Computational Statistics and Data Analysis 56 (2012) 15–24

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

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

A periodic Levinson–Durbin algorithm for entropy maximization

Georgi N. Boshnakov [a], Sophie Lambert-Lacroix [b,*]

[a] School of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

[b] UJF-Grenoble 1 / CNRS / UPMF / TIMC-IMAG, UMR 5525, Grenoble, F-38041, France

ARTICLE INFO

Article history:
Received 18 November 2009
Received in revised form 30 April 2011
Accepted 3 July 2011
Available online 14 July 2011

Keywords:
Maximum entropy method
Partial autocorrelation
Periodically correlated processes
Periodic Levinson–Durbin algorithm

ABSTRACT

A recursive algorithm is presented for the computation of the first-order and second-order derivatives of the entropy of a periodic autoregressive process with respect to the autocovariances. It is an extension of the periodic Levinson–Durbin algorithm. The algorithm has been developed for use at one of the steps of an entropy maximization method developed by the authors. Numerical examples of entropy maximization by that method are given. An implementation of the algorithm is available as an R package.

1. Introduction

The class of periodically correlated processes (pc processes) introduced by Gladishev (1961) is useful in many applications; see Hurd and Miamee (2007) for a thorough exposition of the theory, Franses and Paap (2004) for economic applications, Serpedin et al. (2005) for a comprehensive bibliography, and Hindrayanto et al. (2010) for state space modelling.

The maximum entropy principle provides an appealing framework for the specification of complete models from partial information. It was introduced to stationary time series by Burg in the influential works (Burg, 1975, 1967). Given a contiguous set of autocovariances for lags $0, \ldots, p$, the maximum entropy solution is an autoregressive process of order p with those autocovariances. In this case, the problem is linear, and the solution can be obtained by solving the Yule–Walker equations with the Levinson–Durbin algorithm. This result holds in the multivariate case as well. A generalization to pc processes has been obtained by Lambert-Lacroix (2005). Deep results on this and related problems have been obtained by Alpay et al. (2001) and Castro and Girardin (2002). When the lags are not contiguous, the problem is, in general, non-linear, but the solution is still an autoregression of order equal to the maximum specified lag. For univariate stationary processes, the case of non-contiguous lags has been studied by Politis (1992) and Rozario and Papoulis (1987). A method for the solution of the maximum entropy problem for pc processes in the case of general gap patterns has been developed in Boshnakov and Lambert-Lacroix (2009).

The entropy rate is a very complicated function of the autocovariances. It is hardly possible to write down useful expressions for it and its derivatives with respect to the non-specified autocovariances for general gap patterns. The periodic Levinson–Durbin algorithm (see Sakai (1982) or Lambert-Lacroix (2005)) can be used to calculate the entropy rate. For gradient and Newton-type maximization methods, derivatives are also needed. In this paper, we develop recursions for the first-order and second-order derivatives of the entropy rate. We give also numerical examples that illustrate the behaviour of

* Corresponding author.
E-mail address: Sophie.Lambert@imag.fr (S. Lambert-Lacroix).

0167-9473/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2011.07.001

our method. The R programs implementing the algorithm presented here and the maximum entropy method of Boshnakov and Lambert-Lacroix (2009) are available as the R package `pcme` (Boshnakov and Lambert-Lacroix, 2009).

The paper is organized as follows. Section 2 presents some basic results about the entropy of pc processes. Section 3 gives the algorithm for the calculation of the gradient and the Hessian of the entropy. Numerical results illustrating the maximization of the entropy are presented in Section 4. Positive semi-definite (p.s.d.) solutions are discussed in Section 5.

2. Maximum entropy for periodically correlated processes

Let $\mathbb{N}$ be the set of the non-negative integers. A zero-mean process $\{X_t, t \in \mathbb{N} \setminus \{0\}\}$ is periodically correlated of period T if its autocovariance function $R(u, v) = E\{X_u \overline{X_v}\}$ is T-periodic, i.e.

$$R(u + T, v + T) = R(u, v), \quad \text{for all } (u, v) \in \mathbb{N}^2, \tag{1}$$

(see Gladishev (1961), Hipel and McLeod (1994), and Hurd and Miamee (2007)). It is convenient to think about the autocovariances in terms of the *seasons* $t = 1, \ldots, T$ and the *lags* $k \in \mathbb{N}$. Each pair $(u, v) \in \mathbb{N}^2$ may be represented as $(u, v) = (mT + t, mT + t - k)$ for some $t \in \{1, \ldots, T\}$, $m \in \mathbb{N}$, and integer k. From Eq. (1), it follows that $R(mT + t, mT + t - k)$ does not depend on m. So, we may introduce the notation

$$R_t(k) = R(mT + t, mT + t - k), \quad t \in \{1, \ldots, T\},\ m \in \mathbb{N},\ k\text{—integer}.$$

