HOMEWORK 8

Devansh Goenka
908 335 1354

Solution 1

Solution 1.1

The visualization of the first line from `three.txt` is:

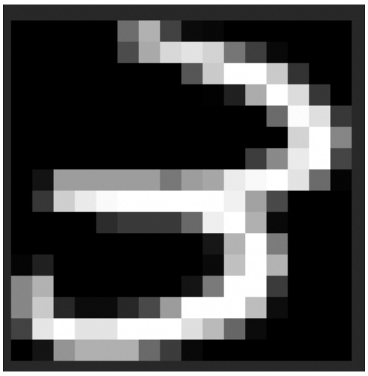

Figure 1: First sample from three.txt

The visualization of the first line from `eight.txt` is:

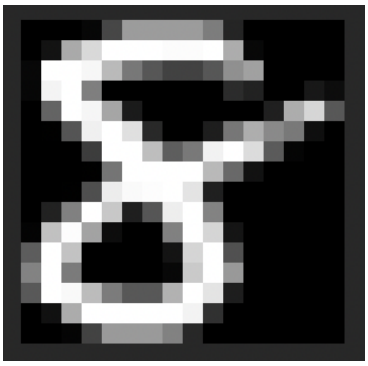

Figure 2: First sample from eight.txt

Solution 1.2

After concatenating the two datasets and computing the sample mean, here is the visualization of the mean:

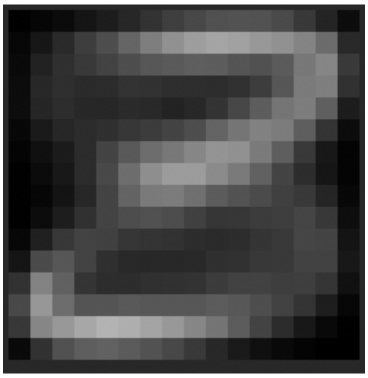

Figure 3: Sample mean of all 400 data points

Solution 1.3

After computing the empirical covariance matrix, here is the sub-matrix S[1 : 5, 1 : 5]:

$$S[1:5, 1:5] = \begin{bmatrix} 59.16 & 142.14 & 28.68 & -7.17 & 14.33 \\ 142.14 & 878.93 & 374.13 & 24.12 & -87.12 \\ 28.68 & 374.13 & 1082.90 & 555.22 & 33.72 \\ -7.17 & 24.12 & 555.22 & 1181.24 & 777.77 \\ -14.33 & -87.12 & 33.72 & 777.77 & 1429.95 \end{bmatrix}$$

Solution 1.4

After performing eigendecomposition of S, we get the following eigenvalues and eigenvectors:

Largest eigenvalue $\lambda_1 = 237155.24$
Second largest eigenvalue $\lambda_2 = 145188.35$

The eigenvectors corresponding to each of these eigenvalues are visualized below. They are first squashed into a range of $[0, 1]$ and then expanded to a range $[0, 255]$ to render these visualizations as gray scale images.

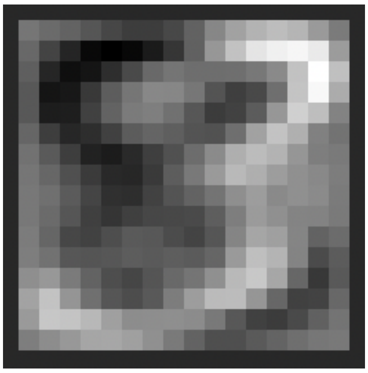

Figure 4: Eigenvector corresponding to the largest eigenvalue

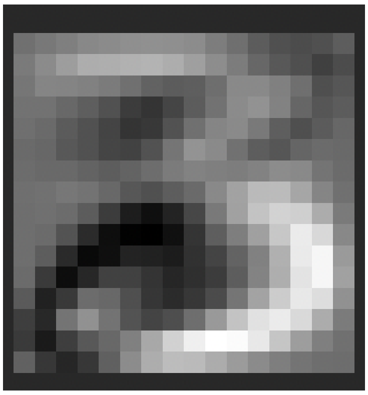

Figure 5: Eigenvector corresponding to the 2nd largest eigenvalue

Solution 1.5

After projecting the dataset X onto the 2 principal components, here are the 2D coordinates for the first line of three and eight respectively:

$$\text{First Sample from Three} = (136.2, 242.6)$$

$$\text{First Sample from Eight} = (-312.6, -649.5)$$

Solution 1.6

The plot for all 400 points in the 2D space is:

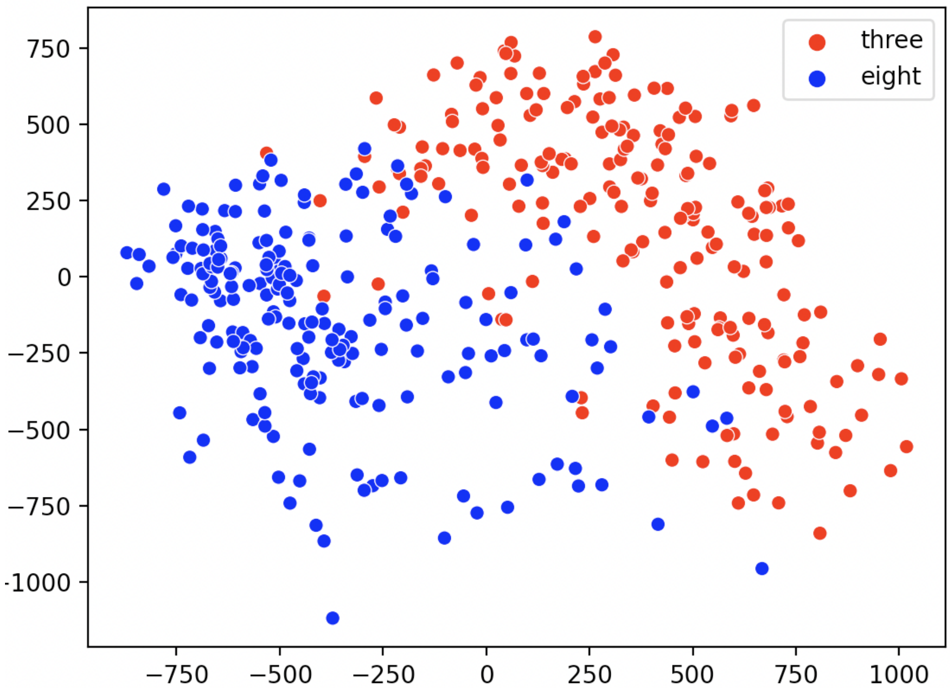

Figure 6: All samples plotted after projection onto 2 principal components

Solution 2

Solution 2.1

The Q-Table at the of 200 steps after following the mentioned deterministic greedy policy is:

State \ Action	Stay	Move
A	0	0
B	0	0

We note that all entries in the Q-Table remain at 0 because when in this case, the tie-breaker action is "Move". When the initial Q-Table has all 0's, the action "Move" is preferred with reward 0, and hence the entries keep on being 0.

If we modify the policy to prefer "Stay" instead, here is the induced Q-Table.

State \ Action	Stay	Move
A	4.999	0
B	0	0

Solution 2.2

Instead of the deterministic greedy policy, we use the ϵ-greedy policy with $\epsilon = 0.5$. Here is the Q-Table after 200 steps:

State \ Action	Stay	Move
A	4.97	3.99
B	4.99	3.94

Solution 2.3

Here, we use a optimal policy to obtain the Q-Table.
Let's assume the true function is Q^* that we get by using the optimal policy.

We then get the recurring Bellman equation as:

$$Q^*(A, Stay) = R(A, Stay) + \gamma \max_{a'} Q^*(A, a')$$
$$Q^*(A, Move) = R(A, Move) + \gamma \max_{a'} Q^*(A, a')$$
$$Q^*(B, Stay) = R(B, Stay) + \gamma \max_{a'} Q^*(B, a')$$
$$Q^*(B, Move) = R(B, Move) + \gamma \max_{a'} Q^*(A, a')$$

We note that the states are symmetric and hence, we have :

$$\forall a \in \{\text{Stay}, \text{Move}\}, Q(A, a) = Q(B, a)$$

Hence, we have:

$$Q^*(S, Stay) = 1 + \gamma Q^*(S, Stay)$$
$$Q^*(S, Move) = 0 + \gamma Q^*(S, Stay)$$

We have $\gamma = 0.8$, and hence the above series converges.

We get:

$$Q^*(S, Stay) = \frac{1}{1-\gamma} = \frac{1}{0.2} = 5$$

$$Q^*(S, Move) = \frac{\gamma}{1-\gamma} = \frac{0.8}{0.2} = 4$$

Thus, we have the final Q-Table as:

State \ Action	Stay	Move
A	5	4
B	5	4