Linear algebra for AI and ML

(September- 16)

Least squares:

$$Ax = b$$

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^{n}$

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^{n}$

find $x \in \mathbb{R}^{m}$ s.t. $Ax = b$

- Existence & uniqueness for $Ax=b$.
- Computation of solution(s) QR, LU/Gaussian.
- Sensitivity analysis

$A \in \mathbb{R}^{n\times n}$

Assumption: Columns of A are linearly independent.

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

square & A invertible.

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

tall A; $m \geq n$ and columns of A are lin. ind.

$Ax = b$

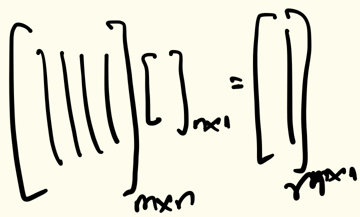
Clearly if $b \in$ col.span(A)

$\Rightarrow$ b can be written as a unique linear combination of columns of A.

What if, $b \notin$ col.span(A)

then clearly, there does <u>NOT</u> exist any $x \in \mathbb{R}^n$ s.t. $Ax = b$

$A \in \mathbb{R}^{m \times n}$
$x \in \mathbb{R}^n$
$b \in \mathbb{R}^m$



There is no solution x to $Ax = b$

$$\underset{x}{\arg\min} \underbrace{\|Ax - b\|_2} = \hat{x}$$

$A \in \mathbb{R}^{m \times n}$, $\boxed{x \in \mathbb{R}^n}$

$b \in \mathbb{R}^m$

$Ax - b \in \mathbb{R}^m$

$r = Ax - b$ = residual vector

$r \in \mathbb{R}^m$

$$r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} = \underbrace{\begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}}_{A} \begin{bmatrix} \\ \\ \end{bmatrix} - \begin{bmatrix} \cdot \\ \cdot \\ \\ \end{bmatrix}$$

$$\|Ax - b\|_2^2 = r_1^2 + r_2^2 + \dots + r_m^2$$

$$= (a_1^T x - b_1)^2 + (a_2^T x - b_2)^2 + \dots + (a_m^T x - b_m)^2$$

$f(a_1^T)$ $f(a_2^T)$ $f(a_m^T)$

there does NOT exist x s.t. $Ax = b$

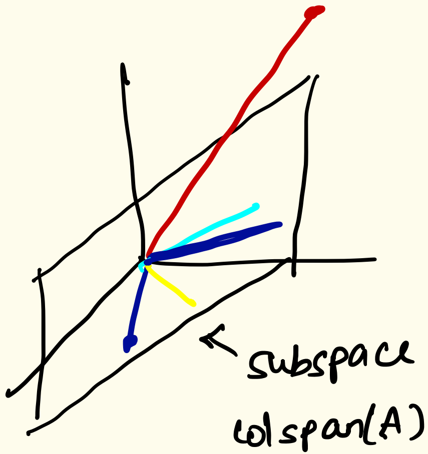
$(b \notin \text{colspan}(A))$

$$Ax = \begin{bmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 A_1 + x_2 A_2 + \cdots + x_n A_n$$

$$\|Ax - b\|_2^2 = \| \overbrace{(x_1 A_1 + x_2 A_2 + \cdots + x_n A_n)} - b\|_2^2$$

$$\hat{x} = \text{arg}\boxed{\min_{x} \|Ax - b\|_2^2}$$

$$A\hat{x} = \hat{x}_1 A_1 + \hat{x}_2 A_2 + \cdots + \hat{x}_n A_n$$



Solution: Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, columns of A are linearly indep.

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$$

Let $f(x) = \|Ax - b\|_2^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \underbrace{A_{ij} x_j} - b_i \right)^2$

where $A = [A_{ij}]$: A_{ij} : $(i,j)^{th}$ entry of A.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix}$$

$$\underbrace{(\nabla f(x))_k} = \frac{\partial}{\partial x_k} f(x)$$

$$(\nabla f(x))_k = \frac{\partial}{\partial x_k} f(x)$$

$$= \frac{\partial}{\partial x_k} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij} x_j - b_i \right)^2$$

$$= \sum_{i=1}^{m} 2 \left(\sum_{j=1}^{n} A_{ij} x_j - b_i \right) (A_{ik})$$

$$= \sum_{i=1}^{m} 2 \underbrace{(A^T)}_{B}{}_{ki} \underbrace{(Ax - b)}_{z}{}_{i}$$

$$\begin{bmatrix} Bz \end{bmatrix} \leftarrow k^{th}$$

$$(\nabla f(x))_k = 2 \left[A^T (Ax - b) \right]_k$$

$$\Rightarrow \quad \nabla f(x) = 2A^T (Ax - b)$$

Let $\hat{x}$ be any minimizer of $\|Ax-b\|_2^2$.

$$\hat{x} \in \underset{x}{\operatorname{argmin}} \|Ax-b\|_2^2$$

$$\Rightarrow \nabla f(\hat{x}) = 0$$

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

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

normal eq$^{\underline{ns}}$ ($n \times n$: square system)

$\because$ cols of A are linearly indep. (A^TA) is invertible.

$$\Rightarrow \boxed{\hat{x} = (A^TA)^{-1}A^Tb}$$

<u>Recall</u>: $(A^TA)^{-1}A^T$ is a special left inverse of A. (pseudo inverse)

Denote the pseudo-inverse of A as A^{+}.

$$A^{+} = (A^{T}A)^{-1} A^{T}$$

and the LS solution to $Ax = b$ is

$$\boxed{\hat{x} = A^{+} b}$$

Direct verification of LS solution.

Suppose $\hat{x} = (A^{T}A)^{-1} A^{T} b$ is a solution to the LS problem.

$$\left(\min_{x} \|Ax - b\|_2^2 \right) : \text{LS problem}$$

Verify !!

$$\|Ax-b\|_2^2 = \|(Ax - A\hat{x}) + (A\hat{x} - b)\|_2^2$$

$$= \|Ax - A\hat{x}\|_2^2 + \|A\hat{x} - b\|_2^2$$

$$+ 2(Ax - A\hat{x})^T(A\hat{x} - b)$$

$$(\|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2 + 2u^Tv)$$

Observe: $$2(Ax - A\hat{x})^T(A\hat{x} - b) = 2(x - \hat{x})^T \boxed{A^T(A\hat{x} - b)}$$

$$= 2(x - \hat{x})^T\, 0$$

$$= 0$$

$$\boxed{\|Ax-b\|_2^2} \geq \quad = \underbrace{\boxed{\|A(x - \hat{x})\|_2^2}}_{\geq 0} + \boxed{\|A\hat{x} - b\|_2^2}$$