

Stats 312, Exercise 6

Due Mar. 10th

In the sixth assignment we explore spike trains from a multi-electrode array. This assignment is a modified version (with permission) of Prof. Byron Yu from CMU, and the data comes from Krishna Shenoy's lab in Stanford.

The data can be found in `/stat312/dataset_2/ps3_realdata.mat`.

Some manipulations for this data set can be found in the script I uploaded to Coursework.

Classification

- I posted the code for classification using the naïve Bayes algorithm assuming marginal Poissons. Please implement (1) a Gaussian generative model, and (2) a nearest-neighbor model. For the Gaussian model, try both shared and separate covariances. For the nearest model, decide on the value of k , and discuss your choice of distance function.
- For all algorithms, compare the prediction on the test set and the training set.
- Compare the confusion matrices of the two algorithms; are the same errors being made? Look at one wrong case from each algorithm, and see if you can understand what caused the mistake. Were you able to observe this is a mistake?

Clustering

- Run a clustering algorithm; Do the neurons cluster into several groups? Is there a hierarchy between the groups?
- Build the distance matrix between neuronal responses. What are the main directions for the matrix? Can you explain the matrix in terms of the classes of trials?
- Find a visual display to convey what you have learned.
- Display a neuron-by-trial matrix. Think about the sorting of column and rows. What can you see?