Statistical Analysis of Dynamic Systems

Analisi Statistica dei Sistemi Dinamici

Gianluca Cubadda
Università degli Studi del Molise
gcubadda@unimol.it

Cira Perna
Università degli Studi di Salerno
perna@unisa.it

Riassunto: Questo lavoro rivisita alcuni recenti sviluppi dell'analisi statistica dei modelli dinamici, con particolare riferimento ai modelli stocastici univariati di serie storiche non-lineari e al ruolo delle caratteristiche comuni nei modelli multivariati di serie storiche non stazionarie.

Keywords: Dynamic Systems, Non linear time series models, Multivariate time series analysis.

1. Introduction

Visual inspection of time series suggests that the usual assumptions of linearity and stationarity may often be violated in practice. Indeed, macroeconomic time series usually display strong trending behaviors, financial time series are characterized by unstable second moments, asymmetry is often present in ecological time series, and so on. Modeling such dynamic characteristics represents an hard challenge for modern time series analysis.

The purpose of this paper is twofold. First, we review some of the most important theoretical and practical issues arising in the statistical analysis of univariate non linear time series, focusing on the connections between this approach and that based on the theory of chaos. In our opinion, the advantages of studying non linear time series from this perspective, already discussed in the seminal work by Tong (1995), has not yet been fully explored in the statistical literature. Second, we reconsider some recent developments in the analysis of non-stationary multivariate systems in the light of the notion of Common Features (Engle and Kozicki, 1993). Such notion is inspired by the evidence that time series often display similar dynamic characteristics, which can be removed by linearly combining them. Modeling such similarities leads to parsimonious representations of multivariate dynamic systems.

In Section 2 we briefly introduce deterministic and chaotic systems. Section 3 presents an overview of some the most recent developments in the field of non linear univariate stochastic model. Finally, Section 4 deals with the analysis of common features among non-stationary time series.

2. Deterministic and stochastic systems

A dynamical system can be defined (Diks, 1999) as a triple (Ω, ϕ, T) where $\Omega \subseteq \Re^d$ is a *state space* or *phase space* representing all possible states $\mathbf{x}(t)$ of the system, $\phi : \Omega \times T \to \Omega$ is an evolution operator and T denotes the set of possible times.

For each fixed t, $\phi(\mathbf{x}(t),t)$ defines a map from Ω to Ω called the *flow* over time t while, for fixed $\mathbf{x}\in\Omega$, $\phi(\mathbf{x},t)$ is called the *evolution* of $\mathbf{x}$ or the *trajectory* through $\mathbf{x}$.

In the following we consider the case of a discrete time dynamical system ($T=Z$); it can be defined in term of the map ϕ and evolutions can be obtained by successive application; that is $\mathbf{x}_t=\phi^t(\mathbf{x}_0)$ with $\mathbf{x}_0=\mathbf{x}(0)\in\Omega$, $t\geq 1$, ϕ^t being the t-fold composition of ϕ.

From a mathematical viewpoint, the dynamical system can be explained in term of its long run position, so looking at the set of points toward which the trajectories converge as $t\to\infty$. This defines the so called attractor which allows to evaluate the stability of the system and to characterize a chaotic system

The nature of chaos is surprisingly simple: it comes if the dynamical system is globally bounded and there is sensitivity to the initial value x_0 when iterating with ϕ. The presence of chaos can be also conjectured by examining constants associated with dynamical systems called the Lyapunov exponents (Wolff, 1992).

In practice, one rarely has the advantage of observing the state of the system, let alone knowing the actual functional form that generate the dynamics. The *reconstruction theorem* (Takens, 1981) enables the reconstruction of the asymptotic dynamics of a dynamical system from an observed deterministic time series and suggests that time lags of a single variable can serve as surrogates for the unobserved variables of the system. In particular, if $(x_i,x_{i-1},\dots x_{i-m})$ is the delay vector of embedding dimension m, under general conditions, Takens (1981) proved that for m>2d, smooth dynamics are induced on delay vectors, so that there is a smooth map $f:\Re^m\to\Re$ such that $x_t=f(x_{t-1},x_{t-2},\dots x_{t-m})$ for all integers t. Putting $\xi_t=(x_t,x_{t-1},\dots x_{t-m+1})'$ and $\mathbf{F}(\xi_t)=\left(f(\xi_t),x_t,\dots,x_{t-m+2}\right)'$, the previous equation can be also written as $\xi_t=\mathbf{F}(\xi_{t-1})$.

