

スパースモデリングに基づく神経ネットワークの構造推定

Neural Network Estimation Based on Sparse Modeling

政廣 蓮¹

Ren Masahiro

大森 敏明¹

Toshiaki Omori

1 はじめに

近年、神経科学に基づく人工知能の発展に伴い、脳研究の重要性が高まっている [1]. 特に、情報処理の役割を担うと考えられている神経ネットワークの解明は脳研究において重要課題の一つとして位置付けられている.

しかし、観測データに基づく神経ネットワークの解明は容易なものではない. 神経ネットワークの解明が困難である理由の一つとして、ネットワークを構成するニューロン数に対して、推定すべきシナプス結合の数が非常に多いことが考えられる. 例えば、ニューロン数が n 個であるとき、推定すべきシナプス結合の数は n^2 となる. その他の理由として、神経ネットワークの推定に用いることのできる情報が限られていることが挙げられる. ニューロンの活動は膜電位やチャネル変数等に依存するのだが、これら全てを観測することは現段階では不可能である.

本研究では、膜電位データのみから神経ネットワークを推定するスパースモデリング手法を提案する (Fig. 1). 提案手法では、逐次モンテカルロ法 [2] と LASSO (Least Absolute Shrinkage and Selection Operator) [3] を融合することで神経ネットワーク内の主要な結合の抽出を実現する. 既存手法と推定結果を比較することで、我々の提案法の有効性を示す.

2 神経ネットワークのスパース推定

ニューロンの神経活動は非線形のダイナミクスに従うことが知られており、また、神経ネットワーク内の結合にはスパース性があることが知られている [4]. これらの知見から、本節では、神経ネットワークの推定手法として、逐次モンテカルロ法と LASSO を融合させた手法を提案する.

2.1 逐次モンテカルロ法による神経回路ダイナミクスの推定

本研究では、神経ネットワークの非線形ダイナミクスが Hodgkin-Huxley (HH) モデル [5] に従うものとする. HH モデルでは、神経の内部状態が潜在変数と呼ばれる膜電位 V とチャネル変数 m, h, n に依存する. この時、神経ネットワークを構成する i 番目のニューロンを例にとり、その内部状態を HH モデルを用いて表すと次のようになる.

$$C \frac{dV^{(i)}}{dt} = I_{ext} - I_{ion} \left(V^{(i)}, m^{(i)}, h^{(i)}, n^{(i)} \right) - \sum_j g_{syn}^{(j \rightarrow i)} s^{(j)} \left(V^{(i)} - V_{syn}^{(j)} \right) \quad (1)$$

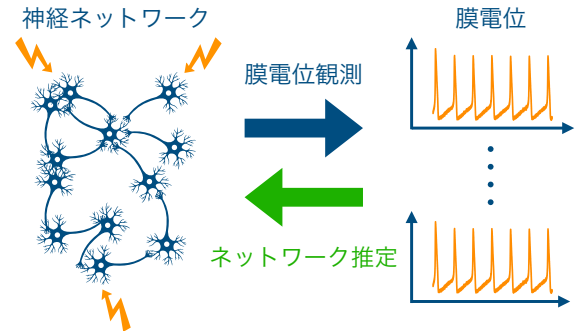


Fig. 1. 研究概要. 神経ネットワークを構成するニューロンに対して外部電流を流入させる. この時、各ニューロンの膜電位の時間変化を観測する. 得られた膜電位データを用いて神経ネットワークの内部状態を推定し、推定した内部状態をもとに神経ネットワークの結合を逆問題的に推定する.

ただし、 $g_{syn}^{(j \rightarrow i)}$ はシナプスコンダクタンスと呼ばれ、後述するスパースモデリングによるネットワーク推定とはシナプスコンダクタンスの値を推定することと等価である.

