DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, IIT MADRAS

CS5691: Assignment 1B
Team 1

Sheth Dev Yashpal
CS17B106

Harshit Kedia
CS17B103

1 1 DIMENSIONAL DATA

1.1 LEAST SQUARE REGRESSION

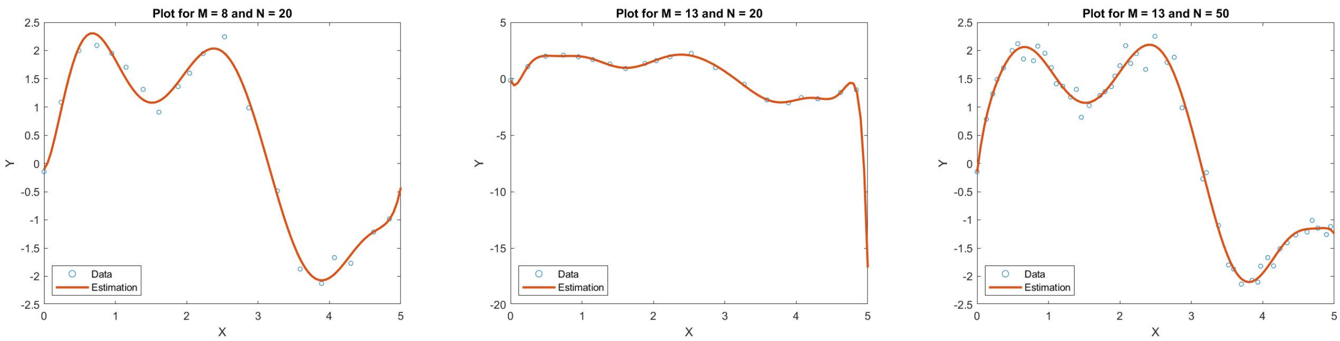

Figure 1.1: Plots for $N = 20$ Data Points for different Model Complexities $M = 8$ and $M = 13$. Plots for $M = 8$ Model Complexity for different Number of Data points $N = 20$ $N = 50$.

- First we plot our Least Square Regression model for different Complexities keeping the number of Data Points constant. We observe that, for lower values of M the curve fails to capture the shape of the curve and for higher values of M, the curve starts over-fitting. Next we plot the estimated function for a particular model complexity but varying No. of Data points and observe that as we increase N, the over-fitting is resolved.

1.2 RIDGE REGRESSION

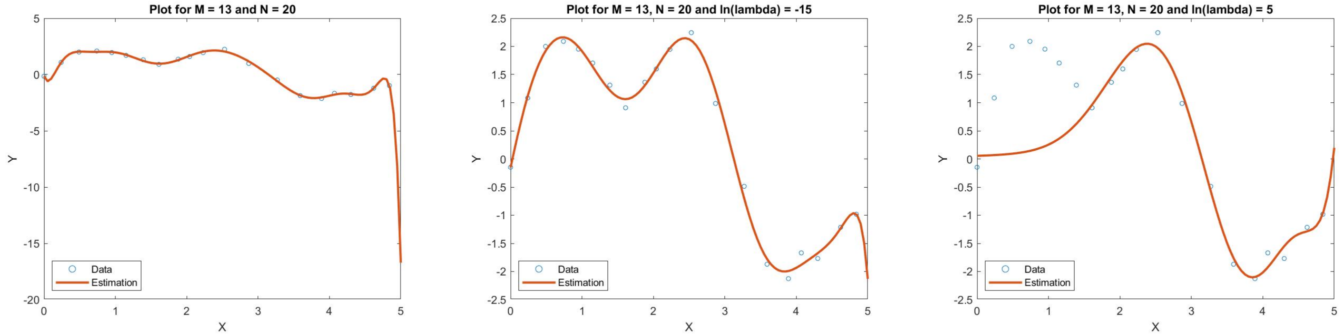

Figure 1.2: Plots for $M = 13$ Model Complexity and $N = 20$ Data points for different values of Regularization Parameter $\ln\lambda = -\infty$, $\ln\lambda = -15$ and $\ln\lambda = 5$.

