LINEAR REGRESSION MODELS W4315

HOMEWORK 6 ANSWERS

October 29, 2009

Instructor: Frank Wood (10:35-11:50)

1. (30 points) 6.18 (c) (e) in the textbook

Answer:

(c) $\widehat{Y} = 12.2006 - 0.1420X_1 + 0.2820X_2 + 0.6193X_3 + 0.0000079X_4$.
(e) See the end.

2. (20 points) 6.24 in the textbook

Answer:

(a) Minimize $Q = \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i1}^2 - \beta_3 X_{i2})^2$.
Normal equations:

$$\begin{aligned}
&\sum Y_i - nb_0 - b_1 \sum X_{i1} - b_2 \sum X_{i1}^2 - b_3 \sum X_{i2} = 0 \\
&\sum Y_i X_{i1} - b_0 \sum X_{i1} - b_1 \sum X_{i1}^2 - b_2 \sum X_{i1}^3 - b_3 \sum X_{i1} X_{i2} = 0 \\
&\sum Y_i X_{i1}^2 - b_0 \sum X_{i1}^2 - b_1 \sum X_{i1}^3 - b_2 \sum X_{i1}^4 - b_3 \sum X_{i1}^2 X_{i2} = 0 \\
&\sum Y_i X_{i2} - b_0 \sum X_{i2} - b_1 \sum X_{i1} X_{i2} - b_2 \sum X_{i1}^2 X_{i2} - b_3 \sum X_{i2}^2 = 0.
\end{aligned}$$

(b) $L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(Y_i - \beta_0 - \beta_1 X_{i1} - \beta_2 X_{i1}^2 - \beta_3 X_{i2})^2]$.
Under the normality assumption of the error terms, they are equivalent.

3. (30 points) 6.27 in the textbook

Answer:

(a)

$$\begin{bmatrix} 33.9321 \\ 2.7848 \\ -0.2644 \end{bmatrix}$$

(b)

$$\begin{bmatrix} -2.6996 \\ -1.2300 \\ -1.6374 \\ -1.3299 \\ -0.0900 \\ 6.9868 \end{bmatrix}$$

(c)

$$\begin{bmatrix} 0.2314 & 0.2517 & 0.2118 & 0.1489 & -0.0548 & 0.2110 \\ 0.2517 & 0.3124 & 0.0944 & 0.2663 & -0.1479 & 0.2231 \\ 0.2118 & 0.0944 & 0.7044 & -0.3192 & 0.1045 & 0.2041 \\ 0.1489 & 0.2663 & -0.3192 & 0.6143 & 0.1414 & 0.1483 \\ -0.0548 & -0.1479 & 0.1045 & 0.1414 & 0.9404 & 0.0163 \\ 0.2110 & 0.2231 & 0.2041 & 0.1483 & 0.0163 & 0.1971 \end{bmatrix}$$

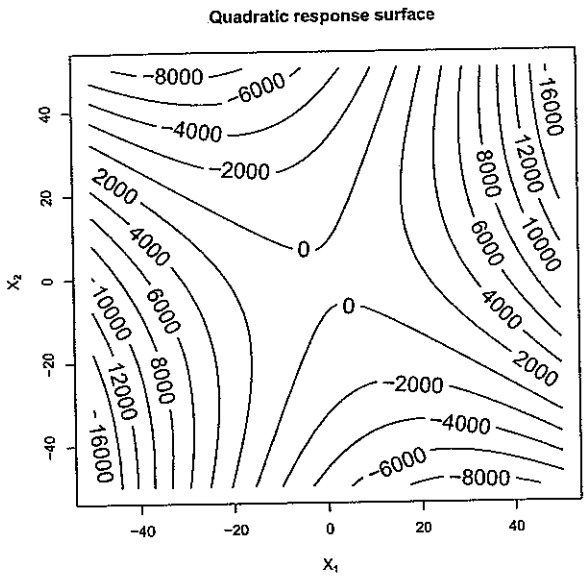
(d) $3,009.926$

(e)

$$\begin{bmatrix} 715.4711 & -34.1589 & -13.5949 \\ -34.1589 & 1.6617 & 0.6441 \\ -13.5949 & 0.6441 & 0.2625 \end{bmatrix}$$

