

# リザーコンピューティングに基づくてんかん発作判別

## Epileptic Seizure Detection Based on Reservoir Computing

岸本 勇太 (Yuta Kishimoto)<sup>1</sup>, 行木 孝夫 (Takao Namiki)<sup>2</sup>

<sup>1</sup> 北海道大学 大学院理学院 (Graduate School of Science, Hokkaido University), <sup>2</sup> 北海道大学  
大学院理学研究院数学部門 (Faculty of Science, Hokkaido University)  
e-mail : kishimoto.yuta.b8@elms.hokudai.ac.jp

### 1 概要

てんかんは脳内の異常な電気活動による反復的な発作を特徴とする神経疾患で、世界中で約 5,000 万人が患っている最も一般的な神経疾患のひとつである。てんかん発作は、意識喪失、異常運動、記憶障害など多様な症状を引き起こし、予測が困難なため、患者の日常生活に大きな影響を与える。さらに、てんかん患者の約 30% は標準的な抗てんかん薬では発作を抑制できない「難治性てんかん」と診断される。てんかん発作を引き起こす神経活動の異常が最初に発生する脳の特定の部位をてんかん焦点といい、この部位で発生した異常な電気信号が、周囲の脳領域に広がることで発作が引き起こされる。難治性てんかんの患者には、てんかん焦点の正確な特定に基づく外科手術が検討される。

先行研究 [1] では、Long Short Term Memory (LSTM) を活用した発作前期と発作間欠期の判別モデルが提案された。LSTM モデルは、時系列データの判別において高い性能を発揮する一方で、複雑な内部構造ゆえに学習には多くの時間とメモリを要するという問題がある。

てんかん発作の判別や予測において学習時間の短縮は重要な課題であるため、本講演では、計算コストの小さい機械学習手法として知られるリザーコンピューティングを用いたてんかん発作判別およびてんかん焦点領域の同定手法について紹介する。

### 2 手法

本研究における発作判別手法は、以下の 4 つのステップで構成される。

- 1) 脳波データの特徴量抽出
- 2) 訓練データを用いた ESN モデルのハイパーパラメータ調整
- 3) テストデータの発作判別
- 4) 性能評価

まず、脳波データを 1 秒 (2000 ステップ) ごとにセグメント分割し、各セグメントから抽出した特徴量を用いて ESN モデルの入力ベクトルを構成した。出力データは発作時に 1、非発作時に 0 をとる 2 クラスラベルである。次に、訓練データを用いて ESN モデルのハイパーパラメータを最適化し、テストデータを使用して発作判別を行った。評価指標には混合行列から得られる感度、特異度、精度の 3 指標を用いた。Leave-One-Out 交差検証により得られる各指標の平均値を ESN モデルの判別性能として評価する。

### 3 リザーコンピューティング

リザーコンピューティングはリカレントニューラルネットワーク (RNN) から派生して生まれた機械学習の枠組みのひとつである。リザーコンピューティングは、中間層から出力層への結合荷重  $W^{out}$  の学習に限定し、従来の RNN が抱えていた学習上の課題を解消する。

ESN モデルは, Jaeger によって 2001 年に提案されたリザバーコンピューティングの代表的なモデルの一つである [2]. 入力層・リザバー層・出力層の 3 層から構成される RNN であり, 非線形ダイナミクスを持つリザバー層を用いて時系列データの予測や判別を行うことが可能である. 時刻  $n$  における  $N_u$  次元の入力ベクトル,  $N_x$  次元のリザバー状態ベクトル,  $N_y$  次元の出力ベクトルをそれぞれ  $u(n)$ ,  $x(n)$ ,  $y(n)$  とする. リザバー層の状態は, 入力層からの情報と前時刻のリザバー状態および出力層からのフィードバックによって更新される. その後, 出力層はリザバー層を用いて更新が行われる. リザバー層, 出力層の更新式は以下の式で定義される.

$$x(n+1) = (1 - \alpha)x(n) + \alpha \tanh(W^{\text{in}}u(n+1) + W^{\text{res}}x(n)) \quad (1)$$

$$y(n+1) = W^{\text{out}}x(n+1) \quad (2)$$

ここで,  $W^{\text{in}} \in \mathbb{R}^{N_x \times N_u}$  は, 入力層からリザバー層への入力結合重み行列,  $W^{\text{res}} \in \mathbb{R}^{N_x \times N_x}$  は, リザバー層のニューロン同士のリカレント結合重み行列, また,  $\alpha \in [0, 1]$  はリーク率と呼ばれる定数である. リーク率が小さいほど現在の状態に過去の状態が強く反映されるため, 長期依存性の高い時系列信号の分析が可能になる.  $W^{\text{out}} \in \mathbb{R}^{N_y \times N_x}$  は, リザバー層から出力層への結合重み行列である. それぞれの結合重み行列には, 結合強度および結合密度が設定される.

