



NICE du 20 au 24 MAI 1985

Adaptive Identification of an Auto-
Regressive Non-Linear Filter

Luis F. Rocha

Instituto de Ingeniería Biomédica-Fac.de Ingeniería-UBA-Paseo Colón 850-1063 Buenos Aires

RESUME

L'identification des systèmes de Volterra (c'est à dire des systèmes constitués par deux blocs en cascade: L'un recurrent, et L'autre, une fonction nonlineale) est un problème encore en evolution.

Ici, nous présentons quelques resultats obtenues avec L'emploi d'un "filtre invers" placé en cascade avec le système en course d'être identifié.

De la même façon, le filtre invers est composé par la cascade d'un transfert invers (un polynome de Taylor) avec un filtre lineal invers.

Les coefficients du transfert non-lineal et du filtre lineal sont variés d'accord avec un algorithme de convergence adaptative.

Aussi, nous présentons quelques exemples.

SUMMARY

A linear AR filter followed in cascade by a non-linear-no-memory function transfer is a special class of a Volterra system whose identification is of significative importance in many fields.

In this approach we use the following method: first we adaptively linearize the non-linear transfer, and proceed afterwards to the identification of the AR filter using the standard moving averages or lattice techniques.

Using this method it is possible to identify Volterra systems with high contents of high order exponents in the non-linear portion of the system with great precision.

In the new method, the system to be identified is excited with a known Gaussian, zero mean random sequence, to obtain the corresponding output of the ARNLIF. The identifier is placed in cascade with the system and a certain number of adaptive coefficients, each one controlling a successive exponent of the system's output are varied to linearize the non linear transfer of the system. At the same time, the MA (or lattice) portion of the identifier is adapted to reduce the final error.

Many different types of systems have been identified using normal and fast adaptive algorithms.-



1.- Introduction

Although the techniques for the identification of linear systems are now well established, and have been widely used, the identification of nonlinear systems is still in process of development. Some surveys on this subject can be found in the references [1-5]. In this work we deal with the case known as the "block-oriented system" consisting in a cascade of a linear system (an autoregressive AR filter), followed by a nonlinear gain. A more general system includes another linear filter in cascade with the mentioned two blocks and receives for this reason the name of "sandwich" Volterra system.

In our case the first linear filter is excited by an input signal that will be a discrete-time stationary zero-mean process, and the nonlinear gain is a single-valued nonlinearity.

For the case of finite impulse response filter (corresponding to a "moving averages" MA filter), and finite term polynomial expansion of the nonlinear transfer function, we have presented [6] an adaptive LMS method of parameter identification based on Volterra kernels expansion, resulting in a great but still finite number of identified kernels. In the same work we proposed and tested an approximate solution for an AR (infinite impulse response) filter and weak nonlinearity using the same method.

Obviously as the exact model would require an infinite number of parameters the real model is only approximated, although the LMS method assures minimum least square error between the system and the identification model.

In this paper identification is faced as an "inverse problem" solution linking in cascade with the system to be identified an identifier model builded also as a cascade of an inverse transfer that linearizes the effect of the nonlinear block of the system, followed by the inverse of the AR filter of the system, that is to say a MA or lattice filter. The input sequence $\{u(n)\}$ entering to the cascade of these four blocks emerges as the same sequence $\{u(n)\}$ when the adaptive identifier has evolved to the right

parameters. In the adaptive process the output is an estimated $\hat{u}(n)$, and the difference or error $(u(n) - \hat{u}(n))$ is used to correct the parameters of the identifier using normal or new gradient algorithms [7].

In order to see how the identification works, we will study first how the LMS or steepest descent algorithm can be used to find the inverse polynomial of the nonlinear gain block and afterwards apply the method to the identification of the "memory block" (linear filter) and the nonlinear block.

