Из проведенных тестов получаем следующую зависимость параметров для исследуемой схемы:

$$\alpha = 0.5;\quad \upsilon = 0.33;\quad k = 0.67\cdot E. \qquad (3)$$

Как видно из эксперимента 1.б., точность полученной схемы с оптимальными параметрами невысока даже для сравнительно большого числа частиц. Это связано с получением больших погрешностей на модельных задачах с углами γ, не совпадающими с направлениями ребер регулярной сетки. Из полученных результатов можно предположить, что данную схему следует применять только в задачах, для которых известно, что вектора главных напряжений и направлений ребер сетки [illegible].

Заключение. Задачи, решаемые в современном компьютерном моделировании, постоянно усложняются за счет совершенствования аппаратной и программной базы параллельных вычислительных систем. В данной работе предлагается способ задействовать вычислительные ресурсы также и для исследования и калибровки схем, дальнейшее использование которых непосредственно определяет качество численного моделирования.

Предлагаемый подход продемонстрирован для структурно простой LS-схемы, использующей регулярную сетку и нормальную линейно деформируемую пружину. Несмотря на простоту, подобные схемы могут иметь практическое применение. Мотивация использования регулярной сетки связана с удобством реализации эффективных алгоритмов с помощью массово-параллельных систем. Кроме того, благодаря специальной структуре сетки, с ребрами, расположенными под фиксированными углами, схема может быть расширена для моделирования анизотропных свойств среды.

Для исследуемой LS-схемы процедуры получения параметров не потребовали больших вычислительных затрат, чего нельзя утверждать для произвольной LS-схемы. Цель дальнейшей работы состоит в получении эффективных методов для исследования и калибровки структурно более сложных LS-схем.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Краснопрошин, В. В. Алгоритм расчета геометрических параметров плоских геометрических трещин / В. В. Краснопрошин, О. Л. Коновалов, В. В. Чайко // Вестник Брестского государственного технического университета. – 2017. – № 5 : Физика, математика, информатика – С. 23–26.
2. Chaiko, V. V. DEM-FVM conjugated parallel solver for hydraulic fracturing / V. V. Chaiko, O. L. Konovalov // 2nd International Discrete Fracture Network Engineering Conference, 20-22 June 2018, Seattle, Washington, USA / Publisher: American Rock Mechanics Association.
3. Theoretical Background // Granular Dynamics, Contact Mechanics and Particle System Simulations: A DEM study / Thornton, Colin – Springer, 2015. – Ch. 2. – P. 13–25.
4. Zhao, G. Development of Micro-Macro Continuum-Discontinuum Coupled Numerical Method // Thesis, Ecole Polytechnique Fédérale de Lausanne – 2010.
5. Gao, H. Numerical simulation of crack growth in an isotropic solid with randomized internal cohesive bonds / H. Gao, P. Klein // Journal of the Mechanics and Physics of Solids – Vol. 46 – Is. 2 – 1998. – P. 187–218.
6. Etzmuss, O. Deriving a particle system from continuum mechanics for the animation of deformable objects / O. Etzmuss, J. Gross and W. Strasser / IEEE Transactions on Visualization and Computer Graphics. – 2003. – Vol. 9 – No. 4 – P. 538–550.
7. Liu, Y. Viscoelastic model for discrete element simulation of asphalt mixtures / Y. Liu, Q. Dai, Z. You // Journal of Engineering Mechanics – 2009. – Vol. 135 – Is 4. – P. 324–333.
8. Zhang, Z. A new quasi-continuum constitutive model for crack growth in an isotropic solid / Z. Zhang, X. Ge // European Journal of Mechanics - A/Solids – Vol. 24 – Is. 2 – 2005 – P. 243–252.
9. Baker, G. An exact solution to the two-dimensional elasticity problem with rectangular boundaries under arbitrary edge forces / G. Baker, M. N. Pavlović, N. Tahan // Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences. – 1993. – Vol. 343 – Is. 1668 – P. 307–335.
10. Sarmadivaleh, M. Numerical and experimental investigation of the interaction of natural and propagated hydraulic fracture / H. Fatahi, M. Hossain, M. Sarmadivaleh // Journal of Natural Gas Science and Engineering – Vol. 37 – 2017 – P. 409–424.
11. Fu, P. Generalized displacement correlation method for estimating stress intensity factors / P. Fu, S. Johnson, R. Settgast, C. Carrigan // Engineering Fracture Mechanics – Vol. 88 – July 2012 – P. 90–107.

Материал поступил в редакцию 31.01.2020

CHAIKO V. V., KONOVALOV O. L., KRASNOPROSHIN V. V. Computational scheme for modeling of stress-strain state based on lattice spring method

The paper proposes an approach for calibration and verification of the computational scheme based on the lattice spring method. Using the proposed algorithm we study the scheme with the regular grid and normal linear spring. The optimal parameters and accuracy of the scheme are calculated.

УДК 004.021:032.26

Махнист Л. П., Гладкий И. И., Каримова Т. И.

[illegible]

Ведение. Задача обучения нейронной сети прямого распространения состоит в нахождении весовых коэффициентов w_{ij} и порогов T_j нейронной сети, которые минимизируют функцию ошибки сети

$$E\left(w_{11}, w_{21}, \ldots, w_{m1}, T_1, w_{12}, w_{22}, \ldots, w_{m2}, T_2, \ldots, w_{1n}, w_{2n}, \ldots, w_{mn}, T_n\right) = \frac{1}{2}\sum_{j=1}^{n}\left(y_j - t_j\right)^2,$$

где $y_j = F\left(S_j\right)$ – значение функции активации j-го выходного нейрона сети, $S_j = \sum_{i=1}^{m} w_{ij} x_i - T_j$, x_i – выходное значение i-го

Махнист Леонид Петрович, *к. т. н., доцент, заведующий кафедрой высшей математики Брестского государственного технического университета.*

Гладкий Иван Иванович, *старший преподаватель кафедры высшей математики Брестского государственного технического университета.*

Каримова Татьяна Ивановна*, к. физ.-мат. н., доцент кафедры высшей математики Брестского государственного технического университета.*

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

нейрона предыдущего слоя, t_j – ожидаемый выход j -го выходного нейрона ($i = \overline{1, m}$, $j = \overline{1, n}$).

