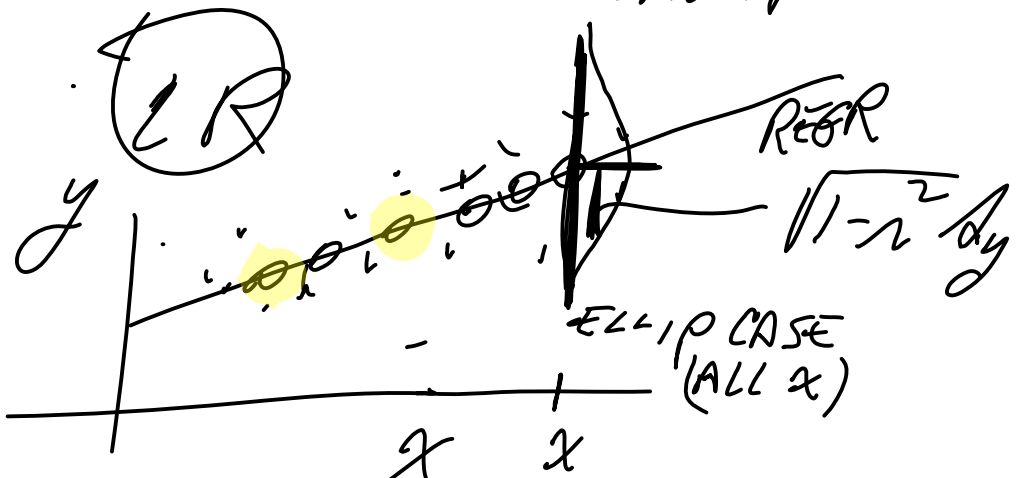
STT 200a 2-23-09

GO OVER MATERIAL FOR TOMORROW.

IDEA: TREAT MLR AS DIRECT EXTENSION OF LR.

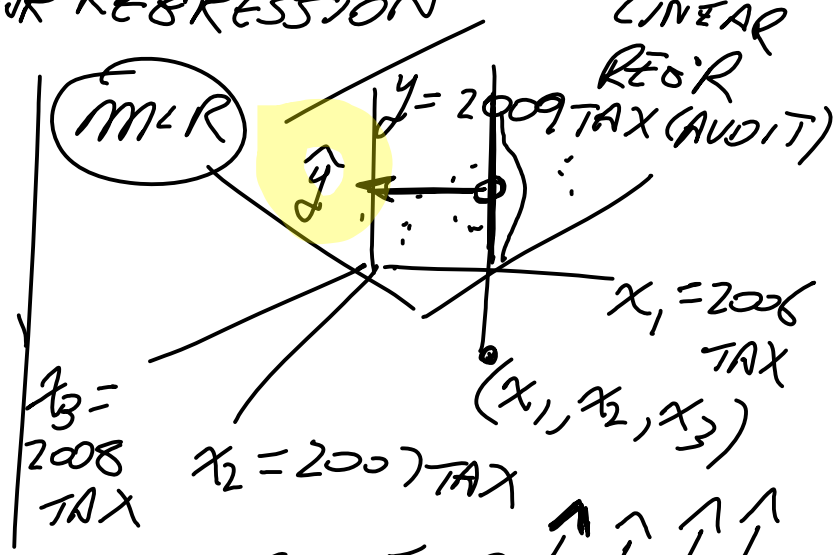
MULTIPLE LINEAR REGRESSION



∞ VERT STRIPAVGS FOR ELLIPTICAL

ALWAYS (ELLIP OR NOT)

$r^2[x,y]$ = FRAC OF σ_y^2 ACCOUNTED FOR BY REGR.



USE MLR TO FIND $\hat{b}_0\ \hat{b}_1\ \hat{b}_2\ \hat{b}_3$

MODEL $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$

PREDICT $\hat{y} = \hat{b}_0 + \hat{b}_1x_1 + \hat{b}_2x_2 + \hat{b}_3x_3$

IF ELLIPTICAL PLOT THEN y AT THIS x IS NORMAL MEAN $\hat{y}$ + SD $\sqrt{1-R^2}\,\sigma_y$.

EXAMPLE 10. SUPPOSE WE'VE GOT DATA FROM RANDOM SAMPLE OF PAVEMENT UNITS i. ≤ 100

$\hat{y} = 6.2(1) + 3(58) + \cdots + .3(36)$

$\sqrt{1-.8^2}\,16.31\,(1.96)$

ERROR OF $\hat{y}$

x_1 = TEMP WHEN POURED

x_2 = COARSENESS OF AGGREGATE

x_3 = AMT OF ADDITIVE

x_4 = TIME WAITING TO POUR.

