_2} \begin{bmatrix} 1 & -4 & 4 & 13 & | & -2 \\ 2 & 2 & -2 & -4 & | & 6 \end{bmatrix}$$

$$\xrightarrow{(iii) A_2 \leftrightarrow A_2 - 2A_1} \begin{bmatrix} 1 & -4 & 4 & 13 & | & -2 \\ 0 & 10 & -10 & -30 & | & 10 \end{bmatrix}$$

$$\xrightarrow{(ii)} \begin{bmatrix} 1 & -4 & 4 & 13 & | & -2 \\ 0 & 1 & -1 & -3 & | & 1 \end{bmatrix}$$

$$\xrightarrow{(iii) A_1 \leftrightarrow A_1 + 4A_2} \begin{bmatrix} 1 & 0 & 0 & 1 & | & 2 \\ 0 & 1 & -1 & -3 & | & 1 \end{bmatrix} = [A'|b']$$

⇒ Ax = b has same sols S as A'x = b'.

$$\Rightarrow S = \left\{ x \in \mathbb{R}^4 \mid \begin{matrix} x_1 + x_4 = 2 \\ x_2 - x_3 - 3x_4 = 1 \end{matrix} \right\}$$

Whatever values we assign to x_3, x_4 , we can solve for x_1, x_2 successfully

In particular,

$$x_3 = x_4 = 0 \Rightarrow \begin{matrix} x_1 = 2 \\ x_2 = 1 \end{matrix} \Rightarrow \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \in S.$$

a particular solution x_0

think of as parameters

general solution: x_3, x_4 free variables — can take on any values

$$\begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_3 \\ x_3 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} -x_4 \\ 3x_4 \\ 0 \\ x_4 \end{bmatrix}$$

so $x \in S \Leftrightarrow x = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_3 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} -x_4 \\ 3x_4 \\ 0 \\ x_4 \end{bmatrix}$

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

$\uparrow \quad \uparrow \quad \uparrow$
 $-u_1 \quad -u_2 \quad x_0$

$$= \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

$$= x_0 + t_1 u_1 + t_2 u_2$$

Def: A matrix is in echelon form if

1. the leading (leftmost nonzero) entries progress to the right from each row to the next;

and 2. all 0 rows are at the bottom.

$$\begin{bmatrix} 0 & 0 & \dots & 0 & * & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & * & \dots \end{bmatrix}$$

The echelon form is reduced if, in addition,

3. every leading entry is 1;

pivot

pivot column

and 4. in each column containing a leading entry, all other entries are 0.

E.g. just did one!

$$\begin{bmatrix} 1 & -4 & 4 & 13 & | & -2 \\ 0 & 10 & -10 & -30 & | & 10 \end{bmatrix} \text{ echelon form}$$

$$\begin{bmatrix} 1 & 0 & 0 & 1 & | & 2 \\ 0 & 1 & -1 & -3 & | & 1 \end{bmatrix} \text{ reduced echelon form}$$

↑ ↑
pivot columns

Thm 4.3: Each matrix has unique reduced echelon form.

Pf: Exercise 16 which you aren't asked to do, but you could. \square

How to find it?

1. Algorithm produces one.
2. Different reduced echelon forms have different sols.

Algorithm (Gaussian elimination)

Echelon form

Init: $i=1$

While: there is a nonzero entry in some row $\geq i$

Do: 1. pick row $\geq i$ with a leftmost such entry

2. swap that row with row i

3. add multiples of row i to rows $> i$ to cancel entry in pivot column

4. $i \rightarrow i+1$

Output: the resulting matrix

Reduced echelon form given any echelon form

Init: $i=1$

While: there is a nonzero entry in some row $\geq i$

Do: 1. rescale row i so pivot is 1

2. add multiples of row i to rows $< i$ to cancel entries in pivot column

3. $i \rightarrow i+1$

Output: the resulting matrix

5. Rows vs. columns

Note: everything today works over any field.

$$A \in \mathbb{R}^{m \times n} \Rightarrow A = \begin{bmatrix} - A_1 - \\ \vdots \\ - A_i - \\ \vdots \\ - A_m - \end{bmatrix}$$

$$A_i = [a_{i1} \dots a_{in}] \in \mathbb{R}_{row}^n \quad i = \text{generic row index}$$

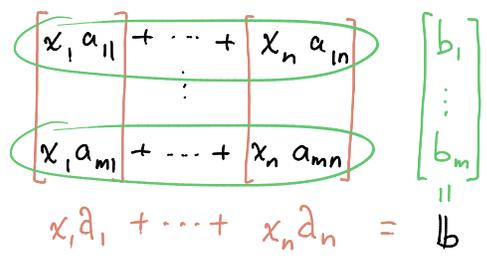
$$= \begin{bmatrix} | & & | \\ a_1 & \dots & a_j & \dots & a_n \\ | & & | \end{bmatrix}$$

$$a_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{R}_{col}^m \quad j = \text{generic column index}$$

$$Ax = b \Leftrightarrow A_i x = b_i \quad \text{for } i = 1, \dots, m$$

but also

$$Ax = b \Leftrightarrow x_1 a_1 + \dots + x_n a_n = b$$

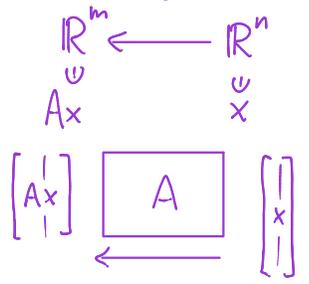


- $\Leftrightarrow Ax$ is a linear combin of the a_j
- $\Leftrightarrow b$ is a linear combin of the columns of A
- $\Leftrightarrow b$ lies in span of cols of A
- $\Leftrightarrow b \in \text{span}(a_1, \dots, a_n)$

Def: $C(A) = \text{column space}$ of A
or image of M_A

why? $A \mapsto$ function taking x 's to b 's:

$\Leftrightarrow x$ lists the coefficients in a linear combination of the columns of A that equals b



E.g. $\begin{bmatrix} -3 & 6 & 4 & e \\ 1 & 2/3 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \\ \pi \end{bmatrix} \begin{matrix} \leftarrow \\ \leftarrow \\ \leftarrow \\ \leftarrow \end{matrix}$

$$= 3 \begin{bmatrix} -3 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 6 \\ 2/3 \end{bmatrix} + 4 \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \pi \begin{bmatrix} e \\ 0 \end{bmatrix}$$

↑ ↑ ↑ ↑
have class fill these in

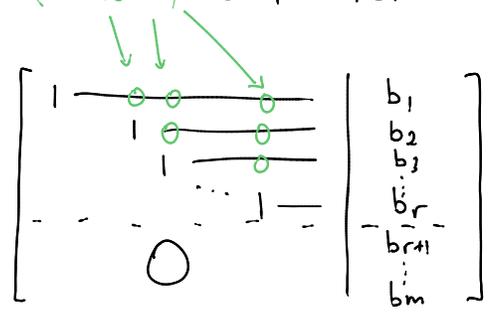
$$\begin{bmatrix} -3 & 6 & 4 & e \\ 1 & 2/3 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \\ \pi \end{bmatrix} = \begin{bmatrix} 19 + \pi e \\ 2/3 \end{bmatrix}$$

Q. When does $Ax = b$ have a solution?
i.e., When is $Ax = b$ consistent?

- ✓ When there exists an x listing the coefficients in a linear combination of the columns of A that equals b .
- ✓ When b lies in span of columns of A
- ✓ When $b \in C(A)$
- ✓ When b lies in the image of the function "multiplication by A " M_A

"maps to"

In (reduced) echelon form:



Lemma 1: consistent $\Leftrightarrow \begin{matrix} b_{r+1} = 0 \\ \vdots \\ b_m = 0 \end{matrix}$

Pf: \Rightarrow : $A_i = 0$ for $i > r$ so $Ax = b \Rightarrow A_i x = b_i$ for $i > r$.
 \Leftarrow : For any values of the free variables, the values of the pivots are determined from bottom to top.

E.g.
$$\begin{bmatrix} 1 & \pi & 4 & 1 & 7 \\ 0 & 0 & 1 & 3 & -11/5 \\ 0 & 0 & 0 & 0 & ? \end{bmatrix}$$

pivot x_1 *free* x_2 *free* x_3 x_4

$x_1 + \pi x_2 + 4x_3 + x_4 = 7$
 $x_3 + 3x_4 = -11/5$
 $0 + 0 + 0 + 0 = ?$ consistent if $? = 0$, inconsistent if $? \neq 0$

make whatever you want!

Def: The rank of $A \in \mathbb{R}^{m \times n}$ is $\dim C(A) = \min \# \text{ columns required to span } C(A)$ (recall)

E.g. $\text{rank} \begin{bmatrix} 1 & \pi & 4 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} = ? 2$ $\text{rank} \begin{bmatrix} 1 & \pi & 4 & 1 & 7 \\ 0 & 0 & 1 & 3 & -11/5 \\ 0 & 0 & 0 & 0 & ? \end{bmatrix} = ? \begin{matrix} 2 \text{ if } ? = 0 \\ 3 \text{ if } ? \neq 0 \end{matrix}$

This is the "correct" definition that I mentioned on day 1. Learn it. Use it. Reduce all questions about rank to it. You will be tested on it. It should be your go-to def.

Lemma 2: Fix v_1, \dots, v_r cols of A with corresponding cols v'_1, \dots, v'_r A' , where $[A|b] \xrightarrow{\text{row ops}} [A'|b']$.

Cor: $C(A) = \mathbb{R}^m \Rightarrow C(A') = \mathbb{R}^m$.
 Pf: $b' \in \mathbb{R}^m \Rightarrow [A'|b'] \rightsquigarrow [A|b]$ for some $b \in C(A)$. \mathbb{R}^m .

Then $b = c_1 v_1 + \dots + c_r v_r \Leftrightarrow b' = c_1 v'_1 + \dots + c_r v'_r$. Note: same coeffs c_i .

Pf: Check type (i), (ii), (iii) row ops directly. And observe that $[A|b] \rightsquigarrow [A'|b'] \Leftrightarrow [A'|b'] \rightsquigarrow [A|b]$. \square

Prop 1: v_1, \dots, v_r span $C(A) \Leftrightarrow v'_1, \dots, v'_r$ span $C(A')$ if $A \rightsquigarrow A'$.

Pf: \Leftarrow : $b \in C(A) \Rightarrow b' \in C(A')$ by Lemma 2
 $\Rightarrow b' = c_1 v'_1 + \dots + c_r v'_r$ by hypothesis
 $\Rightarrow b = c_1 v_1 + \dots + c_r v_r$ by Lemma.
 \Rightarrow : same (i.e., swap the primes and un-primes) \square

This is why pivot columns are important.

Corollary: $\text{rank } A = \text{rank } A'$ if $A \rightsquigarrow A'$. Why is this useful?

Prop 2: $\text{rank } A = \# \text{ pivots in any echelon form, and pivot cols of } A \text{ minimally span } C(A)$.

Pf: By Cor and Prop 1, need only prove for $A = U$, reduced echelon form.

Lemma 1 \Rightarrow pivot cols $\begin{pmatrix} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & 1 \\ e_1, e_2, \dots, e_r \end{pmatrix}$ span $C(U)$, and \mathbb{R}^r can't be spanned by less than r vectors: $C\left(\begin{matrix} < r \\ r \\ \square \end{matrix}\right) = \mathbb{R}^r \Rightarrow$ reduced echelon form has column space \mathbb{R}^r by (Cor of) Lemma 2, but last row is 0, so that is impossible. \square

Prop 3: $Ax=b$ consistent \Leftrightarrow rank $A =$ rank $[A|b]$. *One direction is easy; which?*

pf: \Rightarrow : consistent $\Rightarrow b \in C(A) \Rightarrow C([A|b]) = C(A)$ and same pivot cols span minimally.
 \Leftarrow : pivot cols of A span $C([A|b]) \Rightarrow b$ lies in their span. \square

Parametric to implicit *implicit to parametric means: solve $Ax=b$.*

Given x_0 and $v_1, \dots, v_k \in \mathbb{R}^n$, find linear equations $Ax=b$ so that $\text{sols}(Ax=b) = x_0 + \text{span}(v_1, \dots, v_k)$.

E.g. Find implicit equation(s) for the plane

$$x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \text{ in } \mathbb{R}^3.$$

Equivalently, which vectors x are expressible as $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$?

Solve $x = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, i.e.

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}}_A \underbrace{\begin{bmatrix} s \\ t \end{bmatrix}}_x = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}}_b = \begin{bmatrix} x_1 - 1 \\ x_2 - 2 \\ x_3 - 1 \end{bmatrix}$$

This is a linear system!

cancel $\left(\begin{array}{cc|c} 1 & 2 & x_1 - 1 \\ 0 & 1 & x_2 - 2 \\ 1 & 1 & x_3 - 1 \end{array} \right) \rightsquigarrow \begin{array}{cc|c} 1 & 2 & x_1 - 1 \\ 0 & 1 & x_2 - 2 \\ 0 & -1 & x_3 - x_1 \end{array}$ cancel

$$\begin{bmatrix} 1 & 2 & | & x_1 - 1 \\ 0 & 1 & | & x_2 - 2 \\ 0 & 0 & | & x_3 - x_1 + x_2 - 2 \end{bmatrix}$$

has solution \Leftrightarrow every 0 row in echelon form has corresponding 0 on RHS, so x is expressible $\Leftrightarrow x_3 - x_1 + x_2 - 2 = 0$.

6.

(rest of) today: $A \in \mathbb{R}^{m \times n}$

Def: The (system of) equation(s) $Ax=b$ is inhomogeneous if $b \neq 0$; the corresponding equation(s) $Ax=0$ is the associated homogeneous (system of) equation(s).

Lemma: For vectors $x, y \in \mathbb{R}^n$ and scalar $c \in \mathbb{R}$,

$$A(x+cy) = Ax + cAy.$$

Pf: $A(x+cy) = (x_1+cy_1)a_1 + \dots + (x_n+cy_n)a_n$
 $= \underbrace{x_1a_1 + \dots + x_na_n}_{Ax} + \underbrace{cy_1a_1 + \dots + cy_na_n}_{cAy} \quad \square$

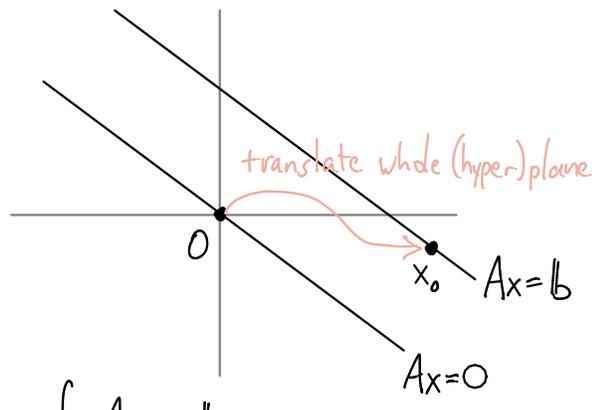
could omit this word without changing anything

Thm 5.3: Assume $Ax=b$ has a "particular" solution x_0 .

v is a solution of $Ax=b \iff v$ has the form

$$v = x_0 + u$$

for some solution u of $Ax=0$.



Pf: $\Leftarrow: v = x_0 + u \implies Av = A(x_0 + u)$
 $= Ax_0 + Au$ by Lemma
 $= b + 0$
 $= b \implies v$ is a solution of $Ax = b$.

\implies : Assume v solves $Ax=b$; i.e. assume $Av=b$. Then $v = x_0 + u$ for some u , namely $u = v - x_0$, and

$$A(v-x_0) = Av - Ax_0 \text{ by Lemma}$$

$$= \underset{b}{b} - \underset{b}{b}$$

$$= 0. \quad \square$$

What happens without this hypothesis?

Corollary: A consistent system $Ax=b$ has a unique solution

$\iff Ax=0$ has only the trivial solution $x=0$.

Prop. 5.4: $Ax=0$ has unique solution $\iff \text{rank } A = n$.

Pf: Equivalent: $Ux=0$ " " " " $\text{rank } U = n$ for all U in reduced echelon form.

Why? If $A \rightsquigarrow U$ then $\text{sols } A = \text{sols } U$ and $\text{rank } A = \text{rank } U$.

So let U be in r.e.f. Then

$\text{rank } U < n \Rightarrow U$ has $< n$ pivots \Rightarrow some column of U has no pivot

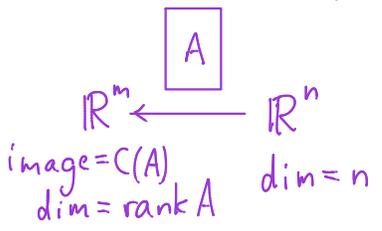
\Rightarrow some variable is free

$\Rightarrow Ux = 0$ has (at least) \mathbb{R} -many solutions.

On the other hand, $\text{rank } U = n \Rightarrow U = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \Rightarrow$ the only sols have $x_1 = 0$

$x_2 = 0$
 \vdots
 $x_n = 0. \quad \square$

Geometrically, why should this (Prop 5.4) be?

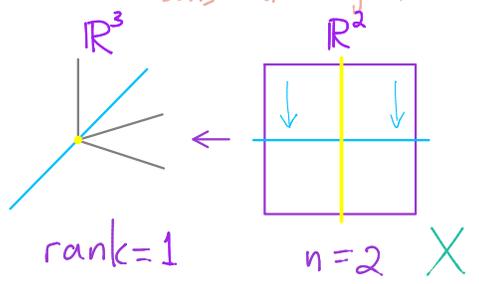
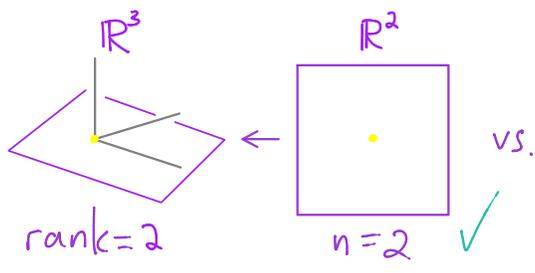
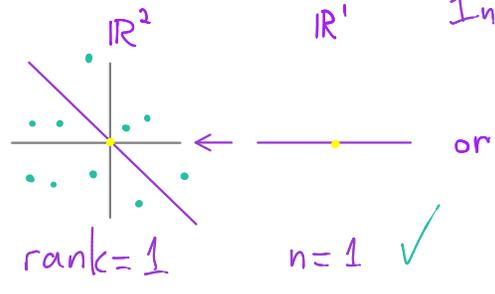


So $\text{rank } A = n$ means A preserves dimension of \mathbb{R}^n !

M_A sticks \mathbb{R}^n into \mathbb{R}^m without compression: $x \neq y \Rightarrow Ax \neq Ay$

In particular: $x \neq 0 \Rightarrow Ax \neq A0 = 0$.

Need more experience under our belts to do this justice.



Q1. For which A does $Ax = b$ have unique solution for all $b \in \mathbb{R}^m$?

Q2. " " " is " consistent " " " ?

A. First look at the pictures: line misses $\therefore \therefore \therefore$. Why? rank $< m$!

general: for all $b \in \mathbb{R}^m$, $\left. \begin{array}{l} \bullet b \text{ has the form } Ax \\ \bullet b \in C(A) \\ \bullet b \in \text{image of } M_A \end{array} \right\} \text{equivalent}$

i.e. A2. Prop: $Ax = b$ consistent for all $b \in \mathbb{R}^m \Leftrightarrow \text{rank } A = m$.

$\mathbb{R}^m = C(A) \quad \square$

A1. $\Leftrightarrow Ax = b$ is consistent for all b and, by Cor, $\Leftrightarrow \text{rank } A = m$

$\bullet Ax = 0$ has only the trivial solution $x = 0$. $\Leftrightarrow \text{rank } A = n$
Prop. 5.4

$\Leftrightarrow \text{rank } A = m = n$.

Def: A is nonsingular (or ~~invertible~~) if $m=n=\text{rank } A$. defined later

A is singular if $m=n$ and $\text{rank } A < n$.

E.g. The $n \times n$ identity matrix $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ is nonsingular.

General: A nonsingular \Leftrightarrow A has r.e.f. $\neq I_n$.

(Do in class if there is time:)

Application: curve fitting

Given 3 points $(x_1, y_1), (x_2, y_2), (x_3, y_3) \in \mathbb{R}^2$ with x_1, x_2, x_3 distinct, find a parabola $y = ax^2 + bx + c$ through them.

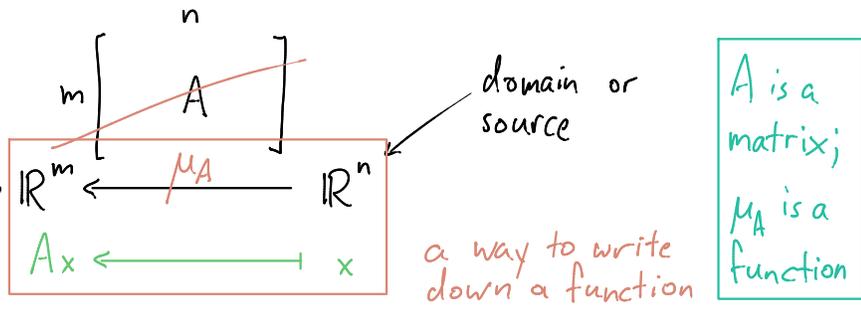
(HW: v_1, v_2, v_3 not collinear \Rightarrow parabola exists and is unique.)

