# Problem Set 6

Debendro Mookerjee
GitHub Account Name: Debendro-dm5790

October 31, 2023

**Abstract**

In this problem set we explore linear algebra and eigensystems in python by performing principal component analysis (PCA) on the radiation spectrum of several galaxies. We learn about various techniques needed to perform PCA such as preparing the data by normalizing and scaling it, determining the covariant matrix of the rescaled data, and determining the eigenvectors of the covariant matrix.

## 1 Introduction

Principal component analysis is a dimensionality reduction technique performed on a large dataset to create predictive models. The reduction is accomplished by transforming to a new coordinate system where most of the variation in the dataset can be described by fewer dimensions. The vectors pointing in the direction of the axes in this new coordinate system are called the principal components and are eigenvectors of the covariance matrix. The total number of selected principal components can be called p and is less than the total number of eigenvectors of the covariance matrix. Intuitively, we can think of PCA as fitting a p-dimensional ellipsoid to the data.

In this report we describe our methods as well as give our outputs.

## 2 Methods

First we read the spectrum data using astropy.io.fits. We define two arrays to store the data. One is called flux, which contains the spectrum values for several galaxies, and the other is called logwaves, which stores the wavelength range as the common logarithm of the wavelength in Angstroms. We plot some galaxy spectra. Plotting these spectra will reveal their similarities with the Balmer series of hydrogen.

Before we perform PCA, we normalize each spectrum and then subtract from it the mean of the normalized data. To perform the integration of the spectrum over the wavelength range, we perform a rectangular approximation. We convert the logwave data into wavelength data by exponentiating. We make a plot of these normalized data, which are also called residuals.

We then calculate the covariance matrix by multiplying the transpose of the residuals with the

residuals and dividing the resulting matrix by the number of galaxies. We use the eigh() function in numpy's linear algebra module to determine the eigenvectors and we sort the eigenvectors in descending order. The eig() function did not work; it always computed the eigenvectors to be arrays of zeros. We choose the five eigenvectors corresponding to the five largest eigenvectors. We plot these eigenvectors and see that each of them has a funnel like shape.

We then perform singular value decomposition (SVD) on the residuals to determine the 5 eigenvectors and compare with the ones found with the eigh() function. We also compared the execution times by using Python's timeit module. The computer had trouble performing SVD on the large array of residuals and outputted an array of zeros for each eigenvector and so we had to truncate the data to perform SVD and diagonalization via eigh().

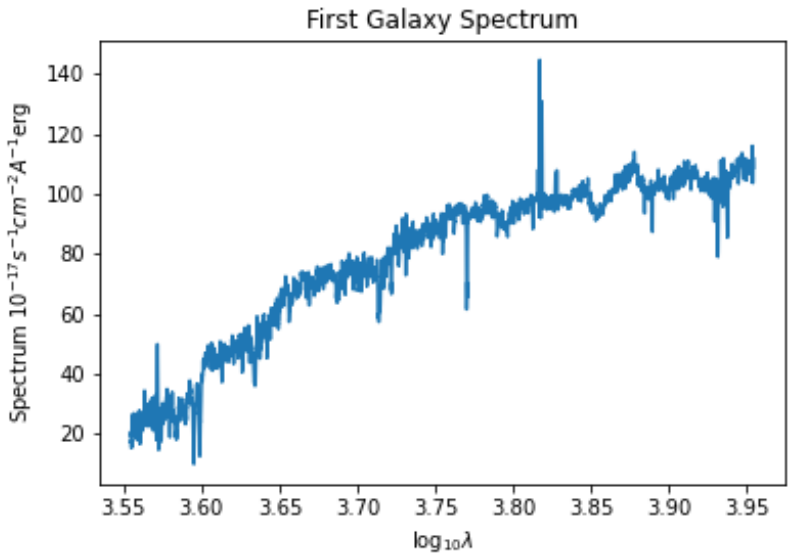
We used the five principle eigenvectors of the covariance matrix of the original data to compute the approximate residuals. We also determine the coefficients in front of the principal components. We make plots of some of these coefficients.

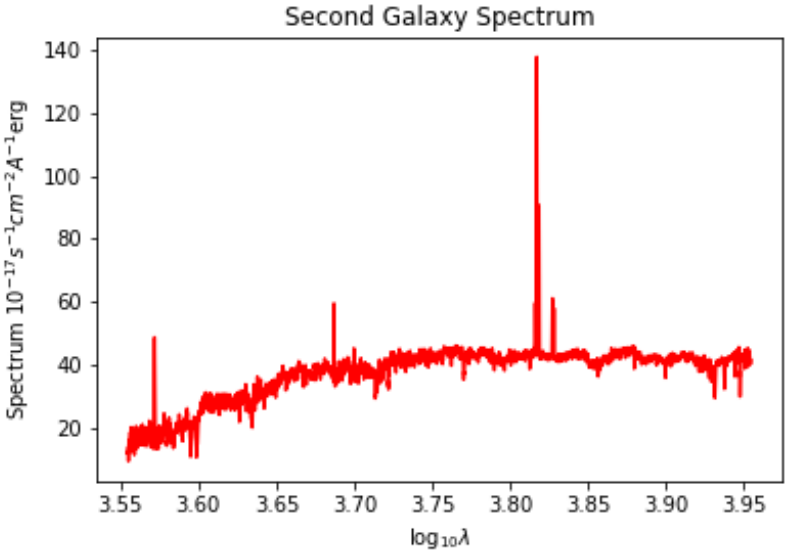
We repeat the above procedure for N principal components with N varying from 1 to 20. We also make plots of the fractional residuals.

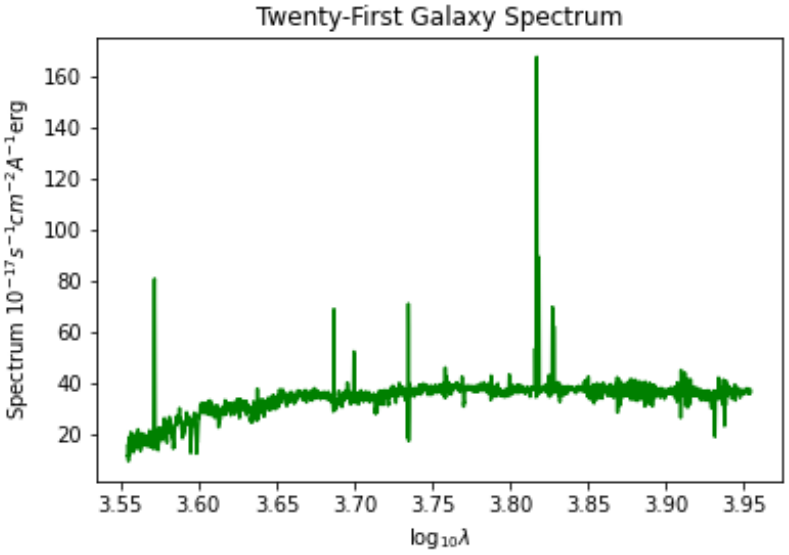
# 3 Results

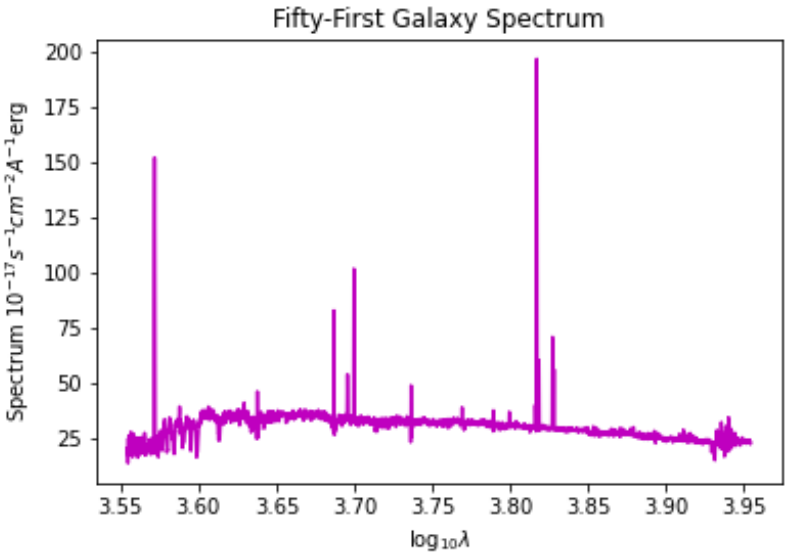
## 3.1 Part A:

Here are the plots of the wavelength spectra of a handful of galaxies.




Second Galaxy Spectrum
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$
140
120
100
80
60
40
20
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95


Twenty-First Galaxy Spectrum
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$
160
140
120
100
80
60
40
20
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95

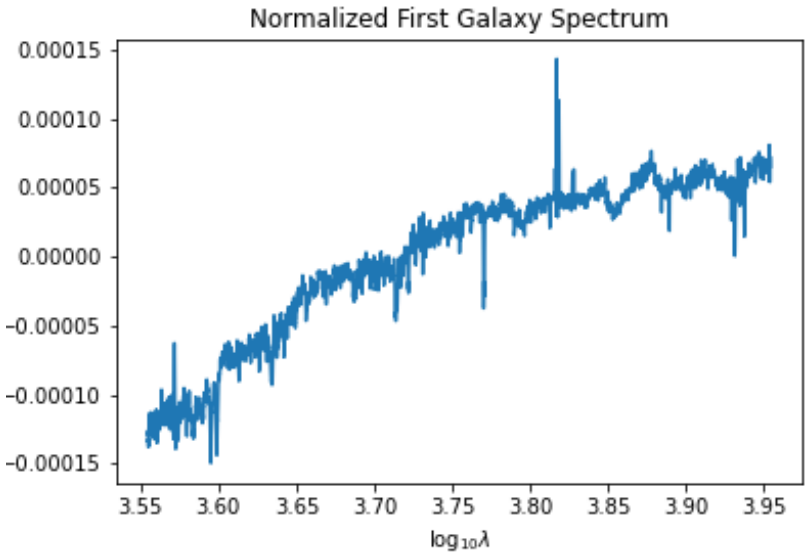


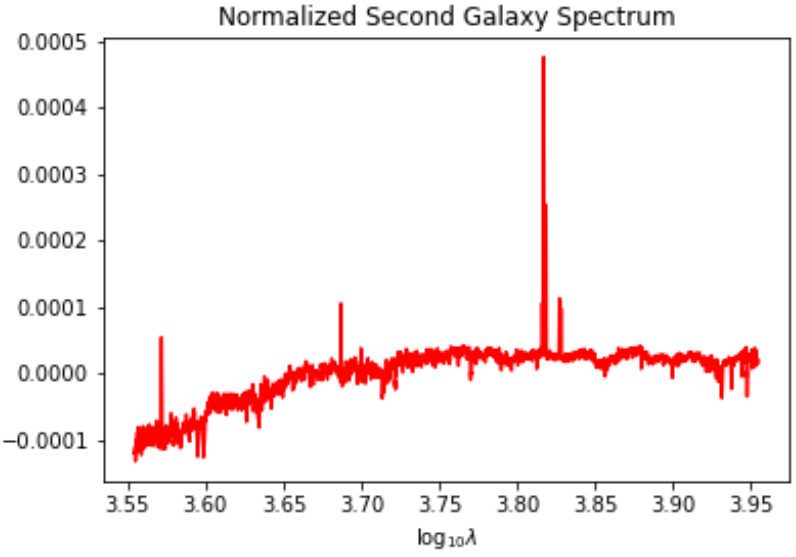

Spectra plots for a handful of galaxies. Wavelengths are in Angstroms, but the units of the x-axis seem weird to me(log of an Angstrom?).

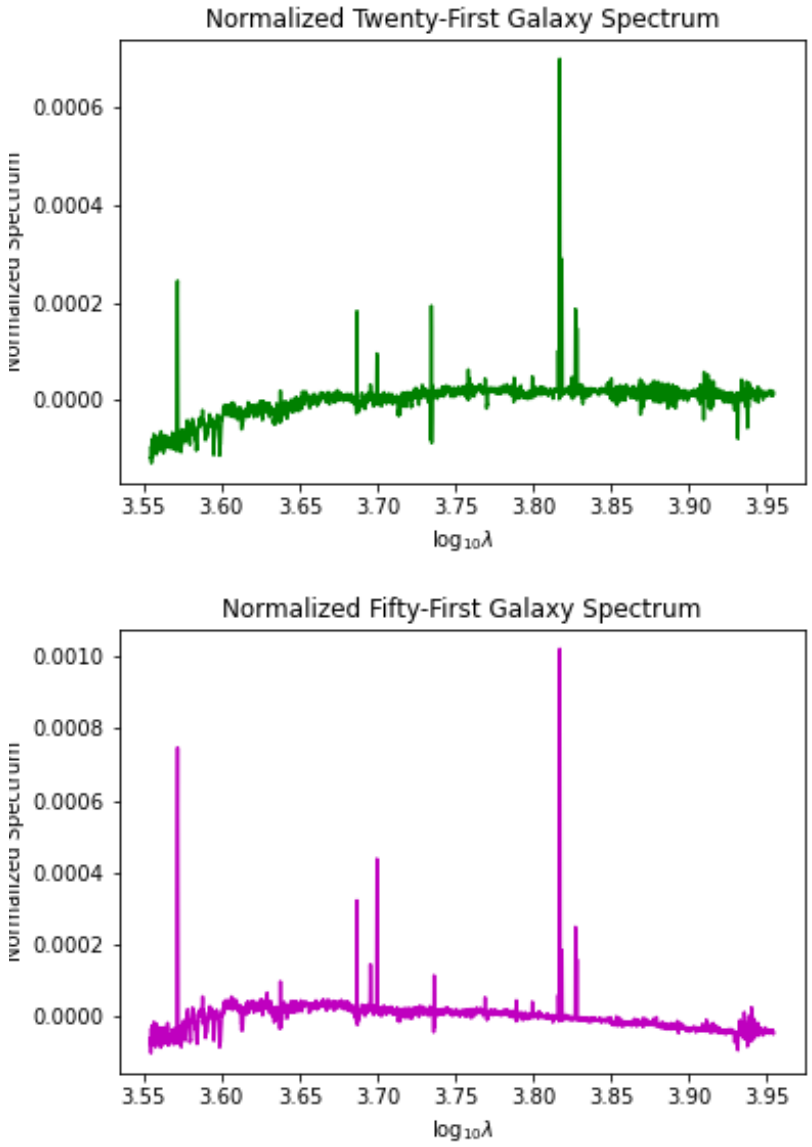
We see that these plots have peaks around $\log_{10}(\lambda) = 3.55$ and 3.80, where $\lambda$ is measured in Angstroms. The corresponding wavelengths in nanometers is about $10^{3.55}/10 \approx 354.8$ nm and $10^{3.8}/10 \approx 630.96$ nm. **These wavelengths are comparable to the Balmer transitions in hydrogen**, as shown in https://en.wikipedia.org/wiki/Balmer_series.

### 3.2 Parts B and C:

Here are the plots where the spectra have been normalized and the mean of the normalized spectra have been subtracted away. We computed the normalization constant by using the rectangular approximation of an integral. During the approximation process, we converted the independent variable into wavelengths, measured in Angstroms.


Normalized First Galaxy Spectrum
0.00015
0.00010
0.00005
0.00000
−0.00005
−0.00010
−0.00015
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

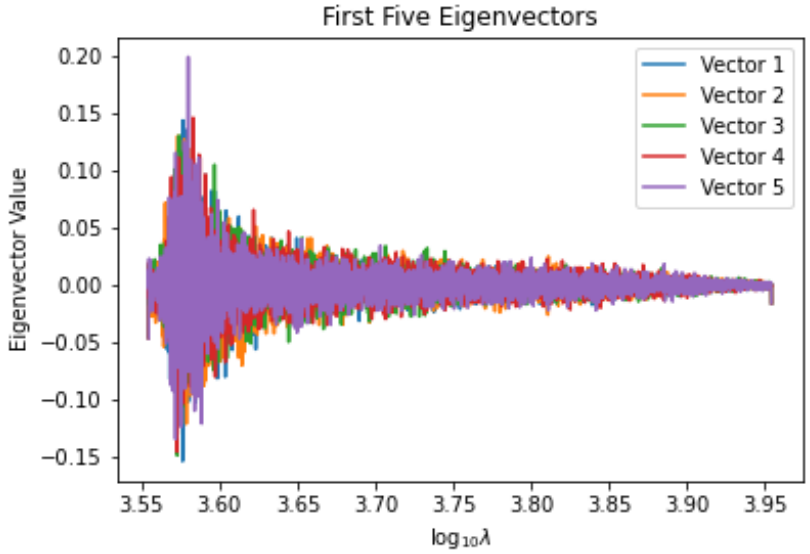

Normalized Second Galaxy Spectrum
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
−0.0001
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ




Normalized Spectra plots for a handful of galaxies. Wavelengths are in Angstroms.

