

神経回路モデルのダイナミクスのデータ駆動的解析

A Data-driven approach for analyzing dynamics of a neural network model

村田伸^{*1} 岡田真人^{*1*2}
Shin MURATA Masato OKADA

^{*1}東京大学新領域創成科学研究科
Graduate School of Frontier Sciences, The University of Tokyo

^{*2}理化学研究所脳科学総合研究センター
RIKEN Brain Science Institute

The associative memory model is a neural network model handling a retrieval of a memory pattern. Spurious states corresponding to failure in memory retrieval exists in the model storing many memory patterns. Analyzing the spurious state theoretically is known as a difficult problem. We introduced data driven approach to analyze the state, and found that a few dynamic modes explain the original high dimensional dynamics when synaptic depression is concerned.

1. 序論

連想記憶モデルは神経回路モデルの一種であり、複数の記憶パターンをニューロン間の結合強度として保持することが可能である。連想記憶モデルの磁性体との類似から、統計力学的な手法を用いて記憶容量などの性能が理論的に解析されている [Hopfield 82, Amit 85, Amit 87]。一方で、多数の記憶パターンを埋め込んだとき、偽記憶状態と呼ばれる記憶パターンと異なる準安定状態が生じることが知られている。偽記憶状態の理論解析は非平衡の統計力学を適用する必要があり困難である [Okada 96]。我々はデータ駆動的なアプローチに着目し、数値実験の結果から連想記憶モデルの性質を系統的に明らかにすることが可能であると考えた。

生理学実験から、シナプス前ニューロンの発火に応じて後ニューロンへの伝達効率が短時間的に減少する、シナプス抑圧という現象が発見されている [Abbott 97]。このシナプス抑圧を連想記憶モデルに組み込むことで、偽記憶状態が主に不安定化され、リミットサイクルが形成されることが主成分分析より示唆されている [Murata 14]。また、リミットサイクルは主に記憶パターンで張られるような部分空間上に存在することが確率的な主成分分析に基づく次元削減により明らかとなった。

一方で、主成分分析は各データが独立であることを仮定した手法であり、リミットサイクルのような時間相関を持つようなデータに対して、その動的な情報を抽出することは困難である。そこで本研究では、流体力学分野で発展している手法である、動的モード分解 (Dynamic Mode Decomposition, DMD) に着目する。DMD はデータを動的成分の足し合わせとして分解する手法である [Schmid 10]。また、各動的成分を l_1 正則化を用いてスパース推定することが可能である [Jovanović 14]。

本研究では、シナプス抑圧を導入した連想記憶モデルの偽記憶状態の数値シミュレーション結果に対して DMD を適用する。また、スパース推定を元に主要な動的成分を推定する。結果として、連想記憶モデルが少数の動的成分へと分解可能であることがわかった。

本予稿では、2章でシナプス抑圧を持つ連想記憶モデルを説明する。次に3章で本研究で用いた手法を紹介する。4章で数値実験の結果を説明する。5章で議論と今後の展望について述べる。

2. モデル

本研究では、 N 個のバイナリニューロン s_i が相互に結合するリカレントニューラルネットワークを考える。それぞれのニューロンは $0, 1$ の二値を取り、 $s_i = 1$ が発火状態に、 $s_i = 0$ が非発火状態に対応する。各ニューロン $s_i(t)$ は次式 (1) に従い確率的に時間発展する。

$$P[s_i(t+1) = 1] = 1 - P[s_i(t+1) = 0] = f[\beta h_i(t)] \quad (1)$$

ここで、 β は逆温度であり、 $f[\beta h] = (1 + \tanh(\beta h))/2$ とした。各ニューロンへの入力 $h_i(t)$ は、

$$h_i(t) = \sum_{j \neq i}^N J_{ij}(t) s_j(t) \quad (2)$$

となる。結合強度 $J_{ij}(t)$ は静的な結合強度 \tilde{J}_{ij} と、シナプスの伝達効率 $x_j(t)$ の積として表される。すなわち、

$$J_{ij}(t) = \tilde{J}_{ij} x_j(t) \quad (3)$$

である。静的な結合強度 \tilde{J}_{ij} はヘブ則を用いて、

$$\tilde{J}_{ij} = \frac{1}{N} \sum_{\mu=1}^{\alpha N} \xi_i^\mu \xi_j^\mu \quad (4)$$

