

A SIMPLE METHOD FOR THE STEADY-STATE IDENTIFICATION OF 2nd ORDER NONLINEARITIES

By

L. KEVICZKY and Cs. BÁNYÁSZ*

Department of Automation, Technical University, Budapest

(Received September 22, 1972)

Presented by Prof. Dr. F. CsÁKI

1. Introduction

The complete quadratic form is used with a special linking for the description of steady-state operation of nonlinear processes in the case of large-scale changing of signals. As far as this description is valid in the environment of a given working point, a model of parameters of minimum number can be obtained, helping to estimate the place of extremum, and the arrangement of response surface as well by means of major axes of quadratic form.

Let us consider the system model of n variables shown in Fig. 1 where z , u , x are the $(n \times 1)$ vectors of input variables, of the values in the working

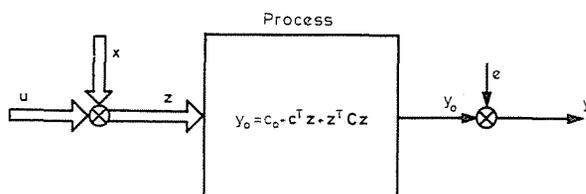


Fig. 1

point and of the disturbing signals at the input, respectively. y_0 is the theoretical value of output, e is the noise at the output and y is the measurable output signal.

Let a complete quadratic form describe the static characteristic of such a process in the following way:

$$y_0(z) = c_0 + c^T z + z^T C z \quad (1)$$

where c_0 , $(n \times 1)$ vector c and $(n \times n)$ matrix C contain the coefficients of static characteristic. (Here T denotes transposition.)

* Automation Research Institute of the Hungarian Academy of Sciences.

Relationship (1) can be written as a second-order Taylor-series, being valid in the environment of the working point $\mathbf{z} = \mathbf{u}$ of the surface, that is

$$y_0(\mathbf{z}) = y_0(\mathbf{x}) = y_0(\mathbf{u}) + \nabla^T y_0(\mathbf{u}) \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} \quad (2)$$

where

$$\mathbf{z} = \mathbf{u} + \mathbf{x} \quad (3)$$

on the basis of Fig. 1; in addition

$$\nabla y_0(\mathbf{u}) = 2\mathbf{C}\mathbf{u} + \mathbf{c} \quad (4)$$

is the gradient vector in the working point and

$$\mathbf{H} = 2\mathbf{C} \quad (5)$$

is the Hessian matrix of second order derivatives.

Exceeding (3) it is assumed that in this measuring situation

$$y = y_0 + e. \quad (6)$$

In connection with input and output disturbance it is assumed that

$$M\{\mathbf{x}\} = \mathbf{0} \quad \text{and} \quad M\{e\} = 0 \quad (7)$$

i.e. they have zero expected values, moreover

$$M\{\mathbf{x}\mathbf{x}^T\} = \mathbf{D} = \langle \sigma_1^2, \dots, \sigma_n^2 \rangle; \quad M\{e\mathbf{x}\} = 0 \quad (8)$$

that is, the input noises are uncorrelated with the output noise and with each other as well ($\langle \dots \rangle$ denotes diagonal matrix). The input noises are also assumed to be symmetrically distributed.

2. Estimations for the determination of coefficients of steady-state model

Let us consider some statistics by means of which an estimation can be given for the derivatives of process at the working point and for the various parameters of static characteristic, respectively.

On the basis of Appendices 1 through 4, evaluating $b_0, \mathbf{b}, \mathbf{d}, \mathbf{B}$, the

various functions of parameters can be calculated, namely:

$$M\{y\} = b_0 = y_0(\mathbf{u}) + \frac{1}{2} \text{tr}(\mathbf{HD}) = y_0(\mathbf{u}) + \frac{1}{2} \sum_{i=1}^n \sigma_i^2 \frac{\partial^2 y_0(\mathbf{u})}{\partial x_i^2} \quad (9)$$

$$M\{xy\} = \mathbf{b} - \mathbf{D}\nabla y_0(\mathbf{u}) \quad (10)$$

$$M\{zy\} = \mathbf{d} = \mathbf{b} + \mathbf{u}b_0 \quad (11)$$

$$M\{xy\mathbf{x}^T\} = \mathbf{B} = \mathbf{D}y_0(\mathbf{u}) + \frac{1}{2} M\{\mathbf{xx}^T \mathbf{H} \mathbf{xx}^T\}. \quad (12)$$

