Introduction

~ Sam Fairchild

~ samantha.fairchild@mis.mpg.de
↳ email me to be put on mailing list

~ office hours → wednesday after class
↳ by email appt.

Grading
- Homework – due every other week
- Project → due 18 January
- Exam → Feb 1?

~> Homework via email? ....
↳ Homework sheet + Julia nb – 01

Course outline

↳ Linear algebra + Prob. Theory

- Network Analysis
- Machine Learning
- TDA
- Matrices + Tensors

Today Orthogonality + Pseudoinverse.

1.1 Linear Algebra

$$A = (a_{ij}) \in \mathbb{R}^{m \times n}$$

$1 \le i \le m$

$1 \le j \le n$

Column vector $a_j := (a_{ij})_{i=1}^{m}$

$$A = [a_1, \dots, a_n]$$

For $x \in \mathbb{R}^n$ $\quad Ax = x_1 a_1 + \dots + x_n a_n$

Matrices are useful ways to organize data

~ List of vectors in $\mathbb{R}^n$ (resp in $\mathbb{R}^m$)

~ linear maps $x \mapsto Ax$
$\mathbb{R}^n \to \mathbb{R}^m$ $y \mapsto A^T y$
$\mathbb{R}^m \to \mathbb{R}^n$

~ bilinear maps $(x, y) \mapsto y^T A x$.
$\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

$\boxed{\text{Geometry of Matrices}}$

$\boxed{\text{Def}}$ Let $A \in \mathbb{R}^{m \times n}$.

The <u>image</u> of $A = \operatorname{Im}(A) = \{Ax \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$

$\operatorname{Im}(A^T) = \{A^T y \mid y \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$

The <u>kernel</u> of $A = \ker(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} \subseteq \mathbb{R}^n$

Euclidean inner product: Given $x, y \in \mathbb{R}^n$

$$\langle x, y \rangle := x^T y \quad \text{(see bilinear map)}$$

$\boxed{\text{Def}}$ Let U, V be subspaces. Then U is <u>perpendicular</u> (<u>orthogonal</u>) to V if

$$\forall\, u \in U,\ v \in V,\quad \langle u, v\rangle = 0.$$

We write $U \perp V$.

The <u>orthogonal complement</u> of U is

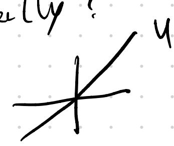
$$U^{\perp} = \{ v \in \mathbb{R}^n \mid \langle u, v \rangle = 0 \ \forall u \in U \}.$$

$\boxed{\text{Exercise}}$ <u>$\mathbb{R}^2$ - reality check:</u>

In pairs

- Write linear map $\mathbb{R}^2 \to \mathbb{R}^2$ $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
- Write euclidean inner product as bilinear map
- Compute $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$, $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$, $\left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$.

  ~ what does this mean geometrically?
- Compute $U^{\perp}$ where $U = \mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

