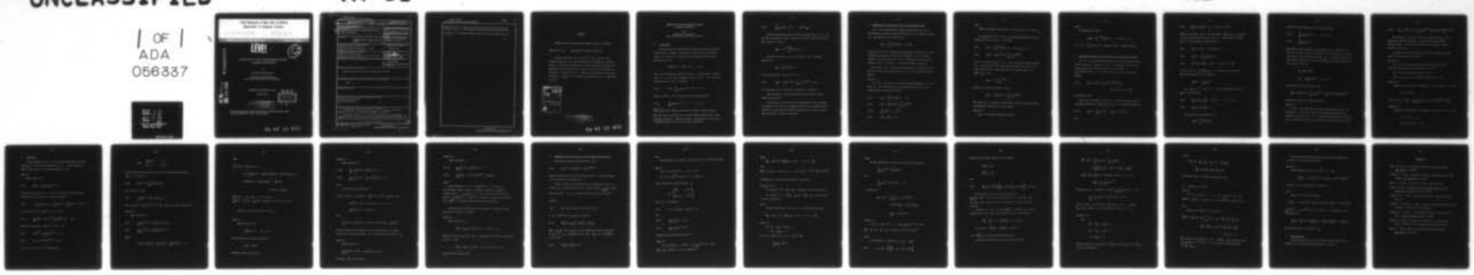
AD-A056 337 STATE UNIV OF NEW YORK AT BUFFALO AMHERST STATISTICA--ETC F/G 12/1
CONSISTENCY OF THE AUTOREGRESSIVE METHOD OF DENSITY ESTIMATION.(U)
MAR 78 J CARMICHAEL N00014-75-C-0734

UNCLASSIFIED TR-61 NL

1 OF 1
ADA
056337



END
DATE
FILMED
8-78
DDC

NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

State University of New York at Buffalo
Department of Computer Science
STATISTICAL SCIENCE
4230 Ridge Lea Road, Room A33; Amherst, New York 14226 Telephone: 716-831-1232

LEVEL II

(12)

AD A056337

CONSISTENCY OF THE AUTOREGRESSIVE METHOD OF DENSITY ESTIMATION*

by

Jean-Pierre Carmichael

Statistical Science Division
State University of New York at Buffalo

TECHNICAL REPORT NO. 61

March 1978

*Research supported in part by the Office of Naval Research under Contract N00014-75-C-0734 (NR 042-234).

78 07 12 011

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. 61 (TR-61) (14)	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) Consistency of the Autoregressive Method of Density Estimation.		5. TYPE OF REPORT & PERIOD COVERED: (9) Technical Rept.
		6. PERFORMING ORG. REPORT NUMBER (15)
7. AUTHOR(s): (10) Jean-Pierre/Carmichael		8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0734
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Statistical Science Division; State University of New York at Buffalo; Amherst, New York 14226		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 042-234
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research; Statistics and Probability Program Code 436; Arlington, Virginia 22217		12. REPORT DATE: (11) Mar [redacted] 78
		13. NUMBER OF PAGES: (12) 25 p. 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Autoregressive Schemes

Density Estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consistency of the autoregressive method of density estimation.
Abbreviated Title: Autoregressive Estimator of density.

A density function $f(\cdot)$, with $1/f^{-1}(\cdot)$ and $\log f(\cdot)$ both Lebesgue-integrable, has a representation as an autoregressive spectral density. We use this representation to obtain new density autoregressive estimators of (is used)

(continued on page 2)

409 511

different orders p based on the empirical characteristic function of a sample of size n . We prove the consistency of these new estimators as $\lim_{n\to\infty} p^3 n^{-1} = 0$ under varying conditions on the smoothness of $f(\cdot)$.

SUMMARY

Consistency of the autoregressive method of density estimation.

abbreviated title: autoregressive estimator of density.

A density function $f(\cdot)$, with $f^{-1}(\cdot)$ and $\log f(\cdot)$ both Lebesgue-integrable, has a representation as an autoregressive spectral density. We use this representation to obtain new density autoregressive estimators of different orders p based on the empirical characteristic function of a sample of size n. We prove the consistency of these new estimators as $\lim_{n\to\infty} p^3 n^{-1} = 0$ under varying conditions on the smoothness of $f(\cdot)$.

ACCESSION for		
NTIS	White Section	☒
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL. and/or SPECIAL	
A		

CONSISTENCY OF THE AUTOREGRESSIVE METHOD OF DENSITY ESTIMATION

by

Jean-Pierre Carmichael[1]

State University of New York at Buffalo

1. Introduction

The autoregressive method has been used so far only in the context of time series. Consider, for instance, a discrete time real process $\{X(t), t \in Z\}$ (where Z is the set of all integers) with stationary covariance function

$$\mathrm{Cov}\Big(X(t), X(t+v)\Big) = R(v) , \quad v \in Z .$$

$X(\cdot)$ is an autoregressive process of *order* p *if there exist a sequence* $\{\alpha_{jp}, j = 1,\ldots,p\}$ and an orthogonal process $\{\eta(t) , t \in Z\}$ with mean zero and variance $\sigma_\eta^2 > 0$ such that

$$X(t) + \sum_{j=1}^{p} \alpha_{jp} X(t-j) = \eta(t) , \quad t \in Z . \tag{1.1}$$

When (1.1) holds, $R(\cdot)$ satisfies the Yule-Walker equations

$$\sum_{\ell=0}^{p} \alpha_{\ell p} R(\ell - j) = 0 , \qquad j = 1,\ldots,p \tag{1.2}$$

[1]This research was supported in part by the Office of Naval Research

AMS 1970 subject classifications. Primary 62G05; Secondary 42A52, 60F05.

