An Approach for the Determination of Predictable Time Series

Anja M. Busse *

Fachbereich Statistik, Universität Dortmund, Germany

Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany

Detlef Steuer

Fachbereich WOW, Universität der Bundeswehr Hamburg, Germany

Claus Weihs

Fachbereich Statistik, Universität Dortmund, Germany

March, 2001

Abstract

The forecasting of time series is one of the primary tasks in the analysis and modeling of unknown processes. Knowledge of predictability of a given time series can also be used to initially introduce a coarse classification for the modeling of the underlying processes. The main aim is to find a facility for the separation into well-predictable and not-well-predictable processes without any information about the functional relationship of the process and with only the given data. We proposed a criterion for determination of predictable time series, which is given by the size of the Lyapunov exponent.

1 Introduction

For the description and the analysis of time series it is useful to initially introduce a coarse classification in order to be able to choose the most appropriate tools for the more detailed analysis.

*email:busse@statistik.uni-dortmund.de

One important classification is to discriminate between well-predictable and not-well-predictable processes. Information about the predictability of a process facilitates e.g. a sensible choice of the forecasting window. In the case of chaotic time series the prediction accuracy can decrease considerably already after only a few time-steps in contrast to a stationary stochastic process (Abarbanel 1996), (Casdagli and Eubank 1992).

A formal identification of predictable time series can be achieved by analyzing the Lyapunov spectrum or the largest Lyapunov exponent of the time series (this is often just referred to as the Lyapunov exponent). Originally, the Lyapunov exponent was defined for non-stochastic, deterministic systems. Anyhow, the concept behind the Lyapunov exponent can be embedded into a statistical framework.

The remainder of this article is organized as follows. After a short outline of the Lyapunov exponent and its mathematical derivation in Section 2 its interpretation with respect to the determination of predictability will be introduced (Sec. 3). Experimental results illustrate the method to determine the Lyapunov exponent in Section 4.

2 Lyapunov exponent

One possibility to distinguish between well-predictable and not-well-predictable time series is given by the computation of the largest Lyapunov exponent (often briefly called the Lyapunov exponent). Firstly, it will be introduced for deterministic processes (see Sec. 2.1). In Section 2.2, it will be shown that it can also be derived for stochastic processes with additive noise.

2.1 Lyapunov exponent in a deterministic context

The dynamics of deterministic processes is defined by

$$x_{t+1} = f_t(x_0) = f(x_t) \ , \tag{1}$$

with initial point or initial state $x_0 \in \mathbb{R}^k$, x_t describes the state at time t. The sequence of states or observations x_t is called *flow* or *observation series* in order to distinguish it from a *time series* of a random process. The functional relationship is described by f and it is assumed that f is differentiable everywhere. In this work, we set $k = 1$ in equation (1).

A specific deterministic process is a chaotic process, if its asymptotic behavior is locally unstable in contrast to a regular deterministic or to an ergodic stochastic system (Abarbanel 1996) (Tong 1993) (Eckmann and Ruelle 1982). In this respect unstable means non-uniform asymptotic behavior.

The Lyapunov exponent $\lambda(x_0)$ of a flow $\{x_t\}$ is formally defined by (Eckmann and Ruelle 1982):

$$\lambda(x_0) := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)| \tag{2}$$

This characteristic feature describes the long time behavior of the average logarithmic derivative. It illustrates the divergence of two different trajectories. This can be motivated as follows:

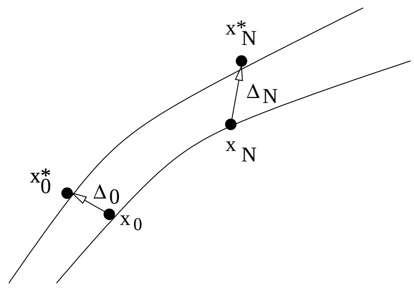

Figure 1: Two trajectories are regarded over time, in order to observe the convergence or divergence of a process.

