METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.pg.gda.pl

LEVENBERG-MARQUARDT ALGORITHM FOR OPTIMIZATION OF MATHEMATICAL MODELS ACCORDING TO *MINIMAX* OBJECTIVE FUNCTION OF MEASUREMENT SYSTEMS

Krzysztof Tomczyk

Cracow University of Technology, Faculty of Electrical and Computer Engineering, Warszawska 24, 31-155 Kraków, Poland
(✉ petomczy@cyf-kr.edu.pl, +48 12 628 2530*)*

Abstract

The paper presents an application of Levenberg-Marquardt algorithm to parametric optimization of the *minimax* type of measurement systems. For the assumed objective function given by the integral square error, optimization of the third-order model is carried out and explained in detail. The optimization procedure is realized in three stages. The optimization method presented in the paper can find broad application in the process of determining optimum models of systems, especially for those that operate in dynamic states.

Keywords: *minimax* optimization, Levenberg-Marquardt algorithm, dynamic error.

1. Introduction

The dynamic development of advanced numerical techniques observed in recent years together with constantly increasing computational efficiency of computers have a great influence on the development of various methods of mathematical modeling of measurement systems [2, 8]. The aim of such modeling is to synthesize models that map precisely the dynamic properties of real systems. However, it is not possible to map precisely the properties of a real system by its model. Application of higher-order models usually gives better mapping, but on the other hand, the analysis of dynamic properties of such models is most often difficult and time-consuming. Hence a tendency has developed towards replacing higher-order models with simplified models [5]. The class and order of such models can be determined by means of approximated knowledge of properties and characteristics of the modeling system, which often results from the modeler's experience, whereas their parameters are determined by means of methods that minimize the mapping error, taking into account all possible input signals. As it is impossible to analyze the full set of all imaginable input dynamic signals, it is suggested to solve this problem by using one signal which maximizes the assumed objective function [2, 3, 7]. In this way, the mapping error being determined is credible for any input. The paper presents such a *minimax* procedure by means of the Levenberg-Marquardt optimization algorithm.

2. *Minimax* optimization of integral square error

Minimax optimization of integral square error includes three main numerical computation stages. At the first stage the standard of the mathematical model is determined. Its order is equal to the order of the modeled system. At the second stage, from among all possible inputs $u(t)$, signal $u_0(t)$ maximizing the integral square error $I_2(u_0)$ (1) is determined - Fig. 1 [2, 3, 7].

$$I_2(u_0) = \max \int_0^T y^2(t)dt. \tag{1}$$

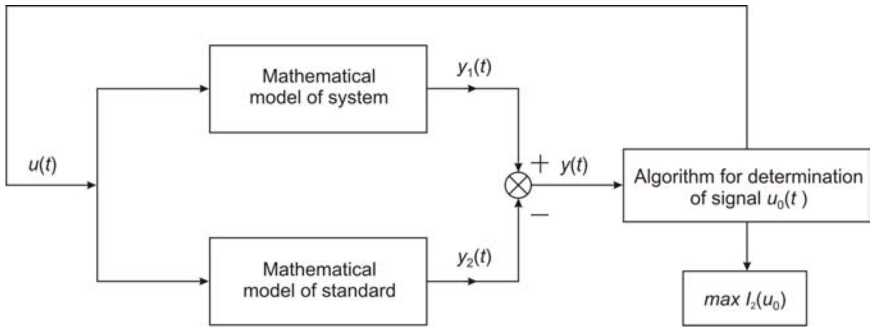

Fig. 1. Block diagram of system for determination of signal $u_0(t)$ maximizing the integral square error.

The output of the system is defined by:

$$y_1(t) = \int_0^t k_1(t-\tau)u(\tau)d\tau \tag{2}$$

and the output of standard:

$$y_2(t) = \int_0^t k_2(t-\tau)u(\tau)d\tau. \tag{3}$$

The error equals:

$$y(t) = y_1(t) - y_2(t) = \int_0^t k(t-\tau)u(\tau)d\tau, \tag{4}$$

where:

$$k(t) = k_1(t) - k_2(t). \tag{5}$$

$k_1(t)$, $k_2(t)$ – impulse responses of system and standard respectively.

At the third stage the class and order of the simplified model are assumed. For this model, by means of the signal $u_0(t)$, parameters minimizing the integral square error are determined. Fig. 2 presents the block diagram of *minimax* optimization by means of the Levenberg-Marquardt algorithm.

