SYSTEM IDENTIFICATION THROUGH FIRST AND SECOND ORDER INFORMATION FROM THE PERIODOGRAM

M. Amengual, G. Vazquez, M. A. Lagunas.

Dept. of Signal Theory and Communications
E.T.S.I. Telecomunicacion. Apdo. 30.002.
08080 Barcelona, SPAIN

ABSTRACT

When approximating linear systems by means of a finite set of parameters, it can be useful the use of first and second order information; that is, information of impulse response and correlation. When that information is not available but one disposes of some samples of a random process which are the result of filtering white noise through the system, some approach must be carried out. Besides the approach taken into account, it seems to be a logical election, as first step, that of estimating the autocorrelation function associated to the random process. This paper deals with one of such approaches which, starting from one estimate of the autocorrelation function, gives rise to an ARMA model for the underlined system. The derivation of the model is achieved from an optimization point of view.

I.- INTRODUCTION

In this paper, it will be assumed that a windowed realization of a widely sense random process is given, obtained by filtering white gaussian noise through the linear system we are trying to identify. From that sequence of random numbers, it is desired to obtain certain information concerning the system and this information must be as deterministic as possible.

In order to obtain a more deterministic information, the first guess we think over is the autocorrelation function r(n).

If x(n) are the data values of the random process at lags n = 0,1,...,N-1, one way of estimating the autocorrelation function is by means of the biased estimator (1):

$$\tilde{r}(n) = \sum_{m=0}^{N-1-|n|} \frac{x(m)x(m+|n|)}{N} \qquad (1)$$

where $\tilde{r}(n)$ is the estimated value and N the number of available data points.

So, it will be assumed that the starting point is the autocorrelation estimate underlined in (1); then, speaking either of periodogram or autocorrelation estimate will be the same, because they are related trough the Fourier transform.

This work has been supported by CAICYT by grant number 21096/84

If h(n) denotes the impulse response of the linear system, the next identity will hold:

$$r(n) = h(n) * h(-n) \qquad (2)$$

where r(n) is the actual autocorrelation function.

Thus, given $\tilde{r}(n)$, one way of obtaining an estimate $\tilde{h}(n)$ of the impulse response is by using any deconvolution technique.

Although when h(n) is given, the associated r(n) is unique, it does not appears to be the same when estimating h(n) from r(n); so, when performing deconvolution, it uses to be quite common to regard as solution the minimum phase one. That is, among all possible sequences $h_i(n)$ such that $r(n) = h_i(n) * h_i(-n)$, select the minimum phase one. Then, it can be performed by the homomorphic deconvolution algorithm, quite easy to use /1/.

An important point that must be taken into account is that of the statistical stability and feasibility of the current information, i.e. $\tilde{r}(n)$ and $\tilde{h}(n)$. It seems to be logical that, besides the bias of the estimator, the uncertainty of the values $\tilde{r}(n)$ grows up as the index "n" becomes larger; it is because when the index "n" becomes greater, the number of averages in (1) decreases. So, first values of r(n) are better estimated than the others; however, it is not clear at all, which is the influence of errors in $\tilde{r}(n)$ over the $\tilde{h}(n)$ estimate .

One approach that can be carried out, in order to improve the statistical behaviour is that of matching an AR model of great order to the spectra of the time series. Thus, if it is considered that the autoregresive model obtained from M values of the autocorrelation (M < N) can approximate well enough the spectra, it will be possible to obtain an estimate of the impulse response with less variance (because the impulse response will just depend on M parameters). These two ways of obtaining the first order information will be compared in the results section.

The organization of the paper is as follows. In section II, it will be studied the interest of using first and second order information in the characterization of linear systems, and necessary conditions will be stated. Section III deals with the specific procedure proposed to identify the system. The method developed is enclosed into the variational framework: among all the rational transfer functions that agree with some impulse response and correlation values,

CH2458-8/88/0000-1341$1.00 © 1988 IEEE

it is selected that one which achieves a minimum for the r(0) value. In section IV, it is derived the algorithm to obtain the solution to the procedure proposed in the previous section. In sections V and VI some examples will be presented as well as the bibliographic references.

