УДК 519.3

Статья представлена дочерью А. И. Каплинского Nadia Udler

О ПОСТРОЕНИИ БЫСТРО СХОДЯЩИХСЯ МЕТОДОВ НЕГЛАДКОЙ ОПТИМИЗАЦИИ МНОГОСЛОЙНОГО ПЕРСЕПТРОНА

А. И. Каплинский, **А. М. Песин**

Воронежский государственный университет

Поступила в редакцию 01.12.2018 г.

Аннотация. В статье предлагается подход к построению поисковых методов, обладающих высокой скоростью сходимости, для оптимизации критерия качества настройки многослойного персептрона. Рассматриваемая задача имеет непосредственное отношение к построению искусственных нейронных сетей и разработке нейрокомпьютеров.

Ключевые слова: персептрон, условие локального улучшения, потенциал.

Annotation. The paper proposes an approach to the construction of search methods with a high convergence rate to optimize the quality criterion of setting a multilayer perceptron. The considered problem is directly related to the building of artificial neural networks and the development of neural computers.

Keywords: perceptron, condition of local improvement, potential.

ВВЕДЕНИЕ

В работах [1–5] предложены и исследованы методы оптимизации скалярных негладких и многоэкстремальных функций векторного аргумента. В [7] данный подход развивается для построения быстро сходящихся методов негладкой и нелокальной оптимизации. В [6, 8] изучались вопросы выхода оптимизируемой функции из области локального экстремума. В настоящей работе предложенный подход используется для решения задачи идентификации на примере оптимизации многослойного персептрона [12]. Задача оптимизации многослойного персептрона имеет непосредственное отношение к построению вычислительных схем, названных искусственными нейросетями [9, 10, 12], и построению вычислительных устройств, названных нейрокомпьютерами [9, 10, 12].

1. ПОСТАНОВКА И АНАЛИЗ ЗАДАЧИ

Пусть требуется минимизировать скалярную, возможно негладкую и многоэкстремальную функцию $\Phi\left[Y\right]$ векторного аргумента Y, $n \geq 1$, $Y \in R^n$. Функция $\Phi\left[Y\right]$ является критерием качества настройки многослойного персептрона [12].

Векторы входов Y_N имеют следующий вид:

$$Y_N = f\left[W_N Y_{N-1}\right],$$

где $N \geq 1$, $f = \left(f_1, ..., f_n\right)$, $f_i\left(i = \overline{1, n}\right)$ – нелинейные сигмоидные отображения [10, 12]. Здесь W_N $\left(N \geq 1\right)$ – это числовые матрицы размерности $n \times n$, которые выбираются с целью решения следующей задачи оптимизации

$$\Phi\left[f\left[W_N Y_{N-1}\right]\right] \to \min_{(W_N, W_{N-1}, ..., W_0)}. \quad (1)$$

Введем обозначение $z_N = W_N Y_{N-1}$, где $N \geq 1$. Тогда настраиваемая модель со структурой многослойного персептрона имеет вид

$$\Phi\left[f\left(z_N\right)\right] \to \Phi\left[Y_N\right] \min_{(W_N, W_{N-1}, ..., W_0)}. \quad (2)$$

В частном случае можно рассматривать минимизацию среднеквадратического критерия качества вида

$$\Phi\left[Y_N\right] \to M\left[Y_N - Y^*\right]^2 \min_{(W_N, W_{N-1}, \ldots, W_0)},$$

где матрицы W_N настраиваемой модели (2) могут выбираться стохастическими [11].

