Math 644, Fall 2012

Solution for Homework 7

Problem 1.

(a) (145)

(b) (0)–0.9596; (4)–0.7338; (45)–(-0.1927); (145)–(-4.6600)

The model calculated: (0), (1), (2), (3), (4), (5), (14), (24), (34), (45), (145), (245), (345), (1245), (1345).

(c) (12345)–(-4.5600), (1345)–(-4.6157), (145)–(-4.6600)

The model calculated: (12345), (2345), (1345), (1245), (1235), (1234), (345), (145), (135), (134), (45), (15), (14).

(d)

$$SSE(F) = 0.2406 \; df = 34,$$

$$SSE(R) = 0.2618 \; df = 46,$$

$$F = \frac{(0.2618 - 0.2406)/2}{0.2406/34} = 1.4979 < F(0.95, 2, 34) = 3.3158$$

Thus, they can be removed.

Problem 2.

See the following R code:

```
xy = read.table('CarGasoline.txt')

y = xy[,1]
x1 = xy[,2]
x2 = xy[,3]
x3 = xy[,4]
x4 = xy[,5]
x5 = xy[,6]
x6 = xy[,7]
x7 = xy[,8]
```

```
x8 = xy[,9]
x9 = xy[,10]
x10 = xy[,11]
x11 = xy[,12]



r0 = lm(y~x1+x2+x3+x4+x5+x6+x7+x8+x9+x10+x11)
summary(r0)


# (a), no, some of the coefficients are obviously not signifciant differet
from 0



# (b)

r1 = lm(y~x1)
summary(r1)

r2 = lm(y~x10)
summary(r2)

r3 = lm(y~x1+x10)
summary(r3)

r4 = lm(y~x2+x10)
summary(r4)

r5 = lm(y~x8+x10)
summary(r5)

r6 = lm(y~x5+x8+x10)
summary(r6)


# (b) comparing the Ra^2, we prefer the last model


step(r0, direction="both")


# (c) we also select model y~x5+x8+x1


#(d)

par(mfrow = c(2, 2))
plot(x1, y)
plot(x2, y)
plot(x8, y)
plot(x10, y)

# yes, there is nonlinearity

# (e)
```

```
w = 100/y
par(mfrow = c(2, 2))
plot(x1, w)
plot(x2, w)
plot(x8, w)
plot(x10, w)

# (e) yes, after transformaion, the nonliearity disappera


# (f)

rw = lm(w~x1+x2+x3+x4+x5+x6+x7+x8+x9+x10+x11)
step(rw, direction="both")


# we select x8 nad x10


# (g)

x13 = x8/x10

rfinal = lm(y~x13)
summary(rfinal)

# which has the biggest Ra^2 among all the models
```

Problem 3.

For model $Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$, if we estimate the coefficients by minimizing

$$Q(b_1, b_2, b_3) = \sum_{i=1}^{n} \{Y_i - b_1 X_{i1} - b_2 X_{i2} - b_3 X_{i3}\}^2 + \lambda(b_1^2 + b_2^2 + b_3^2)$$

Give the estimator of (b_1, b_2, b_3) in matrix, where $\lambda > 0$ is a constant.

Taking derivatives and letting them to be 0, we have

$$\sum_{i=1}^{n} \{Y_i - b_1 X_{i1} - b_2 X_{i2} - b_3 X_{i3}\} X_{i1} + \lambda b_1 = 0$$
$$\sum_{i=1}^{n} \{Y_i - b_1 X_{i1} - b_2 X_{i2} - b_3 X_{i3}\} X_{i2} + \lambda b_2 = 0$$
$$\sum_{i=1}^{n} \{Y_i - b_1 X_{i1} - b_2 X_{i2} - b_3 X_{i3}\} X_{i3} + \lambda b_3 = 0$$

i.e.

$$\sum_{i=1}^{n} X_{i1}Y_i - b_1\sum_{i=1}^{n} X_{i1}^2 - b_2\sum_{i=1}^{n} X_{i1}X_{i2} - b_3\sum_{i=1}^{n} X_{i1}X_{i3} + 2\lambda b_1 = 0$$
$$\sum_{i=1}^{n} X_{i2}Y_i - b_1\sum_{i=1}^{n} X_{i1}X_{i2} - b_2\sum_{i=1}^{n} X_{i2}^2 - b_3\sum_{i=1}^{n} X_{i2}X_{i3} + 2\lambda b_2 = 0$$
$$\sum_{i=1}^{n} X_{i3}Y_i - b_1\sum_{i=1}^{n} X_{i1}X_{i3} - b_2\sum_{i=1}^{n} X_{i3}X_{i2} - b_3\sum_{i=1}^{n} X_{i3}^2 + 2\lambda b_3 = 0$$

or

$$b_1\sum_{i=1}^{n} X_{i1}^2 + b_2\sum_{i=1}^{n} X_{i1}X_{i2} + b_3\sum_{i=1}^{n} X_{i1}X_{i3} + 2\lambda b_1 = \sum_{i=1}^{n} X_{i1}Y_i$$
$$b_1\sum_{i=1}^{n} X_{i1}X_{i2} + b_2\sum_{i=1}^{n} X_{i2}^2 + b_3\sum_{i=1}^{n} X_{i2}X_{i3} + 2\lambda b_2 = \sum_{i=1}^{n} X_{i2}Y_i$$
$$b_1\sum_{i=1}^{n} X_{i1}X_{i3} + b_2\sum_{i=1}^{n} X_{i3}X_{i2} + b_3\sum_{i=1}^{n} X_{i3}^2 + 2\lambda b_3 = \sum_{i=1}^{n} X_{i3}Y_i$$

or

$$\begin{pmatrix} \sum_{i=1}^n X_{i1}^2 & \sum_{i=1}^n X_{i1}X_{i2} & \sum_{i=1}^n X_{i1}X_{i3} \\ \sum_{i=1}^n X_{i1}X_{i2} & \sum_{i=1}^n X_{i2}^2 & \sum_{i=1}^n X_{i2}X_{i3} \\ \sum_{i=1}^n X_{i1}X_{i3} & \sum_{i=1}^n X_{i3}X_{i2} & \sum_{i=1}^n X_{i3}^2 \end{pmatrix} b + \lambda I b = \begin{pmatrix} \sum_{i=1}^n X_{i1}Y_i \\ \sum_{i=1}^n X_{i2}Y_i \\ \sum_{i=1}^n X_{i3}Y_i \end{pmatrix}$$

we have

$$b = \begin{pmatrix} \sum_{i=1}^n X_{i1}^2 + \lambda & \sum_{i=1}^n X_{i1}X_{i2} & \sum_{i=1}^n X_{i1}X_{i3} \\ \sum_{i=1}^n X_{i1}X_{i2} & \sum_{i=1}^n X_{i2}^2 + \lambda & \sum_{i=1}^n X_{i2}X_{i3} \\ \sum_{i=1}^n X_{i1}X_{i3} & \sum_{i=1}^n X_{i3}X_{i2} & \sum_{i=1}^n X_{i3}^2 + \lambda \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n X_{i1}Y_i \\ \sum_{i=1}^n X_{i2}Y_i \\ \sum_{i=1}^n X_{i3}Y_i \end{pmatrix}$$