Key words and phrases. Density estimation, autoregressive representation, orthogonal polynomials on the unit circle, consistency.

$$(1.3) \qquad \sum_{\ell=0}^{p} \alpha_{\ell p} R(\ell) = \sigma_\eta^2 , \qquad \text{with } \alpha_{0p} = 1 .$$

In the autoregressive method, we obtain estimates $\{\hat{\alpha}_{jp}, j = 1,\ldots,p\}$ and $\hat{\sigma}_\eta^2$ from a sample $\{X(1),\ldots,X(T)\}$ by solving (1.2) and (1.3) with $R(\cdot)$ replaced by $R_T(\cdot)$,

$$R_T(v) = T^{-1} \sum_{t=1}^{T-|v|} X(t)\, X(t+v) .$$

The spectral density $f(\cdot)$ of the process $X(\cdot)$ is defined implicitly by

$$R(v) = \int_{-\pi}^{\pi} e^{ivx} f(x)\, dx .$$

For an autoregressive process of order p,

$$(1.4) \qquad f_p(x) = \sigma_\eta^2 (2\pi)^{-1} \left| 1 + \sum_{j=1}^{p} \alpha_{jp} e^{ijx} \right|^{-2}, \ x \in [-\pi,\pi] .$$

It is estimated using the estimated autoregressive parameters.

The consistency of these procedures has been studied by Kromer (1969) and Berk (1974).

In this paper, we present the first application of the autoregressive method in the context of independent identically distributed random variables as a new method of density estimation and we study its convergence properties in terms of consistency.

2. Probabilistic interpretation of the autoregressive method.

Let X be a bounded random variable taking values in $[-\pi,\pi]$ with absolutely continuous distribution function $F(\cdot)$, density function $f(\cdot)$ and characteristic function $\phi(\cdot)$,

$$\phi(v) = \int_{-\pi}^{\pi} e^{ivx} f(x)\, dx , \quad v \text{ real}$$

It follows that $\phi(\cdot)$ is Hermitian and nonnegative definite, i.e. $\phi(\cdot)$ satisfies the necessary and sufficient conditions to be a covariance function. It also follows that $f(\cdot)$ is completely determined by the sequence $\{\phi(v) , v = 0,1,2,\ldots\}$ (see Feller (1966), chapter 19).

We think of $\phi(\cdot)$ as the covariance function of a stationary complex time series (unobservable) and of $f(\cdot)$ as its associated spectral density.

Theorem 2.1

If $f^{-1}(\cdot)$ and $\log f(\cdot)$ are both Lebesque-integrable on $[-\pi,\pi]$, then $f(\cdot)$ can be represented as the spectral density $\tilde{f}(\cdot)$ of an infinite order autoregressive process:

$$\text{(2.1)} \qquad \phi(v) = \int_{-\pi}^{\pi} e^{ivx}\tilde{f}(x)\, dx , \quad v \in Z$$

$$\text{(2.2)} \qquad \tilde{f}(x) = \sigma_\infty^2 (2\pi)^{-1} \left|1 + \sum_{j=1}^{\infty} \alpha_j e^{ijx}\right|^{-2}$$

$$\text{(2.3)} \qquad \sum_{j=1}^{\infty} |\alpha_j|^2 < \infty , \quad \sigma_\infty^2 > 0$$

$$\text{(2.4)} \qquad f(x) = \tilde{f}(x) , \quad \text{a.e.}$$

Proof.

Apply the argument of Doob (1953, p. 577) to $f(\cdot)$ and $f^{-1}(\cdot)$. □

We approximate $\tilde{f}(\cdot)$ by $\tilde{f}_p(\cdot)$, the spectral density of an autoregressive process of order p such that

$$\phi(v) = \int_{-\pi}^{\pi} e^{ivx}\tilde{f}_p(x)\,dx , \quad v = 0, \pm 1, \ldots, \pm p \tag{2.5}$$

$$\tilde{f}_p(x) = \sigma_p^2 (2\pi)^{-1} \left|1 + \sum_{j=1}^{p} \alpha_{jp} e^{ijx}\right|^{-2} . \tag{2.6}$$

Now, we can estimate $\{\alpha_{jp} , j = 1,\ldots,p\}$ and σ_p^2 for different values of p as in Section 1 by solving the Yule-Walker equations with $\phi(\cdot)$ being estimated by the sample characteristic function $\phi_n(\cdot)$ of a sample $\{X_1,\ldots,X_n\}$

$$\phi_n(v) = n^{-1} \sum_{k=1}^{n} e^{ivX_k} .$$

Finally, the estimated density is $\hat{f}_p(\cdot)$,

$$\hat{f}_p(x) = \hat{\sigma}_p^2 (2\pi)^{-1} \left|1 + \sum_{j=1}^{p} \hat{\alpha}_{jp} e^{ijx}\right|^{-2}$$

Note that $\phi_n(\cdot)$ is usually a complex-valued function and the estimated autoregressive coefficients are also complex.

