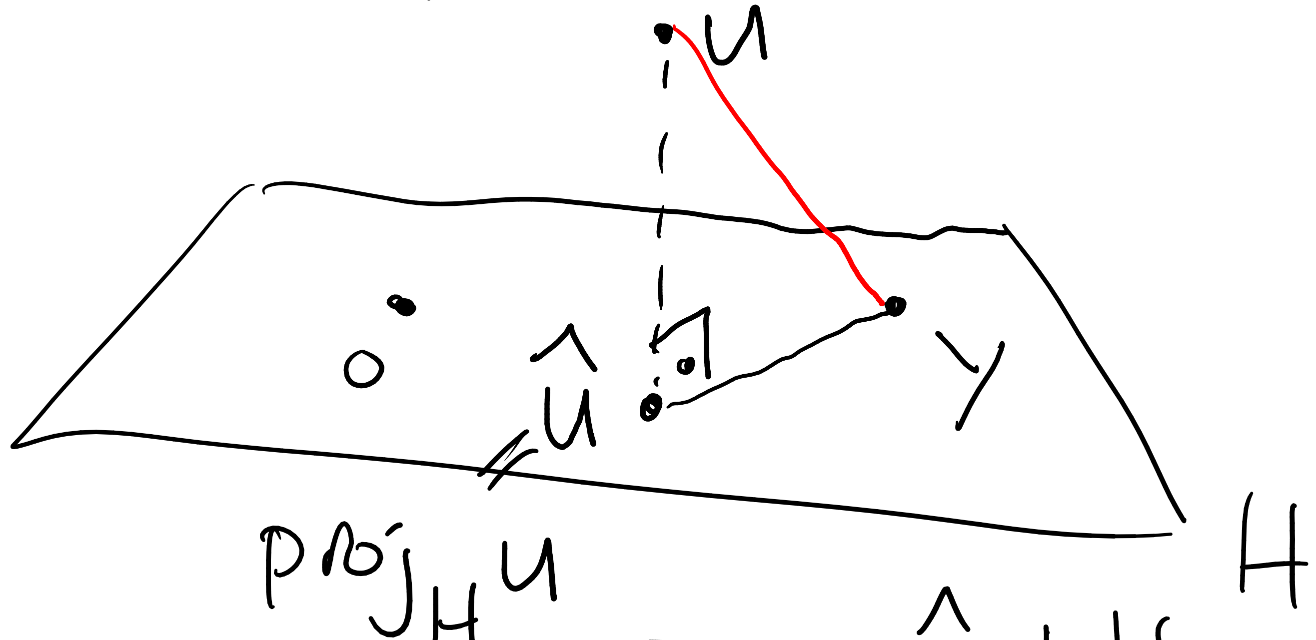
- Given $H \subset \mathbb{R}^n$ a subspace; $(k \leq n)$ $A_{n\times k}$ with $\text{col}(A) = H$ & colums linearly independent.

$\underset{n\times n}{P} = A(A^TA)^{-1}A^T$ projects orthogonally vectors in $\mathbb{R}^n$ to H.

- $P^2 = P$ & $P^T = P$; i.e. P is a projection matrix.
- So $\text{col}(P) = \text{col}(A)$.

  & $\text{rank}(P) = \text{rank}(A) = k$.

- $\dim \text{null}(P) = n-k$.
- If $Pu = \hat{u}$ then for any $n \in \text{null}P$, $u+n$ projects orthog to $\hat{u}$ too.
- 

so $u - \hat{u} \perp H$.

Therefore for any $y \in H$, the distance from u to $y \geq$ the distance from u to $\hat{u}$.

thm1. $\text{proj}_H u$ is the closest point in H to u (in the Euclidean distance)

thm2. Any projection matrix $Q_{n \times n}$ orthogonally projects to a subspace of $\mathbb{R}^n$.

proof. For any x, $Qx \in H = col(Q)$

We claim that $Qx = \text{proj}_H x$.

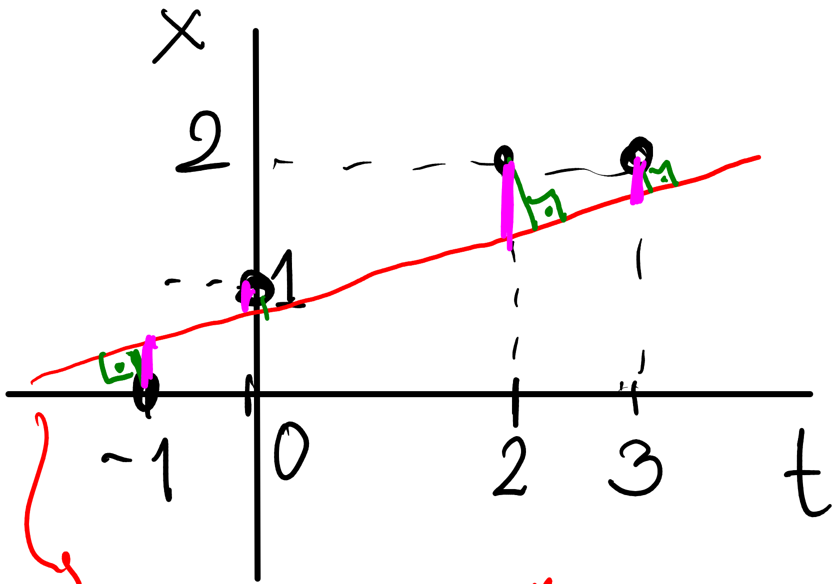
$Qx \in H$ already. We must prove:

