

スパイクトリガー平均のスパース推定について

Sparse Estimation of Spike-Triggered Average

四倉 晋平*¹ 大森 敏明*² 永田 賢二*¹ 岡田 真人*¹
 Shimpei Yotsukura Toshiaki Omori Kenji Nagata Masato Okada

*¹東京大学大学院新領域創成科学研究科
 Graduate School of Frontier Sciences, The University of Tokyo

*²神戸大学大学院工学研究科
 Graduate School of Engineering, Kobe University

Spike-triggered average (STA) characterizes response properties of single neurons. In this study, we propose a method to estimate the STA by means of sparse linear regression with Fourier and polynomial basis functions. By sparse estimation with L1 regularization, we extract substantial basis functions for the STA in the proposed method. Using numerically simulated data, we show that the proposed method achieves more accurate estimation of the STA than simple trial average used in conventional method.

1. 研究背景と概要

神経細胞は入力された情報を、スパイクとよばれるパルス状の出力に符号化することにより、情報処理を行っている。スパイクを引き起こす入力時系列の統計的性質を推定する方法として、スパイクトリガー平均 (Spike-Triggered Average, STA) があり、生理実験に基づいた神経細胞の符号化特性の推定に用いられている [1]。しかし、実験で得られる STA のデータにはノイズが重畳され、正確に求めるのは困難である。

本研究では、スパース推定に基づいて STA を推定する方法を提案する。L1 正則化 [2] に基づいた線形回帰により、STA に対して本質的な基底関数を抽出するアルゴリズムを構成し、試行平均に基づく従来法と比べて、STA が精度良く推定できることを示す。

2. スパイクトリガー平均

神経細胞へ入力電流 $I(t)$ を以下のように与える。

$$I(t) = I_0 + \xi(t) \quad (1)$$

ここで I_0 は定常入力であり、 $\xi(t)$ は雑音入力である。 I_0 は、雑音入力 $\xi(t)$ が神経細胞に与えられない場合でも、スパイクが生じるほど十分大きいとする。図 1 に示すように神経細胞が時刻 t_k ($k = 1, 2, \dots, K$) でスパイクを生成するとする。このような状況で STA を、

$$C_0(\tau) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K \xi(t_k - \tau) \quad (2)$$

と定義する [3, 4]。STA はスパイクが起きる τ だけ前に、神経細胞に印加される雑音入力の平均値を表している。

生理実験で STA を求める際には、有限のスパイク数 K で打ち切って近似的に STA を求めるため、得られる STA のデータ点にはノイズが重畳され、神経細胞の性質を求めるのは困難である。

3. 提案法

生理実験により STA のデータ点 $\{(\tau_i, C(\tau_i))\}_{i=1 \dots N}$ が得られたとする。 $C_0(\tau)$ を基底関数 $\{f_j\}_{j=1 \dots M}$ と実数 $\{a_j\}_{j=1 \dots M}$

連絡先: yotsukura@mns.k.u-tokyo.ac.jp

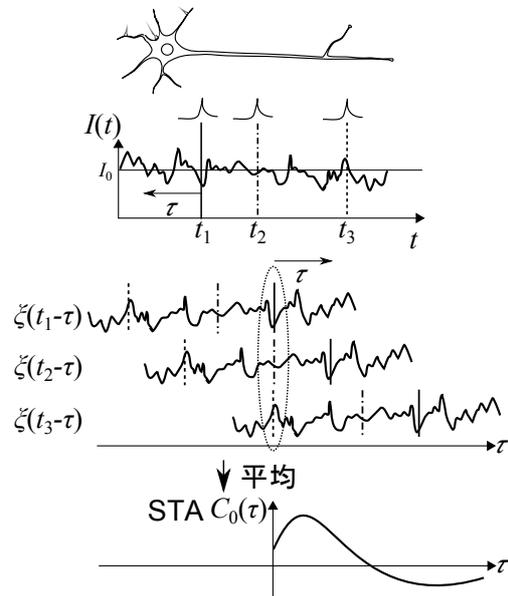


図 1: STA の概略図。神経細胞は入力を受け取り、スパイクを発生する。STA はスパイク発火時刻 $\{t_1, t_2, t_3, \dots\}$ を基準に雑音入力を整列し、発火時刻から τ だけ前の雑音入力を平均することで得られる。

