Journal of Econometrics 186 (2015) 245–257

Contents lists available at ScienceDirect

Journal of Econometrics

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

The generalised autocovariance function

Tommaso Proietti[a,b,*], Alessandra Luati[c]

[a] Department of Economics and Finance, University of Rome Tor Vergata, Italy
[b] CREATES, Aarhus University, Denmark
[c] Department of Statistics, University of Bologna, Italy

ARTICLE INFO

Article history:
Received 10 January 2013
Received in revised form
23 March 2014
Accepted 24 July 2014
Available online 4 August 2014

Keywords:
Stationary processes
Spectral estimation
White noise tests
Feature matching
Discriminant analysis

ABSTRACT

The generalised autocovariance function is defined for a stationary stochastic process as the inverse Fourier transform of the power transformation of the spectral density function. Depending on the value of the transformation parameter, this function nests the inverse and the traditional autocovariance functions. A frequency domain non-parametric estimator based on the power transformation of the pooled periodogram is considered and its asymptotic distribution is derived. The results are employed to construct classes of tests of the white noise hypothesis, for clustering and discrimination of stochastic processes and to introduce a novel feature matching estimator of the spectrum.

1. Introduction

The temporal dependence structure of a stationary stochastic process is characterised by the autocovariance function, or equivalently by its Fourier transform, the spectral density function. We extend this important concept, by introducing the generalised autocovariance function (GACV), which we define as the inverse Fourier transform of the pth power of the spectral density function, where p is a real parameter. The GACV depends on two arguments, the power parameter p and the lag k. Dividing by the GACV at lag zero for p given yields the generalised autocorrelation function (GACF).

For $k = 0$ the GACV is related to the variance profile, introduced by Luati et al. (2012) as the Hölder mean of the spectrum. For $p = 1$, it coincides with the traditional autocovariance function, whereas for $p = -1$ it yields the inverse autocovariance function, as k varies. The extension to any real power parameter p is fruitful for many aspects of econometrics and time series analysis. We focus in particular on model identification, time series clustering and discriminant analysis, the estimation of the spectrum for cyclical time series, and on testing the white noise hypothesis and goodness of fit.

The underlying idea, which has a well established tradition in statistics and time series analysis (Tukey, 1957; Box and Cox, 1964), is that taking powers of the spectral density function allows one to emphasise certain features of the process. For instance, we illustrate that setting $p > 1$ is useful for the identification of spectral peaks, and in general for the extraction of signals contaminated by noise. Moreover, fractional values of $p \in (0, 1)$ enable the definition of classes of white noise tests with improved size and power properties, with respect to the case $p = 1$, as the finite sample distribution can be made closer to the limiting one by the transformation that is implicit in the use of the GACV. Finally, by solving a generalised Yule–Walker system of equations based on the GACV, we can estimate a general spectral model that, according to the value of p, encompasses both autoregressive and moving average spectral models.

For given stochastic processes the GACV can be analytically evaluated in closed form in the time domain by constructing the standard autocovariance function of an auxiliary stochastic process, whose Wold representation is obtained from the original one, by taking a power transformation of the Wold polynomial.

As far as estimation from a time series realisation is concerned, we consider a nonparametric estimator based on the power transformation of the pooled periodogram. For a given p, the estimator is asymptotically normally distributed around the population value, with a variance that depends on the GACV evaluated at $2p$; as a result, a consistent estimator of the asymptotic variance is readily available. We also show that Bartlett's formula generalises to any

* Correspondence to: Department of Economics and Finance, via Columbia 2, 00133, Rome, Italy. Tel.: +39 06 72595941.
E-mail addresses: tommaso.proietti@uniroma2.it (T. Proietti), alessandra.luati@unibo.it (A. Luati).

http://dx.doi.org/10.1016/j.jeconom.2014.07.004
0304-4076/

value of p. As a related result we derive the asymptotic distribution of a ratio estimator of the GACF.

These results open the way to the application of the GACV for the analysis of stationary time series. In addition to the possible uses hinted above (model identification, testing for white noise, and feature extraction), we consider the possibility of defining measures of pairwise distance based on the GACV or GACF, encompassing the Euclidean and the Hellinger distances, and we illustrate their use for discriminant and cluster analysis of time series. Negative values can be relevant as they nest the Euclidean and the Hellinger distances based on the inverse autocorrelation functions.

The structure of the paper is the following. The GACV and the GACF are formally defined in Section 2. The interpretation in terms of the autocovariance function of a suitably defined power-transformed process is provided in Section 3. This is used for the analytical derivation of the GACV for autoregressive (AR) and moving average (MA) processes, as well as long memory processes (Section 4). Estimation is discussed in Section 5. Sections 6–8 focus on three main uses of the GACV and the GACF. The first deals with testing for white noise: two classes of tests, generalising the Box and Pierce (1970) test and the Milhøj (1981) statistics, are proposed and their properties discussed. It emerges that fractional values of p in the (0, 1) interval provide finite sample tests of the white noise hypothesis with improved size properties. A generalised Yule–Walker estimator of the spectrum based on the GACV is presented in Section 7: in particular, the GACV for $p > 1$ will highlight the cyclical features of the series; this property can be exploited for the identification and estimation of spectral peaks. We finally consider measures of distance between two stochastic processes based on the GACV or GACF and we illustrate their use for time series discriminant analysis. In Section 9 we provide some conclusions and directions for future research.

