

DETECTION OF ABRUPT CHANGES IN THE MODAL CHARACTERISTICS
OF NONSTATIONARY VECTOR SIGNALS.(*)

M. Basseville, A. Benveniste, G. Moustakides, A. Rougée
IRISA
Campus Universitaire de Beaulieu
Avenue du Général Leclerc
35042 Rennes Cédex
FRANCE

An Instrumental Test Statistics is presented to detect changes in the pole part of an ARMA process with time varying MA coefficients; the statistical and system theoretic properties of this test are studied. This method is used in vibration monitoring to detect fatigues in offshore structures through accelerometer measurements.

I. INTRODUCTION

The need for identification as well as change detection procedures for the AR parameters of a vector ARMA process arises in many applications. This is the case for example in vibration monitoring where one wishes to identify and then monitor the vibration characteristics of mechanical systems subject to unknown natural excitation such as swell, fluid, wind, earthquakes... As a matter of fact, in such cases the MA parameters of the process, which reflect the unknown excitation, have to be considered as nuisance parameters, and methods are needed which are robust with respect to those nuisance parameters, and possible changes of them.

In the present paper, the following problem is addressed. Consider a vector ARMA process

$$y_t = \sum_{i=1}^n A_i y_{t-i} + \sum_{j=0}^{n-1} B_j(t) e_{t-j} \quad (1)$$

where e_t is a vector standard white noise, and the MA parameters $B_j(t)$ are time varying. Equivalently, we shall make use of the following state-space form for (y_t) .

$$\begin{aligned} x_{t+1} &= Fx_t + V_{t+1}, \quad \text{cov}(V_{t+1}) = Q_t \\ y_t &= Hx_t \end{aligned} \quad (2)$$

where V_t is a nonstationary white noise with time varying covariance matrix Q_t . The AR parameters A_i of (1) are related as usual to the pair (H,F) in (2). The relevance of those models to some problems in vibration monitoring is recognized, and is discussed in [Prévosto et al., 1982].

(*) This work is supported by IFREMER under contract n° 84/7392 and by CNRS, GRECO 69 "Systèmes Adaptatifs".

statistics U_S under both hypotheses H_0 and H_1 . This will allow us to use Gaussian hypothesis testing procedures. To get our Central Limit Theorem, we have to expand U_S into a vector; this is obtained by replacing U_S by the Kronecker product.

$$U_S = \sum_{t=1}^S Z_t \otimes w_t \quad (14)$$

The proper translation of (12) is then as follows. Rearrange $\{\theta(i,j)\}_{1 \leq i \leq n, 1 \leq j \leq d}$ as

$$\Theta : \Theta(i+d, j) = \theta(i, j) \quad (15)$$

which is the vector obtained by reordering the rows of θ as superimposed columns. Then (12) is replaced by

$$-E_1 U_S = (E_1 \mathcal{H}_S^T(n, N) \otimes I_d) \cdot \frac{\delta \Theta}{\sqrt{S}} \quad (16)$$

On the other hand, we have

$$\begin{aligned} E_0 U_S U_S^T &= \sum_{t=1}^S \sum_{i=1-n}^{n-1} E_0 ((Z_t \otimes w_t) (Z_{t-i} \otimes w_{t-i})^T) \\ &= \sum_{t=1}^S \sum_{i=1-n}^{n-1} E_0 ((Z_t \cdot Z_{t-i}^T) \otimes (w_t \cdot w_{t-i}^T)) \end{aligned} \quad (17)$$

and we shall introduce the corresponding estimate

$$\hat{\Sigma}_S = \sum_{t=1}^S \sum_{i=1-n}^{n-1} (Z_t \cdot Z_{t-i}^T) \otimes (w_t \cdot w_{t-i}^T) \quad (18)$$

Introduce now the following assumptions on the nominal system (3):

A1: The matrix F_0 is full rank and asymptotically stable

A2: There exists a scalar $K > 0$ such that, for every vector λ and every integer k we have

$$E_0 (\lambda^T V_t)^4 \leq K \cdot \|\lambda\|^4$$

A3: The pair (H_0, F_0) is observable.

