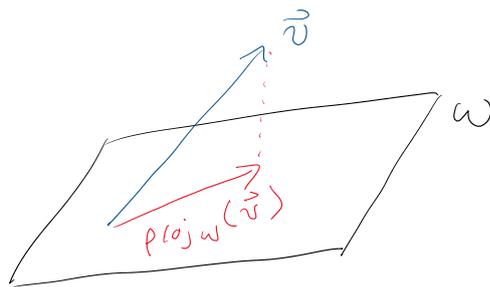


Section 7.3: Least Squares Approximation

Friday, December 01, 2023 12:56 PM

Best Approximation Theorem:

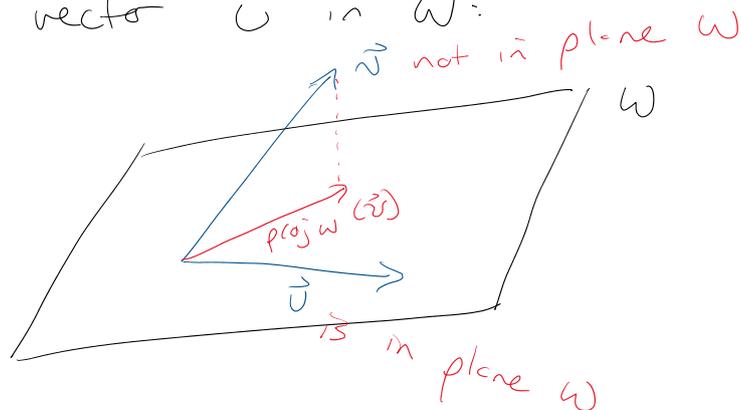
if W is a subspace of \mathbb{R}^n and \vec{v} is a vector in \mathbb{R}^n (which may or may not be within subspace W) then the best approximation to \vec{v} in W is the projection of \vec{v} onto W



note: Euclidean distance from \vec{v} to W is:

$$\underbrace{\|\vec{v} - \text{proj}_W(\vec{v})\|}_{\text{perp}_W(\vec{v})}$$

for any other vector \vec{u} in W :



$$\|\vec{v} - \text{proj}_W(\vec{v})\| \leq \|\vec{v} - \vec{u}\|$$

LHS is the vertical distance to the plane

RHS is the distance from tip of \vec{v} to any point in the plane.

distance to the
plane

tip of \vec{v} to any point
in the plane

why do we care? we can use this idea to
get an approximate solution to inconsistent
systems

example: let $\vec{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, $\vec{w}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}$

Find the best approximation to \vec{v} in the plane
 $W = \text{span}(\vec{w}_1, \vec{w}_2)$ and find the Euclidean
distance from \vec{v} to W .

answer: we first observe that $\vec{w}_1 \perp \vec{w}_2$

then

$$\text{proj}_W(\vec{v}) = \text{proj}_{\vec{w}_1}(\vec{v}) + \text{proj}_{\vec{w}_2}(\vec{v})$$

only true for \vec{w}_1, \vec{w}_2 an orthogonal set

-if not orthogonal, must Gram Schmidt
and make it orthogonal

$$= \frac{\vec{w}_1 \cdot \vec{v}}{\vec{w}_1 \cdot \vec{w}_1} \vec{w}_1 + \frac{\vec{w}_2 \cdot \vec{v}}{\vec{w}_2 \cdot \vec{w}_2} \vec{w}_2$$

$$= \frac{10}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \frac{9}{6} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$= 2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \frac{3}{2} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ 7 \\ 3 \end{bmatrix}$$

$$= \begin{matrix} \text{L0} & \text{L1} \\ \left[\begin{array}{c} 1/2 \\ -1 \\ 3/2 \end{array} \right] & \leftarrow \text{best approximation to } \vec{v} \text{ in } W \end{matrix}$$

Euclidean distance

$$\vec{v} - \text{proj}_W(\vec{v}) = \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix} - \begin{bmatrix} 1/2 \\ -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \\ 5/2 \end{bmatrix}$$

$$\begin{aligned} \|\vec{v} - \text{proj}_W(\vec{v})\| &= \sqrt{1/4 + 1 + 25/4} \\ &= \frac{\sqrt{30}}{2} \end{aligned}$$

least squares approximation

consider a system $A\vec{x} = \vec{b}$

$$A = [\vec{a}_1 \mid \vec{a}_2 \mid \vec{a}_3 \mid \dots \mid \vec{a}_n]$$

$\in A$ is made up of a bunch of column vectors

then

$$A\vec{x} = [\vec{a}_1 \mid \vec{a}_2 \mid \vec{a}_3 \mid \dots \mid \vec{a}_n] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 + \dots + x_n\vec{a}_n$$

the RHS is a linear combination of the columns of A

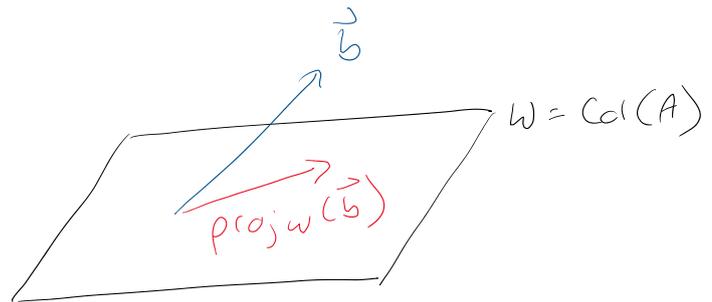
if $A\vec{x} = \vec{b}$ has a solution, then \vec{b} belongs to $\text{Col}(A)$

