Parameter sensitivity of noisy chaotic time series

Elmer S. Hung
Massachusetts Institute of Technology, Rm. 39-665, Cambridge, Massachusetts 02139
(Received 26 September 1996)

We examine the sensitivity of noisy time series of data from a chaotic system to changes in the parameters of the system. The investigation yields two insights: (1) A small fraction of the data contains most of the information about the parameters. (2) For one-parameter families of systems, there is often a preferred direction in parameter space governing how easily trajectories of nearby systems shadow each other. A parameter estimation algorithm is presented that leverages these properties of chaotic systems to yield estimate errors that decrease as $1/N^2$, where N is the number of state samples used. [S1063-651X(97)04507-8]

PACS number(s): 05.45.+b, 05.40.+j

How easy is it to distinguish a time series of noisy state measurements from two slightly different chaotic dynamical systems? It is known that small perturbations in the parameters of a chaotic system can have dramatic effects on the statistical properties of the system [1], and that the long term behavior of slightly different systems can often be easily distinguished, even in the presence of measurement noise [2]. However, for experimental purposes it is also important to know how long a system must be observed for changes in the system parameters to become recognizable and how to go about determining the parameters given a finite series of observations.

This paper investigates these questions by attempting to estimate the parameters of one-parameter families of chaotic systems given a series of noisy full-state observations. This idealized problem turns out to be interesting for a number of reasons. First, the problem is simple enough that one can closely examine the dynamical mechanisms that affect the sensitivity of state trajectories to changes in the system parameter [3]. In addition, parameter estimation techniques are potentially useful in aiding recent efforts to apply chaotic systems for purposes of control and communication [4] and may also have applications for high precision measurement [5], since chaotic systems can be extremely sensitive to small parameter variations.

Here is the problem: Suppose that we are given a one-parameter family of mappings, $f_p(x)$, where x is the state vector of the system, p is a scalar parameter of the system, and f is assumed to vary continuously with x and p. For a specific parameter value $p=p_0$ we are also given a sequence of observations $\{y_n\}_{n=0}^N$ of a certain state trajectory $\{x_n\}_{n=0}^N$, where

$$x_{n+1}=f_{p_0}(x_n)$$

and

$$y_n=x_n+e_n,$$

for $n\in\{0,\ldots,N\}$ where the e_n's represent measurement errors in the data stream. We are interested in estimating the value of p_0 given the data stream $\{y_n\}_{n=0}^N$ when f_{p_0} is chaotic. For analytical purposes we will first assume that the magnitude of the measurement errors is bounded so that for each n, $|e_n|<\epsilon$ for some $\epsilon>0$.

Note that $\{y_n\}_{n=0}^N$ represents part of an actual trajectory of f_{p_0} with ϵ measurement errors added in. Thus, if no trajectory of $f_{p'}$ follows $\{y_n\}_{n=0}^N$ for all $n\in\{0,\ldots,N\}$ within ϵ error for the parameter value $p=p'$, then p' cannot be the actual parameter value of the system that is being observed. In other words, in order to estimate the parameter of f_p given a noisy sequence of data $\{y_n\}_{n=0}^N$, we want to know the values of p for which f_p has at least one trajectory [6], $\{f_p^n(z_0)\}_{n=0}^N$, that *shadows* $\{y_n\}_{n=0}^N$ within ϵ error for some z_0, so that $|f_p^n(z_0)-y_n|<\epsilon$ for $n\in\{0,\ldots,N\}$.

To examine shadowing orbits or trajectories we first need to distinguish between two types of chaotic systems: *uniformly hyperbolic* and *nonuniformly hyperbolic*. Uniformly hyperbolic systems [7] can be locally decomposed into stable and unstable spaces that uniformly contract and expand at an exponential rate. It is known that uniformly hyperbolic systems satisfy a structural stability property [7] so that there exists a one-to-one correspondence between orbits of nearby systems in parameter space. Thus, if f_{p_0} is uniformly hyperbolic, then given any $\epsilon>0$ and any orbit, $\{x_n\}_{n=-\infty}^{\infty}$, of f_{p_0}, there exists an orbit of f_p that shadows $\{x_n\}_{n=-\infty}^{\infty}$ within ϵ for all n for any p sufficiently close to p_0. Consequently, for uniformly hyperbolic systems, there is a limit on the amount of information about the parameters of the system available purely from the dynamics.

However, virtually all physical chaotic systems are not uniformly hyperbolic, so their orbits are not necessarily shadowed by nearby systems in parameter space. However, most orbits of nonuniformly hyperbolic systems still exhibit local hyperbolicity. There exist subspaces that expand or contract exponentially, but the exponential rates of expansion and contraction are not uniformly bounded for every iterate along an infinite orbit. Thus we find that some iterates approach hyperbolic degeneracy conditions.