Answer:	$ax_1^2 + bx_1 + c = y_1$	an inhomogeneous linear system!
	$ax_2^2 + bx_2 + c = y_2$	Solution = coeffs a, b, c on parabola
	$ax_3^2 + bx_3 + c = y_3$	through v_1, v_2, v_3

class selects points; we all solve
(Ensure one pt. has $x=0$, for ease of row reduction.)

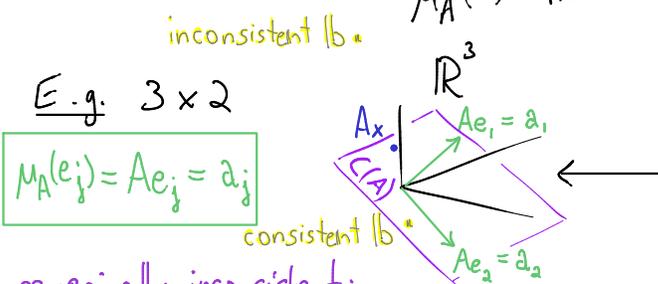
7. Chapter 2 Matrix Algebra \mathbb{R} could be any field

matrix $A \in \mathbb{R}^{m \times n}$
vector $x \in \mathbb{R}^n_{col}$ } $\Rightarrow Ax \in \mathbb{R}^m_{col}$



$A \rightsquigarrow$ function $\mu_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$
 $\mu_A(x) = Ax$

image of μ_A is $\{\mu_A(x) \mid x \in \text{domain}\} \subseteq \text{range}$
 $= \mu_A(\text{domain})$
 $= \mu_A(\mathbb{R}^n) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$
 $= C(A)$.



generically inconsistent:
 $b \notin \text{span}(a_1, a_2)$
"Pf:" row reduce A
 \rightsquigarrow row of 0's

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} | \\ | \\ | \end{bmatrix} = \begin{bmatrix} | & | \\ | & | \\ | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

lin. combin. of a_1 and a_2 with coeff. x_1, x_2

image of μ_A could be
rank: 0 1 2
Note: don't need \angle or \square to know rank!

$Ax = b$ consistent when...?
 $\Leftrightarrow b \in \text{image of } \mu_A$.

$Ax = b$ consistent $\forall b \Leftrightarrow C(A) = \mathbb{R}^m$
 $\Leftrightarrow \text{rank } A = m$.

Def: μ_A is surjective (onto)

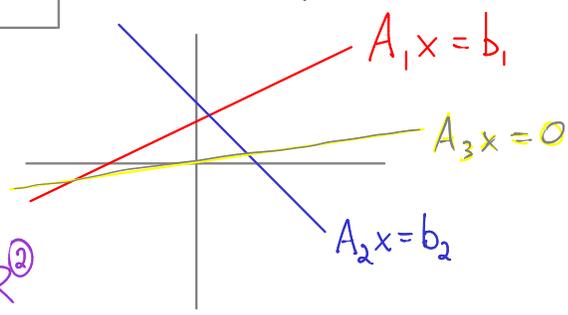
$Ax = b$ has unique sol when...?
• $Ax = b$ is consistent ($b \in C(A)$) and
• no collapsing from \mathbb{R}^n to \mathbb{R}^m : $x \neq y \Rightarrow Ax \neq Ay$
 $\dim(\text{source}) = \dim(\text{image})$
 $n = \text{rank } A$

E.g. 3×2
 $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -A_1^- \\ -A_2^- \\ -A_3^- \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}$

generically inconsistent:
• $x \notin H_1 \cap H_2 \cap H_3$ in \mathbb{R}^2

Def: μ_A is injective (into or one-to-one)

Think in \mathbb{R}^n now:



higher dim:
 m hyperplanes
in \mathbb{R}^n

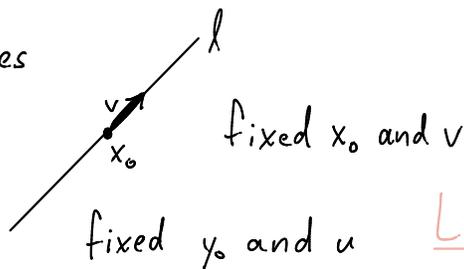
Def: A function $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear (or a linear transformation)

if $T(x+cy) = Tx + cTy$ different "+"!
" operator
" map

E.g. Lemma, last time: μ_A is linear.

Geometrically: linear \Rightarrow takes lines to lines

$$l = \{x_0 + cv \mid c \in \mathbb{R}\}$$



$$T(l) = \left\{ \underbrace{T(x_0)}_{y_0} + c \underbrace{T(v)}_u \mid c \in \mathbb{R} \right\}$$

Linear Algebra!

Prop: Linear maps preserve linear combinations.

Pf: Assume T is linear. Then

$$\begin{aligned} T(c_1x_1 + \dots + c_kx_k) &= T(c_1x_1 + \dots + c_{k-1}x_{k-1}) + c_kTx_k \text{ by linearity} \\ &= c_1Tx_1 + \dots + c_{k-1}Tx_{k-1} + c_kTx_k \text{ by induction on } k (!) \\ &= \sum_{i=1}^k c_iTx_i. \end{aligned}$$

$$k=1: T(c_1v_1) = \underbrace{T(0)} + c_1T(v_1) = 0 + c_1Tv_1 = c_1Tv_1 \checkmark$$

$$T(0) = T(0+0) = T(0) + 1T(0) \Rightarrow 0 = 1T(0) = T(0). \square$$

Q. $\mu_A \left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right) = ?$ $a_1 =$ left column of A

$$\mu_A(e_1) = a_1$$

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{row } j \Rightarrow \boxed{\mu_A(e_j) = a_j}$$

So what? Combine with Prop!

Q. Given linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, what can it look like?

A. $x \in \mathbb{R}^n \Rightarrow x \in \text{span}(e_1, \dots, e_n)$

$\Rightarrow Tx$ determined by what it does to $e_1, \dots, e_n!$

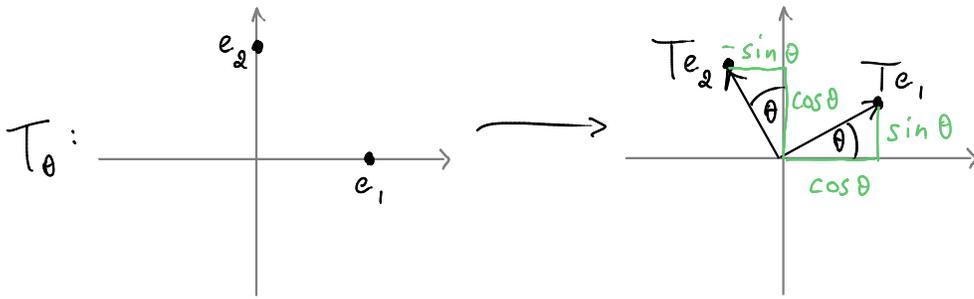
Say $a_1 = Te_1, \dots, a_n = Te_n$. What's Tx ?

$x = x_1e_1 + \dots + x_n e_n$, so Prop \Rightarrow

$$Tx = x_1a_1 + \dots + x_n a_n = \mu_A(x) \text{ for } A = \begin{bmatrix} | & & | \\ Te_1 & \dots & Te_n \\ | & & | \end{bmatrix}$$

Hence every linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is μ_A for some $A \in \mathbb{R}^{m \times n}!$

E.g. rotation of \mathbb{R}^2 by angle θ



Note: $T_\theta(cx) = cT_\theta(x)$

$T_\theta(x+y) = T_\theta(x) + T_\theta(y)$

rotation • preserves circles around 0

• takes parallelograms to congruent parallelograms

$T_\theta\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = ? \quad \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$

$T_\theta\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = ? \quad \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$

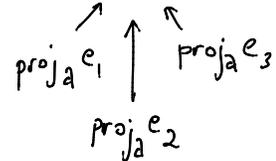
$\Rightarrow T_\theta = M_A$ for $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$



no need to memorize; can reconstruct by picture

E.g. $\text{proj}_{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Set $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. $\text{proj}_a x = \frac{x \cdot a}{\|a\|^2} a = \frac{x_1 + x_2 + x_3}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

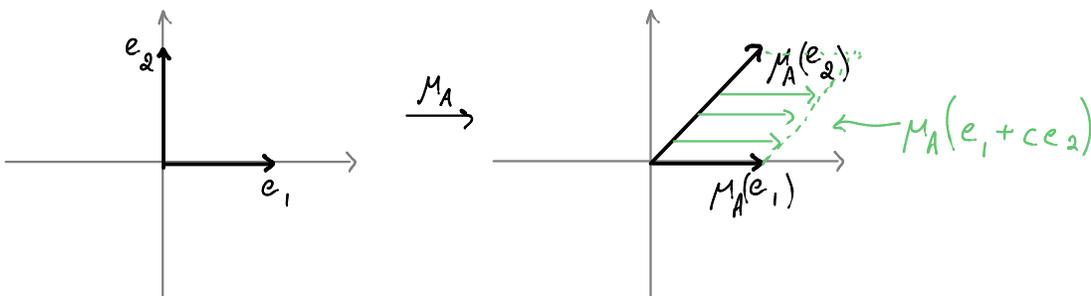
is linear as a function of $x \Rightarrow \text{proj}_a = M_C$ for $C = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

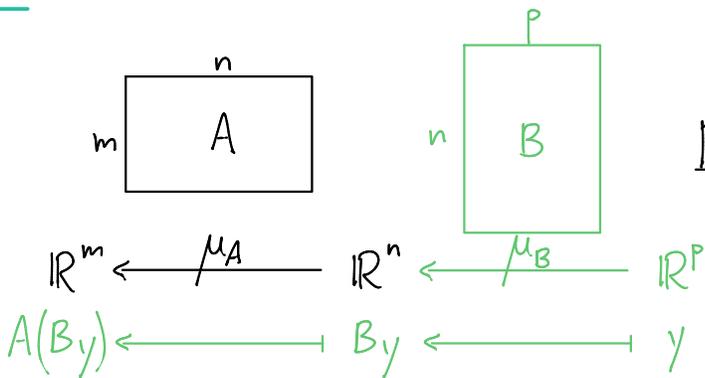


E.g. $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \Rightarrow M_{I_n} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is ...? identity map

E.g. $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow M_A = ?$ reflection across x-axis

E.g. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \Rightarrow M_A = ?$ shear





Lemma: $T \circ S$ is linear if T and S are. (21)

Pf: $T \circ S(y + tz) = T(Sy + tSz) = (T \circ S)y + t(T \circ S)z. \square$

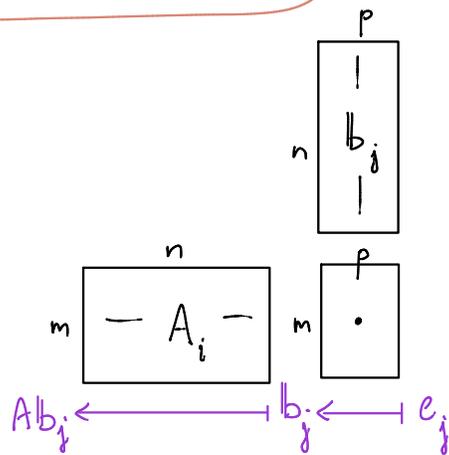
Cor: $M_A \circ M_B$ is linear, so $M_A \circ M_B = M_C$ for some $C!$

Q. What is C ? $C \in \mathbb{R}^{m \times p}$
 A. $M_C(y) =$ linear combination of columns c_1, \dots, c_p with coeffs y_1, \dots, y_p . So
 Q'. What are c_1, \dots, c_p ?

$A(y_1 b_1 + \dots + y_p b_p) \leftarrow y_1 b_1 + \dots + y_p b_p$
 $y_1 A b_1 + \dots + y_p A b_p$

Def: For an $m \times n$ matrix A and $n \times p$ matrix B , their product is the $m \times p$ matrix AB satisfying $M_{AB} = M_A \circ M_B$.

Lemma: AB has columns Ab_1, \dots, Ab_p .
 Equivalently, $(AB)_{ij} = A_i b_j$
 or $(AB)_i = A_i B =$ linear combination of rows of B with coeffs from A_i .



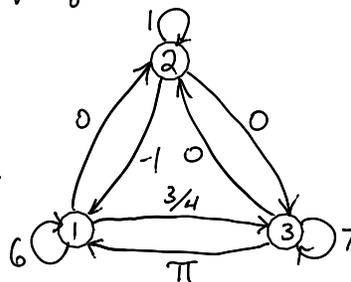
E.g. $A = m [2 \ 3 \ 5]$ $B = n \begin{bmatrix} 7 \\ 11 \\ \pi \end{bmatrix} \Rightarrow AB = ?$ $2 \cdot 7 + 3 \cdot 11 + 5 \cdot \pi = 14 + 33 + 5\pi = 47 + 5\pi$

Q. Is BA defined here? Yes: $BA = ?$ different shape than AB
 in general? No: $C = \begin{bmatrix} 7 & 1 \\ 11 & 1 \\ \pi & 1 \end{bmatrix} \Rightarrow AC$ defined but CA not.

Q. If A and B square (say $n \times n$) is $AB = BA$? Why? $\mathbb{R}^3 \xleftarrow{C} \mathbb{R}^2 \neq \mathbb{R}^1 \xleftarrow{A} \mathbb{R}^3$
 A. No: $A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$ $B = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$ — or just about any square $A, B!$

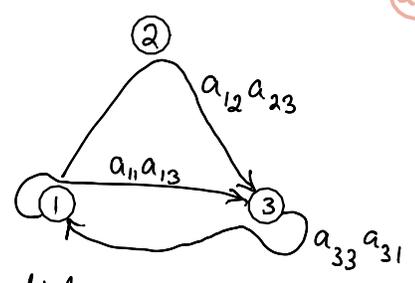
E.g. powers of $n \times n$ A
 directed graph with edge labels

from i to j
 $\begin{bmatrix} 6 & 0 & 3/4 \\ -1 & 1 & 0 \\ \pi & 0 & 7 \end{bmatrix}$



CAN OMIT

$(A)_{ij}$ = "distance" from i to j = "length" of edge from i to j
 $(A^2)_{ij}$ = sum of "lengths" of 2-step paths from i to j
 "length" = $\text{length}_1 \cdot \text{length}_2$



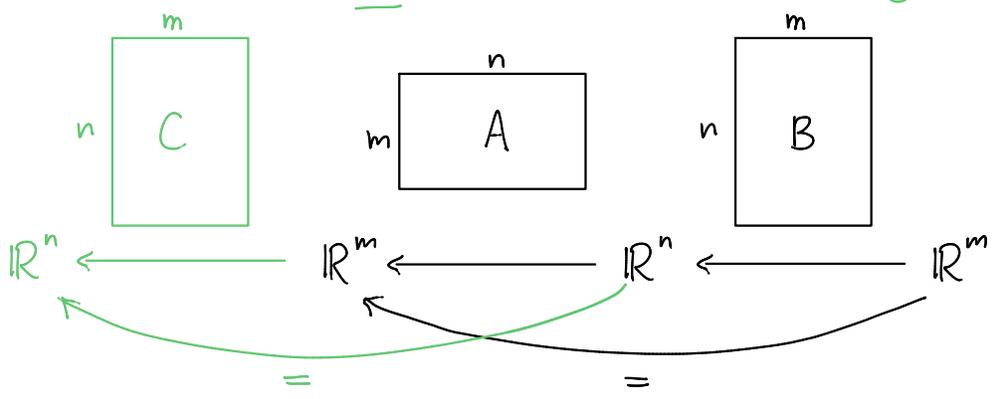
- Rules $A, B \in \mathbb{R}^{m \times n} \Rightarrow$
- $\cdot A+B = B+A$
 - $\cdot c(dA) = (cd)A$
 - $\cdot (A+B)+C = A+(B+C)$
 - $\cdot c(A+B) = cA+cB$
 - $\cdot 0+A = A$
 - $\cdot (c+d)A = cA+dA$
 - $\cdot A+(-A) = 0$
 - $\cdot 1A = A$

(Proof: Matrices are just vectors drawn as rectangles. \square) and you already know these rules for those.

$A \in \mathbb{R}^{m \times n}, A' \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times q}, t \in \mathbb{R} \Rightarrow$

- $\cdot A I_n = A = I_m A$
- $\cdot (tA)B = t(AB) = A(tB)$
- $\cdot (A+A')B = AB+A'B$
- $\cdot (AB)C = A(BC)$ (Pf: Both are the matrix for $\mu_A \circ \mu_B \circ \mu_C$)

Def: For $A \in \mathbb{R}^{m \times n}$, a right inverse is an $n \times m$ matrix B with $AB = I_m$.
 a left inverse is an $n \times m$ matrix C with $CA = I_n$.



A is invertible if A is square and there is a matrix B with $AB = I_n$ and $BA = I_n$. Notation: $B = A^{-1}$.

E.g. $\begin{bmatrix} 1 & 5 \\ 1 & 3 \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} -3 & 5 \\ 1 & -1 \end{bmatrix} : \begin{bmatrix} 1 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -3 & 5 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} [-3]+[5] & [5]-[3] \\ [-3]+[3] & [5]-[3] \end{bmatrix}$
 $= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \checkmark$

Thm 3.2: A is invertible $\Leftrightarrow A$ is nonsingular.

Pf:

reduced echelon form of A is I_n
rank = $n = m$

$A \rightsquigarrow I_n \Leftrightarrow$

(i.e. solve $Ax_1 = e_1$

$[A | I_n] \rightsquigarrow [I_n | B]$ for some $n \times n$ B .

has same sols!

"for all"

$\Leftrightarrow Ax_j = e_j \forall j$ has sols $x_j = b_j \forall j$

$\Leftrightarrow Ab_j = e_j \forall j$

$\Leftrightarrow AB = I_n$.

$Ax_2 = e_2$
 \vdots
 $Ax_n = e_n$

$I_n x_1 = b_1$
 $I_n x_2 = b_2$
 \vdots
 $I_n x_n = b_n$

But also $[A | I_n] \rightsquigarrow [I_n | B] \Rightarrow [I_n | B] \rightsquigarrow [A | I_n]$

$\Rightarrow [B | I_n] \rightsquigarrow [I_n | A]$ Same row operations!

$\Rightarrow BA = I_n. \square$

Cor 3.3: If A, B are $n \times n$ and $BA = I_n$ then $B = A^{-1}$ and $A = B^{-1}$.

Caution: $n \times n$: B right inverse of $A \Rightarrow B$ left inverse of A

$m \times n$ with $m \neq n$: FALSE! Do not make this error!

Pf: A nonsingular \Leftrightarrow rank $A = \#$ cols $\Leftrightarrow Ax = 0$ has only trivial sol.

Assume $BA = I_n$. Then $Ax = 0 \Rightarrow 0 = B0 = BAx$

has only the trivial sol!

$= (BA)x = I_n x = x$

Thus, by Thm 3.2, A has an inverse A^{-1} .

But then $BA = I_n \Rightarrow BAA^{-1} = I_n A^{-1}$
 $\parallel \quad \parallel$
 $B \quad A^{-1}. \square$

Geometrically:

$\Leftrightarrow \mathbb{R}^n \xrightarrow{MA} \mathbb{R}^n$ injective

$\Leftrightarrow \mathbb{R}^n \xrightleftharpoons[M^{-1}]{MA} \mathbb{R}^n$ bijective

Def: injective and surjective

9. Review

$$A \begin{bmatrix} e_1 \\ 0 \\ 0 \end{bmatrix} = ? \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

$$A \begin{bmatrix} 0 \\ e_2 \\ 0 \end{bmatrix} = ? \begin{bmatrix} 3 \\ 11 \\ \pi \end{bmatrix}$$

$$A e_3 = \begin{bmatrix} 5 \\ \pi \end{bmatrix}$$

$$A \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} = ? A(x_1 e_1) = x_1 A e_1 = x_1 \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}$$

$$A(x_2 e_2) = x_2 \begin{bmatrix} 3 \\ 11 \\ \pi \end{bmatrix}$$

$$A(x_3 e_3) = x_3 \begin{bmatrix} 5 \\ \pi \end{bmatrix}$$

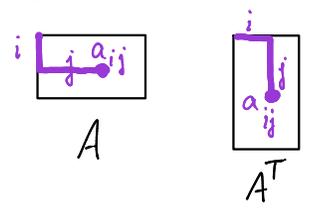
$$\begin{aligned} Ax &= A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A(x_1 e_1 + x_2 e_2 + x_3 e_3) \\ &= x_1 A e_1 + x_2 A e_2 + x_3 A e_3 \\ &= x_1 \begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 11 \\ \pi \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ \pi \end{bmatrix} \\ &= x_1 a_1 + x_2 a_2 + x_3 a_3 = \text{linear combin of cols of } A \text{ with coeffs from } x \end{aligned}$$

Q. $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ 3 & 11 \\ 5 & \pi \end{bmatrix} = ? x_1 [2 \ 7] + x_2 [3 \ 11] + x_3 [5 \ \pi] = \uparrow$

Def: The transpose of a matrix A with a_{ij} in row i and col j is A^T $(A)_{ij} = (A^T)_{ji}$

E.g. $A, A^T \rightarrow$ general: swap rows and cols

E.g. a is col vector $\Rightarrow a^T$ is row vector $a^T a = ? \|a\|^2$
 $x^T a = ? x \cdot a$



$a a^T = ?$... some square matrix...

