THE IDENTIFICATION OF THE DISCRETE-TIME HAMMERSTEIN MODEL

By

R. HABER and L. KEVICZKY

Department of Automation, Technical University Budapest
(Received July 10, 1973)

Presented by Prof. Dr. F. CSÁKI

Introduction

Nonlinear systems have many different types. Since there is no simple mathematical method for the description of different structures (and is unlike to be in the future) only special tasks of nonlinear systems parameter estimation could be solved. In the same way as the impulse response for linear systems the VOLTERRA series expansion means a non-parametric system description for a wide class of nonlinear systems. During parametrizing this linear form, approximating the series of "twice infinite size" (infinite in time and in order of expansion), a functional relationship — linear in parameters — can be obtained which considerably simplifies the identification procedure. Unfortunately, the number of necessary parameters is too great for many practical cases [2].

The methods elaborated for the identification of linear discrete-time systems can simply be extended for a special class of nonlinear systems — i.e. for the HAMMERSTEIN model where the zero memory nonlinearity is followed by linear dynamics.

Considering a discrete-time system, assuming second-order polynomial as a nonlinearity and using the impulse transfer function, the HAMMERSTEIN model is shown in Fig. 1. The equation of the no-memory nonlinearity is

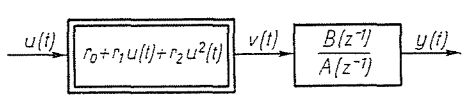

Fig. 1

$$v(t) = r_0 + r_1 u(t) + r_2 u^2(t) \tag{1}$$

and the difference equation for the discrete-time system is (assuming unit sampling time)

$$A(z^{-1})y(t) = B(z^{-1})v(t) \tag{2}$$

where

$$A(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n} \tag{3}$$

and

$$B(z^{-1}) = 1 + b_1 z^{-1} + \ldots + b_m z^{-m}; \qquad m \leqq n. \tag{4}$$

(Here z^{-1} is a backward shift operator: $z^{-1}u(t) = u(t-1)$. Notice that the first term of the polynomial $B(z^{-1})$ is taken as unity for the sake of a safely unambiguous description.) On the basis of (1) and (2) input and output are related by:

$$\begin{aligned} y(t) &= \sum_{i=0}^{m} b_i v(t-i) - \sum_{i=1}^{n} a_i y(t-i) = \\ &= r_0 \sum_{i=0}^{m} b_i + r_1 \sum_{i=0}^{m} b_i u(t-i) + r_2 \sum_{i=0}^{m} b_i u^2(t-i) - \\ &- \sum_{i=1}^{n} a_i y(t-i). \end{aligned} \tag{5}$$

This equation formally corresponds to a "multiple input" single output system as illustrated in Fig. 2. Introducing the $(n+2m+3)$ vector $\mathbf{g}$:

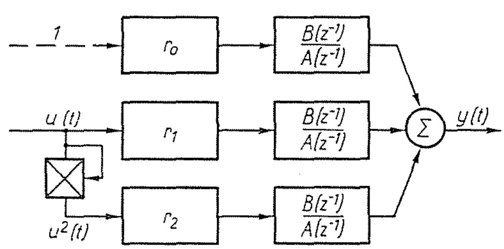

Fig. 2

$$\begin{aligned} \mathbf{g}(t) = [1, u(t), \ldots, u(t-m), u^2(t), \ldots, u^2(t-m), \\ -y(t-1), \ldots, -y(t-n)]^T \end{aligned} \tag{6}$$

and the parameter vector $\mathbf{p}$:

$$\mathbf{p} = [r_0^*, r_1, r_1 b_1, \ldots, r_1 b_m, r_2, r_2 b_1, \ldots, r_2 b_m, a_1, \ldots, a_n]^T \tag{7}$$

where

$$r_0^* = r_0 \left(1 + \sum_{i=1}^{m} b_i\right). \tag{8}$$

Eq. (5) can be given as a scalar product

$$y(t) = \mathbf{g}^T(t)\mathbf{p}. \tag{9}$$

(Here T means the transposition.) Thus we get a system equation having linearity in parameters, the vector $\mathbf{p}$ has, however, redundant elements, since

$$b_i = \frac{p_{2+i}}{p_2} = \frac{p_{3+m+i}}{p_{3+m}}; \qquad 1 \leq i \leq m. \tag{10}$$

If the full parameter vector is estimated then the estimates of numerator of dynamics may be different for the linear part

$$B_1(z^{-1}) = 1 + b_{11} z^{-1} + \ldots + b_{1m} z^{-m} \tag{11}$$

and for the quadratic part

$$B_2(z^{-1}) = 1 + b_{21} z^{-1} + \ldots + b_{2m} z^{-m}. \tag{12}$$

Further generalisation and extension of the model are presented in [3].

For practical applications the measurement noises must be taken into account. The majority of methods — in lack of sufficient a priori information —

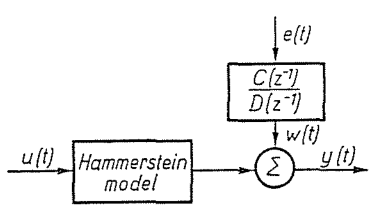

Fig. 3

require to measure the input signal without noise. At the output, however, a linear noise model having rational spectrum is assumed according to practical experiences, this situation is presented in Fig. 3. Here

$$C(z^{-1}) = 1 + c_1 z^{-1} + \ldots + c_k z^{-k}; \qquad k \leq l \tag{13}$$

and

$$D(z^{-1}) = 1 + d_1 z^{-1} + \ldots + d_l z^{-l}. \tag{14}$$

The source noise $e(t)$ is assumed to be normally distributed white noise with variance one and independent of $u(t)$.

The system to be identified is considered as a structurally stable one with constant parameters and that every root of polynomial $z^k C(z^{-1})$ is assumed to lie inside the unit circle. In this paper the possibilities of the off-line parameter estimation are considered, i.e. for N values of $(u(t), y(t))$, where $u(t)$ is a "persistently and sufficiently exciting" signal [2].

1. Iterative Estimation Technique

NARENDRA and GALLMANN (1966) suggested an iterative technique for the identification of the HAMMERSTEIN model [9]. Arrange the N values $u(t)$ and $y(t)$ $(t = 1, 2, \ldots, N)$ to $N \times 1$ vector $\mathbf{u}$ and $\mathbf{y}$ and define the vector $\mathbf{v}$ similarly. Furthermore, introduce the vectors

$$\mathbf{p}_1 = [r_0, r_1, r_2]^T \quad (15)$$

$$\mathbf{g}_1(t) = [1, u(t), u^2(t)]^T \quad (16)$$

and

$$\mathbf{p}_2 = [b_0, b_1, \ldots, b_m, a_1, a_2, \ldots, a_n]^T \quad (17)$$

$$\begin{aligned}\mathbf{g}_2(t) = [&v(t), v(t-1), \ldots, v(t-m), \\ &-y(t-1), -y(t-2), \ldots, -y(t-n)]^T .\end{aligned} \quad (18)$$

Determination of the parameters consists of the following steps:

(a) Assume that $B(z^{-1})/A(z^{-1}) = 1$, thus $\hat{v}_1(t) = y(t)$, where $\hat{v}$ refers to the estimated value.

(b) A simple least-squares (LS) estimation is made for $\mathbf{p}_1$ on the basis of $\mathbf{u}$ and $\hat{\mathbf{v}}_1$, i.e.

$$\hat{\mathbf{p}}_1 = (\mathbf{G}_1^T \mathbf{G}_1)^{-1} \mathbf{G}_1^T \hat{\mathbf{v}}_1 . \quad (19)$$

Here

$$\mathbf{G}_1 = \begin{bmatrix} \mathbf{g}_1^T(1) \\ \vdots \\ \mathbf{g}_1^T(N) \end{bmatrix} . \quad (20)$$

(c) Estimate $v(t)$ again:

$$\hat{v}_2(t) = \mathbf{g}_1(t)\hat{\mathbf{p}}_1 . \quad (21)$$

(d) Make a LS estimation for $\mathbf{p}_2$ on the basis of $\hat{\mathbf{v}}_2$ and $\mathbf{y}$:

$$\hat{\mathbf{p}}_2 = (\mathbf{G}_2^T \mathbf{G}_2)^{-1} \mathbf{G}_2^T y \quad (22)$$

where

$$\mathbf{G}_2 = \begin{bmatrix} \mathbf{g}_2^T(1) \\ \vdots \\ \mathbf{g}_2^T(N) \end{bmatrix} . \quad (23)$$

Naturally, now $\hat{v}_2$ is replacing v in $\mathbf{g}_2(t)$.

(e) After these, estimate $v(t)$ from $y(t)$ by $\hat{\mathbf{p}}_2$:

$$\hat{v}_1(t) = y(t) + \sum_{i=1}^{n} \hat{a}_i y(t-i) - \frac{1}{\hat{b}_0} \sum_{i=1}^{m} \hat{b}_i v_1(t-i) . \quad (24)$$

Here the unambiguity is guaranteed with division by $\hat{b}_0$. The iteration is continued from (b) up to a sufficient accuracy or a given iteration number. Having some a priori information about the parameters of dynamic or nonlinear part of the system, the iteration can be joined at an other point, to the meaning.

2. Estimation Technique with Restrictions

Besides the previous method, iterative technique can also be elaborated for the estimation of the HAMMERSTEIN-model parameters [7]. Let us redraw Fig. 2 to be equivalent for the input and output points, Fig. 4. Let us introduce the vectors:

$$\mathbf{p}_3 = [r_0^*, r_1, r_2, a_1, \ldots, a_n]^T \tag{25}$$

$$\mathbf{g}_3(t) = [1, u^*(t), u^{2*}(t), -y(t-1), \ldots, -y(t-n)]^T \tag{26}$$

where

$$u^*(t) = B(z^{-1})\, u(t) = \sum_{i=0}^{m} b_i\, u(t-i) \tag{27}$$

$$u^{2*}(t) = B(z^{-1})\, u^2(t) = \sum_{i=0}^{m} b_i\, u^2(t-i). \tag{28}$$

The steps of iteration:

(a) The initial estimation of $\hat{v}_1(t)$. (For example, according to iterative estimation technique (a), (b), and (c).)

(b) A simple least-squares estimation for $\mathbf{p}_2$ in knowledge of $\hat{\mathbf{v}}_1$ and $\mathbf{y}$:

$$\hat{\mathbf{p}}_2 = (\mathbf{G}_2^T\, \mathbf{G}_2)^{-1}\, \mathbf{G}_2^T\, y\,.$$

(c) $u^*(t)$ and $u^{2*}(t)$ are estimated according to (27) and (28), where $\hat{B}(z^{-1})$ is determined from $\hat{\mathbf{p}}_2$ by (17).

(d) Least-squares estimation of $\mathbf{p}_3$ on the basis of measured values $\mathbf{y}$ and computed values $\hat{\mathbf{u}}^*$ and $\hat{\mathbf{u}}^{2*}$:

$$\hat{\mathbf{p}}_3 = (\mathbf{G}_3^T\, \mathbf{G}_3)^{-1}\, \mathbf{G}_3^T\, y \tag{29}$$

where

$$\mathbf{G}_3 = \begin{bmatrix} \mathbf{g}_3^T(1) \\ \vdots \\ \mathbf{g}_3^T(N) \end{bmatrix}. \tag{30}$$

(e) $\mathbf{p}_1$ can be computed from $\hat{\mathbf{p}}_2$ and $\hat{\mathbf{p}}_3$ (see (15), (17) and (25)), so

$$\hat{v}_2(t) = \mathbf{g}_1(t)\hat{\mathbf{p}}_1 \tag{31}$$

repeating these steps from (b).

