On the computation of the MA parameters in ARMA models

Manuel Duarte Ortigueira

Inesc - R. Alves Redol, 9 2º 1000 Lisbon and Instituto Superior Técnico - Av. Rovisco Pais Portugal

RESUME

L' estimation des paramètres MA est la partie la plus dificile en modelization ARMA. Tandis qu' il y a assez d' algorithmes efficients pour le calcul des paramètres AR, il n' existe point d' algorithme avec une efficacité semblable pour l' estimation des paramètres MA. Ceci se passe car le calcul MA est essentiellement un problème non linéaire. D' autre côté dans la plupart des procédés, l' estimation des paramètres MA est faite après une filtrage par le filtre AR. Ceci a, comme consequence une grande influence de la partie AR sur la MA. Le méthode le plus commun est réporté avec ses difficultés principales. D' autres deux algorithmes sont décrits qui donnent des solutions pour quelques désavantages qu' ils présentent.

SUMMARY

The estimation of the MA parameters is the most difficult task in ARMA modeling. While there are a lot of efficient algorithms for the computation of the AR parameters, no algorithm exists with equivalent efficiency for the estimation of the MA parameters. This happens because MA computation is an essentially nonlinear problem. On the other hand, in most procedures, the estimation of the MA parameters is performed after a filtering by the AR filter. This has as consequence a great influence of the AR part over the MA one. The usual approach is reported together with its main difficulties. Other two algorithms are described, which solve some of the drawbacks presented by those.

1. INTRODUCTION

The usual approach is based on the high order Yule-Walker equations, which we are going to obtain.

Let x_n be the output of a time-invariant ARMA(N,M) system driven by stationary white noise. The system is described by the difference equation:

$$\sum_{i=0}^{N} a_i \cdot x_{n-i} = \sigma^2 \sum_{i=0}^{M} b_i \cdot \varepsilon_{n-i} \qquad (1.1)$$

where a_i (i=0,...,N) , with $a_0=1$, are the AR parameters, b_i (i=0,...,M) , with $b_0=1$, are the MA parameters and ε_n is the white noise with variance σ^2. Following a common practice, x_n is assumed to be a zero mean signal. We will assume also the system to be stable and minimum phase. We begin by multiplying both sides in (1.1) by x_{n-j} (j=0,1,2,... ...) and taking expectation values to obtain:

$$\sum_{i=0}^{N} a_i R(j-i) = v_j \qquad j\geq 0 \qquad (1.2)$$

where R is the autocorrelation function (ACF) of x_n and v_j is given by:

$$v_j = \sigma^2 \sum_{i=j}^{M} b_i \cdot h_{i-j} \qquad 0\leq j\leq M$$
$$= 0 \qquad j>M \qquad (1.3)$$

where h_i is the impulse response of the ARMA system. The AR coefficients are obtained by solving the high order Yule-Walker equations obtained from (1.2) with j=M to j=N+M. The correlation of v_j with a_j gives the ACF of the MA part:

$$R_{MA}(k) = \sigma^2 \sum_{i=0}^{M-k} b_i \cdot b_{i+k}$$
$$= \sum_{i=0}^{N} \sum_{j=0}^{N} a_i\, a_j\, R[k-(i-j)] \qquad k=0, \ldots, M \qquad (1.4)$$

The computation of the MA parameters is achieved by spectral factorization.

A close look into eq. (1.2) and (1.4) shows that:

a) To compute the ARMA(N,M) parameters and the white noise variance we only need the first N+M+1 ACF values { R(n), n=0, ... , N+M}.

b) The AR parameters do not depend directly on the MA parameters, but only on the MA order.

c) The MA parameters are strongly AR-dependent.

We are going to pay a special attention to the last statement, because it is the main source of problems in ARMA modeling. In fact, most methods use, directly or indirectly, equations (1.2) and (1.4). This means that the estimation errors will affect the ACF of the MA part and, as consequence, the MA parameters. We will consider methods with such base as Yule-Walker methods (sec.2). The associated factorization problem will be treated too.

