

11

LECTURE 25
 03/26/12

PROPERTY (6.3):

- IF $V = \mathbb{C}^n$ w/ STD. INNER PROD. $\langle \underline{x} | \underline{y} \rangle = \underline{x}^* \underline{y}$, THEN
 $\langle \underline{y} | \underline{y} \rangle = \underline{y}$ AND $\langle \underline{x} | \underline{x} \rangle = \underline{x}^*$.

Q: WHAT IF $V = \mathbb{C}^n$ w/ NONSTANDARD INNER PROD.?

EX. $V = \mathbb{C}^2$, $\langle \underline{x} | \underline{y} \rangle = 2\bar{x}_1 y_1 + 3\bar{x}_2 y_2$.

$\Rightarrow |\underline{y}\rangle = (y_1, y_2)^T \in \mathbb{C}^2$.

$\langle \underline{x} | = (2\bar{x}_1, 3\bar{x}_2)^T \in (\mathbb{C}^2)'$.

- IF V GENERAL INNER PROD. SPACE w/ BASIS $\mathcal{B} = \{b_i\}_{i=1}^n$,

$$\langle \underline{x} | \underline{y} \rangle_{\mathcal{B}} = \left\langle \sum_{i=1}^n a_i \underline{b}_i \mid \sum_{j=1}^n c_j \underline{b}_j \right\rangle$$

$$= \sum_{i=1}^n \bar{a}_i \sum_{j=1}^n c_j \underbrace{\langle \underline{b}_i | \underline{b}_j \rangle}_{\substack{n \times n \text{ HERMITIAN} \\ \text{MATRIX,} \\ \text{CAN } G_{\mathcal{B}}}}$$

$$= \overline{(a_1, \dots, a_n)} G_{\mathcal{B}} (c_1, \dots, c_n)^T$$

$$= \overline{[\underline{x}]_{\mathcal{B}}}^T G_{\mathcal{B}} [\underline{y}]_{\mathcal{B}}$$

WE CAN G_B THE MATRIX, BY DEFINITION,

$$(G_B)_{ij} = \langle \underline{b}_i | \underline{b}_j \rangle$$

SO $G_B^* = G_B$. BY THE EXPRESSION ABOVE, WE HAVE THAT

$$|y\rangle_B = [\underline{y}]_B \in \mathbb{C}^n$$

$${}_B\langle x| = \overline{[\underline{x}]_B}^T G_B \in (\mathbb{C}^n)'$$

NOTE: IF $B = \{\underline{b}_i\}_{i=1}^n$ IS AN ORTHONORMAL BASIS THEN

$$G_B = I \quad \text{AND} \quad {}_B\langle x| = \overline{[\underline{x}]_B}^T = [\underline{x}]_B^*$$

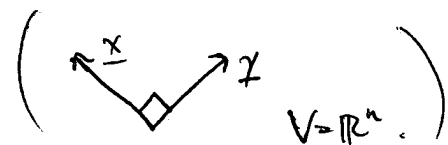
IN THIS CASE, THE INNER PRODUCT IS STANDARD.

Q: WHAT DO WE MEAN BY AN ORTHONORMAL BASIS?

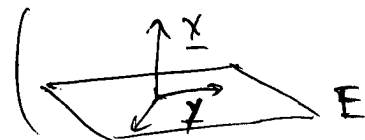
ORTHOGONALITY AND ORTHONORMAL BASES (6.4):

DEF. FOR $\underline{x}, \underline{y} \in V$, \underline{x} IS ORTHOGONAL TO \underline{y} (DENOTE $\underline{x} \perp \underline{y}$)

$$\text{IF } \langle \underline{x} | \underline{y} \rangle = 0.$$



• IF $E \subset V$ SUBSPACE, $\underline{x} \perp E$ IF $\underline{x} \perp \underline{y} \quad \forall \underline{y} \in E$.



