Notes on linear-nonlinear-Poisson model

Ji Xia

January 2022

Contents

1 Linear-nonlinear-Poisson model (single-filter)

One of the common functional models for single neuronal responses is the linear-nonlinear-Poisson (LNP) model. LNP model define neural responses as a result of a cascade of linear, nonlinear, and Poisson spiking stages.

Given stimulus $\mathbf{x}(t)$, linear filter $\mathbf{k}$ and a nonlinear function $f(\cdot)$, the number of spikes $y(t)$ at time window $(t, t + \Delta t)$ is modeled as:

$$\lambda = f(\mathbf{k} \cdot \mathbf{x}(t)) \tag{1}$$

$$y(t) \sim \text{Poiss}(\lambda \Delta t) \tag{2}$$

Here $\mathbf{x}$ and $\mathbf{k}$ are D-dimensional vectors (e.g., for image stimulus with 100 pixels, $\mathbf{x}$ and $\mathbf{k}$ would be 100-dimensional vectors). ”·” denotes the inner product of two vectors. For simplicity of notations, we assume that Δt is small enough so that y = 0 or 1 at a given t.

In the experiment, we can choose the stimulus $\mathbf{x}(t)$ and record $y(t)$ for a given neuron. If the neuron was well-defined by an LNP model, how should we find out the linear filter $\mathbf{k}$ and the nonlinear function $f(\cdot)$?

2 STA: Gaussian stimulus

If the stimulus $\mathbf{x}$ follows a multivariate Gaussian distribution with zero mean (each $\mathbf{x}(t)$ is a random sample from the Gaussian distribution), and if $f(\cdot)$ leads to a change in the mean of spike-triggered ensemble compared to the raw stimulus ensemble, the spike-triggered average (STA) offers an unbiased estimate for the linear filter $\mathbf{k}$. In addition, the ratio between $P(\mathbf{k} \cdot \mathbf{x}|y = 1)$ and $P(\mathbf{k} \cdot \mathbf{x})$ leads to a maximum likelihood estimate of $f(\cdot)$.

2.1 Prove $\mathbf{k} \propto$ STA

Bussgang’s theorem states that for $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, C)$ and any arbitrary nonlinear function $f(\cdot)$, we have the expectation value

$$E[f(\mathbf{k} \cdot \mathbf{x})\mathbf{x}] = E[\mathbf{x}\mathbf{x}^T]\, \mathbf{k}\, E[f'(\mathbf{k} \cdot \mathbf{x})]$$

(For the derivation of Bussgang’s theorem, please see Appendix.)
Because $E[f'(\mathbf{k} \cdot \mathbf{x})]$ is just a constant depending on the form of $f(\cdot)$, we know that

$$\mathbf{k} \propto E[\mathbf{x}\mathbf{x}^T]^{-1} E[f(\mathbf{k} \cdot \mathbf{x})\mathbf{x}] = C^{-1} E[\lambda \mathbf{x}] \tag{3}$$

Furthermore, given a very long experiment, we can approximate $E[\lambda\Delta t\mathbf{x}]$ with STA:

$$E[\lambda\Delta t\mathbf{x}] = \sum_t \lambda\Delta t\mathbf{x} \tag{4}$$

$$= \sum_s \sum_{trial} y\mathbf{x} \tag{5}$$

$$= \sum_i \mathbf{x}_i \tag{6}$$

Here, s denotes index for different stimuli, $\mathbf{x}_i$ is the stimulus precedes i^{th} spike. From equation 3 and 6, we conclude that $\mathbf{k} \propto C^{-1}\sum_i \mathbf{x}_i$, which is STA rotated by the covariance of stimulus distribution.

