Identification of Autoregressive Moving-Average Parameters of Time Series

D. GRAUPE, MEMBER, IEEE, D. J. KRAUSE, MEMBER, IEEE, AND J. B. MOORE

Abstract—**A procedure for sequentially estimating the parameters and orders of mixed autoregressive moving-average signal models from time-series data is presented. Identification is performed by first identifying a purely autoregressive signal model. The parameters and orders of the mixed autoregressive moving-average process are then given from the solution of simple algebraic equations involving the purely autoregressive model parameters.**

I. Introduction

Many control system design algorithms and filtering algorithms in the literature assume knowledge of the parameters of the signal process model. In practice rarely is there *a priori* knowledge of these parameters, and so there exists the need to identify a signal model first. For stationary stochastic time-series an autoregressive moving-average (ARMA) model is frequently used since it is the minimum parameter linear model of such time series.

Important contributions [1]–[6] to the problem of identifying the parameters of an ARMA model have been made in the last few years. The text of Box and Jenkins [1] is probably the most complete book to date on the identification of stationary and nonstationary ARMA models. In particular, Durbin [3] has treated the problem of identifying the autoregressive (AR) parameters of an ARMA model given the moving average (MA) parameters (and vice versa). His work is based on the important studies of purely AR models by Mann and Wald [7]. However, a limitation of his algorithms is that he requires an assumption of the order for the AR or ARMA process. Lee [4] and Gersch [5] also achieved results for the problem of estimating the AR parameters of a mixed ARMA model of given order assuming a knowledge of the MA parameters.

Mehra [6] has presented a method for identifying the state variable model for a Gaussian process which can be executed in a recursive manner. His method is computationally convenient for estimating AR parameters of an ARMA model, but is rather complex when dealing with the MA parameters.

This paper is an extension of earlier work by the authors [8], [9]. It is concerned with the estimation of parameters and orders of stationary mixed ARMA time series without *a priori* assumptions on parameters or on orders. The resulting models may be further transformed to yield the parameters of a linear state space model [10]. In this respect, the algorithms of this paper may be viewed as an alternative approach to that taken by Mehra [6], having their main advantage in the estimation of the MA parameters of the ARMA model.

Manuscript received June 23, 1972; revised August 16, 1973. Paper recommended by H. W. Sorenson, Chairman of the IEEE S-CS Estimation and Identification Committee.

D. Graupe is with the Department of Electrical Engineering, Colorado State University, Fort Collins, Colo.

D. J. Krause was with the Department of Electrical Engineering, Colorado State University, Fort Collins, Colo. He is now with the Department of Electrical Engineering, North Dakota State University, Fargo, N.D.

J. B. Moore is with the Department of Electrical Engineering, University of Newcastle, Newcastle, N.S.W., Australia and the Department of Electrical Engineering, Colorado State University, Fort Collins, Colo.

II. Mixed Autoregressive Moving-Average Model

A well-known property of stationary time sequences is that they may be represented by a linear filter model driven by white noise [10]. Let us consider the ARMA signal model

$$y_k = \sum_{j=1}^{n} \phi_j y_{k-j} + \sum_{i=0}^{m} \theta_i w_{k-i} \tag{2.1}$$

where $\{w_k\}$ is the input white noise sequence with zero mean and variance σ_w^2; $\{y_k\}$ is the output sequence; $\phi_1, \phi_2, \cdots, \phi_n$ are the AR parameters; and $\theta_1, \theta_2, \cdots, \theta_m$ are the MA parameters. Without loss of generality [8] we take $\theta_0 = 1$. Equation (2.1) is commonly termed the mixed ARMA model and may be written in operator notation as

$$\phi(B)y_k = \theta(B)w_k; \qquad B^i w_k = w_{k-i} \tag{2.2}$$

where B is a delay operator.

The system is assumed to be stable and invertible. That is, all the roots of both $\phi(B)$ (for stability) and $\theta(B)$ (for invertibility) lie outside the unit circle to guarantee that both $\phi^{-1}(B)\theta(B)$ and $\theta^{-1}(B)\phi(B)$ form convergent series [2]. We comment that in cases of noninvertible processes, an invertible equivalent model, having the same first and second order statistics of the original process output, will always be identified and is in fact the only model that can be identified. (The roots of the MA polynomial of this equivalent ARMA model will be the reciprocal of those roots of the original process that are inside the unit circle, whereas all other roots will be as in the original process.)

