

振動するニューラル常微分方程式におけるほとんど不変な集合を用いた学習 Learning with Almost Invariant Sets in Neural Oscillatory ODEs

成井 祐人¹⁾ 小西 卓哉¹⁾²⁾ 河原 吉伸¹⁾²⁾
Yuto Inui Takuya Konishi Yoshinobu Kawahara

1 概要

ニューラル常微分方程式は、ニューラルネットワークにより記述される常微分方程式を応用した深層学習モデルであり、その特徴を活かして幅広い問題に応用されている。ニューラル常微分方程式では、経験的にその力学系による軌跡は振動せず安定した平衡点に収束することが知られているが、神経科学などの分野では周期的で振動する力学系が重要な役割を果たす様々な事例が報告されている。本研究では周期的かつ振動する力学系を特徴づけるほとんど不変な集合 (AIS) を特徴として利用する新たな機構を提案する。画像分類の数値実験により、提案手法の有効性とその挙動を評価する。

2 はじめに

近年、深層学習と力学系とのつながりが機械学習やその関連分野において活発に研究されている [1]。そうした中で、ニューラル常微分方程式 [6] は連続時間力学系として表現される深層学習モデルとして注目されている。近年主流となっているフィードフォワード型のニューラルネットワークは多数の隠れ層を経由することで入力を出力へ変換するが、ニューラル常微分方程式は隠れ層における変換をニューラルネットワークによって記述された常微分方程式として与える。こうしたモデル化は連続的な変化を伴う問題の記述に適しており、時系列予測 [16, 31] や、密度関数の推定 [10]、物理法則の学習 [11] など幅広く応用されている。また、確率微分方程式への拡張 [15] や、硬い方程式への応用 [17] など、様々な発展的な手法が提案されている。

本研究では、非線形力学系としてのニューラル常微分方程式の表現能力に着目する。表現能力を改善する研究はこれまでも行われてきたが [7, 24]、これら既存手法では、ニューラル常微分方程式が従う力学系に振動がなく安定した平衡点へ収束することを前提としてしている。これに対して、再帰的ニューラルネットワークにおける研究では、振動などより広い意味での収束が性能の改善に寄与することが報告されている [4, 32]。また、神経科学においても、ニューロンや人の学習過程において周期的な活動や振動現象の重要性が指摘されている [28, 34]。既存のニューラル常微分方程式はこうした力学系の周期性を考慮することができず、力学系の限られた情報しか表現することができない。

こうした周期的かつ振動する力学系から情報を引き出すために、本研究では、ほとんど不変な集合 (AIS; Almost Invariant Set) を特徴として利用する新たな機構を提案する。ほとんど不変な集合は、力学系が作用する状態空間上の領域であり、力学系の作用による軌跡が異なる領域にほとんど入ることがない領域として特徴付け

られる。提案するモデルは、周期的かつ振動する力学系においても、単にその値を利用するのではなく、ほとんど不変な集合を介することでより表現力のある特徴表現を獲得する。画像分類による数値実験では、提案モデルが、他のニューラル常微分方程式 [7, 24] よりも同等もしくは優れた性能を達成することが確認された。

本論文の構成は、以下の通りである。3 章では、ニューラル常微分方程式について概説し、また提案手法の導出のために重要となるクープマン作用素と転送作用素を概説する。4 章では、提案モデルについての説明を行い、更にランダム・フーリエ特徴を用いた実装方法を示す。5 章では、関連研究について述べる。6 章では、提案手法を画像分類へ応用した数値実験を行い、提案手法の比較評価、そしてその詳細な挙動をスペクトル解析によって検証する。最後に、7 章では提案手法と実験結果を考察し、8 章で本研究の結論を述べる。

3 背景

3.1 ニューラル常微分方程式

入力 $x \in \mathbb{X} \subset \mathbb{R}^{n_x}$ から出力 $y \in \mathbb{Y}$ を予測する教師あり学習問題を考える (例えば、回帰の場合なら $\mathbb{Y} \subset \mathbb{R}^{n_y}$ 、分類の場合なら $\mathbb{Y} = \{1, \dots, n_y\}$ を各々想定する。ただし $n_x, n_y \in \mathbb{N}$)。ニューラル常微分方程式では、次の常微分方程式の初期値問題を考える。

$$\frac{dz(t)}{dt} = f_\theta(z(t)), \quad z(0) = x$$

ここで、 $z(t) \in \mathbb{Z} \subset \mathbb{R}^{n_z}$ ($n_z \in \mathbb{N}$) は時刻 $t \in \mathbb{R}$ における隠れ状態であり、 $f_\theta: \mathbb{Z} \rightarrow \mathbb{Z}$ はパラメータ θ をもつニューラルネットワークである。時刻 T における解は以下のように与えられる。

$$z(T) = z(0) + \int_0^T f_\theta(z(t)) dt$$

ニューラル常微分方程式では、まず上記の初期値問題を解くことで解 $z(T)$ を求め、得られた $z(T)$ を出力の予測 \hat{y} を計算するために使用する。ただし、 f_θ はニューラルネットワークであり非線形性をもつため、一般的に初期値問題を解析的に解くことは困難である。ニューラル常微分方程式は、常微分方程式の数値解法を用いることで近似的に $z(T)$ を計算する。

