August 1986 (revised)
January 1985

LIDS-P-1588

Arma Identification[1,2]

G. Alengrin[3], R. S. Bucy[4], J. M. F. Moura[5], J. Pagés[6], and M. I. Ribeiro[7]

Abstract: In view of recent results on the asymptotic behavior of the prediction error covariance for a state variable system, see Ref. 1, an identification scheme for AutoRegressive Moving Average (ARMA) processes is proposed. The coefficients of the d-step predictor determine asymptotically the system moments $U_0, \ldots, U_{d-1}$. These moments are also nonlinear functions of the coefficients of the successive 1-step predictor. Here, we estimate the state variable parameters by the following scheme. First, we use the Burg technique, see Ref. 2, to find the estimates of the coefficients of the successive 1-step predictors. Second, we compute the moments by substitution of the estimates provided by the Burg technique for the coefficients in the nonlinear functions relating the moments with the 1-step predictor coefficients. Finally, the Hankel matrix of moment estimates is used to determine the coefficients of the characteristic polynomial of the state transition matrix, see Refs. 3 and 4.

A number of examples for the state variable systems corresponding to ARMA(2,1) processes are given which show the efficiency of this technique when the zeros and poles are separated. Some of these examples are also studied with an alternative technique, see Ref. 5, which exploits the linear dependence between successive 1-step predictors and the coefficients of the transfer function numerator and denominator polynomials.

In this paper, the problems of order determination are not considered; we assumed the order of the underlying system. We remark that the Burg algorithm is a robust statistical procedure. With the notable exception of Ref. 6 that uses canonical correlation methods, most identification procedures "in control" are based on a deterministic analysis and consequently are quite sensitive to errors. In general, spectral identification based on the windowing of data lacks the resolving power of the Burg technique, which is a super resolution method.

Key words: Autoregressive, Moving Average, ARMA, Identification, Spectral Estimation, Poles, Zeros.

1 Work supported by a NATO research grant # 585/83, by University of Nice, by Thomson CSF-DTAS, and by Instituto Nacional de Investigação Científica. The work of the third author was also partially supported by the Army Research Office under contract DAAG-29-84-k-005.
2 Simple ARMA(2,1) Basic Language analysis programs to construct random data were written by the second author and Dr. K. D. Senne of M.I.T. Lincoln Laboratory. Lack of stability of the direct estimation was observed at TRW with the help of Dr. Gerry Buttler. Analysis programs in Fortran for ARMA(p,q) were written and debugged at CAPS, by the fifth author. Our research was helped by access to VAXs at Thomson CSF-DTAS, Valbonne, France, at CAPS, Instituto Superior Técnico, and at Universidat Politécnica de Barcelona. In particular, the authors explicitly acknowledge Thomson CSF-DTAS, and Dr. Gauthiers, for extending the use of their facilities to all the authors from September 1983 till June 1984, when the examples presented were simulated.
3 Professor, University of Nice (LASSY).
4 Professor, Thomson CSF-DTAS, Valbonne, University of Nice, on leave from University of Southern California.
5 Professor, Carnegie-Mellon University; formerly visiting at Massachusetts Institute of Technology and Laboratory for Information and Decision Systems, while on leave from Instituto Superior Técnico, Lisbon, Portugal.
6 Professor, Universidat Politécnica, Barcelona, Spain.
7 Teaching Assistant, CAPS, Instituto Superior Técnico, Lisbon, Portugal.

1. Introduction

We will be concerned with determining a state variable description of a linear system which produces as output a random s-vector valued process $[y_n^*]$, $n \leq 0$. This process is stationary Gaussian and identical in law to the s-vector valued stationary Gaussian stochastic process $[y_n]$, $n \geq 0$ which is available as input to be processed for identification.