次に、神経ネットワークに対する状態空間モデルを導出する. 式(1)のネットワークモデルを離散化し、時刻 t における各ニューロンの潜在変数を \mathbf{x}_t 、観測される膜電位の集合を \mathbf{y}_t と定義すると、HHモデルに従う状態空間モデルは次のように表される.

$$\mathbf{x}_{t+\Delta t} = \mathbf{f}(\mathbf{x}_t) + \mathbf{v}_t \quad (2)$$

$$\mathbf{y}_{t+\Delta t} = \mathbf{V}_{t+\Delta t} + \mathbf{w}_{t+\Delta t} \quad (3)$$

ただし、 \mathbf{f} は HH モデルに従う非線形関数、 \mathbf{V}_t は時刻 t における真の全ニューロンの膜電位の集合、 $\mathbf{v}_t, \mathbf{w}_{t+\Delta t}$ はそれぞれシステムノイズと観測ノイズを表す. 式(2), (3)に確率モデルを適用すれば、それぞれ $p(\mathbf{x}_{t+\Delta t} | \mathbf{x}_t)$ と $p(\mathbf{y}_{t+\Delta t} | \mathbf{x}_{t+\Delta t})$ となる. これらの確率分布を用いて、各時刻に対して逐次モンテカルロ法の主操作となる一期先予測とフィルタリングを行う. 一期先予測とフィルタリングの操作とはそれぞれ、次の各確率分布を求めることを意味する.

$$p(\mathbf{x}_{t+\Delta t} | \mathbf{y}_{0:t}) = \int p(\mathbf{x}_{t+\Delta t} | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{0:t}) d\mathbf{x}_t \quad (4)$$

$$p(\mathbf{x}_{t+\Delta t} | \mathbf{y}_{0:t+\Delta t}) = \frac{p(\mathbf{y}_{t+\Delta t} | \mathbf{x}_{t+\Delta t}) p(\mathbf{x}_{t+\Delta t} | \mathbf{y}_{0:t})}{p(\mathbf{y}_{t+\Delta t} | \mathbf{y}_{0:t})} \quad (5)$$

一期先予測とフィルタリングを逐次的に行う、すなわち式(4), (5)を交互に求めることにより、神経回路ダイナミクスに関連する潜在変数の推定を実現する.

¹ 神戸大学大学院工学研究科, Graduate School of Engineering, Kobe University

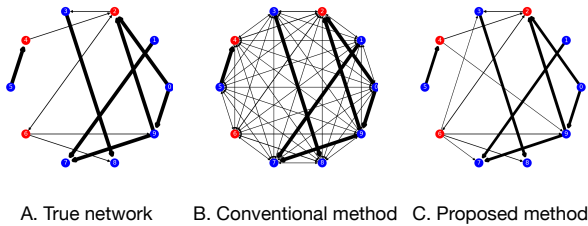


Fig. 2. 神経ネットワークの推定結果. (A). 推定対象となる真の神経ネットワーク. 赤点は興奮性ニューロン, 青点は抑制性ニューロンを表す. (B). 既存手法による推定結果. 本来ならばシナプス結合が存在しない場所でも結合が存在すると推定されている. (C). 提案手法による推定結果. 主要な結合を抽出することにより, 多くの箇所でもシナプス結合を正確に推定できている.

2.2 LASSOによる神経ネットワークの推定

神経ネットワークを推定するという事は, 式(1)のシナプスコンダクタンス g_{syn} を推定することと同値である. このシナプスコンダクタンスは式(2)にのみ現れる. 従って, 式(2)に基づいてニューロン i の膜電位に関する誤差関数を次のように定義する.

$$\begin{aligned} E^{(i)} &= \sum_t \left| V_{t+\Delta t}^{(i)} - f_V(\mathbf{x}_t) \right|^2 \\ &= \left\| \mathbf{z}^{(i)} - A^{(i)} \mathbf{g}_{syn}^{(i)} \right\|_2^2 \end{aligned} \quad (6)$$

神経ネットワークのスパース推定を実現するために, 式(6)に対してシナプスコンダクタンスに関する正則化項を加え, 誤差関数を再定義する.

$$E_{l_{asso}}^{(i)} = \left\| \mathbf{z}^{(i)} - A^{(i)} \mathbf{g}_{syn}^{(i)} \right\|_2^2 + \lambda \left\| \mathbf{g}_{syn}^{(i)} \right\|_1 \quad (7)$$

ただし, λ は正則化係数である. L_1 正則化を適用した式(7)を最小化するシナプスコンダクタンスを求めることで, 主要な結合のみを抽出することができる. これにより, スパースモデリングに基づく神経ネットワークを実現する.