2.2 Estimate $f(\cdot)$

From the definition of LNP model, when Δt is small enough so that $y = 0$ or 1, then we have

$$\begin{aligned} f(\mathbf{k}\cdot\mathbf{x})\Delta t &= \lambda\Delta t \\ &= P(\text{spike}|\mathbf{k}\cdot\mathbf{x}) \\ &= P(\text{spike})P(\mathbf{k}\cdot\mathbf{x}|\text{spike})/P(\mathbf{k}\cdot\mathbf{x}) \end{aligned}$$

Thus, given each value of $\mathbf{k}\cdot\mathbf{x}$, we can estimate the value of $f(\mathbf{k}\cdot\mathbf{x})$ as $\frac{1}{\Delta t}P(\text{spike})\frac{P(\mathbf{k}\cdot\mathbf{x}|\text{spike})}{P(\mathbf{k}\cdot\mathbf{x})}$.

3 STA: spherically symmetric distributed stimulus

If the stimulus $\mathbf{x}$ follows a spherically symmetric distribution (e.g., Gaussian white distribution), and if $f(\cdot)$ leads to a change in the mean of spike-triggered ensemble compared to the raw stimulus ensemble, STA offers an unbiased estimate for the linear filter $\mathbf{k}$ [1]. $f(\cdot)$ can be estimated in the same way as in 2.2.

3.1 Prove $\mathbf{k} \propto$ STA

Here we are going to show that if $P(\mathbf{x})$ is spherically symmetric, then STA $\propto \mathbf{k}$.
From equation 4, 5, 6, we know that STA $\propto E(f(\mathbf{k}\cdot\mathbf{x})\mathbf{x})$. By the definition of spherically symmetric distribution, we know that if $|\mathbf{x}| = |\mathbf{x}^*|$, then $P(\mathbf{x}) = P(\mathbf{x}^*)$. Furthermore, for any $\mathbf{x}$, we can always find a $\mathbf{x}^*$ such that they are symmetric with respect to $\mathbf{k}$ (in other words,

$\mathbf{x} + \mathbf{x}^* \propto \mathbf{k}$ and $\mathbf{k} \cdot \mathbf{x} = \mathbf{k} \cdot \mathbf{x}^*$). Therefore, we can rewrite STA as the following:

$$\begin{aligned}
\text{STA} &\propto E(f(\mathbf{k} \cdot \mathbf{x})\mathbf{x}) \\
&= \sum_{\mathbf{x}} f(\mathbf{k} \cdot \mathbf{x})\mathbf{x}P(\mathbf{x}) \\
&= \sum_{\mathbf{x},\mathbf{x}^*} f(\mathbf{k} \cdot \mathbf{x})\mathbf{x}P(\mathbf{x}) + f(\mathbf{k} \cdot \mathbf{x}^*)\mathbf{x}^*P(\mathbf{x}^*) \\
&= \sum_{\mathbf{x},\mathbf{x}^*} (\mathbf{x} + \mathbf{x}^*) f(\mathbf{k} \cdot \mathbf{x})P(\mathbf{x}) \\
&\propto \mathbf{k} \sum_{\mathbf{x}} f(\mathbf{k} \cdot \mathbf{x})P(\mathbf{x})
\end{aligned}$$

If $\sum_{\mathbf{x}} f(\mathbf{k} \cdot \mathbf{x})P(\mathbf{x})$ is not zero, then STA is proportional to $\mathbf{k}$.

4 Linear-nonlinear-Poisson model (multi-filter)

We can extend the single filter LNP model to the multi-filter LNP model by modifying equation 1.

$$\lambda = f(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x}) \tag{7}$$

It can be shown that if stimulus follows Gaussian or spherically symmetric distribution, then STA lies in the subspace defined by the set of linear filters $(\mathbf{k_1}, \ldots, \mathbf{k_m})$.

5 STC: Gaussian stimulus

If the stimulus x follows a multivariate Gaussian distribution with zero mean, and if $f(\cdot)$ leads to a change in the variance of spike-triggered ensemble compared to the raw stimulus ensemble, we can use spike-triggered covariance (STC) to identify the relevant subspace spanned by the set of linear filters in the LNP model [2].