The work of Grebogi *et al.* [8] demonstrates that away from hyperbolic degeneracies, one can iterate locally hyperbolic finite sections of an orbit with roundoff noise added in and still find close shadowing orbits of the same system. We find, however, that the hyperbolic degeneracies hold the key to performing parameter estimation because they are often accompanied by a local ‘‘folding’’ in state space and a lack of closely shadowing orbits for nearby systems in parameter space. This results in short sections of orbits that are extremely sensitive to small changes in parameters. In particular, we find the following: (1) Most data in a time series of state observations contribute very little information about the parameters. Only the small amount of data taken from stretches of an orbit near a degeneracy really matters. (2) For one-parameter families of systems, the folding action near a hyperbolic degeneracy often results in a preferred direction in parameter space governing how easily orbits of one system shadow orbits of nearby systems in parameter space.

These ideas are most easily illustrated with a one-dimensional example. Consider the family of quadratic maps

$$f_p(x) = px(1-x),$$

for $x \in [0,1]$ and $p \in [0,4]$. We first present numerical experiments demonstrating some rather surprising shadowing results for the quadratic map. We then attempt to explain the results and use them to develop parameter estimation algorithms.

Pick $p_0 = 3.9$ and iterate an orbit, $\{f_{p_0}^n(x_0)\}_{n=0}^N$, of f_{p_0} starting with the initial condition $x_0 = 0.3$. Numerically, the resulting orbit exhibits chaotic properties (e.g., positive Lyapunov exponent). Now consider the question: ‘‘What parameter values p produce orbits that shadow $\{f_{3.9}^n(0.3)\}_{n=0}^N$?’’ We can examine this question numerically by picking values of p close to 3.9 and finding the closest shadowing orbits for each value of p [9].

In order to measure how closely maps with different parameters can shadow the orbit, $\{f_{3.9}^n(0.3)\}_{n=0}^N$, we define $e(N,p)$ to be the maximal distance between that orbit and the closest shadowing orbit of f_p. Let

$$\hat{e}(N,p,p_0,x_0) = \min_{z \in [0,1]} \max_{0 \leq n \leq N} |f_p^n(z) - f_{p_0}^n(x_0)| \qquad (1)$$

and define $e(N,p) = \hat{e}(N,p,3.9,0.3)$.

Figure 1(a) shows the result of computing $e(N,p)$ with respect to $p-3.9$ for three values of N. The three V-shaped traces in the figure represent plots of $e(N,p)$ for $N=61$, $N=250$, and $N=1000$, progressing inward toward the vertical $p=3.9$ axis as N increases. Note the distinct asymmetry of the graph between values of p greater than 3.9 and less than 3.9. Maps with higher parameter values ($p>3.9$) can apparently shadow the original orbit, $\{f_{3.9}^n(0.3)\}_{n=0}^N$, more ‘‘easily’’ than maps with $p<3.9$. Plotting a small part of this graph near the origin and magnifying the horizontal axis [Fig. 1(b)] reveals a sharp transition in the shadowing behavior of maps in parameter space. There is a small interval in parameter space for $p<3.9$ corresponding to maps that shadow the original orbit well; however, as one decreases the parameter value, there is a sudden transition to maps that cannot shadow the original orbit closely at all.

We can examine the situation further by keeping $e(N,p)$ constant at 0.01 and tracking how the V-shaped curves in Fig. 1(a) vary with N. Specifically, define $a(N)>0$ and $b(N)>0$ so that $I(N)=[3.9-a(N),3.9+b(N)]$ is the interval in parameter space such that $e(N,p) \leq 0.01$. A plot of $a(N)$ and $b(N)$ is shown in Fig. 1(c). The plot illustrates two important points: (1) $a(N)$ is several orders of magnitude smaller than $b(N)$ for large N, illustrating the asymmetry in parameter space. (2) $a(N)$ is approximately constant for large stretches of N, but decreases in large increments over a small number of iterates that are especially sensitive to parameters.