と表される。ここで、 $\xi_i^\mu = 1$ は、 μ 番目の記憶パターンの i 番目のニューロンの発火を、 $\xi_i^\mu = -1$ は非発火を表す。このように結合強度を作成することで、各記憶パターン ξ^μ がネットワーク上の不動点として保持される。 α は記憶率と呼ばれ、ネットワークに保持した記憶パターン数のニューロン数に対する比率を表す。また、 $x_j(t)$ は次式に従い時間変化する。

$$x_j(t+1) = x_j(t) + \frac{1 - x_j(t)}{\tau} - U x_j(t) s_j(t) \quad (5)$$

ここで τ は伝達物質の回復の時定数であり、 U は伝達物質の放出量である。

また、数値実験を行うにあたり、初期状態は次式で確率的に与える。

$$P[s_i(0) = 1] = 1 - P[s_i(0) = 0] = \frac{1 + M_0^\mu \xi_i^\mu}{2} \quad (6)$$

このようにして生成された初期状態は、 $N \rightarrow \infty$ の極限で、 μ 番目の記憶パターンとのオーバーラップ、

$$M^\mu(t) = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu [2s_i(t) - 1] \quad (7)$$

が M_0^μ となる。初期オーバーラップ M_0^μ にはある値 M_c が存在し、 $M_0^\mu \geq M_c$ であればモデルは記憶状態へと収束する。

3. 手法

本章では DMD のアルゴリズムを簡単に説明する [Schmid 10, Tu 14]. また、動的成分を選択するためにスパース性を導入した DMD についても説明する [Jovanović 14].

3.1 DMD

DMD では、データが線形のダイナミクス

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t, \quad t = 0, 1, \dots, T-1 \quad (8)$$

に従い生成されると考える。ここで、 $\mathbf{X}_0 = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{T-1}]$, $\mathbf{X}_1 = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$ とおく。このとき、 $\mathbf{X}_1 = \mathbf{A}\mathbf{X}_0$ となる。 $\mathbf{X}_0 = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ と特異値分解し $\mathbf{A} = \mathbf{U}\mathbf{F}\mathbf{U}^*$ と考えると、

$$\begin{aligned} \mathbf{F}_{dmd} &= \arg \min_{\mathbf{F}} \|\mathbf{X}_1 - (\mathbf{U}\mathbf{F}\mathbf{U}^*)(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*)\|^2 \\ &= \mathbf{U}^* \mathbf{X}_1 \mathbf{V} \mathbf{\Sigma}^{-1} \end{aligned} \quad (9)$$

となる。この \mathbf{F}_{dmd} を固有値分解する。行列 \mathbf{X}_0 のランクを r , \mathbf{W} を固有ベクトルを並べた行列 $\mathbf{W} = (\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_r)$ とし、 \mathbf{D}_λ を、対角成分に固有値の並んだ対角行列 $\mathbf{D}_\lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$ とし、 $\mathbf{F}_{dmd} = \mathbf{W}\mathbf{D}_\lambda\mathbf{W}^{-1}$ とする。 t 番目のデータ \mathbf{x}_t は式 (8) から、

$$\mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 \quad (10)$$

$$= (\mathbf{U}\mathbf{F}_{dmd}\mathbf{U}^*)^t \mathbf{x}_0 \quad (11)$$

$$= \mathbf{U}\mathbf{W}\mathbf{D}_\lambda^t\mathbf{W}^{-1}\mathbf{U}^* \mathbf{x}_0 \quad (12)$$