COULD USE

$x_5 = x_1^2$

$x_6 = x_2^2$

$x_7 = x_3^2$

$x_8 = x_4^2$

$x_9 = x_1 x_2$

$x_{10} = x_1 x_3$

$x_{11} = x_1 x_4$ etc

FROM DATA (y = STRENGTH) & x's

FIND BY L.S. (MLR) FIT OF MODEL

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + \text{error}$$

$\hat{b}_0 = 6.2 \quad \hat{b}_1 = 3.0 \quad \hat{b}_2 = -1.6 \quad \hat{b}_3 = 8.2 \quad \hat{b}_4 = 0.3$

R (MULT CORRELATION) $= 0.8$ (FROM DATA); $s_y = 16.31$

PREDICT y BY $\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \hat{b}_3 x_3 + \hat{b}_4 x_4 \pm \sqrt{1-R^2}\, s_y\,(1.96)$

SO FOR $x_1 = 58°$, $x_2 = .3$, $x_3 = .06$, $x_4 = 36$ FOR ELLIPSE —

KNOW IDEA — $\hat{y} \pm 1.96\sqrt{1-R^2}\,\sigma_y$

$$\hat{y} = \sum_{j=0}^{d} \hat{\beta}_j x_{ji}$$

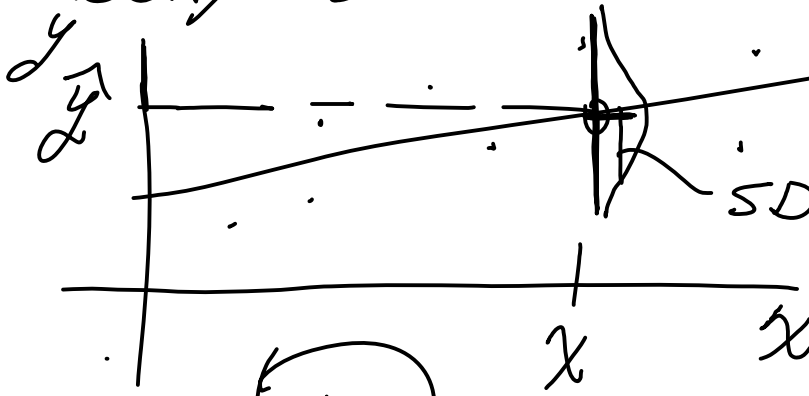
$\hat{y}$ — PRED FOR DATA ($\hat{\beta} \equiv$) & INPUTS $x_1 \cdots x_d$

i — i^{th} INDIVIDUAL (OR ANY SETTING)

STT 2006 2-23-09

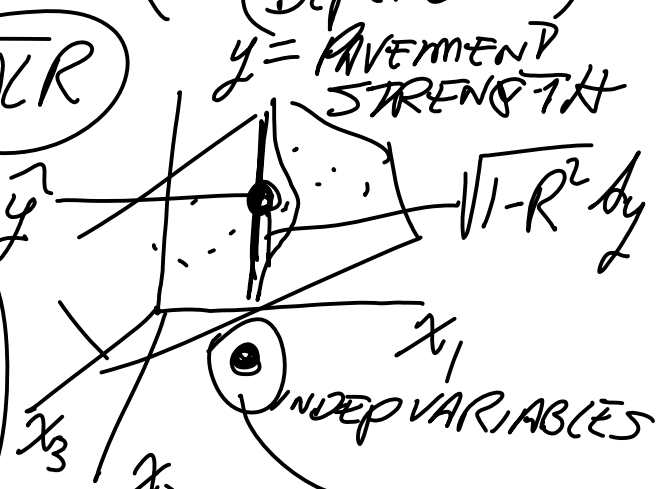
MLR = MULT LIN REGR
LR = ST LINE REGR

TODAY: SEE THAT MLR PARALLELS LR



LR

IF ELLIPSE SEE GALTON'S DATA



r^2 = FRAC OF A_y^2 EXPLAINED BY REGR ON x

$\hat{y}$ FOR GIVEN x

SD OF ALL y FOR GIVEN x IS (ELLIPSE)

$\sqrt{1-r^2}\,A_y$

$\hat{y} = \hat{b}_0 1 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \hat{b}_3 x_3$

x_1 = aggregate x_2 = ADDITIVE

x_3 = CURE TIME

$x_4 = x_1 x_2$, $x_5 = x_1 x_3$, $x_6 = x_2 x_3$

$x_7 = x_1^2$, $x_8 = x_2^2$, $x_9 = x_3^2$

MODEL $y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \text{err}$

MLR $\Rightarrow$ $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3$ SAY $\hat{b}_0 = 4, \hat{b}_1 = 3$ etc

IF (SAY) GIVEN PROPOSED MIX $x_1 = .3, x_2 = .1, x_3 = 36$ ⊙

$\Rightarrow \hat{y} = 4 + 3(.3) + (2.6)(.1) + (6.4)(36)$

EXAMPLE: WE'VE DATA ON y = TIME TO RELAPSE

x_1 = age, x_2 = severity of condition.

MODEL $y = b_0 + b_1 x_1 + b_2 x_2$

(x_1 ↑ age, x_2 ↑ SEVERITY)

DATA $\Rightarrow$ MLR

$\Rightarrow R$ = MULTI CORRELATION $= 0.8$

$\Rightarrow s_y$ (days) $= 162$

$\Rightarrow$ FIT $\hat{b}_0, \hat{b}_1, \hat{b}_2$

$30, 1.6, 0.6$

MODEL WRONG HENCE SINCE IT POSTULATES TIME TO RELAPSE INCR W. x_1

THEN PREDICT $\hat{y} = \hat{b}_0 1 + \hat{b}_1 x_1 + \hat{b}_2 x_2$

FOR SPECIFIED $x_1 = 62$, $x_2 = 3.3$

PREDICT $\hat{y} = 30(1) + 1.6(62) + .6(3.3)$

$\Rightarrow$ EMOE $\sqrt{1-R^2}\, s_y (1.96) = \sqrt{1-.8^2}\, 162 (1.96)$ $\Leftarrow$

(FOR – ELLIPTICAL)