MATH 312
Lecture 14: "Fit"

LEAST - SQUARES

Last time: Projection onto subspaces.

- If $\vec{w}$ is a vector in $\mathbb{R}^n$ and V is some subspace, then we saw how to project $\vec{w}$ onto V using orthogonality:

→ Find a basis $\vec{v}_1, \dots, \vec{v}_k$ of V

→ Construct the matrix $A = [\vec{v}_1 \cdots \vec{v}_k]$, of size $n \times k$.

→ Attack $\vec{w}$ on the left by:

$$P = \left[A\,(A^T A)^{-1} A^T\right].$$

→ Done!

Two things about $P = [A(A^TA)^{-1}A^T]$

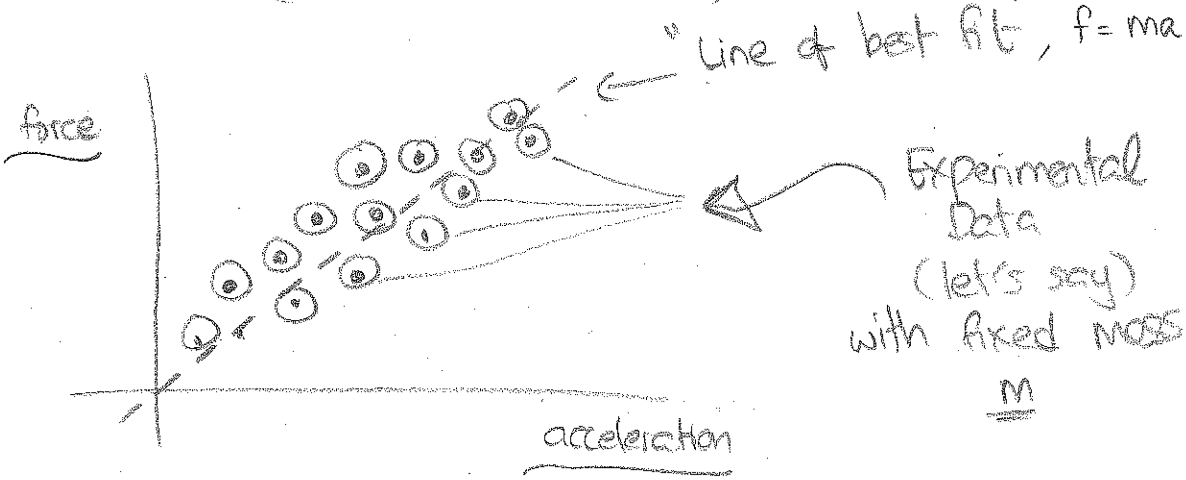
a) $P^2 = P$ (Proj of Proj = Proj)

b) $P^T = P$ (Symmetric about diagonal)

Fact: ANY symmetric matrix P with $P^2 = P$ is one of these projections!

If V is all of $\mathbb{R}^n$, then both A and A^T are invertible (A's columns must form a basis for $\mathbb{R}^n$!), so in this case (and ONLY IN THIS CASE) we can use $(A^TA)^{-1} = A^{-1} \cdot (A^T)^{-1}$ to get P = identity! Proj. onto whole space changes nothing: you're already there.

LEAST SQUARES (DATA - FITTING)



OR, "What to do when $Ax = b$ has no solutions".

Part 1 Finding LINES of best fit

Let's say $f_1, \dots, f_k$ were forces measured (for particles of mass m) corresponding to accelerations $a_1, \dots, a_k$. We are searching for the best "linear relationship", i.e.,

$$f = c\,a + d$$

(WANT: c and d) to best describe these observations. In a perfect world, we'd get

$$\begin{aligned} f_1 &= c a_1 + d \\ f_2 &= c a_2 + d \\ &\vdots \\ f_k &= c a_k + d, \end{aligned}$$

i.e

$$\begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \\ 1 & a_k \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_k \end{bmatrix}$$

Linear system, eg $A\vec{x} = \vec{b}$

BUT can't expect to find solutions even when $k=3$:

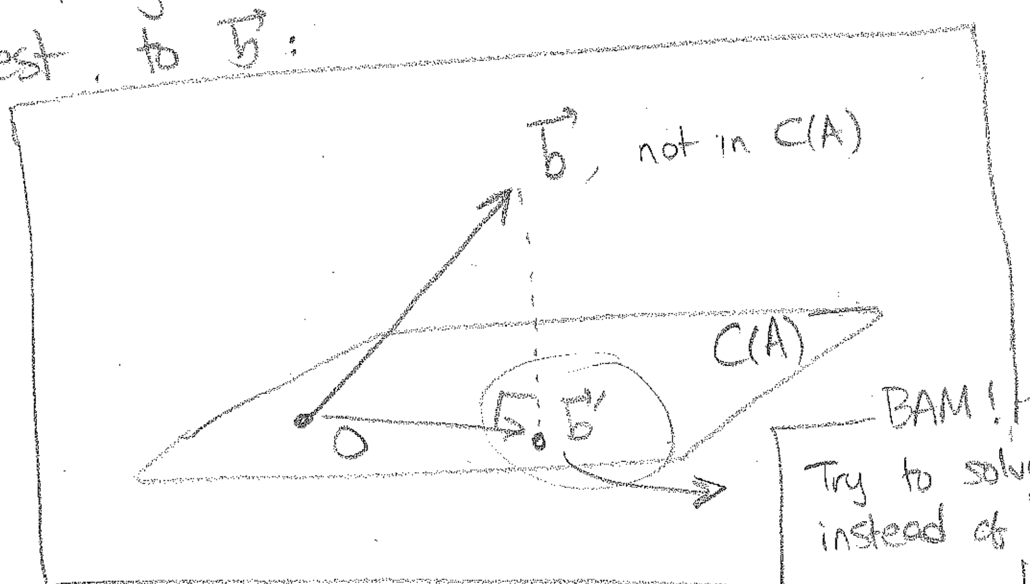
$$\text{eg} \quad \overset{(A)}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}} \overset{(x)}{\begin{bmatrix} c \\ d \end{bmatrix}} = \overset{(b)}{\begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}}$$

$$\Rightarrow \begin{bmatrix} c+d \\ c+2d \\ c+3d \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}$$

"OVERDETERMINED SYSTEM"

Three equations, only two unknowns. Good luck.

- When we can't find solutions for $Ax=b$, i.e., when b is NOT in $C(A)$, then the next best thing is to find the vector $\vec{b}'$ in $C(A)$ nearest to $\vec{b}$:



BAM!

Try to solve $A\vec{x} = \vec{b}'$ instead of $A\vec{x} = \vec{b}$!

But first, let's try CALCULUS!

Returning to $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}$, let's find $\begin{bmatrix} c \\ d \end{bmatrix}$ that minimize the "discrepancy" or SQUARED ERROR!!!

$$E(c,d) = \left\| \begin{bmatrix} c+d \\ c+2d \\ c+3d \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix} \right\|^2 =$$

This requires:

- Computing $E(c,d) =$

  $(c+d-3)^2 + (c+2d)^2 + (c+3d-5)^2$

  $=$ something...

- Taking partial derivatives, setting $= 0$

  $\partial E/\partial c = 2(c+d-3) + 2(c+2d) + 2(c+3d-5) = 0$

  $\partial E/\partial d = 2(c+d-3) + 4(c+2d) + 6(c+3d-5) = 0$

Also have to check second derivatives...

- These are LINEAR EQUATIONS in c and d!

  $6c + 12d = 16$
  $12c + 28d = 36$

  Or, $\begin{bmatrix} 6 & 12 \\ 12 & 28 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 16 \\ 36 \end{bmatrix}$.

  (We'll come back to this!)

The $E(c,d)$ for these optimal values is called "least squares error!"

