Computational Physics HW6

Siyuan Chen

October 31, 2023

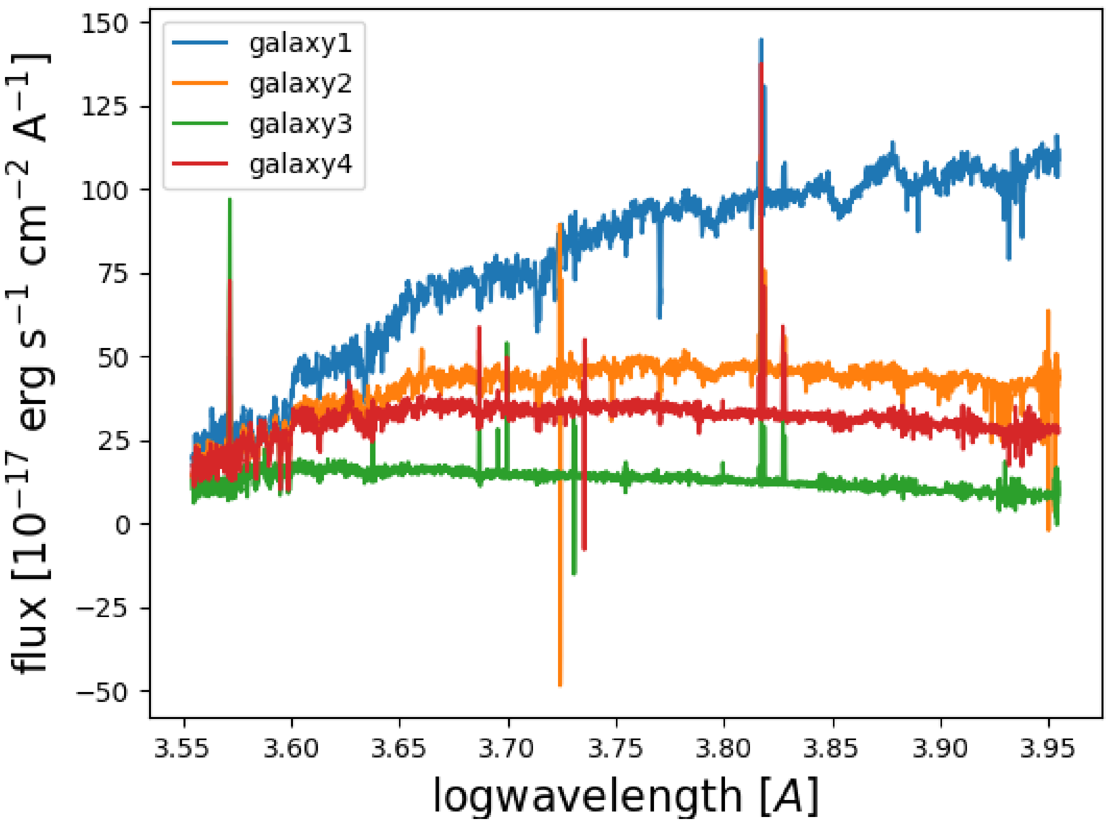

Figure 1: **(a): Flux over wavelength for several galaxies.** From the plot, we can see some striking peaks over certain wavelength, which suggests the components of the atmosphere of the galaxies. Comparing the data with Lyman series for hydrogen atom, we can see the lowest wavelengths for chosen galaxies are above 3500A, which is much higher than hydrogen spectra, which suggests the main components for these galaxies are not hydrogen atoms.

