

GAN を用いて生成した擬似カオス時系列に対する非線形時系列解析

田中 悠貴¹, 山口 裕²¹ 福岡工業大学大学院工学研究科, ² 福岡工業大学情報工学部

e-mail : mfm21106@bene.fit.ac.jp

1 概要

時間的反复を伴わない畳み込みネットワークを用いてカオス的な時系列を生成できるか, 生成した時系列はどの程度カオス的といえるかを明らかにする目的で研究を行った. 具体的にはロジスティック写像のカオス時系列を訓練データとして敵対的生成ネットワーク (GAN) を学習させる事で, GAN がロジスティック写像のカオス時系列に類似した擬似時系列を生成した. さらに, 生成した擬似時系列に決定論的性質があるか, また疑似時系列は軌道不安定性を示すか, という2つの問題を, 非線形時系列解析手法を用いて調べた.

2 モデル

2.1 ネットワーク

GAN とは生成器 (G) と判別器 (D) の二つのネットワークを持ち, それらが敵対的に学習することで, 訓練データに類似したデータを生成することができる生成モデルである [1]. GAN の目的関数は式 (1) により与えられる.

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (1)$$

ここで $x \sim p_{\text{data}}(x)$ は訓練データ集合からサンプルされたデータ, $z \sim p_z(z)$ は潜在空間からサンプルされた z , $D(x)$ は判別器に訓練データが入力されたときのスコア, $G(z)$ は生成器が生成したデータ, $D(G(z))$ は判別器に生成データが入力されたときのスコアである.

実験に使用した GAN の G と D は, それぞれ一次元の畳み込みネットワーク (1D-CNN) を利用して構成した. そしてカオスを再現するために局所的な特徴を抽出できるように同じフィルターサイズの畳み込み層を重ねた. また各層の活性化関数には ELU (Exponential Linear Unit) を使用し, 畳み込み演算後の正規化は行っていない.

2.2 学習

実験では乱数で決定した 100 個の各初期値より, 長さ 1098 の軌道を $a = 4$ で固定したロジスティック写像 ($x_{n+1} = ax_n(1 - x_n)$) を用いて計算し, 訓練データとして使用した. モデルの訓練は, 上記の 100×1098 のデータでの訓練を 1 エポックとして, 20 万エポック訓練させた.

3 結果

まず今回の非線形時系列解析を用いる生成データのリターンマップとヒストグラムを図 3 に示す.

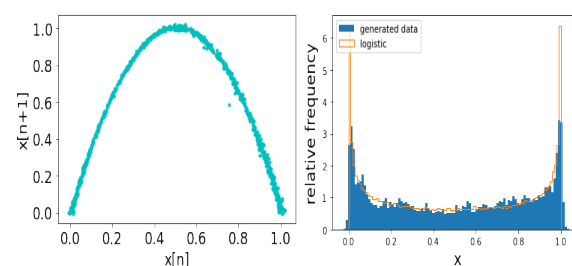


図 1. (左) 生成データのリターンマップ. (右) ヒストグラム. オレンジは訓練データ, 青は生成データ.

次に生成データと数値計算した時系列データから $m (= 1, \dots, 5)$ 次元の埋め込みベクトルを構成し, Wayland の手法 [2] により決定論的性質を評価する. 統計的評価のためにサロゲート法を用いた.

Wayland のアルゴリズムでは, まずある時刻 t_0 における値 $x(t_0)$ について K 個 (本実験では 50 個) の最近接ベクトルを見つける. そしてベクトル間の距離はユークリッド距離で測る. これらのベクトルを $x(t_i) (i = 0, 1, \dots, K)$ とし, 時間の経過に伴う各軌道の変化を

$$v(t_i) = x(t_i + T\Delta t) - x(t_i) \quad (2)$$

とする. 本実験では $T\Delta t = 1$ とした.