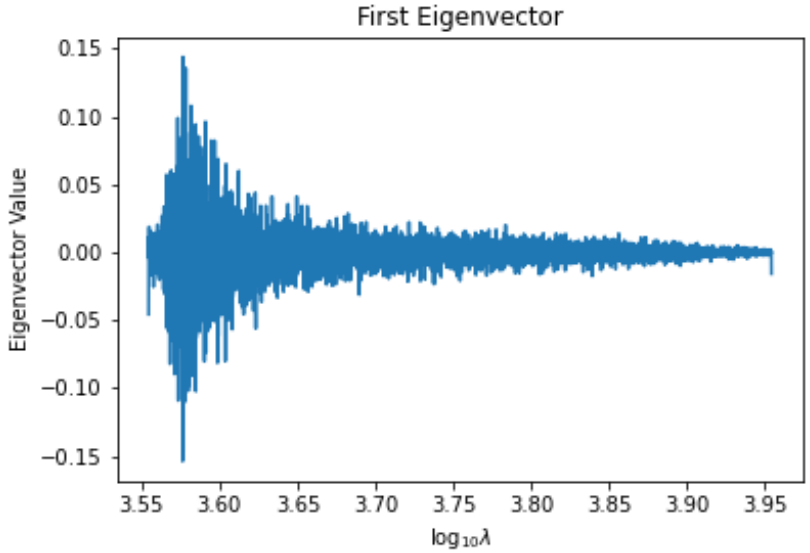
### 3.3 Part D:

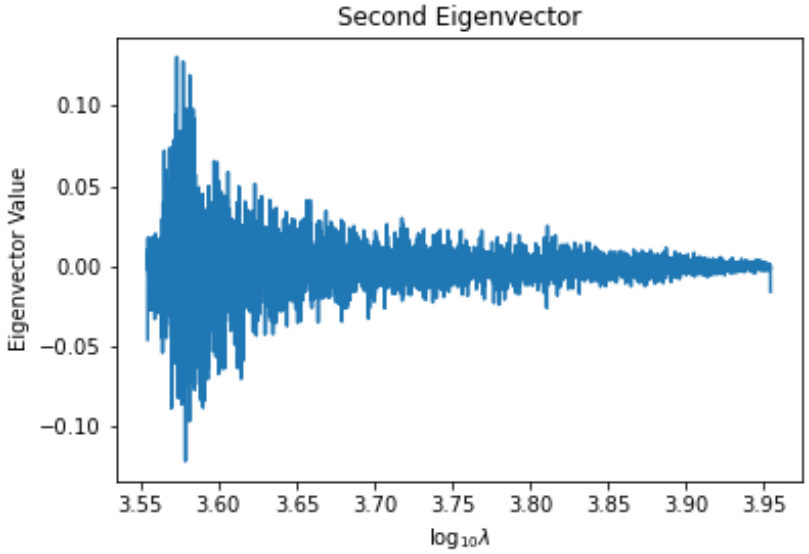
Here is a single plot of the first five eigenvectors of the covariance matrix. These corresponding eigenvalues are the five largest eigenvalues of the covariance matrix. I used eigh() since eig() always gave me eigenvectors of zeros.

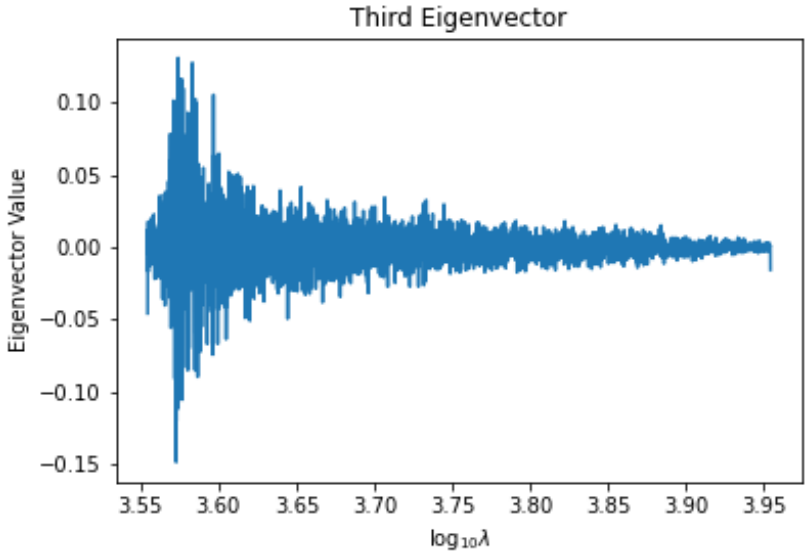


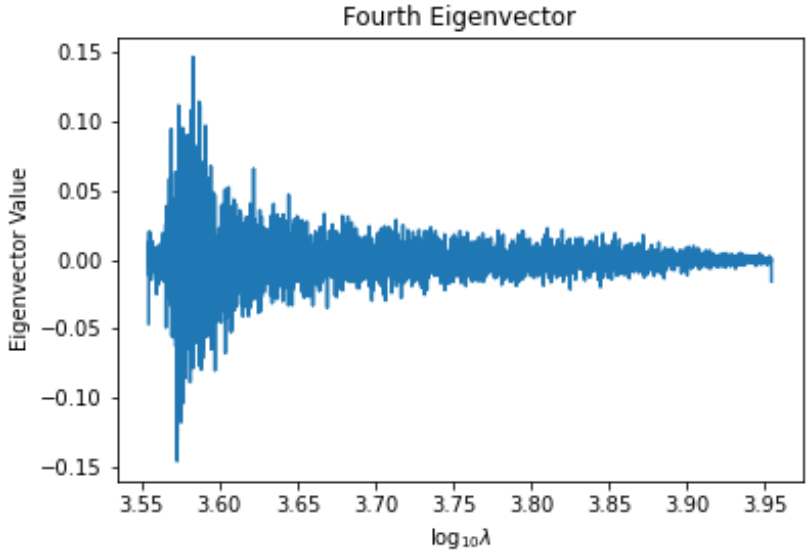

Wavelengths are in Angstroms.

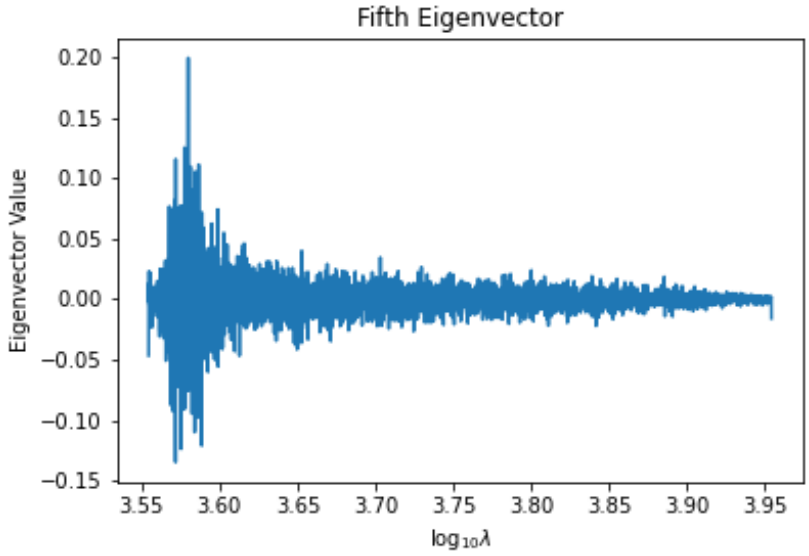
Here are separate plots.




Second Eigenvector
Eigenvector Value
0.10
0.05
0.00
−0.05
−0.10
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Third Eigenvector
Eigenvector Value
0.10
0.05
0.00
−0.05
−0.10
−0.15
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ




Wavelengths are in Angstroms.

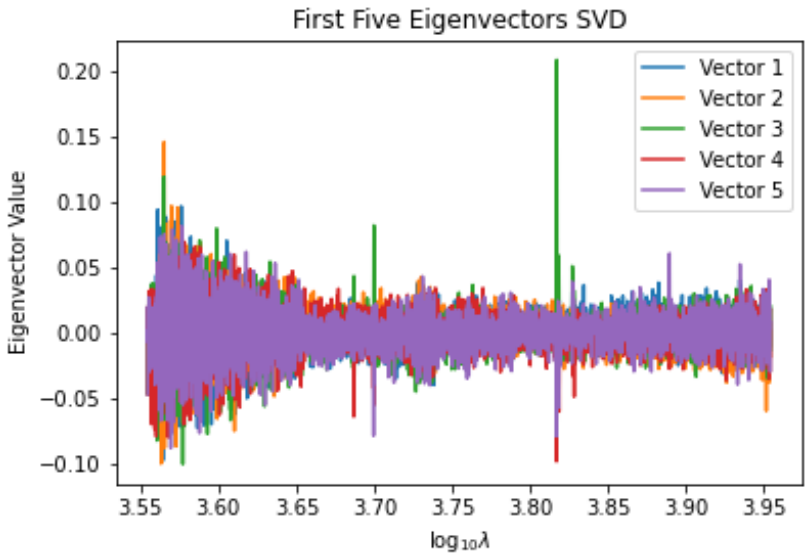
### 3.4 Part E:

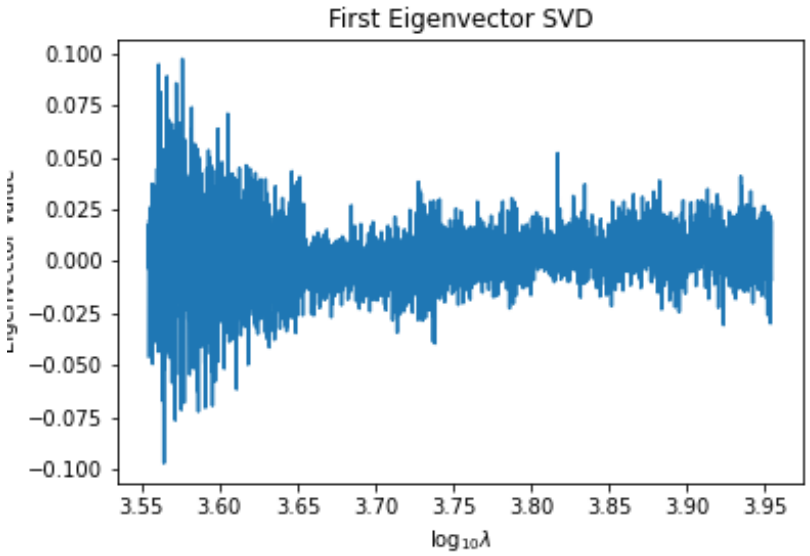
My computer could not compute the SVD of the scaled data because the data was too large. See images below. As a result, I had to truncate the data to include the first 1000 galaxies. Furthermore, I had to click the run button several times before I got decent agreement between the five eigenvectors obtained via SVD and the five obtained from eigh(). By decent I mean not pathetic.

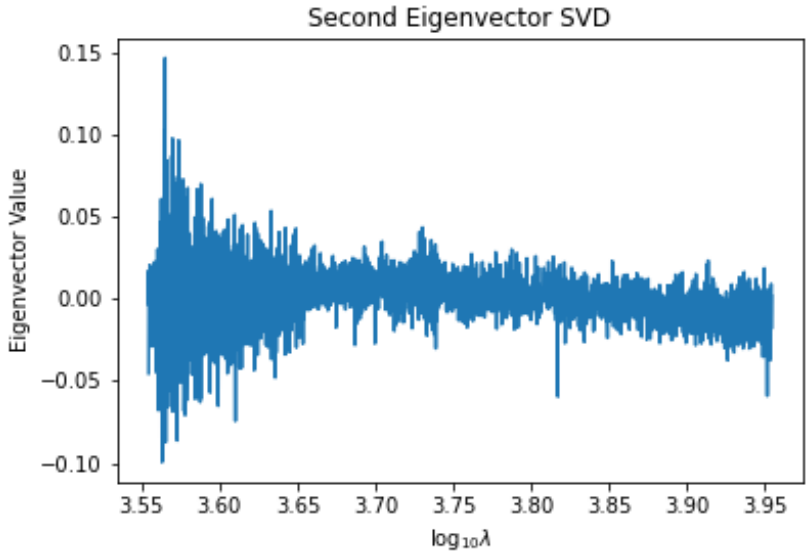
```
In [6]: runfile('C:/Users/mooke/aGradComp/ps-6/PCAPartE.py', wdir='C:/Users/mooke/aGradComp/ps-6')
Reloaded modules: PCA
Traceback (most recent call last):

  File "C:\Users\mooke\aGradComp\ps-6\PCAPartE.py", line 9, in <module>
    e1, e2, e3, e4, e5 = PCA.EigenVectorsThroughSVD(scaledFlux)

  File "C:\Users\mooke\aGradComp\ps-6\PCA.py", line 82, in EigenVectorsThroughSVD
    (U,D,VT) = np.linalg.svd(scaledFlux, full_matrices = False)

  File "<__array_function__ internals>", line 5, in svd

  File "C:\Users\mooke\anaconda3\lib\site-packages\numpy\linalg\linalg.py", line 1626, in svd
    u, s, vh = gufunc(a, signature=signature, extobj=extobj)

MemoryError: Unable to allocate 296. MiB for an array with shape (9713, 4001) and data type float64
```

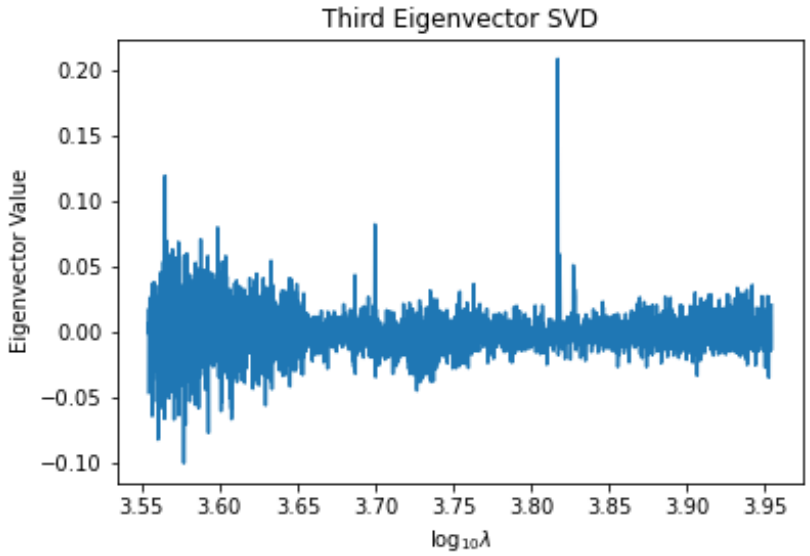
```
In [9]: runfile('C:/Users/mooke/aGradComp/ps-6/PCAPartE.py', wdir='C:/Users/mooke/
aGradComp/ps-6')
Reloaded modules: PCA
Traceback (most recent call last):

  File "C:\Users\mooke\aGradComp\ps-6\PCAPartE.py", line 9, in <module>
    e1, e2, e3, e4, e5 = PCA.EigenVectorsThroughSVD(scaledFlux)

  File "C:\Users\mooke\aGradComp\ps-6\PCA.py", line 82, in EigenVectorsThroughSVD
    (U,D,VT) = np.linalg.svd(scaledFlux, full_matrices = False)

  File "<__array_function__ internals>", line 5, in svd

  File "C:\Users\mooke\anaconda3\lib\site-packages\numpy\linalg\linalg.py", line 1626, in
svd
    u, s, vh = gufunc(a, signature=signature, extobj=extobj)

MemoryError: Unable to allocate 222. MiB for an array with shape (9713, 3000) and data
type float64
```

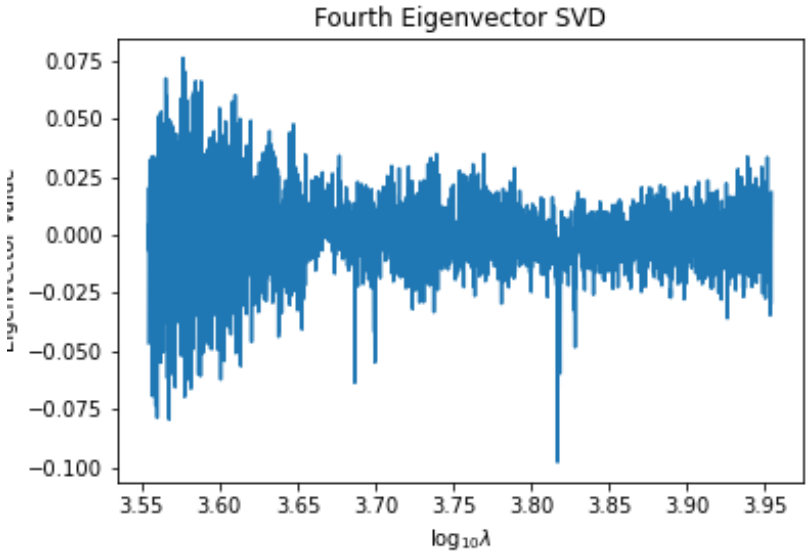
Here are the eigenvectors obtained by performing SVD.

