# Question 3 (20 pts) --COMPLETE

Describe the benefit of a gradient approach w.r.t. the SVD approach for the PCA transformation.

From April 5th lecturer re-question :::

Describe the benefit of the power method and the svd decomposition for implementing PCA

-------------------------------------------------------------------------------------------------------------------------------

To my understanding this question requires no code so I will answer this question in the following breakdown: first I will describe what PCA is. In effect of that I will then expand and explain how power methods or svd can be used to help with PCA.

Principal component analysis (PCA) is a machine learning technique of a linear model that is used to evaluate and map k-dimensional features into latent space (Additionally, PCA is good for finding patterns/structure in multidimensional data). It does so by maximizing the Eigenvalue for each specific feature point. After the maximum sum of squares (eigenvalue) is found for each feature point. Singular value decomposition (SVD) can be used with PCA. Using SVD allows the model to scale its found linear combination of variables so that it is on a scale of 1. Since the ratio remains the same during this time, nothing has changed drastically about our found points. It allows the model to work in faster time keeping these values scaled.

Power method also known as power iteration also helps in the PCA process. Using the Power method as a way to implement PCA allows us to find the top k principal components in a dataset. By using a matrix M where the data in M represents the components in our data (feature points). Manipulation of our feature points by multiplying by the dot product of say X transpose X, allows us to form an ellipse. The longest point of our ellipse composed by our matrix gives us the most important principal component for that set. This is effectively producing our eigenvalue that we found from using PCA with SVD. The benefit of using power iteration for PCA is that it gives us a means to find the patterns and structure in k-dimensional data that someone might not automatically see when data is in it raw form. By mapping the data into latent space (if you are able to minimize the needed data into around 2-3 dimensions) you can visualize where there is a clear line in the separation for classification of the data.