Problem Set V – Model Fitting Part I

1. Let's consider the Linear-Poisson model. Assuming that the firing rate at time i is the product of the stimulus at time i and a parameter ϕ, and that spikes follow a Poisson noise model, formulate the likelihood expression for the encoding model. Now, find the value of ϕ that maximizes likelihood (follow the procedure outlined on the relevant slide in the lecture).

2. Let's say you record the firing rate of one neuron as a monkey performs a reaching task, sort of like this picture:
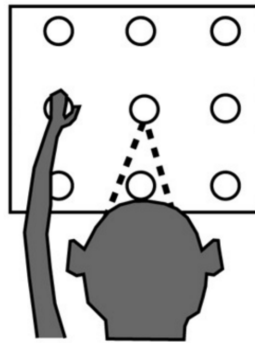
In this task, the animal's hand starts in the center, then reaches to a cued location (the circles around the edge of the square in the picture), and then returns to the center. The cued location is chosen pseudo-randomly, and each of the 8 points is visited 100 times in total over the course of the experiment. For each reach, you record both the average firing rate during the reach, and the reach angle (in other words, you have 2 numbers for each trial).

The data for the neuron looks something like this:
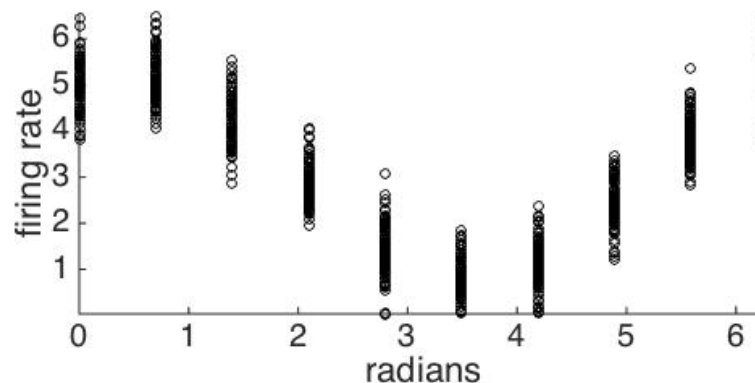
Call Y the vector of all firing rates (100 trials * 8 locations = 800 elements in Y), and S the vector of reach angles. Now, you want to fit a parametric function that describes your data. Clearly, the firing rate changes with the reach angle, so you must think of some function that captures this relationship. You decide a good function to try is $\widehat{y_i} = a * \cos(s_i) + b *$

$\sin(s_i) + c$, where $\hat{y}_i$ is our estimate of the firing rate for a reach i, s_i is the reach angle for reach i, and a, b, and c are parameters that you can change to find the best fit to data. Now, you have to decide on a noise model.

Part 1: Of the noise models we discussed in class, which one best fits the data here? Why?

Part 2: Precisely define the function you want to maximize (or minimize). Either solve to get an analytic expression for the parameters, or write down the expression that you would plug into MATLAB or Python (write it down in the correct syntax) to minimize.

Part 3: Once we choose the parameters, we should report our model fit. Let's say we choose mean squared error ($MSE = \frac{1}{M} \sum_{i=1}^{M}(y_i - \hat{y}_i)^2$, where y_i is the actual firing rate and $\hat{y}_i$ is our estimate of the firing rate) as our quantity to report. Being the rigorous scientists that we are, we want to report our model fit on held-out data. How should we segment our data to make this possible? Describe how you would split this data (800 trials total) into training and test data sets. Do you expect MSE to be higher on the test data set or the training data set?

Part 4: Let's say you decided you wanted to add a penalty on the magnitude of the coefficients (say in one of the ways discussed in class). Do you expect MSE on the testing data set to increase or decrease?