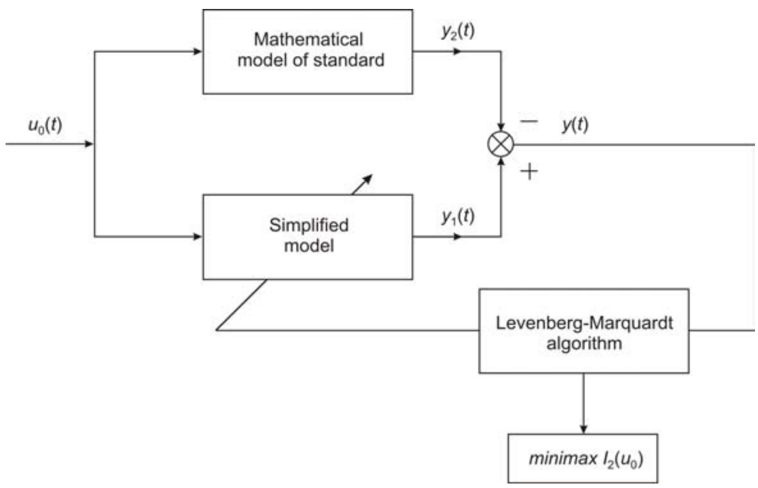

Fig. 2. Block diagram of *minimax* optimization of measurement systems by means of the Levenberg-Marquardt algorithm.

3. Development of standard model

In the presented *minimax* method, the class and order of the standard model corresponds with the class and order of the real system, whereas the selection of the standard parameters can be made by optimization methods [6]. Non-deformation transformation is assumed in this case. The problem of optimization of the standards parameters is then reduced to determination of such values which ensure that their characteristics are close to the characteristics of the non-deformation system. In the broadest range of frequencies their magnitude characteristics should be flat, while phase characteristics should be linear. For ideal non-deformation systems the relation between input y(t) vs. output u(t) is:

$$y(t) = a \cdot u(t), \tag{6}$$

In practice, it is impossible to realize (6). Therefore the relation (7) is adopted more frequently:

$$y(t) = a \cdot u(t - \tau), \tag{7}$$

where τ - time delay, a - amplification coefficient.

For standard model development an upper boundary of frequency ω_{gs} is assumed. The ranges of constancy ω_A of the magnitude-frequency characteristic and linearity of the phase-frequency characteristic ω_φ are most often determined by means of a particular frequency for which these characteristics do not differ more than Δ_A and Δ_φ – Figs 3, 4.

The selection of standard model parameters should ensure maximum ranges of both constancy ω_A and linearity of ω_φ. This range is defined by coefficient η as follows:

$$\eta = \omega_A + \omega_\phi - \left|\omega_A - \omega_\phi\right|. \tag{8}$$

The solution for which coefficient η reaches a maximum value is recognized as the optimal one [6].

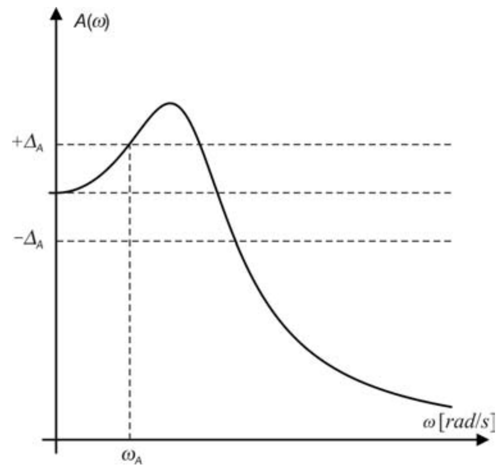

Fig. 3. The range of constancy of magnitude-frequency characteristic.

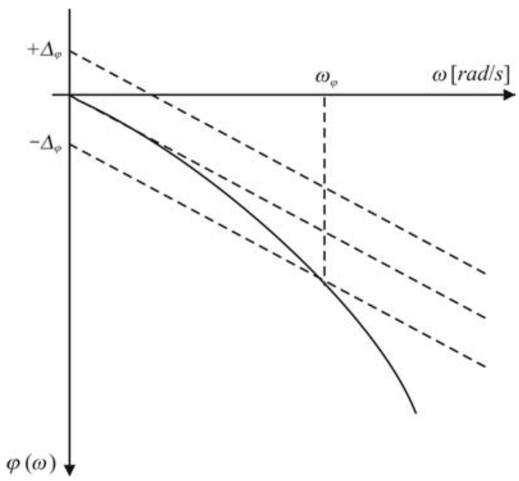

Fig. 4. The range of linearity of phase-frequency characteristic.

