CS754 - Course Project: Estimation of the sample covariance matrix from compressive measurements

Adit Akarsh : 19D070003 | Darin Jeff: 190260016

May 2022

1 Experiments

1. Reconstructed an image from the set using the top k eigenvectors of the true covariance matrix and the estimated covariance matrix, and measured their accuracies.
2. Evaluated the mean estimated covariance matrix after performing the experiment n times, and measured the relative MSE error from the true sample covariance matrix for different values of n.
3. Varied the compression factor $\gamma = m/s$, where m is the compressed dimension and $1/s$ is the probability of -1 or 1 in the random sampling matrix $\mathbf{R}$, and plotted the variation of reconstruction accuracy with γ.
4. Performed classification of the digits using PCA with the estimated covariance matrices for parameters $m = 98$ and $m/s = 0.67$.

Datasets

- MNIST data set

2 Results

2.1 PCA-based Reconstruction

Since we use $\hat{\Sigma}$ as an unbiased estimator of Σ, we compare how $\hat{\Sigma}$ performs against Σ for PCA-based image reconstruction (where covariance matrices are constructed for each digit). For this, we compare the reconstruction errors of the reconstructed images obtained by using the eigenvectors of Σ and $\hat{\Sigma}$ with the k largest eigenvalues. We show these results for different values of k.

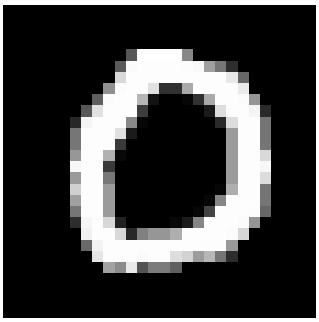

Figure 1: Original image for the class '0'

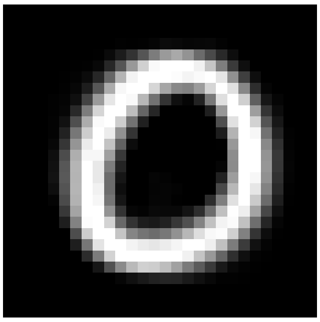

Figure 2: Reconstruction using top 10 true eigenvectors (RMSE = 0.1610)

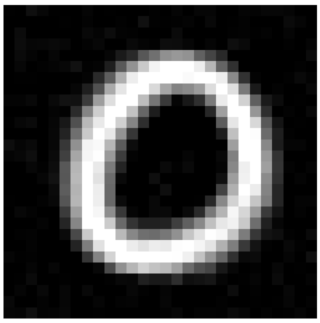

Figure 3: Reconstruction using top 10 estimated eigenvectors (RMSE = 0.1603)

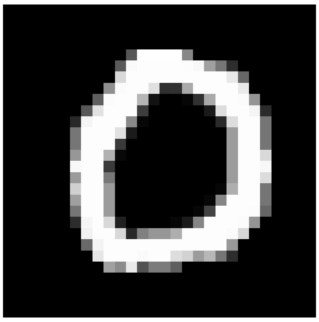

Figure 4: Original image for the class '0'

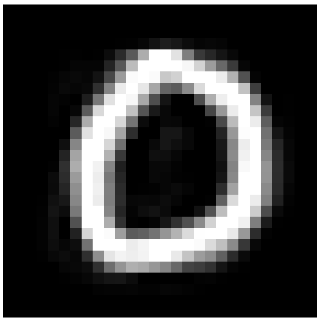

Figure 5: Reconstruction using top 50 true eigenvectors (RMSE = 0.1110)

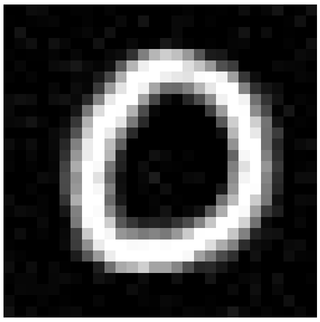

Figure 6: Reconstruction using top 50 estimated eigenvectors (RMSE = 0.1251)

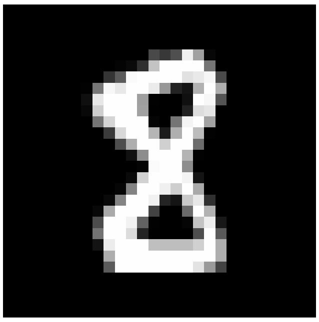

Figure 7: Original image for the class '8'

