

On strong consistency of estimators for infinite variance time series

by

Thomas Mikosch¹ and Claudia Klüppelberg
Victoria University Wellington and ETH Zürich.

Abstract

This paper is a continuation of the authors' work on parameter estimation for heavy-tailed models. Here we review the sample autocorrelation and the Whittle estimator for linear processes in i.i.d infinite variance r.v.'s from the point of view of a.s. convergence. We consider examples which show that the sample autocorrelations do not converge a.s., but they do so along a specified non-random subsequence. This fact can then be used to obtain a.s. convergence of the Whittle estimator along a specified non-random subsequence.

1. Research supported in part by New Zealand FRST Grant (93-VIC-36-039).

AMS 1980 Subject Classifications: Primary 62M15, 62F10;

Secondary 60G10, 62M10

Key words and phrases: moving average process, general linear model, ARMA process, spectral theory, periodogram, transfer function, Whittle estimator, parameter estimation, a.s. convergence, autocorrelation.

1. Introduction

In this paper two closely related, but distinct, subjects are studied: the strong consistency of the sample autocorrelations and of the Whittle estimator for a stationary process. Throughout we consider the discrete moving average process

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z}, \quad (1.1)$$

where $(Z_t)_{t \in \mathbb{Z}}$ is a noise sequence of i.i.d. random variables (r.v.'s) having not necessarily a finite variance. In three preceding papers (Klüppelberg and Mikosch (1993, 1994) and Mikosch, Gadrich, Klüppelberg, Adler (1994)) we studied the asymptotic behaviour of periodogram-type estimators for the process $(X_t)_{t \in \mathbb{Z}}$ under the condition that $Z = Z_0$ is in the domain of normal attraction of an α -stable law for some $\alpha \in (0, 2)$.

In Klüppelberg and Mikosch (1993, 1994) it was shown that the self-normalised periodogram

$$\tilde{I}_{n,X}(\lambda) = \left| \sum_{t=1}^n X_t e^{-i\lambda t} \right|^2 / \sum_{t=1}^n X_t^2, \quad -\pi < \lambda \leq \pi,$$

converges in distribution to

$$\frac{|\psi(e^{-i\lambda})|^2}{\psi^2} \frac{\alpha^2(\lambda) + \beta^2(\lambda)}{Y_0}$$

where $\psi(e^{-i\lambda}) = \sum_{j=-\infty}^{\infty} \psi_j e^{-i\lambda j}$ is the transfer function of the linear filter $(\psi_j)_{j \in \mathbb{Z}}$, $\psi^2 = \sum_{j=-\infty}^{\infty} \psi_j^2$, and the vector $(\alpha(\lambda), \beta(\lambda), Y_0)$ has a mixed stable distribution such that $(\alpha(\lambda), \beta(\lambda))$ are jointly α -stable and Y_0 is positive $\alpha/2$ -stable. Furthermore, the vector of periodogram ordinates $(\tilde{I}_{n,X}(\lambda_i))_{i=1, \dots, m}$ at distinct frequencies $0 < \lambda_1 < \dots < \lambda_m < \pi$ converges weakly, and the components of the limit vector have exponentially fast decreasing tails and are uncorrelated. We also studied smoothed versions of the self-normalised

periodogram and established their weak convergence to the self-normalised power transfer function $\left| \psi \left(e^{-i\lambda} \right) \right|^2 / \psi^2$.

In the present paper we weaken the above assumptions on Z considerably: we only require that $E |Z|^d < \infty$ for some $d > 0$ and that $\left(\sum_{t=1}^n Z_t^2 \right)_{n \in \mathcal{N}}$ satisfies some mild tightness condition. The price we have to pay for this is that we can in general not derive a rate of convergence.

The paper is organised as follows: Assumptions and notations are introduced in Section 2. In Section 3. we study the sample autocorrelations. We review results from Davis and Resnick (1986) and from Mikosch et al. (1994). We show that the sample autocorrelations do in general not converge a.s., but they do so along a specified subsequence of the integers. In Section 4. we deal with parameter estimation for ARMA processes. We recall the notion of the Whittle estimate and some related results in the literature. Using the strong consistency of the sample autocorrelations we show a.s. convergence of the Whittle estimate along a subsequence under very general conditions.