A4: There exists a nonzero vector G such that for every t we have $Q_t \geq GG^T$.

Note that A.4 is not a controllability condition.

Then, despite the fact that y_t is a nonstationary process, the following theorem holds:

BASIC THEOREM: The assumptions A1 to A4 hold.

(i) nonstationary law of large numbers: for S large, $\frac{1}{S} \hat{\Sigma}_S$ is uniformly positive definite and bounded, and $\hat{\Sigma}_S$ is a consistent estimate of Σ_S , i.e.

$$\Sigma_S^{-1} \hat{\Sigma}_S \rightarrow I \text{ w.p. } 1 \quad (19)$$

under the hypothesis H_0 .

(ii) nonstationary Central Limit Theorem

Under the hypothesis H_0

$$\Sigma_S^{-1/2} \cdot U_S \xrightarrow{D} \mathcal{N}(0, I) \quad (20)$$

whereas, under the hypothesis H_1

$$\Sigma_S^{-1/2} (U_S - (\mathcal{H}_S^T(n, N) \otimes I_d) \cdot \frac{\delta \Theta}{\sqrt{S}}) \xrightarrow{D} \mathcal{N}(0, I) \quad (21)$$

PROOF: see [Moustakides and al., 1985] for the scalar case; most of the material needed for the vector case can be found in this paper. The needed additional results are rather technical, and are currently written.

II.2 - Instrumental statistics

We shall make an extensive use of the following fact about Gaussian hypothesis testing. Assume a statistic U is distributed as $\mathcal{N}(\mu, \Sigma)$.

For testing $\mu = 0$ against $\mu \in \text{range}(A)$, where A is some full column rank matrix, the wellknown Generalized Likelihood Ratio approach ([A.S. Willksy, 1976]) corresponds to the following χ^2 -statistic

$$U^T \Sigma^{-1} A(A^T \Sigma^{-1} A)^{-1} A^T \Sigma^{-1} U$$

which is nothing but the maximum value, with respect to v , of the negative loglikelihood ratio between H_1 and H_0 with $\mu = Av$.

We shall now use the statistics U_S as if we had equality in (20.21) and shall use the above mentioned approach for Gaussian hypothesis testing. Let us investigate what are the possible changes in the mean of the statistic U_S . For this purpose introduce now the following additional assumption. Given a matrix M , we shall denote by

$$\sigma_1(M) \geq \sigma_2(M) \geq \dots \geq 0$$

its singular values. The additional assumption is

$$\left[\text{A5} \quad \liminf_{S \rightarrow \infty} \sigma_r \left(\frac{1}{S} \mathcal{H}_S^T(n, N) \right) \geq \sigma > 0 \text{ w.p. } 1; \right.$$

where r is the dimension of the state in (3).

Together with (A.3), the assumption (A.5) guaranties the uniform minimality of the state space model (3) to represent the process (y_t) . The assumptions (A.1) to (A.5) imply the assumptions (C.1) to (C.4) of [Benveniste and Fuchs, 1985] since

the measure of energy A_S introduced there in (formula (II.5)) is here asymptotically of the order of S . As a consequence of (III.9) and (III.11) of that paper, we get the result

$$\frac{1}{S} \mathcal{H}_S(n, N) = \mathcal{O}_n(H_0, F_0) \cdot \frac{1}{S} \mathcal{C}_N(F_0, G_S) + \epsilon(S) \quad (23)$$

where the pair $(F_0, \frac{1}{S} G_S)$ is uniformly controllable, $\epsilon(S)$ tends to zero w.p.1 under H_0 and H_1 , and $\mathcal{O}_n(H, F)$ denotes the observability matrix

$$\mathcal{O}_n(H, F) = \begin{pmatrix} H \\ HF \\ HF^2 \\ \vdots \\ HF^{n-1} \end{pmatrix}$$

whereas $\mathcal{C}_N(F, \frac{1}{S} G_S)$ is the corresponding controllability matrix, which is shown to be uniformly of full rank. As a consequence of (23) and (12), we can detect any change of the statistic U_S in the range of the hankel matrix $\frac{1}{S} \mathcal{H}_S^T(n, N)$, and we shall not be able to detect any change on θ such that

$$\mathcal{O}_n^T(H_0, F_0) \theta = 0 \quad (24)$$

but such a change does not correspond in fact to any change in the minimal representation (3) of y_t , so that, thanks to (A.5), the statistic U_S , or equivalently U_S , will allow us to detect any change in the pair (H, F) corresponding to the minimal representation (3). This is the best we can expect.

