Local Lyapunov Exponents:
A new way to predict chaotic systems[*]

Dominique Guégan[†] and Justin Leroux[‡]

September 29, 2008

Abstract

We propose a novel methodology for forecasting chaotic systems which is based on exploiting the information conveyed by the local Lyapunov exponent of a system. We show how our methodology can improve forecasting within the attractor and illustrate our results on the Lorenz system.

[†]PSE, CES-MSE, Université Paris-I Panthéon Sorbonne, 106 boulevard de l'Hôpital, 75013 Paris, France (Email: Dominique.Guegan@univ-paris1.fr)

[‡]Institute for Applied Economics, HEC Montréal and CIRPÉE, 3000 chemin de la Côte-Ste-Catherine, Montréal, QC H3T 2A7, Canada. (Email: justin.leroux@hec.ca, Fax: (1) 514-340-6469, _corresponding author_)

1 Introduction

It is well known that some deterministic chaotic systems can be difficult to forecast accurately. Indeed, their extreme sensitivity to initial conditions amplifies slight deviations from a trajectory in the state space into dramatic changes in future behavior.

In this chapter, we propose a novel methodology for forecasting deterministic series which corrects for the inevitable bias of most non-parametric predictors (such as the ones based on kernels, radial functions, neural nets, wavelets, etc.; see [1] and [2]) by incorporating additional information on the local chaoticity of the system via the so-called local Lyapunov exponent (LLE). To the best of our knowledge, while several works exist on the forecasting of chaotic systems (see, e.g., [3] [4], [5] and [6]), none exploit the information conveyed by the LLE. The general intuition of the methodology we propose in this chapter can be viewed as a complement to existing forecasting methods, and can be extended to chaotic time series. For illustrative purposes, we describe how our methodology can be used to improve upon the well-known nearest-neighbor predictor on the Lorenz system.

The nearest-neighbor predictor has proved to be a simple yet useful tool for forecasting chaotic systems ([7] and [8]). In the case of a one-neighbor predictor, it takes the observation in the past which most resembles today's state and returns that observation's successor as a predictor of tomorrow's state. The rationale behind this nearest-neighbor predictor is quite simple: given that the system is assumed to be deterministic and ergodic, one obtains a sensible prediction of the variable's future by looking back at its evolution from a similar, past situation. For predictions more than one step ahead, the procedure is iterated by successively merging the predicted values with the observed data.

We show that we can improve predictions in a chaotic system by incorporating information carried by the system's LLE. The LLE (see [9], [10]) represents the local dispersion rate of a system at a given point: a positive value meaning that two nearby points in the state space tend to grow apart over time, while a negative value indicates that nearby points will come closer together in the near future (but may diverge later on). In other words, the LLE is a measure of local chaoticity of a system, as sensitivity to initial conditions is characteristic of chaotic systems.

By definition, the LLE tells us precisely by how much the distance between the current state and its nearest neighbor will expand (or contract) over time, so

that we can easily obtain the distance between the nearest-neighbor predictor (i.e., the neighbor's successor) and the future we are predicting (tomorrow's state). Thus, we know exactly by how much to correct the prediction of the nearest-neighbor predictor. We use this fact to develop a new methodology.

The rest of the chapter is organized as follows. In Section 2, we develop our methodology by first pointing out why the nearest-neighbor predictor is biased and then suggesting how to correct this bias using information carried by the system's LLE. In Section 3, we present simulations carried out on the well-known Lorenz system to illustrate the extent of the substantial potential accuracy gains our methodology yields. In Section 4, we present how one can achieve these potential gains almost perfectly thanks to a simple rule of thumb to solve the selection problem arising from our methodology. Section 5 concludes by pointing to directions in future work in order to refine our selection process and our methodology in general.

2 Methodology

Consider a one-dimensional series of T observations from a chaotic system, $(x_1, ... x_T)$, whose future values we are trying to forecast, d is the embedding dimension used to detect the attractor. A possible embedding method, involving building a d-dimensional orbit, (X_t), with $X_t = (x_t, x_{t-1}, ..., x_{t-(d-1)})$, is described in [11].