3 実験結果

本研究では, 興奮性ニューロンと抑制性ニューロンからなる神経ネットワークを想定した. また, 各ニューロンを結ぶシナプスコンダクタンスの値は, 興奮性ニューロンから他のニューロンへと繋がるものは0.2, 抑制性ニューロンから他のニューロンへと繋がるものは1.0と定めた.

提案手法に従い, 膜電位のみを用いて神経ネットワークを推定した結果をFig. 2に示す. Fig. 2Bに見られるように, 既存手法による神経ネットワークの推定結果は, 本来ニューロン同士の結合が存在しない箇所までも結合が存在するとし, シナプスコンダクタンスの推定値が非0を示した. 一方で, Fig. 2Cに示すように, 提案手法では, 既存手法の推定結果と比較すると, シナプスコンダクタンスが存在しない多くの箇所でも, シナプスコンダクタンスの推定値が0と推定された. 即ち既存手法に比べ高い精度でネットワークの構造推定が行われていたことが分かる. さらに, Fig. 3に示すように, テストデータに対する膜電位応答においても提案手法の優位性が確認できた. 従って, 汎化性能の観点でも, 提案手法の優位性が示された.

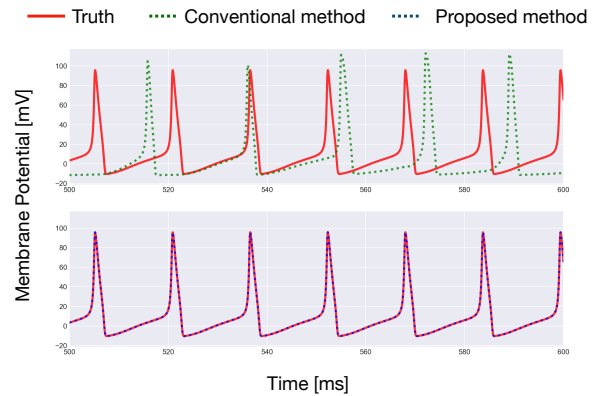


Fig. 3. シナプス結合の推定値を用いた膜電位の再現応答. (上) 既存手法に比べて, (下) 提案手法では高い精度で膜電位応答が再現された.

4 結論

本研究では, 観測された膜電位のみから神経ネットワークを推定するために, 逐次モンテカルロ法とLASSOを反復的に適用する手法を提案した. ニューロン間の結合の有無を決定するシナプスコンダクタンスに対して制約条件を付加した誤差関数を定義し, その誤差関数を最小化するシナプスコンダクタンスを求めることで, 神経ネットワークのスパース推定を実現した. そして, 既存手法と推定結果を比較する事で我々の提案手法の有効性を示した.

謝辞

本研究の一部は, 文部科学省 科学研究費補助金 新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成」[No. 25120010], 国際共同研究加速基金(国際共同研究強化) [No. 15KK0010], 国立研究開発法人 科学技術振興機構 戦略的創造研究推進事業 CRESTの支援のもとで行われた.

参考文献

- [1] D. Hassabis, D. Kumaran, C. Summerfield, M. Botvinick, “Neuroscience-Inspired Artificial Intelligence”, *Neuron*, 95, pp.245-258, (2017).
- [2] G. Kitagawa, “Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Model”, *Journal of Computational and Graphical Statistics*, Vol.5, No.1, pp.1-25, (1996).
- [3] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso”, *Journal of the Royal Statistical Society*, Vol.58, No.1, pp.267-288, (1996).
- [4] C. van Vreeswijk, H. Sompolinsky, “Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity”, *Science*, Vol.274, pp.1724-1726, (1996).
- [5] A. L. Hodgikin, A. F. Huxley, “A Quantitative Description of Membrane Current and Its Application to Conductance and Excitation in Nerve”, *J. Physiol*, Vol. 117, No. 4, pp.500-544, (1952).