$Ax = b$ revisited:

- over/under-determined cases
- range/nullspaces

Why do we care?

- fitting a fct (HW05)
- linear models: $\underbrace{y}_{\text{sensor data}} = \underbrace{C}_{\text{model}} \cdot \underbrace{x}_{\text{robot state}}$

Problem: Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, we seek solution(s) $x \in \mathbb{R}^n$ s.t. $A \cdot x = b$

More generally, given $(\mathcal{X}, \mathcal{F}), (\mathcal{Y}, \mathcal{F})$, $x \in \mathcal{X}$,

linear operator $\mathcal{L}(x) : \underbrace{\mathcal{X}}_{\text{"domain"}} \rightarrow \underbrace{\mathcal{Y}}_{\text{"codomain"}}$

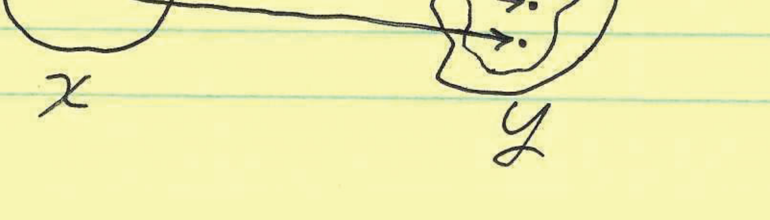
Looking at $\mathcal{L}(x) = A \cdot x$, $\mathcal{L} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where $\mathcal{L}(x) = b \in \mathbb{R}^m$

☆ There might be one sol'n, many sol'ns, or no sol'ns! How do we find them?

But first, more linear algebra...

<u>Def</u>: The <u>Image</u> of $\mathcal{L}(x) : \mathcal{X} \rightarrow \mathcal{Y}$ is

$$\mathrm{im}(\mathcal{L}) := \{ y \in \mathcal{Y} \mid y = \mathcal{L}(x),\ x \in \mathcal{X} \}$$



<u>Def</u>: The <u>kernel</u> of $\mathcal{L} : \mathcal{X} \to \mathcal{Y}$ is

$$\ker(\mathcal{L}) := \{ x \in \mathcal{X} \mid \mathcal{L}(x) = 0 \}$$

☆ Fact: Image & kernel are subspaces (exercise)

<u>For matrices</u>... $A \in \mathbb{R}^{m \times n}$: - linear operator $\mathcal{L}(x) = A \cdot x$
- matrix representation of $\mathcal{L}$ (depends on bases)

<u>Def</u>: The range space of A is:

$$\mathcal{R}(A) := \{ y \in \mathbb{R}^m \mid y = A \cdot x, \ x \in \mathbb{R}^n \}$$

or if we write $A = [A_1 | A_2 | \cdots | A_n]$

$$\mathcal{R}(A) = \text{span}\{A_1, \dots, A_n\}$$

We call this the "image of A" or the "column space of A". $\mathcal{R}(A^T)$, the "row space".

Note: $\text{rank}(\mathcal{L}) = \dim(\overbrace{\text{im}(\mathcal{L})}^{\text{subspace}})$

$\text{rank}(A) = \dim(\underbrace{\mathcal{R}(A)}_{\text{subspace}}) = \dim(\mathcal{R}(A^T))$

<u>Def</u>: The null space of A is

$$\mathcal{N}(A) := \{ x \in \mathbb{R}^n \mid Ax = 0 \}$$

also called the "kernel of A".

Def: The <u>nullity</u> is the dim. of the kernel:

$$\text{nullity}(\mathcal{L}) = \dim(\ker(\mathcal{L}))$$

$$\text{nullity}(A) = \dim(\mathcal{N}(A))$$

☆ Fact: range space & nullspace are subspaces of $\mathcal{X}$ or $\mathcal{Y}$.

Thm: (rank-nullity) $\dim(\mathcal{X}) = \text{rank}(\mathcal{L}) + \text{nullity}(\mathcal{L})$
(proof later)

For rest of lecture, stick to <u>matrices</u> ($\mathcal{R} \leftrightarrow \text{im}$, $\mathcal{N} \leftrightarrow \ker$ for general case)

Thm: 1) $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$ and 2) $\mathcal{N}(A)^\perp = \mathcal{R}(A^T)$

Proof (sketch): ② For $x \in \mathcal{N}(A)$, $A \cdot x = 0 \Leftrightarrow$ rows of $A \perp x$ $(\langle A_i^T, x \rangle = 0)$

$\therefore$ All $x \in \mathcal{N}(A)$ are orthog. to all $\underbrace{y}_{\text{linear comb. of rows of } A} \in \mathcal{R}(A^T)$