と表される。 $\Phi = [\phi_1 \phi_2 \dots \phi_r] \equiv \mathbf{U}\mathbf{W}$, $\alpha \equiv \mathbf{W}^{-1}\mathbf{U}^* \mathbf{x}_0$ とおくと、

$$\mathbf{x}_t = \Phi \mathbf{D}_\lambda^t \alpha = \sum_{i=1}^r \alpha_i \lambda_i^t \phi_i \quad (13)$$

と分解できる。ここで α_i が i 番目のモードの係数、 ϕ_i が i 番目の DMD モード、 λ_i^t がモードの時間発展を表す。

実際に係数 α を決める際には、式 (13) から、

$$\begin{aligned} \alpha_{opt} &= \arg \min_{\alpha} \|\mathbf{X}_0 - \Phi \mathbf{D}_\alpha \mathbf{M}\|^2 \\ &= \arg \min_{\alpha} \|\mathbf{\Sigma}\mathbf{V}^* - \mathbf{W}\mathbf{D}_\alpha \mathbf{M}\|^2 \end{aligned} \quad (14)$$

として求める。ただし、 $\mathbf{D}_\alpha = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_r)$,

$$\mathbf{M} = \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{T-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{T-1} \end{pmatrix} \quad (15)$$

である。式 (14) を変形すると、

$$\alpha_{opt} = \arg \min_{\alpha} (\alpha^* \mathbf{P} \alpha - \mathbf{q}^* \alpha - \alpha^* \mathbf{q} + s) \quad (16)$$

ただし、 $\mathbf{P} = (\mathbf{W}^* \mathbf{W}) \circ (\overline{\mathbf{M}\mathbf{M}^*})$, $\mathbf{q} = \overline{\text{diag}(\mathbf{M}\mathbf{V}\mathbf{\Sigma}^* \mathbf{W})}$, $s = \text{trace}(\mathbf{\Sigma}^* \mathbf{\Sigma})$ である。したがって、 $\alpha_{opt} = \mathbf{P}^{-1} \mathbf{q}$ で求まる。

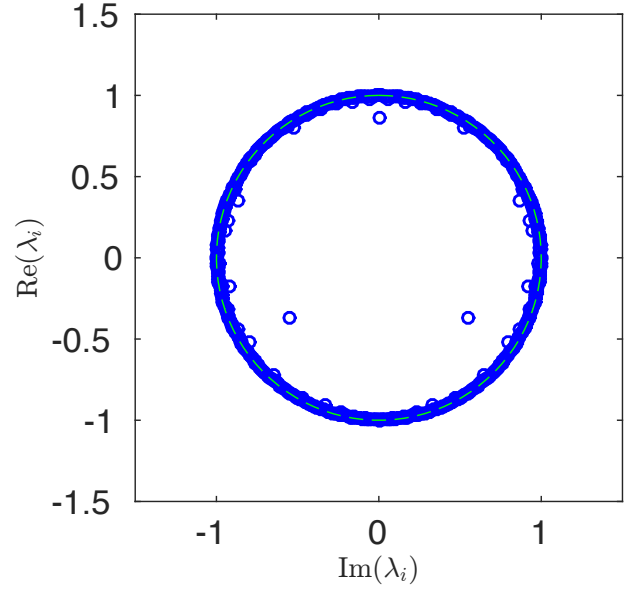


図 1: DMD の各モードに対応する固有値。横軸に各固有値の虚部を、縦軸に実部を取っている。実線は半径 1 の円を表す。

3.2 スパース DMD

式 (16) に ℓ_1 正則化項を加え、各モードの係数 α_i をスパースに刈り込むことを考える。すなわち、

$$\alpha_{sp} = \arg \min_{\alpha} [E(\alpha) + \gamma \|\alpha\|_1] \quad (17)$$

ここで、 $E(\alpha) \equiv \alpha^* \mathbf{P} \alpha - \mathbf{q}^* \alpha - \alpha^* \mathbf{q} + s$ であり、 γ は正則化項の重みを表す。式 (17) の最適化には交互方向乗数法を用いた [Boyd 10].

