HW5 due Thurs before class.

Discuss Problem 2(a,b).

We've seen the <u>geometry</u> of projection.
This problem is about the <u>arithmetic</u> of projection (strange...).

Say matrix P is an orthogonal <u>projection</u> when $P^T = P$ & $P^2 = P$.

(a) If P is a projection then

$Q = I - P$ is a projection.

What do we need to check?

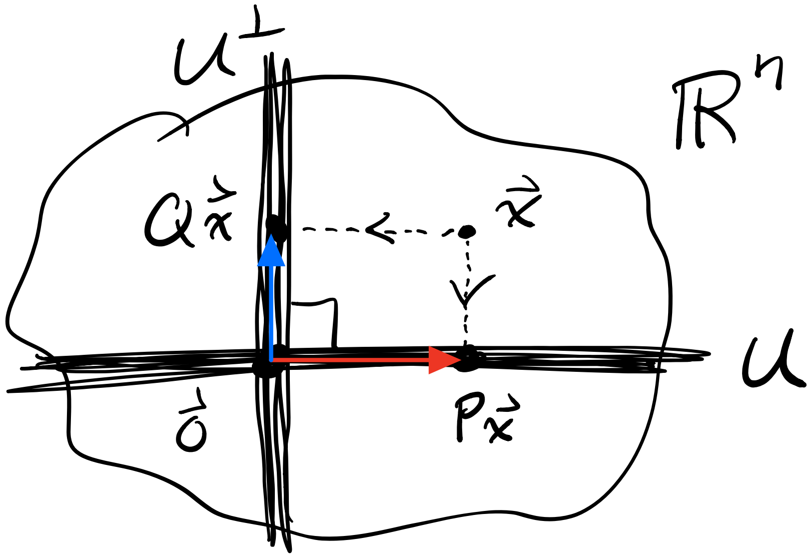
$$
\begin{aligned}
Q^T &= (I-P)^T \\
&= I^T + (-P)^T && [(A+B)^T = A^T + B^T] \\
&= \underline{I}^T - P^T && [(\lambda A)^T = \lambda A^T] \\
&= \underline{I} - P && [\underline{I}^T = \underline{I} \text{ \& } P^T = P] \\
&= Q \quad \checkmark
\end{aligned}
$$

$$
\begin{aligned}
Q^2 &= (I-P)^2 \\
&= (I-P)(I-P) \\
&= II - IP - PI + PP \\
&= I - P - P + P^2 \\
&= I - 2P + P \qquad [P^2 = P] \\
&= I - P \\
&= Q \quad \checkmark
\end{aligned}
$$

Done.

But what does this <u>mean</u>?

Suppose P is the projection onto a subspace $U \subseteq \mathbb{R}^n$:



Let Q be the projection onto the orthogonal complement subspace $U^\perp \subseteq \mathbb{R}^n$. Observe that

$$\vec{0}, \; P\vec{x}, \; Q\vec{x}, \; \vec{x}$$

are the four vertices of a rectangle (which is a parallelogram), hence

$$P\vec{x} + Q\vec{x} = \vec{x}$$
$$(P+Q)\vec{x} = \vec{x}$$

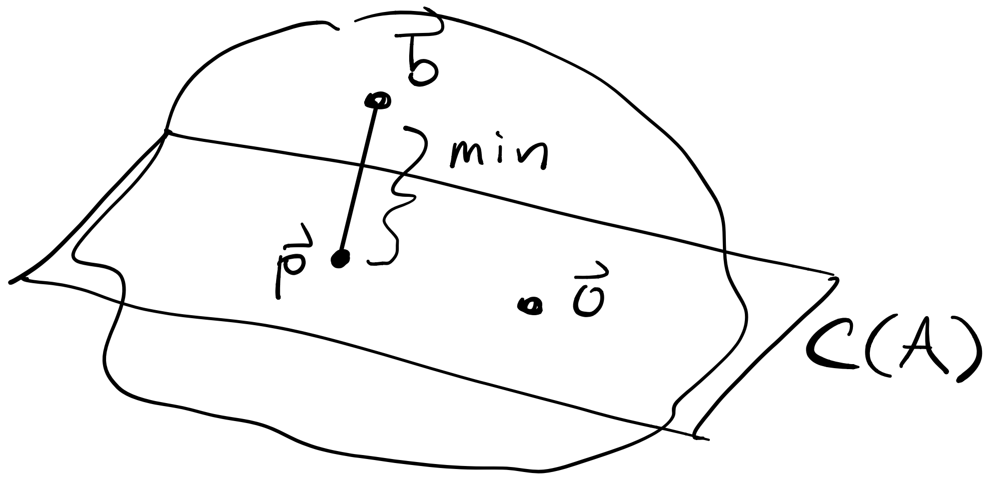
So $P+Q$ is the matrix that does nothing, i.e., $P+Q = I$.

