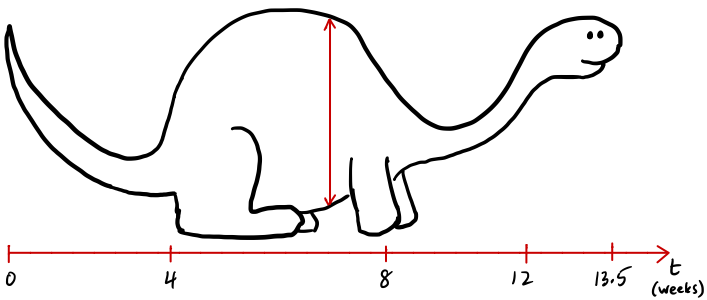
From The Book of Strang

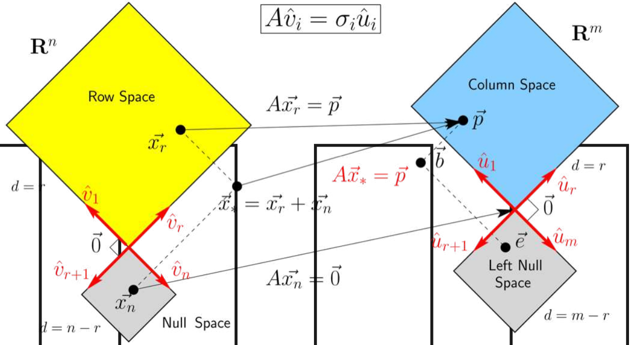
Matrixology
(linear algebra)

Prof Peter Sheridan Dodds

Recorded in 2016

Melvin the Course Difficulty Dinosaur:





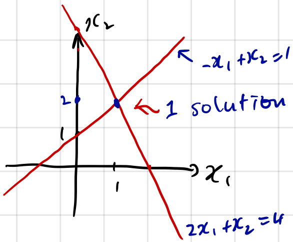
E1ap1

The Central problem of Matrixology:

> Given a matrix $\mathbb{A}_{m\times n}$ and a vector $\vec{b}_{m\times 1}$
> find all $\vec{x}_{n\times 1}$ such that
> $$\underset{m\times n}{\mathbb{A}}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$$

$$\left.\begin{aligned} -x_1 + x_2 &= 1 \quad \cdots (1) \\ 2x_1 + x_2 &= 4 \quad \cdots (2) \end{aligned}\right\}$$ system of linear equations

2 × 2 system
- m rows = # equations
- n columns = # variables



Row picture

Usual way:

$-x_1 + x_2 = 1 \quad \cdots (1)$

$2x_1 + x_2 = 4 \quad \cdots (2)$

eq (3) = eq (2) + 2 eq (1)

$3x_2 = 6 \quad \cdots (3)$

$\Rightarrow x_2 = 2$

substitute into eq (1)

$-x_1 + 2 = 1$

$\Rightarrow x_1 = 1$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

notes
- Found intersection of two lines
- Both equations are true at this one point

Algebra ⇄ Geometry (yes!)

Three possibilities:

a) 1 soln ; b) no soln ; c) same line. ∞ly many soln.

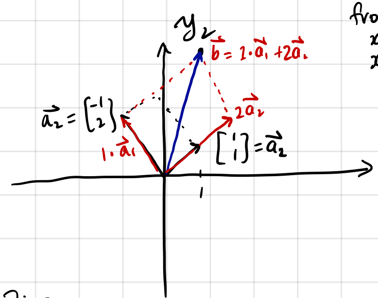
Column picture

Rewrite system as

$$x_1 \begin{bmatrix} -1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$\vec{a}_1$, $\vec{a}_2$: building blocks; $\vec{b}$

from before:
$x_1 = 1$
$x_2 = 2$

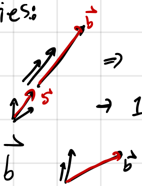


Three possibilities:

a) $\vec{a}_1 \parallel \vec{a}_2 \parallel \vec{b}$ ⇒ ∞ly many solns

b) $\vec{a}_1 \nparallel \vec{a}_2$ → 1 soln (can always make $\vec{b}$ in one way only)

c) $\vec{a}_1 \parallel \vec{a}_2 \nparallel \vec{b}$ ⇒ 0 solns.



E1ap2

$$\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$$

$$\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

Matrix Picture:

Two ways to multiply matrices.

① dot products of rows and columns

$$\underset{2\times 2}{\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}}\underset{2\times 1}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} = \underset{2\times 1}{\begin{bmatrix} -x_1 + x_2 \\ 2x_1 + x_2 \end{bmatrix}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

Row picture: $-x_1 + x_2 = 1$

$2x_1 + x_2 = 4$

②

$$\left[\begin{bmatrix} -1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} -1 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

row vector
1x2

3x3 example: $A\vec{x} = \vec{b}$

Matrix picture:

$$\underbrace{\begin{bmatrix} 2 & 1 & 0 \\ -1 & 1 & 2 \\ 0 & 3 & 1 \end{bmatrix}}_{\substack{3\times 3 \\ m \times n}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\substack{n\times 1 \\ 3\times 1}} = \underbrace{\begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}}_{\substack{m\times 1 \\ 3\times 1}}$$

← find $\vec{x}$ such that A transforms $\vec{x}$ into $\vec{b}$

Multiply out:

Row picture:

$$\begin{aligned} 2x_1 + x_2 &= 2 \\ -x_1 + x_2 + 2x_3 &= 2 \\ 3x_2 + x_3 &= 6 \end{aligned}$$

← equations of planes in 3-d

More sneakily

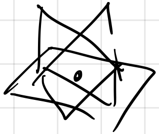
Column picture:

$$x_1 \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}$$

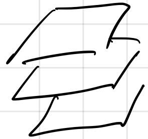
0   2   0

See: $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ (Always takes more work than this!!!)

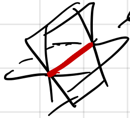
Row picture:



1 soln



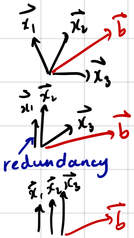
0 soln



← ∞ly many solns

↑ way too hard in 4-d and above...

Column Picture:



1 soln.

0 or ∞ly many soln, } depends on $\vec{b}$.

0 solns or ∞ly many }

← easy in many dimensions

E2ap1

Story: We (people + computers) solve systems of linear equations by "Elimination" ← Gaussian & Gauss-Jordan

Menu:
- Perform Elimination using Row Operations
- Anatomy of Row operations • Triangles !!
- Back Substitution
- key: Pivots D_i, multipliers l_{ij}, upper triangular matrix U, Augmented Matrix
- when things go "wrong"

$$A\vec{x} = \vec{b}$$

$$\begin{cases} -x_1 + 3x_2 = 1 & \dots (1) \; R_1 \\ 2x_1 + x_2 = 5 & \dots (2) \; R_2 \end{cases}$$

(2 ← eliminate; -1)

2x2 system: $m = 2$ = # eqs, $n = 2$ = # variables

$D_1 = -1$:

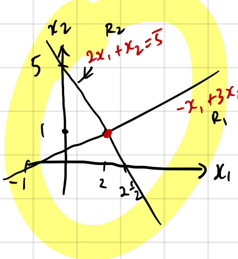
$-x_1 + 3x_2 = 1 \quad \dots R_1$

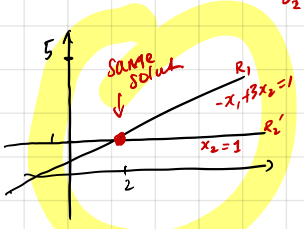
$0x_1 + 7x_2 = 7 \quad \dots$ (upper triangular; D_2)

$$R_2' = R_2 - \left(\frac{2}{-1}\right) R_1$$

l_{21} ← always use this form; $+2$; D_1 = first pivot without division; $D_2 = 7$

R_2 (drop primes): $x_2 = 1$





We have $x_2 = 1$, now solve for x_1 using back substitution:

$$-x_1 + 3x_2 = 1 \Rightarrow -x_1 + 3 = 1$$

$$x_1 = 2$$

(↖ $x_2 = 1$)

solution: $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ✓

For later, we can go further and avoid back substitution. Gauss-Jordan elimination:

$$\begin{cases} -x_1 + 3x_2 = 1 & \dots R_1 \\ 0 + 1x_2 = 1 & \dots R_2 \end{cases}$$

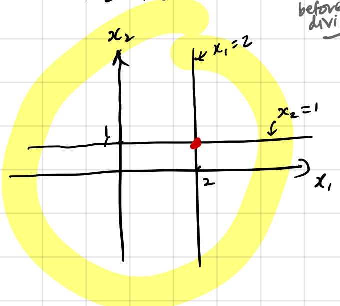
$$R_1' = R_1 - \left(\frac{3}{1}\right) R_2$$

$D_2 = 7$ before division

$$\begin{cases} -x_1 + 0 = -2 \\ 0 + x_2 = 1 \end{cases}$$

$R_2' = R_2$

$$\Rightarrow \; +x_1 = 2$$
$$x_2 = 1$$



E2ap2

Basic Elimination rules:

1. Create upper triangular system by systematic by row operations
2. Swap rows if needed when pivots = 0

$$\begin{aligned} 0 + x_2 &= 3 \\ 3x_1 - 7x_2 &= 0 \end{aligned} \quad \underset{R_1 \leftrightarrow R_2}{\sim} \quad \begin{aligned} 3x_1 - 7x_2 &= 0 \\ x_2 &= 3 \end{aligned}$$

Augmented Matrix approach:

$$\begin{aligned} -x_1 + 3x_2 &= 1 \\ 2x_2 + x_2 &= 5 \end{aligned} \Rightarrow \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Row picture; $A\vec{x} = \vec{b}$, matrix

$D_1 \rightarrow$

$$\left[\begin{array}{cc|c} -1 & 3 & 1 \\ 2 & 1 & 5 \end{array}\right] \sim \quad R_2' = R_2 - \left(\frac{2}{-1}\right) R_1 \quad \left[\begin{array}{cc|c} -1 & 3 & 1 \\ 0 & 7 & 7 \end{array}\right]$$

$[A \mid \vec{b}]$; D_1 ; D_2

$\sim$ means systems have same solut.

E2b p1

Menu:

- 3x3 example of solving $A\vec{x}=\vec{b}$ with Elimination and Row Swaps
- Turn $A\vec{x}=\vec{b}$ into $U\vec{x}=\vec{c}$ (U: upper triangul)

Row picture:

$$2x_1 - 3x_2 + 0\cdot x_3 = 3 \quad \text{eq1}$$
$$4x_1 - 5x_2 + 1\cdot x_3 = 7 \quad \text{eq2}$$
$$2x_1 - 1x_2 - 3x_3 = 5 \quad \text{eq3}$$

Three planes

Column Picture:

$$x_1\begin{bmatrix}2\\4\\2\end{bmatrix} + x_2\begin{bmatrix}-3\\-5\\-1\end{bmatrix} + x_3\begin{bmatrix}0\\1\\-3\end{bmatrix} = \begin{bmatrix}3\\7\\5\end{bmatrix}$$

($\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$, $\vec{b}$)

Matrix Picture

$$\begin{bmatrix}2 & -3 & 0\\4 & -5 & 1\\2 & -1 & -3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}3\\7\\5\end{bmatrix} \qquad A\vec{x}=\vec{b}$$

Augmented Matrix version of row picture:

$$\left[\begin{array}{ccc|c}2 & -3 & 0 & 3\\4 & -5 & 1 & 7\\2 & -1 & -3 & 5\end{array}\right] \begin{matrix}R_1\\R_2\\R_3\end{matrix}$$

$[A \mid \vec{b}]$: A is $m \times n$, $\vec{b}$ is $m \times 1$, $[A \mid \vec{b}]$ is $m \times (n+1)$

(1), (2), (3) = order of eliminati.

multiplier

$$R_2' = R_2 - \left(\frac{4}{2}\right)R_1 \quad (l_{21},\ D_1)$$

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

$$R_3' = R_3 - \left(\frac{2}{2}\right)R_1 \quad (l_{31},\ D_1)$$

$$\left[\begin{array}{ccc|c}2 & -3 & 0 & 3\\0 & 1 & 1 & 1\\0 & 2 & -3 & 2\end{array}\right] \quad (D_1,\ D_2)$$

$$R_3' = R_3 - \left(\frac{2}{1}\right)R_2 \quad (l_{32},\ D_2)$$

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

main diagonal

$$U\vec{x} = \vec{c} \Rightarrow \begin{bmatrix}2 & -3 & 0\\0 & 1 & 1\\0 & 0 & -5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}3\\1\\0\end{bmatrix}$$

(U from A, $\vec{c}$ from $\vec{b}$)

easy to solve with back substitut

Back substitution:

Step back out to equations and work upwards:

R_3: $-5x_3 = 0 \Rightarrow x_3 = 0$

R_2: $x_2 + x_3 = 1 \Rightarrow x_2 = 1$ ($x_3 = 0$)

R_1: $2x_1 - 3x_2 = 3 \Rightarrow 2x_1 - 3 = 3$ ($x_2 = 1$)

$2x_1 = 6$

$x_1 = 3$

Solution:

$$\vec{x} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}3\\1\\0\end{bmatrix}$$

Important:

$l_{21} = 2$ $\quad l_{32} = 2$

$l_{31} = 1$ $\quad$ multipliers

Pivots: find in U:

$D_1 = 2,\ D_2 = 1,\ D_3 = 5$

E2cp1

Menu:

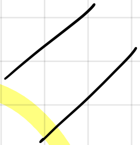
- what can happen when a pivot is zilch....
- Singular system

ex 1

$-x_1 + x_2 = 1$ ....R_1

$x_1 - x_2 = 5$ ...R_2 (+1)

parallel



$\Rightarrow \left[\begin{array}{cc|c} -1 & 1 & 1 \\ 1 & -1 & 5 \end{array}\right]$

"has same solution as"

$R_2' = R_2 - \left(\frac{1}{-1}\right) R_1$ (l_{21}, D_1)

$\left[\begin{array}{cc|c} -1 & 1 & 1 \\ 0 & 0 & 6 \end{array}\right]$

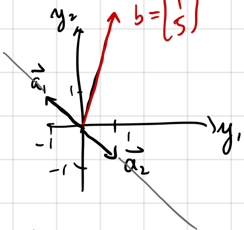
0 pivot — pivot is "missing"

R_2: $0x_1 + 0x_2 = 6$ not true!

$0 = 6$

Column picture:

$x_1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ ($\vec{a}$, $\vec{b}$, $\vec{b}$)



example of Singular System

↳ no unique solution — may have 0 or ∞ly many

ex 2

$-x_1 + x_2 = 1$ ..... R_1

$2x_1 - 2x_2 = -2$ ..... R_2

$\left[\begin{array}{cc|c} -1 & 1 & 1 \\ 2 & -2 & -2 \end{array}\right]$ $A\vec{x} = \vec{b}$

$R_2' = R_2 - \left(\frac{2}{-1}\right) R_1$ ($l_{21} = -2$, D_1)

$\left[\begin{array}{cc|c} -1 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$ $U\vec{x} = \vec{c}$ ← singular matrix

eqs: $-x_1 + x_2 = 1$ ... R_1

$0 = 0$ ... R_2

x_1: later pivot variable; x_2: free variable.

Let $x_2 \in \mathbb{R}^1$ (← real numbers) → x_1 now depends on x_2

$-x_1 = 1 - x_2$

$x_1 = x_2 - 1$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 - 1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}$

($1 \cdot x_2 - 1$, $1 \cdot x_2 + 0$)

replace pivot variables with free variables

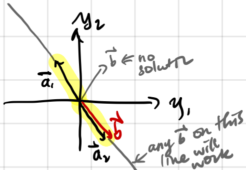
dialable piece, fixed

where $x_2 \in \mathbb{R}^1$

Column pic

$x_1 \begin{pmatrix} -1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ ($\vec{a}_1$, $\vec{a}_2$, $\vec{b}$)

$\vec{a}_1 = -\vec{a}_2$



E2dp1

- Our task: Solve systems of linear equations
- Three pictures: row (where solving happens), column (understanding), & matrix (deep understanding).

2x2 example from Episode 2

$$-x_1 + 3x_2 = 1 \quad \text{(Row 1)}$$
$$2x_1 + x_2 = 5 \quad \text{(Row 2)}$$

row pic $\Leftrightarrow$ $x_1\begin{bmatrix}-1\\2\end{bmatrix} + x_2\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}1\\5\end{bmatrix}$ (column pic; $\vec{a}_1$, $\vec{a}_2$, $\vec{b}$) $\Leftrightarrow$ $\begin{bmatrix}-1 & 3\\2 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1\\5\end{bmatrix}$ (matrix pic; $\mathbb{A}$, $\vec{x}$, $\vec{b}$)

Solve by Gaussian Elimination

Equations (first) $\longleftrightarrow$ (same) Augmented Matrix (second) (tidy)

$-x_1 + 3x_2 = 1$ ... R1

$2x_1 + x_2 = 5$ ... R2

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

multiplier l_{21}

$\Rightarrow$ $-x_1 + 3x_2 = 1$ ... R1

$0x_1 + 7x_2 = 7$ ... $R2' = R2 - \left(\frac{2}{-1}\right)R_1$

↑ echelon form

$R_2' = R_2 - \left(\frac{2}{-1}\right)R_1$ — $D_1 = -1$ = first pivot

multiplier $l_{21} = -2$

$$\left[\begin{array}{cc|c}-1 & 3 & 1\\0 & 7 & 7\end{array}\right]$$

↑ echelon form

Matrix picture: $\mathbb{A}\vec{x} = \vec{b} \Rightarrow \mathbb{U}\vec{x} = \vec{c}$, with $\mathbb{U} = \begin{bmatrix}-1 & 3\\0 & 7\end{bmatrix}$, $\vec{c} = \begin{bmatrix}1\\7\end{bmatrix}$

The Gaussian Eliminator 9000:

Augmented Matrix for $\mathbb{A}\vec{x} = \vec{b}$:

$$\left[\begin{array}{ccc|c}2 & -3 & 0 & 3\\4 & -5 & 1 & 7\\2 & -1 & -3 & 5\end{array}\right]$$

(① 4, ② 2, ③ −1)

multiplier $l_{21} = 2$

$R_2' = R_2 - \left(\frac{4}{2}\right)R_1$ ($2 = D_1$; big tilde: equivalent to)

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

$l_{31} = 1$

$R_3' = R_3 - \left(\frac{2}{2}\right)R_1$ ($2 = D_1$)

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

rowoperatorify

$l_{32} = 2$

$R_3' = R_3 - \left(\frac{2}{1}\right)R_2$ ($1 = D_2$)

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

D_1, D_2, D_3 (pivots); #eliminated; echelon form

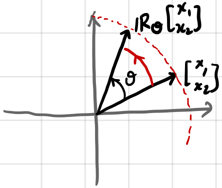
$\mathbb{U}\vec{x} = \vec{c}$ ⇓ easy to solve with back substitution

Menu:
- Using Elimination matrices to do the work for us
- Surprising help for our understanding will be possible
- Somehow, elimination makes two triangles.

Observation: Matrices can do sneaky, gadgety things for us

ex Rotate a vector in 2-d through θ radians

$$\underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\mathbb{R}_\theta} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



ex Permute entries in a vector:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_4 \\ x_1 \end{bmatrix}$$ ← cycle by 1.

Plan: encode row operations as
(1) elimination matrices ← normal elimination steps
& (2) permutation matrices ← row swaps

E3ap1

Augmented Matrix approach:

$$\begin{aligned} -x_1 + 3x_2 &= 1 \\ 2x_2 + x_2 &= 5 \end{aligned} \Rightarrow \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Row picture $\qquad A\vec{x} = \vec{b}$, matrix

$D_1 \rightarrow$

$$\left[\begin{array}{cc|c} -1 & 3 & 1 \\ 2 & 1 & 5 \end{array}\right] \qquad \sim \qquad R_2' = R_2 - \left(\frac{2}{-1}\right) R_1 \qquad \left[\begin{array}{cc|c} -1 & 3 & 1 \\ 0 & 7 & 7 \end{array}\right]$$

$[A \mid \vec{b}\,]$

$\ell_{21} = -2$, D_1, D_2

$\sim$ means systems have same solut.

replace w. matrix multiplication

$\mathbb{E}_{21}$ = elimination matrix that removes the 2,1 entry in A or 1st entry in 2nd row.
(2 = row, 1 = column)

here

$$\mathbb{E}_{21} = \begin{bmatrix} 1 & 0 \\ -\frac{2}{-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

($-\frac{2}{-1} = -\ell_{21}$)

Let's see how this works:

E3ap2

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$\mathbb{E}_{21}$ $\quad \mathbb{A} \quad \vec{x} \quad \mathbb{E}_{21} \quad \vec{b}$

↑ premultiply both sides

$$\begin{bmatrix} -1 & 3 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

$\mathbb{U} \quad \vec{x} \quad = \vec{c}$

Anatomy of $\mathbb{E}_{21}$:

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

keep copy of first row ($R_1' = R_1$)

add 2x first row to second row to make new second row

$$R_2' = R_2 - \ell_{21} R_1$$

(ℓ_{21} ↖ 2)

3x3 example:

We need $\mathbb{E}_{21}$ ✓, $\mathbb{E}_{31}$, & $\mathbb{E}_{32}$
(ℓ_{21}), (ℓ_{31}), (ℓ_{32})

ex

$$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 1 \\ 3 & 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 4 \end{bmatrix}$$

Row op | Eliminat matrix

$$R_2' = 1 \cdot R_2 - \left(\frac{2}{1}\right) R_1$$

(ℓ_{21} = 2, D_1 = 1)

$$\mathbb{E}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

← mostly indentity matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$-2 = -\ell_{21}$

$\ell_{31} = 3$

$$R_3' = 1 \cdot R_3 - \left(\frac{3}{1}\right) R_1$$

(D_2 = 1)

$$\mathbb{E}_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$-3 = -\ell_{31}$

Must use elimination matrices to get to $\mathbb{E}_{32}$

E3ap3

$$\begin{bmatrix}1&0&0\\-2&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&-1\\2&3&1\\3&6&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&0&0\\-2&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}-2\\2\\4\end{bmatrix}$$

$\mathbb{E}_{21}$ (under the first matrix, written $\mathbb{E}_2$), $\mathbb{A}$, $\vec{x}$, $\mathbb{E}_{21}$, $\vec{b}$

$$\Rightarrow \begin{bmatrix}1&0&-1\\0&3&3\\3&6&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-2\\6\\4\end{bmatrix}$$

next: premultiply by $\mathbb{E}_{31}=\begin{bmatrix}1&0&0\\0&1&0\\-3&0&1\end{bmatrix}$

$$\Rightarrow \begin{bmatrix}1&0&-1\\0&3&3\\0&6&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-2\\6\\10\end{bmatrix}$$

Important: can now see next row op

$$R_3' = R_3 - \left(\frac{6}{3}\right)R_2 \iff \mathbb{E}_{32}=\begin{bmatrix}1&0&0\\0&1&0\\0&-2&1\end{bmatrix}$$

$\ell_{21}=2$; D_2; $-\ell_{32}$

As before
Premultiply by elimination matrix $\mathbb{E}_{32}$

LHS:

$$\mathbb{E}_{32}\mathbb{E}_{31}\mathbb{E}_{21}\mathbb{A} = \mathbb{U} =$$

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

$\mathbb{U}$ ← pivots

$D_1 = 1$
$D_2 = 3$
$D_3 = -1$

RHS

$$\mathbb{E}_{32}\mathbb{E}_{31}\mathbb{E}_{21}\vec{b}$$

$$= \begin{bmatrix}-2\\6\\-2\end{bmatrix}$$

To find solution, now use back substitution

E3ap4

Note: $\mathbb{E}_{ij}$ are always $m \times m$ lower triangular matrices (0's above main diagonal).

Sometimes row swaps are necessary.

ex: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}$

$\mathbb{P}_{12}$

ex: 3×3 that swaps rows 2 & 3.

$$\mathbb{P}_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

keep: $R_1' = R_1$

$R_2' = R_3$

$R_3' = R_2$

Usually, do row swaps first

3x3 example

$$\mathbb{U} = \mathbb{E}_{32}\mathbb{E}_{31}\mathbb{E}_{21}\mathbb{P}\mathbb{A}$$

↑ row swaps (at $\mathbb{P}$)

$\parallel$

$$\vec{c} = \mathbb{E}_{32}\mathbb{E}_{31}\mathbb{E}_{21}\mathbb{P}\vec{b}$$

Menu:
- Matrix operations
- How to add, scale, and multiply
- The Sneakiness of Matrix multiplication

E4ap1

(1) Scalar multiplication:

$$3\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3\cdot 2 & 3\cdot 1 \\ 3\cdot(-1) & 3\cdot 3 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ -3 & 9 \end{bmatrix}$$

$\mathbb{B} = c\mathbb{A}$

$b_{ij} = ca_{ij}$

Notation: write ij th entry of $\mathbb{A}$ ($m \times n$) as a_{ij} (row index $1 \le i \le m$; column index $1 \le j \le n$)

$\mathbb{B}$ as b_{ij}

Sometimes: $\mathbb{A} = [a_{ij}]$

(2) Addition:

$\mathbb{A} + \mathbb{B}$ is only possible if $\mathbb{A}$ & $\mathbb{B}$ are the same shape

ex.

$$\underset{3\times 2}{\overset{\mathbb{A}}{\begin{bmatrix} 2 & 3 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}}} + \underset{3\times 2}{\overset{\mathbb{B}}{\begin{bmatrix} 1 & -1 \\ 2 & -1 \\ 0 & 1 \end{bmatrix}}} = \underset{3\times 2}{\overset{\mathbb{C}}{\begin{bmatrix} 3 & 2 \\ 3 & -1 \\ -1 & 2 \end{bmatrix}}}$$

entrywise addition

$$c_{ij} = a_{ij} + b_{ij}$$

(3) Multiplication:

$\underset{m\times k}{\mathbb{A}}\ \underset{k\times n}{\mathbb{B}}$ is only possible if inner dimensions match

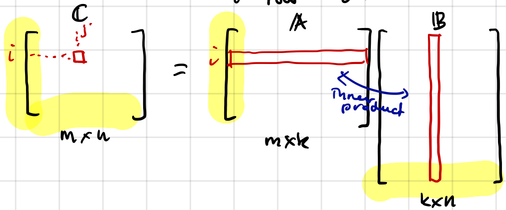
$$\underset{m\times n}{\mathbb{C}} = \underset{m\times k}{\mathbb{A}}\ \underset{k\times n}{\mathbb{B}}$$

Defn:

* c_{ij}, the entry for $\mathbb{C}$ in the i^{th} row and j^{th} column is the dot (inner) product of the i^{th} row of $\mathbb{A}$ and the j^{th} row of $\mathbb{B}$

* $$c_{ij} = \sum_{\ell=1}^{k} a_{i\ell} b_{\ell j}$$

($a_{i\ell}$: i^{th} row; $b_{\ell j}$: j^{th} column)



Rules matrix operations are pretty normal...

$$\mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}$$ commutative law for addition

$$\mathbb{A}\mathbb{B}\mathbb{C} = (\mathbb{A}\mathbb{B})\mathbb{C} = \mathbb{A}(\mathbb{B}\mathbb{C})$$

One bananapoints exception:

> $\mathbb{A}\mathbb{B}$ most often does not equal $\mathbb{B}\mathbb{A}$!!!!

~~Two~~ Three problems

(1) $\mathbb{A}\ \mathbb{B}$ ($m \times k$) ($k \times n$) ✓

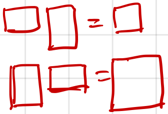
$\mathbb{B}\ \mathbb{A}$ ($k \times n$) ($m \times k$)

if $n \neq m$, $\mathbb{B}\mathbb{A}$ does not make sense

(2) If $n = m$, products are both ok.

$\mathbb{A}\ \mathbb{B}$: ($m \times k$) ($k \times m$) → $m \times m$

$\mathbb{B}\ \mathbb{A}$: ($k \times m$) ($m \times k$) → $k \times k$



if $k \neq m$, no good either

(3) So $m = n = k$ is required for us to even have a chance that $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$

Observe: Only possible for $n \times n$ square matrices

Even then, $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$ often

ex/

$$\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 6 & 5 \end{bmatrix}$$

$\neq$

$$\begin{bmatrix} 0 & -1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 1 & 10 \end{bmatrix}$$

If $\underset{n\times n}{\mathbb{A}}\ \underset{n\times n}{\mathbb{B}} = \underset{n\times n}{\mathbb{B}}\ \underset{n\times n}{\mathbb{A}}$, we get very excited and say $\mathbb{A}$ & $\mathbb{B}$ commute

↑ special. spesh

> **Warning:** Never slide matrices around in products and always be careful with order. //

Menu:

- Wizard-level matrix multiplication skills
- Inner and outer products (dot)
- $A\vec{x}$, $\vec{y}^T B$, AB
- Block multiplication in general

E4b p1

from before:

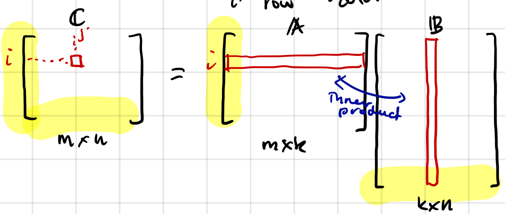
$$\mathbb{C} = A B$$

$m \times n \quad m \times k \quad k \times n$

Defn:

* c_{ij}, the entry for $\mathbb{C}$ in the i^{th} row and j^{th} column is the dot (inner) product of the i^{th} row of A and the j^{th} row of B

* $c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}$ (a_{il}: i^{th} row; b_{lj}: j^{th} column)



ex 1

$$\overset{A}{\begin{bmatrix} 3 & 0 & 2 \\ 1 & -2 & 2 \end{bmatrix}} \overset{B}{\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}} = \begin{bmatrix} -3 & 4 \\ -5 & 2 \end{bmatrix}$$

2×3 3×2 2×2

(c_{11}, c_{12}, c_{21}, c_{22} label the entries of the result)

$$c_{11} = \begin{bmatrix} 3 & 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

(1st row of A; 1st col of B)

$$c_{12} = \begin{bmatrix} 3 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

$$c_{21} = \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = -5$$

$$c_{22} = \begin{bmatrix} 1 & -2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = 2$$

ex2

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

1×3 ✓ $\qquad 1\times1$

inner product

ex3

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

3×1 ✓ $\qquad 3\times1 \qquad 1\times3 \qquad 3\times3$

amazingly important construction

outer product

later: see this is a rank $r=1$ matrix.

See

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

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

ex4 block multiplication: E4bp2

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

row of 2×1's

column of 1×2's

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

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

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

ex5

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

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

$2\times1 \quad 1\times2 \quad 2\times2$

E4bp3

$A\vec{x}$: ($m\times n$, $n\times 1$) — A: m rows, n columns; $\vec{x}$: n rows.

See this as the columns of A being combined with weights in vector $\vec{x}$:

$$A\vec{x} = \begin{bmatrix} | & | & & | \\ \vec{a}_{*1} & \vec{a}_{*2} & \cdots & \vec{a}_{*n} \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

↖ column vectors inside A

$*$ = run over all indices

$\vec{a}_{*1}$ ← first column vector

$$= x_1 \begin{bmatrix} | \\ \vec{a}_{*1} \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ \vec{a}_{*2} \\ | \end{bmatrix} + \dots + x_n \begin{bmatrix} | \\ \vec{a}_{*n} \\ | \end{bmatrix}$$

ex

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

(A is 2×3, vector is 3×1; scalings: (-1), (2), (0))

$$= \begin{bmatrix} -3 \\ -5 \end{bmatrix}_{2\times 1}$$

$\vec{y}^T A$: ($1\times m$, $m\times n$) "transpose"

$$\underbrace{[\,y_1 \; y_2 \cdots y_m\,]}_{\vec{y}^T,\ 1\times m} \underbrace{\begin{bmatrix} - & \vec{a}_{1*} & - \\ - & \vec{a}_{2*} & - \\ & \vdots & \\ - & \vec{a}_{m*} & - \end{bmatrix}}_{A,\ m\times n}$$

n.b. $\vec{y}$ is $m\times 1$ ← column vector

↖ row vectors inside A

$\vec{a}_{1*}$ ← first row vector in a

$$\vec{y}^T A = y_1[-\ \vec{a}_{1*}\ -] + y_2[-\ \vec{a}_{2*}\ -] + \dots + y_m[-\ \vec{a}_{m*}\ -]$$

See this as the rows of A being combined with weights in vector $\vec{y}^T$.

ex

$$\underbrace{[\,3 \;\; 0 \;\; 2\,]}_{1\times 3} \underbrace{\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}}_{3\times 2} = \underbrace{[-3 \;\; 4]}_{1\times 2}$$

$$= (3)[-1 \;\; 0] + (0)[2 \;\; 1] + (2)[0 \;\; 2] = [-3 \;\; 4]$$

E4bp4

$\mathbb{C} = \mathbb{A}\mathbb{B}$

$\mathbb{C}$: $m \times k$, $\mathbb{A}$: $m \times n$, $\mathbb{B}$: $n \times k$

2 views

break $\mathbb{B}$ into columns →

$$\mathbb{A}\begin{bmatrix} | & | & & | \\ \vec{b}_{*1} & \vec{b}_{*2} & \cdots & \vec{b}_{*k} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ (\mathbb{A}\vec{b}_{*1}) & (\mathbb{A}\vec{b}_{*2}) & \cdots & (\mathbb{A}\vec{b}_{*k}) \\ | & | & & | \end{bmatrix}$$

column vector: $\mathbb{A}\vec{b}_{*1}$ is ($m \times n$)($n \times 1$)

$\mathbb{C}$'s columns are made up of $\mathbb{A}$'s columns

break $\mathbb{A}$ into rows ↓

$$\begin{bmatrix} — \vec{a}_{1*} — \\ — \vec{a}_{2*} — \\ \vdots \\ — \vec{a}_{m*} — \end{bmatrix}\mathbb{B} = \begin{bmatrix} — (\vec{a}_{1*}\mathbb{B}) — \\ — (\vec{a}_{2*}\mathbb{B}) — \\ \vdots \\ — (\vec{a}_{m*}\mathbb{B}) — \end{bmatrix}$$

$\vec{a}_{m*}\mathbb{B}$: ($1 \times n$)($n \times k$) → $1 \times k$

$\mathbb{C}$'s rows are made up of $\mathbb{B}$'s rows

ex

$$\overset{\mathbb{A}}{\begin{bmatrix} 3 & 0 & 2 \\ 1 & -2 & 2 \end{bmatrix}}\overset{\mathbb{B}}{\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}}$$

2×3, 3×2

Columns view:

$$= \begin{bmatrix} \overset{\mathbb{A}}{\begin{bmatrix} 3 & 0 & 2 \\ 1 & -2 & 2 \end{bmatrix}}\overset{\vec{b}_{*1}}{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}} & \overset{\mathbb{A}}{\begin{bmatrix} 3 & 0 & 2 \\ 1 & -2 & 2 \end{bmatrix}}\overset{\vec{b}_{*2}}{\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}} \end{bmatrix}$$

2×3, 3×1

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

2×1

Rows view:

$$= \begin{bmatrix} \overset{\vec{a}_{1*}}{\begin{bmatrix} 3 & 0 & 2 \end{bmatrix}}\overset{\mathbb{B}}{\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix}} \\ \overset{\vec{a}_{2*}}{\begin{bmatrix} 1 & -2 & 2 \end{bmatrix}}\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \end{bmatrix}$$

1×3, $\mathbb{B}$: 3×2

$$= \begin{bmatrix} [-3 \quad 4] \\ [-5 \quad 2] \end{bmatrix}$$

1×2, 1×2

$$= \begin{bmatrix} -3 & 4 \\ -5 & 2 \end{bmatrix} \checkmark$$

Matrix Inverses ← delicious

Square Matrices Only

Menu:

- General goodness of inverses
- Identity matrix
- Solving $A\vec{x} = \vec{b}$
- Advanced goodness
- How to w. example

Ex

$$\underbrace{\frac{1}{8}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}}_{A^{-1}} \underbrace{\begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}}_{A} = \frac{1}{8}\begin{bmatrix} 4\times3+2\times-2 & 4\times-2+2\times4 \\ 2\times3+3\times-2 & 2\times-2+3\times4 \end{bmatrix}$$

2×2   2×2

(8, 0, 0, 8)

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}$$

↑ "1" for 2×2 matrices

Identity matrix $\mathbb{I}_n$

↑ n×n

leaves matrices unchanged under multiplication

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

⇑ $R_1' = R_1$, $R_2' = R_2$

$$\mathbb{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbb{I}_4 = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

lots of zeros ↓

E5ap1

Given A (n×n), if A^{-1} exists then

$$A^{-1}A = AA^{-1} = \mathbb{I}$$

↖ order doesn't matter

Selection of Big Deals for A^{-1}:

- Even knowing if A^{-1} exists is valuable
  ↖ may not have actual form.
- If A^{-1} exists then:

$$A^{-1}A\vec{x} = A^{-1}\vec{b}$$

↙ premultiply both sides

(n×n, n×1, n; ↑ m=n)

$$\Rightarrow \mathbb{I}\vec{x} = A^{-1}\vec{b}$$

(n×n, n×1)

$$\Rightarrow \vec{x} = A^{-1}\vec{b}$$

Done!!

But wait:

- Many systems have rectangular A's ($m \neq n$)
- A^{-1} may not exist
- Even if A^{-1} exists, computing A^{-1} is very hard work computationally (grows badly with n)

E5ap2

- if A^{-1} exists, then $A\vec{x} = \vec{b}$ has only one solution, always. (for all $\vec{b}$)

Simply: $\vec{x} = A^{-1}\vec{b}$

- if A^{-1} does not exist then we may have 0 or ∞ly many solutions
  ↖ more later

- If $\exists\, \vec{x} \neq \vec{0}$ (there exists an $\vec{x} \neq \vec{0}$) such that $A\vec{x} = \vec{0}$ " $\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}$ "

  ↖ "A maps $\vec{x}$ to $\vec{0}$"
  "A crushes $\vec{x}$."

then A^{-1} does not exist

Proof

$\underbrace{A\vec{x} = \vec{0}}_{\text{true}} \Rightarrow \underbrace{A^{-1}A}_{n\times n\ \ n\times n}\vec{x}_{n\times 1} = A^{-1}_{n\times n}\vec{0}_{n\times 1}$

$\Rightarrow I\vec{x} = \vec{0}$

$\vec{x} = \vec{0}$

contradiction!

$\Rightarrow A^{-1}$ cannot exist

- Foreshadowing: if $A\vec{x} = \vec{0}$ we say $\vec{x} \in N(A)$
  ↑ null space of A

- $(\underset{n\times n}{A}\underset{n\times n}{B})^{-1} = B^{-1}A^{-1}$

See $B^{-1}\underbrace{A^{-1}(A}B) = B^{-1}(IB) = B^{-1}B$ "B"

$(A\underbrace{B)B^{-1}}A^{-1} = \underset{A}{(AI)}A^{-1} = AA^{-1}$
$= I$

$(ABCD)^{-1}$
$= D^{-1}C^{-1}B^{-1}A^{-1}$

- If we have $\underset{n\times n}{A}, \underset{n\times n}{Z_L}, \underset{n\times n}{Z_R}$ such that

$AZ_R = I$ (↙ right inverse) & $Z_L A = I$ (↙ left inverse)

then $A^{-1} = Z_R = Z_L$

Reason

$\underbrace{Z_L}_{\text{premultiply}}\underbrace{(AZ_R)}_{I} = \underbrace{Z_L(I)}$

$I Z_R = Z_L$

E5bp1

Using Gauss-Jordan Elimination to find A^{-1}

- general story (it's $A\vec{x}=\vec{b}$ again!)
- example

Game: given $\underset{n\times n}{A}$, find $\overset{n\times n}{A^{-1}}$

$$A^{-1}A = AA^{-1} = I$$

$A\vec{x}=\vec{b}$ ish ($m\times n$, $n\times 1$, $m\times 1$)

Consider: $\overset{n\times n}{A}\ \overset{n\times n}{Z} = \overset{n\times n}{I}$

2x2 general ex

$$\underset{2\times 2}{A}\begin{bmatrix} | & | \\ \vec{z}_1 & \vec{z}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(2×1)

⇓ wrangling

$$\begin{bmatrix} | & | \\ \underset{2\times 2}{A}\underset{2\times 1}{\vec{z}_1} & \underset{2\times 2}{A}\underset{2\times 1}{\vec{z}_2} \\ | & | \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix}$$

(2×1, 2×1; $\vec{e}_1$, $\vec{e}_2$)

$\Rightarrow$ Solve $A\vec{z}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ & $A\vec{z}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

↖ $A\vec{x}=\vec{b}$!!

Note: we would make A become U w. row reduction for both equation.

