MCEN 3030

20 Feb 2024

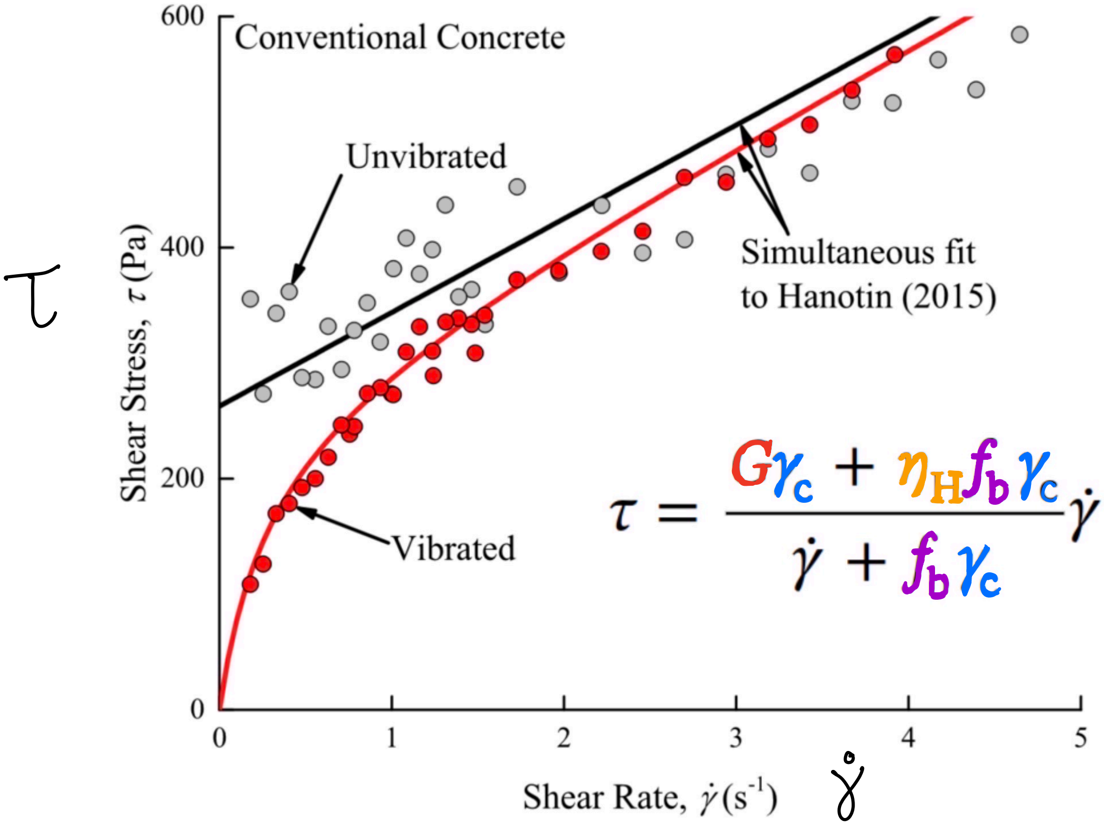
HW #4 Due Monday (mostly: building LU)