2.- Adaptive polynomial linearizing

We assume that the nonlinear block, with input x and output y is a single-valued nonlinear transfer $y = P(x)$, where P is a Taylor (or McLoughlin) polynomial such as

$$y = a_0 + a_1 x + a_2 x^2 + \dots$$

Suppose now that another transfer function $z = Q(y)$ is applied on y in order to obtain $z \approx x$, where $z = b_0 + b_1 y + b_2 y^2 + \dots$ given $P(x)$, we want to get the coefficients of $Q(y)$. That is $z = Q[P(x)]$ such that $z \approx x$ using the adaptive LMS steepest descent algorithm of Widrow-Hoff for each coefficient

$$b_i(n+1) = b_i(n) + \mu_i \varepsilon(n) y^i(n) \quad 0 \leq i \leq N$$

where $b_i(n+1)$ is the i -th coefficient of $Q(y)$ at the instant $(n+1)$, μ_i is a constant and $\varepsilon(n) = (x(n) - z(n))$ is the error.

Although the terms of the Taylor Polynomial $P(x)$ are not orthogonal slowing the convergence of the coefficients, the use of a gaussian noise at the input x , makes it possible to find the coefficients b_i after a certain number of iterations.

Lets show an example: $y = P(x) = 0.3 + 1.5x + 0.5x^2$ and $z = b_0 + b_1 y + b_2 y^2 + b_3 y^3 + b_4 y^4$

The polynomial degree of $Q(y)$ has to be in general higher than the degree of $P(x)$.

We use the following scheme (Fig. 1). The sequence $x(n)$ is zero-mean gaussian.

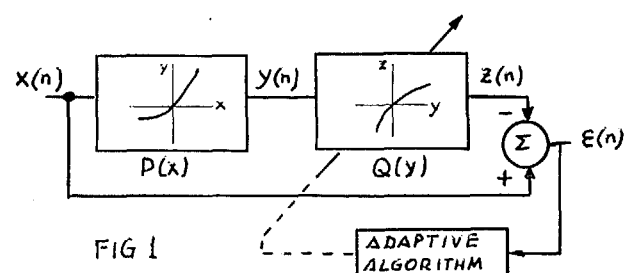


FIG 1

After 2000 iterations we get the following results: $b_0 = -0.217143$; $b_1 = 0.775704$
 $b_2 = -0.223506$; $b_3 = 0.122899$; $b_4 = -0.0482213$

No higher exponents were used, so $Q(y) = b_0 + b_1 y + b_2 y^2 + b_3 y^3 + b_4 y^4$

is the "almost inverse" polynomial of $P(x)$ in the sense of $Q[P(x)] \approx x$

3.- Adaptive AR inversion

Following the same idea, the adaptive inversion of an autoregressive linear process is a well established method. The problem can be seen as the determination of the coefficients α_i of an AR filter of order N having the form

$$x(n) = \sum_{i=1}^N \alpha_i x(n-i) + u(n)$$

knowing the input and output sequences $\{u(n)\}$ and $\{x(n)\}$ respectively.

In the special case of a white noise gaussian input sequence, it is not necessary to know the samples, and the coefficients α_i

can be uniquely determined by the second-order statistic of the process, solving the classical Yule-Walker equations.

Nevertheless in our case the solution can be seen as the linear inverse of an AR filter of order N (having N poles) which is the well known MA filter (having the same N zeros)

The method used to find the coefficients is similar to the preceding one (Fig. 2)

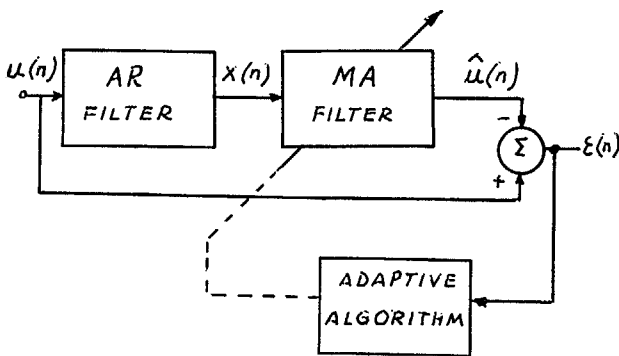


FIG 2

where the estimated coefficients $\hat{\alpha}_i$ are adapted by the algorithm:

$$\hat{\alpha}_i(n+1) = \hat{\alpha}_i(n) + \mu \cdot \epsilon(n) \cdot x(n-i) \quad 0 \leq i \leq N$$

$$\epsilon(n) = u(n) - \hat{u}(n)$$

and the MA adaptive filter is

$$\hat{u}(n) = \sum_{i=0}^N \hat{\alpha}_i \cdot x(n-i) \quad N_1 \geq N$$

If the order N of the AR filter is unknown, the identifier filter MA can be started with an estimated $N_1 > N$ but the convergence speed will decrease.

