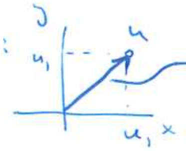
MATH 312 LECTURE 8 : Orthogonality and Projections.

Chapter 4: Orthogonality.

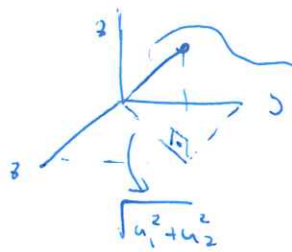
I) Orthogonal complements (Sec. 4.1)

Length of a vector: $\|u\| = (u_1^2+u_2^2+\dots+u_n^2)^{1/2} = (u\cdot u)^{1/2} = (u^T u)^{1/2}$.

$\llcorner$ Pythagoras's th.:

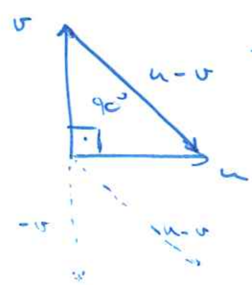


$\longrightarrow \text{length}^2 = u_1^2 + u_2^2$



$\longrightarrow \text{length}^2 = u_3^2 + \left(\sqrt{u_1^2+u_2}\right)^2 = u_1^2+u_2^2+u_3^2$.

• Orthogonal vectors:



$\Rightarrow \|u-v\|^2 = \|u\|^2 + \|v\|^2 \Rightarrow$

$$\left.\begin{aligned} \Rightarrow\ & u_1^2+u_2^2+\dots+u_n^2 + \\ & + v_1^2+v_2^2+\dots+v_n^2 = \underbrace{(u_1-v_1)^2+\dots+(u_n-v_n)^2} \\ & \qquad = u_1^2+v_1^2-2u_1v_1+\dots+u_n^2+v_n^2-2u_nv_n \end{aligned}\right\} \Rightarrow$$

$$\Rightarrow -2u_1v_1 - 2u_2v_2 - \dots - 2u_nv_n = 0 \text{, i.e., } u\cdot n = 0 \,!$$

So, two vectors u, v are orthogonal (or perpendicular) if and only if $u \cdot v = 0$.

<u>Remark</u>: Recall that $u \cdot v = u^T v = v^T u$.

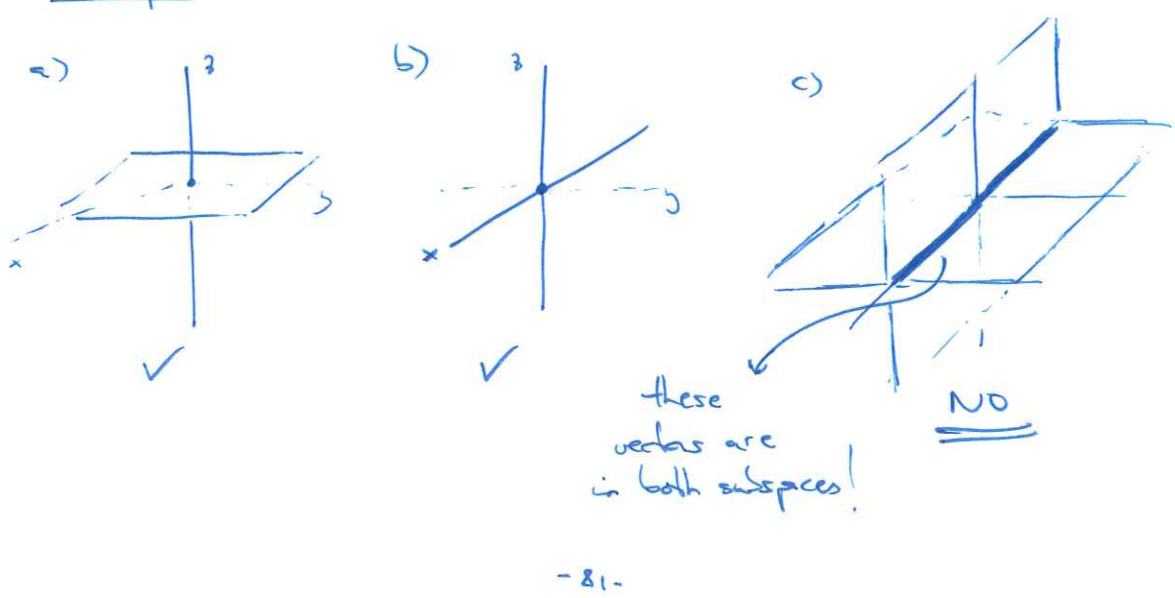
$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \dots + u_n v_n .$$

<u>Remark</u>: With this definition the zero vector is orthogonal to any vector.

