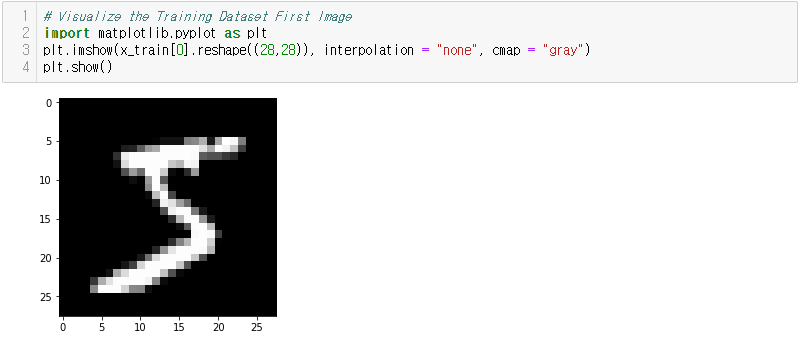
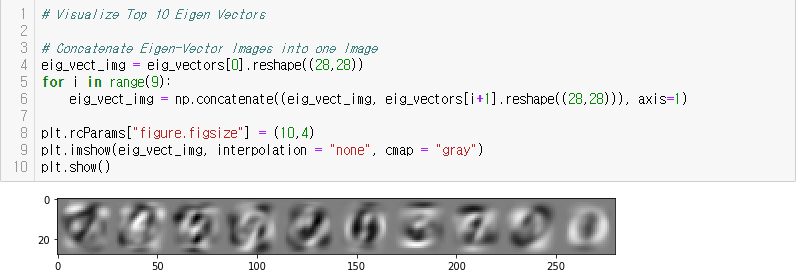
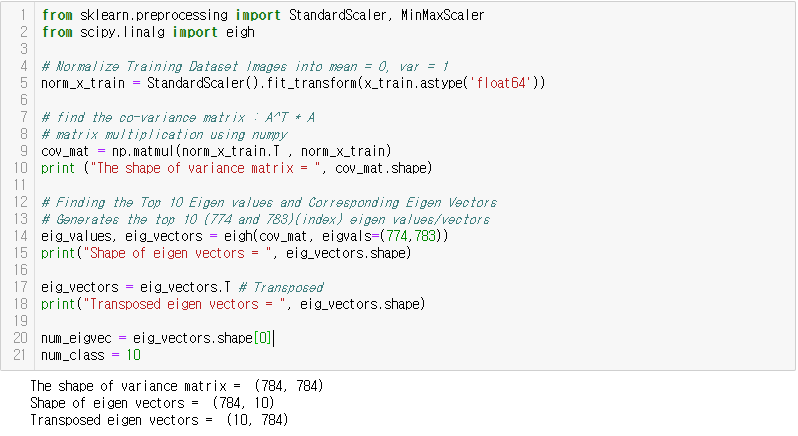
Code (Python, jupyter Notebook)

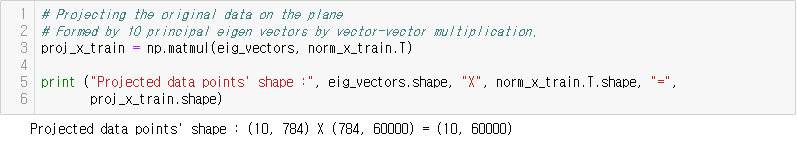
1. Dataset 불러오기



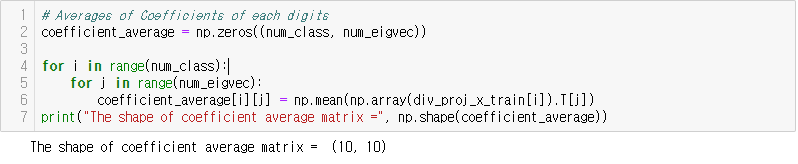
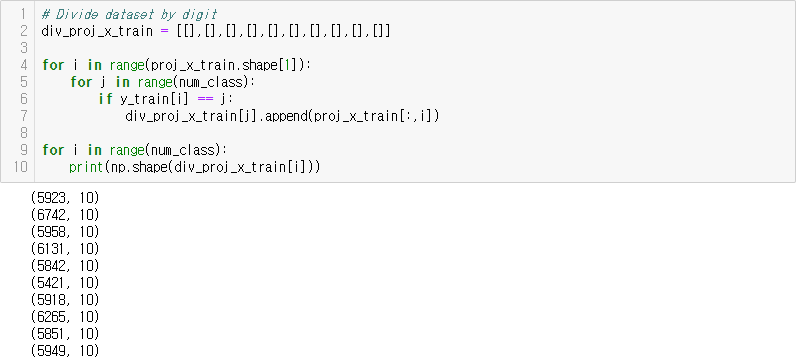
2. Eigen Vector 구하기 (Top 10 Eigen Vectors)



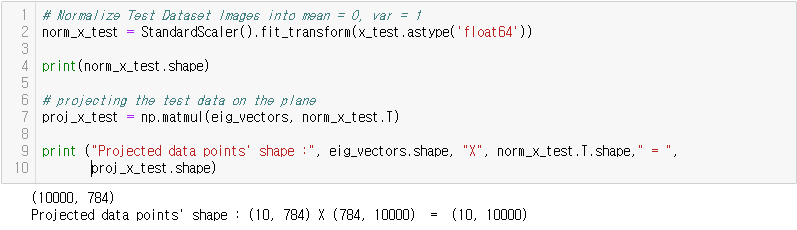
3. Training data를 Eigen Vector로 Projection



4. 각 Digit별 Coefficient 평균 구하기



5. Test data를 Eigen Vector로 Projection



6. MSE/MMSE 정의 및 Test Results



PCA로 784차원의 MNIST data를 10차원의 Eigen Vector의 Coefficient로 줄이고 10개 Digit 각각 10차원의 Coefficient 평균을 구한다.

Test Data 또한 위의 Eigen Vector를 통해 10차원 Coefficient로 바꾸고 그 Coefficient가 10개 평균 중 어느 평균과 가장 가까운지 MMSE를 통해 구한 뒤 Test Data의 해당 Label과 비교하여 분류의 정확도를 계산하였다.

Result & Discussion

Top 5 Eigen Vectors : Accuracy 0.6330

Top 10 Eigen Vectors : Accuracy 0.7291

Top 20 Eigen Vectors : Accuracy 0.7712

Top 50 Eigen Vectors : Accuracy 0.7980

Top 100 Eigen Vectors : Accuracy 0.8039

All Eigen Vectors : Accuracy 0.8089

Eigen Vector를 많이 사용할수록 분류의 정확도가 올라감을 알 수 있으나 모든 Eigen Vector를 써도 정확도가 80%밖에 되지 않는다.

Eigen Vector를 50개 이상 사용한 경우 정확도의 변화가 미미하므로 20 ~ 50개의 Eigen Vector를 사용 하는 것이 적합하다.