ニューラル常微分方程式の表現力を改善するために、入力と隠れ状態の空間を拡張する手法が提案されている。ニューラル常微分方程式では、隠れ状態の次元 n_z がモデルの性能を左右する。例えば、 n_x と n_z が同じであるならば、ニューラル常微分方程式は入力空間の位相が保存されるような制約された表現を学習することになる。この問題に対して、Dupont らは拡張ニューラル常微分方程式を提案している [7]。このモデルでは、入力 x に $\mathbf{0}$ を連結させる、つまり初期時刻の隠れ状態を $z(0) = [x, \mathbf{0}]$ とする。また Massaroli らは、より一般的

- 1) 大阪大学大学院情報科学研究科 Graduate School of Information Science and Technology, Osaka University
- 2) 理化学研究所革新知能統合研究センター Center for Advanced Intelligence Project, RIKEN

なアプローチとして、入力を変換する入力ネットワーク $h_x: \mathbb{X} \rightarrow \mathbb{Z}$ を提案している [24].

$$z(0) = h_x(x)$$

ここで、 h_x は典型的には一層の全結合層または畳み込み層である。このアプローチは、入力層における拡張と呼ばれる。さらに Massaroli らは、同研究において解 $z(T)$ を出力に変換するための出力ネットワーク $h_y: \mathbb{Z} \rightarrow \mathbb{Y}$ を合わせて提案している。

$$\hat{y} = h_y(z(T))$$

3.2 ほとんど不変な集合の同定

ほとんど不変な集合 (AIS; Almost Invariant Set) は測度保存力学系における遅い緩和を担う状態空間における領域であり、状態が一度その領域に入ると抜け出すことが難しくなる。AIS は、非線形力学系の解析に用いられるクープマン作用素の共役である転送作用素のスペクトルを調べることで同定することができる。

まずニューラル常微分方程式によって記述される力学系を調査する。ここで、時間方向に沿って一定間隔 τ だけ軌跡を積分することで得られる、離散時間の系であるフローを考える。フローは次式で定義される。

$$\mathcal{A}^\tau(z(t)) = z(t + \tau) \quad (1)$$

ここで $\mathcal{A}^\tau: \mathbb{Z} \rightarrow \mathbb{Z}$ であり、状態を一定の時間 τ だけ進める写像である。

次に、 $\mathbb{G}(\mathbb{Z})$ を \mathbb{Z} 上の関数空間とし、 $g \in \mathbb{G}(\mathbb{Z})$ を系の観測量とする。このとき、クープマン作用素 [19] (または合成作用素) $\mathcal{K}^\tau: \mathbb{G}(\mathbb{Z}) \rightarrow \mathbb{G}(\mathbb{Z})$ は、任意の $z \in \mathbb{Z}$ に対して次のように定義することができる。

$$\mathcal{K}^\tau g(z) = g(\mathcal{A}^\tau(z)) \quad (2)$$

つまり、クープマン作用素は、元の系を観測量 g によりリフトした表現を与え、関数空間において線形な時間遷移を表現している。また、転送作用素 (またはペロン-フロベニウス作用素) は状態密度の遷移に関する表現を与える。密度が汎関数の空間 $\Phi: \mathbb{Z} \rightarrow \mathbb{G}(\mathbb{Z})$ に埋め込まれているとすると、転送作用素は次のように定義される。

$$\mathcal{P}^\tau \phi(\cdot) = \phi(\mathcal{A}^\tau(\cdot)) \quad (3)$$

ここで $\phi \in \Phi$ である。このとき、転送作用素の固有値 1 に対応する固有関数の値は、AIS の内部で 0 でない値をとる [2]。また AIS は、対応する固有値が複素平面上での単位円の内部に広く散らばる可能性があるため点スペクトルとはならず、転送作用素の代わりにクープマン作用素のスペクトルを通して解析することは一般に難しい。