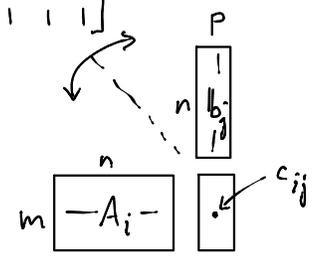
scalar vector $n \times 1$ 1×1

$$\text{proj}_a x = \frac{x \cdot a}{\|a\|^2} a = a \left(\frac{a \cdot x}{\|a\|^2} \right) = a \left(\frac{a^T x}{\|a\|^2} \right) = \frac{a a^T x}{\|a\|^2} = \frac{a a^T}{a^T a} x$$

so $\text{proj}_a = \frac{a a^T}{a^T a}$! E.g. $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \text{proj}_a = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Q. $(AB)^T = ? B^T A^T$ flip page over across

E.g. $(Ax) \cdot y = y^T (Ax) = (y^T A) x = x \cdot (A^T y)$



Important: E is invertible!

• E elementary $\Rightarrow E^{-1}$ elementary of same type.

Pf: (i) $E^{-1} = E$

(ii) replace c by $\frac{1}{c}$

(iii) $-c$

• $E^{-1} = (E_l E_{l-1} \dots E_1)^{-1} = E_1^{-1} \dots E_{l-1}^{-1} E_l^{-1}$.

E.g. If only type (iii) occurs, then E looks like $L = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ * & & & \ddots & \\ & & & & 1 \end{bmatrix}$, lower-triangular with 1's on the diagonal.

In this case $A = LU$ is the LU-decomposition of A .

"Most" but not all matrices A have LU decompositions
↑ "lower" ↑ "upper" (-triangular)
"open Schubert cell" - don't actually say this

Finding constraint equations

For which $b \in \mathbb{R}^4$ does $Ax = b$ have a solution

($b \in$ image of μ_A)

when

$$A = \begin{bmatrix} 1 & 1 & 3 & -1 & 0 \\ -1 & 1 & 1 & 1 & -2 \\ 0 & 1 & 2 & 2 & -1 \\ 2 & -1 & 0 & 1 & -6 \end{bmatrix} ?$$

$$\text{Answer: } [A|b] \rightsquigarrow \left[\begin{array}{ccccc|c} 1 & 1 & -3 & -1 & 0 & b_1 \\ 0 & 2 & 4 & 0 & 2 & b_1 + b_2 \\ 0 & 0 & 0 & 2 & -2 & -\frac{1}{2}b_1 - \frac{1}{2}b_2 + b_3 \\ 0 & 0 & 0 & 0 & 0 & b_1 + 9b_2 - 6b_3 + 4b_4 \end{array} \right] = [U|c]$$

constraint eqn.

Q. $U = EA$; what's E ?

10. Chapter 3: Vector spaces

\mathbb{R}^n is a (real) vector space (over \mathbb{R}) *or:* there are others; general def' after midterm

\mathbb{C}^n	(complex)	\mathbb{C}
\mathbb{Q}^n	(rational)	\mathbb{Q}
\mathbb{F}_2	(binary)	\mathbb{F}_2

These are the "flat things" we've been talking about all along, though only the ones through 0.

Def: A subset $V \subseteq \mathbb{R}^n$ is a subspace if

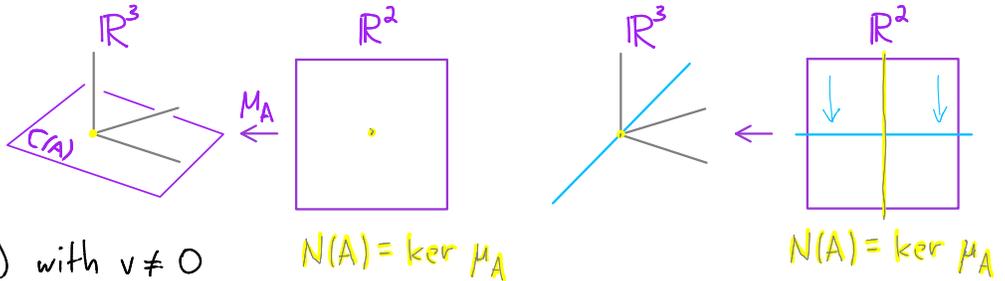
- $V \neq \emptyset$ and
- $v, w \in V$ and $c \in \mathbb{R} \Rightarrow v + cw \in V$.

Remark 1. can use instead: $0 \in V$ ($v=w; c=-1$) *but it's not easier to check*

2. $\Rightarrow V$ closed under arbitrary linear combinations (Pf: induction.)

Examples

- $V = \{0\}$
- $V = \mathbb{R}^n$



- line $l = \text{span}(v)$ with $v \neq 0$
 - hyperplane $H = \{x \in \mathbb{R}^n \mid ax = 0\}$ for some fixed $a \in \mathbb{R}^{\text{row}}$
 - the nullspace of any matrix $A \in \mathbb{R}^{m \times n}$ *Pf: $A0 = 0$ and μ_A is linear!*
- $N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$ (= kernel of μ_A) = sols $[A|0]$

- $\text{span}(v_1, \dots, v_k)$ for any $\begin{bmatrix} v_1, \dots, v_k \\ 1 \dots 1 \end{bmatrix} \in \mathbb{R}^{n \times k}$ *Pf: 1. $0 = 0v_i$. 2. $v = c_1v_1 + \dots + c_kv_k$ and $w = d_1v_1 + \dots + d_kv_k \Rightarrow v + cw = (c_1 + cd_1)v_1 + \dots + (c_k + cd_k)v_k \in V$.*

The point is that none of this is new; you already know what subspaces look like and how to describe all of them — yes, all:

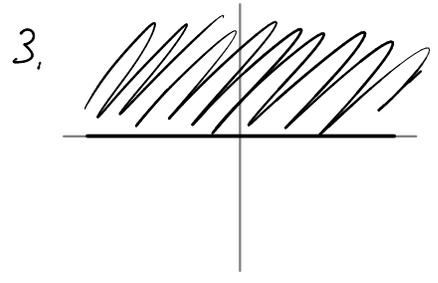
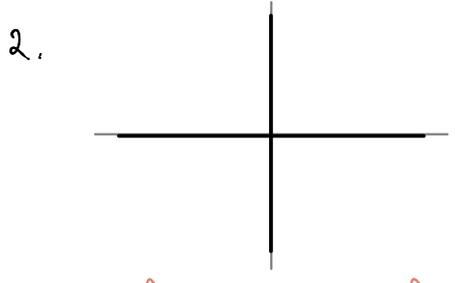
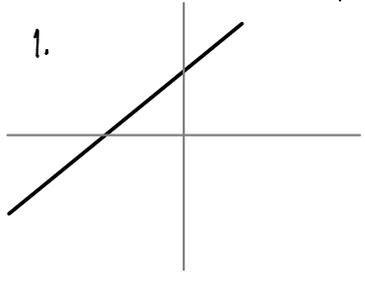
Thm: Every subspace of \mathbb{R}^n has the form of #5. #6.

Pf: For #6, keep finding "independent" vectors that increase dimension until you can't anymore: if $\exists v_1 \neq 0$ in V , then $\text{span}(v_1) \subseteq V$.
 $v_2 \notin \text{span}(v_1)$ $(v_1, v_2) \in V$
 \vdots
 stops at v_k for some $k \leq n$

because $\text{span}(v_1, \dots, v_k) = C(A)$ for $A = \begin{bmatrix} v_1 & \dots & v_k \\ 1 & \dots & 1 \end{bmatrix}$ is spanned by r vectors, where $r = \text{rank } A \leq n$.

For #5: find constraint equations, given v_1, \dots, v_k . \square

Q. Is this a subspace?



4. $x_0 + \text{span}(v_1, \dots, v_k)$ yes, if $x_0 \in \text{span}$; no if not

5. $U+V = \{u+v \mid u \in U \text{ and } v \in V\}$ for subspaces U and V

Yes: $1. 0+0 \in U+V; \quad 2. \left. \begin{matrix} u+v \in U+V \\ u'+v' \in U+V \\ c \in \mathbb{R} \end{matrix} \right\} \Rightarrow u+v+c(u'+v') = (u+cu') + (v+cv') \in U+V.$

Note: $U+V$ is the smallest subspace containing U and V .

E.g. $x\text{-axis} + y\text{-axis} = xy\text{-plane}.$

Def: For any subset $C \subseteq \mathbb{R}^n_{\text{col}}$ set $C^\perp = \{u \in \mathbb{R}^n_{\text{row}} \mid ux = 0 \forall x \in C\}$

$R \subseteq \mathbb{R}^n_{\text{row}}$ set $R^\perp = \{x \in \mathbb{R}^n_{\text{col}} \mid ux = 0 \forall u \in R\}$

WARNING: The book calls C^\perp what I call $(C^\perp)^\top$; book thinks every vector is a column! C^\perp in book: collection of columns
 will make even more sense when we get to row space $R(A)$ C^\perp for us: collection of rows

Prop: C^\perp is a subspace of $\mathbb{R}^n_{\text{row}}$

R^\perp " " " " $\mathbb{R}^n_{\text{col}}$

Pf: 1. $0x = 0. \checkmark$

2. $vx = 0$ and $wx = 0 \Rightarrow (v+cw)x = vx+cwx = 0+c0 = 0 \forall c \in \mathbb{R}.$

For R^\perp , transpose the argument. \square

Lemma: $V^\top \subseteq \mathbb{R}^n_{\text{row}}$ is a subspace $\Leftrightarrow V \subseteq \mathbb{R}^n$ is a subspace.

Pf: Transpose is linear. \square check! $(A+cB)^\top = A^\top + cB^\top \forall A, B \in \mathbb{R}^{m \times n}, c \in \mathbb{R}$

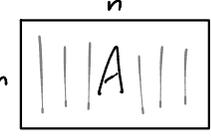
Def: $V \subseteq \mathbb{R}^n_{\text{col}}$ and $W \subseteq \mathbb{R}^n_{\text{row}}$ are orthogonal, written $V \perp W$,

- if $WV = 0$. Equivalently:
- $wv = 0 \forall w \in W \text{ and } v \in V$
 - $W \subseteq V^\perp$
 - $V \subseteq W^\perp$

Also, $V \in \mathbb{R}_{col}^m$ and $W \in \mathbb{R}_{col}^n$: $V \perp W$ if $\cdot W^T V = 0$ $W^T \perp V$
 $\cdot V^T W = 0$ $V^T \perp W$
 $\cdot V \cdot W = 0$



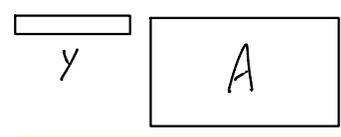
E.g. $C(A)$ = column space of A
 = image of M_A subspace as in #6.



$L(A)$ = left nullspace of $A = \{y \in \mathbb{R}_{row}^m \mid yA = 0\}$
 = kernel of $\rho_A: \mathbb{R}_{row}^m \rightarrow \mathbb{R}_{row}^n$ y times any column of A is 0.

book omits this
 $= N(A^T)^\perp$

$y \mapsto yA$
 right multiplication by A



WARNING: In the book's notation, $L(A) = N(A^T)^\perp$

Prop: $L(A) = C(A)^\perp$.

CAN SKIP:
 $N(A) = R(A)^\perp$
 PROVED IN NEXT LECTURE

Pf: $y \in L(A) \Rightarrow yA = 0 \Rightarrow y a_j = 0 \forall j = 1, \dots, n$
 $\Rightarrow c_1 y a_1 + \dots + c_n y a_n = 0 \forall c_1, \dots, c_n \in \mathbb{R}$ so
 $\Rightarrow y(c_1 a_1 + \dots + c_n a_n) = 0$ "
 $\Rightarrow y \in C(A)^\perp$. Hence $L(A) \subseteq C(A)^\perp$.

But $y \in C(A)^\perp \Rightarrow y a_j = 0 \forall j$ by def. (since $a_j \in C(A)$)
 so $yA = 0$; hence $C(A)^\perp \subseteq L(A)$. \square

Similarly:

Def: $R(A)$ = row space of $A = \text{span}(A_1, \dots, A_m) \subseteq \mathbb{R}_{row}^n$

\swarrow the rows of A

Cor: $N(A) = R(A)^\perp$.

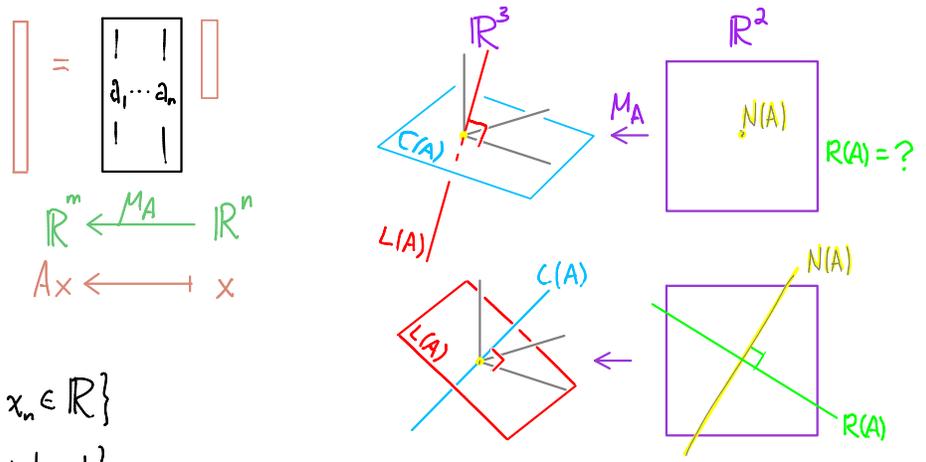
$$\begin{matrix} N(A) & R(A)^\perp \\ \parallel & \parallel \\ L(A^T)^\perp & = (C(A^T)^\perp)^\perp \end{matrix}$$

Pf: Transpose previous Prop. \square

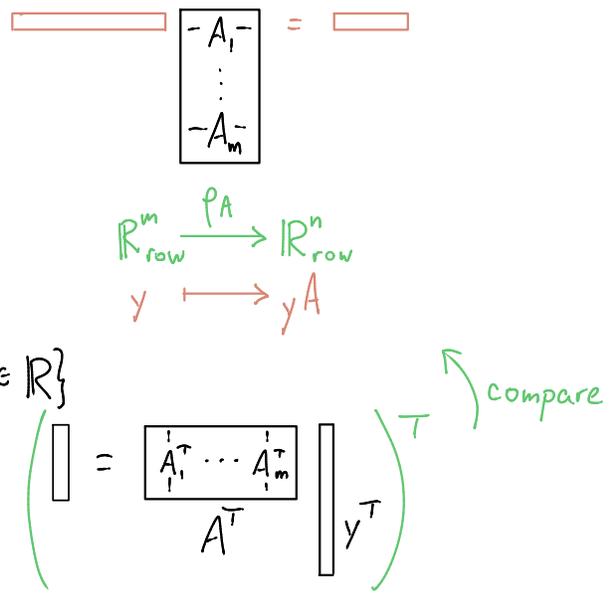
WARNING: The book calls $R(A)^T$ what I call $R(A)$; book thinks every vector is a column!

11. Four subspaces, given $A \in \mathbb{R}^{m \times n}$ $\begin{matrix} n \\ \boxed{A} \\ m \end{matrix}$ or $\begin{matrix} n \\ \boxed{A} \\ m \end{matrix}$ both possible, but let's proceed with this, so $m \geq n$.
 (\mathbb{R} could be any field)

- $C(A)$ = column space
 = $\text{span}(a_1, \dots, a_n)$
 = $\text{image}(\mu_A)$
 = $\mu_A(\mathbb{R}^n)$
 = $\{Ax \mid x \in \mathbb{R}^n\}$
 = $\{x_1 a_1 + \dots + x_n a_n \mid x_1, \dots, x_n \in \mathbb{R}\}$
 = $\{b \in \mathbb{R}^m \mid Ax = b \text{ is consistent}\}$



- $R(A)$ = row space
 = $\text{span}(A_1, \dots, A_m) \in \mathbb{R}_{\text{row}}^n$
 = $\text{image}(p_A)$
 = $p_A(\mathbb{R}_{\text{row}}^m)$
 = $\{yA \mid y \in \mathbb{R}_{\text{row}}^m\}$
 = $\{y_1 A_1 + \dots + y_m A_m \mid y_1, \dots, y_m \in \mathbb{R}\}$
 = $(\underbrace{\text{image}(\mu_{A^T})}_{\subseteq \mathbb{R}_{\text{col}}^n})^T \subseteq \mathbb{R}_{\text{row}}^n$



- $N(A)$ = nullspace
 = $\{x \in \mathbb{R}^n \mid Ax = 0\}$
 = $\text{kernel}(\mu_A)$
 = $R(A)^\perp$ by Prop 2.2.

recall Pf: $x \in N(A) \Rightarrow Ax = 0 \Rightarrow (yA)x = 0 \forall y \in \mathbb{R}_{\text{row}}^m \Rightarrow x \in R(A)^\perp$.
 $x \in R(A)^\perp \Rightarrow A_i x = 0 \forall i = 1, \dots, m$ since $A_i \in R(A) \Rightarrow x \in N(A)$. \square

E.g. Find vectors that span $N(A)$ for $A = \begin{bmatrix} 1 & -1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & -1 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 3 & 0 & -3 \end{bmatrix}$
 $A_2 = A_3 - A_1$
 $A_3' = A_3 - A_1$
 pivots
 (iii) $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$
 $A_1' = A_1 + A_2$
 $A_3' = A_3 - 3A_2$
 free variables

Sol: Reduced echelon form of A is

$$Ax = 0 \Leftrightarrow Ux = 0 \quad (N(A) = N(U))$$

$$\Leftrightarrow \begin{cases} x_1 + x_3 + x_4 = 0 \\ x_3 - x_4 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1 = -x_3 - x_4 \\ x_2 = x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases}$$

$$\Leftrightarrow x = x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$v_1 \qquad v_2$

So $N(A) = N(U) = \text{span}(v_1, v_2)$.

E.g. Find constraint equations for $C(A)$.

sol: $[A|b] \rightsquigarrow [U|c] = E[A|b]$. What's E ?

(iii) $\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 3 & 0 & -3 \end{bmatrix}$
 $A_3' = A_3 - A_1$
 $A_3' = A_3 - A_1$
 (ii) $\begin{bmatrix} 1 & -1 & 1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 3 & 0 & -3 \end{bmatrix}$
 $A_1' = A_1 + A_2$
 $A_3' = A_3 - 3A_2$

$$E = E_2 E_1 = \begin{bmatrix} 1 & 1 & & \\ & 1 & & \\ & & -3 & 1 \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ -1 & & 1 & \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \\ -1 & 1 & 0 & \\ 2 & -3 & 1 & \end{bmatrix}$$

$E_2 \qquad E_1$

$$b \in C(A) \Leftrightarrow Eb \in C(U)$$

$$\Leftrightarrow E_i b = 0 \text{ whenever } E_i A = 0$$

Constraint equations: $\boxed{\{E_i b = 0 \mid E_i A = 0\}}$
 $E_i \in L(A)$

Here, $E_i A = (EA)_i = 0$ only for $i=3$, so constraint equations are

$$\boxed{[2 \ -3 \ 1] b = 0}$$

or $2b_1 - 3b_2 + b_3 = 0$

- 4. $L(A)$ = left nullspace
- = $\{y \in \mathbb{R}^m \mid yA = 0\}$
- = $\ker(p_A)$
- = $N(A^T)^T$
- = subspace of constraints on $\text{image}(\mu_A) = C(A)$

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{bmatrix} A_1^T & \cdots & A_m^T \\ \hline A^T \end{bmatrix} \begin{bmatrix} y^T \end{bmatrix}$$

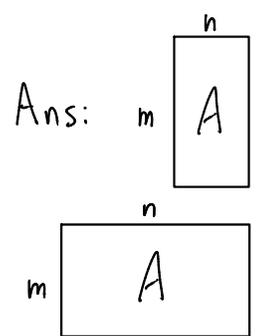
Prop 2.4: $C(A) = L(A)^\perp$. Compare: $L(A) = C(A)^\perp$ last time!

pf: $b \in C(A) \Rightarrow b = Ax$ for some $x \in \mathbb{R}^n$
 $\Rightarrow yb = (yA)x = 0 \ \forall y \in L(A) \Rightarrow b \in L(A)^\perp$. Thus $C(A) \subseteq L(A)^\perp$.

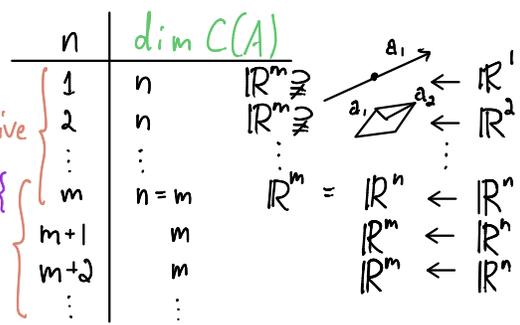
For the reverse containment, \exists constraints y_1, \dots, y_k such that $b \in C(A) \Leftrightarrow y_1 b = \dots = y_k b = 0$.
 Taking $b \in \{a_1, \dots, a_n\}$ yields $y_1, \dots, y_k \in L(A)$. Hence $\{y_1, \dots, y_k\}^\perp = C(A) \Rightarrow L(A)^\perp \subseteq C(A)$. \square
 to be in $L(A)^\perp$ you've got to be \perp to (perhaps) more stuff

12. Q. $A \in \mathbb{R}^{m \times n} \Rightarrow \dim C(A) = ?$

no coincidence
m fixed



n , unless there's some coincidence *injective*
 m , " " " " " *surjective*
both



injective means no collapsing: $T(x) = T(y) \Rightarrow x = y$ "into", "one-to-one"

surjective means target covered: $\forall b \in \text{target} \Rightarrow T(x) = b$ for some $x \in \text{source}$ "onto"

Q \Leftrightarrow Q': Given n vectors $v_1, \dots, v_n \in \mathbb{R}^m$, what is $\dim \text{span}(v_1, \dots, v_n)$?

coincidence: span has dim less than possible

Def: vectors v_1, \dots, v_k are linearly dependent if $c_1 v_1 + \dots + c_k v_k = 0$

for some $c_1, \dots, c_k \in \mathbb{R}$ not all 0. "some nontrivial linear combination vanishes"

Prop: v_1, \dots, v_k linearly dependent $\Leftrightarrow \dim \text{span}(v_1, \dots, v_k) < k \Leftrightarrow v_i \in \text{span}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k)$ for some i .