By definition, the local Lyapunov exponent (or LLE) of a dynamic system characterizes the rate of separation of infinitesimally close points of an orbit. Quantitatively, two neighboring points in phase space with initial separation δX_0 are separated, t periods later, by the distance:

$$|\delta X| \approx |\delta X_0| e^{\lambda_0 t},$$

where $|\cdot|$ represents the modulus of the considered vectors and λ_0 is the local Lyapunov exponent of the system in the vicinity of the initial points. Typically, this local rate of divergence (or convergence, if $\lambda_0 < 0$) depends on the orientation of the initial vector δX_0. A dynamic system is considered to be (locally) chaotic if $\lambda_0 > 0$, and (locally) stable if $\lambda_0 < 0$. (see, e.g., [12]).

Our goal is to exploit the local information carried by the LLE to improve upon existing methods of reconstruction and prediction. We propose a methodology which builds upon the classical nearest-neighbor predictor, which we now

recall.

Consider an orbit $(X_1, ..., X_T)$ whose one-step-ahead future, X_{T+1}, we are trying to predict. The nearest-neighbor predictor returns $\hat{X}_{T+1} = X_{i+1}$, where X_i is the element of the orbit with minimal distance to X_T. Because the dynamic system at hand is aperiodic (or else, forecasting would not be an issue), the nearest-neighbor predictor is inevitably biased. Indeed, because $|X_T - X_i| > 0$, it must also be the case that:

$$|\hat{X}_{T+1} - X_{i+1}| \approx |X_T - X_i|e^{\lambda_i} > 0, \tag{1}$$

where λ_i can be approximated in practice by the following expression:

$$\hat{\lambda}_i = \ln \frac{|X_{i+1} - X_{j+1}|}{|X_i - X_j|} \quad \text{with} \quad X_j = \arg \min_{t \neq i, T} |X_t - X_i| \tag{2}$$

It follows from Expression (1) that knowing the distance between the predictee and the nearest neighbor as well as the LLE at the nearest neighbor allows us to predict the distance of the predictee's image to the neighbor's image. Note that this is true regardless of the sign of λ_i; i.e., regardless of whether the system is locally chaotic or locally stable. Moreover, because the orbit considered results from the embedding of a one-dimensional series, we also know all but the first coordinate of $X_{T+1} = (x_{T+1}, x_T, ..., x_{T-d+2})$. Hence, X_{T+1} lies at the intersection of the sphere of radius $|X_T - X_i|e^{\hat{\lambda}_i}$ centered on X_T and the line defined by $\{(z, x_T, ..., x_{T-d+2}) | z \in \mathbb{R}\}$ which, in the Euclidean space, amounts to solving the following polynomial for $z \in \mathbb{R}$:

$$\sqrt{(z - x_{i+1})^2 + (x_T - x_i)^2 + ... + (x_{T-d+2} - x_{i-d+2})^2} - |X_T - X_i|e^{\hat{\lambda}_i} = 0 \tag{3}$$

Typically, two candidates emerge, $\hat{x}_{T+1}^-$ and $\hat{x}_{T+1}^+$, respectively underestimating and overestimating the true value of observation x_{T+1}(see Figure 1)[1].

[FIGURE 1]

One difficulty lies in determining when the nearest-neighbor predictor overestimates or underestimates the true value to be predicted. Being able to discriminate accurately between $\hat{x}_{T+1}^-$ and $\hat{x}_{T+1}^+$ may significantly improve the accuracy

---

[1] The situation whereby Expression (3) has no real solution would only arise if λ_i had been greatly underestimated, which never occurred to us in practice using Expression (2).

of the nearest-neighbor predictor, as we next illustrate. We discuss this point in Section 4.