In Fig. 1 the behavior of two nearby trajectories is shown. The trajectories of the observation series, which started from two different initial points, converge or diverge when N grows to infinity. x_0 and x_0^* denote the two different initial points for the flow (see equation (1)). The starting point x_0^* is "nearby" but displaced from x_0. Furthermore, the trajectories follow the same functional relationship. The distance between x_0 and x_0^* is given by

$$\Delta_0 = |x_0^* - x_0|. \tag{3}$$

Hence, the distance after one iteration can be approximated by applying the first order Taylor expansion follows as:

$$\Delta_1 = |x_1^* - x_1| = |f(x_0^*) - f(x_0)| \approx |f'(x_0)| \cdot |x_0^* - x_0|.$$

For the analysis of the long time behavior of the flow it is important to evaluate the distance of the two trajectories after N iterations. Using equations (1) and (3) and the chain rule this leads to:

$$\begin{aligned}
\Delta_N &= |x_N^* - x_N| \\
&\approx \left|\frac{\mathrm{d}}{\mathrm{d}x} f^N(x_0)\right| \cdot |x_0^* - x_0| \\
&= \left|\frac{\mathrm{d}}{\mathrm{d}x} f^N(x_0)\right| \cdot \Delta_0 \\
&= |f'(x_0)| \cdot |f'(x_1)| \ldots |f'(x_{N-1})| \cdot \Delta_0 \\
&= \prod_{i=0}^{N-1} |f'(x_i)| \cdot \Delta_0 \ ,
\end{aligned}$$

where $f^N = \underbrace{f \circ \cdots \circ f}_{N\ times}$.

Obviously, the expansion rate of the trajectories can be expressed by

$$\frac{\Delta_N}{\Delta_0} = \prod_{i=0}^{N-1} |f'(x_i)| \tag{4}$$

$$=: e^{N \cdot \lambda_N(x_0)} \ , \tag{5}$$

where λ_N is the characteristic value dependent on time N and x_0. This expansion rate illustrates the behavior of the trajectories after N iterations in dependence of Δ_0 and x_0.

Now, it is possible to define the asymptotic Lyapunov exponent by considering the asymptotic behavior $N \to \infty$ and by transforming equation (5):

$$\lambda(x_0) := \lim_{N\to\infty} \frac{1}{N} \ln\left(\frac{\Delta_N}{\Delta_0}\right) \tag{6}$$

$$\approx \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)|. \tag{7}$$

The Lyapunov exponent measures the asymptotic average logarithmic expansion rate along two trajectories. Note that $\lambda(x_0)$ depends on the initial point x_0.

The derivative f' of the function f is often unknown. It has to be evaluated from the given observation series. On the one hand f' can be numerically evaluated, on the other hand the divergence of the two nearby trajectories can be graphically analyzed. Various approaches for the evaluation of λ have been suggested in the literature (for more details see Abarbanel (1996), Eckmann and Ruelle (1982), Sano and Sawada (1985), Gencay (1996), Wolf et al. (1985), Kantz and Schreiber (1997)). However, the Lyapunov exponent is an attribute of the attractor, which does not depend on time. This property implies that observations used for the evaluation of λ must not be taken in the so-called *transient state*. The series has to be in the asymptotic state for the data to be used for the evaluation of f'. An inadequate evaluation of f' may be due to the observation series still lasting in the transient state.

2.2 Lyapunov exponent in a statistical context

In contrast to the deterministic processes a stochastic process is a functional relationship with random noise ϵ, which reads

$$X_{t+1} = f_t(X_0) + \epsilon = f(X_t) + \epsilon. \tag{8}$$

It is assumed that such a process is a sequence of random variables, where X_0 is the random variable realized in the initial point x_0. The random variable X_t describes the state at time t, the realization or observation of which is denoted by x_t. In this work the functional relationship f is stochastically disturbed with additive noise ϵ. The asymptotic behavior of a stochastic process should be independent of the initial state. In contrast to the *flow* the realizations of the sequence of random variables X_t is called *time series*. If the stochastic process is stationary, its asymptotic behavior should be uniform (so-called "stable") and independent of the initial state (Kendall and Ord (1990)).

Note that it is conceptually possible to transfer a deterministic observation series into a stochastic time series by assuming a functional relationship and a noise ϵ with a one-point distribution.

Moreover, if we compose the functional relationship f and the stochastic error ϵ, we obtain

$$X_{t+1} = g(x), \quad \text{with} \quad g(x) = f(x) + \epsilon \quad \text{and} \quad g'(x) = f'(x). \tag{9}$$

By generalizing equation 2 in an obvious way the Lyapunov exponent for stochastic processes is given by:

$$\lambda(X_0) := \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |g'(X_i)| \tag{10}$$

$$= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(X_i)| \tag{11}$$

The additive error disappears by the use of the derivative.

