Wednesday Nov 6

- Homework 9 due today
- Midterm 2 on Monday
- Midterm 2 Review posted

Example from Monday

System of equations

$$\begin{aligned} x_1 + x_2 &= 5 \\ 2x_1 \qquad &= 2 \\ -x_2 &= -1 \end{aligned}$$

has no solution.

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}}_{b}$$

Can we find an approximate solution? A "good" one?

Least Squares is one definition of good.

<u>Language</u>: Assuming an $m \times n$ matrix A such that $m \geq n$. (So A has more equations than variables.)

- <u>residual</u> $r = Ax - b$ ← difference between output Ax and desired soln b.

- Least squares strategy chooses to minimize

$$\|Ax-b\|^2$$

$\boxed{\text{Ex}}$ $x=(1,1)$ $\left\|\begin{bmatrix}1 & 1\\ 2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}-\begin{bmatrix}5\\2\\-1\end{bmatrix}\right\| = \left\|\begin{bmatrix}2\\2\\-1\end{bmatrix}-\begin{bmatrix}5\\2\\-1\end{bmatrix}\right\|$

$A \quad x \quad - \quad b$

$$= \left\|\begin{bmatrix}-3\\0\\0\end{bmatrix}\right\| = 3$$

$\boxed{\text{Ex}}$ $x=(1,2)$ $\left\|\begin{bmatrix}1 & 1\\ 2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}-\begin{bmatrix}5\\2\\-1\end{bmatrix}\right\| = \left\|\begin{bmatrix}3\\2\\-2\end{bmatrix}-\begin{bmatrix}5\\2\\-1\end{bmatrix}\right\|$

$= \left\|\begin{bmatrix}-2\\0\\-1\end{bmatrix}\right\| = \sqrt{5}$ better than $x=\begin{bmatrix}1\\2\end{bmatrix}$

Can we do better?

- $\hat{x}$ is a solution to the least squares problem if

$$\|A\hat{x}-b\|^2 \le \|Ax-b\|^2 \text{ for all } x$$

- $\hat{x}$ makes the residual as small as possible

- $\hat{x}$ is a least squares approximation of a solution to $Ax=b$

How to minimize $\|Ax-b\|^2$ and find $\hat{x}$?

<u>Answer</u>: Realize that $\|Ax-b\|^2$ is just a function from $\mathbb{R}^n$ to $\mathbb{R}$ and use Calculus III.

Return to our example:

$$\begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

$$A \qquad x \quad = \quad b$$

$$Ax-b = \begin{bmatrix} x_1+x_2-5 \\ 2x_1 \quad -2 \\ -x_1 \; +1 \end{bmatrix} \leftarrow \text{vector}$$

$$* \quad \|Ax-b\|^2 = (x_1+x_2-5)^2+(2x_1-2)^2+(-x_2+1)^2 = f(x_1,x_2)$$

How to find the minimum? Want $\nabla f(x_1,x_2)=0$.

$$\frac{\partial f}{\partial x_1} = 2(x_1+x_2-5)+4(2x_1-2) = 10x_1+2x_2-18 = 0$$

$$\frac{\partial f}{\partial x_2} = 2(x_1+x_2-5)-2(1-x_2) = 2x_1+4x_2-12 = 0$$

$x_1=\frac{4}{3}$, $x_2 = \frac{7}{3}$

$$\left\| A \begin{bmatrix} 4/3 \\ 7/3 \end{bmatrix} - b \right\|^2 = \left\| \begin{bmatrix} 11/3 \\ 8/3 \\ -7/3 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} \right\|^2 = 4$$

So residual is $\underline{\underline{2}}$.

- $\|Ax-b\|^2 \leftarrow$ the objective function.

- take a column view of Ax, then

$$\|Ax-b\|^2 = \Big\| \underbrace{\begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} x_1 + \begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} x_2 + \dots + \begin{bmatrix} | \\ a_n \\ | \end{bmatrix} x_n}_{\text{linear combination of column vectors}} - b \Big\|^2$$

So $\hat{x}$ would be lin. combo closest to b.

- Suppose A has linearly independent columns, the least squares approximation to $Ax = b$ is

$$\hat{x} = (A^TA)^{-1}A^Tb = A^+ b$$

$\uparrow$ pseudo inverse

- On Monday: $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -1 \end{bmatrix}$, $A^+ = (A^TA)^{-1}A^T = \begin{bmatrix} \frac{1}{9} & \frac{4}{9} & \frac{1}{9} \\ \frac{4}{9} & \frac{-2}{9} & \frac{-5}{9} \end{bmatrix}$

So $A^+ b = \begin{bmatrix} \frac{1}{9} & \frac{4}{9} & \frac{1}{9} \\ \frac{4}{9} & \frac{-2}{9} & \frac{-5}{9} \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{5}{9} + \frac{8}{9} - \frac{1}{9} \\ \frac{20}{9} - \frac{4}{9} + \frac{5}{9} \end{bmatrix} = \begin{bmatrix} \frac{12}{9} \\ \frac{21}{9} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ \frac{7}{3} \end{bmatrix}$ ✓

Why is $\hat{x} = A^{\dagger} b$?

component-wise

Now, $$f(x) = \|Ax-b\|^2 = \sum_{i=1}^{m} \Big(\underbrace{\sum_{j=1}^{n} A_{ij} x_j - b_i}_{i^{th} \text{ component of the vector } Ax-b} \Big)^2$$

work ✓

and $$\nabla f(x) = 2A^T(A\hat{x} - b) = 0$$

So $$2A^TA\hat{x} - 2A^Tb = 0$$

$$A^TA\hat{x} = A^Tb$$

A^TA is invertible

$$\hat{x} = (A^TA)^{-1}A^Tb = A^{\dagger}b \checkmark$$