The reconstructed attractor has some geometrical structure, endowed with a measure related to the relative frequencies with which different parts of the attractor are visited.

However, dynamical systems as well as reconstruction formulations are, by definition, deterministic while empirical and experimental time series are never entirely noise-free.

Thus a more “reasonable” approach is to model “chaotic” data by a non linear dynamical system with dynamic noise (Cheng and Tong, 1992, *inter alia*)

If noise sources are considered to be the essential elements for models of erratic time series, stochastic models have to be used. A natural way to incorporate stochastic perturbations in a system and to find the connection between the theory of dynamical systems and of stochastic processes is to enlarge the trajectory defined through the state vector ξ_t (to which a Dirac-δ-measure is associated) to a trajectory $(\mu_0,\mu_1,\dots)$ in the space of probability measure on $\Re^d$. Here $\mu_r=\mu_r(\mu_0)$ denotes the probability measure of ξ_r given that ξ_0 has a probability measure μ_0. Let us now impose a Markovian assumption such that $\{\xi_t\}$ follows a Markov chain on $\Re^d$ and that $\mu_{s+t}(\mu_0)=\mu_s(\mu_t(\mu_0))$.

It is reasonable to represent the Markov chain $\{\xi_t\}$ on $\Re^d$ in the form: $\xi_{t+1}=F(\xi_t)+\varepsilon_{t+1}$ where $\{\varepsilon_t\}$ is a sequence of *i.i.d.* d-dimensional random vectors and ε_t is independent of ξ_s $0 \le s < t$. Using the previous definition of the vector ξ_t and by an abuse of notation on the function *f*, we obtain the well known non linear autoregressive model of order *d*:

$$X_t = f(X_{t-1}, \ldots, X_{t-d}) + \varepsilon_t \quad \text{t=0,1,...} \tag{1}$$

In model (1) the independence assumption on $\{\varepsilon_t\}$ can be relaxed and a mixing dependence structure can be adopted. In this way, the process is still sufficiently well behaved to allow laws of large numbers and central limit theorems to be established, thus making possible a satisfactory asymptotic theory of estimation and inference.

3. Non-linear univariate stochastic models

Due to the great variety of ways in which non linearity can arise in observed time series, the theory of non linear models is generally very complex and it is not possible to give general conditions for stationarity, ergodicity and invertibility which are available only for certain special case. Chan ant Tong (1985, 1994) derived conditions under which a skeleton can yield an ergodic stochastic process while stationarity, a condition always required for non linear processes, can be verified using the Markov chain representation of the process. Invertibility is also a property of the model that allows the evaluation of point predictions for non linear models but it is not easy to verify in practice.
A natural generalization of univariate ARIMA models is the class of Bilinear models (Subba Rao, 1981) which are able to describe stationary level and occasional sharp spikes, a typical feature of financial time series. A complete bilinear time series model *BL(p,r,m,k)* has the form:

$$X_t + \sum_{i=1}^{p} a_i X_{t-i} = \sum_{i=0}^{r} c_i \varepsilon_{t-i} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-i} \varepsilon_{t-j} \tag{2}$$

where $c_0 = 1$ and $\{e_t\}$ is a white noise process with zero mean and finite variance σ^2.
Tong (1990) discussed the relationships between the (deterministic) bilinear system and the model (2) while Pham Dinh Tuan (1985) gave the Markovian form of the class of subdiagonal bilinear models. Sufficient conditions for stationarity, invertibility and ergodicity are also available for certain special cases (Quinn, 1988).
A different approach to model non linearity is the class of Threshold Autoregressive (TAR) models (Tong,1990), widely applied in modeling of nonlinear time series which exhibit asymmetry, jump phenomena and limit cycles. These models have a strong connection with the theory of dynamical systems through the *threshold principle* which allows the decomposition of a complex stochastic system into simpler subsystems. The canonical form of a threshold autoregressive model is given by:

$$X_t = B^{(J_t)} X_t + A^{(J_t)} X_{t-1} + H^{(J_t)} \varepsilon_t + C^{(J_t)} \tag{3}$$

where for J_t=j, $A^{(j)}$, $B^{(j)}$ and $H^{(j)}$ are (k,k) matrix coefficients , $C^{(j)}$ is a (k,1) vector of constants and $\{\varepsilon_t\}$ is a sequence of i.i.d. k-dimensional random vectors with zero mean and a constant covariance matrix. The class of models defined in (3) includes as special cases the SETAR models, the SETARMA models, the TARSO models (for a review, see Tong, 1990 and references therein) and the Switching Regime models (Hamilton, 1989). Probabilistic properties of the class of TAR models concerning stationarity, invertibility and ergodicity have been proved only for some specific cases.
More recently, researches have moved their interest towards the study of the implications of non linearity for the behavior of high order moments of the process. This motivated the development of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models (Bollerslev, 1986), which consider a multiplicative model for X_t

$$X_t = h_t^{1/2}\varepsilon_t \text{ with } h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}$$

where $\{\varepsilon_t\}$ are i.i.d. random variables with standard Normal distribution, q and p are non negative integer, α_i and β_j are non negative parameters $\forall i, j$ ($\alpha_0 > 0$) If $\sum_{i=1}^{q}\alpha_i + \sum_{j=1}^{p}\beta_j < 1$ the model has finite variance.
By recognizing $\{X_t^2\}$ as a special case of a bilinear time series, it is possible to derive its Markovian representation while theoretical results about stationarity, autocorrelation structure, estimation and identification are provided in Bollerslev (1986) and He and Terasvirta, (1999), *inter alia*.
A completely different strategy for the analysis of model (1), which does not assume any a priori parametric forms for *f(.)*, is the non parametric approach (Tjostheim, 1994; Härdle *et al.*, 1997). The works of Casdagli (1992) and Yao and Tong (1994) focus on this approach from a dynamic system perspective.
However, because of the so called *curse of dimensionality* due to sparse data in high dimension, local techniques, such as local regression methods and kernel, typically require d to be 1 or 2. Considerably more flexible non linear time series modeling methodologies seem to be the additive autoregressive models (Chen and Tsay, 1993), the multivariate adaptive regression splines (Friedman, 1991) and the adaptive spline threshold autoregression (Lewis and Stevens, 1991). A more general approach, which circumvents the necessity of additive hypotheses and allows to approximate complex non linear stochastic models, is the approach based on neural networks (Refenes and Zapranis, 1999 and references therein) and more recently on stochastic neural networks (Lai and Wong, 2001).

4. Multivariate stochastic models

Multiple time series analysis is often subject to the *curse of dimensionality* due to the large number of parameters to be estimated. Indeed, let us consider a fairly simple multivariate stochastic dynamic system such as the Vector Auto Regressive (VAR) model:

$$A(L)\mathbf{X}_t = \Phi D_t + \boldsymbol{\varepsilon}_t , \tag{4}$$

where $\mathbf{X}_t$ is a *n*-vector time series with fixed values $\mathbf{X}_{-p+1},...,\mathbf{X}_0$, $A(L) = I_n - \sum_{i=1}^{p} A_i L^i$, L is the lag operator, $\boldsymbol{\varepsilon}_t$ is a sequence of i.i.d. $N_n(0,\Sigma)$ variables, and D_t is a *m*-vector of deterministic elements.

Notice that the conditional mean of time series $\mathbf{X}_t$ depends on $n(pn+m)$ parameters, which is typically a quite large number in empirical applications. Hence, a large body of research has recently focused on parsimonious representations of multivariate dynamic systems. Engle and Kozicki (1993) proposed the general notion of Common Features (CF) which encompasses many of the developments in this field.

