MAST30025: Linear Statistical Models

Akira Wang, Student ID: 913391
Assignment 1, 2019

March 29, 2019

All R output is provided at the end of the assignment. Written using LaTeXon Overleaf.

1. Suppose that A is a symmetric matrix with $A^k = A^{k+1}$ for some integer $k \geq 1$. Show that A is idempotent.

   Given A is a symmetric matrix, then an orthogonal matrix P exists such that

$$\Rightarrow P^T AP = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix} \text{where } \lambda_i, i = 1, 2, \dots, k, \text{are the eigenvalues of } A.$$

   Now, let $A = P^T \Lambda P$. So

$$\begin{aligned} \Rightarrow A^2 &= P^T \Lambda P P^T \Lambda P \\ &= P^T \Lambda^2 P, \text{ since P is orthogonal} \end{aligned}$$

   Each diagonal entry of Λ are either 0 or 1, so $\Lambda^2 = \Lambda$ and $A^2 = A$.

$$\begin{aligned} \Rightarrow A\mathbf{x} &= \lambda\mathbf{x} \\ \therefore A^{k+1}\mathbf{x} &= A^k\mathbf{x} \\ \lambda^{k+1} &= \lambda^k \\ \lambda(\lambda^k) &= \lambda^k \\ \lambda^k - \lambda(\lambda^k) &= \mathbf{0} \\ \lambda^k(1-\lambda) &= \mathbf{0} \end{aligned}$$

   We find that $\lambda_1 = 0, \lambda_2 = 1$ which satisfies the condition that the eigenvalues of an idempotent matrix are always either 0 or 1.

   Therefore we can say that the symmetric matrix A is idempotent.

2. Let $A_1, A_2, \ldots, A_m$ be a set of symmetric $k \times k$ matrices. Suppose that there exists an orthogonal matrix P such that $P^T A_i P$ is diagonal for all i. Show that $A_i A_j = A_j A_i$ for every pair $i, j = 1, 2, \ldots, m$.

Let A_n be the nth matrix in the set $A_1, A_2, \ldots, A_m$. Then

$$\Rightarrow P^T A_n P = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_k \end{bmatrix} \text{where } \lambda_i, i = 1, 2, \ldots, k, \text{are the eigenvalues of } A_n.$$

Since P is the diagonalizing matrix, we know that the columns of P correspond to the orthogonal eigenvectors of A_n.

Furthermore, we know that $P^T P = I$, and $P^{-1} = P^T$ since P is orthogonal. So for every pair $i, j = 1, 2, \ldots, m$ we have

$$\begin{aligned} \Rightarrow P^T A_i P P^T A_j P &= \begin{bmatrix} \lambda_{i,1} & 0 & \ldots & 0 \\ 0 & \lambda_{i,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{i,k} \end{bmatrix} \begin{bmatrix} \lambda_{j,1} & 0 & \ldots & 0 \\ 0 & \lambda_{j,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{j,k} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_{i,1}\lambda_{j,1} & 0 & \ldots & 0 \\ 0 & \lambda_{i,2}\lambda_{j,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{i,k}\lambda_{j,k} \end{bmatrix} \\ &= P^T A_i A_j P \\ &= A_i A_j \end{aligned}$$

$$\begin{aligned} \Rightarrow P^T A_j P P^T A_i P &= \begin{bmatrix} \lambda_{j,1} & 0 & \ldots & 0 \\ 0 & \lambda_{j,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{j,k} \end{bmatrix} \begin{bmatrix} \lambda_{i,1} & 0 & \ldots & 0 \\ 0 & \lambda_{i,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{i,k} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_{i,1}\lambda_{j,1} & 0 & \ldots & 0 \\ 0 & \lambda_{i,2}\lambda_{j,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{i,k}\lambda_{j,k} \end{bmatrix} \\ &= P^T A_j A_i P \\ &= A_j A_i \end{aligned}$$

Therefore $A_i A_j = A_j A_i$ for every pair $i, j = 1, 2, \ldots, m$.

3. Show directly that for any random vector $\mathbf{y}$ and compatible matrix A, we have $\text{var}(A\mathbf{y}) = A\text{var}(\mathbf{y})A^T$.

