

多目的最適化手法を用いた Reservoir Computing の性能向上

伊藤美賀 (指導教員: Natanael Aubert-Kato)

1 はじめに

Reservoir Computing (RC) とは再帰的ニューラルネットワーク (RNN) の特殊なモデルを一般化した概念で、時系列情報処理に適した機械学習の枠組みの一つである。近年では、時間的パターンの分類、予測、生成など、多くの問題に適用されている。RC の計算性能を向上させるには、RNN ベースの Reservoir を最適に設計する必要がある。本研究では RC の特性に焦点を置き、多目的最適化アルゴリズムを用いることで RC の性能向上を目指す。

2 Reservoir Computing

2.1 Reservoir Computing の概要

従来型 RNN は、ノードを完全に相互接続し、すべての接続の重みを更新することで学習が行われるため、ネットワークのサイズが大きくなるにつれて、学習時間や結合重みを記憶するためのメモリが爆発的に増大してしまうという問題があった。しかし RC は、入力信号を reservoir で変換した後、読み出し部のみを簡便なアルゴリズムで訓練するので、極めて学習が高速である。また、適応更新を伴わないので、様々な物理システム、基板、およびデバイスを使用したハードウェア実装が可能である。実際このような物理的な RC は、様々な研究分野で注目を集めている。

2.2 Echo State Network (ESN)

ESN [1] とは、RC の代表的なモデルの一つで、図 1 のような構造をとる。主な考え方は、まず入力信号を用いてランダムでスパースな再帰的ネットワークを駆動する。それによってネットワーク内の各ニューロンに非線形応答信号を誘導し、過去の入力情報を反響 (Echo) させることで、時系列入力のパターンを学習する。適切な学習を行うには、各時刻での入力の影響がリザーバー内で時間とともに徐々に消失していくことが必要である。このような性質を、Echo state property と呼ぶ。

典型的な ESN の更新式は以下に与えられる。

$$\tilde{\mathbf{x}}(n) = \tanh(\mathbf{W}^{\text{in}}[1; \mathbf{u}(n)] + \mathbf{W}\mathbf{x}(n-1)) \quad (1)$$

$$\mathbf{x}(n) = (1 - \alpha)\mathbf{x}(n-1) + \alpha\tilde{\mathbf{x}}(n) \quad (2)$$

ここで $\mathbf{x}(n)$ は reservoir 内のニューロン活性化ベクトルであり、 $\tilde{\mathbf{x}}(n)$ で更新を行う。 \mathbf{W}^{in} と \mathbf{W} はそれぞれ入力重み行列と結合重み行列を表し、 α は漏れ率である。

出力層は以下のように定義される。

$$\mathbf{y}(n) = \mathbf{W}^{\text{out}}[1; \mathbf{u}(n); \mathbf{x}(n)] \quad (3)$$

ここで、 $\mathbf{y}(n)$ はネットワーク出力であり、 \mathbf{W}^{out} は出力重み行列を表している。

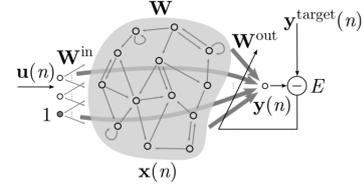


図 1: Echo State Network

3 Reservoir Computing の性能指標

最適化を行う上で RC が達成すべき性能を定義する必要がある。ここでは一般的な性能指標として、記憶容量、カーネル特性、汎化能力、reservoir サイズを目的関数として採用した [2]。

記憶容量 (Memory Capacity) 記憶容量とは入力情報をどれくらいの期間記憶しておけるかということである。RC は適切な学習を行うために、各時刻での入力の影響が reservoir 内で時間とともに徐々に消失していく性質を持つ。これは以下のように、入力信号 $\mathbf{u}(n)$ に対して、 $\mathbf{u}(n-k)$ を記憶するタスクで学習された出力信号 $\mathbf{y}_k(n)$ と、 k によって遅延された入力信号 $\mathbf{u}(n)$ の間の二乗相関係数で計算される。

$$MC = \sum_{k=1}^{\infty} \frac{\text{cov}^2(\mathbf{u}(n-k), \mathbf{y}_k(n))}{\sigma^2(\mathbf{u}(n))\sigma^2(\mathbf{y}_k(n))} \quad (4)$$

カーネル特性 (Kernel Rank) 次にカーネル特性とは、reservoir が異なる入力パターンを区別する能力の尺度である。これによって、入力 $\mathbf{u}(n)$ や過去の情報 $\mathbf{u}(n-1), \mathbf{u}(n-2) \dots$ といった非線形表現を適切に生成する能力を測定することができる。多くの実用的なタスクは線形分離不可能であるため、reservoir は通常入力情報の非線形変換を必要とする。すなわち、reservoir がどの程度複雑で多様な非線形変換操作が可能なのかを測りたいのである。カーネル特性を測るには、分散の大きいデータを入力信号として用い、reservoir 内の状態 $\mathbf{x}(n)$ を特異値分解した値を計算する。この値が大きいほうが高いカーネル特性を持つ。

汎化能力 (Generalization Rank) 汎化能力とは、類似した入力情報が与えられた場合の一般化能力のことである。汎化能力が劣っていると、訓練データに対しては学習がされているが、未知データ (テストデータ) に対しては適合できていないという状況に陥ってしまう。reservoir は学習に使用した入力だけでなく、これまでに見たことがないような新たな入力に対しても良い性能を発揮する必要がある。汎化能力を測るには、ノイズの大きいデータを入力信号として用い、reservoir 内の状態 $\mathbf{x}(n)$ を特異値分解した値を計算する。この値が小さいほうが高い汎化能力を持つ。

reservoir サイズ reservoir サイズとは, reservoir 内のニューロン数である.

