Northeastern University, Department of Mathematics

MATH 4570 - Matrix Methods in Data Analysis and Machine Learning

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§~~5~~ 6 ~~Inner product spaces~~

• $\boxed{\text{Mid I}}$

Next Thursday!

In class (with

· lecture note

· one book!)

NO: homework!

CONTENTS

Review: V $\langle , \rangle$

1. Inner Product Space (vector space with an inner product)

Example: 1. $\mathbb{R}^n$ (weighted) dot product

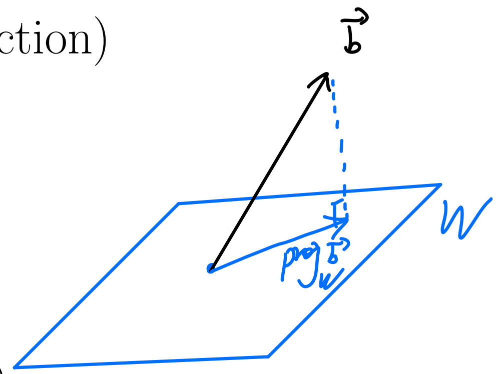
2. $C[a,b]$. $\langle f, g\rangle := \int_a^b f(x)g(x)\,dx$

$\Downarrow$ Geometry $\to$ norm, $\to$ angle

2. Use orthogonal basis (to find orthogonal projection)

$\mathcal{B} = \{\vec{v}_1 \cdots \vec{v}_s\}$ of $W \subset V$

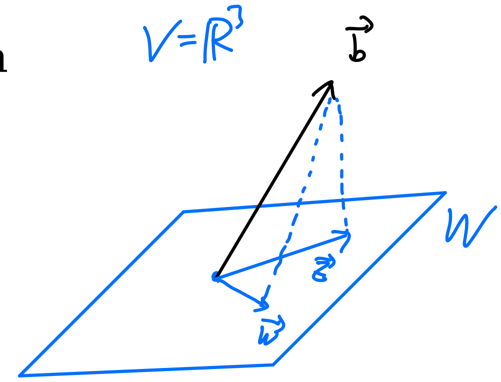
$$\text{proj}_W \vec{b} = \frac{\langle \vec{b}, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{v}_1 + \cdots + \frac{\langle \vec{b}, \vec{v}_s\rangle}{\langle \vec{v}_s, \vec{v}_s\rangle}\vec{v}_s$$



3. Find orthogonal basis (Gram-Schmidt process)

A basis $\{\vec{b}_1 \cdots \vec{b}_s\}$ of $W \longmapsto$ orthogonal basis $\{\vec{v}_1 \cdots \vec{v}_s\}$ of W.

1. Least Squares Problem



Set up: · V : inner product space

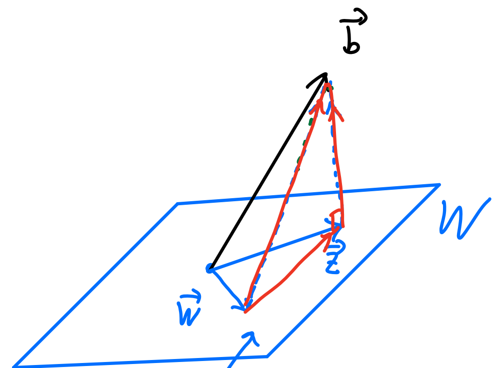
· $W \subset V$ subspace.

· $\vec b \in V$, $\vec b \notin W$.

Question: What is the "closest" vector $\vec z \in W$ to $\vec b$?

$\iff$ Find $\vec z \in W$ s.t. $\|\vec b - \vec z\| \le \|\vec b - \vec w\|$ for any $\vec w \in W$

Answer: Thm: $\vec z = \text{proj}_W \vec b$



· proof:

$$\|\vec b - \vec w\|^2 = \|\vec b - \vec z\|^2 + \|\vec z - \vec w\|^2$$

Calculation:

Method (1) If W has an orthogonal basis $\{\vec v_1 \cdots \vec v_p\}$, then

★ $$\text{proj}_W \vec b = \frac{\langle \vec b, \vec v_1\rangle}{\langle \vec v_1, \vec v_1\rangle}\vec v_1 + \cdots + \frac{\langle \vec b, \vec v_p\rangle}{\langle \vec v_p, \vec v_p\rangle}\vec v_p$$

Method (2) If W has "basis" $\{\vec w_1, \cdots, \vec w_p\}$, then $\vec z = \cancel{x_1}\vec w_1 + \cdots + \cancel{x_p}\vec w_p = ?$

$= \text{proj}_W \vec b$

Q: how to find $x_1 \cdots x_p$?