Do all at once with a super augmented matrix:

$$\left[\ A \ \middle|\ I\ \right]$$

$n\times n$, $n\times n$, $n\times 2n$

#awesome

Use row ops to turn A into U then I

I will change into A^{-1}

- actually: finding right inverse of A; later we show it's the true inverse
- only works if A has n pivots

Example:

$$\mathbb{A} = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}$$

$$[\mathbb{A} \mid \mathbb{I}] = \left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ -2 & 4 & 0 & 1 \end{array}\right]$$

($3 \leftarrow D_1$)

$R_2' = R_2 - \left(\frac{-2}{3}\right) R_1$ ($3 \leftarrow D_1$; $\ell_{21} = \frac{-2}{3}$)

$$\left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ 0 & 8/3 & 2/3 & 1 \end{array}\right]$$

($3 \leftarrow D_1$, $-2 \leftarrow \mathbb{U}$, $8/3 \leftarrow D_2$; $4 - \frac{4}{3}$, $12/3$)

$R_1' = R_1 - \left(\frac{-2}{8/3}\right) R_2$ ($8/3 \leftarrow D_2$; $+\frac{3}{4}$)

$$\left[\begin{array}{cc|cc} 3 & 0 & 3/2 & 3/4 \\ 0 & 8/3 & 2/3 & 1 \end{array}\right]$$

($1 + \frac{2}{3} \cdot \frac{3}{4}$)

E5bp2

divide by pivots

$R_1' = \frac{1}{3} R_1$ ($3 \leftarrow D_1$)

$R_2' = \frac{1}{8/3} R_2$ ($8/3 \leftarrow D_2$)

$$\left[\begin{array}{cc|cc} 1 & 0 & 1/2 & 1/4 \\ 0 & 1 & 1/4 & 3/8 \end{array}\right]$$

$\mathbb{I}$ | $\mathbb{A}^{-1}$

tidying up

$$\mathbb{A}^{-1} = \frac{1}{8} \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

see E5ap1.

turns out this is a very special number for $\mathbb{A}$

notation
- Determinant of $\mathbb{A}$
- Det($\mathbb{A}$)
- $|\mathbb{A}|$

More later!!

E5bp3

3x3 plan

$$\left[\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 1 & -1 & 7 & 0 & 1 & 0 \\ 13 & 2 & 17 & 0 & 0 & 1 \end{array}\right]$$

$A \qquad\qquad I$

order of elimination: ① 1, ② 13, ③ 2, ④ 7, ⑤ 4, ⑥ 3

Hidden Secrets of Inverses:

- A^{-1} and elimination matrices
- Inverses of elimination matrices
- Missing pivots → A^{-1} does not exist

Pratchett upon learning more about inverses →

- Curious things about columns...

$A = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}$ ↑ aside $A = A^T$ ↖ transpose

$A^{-1} = \frac{1}{8}\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$ ↖ $(A^{-1})^T = A^{-1}$

Row reduction
⇒ Elimination matrices

for Solving $[A \mid I] \Leftrightarrow AZ = I$

row op 1

$E_{21} = \begin{bmatrix} 1 & 0 \\ 2/3 & 1 \end{bmatrix}$ (row 2, row 1; $2/3$ ← $-\ell_{21}$)

row op 2 — multiplier ↓

$E_{12} = \begin{bmatrix} 1 & 3/4 \\ 0 & 1 \end{bmatrix}$

ESc p1

× pivot matrix

$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 8/3 \end{bmatrix}$

$D^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 3/8 \end{bmatrix}$ ← undo each other ↑

$\left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ -4 & 4 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 1/2 & 1/4 \\ 0 & 1 & 1/4 & 3/8 \end{array}\right]$

(−4 → −2) [note: transcribed incorrectly in video]

$\underbrace{D^{-1}E_{12}E_{21}}_{A^{-1}} A Z = \underbrace{D^{-1}E_{12}E_{21}}_{A^{-1}} I$

$I Z_1 = A^{-1}$

↑ found by row operations

↖ made by E_{ij} & D matrices.

Big Deal:
See A^{-1} is a product of E_{ij}'s, D^{-1}, P
(↑ pivots; ↖ permutation for row swaps)

⇑ Huge: Demonstrates that A^{-1} is a left and right inverse

Next: Elimination matrices have simple inverses.

$$\mathbb{E}_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \longleftrightarrow R_3' = R_3 - 2R_1$$

(-2 is $-\ell_{31}$)

undo with $R_3' = R_3 + 2R_1$

$$\Rightarrow \mathbb{E}_{31}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +2 & 0 & 1 \end{bmatrix}$$

($+2$ is $+\ell_{31}$) swap sign

check $\mathbb{E}_{31}\mathbb{E}_{31}^{-1} = \mathbb{E}_{31}^{-1}\mathbb{E}_{31} = \mathbb{I}$.

In general flip sign of (one) off diagonal element to turn $\mathbb{E}_{ij}$ into $\mathbb{E}_{ij}^{-1}$

Monks made us do this... Sneaky plan.

Permutation matrices:

$$\mathbb{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \leftarrow \begin{array}{l} R_1' = R_1 \\ R_2' = R_3 \\ R_3' = R_2 \end{array}$$

here: $\mathbb{P}^{-1} = \mathbb{P}$

But in general $\mathbb{P}^{-1} = \mathbb{P}^T$

E5cp2

Missing pivots.

What if $[\mathbb{A} \mid \mathbb{I}] \to$ one or more rows of zeros on left? i.e., missing pivots?

from before: if $\vec{x} \neq \vec{0}$ solves $\mathbb{A}\vec{x} = \vec{0}$ then $\mathbb{A}^{-1}$ cannot exist

so $[\mathbb{A} \mid \vec{0}] \sim$ (row ops) $[\mathbb{U} \mid \vec{0}]$

($\vec{0}$ on the left is $\vec{b}$; $\mathbb{U}$ is upper triangular; $\vec{0}$ on the right is $\vec{c}$)

Row of 0's in $\mathbb{U} \to \infty$ly many solns
$\to \mathbb{A}\vec{x} = \vec{0}$ is solved by $\vec{x} \neq \vec{0}$
$\to \mathbb{A}^{-1}$ does not exist

ex

$$\left[\begin{array}{cc|c} 3 & 2 & 0 \\ 6 & 4 & 0 \end{array}\right] \underset{R_2' = R_2 - \left(\frac{6}{3}\right)R_1}{\sim} \left[\begin{array}{cc|c} 3 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

(the circled 0 in row 2: missing pivot)

Upshot:

$\mathbb{A}^{-1}$ exists ($\mathbb{A}$ is $n \times n$)

$\Longleftrightarrow \mathbb{A}$ has n pivots

$\Longleftrightarrow \mathbb{A}\vec{x} = \vec{0}$ has only $\vec{x} = \vec{0}$ as a solution

$\Longleftrightarrow \det(\mathbb{A}) \neq 0$

($\det$: later; parallelograms will be involved)

If $\mathbb{A}_{2\times 3}$ has column 1 + column 2 = column 3

show $\mathbb{A}^{-1}$ does not exist.. (weird)

(a) See $\mathbb{A}\begin{bmatrix}1\\1\\-1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$ ⇐ non-zero solut to $\mathbb{A}\vec{x}=\vec{0}$

⇓

$\mathbb{A}^{-1}$ does not exist ⇒ pivot missing.

column picture: $1\vec{a}_1 + 1\cdot\vec{a}_2 - 1\cdot\vec{a}_3 = \begin{bmatrix}0\\0\\0\end{bmatrix}$ (3×1)

(b) Another aspect:

Row operations destroy rows But Column relationships are unchanged.

row reduct → $\begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & \xi \end{bmatrix}$ ← must be 0 → 3rd pivot is missing

$C_1 + C_2 = C_3 \Rightarrow \xi = 0 + 0 = 0.$

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

$\vec{a}_1 \quad \vec{a}_2 \quad \vec{a}_3 = \vec{a}_1 + \vec{a}_2$

col 1 + col 2 = col 3

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

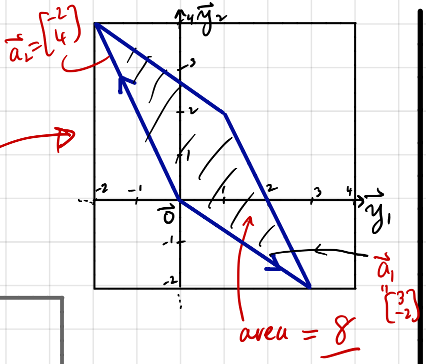
columns are "linearly dependent"

⇒ connects to $\mathbb{A}^{-1}$ not existing

E5dp1

Foreshadowing:

Determinant matrix.



from p5bp2

$$[A \mid I] = \left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ -2 & 4 & 0 & 1 \end{array}\right]$$

$R_2' = R_2 - \left(\frac{-2}{3}\right) R_1$, $\ell_{21} = -\frac{2}{3}$

$$\left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ 0 & 8/3 & 2/3 & 1 \end{array}\right]$$

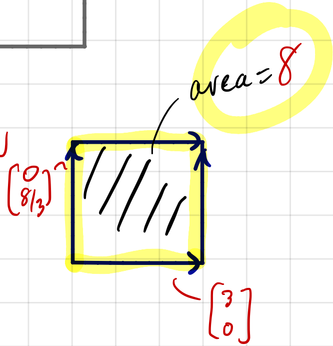
($D_1 = 3$, $U = -2$, $D_2 = 8/3$; $4 - \frac{4}{3} = \frac{12}{3} - \frac{4}{3}$)

$R_1' = R_1 - \left(\frac{-2}{8/3}\right) R_2$, $+\frac{3}{4}$

$$\left[\begin{array}{cc|cc} 3 & 0 & 3/2 & 3/4 \\ 0 & 8/3 & 2/3 & 1 \end{array}\right]$$

($1 + \frac{2}{3} \cdot \frac{3}{4} = 3/2$)

$$|A| = 3 \times 4 - (-2)(-2) = 12 - 4 = 8$$



Triangle × Triangle = Rectangle

Menu:

- Our first factorization: $\mathbb{L}\,\mathcal{U}$ ← t-shirt for each factorization
- Method first
- The l_{ij}'s serve us well (as promised by mysterious monks)

Ex

$$x_1 \begin{bmatrix}1\\1\\1\end{bmatrix} + x_2 \begin{bmatrix}1\\2\\3\end{bmatrix} + x_3 \begin{bmatrix}1\\3\\6\end{bmatrix} = \begin{bmatrix}5\\7\\11\end{bmatrix}$$

$\vec{a}_1 \quad \vec{a}_2 \quad \vec{a}_3 \quad \vec{b}$

$A\vec{x} = \vec{b}$ (3×3, $m\times n$; 3×1; 3×1)

Normal plan:

set up $[A \mid \vec{b}] \xrightarrow{\text{row ops}} [\mathcal{U} \mid \vec{c}] \xrightarrow{\text{back sub}} \vec{x}$

Now: focus on reducing A by itself ↖ very good if $\vec{b}$ is changed.

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

$R_2' = R_2 - \left(\frac{1}{1}\right)R_1$, $l_{21} = 1$

$R_3' = R_2 - \left(\frac{1}{1}\right)R_1$, $l_{31} = 1$

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

$R_3' = R_3 - \left(\frac{2}{1}\right)R_2$, $l_{32} = 2$

$$\begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\\0 & 0 & 1\end{bmatrix} \leftarrow \mathcal{U}$$

$D_1 = 1$, $D_2 = 1$, $D_3 = 1$ ← not exciting

Elimination matrix story

E6ap1

$A \to \mathcal{U} = E_{32} E_{31} E_{21} A$

powerful encoding of our row operations

Monks whisper: "invert E_{ij}'s"

RHS, LHS

$$E_{21}^{-1} E_{31}^{-1} E_{32}^{-1} E_{32} E_{31} E_{21} A = E_{21}^{-1} E_{31}^{-1} E_{32}^{-1} \mathcal{U}$$

($E_{32}^{-1}E_{32} = I$; $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}E_{32}E_{31}E_{21} = I$)

$$\Rightarrow A = E_{21}^{-1} E_{31}^{-1} E_{32}^{-1} \mathcal{U}$$

← upper triangular

$E_{21}^{-1} E_{31}^{-1} E_{32}^{-1}$: row operations in reverse

Tells us how to combine rows of $\mathcal{U}$ (simple-ish) to make rows of A

Amazingly: we'll see

$$\underset{m\times n}{A} = \underset{\text{square } m\times m}{\mathbb{L}}\ \underset{m\times n}{\mathcal{U}}$$

lower triangular:

$$\mathbb{L} = \begin{bmatrix}1 & 0 & 0\\ +l_{21} & 1 & 0\\ +l_{31} & +l_{32} & 1\end{bmatrix}_{m\times m}$$

upper triangular:

$$\mathcal{U} = \begin{bmatrix}D_1 & * & *\\ 0 & D_2 & *\\ 0 & 0 & D_3\end{bmatrix}$$

Now

$$\mathbb{A} = \underbrace{\mathbb{E}_{21}^{-1}\mathbb{E}_{31}^{-1}\mathbb{E}_{32}^{-1}}_{\text{know these are simple}}\mathbb{U}$$

($\mathbb{E}_{21}^{-1}\mathbb{E}_{31}^{-1}\mathbb{E}_{32}^{-1} = \mathbb{L}$)

ex recall

$$\mathbb{E}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ +l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$R_2' = R_2 - l_{21}R_1$ $\qquad$ $R_2' = R_2 + l_{21}R_1$

Big Deals:

(1) $\mathbb{E}_{ij}^{-1}$ is $\mathbb{E}_{ij}$ with single off diagonal element flipped in sign

(2) $\mathbb{E}_{ij}$'s & $\mathbb{E}_{ij}^{-1}$'s are all lower triangular

(3) $\mathbb{E}_{ij}$ is $\mathbb{I}$ with $-l_{ij}$ replacing 0 in ij position

(4) Remarkably:

$$\mathbb{E}_{21}^{-1}\mathbb{E}_{31}^{-1}\mathbb{E}_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ +l_{21} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +l_{31} & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & +l_{32} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{32} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} = \mathbb{L}$$

$\mathbb{L}$ always has 1's on the diagonal.

Back to example:

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

LE6ap2

So:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} = \mathbb{L}\,\mathbb{U}$$

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

(ℓ_{21}, ℓ_{31}, ℓ_{32} marked in $\mathbb{L}$; pivots D_1, D_2, D_3 marked in $\mathbb{U}$)

Now solve $A\vec{x} = \vec{b}$ if $\vec{b} = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix}$ by solving two (easy) triangular systems

$$\mathbb{L}\,\underbrace{\mathbb{U}\,\vec{x}}_{\vec{c}} = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix} \leftarrow \vec{b}.$$

From before $\mathbb{U}\vec{x} = \vec{c}$

$\rightarrow$ $\mathbb{L}\,\vec{c} = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix}$ ($\mathbb{L}$ ✓, $\vec{c}$?, $\vec{b}$ ✓)

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 11 \end{bmatrix} \leftarrow \text{Solve by forward substitution}$$

R1: $c_1 = 5$

R2: $c_1 + c_2 = 7 \Rightarrow c_2 = 2$ ($c_1 = 5$)

R3: $c_1 + 2c_2 + c_3 = 11 \Rightarrow c_3 = 2$ ($c_1 = 5$, $c_2 = 2$)

$$\vec{c} = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$$

Now solve $\mathbb{U}\vec{x} = \vec{c}$ with back substitution ($\mathbb{U}$ ✓, $\vec{x}$?, $\vec{c}$ ✓) E6ap3

$$\uparrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$$

R_3: $x_3 = 2$

R_2: $x_2 + 2x_3 = 2 \rightarrow x_2 = -2$ ($x_3 = 2$)

$x_1 + x_2 + x_3 = 5 \rightarrow x_1 = 5$ ($x_2 = -2$, $x_3 = 2$)

$$\vec{x} = \begin{bmatrix} 5 \\ -2 \\ 2 \end{bmatrix}$$

done

Big deal: Swap $\vec{b}$, easy to solve. Row reduction is done once and is encoded in $\mathbb{L}$ & $\mathbb{U}$.

Extra pieces:

Our A was special b/c $A = A^T$
$\Rightarrow$ $\mathbb{L}$ and $\mathbb{U}$ are transposes of each other

But only b/c $D_1 = D_2 = D_3 = 1$

LE6ap4

Also very useful:
Separate out pivots

$$\underbrace{\begin{bmatrix} 2 & 1 & 3 \\ 2 & 0 & 0 \\ 4 & 3 & 13 \end{bmatrix}}_{\mathbb{A}\ (m\times n)} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}}_{\mathbb{L}\ (m\times m \text{ square})} \underbrace{\begin{bmatrix} 2 & 1 & 3 \\ 0 & -1 & -3 \\ 0 & 0 & 4 \end{bmatrix}}_{\mathbb{U}}$$

(diagonal of $\mathbb{L}$: always 1s)

$l_{21} = 1$
$l_{31} = 2$
$l_{32} = -1$

$D_1 = 2,\ D_2 = -1,\ D_3 = 4$

Alternate factorization ($\mathbb{U}$ is different!)

$$\mathbb{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}}_{m\times m} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix}}_{m\times m} \begin{bmatrix} 1 & 1/2 & 3/2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \mathbb{L}\ \mathbb{D}\ \mathbb{U}$$

* If $\mathbb{A} = \mathbb{A}^T$ then $\mathbb{A} = \mathbb{L}\mathbb{D}\underbrace{\mathbb{L}^T}_{\mathbb{U}}$

(↑ next)

Bad notation:
$\mathbb{U}$ in $\mathbb{L}\mathbb{U}$ $\neq$ $\mathbb{U}$ in $\mathbb{L}\mathbb{D}\mathbb{U}$

Must state which form we're using from the start.

Last thing: **Row Swaps** (do at the start)

$$\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{U}$$
or
$$\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{D}\mathbb{U}$$

possible for every matrix $\mathbb{A}$

Amazing!!

E6bp1

Why $\mathbb{L}\mathbb{U}$ works:

mad!

Claim: $\mathbb{E}_{ij}$ matrices always combine to produce a lower triangular matrix with l_{ij}'s in the right spots & 1's along the main diagonal

$$\mathbb{L} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$$

Why does $\mathbb{E}_{21}^{-1}\mathbb{E}_{31}^{-1}\mathbb{E}_{32}^{-1}$ work so simply?

Reason:

As we uncover $\mathbb{U}$ with row operations, we only use rows of $\mathbb{U}$ to modify lower rows of $\mathbb{A}$.

$\Rightarrow$ $\mathbb{L}$ will be lower triangular

"combining" matrix tells us how to combine $\mathbb{U}$'s rows to produce $\mathbb{A}$.

3x3 example (ignoring row swaps):

Row 1 of $\mathbb{U}$ = Row 1 of $\mathbb{A}$

Row 2 of $\mathbb{U}$ = Row 2 of $\mathbb{A}$ $- l_{21} \times$ Row 1 of $\mathbb{U}$ (Row 1 of $\mathbb{A}$)

Row 3 of $\mathbb{U}$ = Row 3 of $\mathbb{A}$ $- l_{31} \times$ Row 1 of $\mathbb{U}$ $- l_{32} \times$ Row 2 of $\mathbb{U}$

Invert

Row 1 of $\mathbb{A}$ = 1x Row 1 of $\mathbb{U}$

Row 2 of $\mathbb{A}$ = 1x Row 2 of $\mathbb{U}$ $+ l_{21} \times$ Row 1 of $\mathbb{U}$

Row 3 of $\mathbb{A}$ = 1x Row 3 of $\mathbb{U}$ $+ l_{31} \times$ Row 1 of $\mathbb{U}$ $+ l_{32} \times$ Row 2 of $\mathbb{U}$

RHS is simple

Transposes and Symmetric Matrices ↖ ♡

Menu:

- Transposes
- Symmetric matrices
- Properties of peculiar nature

//

Defn: A^T = the transpose of A (A^T: $n \times m$, A: $m \times n$)

= A flipped about the main diagonal

[A^T's columns are the rows of A
& " rows " " columns of A

ex* $\begin{bmatrix} 2 & 7 & 3 \\ -1 & 2 & 4 \end{bmatrix}^T = \begin{bmatrix} 2 & -1 \\ 7 & 2 \\ 3 & 4 \end{bmatrix}$ (2×3, $m \times n$ → 3×2)

ex. $\begin{bmatrix} 3 & 9 \\ 17 & 23 \end{bmatrix}^T = \begin{bmatrix} 3 & 17 \\ 9 & 23 \end{bmatrix}$

Defn again: $(A^T)_{ij} = (A)_{ji}$

*$(A)_{21} = (A^T)_{12} = -1$

LE 7a p1

Big Deal:
The transpose of A will matter ridiculously everywhere and especially for solving $A\vec{x} = \vec{b}$ ← our pal.

//

Defn: if $A = A^T$ (means A must be square)
then we say A is symmetric and we are very happy.
↙
Super, super special matrices

//

Properties:

$(A+B)^T = A^T + B^T$ ✓

$(AB)^T \overset{?}{=} A^T B^T$ ~~(A: $m \times q$, B: $q \times n$, AB: $m \times n$, $(AB)^T$: $n \times m$; A^T: $q \times m$, B^T: $n \times q$)~~ can't be general...

$= B^T A^T$ ← always true
(B^T: $n \times q$, A^T: $q \times m$ → $n \times m$) ← right dimensions.

$$(AB)^T = B^T A^T$$ ← proof later

What about $(A^{-1})^T$?

know
$$A^{-1}A = I = AA^{-1}$$

take transposes:
$$(A^{-1}A)^T = I^T = (AA^{-1})^T$$
$$(A^T)(A^{-1})^T = I = (A^{-1})^T(A)^T$$
$$\Rightarrow (A^T)^{-1} = (A^{-1})^T$$

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

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

So if A is symmetric, so its inverse

E7ap2

Crazily important objects:

$A^T A$ & AA^T

$A^T A$: $n \times m$, $m \times n$ → $n \times n$

AA^T: $m \times n$, $n \times m$ → $m \times m$

Square matrices: A does not have to be square

$$(A^T A)^T = (A)^T (A^T)^T = A^T A$$ (undo)

So $A^T A$ is always symmetric

Check: true for AA^T as well.

ex
$$A = \begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$
3×2

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$
2×2

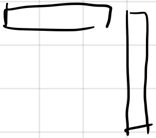
$$AA^T = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$
3×3

#awesome

E7ap3

Inner product:

row vector → $\vec{a}^T \vec{b}$



what happens with?:

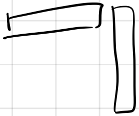
$(\vec{a}^T \vec{b})^T$

$= \vec{b}^T (\vec{a}^T)^T$

$= \vec{b}^T \vec{a}$

1×n n×1

1×1



More advanced inner producting:

$A\vec{x}$ & $\vec{y}$

transformation of $\vec{x}$

ex: $(A\vec{x})^T \vec{y} = \vec{x}^T A^T \vec{y}$

$= \vec{x}^T (A\vec{y})$

Inner product of $\vec{x}$ & $A^T \vec{y}$

Transform $\vec{y}$ with A^T first

More on the Transpose

E7bp1

menu:
- example of $(\mathbb{A}\mathbb{B})^T = \mathbb{B}^T\mathbb{A}^T$
- three different proofs

ex

$$\left(\underbrace{\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}}_{\mathbb{A}}\underbrace{\begin{bmatrix} -1 & 4 \\ 3 & 1 \end{bmatrix}}_{\mathbb{B}}\right)^T = \begin{bmatrix} 7 & 11 \\ -4 & 3 \end{bmatrix}^T = \begin{bmatrix} 7 & -4 \\ 11 & 3 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} -1 & 4 \\ 3 & 1 \end{bmatrix}^T}_{\mathbb{B}^T}\underbrace{\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}^T}_{\mathbb{A}^T}$$

$$= \begin{bmatrix} -1 & 3 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 7 & -4 \\ 11 & 3 \end{bmatrix}$$

$$(\mathbb{A}\mathbb{B})^T_{ij}$$

$$= (\mathbb{A}\mathbb{B})_{ji} = \sum_{k=1}^{q} a_{jk} b_{ki}$$

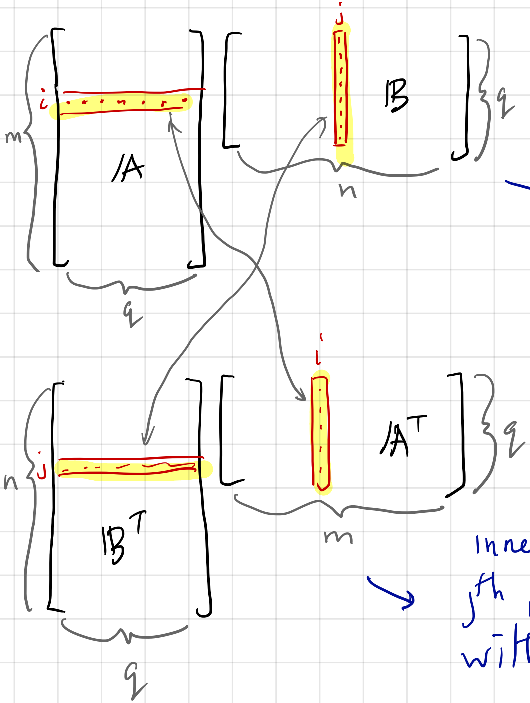
$\mathbb{A}$: $m \times q$, $\mathbb{B}$: $q \times n$

inner product of jth row of $\mathbb{A}$ & ith column of $\mathbb{B}$

$$= \sum_{k=1}^{q} (\mathbb{A}^T)_{kj} (\mathbb{B}^T)_{ik}$$

$$= \sum_{k=1}^{q} (\mathbb{B}^T)_{ik} (\mathbb{A}^T)_{kj}$$

$$= (\mathbb{B}^T\mathbb{A}^T)_{ij}$$



→ inner product of i^{th} row of A with j^{th} column of B

$$= (AB)_{ij}$$

$$= ((AB)^T)_{ji}$$

→ inner product of j^{th} row of B^T with i^{th} column of A^T

$$= (B^T A^T)_{ji}$$

$$(AB)^T = B^T A^T$$

Yet another way:

$$(A\vec{x})^T = \left(x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n\right)^T$$

A: $m\times n$, $\vec{x}$: $n\times 1$ — column picture

$$= x_1\vec{a}_1^T + \dots + x_n\vec{a}_n^T$$

($\vec{a}_n^T$: row vector)

$$= [x_1 \dots x_n]\begin{bmatrix} — \vec{a}_1^T — \\ — \vec{a}_2^T — \\ \vdots \\ — \vec{a}_n^T — \end{bmatrix}$$

$$= \vec{x}^T A^T$$

use here

$$(AB)^T$$

A: $m\times q$, B: $q\times n$

$$= \left(A\begin{bmatrix} | & | & & | \\ \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \\ | & | & & | \end{bmatrix}\right)^T$$

$$= \left(\begin{bmatrix} | & | & & | \\ A\vec{b}_1 & A\vec{b}_2 & \dots & A\vec{b}_n \\ | & | & & | \end{bmatrix}\right)^T$$

E7bp3

$$= \begin{bmatrix} — (A\vec{b}_1)^T — \\ — (A\vec{b}_2)^T — \\ \vdots \\ — (A\vec{b}_n)^T — \end{bmatrix}$$

$$= \begin{bmatrix} — \vec{b}_1^T A^T — \\ — \vec{b}_2^T A^T — \\ \vdots \\ — \vec{b}_n^T A^T — \end{bmatrix}$$

$$= \begin{bmatrix} — \vec{b}_1^T — \\ \vdots \\ — \vec{b}_n^T — \end{bmatrix} A^T$$

$$= B^T A^T$$

Yes!!

"Paging Dr. Vector Spaceman"

Menu:

- Our new plan for $A\vec{x}=\vec{b}$
- Vector spaces, introduction to

The Column picture for $A\vec{x}=\vec{b}$

$$A = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 0 & 4 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$

We solve
$$x_1 \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

for $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Game:

Find out how we may combine column vectors of A to create / generate / reach $\vec{b}$

New idea:

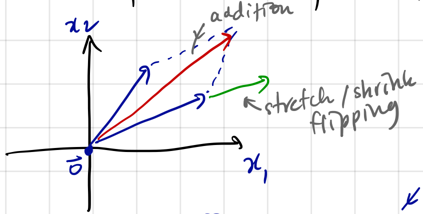
Understand the places ("spaces") where $\vec{a}$'s, $\vec{x}$'s, and $\vec{b}$'s live

E8op1

Big things coming:

- Null space of A
- Column Space of A
- Row Space of A
- Left Null Space of A
- Beautiful connection to A^T, A's transpose

Example Vector Space: Idealized plane



$$V = R^{②} = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ; x_1, x_2 \in R \right\}$$

basic structure; restrictions; curly braces; real numbers $R = R^1$

Two (pretty obvious) features of vector spaces:

They are closed under addition and scalar multiplication.

(1) If we add any two vectors in V, we get another vector that's still in V

(2) If we multiply a vector in V by a scalar (for us: a real number), the result is still in V.

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 773 \\ 6 \end{bmatrix} = \begin{bmatrix} 776 \\ 10 \end{bmatrix}$$

$\in R^2$ $\in R^2$ $\in R2$

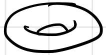
↑ "is an element of"

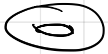
$$7 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 21 \\ 28 \end{bmatrix}$$

$7 \in R$ (real numbers) $\begin{bmatrix} 3 \\ 4 \end{bmatrix} \in R^2$ $\begin{bmatrix} 21 \\ 28 \end{bmatrix} \in R^2$

(E8ap2)

Examples of things that are and are not Vector spaces:

(1) $\vec{V}_1 = 3$  $+ 4$ 

$\vec{V}_2 = 2$  $+ 1.3$ 

$V = \{ [$   $] \}$

(2) $V = \{ f(x) = c_2 x^2 + c_1 x + c_0 \mid c_2, c_1, c_0 \in R \}$.

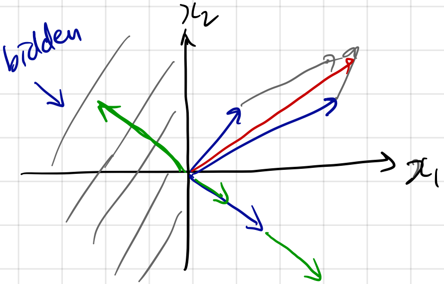
(annotations: $c_2 \to x$, $c_1 \to y$, $c_0 \to z$, $\mid$ = "such that")

$f_1(x) = 2x^2 + 3$

$f_2(x) = -7x^2 + 3x + 4$

$f_1(x) + f_2(x) = -5x^2 + 3x + 7$

(3)



$$V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;\middle|\; x_2 \in \mathbb{R},\; x_1 \geq 0 \right\}$$

$-\infty < x_2 < \infty$.

Not a vector space

Observe:

Addition works

if $\vec{v}_1, \vec{v}_2 \in V$ then $\vec{v}_1 + \vec{v}_2 \in V$

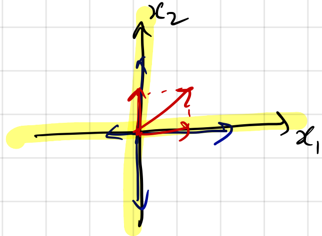
Scalar multiplication fails!

$$-3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix}$$

$\in V$ $\quad$ $\notin V$

Note: we're starting to talk about subspaces

(4) All points on axes of $\mathbb{R}^2$ ∠E8ap3



Now see

Scalar multiplication works but addition fails

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$\in V$ $\quad$ $\in V$ $\quad$ not in V

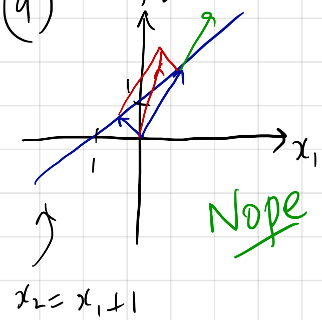
(5) All $m \times n$ matrices form a vector space $\equiv \mathbb{R}^{mn}$

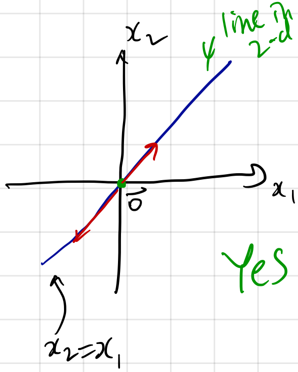
(6) What about the integers ✗

(7) " " rational numbers ✗

(8) " " real numbers ✓

(a)





LE 8ap 4

Vector Spaces Inside Vector Spaces

Menu:

- vector space requirements
- subspace requirements

General Requirements of a Vector Space:

VSP1 if $\vec{x}_1, \vec{x}_2 \in V$ then $\vec{x}_1 + \vec{x}_2 \in V$

VSP2 if $\vec{x} \in V$ then $c\vec{x} \in V$ for all $c \in \mathbb{R}$

VSP3 $\vec{0} \in V$ and $\vec{x} + \vec{0} = \vec{x}$ for all $\vec{x}$ in V

(vector space property)

+ a series of increasingly boring conditions such as $c(\vec{x}_1 + \vec{x}_2) = cx_1 + cx_2$

E8bp1

Our focus: $\mathbb{R}^n$, $n = 0, 1, 2, \ldots$

Super big deal:

Vector Spaces have vector spaces inside them and we call these subspaces

Need three properties for a subset S of V to be a subspace:

SSP1 if $\vec{x}_1$ & $\vec{x}_2 \in S$ then $\vec{x}_1 + \vec{x}_2 \in S$

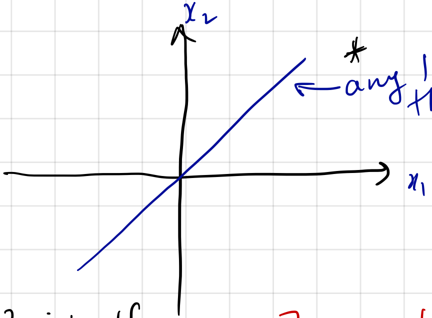
SSP2 if $\vec{x} \in S$ then $c\vec{x} \in S$ for all $c \in \mathbb{R}$

SSP3 $\vec{0} \in S$

E86p2

Examples of subspaces:

$\mathbb{R}^2$



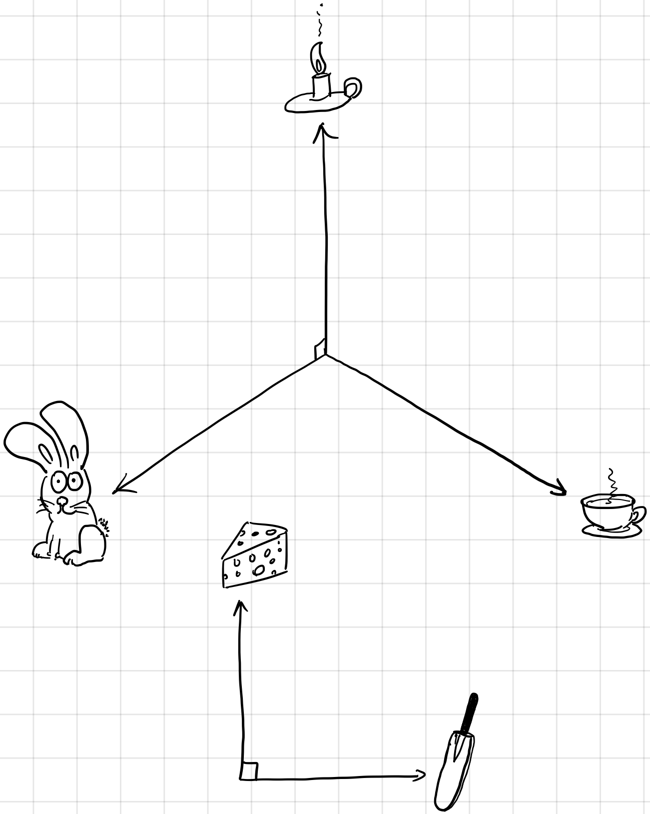
* any line through $\vec{0}$ (∞ly many)

* $\mathbb{R}^2$ itself
* $\{\vec{0}\}$ works too

] important!

$\mathbb{R}^3$: Subspaces
* $\mathbb{R}^3$ itself
* $\{\vec{0}\}$
* any line through $\vec{0}$
* any plane " $\vec{0}$

Bonus very silly Spaces:



"Danger Will Robinson! We are entering column space!!"

E8cp1

Menu: • Column space for $A\vec{x}=\vec{b}$
→ the first of four awesome subspaces

Our [beloved / belated] problem $\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$

delete as applicable

(m = rows, n = columns)

The column picture:

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \vec{b}$$

↑ each column has m entries

Observation / Big deal:

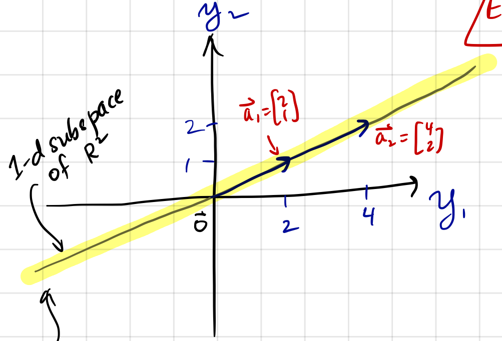
Columns of A and $\vec{b}$ live in $\mathbb{R}^m$ (not $\mathbb{R}^n$ ← $\vec{x}$ lives here)

ex

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

$$\vec{a}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

↑ live in y-space



all linear combinations of $\vec{a}_1$ & $\vec{a}_2$, i.e. $x_1\vec{a}_1 + x_2\vec{a}_2$, live on this line which is a subspace of $\mathbb{R}^2$

Huge:

← notation

$$C(A) = \text{Column Space of } A$$
$$= \text{Subspace of } \mathbb{R}^m$$

(A is $m\times n$)

Here:

$$C(A) = \left\{ \vec{y} \in \mathbb{R}^2 \;\middle|\; \vec{y} = c\begin{bmatrix} 2 \\ 1 \end{bmatrix},\ c \in \mathbb{R} \right\}$$

↑ by eye (at $C(A)=$); ↑ "such that" (at $|$)

Big deal:

$A\vec{x} = \vec{b}$ has a solution (1 or ∞ly many) only if $\vec{b} \in C(A)$

" $\vec{b}$ lives in the column of A ".

$\Rightarrow$ If $\vec{b} \notin A$, $A\vec{x} = \vec{b}$ has no solution.

For $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$

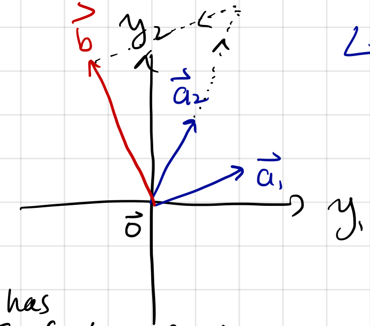
$\vec{b} = \begin{bmatrix} 38 \\ 19 \end{bmatrix} \in C(A)$ ← we'll see means there are ∞ly many solut.

$\vec{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \notin C(A)$

↑ so no solution to $A\vec{x} = \vec{b}$

ex

$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ (columns $\vec{a}_1$, $\vec{a}_2$)

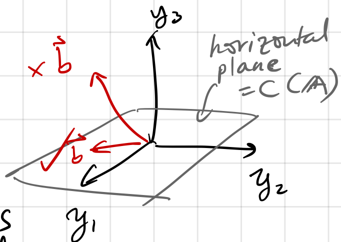


See that $A\vec{x} = \vec{b}$ always has a solution. In fact, only 1.

$\Rightarrow C(A) = \mathbb{R}^2$

ex $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ (3×2, fall)

$\hat{i}$ $\hat{j}$ ← 2 vectors in 3-d.



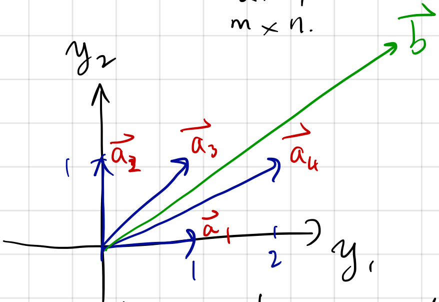
↓

$C(A) \neq \mathbb{R}^3$ ← $m = 3$

ex: $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{bmatrix}$

wide A →

2×4
$m \times n$.



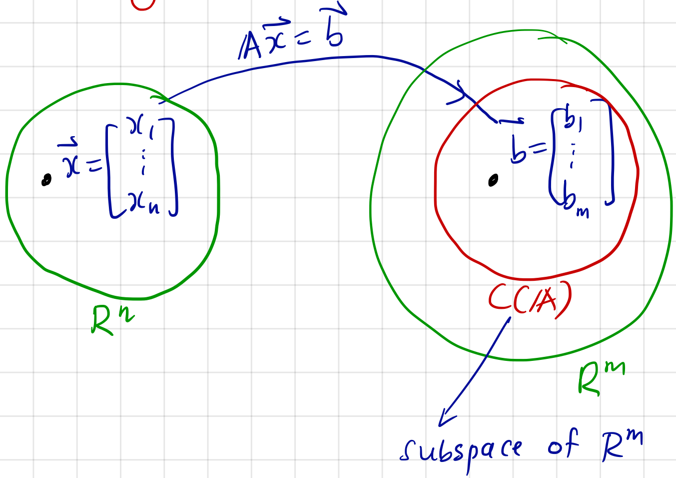
See: any two column vectors of A will work.

⇒ ∞ly many solutions

Again $C(A) = R^2$ ← m=2

Emerging Picture

LE8cp3



A new realm opens up: Null Space

Menu:
- Definition of the Null space of A, $N(A)$
- what $C(A)$ & $N(A)$ mean for $A\vec{x} = \vec{b}$

Consider $A\vec{x} = \vec{0}$ ← very special $\vec{b}$

↑ called Nullspace Equation or Homogeneous Equation

how can we combine columns of A to produce nothing?