優れた RC は大きな記憶容量, 高いカーネル特性, 高い汎用能力を持っていると言える. reservoir サイズが小さいものほど物理的な実験と計算機上のシミュレーション結果での差異が小さくなるため, 上記の3つの性能指標を満たす最小の reservoir を得たい.

4 多目的最適化手法の実装

4.1 実験設定

本研究では, より高性能な reservoir を実現するための設計指針として3つの手法を用いた. 第一の方法 (approach 1) は, 入力重み行列 \mathbf{W}^{in} と結合重み行列 \mathbf{W} はいくつかのパラメータに従ってランダムに生成されるので, それらを事前に最適化することで, 予め良い reservoir を生成するというものである. そこで, 最適化するパラメータとして reservoir サイズ, 漏れ率 α , ネットワークの結合重み行列のスペクトル半径を使用した. 第二の方法 (approach 2) は, 直接的なアプローチとして, 入力重み行列と結合重み行列の最適化を行った. この際, 漏れ率とスペクトル半径は経験的に良いとされる数値を使用した. 最後に, 第三の方法 (approach 3) として reservoir サイズ, 漏れ率 α , ネットワークの結合重み行列のスペクトル半径の最適化と, 入力重み行列, 結合重み行列の最適化を同時に行った. また本研究では, reservoir として ESN を使用し, 多目的最適化手法として Nondominated Sorting Genetic Algorithm II (NSGA-II) [3] を用いた. NSGA-II は遺伝的アルゴリズムの一種であり, 支配されている数に応じてソートを行い次世代の個体を選んでいくというアルゴリズムである.

4.2 実験結果

実験を通して得られた Pareto Front をそれぞれ2種類ずつの性能指標に注目して図2に表した. Pareto Front とは最適解集合により形成される曲線のことである. approach 2では記憶容量, カーネル特性を最大化するのに適している一方で, 汎化能力が劣っている. そして, サイズとカーネル特性が強い相関を持っていることが確認できた. approach 1では, 他の手法に比べ, 高い汎化能力を得ることに成功しており, 記憶容量に対してカーネル特性を最大化するのに適していることがわかる.

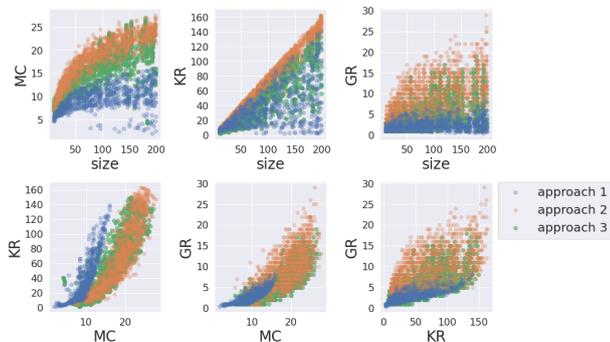


図2: Pareto Front の分布

また図3には hypervolume の推移を表した. hyper-

volume が大きいほど多様性の大きい解集合が獲得されたことを示す. 今回の実験の条件下では approach 2 が最も広範囲の解集合が得られたことがわかる. また図4には評価における各目的関数の中央値の推移を表した. approach 2において, 他手法よりも小さいサイズの reservoir で高い記憶容量, カーネル特性を達成しているが, これと同時に汎化能力が低下している. 一方で, approach 1は記憶容量, カーネル特性が共に低い値をとるものの, 高い汎化能力を得ている.

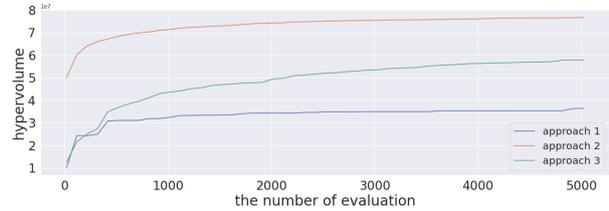


図3: hypervolume の推移

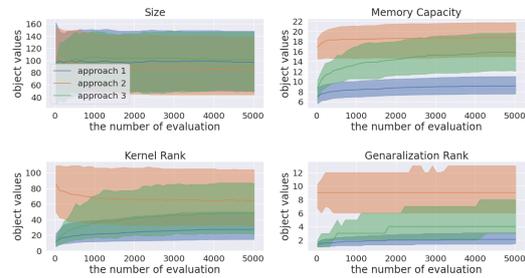


図4: 評価における目的関数の値の推移

5 まとめと今後の課題

4つの性能指標に焦点を置き, 多目的最適化アルゴリズムを用いて RC の性能向上に向けて実験を行った. 今回行った3つの手法によって, 様々な特性を持った最適解に近い reservoir を生成することができた. reservoir を使用する目的やタスクに合わせて適切な reservoir を選択することができるだろう. 今後は Quality-Diversity などの別のアルゴリズムを用いて, 結果を比較したい. また ESN 以外の reservoir を用いた RC の実装も実現したいと考えている.

参考文献

- [1] M Lukoševičius. A practical guide to applying echo state networks. In *Neural networks: Tricks of the trade*, pp. 659–686. Springer, 2012.
- [2] M Dale, JF Miller, S Stepney, and MA Trefzer. A substrate-independent framework to characterize reservoir computers. arXiv 2018. *arXiv preprint arXiv:1810.07135*.
- [3] K Deb, A Pratap, S Agarwal, and TMT Meiyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, Vol. 6, No. 2, pp. 182–197, 2002.