We assume that $[y_n]$, $n \geq 0$ is the output of a linear autonomous system. We consider all triplets (H,Φ,G) where H, Φ and G are sxp, pxp, pxr matrices with real number entries. With each triple, we associate the following dynamical system

$$x_{n+1} = \Phi x_n + G w_n \tag{1a}$$

$$y_n = H x_n \tag{1b}$$

where w_n is an r-vector valued Gaussian white noise sequence, x_o is Gaussian of zero mean and covariance Σ, independent of w_n. We assume that

$$\Sigma = \Phi \Sigma \Phi' + G G'. \tag{2}$$

If $R(\tau)$ denotes the covariance matrix of $[y_n]$, $n \geq 0$

$$R(\tau) = E\, y_{\tau+j}\, y_j' \quad \tau \geq 0 . \tag{3}$$

Define the following equivalence relation between (H,Φ,G) and (C,A,B).

Definition 1.1: $(H,\Phi,G) \cong (C,A,B)$ iff

$$H \Phi^k \Sigma_1 H' = C A^k \Sigma_2 B \quad , k \geq 0$$

where

$$\Sigma_1 = \Phi \Sigma_1 \Phi' + G G' \tag{4a}$$

$$\Sigma_2 = A \Sigma_2 A' + B B' \quad . \tag{4b}$$

We are only concerned with the equivalence class containing $(H,\emptyset,G)$ with

$$H \emptyset^{\tau} \Sigma_1 H' = R(\tau) , \qquad \tau \geq 0 .$$

Remark 1.1: We assume that $[y_n]$, $n \geq 0$ is generated by sistem (1a) and (1b). To avoid trivial complications, Σ is taken positive definite.

Since our problem consists simply on the determination of the above equivalence class, we will select a canonical element to represent the class, with the following properties;

1) $(H,\emptyset,G)$ is completely controllable and completely observable,

2) det $(zI-\emptyset)$ is a strictly Hurwitz polynomial, i.e., all roots lie within the unit disk,

3) $r = s$, the number of inputs is equal to the number of outputs,

see Refs. 7 - 9.

The following assumption serves as a regularity condition:

Assumption 1.1: The canonical element $(H,\emptyset,G)$ of the equivalence classes of $[y_n]$, $n \geq 0$ is such that $U_o = H G$ is invertible.

Note that all the assumptions of Ref. 1 are valid. Consequently, there exists a matrix P, positive semi - definite, (it is the prediction error covariance), so that the following is true: the conditional distribution of x_{n+1} given $y_t,\ldots,y_n$ is normal, with mean

$$\hat{x}_{n+1|n} = E(x_{n+1}|y_n,\ldots,y_t)$$

and covariance P_{n+1}. Furthermore,

$$P_n \xrightarrow[t \longrightarrow -\infty]{} P = G\,G',$$

the limit being attained by a monotone non- increasing sequence in the natural ordering of the cone of positive semi - definite matrices, see Ref. 1.

Of course, "the" representative of the equivalence class of $[y_n]$, $n \geq 0$ is in fact a set of system moments,

$$U_o, U_1, \ldots$$

and coefficients

$$\alpha_i, \; i = 1, \ldots, p$$

where $U_i = H \, \emptyset^i \, G$ and $\det (sI - \emptyset) = s^p - \sum_{i=1}^{p} \alpha_i \, s^{p-i}$.

Each equivalence class is the set of systems with the same z transform.

2. Representation of the d-Step Predictor.

We define the d-step predictor as

$$\hat{y}_{k|m} = E(y_k \mid y_m, \ldots, y_t) = - \sum_{j=t}^{m} b_m^k (j) y_j \tag{5}$$

where $k = d + m$, $d > 0$. Further, the 1-step predictor coefficients $a^{m+1}(j)$ are

$$a^{m+1}(j) = b_m^{m+1}(j) \tag{6}$$

$$b_m^k(k) = 1 \text{ and } b_m^k(j) = 0 \text{ for } j > k.$$

The following theorem relates the d-step predictor to the 1-step predictor.

