Dynamic parameters for Brazilian financial time series

Gerson Francisco[§]
Cláudio Paiva[¤]
Rogério Rosenfeld[‡]

RESUMO

Neste estudo analisamos o IBOVESPA, índice da Bolsa de Valores de S. Paulo, usando técnicas da teoria de sistemas dinâmicos e processos estocásticos. Discutimos o expoente de Lyapunov, a dimensão de correlação, a complexidade de Lempel-Ziv, o expoente de Hurst e a estatística BDS. Comparamos este estudo com outras séries temporais incluindo preços de ações e sistemas determinísticos. Concluímos que o IBOVESPA é um processo estocástico linear que exibe o fenômeno de persistência, isto é, possui memória de longo prazo. As ações são descritas por processos estocásticos não-lineares tornando difícil sua simulação com modelos determinísticos, tais como as arquiteturas usuais de redes neurais.

Palavras-chave: determinístico, estocástico, reconstruções no espaço de fase , complexidade.

ABSTRACT

In this study we analyse a Brazilian stock index called IBOVESPA using techniques from dynamical systems theory and stochastic processes. We discuss the Lyapunov exponent, the correlation dimension, the Lempel-Ziv complexity, the Hurst exponent and the BDS statistics. We compare this study with other time series including stock prices and deterministic systems. We conclude that the IBOVESPA is a linear stochastic process that exhibits the phenomenon of persistence, that is, it has long term memory. The stocks are described by nonlinear stochastic processes making it impossible to be simulated with deterministic models such as the usual neural networks architectures.

Key words: deterministic, stochastic, reconstruction in phase space, complexity.

JEL classification: G100, G140.

---

§ Universidade Estadual Paulista - Instituto de Física Teórica. gerson@ift.unesp.br.

¤ Analitix Soluções em Finanças Ltda. claudio.paiva@analitix.com.br.

‡ Universidade Estadual Paulista Instituto de Física Teórica. rosenfel@ift.unesp.br.

Recebido em junho de 2000. Aceito em janeiro de 2002.

1 Introduction

Dynamical systems theory and stochastic processes are the two prominent tools used to investigate the behaviour of complex time evolution in the physical sciences. These methods are being increasingly employed in systems obtained from economic and financial time series. One of the main points in such studies is to determine the type of evolution mechanism that generated the motion. Such an understanding is important in order to determine the most adequate set of tools in the modelling, and eventual prediction, of the series. (Weigend and Gershenfeld, 1994) In many problems of interest the explicit mechanism that generated the motion, whether deterministic or stochastic, is unknown. This situation is common in the financial and economic systems together with the fact that they lack sufficient amounts of data for the full use of some of the tools proposed herein. In spite of these shortcomings the methods here analysed can be used to obtain valuable insights into the problem when proper care is exercised.

Here we adopt an exploratory approach to the problem of understanding and classifying the behaviour of a financial time series describing the IBOVESPA, the main Brazilian stock market index. This index, together with other Brazilian stocks, were analysed in Denisard, Brundo and Francisco (2000) using the BDS statistics. It was found that although some form of nonlinearity could be detected for the most important stock indices, no trace of it was present in the IBOVESPA. In the present study we analyse the IBOVESPA further using some diagnostic parameters that will help to determine the nature of its generating mechanism. We compare these results with that of other systems, as for example chaotic models, and conclude that fundamental differences exist between the several modes of evolution.

This paper is organised as follows. The Lyapunov exponent is presented in §2 in the context of dynamical systems theory and it is stressed that this parameter is not adequate to distinguish between chaoticity and stochasticity, although it is an important measure of unpredictability in deterministic systems. Next we discuss the limitations of the correlation sum in §3, together with the reconstruction of the phase space. We also introduce the concept of correlation dimension and discuss its bearing on the classification of stochastic and deterministic behaviour. In order to substantiate the conclusion of the previous paragraph, we analyse the system using another parameter, the Lempel-Ziv complexity in §4, which endorses the random character of the index. The Hurst exponent is presented in §5 and its use in determining the persistence, or anti persistence, is discussed. We choose R/S analysis as the method for computing the Hurst exponent although other alternatives exist. Finally, the BDS statistics is briefly mentioned in §6. All parameters discussed are calculated for the IBOVESPA and their values compared with those of the well known Lorenz attractor and with a selection of stocks. We conclude with some remarks on the results obtained.

