LAST TIME: LS solutions of $A\vec{x} = \vec{b}$ (an inconsistent system)

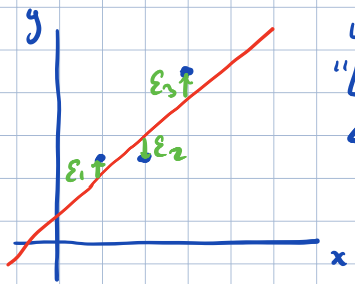
are solutions $\hat{\vec{x}}$ of $\boxed{A^T A \hat{\vec{x}} = A^T \vec{b}}$ – "normal equations"

$A\hat{\vec{x}}$ – LS approximation

$\|\vec{b} - A\hat{\vec{x}}\|$ – LS error

# Application of LS solutions: LS approximation.

**Ex:** given data points $(1,2)$, $(2,2)$, $(3,4)$ find the line $y = a + bx$ which is the "best fit" for the points.




↑ want to have the "LS error" $\Delta = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2}$ as small as possible

$$\left.\begin{aligned} \varepsilon_1 &= 2 - (a + b \cdot 1) \\ \varepsilon_2 &= 2 - (a + b \cdot 2) \\ \varepsilon_3 &= 4 - (a + b \cdot 3) \end{aligned}\right\} \text{errors}$$

So: we want a LS solution of

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}}_{\vec{b}}, \quad \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \vec{b} - A\vec{x}$$

↑ "error vector" – trying to minimize its norm

**normal eq.:** $\underbrace{A^T A}\hat{\vec{x}} = \underbrace{A^T \vec{b}}$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \qquad \begin{bmatrix} 8 \\ 18 \end{bmatrix}$$

Aug. mat.: $\left[\begin{array}{cc|c} 3 & 6 & 8 \\ 6 & 14 & 18 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 2 & 8/3 \\ 0 & 2 & 2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 2/3 \\ 0 & 1 & 1 \end{array}\right] \Rightarrow \hat{\vec{x}} = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$ ← a, ← b

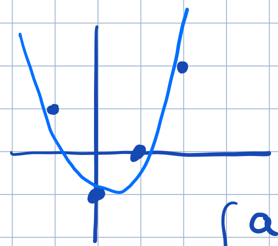
$\Rightarrow$ best fitting line: $y = \frac{2}{3} + x$

("least squares approximating line")

**LS error:** $\Delta = \|\vec{b} - A\hat{\vec{x}}\| = \left\| \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 5/3 \\ 8/3 \\ 11/3 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 1/3 \\ -2/3 \\ 1/3 \end{bmatrix} \right\| = \frac{\sqrt{6}}{3} = \sqrt{\frac{2}{3}}$

$$A\hat{\vec{x}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2/3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 8/3 \\ 11/3 \end{bmatrix}$$

**Ex:** Find the parabola that gives the best LS approximation to the points $(-1,1)$, $(0,-1)$, $(1,0)$, $(2,2)$



**Sol:** $y = a + bx + cx^2$ - approximating parabola

substituting the given points, we get:

$$\begin{cases} a - b + c = 1 \\ a = -1 \\ a + b + c = 0 \\ a + 2b + 4c = 2 \end{cases} \quad \text{or:} \quad \underbrace{\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}}_{\vec{b}}$$

**LS sol:** normal eq. $\underbrace{\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix}}_{A^TA} \hat{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 3 \\ 9 \end{bmatrix}}_{A^T\vec{b}} \rightsquigarrow \hat{\vec{x}} = \begin{bmatrix} -7/10 \\ -3/5 \\ 1 \end{bmatrix}$

So, the least squares approximating parabola: $\boxed{y = -\frac{7}{10} - \frac{3}{5}x + x^2}$

**LS error:** $\Delta = \|\vec{b} - \underbrace{A\hat{\vec{x}}}_{\begin{bmatrix} 3/10 \\ -7/10 \\ -3/10 \\ 21/10 \end{bmatrix}}\| = \left\| \begin{bmatrix} -1/10 \\ -3/10 \\ 3/10 \\ -1/10 \end{bmatrix} \right\| = \frac{\sqrt{20}}{10} = \frac{1}{\sqrt{5}}$

---

**Another way to construct LS solutions**

**Thm** If $A$ is $m \times n$ matrix with lin. indep. columns and $A = QR$ - QR factorization, then for each $\vec{b} \in \mathbb{R}^m$ the LS sol. of $A\vec{x} = \vec{b}$ is: $\boxed{\hat{\vec{x}} = R^{-1}Q^T\vec{b}}$.

**Idea:** normal eq. $\underbrace{A^TA}_{R^T\underbrace{Q^TQ}_{I}R}\hat{\vec{x}} = \underbrace{A^T\vec{b}}_{R^TQ^T\vec{b}} \iff R^TR\vec{x} = R^TQ^T\vec{b}$

$\underset{(R^T)^{-1}\cdot}{\iff} R\vec{x} = Q^T\vec{b}$

$\underset{R^{-1}\cdot}{\iff} \vec{x} = R^{-1}Q^T\vec{b}$

- for $A$ an $m\times n$ matrix with lin. indep. columns, the unique LS sol. of $A\vec{x}=\vec{b}$ is $\hat{x}=\underbrace{(A^TA)^{-1}A^T}_{\substack{A^+ \\ n\times m}}\vec{b}$

<u>def</u> For $A$ an $m\times n$ matrix with lin. indep. columns, the $n\times m$ matrix $A^+ := (A^TA)^{-1}A^T$ is called the <u>pseudoinverse</u> of $A$.

<u>Ex:</u> for $A=\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix}$, find the pseudoinverse

<u>Sol:</u> $A^TA=\begin{bmatrix}3&6\\6&14\end{bmatrix} \rightsquigarrow (A^TA)^{-1}=\frac{1}{6}\begin{bmatrix}14&-6\\-6&3\end{bmatrix} \rightarrow$

$\rightarrow A^+=\frac{1}{6}\begin{bmatrix}14&-6\\-6&3\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&3\end{bmatrix}=\frac{1}{6}\begin{bmatrix}8&2&-4\\-3&0&3\end{bmatrix}$

<u>Properties:</u> 1) for $A$ $n\times n$ (with lin. indep. columns), $A^+=A^{-1}$

2) $A^+A=I$

3) $AA^+$ is the matrix of projection onto $W=\text{col}(A)\subset\mathbb{R}^m$, i.e. $\text{proj}_W(\vec{v})=AA^+\vec{v}$ for any $\vec{v}\in\mathbb{R}^m$

4) if $A=QR$ - QR-factorization, then $A^+=R^{-1}Q^T$ and $AA^+=QQ^T$ - matrix of projection

↑ this is a <u>symmetric</u> matrix

<u>Ex:</u> find the matrix of $\text{proj}_W:\mathbb{R}^3\to\mathbb{R}^3$, $W=\text{span}\left(\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\2\\3\end{bmatrix}\right)$

<u>Sol:</u> $W=\text{col}\Big(\underbrace{\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix}}_{A}\Big)$

matrix of projection: $AA^+=\begin{bmatrix}1&1\\1&2\\1&3\end{bmatrix}\underbrace{\frac{1}{6}\begin{bmatrix}8&2&-4\\-3&0&3\end{bmatrix}}_{\text{computed before}}$

$=\frac{1}{6}\begin{bmatrix}5&2&-1\\2&2&2\\-1&2&5\end{bmatrix}$