5.1 Prove the relation between STC and the relevant subspace

STC is defined as the covariance of spike-triggered ensemble.

$$\text{STC} = E[(\mathbf{x} - \text{STA})(\mathbf{x} - \text{STA})^T|\text{spike}] \tag{8}$$

$$= E[\mathbf{x}\mathbf{x}^T|\text{spike}] - \text{STA STA}^T \tag{9}$$

We can rewrite the first term:

$$E[\mathbf{x}\mathbf{x}^T|\text{spike}] = \sum_{\mathbf{x}} \mathbf{x}\mathbf{x}^T P(\mathbf{x}|\text{spike}) \tag{10}$$

$$= \sum_{\mathbf{x}} \mathbf{x}\mathbf{x}^T \frac{P(\text{spike}|\mathbf{x})P(\mathbf{x})}{P(\text{spike})} \tag{11}$$

$$= \frac{1}{P(\text{spike})} \sum_{\mathbf{x}} \mathbf{x}\mathbf{x}^T \lambda \Delta t P(\mathbf{x}) \tag{12}$$

$$= \frac{1}{\sum_{\mathbf{x}} \lambda \Delta t P(\mathbf{x})} \sum_{\mathbf{x}} \mathbf{x}\mathbf{x}^T \lambda \Delta t P(\mathbf{x}) \tag{13}$$

To further relate this first term with the linear filters, we need to use Bussgang's theorem. Bussgang's theorem states that for D-dimensional vector $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, C)$ and any arbitrary nonlinear function $g(\cdot)$, we have:

$$E[x_i g(\mathbf{x})] = \sum_{j=1}^{D} C_{ij} E[\frac{\partial g(\mathbf{x})}{\partial x_j}]$$

If we apply Bussgang's theorem twice, we have:

$$\begin{aligned} E[x_i x_j g(\mathbf{x})] &= \sum_{k=1}^{D} C_{ik} E[\frac{\partial [x_j g(\mathbf{x})]}{\partial x_k}] \\ &= \sum_{k=1}^{D} C_{ik} (\delta_{j,k} E[g(\mathbf{x})] + E[x_j \frac{\partial g(\mathbf{x})}{\partial x_k}]) \\ &= C_{ij} E[g(\mathbf{x})] + \sum_{k,m=1}^{D} C_{ik} C_{jm} E[\frac{\partial^2 g(\mathbf{x})}{\partial x_k \partial x_m}] \end{aligned}$$

We can then divide both sides of the equation by $E[g(\mathbf{x})]$, then the left-hand side equals to equation 13, if $g(\mathbf{x}) = \lambda \Delta t = f(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x}) \Delta t$. Thus, we have:

$$\begin{aligned} \text{STC} + \text{STA STA}^T &= C + CGC^T \\ G_{ij} &= E[\frac{\partial^2 g(\mathbf{x})}{\partial x_i \partial x_j}] \\ &= E[\sum_{\alpha} \sum_{\beta} \frac{\partial^2 g(\mathbf{s})}{\partial s_\alpha \partial s_\beta} \mathbf{k}_\alpha(i) \mathbf{k}_\beta(j)] \end{aligned}$$

Here $s_\alpha = \mathbf{k}_\alpha \cdot \mathbf{x}$. Let's define $\Delta C = \text{STC} + \text{STA STA}^T - C$, we have:

$$\begin{aligned} \Delta C_{i,j} &= \sum_{m,\alpha,\beta,l} [C_{i,m} \mathbf{k}_\alpha(m)] A_{\alpha,\beta} [\mathbf{k}_\beta(l) C_{l,j}] \\ A_{\alpha,\beta} &= E[\frac{\partial^2 g(\mathbf{s})}{\partial s_\alpha \partial s_\beta}] \end{aligned}$$