4. Procedure for determining the signal maximizing the integral square error

In relevant literature it has been proved that an input signal with imposed constraint on magnitude A and constraint on rate of change ϑ maximizes the integral square-error criterion only when it reaches the imposed constraints over [0, T]. If the only constraint imposed on a signal is the magnitude constraint, then it is always of the "bang-bang" type. In the case when there are two simultaneous constraints, this signal can be only in the form of triangles or of trapezoids with the slopes inclination resulting from ϑ and A [2, 3].

For the integral square-error and for a signal with two constraints A and ϑ imposed simultaneously, the analytical solution describing its shape has not been worked out so far. Additionally, the function space of a possible solution is the space of infinite power of a set and of infinite dimension [4].

A good result can be obtained if the evolutional algorithm is applied [7]. It assures that the solutions are correct and received in a very short calculation time. In the evolutional algorithm we can assume that the switching moments for $t_i < t_{i-1}$ or $t_i > t_{i+1}$, i=1, 2, …, n-1 cannot appear in the generated sequence of signals $u(t)$.

5. Application of the Levenberg-Marquardt algorithm for *minimax* optimisation

The Levenberg-Marquardt algorithm combines the steepest descent method with the Gauss-Newton method and operates correctly in search for parameters both far from and close to the optimum one. In the former case the algorithm of the linear model of steepest descent is used, and in the latter one - squared convergence. Fast convergence is an additional advantage of the algorithm [1].

The Levenberg-Marquardt algorithm is an iterative method, in which the vector of unknown parameters is determined during step $k+1$ by the equation:

$$\mathbf{a}_{k+1} = \mathbf{a}_k^{\,T} - \left[\mathbf{J}^T(\mathbf{a}_k,t)\mathbf{J}(\mathbf{a}_k,t) + \mu_k I\right]^{-1} \mathbf{J}^T(\mathbf{a}_k,t)\,y(\mathbf{a}_k,t) \tag{9}$$

with the error:

$$I_2 = \int_0^T y^2(\mathbf{a}_k,t)\,dt, \tag{10}$$

where:

$$y(\mathbf{a}_k,t)=\int_0^t k(t-\tau)u(\tau)d\tau, \tag{11}$$

$$\mathbf{J}(\mathbf{a}_k,t)=\begin{bmatrix} \frac{\partial y(\mathbf{a}_k,t_1)}{\partial \mathrm{a}_1} & \frac{\partial y(\mathbf{a}_k,t_1)}{\partial \mathrm{a}_2} & \cdots & \frac{\partial y(\mathbf{a}_k,t_1)}{\partial \mathrm{a}_m} \\ \frac{\partial y(\mathbf{a}_k,t_2)}{\partial \mathrm{a}_1} & \frac{\partial y(\mathbf{a}_k,t_2)}{\partial \mathrm{a}_2} & \cdots & \frac{\partial y(\mathbf{a}_k,t_2)}{\partial \mathrm{a}_m} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y(\mathbf{a}_k,t_n)}{\partial \mathrm{a}_1} & \frac{\partial y(\mathbf{a}_k,t_n)}{\partial \mathrm{a}_2} & \cdots & \frac{\partial y(\mathbf{a}_k,t_n)}{\partial \mathrm{a}_m} \end{bmatrix}. \tag{12}$$

The notations in (9-12) are as follows:

- $k=1,2,...,p,\ \ p$ – number of iteration loops;
- $\mathbf{J}_{(n\,x\,m\,)}(\mathbf{a}_k,t)$ – Jacobian matrix;
- $\mathbf{I}_{(m\,x\,m\,)}$ – unit matrix;
- μ_k – scalar, its value changes during iteration;
- $\mathbf{a}=[\mathrm{a}_1,\mathrm{a}_2,...,\mathrm{a}_m]$ – model parameters searched for.

The Levenberg-Marquardt algorithm is used for computation in the following stages:

Stage 1, for k=1

- Assume the initial values of the parameters of vector $\mathbf{a}_k$;
- Assume the initial value of the coefficient μ_k (*e.g.* $\mu_k = 0.1$);
- Solve the matrix equation (12) and calculate (11);
- Calculate the value of error (10);
- Determine the parameters of vector $\mathbf{a}_{k+1}$, following (9).