Ex Find all vectors orthogonal to $\underline{x} = (i, 2, 1+i)^T$.

$$\langle \underline{x} | \underline{y} \rangle = \underline{x}^* \underline{y} = 0 \Rightarrow (-i, 2, 1-i) (\gamma_1, \gamma_2, \gamma_3)^T = 0.$$

Ex Find all vectors orthogonal to both $\underline{x}_1 = (i, 2, 1+i)^T$

and $\underline{x}_2 = (3, 2-i, i)^T$.

$$\langle \underline{x}_1 | \underline{y} \rangle = \underline{x}_1^* \underline{y} = 0$$

$$\langle \underline{x}_2 | \underline{y} \rangle = \underline{x}_2^* \underline{y} = 0$$

$$\Rightarrow A^* \underline{y} = \underline{0} \quad \text{where}$$

$$A = (\underline{x}_1, \underline{x}_2) \in M_{3,2}(\mathbb{C}).$$

$$= \begin{pmatrix} i & 3 \\ 2 & 2-i \\ 1+i & i \end{pmatrix}.$$

• $\{\underline{b}_1, \dots, \underline{b}_n\}$ orthogonal if $\underline{b}_i \perp \underline{b}_j \quad \forall i \neq j$.

$\{\underline{b}_1, \dots, \underline{b}_n\}$ orthonormal if " " and $\|\underline{b}_i\| = 1 \quad \forall i=1, \dots, n$.

Thm (i) (Pythagorean thm.)

$$\{\underline{b}_i\} \text{ orthogonal} \Rightarrow \|\underline{b}_1 + \dots + \underline{b}_n\|^2 = \|\underline{b}_1\|^2 + \dots + \|\underline{b}_n\|^2.$$

PF (n=2) $\|\underline{x} + \underline{y}\|^2 = \langle \underline{x} + \underline{y} | \underline{x} + \underline{y} \rangle$
 $= \langle \underline{x} | \underline{x} \rangle + \langle \underline{x} | \underline{y} \rangle + \langle \underline{y} | \underline{x} \rangle + \langle \underline{y} | \underline{y} \rangle$
 $= \|\underline{x}\|^2 + \|\underline{y}\|^2.$

(ii) $\{\underline{b}_i\}$ orthogonal $\Rightarrow \{\underline{b}_i\}_{i=1}^n$ are linearly indep.

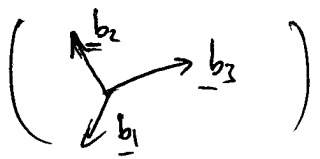
PF $c_1 \underline{b}_1 + \dots + c_n \underline{b}_n = \underline{0} \Rightarrow \|\underline{0}\|^2 = |c_1|^2 \|\underline{b}_1\|^2 + \dots + |c_n|^2 \|\underline{b}_n\|^2$

$$\Rightarrow c_i = 0 \quad \forall i=1, \dots, n.$$

\Rightarrow linearly indep.

EXPANSION IN ORTHOGONAL BASIS :

SUPPOSE $\mathcal{B} = \{ \underline{b}_1, \dots, \underline{b}_n \}$ IS ORTHOGONAL AND IS A BASIS OF V .



$$\underline{x} = a_1 \underline{b}_1 + \dots + a_n \underline{b}_n \iff [\underline{x}]_{\mathcal{B}} = (a_1, \dots, a_n)^T.$$

Q: HOW TO FIND a_i ?

$$\begin{aligned} \underline{A}: \langle \underline{b}_i | \underline{x} \rangle &= \langle \underline{b}_i | a_1 \underline{b}_1 + \dots + a_n \underline{b}_n \rangle \\ &= a_1 \langle \underline{b}_i | \underline{b}_1 \rangle + \dots + a_n \langle \underline{b}_i | \underline{b}_n \rangle \\ &= a_i \langle \underline{b}_i | \underline{b}_i \rangle = a_i \|\underline{b}_i\|^2 \end{aligned}$$

$$\Rightarrow \left[a_i = \frac{\langle \underline{b}_i | \underline{x} \rangle}{\|\underline{b}_i\|^2} \right]$$

THEFORE,

$$\underline{x} = \sum_{i=1}^n \frac{\langle \underline{b}_i | \underline{x} \rangle}{\|\underline{b}_i\|^2} \underline{b}_i. \quad \left(\underline{x} = \sum_{i=1}^n \langle \underline{b}_i | \underline{x} \rangle \underline{b}_i \right)$$

IF $\{ \underline{b}_i \}$ ORTHOGONAL

NOTE: IN BRA-KET NOTATION, THIS IS

$$|\underline{x}\rangle = \sum_{i=1}^n \frac{\langle \underline{b}_i | \underline{x} \rangle}{\langle \underline{b}_i | \underline{b}_i \rangle} |\underline{b}_i\rangle = \left(\sum_{i=1}^n \frac{|\underline{b}_i\rangle \langle \underline{b}_i|}{\langle \underline{b}_i | \underline{b}_i \rangle} \right) |\underline{x}\rangle,$$

SO $\mathbb{I} = \sum_{i=1}^n \frac{|\underline{b}_i\rangle \langle \underline{b}_i|}{\langle \underline{b}_i | \underline{b}_i \rangle}$ IS THE IDENTITY OPERATOR!