Theorem 2.1:

$$\hat{y}_{m+d|m} = \sum_{j=t}^{m} E(y_{m+d}|\, I_j); \qquad I_j = y_j - \hat{y}_{j|j-1} \tag{7}$$

and, for $t \longrightarrow -\infty$,

$$\tilde{y}_{m+d|m} = \tilde{y}_{m+d|m+d-1} + \sum_{j=m+1}^{m+d-1} U_{m+d-j}\, U_o^{-1}\, (y_j - \hat{y}_{j|j-1}) \tag{8a}$$

where

$$\tilde{y}_{a|b} = y_a - \hat{y}_{a|b}. \tag{8b}$$

Proof: The first result is an immediate consequence of the independence of the innovations sequence, see Ref. 9.

The second relation is a consequence of the innovation expansion

$$\hat{y}_{m+d|m+d-1} = \hat{y}_{m+d|m} + \sum_{j=m+1}^{m+d-1} E(y_{m+d}\, I_j')(E\, I_j I_j')^{-1} I_j \tag{9}$$

Now, as $t \longrightarrow -\infty$, using results of Ref. 1,

$$E(y_{m+d}\, I_j') = H\, \phi^{m+d-j}\, P_j\, H' \longrightarrow U_{m+d-j}\, U_o',$$

and, also,

$$E(I_j I_j') = H\, P_j\, H' \longrightarrow U_o\, U_o'$$

so that equations (8a) and (8b) are a consequence of equation (9).

□

Corollary 2.1:

As $t \longrightarrow -\infty$

$$b_m^{m+d}(j) = a^{m+d}(j) + \sum_{i=1}^{m+d-j} U_i\, U_o^{-1}\, a^{m+d-i}(j). \tag{10}$$

Proof: This follows by expanding both sides of equations (8a) and (8b) as a series in the y's and equating coefficients. Notice that $b_d^k(j)=0$ for $j=d+1,\ldots,k-1$.

In matrix form

$$[I\ B_1\ldots B_{d-1}]\begin{bmatrix} a^{m+d}(m+d-1) & a^{m+d}(m+d-2) & \cdots & a^{m+d}(m+1) \\ I & a^{m+d-1}(m+d-2) & \cdots & a^{m+d-1}(m+1) \\ & I & & \vdots \\ & & \ddots & \\ \underline{0} & & & a^{m+2}(m+1) \\ & & & \\ & & & I \end{bmatrix} = 0 \qquad (11)$$

where $B_j = U_j\ U_o^{-1}$, or succintly,

$$[I\ B_1\ldots B_{d-1}]\ K_m^{d-1}(a'(.)) = 0. \qquad (12)$$

Notice that equations (11) and (12) are consequences of equations (9a) and (8b) and the fact that $b_{m-1}^{m+d}(j) = 0$, $j = m,\ldots,m+d-1$.

□

In order to complete the description of the representative system, the following theorem due to Padé when s = 1, see Ref. 3, and to Ho and Kalman in the general case, see Ref. 4, determines the transfer matrix denominator polynomial.

Theorem 2.2:

$$\text{Let} \quad H_n(U) = \begin{bmatrix} U_o & U_1 & \cdots & U_{n-1} & U_n \\ U_1 & & & U_n & U_{n+1} \\ \vdots & & \iddots & \vdots & \vdots \\ U_{n-1} & U_n & & & \\ U_n & U_{n+1} & \cdots & U_{2n} & \end{bmatrix}$$

The system $(H,\emptyset,G)$ has state dimension p iff

$$\det H_k \neq 0 \qquad k = 1,2,\ldots,p-1$$
$$\det H_k = 0 \qquad k \geq p$$

Proof: see Refs. 3 and 4. □

Corollary 2.2:

$$H_p(U)\,[-\alpha_p \;\; -\alpha_{p-1} \;\ldots\; -\alpha_1 \;\; 1]' = 0. \tag{13}$$

□

3. Identification

In general, we require a numerical procedure pass from the sequence $[y_n]$ to a 1-step predictor sequence

$$[\,\hat{y}_{n|n-1}\,]$$

or, more precisely, to $\hat{a}^n(j)$, n=0,1,... j<n. The coefficients $\hat{a}^n(j)$ satisfy

$$\hat{y}_{n|n-1} = -\sum_{j=t}^{n-1} \hat{a}^n(j)\, y_j\,. \tag{14}$$

Notice that $\hat{y}_{n|n-1}$ and $\hat{a}^n(j)$ are estimates of the 1-step predictor and its coefficients.