Stage 2 and further steps, for $k=2,3,...,p$

- Update the values of the parameters of vector $\mathbf{a}_k$;
- Solve the matrix equations (12) , calculate (11) and (9);
- Calculate the value of error (10);
- Compare the values of error (10) for the step k and the step k–1 If the result is $y(\mathbf{a}_k,t)\geq y(\mathbf{a}_{k-1},t)$, multiply μ_k by the specified value $\lambda\in\Re$ (*e.g.* $\lambda=10$) and return to step 2 of stage 2. If the result is $y(\mathbf{a}_k,t)<y(\mathbf{a}_{k-1},t)$ divide μ_k by the value λ and return to step 1 of stage 2.

The initial parameters of vector **a** are assumed in an arbitrary way, *e.g.* **a** =[1, 1,...,1].

If in a consecutive stage the decrease in the value of error (10) is very small and insignificant, we then finish the iteration process. We fix $\mu_k=0$ and determine the final result for the parameters of vector **a** If the value of coefficient μ_k is high, it means that the solution is not satisfactory [1]. The values of parameters of the vector are not optimum ones, and the value of error (10) is not at the minimum level. At this point:

$$\mathbf{J}^T(\mathbf{a}_k,t)\mathbf{J}(\mathbf{a}_k,t)<<\mu_k\mathbf{I}, \tag{13}$$

can be assumed and this leads to the steepest descent method, for which we have:

$$\mathbf{a}_{k+1} = \mathbf{a}_k^T - \frac{1}{\mu_k}\mathbf{J}^T(\mathbf{a}_k,t)y(\mathbf{a}_k,t). \tag{14}$$

If the value of coefficient μ_k is small, it means that the values of the parameters of vector **a** are close to the optimum solution. Then:

$$\mathbf{J}^T(\mathbf{a}_k,t)\mathbf{J}(\mathbf{a}_k,t) >> \mu_k\mathbf{I}, \tag{15}$$

which means that the Levenberg-Marquardt algorithm is reduced to the Gauss-Newton method:

$$\mathbf{a}_{k+1} = \mathbf{a}_k^T - \left[\mathbf{J}^T(\mathbf{a}_k,t)\mathbf{J}(\mathbf{a}_k,t)\right]^{-1}\mathbf{J}^T(\mathbf{a}_k,t)y(\mathbf{a}_k,t). \tag{16}$$

The selection of coefficient values μ and λ depends on the programmer's experience and in practical solutions they are usually assumed as: $\mu = 0.1$ and $\lambda = 10$ [1].

6. Application Example

Let us examine an exemplary model:

$$K_6(s) = \frac{1}{(s^2+1.25\cdot s+1)\cdot(s^2+1.55\cdot s+1)\cdot(s^2+1.75\cdot s+1)}. \tag{17}$$

The mathematical model of the standard was obtained by means of the Levenberg-Marquardt optimization algorithm and procedure presented in Point 3. This model is:

$$K_s(s) = \frac{1}{(s^2+1.91\cdot s+1)\cdot(s^2+1.42\cdot s+1)\cdot(s^2+0.01\cdot s+1)}$$

$$\text{for } \omega_A = 0.55\ \frac{\text{rad}}{\text{s}},\ \omega_\varphi = 0.45\ \frac{\text{rad}}{\text{s}}. \tag{18}$$

For optimization needs, the following parameters have been assumed:

1. Digitization step 0.01s;
2. Upper boundary of frequency $\omega_{gs} = 10\ \frac{\text{rad}}{\text{s}}$;
3. The value of Δ_A and Δ_φ on the level of 5% of $A(0)$.

In order to determine signal $u_0(t)$ maximizing (1), for two constraints imposed, the procedure described in Point 4 was applied. This signal was obtained by means of evolutionary methods [7]. The magnitude A = 1 was assumed and the rate of change ϑ = 0.40 was calculated as the maximum of the impulse response $k_6(t)$ of the system (17).

Signal $u_0(t)$ is in the form:

$u_0 \Rightarrow \vartheta_+$[0.0, 1.66], ϑ_-[1.66, 5,82], -1[5.82, 6.08], ϑ_+[6.08, 11.08], +1[11.08, 11.72], ϑ_-[11.72, 16.72], -1[16.72, 20.00] (19)

and it generates the maximum value of the error equal $I_2(u_0) = 4.86\ \text{V}^2\cdot\text{s}$.

