[Кириченко, Крак, Полищук,2004] Н.Ф.Кириченко, Ю.В. Крак.А.А. Полищук Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей.// Кибернетика и системный анализ –2004.–№3.

[Donchenko, Kirichenko,2005.] V.S Donchenko, N.F. Kirichenko Generalized Inverse in Control with the constraints.// Cybernetics and System Analysis.– v.39 (6) November-December– 2003.–p.854-861(USA).

[Линник,1962] Ю.В.Линник Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. Изд. 2-е. – М.:Физматгиз.– 1962.– 349 с.

[Вапник,1979] В.Н. Вапник. Восстановление зависимостей по эмпирическим данным. – М.: Наука. – 1979. – 447 с.

[Ивахненко,1969] А.Г. Ивахненко Самоорганизующиеся системы распознавания и автоматического управления. – К.: Техника. – 1969. – 395 с.

Authors' Information

Volodymyr Donchenko – Kyiv National Taras Shevchenko University (Ukraine), Professor, e-mail: voldon@unicyb.kiev.ua

Mykola Kirichenko – Institute of the Cybernetics, National Academy of Sciences (Ukraine), Professor

Yuriy Krivonos – Institute of the Cybernetics, National Academy of Sciences (Ukraine), Member Correspondent of the National Academy of Sciences.

NEURAL NETWORK BASED OPTIMAL CONTROL WITH CONSTRAINTS

Daniela Toshkova, Georgi Toshkov, Todorka Kovacheva

***Abstract**: In the present paper the problems of the optimal control of systems when constraints are imposed on the control is considered. The optimality conditions are given in the form of Pontryagin's maximum principle. The obtained piecewise linear function is approximated by using feedforward neural network. A numerical example is given.*

***Keywords**: optimal control, constraints, neural networks*

***ACM Classification Keywords**: I.2.8 Problem Solving, Control Methods, and Search*

Introduction

The optimal control problem with constraints is usually solved by applying Pontryagin's maximum principle. As it is known the optimal control solution can be obtained computationally. Even in the cases when it is possible an analytical expression for optimal control function to be found, the form of this function is quite complex. Because of that reason the possibilities of using neural networks for solving the optimal control problem are studied in the present paper.

The ability of neural networks to approximate nonlinear function is central to their use in control. Therefore it can be effectively utilized to represent the regulator nonlinearity. Other advantages are their robustness, parallel architecture.

Lately, different approaches are proposed in the literature treating the problem of constrained optimal control for using neural networks. In [Ahmed 1998] a multilayered feedforward neural network is employed as a controller. The training of the neural network is realized on the basis of the so called concept of Block Partial Derivatives. In [Lewis 2002] a closed form solution of the optimal control problem with constraints is obtained solving the associate Hamilton-Jacobi-Bellman (HJB) equation. The solution of the value function of HJB equation is approximated by using neural networks.

In the present paper the problem of finding the optimal control with constraints is considered. A numerical example is given.

Problem Statement

The control system, described by following differential equations is considered:

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij} x_i + b_i u \qquad (i = 1, 2, \ldots, n) \tag{1}$$

where x_j are phase coordinates of the system, function u describes the control action and a_{ij} are constant coefficients. The admissible control u belonging to the set U of piecewise linear functions is constrained by the condition

$$|u(t)| \le 1 \tag{2}$$

Following problem for finding the optimal control is formulated. To find such a control function $u(x_1,\ldots, x_n)$ for the system (1) among all the admissible controls that the corresponding trajectory $(x_1(t),\ldots, x_2(t))$ of the system (1) starting from any initial state $(x_1(0),\ldots, x_2(0))$ to tend to zero at $t \to \infty$ and the performance index

$$J = \int_0^{\infty} \left(\sum_{i=1}^{n} q_i x_i^2 + r u^2 \right) dt \tag{3}$$

to be converging and to take its smallest possible value. The coefficient qi and r are positive weight constants.

