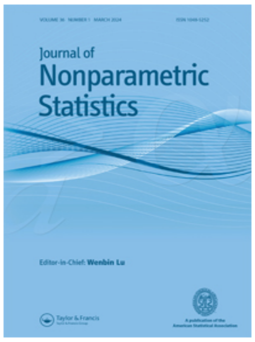

Taylor & Francis
Taylor & Francis Group

Journal of Nonparametric Statistics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gnst20

Efficient nonparametric estimation of generalised autocovariances

Alessandra Luati, Francesca Papagni & Tommaso Proietti

To cite this article: Alessandra Luati, Francesca Papagni & Tommaso Proietti (2024) Efficient nonparametric estimation of generalised autocovariances, Journal of Nonparametric Statistics, 36:1, 23-38, DOI: 10.1080/10485252.2023.2252527

To link to this article: https://doi.org/10.1080/10485252.2023.2252527

Published online: 02 Sep 2023.

JOURNAL OF NONPARAMETRIC STATISTICS
2024, VOL. 36, NO. 1, 23–38
https://doi.org/10.1080/10485252.2023.2252527

Efficient nonparametric estimation of generalised autocovariances

Alessandra Luati[a,b], Francesca Papagni[c] and Tommaso Proietti[d]

[a]Imperial College London London UK; [b]University of Bologna Bologna Italy; [c] Free University of Bozen, Bolzano, Italy; [d]University of Rome Tor Vergata Rome Italy

ABSTRACT

This paper provides a necessary and sufficient condition for asymptotic efficiency of a nonparametric estimator of the generalised autocovariance function of a stationary random process. The generalised autocovariance function is the inverse Fourier transform of a power transformation of the spectral density and encompasses the traditional and inverse autocovariance functions as particular cases. A nonparametric estimator is based on the inverse discrete Fourier transform of the power transformation of the pooled periodogram. We consider two cases: the fixed bandwidth design and the adaptive bandwidth design. The general result on the asymptotic efficiency, established for linear processes, is then applied to the class of stationary ARMA processes and its implications are discussed. Finally, we illustrate that for a class of contrast functionals and spectral densities, the minimum contrast estimator of the spectral density satisfies a Yule–Walker system of equations in the generalised autocovariance estimator.

ARTICLE HISTORY
Received 5 April 2023
Accepted 22 August 2023

KEYWORDS
Cramér–Rao lower bound; frequency domain; minimum contrast estimation; periodogram

The autocovariance function and its Fourier transform, the spectral density function, characterise the temporal dependence structure of a stationary stochastic process and are of fundamental importance in time series analysis and prediction. For Gaussian stationary processes they provide, along with the mean, a complete characterisation of the probability distribution of the process and, for linear processes, the basic ingredients for optimal (minimum mean squared error) prediction, based on time series observations.

The autocovariance function is estimated nonparametrically by the sample autocovariance function. This estimator has a long tradition in time series analysis, and its properties are demonstrated and discussed in time series textbooks, such as, for instance, Hannan (1970), Anderson (1994), and Brockwell and Davis (1991, ch. 7), where it is shown that, under regularity conditions, it has an asymptotically normal distribution and that the elements of the asymptotic covariance matrix are given by the celebrated Bartlett's formula. Wu (2011) has extended the theory to a class of nonlinear processes.

The literature has further addressed the important question as to what classes of parametric linear processes admit the sample autocovariance as an asymptotically efficient

CONTACT Alessandra Luati a.luati@imperial.ac.uk

estimator, i.e. an estimator whose asymptotic variance equals the Cramér–Rao lower bound arising in the parametric context of maximum likelihood or quasi maximum likelihood estimation.

The reverse problem of determining for which spectral models a given fundamental statistics (the sample covariance) is asymptotically efficient has been investigated by Porat (1987) for Gaussian autoregressive (AR) moving average (MA) mixed processes, based on state–space representations and matrix Lyapunov equation theory. For Gaussian ARMA(r, q) processes with $r \geq q$ the sample autocovariances are asymptotically efficient only in a restricted number of cases, while if $q > r$ none of the sample autocovariances is asymptotically efficient. See also Walker (1995) for an alternative derivation of this result. This implies that the variance and the first r autocovariances of a pure AR(r) process are efficiently estimated by the sample autocovariances, while for a pure MA process none of the sample autocovariances is asymptotically efficient. Kakizawa and Taniguchi (1994) derived, in the frequency domain, a necessary and sufficient condition for asymptotic efficiency of the sample autocovariances that applies to the more general class of Gaussian stationary processes. Kakizawa (1999) extended the previous results to the case of vector processes. Boshnakov (2005) studied the efficiency of the sample autocovariances for processes obtained by a finite linear transformation of a pure autoregressive process.

The generalised autocovariance (GACV) function was defined in Proietti and Luati (2015) as the inverse Fourier transform of the pth power of the spectral density function. It encompasses the traditional autocovariance function ($p = 1$) and the inverse autocovariance function (Cleveland 1972), which is the sequence of coefficients associated with the Fourier expansion of the inverse spectrum ($p = -1$). The GACV with noninteger p is a powerful tool for various purposes, including feature extraction and model identification, in the same spirit as in the approach proposed by Xia and Tong (2011). It can also be used for the specification of white noise tests and goodness of fit tests, related to the Hong (1996) and Chen and Deo (2004) test statistics, as well as for clustering and discriminant analysis of time series, in that a distance measure, nesting the Euclidean distance ($p = 1$) and the Hellinger distance ($p = 1/2$), can be defined based on the generalised autocovariances of two different stochastic processes, see Proietti and Luati (2015).

Following Hannan and Nicholls (1977) and Luati, Proietti, and Reale (2012), a nonparametric estimator of the GACV was defined in Proietti and Luati (2015), based on the powers of the pooled periodogram over m non-overlapping consecutive frequencies, where m is a fixed pooling parameter. Consistency and asymptotic normality of the estimator were established for linear processes. In this paper, we discuss the efficiency of this estimator and prove that its asymptotic variance equals the Cramér–Rao lower bound only for Gaussian processes and for $p = 1$. This limitation is due to an inefficiency factor that depends on the combination of the power p and the pooling parameter m, and can be controlled in finite samples.

To reach asymptotic efficiency, we introduce a novel estimator that can be interpreted as a more general formulation of the one defined in Proietti and Luati (2015), where the pooling parameter is allowed to grow with the sample size. We establish a necessary and sufficient condition for asymptotic efficiency in terms of the spectral density and its derivatives for linear processes, which nests as a particular case the result of Kakizawa and Taniguchi (1994), which holds for $p = 1$ and for Gaussian processes. The results also show that the asymptotic variance of the nonparametric estimator equals the Cramér–Rao lower bound for $p = -1$, i.e. it estimates efficiently the first q inverse autocovariances when the true generating process is pure MA(q), thereby complementing the results by Bhansali (1980) and Battaglia (1988). The inverse autocovariance function is useful in interpolation problems and for the identification of ARMA models.

As a further contribution of the paper, the asymptotic distribution of the estimator in the adaptive-bandwidth design is derived, following Bhansali (1980).