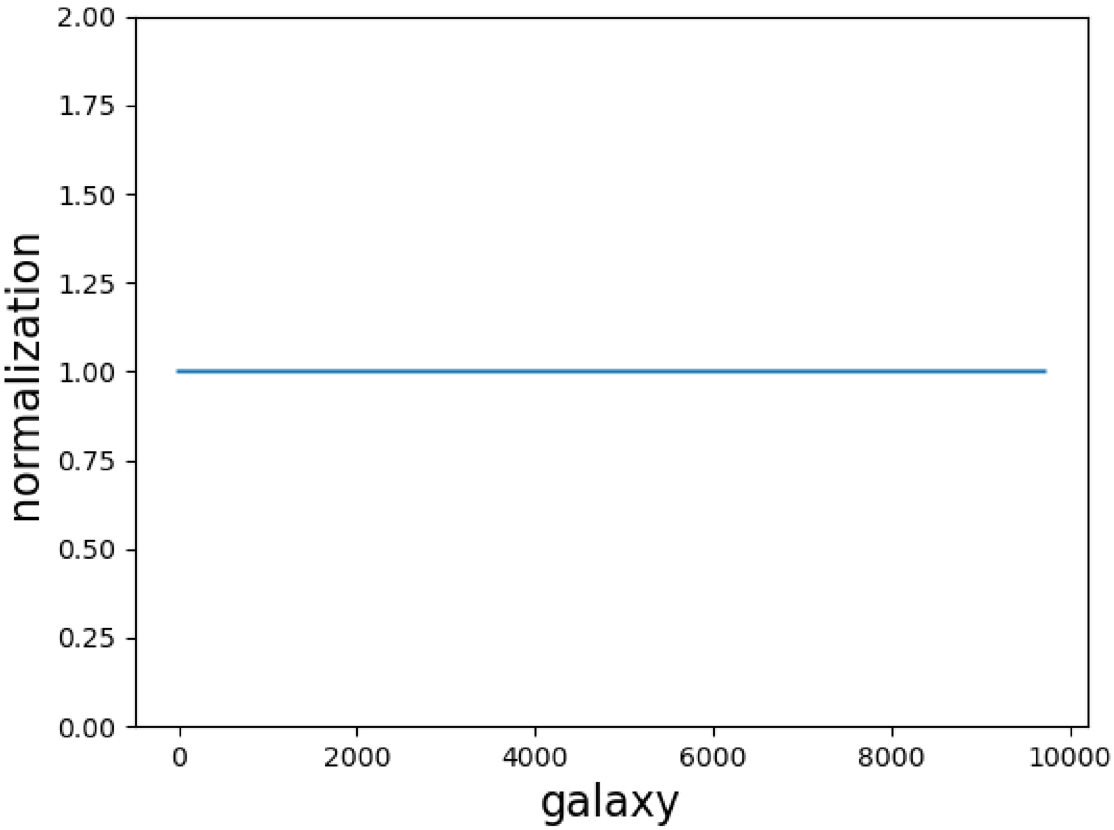

Figure 2: **(b): Normalize flux for all galaxies to be 1.** Since we have huge amount of data over wavelength, it is good enough for us to approximate the integral for flux as a summation over all data points. Get the normalized flux by dividing the summation. Check they are normalized by plotting the new sums. Plot shows normalized flux sums to 1 for all galaxies.

