INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mathématique

Amine Amimour, Karima Belaide and Ouagnina Hili

Minimum Hellinger distance estimates for a periodically time-varying long memory parameter

Volume 360 (2022), p. 1153-1162

https://doi.org/10.5802/crmath.381

Les Comptes Rendus. Mathématique *sont membres du Centre Mersenne pour l'édition scientifique ouverte*
www.centre-mersenne.org
e-ISSN : 1778-3569

Comptes Rendus
Mathématique
2022, Vol. 360, p. 1153-1162
https://doi.org/10.5802/crmath.381

Probability theory, Statistics / *Probabilités, Statistiques*

Minimum Hellinger distance estimates for a periodically time-varying long memory parameter

Amine Amimour*, [a], Karima Belaide [a] and Ouagnina Hili [b]

[a] Department of Mathematics, Applied Mathematics Laboratory, University of Bejaia, Bejaia Algeria

[b] Laboratory of Mathematics and New Technologies of Information, National Polytechnic Institute Felix HOUPHOUET-BOIGNY, Yamoussoukro, P.O. Box 1093, Ivory Coast

E-mails: amineamimour@gmail.com, k_tim2002@yahoo.fr, o_hili@yahoo.fr

Abstract. We consider a purely fractionally deferenced process driven by a periodically time-varying long memory parameter. We will build an estimate for the vector parameters using the minimum Hellinger distance estimation. The results are investigated through simulation studies.

2020 Mathematics Subject Classification. 37M10, 62M10, 91B84.

Manuscript received 24 November 2020, accepted 19 May 2022.

1. Introduction

One of the major prominent periodic invertible and causal process in time series analysis is the PARMA model (see [23]) which generalizes the ARMA model (see [6]). $(X_t, t \in \mathbb{Z})$, is said to be PARMA model if it satisfies the difference equation

$$\sum_{j=0}^{P} \phi_{i,j} X_{i+pm-j} = \sum_{k=0}^{Q} \theta_{i,j} \varepsilon_{i+pm-k}, m \in \mathbb{Z}, \tag{1}$$

where, for each season i ($i = 1, \ldots, p$), where p is the period, P and Q are the AR and MA orders respectively, and the coefficients satisfy $\phi_{t+p,j} = \phi_{t,j}$ for $j = 1, ..P$ and $\theta_{t+p,k} = \theta_{t,k}$ for $k = 1, \ldots, Q$. The sequence $(\varepsilon_t)_{t\in\mathbb{Z}}$ is zero-mean and uncorrelated with finite variance σ_t^2, the variance is periodic in t such that $\sigma^2_{t+pm} = \sigma^2_t$. Troutman [22] considered a periodic autoregressive process which is defined in (1) with null moving average order and inspected divers properties of this model by considering the associated stationary multivariate autoregressive process. Bentarzi

---

* Corresponding author.

and Hallin [4] investigated this idea. They consider the random variables in the periodic non stationary process as the elements of a multivariate stationary process which is an interest of periodic models, in addition to allowing the study of seasonal phenomena. It is that they can be exploited in the context of the analysis of the stationary multidimensional time series, to cut substantially the number of the parameters to estimate, especially for multivariate autoregressive processes (see [8,19] and [17]). However, the zero mean purely fractional AFRIMA$(0,d,0)$ process $(X_t, t \in \mathbb{Z})$ represented by

$$(1-B)^d X_t = \varepsilon_t, \tag{2}$$

is also an extension of the ARMA model, where B denotes the backward shift operator and d can take any real number. Establishing a relationship between the fractional integration and long memory, this model goes back to Hosking [11], showing that the long memory process is invertible for $d > -\frac{1}{2}$ and stationary for $d < \frac{1}{2}$. Odaki [18] noticed that this process is invertible even when $-1 < d \leq -\frac{1}{2}$, these conditions concerning the univariate case. Ching-fan [7] used a VARFIMA model to define the invertibility condition of multivariate p-dimensional stationary process in the sense of Hosking [11]. In this article, we frame the model (2) with periodically time varying memory parameter d with period p by accommodating a long memory stationary model for the p-variate process. It is an interesting topic due to the importance in one hand of the ARFIMA models and on the other hand the periodic phenomenon (see [1,2]). Furthermore, the model of Hosking has found its potential in long-term forecasting, and so it has turn into one of the basic famous parametric long memory models in the statistical literature. For this model, the parametrical estimation of the memory parameter has been widely used as in Yajima [25] determined the estimation of d and σ^2 for the model defined in (2), using two methods, the least squares estimates and the maximum likelihood estimate by calculating the spectral density of the sequences ε_t and by the density function of the maximum likelihood estimator respectively and assuming that this is a leading note in identification and estimation procedure of a general ARFIMA (p,d,q). Gupta [9] proposed a regression method for estimating the d factor in this general model and proved that this estimator have mean square consistency and compared its performances with some known results such as Yajima [25] and others, concluding that this method ables to estimate d and the ARMA parameters, Sowell [21] considered this general model and derived the unconditional exact likelihood function, assuming that this method allows the simultaneous estimation of all the parameters of the model by exact maximum likelihood. Mayoral [15] used a minimum distance by a new method for estimating the parameters of stationary and non-stationary ARFIMA(p,d_0,q) process for $d_0 < 0,75$. The quasi-maximum likelihood approach for a non-stationary multivariate ARFIMA process is derived by Kamagate and Hili [13]. Kamagate and Hili [12] determined the minimum Hellinger distance estimate (MHDE) of a general ARFIMA model. Mbeke and Hili [16] extended this work to the multivariate case. They constructed an estimate for a vector parameters, so these known results may be used to infer the desired property of the model considered here. Although we deal only with the characteristics of the purely fractionally differenced process defined in (2) with periodic long memory parameter. Indeed, Amimour and Belaide [3] have proved recently that this model has a local asymptotic normality property. The cause of selecting the MHD method is that its estimate own exquisite robustness properties of the MHD estimators. It has been pioneered by Beran [5] for independent and identically distributed observations, Hili [10] extended these results to the case of dependent observations of non-linear time series.