DMD のアルゴリズムはデータ \mathbf{x}_t と \mathbf{x}_{t+1} の間の線形のダイナミクスを仮定している。そこで、 $d_t = (\mathbf{x}_t, \mathbf{x}_{t+1})$, $t = 0, 1, \dots, T-1$ としてデータの組を考え、 $\{d_t\}_{t=0}^{T-1}$ を訓練データとテストデータに分割し交差検定法で γ の値を決める。

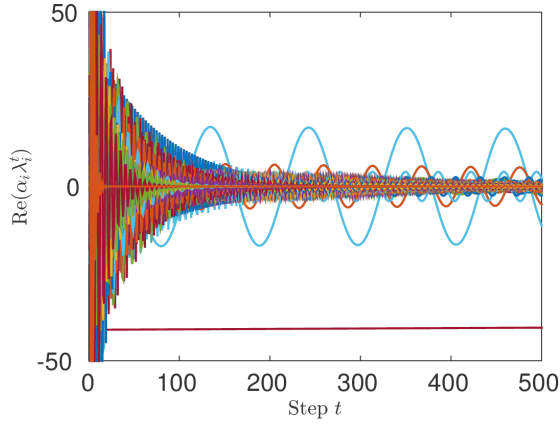
4. 結果

本研究で用いたパラメータは次の通りである。ニューロン数 $N = 5000$, 記憶率 $\alpha = 0.04$, 逆温度 $\beta = 10$, シナプス抑圧のパラメータ $\tau = 40$, $U = 0.0125$ としている。初期値、 $M_0^\mu = 0.1$ として、15000 ステップの数値実験を行った。DMD には最後の 625 ステップ分のデータを用いた。

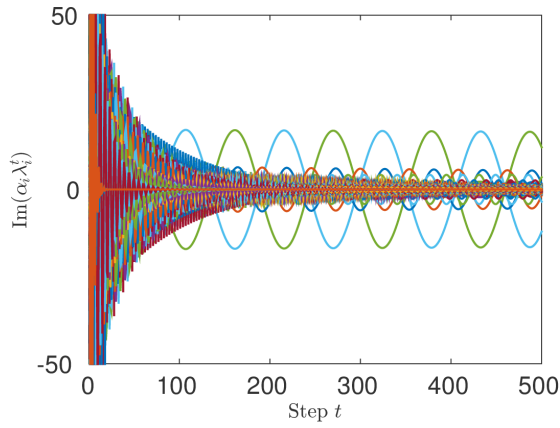
4.1 DMD による結果

図 1 は、横軸に各モードの固有値 λ_i の虚部を、縦軸に実部を取ったものである。点線で、半径 1 の円を示している。円周上の固有値に対応するモードは振動成分を表し円の内側に固有値に対応するモードは時間とともに減衰する。また、 $\lambda_i = 1$ に対応するモードは時間変化しない、平均値に対応する成分である。

図 2(a)(b) に、各モードの $\alpha_i \lambda_i^t$ の実部と虚部を示している。図から、指数的に減衰する成分と、ほとんど減衰しない振幅の大きい振動成分、時間変化しない成分に分解されていることがわかる。このうち、減衰成分や周波数の大きい成分、振幅の小さい成分などは本質的ではないと考えることができる。そこで、3.2 節で説明したスパース DMD を用いて、不要な成分を刈り込む。



(a) 実部.



(b) 虚部.

図 2: 時間成分と係数の積 $\alpha_i \lambda_i^t$ の変化.