を用い以下のように線形に回帰を行う。

$$C_0(\tau) = \sum_{j=1}^M a_j f_j(\tau) \quad (3)$$

式 (3) において係数 $\{a_j\}$ を STA のデータ点 $\{(\tau_i, C(\tau_i))\}$ に基づいて決定し、STA の回帰モデルを構成する。

式 (3) の回帰では、どのような基底関数を用いるかが回帰の精度を決定する。Galán らはフーリエ基底により、位相応答曲線とよばれる周期関数の線形回帰を行い、ノイズが重畳されたデータから位相応答曲線を推定するうえで有効であることが示されている [5]。また Ermentrout らは STA が PRC の微分を用いて表現されることを示した [4]。そのためフーリエ基底による回帰が STA にも同様に有効であることが期待される。し

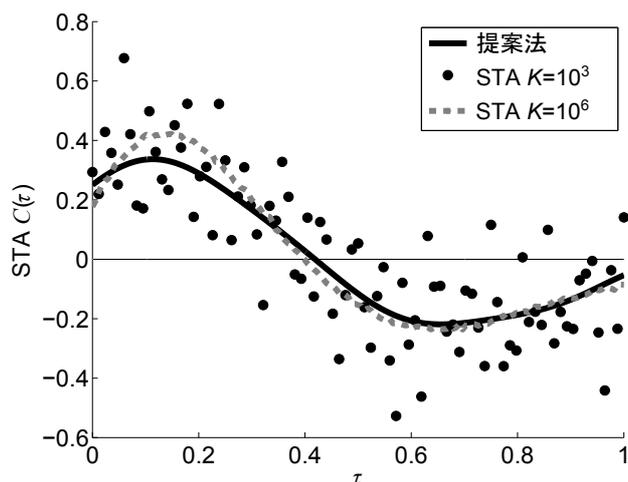


図 2: 数値実験で得た STA データに提案法を適用した結果. 点
はノイズが重畳された STA のデータ点 ($K = 10^3$) を表し, 実
線はデータ点に提案法を適用した推定結果をそれぞれ表す. ま
た破線はノイズをほとんど含まない STA のデータ ($K = 10^6$)
を示す.

しかし STA は周期関数であるが, $C_0(0) \neq C_0(1)$ であるため不連続性を持つ. そのためフーリエ基底のみを用いて回帰を行うと, 不連続性を表現するために STA に本来含まれていないはずの高周波成分が必要となることが予測される. そこで STA の不連続性を表現するための基底関数として, 多項式基底をフーリエ基底と合わせて用いる.

フーリエ基底と多項式を用いるうえで, どの次数までの基底を用いて回帰するのが良いのかは明らかではない. 線形回帰の際に冗長な基底関数を用いると, ノイズまで回帰モデルに組み込むため, オーバーフィットが生じ, 推定の精度が悪化することが知られている. したがって測定ノイズによる影響を受けない形で回帰モデルを得るため, STA に対して本質的な基底関数のみを抽出して回帰を行う必要がある. このような基底関数の抽出を行う方法として, L1 正則化がある [2]. L1 正則化は次の目的関数

$$E = \sum_{i=1}^N \left(C(\tau_i) - \sum_{j=1}^M a_j f_j(\tau_i) \right)^2 + \sum_{j=1}^M \lambda_j |a_j| \quad (4)$$

を最小化するような係数を選択する. 目的関数 E の第一項は回帰モデルと測定データの精度を測る項であり, 第二項は係数の絶対値が大きくなることを防ぐ重みづけ罰則項である. L1 正則化ではデータの回帰に適さない冗長な基底関数は, 対応する係数が 0 になる傾向が強い. 係数が 0 になっている基底は回帰に用いられなかったと考えることができる. そのため L1 正則化により, 回帰の誤差は減少させると同時に, 本質的な基底関数を抽出し適切な回帰モデルを選択するという, スパース推定の枠組みを実現できる.