II.- FIRST AND SECOND ORDER INFORMATION /2/

When one is trying to approximate the impulse response of a discrete linear system just regarding some values: h(0),h(1),...,h(Q) of such response, it is clear that many solutions can be achieved. Because no information is reported concerning the "tail" of the h(n) sequence, stability is no guaranteed "a priori". So, it appears, as a consequence, the necessity of taking into account the r(0) value; then, if a finite value of r(0) is fulfilled, stability is guaranteed. Besides that value, the inclusion of more information: r(1),... can report knowledge of the global behaviour of the h(n) sequence.

There are, however, two particular cases where just one kind of information can be enough:

* When an AR model is desired, second order information becomes sufficient.

* When a MA model is desired, first order information becomes sufficient.

In the more general case of considering as constraints (i.e. parameters to be fulfilled) the values: h(0,...,h(Q),r(0),...,r(P) and if it is desired a rational model (3) for the transfer function H(z) of the discrete system, some consistency test must be carried out over the available information.

$$H(z) = \frac{b(0)+b(1)z^{-1}+\dots+b(Q)z^{-Q}}{1+a(1)z^{-1}+\dots+a(P)z^{-P}} \quad (3)$$

The consistency test sets the necessary conditions that must be fulfilled for the pretended parametric model, in order to a solution exist.

To see which one must be the consistency test, let us consider that the Q+1 impulse response measures and the P+1 of correlation belong effectively to a system which transfer function is given by (3). If the numerator and denominator polynomials are B(z) and A(z) respectively, H(z) will be:

$$H(z) = \frac{B(z)}{A(z)} = h(0) + h(1)z^{-1} + \dots \quad (4)$$

Matching coefficients, in the previous identity, a matrix expression can be obtained which relates the numerator and denominator vectors. The mentioned identity will be the following:

$$\begin{bmatrix} b(0) \\ b(1) \\ \vdots \\ b(Q) \end{bmatrix} = \begin{bmatrix} h(0) & & \Phi & \\ h(1) & \ddots & & \\ \vdots & & \ddots & \\ h(Q) & h(Q-1) & \dots & h(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ \vdots \\ a(P) \end{bmatrix} \quad (5)$$

where it has been assumed that the number of poles P is greater than the number of zeroes Q and, because of this fact, there are P-Q columns of zeroes in the matrix.

If $\underline{\underline{G}}$ denotes the impulse response matrix and vectors $\underline{b}$ and $\underline{a}$ are given by $\underline{b} = (b(0),b(1),\dots,b(Q))^T$ and $\underline{a} = (1,a(1),\dots,a(P))^T$ (where "T" denotes transpose), equation (5) can be written down in a more compact fashion as:

$$\underline{b} = \underline{\underline{G}}\,\underline{a} \quad (6)$$

To find a solution for the $\underline{b}$ and $\underline{a}$ vectors, another matrix equation is needed relating both vectors. The next identification that can be carried out becomes from the correlation:

$$R(z) = H(z)H(1/z) = \frac{B(z)B(1/z)}{A(z)A(1/z)} \quad (7.a)$$

$$R(z)A(z) = B(z)\frac{B(1/z)}{A(1/z)} \quad (7.b)$$

Identifying coefficients with equal powers of z in (7.b), expression (8) is obtained.

$$\begin{bmatrix} r(0) & r(1) & \dots & r(P) \\ r(1) & r(0) & \dots & r(P-1) \\ \vdots & & \ddots & \vdots \\ r(P) & r(P-1) & \dots & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ \vdots \\ a(P) \end{bmatrix} = \begin{bmatrix} h(0) & h(1) & \dots & h(Q) \\ & h(0) & \dots & h(Q-1) \\ & \Phi & \ddots & \vdots \\ & & & h(0) \end{bmatrix} \begin{bmatrix} b(0) \\ b(1) \\ \vdots \\ b(Q) \end{bmatrix} \quad (8)$$

This last equation can be written in a more compact form as:

$$\underline{\underline{R}}\,\underline{a} = \underline{\underline{G}}^T\underline{b} \quad (9)$$

where $\underline{\underline{R}}$ is regarded as the correlation matrix.