(f) 53.8471

(g) 5.4247



4. (10 points) 8.1 in the textbook

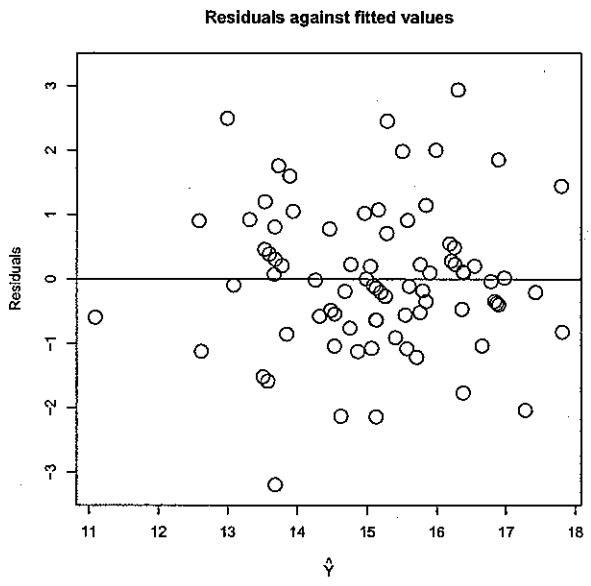
Answer:

A saddle surface.

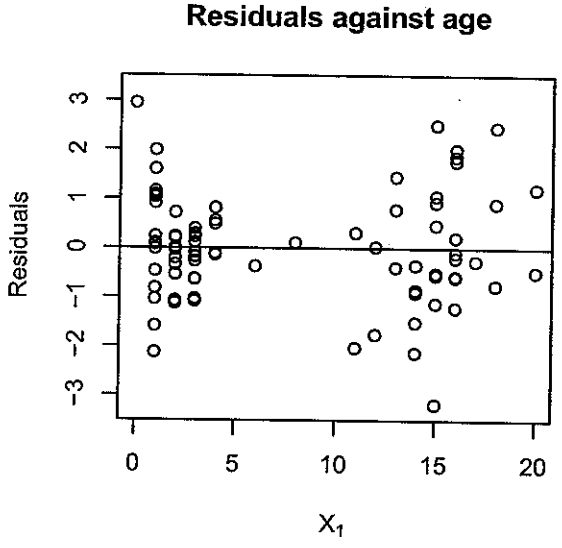
5. (10 points) 8.3 (a) in the textbook (in no more than three sentences)

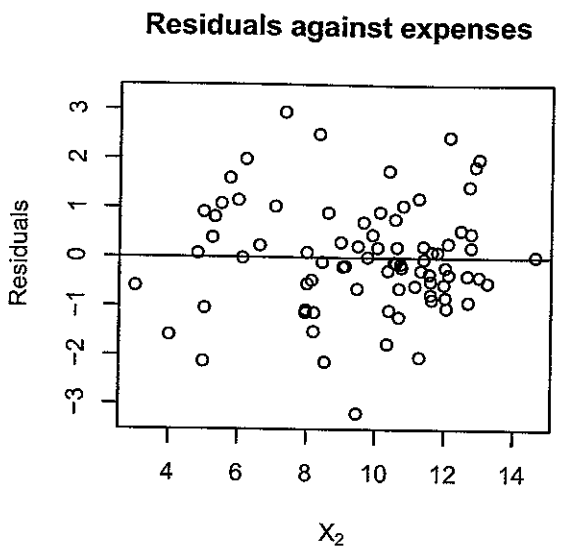
Answer:

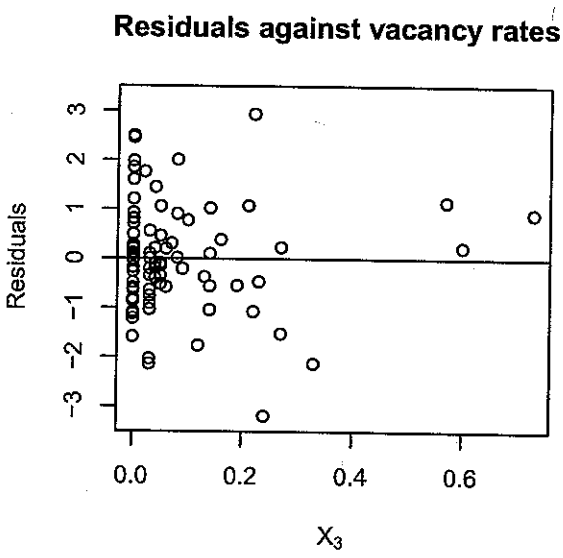
R^2 increases with the number of variables included in the model. However there is a variance-bias tradeoff.