Immediate Observation: $A\vec{0} = \vec{0}$

So: Always a solution → $\vec{0} \in C(A)$ always

→ May be 1 or ∞ly many

Example: Solve $A\vec{x} = \vec{0}$ for:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -3 & -2 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{(row ops: pain \& suffering)} \dashrightarrow \quad \vec{x} = c\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \text{ where } c \in \mathbb{R} \checkmark$$

$m^{3\times 3}n$

$$N(A) = \left\{ \vec{x} \in \mathbb{R}^3 \;\middle|\; \vec{x} = c\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, c \in \mathbb{R} \right\}$$

↑ Null space of A (3 = n) ← subspace of $\mathbb{R}^3$.

E9ap1

Notation:

$\vec{x}_n$ for a nullspace vector

Also $\vec{x}_h$ ← homogeneous

Now solve:

$$A\vec{x} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -3 & -2 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \leftarrow \vec{b}$$

non-zero interesting $\vec{b}$

more pain and suffering (row operations)

find

$$\vec{x}_r = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

← could be called $\vec{x}_p$ where p is "particular"

we'll write $\vec{x}_r$ because some dying monk said we should

So $A\vec{x}_n = 0$ & $A\vec{x}_r = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} \leftarrow \vec{b}$

⇒ See $\vec{x}_r$ is not a unique solution b/c $A(\vec{x}_r + \vec{x}_n) = \vec{b}$

More:

$$\begin{bmatrix}1&2&3\\1&-3&-2\\0&1&1\end{bmatrix}\left(\begin{bmatrix}1\\0\\1\end{bmatrix}+c\begin{bmatrix}1\\1\\-1\end{bmatrix}\right)$$

(A, $\vec{x}_r$, $\vec{x}_n$)

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

$$=\begin{bmatrix}4\\-1\\1\end{bmatrix}+\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$=\begin{bmatrix}4\\-1\\1\end{bmatrix}=\vec{b}$$

$\Rightarrow$ There are infinitely many solutions b/c $N(A)\neq\{\vec{0}\}$

In general:

$$A(\vec{x}_r+\vec{x}_n)=A\vec{x}_r+A\vec{x}_n$$

($A\vec{x}_r = \vec{b}$, $A\vec{x}_n = \vec{0}$)

E9ap2

The Big Deals:

(1) All vectors $\vec{x}\in R^n$ for which $A\vec{x}=\vec{0}$ form a subspace of R^n

SSP1 $A(\vec{x}_1+\vec{x}_2)=A\vec{x}_1+A\vec{x}_2=\vec{0}$

SSP2 $A(c\vec{x})=c(A\vec{x})=c\vec{0}=\vec{0}$

SSP3 $\vec{0}\in N(A)$ b/c $A\vec{0}=\vec{0}$

(2)
- If $N(A)=\{\vec{0}\}$ and $\vec{b}\in C(A)$ then $A\vec{x}=\vec{b}$ has one, unique solution
- If $N(A)\neq\{\vec{0}\}$ and $\vec{b}\in C(A)$ then $A\vec{x}=\vec{b}$ has infinitely many solutions
- if $\vec{b}\notin C(A)$ then $A\vec{x}=\vec{b}$ has no solutions

Row Reduction, as you wish.

menu:

- Turning $\mathbb{A}\vec{x}=\vec{b}$ into $\mathbb{R}_{\mathbb{A}}\vec{x}=\vec{d}$
- Reduced Row Echelon Forms (RREFs)
- Pivot and free variables
- The rank r of a matrix
- Fezzik

so much winning

$$\mathbb{A}=\begin{bmatrix}2&4&3&4\\2&4&6&10\\6&12&12&18\end{bmatrix} \qquad \vec{b}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$$

Fezzik, 3×4, $\vec{x}\in\mathbb{R}^4$, arbitrary

Monks tell us:

Solve $\mathbb{A}\vec{x}=\vec{b}$ for general $\vec{b}$

$$[\mathbb{A}\,|\,\vec{b}]=\left[\begin{array}{cccc|c}2&4&3&4&b_1\\2&4&6&10&b_2\\6&12&12&18&b_3\end{array}\right]$$

D_1

$R_2'=1\cdot R_2-\left(\frac{2}{2}\right)R_1$ ($2 \leftarrow D_1$)

$R_3'=1\cdot R_3-\left(\frac{6}{2}\right)R_1$ ($2 \leftarrow D_1$)

$$\left[\begin{array}{cccc|c}2&4&3&4&b_1\\0&0&3&6&b_2-b_1\\0&0&3&6&b_3-3b_1\end{array}\right]$$

D_1, D_2

$R_3'=1\cdot R_3-\left(\frac{3}{3}\right)$

pivot column, free column, pivot column, free column

E9bp1

$$\left[\begin{array}{cccc|c}2&4&3&4&b_1\\0&0&3&6&b_2-b_1\\0&0&0&0&b_3-b_2-2b_1\end{array}\right]$$

$U\vec{x}=\vec{c}$.

$(b_3-3b_1)-(b_2-b_1)$

Keep Going!! (as with inverses).

$R_1'=1\cdot R_1-\left(\frac{3}{3}\right)R_2$

$$\left[\begin{array}{cccc|c}2&4&0&-2&2b_1-b_2\\0&0&3&6&b_2-b_1\\0&0&0&0&b_3-b_2-2b_1\end{array}\right]$$

last step: divide through by pivots

$R_1'=\frac{1}{2}R_1$

$R_2'=\frac{1}{3}R_2$

1st 2 columns of Identity matrix

$$\left[\begin{array}{cccc|c}1&2&0&-1&\frac{1}{2}(2b_1-b_2)\\0&0&1&2&\frac{1}{3}(b_2-b_1)\\0&0&0&0&b_3-b_2-2b_1\end{array}\right]$$

$$=\left[\,\mathbb{R}_{\mathbb{A}}\,\middle|\,\vec{d}\,\right]$$

Reduced Row Echelon Form of $\mathbb{A}$

LE9bp2

Big Deal Things:

- We can't reduce any further
- $\mathbb{R}_{\mathbb{A}}$ is unique for any $\mathbb{A}$
- Row swaps are still part of the game
- New — pivots may appear irregularly

- Pivot columns match Identity Matrix Columns
- We call x_j that match up with pivot columns, the pivot variables.
- Similarly: free columns ↔ free variables

For Fezzik x_1 & x_3 are pivot variables
x_2 & x_4 are free variables

Very, very big deal:

Definition:

pivot columns in $\mathbb{R}_{\mathbb{A}}$ = rank of $\mathbb{A}$

Notation: rank of $\mathbb{A} = r$

For Fezzik: $r=2$

→ $m=3$, $n=4$, $r=2$

Huge idea: Inside every matrix $\mathbb{A}_{m \times n}$ there is an invertible $r \times r$ square matrix

The Search for Column Space...

← $C(A)$

Our Story:

$C(A)$ = all $\vec{b}$'s for which $A\vec{x} = \vec{b}$ has a solution.

↓ subspace of R^m

Method 1 of 3 for finding $C(A)$:

reduce $[A \mid \vec{b}]$ ← general $\vec{b}$ to $[R_A \mid \vec{d}]$ and for all rows of 0 in R_A, set matching entries in $\vec{d}$ to 0.

Our friend Fezzik:

$$[A \mid \vec{b}] = \left[\begin{array}{cccc|c} 2 & 4 & 3 & 4 & b_1 \\ 2 & 4 & 6 & 10 & b_2 \\ 6 & 12 & 12 & 18 & b_3 \end{array}\right]$$

previous row reduction suffering ↓ 〰

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & -1 & \frac{1}{2}(2b_1 - b_2) \\ 0 & 0 & 1 & 2 & \frac{1}{3}(b_2 - b_1) \\ 0 & 0 & 0 & 0 & b_3 - b_2 - 2b_1 \end{array}\right] = [R_A \mid \vec{d}]$$

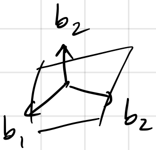
← $\vec{d}$

[E9cp1]

$$\underbrace{0x_1 + 0x_2 + 0x_3 + 0x_4}_{0} = \underbrace{b_3 - b_2 - 2b_1}_{\text{must be } 0}$$

$\Rightarrow$ $b_3 - b_2 - 2b_1 = 0$

↖ eq of a plane in



true but not useful.

Better (but not only way):

Set $b_3 = b_2 + 2b_1$, where $b_1, b_2 \in R$

↖ b_3 depends on ↗

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_2 + 2b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

always do this. ← fixed vectors →

where $b_1, b_2 \in R$.

Formal result:

$$C(A) = \left\{ \vec{b} \in R^3 \;\middle|\; \vec{b} = b_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},\ b_1, b_2 \in R \right\}$$

Big deal: See $C(A)$ is a 2-d subspace (plane) of R^3 $(m=3)$

rank of A, r

#awesome

Notes

* Because $C(A)$ does not fill up R^3 then $A\vec{x}=\vec{b}$ may or may not have solutions

* Nothing super special about $\begin{bmatrix}1\\0\\2\end{bmatrix}$ & $\begin{bmatrix}0\\1\\1\end{bmatrix}$

$\times 2$ $\times -3$

$\begin{bmatrix}2\\0\\4\end{bmatrix}$, $\begin{bmatrix}0\\-3\\-3\end{bmatrix}$ would work

*we had $b_3 - b_2 - 2b_1 = 0$

$b_2 = b_3 - 2b_1$

↑ dependent var

LE9cp2

very nutritious

* $b_3 - b_2 - 2b_1 = 0$

$$\begin{bmatrix}-2 & -1 & 1\end{bmatrix}\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix} = \begin{bmatrix}0\end{bmatrix}$$

1×3 3×1 1×1

$A \quad \vec{x} = \vec{0}$

Solve a nullspace problem to find $C(A)$

The Search for Null Space, $N(A)$:

Quest: find all $\vec{x}$ such that $A\vec{x} = \vec{0}$

$\vec{x}$ is on right of A ⇒ Right null space

Again, with Fezzik:

$[A \mid \vec{b}] =$ (with $\vec{b}$ crossed out and replaced by $\vec{0}$)

$$\left[\begin{array}{cccc|c} 2 & 4 & 3 & 4 & b_1 \\ 2 & 4 & 6 & 10 & b_2 \\ 6 & 12 & 12 & 18 & b_3 \end{array}\right]$$

previous row reduction suffering

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & -1 & \frac{1}{2}(2b_1 - b_2) \\ 0 & 0 & 1 & 2 & \frac{1}{3}(b_2 - b_1) \\ 0 & 0 & 0 & 0 & b_3 - b_2 - 2b_1 \end{array}\right] = [R_A \mid \vec{d}]$$

($\vec{d} \to \vec{0}$)

For $N(A)$, set $\vec{b} = \vec{0}$ (or start with $\vec{b} = \vec{0}$)

that's not in video — the shame

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

E9d p1

Usual story:

$$x_1 + 2x_2 \quad - x_4 = 0$$
$$x_3 + 2x_4 = 0$$

free variables: x_2, x_4

pivot variables: x_1, x_3

Always do the following well defined procedure:
Express pivot variables (dependent) in terms of the free variables (independent)

$$\Rightarrow \quad x_1 = -2x_2 + x_4$$
$$x_3 = 0 \quad -2x_4$$

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

where $x_2, x_4 \in \mathbb{R}$

$A\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$, $A\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} = \vec{0}$

always do this!

replace pivot vars with free vars

L E9d p2

Formally:

$$N(A) = \left\{ \vec{x} \in \mathbb{R}^{4} \;\middle|\; \vec{x} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right\} \quad x_2, x_4 \in \mathbb{R}$$

(4 circled: ↙ n)

↑ For Fezzik, this is a plane in 4 dimensions

Notes:

* b/c $N(A) \neq \{\vec{0}\}$, if $A\vec{x} = \vec{b}$ has a solution (i.e., $\vec{b} \in C(A)$) then there a ∞ly many solutions

* Soon we we'll see that the dimension of $N(A)$ is

$$\dim N(A) = n - r$$

(n ↑ # columns; r ↖ rank of A)

Fezzik: $4 - 2 = 2$ ✓

Solving $A\vec{x}=\vec{b}$ the Subspace Way:

- Fezzik with $\vec{b}$ in $C(A)$ & $\vec{b}\neq\vec{0}$

E9ep1

From before:

$$[A\,|\,\vec{b}] = \left[\begin{array}{cccc|c} 2 & 4 & 3 & 4 & b_1 \\ 2 & 4 & 6 & 10 & b_2 \\ 6 & 12 & 12 & 18 & b_3 \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|c} \overset{p}{1} & \overset{F}{2} & \overset{p}{0} & \overset{F}{-1} & \frac{1}{2}(2b_1-b_2) \\ 0 & 0 & 1 & 2 & \frac{1}{3}(b_2-b_1) \\ 0 & 0 & 0 & 0 & b_3-b_2-2b_1 \end{array}\right] = [R_A\,|\,\vec{d}]$$

example: $\vec{b} = \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}$ (b_1, b_2, b_3) ← see this is the first column!

$$[R_A\,|\,\vec{d}] = \left[\begin{array}{cccc|c} 1 & 2 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \leftarrow \checkmark$$

$\vec{b} \in C(A)$

Plan: Use same steps as for finding $N(A)$

$$\begin{cases} \overset{p}{x_1} + \overset{F}{2x_2} \qquad - \overset{F}{x_4} = 1 \\ \qquad\qquad \overset{p}{x_3} + 2x_4 = 0 \end{cases}$$

each pivot variable appears only once in all equations

put free variables on RHS

$$x_1 = 1 - 2x_2 + x_4$$
$$x_3 = 0 + 0x_2 - 2x_4$$

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

where $x_2, x_4 \in \mathbb{R}$

$\vec{x}_p$, p = particular; $\vec{x}_h$, h = homogeneous

replace pivot vars w. free var expression

General Story:

$$A(\vec{x}_p + \vec{x}_h) = A\vec{x}_p + A\vec{x}_h = \vec{b}$$

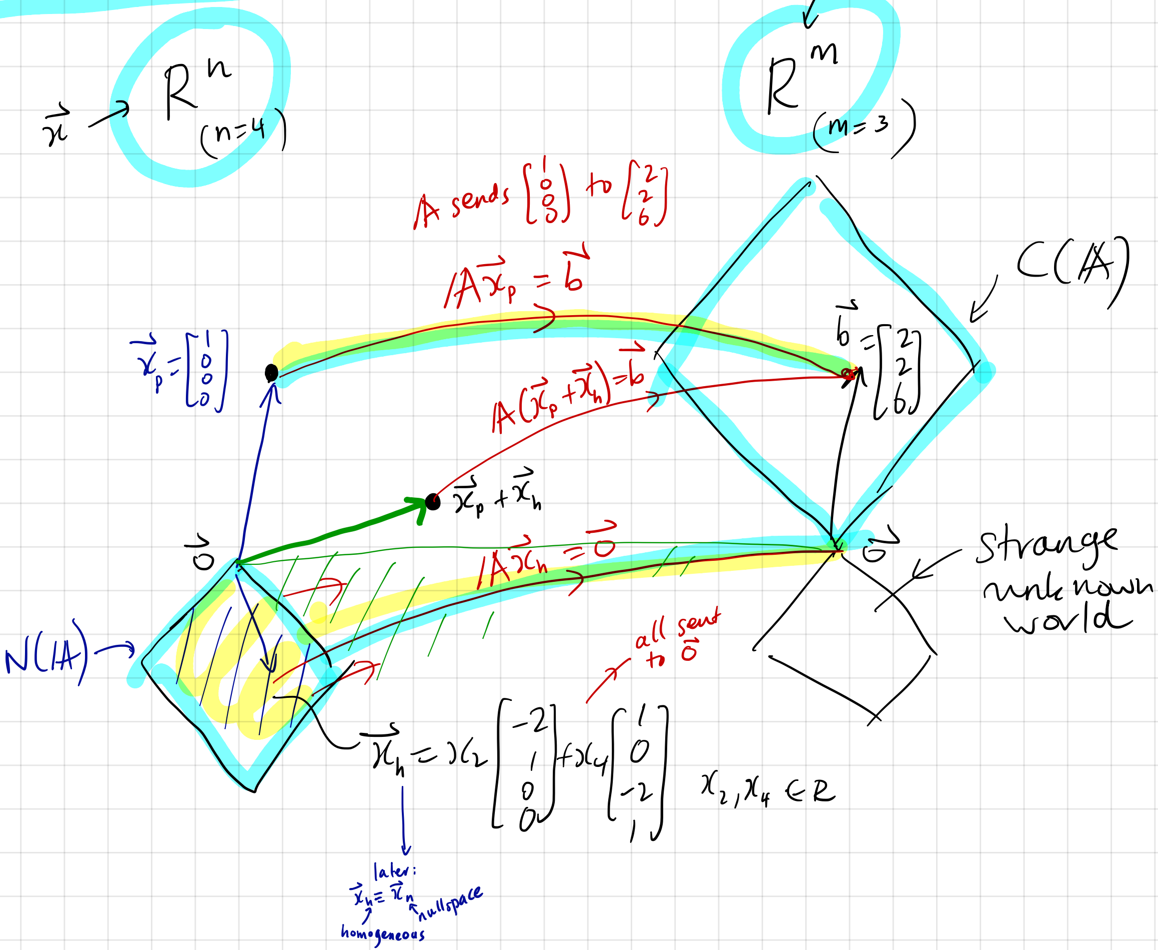
($A\vec{x}_p \to \vec{b}$, $A\vec{x}_h \to \vec{0}$; $\vec{x}_p + \vec{x}_h = \vec{x}$; $\vec{x}_p$ not unique; not necessarily the same)

Later:

$$A(\vec{x}_r + \vec{x}_n) = A\vec{x}_r + A\vec{x}_n$$

($\vec{x}_r + \vec{x}_n = \vec{x}$; $A\vec{x}_r \to \vec{b}$, $A\vec{x}_n \to \vec{0}$; $\vec{x}_r$: unique, row — think of as most excellent $\vec{x}_p$; $\vec{x}_n$: null)

Big picture emerging for $A\vec{x} = \vec{b}$.



"I see null vectors"

E9fp1

- Jumping to the form of $N(A)$ from R_A

Fezzik's R_A:

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

P F P F

pivot columns: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$ ↑ $r=2$

Our two null vectors from our earlier solution for $N(A)$:

make a matrix

$$N = \begin{matrix} P \\ F \\ P \\ F \end{matrix} \begin{bmatrix} -2 & 1 \\ 1 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix}$$

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

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

See $R_A N = 0$

R_A: $m \times n$, N: $n \times (n-r)$ ↑ secret, 0: $m \times (n-r)$ ↑ all zeros = kevin the destroyer of words

also $A N = 0$

↳ A & R_A have the same Nullspace.

General Story

$$R_A = \begin{bmatrix} I & F \\ 0 & 0 \cdots 0 \end{bmatrix}$$

R_A: $m \times n$; I: $r \times r$; F: $r \times (n-r)$ ← permutation of x_i's; 0 ← all zeros, $(m-r) \times n$

m−r rows of 0s

$$N = \begin{bmatrix} -F \\ I \end{bmatrix}$$

N: $n \times (n-r)$

$n-r = \dim N(A)$

$$R_A N = \left[-IF + FI\right] = 0$$

R_A: $m \times n$, N: $n \times (n-r)$; IF: $r \times (n-r)$, FI: $r \times r$; 0: $(m-r) \times (n-r)$ — absent in videos due to cavalier attitude; 0: $m \times (n-r)$

E9gp1

Solving $A\vec{x}=\vec{b}$ the Subspace way: simpler examples

(1) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ← the identity matrix ($m \times n$)

$A\vec{x}=\vec{b}$ is always solvable!

($A = I$ → $\vec{x}=\vec{b}$)

see $R_A = A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (p, p)

* Column Space

Solve $\left[\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 1 & b_2 \end{array}\right]$ for $C(A)$ (p, p) ← R_A already here

$\Rightarrow b_1, b_2 \in R$ (no restrictions)

$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (always) $b_1, b_2 \in R$

$\Rightarrow C(A) = R^2$ ← $m=2$

* Null Space

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

$\Rightarrow$ $x_1 = 0$, $x_2 = 0$ $\Rightarrow$ $\vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$ is only solution

$N(A) = \{\vec{0}\}$.

Delectable Upshots

- Every $\vec{b} \in C(A)$ so $A\vec{x}=\vec{b}$ is always solvable
- Because $N(A)=\{\vec{0}\}$, every solution is unique.

$\dim C(A) = 2$ ($= r$) later

$\dim N(A) = 0$ ($= n - r$) ($2-2$)

(2)

$\mathbb{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ with $\vec{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ & then $\vec{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

* see E8cp1 for first examination of this $\mathbb{A}$

E9gp2

see columns are multiples of each other as are rows

Find $C(\mathbb{A})$:

$$[\mathbb{A} \mid \vec{b}] = \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 2 & 4 & b_2 \end{array}\right]$$

D_1

$$\sim \quad R_2' = R_2 - \left(\frac{2}{1}\right) R_1$$

D_1

$$\left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & 0 & b_2 - 2b_1 \end{array}\right]$$

P F ; rank $r = 1$; $\mathbb{R}_{\mathbb{A}}$; d ; $= 0$

$\Rightarrow$ $b_2 - 2b_1 = 0$

$b_2 = 2b_1$ where $b_1 \in R$

dependent ; independent.

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ 2b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ where } b_1 \in R$$

Find $N(\mathbb{A})$:

$$[\mathbb{A} \mid \vec{0}] \sim \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

P F

$\Rightarrow$ $x_1 + 2x_2 = 0$

P ; F

$\Rightarrow$ $x_1 = -2x_2$

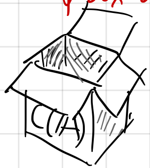
$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} \text{ where } x_2 \in R$$

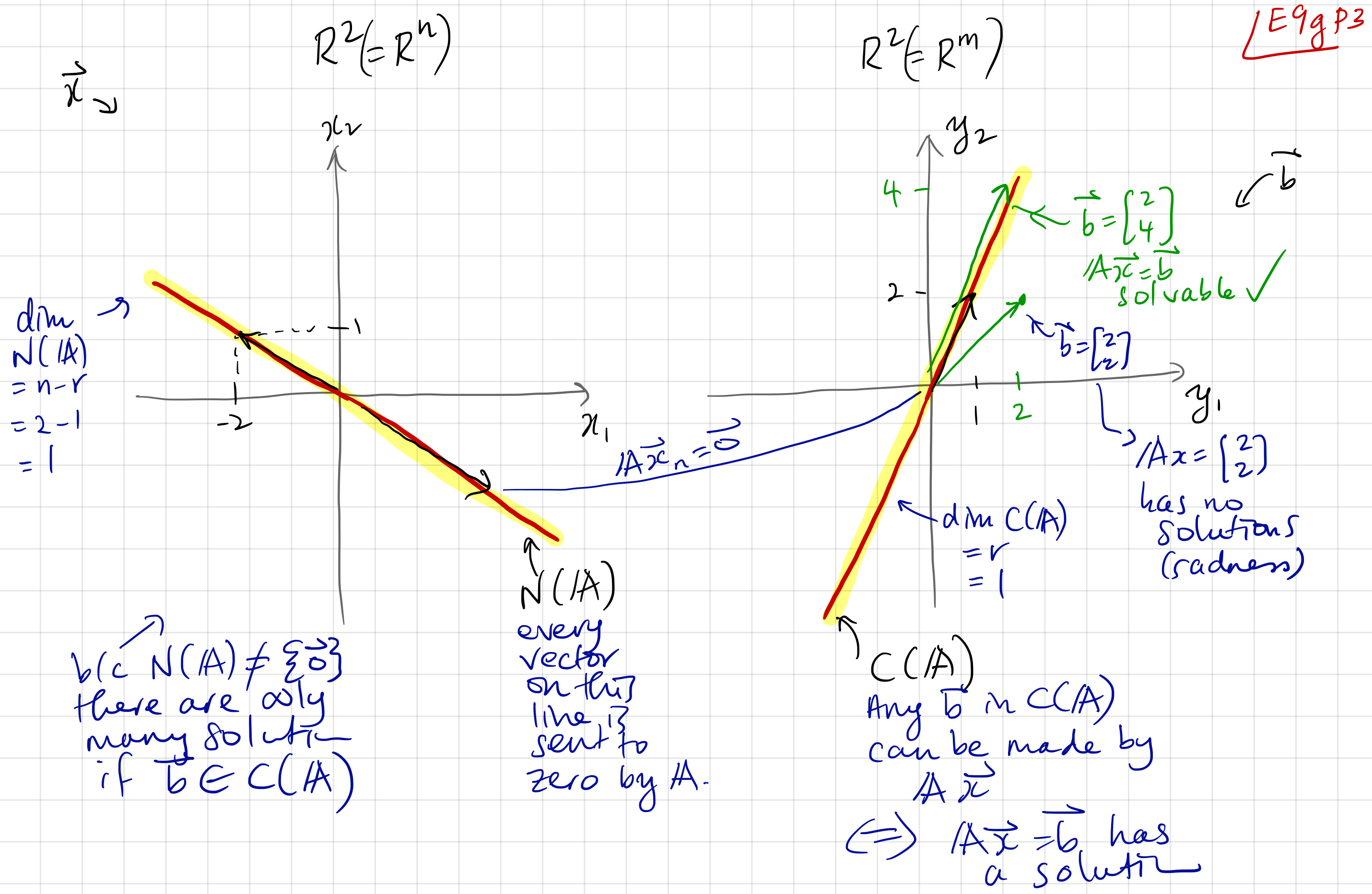
replace pivot var

$$C(\mathbb{A}) = \left\{ \vec{b} \in R^2 \,\middle|\, \vec{b} = b_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ b_1 \in R \right\}$$

$$N(\mathbb{A}) = \left\{ \vec{x} \in R^2 \,\middle|\, \vec{x} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix},\ x_2 \in R \right\}$$

box of vectors




E9gP3
$R^2 (= R^n)$
$R^2 (= R^m)$
$\vec{x}$
$\vec{b}$
x_2
x_1
y_2
y_1
-2
1
2
4
dim N(A) = n-r = 2-1 = 1
$A\vec{x}_n = \vec{0}$
N(A)
every vector on this line is sent to zero by A.
b/c N(A) ≠ {$\vec{0}$} there are ∞ly many solution if $\vec{b} \in C(A)$
$\vec{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$
$A\vec{x} = \vec{b}$ solvable ✓
$\vec{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$
$Ax = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ has no solutions (sadness)
dim C(A) = r = 1
C(A)
Any $\vec{b}$ in C(A) can be made by $A\vec{x}$
⟺ $A\vec{x} = \vec{b}$ has a solution

What $\mathbb{R}_A$ tells us

Menu:

- The four basic kinds of $\mathbb{R}_A$
- How these forms for $\mathbb{R}_A$ dictate $A\vec{x}=\vec{b}$'s solution
- Later: $\mathbb{R}_{A^T}$ gives us the rest of what we need to know

---

The Story so far:

$\mathbb{R}_A$ provides us with

(1) The rank r of A (# pivot columns)

(2) Nullspace $N(A)$ (solve $\mathbb{R}_A\vec{x}=\vec{0}$)

(3) The number of possible solutions to $A\vec{x}=\vec{b}$

(1),(2) → (3) because:

- If $r < m$, one or more rows of $\mathbb{R}_A$ are all 0's and therefore some $\vec{b}$'s will lead to no solution for $A\vec{x}=\vec{b}$
- If $N(A) \neq \{\vec{0}\} \iff \mathbb{R}_A$ has one or more free columns then $A\vec{x}=\vec{b}$ will have ∞ly many solutions if it is solvable (i.e., if $\vec{b} \in C(A)$)

Four examples:

E10ap1

(i)

$$\mathbb{R}_A = \begin{bmatrix} 1 & \omega_{12} & \omega_{13} & 0 & \omega_{15} & 0 \\ 0 & 0 & 0 & 1 & \omega_{25} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(columns: pivot, free, f, p, f, p.)

from a wide A

no row of zeros

$m=3$, $n=6$, $r=3$

See: always a solution to $A\vec{x}=\vec{b}$

$\Rightarrow C(A) = R^3 = R^m = R^r$

Also $N(A)$ is a 3-d subspace of R^6 ($\leftarrow n=6$)

$\vec{x}$ lives in R^6; to be proven

Know $N(A) \neq \{\vec{0}\}$

So: $A\vec{x}=\vec{b}$ always has a solut (C(A)) and there are always ∞ly many (N(A))

Note: Wide A's always have at least 1 free variable $\Rightarrow N(A) \neq \{\vec{0}\}$

E10ap2

(ii) $\mathbb{R}_{/A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbb{I}$ (columns: p p p)

$m = n = r = 3$

$\Rightarrow$ No free variables

See: $A\vec{x} = \vec{b}$ is always solvable

and $N(A) = \{\vec{0}\}$

So: $A\vec{x} = \vec{b}$ always has 1, unique solut.

For square invertible matrices ($n \times n$)

$\mathbb{R}_{/A} = \mathbb{I}$ always.

↖ 1-1 mapping from $R^n \to R^n$

(iii) $\mathbb{R}_{/A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ $m = 4$, $n = r = 3$

tall →

↖ tall matrices must always have a row of zeros

see:

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

+ Possible: no solutions

$A\vec{x} = \vec{b}$ has 0 or 1 solutn

(0 → $\vec{b} \notin C(A)$; 1 → $b \in C(A)$)

(iv)

m and $n > r$

$\begin{bmatrix} 1 & \omega_{12} & 0 & \omega_{14} \\ 0 & 0 & 1 & \omega_{24} \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (columns: P F P F)

$m = 3$

$n = 4$

$r = 2$

↖ later: really a 2 by 2 matrix in a 3×4 matrix

A sends a plane in R^4 to a plane in R^3

See: $N(A) \neq \{\vec{0}\}$ (2 free variables)

$C(A)$ may or may not contain $\vec{b}$ (row of 0's in $\mathbb{R}_{/A}$)

$\Rightarrow$ either 0 or ∞ly many solutn.

E10ap3

Case:	example R_A	# Solutions
(i) $m=r$, $n=r$ — square	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (p p p)	1 always: $C(A)=\mathbb{R}^m$, $N(A)=\{\vec{0}\}$
(ii) $m=r$, $n>r$ — wide	$\begin{bmatrix} 1 & -2 & 0 & -4 & 0 \\ 0 & 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ (P F P F P)	∞ always: $C(A)=\mathbb{R}^m$, $N(A)\neq\{\vec{0}\}$
(iii) $m>r$, $n=r$ — tall	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (p p p)	0 or 1: $C(A)\neq\mathbb{R}^m$, $N(A)=\{\vec{0}\}$
(iv) $m>r$, $n>r$ — many possibilit.	$\begin{bmatrix} 1 & -2 & 0 & 12 \\ 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	0 or ∞: $C(A)\neq\mathbb{R}^m$, $N(A)\neq\{\vec{0}\}$

Next:
- find bases for $C(A)$ & $N(A)$
- $\dim C(A)=r$, $\dim N(A)=n-r$

E11aP1

Getting to know your subspaces:

Menu:
- Care and feeding
- Spanning sets
- Bases
- Dimensions

New friend: Inigo $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}$

Plan: explore $A\vec{x} = \vec{b}$ with Inigo

First ~ find C(A) and N(A)

Solve $\left[\begin{array}{ccc|c} 1 & 2 & 1 & b_1 \\ 2 & 4 & 2 & b_2 \end{array}\right] = [A \mid \vec{b}]$ (D_1 = pivot 1)

$R_2' = R_2 - \left(\frac{2}{1}\right) R_1$ (ℓ_{21}; D_1)

$\left[\begin{array}{ccc|c} 1 & 2 & 1 & b_1 \\ 0 & 0 & 0 & b_2 - 2b_1 \end{array}\right] = [R_A \mid \vec{d}]$ (P F F)

m = 2 rows, n = 3 columns, r = 1, rank

C(A): Must have $b_2 - 2b_1 = 0$ for solution to be possible

$\Rightarrow b_2 = 2b_1$ (b_2 dependent, b_1 independent)

$\Rightarrow \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ 2b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$C(A) = \left\{ \vec{y} \in R^2 \;\middle|\; \vec{y} = c \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ c \in R \right\}$$

line through origin; r = 1; 1-d subspace of R^2 — m = 2

N(A): Solve $A\vec{x} = \vec{0}$

$\Rightarrow$ Set $\vec{b} = 0$ in previous

$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = [R_A \mid \vec{0}]$ (P F F)

$\Rightarrow x_1 + 2x_2 + x_3 = 0 \Rightarrow x_1 = -2x_2 - x_3$ (P F F)

express pivot variables in terms of the free variables

replace pivot variables

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

where $x_2, x_3 \in R$

always

$$N(A) = \left\{ \vec{x} \in R^3 \;\middle|\; \vec{x} = c_1 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix};\ c_1, c_2 \in R \right\}$$

(n = 3) plane in R^3

Always true:

$C(A) \subset R^m$ & $N(A) \subset R^n$

($\vec{b}$'s in R^m, $\vec{x}$'s in R^n; $\subset$ "is a subspace of")

Boring but important:

How do we know C(A) & N(A) are really subspaces and not some wheedly subsets?

N(A) for Inigo comprises all linear combinations of $\begin{bmatrix}-2\\1\\0\end{bmatrix}$ & $\begin{bmatrix}-1\\0\\1\end{bmatrix}$

(all linear combinations → automatic subspacification)

Check supspace properties:

(SSP1) if $\vec{x}_1, \vec{x}_2 \in N(A)$, $\vec{x}_1 + \vec{x}_2 \in N(A)$

$\vec{x}_1 = c_{11}\begin{bmatrix}-2\\1\\0\end{bmatrix} + c_{12}\begin{bmatrix}-1\\0\\1\end{bmatrix}$, $\vec{x}_2 = c_{21}\begin{bmatrix}-2\\1\\0\end{bmatrix} + c_{22}\begin{bmatrix}-1\\0\\1\end{bmatrix}$

$\Rightarrow \vec{x}_1 + \vec{x}_2 = (c_{11}+c_{21})\begin{bmatrix}-2\\1\\0\end{bmatrix} + (c_{12}+c_{22})\begin{bmatrix}-1\\0\\1\end{bmatrix}$ = Vector in N(A)

(SSP2) if $\vec{x}_1 \in N(A)$, $c\vec{x}_1 \in N(A)$ for all $c \in \mathbb{R}$

Yes: $\underbrace{c \times c_{11}}_{\text{still a real number}}\begin{bmatrix}-2\\1\\0\end{bmatrix} + \underbrace{c \times c_{21}}_{\text{same}}\begin{bmatrix}-1\\0\\1\end{bmatrix}$

(SSP3) $\vec{0} \in N(A)$ E11ap2

Yes: set $c_{11} = c_{12} = 0 \Rightarrow \vec{x}_1 = \begin{bmatrix}0\\0\\0\end{bmatrix}$

General Story:

Sets made up of all linear combinations of some set of vectors are automatically subspaces.

Terminology:

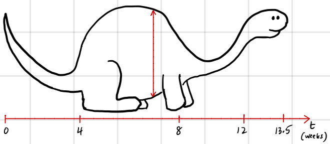
We say $\begin{bmatrix}-2\\1\\0\end{bmatrix}$ & $\begin{bmatrix}-1\\0\\1\end{bmatrix}$ "span" the nullspace of A

and that $\left\{\begin{bmatrix}-2\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$ are a "Spanning set" for N(A)

"All your bases are belong to us"

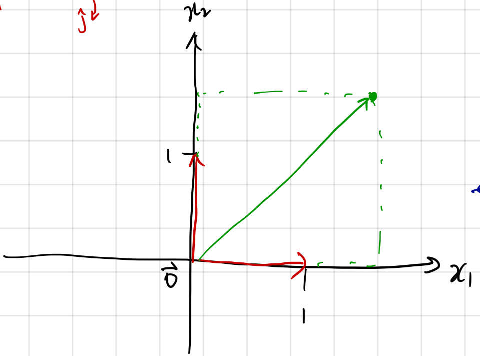
Menu:
- Spanning sets for vector spaces & subspaces
- Bases for vector spaces & subspaces
- How bases are all about $\mathbb{A}\vec{x}=\vec{0}$ and the Nullspace of $\mathbb{A}$
- The dimensions of subspaces
- And this:

Melvin the Course Difficulty Dinosaur:



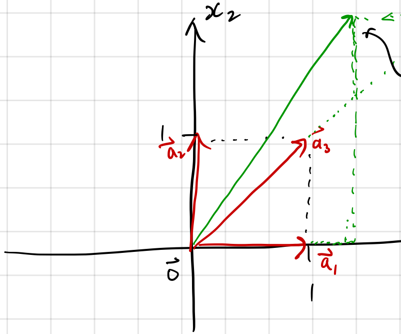
Three Examples of Spanning sets for $\mathbb{R}^2$

(1) $\left\{ \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix} \right\}$ ($\vec{a}_1 = \hat{i}$, $\vec{a}_2 = \hat{j}$) — basis



2 non-parallel vectors give unique representation of each vector

(2) $\left\{ \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix} \right\}$ ($\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$) — not a basis



← only many ways to create this vector from a linear combinat of $\vec{a}_1, \vec{a}_2$, and $\vec{a}_3$

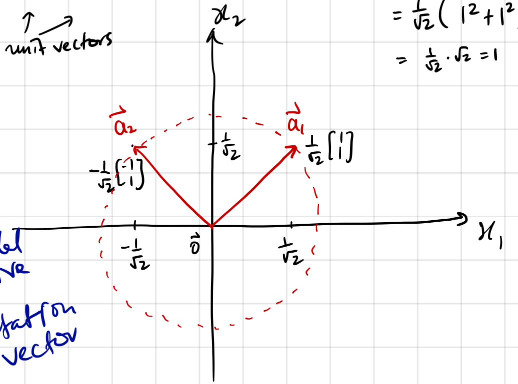
(3) $\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix} \right\}$ ($\vec{a}_1$, $\vec{a}_2$) ← basis.

↑ unit vectors

note: $\|\vec{a}_1\|$ ← length

$= \frac{1}{\sqrt{2}}(1^2+1^2)^{1/2}$

$= \frac{1}{\sqrt{2}} \cdot \sqrt{2} = 1$



Observe:

- Examples (1) & (3) are special because we need both vectors
- For (2), we could take any one vector away, and the remaining two would still span R^2

The right words for the above:

(1) & (3) have linearly independent sets of vectors

(2) has a linearly dependent set of vectors

Big Deal time:

Defn: A set of vectors $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ in R^m is linearly independent if

ER → $x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n = \vec{0}$ only has $x_1 = x_2 = \ldots = x_n = 0$ as a solution

Nullspace Equation $A\vec{x} = \vec{0}$

(x_i is a scalar)

Why? If one or more $x_i \neq 0$, then we can express one vector in terms of the others

ex (2) $\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}$

$\vec{a}_3 \quad \vec{a}_1 \quad \vec{a}_2$

$1\cdot\vec{a}_1 + 1\cdot\vec{a}_2 - 1\cdot\vec{a}_3 = \vec{0}$

(x_1, x_2, x_3 point to the coefficients)

Seeing things:

$\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ are linearly independent

$\iff$ $\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} \vec{x} = \vec{0}$

$\mathbb{A}$ ($m \times n$), $\vec{x}$ ($n \times 1$)

has only $\vec{x} = \vec{0}$ as a solution

$\iff$ $N(\mathbb{A}) = \{\vec{0}\}$

So exciting...

Defn: A spanning set that is linearly independent is called a **basis**

(plural: bases) ← "bay sieze"

Note: Bases are not unique (see ex(1) & (3) above) but some bases are better than others, and some are totally awesome

General goodness: Bases give us a unique representation of every point in the space they span.

Defn:

The **dimension** of a space is the number of vectors in any basis for that space

E11cp1

Why the dimension of C(A) is the rank of A, r

menu
- Including a second way to find C(A)
- Inigo & Fezzik

Claim: $\dim C(A) = r = \#$ pivot columns in R_A (← RREF)

Two key points:

#1 When we perform row operations on a matrix, the relationships between the columns do not change.

Fezzik:

$$A = \begin{bmatrix} 2 & 4 & 3 & 4 \\ 2 & 4 & 6 & 10 \\ 6 & 12 & 12 & 18 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R_A$$

(columns labelled P F P F; pivot columns of R_A contain the identity matrix)

Observations: x_1 & x_3 are pivot variables; x_2 & x_4 are free variables

$c_2 = 2c_1$ in both A & R_A

$c_4 = -c_1 + 2c_3$ " " " " "

easiest to see relationships in R_A ⇓ Identity matrix in pivot columns is key

#2 Follows that in A, the free columns can be made out of the pivot columns, and the pivot columns have to be linearly independent.

⇒ The pivot columns of A form a basis for $C(A)$

⇒ Because there a r pivot columns, then $\dim C(A) = r$.

