Applied Math 166 Problem Set 8 4/11/03

11.54 Using the matrix notation of Section 11.10, write

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad Y = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \qquad \text{so that} \qquad X'X = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} \qquad (X'X)^{-1} = \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{10} \end{bmatrix}$$

Then

$$\widehat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} 1.5 \\ -.6 \end{bmatrix}$$

so that the least squares straight line is $\widehat{y} = 1.5 - .5x$. The observed points and the fitted line were shown in Figure 11.1.

11.55

$$x = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad Y = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \qquad X'Y = \begin{bmatrix} 7.5 \\ 1.5 \end{bmatrix} \qquad X'X = \begin{bmatrix} 5 & 5 \\ 5 & 15 \end{bmatrix}$$

The student may verify that

$$(X'X)^{-1} = \begin{bmatrix} .3 & -.1 \\ -.1 & .1 \end{bmatrix} \quad \text{and} \quad \widehat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} 2.1 \\ -.6 \end{bmatrix}$$

so that the least squares line is

$\widehat{y} = 2.1 - .6x$

The line is shown in Figure 11.8. Notice that for this exercise $X'X$ was no longer diagonal and hence the calculation of $(X'X)^{-1}$ was a bit more tedious.

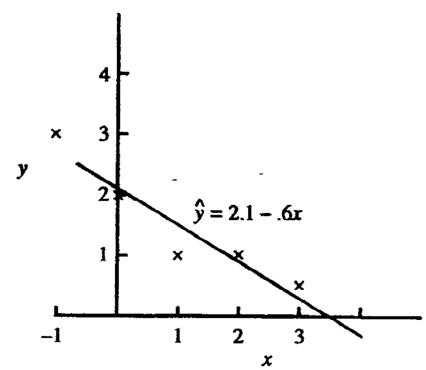

Figure 11.8

11.57 **a.** Using the model $y = \beta_0 + \beta_1 x + \epsilon$, calculate

$$(X'X) = \begin{bmatrix} 10 & 0 \\ 0 & 330 \end{bmatrix} \qquad X'Y = \begin{bmatrix} 299.9 \\ 458.3 \end{bmatrix}$$

$$\widehat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} 29.99000 \\ 1.3887879 \end{bmatrix}$$

and the least squares line is

$$\widehat{y} = 29.99 + 1.39x.$$

b. Using the model $y = \beta_0 + \beta_1 x + \beta_2 c^2 + \epsilon$, calculate

$$X'X = \begin{bmatrix} 10 & 0 & 330 \\ 0 & 330 & 0 \\ 330 & 0 & 19{,}338 \end{bmatrix} \quad X'Y = \begin{bmatrix} 299.9 \\ 458.3 \\ 8220.7 \end{bmatrix}$$

$$\widehat{\beta} = (X'X)^{-1}(X'Y) = \begin{bmatrix} 36.54 \\ 1.59 \\ -.20 \end{bmatrix}.$$

Hence the least squares line is

$$\widehat{y} = 36.54 + 1.59x - .20x^2.$$

11.59 Since the vector $\boldsymbol{a}_i$ is a vector of k 0's and one 1 (in the jth position), we can write

$$\beta_i = \boldsymbol{a}_i'\boldsymbol{\beta} \qquad \text{and} \qquad \widehat{\beta}_i = \boldsymbol{a}_i'\widehat{\boldsymbol{\beta}}$$

in vector notation. Then with $U = \boldsymbol{a}_i'\widehat{\boldsymbol{\beta}}$,

$$E(\widehat{\beta}_i) = E\left(a_i'\widehat{\boldsymbol{\beta}}\right) = \boldsymbol{a}_i' E(\widehat{\boldsymbol{\beta}}) = \boldsymbol{a}_i'\boldsymbol{\beta} = \beta_i$$

and

$$V(\widehat{\beta}_i) = V\left(\boldsymbol{a}_i'\widehat{\boldsymbol{\beta}}\right) = [\boldsymbol{a}_i'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{a}_i]\,\sigma^2.$$

But

$$\boldsymbol{a}_i'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{a}_i = [c_{0i} \quad c_{1i} \quad c_{2i} \quad \cdots \quad c_{ki}] \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 1 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = c_{ii}$$

so that $V(\widehat{\beta}_i) = c_{ii}\sigma^2$.

11.60 **a.** From Exercise 11.57, $\widehat{\beta}_2 = -.19839$, $c_{22} = .0001184712$. Then

$$\text{SSE} = \boldsymbol{Y}'\boldsymbol{Y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} = 245.68562$$

using the matrices given in Exercise 11.57. The hypothesis to be tested is

$$H_0\text{: } \beta_2 = 0 \qquad \text{vs.} \qquad H_a\text{: } \beta_2 \neq 0$$

and the test statistic is

$$t = \frac{\widehat{\beta}_2 - 0}{\sqrt{s^2(c_{22})}} = \frac{-.19839}{\sqrt{\left(\frac{\text{SSE}}{7}\right)c_{22}}} = -3.08$$

The rejection region with $\alpha = .10$, and 7 degrees of freedom is $|t| > 1.895$, and H_0 is rejected. There is evidence of a quadratic effect.

b. The 90% confidence interval for β_2 is

$$\widehat{\beta}_2 \pm t_{.05}\sqrt{s^2 c_{22}} = -.19839 \pm 1.895(.064456) = -.20 \pm .12$$

or $[-.32, -.08]$.