Задача (1)–(2) может рассматриваться как задача идентификации заданного вектора входов Y^* векторами выходов Y_N настраиваемой модели (2), при этом вектор входов Y^* также может считаться случайным. В данной статье рассматривается задача (1)–(2) применительно к возможно негладкой и многоэкстремальной функции $\Phi\left[Y\right]: R^n \to R^1$, причем упор делается на построение именно поисковых методов оптимизации. Этот выбор обусловлен тем, что согласно [9, 10, 12], вектор-функция $f = \left(f_1, \ldots, f_n\right)$ состоит из сигмоидных нелинейных одномерных отображений f_i [9, 10, 12]. При использовании методов гладкой оптимизации в случае, например, критерия $\Phi\left[Y_N\right] \to M\left[Y_N - Y^*\right]^2$ с необходимостью участвует производная f_i' сигмоидного отображения. Дифференцирование отображений f_i ввиду их структуры приводит к обнулению достаточно больших массивов коэффициентов матриц W_N, что может привести к отсутствию работоспособности настраиваемого многослойного персептрона [9, 10, 12].

2. ПОСТРОЕНИЕ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ (1)–(2)

Согласно работам [1, 4, 5, 11], построение эффективных на практике методов негладкой оптимизации многоэкстремальных функций заключается в предварительной рандомизации исходной задачи (1)–(2). Для этого рассмотрим матричный линейный стохастический процесс

$$\begin{cases} W_N = W_{N-1} + \varepsilon_N \theta_N \quad \left(N \geq 1\right), \\ \left\|M\left(W_0\right)\right\| < \infty, \left\|\left(\theta_1\right)\right\| < \infty, 0 < \varepsilon_N < \infty, \\ \left(\varepsilon_1 + \varepsilon_2 + \ldots\right) = \infty. \end{cases} \tag{3}$$

Согласно введенным ранее обозначениям, имеет место выражение:

$$Z_N = W_N Y_{N-1} + \varepsilon_N \cdot \theta_N \cdot Y_{N-1}. \tag{4}$$

Векторный поисковый случайный процесс (4) приводит к рандомизации следующего критерия оптимизации:

$$F(\varepsilon_N) = \int_{R^N} \hat{\Phi}[f(z)] p_{Z_N}(z) dz,$$

где $\hat{\Phi}[f(z)] = \Phi[f(z)] - c_N$, $c_N = \text{const}$.

Задача (1)–(2) преобразуется в следующую задачу:

$$F(\varepsilon_N) \to \min_{(\varepsilon_N)}. \tag{5}$$

Заметим, что разложение (4) Z_N в сумму двух случайных векторов позволяет полностью обосновать математическую корректность перехода от задачи (1)–(2) к задаче (5) [например, 1, 3, 4].

Предположим, что существует плотность распределения p_{Z_N} случайного вектора Z_N для каждого $N = 1, 2, 3, \ldots$. Пусть выполняется условие знакоопределенности

$$\hat{\Phi}[f(z)] = \Phi[f(z)] - c_N \geq 0,$$

где c_N – соответствующие константы, такие что $0 < c_N < \infty$, $\left(N = 1, 2, 3, \ldots\right)$.

При заданном на шаге N векторе входов Y_{N-1} должна выбираться из условия (6) матрица θ_N, задающая вектор $\theta_N Y_{N-1}$ направления движения для решения задач (1)–(2) или (5). Условие невозрастания критерия оптимизации $F(\varepsilon_N)$ на каждом шаге $N = 1, 2, 3, \ldots$ названо *условием локального улучшения* (УЛУ) [1, 2, 3, 4]. Оно имеет следующий вид:

$$\left.\frac{dF}{d\varepsilon}\right|_{\varepsilon=0} \leq 0. \tag{6}$$

Для задач (1)–(2) и (5) УЛУ (6) имеет вид

$$\begin{gathered} \left.\frac{dF}{d\varepsilon}\right|_{\varepsilon=0} \leq 0 = \int_{R^N} \hat{\Phi}[f(z)] \cdot div_Z[p_{W_{N-1}Y_{N-1}}(z) \times \\ \times M(\theta_N Y_{N-1} / W_{N-1} Y_{N-1} = z) dz, \end{gathered} \tag{7}$$

где $M(\theta_N Y_{N-1} / W_{N-1} Y_{N-1} = z)$ – вектор условного математического ожидания случайного вектора $\theta_N Y_{N-1}$.