~~B.c. N(A) and R(A^T) are subspaces & orthog.~~

$\mathcal{N}(A) = \{A_1^T, \dots, A_m^T\}^\perp = \text{span}\{A_1^T, \dots, A_m^T\}^\perp = \mathcal{R}(A^T)^\perp$ □

① Follows similarly

Thm: 1) $\mathcal{R}(A) \oplus \mathcal{N}(A^T) = \mathbb{R}^m$ (codomain)

2) $\mathcal{R}(A^T) \oplus \mathcal{N}(A) = \mathbb{R}^n$ (domain)

Proof ①: $\mathbb{R}^m = \overbrace{\mathcal{R}(A)}^{\text{subspace}} \oplus \mathcal{R}(A)^\perp$ from last lecture

and $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$ from prior theorem.

$\therefore \mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$ ✓ □

② Follows similarly

Note: For a square matrix $A \in \mathbb{R}^{n \times n}$, nullspace gives us a new tool to check if A^{-1} exists!

If $\mathcal{N}(A) = \{0\} \Rightarrow \text{nullity}(A) = 0$

$\therefore \text{rank}(A) = n \Rightarrow A$ full rank $\Rightarrow A^{-1}$ exists

TFAE for $A^{n \times n}$:

1. $\mathcal{N}(A) = \{0\}$
2. A is full rank
3. $\det(A) \neq 0$ $\qquad \det(A) = \prod_i \lambda_i$
4. A^{-1} exists

Have we seen this before?

Back to eigenvalues/vectors: $A \cdot v = \lambda \cdot v$

$$\Leftrightarrow (A - \lambda \cdot I) \cdot v = 0$$

Because $v \neq 0$, $v \in \mathcal{N}(A - \lambda \cdot I)$ ← also called "eigenspace"

$$\Rightarrow \det(A - \lambda I) = 0 \text{ from TFAE.}$$

characteristic eqn!

Back to $A \cdot x = b$:

Given $A \in \mathbb{R}^{m \times n}$, A full rank $(\text{rank}(A) = \min(n, m))$, $b \in \mathbb{R}^m$, we seek $x \in \mathbb{R}^n$ s.t. $A \cdot x = b$

case 1: $m = n$. Then $\mathcal{R}(A) = \mathbb{R}^n = \mathcal{R}(A^T)$, $b \in \mathcal{R}(A)$ and $x \in \mathcal{R}(A) = \mathbb{R}^n$

one solution $\Rightarrow x = A^{-1} \cdot b$, this is the "critical" case.

$b \in \mathbb{R}^m,\ x \in \mathbb{R}^n,\ A \in \mathbb{R}^{m \times n}$ "wide"

<u>Case 3</u>: $n > m$ (underdetermined) $A = [\quad\quad]$

↳ less equations than unknowns

↳ many solutions!

Recall $x \in \mathbb{R}^n$ and $\mathbb{R}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$

Idea: decompose $\hat{x}$ into components in $\mathcal{R}(A^T)$ & $\mathcal{N}(A)$

$$\hat{x} = \hat{x}_{\mathcal{R}(A^T)} + \hat{x}_{\mathcal{N}(A)}, \quad \text{where } \hat{x}_{\mathcal{R}(A^T)} \in \mathcal{R}(A^T)$$

$$\hat{x}_{\mathcal{N}(A)} \in \mathcal{N}(A)$$

$$A(\hat{x}_{\mathcal{R}(A^T)} + \hat{x}_{\mathcal{N}(A)}) = b$$

$$A\hat{x}_{\mathcal{R}(A^T)} + \cancelto{0}{A \cdot \hat{x}_{\mathcal{N}(A)}} = b$$

Choose $\hat{x} = \hat{x}_{\mathcal{R}(A^T)} \in \mathcal{R}(A^T) \implies \hat{x} = A^T \cdot \alpha$ (linear comb. of rows of $\mathbf{A}$)

Then, $A\hat{x} = A \cdot A^T \alpha = b$

$$\implies \alpha = \underbrace{(AA^T)^{-1}}_{\text{exists b.c. } A \text{ full rank}} b$$

$$\therefore \boxed{\hat{x} = A^T(AA^T)^{-1}b}$$

This is the minimum norm solution!

b.c.

$$\|\hat{x}\|^2 = \|\hat{x}_{\mathcal{R}(A^T)} + \cancelto{0 \text{ by choice}}{\hat{x}_{\mathcal{N}(A)}}\|^2 = \|\hat{x}_{\mathcal{R}}\|^2 + \cancelto{0}{\|\hat{x}_{\mathcal{N}}\|^2}$$

~~Also~~

~~by Pythag. thm ($\hat{x}_R \perp \hat{x}_N$)~~