HW02

caojilin

9/14/2018

Problem 1

1. let $a = (a_1, \cdots, a_n)$ $\beta_1 = \mathbb{E}(a^T Y)$ we have $\mathbb{E}(Y) = \mathbb{E}(X\beta + e) = X\beta$

so $a^T X \beta = \beta_1$ this can be written $(X^T a)^T \beta = \beta_1$ let $X^T a = \lambda$

By definition of estimability, $X^T a \in C(X^T)$, so β_1 is estimable

Problem 2

a and b

2. a) $Y = X\beta + e$ where $X = \underset{n \times 2}{\begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix}}$, $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$

b) normal equations : $X^T X \beta = X^T y$

$$\Downarrow$$

$$\begin{pmatrix} n & n \\ n & n \end{pmatrix} \beta = \begin{pmatrix} y_1 + \cdots + y_n \\ y_1 + \cdots + y_n \end{pmatrix}$$

$$\Downarrow$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \beta = \begin{pmatrix} \bar{y} \\ \bar{y} \end{pmatrix} \Rightarrow \begin{pmatrix} \beta_0 + \beta_1 \\ \beta_0 + \beta_1 \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \bar{y} \end{pmatrix}$$

one equation with two unknown, so solutions are not unique

one solution might be $\beta_0 = \frac{1}{2}\bar{y}$, $\beta_1 = \frac{1}{2}\bar{y}$

c

c) from last part, we see $X^TX \doteq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ which is not invertible, so any solutions of $X^TX\beta = X^Ty$ is a least squares estimator of β

We solved $\beta_0 + \beta_1 = \bar{y}$, so least square estimator of $\beta_0 + \beta_1 = \bar{y}$

d and e

d) $\beta_1 = (0,1)\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix} = (0,1)\beta = \begin{pmatrix}0\\ 1\end{pmatrix}^T\beta$ let $\lambda = \begin{pmatrix}0\\ 1\end{pmatrix}$

and $\lambda \notin C(X^T)$ so β_1 is not estimable

e) now $X = \underbrace{\begin{pmatrix}1 & 1\\ \vdots & \vdots\\ 1 & 1\\ 1 & 2\end{pmatrix}}_{(n+1)\times 2}$ $\beta = \begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}$ $\hat{\beta}_{LS} = (X^TX)^{-1}X^TY$

$$X^TX = \begin{pmatrix}1 & \cdots & 1 & 1\\ 1 & \cdots & 1 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ \vdots & \vdots\\ 1 & 1\\ 1 & 2\end{pmatrix} = \begin{pmatrix}n+1 & n+2\\ n+2 & n+4\end{pmatrix}$$

$$\det(X^TX) = (n+1)(n+4) - (n+2)^2 = n^2+5n+4 - n^2-4n-4 = n$$

and $n>0$ thus X^TX is invertible

$$(X^TX)^{-1} = \frac{1}{n}\begin{pmatrix}n+4 & -(n+2)\\ -(n+2) & n+1\end{pmatrix}$$

$$X^TY = \begin{pmatrix}1 & \cdots & 1 & 1\\ 1 & \cdots & 1 & 2\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_{n+1}\end{pmatrix} = \begin{pmatrix}\sum_{i=1}^{n+1} y_i\\ y_{n+1} + \sum_{i=1}^{n+1} y_i\end{pmatrix}$$

$$\hat{\beta}_{LS} = \begin{pmatrix}\dfrac{(n+4)\sum_{i=1}^{n+1} y_i - (n+2)(y_{n+1} + \sum_{i=1}^{n+1} y_i)}{n}\\ \dfrac{-(n+2)\sum_{i=1}^{n+1} y_i + (n+1)(y_{n+1} + \sum_{i=1}^{n+1} y_i)}{n}\end{pmatrix}$$

Problem 3

3. a) first we consider model $Y = X\beta + e$, So

X can be written $\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix}$ $X^T = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & \vdots \\ 0 & 1 & 0 & \cdots & \vdots \\ \vdots & \vdots & 1 & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$

a.

$n \times (n+1)$

$\beta_0 + \beta_2 = (1,0,1,\cdots,0)\begin{pmatrix}\beta_0 \\ \vdots \\ \beta_n\end{pmatrix} = \begin{pmatrix}1 \\ 0 \\ 1 \\ \vdots \\ 0\end{pmatrix}^T \beta$, $\lambda = \begin{pmatrix}1 \\ 0 \\ 1 \\ \vdots \\ 0\end{pmatrix}$

$\lambda \in C(X^T)$ thus $\beta_0 + \beta_2$ is estimable

b) $\beta_1 = (0,1,\cdots,0)\beta = \begin{pmatrix}0 \\ 1 \\ \vdots \\ 0\end{pmatrix}^T \beta$, let $\lambda = \begin{pmatrix}0 \\ 1 \\ \vdots \\ 0\end{pmatrix}$

b.