Lecture 26
03/28/12

LAST TIME WE SAW THAT (IN BRACKET NOTATION) :

$$|x\rangle = \left(\sum_{i=1}^n \frac{|b_i\rangle\langle b_i|}{\langle b_i|b_i\rangle} \right) |x\rangle$$

I

WHAT IS $P_{b_i} = \frac{|b_i\rangle\langle b_i|}{\langle b_i|b_i\rangle}$?

PROJECTIONS AND GRAM-SCHMIDT PROCESS (6.5-6.6) :

DEF. $\underline{v} \in V$, $\underline{v} \neq \underline{0}$. LET $P_{\underline{v}} \doteq \frac{|\underline{v}\rangle\langle\underline{v}|}{\langle\underline{v}|\underline{v}\rangle}$

BE THE PROJECTION OPERATOR IN DIRECTION OF \underline{v} .

THM. FOR ANY $\underline{x} \in V$, CAN WRITE $\underline{x} = \underline{w} + \underline{y}$
WHERE FOR SOME GIVEN $\underline{v} \neq \underline{0}$, $\underline{w} \parallel \underline{v}$ AND $\underline{y} \perp \underline{v}$.

PR. $\underline{x} = \underline{I} \underline{x} = \left[P_{\underline{v}} + (\underline{I} - P_{\underline{v}}) \right] \underline{x}$
 $= \underbrace{P_{\underline{v}} \underline{x}}_{\text{CALL } \underline{w}} + \underbrace{(\underline{I} - P_{\underline{v}}) \underline{x}}_{\text{CALL } \underline{y}}$

LET $W \subset V$ SUBSPACE WITH ORTHOGONAL BASIS $\{d_1, \dots, d_m\}$. L2

DEF. $W^\perp \doteq$ SPACE OF ALL VECTORS \perp TO W .

• NOTE THAT $\underline{x} \in W^\perp \Leftrightarrow P_W \underline{x} = \underline{0}$
 $\Leftrightarrow \underline{x} \in \text{Ker}(P_W)$

WHERE

$$P_W = \sum_{i=1}^m P_{d_i} = \sum_{i=1}^m \frac{|d_i\rangle\langle d_i|}{\langle d_i | d_i \rangle}$$

IS THE PROJECTION ONTO SUBSPACE W .

• ANY $\underline{x} \in V$ IS $\underline{x} = \underline{w} + \underline{y}$ WHERE $\underline{w} \in W, \underline{y} \in W^\perp$.

SINCE

$$\underline{x} = P_W \underline{x} + \underbrace{(\mathbf{I} - P_W) \underline{x}}_{P_W^\perp \doteq P_{W^\perp}}$$

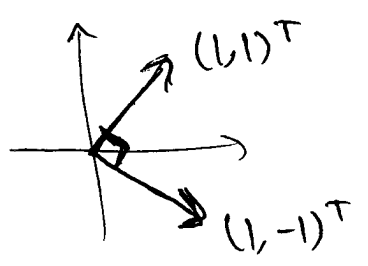
EX. $V = \mathbb{R}^2, W = \text{SPAN} \{ \underbrace{(1, 1)^T}_{\text{CALL THIS } \underline{b}} \}$.
WHAT IS W^\perp ?

$W^\perp =$ ALL $\underline{x} \in V$ S.T. $P_W \underline{x} = \underline{0}$.

$$P_W = \frac{|b\rangle\langle b|}{\langle b | b \rangle} = \frac{(1, 1)^T (1, 1)}{\|(1, 1)^T\|^2} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

$$\Rightarrow \text{Ker}(P_W) = \text{SPAN} \{ (1, -1)^T \}$$

THIS MAKES SENSE, AS WE SEE BY A PICTURE:



Ex. $V = \mathbb{C}^2$, $W = \text{SPAN} \{ \underbrace{(1-i, 2i)^T}_{\text{call this } \underline{b}}$

$$P_W = \frac{(1-i, 2i) \cdot (1+i, -2i)}{\|(1-i, 2i)\|^2}$$

$$= \frac{1}{6} \begin{pmatrix} 2 & -2-2i \\ -2+2i & 4 \end{pmatrix}$$

NOW FIND $\text{Ker}(P_W)$ TO OBTAIN W^\perp .