Occasionally the linear filter preceding the nonlinear gain (or transfer) has AR and MA coefficients (ARMA filter) such as:

$$x(n) = \sum_{i=1}^N \alpha_i x(n-i) + \sum_{j=0}^M \beta_j u(n-j)$$

This filter has N poles M zeros in its transfer function, and its inverse (an MA-AR filter) has to have N zeros and M poles superimposed and in opposition to those of the ARMA filter.

The method of adaptive LMS (steepest descent) works well in this case. The coefficients $\hat{\alpha}_i$ and $\hat{\beta}_j$ of the MA-AR filter

$$\hat{u}(n) = \sum_{i=0}^N \hat{\alpha}_i x(n-i) + \sum_{j=1}^M \hat{\beta}_j \hat{u}(n-j)$$

are computed from:

$$\hat{\alpha}_i(n+1) = \hat{\alpha}_i(n) + \mu_1 \epsilon(n) \cdot x(n-i) \quad 0 \leq i \leq N$$

$$\hat{\beta}_j(n+1) = \hat{\beta}_j(n) + \mu_2 \epsilon(n) \cdot \hat{u}(n-j) \quad 1 \leq j \leq M$$

where $\epsilon(n) = u(n) - \hat{u}(n)$

4.- Adaptive identification of an AR nonlinear system:

Previous works on the identification of parameters in non-linear systems can be grouped in 1) non-linear filtering problem [3] 2) difference equations where the nonlinear parameters are operated linearly [9], 3) state-space approach with full input-output measurements [10]. In all the cases a "parallel" identifier is used (see Fig. 3) where the same (or noise corrupted) input is used to excite an "estimated" coefficients model in order to obtain a similar "plant" output.-



Adaptive Identification of an Auto-Regressive Non-Linear Filter

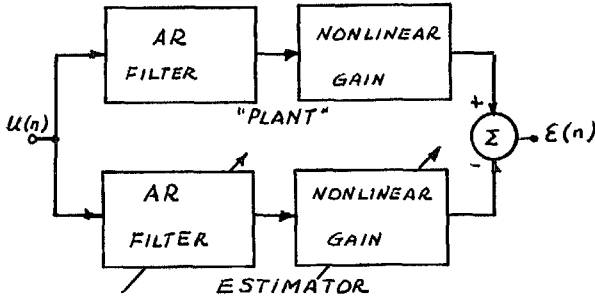


FIG 3

In this approach, we are looking for an inverse system where the "plant" output will be transformed into an estimated "plant" input (see Fig. 4).

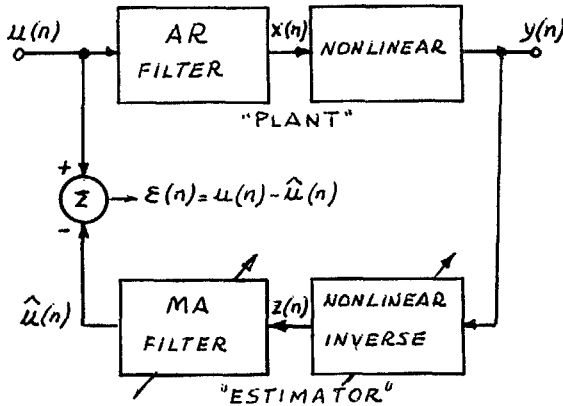


Fig. 4

This is a limited case of the general inverse problem: to find the operator T^* such $T^*[T(u)] = u$