4 提案手法

ニューラル常微分方程式のための新たな機構であるニューラル AIS を提案し、ランダム・フーリエ特徴を用いた実装方法について説明する。

4.1 ニューラル AIS

既存研究ではニューラル常微分方程式の優れた性能が報告されているが、ニューラル常微分方程式が従う力学系は、経験的に振動せず安定で平衡点に収束していくこ

とが知られている。これに対して、神経科学から着想を得た数理モデルでは、脳活動で観察されるようなニューロンの高頻度の発火によって振動が発生する [12, 30]。しかしこのような振動するモデルは、相対的に良い性能を発揮できずこれまであまり用いられてこなかった。その主な理由の一つとして振動パターンを特徴として活用することの難しさが挙げられる。本研究では、ニューラル常微分方程式の力学系がもつ特徴を、AIS を介することで、より表現力のある特徴として取り出すことができるニューラル AIS を提案する。

ニューラル AIS はニューラル常微分方程式に次の二つの変更を加える。まず、神経科学における振動を伴うモデルを参考に、次のような常微分方程式を定義する。

$$f'_\theta(z(t)) = f_\theta(z(t)) - \gamma z(t) \quad (4)$$

ここで、 $\gamma \in \mathbb{R}_{>0}$ は正のスカラ値である。式 (4) における右辺の第二項 $-\gamma z(t)$ は力学系に振動を発生させる。この変更は元のモデル $f_\theta(z(t))$ に $-\gamma z(t)$ を加えるだけであり、従来のニューラル常微分方程式に容易に適用することができる。

さらに、上記の振動する力学系を活用するために、出力ネットワークを次のように定義する。

$$h_y(z(T)) = \mathbf{A}\phi'_k(z(T)) + \mathbf{b} \quad (5)$$

ここで、 $\phi'_k: \mathbb{Z} \rightarrow \mathbb{R}^M$ ($M \in \mathbb{N}$) は、式 (3) における ϕ の近似であり、 $\mathbf{A} \in \mathbb{R}^{n_y \times M}$ は重み、 $\mathbf{b} \in \mathbb{R}^{n_y}$ はバイアスをそれぞれ表す。次の 4.2 節では、どのようにして ϕ を有限次元で近似するか説明する。ここで、 ϕ'_k はすでに有限次元で近似されているとすると、転送作用素の 1 でない固有値に対応する ϕ'_k の成分は最終時刻 T にはほとんど消失しており、 $\phi'_k(z(T))$ は AIS を介した振動する力学系の特徴を表している。最後に、 $\phi'_k(z(T))$ を一層の全結合層で変形することで予測を出力する。

4.2 ランダム・フーリエ特徴を用いた実装

まず、内積 $\langle \cdot, \cdot \rangle$ と正定値カーネル k を備えた再生核ヒルベルト空間 (RKHS) を Φ_k とする。このとき、式 (3) における ϕ として特徴写像 $\phi_k: \mathbb{Z} \rightarrow \Phi_k$ を用いる。依然として ϕ_k は無限次元であるため、ここでランダム・フーリエ特徴 (RFF; Random Fourier Feature) [29] で ϕ_k を有限次元近似する。ランダム・フーリエ特徴は、正定値かつシフト不変であるカーネル k において、特徴写像の近似である写像 $\phi'_k: \mathbb{Z} \rightarrow \mathbb{R}^M$ を生成する。この近似的な特徴写像は $\phi'_k(x)^\top \phi'_k(y) \approx \langle \phi_k(x), \phi_k(y) \rangle = k(x, y)$ を満たし、カーネル k に依存する。本研究ではガウスカーネルを用いる。このとき、近似的な特徴写像は次のように表せる。

$$\begin{aligned} \phi'_k(x) &= \frac{1}{\sqrt{M/2}} (\sin(\omega_1^\top x), \cos(\omega_1^\top x), \dots, \cos(\omega_{M/2}^\top x))^\top \\ \omega_j &\sim \mathcal{N}(0, \sigma^2 \mathbf{I}) \end{aligned}$$

ここで、 $\omega_j \in \mathbb{X}$ は正規分布に従う確率ベクトルであり、一度サンプリングによって得られた後は固定される。 σ^2 は正規分布の分散である。 M および σ^2 はハイパーパラメータとして扱う。

5 関連研究

力学系と関連したニューラルネットワーク

ニューラル常微分方程式は、深層学習と力学系の2つの見方ができるため様々な側面から研究されている。Dupont らは隠れ状態の次元を拡張する拡張ニューラル常微分方程式を提案した [7]。拡張ニューラル常微分方程式をもとに拡張方法の一般化 [24] や、2階常微分方程式との関係を解析した研究 [26] が行われている。また、ニューラル常微分方程式は畳み込みニューラルネットワークと比べて敵対的攻撃に頑強であることが報告されている [36]。さらに安定性理論の観点から敵対的攻撃に対する防御の面を改善した手法も提案されている [14]。