In other words,

$$Q = I - P.$$

Review: Suppose $A\vec{x} = \vec{b}$ has no solution. We want to find $\hat{x}$ such that $\|A\hat{x}-\vec{b}\|^2$ is MINIMIZED.

Picture:



The shortest length $\|\vec{p}-\vec{b}\|^2$ happens when $\vec{p}-\vec{b}$ is $\perp$ to the column space. Observe that

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

$\perp$ every column of A

$\perp$ every row of A^T

In other words:

(1) $A^T(\vec{p}-\vec{b}) = \vec{0}$

We also want $\vec{p}$ to be <u>in</u> the column space, so that

(2) $\vec{p} = A(\text{something})$
$= A\hat{x}$

for some vector $\hat{x}$.

Put (1) & (2) together:

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

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

Called the "normal equation" because it expresses a bunch of right angles.

If you want an <u>expression</u> for

$$\vec{p} = A\hat{x},$$

Let's assume A^TA is invertible so that

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

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

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

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

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

The orthogonal projection of point $\vec{b}$ onto column space of A.

Corresponding matrix:

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

Example: Project the point $(1,2,3)$ onto the line $t(1,1,2)$ &

onto the plane $x+y+2z=0$.

Let P be the matrix that projects onto the line $t(1,1,2)$, i.e., the column space of $\vec{a}=\begin{pmatrix}1\\1\\2\end{pmatrix}$.

$$P=\vec{a}\,(\vec{a}^T\vec{a})^{-1}\,\vec{a}^T$$

$$=\begin{pmatrix}1\\1\\2\end{pmatrix}\left[\begin{pmatrix}1&1&2\end{pmatrix}\begin{pmatrix}1\\1\\2\end{pmatrix}\right]^{-1}\begin{pmatrix}1&1&2\end{pmatrix}$$

$$=\begin{pmatrix}1\\1\\2\end{pmatrix}\left[1+1+4\right]^{-1}\begin{pmatrix}1&1&2\end{pmatrix}$$

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

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

Project $(1,2,3)$:

$$P\begin{pmatrix}1\\2\\3\end{pmatrix}=\frac{1}{6}\begin{pmatrix}1&1&2\\1&1&2\\2&2&4\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 1+2+6 \\ 1+2+6 \\ 2+4+12 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 9 \\ 9 \\ 18 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 3/2 \\ 3 \end{pmatrix}. \qquad ///$$

Now project onto the plane $x+y+2z=0$.

There are two ways to do this.

Hard Way: Express the plane as the column space of a matrix.

Find two vectors in the plane ....

$$\{ x + y + 2z = 0$$

pivot: x

Free: y & z. Let $s=y$ & $t=z$.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -s-2t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

two vectors in the plane.

Plane is col. space of $A = \begin{pmatrix} -1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Compute the projection matrix

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

$\vdots$ computer

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

Project $(1,2,3)$ onto the plane:

$$Q\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5 & -1 & -2 \\ -1 & 5 & -2 \\ -2 & -2 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 5-2-6 \\ -1+10-6 \\ -2-4+6 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ 0 \end{pmatrix}.$$

Easy Way: The line $t(1,1,2)$ & plane $x+y+2z=0$ are orthogonal complements, hence

$$P + Q = I.$$

$$P\begin{pmatrix}1\\2\\3\end{pmatrix} + Q\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}$$

$$Q\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix} - P\begin{pmatrix}1\\2\\3\end{pmatrix}$$

$$= \begin{pmatrix}1\\2\\3\end{pmatrix} - \begin{pmatrix}3/2\\3/2\\3\end{pmatrix}$$

$$= \begin{pmatrix}-1/2\\1/2\\0\end{pmatrix}$$

Application: Fitting a line to data points. We want the best fit line $y = mx + b$ for n arbitrary data points:

$$(x,y) = (x_1,y_1), (x_2,y_2), \dots, (x_n,y_n).$$

Each data point gives a linear equation in 2 unknowns m & b.

$$y_i = mx_i + b.$$

Get a system of n equations, 2 unknowns:

$$\begin{cases} mx_1 + b = y_1 \\ mx_2 + b = y_2 \\ \quad \vdots \\ mx_n + b = y_n \end{cases} \rightarrow \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

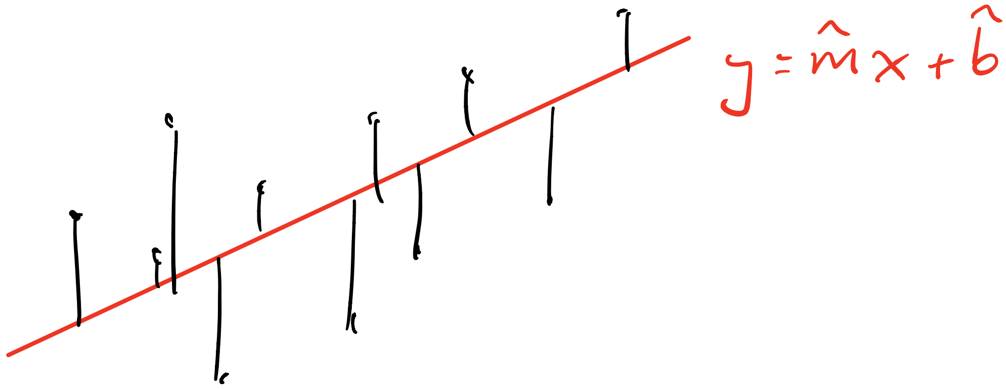
This has (probably) NO SOLUTION.