The outline of our paper is as follows. Section 2 describes the model, provides some required assumptions and carry out the indispensable propositions. Section 3 is devoted to state the main results. To see how the MHD method applies, we conduct some simulations study in Section 4.

2. Notation and basic assumptions

2.1. *Definition and notation*

We shall consider the periodic autoregressive fractionally integrated moving average processes $(X_t, t \in \mathbb{Z})$ with period p, denoted here by PtvARFIMA$_p$ $(0, d_t, 0)$, which are proposed by [3]. The model is given by

$$(1-B)^{d_t} X_t = \varepsilon_t \iff (1-B)^{d_i} X_{i+pm} = \varepsilon_{i+pm}, \tag{3}$$

where for all $t \in \mathbb{Z}$, there exists $i = \{1, \ldots, p\}$, $m \in \mathbb{Z}$, such that $t = i + pm$, d_i is the long memory parameter which varies over time, whose values lie in $(0, \frac{1}{2})$, and $(\varepsilon_t, t \in \mathbb{Z})$ is a zero mean white noise with finite variance σ_t^2, the variance is periodic in t such that $\sigma_{t+pm}^2 = \sigma_t^2$.

When $d_i > 0$. The process (3) is invertible and has an infinite autoregressive representation as follows:

$$\varepsilon_{i+pm} = (1-B)^{d_i} X_{i+pm} = \sum_{j=0}^{\infty} \pi_j^i X_{i+pm-j}, \tag{4}$$

where

$$\pi_j^i = \frac{\Gamma\left(j - d_i\right)}{\Gamma(j+1)\Gamma(-d_i)}.$$

$\Gamma(.)$ is the gamma function.

When $d_i < \frac{1}{2}$, the process (3) is causal and has an infinite moving-average representation as follows:

$$X_{i+pm} = (1-B)^{-d_i} \varepsilon_{i+pm} = \sum_{j=0}^{\infty} \psi_j^i \varepsilon_{i+pm-j}, \tag{5}$$

where

$$\psi_j^i = \frac{\Gamma\left(j + d_i\right)}{\Gamma\left(j+1\right)\Gamma\left(d_i\right)}.$$

The convergence of infinite sums, in (4) and (5), is to be understood in the quadratic mean sense, for $i = 1, \ldots, p$.

We consider the parameters vector $d = (d_1, \ldots, d_p)$ a p-dimensional real vector $d \in \Theta$, where Θ is a compact subset of $\mathbb{R}^p$. We assume that we have an realization of size n, $(X_1, \ldots.., X_n)$ of the solution of equation (3). Suppose, for simplicity of notation reasons, that the size n is a multiple of p, i.e. $n = pn'$. Let $i = 1, \cdots, p$ and $m = 0, 1, \cdots, n' - 1$.

The related multivariate stationary process of the PtvARFIMA$_p$ $(0, d_i, 0)$ model is given by

$$\begin{pmatrix} \sum_{j=0}^{\infty} \pi_j^1 & & 0 \\ & \ddots & \vdots \\ 0 & \cdots & \sum_{j=0}^{\infty} \pi_j^p \end{pmatrix} \begin{pmatrix} X_{1+pm-j} \\ X_{2+pm-j} \\ . \\ . \\ X_{p+pm-j} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1+pm} \\ \varepsilon_{2+pm} \\ . \\ . \\ \varepsilon_{p+pm} \end{pmatrix} = \epsilon_m.$$

2.2. *Main assumptions and propositions*

In order to deal with MHD estimation based on [5, the Theorems 2 and 4] and [10, Lemma 3.1], our starting point is to introduce further notations and assumptions that are necessary in the sequel. For more detail, we refer the reader to [16].