再帰的ニューラルネットワークは離散時間力学系として解釈することができ、力学系としての拡張や発展手法が研究されている [4, 13, 32]。特に、Rusch と Mishra は 2 階常微分方程式として記述される振動する力学系に基づく再帰的ニューラルネットワークを提案している [32]。このモデルは我々と同じく振動する系を活用しているが、その常微分方程式としての記述には振動の減衰を課す項も含まれており、我々とは異なる力学系を扱っている。

また、力学系と関連するニューラルネットワークに深層平衡モデル (DEQ) が挙げられる [1]。DEQ は重みが共有された変換を繰り返し作用させたときの平衡点として隠れ層から出力を得る。近年、DEQ をクープマン作用素によって近似し、さらに転送作用素に基づきより広い力学系への収束を想定したモデルが提案されている [18]。

クープマン作用素

クープマン作用素は作用素論の観点から非線形力学系を解析するための重要なツールとなっており、理論的な解析だけでなく応用も多岐に渡る [3, 19]。また、AIS として知られる力学系における遅い減衰の検知は、流体力学 [8]、海洋力学 [9] などの分野で活用されている。AIS によって定義される遅く減衰する領域は力学系の大域的な情報を持っており、転送作用素はそれを獲得するために用いられている。

6 実験

二つの数値実験を通して提案手法を評価する。以降の実験では、モデルと訓練アルゴリズムを PyTorch [27] と torchdiffeq [5] を用いて実装した。またすべての実験を AMD EPYC 7542 と NVIDIA A100 を搭載する計算機上で行った。

6.1 画像分類

本節では画像分類にニューラル AIS を応用することで既存のニューラル常微分方程式と比較する。具体的には、拡張ニューラル常微分方程式 (ANODE) [7] と、入力層拡張ニューラル常微分方程式 (IL-NODE) [24] の二つを比較対象とした。4.1 節で述べたように、ニューラル AIS は、ANODE と IL-NODE の入力ネットワーク h_x と f_θ を変更することなく応用することができる (ただし、ニューラル AIS の f'_θ は $-\gamma z(t)$ の項を含む)。そこで、それぞれの研究で提案されている実装に対して別々にニューラル AIS を適用することで改善が見られるか評価した。

6.1.1 データセット

CIFAR-10 [21], SVHN [25], MNIST [20] の三つのデータセットを使用した。いずれのデータセットも訓練データとテストデータが含まれるが、元の訓練データの 90% を訓練データに、10% を検証データとして使用し、テストデータについては元のテストデータをそのまま使用した。また、前処理については公平な比較のために比較するモデルと同じ処理を行った。表 1 に三つのデータセットの統計量を示す。

6.1.2 モデルの訓練および評価

いずれの訓練と評価についても初期値問題を数値的に解くために Runge-Kutta 法の 1 つであり適応的なステップ幅を利用する Dormand-Prince 法 (dopri5) を使用した。また、勾配の計算には常微分方程式の数値解法にそのまま自動微分を適用することで計算する discretize-then-optimize (DTO) を用いた。モデルの訓練と評価を行う前に、訓練データと検証データを用いてハイパーパラメータの調節を行った。ハイパーパラメータを探索するために Tune [23] を使用した。具体的には探索アルゴリズムにランダム探索を設定し、探索プロセスのスケジューリングアルゴリズムに asynchronous successive halving algorithm (ASHA) [22] を使用した。次に探索によって見つかった最良のハイパーパラメータを設定したモデルを訓練データと検証データの両方で訓練し、最後にテストデータで訓練したモデルを評価した。

6.1.3 結果

表 2 および表 3 に ANODE と IL-NODE それぞれに対してニューラル AIS を応用したときのテストデータに対する正解率の平均値を示す。比較モデルの ANODE と IL-NODE についてはそれぞれの論文において報告された数値を掲載している。ただし、IL-NODE については SVHN での実験が報告されていなかったため、ニューラル AIS と同様の手順でハイパーパラメータの探索、訓練および評価を行い、結果を示している。表 2 と表 3 はニューラル AIS がすべてのデータセットで比較モデルよりも高い正解率を達成したことを示している。また、それぞれのモデルの学習可能なパラメータ数も合わせて示しており、ニューラル AIS がそれぞれ ANODE よりも小さく、IL-NODE よりも大きいことがわかる。これは、 h_x と f_θ は比較モデルと同じモデルを使用しているが、ニューラル AIS はランダムフーリエ特徴の次元 M を十分大きく設定する必要があり、 M と対応する全結合層の重み A を調整する必要があったためである。

