5.4: Least Squares.

<u>Ex1</u>: Find a linear model to describe height as a fct of armspan.

$$c_0 R + c_1 = H$$

collect student data.

$$\begin{array}{ccc} A & \vec{x} & = & \vec{b} \\ \begin{bmatrix} R_1 & 1 \\ R_2 & 1 \\ R_3 & 1 \\ & \vdots \end{bmatrix} & \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} & = & \begin{bmatrix} H_1 \\ H_2 \\ H_3 \\ \vdots \end{bmatrix} \end{array}$$

can we solve $A\vec{x} = \vec{b}$? ... no, it's inconsistent.

since we can't find a perfect fit, let's look for the "best possible" $\vec{x}$ ... call it $\vec{x}^*$.

$\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$ or the solution to

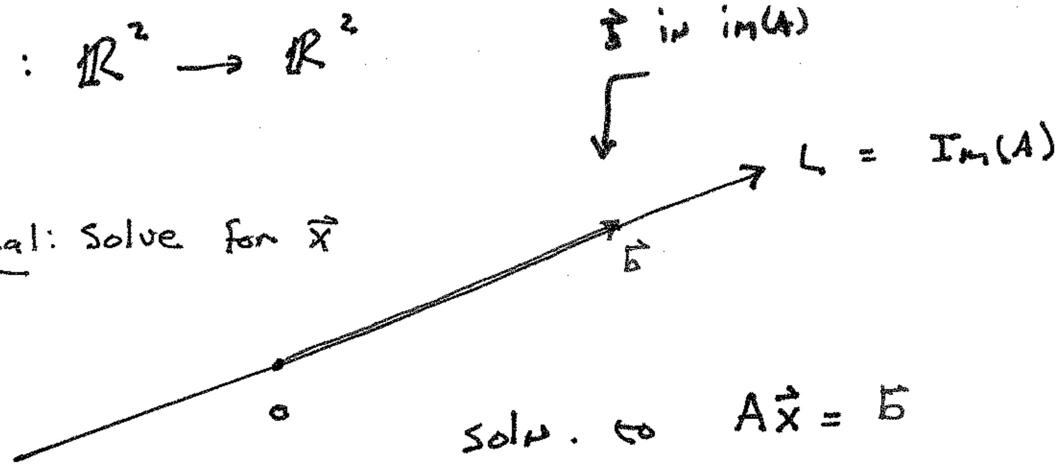
$A^T A \vec{x}^* = A^T \vec{b}$ ....

compare to the least squares sol. on the calculator...

wow! where did this come from?

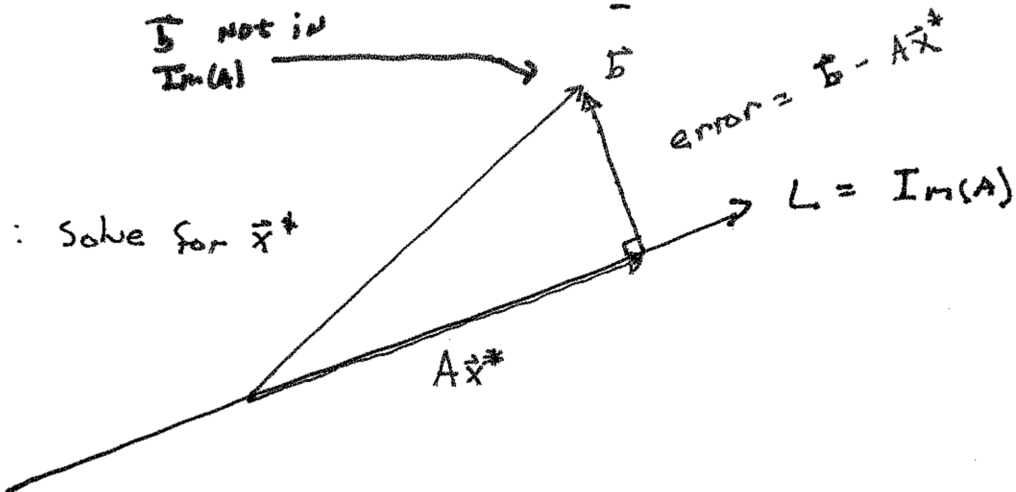
In 2D: $A : \mathbb{R}^2 \to \mathbb{R}^2$

Goal: Solve for $\vec{x}$



soln. to $A\vec{x} = \vec{b}$
is the $\vec{x}$ s.t. $A\vec{x} = \vec{b}$

Goal: Solve for $\vec{x}^*$

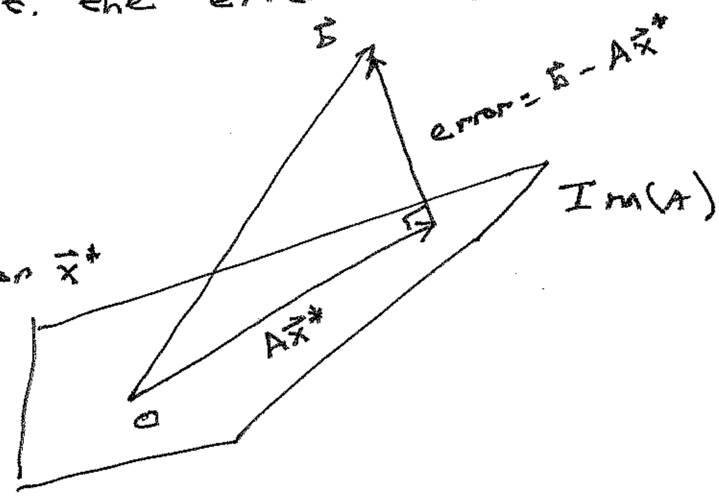


No soln. to $A\vec{x} = \vec{b}$

The closest vector to $\vec{b}$ in Im(A) is $A\vec{x}^*$ and so we want to find this $\vec{x}^*$ s.t. the error vector $\vec{b} - A\vec{x}^*$ is minimized.

In 3D: $A : \mathbb{R}^3 \mapsto \mathbb{R}^3$

Goal: solve for $\vec{x}^*$



consider $V = \text{Im}(A)$ where $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$

$$V^\perp = \{\vec{x} \in \mathbb{R}^n \mid \vec{v}_i^T \vec{x} = 0 \quad \text{for } i = 1, \dots, m\}$$

The solutions to

$$\Rightarrow \begin{bmatrix} - & \vec{v}_1^T & - \\ & \vdots & \\ - & \vec{v}_m^T & - \end{bmatrix} \vec{x} = \vec{0} \quad \text{make up } V^\perp.$$

$$\Rightarrow \quad V^\perp = (\text{Im}(A))^\perp = \text{Ker}(A^T).$$

Thm:

(a) if A is an $n \times m$ matrix, $\ker(A) = \ker(A^T A)$.

* key concept in the proof:

If $\ker(A) \subseteq \ker(A^T A)$ and $\ker(A^T A) \subseteq \ker(A)$

Then $\ker(A) = \ker(A^T A)$.

(b) if A is an $n \times m$ matrix w/ $\ker(A) = \{\vec{0}\}$

Then $A^T A$ is invertible.

* This follows from A.

Dfn: Consider a lin. sys. $A\vec{x} = \vec{b}$ where A is an $n \times m$ matrix. A vector $\vec{x}^* \in \mathbb{R}^m$ is called a least-squares sol. of this system if

$$\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\| \quad \text{for all } \vec{x} \in \mathbb{R}^m$$

Q: Why is this called the "least-squares" soln.?

Logic

We want $\vec{x}^*$ sol. to $A\vec{x} = \vec{b}$

$\Updownarrow$

$\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$

$\Updownarrow$ Note: $A\vec{x}^* = \text{proj}_V \vec{b}$

$\vec{b} - A\vec{x}^* \in V^\perp = (\text{Im}(A))^\perp = \ker(A^T)$

$\Updownarrow$

$A^T(\vec{b} - A\vec{x}^*) = \vec{0}$

$\Updownarrow$

$A^T A x^* = A^T \vec{b}$. (the normal eqt to $A\vec{x} = \vec{b}$).

If the cols of $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$ form a basis for $V \Rightarrow \ker(A) = \{\vec{0}\} \Rightarrow A^T A$ is invertible.

$\Rightarrow \vec{x}^* = (A^T A)^{-1} A^T \vec{b}$

and $\text{proj}_V \vec{b} = Ax^* = A(A^T A)^{-1} A^T \vec{b}$.

<u>Thm:</u> Consider a subspace V of $\mathbb{R}^n$ w/ basis $\vec{v}_1, \dots, \vec{v}_m$. Let $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$.

Then the matrix of the of the orthogonal projection onto V is $A(A^TA)^{-1}A^T$