

ADAPTIVE SYSTEM IDENTIFICATION BASED ON HIGHER-ORDER STATISTICS

José A. RODRIGUEZ-FONOLLOSA, Josep VIDAL, Enrique MASGRAU

Department of Signal Theory and Communications, UPC
E.T.S.E. Telecomunicació, Apdo. 30.002
08080 Barcelona, Spain.

E-mail: adrian@tsc.upc.es

ABSTRACT

The problem of estimating the AR parameters of a causal ARMA(p,q) process using higher order statistic is addressed. Methods based on a single 1-D slice of the cumulants give a Toeplitz matrix that can be solve with fast lattice algorithms. Nevertheless, it has been shown recently that these methods cannot assure, in general, the consistency of the estimation. Lattice implementations of these algorithms suffer from one additional problem, they fail if any of the diagonal principal minors is singular. In this paper it is shown that there is always a linear combination of p+1 slices that gives a full-rank Toeplitz matrix. This derivation proves with a new approach that consistent estimates can always be obtained with this set of p+1 1-D slices. These results lead to the development of a new adaptive lattice algorithm with improved performance. Some results are presented comparing this scheme with previous algorithms based on a single slice. Estimation of the MA parameters of the obtained AR-compensated sequence completes the identification of the system. As this method is based on cumulants, the estimation will be unbiased even in the presence of colored gaussian noise.

I. INTRODUCTION

Most of the efforts made up to date in system identification and ARMA modelling have been based on the properties of the autocorrelation function related to gaussian noise and linearity hypothesis. However, in many applications where non-gaussian processes or nonlinearities are present, the autocorrelation (and hence, the power spectrum) fails to provide proper estimations of the system to identify, or of the ARMA parameters of a process. Cumulants, and their Fourier transform, polyspectra, do contain information about deviations from normality and nonlinearities. Another important property of cumulants is that they are phase sensitive, so that it is possible to identify two systems which differ only in their phase response Estimation methods based on these statistics are also interesting because they give unbiased results even when the observed signal is corrupted by additive colored Gaussian noise. Therefore, cumulants have found application in problems as blind equalization, speech and array processing, noise and echo cancellation; where accurate system identification is needed.

A great number of linear approaches has been proposed in the literature to estimate the parameters of a moving average (MA) or autoregressive (AR) model given the output. Nevertheless, works related to higher-order statistics techniques [1,3,6] are mainly off-line and suitable for stationary conditions, while those applications require adaptive algorithms for real-time processing. We have concentrated our attention in obtaining adaptives algorithms based on the normal equations and in this paper a new adaptive algorithm is developed, with the aim of obtaining simple estimators with tracking capabilities and assured convergence.

This work was supported by PRONTIC 105/88

II. AR SYSTEM IDENTIFICATION

Several approaches can be found in the literature for AR system identification using cumulants. The adaptive algorithms presented on this paper are based on the linear approaches in which a linear system must be solved to obtain the AR parameters.

Consider a causal real autoregressive AR(p) process x(n) described by

$$\sum_{i=0}^p a(i)x(n-i) = u(n), \quad a(0) = 1 \quad (1)$$

where the input u(n) is an i.i.d. process with $E\{u(n)\} = 0$ and m-th order cumulants γ_m .

The impulse response h(n) of this system satisfies the following recursion

$$\sum_{i=0}^p a(i)h(n-i) = \delta(n) \quad (2)$$

Using this recursion and the Brillinger-Rosenblatt summation formula that relates the m-th order cumulant of x(n) to the impulse response

$$C_{m,x}(i_1, i_2, \dots, i_{m-1}) = \gamma_m \sum_{n=-\infty}^{\infty} \prod_{k=0}^{m-1} h(n+i_k) \quad i_0 = 0 \quad (3)$$

it can be shown that cumulants satisfy a similar recursion

$$\sum_{i=0}^p a(i)C_{m,x}(t-i, i_2, \dots, i_{m-1}) = \gamma_m h(-t) \prod_{k=2}^{m-1} h(-t+i_k) \quad (4)$$

