Sensory coding

Quan Wen

04 Dec 2019

Kernel method

In general, the response of a neuron or an animal at time t can be viewed as a functional r of the sensory stimuli s, where s itself is also a function. Functional r is a scalar, and its value at time t, that is $r(t)$, depends on the entire stimulus history of $s(t-\tau)$, where $\tau \in [0, \infty]$. We can discretize time into bins with bin size Δt, whose centers are at positions $\tau_i, i = 0, 1, 2, ...,$, and define $\tilde{r}(s_i) = \tilde{r}[s(t-\tau_i)]$. Now $\tilde{r}$ is a function, and its value depends on the multivariables s_i. By Taylor series expansion, we have

$$\tilde{r}(\mathbf{s_0} + \delta \mathbf{s}) = \tilde{r}(\mathbf{s_0}) + \sum_i \frac{\partial \tilde{r}}{\partial s_i} \delta s_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 \tilde{r}}{\partial s_i \partial s_j} \delta s_i \delta s_j + ..., \tag{1}$$

We can think the stimulus has a time-independent mean s_0, and a time-dependent fluctuation δs. For simplicity, we may set $s_0 = 0$. To take off the hat and regain the functional r, we can replace $\sum_i$ by $\sum_i \Delta t \to \int dt$. This can be done in the following way

$$\begin{aligned} r[s + \delta s] = r(s) + \int d\tau \frac{\delta r}{\delta s(t-\tau)} \delta s(t-\tau) \\ + \int d\tau d\tau' \frac{\delta^2 r}{\delta s(t-\tau) \delta s(t-\tau')} \delta s(t-\tau) \delta s(t-\tau') + ... \end{aligned} \tag{2}$$

Here

$$\begin{aligned} \frac{\delta r}{\delta s(t-\tau)} &= \lim_{\Delta t \to 0} \frac{\partial \tilde{r}}{\partial s_i} \frac{1}{\Delta t} \\ \frac{\delta^2 r}{\delta s(t-\tau) \delta s(t-\tau')} &= \lim_{\Delta t \to 0} \frac{\partial^2 \tilde{r}}{\partial s_i \partial s_j} \frac{1}{\Delta t^2}. \end{aligned} \tag{3}$$

As we can see, if the functional derivative of r does not depend on the choice of t, then

$$\frac{\delta r}{\delta s(t-\tau)} = D(\tau), \tag{4}$$

$D(\tau)$ is indeed the functional derivative of the response function, and it is also called the first Wiener Kernel in the literature.

We would like to find the optimal kernel $D(\tau)$ to characterize the response function r. The estimated response function has the following form (only consider the first order),

$$r_{est}(t) = r_0 + \int_0^{\infty} D(\tau)s(t-\tau)d\tau \tag{5}$$

We have replaced $\delta s(t-\tau)$ with $s(t-\tau)$ by assuming that the mean of the stimulus is zero.

A formal definition of functional derivative in the physics literature is given by

$$\frac{\delta r[s(t-\tau)]}{\delta s(t-\tau')} = \lim_{\epsilon \to 0} \frac{r[s(t-\tau)+\epsilon\delta(\tau-\tau')] - r[s(t-\tau)]}{\epsilon} \tag{6}$$

where $\delta(\tau-\tau')$ is the delta function.

We would like to minimize the error function with respect to the function D, and find the optimal kernel. Functional derivative is very similar to function derivative, and there are convenient rules we would like to follow. For example

$$\begin{aligned} \frac{\delta f(x)}{\delta f(y)} &= \delta(x-y) \\ \frac{\delta h(f(x))}{\delta f(y)} &= h' \frac{\delta f(x)}{\delta f(y)} = h'\delta(x-y) \end{aligned} \tag{7}$$

Now let us define an error function,

$$E = \int_0^T dt \left(r_{est}(t) - r(t)\right)^2 \tag{8}$$

Therefore, we have

$$\begin{aligned} \frac{\delta E}{\delta D} = 0 &= \int_0^T dt(r_{est}(t) - r(t))\frac{\delta r_{est}}{\delta D}. \\ \frac{\delta r_{est}}{\delta D(\tau)} &= \int_0^T d\tau' s(t-\tau')\delta(\tau'-\tau) = s(t-\tau). \end{aligned} \tag{9}$$

