$C(q^{-1})$ [1]. In a previous paper [5], we have shown that in fact the coefficients of the polynomial G_k can be characterized very simply for all $k \geqslant 1$. The purpose of this note is to give a similar characterization for the coefficients of the polynomial F_k for all $k \geq 1$. It will turn out that the characterization of F_k is directly obtained as an intermediate step to determining the coefficients of G_k. This result, together with those of [5], give a very simple formula for determining optimal ARMAX predictors.

II. The Characterization of the Polynomial G_k

In this section, we introduce some notation and review the formula for the coefficients of G_k. These will be used later.

Let

$$A=\begin{bmatrix} 0 & 0 & \cdots & & 0 & -a_n \\ 1 & 0 & \cdots & & 0 & -a_{n-1} \\ 0 & \cdot & & & \vdots & \\ \vdots & & \cdot & & 0 & \vdots \\ 0 & & \cdots & 0 & 1 & -a_1 \end{bmatrix}$$

$$K=\begin{bmatrix} c_n-a_n \\ c_{n-1}-a_{n-1} \\ \vdots \\ c_1-a_1 \end{bmatrix}$$

$$C=[0 \ \cdots \ 0 \ 1].$$

These matrices will be recognized as those which feature in the observable representation of the ARMAX equation (1) [4]. It is proved in [5] that if we denote $G_k(q^{-1})$ by

$$G_k(q^{-1})=g_{k,1}+g_{k,2}q^{-1}+\cdots+g_{k,n}q^{-(n-1)} \qquad k\geq 1$$

then the coefficients $g_{k,i}$ satisfy the formula

$$A^{k-1}K=\begin{bmatrix} g_{k,n} \\ g_{k,n-1} \\ \vdots \\ g_{k,1} \end{bmatrix}. \tag{4}$$

We note also the following relations, which are used in [5] and also can be easily verified

$$\frac{C(q^{-1})}{A(q^{-1})}=1+C(qI-A)^{-1}K \tag{5}$$

$$\frac{G_k(q^{-1})}{A(q^{-1})}=qC(qI-A)^{-1}A^{k-1}K. \tag{6}$$

III. The Characterization of the Polynomial F_k

From (2), it is readily seen that the polynomial F_k can be written as

$$F_k(q^{-1})=1+f_{k,1}q^{-1}+\cdots+f_{k,k-1}q^{-(k-1)}. \tag{7}$$

Our task is then to determine $f_{k,i}$, $i = 1, \cdots, k-1$. From (2), we have

$$F_k(q^{-1})=\frac{C(q^{-1})}{A(q^{-1})}-\frac{q^{-k}G_k(q^{-1})}{A(q^{-1})}. \tag{8}$$

Using (5) and (6), we get

$$F_k(q^{-1})=1+C(qI-A)^{-1}K-q^{-(k-1)}C(qI-A)^{-1}A^{k-1}K. \tag{9}$$

However, we also have

$$C(qI-A)^{-1}K=Cq^{-1}(I-q^{-1}A)^{-1}K$$

$$=Cq^{-1}\sum_{j=0}^{\infty}q^{-j}A^jK. \tag{10}$$

Similarly,

$$C(qI-A)^{-1}A^{k-1}K=Cq^{-1}\sum_{j=0}^{\infty}q^{-j}A^{j+k-1}K. \tag{11}$$

Substituting (10) and (11) into (9) gives

$$F_k(q^{-1})=1+C\sum_{j=0}^{\infty}q^{-(j+1)}A^jK-C\sum_{j=0}^{\infty}q^{-(j+k)}A^{j+k-1}K$$

$$=1+C\sum_{j=1}^{k-1}q^{-j}A^{j-1}K. \tag{12}$$

On comparing (7) and (12), we see that

$$f_{k,j}=CA^{j-1}K \qquad 1\leq j\leq k-1. \tag{13}$$

This simple formula completely determines the polynomial F_k.

On comparing (13) and (4), we see that the coefficients $f_{k,j}$ are determined on route to the determination of $g_{k,j}$. We also see that the two sets of coefficients can be determined by the solution of the equation

$$p(t+1)=Ap(t) \qquad t\geq 1$$

$$p(1)=K. \tag{14}$$

In particular,

$$p(k)=\begin{bmatrix} g_{k,n} \\ \vdots \\ g_{k,1} \end{bmatrix} \tag{15}$$

and

$$f_{k,j}=Cp(j), \qquad 1\leq j\leq k-1. \tag{16}$$

References

[1] K. J. Astrom, *Introduction to Stochastic Control Theory.* New York: Academic, 1970.

