[WS-GRAM, Int] WS GRAM Public Interfaces // Available at http://www.globus.org/toolkit/docs/4.0/execution/wsgram/WS_GRAM_Public_Interfaces.html

[Condor, 2005] Condor Project Home // http://cs.wisc.edu/condor/

[MDS, 2005] Information Services (MDS) : Key Concepts // http://globus.org/toolkit/docs/4.0/info/key-index.html#id2770162

[Ruby] Ruby Programming Language // http://www.ruby-lang.org

[NCFTP-Client] A command line ftp client for Unix // Available at http://www.ncftp.com/ncftp/

[Miller, 2005] Naill Miller. How to write a simple Execution Aggregator Information Provider for MDS4 // Available at http://www-unix.mcs.anl.gov/~neillm/mds/writing_a_gt4_information_provider.html

[Raicu, 2005] Ioan Raicu. A performance study of the Globus Toolkit and Grid services via DiPerF, an automated distributed performance testing framework // M.S. Thesis, The University of Chicago, May 2005

Author's Information

Andrey Y. Shelestov, PhD. – senior scientific researcher, Space Research Institute NASU-NSAU; 40 Glushkov Ave 03187, Kyiv, Ukraine; e-mail: inform@ikd.kiev.ua

Alexander G. Lobunets, MSc. – junior scientific researcher, Space Research Institute NASU-NSAU; 40 Glushkov Ave 03187, Kyiv, Ukraine; e-mail: alexander.lobunets@gmail.com

Michael B. Korbakov, Bcs. – system developer, Space Research Institute NASU-NSAU; 40 Glushkov Ave 03187, Kyiv, Ukraine; e-mail: alexander.lobunets@gmail.com

REPRESENTATION OF NEURAL NETWORKS BY DYNAMICAL SYSTEMS

Volodymyr Donchenko, Denys Serbaev

Abstract: Representation of neural networks by dynamical systems is considered. The method of training of neural networks with the help of the theory of optimal control is offered.

Keywords: neural nets, dynamical systems, training.

ACM Classification Keywords: G.1.6 Optimization, G.1.2 Approximation, I.2 Artificial Intelligence

Introduction

At present time neural networks have received the wide circulation and are successfully applied to the decision of various complicated problems such as, for example, control and identification of nonlinear systems, the analysis of the financial markets, modelling of signals etc. Quality of work of neural networks depends from efficiency of the chosen algorithm of definition of weights of a network for achievement of required accuracy on training and test samples. The method of adjustment of weights of neural net on the basis of the theory of optimal control and representation of neural network by dynamical system is offered below.

Representation of a Network by the System of Recurrent Equations

As is known [4], neural network - is set of the same elements - neurons, - divided in to the parts-layers consistently connected among themselves. Each of neurons from, actually, answers scalar function of vector argument $y=F(w^T x)$, that is superposition of the linear form with a vector of the linear form w which name as vector of weights, - and scalar function F. Function F is activation function of neuron. Vector argument x is an input of the neuron, scalar value y - an output. Inputs of the neuron belong to this or that layer on which the network is broken. These layers are ordered consistently so, that outputs of all neurons of the previous layer move on inputs of any of neurons of the following layer. An input of the first layer is the signal which is input for whole network. For standardization of designations, we shall consider that the input forms a layer with number 0. This layer does not contain neurons and, actually, sets an input signal $x(0)=x_0$ which consists from l_0 components. The layer with number N is output. Each of layers with appropriate numbers k: $k=0,\dots,N$ has l_k neurons. Scalar outputs of the neurons of one layer are united in one vector $x(k)$, $k=0,\dots,N$, which we shall consider as output of the appropriate layer. Dimension of such vector coincides with quantity l_k, $k=0,\dots,N$ of the neurons in the appropriate layer.

Let's consider that all neurons of the same layer have identical weights. General weight for all neurons of one layer we shall designate accordingly by number of a layer $w\ (k)$: $w(k)=(w(k)_1,\dots,w(k)_{l_{k-1}})^T$, k=1,...,N. Dimension of a vector of weights, naturally, coincides with quantity l_{k-1} neurons of the layer-predecessor.

For the activation functions of any of neurons of the appropriate layer, we shall consider, that they are different for every neuron and will be designated $F_i^{(k)}(z)$, $i=1,\dots,l_k$, $k=1,\dots,N$. We shall remind, that activation functions are scalar functions of scalar arguments.