3 Prediction characterized by the Lyapunov exponent

Now the Lyapunov exponent can be used to distinguish between well-predictable and not-well-predictable processes. For this it is necessary to discuss the relation between the Lyapunov exponent and the loss of information.

It is assumed that x_0 is in a ball I_0, at time n x_n is in I_n and at time $n+1$ it is in I_{n+1} (see Fig. 2).

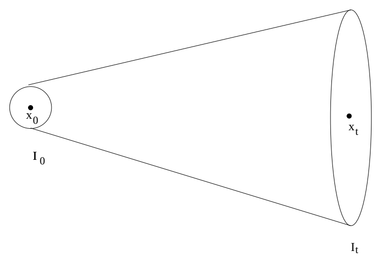

Figure 2: Information loss or increase from I_0 to I_t

In addition, in a small interval more information about the true position of x_n in I_n exists in contrast to a larger interval (see Beck and Schlögl (1993)). As a measure of the information content the (exponential) bit-number b_n is suggested, whereas b_n is defined by

$$b_n := \ln \frac{1}{\Delta_n} = -\ln(\Delta_n). \tag{12}$$

Thus, for the length Δ_n of the interval I_n follows by transformation of equation 12:

$$\Delta_n = \exp(-b_n). \tag{13}$$

Now, the information loss is the difference of the bit-numbers before and after the iteration step:

$$b_n - b_{n+1} = \ln \Delta_{n+1} - \ln \Delta_n \approx \ln |f'(x_n)|. \tag{14}$$

Proof:

$$\begin{aligned} b_n - b_{n+1} &= -\ln \Delta_n + \ln \Delta_{n+1} \\ &= \ln \Delta_{n+1} - \ln \Delta_n = \ln \left(\frac{\Delta_{n+1}}{\Delta_n} \right) \\ &\approx \ln |f'(x_n)| \end{aligned}$$

because of

$$\begin{aligned} \Delta_1 &\approx |f'(x_0)| \cdot \Delta_0 \quad \text{and thus} \\ \Delta_{n+1} &\approx |f'(x_n)| \cdot \Delta_n. \end{aligned}$$

q.e.d

The difference (14) between the bit-numbers can be interpreted either as a loss or an increase of information. Consequently, the Lyapunov exponent can be interpreted as the average information loss or information increase, since

$$\begin{aligned} \lambda(X_0) &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(X_i)| \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} (b_i - b_{i+1}). \end{aligned}$$

The separation between well-predictable and not-well-predictable processes is given by the sign of the Lyapunov exponent because of a positive or negative expansion.

- If $\lambda(x_0) < 0 \Leftrightarrow \Delta_N < \Delta_0$, the trajectory of the given time or observation series asymptotically reaches a stable limit point or a stable limit cycle. Convergence is given in this case. The information increases. Consequently, the process is well-predictable.

- If $\lambda(x_0) \approx 0 \Leftrightarrow \Delta_N \approx \Delta_0$, the process behaves like a random walk, since for a random walk $\Delta_1 \approx 1 \cdot \Delta_0$, and thus $\Delta_N \approx \Delta_0$.
- If $\lambda(x_0) > 0 \Leftrightarrow \Delta_N > \Delta_0$, the trajectories of two different, nearby initial points diverge exponentially on average by a factor of $e^{N\lambda}$ after N iterations. In this case, the limiting behavior is not uniform, it is unstable and in literature it is denoted by *strange attractor* (Eckmann and Ruelle 1982), (Grassberger and Procaccia 1983a) (Grassberger and Procaccia 1983b). The information decreases, as a result of which the process is not-well-predictable.

This separation criterion can be used without any modeling of the process. It is possible to distinguish well-predictable and not-well-predictable processes by only the given time or observation series.

4 Experimental results

Neither the original time series nor its spectral density give much information, which can be used to distinguish between well-predictable and not-well-predictable time series (see Fig. 3). The Lyapunov exponent achieves this classification. It can be easily computed, if the functional relationship of the time series is given. However, the underlying process is generally unknown in case of real-world problems. Consequently, it is necessary to evaluate a proper estimator $\hat{\lambda}$ from the given time series. The method, which is based on work of Sano and Sawada (1985) and which was implemented by Kantz and Schreiber (1997), is applied to the estimation of the Lyapunov exponent in the following examples.