[2] K. J. Astrom and B. Wittenmark, *Computer Controlled Systems: Theory and Design.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[3] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[4] M. H. A. Davis and R. B. Vinter, *Stochastic Modelling and Control.* London: Chapman and Hall, 1985.

[5] R. H. Kwong, "A simple characterization of optimal ARMA predictors," *Syst. Contr. Lett.*, vol. 6, pp. 353-355, Jan. 1986.

Optimum Robust Detection of Changes in the AR Part of a Multivariable ARMA Process

ANNE ROUGÉE, MICHÈLE BASSEVILLE, ALBERT BENVENISTE, AND GEORGES V. MOUSTAKIDES

Abstract—**We investigate the theoretical properties of new instruments-based test statistics recently proposed [3] for detection and diagnosis of changes in the AR part of a multivariable ARMA process. The design flexibilities are analyzed, and the optimum design of the test is exhibited. The connection with the accuracy of the I.V. identification method [14] is established, and the comparison with the local likelihood ratio tests is done. These tests have been developed as a solution to the problem of vibration monitoring for offshore platforms.**

I. Introduction and Problem Statement

Consider a multivariable process, described either by the state-space representation

Manuscript received February 4, 1987; revised April 13, 1987. This work was supported by IFREMER under Grant 84/7392 and by CNRS Greco Sarta.

A. Rougée is with IRISA, 35042 Rennes Cedex, France.

M. Basseville is with IRISA/CNRS, 35042 Rennes Cedex, France.

A. Benveniste and G. V. Moustakides are with IRISA/INRIA, 35042 Rennes Cedex, France.

IEEE Log Number 8715600.

0018-9286/87/1200-1116$01.00 © 1987 IEEE

$$X_{t+1}=FX_t+V_{t+1}$$

$$Y_t=HX_t \quad (1\text{-}1)$$

where $X_t \in R^n$, $Y_t \in R^r$, cov $(V_{t+1}) = Q$, or equivalently by the ARMA representation

$$A(q^{-1})Y_t=B(q^{-1})E_t \quad (1\text{-}2)$$

where (E_t) is a white noise with covariance matrix Λ, and

$$A(q^{-1})=I-\sum_{i=1}^{p} A_i q^{-i}$$

$$B(q^{-1})=I+\sum_{j=1}^{p-1} B_j q^{-j}. \quad (1\text{-}3)$$

In [2] and [3], a procedure, called instrumental test (I.T.) has been introduced to detect and isolate changes in the state transition matrix F (respectively, the AR parameters $(A_i)_{1 \le i \le p}$) while the state covariance matrix Q (respectively, Λ and the MA coefficients $(B_j)_{1 \le j \le p-1}$) are unknown. We refer the reader to [3] for the application of I.T. to vibration monitoring whose features motivated the introduction of our instrumental test: in this application, Q is moreover time-varying and the use of I.T. in such a case has been justified both theoretically in [9] and experimentally in [2] and [3].

As the reader may guess, the I.T. is related to the instrumental variable (I.V.) method of Stoica *et al.* [14] in a way which is enlightened in [6]. Following the lines of [14], the purpose of the present note is, in the case of a time-invariant Q, to explore the design flexibilities of the I.T. (Section II), to investigate the relationships between I.T. and I.V. (Section III), and to design I.T.'s that are optimal in a robust sense and compare favorably with the min-max optimal local likelihood ratio tests (Section IV).

II. The Instrumental Test: Design Flexibilities

Introduce the pr^2 vector

$$\theta=\text{col}\,(A_p, \cdots, A_1) \quad (2\text{-}1)$$

obtained by stacking the pr columns of $(A_p, \cdots, A_1)$ on top of each other. Assume we have a nominal model θ^0. For detecting changes in θ^0, we follow the asymptotic local approach of [8], [7], [10], [6], i.e., given a new record $Y_1, \cdots, Y_s$, we want to decide between the hypotheses

$$H_0: \theta=\theta^0 \text{ (no change occurred)}$$

$$H_1: \theta=\theta^0+\frac{\delta\theta}{\sqrt{s}} \quad (2\text{-}2)$$

where $\delta\theta \neq 0$ is an unknown possible direction of change. In the sequel, we shall, respectively, denote by

$$P_0 \text{ and } P_{\delta\theta,s} \quad (2\text{-}3)$$

the law corresponding to the hypotheses H_0 and H_1, and by E_0, $E_{\delta\theta,s}$ the corresponding expectations.