Collecting (4) for t=0, ..., p we obtain the following system of linear equations

$$\begin{aligned} R_m \mathbf{a} &= \mathbf{g} \\ \mathbf{a} &= [1, a(1), \dots, a(p)]^T \\ \mathbf{g} &= [\gamma_m, 0, \dots, 0] \end{aligned} \quad (5)$$

where the matrix R_m is a (p+1) by (p+1) matrix with $C_{m,x}(j-i, i_2, \dots, i_{m-1})$ as its (j,i) element. Note that R_m is Toeplitz but in general is not symmetric.

Several authors have derived algorithms based on 1-D slices, i. e., cumulants with fixed lags i_k (k=2, ..., m-1) [3,6]. The Toeplitz structure of R_m allows to use fast algorithms to solve the linear equations (5) as in the well-known autocorrelation (m=2) case. These efficient algorithms are in close relation with the lattice filter, and in [5] a generalised adaptive lattice structure based in these higher order normal equations was derived.

Nevertheless, in general, consistent estimates cannot be obtained from a single 1-D slice [4], i. e., the rank of R_m may be less than p . When trying to use the fast recursive procedures for solving Toeplitz problems, one additional problem is found. The algorithm will fail if any of the diagonal principal minors of the matrix vanish.

For example, consider an AR(2) model

$$H(z) = \frac{1}{1 + a(1)z^{-1} + a(2)z^{-2}} \quad (6)$$

The numerator of its third-order cumulant $C_{3,x}(0,0)$ is given by

$$N_{3,x}(0,0) = 1 - 2a(1)a(2) + a(2)^3 \quad (7)$$

It can be shown that, for AR(2) processes, $C_{3,x}(t,0)$ is always a full-rank slice, but the diagonal term $C_{3,x}(0,0)$ may be equal to zero. The solid curve in figure 1 illustrates the family of stable AR(2) processes for which this cumulant vanishes. In those cases, (adaptive) lattice algorithms will fail because they need to perform a division by $C_{3,x}(0,0)$ to compute the first order solution (first backward and forward reflection coefficients).

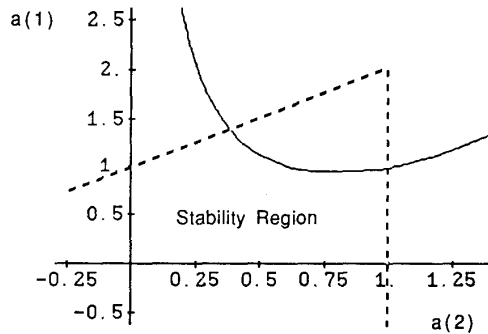


Figure 1. Family of AR(2) processes with $C_{3,x}(0,0) = 0$

The algorithms presented in this paper give solutions to the above mentioned problems while conserving the simplicity of the lattice structure proposed in [5].

In the appendix it is proved that, for a general ARMA(p,q) process, there is always a linear combination of $p+1$ slices that gives a full-rank Toeplitz problem. We will call a linear combination of 1-D slices a w-slice. From equation (4) it is easy to see that the w-slice,

$$W_{m,x}(-t; 0) := \sum_{i=0}^p w(i) C_{m,x}(t, -i, 0, \dots, 0) \quad (8)$$

satisfies the same recursion and, for an AR(p) process,

$$\sum_{i=0}^p a(i) W_{m,x}(-t-i; 0) = 0, \quad t > 0 \quad (9)$$

independently of the weights values $w(i)$.

When these weights are the AR coefficients it is shown in the appendix (24) that $W_{m,x}(-t; 0) = \gamma_m h(t)$. Therefore, using this w-slice in the normal equations (5) we will obtain a Lower Triangular Toeplitz matrix with diagonal elements $W_{m,x}(0; 0) = \gamma_m$. That is, a full-rank Toeplitz Matrix with full-rank diagonal minors.