4.1 Experiments with well-predictable and not-well-predictable sets

We applied the method of Lyapunov exponent estimation to the function (15), which generates a well-predictable process, as well as to the function (16), which does not create a well-predictable behavior. The well-predictable time series is a stationary stochastic time series and the not-well-predictable process is created by an chaotic observation series. Additionally, the exact Lyapunov exponent of the processes can be evaluated analytically.

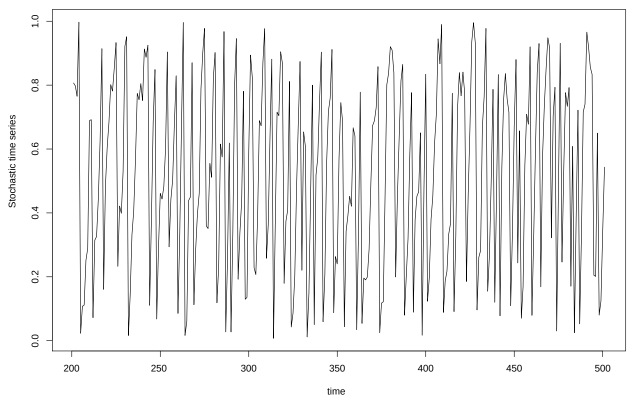

Figure 3: A part of the stochastic time series from equation (15) is shown. This process is an example for a well-predictable process.

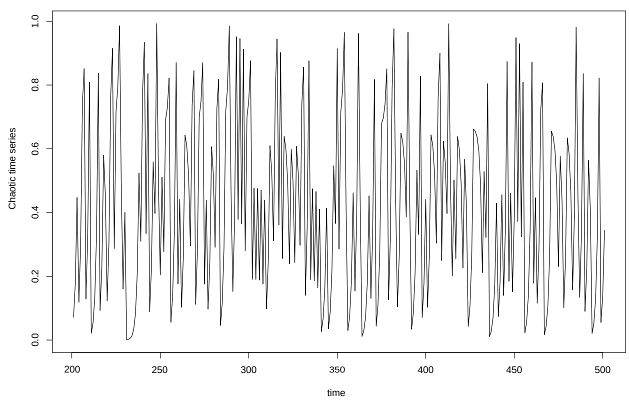

Figure 4: The Modulo function from equation (16) generates a chaotic time series, an example for a not-well-predictable observation series.

To generate a well-predictable time series a uniformly distributed noise term, $U[0,1]$, is used in the following way:

$$x_t = (0.9x_{t-1} + 0.5\epsilon), \epsilon \sim U[0,1]. \tag{15}$$

The not-well-predictable time series is generated by

$$x_t = (2.5x_{t-1}) \mod 1. \tag{16}$$

In both cases the initial point $x_0 = 0.699$ is used and the sample size is 1024.

Corresponding trajectories for the time series of equation (15) and equation (16) are illustrated in Fig. 3 and 4.

No cycles or trends can be detected in the figures. Both the time series and the observation series look like similar random processes.

If we observe the spectral densities of the time series (see Fig. 5 and 6), we come to the same result. There are no obvious differences between the stochastic and the chaotic time series. Additional usual time series characteristics (e.g. autocorrelation function or partial autocorrelation function) yielded likewise no separation.

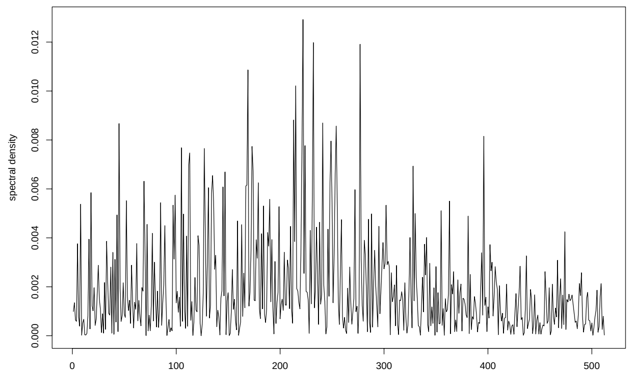

Figure 5: The well-predictability is not reflected in the spectral density of the process of equation (15).