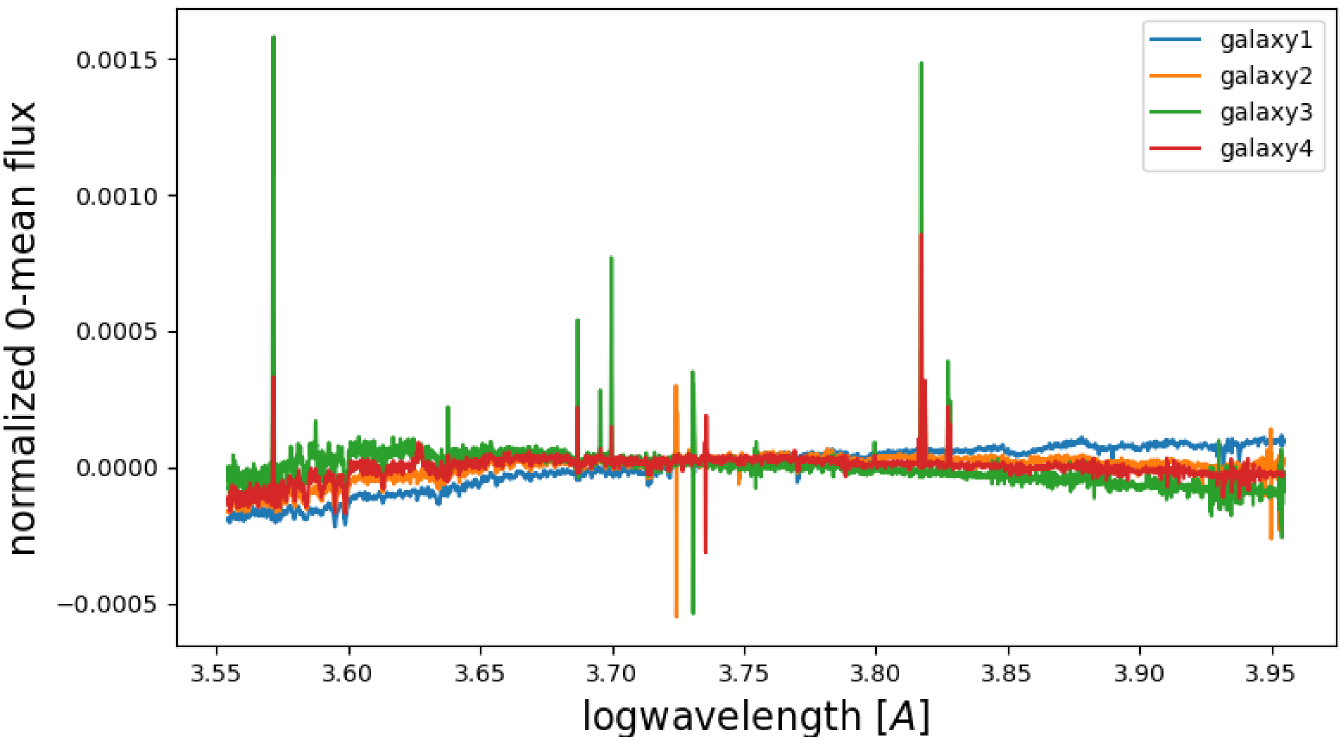

Figure 3: **(c): Plot of residuals of normalized flux from mean.** We subtract normalized mean flux from the normalized flux and get the new spectra for the same galaxies as figure 1.