Для дальнейших рассуждений введем обозначение $y_N(z) = M(\theta_N Y_{N-1} / W_{N-1} Y_{N-1} = z)$.

С целью выполнения УЛУ (7) выберем вектор-функцию $y_N(z)$ как решение следующего уравнения:

$$\begin{gathered} div_Z[p_{W_{N-1}Y_{N-1}}(z) \cdot y_N(z)] = \\ = -p_{u_N}(z)[\Phi[f(z)] - c_N], \end{gathered} \tag{8}$$

при этом параметры c_N, названные константами уровней, выбираются из условия «неразрывности» [4, 5, 11, 8]

$$\int_{R^N} div_Z[p_{W_{N-1}Y_{N-1}}(z)\cdot y_N(z)]dz=0.$$

В (8) функция $p_{u_N}(z)$ имеет смысл плотности распределения вероятности вспомогательного случайного вектора u_N $(N=1,2,3,...)$, функция $p_{u_N}(z)$ определяет степень свободы при решении уравнения (8) [3, 5, 1, 8, 11].

При некоторых условиях [1, 2, 3, 4] векторное поле $p_{W_{N-1}Y_{N-1}}(z)\cdot y_N(z)$ можно разложить следующим образом

$$p_{W_{N-1}Y_{N-1}}(z)\cdot y_N(z)=\nabla_Z v_N(z)+\chi_N(z). \quad (9)$$

Здесь $v_N(z):R^n\to R^1$ для каждого значения $N=1,2,3,...$ является потенциалом векторного поля $p_{W_{N-1}Y_{N-1}}(z)$, а $\chi_N(z):R^n\to R^1$ есть бездивергентная составляющая разложения, то есть $div_Z[\chi_N(z)]=0$. Составляющая $\chi_N(z)$ определяет степень свободы в (9), причем считается, что $\chi_N(z)=0,(N=1,2,3,...)$.

Подставляя разложение (9) в уравнение (8), получим уравнение Пуассона на потенциале $v_N(z)$ вила

$$\Delta_Z v_N(z)=-p_{u_N}(z)[\hat{\Phi}[f(z)]], \quad (10)$$

где $\Delta=\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2}$.

Решение уравнения (10) имеет следующий вид

$$v_N(z)=\int_{R^N} \mathrm{E}(z,s)\hat{\Phi}[f(s)]\cdot p_{u_N}(s)ds, \quad (11)$$

где $E(z,s)=\begin{cases}(n-2)^{-1}\cdot\omega_n^{-1}|z-s|^{2-n},\ n>2,\\ \frac{1}{2\pi}\log|z-s|^{-1},\ n=2;\end{cases}$

$\omega_n=\frac{2(\sqrt{\pi})^n}{\Gamma(n/2)}$ – площадь поверхности единичной сферы в R^n.

Таким образом, функция (11) является потенциалом ньютоновского векторного поля (9) [1–4]. С целью ускорения сходимости предлагаемых методов оптимизации в работе [7] развивался подход, связанный с использованием собственных функций линейных операторов. Применим здесь этот подход и получим другой способ выполнения УЛУ (7).

В условиях [1–4] векторное поле $p_{u_N}(z)y_N(z)$ можно представить следующим образом

$$p_{u_N}(z)y_N(z)=\nabla_Z[\varphi_N(z)], \quad (12)$$

где

$$\Delta\varphi_N+q_N^2\varphi_N= =-p_{u_N}(z)[\hat{\Phi}[f(z)]],\ N=1,2,3,.. \quad (13)$$

$0<q_N<\infty$ – константы.