2 Unpredictable deterministic systems

The evolution of systems which exhibit complex behaviour are described either by deterministic or stochastic propagation in time. Amongst the deterministic modes of evolution those systems known as chaotic will be the focus of our study. The trajectories are generated by a well defined mechanism, such as differential equations or some nonlinear mapping. This is completely different from a stochastic process where trajectories are generated by random variables defined a priori. For example, the geometric Brownian motion, which models the price trajectories for the calculation of contingent claims in the Black-Scholes theory, are built from the start with gaussian random processes and independent increments. Chaotic systems on their turn are described by trajectories highly sensitive to minute fluctuations in the specification of their initial conditions. This unstable behaviour will ultimately lead to unpredictable dynamics since such fluctuations are unavoidable in practice and also in computer simulations. More specifically, trajectories of chaotic systems diverge exponentially in phase space at a rate known as the Lyapunov exponent. A positive exponent indicates that a deterministic system will be unpredictable, that is, the time horizon for which predictions can be made is restricted to a certain characteristic interval. Beyond this horizon only statistical estimates can be made and in this sense there is some resemblance with stochastic process, although a chaotic system has an attractor with well defined geometry and dimension, features never found in stochastic processes. The attractor is a confined region of phase space where trajectories will be trapped. A chaotic attractor exhibits both confinement and local divergence of orbits. The coexistence of such apparently paradoxical behaviour is possible because families of orbits bend on each other in such a way as to satisfy these requirements and resulting in the fractal dimensionality observed in many attractors.

Rigorous treatments of Lyapunov exponents in terms of eigenvalues and eigenvectors of infinite products of matrices can be found in the literature. (Mañe, 1983; Eckmann and Ruelle, 1985) In general one will find up to n distinct exponents in n-dimensional spaces and the important task in applications is to find the maximal exponent. The numerical implementation contained in Wolf *et al.* (1985) can be used with the proviso that this method **assumes** that the system is deterministic and it will not in general provide correct results for stochastic processes. A code which circumvents the difficulties with this algorithm is presented in Kantz and Schreiber (1999). For simulation purposes the numerical orbit will always align during evolution in a direction which produces the largest exponent. If this exponent is positive the system is called chaotic, if not we have determinism without chaos and this possibility is uninteresting in our study. A one-dimensional example which shows clearly how to estimate the time horizon for predictability using the Lyapunov exponent is given in Shaw (1980). For higher dimensional systems the sum of all positive exponents needs to be used. (Schuster, 1988)

3 Phase space reconstruction and correlation dimension

The main procedure used in this section is a method that allows the construction of phase space trajectories of a system using just one of its observable component. This approach is useful even when the evolution mechanism is unknown. For example, the price of an asset may be described by an unknown dynamics in terms of variables not immediately assessed or identifiable. Under some conditions, using the method discussed herein, one can determine the number of variables that generated the motion by considering just the price series or the series corresponding to a single component of the system. When trajectories of dynamical systems are evolved to the future, they tend to accumulate near attractors. The method also provides, in some cases, a way to obtain the fractal dimension of the attractor. (Mañe, 1981; Takens, 1981) Two main limitations for the application of the method occur when there is excessive presence of noise and when large amounts of data is not available. We will not discuss here how to filter the data but refer to Kostelich and Screiber (1993) and Davies (1997). A critical issue is the amount of data required for meaningful determination of parameters and an inequality giving an upper bound for the dimension calculation will be discussed.