A feature is defined as a time-series property of the data which respects the following axioms:

1. The vector series $\mathbf{X}_t$ has (does not have) the feature if any non-singular linear transformation of $\mathbf{X}_t$ still has (does not have) such feature;
2. If a *n*-vector time series $\mathbf{Y}_t$ does not have the feature and $\mathbf{X}_t$ does not have the feature then $\mathbf{X}_t + \mathbf{Y}_t$ does not have the feature;
3. If $\mathbf{Y}_t$ does not have the feature but $\mathbf{X}_t$ has the feature then $\mathbf{X}_t + \mathbf{Y}_t$ has the feature.

Examples of features include serial correlation, non-stationarity, seasonality, and volatility. A feature, which is present in set of series, is said to be common when there exists a linear combination of these series which does not have the feature. The practical advantage of taking into account the presence of CF is that less parameters are needed to model series which share a given dynamic property. In the sequel we will review some relevant cases of CF.

4.1. Common unit roots

Since many observed time series display non-stationary characteristics, a large body of research has recently arisen from cointegration analysis (Engle and Granger, 1987). The basic idea underlying cointegration is that a set of non-stationary time series may be linearly combined to get a stationary time series. In particular, a *n*-vector time series $\mathbf{X}_t$ is defined integrated of order 1 [I(1)] if it is generated by the following stochastic process

$$\Delta\mathbf{X}_t = \Theta D_t + C(L)\boldsymbol{\varepsilon}_t , \tag{5}$$

where $\Delta = (1-L)$, $C(L) = I_n + \sum_{i=1}^{\infty} C_i L^i$, $\sum_{i=1}^{\infty} i\left|C_i\right| < \infty$, and $C(1) \neq 0$.

Notice that the above definition implies that the process $(\Delta\mathbf{X}_t - \Theta D_t)$ is stationary. Consequently, we define the increments $\Delta\mathbf{x}_t$ as being integrated of order zero [I(0)].

An I(1) time series vector admits the Beveridge-Nelson (1981) [BN] decomposition

$$\mathbf{X}_t = \Theta^* D_t^* + \sum_{i=0}^{\infty} C(1)\boldsymbol{\varepsilon}_{t-i} + C^*(L)\boldsymbol{\varepsilon}_t , \tag{6}$$

where $\Delta\Theta^* D_t^* = \Theta D_t$, $C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i$, and $C_i^* = -\sum_{j>i} C_i$ for all i.

The BN representation decomposes series $(\mathbf{X}_t - \Theta^* D_t^*)$ in a random-walk component or trend, that is the long-term forecast of such series, and in a stationary remainder or cycle.

Suppose that there exists a $n \times r$-matrix β with rank equal to r such that $\beta' C(1) = 0$. In this case, series $\mathbf{X}_t$ are driven by $(n-r)$ common stochastic trends (Stock and Watson, 1988). Premultiplying both sides of Equation (6) by β', we see that series $\beta' \mathbf{X}_t$ are I(0). Engle and Granger (1987) defined series $\mathbf{X}_t$ as cointegrated of order one [C(1,1)]. Moreover, they proved that if series $\mathbf{X}_t$ admit a VAR representation, Equation (4) may be rewritten in the following Vector Error Correction Model (VECM) is

$$\Gamma(L)\Delta\mathbf{X}_t = \Phi D_t + \alpha\beta'\mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{7}$$

where $\Phi D_t = A(L)\Theta^* D_t^*$, α is a $n \times r$-matrix with rank equal to r such that $\alpha\beta' = A(1)$, $\Gamma(L) = I_n - \sum_{i=1}^{p-1} \Gamma_i L^i$, and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$ for $i = 1,2,\ldots, p-1$.

The VECM is formally analogous to a VAR model of series $\Delta\mathbf{X}_t$ augmented with the exogenous terms $\beta'\mathbf{X}_{t-1}$. Such additional terms represent the stationary deviations from the attractor of the system.

