

多点計測スパイクデータに基づく神経回路網の結合推定

村田 昇 (早稲田大学)

1 はじめに

近年、多電極同時記録法などの計測技術の発展により、多数の神経細胞を数時間から 24 時間に渡って同時に計測できるようになった [1]. この計測により、スパイクデータと呼ばれる細胞の発火を記録した点過程データを得ることができ、スパイクデータを解析対象とした研究が盛んに行われてきた [2]. 古典的には、2 つの細胞の組全てに対して Cross-Correlation を算出し、その相関係数を細胞間の結合の強度とみなす方法が用いられてきた [3] が、相関のみに注目した手法は本来存在しない結合を推定する擬相関による推定誤りが生じる。これを回避する解析手法として、グラフ構造推定や情報幾何を用いた解析手法 [4]–[6] などが提案されているが、無向グラフで情報伝達の方向性を考慮できない、あるいは興奮・抑制性を表現できないなどの問題点が指摘されている。

本発表では、高次の相関の影響も含めたスパイクデータの生成モデルを仮定することで、細胞間の結合と細胞間の相関の関係を定式化し、上記の問題を解決する方法を提案する*1.

2 モデル

観測される細胞の数を N とし、細胞 i の時刻 t における発火を $X_i(t) \in \{0, 1\}$ で表す。細胞の発火は、膜電位に相当する内部状態 $U_i(t) \in \mathbb{R}$ と活性化関数 Φ によって確率的に決まるとする。

$$\Pr(X_i(t) = 1) = \Phi_{\sigma^2}(U_i(t)), \quad \Phi_{\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{z^2}{2\sigma^2}\right) dz. \quad (1)$$

信号伝達の時間遅れによる相関を考えるために、区間 $t_{\Delta} = [t - \Delta, \dots, t - 1]$ における細胞の発火を $X_i[t_{\Delta}] \in \{0, 1\}$ で表すこととし、内部状態は結合した細胞の発火状態の線形和に依存するものとする。

$$U_i(t) = B_i(t) + \sum_{j=1}^N \lambda_{ij} X_j[t_{\Delta}], \quad (2)$$

ただし、 $B_i(t)$ は観測されない細胞から細胞 i への入力を表し、係数 λ_{ij} は細胞 j が発火した時に細胞 i に与える影響の大きさを表す。なお、 λ_{ij} は細胞 j から細胞 i への直接の結合 w_{ij} による影響だけでなく、他の細胞を介した間接的な影響を含んでおり、観測期間中は時不変な定数として扱う。以後、この間接的な影響を含めた係数 λ_{ij} を擬結合と呼ぶ。

*1 本発表は、岩崎泰士、樋口翔 (早稲田大学)、日野英逸 (筑波大学)、赤穂昭太郎 (産業技術総合研究所)、龍野正実 (Lethbridge 大学) との共同研究に基づくものである [7].

3 提案手法

まず観測されない細胞が個々の細胞に平均的に与える影響の大きさを推定し、それを用いてスパイクデータから内部状態に含まれる擬結合 λ_{ij} を推定する手法を提案する。次に擬結合 λ_{ij} から結合 w_{ij} を推定する手法について述べる。

3.1 擬結合の推定

観測されない細胞の影響を取り除くために、独立な確率変数の平均に関する以下の定理を利用する。

定理 1. 確率変数 X, Y は独立とする。任意の有界な関数 g に対して

$$\mathbb{E}[g(X + Y)] = \mathbb{E}[h(X + \mathbb{E}[Y])] \quad (3)$$

が成り立つ。ただし、 f_Y は Y の確率密度で、 h は以下で与えられる。

$$h = g * f_Y^-, \quad f_Y^-(x) = f_Y(\mathbb{E}[Y] - x) \quad (4)$$

これは [8] の一般形である。特に活性化関数が正規分布の累積確率で、観測されない細胞の影響が正規分布である場合には、以下の簡単な関係が成り立つ

系 2. 関数 g が Φ_{σ^2} 、確率変数 X が定数 x 、また Y が平均 $\mathbb{E}[Y]$ 、分散 τ^2 の正規分布に従うとする。このとき

$$\mathbb{E}[\Phi_{\sigma^2}(x + Y)] = \Phi_{\sigma^2 + \tau^2}(x + \mathbb{E}[Y]) \quad (5)$$

が成り立つ。

まず $X_j[t_\Delta] = 1$ の場合を考える。項を整理して

$$U_i(t | X_j[t_\Delta] = 1) = B_i(t) + \lambda_{ij} X_j[t_\Delta] + \sum_{k \neq j} \lambda_{ik} X_k[t_\Delta] = \lambda_{ij} + C_{ij}(t | X_j[t_\Delta] = 1).$$

