Change of Plans:

HW5 will be due before Mondy's class, posted before tomorrow's class.

Quiz 5 will be at beginning of class on Tues, June 23.

Wed, June 24: Review

Thurs, June 25: NO CLASS

Friday, June 26: Final Project Due.

[Final Project: Write a summary of what you learned in this course.
Min 2 pages, Max 10 pages.]

Topic: Least Squares Approximation

Invented by Gauss in order to estimate the orbital parameters of asteroid Ceres.

Suppose linear system

$$A\vec{x} = \vec{b}$$

has no exact solution. Example: Several measurements of asteroid's position do not exactly fit an elliptical orbit because the measurements have errors. In this case we want to find an approximate solution $\vec{x}$ such that

$$\|A\vec{x} - \vec{b}\| \text{ is minimized.}$$

"Least Squares Approximation"

Gauss' solution in modern language says that the best approximate solution of $A\vec{x} = \vec{b}$ is an exact solution of

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

"The Normal Equation(s)"

Furthermore, if A has independent columns then the square matrix $A^T A$ is invertible and we can write

$$\vec{x} = (A^T A)^{-1} A^T \vec{b} .$$

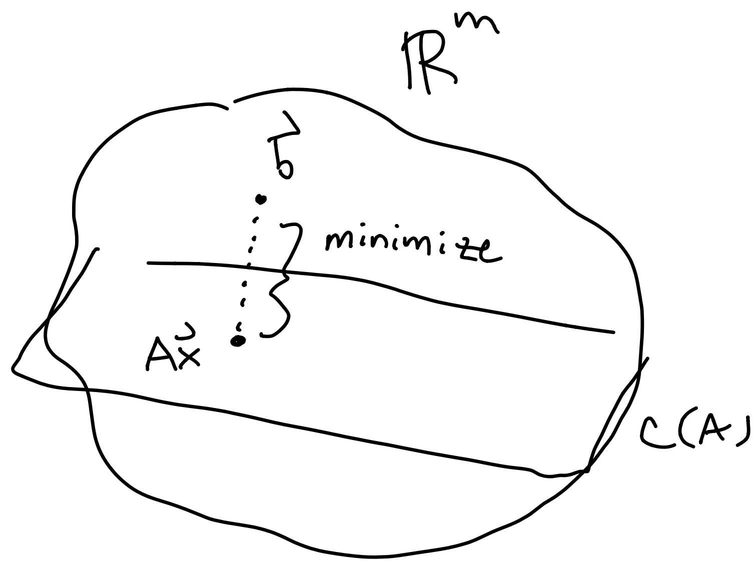
Today: Why does this work?

Recall that the column space of A is set of points of the form $A\vec{x}$ for some $\vec{x}$:

$$A\vec{x} = x_1(1\text{st col}) + \cdots + x_n(n\text{th col}).$$

The equation $A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b}$ is in the column space.

Say A is $m \times n$:

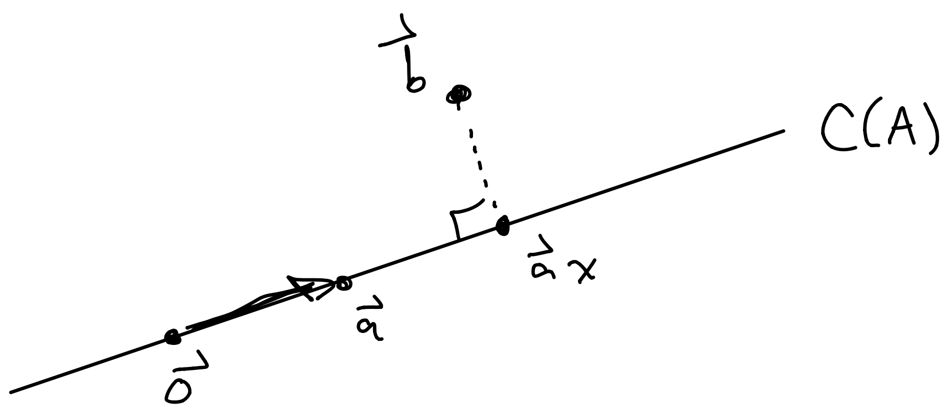


If $\vec{b}$ is not in the column space, we want to find some point $A\vec{x}$ in the column space such that the distance

$\|A\vec{x} - \vec{b}\|$ is MINIMIZED.