3   Simulations

We illustrate our point by simulating a well-known chaotic system: the Lorenz system (see [13]), which is characterized by the following system of differential equations:

$$\begin{cases} \frac{dx}{dt} = \sigma(y - x) \\ \frac{dy}{dt} = x(R - z) - y \\ \frac{dz}{dt} = xy - bz. \end{cases}$$

We simulated this system with values $\sigma = 16$, $R = 45.92$ and $b = 4$, initial values $x_0 = -10$, $y_0 = -10$ and $z_0 = 30$, and a step size of 0.01. Taking 5,000 observations, deleting the first 1,000 ensure that we are working within the attractor and considering the values on the x-coordinate as its own series, we successively predicted the last 1,000 in-sample observations. Each prediction was carried out with the full—and true—information set leading up to it, each time using the best of the two candidates, $\hat{x}_{T+1}^-$ and $\hat{x}_{T+1}^+$ (measured in distance to the—known—successor). We obtain results which are always better than with the nearest-neighbor predictor and a mean-squared error which is roughly two-thirds that of the nearest-neighbor predictor: 0.1552 for the raw nearest-neighbor predictor, versus 0.1056 after LLE-correction.

This example shows the interest of the method and the great potential in improving upon the accuracy of the nearest-neighbor predictor by incorporating the information contained in local Lyapunov exponents as in Expression (1). Another example based on the logistic map is available in [14].

4   Solving the selection problem

Several aspects of the implementation are still to be refined. Consistently discriminating between the two candidates, $\hat{x}_{T+1}^-$ and $\hat{x}_{T+1}^+$, can prove to be a difficult task due to the inherent chaotic nature of the systems at hand. As a first guess, one can select the candidate which maximizes the colinearity between the $X_{i+1} - X_T$ vector and the vector $\hat{X}_{T+1} - X_T$ (with $\hat{X}_{T+1}$ standing for $\hat{X}_{T+1}^-$ or $\hat{X}_{T+1}^+$). With the simulation of the Lorenz system described above, we

achieve 94.3% accuracy over 1000 predictions, suggesting that this method of discrimination might be reasonable. Results are summarized in Table 1 below.

Notice that the LLE ranges from -1.4353 to 1.4580 on the portion studied. In other words, the attractor contains stable and chaotic regions. Our predictions are more accurate on 1000 predictions than on 100, which is likely due to the fact that the corresponding latter portion of the attractor is more chaotic, which is confirmed by the fact that the LLE is larger on average in that region (0.2756 versus 0.1940).

Table 1: Lorenz system		
Number of predictions	100	1000
Success rate	87%	94.30%
Average error	0.0738	0.0390
NNP average error	0.1075	0.0548
Best candidate average error	0.0689	0.0372
otherSE	0.2041	0.1036
mean LLE (min;max)	0.2756(-0.9861;1.3639)	0.1940 (-1.4353;1.4580)
mean LLE on mistakes (min;max)	0.4685(-0.4020;1.3639)	0.4354 (-0.5142;1.3639)

Table 2 provides the size of errors on small ranges of LLE on 1000 predictions.

Table 2: Average errors and success rates						
	Average errors			Success rates		
LLE	success	failure	raw 1-NN	success	failure	success rate
-1.5,1.3	0.0077	-	0.0077	1	0	100%
-1.3,-1.1	-	-	-	-	-	-
-1.1,-0.9	0.0335	-	0.0335	3	0	100%
-0.9,-0.7	0.0286	-	0.0286	3	0	100%
-0.7,-0.5	0.0451	0.0169	0.0479	67	1	98.51%
-0.5,-0.3	0.0541	0.005	0.0594	92	1	98.91%
-0.3,-0.1	0.0505	0.0899	0.0626	98	2	97.96%
-0.1,0	0.0477	0.0347	0.0580	47	4	91.49%
0,0.1	0.0249	0.027391	0.0365	61	4	93.44%
0.1,0.3	0.0342	0.0876	0.0539	109	7	93.58%
0.3,0.5	0.0290	0.0594	0.0626	195	8	95.90%
0.5,0.7	0.0325	0.0847	0.0560	223	22	90.13%
0.7,0.9	0.0236	0.0562	0.0267	18	2	88.89%
0.9,1.1	0.0152	0.396	0.0273	13	3	76.92%
1.1,1.3	0.0433	0.0632	0.0333	9	2	77.88%
1.3,1.5	0.1594	0.0432	0.0339	4	1	75%
Overall	0.0373	0.0667	0.0548	943	57	94.3%

We observe that errors can be small even when we make mistakes, but can also be relatively large with accurate selection. However, the size of errors is relatively stable over the range of LLEs when selection is successful. This seems to indicate that our method accurately corrects for the dispersion of neighboring trajectories as measured by the value of the LLE. If this were not the case, one would expect the size of errors to be larger for larger values of LLEs. In fact, errors become large only for values of the LLE near the upper end of their range. A possible reason for this sudden increase may be that our estimator for the value of the LLE is not sufficiently robust in regions of high chaoticity. We expect that a more sophisticated estimation method for the LLE may solve this issue.