## 4 主結果

難治性てんかん患者の発作識別における ESN モデルの有効性を評価し, 以下の成果を得た.

- 1) ESN モデルは高い感度および特異度を示し, 特にてんかん焦点電極では 90% を超える高感度で発作を検出できることが確認された. 一方, 非焦点電極では発作時の特徴を十分に捉えることができず, 感度が低下する傾向が見られた.
- 2) 各電極における発作検出感度を分析した結果, ESN モデルはてんかん焦点に近い電極で顕著に高い感度を示し, 焦点から離れるにつれて感度が低下することが分かった. このことから, 感度の分布が焦点の特定に有用である可能性が示唆された.

謝辞 本研究にあたり, 脳波データをご提供くださった京都大学大学院医学研究科てんかん・運動異常生理学講座の池田先生, 松橋先生に深く感謝いたします. また, 有益な議論を交わしていただいた札幌市立大学 AIT センターの津田先生にも心より感謝いたします. なお, 本研究は JST 次世代研究者挑戦的研究プログラム (JPMJSP2119) の支援を受けて実施されたものです.

## 参考文献

- [1] Tsiouris, Kostas M., et al. "A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals." *Computers in biology and medicine* 99 (2018): 24-37.
- [2] Jaeger, Herbert. "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note." Bonn, Germany: German National Research Center for Information Technology GMD Technical Report 148.34 (2001): 13.

# マルコフ連鎖準モンテカルロ法によるベイズ計算と高次元問題

## High-dimensional Bayesian inference using Markov chain quasi-Monte Carlo

原瀬 晋 (Shin Harase)<sup>1</sup>

<sup>1</sup> 立命館大学 (Ritsumeikan University)

e-mail : harase@fc.ritsumei.ac.jp

### 1 概要

マルコフ連鎖モンテカルロ (MCMC) 法を用いた期待値計算を考える。乱数を用いた計算は収束が非常に遅いため、より高い一様性を有した準乱数に置き換えて高速化を図る準モンテカルロ法を適用したい。しかるに、従来の準乱数はそのまま適用することが出来ない。2000 年代に入って、Owen ら [1] により、CUD 列 (completely uniformly distributed sequence) と呼ばれる点列を用いた際に期待値計算の一致性が示された。CUD 列とは、1 次元無限点列  $u_0, u_1, u_2, \dots \in (0, 1)$  について、すべての次元  $s \geq 1$  に対して、連続した  $s$  個の組

$$(u_l, \dots, u_{l+s-1}), l = 0, 1, 2, \dots \quad (1)$$

が一様分布している点列  $\{u_l\}$  のことである。ここで、CUD 列となっている準乱数の計算機上の実装が問題となる。Owen ら [1] は短い周期の擬似乱数発生法を準備して、1 周期使い切った際に現れる格子構造を準乱数と見なして、CUD 近似点列として用いる方法を提案した。原瀬 [2, 3] は、この枠組みにおいて、線形フィードバックシフトレジスタ型擬似乱数発生法に着目し、次元  $s = 1, 2, 3, \dots$  の順番に高い一様性をもったパラメータを探索して、新しい CUD 近似点列を開発した。本講演では、このような準乱数がベイズ統計の高次元問題に威力を発揮する場合があることを紹介したい。

