

One more short HW this week

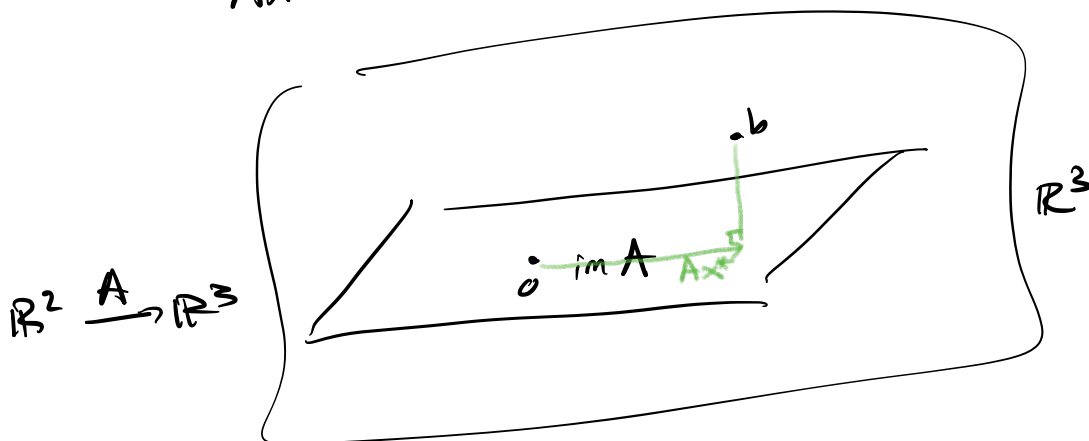
Exam ^{monday}, review ^{wednesday},
next week this week.

least squares approximation

Suppose we're trying to solve a system
 $Ax = b$ which is inconsistent.

We say that x^* is a least square solution if

Ax^* is the closest vector to b in the range of A .



Fact: x^* is a least $\square$ soln if and only if $(b - Ax^*)$
is perp to $\text{im } A$.

$$v \cdot w = v^T w$$

Lemma: $(\text{im } A)^\perp = \ker A^T$

Pf: $(\text{im } A)^\perp$ are the set of v 's such that $v \cdot Ax = 0$ all x

$$v \cdot Ax = 0 \Leftrightarrow v^T A x = 0 \quad \text{all } x \Leftrightarrow v^T A = 0$$

$$(v^T A)^T = 0$$

$$\Downarrow$$

$$A^T v = 0 \quad \text{i.e. } v \in \ker A^T$$

Using this, we see:

x^* is a least $\square$ soln to $Ax = b$

$$(\text{is the same as}) \Leftrightarrow (b - Ax^*) \in (\text{im } A)^\perp = \ker A^T$$

$$\text{i.e. } A^T(b - Ax^*) = 0 \quad \text{or} \quad A^T b = A^T A x^*$$

"normal equations"

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

x^* is a soln to $\nearrow$

Fact this is always consistent and will have a unique solution whenever A is 1-1 (for $A=0$)

ex: $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad b = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}$

$$A: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 3 & 6 \end{array} \right] \rightarrow X$$

$$Ax = b \text{ inconsistent}$$

$$\downarrow$$

normal equations
 $A^T A x = A^T b$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad A^T b$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix} = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} x = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 6 & 6 \\ 6 & 14 & 18 \end{array} \right]$$

$$\left\{ \begin{array}{cc|c} 1 & 2 & 2 \\ 3 & 7 & 9 \end{array} \right\}$$

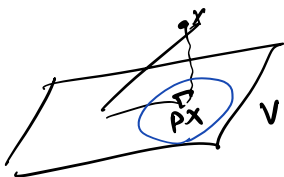
$$\left\{ \begin{array}{cc|c} 1 & 2 & 2 \\ 0 & 1 & 3 \end{array} \right\}$$

$$\left\{ \begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & 3 \end{array} \right\}$$

$$(-4, 3)$$

Suppose $V \subseteq \mathbb{R}^n$ is a subspace w/ basis $v_1, \dots, v_m$

want to describe the map $\mathbb{R}^n \xrightarrow{P} V$
 $x \mapsto$ closest vector in V to x .



Sketch: $A = [v_1 | \dots | v_m]$ then $\boxed{\text{im } A = V}$
 $\{Ay \mid y \in \mathbb{R}^m\}$

Start w/ $x \in \mathbb{R}^n$
want closest to x in $V = \text{im } A$ $A \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = y_1 v_1 + \dots + y_m v_m$

want to solve $p x = \text{closest to } x \text{ in } \text{im } A$

$p x$ is a soln to " $A y = x$ " inconsistent

$$A y = p x \iff A^T A y = A^T x$$

y as here satisfies $A y$ close as possible to x

$$\{ \\ p x = A y = A (A^T A)^{-1} A^T x$$

Projection from $\mathbb{R}^n$ to V , V has basis $v_1, \dots, v_m$
then we set $A = [v_1 \dots v_m]$

and we ^{can} define the projection $p: \mathbb{R}^n \rightarrow V \subseteq \mathbb{R}^n$

$$\text{formula is } p x = \underbrace{A (A^T A)^{-1} A^T}_\text{matrix for the projection} x$$

Sidebar: why is $A^T A$ invertible?