In the case s=r=1, the Burg technique, see Ref. 2, provides the estimates $[\hat{a}^n(j)]$, n>0. Our method then proceeds as follows: we estimate B_i as the solution of

$$[\,I \;\; \hat{B}_1 \;\; \hat{B}_2 \;\ldots\; \hat{B}_j\,]\, K_n^j(\hat{a}^{\cdot}(.)) = 0, \tag{15}$$

and finally determine the estimates of the coefficients of the characteristic equation of $\emptyset$ as

$$H_p(\hat{B})\,[\,-\hat{\alpha}_p \;\; -\hat{\alpha}_{p-1} \;\ldots\; -\hat{\alpha}_1 \;\; 1\,]' = 0. \tag{16}$$

The choice of n depends on the settling time of the estimates, which in turn depends on the zero locations, see Ref. 1. Hence, an increasing sequence of n should be used with equations (15) and (16), n being large enough when the estimates settle down. A similar method can be used to determine the number of samples necessary to fing "good" estimates of a'(.). In the next section, we will outline our experience using this scheme in the case s=r=1 and where the process to be identified is an Autoregressive Moving Average with 2 poles and 1 zero, i.e., an ARMA(2,1).

4. Examples

For an ARMA(2,1) process, the spectrum of $[y_n]$ is

$$F(s).F(1/s) \tag{17}$$

where

$$F(s) = \frac{U_0 s + \alpha_2 U_{-1}}{s^2 - \alpha_1 s - \alpha_2} \tag{18}$$

In this case, if $a^n(j)$ are the predictor coefficients,

$$\hat{B}_1 = - \hat{a}^n(n-1) \tag{19a}$$

$$\hat{B}_2 = - \hat{a}^n(n-2) + \hat{a}^n(n-1)\ \hat{a}^{n-1}(n-2) \tag{19b}$$

$$\hat{B}_3 = - \hat{a}^n(n-3) + \hat{a}^n(n-1)\ \hat{a}^{n-1}(n-3) + \hat{a}^n(n-2)\ \hat{a}^{n-2}(n-3) - \hat{a}^n(n-1)\ \hat{a}^{n-1}(n-2)\ \hat{a}^{n-2}(n-3) \tag{19c}$$

$$\hat{\alpha}_1 = (\hat{B}_1 \hat{B}_2 - \hat{B}_3) / (\hat{B}_1 \hat{B}_1 - \hat{B}_2) \tag{20a}$$

$$\hat{\alpha}_2 = (\hat{B}_1 \hat{B}_3 - \hat{B}_2 \hat{B}_2) / (\hat{B}_1 \hat{B}_1 - \hat{B}_2). \tag{20b}$$

The examples displayed in Table 1 were investigated using simulation. A program was written which had as input the zeros and the poles and as output the sequence $[y_n]$, $n \geq 0$ via equation (1) and (2), with the white noise

generated by the random number generator described in Ref. 10. This sequence was input to a program which found the approximate 1-step predictors based on finite sequences of $[y_n]$. Using the relations (15) and (16), estimates of the zeros and poles were obtained.