### 2 CUD 近似点列と一様性の評価指標 $t$ -値

原瀬 [2, 3] によって探索された CUD 近似点列について簡単に紹介する。ある自然数  $m$  に対して  $N = 2^m$  とおく。1 次元点列  $\{u_l\}$  は周期  $N - 1$  をもつ、すなわち、 $u_0, u_1, \dots, u_{N-2}, u_{N-1} = u_0, \dots \in (0, 1)$  となっているものとする。このとき、性質 (1) にすり合わせるため、 $s$  次元の点  $\mathbf{u}_0 = (u_0, \dots, u_{s-1}), \mathbf{u}_1 = (u_1, \dots, u_s), \dots, \mathbf{u}_{N-2} = (u_{N-2}, u_0, \dots, u_{s-2})$  を作る。原点  $\{\mathbf{0}\}$  を加えた点集合を  $P_s = \{\mathbf{0}\} \cup \{\mathbf{u}_l\}_{l=0}^{N-2}$  とおく。これに対して、準乱数の一様性の尺度として  $t$ -値と呼ばれる非負整数を考える。 $t$ -値が小さいほど一様性が高く、 $t$ -値が 0 のとき最適値となる。

**定義 1 ( $t$ -値)**  $s \geq 1, t$  を  $0 \leq t \leq m$  を満たす整数とする。 $N = 2^m$  個からなる  $s$  次元点集合  $P_s \subseteq [0, 1)^s$  が、体積が  $2^{t-m}$  となる  $[0, 1)^s$  内のすべての 2 進部分区間  $E = \prod_{j=1}^s [a_j/2^{c_j}, (a_j+1)/2^{c_j})$  にちょうど  $2^t$  個ずつ含まれるときに、 $P_s$  は  $(t, m, s)$ -net であるという。ただし、各  $1 \leq j \leq s$  に対して、 $c_j, a_j$  は  $c_j \geq 0, 0 \leq a_j < 2^{c_j}$  となる整数とする。この性質を満たす最小の  $t$  を  $t$ -値という。

原瀬 [2] は次元  $s = 1, 2$  で  $t$ -値 0、次元  $s \geq 3$  においても小さい  $t$ -値をもつ点列  $\{u_l\}$  を実装した。

### 3 プロビットモデルのギブスサンプリング

高次元問題として、二項プロビットモデル  $P(Y_i = 1) = \Phi(\mathbf{x}_i^\top \boldsymbol{\beta}), i = 1, \dots, n$  のベイズ推定を扱う。 $Y_i \in \{0, 1\}$  および  $\mathbf{x}_i, \boldsymbol{\beta} \in \mathbb{R}^k$  は各々 2 値応答変数、説明変数ベクトル、回帰係数ベクトルであ

パラメータ	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$
乱数	1.23e-05	1.15e-09	3.62e-08	1.70e-08	1.70e-11	3.55e-09	7.91e-07	8.80e-08
準乱数 [2]	1.09e-09	6.15e-14	2.01e-12	1.06e-12	1.13e-15	3.94e-13	1.49e-10	7.05e-12

表 1. 推定値の不偏分散： $N = 2^{16}$  個の乱数・準乱数により事後平均を 300 回計算

る。いま、 $Y_i$  を不完全データとみなして、潜在変数  $Z_i$  を導入して

$$Z_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \epsilon_i \quad (\epsilon_i \sim \mathcal{N}(0, 1)),$$

$$Y_i = \mathbf{1}_{Z_i > 0}$$

とおく。計画行列  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  は階数  $k < n$  (フルランク) とし、 $\boldsymbol{\beta}$  に無情報事前分布を仮定する。このとき、条件付き事後分布からギブスサンプリングにより  $Z_1, \dots, Z_n$  と  $\boldsymbol{\beta}$  を発生させる：

$$Z_i | \boldsymbol{\beta}, \text{data} \sim \begin{cases} \mathcal{N}(\mathbf{x}_i^\top \boldsymbol{\beta}, 1) | Z_i > 0 & \text{if } Y_i = 1, \\ \mathcal{N}(\mathbf{x}_i^\top \boldsymbol{\beta}, 1) | Z_i \leq 0 & \text{if } Y_i = 0, \end{cases}$$

$$\boldsymbol{\beta} | \mathbf{Z}, \text{data} \sim \mathcal{N}((\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Z}, (\mathbf{X}^\top \mathbf{X})^{-1}), \quad \mathbf{Z} = (Z_1, \dots, Z_n)^\top.$$

状態ベクトル  $(\boldsymbol{\beta}^\top, Z_1, \dots, Z_n)^\top$  は  $(k + n)$  次元となり、標本サイズ  $n$  が増えると次元が増加する。