The introduced functions $\text{sgn}(x_i)$ and $k(x_i)$ are seen in Fig. 2 where h_i is obtained from the equality of probabilities $P[k(x_i) > 0] = P[k(x_i) < 0]$ and introducing notations

$$\text{sgn}(\mathbf{x}) = [\text{sgn}(x_1), \dots, \text{sgn}(x_n)]^T \quad (13)$$

$$\mathbf{k}(\mathbf{x}) = [k(x_1), \dots, k(x_n)]^T \quad (14)$$

further estimation possibilities are shown in Appendices 5 through 7 by determination of values \mathbf{a} , \mathbf{A} and \mathbf{q} .

$$M\{\text{sgn}(\mathbf{x})y\} = \mathbf{a} = \mathbf{S}\nabla y_0(\mathbf{u}) \quad (15)$$

$$M\{\text{sgn}(\mathbf{x})y \text{sgn}^T(\mathbf{x})\} = \mathbf{A} = y_0(\mathbf{u}) \mathbf{E} + \frac{1}{2} M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{H} \mathbf{x} \text{sgn}^T(\mathbf{x})\} \quad (16)$$

$$M\{\mathbf{k}(\mathbf{x})y\} = \mathbf{q} = \frac{1}{2} \mathbf{Q}\mathbf{p}. \quad (17)$$

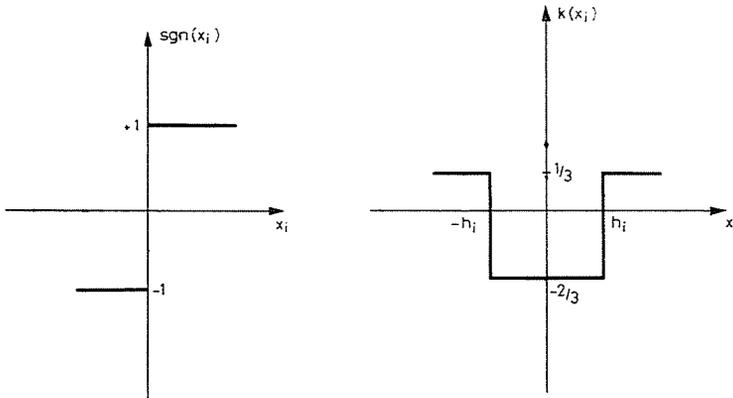


Fig. 2

By means of these estimation methods, elements of an $(n \times n)$ matrix $\mathbf{G} = [g_{ij}]$ and an $(n \times 1)$ vector $\mathbf{g} = [g_i]$ can be calculated so that the following identities hold:

$$\frac{\partial^2 y_0(\mathbf{u})}{\partial x_i^2} = g_{ii} \frac{1}{\lambda_{ii}} \quad (18)$$

$$\frac{\partial^2 y_0(\mathbf{u})}{\partial x_i \partial x_j} = g_{ij} \frac{1}{\lambda_i \lambda_j} \quad (19)$$

$$\frac{\partial y_0(\mathbf{u})}{\partial x_i} = \frac{1}{\lambda_i} g_i \quad (20)$$

Be then

$$\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2 \quad (21)$$

where

$$\mathbf{G}_1 = \begin{bmatrix} 0 & g_{12} & \cdots & g_{n1} \\ g_{21} & 0 & \cdots & g_{r,2} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ g_{n1} & g_{n2} & \cdots & 0 \end{bmatrix} \quad (22)$$

and

$$\mathbf{G}_2 = \langle g_{11}, \dots, g_{nn} \rangle \quad (23)$$

Let us introduce

$$\mathbf{L}_1 = \left\langle \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n} \right\rangle \quad (24)$$

and

$$\mathbf{L}_2 = \left\langle \frac{\lambda_1^2}{\lambda_{11}}, \dots, \frac{\lambda_n^2}{\lambda_{nn}} \right\rangle \quad (25)$$