Потенциал $\varphi_N(z)$ является решением уравнения (13) типа Гельмгольца и имеет вид

$$\varphi_N(z)=\int_{R^N}\mathrm{E}_q(z,s)\hat{\Phi}[f(s)]\cdot p_{u_N}(s)ds. \quad (14)$$

В решении (14) ядро $\mathrm{E}_q(z,s)$ – это сферически-инвариантное (или фундаментальное) решение уравнения $(\nabla_Z+q_N^2)\mathrm{E}_q=0$ [7, 8].

Функция $\mathrm{E}_q(z,s)$ имеет следующий вид:

а) для четных n

$$\mathrm{E}_q(z,s)=|z-s|^{-\frac{(n-2)}{2}}\cdot J_{-\frac{(n-2)}{2}}(q_N|z-s|), \quad (15)$$

б) для нечетных n

$$\mathrm{E}_q(z,s)=|z-s|^{-\frac{(n-2)}{2}}\cdot J_{\frac{(n-2)}{2}}(q_N|z-s|).$$

В выражении (15) $J_{-\beta}$ и N_β обозначают соответствующие функции Бесселя и Неймана β-го порядка [7, 8].

Таким образом, функция $\varphi_N(z)$ (14) является потенциалом векторного поля (12).

3. УСКОРЕНИЕ СХОДИМОСТИ

Выберем функцию $p_{u_N}(z)$ в (13) как решение интегрального уравнения

$$\hat{\Phi}[f(z)]p_{u_N}(z)= =U_N(z)\int_{R^N}\mathrm{E}_q(z,s)\hat{\Phi}[f(s)]\cdot p_{u_N}(s)ds, \quad (16)$$

где $U_N(z):R^n\to R^1$.

Тогда, учитывая (14), получим следующее равенство

$$\hat{\Phi}[f(z)]p_{u_N}(z)dz=-U_N(z)\varphi_N(z).$$

Тогда уравнение (13) перепишется в виде

$$\Delta\varphi_N+q_N^2\varphi_N=U_N\varphi_N,\ N=1,2,3... \quad (17)$$

Уравнение (17) при определенном способе выбора констант q_N и функций U_N представляет собой уравнение Шредингера на стационарные состояния системы с «потенциальной энергией» $U_N(z)$.

Выберем теперь функцию $U_N(z)$ на каждом шаге $N = 1,2,3,\ldots$ таким образом, чтобы обеспечить ускорение сходимости конструируемых методов оптимизации. Для этого потребуем выполнения «соотношения оптимальности» для функций $\varphi_N(z)$

$$\hat{\Phi}[f(z)] = \varphi_N(z). \qquad (18)$$

Тогда из формул (14)–(16) получим $U_N(z) = p_{u_N}(z)_N$. Соотношение оптимальности (18) означает, что точки, где расположены экстремумы функции $\hat{\Phi}[f(z)]$, совпадают с точками расположения экстремумов функции $\varphi_N(z)$ для всех N. Соотношение (18) может выполняться лишь в окрестностях точек, где функция $\hat{\Phi}[f(z)]$ обладает для этого достаточной гладкостью. В других точках пространства R^n соотношение (18) является оптимальным сглаживанием функции $\hat{\Phi}[f(z)]$ [7]. Таким образом, для ускорения сходимости конструируемых методов оптимизации предлагается выбирать степени свободы $p_{u_N}(z)$ в (14) как решение интегрального уравнения (16) с учетом «соотношения оптимальности» (18).