В работе рассматриваются различные подходы к выбору шага обучения нейронной сети, производится их сравнительный анализ с точки зрения сходимости алгоритма обучения с использованием метода наискорейшего спуска.

Выбор шага обучения и оценки скорости сходимости. Введем обозначения:

$\overline{W} = \left(w_{11}, w_{21}, \ldots, w_{m1}, T_1, w_{12}, w_{22}, \ldots, w_{m2}, T_2, \ldots, w_{1n}, w_{2n}, \ldots, w_{mn}, T_n\right)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порогов T_j нейронной сети, а $\overline{W}_j = \left(w_{1j}, w_{2j}, \ldots, w_{mj}, T_j\right)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порога T_j , связанных с j -м выходным нейроном сети, $E\left(\overline{W}\right) = \frac{1}{2}\sum_{j=1}^{n}\left(y_j - t_j\right)^2$ – функция ошибки сети, $E\left(\overline{W}_j\right) = \frac{1}{2}\left(y_j - t_j\right)^2$ – функция ошибки j -го выходного нейрона сети.

Очевидно, что функция ошибки сети $E\left(\overline{W}\right) = \frac{1}{2}\sum_{j=1}^{n}\left(y_j - t_j\right)^2$ и функции ошибок выходных нейронов сети связаны соотношением

$E\left(\overline{W}\right) = \sum_{j=1}^{n} E\left(\overline{W}_j\right)$.

Предполагается, что функция активации нейронной сети F является дважды дифференцируемой функцией.

Обучение нейронной сети с использованием метода наискорейшего спуска состоит в изменении весовых коэффициентов w_{ij} и порогов T_j нейронной сети на каждом шаге обучения $(t+1)$ ($t = 1, 2, \ldots$) в соответствии с формулами:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha_j(t)\frac{\partial E\left(\overline{W}_j(t)\right)}{\partial w_{ij}}, \qquad (1)$$

$$T_j(t+1) = T_j(t) - \alpha_j(t)\frac{\partial E\left(\overline{W}_j(t)\right)}{\partial T_j}, \qquad (2)$$

если шаг обучения $\alpha_j(t)$ выбирается только для весовых коэффициентов w_{ij} и порога T_j , связанных только с j -м выходным нейроном сети, для минимизации функции ошибки сети j -го выходного нейрона $E\left(\overline{W}_j\right)$;

$$w_{ij}(t+1) = w_{ij}(t) - \alpha(t)\frac{\partial E\left(\overline{W}(t)\right)}{\partial w_{ij}}, \qquad (3)$$

$$T_j(t+1) = T_j(t) - \alpha(t)\frac{\partial E\left(\overline{W}(t)\right)}{\partial T_j}, \qquad (4)$$

если шаг обучения $\alpha(t)$ выбирается для всех весовых коэффициентов w_{ij} и порогов T_j для минимизации функции ошибки сети $E\left(\overline{W}\right)$ (например, в [1–3]).

Частные производные первого порядка функции ошибки сети j -го выходного нейрона $E\left(\overline{W}_j\right)$ по соответствующим переменным w_{ij} и T_j определяются соотношениями:

$$\frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}} = \frac{\partial E\left(\overline{W}_j\right)}{\partial S_j}\cdot\frac{\partial S_j}{\partial w_{ij}} = \left(y_j - t_j\right)F'\left(S_j\right)x_i, \quad \frac{\partial E\left(\overline{W}_j\right)}{\partial T_j} = \frac{\partial E\left(\overline{W}_j\right)}{\partial S_j}\cdot\frac{\partial S_j}{\partial T_j} = -\left(y_j - t_j\right)F'\left(S_j\right), \quad (i = \overline{1, m}).$$

Очевидно, что частные производные первого порядка функции ошибки сети $E\left(\overline{W}\right)$ по соответствующим переменным w_{ij} и T_j удовлетворяют соотношениям:

$$\frac{\partial E\left(\overline{W}\right)}{\partial w_{ij}} = \frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}} = \left(y_j - t_j\right)F'\left(S_j\right)x_i, \quad \frac{\partial E\left(\overline{W}\right)}{\partial T_j} = \frac{\partial E\left(\overline{W}_j\right)}{\partial T_j} = -\left(y_j - t_j\right)F'\left(S_j\right), \quad (i = \overline{1, m}, j = \overline{1, n}).$$

Тогда соотношения (1) и (2) можно записать в виде

$$w_{ij}(t+1) = w_{ij}(t) - \alpha_j(t)\left(y_j(t) - t_j\right)F'\left(S_j(t)\right)x_i, \qquad (5)$$

$$T_j(t+1) = T_j(t) + \alpha_j(t)\left(y_j(t) - t_j\right)F'\left(S_j(t)\right), \qquad (6)$$

а соотношения (3) и (4) в виде:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha(t)\left(y_j(t) - t_j\right)F'\left(S_j(t)\right)x_i, \qquad (7)$$

$$T_j(t+1) = T_j(t) + \alpha(t)\left(y_j(t) - t_j\right)F'\left(S_j(t)\right), \tag{8}$$

Введем обозначения $\nabla E\left(\overline{W}_j\right) = \left(\dfrac{\partial E\left(\overline{W}_j\right)}{\partial w_{1j}}, \dfrac{\partial E\left(\overline{W}_j\right)}{\partial w_{2j}}, \ldots, \dfrac{\partial E\left(\overline{W}_j\right)}{\partial w_{mj}}, \dfrac{\partial E\left(\overline{W}_j\right)}{\partial T_j}\right)^T$ – градиент функции ошибки сети j -го выходного нейрона $E\left(\overline{W}_j\right)$.