Thus, transformation of an input signal $x(0)=x_0$ by consecutive layers of the neural network is described by system of recurrent equations:

$$x(k+1)=\begin{pmatrix} F_1^{(k+1)}(w(k+1)^T x(k)) \\ \dots \\ F_{l_{k+1}}^{(k+1)}(w(k+1)^T x(k)) \end{pmatrix}, \text{ k=0,...,N–1...} \tag{1}$$

Having designated vector function $(F_1^{(k+1)}(z),\dots,F_{l_k}^{(k+1)}(z))$ through $g(z,k+1)$, we shall rewrite (1) in more compact kind:

$$x(k+1)=g(w(k+1)^T x(k),\ k+1),\ k=0,\dots,N-1. \tag{2}$$

Taking into account that each layer of a network carries out mapping from one linear space in another according to (1) or (2), the generalized scheme of the neural network may be represented by figure 1.

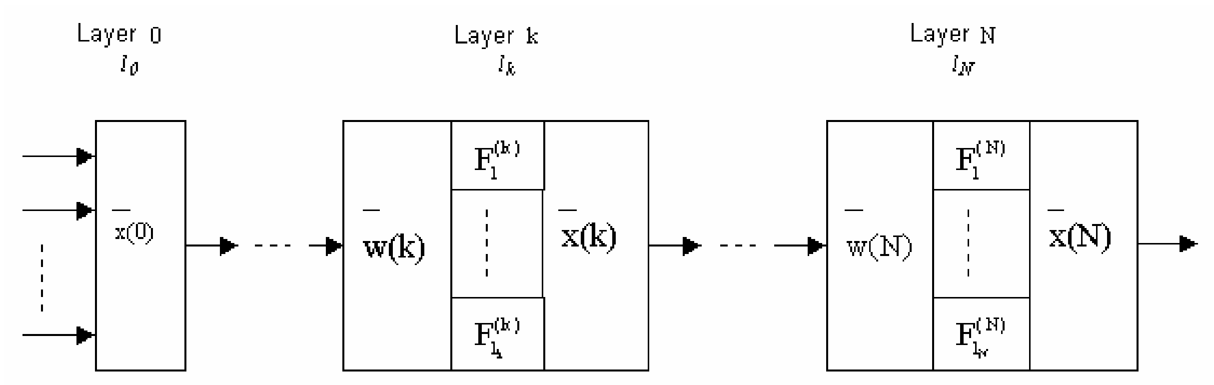

Figure 1.

Layers of the network are represented by rectangulars, which carry out transformation according to (1) or (2) in figure 1. Selected parts, in every such rectangular, which have natural interpretation: $w(k)$ answers input synapses, parts with activation functions $F_1^k,\dots,F_{l_k}^k$ answer neurons, the part with $x(k)$ is responsible for concentration of outputs of neurons of a layer in a single output of the whole layer.

The Task of Training of Neural Networks

The problem of training of the neural networks is to adjust weights $w(k)$, $k=1,\dots,N$ of the layers of the network so that on the given training sample: for sequences of pairs $(x_0^{(1)},y^{(1)}),\dots,(x_0^{(M)},y^{(M)})$, $x_0^{(i)}\in R^{l_0}, y^{(i)}\in R^{l_N}$,

$i=\overline{1,N}$, in which first component is interpreted as one of variants of an input of a network and has dimension of an input layer l_0, and the second – as the desirable output of a network and has dimension of an output layer l, – the least deviation of output signals of a network from desirable was reached. Thus training of the neural networks lies in minimization of functional J(w(1),…,w(N)), which is determined by following equation:

$$J(w(1),\dots,w(N))= \sum_{i=1}^{M} \| y^{(i)} - x^{(i)}(N) \|^2 , \qquad (3)$$

where $x^{(i)}(N)$, $i=\overline{1,N}$ – an output of a network for element i of a training sample:

the united output of neurons of the last layer of neural network when appropriate input value of an element of training sample moves on an input.

Let's note, that if training sample consists of one element i.e. that if a signal moves on an input, and the network should be trained on output signal, cost functional of training will be following:

$$J(w(1),\dots, w(N))= \| y - x(N) \|^2 . \qquad (4)$$

The Standard Approach for Solving of a Task of Training of Neural Network

The standard approach to training of neural network is approach according to which on the fixed input from possible variants of an input of training sample and previously fixed at any level weights of layers there is a consecutive change of these weights of each of layers in a direction opposite to a gradient of functional J(w(1),…,w(N)) by weight of the appropriate layer w(k), k=1,…,N. Coefficients which define length of a step in the appropriate direction, should be taken rather small.