During the identification from noisy measurements it must be taken into account that — in these algorithms — the necessary condition of unbiased parameter estimation is $C(z^{-1}) = 1$ and $D(z^{-1}) = A(z^{-1})$. Namely, in this case the "equation error" in (22) is white noise (equals $e(t)$) and the LS estimation coincides with the maximum likelihood (ML) one [2]. Though this noise structure is necessary for an unbiased estimation, but this is not a sufficient

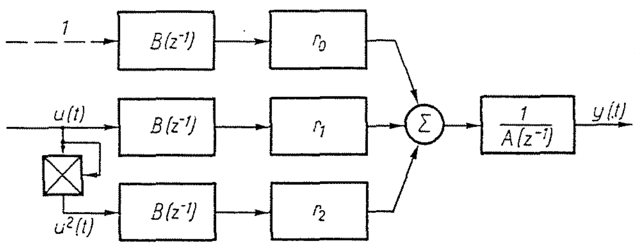

Fig. 4

condition in both previously iterative techniques presented above since e. g. the "input" signal and the equation error are correlated in (22) because of (24). Additional problem is that the solution obtained by these procedures is not surely the best one (only a local minimum).

3. Direct (Noniterative) Estimation Technique

Hsia (1968) and Chang (1971) suggested a direct method instead of the iterative technique for a special case and for the Hammerstein model, respectively [5]. It is shown in the previous item that both the simple and the generalized Hammerstein model can be described by Eq. (9), linear in parameters. This means that in case of $C(z^{-1}) = 1$ and $D(z^{-1}) = A(z^{-1})$ the ordinary LS estimation gives an unbiased estimation of $\mathbf{p}$, i.e.

$$\hat{\mathbf{p}} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T y . \tag{32}$$

Here

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}^T(1) \\ \vdots \\ \mathbf{g}^T(N) \end{bmatrix} . \tag{33}$$

This method permits considerable saving in computation time compared to the iterative technique. In each case it has to be investigated which the better estimation of b_i is because the parameter vector (7) in the linearized equation of the simple Hammerstein model has redundant elements, see (10). This can be performed by the investigation of covariance matrix $(\mathbf{G}^T\mathbf{G})^{-1}$ or by comparison of squared sums of residuals. Estimation techniques can be elaborated for other noise structures, too, applying the methods used for the identification of linear discrete-time systems. In the next item the generalized-least-squares (GLS) method [6] and the ML method [1] are extended for the HAMMERSTEIN model.

4. Generalized-Least-Squares Method

The GLS method suggested by CLARKE is actually a special case of AITKEN estimation for the linear discrete-time systems. The estimation procedure consists of the following steps [6]:

(a) An ordinary LS estimation for the parameters of (9) according to (32).

(b) Computation of the residuals (equation errors)

$$f(t) = y(t) - \mathbf{g}^T(t)\mathbf{p}\ ; \qquad t = 1, 2, \ldots, N \tag{34}$$

by the estimated parameters of $\mathbf{p}$.

(c) Assume that the equation of noise model is

$$f(t) = \mathbf{q}^T(t)\,\mathbf{h}\,, \tag{35}$$

i.e. linear in parameters, where

$$\mathbf{q}(t) = [f(t-1), f(t-2), \ldots, f(t-s)]^T \tag{36}$$

and

$$\mathbf{h} = [h_1, h_2, \ldots, h_s]^T\,. \tag{37}$$

Thus the LS estimation of parameter vector $\mathbf{h}$ can be determined on the basis of values $f(t)$ computed according to (34):

$$\hat{\mathbf{h}} = (\mathbf{F}^T\,\mathbf{F})^{-1}\,\mathbf{F}^T\,\mathbf{f} \tag{38}$$

where

$$\mathbf{F} = \begin{bmatrix} \mathbf{q}^T(1) \\ \vdots \\ \mathbf{q}^T(N) \end{bmatrix} \tag{39}$$

and vector $\mathbf{f}$ contains the N values of $f(t)$. Eq. (35) is valid only for $C(z^{-1}) = 1$ and $D(z^{-1}) = A(z^{-1})\,H(z^{-1})$, where

