Simulation of Single-layer Neural Networks

https://doi.org/10.31713/MCIT.2019.70

Ibraim Didmanidze
Batumi Shota Rustaveli State University
Batumi, Georgia
ibraim.didmanidze@bsu.edu.ge

Guram Chachanidze
Batumi Shota Rustaveli State University
Batumi, Georgia
guramchachanidze@yahoo.com

Gregory Kakhiani
Batumi Shota Rustaveli State University
Batumi, Georgia
gkakhiani@gmail.com

Abstract— **The task of optimal management of artificial neuron network is represented. Showed that it is possible to use the Lagrange method with the help of the Lagrange new function for solving the task.**

Keywords— artificial neuron network; optimal management

I. Introduction

A criteria for the forecasting quality of the stock market should always take into account the deviation of predicted values from real data as well as the correspondence of the direction of the existing trend.

Let's consider a neuron network that consists of n neurons. Characterize i-th state of the neuron by a function $x_i(t), i=\overline{1,n}$, which dynamics is described by the following system of differential equations:

$$x_i'(t)=h_i(x(t))+\sum_{j=1}^{n}\left[w_{ij}^0+w_{ij}(t)\right]g_j(x(t)), i=\overline{1,n} \quad (1)$$

here $h_i(x), g_i(x)$ – are given as continuously differentiated functions on R^n , $x(t)$ is a continuous function, w_{ij}^0 and $w_{ij}(t)$ are the weight coefficients which characterize the effect of i-th neuron on j-th neuron, in the initial conditions $x_i(0)=a_i, i=\overline{1,n}$.

The differential equation system can be recorded in the following vector form:

$$x_i'(t)=h(x(t))+(W_0+W(t))g(x(t)). \quad (2)$$

With the following initial conditions:

$$x(0)=a, \quad (3)$$

here $a=(a_1,a_2,\dots,a_n)$; W_0 is a continuous $n\times n$ matrix which consists of w_{ij}^0 elements; the matrix of $n\times n$ dimension, composed of weight coefficients W(t)-$w_{ij}(t)$,which will subsequently be considered as management functions and which can be extended to the following limitations: $|w_{ij}(t)|\le A_{ij}, i,j=\overline{1,n}.$ Let's choose the optimal values of neuronal network weight coefficients according to the minimum condition of target function (4)

$$J(w(\cdot))=\Phi(x(T))+\int_0^T\left[Q(x(t)+\frac{1}{2}\sum_{i,j=1}^{n}r_{ij}w_{ij}^2(t)\right]dt \quad (4)$$

In this model the process time T > 0 is fixed, Q (x) and Φ (x) are scalar continuous differentiating functions, while $r_{ij}\ge 0, i,j=\overline{1,n}$ coefficients form a symmetric positive R matrix.

II. Main Part

The task of optimal management of artificial neuron network is represented by such a selection of $w_{i,j}(t),\ i,j=\overline{1,n}$ weight coefficients where the image (4) is taken to the minimum value. This task belongs to the class of optimal management task. Let's consider the weight coefficients as management functions and use the theorem about the necessary conditions of optimality [1].

Let's introduce the switching functions $\Phi_{ij}(t)$ taking into account the following rule:

$$\Phi_{ij}(t)=p_i(t)g_j(x(t))r_{ij}^{-1},\ i,j=\overline{1,n}. \quad (5)$$

Consider just a regular case to build $\overline{w_{ij}}(t)$ optimal management:

$$\bar{w}_{ij}(t)=\begin{cases}A_{ij}, \Phi_{ij}(t)>A_{ij},\\ \Phi_{ij}, |\Phi_{ij}|\le A_{ij}\\ -A_{ij}, \Phi_{ij}(t)<-A_{ij}\end{cases} \quad (6)$$

The coupling functions $p_i(t), i,j=\overline{1,n}$ satisfy the following system of differential equations:

$$p_i'(t)=\frac{\partial Q(x(t))}{\partial x_i}-\sum_{i=1}^{n}p_i(t)\left(\sum_{j=1}^{n}\left(w_{ij}^0-w_{ij}(t)\right)\frac{\partial g_i(x(t))}{\partial x_i}+\frac{\partial h_i(x(t))}{\partial x_i}\right);$$
$$i=\overline{1,n}, \quad (7)$$

They also meet the following boundary conditions

$$p_i(T)=-\frac{\partial\Phi(x(T))}{\partial x_i}. \quad (8)$$

In the process of studying artificial neuron networks the following task often appears: how to build such a management, as a result of which the neuter network vector x (T) would take a predetermined value θ, i.e. x (T) = θ. Such a situation should lead to the necessity of solving the boundary problem of (1), (3), (6) - (8) maximization. For this purpose, it is necessary to determine the terminal compounds as follows:

$$\Phi\big(x(T)\big) = -\alpha_1\big(\theta, x(T)\big) + \alpha_2(|x(T)-\theta|^2), \quad (9)$$

where α_1 and α_2 are selectively increased by the correspondence of $x(T)$ and θ vectors' directions.