Pf: \Leftrightarrow : by def!

\Rightarrow : $c_1 v_1 + \dots + c_k v_k = 0$ with $c_i \neq 0 \Rightarrow v_i = -\frac{1}{c_i} \sum_{j \neq i} c_j v_j \in V_i$.

\Leftarrow : $v_i = x_1 v_1 + \dots + x_{i-1} v_{i-1} + x_{i+1} v_{i+1} + \dots + x_k v_k = \sum_{j \neq i} x_j v_j \in V_i$
 $\Rightarrow \sum_{j=1}^k x_j v_j = 0$, where $x_i = -1$. \square

E.g. Are $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -4 \end{bmatrix}, \begin{bmatrix} 8 \\ 1 \\ -1 \end{bmatrix}$ linearly dependent?

Sol:

$\Leftrightarrow N(A) \neq \emptyset$ for $A = \begin{bmatrix} 3 & -1 & 8 \\ 0 & 1 & 1 \\ 1 & -4 & -1 \end{bmatrix}$!

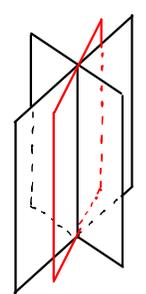
$x_1 v_1 + x_2 v_2 + x_3 v_3 = 0 \Leftrightarrow x \in N(A)$
dependence relation \Leftrightarrow also $x \neq 0$

Moral: $N(A) \setminus \{0\} =$ linear dependence relations on columns of A

$\begin{bmatrix} 3 & -1 & 8 \\ 0 & 1 & 1 \\ 1 & -4 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} \in N(A) \setminus \{0\}$, so: yes.

Q. Given rows A_1, \dots, A_m of $A \in \mathbb{R}^{m \times n}$, what is $\dim(\{A_1 x = 0\} \cap \dots \cap \{A_m x = 0\})$?

A. $n - m$ if $n \geq m$, unless there is some coincidence!
 0 if $n < m$, " " " " " " \boxed{A}



a_1, \dots, a_n linearly independent in \mathbb{R}^m

\Leftrightarrow if $x_1 a_1 + \dots + x_n a_n = 0$ then $x_i = 0 \forall i$

\Leftrightarrow no column of A lies in the span of the others for $A = [a_1 \dots a_n]$

$\Leftrightarrow N(A) = 0$

$\Leftrightarrow 0$ can be expressed uniquely as a linear combination of a_1, \dots, a_n

$\Leftrightarrow b \in \text{span}(a_1, \dots, a_n) = C(A)$ is uniquely a linear combination of a_1, \dots, a_n

$\Leftrightarrow Ax = 0$ has only one solution $[A|b]$ consistent

$\Leftrightarrow Ax = b$ has only one solution when $b \in C(A)$

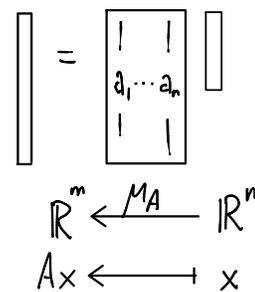
$\Leftrightarrow Ax = b$ has at most one solution $\forall b \in \mathbb{R}^m$

$\Leftrightarrow M_A$ is injective

$\Leftrightarrow M_A$ does not decrease dimension

$\Leftrightarrow \text{rank } A = n$

$\Leftrightarrow A$ has a left inverse



Prop 3.2: Assume v_1, \dots, v_k are linearly independent.

Then v_1, \dots, v_k, v is linearly independent $\Leftrightarrow v \notin \text{span}(v_1, \dots, v_k)$.
dependent $\Leftrightarrow v \in \text{span}(v_1, \dots, v_k)$

Pf: \Rightarrow : Suppose $c_1 v_1 + \dots + c_k v_k + c v = 0$. Then $c \neq 0$ since $c = 0 \Rightarrow$
 $c_1 v_1 + \dots + c_k v_k = 0 \Rightarrow c_1 = \dots = c_k = 0$ because v_1, \dots, v_k are linearly independent.
 So $v = -\frac{1}{c}(c_1 v_1 + \dots + c_k v_k)$.

\Leftarrow : Previous prop. (Doesn't need v_1, \dots, v_k linearly independent.) \square

E.g. Is 0 linearly independent? No: $1 \cdot 0 = 0$

E.g. Is $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ linearly independent? No: $1 \cdot v - 1 \cdot v = 0$. \Rightarrow need multisets technically

E.g. Is $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ $v \notin \text{span}(v_1, v_2)$ but v_1, v_2, v not linearly independent. {...} = set, so I list the vectors with no brackets
 Why doesn't this contradict Prop.?

E.g. Prove that Av_1, \dots, Av_k are linearly independent if v_1, \dots, v_k are linearly independent and $A \in \mathbb{R}^{m \times n}$ has rank n .

Sol. Suppose $c_1 Av_1 + \dots + c_k Av_k = 0$. Then $A(c_1 v_1 + \dots + c_k v_k) = 0$, so $c_1 v_1 + \dots + c_k v_k = 0$ because M_A is injective. Thus $c_1 = \dots = c_k = 0$ since v_1, \dots, v_k are linearly independent.

13. Def: v_1, \dots, v_k is a basis for a subspace $V \subseteq \mathbb{R}^n$ if v_1, \dots, v_k

- (i) span V and not too few
 - (ii) are linearly independent. not too many
- } Goldilocks

Q. Name a basis for \mathbb{R}^n .

A. standard basis e_1, \dots, e_n : (i) $x = x_1 e_1 + \dots + x_n e_n \quad \forall x \in \mathbb{R}^n$. \mathbb{I}_n
 (ii) $x = x_1 e_1 + \dots + x_n e_n = 0 \Rightarrow x = 0$. (or: $\text{rank} \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} = n$)

E.g. Which of these are bases?

1. $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ in $V = \mathbb{R}^3$
 v_1 v_2 v_3

$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 5 & 3 \end{bmatrix} \xrightarrow{A'_3 = A_3 - A_1, A'_2 = A_2 - 2A_1} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 3 & 2 \end{bmatrix} \xrightarrow{EA} \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \in N(A) \setminus \{0\}$
dependence relation
 $v_1 - 2v_2 + 3v_3 = 0$

no: fails (ii) (\Rightarrow fails (i))

2. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ in $V = \mathbb{R}^3$

no: too many vectors

Lemma: a_1, \dots, a_n linearly dependent in \mathbb{R}^m if $n > m$.

equivalent conditions super useful! Pf: $\text{rank } A \leq m < n$. \square rank $< n \Leftrightarrow$ dep.

4. $\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \\ 2 \end{bmatrix}$ in $V = \mathbb{R}^4$

Yes: $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -2 \\ 2 & 1 & 4 & 1 \\ 3 & 1 & 4 & 2 \end{bmatrix} \xrightarrow{m} \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -2 \\ 0 & 1 & 2 & -3 \\ 0 & 1 & 1 & -4 \end{bmatrix} \xrightarrow{m} \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -2 \end{bmatrix} \Rightarrow \text{rank} = 4$

Note: If #vectors = $\dim V$ then (i) \Leftrightarrow (ii) No: they don't span by Prop. 3.2.

v_1, \dots, v_k linearly independent $\Leftrightarrow v_1, \dots, v_k$ is a basis for $\text{span}(v_1, \dots, v_k)$

All conditions equivalent to "linearly independent" work for "basis for their span".

Prop 3.4: $A \in \mathbb{R}^{n \times n}$ nonsingular \Leftrightarrow cols of A form a basis of \mathbb{R}^n .

Pf: A nonsingular $\Leftrightarrow N(A) = 0 \Leftrightarrow \mu_A$ is bijective $\Leftrightarrow \text{rank } A = n$. \square

Thm 3.5: Every subspace $V \subseteq \mathbb{R}^n$ has a basis.

Pf: $V = \text{span}(v_1, \dots, v_k)$ for some $v_1, \dots, v_k \in V$ proved in Lec. 10.

Assume k is minimal. Then $v_i \notin \text{span}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k)$ by assumption (!),

so v_1, \dots, v_k is independent by \square . \square

Thm 4.2: v_1, \dots, v_k and w_1, \dots, w_l are two bases for a subspace $V \subseteq \mathbb{R}^n \Rightarrow k = l$.

Pf: $w_i \in \text{span}(v_1, \dots, v_k) \Rightarrow w_i = Ax_i$ for some $x_i \in \mathbb{R}^k$, where $A = \begin{bmatrix} | & & | \\ v_1 & \dots & v_k \\ | & & | \end{bmatrix} \in \mathbb{R}^{n \times k}$

$$\begin{matrix} \vdots \\ w_l = Ax_l & \text{"} & \text{"} & x_l & \text{"} \end{matrix}$$

$$n \begin{matrix} l \\ \left[\begin{array}{c|c} | & | \\ w_1 & \dots & w_l \\ | & & | \end{array} \right] \end{matrix} = n \begin{matrix} k \\ \left[\begin{array}{c|c} | & | \\ v_1 & \dots & v_k \\ | & & | \end{array} \right] \end{matrix} \begin{matrix} l \\ \left[\begin{array}{c|c} | & | \\ x_1 & \dots & x_l \\ | & & | \end{array} \right] \end{matrix} k$$

$$W = AX \quad \text{Want: } X \text{ is square.}$$

$l > k \Rightarrow N(X) \neq \{0\} \Rightarrow \exists y \in N(X) \setminus \{0\}$

$$\Rightarrow Wy = (AX)y = A(Xy) = A0 = 0$$

so $y \in N(W) \setminus \{0\}$. Thus w_1, \dots, w_l are linearly dependent.

Therefore $l \leq k$. By symmetry, $k \leq l$. Thus $k = l$. \square

Cor: $\dim V = \text{size of any basis of } V$.

E.g. $\dim C(A) = \text{rank } A : C(A) = \text{span}(v_1, \dots, v_r)$ for cols v_1, \dots, v_r of A with r minimal $\Rightarrow v_1, \dots, v_r$ linearly independent.

$\dim C(A) \leq n$. Why? $\text{image}(\mu_A)$ spanned by $\mu_A(e_1), \dots, \mu_A(e_n)$
 a_1, \dots, a_n

$\mu_A(e_1), \dots, \mu_A(e_n)$ might be linearly dependent, but they do still span $\text{image}(\mu_A)$.

Grammar: v_1, \dots, v_k is a basis \mathcal{B}

$v_i \in \mathcal{B}$ is a basis element (or basis vector)

v_1, \dots, v_k and w_1, \dots, w_k are bases

14. Bases for $R(A)$, $C(A)$, $N(A)$, $L(A)$

Thm 4.5: Fix $A \in \mathbb{R}^{m \times n}$ and $U = EA$ the reduced echelon form of A , with E invertible. *e.g. $E \neq 0$ and doesn't even have any 0 rows*

E.g.

$$\begin{matrix} \begin{bmatrix} 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ -1 & -1 & 1 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ 1 & 2 & 1 & 1 & 6 \\ 0 & 1 & 1 & 1 & 3 \\ 2 & 2 & 0 & 1 & 7 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ E & A & & U \end{matrix}$$

$R(A)$: The nonzero rows of U form a basis.

E.g. basis for $R(A)$ is

$$\begin{bmatrix} \textcircled{1} & 0 & -1 & 0 & 1 \\ 0 & \textcircled{1} & 1 & 0 & 2 \\ 0 & 0 & 0 & \textcircled{1} & 1 \end{bmatrix} \begin{matrix} U_1 \\ U_2 \\ U_3 \end{matrix}$$

Pf: The rows of U are linear combinations of the rows of A with coefficients from (rows of) E . Thus $\text{rows}(U) \subseteq R(A)$, so $R(U) \subseteq R(A)$.

But $EA = U \Rightarrow A = E^{-1}U \Rightarrow R(A) \subseteq R(U)$, so $R(U) = R(A)$. *Now: are they a basis?*

The pivot rows U_1, \dots, U_r of U are independent because

(*) $c_1U_1 + \dots + c_rU_r$ has entries c_1, \dots, c_r in the pivot columns

$$c_1U_1 + c_2U_2 + c_3U_3 = \begin{bmatrix} c_1 & c_2 & c_2 - c_1 & c_3 & c_1 + 2c_2 + c_3 \end{bmatrix}$$

$\uparrow \quad \uparrow \quad \quad \uparrow$

so it equals 0 $\Rightarrow c_1 = \dots = c_r = 0$. \square

$L(A)$: The rows of E corresponding to zero-rows of U form a basis

E.g. basis for $L(A)$ is $[-1 \ -1 \ 1 \ 1] = E_4$

Pf: First compute $L(U) = \text{span}(e_{r+1}^T, \dots, e_m^T)$, where $r = \text{rank } A$, which holds

because $c_1U_1 + \dots + c_mU_m = 0 \Leftrightarrow c_1U_1 + \dots + c_rU_r = 0$ (since $U_{r+1}, \dots, U_m = 0$)

$\Rightarrow c_1 = \dots = c_r = 0$ (by (*) or better, by basis for $R(A)$).

Compare $L(U)$ to $L(A)$: $y \in L(U) \Leftrightarrow yU = 0$

$$\Leftrightarrow yEA = 0$$

$$\Leftrightarrow yE \in L(A),$$

so $L(A) = L(U)E$

$= \text{span}(e_{r+1}^T, \dots, e_m^T) E$
 (i) $= \text{span}(E_{r+1}, \dots, E_m)$ independent because E is invertible. \square (ii)

$N(A)$: Make U into an $n \times n$ matrix U' as follows.

1. Move rows down so all pivots sit on diagonal.
2. Add or delete 0's to ensure $n \times n$.
3. Set $\text{diag} = -1$. Note: suffices to change all 0's on diag to -1

The free-variable columns of U' are a basis for $N(A)$.

E.g. $U = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow U' = \begin{bmatrix} -1 & 0 & -1 & 0 & 1 \\ 0 & -1 & 1 & 0 & 2 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$

Pf: Solve equation $Ux = 0$ ($\Leftrightarrow Ax = 0$):

(pivot var)₁ + later terms = 0 \Leftrightarrow later terms = -(pivot var)₁
 \vdots
 (pivot var)_r + later terms = 0 \Leftrightarrow later terms = -(pivot var)_r.
 only involve free vars!

So insert rows - free var = - free var. \square

$x_1 - x_3 + x_5 = 0 \Leftrightarrow 0 + \begin{bmatrix} -1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} +1 \\ +2 \\ +0 \\ +1 \\ -1 \end{bmatrix} x_5 = -x_1$
 v_1, v_2 is a basis for $N(A)$.

$\uparrow \uparrow$ pivot cols don't matter! Hence -1 vs. 0 okay.

$C(A)$: The pivot columns of A form a basis.

E.g. $A = \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ 1 & 2 & 1 & 1 & 6 \\ 0 & 1 & 1 & 1 & 3 \\ 2 & 2 & 0 & 1 & 7 \end{bmatrix} \Rightarrow C(A) = \text{span} \left(\begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right)$. v_1, v_2 say that the free var cols of A lie in $\text{span}(\text{pivot var cols})$.

Pf: Our basis of $N(A)$ says that $a_j \in \text{span}(\text{pivot cols})$ if j is a free var col.

(E.g. v_2 says $a_1 + 2a_2 + a_4 - a_5 = 0$)

Thus the pivot cols span $C(A)$. But #pivot cols = $\text{rank } A = \dim C(A)$. \square

Cor 4.6: 1. $\dim R(A) = \dim C(A) = \text{rank } A$. *any field rank - nullity theorem*

2. $\dim N(A) = \# \text{cols} - \text{rank } A$: $A \in \mathbb{R}^{m \times n} \Rightarrow \dim C(A) + \dim N(A) = n$

$T \text{ linear} \Rightarrow \dim(\ker T) + \dim(\text{im } T) = \dim(\text{source } T)$

3. $\dim L(A) = m - \text{rank } A$.

This is how rank-nullity is usually used.

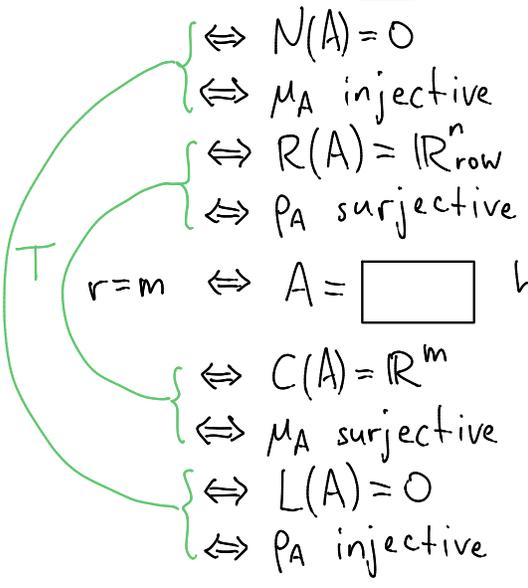
Pf: $\text{rank } A = \# \text{pivots}$. \square

Prop 4.8: $V \subseteq \mathbb{R}^n$ subspace of $\dim k \Rightarrow \dim V^\perp = n - k$.

Pf: V has basis v_1, \dots, v_k . Let $A \in \mathbb{R}^{n \times k}$ have these cols. $V = C(A)$ and $V^\perp = L(A)$. \square

Summary $A \in \mathbb{R}^{m \times n}$ $r = \text{rank } A$

$r = n \Leftrightarrow A = \begin{bmatrix} \square \\ \square \\ \square \\ \square \end{bmatrix}$ has \cdot all columns independent
 \cdot n rows independent \leftarrow somewhere in A



$r = m \Leftrightarrow A = \begin{bmatrix} \square & \square & \square & \square \end{bmatrix}$ has \cdot all rows independent
 \cdot m columns independent

$m = r = n \Leftrightarrow A = \begin{bmatrix} \square & \square & \square \\ \square & \square & \square \\ \square & \square & \square \end{bmatrix}$ has all $m = n$ rows and columns independent

$\Leftrightarrow A$ nonsingular \Leftrightarrow invertible

$\Leftrightarrow N(A) = \{0\}$

$\Leftrightarrow C(A) = \mathbb{R}^n$

$\Leftrightarrow \mu_A$ bijective $\Leftrightarrow \mu_A$ injective $\Leftrightarrow \mu_A$ surjective

$\Leftrightarrow \rho_A$ bijective $\Leftrightarrow \rho_A$ surjective $\Leftrightarrow \rho_A$ injective

15. §3.6 Abstract vector spaces ← any field

Def: A vector space over \mathbb{R} is a set V with two operations

- vector addition: $u, v \in V \mapsto u+v \in V$
- scalar multiplication: $v \in V, c \in \mathbb{R} \mapsto cv \in V$

axioms for "field" are similar

satisfying

and $\forall u, v \in V$ and $c, d \in \mathbb{R}$,

V is an abelian group

1. $u+v = v+u \forall u, v \in V$; [commutative]
2. $(u+v)+w = u+(v+w) \forall u, v, w \in V$; [associative]
3. $\exists 0 \in V$ with $0+v = v \forall v \in V$;
4. for each $v \in V \exists -v \in V$ with $v+(-v) = 0$;

5. $c(dv) = (cd)v$;
6. $c(u+v) = cu+cv$;
7. $(c+d)v = cv+dv$;
8. $1v = v$.

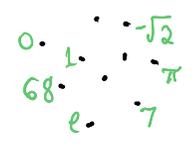
E.g. 1. \mathbb{R}^n or any subspace

2. $\mathbb{R}^{m \times n}$ (or \mathbb{C} or \mathbb{Q} or \mathbb{F}_2)

3. \mathcal{I} any set, $\mathcal{F}(\mathcal{I}) = \text{functions } \mathcal{I} \rightarrow \mathbb{R} \Rightarrow \mathcal{F}(\mathcal{I})$ is a vector space:

$$(cf)(t) = c f(t) \quad (f+g)(t) = f(t) + g(t)$$

Axiom check: not interesting; uses corresponding properties of \mathbb{R} .

