

$a_1 - a_2 = \pm 2r$, и минимальное (отрицательное) значение на границе области $|a_1 - a_2| \geq 2r$, если $a_1 - a_2 = \pm 2ri$. Отсюда и следует утверждение теоремы.

Работа выполнена при поддержке Белорусского республиканского фонда фундаментальных исследований.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Cherkaev A. Variational Methods for Structural Optimization. N.-Y. etc.: Springer Verlag, 2000.
2. Лурье К. А., Черкаев А. В. Эффективные характеристики композиционных материалов и оптимальный дизайн

- структурных элементов // Успехи математических наук, т. 9 (1986), 3-81.
3. Bendse M. P. Optimization of Structural Topology, Shape and Material. Berlin etc.: Springer Verlag, 1995.
4. Buttazzo G. On existence of minimizing domains for some shape optimization problem// ESAIM: Proceedings, v. 3 (1998), 51-64.
5. Mityushev V. V., Rogosin S. V. Constructive Methods for Linear and Nonlinear Boundary Value Problems for Analytic Functions. Boca Raton – London: Chapman & Hall / CRC Press, 2000.
6. Макарук С. Ф. Смешанная краевая задача для аналитических функций со свободной частью границы // Труды института математики НАНБ, т. 9 (2001), 101-104.

УДК 681.324:519.711.7

Маньяков Н.В., Махнист Л.П.

МАТРИЧНАЯ АЛГОРИТМИЗАЦИЯ ОБУЧЕНИЯ МНОГОСЛОЙНЫХ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ ГРАДИЕНТНЫХ МЕТОДОВ

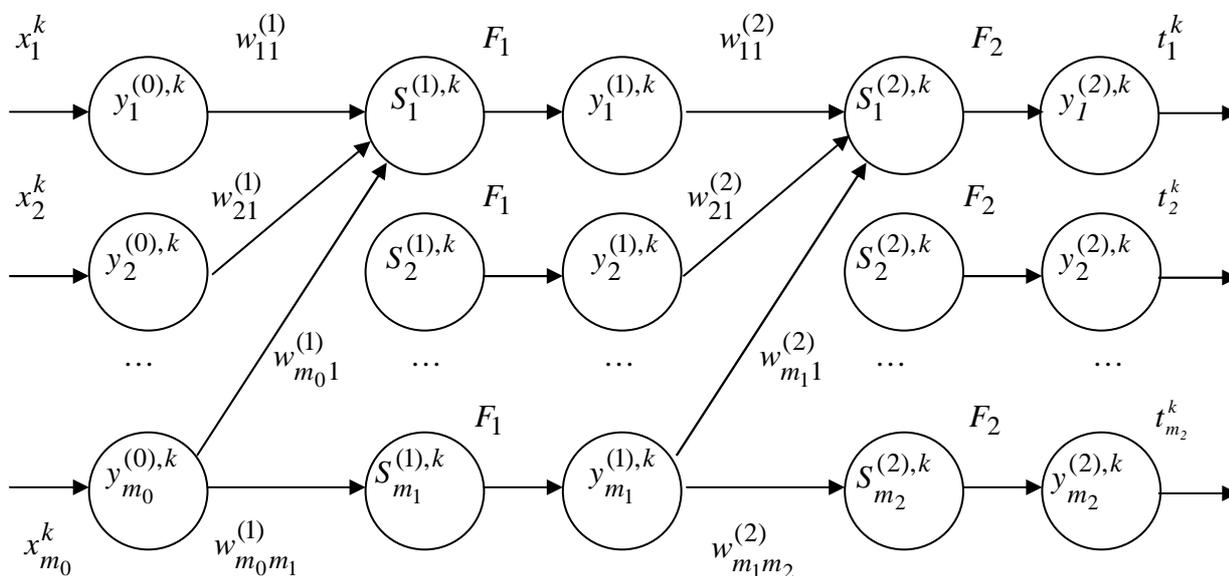


Рис.1. Схема функционирования нейронной сети

1. ПОСТАНОВКА ЗАДАЧИ ОБУЧЕНИЯ ДВУХСЛОЙНОЙ НЕЙРОННОЙ СЕТИ

Рассмотрим двухслойную гетерогенную нейронную сеть, состоящую из m_0 нейронных элементов распределительного слоя, m_1 нейронов скрытого слоя и m_2 – выходного слоя (рис.1).