Theorem 2.2

$\hat{f}_p(\cdot)$ is a probability density function.

Proof.

By definition of $\hat{f}_p(\cdot)$,

$$\phi_n(v) = \int_{-\pi}^{\pi} e^{ivx} \hat{f}_p(x)\, dx \quad , \quad v = 0, \pm 1, \ldots, \pm p \quad .$$

For $v = 0$, $\int_{-\pi}^{\pi} \hat{f}_p(x)\, dx = 1$. Finally, $\hat{\sigma}_p^2 > 0$ (Pagano (1973)). □

3. Approximation theory interpretation of the autoregressive method.

Let $F(\cdot)$ be a nondecreasing bounded function with infinitely many points of increase, defined on $[-\pi,\pi]$. We denote by L_F^2 the space of measurable complex-valued functions $u(\cdot)$ such that $\int_{-\pi}^{\pi} |u(e^{ix})|^2\, dF(x) < \infty$. It is well known that L_F^2 with the inner product

$$\text{(3.1)} \qquad \Big(u(\cdot), v(\cdot)\Big)_F = \int_{-\pi}^{\pi} u(e^{ix})\, \overline{v(e^{ix})}\, dF(x)$$

$u(\cdot)$ and $v(\cdot)$ in L_F^2

is a Hilbert space.

From the set of powers $\{1,z,z^2,\ldots\}$ in L_F^2 , we obtain a set of orthonormal polynomials $\{g_0(\cdot),g_1(\cdot),g_2(\cdot),\ldots\}$ uniquely determined by

$$\text{(3.2)} \qquad g_\ell(z) = \sum_{j=0}^{\ell} a_{j\ell}\, z^{\ell-j} \quad , \quad a_{0\ell} > 0 \text{ , for all } \ell$$

and

(3.3) $$\left(g_j(\cdot),\, g_k(\cdot)\right)_F = \delta_{jk}\,, \quad \text{for all } j \text{ and } k\,.$$

Consider the subspace L_p^2 of L_F^2 generated by $\left(g_0(\cdot), g_1(\cdot), \ldots, g_p(\cdot)\right)$. It is a reproducing kernel Hilbert space, i.e. there exists a function $K_p(\cdot,\cdot)$ of two complex variables such that

(3.4) $$K_p(\cdot,y) \in L_p^2\,, \quad \text{for any fixed } y$$

(3.5) $$K_p(\cdot,y) = \sum_{j=0}^{p} k_{jp}(y)\, g_j(\cdot)$$

(3.6) $$\left(u(\cdot),\, K_p(\cdot,y)\right)_F = u(y)\,, \quad \text{for all } u(\cdot) \in L_p^2\,.$$

In fact, $K_p(\cdot,y) = \sum_{j=0}^{p} \overline{g_j(y)}\, g_j(\cdot)$. We restrict our attention to $K_p(\cdot,0)$ and express it as a polynomial

$$K_p(z,0) = \sum_{j=0}^{p} b_{jp}\, z^j$$

Let $u_j(z) = z^j$, $j = 0,1,\ldots,p$. By the reproducing property of $K_p(\cdot,\cdot)$, we have that

(3.7) $$\left(u_j(\cdot),\, K_p(\cdot,0)\right)_F = u_j(0) = 0\,, \quad j = 1,\ldots,p$$

(3.8) $$\left(u_0(\cdot),\, K_p(\cdot,0)\right)_F = 1\,.$$

If we introduce the notation $\phi(\cdot)$

$$\phi(v) = \int_{-\pi}^{\pi} e^{ivx}\, dF(x)$$

then the system of equations (3.7) and (3.8) becomes

(3.9) $$\sum_{\ell=0}^{p} \overline{b}_{\ell p}\, \phi(j-\ell) = 0 \ , \qquad j = 1,\dots,p$$

(3.10) $$\sum_{\ell=0}^{p} \overline{b}_{\ell p}\, \phi(-\ell) = 1 \ .$$

Upon taking complex conjugates and identifying $\phi(\cdot)$ with $R(\cdot)$, we see that (3.9) and (3.10) are equivalent to (1.2) and (1.3) up to a constant factor, the difference being that $\alpha_{0p} = 1$. Thus we divide (3.9) and (3.10) by $K_p(0,0)$ and make the following identification between the two systems of equations:

$$\sigma_\eta^2 = \left(K_p(0,0)\right)^{-1}$$

$$\alpha_{jp} = b_{jp}\left(K_p(0,0)\right)^{-1} \ , \qquad j = 0,\dots,p$$

from which identification follows that

$$\tilde{f}_p(x) = \sigma_\eta^2 (2\pi)^{-1} \left|1 + \sum_{j=1}^{p} \alpha_{jp} e^{ijx}\right|^{-2} = K_p(0,0)(2\pi)^{-1}\left|K_p(e^{ix},0)\right|^{-2} .$$

We have thus proved the following theorem.