## 4 数値例

Mroz[4] の女性の労働供給データセットに適用する ( $n = 753, k = 8$ )。これは 753 人の既婚女性について働いているかを調べたデータである。応答変数は INLF (女性が働いていれば 1, 働いていなければ 0) であり、教育年数や年齢など 8 つの説明変数を用いたプロビットモデルをベイズ推定する：

$$P(\text{INLF} = 1) = \Phi(\beta_0 + \beta_1 \text{NWIFINC} + \beta_2 \text{EDUC} + \beta_3 \text{EXPER} \\ + \beta_4 \text{EXPER2} + \beta_5 \text{AGE} + \beta_6 \text{KIDLT6} + \beta_7 \text{KIDGE6}).$$

乱数による 5000 回の burn-in の後に  $N = 2^{16}$  個の乱数・準乱数を用いて  $\boldsymbol{\beta} = (\beta_0, \dots, \beta_7)^\top$  の事後平均  $E[\beta_j]$  を推定する。初期値を 300 回変えて、推定値の不偏分散を計算した結果が表 1 である。

761 次元の問題に対して準モンテカルロ法が効果を発揮する一つの理由として、 $\boldsymbol{\beta}$  のサンプリングの他に、潜在変数  $Z_i$  を発生させるステップにおいて独立な 1 次元サンプリングを繰り返し行っており、この部分において準乱数のもつ一様性が効いて全体の収束性が向上した可能性がある。

**謝辞** 本研究は JSPS 科研費 22K11945 の助成を受けたものです。

## 参考文献

- [1] A. B. Owen and S. D. Tribble, A quasi-Monte Carlo Metropolis algorithm, Proc. Natl. Acad. Sci. USA, 102(25):8844–8849, 2005.
- [2] S. Harase, A table of short-period Tausworthe generators for Markov chain quasi-Monte Carlo, J. Comput. Appl. Math. 384 (2021), 113136, 12 pp.
- [3] S. Harase, A search for short-period Tausworthe generators over  $\mathbb{F}_b$  with application to Markov chain quasi-Monte Carlo, J. Stat. Comput. Simul. 94 (2024), no. 9, 2040–2062.
- [4] T. Mroz, The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions, Econometrica, vol. 55, no. 4 (1987): 765–99.

# 加重平均下での多腕バンディット問題におけるリグレット上限の短期間での改善

## An improvement of upper bounds of the regret in multi-armed bandit problem with weighted averages in short intervals

加藤寛之 (Hiroyuki Kato)

嘉悦大学 経営経済学部 (Kaetsu University, Department of Management and Economics)

e-mail : hiroyuki-kat0@kaetsu.ac.jp

### 1 概要

異なる確率で報酬がでる複数のスロットマシンの中から一つずつを、それらの確率が分からない状態で多期間に渡って選びながら、最大の報酬を得ようとする問題を一般に多腕バンディット問題という。最大の確率のマシンを選び続ける時の報酬と、ある方策の下で選ぶ時の報酬の差をリグレットといい、この期待値をできるだけ小さくする方策を考える問題となる。ここまでに多くの研究があり、選択回数  $N$  に対し、 $A \log N + B$  のような形の上限を持つことが分かっている。本研究は、UCB として知られる‘報酬平均 +  $\log N$  の関数が最大となるマシンを選ぶ’という方策を一般化し、適当な加重列を考え‘報酬の加重平均 +  $\log N$  の関数が最大となるマシンを選ぶ’ことで、リグレット上限を短期間で低めにする方策を考える。シミュレーションによると、初期の上限を下げる加重列のもとでは、期間を経るにつれてリグレットが改善される例が観測された。

### 2 設定

$\mathbb{N}$  を自然数の集合、 $K$  をスロットマシンの数とし、一般性を失うことなく  $K = 3$  とする。スロットマシンは報酬 1 か 0 を出すものとする。 $X_\lambda, Y_\lambda, Z_\lambda \in \{0, 1\}$ 、 $\lambda \in \mathbb{N}$  を適当な確率空間上の独立な確率過程とし、それぞれマシン  $X, Y, Z$  と名付ける。 $Z$  を最も高い確率で報酬が出るマシンとし、 $E[Z] > \max\{E[X], E[Y]\}$  とする。 $N \in \mathbb{N}$  をマシンを回した総回数とする。 $l \in \mathbb{N}$  を

$$l := \min \left\{ i \in \mathbb{N} \mid i \geq \frac{8 \log N}{(E[Z] - E[X])^2} \right\}.$$