4.2 スパース推定の結果

DMD モード Φ 並びに、時間成分 M を 3.1 節で説明した方法で推定したのち、係数 α を 3.2 節で説明した手法でスパース推定する。スパース性の強さ γ を 10^{-1} から 10^5 まで、等比数列で 100 個用意する。それぞれの γ の値を固定して式 (17) を最小化するように求める。その際に、5 分割交差検証法でテスト誤差の値を求め、 γ を決めることを考える。

図 3 にスパース性の強さ γ と、非ゼロ要素の数との関係を赤実線で、テスト誤差との関係を青実線で示す。破線が平均テスト誤差最小の γ_{min} を表し、一点鎖線が平均テスト誤差最小から 1 標準偏差内で最大の γ_{sparse} を表す。このとき、非ゼロ要素数の中央値はそれぞれ 11 個、3 個であった。テスト誤差の変化が大きい点が二つ存在している。一つは、振動成分を全て刈り込み、時間変化しない成分だけで元のデータを再現した場合である。もう一つは全ての成分を刈り込んだ点である。それらの点を除くと、非ゼロ要素数が減少してもテスト誤差の増加はほとんどないことがわかる。

図 4(a)(b) は、 γ_{sparse} での $\alpha_i \lambda_i^t$ の値を実部、虚部ごとに示したものである。図 2 と比較すると、指数的に減衰する成分や周期の短い成分、振幅の小さな成分がスパース推定で刈り込まれ、時間変動しない成分と、振幅の大きい振動成分が残っていることがわかる。

これらの結果から、DMD とそのスパース推定を用いてデータから主要な振動モードと平均値を抽出可能なことがわかつ

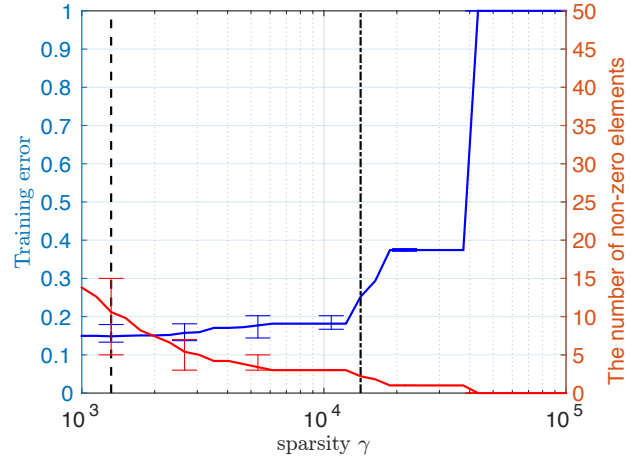


図 3: 交差検定法の結果. スパース性の強さ γ と、非ゼロ要素の数、テスト誤差の関係を示している。破線が平均テスト誤差最小の γ を表し、一点鎖線が平均テスト誤差最小から 1 標準偏差内で最大の γ を表す。エラーバーは最小・最大を表す。

た。その結果、主成分分析による次元削減で示唆されたように、シナプス抑圧を持つときの偽記憶状態が、実際にリミットサイクルのように振動的に振舞っていることがわかった。

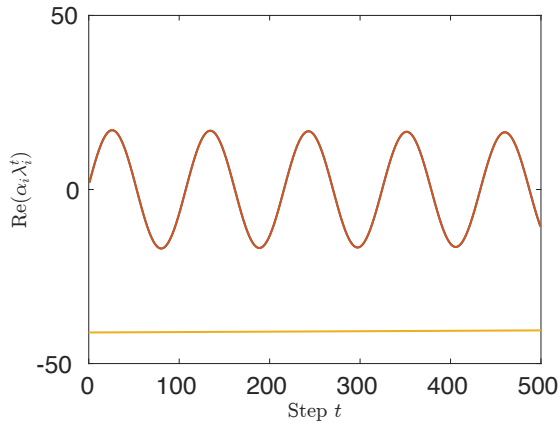
5. 議論・結論

本研究は、神経回路モデルのダイナミクスをデータ駆動的アプローチで解明することを目的とし行った。連想記憶モデルにおいて、シナプス抑圧を導入することで偽記憶状態にリミットサイクルが形成されることが、先行研究での主成分分析の結果から示唆されていた。しかしながら主成分分析はデータが独立であることを想定しているため、時間構造を考慮しない解析となっていた。本研究では、DMD という手法を用いて、データを複数の動的成分に分解可能なことを示した。さらに、 ℓ_1 正則化を用いて、実際に必要な少数の動的成分を抽出し、それらが振動成分と平均値であることが示唆され、ネットワークがリミットサイクルのように運動していることがわかった。