   Let $E[\mathbf{y}] = \boldsymbol{\mu}$,

$$\begin{aligned}
\Rightarrow \text{var}(A\mathbf{y}) &= E[(A\mathbf{y} - E[A\mathbf{y}])(A\mathbf{y} - E[A\mathbf{y}])^T] \\
&= E[A(\mathbf{y} - E[\mathbf{y}])(\mathbf{y} - E[\mathbf{y}])^T A^T] \\
&= AE[(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T]A^T \\
&= A\text{var}(\mathbf{y})A^T, \text{ as required.}
\end{aligned}$$

4. Let $\mathbf{y}$ be a 3-dimensional multivariate normal random vector with mean and variance

$$\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}, V = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

   Let

$$A = \frac{1}{10}\begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

   (a) Describe the distribution of $A\mathbf{y}$.
   Any combination of multivariate normals is in turn a multivariate normal. Then $A\mathbf{y} \sim MVN(A\boldsymbol{\mu}, AVA^T)$, where $V = \text{var}(\mathbf{y})$

$$\Rightarrow A\mathbf{y} \sim MVN\left(\begin{bmatrix} 1.2 \\ -0.6 \\ -2.0 \end{bmatrix}, \begin{bmatrix} 0.4 & -0.2 & 0 \\ -0.2 & 0.1 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}\right), \text{ using R.}$$

   (b) Find $E[\mathbf{y}^T A\mathbf{y}]$.

$$\begin{aligned}
\Rightarrow E[\mathbf{y}^T A\mathbf{y}] &= tr(AV) + \boldsymbol{\mu}^T A\boldsymbol{\mu} \\
&= tr\left(\frac{1}{10}\begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) + \frac{1}{10}\begin{bmatrix} 3 & 0 & -2 \end{bmatrix}\begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} \\
&= \frac{1}{10}tr\left(\begin{bmatrix} 8 & -4 & 0 \\ -4 & 2 & 0 \\ 0 & 0 & 10 \end{bmatrix}\right) + \frac{1}{10}\begin{bmatrix} 3 & 0 & -2 \end{bmatrix}\begin{bmatrix} 12 \\ -6 \\ -20 \end{bmatrix} \\
&= \frac{1}{10}(8 + 2 + 10) + \frac{1}{10}(36 + 0 + 40) \\
&= 9.6
\end{aligned}$$

(c) Describe the distribution of $\mathbf{y}^T A\mathbf{y}$.

$\mathbf{y}^T A\mathbf{y}$ will have a non-central χ^2 distribution with k degrees of freedom iff AV is idempotent and has rank k.

Using R, we find that AV is indeed idempotent with rank k.

$$\begin{aligned}\Rightarrow k &= r(AV) \\ &= r\left(\frac{1}{10}\begin{bmatrix}4 & -2 & 0\\ -2 & 1 & 0\\ 0 & 0 & 10\end{bmatrix}\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}\right) \\ &= 2\end{aligned}$$

$$\begin{aligned}\Rightarrow \lambda &= \frac{1}{2}\boldsymbol{\mu}^T A\boldsymbol{\mu} \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{10}\right)(76),\ \text{from part (b)} \\ &= 3.8\end{aligned}$$

Therefore

$$\Rightarrow \mathbf{y}^T A\mathbf{y} \sim \chi^2_{2,3.8}.$$

(d) Find all linear combinations of $\mathbf{y}$ elements which are independent of $\mathbf{y}^T A\mathbf{y}$.

A random vector $\mathbf{a}$ is linearly independent if the only solution to $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_k\mathbf{x}_k = \mathbf{0}$, is when all α are zero. Let $B = \mathbf{a}^T$, a 1×3 random vector, then $B\mathbf{y}$ will be independent of $\mathbf{y}^T A\mathbf{y}$ iff $BVA = \mathbf{0}$. We have

$$\begin{aligned}\Rightarrow BVA &= (BVA)^T \\ &= (VA)^T B^T \\ &= AV\mathbf{a} \\ &= \frac{1}{10}\mathbf{a}\begin{bmatrix}8 & -4 & 0\\ -4 & 2 & 0\\ 0 & 0 & 10\end{bmatrix}.\end{aligned}$$

Then we reduce AV,

$$\Rightarrow AV\begin{bmatrix}8 & -4 & 0\\ -4 & 2 & 0\\ 0 & 0 & 10\end{bmatrix} R_1 \to r_1 - 2r_2 \begin{bmatrix}0 & 0 & 0\\ -4 & 2 & 0\\ 0 & 0 & 10\end{bmatrix} R_2 \to \frac{1}{2}r_2 \begin{bmatrix}0 & 0 & 0\\ -2 & 1 & 0\\ 0 & 0 & 10\end{bmatrix}. \tag{1}$$