Let $\widehat{d}_n$ be the MHD estimate of d, which minimizes the Hellinger distance between f_n and f_d. That is

$$\widehat{d}_n = \arg\min_{d \in \Theta} H_2\left(f_n, f_d\right), \tag{6}$$

where $H_2(f_n, f_d)$ is the Hellinger distance between f_n and f_d defined by

$$H_2\left(f_n, f_d\right) = \left(\int_{\mathbb{R}^p} \left| f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x) \right|^2 dx\right)^{\frac{1}{2}}, \tag{7}$$

where $f_d(.)$ is the theoretical probability density of ϵ_m, with $f_d : \mathbb{R}^p \to \mathbb{R}_+$.
$f_n(.)$ is the random function of $\widehat{\epsilon}_m$ given by

$$f_n(x) = \frac{1}{nh_n^p} \sum_{m=0}^{n'-1} K\left(\frac{x - \widehat{\epsilon}_m}{h_n}\right), \quad x \in \mathbb{R}^p,$$

where

$$\begin{pmatrix} \sum_{j=0}^{n} \pi_j^1 & & 0 \\ & \ddots & \vdots \\ 0 & \cdots & \sum_{j=0}^{n} \pi_j^p \end{pmatrix} \begin{pmatrix} X_{1+pm-j} \\ X_{2+pm-j} \\ . \\ . \\ X_{p+pm-j} \end{pmatrix} = \begin{pmatrix} \widehat{\varepsilon}_{1+pm} \\ \widehat{\varepsilon}_{2+pm} \\ . \\ . \\ \widehat{\varepsilon}_{p+pm} \end{pmatrix} = \widehat{\epsilon}_m, \ m = 0, \cdots, n' - 1, \tag{8}$$

and $K : \mathbb{R}^p \to \mathbb{R}_+$ is the kernel density function and h_n the bandwidth.
Let $\widetilde{f}_n(.)$ denote the kernel density estimation of f_d of ϵ_m, such that

$$\widetilde{f}_n(x) = \frac{1}{nh_n^p} \sum_{m=0}^{n'-1} K\left(\frac{x - \epsilon_m}{h_n}\right), \quad x \in \mathbb{R}^p,$$

In the whole of the article, we consider the following assumptions:

Assumption 1. *The process given in* (3) *satisfies the sufficient condition of invertibility and causality with* $0 < d_i < \frac{1}{2}$.

Assumption 2. $E(|\epsilon_m|^t) < +\infty$ *for* $t \geq 1$.
For all $(u; v) \in \mathbb{R}^{2p}$, *we have*

$$\int_{\mathbb{R}^p} K^2(u)du < \infty, \qquad \int_{\mathbb{R}^p} u_i K(u)du = 0 \qquad \text{for} \quad 1 \leq i \leq p.$$
$$\int_{\mathbb{R}^p} u_i u_j K(u)du = 0, \qquad \int_{\mathbb{R}^p} u_i^2 K(u)du < \infty \qquad \text{for} \quad 1 \leq j \leq p.$$

There exists $c > 0$ *such that* $\sup_{u \in \mathbb{R}^p} |K(u+v) - K(u)| \leq c|v|$.

Assumption 3. ϵ_m *admits a density absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^p$. *For all* $d \in \Theta$ *and* $x \in \mathbb{R}^p$, *the functions* $x \to f_d(x)$ *and* $x \to f_d^{\frac{1}{2}}(x)$ *are continuously differentiable.*

Assumption 4. *For all* $x \in \mathbb{R}^p$, *the functions* $d \to \frac{\partial}{\partial d_i} f_d^{\frac{1}{2}}(x)$, *for* $1 \leq i \leq p$ *and* $d \to \frac{\partial^2}{\partial d_i \partial d_k} f_d^{\frac{1}{2}}(x)$, *for* $1 \leq j, k \leq p$, *are bounded, continuous and defined in* $L^2(\mathbb{R}^p)$.

Assumption 5. $h_n = n^\alpha \ell(n)$, $-1 < \alpha < 0$ *with* $\ell(.)$ *a slowly varying function,*

$$\lim_{n \to \infty} h_n = 0, \quad \lim_{n \to \infty} nh_n = \infty, \quad \lim_{n \to \infty} \frac{\ell(an)}{\ell(n)} = 1, a > 0.$$

For all

$$d \in \Theta, \sup_{x \in \mathbb{R}^p} \left| \frac{\partial^j f_d}{\partial x_k^j}(x) \right| < \infty, \quad j = 0, 1, 2, \ldots. \quad \text{and} \quad k = 1, \ldots\ldots, p.$$

Assumption 6. *For* $d, d' \in \Theta$, $d \neq d'$ *implies that* $\{x \in \mathbb{R}^p / f_d(x) \neq f_{d'}(x)\}$ *is a set of positive Lebesgue measure.*

Assumption 7. *There exists a constant M such that* $\sup_{x\in\mathbb{R}^p} f_n(x) \le M < \infty$.