Заметим, что $\nabla E\left(\overline{W}_j\right) = \left(y_j - t_j\right)F'\left(S_j\right)\cdot\left(x_1, \quad x_2, \quad \ldots \quad x_m, \quad -1\right)^T$.

Тогда соотношения (1), (2) и (3), (4) можно, соответственно, записать в виде:

$$\overline{W}_j(t+1) = \overline{W}_j(t) - \alpha_j(t)\nabla E\left(\overline{W}_j(t)\right), \tag{9}$$

$$\overline{W}_j(t+1) = \overline{W}_j(t) - \alpha(t)\nabla E\left(\overline{W}_j(t)\right). \tag{10}$$

Рассмотрим случай, когда шаг обучения $\alpha_j(t)$ выбирается только для весовых коэффициентов w_{ij} и порога T_j, связанных только с j -м выходным нейроном сети для минимизации функции ошибки сети j -го выходного нейрона $E\left(\overline{W}_j\right)$.

Введем обозначения $\Delta w_{ij}(t) = w_{ij}(t+1) - w_{ij}(t)$, $\Delta T_j(t) = T_j(t+1) - T_j(t)$ ($i = \overline{1, m}$, $j = \overline{1, n}$).

Заметим, что разложение функции $E\left(\overline{W}_j(t+1)\right)$ в ряд Тейлора в точке $\overline{W}_j(t)$, с точностью до частных производных второго порядка, вычисляемых, как и частные производные первого порядка, на шаге t, может быть представлено в виде:

$$\begin{aligned}
E\left(\overline{W}_j(t+1)\right) &= E\left(\overline{W}_j(t)\right) + \sum_{i=1}^{m}\frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}}\Delta w_{ij}(t) + \frac{\partial E\left(\overline{W}_j\right)}{\partial T_j}\Delta T_j(t) + \\
&+ \frac{1}{2}\cdot\sum_{i=1}^{m}\sum_{k=1}^{m}\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial w_{kj}}\Delta w_{ij}(t)\Delta w_{kj}(t) + \frac{1}{2}\cdot\sum_{i=1}^{m}\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial T_j}\Delta w_{ij}(t)\Delta T_j(t) + \\
&+ \frac{1}{2}\cdot\sum_{k=1}^{m}\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j\partial w_{kj}}\Delta T_j(t)\Delta w_{kj}(t) + \frac{1}{2}\cdot\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j^2}\Delta T_j(t)^2 = \\
&= E\left(\overline{W}_j(t)\right) + \sum_{i=1}^{m}\frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}}\Delta w_{ij}(t) + \frac{\partial E\left(\overline{W}_j\right)}{\partial T_j}\Delta T_j(t) + \\
&+ \frac{1}{2}\cdot\sum_{i=1}^{m}\Delta w_{ij}(t)\left(\sum_{k=1}^{m}\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial w_{kj}}\Delta w_{kj}(t) + \frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial T_j}\Delta T_j(t)\right) + \\
&+ \frac{1}{2}\cdot\Delta T_j(t)\left(\sum_{k=1}^{m}\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j\partial w_{kj}}\Delta w_{kj}(t) + \frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j^2}\Delta T_j(t)\right) = \\
&= E\left(\overline{W}_j(t)\right) + \left(\nabla E\left(\overline{W}_j(t)\right), \Delta\overline{W}_j(t)\right) + \frac{1}{2}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\Delta\overline{W}_j(t), \Delta\overline{W}_j(t)\right),
\end{aligned}$$

где $\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\Delta\overline{W}_j(t)$ – произведение матрицы Гессе $\nabla^2 E\left(\overline{W}_j(t)\right)$ функции $E\left(\overline{W}_j(t)\right)$ и вектора $\Delta\overline{W}_j(t)$, $\left(\nabla E\left(\overline{W}_j(t)\right), \Delta\overline{W}_j(t)\right)$ – скалярное произведение вектора $\nabla E\left(\overline{W}_j(t)\right)$ и вектора $\Delta\overline{W}_j(t) = \overline{W}_j(t+1) - \overline{W}_j(t)$, а $\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\Delta\overline{W}_j(t), \Delta\overline{W}_j(t)\right)$ – скалярное произведение вектора $\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\Delta\overline{W}_j(t)$ и вектора $\Delta\overline{W}_j(t)$.

Тогда, учитывая (9), получим

$$\begin{aligned}
E\left(\overline{W}_j(t+1)\right) &= E\left(\overline{W}_j(t)\right) + \left(\nabla E\left(\overline{W}_j(t)\right), -\alpha_j(t)\nabla E\left(\overline{W}_j(t)\right)\right) + \\
&+ \frac{1}{2}\left(-\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\alpha_j(t)\nabla E\left(\overline{W}_j(t)\right), -\alpha_j(t)\nabla E\left(\overline{W}_j(t)\right)\right) = \\
&= E\left(\overline{W}_j(t)\right) - \alpha_j(t)\left(\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right) + \frac{\alpha_j^2(t)}{2}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right).
\end{aligned}$$

Рассмотрим функцию $E\left(\overline{W}_j(t+1)\right) = E\left(\alpha_j(t)\right)$ как функцию аргумента $\alpha_j(t)$.

Заметим, что производная этой функции равна

$$E'\left(\alpha_j(t)\right) = -\left(\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right) + \alpha_j(t)\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot \nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right).$$

Предположим, что вторая производная функции $E\left(\alpha_j(t)\right)$ удовлетворяет неравенству $E''\left(\alpha_j(t)\right) = \left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot \nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right) > 0$ [4].

