Problem Set 1: Matrix Algebra Review

Donggyun Kim
27008257
Lab 102

2/1/2018

Problem 1

```
X <- matrix(c(2, 3, -1, 4), nrow = 2, byrow = 1)
X
```

```
##      [,1] [,2]
## [1,]    2    3
## [2,]   -1    4
```

```
Y <- matrix(c(2, 0, 1, 1, -2, 3), nrow = 2, byrow = 1)
Y
```

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

```
Z <- matrix(c(1, 1, -1, 1, 0, 2), nrow = 3, byrow = 1)
Z
```

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

```
W <- matrix(c(1, 0, 8, 3), nrow = 2, byrow = 1)
W
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]    8    3
```

```
I <- diag(2)
I
```

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

Problem 2

a. $\mathbf{X} + \mathbf{Y}$

```
X + Y
```

```
## Error in X + Y: non-conformable arrays
```

```r
dim(X)
```

```
## [1] 2 2
```

```r
dim(Y)
```

```
## [1] 2 3
```

Operation cannot be performed. **X** and **Y** have different dimensions.

b. $\mathbf{X} + \mathbf{W}$

```r
X + W
```

```
##      [,1] [,2]
## [1,]    3    3
## [2,]    7    7
```

c. **X** - **I**

```r
X - I
```

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

d. **XY**

```r
X %*% Y
```

```
##      [,1] [,2] [,3]
## [1,]    7   -6   11
## [2,]    2   -8   11
```

e. **XI**

```r
X %*% I
```

```
##      [,1] [,2]
## [1,]    2    3
## [2,]   -1    4
```

f. $\mathbf{X} + (\mathbf{Y} + \mathbf{Z})$

```r
X + (Y + Z)
```

```
## Error in Y + Z: non-conformable arrays
```

```r
dim(X)
```

```
## [1] 2 2
```

```r
dim(Y)
```

```
## [1] 2 3
```

```r
dim(Z)
```

```
## [1] 3 2
```

Operation cannot be performed. **X**, **Y**, and **Z** have different dimensions.

g. $\mathbf{Y}(\mathbf{I} + \mathbf{W})$

```r
Y %*% (I + W)
```

```
## Error in Y %*% (I + W): non-conformable arguments
```

```r
dim(I + W)
```

```
## [1] 2 2
```

```r
dim(Y)
```

```
## [1] 2 3
```

$\mathbf{Y}_{2\times 3}$ and $(\mathbf{I}+\mathbf{W})_{2\times 2}$ do not match dimensions for matrix products.

Problem 3

a. Every orthogonal matrix is nonsingular: TRUE
   For orhogonal matrices,
   $\mathbf{Q}\mathbf{Q}^\top = \mathbf{I} \implies \mathbf{Q}^{-1} = \mathbf{Q}^\top$
   A square matrix is singular if and only if its determinant is 0. If its determinant is 0, then it does not have an inverse matrix.
b. Every nonsingular matrix is orthogonal: FALSE
   Let $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
   $\det(\mathbf{A}) = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2 \neq 0$
   $\mathbf{A}$ is a nonsingular matrix. However,
   $\mathbf{A}^\top = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^\top = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
   $\mathbf{A}\mathbf{A}^\top = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1+4 & 3+8 \\ 3+8 & 9+16 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 11 & 25 \end{bmatrix} \neq \mathbf{I}$
c. Every matrix of full rank is square: TRUE
d. Every square matrix is of full rank: FALSE
e. Every nonsingular matrix is of full rank: TRUE

Problem 4

$(\mathbf{A}_{ij})^\top = \mathbf{A}_{ji} \Leftrightarrow (\mathbf{A}^\top)_{ij} = \mathbf{A}_{ji}$
For $\mathbf{X}_{n\times p}$, $\mathbf{Y}_{p\times q}$, and $\mathbf{Z}_{q\times r}$,

$$(\mathbf{XYZ})_{ij} = \sum_{k=1}^{p}\sum_{l=1}^{q} \mathbf{X}_{ik}\mathbf{Y}_{kl}\mathbf{Z}_{lj}$$

$$\begin{aligned}
((\mathbf{XYZ})_{ij})^\top &= (\mathbf{XYZ})_{ji} \\
&= \sum_{k=1}^{p}\sum_{l=1}^{q} \mathbf{X}_{jk}\mathbf{Y}_{kl}\mathbf{Z}_{li} \\
&= \sum_{k=1}^{p}\sum_{l=1}^{q} \mathbf{Z}_{li}\mathbf{Y}_{kl}\mathbf{X}_{jk} \\
&= \sum_{k=1}^{p}\sum_{l=1}^{q} (\mathbf{Z}^\top)_{il}(\mathbf{Y}^\top)_{lk}(\mathbf{X}^\top)_{kj} \\
&= (\mathbf{Z}^\top\mathbf{Y}^\top\mathbf{X}^\top)_{ij}
\end{aligned}$$

$\therefore (\mathbf{XYZ})^\top = \mathbf{Z}^\top\mathbf{Y}^\top\mathbf{X}^\top$

Problem 5

From linear algebra class, we know that
$\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}^* \mathbf{y} \rangle$, where $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$ and $\mathbf{A}^*$ is conjugate traspose of $\mathbf{A}$
Since $\mathbf{A}_{n \times n}$ is symmetric and real value matrix,
$\mathbf{A}^* = \mathbf{A}^\top = \mathbf{A}$
$\lambda_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \langle \lambda_i \mathbf{v}_i, \mathbf{v}_j \rangle = \langle \mathbf{A}\mathbf{v}_i, \mathbf{v}_j \rangle = \langle \mathbf{v}_i, \mathbf{A}\mathbf{v}_j \rangle = \langle \mathbf{v}_i, \lambda_j \mathbf{v}_j \rangle = \lambda_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle$.
$\lambda_i \neq \lambda_j \implies \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0, \forall i \neq j, 1 \leq i, j \leq n$
Therefore, $\mathbf{v}_i$ and $\mathbf{v}_j$ associated with two distinct eigenvalues λ_i and λ_j of $\mathbf{A}$ are mutually orthogonal.

Problem 6

Function `inner_product`

```
inner_product <- function(v, u) {
  if(length(v) != length(u)) {
    stop("Vector v and u must have the same length!")
  }
  as.numeric(t(v) %*% u)  # display purpose
}
v <- c(1, 3, 5)
u <- c(1, 2, 3)

inner_product(v, u)
```

```
## [1] 22
```

Function `projection`

```
projection <- function(v, u) {
  (inner_product(v, u) / inner_product(u, u)) * u
}

v <- c(1, 3, 5)
u <- c(1, 2, 3)

projection(v, u)
```

```
## [1] 1.571429 3.142857 4.714286
```

Problem 7

Gram-Schmit orthonormalization

```
x <- c(1, 2, 3)
y <- c(3, 0, 2)
z <- c(3, 1, 1)

u1 <- x
u1
```

```
## [1] 1 2 3
```

```r
u2 <- y - projection(y, u1)
u2
```

```
## [1]  2.35714286 -1.28571429  0.07142857
```

```r
u3 <- z - projection(z, u2) - projection(z, u1)
u3
```

```
## [1]  0.5148515  0.9009901 -0.7722772
```

```r
e1 <- u1/as.numeric(sqrt(inner_product(u1, u1)))
e1
```

```
## [1] 0.2672612 0.5345225 0.8017837
```

```r
e2 <- u2/as.numeric(sqrt(inner_product(u2, u2)))
e2
```

```
## [1]  0.87758509 -0.47868278  0.02659349
```

```r
e3 <- u3/as.numeric(sqrt(inner_product(u3, u3)))
e3
```

```
## [1]  0.3980149  0.6965260 -0.5970223
```

Problem 8

l_p-norm

```r
lp_norm <- function(x, p = 1) {
  if (p == "max") {
    max(abs(x))
  } else {
    (sum(abs(x)^p))^(1/p)
  }
}

x <- c(-1, 2, -5)
lp_norm(x)
```

```
## [1] 8
```

```r
lp_norm(x, p = 2)
```

```
## [1] 5.477226
```

```r
lp_norm(x, p = "max")
```

```
## [1] 5
```

Problem 9

a.

```r
zero <- rep(0, 10)
lp_norm(zero)
```

```
## [1] 0
```

b.

```r
ones <- rep(1, 5)
lp_norm(ones, p = 2)
```

```
## [1] 2.236068
```

c.

```r
u <- rep(0.4472136, 5)
lp_norm(u, p = 2)
```

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

d.

```r
u <- 1:500
lp_norm(u, p = 100)
```

```
## [1] 508.5663
```

e.

```r
u <- 1:500
lp_norm(u, p = "max")
```

```
## [1] 500
```

Problem 10

a. $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \implies (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \implies \det(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0}$
$b\mathbf{A}\mathbf{v} = b\lambda\mathbf{v} \implies b(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \implies (b\mathbf{A} - b\lambda\mathbf{I})\mathbf{v} = \mathbf{0}$
$\det(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0} \implies \det(b\mathbf{A} - b\lambda\mathbf{I}) = \mathbf{0}$
$b\mathbf{A} = \mathbf{B}, b\lambda = \lambda_b \implies (\mathbf{B} - \lambda_b\mathbf{I})\mathbf{v} = \mathbf{0}$
$\mathbf{B}\mathbf{v} = \lambda_b\mathbf{v}$ $\mathbf{B}$ has an eigenvalue λ_b with eigenvector $\mathbf{v}$
Therefore, $b\mathbf{A}$ has eigenvalue $b\lambda$ with eigenvector $\mathbf{v}$

b. $\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \implies (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \implies \det(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0}$
$(\mathbf{A} + c\mathbf{I})\mathbf{v} = \lambda\mathbf{v} + c\mathbf{v} = (\lambda + c)\mathbf{v} \implies (\mathbf{A} + c\mathbf{I})\mathbf{v} - (\lambda + c)\mathbf{v} = \mathbf{0}$
$\implies (\mathbf{A} + c\mathbf{I} - \lambda\mathbf{I} - c\mathbf{I})\mathbf{v} = \mathbf{0} \implies (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$
$\mathbf{A} + c\mathbf{I} = \mathbf{C}, \lambda + c = \lambda_c \implies (\mathbf{C} - \lambda_c\mathbf{I})\mathbf{v} = \mathbf{0}$ and $\det(\mathbf{C} - \lambda_c\mathbf{I}) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0$
$\mathbf{C}\mathbf{v} = \lambda_c\mathbf{v}$ $\mathbf{C}$ has an eigenvalue λ_c with eigenvector $\mathbf{v}$
Therefore, $\mathbf{A} + c\mathbf{I}$ has an eigenvalue $\lambda + c$ with eigenvector $\mathbf{v}$

Problem 11

a.

```r
X <- as.matrix(state.x77[, 1:5])
head(X, 10)
```

```
##            Population Income Illiteracy Life Exp Murder
## Alabama          3615   3624        2.1    69.05   15.1
## Alaska            365   6315        1.5    69.31   11.3
## Arizona          2212   4530        1.8    70.55    7.8
```

```
## Arkansas          2110    3378         1.9    70.66    10.1
## California       21198    5114         1.1    71.71    10.3
## Colorado          2541    4884         0.7    72.06     6.8
## Connecticut       3100    5348         1.1    72.48     3.1
## Delaware           579    4809         0.9    70.06     6.2
## Florida           8277    4815         1.3    70.66    10.7
## Georgia           4931    4091         2.0    68.54    13.9
```

```r
n <- nrow(X)
n
```

```
## [1] 50
```

```r
p <- ncol(X)
p
```

```
## [1] 5
```

b.

```r
D <- diag(n) / n
sum(diag(D))
```

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

c.

```r
g <- crossprod(X, D) %*% rep(1, n)
g
```

```
##                   [,1]
## Population 4246.4200
## Income     4435.8000
## Illiteracy    1.1700
## Life Exp     70.8786
## Murder        7.3780
```

d.

```r
Xc <- X - rep(1, n) %*% t(g)
colMeans(Xc)
```

```
##    Population        Income    Illiteracy      Life Exp        Murder
## -3.637979e-14  7.275958e-13 -5.950795e-16 -5.968559e-15  7.283063e-16
```

e.

```r
V <- crossprod(X, D) %*% X - tcrossprod(g)
V
```

```
##                  Population      Income  Illiteracy    Life Exp      Murder
## Population 19533050.0836 559805.1840  287.010600 -399.685612 5550.253240
## Income       559805.1840 370021.8400 -160.428000  275.049920 -511.456400
## Illiteracy      287.0106   -160.4280    0.364100   -0.471882    1.550140
## Life Exp       -399.6856    275.0499   -0.471882    1.765980   -3.792091
## Murder         5550.2532   -511.4564    1.550140   -3.792091   13.354916
```

f.

```r
Ds <- diag(1 / sqrt(diag(V)))
diag(Ds)
```

```
## [1] 0.0002262637 0.0016439414 1.6572562309 0.7525010403 0.2736398979
```

g.

```r
Z <- Xc %*% Ds
colMeans(Z)
```

```
## [1] -1.755540e-17  1.183567e-15 -9.983681e-16 -4.484993e-15  1.925890e-16
```

```r
apply(Z, 2, var)
```

```
## [1] 1.020408 1.020408 1.020408 1.020408 1.020408
```

h.

```r
R <- Ds %*% V %*% Ds
R
```

```
##             [,1]       [,2]       [,3]        [,4]       [,5]
## [1,]  1.00000000  0.2082276  0.1076224 -0.06805195  0.3436428
## [2,]  0.20822756  1.0000000 -0.4370752  0.34025534 -0.2300776
## [3,]  0.10762237 -0.4370752  1.0000000 -0.58847793  0.7029752
## [4,] -0.06805195  0.3402553 -0.5884779  1.00000000 -0.7808458
## [5,]  0.34364275 -0.2300776  0.7029752 -0.78084575  1.0000000
```

i.

```r
R <- crossprod(Z, D) %*% Z
R
```

```
##             [,1]       [,2]       [,3]        [,4]       [,5]
## [1,]  1.00000000  0.2082276  0.1076224 -0.06805195  0.3436428
## [2,]  0.20822756  1.0000000 -0.4370752  0.34025534 -0.2300776
## [3,]  0.10762237 -0.4370752  1.0000000 -0.58847793  0.7029752
## [4,] -0.06805195  0.3402553 -0.5884779  1.00000000 -0.7808458
## [5,]  0.34364275 -0.2300776  0.7029752 -0.78084575  1.0000000
```