• <u>Def</u>: Two subspaces U and V of $\mathbb{R}^n$ are (perpendicular) orthogonal if every vector $u \in U$ is orthogonal to every vector $v \in V$.

$U \perp V$ if $\forall u \in U, \forall v \in V$, then $u \cdot v = 0$.

<u>Examples</u>: Q:



• Def: Orthogonal Complement

Let V be a subspace of $\mathbb{R}^n$. Then the collection of all vectors orthogonal to V is also a subspace of $\mathbb{R}^n$, denoted by $V^\perp$.

Examples: 2: a), but $\underbrace{\text{not b)}}_{!}$ and not c)

V and V^T fill $\mathbb{R}^n$.

The dimension of a subspace + dimension of its orthog. complement = dimension of ambient vector space.

((*) Exercise pág. 85)

• Fundamental Theorem of Linear Algebra

Let A $m\times n$ matrix. Then

1) $N(A)$ and $C(A^T)$ are subspaces of $\mathbb{R}^n$

2) $N(A^T)$ and $C(A)$ are subspaces of $\mathbb{R}^m$,

$\dim N(A) + \dim C(A^T) = n$

$\dim C(A) + \dim N(A^T) = m$

and

$$\boxed{\begin{aligned} N(A) &= C(A^T)^\perp \\ C(A) &= N(A^T)^\perp \end{aligned}}$$

Why the second part?

Let's see:

- $\underline{N(A) = C(A^T)^\perp}$:

$N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$, $C(A^T) = \{\text{span of rows of } A\}$.

So, if $x \in N(A)$, then

$$Ax = 0 \rightsquigarrow \begin{bmatrix} - \text{row}1 - \\ \vdots \\ - \text{row}\, m - \end{bmatrix} \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ that is,}$$

$$\left.\begin{array}{l} \text{row}1 \cdot x = 0 \\ \text{row}2 \cdot x = 0 \\ \vdots \\ \text{row}\,m \cdot x = 0 \end{array}\right\} \to$$

x is orthogonal to __all__ the rows of A, therefore x is orthogonal to all their linear combinations, thus x is orthogonal to $C(A^T)$.

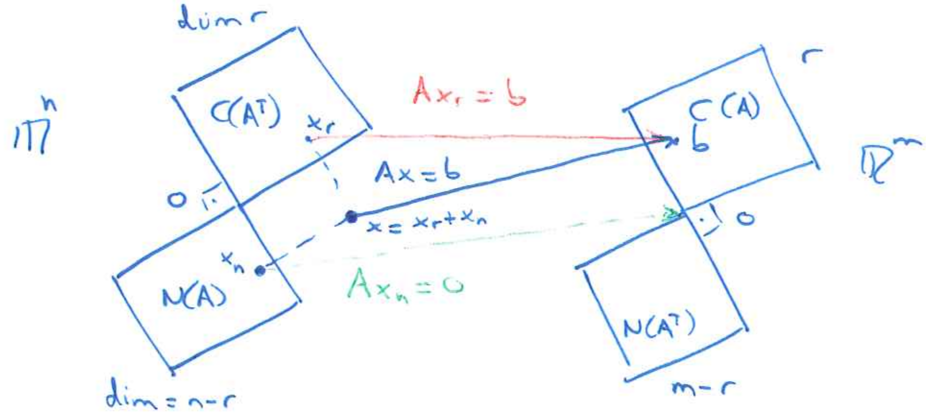
Since we have check that $N(A)$ is orthogonal to $C(A^T)$ __and__ we know that

$$\dim N(A) + \dim C(A^T) = n,$$

we conclude that they are orthogonal complements//.

- $C(A)$ and $N(A^T)$ exercise for home//.

Full picture



Given any x in $\mathbb{R}^n$, now we can split it $x = \underbrace{x_r}_{\text{in } C(A^T)} + \underbrace{x_n}_{\text{in } N(A)}$

Important: Given a b in $C(A)$, there might be many x in $\mathbb{R}^n$ such that $Ax = b$ (if $N(A)$ is not just the zero).

However, there is always <u>just one</u> x_r in $C(A^T)$ such that $Ax_r = b$.

I.e.,

We will be able to define an "inverse" from $C(A)$ to $C(A^T)$.

• Exercise: Let V be the subspace define by $x+y+z=0$ (plane).

Find a basis for V and for $V^\perp$. (*) (Hacer antes de FTLA).

Q: $\dim V^\perp$?