Let's analyze the case when it is necessary to store one n-dimensional θ vector in associative memory. This vector is called from the memory when the current state of the neuron network is close to θ [2].

For the case under consideration let's suppose that $h_i(x) = -\beta_i x_i$, where $\beta_i > 0, i = \overline{1,n}$ – are the given numbers, $g_j(x(t)) = x_j(t)$, $Q = O$, $W_0 = O$, $R = I$, $\Phi(x) = -(\theta, x)$, here O, I are correspondingly $n \times n$ dimensional zero and single matrix. There the initial (1) - (4) task will take the following form:

$$J\big(w(\cdot)\big) = -\big(\theta, x(T)\big) + \frac{1}{2}\int_0^T \sum_{i,j=1}^n w_{ij}^2(t)dt; \quad (10)$$

$$x_i'(t) = -\beta_i x_i(t) + \sum_{j=1}^n w_{ij}(t)x_j(t),\ i = \overline{1,n}; \quad (11)$$

$$|w_{ij}(t)| \le A_{ij},\ i,j = \overline{1,n}, w_{ij} = 0 \quad (12)$$

$$x_i(0) = a_i, i = \overline{1,n}.$$

In this task the image (θ, x (T)) represents the correlation between the vector x (t) and the desired vector θ of the final position of the neural network. We should note that minimizing the negative value of the correlation is equivalent to the task of correlation maximizing.

In this case, the optimal values of weight coefficients are determined by the same (6) formula, but here we should take into account the fact that the switching function $\Phi_{ij}(t), i,j = \overline{1,n}$ has the following form:

$$\Phi_{ox}(t) = p_i(t)x_j(t), i \ne j, \quad (13)$$

and the coupling functions $p_i(T) = \theta, i = \overline{1,n}$ satisfy the following system of differential equations

$$p_i'(t) = -\sum_{i=0}^n p_i(t)w_{ij}(t) + \beta_i p_i(t), i = \overline{1,n}, \quad (14)$$

with the following boundary conditions

$$p_i(T) = \theta_i, p_i(t) = 0, t > T, i = \overline{1,n}. \quad (15)$$

A a discrete model (i.e. a model that describes the artificial neuron network state in the fixed moments of time [3,4]) of the neuron network can be constructed by analogy of the continuous model of artificial neuron network.

Let's consider the discrete task of optimal management. It is necessary to minimize the following function

$$I([u]) = \sum_{k=0}^{q-1} f_k^0\big(x^k, u^k\big) + \Phi\big(x^{ij}\big) \quad (16)$$

with the restrictions on the following recurring correlations

$$x^{k+1} = f_k\big(x^k, u^k\big), k = \overline{0, q-1} \quad (17)$$

$$x^0 = b^0 \in R, \quad (18)$$

Let's introduce the following designations:

$$z^k = (x^k, u^k);$$

$$[u] = [u^0, \cdots, u^{q-1}],$$

$$[\tilde{u}] = [u^0, \cdots u^{i-1}, u^i + \Delta, u^{i+1}, \cdots, u^{q-1}]; \quad (19)$$

$$[x] = [x^0, \cdots, x^{q-1}], [\tilde{x}] = [x^0, \cdots, x^{i-1}, x^i + \Delta, x^{i+1}, \cdots, x^{q-1}]$$

Let's determine a coupling vector

$$p^i = \frac{dI([x],[u])}{dx^t} = \lim_{\Delta\to 0}\frac{I([\tilde{x}],[\tilde{u}]) - I([x],[u])}{\Delta}. \quad (20)$$

if $i = q$ then from (20) it follows that

$$p^q = \frac{dI([x],[u])}{dx^q} = \frac{\partial\Phi(x^q)}{\partial x^q} \quad (21)$$

if $i = q-1$ then

$$p^{q-1} = \frac{\partial I}{\partial x^{q-1}} + \left(\frac{\partial x^q}{\partial x^{q-1}}\right)^q \frac{dI}{dx^{q-1}} = \frac{\partial}{\partial x^{q-1}}\left[f_{q-1}^0(z^{q-1}) + \big(p^q, f_{q-1}(z^{q-1})\big)\right] \quad (22)$$

