Least - Squares Problems

Q/ : If $A\underline{x} = \underline{b}$ does not admit a solution what is the best approximate solution?

<u>Definition</u> A - $m \times n$ matrix and $\underline{b}$ in $\mathbb{R}^m$, a <u>least-squares</u> solution to $A\underline{x} = \underline{b}$ is an $\hat{\underline{x}}$ in $\mathbb{R}^n$ such that

$$\|\underline{b} - A\hat{\underline{x}}\| \leq \|\underline{b} - A\underline{x}\| \quad \text{for all } \underline{x} \text{ in } \mathbb{R}^n.$$

called least-squares error

<u>Remark</u>

If $A\underline{x} = \underline{b}$ is consistent then we can choose $\hat{\underline{x}}$ any solution, ie $A\hat{\underline{x}} = \underline{b}$. If not then $\hat{\underline{x}}$ is the "best approximate solution" to $A\underline{x} = \underline{b}$.

Recall :

1/ $W \subset \mathbb{R}^m$ a subspace, $\underline{b}$ in $\mathbb{R}^m$ and $\hat{\underline{b}}$ in W, then

$$\|\underline{b} - \hat{\underline{b}}\| \leq \|\underline{b} - \underline{w}\| \text{ for all } \underline{w} \text{ in } W \iff \hat{\underline{b}} = \mathrm{Proj}_W(\underline{b}) \iff \underline{b} - \hat{\underline{b}} \text{ in } W^\perp$$

2/ $\mathrm{Col}(A) = \{ A\underline{x} \text{ such that } \underline{x} \text{ in } \mathbb{R}^n \}$

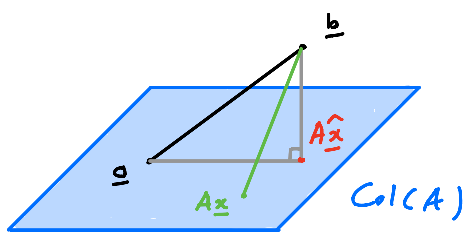
Hence

in Col(A)

$$\|\underline{b} - A\hat{\underline{x}}\| \leq \|\underline{b} - A\underline{x}\| \text{ for all } \underline{x} \text{ in } \mathbb{R}^n \iff A\hat{\underline{x}} = \underbrace{\mathrm{Proj}_{\mathrm{Col}(A)}(\underline{b})}_{=\hat{\underline{b}}}$$

$$\iff \underline{b} - A\hat{\underline{x}} \text{ in } (\mathrm{Col}(A))^\perp$$

<u>Picture in $\mathbb{R}^3$</u> :



$A = (\underline{a}_1 \cdots \underline{a}_n) \Rightarrow \mathrm{Col}(A) = \mathrm{Span}(\underline{a}_1, \ldots \underline{a}_n)$

$\underline{b} - A\hat{\underline{x}}$ in $(\mathrm{Col}(A))^{\perp} \Leftrightarrow \underline{a}_i \cdot (\underline{b} - A\hat{\underline{x}}) = 0$ for all i

$\left(A^T = \begin{pmatrix} \underline{a}_1^T \\ \vdots \\ \underline{a}_n^T \end{pmatrix}\right)$ $\Leftrightarrow \underline{a}_i^T(\underline{b} - A\hat{\underline{x}}) = 0$ for all i

$\Leftrightarrow A^T(\underline{b} - A\hat{\underline{x}}) = \underline{0}$

$\Leftrightarrow A^T A \hat{\underline{x}} = A^T \underline{b}$

__Remark__ This also follows from the fact that $(\mathrm{Col}(A))^{\perp} = \mathrm{Nul}(A^T)$

__Conclusion__

$\hat{\underline{x}}$ a least-squares solution to $A\underline{x} = \underline{b} \Leftrightarrow A\hat{\underline{x}} = \mathrm{Proj}_{\mathrm{Col}(A)}(\underline{b})$

$\Leftrightarrow \underline{b} - A\hat{\underline{x}}$ in $(\mathrm{Col}(A))^{\perp}$

$\Leftrightarrow \underline{b} - A\hat{\underline{x}}$ in $\mathrm{Nul}(A^T)$

$\Leftrightarrow \hat{\underline{x}}$ a solution to

$A^T A \underline{x} = A^T \underline{b}$.

Called the normal equations for $A\underline{x} = \underline{b}$

__Example__

$A = \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}, \quad \hat{\underline{x}} = ?$

$\Rightarrow A^T A = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 6 & -11 \\ -11 & 22 \end{pmatrix}$

$A^T \underline{b} = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ 11 \end{pmatrix}$

$$\left(\begin{array}{cc|c} 6 & -11 & -4 \\ -11 & 22 & 11 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 6 & -11 & -4 \\ 1 & -2 & -1 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & -2 & -1 \\ 6 & -11 & -4 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & -2 & -1 \\ 0 & 1 & 2 \end{array}\right)$$

$$\downarrow$$

$$\hat{\underline{x}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \Leftarrow \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array}\right)$$

least squares solution to $A\underline{x} = \underline{b}$

Remark

$A\underline{x} = \underline{b}$ has a unique least squares solution

$\iff A\underline{x} = \hat{\underline{b}}$ has a unique solution

$\iff$ Columns of A are L.I.

$\iff A^T A$ invertible

In this case $\quad \underline{\hat{x}} = (A^TA)^{-1}A^T\underline{b}$

Recall: A – $m\times n$ with L.I. columns $\Rightarrow A = \underset{\substack{m\times n \text{ with} \\ \text{orthonormal columns}}}{Q}\ \underset{\substack{n\times n \\ \text{upper} \\ \text{triangular}}}{R}$

$$\Rightarrow \quad A^T = R^TQ^T \quad \Rightarrow \quad A^TA = R^TQ^TQR = R^TR$$

$$\Rightarrow \quad \underline{\hat{x}} = \overset{A^TA}{\overset{\parallel}{(R^TR)}}{}^{-1}\overset{A^T}{\overset{\parallel}{(R^TQ^T)}}\underline{b} = R^{-1}(R^T)^{-1}R^TQ^T\underline{b} = R^{-1}Q^T\underline{b}$$

Conclusion $\quad A$ – $m\times n$ matrix with L.I. columns, $\underline{b}$ in $\mathbb{R}^m$

$A = \underset{\substack{m\times n \text{ with} \\ \text{orthonormal columns}}}{Q}\ \underset{\substack{n\times n \\ \text{upper} \\ \text{triangular}}}{R} \quad \Rightarrow$ Unique Least-squares solution is $\underline{\hat{x}} = R^{-1}Q^T\underline{b}$

$(\iff R\underline{\hat{x}} = Q^T\underline{b})$

Example

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} 3 \\ 5 \\ 7 \\ -3 \end{pmatrix}$$

Gram-Schmidt $\Rightarrow$
$$A = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{pmatrix}\begin{pmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\Rightarrow \quad Q^T\underline{b} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{pmatrix}\begin{pmatrix} 3 \\ 5 \\ 7 \\ -3 \end{pmatrix} = \begin{pmatrix} 6 \\ -6 \\ 4 \end{pmatrix}$$

$\underline{\hat{x}}$ is the unique solution to the linear system

$$\left(\begin{array}{ccc|c} 2 & 4 & 5 & 6 \\ 0 & 2 & 3 & -6 \\ 0 & 0 & 2 & 4 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 10 \\ 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 2 \end{array}\right) \Rightarrow \underline{\hat{x}} = \begin{pmatrix} 10 \\ -6 \\ 2 \end{pmatrix}$$