if $A\vec{x} = \vec{b}$ has a solution, then \vec{b} belongs to $\text{Col}(A)$

if $A\vec{x} = \vec{b}$ has no solution, then \vec{b} does not belong to $\text{Col}(A)$ and the closest we can come is

$$A\vec{x} = \text{proj}_W(\vec{b})$$

which will have a solution, called the least squares solution



the reason this is called "least squares" is because it minimizes

$$\|\vec{b} - A\vec{x}\|^2$$

then

$$A\vec{x}_{LS} = \text{proj}_W(\vec{b})$$

↑

the \vec{x} that minimizes the distance between the actual \vec{b} and the column space of A

but then

$$\vec{b} - \text{proj}_W(\vec{b}) = \vec{b} - A\vec{x}_{LS}$$



must be orthogonal to W where $W = \text{Col}(A)$

so that $\vec{b} - A\vec{x}_{LS}$ is in $\text{Null}(A^T)$

so that $\vec{b} - A\vec{x}_{LS}$ is in $\text{Null}(A^T)$ where $W = \text{Col}(A)$

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

$$A^T\vec{b} - A^T A\vec{x}_{LS} = \vec{0}$$

$$A^T A\vec{x}_{LS} = A^T\vec{b}$$

and if A has linearly independent columns, then $A^T A$ is invertible and

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

on final exam formula sheet

example: Find the least-squares solution for the inconsistent system $A\vec{x} = \vec{b}$ where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Also, calculate the least-squares error.

answer: $\vec{x}_{LS} = (A^T A)^{-1} A^T \vec{b}$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

so

$$\begin{aligned} \vec{x}_{LS} &= (A^T A)^{-1} A^T \vec{b} \\ &= \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} \\ &= \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

now for the least squares error $\|\vec{b} - A \vec{x}_{LS}\|$

$$\begin{aligned} \vec{b} - A \vec{x}_{LS} &= \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix} \end{aligned}$$

$$\|\vec{b} - A \vec{x}_{LS}\| = \sqrt{84} = 2\sqrt{21}$$

section 7.3: cont'd : 2023/12/04

example: Find the least squares solution to the

example: Find the least squares solution to the following inconsistent system.

$$\begin{cases} x_1 - x_2 = 4 \\ 3x_1 + 2x_2 = 1 \\ -2x_1 + 4x_2 = 3 \end{cases}$$

note: RREF of augmented matrix:

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$$

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

$$\begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

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

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{285} \begin{bmatrix} 21 & 3 \\ 3 & 14 \end{bmatrix}$$

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

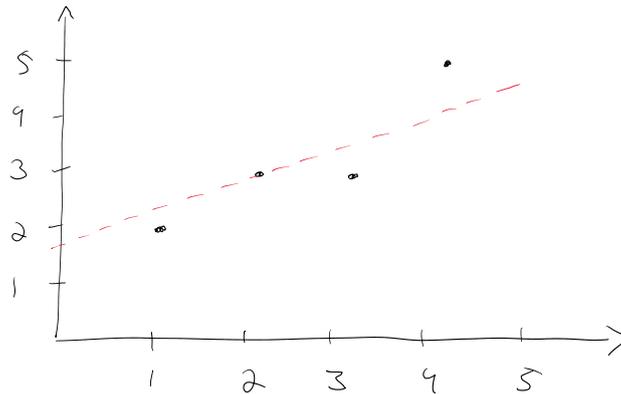
$$\vec{x}_{LS} = (A^T A)^{-1} A^T \vec{b}$$

$$= \frac{1}{285} \begin{bmatrix} 21 & 3 \\ 3 & 14 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

$$= \frac{1}{285} \begin{bmatrix} 51 \\ 143 \end{bmatrix} \approx \begin{bmatrix} 0.18 \\ 0.50 \end{bmatrix}$$

example: Find a least squares linear fit to the points $(1, 2), (2, 3), (3, 3), (4, 5)$

x	y
1	2
2	3
3	3
4	5



what is
the best-fit
line?

we want to fit $y = mx + b$ to these data points
(textbook uses $y = a + bx$)

$$\vec{y} = m\vec{x} + b$$

$$\begin{cases} 2 = 1m + b \\ 3 = 2m + b \\ 3 = 3m + b \\ 5 = 4m + b \end{cases}$$

← plug all pairs (x, y) into
 $y = mx + b$ to get
these equations

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 3 \\ 5 \end{bmatrix}$$

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

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 30 & 10 \\ 10 & 4 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{20} \begin{bmatrix} 4 & -10 \\ -10 & 30 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 2 & -5 \\ -5 & 15 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 37 \\ 13 \end{bmatrix}$$

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

$$= \frac{1}{10} \begin{bmatrix} 2 & -5 \\ -5 & 15 \end{bmatrix} \begin{bmatrix} 37 \\ 13 \end{bmatrix}$$

$$\begin{bmatrix} m \\ b \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 9 \\ 10 \end{bmatrix}$$

$$\text{so } m = 9/10 \\ b = 1$$

$$y = \frac{9}{10}x + 1$$

note: not tested

in general, if you are fitting $y = mx + b$, then you get

$$\begin{bmatrix} x_1 & | \\ x_2 & | \\ x_3 & | \\ x_4 & | \end{bmatrix} \begin{bmatrix} m \\ y \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

suppose you want to fit $y = ax^2 + bx + c$

$$\begin{bmatrix} x_1^2 & x_1 & | \\ x_2^2 & x_2 & | \\ x_3^2 & x_3 & | \\ x_4^2 & x_4 & | \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$