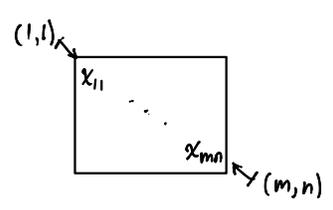


Think: real vectors with entries indexed by \mathcal{I}

e.g. (i) $\mathcal{I} = \{1, \dots, n\}$

$$\begin{matrix} 1 \mapsto x_1 \\ 2 \mapsto x_2 \\ \vdots \\ n \mapsto x_n \end{matrix}$$

(ii) $\mathcal{I} = \{(i, j) \mid i \in \{1, \dots, m\} \text{ and } j \in \{1, \dots, n\}\}$



(iii) $\mathcal{I} = V \Rightarrow \mathcal{F}(V)$ is a vector space

dual vector space $V^* = \{ \text{linear functions } V \rightarrow \mathbb{R} \}$ better reflects structure of V

(iv) $\mathcal{I} = \{1, 2, \dots\} \Rightarrow \mathcal{F}(\mathcal{I}) = \mathbb{R}^\omega$ $\omega = \# \text{positive integers}$

$$x \in \mathbb{R}^\omega \Rightarrow x = (x_1, x_2, \dots)$$

$$cx = (cx_1, cx_2, \dots)$$

$$x + y = (x_1 + y_1, x_2 + y_2, \dots)$$

Def: Let V be a vector space. $U \subseteq V$ is a subspace if $U \neq \emptyset$ and

$v + cw \in U \forall v, w \in U$ and $c \in \mathbb{R}$. (same def as before)

Prop 6.1: $U \subseteq V$ subspace $\Rightarrow U$ is a vector space.

Pf: Already have + and scalar multiplication in V. Def $\Rightarrow u+v \in U$ and $cu \in U$ $\forall u, v \in U$ and $c \in \mathbb{R}$. What's needed: $\forall u, v, w \in U$ and $c, d \in \mathbb{R}$, every equality in the 8 axioms holds in U. But each one already holds in V! \square

e.g. $u+v = v+u$ in V, but both sides lie in U

E.g. 4. $I \subseteq \mathbb{R}$ an interval \Rightarrow {continuous functions $I \rightarrow \mathbb{R}$ } = $C^0(I)$ is a vector space.

Pf: $F(I)$ is a vector space. $0 \in C^0(I)$, and

\cup
 $C^0(I)$ is a subspace: f, g continuous $\Rightarrow f+g$ and cf continuous.

\cup
 $\hat{C}^0(I) =$ {differentiable functions $I \rightarrow \mathbb{R}$ } is a subspace: f, g differentiable $\Rightarrow f+g$ and cf "

\cup
 $C^1(I) =$ {continuously differentiable functions $I \rightarrow \mathbb{R}$ }

\cup
 $C^2(I) =$ { $f \in C^1(I) \mid f' \in C^1(I)$ }

\cup
 $C^3(I) =$ { $f \in C^2(I) \mid f' \in C^2(I)$ }

\cup
 \vdots
 \cup
 $C^\infty(I) = \bigcap_{i=0}^\infty C^i(I)$

5. $\mathcal{P} =$ {polynomial functions $p(t) = a_k t^k + a_{k-1} t^{k-1} + \dots + a_1 t + a_0$ }

\cup
 \mathcal{P}_k subspace $\Rightarrow \mathcal{P}$ is a vector space.

Q. Is {polynomials of degree k} a vector space?

Ans: $(t^2 + 2t - 1) + (-t^2 + 2) = 2t + 1$ so No.
deg: \uparrow_2 \uparrow_2 \uparrow_1

{ $p \in \mathcal{P} \mid \deg p \leq k$ } polynomials of degree $\leq k$

Q. $\dim \mathcal{P} = ?$ ω

$\dim \mathcal{P}_k = ?$ $k+1$

Prop 6.3: \mathcal{P}_k has a basis $1, t, \dots, t^k$.

Pf: $\mathcal{P}_k = \text{span}(1, t, \dots, t^k)$ by definition.

But $p(t) = 0 \Rightarrow a_i = 0 \forall i$. Why?

1. $\deg p \leq k \Rightarrow p$ has $\leq k$ roots if $p \neq 0$.

2. $p(0) = 0 \Rightarrow a_0 = 0$

$p'(0) = 0 \Rightarrow a_1 = 0$

\vdots

$p^{(k)}(0) = 0 \Rightarrow k! a_k = 0 \Rightarrow a_k = 0. \square$

span, linear (in)dependence, basis, dimension — all make sense for arbitrary vector spaces.

6. $V = \{f \in C^1(\mathbb{R}) \mid f' = f\} \subseteq C^1(\mathbb{R})$ subspace: $0' = 0 \checkmark$

$e^x \in V$.

$$(f+cg)' = f' + cg' \\ = f+cg \text{ if } f, g \in V.$$

Claim: e^x is a basis for V .

Pf: Given $f \in V$, let $g(x) = f(x)e^{-x} = \frac{f(x)}{e^x}$.

Want: $g(x) = c$ constant, so $f(x) = ce^x$.

$$\text{Calculate: } g'(x) = f'(x)e^{-x} + f(x)(-e^{-x}) \\ = \underbrace{(f'(x) - f(x))}_0 e^{-x} \\ = 0.$$

Thus $g \equiv c$, so $f(x) = ce^x. \square$

7. $\mathbb{R}^{n \times n}$

$\mathcal{U} = \left\{ \begin{bmatrix} * & & \\ & \Delta & \\ & & * \end{bmatrix} \right\}$ upper-triangular matrices

$\mathcal{L} = \mathcal{U}^T =$ lower-triangular matrices

$\mathcal{D} = \left\{ \begin{bmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{bmatrix} \right\}$ diagonal matrices

$\left\{ \begin{bmatrix} * & & 0 \\ & \Delta & \\ & & * \end{bmatrix} \right\}$

16. Inner products and projections can be made to work \mathbb{C} or $\mathbb{F} \in \mathbb{R}$

Def: Let V be a vector space \mathbb{R} . An inner product on V assigns to each pair $u, v \in V$ a number $\langle u, v \rangle \in \mathbb{R}$ such that $\forall u, v, w \in V$ and scalars c ,

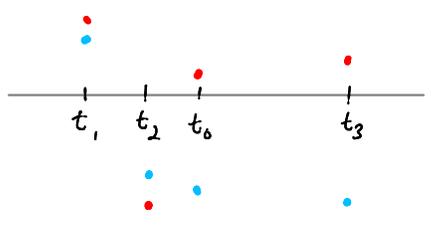
- symmetric 1. $\langle u, v \rangle = \langle v, u \rangle$
- bilinear $\left\{ \begin{array}{l} 2. \langle cu, v \rangle = c \langle u, v \rangle \\ 3. \langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \end{array} \right. \Rightarrow \left. \begin{array}{l} \text{lengths: } \|v\|^2 = \langle v, v \rangle \\ \text{angles: } \cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|} \end{array} \right\}$
- needs $\mathbb{R} \rightarrow$ 4. $\langle v, v \rangle \geq 0$ and $= 0 \Leftrightarrow v = 0$.
positive ——— definite

E.g. (a) $V = \mathbb{R}^n$ $\langle x, y \rangle = y^T x = x \cdot y$

(b) $V = \mathcal{P}_k$ $t_0, \dots, t_k \in \mathbb{R}$ $\langle p, q \rangle = \sum_{i=0}^k p(t_i) q(t_i)$

- 1. \checkmark 2. \checkmark 3. \checkmark

4. $\geq 0 \checkmark$; $\langle p, p \rangle = 0 \Rightarrow \sum_{i=0}^k p(t_i)^2 = 0 = \begin{bmatrix} p(t_0) \\ \vdots \\ p(t_k) \end{bmatrix} \cdot \begin{bmatrix} q(t_0) \\ \vdots \\ q(t_k) \end{bmatrix}$



$\Rightarrow p(t_i) = 0 \forall i$

$\Rightarrow p \equiv 0$ (remember last lecture?)

(c) $V = C^0(I)$ for $I = [a, b]$ $\langle f, g \rangle = \int_a^b f(t)g(t) dt$

same picture but "using all $t \in [a, b]$ "

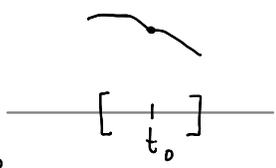
- 1. \checkmark

2. $\int c \dots dt = c \int \dots dt$

3. $\int (f+g)h dt = \int (fh + gh) dt = \int fh dt + \int gh dt$ } \int_a^b is linear!

4. $\int_a^b \underbrace{f(t)^2}_{\geq 0} dt \geq 0$ $f(t_0)^2 \neq 0 \Rightarrow f(t) > \frac{1}{2} f(t_0)^2$ on $[t_0 - \epsilon, t_0 + \epsilon]$

$\Rightarrow \int_a^b f(t)^2 dt > 2\epsilon \cdot \frac{1}{2} f(t_0)^2 = \epsilon f(t_0)^2 > 0$.



Lemma: $v_1, \dots, v_d \in V$ mutually orthogonal of length 1 under any inner product

$\langle v_i, v_j \rangle = 0$ for $j \neq i$, $\langle v_i, v_i \rangle = 1 \forall i$ orthonormal \Rightarrow linearly independent.

Pf: $v = c_1 v_1 + \dots + c_d v_d = 0 \Rightarrow \langle v, v_i \rangle = 0 \forall i$
 \parallel
 $c_i = 0$. \square

Application:

Thm 6.4: Given $k+1$ points $(t_0, a_0), \dots, (t_k, a_k)$ in \mathbb{R}^2 with t_0, \dots, t_k distinct,

$\exists!$ $p \in \mathcal{P}_k$ whose graph passes through the points.
there exists unique

Pf: Construct orthonormal $p_0, \dots, p_k \in \mathcal{P}_k$ under inner product (b):

$p_i(t_i) = 1, p_i(t_j) = 0$ for $j \neq i$. Set $\Delta(t) = (t-t_0)(t-t_1)\dots(t-t_k)$

$$\Delta_i(t) = \frac{\Delta(t)}{(t-t_i)} \quad (\text{omit the } t-t_i \text{ factor}).$$

Then $\Delta_i(t_j) = 0$ for $j \neq i$ and $\Delta_i(t_i) \neq 0$ since t_0, \dots, t_k distinct, so

$p_i(t) = \frac{\Delta_i(t)}{\Delta_i(t_i)}$ has p_0, \dots, p_k orthonormal. Lemma \Rightarrow basis

\Rightarrow every $f \in \mathcal{P}_k$ has unique expression $f = c_0 p_0 + \dots + c_k p_k$.

Note that $f(t_i) = 0 + \dots + 0 + c_i \cdot 1 + 0 + \dots + 0 = c_i$.

Take $p = a_0 p_0 + \dots + a_k p_k$. \square

Pf:

Cor: t_0, \dots, t_k distinct \Rightarrow
$$\begin{bmatrix} 1 & t_0 & t_0^2 & \dots & t_0^k \\ 1 & t_1 & t_1^2 & \dots & t_1^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_k & t_k^2 & \dots & t_k^k \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}$$
 is nonsingular.

evaluations of p at t_0, \dots, t_k
coeffs on the polynomial $p(t)$

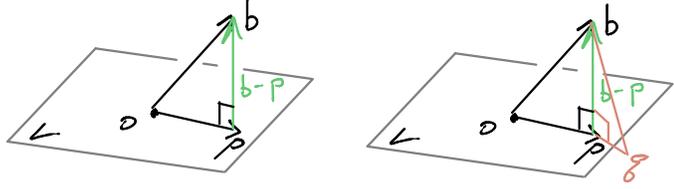
Thm 6.4 \Rightarrow linear system has ! sol.

\Rightarrow (square!) matrix is nonsingular. \square

§4.1

need IR here: using dot product and length

Def: Fix a subspace $V \subseteq \mathbb{R}^m$ and $b \in \mathbb{R}^m$. The (orthogonal) projection of b onto V is the unique vector $p = \text{proj}_V b \in V$ such that $b-p \in V^\perp$ under \cdot .



Lemma 1.1: Set $p = \text{proj}_V b$. Then $\|b-p\| \leq \|b-q\| \forall q \in V$.

p is closest to b in V.

Pf: $\|b-q\|^2 = \|b-p\|^2 + \|p-q\|^2$. \square

How to find p? Fix basis v_1, \dots, v_n for V .

Need $v_i \cdot (b-p) = 0, \dots, v_n \cdot (b-p) = 0$
 $v_i^T (b-p) = 0, \dots, v_n^T (b-p) = 0$

$\Leftrightarrow A^T (b-p) = 0$
 $\Leftrightarrow A^T b = A^T p$

$A = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix}$, so $\begin{bmatrix} -v_1^T \\ \vdots \\ -v_n^T \end{bmatrix} \begin{bmatrix} | \\ b-p \\ | \end{bmatrix} = 0$.

$\Leftrightarrow p \in b + V^\perp$

Q. Is that enough? No: Need also $p \in V$ — i.e. $p = Ax$ "p is a linear combination of the columns of A"

Prop: Given an $m \times n$ matrix of rank n , the normal equation $A^T A x = A^T b$ has a unique solution $\bar{x} \in \mathbb{R}^n$, the least squares solution of $Ax = b$.

Pf: $A^T A$ is $n \times n$.

Lemma: $M_{A^T A}$ is injective.

Lemma $\Rightarrow A^T A$ is nonsingular. \square

Pf of Lemma: $C(A)^\perp = N(A)$

$\Rightarrow C(A) \perp \underbrace{L(A)^T}_{= N(A^T)}$ under dot product in $\mathbb{R}^m_{\text{col}}$

$\Rightarrow C(A) \cap N(A^T) = 0$.

$Ax \in C(A)$

Note: \nleftrightarrow

$A^T A x = 0 \Leftrightarrow A^T (Ax) = 0$

$x \in \ker(M_{A^T A}) \Leftrightarrow \mu_A(x) \in \ker(\mu_{A^T})$

$\Leftrightarrow Ax \in \underbrace{N(A^T) \cap C(A)}_0$

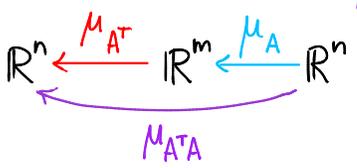
$\Leftrightarrow \mu_A(x) = 0$

$\Leftrightarrow Ax = 0$

since $\text{im}(\mu_A) \cap \ker(\mu_{A^T}) = 0$

$\Leftrightarrow x = 0. \square$

$\Leftrightarrow x = 0. \square$



calculate nullspace (kernel)

17. Prop 1.2: $\text{proj}_V b = \underbrace{A(A^T A)^{-1} A^T}_{P_V \text{ projection matrix}} b$ where $A = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix}$ for a basis v_1, \dots, v_n of V .

Pf: $p = \text{proj}_V b \Rightarrow p = A\bar{x}$, where $A^T A \bar{x} = A^T b$ uniquely by previous Prop. \bar{x} = coeffs. on v_1, \dots, v_n needed to express $p \in V$

$$\Rightarrow A\bar{x} = A(A^T A)^{-1} A^T b. \square$$

E.g. Find a point $p = Ax$ closest to $b = \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix}$ for $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}$.
How? It's $A(A^T A)^{-1} A^T b = \text{proj}_{C(A)} b$.

$$A^T A = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \Rightarrow (A^T A)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix}$$

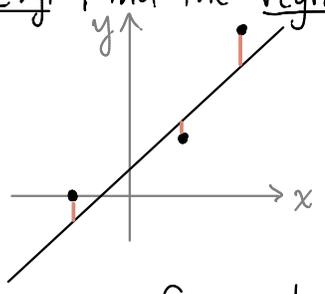
don't need to compute; why? $(A^T A)^T = A^T A^T$ symmetric

$$A^T b = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$(A^T A)^{-1} A^T b = \frac{1}{10} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 3 \\ 11 \end{bmatrix}$$

$$A \left(\frac{1}{10} \begin{bmatrix} 3 \\ 11 \end{bmatrix} \right) = \frac{1}{10} \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 11 \end{bmatrix} = \frac{1}{10} \left(\begin{bmatrix} 6 \\ 3 \\ 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 11 \\ 11 \\ 11 \\ -11 \end{bmatrix} \right) = \frac{1}{10} \begin{bmatrix} 17 \\ 14 \\ 11 \\ -8 \end{bmatrix}$$

E.g. Find the regression line (least squares line) for the (data) points $\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.



Want (but can't have!) $y = ax + b$ with

$$\begin{aligned} 0 &= -1a + b \\ 1 &= 1a + b \\ 3 &= 2a + b \end{aligned} \Leftrightarrow \underbrace{\begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}}_A \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

Can get $y = \bar{a}x + \bar{b}$ so that $\bar{y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \end{bmatrix} = A \begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}$ as close to $\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ as possible.

$\bar{y} = A \begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}$ is the least squares solution, minimizes $\| \epsilon \|$ for $\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_1 - 0 \\ \bar{y}_2 - 1 \\ \bar{y}_3 - 3 \end{bmatrix}$ error vector

looking for $\begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}$ but closeness condition is on the y 's.

we don't actually care about these!

no leading "A"
Solve: $\begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix} = (A^T A)^{-1} A^T \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$.

$$(A^T A)^{-1} = \left(\begin{bmatrix} -1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}^{-1} = \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix}$$

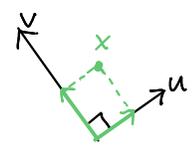
$$\begin{aligned} (A^T A)^{-1} A^T \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} &= \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} -1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \\ &= \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \end{bmatrix} \\ &= \frac{1}{14} \begin{bmatrix} 13 \\ 10 \end{bmatrix} \Rightarrow \text{regression line is } y = \frac{13}{14}x + \frac{5}{7} \quad (\text{check against reality}) \end{aligned}$$

E.g. $\dim V = 1 \Rightarrow A = a \in \mathbb{R}^m$

$$\Rightarrow P_V = a(a^T a)^{-1} a^T = a \frac{1}{\|a\|^2} a^T = \frac{1}{\|a\|^2} a a^T = \frac{a a^T}{a^T a}$$

from early HW

Recall $u \perp v$ under \cdot and $x \in \text{span}(u, v) \Rightarrow x = \text{proj}_u x + \text{proj}_v x$



Rewrite: $u \perp v$ and $V = \text{span}(u, v) \Rightarrow \text{proj}_V = \text{proj}_u + \text{proj}_v$

special case of product of

block matrices: $\begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix} \begin{bmatrix} \square \\ \square \end{bmatrix}$

$$\begin{aligned} &= \frac{1}{\|u\|^2} u u^T + \frac{1}{\|v\|^2} v v^T \\ &= \begin{bmatrix} | & | \\ u & v \\ \|u\| & \|v\| \end{bmatrix} \begin{bmatrix} -u^T / \|u\| & - \\ -v^T / \|v\| & - \end{bmatrix} \\ &= \text{blue block} + \text{red block} \end{aligned}$$

General: v_1, \dots, v_n orthogonal $\Rightarrow P_V =$

can be taken as def. of "orthogonal set"

$$\begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \frac{1}{\|v_1\|^2} & & \\ & \ddots & \\ & & \frac{1}{\|v_n\|^2} \end{bmatrix} \begin{bmatrix} -v_1^T & - \\ \vdots & \\ -v_n^T & - \end{bmatrix}$$

$A \quad (A^T A)^{-1} \quad A^T$

Q. $(A^T A)^{-1} = ? \begin{bmatrix} \frac{1}{\|v_1\|^2} & & \\ & \ddots & \\ & & \frac{1}{\|v_n\|^2} \end{bmatrix}^{-1}$

Algorithm 2.4 (Gram-Schmidt)

Input: basis v_1, \dots, v_n for inner product space V

Output: orthogonal basis of V

Initialize: $w_1 = v_1$

$W_1 = \text{span}(w_1)$
 $i = 1$

While: $i < n$

Do: $w_{i+1} = v_{i+1} - \text{proj}_{W_i} v_{i+1}$
 $W_{i+1} = \text{span}(w_1, \dots, w_{i+1})$ ($= \text{span}(v_1, \dots, v_{i+1})!$) but think in terms of the w_i
 $i \leftarrow i+1$

Return: w_1, \dots, w_n

To get orthonormal basis q_1, \dots, q_n set $q_i = \frac{w_i}{\|w_i\|}$
 Return: q_1, \dots, q_n

E.g. $v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 3 \\ 3 \\ 3 \end{bmatrix} \Rightarrow w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$w_2 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix} - \frac{(3, 1, -1, 1) \cdot (1, 1, 1, 1)}{\|(1, 1, 1, 1)\|^2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix} - \frac{4}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \end{bmatrix}$$

$$w_3 = v_3 - \text{proj}_{w_1} v_3 - \text{proj}_{w_2} v_3 = v_3 - \left(\text{proj}_{w_1} v_3 + \text{proj}_{w_2} v_3 \right)$$

$$= v_3 - \left(\frac{v_3 \cdot w_1}{\|w_1\|^2} w_1 + \frac{v_3 \cdot w_2}{\|w_2\|^2} w_2 \right)$$

orthonormal basis: columns of

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

A Q R

$$= \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix} - \frac{8}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-4}{8} \begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

By construction, each w_i (or q_i) lies in span of v_1, \dots, v_i . Thus

$$\begin{bmatrix} | & & | \\ q_1 & \dots & q_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} * & * & \dots & * \\ 0 & * & & * \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & & * \end{bmatrix} \quad U^{-1} = R \Rightarrow QR = A$$

Q A U upper-triangular

QR decomposition
 expresses A as product of
 (matrix with orthonormal cols)
 and (upper-triangular matrix)

E.g. $A = QR$

Questions: 1. $r_{ij} = ?$ coeff. of q_i in $v_j = q_i \cdot v_j$ since q_1, \dots, q_n orthonormal

2. What's P_V in terms of Q? $P_V = QQ^T = \text{proj}_{q_1} + \dots + \text{proj}_{q_n}!$

pf: $P_V = A(A^T A)^{-1} A^T = QR(R^T Q^T QR)^{-1} R^T Q^T$ Note: $R^{-1} = U$

$$= QR(R^T R)^{-1} R^T Q^T = QR R^{-1} (R^T)^{-1} R^T Q^T = QQ^T \square$$

18. Recall: for vector spaces V and W , the transformation map function operator $T: V \rightarrow W$ is linear

if $T(u+cv) = Tu + cTv \quad \forall u,v \in V$ and scalars c .

$$T\left(\sum_{i=1}^k c_i v_i\right) = \sum_{i=1}^k c_i T v_i$$

E.g. (i) $D: C^1(I) \rightarrow C^0(I) \quad D(f+cg) = (f+cg)' = f' + cg' = Df + cDg$
 $f \mapsto f'$

$D: \mathcal{P}_k \rightarrow \mathcal{P}_{k-1}$ or $D: \mathcal{P}_k \rightarrow \mathcal{P}_k$

rank?	k	k
image?	\mathcal{P}_{k-1}	\mathcal{P}_{k-1}
kernel?	{constants}	{constants}
injective?	no	no
surjective?	yes	no

nullity? 1: rank + nullity = $k+1$ in both cases

Why do these examples? Demonstrate that abstract vector space constructions have concrete interpretations in concrete (!) vector spaces.