この $v(t_i)$ を用いて, translation error を

$$E_{\text{trans}} = \frac{1}{K+1} \sum_{i=0}^K \frac{\|v(t_i) - \hat{v}\|}{\|\hat{v}\|} \quad (3)$$

とする.ここで

$$\hat{v} = \frac{1}{K+1} \sum_{i=0}^K v(t_i) \quad (4)$$

とする.この操作を複数の点で行い E_{trans} の中央値をとる操作を複数回繰り返す, その中央値 \bar{E}_{trans} により決定論性を評価する. 低次元カオスのような決定論的時系列では, 決定論的側面が見えるにつれて $\bar{E}_{trans} \rightarrow 0$ となる [2].

サロゲートデータには, 数値計算した時系列データの埋め込みベクトルに対して, 値の順番をランダムに入れ替えた系列 (ランダムシャッフル) と周波数を変えずに位相をランダムに入れ替えた系列 (位相シャッフル) の二つを用いた. それぞれの 100 個ずつのサロゲートデータと生成データの埋め込みベクトルに Wayland の検定アルゴリズムを適用して比較した結果を図 2 に示す.

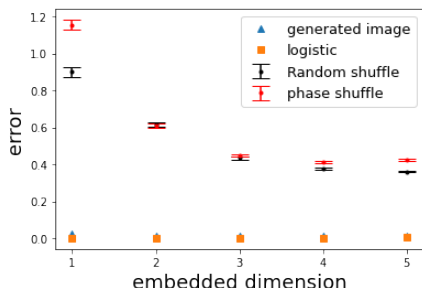


図 2. Wayland のアルゴリズムにより計算した translation error

図 2 から, GAN の生成した擬似時系列はランダムなサロゲートデータよりも決定論性がはるかに高く, 元の時系列と同程度の決定論的性質を示すことがわかった.

次に Kantz の最大リアプノフ指数の推定を行ったアルゴリズム [3] を用いた結果を示す. 埋め込み空間の時系列の点 $x(t_0)$ を選び, その点から半径 ϵ 内にある隣接点を選択する. 相対時間の関数として参照点からの軌道と隣接点からの軌道との間の距離の平均を計算する. これを多くの t_0 の値に対して行う. つまり式 (5) を計算する:

$$S(k) = \frac{1}{N} \sum_{t_0=1}^N \ln \left(\frac{1}{|u(x(t_0))|} \sum_{x(t) \in u(x(t_0))} \|x(t_0+k) - x(t+k)\| \right) \quad (5)$$

k のある範囲で式 (5) がロバストな線形増加を示す場合, その傾きは時間ステップ毎の最大リアプノフ指数の推定値である [3].

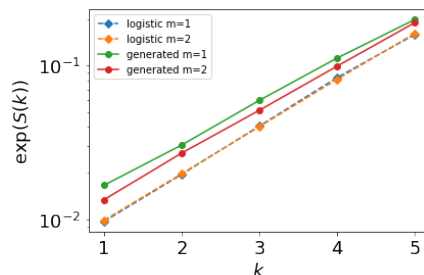


図 3. 最大リアプノフ指数の推定

生成した擬似時系列の $S(k)$ は線形増加を示しており, カオスの不安定性を学習出来ていると考えられる. 推定された最大リアプノフ指数は, 0.62629 ($m=1$) と 0.66007 ($m=2$) であった. なお真の値は $\log(2)$ (0.69315) である.

4 まとめ

本研究では GAN を用いてカオス時系列データの生成を行い, 非線形時系列解析手法を用いて分析を行った. その結果, 数値計算のカオス時系列と類似した性質を示すことが明らかとなった.

謝辞 本研究は JSPS 科研費 20K11985 の助成を受けた.

参考文献

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio. Generative adversarial nets, *Advances in neural information processing systems*, 27, 2672–2680, (2014).
- [2] R. Wayland, D. Bromley, D. Pickett, A. Passamante, *Recognizing determinism in a time series*, *physical review letters*, 70, 580–582, (1993).
- [3] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, (2000).