2. Notation and assumptions

We consider the moving average process $(X_t)_{t \in \mathbb{Z}}$ defined by (1.1). To formulate the conditions on the noise $(Z_t)_{t \in \mathbb{Z}}$ we introduce the following functions for $x > 0$:

$$\begin{aligned} G(x) &= P \left(Z^2 > x \right) \\ K(x) &= x^{-2} E Z^4 I \left(Z^2 \leq x \right) \\ Q(x) &= G(x) + K(x) = E \left[1 \wedge \left(x^{-1} Z^2 \right)^2 \right]. \end{aligned}$$

Since Q is strictly decreasing and continuous on $(0, \infty)$ the identity

$$Q \left(a_n^2 \right) = \frac{1}{n}, \quad n \in \mathcal{N},$$

defines a sequence of positive numbers a_n such that $a_n \uparrow \infty$ as $n \rightarrow \infty$. Furthermore, define

$$\gamma_{n,Z}^2 = a_n^{-2} \sum_{t=1}^n Z_t^2, \quad n \in \mathcal{N}. \quad (2.1)$$

For the moving average process as in (1.1) we introduce the following assumptions: There exists some $d > 0$ such that

$$(A1) \quad E|Z|^d < \infty;$$

$$(A2) \quad \sum_{j=-\infty}^{\infty} |j| |\psi_j|^\delta < \infty \quad \text{for } \delta = 1 \wedge d;$$

$$(A3) \quad n/a_n^{2\delta} \rightarrow 0, \quad n \rightarrow \infty, \quad \text{for } \delta = 1 \wedge d;$$

$$(A4) \quad \lim_{x \rightarrow 0} \limsup_{n \rightarrow \infty} P(\gamma_{n,Z}^2 \leq x) = 0.$$

(A5) There exists a sequence of positive numbers e_n such that

$$\liminf_{n \rightarrow \infty} e_n^{-2} \sum_{t=1}^n Z_t^2 = 1 \quad \text{a.s.}, \quad (2.2)$$

where the norming constants e_n satisfy the following conditions: There exists some $\nu \in \mathcal{N}$ such that for $n_k = k^\nu$, $k \in \mathcal{N}$, $\sum_{k=1}^{\infty} (n_k e_{n_k}^{-2\delta} + e_{n_k}^{-\delta}) < \infty$, for $\delta = 1 \wedge d$, and (e_n/e_{n_k}) is bounded away from 0 and ∞ uniformly for $n \in [n_k, n_{k+1}]$ for all $k \in \mathcal{N}$.

Remarks. 1) (A1) and (A2) imply absolute a.s. convergence of the series (1.1) for every $t \in \mathcal{Z}$. This is a consequence of the three-series theorem.

2) (A2) is obviously satisfied for every ARMA(p, q)-process. In this case the ψ_j decrease exponentially.

3) The conditions $E Z^2 < \infty$, (A3) and (A4) cannot hold together, since (A3) and the SLLN imply that $\gamma_{n,Z}^2 \xrightarrow{a.s.} 0$ contradicting (A4).

4) (A4) is a stochastic compactness condition on $\gamma_{n,Z}^2$. A necessary and sufficient condition for $\gamma_{n,Z}^2$ to be stochastically compact is

$$\liminf_{x \rightarrow \infty} K(x)/G(x) > 0.$$

(e.g. Maller (1981)). Furthermore, if $\gamma_{n,Z}^2$ is stochastically compact, then there exists some constant $c > 0$ such that for all $n \in \mathcal{N}$

$$P\left(\gamma_{n,Z}^2 \leq x\right) \leq c x, \quad x \geq 0,$$

(Griffin (1983)) which implies (A4).

5) A survey of results of type (2.2) can be found in Pruitt (1990, p. 1149). Fristedt and Pruitt (1971) proved under the restriction $E|Z|^d < \infty$ for some $d > 0$ that (2.2) holds with

$$e_n^2 = \frac{\ln \ln n}{\eta(\xi \ln \ln n / n)}$$

for some constant $\xi > 1$ where $\eta(\cdot) = \left(-\ln E e^{-\cdot Z^2}\right)^{\leftarrow}$ denotes the generalised inverse of $-\ln E e^{-\cdot Z^2}$.

A natural class of noise variables to satisfy conditions (A1) and (A3)–(A5) is the domain of attraction of an α -stable random variable, which we denote by $DA(\alpha)$. We also use the abbreviation $DNA(\alpha)$ for domain of normal attraction of an α -stable law. For the definition and properties of α -stable r.v.'s, their domain of attraction and regularly and slowly varying functions see e.g. Feller (1971), Bingham, Goldie and Teugels (1987) or Petrov (1975).