To apply the general principle (22) of Gaussian GLR testing, we must reduce the matrix

$$\frac{1}{S} \mathcal{H}_S^T(n, N) \otimes I_d$$

to a full column rank matrix. To do so, choose any matrix D such that

$$D \cdot \mathcal{O}_n(H_0, F_0) \text{ be invertible} \quad (25)$$

Then, thanks to (23),

$$\frac{1}{S} \mathcal{H}_S^T(n, N) \cdot D^T \otimes I_d \stackrel{\Delta}{=} \underline{\underline{\mathcal{H}}}_S^T \quad (26)$$

has full column rank.

Finally, since $N \geq n$ and D were chosen arbitrarily we derived the following family of Instrumental Test Statistics to detect an unknown change in the nominal pair (H_0, F_0) :

$$x_S = U_S^T \hat{\Sigma}_S^{-1} \underline{\underline{\mathcal{H}}}_S^T (\underline{\underline{\mathcal{H}}}_S \hat{\Sigma}_S^{-1} \underline{\underline{\mathcal{H}}}_S^T)^{-1} \underline{\underline{\mathcal{H}}}_S \hat{\Sigma}_S^{-1} U_S \quad (27)$$

Under the hypothesis H_0 of no change, x_S is approximately a central χ^2 with $N \cdot d^2$ degrees of freedom, whereas x_S is non central if any change occurs. Note that (27) takes the form of an off-line hypothesis testing method (or "model validation"); but there is no problem to design a corresponding on-line change detection method (see [Benveniste, 1985], [Nikiforov, 1985], [André-Obrecht, 1985]).

A special case of Instrumental Test Statistics

The simplest U_S we can design corresponds in fact to use the definition of U_S suggested in (12) (using the empirical Hankel matrix)

$$U_S = \mathcal{H}_S^T(n+1, N) \cdot \begin{bmatrix} \theta_0 \\ -I_d \end{bmatrix} \quad (28)$$

and to reduce it as follows.

STEP 1 : select in $\mathcal{O}_n(H_0, F_0)$ r (=dimension of the state) independent rows to get \mathcal{O}_0 , and assign the value zero to the rows of θ which were not selected in \mathcal{O}_0 ; then, solve the reduced form of (4) to get the reduced nominal model θ_0 .

STEP 2 : select now columns in $\frac{1}{S} \mathcal{H}_S(n, N)$ to get, after row and column reduction, an invertible reduced matrix $\frac{1}{S} \mathcal{H}_S$ of rank r , and set

$$\mathcal{H}_S^{\min} = \begin{bmatrix} \mathcal{H}_S \\ R_n(S), \dots, R_{n+N-1}(S) \end{bmatrix} \quad (29)$$

$$U_S^{\min} = (\mathcal{H}_S^{\min})^T \cdot \begin{bmatrix} \theta_0 \\ -I_d \end{bmatrix} \quad (30)$$

Since \mathcal{H}_S^{\min} is invertible, (27) reduces to

$$x_S^{\min} = (U_S^{\min})^T (\hat{\Sigma}_S^{\min})^{-1} (U_S^{\min}) \quad (31)$$

where U_S^{\min} is obtained from U_S^{\min} by the usual stacking transform, and $\hat{\Sigma}_S^{\min}$ is easily obtained through column and row reductions from $\hat{\Sigma}_S$.

The Instrumental Test Statistics in (31) is the kind of statistics we use in practice for our experimental work on vibration monitoring. But it is by no means optimal (at least theoretically), as the next section will show.