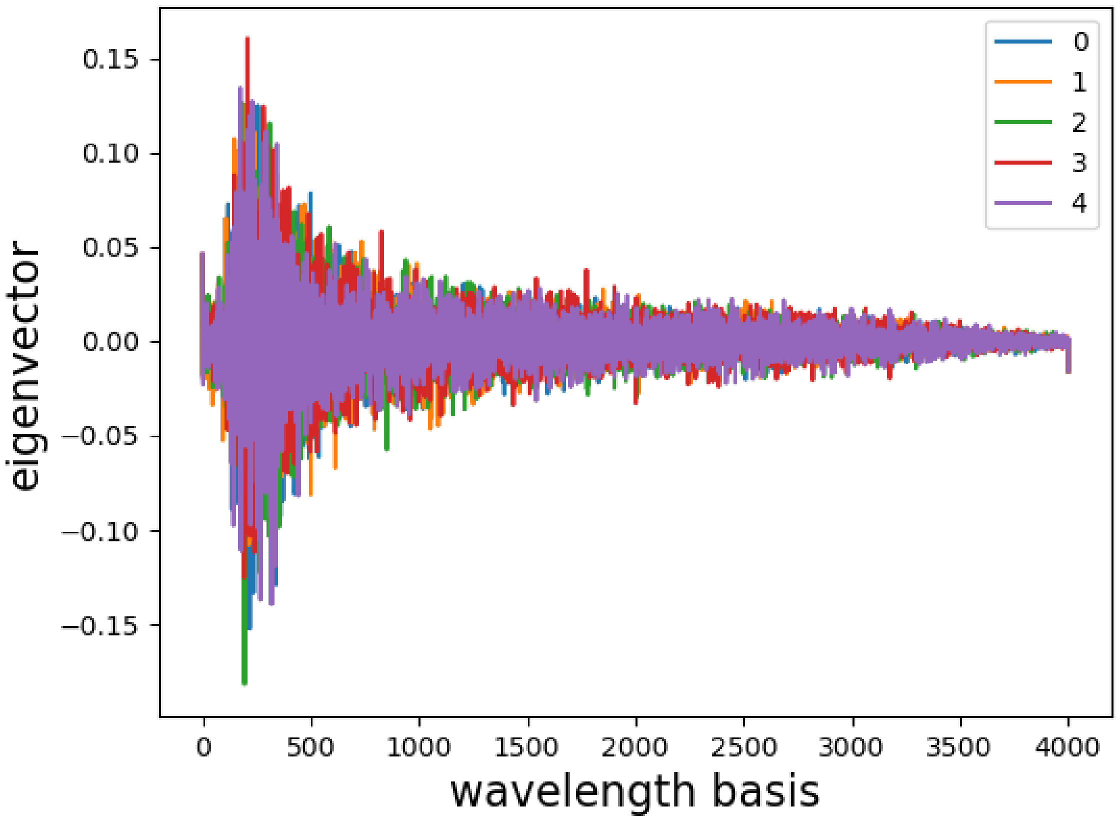

Figure 4: **(d): first 5 eigenvectors for covariance matrix** We already have the R matrix which is the residual matrix of normalized flux from mean for all galaxies. By $C = R^T \cdot R$, we get the covariance matrix C. Check the dimension is 4001×4001, which is what we expected. Find eigenvalues and eigenvectors with np function np.linalg.eig and plot the first 5 eigenvectors.

