HW 7 Solutions

4.3.5. We have 4 data points

$$\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} t_1 \\ 0 \end{pmatrix}, \begin{pmatrix} t_2 \\ 8 \end{pmatrix}, \begin{pmatrix} t_3 \\ 8 \end{pmatrix}, \begin{pmatrix} t_4 \\ 20 \end{pmatrix}$$

[Note: The times won't matter so Strang didn't even tell us what they are.]

Fit these points to a <u>horizontal</u> line of the form $C = b$.

The equations

$$\begin{aligned} C &= 0 \\ C &= 8 \\ C &= 8 \\ C &= 20 \end{aligned} \quad \leftrightarrow \quad \underset{A}{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}} \underset{\vec{x}}{(C)} = \underset{\vec{b}}{\begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}}$$

obviously have no solution, so we try the normal equation $A^T A \hat{x} = A^T \vec{b}$:

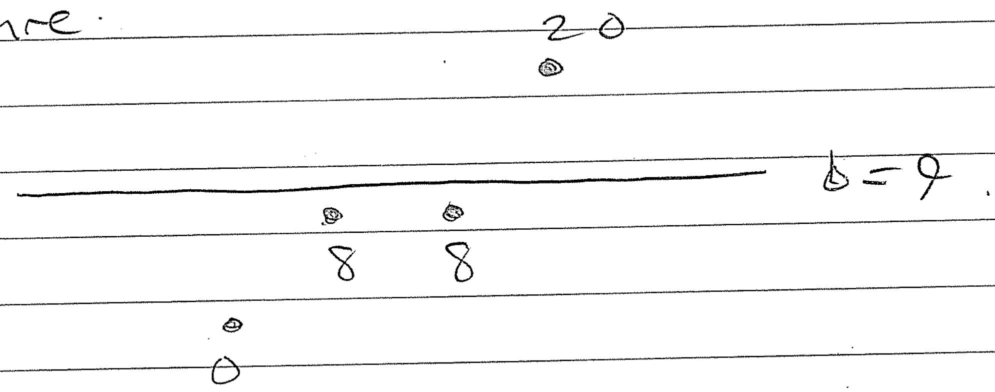
$$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} (C) = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 8 \\ 8 \\ 12 \end{pmatrix}$$

$$4C = 0 + 8 + 8 + 20$$

$$C = \frac{(0 + 8 + 8 + 20)}{4} = 9.$$

This $C = 9$ is just the average of the b values!

Picture:



4.3.7. Find the line $Dt = b$ through the origin closest to

$$\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 \\ 20 \end{pmatrix}$$

[Now Strang tells us the t values]

The silly equations are

$$\begin{aligned} 0D &= 0 \\ 1D &= 8 \\ 3D &= 8 \\ 4D &= 20 \end{aligned} \longrightarrow \begin{pmatrix} 0 \\ 1 \\ 3 \\ 4 \end{pmatrix} (D) = \begin{pmatrix} 0 \\ 8 \\ 8 \\ 20 \end{pmatrix}$$

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

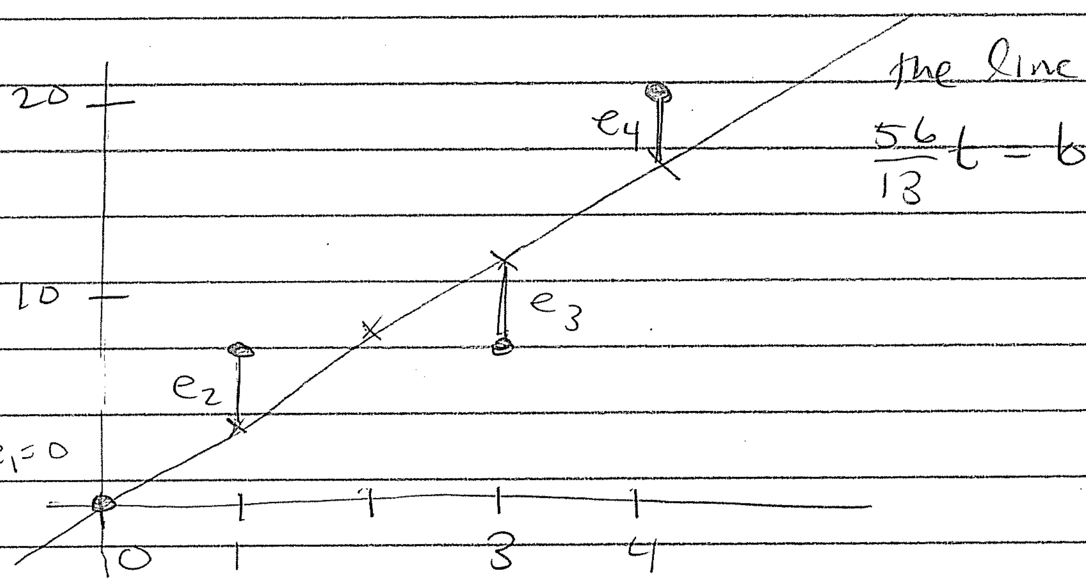
There is no solution, so we solve the normal equation $A^TA\hat{x} = A^T\vec{b}$