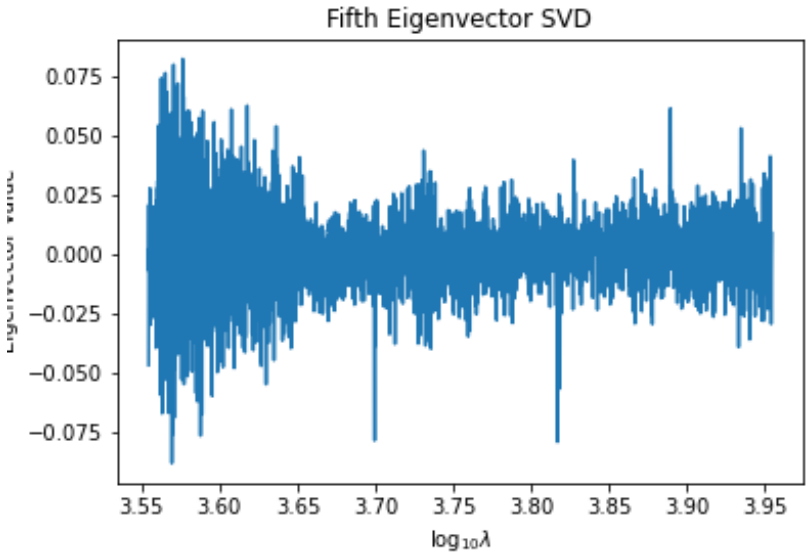



First Eigenvector SVD
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−0.100
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Second Eigenvector SVD
Eigenvector Value
0.15
0.10
0.05
0.00
−0.05
−0.10
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

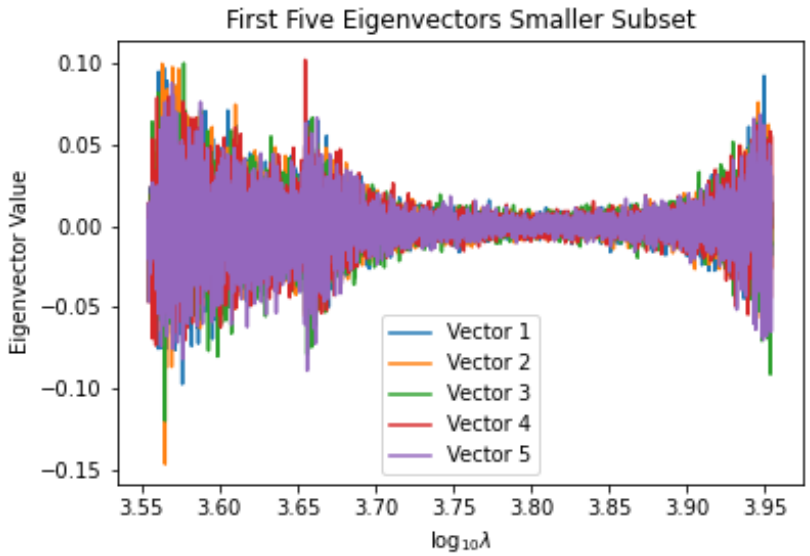

Third Eigenvector SVD
Eigenvector Value
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

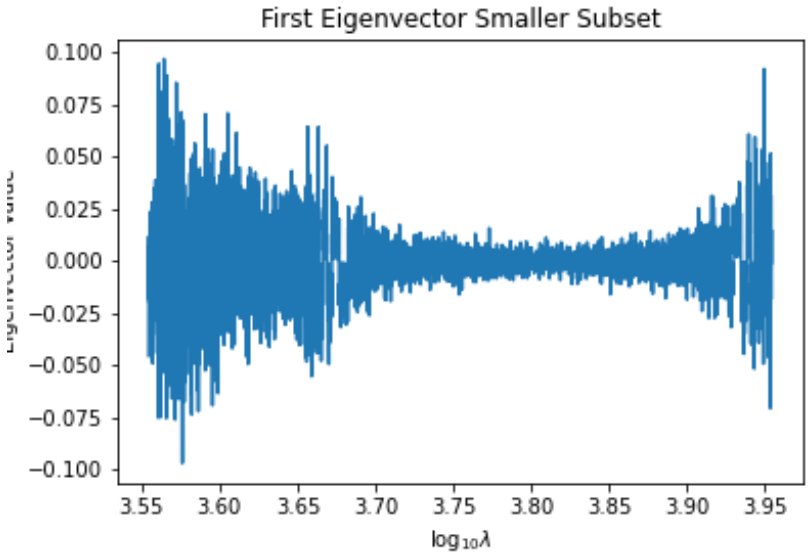

Fourth Eigenvector SVD
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−0.100
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

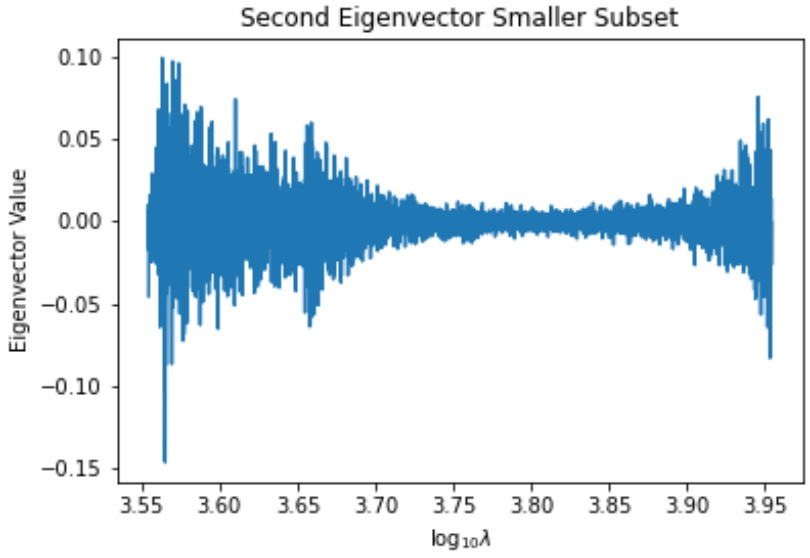


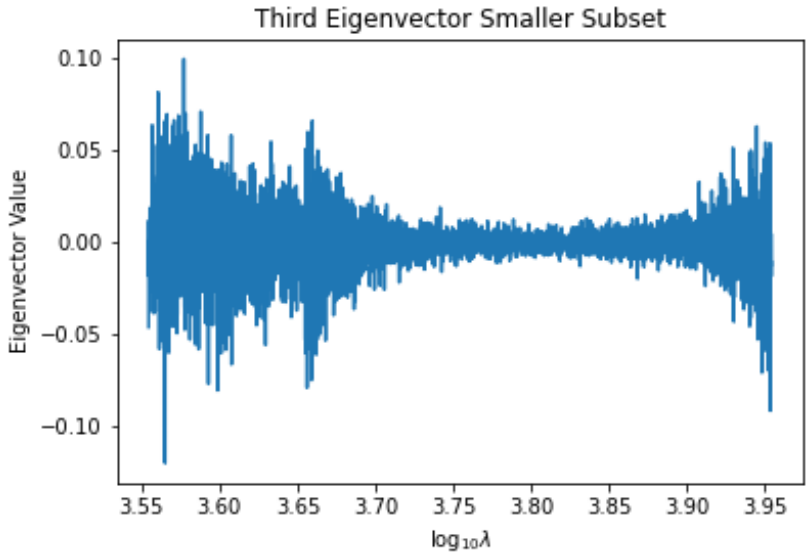

Wavelengths are in Angstroms.

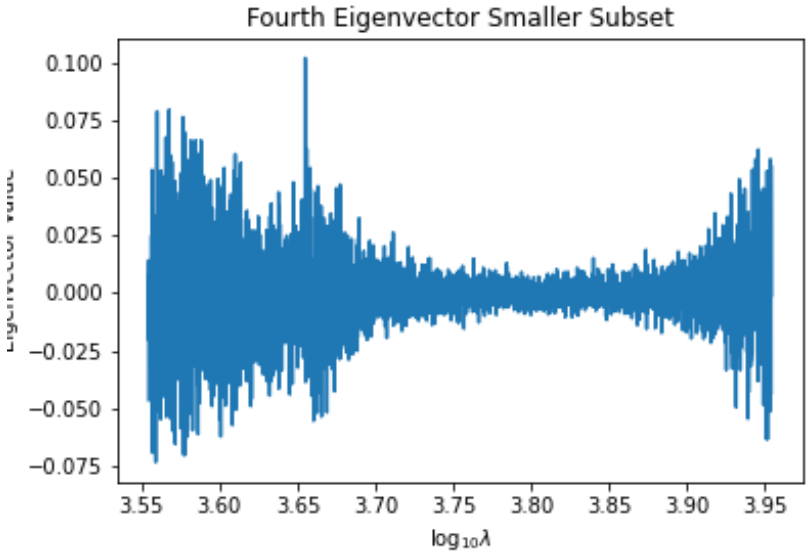
Here are the eigenvectors obtained by diagonalizing the covariance matrix.

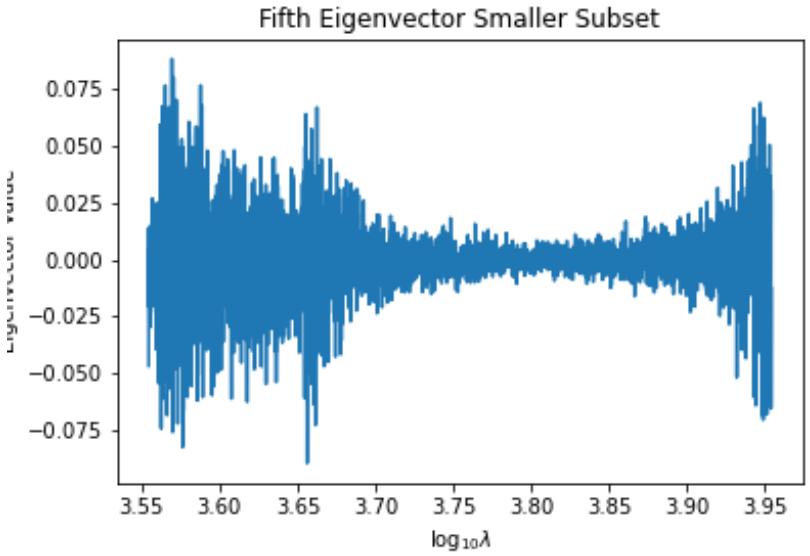



First Eigenvector Smaller Subset
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
−0.100
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Second Eigenvector Smaller Subset
Eigenvector Value
0.10
0.05
0.00
−0.05
−0.10
−0.15
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Third Eigenvector Smaller Subset
Eigenvector Value
0.10
0.05
0.00
−0.05
−0.10
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Fourth Eigenvector Smaller Subset
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ




Wavelengths are in Angstroms.

Below are the execution times for the SVD and diagnolization methods. We see that SVD is much slower than the eigh() function.

```
In [1]: runfile('C:/Users/mooke/aGradComp/ps-6/untitled1.py', wdir='C:/Users/mooke/
aGradComp/ps-6')
Time it took to compute eigenvectors with numpy.linalg.svd is 4.424254300000001
seconds.
Time it took to compute eigenvectors with numpy.linalg.eigh is 0.516422900000002
seconds.
```

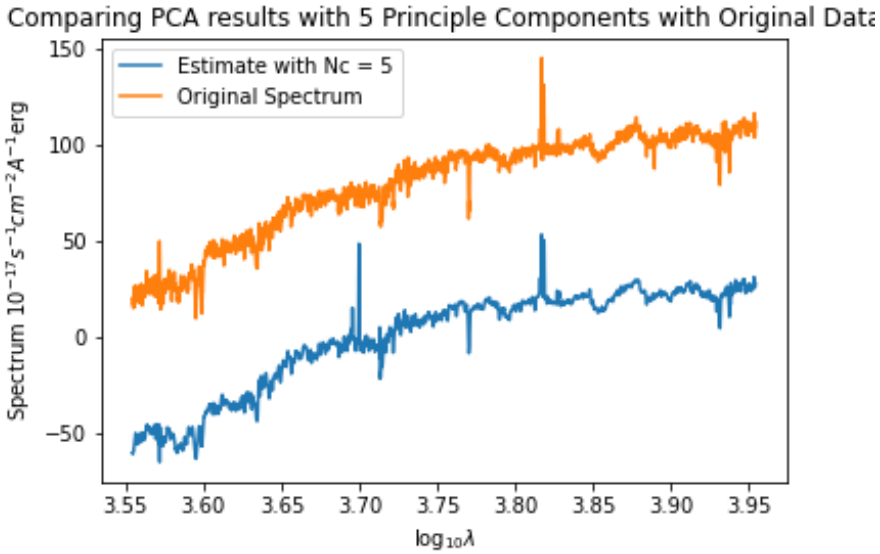
## 3.5 Part F:

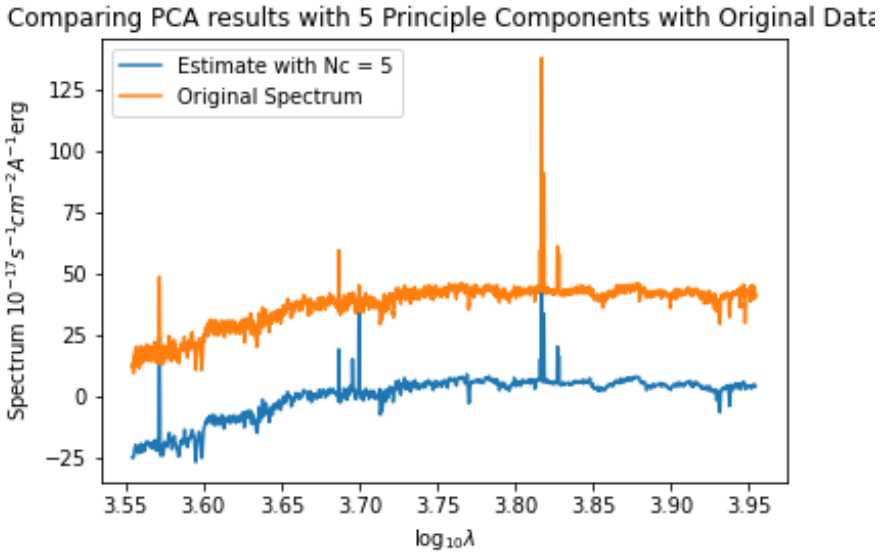
There are several disadvantages of SVD. For me, it is very unstable since it took me several clicks of the run button to get decent eigenvectors. My laptop cannot use it for very large matrices. SVD also takes longer; **however the one advantage is that it has a smaller condition number, which is the ratio of the largest singular value to the smallest one. It is, therefore, more stable to perturbations than the eigh() function.** Here are the conditioned numbers for these two methods.