- Next, for a fixed value of M and N we vary the Regularization Parameter of Ridge Regression. We observe that the over-fitting Curve for Least Square case is now controlled, however if we have very high value of the Regularization Parameter we see that all efforts are in vain!

- We also plot the error of our model for training and development sets vs model complexity and the regularization parameter. In the plot with varying model complexity we see that at first the error decreases initially but later on due to over-fitting it starts increasing again. Similar trend is observed for varying regularization parameter however the plot shifts from over-fit to under-fit.

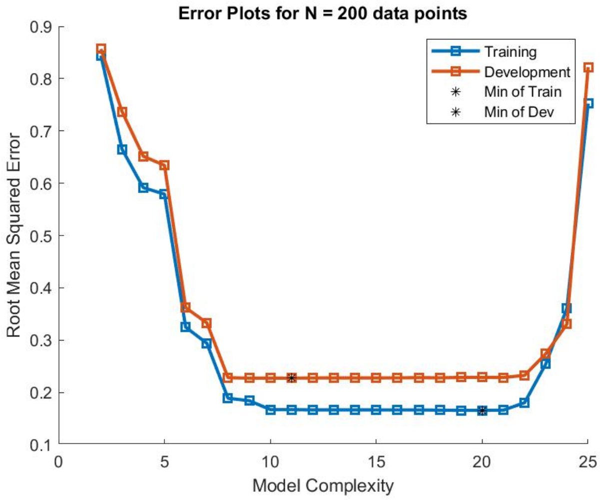

Figure 1.3: Error Plot for Training and Development Data vs Model Complexity for a fixed number of Data Points in case of Least Square Regression.

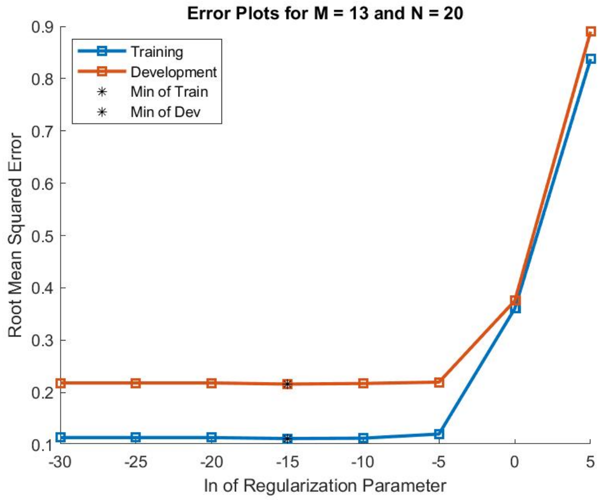

Figure 1.4: Error Plot for Training and Development Data vs Regularization Parameter for a fixed number of Data Points and Model Complexity.

