Computational Physics Homework 6

Daniel Bateman

October 2024

1 Link to Github

https://github.com/danbateman01/phys-ga2000/tree/main

2 Abstract

This problem set involves Principle Component Analysis of a data set from the central optical spectra of 9713 nearby galaxies using the astropy package. I completed the code in Google Colab as it has the necessary installed packages and also utilised the generative text tool for a number errors of errors as the run time was over 10 minutes toward the end and it was taking too long to go back for small adjustments. I also took some chunks of code to ChatGPT for fixes and to do sections individually to get graphs for the report as again it was taking too long to execute. A lot of the code is stuck together and I cannot be sure if it will all work simultaneously in all environments as i resorted to ing out code in sections that were already completed to reduce run time.

3 Astro

3.1 Part A

In this part I used the code given by the text and plotted the first galaxy, altering the axis to get a good visual representation. Thinking of the Hydrogen Atom, it looks like some of the transition lines match with the peaks that you can see.

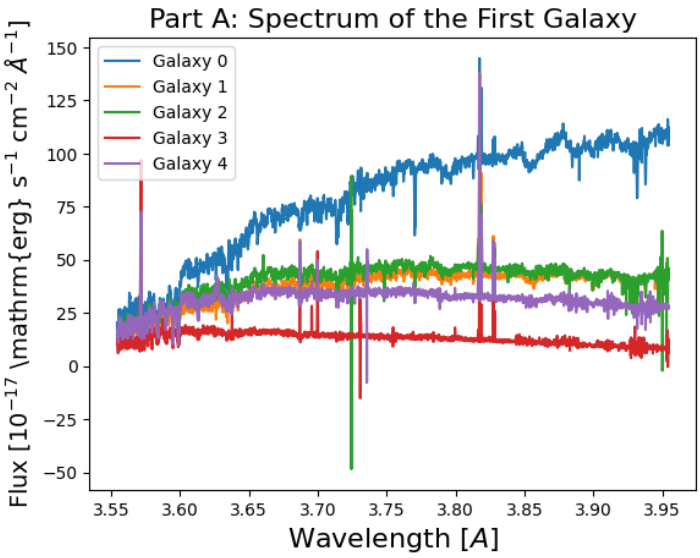

Figure 1: Galaxies

3.2 Part B

Here I normalized all the fluxes so their integrals over wavelength are the same.

3.3 Part C

Here I subtracted the mean to obtain the residuals.

3.4 Part D

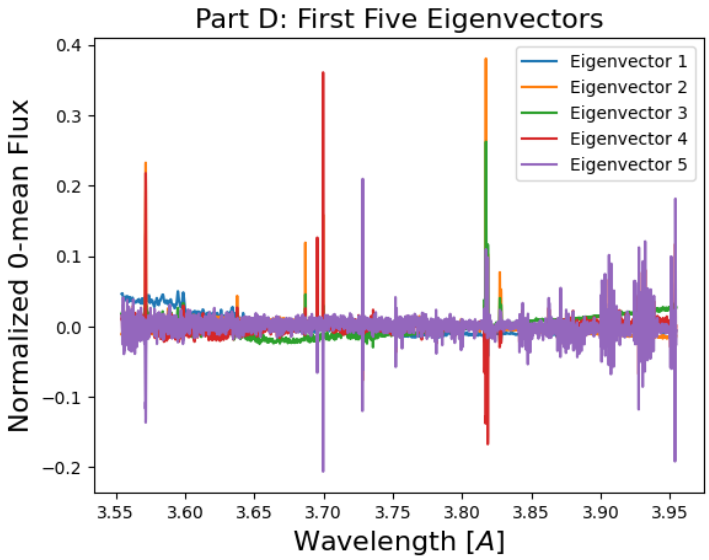

Figure 2: 5 Eignevectors

3.5 Part E

Figure 4 shows that the eigenvectors are the same for the covariance method if you were to plot them. Figure 3 shows the symmetry of the SVD eigenvalues. Both methods take over 20 minutes to compute.

3.6 Part F

Calculating the condition number shows that R is roughly $6.5 * 10^6$ whereas C is about $2.7 * 10^{10}$. The covariance matrix leads to an even bigger condition number and therefore the SVD will be much more stable.