A. Introducing I.T.

Introduce

$$\varphi_t^T=(Y_{t-p}^T, \cdots, Y_{t-1}^T)$$

$$Z_t^T(N)=(Y_{t-p}^T, \cdots, Y_{t-p-N+1})$$

$$W_t=Y_t-\sum_{i=1}^{p} A_j^0 Y_{t-i}. \quad (2\text{-}4)$$

The instrumental statistics is defined as the Nr^2 vector

$$U_s(N)=\frac{1}{\sqrt{s}}\sum_{t=1}^{s} Z_t(N) \otimes W_t \quad (2\text{-}5)$$

where $\otimes$ denotes the Kronecker product. For convenience, in the sequel we shall make use of the following notation, for any matrix A:

$$\bar{A}=A \otimes I_r \quad (2\text{-}6)$$

where I_r is the r-dimensional identity matrix.

The following result is proved in [9].

Theorem 1 [9]: Assume the integer N is such that the $pr \times Nr$ Hankel matrix

$$H_{p,N}=E_0(\varphi_t Z_t^T(N)) \quad (2\text{-}7)$$

is of rank n [cf. (1-1)]. Then we have the following asymptotic normality result:

$$\text{under } H_0: U_s(N) \xrightarrow[s\to\infty]{} N(0, \Sigma_N)$$

$$\text{under } H_1: U_s(N) \xrightarrow[s\to\infty]{} N(\bar{H}_{p,N}^T \cdot \delta\theta, \Sigma_N) \quad (2\text{-}8)$$

where the notation (2-6) has been used, and

$$\Sigma_N=\sum_{i=1-p}^{p-1} E_0[Z_t(N)Z_{t-i}^T(N) \otimes W_t W_{t-i}^T]. \quad (2\text{-}9)$$

Hence, Theorem 1 reduces our problem to a Gaussian hypothesis testing, since any nontrivial change is reflected in a nonzero mean in $U_s(N)$ thanks to assumption (2-7) [12].

Let us now recall some elementary facts about Gaussian hypothesis testing. Let U be a k-dimensional random variable distributed as $N(\mu, \Sigma)$. For testing $\mu = 0$ against $\mu = M\nu$, $\nu \neq 0$, where M is a $k \times j$ matrix (where j is arbitrary), one proceeds as follows. Choose any reduction matrix D such that

$$\text{number of columns of } MD=\text{column rank of } MD=\text{column rank of } M \quad (2\text{-}10)$$

and use the χ^2-test

$$\chi=U^T\Sigma^{-1}\tilde{M}(\tilde{M}^T\Sigma^{-1}\tilde{M})^{-1}\tilde{M}^T\Sigma^{-1}U, \qquad \tilde{M}=MD. \quad (2\text{-}11)$$

Remark 1: χ does not depend on the particular choice of the reduction matrix D satisfying (2-10), cf. [12].

According to Theorem 1, for s large, our desired test I.T. is given by (2-10), (2-11) where

$$U=U_s(N),\ M=\bar{H}_{p,N}^T, \qquad \Sigma=\Sigma_N. \quad (2\text{-}12)$$

For practical implementation, $H_{p,N}$ and Σ_N can be replaced by convenient estimates according to [3] and [9]. Denote these statistics by $\chi_{\text{I.T.}}$.

B. Design Flexibilities and Performance Evaluation

The design choices are

1) the reduction matrix D,
2) the number N of instruments.

To evaluate these possible choices, we introduce the following classical performance index. The I.T. test statistics $\chi_{\text{I.T.}}$ is χ^2-distributed with nr degrees of freedom, with noncentrality parameter equal to zero under H_0, and to γ under H_1, where

$$\gamma=\delta\theta^T\Gamma_{N,D}\delta\theta$$

$$\Gamma_{N,D}=M^T\Sigma^{-1}\tilde{M}(\tilde{M}^T\Sigma^{-1}\tilde{M})^{-1}\tilde{M}^T\Sigma^{-1}M$$

$$\tilde{M}=MD \quad (2\text{-}13)$$

and M, Σ are given by (2-12). Consequently, we choose *the positive symmetric matrix* $\Gamma_{N,D}$ as a performance criterion. The following theorem explores the possible choices in designing I.T.

Theorem 2:

1) $\Gamma_{N,D}$ does not depend upon D

2) $N \to \Gamma_{N,D}$ is an increasing function with respect to the order of positive matrices.

Proof: 1) is a straightforward consequence of Remark 1. Elementary inversion formulas for partitioned matrices [12] show 2). As a matter of fact, it can be shown that, when (Y_t) is AR, then $\Gamma_{N,D}$ is constant for $N \geqq p$.