(ii) $C^0(I) \xrightarrow{M_t} C^0(I)$ "multiplication by t "
 $f(t) \mapsto t f(t)$

injective? yes: $t f(t) = 0 \Rightarrow f \equiv 0$ on I
surjective? need g/t continuous $\forall g \in C^0(I) \Rightarrow \begin{cases} \text{yes if } 0 \notin I \\ \text{no if } 0 \in I \end{cases}$

(iii) $C^0([0,1]) \rightarrow C^0([0,1])$
 $f(t) \mapsto \int_0^t f(s) ds$ i.e. if $F' = f$ then $f \mapsto F - F(0)$

Fundamental Thm of Calculus

injective? Can \int (nonzero function) be the zero function? No: $f \neq 0 \Rightarrow F \neq 0$, so Yes injective.

surjective? No: image is $C^1 \subsetneq C^0$ Pf: $F' = f!$

(iv) $ev_{0,1,3}: C^0([0,4]) \rightarrow \mathbb{R}^3$ "evaluation"

$$f \mapsto \begin{bmatrix} f(0) \\ f(1) \\ f(3) \end{bmatrix} \quad f+cg \mapsto \begin{bmatrix} f(0) \\ f(1) \\ f(3) \end{bmatrix} + c \begin{bmatrix} g(0) \\ g(1) \\ g(3) \end{bmatrix}$$

Q. Is $ev_{0,1,3}: \mathcal{P}_4 \rightarrow \mathbb{R}^3$ injective? surjective? Find kernel and image.

A. $\dim: 5 \quad 3 \Rightarrow \text{rank} \leq 3$. rank-nullity thm $\Rightarrow \dim \ker \geq 2$.
 \Rightarrow not injective.

Thm 3.6.4 \Rightarrow surjective! image = $\mathbb{R}^3 \Rightarrow \text{rank} = 3$
 $\Rightarrow \dim \ker = 2$.

0, 1, 3 roots $\Rightarrow t(t-1)(t-3), t^2(t-1)(t-3) \in \ker$
independent because different degrees \Rightarrow basis for \ker .

(v) $\mathbb{R}^{2 \times 2} \xrightarrow{M \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix}} \mathbb{R}^{2 \times 2}$ any $B \in \mathbb{R}^{2 \times 2}$ would do

$A \mapsto \begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix} A$ $B(A+cA') = BA + cBA' \Rightarrow \mu_B$ is linear

Def: $T: V \rightarrow W$ is an isomorphism if T is bijective. *injective and surjective*

Lemma 4.1: T is an isomorphism $\Leftrightarrow \exists T^{-1}: W \rightarrow V$ with

$T \circ T^{-1} = id_W : w \mapsto w \quad \forall w \in W$
 $T^{-1} \circ T = id_V : v \mapsto v \quad \forall v \in V$

Pf: Exercise. (\Rightarrow : HW, #13) *content: bijective \Leftrightarrow has inverse as map of sets; need linearity.*

Prop: $T: V \rightarrow W$ isomorphism $\Leftrightarrow T(\text{basis for } V) = \text{basis for } W$.

Pf: Exercise. *(not assigned)*

Cor: $\dim V = n \Leftrightarrow V \cong \mathbb{R}^n$.

Def: Let $T: V \rightarrow W$ be linear. If

$\mathcal{V} = (v_1, \dots, v_n)$ is an ordered basis of V

$\mathcal{W} = (w_1, \dots, w_m)$ is an ordered basis of W

then $A = [T]_{\mathcal{V}, \mathcal{W}}$ is the matrix of T with respect to \mathcal{V} and \mathcal{W} if the j -th column of A lists the coefficients on w_1, \dots, w_m in Tv_j .

$Tv_1 = a_{11}w_1 + \dots + a_{m1}w_m = [w_1 \dots w_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}$ *1x1 symbol*

$Tv_n = a_{1n}w_1 + \dots + a_{mn}w_m = [w_1 \dots w_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$

a list of symbols

$[Tv_1 \dots Tv_n] = [w_1 \dots w_m] \begin{bmatrix} | & & | \\ a_1 & \dots & a_n \\ | & & | \end{bmatrix}$

Key point: Tv_j is a linear combination of the w 's; what are the coefficients? Listed in a_j .

E.g. $D: \mathcal{P}_3 \rightarrow \mathcal{P}_2 \quad \mathcal{V} = (1, t, t^2, t^3)$ and $\mathcal{W} = (1, t, t^2) \Rightarrow [D]_{\mathcal{V}, \mathcal{W}} = ?$

$[D1 \ Dt \ Dt^2 \ Dt^3] = [0 \ 1 \ 2t \ 3t^2] = [1 \ t \ t^2] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

19. $T: V \rightarrow W$ linear $\mathcal{V} = (v_1, \dots, v_n)$ ordered basis of V
 $\mathcal{W} = (w_1, \dots, w_m)$ ordered basis of W

compact notation
not in textbook!

$$T[v_1 \dots v_n] = [w_1 \dots w_m] [T]_{\mathcal{V}, \mathcal{W}}$$

defines $[T]_{\mathcal{V}, \mathcal{W}} = A$

1x1 symbols

$$T v_j = a_{1j} w_1 + \dots + a_{mj} w_m = [w_1 \dots w_m] \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Q. Set $A = [T]_{\mathcal{V}, \mathcal{W}} \in \mathbb{R}^{m \times n}$. What does it mean that $Ax = b$?

A. $T[v_1 \dots v_n] x = [w_1 \dots w_m] Ax$ (*) The coefficients x of $v \in V$ on v_1, \dots, v_n get taken to the coefficients b of $w \in W$ on w_1, \dots, w_m

$$T(x_1 v_1 + \dots + x_n v_n) = [w_1 \dots w_m] b = \underbrace{b_1 w_1 + \dots + b_m w_m}_w$$

E.g. $D: \mathcal{P}_3 \rightarrow \mathcal{P}_2$ $\mathcal{V} = (1, t, t^2, t^3)$ and $\mathcal{W} = (1, t, t^2) \Rightarrow [D]_{\mathcal{V}, \mathcal{W}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

$$v = x_0 + x_1 t + x_2 t^2 + x_3 t^3 \Rightarrow Dv = x_1 + 2x_2 t + 3x_3 t^2$$

$$\leftrightarrow \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\leftrightarrow \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = [D]_{\mathcal{V}, \mathcal{W}} x = \begin{bmatrix} x_1 \\ 2x_2 \\ 3x_3 \end{bmatrix}$$

E.g. $V = \mathbb{R}^n$ $W = \mathbb{R}^m$
 $\mathcal{V} = \mathcal{E}_n$ $\mathcal{W} = \mathcal{E}_m$
 $= e_1, \dots, e_n$ $= e_1, \dots, e_m$

$$T = M_A \quad m \begin{array}{|c|} \hline n \\ \hline A \\ \hline \end{array} \Rightarrow [T]_{\mathcal{E}_n, \mathcal{E}_m} = ?$$

$$M_A [e_1 \dots e_n] = [e_1 \dots e_m] [T]_{\mathcal{E}_n, \mathcal{E}_m}$$
$$\Rightarrow [T]_{\mathcal{E}_n, \mathcal{E}_m} = A$$

Why? $M_A e_j = A \begin{bmatrix} | \\ e_j \\ | \end{bmatrix} = a_j = [e_1 \dots e_m] \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$

$Ax = b \Rightarrow$ coefficients of x on e_1, \dots, e_n
 \mapsto coefficients of b on e_1, \dots, e_m

j^{th} column of $[T]_{\mathcal{V}, \mathcal{W}}$ lists the coefficients on w_1, \dots, w_m in Tv_j

Lemma: Fix V with basis v_1, \dots, v_n . If v'_1, \dots, v'_n is another basis of V then $[v'_1 \dots v'_n] = [v_1 \dots v_n] P$ for an invertible $P \in \mathbb{R}^{n \times n}$. Note: $P = [id_V]_{\mathcal{V}', \mathcal{V}}$

Pf: $P = [id_V]_{\mathcal{V}', \mathcal{V}}$ by def; need P invertible. Enough: $N(P) = \{0\}$. But $x \in N(P) \Rightarrow [v'_1 \dots v'_n] x = [v_1 \dots v_n] P x = 0 \Rightarrow x = 0$ because v_1, \dots, v_n independent. \square

E.g. $V = \mathcal{P}_3$ $\mathcal{V} = 1, t, t^2, t^3$ $\mathcal{V}' = 1, t-1, t^2-t, t^3-t^2 \Rightarrow P = ?$

$$[1 \ t-1 \ t^2-t \ t^3-t^2] = [1 \ t \ t^2 \ t^3] \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \leftarrow P. \text{ change-of-basis matrix}$$

Theorem 4.2 (change-of-basis formula): Fix a linear map $T: V \rightarrow W$,

- ordered bases $\mathcal{V}, \mathcal{V}'$ for V with $[v'_1 \dots v'_n] = [v_1 \dots v_n] P$
- $\mathcal{W}, \mathcal{W}'$ for W $[w'_1 \dots w'_m] = [w_1 \dots w_m] Q$,
- matrices $A = [T]_{\mathcal{V}, \mathcal{W}}$ and $A' = [T]_{\mathcal{V}', \mathcal{W}'}$. Then $A' = Q^{-1}AP$.

Pf: $T[v_1 \dots v_n] = [w_1 \dots w_m] A \Rightarrow T[v_1 \dots v_n] P = [w_1 \dots w_m] AP$
 $T[v'_1 \dots v'_n] = [w'_1 \dots w'_m] Q^{-1}AP$ defines A' . \square

E.g. $T: \mathcal{P}_3 \rightarrow \mathbb{R}^2$ $T = \text{ev}_{0,1}$ $\mathcal{W} = \mathcal{E}_2 = (e_1, e_2) = \mathcal{W}'$

$$f \mapsto \begin{bmatrix} f(0) \\ f(1) \end{bmatrix}$$

Calculate $A = [T]_{\mathcal{V}, \mathcal{W}}$ and $A' = [T]_{\mathcal{V}', \mathcal{W}'}$ and verify the relation between them.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \text{ evaluate basis vectors at 0 and 1}$$

$$A' = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} v'_3 \text{ and } v'_4 \in \ker T \\ -1e_1 + 0e_2 \\ 1e_1 + 1e_2 \end{matrix}$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$$Q^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = A'$$

aside: $\ker T$ more visible in A' : $\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

v'_3 and $v'_4 \mapsto \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow v'_3, v'_4$ form a basis of $\ker T$

image T : Tv'_1 and Tv'_2 span image $T \Rightarrow$ they form a basis for $\text{im } T$

$$\text{rank-nullity: } 2 + 2 = 4$$

20.

$$T [v_1 \dots v_n] = [w_1 \dots w_m] [T]_{\mathcal{W}, \mathcal{V}}$$

$$[v'_1 \dots v'_n] = [v_1 \dots v_n] P$$

$$[w'_1 \dots w'_m] = [w_1 \dots w_m] Q$$

$$A = [T]_{\mathcal{W}, \mathcal{V}}$$

$$A' = [T]_{\mathcal{W}', \mathcal{V}'} \Rightarrow A' = Q^{-1} A P$$

$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ lists coefficients of $v \in V$ on basis v'_1, \dots, v'_n

$$[v'_1 \dots v'_n] x = [v_1 \dots v_n] P x$$

$\Rightarrow P x$ lists coefficients of $v \in V$ on basis v_1, \dots, v_n

Def: For a basis \mathcal{B} of V and $T: V \rightarrow V$ linear, set $[T]_{\mathcal{B}} = [T]_{\mathcal{B}, \mathcal{B}}$. Reiterate (*)

E.g. $V = \mathbb{R}^n \Rightarrow [M]_{\mathcal{E}_n} = A$.

Cor: Bases $\mathcal{B}, \mathcal{B}'$ for V with $[v'_1 \dots v'_n] = [v_1 \dots v_n] P \Rightarrow [T]_{\mathcal{B}'} = \underbrace{P^{-1} [T]_{\mathcal{B}} P}_{\text{conjugate of } [T]_{\mathcal{B}} \text{ by } P}$.

In particular,

$$P = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} \Rightarrow [T]_{\mathcal{B}} = P^{-1} [T]_{\mathcal{E}_n} P.$$

$[T]_{\mathcal{B}'}$ and $[T]_{\mathcal{B}}$ are similar

Pf: $V=W, Q=P$. \square

E.g. Fix orthonormal v_1, v_2, v_3 in \mathbb{R}^3 . Describe the linear map $L: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that

1. sends v_1, v_2, v_3 to e_1, e_2, e_3
2. rotates by $\pi/3$ around z -axis
3. sends e_1, e_2, e_3 to v_1, v_2, v_3 .

Answer: $[L]_{\mathcal{E}} = \underset{3}{P} \underset{2}{R} \underset{1}{P}^{-1} \Rightarrow P^{-1} [L]_{\mathcal{E}} P = R \Rightarrow R = [L]_{\mathcal{B}} \Rightarrow L$ rotates around v_3 by $\pi/3$!

where $P = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}$ and $R = \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$v_1 \mapsto \begin{bmatrix} -\frac{1}{2} v_1 \\ + \frac{\sqrt{3}}{2} v_2 \\ + 0 v_3 \end{bmatrix} \quad v_2 \mapsto \begin{bmatrix} -\frac{\sqrt{3}}{2} v_1 \\ + \frac{1}{2} v_2 \\ + 0 v_3 \end{bmatrix} \quad v_3 \mapsto \begin{bmatrix} 0 v_1 \\ + 0 v_2 \\ + 1 v_3 \end{bmatrix}$$

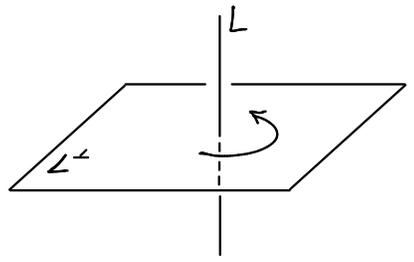
$$= [v_1 \ v_2 \ v_3] \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix} = [v_1 \ v_2 \ v_3] \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} = [v_1 \ v_2 \ v_3] \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Crucial note $P = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}$ has two tellingly different interpretations:

- multiplication by P is a linear map $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ that takes e_1, e_2, e_3 to v_1, v_2, v_3
 - takes the coefficients $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ of a fixed $v \in \mathbb{R}^3$ on v_1, v_2, v_3 to the coefficients $P x$ of the same $v \in \mathbb{R}^3$ on e_1, e_2, e_3
- $$[v_1 \ v_2 \ v_3] x = [e_1 \ e_2 \ e_3] \underbrace{P x}_{= v!} \leftarrow$$

Def: Fix subspace $L \subseteq \mathbb{R}^n$ with $\dim L = n-2$. The rotation by angle α around L is the linear map rot_α^L determined by

- $\text{rot}_\alpha^L(L) = L$
- (*) • $\text{rot}_\alpha^L(L^\perp)$ is usual rotation of \mathbb{R}^2 by α .



Q. Why set $\dim L = n-2$?

A. $\dim L^\perp = \cancel{?} 2$

Q. $L \cap L^\perp = \cancel{?} 0$ What kind of "0" is this?

Q. How is rot_α^L "determined by" (*)?

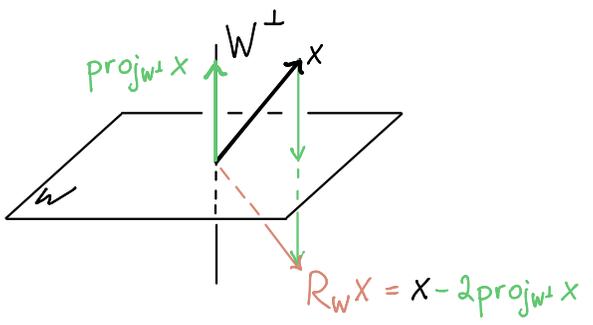
A. Choose • basis v_3, \dots, v_n for L
 • orthonormal basis x, y for L^\perp .

$B = x, y, v_3, \dots, v_n \Rightarrow [\text{rot}_\alpha^L]_B =$

$\cos \alpha$	$-\sin \alpha$	0			
$\sin \alpha$	$\cos \alpha$	0			
0	0	1			
0	0	0	\ddots	0	
0	0	0	0	1	
0	0	0	0	0	1

rotation of x into y by α
 no mixing of L with L^\perp
 $v_3 \mapsto v_3, \dots, v_n \mapsto v_n$

Def: Fix subspace $W \subseteq V$. The reflection across W is $R_W = \text{id}_V - 2 \text{proj}_{W^\perp}$



E.g. $V = \mathbb{R}^3, w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, w_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

$W = \text{span}(w_1, w_2)$. Find $[R_W]_{\mathcal{E}}$.

Need proj_{W^\perp} . $\dim W = 2$ (proof?) $\Rightarrow \dim W^\perp = 1$. $v_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \in W^\perp$
 $\Rightarrow W^\perp = \text{span}(v_3)$.

2proj_{W^\perp} has matrix $2 \frac{v_3 v_3^T}{v_3^T v_3} = 2 \frac{1}{6} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 1 \end{bmatrix}$
 $= 2 \frac{1}{6} \begin{bmatrix} 1 & -2 & -1 \\ -2 & 4 & 2 \\ -1 & 2 & 1 \end{bmatrix}$

$\Rightarrow \text{id}_{\mathbb{R}^3} - 2 \text{proj}_{W^\perp}$ has matrix $\frac{1}{3} \begin{bmatrix} 3 & & \\ & 3 & \\ & & 3 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & -2 & -1 \\ -2 & 4 & 2 \\ -1 & 2 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{bmatrix} = [R_W]_{\mathcal{E}}$.

E.g. $V = \mathbb{R}^3$, $\mathcal{B} = (w_1, w_2, v_3)$. Find $[R_W]_{\mathcal{B}}$

Go back to def: $R_W[w_1, w_2, v_3] = [w_1, w_2, v_3][R_W]_{\mathcal{B}}$ *W is fixed*

But $[R_W w_1, R_W w_2, R_W v_3] = [w_1, w_2, -v_3]$ *$W^\perp \rightarrow -W^\perp$*

$$[w_1 \ w_2 \ v_3] \underbrace{\begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix}}_{[R_W]_{\mathcal{B}}}$$

E.g. Use $[R_W]_{\mathcal{B}}$ to compute $[R_W]_{\mathcal{E}}$.

$$P = \begin{bmatrix} 1 & 1 & 1 \\ w_1 & w_2 & v_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \Rightarrow [R_W]_{\mathcal{B}} = P^{-1}[R_W]_{\mathcal{E}}P$$

$$\Rightarrow P[R_W]_{\mathcal{B}}P^{-1} = [R_W]_{\mathcal{E}}$$

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

columns of P are orthogonal!

$$= \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix} \begin{bmatrix} 1/2 & & \\ & 1/3 & \\ & & 1/6 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ -1 & 2 & 1 \end{bmatrix}}_{P^T}$$

$$= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -2 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & -1/3 \\ -1/6 & 1/3 & 1/6 \end{bmatrix}$$

$$= \begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 2/3 & -1/3 & 2/3 \\ -2/3 & 2/3 & 2/3 \end{bmatrix} \quad \checkmark$$

Summary: $T[v_1 \dots v_n] = [v_1 \dots v_n][T]_{\mathcal{B}}$ (def) *entries v_j are 1×1 symbols*

abstract $[T]_{\mathcal{E}}P = P[T]_{\mathcal{B}}$

columns $\begin{bmatrix} v_j \\ 1 \end{bmatrix}$ are $n \times 1$

in coordinates $[T]_{\mathcal{E}} \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} [T]_{\mathcal{B}}$

$$\Leftrightarrow [T]_{\mathcal{B}} = P^{-1}[T]_{\mathcal{E}}P$$

$$\Leftrightarrow P[T]_{\mathcal{B}}P^{-1} = [T]_{\mathcal{E}}$$

21. Chapter 5: Determinants $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
 can be any F

Thm: $\exists!$ function $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ satisfying

- alternating 1. $\det A = 0$ if A has two equal adjacent rows $\rightarrow \det \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} = 0$
- multilinear 2. $\det A' = c \det A$ if A' is obtained by multiplying a row of A by c . $\det \begin{bmatrix} \text{---} \\ cR \\ \text{---} \\ \text{---} \end{bmatrix} = c \det \begin{bmatrix} \text{---} \\ R \\ \text{---} \\ \text{---} \end{bmatrix}$
3. $\det A = \det A' + \det A''$ if A, A', A'' agree in all rows except row i , where $A_i = A'_i + A''_i$. $\det \begin{bmatrix} \text{---} \\ + \\ \text{---} \\ \text{---} \end{bmatrix} = \det \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} + \det \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix}$
4. $\det I_n = 1$.

