6.4 Best approximations; Least Squares

Exercise # 2) b)

Solve the system $A\vec{x}=\vec{b}$ $\qquad$ (1)

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

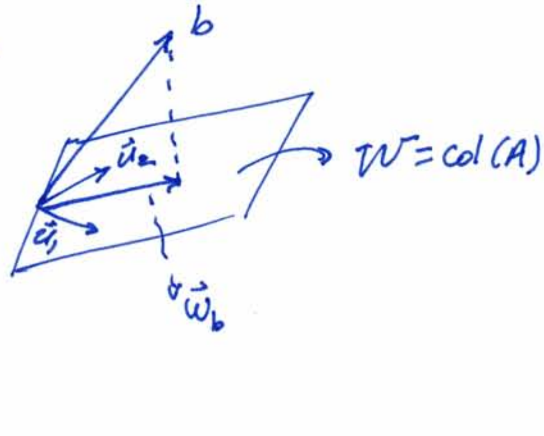
$$\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & -1 \\ 3 & 1 & 1 \end{bmatrix}\sim\begin{bmatrix} 1 & 1 & -1 \\ 2 & -2 & 2 \\ 3 & 1 & 1 \end{bmatrix}\sim\begin{bmatrix} 1 & 1 & -1 \\ 0 & -4 & 4 \\ 0 & -2 & 4 \end{bmatrix}\sim\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & -2 & 4 \end{bmatrix}\sim$$

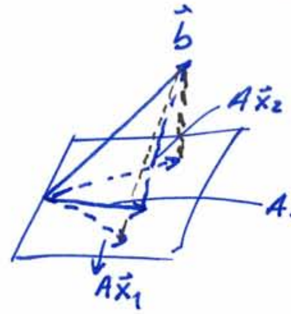
$$\sim\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & \boxed{0 \quad 2} \end{bmatrix} \rightarrow \text{inconsistent!}$$

This means that system (1) does not have a solution.

Then, the question or the vector $\vec{b}$ is not in $Col(A) =$

$$\text{Span}\left\{ \underset{\vec{u}_1}{\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}}, \underset{\vec{u}_2}{\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}} \right\} = W$$





Therefore, for any $\vec{x} \in \mathbb{R}^2$

$$\|A\vec{x} - \vec{b}\| > 0$$

The question is if there is $\vec{x}^* \in \mathbb{R}^2$ Such that

$$\|A\vec{x}^* - \vec{b}\| \text{ is minimum}$$

or $0 < \|A\vec{x}^* - \vec{b}\| \leq \|A\vec{x} - \vec{b}\|$, for any $\vec{x} \neq \vec{x}^*$ (2.1)

The answer is "yes" and the vector $\vec{x}^*$ happens to be a Solution of $\boxed{A\vec{x} = \text{proj}_W \vec{b},}$ where $W = \text{col}(A)$

This is the statement of next theorem.

Thm 6.4.1 (Best approx. thm)

- $W \subset V$ finite-dimensional subspace. of I.P.S. V
- $\vec{b} \in V$

Then, $\text{proj}_W \vec{b}$ is the best approximation to $\vec{b}$ from W

or $\|\vec{b} - \text{proj}_W \vec{b}\| < \|\vec{b} - \vec{w}\|$, for any $\vec{w} \in W$ Such that $\vec{w} \neq \text{proj}_W \vec{b}$. (2.2)

Since $A\vec{x}^* = \text{proj}_W \vec{b}$ and for each $\vec{w} \in W$ there is at least ~~one~~ $\vec{x}_w$ Such that $A\vec{x}_w = \vec{w}$, then (2.1) and (2.2) are equivalent.

Proof.- the proof is based on the orthogonality of the vectors $(\vec{b} - proj_W \vec{b})$ with any vector $\vec{w} \in W$. (thm 6.3.3)