III. Relation with the I.V. Method

According to [14], the I.V. estimate of given the record $Y_1, \cdots, Y_s$ is obtained by minimizing

$$J_s(\theta)=\left\| \text{col} \left[\sum_{t=1}^{s} Z_t(N) W_t^T \right] \right\|_Q^2 \tag{3-1}$$

where θ, $Z_t(N)$, W_t are given by (2-1) and (2-4) with A_i instead of A_i^0, and Q is a $Nr^2 \times Nr^2$ symmetric positive definite matrix to be chosen. Setting

$$M_s=\frac{1}{s}\sum_{t=1}^{s} Z_t(N)\varphi_t^T \otimes I_r \tag{3-2}$$

the I.V. estimate is given by

$$\hat{\theta}_{N,Q}(s)=(M_s^TQM_s)^{-1}M_s^TQ\left(\frac{1}{s}\sum_{t=1}^{s} Z_t(N) \otimes Y_t\right). \tag{3-3}$$

Under P_0 we have, for s large, $M_s \approx \bar{H}_{p,N}$, which we shall assume in the sequel to be of full column rank.[1] From (2-5), (3-3) and the fact that $M_s \to \bar{H}_{p,N}$, we get under P_0 that, for s large,

$$\sqrt{s}[\hat{\theta}_{N,Q}(s)-\theta_0] \approx (M^TQ^{-1}M)^{-1}M^TQU_s(N)$$
$$M=\bar{H}_{p,N}^T \tag{3-4}$$

so that the central-limit theorem of [14] is reobtained

$$P_{IV}(N)=(M^TQM)^{-1}M^TQ\Sigma QM(M^TQM)^{-1} \tag{3-5}$$

where M, Σ are as in (2-12) and $P_{\text{I.V.}}(N)$ is the asymptotic covariance matrix of the I.V. estimate. From (2-13), (3-4), and (3-5) one easily gets

$$P_{IV}(N) \geqq \Gamma_N^{-1} \tag{3-6}$$

with equality iff $Q = \Sigma_N^{-1}$. Recall that the reduction matrix D is no more useful, hence the notation Γ_N.

An important consequence is that, when M is of full column rank, the Cramer-Rao inequality can be applied to $P_{\text{I.V.}}(N)$ to yield

$$\Gamma_\infty=\lim_N \Gamma_N \leqq F_{11}-F_{12}F_{22}^{-1}F_{21} \tag{3-7}$$

where

$$F=\begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \tag{3-8}$$

is the Fisher information matrix of the ARMA process Y_t, partitioned according to the AR and MA coefficients.

Remark: We discuss in [6] the reason for which the I.T. is more efficient than comparing the confidence ellipsoid to the I.V. estimate through a χ^2-test.

[1] This is always the case for scalar signals when p is the true AR order.

IV. Robust Optimality and Comparison with the Local Likelihood Ratio Test

In this section, we investigate the connection between local likelihood tests and instrumental tests. Let us first emphasize for which reason the connection between likelihood and instrumental approaches is more difficult to prove for our testing problem than for identification as done in [14]. In the two situations, the problem is to deal with the AR part of an ARMA process, and the key difficulty lies in the coupling effect which exists between the AR and the MA parts: the Fisher information matrix is not block-diagonal.

In [14], Stoica *et al.* consider the I.V. identification method for estimating the AR part and the likelihood method for estimating both the AR and MA parts; and they compare the accuracy, on the AR part, of the two methods. As this accuracy is related to the *inverse* of the Fisher information matrix, the coupling effect is implicitly taken into account (inversion lemma for partitioned matrices).

In our testing problem, we want to test for changes in the AR part, *without knowing* possible changes in the MA part. The likelihood ratio test for testing changes in both the AR and MA parts involves a nontrivial dependency with respect to the MA part. Therefore, for testing for the AR part only, we follow a min-max approach, i.e., we consider the least favorable case for changes in the MA part, in order to eliminate these nuisance parameters. Thus, we first consider the test based upon an asymptotic local expansion of the likelihood ratio test, and then apply it to the present problem of detecting changes in the AR part of the process (Y_t) with elimination of the nuisance parameters due to changes in the MA part.