To illustrate our results, we consider the case of stationary ARMA(r, q) processes, in which case some numerical examples highlight the rate of convergence to the Cramér–Rao bound. The results obtained include, as a special case, the results for the sample autocovariance function by Porat (1987) and Kakizawa and Taniguchi (1994). Finally, we show that for a class of contrast functionals and spectral densities, the minimum contrast estimator of the spectral density satisfies a Yule–Walker system of equations in the original generalised autocovariance estimator.

This paper is organised as follows. Section 1 states the main assumptions concerning the generating process, recalls the definition of the estimator of the GACV based on the fixed bandwidth, introduces the GACV estimator based on the adaptive bandwidth and derives the Cramér–Rao lower bound in the two cases. The asymptotic properties of the estimator in the case of adaptive bandwidth design are also derived in Section 1. Section 2 contains the main result of this paper, establishing a necessary and sufficient condition for asymptotic efficiency and discussing its positioning in the literature. The asymptotic efficiency when series are generated by ARMA processes is discussed in Section 3, with some numerical illustrations. Section 4 provides an interpretation of the nonparametric GACV estimator as a minimum contrast estimator (Taniguchi 1987). Proofs are deferred to the Appendix.

1. Basic definitions and assumptions

Let $\{X_t\}_{t\in T}$, $T \subseteq \mathbb{Z}$, denote the linear causal process $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, $\psi_0 = 1$, $\sum_j |\psi_j| < \infty$, where $\{\epsilon_t\}_{t\in T}$ is a sequence of independent and identically distributed (i.i.d.) random variables with zero mean and variance equal to σ^2, $\mathrm{E}(\epsilon_t^s) < \infty$, $s \geq 1$. Let $\gamma_k = \mathrm{E}(X_t X_{t-k})$, $k \in \mathbb{Z}$, and $f(\omega) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \gamma_k \,\mathrm{e}^{-\imath\omega k}$, $\omega \in [-\pi, \pi]$.

Assumption 1.1: *The autocovariance function of X_t, γ_k, and its spectral density function, $f(\omega)$, can be expressed as functions of a $s \times 1$ vector of parameters $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{s-1})' \in \mathbb{R}^s$. For the spectral density function, denoted as $f_\theta(\omega)$, we assume that there exists two positive constants, $\underline{c}$ and $\bar{c}$, such that $0 < \underline{c} \leq f_\theta(\omega) \leq \bar{c} < \infty$, for $\omega \in [-\pi, \pi]$.*

For $p \in \mathbb{R}$ the generalised autocovariances, denoted as γ_{pk}, are defined as the sequence of Fourier coefficients of $[2\pi f_\theta(\omega)]^p$, i.e.

$$[2\pi f_\theta(\omega)]^p = \sum_{k=-\infty}^{\infty} \gamma_{pk} \,\mathrm{e}^{-\imath\omega k},$$

or, equivalently,

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

The boundedness of $f_\theta(\omega)$ implies that $\int_{-\pi}^{\pi} [f_\theta(\omega)]^p \, d\omega < \infty$. Obviously, $\gamma_{1k} = \gamma_k$, while $\gamma_{-1,k}$ is the inverse autocovariance function, see Cleveland (1972) and Battaglia (1983).

Throughout the paper, we make the following additional assumptions.

Assumption 1.2: *The partial derivatives of the generalised autocovariances,* $\partial\gamma_{pk}/\partial\theta_j$, $j = 0, \dots, s-1$, *satisfy the summability conditions* $\sum_{k=1}^{\infty} k|\partial\gamma_{pk}/\partial\theta_j| < \infty$.

Assumption 1.3: *The* $s \times s$ *matrix*

$$\frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}'} \frac{d\omega}{f_\theta^2(\omega)}$$

is positive definite.

Remark 1.1: Assumption 1.1 ensures the existence of a bounded parametric spectral density function, $f_\theta(\omega)$. Assumption 1.2 implies that $f_\theta(\omega)$ is differentiable with respect to θ_j, and $\partial f_\theta(\omega)/\partial\theta_j$ is continuous and differentiable with respect to ω, with continuous derivative. Assumption 1.3 ensures invertibility of the Fisher information matrix and thus existence of the Cramér–Rao lower bound (Definition 1.1). We restrict our attention to short memory processes, ruling out long memory and non-invertible models, see, e.g. Hassler (2018).

Remark 1.2: Assumptions 1.1–1.3 are related to a specific parameterisation of the spectral density function and, consequently, of the expected Fisher information matrix determining the Cramér–Rao lower bound in Definition 1.1. The parameterisation encompasses the direct representation of the spectral density function in terms of the autocovariance sequence, and that in terms of the inverse autocovariance sequence, among others, depending on the power parameter. The assumptions imply that the parameterisation is one-to-one and continuously differentiable. Alternative one-to-one continuously differentiable parameterisations have been proposed for different purposes, see, e.g. Barndorff-Nielsen and Schou (1973).

Given a time series of N consecutive observations, $\{x_t, t = 1, 2, \dots, N\}$, and their sample mean $\bar{x}_N = N^{-1}\sum_{t=1}^{N} x_t$, we define the periodogram

$$I(\omega_j) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N}(x_t - \bar{x}_N)\exp(-i\omega_j t)\right|^2,$$

where ω_j is the Fourier frequency $\omega_j = \frac{2\pi j}{N} \in (0, \pi)$, $1 \le j \le \lfloor\frac{N-1}{2}\rfloor$, and $\lfloor\cdot\rfloor$ denotes the largest integer not greater than the argument.

Based on Hannan and Nicholls (1977) and Luati et al. (2012), Proietti and Luati (2015) proposed the following nonparametric estimator of the generalised autocovariances, obtained as the inverse discrete Fourier transform of the pth power of the bias-corrected pooled periodogram,

$$\hat{\gamma}_{pk} = \frac{1}{M}\sum_{j=0}^{M-1} Y_j^{(p)} \cos(\bar{\omega}_j k), \tag{2}$$

with $M = \lfloor \frac{N-1}{2m} \rfloor$ and where

$$Y_j^{(p)} = (2\pi \bar{I}_j)^p \frac{\Gamma(m)}{\Gamma(m+p)},$$

and

$$\bar{I}_j = \sum_{l=1}^{m} I(\omega_{jm+l})$$

is the pooled periodogram over $m \geq 1$ non-overlapping consecutive frequencies, while $\bar{\omega}_j = \omega_{jm+(m+1)/2}$ is a mid-range frequency; $\Gamma(\cdot)$ is the usual Gamma function and m denotes the pooling parameter.

The asymptotic distribution of the estimator $\hat{\gamma}_{pk}$ has been derived in Proietti and Luati (2015) under the constraints embodied in the following assumption that incorporates Assumptions A1, A3, and A4 in Proietti and Luati (2015). Specifically, asymptotic normality is derived using the Bartlett's decomposition of the periodogram of a linear process and a central limit theorem for the powers of the periodogram of an i.i.d. process, based on the method of moments and on Edgeworth expansions, which require the Cramér conditions (ii)–(iii) in Assumption 1.4. Assumption 1.4(i) serves to prove the negligibility of the term involving the remainder in the Bartlett's decomposition, based on bounds in probability.

Assumption 1.4: *The process* $\{X_t\}_{t\in T}$ *is such that (i)* $\sum_{j=0}^{\infty} j^{\delta}\,|\psi_j| < \infty, \delta > \frac{3}{4}$, *(ii)* $\sup_{|s|\geq s_0} |E(\mathrm{e}^{\imath s\epsilon_t})| = \delta(s_0) < 1$ *(iii)* $\int_{-\infty}^{\infty} |E(\mathrm{e}^{\imath s\epsilon_t})|^r \, ds < \infty$ *for some* $r \geq 1$.