In (19) the following notation is used: ϑ_+ - signal increasing in the interval, ϑ_- - signal decreasing in the interval, ± 1 - a constant signal in [0,20] [2, 3, 7]. Time switchings of $u_0(t)$, presented by means of (19) are expressed in seconds.

Eq. 20 presents the structure of the third-order model assumed for *minimax* optimization,

$$K_3(s) = \frac{1}{(s^2 + a_1 s + 1) \cdot (a_2 s + 1)}. \tag{20}$$

As the result of optimization, the following parameters have been obtained: $a_1 = 1.37$, $a_2 = 1.71$ and *minimax* $I_2(u_0) = 2.68 V^2 \cdot s$.

Fig. 5 presents signal $u_0(t)$, errors $y(t)$ and $y_{opt}(t)$ corresponding to it.

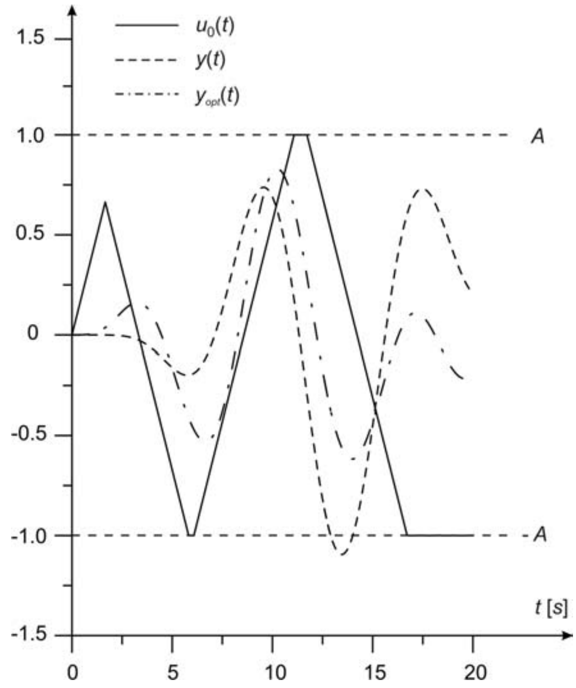

Fig. 5. Signal $u_0(t)$, and corresponding errors $y(t)$ and $y_{opt}(t)$.

7. Conclusion

The paper presents an application of the Levenberg-Marquardt algorithm to the optimization of parameters of a mathematical model of measurement systems according to *minimax* of integral square error. The presented method minimizes the mapping error of a simplified model with reference to a higher-order model, through application of an indirect standard model. The model of the standard is obtained by way of optimization of its parameters with respect to non-deformation transformation.

The presented results were obtained by means of computer programs implemented in C/C++.

As an example, the optimization of parameters of a third-order model with reference to a sixth-order system was presented.

References

[1] A. Bobon, J. Kudla, C. Ondrusek: „Zastosowanie algorytmu genetycznego i metody Levenberga-Marquardta do aproksymacji indukcyjnosci widmowych maszyny synchronicznej". *Zeszyty Naukowe Politechniki Slaskiej Elektryka*, vol. 168, Gliwice, 1999, pp. 113-124. (in Polish)

[2] E. Layer: *Modelling of Simplified Dynamical Systems. Springer-Verlag.* Berlin Heidelberg New York, 2002.

[3] E. Layer: "Non-standard input signals for the calibration and optimization of the measuring systems". *Measurement,* vol.34, issue 2, 2003, pp.179-186.

[4] N.K. Rutland: "The Principle of Matching: Practical Conditions for Systems with Inputs Restricted in Magnitude and Rate of Change". *IEEE Trans. Autom. Control*., vol. 39, 1994, pp. 550-553.

[5] N.K. Sinha, G.T. Bereznai: "Optimum approximation of high-order systems by low order model". *Int. J. Control.*, no. 21, 1971, pp. 951-959.

[6] K. Tomczyk: „Optymalizacja parametrow matematycznych modeli wzorcow ze względu na transformacje niezniekształcajaca". *Proc. V Symp. Dynamical Measurement,* Szczyrk, 2005, pp.119-128. (in Polish)

[7] K. Tomczyk: "Application of genetic algorithm to measurement system calibration intended for dynamic measurement". *Metrol. Meas. Syst.,* vol. XIII, no. 1, 2006, pp. 193-103.

[8] V. Zakian: "Perspectives of the Principle of Matching and the Method of Inequalities". *Int. J. Control*, vol. 65, 1996, pp. 147-175.