Equations (6) and (9) conform a matrix equations system. Substituting (6) into (9) yields

$$\underline{\underline{R}}\,\underline{a} = \underline{\underline{G}}^T\underline{\underline{G}}\,\underline{a} \quad (10.a)$$

$$(\underline{\underline{R}} - \underline{\underline{G}}^T\underline{\underline{G}})\,\underline{a} = \underline{0} \quad (10.b)$$

Equation (10.b) implies that vector $\underline{a}$, which defines the denominator of the model, must belong to the kernel of the $\underline{\underline{K}}(P,Q)$ matrix:

$$\underline{\underline{K}}(P,Q) = (\underline{\underline{R}} - \underline{\underline{G}}^T\underline{\underline{G}}) \quad (11)$$

The numerator of the model would be obtained through (6).

Before making system identification by using impulse response and correlation estimated values other than exact ones, some considerations must be taken into account concerning $\underline{\underline{K}}(P,Q)$ matrix.

It can be readily shown that $\underline{\underline{K}}(P,Q)$ matrix can be expressed as underlined in (12):

$$\underline{\underline{K}}(P,Q) = \sum_{k=Q+1}^{\infty} \begin{bmatrix} h(k) \\ \cdot \\ \cdot \\ \cdot \\ h(k-P) \end{bmatrix} \left[h(k),\ldots,h(k-P)\right] \qquad (12)$$

Therefore, $\underline{\underline{K}}(P,Q)$ matrix must be a semidefinite positive matrix. This means that when estimated values are used as entries in (11), solutions can not exist if the semidefinite positive condition is not achieved. It is easy to verify that $\underline{\underline{K}}(P-1,Q-1)$ is the matrix that would result when suppressing the first row and the first column in $\underline{\underline{K}}(P,Q)$. Remind that $\underline{\underline{K}}(P-1,Q-1)$ would be formed as $\underline{\underline{K}}(P,Q)$ but regarding just the first Q-1 impulse response values and the first P-1 correlation ones. It can be shown /3/ that when $\underline{\underline{K}}(P,Q)$ is semidefinite positive (necessary condition) and $\underline{\underline{K}}(P-1,Q-1)$ is definite positive then, all the roots of the A(z) polynomial lie inside the unit circle and may be, someone of them, on the unit circle. Clearly, the solution for the $\underline{a}$ vector will be unique in the situation just denoted. However, constraints will only be verified if the A(z) polynomial is strictly minimum phase. In the case that A(z) displays some of its roots on the unit circle, it is clear that not all the constraints can be fulfilled although equations (6) and (9) are still holding; in this situation, it can not be achieved a finite value for r(0) and hence the r(0) constraint can not be fulfilled. If the matrix equations are verified but solution does not exist, means that pole-zero cancellations must occur on the unit circle.

From these considerations, it turns out to be clear which must be the necessary and sufficient conditions such that a set of Q+1 impulse response measures and P+1 correlation ones, can be fulfilled by a model of P poles and Q zeroes.

It is important to remark that $\underline{\underline{K}}(P,Q)$ matrix contains P+Q+2 constraints while the model depends only over P+Q+1 parameters. So if it is not included some freedom degree to the issue of the constraints, the necessary condition noted before will rarely occur. Thus, knowing the necessity for some freedom degree, the problem of finding a model could be stated as the optimization of an objective function subject to constraints. Then, it would be desirable that the value of the objective function (freedom degree) were the parameter which could guarantee the necessary condition of the $\underline{\underline{K}}(P,Q)$ matrix for a solution to exist. The strictly minimum phase condition for the denominator polynomial is not controllable "a priori" just regarding the $\underline{\underline{K}}(P,Q)$ matrix. In the next section, it will be developed a system identification procedure from the variational point of view.