とまとめ、 $C_{ij} \sim \mathcal{N}(\bar{C}_{ij}, \tau^2)$ を仮定して系を用いると

$$\begin{aligned} \mathbb{E}[X_i(t) | X_j[t_\Delta] = 1] &= \mathbb{E}[\Phi_{\sigma^2}(U_i(t | X_j[t_\Delta] = 1))] = \mathbb{E}[\Phi_{\sigma^2}(\lambda_{ij} + C_{ij}(t | X_j[t_\Delta] = 1))] \\ &= \Phi_{\rho^2}(\lambda_{ij} + \bar{C}_{ij}), \quad (\rho^2 = \sigma^2 + \tau^2) \end{aligned}$$

が成り立つ。更に、2 値確率変数の平均と確率値の関係を用いると

$$\Phi_{\rho^2}(\lambda_{ij} + \bar{C}_{ij}) = \Pr(X_i(t) = 1 | X_j[t_\Delta] = 1), \quad \Leftrightarrow \quad \lambda_{ij} + \bar{C}_{ij} = \rho \cdot \Phi_1^{-1}(\Pr(X_i(t) = 1 | X_j[t_\Delta] = 1))$$

と書き直すことができる。 $X_j[t_\Delta] = 0$ の場合も同様に考えると

$$\begin{aligned} U_i(t | X_j[t_\Delta] = 1) &= \lambda_{ij} + C_{ij}(t | X_j[t_\Delta] = 1), \\ U_i(t | X_j[t_\Delta] = 0) &= C_{ij}(t | X_j[t_\Delta] = 0). \end{aligned}$$

となるので、 $C_{ij}(t | X_j[t_\Delta] = 1), C_{ij}(t | X_j[t_\Delta] = 0) \sim \mathcal{N}(\bar{C}_{ij}, \tau^2)$ を仮定すると、 λ_{ij} は以下の式で表すことができる。

$$\lambda_{ij} = \rho \{ \Phi_1^{-1}(\Pr(X_i(t) = 1 | X_j[t_\Delta] = 1)) - \Phi_1^{-1}(\Pr(X_i(t) = 1 | X_j[t_\Delta] = 0)) \}. \quad (6)$$

したがって、以下の標本平均を用いて擬結合 λ_{ij} の推定を行うことができる。

$$\Pr(X_i(t)=1 | X_j[t_\Delta]=1) = \frac{1}{|T_1|} \sum_{t \in T_1} X_i(t), \quad T_1 = \{t | X_j[t_\Delta]=1\}$$

$$\Pr(X_i(t)=1 | X_j[t_\Delta]=0) = \frac{1}{|T_0|} \sum_{t \in T_0} X_i(t), \quad T_0 = \{t | X_j[t_\Delta]=0\}.$$

3.2 結合と擬結合の関係

次に擬結合 λ_{ij} と真の結合 w_{ij} の関係をモデル化する。2つの関係は適当な時間遅れ δ, δ' を用いると

$$\lambda_{ij} = w_{ij} + \sum_k w_{ik} \Pr(X_k(t-\delta)=1 | X_j(t-\delta')=1) + (\text{higher order terms})$$

で十分良く近似できる。ここで、適当な区間 t_δ を用いた仮想的な確率を

$$\theta_{ij} = \Pr(X_i(t)=1 | X_j[t_\delta]=1),$$

と置くと、擬結合 λ_{ij} は

$$\lambda_{ij} = w_{ij} + \sum_k w_{ik} \theta_{kj} + \sum_{k,l} w_{ik} \theta_{kl} \theta_{lj} + \sum_{k,l,m} w_{ik} \theta_{kl} \theta_{lm} \theta_{mj} + \dots$$

と分解できると考えられる。これは、 $\Lambda = (\lambda_{ij})$, $W = (w_{ij})$, $\Theta = (\theta_{ij})$ として行列表示すると

$$\Lambda = W(I + \Theta + \Theta^2 + \Theta^3 + \dots) = W(I - \Theta)^{-1},$$

となる。また θ_{ij} と w_{ij} の間には

$$\theta_{ij} = \Pr(X_i(t)=1 | X_j[t_\delta]=1) = \mathbb{E}[\Phi_{\sigma^2}(w_{ij} + C'_{ij})] = \Phi_{\rho^2}(w_{ij} + \mathbb{E}[C'_{ij}])$$

という関係があるので、 $C'_{ij} \sim \mathcal{N}(\bar{C}_{ij}, \tau^2)$ を仮定すれば、 w_{ij} を用いて θ_{ij} を計算することができる。

$$\theta_{ij} = \Phi_{\rho^2}(w_{ij} + \bar{C}_{ij}),$$

$$\bar{C}_{ij} = \rho \cdot \Phi_1^{-1}(\Pr(X_i(t)=1 | X_j[t_\Delta]=0)).$$

以上より Λ と W の関係式が得られ、与えられた Λ から W を数値的に解くことができる。

3.3 興奮性および抑制性の推定

神経細胞には興奮性細胞と抑制性細胞があり、それぞれ正の結合か負の結合しかもたないと考えられている。生理学的な知見から細胞の性質がわかっている場合もあるが、一般には未知なので、上記のモデルに基づいて推定を行うことになる。細胞 i のラベル z_i を考え、興奮性を 1、抑制性を 0 で表すとすれば、Bayes の定理を用いてラベル $\mathbf{z} = (z_i)$ の事後確率を求め、この値に基づいて判別することができる。

$$\Pr(\mathbf{z} | \text{Data}, W) \propto \Pr(\text{Data} | W, \mathbf{z})$$

ただし、この計算は解析的に実行することが難しいので、Gibbs サンプリングを用いた近似計算を行う。

以上、各ステップを統合し、ラベル \mathbf{z} の推定と W の推定を交互に行うことにより、結合 w_{ij} の推定を行うことができる。以下に興奮性か抑制性かのラベル \mathbf{z} を与えたときの W の推定アルゴリズムをまとめる。