Solution:

V is the nullspace of $A=[1\ 1\ 1]$:

$[1\ 1\ 1]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$. So a basis for A is easy to find:

$[\boxed{1}\ 1\ 1]$ ↑↑ free

$\left.\begin{array}{l} x_2=\alpha \\ x_3=\beta \end{array}\right\} \Rightarrow x_1=-\alpha-\beta$

Basis for $V = \left\{ \underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}} \right\}$

(of course two vectors since it was a plane!).

What is $V^\perp$?

All vectors orthogonal to all vectors in V. So, let $x \in V^\perp$.

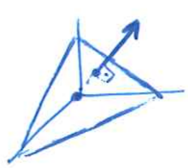
If x is perpendicular to the basis vectors of V, it will be perp. to all, so the equations for $V^\perp$ are:

$x \cdot \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = 0 \rightsquigarrow -x_1+x_2=0$

$x \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0 \rightsquigarrow -x_1+x_3=0$

And a basis is given by the nullspace of:

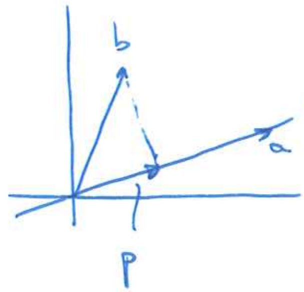
$$\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & -1 \end{bmatrix} \rightsquigarrow \text{Basis for } V^\perp = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$



(we could have seen this from the very start!).

II) Projection (Sec. 4.2)

1) Projection onto a line:



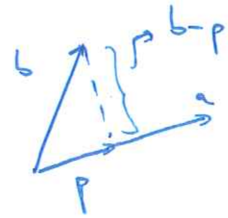
$V \equiv$ subspace spanned by a (a line)

Given b and a, how to find $p \equiv$ projection of b onto a?

Well, p is on the line, so it has to be some multiple of a,

$p = \hat{x} a$, where $\hat{x} \in \mathbb{R}$ is what we want to find.

- How?



We see that $b-p$ is perpendicular to a, that is,

$$\left.\begin{array}{l} a \cdot (b-p) = 0 \\ p = \hat{x} a \end{array}\right\} \Rightarrow a \cdot (b - \hat{x} a) = 0 \Rightarrow a \cdot b = \hat{x} \underbrace{a \cdot a}_{\text{number} \neq 0} \Rightarrow$$

$$\Rightarrow \hat{x} = (a \cdot a)^{-1} a \cdot b = \frac{a \cdot b}{a \cdot a}$$

So our projection is $p = \dfrac{a \cdot b}{a \cdot a} a$, or using matrix notation

$$p = \underbrace{\frac{a^T b}{a^T a}}_{\text{number}} \underset{\text{vector}}{a} = a \frac{a^T b}{a^T a} = \underbrace{\frac{a a^T}{a^T a}}_{\text{matrix}} \underset{\text{vector}}{b}$$

• In summary, the projection of b onto a line given by a is:

$$p = \frac{a a^T}{a^T a} b = \frac{a \cdot b}{\|a\|^2} a$$, where $P = \frac{a a^T}{a^T a}$ is the so-called projection matrix.

• <u>Example:</u> Find the projection of $b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ onto the line given by $\left.\begin{matrix} x - y = 0 \\ x - z = 0 \end{matrix}\right\}$

<u>Sol:</u> We first find a basis for that line (i.e. our vector "a").

$$\left.\begin{matrix} x - y = 0 \\ x - z = 0 \end{matrix}\right\} \to \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad z = \alpha \to x = \alpha, y = \alpha$$

So our line is given by $\underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{a}$ (we can write this directly from the eqr.).

So the projection is:

$$p = \frac{a \cdot b}{\|a\|^2} a = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

• What is in the example the projection matrix?

$$P = \frac{aa^T}{a^Ta} = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}$$

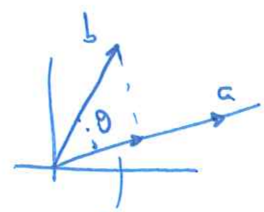
• What is $N(P)$, $C(P)$? P^2?

Nullspace → The orthogonal complement of the line!

Since we are in $\mathbb{R}^3$, it will be a plane: $x+y+z=0$.

Column space → The line! (of course, all vectors get projected onto the line when multiplied by P).

