Adaptive Control System with a Multilayer Neural Network under Parametric Uncertainty Condition

Nadirbek Yusupbekov [a], Husan Igamberdiev[a], Uktam Mamirov[a]

[a] *Tashkent state technical university, Tashkent, 100095, Uzbekistan*

Abstract

The article deals with the synthesis of algorithms for constructing an adaptive control system with a multilayer neural network in conditions of parametric uncertainty. To solve the problem of synthesizing an adaptive control system, an extended system of equations for the basic structure of a generalized custom object has been compiled. Stable iterative algorithms for calculating control actions based on the methods of minimal iterations and partitioning of matrices into blocks are presented. The above algorithms make it possible to produce stable pseudo-inversion of matrices and thereby increase the accuracy of adaptive control systems with a multilayer neural network under conditions of parametric uncertainty.

Keywords 1

Adaptive control system, multilayer neural network, parametric uncertainty, iterative algorithms, matrix pseudo-inversion.

1. Introduction

At present, a large number of methods for analysis and synthesis of nonlinear automatic control systems have been developed, but each of them gives a solution only for objects of a certain class. Taking into account the influence of nonlinearities in any automatic control system encounters great difficulties, since one has to deal with the solution of nonlinear differential equations of high orders. The choice of this or that method depends on the formulation of the research problem, the type of nonlinearity and the order of the differential equation describing the system [1,2]. If the control system is described by a differential equation of the first, second or third order, then methods based on the study of processes in the phase space are used for the analysis and synthesis of nonlinear systems [1-3].

The inclusion of a multilayer neural network in a nonlinear dynamic control system expands the phase space of this system. This expansion is due to the dynamic nature of the processes of information transformation and training of the neural network by the method of back propagation of errors [4-6].

According to the theory of synergetic control [7], the expansion of the phase space in complex nonlinear systems creates prerequisites for the emergence of self-organization processes in them. In such a system, the initial external influences in relation to the original control object become the internal forces of the "extended" system. The system becomes open in the thermodynamic sense and energy, matter, or, as in our case, information from the corresponding source flow through it. Synthesized controls are the carriers of this information. Therefore, the control problem in the "extended" system is formulated as the problem of forming an algorithm for interaction between the components of the extended system, which ensures the occurrence of self-organization processes. This algorithm is a multi-layer neural network learning algorithm that minimizes some functional $Q(\sigma)$, where σ – generalized error of both training and management [4]. The structure and

Russian Advances in Fuzzy Systems and Soft Computing: selected contributions to the 8-th International Conference on Fuzzy Systems, Soft Computing and Intelligent Technologies (FSSCIT-2020), June 29 – July 1, 2020, Smolensk, Russia
EMAIL: dodabek@mail.ru; ihz.tstu@gmail.com; uktammamirov@gmail.com
ORCID: 0000-0002-2184-2368; 0000-0003-3047-5261; 0000-0002-5596-2819

properties of the system are determined both by the method of generating this error, depending on the available measurement information, and by the choice of the functional $Q(\sigma)$.

The process of structural synthesis of a control system for a dynamic object using a neural network is based on the application of the procedure for analytical synthesis of systems with multilayer neural networks based on specified invariant manifolds. The initial data for the synthesis of a control system with a neural network include a mathematical model of a control object, a model of the external environment and a control goal [6, 8].

2. Formation of an extended system of differential equations

The mathematical model of the object is written in the form of equations of state that determine the change in time of the state vector $x \in R^n$ under the action of measured disturbances $v \in R^\mu$ and control actions $u \in R^m$. It is assumed that a subset of the state vector is available for measurement $y \in R^v$. In what follows, we restrict ourselves to considering the class of control objects described by a system of differential equations linear in control:

$$\begin{gathered} x_{k+1} = f(x_k, v_k) + g(x_k, v_k)u_k, \\ y_k = h(x_k, v_k), \end{gathered} \tag{1}$$

where vector smooth functions $f(\cdot)$,$g(\cdot)$,$h(\cdot)$ are defined up to a set of parameters.