A. Local Likelihood Ratio Test: Min-Max Approach

To parameterize the ARMA model (1-2), (1-3), introduce the partitioned vector

$$\Psi=\begin{pmatrix} \theta \\ \beta \end{pmatrix}, \ \beta=\text{col}\,(B_{p-1}, \cdots, B_1) \tag{4-1}$$

and denote by Ψ_0 a given nominal model. Consider the normalized gradient of the log likelihood

$$\Delta_s(\Psi)=\frac{1}{\sqrt{s}}\frac{d}{d\Psi}\log L(Y_1, \cdots, Y_s|\Psi). \tag{4-2}$$

From [7] and [13] we know that

$$\text{under } \Psi_0+\frac{\delta\Psi}{\sqrt{s}}, \ \Delta_s(\Psi_0) \underset{s\to\infty}{\longrightarrow} N(F \cdot \delta\psi, F) \tag{4-3}$$

where F is the Fisher information matrix (3-8) under the nominal model Ψ_0, and the associated χ^2 statistics (with dim Ψ degrees of freedom) $X_s = \Delta_s(\Psi_0)^T \cdot F^{-1} \cdot \Delta_s(\Psi_0)$ is the asymptotically uniformly most powerful (UMP) test to detect deviations from this nominal model Ψ_0.

However, since we are interested in monitoring the AR parameters θ only, we shall follow a min-max robust approach by considering the MA parameters β as nuisances. Consequently, to *each possible change* $\delta\theta$, we associate the corresponding least favorable change $\delta\beta_*$ and consider $\delta\Psi^T = (\delta\theta^T, \delta\beta_*^T)$ as a relevant candidate for a possible change in (4-3). It is known [7] that, for a fixed level, the power of the above-mentioned χ^2-test is an increasing function of the parameter

$$(\delta\theta^T, \delta\beta^T)F\begin{pmatrix} \delta\theta \\ \delta\beta \end{pmatrix}.$$

Accordingly, to $\delta\theta$ we associate

$$\delta\beta_*=\arg\min_{\delta\beta}\,(\delta\theta^T, \delta\beta^T)F\begin{pmatrix} \delta\theta \\ \delta\beta \end{pmatrix}=-F_{22}^{-1}F_{21}\delta\theta \tag{4-4}$$

where F is partitioned according to (3-8). According to (4-3), (4-4), and (2-11), the min-max robust likelihood ratio test is given by

$$\chi_{L.R.}(s)=\tilde{\Delta}_s^T(\Psi_0)\tilde{F}^{-1}\tilde{\Delta}_s(\Psi_0)$$

$$\tilde{\Delta}_s(\Psi)=(I-F_{12}F_{22}^{-1})\Delta_s(\Psi)$$

$$\tilde{F}=F_{11}-F_{12}F_{22}^{-1}F_{21} \tag{4-5}$$

where $\Delta_s(\Psi)$ is defined in (4-2). It is easy to see that, under Ψ_0, $\chi_{L.R.}$ is a centered χ^2 (with dim θ degrees of freedom) while under $\Psi_0 + \delta\Psi/\sqrt{s}$ (with $\delta\beta$ arbitrary) $\chi_{L.R.}$ has noncentrality parameter equal to

$$\delta\theta^T\tilde{F}\delta\theta, \tag{4-6}$$

i.e., *independent* of $\delta\beta$. The following theorem results from the previous discussion.

Theorem 3: Consider again $Y_1, \cdots, Y_s$ and choose a fixed level. Denote by $\pi(\chi(s)|\delta\beta)$ the power of a test $\chi(s)$ to test H_0 against H_1 with any possible change $\delta\beta/\sqrt{s}$. Then the following relationships:

$$\lim_{s\to\infty} \pi(\chi_{L.R.}(s)|\delta\beta)=\lim_{s\to\infty} \pi(\chi_{L.R.}(s)|\delta\beta_*) \geqq \lim_{s\to\infty} \pi(\chi(s)|\delta\beta_*) \tag{4-7}$$

hold for any other test $\chi(s)$, where $\delta\beta_*$ is given by (4-4).

B. Min-Max Robust Optimality of the I.T. in the Case of a Scalar Y_t

First, recall that no reduction matrix D is required in I.T. in the scalar case. The purpose of this paragraph is to prove the following theorem, where we refer to (4-5) for the undefined objects:

Theorem 4: i) The following relationship holds for s large, under both H_0 and H_1:

$$\tilde{\Delta}_s(\Psi)=\frac{1}{\sqrt{s}}\sum_{t=1}^{s}\begin{pmatrix} G_1(q^{-1})Y_{t-p} \\ \cdots \\ G_p(q^{-1})Y_{t-p}\end{pmatrix}\cdot W_t. \tag{4-8}$$

In (4-8), the transfer functions $G_i(q^{-1})$ are given by

$$G_{p-i+1}(q^{-1})=\frac{1}{\sigma^2B^2(q^{-1})}Q_i(q^{-1}) \tag{4-9}$$

where σ^2 is the variance of the innovation, while the pair (K_i, Q_i) is the unique solution of the polynomial equation

$$q^{1-i}B(q^{-1})-K_i(q^{-1})A(q^{-1})=q^{1-p}B(q)Q_i(q^{-1}) \tag{4-10}$$

such that $d^0K_i \leqq p-2$, $d^0Q_i \leqq p-1$.