В результате получим интегральное уравнение Фредгольма для выбора оптимальной функции $p_u^*(s)$

$$\hat{\Phi}[f(z)] = \int_{R^N} \mathrm{E}_q(z,s)\hat{\Phi}[f(s)]p_u^*(s)ds. \qquad (19)$$

Отметим, что получение решения $p_u^*(s)$ уравнения (19) требует привлечения входной информации лишь о значениях оптимизируемой функции $\hat{\Phi}[f(z)]$. Затем функция $p_u^*(s)$ подставляется в (14) и получается оптимальный потенциал $\varphi^*(z)$ вида

$$\varphi^*(z) = -\int_{R^N} \mathrm{E}_q(z,s)\hat{\Phi}[f(s)]p_u^*(s)ds. \qquad (20)$$

4. МЕТОДЫ ОПТИМИЗАЦИИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ 1–2

Многослойный персептрон (2) настраивается процессом (3) , где матрицы θ_N, задающие «направления матричного движения», определяются своими условными математическими ожиданиями $y_N(z)$ (9) или (12) на каждом шаге $N = 1,2,3\ldots$

Из соотношения (9) следует, что

$$y_N(z) = p^{-1}_{W_{N-1}Y_{N-1}}(z)\nabla_z v_N(z),$$

где $v_N(z)$ задается формулой (11). Из соотношения (12) следует другое решение

$$y^*(z) = (p_u^*(z))^{-1}\nabla_z\varphi^*(z),$$

где функция $p_u^*(z)$ является решением уравнения (19), а функция $\varphi^*(z)$ задается формулой (20). Решение $y^*(z)$ используется с целью получения методов оптимизации, обладающих быстрой сходимостью [7, 8].

Настройка констант уровня c_N проводится в соответствии с условием «неразрывности» и имеет следующий вид [4, 5, 11]

$$\begin{cases} c_N = \int_{R^N} \hat{\Phi}[f(s)]p_{u_N}(s)ds = M[\hat{\Phi}(f(u_N))], \\ c^* = \int_{R^N} \hat{\Phi}[f(s)]p_u^*(z)ds = M[\hat{\Phi}(f(u))], \\ \quad N = 1,2,3\ldots \end{cases} \qquad (21)$$

Уравнения «типа регрессии» (21) могут быть решены, например, следующим образом

$$\begin{cases} c_{N,K} = c_{N,K} + \dfrac{1}{k}\{\hat{\Phi}[f(s_{N,K})]\} - c_{N,K-1}, \\ c_K^* = c_{K-1}^* + \dfrac{1}{k}\{\hat{\Phi}[f(s_K^*)] - c_{K-1}^*\}, \end{cases} \qquad (22)$$

где $(k = 1,2,\ldots)$, $|c_{1,0}| < \infty$, $|c_0| < \infty$.

В формулах (22) через $c_{N,K}$ и c_K^* обозначены оценки констант уровня c_N и c^* соответственно, через $s_{N,K}$ и s_K^* обозначены реализации случайных векторов u_N и u^* с плотностями распределений $p_{u_N}(s)$ и $p_u^*(s)$, $k = 1,2,\ldots$ для каждого N.

Аппроксимируем условное математическое ожидание одной реализацией соответствующего случайного вектора

$$y_N(z) = M[\theta_N Y_{N-1} / W_{N-1}Y_{N-1} = z] \approx \theta_N Y_{N-1}.$$

Тогда

$$\theta_N Y_{N-1} \approx \nabla_z v_N(z) p^{-1}_{W_{N-1}Y_{N-1}}(z), \qquad (23)$$

где $z = W_{N-1}Y_{N-1}$.

Соответственно получим

$$\theta_N^* Y_{N-1} \approx \nabla_z\varphi^*(z)[p^{-1}_{W^*_{N-1}Y_{N-1}}(z)]^{-1}, \qquad (24)$$

где $z = W^*_{N-1}Y_{N-1}$.