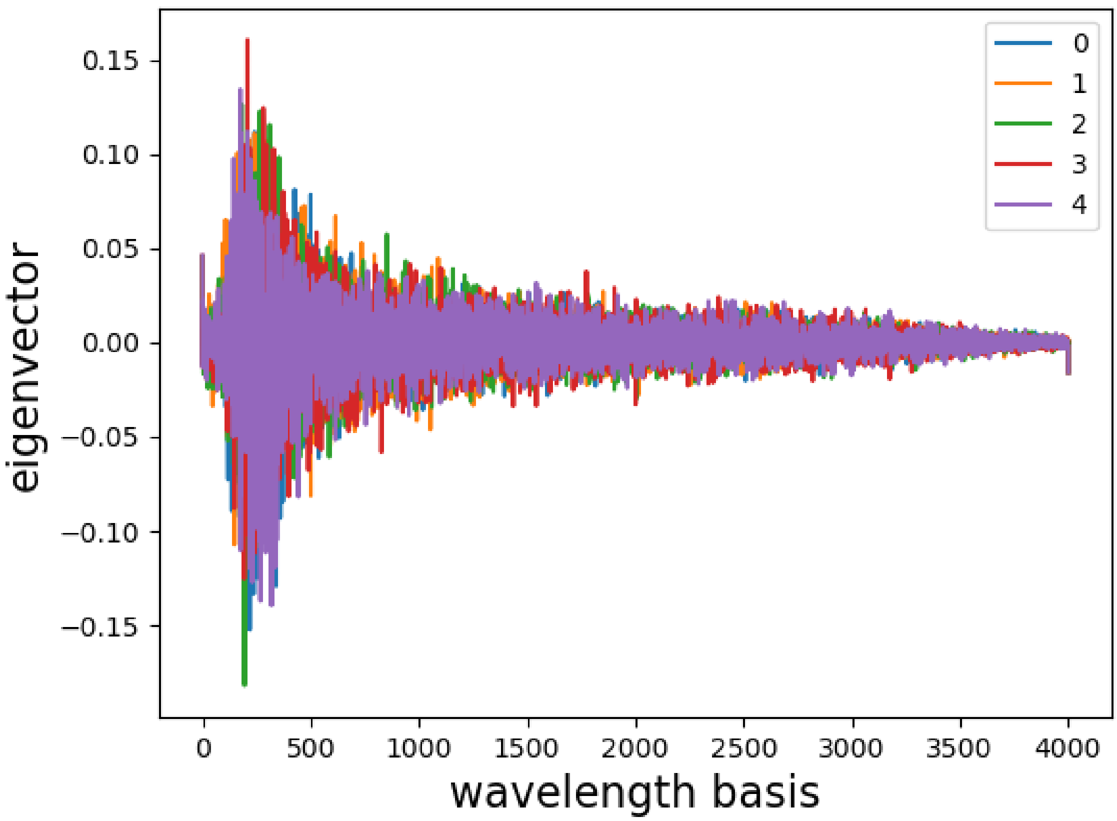

Figure 5: **(e): first 5 eigenvectors for covariance matrix by SVD** By equation (3) in the problem set, we see that $C = R^T \cdot R = V \cdot W \cdot U^T \cdot U \cdot W \cdot V^T = V \cdot W^2 \cdot V^T$, where V is unitary whose columns are eigenvectors and the central structure is a diagonal matrix W^2 whose entries are eigenvalues. We plot the first 5 eigenvectors and find them to be the same as figure 4. Notice that, in general, the eigenvectors may be different by a sign as the SVD decomposition is unique up to a sign in the singular matrices. Comparing the running times, we can see np method is much efficient than the SVD method.

(f): Notice the diagonal matrix for R is W, while for C is W^2, the condition number for C is also square of condition number for R, which suggests C is less stable.