Q: IN ORDER TO USE THE FORMULA

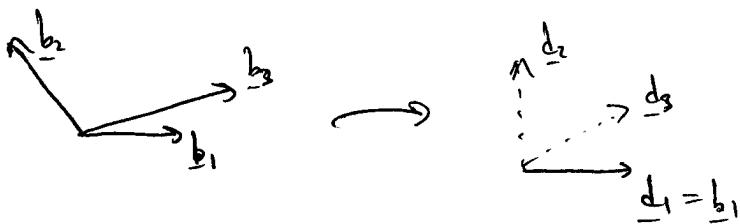
$$\underline{x} = \sum_{i=1}^n \frac{\langle \underline{b}_i | \underline{x} \rangle}{\|\underline{b}_i\|^2} \underline{b}_i, \quad \text{WE HAD TO ASSUME}$$

THAT $\{ \underline{b}_i \}$ WERE ORTHOGONAL.

GIVEN SOME BASIS $\{ \underline{b}_i \}$ OF V , HOW TO FIND AN ORTHOGONAL (OR ORTHONORMAL) BASIS OF V ?

A: GRAM-SCHMIDT PROCEDURE.

IDEA: GIVEN BASIS $\mathcal{B} = \{\underline{b}_i\}_{i=1}^n$ OF V , START WITH FIRST VECTOR AND ITERATIVELY FIND "NEW" DIRECTIONS.



PROCEDURE: DEFINE NEW BASIS $\{\underline{d}_i\}_{i=1}^n$ BY:

$$1) \underline{d}_1 = \underline{b}_1$$

$$2) \underline{d}_2 = \underbrace{(\mathbf{I} - P_{\underline{d}_1})}_{P_{\underline{d}_1}^\perp} \underline{b}_2$$

$$3) \underline{d}_3 = \underbrace{(\mathbf{I} - P_{\underline{d}_1} - P_{\underline{d}_2})}_{P_{\text{SPAN}\{\underline{d}_1, \underline{d}_2\}}^\perp} \underline{b}_3$$

⋮

$$n) \underline{d}_n = \underbrace{(\mathbf{I} - \sum_{i=1}^{n-1} P_{\underline{d}_i})}_{P_{\text{SPAN}\{\underline{d}_1, \dots, \underline{d}_{n-1}\}}^\perp} \underline{b}_n$$

$\Rightarrow \mathcal{D} = \{\underline{d}_i\}_{i=1}^n$ ORTHOGONAL BASIS OF V !

IF WE LET $\underline{e}_i = \frac{\underline{d}_i}{\|\underline{d}_i\|}$ FOR EACH $i=1, \dots, n$,

THEN $\mathcal{E} = \{\underline{e}_i\}_{i=1}^n$ IS AN ORTHONORMAL BASIS OF V .

Ex. $V = \mathbb{R}^3$, $\mathcal{B} = \{(1, 1, 0)^T, (3, 1, 1)^T, (1, 1, 3)^T\}$.

TO FIND AN ORTHONORMAL BASIS OF \mathbb{R}^3 , WE USE THE GRAM-SCHMIDT PROCEDURE ON \mathcal{B} :

$$1) \underline{d}_1 = \underline{b}_1 = (1, 1, 0)^T.$$

$$2) \underline{d}_2 = (\mathbf{I} - P_{\underline{d}_1}) \underline{b}_2 = \underline{b}_2 - P_{\underline{d}_1} \underline{b}_2$$

$$= (3, 1, 1)^T - \frac{\langle (1, 1, 0)^T | (3, 1, 1)^T \rangle}{\|(1, 1, 0)^T\|} (1, 1, 0)^T$$

$$= (1, -1, 1)^T.$$

$$3) \underline{d}_3 = (\mathbf{I} - P_{\underline{d}_1} - P_{\underline{d}_2}) \underline{b}_3 = \underline{b}_3 - P_{\underline{d}_1} \underline{b}_3 - P_{\underline{d}_2} \underline{b}_3$$

$$= \underline{b}_3 - \frac{\langle \underline{d}_1 | \underline{b}_3 \rangle}{\|\underline{d}_1\|^2} \underline{d}_1 - \frac{\langle \underline{d}_2 | \underline{b}_3 \rangle}{\|\underline{d}_2\|^2} \underline{d}_2$$

$$= (-1, 1, 2)^T.$$

$\Rightarrow \mathcal{D} = \{(1, 1, 0)^T, (1, -1, 1)^T, (-1, 1, 2)^T\}$. ORTHONORMAL BASIS.