$x - Qx \perp H$ i.e. $x - Qx \perp$ every column of Q

$$\Leftrightarrow Q^T(x - Qx) = 0 \longleftrightarrow 0$$

$$Q^T(x-Qx) = Q^T x - Q^T Q x \overset{Q^T = Q}{=} Qx - Q^2 x \overset{Q^2 = Q}{=} Qx - Qx = 0$$

LEAST SQUARES APPROXIMATION



t	x
-1	0
0	1
2	2
3	2

L is the "best fitting" line.

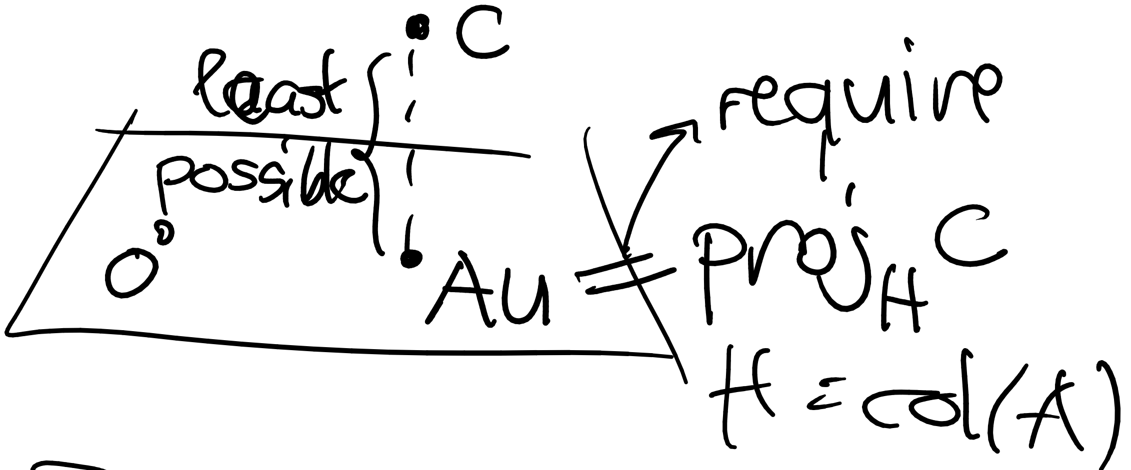
$L: x = a + bt$; a, b unknown.

Find them, so that the sum of the square of the vertical distances of the data points to L is minimum.

We fancy a line L:

$$\begin{aligned} a + (-1)b &= 0 \\ a + 0 \cdot b &= 1 \\ a + 2b &= 2 \\ a + 3b &= 2 \end{aligned} \iff \overbrace{\begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}^{A} \overbrace{\begin{pmatrix} a \\ b \end{pmatrix}}^{u} = \overbrace{\begin{pmatrix} 0 \\ 1 \\ 2 \\ 2 \end{pmatrix}}^{c}$$

There is no soln for u because apparently the data points are not collinear. Instead we are content to solve u such that Au is closest to c.



This solves the minimizing the sum of the square of vertical distances problem.

Let's solve.

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

$$= A \cdot \begin{pmatrix} 4 & 4 \\ 4 & 14 \end{pmatrix}^{-1} \cdot A^T = A \cdot \frac{1}{40}\begin{pmatrix} 14 & -4 \\ -4 & 4 \end{pmatrix} A^T$$

$$= A \cdot \frac{1}{40} \cdot \begin{pmatrix} 18 & 14 & 6 & 2 \\ -8 & -4 & 4 & 8 \end{pmatrix}$$

Thus solve $\underset{4\times 2}{A} u = Pc \in col(A)$

The soln for u is unique because $A_{4\times 2}$ has rank 2.