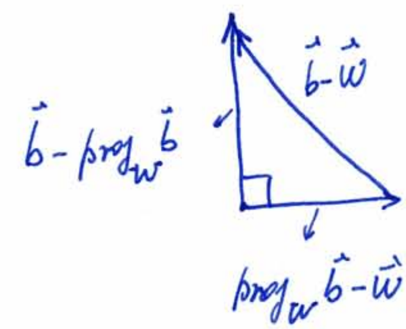
In fact,

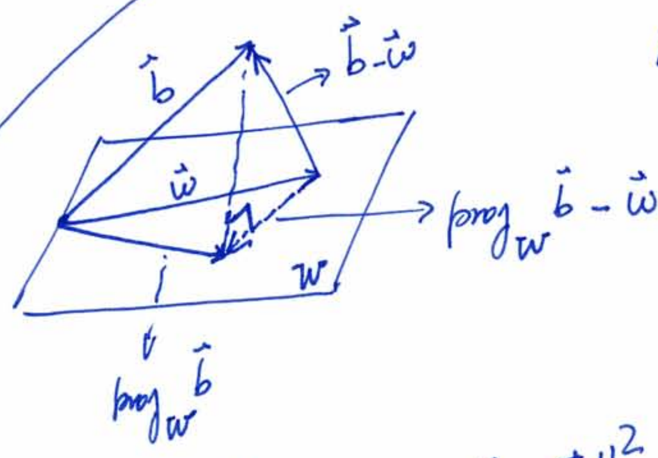
$$\vec{b} - \vec{w} = (\vec{b} - proj_W \vec{b}) + (proj_W \vec{b} - \vec{w})$$

Now, $(proj_W \vec{b} - \vec{w}) \in W$. Therefore, $(\vec{b} - proj_W \vec{b}) \perp (proj_W \vec{b} - \vec{w})$

Applying Pithagoras thm

$$\|\vec{b} - \vec{w}\|^2 = \|\vec{b} - proj_W \vec{b}\|^2 + \|proj_W \vec{b} - \vec{w}\|^2.$$





$$\Rightarrow \quad \|\vec{b} - proj_W \vec{b}\| < \|\vec{b} - \vec{w}\|, \quad \text{for any } \vec{w} \neq proj_W \vec{b}$$

Back to exercise # 2) b)

$\rightarrow$ $\vec{b} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ and $W = Col(A) = span\left\{ \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$

It can be shown (verify it!) that

$proj_W \vec{b} = \begin{pmatrix} 10/21 \\ -5/21 \\ 13/21 \end{pmatrix}$ (It requires finding an orthogonal basis for Col(A) using G-S process and then using the formula for $proj_W \vec{b}$ given by thm 6.3.4)

Therefore, a vector $\vec{x}^* \in \mathbb{R}^2$ that minimizes

$0 < \|A\vec{x} - \vec{b}\|$

can be obtained by solving the system of equations

$$\begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 10/21 \\ -5/21 \\ 13/21 \end{bmatrix}$$

$\Rightarrow x_1^* = 3/7, \quad x_2^* = -2/3$, unique soln.

Least squares Problem:

Given a linear system $A\vec{x} = \vec{b}$, A $m \times n$

Find $\vec{x}^*$ if possible that minimizes $\|A\vec{x} - \vec{b}\|$

with respect to the Euclidean Inner Product.

Such $\vec{x}^*$ is called a least squares solution of the system and $\|b - A\vec{x}^*\|$ is called the least squares error.

In general, consider the system

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

If $\qquad W = col(A)$

Then, from the best approx. thm the closest vector in W to $\vec{b}$ is

$$proj_W \vec{b}$$

Therefore, any least squares solution $\vec{x}^*$ of $A\vec{x} = \vec{b}$ must satisfy

$$A\vec{x}^* = proj_W \vec{b} \qquad (5.0)$$

Notice that

$$\|\vec{b} - proj_W \vec{b}\| < \|\vec{b} - \vec{w}\| \qquad (5.1)$$

is equivalent to

$$\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\|, \quad \text{for all } \vec{x}. \qquad (5.2)$$

The equality sign in (5.2) happens when there is more than one soln. to (5.0).

Alternative to solve least squares problems

Given $A\vec{x}=\vec{b}$ consistent or not, (6.1)

its least squares soln. is obtained by solving the least squares problem: $A\hat{x} = \text{proj}_W \vec{b}$, $W = \text{col}(A)$. (6.2)

We can avoid the comp. of $\text{proj}_W \vec{b}$. In fact,

from (6.2)

$$\vec{b} - A\hat{x} = \vec{b} - \text{proj}_W \vec{b}$$

Applying A^T to both sides

$$A^T(\vec{b} - A\hat{x}) = A^T(\vec{b} - \text{proj}_W \vec{b})$$

Since $(\vec{b} - \text{proj}_W \vec{b}) \in \text{Col}(A)^{\perp} = \text{Null}(A^T)$, $= (\text{row}(A^T))^{\perp}$

$$\Rightarrow \quad A^T(\vec{b} - \text{proj}_W \vec{b}) = 0$$

$$\Rightarrow \quad A^T(\vec{b} - A\hat{x}) = 0 \iff \boxed{A^TA\hat{x} = A^T\vec{b}} \quad (6.3)$$

Normal equs. or Assoc. normal syst. to $A\hat{x}=\vec{b}$.