III. Estimation of Mixed ARMA Processes

A. Pure AR Processes

The problem of consistent least squares estimation of the AR parameters in a purely AR process [(2.1) with $m=0$] has already been solved by Mann and Wald [7]. For Gaussian sequences, their algorithms are, moreover, asymptotically efficient. A recursive version of the algorithm is described by Lee [4] (see also [11]) using sequential regression and will not be repeated here. Consistency of the AR model identification can also be proven using stochastic approximation theory [12] for stable signal model cases.

B. Mixed ARMA Processes

When a process involves $m \geqslant 1$ MA terms, we proceed as follows.

Cross multiplying both sides of (2.2) by $\theta^{-1}(B)$ as in [13], [14], we obtain

$$\theta^{-1}(B)\phi(B)y_k = \gamma(B)y_k = w_k. \tag{3.1}$$

Here $\gamma(B) = \theta^{-1}(B)\phi(B)$ is an *infinite* power series that is convergent for

$|B|<1$ as noted in Section II. Consequently, (y_k) may be represented by

$$y_k=\sum_{i=1}^{\infty}\gamma_i y_{k-i}+w_k+e_k=\sum_{i=1}^{s}\hat{\gamma}_i y_{k-i}+\hat{w}'_k,\qquad \hat{\gamma}=\text{estimate of }\gamma \quad (3.2)$$

where $E[e_k^2]$ can be made arbitrarily small by choosing a sufficiently high finite order s. Actually, an upper bound β_s on $E[e_k^2]$ that vanishes with increasing s can be evaluated by expanding $E[\Sigma_{i=1}^{s}(\gamma_i-\hat{\gamma}_i)y_{k-i}+\Sigma_{j=s+1}^{\infty}\gamma_j y_{k-j}]^2$, as follows [15]:

$$\beta_s\simeq|E[\hat{w}_k'^2]-E[y_k\hat{w}'_k]|+|\hat{\gamma}_s E[y_k y_{k-s}]|\frac{1}{1-\exp(-\Omega)}>E[e_k^2] \quad (3.3)$$

where Ω is the largest time constant of the envelope of $E[y_k y_{k-i}]$, given that $i\gg 1$. As shown in [16], such an envelope always exists for wide-sense stationary y_k.

Since (3.1) and (3.2) represent a purely AR process whose residuals converge to discrete white noise w_k, we may employ sequential regression as in Section III-A to consistently identify γ_i (this identification being asymptotically efficient for Gaussian y_k). The order s may be chosen as some large integer and checked by computing the autocorrelation of $\hat{w}'_k$. As long as the sequence $\{\hat{w}'_k\}$ is correlated for a given order s, a larger value for s must be chosen. We note that the sequential regression estimation of γ_i is extremely fast, even for large s, (in examples worked, a value of $s=20$ is more than adequate). Furthermore, with s sufficiently large, slight changes in s are usually of little consequence, as is indicated by Tables II and III.

C. *Derivation of ARMA Parameters and Orders from AR Model*

Once the parameters of the pure AR model have been consistently identified, an ARMA model of the same process can be derived. The ARMA parameters and order can be obtained directly from the parameters of the purely AR model, noting the relation

$$\frac{y_k}{w_k}=\frac{1}{1+\gamma_1B+\gamma_2B^2+\cdots}=\frac{1+\theta_1B+\theta_2B^2+\cdots+\theta_mB^m}{1+\phi_1B+\phi_2B^2+\cdots+\phi_nB^n} \quad (3.4)$$

where γ_i are the parameters of the purely AR model and ϕ_i, θ_j are the AR and the MA parameters of the mixed ARMA model of the same process. Cross multiplication and equating the coefficients of like powers of B yields

$$\phi_1=\theta_1+\gamma_1$$
$$\phi_2=\theta_2+\theta_1\gamma_1+\gamma_2$$
$$\vdots$$
$$\phi_n=\theta_n+\theta_{n-1}\gamma_1+\cdots+\gamma_n$$
$$0=\theta_n\gamma_i+\theta_{n-1}\gamma_{i+1},\cdots,\gamma_{n+i},\qquad \text{for } i=1,2,\cdots$$

where $\theta_j=0$ for $j=m+1,\ m+2,\cdots,n$. The preceding relationships give us the useful matrix equations