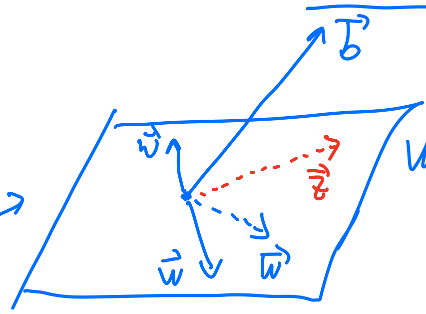
$\vec b^\perp = \vec b - \vec z \perp W$

Solve:

$$\begin{cases} \langle \vec{b} - \vec{z}, \vec{w}_1 \rangle = 0 \\ \vdots \\ \langle \vec{b} - \vec{z}, \vec{w}_p \rangle = 0 \end{cases}$$

2. Approximate Solutions to Inconsistent Systems

$A\vec{x} = \vec{b}$



$W = \{\text{all } A\vec{x}\} = \text{im} A$

Set up:

Let A be an $n \times m$ matrix. $\mathbb{R}$

Let $\vec{b} \in \mathbb{R}^n$.

Suppose $A\vec{x} = \vec{b}$ has no solution. $\vec{b} \notin \text{im} A$

Q: Find the (best) solution for $A\vec{x} = \vec{b}$.

- Consider $\mathbb{R}^n$ with any inner product. $\Leftrightarrow$ Find the "closest" vector $\vec{z}$ to $\vec{b}$. $\vec{z} = A\vec{x}_*$

e.g.① dot prod e.g.② weighted dot prod.

[Least-Squares Problem/Solution for $A\vec{x} = \vec{b}$]

Problem: Find the vector(s) $\vec{x}_* \in \mathbb{R}^m$ such that for all $x \in \mathbb{R}^m$,

$$||\overset{\vec{z}}{A\vec{x}_*} - \vec{b}|| \leq ||\overset{\vec{w}}{A\vec{x}} - \vec{b}||$$

Solutions:

$$A\vec{x}_* = \text{proj}_{\text{im} A} \vec{b}$$

"dot product"

Example 1. Find the least-squares solutions for $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} -1 & 4 \\ 1 & 8 \\ -1 & 4 \end{bmatrix}$

and $\vec{b} = \begin{bmatrix} 14 \\ -4 \\ 0 \end{bmatrix}$

- $\vec{a_1} \cdot \vec{a_2} = 0 \implies \{\vec{a_1}, \vec{a_2}\}$ is orthogonal basis for im A.

- $\text{proj}_{\text{im} A} \vec{b} = \dfrac{\vec{b} \cdot \vec{a_1}}{\vec{a_1} \cdot \vec{a_1}} \vec{a_1} + \dfrac{\vec{b} \cdot \vec{a_2}}{\vec{a_2} \cdot \vec{a_2}} \vec{a_2} = \begin{bmatrix} 7 \\ -4 \\ 7 \end{bmatrix}$

- Solve $\left[\begin{array}{cc|c} -1 & 4 & 7 \\ 1 & 8 & -4 \\ -1 & 4 & 7 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & & -6 \\ & 1 & 1/4 \end{array}\right]$

- $\vec{x}_* = \begin{bmatrix} -6 \\ 1/4 \end{bmatrix}$

In particular, if we consider "**dot product**" on $\mathbb{R}^n$, we have the following formula.

Any $n\times m$ matrix.

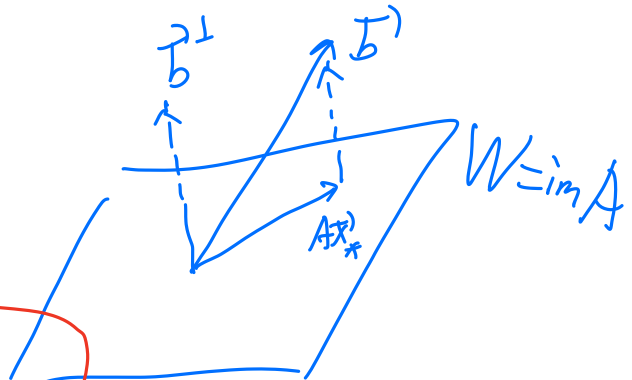
$A\vec{x}_* = \text{proj}_{\text{im}A}\vec{b}$

Theorem 2. *(Normal Equation) The Least-Square solutions of $A\vec{x} = \vec{b}$ coincide with the solutions of of* ***normal equations***

$$(\underbrace{A^T}_{m\times n}\underbrace{A}_{n\times m})\vec{x} = A^T\vec{b}.$$

$(m\times m)$

- $A\vec{x}_* = \text{proj}_{\text{im}A}\vec{b}$



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

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

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

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

Assume A^TA is invertible.

More generally, we can also consider **weighted dot product** on $\mathbb{R}^n$,

$$\langle \vec{u}, \vec{v} \rangle_W := \vec{u}^T W \vec{v}$$

where W is a "positive-definite" "symmetric" matrix.

e.g. $W = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ & 2 & & & \\ & & 3 & & \\ & & & \ddots & \\ & & & & n \end{bmatrix}$ $\quad c_i > 0$

$\langle \vec{u}, \vec{v} \rangle_W = c_1u_1v_1 + \cdots + c_nu_nv_n$

- $A\vec{x} = \vec{b}$ no slu.
- Slution $A\vec{x}_* = \text{proj}^W_{\text{im}A}\vec{b}$