Note also that

$$g_d(x) = f_d^{\frac{1}{2}}(x), \quad g'_d(x) = \frac{\partial g_d}{\partial_d}(x),$$
$$g''_d(x) = \frac{\partial^2 g_d}{\partial d \partial d^t}(x).U_d(x) = \left[\int_{\mathbb{R}^p} g'_d(x)\left[g'_d(x)\right]^t dx\right]^{-1} g'_d(x).$$

Here and in what follows t denotes the transpose, and $\xrightarrow{a.s}$ the convergence with probability one.

Condition 8. *The components of* g'_d *and* g''_d *are in* L_2 *and the norms of the components are continuous functions at* d.

Condition 9. $\int_{\mathbb{R}^p} g''_d(x) g_d(x)dx$ *is a non-singular* $(p\times p)$*-matrix.*

Proposition 10. *Under the assumptions 1-3 we have for any* $p \ge 2$,

$$f_n(x) - f_d(x) \xrightarrow{a.s} 0, \ as \ n \to \infty. \tag{9}$$

Proof of Proposition 10. Using the triangular inequality, we have:

$$\sup_{x\in\mathbb{R}^p}\left|f_n(x)-f_d(x)\right| \le \sup_{x\in\mathbb{R}^p}\left|f_n(x)-\widetilde{f}_n(x)\right| + \sup_{x\in\mathbb{R}^p}\left|\widetilde{f}_n(x)-E\left(\widetilde{f}_n(x)\right)\right| + \sup_{x\in\mathbb{R}^p}\left|E\left(\widetilde{f}_n(x)\right)-f_d(x)\right|.$$

Now, we will study the almost sure convergence of every term.
For the first term

$$\sup_{x\in\mathbb{R}^p}\left|f_n(x)-\widetilde{f}_n(x)\right| \xrightarrow{a.s} 0, \tag{i}$$

$$\sup_{x\in\mathbb{R}^p}\left|f_n(x)-\widetilde{f}_n(x)\right| = \frac{1}{nh_n^{p+1}}\sum_{m=0}^{n'-1} K\left|\epsilon_m - \widehat{\epsilon}_m\right|.$$

Denoted by $\varepsilon_m - \widehat{\varepsilon}_m = Z_m$, such that

$$\begin{pmatrix} \sum\limits_{j=n+1}^{\infty}\pi_j^1 & & 0 \\ & \ddots & \vdots \\ 0 & \cdots & \sum\limits_{j=n+1}^{\infty}\pi_j^p \end{pmatrix}\begin{pmatrix} X_{1+pm-j} \\ X_{2+pm-j} \\ . \\ . \\ X_{p+pm-j}\end{pmatrix} = \begin{pmatrix} z_{1+pm} \\ z_{2+pm} \\ . \\ . \\ z_{p+pm}\end{pmatrix}. \tag{10}$$

By Assumption 2

$$\sup_{x\in\mathbb{R}^p}\left|f_n(x)-\widetilde{f}_n(x)\right| \le \frac{c}{nh_n^{p+1}}\sum_{m=0}^{n'-1}\left|Z_m\right|,$$

with $c > 0$

$$E\left(\frac{1}{nh_n^{p+1}}\sum_{m=0}^{n'-1}\|Z_m\|\right)^2 = \frac{1}{n^2h_n^{2p+2}}E\left(\sum_{m=0}^{n'-1}(\|Z_m\|)^2 + \sum_{\substack{t=0\\t\neq m}}^{n'-1}(\|Z_m\|)(\|Z_t\|)\right)$$
$$\le \frac{1}{n^2h_n^{2p+2}}\left(2\sum_{m=0}^{n'-1}E(\|Z_m\|)^2 + \sum_{\substack{t=0\\t\neq m}}^{n'-1}E(\|Z_t\|)^2\right).$$

The coefficients π_j^i for $i = 1,\ldots,p$ are square summable, (see [18] for the fixed (i)).
The almost sure convergence of (i) is then verified.
Next, for the second term

$$\sup_{x\in\mathbb{R}^p}\left|\widetilde{f}_n(x)-E\left(\widetilde{f}_n(x)\right)\right| \xrightarrow{a.s} 0. \tag{ii}$$