Therefore, any solution $\vec{x}$ of

$$A\hat{x} = \mathrm{proj}_W \vec{b} \qquad (7.1)$$

is also a soln. of

$$A^T A \hat{x} = A^T \vec{b} \qquad (7.2)$$

Conversely, if $\hat{x}$ is a soln. of (7.2), then

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

then, $(\vec{b} - A\hat{x}) \in Null(A^T) = [Col(A)]^{\perp}$

If we call $\vec{\omega}_2' = \vec{b} - A\hat{x}$

then $$\vec{b} = \vec{\omega}_2' + A\hat{x}$$

, $\vec{\omega}_2' \in [Col(A)]^{\perp}$ and $A\hat{x} \in Col(A) = W$

Now, from the "projection thm"

$$\vec{b} = \vec{\omega}_1 + \vec{\omega}_2\,, \quad \text{where } \vec{\omega}_1 = \mathrm{proj}_W \vec{b} \text{ and } \vec{\omega}_2 = \vec{b} - \mathrm{proj}_W \vec{b}$$

and this decomp. is unique.

then, $A\hat{x} = \vec{\omega}_1 = \mathrm{proj}_W \vec{b}$ and $\omega_2' = \vec{\omega}_2$

Therefore, $\hat{x}$ is a soln. of (7.1)

The above results can be summarized in the following theorem.

<u>Theorem 6.4.2</u> (Diff. than book)

Given the linear system $A\vec{x}=\vec{b}$ and calling $W = col(A)$.

The two linear systems

$$\text{i)} \quad A^T A\vec{x} = A^T\vec{b}$$

(Associated linear system to $A\vec{x}=\vec{b}$ or normal system)

and

$$\text{ii)} \quad A\vec{x} = proj_W \vec{b}$$

are equivalent. They are also consistents and all solutions are least squares soln. of $A\vec{x}=\vec{b}$.

The systems are consistents, because $proj_W \vec{b} \in col(A) = W$

Also, from best approx. theorem, any soln. of (7.1) is a least squares soln. of $A\vec{x}=\vec{b}$.

Back to our original problem #2) (b)

Solve the system

$$\begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad W = \mathrm{Col}(A)$$

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

We found this system inconsistent!

We also found that there is a unique least squares soln. $\vec{x}^*$ satisfying the least squares problem

$$A\vec{x}^* = \mathrm{proj}_W \vec{b}$$

or

$$\begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 10/21 \\ -5/21 \\ 13/21 \end{bmatrix}$$

$$\Rightarrow \quad x_1^* = 3/7, \quad x_2^* = -2/3.$$

Using thm 6.4.2, the soln. $\vec{x}^*$ can also be obtained by solving the system

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

or

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

or

$$\begin{bmatrix} 14 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \end{bmatrix} \Rightarrow \quad x_1^* = 6/14 = 3/7 \checkmark$$

and

$$x_2^* = -4/6 = -2/3. \checkmark$$

This procedure is much simpler than the previous one of solving $A\vec{x}^* = \text{proj}_W \vec{b}$.

To obtain $\text{proj}_W \vec{b}$, we compute

$$A\vec{x}^* = \begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 3/7 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 10/21 \\ -5/21 \\ 13/21 \end{bmatrix}$$

and now, we can evaluate the least squares error.

$$\|A\vec{x}^* - \vec{b}\| = \left\| \begin{bmatrix} 10/21 - 2 \\ -5/21 + 1 \\ 13/21 - 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -32/21 \\ 16/21 \\ -8/21 \end{bmatrix} \right\|$$

$$= \sqrt{\left(\frac{32}{21}\right)^2 + \left(\frac{16}{21}\right)^2 + \left(\frac{8}{21}\right)^2}$$

Exercise #9 ⓐ

$\vec{u} = (6,3,9,6)$, $\vec{v}_1 = (2,1,1,1)$, $\vec{v}_2 = (1,0,1,1)$
$\vec{v}_3 = (-2,-1,0,-1)$

Find orthogonal projection of $\vec{u}$ on $W = \text{span}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\})$.