Theorem 3.1

If $F(\cdot)$ is an absolutely continuous distribution function, $f(\cdot)$ its density, $f^{-1}(\cdot)$ and $\log f(\cdot)$ are Lebesgue-integrable and we replace in (3.9) and (3.10) $\phi(\cdot)$ by the sample characteristic function $\phi_n(\cdot)$ as defined in Section 2, we obtain two representations for $\hat{f}_p(\cdot)$, (3.11) $$\hat{f}_p(x) = \hat{\sigma}_p^2(2\pi)^{-1}\left|1 + \sum_{j=1}^{p} \hat{\alpha}_{jp}e^{ijx}\right|^2 = \hat{K}_p(0,0)(2\pi)^{-1}\left|\hat{K}_p(e^{ix},0)\right|^2 .$$

The properties of $K_p(\cdot,0)$ have been studied extensively in the approximation theory literature, e.g. Grenander and Szegö (1958), Geronimus (1961). We quote known results from these sources without proof as we need them.

We study the convergence properties of the autoregressive method under the following sets of conditions:

Conditions A :

$F(\cdot)$ is an absolutely continuous distribution function with infinitely many points of increase, defined on $[-\pi,\pi]$.

$f(\cdot)$ is the corresponding density function.

$f^{-1}(\cdot)$ and $\log f(\cdot)$ are Lebesgue-integrable on $[-\pi,\pi]$.

Conditions B :

Conditions A are satisfied. Furthermore, $0 < m \le f(x) \le M < \infty$, a.e.

$$f(\cdot) \in \text{Lip}\ (\tfrac{1}{2},2)$$

where $$\text{Lip}\ (\alpha,2) = \left\{u(\cdot) : \sup_{|h|<\delta}\left(\int_{-\pi}^{\pi} |u(x+h) - u(x)|^2\, dx\right)^{\frac{1}{2}} = O(\delta^{\alpha})\right\} .$$

Conditions C :

Conditions A are satisfied. Furthermore, $0 < m \le f(x) \le M < \infty$, a.e.

$$f(\cdot) = d(\cdot)\ , \quad \text{a.e.}$$

$$d(\cdot) \in \text{Lip}\ (\alpha,2)\ , \quad \alpha > \tfrac{1}{2}$$

4. Bias study

In the estimation of $f(\cdot)$, we have used an approximate representation $\tilde{f}_p(\cdot)$ that we have estimated by $\hat{f}_p(\cdot)$. In this section we study how good $\tilde{f}_p(\cdot)$ is as an approximation to $f(\cdot)$.

Lemma 4.1

Define $g_p^*(\cdot)$ by

$$(4.1) \qquad g_p^*(z) = K_p(z,0)\, K_p(0,0)^{-\frac{1}{2}} .$$

The Lebesgue-integrability of $\log f(\cdot)$ is a necessary and sufficient condition for the existence of the following limits

$$(4.2) \qquad 0 < \lim_{p\to\infty} K_p(0,0) = \lim \sum_{j=0}^{p} |g_j(0)|^2 = \sum_{j=0}^{\infty} |g_j(0)|^2 = K(0,0) < \infty ;$$

(we will now use $K_p = K_p(0,0)$ and $K = K(0,0)$)

$$(4.3) \qquad \lim_{p\to\infty} g_p^*(z) = g(z) = K^{-\frac{1}{2}} \sum_{j=1}^{\infty} \overline{g_j(0)}\, g_j(z) , \quad |z| < 1$$

where the convergence is uniform in $|z| \le r < 1$;

$$(4.4) \qquad g(e^{ix})^{-1} = \lim_{r\to 1^-} g(re^{ix})^{-1} , \quad \text{a.e.};$$

$$(4.5) \qquad f(x) = (2\pi)^{-1} |g(e^{ix})|^{-2} , \quad \text{a.e.};$$

for $E = \{x : 0 < f(x) < \infty\}$, define $\tilde{g}(\cdot)$,

$$\tilde{g}(x) = \begin{cases} g(e^{ix}) & , \quad x \in E \\ 0 & , \quad x \notin E \end{cases} .$$

$\tilde{g}$ has the following expansion in terms of the orthonormal polynomials $\{g_j(\cdot) \ , \ j = 0,1,2,\dots\}$

$$(4.6) \qquad \tilde{g}(x) \sim K^{-\frac{1}{2}} \sum_{j=0}^{\infty} \overline{g_j(0)}\, g_j(e^{ix})$$

that converges in L_F^2 ;

$$(4.7) \qquad \lim_{p\to\infty} \|\tilde{g}(\cdot) - K_p^{\frac{1}{2}} K^{-\frac{1}{2}} g_p^*(\cdot)\|_F = 0 .$$

where $\|u(\cdot)\|_F = \left(\int_{-\pi}^{\pi} |u(e^{ix})|^2 f(x)\, dx\right)^{\frac{1}{2}}$ (Geronimus (1961), Chapter II).