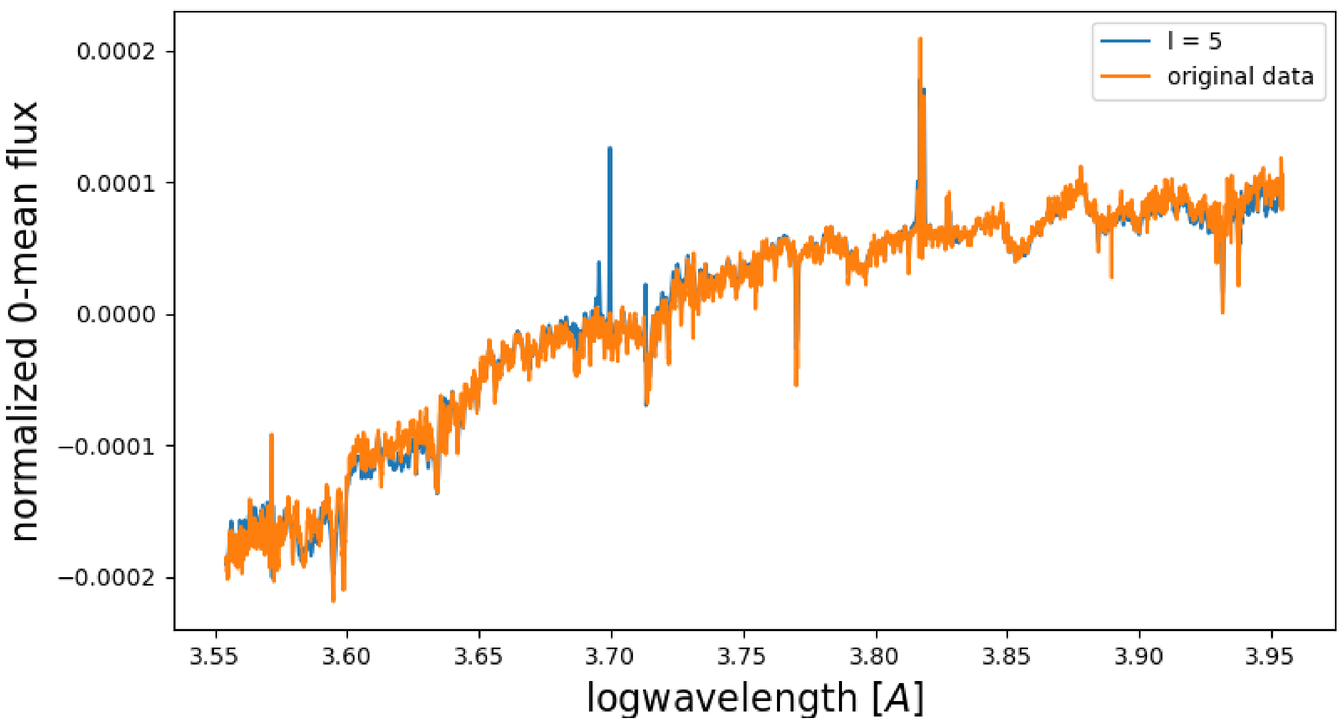

Figure 6: **(g): approximate spectra based on first 5 eigenvectors vs original spectra for galaxy1** Project wavelength data to the first 5 eigenvectors to get the approximate spectra.

