MAST90104: A First Course in Statistical Learning

Week 3 Workshop/Lab

R exercises

Questions 1-4 are taken from Chapter 3 of spuRs (Introduction to Scientific Programming and Simulation Using R).

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

x	≤ 0	$\in (0, 1]$	> 1
$f(x)$	$-x^3$	x^2	$\sqrt{x}$

   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?

2. 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.

3. 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.

4. 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 $x = 1$.

5. We return to the house price example in Module 4.

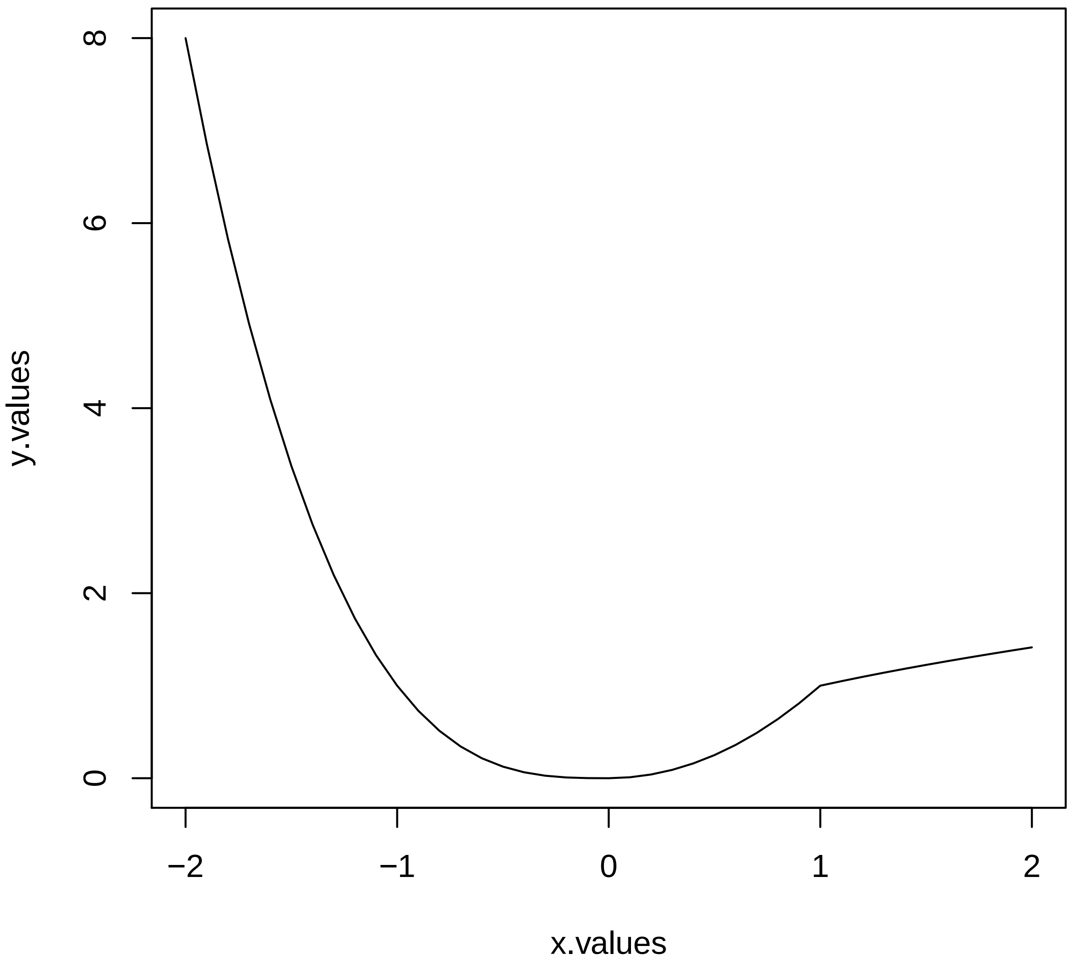

Figure 1: The graph produced by Exercise 1.

Price (× \$10k)	Age (years)	Area ($\times 100m^2$)
50	1	1
40	5	1
52	5	2
47	10	2
65	20	3

Recall that we want to predict a house price based on its age and area. The linear model is of the form $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

Create a *data frame* in R with this data. Fit the linear model using function `lm()` in R.

Workshop questions

1. Let X be a 10×5 matrix of full rank and let $H = X(X^T X)^{-1} X^T$.
   Find $tr(H)$ 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 $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T$ be a random vector with mean $\boldsymbol{\mu} = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix}^T$, and assume that $\operatorname{Var} y_i = 4$ and $\operatorname{Cov}(y_i, y_j) = 0$.
   (a) Write down $\operatorname{Var} \mathbf{y}$.
   (b) Let
$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & 0 \\ -1 & 6 & 1 \end{bmatrix}$$
   and find $\operatorname{Var} A\mathbf{y}$ and $\mathbb{E}[\mathbf{y}^T A \mathbf{y}]$

4. Let $\mathbf{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$ be a normal random vector with mean and variance
$$\mathbf{M} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} \text{ and } V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
   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 $\mathbf{y}^T A \mathbf{y} + \mathbf{y}^T B \mathbf{y}$?

5. Let $\mathbf{Y} \sim \text{MVN}(\mathbf{0}, \mathbf{I})$ be a $n \times 1$ random vector and let $\mathbf{A}$ be a n be a symmetric square matrix of order n (n by n symmetric matrix). Show that $\mathbf{Y}^T \mathbf{A} \mathbf{Y}$ has a (central) χ^2 distribution with k degrees of freedom if and only if $\mathbf{A}$ is idempotent and has rank k.

6. Let $\mathbf{Y} \sim \text{MVN}(\boldsymbol{\mu}, \sigma^2 \mathbf{I})$ be a $n \times 1$ random vector and let $\mathbf{A}$ be a n be a symmetric square matrix of order n. Show that $\frac{1}{\sigma^2} \mathbf{Y}^T \mathbf{A} \mathbf{Y}$ has a noncentral χ^2 distribution with k degrees of freedom and non-centrality parameter $\lambda = \frac{1}{2\sigma^2} \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}$ if and only if $\mathbf{A}$ is idempotent and has rank k.