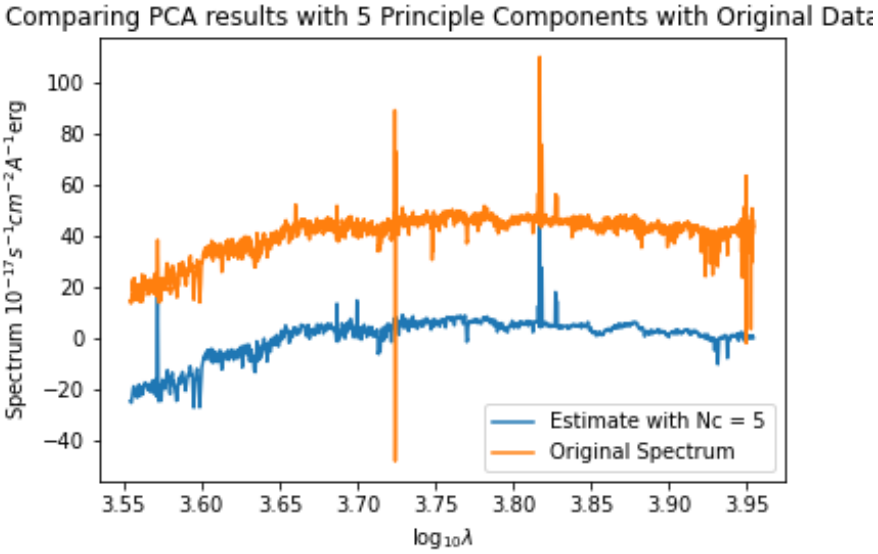
```
aGradComp/ps-6')
Reloaded modules: PCA
Condition number for R is
1280.3666
Condition number for the Covariance matrix is
nan
```

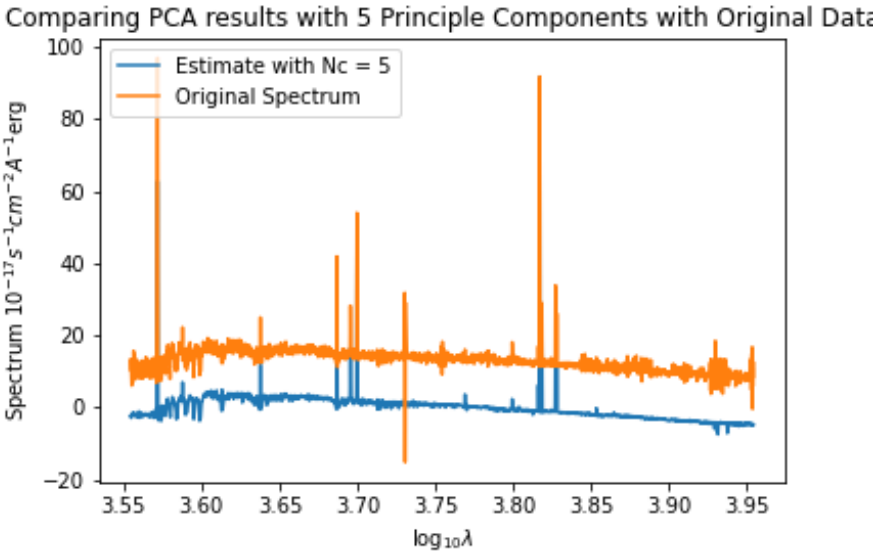
### 3.6 Part G

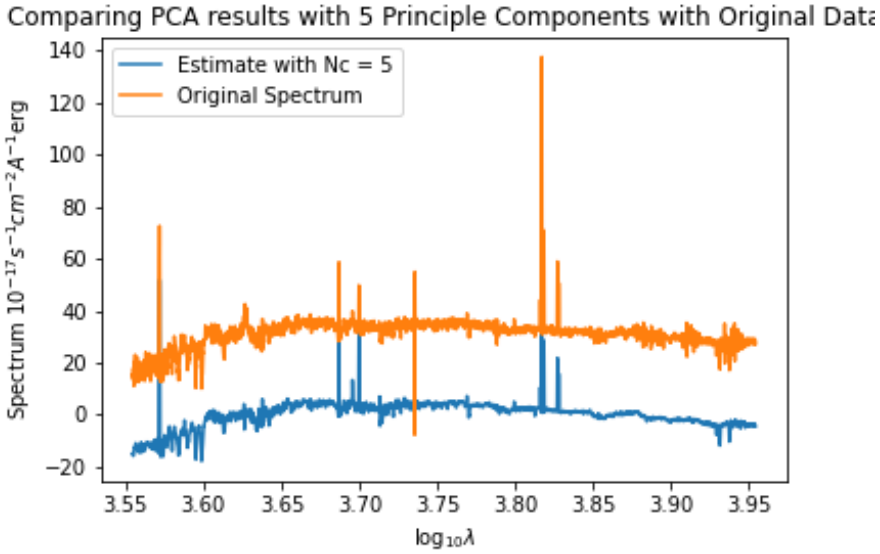
Here are the PCA approximations to the actual original spectra, not the normalized and re-scaled original spectra. These each use five principal components.

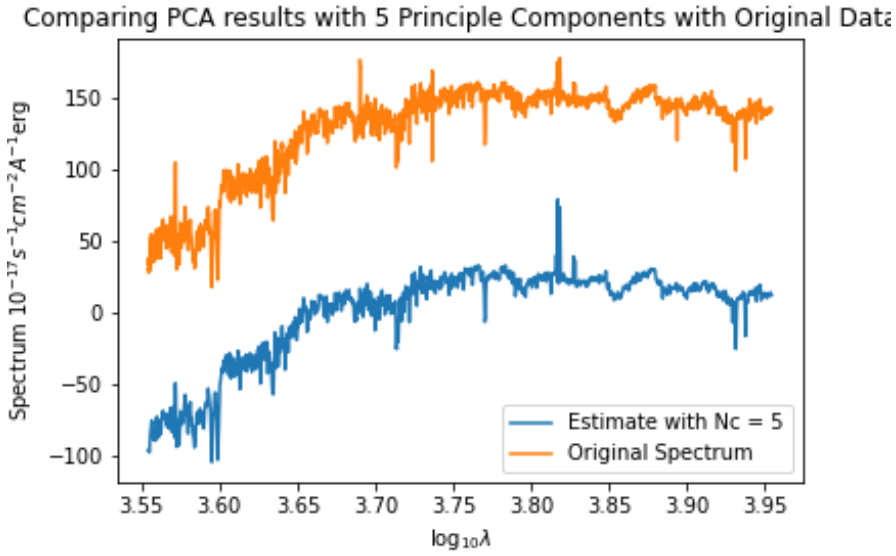


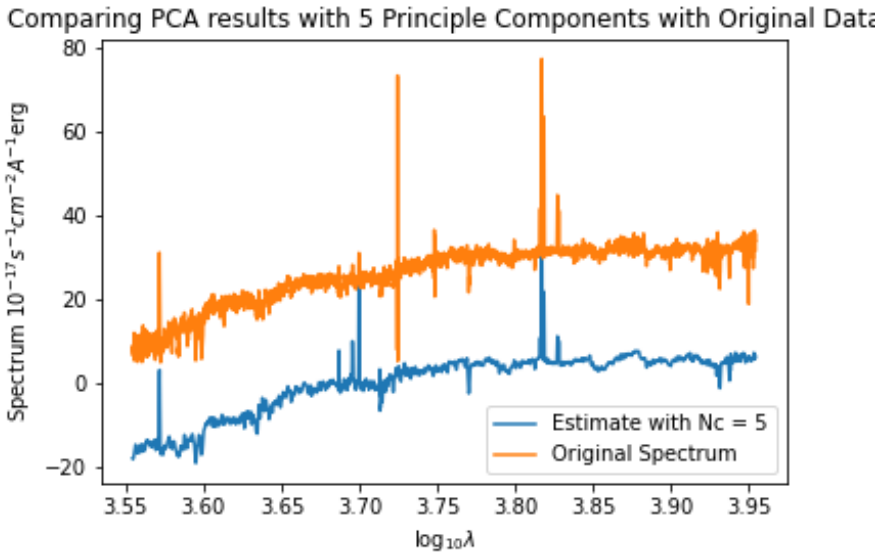




Comparing PCA results with 5 Principle Components with Original Data
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$
Estimate with Nc = 5
Original Spectrum


Comparing PCA results with 5 Principle Components with Original Data
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$
Estimate with Nc = 5
Original Spectrum

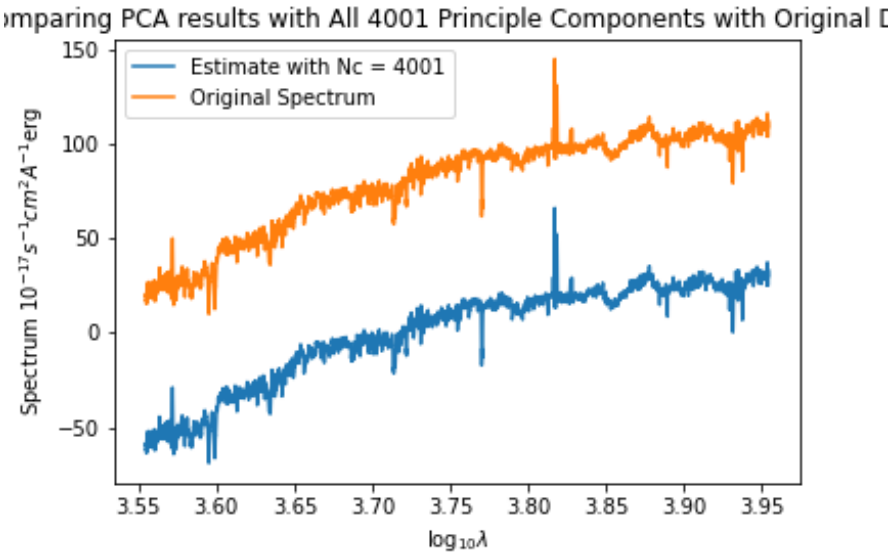

Comparing PCA results with 5 Principle Components with Original Data
Estimate with Nc = 5
Original Spectrum
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$

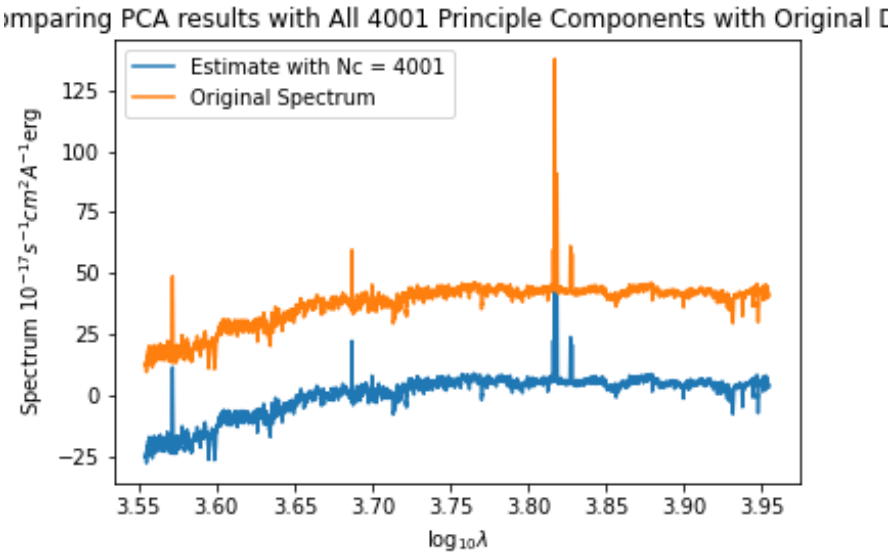

Comparing PCA results with 5 Principle Components with Original Data
Estimate with Nc = 5
Original Spectrum
Spectrum $10^{-17} s^{-1} cm^{-2} A^{-1} erg$
$\log_{10}\lambda$

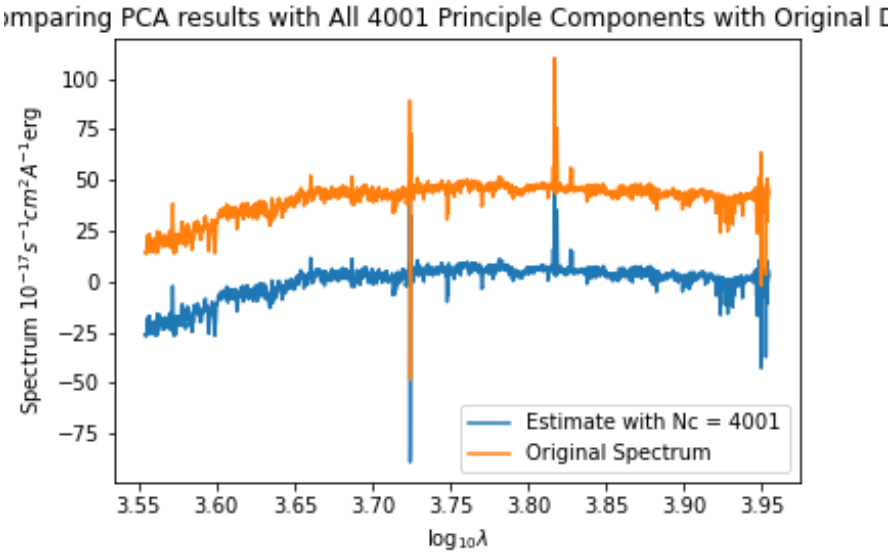


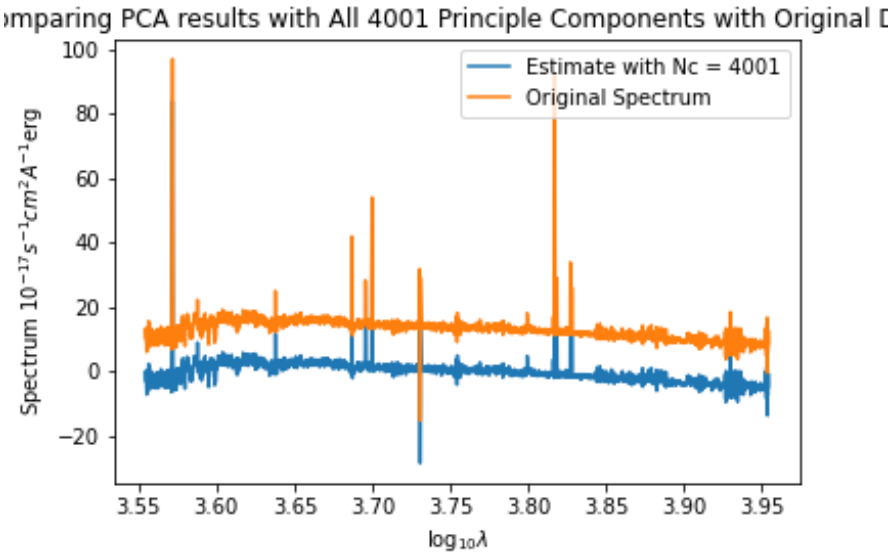

Wavelengths are in Angstroms.

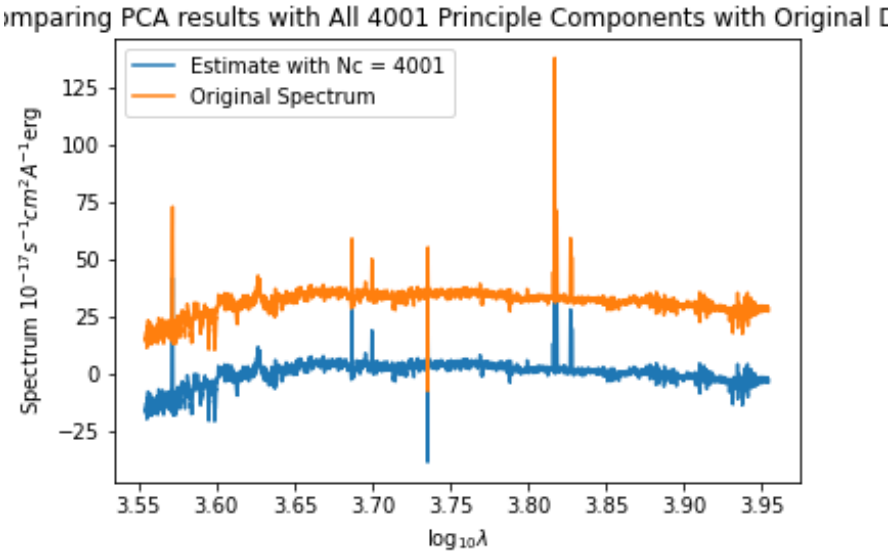
Here are the plots when all 4001 principal components are used.