Exam #1 Tues Feb 27

Last time: $Ax = b \rightarrow Ux = d$

Today: Least-Squares Regression

$$\tau = G\gamma_c + \eta_H \dot{\gamma}$$



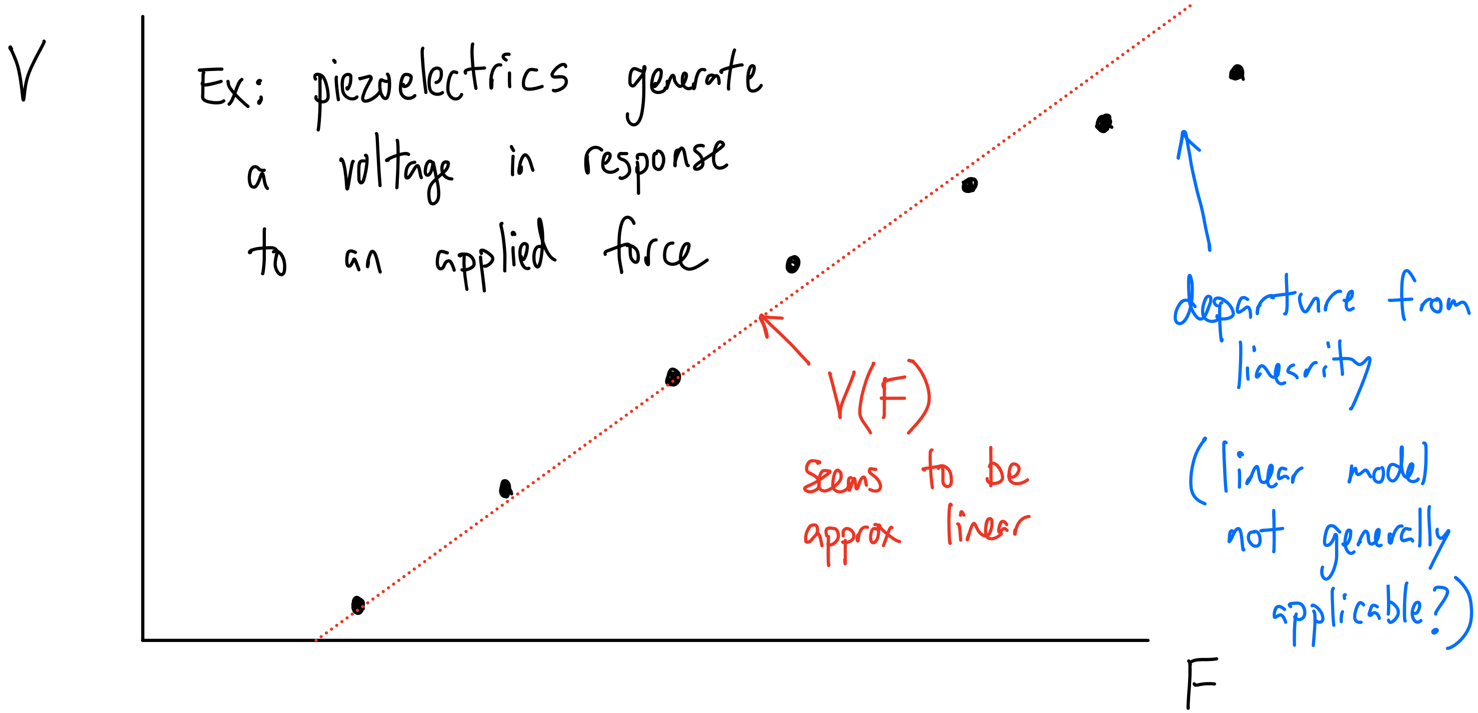
$$\tau = \frac{G\gamma_c + \eta_H f_b \gamma_c}{\dot{\gamma} + f_b \gamma_c}\dot{\gamma} + \frac{\eta_H}{\dot{\gamma} + f_b \gamma_c}\dot{\gamma}^2$$

Vibration of fresh concrete understood through the paradigm of granular physics
JA Koch, DI Castaneda, RH Ewoldt, DA Lange
Cement and Concrete Research 115, 31-42

Modeling / Data Fitting

The equations we use in engineering are at best approximations of reality, and experimental data is always flawed.

Ex: piezoelectrics generate a voltage in response to an applied force



$$V(F) = \underset{\uparrow}{a} + \underset{\uparrow}{b} F$$

model parameters

Q: How to determine the best/most accurate/most useful values of parameters a & b?