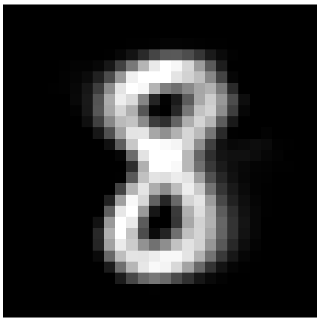

Figure 8: Reconstruction using top 10 true eigenvectors (RMSE = 0.2359)

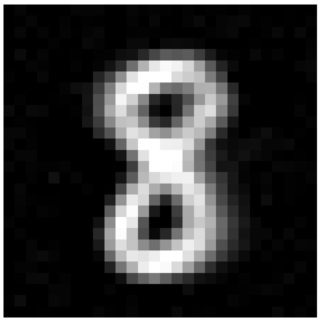

Figure 9: Reconstruction using top 10 estimated eigenvectors (RMSE = 0.2438)

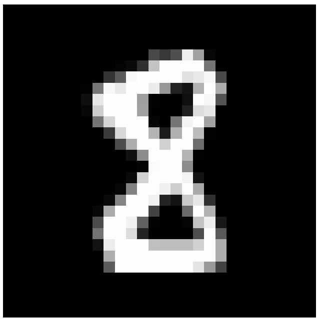

Figure 10: Original image for the class '8'

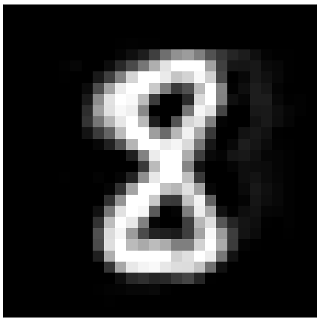

Figure 11: Reconstruction using top 50 true eigenvectors (RMSE = 0.1322)

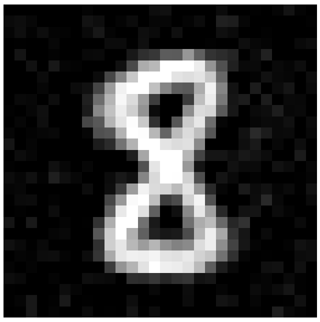

Figure 12: Reconstruction using top 50 estimated eigenvectors (RMSE = 0.1600)

2.2 Mean Square Error

Since the estimated covariance matrix, $\hat{\Sigma_n}$ is a random variable, with an expected value equal to the true sample covariance matrix, Σ_n, we expect the average estimated covariance matrix $(\hat{\Sigma}_{n,N})$ to tend to the true sample covariance matrix as the number of runs of the experiment (N) increases, from the law of large numbers. We verify this by plotting the relative mean square error between $\hat{\Sigma}_{n,N}$ and the true sample covariance matrix, for different number of runs, N, of the experiment.