Из соотношений (23), (24) получаются матрицы θ_N и θ^*_N вида

$$\begin{cases} \theta_N \approx [Y_N^T Y_N]^{-1}[Y_N^T \nabla_z v_N(z)] p^{-1}_{W_{N-1}Y_{N-1}}(z), \\ \theta^*_N \approx [Y_N^T Y_N]^{-1}[Y_N^T \nabla_z \varphi^*(z)](p^{-1}_{W^*_{N-1}Y_{N-1}}(z))^{-1}. \end{cases} \quad (25)$$

Таким образом, получены следующие формулы, позволяющие организовать оптимизационную процедуру

$$\text{а) } W_N = W_{N-1} + \varepsilon_N [Y_N^T Y_N]^{-1} \times \times [Y_N^T \nabla_z v_N(z)] p^{-1}_{W_{N-1}Y_{N-1}}(z), \quad (26)$$

где $z = W_{N-1}Y_{N-1}$, $Y_i = f(W_{i-1}Y_{i-1})$, $(i = 1, 2, 3, \ldots)$; $\|M(W_1)\| < \infty$; потенциал $v_N(z)$ задается формулой (11),

$$\nabla_z v_N(z) = \int_{R^N} p_{u_N}(s) \nabla_z E(z, s)[\Phi(f(s)) - c_N] ds;$$

c_N определяется процедурой (22);

$$\text{б) } W^*_N = W^*_{N-1} + \varepsilon_N [Y_N^T Y_N]^{-1} \times \times [Y_N^T \nabla_z \varphi^*(z^*)] p^{-1}_{W^*_{N-1}Y_{N-1}}(z^*), \quad (27)$$

где $z^* = W^*_{N-1}Y_{N-1}$, $Y_i = f(W^*_i Y_{i-1})$ $(i = 1, 2, 3, \ldots)$; $\|M(W^*_1)\| < \infty$; оптимальный потенциал $\varphi^*(z)$ задан формулой (20);

$$\nabla_z \varphi^*(z) = \int_{R^N} \nabla_z E_q(z, s) p^*_u(s)[\hat{\Phi}(f(s)) - c^*] ds;$$

c^* определяется процедурой (22).

Выбор числовой последовательности ε_N в рамках условий процедуры (3) может производиться различными способами [4, 5, 3, 11].

В качестве примера в [11] был предложен следующий способ

$$\varepsilon_{N+1} = \begin{cases} \varepsilon_N \gamma_1, \text{ если } \hat{\Phi}(f(z_N)) < \hat{\Phi}(f(z_{N-1})), \\ \varepsilon_N \gamma_2, \text{ если } \hat{\Phi}(f(z_N)) \geq \hat{\Phi}(f(z_{N-1})), \end{cases}$$

где $0 < \gamma_1 < 1$; $0 < \gamma_2 < 1$; γ_1 выбирается эмпирически; γ_2 выбирается с учетом γ_1 из соотношения $\ln \gamma_2 = \lambda \ln \gamma_1$, где $\lambda < 0$.

Последовательность ε_N должна с необходимостью удовлетворять (3).

Условия сходимости и другие свойства методов оптимизации типа (26), (27) рассматривались в работах [7, 11, 8, 6].

ЗАКЛЮЧЕНИЕ

Основными преимуществами предлагаемых методов оптимизации (26), (27) перед традиционными, согласно [7, 11, 5, 8, 6, 4], являются их возможности, связанные с заменой функции $\hat{\Phi}$ на потенциалы v и φ. Это позволило сконструировать градиентные методы (26) и (27) в негладкой задаче оптимизации (1)–(2). Здесь оптимальная настройка многослойного персептрона в смысле решения задачи (2) осуществляется процессом перестройки матриц W_N и W^*_N. Остальные соотношения являются необходимыми для построения таких матричных процессов. Структура многослойного персептрона (2) входит в (26) и (27) через аргументы z и z^* и плотности распределения $p_{W_{N-1}Y_{N-1}}(z)$ и $p^*_{W_{N-1}Y_{N-1}}(z)$. Поскольку при построении (27) использовалось соотношение оптимальности (18), то возможно достижение высокой скорости сходимости оценок (27) к своим предельным значениям.

