Deep learning algorithms for estimating Lyapunov exponents from observed time series in discrete dynamic systems

Andrey V. Makarenko
avm.science@mail.ru

Constructive Cybernetics Research Group
Moscow, Russia, www.rdcn.org

Institute of Control Sciences RAS
Moscow, Russia

XIV International Conference on Stability and Oscillations of Nonlinear Control Systems

01 June 2018

ICS RAS, Moscow, Russia

Outline

Outline section

Dynamical systems

Assume a dynamic system described by a discrete map:

$$\begin{aligned} &\mathbf{s}_{k+1} = \mathbf{f}(\mathbf{s}_k,\, \mathbf{p}), \\ &\boldsymbol{\phi}_{\mathbf{p}} : \mathrm{S} \times \mathrm{K} \to \mathrm{S}, \quad \boldsymbol{\phi}_{\mathbf{p}}(\mathbf{s},\, k) \equiv \mathbf{f}^{k}\left(\mathbf{s},\, \mathbf{p}\right), \end{aligned} \tag{1}$$

with properties:

$$\begin{aligned} &\mathbf{s} \in \mathrm{S} \subset \mathbb{R}^{N}, \quad \mathbf{p} \in \mathrm{P} \subset \mathbb{R}^{L}, \quad \mathbf{f} \in \mathrm{C}^{0}(\mathrm{S} \times \mathrm{P}), \\ &k \in \mathrm{K} \subseteq \mathbb{Z}, \quad n \in \overline{1,\, N}, \quad l \in \overline{1,\, L}, \end{aligned}$$

where $\mathbf{s}$ – map state variable, $\mathbf{p}$ – vector of map parameters, k – discrete time.

With the system (1) we also associate the trajectory $\{\mathbf{s}_k\}_{k \in \mathrm{K}}$ of the evolution of its state $\mathbf{s}$.

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

It estimated the local exponential divergence of two close trajectories and can be calculated from the eigenvalues of matrix

$$\mathbf{M}_{ij}^{k} = \big[\mathbf{G}(\mathbf{s}_{k-1})\,\mathbf{G}(\mathbf{s}_{k-2})\ldots\mathbf{G}(\mathbf{s}_0)\big]_{ij},$$

$$\mathbf{G}_{ij}(\mathbf{s}_k) = \frac{\partial\,\mathbf{f}_i}{\partial\,\mathbf{s}^{(j)}}\bigg|_{\mathbf{s}=\mathbf{s}_k},$$

where $\mathbf{G}$ – Jacobian matrix of system (1).

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

It estimated the local exponential divergence of two close trajectories and can be calculated from the eigenvalues of matrix

$$\mathbf{M}_{ij}^{k} = \big[\mathbf{G}(\mathbf{s}_{k-1})\,\mathbf{G}(\mathbf{s}_{k-2})\ldots\mathbf{G}(\mathbf{s}_0)\big]_{ij},$$

$$\mathbf{G}_{ij}(\mathbf{s}_k) = \frac{\partial\,\mathbf{f}_i}{\partial\,\mathbf{s}^{(j)}}\bigg|_{\mathbf{s}=\mathbf{s}_k},$$

where $\mathbf{G}$ – Jacobian matrix of system (1).

It follows, that to calculate Λ it is required to have at least an adequate model of system (1) and a Jacobian matrix $\mathbf{G}$. As a rule, in real-life applications, different variations of the Benettin algorithm [1980] are used in such cases.

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

It estimated the local exponential divergence of two close trajectories and can be calculated from the eigenvalues of matrix

$$\mathbf{M}_{ij}^{k} = \left[\mathbf{G}(\mathbf{s}_{k-1})\,\mathbf{G}(\mathbf{s}_{k-2})\ldots\mathbf{G}(\mathbf{s}_0)\right]_{ij},$$

$$\mathbf{G}_{ij}(\mathbf{s}_k) = \frac{\partial\,\mathbf{f}_i}{\partial\,\mathbf{s}^{(j)}}\bigg|_{\mathbf{s}=\mathbf{s}_k},$$

where $\mathbf{G}$ – Jacobian matrix of system (1).