Of course, these *optimal* weights are not known a priori, because they are the same AR coefficients that we want to estimate. Nevertheless, in an adaptive implementation of this w-slice algorithm we can expect good results if we use the currently estimated parameters $\hat{a}(i)$.

III. THE W-SLICE ALGORITHM

To simplify the derivation of the adaptive algorithm we will consider $m=3$ with $\gamma_3 \neq 0$. In this case the normal equations (5) based on the third order cumulants can be expressed as

$$\begin{aligned} E\{z(n)x(n)^T a\} &= E\{z(n)e(n)\} = g \\ z(n) &= [z(n), z(n-1), \dots, z(n-p)]^T \\ x(n) &= [x(n), x(n-1), \dots, x(n-p)]^T \end{aligned} \quad (10)$$

where $z(n)$ is $x(n)x(n+2)$. This equation allows an interpretation of the problem in terms of linear prediction of $x(n)$ with $z(n)$ as *instrumental* process [5].

The adaptive w-slice algorithm uses a linear combination of 1-D slices. Therefore, for $m=3$, $z(n)$ is now

$$z(n) = \sum_{i=0}^p \hat{a}(i) x(n)x(n-i) = x(n)e_p^f(n) \quad (11)$$

where $e_p^f(n)$ is the output of the adaptive lattice filter (Fig. 2).

The coefficients of the lattice filter can be adapted with a generalised version of the gradient lattice algorithms [8], with two reflection coefficients per section (GAL2). The resulting equations are the following

$$\begin{aligned} e_j^f(n) &= e_{j-1}^f(n) + k_j^f(n)e_{j-1}^b(n-1) \\ e_j^b(n) &= e_{j-1}^b(n-1) + k_j^b(n)e_{j-1}^f(n) \\ r_j^f(n) &= \beta r_j^f(n-1) + e_{j-1}^b(n-1)z(n-j) \\ r_j^b(n) &= \beta r_j^b(n-1) + e_{j-1}^f(n)z(n) \\ k_j^f(n+1) &= k_j^f(n) + e_j^f(n)z(n-j)/r_j^f(n) \\ k_j^b(n+1) &= k_j^b(n) + e_j^b(n)z(n)/r_j^b(n) \end{aligned} \quad (12)$$

where $e_j^f(n)$ and $e_j^b(n)$ are the forward and backward prediction errors at the j -stage, respectively, and $x(n)$ is the AR sequence that is being filtered with the FIR lattice. Note that an additional lattice for the instrumental process $z(n)$ is not required as figure 2 illustrates.

Another interesting property of the derived w-slice algorithm is that, for a pure AR(d) process, with $d \leq p$, the j -th reflection coefficient will converge to $a(j)$, and all the backward coefficients to zero.

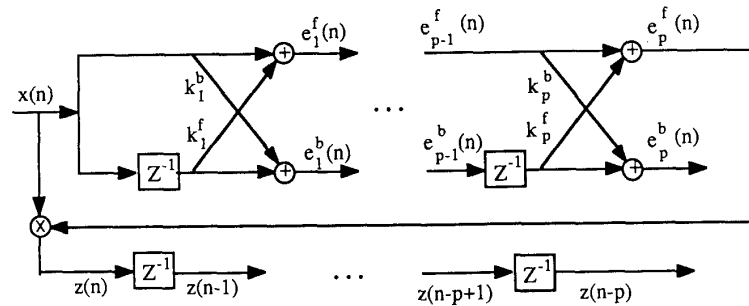


Figure 2. Lattice structure for the w-slice algorithm

IV. ARMA PROCESSES

The w-slice algorithm can be extended in order to estimate the AR parameters of an ARMA(p,q) process. From (24) we see that the w-slice must be the linear combination of the 1-D slices that goes from q-p to q. For m=3 the instrumental z(n) is given by

$$z(n) = \sum_{i=0}^p \hat{a}(i)x(n)x(n+q-i) \quad (13)$$

After convergence, the output of the lattice filter $e_p^f(n)$ will be the AR-compensated sequence. Then, the identification of the zeros of this MA process can be performed with any adaptive version of the cumulant-based methods [1,2,5].