We then solve for $AV\mathbf{a} = \mathbf{0}$.

$$AV\mathbf{a} = \mathbf{a}\left[\begin{array}{ccc|c}0 & 0 & 0 & 0\\ -2 & 1 & 0 & 0\\ 0 & 0 & 10 & 0\end{array}\right] \tag{2}$$

$$\Rightarrow 2\alpha_1 - \alpha_2 = 0 \qquad \Rightarrow \alpha_3 = 0s \tag{3}$$

Therefore

$$\mathbf{a} = \alpha\begin{bmatrix}1\\ 2\\ 0\end{bmatrix} \tag{4}$$

Finally, all linear combinations of $\mathbf{y}$ elements are

$$\Rightarrow \alpha y_1 + 2\alpha y_2. \tag{5}$$

5. The table below shows prices in US cents per pound received by fishermen and vessel owners for various species of fish and shellfish in 1970 and 1980. (Taken from Moore and McCabe, Introduction to the Practice of Statistics, 1989.)

   (*Refer to assignment spec for table*)

   We will model the 1980 price of fish, based on the 1970 price.

   (a) The linear model is of the form $\mathbf{y} = X\beta + \epsilon$. Write down the matrices and vectors involved in this equation.

$$\mathbf{y} = \begin{bmatrix} 27.3 \\ 42.4 \\ 38.7 \\ 4.5 \\ 23.0 \\ 166.3 \\ 109.7 \\ 80.1 \\ 150.7 \\ 20.3 \\ 189.7 \\ 131.3 \\ 404.2 \\ 149.0 \end{bmatrix}, X = \begin{bmatrix} 1 & 13.1 \\ 1 & 15.3 \\ 1 & 25.8 \\ 1 & 1.8 \\ 1 & 4.9 \\ 1 & 55.4 \\ 1 & 39.3 \\ 1 & 26.7 \\ 1 & 47.5 \\ 1 & 6.6 \\ 1 & 94.7 \\ 1 & 61.1 \\ 1 & 135.6 \\ 1 & 47.6 \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{14} \end{bmatrix}. \tag{6}$$

   (b) Find the least squares estimates of the parameters.
   Let $\mathbf{b}$ be the least squares estimator for β,

$$\begin{aligned} \Rightarrow \mathbf{b} &= (X^T X)^{-1} X^T \mathbf{y} \\ &= \begin{bmatrix} -1.233836 \\ 2.701553 \end{bmatrix}, \text{ using R.} \end{aligned}$$

   (c) Calculate the sample variance s^2.
   Let $SS_{Res} = (\mathbf{y} - X\mathbf{b})(\mathbf{y} - X\mathbf{b})^T$,

$$\begin{aligned} \Rightarrow s^2 &= \frac{SS_{Res}}{n - (k+1)} \\ &= \frac{9325.833}{14 - (1+1)} \\ &= 777.1528, \text{ using R.} \end{aligned}$$

(d) A fisherman sold ocean trout for 18 cents per pound in 1970. Predict the price for ocean trout in 1980.

Using our least squares estimator $\mathbf{b}$, let $\mathbf{t}^T = \begin{bmatrix}1 & 18\end{bmatrix}$. Then the unbiased cost of Ocean Trout in 1980 is $\mathbf{t}^T\mathbf{b}$,

$$\begin{aligned}\Rightarrow \text{Ocean Trout in 1980 } &= \mathbf{t}^T\mathbf{b} \\ &= 47.39412 \text{ cents, using R.}\end{aligned}$$

(e) Calculate the standardised residual for sea scallops.

Let $\mathbf{e} = \mathbf{y} - X\mathbf{b}$, then $\text{var}(\mathbf{e}) = \sigma^2(I - H)$ is the variance of the residuals, where we define $H = X(X^TX)^{-1}X^T$.

Sea Scallops is at $i = 13$.

$$\begin{aligned}\Rightarrow z_{13} &= \frac{e_{13}}{\sqrt{s^2(1 - H_{13,13})}}, \\ &= \frac{39.10322}{\sqrt{777.1528(1 - 0.5559672)}}, \text{ from part (c)} \\ &= 2.104999, \text{ using R.}\end{aligned}$$

(f) Calculate the Cook's distance for sea scallops.