So consider the normal equation:

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \hat{m} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} \sum x^2 & \sum x \\ \sum x & n \end{pmatrix} \begin{pmatrix} \hat{m} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \sum xy \\ \sum y \end{pmatrix}$$

This can be easily solved to obtain $\hat{m}, \hat{b}$ and the best fit line:



Sum of squares of vertical errors is as small as possible.

One Step Further: Find the best fit parabola $y = ax^2 + bx + c$ for n data points $(x_1, y_1), \ldots, (x_n, y_n)$.

Have n linear eqns in 3 unknowns:

$$\begin{cases} a x_1^2 + b x_1 + c = y_1 \\ \qquad \vdots \\ a x_n^2 + b x_n + C = y_n \end{cases}$$

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

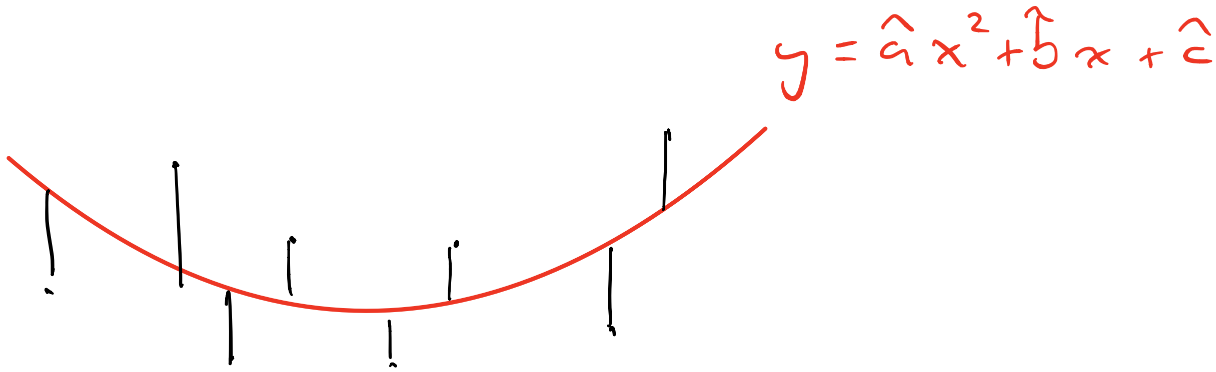
$$\Big\downarrow \text{ normal equation}$$

$$\begin{pmatrix} x_1^2 & \cdots & x_n^2 \\ x_1 & \cdots & x_1 \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} x_1^2 & \cdots & x_n^2 \\ x_1 & \cdots & x_1 \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

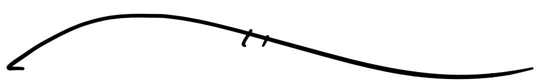
$$\begin{pmatrix} \Sigma x^4 & \Sigma x^3 & \Sigma x^2 \\ \Sigma x^3 & \Sigma x^2 & \Sigma x \\ \Sigma x^2 & \Sigma x & n \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} \Sigma x^2 y \\ \Sigma x y \\ \Sigma y \end{pmatrix}$$

Now solve for $\hat{a}, \hat{b}, \hat{c}$ to obtain the best fit parabola:

$$y = \hat{a} x^2 + \hat{b} x + c.$$



Sum of squares of verticall errors is as small as possible.



Example: Find best fit parabola for 5 data points (x, y):

$$(-2,1),\ (-1,5),\ (0,5),\ (1,2),\ (2,1)$$

Plug data into $y = ax^2 + bx + c$ to get 5 linear eqns in 3 unknowns a, b, c:

$$\begin{cases} 4a - 2b + c = 1 \\ 1a - b + c = 5 \\ 0a + 0b + c = 5 \\ 1a + b + c = 2 \\ 4a + 2b + c = 1 \end{cases}$$

$$\begin{pmatrix} 4 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \\ 4 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 5 \\ 2 \\ 1 \end{pmatrix}$$

$$\Big\downarrow \text{ normal equation}$$

$$\begin{pmatrix} 34 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 5 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} 15 \\ -3 \\ 14 \end{pmatrix}$$

$$\Big\downarrow \text{ solve}$$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{pmatrix} = \begin{pmatrix} -13/14 \\ -3/10 \\ 163/35 \end{pmatrix}$$

Picture: