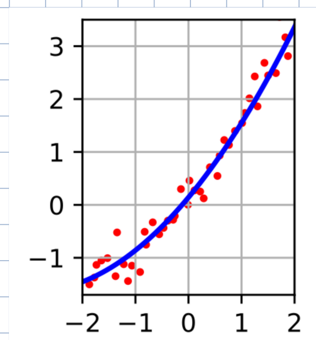
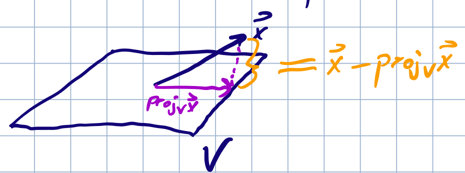
HW 7 due 3/11
5.1: 16, 26, 28
5.2: 4, 6, 18, 29
5.3: 2, 6, 8, 10

HW 8 due 3/18
5.4: 20, 36, 38



Thm For $\vec{x} \in \mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$, $\text{proj}_V \vec{x}$ is the closest vector in V to $\vec{x}$, i.e.,

$$\|\vec{x} - \text{proj}_V \vec{x}\| < \|\vec{x} - \vec{v}\| \text{ for all } \vec{v} \in V \text{ where } \vec{v} \neq \text{proj}_V \vec{x}$$



Def For a linear $A\vec{x} = \vec{b}$, a vector $\vec{x}^*$ is a <u>least squares solution</u> of this system if $\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\|$ for all vectors $\vec{x}$.

$\sqrt{(\ \ )^2 + (\ \ )^2 + \cdots + (\ \ )^2}$

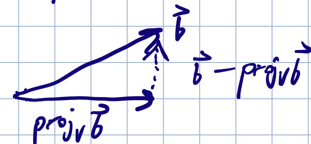
Eg $A = \begin{pmatrix} 1 & 1.5 & 2.25 \\ 1 & 1.6 & 2.56 \\ 1 & 1.7 & 2.89 \\ 1 & 1.8 & 3.24 \end{pmatrix}$ $\quad \vec{b} = \begin{pmatrix} 53 \\ 58 \\ 64 \\ 72 \end{pmatrix}$ $\quad A\vec{x} = \vec{b}$ is inconsistent

Want $\vec{x}^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix}$ such that $\|\vec{b} - A\vec{x}^*\|$ is minimized.

How to compute?

$\|\vec{b} - A\vec{x}^*\|$ is minimized $\iff A\vec{x}^* = \text{proj}_V \vec{b}$ for $V = \text{im}(A)$

$\iff \vec{b} - A\vec{x}^* \in (\text{im} A)^\perp$



Thm $(\operatorname{im} A)^\perp = \ker(A^T)$

Pf

$$A = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix} \qquad V = \operatorname{im} A$$

$$V^\perp = \{\vec{x} \in \mathbb{R}^n : \underbrace{\vec{v}_i \cdot \vec{x}}_{} = 0 \text{ for all } i\}$$

matrix multip $\longrightarrow$ $\vec{v}_i^T \vec{x} = 0$

$$\Leftrightarrow \quad \underbrace{\begin{pmatrix} - & \vec{v}_1^T & - \\ & \vdots & \\ - & \vec{v}_n^T & - \end{pmatrix}}_{A^T} \vec{x} = \vec{0}$$

$$(\operatorname{im} A)^\perp = V^\perp = \{\vec{x} \in \mathbb{R}^n : A^T\vec{x} = \vec{0}\} = \ker(A^T) \quad \square$$

Using this theorem, $\|\vec{b} - A\vec{x}^*\|$ is minimized $\Leftrightarrow \vec{b} - A\vec{x}^* \in \ker(A^T)$

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

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

$\vec{v} \in \ker A^T \Leftrightarrow A^T\vec{v} = \vec{0}$

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

Eg

$$A = \begin{pmatrix} 1 & 1.5 & 2.25 \\ 1 & 1.6 & 2.56 \\ 1 & 1.7 & 2.89 \\ 1 & 1.8 & 3.24 \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} 53 \\ 58 \\ 64 \\ 72 \end{pmatrix} \qquad A\vec{x} = \vec{b} \text{ is inconsistent}$$