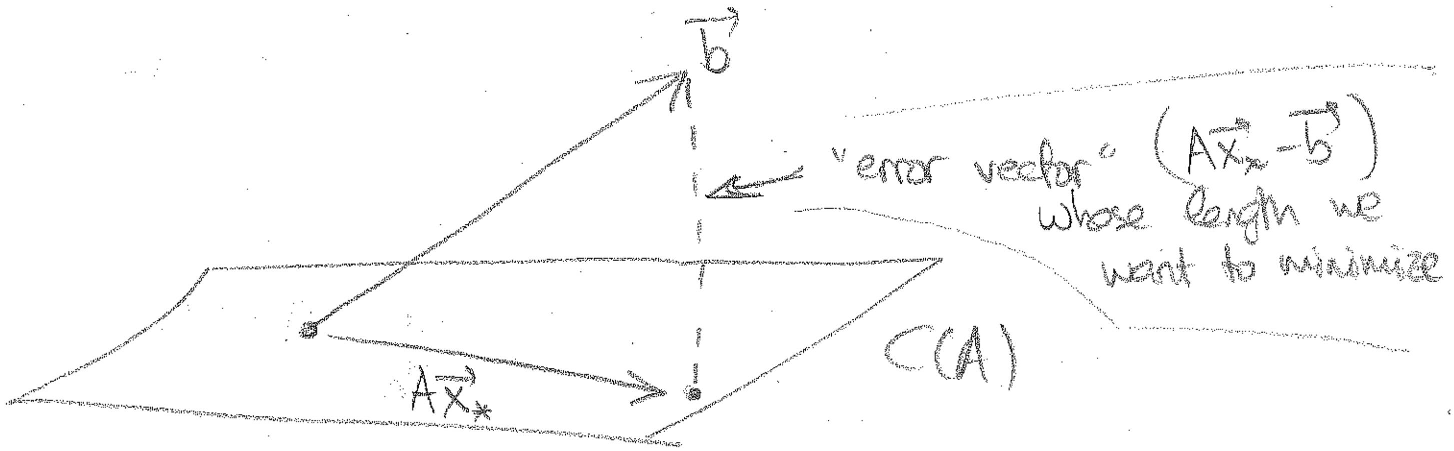
- An EQUIVALENT picture comes from GEOMETRY!

We are "approximating" $A\vec{x} = \vec{b}$, where

$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}$.

We want to determine the best approximation $\vec{x}_*$ so that $A\vec{x}_* =$ Proj of $\vec{b}$ on $C(A)$.

Picture, again:



The error $(A\vec{x}_* - \vec{b})$ must be orthogonal to $C(A)$, so $(A\vec{x}_* - \vec{b})$ lies in $C(A)^\perp = N(A^T)$: left nullspc

Then,

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

$$\Rightarrow \boxed{(A^T A)\,\vec{x}_* = A^T\vec{b}}$$ ← Called the "normal equation"

When $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}$, we get

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad \text{and} \quad A^T\vec{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$$

So, solve: $\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 8 \\ 16 \end{bmatrix}$

SAME as "calculus" eqns

Remarks:

1. This works in ANY dimension: you can also find PLANES of best fit through data points!

2. If $\vec{b}$ is actually in $C(A) = N(A^T)^\perp$ then the error $A\vec{x}_* - \vec{b}$ is the zero vector!

3. You don't HAVE to restrict to linear settings. For instance, say you want to find the parabola of best fit: $\underbrace{cx^2 + dx + e}_{\text{(QUADRATIC !!)}}$, then set up

$A^TA x = A^T b$

$$\begin{bmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ & \vdots & \\ 1 & a_n & a_n^2 \end{bmatrix} \begin{bmatrix} c \\ d \\ e \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

Still, you need to know what you're looking for.

And use the least-square method here!

3.3 #12 (Strang)

V is the subspace of $\mathbb{R}^4$ spanned by $(1,1,0,1)$ and $(0,0,1,0)$. Find

a) A basis for $V^\perp$

b) Projection matrix P_V

c) The vector in V closest to $(0,1,0,-1)$ in $V^\perp$