Then

$$\nabla y_0(\mathbf{u}) = \mathbf{L}_1 \mathbf{g} \quad (26)$$

as well as

$$\mathbf{H} = \mathbf{L}_1 \mathbf{G}_2 \mathbf{L}_2 \mathbf{L}_1^T + \mathbf{L}_1 \mathbf{G}_1 \mathbf{L}_1^T = \mathbf{L}_1 (\mathbf{G}_1 + \mathbf{G}_2 \mathbf{L}_2) \mathbf{L}_1 \quad (27)$$

Then the extremum of quadratic form (as it is known) can be expressed in the following way:

$$\mathbf{u}^* = \mathbf{H}^{-1} \nabla y_0(\mathbf{u}) = \mathbf{L}_1^{-1} (\mathbf{G}_1 + \mathbf{G}_2 \mathbf{L}_2)^{-1} \mathbf{b} \quad (28)$$

Checking Eqs (4), (5) and (26), (27) the coefficients of quadratic form can be obtained:

$$\mathbf{C} = \frac{1}{2} \mathbf{L}_1 (\mathbf{G}_1 + \mathbf{G}_2 \mathbf{L}_2) \mathbf{L}_1 \quad (29)$$

$$\mathbf{c} = \mathbf{L}_1 [\mathbf{g} - (\mathbf{G}_1 + \mathbf{G}_2 \mathbf{L}_2) \mathbf{L}_1 \mathbf{u}] \quad (30)$$

and finally

$$c_0 = y_0(\mathbf{u}) - \mathbf{g}^T \mathbf{L}_1 \mathbf{u} + \frac{1}{2} \mathbf{u}^T \mathbf{L}_1 (\mathbf{G}_1 + \mathbf{G}_2 \mathbf{L}_2) \mathbf{L}_1 \mathbf{u}. \quad (31)$$

Parameters are rather easy to evaluate, mainly restricted to the constitution of expected value and then to correct with a scale-factor, as it is shown above.

3. Adaptive algorithms for the determination of the coefficients

In estimations presented in the previous chapter, the parameters were determined by statistical averaging. The sequential performance of averaging corresponds to an adaptive stochastic approximation algorithm which is optimal in the case of weighting function $\gamma[k] = 1/k$, too, and provides minimum variance estimation.

Let us consider the determination of process parameters by relationships (15), (16), (17) as an example. According to these

$$\nabla y_0(\mathbf{u}) = \mathbf{S}^{-1} \mathbf{a} = \mathbf{L}_1 \mathbf{g} = \mathbf{L}_1 M\{\text{sgn}(\mathbf{x}) y\} \quad (32)$$

and

$$\mathbf{S} = \langle M\{|x_1|\}, \dots, M\{|x_n|\} \rangle = \mathbf{L}_1^{-1}. \quad (33)$$

In addition,

$$\mathbf{G}_1 = \langle M\{k_1(x_1) y\}, \dots, M\{k_n(x_n) y\} \rangle \quad (34)$$

$$\mathbf{L}_2 = 2 \left\langle \frac{M^2\{|x_2|\}}{M\{k_1(x_1) x_1^2\}}, \dots, \frac{M^2\{|x_n|\}}{M\{k_n(x_n) x_n^2\}} \right\rangle \quad (35)$$

as well as for the elements of \mathbf{G}_2 :

$$g_{ij} = M\{\text{sgn}(x_i) \text{sgn}(x_j) y\} \quad i \neq j. \quad (36)$$

So \mathbf{H} and $y_0(\mathbf{u})$ are formed according to (27) and (9), respectively. The coefficients of quadratic form are given by (29), (30), (31).

Two different situations can be distinguished at the constitution of adaptive algorithms. In the first one \mathbf{x} — as an external testing signal — is given to the system so its every property is known in advance just as are the scale factors \mathbf{L}_1 , \mathbf{L}_2 and \mathbf{D} .

Applying the adaptive averaging for the determination of a the following algorithm is obtained [1]:

$$\mathbf{a}[k] = \mathbf{a}[k-1] + \frac{1}{k} (\text{sgn}(\mathbf{x}[k]) y[k] - \mathbf{a}[k-1]). \quad (37)$$

Here k means the k -th step of the sequential estimation i.e. the k -th time during the on-line data-processing. Similar algorithms give solutions for G_1 and G_2 . Flow chart of adaptive estimation based on this principle is shown in Fig. 3 for the case where the input noise is generated by us (with known features). The notations in the figure are the same as in [1, 2].