$$\Rightarrow \mathcal{E} = \left\{ \frac{1}{\sqrt{2}} (1, 1, 0)^T, \frac{1}{\sqrt{3}} (1, -1, 1)^T, \frac{1}{\sqrt{6}} (-1, 1, 2)^T \right\}$$

ORTHONORMAL BASIS

REMARK: IN GRAM SCHMIDT PROCEDURE, REMEMBER THAT

$$\underline{d}_i = \left(\mathbf{I} - P_{\underline{d}_1} - \dots - P_{\underline{d}_{i-1}} \right) \underline{b}_i \quad \checkmark$$

AND NOT

~~$$\underline{d}_i = \left(\mathbf{I} - P_{\underline{b}_1} - \dots - P_{\underline{b}_{i-1}} \right) \underline{b}_i$$~~

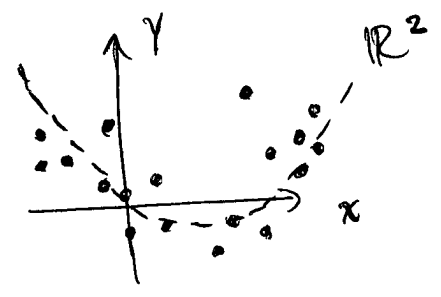
THAT IS, WE MUST USE THE ITERATIVE PROCEDURE TO OBTAIN A CORRECT ANSWER.

Lecture 27
03/30/12

$A\underline{x} = \underline{b}$ has sol'n $\underline{x} \iff \underline{b} \in \text{Ran}(A)$
(column space of A).

Q: WHAT IF $\underline{b} \notin \text{Ran}(A)$? CAN WE FIND \underline{x} THAT "ALMOST" SOLVES $A\underline{x} = \underline{b}$? WHAT DO WE MEAN BY "ALMOST"?

MOTIVATION: FITTING CURVES TO DATA.



HAVE DATA $\{(x_i, y_i)\}_{i=1}^m$.

SUPPOSE WE EXPECT THE OBSERVED DATA TO FIT THE MODEL $y = cx + dx^2$.
↑ unknown.

WHAT CHOICE FOR c, d ? ACCORDING TO MODEL,

$$\begin{aligned} cx_1 + dx_1^2 &= y_1 \\ cx_2 + dx_2^2 &= y_2 \\ &\vdots \\ cx_m + dx_m^2 &= y_m \end{aligned}$$

$$\begin{pmatrix} x_1 & x_1^2 \\ \vdots & \vdots \\ x_m & x_m^2 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

$A \in M_{m,2}(\mathbb{R})$ $\underline{x} \in \mathbb{R}^2$ $\underline{y} \in \mathbb{R}^m$

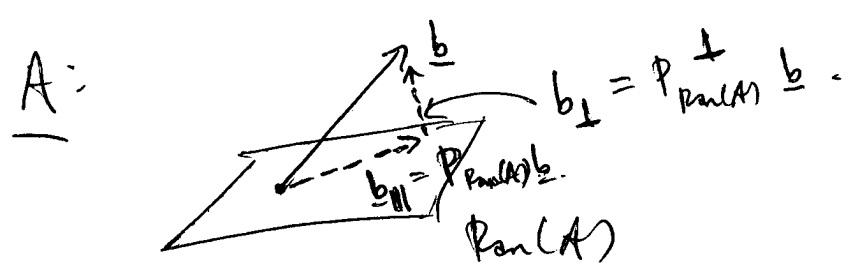
LEAST SQUARES (6.7) :

DEFINE ERROR $E(\underline{x}) = \|A\underline{x} - \underline{b}\|^2$.

NOTE THAT $E(\underline{x}) = 0$ IFF \underline{x} EXACTLY SOLVES $A\underline{x} = \underline{b}$!

DEF. \underline{x} IS A LEAST SQUARES SOLN TO $A\underline{x} = \underline{b}$ IF $E(\underline{x})$ IS MINIMIZED AT \underline{x} .