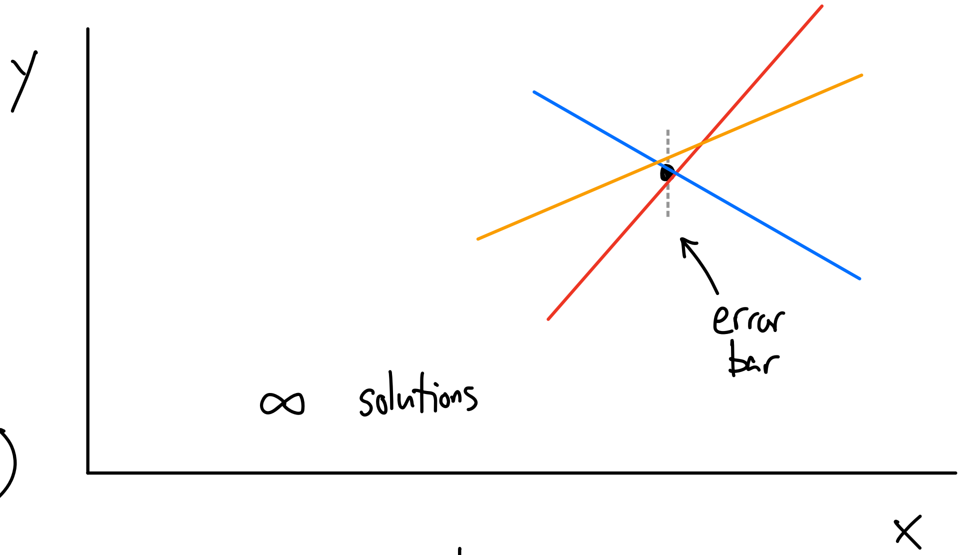
A: Model fitting of experimental data

We will likely need to determine the parameters experimentally/numerically. Suppose we have N measurements.

model $y = a + bx$ allegedly governs them all

$N = 1$

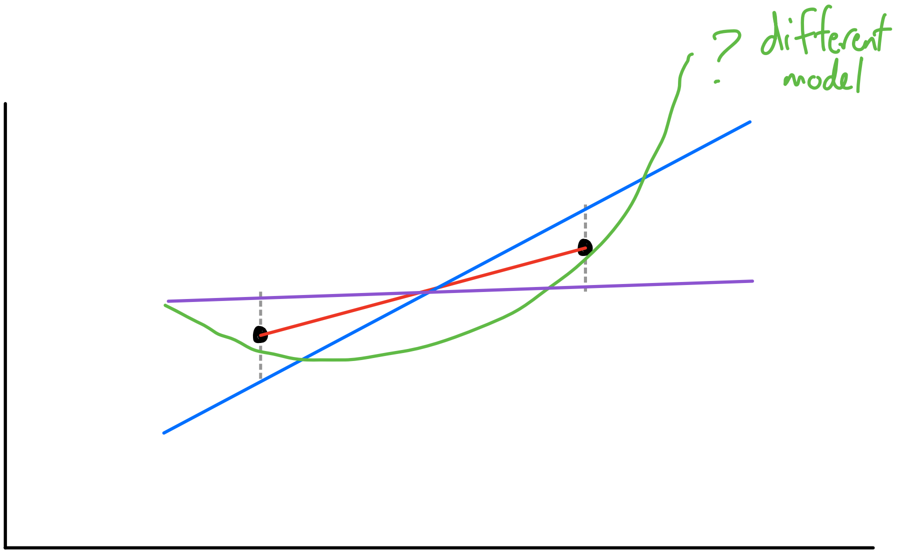
(we can tell, possibly, the order of magnitude of the respose/ demonstrate phenomenon)