Recently, a new method appeared which computes both the AR and MA parameters in a decoupled way. We will call it "Dual" method and will present it in sec.3. At last a method proposed by the author and called here recursive method is presented. In this case, two "MA" polynomials are computed simultaneously with the AR polynomial. The MA polynomial is computed as the asymptotic limit of one of those polynomials. The limiting operation is performed in a recursive way (sec.4). In sec. 5 we will present some conclusions.

2- Yule-Walker methods

a) AR computation

Most ARMA methods are based on equations (1.2) and (1.4). As referred before, the main problem in this approach is the influence of the AR part over the MA. This cannot be avoided in this context, but it can be alleviated by better AR estimates. The use of overdimensioned high order Yule-Walker equations together with SVD algorithms (see [3],[4] and [5]) could increase slightly the performance of the estimator.

b) MA autocorrelation computation

Recently Moses and Beex published an excellent survey about MA computation. From it, one can conclude that the best methods used in this context are obtained by computing the biased ACF estimate of the forward and backward time series obtained by filtering the original series by the AR filter. If the length of the series is not enough high, a lag window must be used to prevent against a nonpositive semi definite autocorrelation. Instead of a window, we suggest the use of the unbiased autocorralation estimate and slight increases of R(0) when the factorization algorithm fails to converge pointing out an ACF matrix non semi definite positive.

c) Spectral factorization

Once we computed the ACF a spectral factorization is needed to compute the MA parameters. We are going to describe briefly two of the known methods: Wilson [1] and Le Roux-Grenier [9] methods.

c1) Wilson algorithm

This algorithm, described by Box and Jenkins [1] computes the MA parameters { b_i , i=1,M} and σ^2 iteratively by adjusting them in order to minimize the error energy:

$$E=\sum_{n=0}^{M}[R(n)-\sigma^2.\sum_{i=0}^{M-n}b_i\,b_{i+n}]^2 \qquad (2.1)$$

using a quadratically convergent Newton-Raphson algorithm.

c2) Le Roux-Grenier algorithm

This a very ellegant factorization algorithm. It is a consequence of Levinson recursion. Let us describe it.

Define two polynomials $B^M(z)$ and $\beta^M(z)$ by:

$$B_N^M(z)=\sum_{i=0}^{M}b_i\,z^{-i} \qquad (2.2)$$
$$\beta_N^M(z)=\sum_{i=0}^{M}\beta_i\,z^{-i} \qquad (2.3)$$

with

$$b_i=\beta_i=R(i)/R(0) \qquad i=1,...,M \qquad (2.4)$$
$$b_0=1,\ \beta_0=0$$

These polynomials converge asymptotically ($N\longrightarrow\infty$) to the MA polynomial and to zero [9], respectively and verify the following recursions:

$$B_N^M(z)=[B_{N-1}^M(z)-K_0^N\,z.\beta_{N-1}^M(z)]/q_0^N \qquad (2.5)$$
$$\beta_N^M(z)=[z.\beta_{N-1}^M(z)-K_0^N\,B_{N-1}^M(z)]/q_0^N \qquad (2.6)$$

where K_0^N is the reflection coefficient (RC) which is computed in such a way that the zeroth order coefficients of $B_N^M(z)$ and $\beta_N^M(z)$ are one and zero, respectively:

$$K_0^N=z.\beta_{N-1}^M(z)\Big|_{z=\infty} \qquad (2.7)$$

or, the 1st order coefficient of $\beta_{N-1}^M(z)$ and q_0^N is given by:

$$q_0^N=1-(K_0^N)^2 \qquad (2.8)$$

Result (2.7) is very important since it shows that the RC sequence is extrapolated. The convergence can be assured and it emerges as a consequence of the increasing order predictors [12]. Another interesting result can be obtained from (2.5) and (2.6). Define a spectral function $S_N^M(z)$ by:

$$S_N^M(z)=B_N^M(z).\,B_N^M(z^{-1})-\beta_N^M(z).\beta_N^M(z^{-1}) \qquad (2.9)$$

If we use (2.5) and (2.6), we obtain:

$$S_N^M(z)=S_{N-1}^M(z)/(q_0^N)^2 \qquad (2.10)$$

Both members of (2.10) are, aside a constant, $S_{MA}(z)$ the spectral function of the MA part. We arrive to this conclusion by noting that (2.10) is valid for all N and

$$S_\infty^M(z)=B_\infty^M(z).\,B_\infty^M(z^{-1})=B(z).\,B(z^{-1})=S_{MA}(z).$$

The extrapolation of the RC sequence allows us to compute the white noise variance by:

$$\sigma^2=R(0).\prod_{i=1}^{\infty}[1-(K_0^i)^2] \qquad (2.11)$$

3 - Dual method

Consider eq. (1.2) again. As seen previously, this equation allows us to compute the parameters of an ARMA minimum phase model. We are going to do some algebraic manipulations to obtain an equivalent, but more useful, set of equations. We begin by remembering that the inverse of the autocorrelation matrix can be expressed as:

$$R^{-1}=C.\,D.\,C^T \qquad (3.1)$$

where C is a lower triangular matrix whose columns are the coefficients of the predictor error filters of decreasing orders and D a diagonal matrix of the inverses of the prediction error powers in decreasing orders [2]. The substitution of (3.1) into (1.2) allows us to obtain:

$$\sum_{i=0}^{M}w_N^M(i).\,C_{j,i}=a_M^N(j) \qquad j=0,......,N$$
$$=0 \qquad j=N+1,\ldots\ldots N+M \qquad (3.2)$$

where $C_{j,i}=p_{j-i}^{N+M-i}$ for $i\le j$ and $C_{j,i}=0$ for $i>j$, p_{j-i}^{N+M-i} are the (N+M-i)th order predictor coefficients, $a_M^N(j)$ are the coefficients of the (N,M) AR polynomial and $w_N^M(i)$ ($w_N^M(0)=1$) are the components of a vector w given by:

$$w=D\ C^T\,v \qquad (3.3)$$

with v given by (1.3). It is not hard to prove that $w_N^M(i)$

(i=1,...,M) converge asymptotically to the MA parameters ([10] and [13]). They are the solution of the system formed by the last M+1 equations in (3.2). This algorithm has two advantages over Yule-walker approach:

1- The substitution the ACF by predictors or, equivalently, by RC's. This is an advantage of (3.2) over (1.2) since we know several efficient algorithms of linear prediction [1], [2], [6] and [7]. The problems posed by the ACF are well known [2].

2- We avoided the spectral factorization to find the MA parameters.

In a similar way, one can easily obtain:

$$\sum_{i=0}^{N} a_M^N(i).\, F_{j,i} = w_N^M(j) \qquad j=0,\ldots\ldots,N$$
$$= 0 \qquad j=N+1,\ \ldots\ldots\ N+M \qquad (3.4)$$

where $F_{j,i}$ are the elements of a matrix F given by:

$$F = C^{-1} \qquad (3.5)$$

Now the AR parameters are the solution of the last N+1 equations in (3.4). In [10] and [13] overdimensioned matrices are used in (3.2) and (3.4). Their least-squares solutions can be obtained in a "lag" recursive way. This is valid for both AR and MA estimations. This decoupled computations explains why the method is called "dual".

4 - Recursive method

a) A Fundamental Relationship

The algorithm we are going to describe in the following is based on the fundamental relationship [12]:

$$A_M^N(z) = B_N^M(z).P^{N+M}(z) + \beta_N^M(z)\, z^{-N-M}\, P^{N+M}(z^{-1}) \qquad (4.1)$$

where:

$$A_M^N(z) = \sum_{0}^{N} a_M^N(i).\, z^{-i} \qquad (4.2)$$
$$B_N^M(z) = \sum_{0}^{M} b_N^M(i).\, z^{-i} \qquad (4.3)$$
$$\beta_N^M(z) = \sum_{0}^{M} \beta_N^M(i).\, z^{-i} \qquad (4.4)$$
$$P^{N+M}(z) = \sum_{0}^{N+M} p^{N+M}(i).\, z^{-i} \qquad (4.5)$$

with $a_M^N(0)=1$, $b_N^M(0)=1$ and $\beta_N^M(0)=0$. $P^{N+M}(z)$ is the (N+M)th order linear predictor, $A_M^N(z)$ is the ARMA(N,M) AR polynomial and, if M is the correct MA order, the polynomials $B_N^M(z)$ and $\beta_N^M(z)$ converge to the MA polynomial and to zero respectively, as N goes to infinity [12]. The relation (4.1) shows that the AR polynomial is a vector in the space spanned by the (N+M)th order forward and backward predictors. Eq. (4.1) can be obtained from (1.2) by the use of Gohberg-Semencul formula to express the inverse of the autocorrelation matrix [8].

b) Recursive Computation of the AR Parameters

The algorithm to be presented can be found in [12] and can be obtained from (4.1).

Define $A_0^N(z) = P^N(z)$. The ARMA(N,M) AR polynomial, $A_M^N(z)$, can be computed recursively by (see [12] and [14]):

$$A_M^N(z) = A_{M-1}^{N+1}(z) + \mu_M^N.\, z^{-1}\, A_{M-1}^N(z) \qquad (4.6)$$

where μ_M^N is obtained by forcing the (N+1)th coefficient on the RHS of (4.6) to be equal to zero, leading to:

$$\mu_M^N = -K_{M-1}^{N+1} / K_{M-1}^N \qquad (4.7)$$

with $\mu_M = -K_0^1$; K_M^N is the last coefficient of $A_M^N(z)$ and is called Generalized Reflection Coefficient (GRC) or Generalized Partial ACF (GPAC), [12] and [14].

Eq. (4.6) defines a recursion for the computation of the AR part of an ARMA model from which we can conclude that an AR polynomial for a given (N,M) pair is computed from a set of N+M predictors or RC's.

c) Recursive Computation of the MA Parameters

Both polynomials $B_N^M(z)$ and $\beta_N^M(z)$ can be computed in a recursive way. This algorithm is readily found by substituting (4.1) into (4.6) and using,[2]

$$P^{N+M-1}(z) = [\, P^{N+M}(z) - K_0^{N+M}.\, z^{-N-M}\, P^{N+M}(z^{-1})\,]/q_0^N \qquad (4.8)$$

with

$$q_M^N = 1 - (\, K_0^{N+M})^2 \qquad (4.9)$$

So, inserting (4.6) into (4.4) and rearranging the terms we obtain:

$$B_N^M(z) = B_{N+1}^{M-1}(z) + \gamma_M^N\,[\, z^{-1}\, B_N^{M-1}(z) - K_0^{N+M}\, \beta_N^{M-1}(z)\,] \qquad (4.10)$$
$$\beta_N^M(z) = \beta_{N+1}^{M-1}(z) + \gamma_M^N\,[\, \beta_N^{M-1}(z) - K_0^{N+M}\, z^{-1}\, B_N^{M-1}(z)\,] \qquad (4.11)$$

with

$$B_N^0(z) = 1 \text{ and } \beta_N^0(z) = 0 \qquad (4.12)$$
$$\gamma_M^N = \mu_M^N / q_M^N \qquad (4.13)$$

μ_M^N and q_M^N are given by (4.6) and (4.9).

These recursions allow the computation of the three polynomials for values of M from M=0 to $M=M_0$ and for N from N=0 to $N= N_0+M_0-M$, where N_0 and M_0 are pre-assigned integers that we may suppose to be the correct ARMA orders.

Now we must face the problem of the computation of $B_{\infty}(z)$. First we note that, from (1.2)

$$A_M^N(z) = A_M^{N-1}(z) \qquad (4.14)$$

if $M=M_0$ and $N>N_0$.

