ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Busse, Anja M.

Working Paper

Classification of Processes by the Lyapunov exponent

Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen, No. 2004,70

Provided in cooperation with:
Technische Universität Dortmund

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der
Leibniz
Leibniz-Gemeinschaft

Classification of Processes by the Lyapunov exponent

Anja M. Busse

University of Dortmund[⋆]
Department of Statistics
44221 Dortmund, Germany

Abstract. This paper deals with the problem of the discrimination between well-predictable and not-well-predictable time series. One criterion for the separation is given by the size of the Lyapunov exponent, which was originally defined for deterministic systems. However, the Lyapunov exponent can also be analyzed and used for stochastic time series. Experimental results illustrate the classification between well-predictable and not-well-predictable time series.

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.

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 (1991)).

In addition, in the analysis of stochastic processes there often is the problem that only one time series is available and no previous knowledge about the temporal-functional relationship is given.

Despite these restrictions 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 paper is organized as follows. After an introduction of the Lyapunov exponent (Sec 2) we will show that it can be used as a criterion to discriminate between well-predictable and not-well predictable time series (Sec 3). Experimental results of a *BTA*-deep-hole drilling process
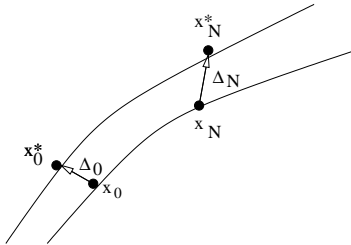
Fig. 1. Two trajectories are regarded over time in order to observe the convergence or divergence of a process.

illustrate the method of separation by the Lyapunov exponent (Sec 4). A conclusion is drawn in Sec. 5.

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). This was originally defined for non-stochastic, deterministic processes. However, the Lyapunov exponent can also be analyzed and used for the stochastic case.

Firstly, it will be introduced for deterministic processes. 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 functional relationship is described by f and it is assumed that f is differentiable everywhere. Hence, the dynamics is entirely deterministic.

The Lyapunov exponent describes the divergence of two different trajectories. This can be motivated as follows:

In Fig. 1 the behavior of two nearby trajectories is shown. 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

$$\Delta_0 = |x_0^* - x_0|. \tag{2}$$

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

$$\Delta_1 = |x_1^* - x_1| = |f(x_0^*) - f(x_0)| \approx |f'(x_0)| \cdot |x_0^* - x_0|. \tag{3}$$

After N iterations the distance between the trajectories arises from using the chain rule:

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

Thus, we are interested in diverging or converging of the trajectories after N iterations in comparison to the beginning. This is estimated by an expansion rate. Obviously, the expansion rate of the trajectories can be expressed by

$$\frac{\Delta_N}{\Delta_0} \approx \prod_{i=0}^{N-1} |f'(x_i)| = e^{N \cdot \lambda_N(x_0)} \, , \tag{5}$$

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.

The consideration of the asymptotic behavior for $N \rightarrow \infty$ yields the definition for the Lyapunov exponent of deterministic processes:

$$\lambda(x_0) := \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)|. \tag{6}$$

It is the long time consideration of the average logarithmic derivation after N equals infinity many iterations. The Lyapunov exponent measures the asymptotic average logarithmic expansion rate along two trajectories.

The derivative f' of the function f is often unknown. It has to be evaluated from the given observation series. Various approaches for the calculation of λ have been suggested in the literature (for more details see for example Sano and Sawada (1985), Kantz and Schreiber (1997)).

If stochastic processes are considered, two cases have to be distinguished separately: The random effect is additive in the functional equation and the random effect is not necessarily additive.

First the case with an additive noise is considered. The dynamics of stochastic processes with an additive random effect is defined by

$$X_{t+1} = f(X_t) + \epsilon_t. \tag{7}$$

By transforming

$$X_{t+1} = g(X_t, \epsilon_t), \quad \text{with} \quad g(X_t, \epsilon_t) = f(X_t) + \epsilon_t \tag{8}$$

we obtain the same derivatives of g and f so that the definition of the Lyapunov exponent for stochastic processes with an additive noise is directly derived from the deterministic case. The function g is inserted in the definition of the Lyapunov exponent for deterministic processes and the definition for stochastic processes with an additive noise is obtained:

$$\lambda(X_0) := \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln ||\frac{d}{dX_t}g(X_t(x_i), \epsilon_t)|. \tag{9}$$