6.2 ニューラル AIS のスペクトル解析

ニューラル AIS が従う力学系の挙動を詳細に検証するためにスペクトル解析を行った。IL-NODE にニューラル AIS を応用したモデルを CIFAR-10 データセットで訓練し、そのクープマン作用素と転送作用素の固有値を推定した。モデルが想定通りの力学系に従うならば、複素平面上でクープマン作用素の固有値の多くは単位円上に存在し、転送作用素の固有値のいくつかは 1 に近い値を取ることが期待される。クープマン作用素および転送作用素の固有値を推定するために動的モード分解 (DMD) を使用した [33, 35]。DMD を正確に実行するために、常微分方程式の数値解法として固定サイズの

表 1 画像分類における三つのデータセットの統計量.

	訓練データ数	検証データ数	テストデータ数	入力次元	出力次元
CIFAR-10	45,000	5,000	10,000	$3 \times 32 \times 32$	10
SVHN	65,931	7,326	26,032	$3 \times 32 \times 32$	10
MNIST	54,000	6,000	10,000	$1 \times 28 \times 28$	10

表 2 画像分類における 5 回の実行に対するテストデータに対する正解率の平均と, ANODE とニューラル AIS の学習可能なパラメータ数. ANODE の結果は文献 [7] で報告された値を示した.

	CIFAR-10		SVHN		MNIST	
	正解率	パラメータ数	正解率	パラメータ数	正解率	パラメータ数
ANODE	60.6	171.8K	83.5	171.8K	98.2	84.8K
ニューラル AIS (提案法)	65.7	120.6K	86.4	120.6K	99.1	78.7K

ステップ幅を使用する 4 次のルンゲ・クッタ法 (rk4) を用いた. なおステップ幅 τ は 0.025 とした. 固定のステップ幅の数値解法を使用するため画像分類と異なる設定だが, rk4 を使用した場合も, CIFAR-10 データセットに対して dopri5 を使用した場合とほぼ同等のテストデータの正解率を達成することを確認した.

本実験では, exact DMD [35] に従ってクープマン作用素の固有値を次のように推定した. まず, 訓練データと検証データから選択した 20 個の画像についてニューラル AIS の順方向の計算を行った. それぞれ画像の隠れ状態を計算することで, 20 個の系列からなる集合 $\{(z_i(0), z_i(\tau), z_i(2\tau), \dots, z_i(40\tau))\}_{i=1}^{20}$ を得た. ここで, $z_i(t)$ は i 番目の画像の時刻 $t \in \{0, \tau, 2\tau, \dots, 40\tau\}$ における隠れ状態を表し, $z_i(t)$ は全てベクトルに変形されているとする. 次に, 二つのデータ行列 $\mathbf{X}_K \in \mathbb{R}^{43008 \times 800}$ と $\mathbf{Y}_K \in \mathbb{R}^{43008 \times 800}$ を次のように作成した.

$$\begin{aligned} \mathbf{X}_K &= [z_1(0) \cdots z_1(39\tau) \cdots z_{20}(0) \cdots z_{20}(39\tau)] \\ \mathbf{Y}_K &= [z_1(1\tau) \cdots z_1(40\tau) \cdots z_{20}(1\tau) \cdots z_{20}(40\tau)] \end{aligned}$$

これらデータ行列について, exact DMD では以下の最小化問題を解くことで行列 \mathbf{K} を求める.

$$\mathbf{K} = \underset{\mathbf{K}}{\operatorname{argmin}} \|\mathbf{Y}_K - \mathbf{K}\mathbf{X}_K\|_F$$

この \mathbf{K} に固有値分解を行い, その固有値を求めることで, クープマン作用素の固有値の推定することができる. 実用上は, 直接 \mathbf{K} の固有値分解を計算する代わりに, 行列 $\mathbf{K}' = \mathbf{U}^H \mathbf{K} \mathbf{U}$ の固有値分解を計算することで効率的に計算する方法が用いられる. ここで, \mathbf{U} は \mathbf{X}_K を特異値分解した際の左特異ベクトルからなる行列であり, \mathbf{U}^H はそのエルミート転置である. \mathbf{U} は直交行列であり, \mathbf{K}' の固有値は \mathbf{K} の先頭の固有値と一致する. 本実験でも \mathbf{K}' の固有値分解を行うことで, クープマン作用素の固有値の推定を行った. 転送作用素については, 先ほどの隠れ状態をランダムフーリエ特徴で変換した $\phi'_k(z_i(t))$ で構成される系列の集合 $\{(\phi'_k(z_i(0)), \dots, \phi'_k(z_i(40\tau)))\}_{i=1}^{20}$ を用意した. この系列集合に対して上記のクープマン作用素と同じ推定方法によって固有値を推定することで, 転送作用素の固有値を推定した.