Let us the Prakasa Rao [20] inequality, modified for the case of multivariate stationary process (see [16]). Posing

$$\delta_m(x,\epsilon_m) = \frac{1}{h_n^p} K\left(\frac{x-\epsilon_m}{h_n}\right),$$

and $s_n = nh_n$, there exists $c_0 > 0$ such that

$$\sup_{x\in\mathbb{R}^p} \delta_m(x,\epsilon_m) \le c_0 s_n.$$

$$\mathbb{P}\left(\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon\sqrt{\frac{s_n \log(n)}{n}}\right) \le 2\exp\left(-\frac{s_n\log(n)\epsilon^2}{8c_0 M}\right),$$

$$\mathbb{P}\left(\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon n^{\frac{\alpha}{2}}\log(n)\right) \le 2\exp\left(-\frac{n^{\alpha+1}\log^2(n)\epsilon^2}{8c_0 M}\right),$$

$$\mathbb{P}\left(n^{1/4}\sup_{x\in\mathbb{R}^p}\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon n^{\frac{2\alpha+1}{4}}\log(n)\right) \le 2\exp\left(-\frac{n^{\frac{4\alpha+5}{4}}\log^2(n)\epsilon^2}{8c_0 M}\right),$$

Hence $\lim_{n\to\infty} \frac{s_n\log(n)}{n} = 0$. Moreover, let a_n be the sequence such that $a_n = \beta\log(n)$, where $\beta \ge 2$. Since

$$0 < \frac{\epsilon^2 n^{\frac{4\alpha+5}{4}}\log^2(n)}{8c_0 M} < \infty, \quad \text{we have} \quad \exp\left(-\frac{\epsilon^2 n^{\frac{4\alpha+5}{4}}\log^2(n)}{8c_0 M}\right) < n^{\beta}.$$

$$\mathbb{P}\left(n^{1/4}\sup_{x\in\mathbb{R}^p}\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon n^{\frac{2\alpha+1}{4}}\log(n)\right) \le \frac{2}{n^\beta},$$

$$\sum_{n>1}\mathbb{P}\left(n^{1/4}\sup_{x\in\mathbb{R}^p}\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon n^{\frac{2\alpha+1}{4}}\log(n)\right) \le \sum_{n>1}\frac{2}{n^\beta}.$$

Since the serie $\sum_{n>1}\frac{2}{n^\beta}$ converges, by Borel-cantelli Lemma, it follows that

$$\sup_{x\in\mathbb{R}^p}\left|\tilde{f}_n(x) - E\left(\tilde{f}_n(x)\right)\right| > \epsilon = o\left(n^{\frac{2\alpha+1}{4}}\log(n)\right), \tag{11}$$

almost surely when $n\to\infty$. Hence $\sup_{x\in\mathbb{R}^p}|\tilde{f}_n(x) - E(\tilde{f}_n(x))|$ converges a.s to 0.
Finally, for the third term

$$\sup_{x\in\mathbb{R}^p}\left|E\left(\tilde{f}_n(x)\right) - f_d(x)\right| \xrightarrow{a.s} 0. \tag{iii}$$

We have

$$\begin{aligned} E\left(\tilde{f}_n(x)\right) &= \frac{1}{h_n^p}E\left(K\left(\frac{x-\epsilon_1}{h_n}\right)\right)\\ &= \frac{1}{h_n^p}\int_{\mathbb{R}^p} K\left(\frac{x-z}{h_n}\right) f_d(z)dx\\ &= \frac{1}{h_n^p}\int_{\mathbb{R}^p} K(u) f_d(x-uh_n)\,du, \end{aligned}$$

by a simple calculus and using the generalized developement of Taylor we get

$$\begin{aligned} E\left(\tilde{f}_n(x)\right) - f_d(x) &= \int_{\mathbb{R}^p} K(u)\left[\left|\sum_{k=1}^p \frac{\partial f_d}{\partial x_k}(x)\right|(-h_n)\,u_k + \frac{h_n^2}{2}\left|\sum_{j=1}^p\sum_{k=1}^p \frac{\partial^2 f_d}{\partial x_j x_k}(x)\right| u_j u_k + o\left(h_n^2\right)\right] du\\ &= \int_{\mathbb{R}^p} K(u)\left[\frac{h_n^2}{2}\left|\sum_{k=1}^p \frac{\partial^2 f_d}{\partial x_k^2}(x)\right| u_k^2 + o\left(h_n^2\right)\right] du \end{aligned}$$

$$\sup_{x\in\mathbb{R}^P}\left|E\left(\tilde{f}_n(x)\right) - f_d(x)\right| \le \frac{h_n^2}{2}\sum_{k=1}^p \sup_{x\in\mathbb{R}^P}\left|\frac{\partial^2 f_d}{\partial x_k^2}(x)\right|\int_{\mathbb{R}^p} K(u)\left[u_k^2 + o(1)\right] du.$$

By the assumptions 2 and 5 we have

$$\lim_{n\to\infty} \frac{h_n^2}{2} \sum_{k=1}^{p} \sup_{x\in\mathbb{R}^p} \left| \frac{\partial^2 f_d}{\partial x_k^2}(x) \right| \to 0, \int_{\mathbb{R}^p} K(u)\left[u_k^2 + o(1)\right] du < \infty, \text{ for every } x \in \mathbb{R}^p.$$

Therefore, $\sup_{x\in\mathbb{R}^p} |E\left(\tilde{f}_n(x)\right) - f_d(x)| \to 0$, as $n \to \infty$.