If we define $\mathbf{K}$ as a $m \times D$ matrix whose rows correspond to different linear filters $\mathbf{k}_\alpha$. Then we can write $\Delta C = C\mathbf{K}^T A\mathbf{K}C$. The $D \times D$ matrix ΔC is determined by a $m \times m$ matrix A, meaning that ΔC will only have m nonzero eigenvalues. The eigenvectors of ΔC associated with non-zero eigenvalues, $\{\mathbf{u}_\alpha\}$, are linear combinations of the linear filters $\{\mathbf{k}_\alpha\}$, blurred by the stimulus covariance. In other words, $\{C^{-1}\mathbf{u}_\alpha\}$ span the relevant subspace defined by $\mathbf{K}$.

6 STC: spherically symmetric distributed stimulus

If the stimulus x follows a spherically symmetric distribution, and if $f(\cdot)$ leads to a change in the variance of spike-triggered ensemble compared to the raw stimulus ensemble, we can use STC to identify the relevant subspace spanned by the set of linear filters in the LNP model[3].

6.1 Prove the relation between STC and the relevant subspace

First, we want to show that any vector perpendicular to the relevant space ($\mathbf{v} \in \mathbf{K}_\perp$) is an eigenvector of STC: STC $\mathbf{v} \propto \mathbf{v}$

Let's derive the expression for STC $\mathbf{v}$. Because STA lies in the relevant subspace, we have STA$^T\mathbf{v} = 0$. According to equation 9, we have:

$$\text{STC } \mathbf{v} = E[\mathbf{x}\mathbf{x}^T|\text{spike}]\mathbf{v} \tag{14}$$

Similar to the derivation in section 3.1, for any stimulus $\mathbf{x}$, we can find a stimulus $\mathbf{x}^*$ such that $\mathbf{x}$ and $\mathbf{x}^*$ are symmetric with respect the hyperplane $\mathbf{v}_\perp$. Because $\mathbf{K}$ is a subspace of $\mathbf{v}_\perp$, we have $P(\text{spike}|\mathbf{x}) = P(\text{spike}|\mathbf{x}^*)$. In addition, by the definition of spherically symmetric distribution, we choose $|\mathbf{x}^*| = |\mathbf{x}|$ so that $P(\mathbf{x}^*) = P(\mathbf{x})$. According to equation 11, we can rewrite equation 14 as the following:

$$\begin{aligned}
\text{STC } \mathbf{v} &= \frac{1}{P(\text{spike})}\sum_{\mathbf{x}} P(\text{spike}|\mathbf{x})P(\mathbf{x})\mathbf{x}\mathbf{x}^T\mathbf{v} && (15)\\
&\propto \sum_{\mathbf{x},\mathbf{x}^*} P(\text{spike}|\mathbf{x})P(\mathbf{x})\mathbf{x}\mathbf{x}^T\mathbf{v} + P(\text{spike}|\mathbf{x}^*)P(\mathbf{x}^*)\mathbf{x}^*\mathbf{x}^{*T}\mathbf{v} && (16)\\
&= \sum_{\mathbf{x},\mathbf{x}^*} P(\text{spike}|\mathbf{x})P(\mathbf{x})(\mathbf{x}\mathbf{x}^T\mathbf{v} + \mathbf{x}^*\mathbf{x}^{*T}\mathbf{v}) && (17)
\end{aligned}$$

Because $\mathbf{x}$ and $\mathbf{x}^*$ are symmetric with respect the hyperplane $\mathbf{v}_\perp$, we have $\mathbf{x}^T\mathbf{v} = -\mathbf{x}^{*T}\mathbf{v}$ and $\mathbf{x} - \mathbf{x}^* \propto \mathbf{v}$. Thus, we can rewrite equation 17 as the following:

$$
\begin{aligned}
\text{STC } \mathbf{v} &\propto \sum_{\mathbf{x},\mathbf{x}^*} P(\text{spike}|\mathbf{x})P(\mathbf{x})(\mathbf{x}\mathbf{x}^T\mathbf{v} - \mathbf{x}^*\mathbf{x}^T\mathbf{v}) \\
&\propto \sum_{\mathbf{x},\mathbf{x}^*} P(\text{spike}|\mathbf{x})P(\mathbf{x})\mathbf{v}\mathbf{x}^T\mathbf{v} \\
&\propto \mathbf{v}
\end{aligned}
$$