The purpose of the control and the procedure for adjusting the weighting coefficients of the synaptic connections of the neural network are specified by the local functional *Q=Q(y, r)*, or by the desired dynamic characteristics of the synthesized system in the form of a reference model $x_{k+1}^M = f_M(x_{k+1}^M, r_k)$, and the function $f_M(\cdot)$, is chosen so that the change in x^M under the action of the setting influence r_k has the desired character. Taking into account the dimension of the vectors of the input r_k and output y signals, the approximate number of network layers, the number of basic elements in the layer and the dimension of the vector of synaptic connections for each element are determined. The initial conditions of the weighting coefficients of synaptic connections of the network are set [9, 11]. At the first stage of the synthesis, an extended set of equations for the basic structure of the generalized custom object is constructed.

We write the original system of differential equations of the object (1) without taking into account the disturbances v of the external environment:

$$x_{k+1} = f(f_k, \theta_k) + g(x_k)u_k, \tag{2}$$

where $x_k \in R^n$ – state vector of the control object; $u_k \in R^m (m \leq n)$ - control action; $f(\cdot)$,$g(\cdot)$ – smooth functions of their arguments, additional requirements for which are formulated as needed; θ– vector of unknown parameters.

Then we will compose a system of differential equations that determine the dynamics of the measurement and observation system, suppression of disturbances and the actual model of disturbing influences as particular solutions of our differential equations, a system for reproducing and calculating control actions:

$$\omega_{k+1} = p(\omega_k), \tag{3}$$

where, $\omega \in R^\mu$ the state vector of dynamical systems external to the object generating disturbances.

Let us write the coupling equations of the systems of equations (2) and (3) in the form:

$$h(x, u, \omega) = 0, \tag{4}$$

Based on equations (2) - (4), we write down an extended system of differential equations:

$$\tilde{x}_{k+1} = \tilde{f}(\tilde{x}_k, \theta_k) + \tilde{g}(\tilde{x}_k)\tilde{u}_k, \tag{5}$$

Where $\tilde{x}_{k+1} \in R^q$ $(q = n + \mu)$ – extended system state vector; $\tilde{u}_{k+1} \in R^p$ – a new vector of control of the extended system. Equations (5) and are an extended system of nonlinear differential equations of the generalized tunable object.

3. Adaptive control system with multilayer neural network

Consider the problem of synthesizing an adaptive control system with nonlinear affine equations for object control (2) and using a multilayer neural network as an adaptive controller. The behavior of the control object (2) depends on the parameters θ. A set Ω_θ of possible values of θ is also given, which determines the class of adaptability. The desired dynamic characteristics of the synthesized adaptive system with a reference model are determined by a system of linear differential equations:

$$x_{k+1}^M = A^M x_k^M + B_{r_k}^M, \tag{6}$$

where $x_k^M \epsilon R^n$ – state vector of the reference model; A^M, B^M – matrices of corresponding dimensions, the coefficients of which are known and thereby determine the desired state vector of the model x_k^M under the input action r. In accordance with the stages of the synthesis of adaptive control systems [12-14] for a given system of equations of the object (2) and its reference model (6), the structure and algorithm of the regulator of the main loop of the system, the functional diagram of the loop and the adaptation algorithm are determined.

The adaptation algorithm changes the vector of adjustable parameters $w_j^{(l)}$ in such a way as to ensure the achievement of the control goal for any vector of unknown parameters θ belonging to a given class Ω_θ. Let us choose the backpropagation algorithm as such an adaptation algorithm for tuning a multilayer neural network [15-17].

It can be shown [12, 18] that the control goal is achieved by minimizing the quadratic learning functional $J = 0{,}5\sigma^T\sigma$ by the adaptation algorithm using back propagation of the error. In this case, the generalized control error σ is chosen as the left side of the homogeneous system of equations, the solution of which is satisfied by the desired control actions:

$$x_{k+1}^M - x_{k+1} = 0. \tag{7}$$

Substituting into (7) the value of the derivative of the state vector of the model (6) and replacing the state vector of the model x^M with the measured state vector x of the control object (3), we obtain an expression for the generalized error function in the form:

$$\sigma_k = A^M x_k + B^M r_k - x_{k+1}. \tag{8}$$

The control action u corresponding to (8) can be obtained from the condition $\sigma = 0$ taking into account the equations of the object (2):

$$g(x)u = b, \tag{9}$$

where $b = A^M x_k + B^M r_k - f(x)$.