After substituting (4.1) into (4.14), and using (4.8), it is a simple task to find:

$$B_N^M(z) = [\, B_{N-1}^M(z) - K_0^{N+M}\, z.\beta_{N-1}^M(z)\,] / q_M^N \qquad (4.15)$$

$$\beta_N^M(z) = [z.\, \beta_{N-1}^M(z) - K_0^{N+M}\, B_{N-1}^M(z)\,] / q_M^N \qquad (4.16)$$

As the zeroth order coefficients of $B_N(z)$ and $\beta_N(z)$ are one and zero, respectively, we conclude that:

$$K_0^{N+M} = z.\, \beta_{N-1}^M(z)\Big|_{z=\infty} \qquad (4.17)$$

or, the 1st order coefficient of $\beta_{N-1}^M(z)$. The similarity between (4.15) to (4.17) and (2.5) to (2.7) is clear. It can be proven that the results (2.9) to (2.11) remain valid here. One must remark that:

1- We only need to estimate N_0+M_0 RC's to compute the ARMA(N_0,M_0) parameters.

2- The convergence is can be warranted and it emerges as a consequence of the convergence of the predictors. This happens if the RC have magnitudes inferior to unity.

6- Conclusions

We faced the problem of the computation of MA parameters in ARMA models. Three different approaches were presented: Yule-Walker, dual and recursive methods. It seems that the last two are better than the former. Only an exaustive simulation can show the comparative performance of the methods. However it is clear that the last two open new ways which may lead to high performance algorithms.

REFERENCES

[1] G. E. P. Box and G. H. Jenkins, "Time Series Analysis: Forecasting and Control," Holden-Day, Ed. 1970.

[2] J. P. Burg, "Maximum Entropy Spectral Analysis," Ph. D. dissertation, Dep. Geoph. Stanford University, Stanford, CA, May 1975.

[3] J. A. Cadzow, " Spectral Estimation: An Overdetermined Rational Model Equation Approach," Proc. IEEE, 70, pp. 907-939, Sep. 1982.

[4] J. A. Cadzow, B. Baseghi and T. Hsu, " Singular Value Decomposition to Time Series Modeling," IEE Proc. Vol. 130, Pt. F, No. 3, April 1983.

[5] J. A. Cadzow and O. M. Solomon, " ARMA System Identification: An Algebraic Approach," IEEE Proc. ICASSP, 1984.

[6] D. G. Childers (Ed.), " Modern Spectrum Analysis", IEEE Press, 1978.

[7] S. Haykin," Nonlinear Methods of Spectral Analysis", Springer-Verlag, 1979.

[8] T. Kailath, A. Vieira, and M. Morf, " Inverses of Toeplitz operators, innovations, and orthogonal polynomials," SIAM Review, Vol.20, No.1, Jan. 1978.

[9] J. Le Roux and Y. Grenier, " An Iterative Procedure for Moving Average Models Estimation, Proc. of the IEEE ICASSP, 1980.

[10] J. M. F. Moura and M. I. Ribeiro, " Parametric spectral estimation for ARMA processes", Proc. of the IEEE third ASSP Workshop on Spectrum Estimation and Modeling, Boston, 1986.

[11] M. D. Ortigueira, "A Predictor Based ARMA Model," Proc. of the Mediterranean Electrotechnical Conference (Melecon), Madrid, Spain, 1985.

[12] M. D. Ortigueira, "A Reflection Sequence Based Recursive Algorithm for the Computation Of the ARMA Parameters", to be published.

[13] M. I. Ribeiro and J. M. F. Moura, " Dual estimation of Poles and Zeros of an ARMA(p,q) process", CAPS, Technical Report 02/84, Nov. 1984; revised as LIDS Technical Report No. LIDS-P-1521, MIT, Sept. 1985, submitted for publication.

[14] R. S. Tsay and G. C. Tiao ", Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models," J. of American Statistical Association, Vol. 79, No. 385, March 1984.