とする。 $0 < \bar{\varepsilon} < 1$  とし、加重列  $\varphi : \mathbb{N} \rightarrow (0, 1)$  を以下のように定義する；

$$\sum_{\lambda=1}^i \varphi(\lambda) \geq \frac{8 \bar{\varepsilon} \log N}{(E[Z] - E[X])^2}, \quad i \geq l.$$

$0 < \varepsilon \leq \bar{\varepsilon}$ 、 $X$  を回した回数  $i \in \mathbb{N}$  に対し、 $X$  の指標を以下のように定義する；

$$\mathbb{X}_i^N := \frac{\sum_{\lambda=1}^i \varphi(\lambda) X_\lambda}{\sum_{\lambda=1}^i \varphi(\lambda)} + \sqrt{\frac{2 \varepsilon \log N}{\sum_{\lambda=1}^i \varphi(\lambda)}},$$

$Y, Z$  の指標、 $\mathbb{Y}_j^N, \mathbb{Z}_k^N$  も同様に定義する。方策 (policy) を、以下とする：最初に全てのマシンを 1 回ずつ回し、その後は、 $X, Y, Z$  の指標のうち、最大の値のマシンを回すことにする。 $X, Y$  が回される回数 (確率変数) をそれぞれ  $T_X(N), T_Y(N)$  とし、以下をリグレットと定義する；

$$R_N := (E[Z] - E[X])T_X(N) + (E[Z] - E[Y])T_Y(N).$$

この期待値  $E[R_N]$  の上限をできるだけ小さくする問題を考える。

### 3 結果

$T_X(N)$  の期待値をとると、以下の上限を持つことが示せる；

$$E[T_X(N)] \leq \frac{8 \log N}{(E[Z] - E[X])^2} + 1 + 2 \sum_{n=1}^N \sum_{i=1}^n \left( \frac{1}{K + n - 1} \right)^{4\varepsilon\Phi(i)-1},$$

但し、

$$\Phi(i) := \frac{\sum_{\lambda=1}^i \varphi(\lambda)}{\sum_{\lambda=1}^i \varphi^2(\lambda)}.$$

主要な先行研究である Auer et al. (2002) では、 $\varepsilon\Phi(i) \equiv 1$  である一方で、本研究では、 $\varepsilon\Phi(i) > 1$  を探る。仮に  $\varphi(i) \equiv \bar{\varepsilon}$  とすると、 $\Phi(i) = 1/\bar{\varepsilon}$  となり、 $\varepsilon\Phi(i) > 1$  は不可能となるため、比重を変化させる加重列を考える。以下のような、 $\Phi$  を一時的に大きく取れるよう、小さめの加重列  $\varphi$  と  $\varepsilon$  を考える。上限の級数部分は収束するよう、 $\lim_{i \rightarrow \infty} (4\varepsilon\Phi(i) - 2) > 1 + h$ ,  $h > 0$  となるようにしておく。

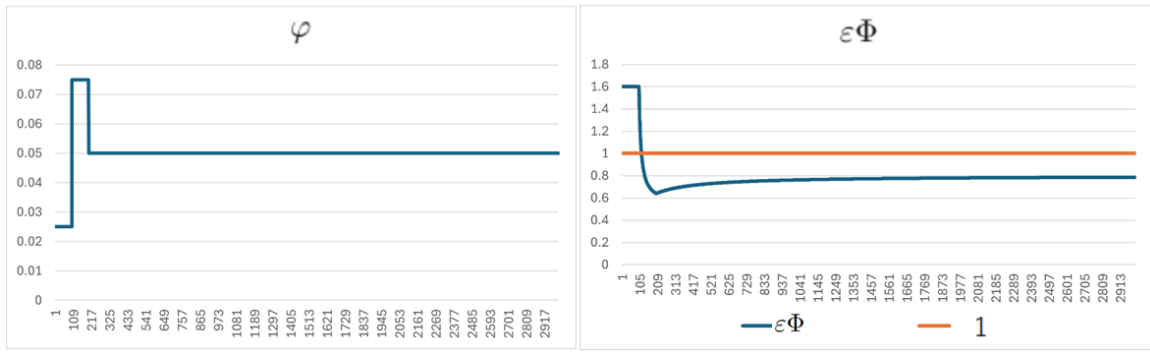


図 1. 左図は加重列  $\varphi$ 。右図はその時の  $\varepsilon = 0.04$  とした時の  $\varepsilon\Phi$

この加重列に基づいていくつかのパターンの乱数から、 $N = 3000$  までシミュレーションを行った。30 試行の平均をとった。以下に見られるように、上限を抑えた分はむしろ探索が行われ、リグレットの期待値は一時的に大きくなり、その後、リグレットの期待値は小さくすることができている。