Second way of finding C(A)

Fezzik:

$$C(A) = \left\{ \vec{y} \in \mathbb{R}^3 \;\middle|\; \vec{y} = c_1 \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 6 \\ 12 \end{bmatrix},\ c_1, c_2 \in \mathbb{R} \right\}$$

basis for $C(A)$: $\left\{ \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 12 \end{bmatrix} \right\}$

E11cp2

Inigo:

$$\mathbb{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix} \sim \mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

(columns of $\mathbb{R}_{\mathbb{A}}$ marked P F F; first column of $\mathbb{A}$ circled: pivot column)

Basis for $C(\mathbb{A}) = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

$\dim C(\mathbb{A}) = r = 1$

Notes: $C(\mathbb{A}) \neq C(\mathbb{R}_{\mathbb{A}})$ ← in general

$N(\mathbb{A}) = N(\mathbb{R}_{\mathbb{A}})$

↑ about $\vec{x}$'s

($\mathbb{A}\vec{x} = \vec{b}$ has same solutions as $\mathbb{R}_{\mathbb{A}}\vec{x} = \vec{d}$)

Why the dimension of $N(A)$ is $n-r$

- see /E9fp1 (p59ish)

Big Deal:

Our one true method of finding nullspace always produces a set of vectors that are linearly independent and are therefore a basis for $N(A)$

Inigo: ← vectors span $N(A)$

$$\mathbb{N} = \begin{matrix} P \\ F \\ F \end{matrix} \begin{bmatrix} -2 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ \multicolumn{2}{c}{I} \end{bmatrix}$$

I appears in free variable **rows**

Fezzik:

$$\mathbb{N} = \begin{matrix} P \\ F \\ P \\ F \end{matrix} \begin{bmatrix} -2 & 1 \\ 1 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$n-r$ = # free variables (monks)

b/c we express pivot variables in terms of the free variables when finding $N(A)$ (always)

/E11dp1

$\mathbb{N}$ always has $n-r$ columns that are linearly independent

$\Rightarrow$ form a basis for $N(A)$

$\Rightarrow$ $\dim N(A) = n-r$.

Inigo: $\dim N(A) = 3-1 = 2$ ✓

Fezzik: $\dim N(A) = 4-2 = 2$ ✓

"It came from Row Space!"

E12ap1

- the row space of $\mathbb{A}$ is a thing
- what this means for $\mathbb{A}\vec{x} = \vec{b}$
- Many big deals
- The Big Picture

Story: Row Space of $\underset{m\times n}{\mathbb{A}}$ = all linear combinations of the rows of $\mathbb{A}$.

= subspace of $\mathbb{R}^n$ ← where $N(\mathbb{A})$ lives

Contrast: $C(\mathbb{A})$ = subspace of $\mathbb{R}^m$

Big Deal

BD #1: If $\vec{x} \in \mathbb{A}$'s Row Space, then $\mathbb{A}\vec{x} \neq \vec{0}$ unless $\vec{x} = \vec{0}$

← $\vec{x} \notin N(\mathbb{A})$

Example Inigo:

$\mathbb{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}$ ⇒ clearly, Row Space of $\mathbb{A}$

$= \left\{ \vec{x} \in \mathbb{R}^m \;\middle|\; \vec{x} = c\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},\ c \in \mathbb{R} \right\}$

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix} \times c\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = c\begin{bmatrix} 6 \\ 12 \end{bmatrix} = 6c\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in C(\mathbb{A})$$

↑ has to be true

Recall $\mathbb{A}\vec{x} = \vec{b}$

$$\vec{x} = \vec{x}_p + \vec{x}_h = \vec{x}_r + \vec{x}_n$$

$\vec{x}_p$ ↘ particular; $\vec{x}_h$ ↗ homogeneous; $\vec{x}_r$ ↘ row; $\vec{x}_n$ ↘ null

note $\vec{x}_p \neq \vec{x}_r$ necessarily

$$\mathbb{A}(\vec{x}_p + \vec{x}_h) = \underbrace{\mathbb{A}\vec{x}_p}_{\vec{b}} + \underbrace{\mathbb{A}\vec{x}_h}_{\vec{0}}$$

$\vec{x}_h$ ← may be infinitely many $\in N(\mathbb{A}) \neq \{\vec{0}\}$

↑ $\vec{x}_p$ must partly live in row space of $\mathbb{A}$.

BD #2

Any $\vec{x}$ in Row Space of $\mathbb{A}$ is $\perp$ / orthogonal / at right angles to any $\vec{x}$ in Null space of $\mathbb{A}$.

$$\underset{\mathbb{A}}{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}} \left(\underbrace{c_1 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{\in N(\mathbb{A}) \text{ from before}} \right) = \vec{0}$$

BD#3 The row space of $\mathbb{A}$ is the same as the row space of $\mathbb{R}_{\mathbb{A}}$

⇒ beautiful basis for row space of $\mathbb{A}$ = non-zero rows of $\mathbb{R}_{\mathbb{A}}$.

ex Inigo: $\mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ → basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$

Fezzik: $\mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ → basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

↖ rows are linearly independent b/c of $\mathbb{I}$ sitting in pivot columns

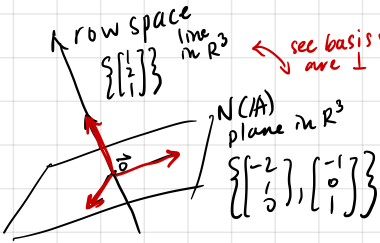
E12ap2

BD#4 Dim Row Space of $\mathbb{A}$
$= \text{Dim Row Space of } \mathbb{R}_{\mathbb{A}}$ (because samesies)
$= r$ = non-zero rows

(same as dim $C(\mathbb{A})$) amazing!!

BD#5 Dims of Row Space of $\mathbb{A}$ and Nullspace of $\mathbb{A}$ add up to n $(= r + (n-r))$.

Inigo:



see basis vectors are ⊥ to each other

dim RowSpace of $\mathbb{A} = r = 1$

dim $N(\mathbb{A}) = n - r = 3 - 1 = 2$

$1 + 2 = 3$

(soon: "othogonal complements")

(Imagine loud organ music and lightning)

Row Space & $N(\mathbb{A})$ neatly divide up R^n

BD#6

Consider $\mathbb{A}^T$ for Inigo

$$\mathbb{A}^T = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}$$

⇒ now see Row Space of $\mathbb{A}$ is also the Column Space of $\mathbb{A}^T$

wow!

Notation:

$C(\mathbb{A}^T)$ = Row Space of $\mathbb{A}$

cat in a box:

repurpose b/c deep connection.

BD#7: We find a 3rd and final and totally bestest way for finding $C(\mathbb{A})$.

Find $\mathbb{R}_{\mathbb{A}^T}$ and then read off basis vectors for $C(\mathbb{A})$

row space of $\mathbb{A}^T$ = column space of $\mathbb{A}$

awesome!!

Note: $\mathbb{R}_{\mathbb{A}^T} \neq (\mathbb{R}_{\mathbb{A}})^T$ in general — eek!!

E12ap3

ex: $\mathbb{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}$ $\quad \mathbb{A}^T = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}$

Inigo:

$\mathbb{A}^T \sim \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ is a basis for $C(\mathbb{A})$

$R_2' = R_2 - \left(\frac{2}{1}\right) R_1$

$R_3' = R_3 - \left(\frac{1}{1}\right) R_1$

Fezzik:

$\mathbb{A}^T = \begin{bmatrix} 2 & 2 & 6 \\ 4 & 4 & 12 \\ 3 & 6 & 12 \\ 4 & 10 & 18 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

row ops + tears

$\Rightarrow \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ is a basis for $C(\mathbb{A})$

$\mathbb{I}$ matrix now always appears

The Known Unknowns of Left Nullspace

- What Left Nullspace is
- Connection to the other three subspaces

E13ap1

We have $C(A)$, $N(A)$, and $C(A^T)$ (row space of A)

What about $N(A^T)$?

↑ Left Nullspace of A

Reason: if $\vec{y} \in N(A^T) \subset \mathbb{R}^m$

then $A^T \vec{y} = \vec{0}$ ← defn.

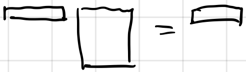
(A^T: $n \times m$; $\vec{y}$: $m \times 1$, $\in \mathbb{R}^m$; $\vec{0}$: $n \times 1$, $\in \mathbb{R}^n$ (where the $\vec{x}$'s are))

Take transpose of both sides:

$$(A^T \vec{y})^T = (\vec{0})^T \quad \leftarrow (BC)^T = C^T B^T$$

$$\vec{y}^T A = \vec{0}^T$$

($\vec{y}^T$: $1 \times m$ row vector; A: $m \times n$; $\vec{0}^T$: $1 \times n$ row vector)



$\vec{y}^T$ multiplies A from the left

So in fact $N(A)$ is the Right Nullspace of A

$$A \vec{x} = \vec{0}$$

↑ on the right.

Know immediately:

$$\dim N(A^T) = \#\text{ columns of } A^T - \text{rank of } A^T$$
$$= m - r$$

(for $N(A)$, we have $n-r$)

We find $N(A^T)$ just as we would find $N(A)$

Solve $A^T \vec{y} = \vec{0}$

($\vec{y}$ in $\mathbb{R}^m$)

Ex: Intyo.

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 1 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} \overset{P}{1} & \overset{F}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

↑ A^T ↑ $\vec{0}$

Express pivot vars in terms of free

$$y_1 + 2y_2 = 0 \Rightarrow y_1 = -2y_2$$

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -2y_2 \\ y_2 \end{bmatrix} = y_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \; y_2 \in R$$

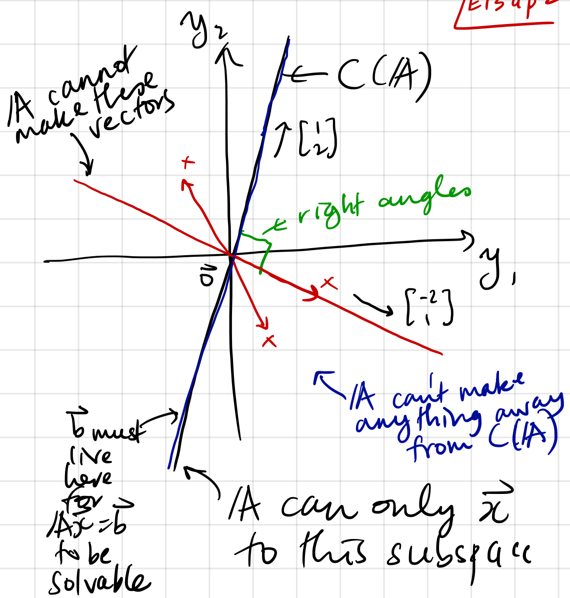
↑ basis vector

Just as $N(A)$ & $C(A^T)$ divide up R^n so too do $N(A^T)$ & $C(A)$

$N(A^T)$ ↕ basis $\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$

$C(A)$ ↕ basis $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

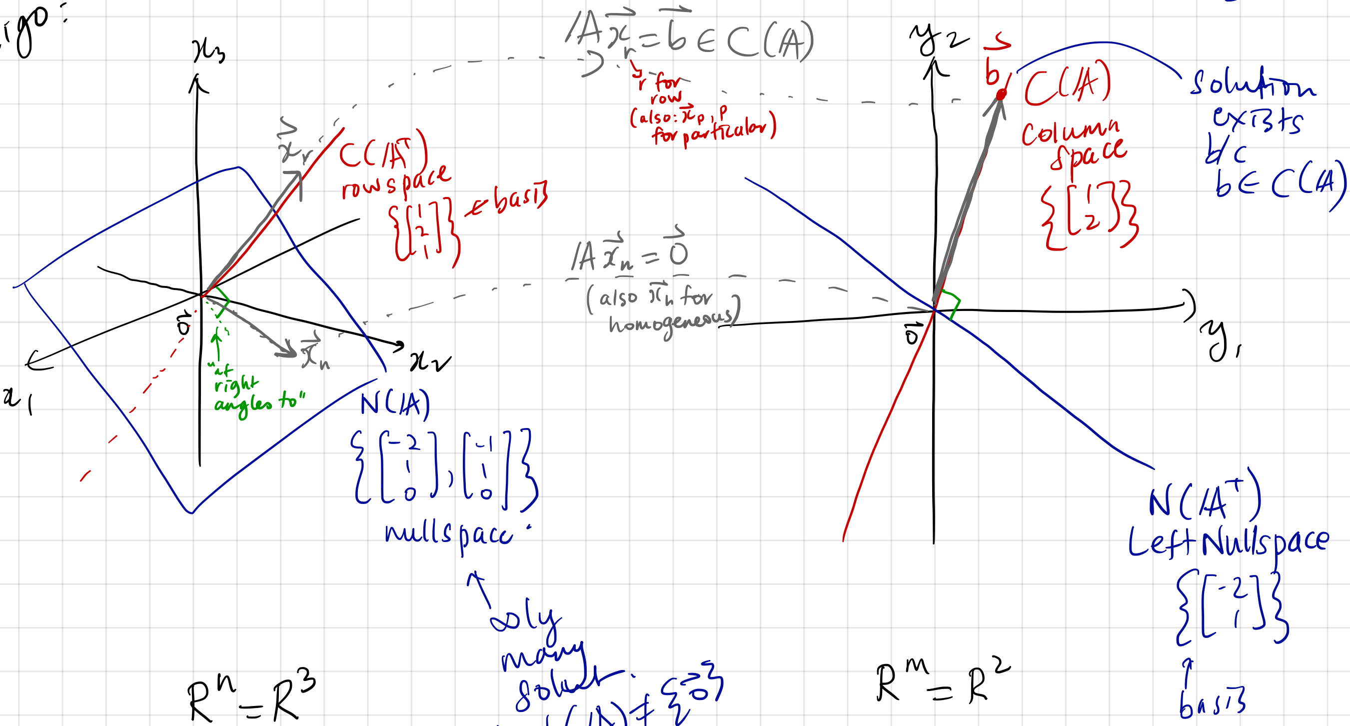
see these are 1-d orthogonal.



The Fundamental Theorem of Matrixology (almost)

$$A\vec{x} = A(\vec{x}_r + \vec{x}_n) = \vec{b}$$

Big picture for Inigo:



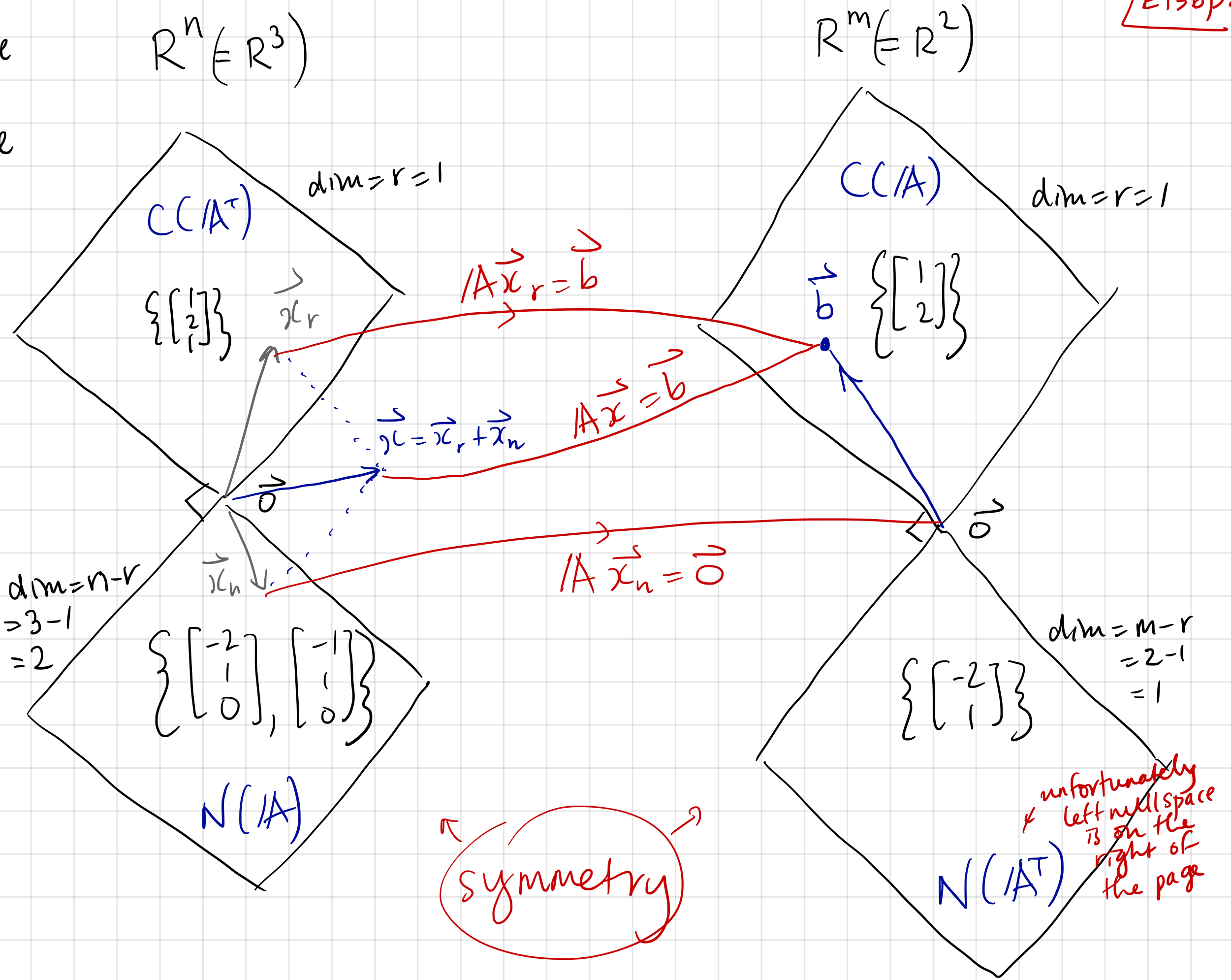
key:

Inigo sends a line to a line

Later: see Inigo ~ $\sqrt{30}$

E13bp2

Abstract Big picture with Inigo's structure



Definitions we need:

(1) (old) if $\vec{x}^T\vec{y} = 0$ we say $\vec{x}$ & $\vec{y}$ are orthogonal

$\vec{y}^T\vec{x}$ (//)

(2) We say two supspaces S_1 & S_2 are orthogonal if all vectors in S_1 are orthogonal to all vectors in S_2

(3) If two subspaces (S_1 & S_2) in a vector space V (dim n) are orthogonal and their dimensions add to n, we say S_1 & S_2 are orthogonal complements of each other

Notation: S and $S^\perp$

and $S \oplus S^\perp = V$

⇑

any vector in V can be written as a sum of a vector in S and a vector in $S^\perp$

E13bp3

Fundamental Theorem of Matrixology (mostly):

- $\dim C(A) = r$ — column space
- $\dim N(A^T) = m - r$ — left nullspace
- $\dim C(A^T) = r$ — row space
- $\dim N(A) = n - r$ — nullspace
- $C(A)$ and $N(A^T)$ are orthogonal complements in R^m
- $C(A^T)$ and $N(A)$ are orthogonal complements in R^n
- The bases of $C(A)$ & $N(A^T)$ combine to give a basis of R^m
- The bases of $C(A^T)$ & $N(A)$ combine to give a basis of R^n

More near the end of course

E14ap1

Matrix-fu

The Man in Black, Westley:

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

2×2 (m, n)

$m=2$ rows
$n=2$ cols

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

$R_2' = R_2 - \left(\frac{3}{1}\right) R_1$
$R_1' = R_1$

$$\begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} = \mathbb{R}_{\mathbb{A}}$$

P (under column 1), F (under column 2)

see rank $r=1$

x_1 is a pivot variable
x_2 is a free variable

$m=2, n=2, r=1$

$$\mathbb{A}^T = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}$$

$R_2' = R_2 - \left(\frac{-2}{1}\right) R_1$

$$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} = \mathbb{R}_{\mathbb{A}^T}$$

y_1 y_2

Dimensions:

$\dim C(\mathbb{A}) = r = \dim C(\mathbb{A}^T)$

column space; $r = 1$; row space

$\dim N(\mathbb{A}) = n - r = 2-1 = 1.$

$\dim N(\mathbb{A}^T) = m - r = 2-1 = 1.$

(Right) Nullspace; Left Nullspace

Bases:

Nullspaces:

$\mathbb{A}\vec{x} = \vec{0} \iff \mathbb{R}_{\mathbb{A}}\vec{x} = \vec{0}$

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

$x_1 - 2x_2 = 0$

P, F

$\Rightarrow x_1 = 2x_2$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

replace pivot variables

$x_2 \in \mathbb{R}$

$N(\mathbb{A}) =$

$$\left\{ \vec{x} = x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \; x_2 \in \mathbb{R} \right\}$$

∞ly many points

A basis for $N(\mathbb{A})$ is $\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$

Also good

$\begin{bmatrix} 2\infty \\ \infty \end{bmatrix}$, $\begin{bmatrix} -2 \\ -1 \end{bmatrix}$, $\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (unit vector)

Left Nullspace:

Solve $\mathbb{A}^T \vec{y} = \vec{0}$

$\iff \mathbb{R}_{\mathbb{A}^T} \vec{y} = \vec{0}$

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

$\Rightarrow y_1 + 3y_2 = 0.$

(y_1: P, y_2: F)

$y_1 = -3y_2$

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -3y_2 \\ y_2 \end{bmatrix} = y_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

(y_2 free)

$$N(\mathbb{A}^T) = \left\{ \vec{y} = y_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix}, y_2 \in \mathbb{R} \right\}$$

A basis for $N(\mathbb{A}^T)$ is

$$\left\{ \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right\}.$$

Column space $C(\mathbb{A})$.

(1) of 3. Solve $\mathbb{A}\vec{x} = \vec{b}$ for general $\vec{b}$.

$$\left[\begin{array}{cc|c} 1 & -2 & b_1 \\ 3 & -6 & b_2 \end{array}\right]$$

$$\overset{\sim}{R_2' = R_2 - \left(\tfrac{3}{1}\right)R_1} \left[\begin{array}{cc|c} 1 & -2 & b_1 \\ 0 & 0 & b_2 - 3b_1 \end{array}\right]$$

$0 = b_2 - 3b_1$ must hold if $\vec{b} \in C(\mathbb{A})$.

$\Rightarrow b_2 = 3b_1$

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ 3b_1 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$C(\mathbb{A}) = \left\{ \vec{y} = b_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix}, b_1 \in \mathbb{R} \right\}.$$

(1-d)

basis: $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$.

(2) of 3. Find pivot columns through $\mathbb{R}_{\mathbb{A}}$

$\Rightarrow$ Same columns in $\mathbb{A}$ form a basis for $C(\mathbb{A})$

here 1st column: $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$

(3) of 3. Take non-zero rows of $\mathbb{R}_{\mathbb{A}^T} = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$

Again: $\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$.

best way

E14ap2

Row space: take non-zero rows of $\mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$

$\Rightarrow \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$

E14 ap3

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

Bases $\mathbb{R}^m \to C(\mathbb{A})$: $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $N(\mathbb{A}^T)$: $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$

$\mathbb{R}^n \to C(\mathbb{A}^T)$: $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $N(\mathbb{A})$: $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\mathbb{A}\vec{x}$: ① $\mathbb{A}$ sends any multiple of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ to some multiple of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

② $\mathbb{A}$ sends any multiple of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (N(A)) to $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

③ $\mathbb{A}$ cannot make any vector which has some non-zero part of $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$.

Complementary Orthogonality of Subspaces

$C(\mathbb{A}) \oplus N(\mathbb{A}^T) = \mathbb{R}^2 \leftarrow m=2$

$C(\mathbb{A}^T) \oplus N(\mathbb{A}) = \mathbb{R}^2 \leftarrow n=2$.

$\begin{bmatrix} 1 & 3 \end{bmatrix}\begin{bmatrix} -3 \\ 1 \end{bmatrix} = 0$ $\quad$ $\begin{bmatrix} 1 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 0$.

how $\mathbb{A}$ functions:

$\mathbb{A}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

in C(A)

length $= \sqrt{1^2 + (-2)^2} = \sqrt{5}$.

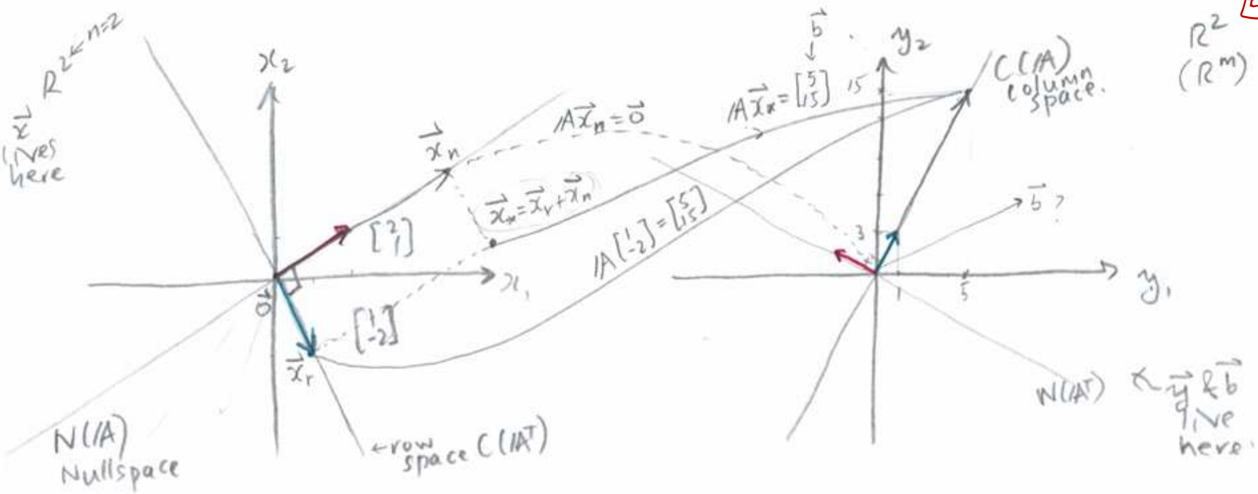
length $5\sqrt{1+3^2} = 5\sqrt{10}$

Stretch factor: $\frac{5\sqrt{10}}{\sqrt{5}} = \sqrt{5}\sqrt{10} = 5\sqrt{2}$.

So Westley is like $y = 5\sqrt{2}\,x$ (1×1, $r=1$)

but only for vectors in row & column space (1-d subspace).

$\mathbb{A}$ is $r\times r$ invertible in these subspaces

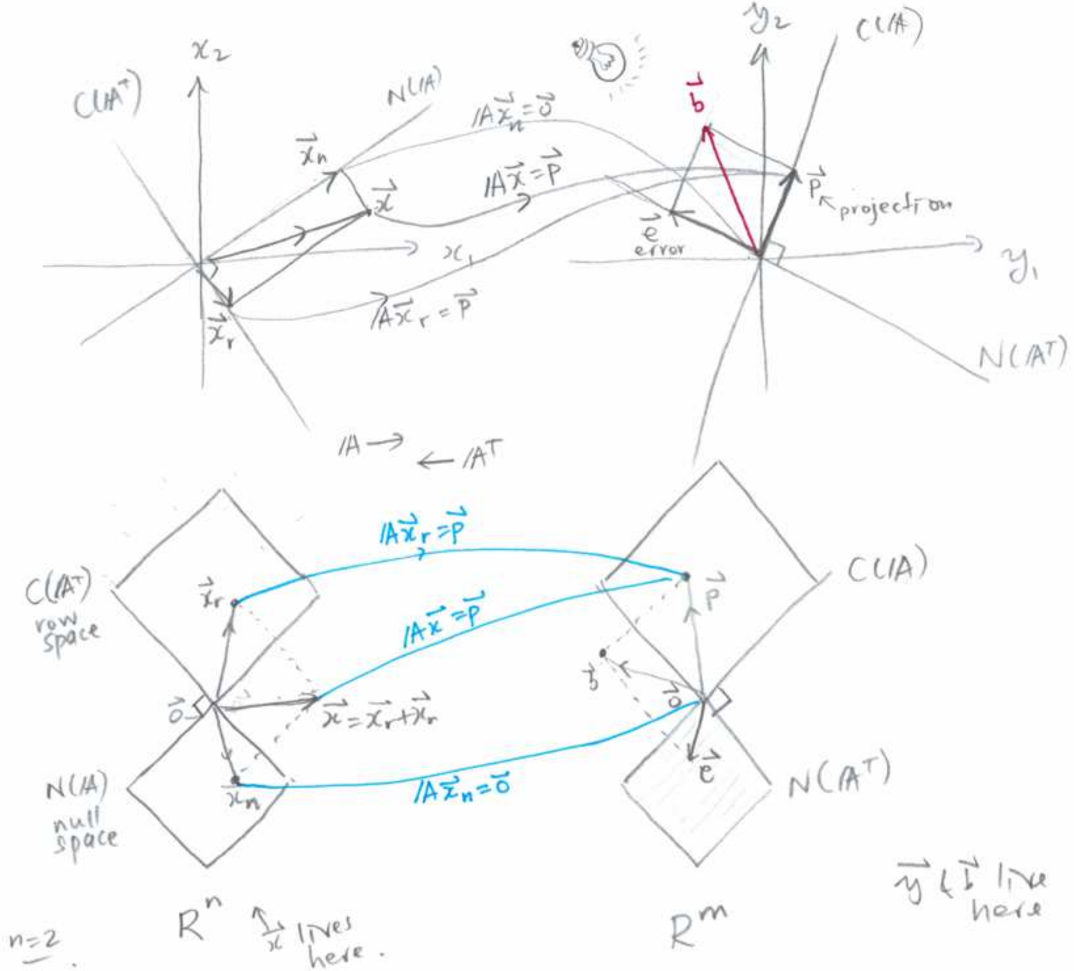


- $A\vec{x}=\vec{b}$ is solvable if $\vec{b} \in C(A)$.
- if $\vec{b} \in C(A)$, then there is one solut. if $N(A)=\{\vec{0}\}$ and ∞ly many otherwise. $\dim N(A) \geq 1$.

Ex// if $\vec{b} = 5\begin{bmatrix}1\\3\end{bmatrix}$ then $\vec{x} = \underbrace{\begin{bmatrix}1\\-2\end{bmatrix}}_{C(A^T)}^{\vec{x}_r} + c\underbrace{\begin{bmatrix}2\\1\end{bmatrix}}_{N(A)}^{\vec{x}_n}$, $c \in R$.

Ex if $\vec{b} = 5\begin{bmatrix}1\\3\end{bmatrix} + 4\begin{bmatrix}-3\\1\end{bmatrix} = \begin{bmatrix}-7\\19\end{bmatrix}$, no solutn.
$\in N(A^T)$
#inconceivable

E14ap5



x2
C(A^T)
N(A)
x_n
x
x1
x_r
A x_n = 0
A x = p
A x_r = p
b
y2
C(A)
p ← projection
e
error
y1
N(A^T)
A →
← A^T
A x_r = p
C(A^T)
row space
x_r
A x = p
0
x = x_r + x_n
N(A)
null space
x_n
A x_n = 0
C(A)
p
y
0
e
N(A^T)
R^n
x lives here.
n=2
R^m
y & b live here

Inigo:

$\underset{2\times 3}{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{bmatrix}$ $m=2$, $n=3$

$\underset{3\times 2}{\mathbb{A}^T} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}$

$\mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (P F F) $r=1$

$\mathbb{R}_{\mathbb{A}^T} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ (P F)

Bases

$C(\mathbb{A}): \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

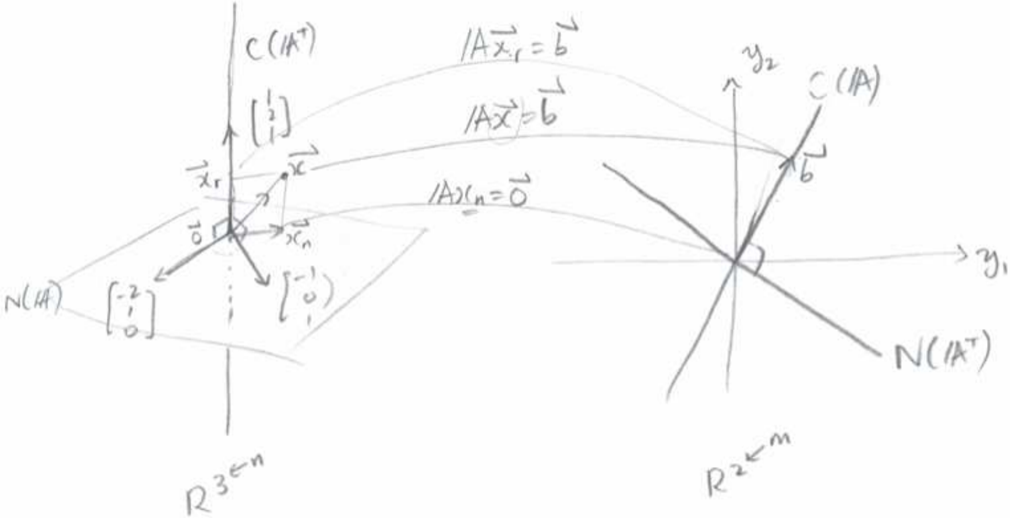
$C(\mathbb{A}^T): \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$

$N(\mathbb{A}^T): \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$ $m-r=2-1=1$

$N(\mathbb{A}): \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$

↗ dim: $n-r = 3-1 = 2$

not a beautiful basis #morelater.



row space ⟷ colspace.

Inigo is "1×1" matrix, equivalent to $\sqrt{30}$.
($r \times r$)

E14cp1

Fezzik:

$$\mathbb{A}_{3\times4} = \begin{bmatrix} 2 & 4 & 3 & 4 \\ 2 & 4 & 6 & 10 \\ 6 & 12 & 12 & 18 \end{bmatrix} \quad m=3,\ n=4$$

$$\mathbb{A}^T_{4\times3} = \begin{bmatrix} 2 & 2 & 6 \\ 4 & 4 & 12 \\ 3 & 6 & 12 \\ 4 & 10 & 18 \end{bmatrix}$$

$$\mathbb{R}_{\mathbb{A}} = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad r=2$$

P F P F

$$\mathbb{R}_{\mathbb{A}^T} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

P P F

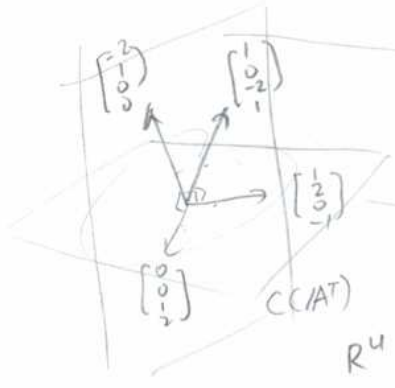
bases

$C(\mathbb{A})$: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ dim = r = 2

$N(\mathbb{A}^T)$: $\left\{ \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \right\}$ dim m−r = 3−2 = 1

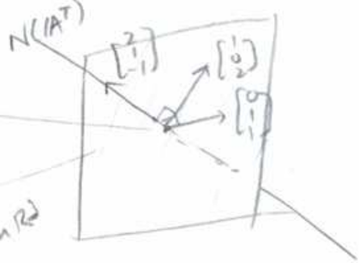
$C(\mathbb{A}^T)$: $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$ dim r = 2

$N(\mathbb{A})$: $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \right\}$ dim n − r = 4 − 2 = 2



$\mathbb{A}\vec{x}_n = \vec{0}$

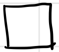
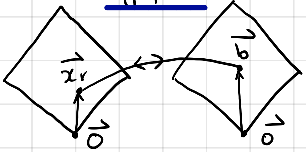
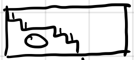
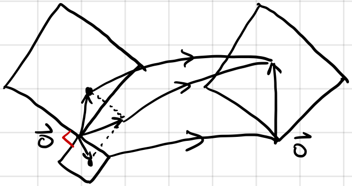
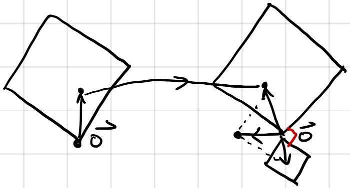
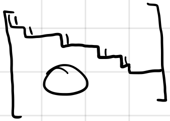
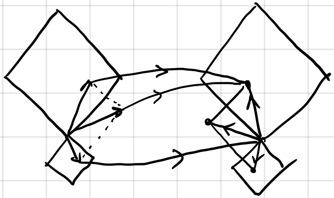
$\mathbb{A}\vec{x}_r = \vec{b}$ maps 2-d plane in R^4 to 2-d plane in R^3



R^4 (eek).

$R^3 \leftarrow m$.

Everything hinges on $\mathbb{R}_A$ & $\mathbb{R}_{A^T}$ // Four main kinds of A.

Shape/rank: (full rank)	big picture: (m, n, r)	# solutions to $A\vec{x}=\vec{b}$:	dim $N(A)$: (how many solutions)	dim $N(A^T)$: (whether solution is possible)
$m=n=r$ square  invertible		1	0	0
$m=r$, $n>r$  wide		∞	$\geqslant 1$	0
$m>r$, $n=r$  tall		1 or 0 $\Updownarrow$ $b \in C(A)$	0	$\geqslant 1$
 $m, n > r$		∞ or 0	$\geqslant 1$	$\geqslant 1$

Projections:

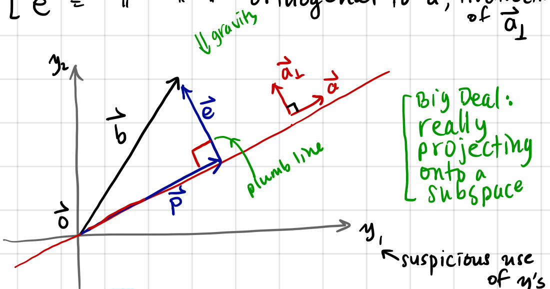
menu:
- Project a vector onto a line
- Notion of an error vector $\vec{e}$
- Goal: Handle $A\vec{x}=\vec{b}$ when no solutions are possible. Big idea: Best approximation

---

Idea: Given a vector $\vec{b}$ and a direction described by a vector $\vec{a}$, break $\vec{b}$ into two components $\vec{b}=\vec{p}+\vec{e}$:

- $\vec{p}$ = piece of $\vec{b}$ in direction of $\vec{a}$
- $\vec{e}$ = " " " orthogonal to $\vec{a}$, in direction of $\vec{a}_\perp$

Picture: (for $\mathbb{R}^2$ but works in $\mathbb{R}^m$)



Big Deal: really projecting onto a subspace

- $\vec{p}$ = projected component
- $\vec{e}$ = error

One reason for doing this:

In solving $A\vec{x}=\vec{b}$, if $\vec{b}\notin C(A)$, we can still solve $A\vec{x}_*=\vec{p}$ where $\vec{p}$ is $\vec{b}$'s projection onto Column Space.

- Best Approximation
- Left Nullspace will matter!

How to find $\vec{p}$ & $\vec{e}$ given $\vec{b}$ & $\vec{a}$:

/E15ap1

We want $\vec{p} \parallel \vec{a}$ and $\vec{e}\perp\vec{a}$

Mathematically: $\vec{p}=x_*\vec{a}$ (↑ some number $\in\mathbb{R}$) $\qquad \vec{e}^T\vec{a}=\vec{a}^T\vec{e}=0$ (↖ inner (dot) product)

Monks whisper: "Use the orthogonality..."

$$\vec{b}=\vec{p}+\vec{e}$$

Sneakiness:

$$\vec{a}^T(\vec{b}) = \vec{a}^T(\vec{p}+\vec{e})$$

$$\underbrace{\vec{a}^T\vec{b}}_{\text{number}} = \vec{a}^T\vec{p}+\vec{a}^T\vec{e} \quad (\vec{a}^T\vec{e}\to 0)$$

$$= \vec{a}^T(x_*\vec{a})$$

$$= x_*\underbrace{\vec{a}^T\vec{a}}_{\text{number}}$$

$$\Rightarrow x_* = \frac{(\vec{a}^T\vec{b})}{(\vec{a}^T\vec{a})}$$

$$\vec{p}=x_*\vec{a}=\underbrace{\frac{(\vec{a}^T\vec{b})}{(\vec{a}^T\vec{a})}}_{\text{some scaling of }\vec{a}}\vec{a}$$

$$\vec{e}=\vec{b}-\vec{p}$$

done.