Тогда наименьшее значение функции $E\left(\alpha_j(t)\right)$ достигается, если для шага обучения $\alpha_j(t)$ выполняется соотношение $\alpha_j(t) = \dfrac{\left(\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)}{\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot \nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)}$ или

$$\alpha_j(t) = \frac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot \nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)}, \qquad (11)$$

где $\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|$ – длина вектора градиента $\nabla E\left(\overline{W}_j(t)\right)$, связанная со скалярным произведением $\left(\nabla E\left(\overline{W}_j(t)\right), \nabla E\left(\overline{W}_j(t)\right)\right)$, так как $\nabla E\left(\overline{W}_j\right) = \left(y_j - t_j\right)F'\left(S_j\right)\cdot\left(x_1,\ x_2,\ \ldots\ x_m,\ -1\right)^T$, то $\left\|\nabla E\left(\overline{W}_j\right)\right\|^2 = \left(y_j - t_j\right)^2\left(F'\left(S_j\right)\right)^2\cdot\left(\sum_{i=1}^{m} x_i^2 + 1\right)$ или $\left\|\nabla E\left(\overline{W}_j\right)\right\|^2 = \left(E'\left(\overline{W}_j\right)\right)^2\cdot\left(\sum_{i=1}^{m} x_i^2 + 1\right)$, где $E'\left(\overline{W}_j\right) = \left(y_j - t_j\right)F'\left(S_j\right)$.

Частные производные второго порядка функции ошибки сети j -го выходного нейрона $E\left(\overline{W}_j\right)$ по соответствующим переменным определяются соотношениями:

$$\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial w_{kj}} = \frac{\partial}{\partial w_{kj}}\left(\frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}}\right) = \frac{\partial\left(\left(y_j - t_j\right)F'\left(S_j\right)\cdot x_i\right)}{\partial w_{kj}} = \left(\frac{\partial\left(y_j - t_j\right)}{\partial w_{kj}}F'\left(S_j\right) + \left(y_j - t_j\right)\frac{\partial F'\left(S_j\right)}{\partial w_{kj}}\right)\cdot x_i =$$

$$= \left(\left(F'\left(S_j\right)\right)^2 x_k + \left(y_j - t_j\right)F''\left(S_j\right)x_k\right)\cdot x_i = \left(\left(F'\left(S_j\right)\right)^2 + \left(y_j - t_j\right)F''\left(S_j\right)\right)\cdot x_i x_k,$$

$$\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial T_j} = \frac{\partial}{\partial T_j}\left(\frac{\partial E\left(\overline{W}_j\right)}{\partial w_{ij}}\right) = \frac{\partial\left(\left(y_j - t_j\right)F'\left(S_j\right)\cdot x_i\right)}{\partial T_j} = \left(\frac{\partial\left(y_j - t_j\right)}{\partial T_j}F'\left(S_j\right) + \left(y_j - t_j\right)\frac{\partial F'\left(S_j\right)}{\partial T_j}\right)\cdot x_i =$$

$$= \left(-\left(F'\left(S_j\right)\right)^2 - \left(y_j - t_j\right)F''\left(S_j\right)x_k\right)\cdot x_i = -\left(\left(F'\left(S_j\right)\right)^2 + \left(y_j - t_j\right)F''\left(S_j\right)\right)\cdot x_i,$$

$$\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j^2} = \frac{\partial}{\partial T_j}\left(\frac{\partial E\left(\overline{W}_j\right)}{\partial T_j}\right) = \frac{\partial\left(-\left(y_j - t_j\right)F'\left(S_j\right)\right)}{\partial T_j} = -\left(\frac{\partial\left(y_j - t_j\right)}{\partial T_j}F'\left(S_j\right) + \left(y_j - t_j\right)\frac{\partial F'\left(S_j\right)}{\partial T_j}\right) =$$

$$= -\left(-\left(F'\left(S_j\right)\right)^2 - \left(y_j - t_j\right)F''\left(S_j\right)\right) = \left(F'\left(S_j\right)\right)^2 + \left(y_j - t_j\right)F''\left(S_j\right),\ (i = \overline{1,m},\ j = \overline{1,n}).$$

Следовательно, матрицу Гессе $\nabla^2 E\left(\overline{W}_j\right)$ можно записать в виде:

$$\nabla^2 E\left(\overline{W}_j\right) = E''\left(\overline{W}_j\right)\cdot\begin{pmatrix} x_1^2 & x_1x_2 & \ldots & x_1x_m & -x_1 \\ x_2x_1 & x_2^2 & \ldots & x_2x_m & -x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_mx_1 & x_mx_2 & \ldots & x_m^2 & -x_m \\ -x_1 & -x_2 & \ldots & -x_m & 1 \end{pmatrix},$$

где $E''\left(\overline{W}_j\right) = \left(F'\left(S_j\right)\right)^2 + \left(y_j - t_j\right)F''\left(S_j\right)$.

Тогда произведение матрицы $\nabla^2 E\left(\overline{W}_j(t)\right)$ и вектора $\nabla E\left(\overline{W}_j\right)$ определяется соотношением:

$$\nabla^2 E\left(\overline{W}_j\right)\cdot\nabla E\left(\overline{W}_j\right)=E'\left(\overline{W}_j\right)\cdot E''\left(\overline{W}_j\right)\cdot\begin{pmatrix} x_1^2 & x_1x_2 & \dots & x_1x_m & -x_1 \\ x_2x_1 & x_2^2 & \dots & x_2x_m & -x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_mx_1 & x_mx_2 & \dots & x_m^2 & -x_m \\ -x_1 & -x_2 & \dots & -x_m & 1 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ -1 \end{pmatrix}=E'\left(\overline{W}_j\right)\cdot E''\left(\overline{W}_j\right)\cdot\begin{pmatrix} x_1^3+x_1x_2^2+\ldots+x_1x_m^2+x_1 \\ x_2x_1^2+x_2^3+\ldots+x_2x_m^2+x_2 \\ \vdots \\ x_mx_1^2+x_mx_2^2+\ldots+x_m^3+x_m \\ -x_1^2-x_2^2-\ldots-x_m^2-1 \end{pmatrix}.$$

