Quiz 1

September 22, Fall 2017

1. (40 points)

This question uses data file `dataset_1.csv`. Each column is a variable and each row is an observation.

Write your Python script to:

- calculate the variance of every variable in the data file.
- calculate the covariance between x and y, and between y and z
- do PCA of all the data in the given data file using your own PCA module.

2. (30 points)

Suppose $S = \{v_1, v_2, \cdots, v_n\}$ is an *orthonormal* basis for $\mathbb{R}^n$. Then we can write any vector v as

$$v = \sum_{i=1}^{n} c_i v_i$$

for some constants $v_i, i = 1, 2, \cdots, n$. Find a formula for the constants in terms of S.

3. (30 points)

(1) Manually find the eigenvalue and eigenvectors of the following matrix:

$$A = \begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix}$$

Include the detailed calculation procedure in your solution.

(2) Use the `linalg` module in `numpy` to find the eigenvalues and eigenvectors. Are they the same as your manual solution?