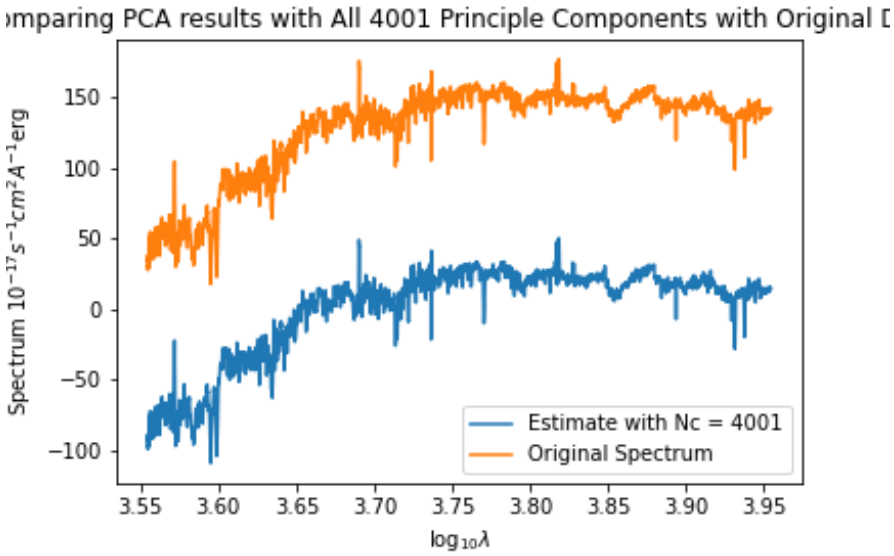


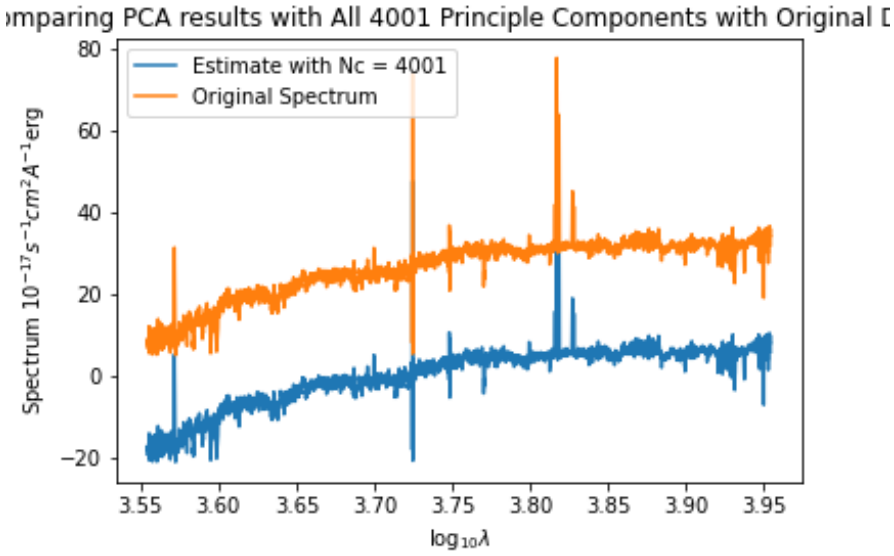

mparing PCA results with All 4001 Principle Components with Original D
Estimate with Nc = 4001
Original Spectrum
Spectrum $10^{-17} s^{-1} cm^2 A^{-1} erg$
125
100
75
50
25
0
−25
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
$\log_{10}\lambda$


mparing PCA results with All 4001 Principle Components with Original D
Spectrum $10^{-17} s^{-1} cm^2 A^{-1} erg$
100
75
50
25
0
−25
−50
−75
Estimate with Nc = 4001
Original Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
$\log_{10}\lambda$


mparing PCA results with All 4001 Principle Components with Original D
Estimate with Nc = 4001
Original Spectrum
Spectrum $10^{-17} s^{-1} cm^2 A^{-1} erg$
$\log_{10}\lambda$


mparing PCA results with All 4001 Principle Components with Original D
Estimate with Nc = 4001
Original Spectrum
Spectrum $10^{-17} s^{-1} cm^2 A^{-1} erg$
$\log_{10}\lambda$

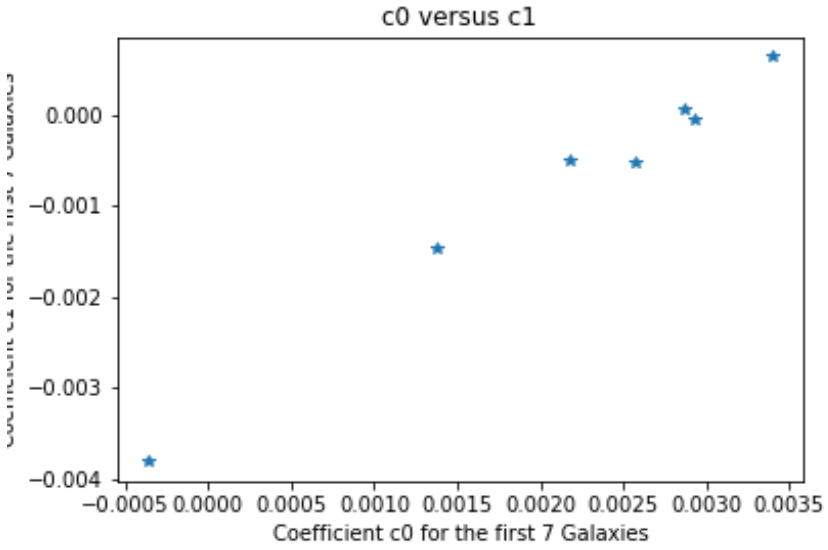


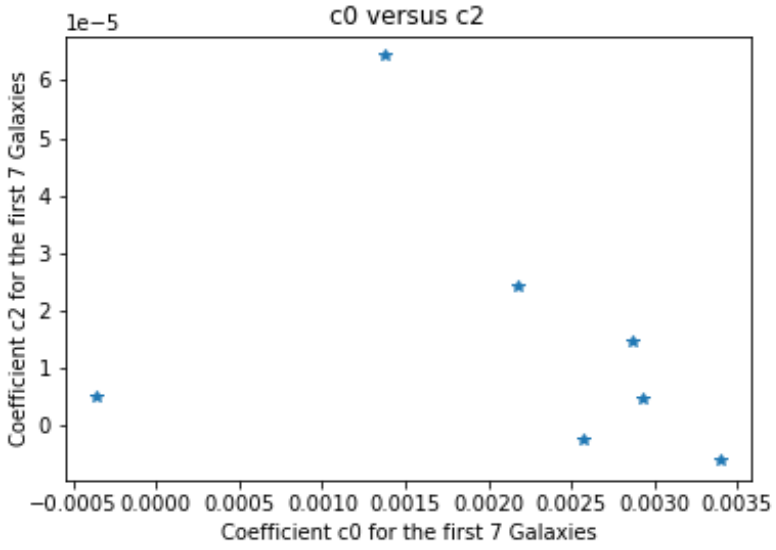




Wavelengths are in Angstroms.

## 3.7 Part H:

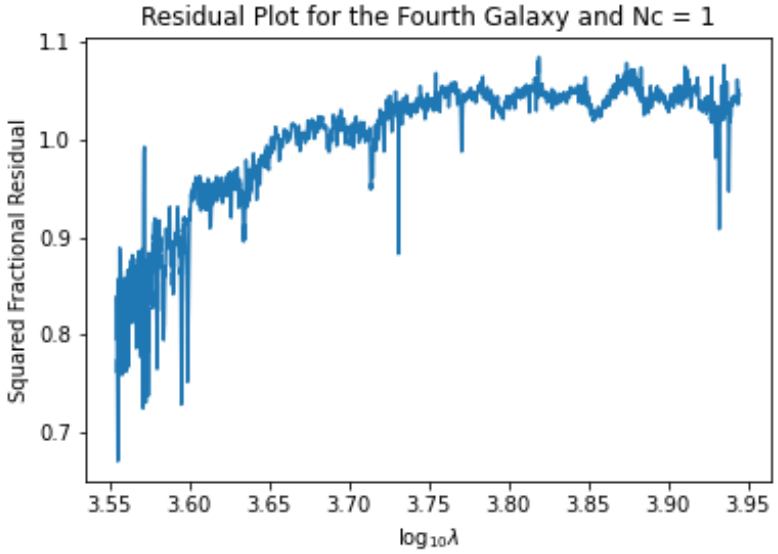
Here are the plots of $c_0$ versus $c_1$ and $c_0$ versus $c_2$

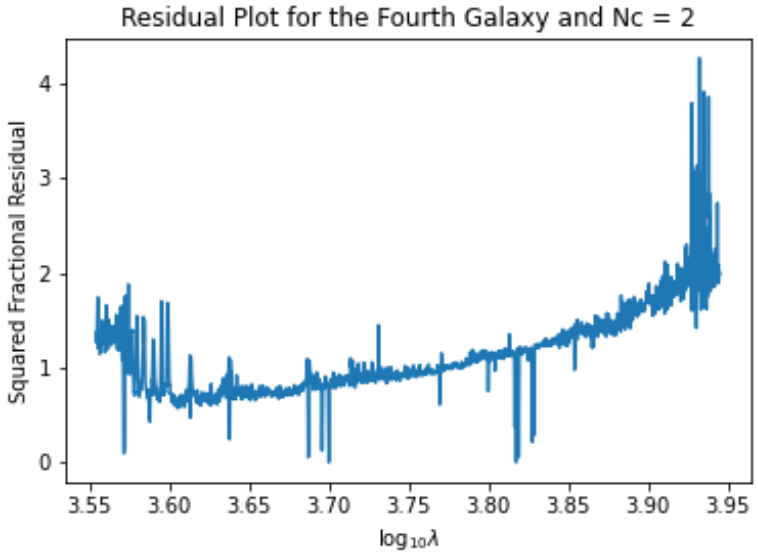


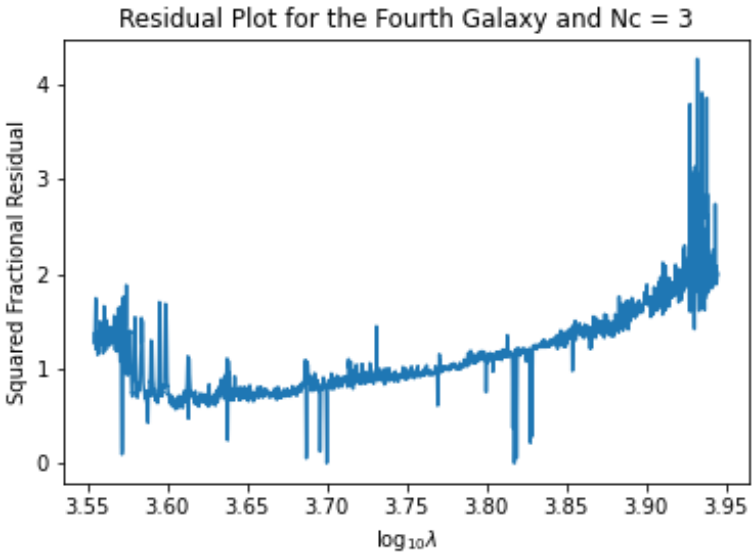



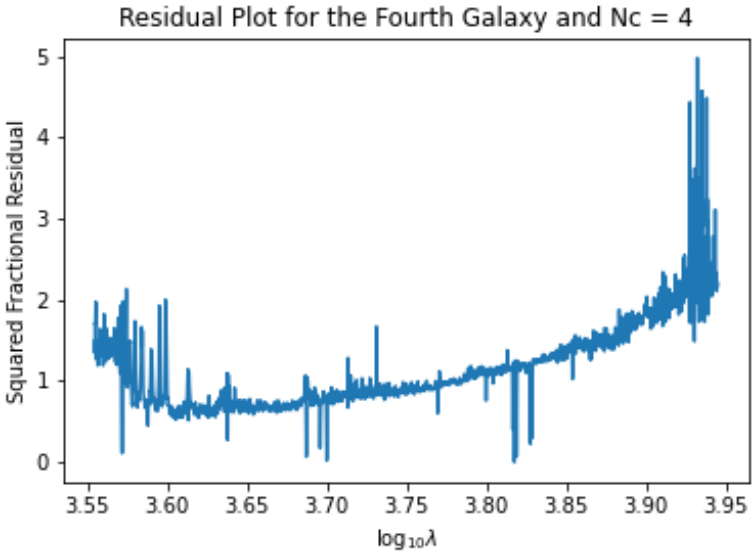

### 3.8 Part I:

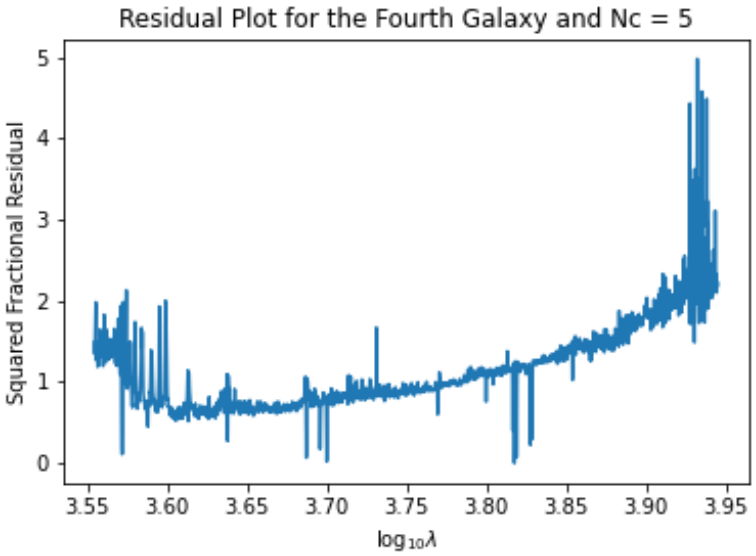
Finally, here are the plots of the squared fractional residuals as the number of principle components varied from 1 to 20. I noticed a huge residual at the end of each plot which made the remaining residuals look small. I removed that large residual from the each plot. Here are these plots without this large residual. The fractional residuals squared have a value of about 5 when 20 components were used.

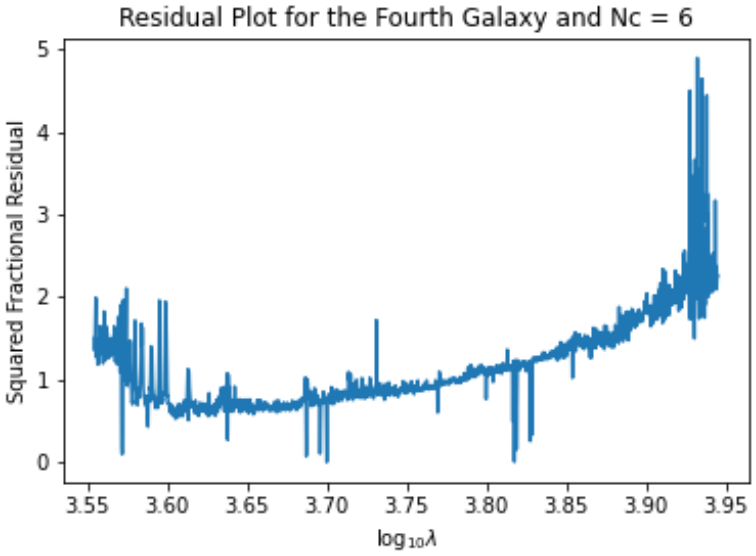

Residual Plot for the Fourth Galaxy and Nc = 1
Squared Fractional Residual
1.1
1.0
0.9
0.8
0.7
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

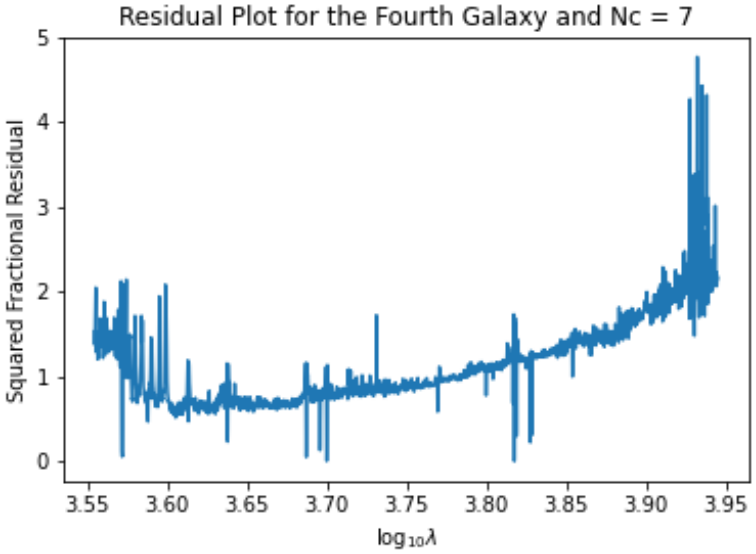

Residual Plot for the Fourth Galaxy and Nc = 2
Squared Fractional Residual
4
3
2
1
0
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

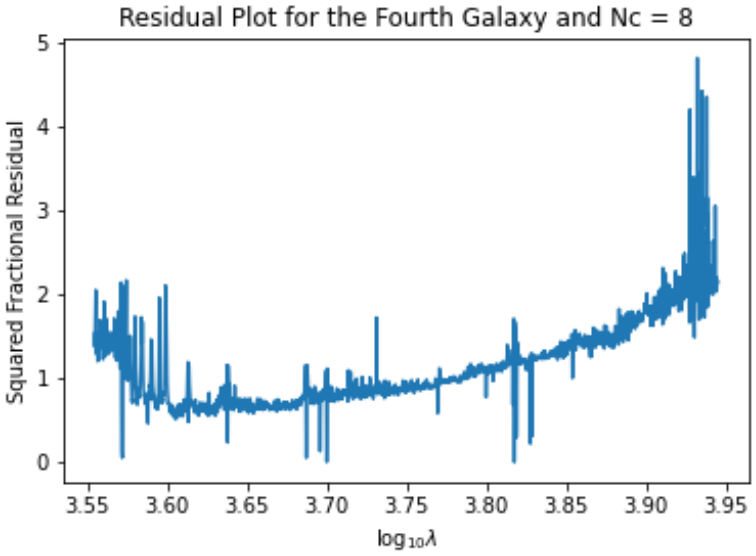

Residual Plot for the Fourth Galaxy and Nc = 3
Squared Fractional Residual
0
1
2
3
4
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