$$H(z^{-1}) = 1 + h_1 z^{-1} + \ldots + h_s z^{-s}\,. \tag{40}$$

Thus the GLS method gives an asymptotically unbiased estimation. In general, the noise structure assumed above fairly approximates other noise structures. The approximation is the better, the more exactly it realizes the condition:

$$H(z^{-1}) \cong D(z^{-1})/[A(z^{-1})C(z^{-1})] .$$

(d) Compute the filtered values multiplying both sides of Eq. (9) by $H(z^{-1})$:

$$u_1^F(t) = H(z^{-1})u(t) \; ; \quad u_2^F(t) = H(z^{-1})u^2(t)$$
$$y^F(t) = H(z^{-1})y(t) . \tag{41}$$

For the HAMMERSTEIN model the system equation with the filtered values is

$$y^F(t) = \mathbf{g}_F^T(t)\,\mathbf{p}_F \tag{42}$$

where

$$\mathbf{g}_F(t) = [1^F, u_1^F(t), \ldots, u_1^F(t-m), u_2^F(t), \ldots, u_2^F(t-m), \\ -y^F(t-1), \ldots, -y^F(t-n)]^T \tag{43}$$

and

$$\mathbf{p}_F = [r_0^{\bullet}, r_1, \ldots, r_1 b_{1m}, r_2, \ldots, r_2 b_{2m}, a_1, \ldots, a_n]^T = \mathbf{p} . \tag{44}$$

Here

$$1^F = \left(1 + \sum_{i=1}^{s} h_i\right) . \tag{45}$$

(e) Constituting vector $\mathbf{y}_F$ from values $y^F(t)$ and matrix

$$\mathbf{G}_F = \begin{bmatrix} \mathbf{g}_F^T(1) \\ \vdots \\ \mathbf{g}_F^T(N) \end{bmatrix} \tag{46}$$

the GLS estimation of $\mathbf{p}_F$ is:

$$\hat{\mathbf{p}} = \hat{\mathbf{p}}^F = (\mathbf{G}_F^T \mathbf{G}_F)^{-1}\,\mathbf{G}_F^T y_F . \tag{47}$$

The procedure is continued from point (b). (Notice that the second filtering method of CLARKE can also be applied but in general it is of poorer convergence [6].)

5. Maximum Likelihood Method

In the previous methods the noise model is to be assumed a special case of the general one shown in Fig. 3, and reduce the essentially nonlinear estimation problem to a linear one in parameters via "quasi-linearization". Åström (1965) has developed the ML method for linear discrete-time systems for $D(z^{-1}) = A(z^{-1})$ solving the nonlinear estimation problem and has given its computation technique [1]. (Since the system model and the noise model can be reduced to common denominator in every case, the condition $D(z^{-1}) = A(z^{-1})$ is not too severe.) Let us investigate the application of ML method for the Hammerstein model.

Residual $\varepsilon(t)$ is defined by equation

$$C(z^{-1})\varepsilon(t) = A(z^{-1})y(t) - r_0^* - B_1(z^{-1})u(t) - B_2(z^{-1})u^2(t)\,, \tag{48}$$

and it is the estimation of source noise $\lambda e(t)$ shown in Figs 2, 3.

The logarithm of likelihood function becomes

$$L = -\frac{1}{2\lambda^2}\sum_{t=1}^{N}\varepsilon^2(t) - N\log\lambda + \text{const.} \tag{49}$$

Maximizing this function is equivalent to minimizing the loss function:

$$V(p) = \frac{1}{2}\sum_{t=1}^{N}\varepsilon^2(t)\,. \tag{50}$$

The ML estimation of λ can be obtained by $\hat{\mathbf{p}}$ belonging to the minimum of the loss function:

$$\hat{\lambda} = \sqrt{\frac{2}{N}\,V(\hat{\mathbf{p}})}\,. \tag{51}$$

Now $\mathbf{p}$ has the following form:

$$\mathbf{p} = [r_0^*, r_1, \ldots, r_1 b_{1m}, r_2, \ldots, r_2 b_{2m}, a_1, \ldots, a_n, c_1, \ldots, c_k]^T\,. \tag{52}$$

The ML estimate is consistent, asymptotically normal and efficient under mild conditions [1].