Let $\boldsymbol{\gamma}_p = [\gamma_{p0}, \gamma_{p1}, \ldots, \gamma_{pK}]'$ be the vector of the generalised autocovariances up to lag K and $\hat{\boldsymbol{\gamma}}_p = [\hat{\gamma}_{p0}, \hat{\gamma}_{p1}, \ldots, \hat{\gamma}_{pK}]'$ the corresponding estimator in (2). Under Assumption 1.4, it is shown in Proietti and Luati (2015) that, for m a fixed positive integer such that $m + 4(p-1) > 0$,

$$\sqrt{N}(\hat{\boldsymbol{\gamma}}_p - \boldsymbol{\gamma}_p) \underset{d}{\rightarrow} N(\mathbf{0}, \boldsymbol{V}),$$

where $\boldsymbol{V} = \{v_{kl}, k, l = 1, \ldots, K\}$, with

$$v_{kl} = m\left(C(m;p,p) - 1\right)\frac{1}{\pi}\int_{-\pi}^{\pi} [2\pi f_\theta(\omega)]^{2p}\cos(\omega k)\cos(\omega l)\,\mathrm{d}\omega + p^2\kappa_4\gamma_{pk}\gamma_{pl}, \tag{3}$$

where κ_4 is the fourth cumulant of ϵ_t and

$$C(m;p,p) = \frac{\Gamma(m+2p)\Gamma(m)}{\Gamma^2(m+p)}.$$

The following definition is discussed in p. 554 of Taniguchi and Kakizawa (2000).

Definition 1.1: Under Assumptions 1.1–1.3, the GACV estimator $\hat{\gamma}_{pk}$ in Equation (2) is asymptotically efficient, in the sense of Bahadur, if its asymptotic variance, v_{kk}, equals the *Cramér–Rao lower bound*

$$CRB\{\hat{\gamma}_{pk}\} = \frac{\partial \gamma_{pk}}{\partial \boldsymbol{\theta}'} \mathfrak{I}_N^{-1}(\boldsymbol{\theta}) \frac{\partial \gamma_{pk}}{\partial \boldsymbol{\theta}}, \tag{4}$$

with

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

and where $\mathfrak{I}(\boldsymbol{\theta})$ is the Fisher information matrix

$$\mathfrak{I}(\boldsymbol{\theta}) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}'} \frac{1}{f_\theta^2(\omega)}\, \mathrm{d}\omega.$$

In the particular case of a Gaussian process, the term κ_4 in Equation (3) is null, $\kappa_4 = 0$, and the Cramér–Rao inequality for the estimator in Equation (2) becomes

$$\begin{aligned} & m(C(m; p, p) - 1)(2\pi)^{2p} \frac{1}{\pi} \int_{-\pi}^{\pi} [f_\theta(\omega)]^{2p} \cos^2(\omega k)\, \mathrm{d}\omega \\ & \geq \left\{ (2\pi)^{p-1} \int_{-\pi}^{\pi} \frac{\partial [f_\theta(\omega)]^p}{\partial \boldsymbol{\theta}'} \cos(k\omega)\, \mathrm{d}\omega \right\} \\ & \times \left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}} \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}'} \frac{1}{f_\theta^2(\omega)}\, \mathrm{d}\omega \right\}^{-1} \left\{ (2\pi)^{p-1} \int_{-\pi}^{\pi} \frac{\partial [f_\theta(\omega)]^p}{\partial \boldsymbol{\theta}'} \cos(k\omega)\, \mathrm{d}\omega \right\}' \end{aligned} \tag{5}$$

or, equivalently,

$$\begin{aligned} & \frac{m(C(m; p, p) - 1)}{p^2} \int_{-\pi}^{\pi} [f_\theta(\omega)]^{2p} \cos^2(\omega k)\, \mathrm{d}\omega \geq \left\{ \int_{-\pi}^{\pi} [f_\theta(\omega)]^p \frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'} \cos(k\omega)\, \mathrm{d}\omega \right\} \\ & \times \left\{ \int_{-\pi}^{\pi} \frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}} \frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'}\, \mathrm{d}\omega \right\}^{-1} \left\{ \int_{-\pi}^{\pi} [f_\theta(\omega)]^p \frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'} \cos(k\omega)\, \mathrm{d}\omega \right\}'. \end{aligned}$$

We show below (Remark 2.1) that the estimator $\hat{\gamma}_{pk}$ can be efficient, under the circumstances of Theorem 2.1, only if $\kappa_4 = 0$, as in the case of a Gaussian process, and either $p = 1$, that implies that $C_{m,p} = m(C(m; p, p) - 1)/p^2 = 1$, or m large enough so that $C_{m,p}$ is close to 1. When this is not the case, under Gaussianity, $C_{m,p}$ can be interpreted as an inefficiency factor. Table 1 displays its values for different combinations of p and m: values of $C_{m,p}$ close to one are ensured when $m \geq 5$ for positive p, while it is required that $m \geq 30$, for negative values of p. In Appendix A, it is shown that $C_{m,p}$ converges to 1 if m tends to infinity.

1.1. Nonparametric GACV estimation with adaptive bandwidth

We consider now a different estimation framework, in which m increases with N, hence we write m_N, but at a slower rate, i.e. $m_N = O(N^{1-\alpha})$, $0 < \alpha < 1$. Correspondingly, the

Table 1. Values of $C_{m,p} = m[C(m,p,p) - 1]/p^2$ for combinations of values of p and m.

	$m=1$	$m=2$	$m=5$	$m=10$	$m=15$	$m=20$	$m=30$	$m=50$
$p=-2$					1.42	1.29	1.18	1.10
$p=-1$				1.25	1.15	1.11	1.07	1.04
$p=-\frac{1}{2}$				1.13	1.08	1.06	1.04	1.02
$p=\frac{1}{2}$			1.02	1.01	1.01	1.01	1.00	1.00
$p=1$	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$p=2$	1.25	1.17	1.08	1.05	1.03	1.02	1.02	1.01

number of mid-range frequencies, M_N, is $O(N^\alpha)$, if we take $M_N = \lfloor (N-1)/(2m_N) \rfloor$. A more general case is one in which $M_N = O(N^\beta)$, $0 < \beta \le \alpha$, which allows for an overlapping design. In this setting, a different large sample theory is available, which is essentially derived from Bhansali (1980). We first note that, by Stirling's approximation, $\frac{\Gamma(m_N)}{\Gamma(m_N+p)} \approxeq m_N^{-p}$ as $m_N \to \infty$, and thus $Y_j^{(p)} \approxeq (\frac{2\pi \bar{I}_j}{m_N})^p$ is interpreted as the pth power of the Daniell spectral estimator of $2\pi f(\bar{\omega}_j)$.

