# Problem Set 4

Due Wednesday, Jan 6, 2021

## Linear-nonlinear model

The optimal kernel derived from a linear model has two problems. First, there is nothing to prevent the predicted response to become negative. Neuron is a nonlinear device. The response of a neuron is typically characterized by the firing rate. This number *cannot* be negative. Moreover, in a linear model, the predicted response does not saturate. As the magnitude of the response increases, the response would also increase without bound. If we use $L$ to represent the linear term we have been discussing thus far:

$$L(t) = \int_0^\infty d\tau D(\tau) s(t-\tau) \tag{1}$$

The modification is to replace the linear prediction $R_{est}(t) = R_0 + L(t)$ with the generalization

$$r_{est}(t) = r_0 + F[L(t)] \tag{2}$$

For example, one can choose $F(x) = [x]^+ = \mathbf{max}(x, 0)$, one can also set an upper bound so that the function saturates for large $x$, i.e., $F(x) = \mathbf{min}(x, F_0)$.

However, when nonlinearity is added, there is no guarantee that the derived kernel is optimal. A self-consistent solution for the optimal kernel should satisfy

$$\begin{aligned} D(\tau) = \frac{Q_{rs}(-\tau)}{\sigma^2} &= \frac{1}{\sigma^2 T}\int_0^T r(t)s(t-\tau)dt \\ &\approx \frac{1}{\sigma^2 T}\int_0^T r_{est}(t)s(t-\tau)dt \\ &= \frac{1}{\sigma^2 T}\int_0^T F[L(t)]s(t-\tau)dt \end{aligned} \tag{3}$$

In general, the above equation does not hold. There is one exception. if the stimulus is Gaussian white noise, please show that the expected value of the integral satisfies

$$\frac{1}{\sigma^2 T}\int_0^T F[L(t)]s(t-\tau)dt = \frac{D(\tau)}{T}\int_0^T dt \frac{dF(L(t))}{dL} \tag{4}$$

The integral on the right hand side is a normalization condition. By properly scaling $F$, we can make $\frac{1}{T}\int_0^T dt \frac{dF(L(t))}{dL} = 1$.

**Hint:**

(1) For a Gaussian random variable $x$ with zero mean and standard deviation $\sigma$, prove using integration by part that

$$\langle xF(\alpha x)\rangle = \alpha\sigma^2\langle F'(\alpha x)\rangle, \tag{5}$$

where $F$ is any function, $\alpha$ is a constant, and $\langle ...\rangle$ denotes the gaussian weighted average (or expected value),

$$\langle g(x)\rangle = \int dx g(x) \frac{1}{\sqrt{2\pi\sigma^2}} \exp(\frac{-x^2}{2\sigma^2}). \tag{6}$$

Extend your argument to multivariate functions and then to functional. For those who do not understand functional, please read my lecture notes carefully.

(2) Gaussian white noise has a gaussian probability density, with zero mean and variance $\sigma^2/\Delta t$, where $\Delta t$ is the size of the time window.

# Entropy and Mutual Information

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

$$H(X) = \int p(X) \ln p(X) \tag{7}$$

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

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

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

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

$$\sum_i^N \langle x_i^2\rangle = N\sigma^2 \tag{8}$$

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

$$H(\mathbf{X}) = \frac{1}{2}\ln|\mathbf{\Sigma}| + \frac{D}{2}(1 + \ln(2\pi)) \tag{9}$$

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

## Hassentein-Reichardt correlator

Please take a close look at the Reichardt detector model I have discussed in the class. For a grading stimulus with defined spatial frequency ($k = 2\pi/\lambda$) and temporal frequency $\omega_0$, the light intensity signal received by two neighboring channels (i.e., two photoreceptors) have the following form:

$$\begin{aligned} s_1(t) = \Delta I \sin(\omega_0 t) = \text{Im}\left[\Delta I e^{i\omega_0 t}\right]; \\ s_2(t) = \Delta I \sin(\omega_0 t - k\Delta x) = \text{Im}\left[\Delta I e^{i(\omega_0 t - k\Delta x)}\right]. \end{aligned} \tag{10}$$

In the simplest model, we can think that the response of a neuron is a low passed filter of the sensory input with some Kernel $D_1(t)$ and $D_2(t)$. As a result, the response function might be written as

$$\begin{aligned} r_1(t) &= \int_{-\infty}^{\infty} s_1(t-\tau)D_1(\tau)d\tau; \\ r_2(t) &= \int_{-\infty}^{\infty} s_2(t-\tau)D_2(\tau)d\tau. \end{aligned} \tag{11}$$

Similar responses could be written down for $r_3(t)$ and $r_4(t)$. The motion detection output signal is defined as

$$R(t) = r_1(t)r_2(t) - r_3(t)r_4(t) \tag{12}$$

And the steady state solution $\langle R \rangle_t$ is given by averaging over the time period $2\pi/\omega_0$.

- Show that

  $$\langle R \rangle_t = \|\tilde{D_1}(\omega_0)\|\|\tilde{D_2}(\omega_0)\| \sin[\phi_1(\omega_0) - \phi_2(\omega_0)]\Delta I^2 \sin(k\Delta x), \tag{13}$$

  where the fourier transform of the kernels are defined as

  $$\tilde{D_1}(\omega_0) = \|\tilde{D_1}(\omega_0)\| e^{i\phi_1(\omega_0)},$$

  $$\tilde{D_2}(\omega_0) = \|\tilde{D_1}(\omega_0)\| e^{i\phi_2(\omega_0)},$$

  If you cannot solve this problem, use this result and solve the next two questions.

- Consider a simple kernel $D_1(t) = \frac{1}{\tau}\exp(-t/\tau)$, and $D_2(t) = \delta(t)$, we find $\tilde{D_1}(\omega_0) = \frac{1}{1+i\omega_0\tau}$ , and show that

  $$\langle R \rangle \sim \frac{\omega_0 \tau}{\omega_0^2 \tau^2 + 1} \tag{14}$$

  This function has a maximum when $\omega_0 = 1/\tau$.

- If the filters on both arms are first-order low-pass, so that $D_1(t) = \frac{1}{\tau}\exp(-t/\tau_1)$, $D_2(t) = \frac{1}{\tau_2}\exp(-t/\tau_2)$, show that the steady state response is given by

$$\langle R \rangle \sim \frac{\omega_0(\tau_2 - \tau_1)}{(1+\omega_0^2\tau_1^2)(1+\omega_0^2\tau_2^2)} \tag{15}$$

*Hint*: The analytical form of $r_1(t)$ and $r_2(t)$ can be computed by taking the fourier transform of the convolution, and then performing an inverse fourier transform. As a first step:

$$\begin{aligned} \tilde{r_1}(\omega) &= \sqrt{2\pi}\delta(\omega-\omega_0)\tilde{D}_1(\omega); \\ \tilde{r_2}(\omega) &= e^{-ik\Delta x}\sqrt{2\pi}\delta(\omega-\omega_0)\tilde{D}_2(\omega) \end{aligned} \tag{16}$$