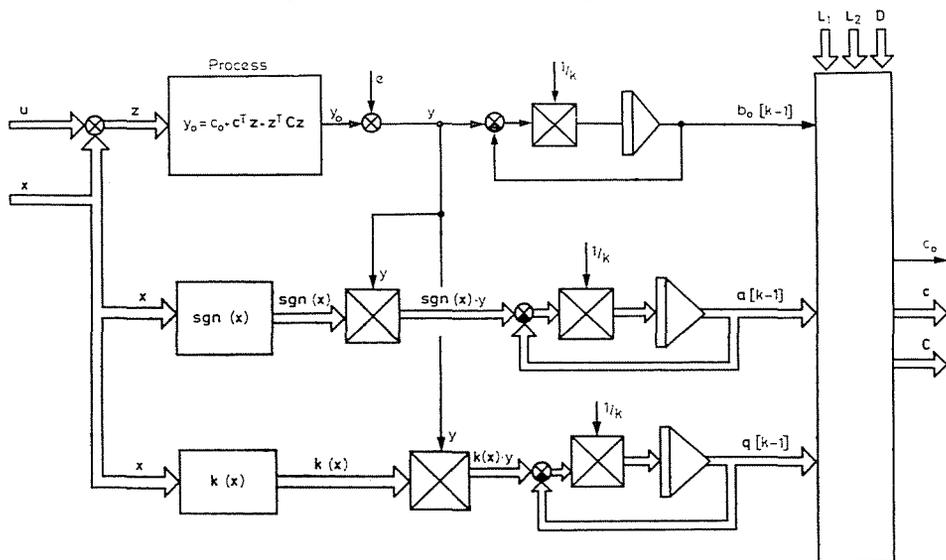


Fig. 3

The flow chart of adaptive estimation can also be established when the features of x are not known, i.e. x is not generated by us. In this situation additional adaptive cycles must be included to the procedure for estimating the values in L_1, L_2, D . In these cases the convergence slows down significantly as the parameters of transformations $\text{sgn}(x)$ and $k(x)$ depend on x . In this situation z can only be measured and the suitable information must be separated from it.

Results of simulation

A program simulating the measuring situation was run on a digital computer for studying this method. The process was assumed in form (1) and its parameters were:

$$c_0 = 10; \quad c^T = [10, -8]; \quad C = \begin{bmatrix} 6 & 2,5 \\ 2,5 & -4 \end{bmatrix}$$

The investigation was performed by the method presented in chapter 3, for the working point $u = 0$ under input variances $\sigma_1^2 = \sigma_2^2 = 1$ in case of noises

of normal distribution. Results of on-line estimation are seen in Fig. 4. In evaluating the coefficient values it must be taken into account that in case of normal distribution

$$\frac{1}{2} M\{k_i(x_i) x_i^2\} = 0,157 \sigma_i^2 \quad \text{and} \quad M\{|x_i|\} = 0,798 \sigma$$

An optimal control performed by the same algorithm is demonstrated in Fig. 5 for a positive definite form of extremum $u^* = [-2, 4]^T$ by means of algorithm (28).

