WHY IT IS IMPORTANT TO HIGHLIGHT THE CHAOTIC BEHAVIOUR OF TIME SERIES

CIOBANU DUMITRU
PH.D. STUDENT, UNIVERSITY OF CRAIOVA,
e-mail:ciobanubebedumitru@yahoo.com

Abstract

Chaos Theory says that many apparently random data are, in fact, the results of deterministic process, and try to find the underlying order from data. Determining if a process is or not chaotic may supply valuable information about how to deal with that process. Because there is no single test that completely identifies chaos, is better to perform many tests that provide indices of chaos. I applied more tests on the exchange rate EUR-LEU and I found evidences of chaos.

Keywords: Chaos tests, Time Series, Exchange rate.

JEL Classification: C53, G17.

1. Introduction

Chaotic systems are in fact complex deterministic systems with a large number of variables that influence the evolution of the process making it impossible for humans to simulate it and therefore making them seems unpredictable. This, also, makes it impossible to determine the initial state of the system knowing just the final state.

Most processes and systems that are found in nature involve the interaction of many factors, which allows us to catalog them as chaotic systems. Thus, chaos is met in: solar system dynamics, evolution of populations, the weather, chemical reactions, etc.

The economy can be seen as a chaotic system, a factor that brings a huge number of variables is direct involvement of people. The chaos from complex systems is known as chaos deterministic.

For the chaotic systems, the fact that they are deterministic does not make them predictable. However, the predictive power in the case of chaotic systems can be improved and illustrate this with weather system for which predictions for short periods have reached to a very good accuracy.

We must emphasize the fact that the emergence and development of chaos theory could not have to take place before the invention of computers as simulation of complex systems with many variables could not have done without their help.

An important feature of chaotic systems is Sensitive Dependence on Initial Conditions (SDIC). This tells us that two initially close trajectories depart exponentially in a finite number of iterations, sometimes very quickly. In such a system, prediction is impossible except maybe the prediction for very short periods. The most used tool for identifying these processes from dynamical systems theory or experimental series is Lyapunov characteristic exponent (LCE).

2. Tests of chaos for time series

The working mode for highlighting chaos in time series is embedding them into a multidimensional space.

Transition from one-dimensional time series to the corresponding $d$-dimensional series in state space is done using Takens theorem.

Takens Theorem represent a technique of reconstruction of an approximation for an unknown vector $\{x^d_t\}$ in $d$-dimensional phase space by delaying and embedding the observed time series $\{x_t\}$ [18]. The process of reconstruction suggested by Takens has four steps:

1. Suppose it’s available a time series with $N$ observations: $\{x_1, x_2, \ldots, x_N\}$.
2. Is determined an appropriate time delay $\tau$.
3. Is determined the embedding dimension $d$.
4. Is embedded the 1-dimensional series in $d$-dimensional space by constructing the vectors of length $d$: $X^d_t = (x_t, x_{t+\tau}, \ldots, x_{t+(d-1)\tau}), \ t = 1, 2, \ldots, N - \tau(d - 1)$. 
This approximate reconstruction gives us the state vectors which are composed by time delays of time series \( \{x_t\} \), where \( \tau \) is the number of time delays and \( d \) is the embedding dimension of system. Accurate calculation of \( d \) and \( \tau \), guarantees, according to the Embedding Theorem [1], that the sequential order of reconstructed state vectors \( x^d_t \rightarrow x^d_{t+1} \) is topological equivalent with generation of state vectors \( x_t \rightarrow x_{t+1} \), allowing that \( \{x^d_t\} \) to represent without ambiguity the initial source, the observed time series \( \{x_t\} \).

We embed one-dimensional series in a \( d \)-dimensional space by building vectors of length \( d \) as follows:

\[
x^d_t = (x_t, x_{t+\tau}, \ldots, x_{t+(d-1)\tau}), \quad t = 1, \ldots, N - \tau(d - 1),
\]

where \( \tau \) is the number of time delays.

The accurate determination for time delay \( \tau \) guarantees that the coordinate vectors \( \{x^d_t\} \) from state space are independent to each. If is chosen a too small value for \( \tau \), the data from the state space are clustered. Choosing a too high value for \( \tau \), lead to the disappearance of relations between points and attractor.

The relationship with which is calculated the mutual information function between two coordinates of \( \{x^d_t\} \), for example \( x^d_t \) and \( x^d_{t+\tau} \) is

\[
MI(\tau) = \log_2 \frac{P(x^d_t, x^d_{t+\tau})}{P(x^d_t)P(x^d_{t+\tau})}
\]

where \( P(x^d_t, x^d_{t+\tau}) \) is the joint probability distribution function between \( x^d_t \) and \( x^d_{t+\tau} \).