$\lambda \notin C(X^T)$, thus β_1 is not estimable, so no least squares estimate

c) $\beta_1 - \beta_2 = (0,1,-1,0,\cdots,0)\begin{pmatrix}\beta_0 \\ \vdots \\ \beta_n\end{pmatrix} = \begin{pmatrix}0 \\ 1 \\ -1 \\ \vdots \\ 0\end{pmatrix}^T \beta$, let $\lambda = \begin{pmatrix}0 \\ 1 \\ -1 \\ \vdots \\ 0\end{pmatrix}$

c. let x_1, x_2 be columns of X^T

$\begin{pmatrix}1 \\ 1 \\ 0 \\ \vdots \\ 0\end{pmatrix} + (-1)\begin{pmatrix}1 \\ 0 \\ 1 \\ \vdots \\ 0\end{pmatrix} = \begin{pmatrix}0 \\ 1 \\ -1 \\ \vdots \\ 0\end{pmatrix} = \lambda$ so $\lambda \in C(X^T)$, thus $\beta_1 - \beta_2$ is estimable d.

d) $\beta_1+\beta_2+\beta_3-3\beta_4 = (0,1,1,1,-3,0,\cdots 0)\begin{pmatrix}\beta_0\\ \beta_1\\ \vdots\\ \beta_n\end{pmatrix} = \begin{pmatrix}0\\ 1\\ 1\\ 1\\ -3\\ \vdots\\ 0\end{pmatrix}^T\beta$

let $\begin{pmatrix}0\\ 1\\ 1\\ 1\\ -3\\ \vdots\\ 0\end{pmatrix}$ be λ

$$\begin{pmatrix}1\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}+\begin{pmatrix}1\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix}+\begin{pmatrix}1\\ 0\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix}+(-3)\begin{pmatrix}1\\ 0\\ 0\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix}=\begin{pmatrix}0\\ 1\\ 1\\ 1\\ -3\\ \vdots\\ 0\end{pmatrix}$$

thus $\lambda \in C(X^T)$, thus $\beta_1+\beta_2+\beta_3-3\beta_4$ is estimable lease squares estimates

3. least squares estimate part

$$Q(\beta) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_i)^2$$

$$\frac{\partial Q}{\partial \beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_i) = 0 \Rightarrow$$

$$\sum_{i=1}^{n} y_i = n\beta_0 + \sum_{i=1}^{n}\beta_i$$

$$\frac{\partial Q}{\partial \beta_1} = \frac{\partial\left((y_1-\beta_0-\beta_1)^2 + (y_2-\beta_0-\beta_2)^2 + \cdots + (y_n-\beta_0-\beta_n)^2\right)}{\partial \beta_1}$$

$$= -2(y_1 - \beta_0 - \beta_1)$$

$$\frac{\partial Q}{\partial \beta_n} = -2(y_n - \beta_0 - \beta_n)$$

thus normal equations are

$$n\beta_0 + \sum_{i=1}^{n}\beta_i = \sum_{i=1}^{n} y_i$$

$$\beta_0 + \beta_1 = y_1$$

$$\vdots$$

$$\beta_0 + \beta_n = y_n$$

We notice if $\beta_0 = 0$, then $Q(\beta) = 0$ the smallese possible value, so

$$\hat{\beta} = \begin{pmatrix} 0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$ is lease squares estimate

	LSE
$\beta_0 + \beta_2$	y_2
β_1	not estimable
$\beta_1 - \beta_2$	$y_1 - y_2$
$\beta_1 + \beta_2 + \beta_3 - 3\beta_4$	$y_1 + y_2 + y_3 - 3y_4$

Problem 4

a)

First estimate

```
body = read.csv("Bodyfat.csv")

xmat = matrix(0, nrow(body), 5)
xmat[,1] = rep(1, nrow(body))
xmat[,2] = body$Age
xmat[,3] = body$Weight
xmat[,4] = body$Height
xmat[,5] = body$Age + 10*body$Weight + 3*body$Height


lmod1 = lm(bodyfat ~ Age + Weight + Height + I(Age + 10*Weight + 3*Height), data = body)
summary(lmod1)$coefficients
```

```
##                Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 17.7673848 7.47935066  2.375525 1.828479e-02
## Age          0.1697902 0.02956033  5.743853 2.698997e-08
## Weight       0.1981519 0.01312664 15.095402 5.829476e-37
## Height      -0.5943339 0.10690038 -5.559698 6.972685e-08
```

using `lm` function to get least squares esitmates BODYFAT $= 17.7673848 + 0.1697902Age + 0.1981519Weight - 0.5943339Height$

