Problem Set 3

Due November 26, 2023

There are three problem sets, which cover the basic concepts we have introduced in the course on sensory coding. Please select any two problem sets. (*) indicates that this part has not been covered by our lectures. But solving it would give you extra credit. If you are willing to solve all three problem sets, that would be great, since the TF will give you *super extra* credit.

Analyze Motion Detection Neuron in Fly

The MATLAB file H1.mat contains data collected and provided by Rob de Ruyter van Steveninck from a fly H1 neuron responding to an approximate white-noise visual motion stimulus. Data were collected for 20 minutes at a sampling rate of 500 Hz. In the file, *rho* is a vector that gives the sequence of spiking events or nonevents at the sampled times (every 2 ms). When an element of rho is one, this indicates the presence of a spike at the corresponding time, whereas a zero value indicates no spike. The variable *stim* gives the sequence of stimulus values at the sampled times.

(a) Calculate and plot the spike-triggered average from these data over the range from 0 to 300 ms (150 time steps).

(b) Use the spike-triggered average to construct a linear kernel, as we discussed in the class, to provide a model of the response of the H1 neuron. Choose r_0 so that the average firing rate predicted by the model in response to the stimulus used for the data matches the actual average firing rate.

(c) Use a Poisson generator with the computed rate to generate a synthetic spike train from this linear estimate of the firing rate in response to the stimulus. Plot examples of the actual and synthetic spike trains. How are they similar and how do they differ?

(d) Plot the autocorrelation function of the actual and the synthetic spike trains over the range 0 to 100 ms. Why is there a dip at a lag of 2 ms in the autocorrelation of the actual spike train? Is there a dip for the synthetic train too? Plot the interspike interval histogram for both spike trains. Why is there a dip below 6 ms in the histogram for the actual spike train? What are the coefficents of variation (see lecture note) for the two spike trains and why might they differ?

Maximization of Entropy under Constraints

The Entropy of a variable X drawn from a distribution $p(X)$ is given by the following formula

$$H(X) = - \int dX p(X) \ln p(X) \tag{1}$$

Use the Lagrange Multiplier method to evaluate the maximum entropy probability distribution $p(X)$ in the following cases:

(a) X is a one dimensional continuous random variable, which takes only positive values and its maximum value is fixed. Hint: you should also take into account the constraint imposed by the normalization of p.

(b) X is a one dimensional continuous random variable, which takes only positive values and its mean is fixed.

(c) There is no constraint on the range of X but its variance is given.

(d) X is an N-dimensional continuous random variable with constraint on the total variance,

$$\sum_i^N \langle x_i^2 \rangle = N \sigma^2 \tag{2}$$

(e*) Show that the entropy of the multivariate Gaussian $N(\mathbf{X}|\mu, \mathbf{\Sigma})$ is given by

$$H(\mathbf{X}) = \frac{1}{2} \ln |\mathbf{\Sigma}| + \frac{D}{2}(1 + \ln(2\pi)) \tag{3}$$

where D is the dimensionality of $\mathbf{X}$, $|\mathbf{\Sigma}|$ is the determinant of the covariance matrix $\mathbf{\Sigma}$.

Fisher Information and Mutual Information

Consider a group of N cells responding to a stimulus x. The response of each neuron i is given by $r_i = wx + z_i$ where z_i varies from trial to trial as a gaussian random variable with zero mean and variance σ_i^2. The z_i are uncorrelated: $\langle z_i z_j \rangle = \delta_{i,j}$. The stimulus x is drawn from a gaussian distribution with zero mean and variance σ_0.

(a) Compute the Fisher Information of the system.

(b) Recall what I discussed in the class about the Maximum Likelihood (ML) estimate

$$\underset{x}{\operatorname{argmax}} \, p(r|x).$$

Give an expression for the ML estimator of $\hat{x}$. Evaluate the mean square error of the estimate $\langle (x - \hat{x})^2 \rangle$, and compare with the result of (a).

(c*) Instead of using ML estimate, now let us consider the Bayes rule:

$$p(x|r) = \frac{p(r|x)p(x)}{p(r)}. \tag{4}$$

We would like to find a Bayesian estimator $\hat{x}$ that would maximize the posterior distribution $p(x|r)$. Evaluate its mean square error and explain its dependence on σ_0. Compare your results with (b) and Under what conditions you expect that the two estimators will differ significantly? Explain the relationship between the results and the Cramer-Rau Bound.

(d) Compute the Mutual Information between the neuronal responses and the stimulus. What is the relationship between your result and the Fisher Information? Assume now that all $\sigma_i = \sigma$. How do both quantities depend on the population size?

Guide: In the above, the mean square error average $\langle ... \rangle$ is defined as an average over both neuronal noise and stimulus values.