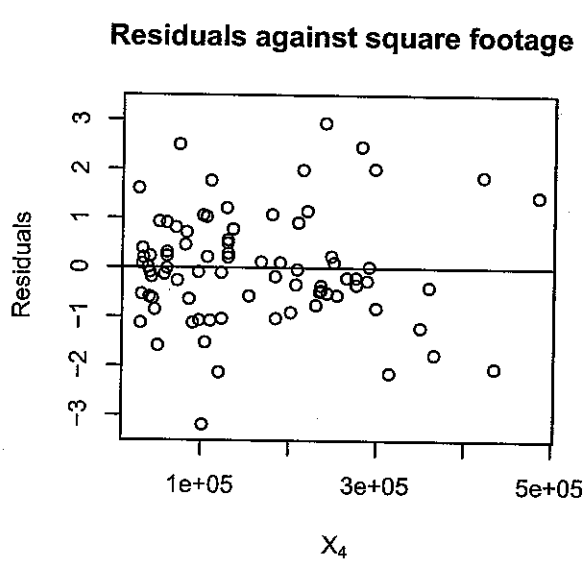


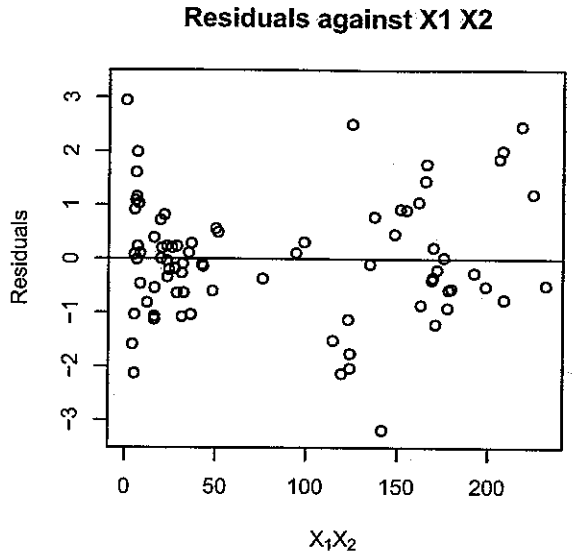
There does not seem to be a systematic relationship between the residuals and $\hat{Y}$. The variance also seems constant and does not vary with $\hat{Y}$.