We've seen bilinear: $\langle \cdot, \cdot \rangle$
 linear in each variable

Def: $\det A$ is the determinant of A .

Pf: \exists uses cofactors - next class.

!: assume $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ is some function satisfying #1 - #4.

Lemma: $\det A' = -\det A$ if A' has two rows swapped from A .

Pf: $0 \stackrel{\#1}{=} \det \begin{bmatrix} \vdots \\ -x+y \\ -x+y \\ \vdots \end{bmatrix} \stackrel{\#3}{=} \det \begin{bmatrix} \vdots \\ -x \\ -x+y \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ -y \\ -x+y \\ \vdots \end{bmatrix}$

$\stackrel{\#3}{=} \det \begin{bmatrix} \vdots \\ -x \\ -x \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ -x \\ -y \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ -y \\ -x \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ -y \\ -y \\ \vdots \end{bmatrix}$

so done if swapped rows adjacent. If not adjacent, then

$i \begin{bmatrix} \vdots \\ -A_i \\ \vdots \\ -A_j \\ \vdots \end{bmatrix} \xrightarrow{j-i \text{ steps}} i+1 \begin{bmatrix} \vdots \\ -A_i \\ \vdots \\ -A_j \\ \vdots \end{bmatrix} \xrightarrow{1 \text{ step}} i+1 \begin{bmatrix} \vdots \\ -A_j \\ \vdots \\ -A_i \\ \vdots \end{bmatrix} \xrightarrow{j-i \text{ steps}} j \begin{bmatrix} \vdots \\ -A_j \\ \vdots \\ -A_i \\ \vdots \end{bmatrix}$ and $(-1)^{2(j-i)+1} = -1. \square$

Cor: $\det A = 0$ if any two rows are equal.

Pf: Swap 'til you drop. \square

Prop 1.4: $\det(EA) = \det E \det A$ if E is elementary.

Pf: E swaps rows: $\det(EA) = -\det A = \det E \det A$
Lemma #4 + Lemma

E multiplies row by scalar c : $\det(EA) = c \det A = \det E \det A$
#2 #2 + #4

E replaces A_i with $A_i + cA_j$:
 $\det \begin{bmatrix} \vdots \\ A_i + cA_j \\ \vdots \\ -A_j \\ \vdots \end{bmatrix} \stackrel{\#3}{=} \det \begin{bmatrix} \vdots \\ -A_i \\ \vdots \\ -A_j \\ \vdots \end{bmatrix} + \det \begin{bmatrix} \vdots \\ cA_j \\ \vdots \\ -A_j \\ \vdots \end{bmatrix}$
 $\stackrel{\#2}{=} \det A + c \det \begin{bmatrix} \vdots \\ -A_j \\ \vdots \\ -A_j \\ \vdots \end{bmatrix}$ *0 by Cor.*

$A = I_n \Rightarrow \det E = 1 (!) \Rightarrow \det E \det A = \det A$. \square

Thm 1.2: $A \in \mathbb{R}^{n \times n}$ singular $\Leftrightarrow \det A = 0$.

If one row is a linear combination of the others, expanding by multilinearity yields 0 in every summand by alternation (#1).

Pf: Write $U = E_k \cdots E_1 A$ reduced echelon form.

$\det U = \det E_k \det(E_{k-1} \cdots E_1 A)$ by Prop 1.4.
 $= \det E_k \det E_{k-1} \det(E_{k-2} \cdots E_1 A)$
 \vdots
 $= \det E_k \det E_{k-1} \det E_{k-2} \cdots \det E_1 \det A$.

any function $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ satisfying #1 - #4

A singular $\Rightarrow U$ has zero-row U_n (bottom row)

$\Rightarrow \det U = 0 \Rightarrow \det A = 0$ by #2: $U_n = 0 U_n$. But

$\det E = \begin{cases} -1 & \text{if } E \text{ has type (i)} \\ c & \text{(ii)} \\ 1 & \text{(iii)} \\ \neq 0 & \end{cases} \Rightarrow \det A = 0$

A nonsingular $\stackrel{\#4}{\Rightarrow} 1 = \det U = \det E_k \det E_{k-1} \det E_{k-2} \cdots \det E_1 \det A$
 $\Rightarrow \det A \neq 0$. \square

To finish proof of !, $\det A = 0$ if A singular

At most one function can do this. Given that one exists, there only be one.

$\det A = \frac{\det I_n}{\det E_k \cdots \det E_1}$ if A nonsingular. \square

E.g. $\det \begin{bmatrix} 2 & 4 & 6 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \stackrel{(i)}{=} -\det \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 2 & 4 & 6 \end{bmatrix} \stackrel{(iii)}{=} -\det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 4 & 4 \end{bmatrix} \stackrel{(iii)}{=} -\det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 12 \end{bmatrix} \stackrel{(ii)}{=} -12 \det \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \stackrel{(iii)}{=} -12 \det I_3 = -12$

Consequences (of $\exists!$, etc.)

Thm 1.5: $A, B \in \mathbb{R}^{n \times n} \Rightarrow \det(AB) = \det A \det B$.

Pf: A singular $\Rightarrow \det A = 0 \Rightarrow \det A \det B = 0$.

\downarrow
 $L(A) \neq 0 \Rightarrow L(AB) \neq 0$ since $L(AB) \geq L(A)$: $\boxed{A} = 0 \Rightarrow \boxed{A} \boxed{B} = 0$
 $\Rightarrow AB$ singular $\Rightarrow \det(AB) = 0$. \checkmark

A nonsingular $\Rightarrow \det(AB) = \det(E_1' \dots E_k' B)$
 $= \det E_1' \dots \det E_k' \det B$
 $= \det A \det B$. \square

Cor 1.6: $\det(A^{-1}) = \frac{1}{\det A}$.

Pf: $\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det I = 1$. \square

Cor: A similar to $A' \Rightarrow \det A = \det A'$.

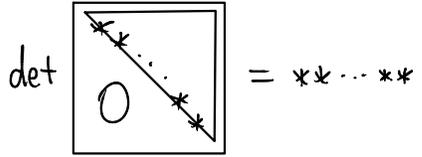
Note: not $\Leftrightarrow!$ e.g. $A \sim I \Rightarrow A = ?$
but $\det E = 1$ for type (iii)

Pf: $\det(PAP^{-1}) = \det P \det A \det P^{-1}$
 $= \det A$. \square

Prop 1.7: $\det A^T = \det A$.

Pf: Check for elementary matrices when A is nonsingular: E has same type as E^T .
Both = 0 if A is singular. \square

Prop 1.3: A upper-triangular $\Rightarrow \det A = a_{11} \dots a_{nn} =$ product of main diagonal entries.
or lower -



Pf: A nonsingular $\Rightarrow A \rightsquigarrow I_n$ by pulling out factors a_{11}, \dots, a_{nn} and then type (iii) operations, which have $\det 1$.

A singular $\Leftrightarrow < n$ pivots $\Leftrightarrow a_{ii} = 0$ for some i . \square

22.

Today: \exists det satisfying #1-#4

Def: $\det: \mathbb{R}^{1 \times 1} \rightarrow \mathbb{R}$
 $[a] \mapsto a$



$\det: \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mapsto a_{11} \det A_{11} - a_{21} \det A_{21} \\ = a_{11} \det [a_{22}] - a_{21} \det [a_{12}] \\ = a_{11} a_{22} - a_{21} a_{12}$$

\vdots
 $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$

recursive definition: def for n in terms of def for n-1

$$A \mapsto a_{11} \det A_{11} - a_{21} \det A_{21} + \dots + (-1)^{n+1} a_{n1} \det A_{n1} = a_{11} C_{11} + a_{21} C_{21} + \dots + a_{n1} C_{n1}$$

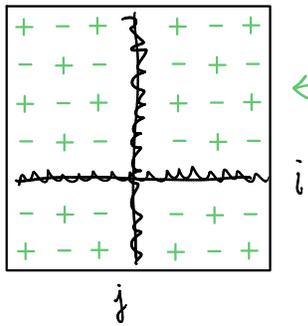
E.g. $\det \begin{bmatrix} 2 & 1 & 3 \\ 1 & -2 & 3 \\ 0 & a & 1 \end{bmatrix} = 2 \det \begin{bmatrix} -2 & 3 \\ a & 1 \end{bmatrix} - 1 \det \begin{bmatrix} 1 & 3 \\ a & 1 \end{bmatrix} + 0 \det \begin{bmatrix} 1 & 3 \\ -2 & 3 \end{bmatrix}$

$$= 2(-8) - 1(-5) + 0 \\ = -11.$$

- 1. Don't use to compute $\geq 3 \times 3$
- 2. Use for $\det(\nabla) = \prod \text{diag}$

Def: For $A \in \mathbb{R}^{n \times n}$ with $n \geq 2$, get $(n-1) \times (n-1)$ matrix A_{ij} by deleting row i and column j .
The ij^{th} cofactor of A is $C_{ij} = (-1)^{i+j} \det A_{ij}$.

$$= a_{11} C_{11} + a_{21} C_{21} + \dots + a_{n1} C_{n1}$$

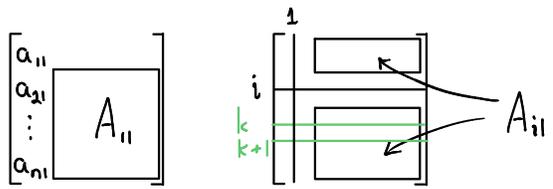


E.g. $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -2 & 3 \\ 0 & a & 1 \end{bmatrix} \Rightarrow A_{13} = \begin{bmatrix} 1 & -2 \\ 0 & a \end{bmatrix} \quad A_{22} = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \quad A_{32} = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix}$
 $C_{13} = +2 \quad C_{22} = +2 \quad C_{32} = -(2 \cdot 3 - 3 \cdot 1) = -3$

Thm: $\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ satisfies #1-#4.

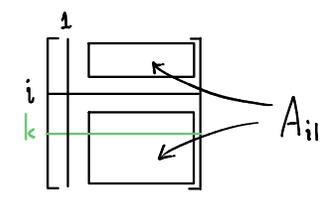
Pf: $n=2$: do it yourself.

Assume $n \geq 3$ and prove by induction on n .

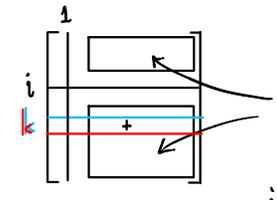


#1: rows k and $k+1$ equal
 $\Rightarrow \det A_{ii} = 0$ for $i \neq k, k+1$, and

$a_{k+1} \det A_{k+1} = a_{(k+1)1} \det A_{(k+1)1}$ but $-(-1)^{k+1} = (-1)^{k+1+1}$
 so these terms cancel.



#2: $A \xrightarrow{A_k \rightsquigarrow cA_k} A' \xrightarrow{\text{induction}} \det A'_{ii} = c \det A_{ii}$ but $a'_{ii} = a_{ii}$ if $i \neq k$
 $\det A'_{k1} = \det A_{k1}$ but $a'_{ii} = ca_{ii}$, so
 $a'_{ii} \det A'_{ii} = ca_{ii} \det A_{ii}$ either way.



#3: A, A', A'' agree in all rows $\neq k$, where $A_k = A'_k + A''_k$
 $i \neq k: \det A_{ii} = \det A'_{ii} + \det A''_{ii} \Rightarrow a_{ii} \det A_{ii} = a_{ii} (\det A'_{ii} + \det A''_{ii})$
 $a_{ii} = a'_{ii} + a''_{ii} \Rightarrow = a_{ii} \det A'_{ii} + a_{ii} \det A''_{ii}$
 $= a'_{ii} \det A'_{ii} + a''_{ii} \det A''_{ii}$
 $i = k: \det A_{k1} = \det A'_{k1} + \det A''_{k1} \Rightarrow a_{k1} \det A_{k1} = (a'_{k1} + a''_{k1}) \det A_{k1}$
 $a_{k1} = a'_{k1} + a''_{k1} \Rightarrow = a'_{k1} \det A_{k1} + a''_{k1} \det A_{k1}$
 $= a'_{k1} \det A'_{k1} + a''_{k1} \det A''_{k1}$

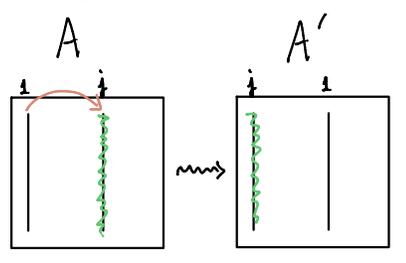
$\Rightarrow \det A = \det A' + \det A''$ term by term.

#4: $\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} A_{ii} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} I_{n-1} \Rightarrow \det I_n = 1 \det I_{n-1} - 0 + 0 - 0 + \dots \pm 0$
 $= 1(1)$ by induction
 $= 1. \square \Rightarrow \exists! \det$

Prop 2.2: $\det A = a_{1j} C_{1j} + \dots + a_{nj} C_{nj}$ for any fixed j . expand along any column

Pf: Swap columns 1 and j of A to get A' . Then $\det A' = -\det A$.

But $a'_{i1} = a_{ij} \forall i$, and $A_{ij} \rightsquigarrow A'_{i1}$ by moving leftmost column across $j-2$ columns to column $j-1$. Hence



$C'_{i1} = (-1)^{i+1} \det A'_{i1} = (-1)^{i+1} (-1)^{j-2} \det A_{ij}$
 $= -(-1)^{i+j} \det A_{ij} = -C_{ij} \forall i$, so $a'_{i1} C'_{i1} = -a_{ij} C_{ij}. \square$

Cor 2.1: $\det A = a_{i1} C_{i1} + \dots + a_{in} C_{in}$ for any fixed i . expand along any row

The rules for swapping columns are the same as those for rows.

Pf: $\det A = \det A^T + \text{Prop 2.2. } \det A' = \det (A')^T = -\det A^T = -\det A. \square$

§5.3 Geometric interpretation

det(→) = area(▭) 2D area = 2D volume

over R

in R^n: det(→) = volume of parallelepiped. Why?

The axioms we use to define determinants are the same as those we use to define volume.

- 1. v1, ..., vn dependent => dim(span) < n => flat => vol = 0
2. scale edge by c => vol -> c vol
3. vol(≡) = vol(≡) Cavalieri's Principle
4. vol(unit hypercube) = 1.
} volume is multilinear!

Thm 2.3 (Cramer's rule): Ax = b with A nonsingular => xi = det Bi / det A, where A -> Bi.

Pf: b = Ax = x1a1 + ... + xn an => det Bi = det [a1 ... a_{i-1} x1a1 + ... + xn an a_{i+1} ... an] = det [a1 ... a_{i-1} xi ai a_{i+1} ... an] = xi det A. □

E.g. [2 3; 4 7] [x1; x2] = [3; -1] => B1 = [3 3; -1 7] B2 = [2 3; 4 -1]
x1 = (21+3)/(14-12) = 12
x2 = (-2-12)/(14-12) = -7 => x = [12; -7]

Magic! Look up "exterior algebra".

Thm 2.3: C = [Cij] = cofactor matrix of A => AC^T = (det A) In.

(i.e. A nonsingular => A^-1 = 1/det A C^T)

Pf: The diagonal entries of AC^T are precisely the sums in Cor 2.1.

Define a matrix Dij by copying row i of A into row j, so A -> Dij.

The ij entry of AC^T is det Dij as expanded along row j.

But det Dij = det A if i=j (Dij = A)
0 if i != j (Dij has Ai repeated). □

23. §6.1 Def: Fix $T: V \rightarrow V$. $v \in V$ is an eigenvector if $v \neq 0$ and

$Tv = \lambda v$ for some scalar λ , called an eigenvalue of T . *works over any field*

E.g. $A \in \mathbb{C}^{n \times n}$ with $\text{rank } A < n$ has eigenvalue 0: $v \in N(A) \setminus \{0\}$ has $Av = 0v$.

Prop 1.1: \mathcal{B} is a basis of eigenvectors of $T: V \rightarrow V$ $\Leftrightarrow [T]_{\mathcal{B}}$ is diagonal. ^{§4.3}

Pf: $T[v_1 \dots v_n] = [\lambda_1 v_1 \dots \lambda_n v_n] = [v_1 \dots v_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \Leftrightarrow [T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$. \square
 v_1, \dots, v_n eigenvectors *$[T]_{\mathcal{B}}$ diagonal*

Def: $T: V \rightarrow V$ is diagonalizable if $[T]_{\mathcal{B}}$ is diagonal for some basis \mathcal{B} of V .

An $n \times n$ matrix A is diagonalizable if μ_A is.

Lemma: A is diagonalizable $\Leftrightarrow A$ is similar to a diagonal matrix. *A is similar to Λ*

Pf: $\mu_A[v_1 \dots v_n] = [v_1 \dots v_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Leftrightarrow A \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Leftrightarrow A = P \Lambda P^{-1}$. \square

E.g. not diagonalizable: $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ (proof in a bit)

Note: Why is $A = P \Lambda P^{-1}$ useful?

$$A^k = (P \Lambda P^{-1})(P \Lambda P^{-1}) \dots (P \Lambda P^{-1}) = P \Lambda^k P^{-1} = P \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} P^{-1}$$

Aside: $e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$ (and this converges) - see Math 403

Lemma 1.2: λ is an eigenvalue of $T: V \rightarrow V$ or $n \times n$ A

$$\Leftrightarrow \underbrace{\ker(T - \lambda I) \neq 0}_{\lambda\text{-eigenspace of } T} \quad \underbrace{N(A - \lambda I) \neq 0}_{\dots \text{ of } A \quad E(\lambda)}$$

Pf: $v \neq 0$ satisfies $\underset{A}{T}v = \lambda v \Leftrightarrow \underset{A}{T}v - \lambda v = 0$
 $\Leftrightarrow \underset{A}{(T - \lambda I)}v = 0$. \square

$v \in E(\lambda) \setminus \{0\} \Leftrightarrow \underset{A}{T}v = \lambda v$ and $v \neq 0$
 $\Leftrightarrow v$ is an eigenvector with eigenvalue λ

Prop: λ is an eigenvalue of $A \Leftrightarrow \det(A - \lambda I) = 0$.

Pf: $N(A - \lambda I) \neq 0 \Leftrightarrow A - \lambda I$ singular
 $\Leftrightarrow \det(A - \lambda I) = 0$. \square

E.g. Eigenvalues of $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ are the roots of

$$\det\left(\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} - t \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 2-t & 1 \\ 0 & 2-t \end{bmatrix}\right) = (2-t)^2$$

so λ eigenvalue $\Leftrightarrow \lambda = 2$. But $E(2) = N\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) \neq 0$ and $\neq \mathbb{R}^2$

$\Rightarrow \dim E(2) = 1 \Rightarrow$ no basis of eigenvectors.

E.g. Find eigenvalues and eigenvectors of $\begin{bmatrix} 3 & 1 \\ -3 & 7 \end{bmatrix}$.

$$\det(A - tI) = \det \begin{bmatrix} 3-t & 1 \\ -3 & 7-t \end{bmatrix} = (3-t)(7-t) + 3$$

$$= 21 - 7t - 3t + t^2 + 3 = t^2 - 10t + 24$$

$$= (t-4)(t-6)$$

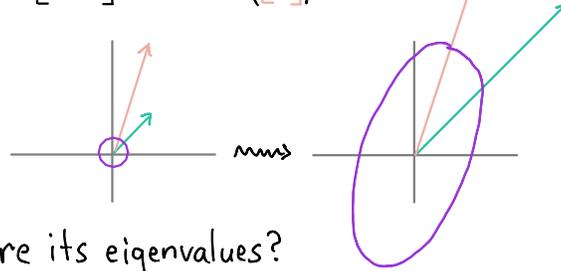
λ eigenvalue $\Leftrightarrow \lambda = 4$ or $\lambda = 6$

$$E(4) = N(A - 4I) = N \begin{bmatrix} -1 & 1 \\ -3 & 3 \end{bmatrix} = \text{span} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

eigenvector check $A \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$E(6) = N(A - 6I) = N \begin{bmatrix} -3 & 1 \\ -3 & 1 \end{bmatrix} = \text{span} \left(\begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$$

check $A \begin{bmatrix} 1 \\ 3 \end{bmatrix}$



You can "see" what μ_A does!

E.g. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. What are its eigenvalues?

$$\det(A - tI) = \det \begin{bmatrix} -t & -1 \\ 1 & -t \end{bmatrix} = t^2 + 1$$

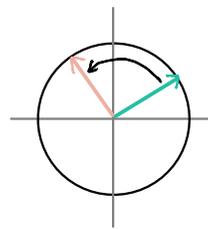
$$= 0 \Leftrightarrow t^2 = -1$$

$\Leftrightarrow t = i$ or $t = -i \Rightarrow$ no real eigenvalues!

Q. Why not? What map does A represent?

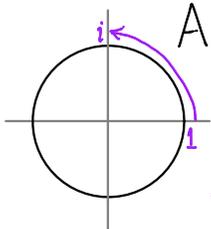
A. rotation by $\pi/2$ has no real eigenvectors:

$v \neq 0$ moves to a different (orthogonal) line, not a scalar multiple



But: $A \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}$! eigenvector with eigenvalue $-i$!

$A \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}$ eigenvector with eigenvalue i !