Now if $Z \in DA(\alpha)$ for some $\alpha \in (0, 2)$, then $Z^2 \in DA(\alpha/2)$ and

$$\lim_{x \rightarrow \infty} G(x)/K(x) = (4 - \alpha)/\alpha.$$

Then G is a regularly varying function with index $-\alpha/2$, and an alternative choice for the norming constants in (2.1) is given by

$$a_n^2 = G^{\leftarrow}(n^{-1}) = \inf \left\{x : G(x) < n^{-1}\right\},$$

where G^{\leftarrow} is the generalized inverse of G . This implies that $a_n^2 = n^{2/\alpha} L(n)$ where L is a slowly varying function and $\gamma_{n,Z}^2 \xrightarrow{d} Y_0$ for some positive $\alpha/2$ -stable r.v. Y_0 . Furthermore, $E|Z|^d < \infty$ for $d < \alpha$.

In the following lemma we summarise these relations.

Lemma 2.1. *Suppose $Z \in DA(\alpha)$ for some $\alpha \in (0, 2)$, then (A1), (A3) and (A4) hold for some $0 < d < \alpha$ and $a_n^2 = n^{2/\alpha} L(n)$ where L is a slowly varying function. Moreover, (A5) is satisfied for $d < \alpha$ and $e_n^2 = n^{2/\alpha} \tilde{L}(n)$ for some slowly varying function \tilde{L} . The number ν in (A5) can be chosen to satisfy $\nu > \alpha/(2\delta - \alpha) \vee (\alpha/\delta)$ provided $\delta > \alpha/2$. \square*

The following notation will be used throughout the paper: Let $(A_t)_{t \in \mathbb{Z}}$ be any of the sequences $(Z_t)_{t \in \mathbb{Z}}$ or $(X_t)_{t \in \mathbb{Z}}$ and choose (a_n) as in (2.1). Then define

$$\begin{aligned} \gamma_{n,A}^2 &= a_n^{-2} \sum_{t=1}^n A_t^2, \\ I_{n,A}(\lambda) &= a_n^{-2} \left| \sum_{t=1}^n A_t e^{-i\lambda t} \right|^2, \quad \lambda \in (-\pi, \pi], \\ \tilde{I}_{n,A}(\lambda) &= I_{n,A}(\lambda) / \gamma_{n,A}^2 = \left| \sum_{t=1}^n \tilde{A}_t e^{-i\lambda t} \right|^2, \quad \lambda \in (-\pi, \pi]. \end{aligned}$$

3. Consistency of the sample autocorrelation function

For $h \in \mathbb{Z}, h \neq 0$ define

$$\begin{aligned} \tilde{\gamma}_{n,X}(h) &= \gamma_{n,X}(h) / \gamma_{n,X}^2 \\ \tilde{\gamma}(h) &= \gamma(h) / \psi^2 \end{aligned}$$

where

$$\begin{aligned} \gamma_{n,X}(h) &= a_n^{-2} \sum_{t=1}^{n-|h|} X_t X_{t+|h|} \\ \gamma(h) &= \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+|h|}. \end{aligned}$$

If $E Z^2 < \infty$ it is well known that $\tilde{\gamma}_{n,X}(h)$ is a consistent estimator of the autocorrelation function $\tilde{\gamma}(h)$ of $(X_t)_{t \in \mathcal{Z}}$. In extension of this result, Davis and Resnick (1986) proved the following: For $Z \in DA(\alpha)$, $\alpha \in (0, 2)$, Z symmetric

$$\begin{aligned} & \left((nL(n))^{1/\alpha} \left(\tilde{\gamma}_{n,X}(h) - \tilde{\gamma}(h) \right) \right)_{h=1, \dots, m} \xrightarrow{d} \\ & \left(\sum_{j=1}^{\infty} (\tilde{\gamma}(j+h) - \tilde{\gamma}(j-h) - 2\tilde{\gamma}(j)\tilde{\gamma}(h)) \frac{Y_j}{Y_0} \right)_{h=1, \dots, m}, \end{aligned} \quad (3.1)$$

where L is a slowly varying function, Y_0, Y_1, Y_2, \dots are independent r.v.'s, Y_0 is positive $\alpha/2$ -stable and $(Y_j)_{j \in \mathcal{N}}$ are i.i.d. symmetric α -stable. (3.1) implies that $\tilde{\gamma}_{n,X}(h)$ is weakly consistent with limit $\tilde{\gamma}(h)$ and the rate of convergence in (3.1) compares favourably with \sqrt{n} in the finite variance case. Under our more general conditions (A1)-(A4) a precise result as (3.1) cannot be expected. In Mikosch et al. (1994) we proved the following weak consistency result:

Proposition 3.1. *Suppose $(X_t)_{t \in \mathcal{Z}}$ satisfies (A1)-(A4), then*

$$\tilde{\gamma}_{n,X}(h) \xrightarrow{P} \tilde{\gamma}(h), \quad h \in \mathcal{N}, \quad n \rightarrow \infty. \quad \square$$

Convergence in probability for $\tilde{\gamma}_{n,X}(h)$ can be strengthened to a.s. convergence provided the second moment of Z exists. This is not true in the infinite variance case as the following simple example shows. Similar examples can be constructed for any finite moving average process and any sample autocorrelation at lags greater than 1.

Example. Consider the $MA(1)$ process

$$X_t = Z_t + \theta Z_{t-1}, \quad t \in \mathcal{Z}, \quad |\theta| < 1,$$

for a symmetric α -stable Z , $\alpha \in (0, 2)$. Then, as mentioned in Section 2, (A1)-(A5) are satisfied for some $0 < d < \alpha$ and (e_n) can be chosen as

$$e_n = n^{1/\alpha} (\ln \ln n)^{-(2-\alpha)/(2\alpha)}.$$

Now consider

$$\begin{aligned}\tilde{\gamma}_{n,X}(1) &= \frac{\sum_{t=1}^{n-1} (Z_t + \theta Z_{t-1}) (Z_{t+1} + \theta Z_t)}{\sum_{t=1}^n (Z_t + \theta Z_{t-1})^2} \\ &= \frac{\sum_{t=1}^{n-1} Z_t Z_{t+1} + \theta \sum_{t=1}^{n-1} Z_{t-1} Z_{t+1} + \theta \sum_{t=1}^{n-1} Z_t^2 + \theta^2 \sum_{t=1}^{n-1} Z_{t-1} Z_t}{\sum_{t=1}^n Z_t^2 + \theta^2 \sum_{t=1}^n Z_{t-1}^2 + 2\theta \sum_{t=1}^n Z_{t-1} Z_t}.\end{aligned}$$

It follows from Rosinski and Woyczynski (1987) that for some $c > 0$

$$P(Z_1 Z_2 > x) \leq c x^{-\alpha} (1 + \ln^+ x^{-1}),$$

where $\ln^+ y = \max(0, \ln y)$, $y > 0$. Similar arguments as in the proof of Heyde's SLLN (see Stout (1974)) and the fact that

$$\sum_{h=1}^{\infty} P(Z_1 Z_2 > e_n^2) < \infty$$

imply that

$$\lim_{n \rightarrow \infty} \frac{e_n^{-2} \sum_{t=1}^{n-1} (Z_t Z_{t+1} + \theta^2 Z_t Z_{t-1} + \theta Z_{t-1} Z_{t+1})}{e_n^{-2} \sum_{t=1}^{n-1} Z_t^2} = 0 \quad \text{a.s.}$$

Thus

$$\tilde{\gamma}_{n,X}(1) = \frac{\theta + o(1)}{\theta^2 + 1 + (Z_n^2 / \sum_{t=1}^{n-1} Z_t^2) + o(1)} \quad \text{a.s.}$$

We shall show that every real number between 0 and ∞ is a.s. limit point of the sequence $(Z_n^2 / \sum_{t=1}^{n-1} Z_t^2)$ thus implying that the set of a.s. limit points of $(\tilde{\gamma}_{n,X}(1))$ is the interval $[0, \theta/(1 + \theta^2)]$ if $\theta > 0$ or $[\theta/(1 + \theta^2), 0]$ if $\theta < 0$. (Note that $\tilde{\gamma}(1) = \theta/(1 + \theta^2)$.)