<u>underdetermined</u> system $y_1 = a + bx_1$

$N = 2$

$$y_1 = a + bx_1$$
$$y_2 = a + bx_2$$

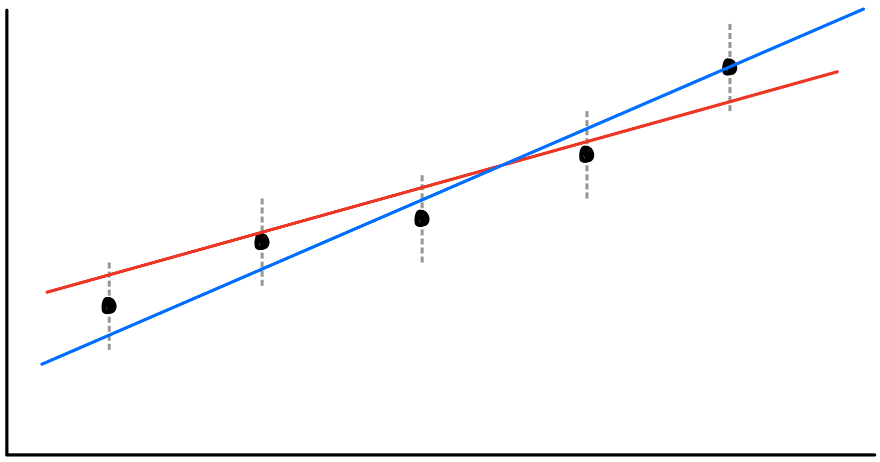


Linear algebraically we feel good – 2 eqns, 2 unk.

Statistically, we have doubts.

$N=5$?

Statistically, we feel better about 5 data points.



$$\left.\begin{aligned}
y_1 &= a + bx_1 \\
y_2 &= a + bx_2 \\
y_3 &= a + bx_3 \\
y_4 &= a + bx_4 \\
y_5 &= a + bx_5
\end{aligned}\right\}$$

over-determined

no solution?