"multiplication by i is rotation by $\pi/2$ "!

Def: The characteristic polynomial of an $n \times n$ matrix A is $p_A(t) = \det(A - tI)$.

So root of $p_A(t) \Leftrightarrow$ eigenvalue of A .

Lemma 1.4: A similar to $B \Rightarrow p_A(t) = p_B(t)$.

Pf: Suppose $B = QAQ^{-1}$. Then $p_B(t) = p_{QAQ^{-1}}(t) = \det(QAQ^{-1} - tI)$

$$= \det(QAQ^{-1} - Q(tI)Q^{-1})$$

$$= \det(Q(A - tI)Q^{-1})$$

$$= \cancel{\det Q} \det(A - tI) \cancel{\det Q^{-1}}$$

$$= p_A(t). \square$$

Def: $T: V \rightarrow V$ has characteristic polynomial $p_T = p_A$ for some (hence every) $A = [T]_{\mathcal{B}}$.
 Well defined by Lemma 1.4.

$$p_A(t) = c_n t^n + c_{n-1} t^{n-1} + \dots + c_1 t + c_0.$$

Q. What's c_n ?

A. $(-1)^n$

Q. What's c_0 ?

A. $c_0 = p_A(0) = \det(A - 0I) = \det A$.

A': A diagonalizable, say $A \sim \Lambda$, $\Rightarrow p_A(t) = p_{\Lambda}(t) = \det \begin{bmatrix} \lambda_1 - t & & \\ & \ddots & \\ & & \lambda_n - t \end{bmatrix}$

$$= (\lambda_1 - t) \dots (\lambda_n - t)$$

has $c_0 = \lambda_1 \dots \lambda_n$.

General (next lecture): $\det A =$ product of eigenvalues with multiplicities counted correctly.

A singular $\Leftrightarrow 0$ is an eigenvalue

\Leftrightarrow product of eigenvalues is 0

$\Leftrightarrow \det A = 0$

} explains why det detects singularity

§6.2 Recall: λ eigenvalue of $T: V \rightarrow V$ or $n \times n$ $A = [T]_{\mathcal{B}}$

$\Leftrightarrow Tv = \lambda v$ for some $v \neq 0$ eigenvector

$\Leftrightarrow \lambda$ root of characteristic polynomial $\det(A - tI) = p_T(t)$

Today: how many linearly independent eigenvectors with eigenvalue λ can A have?

Def: geometric multiplicity $g(\lambda) = \dim E(\lambda) = \dim \ker(T - \lambda I)$

Compare: algebraic multiplicity $a(\lambda) = \# \text{ times } t - \lambda \text{ divides } p_T(t)$

$p_T(\lambda) = 0 \Rightarrow p_T(t) = (t - \lambda)q(t)$, $q(\lambda) = 0 \Rightarrow q(t) = (t - \lambda)r(t)$, ... until λ not a root

E.g. $A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \Rightarrow p_A(t) = \det \begin{bmatrix} \lambda - t & 1 \\ 0 & \lambda - t \end{bmatrix} = (\lambda - t)^2 \Rightarrow a(\lambda) = 2$

$E(\lambda) = N(A - \lambda I) = N \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right)$ has $\dim 1 = g(\lambda)$

Theorem 2.1 (strengthened): Suppose B is a set of eigenvectors for a linear map $T: V \rightarrow V$.

If $B \cap E(\lambda)$ is linearly independent for every (eigenvalue) λ , then B is linearly independent.

Pf: Let $c_1 v_1 + \dots + c_k v_k = 0$ with $v_1, \dots, v_k \in B$. Need $c_i = 0 \forall i$.

$k = 1 \Rightarrow c_1 = 0$ since $v_1 \neq 0$. Assume, by induction on k , that all size $k-1$ subsets of B are independent.

Suppose $v_i \in E(\lambda_i) \forall i$. Then $0 = (T - \lambda_1 I)(c_1 v_1 + \dots + c_k v_k)$

$$= c_1 (T v_1 - \lambda_1 v_1) + \dots + c_k (T v_k - \lambda_1 v_k)$$

$$= c_1 (\lambda_1 v_1 - \lambda_1 v_1) + \dots + c_k (\lambda_k v_k - \lambda_1 v_k)$$

$$= c_1 (\cancel{\lambda_1 - \lambda_1}) v_1 + \dots + c_k (\lambda_k - \lambda_1) v_k$$

$$= c_2 (\lambda_2 - \lambda_1) v_2 + \dots + c_k (\lambda_k - \lambda_1) v_k$$

$$\Rightarrow c_i (\lambda_i - \lambda_1) = 0 \forall i \geq 2 \text{ by induction}$$

$$\Rightarrow c_i = 0 \forall i \text{ such that } \lambda_i \neq \lambda_1$$

$$\Rightarrow c_i = 0 \forall i \text{ by linear independence of } B \cap E(\lambda_1). \quad \square$$

Corollary 2.2: $\dim V = n$ and $T: V \rightarrow V$ has n distinct eigenvalues in $\mathbb{F} \Rightarrow T$ is diagonalizable over \mathbb{F} .

Pf: $E(\lambda) \neq 0$ if $p_T(\lambda) = 0$, so T has n eigenvectors with distinct eigenvalues.

These are independent by Thm 2.1 since $\#(B \cap E(\lambda)) = 1 \forall$ eigenvalues λ and

hence form a basis because $\dim V = n$. \square

Q. Can hypothesis of Cor. 2.2. fail?

A. Yes 1. repeated roots, e.g. $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$

2. complex roots, e.g. $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = [\text{rot}_{\pi/4}]_E$ $p_A(t) = (t - \frac{1}{\sqrt{2}})^2 + \frac{1}{2}$

$= t^2 - t\sqrt{2} + 1$ has distinct roots

$$\frac{\sqrt{2} \pm \sqrt{2-4}}{2} = \frac{\sqrt{2} \pm i\sqrt{2}}{2} = \frac{1}{\sqrt{2}} \pm i\frac{1}{\sqrt{2}} \notin \mathbb{R}.$$

E.g. $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ & 3 & 1 \\ & 0 & 3 \end{bmatrix}$ vs. $B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \\ & 3 & 0 \\ & 0 & 3 \end{bmatrix}$

$$p_A(t) = (2-t)^2(3-t)^2 = p_B(t) \Rightarrow a(2) = 2$$

$$g(2) = 2 \qquad 1 \qquad a(3) = 2$$

$$g(3) = 1 \qquad 2$$

Prop. 2.3: $1 \leq g(\lambda) \leq a(\lambda)$.

Pf: Pick a basis v_1, \dots, v_g of $E(\lambda)$, so $g = g(\lambda)$. Extend to a basis $\mathcal{B} = v_1, \dots, v_n$ of V .

Then $[T]_{\mathcal{B}} = \begin{bmatrix} \lambda I_g & B \\ O & C \end{bmatrix}$ so $p_A(t) = p_{\lambda I_g}(t) p_C(t) = (\lambda - t)^g p_C(t) \Rightarrow g \leq a(\lambda)$. \square

Thm 2.4: $T: V \rightarrow V$ diagonalizable \iff all eigenvalues lie in \mathbb{F} and $g(\lambda) = a(\lambda) \forall \lambda$.

Pf: T diagonalizable $\Rightarrow [T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = A$, where \mathcal{B} is a basis of eigenvectors each $\lambda_i \leftrightarrow$ basis vector v_i Prop 2.3

v_1, \dots, v_n with eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. $a(\lambda) = \# \text{ times } \lambda \text{ appears in } A \leq g(\lambda) \leq a(\lambda)$

$\Rightarrow g(\lambda) = a(\lambda)$. On the other hand, assume $a(\lambda) = g(\lambda) \forall \lambda$ and that all eigenvalues $\in \mathbb{F}$.

Fix basis \mathcal{B}_λ for $E(\lambda)$. The union $\mathcal{B} = \bigcup_{\lambda} \mathcal{B}_\lambda$ of these bases is linearly independent by Thm. 2.1, so \mathcal{B} is a basis because $\sum_{\lambda} a(\lambda) = n$ by "all eigenvalues $\in \mathbb{F}$ ". \square

E.g. $A = \begin{bmatrix} -1 & 4 & 2 \\ -1 & 3 & 1 \\ -1 & 2 & 2 \end{bmatrix}$ vs. $B = \begin{bmatrix} 0 & 3 & 1 \\ -1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ have $p_A(t) = p_B(t) = (1-t)^2(2-t)$

$a(1) = 2 \quad a(2) = 1. \quad g(1) = ?$

$A - I = \begin{bmatrix} -2 & 4 & 2 \\ -1 & 2 & 1 \\ -1 & 2 & 1 \end{bmatrix}$ $B - I = \begin{bmatrix} -1 & 3 & 1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ rows not multiples of one row
 can't = 3 since $\det(B - I) = 0$

$\text{rank} = 1 \Rightarrow g(1) = 2$ $\text{rank} > 1 \Rightarrow \text{rank} = 2 \Rightarrow g(1) = 1.$

§6.4 What kinds of real matrices are symmetric? Projections!

$P_V = QQ^T$ if cols q_1, \dots, q_n of Q are \perp normal basis for $V \subseteq \mathbb{R}^m$.

$(QQ^T)^T = (Q^T)^T Q^T = QQ^T = q_1 q_1^T + \dots + q_n q_n^T = \text{proj}_{q_1} + \dots + \text{proj}_{q_n}$

also symmetric: $\lambda_1 q_1 q_1^T + \dots + \lambda_n q_n q_n^T$ any linear combination of projections

$Q \Lambda Q^T$

$\begin{bmatrix} | & & | \\ q_1 & \dots & q_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} -q_1^T \\ \vdots \\ -q_n^T \end{bmatrix}$

Def: The spectrum of a matrix is its (multi)set of eigenvalues. (atoms)

Thm 4.1 (spectral theorem): All real symmetric matrices arise this way[↑]:

$A \in \mathbb{R}^{n \times n}$ symmetric \Rightarrow 1. A has real eigenvalues (\Rightarrow diagonalizable/ \mathbb{R})

$Q^{-1} = Q^T$ 2. \mathbb{R}^n has orthonormal basis q_1, \dots, q_n of eigenvectors of $A \Rightarrow$

1+2 $\Leftrightarrow \exists$ orthogonal matrix Q with $Q^{-1} A Q = \Lambda$ diagonal

$A = Q \Lambda Q^T$

Lemma: H symmetric $\Rightarrow H y \cdot x = y \cdot H x \forall x, y$.

Pf: $H y \cdot x = (x^T H y)^T = y^T H^T x = H^T x \cdot y$. But $H = H^T$. \square *complex conjugate*

Pf of Thm: 1. Suppose $\lambda = a + bi$ eigenvalue of A , so $\bar{\lambda} = a - bi$ is, too.

finally! $\mathbb{C}^{n \times n}$ Set $S = (A - \lambda I)(A - \bar{\lambda} I) = A^2 - (\lambda + \bar{\lambda})A + \lambda \bar{\lambda} I = A^2 - 2aA + (a^2 + b^2)I \in \mathbb{R}^{n \times n}$ after all.

$\det S = 0$ since $A - \lambda I$ is singular. Pick $x \in N(S) \setminus 0$:

$0 = Sx \cdot x = (A^2 - 2aA + a^2 I)x \cdot x + (b^2 I)x \cdot x$
 $= (A - aI)^2 x \cdot x + b^2 x \cdot x$

$\stackrel{\text{Lemma}}{=} (A - aI)x \cdot (A - aI)x + bx \cdot bx$

$= \|(A - aI)x\|^2 + \|bx\|^2 \Rightarrow (A - aI)x = 0$ and $bx = 0$

$x \neq 0 \Rightarrow x \in E(a)$ and $b = 0$.

2. Pick unit vector $q_1 \in E(\lambda_1)$ and v_2, \dots, v_n orthonormal basis for q_1^\perp . Set $\mathcal{B} = q_1, v_2, \dots, v_n$

$[\mu_A]_{\mathcal{B}} = B = \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & & \\ 0 & & & C \end{bmatrix}$ *who knows?* Change-of-basis formula: $B = P^{-1} A P$ where $P = \begin{bmatrix} | & & | \\ q_1 & v_2 & \dots & v_n \\ | & & & | \end{bmatrix} = P^T A P$

$\Rightarrow B^T = P^T A^T (P^T)^T = P^T A^T P$ symmetric

$\Rightarrow * = 0 \forall *$ and C symmetric of size $n-1$

\Rightarrow done by induction, the case $n=1$ being very easy. \square

E.g. Suppose $p_A(t) = (7-t)^5$

$\dim N(A-7I) = 2$ $e_1 \mapsto 0, e_2 \mapsto 0$ What is $J = J(A)$?

$\dim N((A-7I)^2) = 4$ $e_3 \mapsto e_1, e_4 \mapsto e_2$ $\dim N((A-7I)^3) = 5$

Ans: $n \times n$ for $n = 5$ 2 Jordan blocks

only two choices: sizes

$$1+4: \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} \text{ or } 2+3: \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} = J-7I \Rightarrow J = \begin{bmatrix} 7 & 1 & & & \\ & 7 & & & \\ & & 7 & 1 & \\ & & & 7 & 1 \\ & & & & 7 \end{bmatrix}$$

General meaning of a Jordan basis \mathcal{B} :

$$\begin{aligned} v_1, \dots, v_d \leftrightarrow \text{a block} &\Rightarrow Tv_1 = \lambda v_1 && \Leftrightarrow (T-\lambda I)v_1 = 0 \\ Tv_2 = \lambda v_2 + v_1 &&& (T-\lambda I)v_2 = v_1 \neq 0, \text{ but } (T-\lambda I)^2 v_2 = 0 \\ Tv_3 = \lambda v_3 + v_2 &&& (T-\lambda I)v_3 = v_2 \neq 0, \text{ and } (T-\lambda I)^2 v_3 \neq 0 \\ &&& \text{but } (T-\lambda I)^3 v_3 = 0 \\ \vdots &&& \vdots \\ Tv_d = \lambda v_d + v_{d-1} &&& (T-\lambda I)v_d = v_{d-1} \neq 0, \dots, (T-\lambda I)^{d-1} v_d \neq 0 \\ &&& \text{but } (T-\lambda I)^d v_d = 0 \end{aligned}$$

$$T[v_1 \dots v_d] = [v_1 \dots v_d] \begin{bmatrix} \lambda & & & & \\ & \lambda & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & \lambda & \\ & & & & & \lambda \end{bmatrix} \Leftrightarrow \dots$$

$$\dim N((A-\lambda I)^r) - \dim N((A-\lambda I)^{r-1}) = \# \text{ blocks of size } \geq r.$$

Note: If T has only one Jordan block, then

$v_d =$ almost any element of V works!

Need only $v_d \notin \ker((T-\lambda I)^{d-1})$.

26. § 7.3 Systems of ODE Ordinary Differential Equations

Recall: $f = f(t)$ satisfies $f' = af \Leftrightarrow f = Ce^{at}$,
 $f(0) = C$

Q. If $x = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$ satisfies $x'(t) = Ax(t)$ with $A \in \mathbb{R}^{n \times n}$,

$$\text{so } \begin{aligned} x_1'(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) \\ &\vdots \\ x_n'(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t) \end{aligned}$$

in what sense is "the" solution Ce^{tA} ?

Recall: $e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots + \frac{a^k}{k!} + \dots$

Def: $A \in \mathbb{R}^{n \times n} \Rightarrow e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots + \frac{A^k}{k!} + \dots$

$$e^{tA} = I + tA + t^2 \frac{A^2}{2!} + \dots + t^k \frac{A^k}{k!} + \dots$$

convergence issues: sequences in $\mathbb{R}^{n \times n}$

E.g. 1. $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Rightarrow e^\Lambda = \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{bmatrix}$ and $e^{t\Lambda} = \begin{bmatrix} e^{t\lambda_1} & & \\ & \ddots & \\ & & e^{t\lambda_n} \end{bmatrix}$

2. $P^{-1}AP = \Lambda \Rightarrow A = P\Lambda P^{-1}$
 $\Rightarrow A^k = P\Lambda^k P^{-1}$
 $\Rightarrow e^A = Pe^\Lambda P^{-1}$
 $e^{tA} = Pe^{t\Lambda} P^{-1}$

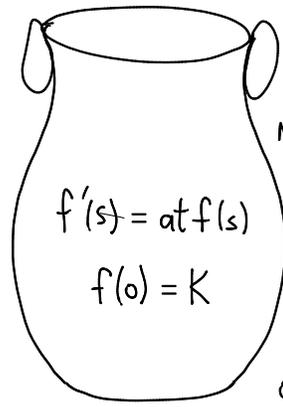
3. $A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \Rightarrow e^{tA} = \begin{bmatrix} e^{2t} & \\ & e^{-t} \end{bmatrix}$
 $\Rightarrow e^{tA} = Pe^{t\Lambda}P^{-1} = \begin{bmatrix} e^{2t} & 0 \\ e^{2t} - e^{-t} & e^{-t} \end{bmatrix}$

General: $v \in \mathbb{R}^n \Rightarrow$ entries of $e^{tA}v$ are functions of t

$C(e^{tA}) =$ analogue of Ce^{at} : *typographical pun!*

vectors $e^{tA}v$ are solutions of $x'(t) = Ax(t)$.

Thm 3.3: For $A \in \mathbb{R}^{n \times n}$, solutions set of $x'(t) = Ax(t)$ is the vector space $C(e^{tA})$ of dim n .



Riddle: what is this?
 Note: the solution is a hint.

$$\Rightarrow f(s) = Ke^{ats}$$

Answer:
 ODE on a Grecian urn!

E.g. $A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}$ and $x' = Ax \Rightarrow x = e^{tA} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$= v_1 \begin{bmatrix} e^{2t} \\ e^{2t} - e^{-t} \end{bmatrix} + v_2 \begin{bmatrix} 0 \\ e^{-t} \end{bmatrix}$ *eigenvalues of A*

$= v_1 e^{2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (v_2 - v_1) e^{-t} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *eigenvectors of A*

Lemma: $(e^{tA})' = (I + tA + t^2 \frac{A^2}{2!} + \dots + t^k \frac{A^k}{k!} + \dots)'$
 $= 0 + A + tA^2 + t^2 \frac{A^3}{2!} + \dots + t^k \frac{A^{k+1}}{k!} + \dots$
 $= A(I + tA + t^2 \frac{A^2}{2!} + \dots + t^k \frac{A^k}{k!} + \dots)$
 $= Ae^{tA} = e^{tA}A. \quad \square$

Pf of Thm: $v \in \mathbb{R}^n \Rightarrow (e^{tA}v)' = (e^{tA})'v$ because $\frac{d}{dt}$ is linear
 $= (Ae^{tA})v$ by Lemma
 $= A(e^{tA}v). \quad \checkmark$

Need: every sol is $e^{tA}v$ for some $v \in \mathbb{R}^n$.

Assume $x'(t) = Ax(t)$. Set $y(t) = e^{-tA}x(t)$.

*Amazing: same proof as sols ($f' = f$) = span(e^{at}):
Divide purported sol by e^{at} and conclude
by given equation that quotient is constant.*

Then $y'(t) = (e^{-tA}x(t))'$
 $= (e^{-tA})'x(t) + e^{-tA}x'(t)$
 $= -Ae^{-tA}x(t) + e^{-tA}Ax(t)$ by Lemma + hypothesis
 $= (-Ae^{-tA} + \underbrace{e^{-tA}A})x(t)$
 $= 0$ *Ae^{-tA} by Lemma*

$\Rightarrow y_1'(t) = 0 \Rightarrow y(t) = v \in \mathbb{R}^n$ is constant
 \vdots
 $y_n'(t) = 0$

$\Rightarrow x(t) = e^{tA}y(t)$
 $= e^{tA}v. \quad \checkmark$

dim C(e^{tA}) = n:

entries of e^{tA} are functions, not scalars, so can't ask e^{tA} to be invertible.

Check that cols of e^{tA} are indep.

Need e^{tA} v = 0 \Rightarrow v = 0.

But e^{tA} v = 0 \Rightarrow 0 = (e^{tA} v)|_{t=0} = e^{0A} v = I v = v. \square

cols of e^{tA}|_{t=0} are indep.

Cor 3.4: Solution set of general order n ODE

(*) f^{(n)}(t) + a_{n-1} f^{(n-1)}(t) + \dots + a_2 f''(t) + a_1 f'(t) + a_0 f(t) = 0

with constant coeffs a_{n-1}, \dots, a_0 is an n-dim subspace of C^\infty(\mathbb{R}).

E.g. n = 2: f'' + f = 0

\Rightarrow all sols are linear combinations of sin and cos.

Pf: Set x(t) = [f(t), f'(t), \dots, f^{(n-1)}(t)]^T and A = [0 1 0 \dots 0, 0 0 1 \dots 0, \dots, 0 \dots 0 -a_{n-1}]

Then x'(t) = Ax(t) \Leftrightarrow (*)

Sols of (*) are the top entries of sols of x' = Ax.

e.g. e^{tA} has top row [f_1(t) \dots f_n(t)]

with f_j(t) a sol of (*) \forall j.

Moreover, c_1 f_1 + \dots + c_n f_n = 0 \Rightarrow c_1 [f_1, f_1', \dots, f_1^{(n-1)}]^T + \dots + c_n [f_n, f_n', \dots, f_n^{(n-1)}]^T = 0

cols of e^{tA} indep. \Rightarrow c_1 = \dots = c_n = 0,

so f_1, \dots, f_n are independent in C^\infty(\mathbb{R}). \square