Define for positive $\varepsilon_1 < \varepsilon_2$ and $\varepsilon_3 < \varepsilon_4$

$$A_n := \left\{ \varepsilon_2 > Z_n^2 n^{-2/\alpha} > \varepsilon_1 \right\}, \quad B_n := \left\{ \varepsilon_4 > n^{2/\alpha} / \sum_{t=1}^{n-1} Z_t^2 > \varepsilon_3 \right\},$$

then

$$P\left(\varepsilon_2\varepsilon_4 > Z_n^2 \Big/ \sum_{t=1}^{n-1} Z_t^2 > \varepsilon_1\varepsilon_3 \text{ i.o.}\right) \geq P(A_n \cap B_n \text{ i.o.}) .$$

Note that $\sum_{n=1}^{\infty} P(A_n) = \infty$ and $\liminf_{n \rightarrow \infty} P(B_n) > 0$. Since A_n and $\{B_1, B_2, \dots, B_n\}$ are independent for each $n \geq 1$, an application of a standard Borel-Cantelli lemma (e.g. Petrov (1975), Lemma 5, Section IX.2) yields $P(A_n \cap B_n \text{ i.o.}) > 0$, hence

$P(A_n \cap B_n \text{ i.o.}) = 1$ which implies that $Z_n^2 / \sum_{t=1}^{n-1} Z_t^2$ visits infinitely often with probability one any finite interval $(\varepsilon_1\varepsilon_3, \varepsilon_2\varepsilon_4)$. This shows that any positive real number is an a.s. limit point. A modification of the above arguments yields that 0 and ∞ must be a.s. limit points as well. \square

We shall show in the following, by replacing conditions (A3) and (A4) by the slightly more restrictive condition (A5), that it is possible to obtain a.s. convergence of $\tilde{\gamma}_{n,X}(h)$ to $\tilde{\gamma}(h)$ along some given subsequence (n_k) which is defined in (A5). In particular, this condition is satisfied for $Z \in DA(\alpha)$, $\alpha \in (0, 2)$.

The following result complements Proposition 3.1.

Proposition 3.2. *Suppose $(X_t)_{t \in \mathbb{Z}}$ satisfies (A1), (A2) and (A5). Then*

$$\tilde{\gamma}_{n_k,X}(h) \xrightarrow{\text{a.s.}} \tilde{\gamma}(h), \quad h \in \mathcal{N}, \quad n \rightarrow \infty .$$

Proof. We have the decomposition

$$\begin{aligned} \sum_{t=1}^n X_t X_{t+h} - \tilde{\gamma}(h) \sum_{t=1}^n X_t^2 &= \sum_{t=1}^n \sum_{i \neq j} \psi_i \left(\psi_{j+h} - \tilde{\gamma}(h) \psi_j \right) Z_{t-i} Z_{t-j} \\ &+ \sum_{t=1}^n \sum_i \psi_i \left(\psi_{i+h} - \tilde{\gamma}(h) \psi_i \right) \left(Z_{t-i}^2 - Z_t^2 \right) =: V_1 + V_2 , \end{aligned} \tag{3.2}$$

where we used the fact that $\sum_i \psi_i (\psi_{i+h} - \tilde{\gamma}(h) \psi_i) = 0$. Thus we obtain that

$$\max_{n \in [n_k, n_{k+1}]} |V_1| \leq \sum_{t=1}^{n_{k+1}} \sum_{i \neq j} \left| \psi_i \left(\psi_{j+h} - \tilde{\gamma}(h) \psi_j \right) \right| \left| Z_{t-i} Z_{t-j} \right| .$$

By (A1), (A2) and (A5) we obtain for all $\varepsilon > 0$

$$\begin{aligned} \sum_{k=1}^{\infty} P \left(\max_{n \in [n_k, n_{k+1}]} |V_1| > \varepsilon e_{n_k}^2 \right) &\leq c_1 \sum_{k=1}^{\infty} e_{n_k}^{-2\delta} E \max_{n \in [n_k, n_{k+1}]} |V_1|^\delta \\ &\leq c_2 \sum_{k=1}^{\infty} e_{n_k}^{-2\delta} n_{k+1} < \infty \end{aligned}$$

for some $c_1, c_2 > 0$ depending on ε . A Borel-Cantelli argument yields

$$\lim_{k \rightarrow \infty} \max_{n \in [n_k, n_{k+1}]} |V_1| e_n^{-2} = 0 \quad \text{a.s.}$$

Now to estimate V_2 set

$$f_i = \psi_i (\psi_{i+h} - \tilde{\gamma}(h) \psi_i) .$$

Then

$$\begin{aligned} V_2 &= \sum_{i>0} f_i \sum_{t=1}^n (Z_{t-i}^2 - Z_t^2) + \sum_{i<0} f_i \sum_{t=1}^n (Z_{t-i}^2 - Z_t^2) \\ &= V_3 + V_4 . \end{aligned}$$