Observations on n-dimensional trajectories constitute a sequence of numbers defined on points on these orbits. Suppose that the only information about the system is a series of data values $a(t)$ representing some measured empirical observation which may correspond to indices, price returns, exchange rates, etc. The time index depend on the problem at hand and can be minute, day, week, etc. Choose a time delay[1] T and consider an N-dimensional model of phase-space with trajectories defined by points $a^N(t) = (a(t+T), a(t+2T),..., a(t+NT))$. Increasing t implies in the evolution of the N-dimensional orbit $a^N(t)$ and this provides an image of the attractor. The crucial parameter to be determined is the minimum dimension n, called embedding dimension, of the space where this process occurs. The method for the determination of n from the N-dimensional space requires the consideration of the correlation sum $C(r)$. The function $C(r)$ is defined as the normalised average number of pairs of points inside N-dimensional spheres of radius r along the trajectory of the system. (Grassberger and Procaccia, 1983; Abarbanel *et al.*, 1993). The main point here is to identify a scaling region in the logarithmic plots of r by $C(r)$ for several values of N and r where an unambiguous slope is clearly visible (see Kantz and Schreiber, 1999). At each increment of N we note that the slope of the linear part of the plot tends to increase until a certain maximal stabilisation value is reached.

1 There are well defined methods to choose this delay (see Abarbanel). For example, T could be chosen as the time necessary for the autocorrelation to decay to $1/e$ of its initial value. In our case, since data are scarce, it is a common procedure to choose the delay as one unit of time.

The first N for which the slope saturates and stops increasing is the embedding dimension n of the system, and the corresponding slope is the fractal dimension of the attractor, in this case called the correlation dimension. Another approach to determine the embedding dimension is the notion of false nearest neighbours as discussed in Kennel, Brown and Abarbanel (1992). A distinction between deterministic chaotic evolution and stochastic processes is that the embedding dimension for stochastic evolution always increases and this motion will never be confined to lower dimensional spaces.

In this discussion, and in what follows, the delay is taken to be T=1, although the method of mutual information suggested in Abarbanel *et al.* (1993) should have been used. We do not have enough data for discarding points and our conclusions will be corroborated by the use of the Lempel-Ziv complexity which does not require embeddings. The Lorenz attractor is a prototype for chaotic deterministic behaviour and its computed maximal Lyapunov exponent is λ=2.16.[2] The same method applied to the IBOVESPA logarithmic returns between May 2, 1994 and July 27, 1999, a series comprised of 1292 valid records, would also give a positive exponent λ=0,63. Two remarks need to be made concerning the value of this exponent. Firstly, it should be stressed that we have considered the exponent over the whole period. However the Brazilian stock market can have large local variations, sometimes in a few days, and this is not apparent in a long-term average of 1292 records. For practical uses of this concept we recommend the calculation of the local Lyapunov exponent, for example computed weekly, a technique called short time average which has successfully been applied in fluid turbulence (Tavakol and Tworkowski, 1988) and chaotic cosmological models. (Burd, Buric and Ellis, 1990) Secondly, the algorithm by Wolf et al cannot be recommended as the sole criterion for chaoticity since it does not provide any indication that a deterministic mechanism is operating in the dynamics.[3] In other words, although the index system has a positive Lyapunov exponent we cannot infer it is chaotic based on just this information since chaos is a manifestation of a deterministic process. More has to be explored to arrive at a classification of deterministic or nondeterministic system. In the easier case of the Lorenz attractor its correlation dimension is 2.03 with an embedding dimension 3 because increasing N beyond this value the correlation dimension does not alter significantly from that value. If nothing else was known about the

2 See Wolf *et al.* (1985) for the constant parameters used for the Lorenz attractor. There are other possible choices, giving different exponents.

3 The algorithm presented in Wolf *et al.* (1985) assumes determinism, but this is what we are trying to assess (see Kantz and Schreiber, 1999).

Lorenz system these results strongly suggest that its generating mechanism is deterministic. As regards the financial index, if we successively increase N from 2 to 5 the correlation function increases, respectively, as 2.06, 3.06, 3.76 and 4.46. This is an indication that there is no saturation of the slope and no definite fractal dimension can be inferred. This is an indication that the index is not deterministic. Further insights will be provided by the Lempel-Ziv complexity in the next section; below we present additional discussion on the determination of the correlation dimension.

