

9.641 Neural Networks

Problem Set 8: PCA, k-means and Ring network

(Due before class on Thursday, Apr. 21)

1. Double-ring network model of a neural integrator.

Consider a network with the following dynamics:

$$\begin{aligned}\dot{\mathbf{x}}_r + \mathbf{x}_r &= [\mathbf{W}^r \mathbf{x}_r + \mathbf{W}^l \mathbf{x}_l + b_r]^+ \\ \dot{\mathbf{x}}_l + \mathbf{x}_l &= [\mathbf{W}^r \mathbf{x}_r + \mathbf{W}^l \mathbf{x}_l + b_l]^+\end{aligned}$$

where \mathbf{x}_r and \mathbf{x}_l are vectors representing the neural activities of the right and left rings respectively. The synaptic interaction matrices \mathbf{W}^r and \mathbf{W}^l are defined as:

$$\begin{aligned}W_{ij}^r &= J_0 + J_2 \cos(2\pi(i-j)/N - \phi) \\ W_{ij}^l &= J_0 + J_2 \cos(2\pi(i-j)/N + \phi)\end{aligned}$$

Simulate this network in MATLAB using $N = 100$, $J_0 = -9.8/N$, $J_2 = 10/N$, and $\phi = \pi/6$. Plot the activity of the network using two bar graphs, one for \mathbf{x}_r and another for \mathbf{x}_l . Show the simulation results for the following two cases:

- Take $b_r = b_l = 1$.
- Take $b_r = 1$ and vary b_l . You should see a travelling pulse whenever $b_l \neq b_r$. Plot the relationship between the velocity of the travelling pulse and the input b_l .

2. Principal component analysis using the CBCL face dataset.

Download the subset of the CBCL face dataset (`faces.mat`). The faces are the columns of the matrix. (You can try to visualize faces by reshaping the columns of the matrix to be 19x19 with the `reshape` command.)

- Use the `cov` command to compute the covariance of all the faces in `faces.mat`. Then use the `eig` command to find the eigenvectors and eigenvalues. Sort the eigenvectors by decreasing eigenvalues and display all eigenvectors as images (you can use the command `montage`). Describe what the top eigenvectors are coding for. What can you say about the eigenvectors as the eigenvalues decrease?
- Using the `eig` command is inefficient because it finds all the eigenvectors instead of just the ones that you want. Show that iterating

$$x := \frac{Cx}{|Cx|}$$

converges to the principal eigenvector from almost all initial conditions. Simulate it numerically and verify that the result is the same as the principal eigenvector obtained using `eig`.

3. Clustering with K-means.

Download the matlab file called `kmeans.mat`. The data to cluster is contained in the vector `X` and the true data label in the vector `label`.

- Implement the kmeans algorithm (do not use the built-in matlab function) and run it on the data with $k = 3$. How similar is the result of your algorithm to the true labels? What happens if you set $k = 2, 4, 5$? Submit your code and relevant plots.

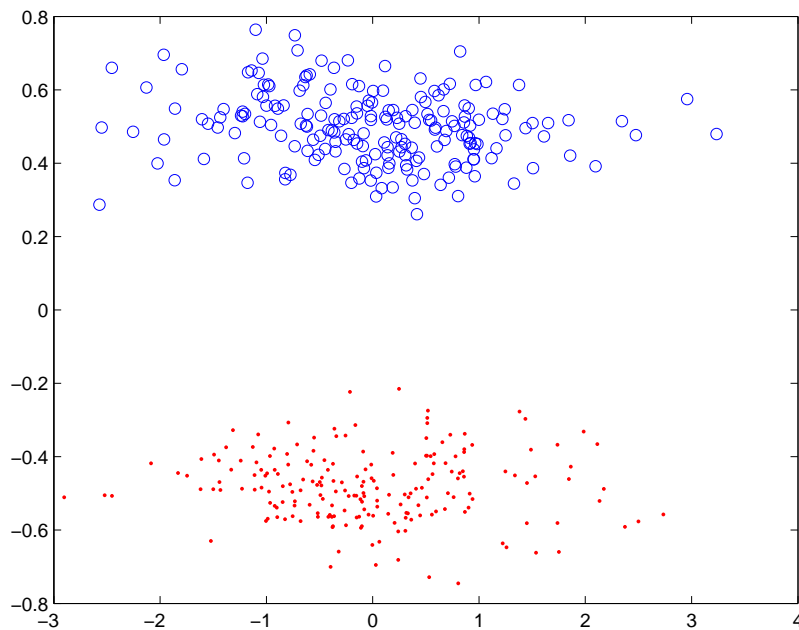


Figure 1: Data to be clustered

- (b) Consider Fig. 1. What would happen if you were to run kmeans on this dataset? Explain your answer.
(c) Give examples of datasets where kmeans has multiple local minima.

Hint: Consider a toy example of three points.

4. Some statistics

- (a) For two random variables x and y , prove that the magnitude of the correlation is upper bounded by the square root of the product of the second moments,

$$|\langle xy \rangle| \leq \sqrt{\langle x^2 \rangle} \sqrt{\langle y^2 \rangle}$$

Hint: Use the Cauchy-Schwarz inequality.

- (b) Pearson's famous correlation coefficient r is defined as the ratio of the covariance to the product of the standard deviations. Show that the above inequality implies that

$$-1 \leq r \leq 1$$

Under what conditions is $r = 1$ or $r = -1$ exactly? The correlation coefficient assesses the quality of a least squares fit of a linear model relating x and y .