Lecture #17

02/17/2023

Solving tall ($m > n$) linear system:

$$\underbrace{A}_{m \times n} \underbrace{\underline{x}}_{n \times 1} = \underbrace{\underline{b}}_{m \times 1}$$

unknown

known

Example: (Usually tall linear systems have NO (exact) solution)

$$\underbrace{\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}}_{3 \times 2} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{2 \times 1} = \underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}_{3 \times 1} \iff A \underline{x} = \underline{b}$$

$$\begin{aligned}
\Longleftrightarrow \quad 2x_1 &= 1 \implies x_1 = 0.5 \\
-x_1 + x_2 &= 0 \implies x_2 = x_1 = 0.5 \\
2x_2 &= -1 \implies \underbrace{2 \times 0.5}_{1} = -1
\end{aligned}$$

← Impossible !

$\therefore$ There does NOT exist any $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ such that

$$\begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Therefore, it makes sense to compute vector $\underline{x}$ such that the error $\underbrace{\| A\underline{x} - \underline{b} \|_2}_{\text{is called error / "residual"}}$ is minimized.

IDEA:

Compute
$$\hat{\underline{x}} := \underset{\underline{x} \in \mathbb{R}^n}{\operatorname{argmin}} \| A\underline{x} - \underline{b} \|_2 = \underset{\underline{x} \in \mathbb{R}^n}{\operatorname{argmin}} \| A\underline{x} - \underline{b} \|_2^2$$

This problem is called the "least squares problem"

The solution $\hat{\underline{x}}$ of the least squares problem, in general, will fail to satisfy the exact equality:

$A\hat{\underline{x}} = \underline{b}$.

<u>Example:</u> (same as the previous 3X2 example)

The minimization problem, for this example, becomes:

$$\underset{\underline{x} \in \mathbb{R}^2}{\operatorname{argmin}} \left\| \underbrace{\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\underline{x}} - \underbrace{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}_{\underline{b}} \right\|_2^2$$

$$\Leftrightarrow \underset{x_1, x_2}{\operatorname{argmin}} \left\{ (2x_1 - 1)^2 + (-x_1 + x_2)^2 + (2x_2 + 1)^2 \right\}$$

So to minimize the expression within curly braces, we set:

$$\left.\begin{aligned} \frac{\partial}{\partial x_1}\{\cdot\} &= 0 \\ \frac{\partial}{\partial x_2}\{\cdot\} &= 0 \end{aligned}\right\} \Rightarrow \underline{\hat{x}} = \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix}.$$

Now, the error: $A\underline{\hat{x}} - \underline{b}$

$$= \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2/3 - 1 \\ -1/3 - 1/3 - 0 \\ -2/3 + 1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ -2/3 \\ +1/3 \end{pmatrix}$$

$$\Rightarrow \|A\underline{\hat{x}} - \underline{b}\|_2^2 = \left(-\tfrac{1}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(\tfrac{1}{3}\right)^2$$
$$= \tfrac{2}{3}.$$

Solution of the least squares problem in general:

Claim: The vector $\underline{\hat{x}}$ solving the least squares problem, solves:

square linear system

$$\underbrace{(A^T A)}_{n \times n} \underline{\hat{x}} = \underbrace{\underset{n \times m}{A^T} \underset{m \times 1}{\underline{b}}}_{n \times 1}$$

called "normal equation"

Please review the linear algebra recap notes in CANVAS File Section: folder: "Supplementary Notes" → LinearAlgebraRecap.

linear independence, left inverse, right inverse, inverse, pseudo-inverse, QR factorization.

please review if and when needed

If A is left-invertible

$$\Updownarrow$$

A has linearly independent columns

$$\Updownarrow$$

$A^T A$ is invertible / non-singular

Then

$$\underline{\hat{x}} = \underbrace{(A^T A)^{-1} A^T}_{A^+} \underline{b}$$

$$= A^+ \underline{b}$$

A^+ ← pseudo-inverse of A.

This $\underline{\hat{x}}$ is the unique solution for the least sq. problem.