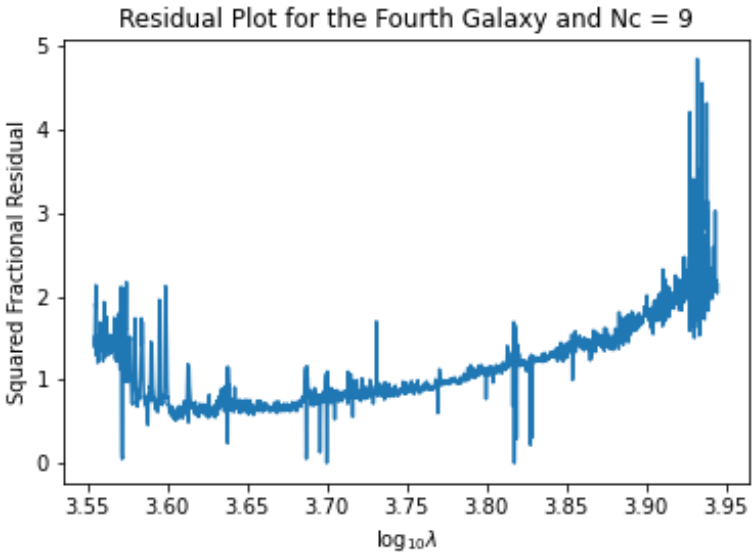

Residual Plot for the Fourth Galaxy and Nc = 4
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

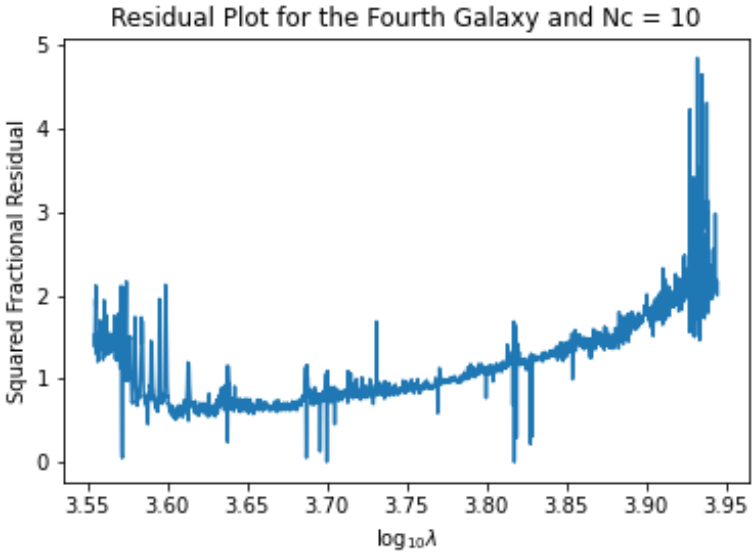

Residual Plot for the Fourth Galaxy and Nc = 5
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

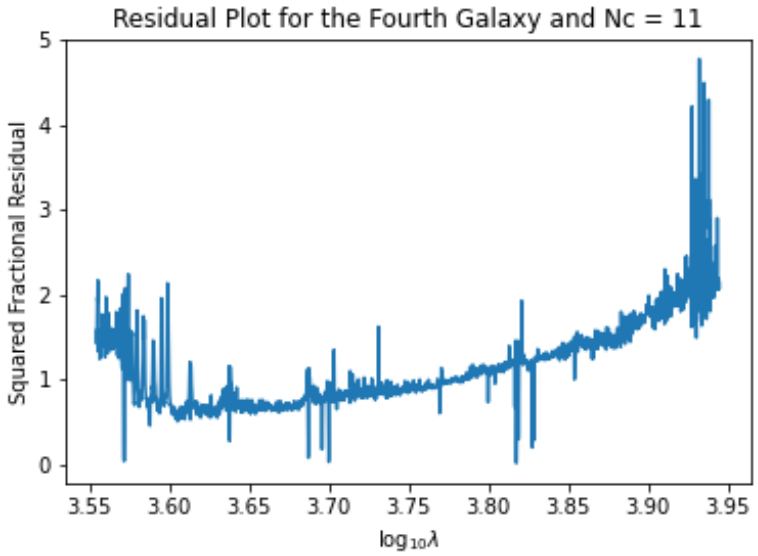

Residual Plot for the Fourth Galaxy and Nc = 6
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

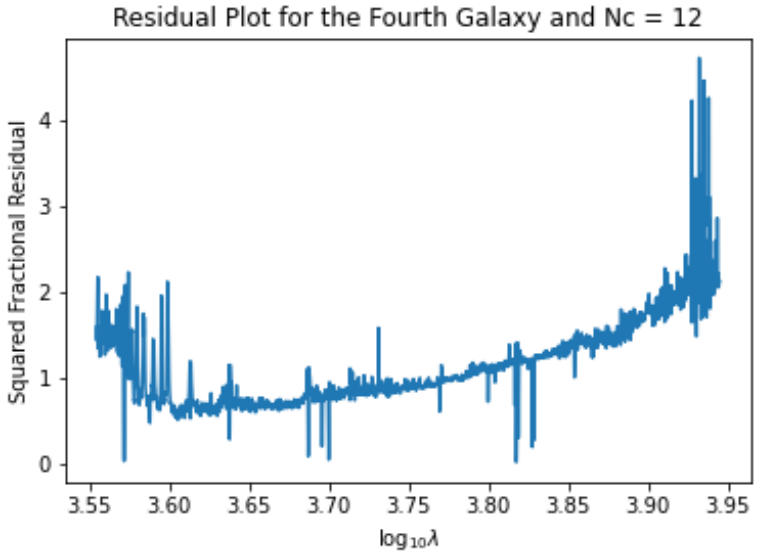

Residual Plot for the Fourth Galaxy and Nc = 7
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

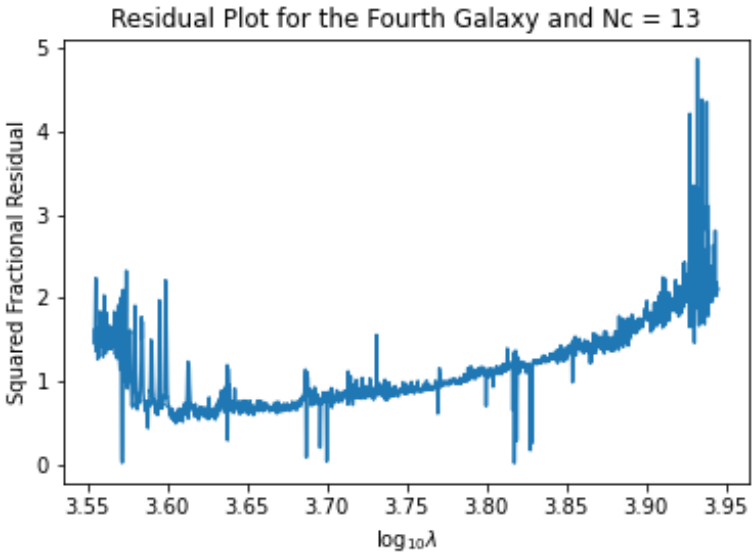

Residual Plot for the Fourth Galaxy and Nc = 8
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

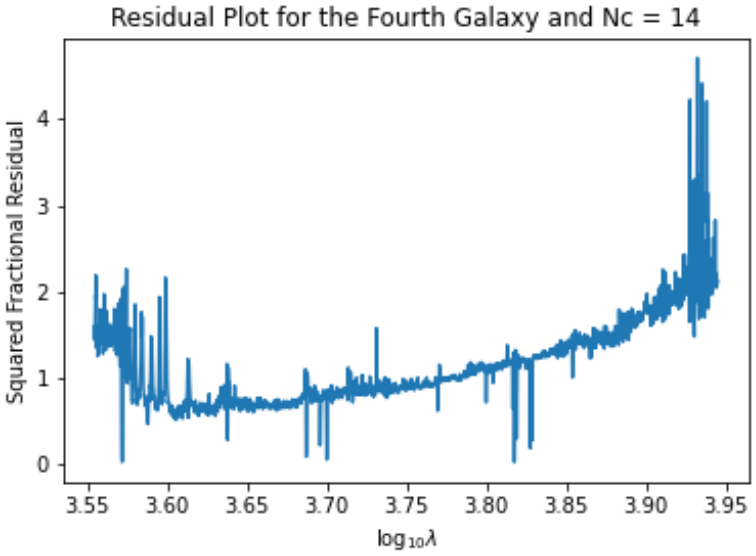

Residual Plot for the Fourth Galaxy and Nc = 9
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

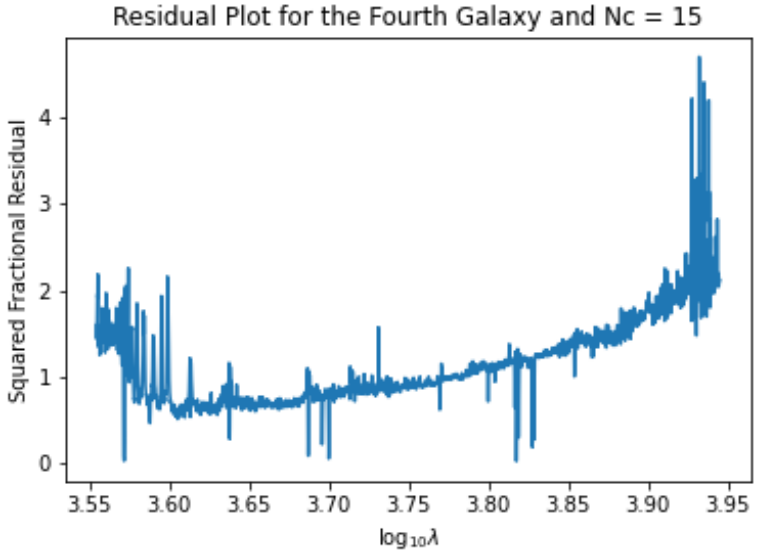

Residual Plot for the Fourth Galaxy and Nc = 10
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

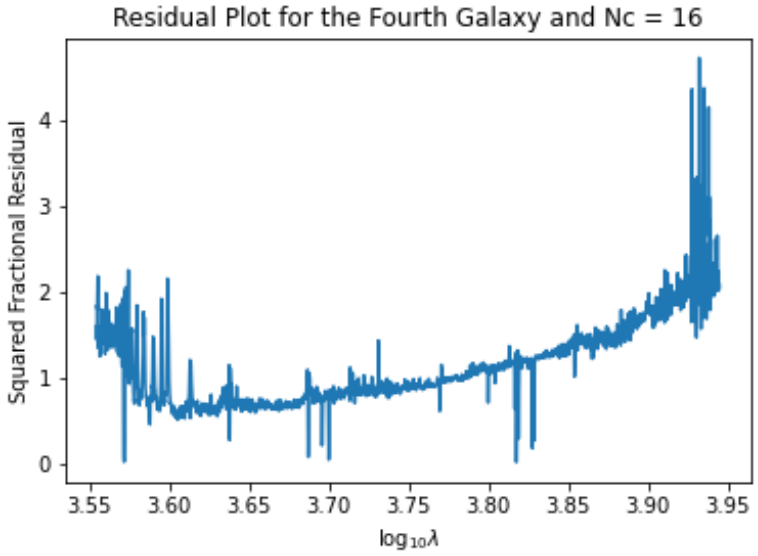

Residual Plot for the Fourth Galaxy and Nc = 11
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

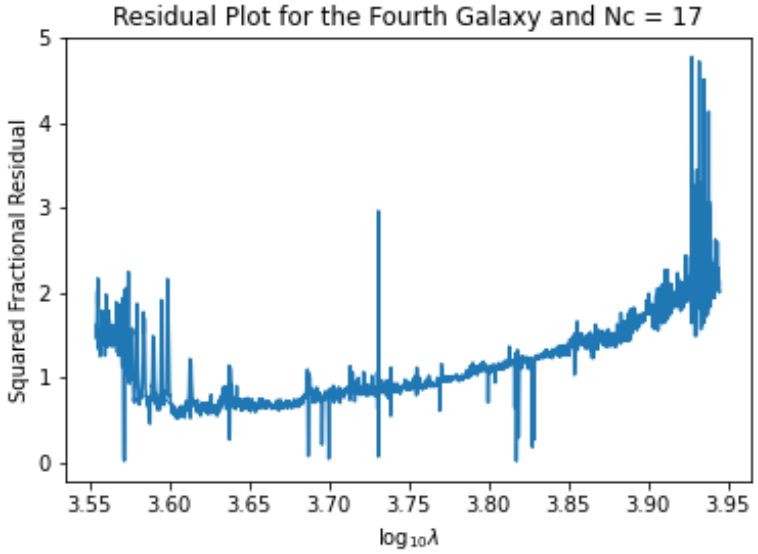

Residual Plot for the Fourth Galaxy and Nc = 12
Squared Fractional Residual
0
1
2
3
4
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

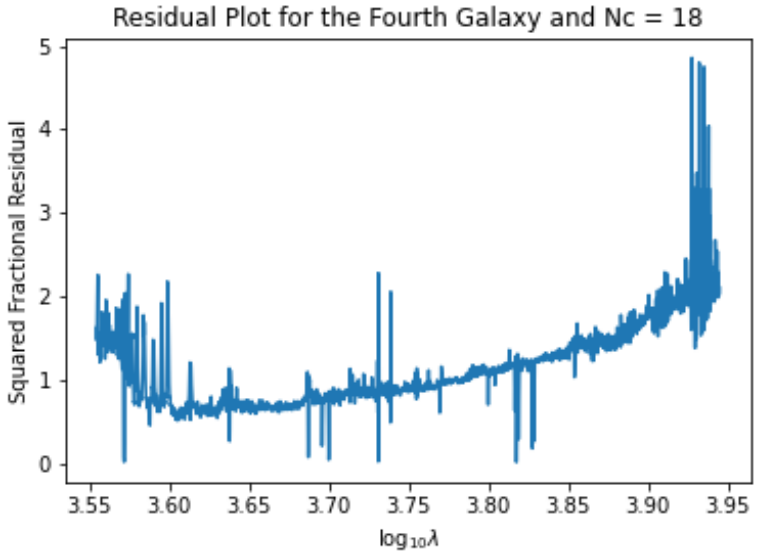

Residual Plot for the Fourth Galaxy and Nc = 13
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

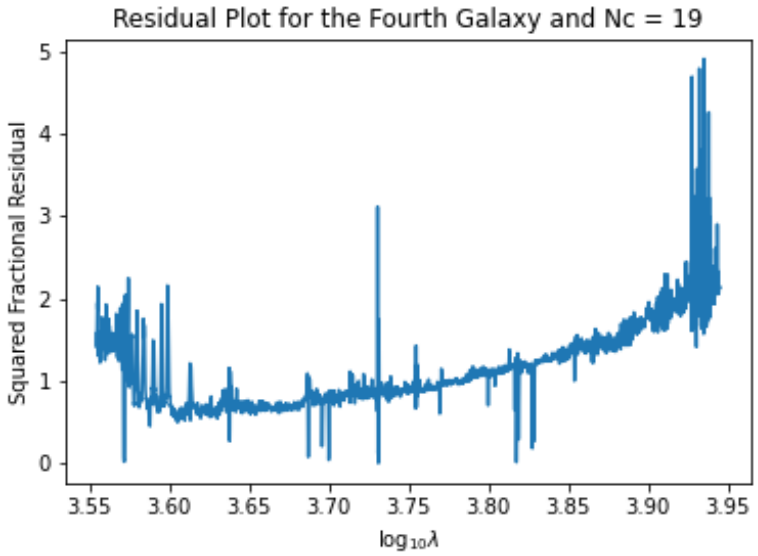

Residual Plot for the Fourth Galaxy and Nc = 14
Squared Fractional Residual
0
1
2
3
4
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