L1 正則化を適用するうえで, 係数の絶対値に対する罰則の重み λ_j を決定する必要がある. STA のデータにはノイズが重畳されるため, フーリエ基底の高周波成分がノイズの影響により過剰に選択される可能性があると考えられる. そこで高周波のフーリエ基底が抽出されることに強い罰則をかける. d 次のフーリエ基底 $\cos 2\pi d\tau$, $\sin 2\pi d\tau$ に対応する重みは周波数に比

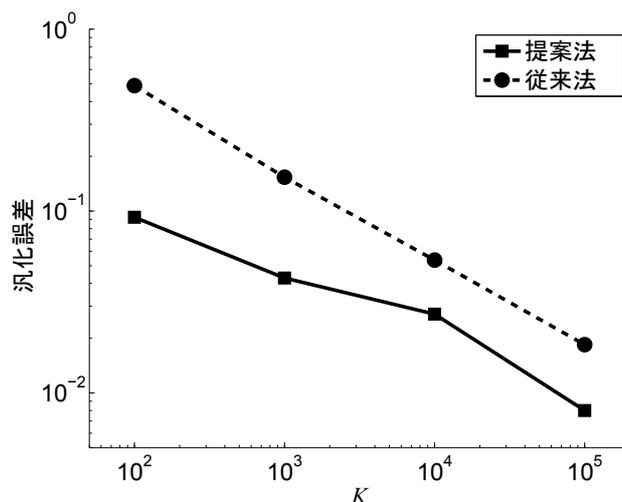


図 3: 従来法と提案法の汎化誤差を比較した図. 横軸はスパイク数 K を示し, 縦軸は汎化誤差を示している.

例させ $\lambda_j = d\lambda$ とする. また d 次の多項式基底 τ^d に対応する重みは $\lambda_j = \lambda$ とする. ただし λ は正の定数で, 10 分割交差検定を用いて決定する.

4. 実験結果

Morris-Lecar モデル [6] を用いた数値実験により, STA のデータ点を求め, 提案法を適用する. フーリエ基底は次数 25 次までの 50 個, 多項式基底は 50 次までを用いる.

図 2 に提案手法を適用した回帰の結果を示す. 図中の点はノイズが重畳された STA のデータ点 ($K = 10^3$) を表し, 実線はデータ点に提案法を適用した推定結果をそれぞれ表す. また破線はノイズをほとんど含まない STA のデータ ($K = 10^6$) を示す. 提案法を適用した結果, 多項式基底は 2 次の基底関数が 1 つ抽出され, フーリエ基底は最高 3 次までの合計 5 つの基底が抽出された. 多項式基底が選択されたことで, STA の不連続性 $C_0(0) \neq C_0(1)$ を表現できていることが分かる. また主要な低周波成分を上手く抽出することができ, 少数の基底で精度の良い回帰モデルが提案法により実現できている.

また試行平均に基づく従来法と提案法の汎化誤差を比較したグラフを図 3 に示す. 提案法による汎化誤差は試行回数が 10 倍の時の従来法と同程度の誤差となる結果が得られた. すなわち試行回数を 10 分の 1 に減らしても, 提案法を用いることで同じ精度で STA を求めることができる. そのため生理実験にかかるコストを大幅に削減できることが期待される.

参考文献

- [1] E. P. Simoncelli et al., The Cognitive Neurosciences, Gazzaniga (ed.), MIT Press, 2004.
- [2] R. Tibshirani, J. Royal. Statist. Soc. B, 1996.
- [3] T. Omori et al., Phys. Rev. E, 2010.
- [4] G. B. Ermentrout et al., Phys. Rev. Lett., 2007.
- [5] R. F. Galán et al., Phys. Rev. Lett., 2005.
- [6] C. Morris and H. Lecar, Biophys. J., 1981.