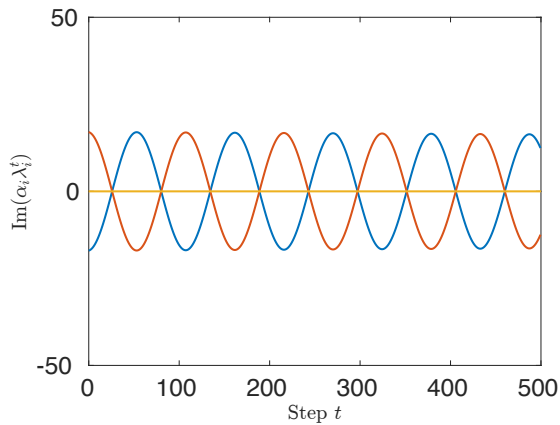
連想記憶モデルにおいて、温度やシナプス抑圧の強さを変化させることで、ネットワークの平衡状態が変わることが知られている。今後の課題として、モデルパラメータの変化が DMD 解析の結果とどのように変化するか、系統的に調べることが挙げられる。また、時間的な性質の推定だけでなく、空間的な性質の推定を行うために、メキシカンハットのような結合を持つニューラルネットワークに適應することなども考えられる。

参考文献

- [Abbott 97] Abbott, L., Varela, J., Sen, K., and Nelson, S. B.: Synaptic depression and cortical gain control, *Science*, Vol. 275, No. 5297, pp. 221–224 (1997)
- [Amit 85] Amit, D. J., Gutfreund, H., and Sompolinsky, H.: Spin-glass models of neural networks, *Physical Review A*, Vol. 32, No. 2, pp. 1007–1018 (1985)
- [Amit 87] Amit, D. J., Gutfreund, H., and Sompolinsky, H.: Statistical mechanics of neural networks near saturation, *Annals of Physics*, Vol. 173, pp. 30–67 (1987)



(a) 実部.



(b) 虚部.

図 4: スパース推定された時間成分と係数の積 $\alpha_i \lambda_i^t$ の変化.

[Boyd 10] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J.: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, *Foundations and Trends in Machine Learning*, Vol. 3, No. 1, pp. 1–122 (2010)

[Hopfield 82] Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the national academy of sciences*, Vol. 79, No. 8, pp. 2554–2558 (1982)

[Jovanović 14] Jovanović, M. R., Schmid, P. J., and Nichols, J. W.: Sparsity-promoting dynamic mode decomposition, *Physics of Fluids*, Vol. 26, No. 024103, pp. 1–22 (2014)

[Murata 14] Murata, S., Otsubo, Y., Nagata, K., and Okada, M.: Oscillations in Spurious States of the Associative Memory Model with Synaptic Depression, *Journal of the Physical Society of Japan*, Vol. 83, No. 124004, pp. 1–8 (2014)

[Okada 96] Okada, M.: Notions of Associative Memory and Sparse Coding, *Neural Networks*, Vol. 9, No. 8, pp. 1429–1458 (1996)

[Schmid 10] Schmid, P. J.: Dynamic mode decomposition of numerical and experimental data, *Journal of Fluid Mechanics*, Vol. 656, pp. 5–28 (2010)

[Tu 14] Tu, J. H., Rowley, C. W., Luchtenburg, D. M., Brunton, S. L., and Kutz, J. N.: On Dynamic Mode Decomposition: Theory and Applications, *Journal of Computational Dynamics*, Vol. 1, No. 2, pp. 391–421 (2014)