Diagnosis on the origin of the change : a sensitivity method

We shall only present the principles of the method. Since our approach searches for small changes in the AR part of the system, we shall use the following sensitivity method. Parametrize the AR part of the system as

$$\Theta = f(\phi) \quad (32)$$

where ϕ is some minimal parametrization of the AR part (for example poles + modes) and f is \mathcal{C}^∞ in a neighbourhood of the nominal system ϕ_0 . To concentrate on a possible change in a subset of the coordinates of ϕ_0 (say one pole and its associated

mode), denote by \mathcal{J} the matrix obtained by selecting the corresponding columns of the Jacobian $f'(\theta_0)$; then apply the same technique as previously, but replacing $\mathcal{H}_S^T(n, N)$ by $\mathcal{H}_S^T(n, N) \cdot \mathcal{J}$. This method is extensively investigated in the scalar case in [Basseville and al., 1984] and currently experimented in the vector case.

III. OPTIMAL CHOICE OF THE INSTRUMENTS

This section is concerned with the stationary case only ($Q_t = Q$ in (3)). Since we had a large flexibility in designing the tests, it is natural to investigate the optimization problem.

To compare the tests, we shall fix a level α to select the threshold λ through the constraint

$$P_0\{\chi_S > \lambda\} < \alpha \tag{33}$$

then we shall try to maximize the power β of the test, defined by

$$\beta = P_1\{\chi_S > \lambda\} \tag{34}$$

Optimizing N and D

To design the statistic χ , we were free to select the reduction matrix D (see (26)) and the number N of instruments. We have the following result

THEOREM 2 [Rougée, 1985]

- (i) For N fixed, $\beta = \beta(N)$ is independent of the reduction matrix D
- (ii) $\beta(N)$ is increasing with N, so that the optimum is reached for $\beta_\infty = \lim \beta(N)$

Using filtered instruments

Instead of using infinitely many instruments, we can rather try to choose N finite, but replace the instrument Z by a filtered version of it.

To do so we proceed as follows. Introduce the left coprime factorization of the nominal model in its innovations form (5) (where the B_j^i s are now stationary), see [Fuhrmann, 1981]

$$\begin{aligned} y_t &= F_0(z^{-1}) e_t \\ F_0(z^{-1}) &= A_0^{-1}(z^{-1}) B_0(z^{-1}) \end{aligned} \tag{35}$$

Then denote by $\check{B}_0(z^{-1})$ the maximum phase polynomial matrix associated to $B_0(z^{-1})$, i.e., see again [Fuhrmann, 1981]

$$\begin{aligned} \check{B}_0(z^{-1}) \check{B}_0^T(z) &= B_0(z^{-1}) B_0(z) \\ \check{B}_0(z^{-1}) &\text{ maximum phase.} \end{aligned} \tag{36}$$

Then, set

$$Z_t^T = (y_{t-n}^T, \dots, y_{t-2n-1}^T)$$

$$w_t = A_0(z^{-1}) y_t$$

$$U_S^{opt} = \sum_{t=1}^S (Z_t \otimes (B_0^{-T}(z^{-1}) \cdot B_0^{-1}(z))) \cdot w_t$$

$$\Sigma_S^{opt} = E_0(U_S^{opt} U_S^{opt T}) = E_0(\hat{\Sigma}_S^{opt})$$

$$\chi_S^{opt} = U_S^{opt T} \Sigma_S^{opt -1} U_S^{opt} \tag{37}$$

Note that the filter acting on the signal Z_t is properly stable and causal. Then, we have the following theorem :

THEOREM 3 [Rougée, 1985]

