MAST90104: A first course in Statistical Learning

Week 1 Practical/Workshop Solutions

Workshop questions

1. Show that $X^T X$ is a symmetric matrix.
   Solution:
$$(X^T X)^T = X^T (X^T)^T = X^T X.$$

2. (a) Let
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
   be a nonsingular 2×2 matrix. Show by direct multiplication that
$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
   Solution:
$$\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} da - bc & db - bd \\ -ca + ac & -cb + ad \end{bmatrix} = I.$$
   (b) Find the inverse of
$$\begin{bmatrix} 3 & 1 \\ 4 & -2 \end{bmatrix}.$$
   Solution: From above, the inverse is
$$-\frac{1}{10} \begin{bmatrix} -2 & -1 \\ -4 & 3 \end{bmatrix}.$$

3. Let
$$A = \begin{bmatrix} -6 & 3 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$
   and
$$B = \begin{bmatrix} 0 & 6 & -1 \\ 8 & -1 & 0 \\ 0 & 2 & 3 \end{bmatrix}$$
   (a) Find the product AB
   Solution:
$$AB = \begin{bmatrix} 24 & -39 & 6 \\ -8 & 9 & 2 \end{bmatrix}$$
   (b) Does BA exist?
   Solution: No, the dimensions do not match
   (c) Can we calculate $B^T A^T$? If so, what is it? **Solution**:
$$B^T A^T = (AB)^T = \begin{bmatrix} 24 & -8 \\ -39 & 9 \\ 6 & 2 \end{bmatrix}$$

4. Let $\mathbf{x} = \begin{bmatrix} 1 & 6 & -4 & 2 \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} 2 & -1 & -1 & 0 \end{bmatrix}^T$
   (a) What is the norm of $\mathbf{x}$?
   Solution: $\|\mathbf{x}\| = \sqrt{1^2 + 6^2 + (-4)^2 + 2^2} = \sqrt{57} \approx 7.55$

(b) Are $\mathbf{x}$ and $\mathbf{y}$ orthogonal ?

Solution: Yes, because $\mathbf{x}^T\mathbf{y} = 0$

5. Prove that for any matrix A

$$r(A) = r(A^T) = r(A^TA).$$

You may use the fact that pre- or post-multiplying by a non-singular matrix does not change the rank.

Solution: Suppose that A is of dimension $m \times n$ and $r(A) = k$. Then the column space of A has dimension k; let $\mathbf{c}_1, \ldots, \mathbf{c}_k$ be a basis for this column space. Let $C = [\mathbf{c}_1|\cdots|\mathbf{c}_k]$. Now every column of A can be expressed as a linear combination of columns of C; this means that there is a $k \times n$ matrix R such that $A = CR$.

But now each row of A is expressed as a linear combination of the k rows of R; hence the row space of A has dimension at most k. In other words, $r(A^T) \leq r(A)$. As we can apply this argument equally to A^T, we have $r(A^T) = r(A)$.

For the second part of the problem, we first note that

$$A\mathbf{x} = \mathbf{0} \text{ if and only if } A^TA\mathbf{x} = \mathbf{0}.$$

The only if part is obvious. For the if part we have

$$A^TA\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x}^TA^TA\mathbf{x} = \mathbf{0} \Rightarrow \|A\mathbf{x}\|^2 = \mathbf{0} \Rightarrow A\mathbf{x} = \mathbf{0}.$$

Let $B = [\mathbf{b}_1|\cdots|\mathbf{b}_n]$, then $B\mathbf{x} = \mathbf{0}$ is equivalent to $x_1\mathbf{b}_1 + \cdots x_n\mathbf{b}_n = \mathbf{0}$. So since $A\mathbf{x} = \mathbf{0}$ if and only if $A^TA\mathbf{x} = \mathbf{0}$, we see that a linear combination of the columns of A is zero precisely when the same linear combination of columns of A^TA is zero. Thus they have the same number of linearly independent columns and so by definition $r(A^TA) = r(A)$.

True or false:

Only (c) is true.

Practical exercises

The following are taken from Chapter 2 of spuRs (Introduction to Scientific Programming and Simulation Using R).

1. Set $a = 1.1$, $b = 1.2$, and $x = 123$. Give R assignment statements that set the variable z to

   (a) x^{a^b}

   (b) $(x^a)^b$

   (c) $3x^3 + 2x^2 + 6x + 1$ (try to minimise the number of operations required)

   (d) the second-to-last digit of x before the decimal point (hint: use `floor(x)` and/or `%%`)

   (e) $z + 1$

   Solution:

```
> x <- 123
> a <- 1.1
> b <- 1.2
> # a
> (z <- x^(a^b))
[1] 220.3624
> (z <- x^a^b)
[1] 220.3624
```

```
> # b
> (z <- (x^a)^b)
[1] 573.6867

> # c
> (z <- 3*x^3 + 2*x^2 + 6*x + 1) #8 operations
[1] 5613598

> (z <- (3*x + 2)*(x^2 + 2) - 3) #6 operations
[1] 5613598

> (z <- sum((x^(3:0))*c(3, 2, 6, 1))) #vectorised
[1] 5613598

> # d
> y <- abs(x)
> (z <- (y %% 100 - y %% 10)/10)
[1] 2

> (z <- floor(y/10 - floor(y/100)*10))
[1] 2

> # e
> (z <- z + 1)
[1] 3
```

2. Give R expressions that return the following matrices and vectors

   (a) $(1, 2, 3, 4, 5, 6, 7, 8, 7, 6, 5, 4, 3, 2, 1)$

   (b) $(1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5)$

   (c) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

   (d) $\begin{pmatrix} 0 & 2 & 3 \\ 0 & 5 & 0 \\ 7 & 0 & 0 \end{pmatrix}$

   Solution:

```
> # a
> c(1:8, 7:1)
 [1] 1 2 3 4 5 6 7 8 7 6 5 4 3 2 1

> # b
> rep(1:5, 1:5)
 [1] 1 2 2 3 3 3 4 4 4 4 5 5 5 5 5

> # c
> matrix(1, 3, 3) - diag(3)
     [,1] [,2] [,3]
[1,]    0    1    1
```

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


> # d
> matrix(c(0,0,7, 2,5,0, 3,0,0), 3, 3)
     [,1] [,2] [,3]
[1,]    0    2    3
[2,]    0    5    0
[3,]    7    0    0
```

3. Use R to produce a vector containing all integers from 1 to 100 that are not divisible by 2, 3, or 7.

   Solution:

```
> x <- 1:100
> idx <- (x %% 2 != 0) & (x %% 3 != 0) & (x %% 7 != 0)
> x[idx]


 [1]  1  5 11 13 17 19 23 25 29 31 37 41 43 47 53 55 59 61 65 67 71 73 79 83 85
[26] 89 95 97
```

4. Which of the following assignments will be successful? What will the vectors x, y, and z look like at each stage?

```
> rm(list=ls())
> x<-1
> x[3]<-3
> x
[1]  1 NA  3
> y <- c()
> y[2]<-2
> y[3] <- y[1]
> y[2] <- y[4]
> y
[1] NA NA NA
> z[1] <- 0
Error: object 'z' not found
```

5. Build a 10×10 identity matrix. Then make all the non-zero elements 5. Do this latter step in at least two different ways.

   Solution:

```
   > Id <- diag(10)
   > Id <- 5*Id # one way, using the fact that Id is the identity
   > Id[Id != 0] <- 5 # another way, using vector indexing of a matrix
   > diag(Id) <- 5
```