Journal of Econometrics 178 (2014) 331–341

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Summability of stochastic processes—A generalization of integration for non-linear processes

Vanessa Berenguer-Rico [a,*], Jesús Gonzalo [b,*]

[a] Department of Economics, University of Oxford and Institute for New Economic Thinking, United Kingdom

[b] Department of Economics, Universidad Carlos III de Madrid, Spain

ARTICLE INFO

Article history:
Available online 5 September 2013

JEL classification:
C01
C22

Keywords:
Co-integration
Co-summability
Integrated processes
Non-linear balanced relationships
Non-linear processes
Summability

ABSTRACT

The order of integration is valid to characterize linear processes; but it is not appropriate for non-linear worlds. We propose the concept of summability (a re-scaled partial sum of the process being $O_p(1)$) to handle non-linearities. The paper shows that this new concept, $S(\delta)$: (i) generalizes $I(\delta)$; (ii) measures the degree of persistence as well as of the evolution of the variance; (iii) controls the balancedness of non-linear relationships; (iv) opens the door to the concept of co-summability which represents a generalization of co-integration for non-linear processes. To make this concept empirically applicable, an estimator for δ and its asymptotic properties are provided. The finite sample performance of subsampling confidence intervals is analyzed via a Monte Carlo experiment. The paper finishes with the estimation of the degree of summability of the macroeconomic variables in an extended version of the Nelson–Plosser database.

1. Introduction

No one doubts that the concepts of integration and co-integration have been and still are very useful in time series econometrics. The former by producing a single parameter that was able to summarize the long-memory properties of a given time series. The latter by linking the existence of common trends to long-run linear equilibrium relationships. Thanks, amongst others, to the work by Dickey and Fuller (1979), Nelson and Plosser (1982), Phillips (1986), Engle and Granger (1987) and Johansen (1991), these two concepts are easily handled theoretically as well as empirically.

In parallel, non-linear time series models from a stationary perspective were introduced in the literature—see Granger and Teräsvirta (1993), Franses and van Dijk (2000), Fan and Yao (2003), and Teräsvirta et al. (2010) for some overviews. The introduction of persistent variables into non-linear models – see Park and Phillips (1999, 2001), de Jong and Wang (2005) or Pötscher (2004) for the study of transformations of integrated processes – produced a natural query: Which is the order of integration of these non-linear transformations? Such a question does not have a clear answer since the existing definitions of integrability do not properly apply. Integration is a linear concept.

This lack of definition has at least two important worrying consequences. First, in univariate terms, it implies that an equivalent synthetic measure of the stochastic properties of the time series, such as the order of integration, is not available to characterize non-linear time series. As pointed out by Granger (1995), this does not only affect econometricians but also economic theorists who need to account for important characteristics of economic variables to construct their theories. Second, from a multivariate perspective, it becomes troublesome to determine whether a non-linear model is balanced or not. Unbalancedness is a symptom of a misspecified model, a feature that is easily likely to occur when managing non-linear transformations of persistent variables. In linear setups, the concept of integrability did a good job dealing with balanced/unbalanced relations. However, in non-linear frameworks, the non-existence of a synoptic quantitative measure makes it difficult to check the balancedness of a postulated model.

Additionally, this implies that a definition for non-linear co-integration is difficult to be obtained from the usual concept of integrability. To clarify this point, suppose $y_t = f(x_t, \theta) + u_t$, where $x_t \sim I(1)$ and $u_t \sim I(0)$. For $f(\cdot)$ non-linear, the order of integration of $f(x_t, \theta)$, and hence that of y_t, may not be properly defined implying that the standard concept of co-integration is difficult to be applied. In fact, the literature on non-linear co-integration – see Park and Phillips (2001), Karlsen et al. (2007), Wang and Phillips (2009) – undertakes the whole analysis assuming the existence of a long-run relationship; something that should be tested in practice.

* Corresponding authors.
E-mail addresses: vanessa.berenguer-rico@nuffield.ox.ac.uk (V. Berenguer-Rico), jesus.gonzalo@uc3m.es (J. Gonzalo).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.031

It was already stated in Granger and Hallman (1991) that extensions of the linear concepts $I(0)$ and $I(1)$ are needed to generalize co-integration to non-linear frameworks. This has led some authors to introduce alternative definitions. For instance, Granger (1995) proposed the concepts of Extended and Short Memory in Mean. However, these concepts are neither easy to calculate nor general enough to handle some types of non-linear long run relationships. And, furthermore, a measure of the order of the extended memory is not available. Dealing with threshold effects in co-integrating regressions, Gonzalo and Pitarakis (2006) faced these problems and proposed, in a very heuristic way, the concept of summability (a re-scaled partial sum of the process being $Op(1)$). However, they did not emphasize the avail of such an idea.

In this paper, we define summability properly and show its usefulness and generality. Specifically, we put forward several relevant examples in which the order of integrability is difficult to be established, but the order of summability can be easily determined. Moreover, we show that integrated time series are particular cases of summable processes, in the sense that the order of summability is the same as the order of integration. Hence, summability is a generalization of integrability. Furthermore, summability does not only characterize some properties of univariate time series, but also allows to easily study the balancedness of a postulated relationship —linear or not. And even more important, non-linear long run equilibrium relationships between non-stationary time series can be properly defined. In particular, the concept of co-summability, which can be applied to extend co-integration to non-linear frameworks, is developed by the authors in Berenguer-Rico and Gonzalo (2013).