$(\text{im}A)^\perp = \ker(WA)^T$

- $A^TWA\vec{x}_* = A^TW\vec{b}$

Example 3. Find the least-squares solutions for $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} -1 & 4 \\ 1 & 8 \\ -1 & 4 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 14 \\ -4 \\ 0 \end{bmatrix}$

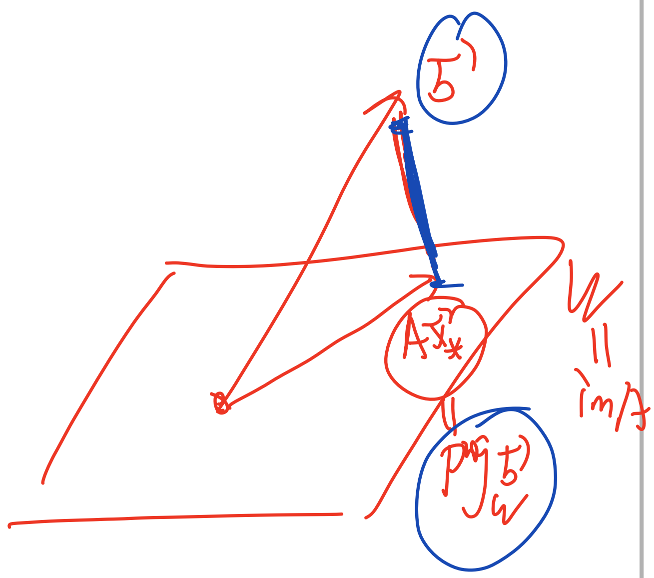
$$A^TA = \begin{bmatrix} -1 & 1 & -1 \\ 4 & 8 & 4 \end{bmatrix} \begin{bmatrix} -1 & 4 \\ 1 & 8 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 96 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} -1 & 1 & -1 \\ 4 & 8 & 4 \end{bmatrix} \begin{bmatrix} 14 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} -18 \\ 24 \end{bmatrix}$$

solve the normal equation $\boxed{A^TA\vec{x} = A^T\vec{b}}$

$$\left[\begin{array}{cc|c} 3 & 0 & -18 \\ 0 & 96 & 24 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & -6 \\ 0 & 1 & 4 \end{array}\right]$$

So $\vec{x}_* = \begin{bmatrix} -6 \\ 4 \end{bmatrix}$ is the least-squares solution



(2) The image $\text{im}(A)$ is a plane in $\mathbb{R}^3$ passing the origin. Find the distance from the vector $\vec{b}$ (or the point $(14, -4, 0)$) to the plane $\text{im}(A)$. "Error"

The distance is given be the norm of $\vec{b}^{\perp} = \vec{b} - \text{proj}_{\text{im}(A)} \vec{b}$.

We know that $\text{proj}_{\text{im}(A)} \vec{b} = Ax_* = \begin{bmatrix} -1 & 4 \\ 1 & 8 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} -6 \\ 1/4 \end{bmatrix} = \begin{bmatrix} 7 \\ -4 \\ 7 \end{bmatrix}$.

So, $\vec{b}^{\perp} = \begin{bmatrix} 14 \\ -4 \\ 0 \end{bmatrix} - \begin{bmatrix} 7 \\ -4 \\ 7 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ -7 \end{bmatrix}$. So the distance is $||\vec{b}^{\perp}|| = 7\sqrt{2}$.

Example 4. Find the least-squares solutions for the system $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$

Step 1. Construct the normal equation $A^TA\vec{x} = A^T\vec{b}$

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 4 \\ 6 \end{bmatrix}$$

Solve the normal equation

$$\left[\begin{array}{ccc|c} 4 & 2 & 2 & 10 \\ 2 & 2 & 0 & 4 \\ 2 & 0 & 2 & 6 \end{array}\right] \to \cdots \to \text{rref} = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$x_1 = 3 - x_3$

$x_2 = -1 + x_3$

x_3 free

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 - x_3 \\ -1 + x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

A technical property:

Proposition 5. *Let A be an $n \times m$ matrix.*

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

$$\Updownarrow$$

$$\operatorname{rank} A = \operatorname{rank} A^T A$$

$m \times m$

Corollary 6. *If* rank $A = m$, *the normal equation* $(A^T A)\vec{x} = A^T \vec{b}$ *has a unique solution:*

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

QR factorization method Suppose A is $n \times m$ matrix with full column rank. Solve the least squares solution using QR factorization $A = QR$ where Q is an orthogonal matrix $n \times m$ and R is an $m \times m$ upper triangular matrix with rank m.

$Q^T = Q^{-1}$

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

Problem: Fitting a function of a certain type of data. We use the following three example to illustrate this application.

