LECTURE 22

10/26/2016

LAST TIME:

* Projections
* Normal equation
* Projection matrices

TODAY

* Least squares
* Fitting a polynomial to given data points

$\boxed{\text{RECAP}}$

For cases when $A\vec{x} = \vec{b}$ has no solution

(when $\vec{b} \notin C(A)$, too many equations, imperfect measurements)

Solve $\boxed{\text{(normal equation)} \quad A^T A \hat{\vec{x}} = A^T \vec{b}}$

* $\hat{\vec{x}}$ minimizes length of the error $\vec{e} = \vec{b} - A\vec{x}$
* $\hat{\vec{x}}$ is called the <u>least squares</u> solution

EXAMPLE: Fit the points $(1,2)$, $(0,3)$, $(2,7)$ and $(3,-1)$ with a line as best as you can.

No straight line $y = ax + b$ passes exactly through these four points

We will set up 4 equations we wish were true

at $x=1$, $\quad 2 = a(1) + b$

$x=0$, $\quad 3 = a(0) + b$

$x=2$, $\quad 7 = a(2) + b$

$x=3$, $\quad -1 = a(3) + b$

write as a linear systems for unknowns a, b

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 3 \\ 7 \\ -1 \end{bmatrix}}_{\vec{y}}$$

(note : $b \notin C(A)$)

We want to solve the system with minimal error;

i.e., find $\vec{x}$ s.t. $A\vec{x} = \vec{y} + \vec{e}$ with smallest possible $\|e\|^2$

(least squares solution)

Here $\vec{e} = \begin{pmatrix} \text{error in } y\text{-value for pt. 1} \\ \text{error in } y\text{-value for pt. 2} \\ \text{error in } y\text{-value for pt. 3} \end{pmatrix}$

We will solve $A^TA\hat{\vec{x}} = A^T\vec{y}$ instead

$$\underbrace{\begin{bmatrix}1&1\\0&1\\2&1\\3&1\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}a\\b\end{bmatrix}}_{\vec{x}}=\underbrace{\begin{bmatrix}2\\3\\7\\-1\end{bmatrix}}_{\vec{y}}$$

$$A^TA = \begin{bmatrix}1&0&2&3\\1&1&1&1\end{bmatrix}\begin{bmatrix}1&1\\0&1\\2&1\\3&1\end{bmatrix} = \begin{bmatrix}14&6\\6&4\end{bmatrix}$$

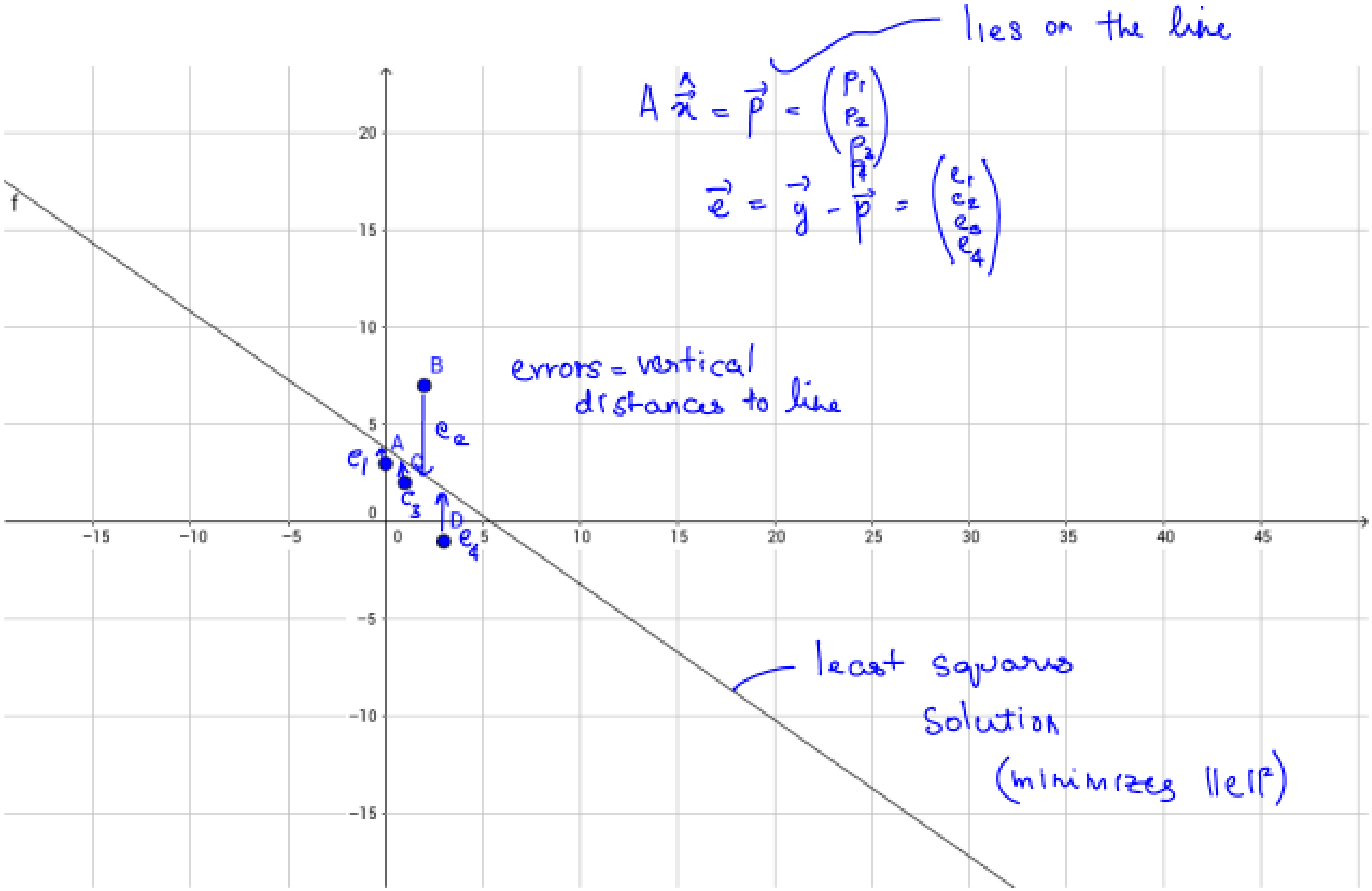
$$A^T\vec{y} = \begin{bmatrix}1&0&2&3\\1&1&1&1\end{bmatrix}\begin{bmatrix}2\\3\\7\\-1\end{bmatrix} = \begin{bmatrix}13\\11\end{bmatrix}$$