Следовательно, скалярное произведение вектора $\nabla^2 E\left(\overline{W}_j\right)\cdot\nabla E\left(\overline{W}_j\right)$ и вектора $\nabla E\left(\overline{W}_j\right)$ определяется соотношением:

$$\left(\nabla^2 E\left(\overline{W}_j\right)\cdot\nabla E\left(\overline{W}_j\right),\nabla E\left(\overline{W}_j\right)\right)=\left(E'\left(\overline{W}_j\right)\right)^2\cdot E''\left(\overline{W}_j\right)\cdot\left(\begin{pmatrix} x_1^3+x_1x_2^2+\ldots+x_1x_m^2+x_1 \\ x_2x_1^2+x_2^3+\ldots+x_2x_m^2+x_2 \\ \vdots \\ x_mx_1^2+x_mx_2^2+\ldots+x_m^3+x_m \\ -x_1^2-x_2^2-\ldots-x_m^2-1 \end{pmatrix},\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ -1 \end{pmatrix}\right)=$$

$$=\left(E'\left(\overline{W}_j\right)\right)^2\cdot E''\left(\overline{W}_j\right)\cdot\left(x_1^2\left(\sum_{i=1}^m x_i^2+1\right)+x_2^2\left(\sum_{i=1}^m x_i^2+1\right)+\ldots+x_m^2\left(\sum_{i=1}^m x_i^2+1\right)+\sum_{i=1}^m x_i^2+1\right)=$$

$$=\left(E'\left(\overline{W}_j\right)\right)^2\cdot E''\left(\overline{W}_j\right)\cdot\left(\sum_{i=1}^m x_i^2+1\right)^2.$$

Тогда $\dfrac{\left\|\nabla E\left(\overline{W}_j\right)\right\|^2}{\left(\nabla^2 E\left(\overline{W}_j\right)\cdot\nabla E\left(\overline{W}_j\right),\nabla E\left(\overline{W}_j\right)\right)}=\dfrac{\left(E'\left(\overline{W}_j\right)\right)^2\cdot\left(\sum_{i=1}^m x_i^2+1\right)}{\left(E'\left(\overline{W}_j\right)\right)^2\cdot E''\left(\overline{W}_j\right)\cdot\left(\sum_{i=1}^m x_i^2+1\right)^2}=\dfrac{1}{E''\left(\overline{W}_j\right)\cdot\left(\sum_{i=1}^m x_i^2+1\right)}$, а соотношение (11) можно записать в виде:

$$\alpha_j(t)=\frac{1}{\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)\cdot\left(\sum_{i=1}^m x_i^2+1\right)}. \tag{12}$$

Таким образом, если шаг обучения $\alpha_j(t)$ выбирается только для весовых коэффициентов w_{ij} и порога T_j, связанных только с j-м выходным нейроном сети, для минимизации функции ошибки сети j-го выходного нейрона $E\left(\overline{W}_j\right)$, то изменение весовых коэффициентов w_{ij} и порогов T_j нейронной сети на каждом шаге обучения производится в соответствии с формулами:

$$w_{ij}(t+1)=w_{ij}(t)-\frac{\left(y_j(t)-t_j\right)F'\left(S_j(t)\right)\cdot x_i}{\left(\sum_{i=1}^m x_i^2+1\right)\cdot\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)}, \tag{13}$$

$$T_j(t+1)=T_j(t)+\frac{\left(y_j(t)-t_j\right)F'\left(S_j(t)\right)}{\left(\sum_{i=1}^m x_i^2+1\right)\cdot\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)}. \tag{14}$$

Заметим, что абсолютное изменение функции ошибки сети j-го выходного нейрона $E\left(\overline{W}_j\right)$ на каждом шаге обучения удовлетворяет соотношению:

$$\left|E\left(\overline{W}_j(t+1)\right)-E\left(\overline{W}_j(t)\right)\right|=\left|-\alpha_j(t)\left(\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)+\frac{\alpha_j^2(t)}{2}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)\right|=$$

$$=\left|-\alpha_j(t)\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2+\frac{\alpha_j(t)}{2}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2\right|=\frac{\alpha_j(t)}{2}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2=$$

$$=\frac{\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2\cdot\left(\sum_{i=1}^{m}x_i^2+1\right)}{2\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)\cdot\left(\sum_{i=1}^{m}x_i^2+1\right)}=\frac{\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2}{2\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)},$$

а абсолютное изменение функции ошибки сети $E\left(\overline{W}\right)$ на каждом шаге обучения удовлетворяет соотношению:

$$\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=\left|\sum_{j=1}^{n}\left(E\left(\overline{W}_j(t+1)\right)-E\left(\overline{W}_j(t)\right)\right)\right|=\frac{1}{2}\sum_{j=1}^{n}\alpha_j(t)\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2=\frac{1}{2}\sum_{j=1}^{n}\frac{\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2}{\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)}.$$

Теперь рассмотрим случай, когда шаг обучения $\alpha(t)$ выбирается для всех весовых коэффициентов w_{ij} и порогов T_j для минимизации функции ошибки сети $E\left(\overline{W}\right)$.