According to (i), (ii) and (iii), we get the almost sure convergence to zero of $f_n(x) - f_d(x)$.

This completes the proof of Proposition 10. □

Proposition 11. *Noting by* $\mathbb{F}$ *the set of all densities with respect to the Lebesgue measure, then for every* $g \in \mathbb{F}$, *the functional* $T : \mathbb{F} \to \Theta$ *is such that:*

$$T(g) = \left\{ d_0 \in \Theta : H_2\left(g, f_{d_0}\right) = \min_{d\in\Theta} H_2\left(g, f_d\right\}\right\}.$$

If such a minimum exists. In case $T(g)$ *is not unique,* $T(g)$ *will mean one of the minimum values selected arbitrarily.*

Proof of Proposition 11. The proof is well detailed by [5, Theorem 1] and by [10, Lemma 3.1]. □

Proposition 12. *Assume that the Assumptions 3-6 and the Conditions 8-9 hold and that* d *lies in interior of* Θ. *So, for any sequence* f_n *converging to* f_d *in the Hellinger metric, we have*

$$T\left(f_n(n)\right) = d + \int_{\mathbb{R}^p} U_d(x)\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right] dx + V_n \int_{\mathbb{R}^p} g_d'(x)\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right] dx. \tag{12}$$

Here V_n *is a non-singular* $p \times p$*-matrix, such that the components of* $\sqrt{n}V_n$ *tend to zero when* $n \to \infty$.

Proof of Proposition 12. See the proof of [5, Theorem 4]. □

Proposition 13. *Assume that the Assumptions 2-3 and 5-7 hold. Then, the limiting distribution of*

$$\sqrt{nh_n}\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right] \quad is \quad N\left(0, f_d(x)\int_{\mathbb{R}^p} K^2(u)du\right).$$

Proof of Proposition 13. The proof of this is similar to the proof of [24, Theorem 3]. □

After these preliminary results, we are ready to establish the MHD estimation. The almost sure convergence of $\widehat{d_n}$ to d and it asymptotic normality are formally stated in the next Section 3.

3. Asymptotic properties of the MHDE for PtvARFIMA

Theorem 14. *Suppose that Assumptions 1-7 hold. Then,*

$$\widehat{d_n} \xrightarrow{a.s} d, \text{ as } n \to \infty. \tag{13}$$

Proof of Theorem 14. By the propositions 10 and 11, the proof can directly be achieved. From proposition 10, we have

$$\mathbb{P}\left\{\lim_{n\to\infty} f_n^{\frac{1}{2}}(x) = f_d^{\frac{1}{2}}(x) \ \forall x\right\} = 1.$$

Consequently

$$H_2\left(f_n, f_d\right) \xrightarrow{a.s} 0 \quad \text{as} \quad n \to \infty.$$

Next, by Proposition 11, $T(f_d) = d$ uniquely on Θ, then the functional T is continuous at f_d in the Hellinger topology. Therefore

$$\widehat{d_n} = T\left(f_n(x)\right) \to T\left(f_d(x)\right) = d,$$

almost surely as $n \to \infty$. □

Theorem 15. *Suppose that the Assumptions 1-7 and Conditions 8 and 9 hold. Then, the limit distribution of* $\sqrt{n}(\widehat{d}_n - d)$ *is*

$$\sqrt{n}\left(\widehat{d}_n - d\right) \to N\left(0, \Sigma^2\right), \tag{14}$$

where

$$\Sigma^2 = \frac{1}{4}\left[\int_{\mathbb{R}^p} g_d'(x)\left[g_d'(x)\right]^t dx\right]^{-1} \int_{\mathbb{R}^p} k^2(u)du. \tag{15}$$

Proof of Theorem 15. By Proposition 12, one can show that

$$\begin{aligned}\sqrt{n}\left(\widehat{d}_n - d\right) &= \sqrt{n}\int_{\mathbb{R}^p} U_d(x)\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right]dx + \sqrt{n}V_n\int_{\mathbb{R}^p} g_d'(x)\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right]dx.\\ &= \sqrt{n}\int_{\mathbb{R}^p} U_d(x)\left[f_n^{\frac{1}{2}}(x) - f_d^{\frac{1}{2}}(x)\right]dx + o_p(1).\end{aligned}$$

With $V_n \to 0$ in probability, $U_d \in L_2$ and $U_d \perp f_d^{\frac{1}{2}}$, where $\perp$ is the orthogonality in L_2.
For $b \geq 0$, $a > 0$, the algebraic identity is given by

$$b^{\frac{1}{2}} - a^{\frac{1}{2}} = \frac{b-a}{2a^{\frac{1}{2}}} - \frac{(b-a)^2}{\left[2a^{\frac{1}{2}}\left(b^{\frac{1}{2}} + a^{\frac{1}{2}}\right)^2\right]}.$$