Optimality Conditions

The notation is introduced [Pontryagin 1983]:

$$f_0(x_1, \ \ldots, \ x_n, \ u) = \sum_{j=1}^{n} q_j x_j^2 \ + \ ru^2 \tag{4}$$

$$f_i(x_1, \ \ldots, \ x_n, \ u) = \sum_{j=1}^{n} a_{ij} x_j^2 \ + \ b_i u \qquad (i = 1,\ldots,n) \tag{5}$$

One more variable θ_0 is added to the state variables $(x_1,\ldots, x_n)$ of the system (1) [Chjan 1961]. It is a solution of the following equation

$$\frac{dx_0}{dt} = f_0(x_1, \ \ldots, \ x_n, \ u) \tag{6}$$

and initial condition $x_0(0) = 0$. Then the quantity J according to (9) becomes equal to the boundary of x(t) when $t \to \infty$. The system of differential equation, which are adjoint to the system (7) is composed with new variables $\Psi=\{\Psi_0, \ \Psi_1, \ \ldots, \ \Psi_n\}$:

$$\frac{d\Psi_0}{dt} = -\sum_{\alpha=0}^{n} \frac{\partial f_\alpha}{\partial x_0} \Psi_\alpha = 0 \tag{7}$$

$$\frac{d\Psi_i}{dt} = -\sum_{\alpha=0}^{n} \frac{\partial f_\alpha}{\partial x_0} \Psi_\alpha = -2q_i\theta_i - \sum_{j=1}^{n} a_{ij} \Psi_j \qquad (i = 1,\ldots,n) \tag{8}$$

After that the Hamilton function is composed:

$$H(\theta,\Psi,u) = \sum_{\alpha=0}^{n} \Psi_\alpha \frac{dx_\alpha}{dt} = \sum_{\alpha=0}^{n} \Psi_\alpha f_\alpha(x_1,\ldots x_n,u) = \Psi_0 \left(\sum_{i=1}^{n} q_i x_i^2 + ru^2 \right) + \sum_{i=1}^{n} \Psi_i \left(\sum_{j=1}^{n} a_{ij} x_j + b_i u \right) \tag{9}$$

In the right-hand side of Eq. (9) the quantity *u* is contained in the expression

$$H_1 = r\Psi_0(t)u^2(t) + u(t)\sum_{i=1}^{n} b_i\Psi_i(t) \tag{10}$$

Because of that the condition for maximum of H coincide with the condition

$$\begin{aligned}\max_{|u|\le 1} H_1 &= \max_{|u|\le 1}\left[r\Psi_0(t)u^2(t) + u(t)\sum_{i=1}^{n} b_i\Psi_i(t)\right] = \\ &= \max_{|u|\le 1}\left\{ r\Psi_0(t)\left[u(t) + \frac{1}{2r\Psi_0}\sum_{i=1}^{n} b_i\Psi_i(t)\right]^2 - \frac{1}{4r\Psi_0}\left[\sum_{i=1}^{n} b_i\Psi_i(t)\right]^2\right\}\end{aligned} \tag{11}$$

Having in mind condition (7) the quantity Ψ_0 is a constant. As its value can be any negative number it is set to $\Psi_0 = -1$.

After placing this value in Eq. (11) the maximum of the expression in the square brackets will be reached when the first negative addend becomes zero if it is possible or takes its minimal absolute value. The expression

$$\left[u(t) - \frac{1}{2r}\sum_{i=1}^{n} b_i\Psi_i(t)\right]^2 \tag{12}$$

will take its minimal absolute value if on condition $|u| \le 1$ a value of the following kind is chosen for u

$$u(t) = \begin{cases} \dfrac{1}{2r}\sum_{i=1}^{n} b_i\Psi_i & \text{at} \quad \left|\dfrac{1}{2r}\sum_{i=1}^{n} b_i\Psi_i\right| \le 1 \\ 1 & \text{at} \quad \left|\dfrac{1}{2r}\sum_{i=1}^{n} b_i\Psi_i\right| \ge 1 \\ -1 & \text{at} \quad \left|\dfrac{1}{2r}\sum_{i=1}^{n} b_i\Psi_i\right| \le -1 \end{cases} \tag{13}$$

The values of $\psi_c(t)$ can be determined if the adjoint equations (7), (8) are solved. This leads to the requirement the initial values of $\psi_0(0)$ to be found beforehand.