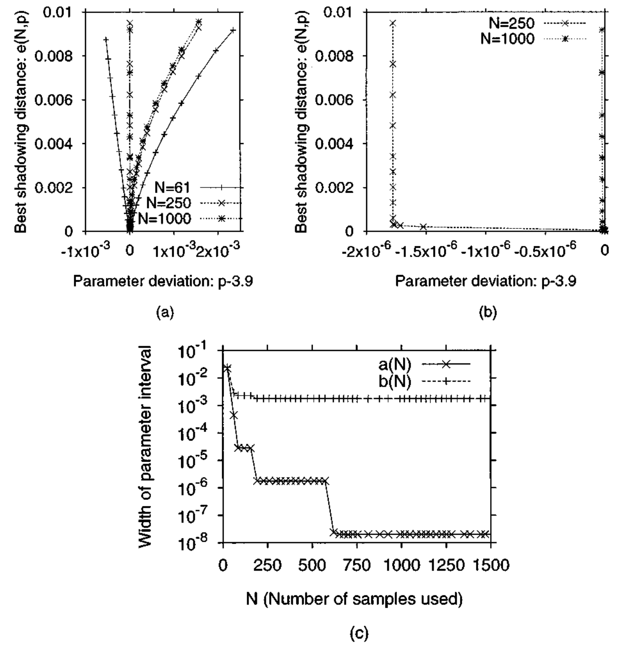

FIG. 1. $e(N,p)$ measures how close the closest shadowing orbit of $f_p = px(1-x)$ comes to shadowing the orbit $\{f_{3.9}^n(0.3)\}_{n=0}^N$. (a) The three V-shaped curves represent $e(N,p)$ for $N=61$, $N=250$, and $N=1000$. (b) The left side of the graph in part (a) for $p<3.9$ is magnified. (c) Graph of $a(N)$ and $b(N)$, where $I(N)=[3.9-a(N),3.9+b(N)]$ is the maximal parameter space interval such that $p \in I(N)$ implies $e(N,p) \leq 0.01$.

Related qualitative results concerning the asymmetry in parameter space have previously been proven analytically [10]. One can also show the following quantitative result [11].

Theorem 1: Let $\hat{e}(N,p,p_0,x_0)$ be as defined in Eq. (1) for the quadratic map $f_p(x) = px(1-x)$ and define

$$\epsilon(p,p_0,x_0) = \lim_{N \to \infty} \hat{e}(N,p,p_0,x_0).$$

For any $\gamma \in (0,1)$, there is a set, $E(\gamma) \subset [0,4]$, of positive Lebesgue measure and constants $\delta>0$, $C>0$, and $K>0$ such that the following holds for any $p_0 \in E(\gamma)$.

(1) f_{p_0} has a positive Lyapunov exponent.

(2) If $x_0 \in [0,1]$, then

$$\epsilon(p,p_0,x_0) < C|p-p_0|^{1/3}$$

for all $p \in (p_0, p_0+\delta)$.

(3) For almost all $x_0 \in [0,1]$,

$$\epsilon(p,p_0,x_0) > K|p-p_0|^{\gamma}$$

for all $p \in (p_0-\delta, p_0)$.

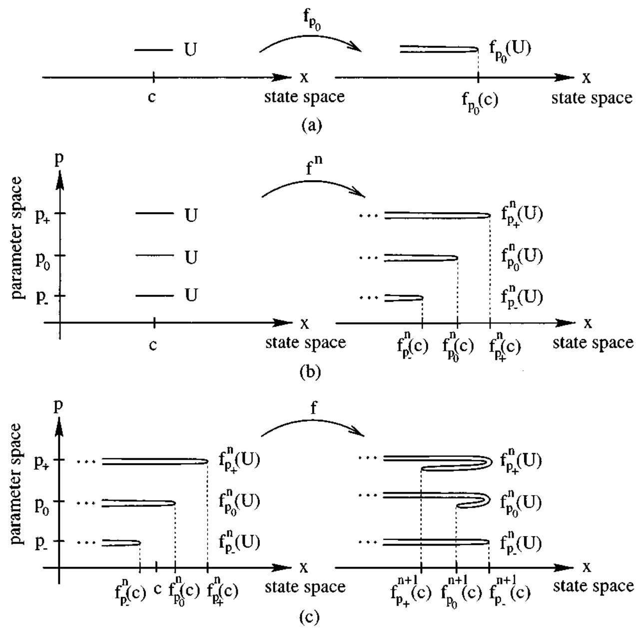

FIG. 2. (a) Depiction of how a state-space neighborhood U, around $c=\frac{1}{2}$ gets folded by the map $f_p = px(1-x)$ for $p=p_0=3.9$. (b) Images of $f_p^n(U)$ get offset from each other exponentially fast for different values of p as n increases. This behavior ultimately leads to asymmetric shadowing in parameter space. (c) When $f_p^n(c)$ returns close to c, offset images are folded back on top of each other, facilitating shadowing.

Thus for the quadratic map there exists a definite parametric asymmetry so that higher parameters are easier to shadow than lower parameters. The proof for the theorem is rather involved; however it is helpful to qualitatively explore the mechanism behind the parametric asymmetry.

