M3/4/5N9 Computational Linear Algebra
Project 1 (20% of the final mark)
Due November 5th 2019 (must submit on Blackboard)

Prof Colin Cotter

Autumn Term 2019

Reminders

1. ****Once you submit this project, you will be registered for M3/4/5N9!**
2. Be sure to follow all project guidelines (separate document on Blackboard).
3. Write your code in Python. You should use the `numpy` Python module to implement fast array operations, but do not use any other Python modules.
4. In terms of code, marks will be given for: clear, accessible, readable, re-useable code that makes sensible use of `numpy` array operations, and is well-organised into suitable files.
5. In terms of reporting, marks will be given for: clear, appropriate description that demonstrates understanding of the material.

1 QR factorisation by Householder reflections

1. Write a Python function that implements the QR factorisation of a $n \times m$ matrix A, returning Q and R.
2. Write a selection of tests for your Python function that verify that it correctly produces the QR factorisation, considering different possible cases for A. The tests should be executable by running a single Python script that you provide.
   In your report, briefly describe your tests and how they should convince the reader that you have correctly implemented the QR factorisation.

2 Polynomial fitting by least squares

This section uses the file `readings.csv` provided. It contains measurements from a species of bacteria that responds to ultraviolet light by changing temperature. The first column of data gives the light intensity (in non-dimensional units) whilst the second column gives the change in temperature (in degrees Centigrade). Our aim is to fit a polynomial to this data.

1. Formulate a least squares problem for the polynomial-fitting process.
2. Write a Python function to compute the least-squares best-fit polynomial coefficients for the data, with arbitrary degrees of polynomial. Your function must make use of the QR factorisation, calling the function written in the previous section. Briefly describe how a QR factorisation can be used to solve a least squares problem in the report.
3. Examine the accuracy of the polynomial fit for different polynomial degrees. What is the most appropriate choice of polynomial degree to model the data? Provide some justification in your report.

3 QR analysis of dataset

This section uses the file `readings2.csv` provided. It contains time series of yeast temperatures for different species of yeast. Each row corresponds to a species of yeast, with the (relative) temperature recorded every minute for ten minutes.

Treating this data as a matrix A, with each row being one species, and each column corresponding to the temperature at one measurement time, use your factorisation code to compute the QR factorisation of A.

1. What do you notice about the matrix R?
2. What does this tell you about the matrix A?
3. What does this tell you about the dataset?