E15ap2

Example:

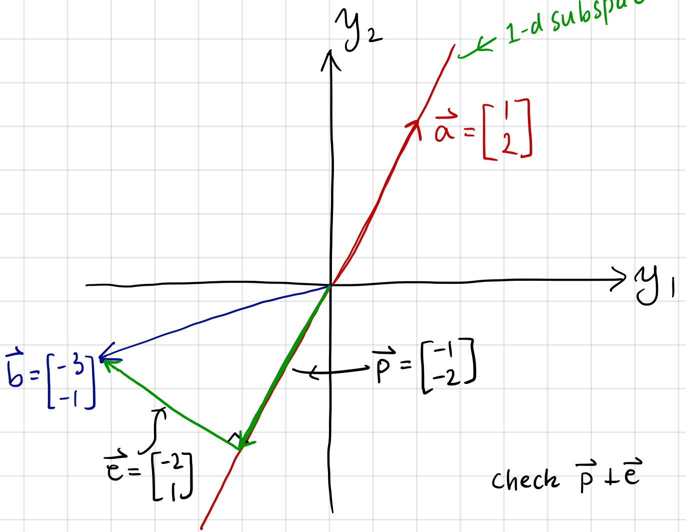
project $\vec{b} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$ onto $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ — direction is all that matters

$$x_* = \frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}} = \frac{\begin{bmatrix} 1 & 2 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix}} = \frac{-5}{5} = -1$$

$$\Rightarrow \vec{p} = x_*\vec{a} = (-1)\vec{a} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

$$\Rightarrow \vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} -3 \\ -1 \end{bmatrix} - \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

note: $\vec{p} \perp \vec{e}$ as required



check $\vec{p} \perp \vec{e}$

More Sneakiness:

We have $\vec{p} = \left(\frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}\right)\vec{a}$ — the bracket is 1×1, a number; $\vec{a}$ is $m\times1$

← not great to have to recalculate for each $\vec{b}$ → really want a gadget matrix that projects $\vec{b}$ as an operator

$$= \frac{(\vec{a}^T\vec{b})}{(\vec{a}^T\vec{a})}\vec{a} = \frac{1}{(\vec{a}^T\vec{a})}\cdot(\vec{a}^T\vec{b})\,\vec{a}$$

($\vec{a}^T\vec{b}$: 1×1; $\vec{a}^T\vec{a}$: 1×1; $\vec{a}$: $m\times1$)

$$= \frac{1}{(\vec{a}^T\vec{a})}\vec{a}(\vec{a}^T\vec{b}) = \frac{1}{(\vec{a}^T\vec{a})}(\vec{a}\,\vec{a}^T)\vec{b}$$

($\vec{a}$: $m\times1$, $\vec{a}^T$: $1\times m$, $\vec{b}$: $m\times1$; $\vec{a}\,\vec{a}^T$: outer product, $m\times m$ square)

$$= \frac{1}{\|\vec{a}\|^2}(\vec{a}\,\vec{a}^T)\vec{b} = \left(\frac{\vec{a}}{\|\vec{a}\|}\frac{\vec{a}^T}{\|\vec{a}\|}\right)\vec{b}$$

($\|\vec{a}\|^2$ = (length of $\vec{a}$)^2; bracket is $m\times m$; $\vec{b}$ is $m\times1$)

$$= \hat{a}\,\hat{a}^T\vec{b} = \mathbb{P}_{\hat{a}}\vec{b}$$

← good: length of $\vec{a}$ cannot matter

$\hat{a}$: unit vector; $\hat{a}\hat{a}^T$: outer product, $m\times m$; $\vec{b}$: $m\times1$

$\mathbb{P}_{\hat{a}}$: **Projection Operator**

E15pa3

Example again:

project $\vec{b} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$ onto $\vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

make unit vector

$$\hat{a} = \frac{1}{\|\vec{a}\|}\vec{a} = \frac{1}{\sqrt{1^2+2^2}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\mathbb{P}_{\hat{a}} = \hat{a}\hat{a}^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

symmetry is guaranteed

So:

$$\vec{p} = \mathbb{P}_{\hat{a}}\vec{b} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -5 \\ -10 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} \checkmark$$

$\vec{e} = \vec{b} - \vec{p}$ as before

Bonus:

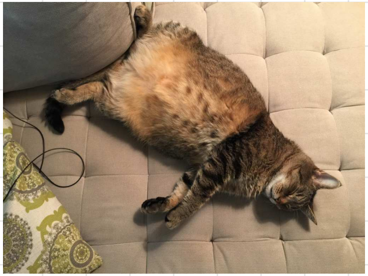
$$\vec{e} = \vec{b} - \vec{p} = \underset{m\times1}{\vec{b}} - \underset{m\times m}{\mathbb{P}_{\hat{a}}}\underset{m\times1}{\vec{b}}$$

$$= \underset{m\times m}{\mathbb{I}}\underset{m\times1}{\vec{b}} - \underset{m\times m}{\mathbb{P}_{\hat{a}}}\underset{m\times1}{\vec{b}} = (\underset{m\times m}{\mathbb{I}} - \underset{m\times m}{\mathbb{P}_{\hat{a}}})\underset{m\times1}{\vec{b}}$$

~~$(\underset{1\times1}{1} - \underset{m\times m}{\mathbb{P}_{\hat{a}}})$~~ wrong.

$(\mathbb{I} - \mathbb{P}_{\hat{a}})$: Extracts $\vec{e}$ part of $\vec{a}$.

much happiness over $\mathbb{P}_{\hat{a}}$



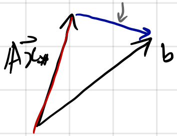
The Amazing Normal Equation:

Menu:
- Find the best approximation to $A\vec{x}=\vec{b}$ when $\vec{b}\notin C(A)$

Before: We just gave up when $A\vec{x}=\vec{b}$ had no solution ⇐ betrayed a lack of ticker

New plan: find $\vec{x}_*$ so that $A\vec{x}_*$ is as close to $\vec{b}$ as possible. (↑ $*$ denotes approximation)

Mathematically: we want $\vec{x}^*$ that minimizes $\|\vec{b}-A\vec{x}_*\|$ ← distance between $\vec{b}$ and $A\vec{x}_*$



Big idea: See $\vec{b}=\vec{p}+\vec{e}$ where

$\vec{p}\in C(A)$
& $\vec{e}\in N(A^T)$ } guarantees $\vec{p}\perp\vec{e}$

We project $\vec{b}$ onto $C(A)$ and solve $A\vec{x}_*=\vec{p}$ instead

How? E15 bp1

Same approach as for simple projections:

We want
$\vec{b}=\vec{p}+\vec{e}$ where $A\vec{x}_*=\vec{p}$ & $A^T\vec{e}=\vec{0}$
(①, ②, ③)

Start with $A^T\vec{e}=\vec{0}$ ← Monks ③

$$\vec{0}=A^T\vec{e}=A^T(\vec{b}-\vec{p})=A^T\vec{b}-A^T\vec{p}$$
(③, ①)

$$\Rightarrow A^T\vec{b}=A^T\vec{p}=A^T(A\vec{x}_*)$$
(② ; group. ; $A\vec{x}_*$ means $\vec{p}$ is some linear combination of A's columns)

Switch sides:

$$(A^TA)\vec{x}_*=A^T\vec{b}$$

A^TA: square, symmetric = awesome; A^T: $n\times m$, A: $m\times n$, A^TA: $n\times n$; $\vec{x}_*$: $n\times 1$ $\in\mathbb{R}^n$; A^T: $n\times m$, $\vec{b}$: $m\times 1$, $A^T\vec{b}$: $n\times 1$ $\in\mathbb{R}^n$

⇑ incredible! always solvable!

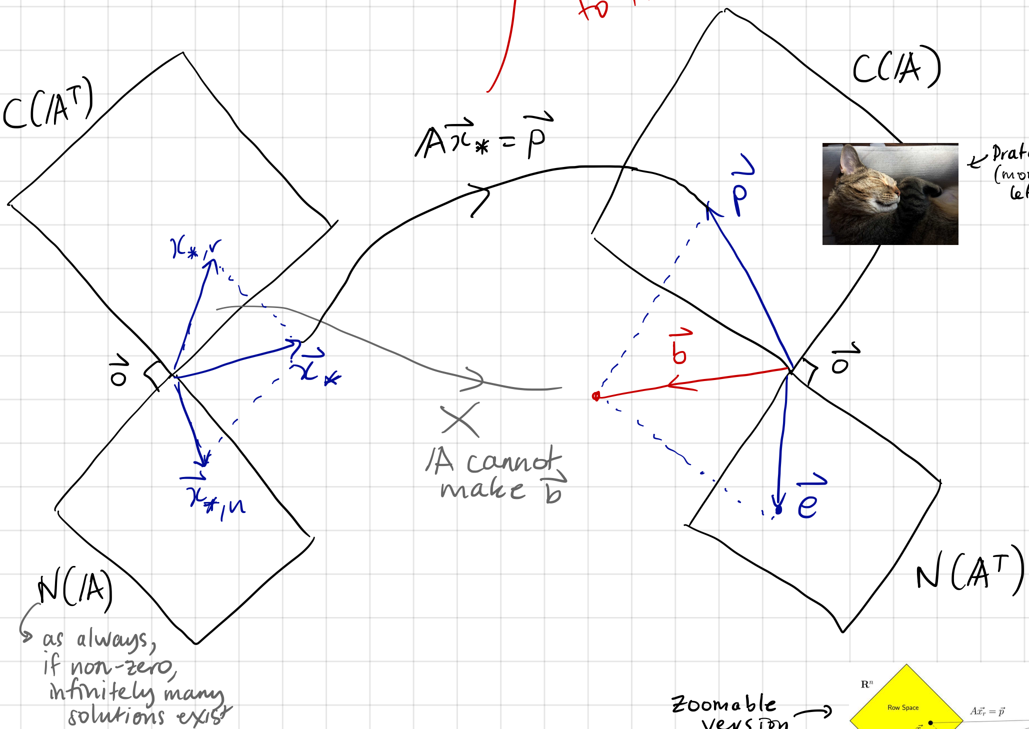
of the form:

$$A'\vec{x}_*=\vec{b}'$$
(A': $n\times n$, $\vec{x}_*$: $n\times 1$, $\vec{b}'$: $n\times 1$) ← prime

Abstract Big Picture

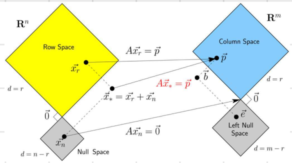
$$A^T A \vec{x}_* = A^T \vec{b}$$

solve normal equations to find $\vec{x}_*$



← Pratchett (more of a left nullspace fan)

Zoomable version →



Example of using the Normal Equation

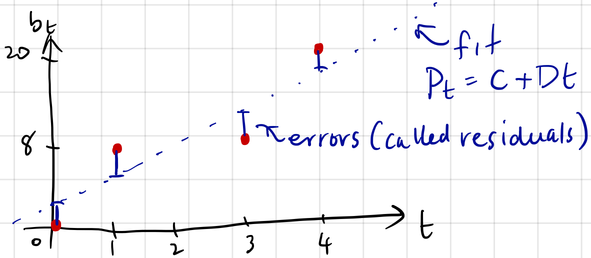
E15cp1

- fitting a straight line to a set of data points ← fundamental scientific activity!

Ex from Strang:

$b_t = 0, 8, 8, 20$

at times $t = 0, 1, 3, 4$



Want fit to be true:

$t=0$	$b_0 = 0 = 1C + D\times 0$
$t=1$	$b_1 = 8 = 1C + D\times 1$
$t=3$	$b_2 = 8 = 1C + D\times 3$
$t=4$	$b_3 = 20 = 1C + D\times 4$

Matrixify:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$$

$\begin{bmatrix} 0 \\ 1 \\ 9 \\ 16 \end{bmatrix}$ for $p_t = C + Dt + Et^2$; E

✓ A 4×2 ; ? $\vec{x}$ 2×1 ; ✓ $\vec{b}$ 4×1

Clear $\vec{b}$ is not in A's Column Space

Solve $A^TA\vec{x}_* = A^T\vec{b}$:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{\vec{x}_*} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}}_{\vec{b}}$$

always 2×2

$$\Rightarrow \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$$

line fit for 10^6 (ex) data pts ⇒ still 2×2 problem

$[A^TA \mid A^T\vec{b}]$

$R_2' = R_2 - \left(\frac{8}{4}\right) R_1$ $\quad \left[\begin{array}{cc|c} 4 & 8 & 36 \\ 0 & 10 & 40 \end{array}\right]$ ← U

Back substitute: $4C_* + 8D_* = 36$ ← $C_* = 1$

$10D_* = 40$ ← $D_* = 4$

$\vec{x}_* = \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ Best fit line $= p_t = 1 + 4t$

$\vec{e} = \vec{b} - \vec{p} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix}$; $\|\vec{e}\|^2 = 1^2 + 3^2 + 5^2 + 3^2 = 44$

The Normal Equation and the Man in Black

Westley, our hero: $\mathbb{A} = \begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix}$

Solve $\begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \leftarrow \vec{b} \notin C(\mathbb{A})$

$\left[\begin{array}{cc|c} 1 & -2 & 5 \\ 3 & -6 & 5 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -2 & 5 \\ 0 & 0 & -10 \end{array}\right] \leftarrow$ eek: $0 = -10$

$R_2 = R_2' - \left(\frac{3}{1}\right)R_1$

inconceivable! (or: no solution)

Time for the Normal Equation: $\mathbb{A}^T\mathbb{A}\vec{x}_* = \mathbb{A}^T\vec{b}$

$\mathbb{A}^T\mathbb{A}$ is always symmetric

$\mathbb{A}^T\mathbb{A} = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix} = \begin{bmatrix} 10 & -20 \\ -20 & 40 \end{bmatrix}$

$n\times m \; m\times n \to n\times n$

$\mathbb{A}^T\vec{b} = \begin{bmatrix} 1 & 3 \\ -2 & -6 \end{bmatrix}\begin{bmatrix} 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 20 \\ -40 \end{bmatrix}$

$n\times m \; m\times 1 \to n\times 1$

Now Solve

$\left[\begin{array}{cc|c} 10 & -20 & +20 \\ -20 & 40 & -40 \end{array}\right] \sim \left[\begin{array}{cc|c} 10 & -20 & +20 \\ 0 & 0 & 0 \end{array}\right]$ ← all good as promised

$R_2' = R_2 - \left(\frac{-20}{10}\right)R_1$

$R_1' = \frac{1}{10}R_1$ $\left[\begin{array}{cc|c} 1 & -2 & +2 \\ 0 & 0 & 0 \end{array}\right]$ (P F)

$\Rightarrow x_{*,1} - 2x_{*,2} = +2$ (P, F)

$\Rightarrow x_{*,1} = +2 + 2x_{*,2}$ (P, F)

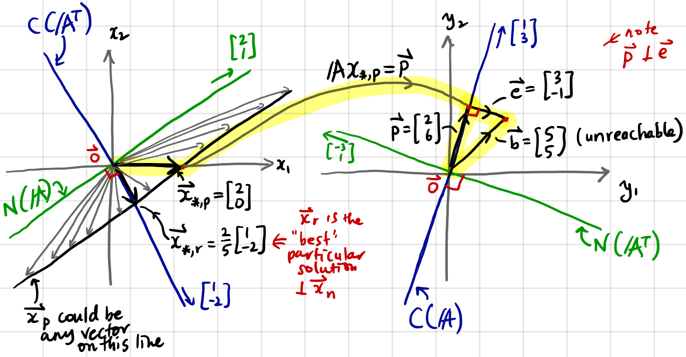
$\vec{x}_* = \begin{bmatrix} x_{*,1} \\ x_{*,2} \end{bmatrix} = \begin{bmatrix} +2 + 2x_{*,2} \\ 0 + x_{*,2} \end{bmatrix} = \begin{bmatrix} +2 \\ 0 \end{bmatrix} + x_{*,2}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$

(P, F; replace pivot vars in terms of free vars; $\vec{x}_{*,p}$ particular, $\vec{x}_{*,h}$ homogeneous)

where $x_{*,2} \in \mathbb{R}$

$\vec{p} = \mathbb{A}\vec{x}_* = \begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} +2 \\ 0 \end{bmatrix} + x_{*,2}\begin{bmatrix} 1 & -2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (crossed out → $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$)

$= \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ ← particular solution does the work



- Note $\vec{x}_p \neq \vec{x}_r$ in general; $\vec{x}_r$ is the special $\vec{x}_p$ that lives in A's row space
- How $\vec{b}$ was built:

$\vec{b} = 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} - 1\begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$

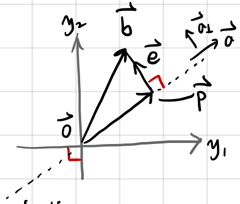
$\vec{p} + \vec{e}$

LE15ep1

Projecting a vector $\vec{b} \in R^m$ onto a subspace of R^m

#excitement

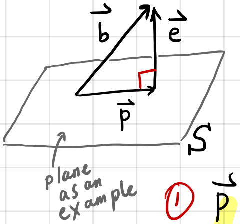
We know how to project a vector $\vec{b}$ onto a line defined by a vector $\vec{a}$:



$$\vec{p} = \underbrace{\underset{m\times 1}{\hat{a}}\ \underset{1\times m}{\hat{a}^T}}_{m\times m} \vec{b} = \mathbb{P}_{\hat{a}} \vec{b}$$

(outer product of unit vectors; $\mathbb{P}_{\hat{a}}$: Projection matrix/operator)

Now: Generalize to an r-dim subspace of R^m



We have ~~a~~ some basis for subspace S:

$$\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r\}$$

① $\underset{m\times 1}{\vec{p}} = \underset{m\times r}{A}\ \underset{r\times 1}{\vec{x}_*} = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_r \\ | & | & & | \end{bmatrix} \vec{x}_*$ $\leftarrow \vec{p} \subset C(A)$

(linearly independent because $\vec{a}_i$ form a basis)

② $A^T \vec{e} = \vec{0}$ $\leftarrow \vec{e} \in N(A^T)$

③ $\vec{b} = \vec{p} + \vec{e}$

Monks:

$$\vec{0} \overset{②}{=} A^T\vec{e} \overset{③}{=} A^T(\vec{b} - \vec{p}) = A^T\vec{b} - A^T\vec{p}$$

($\vec{p} = A\vec{x}_*$ ← ①)

$$\Rightarrow \vec{0} = A^T\vec{b} - A^T A\vec{x}_*$$

$$\Rightarrow A^T A \vec{x}_* = A^T\vec{b}$$

Solve for $\vec{x}_*$ then find $\vec{p}$ as $\vec{p} = A\vec{x}_*$

Special deal: extra tofu knives

b/c A's columns are linearly independent, A^TA is invertible (reason to follow)

$$\Rightarrow \underbrace{(A^TA)^{-1}(A^TA)}_{\mathbb{I}} \vec{x}_* = (A^TA)^{-1}A^T\vec{b}$$

(premultiply both sides by inverse)

$$\Rightarrow \vec{x}_* = (A^TA)^{-1}A^T\vec{b}$$

$$\Rightarrow \vec{p} = A\vec{x}_* = \underset{m\times r}{A}\ \underbrace{(\underset{r\times m}{A^T}\ \underset{m\times r}{A})^{-1}}_{r\times r}\ \underset{r\times m}{A^T}\ \underset{m\times 1}{\vec{b}}$$

$$= \underset{m\times m}{\mathbb{P}}\ \underset{m\times 1}{\vec{b}}$$

Projection Matrix (good for low dimensions)

More goodness: expect $\mathbb{P}^2\vec{b} = \mathbb{P}^3\vec{b} = \cdots = \vec{p}$

or $\mathbb{P} = \mathbb{P}^n$ for all $n \geq 1$

Check: $\mathbb{P}^2 =$

$$A\underbrace{(A^TA)^{-1}A^TA}_{\mathbb{I}}(A^TA)^{-1}A^T$$

$$= A(A^TA)^{-1}A^T = \mathbb{P}$$

cool! (right?)

Our Projection Matrix when we have a basis for S:

$$P = A(A^TA)^{-1}A^T; \quad P\vec{b} = \vec{p} \in S$$

Warning! $(A^TA)^{-1} \neq A^{-1}(A^T)^{-1}$ generally (may be equal sometimes)

A may be rectangular!!!

$\overset{r\times m}{A^T}\overset{m\times r}{A}$ ($r\times r$) is always square and symmetric

Important Truth:

A^TA is invertible iff ("if and only if") A's columns are linearly independent

$N(A^TA) = \{\vec{0}\}$, A^TA ($r\times r$) is full rank r.

$\iff A\vec{x} = \vec{0}$ only has $\vec{x} = \vec{0}$ as a solution

$\iff N(A) = \{\vec{0}\}$

Plan: Show A^TA & A have the same Nullspace always

Need to show that if $\vec{x} \in N(A)$ then $\vec{x} \in N(A^TA)$ and vice versa

LE15ep2

Assume $\vec{x} \in N(A)$: $A\vec{x} = \vec{0}$

$\Rightarrow A^T(A\vec{x}) = A^T(\vec{0})$

$\Rightarrow (A^TA)\vec{x} = \vec{0}$

So $\vec{x} \in N(A^TA)$

Second, if $\vec{x} \in N(A^TA)$ then

$A^TA\vec{x} = \vec{0}$ (by definition)

$\Rightarrow \overset{1\times r}{\vec{x}^T}(A^TA\vec{x}) = \vec{x}^T(\overset{r\times 1}{\vec{0}})$

$\Rightarrow \vec{x}^TA^TA\vec{x} = \underset{1\times 1}{0}$

we $(BC)^T = C^TB^T$

$\Rightarrow (A\vec{x})^T(A\vec{x}) = 0$

$\Rightarrow \|\overset{m\times r}{A}\overset{r\times 1}{\vec{x}}\|^2 = 0$ ($m\times 1$ vector)

if length $= 0$

$\Rightarrow A\vec{x} = \vec{0}$

So $\vec{x} \in N(A)$ // done

Note: we can do the same sort of thing for AA^T

Upshot: if $N(A) = \{\vec{0}\}$ then A^TA is invertible

(A need not be square) (A^TA: square $r\times r$ matrix with rank r)

Orthogonal and Orthonormal bases help us win friends and influence people

Menu: • Motivation for orthogonality
• orthogonal Matrices

Next: • the Gram-Schmidt Process
• What this all means for $A\vec{x}=\vec{b}$

Observation: We've been finding bases for our four fundamental subspaces, and we've so far taken whatever popped out.

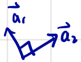
ex: $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ ← Basis for Fezzik's $C(A)$ → Describes 2-d subspace in $\mathbb{R}^3$

Does the job BUT we really like orthogonality in our bases and $\begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 2 \neq 0$

not orthogonal

ex 1 $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$ = basis for a plane in $\mathbb{R}^3$

$\vec{a}_1$, $\vec{a}_2$



$\vec{a}_1^T\vec{a}_2 = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = -2+0+2 = 0$ ✓

We call such a basis Orthogonal

E16ap1

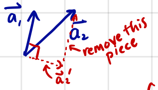
ex 2. (from Monks) $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix} \right\}$

$\vec{a}_1$, $\vec{a}_2$

$\vec{a}_1^T\vec{a}_2 = 1+3+14 = 18 \neq 0$

sneakiness

$\vec{a}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$



← $\vec{a}_2$ contains some of $\vec{a}_1$

[Big idea: Systematically turn a basis into an orthogonal basis by removing non-orthogonal pieces

[Everything is connected: Projections will do the work for us.

Claim: Orthogonality makes a basis a happy basis ← math

• Main reason: Representation of vectors is very clean.

Information contained in each basis vector is distinct.

• Later: We will see we get orthogonal bases for free when working with a certain kind of Totally Awesome Matrices)

- **Bonus:** A set of orthogonal vectors is automatically linearly independent and therefore must form a basis for the subspace they span

"Obvious" but proof is nutritious

↖ dangerous word

Given $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ with $\vec{a}_i^T \vec{a}_j = 0$ for all $1 \le i, j \le n$, $i \ne j$

& $\vec{a}_i \in \mathbb{R}^m$

$\mathbb{A} = [\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n]$ ($m \times n$)

Monks: Presume linear dependence

$\Leftrightarrow$ $\mathbb{A}\vec{x} = \vec{0}$ has a non-zero solution $\vec{x}$

$\Leftrightarrow$ $N(\mathbb{A}) \ne \{\vec{0}\}$

Must have

$$0 = \vec{0}^T\vec{0} = (\mathbb{A}\vec{x})^T(\mathbb{A}\vec{x})$$

($1\times m$, $m \times 1$)

use $(\mathbb{B}\mathbb{C})^T = \mathbb{C}^T\mathbb{B}^T$

$$= \vec{x}^T \mathbb{A}^T \mathbb{A} \vec{x}$$

$$= [x_1 \cdots x_n] \begin{bmatrix} - \vec{a}_1^T - \\ - \vec{a}_2^T - \\ \vdots \\ - \vec{a}_n^T - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

($\vec{x}^T$ … $\vec{x}$)

$$= [x_1 \cdots x_n] \begin{bmatrix} \|\vec{a}_1\|^2 & & 0 \\ & \|\vec{a}_2\|^2 & \\ 0 & & \ddots \\ & & & \|\vec{a}_n\|^2 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

← all zeros; diagonal matrix

$$= x_1^2 \|\vec{a}_1\|^2 + x_2^2 \|\vec{a}_2\|^2 + \cdots + x_n^2 \|\vec{a}_n\|^2$$

($\|\vec{a}_i\| \ne 0$)

$$= 0 \text{ only if } x_1 = x_2 = \cdots = x_n = 0$$

Contradiction

$\Rightarrow N(\mathbb{A}) = \{\vec{0}\}$

E16ap2

Extra happy kind of basis:

An Orthonormal Basis

$\equiv$ Orthogonal Basis made up of unit vectors

Observation: Easy to do!

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\} \Rightarrow \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

orthogonal basis — divide by lengths (easy) — orthonormal basis

Next: How to create an orthogonal basis in the first place ← hard part

Transmuting a basis into an orthogonal one

Menu
- The Gram-Schmidt Process

E16bp1

Idea: Turn $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ (basis, $\subset \mathbb{R}^m$) into $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$ (orthogonal basis)

and then $\{\hat{q}_1, \hat{q}_2, \dots, \hat{q}_n\}$ (orthonormal basis) $\quad \hat{q}_i := \frac{1}{\|\vec{q}_i\|}\vec{q}_i$

by incrementally removing parts of vectors.

$n=3$ general formula: (3d subspace in $\mathbb{R}^m$)

① Set $\vec{q}_1 = \vec{a}_1$

② $\vec{q}_2 = \vec{a}_2 - \frac{\vec{q}_1^{\,T}\vec{a}_2}{\vec{q}_1^{\,T}\vec{q}_1}\vec{q}_1$ ← projection of $\vec{a}_2$ onto direction described by $\vec{q}_1$

③ $\vec{q}_3 = \vec{a}_3 - \frac{(\vec{q}_1^{\,T}\vec{a}_3)}{(\vec{q}_1^{\,T}\vec{q}_1)}\vec{q}_1 - \frac{(\vec{q}_2^{\,T}\vec{a}_3)}{(\vec{q}_2^{\,T}\vec{q}_2)}\vec{q}_2$

projections

ⓝ $\vec{q}_n = \vec{a}_n - (\;\dots\dots\;)$

$n-1$ projections of $\vec{a}_n$ onto $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_{n-1}$

We know $\frac{(\vec{q}_1^{\,T}\vec{a}_2)}{(\vec{q}_1^{\,T}\vec{q}_1)}\vec{q}_1 = \hat{q}_1\hat{q}_1^{\,T}\vec{a}_2$

($\vec{q}_1^{\,T}\vec{a}_2$ ← number; $\vec{q}_1^{\,T}\vec{q}_1$ ← number; $\hat{q}_1\hat{q}_1^{\,T}$: outer product, $P_{\hat{q}_1}$; $\hat{q}_1$: $m\times1$, $\hat{q}_1^{\,T}$: $1\times m$, $\vec{a}_2$: $m\times1$; $\hat{q}_1\hat{q}_1^{\,T}$: $m\times m$)

So above 3 steps can be rewritten as:

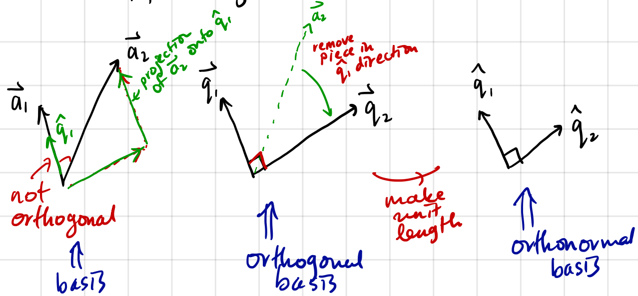
① $\vec{q}_1 = \vec{a}_1 \quad \Rightarrow \quad \hat{q}_1 = \frac{1}{\|\vec{q}_1\|}\vec{q}_1$

② $\vec{q}_2 = \vec{a}_2 - \hat{q}_1\hat{q}_1^{\,T}\vec{a}_2 \quad \Rightarrow \quad \hat{q}_2 = \frac{1}{\|\vec{q}_2\|}\vec{q}_2$

③ $\vec{q}_3 = \vec{a}_3 - \hat{q}_1\hat{q}_1^{\,T}\vec{a}_3 - \hat{q}_2\hat{q}_2^{\,T}a_3 \quad \Rightarrow \quad \hat{q}_3 = \frac{1}{\|\hat{q}_3\|}\hat{q}_3$

↗ good for theory

What's happening:



Example calculation:

$$\left\{ \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}4\\2\\0\end{bmatrix}, \begin{bmatrix}5\\-4\\-7\end{bmatrix} \right\}$$
$\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$

① $\vec{q}_1 = \vec{a}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}$ ← orthogonal →

② $$\vec{q}_2 = \vec{a}_2 - \frac{\vec{q}_1^T\vec{a}_2}{\vec{q}_1^T\vec{q}_1}\vec{q}_1$$

$$= \begin{bmatrix}4\\2\\0\end{bmatrix} - \frac{[1\ 1\ 1]\begin{bmatrix}4\\2\\0\end{bmatrix}}{[1\ 1\ 1]\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}4\\2\\0\end{bmatrix} - \frac{6}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}2\\0\\-2\end{bmatrix}$$

(numerator: 6, denominator: 3)

③ $$\vec{q}_3 = \vec{a}_3 - \frac{\vec{q}_1^T\vec{a}_3}{\vec{q}_1^T\vec{q}_1}\vec{q}_1 - \frac{\vec{q}_2^T\vec{a}_3}{\vec{q}_2^T\vec{q}_2}\vec{q}_2$$

$$= \begin{bmatrix}5\\-4\\-7\end{bmatrix} - \frac{[1\ 1\ 1]\begin{bmatrix}5\\-4\\-7\end{bmatrix}}{[1\ 1\ 1]\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{[2\ 0\ -2]\begin{bmatrix}5\\-4\\-7\end{bmatrix}}{[2\ 0\ -2]\begin{bmatrix}2\\0\\-2\end{bmatrix}}\begin{bmatrix}2\\0\\-2\end{bmatrix}$$

(−6, +3, 24, 8)

$$= \begin{bmatrix}5\\-4\\-7\end{bmatrix} - \frac{-6}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{24}{8}\begin{bmatrix}2\\0\\-2\end{bmatrix} = \begin{bmatrix}1\\-2\\1\end{bmatrix}$$

(+2, −3)

Normalize:

$$\hat{q}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \hat{q}_2 = \frac{1}{\sqrt{8}}\begin{bmatrix}2\\0\\-2\end{bmatrix}, \quad \hat{q}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix}$$

$\sqrt{8} = 2\sqrt{2}$; $\hat{q}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix}$

Note: Gram-Schmidt method tends to produce many square roots

Check $\hat{q}_1^T\hat{q}_2 = \hat{q}_2^T\hat{q}_3 = \hat{q}_3^T\hat{q}_1 = 0$

- test every pair of basis vectors
- must be orthogonal

Next: See $\vec{a}_i$'s can be rebuilt from $\hat{q}_i$'s $\Rightarrow$ $A = QR$

A: $m\times n$, Q: $m\times n$, R: $n\times n$

A new factorization: $\mathbb{A} = \mathbb{Q}\mathbb{R}$

$m\times n \quad m\times n \quad n\times n$

← Q's shape matches $\mathbb{A}$'s

unfortunate namespace overlap → NOT $\mathbb{R}/\mathbb{A}$

E16cp1

Idea: We love to use matrices to encode our methods ← empowering

ex $\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{U}$ ≡ Gaussian Elimination

So: let's turn the Gram-Schmidt Process into a factorization of $\mathbb{A}$

From a few pages back:

(1) $\vec{q}_1 = \vec{a}_1 \Rightarrow \hat{q}_1 = \frac{1}{\|\vec{q}_1\|}\vec{q}_1$

(2) $\vec{q}_2 = \vec{a}_2 - \underbrace{\hat{q}_1\hat{q}_1^T}_{m\times m\ \mathbb{P}_{\hat{q}_1}}\vec{a}_2 \Rightarrow \hat{q}_2 = \frac{1}{\|\vec{q}_2\|}\vec{q}_2$

(3) $\vec{q}_3 = \vec{a}_3 - \underbrace{\hat{q}_1\hat{q}_1^T}_{m\times m}\vec{a}_3 - \underbrace{\hat{q}_2\hat{q}_2^T}_{m\times m} a_3 \Rightarrow \hat{q}_3 = \frac{1}{\|\vec{q}_3\|}\hat{q}_3$

Monks say: Express the $\vec{a}_i$ in terms of the $\hat{q}_i$ using a column picture approach

$\Rightarrow$ Connect $\mathbb{A}$ to $\mathbb{Q} = \begin{bmatrix} | & | & & | \\ \hat{q}_1 & \hat{q}_2 & \cdots & \hat{q}_n \\ | & | & & | \end{bmatrix}$

Rearrange above so $\vec{a}_i = \cdots$:

↑ put $\vec{a}_i$'s on left by themselves

(1) $\vec{a}_1 = \vec{q}_1$

(2) $\vec{a}_2 = \vec{q}_2 + \hat{q}_1\hat{q}_1^T\vec{a}_2$

(3) $\vec{a}_3 = \vec{q}_3 + \hat{q}_1\hat{q}_1^T\vec{a}_3 + \hat{q}_2\hat{q}_2^T\vec{a}_3$

need these to look the same

Sneakiness: See $\vec{q}_i$ as projection of $\vec{a}_i$ onto $\hat{q}_i$ direction

ex (3) above:

$(\hat{q}_3\hat{q}_3^T)\vec{a}_3 = (\hat{q}_3\hat{q}_3^T)(\vec{q}_3 + \hat{q}_1\hat{q}_1^T\vec{a}_3 + \hat{q}_2\hat{q}_2^T\vec{a}_3)$

($\hat{q}_3\hat{q}_3^T = \mathbb{P}_{\hat{q}_3}$; $\perp$, $\perp$)

$= \hat{q}_3(\hat{q}_3^T\vec{q}_3) + 0 + 0$

$= \vec{q}_3$ ($\hat{q}_3^T\vec{q}_3 \to \|\vec{q}_3\|$)

(1) $\vec{a}_1 = \hat{q}_1\hat{q}_1^T\vec{a}_1$

(2) $\vec{a}_2 = \hat{q}_2\hat{q}_2^T\vec{a}_2 + \hat{q}_1\hat{q}_1^T\vec{a}_2$

(3) $\vec{a}_3 = \hat{q}_3\hat{q}_3^T\vec{a}_3 + \hat{q}_1\hat{q}_1^T\vec{a}_3 + \hat{q}_2\hat{q}_2^T\vec{a}_3$

Makes sense: $\vec{a}_3$ has components in $\hat{q}_1$, $\hat{q}_2$, & $\hat{q}_3$ directions

E16cp2

Reorder:

(1) $\vec{a}_1 = \hat{q}_1 \hat{q}_1^T \vec{a}_1$ ($m\times1$, $1\times m$, $m\times1$) — numbers (inner products)

(2) $\vec{a}_2 = \hat{q}_1 \hat{q}_1^T \vec{a}_2 + \hat{q}_2 \hat{q}_2^T \vec{a}_2$

(3) $\vec{a}_3 = \hat{q}_1 \hat{q}_1^T \vec{a}_3 + \hat{q}_2 \hat{q}_2^T \vec{a}_3 + \hat{q}_3 \hat{q}_3^T \vec{a}_3$

Column picture:

(1) $$\vec{a}_1 = \begin{bmatrix} | & | & | \\ \hat{q}_1 & \hat{q}_2 & \hat{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \hat{q}_1^T\vec{a}_1 \\ 0 \\ 0 \end{bmatrix}$$

(2) $$\vec{a}_2 = \begin{bmatrix} | & | & | \\ \hat{q}_1 & \hat{q}_2 & \hat{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \hat{q}_1^T\vec{a}_2 \\ \hat{q}_2^T\vec{a}_2 \\ 0 \end{bmatrix}$$

(3) $$\vec{a}_3 = \begin{bmatrix} | & | & | \\ \hat{q}_1 & \hat{q}_2 & \hat{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \hat{q}_1\vec{a}_3 \\ \hat{q}_2\vec{a}_3 \\ \hat{q}_3\vec{a}_3 \end{bmatrix}$$

TRIUMPHANCY:

$$\mathbb{A} = \begin{bmatrix} | & | & | \\ \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \hat{q}_1 & \hat{q}_2 & \hat{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \hat{q}_1^T\vec{a}_1 & \hat{q}_1^T\vec{a}_2 & \hat{q}_1^T\vec{a}_3 \\ 0 & \hat{q}_2^T\vec{a}_2 & \hat{q}_2^T\vec{a}_3 \\ 0 & 0 & \hat{q}_3^T\vec{a}_3 \end{bmatrix}$$

$m\times n$ = $\mathbb{Q}$ ($m\times n$; $\mathbb{A}$ cleaned up) · $\mathbb{R}$ ($n\times n$; "upper triangular combining matrix")

- $\mathbb{A} = \mathbb{Q}\mathbb{R}$ will help with $\mathbb{A}\vec{x} = \vec{b}$ (next)
- Delicious way to find $\mathbb{R}$:

$$\mathbb{Q}^T\mathbb{A} = \mathbb{Q}^T\mathbb{Q}\mathbb{R} \Rightarrow \mathbb{R} = \mathbb{Q}^T\mathbb{A}$$

↑ premultiply by $\mathbb{Q}^T$; $\mathbb{Q}^T\mathbb{Q} = \mathbb{I}$ because $\mathbb{Q}$'s columns are unit vectors

Return to example:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} \right\} \Rightarrow \left\{ \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

($\vec{a}_1, \vec{a}_2, \vec{a}_3$ → $\hat{q}_1, \hat{q}_2, \hat{q}_3$)

$$\underbrace{\begin{bmatrix} 1 & 4 & 5 \\ 1 & 2 & -4 \\ 1 & 0 & -7 \end{bmatrix}}_{\mathbb{A}} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}}_{\mathbb{Q}} \underbrace{\begin{bmatrix} \sqrt{3} & 2\sqrt{3} & -2\sqrt{3} \\ 0 & 2\sqrt{2} & 6\sqrt{2} \\ 0 & 0 & \sqrt{6} \end{bmatrix}}_{\mathbb{R}}$$

check $\mathbb{A} = \mathbb{Q}\mathbb{R}$ works; check $\mathbb{R} = \mathbb{Q}^T\mathbb{A}$

Find $\mathbb{R}$ by either computing inner products $\hat{q}_i \vec{a}_j, i \le j$

or $\mathbb{R} = \mathbb{Q}^T\mathbb{A}$ ↳ do this!

E16dp1

$\mathbb{A}\vec{x} = \vec{b}$ & $\mathbb{A} = \mathbb{Q}\mathbb{R}$: (presume $r=n$)

$\mathbb{A}$: $m\times n$, $\vec{x}$: $n\times 1$, $\vec{b}$: $m\times 1$; $\mathbb{Q}$: $m\times n$, $\mathbb{R}$: $n\times n$

Solve the normal equation using $\mathbb{A}=\mathbb{Q}\mathbb{R}$: ($\vec{b}\notin C(\mathbb{A})$)

$$\mathbb{A}^T\mathbb{A}\vec{x}_* = \mathbb{A}^T\vec{b}$$

($\mathbb{R}$ is upper triangular)

$$\Rightarrow (\mathbb{Q}\mathbb{R})^T(\mathbb{Q}\mathbb{R})\vec{x}_* = (\mathbb{Q}\mathbb{R})^T\vec{b}$$

$$\Rightarrow \underset{n\times n}{\mathbb{R}^T}\underbrace{\underset{n\times m}{\mathbb{Q}^T}\underset{m\times n}{\mathbb{Q}}}_{\mathbb{I}_{n\times n}}\underset{n\times n}{\mathbb{R}}\underset{n\times 1}{\vec{x}_*} = \underbrace{\underset{n\times n}{\mathbb{R}^T}\underset{n\times m}{\mathbb{Q}^T}\underset{m\times 1}{\vec{b}}}_{n\times 1}$$

$$\Rightarrow \mathbb{R}^T\mathbb{R}\vec{x}_* = \mathbb{R}^T\mathbb{Q}^T\vec{b}$$

$(\mathbb{R}^T)^{-1}\times$ both sides

$$\Rightarrow \mathbb{R}\vec{x}_* = \mathbb{Q}^T\vec{b} \quad \text{because } (\mathbb{R}^T)^{-1} \text{ exists}$$

(square full rank, all n pivots exist)

$$\underset{n\times n}{\mathbb{R}}\underset{n\times 1}{\vec{x}_*} = \underset{n\times m}{\mathbb{Q}^T}\underset{m\times 1}{\vec{b}}$$

$\mathbb{A}'\vec{x}_* = \vec{b}'_{n\times 1}$

upper triangular system! Easy to solve!

c.f. $\mathbb{A} = \mathbb{L}\mathbb{U}$

But

$$\mathbb{A}\vec{x} = \vec{b}$$

$$\Rightarrow (\mathbb{Q}\mathbb{R})\vec{x} = \vec{b}$$

$$\Rightarrow \mathbb{Q}^T\mathbb{Q}\mathbb{R}\vec{x} = \mathbb{Q}^T\vec{b}$$

(premultiply both sides; $\mathbb{Q}^T$ left inverse of $\mathbb{Q}$)

$$\Rightarrow \underset{n\times n}{\mathbb{I}}\mathbb{R}\vec{x} = \mathbb{Q}^T\vec{b}$$

$$\Rightarrow \mathbb{R}\vec{x} = \mathbb{Q}^T\vec{b}$$

not $\vec{x}_*$ eek? ??

really $\mathbb{R}\vec{x}_* = \mathbb{Q}^T\vec{b}$.