- (i) χ_S^{opt} achieves the optimal power β_∞ for every chosen level of the test ;
- (ii) β_∞ is also the power of the classical local test to detect changes in the parameters of the AR process (see [Nikiforov, 1985]) :

$$\check{y}_t = A_0^{-1}(z^{-1}) e_t$$

As a consequence χ_S^{opt} is asymptotically Uniformly Most Powerful. Note that, in the scalar case, (37) reduces to

$$U_S^{opt} = \sum_{t=1}^S (B^{-2}(z^{-1}) \cdot Z_t) \cdot w_t \tag{38}$$

A connection with the work of [Stoica and al., 1984] on the MIV method

The result of the theorems 2 and 3, and especially the formula (38), suggest strong connections with the above mentioned work. To establish such a link in a simple case, assume that $\frac{1}{S} \mathcal{H}_S(n, n)$ is uniformly invertible, i.e. $r=nd$, and we choose the minimum number of instruments (i.e. $N=n$). Let θ_S denote the solution of the IV method ;

$$\mathcal{H}_S^T(n+1, n) \begin{bmatrix} \theta_S \\ -I_d \end{bmatrix} = 0 \tag{39}$$

and denote by Θ_S the vector built from θ_S using (15). Then, for the corresponding design choices in the instrumental statistics U , we have, thanks to (28) and (39),

$$U_S = (\mathcal{H}_S^T(n, n) \otimes I_d) \cdot \tilde{\Theta}_S \quad (40)$$

$$\tilde{\Theta}_S = \Theta_S - \Theta_0$$

so that the connection between U_S and Θ_S under the hypothesis H_0 (i.e. θ_0 is the true system) becomes obvious.

CONCLUSION

We have presented a method to detect changes in the AR part of a vector ARMA process. The instrumental test statistics we have introduced for this purpose have been shown to be robust with respect to the unknown, possibly nonstationary, MA part. We have also investigated the system theoretic aspects of the method, establishing a connection between our approach and the MIV method of identification. The problem of diagnosis has also been investigated. This method is currently used in vibration monitoring, to detect is currently used in vibration monitoring, to detect fatigues or failures inside an offshore structure through accelerometer or strain gauges measurements ; experimental results on large systems will be reported elsewhere.

REFERENCES

- [André-Obrecht, 1985] : R. André-Obrecht "Segmentation du signal de parole, sans reconnaissance". Thèse de 3ème cycle, Univ. Rennes. May 9, 1985.
- [Basseville and al., 1984] : M. Basseville, A. Benveniste, G. Moustakides "Detection and diagnosis of abrupt changes in modal characteristics of nonstationary digital signals". INRIA Research Report n°348. Nov. 84, Submitted to IEEE-IT.
- [Basseville and Benveniste, 1985] : "Detection of abrupt changes in signals and dynamical systems". M. Basseville, A. Benveniste Ed., Advances in Statistical Signal Processing, vol.2, JAI Press.
- [Benveniste and Fuchs, 1985] : A. Benveniste, J.J. Fuchs "Single sample modal identification of a nonstationary stochastic process". IEEE-AC 30 n°1. pp.66-74.
- [Benveniste, 1985] : A. Benveniste "Advanced methods of change detection, an overview", In [Basseville and Benveniste, 1985].
- [Fuhrmann, 1981] : P.A. Fuhrmann "Linear Systems and Operators in Hilbert Spaces" McGraw Hill.
- [Moustakides and al., 1985] : G. Moustakides, A. Benveniste "Detecting changes in the AR parameters of a nonstationary ARMA process". IRISA report n°249, Submitted to Stochastics.
- [Nikiforov, 1985] : I.V. Nikiforov "Sequential detection of changes in Stochastic Systems" In [Basseville and Benveniste, 1985].
- [Prevosto and al., 1982] : M. Prevosto, A. Benveniste, B. Barnouin "Modelling and Identification of the modal characteristics of a vibrating structure". INRIA report n°130. April 1982.
- [Rougée, 1985] : A. Rougée, Thèse 3ème cycle, Univ. Rennes.

[Roussas, 1972] : G.G. Roussas "Contiguity of probability measures, some applications in statistics". Cambridge Univ. Press.

[Stoica and al., 1984] : P. Stoica, T. Soderström, B. Friedlander "Optimal Instrumental Variable estimates of the AR parameters of an ARMA process". SCT report, Palo Alto, Submitted to IEEE-AC.

[Willsky, 1976] : A.S. Willsky "A survey of design methods for failure detection in dynamic systems". Automatica, vol.12, 1976, pp.601-611.