V. RESULTS

In order to test the performance of the two types of instrumental variables, a simulation has been done in the following conditions: 100 Monte Carlo runs of an AR process, driven by exponentially distributed white noise, have been applied to two versions of the lattice algorithm. The first algorithm uses $z(n) = x(n)x(n)$, i.e., it is based on the diagonal slice. The second algorithm is the developed w-slice algorithm that uses the linear combination of 1-D slices. Both algorithms use the mentioned gradient-type equations (12) with $\beta=1$.

Each sequence has 3000 points. The parameters of the AR(2) models have been chosen in order to check the quality of the estimations when the parameters approach the curve shown in figure 1. As we can see in table I, for case I, the behavior is similar in mean and standard deviation for the two methods. However, in case II and III, as the values of the parameters get closer to the curve, the algorithm with $z(n) = x(n)x(n)$ becomes biased and a dramatical increase in the value of the deviations is found. We can check as well that the w-slice algorithm is hardly affected by the nominal values of the parameters, both in mean and variance.

Parameters	True values	Algorithm with $z(n)=x(n)x(n)$	w-slice algorithm $z(n)=e(n)x(n)$
I a(1)	0.5	0.4898±0.0258	0.4931±0.0225
	0.5	0.5308±0.0430	0.5198±0.0318
II a(1)	1.0	0.9651±0.4602	0.9915±0.0280
	0.5	0.5910±0.2227	0.5151±0.0510
III a(1)	1.125	0.8632±0.5415	1.1171±0.0330
	0.5	0.6729±0.2801	0.5188±0.0557

Table I. Comparison of performance of the adaptive lattice algorithm with $z(n)=x(n)x(n)$, and the proposed w-slice algorithm.

Figure 3 shows the a(2) parameter track of the w-slice algorithm for case I. We observe a fast initial convergence and a progressive reduction in variance.

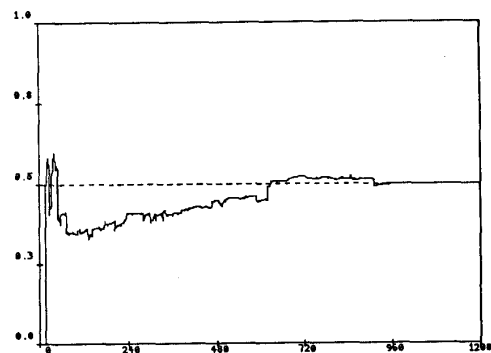


Figure 3. W-slice coefficient track (a(2)).

CONCLUSIONS

It has been shown also that, for a causal ARMA(p,q) process, the linear combination of 1-D slices of the m-th cumulant

$$W_{m,x}(t, q) := \sum_{i=0}^p a(i)C_{m,x}(z; q-i)$$

gives a full-rank *w-slice*. A new adaptive algorithm for system identification using cumulants has been proposed based on this property of cumulant. As this algorithm uses a linear combination of 1-D slices, it allows a simple implementation for real-time applications. It can be used for AR or ARMA non-gaussian processes, providing accurate estimations of the AR parameters even in the presence of additive colored gaussian noise.

This result is at the same time, a simple proof that methods using the normal equations based on this set of slices give consistent estimates of the AR parameters.

APPENDIX

Using a similar notation that in [4], we obtain from (4) that the z-transform of the 1-D slice of a causal, stable ARMA(p,q) process

$$C_{m,x}(t; k) := C_m(-t, k, 0, \dots, 0) \quad (14)$$

is given by

$$C_{m,x}(z; k) := \sum_{l=-\infty}^{\infty} C_m(t; k)z^{-l} = \gamma_m H_{m-1}(z; k)H(z^{-1}) \quad (15)$$

where

$$H_1(z; 0) := H(z) = \frac{\sum_{i=0}^q b(i) z^{-i}}{\sum_{i=0}^p a(i) z^{-i}} = \frac{B(z)}{A(z)} \quad (16)$$

$$a(0) = b(0) = h(0) = 1, \quad b(q) \neq 0, \quad a(p) \neq 0$$

and

$$H_{m-1}(z; k) = \sum_{l=-\infty}^{\infty} h^{m-2}(n)h(n+k)z^{-l} = z^k H(z) * H_{m-2}(z; 0) \quad (17)$$

where * denotes complex convolution.

