# Solutions to Ax=b

ROB 501

Necmiye Ozay

- **Orthogonal projection operator**
- **Ax=b, when does it have a unique exact solution?**
  - **Range (image), nullspace (kernel)**

# Orthogonal Projection Operator

$(\mathcal{X}, \mathbb{R}, \langle \cdot , \cdot \rangle)$ finite dim. inner product space and $M \subset \mathcal{X}$ a subspace. For $x \in \mathcal{X}$ and $\hat{x} \in M$,

TFAE:

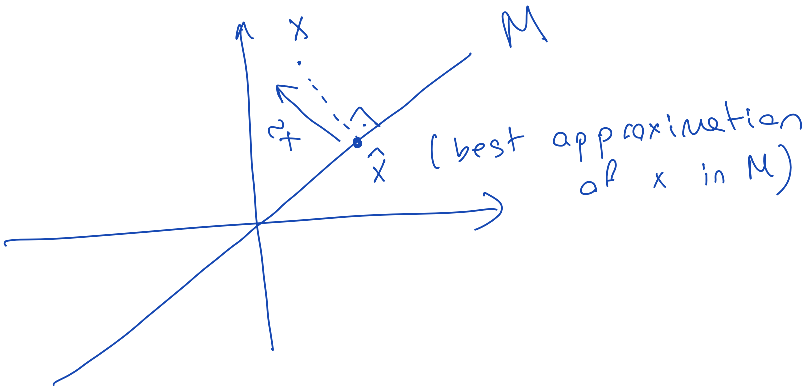
a) $\hat{x} = \underset{m \in M}{\arg\min} \, \| x - m \|$

b) $\underline{x - \hat{x}} \perp M$

c) $\exists \, \underline{\tilde{x}} \in M^{\perp}$ s.t. $x = \underset{\in M}{\hat{x}} + \underset{\in M^{\perp}}{\tilde{x}}$

↳ error vector

__Def:__ $P : \mathcal{X} \longrightarrow M$ by $P(x) = \hat{x}$, where $\hat{x} \in M$ that satisfies (a), (b), or (c), $\hat{x}$ is called the orthogonal projection of $x$ onto $M$.




<u>Exercise</u>: 1) $P : \mathcal{X} \to M$ defined above is a <u>linear operator</u>. (Hint: use (c) in TFAE to prove)

2) Let $\{v^1, \dots, v^k\}$ be an orthonormal basis for M. Then

$$P(x) = \sum_{i=1}^{k} \langle x, v^i \rangle v^i = \hat{x}$$

---

## <u>Ax = b revisited</u>

- over/under determined equations
- range/null space

Why do we care?

- fitting functions exacty (interpolation)
- linear models for sensing: $\underbrace{y}_{\text{sensor data}} = \underbrace{C}_{\text{model of sensor}} \underbrace{x}_{\text{robot state}}$

**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. $Ax = b$.

More generally, given $(\mathcal{X}, \mathbb{F})$, $(\mathcal{Y}, \mathbb{F})$, linear operator $\mathcal{L} : \mathcal{X} \longrightarrow \mathcal{Y}$ and $y \in \mathcal{Y}$. We seek $x \in \mathcal{X}$ s.t. $\mathcal{L}(x) = y$. Special case $\mathcal{L}(x) = Ax$, $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$.

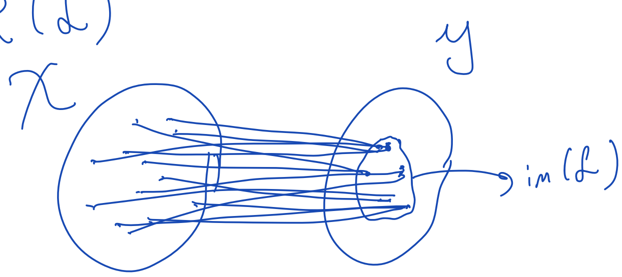
* There might be one sol'n, many sol'ns or no sol'n. How do we find sol'n when they exist?