$$\begin{pmatrix}0 & 1 & 3 & 4\end{pmatrix}\begin{pmatrix}0\\1\\3\\4\end{pmatrix}(D) = \begin{pmatrix}0 & 1 & 3 & 4\end{pmatrix}\begin{pmatrix}0\\8\\8\\20\end{pmatrix}$$

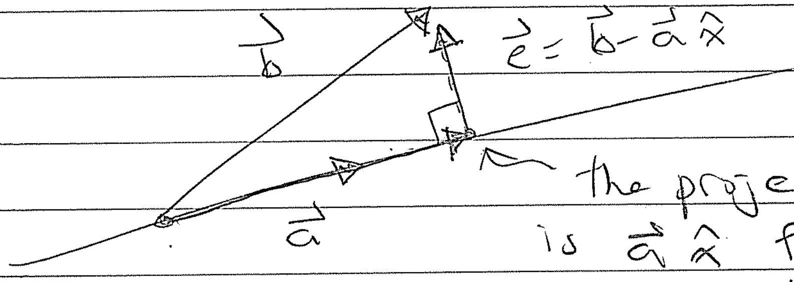
$$26D = 112$$

$$D = 112/26 = 56/13$$

Picture:



4.3.12. Project the data $\vec{b} = (b_1, \ldots, b_m)$ onto the line through $\vec{a} = (1, 1, \ldots, 1)$.



the projection is $\vec{a}\hat{x}$ for some <u>number</u> $\hat{x}$.

By orthogonality, we must have

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

$$\vec{a}^T\vec{b} - \vec{a}^T\vec{a}\,\hat{x} = 0$$

$$\vec{a}^T\vec{b} = \vec{a}^T\vec{a}\,\hat{x}.$$

(a) Solve this. $\vec{a}^T\vec{a}\,\hat{x} = \vec{a}^T\vec{b}$ is

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \hat{x} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$$m\hat{x} = b_1 + \cdots + b_m$$

$$\hat{x} = \frac{b_1 + \cdots + b_m}{m}$$

= the average / mean of the b values.

(b) The error vector is

$$\vec{e} = \vec{b} - \vec{a}\hat{x} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} - \hat{x}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 - \hat{x} \\ b_2 - \hat{x} \\ \vdots \\ b_m - \hat{x} \end{pmatrix}$$

Its squared length is the variance

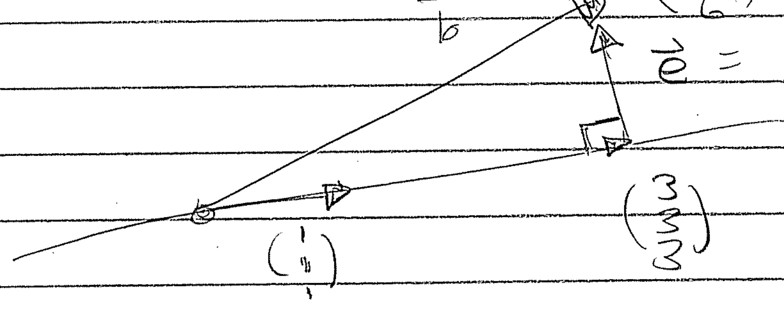
$$\|\vec{e}\|^2 = (b_1 - \hat{x})^2 + (b_2 - \hat{x})^2 + \cdots + (b_m - \hat{x})^2$$

Its length is the standard deviation

$$\|\vec{e}\| = \sqrt{(b_1 - \hat{x})^2 + \cdots + (b_m - \hat{x})^2}$$

(c) Example: $\vec{b} = (1, 2, 6)$

To find the best horizontal line (which is just the mean of the b's) we project $\vec{b}$ onto the line through $\vec{a} = (1, 1, 1)$



$$\vec{e} = \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 3 \end{pmatrix}$$

the projection is $\begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$

The error is $\perp$ to the projection

$$\begin{pmatrix} -2 & -1 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 0.$$

The matrix that peforms the projection onto $\vec{a} = (1,1,1)$ is

$$P = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} = \frac{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}}{\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Check:

$$P\vec{b} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$$

It works.