In a unidimensional case $N = 1$, the problem is reduce and can be calculated directly:

$$\Lambda\big(s_0\big) = \lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\ln\left|\frac{\mathrm{d}}{\mathrm{d}\,s}f^{k}(s_0)\right|.$$

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

In cases where Λ is estimated by the observable realization of a dynamic process, the first approach, chronologically, was proposed by Wolf [1985].

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

In cases where Λ is estimated by the observable realization of a dynamic process, the first approach, chronologically, was proposed by Wolf [1985].

Problems:

- However, various metods are mostly centered around attractor reconstruction (especially for $N = 1$ cases), which is generally unstable when the observed trajectories are short (approximately $K < 10^4$ sequence elements).

Lyapunov exponent

One of the main diagnostic properties that determines whether the investigated system (1) is in a chaotic state is the Lyapunov exponent Λ.

In cases where Λ is estimated by the observable realization of a dynamic process, the first approach, chronologically, was proposed by Wolf [1985].

Problems:

- However, various metods are mostly centered around attractor reconstruction (especially for $N = 1$ cases), which is generally unstable when the observed trajectories are short (approximately $K < 10^4$ sequence elements).

This paper proposes an original approach to solving the issue of estimating the Lyapunov exponent from an observed finite trajectory with $N = 1$ and $K < 10^4$ based on a neural-network estimator.

Outline section

Some theoretical premises

- First, it follows from the assumptions and conclusions of the universal approximation theorem of Cybenko [1989] that deep convolutional neural networks (given their high generalization power) can turn out to be an effective nonlinear approximator of attractor properties that are relevant for estimating the Lyapunov exponent.
- Second, estimating the Lyapunov exponent from observed realization of time series $\{\mathbf{s}_k\}_{k\in\mathrm{K}}$ in case $N = 1$ relies heavily on the Takens theorem of attractor reconstruction from its observed projection. The theorem lends solid mathematical foundations to general ideas of nonlinear autoregression, see the postulates of the method of singular spectrum analysis.

Some theoretical premises

Combining the two aforementioned aspects and applying it to the properties of convolutional neural networks necessitates convolution of time series $\{\mathbf{s}_k\}_{k\in\mathrm{K}}$ into delay matrix $\mathbf{S}$:

$$\mathbf{S} = \begin{bmatrix} s_0 & s_1 & \cdots & s_{Nc-1} \\ s_{Nc} & s_{Nc+1} & \cdots & s_{2Nc-1} \\ \cdots & \cdots & \cdots & \cdots \\ s_{(Nr-1)Nc} & s_{(Nr-1)Nc+1} & \cdots & s_{NrNc-1} \end{bmatrix},$$

where $K = Nr\,Nc$. Matrix $\mathbf{S}$ is, in turn, fed to the neural network with 2D convolutional kernels. This architecture allows for using compact kernels to estimate the $\{\mathbf{s}_k\}_{k\in\mathrm{K}}$ structure at significantly varying scales. The output layer is designed to solve the regression problem to get an estimation of Λ.

CNN estimator development

Pre-training and preliminary testing of our CNN estimator for the Lyapunov exponent used time series generated by the following logical representation:

$$s_{k+1} = f(s_k,\, \lambda) = 4\,\lambda\, s_k\,(1 - s_k),$$

where: $\lambda \in (0,\, 1]$ – control parameter, $s \in (0,\, 1)$ – system phase variable.

CNN estimator development

Pre-training and preliminary testing of our CNN estimator for the Lyapunov exponent used time series generated by the following logical representation:

$$s_{k+1} = f(s_k,\, \lambda) = 4\,\lambda\, s_k\,(1 - s_k),$$

where: $\lambda \in (0,\, 1]$ – control parameter, $s \in (0,\, 1)$ – system phase variable.

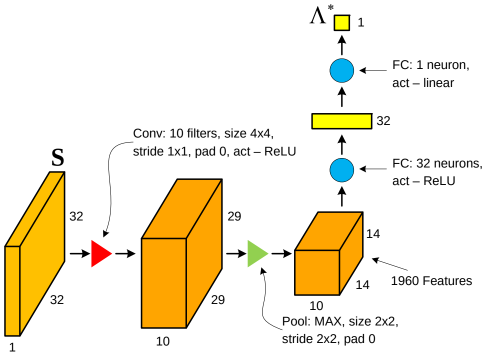
Non-overlapping sets of sequences $\{\mathbf{s}_k\}_{k \in \mathrm{K}}$ were formed to train and preliminary test the neural network estimator. Training was carried out in value range $\lambda \in [0.89,\, 0.94]$ (42 600 chaotic and 7 500 non-chaotic realizations); testing range was $\lambda \in (0.94,\, 1]$ (53 700 chaotic and 6 300 non-chaotic realizations).