This leads to reformulating the GACV estimator as follows:

$$\hat{\gamma}_{pk,m_N} = \frac{1}{M_N} \sum_{j=0}^{M_N-1} Y_{j,B_N}^{(p)} \cos(\tilde{\omega}_j k), \tag{6}$$

where

$$Y_{j,B_N}^{(p)} = \left(\frac{1}{B_N} \int_{-\pi B_N}^{\pi B_N} I(\tilde{\omega}_j + \lambda)\, \mathrm{d}\lambda \right)^p, \tag{7}$$

and $I(\omega)$ is the extended periodogram, which equals $I(\omega_j)$ for $\omega \in (\omega_j - \pi/N, \omega_j + \pi/N]$, while $\tilde{\omega}_j = 2\pi j/M_N, j = 1, \dots, M_N - 1$ and $B_N = \frac{m_N}{N-1}$ is the bandwidth.

Remark 1.3: The new notation for the pth power of the Daniell estimator of the spectral density in Equation (7) is convenient in light of the new design of the estimator of the spectral density at the frequency $\tilde{\omega}_j$. We remark that the interval $(0, 2\pi]$ is divided into $M_N - 1$ subintervals, centred at $\tilde{\omega}_j = \frac{2\pi j}{M_N}, j = 1, \dots, M_N - 1$. Around these frequencies, we consider m_N periodogram ordinates falling in a band of frequencies $\tilde{\omega}_j \pm \pi B_N$, where $B_N = m_N/(N-1)$. As a consequence, the periodogram ordinates used for estimating the spectrum by the Daniell estimator at $\tilde{\omega}_j$ have a partial overlapping with those used for the neighbouring frequencies.

The large sample properties of the GACV estimator in Equation (6) are derived for the linear process $\{X_t\}_{t\in T}$ introduced so far, under an additional regularity condition on the process and an assumption on the design of the GACV estimator.

Assumption 1.5: *The autocovariance of $\{X_t\}_{t\in T}$ satisfies $\sum_{k=-\infty}^{\infty} |k|^3 |\gamma_k| < \infty$.*

Assumption 1.6: *In the GACV estimator (6) let $N \to \infty$, $M_N \to \infty$ and $B_N \to 0$ according to $M_N = O(N^\beta)$, $0 < \beta < 1$ and $B_N = O(N^{-\alpha})$, or equivalently $m_N = O(N^{1-\alpha})$. Moreover, the rates α and β satisfy $\frac{1}{4} < \alpha < 1 - \delta_0$, with $\delta_0 > 0$ arbitrarily small, and $\frac{1-\alpha}{5} < \beta < \frac{1-\alpha}{3}$.*

Theorem 1.1: *Under Assumptions 1.5 and 1.6, $\hat{\gamma}_{pk,m_N} \to_p \gamma_{pk}$, and*

$$N^{\frac{1}{2}(1-\alpha+\beta)}(\hat{\gamma}_{pk,m_N} - \gamma_{pk}) \to_d N\left(0, \frac{p^2}{\pi}\int_{-\pi}^{\pi}[2\pi f_\theta(\omega)]^{2p}\cos^2(\omega k)\,\mathrm{d}\omega\right). \tag{8}$$

Proof: See Appendix B. ■

Remark 1.4: Choosing $\alpha = \frac{1}{4} + \varepsilon$, for $\varepsilon > 0$ arbitrarily small, and $\beta = \frac{1}{4} - \varepsilon'$, $\varepsilon' > \varepsilon/3$, the rate of convergence of the asymptotic variance can be made arbitrarily close to 1. In practice, this implies that we can continue to use the pooled estimator with m_N growing at a rate slightly below $N^{0.75}$ and setting M_N equal to

$$M_N = \left\lfloor \frac{1}{2}B_N^{-1} \right\rfloor = \left\lfloor \frac{N-1}{2m_N} \right\rfloor,$$

which yields the usual GACV design. This choice would also amount to setting M_N equal to the equivalent number of independent spectral estimates, see Jenkins (1961, p. 155) and the discussion in Bhansali (1980, pp. 562–563).

In this adaptive bandwidth setting, Equation (8), gives

$$\begin{aligned}
&\frac{p^2}{\pi}(2\pi)^{2p}\int_{-\pi}^{\pi}[f_\theta(\omega)]^{2p}\cos^2(\omega k)\,\mathrm{d}\omega \\
&\geq \left\{(2\pi)^{(p-1)}p\int_{-\pi}^{\pi}[f_\theta(\omega)]^p\frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'}\cos(k\omega)\,\mathrm{d}\omega\right\} \\
&\quad\times\left\{\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}}\frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'}\,\mathrm{d}\omega\right\}^{-1} \\
&\quad\times\left\{(2\pi)^{(p-1)}p\int_{-\pi}^{\pi}[f_\theta(\omega)]^p\frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}'}\cos(k\omega)\,\mathrm{d}\omega\right\}'.
\end{aligned} \tag{9}$$

2. Asymptotic efficiency of nonparametric estimators of the GACV

The nonparametric estimator $\hat{\gamma}_{pk,m_N}$ of the generalised autocovariance function γ_{pk} is asymptotically efficient if its asymptotic variance attains the Cramér–Rao bound, that is if equality holds in the inequality (9). The following theorem provides a necessary and sufficient condition for asymptotic efficiency of $\hat{\gamma}_{pk,m_N}$ for the linear process $\{X_t\}_{\in T}$.

Theorem 2.1: *Suppose that Assumptions* 1.2, 1.3, 1.6 *are satisfied. Then,* $\hat{\gamma}_{pk,m_N}$ *defined in* (6) *is asymptotically efficient if and only if there exists an s-dimensional vector* **c**, *not depending on ω, such that it holds for all ω:*

$$[f_\theta(\omega)]^{p+1}\cos(k\omega) + \mathbf{c}'\frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}} = 0. \tag{10}$$

Proof: See Appendix C. ■

The above condition can also be stated as

$$[f_\theta(\omega)]^p \cos(k\omega) + \mathbf{c}' \frac{\partial \ln f_\theta(\omega)}{\partial \boldsymbol{\theta}} = 0.$$

The proof of Theorem 2.1 is based on a matrix integral inequality from Kakizawa and Taniguchi (1994), generalising the Cauchy–Schwarz inequality and Holevo's inequality (Kholevo 1969).

Theorem 2.1 provides a necessary and sufficient condition for asymptotic efficiency of $\hat{\gamma}_{pk,m_N}$ which is valid for general linear processes. It is expressed in terms of the spectral density function and is easy to check for various models. This result embodies in a single equation the condition for asymptotic efficiency of the sample autocovariance function ($p = 1$), of the estimator $\hat{\gamma}_{-1,k}$ of the inverse autocovariance function ($p = -1$), which at lag $k = 0$ provides the inverse of the interpolation error variance, and of the estimator $\hat{\gamma}_{pk,m_N}$ for general real powers p.

Remark 2.1: In the Gaussian case, under Assumptions 1.2, 1.3, and 1.4, the estimator $\hat{\gamma}_{pk}$ in Equation (2) is asymptotically efficient if condition (10) holds and either $p = 1$ or, with a degree of approximation summarised in Table 1, if m is sufficiently large so that $m(C(m;p,p) - 1)/p^2$ is approximately equal to 1.