III.- VARIATIONAL STATEMENT

If it is assumed that the measures x(n) (n = 0,1,...,N-1) are corrupted by additive white noise, the value of r(0) will not be accurately estimated.
Then, if one is interested in obtaining an ARMA model for the underlined system, and if no mention is made about the reliability of the h(n) measures (h(n) is obviously affected by the errors in r(0)), it would appear to be a good statement that of wishing a model that minimizing r(0), would verify certain values of impulse response and correlation.

That is, it must be found out a transfer function H(z) such that (13) is minimized subject to the constraints (14):

$$r^{op}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{H(z)H(1/z)\Big|_{z=\exp(jw)}\}\, dw \qquad (13)$$

$$r(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{H(z)H(1/z)\Big|_{z=\exp(jw)}\} e^{jwn}\, dw \qquad (14.a)$$

$$|n| \leqslant P,\ n \neq 0$$

$$h(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \{H(z)\Big|_{z=\exp(jw)}\} e^{jwn}\, dw \qquad (14.b)$$

$$|n| \leqslant Q$$

Where, even r(n) or h(n) would be constraints obtained as commented in section I or in another fashion; and $r^{op}(0)$ is the optimum of the objective function.

By means of the Lagrange multipliers technique, it can be shown that the H(z) function that is solution to the above stated variational problem is like that underlined in (3). It can also be shown that when solution exists, it corresponds to a minimum.

IV.-ALGORITHM

The algorithm to find the model coefficients which are solution of the previous variational problem, selects the value of the objective function in such a way that the corresponding $\underline{\underline{K}}(P,Q)$ matrix becomes semidefinite positive.

If $\underline{\underline{R}}'$ is the same as $\underline{\underline{R}}$ by only changing the entries on the main diagonal by the $r^{op}(0)$ value, the next identity must hold:

$$(\underline{\underline{R}}' - \underline{\underline{G}}^T \underline{\underline{G}})\ \underline{a} = \underline{0} \qquad (15)$$

Writing $r^{op}(0)$ as (16)

$$r^{op}(0) = r(0) - \Delta r(0) \qquad (16)$$

Where r(0) is the measured value of r(n) at origin; equation (15) can be rewritten as:

$$(\underline{\underline{R}} - \underline{\underline{G}}^T \underline{\underline{G}})\ \underline{a} = \Delta r(0)\ \underline{a} \qquad (17)$$

Therefore, the value of the objective function, so as the denominator of the model, depend upon some eigenvalue and associated eigenvector of the $\underline{\underline{R}} - \underline{\underline{G}}^T \underline{\underline{G}}$ data matrix. Since the $\underline{\underline{R}}' - \underline{\underline{G}}^T \underline{\underline{G}}$ matrix must be semidefinite positive, the selected eigenvalue must be the minimum one, although it could seem apparent from (16) that a maximum value for $\Delta r(0)$ was desired.

In such manner, the algorithm would be as follows:

* Estimate in some way the r(n) and h(n) information from the measured values of the time series.
* Construct the $\underline{\underline{R}} - \underline{\underline{G}}^T \underline{\underline{G}}$ matrix with the previous information.

* Look for the eigenvector associated to the minimum eigenvalue; then, it would already be obtained the denominator of the model.
* Find the numerator polynomial through equation (6).

If constraints are obtained as pointed out in section I and solution exists for the model which fulfills the constraints, the model is not guaranteed to be minimum phase: that is, although the sequence $\tilde{h}(n)$ obtained from $\tilde{r}(n)$ is a minimum phase one, the sequence $\tilde{h}(0),\ldots,\tilde{h}(Q)$ is not guaranteed to be it. So, if A(z) is strictly minimum phase, solution exists, constraints are fulfilled, but the final model can be non-minimum phase.