However, it will be shown that the estimation of the Lyapunov exponent yields good results with respect to predictability separation. For the well-predictable process $\hat{\lambda} = -0.0945$ was estimated with a real Lyapunov exponent of $\lambda = -0.11$, i.e. $\hat{\lambda} < 0$, which describes a well-predictable process.

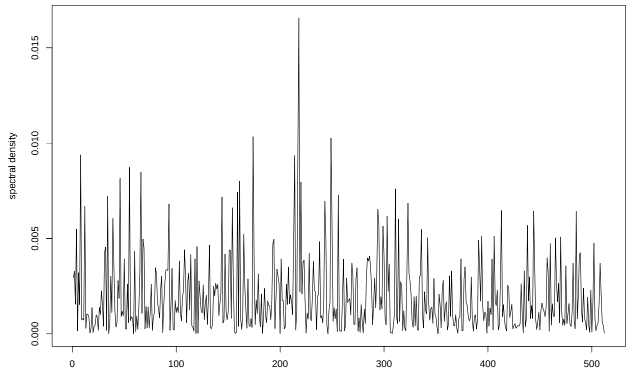

Figure 6: The analysis of the spectral density of the chaotic process provides no results with respect to predictability.

In the not-well-predictable situation the estimation of the Lyapunov exponent leads to $\hat{\lambda} = 0.92$ ($\lambda = 0.92$), i.e. the property of $\hat{\lambda} > 0$ is fulfilled. In contrast to the evaluation of the original time series or the spectral densities Lyapunov exponent estimation can correctly classify the different processes.

5 Conclusion

We analyzed the Lyapunov exponent in the context of the separation between well-predictable and not-well-predictable processes. A classification seems useful since it would facilitate a more detailed analysis of the underlying process with respect to the choice of the appropriate tools. In this work it was suggested the Lyapunov exponent for separation. This criterion describes the "stable" or "unstable" asymptotic behavior of a process.

It was theoretically shown that the Lyapunov exponent can be used for the evaluation of predictability. The Lyapunov exponent distinguishes well-predictable and not-well-predictable processes only by using the given time or observation series and without any information about the functional relationship. In addition, well-predictable and not-well-predictable processes were inspected with respect to separation. It was shown that the estimation yields correct classifications for both forms of processes. In case of not-well-predictable time series the estimate was even exactly evaluated with respect to the true Lyapunov exponent.

Acknowledgements

The first author thanks Peter Stagge for his helpful comments, suggestions and discussions. The financial support of the Deutsche Forschungsgemeinschaft (SFB 475, Reduction for complexity in multivariate data structures) is gratefully acknowledged.

References

Abarbanel, H. D. I. (1996). *Analysis of Observed Chaotic Data.* Institute for Nonlinear Science. New York: Springer Verlag.

Beck, C. and F. Schlögl (1993). *Thermodynamics of chaotic systems.* Cambridge: Cambridge University Press.

Casdagli, M. and S. Eubank (1992). *Nonlinear Modeling and Forecasting.* Adison-Wesley.

Eckmann, J. P. and D. Ruelle (1982). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics 57*, 617–657.

Gencay, R. (1996). A statistical framework for testing chaotic dynamics via lyapunov exponents. *Physica D 89*, 261–266.

Grassberger, P. and I. Procaccia (1983a). Characterization of strange attractors. *Physical Review Letters 50*, 346–349.

Grassberger, P. and I. Procaccia (1983b). Measuring the strangeness of strange attractors. *Physica D 9*, 189–208.

Kantz, H. and T. Schreiber (1997). *Nonlinear time series analysis.* Cambridge: Cambridge University Press.

Kendall, M. and J. K. Ord (1990). *Time Series, Third Edition.* Suffolk: Edward Arnold.

Sano, M. and Y. Sawada (1985). Measurements of the lyapunov spectrum from a chaotic time series. *Physical Review Letters 55*, 1082–1085.

Tong, H. (1993). *Non-linear Time Series, A Dynamical System Approach.* Oxford Statistical Science Series. New York: Oxford University Press.

Wolf, A., J. B. Swinney, and H. L. Vastan (1985). Determining lyapunov exponents from a time series. *Physika D 16*, 285–317.