Remark 2.2: The estimators (2) and (6) can be viewed in the wider context of estimation of functionals of the spectral density, which are related to many important quantities in stationary time series. Setting $m = 1$, for $p > 0$, Y_j^p is the inverse Laplace transform of $[2\pi f(\omega_j)]^{-(p+1)}$ evaluated at $2\pi I(\omega_j)$, proposed by Taniguchi (1980) for estimating $[2\pi f(\omega_j)]^p$. Asymptotic efficiency of this estimator is studied in Taniguchi (1981), who establishes that this estimator is asymptotically efficient if $p = 1$ and the spectral density is constant over $[-\pi, \pi]$. The nonparametric estimators of γ_{pk} further generalise these results to any real power transform, including negative p. As pointed out by Taniguchi (1980, p. 74), replacing the periodogram by a consistent spectral density estimator results in a loss of efficiency. However, as shown above, if the number of frequencies used in the estimation of the functional increases at a sufficiently slow rate, then the asymptotic variance is $O(N^{-1+\alpha-\beta})$, which can be made arbitrarily close to $O(N^{-1})$. Hence, the introduction of the pooling parameter m_N allows us to obtain asymptotically efficient estimates also for $p \neq 1$.

Remark 2.3: By setting the power p and the pooling parameter m to 1, inequality (5) reduces to the asymptotic Cramér–Rao inequality for the sample estimator of the autocovariance function analysed by Kakizawa and Taniguchi (1994). Note also that for $p = 1$, $C_{m,p} = 1$. Hence, if we consider estimation of the traditional autocovariance function, the asymptotic variance of the nonparametric estimator $\hat{\gamma}_{1k}$ does not depend on the pooling parameter m. Indeed, $\hat{\gamma}_{1k}$ is the Riemannian sum approximation over the Fourier frequencies of the sample autocovariance at lag k, denoted by $\tilde{\gamma}_k$,

$$\lim_{N\to\infty} \frac{1}{\lfloor (N-1)/2 \rfloor} \sum_{j=1}^{\lfloor (N-1)/2 \rfloor} 2\pi I(\omega_j) \cos(\omega_j k) = \int_{-\pi}^{\pi} I(\omega) \cos(\omega k)\, \mathrm{d}\omega = \tilde{\gamma}_k,$$

with $I(\omega) = \frac{1}{2\pi}\sum_{|h|<N}\tilde{\gamma}_h \cos(\omega h)$. Hence $\lim_{N\to\infty}\hat{\gamma}_{1k} = \tilde{\gamma}_k$, and their asymptotic variances, as $N \to \infty$, are equivalent. As a matter of fact, by setting $p = 1$, Theorem 2.1 provides the condition for asymptotic efficiency of the sample autocovariances by Kakizawa and Taniguchi (1994).

Remark 2.4: The results derived in the paper hold for fixed values of the power parameter p. The latter is either selected over a grid of possible values, by means of some criterion function, such as the deviance measure based on the Whittle likelihood advocated by Xia and Tong (2011), or is a priori fixed according to the scope of the analysis. For instance, a p-squared distance between two processes can be specified based on their power transformed spectral densities that gives the Hellinger distance ($p = 1/2$) or the quadratic distance defined by Hong (1996) ($p = 1$), see Proietti and Luati (2015). An alternative strategy, not implemented yet, would involve estimation of p based on some smoothness criterion without the need for a parametric model.

We conclude the Section with a Corollary to the main theorem that shows how the first K inverse autocovariances of a moving average process of order K can be efficiently estimated with a nonparametric method, despite the inefficiency of the sample autocovariances for pure moving average processes.

Corollary 2.1: *Consider the process with spectral density function* $f_\theta(\omega) = \frac{1}{2\pi}[\frac{1}{\theta(\omega)}]^{\frac{1}{p}}$, *with* $\theta(\omega) = \theta_0 + 2\sum_{j=1}^{K}\theta_j\cos(\omega j)$, *so that* $\frac{\partial\theta(\omega)}{\partial\boldsymbol{\theta}} = \boldsymbol{q}(\omega) = [1, 2\cos(\omega), 2\cos(2\omega), \ldots, 2\cos(\omega K)]'$. *Then,*

$$\frac{\partial f_\theta(\omega)}{\partial\boldsymbol{\theta}} = -(2\pi)^p\frac{1}{p}[f_\theta(\omega)]^{p+1}\boldsymbol{q}(\omega).$$

Condition (10) *in Theorem* 2.1 *becomes*

$$[f_\theta(\omega)]^{p+1}\left\{\cos(k\omega) - \frac{(2\pi)^p}{p}\mathbf{c}'\boldsymbol{q}(\omega)\right\} = 0,$$

which is satisfied if $\mathbf{c} = [0, 0, \ldots, \frac{p}{2(2\pi)^p}, 0, \ldots, 0]'$. *This implies that for* $p = -1$ *the process is moving-average of order* K *and the first* K *inverse autocovariances* $\boldsymbol{\gamma}_{-1,K} = [\gamma_{-1,1}, \ldots, \gamma_{-1,K}]'$ *and* $\gamma_{-1,0}$ *can be efficiently estimated as* $N \to \infty$ *by the estimator of the GACV* $\hat{\boldsymbol{\gamma}}_{-1,K,m_N}$.

3. Numerical illustrations

Some specific cases of ARMA processes and values of the power parameter are considered to illustrate how the asymptotic variance of the nonparametric estimator discussed in the paper is related to its Cramér–Rao lower bound.

Let $\{X_t\}_{t\in T}$ be a zero mean stationary ARMA(r, q) process, $\phi_r(L)X_t = \theta_q(L)\epsilon_t$, where $\phi_r(L) = 1 - \phi_1 L - \ldots - \phi_r L^r$, $\theta_q(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$, $\epsilon_t \sim$ i.i.d.$(0, \sigma^2)$

Table 2. Relative asymptotic efficiency AV/CRB of $\hat{\gamma}_{pk,m_N}$ for an AR(1) model with $\phi = 0.8$ and $\sigma^2 = 1$ for positive p and increasing values of the lag parameter k.

AV/CRB	$k = 1$	$k = 2$	$k = 4$	$k = 5$	$k = 7$
$p = 2$	1.08	1.10	1.19	1.26	1.51
$p = \frac{3}{2}$	1.03	1.05	1.15	1.25	1.59
$p = 1$	1.00	1.00	1.16	1.31	1.88

Table 3. Relative asymptotic efficiency AV/CRB of $\hat{\gamma}_{pk,m_N}$ for an MA(1) model with $\theta = -0.7, \sigma^2 = 1$.

AV/CRB	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
$p = -1$	1.00	1.06	1.24	1.60	2.24

with spectral density function

$$f_\theta(\omega) = \frac{\sigma^2}{2\pi} \frac{|\theta_q(\mathrm{e}^{-\imath\omega})|^2}{|\phi_r(\mathrm{e}^{-\imath\omega})|^2}.$$

We denote the asymptotic variance of $\hat{\gamma}_{pk,m_N}$ at $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_r, \theta_1, \ldots, \theta_q, \sigma^2)'$ by

$$AV\{\hat{\gamma}_{pk,m_N}\} = \frac{(\sigma^2)^{2p}}{\pi} p^2 \int_{-\pi}^{\pi} \left(\frac{|\theta_q(\mathrm{e}^{-\imath\omega})|^2}{|\phi_r(\mathrm{e}^{-\imath\omega})|^2} \right)^{2p} \cos^2(k\omega)\, \mathrm{d}\omega$$

and observe that, in the (most conservative) Gaussian case, the asymptotic variance of the fixed-bandwidth estimator $\hat{\gamma}_{pk}$ is equal to $C_{mp} AV\{\hat{\gamma}_{pk,m_N}\}$, where the impact of C_{mp} for some combinations of m and p is quantified in Table 1. The latter showed that, except in the case of negative integers, requiring at least $m = 30$ for $C_{m,p}$ being close to unity, in all the other cases, $m \geq 5$ ensures that $C_{m,p}$ is approximately 1.