Разложение функции $E\left(\overline{W}(t+1)\right)$ в ряд Тейлора в точке $\overline{W}(t)$, с точностью до частных производных второго порядка, аналогично, как и для функций $E\left(\overline{W}_j(t+1)\right)$, может быть представлено в виде:

$$E\left(\overline{W}(t+1)\right)=E\left(\overline{W}(t)\right)+\left(\nabla E\left(\overline{W}(t)\right),\Delta\overline{W}(t)\right)+\frac{1}{2}\left(\nabla^2 E\left(\overline{W}(t)\right)\cdot\Delta\overline{W}(t),\Delta\overline{W}(t)\right),$$

где $\nabla E\left(\overline{W}\right)=\left(\nabla E\left(\overline{W}_1\right)^T,\nabla E\left(\overline{W}_2\right)^T,\ldots,\nabla E\left(\overline{W}_n\right)^T\right)^T$ – градиент функция ошибки сети $E\left(\overline{W}\right)$,

$\Delta\overline{W}=\left(\Delta\overline{W}_1{}^T,\Delta\overline{W}_2{}^T,\ldots,\Delta\overline{W}_n{}^T\right)^T$, $\nabla^2 E\left(\overline{W}(t)\right)$ – матрица Гессе $\nabla^2 E\left(\overline{W}\right)$ функции $E\left(\overline{W}\right)$.

Так как $\dfrac{\partial^2 E\left(\overline{W}\right)}{\partial w_{kj}\partial w_{ij}}=\dfrac{\partial^2 E\left(\overline{W}\right)}{\partial w_{ij}\partial w_{kj}}=\dfrac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial w_{kj}}=\left(\left(F'\left(S_j\right)\right)^2+\left(y_j-t_j\right)F''\left(S_j\right)\right)\cdot x_i x_k$,

$$\frac{\partial^2 E\left(\overline{W}\right)}{\partial T_j\partial w_{ij}}=\frac{\partial^2 E\left(\overline{W}\right)}{\partial w_{ij}\partial T_j}=\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial w_{ij}\partial T_j}=-\left(\left(F'\left(S_j\right)\right)^2+\left(y_j-t_j\right)F''\left(S_j\right)\right)\cdot x_i,$$

$$\frac{\partial^2 E\left(\overline{W}\right)}{\partial T_j^2}=\frac{\partial^2 E\left(\overline{W}_j\right)}{\partial T_j^2}=\left(F'\left(S_j\right)\right)^2+\left(y_j-t_j\right)F''\left(S_j\right),\ (i=\overline{1,m},\ j=\overline{1,n}),$$

а другие частные производные второго порядка функции $E\left(\overline{W}\right)$ равны нулю, то матрица $\nabla^2 E\left(\overline{W}\right)$ может быть записана в виде:

$$\nabla^2 E\left(\overline{W}\right)=\begin{pmatrix}\nabla^2 E\left(\overline{W}_1\right) & \mathrm{O} & \mathrm{O} & \ldots & \mathrm{O}\\ \mathrm{O} & \nabla^2 E\left(\overline{W}_2\right) & \mathrm{O} & \ldots & \mathrm{O}\\ \mathrm{O} & \mathrm{O} & \nabla^2 E\left(\overline{W}_3\right) & & \vdots\\ \vdots & \vdots & & & \mathrm{O}\\ \mathrm{O} & \mathrm{O} & \ldots & \mathrm{O} & \nabla^2 E\left(\overline{W}_n\right)\end{pmatrix},$$

где O – матрица размерности $(m+1)\times(m+1)$, все элементы которой равны нулю.

Тогда, учитывая, что $E\left(\overline{W}(t)\right)=\sum_{j=1}^{n}E\left(\overline{W}_j(t)\right)$ и $\left(\nabla E\left(\overline{W}(t)\right),\Delta\overline{W}(t)\right)=\sum_{j=1}^{n}\left(\nabla E\left(\overline{W}_j(t)\right),\Delta\overline{W}_j(t)\right)$, получим соотношение:

$$E\left(\overline{W}(t+1)\right)=\sum_{j=1}^{n}E\left(\overline{W}_j(t)\right)+\sum_{j=1}^{n}\left(\nabla E\left(\overline{W}_j(t)\right),\Delta\overline{W}_j(t)\right)+\frac{1}{2}\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\Delta\overline{W}_j(t),\Delta\overline{W}_j(t)\right).$$

Учитывая (10), получим

$$E\left(\overline{W}(t+1)\right)=\sum_{j=1}^{n}E\left(\overline{W}_j(t)\right)-\alpha(t)\sum_{j=1}^{n}\left(\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)+\frac{\alpha^2(t)}{2}\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right).$$

Рассмотрим функцию $E\left(\overline{W}(t+1)\right)=E\left(\alpha(t)\right)$ как функцию аргумента $\alpha(t)$.

Заметим, что производная этой функции равна

$$E'\left(\alpha_j(t)\right)=-\sum_{j=1}^{n}\left(\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)+\alpha(t)\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right).$$

Предположим, что вторая производная функции $E\left(\alpha(t)\right)$ удовлетворяет неравенству $E''\left(\alpha(t)\right)=\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)>0$ [4].

Тогда наименьшее значение функции $E\left(\alpha(t)\right)$ достигается, если для шага обучения $\alpha(t)$ выполняется соотношение

$$\alpha(t)=\frac{\sum_{j=1}^{n}\left(\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)}{\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)}\ \text{или}\ \alpha(t)=\frac{\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)},\qquad(15)$$

Так как $\left\|\nabla E\left(\overline{W}_j\right)\right\|^2=\left(y_j-t_j\right)^2\left(F'\left(S_j\right)\right)^2\cdot\left(\sum_{i=1}^{m}x_i^2+1\right)$ и $\left(\nabla^2 E\left(\overline{W}_j\right)\cdot\nabla E\left(\overline{W}_j\right),\nabla E\left(\overline{W}_j\right)\right)=$

$=\left(y_j-t_j\right)^2\left(F'\left(S_j\right)\right)^2\left(\left(F'\left(S_j\right)\right)^2+\left(y_j-t_j\right)F''\left(S_j\right)\right)\cdot\left(\sum_{i=1}^{m}x_i^2+1\right)^2$, то соотношение (15) можно записать в виде:

$$\alpha(t)=\frac{\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2}{\left(\sum_{i=1}^{m}x_i^2+1\right)\cdot\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)}.\qquad(16)$$