For a quadratic map with a positive Lyapunov exponent, hyperbolic degeneracies occur for orbit sections that approach the critical point, $c=\frac{1}{2}$. The sharp dips in the graph of $a(N)$ in Fig. 1(c) occur where the orbit $\{f_{3.9}^n(0.3)\}_{n=0}^N$ approaches the critical point. It turns out that the magnitude of $a(N)$ is governed by how closely the orbit $\{f_{3.9}^n(0.3)\}_{n=0}^N$ comes to c. What happens when an orbit approaches c? Consider a neighborhood U of initial conditions in state space near c. Regions of state space near c are folded on top of each other by $f_{p_0=3.9}$ as shown schematically in Fig. 2(a).

Figure 2(b) illustrates what the images of U look like after n applications of f_p for three closely spaced parameter values, $p=p_- < p_0 < p_+$. The folded images of U get stretched in state space and become offset from each other in parameter space exponentially fast with increasing n. This leads to asymmetrical shadowing in parameter space. For example, since the image, $f_{p_-}^n(U)$, ''lags behind'' $f_{p_0}^n(U)$, there is no orbit of f_{p_-} that closely shadows orbits of f_{p_0} with initial conditions near c for more than n iterates. The lack of shadowing for lower parameter values leads to the dips in the graph of $a(N)$ in Fig. 1(c). On the other hand, from Fig. 2(b), since $f_{p+}^n(U)$ ''leads'' or overlaps $f_{p_0}^n(U)$, we expect that there are no orbits of f_{p_0} which are not shadowed closely by at least one orbit of f_{p_+}.

This is not the end of the story. Figure 2(c) illustrates that if $f_{p_0}^n(c)$ returns close to c, then images that lag behind can catch up as regions get refolded. In this case, we see that the forward image of U under $f_{p_0}^{n+1}$ gets folded back onto the corresponding forward image of U under $f_{p_-}^{n+1}$, thus allowing orbits of f_{p_-} to shadow orbits of f_{p_0}. Numerically, this effect of lower parameter value shadowing is generally too small to be seen (but it complicates the statement of Theorem 1).

We now take advantage of the properties illustrated by Fig. 1 to design an algorithm to perform parameter estimation on large data sets of chaotic time series [11]. Since only a few pieces of the data really matter, we quickly scan through most of the data using a locally linearizing square-root extended Kalman filter [12]. Once we locate pieces of data that might be especially sensitive, we analyze this data using a method based on a Monte Carlo approximation to Bayesian estimation.

For increasing numbers of iterates N we progressively eliminate parameter values corresponding to maps which could not have produced the observed data. This leaves us with a parameter interval $I(N)$ containing the parameters that are consistent with the observed data; however, because of the asymmetry in parameter space, we know that the real parameter value is very close to one of the end points of $I(N)$. We choose the appropriate end point to be our parameter estimate. For example, for the quadratic map we always choose the lower end point of $I(N)$ to be the parameter es-

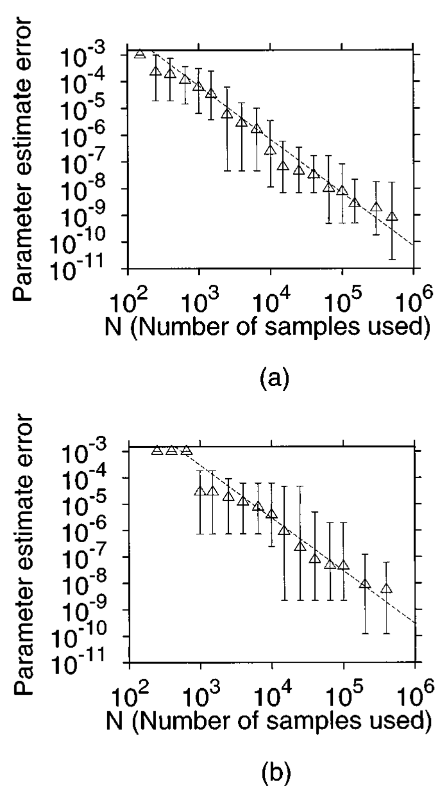

FIG. 3. Results of the parameter estimation algorithm for ten trajectories of (a) the Hénon map and (b) the standard map. Vertical bars delimit the maximum and minimum error results. Triangles indicate log-average errors. The dotted line has a slope of -2 corresponding to a $1/N^2$ convergence rate.

timate so that our estimation error is approximately $a(N)$ in Fig. 1(c). Determining the proper end point to use may be done by testing the algorithm with known parameter values.

