Mathematical Skills for Data Scientists
Lab Exercises 5 – 4 Marks (Due: 14/11/22)

Gibin Powathil

g.g.powathil@swansea.ac.uk

Exercise 1 (Excel)**.** Do linear regression on the following data:

x	y
1.1	2.7
1.2	2.8
2.4	5.2
3.1	6.8

What values for slope and intercept do you get? Try out forcing the intercept to be zero. Does the slope change? If so, why? 1 mark

Exercise 2 (Matlab)**.** Write a Matlab function that take a data matrix to find the covariance matrix (Input: data matrix and output: covariance of data matrix). You need to use the formula (in matrix form) discussed in the lectures (and not inbuilt command, cov). Use this to find the covariance matrix of the data matrix.

$$\begin{pmatrix} 1.1 & 2.7 & 3.2 & 2.1 \\ 1.2 & 2.8 & 3.4 & 1.8 \\ 2.4 & 5.2 & 1.9 & 2.5 \\ 3.1 & 6.8 & 3.5 & 2 \end{pmatrix}$$

- 1 mark

Exercise 3 (Matlab)**.** Start with data $x_i = y_i = i$ for i ranging from 1 to 10. Add random perturbations of size up to 0.1 to each value. Compute the covariance matrix of the result, and obtain the eigenvectors and eigenvalues. Interpret. [Note: here you can use matlab commands for finding covariance ('cov') and eigenvalue/eigenvectors ('eig')]- 2 marks