We are really solving the normal equation... because $\mathbb{Q}^T\vec{b} = \mathbb{Q}^T\vec{p}$ ($\vec{p}$: projection of $\vec{b}$ onto $C(\mathbb{A})$)

$$\mathbb{Q}^T\vec{b} = \begin{bmatrix} — \hat{q}_1^T — \\ — \hat{q}_2^T — \\ \vdots \\ — \hat{q}_n^T — \end{bmatrix}\begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix}$$

$\hat{q}_i$'s span $C(\mathbb{A})$

$$\underset{m\times m}{\mathbb{P}_{\hat{q}_i}} = \underset{m\times 1}{\hat{q}_i}\underset{1\times m}{\hat{q}_i^T}$$ ← projection operator for direction $\hat{q}_i$

What's going on: $\mathbb{Q}^T\vec{b} = \mathbb{Q}^T(\vec{p} + \vec{e})$ ($\vec{p}\in C(\mathbb{A})$, $\vec{e}\in N(\mathbb{A}^T)$)

$$= \mathbb{Q}^T\vec{p} + \mathbb{Q}^T\vec{e}$$ ($\mathbb{Q}^T\vec{e} = \vec{0}$; $\mathbb{Q}$: $C(\mathbb{A})$)

E16e p1

Orthogonal Matrices:
(really "orthonormal")

The Gram-Schmidt Process gave us $\mathbb{A} = \mathbb{Q}\,\mathbb{R}$

$m\times n$ $m \times n$ $n \times n$

$\mathbb{R}$ ← upper triangular combining matrix

$\mathbb{Q}$'s columns $\{\hat{q}_i\}$ form an orthonormal basis for $\mathbb{A}$'s column space;

$$\hat{q}_i^T \hat{q}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

[Note: $\mathbb{A}$'s columns are ideally linearly independent ($n=r$)]

More on what $\mathbb{Q}$ type matrices can do for you:

$$\mathbb{Q}^T \mathbb{Q} = \mathbb{I}$$ ← left inverse for $\mathbb{Q}$

$$\underbrace{\begin{bmatrix} — \hat{q}_1^T — \\ — \hat{q}_2^T — \\ \vdots \\ — \hat{q}_n^T — \end{bmatrix}}_{n\times m} \underbrace{\begin{bmatrix} | & | & & | \\ \hat{q}_1 & \hat{q}_2 & \cdots & \hat{q}_n \\ | & | & & | \end{bmatrix}}_{m \times n} = \underbrace{\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}_{n\times n} = \mathbb{I}$$

If $\mathbb{Q}$ is square, then $m=n=r$

So inverse exists ($N(\mathbb{Q}) = \{\vec{0}\}$)

then

$$\left.\begin{aligned} \mathbb{Q}^T\mathbb{Q} &= \mathbb{I} = \mathbb{Q}\mathbb{Q}^T \\ \mathbb{Q}^{-1}\mathbb{Q} &= \mathbb{I} = \mathbb{Q}\mathbb{Q}^{-1} \end{aligned}\right\} \quad \boxed{\mathbb{Q}^{-1} = \mathbb{Q}^T}$$

Say $\mathbb{Q}$ is an orthogonal matrix

Many other groovy properties:

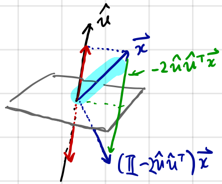
$$\|\mathbb{Q}\vec{x}\| = \|\vec{x}\|$$ ← length is preserved under transformation by $\mathbb{Q}$ (length)

$$(\mathbb{Q}\vec{x})^T(\mathbb{Q}\vec{y}) = \vec{x}^T\vec{y}$$ ← $\mathbb{Q}$ preserves angles

ex 1 $\quad \mathbb{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ ← rotation by θ ↺

ex 2 $\quad \mathbb{Q} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ ← permutat.

ex 3 $\quad \mathbb{Q} = \mathbb{I} - 2\hat{u}\hat{u}^T$ ($\hat{u}\hat{u}^T$: $\mathbb{P}_{\hat{u}}$; project onto $\hat{u}$ and flip)



Three reasons to love arbitrary powers of square matrices

In our journey so far, we've spent a lot of time thinking about one of the Monks' favorite equations: $\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$

Now: The Monks tell us to think about square matrices ($n\times n$) as gadgets, things that transform vectors into new vectors ✓ ✓

$$\underset{n\times 1}{\vec{x}\,'} = \underset{n\times n}{A}\,\underset{n\times 1}{\vec{x}}$$

A might { flip, rotate, stretch, project } $\vec{x}$

Big Question: what happens if we use A to repeatedly transform a vector?

Start with $\vec{x}_0$

$\vec{x}_1 = A\vec{x}_0$, $\vec{x}_2 = A\vec{x}_1$, ..., $\vec{x}_k = A\vec{x}_{k-1}$, ...

$$\Rightarrow \quad \boxed{\vec{x}_k = A^k \vec{x}_0}$$

$$\underset{n\times n}{A^k} = \underbrace{\underset{n\times n}{A}\cdot\underset{n\times n}{A}\cdot \cdots \cdot \underset{n\times n}{A}}_{k}$$

← kth power of A

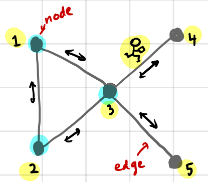
Difficulty: Mindless multiplication of many matrices works but is
(1) computationally expensive;
(2) doesn't give us any understanding of how A^k behaves ← deep story

The Monks whisper that we must understand eigenthings... (vast and wonderful)

But first: three example areas showing the excellence of $\underset{n\times n}{A^k}$...

E17ap2

(1) The distracted texter wandering randomly on a network:



columns must sum to 1

$$\begin{bmatrix} P_{t+1,1} \\ P_{t+1,2} \\ P_{t+1,3} \\ P_{t+1,4} \\ P_{t+1,5} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 & 1 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 \end{bmatrix} \begin{bmatrix} P_{t,1} \\ P_{t,2} \\ P_{t,3} \\ P_{t,4} \\ P_{t,5} \end{bmatrix}$$

$$\vec{P}_{t+1} \qquad \mathbb{A} \qquad \vec{P}_t$$

$\vec{P}_{t+1}$: probability texter is at nodes 1,..,5 at time t+1

$\mathbb{A}$: transition matrix

Natural question:
- where is our texter likely to be as time goes on?
- or, what is $\vec{P}_\infty$?
- or, what is $\mathbb{A}^k$ as $k \to \infty$?

Monks (and soon you) tell us that

$$\mathbb{A} \begin{bmatrix} 2 \\ 2 \\ 4 \\ 1 \\ 1 \end{bmatrix} = 1 \cdot \begin{bmatrix} 2 \\ 2 \\ 4 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \vec{P}_\infty = \frac{1}{10} \begin{bmatrix} 2 \\ 2 \\ 4 \\ 1 \\ 1 \end{bmatrix}$$

1: eigenvalue (later)

$[2,2,4,1,1]^T$: eigenvector — no change

$\frac{1}{10}$: normalized

great result: $P_{\infty,i}$ is proportional to the degree of node i

(2) Solving coupled, linear differential equations

simple

$$\frac{dx}{dt} = 3x \Rightarrow x(t) = x(0)e^{3t}$$

$x(0)$: initial value at t=0

← check this works

Coupled

$$\frac{dx_1}{dt} = 2x_1 - x_2$$

$$\frac{dx_2}{dt} = 3x_1 + 2x_2$$

change depends on current position

← continuous & discrete math crushing it together

Rewrite with matrices:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \frac{d\vec{x}}{dt} = \mathbb{A}\vec{x}$$

($\mathbb{A}$: 2×2)

Solution is of the form: $\vec{x}(t) = e^{\mathbb{A}t}\vec{x}(0)$

($\vec{x}(t)$: 2×1; $e^{\mathbb{A}t}$: 2×2 — what??; $\vec{x}(0)$: 2×1)

It's true!: you can exponentiate matrices!!

$$e^{\mathbb{A}} = \mathbb{I} + \mathbb{A} + \frac{1}{2!}\mathbb{A}^2 + \frac{1}{3!}\mathbb{A}^3 + \ldots + \frac{1}{k!}\mathbb{A}^k + \ldots$$

← Taylor expansion

we need to really understand all powers of $\mathbb{A}$ to cope with this...

$$e^{\mathbb{A}t} = \mathbb{I} + \mathbb{A}t + \frac{1}{2!}\mathbb{A}^2t^2 + \frac{1}{3!}\mathbb{A}^3t^3 + \ldots$$

(t: scalar)

$$e^{\mathbb{G}} = \mathbb{I}_{\mathbb{G}} + \mathbb{G} + \frac{1}{2!}\mathbb{G}^2 + \frac{1}{3!}\mathbb{G}^3 + \ldots$$

#awesome

(3) Solving super fun difference equations

ex $F_{k+2} = F_{k+1} + F_k$ with $F_0 = F_1 = 1$ ← initial conditions

↖ Fibonacci sequence

Monks say try:

$$\vec{f}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

↖ stack consecutive Fibonacci numbers

$$\vec{f}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \vec{f}_{k-2}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \vec{f}_0$$

(the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$: unbelievable matrix... ; $\begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$: ↑ $\vec{f}_{k-1}$)

So if we can compute $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k$ for all k in a clever way, we'll have a formula for the Fibonacci numbers.

↘ later.

Again, understanding and calculating A^k is made possible through the magic of eigenthings

- values
- vectors
- spaces
- functions

The Magic of Eigenthings

an introduction to happiness

Scene:

A Monk hands us a parchment with

$A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}, \quad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ ← note $\vec{v}_1 \perp \vec{v}_2$ ....

and other strange symbols written on it, smiles, and then mysteriously disappears...

Let's try some things...

$$A\vec{v}_1 = \begin{bmatrix} 1 & 1/2 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 3/2 \end{bmatrix} = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{3}{2}\vec{v}_1$$

(2x2; A ↑ symmetric; ↑ $\vec{v}_1$; ← only direction matters; ↑ $\vec{v}_1$ comes back with a 3/2 stretch factor)

$$A^2\vec{v}_1 = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\frac{3}{2}\right)^2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\frac{3}{2}\right)^2\vec{v}_1$$

$$A^k\vec{v}_1 = \begin{bmatrix} 1 & 1/2 \\ \frac{1}{2} & 1 \end{bmatrix}^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\frac{3}{2}\right)^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\frac{3}{2}\right)^k\vec{v}_1$$

↑ grows

A "likes" the direction of $\vec{v}_1$

↳ "eigenvector" (eigen: German for "own")

And we'll call $\lambda_1 = 3/2$ the "eigenvalue" associated with $\vec{v}_1$

E17bp1

$$A\vec{v}_2 = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2}\vec{v}_2$$

$\vec{v}_2$ is also an eigenvector of A

$\lambda_2 = \frac{1}{2}$ is the associated eigenvalue

- Note again: only direction of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ matters

$$A\begin{bmatrix} 4 \\ 4 \end{bmatrix} = \frac{3}{2}\begin{bmatrix} 4 \\ 4 \end{bmatrix} \quad \& \quad A\begin{bmatrix} 17 \\ -17 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 17 \\ -17 \end{bmatrix}$$

(↑ $4\vec{v}_1$; ↑ $17\vec{v}_2$)

- Also: $A^k\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \left(\frac{1}{2}\right)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ ↖ vanishes

One more thing:

$$A\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$ ← different direct

↑ not an eigenvector

$$A^2\begin{bmatrix} 2 \\ 0 \end{bmatrix} = A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 2 \end{bmatrix}$$ ← different again...

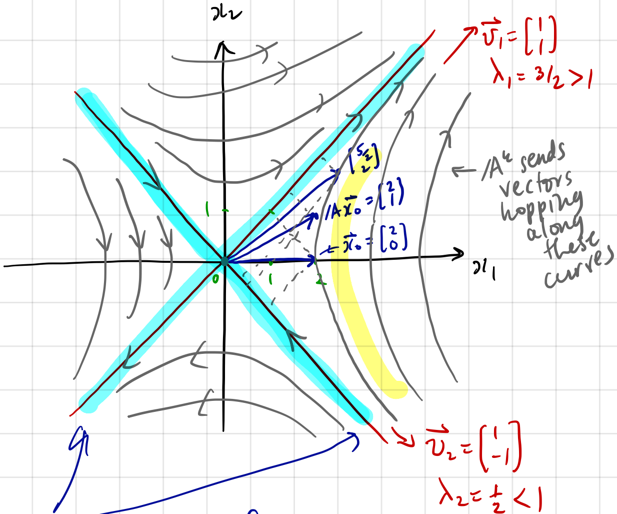
Better:

$$A\begin{bmatrix} 2 \\ 0 \end{bmatrix} = A\left(\underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\vec{v}_1} + \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\vec{v}_2}\right) = \frac{3}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

← A is simple in eigenvector basis

huge: $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$ is a great basis for A

picture for A^k:



Eigenspaces of A
(1-d subspaces of $\mathbb{R}^2$)

possibilities for $A\vec{v} = \lambda\vec{v}$

$\lambda > 1$: growth

$\lambda = 1$: stays the same

$0 < \lambda < 1$: shrinkage

$\lambda < 0$: jumping back and forth across origin, $|\lambda|$ governs growth

λ complex: rotation

E17bp2

Big question: If monks aren't around, how do we find $\vec{v}$'s and λ's? How many are possible if A is $n \times n$?

Game is to solve the Eigenvalue Equation:

$$A\vec{v} = \lambda\vec{v}$$

(A: $n\times n$, $\vec{v}$: $n\times 1$, λ: scalar, $\lambda\vec{v}$: $n\times 1$)

$A\vec{x} = \vec{b}$

$$A\vec{v} - \lambda\vec{v} = \vec{0}$$

($\vec{0}$: $n\times 1$ ← vector!!)

temptation ↓ shame → $(A - \lambda)\vec{v} = \vec{0}$ (A: $n\times n$, λ: 1×1) no!!

→ sneakiness (Monks)

$$A\vec{v} - \lambda I\vec{v} = \vec{0}$$

(A: $n\times n$, $\vec{v}$: $n\times 1$, λ: scalar, I: $n\times n$, $\vec{v}$: $n\times 1$, $\vec{0}$: $n\times 1$)

$$(A - \lambda I)\vec{v} = \vec{0}$$

($A - \lambda I$: $n\times n$, $\vec{v}$: $n\times 1$, $\vec{0}$: $n\times 1$)

Nullspace Equation!!

λ (← lambda) must be such that,

- $\text{rank}(A - \lambda I) < n$
- $N(A - \lambda I) \neq \{\vec{0}\}$
- $A - \lambda I$ has no inverse (A is singular)
- $\det(A - \lambda I) = |A - \lambda I| \equiv 0$ ← new thing

Solve $(\mathbb{A} - \lambda \mathbb{I})\vec{v} = \vec{0}$ for $\mathbb{A} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$

($\lambda\mathbb{I} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$)

Usual way:

$$\begin{bmatrix} 1-\lambda & \frac{1}{2} \\ \frac{1}{2} & 1-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

↑ subtracting λ from diagonal entries of $\mathbb{A}$

Augmented matrix

$$\left[\begin{array}{cc|c} 1-\lambda & \frac{1}{2} & 0 \\ \frac{1}{2} & 1-\lambda & 0 \end{array}\right] \xrightarrow{R_2' = R_2 - \left(\frac{1/2}{1-\lambda}\right)R_1} \left[\begin{array}{cc|c} 1-\lambda & \frac{1}{2} & 0 \\ 0 & 1-\lambda - \frac{(\frac{1}{2})^2}{1-\lambda} & 0 \end{array}\right]$$

↑ eek!!

See $(1-\lambda) - \frac{(\frac{1}{2})^2}{1-\lambda} = 0$

for $r = 1$ (rank), $N(\mathbb{A} - \lambda\mathbb{I}) \neq \{\vec{0}\}$

$\Rightarrow \vec{v}$ is healthy

$$(1-\lambda)^2 = \left(\tfrac{1}{2}\right)^2$$

$\Rightarrow 1 - \lambda = \pm \frac{1}{2}$

$\Rightarrow \lambda = 1 \pm \frac{1}{2}$

$\lambda_1 = 3/2$, $\lambda_2 = \frac{1}{2}$

just as we found...

$\Rightarrow$ next step: find $\vec{v}$ as nullspace vectors (really: basis vectors) for $\mathbb{A} - \frac{3}{2}\mathbb{I}$ & $\mathbb{A} - \frac{1}{2}\mathbb{I}$

Unfortunately, preceding is a very messy way to handle $(\mathbb{A} - \lambda\mathbb{I})\vec{v} = \vec{0}$... There's a better, more illuminating way.

= **Set $\det(\mathbb{A} - \lambda\mathbb{I}) = 0$** ← reason is coming...

and find λ...

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \text{ for } 2\times2\text{s}$$

$$\begin{vmatrix} 1-\lambda & \frac{1}{2} \\ \frac{1}{2} & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - \left(\tfrac{1}{2}\right)^2 = 0$$

↑ same equation as before

- see next episodes for all things determinants
- we return to eigenthings after this strange excursion

$\lambda_1 = 3/2$, $\lambda_2 = 1/2$

Find $\vec{v}_1$ & $\vec{v}_2$

$\lambda_1 = 3/2$:

$$(\mathbb{A} - \tfrac{3}{2}\mathbb{I})\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$ ← standard Nullspace Equation

$$\Downarrow$$

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

$$= \left[\begin{array}{cc|c} -\tfrac{1}{2} & \tfrac{1}{2} & 0 \\ \tfrac{1}{2} & -\tfrac{1}{2} & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} -\tfrac{1}{2} & \tfrac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

(P, F over the columns; $R_2' = R_2 - \left(\frac{+\frac{1}{2}}{-\frac{1}{2}}\right) R_1$)

$$-\tfrac{1}{2}v_1 + \tfrac{1}{2}v_2 = 0$$

$$\Rightarrow v_1 = v_2$$

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_2 \end{bmatrix} = v_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$ ← Eigenspace for $\lambda_1 = 3/2$

Say $\lambda_1 = 3/2$ with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

↑ basis for eigenspace

$\lambda_2 = \frac{1}{2}$:

$$(\mathbb{A} - \tfrac{1}{2}\mathbb{I})\vec{v}_2 = \vec{0}$$

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

$$= \left[\begin{array}{cc|c} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

(P, F over the columns)

$$\Rightarrow \vec{v}_2 = c\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad c \in \mathbb{R}$$ ← Eigenspace

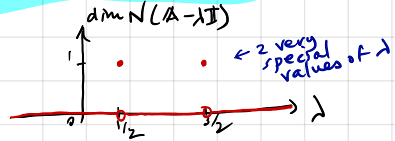
$\lambda_1 = \frac{1}{2}$ has eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Often, unit vectors are best

$$\hat{v}_1 = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \& \quad \hat{v}_2 = \tfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$\lambda_1 = 3/2$ $\qquad$ $\lambda_2 = 1/2$

↗ $\mathbb{A}$'s natural basis

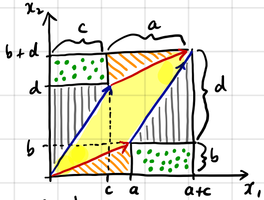


Determinants from the ground up:

2×2's first then n×n's

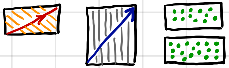
$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

idea: Consider area of parallelogram formed by row vectors of A: $\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix}$



Area of  $= (a+c)(b+d)$

$- ab \quad - dc \quad - 2bc$



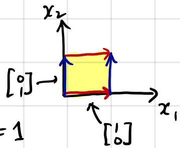
$= ab + ad + cb + cd$
$\quad - ab - dc - 2bc$

$= ad - bc$ → call this A's determinant: $\det(A)$ or $|A|$

Three Observations about this determinant thing for 2×2's:

① $|I| = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$

anchor



area = 1

② If we swap A's rows, we flip the sign of the determinant

$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ but $\begin{vmatrix} c & d \\ a & b \end{vmatrix} = bc - ad$
$= -(ad - bc)$

straight lines for determinant

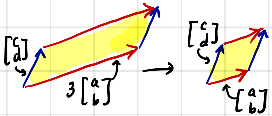
−ve area indicates ordering of vectors

③ $|A| = ad - bc$ is **multilinear** in the rows of A

Two pieces:

3.1 Area Scales:

ex $\begin{vmatrix} 3a & 3b \\ c & d \end{vmatrix} = 3 \begin{vmatrix} a & b \\ c & d \end{vmatrix}$



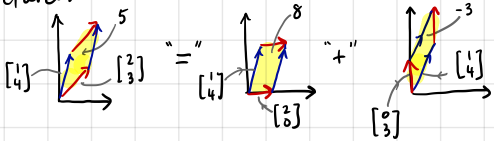
ex $\begin{vmatrix} 2a & 2b \\ 4c & 4d \end{vmatrix} = 2 \cdot 4 \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

3.2 Areas add when single rows add:

ex $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 1 & 4 \end{vmatrix}$

formula: $\underset{5}{2 \cdot 4 - 3 \cdot 1} = \underset{8}{(2 \cdot 4 - 1 \cdot 0)} + \underset{-3}{(0 \cdot 4 - 1 \cdot 3)}$

pictures:



Determinants from the ground up:

The plan: We assert Three Properties for Determinants of n×n matrices

① $|\mathbb{I}| = 1$. ← n×n unit hypercube

[Determinant = ± volume of parallelipiped created by row vectors of $\mathbb{A}$]

② Swapping any two rows of $\mathbb{A}$ changes the sign of the determinant.

③ Determinants are multilinear in their rows.

Big Deal: Can now connect $|\mathbb{A}|$ to $|\mathbb{I}| = 1$ and many, many good things will follow

ok: Let's fully connect $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}$ to $|\mathbb{I}|$

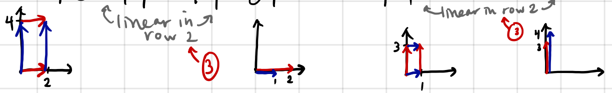
$$|\mathbb{A}| = \begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 1 & 4 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 1 & 4 \end{vmatrix}$$

← linear in row 1 ③

need to introduce 0's to get to $\mathbb{I}$

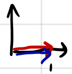
$$= \begin{vmatrix} 2 & 0 \\ 0 & 4 \end{vmatrix} + \begin{vmatrix} 2 & 0 \\ 1 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 0 & 4 \end{vmatrix}$$

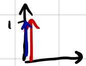
linear in row 2 ③ … linear in row 2 ③



$$= 2\cdot4\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + 2\begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} + 3\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + 3\cdot4\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix}$$

↑ 0 volume — row swap — ↑ 0 volume





$$= 8\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + 3(-1)\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

$$= (8-3)\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 5\,|\mathbb{I}| = 5$$

Next: Show how our standard row operations lead to many results ← $\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{U}$ including $|\mathbb{A}| = \pm|\mathbb{U}|$ & $|\mathbb{A}\mathbb{B}| = |\mathbb{A}||\mathbb{B}|$

#excitement

[Many results for determinants based on three properties:

① $|\mathbb{I}_{n\times n}|=1$, ② Row swap → $\times(-1)$ (sign flip), ③ Multilinearity (rows)

[notation: $\det \mathbb{A} = |\mathbb{A}|$]

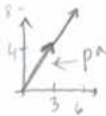
(a) $|t\mathbb{A}^{n\times n}| = t^n |\mathbb{A}|$ ← multilinearity, $t \in \mathbb{R}$

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

(b) If two of $\mathbb{A}$'s rows are the same, then $|\mathbb{A}|=0$.

ex $\begin{vmatrix}3&4\\3&4\end{vmatrix} \underset{R_2\leftrightarrow R_1}{=} -\begin{vmatrix}3&4\\3&4\end{vmatrix}$ (property ②) so $\begin{vmatrix}3&4\\3&4\end{vmatrix}=0$

ex $\begin{vmatrix}3&6&6\\3&6&6\\2&-2&1\end{vmatrix} \underset{R_2\leftrightarrow R_1}{=} -\begin{vmatrix}3&6&6\\3&6&6\\2&-2&1\end{vmatrix}$ (property ②) so must be zilch



(c) Performing a step of standard row reduction doesn't change the value of the determinant #bigdeal

$R_i' = 1 R_i - (l_{ij}) R_j$

Row reduction for a 2×2:

$\mathbb{A} = \begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix} \underset{R_2'=R_2-l_{21}R_1}{\sim} \begin{bmatrix}a_{11}&a_{12}\\a_{21}-l_{21}a_{11}&a_{22}-l_{21}a_{12}\end{bmatrix}$

(tilde; zero this out; $l_{21} = \left(\frac{a_{21}}{a_{11}}\right)$; now find determinant of this new matrix)

$\begin{vmatrix}a_{11}&a_{12}\\a_{21}-l_{21}a_{11}&a_{22}-l_{21}a_{22}\end{vmatrix} =$ ← multilinearity

$\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix} - l_{21}\begin{vmatrix}a_{11}&a_{12}\\a_{11}&a_{12}\end{vmatrix}$ (= 0)

↑ original matrix; ↑ same rows (b)

☺ • Above generalizes to n×n

☺ "∼" for solving $\mathbb{A}\vec{x}=\vec{b}$ ⇒ "=" for determinants

(d) If $\mathbb{A}_{n\times n}$'s rows are linearly dependent, then $|\mathbb{A}|=0$ #bigdeal

Reason:

* We can use row row ops to make one or more rows of zeros
* Now add one non-zero row to a zero row using one more row op

⇒ 2 rows the same (b) ⇒ $|\mathbb{A}|=0$

(d) cont.

Or multiply zero row by any number c

$\Rightarrow$ Multilinearity means determinant should scale by a factor of c

$\Rightarrow$ But row was unchanged $c\vec{0}^T = \vec{0}^T$ so $|A| = 0$.

ex $\begin{vmatrix} 2 & -2 & 1 \\ 3 & 6 & 6 \\ 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 2 & -2 & 1 \\ 3 & 6 & 6 \\ c[0 & 0 & 0] \end{vmatrix} = c \begin{vmatrix} 2 & -2 & 1 \\ 3 & 6 & 6 \\ 0 & 0 & 0 \end{vmatrix}$

only possible if $\begin{vmatrix} 2 & -2 & 1 \\ 3 & 6 & 6 \\ 0 & 0 & 0 \end{vmatrix} = 0$

Connections:

$|A|_{n\times n} = 0$ ← Applies if A is square only

$\Leftrightarrow$ A's rows are linearly dependent

$\Leftrightarrow$ Rank of A, r, is less than n

$\Leftrightarrow$ $N(A)$ & $N(A^T)$ have dim $n-r \geq 1$

$\Leftrightarrow$ A has no inverse

$\Leftrightarrow$ If $A\vec{x} = \vec{b}$ has a solution, then there are ∞ly many solutions

$\Leftrightarrow$ One or more sides of A's parallelipiped have 0 length.

$\Leftrightarrow$ one or more pivots = 0

E186 P2

Example calculation of $|A|$ using row ops:

$|A| = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 2 & 1 & 2 \end{vmatrix}$ ↙ freestyling

$\underset{R_2 \Leftrightarrow R_3}{=} - \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 0 & 3 & 2 \end{vmatrix}$

$\underset{R_2' = R_2 - (\frac{2}{1})R_1}{=} - \begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 3 & 2 \end{vmatrix}$

$\underset{R_3' = R_3 - (\frac{3}{-1})R_2}{=} - \begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{vmatrix}$ ← U from $PA = LU$; d_1, d_2, d_3 pivots!

$\underset{\substack{R_1' = R_1 - (\frac{1}{-1})R_2 \\ R_1'' = R_1' - (\frac{1}{2})R_3}}{=} - \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{vmatrix} = \underbrace{-(1)(-1)(2)}_{2} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$

↖ R_A echelon ↖ new thing; ↑ $R_A = I$

$= 2|I| = 2 \times 1 = 2$

(e) $|\mathbb{A}| = \pm |\mathbb{U}|$ ← as in $\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{U}$ (↖ permutation)

↑ depends on number of row swaps

(f) $|\mathbb{A}| = \pm \prod_{i=1}^{n} d_i$ ← pivots of $\mathbb{A}$

↖ follows from row op goodness ↖(c)

if one or more pivots = 0 $\Rightarrow |\mathbb{A}| = 0$.

(g) $\underset{n\times n}{|\mathbb{A}}\ \underset{n\times n}{\mathbb{B}|} = \underset{n\times n}{|\mathbb{A}|}\ \underset{n\times n}{|\mathbb{B}|}$ ← #bigdeal

Various proofs exist.

Monks say use row reduction:

Know $\mathbb{E}\,\mathbb{P}\,\mathbb{A} = \mathbb{R}^{e}_{\mathbb{A}} = \mathbb{D}_{\mathbb{A}}$ ← reduced row echelon form with pivots still in

- $\mathbb{E}$: all elimination matrices
- $\mathbb{P}$: permutation matrix
- $\mathbb{D}_{\mathbb{A}}$: pivot matrix for $\mathbb{A}$

(first presume all pivots $\neq 0$).

ok $|\mathbb{A}\,\mathbb{B}|$

$= \pm|\mathbb{E}\,\mathbb{P}\,\mathbb{A}\,\mathbb{B}| = \pm|\mathbb{D}_{\mathbb{A}}\,\mathbb{B}|$

(± depends on $\mathbb{P}$; det unchanged with row ops)

$$= \pm \left| \begin{bmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & & \ddots \; d_n \end{bmatrix} \begin{bmatrix} - \vec{b}_{1*} - \\ - \vec{b}_{2*} - \\ \vdots \\ - \vec{b}_{n*} - \end{bmatrix} \right|$$

$$= \pm \begin{vmatrix} - d_1 \vec{b}_{1*} - \\ - d_2 \vec{b}_{2*} - \\ \vdots \\ - d_n \vec{b}_{n*} - \end{vmatrix}$$

$$= \pm \left(\prod_{i=1}^{n} d_i\right) |\mathbb{B}| = |\mathbb{A}|\,|\mathbb{B}|.$$

↑ multilinearity

[Now if one or more pivots = 0, (← $|\mathbb{A}| = 0$) can see same row reductions leads to row of 0's $\Rightarrow |\mathbb{A}\mathbb{B}| = 0$ ✓.

(h) $|\mathbb{A}^{-1}| = \dfrac{1}{|\mathbb{A}|}$ ← from (g)

reason $|\mathbb{A}\,\mathbb{A}^{-1}| = |\mathbb{A}|\,|\mathbb{A}^{-1}|$

$\quad \| $

$|\mathbb{I}| = 1$

Note $|\mathbb{A}| = 0 \Rightarrow |\mathbb{A}^{-1}| = \infty$ ← ouch!

E18bp4

(i) If $\mathbb{A}$ is upper or lower triangular then $|\mathbb{A}|$ = product of entries on $\mathbb{A}$'s main diagonal.

Reason: row reduction on a triangular matrix requires no row swaps and does not change entries on main diag. Plus: zero leads to a zero row $\Rightarrow |\mathbb{A}|=0$.

ex

$$\begin{vmatrix} 4 & 77 & 16 \\ 0 & -3 & 17 \\ 0 & 0 & 2 \end{vmatrix} = (4)(-3)(2) = -24 \; ; \; \begin{vmatrix} 4 & 0 & 0 \\ 77 & 1 & 0 \\ 13 & 99 & 2 \end{vmatrix} = (4)(1)(2) = 8.$$

$$\begin{vmatrix} 4 & 77 & 16 \\ 0 & 0 & 17 \\ 0 & 0 & 2 \end{vmatrix} = (4)(0)(2) = 0.$$

last

(j) $|\mathbb{A}| = |\mathbb{A}^T|$ ← #groovy ← #bigdeal

Means: can use "column ops" in the same way as row ops & all results for rows work for columns too.

Reason

Use $\underset{n\times n}{\mathbb{P}}\,\underset{n\times n}{\mathbb{A}} = \underset{n\times n}{\mathbb{L}}\,\underset{n\times n}{\mathbb{U}}$ (use monks) → $|\mathbb{P}||\mathbb{A}| = |\mathbb{L}||\mathbb{U}|$ ☺*

$$(\mathbb{P}\mathbb{A})^T = (\mathbb{L}\mathbb{U})^T$$

$$\mathbb{A}^T\mathbb{P}^T = \mathbb{U}^T\mathbb{L}^T$$

Take determinants of both sides

$$|\mathbb{A}^T||\mathbb{P}^T| = |\mathbb{U}^T||\mathbb{L}^T|$$

triangular determinant is unchanged by transpose

$$|\mathbb{P}^T||\mathbb{A}^T| = |\mathbb{L}||\mathbb{U}|$$

↖ handle this

$\mathbb{P}$ is a permutation matrix, a shuffling of the identity matrix.

$\Rightarrow |\mathbb{P}| = \pm 1$ ← #row swaps

Also know

$\mathbb{P}^{-1} = \mathbb{P}^T$ so $|\mathbb{P}^T\mathbb{P}| = |\mathbb{I}| = 1$.

$\quad\quad\quad\quad\quad$ ‖ ← (g)

$\quad\quad\quad\quad |\mathbb{P}^T||\mathbb{P}|$

$\Rightarrow$ either $|\mathbb{P}| = |\mathbb{P}^T| = 1$, or $|\mathbb{P}| = |\mathbb{P}^T| = -1$ ← they match.

$\Rightarrow$ ☺* $|\mathbb{P}||\mathbb{A}^T| = |\mathbb{L}||\mathbb{U}| \Rightarrow |\mathbb{A}^T| = |\mathbb{A}|$.

Computing determinants: The way of the cofactor

* Recipe first, understanding later.
* Need a clean way to find determinants for eigenvalue problem $A\vec{v} = \lambda\vec{v}$
* Row operations helped us with results about determinants but are messy...

The story:

* $n \times n$ determinants are sums of n $(n-1)\times(n-1)$ determinants
* 3×3 determinants are sums of 3 2×2 determinants. recursive.

Example to work with:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 2 & 1 & 2 \end{bmatrix}.$$

Defn: M_{ij} is the $(n-1)\times(n-1)$ matrix formed by deleting the i^{th} row and j^{th} column of A; there are n^2 of these "minor matrices!"

exs

$$\begin{bmatrix} \not{1} & \not{1} & \not{1} \\ \not{0} & 3 & 2 \\ \not{2} & 1 & 2 \end{bmatrix} \quad M_{11} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$$

$$\begin{bmatrix} \not{1} & \not{1} & \not{1} \\ 0 & 3 & \not{2} \\ 2 & 1 & \not{2} \end{bmatrix} \quad M_{13} = \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \not{1} & 1 \\ \not{0} & \not{3} & \not{2} \\ 2 & \not{1} & 1 \end{bmatrix} \quad M_{22} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$

← A (3×3) has $3\times3 = 9$ minor matrices

Defn: $|M_{ij}|$ (det) is the ij^{th} **minor** of A

Defn: $C_{ij} = (-1)^{i+j}|M_{ij}|$ is the ij^{th} **cofactor** of A

$(-1)^{i+j} \Rightarrow$ checkerboard of +'s & −'s

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \\ + & - & + & - & \\ - & + & - & + & \\ \vdots & & & & \ddots \end{bmatrix}$$

Theorem:

The Determinant of $\mathbb{A}$ is given by the dot product of $\mathbb{A}$'s cofactors and $\mathbb{A}$'s entries along any one row or column.

ex Using row 1

$$|\mathbb{A}| \stackrel{\downarrow}{=} \sum_{j=1}^{n} C_{1j} a_{1j}$$

$$= \sum_{i=1}^{n} C_{i2} a_{i2}$$ ← column 2

#crazytown bananapants

Ex $\mathbb{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 2 \\ 2 & 1 & 2 \end{bmatrix}$

Let's first try row 1

⇒ Compute C_{11}, C_{12}, C_{13} using $C_{ij} = (-1)^{i+j} |\mathbb{M}_{ij}|$

$\mathbb{M}_{11} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}$; $|\mathbb{M}_{11}| = 3 \cdot 2 - 2 \cdot 1 = 4$; $C_{11} = (-1)^{1+1} \cdot 4 = 4$

$\mathbb{M}_{12} = \begin{bmatrix} 0 & 2 \\ 2 & 2 \end{bmatrix}$; $|\mathbb{M}_{12}| = 0 \cdot 2 - 2 \cdot 2 = -4$; $C_{12} = (-1)^{1+2} \cdot (-4) = 4$

$\mathbb{M}_{13} = \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix}$; $|\mathbb{M}_{13}| = 0 \cdot 1 - 3 \cdot 2 = -6$; $C_{13} = (-1)^{1+3}(-6) = -6$

$|\mathbb{A}| = C_{11} a_{11} + C_{12} a_{12} + C_{13} a_{13}$ ← dot product

$= (4)(1) + (4)(1) + (-6)(1) = 2$ // (as for row ops)

Now We can choose any row or column so let's do all of them at once. #crazy

Create cofactor matrix $\mathbb{C}$:

$$\mathbb{C} = \begin{bmatrix} 4 & 4 & -6 \\ -1 & 0 & 1 \\ -1 & -2 & 3 \end{bmatrix}$$ ← first row as above; } exercise

$$[a_{ij} C_{ij}] = \begin{bmatrix} 1\times4 & 1\times4 & 1\times(-6) \\ 0\times(-1) & 3\times0 & 2\times1 \\ 2\times(-1) & 1\times(-2) & 2\times3 \end{bmatrix} = \begin{bmatrix} 4 & 4 & -6 \\ 0 & 0 & 2 \\ -2 & -2 & 6 \end{bmatrix}$$

↑ direct product of elements

Magic:

$$\begin{bmatrix} 4 & 4 & -6 \\ 0 & 0 & 2 \\ -2 & -2 & 6 \end{bmatrix} \begin{matrix} (2) \\ (2) \\ (2) \end{matrix}$$

← row sums; (2) as above

column sums → (2) (2) (2)

#inconceivable

Cofactor method enables sneakiness: (excellent.)

Choose row or column with most zeros:

ex $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 2 & 1 & 3 \end{vmatrix} = 1 (-1)^{1+1} \begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} = 1 \cdot 1 \cdot (6+1) = 7$

(row 1: best choice; $a_{12} = a_{13} = 0$; $1 = a_{11}$; $(-1)^{1+1}\begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} = C_{11}$)

ex $\begin{vmatrix} 2 & 7 & 0 \\ 0 & 3 & 2 \\ -1 & 4 & 0 \end{vmatrix} = 2 (-1)^{2+3} \begin{vmatrix} 2 & 7 \\ -1 & 4 \end{vmatrix} = 2 \cdot (-1) \cdot (8+7) = -30.$

(col 3; row 2; $2 = a_{23}$; $(-1)^{2+3}\begin{vmatrix} 2 & 7 \\ -1 & 4 \end{vmatrix} = C_{23}$)

No need to compute cofactors associated with 0's in A. → avoid trauma

Fun: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ ← go along top row

$= a_{11} (-1)^{1+1} |a_{22}| + a_{12} (-1)^{1+2} |a_{21}|$

($(-1)^{1+1}|a_{22}| = C_{11}$; $(-1)^{1+2}|a_{21}| = C_{12}$; det of a 1x1 = length)

$= a_{11}a_{22} - a_{12}a_{21}$ ✓

↑ everything works

One more example: 4x4

$\begin{vmatrix} 2 & 2 & 3 & -1 \\ 0 & 7 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & 4 & 0 & 3 \end{vmatrix}$ ← most 0's

$= 7 \cdot (-1)^{2+2} \begin{vmatrix} 2 & 3 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{vmatrix}$ ← most 0's

($7 = a_{22}$)

$= 7 \cdot \left(3 \cdot (-1)^{3+3} \begin{vmatrix} 2 & 3 \\ -1 & 1 \end{vmatrix} \right)$

$= 7 \cdot 3 (2+3) = 7 \cdot 3 \cdot 5 = 105$

* Starting with, say, row 1 would have really hurt...

Determinants & $\mathbb{A}_{n\times n}\vec{x}=\vec{b}$
Cramer's rule and a formula for the inverse of $\mathbb{A}$ #inconceivable

Monks say try this for 3×3 matrices:

$$\mathbb{A}\begin{bmatrix} | & 0 & 0 \\ \vec{x} & 1 & 0 \\ | & 0 & 1 \end{bmatrix} = \begin{bmatrix} | & & \\ \mathbb{A}\vec{x} & \mathbb{A}\begin{bmatrix}0\\1\\0\end{bmatrix} & \mathbb{A}\begin{bmatrix}0\\0\\1\end{bmatrix} \\ | & & \end{bmatrix}$$

↑ $\mathbb{I}$ with $\vec{x}$ in first column

$$= \begin{bmatrix} | & | & | \\ \vec{b} & \vec{a}_2 & \vec{a}_3 \\ | & | & | \end{bmatrix} = \mathbb{B}_1$$

↑ call this

↑ $\mathbb{A}$ with first column replaced by $\vec{b}$.

Similarly

$$\mathbb{A}\begin{bmatrix} 1 & | & 0 \\ 0 & \vec{x} & 0 \\ 0 & | & 1 \end{bmatrix} = \begin{bmatrix} | & | & | \\ \vec{a}_1 & \vec{b} & \vec{a}_3 \\ | & | & | \end{bmatrix} ; \quad \mathbb{A}\begin{bmatrix} 1 & 0 & | \\ 0 & 1 & \vec{x} \\ 0 & 0 & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \vec{a}_1 & \vec{a}_2 & \vec{b} \\ | & | & | \end{bmatrix}$$

(↙ $\mathbb{B}_2$, ↙ $\mathbb{B}_3$)

Monks whisper "take determinants"....

$$|\mathbb{A}| \begin{vmatrix} | & 0 & 0 \\ \vec{x} & 1 & 0 \\ | & 0 & 1 \end{vmatrix} = |\mathbb{B}_1|$$

$\shortparallel$ x_1 (use row reduction on transpose)

$$\Rightarrow x_1 = \frac{|\mathbb{B}_1|}{|\mathbb{A}|}, \quad x_2 = \frac{|\mathbb{B}_2|}{|\mathbb{A}|}, \quad x_3 = \frac{|\mathbb{B}_3|}{|\mathbb{A}|}$$

↑ wait! we just solved $\mathbb{A}_{n\times n}\vec{x}=\vec{b}$!!