Для данной сети каждый нейрон распределительного слоя имеет синаптические связи $w_{i_0 i_1}^{(1)}$ ($i_0 = \overline{1, m_0}, i_1 = \overline{1, m_1}$) со всеми нейронами скрытого слоя, а каждый нейрон скрытого слоя имеет синаптические связи $w_{i_1 i_2}^{(2)}$ ($i_1 = \overline{1, m_1}, i_2 = \overline{1, m_2}$) со всеми нейронами выходного слоя. В качестве нейронов скрытого слоя используются

элементы с функцией активации F_1 , в качестве нейронов выходного слоя – с функцией активации F_2 . На вход сети подаются входные образы – векторы

$\overline{x}^k = (x_1^k, \dots, x_{m_0}^k), (k = \overline{1, L})$. Входами распределительного слоя являются значения $y_{i_0}^{(0),k} = x_{i_0}^k$. При этом формируется вектор $Y^{(0),k} = (y_1^{(0),k} \quad y_2^{(0),k} \quad \dots \quad y_{m_0}^{(0),k} \quad -1)^T$.

Выходное значение i_1 -го нейрона скрытого слоя сети для k -го образа определяется соотношением:

$$y_{i_1}^{(1),k} = F_1(S_{i_1}^{(1),k}),$$

Маньяков Николай Владимирович. Ст. преподаватель каф. высшей математики Брестского государственного технического университета.

Махнист Леонид Петрович. К.т.н., доцент каф. высшей математики Брестского государственного технического университета.

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

где

$$S_{i_1}^{(1),k} = \sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(0),k} - T_{i_1}^{(1)}, \quad i_1 = \overline{1, m_1}, \quad k = \overline{1, L}.$$

При этом формируется вектор $Y^{(1),k} = (y_1^{(1),k} \quad y_2^{(1),k} \quad \dots \quad y_{m_1}^{(1),k} \quad -1)^T$.

Выходное значение i_2 -ого нейрона выходного слоя сети для k -ого образа определяется соотношением:

$$y_{i_2}^{(2),k} = F_2(S_{i_2}^{(2),k}),$$

где

$$S_{i_2}^{(2),k} = \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} y_{i_1}^{(1),k} - T_{i_2}^{(2)}, \quad i_2 = \overline{1, m_2}, \quad k = \overline{1, L}.$$

Задача обучения нейронной сети с фиксированными функциями активации [1] состоит в нахождении весовых коэффициентов $w_{i_0 i_1}^{(1)} \quad (i_0 = \overline{1, m_0}, \quad i_1 = \overline{1, m_1})$,

$w_{i_1 i_2}^{(2)} \quad (i_1 = \overline{1, m_1}, \quad i_2 = \overline{1, m_2})$ и порогов нейронных элементов $T_{i_1}^{(1)} \quad (i_1 = \overline{1, m_1})$, $T_{i_2}^{(2)} \quad (i_2 = \overline{1, m_2})$, которые

минимизируют некоторую ошибку сети E_S , как отклонение выходных значений $y_{i_2}^{(2),k}$ от эталонных значений $t_{i_2}^k$ - i_2 -ого нейрона сети для k -ого образа. В качестве ошибки будем рассматривать усредненное по количеству образов «квадратичное отклонение»

$$E_S = \frac{1}{2L} \sum_{k=1}^L \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k)^2.$$

Матрицы $W^{(1)} = \begin{pmatrix} w_{11}^{(1)} & w_{21}^{(1)} & \dots & w_{m_0 1}^{(1)} \\ w_{12}^{(1)} & w_{22}^{(1)} & \dots & w_{m_0 2}^{(1)} \\ \dots & \dots & \dots & \dots \\ w_{1 m_1}^{(1)} & w_{2 m_1}^{(1)} & \dots & w_{m_0 m_1}^{(1)} \end{pmatrix}_{m_1 \times m_0}$ и

$W^{(2)} = \begin{pmatrix} w_{11}^{(2)} & w_{21}^{(2)} & \dots & w_{m_1 1}^{(2)} \\ w_{12}^{(2)} & w_{22}^{(2)} & \dots & w_{m_1 2}^{(2)} \\ \dots & \dots & \dots & \dots \\ w_{1 m_2}^{(2)} & w_{2 m_2}^{(2)} & \dots & w_{m_1 m_2}^{(2)} \end{pmatrix}_{m_2 \times m_1}$ и векторы

$$\overline{T^{(1)}} = (T_1^{(1)}, T_2^{(1)}, \dots, T_{m_1}^{(1)})^T,$$

$\overline{T^{(2)}} = (T_1^{(2)}, T_2^{(2)}, \dots, T_{m_2}^{(2)})^T$ будем называть приближенным решением или просто решением (по методу наименьших квадратов) системы:

$$F_2 \left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot F_1 \left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(0),k} - T_{i_1}^{(1)} \right) - T_{i_2}^{(2)} \right) = t_{i_2}^k, \quad i_2 = \overline{1, m_2}, \quad k = \overline{1, L}$$