The maximum likelihood analysis of the VECM may be conducted by means of reduced-rank regression. However, both the test statistics for cointegration and the estimator of β have non-standard limit distributions which can be expressed as functionals of multivariate Brownian Motions, see Johansen (1996).

The basic cointegration approach has recently been extended in several directions including the analyses of seasonally integrated VAR models (Johansen and Schaumburg, 1998; Cubadda. 2001), fractional unit-root processes (Marinucci, and Robinson, 2001), and I(2) time series (Johansen, 1997; Paruolo, 2000).

4.2. Common serial correlation

Cointegration is essentially a long-run property of the data. However, many empirical studies are instead devoted to the transitory components of non-stationary time series. A remarkable example is business cycle analysis in economics. Engle and Kozicki (1993) proposed the notion of the Serial Correlation Common Feature (SCCF) as a measure of short-run comovement among I(1) time series. The existence of the SCCF requires that there exists a $n \times s$-matrix δ with rank equal to s such that series $\delta'\Delta\mathbf{X}_t$ are innovations with respect to the σ-field generated by $\{\mathbf{X}_{t-i}; i \geq 1\}$. This is equivalent to require that $\delta' C(L) = \delta'$.

Vahid and Engle (1993) noticed that it must also hold that $\delta' C^*(L) = 0$. Hence, series $\mathbf{X}_t$ share $(n-s)$ common cycles in their BN decomposition. Moreover, if we premultiply both sides of Equation (7) by δ', we see that all the coefficient matrices of the VECM, but Φ, must have the same left-null space. Hence, under the SCCF series $\mathbf{X}_t$ admit the following common factor representation:

$$\Delta \mathbf{X}_t = \Phi D_t + \Lambda \mathbf{F}_t + \boldsymbol{\varepsilon}_t \tag{8}$$

where Λ is a full-rank $n \times (n-s)$-matrix such that $\delta' \Lambda = 0$,

$$\mathbf{F}_t = \tilde{\alpha} \beta' \mathbf{X}_{t-1} + \sum_{i=1}^{p-1} \tilde{\Gamma}_i' \Delta \mathbf{X}_{t-i},$$

$\tilde{\alpha}$ is a $(n-s) \times r$-matrix, and $\tilde{\Gamma}_i$ is a $n \times (n-s)$-matrix for $i = 1,2,..., p-1$.
Recent extensions of the SCCF approach include the notions of codependent cycles (Vahid and Engle, 1997), common stochastic seasonal features (Cubadda, 1999), and polynomial SCCF (Cubadda and Hecq, 2001). Optimal statistical inference on all these models is obtained by reduced-rank regression and standard limit theory is involved.

References

Beveridge S., Nelson C.R. (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of Monetary Economics*, 7, 151-174.

Bollerslev T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-328.

Casdagli M. (1992) Chaos and Deterministic *versus* Stochastic Non-linear Modelling, *J.R.S.S.*, B, 54, 303-328.

Chan K.S., Tong H. (1986) On Estimating thresholds in autoregressive models, *Journal of Time Series Analysis*, 7, 179-190

Chan K.S., Tong H. (1994) A Note on Noisy Chaos, *J.R.S.S.*, B, 56, 301-311.

Chen R., Tsay R. (1993) Nonlinear additive ARX models, *Journal of the American Statistical Association*, 88, 955-967.

Cheng B., Tong H. (1992) On Consistent Nonparametric Order Determination and Chaos, *J.R.S.S.*, B, 54, 427-449.

Cubadda G. (1999) Common cycles in seasonal non-stationary time series, *Journal of Applied Econometrics*, 14, 273-291.

Cubadda G. (2001) Complex reduced rank models for seasonally cointegrated time series, *Oxford Bulletin of Economics and Statistics*, 63, 497-511.

Cubadda G., A. Hecq (2001) On Non-Contemporaneous Short-Run Comovements, *Economics Letters*, 73, 389-397.