because A is 1-1. From last time, enough to check
ker $A = \{0\}$

$$\begin{bmatrix} v_1 & \dots & v_m \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = 0 \Leftrightarrow y_1 v_1 + \dots + y_m v_m = 0$$

but v 's are independent.
 so this means $y_1 = 0 = y_2 = \dots = y_m$
 $\Rightarrow \ker A = 0$

Remark: $v_1, \dots, v_m$ are orthonormal

$$\Leftrightarrow$$

$$v_i \cdot v_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

$$\Leftrightarrow$$

$$\begin{bmatrix} \hline v_1 \\ \vdots \\ \hline v_m \end{bmatrix} \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix} = \text{Id}_m$$

$$\Leftrightarrow$$

if $A = \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix}$ then ~~$A^T = A^{-1}$~~ $ATA = \text{Id}_m$
 $\downarrow$
 orthonormal

Projection from $\mathbb{R}^n$ to V , V has basis $v_1, \dots, v_m$

then we set $A = \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix}$ orthonormal

and we ^{can} define the projection $p: \mathbb{R}^n \rightarrow V \subseteq \mathbb{R}^n$

Formula is $p x = A (A^T A)^{-1} A^T x$
 matrix for the projector.

~~$A^T = A^{-1}$~~ $\Rightarrow$ $A^T A = I_{\mathbb{R}^m}$ $p x = A A^T x$

Abstract vector space perspective

Suppose U, W are vector spaces

Recall: A linear transformation $T: U \rightarrow W$ is a function such that

- $T(x+y) = T(x) + T(y)$
- $T(\lambda x) = \lambda T(x)$

We say T is an isomorphism if there is a lin trans. $T^{-1}: W \rightarrow U$
 s.t. $T T^{-1} = \text{id}_W$ $T^{-1} T = \text{id}_U$

Prop: if T is bijective lin trans. then the inverse function T^{-1}
 is also a lin transformation.

Proof: suppose T^{-1} is the inverse function. so $T T^{-1} = \text{id}_W$
 $T^{-1} T = \text{id}_U$

want to show $T^{-1}(x+y) = T^{-1}(x) + T^{-1}(y)$

$$\& T^{-1}(\lambda x) = \lambda T^{-1}(x)$$

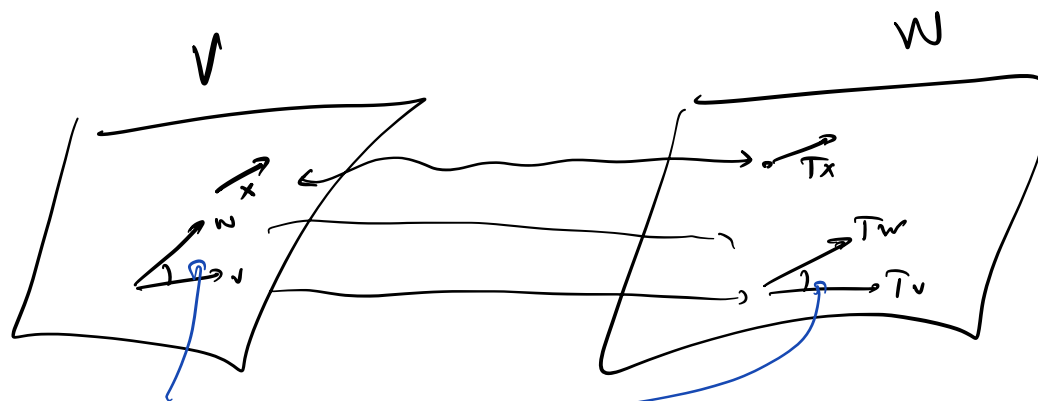
$$\begin{aligned}
 T^{-1}(\lambda + \gamma) &= T^{-1}(T(T^{-1}(\lambda)) + T(T^{-1}(\gamma))) \\
 &= T^{-1}(T(T^{-1}(\lambda) + T^{-1}(\gamma))) \\
 &= (T^{-1} \circ T)(T^{-1}(\lambda) + T^{-1}(\gamma)) \\
 &= \text{id}_V(T^{-1}(\lambda) + T^{-1}(\gamma)) = T^{-1}(\lambda) + T^{-1}(\gamma)
 \end{aligned}$$

$$\begin{aligned}
 T^{-1}(\lambda x) &= T^{-1}(\lambda T(T^{-1}(x))) \\
 &= T^{-1}(T(\lambda T^{-1}(x))) = \underbrace{(T^{-1} \circ T)}_{\text{id}}(\lambda T^{-1}(x)) = \lambda T^{-1}(x).
 \end{aligned}$$

If V & W are inner product spaces w/ inner products $\langle, \rangle_V$
 $\langle, \rangle_W$

an $T: V \rightarrow W$ is an isomorphism,

we say that T is an isometry if $\langle v, w \rangle = \langle T(v), T(w) \rangle$



isometry means same $\angle$

$$v \cdot w = |v||w| \cos \theta$$

in general $\langle v, w \rangle = \sqrt{\langle v, v \rangle} \sqrt{\langle w, w \rangle} \cos \theta$

ex: $\langle 1, x, x^2 \rangle$

inner product

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)g(x)dx$$

$$\|f(x)\|^2 = \int_0^1 f(x)^2 dx$$