The average of mutual information for all coordinates is calculated according to formula:

\[
AMI(\tau) = \sum_{x^d_t, x^d_{t+\tau}} P(x^d_t, x^d_{t+\tau}) \log_2 \frac{P(x^d_t, x^d_{t+\tau})}{P(x^d_t)P(x^d_{t+\tau})}.
\]

The signal reconstruction in state space requires a dimension to ensure that there will be no overlap for orbits of dynamic system. The optimum dimension is obtained by calculating the percentage of False Nearest Neighbors (FNN) between points in state space. The number of false nearest neighbors is calculated using the vectors \( x^d_t \) from reconstructed state space of different embedding dimensions \( d \), but with a constant number of delays \( \tau \). Is generally accepted the idea that the minimum embedding dimension for original state space is reached when the percentage of false nearest neighbors drops to zero and this value guaranteeing that the orbits are unique. Calculation of false nearest neighbors requires measuring the distance \( R_d \), defined as the distance between neighboring vectors in consecutive dimensions.

The Euclidean distance representing \( R_d \) in dimension \( d \) is:

\[
R_d^2(t) = \sum_{m=1}^{d} (x^d_{t+(m-1)\tau} - x^d_{t+(m-1)})^2
\]

where \( t \) is the current index of discrete signal (\( x_t \)) and \( x^d_{t\tau} \) is the nearest neighbor (NN) for \( x_t \).

For \((d+1)\) dimension, the Euclidean distance is:

\[
R_{d+1}^2(t) = \sum_{m=1}^{d+1} (x^d_{t+(m-1)\tau} - x^d_{t+(m-1)})^2 = R_d^2(t) + (x^d_{t+\tau m} - x^d_{t+\tau m+1})^2.
\]

It is considered the criterion according to which a false nearest neighbor is any neighbor for which the change of distance between points in dimension \( d \) and dimension \((d+1)\) exceeds a heuristic threshold \( R_{tol} \).

\[
\sqrt{\frac{R_{d+1}^2(t) - R_d^2(t)}{R_d^2(t)}} = \frac{\|x^d_{t+\tau m} - x^d_{t+\tau m+1}\|}{R_d(t)} > R_{tol}.
\]

Determination of false nearest neighbors depends on how is changed the distance between vectors in state space in consecutive dimensions. If the distance increases significantly with embedding dimension increment, the vectors are false neighbors; this means that points, apparently close because the projection, are separate in larger embedding dimensions. If the gap remains below a certain threshold, the points in state space are real neighbors resulting from the dynamics of the system. Embedding dimension that accurately representing the system is that who eliminates most false neighbors resulting in a system whose state space trajectories are positioned according to system dynamics and not due to space reconstruction [7].
To understand how to choose the best dimension $d$ for embedding is helpful to understand what is happening geometrically. As the size increases, the chaotic attractor unfolds. When the unfolding is complete, a sequential path from one point to another will not intersect herself. If the embedding dimension, $d$, is too small, some paths of the projected attractor will intersect herself.

The method of false nearest neighbors recognizes that where paths of attractor intersect herself, two neighboring points are distant in real embedding space of data series [7]. Using this idea, Cao proposed a method for determining a good embedding dimension [4].

Let

$$E_1(d) = \frac{E(d+1)}{E(d)}$$

with

$$E(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} \left\| x_i^{d+1} - x_i^{d+\text{NN}} \right\|$$

and

$$\left\| x_i^d - x_i^{d,\text{NN}} \right\| = \max_{0 \leq m \leq d-1} \left| x_{i+m\tau} - x_{i+m\tau}^{\text{NN}} \right|,$$

where $N$ is the length of original series of data, $d$ represent the embedding dimension and superscript $\text{NN}$ identify the nearest neighbor of vector. As $d$ increasing, $E_1(d)$ tends to one. The embedding dimension will be given by the value of $d$ for witch $E_1(d)$ stops to modify.

Cao, Mees & Judd [4] have proposed a similar method that verifies if the original time series is random, using a different metric

$$E_2(d) = \frac{E^*(d+1)}{E^*(d)}$$

where

$$E^*(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} \left| x_{i+m\tau} - x_{i+m\tau}^{\text{NN}} \right|.$$  

A random time series will have $E_2(d)$ close to unity for all values of $d$ while a chaotic series will have $E_2(d)$ less than one for small values of $d$.

There are many ways to define a fractal dimension but the most used is correlation dimension.

Let $N$ be the number of elements of time series, $d$ the embedding dimension and $\tau$ the time delay. The embedding of time series of observations in a $d$-dimensional space is achieved by constructing the vectors

$$x_t^d = (x_t, x_{t+\tau}, \ldots, x_{t+(d-1)\tau}), \quad t = 1, 2, \ldots, M,$$

where $M = N - \tau(d-1)$.

