

ЕЛЕНА АРКАДЬЕВНА АНДРЕЕВА

доктор физико-математических наук, профессор, заведующий кафедрой компьютерной безопасности и математических методов управления математического факультета, Тверской государственный университет (Тверь, Российская Федерация)

andreeva.tvgu@yandex.ru

ИРИНА СЕРГЕЕВНА МАЗУРОВА

аспирант кафедры компьютерной безопасности и математических методов управления математического факультета, Тверской государственный университет (Тверь, Российская Федерация)

IrinaSMazurova@gmail.com

## ЧИСЛЕННЫЕ МЕТОДЫ ОБУЧЕНИЯ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ\*

Решается задача моделирования и обучения искусственной нейронной сети достаточно общей топологии, динамика которой описывается системой интегро-дифференциальных уравнений. Задача обучения искусственной нейронной сети рассматривается как задача оптимального управления с нефиксированным временем. Сформулированы необходимые условия оптимальности весовых коэффициентов искусственной нейронной сети и внешнего управляющего воздействия с учетом заданных ограничений и вида функционала. Исходная задача сведена к дискретной задаче оптимального управления, которая решается на основе методологии быстрого автоматического дифференцирования. Проведен анализ влияния параметров задачи на оптимальное решение.

Ключевые слова: искусственные нейронные сети, интегро-дифференциальные уравнения, оптимальное управление

### ВВЕДЕНИЕ

В последние десятилетия большое внимание уделяется исследованию искусственных нейронных сетей, которые эффективно используются в различных системах управления, медицине и экономике. В статье рассмотрена проблема моделирования и обучения искусственных нейронных сетей с достаточно общей топологией, описываемой системой интегро-дифференциальных уравнений. При этом целью обучения является минимизация ошибки обучения, энергии сети и времени обучения, то есть решается многокритериальная задача. Для решения этой задачи используется аппарат математической теории оптимального управления, с помощью которого находятся оптимальные значения весовых коэффициентов дискретной нейронной сети. Заметим, что таким способом может быть решен большой класс задач, включающих создание ассоциативной памяти, модель распознавания образов, прогнозирования и т. д.

### ПОСТАНОВКА ЗАДАЧИ

Рассмотрим математическую модель нейронной сети, в которой динамика отдельного нейрона описывается на основе биологических предположений [1], [5]. Искусственная нейронная сеть состоит из нейронов, каждый из которых взаимодействует со всеми остальными нейронами. Воздействие  $j$ -го нейрона на  $i$ -й нейрон в момент времени  $t$  характеризуется весовым коэффициентом  $\omega_{ij}(t)$ . Пусть  $x_i(t)$  – функция, характеризу-

ющая состояние  $i$ -го нейрона в момент времени  $t$ ,  $i = \overline{1, n}$ . Динамика искусственной нейронной сети описывается системой интегро-дифференциальных уравнений:

$$\dot{x}_i(t) = -\beta_i x_i(t) + f_i \left( \int_0^t \sum_{j=1}^n \omega_{ij}(z) x_j(z) dz \right) + \gamma_i u_i(t), \quad (1)$$

$$t \in [0, T]$$

с начальными условиями  $x_i(0) = \theta_i$ ,  $i = \overline{1, n}$ , (2)

где  $\beta_i$  – параметр, определяющий скорость затухания  $i$ -го нейрона при отсутствии внешнего воздействия,  $f_i$  – функция активации, при этом  $\gamma_i u_i(t)$  – внешнее воздействие на нейрон в момент времени  $t$ . Коэффициенты  $\varepsilon_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $i, j = \overline{1, n}$ , – заданные неотрицательные параметры модели.

Весовые коэффициенты нейронной сети  $\omega_{ij}(t)$  и внешние воздействия  $u_i(t)$  являются функциями управления и ограничены

$$|\omega_{ij}(t)| \leq a_{ij}, \quad |u_i(t)| \leq a_i, \quad i, j = \overline{1, n}, \quad \text{п. в. } t \in [0, T] \quad (3)$$

где  $a_{ij}, a_i$ ,  $i, j = \overline{1, n}$ , заданные положительные параметры модели, полагаем  $\omega_{ii}(t) \equiv 0$ .

Задача оптимального управления заключается в минимизации функционала:

$$J(\omega, u) = \int_0^T \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} \omega_{ij}^2(t) + \sigma_i u_i^2(t) \right) dt + \sum_{i=1}^n M_i (x_i(T) - A_i)^2 + T, \quad (4)$$

где первое слагаемое характеризует энергию рассматриваемой нейронной сети, зависящую

от текущей конфигурации сети,  $\alpha_{ij}$  – параметр, определяющий меру влияния значения весового коэффициента между нейронами  $i$  и  $j$  и на минимизируемый функционал, второе слагаемое характеризует ошибку обучения нейронной сети,  $M_i$  – параметр, определяющий меру воздействия этого слагаемого на минимизируемый функционал,  $A$  – целевой вектор.

Задача оптимального управления заключается в том, чтобы обучить нейронную сеть таким образом, чтобы минимизировать функционал (4) и построить оптимальные значения весовых коэффициентов, внешнего управляющего воздействия. Задача оптимального управления с нефиксированным временем процесса была рассмотрена в работе [2] для модели искусственной нейронной сети, описываемой системой дифференциальных уравнений с запаздывающим аргументом.

### НЕОБХОДИМЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ

Особенностью поставленной задачи является большая размерность управляющих функций, наличие распределенного запаздывания и нефиксированное время процесса.

Для решения задачи (1)–(4) перейдем от задачи оптимального управления с нефиксированным временем процесса к задаче с фиксированным временем процесса. Для этого осуществим следующую параметризацию:

$$t(\tau) = \xi\tau, \quad \tau \in [0, T_0], \quad t(T_0) = \xi T_0 = T$$

и введем обозначения:

$$x(t(\tau)) = \tilde{x}(\tau), \quad \omega(t(\tau)) = \tilde{\omega}(\tau), \quad u(t(\tau)) = \tilde{u}(\tau).$$

В этих обозначениях задача оптимального управления примет вид:

$$J(\tilde{x}(\tau), \tilde{\omega}(\tau), \tilde{u}(\tau), t(\tau)) = \int_0^{T_0} \xi \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}(\tau))^2 + \sigma_i \tilde{u}_i(\tau)^2 \right) d\tau + \sum_{i=1}^n M_i (\tilde{x}_i(T_0) - A_i)^2 + t(T_0), \quad (5)$$

$$\dot{\tilde{x}}_i(\tau) = -\xi \beta_i \tilde{x}_i(\tau) + \xi f_i \left( \xi \int_0^\tau \sum_{j=1}^n \tilde{\omega}_{ij}(\tilde{z}) \tilde{x}_j(\tilde{z}) d\tilde{z} \right) + \xi \gamma_i \tilde{u}_i(\tau), \quad i = \overline{1, n}, \quad \dot{t}(\tau) = \xi.$$

Введем новые переменные:

$$\tilde{y}_i(\tau) = \xi \int_0^\tau \sum_{j=1}^n \tilde{\omega}_{ij}(\tilde{z}) \tilde{x}_j(\tilde{z}) d\tilde{z}, \quad i = \overline{1, n},$$

такие что

$$\dot{\tilde{y}}_i(\tau) = \xi \sum_{j=1}^n \tilde{\omega}_{ij}(\tau) \tilde{x}_j(\tau), \quad i = \overline{1, n}, \quad \tilde{y}_i(0) = 0.$$

В этих переменных исходная система примет вид (6)–(8):

$$\dot{\tilde{x}}_i(\tau) = -\xi \beta_i \tilde{x}_i(\tau) + \xi f_i(\tilde{y}_i(\tau)) + \xi \gamma_i \tilde{u}_i(\tau), \quad i = \overline{1, n}, \quad (6)$$

$$\dot{\tilde{y}}_i(\tau) = \xi \sum_{j=1}^n \tilde{\omega}_{ij}(\tau) \tilde{x}_j(\tau), \quad i = \overline{1, n}, \quad (7)$$

$$\dot{t}(\tau) = \xi. \quad (8)$$

Построим функцию Понтрягина:

$$H(t, \tilde{x}, \tilde{y}, \tilde{\omega}, \tilde{u}, \xi, r, p, s, \lambda_0) = -\lambda_0 \xi \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}(\tau))^2 + \sigma_i \tilde{u}_i(\tau)^2 \right) + \sum_{i=1}^n p_i(\tau) (-\xi \beta_i \tilde{x}_i(\tau) + \xi f_i(\tilde{y}_i(\tau)) + \xi \gamma_i \tilde{u}_i(\tau)) + \sum_{i=1}^n r_i(\tau) \left( \xi \sum_{j=1}^n \tilde{\omega}_{ij}(\tau) \tilde{x}_j(\tau) \right) + s(\tau) \xi. \quad (9)$$

Теорема. Пусть  $\tilde{\omega}_{ij}(\tau)$ ,  $i, j = \overline{1, n}$  – оптимальные весовые коэффициенты,  $\tilde{u}_i(\tau)$ ,  $i = \overline{1, n}$  – оптимальные управляющие внешние воздействия,  $\xi$  – оптимальное значение параметра,  $\alpha_{ij}$  и  $\sigma_i$  отличны от нуля, тогда

$$\tilde{\omega}_{ij}(\tau) = \begin{cases} a_{ij}, & \Phi_{ij}(\tau) > a_{ij} \\ \Phi_{ij}(\tau), & |\Phi_{ij}(\tau)| \leq a_{ij}, \quad i, j = \overline{1, n}, \\ -a_{ij}, & \Phi_{ij}(\tau) < -a_{ij} \end{cases} \quad (10)$$

где

$$\Phi_{ij}(\tau) = \frac{r_i(\tau) \tilde{x}_j(\tau)}{2\lambda_0 \alpha_{ij}}, \quad i, j = \overline{1, n},$$

$$\tilde{u}_i(\tau) = \begin{cases} a_i, & \Psi_i(\tau) > a_i \\ \Psi_i(\tau), & |\Psi_i(\tau)| \leq a_i, \quad i = \overline{1, n}, \\ -a_i, & \Psi_i(\tau) < -a_i \end{cases} \quad (11)$$

где

$$\Psi_i(\tau) = \frac{p_i(\tau) \gamma_i}{2\lambda_0 \sigma_i}, \quad i = \overline{1, n},$$

$$\frac{\partial H}{\partial \xi} = -\lambda_0 \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}(\tau))^2 + \sigma_i \tilde{u}_i(\tau)^2 \right) + \sum_{i=1}^n p_i(\tau) (-\beta_i \tilde{x}_i(\tau) + f_i(\tilde{y}_i(\tau)) + \gamma_i \tilde{u}_i(\tau)) + \sum_{i=1}^n r_i(\tau) \left( \sum_{j=1}^n \tilde{\omega}_{ij}(\tau) \tilde{x}_j(\tau) \right) + s(\tau) = 0, \quad (12)$$

а система дифференциальных уравнений для сопряженных функций имеет вид (13):

$$\dot{p}_i(\tau) = \xi p_i(\tau) \beta_i - \xi \sum_{j=1}^n r_j(\tau) \tilde{\omega}_{ji}(\tau), \quad i = \overline{1, n} \quad (13)$$

$$\dot{r}_i(\tau) = -\xi p_i(\tau) \frac{\partial f(\tilde{y}_i)}{\partial \tilde{y}_i}, \quad \dot{s}(\tau) = 0,$$

с условием трансверсальности на правом конце (14):

$$p_i(T_0) = -2\lambda_0 M_i (\tilde{x}_i(T_0) - A_i), \quad i = \overline{1, n} \\ r_i(T_0) = 0, \quad i = \overline{1, n}, \quad s(T_0) = -\lambda_0. \quad (14)$$

### ДИСКРЕТНАЯ ЗАДАЧА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Для определения оптимального решения рассмотрим дискретную аппроксимацию начальной задачи, в которой используется правило левых

прямоугольников для аппроксимации интеграла и схема Эйлера аппроксимации производных. Вводя следующие обозначения  $\Delta\tau = T_0 / q$ ,  $\tau_k = k\Delta\tau$ ,  $t^q = T_0$ ,  $\tilde{x}_i^k = \tilde{x}_i(\tau_k)$ ,  $\tilde{y}_i^k = \tilde{y}_i(\tau_k)$ ,  $\tilde{\omega}_{ij}^k = \tilde{\omega}_{ij}(\tau_k)$ ,  $\tilde{u}_i^k = \tilde{u}_i(\tau_k)$ ,  $i, j = \overline{1, n}$ ,  $k = \overline{0, q-1}$ , построим дискретную задачу оптимального управления, которая состоит в минимизации функции:

$$I(t, \tilde{x}, \tilde{y}, \tilde{\omega}, \tilde{u}, \xi) = \Delta\tau \xi \sum_{k=1}^{q-1} \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}^k)^2 + \sigma_i \tilde{u}_i^{k2} \right) + \sum_{i=1}^n M_i (\tilde{x}_i^q - A_i)^2 + t^q, \quad (15)$$

при ограничениях

$$\tilde{x}_i^{k+1} = \tilde{x}_i^k - \Delta\tau \xi \beta_i \tilde{x}_i^k + \Delta\tau \xi f_i(\tilde{y}_i^k) + \Delta\tau \xi \gamma_i \tilde{u}_i^k, \quad (16)$$

$$\tilde{y}_i^{k+1} = \tilde{y}_i^k + \Delta\tau \xi \sum_{j=1}^n \tilde{\omega}_{ij}^k \tilde{x}_j^k,$$

$$t^{k+1} = t^k + \Delta\tau \xi,$$

$$i = \overline{1, n}, k = \overline{0, q-1}.$$

На функции управления наложены ограничения

$$|\omega_{ij}^k| \leq a_{ij}, \omega_{ii}^k \equiv 0, |u_i^k| \leq a_i, \quad (17)$$

$$i, j = \overline{1, n}, k = \overline{0, q-1}.$$

Для решения задачи (15)–(17) в работе используется методология быстрого автоматического дифференцирования, разработанная в ВЦ РАН под руководством Ю. Г. Евтушенко, позволяющая с единых позиций определять градиенты для явно и неявно определенных функций и для вычислительных процессов, которые являются результатом дискретизации непрерывных систем, описываемых дифференциальными и интегродифференциальными уравнениями [4]. В соответствии с методологией быстрого автоматического дифференцирования введем множители Лагранжа  $p$ ,  $r$  и  $s$ , тогда функция Лагранжа для данной задачи может быть представлена следующим выражением

$$L(t, \tilde{x}, \tilde{y}, \tilde{\omega}, \tilde{u}, \xi, r, p, s, \lambda_0) = \lambda_0 \Delta\tau \xi \sum_{k=0}^{q-1} \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}^k)^2 + \sigma_i \tilde{u}_i^{k2} \right) + \lambda_0 t^q + \sum_{i=1}^n M_i (\tilde{x}_i^q - A_i)^2 + \sum_{k=0}^{q-1} \sum_{i=1}^n p_i^{k+1} (\tilde{x}_i^{k+1} - \tilde{x}_i^k - \Delta\tau \xi \beta_i \tilde{x}_i^k + \Delta\tau \xi f_i(\tilde{y}_i^k) - \Delta\tau \xi \gamma_i \tilde{u}_i^k) + \sum_{k=0}^{q-1} \sum_{i=1}^n r_i^{k+1} (\tilde{y}_i^{k+1} - \tilde{y}_i^k - \Delta\tau \xi \sum_{j=1}^n \tilde{\omega}_{ij}^k \tilde{x}_j^k) + \sum_{k=0}^{q-1} s^{k+1} (t^{k+1} - t^k - \Delta\tau \xi).$$

Градиент минимизируемой функции

$$\frac{\partial L}{\partial \tilde{u}_i^m} = 2\lambda_0 \Delta\tau \xi \sigma_i \tilde{u}_i^m - \Delta\tau \xi \gamma_i p_i^{m+1}, \quad (18)$$

$$\frac{\partial L}{\partial \tilde{\omega}_{ij}^m} = 2\lambda_0 \Delta\tau \xi \alpha_{ij} \tilde{\omega}_{ij}^m - \Delta\tau \xi r_i^{m+1} \tilde{x}_j^m, \quad (19)$$

$$\frac{\partial L}{\partial \xi} = \lambda_0 \Delta\tau \sum_{k=0}^{q-1} \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}^k)^2 + \sigma_i \tilde{u}_i^{k2} \right) + \sum_{k=0}^{q-1} \sum_{i=1}^n p_i^{k+1} (\Delta\tau \xi \beta_i \tilde{x}_i^k - \Delta\tau \xi f_i(\tilde{y}_i^k) - \Delta\tau \xi \gamma_i \tilde{u}_i^k) + \sum_{k=0}^{q-1} \sum_{i=1}^n r_i^{k+1} (\Delta\tau \xi \sum_{j=1}^n \tilde{\omega}_{ij}^k \tilde{x}_j^k) - \sum_{k=0}^{q-1} s^{k+1} \Delta\tau \xi, \quad (20)$$

где

$$p_i^m = p_i^{m+1} - \Delta\tau \xi p_i^{m+1} \beta_i + \Delta\tau \xi \sum_{j=1}^n r_j^{m+1} \tilde{\omega}_{ji}^m, \quad (21)$$

$$p_i^q = -2\lambda_0 M_i (\tilde{x}_i^q - A_i), \quad (22)$$

$$r_i^m = \Delta\tau \xi p_i^{m+1} \frac{\partial f_i(\tilde{y}_i^m)}{\partial \tilde{y}_i^m} + r_i^{m+1}, \quad r_i^q = 0, \quad (23)$$

$$s^m = s^{m+1}, \quad s^q = -\lambda_0. \quad (24)$$

На основе быстрого автоматического дифференцирования разработан комплекс программ построения оптимального управления и проведен численный эксперимент, результаты которого представлены ниже. Заметим, что основное вычислительное время при расчете градиента по формулам (18)–(20) требуется на нахождение векторов  $p$  и  $r$ , для этого необходимо решить 2 системы уравнений с  $q \cdot (n^2 + n)$  неизвестными.

Проведено исследование зависимости результатов работы метода от параметров задачи.

На рис. 1–2 представлены функции, соответствующие оптимальному состоянию нейронной сети,  $\tilde{x}_i(t)$ ,  $i = \overline{1, 3}$ , и оптимальные весовые коэффициенты,  $\tilde{\omega}_{ij}(t)$ ,  $i, j = \overline{1, 3}$ . Целью работы нейронной сети является перевод ее из состояния  $x_0 = (2, 1, 0)$  в состояние  $A = (1, 0, 5)$ .

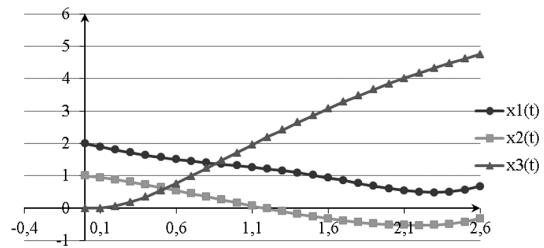


Рис. 1. Функции состояния нейронной сети  $\tilde{x}_i(t)$ ,  $i = \overline{1, 3}$

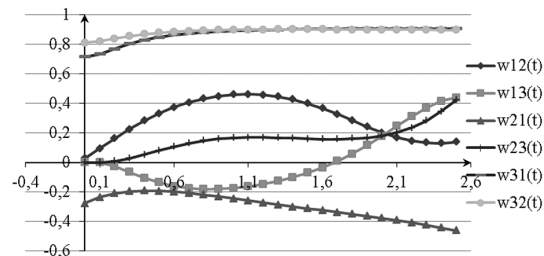


Рис. 2. Весовые коэффициенты нейронной сети  $\tilde{\omega}_{ij}(t)$ ,  $i, j = \overline{1, 3}$

Параметры режима, представленного на рис. 1–2, следующие:  $n = 3$ ,  $T_0 = 5$ ,  $\Delta\tau = 0,1$ ,  $|u_i^k| \leq 0,1$ ,  $\beta_i = 0,5$ ,  $|\omega_{ij}^k| \leq 1$ ,  $\alpha_i = 0,1$ ,  $M_i = 10000$ , точность метода  $\varepsilon = 0,000001$ . За 101841 итерацию значение минимизируемого функционала достигло величины 0,5176, что означает достаточно высокую эффективность метода, при этом оптимальное время процесса  $T = 2,8$ .

В таблице приведены значения минимизируемого функционала при различном значении времени процесса

Значения минимизируемого функционала при различном значении времени процесса

| T          | 2      | 2,2   | 2,4                 | 2,6                 | 2,8                 | 3                   | 3,2                 |
|------------|--------|-------|---------------------|---------------------|---------------------|---------------------|---------------------|
| $I_{opt1}$ | 2002,4 | 7,203 | 0,559               | 0,558               | 0,518               | 0,686               | 0,723               |
| $I_{opt1}$ | 0,9402 | 1,034 | 0,559               | 0,558               | 0,518               | 0,686               | 0,723               |
| $I_{opt2}$ | 2001,5 | 6,169 | $9,9 \cdot 10^{-5}$ | $2,1 \cdot 10^{-5}$ | $7,3 \cdot 10^{-6}$ | $2,8 \cdot 10^{-6}$ | $2,4 \cdot 10^{-6}$ |

$$\text{где } I_{opt1} = \Delta\tau \sum_{k=1}^{q-1} \sum_{i=1}^n \left( \sum_{j=1}^n \alpha_{ij} (\tilde{\omega}_{ij}^k)^2 + \sigma_i \tilde{u}_i^{k2} \right),$$

$$I_{opt2} = \sum_{i=1}^n M_i (\tilde{x}_i^q - A_i)^2, \quad I_{opt} = I_{opt1} + I_{opt2}.$$

Из таблицы следует, что найденное оптимальное время процесса  $T = 2,8$  соответствует наименьшему значению минимизируемого функционала, то есть за время работы нейронной сети

$T = 2,8$  значение величины ошибки обучения и энергии нейронной сети достигает наименьших значений, при уменьшении или увеличении времени процесса увеличивается значение минимизируемого функционала.