Example 7. Find a cubic polynomial $f(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3$ whose graph passes through the points $(0, 5)$, $(1, 3)$, $(-1, 13)$, $(2, 1)$

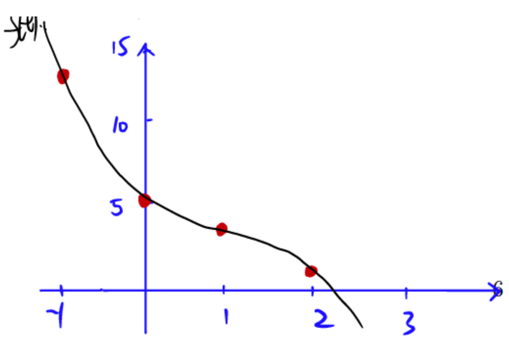
Solution:

$f(0)=5 \iff$

We need to solve the linear system
$$\begin{cases} c_0 & = 5 \\ c_0 + c_1 + c_2 + c_3 & = 3 \\ c_0 - c_1 + c_2 - c_3 & = 13 \\ c_0 + 2c_1 + 4c_2 + 8c_3 & = 1 \end{cases}$$

$$[A|\vec{b}] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 1 & 1 & 1 & 1 & 3 \\ 1 & -1 & 1 & -1 & 13 \\ 1 & 2 & 4 & 8 & 1 \end{array}\right] \to \cdots \to \mathbf{rref}[A|\vec{b}] = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

So, the linear system has the unique solution $\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \\ 3 \\ -1 \end{bmatrix}$ So, the cubic polynomial is $f(t) = 5 - 4t + 3t^2 - t^3$.



perfect fit, but calculation is hard.

Example 8. Fit a quadratic function $\boxed{g(t) = c_0 + c_1 t + c_2 t^2}$ to the four data points $(0, 5)$, $(1, 3)$, $(-1, 13)$, $(2, 1)$

We need to solve the linear system

$$g(0)=5 \quad \begin{cases} c_0 & = 5 \\ c_0 + c_1 + c_2 & = 3 \\ c_0 - c_1 + c_2 & = 13 \\ c_0 + 2c_1 + 4c_2 & = 1 \end{cases}$$

"dot prod."

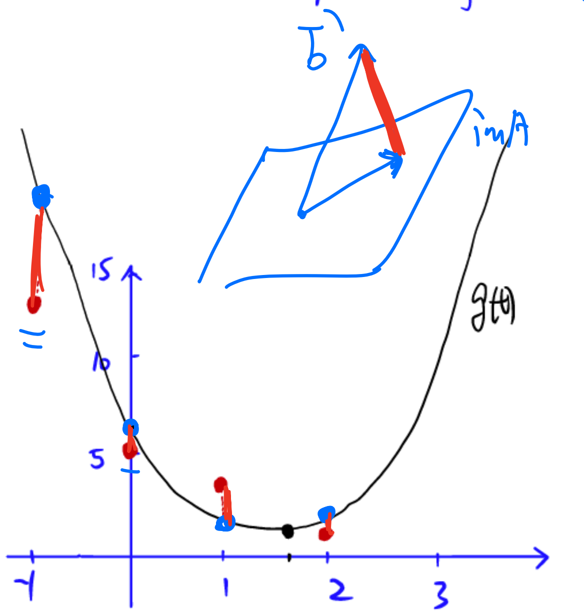
no sln.

As matrix equation $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 2 & 4 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 5 \\ 3 \\ 13 \\ 1 \end{bmatrix}$

$$A^TA = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \qquad A^T\vec{b} = \begin{bmatrix} 22 \\ -8 \\ 20 \end{bmatrix}$$

Solve the normal equation $A^TA\vec{x} = A^T\vec{b}$ $\qquad \vec{x} = \begin{bmatrix} 5.9 \\ -5.3 \\ 1.5 \end{bmatrix} = \vec{c}^*$

So, the quadratic function $g(t) = 5.9 - 5.3t + 1.5t^2$ model.



$$A\vec{c}^* = \begin{bmatrix} g(a_1) \\ g(a_2) \\ g(a_3) \\ g(a_4) \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

$$\|\vec{b} - A\vec{c}^*\|^2 = (b_1 - g(a_1))^2 + (b_2 - g(a_2))^2 + (b_3 - g(a_3))^2 + (b_4 - g(a_4))^2$$

The sum of the vertical distances between graph and data points is minimal.

Example 9. Fit a linear function $h(t) = c_0 + c_1 t$ to the four data points $(0, 5)$, $(1, 3)$, $(-1, 13)$, $(2, 1)$

We need to solve the linear system

$$\begin{cases} \boxed{c_0} & = 5 \\ c_0 + c_1 & = 3 \\ c_0 - c_1 & = 13 \\ c_0 + 2c_1 & = 1 \end{cases}$$

As matrix equation $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 5 \\ 3 \\ 13 \\ 1 \end{bmatrix}$

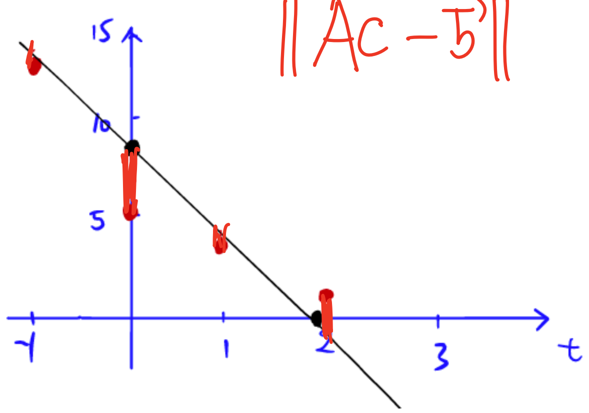
$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 13 \\ 1 \end{bmatrix} = \begin{bmatrix} 22 \\ -8 \end{bmatrix}$$