V.- RESULTS

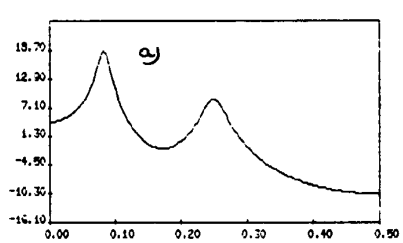
In this section, and just as an example, some numerical results related to the following experiment will be shown. From 640 data points belonging to a windowed realization of an ARMA (4,2) process, 5 estimates are performed by taking into account 128 data points in each one. The initial data points are the result of filtering white gaussian noise of power equal to one through a linear system characterized by the following parameters:

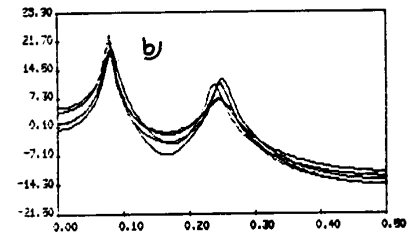
- gain: unity
- zeros: $.6\angle 60^\circ$
- poles: $.95\angle 30^\circ$; $.9\angle 90^\circ$

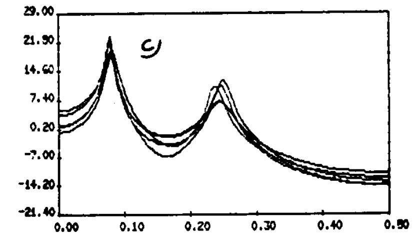
From these data points, correlation values are estimated by means of (1) and then, minimum phase impulse response is evaluated through homomorphic deconvolution. The constraints taken into account in the procedure presented in the previous section are: h(0),h(1),h(2),r(1),r(2), r(3) and r(4).

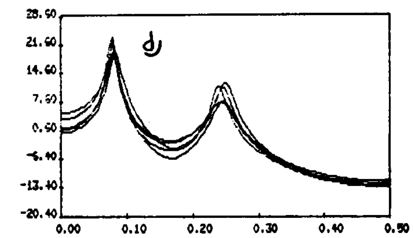
Figure a) depitcs the actual spectra. In figures b), c) and d) are plotted the estimates which result after considering the previous data in the cases of considering no additive noise (b)), and in the cases of considering additive white noise of power 0.01 (c)) and 0.1 (d)). By comparing figures a), b), c) and d), it is concluded that the procedure is quite robust in front of errors in the r(0) value.

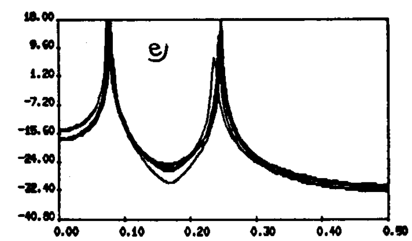
Figures e) and f) are the result of matching an AR model of order 10, before computing the impulse response. In these last two pictures a less variability is observed as well as a greater bias. It is probably due to the fact that errors in r(0) affect to the AR estimate; as less parameters are regarded, the variance seems to be smaller.

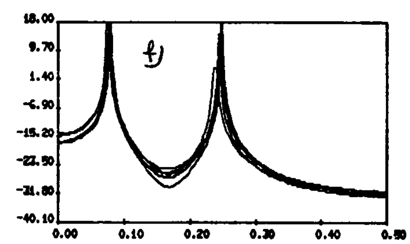












VI.-REFERENCES

/1/ A. V. Oppenheim and R.W. Schafer, "Digital Signal Processing". Prentice Hall, 1975.

/2/ C.T. Mullis and A. Roberts, "The Use of Second-Order Information in the Approximation of Discrete-Time Linear Systems". IEEE Trans. ASSP Vol 24, No 3, June 1976.

/3/ B. R. Musicus and A. M. Kabel, "Maximum Entropy Pole-Zero Estimation". Submitted to the IEEE Trans on ASSP.