As noted in [6, 9, 10, 14, 15], the absence of a regular procedure for choosing network parameters, such as the number of layers, the number of neurons in a layer, the values of the initial conditions of the weighting coefficients of the network, as well as the network tuning factor, which determine the stability and quality of transient processes in an adaptive control system, leads to the need for preliminary modeling of the synthesized adaptive control system. In the course of modeling, the selection of network parameters that satisfy the required quality of transient processes in such a system for any vector of unknown parameters θ belonging to a given class Ω_θ is carried out.

The block diagram of an adaptive control system for a nonlinear dynamic object with an explicit reference model and a dynamic multilayer neural network of the “inverse control model” type as an adaptive controller is shown in the figure. The reference model is also implemented in the form of a dynamic neural network of the “direct learning model” type [9, 14, 15].

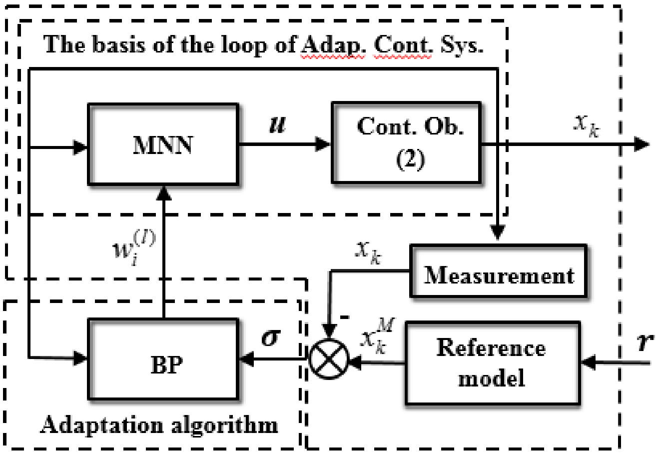

Figure 1: Block diagram of an adaptive system with a multilayer neural network.

4. Stableiterative algorithms for calculating control actions based on the methods of minimal iterations

Direct use of relation (9) can lead to a decrease in the control accuracy of u, since the pseudo-inverse operation is used here. In the case when the matrix *g(x)* is a matrix of incomplete rank, then the problem under consideration is ill-posed. To give numerical stability to the procedure for pseudo-inversion of the matrix *g(x)*, it is expedient to use the concepts of regular methods [19-22]. An extensive literature is devoted to the discussion of the problems of numerical determination of pseudoinverse matrices [23-27].

We will apply the $\left(g^T(x)g(x)\right)$– minimal iteration method to solving the overdetermined system (9) [23, 24, 28].

For some arbitrary vector $s \in C^m$, a sequence of vectors q_1, q_2,…, where $q_i \in C^m$, is constructed using the following formulas:

$$q_1 = s, v_1 = g(x)q_1, q_2 g(x)v_1 - \gamma_1 q_1, \tag{10}$$

then for *i=2,3…* are calculated

$$q_{1+1} = g(x)v_i - \gamma_i q_i - \delta_i q_{i-1}, \tag{11}$$

where

$$\gamma_i = \frac{(g(x)v_i, g(x)v_i)}{(v_i, v_i)}, \delta_i = \frac{(v_i, v_i)}{(v_{i-1}, v_{i-1})}, v_i = g(x)q_i, v_i \in C^n. \tag{12}$$

This method is considered purely iterative. It is assumed that all vectors q_1, q_2,… are nonzero.

In parallel, *i=2,3…* the values are calculated

$$u_0 = 0, u_i = u_{i-1} + \eta_i q_i, u_i \in C^m, \tag{13}$$

where

$$\eta_i = \frac{(b, v_i) - (u_{i-1}, g(x)v_i)}{(v_i, v_i)}. \tag{14}$$

With an exact implementation of the method, if the vectors $q_1,\dots,q_m$ are linearly independent, the computational process ends at the *m*-th step, i.e. the modified method, like the original one, is final; the result is $u_* = u_m$. In [28] it was shown that the error functional, which for this method has the form that is shown below

$$F_1(u) = (g(x)(u - u_*), g(x)(u - u_*)),$$

when using modification during iterations, it is guaranteed not to increase at each step.

However, there is the following complication of the calculation procedure when solving large problems. If the matrix *g(x)* contains elements of large modulus, then its norm is large, and then the vectors q_i and v_i can strongly increase in norm with increasing number *i*.