In general, combined Gauss—Newton and Newton—Raphson algorithms are used to minimize the loss function:

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - \alpha[V_{pp}(\hat{\mathbf{p}}_k)]^{-1}\,V_p(\hat{\mathbf{p}}_k) \tag{53}$$

where $\mathbf{V}_p$ is the gradient of $V(\mathbf{p})$, and $\mathbf{V}_{pp}$ is the matrix of second order partial derivates. The problem involves the restriction that during the minimum seeking every root of $z^k C(z^{-1})$ must lie inside the unit circle. In the vicinity of illegitimate region the factor α has a role that can be computed according to different strategies [4]. The computation of derivatives was detailed in [3]. In lack of a priori information the seeking starts from a LS estimation. The globality of obtained minimum has to be controlled by restarting the seeking from other initial points.

It is obvious from these relationships that Eq. (48) formally corresponds to a triple input, single output linear discrete-time system, where $u_1(t) = 1$, $u_2(t) = u(t)$ and $u_3(t) = u^2(t)$. (This, of course, can be extended to higher order nonlinearities.) The parameters to be estimated are the coefficients of polynomials $C(z^{-1})$, $A(z^{-1})$, $B_1(z^{-1})$, $B_2(z^{-1})$ and $B_3(z^{-1})$. Here $B_1(z^{-1})$ must be assumed of zero order. Thus, if a program is available for the ML identification of a multiple input, single output system then this program can be easily adapted for identifying the HAMMERSTEIN model, as well.

Both the GLS method and the ML one have been considered for the estimation of the redundant parameter vector of the HAMMERSTEIN model. Namely first it is to be seen whether separability, i.e. condition (10), is approximately valid.

In the positive case the problem will be solved under restriction (10) to ensure the unambiguity. A possible simple solution of this problem arises by replacing Eq. (9) by an equivalent one:

$$y(t) = \mathbf{g}_4^T(t)\mathbf{p}_4 \tag{54}$$

where

$$\mathbf{p}_4 = [r_0^*, r_1, r_1 b_1, \ldots, r_1 b_m, r_2, a_1, \ldots, a_n]^T \tag{55}$$

and

$$\mathbf{g}_4(t) = [1, u(t), u(t-1), \ldots, u(t-m), x(t), -y(t-1), \ldots, -y(t-n)]^T \tag{56}$$

where

$$x(t) = u^2(t) + \sum_{i=1}^{m} b_i u^2(t-m)\,. \tag{57}$$

Since both methods GLS and ML contain iterative procedures, b_i values can always be computed between the iteration cycles on the basis of (10) to constitute $x(t)$. In this way the parameter vector becomes unambiguous and corresponding to the simple separable HAMMERSTEIN model.

6. Simulation Results

The programs of every algorithm mentioned in this paper have been made. Their operations and estimation properties, the necessary computing times have been investigated and compared by several simulation examples. We give the estimated parameter values of the following HAMMERSTEIN model for different polynomials $C(z^{-1})$ and $r_0 = 0$ and $r_0 = 2$; $r_1 = 2$; $B(z^{-1}) = 1+0.5z^{-1}$ and $A(z^{-1}) = D(z^{-1}) = 1-1.5z^{-1}+0.7z^{-2}$. The noise/signal ratio at the output was $\psi = 30$ per cent, a rather high value. The input signal was a random sequence with normal distribution, zero mean and variance one. The mean square error value

$$MSE = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}[y_0(\mathbf{p}) - y_0(\hat{\mathbf{p}})]^2}$$

is also given. Here y_0 is the system output without noise. The estimation of parameters of the HAMMERSTEIN-model containing also a constant term ($r_0 = 2$) — is presented in Fig. 5 for the case of iterative technique (IT) suggested by NARENDRA and GALLMAN and for the case of technique suggested here under restrictions (BIT), where the equation error is white noise, $C(z^{-1}) = 1$, $N = 300$,