$$A^TA\hat{\vec{x}} = A^T\vec{y} \iff \begin{bmatrix}14&6\\6&4\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}13\\11\end{bmatrix}$$

(elimination) $\left[\begin{array}{cc|c}14&6&13\\6&4&11\end{array}\right] \longrightarrow \left[\begin{array}{cc|c}14&6&13\\0&10/7&38/7\end{array}\right]$ gives $b = 19/5 = 3.8$

$a = -0.7$

Hence, least-squares solution is the line $y = -0.7x + 3.8$


lies on the line
$A\hat{\vec{x}} = \vec{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix}$
$\vec{e} = \vec{y} - \vec{p} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}$
errors = vertical distances to line
least squares solution
(minimizes $\|e\|^2$)
f
A
B
C
D
e_1
e_2
e_3
e_4

Note:

For our problem, $\underbrace{\|e\|^2}_{\text{residual}} = \| A\hat{x} - \vec{y} \|^2$

$$= (a(1)+b-2)^2 + (a(0)+b-3)^2 + (a(2)+b-7)^2 + (a(3)+b-(-1))^2$$

at

$x=1, \quad 2 = a(1)+b$

$x=0, \quad 3 = a(0)+b$

$x=2, \quad 7 = a(2)+b$

$x=3, \quad -1 = a(3)+b$

setting the (partial) derivatives to 0,

Column 1 of A

$$\frac{\partial \|e\|^2}{\partial a} = 0 \iff 2(a+b-2)\boxed{(1)} + \boxed{0} + 2(2a+b-7)\boxed{(2)} + 2(3a+b+1)\boxed{(3)} = 0$$

$$(2a+2b-4) + (8a+4b-28) + (18a+6b+6) = 0$$

$$28a + 12b = 26 \quad \text{or} \quad \underline{14a+6b=13}$$

eqn ① in $A^TA\hat{x} = A^T\vec{y}$

$$\frac{\partial \|e\|^2}{\partial b} = 0 \iff 2(a+b-2) + 2(b-3) + 2(2a+b-7) + 2(3a+b+1) = 0$$

$$(2a+2b-4) + (2b-6) + (4a+2b-14) + (6a+2b+2) = 0$$

$$12a + 8b = 22 \quad \text{or} \quad \underline{6a+4b=11}$$

eqn ② in $A^TA\hat{x} = A^T\vec{y}$

EXAMPLE 2: Fit the same points with a quadratic as best as you can.
Points are (1, 2), (0, 3), (2, 7) and (3, −1).

Note: In general form, the equation of a quadratic (parabola) is
$y = ax^2 + bx + c$

Set up the equations we would like to be satisfied

$2 = a(1^2) + b(1) + c$ at $x = 1$

$3 = a(0^2) + b(0) + c$ at $x = 0$

$7 = a(2^2) + b(2) + c$ at $x = 2$

$-1 = a(3^2) + b(3) + c$ at $x = 3$

Writing as a linear system,

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\hat{\vec{x}}} = \underbrace{\begin{bmatrix} 2 \\ 3 \\ 7 \\ -1 \end{bmatrix}}_{\vec{y}}$$

$$A^TA = \begin{bmatrix} 1 & 0 & 4 & 9 \\ 1 & 0 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\hat{\vec{x}}} = \underbrace{\begin{bmatrix} 2 \\ 3 \\ 7 \\ -1 \end{bmatrix}}_{\vec{y}}$$