1.3 Best Model

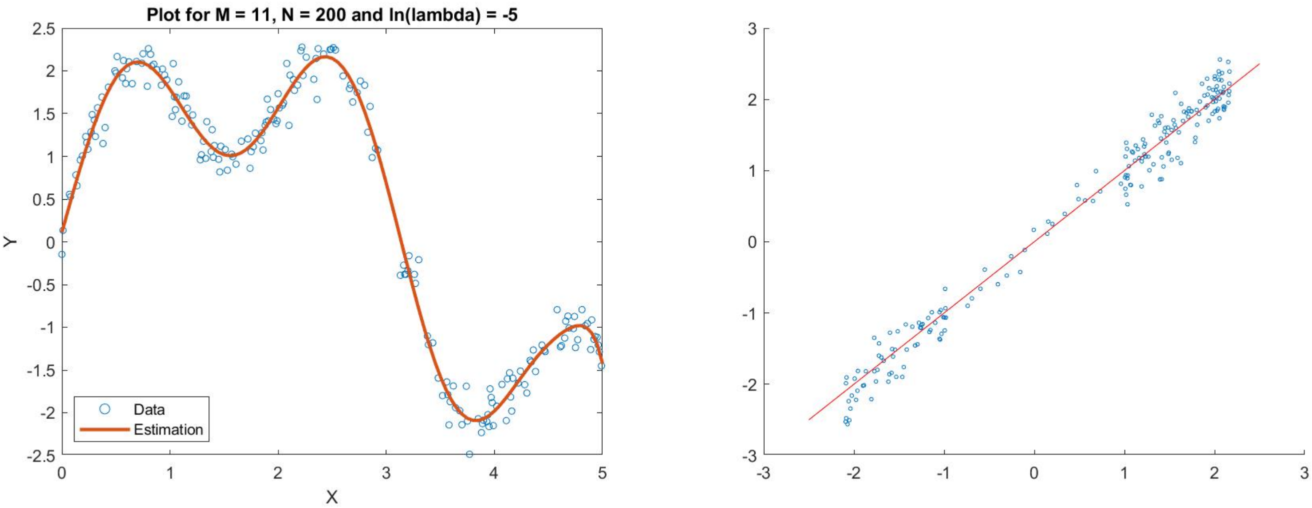

Figure 1.5: The Best Performing Model and the Scatter Plot with our Model Output on $X - axis$ and Target Output on $Y - axis$ for the Development Data Set.

- Finally we plot our Best Trained Model and the corresponding scatter plot with our Model output and the Target output.

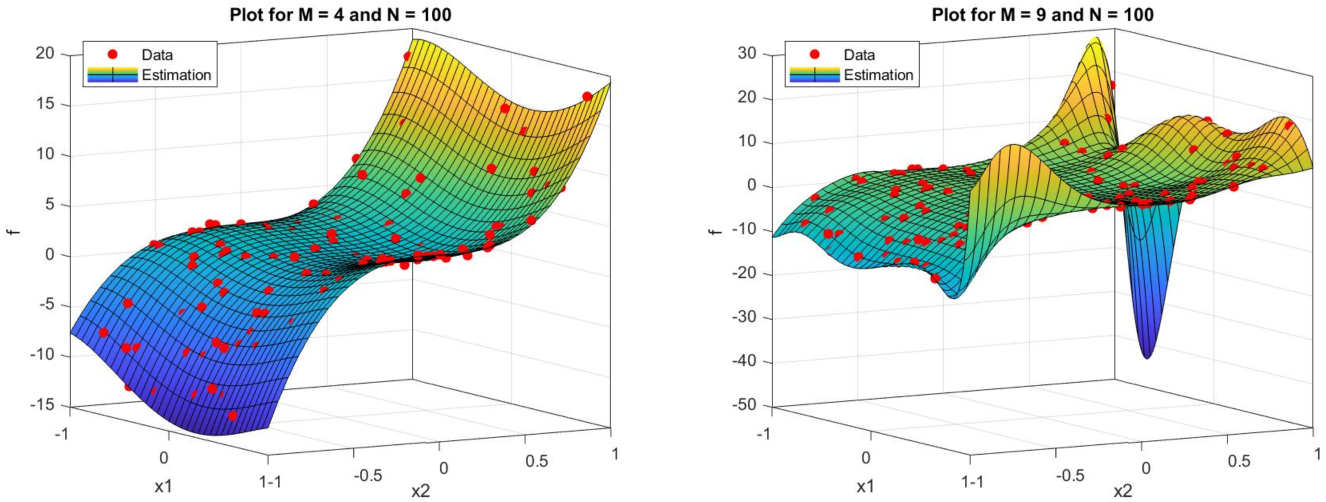

Figure 2.1: Plots for $N = 100$ Data Points for different Model Complexities $M = 4$ and $M = 9$.

2 2 Dimensional Data

2.1 Least Square Regression

- First we plot our estimated function for different Model Complexities keeping the number of Data Points constant. We observe that $M = 4$ Model Complexity gives the best result for almost any value of N. Also we observe that for higher values of M, the curve starts over-fitting.
- Next we plot the estimated function for a particular model complexity but varying No. of Data points and observe that as we increase N, the over-fitting is resolved.