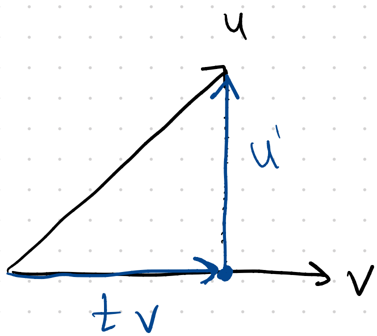


$\boxed{\text{Lemma}}$ Let $U, V \subseteq \mathbb{R}^n$ be two subspaces,

Then $U = V^\perp \iff U \perp V$ and $U \oplus V = \mathbb{R}^n$

$\boxed{\text{Pf}}$ Exercise

Geometric meaning of inner product



$$u = tv + u' \quad \text{where } u' \perp v.$$

Then $\langle u, v \rangle = u^T v = \left(t v^T + (u')^T\right) v$

$$= t v^T v + \underbrace{(u')^T v}_{0 \text{ since } u' \perp v}$$

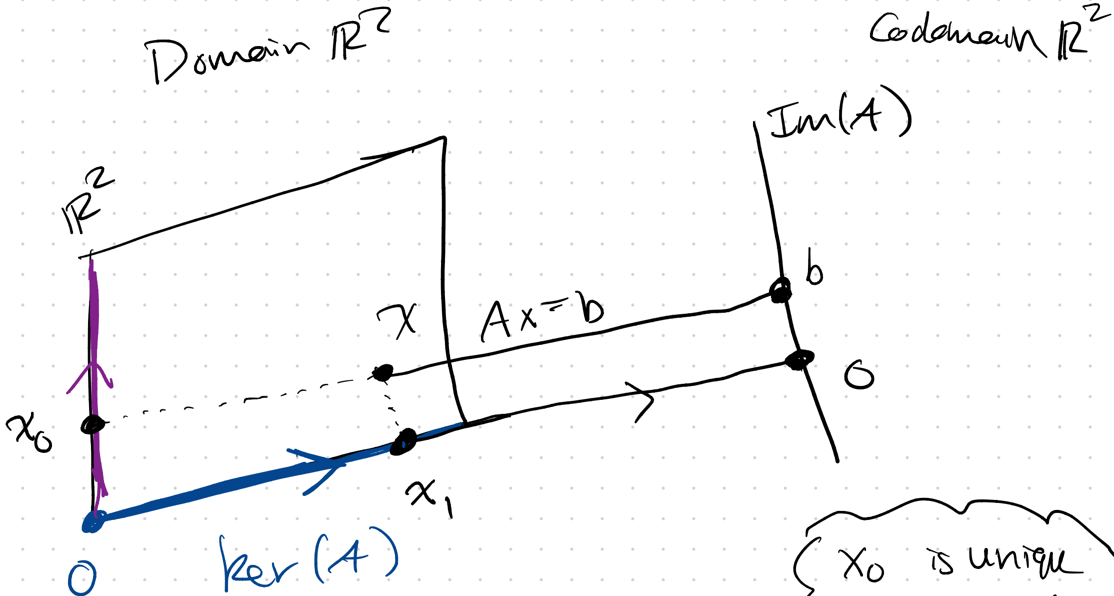
$$= t \langle v, v \rangle$$

Given $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, what are the solutions $x \in \mathbb{R}^n$?

$b \in \mathrm{Im}(A)$

Let $r(A) = \dim(\mathrm{Im}(A))$ be the rank (A).

1. Case 1: $Ax = b$ has a solution $\Leftrightarrow b \in \mathrm{Im}(A)$



x_0 is unique solution upto kernel

Given $x_1 \in \ker(A)$, $Ax_1 = 0$

$x = x_0 + x_1$ $\qquad Ax = Ax_0 + Ax_1 = Ax_0 = b$

Purple line is $\mathrm{Im}(A^T)$, A linear isomorphism on $\mathrm{Im}(A^T)$.

$\boxed{\text{Thm } 1.4}$ Let $A \in \mathbb{R}^{m \times n}$. Then

1. $\mathrm{Im}(A) = \ker(A^T)^\perp$

2. $\mathrm{Im}(A^T) = \ker(A)^\perp$.

$\boxed{\text{Pf}}$ (1) in notes

(2) Recall Rank-Nullity Theorem:

$$r(A^T) = \dim(\mathrm{Im}(A^T))$$

$$r(A^T) + \dim(\ker(A^T)) = n$$

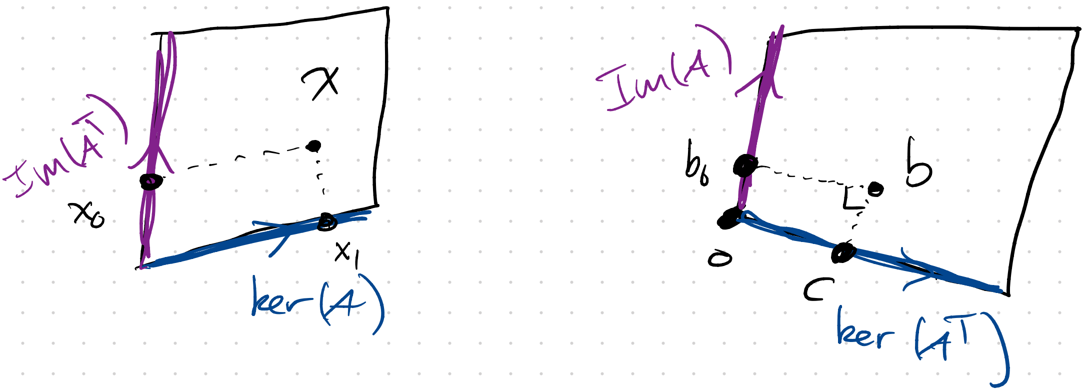
Let $x \in \ker(A)$ and $A^T y \in \mathrm{Im}(A^T)$.

$$\langle x, A^T y \rangle = x^T A^T y = (Ax)^T y = 0$$

$\Rightarrow \mathrm{Im}(A^T) \perp \ker(A)$ and $\mathrm{Im}(A^T) \oplus \ker(A) = \mathbb{R}^n$

$\Rightarrow \mathrm{Im}(A^T) = \ker(A)^\perp$ $\blacksquare$

② **Case 2** What happens when b is not in image?