Example number	T	Real and Estimated Parameters								
		U_1	U_2	U_3	α_1	α_2	x	y	z	
4.1		1.8	1.55	1.1	1	-0.25	0.5	0.5	-0.8	Real
	1000	1.760722	1.443739	0.947579	0.962593	-0.251121	0.481297±j0.139551		-0.798129	Estimated
	5000	1.757674	1.481965	0.987818	1.005	-0.286141	0.5025 ±j0.183398		-0.752674	
	20000	1.757207	1.479813	1.001	0.9944	-0.2675	0.4972 ±j0.14245		-0.762807	
4.2		1.	.43	.239	0.3	0.1	.5	-0.2	-0.7	Real
	1000	1.079234	0.394736	0.199674	0.293943	0.077503	0.318292	-0.024349	-0.785291	Estimated
	5000	1.078453	0.42504	0.217540	0.326341	0.073097	0.478958	-0.152616	-0.752112	
	20000	1.076895	0.422296	0.230920	0.303568	0.095389	0.495507	-0.192507	-0.773895	
4.3		0.1	0.29	0.001	-0.1	0.3	-0.6	0.5	-0.2	Real
	1000	0.114946	0.265280	-8.1539 E-3	-0.153320	0.282903	-0.614042	0.468726	-0.268266	Estimated
	5000	0.105657	0.291954	-8.2857 E-3	-0.139366	0.306679	-0.627836	0.48847	-0.245022	
	20000	0.101993	0.292406	3.5151 E-3	-0.093291	0.301921	-0.598095	0.504804	-0.195284	
4.4		-0.2	0.05	-0.014	-0.4	-0.03	-0.3	-0.1	-0.2	Real
	900	-0.183405	0.0259	-0.0150131	0.8615	0.17861	1.0342	-0.172703	1.04491	Estimated
	5000	-0.194165	0.0504	-0.025	-1.205	-0.1836	-1.02606	-0.178936	-1.01083	
	20000	-0.197951	0.0518	-0.012	-0.149	0.002	0.012392	-0.161392	0.048951	
4.5		0.	-0.02	0.006	-0.3	-0.02	-0.2	-0.1	-0.3	Real
	1000	4.836 E-3	-0.04218	-0.000049	-0.003	-0.04	-0.0015 ±j0.199994		-0.008	Estimated
	5000	5.69 E-3	-0.0183	-0.005	-0.2834	-0.0199	-0.155075	-0.128325	-0.28909	
	8000	1.305 E-3	-0.025	0.00609	-0.2359	-0.025657	-0.11795 ±j0.108373		-0.2377205	
4.6		0.5	0	-0.25	1.	-0.5	0.5 ±j	0.5	0.5	Real
	1000	0.506	-0.0015	-0.255	0.912	-0.476	0.4516 ±j0.517749		0.406	Estimated
	2000	0.517	0.0017	-0.25032	0.947	-0.488	0.4735 ±j0.51361		0.43	
	5000	0.506	0.0071	-0.2536	1.033	-0.51625	0.5165 ±j0.499478		0.527	
4.7		0.5	0.14	-0.04	1.	-0.36	0.5 ±j	0.32	0.5	Real
	1000	0.504481	0.120918	-0.052077	0.846499	-0.306125	0.42325 ±j0.356349		0.342018	Estimated
	5000	0.505500	0.144487	-0.048293	1.092651	-0.407848	0.546326±j0.330721		0.587151	
	10000	0.498139	0.132096	-0.041679	0.926183	-0.329272	0.463091±j0.33885		0.428044	

Table 1

In Table 1, T stands for the number of sample data points, z for the zero location and x and y either for the two real pole locations or for its real and imaginary part if those poles are complex conjugates. The value of U_0=1 in all examples.

5. Alternative Procedure

We briefly discuss an alternative method of identification, see Ref. 5, based again on the statistics provided by the Burg technique. We take here $[y_n]$ to be a scalar process.