A critical issue in all analyses of an unknown dynamics is that the number of data points may impose severe restrictions on the upper limit for the values of N we can use. For example in this case we can not go beyond 5. An inequality (Eckmann and Ruelle, 1992; see also the discussion in Stefanovska, Strle and Kroselj, 1997) sets an upper limit on N beyond which no conclusions can be drawn from a time series. The fractal dimension is related to the number of points M and the relative size ρ of the data (the range for which logarithmic graph of $C(r)$ is approximately a straight line divided by the spread of the data in phase-space). The inequality is given by $D_{max} \leq 2 (\log M) / \log (1/\rho)$. In the case of IBOVESPA, with $\rho \approx 1/10$, we have that $D_{max} \leq 6$. Any calculation which violates this bound is completely meaningless, and even 6 dimensional embeddings should be avoided. This is the reason why we have limited our calculations to 5. Actually, with limited amounts of data it is impossible to distinguish between high dimensional chaos and stochastic processes. In spite of this, within the limited amount of data at our disposal, the index seems to behave as a stochastic process. This phenomenon also happens when we compute the correlation dimension of the stocks of Vale do Rio Doce (mining company) and Telebrás (telephone company). In conclusion, as far as our application of this criterion is concerned, with the data available, the stocks and the index analysed appear to have been generated by a stochastic process.

4 Lempel-Ziv complexity

Our analysis suggested that the index IBOVESPA and other stocks can be considered as generated by a stochastic process and not by a deterministic mechanism. Since this conclusion cannot be considered definitive, we will include another measure of randomness in this discussion that provides further confirmation of that result. In this context one does not need to reconstruct phase-space but interpret the data $a(t)$ as a signal generated by some kind of source. This idea is ever present in communication theory where one wishes to determine the minimum alphabet required to code a source whose signal is to be sent through a noisy channel. Let us consider the length $L(N)$ of the minimal program that reproduces a sequence with N symbols. The Lempel-Ziv algorithm is constructed by a special parsing which splits the sequence into words of the shortest length and that has not appeared previously. For example, the sequence 0011101001011011 is parsed as 0.01.1.10.100.101.1011. One can show that $L(N) \approx N_w(N) \log(N_w(N)+1)$ where $N_w(N)$ is the number of distinct words in a parsing and N the size of the sequence. From this one can see that $L(N)$ contains a measure of randomness where a source that produces a greater number of new words is more random than a source producing a more repetitive pattern. In analogy with dynamical evolution, those systems that are composed of well defined cycles is predictable while chaotic motion and stochastic processes are always producing new kinds of trajectories that never repeat themselves. The Lempel-Ziv complexity is defined as $C = \limsup_{N\to\infty} L(N)/N$. (Lempel and Ziv, 1976; Badii and Politi, 1997) A comparison between chaos and stochasticity can now be obtained. In the former case the Lempel-Ziv complexity is well below 1 while in the latter it is close to one. More specifically, if we consider an oscillatory system such as the well known van der Pol oscillator, then C=0.049. For the Lorenz attractor, C=0.181. The complexity for the index IBOVESPA is approximately C=1.06 thus confirming a higher level of randomness than in a chaotic or oscillatory regime. The stocks from Vale do Rio Doce and Telebrás also present a high value for the complexity providing further evidence of their random character.

5 R/S Analysis and Hurst exponent

Here we discuss an important parameter that can be used for any system, chaotic or stochastic, gaussian or non gaussian. The Hurst exponent provides a measure of long term memory structures for series of data which can be either empirical or explicitly generated by some mechanism. In addition to randomness a given series might possess the tendency to maintain or to revert its previous behaviour. We will quantify this idea using the methods of R/S analysis developed by Hurst. (Mandelbrot, 1988; Moody and Wu, 1996; Peters, 1994) Given a series of data $a(t)$ we first compute its average on an interval of length N from t_0, that is, $m(N,t_0) = N^{-1}\sum_{t=t_0+1}^{t_0+N} a(t)$. The deviation from the mean in an interval of size τ is just $X(N,t_0,\tau) = \sum_{t=t_0+1}^{t_0+\tau}(a(t)-m(N,t_0))$. Then one calculates the maximum and the minimum deviation for $1 \le \tau \le N$ and defines the range as the difference between them: $R(N,t_0) = Max_\tau X(N,t_0,\tau) - Min_\tau X(N,t_0,\tau)$. In order to compare phenomena with widely different scales Hurst reescaled this range using the standard deviation $S(N,t_0)$, obtaining the reescaled range $RS(N,t_0) = R(N,t_0)/S(N,t_0)$. By calculating the reescaled range for contiguous intervals of size N, for example, $(t_0+1, t_0+N); (t_0+N+1, t_0+2N)$, etc. one obtains, for each N, the average $RS(N)$. If a process is such that a scaling law