We restrict ourselves to show that $\lim_{k \rightarrow \infty} e_{n_k}^{-2} V_3 = 0$ a.s., the proof for $e_{n_k}^{-2} V_4$ is similar. We have

$$\begin{aligned} V_3 &= \sum_{i>n} f_i \sum_{t=1-i}^{n-i} Z_{t-i}^2 - \sum_{i>n} f_i \sum_{t=1}^n Z_t^2 + \sum_{1 \leq i \leq n} f_i \sum_{t=1-i}^0 Z_t^2 - \sum_{1 \leq i \leq n} f_i \sum_{t=n-i+1}^n Z_t^2 \\ &= V_5 - V_6 + V_7 - V_8 . \end{aligned}$$

We restrict ourselves to show that $\lim_{k \rightarrow \infty} e_{n_k}^{-2} V_7 = 0$ a.s., the proof for V_5, V_6 and V_8 is analogous. Again by (A1), (A2) and (A5) we have

$$\sum_{k=1}^{\infty} E \left| e_{n_k}^{-2} V_7 \right|^{\delta/2} \leq c \sum_{k=1}^{\infty} e_{n_k}^{-\delta} \sum_{i>0} |f_i|^{\delta/2} |i| < \infty$$

and a Borel-Cantelli argument yields the desired result. Similar arguments show that

$$e_{n_k}^{-2} \sum_{t=1}^{n_k} X_t^2 = e_{n_k}^{-2} \psi^2 \sum_{t=1}^{n_k} Z_t^2 + o(1) \quad \text{a.s.}$$

This, (3.2) and (A5) imply that

$$\begin{aligned} \tilde{\gamma}_{n,X}(h) - \tilde{\gamma}(h) &= \frac{\sum_{t=1}^n X_t X_{t+h} - \tilde{\gamma}(h) \sum_{t=1}^n X_t^2}{\sum_{t=1}^n X_t^2} - \frac{\sum_{t=n-h+1}^n X_t X_{t+h}}{\sum_{t=1}^n X_t^2} \\ &= \frac{V_1 + V_2}{\sum_{t=1}^n X_t^2} - \frac{\sum_{t=n-h+1}^n X_t X_{t+h}}{\sum_{t=1}^n X_t^2} = o(1) \quad \text{a.s.} \end{aligned} \quad (3.3)$$

In relation (3.3) we applied (A1)–(A5) and similar arguments as above to show that $a_n^{-2} \sum_{t=n-h+1}^n X_t X_{t+h} \rightarrow 0$ a.s. for every $h > 0$. \square

4. Parameter estimation for ARMA(p, q) processes

We consider a causal invertible ARMA(p, q) process $(X_t)_{t \in \mathbb{Z}}$ satisfying for every t the ARMA equations

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \dots + \theta_q Z_{t-q}$$

for i.i.d. $(Z_t)_{t \in \mathbb{Z}}$. Denote

$$\beta = \left(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q \right)^T$$

and

$$\begin{aligned} \varphi(z, \beta) &= 1 - \varphi_1 z - \dots - \varphi_p z^p, \\ \theta(z, \beta) &= 1 + \theta_1 z + \dots + \theta_q z^q. \end{aligned}$$

Then the transfer function of the ARMA process has representation $\psi(e^{-i\lambda}, \beta) \equiv \varphi(e^{-i\lambda}, \beta) / \theta(e^{-i\lambda}, \beta)$.

We introduce the parameter set

$$C = \left\{ \beta \in \mathcal{R}^{p+q}; \varphi_p \neq 0, \theta_q \neq 0, \varphi(z) \text{ and } \theta(z) \text{ have no common zeros,} \right. \\ \left. \varphi(z) \theta(z) \neq 0 \text{ for } |z| \leq 1 \right\}.$$

Denote by $g(\lambda, \beta)$ the power transfer function corresponding to $\beta \in C$; i.e.