We first consider the case of a stationary AR(1) process, with $\phi = 0.8$ and $\sigma^2 = 1$. Table 2 focuses on positive values of p. The numerical illustration aims at assessing how asymptotic efficiency deteriorates as long as the power p increases and the lag k increases. The results show that the relative efficiency of the estimator is not strongly affected by increasing p. The same is observed when the first lags are estimated. Conversely, for $p = 2$ and $k = 7$, the asymptotic variance is 1.5 times larger than its lower bound.

The same is observed in the case of a MA(1) process with $p = -1$. We recall that Corollary 2.1 establishes that in an MA(K) processes the first K inverse autocovariances are efficiently estimated by the nonparametric estimator $\hat{\gamma}_{-1k,m_N}$. Table 3 shows that inefficiency arises after lag $k = 4$, while for the first four lags, typically most relevant in empirical analysis based on MA processes, values of the ratio AV/CRB slightly greater than one are observed.

Table 4 shows that the estimation of the inverse autocovariances remains efficient for $k = 1$ and close to efficiency for $k = 2$ even for an ARMA(1,2) process, which improves the results by Porat (1987) and Walker (1995) related to sample autocovariances ($p = 1$) of Gaussian ARMA(r, q) processes, none of which was shown to be asymptotically efficient in the case when $q > r$.

Table 4. Relative asymptotic efficiency AV/CRB of $\hat{\gamma}_{pk,m_N}$ for an ARMA(1,2) model with $\theta_1 = 0.7, \theta_2 = -0.1, \phi = 0.6, \sigma^2 = 1$.

AV/CRB	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
$p = -1$	1.00	1.14	2.44	2.72	3.10

4. Minimum contrast estimation

As in Corollary 2.1, let us consider the process with spectral density function

$$[2\pi f_\theta(\omega)]^p = [\theta(\omega)]^{-1}, \tag{11}$$

where $\theta(\omega) > 0$ is the trigonometric polynomial $\theta_0 + 2\sum_{k=1}^{K} \theta_k \cos(\omega k)$. Writing $\theta(\omega) = \theta_0 |\phi(\mathrm{e}^{-\imath\omega})|^2$, $\phi(\mathrm{e}^{-\imath\omega}) = 1 - \sum_{j=1}^{K} \phi_j \,\mathrm{e}^{-\imath\omega j}$, such that $\theta_k = \theta_0 \sum_{j=1}^{K-k} \phi_j \phi_{j+k}$, and setting $\sigma^2 = \theta_0^{-1}$, it can be seen, by integrating both sides of (11) over $\omega \in [-\pi, \pi]$, that γ_{pk} is the autocovariance function of the AR(K) process $U_t = \sum_{j=1}^{K} \phi_j U_{t-j} + \sigma\epsilon_t, \epsilon_t \sim$ i.i.d. N(0, 1).

Following Taniguchi (1987), let us consider minimum contrast (MC) estimation of the spectral density $f_\theta(\omega)$ using the contrast functional

$$K(z; p) = \ln(z^p) + \frac{1}{z^p},$$

applied to $f_\theta(\omega)/g_N(\omega)$, where, defining

$$Y(\omega) = \frac{1}{2\pi} \sum_{-M+1}^{M-1} \hat{\gamma}_{pk}\, \mathrm{e}^{-\imath\omega k}, \quad \omega \in [-\pi, \pi],$$

so that $\hat{\gamma}_{pk} = \int_{-\pi}^{\pi} Y(\omega)\, \mathrm{e}^{\imath\omega k}\, \mathrm{d}\omega$, we have set $g_N(\omega) = [Y(\omega)]^{1/p}$.

The MC estimator of $(\phi_1, \ldots, \phi_K, \sigma^2)'$ is the minimiser of

$$\begin{aligned} &\int_{-\pi}^{\pi} K\left(\frac{f_\theta(\omega)}{g_N(\omega)}, p\right) \mathrm{d}\omega \\ &\quad = \int_{-\pi}^{\pi} \left\{ \ln \sigma^2 - \ln |\phi(\mathrm{e}^{-\imath\omega})|^2 - \ln Y(\omega) + \frac{1}{\sigma^2} Y(\omega) |\phi(\mathrm{e}^{-\imath\omega})|^2 \right\} \mathrm{d}\omega. \end{aligned}$$

The MC estimator of σ^2 is $\hat{\sigma}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(\omega) |\hat{\phi}(\mathrm{e}^{-\imath\omega})|^2 \,\mathrm{d}\omega$. Plugging this into the contrast function (and noting $\int_{-\pi}^{\pi} |\hat{\phi}(\mathrm{e}^{-\imath\omega})|^2 \,\mathrm{d}\omega = 0$), the MC estimator of $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_s)'$ is the minimiser of the criterion function

$$Q(\boldsymbol{\phi}) = \int_{-\pi}^{\pi} Y(\omega) |\phi(\mathrm{e}^{-\imath\omega})|^2 d\omega.$$

Writing

$$|\phi(\mathrm{e}^{-\imath\omega})|^2 = 1 - 2\boldsymbol{\phi}' \mathbf{b}(\omega) + \boldsymbol{\phi}' \mathbf{B}(\omega) \boldsymbol{\phi}$$

where $\mathbf{b}(\omega) = [\cos\omega, \cos(2\omega), \ldots, \cos(\omega K)]'$ and $\mathbf{B}(\omega) = \{\cos(\omega(h-k)), h, k = 1, 2, \ldots, s\}$, differentiating with respect to $\boldsymbol{\phi}$ and setting the derivatives equal to zero yields

$$\frac{\partial Q}{\partial \boldsymbol{\phi}} = \int_{-\pi}^{\pi} Y(\omega)(\mathbf{b}(\omega) - \mathbf{B}(\omega)\boldsymbol{\phi})\,\mathrm{d}\omega \equiv 0,$$

which is the generalised Yule–Walker system of equations:

$$\hat{\gamma}_{pk} = \sum_{j=1}^{K} \hat{\phi}_j \hat{\gamma}_{p,k-j}, k = 1, 2, \ldots, K.$$

Hence, an asymptotically efficient estimator of $(\boldsymbol{\phi}, \sigma^2)$, and thus of $\boldsymbol{\theta}$, can be obtained by solving a generalised Yule–Walker system based on the GACV estimator (2).

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Anderson, T.W. (1994), *The Statistical Analysis of Time Series*. New York, NY: John Wiley & Sons.

Barndorff-Nielsen, O., and Schou, G. (1973), 'On the Parametrization of Autoregressive Models by Partial Autocorrelations', *Journal of Multivariate Analysis*, 3(4), 408–419.

Battaglia, F. (1983), 'Inverse Autocovariances and a Measure of Linear Determinism for a Stationary Process', *Journal of Time Series Analysis*, 4(2), 79–87.