The following relations hold: (n=0,1,2,3, ...)

$$x(n) = u(n) + \sum_{i=1}^L \alpha_i x(n-i) \quad (\text{AR filter})$$

$$y(n) = \sum_{j=0}^M a_j x^j(n) \quad (\text{Taylor expansion})$$

$$z(n) = \sum_{i=0}^N b_i y^i(n) \quad (\text{Inverse nonlinear block})$$

$$\hat{u}(n) = \sum_{\ell=0}^Q \beta_{\ell} z(n-\ell) \quad (\text{MA inverse filter})$$

$$E(n) = u(n) - \hat{u}(n) \quad (\text{error})$$

The coefficients b_i and β_{ℓ} will be varied in an adaptive way in order to reduce the error to a minimum (mean square criterium)

The gradient algorithm is used to update the coefficients:

$$b_i(n+1) = b_i(n) + \mu_i \{ y(n) \}^i \cdot E(n)$$

$$\beta_{\ell}(n+1) = \beta_{\ell}(n) + \mu_{\ell} \cdot z(n-\ell) \cdot E(n)$$

where μ_i and μ_{ℓ} are constant (or very slow decreasing) convergence factors.

This algorithm is intended to make the coefficients b_i and β_{ℓ} converge to those values that minimize the quadratic mean deviation of the difference $E(n)$, that is, of $E[\{u(n) - \hat{u}(n)\}^2]$ a function of both sets of coefficients.

In spite of its simplicity, the adaptive stochastic gradient algorithm is (in the words of Ljung [11]) "... surprisingly more difficult to analyze". (than the recursive least squares RLS algorithm). Several authors have addressed the convergence problem but some of the assumptions have not been fulfilled in the practice [12-14]. Nevertheless, the algorithm proposed in this communication has worked well in all of the many cases considered.

In order to clarify the method employed let us put an example: the "plant" will consist of a simple AR linear filter followed by a nonlinear (logarithmic) gain. No noise is present.

$$\left. \begin{aligned} x(n) &= u(n) + 1.41x(n-1) - 0.6x(n-2) \\ y(n) &= \ln[x(n) + 2.3] \end{aligned} \right\} \text{plant}$$

the identifier will consist of a (truncated) Taylor expansion of 6 terms followed by a MA linear filter of 3 coefficients.

$$\left. \begin{aligned} z(n) &= b_0 + b_1 y(n) + b_2 y^2(n) + \dots \\ \hat{u}(n) &= \beta_0 z(n) + \sum_{\ell=1}^3 \beta_{\ell} z(n-\ell) \end{aligned} \right\} \text{identifier}$$

A uniform noise distribution is used as input and zero starting coefficients.

The variable $x(n)$ was forced to be > -2.3 to avoid problems with the LOG function, the constants are $\mu_{\ell} = 0.05$ ($\ell=1,2,3$) and $\mu_i = 0.05$ ($i=1,2,3$) and the three last μ_i ($i=4,5,6$) are respectively 0,005, 0,005 and 0,001.-

Every 200 samples the relative error was calculated using

$$\overline{ERROR} = \frac{\sum_{n=200}^{M-1} \epsilon^2(n)}{\sum_{n=200}^{M-1} u^2(n)}$$

and the evolution of the coefficients was listed. In a typical run after 6000 cycles the error falls down from 1 to about $6 \cdot 10^{-3}$ (Fig. 5) remaining in similar levels. In order to decrease it

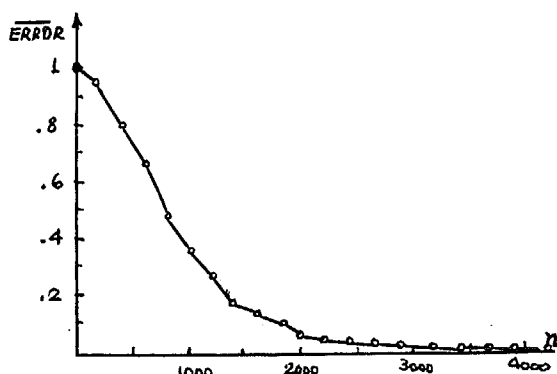


Fig. 5

it will be necessary to increase the number of polynomial terms. There exists the possibility of using a time "window" to weight μ_1 and μ_2 but because our "plant" in this case has not "noise" we prefer to use constants μ'_s avoiding the possibility of biasing the coefficients if the μ 's decrease very fast.