We could use multivariable calculus, but it is much simpler to use Linear Algebra!

Example: Let $A = \vec{a}$ be a column vector. The column space $C(A)$ is the line $\vec{a}x$, where x is any scalar.



Observe that distance $\|\vec{a}x - \vec{b}\|$ is minimized when $\vec{a}x - \vec{b} \perp \vec{a}$. We can use this to compute the scalar x:

$$\vec{a}x - \vec{b} \perp \vec{a}$$

$$\Longrightarrow \quad \vec{a} \cdot (\vec{a}x - \vec{b}) = 0$$

$$\Rightarrow (\vec{a}\cdot\vec{a})x - (\vec{a}\cdot\vec{b}) = 0$$

$$\Rightarrow \quad x = \frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}} \quad \text{DONE.}$$

We call this the orthogonal projection of point $\vec{b}$ onto the line $\vec{a}x$.

$$\text{proj}_{\vec{a}}(\vec{b}) = \left(\frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}}\right)\vec{a}.$$

We can see from this that "projection onto $\vec{a}x$" is a <u>linear</u> function. What is the matrix?

$$\text{proj}_{\vec{a}}(\vec{b}) = P\vec{b}.$$

I claim that

$$\boxed{\text{say } \vec{a}\in\mathbb{R}^m}$$

$$P = \underbrace{\frac{1}{\vec{a}\cdot\vec{a}}}_{\text{scalar}}\ \underbrace{\vec{a}\,\vec{a}^T}_{m\times m \text{ matrix}}$$

Proof: For any $\vec{b} \in \mathbb{R}^m$ we have

$$P\vec{b} = \frac{1}{\vec{a}\cdot\vec{a}}(\vec{a}\vec{a}^T)\vec{b}$$

$$= \frac{1}{\vec{a}\cdot\vec{a}}\vec{a}\,(\vec{a}^T\vec{b})$$

just a scalar

$$\boxed{\vec{a}^T\vec{b} = \vec{a}\cdot b}$$

$$= \frac{1}{\vec{a}\cdot\vec{a}}\vec{a}\,(\vec{a}\cdot\vec{b})$$

$$= \frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}}\vec{a} \quad \checkmark$$

For example: To project onto the line $t(1,1)$, the matrix is

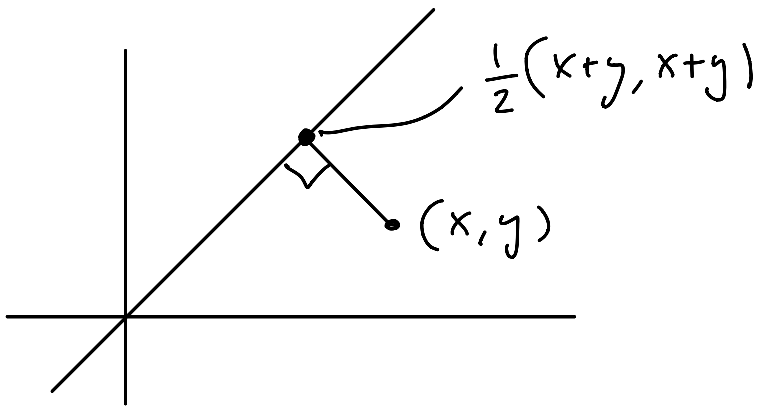
$$P = \frac{1}{\|(1,1)\|^2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$$

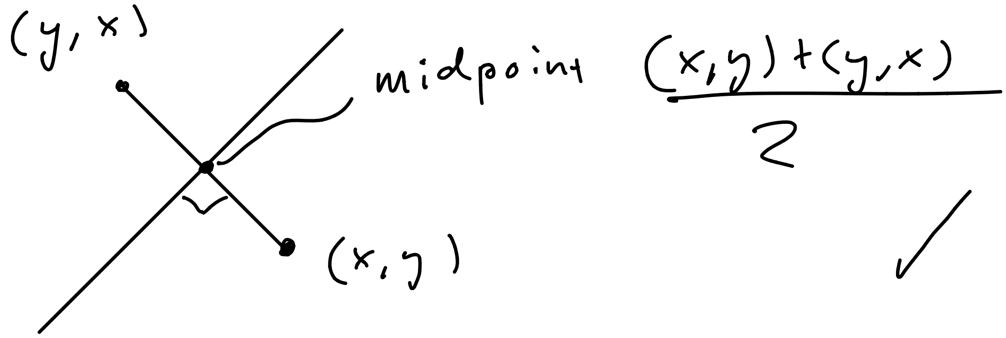
So point $\begin{pmatrix}x\\y\end{pmatrix}$ gets projected to the point $P\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

$$= \frac{1}{2}\begin{pmatrix} x+y \\ x+y \end{pmatrix}.$$

Picture:



Observe how this relates to the reflection across the line:



✓

Another Example: Project onto the line $t(1,-2,1)$ in $\mathbb{R}^3$.

$$P = \frac{1}{\|(1,-2,1)\|^2} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}$$

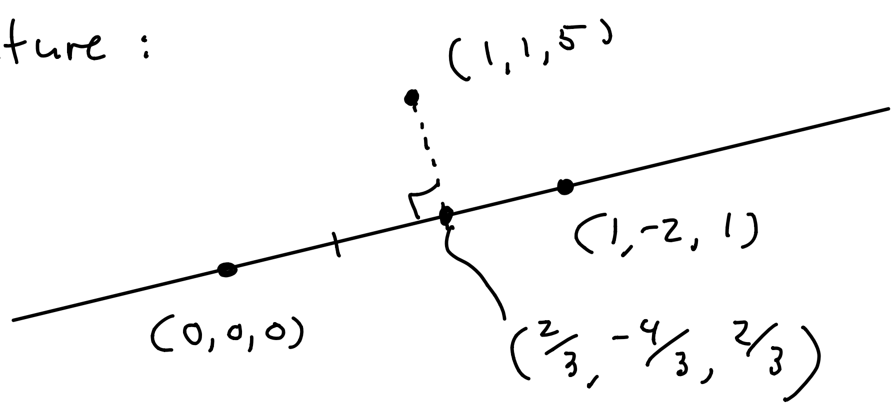
$$= \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Project the point $(1,1,5)$:

$$P\begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}$$

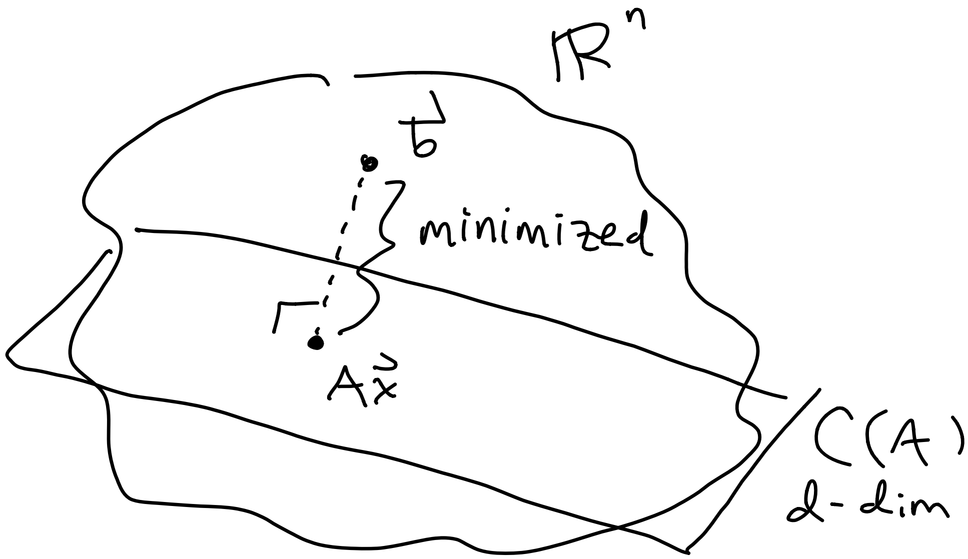
$$= \frac{1}{6} \begin{pmatrix} 4 \\ -8 \\ 4 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ -4 \\ 2 \end{pmatrix}.$$

Picture:



Projection onto a plane in $\mathbb{R}^3$?
Projection onto a d-plane in $\mathbb{R}^n$?
It's all the same!

Any d-plane in $\mathbb{R}^n$ is the column space of some $n \times d$ matrix A with independent columns:



Key Geometric Fact:
Distance $\|A\vec{x} - \vec{b}\|$ is minimized when vector $A\vec{x} - \vec{b}$ is $\perp$ to $C(A)$.