Battaglia, F. (1988), 'On the Estimation of the Inverse Correlation Function', *Journal of Time Series Analysis*, 9(1), 1–10.

Bhansali, R. (1980), 'Autoregressive and Window Estimates of the Inverse Correlation Function', *Biometrika*, 67(3), 551–566.

Boshnakov, G.N. (2005), 'On the Asymptotic Properties of Multivariate Sample Autocovariances', *Journal of Multivariate Analysis*, 92(1), 42–52.

Brockwell, P.J., and Davis, R.A. (1991), *Time Series: Theory and Methods*, Springer Series in Statistics.

Chen, W.W., and Deo, R.S. (2004), 'A Generalized Portmanteau Goodness-of-fit Test for Time Series Models', *Econometric Theory*, 20(2), 382–416.

Cleveland, W.S. (1972), 'The Inverse Autocorrelations of a Time Series and Their Applications', *Technometrics*, 14(2), 277–293.

Erdélyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. (1954), 'Higher Transcendental Functions', *Bulletin of the American Mathematical Society*, 60, 4405–409.

Hannan, E. (1970), *Multiple Time Series*. New York, NY: John Wiley & Sons.

Hannan, E., and Nicholls, D. (1977), 'The Estimation of the Prediction Error Variance', *Journal of the American Statistical Association*, 72(360a), 834–840.

Hassler, U. (2018), *Time Series Analysis with Long Memory in View*. Hoboken, NJ: John Wiley & Sons.

Hong, Y. (1996), 'Consistent Testing for Serial Correlation of Unknown Form', *Econometrica: Journal of the Econometric Society*, 64(4), 837–864.

Jenkins, G.M. (1961), 'General Considerations in the Analysis of Spectra', *Technometrics*, 3(2), 133–166.

Kakizawa, Y. (1999), 'Note on the Asymptotic Efficiency of Sample Covariances in Gaussian Vector Stationary Processes', *Journal of Time Series Analysis*, 20(5), 551–558.

Kakizawa, Y., and Taniguchi, M. (1994), 'Asymptotic Efficiency of the Sample Covariances in a Gaussian Stationary Process', *Journal of Time Series Analysis*, 15(3), 303–311.

Kholevo, A.S. (1969), 'On Estimates of Regression Coefficients', *Theory of Probability & Its Applications*, 14(1), 79–104.

Luati, A., Proietti, T., and Reale, M. (2012), 'The Variance Profile', *Journal of the American Statistical Association*, 107(498), 607–621.
Porat, B. (1987), 'Some Asymptotic Properties of the Sample Covariances of Gaussian Autoregressive Moving-Average Processes', *Journal of Time Series Analysis*, 8(2), 205–220.
Priestley, M.B. (1981), *Spectral Analysis and Time Series*. London: Academic Press Inc.,
Proietti, T., and Luati, A. (2015), 'The Generalised Autocovariance Function', *Journal of Econometrics*, 186(1), 245–257.
Taniguchi, M. (1980), 'On Estimation of the Integrals of Certain Functions of Spectral Density', *Journal of Applied Probability*, 17(1), 73–83.
Taniguchi, M. (1981), 'An Estimation Procedure of Parameters of a Certain Spectral Density Model', *Journal of the Royal Statistical Society: Series B (Methodological)*, 43(1), 34–40.
Taniguchi, M. (1987), 'Minimum Contrast Estimation for Spectral Densities of Stationary Processes', *Journal of the Royal Statistical Society: Series B (Methodological)*, 49(3), 315–325.
Taniguchi, M., and Kakizawa, Y. (2000), *Asymptotic Theory of Statistical Inference for Time Series*. New York, NY: Springer-Verlag.
Walker, A.M. (1995), 'On Results of Porat Concerning Asymptotic Efficiency of Sample Covariances of Gaussian ARMA Processes', *Journal of Time Series Analysis*, 16(2), 237–248.
Wu, W.B. (2011), 'Asymptotic Theory for Stationary Processes', *Statistics and Its Interface*, 4(2), 207–226.
Xia, Y., and Tong, H. (2011), 'Feature Matching in Time Series Modeling', *Statistical Science*, 26(1), 21–46.

Appendices

Appendix 1. Asymptotic behaviour of the factor $C_{m,p}$

The term $C_{m,p} = m(C(m; p, p) - 1)/p^2$ converges to 1 as $m \to \infty$. We start by using a result about the approximation of a quotient of two Gamma functions, obtained by the use of the Stirling's series, see Erdélyi, Magnus, Oberhettinger, and Tricomi (1954):

$$\frac{\Gamma(z+\alpha)}{\Gamma(z+\beta)} = z^{\alpha-\beta}\left[1 + \frac{(\alpha-\beta)(\alpha+\beta-1)}{2z} + O(|z|^{-2})\right],$$

as $z \to \infty$, where α and β are bounded. By using this approximation, we can rewrite

$$\begin{aligned}\frac{\Gamma(m)\Gamma(m+2p)}{[\Gamma(m+p)][\Gamma(m+p)]} &\approx m^{-p}\left[1 + \frac{(-p)(p-1)}{2m}\right] m^{p}\left[1 + \frac{p(3p-1)}{2m}\right] \\ &= \frac{4m^2 + 4mp^2 - 3p^4 + 4p^3 - p^2}{4m^2},\end{aligned}$$

By a change of variable and the De L'Hôpital theorem, it is straightforward to show that $m(C(m; p, p) - 1) \to p^2$ as $m \to \infty$.

Appendix 2. Proof of Theorem 1.1

Let us denote $\hat{f}(\tilde{\omega}_j) = \frac{1}{2\pi B_N}\int_{-\pi B_N}^{\pi B_n} I(\tilde{\omega}_j + \lambda)\, \mathrm{d}\lambda$, the Daniell estimator of $f(\tilde{\omega}_j)$. Then, writing $Y_{j,B_N}^{(p)} = (2\pi \hat{f}(\tilde{\omega}_j))^p$, and letting $N \to \infty$, $M_N \to \infty$ and $B_N \to 0$, $NB_N \to \infty$, as in the statement of Theorem 4.3 of Bhansali (1980), so that $\nu_j = \lim_{N\to\infty} \tilde{\omega}_j$, we have that $Y_{j,B_N}^{(p)}$ is an estimator of $[2\pi f(\nu_j)]^p$. Since the Daniell kernel satisfies Assumption 1 of Bhansali (1980), with characteristic exponent equal to 2, under Assumption 1.6 for the linear process $\{X_t\}_{t\in T}$, it holds that $\hat{f}(\nu)$ converges in probability to $f(\nu)$ uniformly in ν (Bhansali 1980, Lemma 4.1).