As a result, we have

$$\int_0^T dt \int_0^{\infty} d\tau' D(\tau')s(t-\tau')s(t-\tau) = \int_0^T dt(r(t)-r_0)s(t-\tau) \tag{10}$$

Rearranging the integral on the left, we have

$$\int_0^{\infty} D(\tau')d\tau' \int_0^T dts(t-\tau')s(t-\tau) = \int_0^T dt(r(t)-r_0)s(t-\tau) \tag{11}$$

Now, we can define the autocorrelation function of the stimulus as

$$Q_{ss}(\tau) = \frac{1}{T}\int_0^T s(t)s(t+\tau)dt, \tag{12}$$

We can also define the correlation function between stimulus and response as

$$Q_{rs}(\tau) = \frac{1}{T}\int_0^T r(t)s(t+\tau)dt. \tag{13}$$

Then,

$$\int_0^\infty D(\tau')d\tau' Q_{ss}(\tau-\tau') = Q_{rs}(-\tau) \tag{14}$$

Now consider the simplest case in which the stimuli is white noise. In this case $Q_{ss}(\tau) = \sigma^2\delta(\tau)$. Then

$$\int_0^\infty D(\tau')\sigma^2\delta(\tau-\tau')d\tau' = D(\tau)\sigma^2 \tag{15}$$

As a result, we find the optimal kernel has an explicit expression

$$D(\tau) = \frac{Q_{rs}(-\tau)}{\sigma^2} \tag{16}$$

When the stimuli is not white noise, it is hard to calculate the kernel explicitly. However, if we ignore the causality, and make the assumption that response not only depends on the past but also the future of the stimulus, then things become much easier. performing Fourier transform, we found that in the frequency domain

$$\tilde{D}(\omega)\tilde{Q}_{ss}(\omega) = \tilde{Q}_{rs}(-\omega) \tag{17}$$

The Spike-Triggered Average

We are now at a position to measure the receptive field of a neuron. Before doing that, let me introduce the concept of spike triggered average, which is defined as

$$C(\tau) = \left\langle \frac{1}{n}\sum_{i=1}^{n} s(t_i-\tau)\right\rangle \tag{18}$$

Here t_i is the time when a spike occurs. The average $\langle ... \rangle$ is over different trials during which we are presenting exactly the same stimulus. The timing of the spike, however, can be variable. Below I would like to show that the spike triggered average has a direct connection the stimulus response correlation function $Q_{rs}(-\tau)$, provided that r is defined as the firing rate of a neuron. Given the density function

$$\rho(t) = \sum_i \delta(t - t_i) \tag{19}$$

It is easy to see that the

$$\sum_{i=1}^{n} s(t_i - \tau) = \int_0^T \rho(t) s(t - \tau) dt$$

Thus,

$$C(\tau) = \left\langle \frac{1}{n} \int_0^T \rho(t) s(t - \tau) dt \right\rangle$$

Formally, the firing rate of a neuron $r(t)$ is defined as

$$r(t) = \frac{1}{\Delta t} \int_t^{t+\Delta t} d\tau \langle \rho(\tau) \rangle, \tag{20}$$

where the average $\langle ... \rangle$ is also over different trials. Next, we want to make an important approximation by replacing the number of spikes n within a single trial with the mean number of spikes across all trials $\langle n \rangle$ during the period T. Then, because we are using exactly the same stimulus on each trial, the spike triggered average can be reduced to

$$\begin{aligned} C(\tau) &\approx \frac{1}{\langle n \rangle} \int_0^T \langle \rho(t) \rangle s(t - \tau) dt \\ &= \frac{1}{\langle n \rangle} \int_0^T r(t) s(t - \tau) dt \\ &= \frac{T}{\langle n \rangle} Q_{rs}(-\tau) \\ &= \frac{1}{\langle r \rangle} Q_{rs}(-\tau) \end{aligned} \tag{21}$$