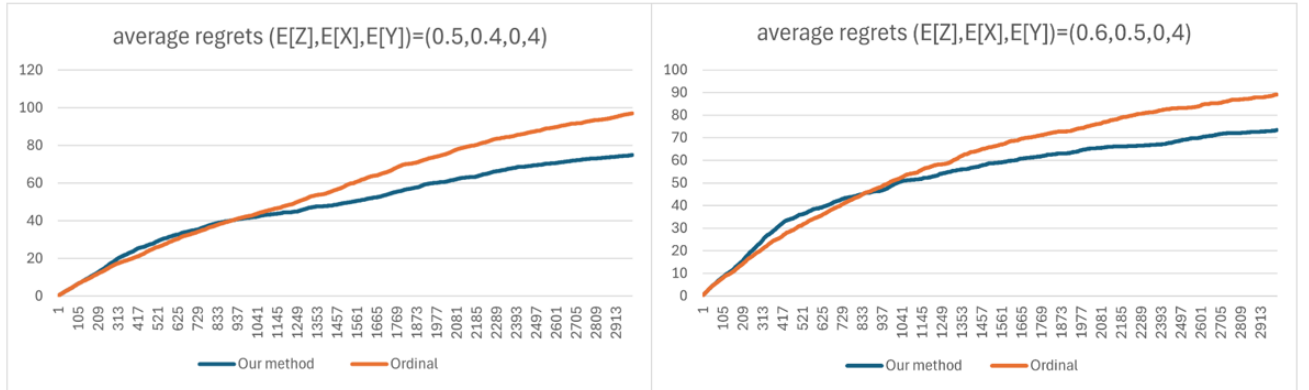


図 2. 青色のグラフが本研究によるリグレット期待値で、オレンジは従来によるもの。

### 参考文献

- [1] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. Machine Learning, 47(2) (2002), 235–256.

# Neural Sheaf Diffusion の普遍近似定理と交通予測への応用

## Neural Sheaf Diffusion for Regression on Graphs: Approximation Theorem and Applications to Traffic Forecasting

軽部 友裕 (Tomohiro Karube)<sup>1</sup>, 矢野 良輔 (Ryosuke Yano)<sup>2</sup>,

<sup>1</sup> 東京大学 数理科学研究科

(Graduate School of Mathematical Sciences, the University of Tokyo),

<sup>2</sup> 東京海上ディーアール 株式会社 (Tokio Marine dR Co. Ltd.),

e-mail : karube-tomohiro803@g.ecc.u-tokyo.ac.jp

### 1 はじめに

グラフ構造をもつデータの解析において、グラフニューラルネットワーク (GNN) は近年、画像認識、交通予測、分子構造解析など幅広い応用で注目を集めている。しかし、GNN の層を深くすることで全てのノードの特徴量が平均化される **oversmoothing 現象**は、GNN の表現力を制限する要因の一つである。この原因は、GNN に現れる行列の固有値分布に起因すると考えられている [1]。

このような背景のもと **Neural Sheaf Diffusion (NSD)** は、グラフ上の層 (sheaf) という数学的構造を用いることで、これらの問題に対処する [2]。NSD では、各辺にスカラー重みではなく線形写像を割り当てることができ、従来の GNN に比べて高い表現力が期待されている。

本研究では、この NSD に対して以下の 2 つの観点から新たな知見を提供する：

- NSD が普遍近似定理 (universal approximation theorem) を満たすことの証明
- NSD の回帰タスクへの応用と実証：交通速度予測への適用

このうち理論面では、NSD の変種である NSD-b に注目し、その近似能力に関する定理を証明する。これは、NSD が oversmoothing の影響下でも十分な表現力を持つことの理論的裏付けとなる。

また応用面では、PeMSD7 に対して NSD を適用し、時間発展する交通速度を高精度に予測できることを示す。特に、RNN や LSTM のような時系列モデル [3, 4] と同様の交通速度の予測精度を達成した。本講演では、NSD の理論的な有効性と実問題への応用可能性の双方を明らかにする。

