# MAST90104: A First Course in Statistical Learning

## Assignment 1, 2022

Due: 11:59 pm Sunday August 14. Please submit a scanned or electronic copy of your work via the Learning Management System. Late submissions will have their score deducted (10% for every 12 hrs late)

*This assignment is worth 5% of your total mark.*

You may use R for this assignment, but only for matrix calculations and plotting the data. If you do, include your R commands and output.

1. (3 pt) Consider the following statements and decide whether they are TRUE of FALSE. If a statement is false, provide a counterexample.
   - Every nonsingular matrix is of full rank
   - Every nonsingular matrix is orthogonal
   - Every matrix of full rank is orthogonal
2. (4 pt) Prove that a (real) symmetric matrix $A$ positive definite if and only if all of its eigenvalues are strictly positive.
3. (6 pt) Let $\mathbf{y}$ be a 3-dimensional multivariate normal random vector with mean and variance

$$\boldsymbol{\mu} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}, \quad V = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 4 \end{bmatrix}.$$

   Let

$$A = \begin{bmatrix} 1 & -4 & 3 \\ -4 & 3 & 6 \\ 3 & 6 & 2 \end{bmatrix}.$$

   (a) Describe the distribution of $A\mathbf{y}$.

   (b) Find $E[\mathbf{y}^T A\mathbf{y}]$.

   (c) Does $\mathbf{y}^T A\mathbf{y}$ have a (noncentral) chi-square distribution ? Explain your answer.

4. (7pt) The Leaning Tower of Pisa is an architectural wonder. The following table gives measurements of the lean of the Tower of Pisa over time, between 1975 and 1987. Only the last two digits of the year were entered into the computer, so year 1975 appears as 75.

   The variable "lean" represents the differences between where a point on the tower would be if the tower were straight and where it actually is, and coded as tenths of a millimetre in excess of 2.9 meters. For example in 1975 the lean measurement was 2.9642 meters, and is recorded in the table as 642.

| Year | 75 | 76 | 77 | 78 | 79 | 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lean | 642 | 644 | 656 | 667 | 673 | 688 | 696 | 698 | 713 | 717 | 725 | 742 | 757 |

   We will model the lean of the tower over time via a linear model.

   (a) Plot the data; is a linear model appropriate? (*Hint: You can use the function* `plot()` *in R*)

   (b) The linear model is of the form $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Write down the matrices and vectors involved in this equation.

   (c) Is this model a full rank model?

   (d) Using matrices, find the least squares estimators of the parameters.