$$\begin{bmatrix}\gamma_n & \gamma_{n-1} & \cdots & \gamma_{n-m+1}\\ \gamma_{n+1} & \gamma_n & \cdots & \gamma_{n-m+2}\\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \gamma_{n+m-1} & & \cdots & \gamma_n\end{bmatrix}\times\begin{bmatrix}\theta_1\\ \theta_2\\ \cdot\\ \cdot\\ \cdot\\ \theta_m\end{bmatrix}=-\begin{bmatrix}\gamma_{n+1}\\ \gamma_{n+2}\\ \cdot\\ \cdot\\ \cdot\\ \gamma_{n+m}\end{bmatrix} \quad (3.5)$$

$$\begin{bmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_i\\ \vdots\\ \phi_n\end{bmatrix}=\begin{bmatrix}\theta_1\\ \theta_2\\ \vdots\\ \theta_j\\ \vdots\\ \theta_n\end{bmatrix}+\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ \theta_1 & 1 & 0 & & 0\\ \theta_2 & \theta_1 & 1 & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \cdot & \cdot & & & \\ \theta_{n-1} & \theta_{n-2} & & & 1\end{bmatrix}\times\begin{bmatrix}\gamma_1\\ \gamma_2\\ \\ \\ \gamma_n\end{bmatrix} \quad (3.6)$$

for $j=1,2,\cdots$, where again it is understood that $\theta_j=0$ for $j=m+1,\ m+2$ and $\phi_j=0$ for $j=n+1,n+2,\cdots$. Clearly, with m, n, and $\gamma_1,\gamma_2,\cdots,\gamma_{n+m}$ known, the parameters θ_i and ϕ_i can be determined from (3.5) and (3.6). Note that the square matrix involving the γ_i in (3.5) must be nonsingular for a unique solution of θ_i. Such will be the case if the m and n specified are minimal orders.

Actually, the determination of the minimal orders m and n from just knowledge of the AR parameters is straightforward. Denoting the square matrix involving the γ_i in (3.5) for $n=\bar{n}$ and $m=\bar{m}$ by $\bar{A}_{\bar{n},\bar{m}}$, the rank of $\bar{A}_{\bar{n},\bar{m}}$ is tested for $[\bar{n},\bar{m}]=[0,0],[1,1],[1,0],[2,2],[2,1],[2,0]$, etc., until for some m and n,

$$\text{rank}\{\bar{A}_{\bar{n},\bar{m}}\}=m,\qquad \text{for }\bar{n}>n,\bar{m}>m$$

or alternatively,

$$|\bar{A}_{\bar{n},\bar{m}}|=0,\qquad \text{for }\ \bar{n}>n,\bar{m}>m. \quad (3.7)$$

In practice, only estimates of the true γ_i and of $|A_{\bar{n},\bar{m}}|$ are available. Hence the condition (3.7) is replaced by a test

$$|\bar{A}_{\bar{n},\bar{m}}|^2<\varepsilon,\qquad \text{for some }\ \bar{n}>n,\bar{m}>m \quad (3.8)$$

and some small $\varepsilon>0$, or better, the values $|\bar{A}_{\bar{n},\bar{m}}|^2$ can be examined for a range of $\bar{n}$ and $\bar{m}$ and the region for which $|\bar{A}_{\bar{n},\bar{m}}|^2$ becomes small to obtain good estimates of m and n. Since the ARMA parameter estimates above are based on estimates of pure AR model parameters which have been shown in Section II to be consistent (if the order of the purely AR model is correct, or otherwise to be within upper bounds as in [15]), and noting [17, theorem 2.3.3], the ARMA parameter estimates will also be consistent (or bounded for unknown AR order). There is likelihood of difficulties with a signal model if m and n are estimated on the low side of the true values since then there is not the possibility of omitting dynamics which are essential in a signal description. On the other hand, if the values for m and n are overestimated, all that happens is that negligibly small extra coefficients are introduced into the model which correspond to the addition of small-magnitude high-frequency terms which for signal models in other than control applications do not usually spell disaster. Some examples will be considered to give some feel for what can happen in working with estimates.

Example 1: Consider a pure ARMA process model

$$\frac{1-0.5B}{1+1.5B+0.625B^2}=\frac{\theta(B)}{\phi(B)};\quad m_0=1;\quad n_0=2.$$

The correct purely AR model for the preceding process is $\gamma(B)=1+2+1.625B^2+\frac{1.625}{2}B^3+\frac{1.625}{4}B^4+\cdots$. Identification of $\gamma(B)$ has yielded $\hat{\gamma}_1=2.018$; $\hat{\gamma}_2=1.639$; $\hat{\gamma}_3=0.801$; $\hat{\gamma}_4=0.402$; $\hat{\gamma}_5=0.198$.