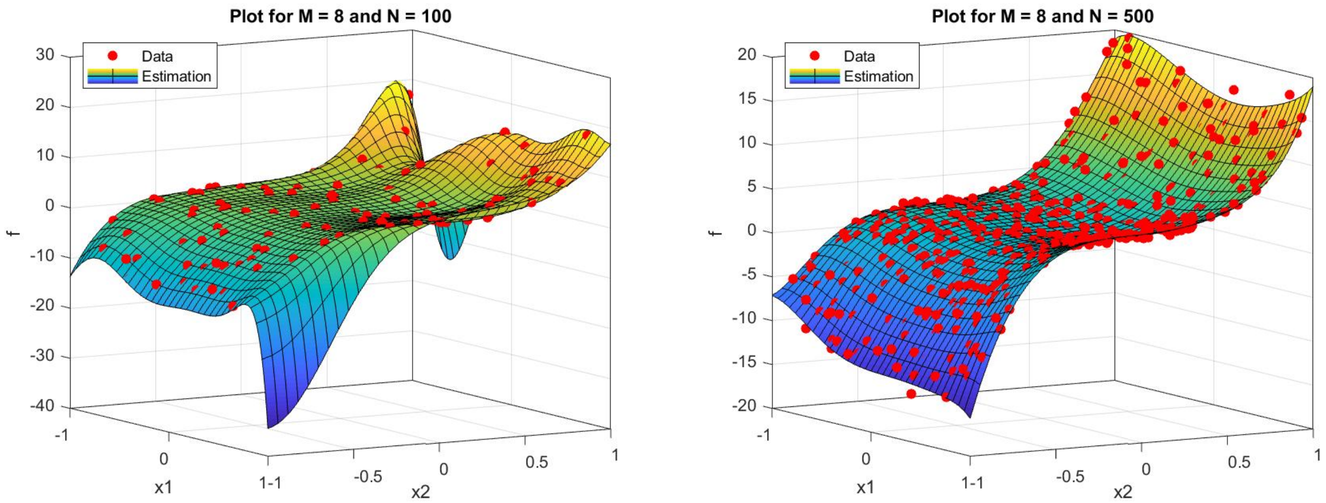

Figure 2.2: Plots for $M = 8$ Model Complexity for different Number of Data Points $N = 100$ and $N = 200$.

2.2 Ridge Regression

- Next, we fix a value of M and N and vary the Regularization Parameter of Ridge Regression. We observe that the curve that was over-fitting for Least Square case, is now controlled however if we abruptly increase the Regularization Parameter then we see that all efforts are in vain!
- We also plot the error of our model for training and development sets vs model complexity and the regularization parameter. In the plot with varying model complexity we see that at first the error decreases initially but later on due to over-fitting it starts increasing again. Similar trend is observed for varying regularization parameter however the plot shifts from over-fit to under-fit.

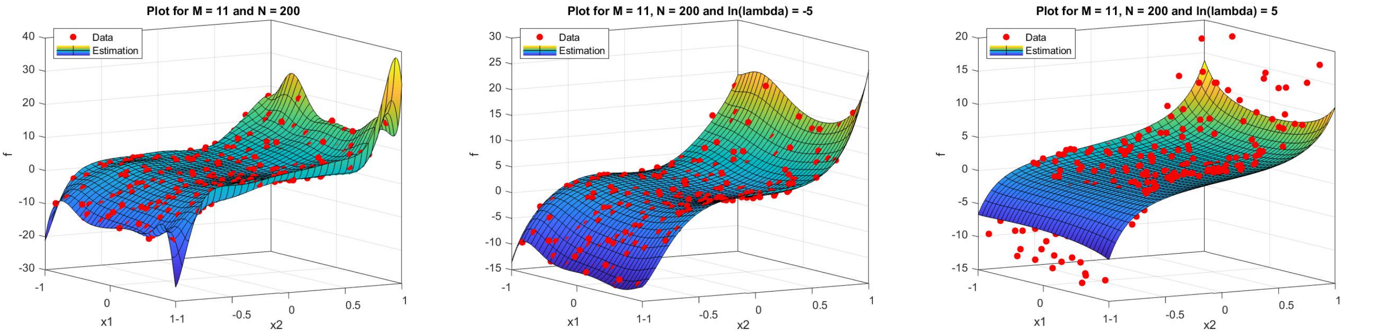

Figure 2.3: Plots for $M = 11$ Model Complexity and $N = 200$ Data Points for different values of Regularization Parameter $\ln\lambda = -\infty$, $\ln\lambda = -5$ and $\ln\lambda = 5$.