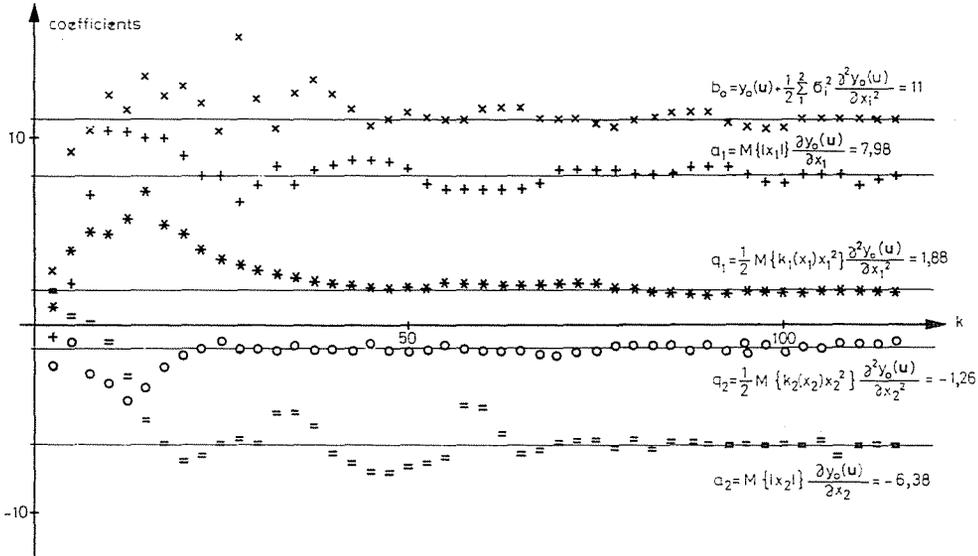


Fig. 4

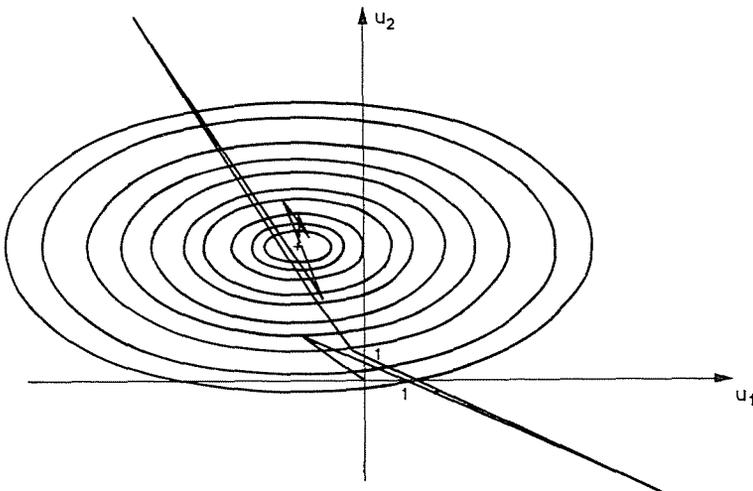


Fig. 5

5. Conclusions

In our paper some new estimation methods were investigated, suggested for the determination of parameters of system description with quadratic polynomial used in general in steady-state operation of large-scale changing of signals. These methods simplify dataprocessing under above detailed conditions of input noises. As conditions for the input noises are very severe, these methods, of course, have nothing of common with the usual method of least squares. Their simplicity makes them, however, suitable for identifying hardware instruments. Such an instrument can be of a great help for constructing steady-state models of high accuracy as a preliminary structure estimator.

There is nothing of vitally new in the presented methods. We can get through to these considerations in other ways, too, for example by generalizing the statistic linearizing methods, by making the designed experiments continuous or by extending the idea of synchronous-detection.

As it is shown, optimal control can also be realized by the presented methods. The construction of external searching signal of extremum-seeking controls can also be discussed from the aspect of specifications requiring the input noises to include suitable deterministic signals.

Appendix

1.

$$b_0 = M\{y\} = M\{y_0 + e\} = M\{y_0\} + M\{e\} = M\{y_0(\mathbf{u}) + \nabla^T y_0(\mathbf{u}) \mathbf{x} + 1/2 \mathbf{x}^T \mathbf{H} \mathbf{x}\} = \\ = M\{y_0(\mathbf{u})\} + \nabla^T y_0(\mathbf{u}) M\{\mathbf{x}\} + 1/2 M\{\mathbf{x}^T \mathbf{H} \mathbf{x}\} = y_0(\mathbf{u}) + 1/2 \text{tr}(\mathbf{H} \mathbf{D}).$$

Namely, the mixed moments in $M\{\mathbf{x}^T \mathbf{H} \mathbf{x}\}$ equal zero, because they are uncorrelated and (7), (8) were used. Here tr denotes trace of matrix.