Using the results and R output from part (e)

$$\begin{aligned}D_{13} &= \frac{z_{13}^2}{k+1}\left(\frac{H_{13,13}}{1 - H_{13,13}}\right) \\ &= \frac{2.105}{1+1}\left(\frac{0.556}{1 - 0.556}\right) \\ &= 2.774008, \text{ using R.}\end{aligned}$$

(g) Does sea scallops fit the linear model? Justify your argument.

Cook's distance grows larger if both the standardised residual and the leverage is large, and we generally consider a value greater than 1 to be large. Since we have a value of $2.774008 > 1$, we can conclude that sea scallops does not fit the linear model well.

R code for Assignment 1. Akira Wang.

Question 4.a)

```
(mu = matrix(c(3,0,-2),3,1))
```

```
##      [,1]
## [1,]    3
## [2,]    0
## [3,]   -2
```

```
(V = matrix(c(2,0,0,0,2,0,0,0,1),3,3))
```

```
##      [,1] [,2] [,3]
## [1,]    2    0    0
## [2,]    0    2    0
## [3,]    0    0    1
```

```
(A = 1/10 * matrix(c(4,-2,0,-2,1,0,0,0,10),3,3))
```

```
##      [,1] [,2] [,3]
## [1,]  0.4 -0.2    0
## [2,] -0.2  0.1    0
## [3,]  0.0  0.0    1
```

```
A%*%mu
```

```
##      [,1]
## [1,]  1.2
## [2,] -0.6
## [3,] -2.0
```

```
A%*%V%*%t(A)
```

```
##      [,1] [,2] [,3]
## [1,]  0.4 -0.2    0
## [2,] -0.2  0.1    0
## [3,]  0.0  0.0    1
```

Question 4.c)

```
M = A%*%V
M%*%M
```

```
##      [,1] [,2] [,3]
## [1,]  0.8 -0.4    0
## [2,] -0.4  0.2    0
## [3,]  0.0  0.0    1
```

```
M
```

```
##      [,1] [,2] [,3]
## [1,]  0.8 -0.4    0
## [2,] -0.4  0.2    0
## [3,]  0.0  0.0    1
```

Question 5.b)

```r
(X = matrix(c(rep(1,14),13.1,15.3,25.8,1.8,4.9,55.4,39.3,26.7,47.5,6.6,94.7,61.1,135.6,47.6),14,2))
```

```
##       [,1]  [,2]
##  [1,]    1  13.1
##  [2,]    1  15.3
##  [3,]    1  25.8
##  [4,]    1   1.8
##  [5,]    1   4.9
##  [6,]    1  55.4
##  [7,]    1  39.3
##  [8,]    1  26.7
##  [9,]    1  47.5
## [10,]    1   6.6
## [11,]    1  94.7
## [12,]    1  61.1
## [13,]    1 135.6
## [14,]    1  47.6
```

```r
(y = c(27.3,42.4,38.7,4.5,23.0,166.3,109.7,80.1,150.7,20.3,189.7,131.3,404.2,149.0))
```

```
##  [1]  27.3  42.4  38.7   4.5  23.0 166.3 109.7  80.1 150.7  20.3 189.7
## [12] 131.3 404.2 149.0
```

```r
(b = solve(t(X)%*%X,t(X)%*%y))
```

```
##           [,1]
## [1,] -1.233836
## [2,]  2.701553
```

Question 5.c)

```r
(e = y - X%*%b)
```

```
##              [,1]
##  [1,]  -6.8565106
##  [2,]   2.3000724
##  [3,] -29.7662361
##  [4,]   0.8710405
##  [5,]  10.9962256
##  [6,]  17.8677893
##  [7,]   4.7627957
##  [8,]   9.2023660
##  [9,]  23.6100596
## [10,]   3.7035852
## [11,] -64.9032511
## [12,] -32.5310639
## [13,]  39.1032233
## [14,]  21.6399042
```

```r
(SSRes = sum(e^2))
```

```
## [1] 9325.833
```

```r
(s2 = SSRes/(14 - (1 + 1)))
```

```
## [1] 777.1528
```