However, an additive noise can not always be justified because this assumption is too restrictive with regard to possible model classes. Thus, the
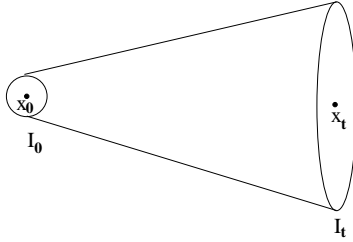
Fig. 2. Information loss of the information area I_0 in comparison to the information area I_t

general case is considered. The dynamics of stochastic processes with a non-necessarily additive noise is defined by

$$X_{t+1} = h(X_t, \epsilon_t). \tag{10}$$

The Lyapunov exponent can be naturally generalized as:

$$\tilde{\lambda}(x_0) := \lim_{N \to \infty} \frac{1}{N} \sum t = 0 N - 1 [\ln |\frac{d}{dX_t} h(X_t(x_0), \epsilon_t)|] \tag{11}$$

In any case, as an estimator of the Lyapunov exponent

$$\hat{\lambda} = \frac{1}{N} \sum_{i=0}^{N-1} \ln |\frac{d}{dX_t} h(X_t, \epsilon_t)|. \tag{12}$$

is taken. For details about the Lyapunov exponent for deterministic and stochastic processes see Busse et al. (2001), Busse (2003), and Busse and Weihs (2004).

3 Well-predictable and not-well-predictable processes

The knowlegde about the quality of prediction of processes is an important property for the interpretation of the predicted results. The greater the information loss in a multi-step-forcasting the greater the decrease in the quality of prediction. Thus, it is interesting to know a measure of information loss for avoiding possible misinterpretations. The Lyapunov exponent can be interpreted as an expansion rate with a direct context to the information loss over time.

If we assume that the true starting point x_0 of a time series is possibly displaced by an ϵ, we know only the information area about the starting point we do not know the proper position of x_0. After t-time steps the time series is in the information area at time t, I_t and after $t + 1$-time steps in the information area I_{t+1} (Fig. 2). If the information area is small, we have more information about the true position of the data point in contrast to a greater

information area (Beck (1993)). As an adequate measure of information the information content b_n of a true position of a data point in an information area I_n of the volume Δ_n is given by:

$$b_n := \ln \frac{1}{\Delta_n} = -\ln(\Delta_n). \tag{13}$$

The connection to the volume of an information area is given by

$$\Delta_n = \exp(-b_n). \tag{14}$$

It can be characterized by the distance between two trajectories of a process at time n. For the evaluation of the quality of prediction we are interested in the information loss from one time to the next. For this the difference of two information contents before and after an iteration step are determined. Thus, the information loss IV about the true position of a data point in one iteration step is given by

$$IV = b_n - b_{n+1} = \ln \Delta_{n+1} - \ln \Delta_n \approx \ln |f'(x_n)|, \tag{15}$$

with $\Delta_{n+1} \approx |f'(x_n)| \cdot \Delta_n$. If the difference is positive, IV describes an information increase, whereas an information loss is given, if I_n is less than I_{n+1}.

The information loss is the logarithmic first derivative of the functional relationship of a process, so that the Lyapunov exponent can be used for the description of the average information loss:

$$\lambda(X_0) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(X_i)| \approx \lim_{N \to \infty} \frac{1}{N} \sum_{i=0}^{N-1} (b_i - b_{i+1}). \tag{16}$$

In contrast to the traditional classification of time series we do not use the given data points but the possible position areas like *k-means clustering* (Hastie (2001)).

The classification of both deterministic and stochastic processes by the Lyapunov exponent is given by:

- $\lambda(x_0) < 0 \Leftrightarrow \Delta_N < \Delta_0 \Rightarrow$ good predictability
  The information about the true position of the data increases due to the reduction of the information area. Consequently, we get a good predictability.
- $\lambda(x_0) \approx 0 \Leftrightarrow \Delta_N \approx \Delta_0 \Rightarrow$ predictability like a random walk
  Here, the information content levels off. We have neither information loss nor information increase.
- $\lambda(x_0) > 0 \Leftrightarrow \Delta_N > \Delta_0 \Rightarrow$ bad predictability
  The information loss about the true position of the data increases over time due to the information area increases. Consequently, we get a bad predictability.
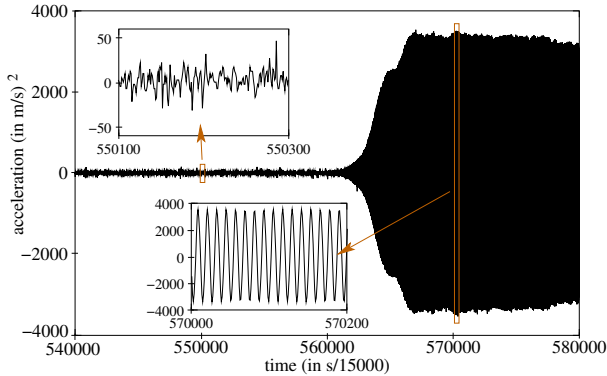
Fig. 3. Acceleration of the drill head in the non-chatter (left) and the chatter (right) area