Solve the normal equation $A^T A \vec{x} = A^T \vec{b}$ $\quad$ $c\ \vec{x} = \begin{bmatrix} 7.4 \\ -3.8 \end{bmatrix}$

So the linear function is $h(t) = 7.4 - 3.8t$

$\|Ac - \vec{b}\|^2$

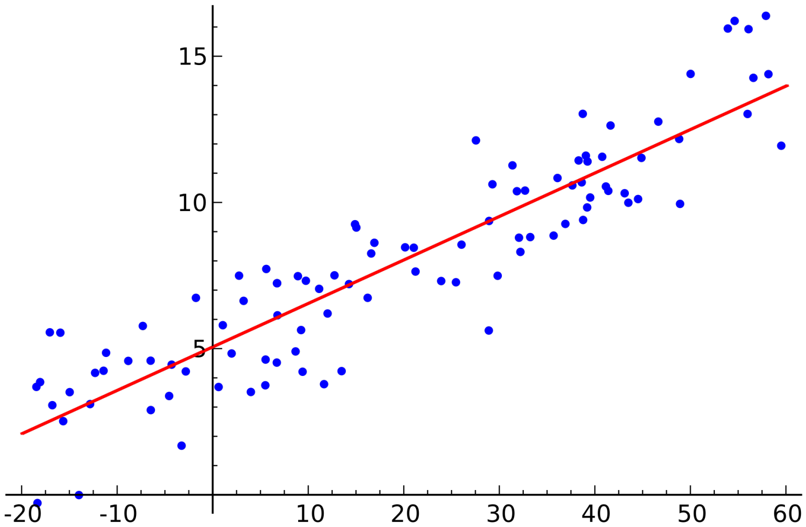


Remark: More generally we can consider n-points $(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)$.

- Find a linear function $h(t) = c_0 + c_1 t$ fits the data by the least squares.

More generally, the following question is very standard in statistics.

Example 10. Consider the data with n points (a_1, b_1), (a_2, b_2), $\ldots$, (a_n, b_n). Find a linear function $h(t) = c_0 + c_1 t$ fits the data by the least squares. (Suppose $a_1 \neq a_2$)



$A\vec{x} = \text{proj}_{\text{im}A} \vec{b}$

dot. prod.

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

We need to solve the least-squares ~~problem for $A\vec{x} = \vec{b}$~~, for $A = \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \\ 1 & a_n \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

$$A^TA = \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \\ 1 & a_n \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a_1 & a_2 & \cdots & a_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n a_i \\ \sum_{i=1}^n a_i & \sum_{i=1}^n a_i^2 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \vdots & \\ 1 & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n b_i \\ \sum_{i=1}^n a_i a_i \end{bmatrix}$$

Since $a_1 \neq a_2$, we know that rank $A = 2$.
The normal equation $A^TA\vec{x} = A^T\vec{b}$ has a unique solution

$$\vec{x}_* = (A^TA)^{-1}A^T\vec{b} = \frac{1}{n\sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2} \begin{bmatrix} \sum_{i=1}^n a_i^2 & -\sum_{i=1}^n a_i \\ -\sum_{i=1}^n a_i & n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n b_i \\ \sum_{i=1}^n a_i a_i \end{bmatrix}$$

$$= \frac{1}{n\sum_{i=1}^n a_i^2 - (\sum_{i=1}^n a_i)^2} \begin{bmatrix} (\sum_{i=1}^n a_i^2)(\sum_{i=1}^n b_i) - (\sum_{i=1}^n a_i)(\sum_{i=1}^n a_i a_i) \\ -(\sum_{i=1}^n a_i)(\sum_{i=1}^n b_i) + n\sum_{i=1}^n a_i a_i \end{bmatrix}$$

$\vec{a}^{(1)} = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1m} \end{bmatrix}$ $\vec{a}^{(2)} = \begin{bmatrix} a_{21} \\ \vdots \\ a_{2m} \end{bmatrix}$ ...

$\vec{a}_1^T$

m points

Example 11. Consider the data with m inputs and 1 output:

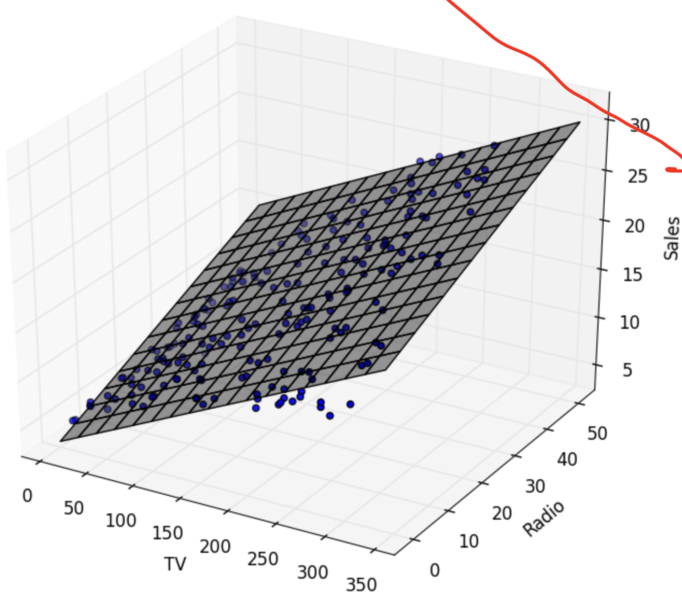
$(a_{11}, a_{12}, ..., a_{1m}, b_1)$, $(a_{21}, a_{22}, ..., a_{2m}, b_2)$, . . . , $(a_{n1}, a_{n2}, ..., a_{nm}, b_n)$.

Find a linear function $h(t_1, t_2, ..., t_n) = c_0 + c_1t_1 + c_2t_2 + \cdots + c_nt_n$ fits the data by the least squares.

$h : \mathbb{R}^n \longrightarrow \mathbb{R}^1$

For example, when $m = 2$,

$h(\vec{a}^{(i)}) = b^{(i)}$



$$\left[\begin{array}{cccc|c} 1 & a_{11} & a_{12} \cdots & a_{1m} & b_1 \\ 1 & a_{21} & a_{22} \cdots & a_{2m} & b_2 \\ \vdots & & \cdots & & \vdots \end{array}\right]$$

$A\vec{x} = \vec{b}$

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

We need to solve the least-squares problem for $A\vec{x} = \vec{b}$, for $A = \begin{bmatrix} 1 & a_{11} & a_{12} & \ldots \\ 1 & a_{21} & a_{22} & \ldots \\ \vdots & \vdots & \vdots & \cdots \\ 1 & a_{n1} & a_{n2} & \ldots \end{bmatrix}$

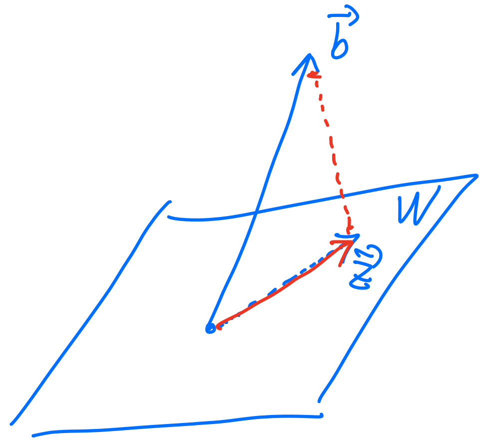
and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

Example 12. Consider the data with m inputs and s outputs:

$(a_{11}, a_{12}, ..., a_{1m}, b_{11}, ..., b_{1s})$, $(a_{21}, a_{22}, ..., a_{2m}, b_{21}, ..., b_{2s})$, . . . , $(a_{n1}, a_{n2}, ..., a_{nm},$

Find a linear function $H(\vec{t}) = \vec{c}_0 + C\vec{t}$ fits the data by the least squares.

- V inner product space
- $W \subset V$ subspace
- $\vec{b} \in V \quad \vec{b} \notin W$



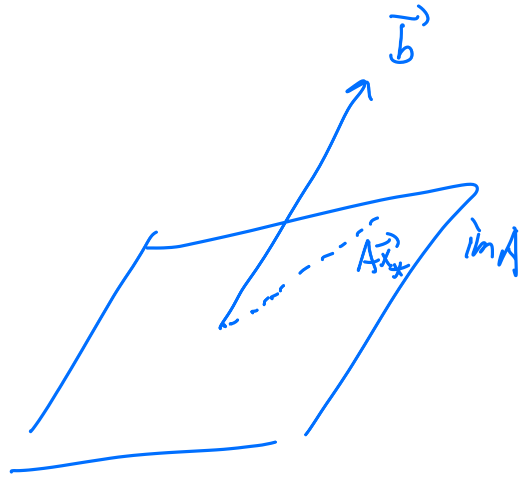
- Least squares problem:
  - Find $\vec{z} \in W$ s.t. $\|\vec{b} - \vec{z}\| \leq \|\vec{b} - \vec{w}\|$ for any $\vec{w} \in W$.

- Answer: $\vec{z} = \text{proj}_W \vec{b}$.