By a Taylor series expansions of $[2\pi\hat{f}(\tilde{\omega}_j)]^p$ about $[2\pi f(\tilde{\omega}_j)]^p$,

$$\begin{aligned}\hat{\gamma}_{pk,m_N} - \gamma_{pk} &= \frac{1}{M_N}\sum_{j=0}^{M_N-1}[2\pi f(\tilde{\omega}_j)]^p\cos(\tilde{\omega}_j k) - \gamma_{pk} \\ &\quad + \frac{2\pi p}{M_N}\sum_{j=0}^{M_N-1}[2\pi f(\tilde{\omega}_j)]^{p-1}\left\{\hat{f}(\tilde{\omega}_j) - f(\tilde{\omega}_j)\right\}\cos(\tilde{\omega}_j k) + O_P(N^{-1}B_N^{-1}) \\ &\leq C|\gamma_{p-1,k}|\max_{\omega}|\hat{f}(\omega) - f(\omega)|,\end{aligned}$$

where C is a positive constant. This follows from the convergence of the Riemann sum $\frac{1}{M_N}\sum_{j=0}^{M_N-1}[2\pi f(\tilde{\omega}_j)]^p\cos(\tilde{\omega}_j k)$ to γ_{pk}, while $\frac{1}{M_N}\sum_{j=0}^{M_N-1}[2\pi f(\tilde{\omega}_j)]^{p-1}\cos(\tilde{\omega}_j k) \to \gamma_{p-1,k}$ as $M_N \to \infty$. By uniform consistency, $P(\max_{\omega}|\hat{f}(\omega) - f(\omega)] < \varepsilon) \to 1$, for all $\varepsilon > 0$ as $N \to \infty$, it follows that $\hat{\gamma}_{pk,m_N} \to_P \gamma_{pk}$.

In light of the Taylor series expansion, the asymptotic normality of $\hat{\gamma}_{pk,m_N}$ is related to that of $\hat{f}(\tilde{\omega}_j)$. The Daniell estimator, $\hat{f}(\omega)$, is evaluated at the frequencies $\tilde{\omega}_j = \frac{2\pi j}{M_N}, j = 1, \ldots, M_N - 1$, where $j \to \infty$ as $N \to \infty$. Assume that $\tilde{\omega}_j$ converges to ν_j as both j and M_N grow. Then, under Assumptions 1.5–1.6, by Theorem 9, page 280, and Theorem 11, p. 289, in Hannan (1970), see also Priestley (1981), Section 6.2.4,

$$\sqrt{NB_N}\left\{\hat{f}(\nu_j) - f(\nu_j)\right\} \to_d \mathrm{N}\left(0, f^2(\nu_j)\right),$$

if $\nu_j \neq (0, \pi)$, whereas the asymptotic variance is doubled if $\nu_j = 0$, or $\nu_j = \pi$. In particular,

$$\mathrm{E}\{\hat{f}(\nu_j) - f(\nu_j)\} = -\frac{\pi^2}{6}B_N^2 f''(\nu_j) + o(1),$$

where $f''(\omega) = d^2 f(\omega)/(\mathrm{d}\omega)^2$. Second,

$$\begin{aligned}N\mathrm{Var}(\hat{f}(\tilde{\omega}_j)) &= \kappa_4\left[\int_{-\pi}^{\pi} f^2(\lambda)W_N(\tilde{\omega}_j - \lambda)\,\mathrm{d}\lambda\right]^2 \\ &\quad + 2\pi\int_{-\pi}^{\pi} f^2(\lambda)W_N(\tilde{\omega}_j - \lambda)\{W_N(\tilde{\omega}_j - \lambda) + W_N(\tilde{\omega}_j + \lambda)\}\,\mathrm{d}\lambda + o(1),\end{aligned}$$

where the first addend is $O(1)$, while the second is $O(B_N^{-1})$, and, when multiplied by B_N, converges to $f^2(\nu_j)$.

Finally, $\hat{f}(\nu_j)$ and $\hat{f}(\nu_k)$ are asymptotically independent. The distance between $\tilde{\omega}_j$ and $\tilde{\omega}_k$ is (at least) $2\pi M_N^{-1}$. To ensure that the corresponding spectral estimates are asymptotically uncorrelated we need that M_N diverges more slowly than B_N^{-1}, as $N \to \infty$. This is indeed the case, as we assumed that $B_N^{-1} = O(N^{\alpha})$ and $M_N = O(N^{\beta})$, with $0 < \beta \leq \alpha$. Thus, if we suppose that as $N \to \infty$ the frequencies $\tilde{\omega}_j$ and $\tilde{\omega}_k$ converge respectively to ν_j and ν_k, as $(j, k, M_N, B_N^{-1}) \to \infty$ with $N \to \infty$, then by Theorem 4.3 in Bhansali (1980),

$$\mathrm{Cov}(\hat{f}(\tilde{\omega}_j), \hat{f}(\tilde{\omega}_k)) = O(N^{-1}B_N^{-1}).$$

Hence, $\sqrt{NB_N M_N}\hat{\gamma}_{pk,m_N} - \gamma_{pk}$ is asymptotically normal with mean zero and variance

$$\lim_{N\to\infty}\left\{NB_N M_N \mathrm{Var}\left(\hat{\gamma}_{pk,m_N} - \gamma_{pk}\right)\right\} = \frac{p^2}{\pi}\int_{-\pi}^{\pi}[2\pi f_{\theta}(\omega)]^{2p}\cos^2(\omega k)\,\mathrm{d}\omega.$$

The conditions (i)–(vi) of Theorem 4.4 in Bhansali (1980) then become: (i) $\alpha + 3\beta < 1 - \delta_0$, for $\delta_0 > 0$ arbitrarily small; (ii) $5\alpha > 1 + \beta$; (iii) $\alpha + 5\beta > 1$; (iv) $\alpha > \beta$; (v) $\alpha > \beta\delta_0/(1 - \delta_0)$; (vi) is implied by (i). By choosing α and β according to Assumption 1.6, all these conditions are satisfied. Hence, result (8) follows. □

Appendix 3. Proof of Theorem 2.1

We first recall the following Lemma by Kakizawa and Taniguchi (1994).

Lemma A.1: *Let* $A(\omega)$ *and* $B(\omega)$ *be* $r \times s, t \times s$ *matrices, respectively, and let* $g(\omega)$ *be a function such that* $g(\omega) > 0$ *almost everywhere on* $[-\pi, \pi]$. *If the matrix*

$$\left\{ \int_{-\pi}^{\pi} \frac{B(\omega)B(\omega)'}{g(\omega)} \, \mathrm{d}\omega \right\}^{-1}$$

exists, then

$$\int_{-\pi}^{\pi} A(\omega)A(\omega)'g(\omega) \, \mathrm{d}\omega \geq \left\{ \int_{-\pi}^{\pi} A(\omega)B(\omega)' \, \mathrm{d}\omega \right\} \left\{ \int_{-\pi}^{\pi} \frac{B(\omega)B(\omega)'}{g(\omega)} \, \mathrm{d}\omega \right\}^{-1} \left\{ \int_{-\pi}^{\pi} A(\omega)B(\omega)' \, \mathrm{d}\omega \right\}'$$

where $\geq$ *means the left-hand side minus the right-hand side results in a positive semi-definite matrix. Equality holds if there exists an* $r \times t$ *matrix* C *which is independent of* ω *such that:*

$$g(\omega)A(\omega) + CB(\omega) = 0.$$

We are now in the position of proving Theorem 2.1.

The Cramér–Rao inequality (9) can be written as in Equation (9). Setting

$$A(\omega) = \cos(k\omega)[f_\theta(\omega)]^{p-1}, \quad B(\omega) = \frac{\partial f_\theta(\omega)}{\partial \boldsymbol{\theta}}, \quad g(\omega) = f_\theta^2(\omega)$$

and applying Lemma A.1, then existence of a vector $\mathbf{c}$ as in Equation (10) ensures that the Cramér–Rao lower bound is attained. □