Заметим, что абсолютное изменение функции ошибки сети $E\left(\overline{W}\right)$ на каждом шаге обучения удовлетворяет соотношению:

$$\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=$$

$$=\left|-\alpha(t)\left(\nabla E\left(\overline{W}(t)\right),\nabla E\left(\overline{W}(t)\right)\right)+\frac{\alpha^2(t)}{2}\left(\nabla^2 E\left(\overline{W}(t)\right)\cdot\nabla E\left(\overline{W}(t)\right),\nabla E\left(\overline{W}(t)\right)\right)\right|=$$

$$=\left|-\alpha(t)\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2+\frac{\alpha(t)}{2}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2\right|=\frac{\alpha(t)}{2}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2=$$

$$=\frac{\left(\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2\right)^2}{2\cdot\sum_{j=1}^{n}\left(y_j(t)-t_j\right)^2\left(F'\left(S_j(t)\right)\right)^2\left(\left(F'\left(S_j(t)\right)\right)^2+\left(y_j(t)-t_j\right)F''\left(S_j(t)\right)\right)}.$$

Установим взаимосвязь между величинами $\alpha_1(t)$, $\alpha_2(t)$,..., $\alpha_n(t)$ и $\alpha(t)$.

Введем обозначение $q_j(t)=\dfrac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}=\dfrac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2}$.

Очевидно, что $\sum_{j=1}^{n}q_j(t)=1$.

Так как, учитывая (11) $\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)=\frac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\alpha_j(t)}$, то $\alpha(t)=\frac{\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\sum_{j=1}^{n}\left(\nabla^2 E\left(\overline{W}_j(t)\right)\cdot\nabla E\left(\overline{W}_j(t)\right),\nabla E\left(\overline{W}_j(t)\right)\right)}=$

$=\frac{\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\sum_{j=1}^{n}\frac{\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2}{\alpha_j(t)}}=\frac{1}{\sum_{j=1}^{n}\frac{q_j(t)}{\alpha_j(t)}}$, т. е. $\alpha(t)$ – взвешенное среднее гармоническое величин $\alpha_1(t)$, $\alpha_2(t)$,..., $\alpha_n(t)$ [5].

Полученное соотношение можно пояснить, например, следующим образом: если путь $\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2=\sum_{j=1}^{n}\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2$ разбит на участки $\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2$, скорость (шаг обучения) на каждом из которых постоянна и равна $\alpha_j(t)$, то средняя скорость (шаг обучения) $\alpha(t)$ по всему пути $\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2$ будет равна взвешенному среднему гармоническому скоростей (шагов обучения) $\alpha_j(t)$ [7].

Сравним абсолютные изменения функции ошибки сети $E\left(\overline{W}\right)$ в рассмотренных двух случаях [6]:

- шаг обучения $\alpha_j(t)$ выбирается только для весовых коэффициентов w_{ij} и порога T_j, связанных только с j-м выходным нейроном сети для минимизации функции ошибки сети j-го выходного нейрона $E\left(\overline{W}_j\right)$;
- шаг обучения $\alpha(t)$ выбирается для всех весовых коэффициентов w_{ij} и порогов T_j для минимизации функции ошибки сети $E\left(\overline{W}\right)$.

В первом случае получим, что абсолютное изменение функции ошибки сети $E\left(\overline{W}\right)$ удовлетворяет соотношению:

$$\left|\Delta_1 E\left(\overline{W}(t)\right)\right|=\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=\left|\sum_{j=1}^{n}\left(E\left(\overline{W}_j(t+1)\right)-E\left(\overline{W}_j(t)\right)\right)\right|=\frac{1}{2}\sum_{j=1}^{n}\alpha_j(t)\left\|\nabla E\left(\overline{W}_j(t)\right)\right\|^2=$$

$$=\frac{1}{2}\left(\sum_{j=1}^{n}q_j(t)\alpha_j(t)\right)\cdot\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2,$$

а во втором случае – соотношению:

$$\left|\Delta_2 E\left(\overline{W}(t)\right)\right|=\left|E\left(\overline{W}(t+1)\right)-E\left(\overline{W}(t)\right)\right|=\frac{\alpha(t)}{2}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2=\frac{1}{2\sum_{j=1}^{n}\frac{q_j(t)}{\alpha_j(t)}}\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2.$$

Рассмотрим неравенство Йенсена.

Пусть функция $f(x)$ является выпуклой на некотором промежутке X и числа q_1, q_2,..., q_n таковы, что q_1, q_2,..., q_n >0 и $q_1+q_2+\ldots+q_n=1$.

Тогда каковы бы ни были числа x_1, x_2,..., x_n из промежутка X, выполняется неравенство:

$f(q_1x_1+q_2x_2+\ldots+q_nx_n)\le q_1f(x_1)+q_2f(x_2)+\ldots+q_nf(x_n)$ или $f\left(\sum_{j=1}^{n}q_jx_j\right)\le\sum_{j=1}^{n}q_jf(x_j)$.Пусть $f(x)=\frac{1}{x}$ (выпуклая функция) и $X=(0;+\infty)$.

Тогда имеем $\frac{1}{\sum_{j=1}^{n}q_jx_j}=f\left(\sum_{j=1}^{n}q_jx_j\right)\le\sum_{j=1}^{n}q_jf(x_j)=\sum_{j=1}^{n}\frac{q_j}{x_j}$ или получаем, что взвешенное среднее гармоническое не превосходит среднего арифметического взвешенного:

$$\frac{1}{\sum_{j=1}^{n}\frac{q_j}{x_j}}\le\sum_{j=1}^{n}q_jx_j\,.$$

Следовательно, при выполнении условия $\min_j\left(\left(y_j-t_j\right)F'\left(S_j\right)\right)^2\cdot\left(\left(F'\left(S_j\right)\right)^2+\left(y_j-t_j\right)F''\left(S_j\right)\right)>0$, получим

$$\max\left(\left|\Delta_1 E\left(\overline{W}(t)\right)\right|,\left|\Delta_2 E\left(\overline{W}(t)\right)\right|\right)=\max\left(\sum_{j=1}^{n} q_j(t)\alpha_j(t),\ \frac{1}{\sum_{j=1}^{n}\frac{q_j(t)}{\alpha_j(t)}}\right)\cdot\frac{\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2}{2}=\left(\sum_{j=1}^{n} q_j(t)\alpha_j(t)\right)\cdot\frac{\left\|\nabla E\left(\overline{W}(t)\right)\right\|^2}{2}.$$

Заключение. При выборе шага обучения использование соотношений (12) будет обеспечивать лучшую скорость сходимости обучения нейронной сети с использованием метода наискорейшего спуска по сравнению с использованием соотношения (16).