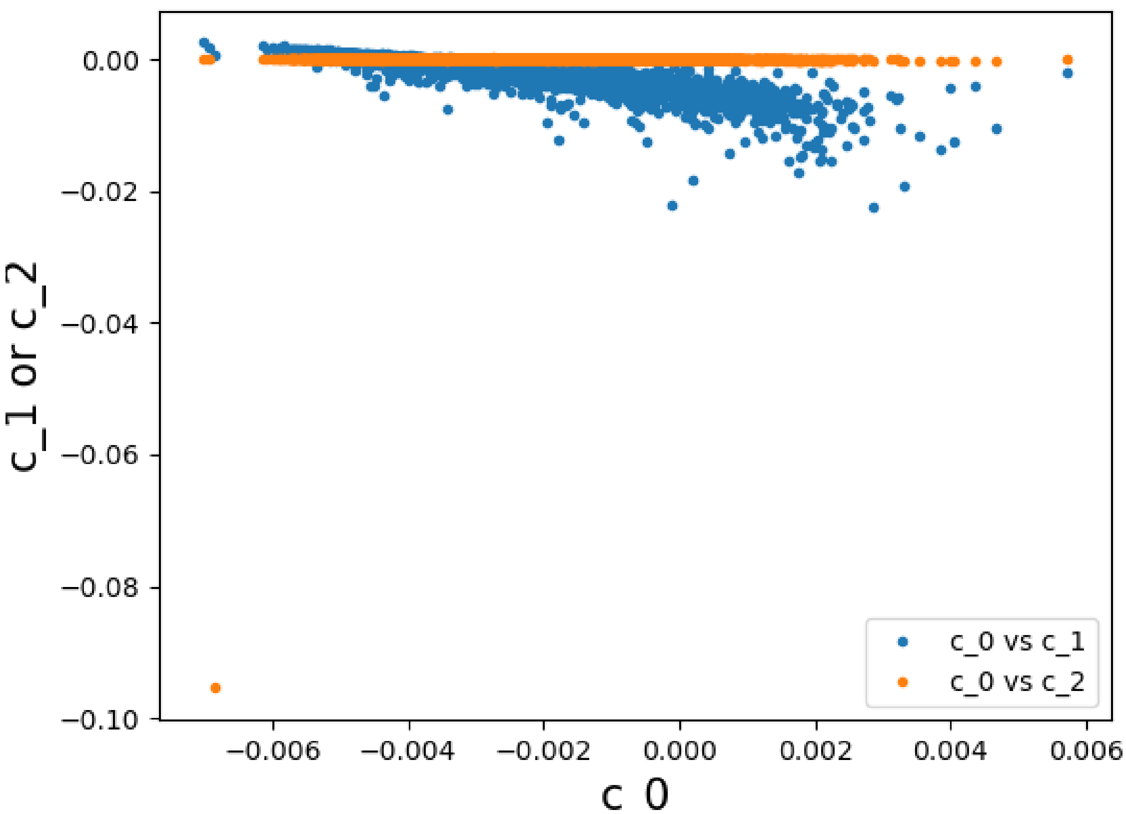

Figure 7: **(h):** C_0 **vs** C_1 **and** C_0 **vs** C_2 **for galaxy1** Project wavelength data to the first 5 eigenvectors to get the approximate spectra. We can also obtain the weight matrix for the projection. Plot the relations between the first three weights coefficients.

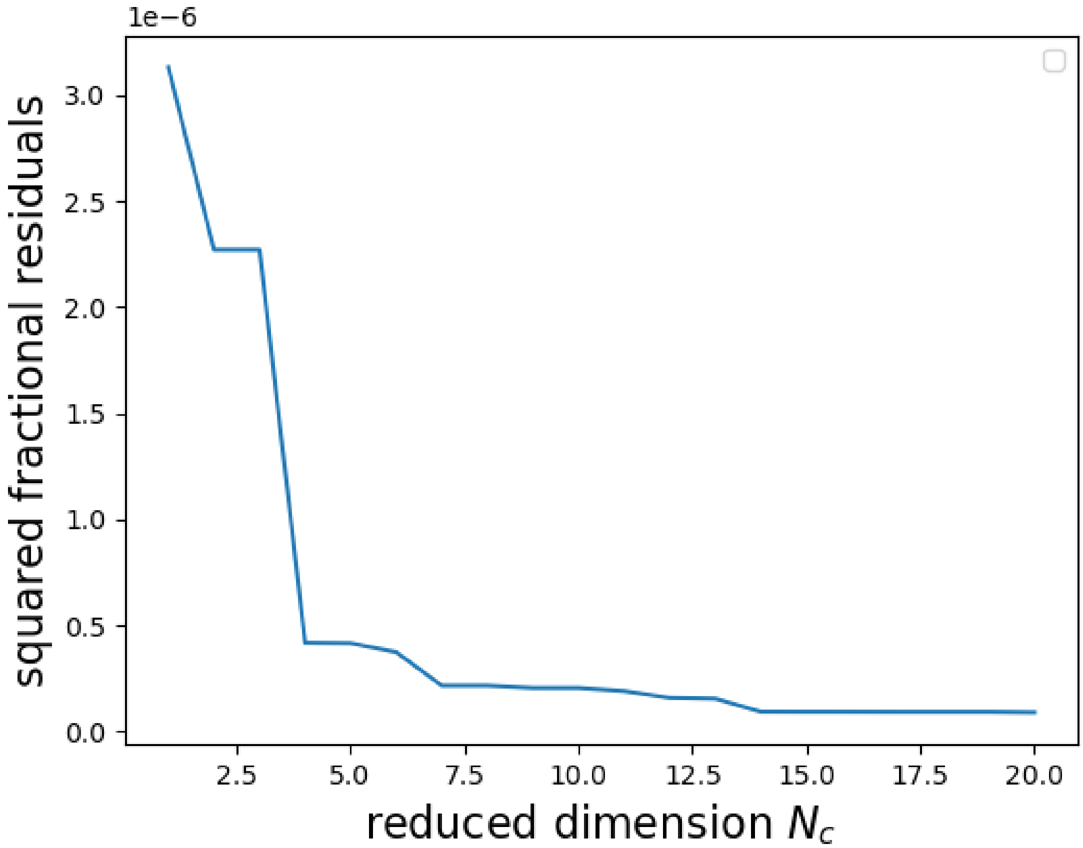

Figure 8: **(i): squared fractional residuals for approximate spectra based on reduced dimension N_c for galaxy1** Project wavelength data to the first N_c eigenvectors to get the approximate spectra. Calculate the residuals and observe it declines as N_c increases as the dimensions filtered out decreases and spectra become more accurate.