Ph21 Assignment 6

Kyriacos Xanthos

March 2020

2-dimensional case

In this part I created a linspace for the dependent variable and added some noise to the independent variable using np.random.uniform. On the plot below it can be seen that there is a high linear correlation between the x and y values.

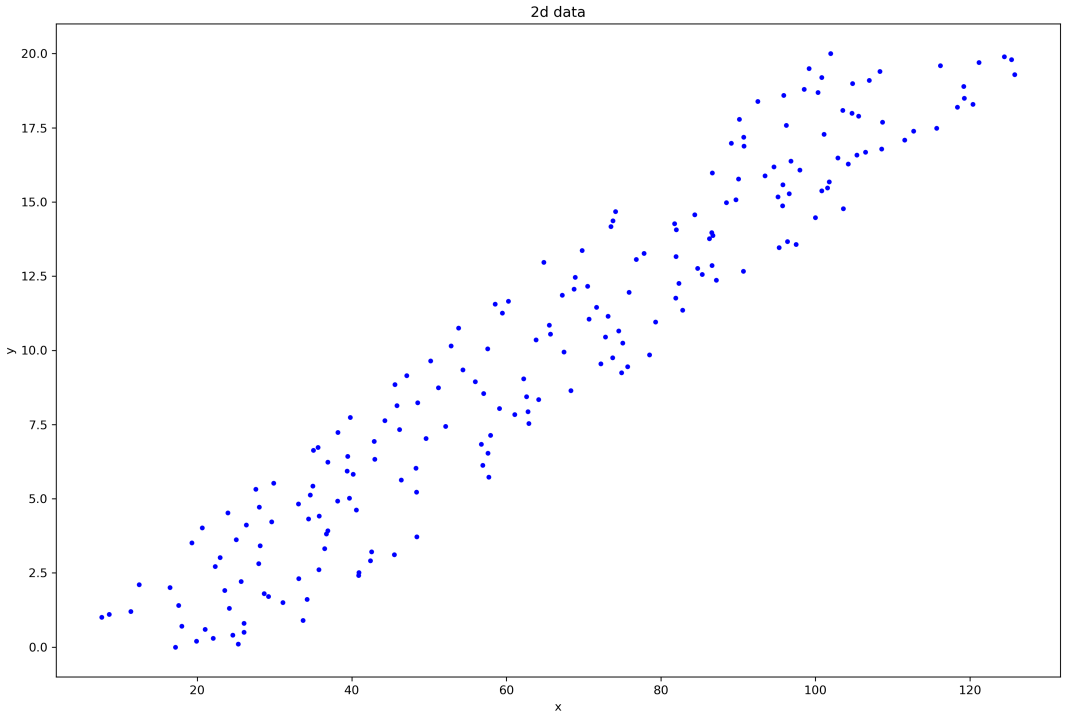

Figure 1: 2-dimentional data

Now performing PCA, the eigenvalues were found to be: $\begin{pmatrix} 961.72594953 \\ 2.76465338 \end{pmatrix}$ and the eigenvectors: $\begin{pmatrix} -0.98366527 \\ -0.18000733 \end{pmatrix}$. As expected there is a high correlation between our data.

3-dimensional case

Now we turn our attention to a 3 dimensional case. This can be interpreted as a system of springs that can move in the x,y,z direction, a system of H2O bonds etc. I have created noise for both x and y variables now. A 3d plot below shows a scatter of all of these points:

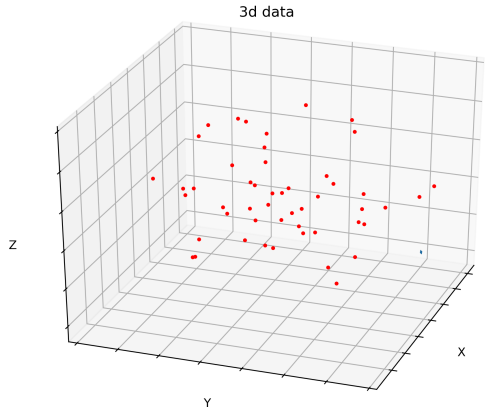

Figure 2: 3-dimensional plot

A better visualisation can be made using a 2d plot and plot each dimension with a different color.

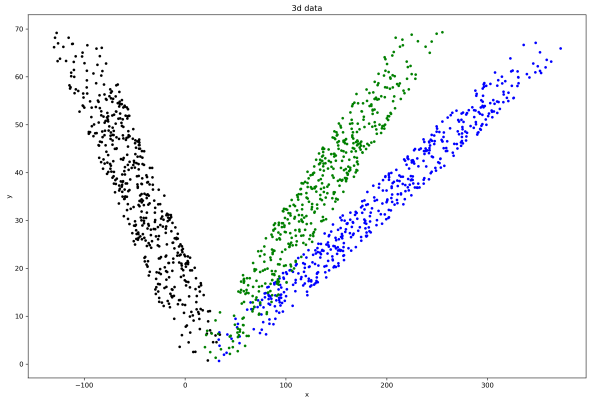

Figure 3: 2-dimensional plot of the 3d case

Performing PCA, the eigenvalues were found to be: $\begin{pmatrix} 6481 \\ 2639 \\ 1315 \\ 39 \\ 19 \\ 7 \\ 2 \end{pmatrix}$. The eigenvectors were: $\begin{pmatrix} -0.97 \\ -0.18 \\ 0.08 \\ -0.03 \\ -0.06 \\ -0.02 \end{pmatrix}$. It can be seen from Figure 3 that the variables are very correlated in the begin and start to spread away to not correlated at all as their x and y values increased. This can be seen from the eigenvectors we found. At first there is a huge correlation but it decreases as the values increase. This shows how successful PCA is in finding the correlation between variables for any number of dimensions.