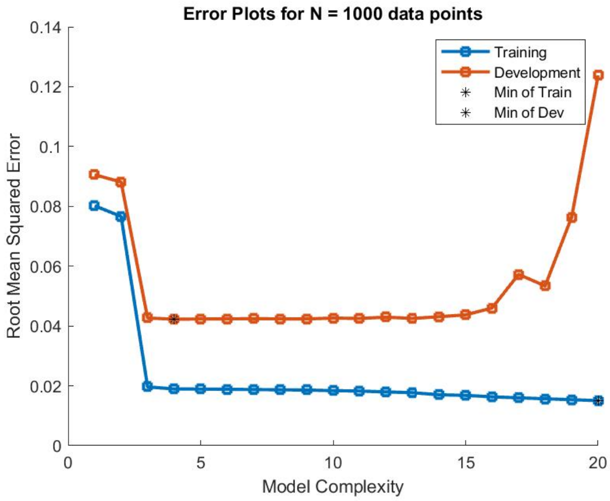

Figure 2.4: Error Plot for Training and Development Data vs Model Complexity for a fixed number of Data Points in case of Least Square Regression.

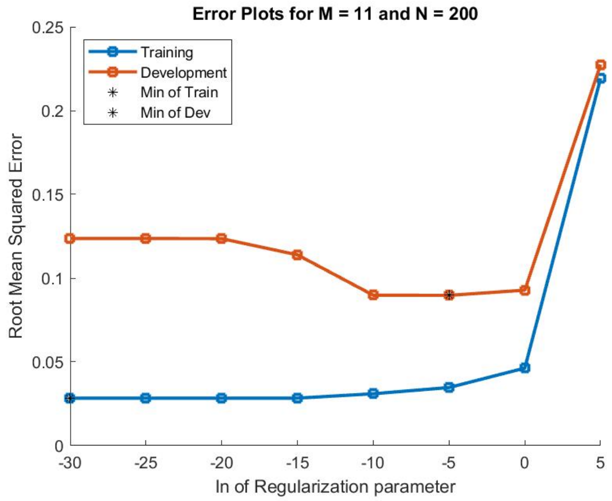

Figure 2.5: Error Plot for Training and Development Data vs Regularization Parameter for a fixed number of Data Points and Model Complexity.

2.3 Best Model

- Finally plot our Best Trained Model and the corresponding scatter plot with our model output and target output. Ideally it should be the line $x = y$, but then in a real world we see that our model performs well and is very close to the expected curve.