• Exercise (Home)

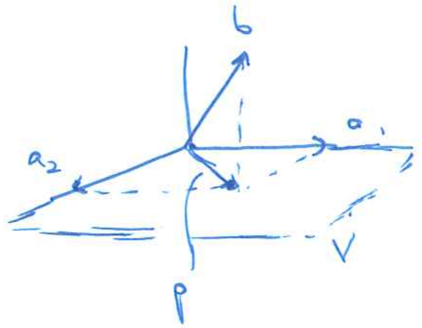


Find the projection of b onto a using what you knew on high school: $\left.\begin{aligned} \cos\theta &= \frac{\|p\|}{\|b\|} \\ a\cdot b &= \|a\|\|b\|\cos\theta \end{aligned}\right\} \to p$?

3) Projection onto a subspace:

Consider first a plane. Choose a basis for the plane: $\{a_1, a_2\}$ in $\mathbb{R}^n$.

column vectors.



Once we know a_1, a_2, given b, what is its projection onto the plane, p?

→ It has to be a linear combination of a_1 and a_2:

$$p = \hat{x}_1 a_1 + \hat{x}_2 a_2 = \begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} \equiv A\hat{x}$$

→ $b - p$ is orthogonal to both a_1 and a_2 (thus to all the plane):

$$a_1 \cdot (b-p) = 0 \rightsquigarrow a_1^T(b - A\hat{x}) = 0 \rightsquigarrow \begin{bmatrix} - & a_1^T & - \end{bmatrix} \begin{bmatrix} b - A\hat{x} \end{bmatrix} = 0$$

$$a_2 \cdot (b-p) = 0 \rightsquigarrow a_2^T(b - A\hat{x}) = 0 \rightsquigarrow \begin{bmatrix} - & a_2^T & - \end{bmatrix} \begin{bmatrix} b - A\hat{x} \end{bmatrix} = 0 \rightsquigarrow$$

$$\rightsquigarrow \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \end{bmatrix} \begin{bmatrix} b - A\hat{x} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \to A^T(b - A\hat{x}) = 0 \Rightarrow$$

$$\Rightarrow \boxed{A^T A \hat{x} = A^T b}$$ We need to solve for $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$

- So, $\hat{x} = (A^TA)^{-1}A^Tb$, and therefore $(p = A\hat{x})$

$$p = A(A^TA)^{-1}A^Tb \quad , \text{ with } \boxed{P = A(A^TA)^{-1}A^T} \text{ projection matrix.}$$

<u>Q</u>: What is the nullspace of P?

In this case, the line orthogonal to the plane!

- <u>Remarks</u>: 1) $P^2 = P$ (obviously, once a vector is projected, the next projection will be the same).

2) A^TA is always a square $n \times n$ matrix and symmetric. (if A is $m \times n$)

↳ Does it always have an inverse?

No, but for practical problem, almost always.

It has an inverse if and only if the columns of A are l.i.

Example: Find the projection matrix P onto the plane $x+y+z=0$, and the projection of $b=\begin{bmatrix}1\\0\\0\end{bmatrix}$ onto the plane.

Solution:

- First, we need a basis of the plane: $\left\{ \underbrace{\begin{bmatrix}1\\-1\\0\end{bmatrix}}_{=a_1}, \underbrace{\begin{bmatrix}0\\1\\-1\end{bmatrix}}_{a_2} \right\}$

- Second, we construct $A=\begin{bmatrix}1&0\\-1&1\\0&-1\end{bmatrix}$.

- Finally, the projection matrix is $P=A(A^TA)^{-1}A^T$,

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

$$\left[\begin{array}{cc|cc}2&-1&1&0\\-1&2&0&1\end{array}\right]\to\begin{bmatrix}2&-1&1&0\\0&3/2&1/2&1\end{bmatrix}\to\begin{bmatrix}2&-1&1&0\\0&1&1/3&2/3\end{bmatrix}\to\begin{bmatrix}2&0&4/3&2/3\\0&1&1/3&2/3\end{bmatrix}\to$$

$$\to\begin{bmatrix}1&0&2/3&1/3\\0&1&1/3&2/3\end{bmatrix}\to(A^TA)^{-1}=\frac{1}{3}\begin{bmatrix}2&1\\1&2\end{bmatrix}$$

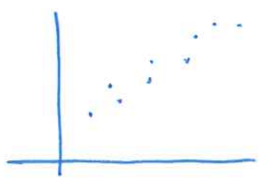
Thus,

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

- The projection of b: $p=\frac{1}{3}\begin{bmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}=\frac{1}{3}\begin{bmatrix}2\\-1\\-1\end{bmatrix}$

III) Least-Squares

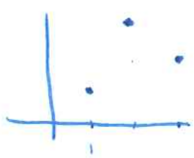
We want to find a model for some data,



for example, a line.