4.3.17. Find the line $C+Dt=b$ closest to the data points

$$\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ 7 \end{pmatrix}, \begin{pmatrix} 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 21 \end{pmatrix}$$

Plug in:

$$\begin{aligned} C + D(-1) &= 7 \\ C + D(1) &= 7 \\ C + D(2) &= 21 \end{aligned} \rightarrow \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 21 \end{pmatrix}.$$

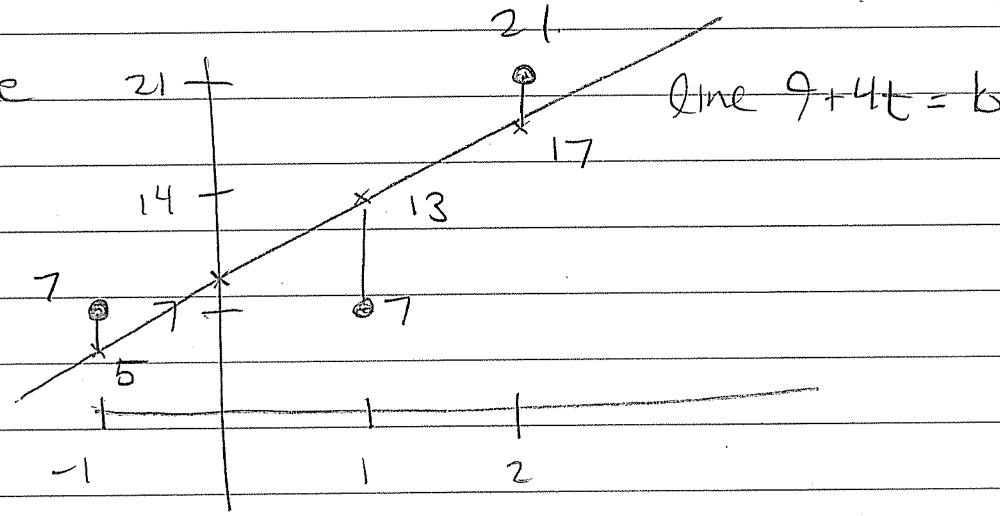
Least Squares:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 7 \\ 21 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 35 \\ 42 \end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \end{pmatrix}.$$

Picture



4.3.22. Find the line $C+Dt=b$ closest to the points $\begin{pmatrix} t \\ b \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

Normal Equation:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -2 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 5 \\ -10 \end{pmatrix}$$

$$\Longrightarrow \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$\boxed{\text{The line is } 1-t=b}$

Additional Problems.

We say that P is a "projection matrix" if $P^T = P$ and $P^2 = P$.

A.1. Show $P = A(A^TA)^{-1}A^T$ is a projection matrix.

$P^2 = P$?

$$P^2 = [A(A^TA)^{-1}A^T][A(A^TA)^{-1}A^T].$$

$$= A(A^TA)^{-1}\cancel{(A^TA)(A^TA)^{-1}}A^T$$

$$= A(A^TA)^{-1}\,I\,A^T$$

$$= A(A^TA)^{-1}A^T = P \quad \checkmark$$

$P^T = P$?

$$P^T = (A(A^TA)^{-1}A^T)^T$$

$$= (A^T)^T[(A^TA)^{-1}]^T(A)^T$$

$$= A[(A^TA)^T]^{-1}A^T$$

$$= A[(A)^T(A^T)^T]^{-1}A^T$$

$$= A(A^TA)^{-1}A^T = P \quad \checkmark$$

A.2. If A is square and invertible, then

$$(A^T A)^{-1} = (A)^{-1}(A^T)^{-1}$$
$$= A^{-1}(A^T)^{-1}$$

Hence

$$P = A(A^T A)^{-1} A^T$$
$$= \cancel{A A^{-1}} \cancel{(A^T)^{-1} A^T}$$
$$= I \cdot I = I.$$

In this case we are projecting onto the full space. How do we project onto the full space? Do nothing! i.e. the identity function

A.3. If $P^T = P$ and $P^2 = P$, then

$$(I-P)^T = I^T - P^T = I - P \quad \checkmark$$ and

$$(I-P)^2 = (I-P)(I-P)$$
$$= II - PI - IP + PP$$
$$= I - P - P + P^2$$
$$= I - P \cancel{- P + P}$$
$$= I - P \quad \checkmark$$

Hence $I-P$ is also a projection.

A.4. The projections P and $I-P$ satisfy

$$P(I-P) = PI - P^2$$
$$= P - P = 0 \quad \left(\text{the matrix of all zeros}\right).$$

We say that the projection functions P and $I-P$ are "orthogonal" to each other. What does this mean?

Example: