УДК 519.97, 519.6, 007.681.5

Е.А. Андреева, В.М. Цирулева

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ УПРАВЛЕНИЯ ДИНАМИЧЕСКОЙ НЕЙРОННОЙ СЕТЬЮ С ЗАПАЗДЫВАНИЕМ

*Тверской государственный университет,
Тверь, Россия*

В настоящее время в мире активно развивается новая прикладная область математики, связанная с исследованием искусственных нейронных сетей. Интерес к ним вызван как теоретическими, так и прикладными достижениями: открылись возможности использования вычислений в сферах, до этого относящихся лишь к области человеческого интеллекта. Актуальность исследований в этом направлении подтверждается многочисленными примерами использования нейронных сетей в системах автоматизации [1], робототехнике процессов распознавания образов [2], адаптивном управлении [3], прогнозировании и создании экспертных систем [4], исследовании ассоциативной памяти [5] и др. В сложных практических задачах обученная нейронная сеть выступает как эксперт. Примером служит медицинская диагностика, где нейронная сеть может учитывать большое количество числовых параметров (электрические импульсы нервных клеток головного мозга и его отделов, фиксируемые с помощью энцефалограмм, давление, вес и т.д.). Целью работы является построение искусственной осцилляторной нейронной сети, которая может применяться при моделировании деятельности мозга: ассоциативной памяти и внимания. Модель формализуется, как многокритериальная задача оптимального управления с запаздывание. Целью управления нейронной сетью является ее обучение, которое включает в себя построение оптимального процесса, удовлетворяющего заданным критериям. Одним из критериев является терминальный критерий, определяющий состояние нейронной сети в конечный момент времени. Условия оптимальности в непрерывной модели получены с помощью принципа максимума для задач с запаздывающим аргументом [6],[7],[8]. Построена краевая задач принципа максимума [9]. Для получения условий оптимальности в дискретной модели, аппроксимирующей непрерывную, используются метод быстрого автоматического дифференцирования и численные методы решения экстремальных задач [9], [10],[11]. Приводятся результаты численного эксперимента.

Ключевые слова: оптимальное управление, осцилляторная нейронная сеть, ансамбль нейронов, математическая модель, многокритериальная задача, принцип максимума с запаздывающим аргументом, дискретная задача оптимального управления.

ВВЕДЕНИЕ. Искусственные нейронные сети индуцированы биологией. Они состоят из элементов, функциональные возможности которых аналогичны большинству элементарных функций биологического нейрона головного мозга. Рассматривая проблемы искусственного интеллекта, естественно исследовать и сопоставить моделируемые процессы человеческого мозга структуре и функциям реальных нейронов, которые связаны с процессом познания, с естественным интеллектом.

Динамические нейронные сети представляют собой сети, построенные из нейронов, состояние которых описывается системой дифференциальных уравнений. Организация сети такова, что каждый нейрон получает сигналы от ансамбля нейронов. С помощью искусственных нейронных сетей можно моделировать работу нелинейных динамических систем. Такие сети можно использовать при моделировании ассоциативной памяти, внимания, задач управления сложными системами, распознавании и классификации образов, нелинейной обработке сигналов и др. Следует отметить, что вследствие широкого спектра различных применений невозможно создать универсальную модель нейронной сети. Каждая такая модель имеет кроме ряда преимуществ и очевидные ограничения, которые необходимо учитывать в практических реализациях.

В статье исследуется динамическая осцилляторная нейронная сеть, которая может применяться при моделировании ассоциативной памяти и внимания. Рассматриваемая искусственная нейронная сеть формализуется как задача оптимального управления динамической системой с запаздыванием и исследуется с помощью методов математической теории оптимального управления и методов оптимизации. Целью управления искусственной нейронной сетью является ее обучение и построение оптимального управления ансамблем нейронов для заданного целевого функционала в многоцелевой задаче оптимального управления. Для построения численного решения применяются метод быстрого автоматического дифференцирования и численные методы оптимизации. В работе получено оптимальное управление системой в зависимости от целевого функционала и реализован численный алгоритм, позволяющий учесть многокритериальность поставленной задачи и исследовать влияние параметров задачи на оптимальное управление и зависимость управления от общей структуры динамической нейронной сети.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ УПРАВЛЕНИЯ ОСЦИЛЛЯТОРНОЙ НЕЙРОННОЙ СЕТЬЮ. Рассматривается задача оптимального управления для динамической системы с запаздыванием, моделирующей динамику искусственной нейронной сети. Нейронная сеть, состоящая из n нейронов, описывается нелинейной системой дифференциальных уравнений второго порядка с запаздыванием:

$$\ddot{x}_i(t) + \varepsilon\left(1 - \beta x_i^2(t)\right) + \omega_i^2 x_i(t) = u(t)\sum_{i=1}^{N}\left(\dot{x}_j\left(t-h_j\right) - \dot{x}_i(t)\right) \qquad (1)$$

здесь и далее $i = 1,..,N, t \in [0,T]$.

Введем следующие обозначения

$$\dot{x}_i(t) = y_i(t), \qquad (2)$$

$$z_j(t) = y_j(t - h_j), \; j = 1,..,N, t \in [0,T] \qquad (3)$$

и перепишем уравнение (1) в виде системы дифференциальных уравнений первого порядка:

$$\dot{x}_i(t) = y_i(t), \; i = 1,..,N, t \in [0,T] \qquad (4)$$

$$\dot{y}_i(t) = -\omega_i{}^2 x_i(t) + \varepsilon(1 - \beta_i x_i{}^2(t)) + u(t) \sum_{j=1}^{N} \left(y_j(t - h_j) - y_i(t)\right) \qquad (5)$$

с начальными условиями

$$x_i(0) = a_i \,, \qquad (6)$$

$$\dot{x}_i(t) = \varphi_i(t), \; i = 1,..,N, \quad t \in [-\max\{h_j\}, 0] \qquad (7)$$

и ограничениями на функцию управления

$$u(t) \epsilon U(t). \qquad (8)$$

Здесь $x_i(t)$ – амплитуда колебаний i-нейрона; $y_i(t)$ – скорость изменения амплитуды колебаний i-нейрона; $u(t)$ – функция управления, характеризующая внешнее воздействие ансамбля нейронов на i-й нейрон; N – количество нейронов; T – фиксированное время протекания процесса; $\varepsilon > 0$ и β – коэффициенты, характеризующие нелинейное воздействие на i-й нейрон ансамбля нейронов, в частности, β влияет на силу затухания колебания i-го нейрона; ω_i – собственная частота колебания i-го нейрона.

Заметим, что уравнение $\ddot{x}_i(t) + \varepsilon(1 - \beta x_i{}^2(t)) + \omega_i{}^2 x_i(t) = 0$ называется осциллятором Ван дер Поля [5] и учитывает нелинейность колебательного процесса, а слагаемое $u(t) \sum_{j=1}^{N} (y_j(t - h_j) - y_i(t))$ характеризует силу воздействия ансамбля нейронов на i-й нейрон.

Характерной особенностью данной математической модели является учет запаздывания при передаче сигнала от одного нейрона к другому.

Целью управления искусственной нейронной сетью является ее обучение, и включает в себя решение следующих задач:

1. построение процесса, при котором в конечный момент времени характеристики нейронов будут совпадать с заданными терминальными условиями;
2. построение оптимального управления $\bar{u}(t)$ ансамблем нейронов.

Задача оптимального управления формализуется как задача минимизации функционала (9)

$$I(u) = R_1 \int_0^T f_0(t, x(t), u(t))dt + R_2 \Phi(x(T)) \to \inf \qquad (9)$$

при ограничениях (4) – (8). Здесь R_1, R_2 – весовые коэффициенты в двухкритериальной задаче оптимального управления.

Целью управления является построение оптимального процесса, минимизирующего функционал (9) при заданных динамических ограничениях (4) – (8).

ПОСТРОЕНИЕ РЕШЕНИЯ. ПРИНЦИП МАКСИМУМА ПОНТРЯГИНА. ОБЩИЙ СЛУЧАЙ. Для определения функции управления в задаче (4) – (9) воспользуемся принципом максимума для динамических систем с запаздывающим аргументом [3], [7], [8].

Функция Понтрягина задачи (4) – (9) имеет вид:

$$H(\lambda_0, x, y, z, p(t), q(t), u) = -\lambda_0 R_1 f_0(t, x, u) + \sum_{i=1}^{N} p_i(t) y_i +$$
$$+\sum_{i=1}^{N} q_i(t)(-\omega_i^2 x_i + \varepsilon\left(1-\beta x_i^2\right) + u \sum_{j=1}^{N}\left(z_j - y_i\right)), \quad (10)$$

где $z_j = y_j\left(t - h_j\right),\ p(t),\ q(t)$ – сопряженные вектор-функции.

Принцип максимума в задаче (4) – (9) запишется в виде (11):

$$-\lambda_0 R_1 f_0(t, \overline{x}, \overline{u}) + \sum_{i=1}^{N} p_i(t) \overline{y}_i +$$
$$+\sum_{i=1}^{N} q_i(t)(-\omega_i^2 \overline{x}_i + \varepsilon\left(1-\beta \overline{x}_i^2\right) + \overline{u} \sum_{j=1}^{N}\left(\overline{z}_j - \overline{y}_i\right)) = \quad (11)$$
$$= \sum_{i=1}^{N} p_i(t) \overline{y}_i + \sum_{i=1}^{N} q_i(t)(-\omega_i^2 \overline{x}_i + \varepsilon\left(1-\beta \overline{x}_i^2\right)) +$$
$$+max_{v\in U}\left[-\lambda_0 R_1 f_0(t, \overline{x}, v) + v \sum_{i=1}^{N} \sum_{j=1}^{N}\left(q_i(t)\left(\overline{z}_j - \overline{y}_i\right)\right)\right].$$

Согласно принципу максимума оптимальное управление определяется из выражения (12):

$$\overline{u}(t) = argmax_{v\epsilon U}\left[-\lambda_0 R_1 f_0(t, \overline{x}, v) + v \sum_{i=1}^{N} \sum_{j=1}^{N}\left(q_i(t)(\overline{z}_j - \overline{y}_i)\right)\right]. \quad (12)$$

Для получения сопряженной системы используем принцип максимума для задачи с запаздывающим аргументом. Согласно этому принципу сопряженные вектор-функции $p(t)$, $q(t)$ удовлетворяют системе дифференциальных уравнений с опережающим аргументом (13)–(14):

$$\dot{p}_k(t) = -\frac{\partial H}{\partial x_k}\left(\lambda_0, \overline{x}(t), \overline{y}(t), \overline{z}(t), p(t), q(t), \overline{u}(t)\right), \quad (13)$$

$$\dot{q}_k(t) = -\frac{\partial H}{\partial y_k}\left(\lambda_0, \overline{x}(t), \overline{y}(t), \overline{z}(t), p(t), q(t), \overline{u}(t)\right) - \quad (14)$$
$$-\frac{\partial H}{\partial z_k}\left(\lambda_0, \overline{x}(t+h_k), \overline{y}(t+h_k), \overline{z}(t+h_k), p(t+h_k), q(t+h_k), \overline{u}(t+h_k)\right).$$

С учетом определения функции Понтрягина (10) система (13)–(14) запишется в виде:

$$\dot{p}_k(t) = \lambda_0 R_1 \frac{\partial f_0(t,\bar{x}(t),\bar{u}(t))}{\partial x_k} + q_k(t)\omega_k^2 + 2q_k(t)\varepsilon\beta\bar{x}_k(t) \qquad (15)$$

$$\dot{q}_k(t) = -p_k(t) + \bar{u}(t)Nq_k(t) - \bar{u}(t+h_k)\textstyle\sum_{i=1}^{N} q_k(t+h_k). \qquad (16)$$

Для сопряженной системы (15)–(16) должны выполняться условия трансверсальности (17)–(19):

$$p_k(T) = -\lambda_0 \frac{\partial \Phi}{\partial x_k}\big(\bar{x}(T)\big), \qquad (17)$$

$$q_k(T) = -\lambda_0 \frac{\partial \Phi}{\partial y_k}\big(\bar{x}(T)\big), \qquad (18)$$

$$q_k(t) = 0, t > T, k = 1,\dots,N. \qquad (19)$$

Заметим, что, принцип максимума позволяет свести задачу построения оптимального процесса к решению краевой задачи [2], которая включает в себя дифференциальные уравнения с отклоняющимся аргументом и граничные условия:

$$\dot{x}_i(t) = y_i(t),$$

$$\dot{y}_i(t) = -\omega_i{}^2 x_i(t) + \varepsilon\big(1 - \beta_i x_i{}^2(t)\big) + \bar{u}(t)\sum_{j=1}^{N}\Big(y_j\big(t-h_j\big) - y_i(t)\Big),$$

$$x_i(0) = a_i,\ \dot{x}_i(t) = \varphi_i(t), \quad t \in \big[-\max\{h_j\}, 0\big],$$

$$\dot{p}_k(t) = \lambda_0 R_1 \frac{\partial f_0(t,x(t),\bar{u}(t))}{\partial x_k} + q_k(t)\omega_k^2 + 2q_k(t)\varepsilon\beta x_k, \qquad (20)$$

$$\dot{h}_k(t) = -p_k(t) + \bar{u}(t)q_k(t) - \bar{u}(t+h_k)\textstyle\sum_{i=1}^{N} q_i(t+h_k),$$

$$p_k(T) = -\lambda_0 \frac{\partial \Phi}{\partial x_k}\big(x(T)\big), q_k(T) = -\lambda_0 \frac{\partial \Phi}{\partial y_k}\big(x(T)\big), q_k(t) = 0, t > T, \text{ где}$$

$$\bar{u}(t) = arg\,max_{v\epsilon U}\left[-\lambda_0 R_1 f_0(t,\bar{x},v) + v\sum_{i=1}^{N}\sum_{j=1}^{N}\big(q_i(t)(\bar{y}_j(t-h_j) - \bar{y}_i)\big)\right].$$

Во всех соотношениях краевой задачи принципа максимума (20), где не оговаривается особо, $i = \overline{1,N},\ k = \overline{1,N},\ t \in [0,T]$.

ПОСТРОЕНИЕ РЕШЕНИЯ. ПРИНЦИП МАКСИМУМА ПОНТРЯГИНА. ЧАСТНЫЙ СЛУЧАЙ. В частном случае при $f_0\big(t,x(t),u(t)\big) = \frac{u^2(t)}{2}$ и $\Phi\big(x(T)\big) = \sum_{i=1}^{N}(x_i(T) - A_i)^2$ минимизируемый функционал примет вид (21):

$$I(u) = R_1\int_0^T \frac{u^2(t)}{2}dt + R_2\sum_{i=1}^{N}(x_i(T) - A_i)^2 \to \inf. \qquad (21)$$

Функция Понтрягина запишется в виде (22):

$$H(\lambda_0,x,y,z,p(t),q(t),u) = -\lambda_0 R_1 \frac{u^2}{2} + \sum_{i=1}^{N} p_i(t) y_i + \\ + \sum_{i=1}^{N} q_i(t)(-\omega_i^2 x_i + \varepsilon\left(1-\beta x_i^2\right) + u \sum_{j=1}^{N}(z_j - y_i)). \qquad (22)$$

Условие максимума примет вид

$$-\lambda_0 R_1 \frac{\bar{u}^2}{2} + \sum_{i=1}^{N} p_i(t) \bar{y}_i + \\ + \sum_{i=1}^{N} q_i(t)(-\omega_i^2 \bar{x}_i + \varepsilon\left(1-\beta \bar{x}_i^2\right) + \bar{u} \sum_{j=1}^{N}\left(\bar{z}_j - \bar{y}_i\right)) = \qquad (23) \\ = \sum_{i=1}^{N} p_i(t)\bar{y}_i + \sum_{i=1}^{N} q_i(t)(-\omega_i^2 \bar{x}_i + \varepsilon\left(1-\beta \bar{x}_i^2\right)) + \\ + max_{v\in U}\left[-\lambda_0 R_1 \frac{v^2}{2} + v \sum_{i=1}^{N}\sum_{j=1}^{N}\left(q_i(t)\left(\bar{z}_j - \bar{y}_i\right)\right)\right].$$

Из (23) получим

$$\bar{u}(t) = \ argmax_{v\epsilon U}\left[-\lambda_0 R_1 \frac{v^2}{2} + v \sum_{i=1}^{N}\sum_{j=1}^{N}\left(q_i(t)(\bar{y}_j\left(t - \ h_j\right) - \bar{y}_i)\right)\right]. \quad (24)$$

Сопряженная система и условия трансверсальности примут вид

$$\dot{p}_k(t) = q_k(t)\omega_k^2 + 2q_k(t)\varepsilon\beta\bar{x}_k(t), \qquad (25)$$

$$\dot{q}_k(t) = -p_k(t) + \bar{u}(t)q_k(t) - \bar{u}(t+h_k)\sum_{i=1}^{N} q_i(t+h_k), \qquad (26)$$

$$p_k(T) = -2\lambda_0(x_k(T) - A_k)R_2, \qquad (27)$$

$$q_k(t) = 0, t \geq T. \qquad (28)$$

Тогда краевая задача принципа максимума запишется в виде (29):

$$\begin{aligned} &\dot{x}_i(t) = y_i(t), \\ &\dot{y}_i(t) = -\omega_i{}^2 x_i(t) + \varepsilon\left(1 - \beta_i x_i{}^2(t)\right) + \bar{u}(t)\textstyle\sum_{j=1}^{N}\left(y_j\left(t-h_j\right) - y_i(t)\right) \\ &x_i(0) = a_i,\ \dot{x}_i(t) = \varphi_i(t),\ \ i = 1,..,N,\ \ t \in \left[-\max\{h_j\},0\right] \\ &\dot{p}_k(t) = q_k(t)\omega_k^2 + 2q_k(t)\varepsilon\beta\bar{x}_k(t), \qquad (29) \\ &\dot{q}_k(t) = -p_k(t) + \bar{u}(t)q_k(t) - \bar{u}(t+h_k)\textstyle\sum_{i=1}^{N} q_i(t+h_k) \\ &p_k(T) = -2\lambda_0(x_k(T) - A_k)R_2, \\ &h_k(t) = 0, t \geq T, \end{aligned}$$

где $\bar{u}(t)$ определяется по формуле (24).

Во всех соотношениях краевой задачи принципа максимума (29), где не оговаривается особо, $i = \overline{1,N},\ k = \overline{1,N},\ t \in [0,T]$.

ДИСКРЕТНАЯ ЗАДАЧА. ПРИМЕНЕНИЕ БАД. ОБЩИЙ СЛУЧАЙ. Разбивая временной интервал $[0,T]$ на n слоев и аппроксимируя интеграл по правилу левых прямоугольников, а систему дифференциальных уравнений по схеме Эйлера, получим дискретную задачу оптимального управления [9], [10]:

$$I(u) = R_1 \sum_{l=0}^{n-1} f_0(t, x^l, u^l)\Delta t + R_2 \Phi(x^n), \qquad (30)$$

$$x_i^{\,l+1} = x_i^{\,l} + \Delta t\, y_i^{\,l}, \qquad (31)$$

$$y_i^{l+1} = y_i^{\,l} + \Delta t\left(-(\omega_i)^2 x_i^{\,l} + \varepsilon\left(1-\beta(x_i^{\,l})^2\right) + u^l \sum_{j=1}^{N}\left(y_j^{\,l-v_j} - y_i^{\,l}\right)\right), \qquad (32)$$

$$x_i^{\,0} = a_i, \qquad (33)$$

$$y_i^{\,l} = \varphi_i^{\,l}, \;\; l \in [-\max\{v_j\}, 0], \qquad (34)$$

$$u^l \epsilon U. \qquad (35)$$

Здесь и далее нижний индекс $i = 1, .., N$ означает номер координаты вектора, а верхний $l = 0, \dots, n$ – номер нейронного слоя в схеме аппроксимации.

Для построения решения дискретной задачи оптимального управления используем метод быстрого автоматического дифференцирования (БАД) [10]. Для дискретной задачи оптимального управления (30) – (35) построим функцию Лагранжа (36):

$$\begin{aligned} L = {} & \lambda_0\left(R_1 \sum_{l=0}^{n-1} f_0(t, x^l, u^l)\Delta t + R_2\Phi(x^n)\right) + \\ & + \sum_{l=0}^{n-1}\sum_{i=1}^{N} p_i^{\,l+1}\left(x_i^{\,l+1} - x_i^{\,l} - \Delta t y_i^{\,l}\right) + \\ & + \sum_{l=0}^{n-1}\sum_{i=0}^{N} q_i^{\,l+1}\left(y_i^{l+1} - y_i^{\,l}\right) - \\ & - \Delta t \sum_{l=0}^{n-1}\sum_{i=0}^{N} q_i^{\,l+1}\left(-(\omega_i)^2 x_i^{\,l} + \varepsilon\left(1-\beta(x_i^{\,l})^2\right) + u^l \sum_{j=1}^{N}\left(y_j^{\,l-v_j} - y_i^{\,l}\right)\right). \end{aligned} \qquad (36)$$

Вычислим градиент минимизируемой функции:

$$\frac{\partial L}{\partial u^m} = \lambda_0 R_1 \Delta t \frac{\partial f_0(t, x^m, u^m)}{\partial u^m} - \Delta t \sum_{i=1}^{N}\sum_{j=1}^{N}\left(q_i^{\,m+1}\left(y_j^{m-v_j} - y_i^m\right)\right). \qquad (37)$$

Здесь и далее *k=1, ..., N, m=0, ..., n-1*.

Условия стационарности по фазовым переменным имеют вид:

$$\frac{\partial L}{\partial x_k{}^m} = \lambda_0 R_1 \Delta t \frac{\partial f_0(t, x^m, u^m)}{\partial x_k{}^m} + p_k{}^m - p_k{}^{m+1} +$$
$$+ q_k{}^{m+1} \Delta t (\omega_k)^2 + 2\Delta t q_k{}^{m+1} \varepsilon x_k{}^m = 0, \quad (38)$$
$$\frac{\partial L}{\partial x_k{}^n} = \lambda_0 R_2 \frac{\partial \Phi(x^n)}{\partial x_k{}^n} + p_k{}^q = 0, \quad (39)$$
$$\frac{\partial L}{\partial y_k{}^m} = -\Delta t\, p_k{}^{m+1} + q_k{}^m - q_k{}^{m+1} -$$
$$-\Delta t\, u^{m+v_k} \sum_{i=1}^{N} q_k{}^{m+v_k+1} - \Delta t q_k{}^{m+1} N u^m = 0, \quad (40)$$
$$\frac{\partial L}{\partial y_k{}^n} = q_k{}^n = 0. \quad (41)$$

Выражая из условий (38)–(41) $p_k{}^m$, $p_k{}^n$, $q_k{}^m$, $q_k{}^n$, получим рекуррентные соотношения для вычисления сопряженных переменных:

$$p_k{}^m = -\lambda_0 R_1 \Delta t \frac{\partial f_0(t, x^m, u^m)}{\partial x_k{}^m} + p_k{}^{m+1} - q_k{}^{m+1} \Delta t (\omega_k)^2 -$$
$$-2\Delta t q_k{}^{m+1} \varepsilon x_k{}^m, \quad (42)$$
$$p_k{}^q = -\lambda_0 R_2 \frac{\partial \Phi(x^n)}{\partial x_k{}^n}, \quad (43)$$
$$q_k{}^m = \Delta t\, p_k{}^{m+1} + q_k{}^{m+1} + \Delta t\, u^{m+v_k} \sum_{i=1}^{N} q_k{}^{m+v_k+1} + \Delta t q_k{}^{m+1} N u^m, (44)$$
$$q_k{}^m = 0, \mathrm{m} \geq \mathrm{n}. \quad (45)$$

Рекуррентные соотношения (42)–(45) используются в численном алгоритме для вычисления градиента минимизируемой функции.

ДИСКРЕТНАЯ ЗАДАЧА. ПРИМЕНЕНИЕ БАД. ЧАСТНЫЙ СЛУЧАЙ.

В частном случае при $f_0(t, x^l, u^l) = \frac{(u^l)^2}{2}$ и $\Phi(x(T)) = \sum_{i=1}^{N} (x_i{}^n - A_i)^2$

минимизируемая функция имеет вид

$$I(u) = R_1 \sum_{l=0}^{n-1} \frac{(u^l)^2}{2} \Delta t + R_2 \sum_{i=1}^{N} (x_i{}^n - A_i)^2. \quad (46)$$

Теперь требуется минимизировать функцию (46) при условиях (31)–(35).

Функция Лагранжа дискретной задачи (46), (31)–(35) запишется в виде (47):

$$L=\lambda_0\left(R_1\sum_{l=0}^{n-1}\frac{(u^l)^2}{2}\Delta t+R_2\sum_{i=1}^{N}({x_i}^n-A_i)^2\right)+$$
$$+\sum_{l=0}^{n-1}\sum_{i=1}^{N}{p_i}^{l+1}({x_i}^{l+1}-{x_i}^l-\Delta t{y_i}^l)+ \qquad (47)$$
$$+\sum_{l=0}^{n-1}\sum_{i=0}^{N}{q_i}^{l+1}(y_i^{l+1}-{y_i}^l)-$$
$$-\Delta t\sum_{l=0}^{n-1}\sum_{i=0}^{N}{q_i}^{l+1}\left(-(\omega_i)^2{x_i}^l+\varepsilon\left(1-\beta({x_i}^l)^2\right)+u^l\sum_{j=1}^{N}\left({y_j}^{l-v_j}-{y_i}^l\right)\right).$$

Градиент минимизируемой функции имеет вид (48):

$$\frac{\partial L}{\partial u^m}=\lambda_0R_1\Delta tu^m-\Delta t\sum_{i=1}^{N}\sum_{j=1}^{N}({q_i}^{m+1}(y_j^{m-v_j}-y_i^m)). \qquad (48)$$

Рекуррентные соотношения для вычисления сопряженных переменных ${p_k}^m$, ${p_k}^n$ записываются в виде (49)–(50):

$${p_k}^m={p_k}^{m+1}-{q_k}^{m+1}\Delta t(\omega_k)^2-2\Delta t{q_k}^{m+1}\varepsilon{x_k}^m, \qquad (49)$$

$${p_k}^n=-2\lambda_0R_2\cdot({x_k}^n-A_k). \qquad (50)$$

Сопряженные переменные ${q_k}^m$ по-прежнему вычисляются с помощью рекуррентных соотношений (44)–(45).

Кроме перечисленных условий в дискретной задаче должны также выполняться условия дополняющей нежесткости, не отрицательности и согласования знаков, а также – условия допустимости решения [9], [12].

При численной реализации метода быстрого автоматического дифференцирования ограничения на управления не вводятся в функцию Лагранжа, а учитываются с помощью метода проекции градиента [9], [10].

Достоверность результатов базируется на применении теоретически обоснованных методов теории оптимального управления и методов оптимизации.

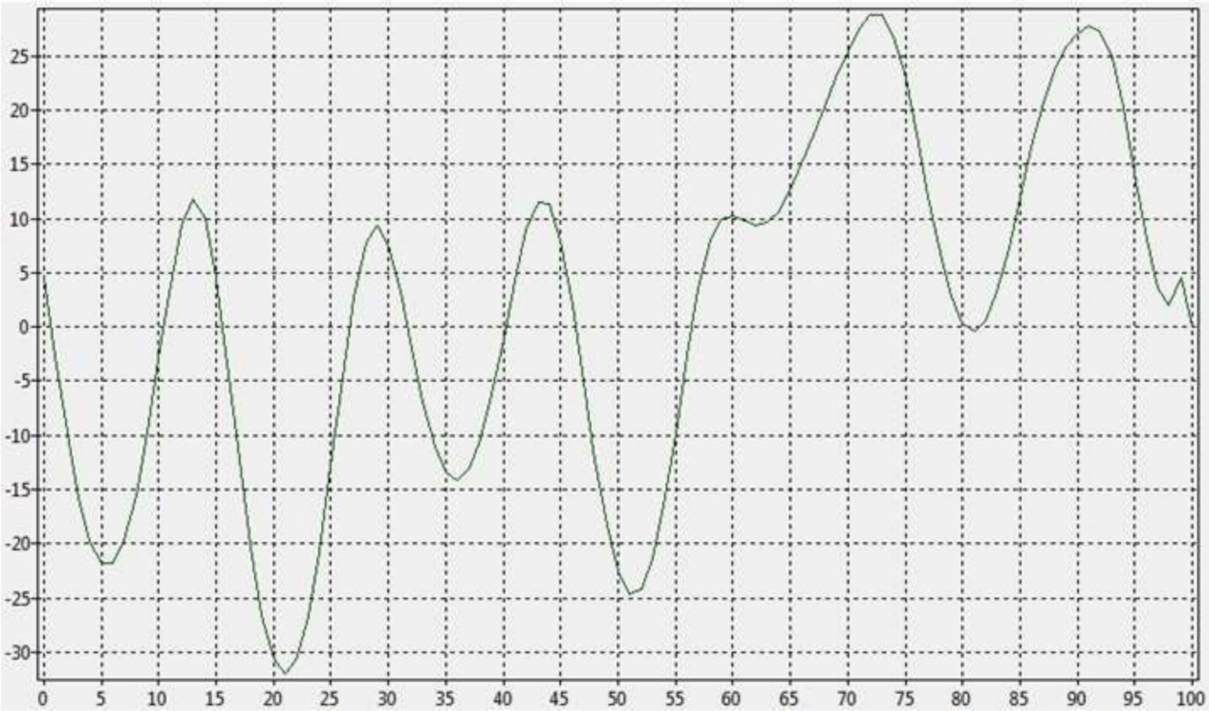

Рисунок 1. График функции управления *u*

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА ЧИСЛЕННОГО РЕШЕНИЯ. Дискретная задача решалась методом проекции градиента [9], [10]. Основными требованиями к пользовательскому интерфейсу программы являются гибкость, простота, возможность введения всех исходных данных и отображения результатов работы алгоритма в форме, удобной для непосредственного восприятия пользователем программы: в виде числовых значений, таблиц и графиков. Для программной реализации был выбран язык программирования Delphi и среда разработки Lazarus. Авторы выражают благодарность Перовой Е.А., выпускнице университета, специальности "Компьютерная безопасность" Тверского государственного за оказанное содействие в проведении числительных программных экспериментов.

Приведем результаты расчетов при следующих значениях параметров модели: количество нейронов N=3; количество слоев n=100;
весовой коэффициент R1=0.00001; весовой коэффициент R2=0.2; коэффициент, характеризующий нелинейное воздействие на i-й нейрон ансамбля нейронов $\varepsilon = 0.1$; коэффициент, влияющий на силу затухания колебания i-го нейрона $\beta = 0.0003$; время протекания процесса *T*=10; собственная частота колебаний i-нейрона $\omega_i = 2, i = 1,2,3$;. амплитуда колебаний нейронов $x_1 = 1, x_2 = 10, x_3 = 3$; $y_i{}^0 = 0,\ \ v_i = 0, i = 1,..,3$; терминальные значения амплитуды колебаний $A_i = 0.1,\ i = 1,2,3$;.

начальный набор управлений для метода градиентного спуска $u^m = 0.18, m = 0, \ldots, 99$.

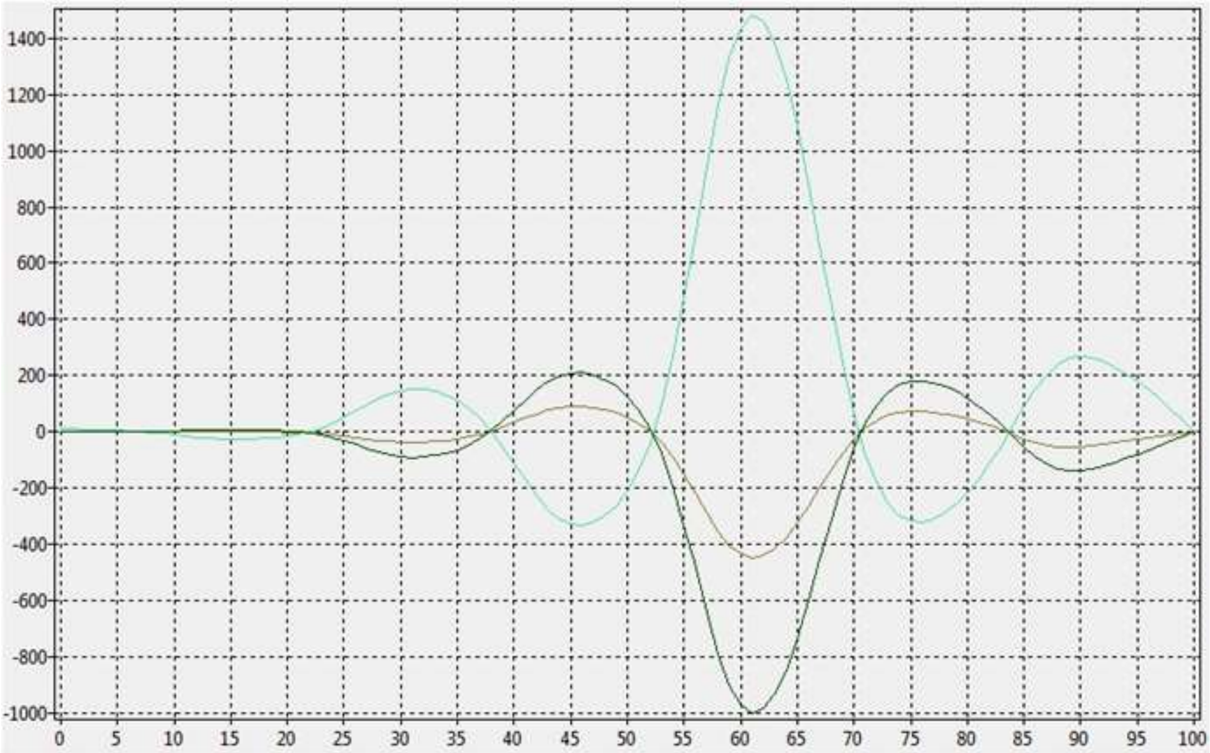

Рисунок 2. График вектор-функции состояния *x*

На Рисунках 1, 2 и 3 изображены графики функций управления, состояния и сопряженной вектор-функции p. Как видно из графиков функций состояния, изображенных на Рисунке 2, построенный процесс, удовлетворяет заданному критерию оптимизации, а именно: в конечный момент времени характеристики нейронов совпадают с заданными терминальными значениями амплитуды колебаний $A_i = 0.1, \ i = 1,2,3$.

ЗАКЛЮЧЕНИЕ. В работе построена и проанализирована динамическая модель искусственной нейронной сети, которая может применяться при моделировании процессов памяти и внимания. Эту модель можно совершенствовать, используя различные функции активации ансамбля нейронов, учитывая зависимость запаздывания от времени и эффект запаздывания в аргументе функции управления.

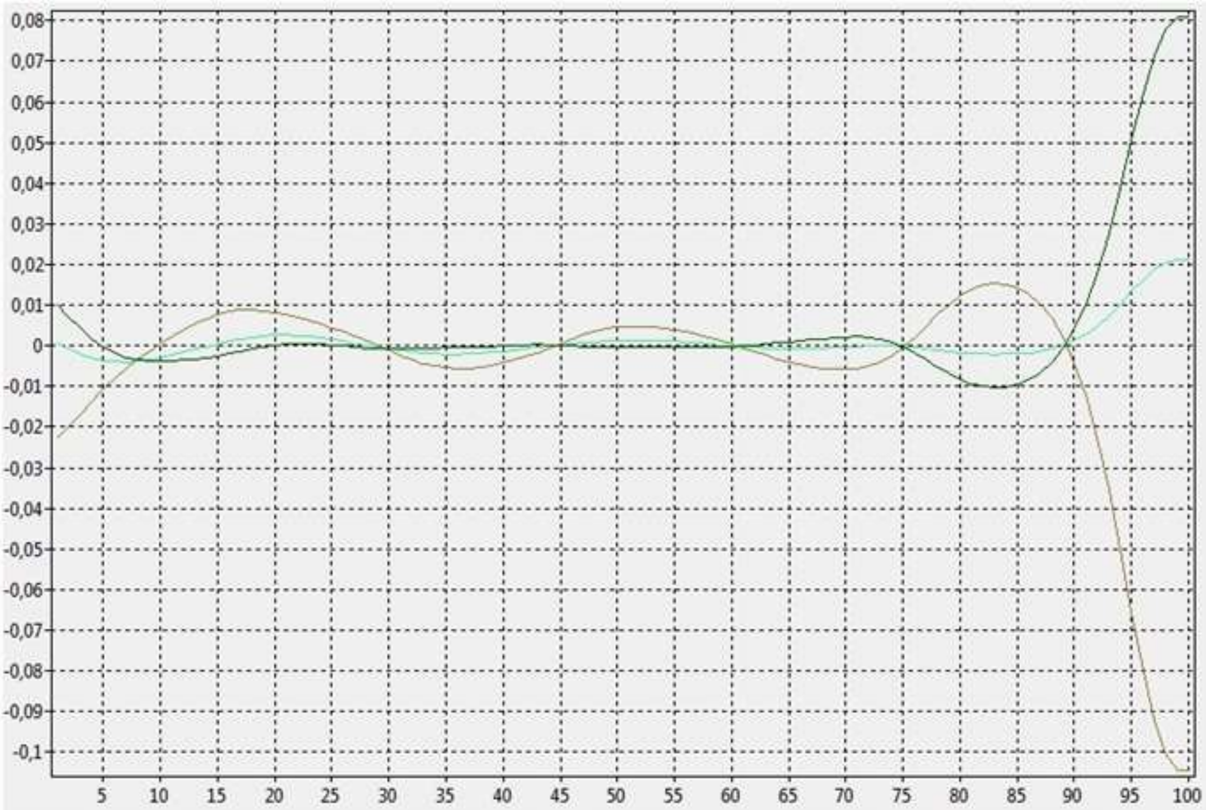

Рисунок 3. График сопряженной вектор-функции *p*

Предложенный в работе алгоритм позволяет находить решение задачи в зависимости от параметров модели, исследовать влияние этих параметров на оптимальный процесс и значение целевого функционала задачи; рассматривать многокритериальные задачи со сверткой критериев; исследовать влияние весовых коэффициентов и различных типов штрафных функций, используемых для учета ограничений на функцию управления и функцию состояния системы.

Проведенные к настоящему времени исследования показывают эффективность математического моделирования динамических процессов, базирующегося на использовании искусственных нейронных сетей.

ЛИТЕРАТУРА

1. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. – М.: Мир, 1992. – 184 с.
2. Андреева Е.А., Кратович П.В. Оптимизация нейронных сетей. – Тверь: Твер. гос. ун-т, 2015. – 116 с.
3. Андреева Е.А., Колмановский В.Б., Шайхет Л.Е. Управление системами с последействием. – М.: Наука, 1992. – 336 с.
4. Галушкин А.И.Нейронные сети. Основы теории. – М.: Горячая линия – Телеком, 2012. – 496 с.

5. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р. Модели динамики нейронной активности при обработке информации мозгом – итоги «десятилетия». //Успехи физических наук. Т. 172, №10, 2002. – С. 1189–1214.
6. Андреева Е.А. Управление динамическими системами. – Тверь: Твер. гос. ун-т, 2016. – 188 с.
7. Андреева Е.А. Оптимальное управление системами с запаздывающим аргументом. // Препринт. – М.: ВЦ АНСССР, 1987. – 32 с.
8. Андреева Е.А., Пустарнакова Ю.А. Численные методы обучения искусственных нейронных сетей с запаздыванием. // Журнал вычислительной математики и математической физики. 2002. – Т. 42. С. 1383–1391.
9. Андреева Е.А., Цирулева В.М. Вариационное исчисление и методы оптимизации. – М.: Высшая школа, 2006. – 584 с.
10. Евтушенко Ю.Г. Методы решения экстремальных задач и их применение в системах оптимизации. – М.: 1982. – 432 с.
11. Андреева Е.А., Мазурова И.С. Обучение искусственных нейронных сетей методом БАД. // Математические методы управления. Сборн. науч. трудов. – Тверь: Твер. гос. ун-т, 2015. – С. 5–18.
12. Андреева Е.А., Цирулева В.М., Кожеко Л.Г. Модель управления процессом рыбной ловли. // Моделирование, оптимизация и информационные технологии. – Воронеж: 2017. – №4 (19). 10 с.

E.A. Andreeva, V.M. Tsiruleva

MATHEMATICAL MODELING OF CONTROL OF A DYNAMIC NEURAL NETWORK WITH DELAY

Tver State University, Tver, Russia

Currently, the world is actively developing a new applied area of mathematics, related to the study of artificial neural networks. Interest in them is caused both by theoretical and applied achievements: the possibilities of using computations in spheres previously related only to the field of human intelligence were opened. The relevance of research in this direction is confirmed by numerous examples of the use of neural networks in automation systems [1], robotics of image recognition processes [2], adaptive control [3], forecasting and creating expert systems [4], research of associative memory [5], etc. In complex practical tasks, the trained neural network acts as an expert. An example is medical diagnostics, where a neural network can take into account a large number of numerical parameters (electrical impulses of the nerve cells of the brain and its parts, recorded by means of encephalograms, pressure, weight, etc.). The aim of the work is to construct an artificial oscillatory neural network that can be used to model the activity of the brain: associative memory and attention. The model is formalized as a multicriteria optimal control problem with delay. The purpose of neural network management is its training, which includes the construction of an optimal process that meets the specified criteria. One of the criteria is the terminal criterion determining the state of the neural network at the final *moment of time. The optimality conditions in the continuous model are obtained with the help of the Maximum principle for problems with delayed argument [6], [7], [8]. The boundary value problem of the maximum principle is constructed [9]. To obtain optimal conditions in a discrete model that approximates a continuous model, the method of rapid automatic differentiation and numerical methods for solving extremal problems are used [9], [10], [11]. The results of a numerical experiment are presented.*

Keywords: optimal control, oscillatory neural network, neuron ensemble, mathematical model, multicriteria problem, maximum principle with delayed argument, discrete optimal control problem.

REFERENCES

1. Wesserman F. Neurocomputer technology: theory and practice. – Moscow: Mir, 1992. – 184 p.
2. Andreeva E.A., Kratovich P.V. Optimization of neural networks. – Tver: Tver state University, 2015. – 116 p.
3. Andreeva E.A., Kolmanovsky V.B., Shaikhet L.Ye. Management of systems with aftereffect. – Moscow: Nauka, 1992. – 336 p.
4. 4. Galushkin A.I. Neural networks. Fundamentals of the theory. – M .: Hot line – Telecom, 2012. – 496 p.
5. Borisyuk G.N., Borisyuk R.M., Kazanovich Ya.B., Ivanitsky G.R. Models of the dynamics of neural activity in the processing of information by the brain – the results of the "decade". // The successes of physical sciences. Vol. 172, No. 10, 2002. – Pp. 1189–1214.
6. Andreeva E.A. Management of dynamic systems. – Tver: Tver state University, 2016. – 188 p.
7. Andreeva E.A. Optimal control of systems with delayed argument. // Preprint. – Moscow: VTS ANSSSR, 1987. – 32 p.
8. Andreeva E.A., Pustarnakova Yu.A. Numerical methods for training artificial neural networks with delay. // Journal of Computational Mathematics and Mathematical Physics. 2002. – Vol. 42. Pp. 1383–1391.
9. Andreeva E.A., Tsiruleva V.M. Variations calculus and optimization methods. – Moscow: Higher School, 2006. – 584 p.
10. Yevtushenko Yu.G. Methods for solving extremal problems and their application in optimization systems. – Moscow: 1982. – 432 p.
11. Andreeva E.A., Mazurova I.S. Training of artificial neural networks by dietary supplements. // Mathematical methods of control. Sborn. sci. works. – Tver: Tver state University, 2015. – Pp. 5–18.
12. Andreeva E.A., Tsiruleva V.M., Kozheko L.G. The model of fisheries management. // Modeling, optimization and information technologies. – Voronezh: 2017. – No. 4 (19).