Find $b_0 \in \mathrm{Im}(A)$ closest to b, so minimize

$$\|b-b_0\| = \sqrt{\langle b-b_0, b-b_0\rangle}.$$

Notation

$$b_0 = \underbrace{\underset{y\in \mathrm{Im}(A)}{\operatorname{argmin}}}\ \|b-y\|$$

argument = value in $\mathrm{Im}(A)$ minimizing quantity $\|b-y\|$.

Claim b_0 is uniquely determined!

$\boxed{\text{Lemma 1.5}}$ Let $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$.

$b_0 = \underset{y \in \mathrm{Im}(A)}{\operatorname{argmin}} \|b - y\|$ is determined by

1. Unique decomposition of Thm 1.4. that

$$b = b_0 + c \quad \text{for} \quad c \in \ker(A^T).$$

2. $A^T b = A^T b_0$

$\boxed{\text{Pf}}$ Uniqueness follows from uniqueness of Thm 1.4. and $c \in \ker(A^T)$

$$\Rightarrow \quad A^T b = A^T b_0 + A^T c = A^T b_0.$$

Now want to show $A^T b = A^T b_0$ gives b_0 as argmin.

$A = [a_1 | \cdots | a_n]$ $\quad b_0 \in \mathrm{Im}(A)$, $b_0 = A x_0$.

Define $\quad \phi(x) = Ax - b = \begin{bmatrix} \phi_1(x) \\ \vdots \\ \phi_m(x) \end{bmatrix}$

goal: minimize $\|\phi(x)\|$

$$0 = \text{Gradient} = \left[\frac{\partial}{\partial x_1}\|\phi(x)\|, \dots, \frac{\partial}{\partial x_n}\|\phi(x)\|\right] \in \mathbb{R}^n$$

$$\frac{\partial}{\partial x_i}\|\phi(x)\| = \frac{\partial}{\partial x_i}\sqrt{\sum_{j=1}^{m}\phi_j^2(x)}$$

$f(x)^{\frac{1}{2}}$

$\frac{1}{2} f(x)^{-\frac{1}{2}} f'(x)$

$$= \frac{1}{2\|\phi(x)\|} \cdot \sum_{j=1}^{m} 2\phi_j(x)\frac{\partial \phi_j}{\partial x_i}(x)$$

HW exercise

$$\boxed{=} \frac{1}{\|\phi(x)\|} a_i^T(Ax-b)$$

$\Rightarrow 0 = A^T(Ax-b)/\|\phi(x)\|$

So either $\|\phi(x_0)\| = 0 \Rightarrow b = b_0$ and done, or

$$0 = A^TAx_0 - A^Tb \Leftrightarrow A^Tb = A^TAx_0 = A^Tb_0.$$

---

① $A\big|_{Im(A^T)}$ is linear isomorphism.

② $\pi_{Im(A)}$ projects b onto $Im(A)$.

$\Rightarrow$ we have a well defined map $\mathbb{R}^m \to \mathbb{R}^n$ by $\left(A\big|_{Im(A^T)}\right)^{-1} \circ \pi_{Im(A)}$ called the pseudo inverse.

$\boxed{\text{Definition}}$ The <u>Pseudo inverse</u> is

$A^+ \in \mathbb{R}^{n \times m}$ is the matrix s.t.

$$A^+ b = x \quad \text{for} \begin{cases} x \in Im(A^T) \\ Ax = b_0 \\ b_0 = \underset{y \in Im(A)}{\operatorname{argmin}} \|b - y\| \end{cases}$$

$\boxed{\text{Properties}}$ $A \in \mathbb{R}^{n \times m}$, $A^+ \in \mathbb{R}^{m \times n}$ its pseudoinverse.

1. If A invertible, $A^{-1} = A^+$

2. AA^+ is the orthogonal projection onto

$\mathrm{Im}(A)$. i.e. $AA^+ = \Pi_{\mathrm{Im}(A)}$.

3. When A has full rank ($r(A) = \min\{m,n\}$)

Then supposing $r(A) = n$,

$$A^+ = (A^T A)^{-1} A^T$$

and $A^+ A = I_n$.

$\boxed{\text{Pf } 3.}$ Let $b \in \mathbb{R}^m$ and $A^+ b = x$. By Lemma 1.5

$$A^T A x = A^T b$$

$$\Rightarrow A^T A A^+ b = A^T b$$

Since $r(A) = n$, $A^T A \in \mathbb{R}^{n \times n}$ is invertible

$$\Rightarrow A^+ b = (A^T A)^{-1} A^T b.$$

$\boxed{\text{example}}$ Compute A^+ of

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

$$A^T A = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \; 1$$

$$A^+ b = 1 \cdot \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$