ii) Theorem 3 can be reinforced as follows: for any change $\delta\beta$, we have

$$\begin{aligned}\lim_{N\to\infty}\lim_{s\to\infty} \Psi(\chi_{I.V.}(N,s)|\delta\beta)&=\lim_{s\to\infty}\Psi(\chi_{L.R.}(s)|\delta\beta)\\ &=\lim_{s\to\infty}\Psi(\chi_{L.R.}(s)|\delta\beta_*)\\ &=\lim_{N\to\infty}\lim_{s\to\infty}\Psi(\chi_{I.V.}(N,s)|\delta\beta_*).\end{aligned} \tag{4-11}$$

COMMENTS:

1) Part i) expresses that the robust likelihood ratio test is in fact an instrumental test, since the vector defined inside the brackets in (4-8) belongs to the linear space spanned by the infinite dimensional instrument $Z_t(\infty)$. Consequently, part ii) is a direct consequence of part i).

2) The formulas (4-8)–(4-10) mean that the statistics $\tilde{\Delta}_s$ do not correspond to the use of filtered instruments in I.T. (see Theorem 5), since the relationship $G_i(q^{-1}) = q^{1-i}G_1(q^{-1})$ does not hold in general as the investigation of the ARMA (2, 1) case shows; see [12].

Basic Steps of the Proof (See [12] for Details): It is known [10], [12] that the gradient of the log likelihood is

$$\Delta_s=\frac{\sigma^{-2}}{\sqrt{s}}\sum_{t=1}^{s}B^{-1}(q^{-1})\begin{pmatrix}\varphi_t\\ \epsilon_t\end{pmatrix}\cdot E_t$$

$$\epsilon_t^T=(E_{t-p+1},\cdots,E_{t-1}) \tag{4-12}$$

and φ_t and E_t are defined in (1-2), (2-4). On the other hand, it is easily seen that

$$F_{12}F_{22}^{-1}\frac{1}{B(q^{-1})}\epsilon_t=E_0\left(\frac{1}{B(q^{-1})}\varphi_t|\mathcal{W}_{t-1}\right)$$

$$\mathcal{W}_{t-1}=\text{span}\left(\frac{1}{B(q^{-1})}E_{t-1},\cdots,\frac{1}{B(q^{-1})}E_{t-p+1}\right) \tag{4-13}$$

where $E_0(\cdot|\cdot)$ denotes conditional expectation. Therefore, according to (4-5), (4-12), (4-13), we have

$$\tilde{\Delta}_s=\frac{\sigma^{-2}}{\sqrt{s}}\sum_{t=1}^{s}\tilde{\varphi}_tE_t$$

$$\tilde{\varphi}_t=\frac{1}{B(q^{-1})}\varphi_t-E_0\left(\frac{1}{B(q^{-1})}\varphi_t|\mathcal{W}_{t-1}\right). \tag{4-14}$$

Then, calculating $\tilde{\varphi}_t$ yields for its ith component

$$\sigma^{-2}\tilde{\varphi}_t^{(i)}=B(q)G_i(q^{-1})Y_{t-p} \tag{4-15}$$

for some transfer function G_i, and (4-14), (4-15) give finally (4-8); a careful use of these formulas gives, on the other hand, the characterization (4-9), (4-10); see [12] for details.

C. Using Filtered Instruments in the Scalar Case

Owing to Theorem 4, we shall now show that it is indeed possible to achieve robust optimality with an I.T. test with *finite dimensional filtered* instruments:

$$U_s(N,G)=\frac{1}{\sqrt{s}}\sum_{t=1}^{s}G(q^{-1})\cdot Z_t(N)\cdot W_t. \tag{4-16}$$

In fact, the following theorem holds (compare to Theorem 4 and [14]).

Theorem 5: $U_s(p, B^{-2})$ corresponds to a robust optimal I.T. test.

The proof given in [12] relies on the fact that $\Gamma_{N-1}(G) = \Gamma_N(G)$ holds for $N \geqq p$ with $G(q^{-1}) = B^{-2}(q^{-1})$, while $\Gamma_\infty(G)$ does not depend on G. Note that this statistic does not correspond to a local likelihood ratio test as the comparison to Theorem 4 shows. Note that knowing B requires knowing the MA part of the true system, which is not fully satisfactory since our goal in designing I.T.'s precisely was to ignore this MA part!