We note that by the innovations expansion, theorem 2.1 of Ref. 1,

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ W_0^1 & 1 & & \underline{0} & \\ W_0^2 & W_1^2 & 1 & & \\ \vdots & \vdots & & \ddots & \\ W_0^N & W_1^N & \cdots & W_{N-1}^N & 1 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix} = W_N \begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix}$$

so that if, $y^N = [y_0\ y_1 \ldots . y_N]'$, then

$$E\, y^N\, y^{N'} = R = W_N \begin{bmatrix} HP_0H' & & & \\ & HP_1H' & \underline{0} & \\ & & \ddots & \\ & \underline{0} & & HP_NH' \end{bmatrix} W_N' .$$

Now, with $\det(sI - \emptyset) = s^p - \sum_{i=1}^{p} \alpha_i s^{p-i}$,

$$\overset{\leftarrow N-p \rightarrow}{[0 \ldots . 0 \ -\alpha_p \ -\alpha_{p-1} \ldots \ -\alpha_1 \ 1]}\ W_N = K \qquad (21a)$$

$$K = \overset{\leftarrow N-q \rightarrow}{[0 \ldots 0 \ b_q^N \ b_{q-1}^N \ldots \ b_0^N]} \qquad (21b)$$

and,

$$b_0^N = 1 \tag{22a}$$

$$b_1^N = -\alpha_1 + W_{N-1}^N \tag{22b}$$

$$b_2^N = -\alpha_2 - \alpha_1 W_{N-2}^{N-1} + W_{N-2}^N \tag{22c}$$

.....

$$b_q^N = -\alpha_q - \alpha_{q-1} W_{N-q}^{N-q+1} - \quad \ldots \quad - \alpha_1 W_{N-q}^{N-1} + W_{N-q}^N \tag{22d}$$

where we assume the process is a scalar Autoregressive Moving Average ARMA(p,q), i.e., $H(sI-\Phi)^{-1}G = \eta(s)/\mu(s)$, degree $\eta(s)=q$, degree $\mu(s)= p$ and $\mu(s)$ is a monic polynomial, H and G are vectors. In fact, the first N-p zeros of the right hand side vector of equation (21b) arise as $W_k^N=H\Phi^{N-k}P_kH'(HP_kH')^{-1}$ and in view of the Cayley-Hamilton theorem. The subsequent zeros in positions N-p+1,..., N-q are zero in view of the theory of the invariant directions of the Riccati equation. In fact, since degree of $\eta(s)=q$, the vectors $\Phi'^{-1}H',\ldots,\Phi'^{-(p-q)}H'$ are invariant directions. This in turn implies that $U_{-1},\ldots,U_{-(p-q)}$ are zero. Because by assumption $U_0\neq 0$, equation (21) follows. Under the assumptions in force in this paper, this result implies

Theorem 5.1 (Moura and Ribeiro, see Ref. 5):

As $N \longrightarrow \infty$

$$b_i^N \longrightarrow \beta_i$$

where

$$\sum_{i=0}^{q} \beta_i s^{q-i} \Big/ 1 - \sum_{i=1}^{p} \alpha_i s^{p-i} = H(sI-\Phi)^{-1}G.$$

□

For details and the vector ARMA process case, see Ref. 5.

Now, equations (21) can be used to identify the poles and zeros of the process, as the elements of $(W_N)^{-1}$ by rows are just the successive 1-step predictors and can be approximated by the Burg predictors.

Table 2 summarizes the results obtained with the present scheme. Examples 4.1, 4.4 and 4.6 displayed in table 1 above are repeated for the case of 1000 samples; the examples in Table 2 have the same system parameters as the corresponding numbered examples in Table 1.

Example Number	T	x	y	z	$\hat{x}$	$\hat{y}$	$\hat{z}$
5.1	1000	0.5	0.5	-0.8	.50098$\pm$j0.13136		-0.80162
5.4	1000	-0.3	-0.1	-0.2	-0.0415$\pm$j0.2131		-0.0605
5.6	1000	0.5$\pm$j0.5		0.5	0.4701$\pm$j0.5112		0.4073

Table 2

The parameters T, x, y, and z, have the same meaning as those in Table 1.

Example 5.1 shows the behavior of the algorithm when there is a significant separation between the zero and the poles. The poles are a double pole. The algorithm solves efficiently for the zero, the estimates for the poles being split into two complex ones about the true pole. Example 5.4 shows the difficulties experienced when a pole-zero cancellation is assumed. Finally, example 5.6 is an intermediate situation, where a real zero is placed between two complex poles. Due to the separation between the zero and the poles, that is larger than in example 5.4, no cancelation occurs, the algorithm separating the zero and the poles. For a larger class of simulation examples, see Ref. 11.