$$RS(N) = const \times e^{HN}$$

can be found then we call it a Hurst process with Hurst exponent H. One can show that $0 < H < 1$, and three important cases arise. For gaussian processes, when H=1/2 the system is just the Brownian motion in continuous time and there are no correlation between past and future. When $H \neq 1/2$ there is long term correlation and the evolution falls into two categories. For H>1/2, the process is called persistent and there is a propensity to maintain past trends, that is, if the motion was increasing/decreasing it keeps increasing/decreasing. When H<1/2, the motion is antipersistent and there is propensity to invert past trend resulting in a motion with larger local variations than for persistent processes. In fact, when a Hurst process is plotted against time its fractal dimension[4] is 2-H, showing that an antipersistent motion has a larger fractal dimension. A Brownian motion has fractal dimension 3/2.

We used R/S analysis to compute the Hurst exponent for the index IBOVESPA. The result we obtained was approximately 0.615, showing that there is memory of past moves and this differs from models in which motion is uncorrelated. When applying these procedures to the Lorenz attractor one obtains $H \approx 0.707$ and in this case R/S analysis indicates a slight periodic tendency in this system. (De Grauwe, Dewachter and Embrechts, 1993) Persistence and antipersistence are phenomena that do not distinguish between determinism and stochasticity but they constitute an important diagnostic parameter leading to a better understanding of empirical data.

6 BDS statistics

The basis for this method is the correlation sum described above where it is used as a test statistic on a time series data. Formally, the BDS test has as its null that the data were generated by a iid process, that is, by an independent and identically distributed stochastic process. (Brock, Dechert and Scheinkman, 1987; Brock, Dechert, Scheinkman and LeBaron, 1996) The BDS test does not specify an alternative hypothesis but a number of Monte Carlo experiments have been performed against a variety of alternatives. (Brock, Hsieh and LeBaron, 1991) In our study we used the BDS statistic as a tool for the specification of models. After adjusting a linear model to the data we applied the BDS statistic on the residue. This procedure was implemented in Denisard, Brundo and Francisco (1999) using data from IBOVESPA,

4 See Mandelbrot (1988) for a discussion of the various notions of fractal dimensions.

Vale do Rio Doce and Telebrás with the following results. The null hypothesis was accepted for the index and rejected for the stocks Vale and Telebrás. This means, as far as these studies are concerned, that the index is best described by a linear process while the stocks contain nonlinearities that might be worthwhile to model.

7 Conclusions

Our analysis of the index IBOVESPA has shown that its behaviour can be considered as that of a persistent linear stochastic processes not generated by a deterministic mechanism. As a consequence, if this index is used as the underlying for pricing derivative securities, then the most adequate model for the trajectories is not the geometric Brownian motion because such process is not persistent. In modelling the index one should avoid using traditional neural networks paradigms or ARIMA in favour of its generalised fractal counterpart ARFIMA. In the case of stocks, the use of nonlinear stochastic process lead to well known difficulties in the choice of a suitable model. An alternative is to implement a neural network as a probability density estimator. (Weigend and Gershenfeld, 1994; Husmeier, 1999) We remark that some codes calculate a positive value for the maximal Lyapunov exponent of stochastic processes but this result can not be used to infer determinism. In our discussion the embedding dimension and the Lempel-Ziv complexity provided the main tools for the characterisation between deterministic and stochastic modes of evolution.

Bibliography

Abarbanel, H. D. I., Brown, R., Sidorowich, J. J.; Tsimring, L. Sh. The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.* 65, 1331, 1993.