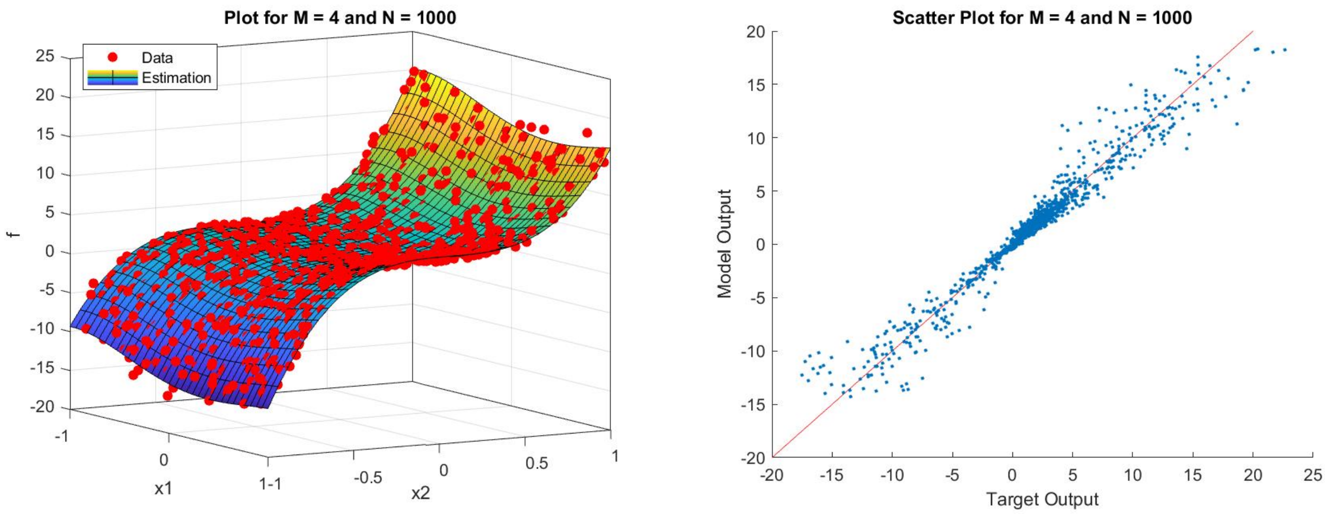

Figure 2.6: The Best Performing Model and the Scatter Plot with our Model Output on $X - axis$ and Target Output on $Y - axis$ for the Development Data Set.

CS5691: Assignment 1A

Sheth Dev Yashpal
CS17B106
August 15, 2019

1 RESULTS

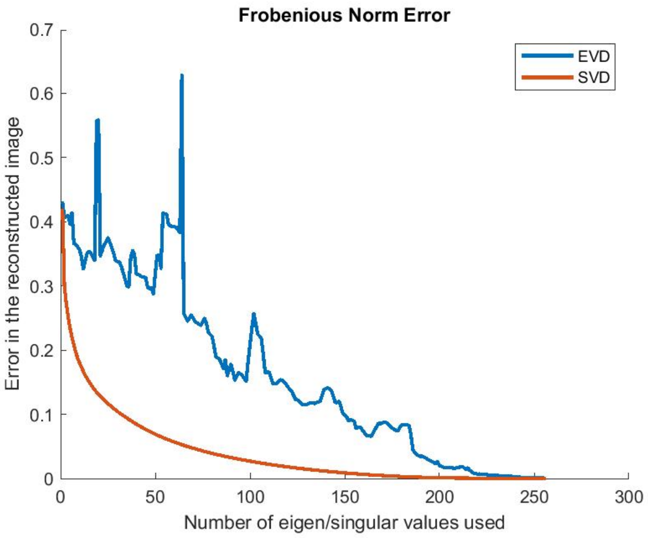

Figure 1.1: Error Plot.

Figure 1.2: Original Image (Number: 29).

2 OBSERVATIONS

- Figure 1.3 is the reconstructed image using the 25 largest eigen values and Figure 1.4 is the corresponding error image. We can observe that the image is blurred, but we can still make out the objects.
- As we increase the number of eigen values, the image becomes clearer. What we would expect is that the image gradually changes from heavily blurred to perfectly clear. But

Figure 2.1: Image for 25 eigen values.

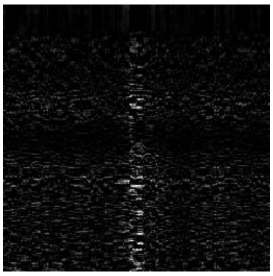

Figure 2.2: Error for 25 eigen values.

Figure 2.3: Image for 252 eigen values.

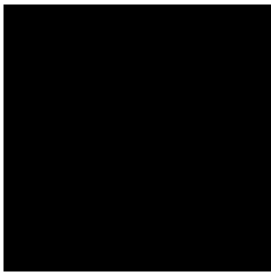

Figure 2.4: Error for 252 eigen values.

that is not the case. When we use the top 24, 32, 81, 107 or even 252 eigen values, we observe only a blurred bottom left quarter of the image as shown in Figure 1.5.

Figure 2.5: Image for 12 singular values.

Figure 2.6: Error for 12 singular values.

Figure 2.7: Image for 64 singular values.

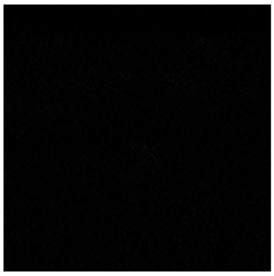

Figure 2.8: Error for 64 singular values.

- Figure 1.7 is the image obtained by using only 12 largest singular values We can clearly observe that it is much clearer than the one with 25 eigen values. Also, if we just use 64 singular values, as shown in Figure 1.9; we get a nearly perfect image with minimal error.
- From the error plot in Figure 1.1, we can observe that as we go on decreasing eigen values, the error increases more or less linearly. However, in case of singular values the error is inversely proportional to the number of singular values used.

