Linear algebra for AI & ML

$Ax = b \qquad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m$

$$r = (Ax - b) \in \mathbb{R}^m \qquad r = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix}$$

$$\min_{x \in \mathbb{R}^n} \|r\|_2^2 = \min_{x \in \mathbb{R}^n} r_1^2 + r_2^2 + \cdots + r_m^2$$

$$\boxed{Ax = b}$$

$$\|Ax - b\|$$

$$A = [a_1 \cdots a_n] \Rightarrow Ax = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

$$\|Ax - b\|_2^2 = \|a_1 x_1 + a_2 x_2 + \cdots + a_n x_n - b\|_2^2$$

$$A = \begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix} \Rightarrow Ax = \begin{bmatrix} \tilde{a}_1^T x \\ \vdots \\ \tilde{a}_m^T x \end{bmatrix}$$

$$\|Ax - b\|_2^2 = \left\| \begin{bmatrix} \tilde{a}_1^T x - b_1 \\ \vdots \\ \tilde{a}_m^T x - b_m \end{bmatrix} \right\|_2^2$$

$$\|Ax-b\|_2^2 = \sum_{i=1}^{m} \left(\tilde{a}_i^T x - b_i\right)^2 = \sum_{i=1}^{m} r_i^2$$

standing assumption: $A \in \mathbb{R}^{m \times n}$ such that columns of A are linearly independent.

$$f = \|r\|_2^2 = \|Ax-b\|_2^2$$

we want to $\min_{x \in \mathbb{R}^n} f(x) = \min_{x \in \mathbb{R}^n} \|Ax-b\|_2^2$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 0$$

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

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

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

$$= 2\left(A^T(Ax-b)\right)_k$$

$$\nabla_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 2A^T(Ax-b)$$

$\hat{x}$ which minimizes $\|Ax-b\|_2^2$ satisfies

$$\nabla_x f(\hat{x}) = 0$$

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

$$\Rightarrow A^TA\hat{x} = A^Tb$$ normal eq^{ny}:

$$\Rightarrow \hat{x} = (A^TA)^{-1}A^T b$$

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

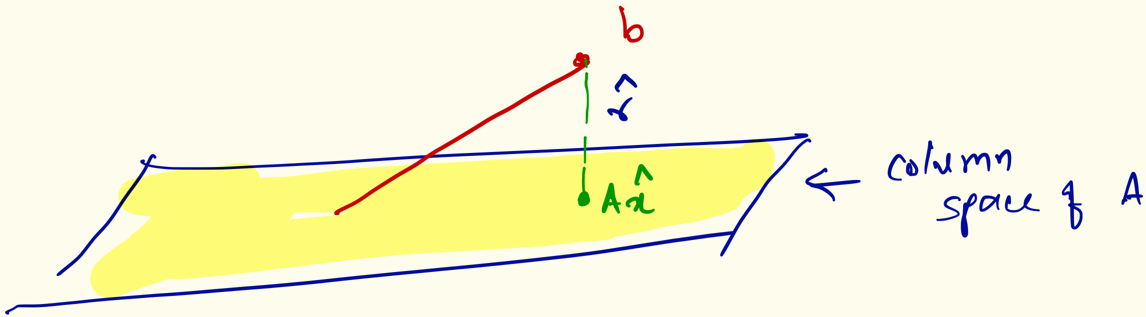
Recall: $Ax=b$ & left inverse exists.

$CA = I$

$Ax = b$

set $\hat{x} = Cb$

$Ax = b$



$\hat{r}$: residual vector at $\hat{x}$

$\quad \parallel$

$A\hat{x} - b$

$\hat{r}$ is orthogonal to the column space of A.

For any $z \in \mathbb{R}^n$, $\quad Az \perp \hat{r}$

$$\Rightarrow (Az)^T \hat{r} = 0$$

$$\Rightarrow \quad z^T A^T (A\hat{x} - b) = 0 \qquad \forall z \in \mathbb{R}^n$$

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

$$\Rightarrow \quad A^T A \hat{x} = A^T b$$

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

Find $\hat{x}$ s.t. $\|A\hat{x} - b\|_2^2$ is minimum.

$$\underbrace{\begin{bmatrix} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix}}_{\tilde{A}} \begin{bmatrix} \\ \\ \end{bmatrix}_{\hat{x}_{m+1}} \qquad \begin{bmatrix} \\ \\ \\ \end{bmatrix}_{\tilde{b}}$$

$\tilde{a}_{m+1}^T$ $\qquad$ b_{m+1}

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

$$= \left(\begin{bmatrix} A \\ \tilde{a}_{m+1}^T \end{bmatrix}^T \begin{bmatrix} A \\ \tilde{a}_{m+1}^T \end{bmatrix} \right)^{-1} \begin{bmatrix} A \\ \tilde{a}_{m+1}^T \end{bmatrix}^T \begin{bmatrix} b \\ b_{m+1} \end{bmatrix}$$

$$= \left(\begin{bmatrix} A^T & \tilde{a}_{m+1} \end{bmatrix} \begin{bmatrix} A \\ \tilde{a}_{m+1}^T \end{bmatrix} \right)^{-1} \begin{bmatrix} A^T & \tilde{a}_{m+1} \end{bmatrix} \begin{bmatrix} b \\ b_{m+1} \end{bmatrix}$$

$$\hat{x}_{m+1} = (A^T A + \tilde{a}_{m+1} \tilde{a}_{m+1}^T)^{-1} (A^T b + \tilde{a}_{m+1} b_{m+1})$$

$$\hat{x} = (A^T A)^{-1} A^T b$$

Matrix inversion lemma

$$(A + UCB)^{-1} = A^{-1} - A^{-1}U(C^{-1} + BA^{-1}U)^{-1}BA^{-1}$$

$$(A^TA + \tilde{a}_{m+1}\tilde{a}_{m+1}^T)^{-1} = (A^TA)^{-1} - (A^TA)^{-1}\tilde{a}_{m+1}(1 + \tilde{a}_{m+1}^T\tilde{a}_{m+1})^{-1}\tilde{a}_{m+1}^T(A^TA)^{-1}$$