Theorem 4.2

Under Conditions A ,

$$(4.8) \qquad \lim_{p\to\infty} \int_{-\pi}^{\pi} |f^{-1}(x) - \tilde{f}_p^{-1}(x)|\, f(x)\, dx = 0$$

$$(4.9) \qquad \lim_{p\to\infty} \int_{-\pi}^{\pi} |f(x) - \tilde{f}_p(x)|\, \tilde{f}_p^{-1}(x)\, dx = 0$$

Proof:

$$|f^{-1}(x) - \tilde{f}_p^{-1}(x)| = 2\pi \left| |g(e^{ix})|^2 - |g_p^*(e^{ix})|^2 \right| \quad , \quad \text{a.e.}$$

Thus,

$$\int_{-\pi}^{\pi} |f^{-1}(x) - \tilde{f}_p^{-1}(x)| \, f(x) \, dx$$

$$\leq 2\pi \int_{-\pi}^{\pi} \left(|g(e^{ix})| + |g_p^*(e^{ix})|\right)\left(|g(e^{ix}) - g_p^*(e^{ix})|\right) f(x) \, dx$$

$$\leq 2\pi \left(\|g(\cdot)\|_F + \|g_p^*(\cdot)\|_F\right)\|g(\cdot) - g_p^*(\cdot)\|_F$$

by Schwarz inequality.

$\|g_p^*(\cdot)\|_F = 1$, $\|g(\cdot)\|_F = 1$ (by (4.6)) and
$\lim_{p\to\infty} \|g(\cdot) - g_p^*(\cdot)\|_F = \lim_{p\to\infty} \|\tilde{g}(\cdot) - g_p^*(\cdot)\|_F = 0$ by (4.7 and (4.2) .

Finally, (4.9) is equivalent to (4.8). □

Lemma 4.3

Under Conditions B ,

$$|g_j^*(z)| \leq C \quad , \quad \text{for} \quad |z| \leq 1 \quad .$$

We can replace the Lipschitz condition by

$$\phi(v) = O(v^{-1}) \quad .$$

(Geronimus (1961), Theorem 3.8).

Theorem 4.4

Under Conditions B ,

(4.10) $$\lim_{p\to\infty} \int_{-\pi}^{\pi} |f^{-1}(x) - \tilde{f}_p^{-1}(x)|^2 \, dx = 0$$

(4.11) $$\lim_{p\to\infty} \int_{-\pi}^{\pi} |f(x) - \tilde{f}_p(x)|^2 \, \tilde{f}_p^{-2}(x) \, dx = 0$$

Proof.

As in the proof of Theorem 4.2,

$$|f^{-1}(x) - \tilde{f}_p^{-1}(x)| \le 2\pi \; |g(e^{ix})| + |g_p^*(e^{ix})| \;\; |g(e^{ix}) - g_p^*(e^{ix})| \text{ , but}$$

$$|g(e^{ix})| \le m^{-\frac{1}{2}} \text{ a.e., as } f(\cdot) \ge m > 0 \text{ , a.e.}$$

$$|g_p^*(e^{ix})| \le C \text{ , by Lemma 4.3.}$$

Thus,

$$\int_{-\pi}^{\pi} |f^{-1}(x) - \tilde{f}_p^{-1}(x)|^2 \, dx \le 4\pi^2 (m^{-\frac{1}{2}} + C)^2 \, m^{-1} \, \|g(\cdot) - g_p^*(\cdot)\|_F^2$$

and the right-hand side converges to zero as in Theorem 4.2. Again, (4.10) and (4.11) are equivalent, $f(\cdot)$ being bounded from below a.e. □

Lemma 4.5

Under Conditions C ,

$$\lim_{p\to\infty} g_p^*(e^{ix}) = g(e^{ix}) \text{ , uniformly on } [-\pi, \pi] \text{ .}$$

(Geronimus (1961), Theorem 5.2).

Theorem 4.6

Under Conditions C ,

(4.12) $$\lim_{p\to\infty} \int_{-\pi}^{\pi} |f(x) - \tilde{f}_p(x)|^2 \, dx = 0$$

(4.13) $$\lim_{p\to\infty} \tilde{f}_p(x) = (2\pi)^{-1} |g(e^{ix})|^{-2} \quad \text{uniformly.}$$

Proof.

Under Conditions C , $0 < a \le |g_p(e^{ix})| \le A < \infty$ for any p (see Ibragimov (1964), Lemma 5), and because $|g_p(e^{ix})| = |g_p^*(e^{ix})|$, it follows that $0 < b \le \tilde{f}_p(x) \le B < \infty$ and (4.12) follows from (4.10). Finally, $\lim_{p\to\infty} |g_p^*(e^{ix})|^{-2} = |g(e^{ix})|^{-2}$ uniformly by Lemma 4.5 and the fact that $|g_p^*(\cdot)|$ is uniformly bounded from above and below. □

Having obtained uniform pointwise convergence, we state the following rates of decrease of the bias.