V. Conclusion and Discussion

We have investigated the asymptotic power of new instrumental tests (I.T.) which we recently proposed to detect and isolate changes in the AR part of a vector ARMA process. The relationships with the instrumental variable method and local likelihood ratio tests have been analyzed. The design flexibilities of this test family have been investigated and robust optimality has been shown for a subset of the I.T.'s. It has been shown that optimality is achieved with a large number of instruments, or with a small number of suitably filtered instruments.

References

[1] M. Basseville and A. Benveniste, Eds., *Detection of Abrupt Changes in Signals and Dynamical Systems* (Lecture Notes in Control and Information Sciences, No. 77). New York: Springer-Verlag, 1986.

[2] M. Basseville, A. Benveniste, and G. Moustakides, "Detection and diagnosis of abrupt changes in modal characteristics of nonstationary digital signals," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 412–417, May 1986.

[3] M. Basseville, A. Benveniste, G. Moustakides, and A. Rougee, "Detection and diagnosis of changes in the eigenstructure of nonstationary multivariable systems," *Automatica*, July 1987.

[4] M. Basseville, A. Benveniste, G. Moustakides, and A. Rougee, "Optimum sensor location for detecting changes in dynamical behavior," Res. Rep. IRISA 285-INRIA 498, Feb. 1986; also in this issue, pp. 1067–1075; also presented at CDC 86, Athens.

[5] A. Benveniste and J. J. Fuchs, "Single sample modal identification of a nonstationary stochastic process," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 66–74, Jan. 1985.

[6] A. Benveniste, M. Basseville, and G. Moustakides, "The asymptotic local approach to change detection and model validation," Res. Rep. IRISA 307-INRIA

564, Aug. 1986; also in *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 583-592, July 1987.
[7] R. B. Davis, "Asymptotic inference in stationary Gaussian time-series," *Adv. Appl. Prob.*, vol. 5, pp. 469-497, 1973.
[8] L. Le Cam, "Locally asymptotically normal families of distributions," Univ. Calif., Publ. Statist., vol. 3, pp. 37-98, 1960.
[9] G. Moustakides and A. Benveniste, "Detecting changes in the AR parameters of a nonstationary ARMA process," *Stochastics*, vol. 16, pp. 137-155, 1986.
[10] I. V. Nikiforov, *Sequential Detection of Changes in Stochastic Systems*. in (Lecture Notes in Control and Information Sciences, No. 77). New York: Springer-Verlag, 1986.
[11] A. Rougee, "Détection de changements dans les paramètres AR d'un processus ARMA vectoriel: application à la surveillance des vibrations" (in French), Thèse 3ème cycle, Univ. Rennes I, France, Sept. 1985.
[12] A. Rougee, M. Basseville, A. Benveniste, and G. Moustakides, "Optimum robust detection of changes in the AR part of a multivariable ARMA process, revised version," Res. Rep. IRISA 346, Jan. 1987.
[13] G. G. Roussas, *Contiguity of Probability Measures; Some Applications in Statistics*. Cambridge, MA: Cambridge University Press, 1972.
[14] P. G. Stoica, T. Söderström, and B. Friedlander, "Optimal instrumental variable estimates of the AR parameters of an ARMA process," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 11, pp. 1066-1074, 1985.

Periodic Tracking Adaptive Control for Multivariable Systems Having More Outputs Than Inputs

FULI WANG AND SHIJUN LANG

Abstract—**This note presents an adaptive control algorithm for multivariable systems in which the number of outputs is greater than the number of inputs. The algorithm can force the outputs to track arbitrary given reference signals periodically. This is the best tracking performance for systems lacking output function controllability. It has been shown that the tracking period is the upper bound on the controllability index of the controlled system. The proposed algorithm is applicable to multivariable systems with arbitrary interactor matrix but no knowledge of the interactor matrix is required.**

I. Introduction

In recent years, many algorithms have been described in the literature for the adaptive control of multivariable systems. But most of them have focused on the case when the system transfer function is square, i.e., the number of inputs is equal to the number of outputs [1]-[7]. There have been few papers [8], [9] concerned with the case when the system transfer function is nonsquare, especially when the number of outputs is greater than the number of inputs. The rationale for this restriction is, as pointed out in [8], that output function controllability requires that the transfer function have rank equal to the number of outputs and a necessary condition for this is that the number of inputs should be greater than or equal to the number of outputs [11].