$$Au = Pc = A \underbrace{(A^T A)^{-1} A^T c}$$

So the unique soln u for u is

$$\boxed{u = (A^T A)^{-1} A^T c}$$

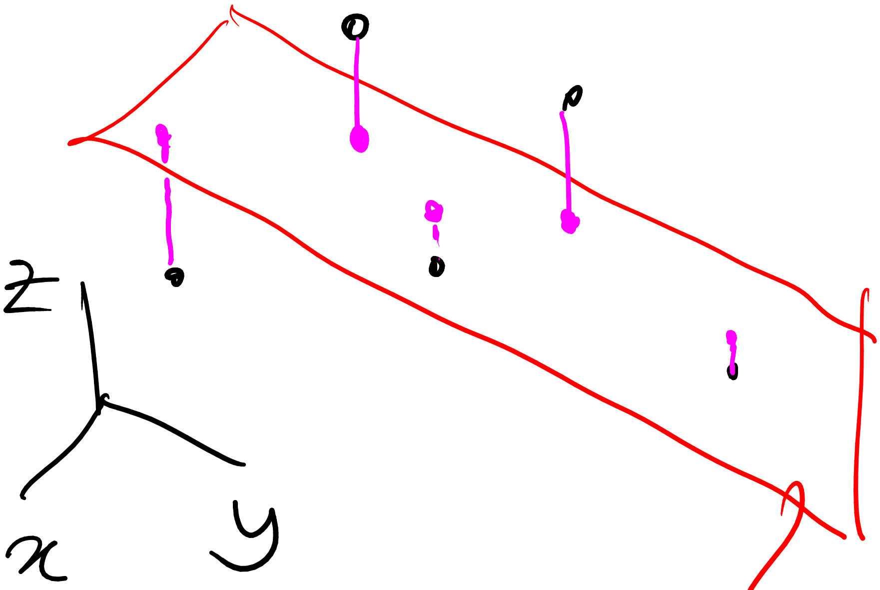
$$= \frac{1}{20} \begin{pmatrix} 9 & 7 & 3 & 1 \\ -4 & -2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ 2 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 15 \\ 10 \end{pmatrix}$$

so that $a = 3/4$, $b = 1/2$.

<u>ex:</u> Measurement:

x	y	z
x_1	$y_1 \to$	z_1
$\vdots$	$\vdots$	$\vdots$
x_k	$y_k \to$	z_k

Here $z = z(x,y)$



best fitting plane

$z = a + bx + cy$

We solve

$$\underbrace{\begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_k & y_k \end{pmatrix}}_{A} \cdot \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{u} = \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix} = z$$

Instead solve $Au = \text{proj}_{\text{col}(A)}\, z$:

$$\boxed{u = (A^TA)^{-1}A^Tz}$$

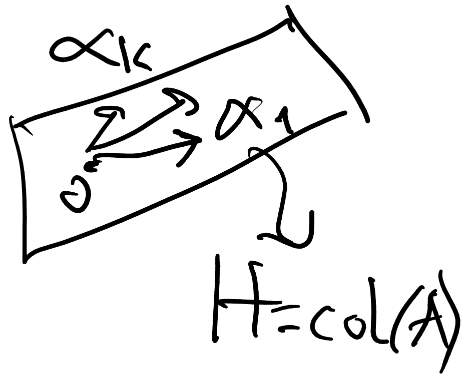
ORTHONORMALITY & ORTHOGONAL MATRICES

Back to projection matrices.

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

with $A = (\alpha_1 \vdots \cdots \vdots \alpha_k)$

& $\{\alpha_1, \dots, \alpha_k\}$ a basis for H. Observe:



$$A^TA = \begin{pmatrix} \cdots \alpha_1^T \cdots \\ \vdots \\ \cdots \alpha_k^T \cdots \end{pmatrix} (\alpha_1 \vdots \cdots \vdots \alpha_k)$$

$$= \begin{pmatrix} \alpha_1^T\alpha_1 & \alpha_1^T\alpha_2 & \cdots \\ \alpha_2^T\alpha_1 & \ddots & \\ \vdots & & \end{pmatrix} = I$$

we'd have

Assume $\alpha_1, \dots, \alpha_k$ is an orthogonal collection with lengths 1,

i.e. $\alpha_i^T\alpha_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Kronecker delta

In that case $P = AA^T$.

<u>defn</u>: An orthogonal collection with all lengths 1 is called an <u>orthonormal</u> collection.

<u>thm 3.</u> Let $\alpha_1, \dots, \alpha_k$ be orthonormal basis for $\mathbb{R}^k$. Then for $u \in \mathbb{R}^k$,

$$u = \text{proj}_{\alpha_1} u + \cdots + \text{proj}_{\alpha_k} u.$$

<u>proof</u>: Let $u = c_1\alpha_1 + \cdots + c_k\alpha_k$.

Then
$$\alpha_1^T u = c_1 \underbrace{\cancel{\alpha_1^T\alpha_1}}_{1} + \cdots + c_k \underbrace{\cancel{\alpha_1^T\alpha_k}}_{0}$$
$$= c_1$$

Similary $c_j = \alpha_j^T u$ so that

$$u = \frac{\alpha_1^T u}{\alpha_1^T\alpha_1}\alpha_1 + \cdots + (\alpha_k^T u)\alpha_k$$
$$= \text{proj}_{\alpha_1} u + \cdots + \text{proj}_{\alpha_k} u.$$