2.

$$\mathbf{b} = M\{\mathbf{xy}\} = M\{\mathbf{xy}_0 + \mathbf{xe}\} = M\{\mathbf{xy}_0\} + M\{\mathbf{xe}\} = M\{\mathbf{xy}_0\} = \\ = M\{\mathbf{xy}_0(\mathbf{u}) + \mathbf{xx}^T \nabla y_0(\mathbf{u}) + 1/2 \mathbf{xx}^T \mathbf{H} \mathbf{x}\} = \\ = y_0(\mathbf{u}) M\{\mathbf{x}\} + M\{\mathbf{xx}^T\} \nabla y_0(\mathbf{u}) + 1/2 M\{\mathbf{xx}^T \mathbf{H} \mathbf{x}\}.$$

Namely, the noises are independent and every third order moment equals zero in consequence of symmetric distribution, so

$$\mathbf{b} = \mathbf{D} \nabla y_0(\mathbf{u}).$$

3.

$$\mathbf{d} = M\{\mathbf{zy}\} = M\{\mathbf{zy}_0 + \mathbf{ze}\} = M\{(\mathbf{x} + \mathbf{u})y_0\} + M\{(\mathbf{x} + \mathbf{u})e\} = \\ = M\{\mathbf{xy}_0\} + M\{\mathbf{uy}_0\} + M\{\mathbf{xe}\} + M\{\mathbf{ue}\} = \mathbf{b} + \mathbf{u} M\{y_0\} + \mathbf{u} M\{e\} = \mathbf{b} + \mathbf{u} b_0 = \\ = \mathbf{D} \nabla y_0(\mathbf{u}) + \mathbf{u} [y_0(\mathbf{u}) + 1/2 \text{tr}(\mathbf{H} \mathbf{D})].$$

4.

$$\mathbf{B} = M\{\mathbf{xyx}^T\} = M\{\mathbf{xy}_0 \mathbf{x}^T\} + M\{\mathbf{xe} \mathbf{x}^T\} = M\{\mathbf{xy}_0 \mathbf{x}^T\}.$$

As the noises are uncorrelated, so $M\{\mathbf{xe} \mathbf{x}^T\} = 0$

$$\mathbf{B} = M\{\mathbf{xy}_0(\mathbf{u}) \mathbf{x}^T\} + M\{\mathbf{xx}^T \nabla y_0(\mathbf{u}) \mathbf{x}^T\} + 1/2 M\{\mathbf{xx}^T \mathbf{H} \mathbf{xx}^T\} = \\ = M\{\mathbf{xx}^T\} y_0(\mathbf{u}) + 1/2 M\{\mathbf{xx}^T \mathbf{H} \mathbf{xx}^T\}.$$

Since every third-order moment equals zero because of symmetrical distribution, hence $M\{\mathbf{xx}^T \nabla y_0(\mathbf{u}) \mathbf{x}^T\} = 0$.

In addition it can be written because of above reasons that

$$\frac{1}{2} M\{\mathbf{xx}^T \mathbf{Hxx}^T\} = \left[(1 + \delta_{ij})^{-1} M\{x_i^2 x_j^2\} \frac{\partial^2 y_0(\mathbf{u})}{\partial x_i \partial x_j} \right]$$

where δ_{ij} is the Cronecker-symbol, x_i or x_j is the corresponding element of \mathbf{x} . Then

$$\mathbf{B} = \mathbf{D}y_0(\mathbf{u}) + 1/2 M\{\mathbf{xx}^T \mathbf{Hxx}^T\}.$$

5.

$$\begin{aligned} \mathbf{a} &= M\{\text{sgn}(\mathbf{x}) y\} = M\{\text{sgn}(\mathbf{x}) y_0\} + M\{\text{sgn}(\mathbf{x}) e\} = \\ &= M\{\text{sgn}(\mathbf{x}) y_0\} = M\{\text{sgn}(\mathbf{x}) y(\mathbf{u})\} + M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \nabla y_0(\mathbf{u})\} + 1/2 M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx}\}. \end{aligned}$$

In this expression $M\{\text{sgn}(\mathbf{x})\} = 0$ for $M\{\mathbf{x}\} = 0$ and it has symmetrical distribution and $M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx}\} = 0$ because the input noises are uncorrelated. Be

$$\mathbf{S} = M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T\} = \langle M\{|x_1|\}, \dots, M\{|x_n|\} \rangle$$

and so

$$\mathbf{a} = \mathbf{S} \nabla y_0(\mathbf{u}).$$

6.