図 1 (左) に複素平面上で推定されたクープマン作用素の固有値を示す. 多くの推定された固有値が単位円上

に分布していることを確認できる. また, 図 1 (右) に推定された転送作用素の固有値を示す. すべての推定された固有値が単位円の内部に存在し, いくつかの固有値は座標 (1, 0) 付近に位置していることが確認できる. これらの結果によりニューラル AIS が従う力学系が周期的であり, AIS を通してその特徴量を抽出できていると考えられる.

7 考察

6.1 節の画像分類の実験では, ニューラル AIS が既存のニューラル常微分方程式よりも高い性能を持つことが確認された. 本研究で着目したより広い収束の有効性が示されたと言える. 興味深い点として, ニューラル AIS の $f'_\theta(z(t))$ のパラメータ数は, 比較モデルの $f_\theta(z(t))$ のパラメータ数と同じであり, ニューラル AIS が力学系からより多くの特徴を効率的に取り出すことができたことが示唆された. なお, 表 3 で示した IL-NODE との比較では, ニューラル AIS の方がパラメータが大きかったが, これは IL-NODE の出力ネットワークは, 畳み込み層や平均プーリング層, 全結合層と画像処理に特化した効率的な隠れ層から構成されているのに対し, ニューラル AIS の出力ネットワークは画像に特化していないことも大きくなった要因であり, モデル化方法についてはさらなる調査の余地がある.

また, 6.2 節ではニューラル AIS のスペクトル解析により, 推定されたクープマン作用素の多くの固有値は複素平面上の単位円上に分布していることを確認した. この結果はニューラル AIS の力学系が周期的であり振動していることを示唆している. また, 推定された転送作用素のいくつかの固有値は 1 に近い値をとるのみで, それ以外は絶対値が 1 以下となっている傾向が見られた. これはつまり, 固有値 1(AIS) 以外の振動は時間が経過すると収束して消失することを意味しており, ニューラル AIS の提案の動機が正しかったことを示している.

8 おわりに

本研究では, 周期的かつ振動する力学系を構成し, その特徴を利用するニューラル常微分方程式のための新たな機構を提案した. ニューラル AIS は, 振動する力学系から特徴を取り出す汎用的な手法であり, スパイキングニューラルネットワークなどにも応用可能であると考えられる.

今後の発展として, まずニューラル AIS をカオスの縁と関連づけた理論解析や敵対的学習に対するロバスト性

表 3 画像分類における 10 回の実行に対するテストデータの正解率の平均と、IL-NODE とニューラル AIS の学習可能なパラメータの数。IL-NODE の MNIST と CIFAR-10 の結果は文献 [24] で報告された値を示した。IL-NODE の SVHN の結果は本研究の実験で得られた値を示した。

	CIFAR-10		SVHN		MNIST	
	正解率	パラメータ数	正解率	パラメータ数	正解率	パラメータ数
IL-NODE	73.4	36.1K	88.3	36.1K	99.1	20.7K
ニューラル AIS (提案法)	76.2	116.8K	94.6	116.8K	99.5	60.9K