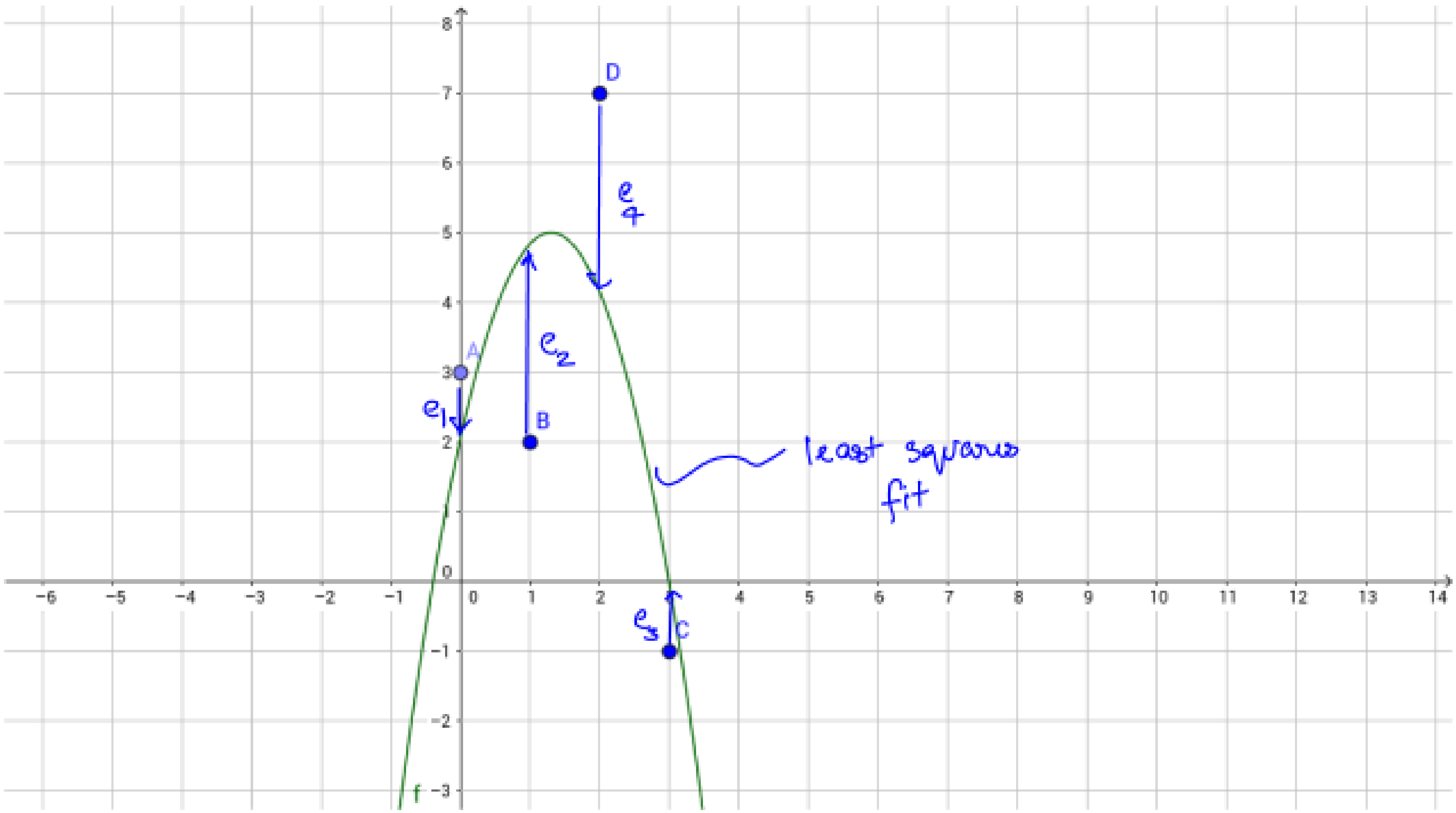
$$\vec{y} \notin C(A)$$

$$A^T\vec{y} = \begin{bmatrix} 1 & 0 & 4 & 9 \\ 1 & 0 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 7 \\ -1 \end{bmatrix} = \begin{bmatrix} 21 \\ 13 \\ 11 \end{bmatrix}$$

Solve $A^TA\hat{\vec{x}} = A^T\vec{y}$

$$\leadsto \begin{bmatrix} 98 & 36 & 14 \\ 36 & 14 & 6 \\ 14 & 6 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 21 \\ 13 \\ 11 \end{bmatrix}$$

We get $\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1.75 \\ 4.55 \\ 2.05 \end{pmatrix}$



A
B
C
D
e_1
e_2
e_3
e_4
least squares fit
f

What happens if we want to fit these points to a cubic function?

EXACT FIT!

In general,

solving $A\hat{\vec{x}} = \vec{y}$

data from x-coordinates

polynomial parameters

y-values

we get a solution $\hat{\vec{x}}$

s.t. $A\hat{\vec{x}} = \vec{y} + \vec{e}$

error in each y-coordinate

So that $\|\vec{e}\|^2$ is minimized

squared error of polynomial from each y-coordinate