MTH3045: Statistical Computing
Example partial coursework

Ben Youngman
(b.youngman@exeter.ac.uk, Laver 817)

Instructions

- Submit your work through BART **by midday on [whatever date is given]**.
- This coursework should be *your* work and not a collaboration.
- You should submit your work as a `.zip` file containing a RMarkdown document in `.Rmd` format that generates a PDF. (You may also include in the `.zip` file the `.pdf` file that it generates, if you wish.)
- See 'Guidance on Courseworks' on the Assessment tile of the MTH3045 ELE page for fuller details on Coursework expectations and submission information.

Example partial coursework

This is an example that is designed to show part of a MTH3045 coursework and how it can be answered. It demonstrates the style of question you can expect.

1. The p-dimensional multivariate Normal distribution, $MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, has pdf

$$f(\mathbf{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})\right\},$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are its mean vector and variance-covariance matrix, respectively.

(a) Assume that the sample $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are independent realisations from the $MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution, where $\mathbf{y}_i$ are p-vectors, for $i = 1, \ldots, n$. The following function, `dmvn(y, my, Sigma, log = TRUE)`, evaluates their log-likelihood, where `y` is a $p \times n$ matrix with ith column $\mathbf{y}_i$, for $i = 1, \ldots, n$, `mu` and `Sigma` are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively, and `log` is a logical indicating whether the log-likelihood (`log = TRUE`) or likelihood (`log = FALSE`) is returned.

```r
dmvn1 <- function(y, mu, Sigma, log = TRUE) {
  # Function to evaluate multivariate Normal pdf
  # y and mu are p-vectors
  # Sigma is a p x p matrix
  # log is a logical
  # Returns scalar, on log scale, if log == TRUE.
  p <- nrow(y)
  res <- y - mu
```

```
  out <- - 0.5 * determinant(Sigma)$modulus - 0.5 * p * log(2 * pi) -
             0.5 * colSums(res * (solve(Sigma) %*% res))
  out <- sum(out)
  if (!log)
    out <- exp(out)
  out
}
```

Write a function `dmvn2(y, mu, Sigma, log)` that modifies `dmvn1()` so that it calculates the multivariate Normal log-likelihood via the QR decomposition.

(b) Use `dmvn1()` and `dmvn2()` to evaluate the log-likelihood, given $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4, \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ below

$$\mathbf{y}_1 = \begin{pmatrix} 0.85 \\ 1.97 \\ 7.35 \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} -0.35 \\ 2.29 \\ 7.39 \end{pmatrix}, \mathbf{y}_3 = \begin{pmatrix} 4.55 \\ 3.85 \\ 11.29 \end{pmatrix}, \mathbf{y}_4 = \begin{pmatrix} 4.46 \\ 0.50 \\ 4.47 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} 3.5 \\ 1.5 \\ 6.2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} 5.0 & 0.2 & 1.1 \\ 0.2 & 1.4 & 1.2 \\ 1.1 & 1.2 & 3.6 \end{pmatrix},$$

and confirm that both give the same result.