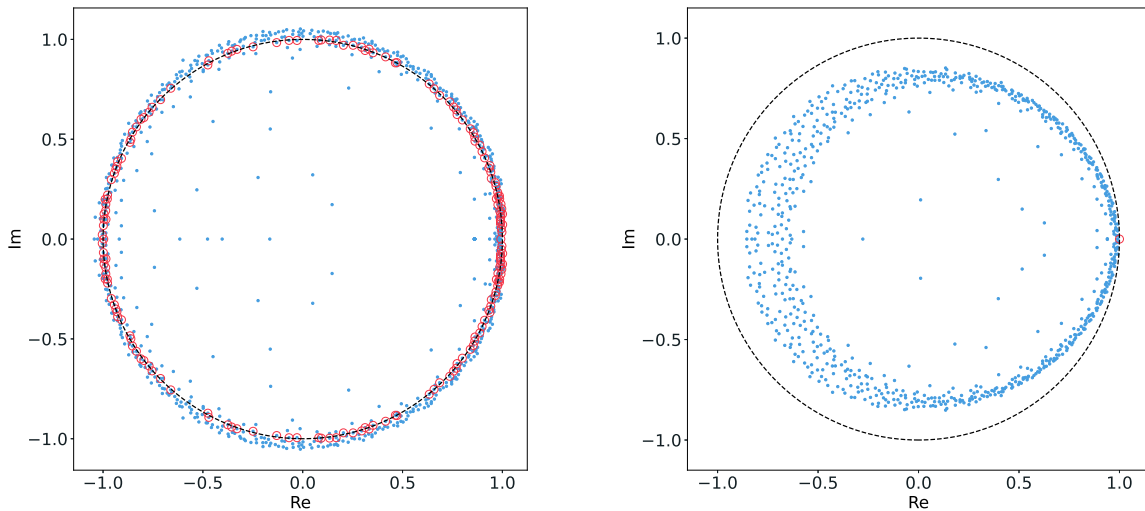


図 1 複素平面上の推定されたクープマン作用素の固有値 (左) と推定された転送作用素の固有値 (右)。点線は複素平面上の単位円を示し、青点は固有値を表す。赤円は絶対値が 1 に近い固有値を囲っている。

の評価を行うことが挙げられる。また、転送作用素が作用する空間としてガウスクERNELによって定まる RKHS を用いたが、より広い空間を仮定したときの性能を検証する必要がある。

常微分方程式のパラメータが時刻に依存しないニューラル常微分方程式とは異なり、従来の離散的に変換を行うモデルの性能に追従するため、ニューラル常微分方程式において常微分方程式を記述するニューラルネットワークのパラメータが時間発展するモデルが提案されている [37]。このモデルの力学系は非定常性を持ち、ニューラル AIS と同様の方法を適用することや、異なる力学系の知識を活かしたアプローチを検討することができる。また、ニューラル常微分方程式の力学系を振動させる方法は本研究の手法以外にも考えられる。今後は、振動させる方法によって力学系から得られる特徴にどのような変化があるのか調査することが考えられる。

謝辞

本研究の成果の一部は、JSPS 科研費 JP22H00516, JP22H05106, 22K17950, 及び JST CREST JPMJCR1913 の助成を受けて実施された研究により得られたものである。また、大阪大学の田中亮悟氏には本研究に関する話題について議論いただいた。理化学研究所の西村能輝氏には計算機に関する技術的な支援をいただいた。ここに深く感謝申し上げる。

参考文献

[1] Bai, S., Kolter, J.Z., Koltun, V.: Deep equilibrium models. In: Advances in Neural Information Processing Systems 32. pp. 688–699 (2019)

[2] Billings, L., Schwartz, I.B.: Identifying almost invariant sets in stochastic dynamical systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **18**(2), 023122 (2008)

[3] Brunton, S.L., Budišić, M., Kaiser, E., Kutz, J.N.: Modern Koopman theory for dynamical systems. *SIAM Rev.* **64**, 229–340 (2021)

[4] Chang, B., Chen, M., Haber, E., Chi, E.H.: AntisymmetricRNN: A dynamical system view on recurrent neural networks. In: 7th International Conference on Learning Representations (2019)

[5] Chen, R.T.Q.: torchdiffeq (2018), <https://github.com/rtqichen/torchdiffeq>

[6] Chen, T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. In: Advances in Neural Information Processing Systems 31. pp. 6572–6583 (2018)

[7] Dupont, E., Doucet, A., Teh, Y.W.: Augmented neural ODEs. In: Advances in Neural Information Processing Systems 32. pp. 3134–3144 (2019)

[8] Froyland, G., Padberg, K.: Almost-invariant sets and invariant manifolds – Connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D: Nonlinear Phenomena* **238**(16), 1507–1523 (2009)

[9] Froyland, G., Padberg, K., England, M.H., Treguier, A.M.: Detection of coherent oceanic structures via transfer operators. *Phys. Rev. Lett.* **98**, 224503 (2007)

[10] Grathwohl, W., Chen, R.T.Q., Bettencourt, J., Sutskever, I., Duvenaud, D.: FFIORD: Free-form continuous dynamics for scalable reversible generative models. In: 7th International Conference on Learning Representations (2019)