4   Experimental results

The Lyapunov exponent achieved a distinction between well-predictable and not-well-predictable time series. We applied this classification to a real-world-problem. The aim was to analyze a *BTA*-deep-hole drilling process and to control at best working conditions (VDI (1974)). *BTA*-deep-hole drilling is used to produce holes with a high length-to-diameter-ratio. But the slenderness of the tool can yield unwanted states, like chatter. This should be avoided, because chatter generates surface discontinuity at the workpiece, noise exposure and increases the wear of cutting edges substantially.

For the analysis we were given a time series of acceleration data with different types of the process. First the non-chatter area with a weakly periodical part and the chatter area with a strongly periodical part. The transition between these areas appears funnel shaped.

The aim was to identify the chatter early to avoid the possible consequences. For this we characterize the transition in time windows of length 1024 data points. We chose the Lyapunov exponent because its ability to distinguish between good and bad predictability makes it possible to estimate the starting point of the transition. For every time window in the transition area the Lyapunov exponent of the given time series was evaluated. For this, equation (6) and equation (12) respectively is estimated by the approach of Kantz and Schreiber (1997). In order to identify the "true" transition it is important whether the Lyapunov exponent is less than 0, that is to classify a good forecast-property or whether λ is greater than 0.

For the interpretation of the resultsnote that amplitude increase is in the time window 559313–560335. The identification of the change between the non-chatter and the chatter area occured one and a half drill rotations earlier than the real amplitude increase (see Table 1, 556240–558287).

Table 1. Lyapunov exponent in time windows in the transition area.

time windows (data points)	Lyapunov exponent	classification decision
550096–551119	0.004	> 0
551120–552143	0.009	> 0
552144–553167	0.008	> 0
553168–554191	0.019	> 0
554192–555215	0.016	> 0
555216–556239	0.012	> 0
556240–557263	0.003	> 0
557264–558287	-0.004	< 0
558288–559312	-0.003	< 0
559313–560335	-0.004	< 0
560336–561359	-0.004	< 0
561360–562383	-0.013	< 0

The distinction between well predictable and not-well predictable processes by the Lyapunov exponent was applied with good results to various time series. For more details about other applications see for example Busse (2003).

5   Conclusion

We analyzed the Lyapunov exponent in the context of the separation between well-predictable and not-well-predictable processes. Such 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 the Lyapunov exponent was suggested for separation. This criterion describes the asymptotical average logarithmic expansion of the model derivative.

It was shown that the Lyapunov exponent can be used for the evaluation of predictability. The Lyapunov exponent as a classification criterion can be used without knowledge stochastics of the process and without knowledge about the temporal-functional relationship, only using the given time series.

In addition, different areas of a *BTA*-deep-hole drilling process were classified by the Lyapunov exponent. Detection of the transition to chatter was possible substantially earlier than the rise in acceleration was visible.

References

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

Arnold, V. I. and A. Avez (1968). *Ergodic problems of classical mechanics*. W. A. Benjamin, New York.

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

Busse, A. M., D. Steuer and C. Weihs (2001). An Approach for the Determination of Predictable Time Series. *Technischer Bericht 12, SFB 475, Fachbereich Statistik, Universität Dortmund, 44221 Dortmund, Germany.*

Busse, A. M. (2003). *Klassifikation von Datenreihen mit Hilfe des Lyapunov–Exponenten*. Dissertation Fachbereich Statistik. Universität Dortmund. eldorado.uni-dortmund.de:8080/FB5/ls7/forschung/2003/Busse.

Busse, A. M. and C. Weihs (2004). Lyapunov exponent for stochastic time series. *Technischer Bericht 37, SFB 475, Fachbereich Statistik, Universität Dortmund, 44221 Dortmund, Germany.*

Casdagli, M. (1991). Chaos and deterministic versus stochastic non-linear modeling. *J. R. Statist. Soc. B 54, 303–328.*

Hastie, T., R. Tibshirani and J. Friedman (2001). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Springer, New York.

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

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

Stout, W. F. (1974). *Almost sure convergence*. Academic Press, New York.

VDI-Richtlinie 3210 (1974, Juni). Tiefbohrverfahren. VDI, Düsseldorf.