According to Assumption 3, the condition $f_d^{\frac{1}{2}}(x) > 0$ and the algebraic identity, we have

$$\sqrt{n}\left(\widehat{d}_n - d\right) = \sqrt{n}\int_{\mathbb{R}^p} U_d(x)\left[\frac{f_n(x) - f_d(x)}{2f_d^{\frac{1}{2}}(x)}\right]dx + A_n.$$

Where

$$A_n = -\sqrt{n}\int_{\mathbb{R}^p} U_d(x)\left[\frac{\left[f_n(x) - f_d(x)\right]^2}{2f_d^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x) + f_d^{\frac{1}{2}}(x)\right)^2}\right]dx.$$

Hence

$$|A_n| \leq 2\delta^{-\frac{3}{2}}\int_{\mathbb{R}^p} |U_d(x)|\sqrt{n}\left[f_n(x) - f_d(x)\right]^2,$$

where $\delta = \inf_{x \in \mathbb{R}^p} f(x)$,

$$2f_d^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x) + f_d^{\frac{1}{2}}(x)\right)^2 > 2\delta^{\frac{3}{2}}.$$

Conditions 8 and 9 imply that $U_d(x)$ is continuous and bounded. So, by proposition (3) and the Vitali's Theorem, $|A_n|$ tends to 0 in probability for $n \to \infty$.

So, we can rewrite $\sqrt{n}(\widehat{d}_n - d)$ as folows

$$\sqrt{n}\left(\widehat{d}_n - d\right) = \sqrt{n}\int_{\mathbb{R}^p} U_d(x)\left[\frac{f_n(x) - f_d(x)}{2f_d^{\frac{1}{2}}(x)}\right]dx + o_p(1).$$

By Proposition 13 and by a simple calculus we deduce the limiting distribution of $\sqrt{n}(\widehat{d}_n - d)$,

$$\int_{\mathbb{R}^p}\left[\frac{U_d(x)}{2f_d^{\frac{1}{2}}(x)}\right]\left[\frac{U_d(x)}{2f_d^{\frac{1}{2}}(x)}\right]^t\int_{\mathbb{R}^p} K^2(u)du f_d(x)dx = \frac{1}{4}\int_{\mathbb{R}^p} U_d(x)(U_d(x))^t dx\int_{\mathbb{R}^p} K^2(u)du.$$

Which proves Theorem 15. □

4. Simulation

In this section, we present a simulation experiment, in order to illustrate, the most significative results for the method proposed in this article, we apply numerically the Minimum Hellinger distance method. We consider two cases:

First case, we assume that the white noise ϵ_m, $\epsilon_m = (\varepsilon_{1+2m}, \varepsilon_{2+2m})^{'}$ is the density function of the standard normal distribution and in this case the kernel density K is also the density of the standard normal distribution.

Second case, we consider ϵ_m and K as the density function of the Cauchy distribution (see [14]).

We generate a PtvARFIMA$_2(0, d, 0)$, $X_m = (X_{1+2m}, X_{2+2m})^{'}$, $m = 0, 1, .., n\prime - 1$, we use three simple size $n = 10$; 50; 100, with two different values of d, such that $d = (0.2, 0.15)$ and $d = (0.49, 0.4)$. We have

$$MSE = \frac{1}{n_r}\sum_{j=1}^{n_r}\left\{\left(\widehat{d}^{j}_{n,1} - d_1\right)^2 + \left(\widehat{d}^{j}_{n,2} - d_2\right)^2\right\},$$

where $n_r = 100$ is the number of replications and $(\widehat{d}^{j}_{n,1}, \widehat{d}^{j}_{n,2})$ denote the estimate of d for the jth replication. The results are displayed in the following tables:

For the first case,

Table 1. MHDE for $d = (0.2, 0.15)$.

n	10	50	100
$MHDE$	$(0.2006, 0.1494)$	$(0.1987, 0.1512)$	$(0.2077, 0.1423)$
MSE	0.0075	0.0071	0.0054

Table 2. MHDE for $d = (0.49, 0.4)$.

n	10	50	100
$MHDE$	$(0.4944, 0.0.3956)$	$(0.4939, 0.4011)$	$(0.4938, 0.401)$
MSE	0.000895	0.000738	0.00064

Second case

Table 3. MHDE for $d = (0.2, 0.15)$.

n	10	50	100
$MHDE$	$(0.2084, 0.1416)$	$(0.1992, 0.1507)$	$(0.1988, 0.1511)$
MSE	0.0084	0.0068	0.0063

Table 4. MHDE for $d = (0.49, 0.4)$.

n	10	50	100
$MHDE$	$(0.4944, 0.3952)$	$(0.494, 0.3992)$	$(0.4943, 0.3973)$
MSE	0.000859	0.000796	0.000758

From tables (Tables 1, 2, 3 and 4) we can deduce that the mean parameter estimates are very close in value to the true value of the parameters, and the normal law provides a high accuracy. However, the MSE is decreasing with the increase in sample size. This happens because the impact of the decrease in variance with increasing sample size.