если «квадратичное отклонение»

$$E_S = \frac{1}{2L} \sum_{k=1}^L \sum_{i_2=1}^{m_2} \left(F_2 \left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot F_1 \left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(1),k} - T_{i_1}^{(1)} \right) - T_{i_2}^{(2)} \right) - t_{i_2}^k \right)^2 = \frac{1}{L} \sum_{k=1}^L E_s^{(k)}$$

достигает своего наименьшего значения, где

$$E_s^{(k)} = \frac{1}{L} \sum_{i_2=1}^{m_2} \left(F_2 \left(\sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot F_1 \left(\sum_{i_0=1}^{m_0} w_{i_0 i_1}^{(1)} y_{i_0}^{(1),k} - T_{i_1}^{(1)} \right) - T_{i_2}^{(2)} \right) - t_{i_2}^k \right)^2$$

2. АЛГОРИТМИЗАЦИЯ ОБУЧЕНИЯ ДВУХСЛОЙНОЙ НЕЙРОННОЙ СЕТИ

Рассмотрим схему алгоритмов обучения двухслойной сети с использованием градиентных методов.

В соответствии с градиентными методами изменение весов и порогов сети производится по формулам:

$$w_{j_0 j_1}^{(1)}(t+1) = w_{j_0 j_1}^{(1)}(t) - \alpha^{(1)} \frac{\partial E_s(t)}{\partial w_{j_0 j_1}^{(1)}},$$

$$j_0 = \overline{1, m_0}, \quad j_1 = \overline{1, m_1}$$

$$T_{j_1}^{(1)}(t+1) = T_{j_1}^{(1)}(t) - \alpha^{(1)} \frac{\partial E_s(t)}{\partial T_{j_1}^{(1)}}, \quad j_1 = \overline{1, m_1}, \text{ и}$$

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha^{(2)} \frac{\partial E_s(t)}{\partial w_{j_1 j_2}^{(2)}},$$

$$j_1 = \overline{1, m_1}, \quad j_2 = \overline{1, m_2}$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha^{(2)} \frac{\partial E_s(t)}{\partial T_{j_2}^{(2)}}, \quad j_2 = \overline{1, m_2},$$

где $\frac{\partial E_s}{\partial z} = \frac{\partial \left(\frac{1}{L} \sum_{k=1}^L E_s^{(k)} \right)}{\partial z} = \frac{1}{L} \sum_{k=1}^L \frac{\partial E_s^{(k)}}{\partial z}$ и $\alpha^{(1)}$, $\alpha^{(2)}$ - параметры обучения, выбираемые постоянными или зависящими от t для достижения наименьшего значения ошибки.

Рассмотрим

$$\frac{\partial E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}} = \frac{\partial \left(\sum_{i_2=1}^{m_2} \frac{1}{2} (y_{i_2}^{(2),k} - t_{i_2}^k)^2 \right)}{\partial w_{j_1 j_2}^{(2)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot \frac{\partial y_{i_2}^{(2),k}}{\partial w_{j_1 j_2}^{(2)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \frac{\partial S_{i_2}^{(2),k}}{\partial w_{j_1 j_2}^{(2)}} =$$

$$= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot y_{j_1}^{(1),k} \cdot \delta_{j_2}^{i_2}$$

где

$$\delta_{j_2}^{i_2} = \begin{cases} 1 & , \quad i_2 = j_2 \\ 0 & , \quad i_2 \neq j_2 \end{cases} \quad - \text{ символ Кронекера.}$$

В матричной форме данная формула примет вид

$$\frac{\partial E_s^{(k)}}{\partial w_{j_1 j_2}^{(2)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot y_{j_1}^{(1),k} \cdot \delta_{j_2}^{i_2} =$$

$$= \varepsilon^k \cdot MF_2 \cdot M_{j_2 j_1}^{(2)} \cdot Y^{(1),k},$$

где

$$\varepsilon^k = \left((y_1^{(2),k} - t_1^k) \quad (y_2^{(2),k} - t_2^k) \quad \dots \quad (y_{m_2}^{(2),k} - t_{m_2}^k) \right),$$

$$MF_2 = \begin{pmatrix} F_2'(S_1^{(2),k}) & 0 & \dots & 0 \\ 0 & F_2'(S_2^{(2),k}) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F_2'(S_{m_2}^{(2),k}) \end{pmatrix} -$$

матрица размерности $m_2 \times m_2$, а матрица $M_{j_2 j_1}^{(2)}$ размерности $m_2 \times (m_1 + 1)$, в которой на пересечении j_2 -ой строки и j_1 -ого столбца расположено число 1, а остальные элементы матрицы равны нулю.