В методах (26) и (27) не используется градиент функции $\hat{\Phi}$, поэтому по отношению к ней они являются поисковыми методами оптимизации. Для подхода, развитого в работах [3, 4, 5, 6, 7, 8, 11], характерно, что предлагаемые методы негладкой и нелокальной оптимизации могут использовать лишь информацию о значениях оптимизируемой функции $\hat{\Phi}$. Таким же свойством обладают и методы (26) и (27). Эти обстоятельства лишают предлагаемые методы недостатков, связанных с дифференцированием сигмоидных отображений и обнулением больших массивов матриц W_N, $(n = 1, 2, 3, \ldots)$.

СПИСОК ЛИТЕРАТУРЫ

1. *Каплинский, А. И.* Об одном способе построения рандомизированных алгоритмов оптимизации / А. И. Каплинский, А. Е. Лимарев // Вопр. кибернетики: Случайный поиск в задачах оптимизации, 1978. – М. – С.13–17.

2. *Каплинский, А. И.* Построение рандомизированных алгоритмов оптимизации / А. И. Каплинский, А. Е. Лимарев, Г. Д. Чернышова // Проблемы случайного поиска. – Рига : Изд-во Зинатне, 1980. – Вып. 8. – С. 63–91.

3. *Каплинский, А. И.* Об одном способе построения рандомизированных алгоритмов / А. И. Каплинский, А. М. Песин. – М. : Из-во Автоматика и телемеханика, 1982. – № 12. – С. 65–75.

4. *Каплинский, А. И.* Вариационный под-

ход к построению нелокальных алгоритмов оптимизации: препринт / А. И. Каплинский, А. И. Пропой. – М. : ВНИИСИ, 1986.

5. *Каплинский, А. И.* О градиентной основе нелокального поиска, использующего теорию потенциала / А. И. Каплинский, А. И. Пропой // Задачи и методы оптимального моделирования. – М. : ВНИИСИ, 1989.

6. *Каплинский, А. И.* Нелокальные алгоритмы поиска оптимальных решений / А. И. Каплинский, А. М. Песин // Оптимизация и моделирование в автоматизированных системах, 1991. – Воронеж : Изд-во Воронеж. политехн. ин-та. – С. 18–26.

7. *Песин, А. М.* Построение новых методов оптимизации на основе уравнений квантовой механики и математической физики / А. М. Песин. – Деп. ВИМИ, 28 февраля 1991 г., №ДО8340. – 30 с.

8. *Каплинский, А. И.* О методах нелокального поиска / А. И. Каплинский, А. М. Песин, А. И. Пропой // Модели и методы оптимизации. – М. : ВНИИСИ, 1991. – Вып. 13. – С. 35–46.

9. Cybenko, G. Approximation by superpositions of a sigmoidal functions. – Univ. of Illinois, 1988.

10. Физика за рубежом // Сборн. статей. Серия «А». – М. : Мир, 1991.

11. *Каплинский, А. И.* Конструирование вычислительных алгоритмов нелокального поиска, использующих теорию потенциала / А. И. Каплинский, А. И. Пропой // препринт. – М. : ВНИИСИ, 1990. – 29 с.

12. *Hunt, K. J.* Neural networks for control systems – A Survey / K. J. Hunt, D. Sbarbaro, R. Zbikowski and P. J. Gawthrop // Automatica, 1992. – Vol. 28, No 6. – P. 1083–1112.

Каплинский А. И. – канд. техн. наук, доцент, доцент кафедры технической кибернетики и автоматического регулирования факультета ПММ Воронежского государственного университета.

Kaplinskyi A. I. – Ph.D., Associate Professor, Department of Technical Cybernetics and Automatic Regulation, Voronezh State University.

Песин А. М. – канд. техн. наук, доцент, доцент кафедры технической кибернетики и автоматического регулирования факультета ПММ Воронежского государственного университета.

Pesin A. M. – Ph.D., Associate Professor, Department of Technical Cybernetics and Automatic Regulation, Voronezh State University.