In practice, it is sometimes needed to control systems having more outputs than inputs. For example, the automatic control system for an artificial heart [8] has three outputs and two inputs. It is clearly impossible to control the system having more outputs than inputs so that the outputs track arbitrary given reference signals at all sampling times since the system lacks output function controllability. However, the control objective can be stated as that the outputs are controlled to track periodically (at regular sampling intervals) given reference signals. The key problem is how to determine the tracking period. It has been shown, in this note, that the tracking period is the upper bound on the controllability index of the controlled system.

Manuscript received January 5, 1987; revised April 22, 1987.
The authors are with the Department of Automatic Control, Northeast University of Technology, Shenyang, Liaoning, People's Republic of China.
IEEE Log Number 8715601.

The organization of the note is as follows. In Section II we discuss the periodic tracking algorithm for known systems and the determination of the tracking period. In Section III, we discuss the adaptive implementation of the algorithm. In Section IV we give a simulation example. Some conclusions are summarized in Section V.

II. Periodic Tracking Algorithm

Consider a process described by the following matrix polynomial ARMAX model

$$A(q^{-1})y(t)=B(q^{-1})u(t)+C(q^{-1})e(t) \tag{2.1}$$

where $u(t) \in R^m$, $y(t) \in R^p$, $e(t) \in R^p$ are the control, output, and disturbance variables, respectively. Disturbance $e(t)$ is assumed to be white with zero mean value. $A(q^{-1})$, $B(q^{-1})$, and $C(q^{-1})$ are polynomial matrices in the delay operator q^{-1}.

$$A(q^{-1})=I+A_1q^{-1}+\cdots+A_{na}q^{-na}, \quad A_i \in R^{p\times p} \quad (i=1, \cdots, n_a)$$

$$B(q^{-1})=B_1q^{-1}+\cdots+B_{nb}q^{-nb}, \quad B_i \in R^{p\times m} \quad (i=1, \cdots, n_b)$$

$$C(q^{-1})=I+C_1q^{-1}+\cdots+C_{nc}q^{-nc}, \quad C_i \in R^{p\times p} \quad (i=1, \cdots, n_c).$$

It is assumed that det $C(q^{-1})$ has all its roots strictly inside the unit circle. In this note, we consider the case when p is greater than m.

We shall first consider the case $C(q^{-1}) = I$. The case $C(q^{-1}) \neq I$ will be treated before closing Section II. In order to obtain the d-step-ahead optimal predictor of output $y(t)$, we introduce the following equality:

$$I=F(q^{-1})A(q^{-1})+q^{-d}G(q^{-1}), \qquad d>1 \tag{2.2}$$

where

$$F(q^{-1})=I+F_1q^{-1}+\cdots+F_{nf}q^{-nf}, \qquad n_f=d-1$$

$$G(q^{-1})=G_0+G_1q^{-1}+\cdots+G_{ng}q^{-ng}, \qquad n_g=n_a-1.$$

Multiplying (2.1) from the left by $F(q^{-1})$ and using (2.2) gives (note $C(q^{-1}) = I$)

$$y(t+d)=G(q^{-1})y(t)+F(q^{-1})B(q^{-1})u(t+d)+F(q^{-1})e(t+d). \tag{2.3}$$

The term $F(q^{-1})B(q^{-1})$ can be partitioned into two parts

$$F(q^{-1})B(q^{-1})=q^{-1}H_1(q^{-1})+q^{-d-1}H_2(q^{-1}) \tag{2.4}$$

where

$$H_1(q^{-1})=H_0+H_1q^{-1}+\cdots+H_{d-1}q^{-d+1}$$

$$H_2(q^{-1})=H_d+H_{d+1}q^{-1}+\cdots+H_{d+nb-2}q^{-nb+2}.$$

Substituting (2.4) into (2.3) gives

$$\begin{aligned} y(t+d)&=G(q^{-1})y(t)+H_1(q^{-1})u(t+d-1)\\ &\quad+H_2(q^{-1})u(t-1)+F(q^{-1})e(t+d)\\ &=G(q^{-1})y(t)+HU(t)+H_2(q^{-1})u(t-1)+F(q^{-1})e(t+d) \end{aligned} \tag{2.5}$$

where

$$H=[H_0, H_1, \cdots, H_{d-1}]$$

$$U(t)=[u(t+d-1)^T, \cdots, u(t)^T]^T.$$

We consider the situation at time t and assume $U(t)$ has been specified as a function of the data up to time t. Since $F(q^{-1})e(t+d)$ represents future noise, the optimal prediction, say $\hat{y}(t+d/t)$, for the quantity $y(t+d)$ can be obtained from (2.5)

$$\hat{y}(t+d/t)=G(q^{-1})y(t)+HU(t)+H_2(q^{-1})u(t-1). \tag{2.6}$$