CNN estimator development

Pre-training and preliminary testing of our CNN estimator for the Lyapunov exponent used time series generated by the following logical representation:

$$s_{k+1} = f(s_k,\, \lambda) = 4\,\lambda\, s_k\,(1 - s_k),$$

where: $\lambda \in (0,\, 1]$ – control parameter, $s \in (0,\, 1)$ – system phase variable.



CNN estimator testing

To estimate coherence of the classifier decisions, an network uncertainty index was introduced

$$\Phi^* = \frac{1}{2}\log\frac{\mathrm{D}\big[\Lambda^*\big]_{s_0}}{\mathrm{D}\big[\Lambda\big]_{s_0}},$$

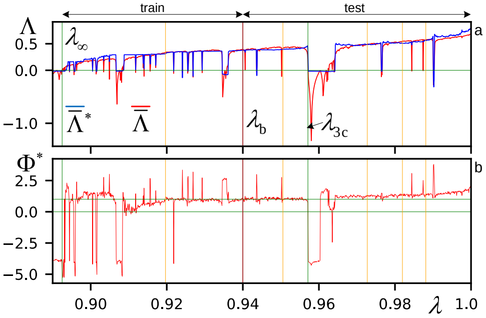
where: D – operator to calculate variance of a random variable, log – decimal logarithm.

CNN estimator testing

To estimate coherence of the classifier decisions, an network uncertainty index was introduced

$$\Phi^* = \frac{1}{2}\log\frac{\mathrm{D}\left[\Lambda^*\right]_{s_0}}{\mathrm{D}\left[\Lambda\right]_{s_0}},$$

where: D – operator to calculate variance of a random variable, log – decimal logarithm.



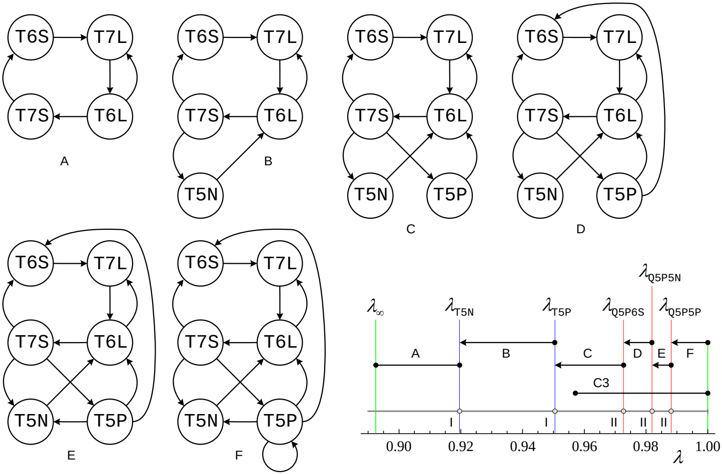
CNN estimator testing

Percentiles of measures MPE (mean percentage error) and MAPE (mean absolute percentage error) scores on the test interval (for chaotic trajectories with $\Lambda > 0$) have the following values:

%	5	25	50	75	95
MPE	-21.75	-6.83	0.83	6.16	13.13
$MAPE$	0.61	3.12	6.39	10.81	45.00

It follows from table that 75% of decisions have an error of less than 11% on the MAPE score.