Diks C. (1999) *Nonlinear Time Series Analysis: Methods and Applications*, World Scientific, Singapore.

Engle R.F., Granger C.W.J. (1987) Cointegration and error correction: representation, estimation and testing, *Econometrica*, 55, 251-78.

Engle R.F., Kozicki S. (1993) Testing for common features, *Journal of Business and Economic Statistics*, 11, 369-395 with discussions.

Friedman J.H. (1991) Multivariate adaptive regression spline (with discussion), *The Annals of Statistics*, 19, 1-141.

Hamilton J.D.A. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 57, 357-384.

Härdle W., Lütkepohl H., Chen R. (1997) A review of nonparametric time series analysis, *International Statistical Review*, 65, 49-72.

He C., Terasvirta T. (1999) Fourth moment structure of the GARCH(p, q) Process, *Econometric Theory*, 15, 824-846.

Johansen S. (1996) *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press, Oxford.

Johansen S. (1997) Likelihood Analysis of the I(2) Model, *Scandinavian Journal of Statistics*, 24, 433-462.

Johansen S., Schaumburg E. (1998) Likelihood analysis of seasonal cointegration, *Journal of Econometrics*, 88, 301-339.

Lay T.L.L., Wong S.P.S. (2001) Stochastic Neural Networks with applications to nonlinear time series, *Journal of the American Statistical Association*, 96, 968-981.

Lewis P.A.W., Stevens J. G. (1991) Nonlinear modeling of time series using multivariate adaptive regression splines (MARS), *Journal of the American Statistical Association*, 86, 864-877.

Marinucci, D., Robinson P.M. (2001), Semiparametric Fractional Cointegration Analysis, *Journal of Econometrics*; 105, 225-47.

Paruolo, P. (2000) Asymptotic Efficiency of the Two Stage Estimator in I(2) Systems, *Econometric Theory*, 16, 524-50.

Quinn B.G. (1988) Stationarity and invertibility of simple bilinear models, *Stochastic Processes and their Applications*, 12, 225-229.

Refenes A.P.N., Zapranis A.D. (1999) Neural Model Identification, Variable Selection and Model Adequacy, *Journal of Forecasting*, 18, 299-332.

Stock, J.H., Watson M.W. (1988) Testing for common trends, *Journal of the American Statistical Association*, 83, 1097-1107.

Subba Rao T. (1981) On the theory of bilinear time series models, *Journal of the Royal Statistical Society*, B, 43, 244-255

Subba Rao T., Gabr M.M. (1984) *An introduction to bispectral analysis and bilinear time series models*, Lecture Notes in Statistics, 24, Springer & Verlag , NY.

Takens F. (1981) Detecting strange attractors in fluid turbolence, in *Dynamical System and Turbolence* (eds. D. Rand and L.S. Young), Springer& Verlag , NY

Tjøstheim D. (1994) Non-linear Time Series: A Selective Review, *Scandinavian Journal of Statistics*, 21, 97-130.

Tong H. (1990) *Non Linear Time Series: A Dynamical System Approach*, Clarendon Press, Oxford.

Tong H. (1995) A Personal Overview of Non-linear Time Series Analysis from a Chaos Perspective, *Scandinavian Journal of Statistics*, 22, 399-445.

Tong H., Lim K.S. (1980) Threshold Autoregression, Limit Cycles and Cyclical Data, *Journal of the Royal Statistical Society*, B, 42, 245-292.

Vahid, F., Engle R.F. (1993) Common trends and common cycles, *Journal of Applied Econometrics*, 8, 341-360.

Vahid, F., Engle R.F. (1997) Codependent cycles, *Journal of Econometrics*, 80, 199-221.

Wolff R.C.L. (1992) Local Lyapunov Exponents: Looking Closely at Chaos, *Journal of the Royal Statistical Society*, B, 54, 353-371.

Yao Q., Tong H. (1994) Quantifying the influence of initial values on non-linear prediction, *Journal of the Royal Statistical Society*, B, 56, 701-725.