- [11] Greydanus, S., Dzamba, M., Yosinski, J.: Hamiltonian neural networks. In: *Advances in Neural Information Processing Systems* 32. pp. 15353–15363 (2019)
- [12] Izhikevich, E.: *Dynamical systems in neuroscience*. MIT Press p. 111 (2007)
- [13] Kag, A., Zhang, Z., Saligrama, V.: RNNs incrementally evolving on an equilibrium manifold: A panacea for vanishing and exploding gradients? In: *8th International Conference on Learning Representations* (2020)
- [14] Kang, Q., Song, Y., Ding, Q., Tay, W.P.: Stable neural ODE with Lyapunov-stable equilibrium points for defending against adversarial attacks. In: *Advances in Neural Information Processing Systems* 34. pp. 14925–14937 (2021)
- [15] Kidger, P., Foster, J., Li, X., Lyons, T.J.: Neural SDEs as infinite-dimensional GANs. In: *Proceedings of the 38th International Conference on Machine Learning*. pp. 5453–5463 (2021)
- [16] Kidger, P., Morrill, J., Foster, J., Lyons, T.J.: Neural controlled differential equations for irregular time series. In: *Advances in Neural Information Processing Systems* 33. pp. 6696–6707 (2020)
- [17] Kim, S., Ji, W., Deng, S., Ma, Y., Rackauckas, C.: Stiff neural ordinary differential equations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **31**, Medium: ED; Size: Article No. 093122 (2021)
- [18] Konishi, T., Kawahara, Y.: Stable invariant models via Koopman spectra. *Neural Networks* **165**, 393–405 (2023)
- [19] Koopman, B.O.: Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences of the United States of America* **17**(5), 315–318 (1931)
- [20] Krizhevsky, A.: Learning multiple layers of features from tiny images (2009)
- [21] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**(11), 2278–2324 (1998)
- [22] Li, L., Jamieson, K.G., Rostamizadeh, A., Gonina, E., Bentzur, J., Hardt, M., Recht, B., Talwalkar, A.: A system for massively parallel hyperparameter tuning. In: *Proceedings of Machine Learning and Systems*. vol. 2, pp. 230–246 (2020)
- [23] Liaw, R., Liang, E., Nishihara, R., Moritz, P., Gonzalez, J.E., Stoica, I.: Tune: A research platform for distributed model selection and training (2018), arXiv preprint arXiv:1807.05118
- [24] Massaroli, S., Poli, M., Park, J., Yamashita, A., Asama, H.: Dissecting neural ODEs. In: *Advances in Neural Information Processing Systems* 33. pp. 3952–3963 (2020)
- [25] Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., Ng, A.Y.: Reading digits in natural images with unsupervised feature learning. In: *NIPS workshop on deep learning and unsupervised feature learning* (2011)
- [26] Norcliffe, A., Bodnar, C., Day, B., Simidjievski, N., Lió, P.: On second order behaviour in augmented neural ODEs. In: *Advances in Neural Information Processing Systems* 33 (2020)
- [27] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: PyTorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems* 32. pp. 8024–8035 (2019)
- [28] Rabinovich, M.I., Varona, P., Selverston, A.I., Abarbanel, H.D.I.: Dynamical principles in neuroscience. *Rev. Mod. Phys.* **78**, 1213–1265 (2006)
- [29] Rahimi, A., Recht, B.: Random features for large-scale kernel machines. In: *Advances in Neural Information Processing Systems* 20. pp. 1177–1184 (2007)
- [30] Rosenblatt, F.: *Principles of neurodynamics; perceptrons and the theory of brain mechanisms*. Spartan Books (1962)
- [31] Rubanova, Y., Chen, R.T.Q., Duvenaud, D.K.: Latent ordinary differential equations for irregularly-sampled time series. In: *Advances in Neural Information Processing Systems* 32. pp. 5321–5331 (2019)
- [32] Rusch, T.K., Mishra, S.: Coupled oscillatory recurrent neural network (coRNN): An accurate and (gradient) stable architecture for learning long time dependencies. In: *9th International Conference on Learning Representations* (2021)
- [33] SCHMID, P.J.: Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **656**, 5–28 (2010)
- [34] Townley, S., Ilchmann, A., Weiß, M.G., McClements, W., Ruiz, A.C., Owens, D.H., Pratzel-Wolters, D.: Existence and learning of oscillations in recurrent neural networks. *IEEE Transactions on Neural Networks* **11**(1), 205–214 (2000)
- [35] Tu, J.H., Rowley, C.W., Luchtenburg, D.M., Brunton, S.L., Kutz, J.N.: On dynamic mode decomposition: Theory and applications. *ACM Journal of Computer Documentation* **1**, 391–421 (2013)
- [36] Yan, H., Du, J., Tan, V.Y.F., Feng, J.: On robustness of neural ordinary differential equations. In: *8th International Conference on Learning Representations* (2020)
- [37] Zhang, T., Yao, Z., Gholami, A., Gonzalez, J.E., Keutzer, K., Mahoney, M.W., Biros, G.: ANODEV2: A coupled neural ODE framework. In: *Advances in Neural Information Processing Systems* 32. pp. 5152–5162 (2019)