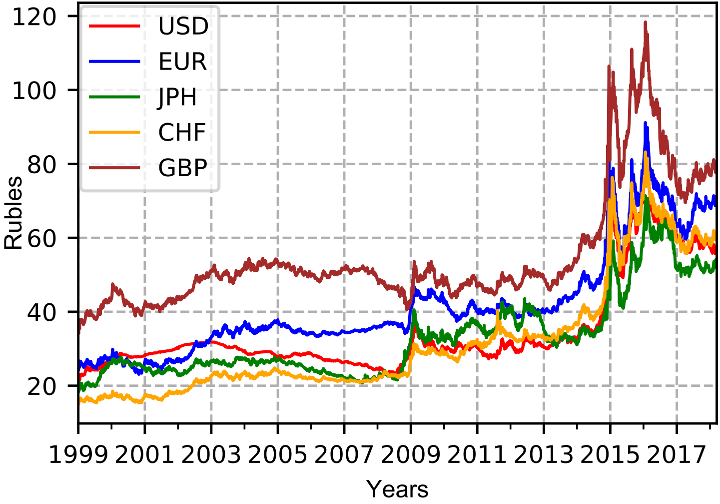
CNN estimator testing



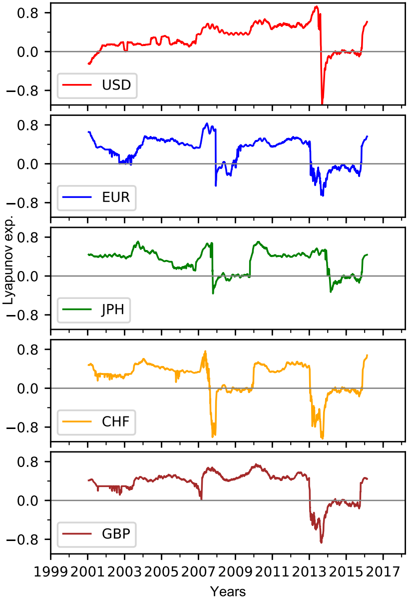
A.V. Makarenko, TQ-bifurcations in discrete dynamical systems: Analysis of qualitative rearrangements of the oscillation mode // Journal of Experimental and Theoretical Physics, 2016, Vol. 123, Issue 4, pp. 666-676.

Outline section

Financial time series



Financial time series



1. the investigated financial time series are generally chaotic;
2. during crisis events the dynamics of exchange rates usually loses its chaotic properties (possibly, due to the exercise of hands-on control over finance and economy during such times);
3. the dynamics of the pair USD/RUB rate is markedly different from the other rates (which conforms with the results of our early study).

Outline section

Summary

- This study has demonstrated effective capabilities of a relatively simple deep convolutional neural network in estimating the Lyapunov exponent..
- Unlike a number of other estimation algorithms, our solution does not require attractor reconstruction (which is generally quite unstable) and operates on rather short signals $K = 1024$ sequence elements in the experiment.
- Moreover, the proposed solution works directly with raw data, automatically synthesized informative features, makes a direct estimation of the Lyapunov exponent Λ.
- Structural analysis of the synthesized and trained convolutional network and its possible functional mechanism as applied to the problem at hand has shown that creating an input signal in the form of delay matrix and the size of 2D convolutional filters in the first hidden layer are rather critical and affect the quality of the final decision.
- This leads us to conclude tentatively that a certain class of problems in computational nonlinear dynamics can be adequately solved by deep learning.
- This study has also shown that deep neural networks are effective in applications involving processes (down to narrowband or broadband stochastic process realizations), as well as dynamic processes that have distinct patterns (e.g. speech, images, etc.).

Summary

- This study has demonstrated effective capabilities of a relatively simple deep convolutional neural network in estimating the Lyapunov exponent..
- Unlike a number of other estimation algorithms, our solution does not require attractor reconstruction (which is generally quite unstable) and operates on rather short signals $K = 1024$ sequence elements in the experiment.
- Moreover, the proposed solution works directly with raw data, automatically synthesized informative features, makes a direct estimation of the Lyapunov exponent Λ.
- Structural analysis of the synthesized and trained convolutional network and its possible functional mechanism as applied to the problem at hand has shown that creating an input signal in the form of delay matrix and the size of 2D convolutional filters in the first hidden layer are rather critical and affect the quality of the final decision.
- This leads us to conclude tentatively that a certain class of problems in computational nonlinear dynamics can be adequately solved by deep learning.
- This study has also shown that deep neural networks are effective in applications involving processes (down to narrowband or broadband stochastic process realizations), as well as dynamic processes that have distinct patterns (e.g. speech, images, etc.).

Thank you for your attention!