To eliminate this drawback, it is often proposed to modify formulas (10) - (14) by applying the normalization of the vectors x at each iteration step [23, 26].

Following [24, 28], we present the corresponding formulas for the modified method with normalization:

$$q_1 = s, v_1 = \frac{1}{\mu_1} g(x) q_1, \mu_1 = |g(x) q_1|, q_2 = g(x) v_1 - \gamma_1 q_1,$$

further, for *i=2, 3, …* we calculate

$$q_{i+1} = g(x) v_1 - \gamma_1 q_1 - \delta_i q_{i-1},$$

where

$$v_i = \frac{1}{\mu_1} g(x) q_i, \mu_i = |g(x) q_i|, \gamma_i = \frac{(g(x) v_i, g(x) v_i)}{\mu_i}, \delta_i = \frac{\mu_i}{\mu_{i-1}}.$$

In parallel for *i=1, 2, …*

$$u_0 = 0, u_i = u_{i-1} + \alpha_i g_i,$$

is computed, where

$$\eta_i = \frac{(b, v_i) - (u_{i-1}, g(x) v_i)}{\mu_i}.$$

To invert the matrix $g(x)$ in (9), it is also advisable to use the method of matrix partitioning according to the Frobenius formula [23, 24].We can divide the redefined matrix $g(x)$ into blocks

$$g(x) = \begin{bmatrix} R & U \\ U^T & S \end{bmatrix}. \tag{15}$$

where$R$ is a nondegenerate square matrix ($|R| \neq 0$) и $r_R = r_{g(x)}$. Then, the equality $S = U^T R^{-1} U$ [26] is true, and therefore

$$g(x) = \begin{bmatrix} R \\ U^T \end{bmatrix} R^{-1} (R \quad U),$$

$$g(x) = \begin{bmatrix} R \\ U^T \end{bmatrix} (I \quad R^{-1} U), (I \quad R^{-1} U) = R^{-1} (R \quad U).$$

Therefore

$$g^+(x) = (R \quad U)^+ (R^{-1}) \begin{bmatrix} R \\ U^T \end{bmatrix}^+ = (R \quad U)^+ R \begin{bmatrix} R \\ U^T \end{bmatrix}^+.$$

Applying the formulas $U^+ = (U^T U)^{-1} U^T$and $(U^T)^+ = U (U^T U)^{-1}$, we finally find

$$g^+(x) = \begin{bmatrix} R \\ U^T \end{bmatrix} (R R^T + U U^T)^{-1} R (R^T R + U U^T)^{-1} (R^T \quad U). \tag{16}$$

Formula (16) gives an explicit expression for the pseudoinverse matrix $g^+(x)$ through the blocks of R, U, U^T.

To impart greater numerical stability when implementing the pseudo-inversion algorithm (16), it is advisable to represent it in the form:

$$g^+(x) = \begin{bmatrix} R \\ U^T \end{bmatrix} Z_\alpha(P) R Z_\beta(D) (R^T \quad U). \tag{17}$$

where $P = R R^T + U U^T, D = R^T R + U U^T, Z_\alpha(P) = (P + \alpha I)^{-1}, Z_\beta(D) = (D + \beta I)^{-1}, \alpha > 0$, $\beta > 0$– regularization parameters.

It is expedient to select the regularization parameters α and β in (17) based on the method of model examples [29].

5. Conclusion

The above algorithms make it possible to produce stable pseudo-inversion of matrices and thereby increase the accuracy of adaptive control systems with a multilayer neural network under conditions of parametric uncertainty.

6. References

[1] T. Ya. Lazareva, Yu. F. Martemyanov, Foundations of the theory of automatic control. Publishing house of Tamb. state tech. University,Tambov, 2004.

[2] V. A. Okhorzin, K. V. Safonov, Control theory. Publishing house Sib. State aerospace Un-t. - Krasnoyarsk, 2011.

[3] O. O. Zaripov, D. A. Akhmedov, U. F. Mamirov, Adaptive estimation algorithms for the state of nonlinear dynamic systems, International Journal of Psychosocial Rehabilitation, 24(3) (2020) 247-253. doi: 10.37200/IJPR/V24I3/PR200776

[4] V. A. Terekhov, V. B. Yakovlev, D. V. Efimov, I. Yu. Tyukin, Control and Learning Algorithms in Intelligent Systems with Multilayer Neural Networks, in: Work of the 2-nd Int. symposium “INTELS'96”, St. Petersburg, Moscow, 1996, 11-14.