Badii, R., Politi, A. *Complexity.* Cambridge University Press, 1997.

Brock, W. A., Dechert, W., Scheinkman, J. A test for independence based on the correlation dimension. *SSRI, Working Papers* n. 8702, Economics, University of Winsconsin-Madison, 1987.

Brock, W. A., Dechert, W., Scheinkman, J.; LeBaron, B. A test for independence based on the correlation dimension. *Econometric Reviews*, 15, p. 197-235, 1996.

Brock, W. A.; Hsieh, D. A.; LeBaron, B. *Nonlinear dynamics, chaos and instability: statistical theory and economic evidence.* MIT Press, 1991.

Burd, A. B., Buric, N.; Ellis, G. F. A numerical analysis of chaotic behaviour in Bianchi IX models. *Gen. Rel. Grav.* 22, p. 349-363, 1990.

Davies, M. E. Reconstructing attractors from filtered time series. *Physica D*101, 195, 1997.

Denisard, A., Brundo, M., Francisco, G. Análise não-linear de séries financeiras. Preprint, IFT/ 2000. Submitted for publication.

De Grauwe, P., Dewachter, H., Embrechts, M. *Exchange rate theory: chaotic models of foreign exchange markets.* Oxford: Blackwell Publishers, 1993.

Eckmann, J.-P., Ruelle, D. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* 57, 617, 1985.

_______ Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. *Physica D* 56, 185, 1992.

Kennel, M. B., Brown, R., Abarbanel, H. D. I. Determining embedding dimension for phase-space reconstruction using a geometric construction. *Phys. Rev. A*45, 3403, 1992.

Grassberger, P.; Procaccia, I. Characterization of strange attractors. *Phys. Rev. Lett.* 50, 346, 1983.

Husmeier, D. *Neural networks for conditional probability estimation: forecasting beyond point predictions.* Springer, 1999.

Kantz, K.; Schreiber, T. *Nonlinear time series analysis.* Cambridge University Press, 1999.

Kostelich, E. J., Schreiber, T. Noise reduction in chaotic time-series data: a survey of common methods. *Phys. Rev. E*48, 1752, 1993.

Lempel, A., Ziv, J. On the complexity of finite sequences. *IEEE Trans. Inform. Theory* 22, 75, 1993.

Mandelbrot, B. *The fractal geometry of nature.* W. H. Feeman, 1988.

Mañe, R. On the dimension of the compact invariant sets of certain nonlinear maps. *In*: *Dynamical systems and turbulence.* Warwick 1980, Lecture Notes in Mathematics, Springer-Verlag, 1981.

Mañe, R. *Teoria ergódica.* English translation by Springer-Verlag, Projeto Euclides, 1983.

Moody, J., Wu, L.-Z. Improved estimates for the reescaled range analysis. *In*: Refenes A.-P. N.; Abu-Mostafa, Y., Moody; J., Weigend, A. (eds.), *Neural networks in financial engineering.* World Scientific, 1996.

Peters, E. *Fractal market analysis.* Wiley, 1994.

Schuster, H. G. *Deterministic chaos.* VCH, 1988.

Shaw, R. Strange attractors, chaotic behaviour, and the information flow. *Z. Naturforsch.* 36a, 80, 1981.

Stefanovska, A., Strle, S., Kroselj, P. On the overestimation of the correlation dimension. *Phys. Lett.* A235, 24, 1997

Takens, F Detecting strange attractors in turbulence. *In*: *Dynamical systems and turbulence.* Warwick 1980, Lecture Notes in Mathematics, Springer-Verlag, 1981.

Tavakol, R. K., Tworkowski, A. S. Fuid intermittency in low dimensional deterministic systems. *Phys. Lett.* A126, p. 318-324, 1988.

Weigend, A. S., Gershenfeld, N. A. (eds.). *Time series prediction: forecasting the future and understanding the past.* Addison Wesley, 1993.

Wolf, A., Swift, J. B.; Swinney, H. L.; Vastano, J. A. Determining the Lyapunov exponent from a time series. *Physica* 16D, 285, 1985.