Considering spheres of radius $r$ around points of embedding space, the average number of points contained in spheres, not counting their centers, is given by relation

$$C(r) = \frac{1}{M(M-1)} \sum_{t=1}^{M} \sum_{s=1}^{M} H(r - \left\| x_t^d - x_s^d \right\|),$$

where $x_t^d$ is the centre of sphere and $H(x)$ is the Heaviside function,

$$H(x) = \begin{cases} 0 & \text{daca } x < 0 \\ 1 & \text{daca } x \geq 0 \end{cases}.$$

The correlation dimension assume that as $r$ approaches to zero, the relationship according to witch $C(r)$ is changing is

$$C(r) = \lim_{r \to 0} kr^{D_C}.$$  

Extracting $D_C$ from previous relation, we obtain

$$D_C = \lim_{r \to 0} \frac{\ln C(r)}{\ln r}.$$  

Because the data set is not continuous, $r$ cannot get too close to zero otherwise the spheres would not contain other points except their centers. In practice, to remove this shortcoming is plotted $\ln C(r)$ versus $\ln r$ and is
identified the apparently linear portion of the graph. The slope of this portion approximates the correlation dimension $D_c$. If $D_c$ is an integer, then the attractor is a usual geometric object, a point for $D_c = 0$, a curve if $D_c = 1$ or a surface when $D_c = 2$. If $D_c$ is not an integer, then the attractor is strange and the system has chaotic behavior [7].

3. Case study the time series of exchange rate EUR-LEU

I used the time series of exchange rate EUR-LEU. Historical values of foreign exchange are available for download on the website of National Bank at http://www.bnr.ro/Baza-de-date-interactiva-604.aspx. For time series modeling and simulations I used MATLAB (R2011a) and tstool Toolbox (time series tools).

![Fig.1. The evolution of exchange rate EUR-LEU over time.](image)

The considered time series contains 3881 records during 04.01.1999 - 25.09.2012 and consists of exchange rate quotations of EUR-LEU established by National Bank of Romania for weekdays.

From the graphical representation (Fig.1.) we can see that in the case of the exchange rate EUR-LEU we deal with strongly nonlinear process. Nonlinearity does not necessarily imply chaos but any chaotic process is nonlinear.
Determining a suitable time delay is achieved by building automutuale information function and finding his first local minimum. I have conducted several simulations for different values of the embedding dimension and obtained values between 12 and 21. The most common value for the delay time was $\tau=19$ (Fig. 2.).

Simulations have indicated, in the case of using the false nearest neighbor method, a minimum embedding dimension between 5 and 7 (Fig. 3.).
Fig. 4. Minimum embedding dimension obtained with Cao method.

Likewise, when I used Cao’s method to determine the embedding dimension I obtained values between 5 and 7 (Fig. 4.). I decided that minimum embedding dimension is 6, the value most often indicated by tests.

I used the two values, time delay $\tau=19$ and minimum embedding dimension $d=6$, to construct the multidimensional series.

Fig. 5. The projection of the time series embedded in a 6-dimensional phase space on a 3-dimensional space.
Fig. 6. The projection of the time series embedded in a 6-dimensional phase space on a 2-dimensional space.

Fig. 7. The determination of correlation dimension.

The linear portion from graphical representation (Fig. 7.) has a slope of about 2.838, noninteger value that indicates chaos. The Takens estimator for the correlation dimension was 2.7708, results obtained with an instrument from the tstool toolbox (Fig. 8.).

Fig. 8. The correlation dimension approximate by the Takens estimator.
Fig. 9. The evolution of the prediction error.

Fig. 10. The largest Lyapunov exponent versus time delay.
The graphic representations (Fig.10.) and (Fig.11.) show the evolution of the largest Lyapunov exponent depending on the delay time respectively embedding dimension space.

Positive values for the largest Lyapunov exponent indicate chaotic nature of the time series of the exchange rate EUR-LEU.

4. Conclusions

In economy, the majority of historical data are available as time series. Detecting chaotic nature of the processes that have provided such data is not an easy task because there is still no way to specify clearly the existence of chaos. Another constraint is the relatively small number of observations that allows us only to issue certain assumptions about the phenomenon studied and to determine estimates of chaos indicators such as largest Lyapunov exponent.

Thus in this uncertainty we can only try to highlight as many aspects that allow us cataloging process as chaotic one.

Due to these weaknesses and others such as difficulty distinguishing between deterministic chaos and noise and limited predictions to just a few steps, economists have lost the enthusiasm displayed upon discovery of chaos theory.

However, there are ideals such as guiding the economy with small impulses applied at appropriate times, to which tend theorists in economics and that could be possible using models based on chaos theory.

In the case of the time series of the exchange rate between the euro and leu simulations indicate the presence of chaos. With a largest Lyapunov exponent of about 0.2, theoretically acceptable predictions are possible for a number of 4-5 steps. Thus, there remains open the problem of determining the model that simulates reasonably well the time series of the exchange rate so that predictions for first steps to be within acceptable error margin.

Determination of chaotic behavior is important from this regard to establish a correct prediction horizon.

5. Bibliography