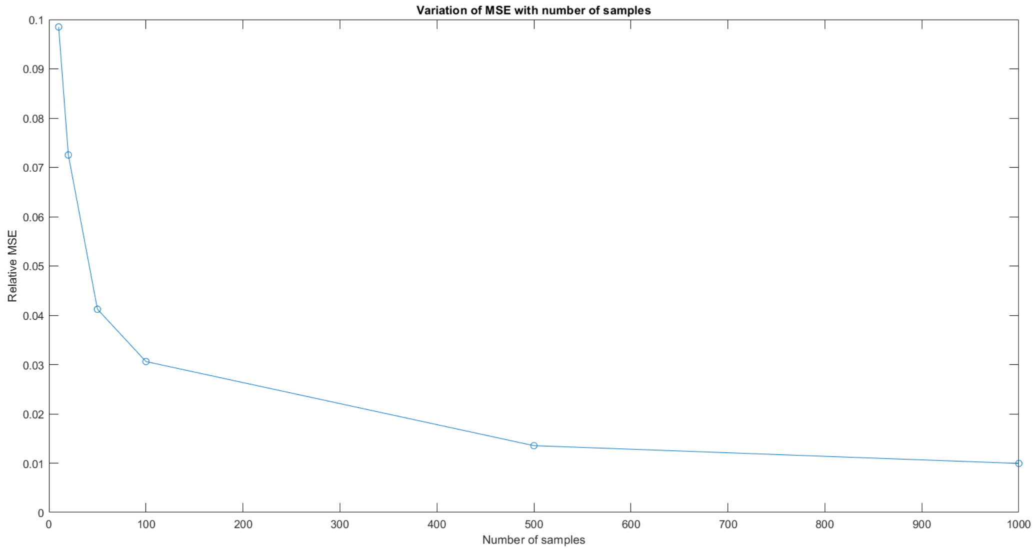

Figure 13: Variation of Relative Mean Squared Error between $\hat{\Sigma}_{n,N}$ and Σ for different values of N

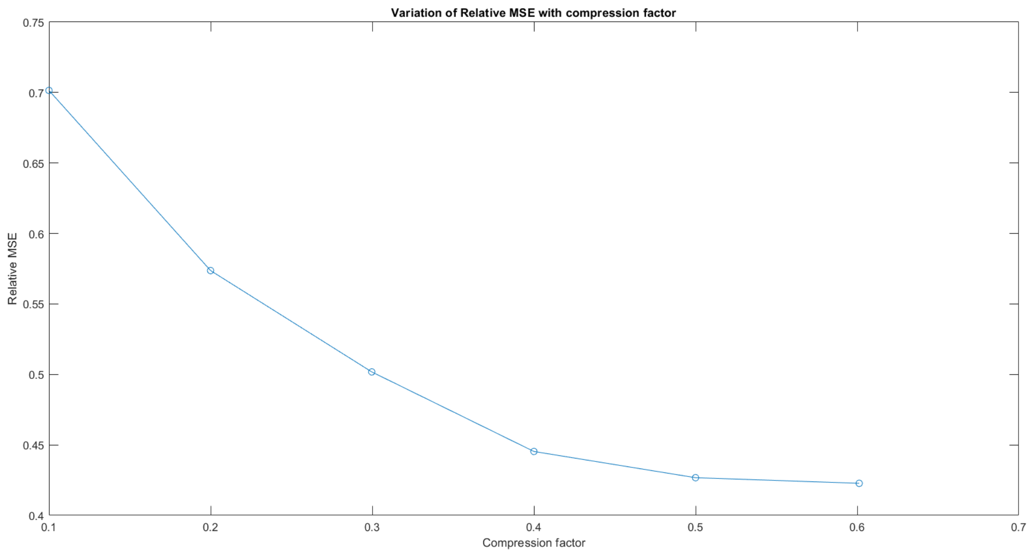

Figure 14: Variation of Relative Mean Squared Error between $\hat{\Sigma}_{n,N}$ and Σ for different values of $\gamma = m/s$

2.3 PCA-based Classification

PCA-based classification is an important image-processing technique which uses the data covariance matrix. We first generate an eigen-basis from the estimated covariance matrices and a mean-vector for each digit from '0' to '9'. We then compute the reconstruction error of the test image using the eigenvectors with the top k eigenvalues from each basis. The test image is finally classified with the digit corresponding to the eigen-basis and mean with the least reconstruction error.
Using the estimated covariance matrix, obtained from compressed samples of 1000 images of each digit from the training set, the technique used above was used to classify 10,000 images, with an accuracy of **94.07%**.

3 Conclusions

- From the experiment 2.2, we see that the estimated covariance matrix, averaged over several runs of the experiment tends to the true covariance matrix. This is in agreement with the the law of large numbers.

- The PCA reconstruction using 10 eigenvectors is comparable to the reconstruction using the true eigenvectors, but the performance is noticeably worse at 50 eigenvectors. This indicates that the eigenvectors with higher eigenvalues are estimated with greater accuracy by the estimation method.
- Classification performance is good and comparable to that using the true covariance matrix. This can be explained by the fact that the error magnitude of the covariance matrix is less than 10%, while the margin between reconstruction performances of digits in different bases is much higher.

References

[1] F. Pourkamali-Anaraki, “Estimation of the sample covariance matrix from compressive measurements,” *IET Signal Processing*, vol. 10, no. 9, pp. 1089–1095, dec 2016. [Online]. Available: https://doi.org/10.1049%2Fiet-spr.2016.0169