

# MAST90104: An First Course in Statistical Learning

## Week 2 Workshop/Lab

1. Let  $X$  be a  $10 \times 5$  matrix of full rank and let  $H = X(X^T X)^{-1} X^T$ .

Find  $\text{tr}(H)$  and  $r(H)$ .

2. Let

$$A = \begin{bmatrix} 3 & 1 & 8 \\ 1 & 0 & 1 \\ 2 & 1 & -4 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

Let  $z = \mathbf{y}^T A \mathbf{y}$ . Write out  $z$  in full, then find  $\frac{\partial z}{\partial \mathbf{y}}$  directly *and* using the matrix formula.

3. Let  $\mathbf{y} = [y_1 \ y_2 \ y_3]^T$  be a random vector with mean  $\boldsymbol{\mu} = [1 \ 3 \ 2]^T$ , and assume that  $\text{Var } y_i = 4$  and  $\text{Cov}(y_i, y_j) = 0$ .

(a) Write down  $\text{Var } \mathbf{y}$ .

(b) Let

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & 0 \\ -1 & 6 & 1 \end{bmatrix}$$

and find  $\text{Var } A\mathbf{y}$  and  $\mathbb{E}[\mathbf{y}^T A \mathbf{y}]$

4. Prove corollaries 3.6 and 3.7 from the lectures.

5. Let  $\mathbf{y} = [y_1 \ y_2]^T$  be a normal random vector with mean and variance

$$\boldsymbol{\mu} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \text{ and } V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Let

$$A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

(a) Find the distributions of  $\mathbf{y}^T A \mathbf{y}$  and  $\mathbf{y}^T B \mathbf{y}$ .

(b) Are  $\mathbf{y}^T A \mathbf{y}$  and  $\mathbf{y}^T B \mathbf{y}$  independent?

(c) What is the distribution of  $\mathbf{y}^T A \mathbf{y} + \mathbf{y}^T B \mathbf{y}$ ?

6. Let  $y_1, \dots, y_n$  be an i.i.d. normal sample. Show that

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \text{ and } s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

are independent. (*Hint*: Express them as a random “vector” and quadratic form respectively.)

## R exercises

The following are taken from Chapter 3 of spuRs (Introduction to Scientific Programming and Simulation Using R).

1. Consider the function  $y = f(x)$  defined by

$$\frac{x}{f(x)} \begin{cases} \leq 0 & \in (0, 1] & > 1 \\ -x^3 & x^2 & \sqrt{x} \end{cases}$$

Supposing that you are given  $x$ , write an R expression for  $y$  using **if** statements.

Add your expression for  $y$  to the following program, then run it to plot the function  $f$ .

```

# input
x.values <- seq(-2, 2, by = 0.1)

# for each x calculate y
n <- length(x.values)
y.values <- rep(0, n)
for (i in 1:n) {
  x <- x.values[i]
  # your expression for y goes here
  y.values[i] <- y
}

# output
plot(x.values, y.values, type = "l")

```

Your plot should look like Figure 1. Do you think  $f$  has a derivative at 1? What about at 0?

- Let  $h(x, n) = 1 + x + x^2 + \cdots + x^n = \sum_{i=0}^n x^i$ . Write an R program to calculate  $h(x, n)$  using a for loop.
- The function  $h(x, n)$  from Exercise 2 is the finite sum of a geometric sequence. It has the following explicit formula, for  $x \neq 1$ ,

$$h(x, n) = \frac{1 - x^{n+1}}{1 - x}.$$

Test your program from Exercise 2 against this formula using the following values

$x$	$n$	$h(x, n)$
0.3	55	1.428571
6.6	8	4243335.538178

You should use the computer to calculate the formula rather than doing it yourself.

- First write a program that achieves the same result as in Exercise 2 but using a **while** loop. Then write a program that does this using vector operations (and no loops).  
If it doesn't already, make sure your program works for the case  $x = 1$ .
- To rotate a vector  $(x, y)^T$  anticlockwise by  $\theta$  radians, you premultiply it by the matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Write a program in R that does this for you.

- Given a vector  $\mathbf{x}$ , calculate its geometric mean using both a for loop and vector operations. (The geometric mean of  $x_1, \dots, x_n$  is  $(\prod_{i=1}^n x_i)^{1/n}$ .)  
You might also like to have a go at calculating the harmonic mean,  $(\sum_{i=1}^n 1/x_i)^{-1}$ , and then check that if the  $x_i$  are all positive, the harmonic mean is always less than or equal to the geometric mean, which is always less than or equal to the arithmetic mean.
- A room contains 100 toggle switches, originally all turned off. 100 people enter the room in turn. The first toggles every switch, the second toggles every second switch, the third every third switch, and so on, to the last person who toggles the last switch only.  
At the end of this process, which switches are turned on?

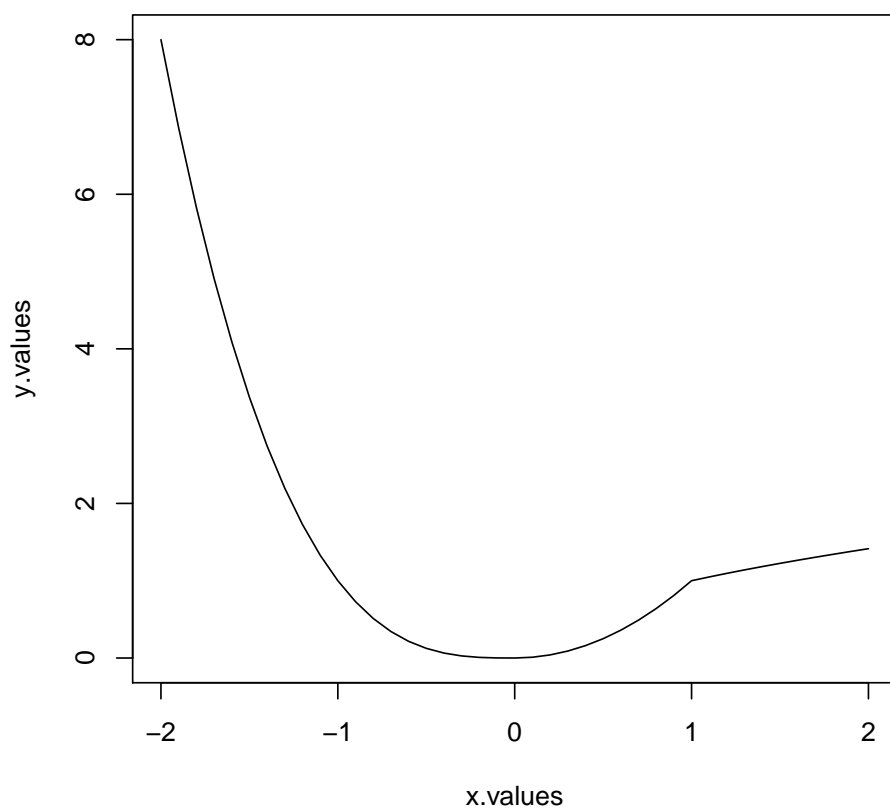


Figure 1: The graph produced by Exercise 1.