11.62 **a.** Defining $x_1 = \frac{T_1-60}{10}, x_2 = \frac{P-15}{5}, x_3 = \frac{C-1.5}{.5}, x_4 = \frac{T_2-150}{50}$ yields the desired result. Notice we are simply subtracting the midpoint (average) and dividing by the distance to the midpoint For example the midpoint for T_1 is $(70 + 50)/2 = 60$ and each point is $(70 - 50)/2 = 10$ from the midpoint.

b. The matrix solutions are

$$\boldsymbol{Y} = \begin{bmatrix} 22.2 \\ 19.4 \\ 22.1 \\ 14.2 \\ 24.5 \\ 24.1 \\ 19.6 \\ 12.7 \\ 24.4 \\ 25.2 \\ 23.5 \\ 19.3 \\ 25.9 \\ 28.4 \\ 16.5 \\ 16.0 \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 \end{bmatrix} \quad \boldsymbol{X'Y} = \begin{bmatrix} 338 \\ -50.2 \\ -19.4 \\ -2.6 \\ -20.4 \end{bmatrix}$$

$$\boldsymbol{X'X} = \begin{bmatrix} 16 & 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix}$$

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{16} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{16} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{16} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{16} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} 21.125 \\ -3.1375 \\ -1.2125 \\ -.1625 \\ -1.275 \end{bmatrix}$$

and the fitted model is

$$\widehat{y} = 21.125 - 3.1375x_1 - 1.2125x_2 - .1625x_3 - 1.275x_4$$

c. Calculate

$$\text{SSE} = \boldsymbol{Y'Y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X'Y} = 7446.52 - 7347.7075 = 98.8125$$

and

$$s^2 = \frac{\text{SSE}}{n-[k+1]} = \frac{98.8125}{16-5} = 8.98.$$

The test of

$$H_0\colon \beta_i = 0 \qquad \text{vs.} \qquad H_a\colon \beta_i \neq 0$$

for $i = 1, 2, 3, 4$, will be based on the test statistic

$$t_i = \frac{\widehat{\beta}_i}{s\sqrt{c_{ii}}} = \frac{4\widehat{\beta}_i}{\sqrt{8.98}} = \frac{\widehat{\beta}_i}{.7493}.$$

For $\alpha = .01$, each null hypothesis will be rejected, $i = 1, 2, 3, 4$, if $|t| > t_{.005,11} = 3.106$. The test statistics are

$$t_1 = \frac{-3.1375}{.7493} = -4.19 \qquad t_3 = \frac{-.1625}{.7493} = -.22$$

$$t_2 = \frac{-1.2125}{.7493} = -1.62 \qquad t_4 = \frac{-1.275}{.7493} = -1.70$$

Hence for $i = 1$, the null hypothesis is rejected, while for $i = 2, 3, 4$, the null hypothesis is not rejected.

For β_1, the p-value is less than .01 since $|t| > t_{.005,11} = 3.106$. For β_2 and β_4, $|t|$ is between $t_{.05,11}$ and $t_{.10,11}$; thus $.1 < p\text{-value} < .2$. For β_3, $|t|$ is less than $1.363 = t_{.10,11}$; thus the p-value $> .2$.

11.63 When $T_1 = 50$, $P = 20$, $C = 1$, and $T_2 = 200$, we have $x_1 = -1$, $x_2 = 1$, $x_3 = -1$, and $x_4 = 1$, so that

$$\boldsymbol{a}' = [1 \quad -1 \quad 1 \quad -1 \quad 1]$$

and the

$$\boldsymbol{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{a} = \tfrac{5}{16} = .3125$$

The estimate of $E(Y)$ at this particular setting is

$$\widehat{y} = 21.125 + 3.1375 - 1.2125 + .1625 - 1.275 = 21.9375$$

and the confidence interval is (based on 11 df)

$$\widehat{y} \pm t_{\alpha/2}\, s\sqrt{\boldsymbol{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{a}} = 21.94 \pm 1.796\sqrt{8.98}\ \sqrt{.3125} = 21.94 \pm 3.01$$

or $[18.93, 24.95]$.

11.64 If the year is 1977, $x = 11$ and $\widehat{y} = 36.54 + 1.39(11) - .20(121) = 27.63$ and

$$\boldsymbol{a}' = [1 \quad 11 \quad 121].$$

Then

$$\boldsymbol{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{a} = [-.243744 \quad .0333 \quad .0104167]\begin{bmatrix}1\\11\\121\end{bmatrix} = 1.3833$$

and

$$\text{SSE} = \boldsymbol{Y}'\boldsymbol{Y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} = 245.68$$

and

$$s^2 = \tfrac{245.68}{7} = 35.0978.$$

Then the 98% prediction interval is

$$27.63 \pm 2.998\sqrt{35.0978(1 + 1.3833)} = 27.63 \pm 27.42,$$

or $[.21, 55.05]$.