Our method selects perfectly for very low values of the LLE. Selection mistakes start to appear past a value of LLE of -0.5, but the success rate does not seem to fall significantly until the value of the LLE increases up to 0.7. The

success rate falls as the LLE grows very large, however, which is in line with the common intuition that the system is then (locally) more chaotic and, thus, more difficult to predict.

5   Concluding comments

We find that our methodology may lead to substantial improvements to existing non-parametric predictors. Despite the candidate selection problem it poses, we were able to reap most of the potential benefits thanks to a simple rule of thumb. However, more sophisticated selection procedure should be considered. For instance, it may be sensible to condition the selection process on the size of the LLE, acting on the intuition that trajectories are likely to be more stable when the LLE is small, and more erratic when it is large. More specifically, the bottom row of Table 1 suggests that the current rule of thumb makes mistakes where the value of the LLE is large. We plan on investigating this line of reasoning in future work.

Such increased precision in short-run predictions may translate into accuracy gains for medium-run predictions, which is currently unsatisfactory with existing techniques. In addition, the general intuition behind the proposed method readily applies to other non-parametric predictors. Next steps include enhancing predictions via better estimations of the LLE, either by using more neighbors, or neural network methods ([15]). Naturally, our ultimate goal is to evaluate how our method holds up when confronted to real data, and particularly to intra-day financial and economic time series.

References

[1] Shintani M and Linton O. Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos, Journal of Econometrics, 2004; 120: 1-33.

[2] Guégan D and Mercier L. Stochastic and chaotic dynamics in high-frequency financial data. Signal Processing and Prediction, Prochazka Eds, 1998; 365-372.

[3] Farmer JD and Sidorowich JJ. Predicting chaotic time series. Physical Review Letters, 1987; 59: 845 – 848.

[4] Doerner R, Hübinger B, Martienssen W, Grossman S and Thomae S, Predictability Portraits for Chaotic Motions, Chaos, Solitons and Fractals, 1991;1:553-571.

[5] Rinaldi S, Casagrandi R and Gragnani A, Reduced order models for the prediction of the time of occurrence of extreme episodes, Chaos, Solitons and Fractals, 2001;12: 313-320.

[6] Guerra FA and Coelho LS, Multi-step ahead nonlinear identification of Lorenz's chaotic system using radial basis neural network with learning by clustering and particle swarm optimization, Chaos, Solitons and Fractals, 2009; 35:967-979.

[7] Yakowitz S, Nearest-neighbour methods for time series analysis, Journal of Time Series Analysis, 1987: 8, 235-247.

[8] Guégan D. Les Chaos en Finance: Approche Statistique. Economica Eds.; 2003, 420 pp.

[9] Abarbanel HDI. Local and global Lyapunov exponents on a strange attractor. Nonlinear Modeling and Forecasting, SFI Studies in the Science of Complexity, Casdagli M and Eubank S Eds, Addison-Wesley, 1992; Proc. Vol. XII: 229-247.

[10] Wolff RCL. Local Lyapunov exponents: looking closely at chaos. Journal of the Royal Statistical Society B, 1992; 54: 353 – 371.

[11] Takens F. Estimation of dimension and order of time series. Progress in Nonlinear Differential Equations and their Applications, 1996; 19: 405-422.

[12] Bailey B. Local Lyapunov exponents: predictability depends on where you are. Nonlinear Dynamics and Economics, Kirman et al. Eds, 1997.

[13] Lorenz EN. Deterministic non-periodic flow. Journal of Atmospheric Science, 1963; 20: 130-141.

[14] Guégan D and Leroux J. Forecasting chaotic systems: The role of local Lyapunov exponents. forthcoming in Chaos, Solitons and Fractals, 2009.

[15] Gençay R and Liu T, Nonlinear modelling and prediction with feedforward and recurrent networks, Physica D, 1997; 108, 119-134.