Algorithm 1 W の推定アルゴリズム

```
1: Input:  $\Lambda, \bar{C}, \mathbf{z}$ 
2: procedure ESTIMATEW( $\Lambda, \bar{C}, \mathbf{z}$ )
3:   Initialization:  $\Theta^{(1)} \leftarrow [0, 1]^{N \times N}, \Lambda^{(1)} \leftarrow \Lambda$ 
4:   for  $\tau \leftarrow 1, T$  do
5:      $W^{(\tau+1)} \leftarrow \Lambda^{(\tau)}(I - \Theta^{(\tau)})$ 
6:     for  $i \leftarrow 1, N$  do
7:       for  $j \leftarrow 1, N$  do
8:          $[\hat{W}(\mathbf{z})^{(\tau+1)}]_{ij} \leftarrow \begin{cases} z_j[W^{(\tau+1)}]_{ij}, & [W^{(\tau+1)}]_{ij} > 0 \\ (1 - z_j)[W^{(\tau+1)}]_{ij}, & [W^{(\tau+1)}]_{ij} < 0 \end{cases}$ 
9:       end for
10:     end for
11:      $[\Theta^{(\tau+1)}]_{ij} \leftarrow \Phi_1([\hat{W}(\mathbf{z})^{(\tau+1)}]_{ij} + \bar{C}_{ij})$ 
12:      $\text{diag}(\Theta^{(\tau+1)}) \leftarrow 0$  ▷ update diagonal elements
13:      $\Lambda^{(\tau+1)} \leftarrow \Lambda^{(\tau)}$ 
14:      $\text{diag}(\Lambda^{(\tau+1)}) \leftarrow \text{diag}(\Lambda^{(\tau)}\Theta^{(\tau+1)})$  ▷ update diagonal elements
15:   end for
16: end procedure
17: Output:  $\hat{W}(\mathbf{z})$ 
```

4 まとめ

本発表では、部分的に観測された細胞のスパイクデータから、観測されない細胞の影響を消去しつつ細胞間の結合を推定する手法を提案した。紙面の都合上割愛したが、人工データ実験によって、提案手法が主要な結合を高精度に抽出できることが示されている。ただし、提案手法は結合強度の小さい抑制性結合の推定精度が低いという問題点がある。これは興奮性細胞と抑制性細胞では発火頻度などの性質が大きく異なるにも拘らず、モデル上は同一の細胞を仮定し、結合のみ異なる扱いをしていることが主な原因と考えられる。推定精度を上げるためには、興奮性細胞と抑制性細胞の違いを考慮したより詳細なモデルが必要である。

References

- [1] M. Tatsuno, P. Lipa, and B. L. McNaughton, “Methodological considerations on the use of template matching to study long-lasting memory trace replay,” *The Journal of Neuroscience*, vol. 26, pp. 10727–10742, 42 Oct. 18, 2006.
- [2] S. Kim, D. Putrino, S. Ghosh, and E. N. Brown, “A granger causality measure for point process models of ensemble neural spiking activity,” *PLoS Computational Biology*, vol. 7, no. 3, Mar. 24, 2011.
- [3] M. A. Wilson and B. L. McNaughton, “Reactivation of hippocampal ensemble memories during sleep,” *Science*, vol. 265, pp. 676–679, 5172 Jul. 29, 1994.
- [4] H. Nakahara and S.-i. Amari, “Information-geometric measure for neural spikes,” *Neural Computation*, vol. 14, no. 10, pp. 2269–2316, Oct. 2002.
- [5] M. Tatsuno, J.-M. Fellous, and S.-i. Amari, “Information-geometric measures as robust estimators of connection strengths and external inputs,” *Neural Computation*, vol. 21, no. 8, pp. 2309–2335, Aug. 2009.
- [6] A. Noda, H. Hino, M. Tatsuno, S. Akaho, and N. Murata, “Intrinsic graph structure estimation using graph laplacian,” *Neural Computation*, vol. 26, no. 7, pp. 1455–1483, Jul. 2014.
- [7] S. Akaho, S. Higuchi, T. Iwasaki, H. Hino, M. Tatsuno, and N. Murata, “Graph structure modeling for multi-neuronal spike data,” *Journal of Physics: Conference Series*, vol. 699, 012012 Apr. 6, 2016.
- [8] A. Hyvärinen, “Gaussian moments for noisy independent component analysis,” *IEEE Signal Processing Letters*, vol. 6, pp. 145–147, 6 Aug. 6, 2002.