Solution of a Task of Training of Neural Network by Application of Results for the Generalized Control System

Principle for the solution of a task of training is a statement of theorems 1, 2 below about representation of a task of training of a network by the generalized control system accordingly for: one trajectory and a beam of such trajectories.

Theorem 1. The task of training of neural network on one input signal represents by itself optimization task for the generalized control system, in which accordingly:

- phase variables $x(k)$, $k=0,...,N$ is outputs of layers with appropriate numbers;
- control $u(k)$ with appropriate number $k=0,...,N-1$ coincides with weight of layer $k+1$ and it is determined by equation: $u(k) = w(k+1), k=0,\dots,N-1;$
- functions f, which describe recurrent connection between values of a phase variable, are determined by functions g of outputs of layers by equations:

$$f(x(k), u(k), k)=g(w(k+1)^T x(k),k+1),\ k=0,\dots,N-1 \qquad (5)$$

- functional $I(u_0,\dots,u_{N-1})$ coincides with J(w(1),…,w(N)) from (2).

The evidence of the theorem is given in [2].

Effect of the theorem 1 is the opportunity to use the theorem of a kind of gradients for the generalized control system for calculation of gradients of cost functional of training in the training of neural network tasks. [2]

Theorem 2. The task of training of neural network on training sample of any quantity M represents by itself optimization task for a beam of dynamics, in which accordingly:

- phase variables X(k), k=0,…,N are matrix and consist from columns $x^{(i)}(k)$, $i=\overline{1,M}$, each of which is an output of a layer with appropriate number if appropriate input element $x_0^{(i)}$, $i=\overline{1,M}$ from training sample;

- control u(k) with appropriate number $k=0,\ldots,N-1$ coincides with weight of a layer $k+1$ and is determined by equation: u(k) = w(k+1), k=0,…,N–1;
- functions $F=F(X(k),u(k),k)$, which describe recurrent connection between values of a phase variable, are matrix and consist from columns $(f(X(k), u(k), k))_i$, $i=\overline{1,M}$, which are determined by functions g of outputs of the appropriate layers of neural network according to equations:

$$f(X(k), u(k), k)_i=g(w(k+1)^T x^{(i)}(k),k+1),\ k=0,\ldots,N-1, \qquad (6)$$

where $x^{(i)}(k)$,$i=\overline{1,M}$ – an output of a layer with number k: $k=\overline{1,M}$, as reaction on element i of the training sample;
- functional $I(u_0,\ldots,u_{N-1})$ coincides with J(w(1),…,w(N)).

The evidence. The evidence will be the same as for previous result and given in [2].

Effect of the theorem 2 is the opportunity to use the theorem of a kind of gradients for the generalized beam of dynamics, for calculation of gradients of cost functional of training in a task of training of neural network. [2] It, actually, is result of the following theorem.

Theorem 3. Gradients of cost functional of training of neural network are determined by equations:

$$grad_{w(k)} J(w(1),\ldots,w(N)) = -grad_{w(k)} \sum_{i=1}^{M} H^{(i)}(x^{(i)}(k), w(k+1), p^{(i)}(k+1), k) \qquad (7)$$

k=1,...,N.

Theorem 3 forms a basis of Error Back Propagation algorithm for training of neural networks.

Conclusions

The adjustment of weights of neural networks method is described in the article. It is based on the theory of optimal control and representation of neural networks by beam of dynamics. Representation of neural networks by beam of dynamics allows to adjust weights of neural networks effectively and thus to solve a task of training.

Literature

1. Кириченко Н. Ф., Крак Ю. В., Полищук А.А. Псевдообратные и проекционные матрицы в задачах синтеза функциональных преобразователей.// Кибернетика и системный анализ –2004.–№3.
2. Кириченко Н.Ф., Донченко В.С., Сербаев Д.П. Нелинейные рекурсивные регрессионные преобразователи: динамические системы и оптимизация. // Кибернетика и системный анализ – 2005.–№1.
3. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals, and Systems, 2, 303–314.
4. Haykin, S. (1999). Neural networks: A comprehensive foundation (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
5. Kohonen, T. (1995). Self-organizing map. Heidelberg, Germany: Springer-Verlag.
6. Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. – К.: Высшая школа, 1975.–328 с.

Authors' Information

Volodymyr Donchenko – professor, Kiev National Taras Shevchenko University, Department of System Analysis and Decision Making Theory, e-mail: vsdon@unicyb.kiev.ua

Denis Serbaev – post graduate course student, Kiev National Taras Shevchenko University, Department of System Analysis and Decision Making Theory, e-mail: sdp@unicyb.kiev.ua