**First some linear algebra:**

**Def.** The image (range) of $\mathcal{L} : \mathcal{X} \to \mathcal{Y}$ is

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

$R(\mathcal{L})$



**Def.** The kernel (nullspace) of

$\mathcal{L} : \mathcal{X} \to \mathcal{Y}$ is

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

$\mathbf{0}$ of the space $\mathcal{Y}$

equivalent notation $\mathcal{N}(\mathcal{L})$

**Fact:** Image and kernel are subspaces.
(proof: left as exercise).

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

**Def.** The range space of $A$ is

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

the image of linear operator
$\mathcal{L}(x) = Ax$

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

$$\mathcal{R}(A) = \operatorname{span}\{A_1, A_2, \dots, A_n\}$$

We call it "range space of $A$" = "image of $A$" = "column space of $A$"

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

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

**Def.** The nullspace of $A$ is

$$\underbrace{\mathcal{N}(A)}_{\text{kernel of } \mathcal{L}(x) = Ax} = \{ x \in \mathbb{R}^n \mid Ax = \mathbf{0} \}$$

aka "kernel of A"

**Note:** The <u>nullity</u> is the dimension of kernel

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

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

**Thm:** (rank-nullity theorem) $\mathcal{L} : \mathcal{X} \to \mathcal{Y}$

$$\dim(\mathcal{X}) = \text{rank}(\mathcal{L}) + \text{nullity}(\mathcal{L})$$

For the rest of the lecture, we will focus on matrices, and show some properties of image and kernel.

Thm 1: $A \in \mathbb{R}^{m \times n}$

1) $\mathcal{R}(A)^{\perp} = \mathcal{N}(A^T)$

$\subseteq \mathbb{R}^m \qquad \subseteq \mathbb{R}^m$

2) $\mathcal{N}(A)^{\perp} = \mathcal{R}(A^T)$

$\subseteq \mathbb{R}^n \qquad \subseteq \mathbb{R}^n$

Proof: 2) For $x \in \mathcal{N}(A) \leftrightarrow Ax = 0$

$\Leftrightarrow$ rows of $A \perp x$

$\Leftrightarrow$ columns of $A^T \perp x$

$\Leftrightarrow$ All $x \in \mathcal{N}(A)$ are orthogonal $y \in \mathcal{R}(A^T)$

$\Leftrightarrow \mathcal{N}(A)^{\perp} = \mathcal{R}(A^T)$

1) proof is similar

Thm 2: 1) $\mathcal{R}(A) \oplus \mathcal{N}(A^T) = \mathbb{R}^m$

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

**Proof:** follows from Thm 1 and that for subspace $M \subset \mathbb{R}^k$, $M \oplus M^{\perp} = \mathbb{R}^k$ and $\mathcal{R}(A)$ and $\mathcal{N}(A)$ are subspaces.

**Note:** For square matrices $A \in \mathbb{R}^{n \times n}$, nullspace gives us another tool to check the invertibility of $A$ (i.e. existence of $A^{-1}$).

$$\mathcal{N}(A) = \{0\} \iff A^{-1} \text{ exists}$$
$$\Updownarrow$$
$$\text{nullity}(A) = 0 \iff \text{rank}(A) = n \iff A \text{ is full rank}$$

TFAE:

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

Note that we already used this when working with eigenvectors.

$v \neq 0 \quad Av = \lambda v \iff (A - \lambda I)v = 0 \quad v \neq 0 \iff$

$v \neq 0 \quad v \in \mathcal{N}(A - \lambda I)$

because 1) $\iff$ 3)

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

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

______ a ______

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. $Ax = b$

- for the existence of an exact sol'n, we need

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

case 1: $m = n$ + A full rank

Then, $\mathcal{R}(A) = \mathbb{R}^n$ and any $b \in \mathcal{R}(A)$

$\Rightarrow$ one solution $x = A^{-1}b$ "$m=n$" is called the critical case.

<u>case 2:</u> $m > n$ (more equations than unknowns)

+ A full rank

option 1) $b \in R(A) \longrightarrow \exists$ unique solution $x = (A^TA)^{-1}A^Tb$

option 2) $b \notin R(A) \longrightarrow$ no exact sol'n but we can solve for an approximate $\hat{x}$ using least squares

<u>case 3:</u> $n > m$ "underdetermined case"

+ A is full rank $\longrightarrow$ many solutions.

# OFFICE HOURS

$$\underbrace{x^T Q x}_{\text{pos. def. quadratic form}} \quad \text{and} \quad Q > 0$$

$\Longrightarrow$ $x^T Q x$ is a convex function

$$\langle x, y \rangle = x^T Q y$$

$$\hookrightarrow \|x\|_Q = \sqrt{x^T Q x}$$

$$\|Cx - y\|_Q^2 \qquad y = Cx$$

take $M \in \mathbb{R}^{n \times n}$

$$M = M_{sym} + M_{anti}$$

$$M_{sym} = \frac{M + M^T}{2} \qquad M_{anti} = \frac{M - M^T}{2}$$

$$\forall x \qquad x^T \left( \frac{M - M^T}{2} \right) x = 0$$

$$\Downarrow \quad \cancel{\frac{x^T M x}{2}} - \cancel{\frac{x^T M^T x}{2}}$$

$$x^T M x = x^T M_{sym} x$$

0