Second estimate

svd

```
pseudo_inverse = pinv(xmat)
# these are least square estimates from beta0 to beta4
esti = pseudo_inverse %*% body$bodyfat;esti
```

```
##              [,1]
## [1,] 17.767384801
## [2,]  0.166472102
## [3,]  0.164971028
## [4,] -0.604288101
## [5,]  0.003318082
```

Third estimates

```
#we notice that the matrix X's rank is 4, because the last column is the linear combinat
ion of the previous columns.
Rank(xmat)
```

```
## [1] 4
```

```
#find the vector that spans the nullspace of X
x5 = nullspace(xmat)
#Then this vector is orthogonal to row space of X
#A new least square estimate could be
esti + x5
```

```
##              [,1]
## [1,] 17.76738480
## [2,]  0.07155630
## [3,] -0.78418697
## [4,] -0.88903550
## [5,]  0.09823388
```

b)

$$\text{x5 is the vector that spans nullspace of X and is orthogonal to row space of X}$$
$$\beta_1 = \lambda^T\beta = (0,1,0,0,0)\beta$$
$$so\ \lambda = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} and\ \ \lambda \cdot x5 = -0.0949158,\ \text{which is not equal to 0}$$
$$\text{This means that } \lambda \text{ is not in the row space of X, so } \beta_1 \text{ is not estimable}$$

c)

$$\text{we assume the model to be}$$
$$BodyFat = \beta_0 + \beta_1 Age + \beta_2 Weight + \beta_3 Height + \beta_4(Age + 10 * Weight + 3 * Height)$$
$$\text{but we can rewrite it as}$$
$$BodyFat = \beta_0 + (\beta_1 + \beta_4)Age + (\beta_2 + 10\beta_4)Weight + (\beta_3 + 3\beta_4)Height$$
$$\text{From part a, we know these results}$$
$$\text{So least sqaures estimates}$$
$$\beta_0 = 17.7673848$$
$$\beta_1 + \beta_4 = 0.1697902$$
$$\beta_2 + 10\beta_4 = 0.1981519$$
$$\beta_3 + 3\beta_4 = -0.5943339$$

d)

Yes, we can just read off estimates from part c

```
summary(lm(bodyfat ~ Age + Weight + Height + I(Age + 10*Weight +
3*Height), data = body))
```

```
##
## Call:
## lm(formula = bodyfat ~ Age + Weight + Height + I(Age + 10 * Weight +
##     3 * Height), data = body)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.3960  -4.5038  -0.0326   3.8324  15.7154
##
## Coefficients: (1 not defined because of singularities)
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        17.76738    7.47935   2.376   0.0183 *
## Age                                 0.16979    0.02956   5.744 2.70e-08 ***
## Weight                              0.19815    0.01313  15.095  < 2e-16 ***
## Height                             -0.59433    0.10690  -5.560 6.97e-08 ***
## I(Age + 10 * Weight + 3 * Height)        NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.809 on 248 degrees of freedom
## Multiple R-squared:  0.524,  Adjusted R-squared:  0.5182
## F-statistic: 90.99 on 3 and 248 DF,  p-value: < 2.2e-16
```

We found that if a column is a linear combination of the other columns, lm function ignores that column.

Problem 5

5. a) $X^TX = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$ $\quad \hat{\beta} = (X^TX)^{-1}X^TY = \frac{1}{8} I \cdot X^TY$

$$\hat{\beta} = \frac{1}{8}\underset{4\times 8}{\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \end{pmatrix}} \underset{8\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_8 \end{pmatrix}} = \frac{1}{8}\begin{pmatrix} \sum_{i=1}^{8} y_i \\ (y_1+y_3+y_5+y_7) - (y_2+y_4+y_6+y_8) \\ (y_1+y_2+y_5+y_6) - (y_3+y_4+y_7+y_8) \\ (y_1+y_4+y_5+y_8) - (y_2+y_3+y_6+y_7) \end{pmatrix}$$

(b)

$$\text{Cov}(\beta) = \text{Cov}((X^TX)^{-1}X^TY) = (X^TX)^{-1}X^T\,\text{Cov}(Y)\,X(X^TX)^{-1} = (X^TX)^{-1}X^T(\sigma^2 I_n)X(X^TX)^{-1}$$
$$= \sigma^2 (X^TX)^{-1}$$

In this case $\text{Cov}(\hat{\beta}) = \sigma^2 \frac{1}{8} I_4$. This means every β_i has individual precision $\frac{\sigma^2}{8}$

We only used 8 observations and have the same precision that has 32 observations.