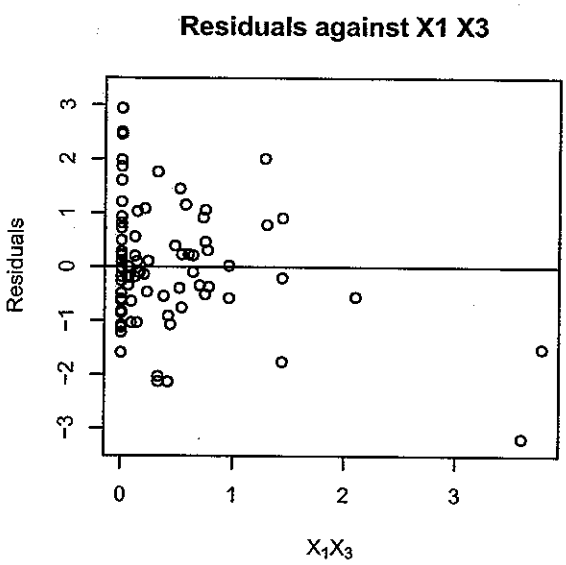

Residuals against age
Residuals
X1

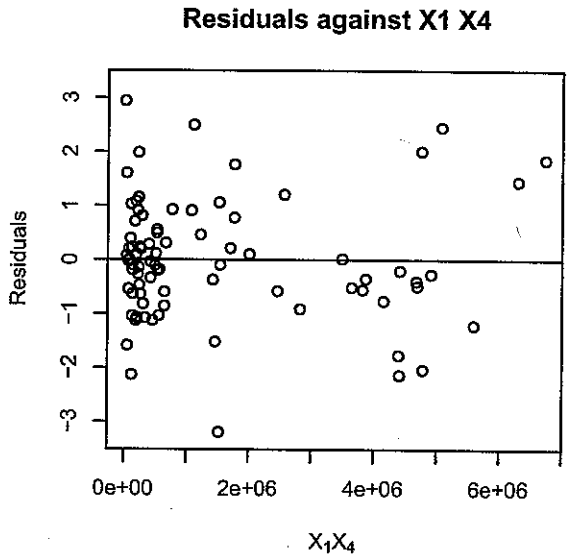

Residuals against expenses
Residuals
X2

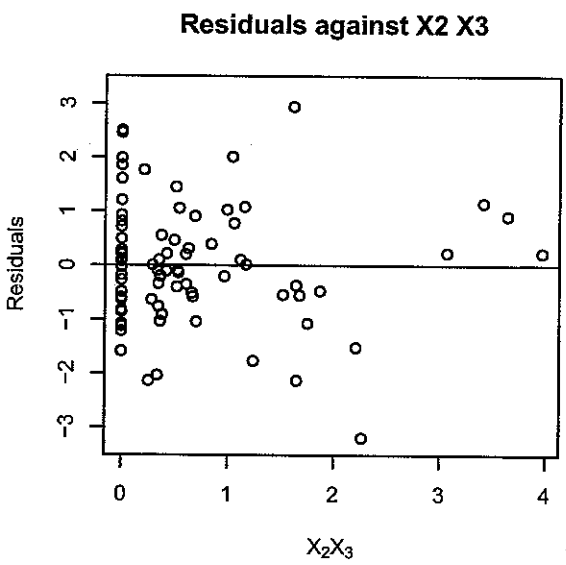

Residuals against vacancy rates
Residuals
X3

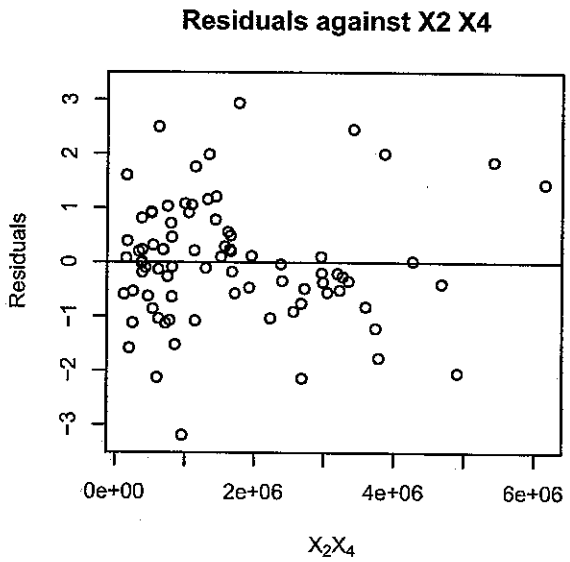

Residuals against square footage
Residuals
X4

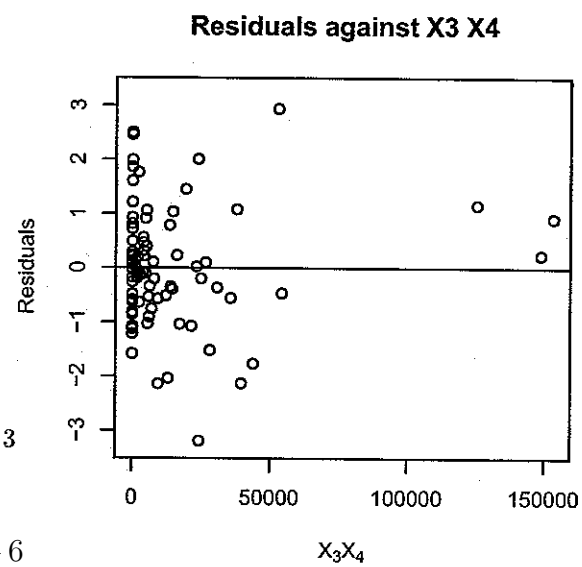

Residuals against X1 X2
Residuals
X1X2

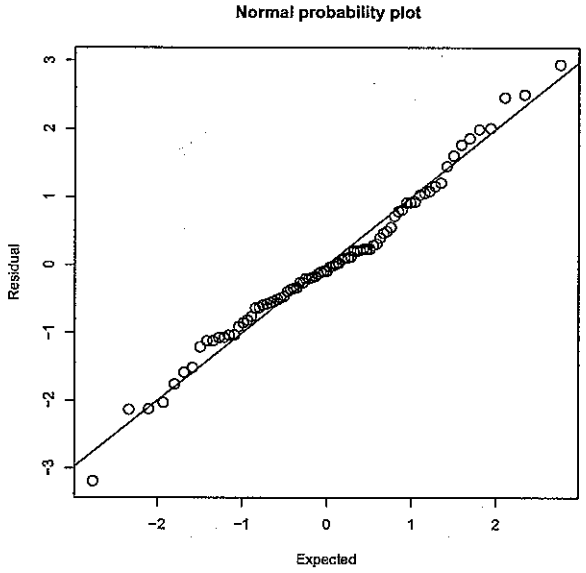

Residuals against X1 X3
Residuals
X1X3


Residuals against X1 X4
Residuals
X1X4


Residuals against X2 X3
Residuals
X2X3


Residuals against X2 X4
Residuals
X2X4


Residuals against X3 X4
Residuals
X3X4


Normal probability plot
Residual
Expected
3
2
1
0
-1
-2
-3
-2
-1
0
1
2