Q: GIVEN A, \underline{b} , HOW TO FIND \underline{x} ? IS IT UNIQUE?



$$\Rightarrow \underline{b} = \underline{b}_{\perp} + \underline{b}_{\parallel}$$

BY PYTHAGOREAN THM.,

$$E(\underline{x}) = \|\underline{b} - A\underline{x}\|^2 = \|\underbrace{\underline{b}_{\perp}}_{\in (\text{Ran}(A))^{\perp}} + \underbrace{(\underline{b}_{\parallel} - A\underline{x})}_{\in \text{Ran}(A)}\|^2$$

$$= \|\underline{b}_{\perp}\|^2 + \|\underline{b}_{\parallel} - A\underline{x}\|^2$$

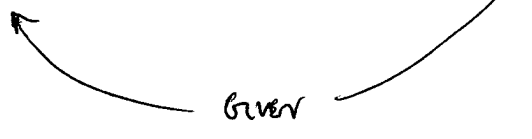
WANT TO MINIMIZE THIS.

SINCE $\underline{b}_{\parallel} \in \text{Ran}(A)$, THERE IS AT LEAST ONE SOLN.

\underline{x} TO $A\underline{x} = \underline{b}_{\parallel}$, SO $\min_{\underline{x}} E(\underline{x}) = \|\underline{b}_{\perp}\|^2$.

so, \underline{x} IS A LEAST SQUARES SOLN IF AND ONLY IF

$$A \underline{x} = \underline{b}_{||} = P_{\text{Ran}(A)} \underline{b} .$$



NOTE: LEAST SQUARES SOLN. IS UNIQUE IF AND ONLY IF A HAS FULL RANK (I.E, RANK OF A EQUALS NUMBER OF COLUMNS).

REMARK: IT IS USUALLY NOT STRAIGHTFORWARD TO COMPUTE $P_{\text{Ran}(A)}$. INSTEAD, WE USE ANOTHER APPROACH.

FIRST, NOTE THAT

$$\underline{b}_{\perp} \perp \text{Ran}(A) \Leftrightarrow \underline{b}_{\perp} \perp \text{columns of } A$$
$$\Leftrightarrow A^* \underline{b}_{\perp} = \underline{0} .$$

THEN, FOR ANY LEAST SQUARES SOLN \underline{x} ,

$$A^*(A \underline{x}) = A^*(\underline{b}_{||}) = A^*(\underline{b} - \underline{b}_{\perp}) = A^* \underline{b}$$

DEF. (NORMAL EQN.) $(A^*A) \underline{x} = A^* \underline{b}$, A, \underline{b} GIVEN.

- THM.
- \underline{x} SOLN. TO NORMAL EQN. $\Leftrightarrow \underline{x}$ LEAST SQUARES SOLN.
 - \underline{x} UNIQUE $\Leftrightarrow A$ HAS RANK $n = \#$ OF COLUMNS.

EX. FIND BEST LINE THROUGH POINTS

$(0, -1), (1, 1), (2, 4), (3, 9)$.

THAT IS, WE SEEK CONSTANTS c, d FOR $y = c + dx$.

WRITING THE EQN'S $y_i = c + dx_i$ FOR $i = 1, \dots, 4$.

IN MATRIX FORM,

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}}_A \underbrace{\begin{pmatrix} c \\ d \end{pmatrix}}_x = \underbrace{\begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \end{pmatrix}}_b$$

$$A^*A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}$$

$$A^*b = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \end{pmatrix} = \begin{pmatrix} 13 \\ 36 \end{pmatrix}$$

NORMAL EQN:

$$\underbrace{\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}}_{A^*A} \underbrace{\begin{pmatrix} c \\ d \end{pmatrix}}_x = \underbrace{\begin{pmatrix} 13 \\ 36 \end{pmatrix}}_b$$

$$\Rightarrow \begin{pmatrix} c \\ d \end{pmatrix} = \frac{1}{10} \begin{pmatrix} -17 \\ 33 \end{pmatrix} \cdot \left(\begin{array}{l} \text{UNIQUE SOLN. SINCE } A \\ \text{HAS FULL RANK.} \end{array} \right)$$

$$\Rightarrow y = -1.7 + 3.3x \quad \text{LINE OF BEST FIT. (REGRESSION)}$$

NOTE: NEED LINEAR EQN. FOR UNKNOWN PARAMETERS
 c_0, c_1, \dots, c_n . MODEL FOR DATA DOESN'T
 NEED TO BE LINEAR!

EX. FIT DATA TO $y = c_0 + c_1 x$. ✓
 LINEAR IN c_i 'S.

EX. FIT DATA TO $c_0 y^2 - c_1 x^2 = c_2$ ✓
 LINEAR IN c_i 'S.

EX. FIT DATA TO $y = c_0 e^{c_1 x}$ ✗
 NOT LINEAR IN c_i 'S