Generally, in (21) and (22) we will similarly get the following

$$p^i = \frac{\partial I}{\partial x^i} + \left(\frac{\partial x^{i+1}}{\partial x^i}\right)^T \frac{dI}{dx^i} = \frac{\partial f_i^0(z^i)}{\partial x^i} + \left[\frac{\partial f_i(z^t)}{\partial x^i}\right]^i p^{i+1}, \quad (23)$$

$$i = \overline{1, q-1}, p^0 = 0, j > q.$$

Let's introduce a function $H_k\big(z^k, p^{k+1}\big) = \lambda_0 f_k^0\big(z^k\big) + \big(p^{k+1}, f_k\big(z^k\big)\big)$, then (23) can be rewritten in the following way

$$p^i = \frac{\partial H_i(z^i, p^{k+1})}{\partial x^i}, i = \overline{q-1, 1}; \quad (24)$$

$$p^q = \frac{\partial\Phi(x^q)}{\partial x^q}.$$

A set of vectors $[x] = [x^1, \cdots, x^q]$ depends on the set of initial vector x^0 of the given condition and the set of admissible vectors $[u] = [u^1, \cdots, u^q]$, so we can write $[x] = [x(u)]$ and minimize the function $I = I([x(u)], [u])$. Let's determine $\frac{dI}{du^i}$

$$\frac{dI([x(u)],[u])}{du^i} = \frac{\partial I([x(u)],[u])}{\partial u^i} + \frac{\partial x^{i+1}[u]}{\partial u^i}\cdot\frac{dI([x(u)],[u])}{dx^{i+1}} = \frac{\partial f_i^0(z^i)}{\partial u^i} + \left[\frac{\partial f_i(z^i)}{\partial u^i}\right]^T p^{i+1} = \frac{\partial}{\partial u^i} H_i\big(z^i, p^{i+1}\big). \quad (25)$$

The Lagrange function for the optimal management task, described in formulas (16)-(19), can be recorded as follows:

$$L([x],[u],[p],\lambda_0) = \sum_{i=0}^{q-1}\left[\lambda_0 f_i^0(z^i) + (p^{i+1}, f_i(z^i) - x^{i+1}\right] + \lambda_0\Phi(x^q).$$

In order to get the required conditions of optimality, it is necessary to calculate the derivatives with respect to variables of the Lagrange function $x^i, i = \overline{1,q}, u^i, i = \overline{1, q-1}$

$$\frac{\partial L}{\partial x^i} = \frac{\partial H_i(z^i, p^{i+1})}{\partial x^i} - p^i = 0, i = \overline{1, q-1},$$

$$\frac{\partial L}{\partial x^q} = \lambda_0\frac{\partial\Phi(x^q)}{\partial x^q} - p^q = 0,$$

$$\frac{\partial L}{\partial u^i} = \frac{\partial H_i(z^i, p^{i+1})}{\partial u^i} = 0, i = \overline{1, q-1}. \quad (26)$$

Comparing formulas (25) and (26) we will get:

$$\frac{\partial L}{\partial u^i} = \frac{\partial I([x(u)],[u])}{\partial u^i}, i = \overline{1, q-1}. \quad (27)$$

According to the image (20) the Langrand's multipliers $p^i, i = \overline{1,q}$, have a value of derivatives with respect to variables of the minimized function $x^i, i = \overline{1,q}$.

The formula (27) can be used for an approximate solution, for example, by graded methods if a task (16)-(19) contains

the restrictions of management vector $u^i \in U_i, i = \overline{0, q-1}$.
For instance:

$$U_i = \{v \in R^r : g_l^i(v) \leq 0, l = \overline{1,k}, h_l^i(v) = 0, l = \overline{k+1, s}\} \tag{28}$$

III. CONCLUSION

It is possible to use the Lagrange method with the help of the Lagrange new function. In this case the Lagrange new function $\tilde{L}$ is determined for the tasks (16) - (19) in the following way:

$$\tilde{L} = L + \sum_{i=0}^{q-1} \left[\sum_{l=0}^{k} \mu_l^i g_l^i(v) + \sum_{l=k+1}^{s} v_l^i h_l^i(v) \right].$$

The tasks (16) - (19) can be also solved by the method of penalty functions.

The discrete task can also be derived from the continuous task if we change the derivation according to the Eiler scheme and the integral, for example, by the rule of the right rectangles.

REFERENCES

[1] Hornik K., Stinchcombe M . and White H . Multilayer feedforward networks are universal approximators//Neum\ Networks. - 1989.-N.2 - P.359-366 .

[2] Didmanidze Ibraim, Kakhiani Grigol, Didmanidze Didar. Multilayer neural network. XXIV international conference PROBLEMS OF DECISION MAKING UNDER UNCERTAINTIES (PDMU-2014). ABSTRACTS. September 1-5, 2014. Cesky Rudolec, Czeck Republic, 2014. P. 32-33.

[3] Didmanidze Ibraim, Kakhiani Gregory, Motskobili Ia. Creation of mathematical model of the analysis of the financial market. The VII International Academic Congress “Innovation in the Modern World”. PAPERS AND COMMENTARIES, VOLUME III. Australia, Sydney, 18-20 May 2015. “Sydney University Press”, 2015. p/ 440-444.

[4] Didmanidze I., Kakhiani G., Dumbadze Z. On the Issue Of Choosing the Neuron Adaptive Educational Speed In the Last Hidden Layer Of a Neural Network. XXX international conference PROBLEMS OF DECISION MAKING UNDER UNCERTAINTIES (PDMU-2015). ABSTRACTS. August 14-19, 2017, Vilnius, Lithuania. p. 41.