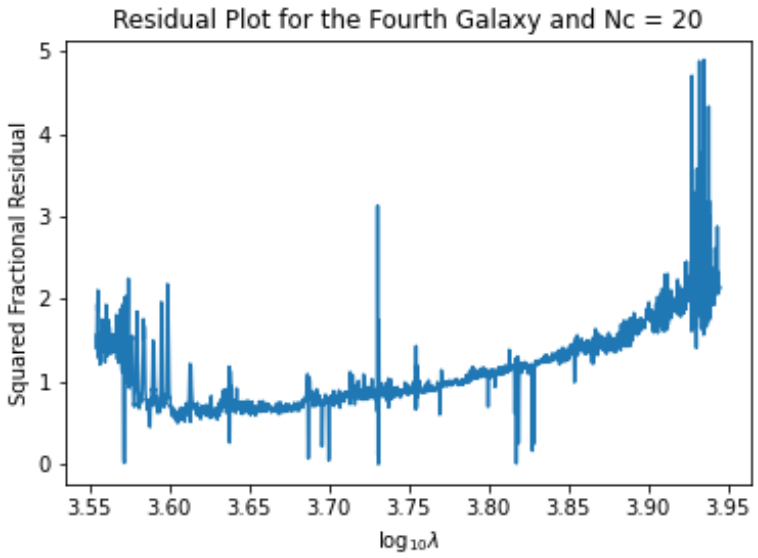

Residual Plot for the Fourth Galaxy and Nc = 15
Squared Fractional Residual
0
1
2
3
4
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Residual Plot for the Fourth Galaxy and Nc = 16
Squared Fractional Residual
0
1
2
3
4
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Residual Plot for the Fourth Galaxy and Nc = 17
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ


Residual Plot for the Fourth Galaxy and Nc = 18
Squared Fractional Residual
0
1
2
3
4
5
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

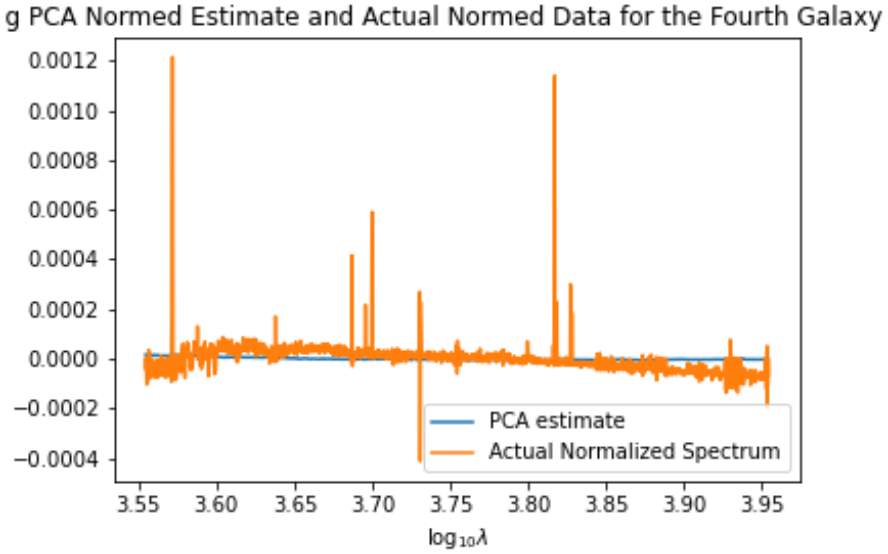


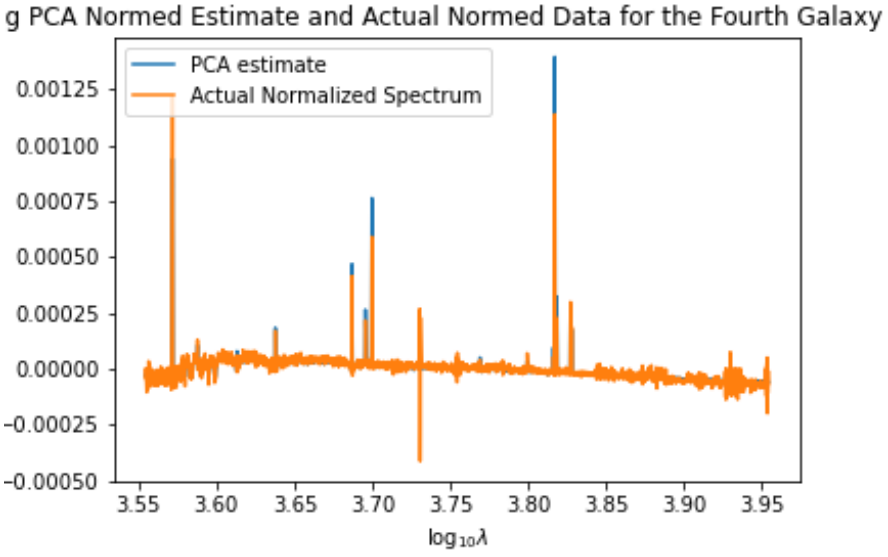



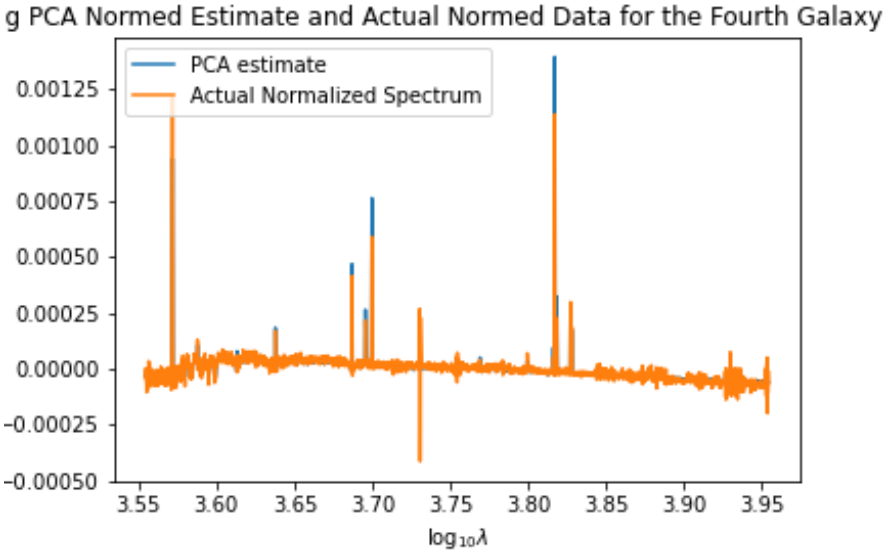

Wavelengths are in Angstroms.

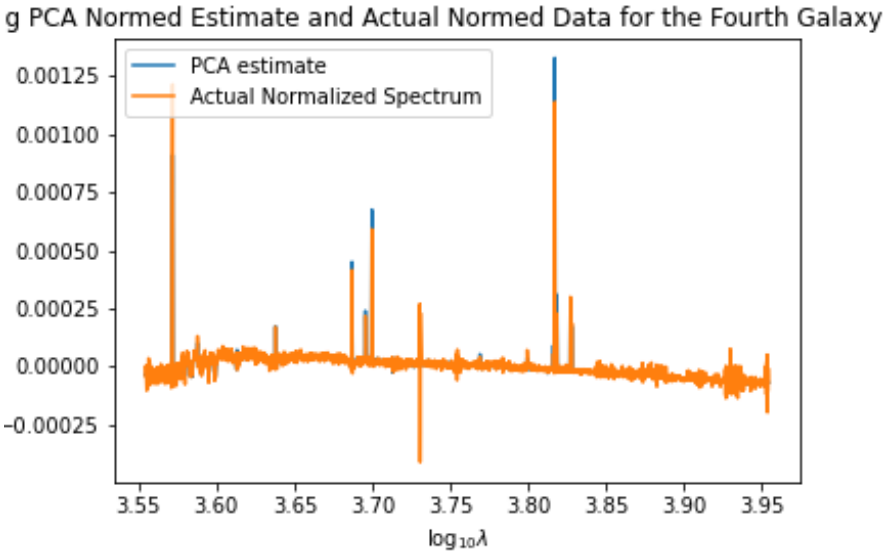
## 3.9 Additional Plots

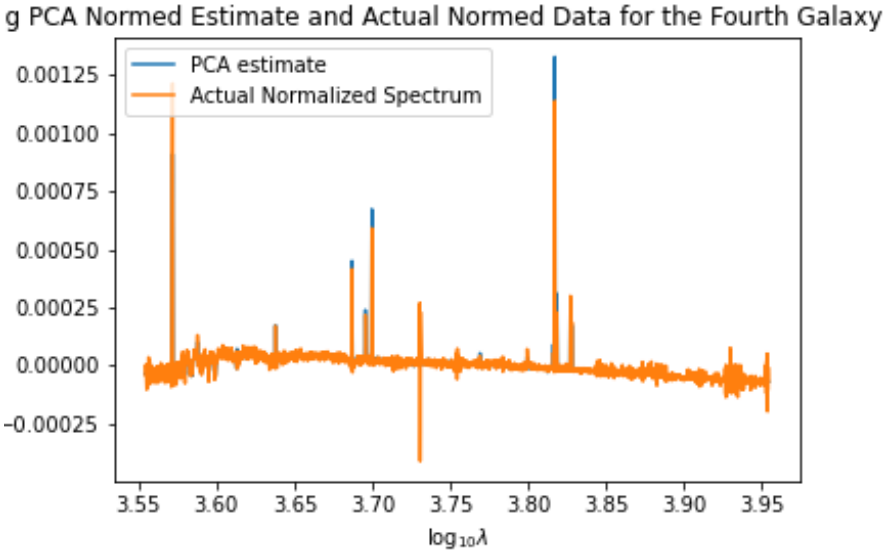
Here we compare out PCA estimates of the normed spectra and the actual normed spectra for the fourth galaxy as the number of principle components vary from 1 to 20. The plots seem to show good agreement. At least there is no constant offset. The squared fractional residual plots are horendous, so I did not show them. They are placed in my github accound.

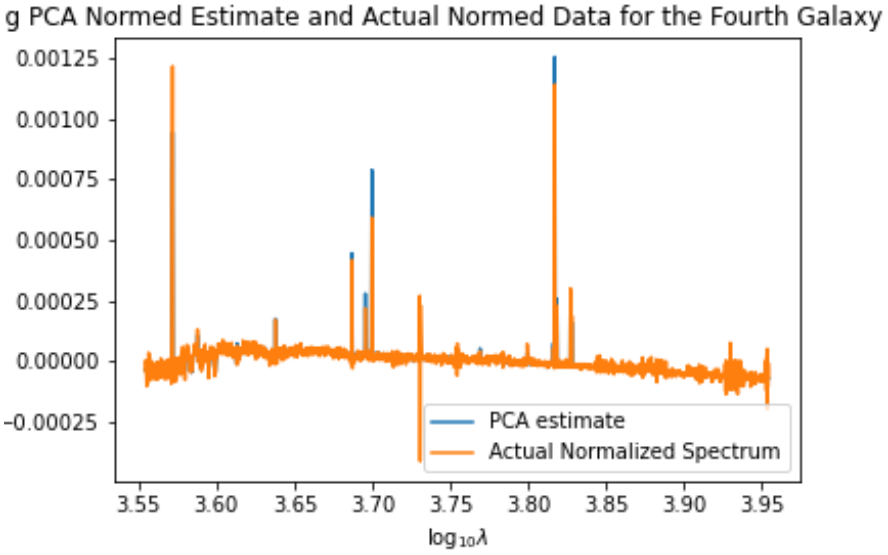

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

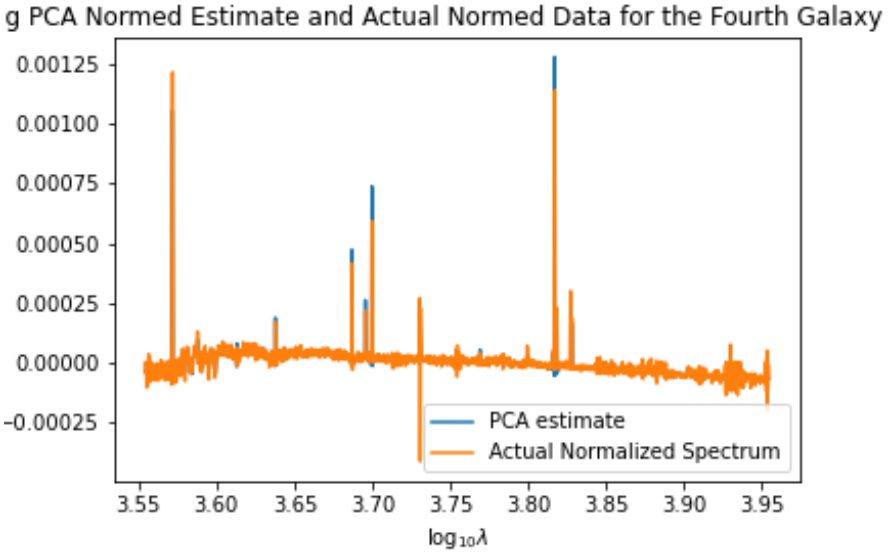

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
PCA estimate
Actual Normalized Spectrum
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
−0.00050
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

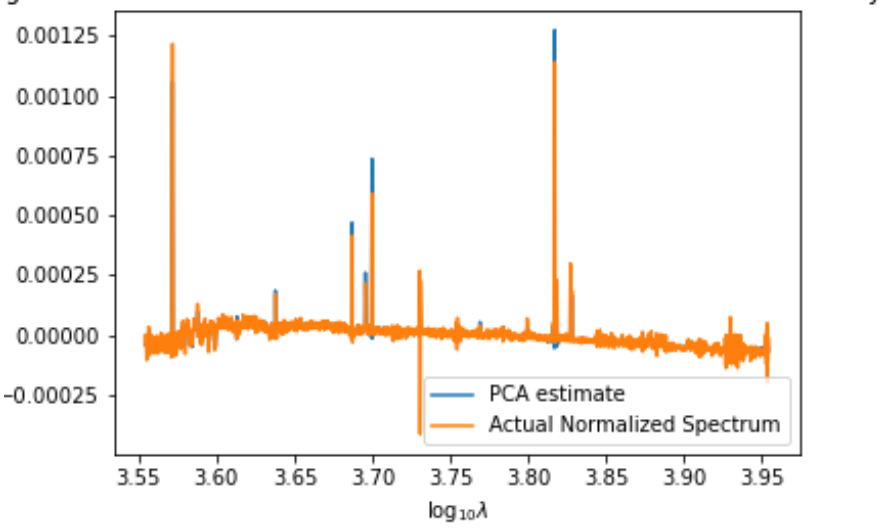

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
PCA estimate
Actual Normalized Spectrum
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
−0.00050
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

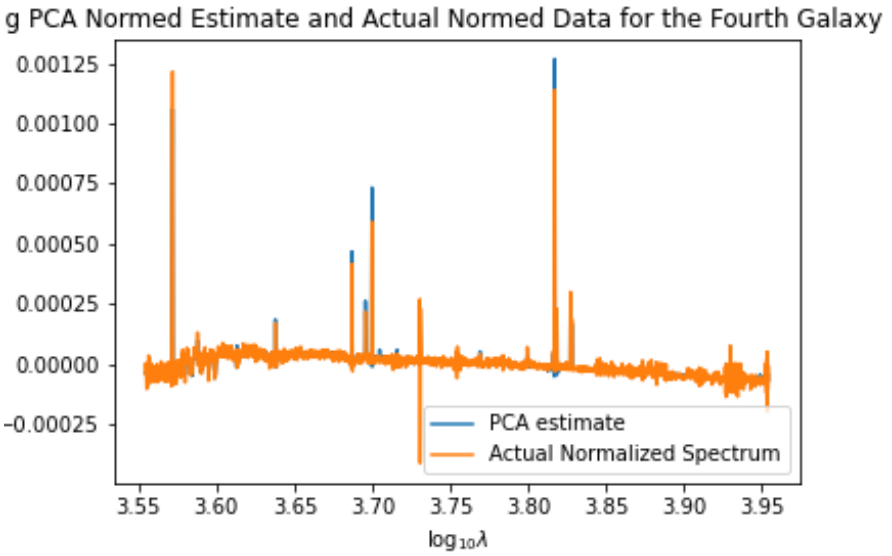

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
PCA estimate
Actual Normalized Spectrum
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

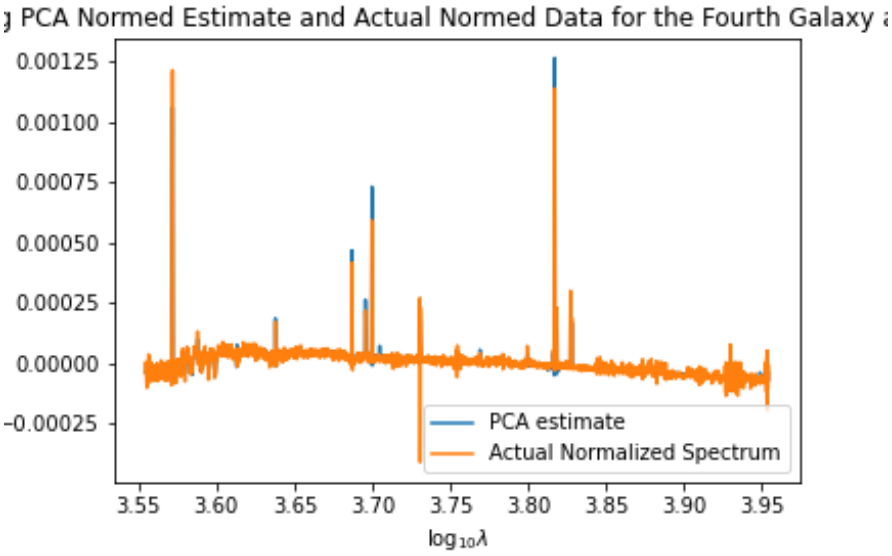

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
PCA estimate
Actual Normalized Spectrum
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