[5] New concepts of general management theory, Sat. scientific. Works, Ed. A.A. Krasovsky. Taganrog Radio Engineering University, Moscow-Taganrog, 1995.

[6] V. A. Terekhov, D. V. Efimov, I. Yu. Tyukin, V. N. Antonov, Neural network control systems. Petersburg University,St. Petersburg, 1999.

[7] A. A. Kolesnikov, Synergetic methods for managing complex systems: The theory of system synthesis. Ed. Stereotype, Moscow, 2019.

[8] V. A. Terekhov, D. V. Efimov, I. Yu. Tyukin, Structural synthesis of adaptive control systems with multilayer neural networks based on the method of Lyaggunov functions, Saint Petersburg State Electrotechnical University, GETU. St. Petersburg, 1997.

[9] S. Haykin, Neural Networks and Learning Machines. - 3rd ed. Prentice Hall, 2009.

[10] S. Omatu, M. Khalid, R. Yusof,Neuro-Control and its Applications. Springer,London, 1995.

[11] T. V. Botirov, S. B. Latipov, B. M. Buranov, “About one synthesis method for adaptive control systems with reference models”,Journal of Physics: Conference Series, 1515, 2 (2020) 1-6. doi:10.1088/1742-6596/1515/2/022078

[12] G. N. Reshetnikova, Adaptive systems. TomskStateUniversity, Tomsk, 2016.

[13] N. R. Yusupbekov, S. M. Gulyamov, M. Y. Doshchanova, Neural identification of a dynamic model of a technological process, in: International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2019, 1-5, doi: 10.1109/ICISCT47635.2019.9011912.

[14] Yu. I. Eremenko, A. I. Glushchenko, On the development of a method for choosing the structure of a neural network for solving the problem of adapting the parameters of linear controllers, Management of large systems. 62 (2016) 75-123.

[15] V. I. Vasiliev, B. G. Ilyasov, Intelligent control systems. Theory and practice. Radiotekhnika, Moscow, 2009.

[16] I. A. Prudnikov, Algorithm for learning a multilayer neural network by the method of back propagation of an error, Eurasian Scientific Journal, 2 (2016) 1-5.

[17] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representations by error propagation, Parallel Distributed Processing. 1, 8 (1986) 318-362.

[18] P. Patrick, Minimisation methods for training feedforward Neural Networks, Neural Networks, 7, 1 (1994) 1-11.

[19] Ch. L. Lawson, R. J. Hanson, Solving least squares problems, Philadelphia: SIAM, 1995.

[20] N. R. Yusupbekov, H. Z. Igamberdiev, U. F. Mamirov, Algorithms of sustainable estimation of unknown input signals in control systems, Journal of Multiple-Valued Logic and Soft Computing, 33, 1-2 (2019) 1-10. https://www.oldcitypublishing.com/pdf/9291

[21] N. R. Yusupbekov, H. Z. Igamberdiev, J. U. Sevinov, Formalization of Identification Procedures of Control Objects as a Process in the Closed Dynamic System and Synthesis of Adaptive Regulators,

Journal of Advanced Research in Dynamical and Control Systems. 12, 6 (2020) 77-88. doi:10.5373/JARDCS/V12SP6/SP20201009.
[22] A. I. Zhdanov, Introduction to methods for solving ill-posed problems. Aerospace University,Samara, 2006.
[23] F. R. Ganthmaher, Matrix theory –M.: Nauka, 1988.
[24] G. H. Golub, Ch. F. Van. Loan, Matrix computations.2nd ed. The Hopkins University, London, 1989.
[25] R. Horn, C. Johnson, Matrix analysis, 1989.
[26] J. W. Demmel, Applied Numerical Linear Algebra. University of California Berkeley, California, 1997.
[27] V. M. Verzhbitsky, Computational linear algebra. Higherschool. Moscow, 2009.
[28] L. F. Yukhno, A modification of conjugate direction methods, Computational Mathematics and Mathematical Physics, 46, 1 (2006) 5-9.
[29] A. F. Verlan, V. S. Sizikov, Integral equations: methods, algorithms, programs. NaukovaDumka, Kiev, 1986.