

ENGR 510 Intro to Machine Learning Homework

Anthony Su

November 5, 2024

Part 1

1.2

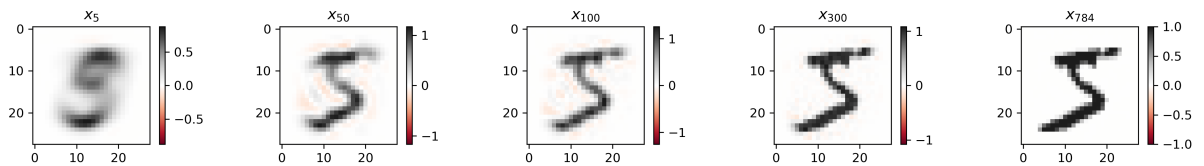


Figure 1: PCA Reconstructions of the first image in the dataset using $r = [5, 50, 100, 300]$.

Note that in Fig. 1, x_{784} is the true image. Also, note that the mean of the data was re-introduced when generating this plot.

1.3

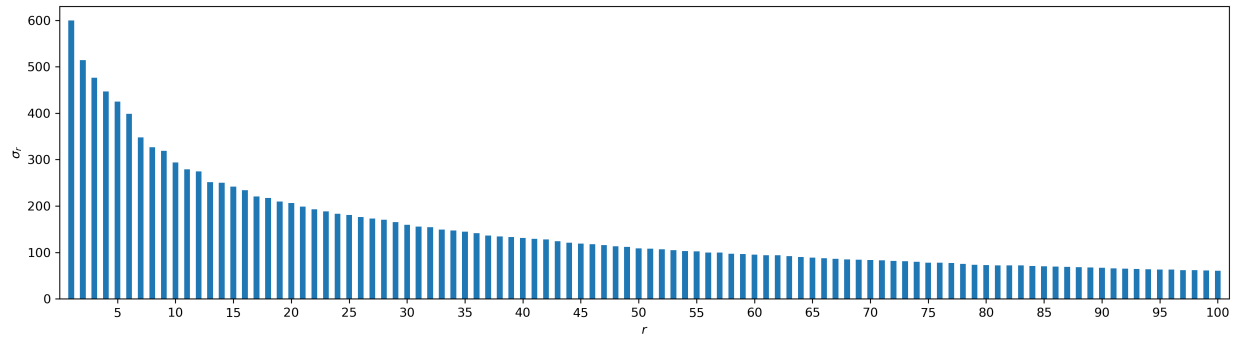


Figure 2: First 100 singular values

1.4

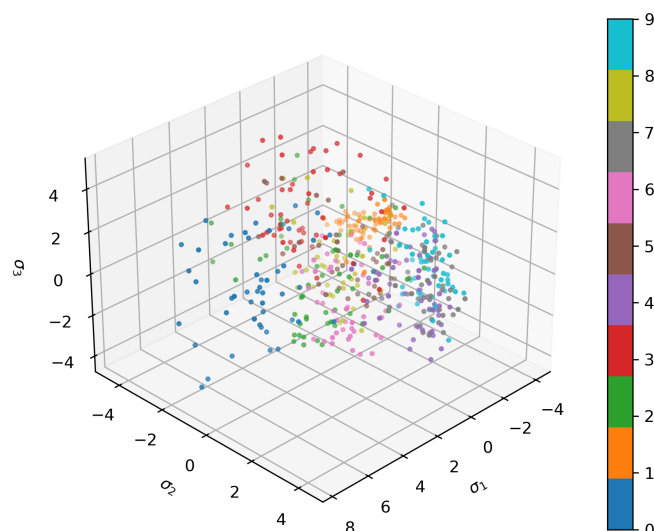


Figure 3: 500 data points in the $r = 350$ PCA space projected into the first three components

The digits 4 and 9 appear difficult to separate as they lie in the same region and have significant overlap. The digits 0 and 9 appear easy to separate as they lie in separate regions with minimal overlap.

Part 2

2.5

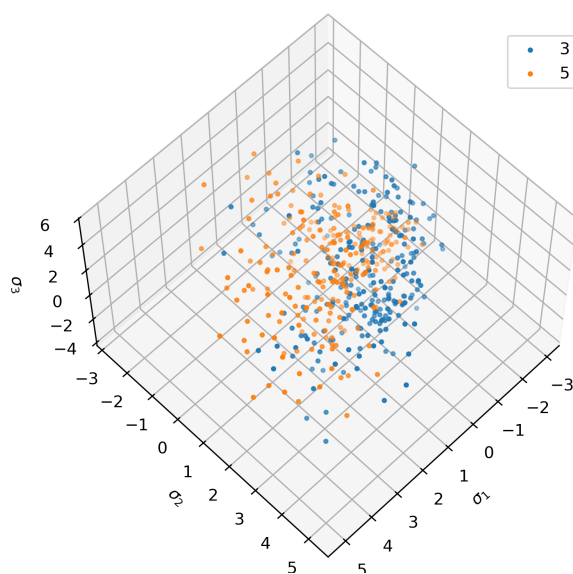


Figure 4: Hardest digit pair for LDA: 3 and 5

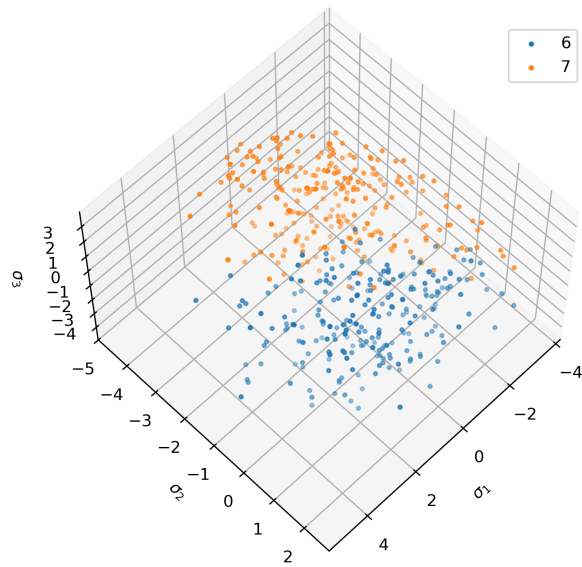


Figure 5: Easiest digit pair for LDA: 6 and 7

In Fig. 4, there is significant overlap between the regions occupied by the two digits. In contrast, in Fig. 5, there is a relatively clean separation between the regions occupied by the two digits.