First u(t) is assumed not to reach its boundary values. Then after placing the upper expression from (13) instead of u(t) in Eqs. (1), (7) и (8) one obtains

$$\begin{aligned}\frac{dx_i}{dt} &= \sum_{j=1}^{n} a_{ij}x_j + \frac{b_i}{2r}\sum_{j=1}^{n} b_j\Psi_j \quad (i = 1,...,n) \\ \frac{d\Psi_i}{dt} &= 2q_i x_i - \sum_{j=1}^{n} a_{ji}\Psi_j\end{aligned} \tag{14}$$

This system of equations has to be solved with the initial conditions $x_1(0),\ldots, x_n(0)$ as well as with the final (boundary) conditions

$$\lim_{t\to\infty} x_1(t) = \lim_{t\to\infty} x_2(t) = \ldots = \lim_{t\to\infty} x_n(t) = 0 \tag{15}$$

It is necessary the appropriate initial conditions $\psi_1(0), \ldots, \psi_n(0)$ to be selected in such a way that the initial and the final conditions for $x_1(t),\ldots, x_n(t)$ to be satisfied.

The relationship between $x_i(0)$ and $\psi_i(0)$ has the following form [4]:

$$\Psi_i(0) = \sum_{j=1}^{n} \chi_{ij} x_i(0) \qquad (i = 1, \ldots, n) \tag{16}$$

These relationships have to be kept in any time, for which one can always assume to be the initial one. Therefore the optimal control *u* within the boundaries is determined and it has the following form:

$$u = \frac{1}{2r}\sum_{i=1}^{n} k_i x_i \tag{17}$$

where $k_i = \sum_{j=1}^{n} b_j \chi_{ji}$

The expression (17) holds only in the cases when the absolute value of the sum $\frac{1}{2r}(k_1x_1 + ... + k_nx_n)$ is not greater than one. When $\left|\frac{1}{2r}(k_1x_1 + ... + k_nx_n)\right| > 1$ the optimal control passes on the boundary i.e. |*u*| = 1, if the right hand boundary conditions are satisfied i.e. the solution of the system (1), which became nonlinear in connection to the nonlinear relationship between *u* and $x_1, \ldots, x_n$, tends to zero. In other words the solution of the system has to be asymptotically stable. Thus the optimal control is defined by the expression

$$u(t) = \begin{cases} \frac{1}{2r}\sum_{i=1}^{n} k_i x_i & \text{at} \quad \left|\frac{1}{2r}\sum_{i=1}^{n} k_i x_i\right| \le 1 \\ 1 & \text{at} \quad \frac{1}{2r}\sum_{i=1}^{n} k_i x_i \ge 1 \\ -1 & \text{at} \quad \frac{1}{2r}\sum_{i=1}^{n} k_i x_i \le -1 \end{cases} \tag{18}$$

Structure and Training of the Neural Network

For the control function realization a feed forward neural network with one hidden layer is used. Thus the necessity of solving a large number of equations for determining the coefficients k_i drops off.

The neural network consists of three layers – an input, output and hidden one. The input and hidden layers have five neurons and the output layer – one. The activation function of the output neuron is piecewise linear. The neural network output is

$$y = \begin{cases} +1 & \varphi(\nu) \ge 1 \\ \varphi(\nu) & |\varphi(\nu)| \le 1 \\ -1 & \varphi(\nu) \le 1 \end{cases} \tag{19}$$

where $\nu = w^T z$. The neural network input is denoted z and w is the neural network weight. The neural network output represents the control u, x – the state vector and weights are the coefficient k.