Theorem 4.7

Under Conditions C ,

$$|f(x) - \tilde{f}_p(x)| = O(p^{-\beta}) \quad \text{a.e.,} \quad 0 \le \beta < \alpha - \tfrac{1}{2}$$

Moreover if the function $d(\cdot)$ has r derivatives and $d^{(r)}(\cdot) \in \text{Lip}(\alpha,2)$, $0 < \alpha \le 1$, then

$$|f(x) - \tilde{f}_p(x)| = O(p^{-\beta}) , \quad \text{a.e.,} \quad 0 \le \beta < r + \alpha - \tfrac{1}{2} .$$

(Kromer (1969), Theorem 3.12).

5. Consistency of the estimators of the autoregressive parameters.

We use the following representation for $\hat{f}_p(\cdot)$,

$$\hat{f}_p(x) = (2\pi \hat{K}_p)^{-1} \left| 1 + \Sigma \hat{\alpha}_{jp} e^{ijx} \right|^{-2} , \tag{5.1}$$

where the estimation is based on a sample of size n . We also consider p as a function of the sample size n .

In what follows, the Euclidean norm of a p-dimensional vector $\underset{\sim}{x}_p$ or matrix X_p is represented by $\|\cdot\|$, whereas $\|X_p\|_H = \sup_{\|\underset{\sim}{x}_p\|=1} \|X_p \cdot \underset{\sim}{x}_p\|$. Also, the symbol ^ on $\hat{\underset{\sim}{x}}_p$ indicates that each element is estimated.

Lemma 5.1

$$\|X_p \cdot \underset{\sim}{x}_p\| \le \|X_p\|_H \cdot \|\underset{\sim}{x}_p\| \le \|X_p\| \cdot \|\underset{\sim}{x}_p\| \quad ; \tag{5.2}$$

if X_p is Hermitian and nonnegative definite,

$$\|X_p\|_H = \lambda_{max}(X_p) , \quad \|X_p^{-1}\|_H = \lambda_{min}^{-1}(X_p) \tag{5.3}$$

where $\lambda_{max}(X_p)$ and $\lambda_{min}(X_p)$ are the maximum and minimum eigenvalues of X_p ; if Y_p is nonsingular and if $\|X_p - Y_p\|_H \le (1-\varepsilon)\cdot\|Y_p^{-1}\|_H^{-1}$, $\varepsilon > 0$, then

$$\|X_p^{-1}\|_H \le \|Y_p^{-1}\|_H \cdot \varepsilon^{-1} \tag{5.4}$$

Proof.

These results are well-known. For a proof of (5.4) see Davies (1973).

Lemma 5.2

Let $0 < m \le f(x) \le M < \infty$, a.e. $[-\pi,\pi]$.

Let $\phi(v) = \int_{-\pi}^{\pi} e^{ivx} f(x)\, dx$, $v = 0, \pm 1, \pm 2, \ldots$

Define the Hermitian Toeplitz matrix R_p .

$$R_p = \begin{bmatrix} \phi(0) & \cdots & \phi(p-1) \\ \vdots & & \vdots \\ \phi(1-p) & \cdots & \phi(0) \end{bmatrix}$$

Then, R_p is nonsingular,

(5.5) $$2\pi\, m \le \lambda_{min}(R_p) \le \lambda_{max}(R_p) \le 2\pi\, M$$

and

(5.6) $$\lim_{p\to\infty} \lambda_{min}(R_p) = 2\pi\, m$$

(5.7) $$\lim_{p\to\infty} \lambda_{max}(R_p) = 2\pi\, M$$

(Grenander and Szegö (1958), Chapter 5).

Lemma 5.3

Let $\underset{\sim}{r}_p = \left(\phi(-1),\ldots,\phi(-p)\right)'$. If $\lim_{n\to\infty} p^2 n^{-1} = 0$, then,

$\|\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p\|$ converges to zero in probability.

Proof.

$$\|\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p\| \le \sqrt{p}\ \max\Big(|\phi_n(v) - \phi(v)|\ ;\quad v = -1,\dots,-p\Big)$$

and

$$P\Big(\sqrt{p}\ \max(|\phi_n(v) - \phi(v)|\ ; v = -1,\dots,-p) \le \varepsilon\Big) \ge 1 - p^2 n^{-1} \varepsilon^{-2}\Big(1 - |\phi(v)|^2\Big)$$

by Bonferroni's inequality and Chebyshev's inequality.

Corollary 5.3.1

If $\lim p^2 n^{-1} = 0$, $\|\hat{R}_p - R_p\|$ converges to zero in probability.

If $\lim p^3 n^{-1} = 0$, $p^{\frac{1}{2}}\|\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p\|$ and $p^{\frac{1}{2}}\|\hat{R}_p - R_p\|$ both converge to zero in probability.

Proof.

Just note that because $\hat{R}_p$ and R_p are Toeplitz and Hermitian,

$$\|\hat{R}_p - R_p\| \le 2\sqrt{p}\ \max\Big(|\phi_n(v) - \phi(v)|\ ,\quad v = -1,\dots,-p\Big) .$$

Lemma 5.4

Let $\underset{\sim}{\alpha}_p = (\alpha_{0p},\dots,\alpha_{pp},0,\dots)$

$\underset{\sim}{\alpha} = (\alpha_0,\alpha_1,\dots.)$.