### 2 グラフ上の層と NSD

近年、グラフ上のデータ構造を捉える手法として、層 (sheaf) という数学的概念が注目されている。層はもともと代数幾何学や微分幾何学などの分野で、空間上の構造を記述・抽象化するための枠組みである。これをグラフに応用したものがグラフ上の層であり、各ノードやエッジにベクトル空間や線形写像を割り当てる (図 1)。こうした枠組みに基づいて設計されたモデルが Neural Sheaf Diffusion (NSD) である。これはグラフ  $G$  上の層  $F$  に対して、 $F$  のラプラシアン行列を導入し、熱拡散方程式を用いて情報伝播を行う。本研究では、NSD の変種である NSD-b に注目する。

**定理 1** グラフ  $G = (V, E)$  とコンパクト部分集合  $I \subset \mathbb{R}$ 、 $f: I \rightarrow \mathbb{R}$  が与えられる。  $\text{ex}: \mathbb{R}^V \rightarrow \mathbb{R}^E$  を次のように定義する。

$$(\text{ex}(b_v)) = \sum_{e=(u,v) \in E} b_e$$

ただし、 $b_v$  は  $\mathbb{R}^V$  の標準的な基底で、 $b_e$  は  $\mathbb{R}^E$  の標準的な基底である。このとき、 $f$  が  $\text{ex}$  を経由すれば、 $f$  は NSD-b で表現できる。

### 3 交通予測への応用

本研究では、実データセットである PeMSD7 を用いて評価を行った。PeMSD7 は、カリフォルニア州の高速道路に設置された 228 個のセンサーから、5 分間隔で収集された交通速度データを含む。我々は過去 1 時間を入力とし、次の 15 分～45 分間の交通速度を予測するタスクを設定した。

NSD はもともと分類タスクに対して設計されていたが、本研究では回帰タスクへ拡張し、時系列モデルのような特別な構造を導入せずに高精度な予測を実現した。図 2 に示すように、NSD はラッシュ時や非ラッシュ時の速度変化を正確に捉えている。また、予測誤差 (RMSE, MAE) についても既存手法に匹敵する結果が得られており、NSD の回帰タスクへの応用可能性を示している。

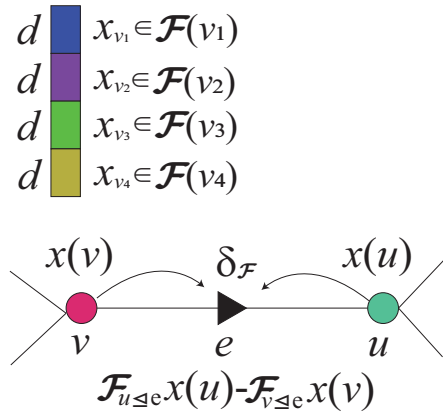


図 1. 層  $F$  の概念図 [2]

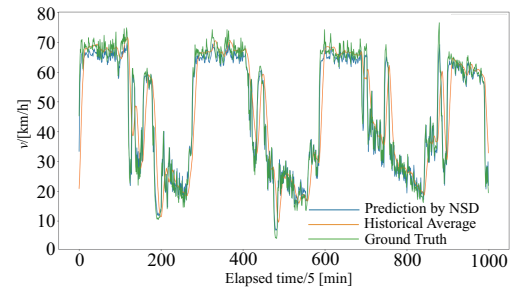


図 2. NSD 予測と実際の交通速度

**謝辞** 研究に有益なコメントをくださった東京大学数理科学研究科の田中雄一郎先生、中村勇哉先生、樋川達郎さん、東北大学災害科学国際研究所の奥村誠先生に深く感謝申し上げます。この研究は JSPS 科研費 24KJ0713 の助成を受けたものです。

### 参考文献

- [1] Kenta Oono and Taiji Suzuki. Graph neural networks exponentially lose expressive power for node classification. In *International Conference on Learning Representations*.
- [2] Cristian Bodnar, Francesco Di Giovanni, Benjamin Chamberlain, Pietro Lio, and Michael Bronstein. Neural sheaf diffusion: A topological perspective on heterophily and over-smoothing in gnns. *Advances in Neural Information Processing Systems*, Vol. 35, pp. 18527–18541, 2022.
- [3] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875*, 2017.
- [4] Weiwei Jiang and Jiayun Luo. Graph neural network for traffic forecasting: A survey. *Expert systems with applications*, Vol. 207, p. 117921, 2022.