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{|\mathbb{A}|}\begin{bmatrix} |\mathbb{B}_1| \\ |\mathbb{B}_2| \\ \vdots \\ |\mathbb{B}_n| \end{bmatrix} \quad !!!$$

Problems: ① only works for $\underline{n\times n}$ ← ok for normal equations
② computing determinants is horribly slow
③ must recompute for new $\vec{b}$

Main utility: theoretical.

E18dp1

Let's use Cramer's rule to find $\mathbb{A}^{-1}$:

Monks say solve these special $\mathbb{A}\vec{x}=\vec{b}$ problems:

$$\mathbb{A}\vec{x}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \leftarrow \vec{b}$$

$$\mathbb{A}\vec{x}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

$$\mathbb{A}\vec{x}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

← generalizes to $n \times n$.

$$\Rightarrow \mathbb{A}\begin{bmatrix}| & | & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \\ | & | & |\end{bmatrix} = \begin{bmatrix}| & | & | \\ \mathbb{A}\vec{x}_1 & \mathbb{A}\vec{x}_2 & \mathbb{A}\vec{x}_3 \\ | & | & |\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = \mathbb{I}.$$

← so must have $\mathbb{A}^{-1}$ here

Use Cramer's rule and work with example $\mathbb{A} = \begin{bmatrix}1&1&1\\0&3&2\\2&1&2\end{bmatrix}$:

Solve $\mathbb{A}\vec{x} = \begin{bmatrix}1\\0\\0\end{bmatrix}$ ($\vec{b}$) first with $\vec{x} = \frac{1}{|\mathbb{A}|}\begin{bmatrix}|\mathbb{B}_1|\\|\mathbb{B}_2|\\|\mathbb{B}_3|\end{bmatrix}$

$$\mathbb{B}_1 = \begin{bmatrix}1&1&1\\0&3&2\\0&1&2\end{bmatrix},\quad \mathbb{B}_2 = \begin{bmatrix}1&1&1\\0&0&2\\2&0&2\end{bmatrix},\quad \mathbb{B}_3 = \begin{bmatrix}1&1&1\\0&3&0\\2&1&0\end{bmatrix}$$

↑ $\mathbb{M}_{11}$ ↑ $\mathbb{M}_{12}$ ↑ $\mathbb{M}_{13}$

$|\mathbb{B}_1| = C_{11}$, $|\mathbb{B}_2| = C_{12}$, $|\mathbb{B}_3| = C_{13}$

← cofactors of $\mathbb{A}$! →

$$\begin{bmatrix}1\\0\\0\end{bmatrix}: \quad \vec{x}_1 = \frac{1}{|\mathbb{A}|}\begin{bmatrix}C_{11}\\C_{12}\\C_{13}\end{bmatrix} \leftarrow \text{top row of } \mathbb{C}$$

$$\begin{bmatrix}0\\1\\0\end{bmatrix}: \quad \vec{x}_2 = \frac{1}{|\mathbb{A}|}\begin{bmatrix}C_{21}\\C_{22}\\C_{23}\end{bmatrix} \leftarrow \text{middle row of } \mathbb{C}$$

$$\begin{bmatrix}0\\0\\1\end{bmatrix}: \quad \vec{x}_3 = \frac{1}{|\mathbb{A}|}\begin{bmatrix}C_{31}\\C_{32}\\C_{33}\end{bmatrix} \leftarrow \text{bottom row of } \mathbb{C}$$

Combine:
$$\mathbb{A}^{-1} = \begin{bmatrix}| & | & | \\ \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \\ | & | & |\end{bmatrix} = \frac{1}{|\mathbb{A}|}\begin{bmatrix}C_{11}&C_{21}&C_{31}\\C_{12}&C_{22}&C_{32}\\C_{13}&C_{23}&C_{33}\end{bmatrix}$$

see transpose of $\mathbb{C}$

$$\Rightarrow \mathbb{A}^{-1} = \frac{1}{|\mathbb{A}|}\mathbb{C}^T$$

#inconceivable

Using earlier calculations

$$\mathbb{A}^{-1} = \frac{1}{2}\begin{bmatrix}4&-1&-1\\4&0&-2\\-6&1&3\end{bmatrix}$$

($2 = |\mathbb{A}|$; matrix $= \mathbb{C}^T$)

Check

$$\frac{1}{2}\begin{bmatrix}4&-1&-1\\4&0&-2\\-6&1&3\end{bmatrix}\begin{bmatrix}1&1&1\\0&3&2\\2&1&2\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2&0&0\\0&2&0\\0&0&2\end{bmatrix} = \mathbb{I}$$

✓

Algebraic & Geometric Multiplicity of eigenvalues

"Some matrices are bad matrices"
— traditional matrix-fu saying.

ex
$$\mathbb{A} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$
$(4-\lambda)(7-\lambda)^2 = 0$

Solve $\mathbb{A}\vec{v} = \lambda\vec{v}$:

$\lambda_1 = 4$, $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$; $\lambda_2 = 7$, $\lambda_3 = 7$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ ← y-z plane

$(\mathbb{A} - 7\mathbb{I})\vec{v} = \vec{0}$ gives a plane of vectors.

$\dim N(\mathbb{A} - 7\mathbb{I}) = 2$

Defn: Algebraic Multiplicity is # times an eigenvalue appears as a root of $|\mathbb{A} - \lambda\mathbb{I}| = 0$ ← characteristic equation of $\mathbb{A}$

Defn: Geometric Multiplicity is the dimension of the eigenspace associated with an eigenvalue λ

$\Leftrightarrow \dim N(\mathbb{A} - \lambda\mathbb{I})$

ex a.m. of $\lambda = 4$ is 1, g.m. is 1

a.m. of $\lambda = 7$ is 2, g.m. is 2 ← healthy

Observation $1 \le$ g.m. $\le$ a.m.

ex
$$\mathbb{A} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

Find eigenvalues: Solve $|\mathbb{A} - \lambda\mathbb{I}| = 0$

$$0 = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 2 & 1-\lambda & 0 \\ 0 & 2 & 1-\lambda \end{vmatrix} = (1-\lambda)^3 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 1$$

$\Rightarrow \lambda = 1$ has algebraic multiplicity of 3.

Find evectors: Solve $(\mathbb{A} - (1)\mathbb{I})\vec{v} = \vec{0}$.

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \vec{v} = c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ only}$$
(p p F; 1-d; $c \in \mathbb{R}$)

$\Rightarrow \lambda = 1$ has geometric multiplicity of 1

* $N(\mathbb{A} - \lambda\mathbb{I})$ not big enough... dim = 1
* $\mathbb{A}$ is a bad matrix... and does not have a full complement of eigenvectors ← basis for eigenspace

Sneaky Monk Tricks (SMTs) for Eigen Stuff:

E19ap1

All about $\underset{n\times n}{A}\,\underset{n\times 1}{\vec{v}} = \underset{\#}{\lambda}\,\underset{n\times 1}{\vec{v}}$

Recap: Solve by
1. Finding λs as roots of $|A-\lambda I|=0$ — Characteristic Equation (use cofactor method)

&

2. For each distinct λ, solving the nullspace equation $(A-\lambda I)\vec{v}=\vec{0}$ for λ's eigenspace

Our helper example:

$A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ (symmetry will be meaningful)

$\lambda_1 = \frac{3}{2} > 1$ $\quad \vec{v}_1 \propto \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = \frac{1}{2} < 1$ $\quad \vec{v}_2 \propto \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

SMT #1

$$|A| = \prod_{i=1}^{n} \lambda_i$$

The determinant of A is equal to the product of its eigenvalues

Check:

$|A| = \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{vmatrix} = 1\cdot 1 - \frac{1}{2}\cdot\frac{1}{2} = 3/4$

$\lambda_1 \cdot \lambda_2 = \frac{3}{2}\cdot\frac{1}{4} = 3/4$ ✓

Why?

general characteristic equation; ex $\left(\frac{3}{2}-\lambda\right)\left(\frac{1}{2}-\lambda\right)$

$|A-\lambda I| = (\lambda_1 - \lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$ (set $\lambda = 0$; λ_1, λ_2 are #)

$\Rightarrow |A| = \prod_{i=1}^{n} \lambda_i$

from before: $|A| = \pm \prod_{i=1}^{n} d_i$ (pivots)

SMT #2

Defn Trace of $A = Tr(A)$ = Sum A's main diagonal elements $= \sum_{i=1}^{n} a_{ii}$

$$Tr(A) = \sum_{i=1}^{n} \lambda_i$$

Our example: $Tr\left(\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\right) = 1+1 = 2$

$\sum_{i=1}^{n} \lambda_i = \frac{3}{2} + \frac{1}{2} = 2$ ✓

General 2x2

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (\lambda_1-\lambda)(\lambda_2-\lambda)$$

$= (a-\lambda)(d-\lambda) - b\cdot c$

$(-\lambda)^2 + (a+d)(-\lambda) + ad - bc$

$= (-\lambda)^2 + (\lambda_1+\lambda_2)(-\lambda) + \lambda_1\lambda_2$

matching:

$\lambda_1 + \lambda_2 = a+d = Tr(A)$

$\lambda_1\lambda_2 = ad-bc = |A|$

E19ap2

General $n \times n$:

$$|A - \lambda I| = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

$$= (-\lambda)^n + \left(\sum_{i=1}^{n} \lambda_i\right)(-\lambda)^{n-1} + \cdots$$

$$= (-\lambda)^n + (\mathrm{Tr}\, A)(-\lambda)^{n-1} + \cdots$$

One thing: Check $\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i$ (easy)

SMT3

$$|A| = \pm \prod_{i=1}^{n} d_i = \prod_{i=1}^{n} \lambda_i$$

(± depends on row swaps required to uncover U)

SMT4

If $\overset{n\times n}{A}\vec{v} = \lambda \vec{v}$ then $A^k \vec{v} = \lambda^k \vec{v}$

$$A^k \vec{v} = A^{k-1}(A\vec{v}) = \lambda A^{k-1}\vec{v} = \cdots = \lambda^k \vec{v}$$

SMT5

If $A\vec{v} = \lambda\vec{v}$ then $(A + tI)\vec{v} = (\lambda + t)\vec{v}$ (shift)

$$(A + tI)\vec{v} = A\vec{v} + tI\vec{v} = \lambda\vec{v} + t\vec{v} = (\lambda + t)\vec{v}$$

SMT6

If $A\vec{v} = \lambda\vec{v}$ then $A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}$ if A^{-1} exists

$$\underbrace{A^{-1}A}_{I}\vec{v} = \lambda A^{-1}\vec{v}$$

$$A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}$$

matches SMT4 for $k = -1$

SMT7 If A's eigenvalues are all different from each other then A's eigenvectors are linearly independent and form a basis for $\mathbb{R}^n$.

Reason:

Assume λ's are distinct and look at $\vec{v}_1$ & $\vec{v}_2$

If dependent, $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ for some $c_1, c_2 \neq 0$

(a) $A(c_1\vec{v}_1 + c_2\vec{v}_2) = A\vec{0}$

$c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 = \vec{0}$ ...... (1)

(b) $\lambda_2 \times (c_1\vec{v}_1 + c_2\vec{v}_2) = \lambda_2 \times \vec{0}$

$c_1\lambda_2\vec{v}_1 + c_2\lambda_2\vec{v}_2 = 0$ .... (2)

(2)−(1): $c_1(\lambda_2 - \lambda_1)\vec{v}_1 = \vec{0}$

$\Rightarrow \lambda_1 \neq \lambda_2$

build up from here for $n \times n$'s

SMT8 eigenvalues & eigenvectors of A & B are not simply related to those of AB & A+B

- If A & B share an eigenvector $\vec{v}$ with eigenvalues λ_A & λ_B then $(A+B)\vec{v} = (\lambda_A + \lambda_B)\vec{v}$ & $(AB)\vec{v} = \lambda_A\lambda_B\vec{v}$

but this is generally not the case.

Why diagonal matrices make us happy

ex

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 17 \end{bmatrix} \Rightarrow A^k = \begin{bmatrix} 3^k & 0 & 0 \\ 0 & (-7)^k & 0 \\ 0 & 0 & (17)^k \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} 3x_1 \\ -7x_2 \\ 17x_3 \end{bmatrix}$$

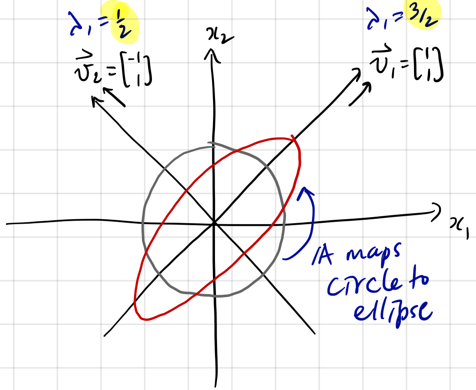
how A changes $\vec{x}$ is simple

$\lambda_1 = 3,\ \lambda_2 = -7,\ \lambda_3 = 17$

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

natural or standard basis for R^3

$$A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$$



- If we could rotate the axes, A's action would be simple
- Big idea: change from standard basis to eigenvector basis and find happiness

Diagonalization is just the best

- Let's assume $A_{n\times n}$ has n linearly independent eigenvectors (form a basis for R^n)

- know $A\vec{v}_i = \lambda_i \vec{v}_i$ for $i=1,\dots,n$

Monk's whisper: Create a new matrix with A's eigenvectors:

$$\underset{n\times n}{S} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}_{n\times n}$$

Consider:

$$\underset{n\times n}{A}\ \underset{n\times n}{S} \overset{\text{matrix-fu}}{=} \begin{bmatrix} | & | & & | \\ A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \\ | & | & & | \end{bmatrix}$$

($A\vec{v}_1$: $n\times n$, $n\times 1$ → $n\times 1$)

$$= \begin{bmatrix} | & | & & | \\ \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \\ | & | & & | \end{bmatrix} \overset{\text{sneaky}}{=} \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{bmatrix}$$

$$= \underset{n\times n}{S}\ \underset{n\times n}{\Lambda} \quad \leftarrow \text{capital } \lambda$$

E20bp1

Let's assume A is a good matrix meaning its eigenvectors form a basis for R^n $\Leftrightarrow$ S^{-1} exists

$$\Rightarrow AS = S\Lambda \Rightarrow$$

(postmultiply by S^{-1})

Diagonalization:

$$\underset{n\times n}{A} = \underset{n\times n}{S}\ \underset{n\times n}{\Lambda}\ \underset{n\times n}{S^{-1}}$$

↑ an amazing factorization

- We say A and Λ are similar ↑ similarity transformation

- We begin (more soon) to see how $A\vec{x}$ works:

$$A\vec{x} = S\Lambda S^{-1}\vec{x}$$

① S^{-1}: changes representation of $\vec{x}$ from standard basis to A's eigenvector basis

② Λ: simple multiplication because Λ is diagonal

③ S: changes back to standard basis representation

Big deal: If A is diagonalizable, then A is really a diagonal matrix when viewed in the right way.

E206p2

Example: $A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$ $\lambda_1 = \frac{3}{2}$ $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

note: symmetry

$\lambda_2 = \frac{1}{2}$ $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

our choice (any multiple would work)

$\Rightarrow S = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

note: $S^{-1} = S^T$

$\Rightarrow S^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ ← $|S| = 2$

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

$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 3/2 & 0 \\ 0 & 1/2 \end{bmatrix}$

$\Rightarrow A = S \Lambda S^{-1}$

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

Let's see how this is super useful for (1) $A\vec{x}$ & (2) A^k

(1) Examine what happens for $\vec{x} = 2\vec{v}_1 + 2\vec{v}_2$

$\vec{x} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$ $A\vec{x}$ in 3 ways ⇓

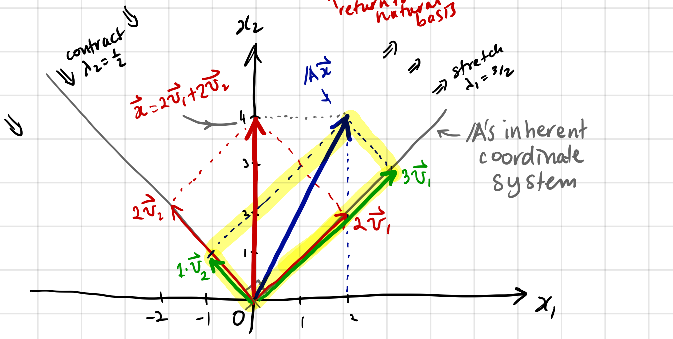
(i) $A\vec{x} = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ ← no great understanding

(ii) $A\vec{x} = A\left(2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right) = \frac{3}{2}\cdot 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\cdot 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ ✓ (may not know this)

(iii) $A\vec{x} = S\Lambda S^{-1}\vec{x} = S\Lambda\left(\frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix}\right) = S\Lambda\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ ← representation of $\vec{x}$ in eigenvector basis

$= S\begin{bmatrix} 3/2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = S\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ ← in eigenvector basis still

$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ ✓ ← return to natural basis



(2) A^k for $k = 0, \pm 1, \pm 2, \ldots$

$$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2 S^{-1}$$

($S^{-1}S = I$)

$$A^3 = (S\Lambda S^{-1})(S\Lambda S^{-1})(S\Lambda S^{-1})$$

$$= S\Lambda^3 S^{-1}$$

$$A^k = S\Lambda^k S^{-1}$$

← super easy!!

Super easy to compute!

$$\Lambda^k = \begin{bmatrix} \lambda_1^k & & 0 \\ & \lambda_2^k & \\ 0 & & \ddots \; \lambda_n^k \end{bmatrix}$$

→ clearly important

$|\lambda_i| < 1$

$|\lambda_i| = 1$

$|\lambda_i| > 1$

← can see that largest eigenvalue will dominate Λ^k and A^k

Ex/

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}^{523}$$

$$= \underbrace{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{S} \underbrace{\begin{bmatrix} \left(\frac{3}{2}\right)^{523} & 0 \\ 0 & \left(\frac{1}{2}\right)^{523} \end{bmatrix}}_{\Lambda^{523}} \underbrace{\frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}_{S^{-1}}$$

($\left(\frac{3}{2}\right)^{523}$ blows up; $\left(\frac{1}{2}\right)^{523} \to 0$)

$$\simeq \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)^{523} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

More goodness:

$2^0 = 1$

- $A^0 = S\Lambda^0 S^{-1} = SIS^{-1} = I$
- $A^{-1} = S\Lambda^{-1}S^{-1}$ works:

$$\underbrace{(S\Lambda^{-1}S^{-1})}_{A^{-1}}\underbrace{(S\Lambda S^{-1})}_{A} = I$$

($S^{-1}S = I$, $\Lambda^{-1}\Lambda = I$, $SS^{-1} = I$)

$A^{\frac{1}{2}} = S\Lambda^{1/2}S^{-1}$ works too!!!

$A^{1/2}A^{1/2} = A^1$ ✓

E20C P1

Fibonacci number finder

note: clearly a monk

From before:

$$F_{k+2} = F_{k+1} + F_k \quad \text{with} \quad F_0 = F_1 = 1$$

← Fibonacci sequence

$$\vec{f}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \vec{f}_0$$

note symmetry... ; $\vec{f}_{k-1}$; $\vec{f}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Mission: Diagonalize $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$

usual thing

$$|A - \lambda I| = 0 \Rightarrow \lambda^2 - \lambda - 1 = 0$$

← very famous

Find $\lambda_1 = \frac{1+\sqrt{5}}{2} = \varphi$ (golden ratio) $\Rightarrow \vec{v}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$ ← good hygiene

$\lambda_2 = \frac{1-\sqrt{5}}{2}$ $\quad \vec{v}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$ ← keep $\sqrt{5}$'s at bay

Note: $\lambda_1 + \lambda_2 = 1 = \mathrm{Tr}(A)$

$\lambda_1 \lambda_2 = -1 = |A|$

Three pieces:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$S = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}$$

$$S^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}$$

aside

$$\varphi = \frac{1+\sqrt{5}}{2} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}}$$

← most irrational number; continued fraction

$$\Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}$$

A ; S ; Λ^k ; S^{-1} — λ_1^k ← grows; λ_2^k ← decays

So:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{bmatrix} (\lambda_1^{k+1} - \lambda_2^{k+1}) \\ (\lambda_1^k - \lambda_2^k) \end{bmatrix}$$

2×1 (not a 2×2)

$$F_k = \frac{1}{\lambda_1 - \lambda_2} \left(\lambda_1^k - \lambda_2^k \right)$$

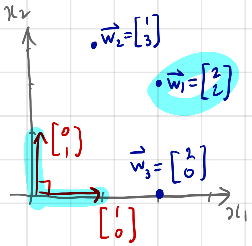
$$= \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right)$$

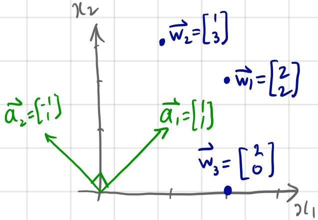
← dominates ; ← disappears

#delicious

Cool beans: $\frac{F_{k+1}}{F_k} \to \frac{1+\sqrt{5}}{2} = \varphi$ as $k \to \infty$ ($= \lambda_1$)

The gentle art of changing basis:





So far, we've expressed all vectors in terms of the standard (or natural) basis.

ex $\vec{w}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

- What's the representation of $\vec{w}_1$, $\vec{w}_2$, and $\vec{w}_3$ in terms of the new basis $\{\vec{a}_1, \vec{a}_2\}$?
- How do we do this systematically?

By solving an $A\vec{x} = \vec{b}$ problem!!!

The set up for $\vec{w}_1$:

$$\vec{w}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = c_1 \vec{a}_1 + c_2 \vec{a}_2 = \begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

$$= A \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \equiv M \vec{w}_1^{(a)}$$

(in natural basis) (column picture) (coordinate of $\vec{w}_1$ in $\{\vec{a}_1, \vec{a}_2\}$ basis) (often used for basis transformation)

E21ap1

$$\Rightarrow \vec{w}_1^{(a)} = M^{-1} \vec{w}_1$$

(inverse takes us from natural to new basis)

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

($\vec{w}_1$ in $\{\vec{a}_1, \vec{a}_2\}$ basis.)

Similarly: $\vec{w}_2^{(a)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{w}_3^{(a)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

To change back: $\vec{w}_i = M \vec{w}_i^{(a)}$

We say:

In basis $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$, $\vec{w}_1$ is represented as $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$

In basis $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$, $\vec{w}_1$ is represented as $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$

(specify same point/vector in space)

Big deal: The vector $\vec{w}_1$ never changes but our representation does.

Big deal:

$$A = S \Lambda S^{-1}$$

- S (M): change basis only (from eigenvector to normal)
- Λ (does the real work): "change" vector
- S^{-1} (M^{-1}): change basis only (from normal to eigenvector)

Symmetry and the Spectral Theorem

We know:

- Diagonalization is joyous and empowering
- $\mathbb{A}_{n\times n}$ can only be diagonalized if it has n linearly independent eigenvectors
- Trouble potentially arises when eigenvalues are repeated (algebraic multiplicity > 1)
  - ↳ May not end up with a full eigenspace

Bonus truths:

- If one or more eigenvalues = 0, $\mathbb{A}^{-1}$ does not exist
  - ↳ $|\mathbb{A}| = 0$
- But $\mathbb{A}$ may still be diagonalizable
  - ↳ depends on eigenvectors

An amazing matrix truth:

If $\mathbb{A}_{n\times n}$ is real & symmetric, i.e. $\mathbb{A} = \mathbb{A}^T$, then (+ a_{ij} is real for all i,j)

- $\mathbb{A}$ always has n linearly independent eigenvectors and is therefore always diagonalizable

even better

1. All of $\mathbb{A}$'s eigenvalues are real (no complex numbers ⇒ no rotations)
2. $\mathbb{A}$'s eigenvectors form an orthogonal basis for R^n !!!

proofs later

EZZap1

We get so excited, we replace $\mathbb{S} = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_n]$ with $\mathbb{Q} = [\hat{v}_1\ \hat{v}_2 \cdots \hat{v}_n]$ (unit vectors)

because we realize we have an orthogonal matrix.

And because $\mathbb{Q}^{-1} = \mathbb{Q}^T$ (← saves a lot of trouble), our diagonalization takes on a new level of majesty:

$$\mathbb{A} = \mathbb{Q}\Lambda\mathbb{Q}^T$$

Wow!!

More amazingness:

$$\mathbb{A} = \mathbb{Q}\Lambda\mathbb{Q}^T = \begin{bmatrix} | & | & & | \\ \hat{v}_1 & \hat{v}_2 & \cdots & \hat{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} - \hat{v}_1^T - \\ - \hat{v}_2^T - \\ \vdots \\ - \hat{v}_n^T - \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \hat{v}_1 & \hat{v}_2 & \cdots & \hat{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} - \lambda_1\hat{v}_1^T - \\ - \lambda_2\hat{v}_2^T - \\ \vdots \\ - \lambda_n\hat{v}_n^T - \end{bmatrix}$$

$\mathbb{A}\vec{x}$ broken into clean pieces

$$= \lambda_1\hat{v}_1\hat{v}_1^T + \lambda_2\hat{v}_2\hat{v}_2^T + \cdots + \lambda_n\hat{v}_n\hat{v}_n^T = \sum_{i=1}^{n} \lambda_i\hat{v}_i\hat{v}_i^T$$

outer products

projection operators!!

for $\mathbb{A}\vec{x}$, each one chops out a piece of $\vec{x}$ and then scales by λ.

Spectral Theorem for Symmetric Matrices

LE22ap2

Example: $A = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} = A^T$ ← symmetric!

Use unit vectors for eigenvectors

$S = Q = \begin{bmatrix} \hat{v}_1 & \hat{v}_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$S^{-1} = Q^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ take transpose! easy!

$$\begin{array}{cccc} & S & \Lambda & S^{-1} \\ A = & Q & \Lambda & Q^T \end{array} =$$

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

and

$$A = \sum_{i=1}^{n} \lambda_i \hat{v}_i \hat{v}_i^T$$

$$= \underbrace{\left(\frac{3}{2}\right)}_{\lambda_1} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\hat{v}_1} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix}}_{\hat{v}_1^T} + \underbrace{\left(\frac{1}{2}\right)}_{\lambda_2} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}}_{\hat{v}_2} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \end{bmatrix}}_{\hat{v}_2^T}$$

$A\vec{x}$: breaks $\vec{x}$ into two orthogonal pieces for which $A\hat{v}_1$ & $A\hat{v}_2$ are very simple. and then recombines.

Why the spectral theorem works

① All of A's eigenvalues are real

(no rotations)

Assume $A = A^T$ and A's entries are real

Given $A\vec{v} = \lambda\vec{v}$, we test to see if λ can be complex: $\lambda = a + bi$, $b \neq 0$ ($i = \sqrt{-1}$)

Denote complex conjugate by over bar: $\overline{a+bi} = a - bi$

Result: $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$

$A\vec{v} = \lambda\vec{v}$ (A: $n\times n$)

monks →

$\overline{A}\,\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}$

↓ real

$A\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}$

take transpose

$\overline{\vec{v}}^T A^T = \overline{\lambda}\,\overline{\vec{v}}^T$

↓ symmetry

$\overline{\vec{v}}^T A = \overline{\lambda}\,\overline{\vec{v}}^T$

↓ postmultiply by $\vec{v}$

Monk sneakiness: monks ↓ premultiply by $\overline{\vec{v}}^T$

$\overline{\vec{v}}^T A\vec{v} = \lambda\overline{\vec{v}}^T\vec{v}$ $\qquad$ $\overline{\vec{v}}^T A\vec{v} = \overline{\lambda}\,\overline{\vec{v}}^T\vec{v}$

See everything matches except λ & $\overline{\lambda}$

$\Rightarrow \lambda = \overline{\lambda}$ so λ is real

E22 bp1

② A's eigenvectors form an orthogonal basis for R^n (crazy!)

- Again have $A = A^T$ and A is real
- We want to show $\vec{v}_i^T\vec{v}_j = 0$ if $i \neq j$
- Work up to full story...

First: If A's eigenvalues are all distinct (i.e., each has algebraic multiplicity 1):

More Monk Sneakiness

$$\vec{v}_1^T A\vec{v}_2 = \vec{v}_1^T(A\vec{v}_2)$$

($1\times n$, $n\times n$, $n\times 1$ — a number)

$$= (A^T\vec{v}_1)^T\vec{v}_2 = \lambda_1\vec{v}_1^T\vec{v}_2 \quad (A^T = A)$$

$$= \vec{v}_1^T(\lambda_2\vec{v}_2) = \lambda_2\vec{v}_1^T\vec{v}_2$$

(what we're interested in...)

but $\lambda_1 \neq \lambda_2$ so these can only be equal if $\vec{v}_1^T\vec{v}_2 = 0$

OK

- What if an eigenvalue is repeated?
- We're worried we won't have enough eigenvectors ....

A suggestive pair of examples:

Not symmetric:

$$\mathbb{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \neq \mathbb{A}^T$$

$\lambda_1 = \lambda_2 = 1$ repeated

only one dimension for eigenspace

$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ #sadness

Symmetric:

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

$\lambda_1 = \lambda_2 = 1$ repeated

$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

2-d eigenspace healthy

Tweaks:

$$\mathbb{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1+\epsilon \end{bmatrix} \neq \mathbb{A}^T$$

(ϵ small)

$\lambda_1 = 1, \lambda_2 = 1+\epsilon$ ← now distinct

$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ✓, $\vec{v}_2 = \begin{bmatrix} 1 \\ \epsilon \end{bmatrix}$ ✓

See as $\epsilon \to 0$, eigenvectors become the same

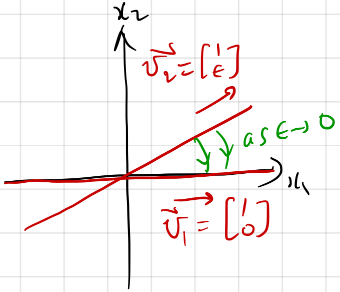
$$\mathbb{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1+\epsilon \end{bmatrix} = \mathbb{A}^T$$

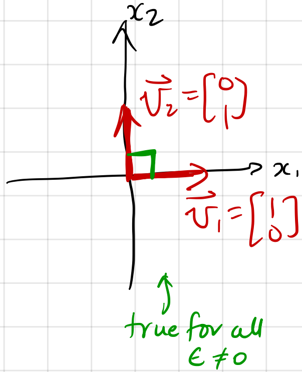
$\lambda_1 = 1, \lambda_2 = 1+\epsilon$ ← distinct

$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ✓, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ✓ (as before)

$\vec{v}_1^T \vec{v}_2 = 0$

Eigenvectors do not budge as $\epsilon \to 0$





Idea: smooth change in tweaks ($\epsilon \to 0$) cannot lead to eigenvectors snapping into orthogonal directions

$\Rightarrow$ orthogonality is preserved //.

Requires more work to show in general but we have the basic story here.

Surprising things about traces

Defn Trace of $\mathbb{A}_{n\times n}$ = $\text{Tr}(\mathbb{A})$
= Sum of the entries of $\mathbb{A}$'s main diagonal
$= \sum_{i=1}^{n} a_{ii}$

ex $\text{Tr}\left(\begin{bmatrix} 3 & 0 & 2 \\ 2 & -1 & -1 \\ 1 & 2 & 4 \end{bmatrix}\right) = 3 + (-1) + 4 = 6$

From earlier: $\text{Tr}(\mathbb{A}) = \sum_{i=1}^{n} \lambda_i$

Now, two more things:

(1) $\text{Tr}(\mathbb{A}_{n\times n}\mathbb{B}_{n\times n}) = \text{Tr}(\mathbb{B}_{n\times n}\mathbb{A}_{n\times n})$

Reason: $\text{Tr}(\mathbb{A}\mathbb{B}) = \sum_{i=1}^{n} (\mathbb{A}\mathbb{B})_{ii}$

$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ji}$ ← from defn of multiplication; inner product of ith row of A & jth column of B

swap everything →

$= \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij} = \text{Tr}(\mathbb{B}\mathbb{A})$

Generalizes:

$\text{Tr}(\mathbb{A}\mathbb{B}\mathbb{C}) = \text{Tr}((\mathbb{A}\mathbb{B})(\mathbb{C})) = \text{Tr}(\mathbb{C}(\mathbb{A}\mathbb{B}))$ (two matrices)

$= \text{Tr}((\mathbb{C}\mathbb{A})(\mathbb{B})) = \text{Tr}((\mathbb{B})(\mathbb{C}\mathbb{A}))$

↑ any cycling leaves Trace unchanged

(2) If $\mathbb{A} = \mathbb{S}\Lambda\mathbb{S}^{-1}$ ← sadly not possible for all matrices

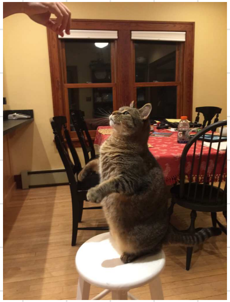
then $\text{Tr}(\mathbb{A}) = \text{Tr}(\mathbb{S}\Lambda\mathbb{S}^{-1})$ — cycle to front

$= \text{Tr}(\mathbb{S}^{-1}\mathbb{S}\Lambda) = \text{Tr}(\Lambda)$ ($\mathbb{S}^{-1}\mathbb{S} = \mathbb{I}$)

$= \sum_{i=1}^{n} \lambda_i$ #delicious

So: a very enjoyable proof of $\text{Tr}(\mathbb{A}) = \sum_{i=1}^{n} \lambda_i$
(but does not work if $\mathbb{A}$ is not diagonalizable)

Pratchett does tricks for treats... →



Positive Definite Matrices

← matrices that are really sure about themselves

E23ap1

defn: A Positive Definite Matrix is a real, symmetric matrix with positive eigenvalues, i.e., $\lambda_i > 0$, $i = 1, \dots, n$

If a matrix is real and symmetric with $\lambda_i \geq 0$ and at least one eigenvalue equal to zero, then we say it is **Semi-Positive Definite**

We recall with alacrity that real, symmetric matrices always have (1) Real eigenvalues ← flipping stretching & shrinking
(2) Eigenvectors that form an orthonormal basis for $\mathbb{R}^n$

$A = Q \Lambda Q^T$

Turns out that also having $\lambda_i > 0$ or $\lambda_i \geq 0$ is an excellent bonus feature ...

Menu
- i/ How to spot a PDM ← surprising results from monks; pivots ⟺ eigenvalues
- ii/ Why we like PDMs (and SPDMs)

Places we'll go, things we'll see:

- * $2x_1^2 + 2x_1x_2 + 2x_2^2 = 1 \Rightarrow$ matrices
- * what elimination really does for symmetric matrices
- * Completing the square

Three example 2x2 matrices:

$$A_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix}; \quad A_3 = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix}$$

For A_1: $\lambda_1 = +3$, $\lambda_2 = +1$ (↑ computing happens elsewhere) ⇑ PDM :)

For A_2: $\lambda_1 = \sqrt{5}$, $\lambda_2 = -\sqrt{5}$:(

For A_3: $\lambda_1 = -1$, $\lambda_2 = -3$:(

Problem: Finding eigenvalues can be pretty hard for real matrices

- We only want to know signs of the eigenvalues
- Could there be a sneaky way? ← especially one that helps computers

SMT #37

If $\mathbb{A}_{n\times n} = \mathbb{A}^T$ & $\mathbb{A}$ is real then:

\# positive eigenvalues = \# positive pivots

\# negative eigenvalues = \# negative pivots

\# zero eigenvalues = zero pivots

↑ #Crazytownbananapants

- Very peculiar: eigenvalues and pivots come from very different parts of matrixology
- Recall we already know for general $\mathbb{A}_{n\times n}$ that $|\mathbb{A}| = \prod_{i=1}^{n} \lambda_i = \pm \prod_{i=1}^{n} d_i$
- SMT#37 says more for real symmetric matrices

Big deal: $\mathbb{A}$ is a PDM if all $d_i > 0$
Pivots are much easier to compute than eigenvalues

LE23ap2

Beautiful reason:

Consider $\mathbb{A}_2 = \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix}$ $\quad \lambda_1 = \sqrt{5}, \quad \lambda_2 = -\sqrt{5}$

find pivots using $\mathbb{L}\mathbb{U}$ decomposition

$R_2' = R_2 - (-\frac{1}{2})R_1$, l_{21}

$$\mathbb{A}_2 = \mathbb{L}\mathbb{U} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & -5/2 \end{bmatrix}$$

($l_{21} = -\frac{1}{2}$; $d_1 = 2$, $d_2 = -5/2$)

Because $\mathbb{A}_2$ is symmetric ($\mathbb{A}_2 = \mathbb{A}_2^T$), we can go further:

$$\mathbb{A}_2 = \mathbb{L}\mathbb{D}\mathbb{L}^T = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$$

Let's think about this parametrized matrix:

$$\mathbb{B}(l_{21}) = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & l_{21} \\ 0 & 1 \end{bmatrix}$$

When $l_{21} = -\frac{1}{2}$, we have $\mathbb{B}(-\frac{1}{2}) = \mathbb{A}_2$

Next: What happens as we move from $l_{21} = -\frac{1}{2}$ to $l_{21} = 0$?

E23ap3

$$\mathbb{B}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -\frac{5}{2} \end{bmatrix}$$

(under the matrices: $\mathbb{I}$, $\mathbb{D}$, $\mathbb{I}$; the result's diagonal entries are labelled d_1 and d_2, and the result $= \mathbb{D}$)

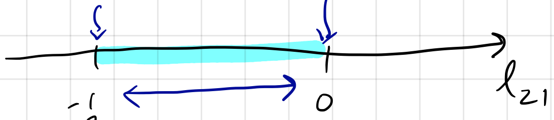
Observations:

- For diagonal matrices, pivots $\equiv$ eigenvalues
- $\mathbb{B}(\ell_{21})$ has pivots $d_1 = 2$ and $d_2 = -\frac{5}{2}$ independent of ℓ_{21}
- $\det(\mathbb{B}(\ell_{21})) = d_1 \cdot d_2 = (2)(-\frac{5}{2}) = -5$ again independent of ℓ_{21}

Big connections:

- We also know $\det(\mathbb{B}(\ell_{21})) = \lambda_1 \cdot \lambda_2$ must $= -5$ for all ℓ_{21} (with $\lambda_1 = \sqrt{5}$, $\lambda_2 = -\sqrt{5}$ for $\ell_{21} = -\frac{1}{2}$)
- As ℓ_{21} changes, the eigenvalues change BUT they cannot pass through 0 as then the determinant would be 0 $(\neq -5)$
- When $\ell_{21} = 0$, $\mathbb{B}(0) = \mathbb{D}$ is diagonal and the pivots and eigenvalues match up: $d_1 \equiv \lambda_1$, $d_2 \equiv \lambda_2$
- Therefore as ℓ_{21} moves away from 0, the eigenvalues must maintain the same signs as the pivots
- Argument assumes all pivots $\neq 0$; proof is tweakable

$\mathbb{B}(-\frac{1}{2}) = \mathbb{A}_2$ $\mathbb{B}(0) = \mathbb{D}$



$\ell_{21} = -\frac{1}{2}$		$\ell_{21} = 0$	
$\lambda_1 = \sqrt{5}$	⟷	$\lambda_1 = 2$	change but signs remain same
$\lambda_2 = -\sqrt{5}$	⟷	$\lambda_2 = -\frac{5}{2}$	
$d_1 = 2$	⟷	$d_1 = 2$	invariant
$d_2 = -\frac{5}{2}$	⟷	$d_2 = -5/2$	

General argument:

Given $\underset{n\times n}{\mathbb{A}} = \mathbb{L}\mathbb{D}\mathbb{L}^T$ create $\hat{\mathbb{L}}(t) = \mathbb{I} + t(\mathbb{L} - \mathbb{I})$

$t = 0$: $\hat{\mathbb{L}}(0) = \mathbb{I}$

$t = 1$: $\hat{\mathbb{L}}(1) = \mathbb{L}$

$$\begin{cases} \mathbb{B}(t) = \hat{\mathbb{L}}(t)\,\mathbb{D}\,\hat{\mathbb{L}}^T(t) \\ \mathbb{B}(0) = \mathbb{D} \ \ \& \ \ \mathbb{B}(1) = \mathbb{A} \end{cases}$$

- As before, pivots don't change as we vary t from 1 to 0
- Same story: Eigenvalues cannot change sign as t varies
- Signs of eigenvalues must match signs of pivots

Positive Definite Matrices in the Wild:

Menu for 23b, c, d:

- $\vec{x}^T A \vec{x}$ and ellipses and other functions
- Completing the Square
- Cholesky factorization

Idea: re-express polynomial functions using matrices.
← especially PDMs

key construct: $\vec{x}^T A \vec{x}$ where $A = A^T$

$[\;]_{1\times n}[\;]_{n\times n}[\;]_{n\times 1}$ → 1×1 ← a scalar

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

General 2×2 example:

$$\vec{x}^T A \vec{x} = \underset{1\times 2}{[x_1 \;\; x_2]} \underset{2\times 2}{\begin{bmatrix} a & b \\ b & c \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}$$

$A = A^T$

$$= \underset{1\times 2}{[x_1 \;\; x_2]} \underset{2\times 1}{\begin{bmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{bmatrix}}$$

inner product

$$= ax_1^2 + bx_1x_2 + bx_1x_2 + cx_2^2$$

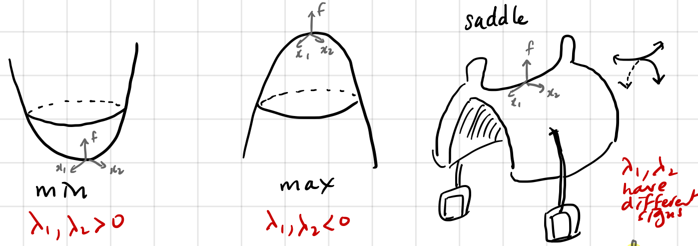
$$= ax_1^2 + 2bx_1x_2 + cx_2^2 = f(x_1, x_2)$$

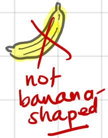
← height

easy to go back this way

E23bp1

$f(x_1, x_2)$ could be:





The Story:
f has a minimum at $x_1 = x_2 = 0$
iff A is Positive Definite

Why? (1) $\vec{x}^T A \vec{x} = 0$ at $\vec{x} = 0$

(2) Consider what happens as $\vec{x}$ moves away from $\vec{0}$

Write $\vec{x} = \sum_{i=1}^{n} c_i \hat{v}_i$ ← A's eigenvector basis

possible because $A = A^T$ ⇒ eigenvectors form an orthonormal basis for $\mathbb{R}^n$

$$\vec{x}^T A \vec{x} = \left(\sum_{i=1}^{n} c_i \hat{v}_i^T\right) A \left(\sum_{j=1}^{n} c_j \hat{v}_j\right) = \left(\sum_{i=1}^{n} c_i \hat{v}_i^T\right)\left(\sum_{j=1}^{n} c_j A \hat{v}_j\right)$$

($A\hat{v}_j = \lambda_j \hat{v}_j$)

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_j c_i c_j \hat{v}_i^T \hat{v}_j = \sum_{i=1}^{n} \lambda_i c_i^2 > 0$$

($\hat{v}_i^T \hat{v}_j$: 1 if $i=j$, 0 otherwise; $c_i^2 \geq 0$)

for all $\{c_i\}$ iff all $\lambda_i > 0$

Ex 1.

