MA 568 – Statistical Analysis of Point Process Data

Problem Set #4
Due November 22, 2016

In class, we discussed the construction of state-space estimation algorithms based on point process observations. In this problem, we will apply the approximate Gaussian filter to the problem of reconstructing a wind stimulus from an ensemble of simulated neurons from the cricket cercal system.

Crickets have pairs of specialized sensory structures on their lower abdomen called cerci that are used to detect air currents such as those generated by rapidly approaching predators. Each of these cerci is covered with up to several thousand hairs of various sizes and shapes that can be deflected in one direction. Each hair grows from a single sensory neuron and each neuron is tuned to respond to different air current dynamics. This system is unique in that the sensory neuron closest to the periphery already fires spikes. This makes it well-suited for studying how neural representations can arise.

It has been shown that for small magnitude wind stimuli, a good model for the firing activity of the neurons closest to the periphery is a Poisson process with rate function $\lambda(t) = \exp\{\alpha + \beta v(t)\}$, where $v(t)$ is the wind speed in the direction to which the neuron responds, and α and β are model parameters for that cell.

Please download the file **`CricketData.mat`** from the course website. Loading this file into MATLAB will produce the following data structures:

- **`trainingStim`** - A 1 second wind stimulus signal for the training period.
- **`trainingSpikes`** - The spiking activity for each of 10 neurons during the training period
- **`testStim`** - The wind stimulus during the test period to be decoded
- **`testSpikes`** - The spiking activity during the test period.

Each signal is recorded at a sampling rate of 1 KHz.

1. Construct a linear Gaussian state model for the wind stimulus, $v_k = Av_{k-1} + \varepsilon_{k-1}$, with $\varepsilon_{k-1} \sim N(0, \sigma^2)$, using the data in **`trainingStim`**. Begin by plotting the stimulus value at each time against its value at the previous time step. Recall that the linear regression estimate for the equation $y = Ax + \varepsilon$ is given by $\hat{A} = \sum(x_i y_i) / \sum x_i^2$. Use linear regression to find the optimal estimate of the parameter A. Plot a histogram of the residuals $v_k - \hat{A}v_{k-1}$, and estimate the state transition variance, $\hat{\sigma}^2$.

2. Estimate the α and β parameters for each of the ten neurons from the **`trainingStim`** data using the MATLAB **`glmfit`** function. Are all of these parameters significant? Construct a KS Plot for one of these neurons. How well does this model fit the observed spiking data?