2. The generalised autocovariance function

Let $\{x_t\}_{t\in T}$ be a stationary zero-mean stochastic process indexed by a discrete time set T, with spectral distribution function $F(\omega)$. We assume that the spectral density function of the process exists, $F(\omega) = \int_{-\pi}^{\omega} f(\lambda)\mathrm{d}\lambda$, that it is positive, that the process is regular (Doob, 1953, p. 564), i.e. $\int_{-\pi}^{\pi} \log f(\omega)\mathrm{d}\omega > -\infty$, and that $\int_{-\pi}^{\pi} f(\omega)^p \mathrm{d}\omega < \infty$.

The generalised autocovariance (GACV) function is defined as the inverse Fourier transform of the pth power of the spectral density function,

$$\gamma_{pk} = \frac{1}{2\pi}\int_{-\pi}^{\pi}[2\pi f(\omega)]^p \cos(\omega k)\mathrm{d}\omega, \quad (1)$$

where we have replaced $\exp(\imath\omega k)$ by $\cos(\omega k)$ since the spectral density and the cosine are even functions while the sine function is odd. Taking the Fourier transform of γ_{pk} gives

$$[2\pi f(\omega)]^p = \gamma_{p0} + 2\sum_{k=1}^{\infty}\gamma_{pk}\cos(\omega k). \quad (2)$$

The coefficients γ_{pk} depend on two arguments, the integer lag k and the real power p. As a matter of fact, for $p = 1$, $\gamma_{1k} = \gamma_k$, the autocovariance of the process at lag k; for $p = 0$, $\gamma_{0k} = 0$, for $k \neq 0$ and $\gamma_{00} = 1$, up to a constant, the autocovariance function of a white noise process; for $p = -1$, $\gamma_{-1k} = \gamma_{ik}$, the inverse autocovariance function (Cleveland, 1972; see also Battaglia, 1983).

Other examples where integrals of powers of the spectral density function are of relevant interest may be found in recent advances on probabilistic approximations to Gaussian processes. As an example, for $k = 0$ and integer $p > 0$, the generalised variance gives the pth cumulant of the sample variance of a stationary zero mean stochastic process (see Nourdin and Peccati, 2012, formula 7.2.2). In turn, the cumulants enter in Berry–Esseen type bounds for the distance between the normalised asymptotic distribution of the sample variance and the standard Gaussian distribution (see Nourdin and Peccati, 2012, formulae 7.3.1 and 9.5.1, which involve fractional powers of the generalised variance).

As defined in (1) and due to the Herglotz theorem, the GACV is a true autocovariance (see also Section 3), and as such it inherits all the well known properties of an autocovariance function: an obvious property is symmetry with respect to the lag, $\gamma_{pk} = \gamma_{p,-k}$; moreover, $\gamma_{p0} > 0$ and $|\gamma_{pk}| \leq \gamma_{p0}$, for all integers k. Non-negative definiteness of the GACV follows from the assumptions on $f(\omega)$. The generalised autocorrelation function (GACF) is defined as

$$\rho_{pk} = \frac{\gamma_{pk}}{\gamma_{p0}}, \quad k = 0, \pm 1, \pm 2, \ldots, \quad (3)$$

taking values in $[-1, 1]$.

Further relevant properties are nested in the following Lemma, which is a consequence of the fact that the spectral density of a convolution is the product of the spectral densities (see corollary 3.4.1.1. in Fuller, 1996).

Lemma 1. *Let γ_{pk} be defined as in (1). Then,*

$$\gamma_{p+q,k} = \frac{1}{2\pi}\int_{-\pi}^{\pi}[2\pi f(\omega)]^{p+q}\cos(\omega k)\mathrm{d}\omega = \sum_{j=-\infty}^{\infty}\gamma_{p,j+k}\gamma_{q,j}. \quad (4)$$

An important special case of Lemma 1, that will be exploited later in the derivation of goodness of fit tests, relates the GACV with transformation parameter $2p$ to the GACV at p and is obtained by setting $p = q$ in Lemma 1:

$$\gamma_{2p,k} = \sum_{j=-\infty}^{\infty}\gamma_{pj}\gamma_{p,j+k}, \quad (5)$$

which for $k = 0$ specialises as

$$\gamma_{2p,0} = \gamma_{p0}^2 + 2\sum_{j=1}^{\infty}\gamma_{pj}^2.$$

Furthermore, setting $q = -p$ in Lemma 1, we obtain

$$\sum_{j=-\infty}^{\infty}\gamma_{pj}\gamma_{-p,j-k} = 1_{\{k=0\}}, \quad (6)$$

where $1_{\{A\}}$ indicates the indicator function on the set A. Property (6) extends the well known orthogonality between the autocovariance function and the inverse autocovariance function (see Pourahmadi, 2001, Theorem 8.12).

3. The power process and its autocovariance function

The function γ_{pk} lends itself to a further interpretation as the autocovariance function of a power process derived from x_t. This interpretation turns out to be useful in the derivation of the analytic form of γ_{pk}, as a function of the parameters that govern the process dynamics, by evaluating an expectation in the time domain, rather than solving (1) directly.

Assuming that $\{x_t\}_{t\in T}$ is purely non-deterministic, its Wold representation will be written as

$$x_t = \psi(B)\xi_t, \quad (7)$$

where $\xi_t \sim \mathrm{WN}(0, \sigma^2)$ and $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots$, with coefficients satisfying $\sum_{j=0}^{\infty}\psi_j^2 < \infty$, and such that all the roots of the characteristic equation $\psi(B) = 0$ are in modulus greater than one; here, $\mathrm{WN}(0, \sigma^2)$ denotes a white noise process, a sequence of