EECS16A DIS 7A

Today's topics

- A numerical least squares example
- Special cases of least squares : orthonormal matrices

Skills

- How to use derivatives to derive least squares solution

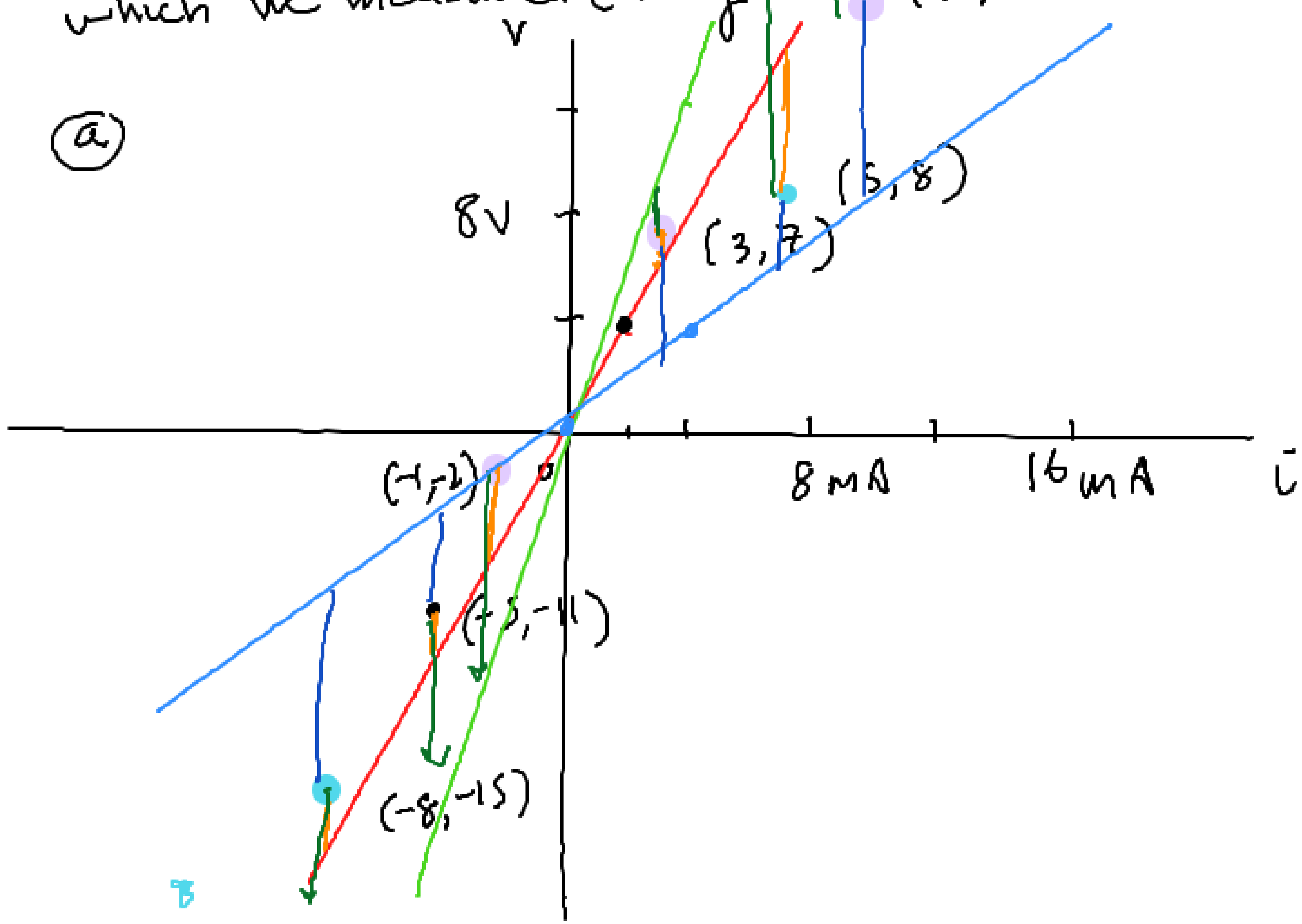
1 Trying to find resistance R

(i_j, v_j)

currents, voltages which we measured (noisy)

Find a best estimate for R

ⓐ



Test	i (mA)	V (V)
1	10	21
2	3	7
3	-1	-2
4	5	8
5	-8	-15
6	-5	-11

Q: Is there a single R?