If we define

$$A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$$

then if $\vec{x}^*$ is a least squares Solution of

$$A\hat{x} = \vec{u}$$

the projection can be computed as

$$A\vec{x}^* = \text{proj}_W \vec{u}$$

In order to find $\vec{x}^*$, we solve the normal system

$$A^T A \vec{x} = A^T \vec{u}$$

Notice

$$A^T A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ -2 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 7 & 4 & -6 \\ 4 & 3 & -3 \\ -4 & -1 & 6 \end{bmatrix}$$

and

$$A^T \vec{u} = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ -2 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 30 \\ 21 \\ -21 \end{bmatrix}$$

Then, we need to find $\vec{x}^*$ satisfying

$$\begin{bmatrix} 7 & 4 & -6 \\ 4 & 3 & -3 \\ -4 & -1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 30 \\ 21 \\ -21 \end{bmatrix}$$

Soln: $x_1 = -6, \quad x_2 = 9, \quad x_3 = -6$

Therefore, $\text{proj}_W \vec{u} = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -6 \\ 9 \\ -6 \end{bmatrix} = \begin{bmatrix} 9 \\ 0 \\ 3 \\ 9 \end{bmatrix}$

<u>Thm 6.4.3</u> A $m \times n$ matrix. The following statements are equivalent.

a) A has linearly indep. column vectors.

b) A^TA is invertible

<u>Proof:-</u>

a) $\Rightarrow$ b)

To show that A^TA is invertible, we will prove the equivalent statement

$$A^TA\vec{x} = \vec{0}$$

only has the trivial soln.

In fact,

if $A^TA\vec{x} = 0 \Rightarrow A\vec{x} \in Null(A^T)$

But also $A\vec{x} \in Col(A)$

From thm. 6.2.4 $Null(A^T) \cap Col(A) = \{\vec{0}\}$

because they are orthogonal complements

$\Rightarrow$ $A\vec{x} = \vec{0}$, but A has n lin. indep column vectors, then $\vec{x} = \vec{0}$ is the only possible soln. of $\boxed{A^TA\vec{x} = \vec{0}}$

b) $\Rightarrow$ a) If $\vec{x}$ is a soln. of $A\vec{x} = \vec{0}$ ($\Leftrightarrow x_1\vec{c}_1 + + x_n\vec{c}_n = \vec{0}$) then $A^TA\vec{x} = A^T(A\vec{x}) = A^T\vec{0} = \vec{0}$. Now, A^TA is invertible, then only possible $\vec{x}$ is $\vec{x} = \vec{0} \Rightarrow \{\vec{c}_1, .., \vec{c}_n\}$ is lin. indep. ✓

Thm 6.4.4 — A $m\times n$ matrix

— A has n linearly indep. columns. $W = \mathrm{col}(A)$

then, the System $A\vec{x} = \vec{b}$

has a unique least squares soln.

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

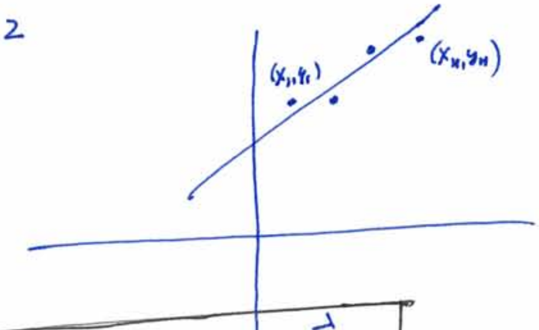
and $\mathrm{proj}_W \vec{b} = A\vec{x} = A\,(A^T A)^{-1} A^T \vec{b}$

Proof.- Follows from 6.4.3.

6.5 Least Squares Fitting Data (LAB)

Consider a set of points:

$$(x_1,y_1),\ldots,(x_n,y_n) \in \mathbb{R}^2$$



Suppose we want to fit a line to these points

$y = a + bx$

$$\begin{cases} y_1 = a + bx_1 \\ y_2 = a + bx_2 \\ \vdots \\ y_n = a + bx_n \end{cases} \longrightarrow \boxed{\overset{M}{\overset{\shortparallel}{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix}}} \overset{\vec{v}}{\overset{\shortparallel}{\begin{bmatrix} a \\ b \end{bmatrix}}} = \overset{\vec{y}}{\overset{\shortparallel}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}}}$$

$$M\vec{v} = \vec{y}$$

For noncollinear points, this system is inconsistent!

It's possible to show that the <u>two columns of M are lin. indep.</u>, except when $(x_1,y_1),\ldots,(x_n,y_n)$ lie on a vertical line.

Therefore, the least squares soln. $\vec{v}^* = \begin{pmatrix} a^* \\ b^* \end{pmatrix}$ is unique and satisfies

$$M^T M \vec{v}^* = M^T \vec{y}$$

The sought line is given by

$$\boxed{y = a^* + b^* x}$$