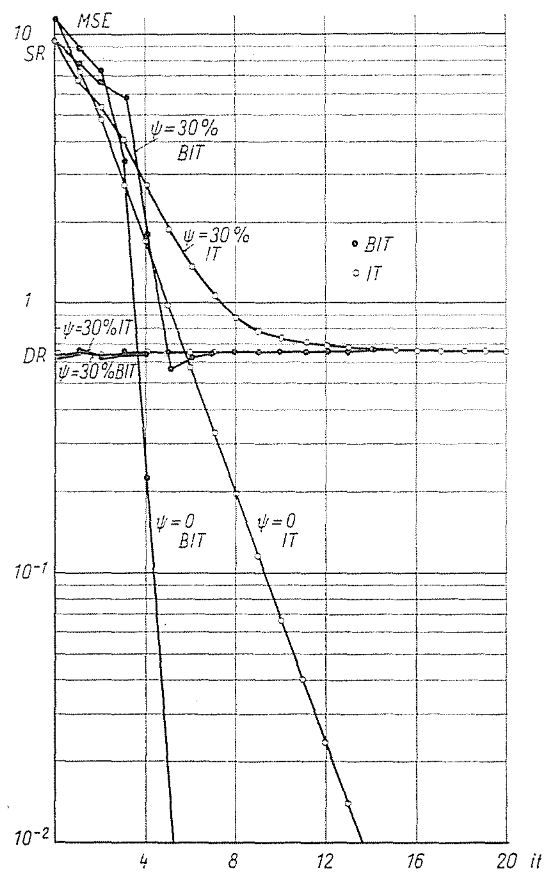

Fig. 5

$\psi = 0$ and $\psi = 30\%$. This latter method is seen to give a higher convergence rate. The iteration started by a static regression (SR) or a direct LS estimation (DR), this latter gives the value of steady state of iteration in one step. The estimates of model with no constant term ($r_0 = 0$) are summarized in Table I under different noise conditions and $N = 500$, $\psi = 30\%$. The following polynomials $C(z^{-1})$ were applied: in case

$$\text{I.} \quad C(z^{-1}) = \frac{1}{1 - 0.6z^{-1}} ;$$

$$\text{II.} \quad C(z^{-1}) = \frac{1}{1 + 0.6z^{-1}} ;$$

$$\text{III.} \quad C(z^{-1}) = 1 - z^{-1} + 0.2z^{-2} ;$$

$$\text{IV.} \quad C(z^{-1}) = 1 - 1.5z^{-1} + 0.7z^{-2} .$$

In the table also the non-iterative (LS), the GLS and the ML methods are compared.

Table 1

	$C(z^{-1})$	λ	r_1	$r_1 b_{11}$	r_2	$r_2 b_{21}$	a_1
			2	1	1	0.5	—1.5
LS			1.952	0.971	1.009	0.499	−1.517
GLS			1.952	0.967	1.010	0.498	−1.518
GLS	(I)	0.5	1.971	1.025	1.002	0.509	−1.500
ML			1.961	1.009	1.002	0.509	−1.509
ML			1.973	1.025	1.005	0.511	−1.504
LS			1.903	1.346	0.955	0.635	−1.411
GLS			1.926	1.135	0.995	0.553	−1.484
GLS	(II)	1.5	1.926	1.088	1.003	0.532	−1.496
ML			1.895	1.117	0.987	0.547	−1.493
ML			1.916	1.115	0.995	0.546	−1.492
LS			1.829	2.110	0.943	0.898	−1.142
GLS	(III)	3.0	1.837	1.470	0.978	0.659	−1.433
ML			1.846	1.246	1.018	0.538	−1.481
LS			1.851	2.639	0.917	1.109	−0.923
GLS	(IV)	3.0	1.974	1.422	1.003	0.657	−1.409
ML			1.879	1.196	0.979	0.565	−1.488

7. Conclusions

The identification technique of a special nonlinear system, the HAMMERSTEIN model, has been shown to closely approach the methods elaborated for linear systems. The published estimation procedures have been reviewed, the relationships pointed out and suggested to extend the GLS and ML methods for the nonlinear case, presenting the formula necessary for the application. The usefulness of the suggested methods is supported by simulation results. Further extensive investigations are needed to study in practical cases whether the HAMMERSTEIN model suit to describe the real nonlinear dynamic systems.