Moreover, it is sufficient to consider $R_t(k)$ for $k \geq 0$. Indeed, if $u - v = k < 0$, i.e. $v > u$, then $(v, u) = (m_1T + s, m_1T + s - |k|)$ for some $s \in \{1, \ldots, T\}$, and $R(u, v)$ can be obtained from the identity $R(u, v) = \overline{R(v, u)} = \overline{R_s(|k|)}$. Similar notation is used by other authors; see Hipel and McLeod (1994). To illustrate this notation, consider a monthly pc process started in January 2000. Let $u = 13$ (January 2001) and $v = 11$ (November 2000). Here, the period is $T = 12$ months, $(u, v) = (1 \times 12 + 1, 1 \times 12 + 1 - 2)$, and $(v, u) = (0 \times 12 + 11, 0 \times 12 + 11 - (-2))$. So, $R(u, v) = R_1(2)$ and $R(v, u) = R_{11}(-2)$. On the other hand, $R_{11}(-2) = R(v, u) = \overline{R(u, v)} = \overline{R_1(2)}$.

If t is one of the seasons, $1, \ldots, T$, and k is a non-negative integer lag, then (t, k) will be called a *season–lag pair*. The T functions $R_1(\cdot), \ldots, R_T(\cdot)$, considered as functions on $\mathbb{N}$, completely parameterize the second-order structure of the pc process in the sense that for each (u, v) there is exactly one season–lag pair (t, k) such that $R(u, v) = R_t(k)$ (if $u \geq v$) or $R(u, v) = \overline{R_t(k)}$ (if $u < v$). In other words, the doubly indexed sequence $\{R_t(k)\}$, $t \in \{1, \ldots, T\}$, $k \in \mathbb{N}$, enumerates the autocovariances in a non-redundant way. An equivalent parameterization is given by the partial autocorrelations (pacfs) $\{\beta_t(k)\}$, $t = 1, \ldots, T$, $k \in \mathbb{N}$ (see Lambert-Lacroix (2005) for details).

Let I be a set of season–lag pairs and $K = \{R_t(k)\}_{(t,k)\in I}$ be a sequence defined on I. Let Γ be the set of all periodic autocovariance sequences whose values coincide with $R_t(k)$ for $(t, k) \in I$ (Γ may be empty). Each element of Γ is a completion (or extension) of K. The maximum entropy extension is the one whose entropy rate is maximal in Γ. We refer to the problem of finding the maximum entropy extension of a sequence K defined on a set of season–lag pairs I as the ME(K, I) problem. A method for the solution of the ME(K, I) problem for arbitrary patterns of the set I has been developed in Boshnakov and Lambert-Lacroix (2009). The method involves maximization of the entropy on season–lags sets of the form

$$E_c(I) = \{(t, k) | t = 1, \ldots, T,\ k = 0, \ldots, p_t\},$$

where $(p_1, \ldots, p_T)$ are the smallest non-negative integers satisfying the constraints $p_1 \leq p_T + 1$ and $p_t \leq p_{t-1} + 1$ for $t = 2, \ldots, T$, and such that $E_c(I) \supseteq I$. We refer to the elements of $E_c(I) \setminus I$ as *gaps* since these are season–lag pairs with non-specified values in K.

To give the entropy rate definition we need additional notation. Let $\{X_t\}$ be a pc process, and let $v_t(k)$ be the variance of the prediction error of X_t in terms of the k previous values $X_{t-1}, \ldots, X_{t-k}$. Then for any given $t \in \{1, \ldots, T\}$ the sequence $\{v_{mT+t}(mT + t - 1)\}_{m=1}^{\infty}$ is convergent as $m \to \infty$, since it is monotonically decreasing and bounded from below by 0. Let

$$\sigma_t^2 = \lim_{m\to\infty} v_{mT+t}(mT + t - 1), \quad t = 1, \ldots, T. \tag{2}$$

An expression for σ_t^2 in terms of the partial autocorrelations is (see Lambert-Lacroix (2005))

$$\sigma_t^2 = R_t(0) \prod_{n=1}^{\infty} (1 - \|\beta_t(n)\|^2), \quad t = 1, \ldots, T.$$

It can be shown (Iqelan, 2007, p. 119) that for a Gaussian not locally deterministic pc process X the entropy rate is equal to

$$h(X) = \frac{1}{2} \log(2\pi e) + \frac{1}{2T} \sum_{t=1}^{T} \log \sigma_t^2,$$

where $\sigma_t^2 > 0$ for $t = 1, \ldots, T$. Since we are considering only second-order properties and the first term is a constant, we can define the entropy rate of a pc process with autocovariance sequence R by

$$h(R) = \frac{1}{T} \sum_{t=1}^{T} \log \sigma_t^2. \tag{3}$$