Question 5.d)

```r
(c(1,18))%*%b
```

```
##          [,1]
## [1,] 47.39412
```

Question 5.e)

```r
(inverse = solve(t(X)%*%X))
```

```
##              [,1]          [,2]
## [1,]  0.163081936 -2.230009e-03
## [2,] -0.002230009  5.425812e-05
```

```r
(H = X %*% inverse %*% t(X))
```

```
##              [,1]         [,2]         [,3]        [,4]        [,5]
##  [1,]  0.11396694  0.110624640  0.094672751  0.13113421  0.12642461
##  [2,]  0.11062464  0.107544949  0.092846423  0.12644305  0.12210349
##  [3,]  0.09467275  0.092846423  0.084129856  0.10405344  0.10147998
##  [4,]  0.13113421  0.126443053  0.104053438  0.15522970  0.14861943
##  [5,]  0.12642461  0.122103488  0.101479976  0.14861943  0.14253059
##  [6,]  0.04970362  0.051410579  0.059557437  0.04093605  0.04334131
##  [7,]  0.07416318  0.073948319  0.072922840  0.07526679  0.07496403
##  [8,]  0.09330545  0.091586549  0.083382721  0.10213433  0.09971225
##  [9,]  0.06170552  0.062469470  0.066115616  0.05778157  0.05885805
## [10,]  0.12384192  0.119723727  0.100068722  0.14499445  0.13919154
## [11,] -0.01000202 -0.003603902  0.026932570 -0.04286508 -0.03384955
## [12,]  0.04104402  0.043431380  0.054825586  0.02878169  0.03214569
## [13,] -0.07213842 -0.060858159 -0.007020536 -0.13007796 -0.11418304
## [14,]  0.06155359  0.062329484  0.066032601  0.05756833  0.05866164
##             [,6]       [,7]         [,8]       [,9]       [,10]
##  [1,] 0.04970362 0.07416318  0.093305447 0.06170552  0.12384192
##  [2,] 0.05141058 0.07394832  0.091586549 0.06246947  0.11972373
##  [3,] 0.05955744 0.07292284  0.083382721 0.06611562  0.10006872
##  [4,] 0.04093605 0.07526679  0.102134329 0.05778157  0.14499445
##  [5,] 0.04334131 0.07496403  0.099712246 0.05885805  0.13919154
##  [6,] 0.08252382 0.07003197  0.060255739 0.07639427  0.04466033
##  [7,] 0.07003197 0.07160437  0.072834942 0.07080352  0.07479800
##  [8,] 0.06025574 0.07283494  0.082679536 0.06642814  0.09838401
##  [9,] 0.07639427 0.07080352  0.066428143 0.07365098  0.05944838
## [10,] 0.04466033 0.07479800  0.098384007 0.05944838  0.13600930
## [11,] 0.11301634 0.06619375  0.029549982 0.09004128 -0.02890555
## [12,] 0.08694639 0.06947528  0.055802232 0.07837361  0.03399047
## [13,] 0.14475029 0.06219926 -0.002405883 0.10424388 -0.10546647
## [14,] 0.07647186 0.07079375  0.066350011 0.07368571  0.05926119
##              [,11]      [,12]        [,13]      [,14]
##  [1,] -0.010002020 0.04104402 -0.072138423 0.06155359
##  [2,] -0.003603902 0.04343138 -0.060858159 0.06232948
##  [3,]  0.026932570 0.05482559 -0.007020536 0.06603260
##  [4,] -0.042865080 0.02878169 -0.130077960 0.05756833
##  [5,] -0.033849550 0.03214569 -0.114183042 0.05866164
```

```
##  [6,]  0.113016338 0.08694639  0.144750286 0.07647186
##  [7,]  0.066193748 0.06947528  0.062199265 0.07079375
##  [8,]  0.029549982 0.05580223 -0.002405883 0.06635001
##  [9,]  0.090041278 0.07837361  0.104243884 0.07368571
## [10,] -0.028905550 0.03399047 -0.105466475 0.05926119
## [11,]  0.227309989 0.12959328  0.346256817 0.09033210
## [12,]  0.129593280 0.09313182  0.173976424 0.07848213
## [13,]  0.346256817 0.17397642  0.555967177 0.10475662
## [14,]  0.090332102 0.07848213  0.104756624 0.07372098
```

```r
i = 13 # index of scallop is 13
(standardised_residuals = e[i]/sqrt(s2*(1 - H[i,i])))
```

```
## [1] 2.104999
```

Question 5.f)

```r
(D = (standardised_residuals^2/(1 + 1))*(H[i,i]/(1 - H[i,i])))
```

```
## [1] 2.774008
```