A: No: they're all not on a line

(b) $\vec{I} = \begin{bmatrix} \vdots \\ \vdots \end{bmatrix}$ $\vec{V} = \begin{bmatrix} \vdots \\ \vdots \end{bmatrix}$

$\vec{V} = R\vec{I}$ • $\vec{V} \in \text{span}\{\vec{I}\}$ for R to be a solution

• for all comp. $\underbrace{i_2 R = V_2}_{\text{this relationship holds}}$

$\vec{V} \notin \text{span}\{\vec{I}\}$

→ Can't get a solution

→ Inconsistent system

(c) Cost - quantity we want to minimize by choosing R

⇒ cost associated with least squares technique

→ $\text{cost}_1(R) = \sum_{j=1}^{6} (V_j - R i_j)^2$ ⇐ minimize

→ $\text{cost}_2(R) = \sum_{j=1}^{6} (V_j - R i_j)$

Difference? First is squared

$V_2 - R i_2 > 0$ and $V_3 - R i_3 < 0$

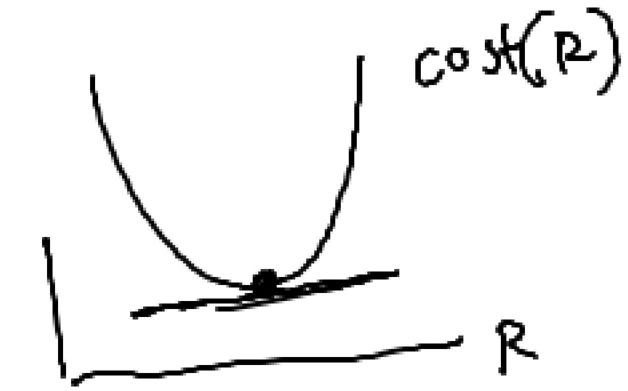
↳ could lower cost in cost_2

could still be large in mag.

Cost_1 is better - cares about abs. difference

(e) How to minimize cost?

↳ Last discussions - used a geometric argument



This discussion: calculus

Minimize cost by choosing R. How? Find location where

$$\frac{d}{dR}\,\text{cost}(R) = 0$$

(d) $\text{cost}(R) \stackrel{?}{=} \langle \vec{V} - \vec{I}R, \vec{V} - \vec{I}R \rangle$

$\text{cost}(R) = \sum_{j=1}^{6} (v_j - R i_j)^2$ $\qquad e_j = v_j - R i_j$

$\vec{e} = \vec{V} - R\vec{i} \leftarrow$

$$\vec{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} v_1 - R i_1 \\ v_2 - R i_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \end{bmatrix} - R \begin{bmatrix} i_1 \\ \vdots \end{bmatrix} = \vec{V} - R\vec{I}$$

Hint: $\sum_{j=1}^{6} (e_j)^2 = \|\vec{e}\|^2$

$$= \|\vec{V} - R\vec{i}\|^2$$

$$= \langle \vec{V} - R\vec{i}, \vec{V} - R\vec{i} \rangle \checkmark$$

(e) Find R s.t. $\frac{d}{dR}\text{cost}(R) = 0$. This is the R that minimizes

$$\frac{d}{dR}\left[\sum_{j=1}^{6}(v_j - Ri_j)^2\right] = 0$$

$\frac{d}{dR}$ dist over +

$$\sum_{j=1}^{6}\frac{d}{dR}(v_j - Ri_j)^2 = 0$$

chain rule

$$\sum_{j=1}^{6} 2\underbrace{(v_j - Ri_j)}_{(\vec{v}-R\vec{I})^T(-\vec{I})}(-i_j) = 0$$

re-express as inner prod.

$$\cancel{2}\langle \vec{V} - R\vec{I}, -\vec{I}\rangle = 0$$

$$-\langle \vec{V}, \vec{I}\rangle + R\langle \vec{I}, \vec{I}\rangle = 0$$

$$\boxed{R = \frac{\langle I, \vec{V}\rangle}{\langle \vec{I}, \vec{I}\rangle} = \frac{\langle \vec{v}, \vec{I}\rangle}{\|\vec{I}\|^2}}$$

$$\frac{d}{dR}\langle \vec{V} - RI, \vec{v} - R\vec{I}\rangle$$

$$\frac{d}{dR}\left(\langle \vec{V}, \vec{V}\rangle - 2R\langle \vec{V}, \vec{I}\rangle + R^2\langle \vec{I}, \vec{I}\rangle\right)$$

$$0 - 2\langle \vec{V}, \vec{I}\rangle + 2R\langle \vec{I}, \vec{I}\rangle = 0$$

(f) $R = \frac{\langle \vec{V}, \vec{I} \rangle}{\|\vec{I}\|^2} = \frac{v_1 i_1 + v_2 i_2 + \cdots + v_6 i_6}{i_1^2 + i_2^2 + \cdots i_6^2}$

$$= \frac{(21V)(10mA) + (7V)(3mA) + \cdots}{(10mA)^2 + (3mA)^2 + \cdots}$$

$$= \boxed{2k\Omega}$$

$\text{cost}(3k\Omega) = \sum_{i=1}^{6} (\_-3k\_)^2 = \text{large}$ $> \text{cost}(2k\Omega)$

$\text{cost}(1k\Omega) = \text{large}$

(g) New data: $(2mA, 4V)$ → since on $2k\Omega$ line, estimate doesn't change

if $(2mA, >4V)$ → R increases

$(2mA, <4V)$ → R decreases

$\boxed{2}$ $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$ A is orthonormal

→ columns are orthogonal : $\langle \vec{a}_i, \vec{a}_j \rangle = 0$

→ columns are norm 1 : $\|\vec{a}_i\| = 1$

$\langle \vec{a}_i, \vec{a}_i \rangle = 1$

$\underset{N \times M}{A}$ $\vec{y} \notin \text{span}\{\vec{a}_1, \ldots, \vec{a}_M\}$