Consequently, for assuming correct orders, the following ARMA parameters where obtained (via $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\gamma}_3$): $\hat{\phi}_1=1.519$ (correct: 1.5); $\hat{\phi}_2=0.649$ (correct: 0.625); $\hat{\theta}_1=-0.489$ (correct: -0.5).

Via $\hat{\gamma}_4$ instead of $\hat{\gamma}_3$, the parameter estimates became $\hat{\phi}_1=1.518$; $\hat{\phi}_2=0.650$; $\hat{\theta}_1=-0.490$, which is very close to the estimates via $\hat{\gamma}_1$ to $\hat{\gamma}_3$.

Underfitting of orders, i.e., assuming $m=1$; $n=1$ will yield for the present example (via $\hat{\gamma}_1$; $\hat{\gamma}_2$) $\hat{\theta}_1=-0.8$ (correct: -0.5); $\hat{\phi}_1=1.218$ (correct: 1.5), whereas using $\hat{\gamma}_1$ to $\hat{\gamma}_{m+n+1}=\hat{\gamma}_3$ instead of $\hat{\gamma}_2$ will yield $\hat{\theta}_1=-.049$; $\hat{\phi}_1=1.528$, which differs considerably from when only $\hat{\gamma}_1$

TABLE I

	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	σ_w^2	s	p	q	Sample Length
True Value	1.8	-1.3	0.4	1.1	0.28	1.0	--	3	2	--
Identified (Gaussian)	1.841	-1.385	0.455	1.012	0.196	0.941	20	3	3	500
True Value	2.0	-1.7	0.5	1.5	0.685	1.0	--	3	2	--
Identified (Gaussian)	2.016	-1.66	0.532	1.466	0.631	1.036	15	3	2	--
True Value	-1.5	-0.625	--	0.41	0.1524	4.09	--	2	2	1000
Identified (Gaussian)	-1.486	-0.612	--	0.391	0.1503	4.161	10	3	2	500
True Value	1.8	-1.3	0.4	1.1	0.28	1.0	--	3	2	--
Identified (Non-Gaussian)	1.78	-1.31	0.416	1.07	0.318	0.980	10	3	2	500

TABLE II

s	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	σ_w^2
5	2.038	-1.762	0.550	1.121	0.513	1.386
10	2.018	-1.638	0.537	1.428	0.609	1.070
15	2.016	-1.661	0.532	1.466	0.631	1.036
20	2.016	-1.677	0.529	1.470	0.623	1.031
True Values	2.0	-1.7	0.5	1.5	0.685	1.0

N = 1000 in all cases

TABLE III

Number of AR Parameters	Number of MA Parameters	$\text{var}[y_k - \hat{y}_k]$
2	1	unstable
2	2	1.423
2	3	1.042
3	1	1.052
3*	2*	1.031
3	3	1.034
4	1	1.039
4	2	1.031
4	3	1.036

s = 20, N = 500 samples *True order

and $\hat{\gamma}_2$ are employed. Note that $\hat{\phi}_2$ is now assumed to be 0 instead of 0.625.

Overfitting of orders, i.e., assuming $m=1$; $n=3$ ($m_0=1$; $n_0=2$), yields very close estimates for ϕ_1, ϕ_2, and θ_1 as when the correct orders are assumed either when $\hat{\gamma}_{m+n}$ is employed or when $\hat{\gamma}_{m+n+1}$ is employed instead of $\hat{\gamma}_{m+n}$. The overfitting above yields (via γ_4) also an estimate for $\hat{\phi}_3$ where $\hat{\phi}_3=-0.0192$ (correct: 0). A similar estimate for $\hat{\phi}_3$ is obtained via $\hat{\gamma}_5$.

Example 2: Consider a process given by:

$$\frac{1-0.5B}{1+1.5B+0.66B^2+0.08B^3}$$

namely, $m_0=1$; $n_0=3$. Here, $\hat{\gamma}_1=1.51$; $\hat{\gamma}_2=1.40$; $\hat{\gamma}_3=0.78$; $\hat{\gamma}_4=0.39$; $\hat{\gamma}_5=0.196$; $\hat{\gamma}_6=0.098$; $\cdots$. The preceding $\hat{\gamma}_i$ yield, for assuming $m=2$; $n=2$, an estimate of $\hat{\theta}_2=-0.0775$ via $\hat{\gamma}_4$ and $\hat{\gamma}_2=-0.060$ via $\hat{\gamma}_5$ (the true θ_2 being 0 in both cases). However, a correct assumption of m and of n yields ARMA parameters that are close to the true ones via either $\hat{\gamma}_4$ and $\hat{\gamma}_5$.