From equation (15) we can express the 1-D slice normal equations in the z domain as

$$C_{m,x}(z; k) A(z^{-1}) = \gamma_m H_{m-1}(z; k) B(z^{-1}) \quad (18)$$

From the above equation we see that, for $m > 2$, if $A(z^{-1})$ and the numerator of $H_{m-1}(z)$ have common factors, this slice will not be sufficient for the identification of the AR parameters, because the recursion will hold with a minimal order less than p. In fact, as it is shown in [4], a full-rank slice may not exist for any k.

Nevertheless, in [4] it is proved that using a set of (p+1) slices ($k=q-p, \dots, q$) we can always obtain a consistent estimation of the AR parameters.

We will show in this appendix that a linear combination of that set of slices gives a full-rank slice, thus giving at the same time, a simpler proof of the above theorem.

Let $k=q-i$ in (14) and form the following linear combination of 1-D slices

$$W_{m,x}(z; k_0) := \sum_{i=0}^p a(i)C_{m,x}(z; k_0-i) \quad (19)$$

from (15) we obtain that

$$W_{m,x}(z; k_0) = \gamma_m \sum_{i=0}^p a(i)[z^{k_0-i}H(z) * H_{m-2}(z; 0)]H(z^{-1}) \quad (20)$$

but from (16)

$$\sum_{i=0}^p a(i)z^{k_0-i}H(z) = z^{k_0}B(z) \quad (21)$$

and from the linearity of the convolution

$$W_{m,x}(z; k_0) = \gamma_m [z^{k_0}B(z) * H_{m-2}(z; 0)]H(z^{-1}) \quad (22)$$

but $z^{k_0}B(z) * H_{m-2}(z; 0)$ is nothing else than the z-transform of

$$\sum_{j=0}^q b(j+k_0)h^{m-2}(j) \quad (23)$$

If $k_0=q$ then (23) is simply equal to $b(q) \neq 0$, and therefore,

$$W_{m,x}(z; q) = \gamma_m b(q)H(z^{-1}) \quad (24)$$

$$W_{m,x}(z; q) A(z^{-1}) = \gamma_m b(q)B(z^{-1}) \quad (25)$$

gives a full-rank *w-slice*. Note that these results are in concordance with the results of [4] because $W_{m,x}(z; q)$ is a linear combination of the 1-D slices that goes from $q-p$ to q .

REFERENCES

- [1] J.K. Tugnait, "Approaches to FIR System Identification with Noisy Data using Higher-Order Statistics", *IEEE Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, 1989.
- [2] G.B. Giannakis, "Cumulants: A powerful tool in signal processing", *Proc. IEEE*, vol. 75, September 1987.
- [3] M.R. Raghuveer and C.L. Nikias, "Bispectrum Estimation: A Parametric Approach", *IEEE Trans. on ASSP*, vol. 33, 1985.
- [4] A. Swami and J. M. Mendel, "AR Identifiability using Cumulants", *IEEE Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, pp 1-6, 1989.
- [5] A. Swami and J. M. Mendel, "Adaptive System Identification Using Cumulants", *Proc. ICASSP*, pp 2248-2251, 1988.
- [6] G.B. Giannakis, J.M. Mendel, "Identification of Nonminimum Phase Systems Using Higher Order Statistics", *IEEE Trans. on ASSP*, March 1989.
- [7] B. Friedlander and B. Porat, "Adaptive IIR Algorithms Based on Higher Order Statistics", *IEEE Trans. on ASSP*, April 1989.
- [8] B. Friedlander, "Lattice Filters for Adaptive Processing", *Proc. IEEE*, vol. 70, No. 8, pp. 829-867, August 1982.