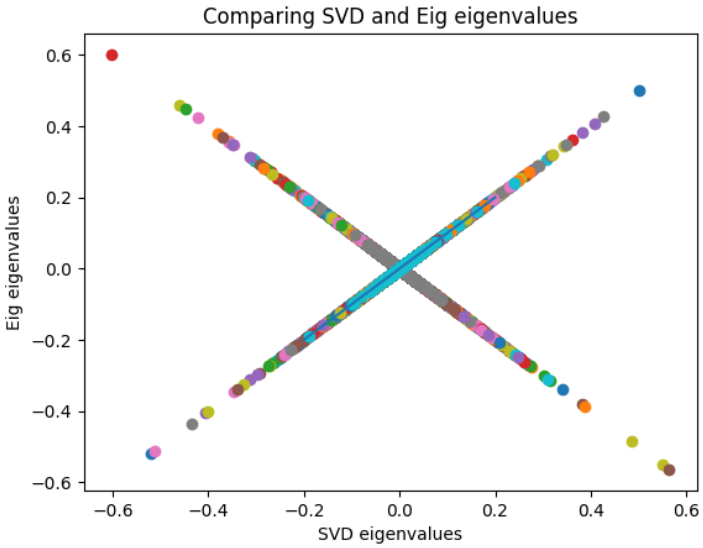

Figure 3: Matricies

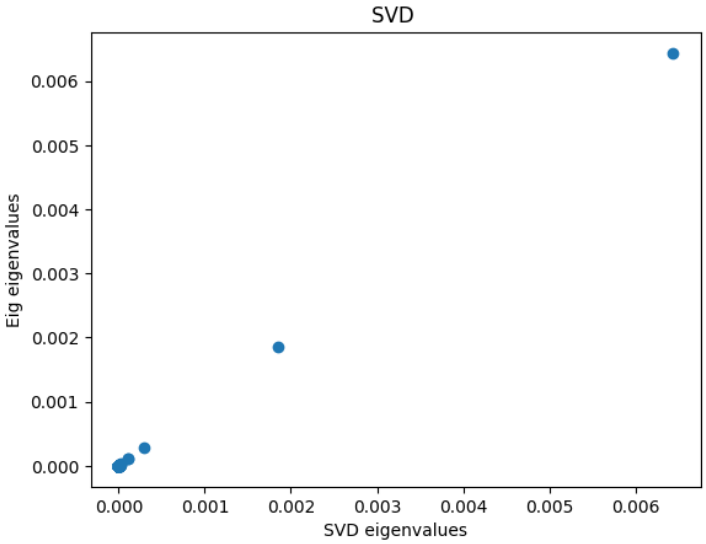

Figure 4: SVD Comparison

3.7 Part G

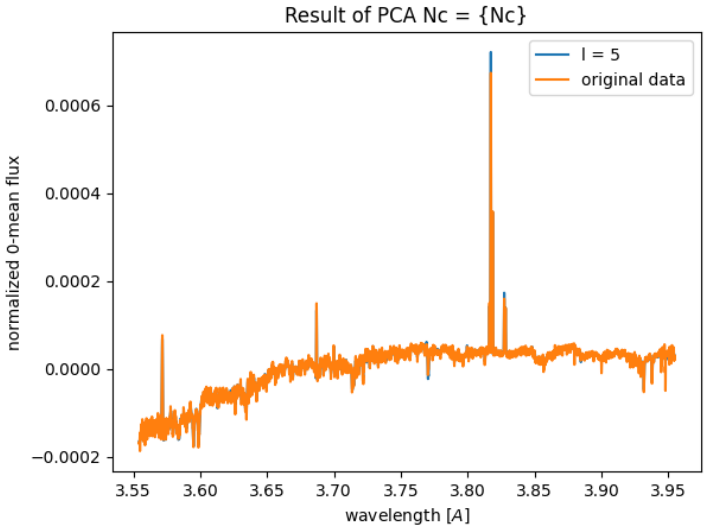

Figure 5: PCA First 5

3.8 Part H

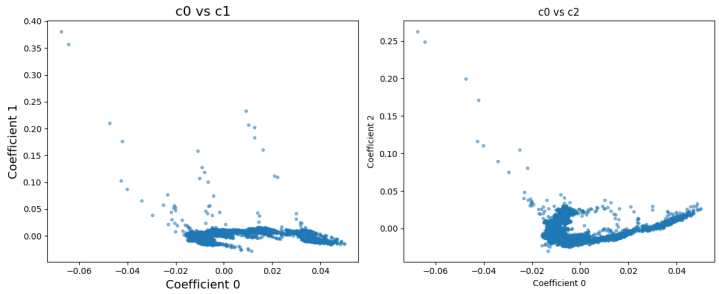

Figure 6: c0 vs c1 and c0 vs c2

3.9 Part I

Squared Fractional Error for Nc = 20: 0.031012

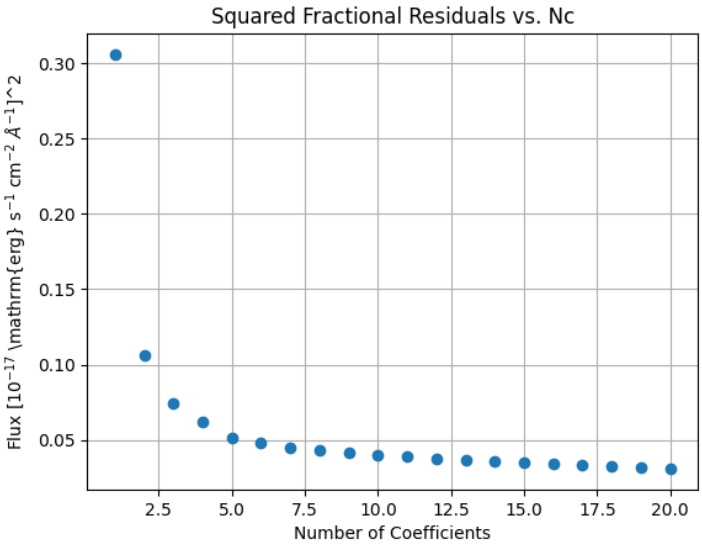

Figure 7: Nc 1 to 20