technically yes, but we can still find the "best" solution

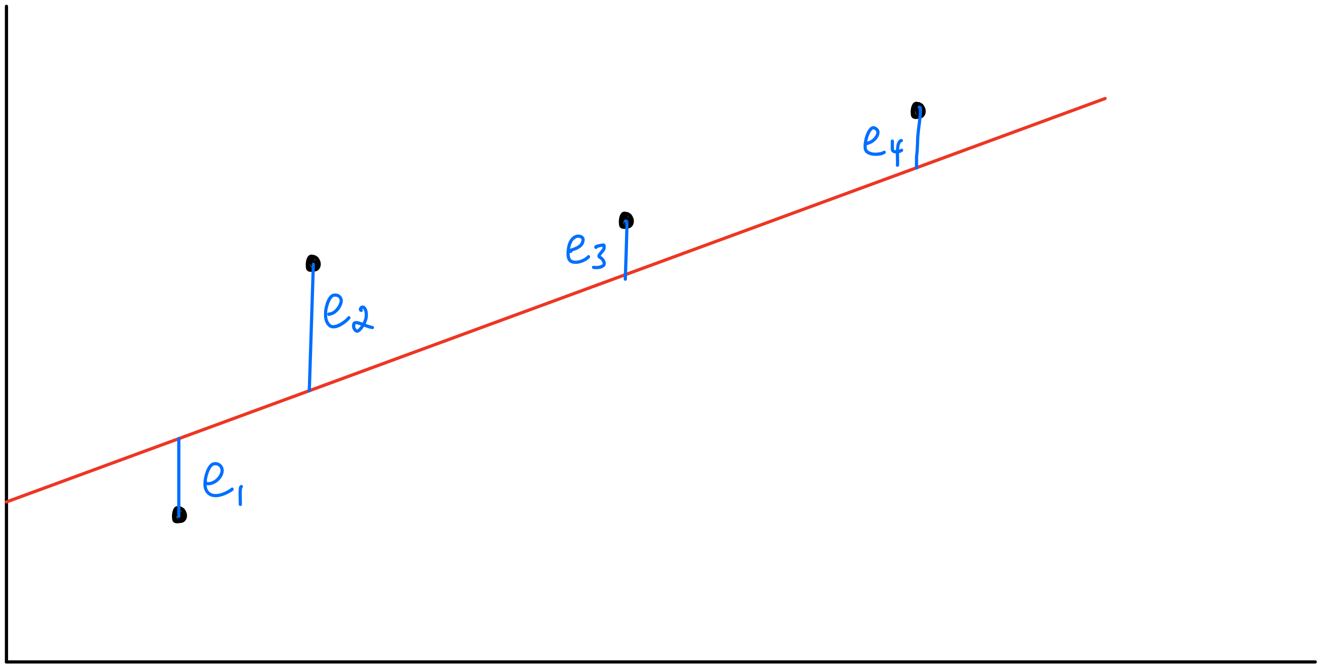
Least - Squares Regression

Idea: Quantify the error/difference of data vs model at each point & determine parameter values a, b, ... that minimize that error

Quantify with "the residual" from each data point i:

$$e_i \equiv \underset{\uparrow \atop \text{data}}{y_1} - \underset{\uparrow \atop \text{model}}{\hat{y}(x_i)}$$

model, $\hat{y}(x_i) = a + bx_i$



We aim to minimize

$$S_r = \sum_{i=1}^{N} \left(y_i - \hat{y}(x_i) \right)^2$$

"Sum of Residuals"

squared so that errors don't cancel out

Goal is to minimize S_r w.r.t. a & b.

$$\Longrightarrow \quad \frac{\partial S_r}{\partial a} = 0 \quad \& \quad \frac{\partial S_r}{\partial b} = 0$$

For a linear fit

$$S_r = \sum_{i=1}^{N} \left(y_i - \hat{y}(x_i)\right)^2 = \sum_{i=1}^{N} \left(y_i - a - bx_i\right)^2$$

Then:

$$\frac{\partial S_r}{\partial a} = \sum_{i=1}^{N} 2\left(y_i - a - bx_i\right)(-1) = 0$$

$$\rightarrow \sum \left(y_i - a - bx_i\right) = 0$$

$$\frac{\partial S_r}{\partial b} = \sum_{i=1}^{N} 2\left(y_i - a - bx_i\right)(-x_i) = 0$$

$$\rightarrow \sum \left(x_i y_i - a x_i - b x_i^2\right) = 0$$

This reduces to "$Ax = b$" form → easy to solve

$$\begin{bmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

i.e. $$\boxed{Na + b\sum x_i = \sum y_i}$$

only unknowns

The above is a totally acceptable way to fit a linear model to data. The following method is a slight modification that will set us up for nonlinear fitting

I called this $y = Za$ originally but changed it to avoid $a \leftrightarrow \underline{a}$ confusion

Data:
$$\begin{aligned} y_1 &= a + bx_1 \\ y_2 &= a + bx_2 \\ &\vdots \end{aligned} \longrightarrow Y = ZA$$

This system is "overdetermined" →

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

But we can write $E = Y - ZA$

looks like I flipped around anyway. This should have been $E = y - Za$, but I had $Y - ZA$ and will now keep it

$$\downarrow$$

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

Then $E^T E = e_1^2 + e_2^2 + \dots + e_N^2$

$$= S_r = \sum e_i^2$$

So, $S_r = E^T E = (Y - ZA)^T (Y - ZA)$

$$= Y^T Y - Y^T Z A - A^T Z^T Y + A^T Z^T Z A$$

We want $\frac{\partial S_r}{\partial a}$ & $\frac{\partial S_r}{\partial b}$. Via some vector witchcraft, this can be compactly represented as

$$\frac{dS_r}{dA} = 0 - Z^TY - Z^TY + 2Z^TZA = 0$$

$$\Longrightarrow \boxed{Z^TZA = Z^TY}$$

That is,

$$A = \underbrace{(Z^TZ)^{-1}Z^TY}_{\text{known}}$$