LECTURE 7

Review problem will now be graded (for attempt). – Week 3.

Q. Find the matrix norm $\|A\|_p$ for $p = \cancel{1}, 2, \cancel{\infty}, \cancel{F}$

where $A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$

A. $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| = \max\{2, 2\} = 2$

$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| = \max\{3, 1\} = 3$

$\|A\|_2 = \left(\max_{i \in [n]} \lambda_i(A^T A)\right)^{1/2}$

$A^T A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ -1 & 1 \end{bmatrix}$

$A^T A - \lambda I = \begin{bmatrix} 5-\lambda & -1 \\ -1 & 1-\lambda \end{bmatrix}$

$\det(A^T A - \lambda I) = (5-\lambda)(1-\lambda) - 1 = 5 - 6\lambda + \lambda^2 - 1 = \lambda^2 - 6\lambda + 4$

$\lambda = \dfrac{6 \pm \sqrt{36 - 4(4)}}{2} = \dfrac{6 \pm \sqrt{20}}{2} = 3 \pm \sqrt{5}.$

$\Rightarrow \|A\|_2 = \sqrt{3 + \sqrt{5}} \approx 2.28825\ldots$

$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} = \sqrt{1 + 4 + 1 + 0} = \sqrt{6} \approx 2.44949$

Prop 3.1 Let $A \in \mathbb{R}^{n \times n}$ and assume A is nonsingular. ⟶ meaning?

Then for any RHS $b \in \mathbb{R}^n$, $Ax = b$ has a unique solution $x_* \in \mathbb{R}^n$.

Pf. If A is nonsingular, then $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^n$ is isomorphism ie

rank $= n$, nullity $= 0$. $\Rightarrow$ col vectors are linearly independent

$\Rightarrow$ form a basis for $\mathbb{R}^n_{\text{target}}$. $\Rightarrow$ unique way to write b with coeffs

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}\begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

(Skipping review material from beginning of Chapter 3)

§3.6 The Least Squares Problem

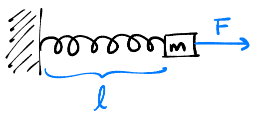
Linear Regression

Suppose you are performing an experiment where you know the relationship b/w two variables is linear. You just want to find the parameters giving the relationship: $y = mx + b$, etc.

eg. Springs and Hooke's law

There will be measurement errors!

variables:



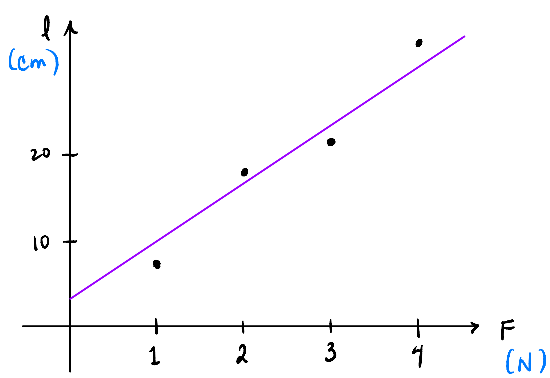
linear relationship: $e + KF = \ell$

params: e, K

($1/K$ is the "spring constant" in Hooke's law, which is derived from diffeqs)

To understand the spring you're using, you perform an experiment

length vs. Force



Overdetermined system: we have more than 2 data points so very likely that our data doesn't actually lie on a single line.

Recall: linear system ⇒ determined if # eqns = # unknowns

In Matrix form:

$$\left.\begin{aligned} e + 1K &= l_1 \\ e + 2K &= l_2 \\ e + 3K &= l_3 \\ e + 4K &= l_4 \end{aligned}\right\} \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} e \\ K \end{bmatrix} = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix}$$

ie.

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} e \\ K \end{bmatrix} - \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(impossible to solve)

Instead, minimize the norm of the residual vector

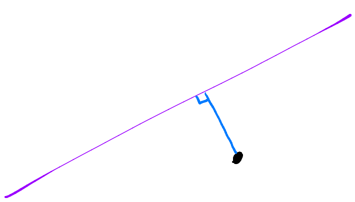
Want to pick e, K so that the purple line

fits the data "the best" ie.

the distance b/w points & the

line is minimized

What is distance b/w point & line?



?

We get to choose our

distance measure.

Goal: Choose line (purple) so that sum-of-squares is minimized

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} e \\ K \end{bmatrix} - \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ l_4 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix}$$

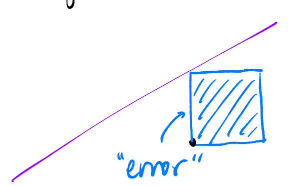
$A \quad x \quad b \quad r$

We want to minimize $\|r\|_2$.

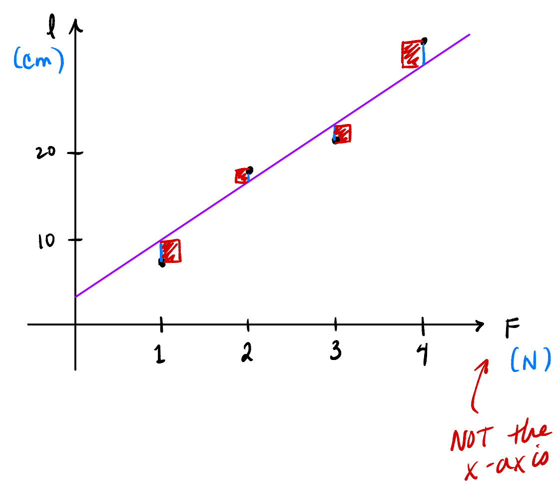
$\|r\|_2 = \sqrt{r_1^2 + r_2^2 + r_3^2 + r_4^2}$ is minimized when $r_1^2 + r_2^2 + r_3^2 + r_4^2$

i.e. the sum of squares is minimized.

Least Squares Method:



Length vs. Force



So now the question becomes:

How do we find $x \in \mathbb{R}^n$ such that the function $\|b - Ax\|_2$ is minimized

or equivalently, $\|b - Ax\|_2^2$ is minimized? x = unknown parameter!

- the unknown x appears linearly in $\min_x \|b - Ax\|_2$

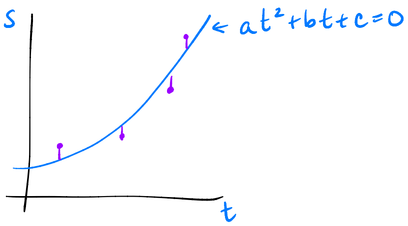
$\Rightarrow$ the <u>linear</u> <u>least squares problem</u>

$*$ $x = \begin{pmatrix} e \\ K \end{pmatrix}$ the parameters!

Yes we are fitting to a line as well but that isn't why it's called the linear least squares problem:

Rmk. It didn't matter that the shape we were trying to fit our data to was a line; we could try to fix our data to

eg. Best fit parabola:



eg. best fit plane in $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}^1$

linear relationships b/w 2 input & 1 output

$$x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

s_i = output from inputting t_i

$$\#\text{ of experiments}\left\{ \begin{array}{c} \begin{array}{ccc} t^2 & t & 1 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix} \end{array} \right. \begin{bmatrix} a \\ b \\ c \end{bmatrix} - \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix}$$

needs to be minimized.

* In this sense, linear least squares problems already cover a lot of regression situations!

eg. Measure gravity constant. (HW03)

Wednesday: How to solve using geom intuition ⇝ normal equations.