If $0 < m \le f(x) \le M < \infty$, a.e. $[-\pi,\pi]$, then

$$\lim_{p\to\infty} \|\underset{\sim}{\alpha}_p - \underset{\sim}{\alpha}\| = 0$$

Proof.

By the representation introduced in Section 3 and Lemma 4.1,

$$\sum_{j=0}^{p} \alpha_{jp} e^{ijx} = K_p^{-\frac{1}{2}} g_p^*(e^{ix})$$

and

$$\sum_{j=0}^{\infty} \alpha_j e^{ijx} = K^{-\frac{1}{2}} g(e^{ix}) , \quad \text{a.e.}$$

it follows that

$$\|\underset{\sim}{\alpha}_p - \underset{\sim}{\alpha}\|^2 = \int_{-\pi}^{\pi} |\sum_{j=1}^{\infty} (\alpha_{jp} - \alpha_j) e^{ijx}|^2 dx$$

$$\leq m^{-1} K^{-1} \|\tilde{g}(\cdot) - K_p^{-\frac{1}{2}} K^{\frac{1}{2}} g_p^*(\cdot)\|_F$$

$$\underset{p\to\infty}{\longrightarrow} 0 , \quad \text{by Lemma 4.1.}$$

Theorem 5.5

If $0 < m \leq f(x) \leq M < \infty$, a.e. and $\lim_{n\to\infty} p^2 n^{-1} = 0$, then $\|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\|$ and $|\hat{K}_p - K_p|$ converge to zero in probability.

Proof.

It is sufficient to consider $\underset{\sim}{\alpha}_p = (\alpha_{1p},\ldots,\alpha_{pp})'$

(5.8) $$\underset{\sim}{\alpha}_p - \hat{\underset{\sim}{\alpha}}_p = \hat{R}_p^{-1}\left[(\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p) + (\hat{R}_p - R_p)\,\underset{\sim}{\alpha}_p\right]$$

because the Yule-Walker equations can be written

$$R_p \underset{\sim}{\alpha}_p = -\underset{\sim}{r}_p$$

$$\hat{R}_p \hat{\underset{\sim}{\alpha}}_p = -\hat{\underset{\sim}{r}}_p \quad .$$

Thus

$$(5.9) \qquad \|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\| \leq \|\hat{R}_p^{-1}\|_H \left[\|\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p\| + \|\hat{R}_p - R_p\| \cdot \|\alpha_p\|\right] \quad \text{by (5.2).}$$

We now bound each term on the right-hand side. $\lim_{p\to\infty} \|\underset{\sim}{\alpha}_p\| = \|\underset{\sim}{\alpha}\| < \infty$ as can be seen from Lemma 5.4. Both $\|\hat{\underset{\sim}{r}}_p - \underset{\sim}{r}_p\|$ and $\|\hat{R}_p - R_p\|$ converge to zero in probability by Lemma 5.3 and its Corollary.

Finally, we use (5.4) to bound $\|\hat{R}_p^{-1}\|_H$. Note that R_p is non-singular and because $\|\hat{R}_p - R_p\|$ converges to zero in probability,

$$\|\hat{R}_p - R_p\| \leq (1 - \varepsilon) \|R_p^{-1}\|_H^{-1} \leq (1 - \varepsilon)\, 2\pi M \quad ;$$

so, by (5.4), $\|\hat{R}_p^{-1}\|_H \leq \varepsilon \|\Upsilon_p^{-1}\|_H \leq \varepsilon (2\pi m)^{-1}$,

i.e. $\|\hat{R}_p^{-1}\|_H$ is bounded with probability one.

To prove the second part of the theorem, we note that

$$\hat{K}_p^{-1} - K_p^{-1} = \sum_{j=1}^{p} \hat{\alpha}_{jp}\,\phi_n(j) - \sum_{j=1}^{p} \alpha_{jp}\,\phi(j)$$

$$= \sum_{j=0}^{p} \left\{ \hat{\alpha}_{jp}(\phi_n(j) - \phi(j)) + \phi(j)(\hat{\alpha}_{jp} - \alpha_{jp}) \right\}$$

$$|\hat{K}_p^{-1} - K_p^{-1}| \le \|\hat{\underset{\sim}{\alpha}}_p\| \sqrt{p}\ \max(|\phi_n(j) - \phi(j)|\ ,\ j = 1,\dots,p)$$

$$+ \|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\| \left(\sum_{j=1}^{p} |\phi(j)|^2\right)^{\frac{1}{2}}$$

This goes to zero in probability provided $\lim_{p\to\infty} \sum_{j=1}^{p} |\phi(j)|^2 < \infty$. But

$$\lim_{p\to\infty} \sum_{j=1}^{p} |\phi(j)|^2 = \int_{-\pi}^{\pi} f^2(x)\,dx \le 2\pi M^2 < \infty$$

because $\{\phi(j)\}$ are the Fourier coefficients of $f(\cdot)$. Finally, $|\hat{K}_p - K_p|$ converges to zero in probability as $K_p < K < \infty$. □

Corollary 5.5.1

Let

$$\hat{\underset{\sim}{\alpha}}_p = (\hat{\alpha}_{1p},\dots,\hat{\alpha}_{pp},0,\dots)$$

$$\underset{\sim}{\alpha}_p = (\alpha_{1p},\dots,\alpha_{pp},0,\dots)$$

$$\underset{\sim}{\alpha} = (\alpha_1,\alpha_2,\dots)$$

Under the conditions of Theorem 5.5, $\|\hat{\alpha}_p - \alpha\|$ and $|\hat{K}_p - K|$ converge to zero in probability.