$$g(\lambda, \beta) = \left| \frac{\varphi(e^{-i\lambda}, \beta)}{\theta(e^{-i\lambda}, \beta)} \right|^2 = \left| \psi(e^{-i\lambda}, \beta) \right|^2,$$

and define

$$\sigma_n^2(\beta) = \int_{-\pi}^{\pi} \frac{\tilde{I}_{n,X}(\lambda)}{g(\lambda, \beta)} d\lambda, \quad \bar{\sigma}_n^2(\beta) = \frac{2\pi}{n} \sum_j \frac{\tilde{I}_{n,X}(\lambda_j)}{g(\lambda_j; \beta)},$$

where the sum is taken over all Fourier frequencies

$$\lambda_j = \frac{2\pi j}{n} \in (-\pi, \pi].$$

Suppose $\beta_0 \in C$ is the true, but unknown parameter vector. Then two natural estimators of β_0 are given by

$$\beta_n = \underset{\beta \in C}{\operatorname{argmin}} \sigma_n^2(\beta), \quad \bar{\beta}_n = \underset{\beta \in C}{\operatorname{argmin}} \bar{\sigma}_n^2(\beta).$$

Given the assumption that $\sigma_n^2(\beta) \sim \bar{\sigma}_n^2(\beta)$, it seems reasonable to assume, as is in fact the case, that $\beta_n \sim \bar{\beta}_n$, and that therefore the two estimators are asymptotically equivalent. It is clear that, in practice, $\bar{\beta}_n$ is the only applicable estimator, since the integral defining $\sigma_n^2(\beta)$ will always have to be evaluated by an approximating sum.

The choice of these estimators is motivated by the fact that the function

$$\int_{-\pi}^{\pi} \frac{g(\lambda, \beta_0)}{g(\lambda, \beta)} d\lambda$$

has its absolute minimum at $\beta = \beta_0$ in C . (cf. Brockwell and Davis (1991), Proposition 10.8.1.) and that $\tilde{I}_{n,X}(\lambda)$ can be considered as an estimator of $g(\lambda, \beta_0)$.

For Gaussian $(X_t)_{t \in \mathbb{Z}}$ the estimator β_n is closely related to least squares and maximum likelihood estimators and it is a standard estimator for ARMA processes with finite variance. The idea goes back to Whittle (1953), see also Dzhaparidze (1986), Fox and Taquq (1986) and Dahlhaus (1989). It is well-known that in the classical case β_n is consistent and asymptotically normal (cf. Brockwell and Davis (1991)). We showed in Mikosch et al. (1994) that β_n is also for ARMA processes with infinite variance a weakly consistent estimator for the true parameter vector β_0 (see also Gadrich (1993) for the case $\bar{\beta}_n$):

Proposition 4.1. *Suppose $(X_t)_{t \in \mathbb{Z}}$ is a causal invertible ARMA(p, q) process and conditions (A1)-(A4) hold. Then*

$$\beta_n \xrightarrow{P} \beta_0 \quad \text{and} \quad \sigma_n^2(\beta_n) \xrightarrow{P} 2\pi\psi^{-2}(\beta_0), \quad n \rightarrow \infty.$$

Furthermore, the same limit relationships hold also for $\bar{\beta}_n$ and $\bar{\sigma}_n^2$. \square

For ARMA(p, q) processes with finite variance β_n is asymptotically normal with rate of convergence of order $n^{1/2}$. An analogous result gives in the case $Z \in DNA(\alpha), \alpha < 2$, a rate of convergence of order $(n/\ln n)^{1/\alpha}$: i.e. the convergence is considerably faster. This is the main result in Mikosch et al. (1994):

Proposition 4.2. *Suppose $(X_t)_{t \in \mathbb{Z}}$ is an ARMA(p, q) process and $(Z_t)_{t \in \mathbb{Z}}$ are i.i.d. symmetric such that $Z \in DNA(\alpha)$ holds for some $\alpha < 2$. Then*

$$\left(\frac{n}{\ln n}\right)^{1/\alpha} (\beta_n - \beta_0) \xrightarrow{d} 4\pi W^{-1}(\beta_0) \frac{1}{Y_0} \sum_{k=1}^{\infty} Y_k b_k, \quad (4.1)$$

where Y_0, Y_1, Y_2, \dots are independent r.v.'s, Y_0 is positive $\alpha/2$ -stable, $(Y_t)_{t \in \mathcal{N}}$, are i.i.d. symmetric α -stable, $W^{-1}(\beta_0)$ is the inverse of the matrix

$$W(\beta_0) = \int_{-\pi}^{\pi} \left[\frac{\partial \ln g(\lambda, \beta_0)}{\partial \beta} \right] \left[\frac{\partial \ln g(\lambda, \beta_0)}{\partial \beta} \right]^T d\lambda,$$

and, for $k \in \mathcal{N}$, b_k is the vector

$$b_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\lambda} g(\lambda, \beta_0) \frac{\partial (1/g(\lambda, \beta_0))}{\partial \beta} d\lambda.$$

