MAST90104: A First Course in Statistical Learning

Week 1 Workshop/Lab

Workshop questions

1. Show that $X^T X$ is a symmetric matrix.
2. (a) Let
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
be a nonsingular 2×2 matrix. Show by direct multiplication that
$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
   (b) Find the inverse of
$$\begin{bmatrix} 3 & 1 \\ 4 & -2 \end{bmatrix}.$$
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
   (b) Does BA exist?
   (c) Can we calculate $B^T A^T$? If so, what is it?
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 $\mathbf{x}$?
   (b) Are $\mathbf{x}$ and $\mathbf{y}$ orthogonal?
5. Prove that for any matrix A
$$r(A) = r(A^T) = r(A^T A).$$
You may use the fact that pre- or post-multiplying by a non-singular matrix does not change the rank.

True or false:

Tick the option(s) that is/are always TRUE.

☐ (EXAMPLE) $2 + 1 = 0$

☑ (EXAMPLE) $1 + 1 = 2$

☐ (a) $\mathbf{A}$ and $\mathbf{C}$ are two nonsingular p by p square matrices. Then, $\mathbf{A} - \mathbf{C} = (\mathbf{A} - \mathbf{C})^{-1}$.

☐ (b) $\mathbf{A}$ is a p by p square matrix such that $\det(\mathbf{A}) < 0$. Then, $\det(\mathbf{A}^{-1}) > 0$.

☐ (c) Consider the design matrix below:

$$\mathbf{X} = \begin{pmatrix} 1 & 5 & -5 \\ 1 & 4 & -4 \\ 1 & 3 & -3 \\ 1 & 2 & -2 \\ 1 & 1 & -2 \\ 1 & 0 & 0 \end{pmatrix},$$

Then, $\mathbf{X}$ has full rank.

Practical questions

Open RStudio on your computer. Set the working directory (using `setwd()`). Open a new `.R` file

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) $z + 1$
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) $\begin{pmatrix} 0 & 2 & 3 \\ 0 & 5 & 0 \\ 7 & 0 & 0 \end{pmatrix}$
3. Use R to produce a vector containing all integers from 1 to 100 that are not divisible by 2, 3, or 7.
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
y <- c()
y[2] <- 2
y[3] <- y[1]
y[2] <- y[4]
z[1] <- 0
```

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.