## ЗАКЛЮЧЕНИЕ

Аналогичный подход распространяется на оптимальное управление моделью искусственной нейронной сети, динамика которой описывается следующей системой интегро-дифференциальных уравнений:

$$\dot{x}_i(t) = -\beta_i x_i(t) + f_i \left( \int_{t-r}^t \sum_{j=1}^n \omega_{ij}(\tau) x_j(\tau - \delta_j) d\tau \right) + \gamma_i u_i(t),$$

$$t \in [0, T]$$

с начальными условиями

$$x_i(t) = \theta_i(t), \quad t \in [-r - \max(\delta_j), 0], \quad i = \overline{1, n},$$

где параметр  $\delta_j$ ,  $j = \overline{1, n}$  учитывает запаздывание при передаче сигнала от одного нейрона к другому, а параметр  $r$  в интегральном слагаемом в аргументе функции активации учитывает тот факт, что внешнее воздействие на нейроны ИНС накапливается на интервале запаздывания  $[t-r, t]$ . Решение задачи оптимального управления с фиксированным временем процесса для такой нейронной сети рассмотрено в работе [3].

\* Работа выполнена в соответствии с Программой поддержки ведущих научных школ (НШ-5264.2012.1).

## СПИСОК ЛИТЕРАТУРЫ

- Андреева Е. А. Оптимизация нейронных сетей: Учеб. пособие. Тверь, 2007.
- Андреева Е. А., Пустырякова Ю. А. Численные методы обучения искусственных нейронных сетей с запаздыванием // ЖВМ и МФ. 2002. Т. 42. С. 1383–1391.
- Большакова И. С., Шаронов Д. А. Обучение нейронной сети с запаздыванием // Программные продукты и системы. 2011. № 2. С. 35–37.
- Евтушенко Ю. Г. Оптимизация и быстрое автоматическое дифференцирование. М., 2013. 144 с.
- Майоров В. В., Мышкин И. Ю. Математическое моделирование нейронов сети на основе уравнений с запаздыванием // Математическое моделирование. 1990. Т. 2. № 11. С. 64–76.

Andreeva E. A., Tver State University (Tver, Russian Federation)  
Mazurova I. S., Tver State University (Tver, Russian Federation)

## NUMERICAL METHODS OF ARTIFICIAL NEURAL NETWORK TRAINING

The problem of artificial neural network modeling is studied. The artificial neural network is characterized by the enough general topology and is described by the system of integral differential equations. Training on the artificial neural network usage is considered as an optimal control problem with variable duration of the process. Considering the defined constraints and the form of the functional, the necessary conditions of optimality of the artificial neural network weighting coefficients incorporating external influence are formulated. The initial current problem is reduced to a discrete optimal control problem, which is solved by the methodology of the fast automatic differentiation. The influence of the task parameters on the optimal solution is analyzed.

Key words: artificial neural networks, integro-differential equations, optimal control

## REFERENCES

- Andreeva E. A. *Optimizatsiya neyronnykh setey* [Optimization of neural network: Educational book]. Tver, 2007.
- Andreeva E. A., Pustyrnakova Yu. A. Numerical methods for training artificial neural networks with delay [Chislennye metody obucheniya iskusstvennykh neyronnykh setey s zapazdyvaniem]. *ZhVM i MF*. 2002. Vol. 42. P. 1383–1391.
- Bolshakova I. S., Sharonov D. A. Training a neural network with delay [Obuchenie neyronnoy seti s zapazdyvaniem]. *Programmnye produkty i sistemy*. 2011. № 2. P. 35–37.
- Evtushenko Yu. G. *Optimizatsiya i bystroie avtomaticheskoe differentsirovanie* [Optimization and fast automation differentiation]. Moscow, 2013. 144 p.
- Mayorov V. V., Myshkin I. Yu. Mathematical modeling of neuronal networks based on equations with delay [Matematicheskoe modelirovanie neyronov seti na osnove uravneniy s zapazdyvaniem]. *Matematicheskoe modelirovanie* [Mathematical modeling]. 1990. Vol. 2. № 11. P. 64–76.

Поступила в редакцию 04.04.2014