Second, we want to show that those vectors perpendicular to the relevant space are associated with the same eigenvalue of STC. Consider two non-parallel vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ that belong to $\mathbf{K}_\perp$. We have STC$\mathbf{v}_1 = a\mathbf{v}_1$ and STC$\mathbf{v}_2 = b\mathbf{v}_2$. The sum of the two vectors also belongs to $\mathbf{K}_\perp$, thus is also an eigenvector of STC: STC$(\mathbf{v}_1 + \mathbf{v}_2) = c(\mathbf{v}_1 + \mathbf{v}_2)$. The identity STC$(\mathbf{v}_1 + \mathbf{v}_2) = a\mathbf{v}_1 + b\mathbf{v}_2 = c(\mathbf{v}_1 + \mathbf{v}_2)$ suggest that $a = b = c$.

In conclusion, irrelevant subspace is spanned by the eigenvectors of STC with the same degenerate eigenvalue, while relevant subspace is spanned by the eigenvectors of STC whose eigenvalues deviate from the degenerate eigenvalue. In practice, if we have the number of linear filters much less than the dimension of the stimulus ($m \ll D$), we can identify the most common eigenvalue or the chance level eigenvalue as the degenerate eigenvalue for irrelevant subspace.

7 The equivalence of MLE and MID

For stimulus follows any arbitrary distribution, there are two methods for estimating the linear filters $\mathbf{K}$: maximum likelihood estimation (MLE) and maximally informative dimension (MID). It turns out the two methods are equivalent to each other [4].

7.1 MLE

We can estimate $\mathbf{K}$ and parameters in $f(\cdot)$ by maximizing the log likelihood of LNP model given data ($\mathcal{L}(\mathbf{K}, f; \mathbf{x}, y)$).

$$
\begin{aligned}
\mathcal{L}(\mathbf{K}, f; \mathbf{x}, y) &= \log P(y|\mathbf{x}, \mathbf{K}, f) && (18) \\
&= \log \prod_t P(y(t)|\mathbf{x}(t), \mathbf{K}, f) && (19) \\
&= \sum_t \log \frac{1}{y(t)!}[f(\cdot)\Delta t]^{y(t)} e^{-f(\cdot)\Delta t} && (20) \\
&= \sum_t y(t) \log[f(\cdot)\Delta t] - \Delta t \sum_t f(\cdot) - \sum_t \log y(t)! && (21) \\
&= \sum_t y(t) \log[f(\cdot)\Delta t] - n_{sp} - \sum_t \log y(t)! && (22)
\end{aligned}
$$

Here $n_{sp} = \sum_t y(t)$ is the total number of spikes during the long experiment.

7.2 MID

We can estimate $\mathbf{K}$ by maximizing the KL Divergence between $q(\mathbf{x}) = P(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x}|\text{spike})$ and $p(\mathbf{x}) = P(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x})$:

$$\begin{align} I &= D_{KL}(q(\mathbf{x})||p(\mathbf{x}) \tag{23} \\ &= \sum_{\mathbf{x}} q(\mathbf{x}) \log \frac{q(\mathbf{x})}{p(\mathbf{x})} \tag{24} \end{align}$$