В статье получены формулы для вычисления шага обучения и ограничения для их использования. Предложенная методика может быть использована в алгоритме обратного распространения ошибки обучения нейронной сети.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Golovko, M. Multilayer neural networks training methodic / M. Golovko, L. Makhnist, N. Maniakov // Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2003) : Proceedings, Lviv, Ukraine, 8–10 Sept. 2003. – Lviv, 2003. – P. 185–190.
2. Maniakov, N. Traing algorithm for forecasting multilayer neural network / N. Maniakov, L. Makhnist, V. Rubanov // Pattern Recognition and Information Processing : Proceedings of The Seventh International Conferences (PRIP'2003), Minsk, Republic of Belarus, 21–23 May 2003 : in 2 vol. – Minsk, 2003. – Vol. 1. – P. 26–30.
3. Makhnist, L. Some Methods of Adaptive Multilayer Neural Network Training / L. Makhnist, N. Maniakov // International Journal of Computing. – 2004. – Vol. 3. – P. 99–106.
4. Maxnist, L. Convergence Analysis of Neural Networks Training Based on Steepest Descent Method / L. Maxnist, A. Doudkin, V. Golovko // Pattern Recognition and Information Processing (PRIP'2007) : Proceedings of the Ninth International Conference, Minsk, Republic of Belarus, 22–24 May 2007 : in 2 vol. – Minsk, 2007. – Vol. 1. – P. 285–289.
5. Гладкий, И. И. О выборе шага обучения искусственных нейронных сетей прямого распространения / И. И. Гладкий, Л. П. Махнист, В. С. Рубанов, Т. Ю. Юхимук // Вычислительные методы, модели и образовательные технологии : сб. материалов VIII Междунар. науч.-практ. конф., Брест, 18 окт. 2019 г. / Брест. гос. ун-т им. А. С. Пушкина ; под общ. ред. А. А. Козинского. – Брест : БрГУ, 2019. – С. 121–122.
6. Махнист, Л. П. О сходимости алгоритмов обучения искусственных нейронных сетей / Л. П. Махнист, А. В. Санюкевич, В. П. Черненко, М. М. Юхимук // Вычислительные методы, модели и образовательные технологии : сб. материалов VIII Междунар. науч.-практ. конф., Брест, 18 окт. 2019 г. / Брест. гос. ун-т им. А. С. Пушкина ; под общ. ред. А. А. Козинского. – Брест : БрГУ, 2019. – С. 133–135.
7. Гладкий, И. И. О скорости сходимости и выборе шага обучения искусственных нейронных сетей прямого распространения / И. И. Гладкий, В. В. Будяков, Б. И. Чикалов, Т. Ю. Юхимук // Современные проблемы математики и вычислительной техники : сб. материалов XI Респ. науч. конф. молодых ученых и студентов, Брест, 21–22 нояб. 2019 г. / Брест. гос. техн. ун-т ; редкол.: В. А. Головко (гл. ред.) [и др.]. – Брест : БрГТУ, 2019. – С. 11–14.

26.02.2020

MAKHNIST L. P., GOLOVKO V. A., HLADKI I. I., KARIMOVA T. I. Estimation of the convergence rate and the choice of the learning step of artificial feed forward neural networks

The paper presents various approaches to the choice of the learning step for a feed forward neural network and compares then in terms of the learning algorithm convergence using the steepest descent approach. The equations for calculating the learning step and the restrictions for their use are obtained. The proposed technique can be used in the back propagation algorithm for training a multilayer perceptron.

338.2: 681.3

Введение. В настоящее время отмечается устойчивая тенденция расширения программ международного сотрудничества по изучению планеты Марс. Начиная с 1965 года, земляне отправили к Марсу 45 космических аппаратов. В 2019 году на орбите Марса находились: орбитальный аппарат Mars Odyssey, зонды Mars Reconnaissance Orbiter и MAVEU, робот-марсоход Curiosity, европейская межпланетная станция Mars Express и индийская Mars Orbiter Mission. В 2020 году в рамках международного проекта «ЭкзоМарс» на планету будет отправлена российская посадочная платформа с европейским марсоходом [1].

Критическим фактором выполнения программ по изучению поверхности планеты является скорость передачи информации по радиоканалам космической связи. Оборудование марсохода генерирует огромное количество научных данных, и не всегда их удается отправлять: часть наименее важных данных иногда приходится удалять [2]. Передача информации роботом-марсоходом на станцию космической связи Земли обычно осуществляется через спутник-ретранслятор, находящийся на орбите Марса. Наиболее экономически целесообразным способом связи является передача информации непосредственно с марсохода на Землю и обратно. Однако, в настоящее время, в открытых литературных источниках не представлены расчеты, позволяющие оценить скорость передачи информации с учетом параметров канала связи (максимального расстояния между планетами, мощности передатчиков, коэффициентов приемо/передающих антенн, модуляции сигнала, несущей частоты и др.). Поэтому целью настоящей работы явилась оценка скорости передачи информации космического канала связи марсоход-станция космической связи Земли при максимальной дальности между ними и минимально допустимой мощности передатчика марсохода.