"Application:" Solve "least squares" slu for $\underset{m\times n}{A}\vec{x} = \vec{b}$ (inconsistent system)

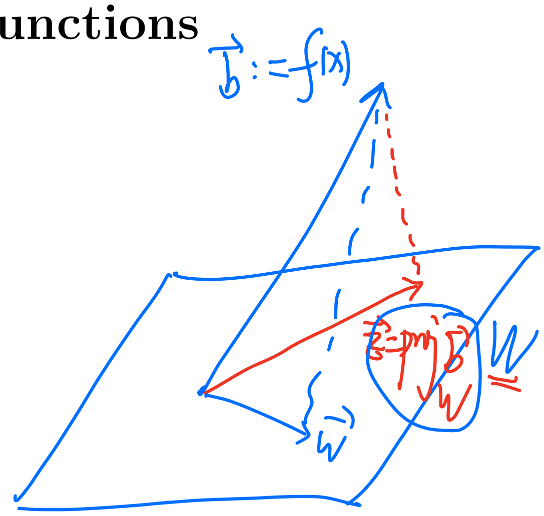
- $W = \text{im} A$
- $V = \mathbb{R}^m$ (some inner product) e.g. "dot prod."
- $\vec{b} \in V \quad \vec{b} \notin W$.

Answer: $\boxed{A\vec{x}_* = \text{proj}_W \vec{b}}$.



4. "Best" Approximation for Functions

"closest"



Set up:

$f: \mathbb{R} \to \mathbb{R}$

Let V be the vector space of {continuous functions.}

Consider the inner product on V:

$\langle f, g \rangle := \int_a^b f \cdot g \, dx \Rightarrow \|f\| = \sqrt{\langle f, f \rangle}$

Let W be the subspace of polynomials of degree $\leq n$.

Consider a function $f(x) \in V$. (e.g., $f(x) = e^x$)

or $\sin x$

Question:

Find the "best" degree n polynomial approximation of $f(x)$.

"closest"

$\Longleftrightarrow$ Find $\vec{z} = a_0 + a_1 x + \cdots + a_n x^n \in W$ s.t. $\|f(x) - \vec{z}\| \leq \|f(x) - \vec{w}\|$

for any $\vec{w} \in W$,

Answer: $\vec{z} = \boxed{\text{proj}_W f(x).}$

Method 1.

- If W has orthogonal basis $\{\vec{v}_0, \cdots, \vec{v}_n\}$, then

$$\vec{z} = \text{proj}_W f = \frac{\langle f, \vec{v}_0 \rangle}{\langle \vec{v}_0, \vec{v}_0 \rangle} \vec{v}_0 + \cdots + \frac{\langle f, \vec{v}_n \rangle}{\langle \vec{v}_n, \vec{v}_n \rangle} \vec{v}_n$$

Method 2 Using standard basis for W

Example: $\{1, x, x^2, \cdots, x^n\}$

- $W = \{a_0 + a_1 x + a_2 x^2\}$ $\quad \langle f, g \rangle = \int_0^1 f \cdot g \, dx$

- $\{1, x, x^2\}$ is a basis for W.

- Gram-Schmidt $\Longrightarrow$ $\{\underbrace{1}_{v_0}, \underbrace{t - \frac{1}{2}}_{v_1}, \underbrace{t^2 - t + \frac{1}{6}}_{v_2}\}$ is orthogonal basis for W

- Q: Find the best polynomial (≤ 2) approx for $f(x) = e^x$ /

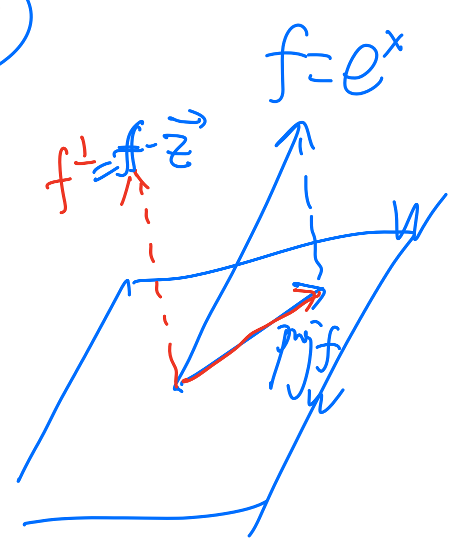
- Method 1: $\text{proj}_W f(x) :=$

- Method 2 Use $\{1, x, x^2\}$ basis for W.