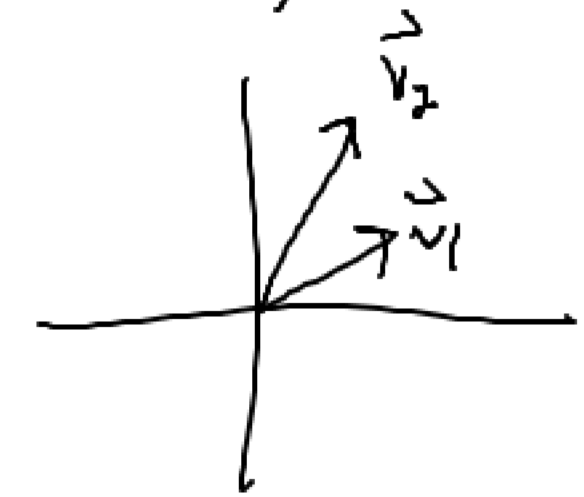
→ $\text{proj}_{C(A)}(\vec{y}) = A(A^TA)^{-1}A^T\vec{y}$ $\text{proj}_{C(A)}(\vec{q}) = A\hat{x}$

(least squares projection formula)

$\hat{x} = (A^TA)^{-1}A^T\vec{y}$

- least squares solution ←

(b) Prove that a square orthonormal matrix has its columns form a basis for $\mathbb{R}^N$



Basis

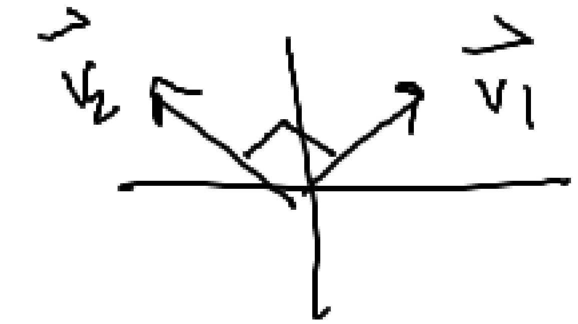
① lin. indep.

② spans space: $\mathbb{R}^N$

$$A = \begin{bmatrix} | & | & & | \\ \vec a_1 & \vec a_2 & \cdots & \vec a_N \\ | & | & & | \end{bmatrix}$$

Orthogonality → lin indep

$A\vec\alpha = \vec 0$ has only $\vec\alpha = \vec 0$



$$\alpha_1 \vec a_1 + \alpha_2 \vec a_2 + \cdots + \alpha_n \vec a_n = \vec 0$$

if only soln. ↑ is $\alpha_i = 0$, lin. indep.

$$\langle \vec a_i, \alpha_1 \vec a_1 + \cdots + \vec a_i \alpha_i + \cdots + \alpha_n \vec a_n \rangle = \langle \vec a_i, \vec 0 \rangle$$

$$\underbrace{\alpha_1 \langle \vec a_i, \vec a_1 \rangle}_{0} + \cdots \alpha_i \langle \vec a_i, \vec a_i \rangle + \cdots = 0$$

$$0 + 0 + \cdots + \alpha_i \cdot 1 + 0 + 0 = 0$$

$\alpha_i = 0$ applies for any i $\quad \vec\alpha = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \vec 0$

(b) Spans $\mathbb{R}^N$

$A\vec{x} = \vec{y} \in \mathbb{R}^N$

$\Rightarrow$ A has lin indep col., A is square

$\vec{x} = A^{-1}\vec{y}$

for every vector in $\mathbb{R}^N$ ($\vec{y}$) we can find it as a lin comb. ($A\vec{x}$)

(c) $\underset{N \times M}{A}$. $\begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_m \\ | & | & & | \end{bmatrix}$ $\updownarrow$ tall

$N \geq M$

orthonormal

$A^T A = I_{M \times M}$

$A^T A = \begin{bmatrix} - & \vec{a}_1^T & - \\ & \vdots & \\ - & \vec{a}_m^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ a_1 & \cdots & a_m \\ | & & | \end{bmatrix}$

$= \begin{bmatrix} \|\vec{a}_1\|^2 & & 0 \\ & \|\vec{a}_2\|^2 & \\ 0 & & \ddots \; \|\vec{a}_m\|^2 \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \end{bmatrix} = I$

off diagonal: $\langle \vec{a}_i, \vec{a}_j \rangle = 0$
$i \neq j$
(bc orthogonal)

(d) If $\underset{N\times M}{A}$ has ~~or~~thonormal col. (c) $A^TA = I_{M\times M}$

$N \geq M$ $\quad proj_{C(A)}\vec{y} = A(A^TA)^{-1}A^T\vec{y}$

$= A\, I_{M\times M}^{-1}\, A^T\vec{y}$

$= A\, I_{M\times M}\, A^T\vec{y}$

$= \underline{AA^T\vec{y}}$ ← Lst. sq proj

$\hat{x} = \underline{A^T\vec{y}}$ ← Lst sq sol

Special case → simplifies formulas