The neural network is trained according to the back-propagation algorithm. Let the training sample $\{z(n), d(n)\}_{n=1}^{N}$ be given where z(n) are the system states and d(n) is the corresponding control, which are known preliminarily. The neural network is trained according to the back-propagation algorithm [Haykin 1999].

Simulation Results

In order to verify the suggested approach for solving the optimal control problem following system is considered:

$$\frac{dx_1}{dt} = x_2$$

$$\frac{dx_2}{dt} = -x_1 - 2x_2 + u$$

and the control is constrained by

$$|u| \leq 1$$

The performance index to be minimized is of the form:

$$\int_{0}^{\infty} [x_1^2(t)+x_2^2(t)+u^2(t)dt$$

The problem is solved by using Pontryagins principle and neural networks. The results, which are obtained by both approaches, are compared. In Fig. 1 the optimal control, obtained by using neural networks is shown. Fig. 2 depicts the corresponding states trajectory. In Fig. 3 and Fig. 4 the optimal control, obtained by applying the maximum principle and the corresponding trajectory are given respectively. By 1 and 2 are denoted x_1 and x_2 respectively.

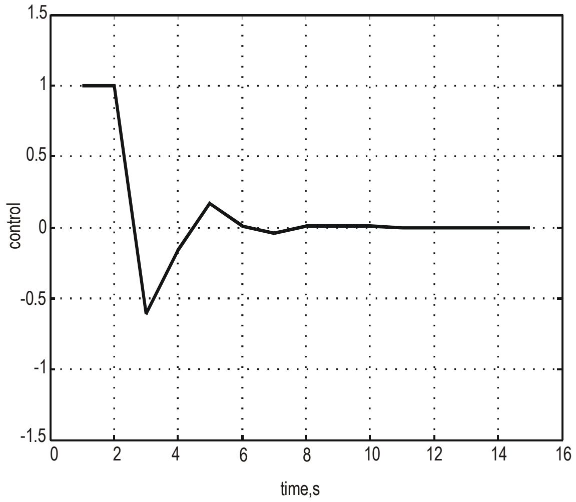

Fig. 1. Optimal control, obtained by using the suggested neural network based approach

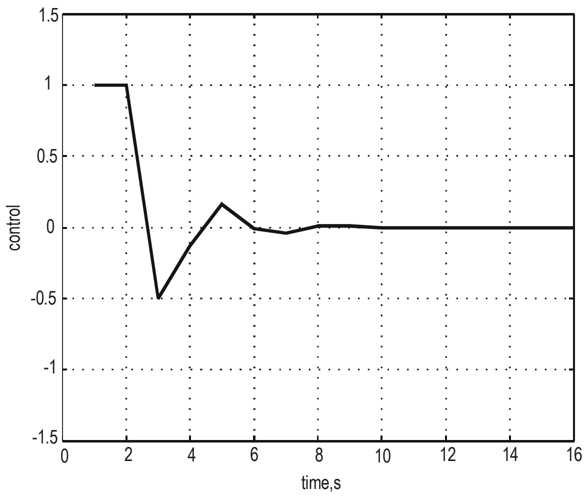

Fig. 2. Optimal control, obtained by applying the maximum principle

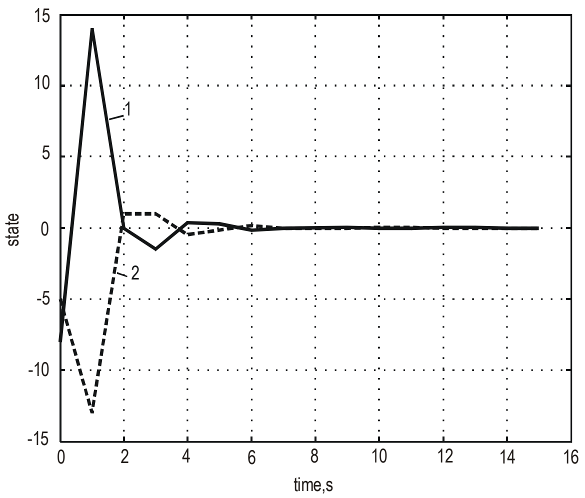

Fig.3 Optimal trajectory (neural network based approach)