Using the estimation algorithm as an analytical tool reveals some interesting results. Below we present results for two different maps, the Hénon map [13]

$$u_{n+1}=u_n+1-av_n^2,$$

$$v_{n+1}=bv_n$$

and the standard map [14]

$$u_{n+1}=(u_n+v_n+k\sin u_n)\bmod 2\pi,$$

$$v_{n+1}=(v_n+k\sin u_n)\bmod 2\pi.$$

Both of these systems are two dimensional where u_n and v_n are the scalar state variables. For the Hénon system, we hold parameter $b=0.3$ fixed and attempt to estimate parameter a ($a=1.4$ is used to generate data). For the standard map, we estimate the parameter k ($k=1$ is used to generate data). The algorithm is run on numerically generated chaotic data: Initial conditions are picked at random, orbits are iterated, and Gaussian white noise with standard deviation $\sigma=0.001$ is added to each iterate.

Both Hénon and standard maps illustrate parametric asymmetry properties similar to those presented in Fig. 1 [11]. In both cases, numerical results show that orbits from maps with lower parameter values tend to be shadowed by orbits from maps with higher parameters but not vice versa.

In Fig. 3 we show the result of applying the parameter estimation procedure to a set of 10 different trajectories. Parameter values are eliminated if there is no trajectory that shadows the data within $\epsilon=8\sigma$. The results, however, are extremely insensitive to ϵ because shadowing properties transition rapidly for parameter values less than the actual parameter value [see, e.g., the sharp slope of $e(N=250,p)$ in Fig. 1(b) for $p-3.9<-1.8\times10^{-6}$].

Note that the estimation accuracy decreases approximately as $1/N^2$ as N gets large. If the accuracy gain was purely statistical, the gain would be at most $1/\sqrt{N}$ [5]. Thus, the dynamics of the systems in question contribute significantly to the improvement in estimation accuracy. The algorithm is also much superior to standard parameter estimation techniques like the extended Kalman filter, which relies on local linearization and has divergence problems near iterates where folding is important. For example, the extended Kalman filter is not able to produce parameter estimates with accuracy better than approximately 10^{-6} on average for the Hénon and standard maps examples shown in Fig. 3 [11].

The author wishes to thank H. Abelson, N. Cohen, K. Koniaris, A. Siapas, G. Sussman, and J. Wisdom, for many useful suggestions. Support for this research was provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract No. N00014-92-J-4097. The author also acknowledges financial support by the Air Force Graduate Fellowship Program.

[1] J. D. Farmer, Phys. Rev. Lett. **55**, 351 (1985).
[2] A. Siapas, Phys. Rev. Lett. **73**, 2184 (1994).
[3] Significant work has been done on more general problems, such as reconstructing the state space of a chaotic time series; see, e.g., M. Casdagli, S. Eubank, J. Farmer, and J. Gibson, Physica D **51**, 52 (1991).
[4] E. Bradley, IEEE Trans. Circuits Syst. **40**, 808 (1993); E. Ott, C. Grebogi, and J. Yorke, Phys. Rev. Lett. **64**, 1196 (1990); K. Cuomo, A. Oppenheim, and S. Strogatz, IEEE Trans. Circuits Syst. **40**, 626 (1993).
[5] J. Wisdom, in *Chaos/Xaoc*, edited by D. K. Campbell (AIP, New York, 1990).
[6] Let $f^{n+1}=f(f^n)$, so that the function f^n refers to the composition of f with itself n times (where f^0 is defined to be the identity function).
[7] R. Bowen, *On Axiom A Diffeomorphisms*, CMBS Regional Conference Series in Mathematics, Vol. 35 (American Mathematical Society, Providence, 1978); M. Shub, *Global Stability of Dynamical Systems* (Springer-Verlag, New York, 1987).
[8] C. Grebogi, S. Hammel, J. Yorke, and T. Sauer, Phys. Rev. Lett. **65**, 1527 (1990).
[9] Despite roundoff errors from iterating long orbits, the example illustrates typical results for the quadratic map, since the important qualitative effects are characterized by local sections of an orbit [11].
[10] H. Nusse and J. Yorke, Commun. Math. Phys. **114**, 363 (1988).
[11] E. Hung, MIT AI-Technical Report No. 1541, May 1995 (unpublished); E. Hung (unpublished).
[12] P. Kaminski *et al.*, IEEE Trans. Autom. Control. **AC-16**, 727 (1971).
[13] M. Hénon, Commun. Math. Phys. **50**, 69 (1976).
[14] B. Chirikov, Phys. Rep. **52**, 263 (1979).