3 Inferences

- In case of Singular values, much of the image can be represented by using the first few values as they hold much more information about the image than the rest.
- One should choose Singular Value Decomposition over Eigenvalue Decomposition because it gives better results while using less values.

CS5691: Assignment 1A

Harshit Kedia, CS17B103

August 27, 2019

1 Results

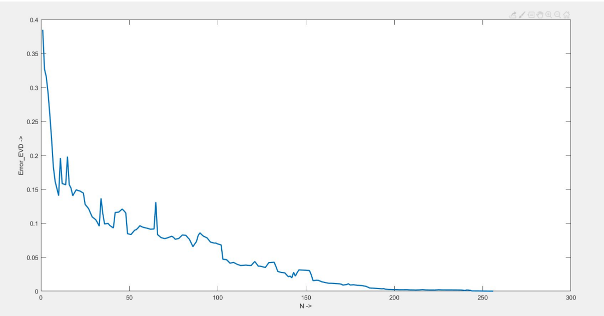

Figure 1.1: EVD Error Plot

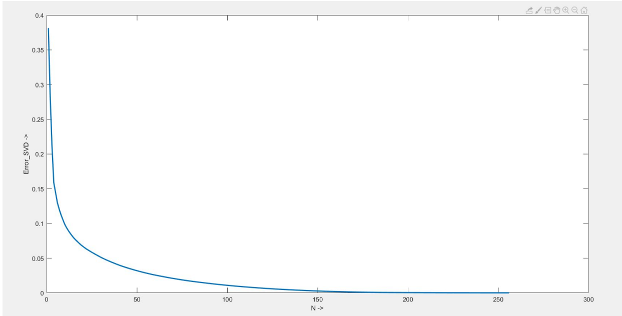

Figure 1.2: SVD Error Plot.

2 Observations

Figure 2.1: Input Image 28.jpg.

Figure 2.2: Image for 150 Eigen values.

Figure 2.3: Image for 128 singular values.

Figure 2.4: Error Image (pitch black).

- Figure 2.2 is the reconstructed image using the 150 largest eigen values and Figure 2.3 is reconstructed using the 128 largest singular valaues is the corresponding error image. The error image is very similar in both cases (full black) as seen on Figure 2.4
- Now we begin by choosing only N eigen values (biggest ones). At N=1, the image is heavily blurred (Figure 2.5) and the error image is similar to input, Figure 2.6. On increasing N, the image becomes less blurry as seen in Figure 2.7, error image Figure 2.8.

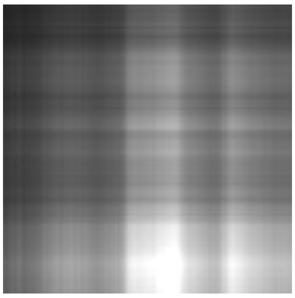

Figure 2.5: Image for 1 eigen values.

Figure 2.6: Error for 1 eigen values.

Figure 2.7: Image for 16 Eigen values.

Figure 2.8: Error for 16 Eigen values.

- Now we begin by choosing only N singular values (biggest ones). At N=2, the image is heavily blurred (Figure 2.5) but the error image is quite familiar Figure 2.6. On increasing N, the image becomes less blurry as seen in Figure 2.7, error image in Figure 2.8.

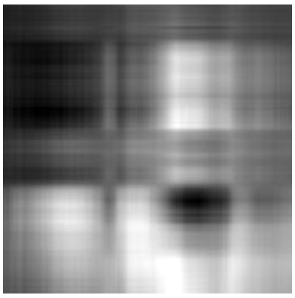

Figure 2.9: Image for 2 singular values.

Figure 2.10: Error for 2 singular values.

Figure 2.11: Image for 8 singular values.

Figure 2.12: Error for 8 singular values.

3 INFERENCES

- On choosing a higher valued N, we get an image similar to input and the error image mostly black. Output is easily identifiable by visual inspection, also error calculated is relatively very small.
- As we decrease the value chosen for N, the reconstructed image gets blurred and not identifiable at very small values of N (approximately $N < 10$), but at even lower values, the error image starts getting similar to input image.
- The SVD graph is much more smoother then the EVD graph, also the reconstructed images are better for lower number of selected values in case of SVD reconstruction.