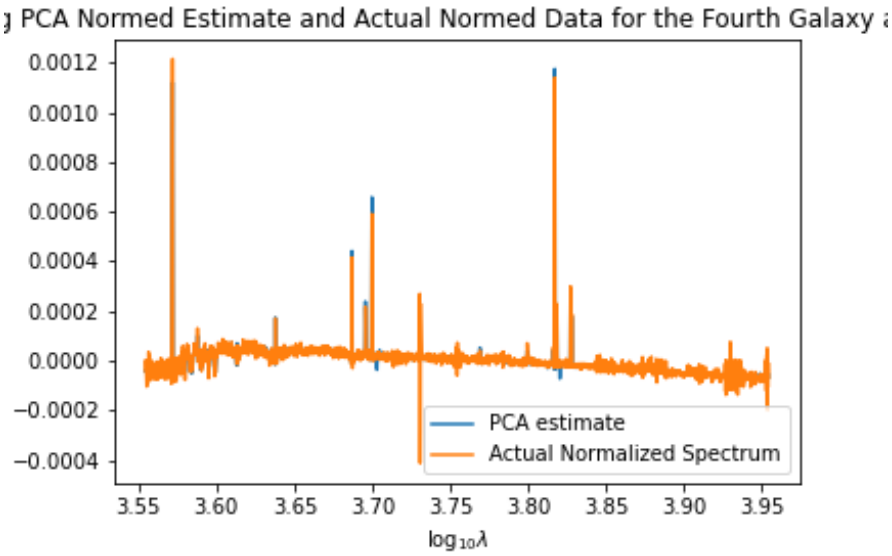

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

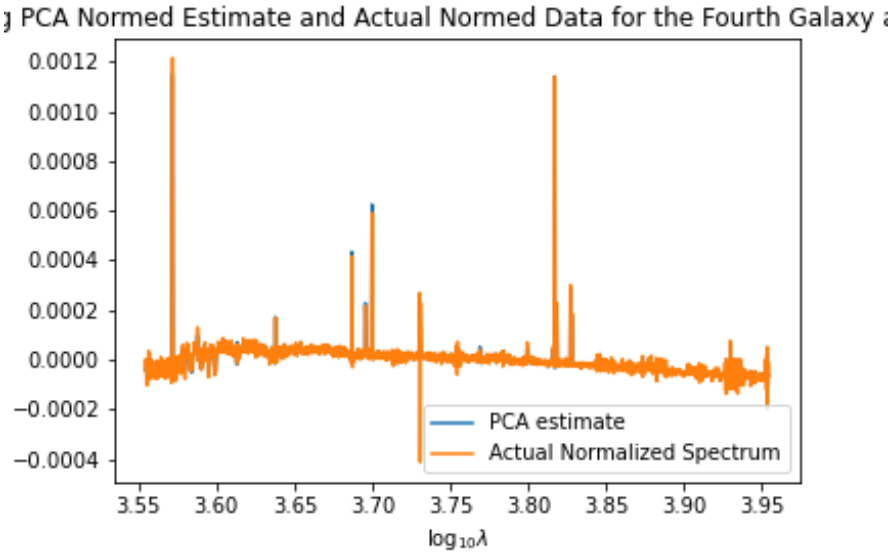

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

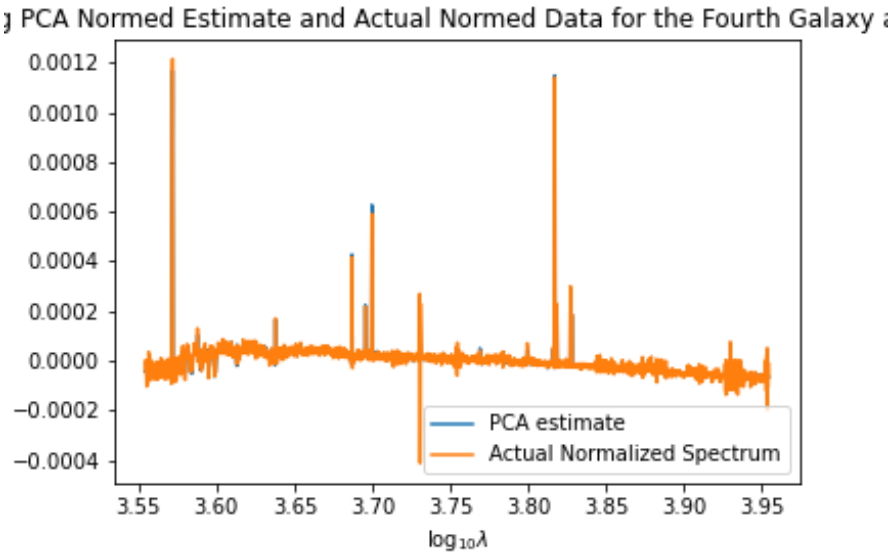

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

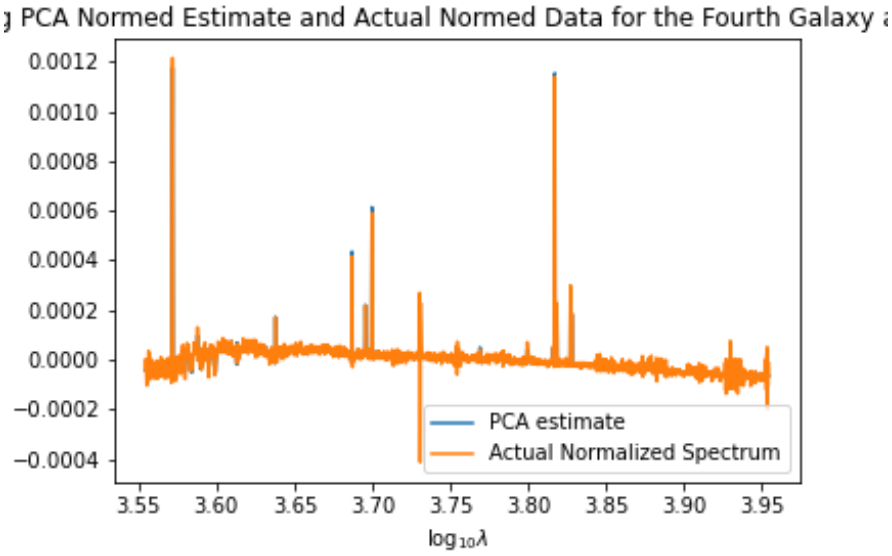

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

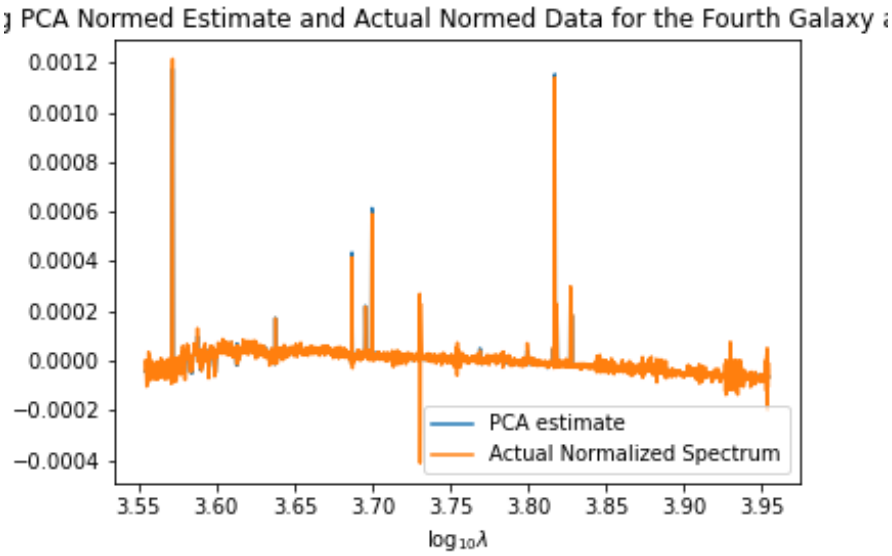

g PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy a
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
−0.00025
PCA estimate
Actual Normalized Spectrum
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ

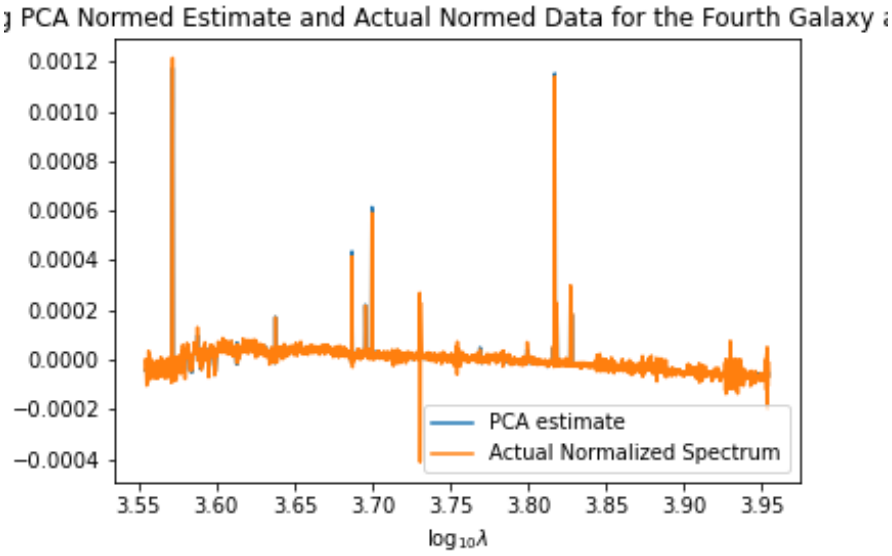

PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum

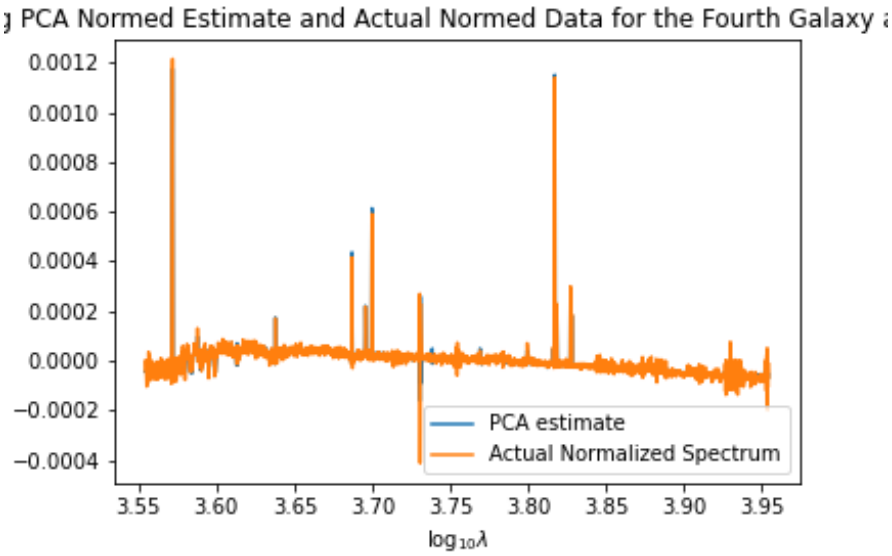

PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum

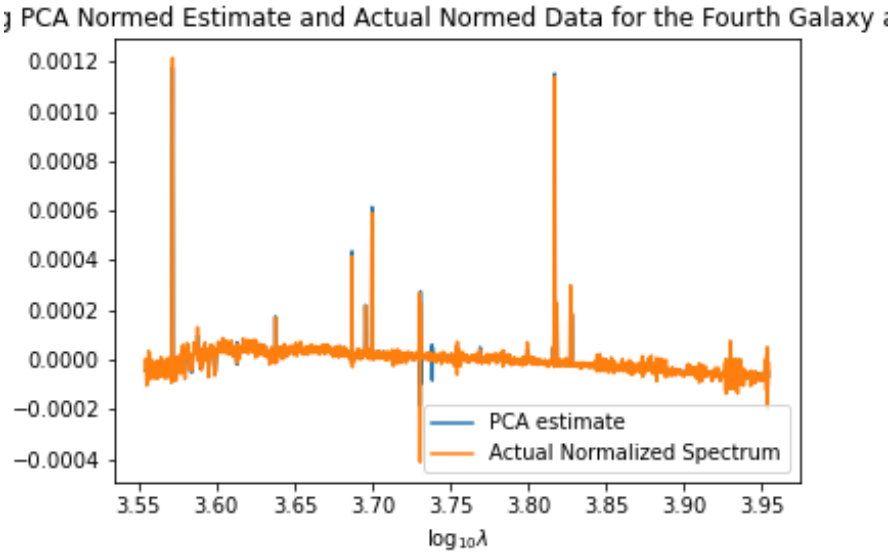

PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum

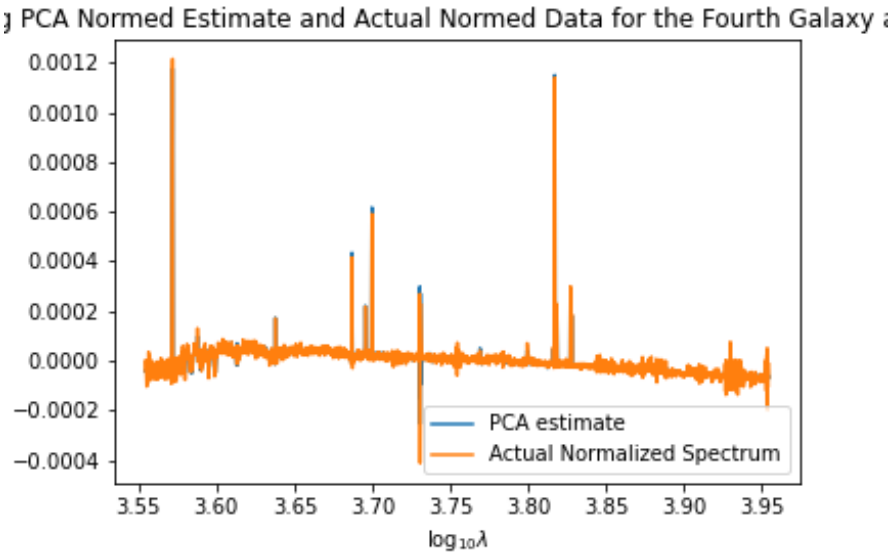

PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum

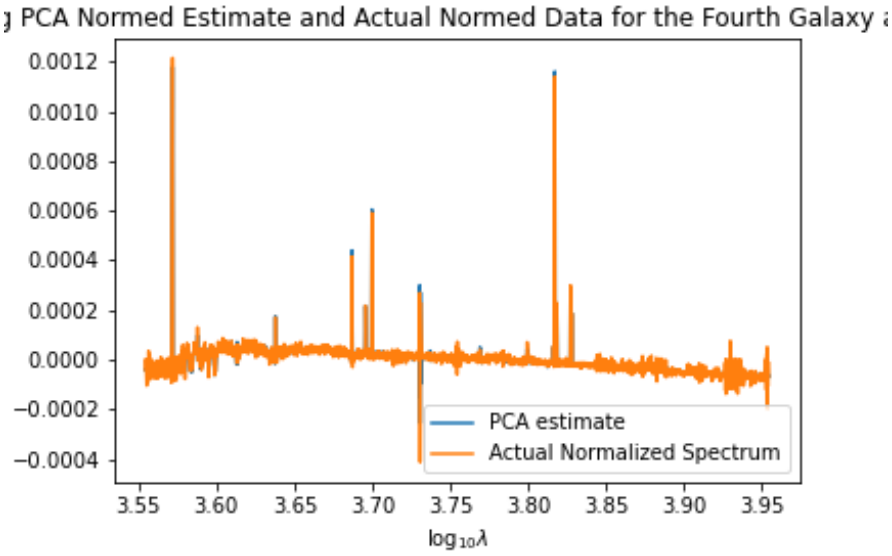

PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum


PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum


PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum


PCA Normed Estimate and Actual Normed Data for the Fourth Galaxy
0.0012
0.0010
0.0008
0.0006
0.0004
0.0002
0.0000
−0.0002
−0.0004
3.55
3.60
3.65
3.70
3.75
3.80
3.85
3.90
3.95
log10λ
PCA estimate
Actual Normalized Spectrum




Wavelengths are in Angstroms.