Want $\vec{x}^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix}$ such that $\|\vec{b} - A\vec{x}^*\|$ is minimized.

$$\Leftrightarrow \underbrace{\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1.5 & 1.6 & 1.7 & 1.8 \\ 2.25 & 2.56 & 2.91 & 3.24 \end{pmatrix} \begin{pmatrix} 1 & 1.5 & 2.25 \\ 1 & 1.6 & 2.56 \\ 1 & 1.7 & 2.89 \\ 1 & 1.8 & 3.24 \end{pmatrix} \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1.5 & 1.6 & 1.7 & 1.8 \\ 2.25 & 2.56 & 2.91 & 3.24 \end{pmatrix} \begin{pmatrix} 53 \\ 58 \\ 64 \\ 72 \end{pmatrix}}_{\text{linear system of equations}}$$

Thm The least squares solutions of the system $A\vec{x} = \vec{b}$ are the exact solutions of the system $A^TA\vec{x}^* = A^T\vec{b}$.

If $\ker A = \{\vec{0}\}$, we can say more

Thm $\ker(A) = \ker(A^TA)$

Pf i) If $\vec{x} \in \ker A$, then $(A^TA)\vec{x} = A^T(A\vec{x}) = A^T\vec{0} = \vec{0}$

ii) If $\vec{x} \in \ker A^TA$, then $A\vec{x} \in \operatorname{im}(A)$

$A^TA\vec{x} = \vec{0} \Rightarrow A\vec{x} \in \ker(A^T) = (\operatorname{im} A)^\perp$

$\Bigg\} \Rightarrow A\vec{x} = \vec{0} \Rightarrow \vec{x} \in \ker A$

$\square$

Cor If $\ker A = \{\vec{0}\}$, then A^TA is invertible because it is square and $\ker(A^TA) = \ker(A) = \{\vec{0}\}$.

$\rightsquigarrow (A^TA)\vec{x}^* = A^T\vec{b}$ with $\ker(A) = \{\vec{0}\}$

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

Thm If $\ker A = \{\vec{0}\}$, then $A\vec{x} = \vec{b}$ has unique least squares solution

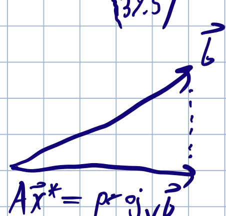
$$\vec{x}^* = (A^TA)^{-1}A^T\vec{b}.$$

Eg $A = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \\ 1 & 4 & 9 \end{pmatrix}}_{\text{columns are linearly independent}}$ $\quad \vec{b} = \begin{pmatrix} 9 \\ 25 \\ 31 \\ 40 \end{pmatrix}$

columns are linearly independent $\Rightarrow \ker A = \{0\}$

Least squares solution: $\vec{x}^* = (A^TA)^{-1}A^T\vec{b} = \begin{pmatrix} -14.5 \\ 31.5 \\ -8 \end{pmatrix}$

$A\vec{x}^* = \begin{pmatrix} 9 \\ 24.5 \\ 32 \\ 39.5 \end{pmatrix}$



$V = \operatorname{im}(A)$

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

$\Longrightarrow \underbrace{A(A^TA)^{-1}A^T}_{\text{projection matrix}}\vec{b} = \operatorname{proj}_V \vec{b}$

Thm For subspace V of $\mathbb{R}^n$ with basis $\vec{v}_1, \ldots, \vec{v}_m$

let $A = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{pmatrix}$. Then, the matrix of proj_V is $A(A^TA)^{-1}A^T$.

Data Fitting

For data $\{(x_1,y_1),\dots,(x_m,y_m)\}$, the polynomial of degree n of best fit, $f(t)$, is given by $f(t) = c_0^* + c_1^* t + c_2^* t^2 + \cdots + c_n^* t^n$ for $\vec{c}^* = \begin{pmatrix} c_0^* \\ c_1^* \\ \vdots \\ c_n^* \end{pmatrix}$ is a least squares solution to $A\vec{x} = \vec{b}$ with

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{pmatrix} \quad \text{and} \quad \vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

f linear (deg 1) $\Longrightarrow$ "linear regression"

Eg $\{(x_i, y_i, z_i)\}$ $\longrightarrow$ see example 5.4 : 6 Least square to find line like $z = c_0 + c_1 x + c_2 y$