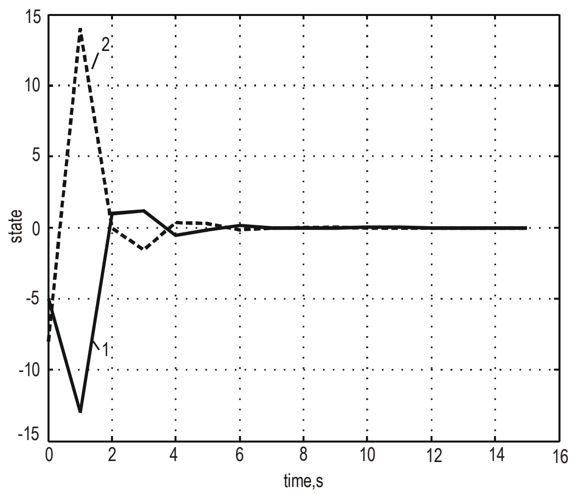

Fig.4 Optimal trajectory (Pontryagin's maximum principle)

Conclusion

In the present paper an approach for optimal constrained control based on using of neural networks is suggested.

On the basis of the simulation experiments one can say that the proposed approach for optimal control is accurate enough for the engineering practice. The suggested approach can be applied for optimal control in real time, where the control is constrained.

Bibliography

[Ahmed 1998] M.S. Ahmed and M.A. Al-Djani. Neural regulator design. Neural Networks, Vol. 11, pp. 1695-1709, 1998

[Chjan 1961] Zh.-V.Chjan. A problem of optimal system synthesis through the maximum principle. Automatization and telemechanics, Vol. XXII, No.1, 1961, pp. 8-12

[Haykin 1999] S. Haykin. Neural Networks: A Comprehensive Foundation., 1999, 2nd ed, Macmillan College Publishing Company, New York

[Lewis 2002] F.L. Lewis and M. Abu-Khalaf. A Hamilton-Jacobi setup for constrained neural network control. Proceedings of the 2003 IEEE International Symposium on Intelligent Control Houston, Texas *October 5-8. 2003

[Pontryagin 1983] L.S. Pontryagin, V.G. Boltyanskiy, R.V. Gamkrelidze and E.F.Mishtenko. Methematical theory of the optimal processes. Nauka, Moskow, 1983, in Russian

Authors' Information

Georgi Toshkov –e-mail: g_toshkov2006@abv.bg

Daniela Toshkova – e-mail: daniela_toshkova@abv.bg

Technical University of Varna, 1, Studentska Str, Varna, 9010, Bulgaria

Todorka Kovacheva – Economical University of Varna, Kniaz Boris Str, e-mail: todorka_kovacheva@yahoo.com

LINEAR CLASSIFIERS BASED ON BINARY DISTRIBUTED REPRESENTATIONS

Dmitri Rachkovskij

Abstract: Binary distributed representations of vector data (numerical, textual, visual) are investigated in classification tasks. A comparative analysis of results for various methods and tasks using artificial and real-world data is given.

Keywords: Distributed representations, binary representations, coarse coding, classifiers, perceptron, SVM, RSC

ACM Classification Keywords: C.1.3 Other Architecture Styles - Neural nets, I.2.6 Learning - Connectionism and neural nets, Induction, Parameter learning

Introduction

Classification tasks consist in assigning input data samples to one or more classes from a predefined set [1]. Classification in the inductive approach is realized on the basis of a training set containing labeled data samples. Usually, input data samples are represented as numeric vectors. Vector elements are real numbers (e.g., some measurements of object characteristics or their function) or binary values (indicators of some features in the input data).

This vector information often isn't explicitly relevant to the classification, therefore some kind of transformation is necessary. We have developed methods for transformation of input information of various kinds (such as numerical [2], textual [3], visual [4]) to binary distributed representations. These representations can then be