ROB 501 - Oct 10, 2019

Today: $Ax = b$ revisited
- range / nullspaces
- over / under determined & critical solutions

Why do we care about $Ax = b$?
- Fitting a f'n (HW05!)
- Linear model: $y = Cx$

sensor data, model, robot state

(if not linear → linearize!)
- Linear sys. of eq'n (circuits, statics, dynamics)

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. $Ax = b$.

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

$$\mathcal{L}(x) : \mathcal{X} \to \mathcal{Y}$$

"domain" "codomain"

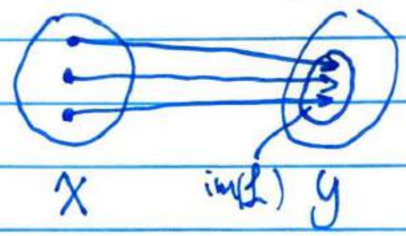
Looking at $\mathcal{L}(x) = Ax$, $\mathcal{L}(x) : \mathbb{R}^n \to \mathbb{R}^m$, when $\mathcal{L}(x) = b \in \mathbb{R}^m$

☆ There might be one sol'n, many or none! How do we find them?

More linear algebra...

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

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

Fact: Image & kernel are subspaces! (proof is exerci

For matrices... $A \in \mathbb{R}^{m \times n}$

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

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

or if we write $A = [A_1 \mid A_2 \mid \dots \mid A_n]$

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

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

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

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

Def The null space of A is:

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

also called the "kernel of A".

Def The nullity is the dim. of the kernel.

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

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

Fact: range space & null space are subspaces.

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

For matrices: (rest of the lecture) ($\mathcal{R} \Leftrightarrow \text{im}$, $N \Leftrightarrow \ker$ general case still exists!)

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

Proof ②: $Ax = 0 \iff$ rows of $A \perp x$ ($\langle A_i, x \rangle = 0$)
$\therefore \forall\ x \in N(A)$ are orthog. to
$\forall\ y \in \mathcal{R}(A^T)$ $\square$

Proof ①: follows similarly

[Recall: $(\mathcal{X}, \mathcal{F})$, $\mathcal{U} \subset \mathcal{X}$, $\mathcal{V} \subset \mathcal{X}$ subspaces of $\mathcal{X}$
$\mathcal{U} \cap \mathcal{V} = \{0\}$, then the direct sum is
$\mathcal{U} \oplus \mathcal{V} = \{u + v \mid u \in \mathcal{U}, v \in \mathcal{V}\}$]

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: $\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$
from before $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$
$\mathbb{R}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$ $\square$

Note: For a square matrix $A \in \mathbb{R}^{n \times n}$, null space gives us a new tool to check if A^{-1} exists! If $\mathcal{N}(A) = \{0\} \Rightarrow$ nullity$(A) = 0$
$\therefore$ rank$(A) = n$ $\therefore$ A is full rank
$\therefore$ A^{-1} exists!

TFAE (for $A^{n \times n}$)
1. $\mathcal{N}(A) = \{0\}$
2. A is full rank
3. $\det(A) \neq 0$
4. A^{-1} exists

Have we seen this before?

Back to eigen: $Av = \lambda v$

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

since $v \neq 0$, $v \in N(A - \lambda I)$ ← also called the "eigenspace" of A, E_λ

$\det(A - \lambda I) = 0$

↑ characteristic equation!

Back to $Ax = b$:

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

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

↳ we have one solution

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

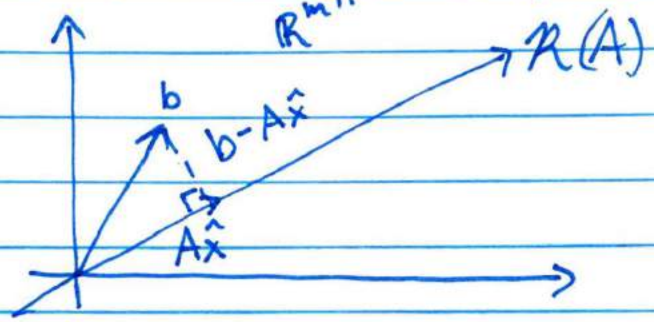
this is the "critical" case.

Case ②: $m > n$ (overdetermined) $\left(A = [\ ]\right)$

We might not have $b \in \mathcal{R}(A)$

↳ no solution!

minimize $\|b - A\hat{x}\|$



$A\hat{x} = \hat{y} \in \mathcal{R}(A)$

$\hat{y} = \underset{y \in \mathcal{R}(A)}{\operatorname{argmin}} \|b - y\|$

⇒ Projection thm!

from proj thm. $b - A\hat{x} \perp A\hat{x}$. Let $\langle x, y \rangle = x^T y$

then:

$$\langle b - A\hat{x}, A\hat{x} \rangle = 0$$

$$\langle b, A\hat{x} \rangle = \langle A\hat{x}, A\hat{x} \rangle$$

$$b^T A \cancel{\hat{x}} = (A\hat{x})^T A\hat{x} = \hat{x}^T A^T A \cancel{\hat{x}}$$

$$(b^T A)^T = (\hat{x}^T A^T A)^T$$

$$A^T b = A^T A \hat{x}$$

$A^T A$ is invertible since A is full rank

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

Note: could also use normal equations!

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

↳ 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 $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$:

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

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

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

Let's pick $\hat{x} = \hat{x}_{\mathcal{R}(A^T)} \in \mathcal{R}(A^T) \rightarrow \hat{x} = A^T\alpha$

$$A\hat{x} = AA^T\alpha = b$$

$$\alpha = (AA^T)^{-1}b$$

exists since A is full rank

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

This is the minimum norm solution!