How to compute $\vec{x}$:

$(A\vec{x} - \vec{b}) \perp C(A)$

$\iff A\vec{x} - \vec{b} \perp$ every col of A.

$\iff A\vec{x} - \vec{b} \perp$ every row of A^T.

$\iff A^T(A\vec{x} - \vec{b}) = \vec{0}$

(Remember This? $N(A^T) = C(A)^\perp$)

Hence we must have

$$A^T A \vec{x} - A^T \vec{b} = \vec{0}$$

$$A^T A \vec{x} = A^T \vec{b}.$$

"The Normal Equation"

It expresses a bunch of right angles.

Observe : $\underset{d\times n}{A^T}\ \underset{n\times d}{A}$ is square $d \times d$.

Is it invertible? PAUSE.

Theorem: $N(A^TA) = N(A)$.

Proof: If $A\vec{x} = \vec{0}$ then

$$(A^TA)\vec{x} = A^T(A\vec{x}) = A^T\vec{0} = \vec{0} \quad ✓$$

Conversely, if $(A^TA)\vec{x} = \vec{0}$, we will show that $A\vec{x} = \vec{0}$. GOOD TRICK:

$$\begin{aligned} A^TA\vec{x} &= \vec{0} \\ \vec{x}^TA^TA\vec{x} &= \vec{x}^T\vec{0} \\ (A\vec{x})^T(A\vec{x}) &= 0 \\ \|A\vec{x}\|^2 &= 0. \end{aligned}$$

So $A\vec{x}$ is a vector of length 0, hence $A\vec{x} = \vec{0}$. ✓ ///

It follows from this fact and the FTLA that

$$\begin{aligned} \operatorname{rank}(A^TA) &= \operatorname{rank}(A) \\ &= \operatorname{rank}(A^T) \\ &= \operatorname{rank}(AA^T). \end{aligned}$$

If A has independent columns, then $\text{rank}(\underset{d\times d}{A^TA}) = \text{rank}(\underset{n\times d}{A}) = d$, and hence A^TA is <u>invertible</u> !

UNPAUSE.

Thefore the normal equation

$$A^TA\vec{x} = A^T\vec{b}$$

has a <u>unique</u> solution:

$$\vec{x} = (A^TA)^{-1}A^T\vec{b}.$$

But recall: $A\vec{x}$ is the projection of $\vec{b}$ onto the column space, so

$$\text{proj}_{C(A)}(\vec{b}) = A\vec{x}$$

$$= \underbrace{A(A^TA)^{-1}A^T}\vec{b}$$

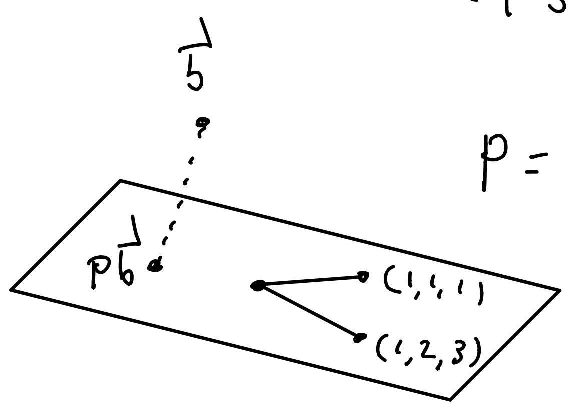
$$= P\vec{b}.$$

Conclusion: If matrix A has independent columns, then

$$P = A(A^TA)^{-1}A^T$$

is the matrix that <u>projects</u> any point onto the column space of A.

Example: $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$.



$$P = A(A^TA)^{-1}A^T.$$

$$A^TA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}$$

$$(A^TA)^{-1} = \frac{1}{6}\begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}$$

$$P = A(A^TA)^{-1}A^T$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$$

$$= \frac{1}{6} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{pmatrix}$$

$$= \frac{1}{6} \begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix} \quad \text{Observe: } P^T = P$$

Observe :

$\frac{1}{6}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$ projects onto line $t\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

$\frac{1}{6}\begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix}$ projects onto plane $r\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + s\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

If we add them :

$$\frac{1}{6}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad !$$

Why did that happen ??