To make this concept empirically operational, we propose a statistical procedure to estimate and carry out inferences on the order of summability of an observed time series. This makes useful the concept of summability not only in theory but also in practice. To estimate the order of summability, we use an estimator introduced by McElroy and Politis (2007) to analyze the rate of convergence of a statistic which is obtained from a simple least squares regression. The inference on the true order of summability is based on the subsampling methodology developed in Politis et al. (1999). It is shown, by simulations, that the subsampling machinery works reasonably well in finite samples given the generality of the approach. Finally, the proposed methodology is used to estimate the order of summability of the macroeconomic time series in an extended version of the Nelson–Plosser database.

The paper is organized as follows. In the next section, the problems of using the order of integration to characterize non-linear processes are highlighted. In Section 3, our proposed solution based on summability is described and its simple applicability showed. Section 4 describes the statistical tools – estimation and inference – to deal empirically with summable processes. In Section 5, an empirical application shows how to determine the order of summability in practice. Finally, Section 6 is devoted to some concluding remarks. All proofs are collected in the Appendix.

A word on notation. We use the symbol "$\Longrightarrow$" to signify convergence in distribution and weak convergence indistinctly, "$\xrightarrow{p}$" to signify convergence in probability. Stochastic processes such as the standard Brownian motion $W(r)$ are defined on $[0, 1]$. Finally, all limits given in the paper are taken as the sample size $n \to \infty$.

2. Order of integration and non-linear processes

2.1. Order of integration

Definition 1. A stochastic process $\{y_t : t \in \mathbb{N}\}$ is said to be an integrated process of order d (in short, an $I(d)$ process) if the process of dth order differences $\Delta^d y_t$ is $I(0)$.

A natural question that arises after reading this definition is: And what is an $I(0)$ process? Attempts to give a formal description of $I(0)$ processes exist in the literature. Engle and Granger (1987) give the following characterization.

Engle and Granger (EG) Characterization. *If $y_t \sim I(0)$ with zero mean then* (i) *the variance of y_t is finite;* (ii) *an innovation has only a temporary effect on the value of y_t;* (iii) *the spectrum of y_t, $f(\omega)$, has the property $0 < f(0) < \infty$;* (iv) *the expected length of time between crossing of $x = 0$ is finite;* (v) *the autocorrelations, ρ_k, decrease steadily in magnitude for large enough k, so that their sum is finite.*

Other characterizations have been used as well. Granger (1995) and Johansen (1995) used autoregressive and moving average representations, respectively. Müller (2008) and Davidson (2009) – among others – define an $I(0)$ as a process that satisfies the functional central limit theorem (FCLT). These latter definitions share the same spirit of our summability definition in Section 3. Nevertheless, in all cases, differences must be taken to discover the order of integration and the intrinsic linearity of the difference operator makes it difficult, if not impossible, to characterize – among others – non-linear processes. Integration is a linear concept.

2.2. Examples

Example 1. Alpha Stable i.i.d. Distributed Processes

Let y_t be *i.i.d.* from some distribution $F \in D(\alpha)$, where $D(\alpha)$ denotes the domain of attraction of an α-stable law with $\alpha \in (0, 2]$. y_t is strictly stationary; however, its second moments may not exist. The fact that such a process is *i.i.d.* could incline to think that this process is $I(0)$. However, if second moments do not exist, EG characterization does not apply. Characterizations based on the FCLT could not be used either since they assume a standard Brownian motion in the limit. Hence, it becomes troublesome to establish the order of integration of y_t.

Example 2. An i.i.d. plus a Random Variable

Consider the following process

$$y_t = z + e_t, \tag{1}$$

where $z \sim N(0, \sigma_z^2)$ and $e_t \sim i.i.d.(0, \sigma_e^2)$ are independent of each other. This process has the following properties:

(i) $E[y_t] = 0$
(ii) $V[y_t] = \sigma_z^2 + \sigma_e^2$
(iii) $\gamma_y(k) = Cov(y_t, y_{t-k}) = \sigma_z^2$ for all $k > 0$.

Since it is a strictly stationary process, one could think that it is $I(0)$. However, the autocovariance function is not absolutely summable and its spectrum does not satisfy the required condition in EG characterization.[1] If y_t is not $I(0)$, to attach any other order of integration to this stochastic process is not obvious. It is controversial to say y_t is $I(1)$ since $\Delta y_t = \Delta e_t$ is generally understood as an $I(-1)$; and it becomes difficult to choose any other number using the above definition of order of integration.

Dealing with non-linear processes similar problems are faced.

[1] The autocovariance of the process in this example can be expressed as

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} \left[\frac{\sigma_z^2 + \sigma_e^2}{2\pi} + \frac{\sigma_z^2}{\pi} \sum_{h=1}^{\infty} \cos(\lambda h) \right] d\lambda.$$

Then, the spectral density is

$$f(\lambda) = \frac{\sigma_z^2 + \sigma_e^2}{2\pi} + \frac{\sigma_z^2}{\pi} \sum_{h=1}^{\infty} \cos(\lambda h),$$

which diverges for all λ.