$$\begin{aligned} \mathbf{A} &= M\{\text{sgn}(\mathbf{x}) y \text{sgn}^T(\mathbf{x})\} = M\{\text{sgn}(\mathbf{x}) y_0 \text{sgn}^T(\mathbf{x})\} + M\{\text{sgn}(\mathbf{x}) e \text{sgn}^T(\mathbf{x})\} = \\ &= M\{\text{sgn}(\mathbf{x}) y_0(\mathbf{u}) \text{sgn}^T(\mathbf{x})\} + M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \nabla y_0(\mathbf{u}) \text{sgn}^T(\mathbf{x})\} + \\ &+ 1/2 M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx} \text{sgn}^T(\mathbf{x})\} = y_0(\mathbf{u}) \mathbf{E} + 1/2 M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx} \text{sgn}^T(\mathbf{x})\}. \end{aligned}$$

Here \mathbf{E} denotes unity matrix and

$$\frac{1}{2} M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx} \text{sgn}^T(\mathbf{x})\} = \left[(1 + \delta_{ij})^{-1} M\{|x_i| |x_j|\} \frac{\partial^2 y_0(\mathbf{u})}{\partial x_i \partial x_j} \right].$$

Hence

$$\mathbf{A} = y_0(\mathbf{u}) \mathbf{E} + 1/2 M\{\text{sgn}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx} \text{sgn}^T(\mathbf{x})\}.$$

7.

$$\begin{aligned} \mathbf{q} &= M\{\mathbf{k}(\mathbf{x}) y\} = M\{\mathbf{k}(\mathbf{x}) y_0\} + M\{\mathbf{k}(\mathbf{x}) e\} = M\{\mathbf{k}(\mathbf{x}) y_0\} = \\ &= M\{\mathbf{k}(\mathbf{x}) y_0(\mathbf{u})\} + M\{\mathbf{k}(\mathbf{x}) \mathbf{x}^T \nabla y_0(\mathbf{u})\} + 1/2 M\{\mathbf{k}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx}\}. \end{aligned}$$

As \mathbf{x} has symmetrical distribution and $P[\mathbf{k}(\mathbf{x}) > 0] = P[\mathbf{k}(\mathbf{x}) < 0]$ by definition, so

$$M\{\mathbf{k}(\mathbf{x})\} = 0 \text{ and } M\{\mathbf{k}(\mathbf{x}) \mathbf{x}^T\} = 0.$$

Finally,

$$\mathbf{q} = 1/2 M\{\mathbf{k}(\mathbf{x}) \mathbf{x}^T \mathbf{Hx}\} - 1/2 \mathbf{Q} \mathbf{p}$$

where

$$\mathbf{p} = [M\{k_1(x_1) x_1^2\}, \dots, M\{k_n(x_n) x_n^2\}]^T$$

and

$$\mathbf{Q} = \left\langle \frac{\partial^2 y_0(\mathbf{u})}{\partial x_1^2}, \dots, \frac{\partial^2 y_0(\mathbf{u})}{\partial x_n^2} \right\rangle.$$

Summary

In our paper a simple method for the steady-state identification of 2nd order nonlinear systems is considered. The algorithm processes the data on-line and the coefficients of static characteristic can be obtained simply by this algorithm in the case of uncorrelated input signals. The model obtained in such way can be used for system optimization. On the basis of this method an equipment can be made, facilitating appropriate structure estimation of steady-state characteristic.

References

1. Цыпкин, Я. З.: Адаптация и обучение в автоматических системах. Изд. «Наука», 1968, Москва.
2. Цыпкин, Я. З.: Основы теории обучающихся систем. Изд. «Наука», 1970, Москва.
3. Мойсюк, Б. Н.: Упрощенный метод математической обработки данных, полученных пассивным экспериментом. «Планирование Эксперимента», Изд. «Наука», 1966 Москва.
4. ALBERT, A.—GARDNER, L. A.: Stochastic Approximation and Nonlinear Regression. MIT Press, Cambridge, 1967
5. CSÁKI, F.: Modern Control Theories. Nonlinear, Optimal and Adaptive Systems. Publishing House of the Hungarian Academy of Sciences, Budapest, 1972

Dr. László KEVICZKY, }
Dr. Csilla BÁNYÁSZ, } 1502 Budapest P.O.B. 91. Hungary