Proof.

$$|\hat{K}_p - K| \le |\hat{K}_p - K_p| + |K_p - K| \ .$$ Similarly

$$\|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}\| \le \|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\| + \|\underset{\sim}{\alpha}_p - \underset{\sim}{\alpha}\| .$$

Just apply Lemma 4.1, Lemma 5.4 and Theorem 5.5. □

6. Consistency of $\hat{f}_p(\cdot)$

Lemma 6.1

If $0 < m \le f(x) \le M < \infty$, a.e. and $\lim_{n\to\infty} p^3 n^{-1} = 0$, then $|\hat{g}_p^*(e^{ix}) - g_p^*(e^{ix})|$ converges to zero in probability uniformly in x.

Proof.

By the same technique as in Theorem 5.5

$$|\hat{g}_p^*(e^{ix}) - g_p^*(e^{ix})| \le \hat{K}_p^{1/2} \sum_{j=1}^{p} |\hat{\alpha}_{jp} - \alpha_{jp}| + |\hat{K}_p^{1/2} - K_p^{1/2}| \sum_{j=0}^{p} |\alpha_{jp}|$$

$$\le \hat{K}_p^{1/2} \sqrt{p}\, \|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\| + \sqrt{p}\, |\hat{K}_p^{1/2} - K_p^{1/2}| \cdot \|\underset{\sim}{\alpha}_p\| ,$$

for all x.

$\hat{K}_p^{1/2}$ converges in probability to $K^{1/2}$; $\sqrt{p}\, \|\hat{\underset{\sim}{\alpha}}_p - \underset{\sim}{\alpha}_p\|$ converges to zero in probability if $\lim p^3 n^{-1} = 0$ by Theorem 5.5, and the same for $\sqrt{p}\, |\hat{K}_p^{1/2} - K_p^{1/2}|$. □

We can now prove consistency theorems analogous to Theorems 4.2, 4.4 and 4.6. We give only one example.

Theorem 6.2

Under Conditions C and if $\lim p^3 n^{-1} = 0$, then

$$\left|\hat{f}_p^{-1}(x) - 2\pi\, |g(e^{ix})|^2\right| \quad \text{and} \quad \left|\hat{f}_p(x) - (2\pi)^{-1}\, |g(e^{ix})|^{-2}\right|$$

converge to zero in probability uniformly in x .

Proof.

We prove only the first statement.

$$\left|\hat{f}_p^{-1}(x) - 2\pi\, |g(e^{ix})|^2\right| \le \left|\hat{f}_p^{-1}(x) - \tilde{f}_p^{-1}(x)\right| + \left|\tilde{f}_p^{-1}(x) - 2\pi\, |g(e^{ix})|^2\right| .$$

$\lim_{p\to\infty} \left|\tilde{f}_p^{-1}(x) - 2\pi\, |g(e^{ix})|^2\right| = 0$, uniformly in x , by Theorem 4.6. On the other hand,

$$|\hat{f}_p^{-1}(x) - \tilde{f}_p^{-1}(x)| \le 2\pi\Big(|\hat{g}_p^*(e^{ix})| + |g_p^*(e^{ix})|\Big)|\hat{g}_p^*(e^{ix}) - g_p^*(e^{ix})| .$$

We now apply Lemma 4.5 and Lemma 6.1. □

7. Acknowledgements

I want to express my thanks to Professor Emanuel Parzen for his guidance, his enthusiasm and his dedication to research.

REFERENCES

Berk, K. N. (1974). Consistent autoregressive spectral estimates. Ann. Statist. 2, 489-502.

Davies, R. B. (1973). Asymptotic inference in stationary Gaussian time-series. Adv. Appl. Prob. 5, 469-497.

Doob, J. L. (1953). Stochastic Processes. Wiley, New York.

Feller, W. (1966). An Introduction to Probability Theory and Its Applications, Volume II. Wiley, New York.

Geronimus, L. Y. (1961). Orthogonal Polynomials. Consultants Bureau, New York.

Grenander, U. and Szegö, G. (1958). Toeplitz Forms and Their Applications. University of California Press, Berkeley.

Ibragimov, I. A. (1964). On the asymptotic behavior of the prediction error. Theor. Probability Appl. 9, 627-633.

Kromer, R. E. (1969). Asymptotic properties of the autoregressive spectral estimator. Ph.D. dissertation, Stanford University.

Pagano, M. (1973). When is an autoregressive scheme stationary? Comm. Statist. 1, 533-545.