Does $f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2$ have a maximum at $x_1 = x_2 = 0$?

Answer: Yes if eigenvalues for f's A are both positive

$\iff$ A's pivots are both positive

(1) Construct $\vec{x}^T A \vec{x}$

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 2x_2^2$$

$$= [x_1 \quad x_2] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

split evenly ← A_1 from before

(2) Determine pivots

$d_1 = 2$, $d_2 = 3/2$ $\Rightarrow$ $\lambda_1 > 0$, $\lambda_2 > 0$ $\Rightarrow$ f has a minimum

Ex 2.

$$f(x_1, x_2) = 2x_1^2 - 2x_1x_2 - 2x_2^2$$

(1) Construct $\vec{x}^T A \vec{x}$

$$f(x_1, x_2) = [x_1 \quad x_2] \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

← A_2 from before

(2) Determine pivots:

$d_1 = 2$, $d_2 = -5/2$ $\Rightarrow$ $\lambda_1 > 0$, $\lambda_2 < 0$ $\Rightarrow$ saddle

Alternate definition:
$A = A^T$ is positive definite iff $\vec{x}^T A \vec{x} > 0$ for all $\vec{x} \neq \vec{0}$

Completing the Square = Gaussian Elimination !!

↑ for square symmetric matrices

E23cp1

Idea We could approach question of determining kinds of stationary points by creating clear squares and then looking at signs.

Ex $f(x_1, x_2) = 2x_1^2 - 2x_1 x_2 + 2x_2^2 = [x_1, x_2]\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ ($\Uparrow$ $\mathbb{A}_1$ from before)

(complete square here; leave x_2 as constant)

$= 2\left(x_1^2 - (x_2)\, x_1\right) + 2x_2^2$

$= 2\left(\left(x_1 - \frac{x_2}{2}\right)^2 - \frac{x_2^2}{4} \right) + 2x_2^2$

$= 2\left(x_1 - \frac{1}{2}x_2\right)^2 + \frac{3}{2}x_2^2$

d_1, l_{21}, d_2 ← what!??

from $\mathbb{A}_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

Ex $f(x_1, x_2) = 2x_1^2 - 2x_1 x_2 - 2x_2^2$ $\Leftarrow \mathbb{A}_2 = \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix}$

$= 2\left(x_1 - \frac{1}{2}x_2\right)^2 - \frac{5}{2}x_2^2$

d_1, l_{21}, d_2 — same thing

In general for 2x2s:

$ax_1^2 + 2bx_1x_2 + cx_2^2$

$= a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(\frac{ac - b^2}{a}\right)x_2^2$

$d_1 = a$, $l_{21} = \frac{b}{a}$, $d_2 = \frac{ac-b^2}{a}$

see $x_1 + \frac{b}{a}x_2$ as a new variable...

Does completing the square always work like this?

Yes! (for symmetric matrices)

$\vec{x}^T \mathbb{A} \vec{x}$ ← any quadratic in n variables ($\mathbb{A} = \mathbb{A}^T$)

$= \vec{x}^T (\mathbb{L}\mathbb{D}\mathbb{L}^T)\vec{x}$ ← see as a variable transformation

$= (\mathbb{L}^T\vec{x})^T \mathbb{D} (\mathbb{L}^T\vec{x})$

$= \vec{y}^T \mathbb{D} \vec{y}$

$= d_1 y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2$ ← if pivots > 0 then $\mathbb{A}$ is a PDM

$\vec{y}^T \; \mathbb{D} \; \vec{y}$

Super bonus: if $\mathbb{A}$ is a PDM then

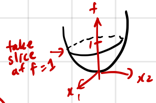
$\mathbb{A} = \tilde{\mathbb{L}}\tilde{\mathbb{L}}^T$ with $\tilde{\mathbb{L}} = \mathbb{L}\mathbb{D}^{\frac{1}{2}}$

↑ Cholesky Factorization

↑ all real numbers lower triangular

- Even better for $\mathbb{A}\vec{x} = \vec{b}$

Principle Axis Theorem



Consider $2x_1^2 + 2x_1x_2 + 2x_2^2 = 1$

↗ Equation of an ellipse oriented at an angle to standard axes

Matrixify:

$$[x_1 \;\; x_2]\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$$

↓ use $A = Q\Lambda Q^T$ ← one of our magic friends

$$[x_1 \;\; x_2]\,\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1$$

(columns of Q: $\hat{v}_1$, $\hat{v}_2$; diagonal of Λ: λ_1, λ_2; rows of Q^T: $\hat{v}_1^T$, $\hat{v}_2^T$)

Q Λ Q^T

$\vec{y}^T = (Q^T\vec{x})^T$ $\vec{y} = Q^T\vec{x}$

$$\Rightarrow [y_1 \;\; y_2]\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 1$$

($\vec{y}^T$, Λ, $\vec{y}$)

$$\Rightarrow 3y_1^2 + y_2^2 = 1$$

Completely clear in y_1-y_2 coordinate system

clearly an ellipse

E23dp1

What's this new coordinate system?:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = Q^T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

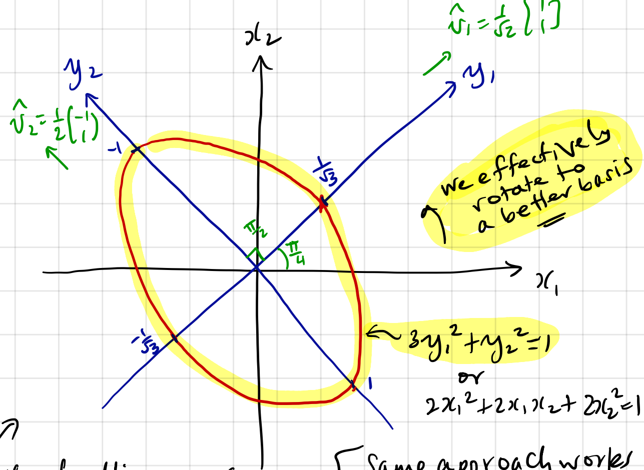
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} x_1 + x_2 \\ -x_1 + x_2 \end{bmatrix}$$

best ↓

Also: See Q as IM ⇒ basis transformat.

$\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ = new basis

($\hat{v}_1$, $\hat{v}_2$)



↗ length of ellipse axes $= \frac{1}{\sqrt{\lambda_i}}$

Same approach works for higher dimensional football

Singular Value Decomposition

Insert ominous organ music

Big deal:

- Matrix Factorizations encode our understanding of problems and greatly enable our methods

$\mathbb{P}\mathbb{A} = \mathbb{L}\mathbb{U}$ $\Rightarrow$ Simultaneous Equations $\underset{m\times n}{\mathbb{A}}\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$ rectangular

$\mathbb{A} = \mathbb{Q}\mathbb{R}$ $\Rightarrow$

$\mathbb{A} = S \Lambda S^{-1}$

$\mathbb{A} = \mathbb{Q}\Lambda\mathbb{Q}^T$

$\Rightarrow$ $\underset{n\times 1}{\vec{x}'} = \overset{n\times n}{\mathbb{A}}\underset{n\times 1}{\vec{x}}$, & $\mathbb{A}^k$, $\mathbb{A}\vec{v} = \lambda\vec{v}$ — square only

- All have limitations
- We love diagonalization for example

but

(1) $\mathbb{A}$ must be $n \times n$

(2) $\mathbb{A}$ must have n linearly independent eigenvectors

(3) Eigenvector basis may not be orthogonal (only guaranteed if $\mathbb{A} = \mathbb{A}^T$)

In attempting to overcome these problems, we'll find a factorization that works for all matrices plus

- helps us identify the most important features of a system (pages on the web, supreme court decisions, data in general, building blocks of images, ...)
- Completes our "Big Picture" story for $\mathbb{A}\vec{x} = \vec{b}$

  → Fundamental Theorem of Linear Algebra

[Theoretical story first, then some nutritious examples

Eigen Story: $\underset{n\times n}{\mathbb{A}}\vec{v} = \lambda\vec{v}$ ← eigenvectors may not form a basis :(

same direction :)

We give this up to (i) accommodate $m \times n$ matrices & (ii) ensure orthogonality of bases

↑ want this

New plan:

$$A\hat{v}_i = \sigma_i \hat{u}_i$$

(A: $m \times n$, $\hat{v}_i$: $n \times 1$, $\hat{u}_i$: $m \times 1$)

- A real
- $\hat{v}_i$ unit vector
- σ_i "singular value"
- $\hat{u}_i$ singular vector, unit vector

where:
- $\hat{v}_i \perp \hat{v}_j$ if $i \neq j$, $\hat{v}_i \in R^n$ (row space + nullspace)
- $\hat{u}_i \perp \hat{u}_j$ if $i \neq j$, $\hat{u}_i \in R^m$ (column space + left nullspace)
- $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_r > 0$ (r = rank)

- the $\hat{v}_i$ form an orthonormal basis for R^n
- the $\hat{u}_i$ form an orthonormal basis for R^m

How $A\vec{x}$ works:

(1) Transform $\vec{x}$ to $\{\hat{v}_i\}$ basis in R^n

(2) Send $\hat{v}_i$'s to $\hat{u}_i$'s and multiply by σ_i

(3) Transform from $\{\hat{u}_i\}$ basis to standard basis in R^m

Yet another great moment in Matrixology: E24ap2

Singular Value Decomposition

$$A = U \Sigma V^T$$

(A: $m \times n$, U: $m \times m$, Σ: $m \times n$, V: $n \times n$)

Monks tell us to believe

c.f. $A = Q \Lambda Q^T$

$$U = \begin{bmatrix} | & | & & | \\ \hat{u}_1 & \hat{u}_2 & \cdots & \hat{u}_m \\ | & | & & | \end{bmatrix}$$

$m \times m$ ← $\vec{b}, \vec{p}, \vec{e}, \vec{y}$

$$V = \begin{bmatrix} | & | & & | \\ \hat{v}_1 & \hat{v}_2 & \cdots & \hat{v}_n \\ | & | & & | \end{bmatrix}$$

$n \times n$ ← $\vec{x}$, $\vec{x}_r, \vec{x}_n$, $\vec{x}_p, \vec{x}_n$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \ 0 \end{bmatrix}$$

$m \times n$

all zeros (off-diagonal)

same shape as A

r = rank

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

E24ap3

Let's see how this all works:

Claim → $A = U\Sigma V^T$ with $A\hat{v}_i = \sigma_i \hat{u}_i$

Monks: "Try A^TA, grasshopper" (symmetric for all A) — okay...

[key: know $(A^TA)^T = A^TA$ for any A]

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T)$$

(A^T: $n\times m$, A: $m\times n$ → $n\times n$)

$$= V^{T^T}\Sigma^T \underbrace{U^TU}_{I}\Sigma V^T$$

$[U^{-1} = U^T,\ V^{-1} = V^T]$

$$= V\Sigma^T\Sigma V^T$$

(Σ^T: $n\times m$, Σ: $m\times n$)

$$= \begin{bmatrix} | & | & & | \\ \hat{v}_1 & \hat{v}_2 & \cdots & \hat{v}_n \\ | & | & & | \end{bmatrix}\begin{bmatrix} \sigma_1^2 & & & & & \\ & \sigma_2^2 & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_r^2 & & \\ & 0 & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}\begin{bmatrix} - \hat{v}_1^T - \\ - \hat{v}_2^T - \\ \vdots \\ - \hat{v}_n^T - \end{bmatrix}$$

(square, $n\times n$)

Looks a lot like: $Q\Lambda Q^T$ (sneaky monks)

But will A^TA always be so wonderfully diagonalizable? ← drama

Monk Joy

A^TA ($n\times m$, $m\times n$ → $n\times n$) is real, symmetric and therefore (1) eigenvalues are real (2) eigenvectors form an orthonormal basis for R^n

Monk Joy ↑ augmented

$$\vec{x}^T(A^TA)\vec{x} = (A\vec{x})^T(A\vec{x}) = \|A\vec{x}\|^2 \geq 0$$

($A\vec{x}$: $n\times 1$; A: $n\times n$, $\vec{x}$: $n\times 1$; ↖ length; ← has to be true)

⇒ A^TA is Semi-Positive Definite

⇒ A^TA's eigenvalues are all ≥ 0

⇒ $\sigma_i = \sqrt{\lambda_i} \geq 0$ is all good

Upshot: Diagonalize A^TA to find σ_i's and $\hat{v}_i$'s

Monks chant: "AA^T! AA^T! AA^T!..."

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$$

($m\times n$, $n\times m$ → $m\times m$; symmetric)

$$= U\Sigma \underbrace{V^TV}_{I}\Sigma^TU^T$$

$$= U\Sigma\Sigma^TU^T$$

($m\times m$)

$$= \begin{bmatrix} | & | & & | \\ \hat{u}_1 & \hat{u}_2 & \cdots & \hat{u}_m \\ | & | & & | \end{bmatrix}\begin{bmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r^2 & \\ & & & & 0 \ddots 0 \end{bmatrix}\begin{bmatrix} - \hat{u}_1^T - \\ - \hat{u}_2^T - \\ \vdots \\ - \hat{u}_n^T - \end{bmatrix}$$

($m\times m$)

surprising!! ↓ A^TA & AA^T must have same non-zero eigenvalues

More upshot: Diagonalize AA^T to find $\hat{u}_i$'s and, again, σ_i's

How do we know $A\hat{v}_i = \sigma_i \hat{u}_i$?

We have: $A^T A \hat{v}_i = \sigma_i^2 \hat{v}_i$

(1) Monks: $\hat{v}_i^T (A^T A \hat{v}_i)$

$$(A\hat{v}_i)^T (A\hat{v}_i) = \hat{v}_i^T (\sigma_i^2 \hat{v}_i)$$

$$\|A\hat{v}_i\|^2 \quad (A\hat{v}_i: m\times 1 \text{ vector}) \qquad = \sigma_i^2 \hat{v}_i^T \hat{v}_i \quad (\hat{v}_i^T \hat{v}_i = 1) \; = \sigma_i^2$$

$\Rightarrow \|A\hat{v}_i\|^2 = \sigma_i^2$

$\Rightarrow \|A\hat{v}_i\| = \sigma_i$

So we have the right length ✓

(2) Monks again

$$A(A^T A \hat{v}_i) = A(\sigma_i^2 \hat{v}_i) \quad \text{(eigenvalue equation)}$$

$$= (AA^T)(A\hat{v}_i) \quad (A\hat{v}_i: m\times 1 \text{ vector})$$

$$= \sigma_i^2 (A\hat{v}_i) \quad (m\times 1 \text{ vector})$$

$\Rightarrow$ $A\hat{v}_i$ is an eigenvector of AA^T with eigenvalue σ_i^2

$\Rightarrow$ $A\hat{v}_i \propto \hat{u}_i$

(1) + (2) $\Rightarrow$ $A\hat{v}_i = \sigma_i \hat{u}_i$

Important details:

- Choose $\hat{u}_i$'s direction to match $A\hat{v}_i$ (makes sure $\sigma_i > 0$)
- If we have found $\hat{v}_i$ already, $\boxed{\hat{u}_i = \frac{1}{\sigma_i} A\hat{v}_i}$ is best way to compute $\hat{u}_i$
- $A\hat{v}_i = \vec{0}$ for $i = r+1, r+2, \ldots, n$ (null space basis)
- $\hat{u}_i$ for $i = r+1, r+2, \ldots, m$ = left nullspace basis

E24ap5

One last piece:

For $\mathbb{A} = \mathbb{A}^T$, we had $\mathbb{Q}\Lambda\mathbb{Q}^T$ and therefore

$$\underset{n\times n}{\mathbb{A}} = \lambda_1 \hat{v}_1 \hat{v}_1^T + \lambda_2 \hat{v}_2 \hat{v}_2^T + \dots + \lambda_n \hat{v}_n \hat{v}_n^T$$

outer products = projection operators

$\mathbb{A}$ = sum of n rank 1 matrices.

For SVD:

$$\mathbb{A} = \underset{m\times m}{\begin{bmatrix} | & | & & | \\ \hat{u}_1 & \hat{u}_2 & \cdots & \hat{u}_m \\ | & | & & | \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & 0 \ddots 0 \end{bmatrix}} \underset{n\times n}{\begin{bmatrix} - & \hat{v}_1^T & - \\ - & \hat{v}_2^T & - \\ & \vdots & \\ - & \hat{v}_n^T & - \end{bmatrix}}$$

$$= \underset{m\times m}{\begin{bmatrix} | & | & & | \\ \hat{u}_1 & \hat{u}_2 & \cdots & \hat{u}_m \\ | & | & & | \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} - & \sigma_1 \hat{v}_1^T & - \\ - & \sigma_2 \hat{v}_2^T & - \\ - & \sigma_r \hat{v}_r^T & - \\ - & \vec{0}^T & - \\ & \vdots & \\ - & \vec{0}^T & - \end{bmatrix}}$$

r rows non-zero

$m-r$ rows of zilchos

$$= \sigma_1 \underset{m\times 1}{\hat{u}_1} \underset{1\times n}{\hat{v}_1^T} + \sigma_2 \underset{m\times 1}{\hat{u}_2} \underset{1\times n}{\hat{v}_2^T} + \dots + \sigma_r \underset{m\times 1}{\hat{u}_r} \underset{1\times n}{\hat{v}_r^T}$$

$$\mathbb{A} = \sum_{i=1}^{r} \sigma_i \hat{u}_i \hat{v}_i^T$$

See $\mathbb{A}$ as a superposition of r outer product rank 1 matrices of diminishing significance

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

- Each rank 1 matrix is a piece of Scottish Tartan
- SVD makes approximation of large matrices vigorous
- Speak of best rank 1, best rank 2, ... approximations

E246p1

SVD Example Calculation #1:

For $\mathbb{A} = \begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix}$ (2×2, $m\times n$) find $\mathbb{A} = \mathbb{U}\Sigma\mathbb{V}^T$ ($\mathbb{U}$: $m\times m$, 2×2; Σ: $m\times n$, 2×2; $\mathbb{V}^T$: $n\times n$, 2×2)

(1) Find $\hat{v}_i$'s and σ_i's using $\mathbb{A}^T\mathbb{A}$

- $\mathbb{A}^T\mathbb{A} = \begin{bmatrix} 3 & 3 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 18 & -6 \\ -6 & 2 \end{bmatrix}$ ← note symmetry
- Solve $|\mathbb{A}^T\mathbb{A} - \lambda\mathbb{I}| = 0$

$$\Rightarrow 0 = \begin{vmatrix} 18-\lambda & -6 \\ -6 & 2-\lambda \end{vmatrix} = (18-\lambda)(2-\lambda) - 36 = 36 - 20\lambda + \lambda^2 - 36 = \lambda(\lambda - 20)$$

(the 36 and -36 are crossed out)

$$\Rightarrow \lambda_1 = 20 = \sigma_1^2, \quad \lambda_2 = 0 = \sigma_2^2 \Rightarrow \sigma_1 = \sqrt{20}, \quad \text{"}\sigma_2\text{"} = 0$$

σ_1 ← row space; σ_2 ← null space $\mathbb{A}\vec{x} = \vec{0}$

- $\lambda_1 = 20$: Solve $(\mathbb{A}^T\mathbb{A} - 20\mathbb{I})\vec{v}_1 = \vec{0}$

$$\Rightarrow \left[\begin{array}{cc|c} -2 & -6 & 0 \\ -6 & -18 & 0 \end{array}\right] \Rightarrow \hat{v}_1 = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

rows must be multiples of each other for 2x2's

- $\lambda_2 = 0$: Solve $(\mathbb{A}^T\mathbb{A} - 0\mathbb{I})\vec{v}_2 = \vec{0}$ (the $0\mathbb{I}$ is crossed out)

$$\Rightarrow \left[\begin{array}{cc|c} 18 & -6 & 0 \\ -6 & 2 & 0 \end{array}\right] \Rightarrow \hat{v}_2 = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

as promised, $\hat{v}_1 \perp \hat{v}_2$

So far:

$$\mathbb{V} = \begin{bmatrix} | & | \\ \hat{v}_1 & \hat{v}_2 \\ | & | \end{bmatrix} = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{20} & 0 \\ 0 & 0 \end{bmatrix}$$

$r = 1$ (↑ rank) $\quad \sigma_1 = \sqrt{20}$

- Need $\mathbb{U}$ as well

Either solve for eigenthings of $\mathbb{A}\mathbb{A}^T$ ($= \begin{bmatrix} 10 & 10 \\ 10 & 10 \end{bmatrix}$) → $\lambda_1 = 20$, $\hat{u}_1$; → $\lambda_2 = 0$, $\hat{u}_2$

- Better: Use $\mathbb{A}\hat{v}_i = \sigma_i\hat{u}_i$

$$\Rightarrow \hat{u}_i = \frac{1}{\sigma_i}\mathbb{A}\hat{v}_i$$

$$\hat{u}_1 = \frac{1}{\sqrt{20}}\mathbb{A}\frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{20}}\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

($\sqrt{20}$: σ_1; $\frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ -1 \end{bmatrix}$: $\hat{v}_1$)

$$= \frac{1}{\sqrt{20}}\frac{1}{\sqrt{10}}\begin{bmatrix} 10 \\ 10 \end{bmatrix} = \frac{1}{\sqrt{2}}\frac{1}{10}\begin{bmatrix} 10 \\ 10 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

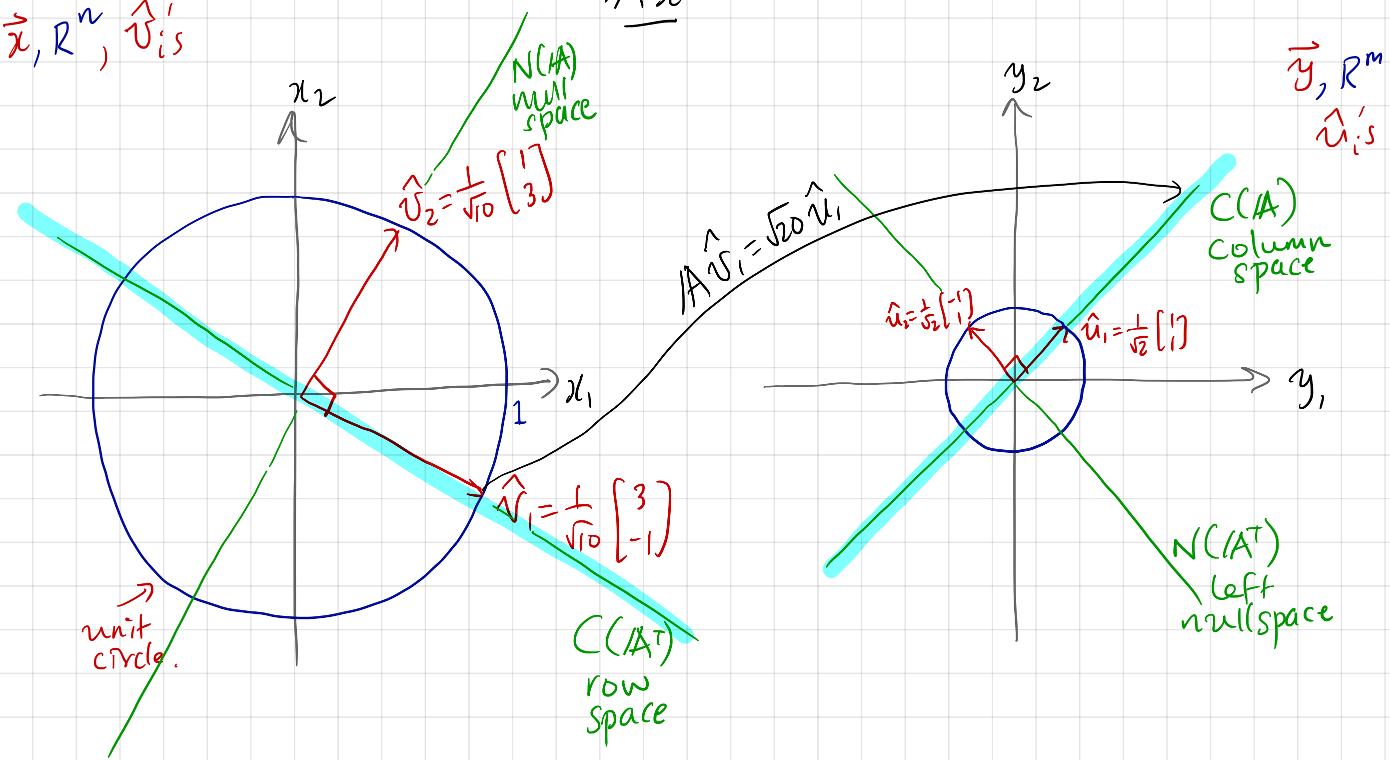
For $\hat{u}_2$, we just need a vector orthogonal to $\hat{u}_1$

By inspection: $\hat{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

← better way to represent $\mathbb{A}$

$$\mathbb{A} = \begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix} = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\mathbb{U}}\underbrace{\begin{bmatrix} \sqrt{20} & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma}\underbrace{\frac{1}{\sqrt{10}}\begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}}_{\mathbb{V}^T}$$

$\mathbb{A}\vec{x}$



- See $\mathbb{A}$ sends $C(\mathbb{A}^T)$ to $C(\mathbb{A})$ with a stretch factor of $\sqrt{20}$.
- $\mathbb{A}$'s action between $C(\mathbb{A}^T)$ & $C(\mathbb{A})$ is invertible

SVD Example Calculation #2

E24cp1

Factorize $A = \frac{1}{5}\begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix}$ as $U \Sigma V^T$

- Diagonalize $A^T A$

$$A^T A = \frac{1}{5}\begin{bmatrix} 2 & 10 \\ 11 & 5 \end{bmatrix} \frac{1}{5}\begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix}$$

symmetric

$$= \frac{1}{25}\begin{bmatrix} 104 & 72 \\ 72 & 146 \end{bmatrix} = \frac{2}{25}\begin{bmatrix} 52 & 36 \\ 36 & 73 \end{bmatrix}$$

SMT #731

if $B\vec{v} = \lambda \vec{v}$
then
$cB\vec{v} = c\lambda \vec{v}$
$\Rightarrow B'\vec{v} = (c\lambda)\vec{v}$

If $\vec{v}$ is an eigenvector of B with eigenvalue λ
then $\vec{v}$ is an eigen vector of cB with eigenvalue $c\lambda$

Find λ's for $\begin{bmatrix} 52 & 36 \\ 36 & 73 \end{bmatrix}$

Solve $|A^T A - \lambda I| = 0$

$$0 = \begin{vmatrix} 52-\lambda & 36 \\ 36 & 73-\lambda \end{vmatrix} = (52-\lambda)(73-\lambda) - (36)^2$$

$$= 3796 - 125\lambda + \lambda^2 - 1296$$

$$= \lambda^2 - 125\lambda + 2500$$

(25 + 100)

$$= (\lambda - 25)(\lambda - 100) \Rightarrow \begin{matrix} \lambda_1 = 100 \\ \lambda_2 = 25 \end{matrix}$$

for $\frac{25}{2} A^T A$

$$\times \frac{2}{25} \Rightarrow \begin{matrix} \lambda_1 = 8 = \sigma_1^2 \\ \lambda_2 = 2 = \sigma_2^2 \end{matrix} \Rightarrow \begin{matrix} \sigma_1 = \sqrt{8} \\ \sigma_2 = \sqrt{2} \end{matrix}$$

for $A^T A$

- $\lambda_1 = 8$: $\frac{2}{25}\left[\begin{matrix} -48 & 36 \\ 36 & -27 \end{matrix}\middle|\begin{matrix} 0 \\ 0 \end{matrix}\right] \Rightarrow \hat{v}_1 = \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

Solve $(A^T A - \lambda I)\vec{v} = \vec{0}$

- $\lambda_2 = 2$: $\frac{2}{25}\left[\begin{matrix} 27 & 36 \\ 36 & 48 \end{matrix}\middle|\begin{matrix} 0 \\ 0 \end{matrix}\right] \Rightarrow \hat{v}_2 = \frac{1}{5}\begin{bmatrix} -4 \\ 3 \end{bmatrix}$

(could choose $\hat{v}_2 = -\frac{1}{5}\begin{bmatrix} -4 \\ 3 \end{bmatrix}$)

now have

$$V = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix} \quad \& \quad \Sigma = \begin{bmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

- Now find $\hat{u}_1$ & $\hat{u}_2$

$$\hat{u}_i = \frac{1}{\sigma_i} \mathbb{A} \hat{v}_i$$ ← best way

$$\hat{u}_1 = \frac{1}{\sqrt{8}} \frac{1}{5} \begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$= \frac{1}{\sqrt{8}} \frac{1}{25} \begin{bmatrix} 50 \\ 50 \end{bmatrix} = \frac{1}{\sqrt{8}} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

unit vectors guaranteed

$$\hat{u}_2 = \frac{1}{\sqrt{2}} \frac{1}{5} \begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix} \frac{1}{5} \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \frac{1}{25} \begin{bmatrix} 25 \\ -25 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbb{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$ ✓.

Could also diagonalize $\mathbb{A}\mathbb{A}^T$: /E24cp2

$$\mathbb{A}\mathbb{A}^T = \frac{1}{5} \begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 2 & 10 \\ 11 & 5 \end{bmatrix}$$

$$= \frac{1}{25} \begin{bmatrix} 125 & 75 \\ 75 & 125 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

Find $\lambda_1 = 8$, $\lambda_2 = 2$

$$\hat{u}_1 = \pm \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \hat{u}_2 = \pm \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

not sure about signs

⇒ Still have to compute $\hat{u}_i = \frac{1}{\sigma_i} \mathbb{A} \hat{v}_i$ ✓.

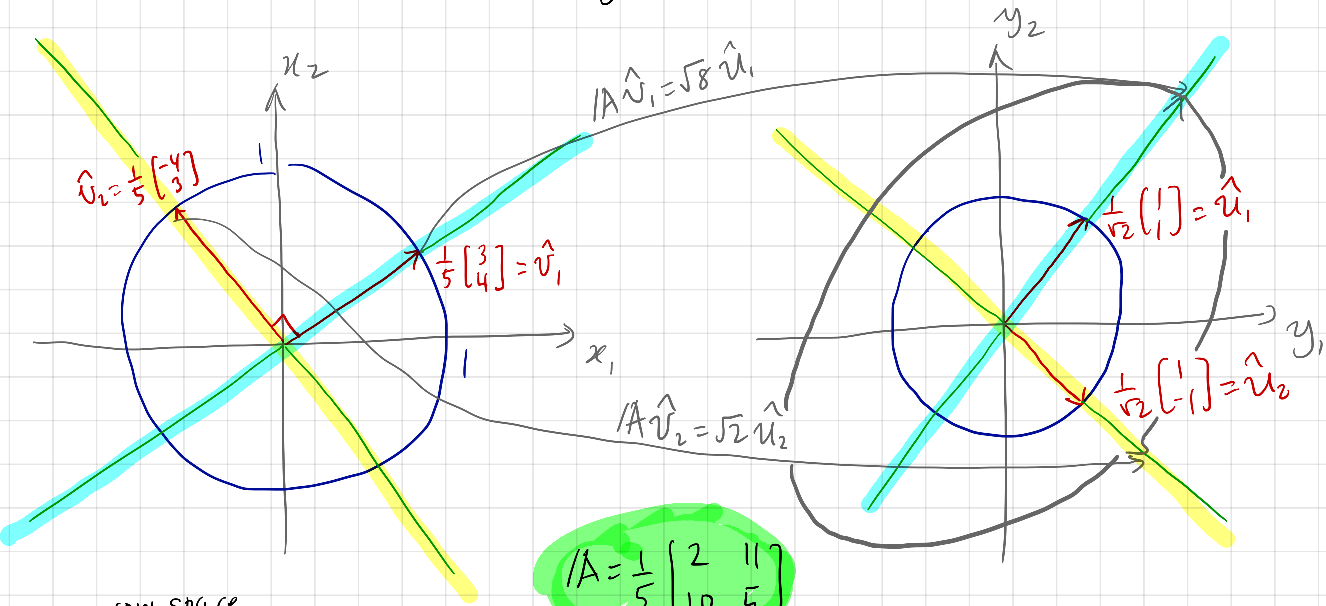
Overall:

$$\mathbb{A} = \frac{1}{5} \begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix} = \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbb{U}} \underbrace{\begin{bmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{5} \begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}}_{\mathbb{V}^T}$$

$R^2, R^n, \vec{x}$'s

$$A\vec{x} = \vec{b}$$
or
$$\vec{y} = A\vec{x}$$

E24cp3



$$A = \frac{1}{5}\begin{bmatrix} 2 & 11 \\ 10 & 5 \end{bmatrix}$$

row space

$$C(A^T) \equiv R^2$$

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

Big deal: Circle $\mapsto$ Ellipse

↑ generalizes to higher dimensions

alphabetic, u, v →

note $U \overset{?}{=} V^T$

but wrong order of operations

E24cp4

how $\vec{y} = \mathbb{A}\vec{x}$ works

$$\mathbb{A} = \mathbb{U}\,\Sigma\,\mathbb{V}^T$$

① Change $\vec{x}$'s representation from standard to $\{\hat{v}\}$ basis ($\mathbb{V}^T$)

② Does the work of $\mathbb{A}$
Stretch/Shrink by σ_i factors
m r dimensions
↗
$C(\mathbb{A}^T)$ & $C(\mathbb{A})$

③ Change from $\{\hat{u}\}$ to standard basis ($\mathbb{U}$)

Fundamental Theorem of Matrixology

E25ap1

From E13bp3:

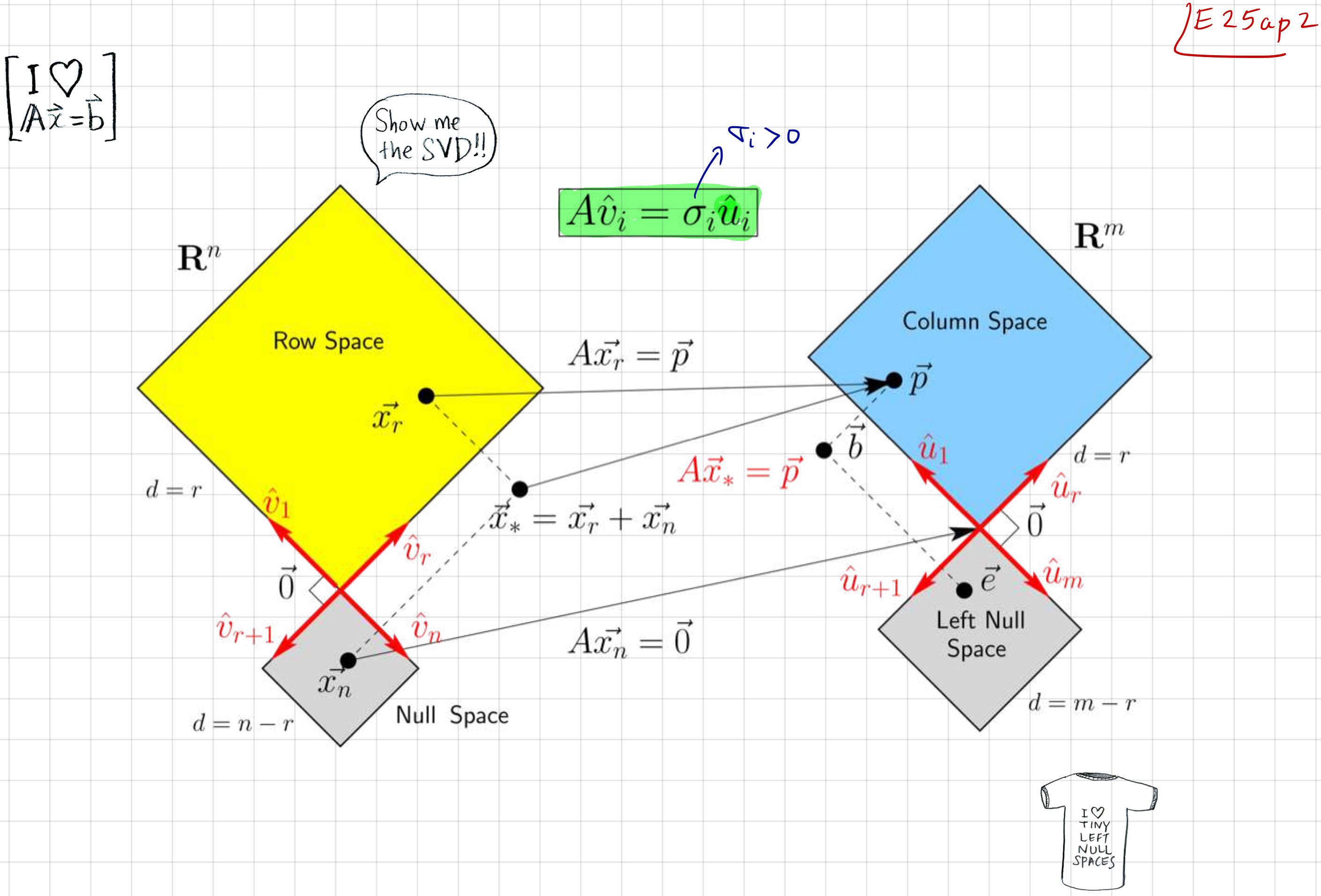
- $\dim C(A) = r$ ← rank — column space
- $\dim N(A^T) = m - r$ — left null space
- $\dim C(A^T) = r$ — row space
- $\dim N(A) = n - r$ — nullspace
- $C(A)$ and $N(A^T)$ are orthogonal complements in R^m ($C(A) \oplus N(A^T)$)
- $C(A^T)$ and $N(A)$ are orthogonal complements in R^n ($C(A^T) \oplus N(A)$)
- The bases of $C(A)$ & $N(A^T)$ combine to give a basis of R^m
- The bases of $C(A^T)$ & $N(A)$ combine to give a basis of R^n

Now we also have:

- Row space has a "natural" orthonormal basis $\{\hat{v}_1, \dots, \hat{v}_r\}$, eigenvectors of A^TA
- Nullspace has a "natural" orthonormal basis $\{\hat{v}_{r+1}, \dots, \hat{v}_n\}$, eigenvectors of A^TA
- Column Space has a "natural" orthonormal basis $\{\hat{u}_1, \dots, \hat{u}_r\}$, eigenvectors of AA^T
- Left Nullspace has a "natural" orthonormal basis $\{\hat{u}_{r+1}, \dots, \hat{u}_m\}$, eigenvectors of AA^T
- The transformation between the "best" bases for row space and column space is diagonal with positive entries:

$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \ddots 0 \end{bmatrix}$$

with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$



E25ap2
I ♥ $A\vec{x}=\vec{b}$
Show me the SVD!!
$A\hat{v}_i = \sigma_i \hat{u}_i$
$\sigma_i > 0$
$\mathbf{R}^n$
$\mathbf{R}^m$
Row Space
Column Space
$A\vec{x}_r = \vec{p}$
$\vec{x}_r$
$\vec{p}$
$\vec{b}$
$A\vec{x}_* = \vec{p}$
$\vec{x}_* = \vec{x}_r + \vec{x}_n$
$d = r$
$d = r$
$\hat{v}_1$
$\hat{v}_r$
$\hat{u}_1$
$\hat{u}_r$
$\vec{0}$
$\vec{0}$
$\hat{v}_{r+1}$
$\hat{v}_n$
$\hat{u}_{r+1}$
$\hat{u}_m$
$\vec{e}$
Left Null Space
$A\vec{x}_n = \vec{0}$
$\vec{x}_n$
Null Space
$d = n - r$
$d = m - r$
I ♥ TINY LEFT NULL SPACES

Time for a nap:

E25ap3