The most extensive application of the suggested methods seems to be — by the moment — for investigating a process wether it can be described by a linear discrete-time model in the vicinity of the working point, in the range of an unchanged input signal or not. In this latter case, — hence for a strong nonlinear character — the linear approximation is only allowed for a lesser changing range of the input signal.

Methods applied for identifying the HAMMERSTEIN model can be used at the adaptive extremum control, as well. Remind that the GLS method suits on-line estimation, while the adaptive system model necessary for the dual control, can be produced by the HAMMERSTEIN model.

a_2 / 0.7	h_1 / c_1	h_2 / c_2	b_{11}	b_{21}	$\hat{\lambda}$	MSE
0.713	—	—	0.497	0.494	—	0.472
0.714	0.034	—	0.496	0.493	—	0.480
0.700	−0.526	−0.086	0.520	0.508	—	0.157
0.705	0.502	—	0.515	0.508	0.534	0.391
0.701	0.568	0.236	0.520	0.508	0.519	0.355
0.618	—	—	0.707	0.664	—	1.459
0.687	0.488	—	0.589	0.555	—	0.406
0.697	0.633	0.115	0.565	0.531	—	0.384
0.694	−0.492	—	0.589	0.554	1.607	0.427
0.694	−0.579	0.239	0.577	0.548	1.511	0.377
0.380	—	—	1.153	0.952	—	3.951
0.649	0.774	0.363	0.800	0.674	—	1.167
0.686	−0.994	0.198	0.676	0.528	3.066	1.029
0.175	—	—	1.426	1.209	—	5.109
0.625	0.997	0.498	0.720	0.655	—	1.368
0.694	−1.519	0.747	0.637	0.577	3.067	0.467

Summary

The possibility of parameter estimation of the HAMMERSTEIN model is investigated in detail in case of quadratic polynomial form and different noise situations. Over and above the well-known and new iterative and noniterative methods, the extensions for nonlinear case of generalised least-squares and maximum likelihood method — used in the linear systems — are presented. These methods are supported by simulation results.

References

1. ASTRÖM, K. J.—BOHLIN, T.: Numerical Identification of Linear Dynamic Systems from Normal Operating Records, IFAC, Teddington, 1965.
2. ASTRÖM, K. J.—EYKHOFF, P.: System Identification — A survey, IFAC, Prague, 1970.
3. BÁNYÁSZ, Cs.—HABER, R.—KEVICZKY, L.: Some Estimation Methods for Nonlinear Discrete Time Identification, IFAC, Hague, 1973.
4. BÁNYÁSZ, Cs.: Investigation of Discrete Identification Methods by Digital Computer, Ph. D. Thesis 1971 Automation Research Institute of the Hungarian Academy of Sciences. 1972/5.
5. CHANG, F. H. I.—LUUS, R.: A Noniterative Method for Identification Using HAMMERSTEIN Model, IEEE Trans. Aut. Control 1971. AC-16.
6. CLARKE, D. W.: Generalized-Least-Squares Estimation of the Parameters of a Dynamic Model, IFAC, Prague, 1967.
7. HABER, R.: Identification of Nonlinear Discrete Processes. Master Thesis 1972 Technical University of Budapest, Automatic Control Department.
8. KEVICZKY, L.—BÁNYÁSZ, Cs.: Identification of Linear Discrete Time Systems. Electrotechnica, Hungary. (To be published.)
9. NARENDRA, K. S.—GALLMANN, P. G.: An Iterative Method for the Identification of Nonlinear Systems Using a HAMMERSTEIN Model. IEEE Trans. Aut. Contr. 1966. AC-11.

Dr. László KEVICZKY
Róbert HABER } H-1521, Budapest,