The coefficients values, after 6000 iterations result: $b_1 = 1.15801$; $b_2 = 0.815005$; $b_3 = 0.452143$

$$b_4 = 0.14179; \quad b_5 = 0.0799121; \quad b_6 = 0.0392296$$

$$\text{and } \beta_0 = 1.03274; \quad \beta_1 = -1.44775; \quad \beta_2 = 0.608411$$

for the inverse polynomial coefficients, -and for the MA inverse filter -respectively.

5.- Conclusion

Many experiments confirm the possibility of using an inverse nonlinear filter to identify a Volterra system "plant" (an AR linear filter cascaded with a nonlinear gain) using the (observable) input and output sequences of the "plant". The algorithm used to find the coefficients of the polynomial expansion of the inverse gain and the MA linear filter is one of the variants of the gradient method of Widrow-Hoff.

This method admits to be seen as a nonlinear deconvolution method, because the inverse filter can be used to deconvolve the output of

the Volterra system. In this case, a two steps operation is necessary: a "learning" mode to find the inverse and a "deconvolving" mode after it.-

6.- Acknowledgments

The author acknowledges the fruitful discussions maintained with Mr. C. Orda and the programming work of Mrs. R. Giennini of our Institute.-

References

- (1) R.K. MEHRA "Nonlinear system identification, selected survey and recent trends" 5th IFAC Symposium 1979, pp 77-83.-
- (2) P.E. CROUCH "Polinomic system theory, a Review" IEE Proc.Contr. Theory and Appl. vol 127,(5), pp. 220-228, 1980.-
- (3) J.F. BARRETT "Bibliography on Volterra series, Hermite function expansions and related subjects" TH Rept. 77-E-71 Eindhoven Univ. 1979.-
- (4) S.A. BILLINGS "Identification of nonlinear systems a survey" IEE Proc. vol 127,n.6 part D, pp.272-285, nov.1980.-
- (5) M.SCHETZEN "Nonlinear System Modelling based on the Wiener Theory" Proc. IEE vol. 69 n.12 pp. 1557-1573 Dec. 1981.-
- (6) L.F. ROCHA "Methode adaptative pour l'identification des systemes Volterra" Congres WOGSC - 84, Paris N. Wo139.-
- (7) E.WALACH and B. WIDROW "The least mean - fourth (LMF) Adaptive algorithm and its family" IEEE Trans. Inf.Theor. IT-30 n.2 pp.275-283 march 1984 (Part I).-
- (8) T.SOEDA and T.YOSHIMURA "A note on the estimation of state variables and unknown parameters in a nonlinear systems" IEEE Trns. Autom.Control AC-14 pp. 585-587 Oct.1969.-
- (9) TC. HSIA and M. Ghandi "An iterative algorithm for nonlinear discrete system identification Int. Jour. Control vol.16,N.4, pp. 763-775, 1972.-



-
- (10) C.H. KNAPP and P.K. PAL "Parameters identification in a class of Nonlinear Systems" IEEE Trans. on Aut. Control AC-28 No.4, pp. 497-503, April 1983.-
- (11) L.LJUNG "Analysis of stochastic gradient algorithms for Linear regression problems" IEEE Trans. Inf. Theory IT-30, n.2 pp. 151-160, March 1984.-
- (12) D. FARDEN "Stochastic approximation with correlated data" IEEE Trans. Inf. Theory IT-27 n.1 pp. 105-113, 1981.-
- (13) E.EWEDA and O. MACCHI "Convergence of an Adaptive Linear Estimation Algorithm" IEEE Trans. and Aut. Contr. AC-29 No 2 pp. 119-127, Feb. 1984.-