Furthermore, (4.1) holds also with β_n replaced by $\bar{\beta}_n$ \square

The limit vector in (4.1) is the ratio of an α -stable $(p+q)$ -dimensional vector over a positive $\alpha/2$ -stable r.v. It is not difficult to see that for $AR(p)$ processes β_n is just the formal analogue of the Yule-Walker estimates. Their weak limit behaviour was derived by Davis and Resnick (1986) using time domain methods.

Reconsidering the proof of the strong consistency of the Whittle estimate in the classical case (see Brockwell and Davis (1991), Chapter 10.8) we see that this result only depends on the strong consistency of the sample autocovariances (equivalently, sample autocorrelations). Thus the same proof applies in case that the assumptions of Proposition 3.2. are satisfied, but now we have to restrict ourselves to convergence along the subsequence (n_k) defined in (A5):

Proposition 4.3. *Suppose $(X_t)_{t \in \mathbb{Z}}$ satisfies (A1), (A2) and (A5). Then*

$$\beta_{n_k} \xrightarrow{a.s.} \beta_0 \quad \text{and} \quad \sigma_{n_k}^2(\beta_{n_k}) \xrightarrow{a.s.} 2\pi\psi^{-2}(\beta_0), \quad k \rightarrow \infty.$$

Furthermore, the same limit relationships hold also for $\bar{\beta}_{n_k}$ and $\bar{\sigma}_{n_k}^2$. \square

References

- BINGHAM, N.H., GOLDIE, C.M. AND TEUGELS, J.L (1987). *Regular Variation*. Cambridge University Press.
- BROCKWELL, P.J. AND DAVIS, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed., Springer, Berlin.
- DAHLHAUS, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17**, 1749-1766.
- DAVIS, R.A. AND RESNICK, S.I. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**, 533-558.
- DZHAPARIDZE, K. (1986). *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*. Springer, New York.
- FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*. Vol. II, 2nd ed., Wiley, New York.
- FOX, R. AND TAQQU, M.S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14**, 517-532.
- FRISTEDT, B.E. AND PRUITT, W.E. (1971). Lower functions for increasing random walks and subordinators. *Z. Wahrsch. verw. Gebiete* **18**, 167-182.
- GADRICH, T. (1993). Parameter Estimation for ARMA Processes with Symmetric Stable Innovations, D.Sc. Thesis, Technion. (In Hebrew)
- GRIFFIN, P.S. (1983). Probability estimates for the small deviations of d -dimensional random walk. *Ann. Probab.* **11**, 939-952.

- KLÜPPELBERG, C. AND MIKOSCH, T. (1993). Spectral estimates and stable processes. *Stoch. Proc. Appl.* **47**, 323-344.
- KLÜPPELBERG, C. AND MIKOSCH, T. (1994). Some limit theory for the self-normalized periodogram of stable processes. *Scand. J. Statist.* To appear.
- MALLER, R.A. (1981). Some properties of stochastic compactness. *J. Austral. Math. Soc. (Series A)* **30**, 264-277.
- MIKOSCH, T., GADRICH, T., KLÜPPELBERG, C., ADLER, R. (1994) Parameter estimation for ARMA processes with infinite variance innovations. *Ann. Statist.* To appear.
- PETROV, V.V. (1975). *Sums of Independent Random Variables*. Springer, New York.
- PRUITT, W.E. (1990). The rate of escape of random walk. *Ann. Probab.* **18**, 1417-1461.
- ROSINSKI, J. AND WOYCZYNSKI, W.A. (1987). Multilinear forms in Pareto-like random variables and product random measures. *Colloquium Mathematicum*, Vol. 51, 303-313.
- STOUT, W.A. (1974). *Almost Sure Convergence*. Academic Press, New York.
- WHITTLE, P. (1953). Estimation and information in stationary time series. *Ark. Mat.* **2**, 423-434.

Claudia Klüppelberg	Thomas Mikosch
Departement Mathematik	Institute of Statistics and Operations Research
ETH Zürich	Victoria University
CH - 8092 Zürich	Wellington
SWITZERLAND	NEW ZEALAND