IV. Computational Results

Table I gives several examples of computational results where the present procedure was employed (using Fortran on a CDC 6400 computer).

Results illustrating the effect of changing the order of the AR model on the ARMA parameter values are given in Table II.

Table III illustrates the effect of various orders m and n on the variance of the one-step prediction error $y_k-\hat{y}_k$, $\hat{y}_k$ being obtained from the ARMA model.

V. Conclusions

A procedure has been presented for identifying the parameters and orders of linear mixed ARMA models of Gaussian and non-Gaussian time series. This procedure differs from that of [6], and from procedures based on [6], in that no computation of covariances and of spectral factorization is required and in the simplicity of deriving the MA parameters. Also the proofs of consistency do not require a Gaussian assumption. Concrete criteria for determining the AR and ARMA orders are given.

Extensions of the method to input–output noise models, and to some nonstationary processes are possible, as is a direct transformation [9] to a state-space formulation, for cases of noise-free and of noisy measurements.

We note that the present approach can use a stochastic approximation subroutine with a scalar correction coefficient (ρ of [18]) instead of employing a sequential least squares regression subroutine as in [9], [11, ch. 6]. However, the convergence rate will inevitably be much slower [18].

The analysis above indicates that ARMA signal model parameter and order estimation can be performed by a sequential pure AR identification followed by a solution of a set of algebraic equations (3.5) and (3.6), the latter requiring only the storing of $m+n+1$ AR parameters. The computation effort is therefore virtually that of the very fast sequential regression identification of the pure AR model, which requires storage of only s measurements where s is the AR model order. Hence, complete identification can be executed with microprocessing hardware at great speed.

References

[1] G. E. P. Box and G. M. Jenkins, *Time Series Analysis, Forecasting and Control.* San Francisco: Holden-Day, 1970.

[2] G. E. P. Box, G. M. Jenkins, and D. W. Bacon, "Models for forecasting seasonal and non-seasonal time series," in *Advanced Seminar on Spectral Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967, p. 271.

[3] J. Durbin, "The Fitting of Time Series Models," *Rev. Inst. Int. de Stat.*, Vol. 28, pp. 233–243, 1960.

[4] R. C. K. Lee, *Optimal Estimation, Identification and Control*, Research Monograph No. 28, Cambridge, Massachusetts, MIT Press, 1964.

[5] W. Gersch, "Estimation of the autoregressive parameters of a mixed autoregressive moving-average time series," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-15, pp. 583–588, Oct. 1970.

[6] R. K. Mehra, "On-line identification of linear dynamic systems with applications to Kalman filtering," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp.12–21, Feb. 1971.

[7] H. B. Mann and A. Wald, "On the statistical treatment of linear stochastic difference equations," *Econometrica*, vol. 11, pp. 173–220, 1943.

[8] D. Graupe and D. J. Krause, "Sequential regression identification procedure for predictor and nonlinear systems parameters," in *Proc. 14th Midwest Symp. Circuit Theory*, 1971, pp. 8.6.1–8.6.10.

[9] D. J. Krause and D. Graupe, "Identification of autoregressive-moving-average predictor parameters," in *Proc. 2nd Symp. Nonlinear Estimation Theory*, San Diego, Calif., 1971.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, ser. D, vol. 82, pp. 34–45.
[11] D. Graupe, *Identification of Systems*. New York: Van Nostrand, 1972.
[12] G. N. Saridis and G. Stein, "Stochastic approximation algorithms for discrete-time system identification," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 515–523, Oct. 1968.
[13] J. Durbin, "Efficient estimation of parameters of moving average models," *Biometrika*, vol. 49, pp. 306–316, 1959.
[14] H. Wold, *A Study in the Analysis of Stationary Time Series*. Uppsala, Almquist and Wiksells, 1938.
[15] D. Graupe,: "Estimation of upper bounds of errors in identifying ARMA models," in *Proc. 4th Symp. Nonlinear Estimation Theory*, San Diego, Calif., 1973.
[16] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953, p. 524.
[17] E. Lukacs, *Stochastic Convergence*. Lexington, Mass.: Heath, 1968.
[18] G. N. Saridis, "Comparison of 5 popular identification algorithms—A survey," in *Proc. Decision and Control Conf.*, New Orleans, La., 1972.