Of course it is impossible in most situation to find a line through all points, but still some lines seem better than others: we look for the "best."

Example:



$(x_1, y_1) = (1,1)$
$(x_2, y_2) = (2,3)$
$(x_3, y_3) = (3,2)$

$y = mx + n$, $\underline{\underline{m, n?}}$

Let's try:
$$\left.\begin{aligned} y_1 &= m x_1 + n \\ y_2 &= m x_2 + n \\ y_3 &= m x_3 + n \end{aligned}\right\} \longrightarrow \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

In the example, $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} n \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$

Solving this system is finding a line going through all three points!

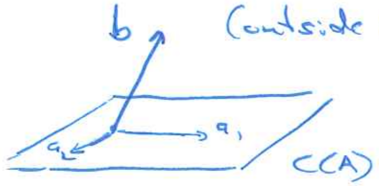
↳ <u>No solution</u>.

That is, if we denote $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}$, $v = \begin{bmatrix} n \\ m \end{bmatrix}$, $b = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$,

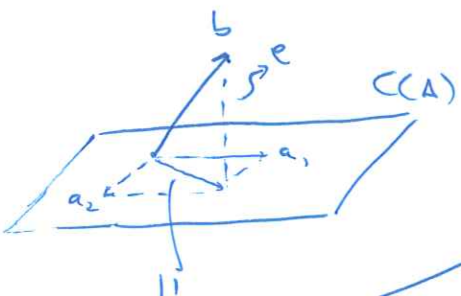
the system $Av = b$ has no solutions.

$\hookrightarrow$ That is, b is not in $C(A)$.

Intuitively we have something that looks like the following



We cannot obtain b as a linear combination of a_1, a_2 (the columns of A). The "best" we can do is to obtain a linear combination of a_1 and a_2 giving a vector b' as close as possible to b: that is, the projection of b onto $C(A)$.



So what we ~~solve~~ want to solve now is $\| Av = b' \|$, instead of $\underbrace{Av = b}_{\text{impossible}}$.

The error will be $e = b - b'$. What is the main characteristic of this error? It is perpendicular to the "plane" (the subspace $C(A)$ in general).

- We are saying that e is in $C(A)^{\perp}$!

But we know that $C(A)^{\perp} = N(A^T)$!!

So, by definition, if $e \in N(A^T)$, that means that

$$A^T e = 0 \rightarrow A^T(b-b') = 0 \rightarrow A^T b = A^T b'.$$

And we were looking for some combination of a_1, a_2 that gives b': $Av = b'$,

so we have the equation $\| A^T A v = A^T b \|$ Normal equation.

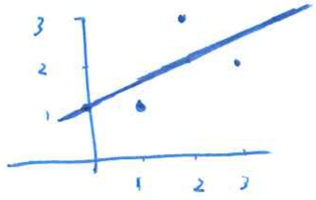
- In our example:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad v = \begin{bmatrix} n \\ m \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \rightarrow (A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix}.$$

So $$\begin{bmatrix} n \\ m \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 14 & -6 \\ -6 & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 13 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$$

So our line was $y = \frac{x}{2} + 1$



$x = 0 \rightarrow y = 1$

$x = 2 \rightarrow y = 1$

- What are the errors?

First point: $(x_1, y_1) = (1, 1)$

What we can obtain: $x_1 = 1 \to \tilde{y}_1 = \frac{1}{2} + 1 = \frac{3}{2}$ $\Big\} \Rightarrow$

$\Rightarrow e_1 = 1 - \frac{3}{2} = \frac{-1}{2}$

Second point: $(x_2, y_2) = (2, 3)$

$x_2 = 2 \to \tilde{y}_2 = \frac{2}{2} + 1 = 2$ $\Big\} e_2 = 3 - 2 = 1$

$\vdots$

Faster: $e = b - \underset{\substack{\uparrow \\ Av}}{b'} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 3/2 \\ 2 \\ 5/2 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix}$

$\to$ What if our data points are in $\mathbb{R}^3$?

We might want to fit a plane instead of a line:

$z = a + bx + cy$

We can proceed similarly to find $A = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ & \vdots & \\ 1 & x_n & y_n \end{bmatrix}$, $v = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $b = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$.

• Why is this called "least squares"?

When we fit our line, we are making the following errors:

$$\left.\begin{aligned} e_1 &= y_1 - m x_1 - n \\ e_2 &= y_2 - m x_2 - n \\ &\vdots \end{aligned}\right\} \text{ So total error } \|e\| .$$

We are indeed minimizing $\|e\|^2 \to$ Calculus.

Let's see it in our example: (... finish in class).