5. conclusion

In this paper, we have addressed the problem of the Minimum Hellinger distance (MHD) in the [5]'s style for a periodically time-varying long-memory parameter. We have constructed an estimate for the vector parameters d, using the MHD method and studied its asymptotic properties by considering the related multivariate stationary model. We have also presented some numerical simulations illustrating our theoretical results.

References

[1] A. Amimour, K. Belaide, "A long memory time series with a periodic degree of fractional differencing", https://arxiv.org/abs/2008.01939, 2020.
[2] ———, "On the invertibility in periodic ARFIMA models", https://arxiv.org/abs/2008.02978, 2020.
[3] ———, "Local asymptotic normality for a periodically time varying long memory parameter", *Commun. Stat., Theory Methods* **51** (2022), no. 9, p. 2936-2952.
[4] M. Bentarzi, M. Hallin, "On the invertibility of periodic moving-average models", *J. Time Ser. Anal.* **15** (1994), no. 3, p. 263-268.
[5] R. Beran, "Minimum Hellinger distance estimates for parametric models", *Ann. Stat.* **5** (1977), no. 3, p. 445-463.
[6] G. E. P. Box, G. M. Jenkins, *Time series analysis: Forecasting and control*, revised ed., Holden-Day Series in Time Series Analysis, Holden Day, 1970.
[7] C. Ching-fan, "Sample means, sample autocovariances, and linear regression of stationary multivariate long memory processe", *Econom. Theory* **18** (2002), no. 1, p. 51-78.
[8] E. G. Gladyshev, "Periodically correlated random sequences", *Sov. Math., Dokl.* **2** (1961), p. 385-388.
[9] S. N. Gupta, "Estimation in long memory time series models", *Commun. Stat., Theory Methods* **21** (1992), no. 5, p. 1327-1338.
[10] O. Hili, "On the estimation of nonlinear time series models", *Stochastics Stochastics Rep.* **52** (1995), no. 3-4, p. 207-226.
[11] J. R. M. Hosking, "Fractional differencing", *Biometrika* **68** (1981), no. 1, p. 165-176.
[12] A. Kamagaté, O. Hili, "Estimation par le minimum de distance de Hellinger d'un processus ARFIMA", *C. R. Acad. Sci. Paris* **350** (2012), no. 13-14, p. 721-725.
[13] ———, "The quasi maximum likelihood approach to statistical inference on a nonstationary multivariate ARFIMA process", *Random Oper. Stoch. Equ.* **21** (2013), no. 3, p. 305-320.
[14] H.-Y. Lee, H.-J. Park, H.-M. Kim, "A clarification of the Cauchy distribution", *Commun. Stat. Appl. Methods* **21** (2014), no. 2, p. 183-191.
[15] L. Mayoral, "Minimum distance estimation of stationary and non-stationary ARFIMA processes", *The Econometrics Journal* **10** (2007), no. 1, p. 124-148.
[16] K. S. Mbeke, O. Hili, "Estimation of a stationary multivariate ARFIMA process", *Afr. Stat.* **13** (2018), no. 3, p. 1717-1732.
[17] H. J. Newton, "Using periodic autoregressions for multiple spectral estimation", *Technometrics* **24** (1982), no. 2, p. 109-116.
[18] M. Odaki, "On the invertibility of fractionally differenced ARIMA processes", *Biometrika* **80** (1993), no. 3, p. 703-709.
[19] M. Pagano, "On periodic and multiple autoregressions", *Ann. Stat.* **6** (1978), no. 6, p. 1310-1317.
[20] B. L. S. Prakasa Rao, *Nonparametric functional estimation*, Probability and Mathematical Statistics, Academic Press Inc., 1983.
[21] F. Sowell, "Maximum likelihood estimation of stationary univariate fractionally integrated time series models", *J. Econom.* **53** (1992), no. 1-3, p. 165-188.
[22] B. M. Troutman, "Some results in periodic autoregression", *Biometrika* **66** (1979), no. 2, p. 219-228.
[23] A. V. Vecchia, "Periodic autoregressive-moving average (PARMA) modeling with application to water resources", *Journal of the American Water Resources Association* **21** (1985), no. 5, p. 721-30.
[24] W. B. Wu, J. Mielniczuk, "Kernel density estimation for linear process", *Ann. Stat.* **30** (2002), no. 5, p. 1441-1459.
[25] Y. Yajima, "On estimation of long-memory time series models", *Aust. J. Stat.* **27** (1985), no. 3, p. 303-320.