6. Conclusions

The Ho-Kalman method proceeds from assumed exact knowledge of the U_i sequence. It became clear that estimation of U_i was a major problem as ad hoc techniques were unstable. The Yule Walker methods proposed in Ref.12

lead to numerical problems as the Hankel system has estimated covariance entries. Our results can be thought of an extension of the Burg technique to processes with numerator zeros. The method we propose is numerically more stable, except when the zero approaches a pole. Example 4.4 shows that certain problems are intrinsically difficult, as they are almost structurally unstable, since pole zero cancellation is a "catastrophy" in the sense of Thom. Example 4.4 is an ARMA (2,1) which is hard to distinguish from an ARMA (1,0). Of course, we have assumed we know the state space dimension, but it seems that canonical correlation techniques as in Ref. 6 are applicable to the K and H matrices we develop (see equation (12) and theorem 2.2). The technique proposed is quite general in that the multi-output case is essentially identical to the cases examined here, except one must find a replacement for the scalar Burg technique.

Presently, we are investigating a scheme where the zeros are identified, removed, and the resultant Auto Regressive (AR) process is analysed using the Burg method. In certain cases, this is more accurate as the estimation errors for the poles and zeros are less coupled. This research will be reported later.

References

1. BUCY, R.S., Identification and Filtering, Mathematical Systems Theory, Vol.16, pp. 308-317, December 1983.
2. BURG, J.P., Maximum Entropy Spectral Analysis, Stanford University, PhD Thesis, 1975.
3. PADÉ, H., Sur la Representation Approchée d'une Fonction par des Fractions Rationelles, Annales de l'École Normale Supérieur, 3rd, 9, pp. 3-93, 1982.
4. HO, B.L., KALMAN, R.E., Effective Construction of Linear State Variable Models from Input/Output Functions, Regelungstechnik, Vol. 14, pp. 545-548, 1966.
5. RIBEIRO, M.I., MOURA, J.M.F., Dual Estimation of the Poles and Zeros of an ARMA(p,q) Process, CAPS, Technical Report No. 02/84, November 1984; revised as LIDS Technical Report No. LIDS-P-1521, M.I.T. September 1985; submitted for publication.

6. AKAIKE, H., Markovian Representation of Stochastic Processes and its Applications to the Analysis of Autoregressive Moving Average Processes, Annals of Mathematical Statistics, Vol. 26, pp. 368-387, 1974.
7. KALMAN, R.E., Mathematical Description of Linear Dynamical Systems, SIAM Journal of Control, Vol. 1, No. 2, pp. 152-192, 1963.
8. KALMAN, R.E., Linear Stochastic Filtering Theory - Reappraisal and Outlook, Proceedings of the Symposium on Systems Theory, J. Fox Editor, Brooklyn Polytechnic Institute, pp. 197-205, 1965.
9. YOULA, D.C., Synthesis of Linear Dynamical Systems from Prescribed Weighting Patterns, SIAM Journal of Applied Mathematics, Vol. 14, pp. 527-549, 1966.
10. SENNE, K.D., A Machine Independent Random Number Generator, Stochastics, Vol. 1, No. 3, pp. 215-238, 1974.
11. RIBEIRO, M.I., Estimação Paramétrica em Processos Autoregressivos de Média Móvel Vectoriais, forthcoming PhD Thesis, Department of Electrical and Computer Engineering, Instituto Superior Técnico, Lisbon, Portugal, 1986.
12. BOX, G.E., JENKINS, G.H., Times Series Analysis, Forecasting and Control, Holden-Day, 1970.

List of special symbols

Σ

τ

$\cong$

$\longrightarrow$

∞

α

$\sum$

$\neq$

$\longleftarrow$

η

μ