7.3 Prove that MLE is equivalent to MID

Let's write out the MLE estimation for $f(\cdot)$ given $\mathbf{K}$. Similar to the derivation in section 2.2, in the multi-filter LNP model:

$$\begin{aligned} f(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x})\Delta t &= P(\text{spike}) \frac{P(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x}|\text{spike})}{P(\mathbf{k_1} \cdot \mathbf{x},\ \mathbf{k_2} \cdot \mathbf{x},\ \ldots,\ \mathbf{k_m} \cdot \mathbf{x})} \\ &= P(\text{spike}) \frac{q(\mathbf{x})}{p(\mathbf{x})} \end{aligned}$$

We can substitute this expression of $\dfrac{q(\mathbf{x})}{p(\mathbf{x})}$ into the expression of KL Divergence I:

$$\begin{aligned} I &= \sum_{\mathbf{x}} q(\mathbf{x}) \log \frac{f(\cdot)\Delta t}{P(\text{spike})} \\ &= \sum_{\mathbf{Kx}} P(\mathbf{Kx}) \frac{P(\text{spike}|\mathbf{Kx})}{P(\text{spike})} \log \frac{f(\mathbf{Kx})\Delta t}{P(\text{spike})} \\ &= \sum_t \frac{1}{T} \frac{y(t)}{n_{sp}/T} \log \frac{f(\mathbf{Kx}(t))\Delta t}{P(\text{spike})} \\ &= \frac{1}{n_{sp}} \sum_t y(t) \log \frac{f(\mathbf{Kx}(t))\Delta t}{P(\text{spike})} \end{aligned}$$

Now let's write I in terms of log likelihood $\mathcal{L}$:

$$\begin{align} I &= \frac{1}{n_{sp}} [\mathcal{L} - (\sum_t y(t) \log P(\text{spike}) - n_{sp} - \sum_t \log y(t)!)] \tag{25} \\ &= \frac{1}{n_{sp}} [\mathcal{L} - \mathcal{L}_{null}] \tag{26} \end{align}$$

We can interpret $\mathcal{L}_{null}$ as the log likelihood of a null LNP model where $f(\cdot)$ doesn't depend on stimulus $\mathbf{x}$. Equation 26 shows that maximize I is equivalent to maximize $\mathcal{L}$.

Appendix

Proof of Bussgang theorem

For $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, C)$ and an arbitrary nonlinear function $f(\cdot)$, we have:

$$
\begin{aligned}
E[f(\mathbf{x})\mathbf{x}] &= \det(2\pi C)^{-1/2} \int f(\mathbf{x})\mathbf{x} e^{-\frac{1}{2}\mathbf{x}^T C^{-1}\mathbf{x}} d\mathbf{x} && (27) \\
&= -\det(2\pi C)^{-1/2} \int f(\mathbf{x}) C d(e^{-\frac{1}{2}\mathbf{x}^T C^{-1}\mathbf{x}}) && (28) \\
&= \det(2\pi C)^{-1/2} C \int e^{-\frac{1}{2}\mathbf{x}^T C^{-1}\mathbf{x}} df(\mathbf{x}) && (29) \\
&= \det(2\pi C)^{-1/2} C \int e^{-\frac{1}{2}\mathbf{x}^T C^{-1}\mathbf{x}} f'(\mathbf{x}) d\mathbf{x} && (30) \\
&= CE[f'(\mathbf{x})] && (31)
\end{aligned}
$$

References

[1] EJ Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: computation in neural systems*, 12(2):199, 2001.

[2] William Bialek and Rob R van Steveninck. Features and dimensions: Motion estimation in fly vision. *arXiv preprint q-bio/0505003*, 2005.

[3] Inés Samengo and Tim Gollisch. Spike-triggered covariance: geometric proof, symmetry properties, and extension beyond gaussian stimuli. *Journal of computational neuroscience*, 34(1):137–161, 2013.

[4] Ross S Williamson, Maneesh Sahani, and Jonathan W Pillow. The equivalence of information-theoretic and likelihood-based methods for neural dimensionality reduction. *PLoS computational biology*, 11(4):e1004141, 2015.