$\Downarrow$

$$\vec{z} = \text{proj}_W f(x) = C_0 + C_1 x + C_2 x^2$$

Use $\boxed{f^\perp \perp W}$

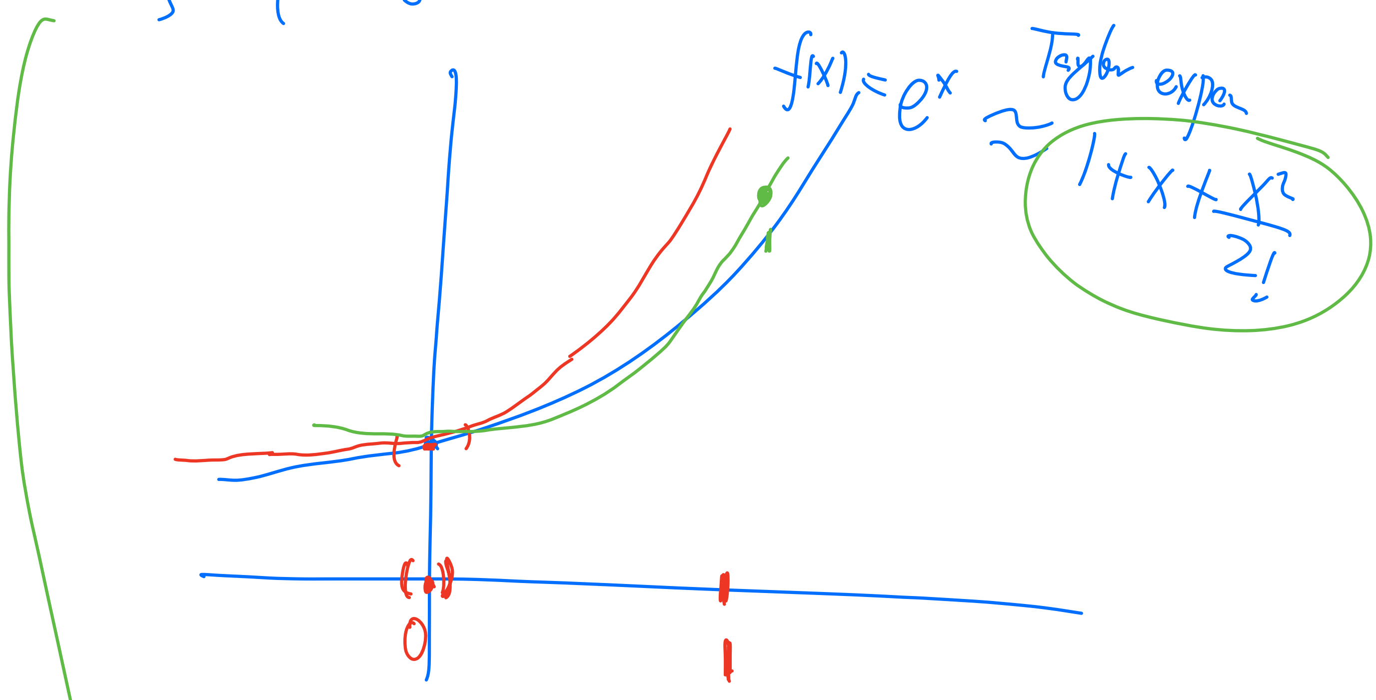


$$\begin{cases} \langle f - \vec{z}, 1 \rangle = 0 \\ \langle f - \vec{z}, x \rangle = 0 \\ \langle f - \vec{z}, x^2 \rangle = 0 \end{cases}$$

$$\begin{cases} \int_0^1 (e^x - c_0 - c_1 x - c_2 x^2)\, 1 = 0 \\ \int_0^1 (e^x - c_0 - c_1 x - c_2 x^2)(x) = 0 \\ \int_0^1 (e^x - c_0 - c_1 x - c_2 x^2)\, x^2 = 0 \end{cases}$$

$$\begin{cases} c_0 + \frac{1}{2} c_1 + \frac{1}{3} c_2 - (e-1) = 0 \\ \frac{1}{2} c_0 + \frac{1}{3} c_1 + \frac{1}{4} c_2 - 1 = 0 \\ \frac{1}{3} c_0 + \frac{1}{4} c_1 + \frac{1}{5} c_2 - (e-2) = 0 \end{cases}$$

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} e-1 \\ 1 \\ e-2 \end{bmatrix}$$



$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1.013 \\ 0.8511 \\ 0.8392 \end{bmatrix} \qquad \vec{z} = 1.013 + 0.8511x + 0.8392x^2$$

HW4: 6

$\dim(V) = 3$

Q:

Question 6. Let $P_2(\mathbb{R})$ be the space of polynomials with degree less or equal than 2. Let S be the subspace of the inner product space $P_2(\mathbb{R})$ generated by the polynomials $1 - x$ and $2 - x + x^2$ where $\langle f, g\rangle$ is defined to be $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. Find a basis for the orthogonal complement of S. (Hint: Let $f(x) = a + bx + cx^2$ be in $P_2(\mathbb{R})$. Use definition of orthogonal complement to set up a linear system.)

$$S = \text{Span}\{1-x,\ 2-x+x^2\} \qquad \dim S = 2$$

Is $\{1-x,\ 2-x+x^2\}$ orthogonal?

$$S^\perp = \{ f(x) = a+bx+cx^2 \in V \mid f(x) \perp S \}$$

$$\begin{cases} \langle f(x), 1-x \rangle = 0 \\ \langle f(x), 2-x+x^2 \rangle = 0 \end{cases} \iff \begin{cases} \int_0^1 (a+bx+cx^2)(1-x)\,dx = 0 \\ \int_0^1 (a+bx+cx^2)(2-x+x^2)\,dx = 0 \end{cases}$$

$$\iff \begin{cases} 40a + 15b + 8c = 0 \\ 110a + 55b + 37c = 0 \end{cases} \Rightarrow \begin{bmatrix} 40 & 15 & 8 \\ 110 & 55 & 37 \end{bmatrix}$$

?

$$\begin{cases} a = 23t \\ b = -120t \\ c = 110t \end{cases}$$

$$f(x) = t\left(\boxed{23 - 120x + 110x^2}\right)$$

$\{23 - 120x + 110x^2\}$ is a basis for $S^{\perp}$