Аналогично

$$\begin{aligned} \frac{\partial E_s^{(k)}}{\partial T_{j_2}^{(2)}} &= \frac{\partial \left(\sum_{i_2=1}^{m_2} \frac{1}{2} (y_{i_2}^{(2),k} - t_{i_2}^k)^2 \right)}{\partial T_{j_2}^{(2)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot \frac{\partial y_{i_2}^{(2),k}}{\partial T_{j_2}^{(2)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \frac{\partial S_{i_2}^{(2),k}}{\partial T_{j_2}^{(2)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot (-1) \cdot \delta_{j_2}^{i_2} \end{aligned}$$

В матричной форме данная формула примет вид

$$\begin{aligned} \frac{\partial E_s^{(k)}}{\partial T_{j_2}^{(2)}} &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot (-1) \cdot \delta_{j_2}^{i_2} = \\ &= \varepsilon^k \cdot MF_1 \cdot M_{j_2(m_1+1)}^{(2)} \cdot Y^{(1),k}. \end{aligned}$$

Тогда формулы для изменения весов и порогов второго слоя в матричной форме примут вид:

$$w_{j_1 j_2}^{(2)}(t+1) = w_{j_1 j_2}^{(2)}(t) - \alpha^{(2)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L \varepsilon^k \cdot MF_2 \cdot M_{j_2 j_1}^{(2)} \cdot Y^{(1),k},$$

$$j_1 = \overline{1, m_1}, \quad j_2 = \overline{1, m_2}$$

$$T_{j_2}^{(2)}(t+1) = T_{j_2}^{(2)}(t) - \alpha^{(2)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L \varepsilon^k \cdot MF_2 \cdot M_{j_2(m_1+1)}^{(2)} \cdot Y^{(1),k},$$

$$j_2 = \overline{1, m_2}$$

Для первого слоя соответствующие компоненты градиента примут вид:

$$\begin{aligned} \frac{\partial E_s^{(k)}}{\partial w_{j_0 j_1}^{(1)}} &= \frac{\partial \left(\sum_{i_2=1}^{m_2} \frac{1}{2} (y_{i_2}^{(2),k} - t_{i_2}^k)^2 \right)}{\partial w_{j_0 j_1}^{(1)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot \frac{\partial y_{i_2}^{(2),k}}{\partial w_{j_0 j_1}^{(1)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \frac{\partial S_{i_2}^{(2),k}}{\partial w_{j_0 j_1}^{(1)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot \frac{\partial y_{i_1}^{(1),k}}{\partial w_{j_0 j_1}^{(1)}} = \end{aligned}$$

$$\begin{aligned} &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \times \\ &\times \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot F_1'(S_{i_1}^{(1),k}) \cdot y_{j_0}^{(0),k} \cdot \delta_{j_1}^{i_1} = \\ &= \varepsilon^k \cdot MF_2 \cdot W^{(2)} \cdot MF_1 \cdot M_{j_1 j_0}^{(1)} \cdot Y^{(0),k}. \\ \frac{\partial E_s^{(k)}}{\partial T_{j_1}^{(1)}} &= \frac{\partial \left(\sum_{i_2=1}^{m_2} \frac{1}{2} (y_{i_2}^{(2),k} - t_{i_2}^k)^2 \right)}{\partial T_{j_1}^{(1)}} = \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot \frac{\partial y_{i_2}^{(2),k}}{\partial T_{j_1}^{(1)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \frac{\partial S_{i_2}^{(2),k}}{\partial T_{j_1}^{(1)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot \frac{\partial y_{i_1}^{(1),k}}{\partial T_{j_1}^{(1)}} = \\ &= \sum_{i_2=1}^{m_2} (y_{i_2}^{(2),k} - t_{i_2}^k) \cdot F_2'(S_{i_2}^{(2),k}) \cdot \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(2)} \cdot F_1'(S_{i_1}^{(1),k}) \cdot (-1) \cdot \delta_{j_1}^{i_1} = \\ &= \varepsilon^k \cdot MF_2 \cdot W^{(2)} \cdot MF_1 \cdot M_{j_1(m_0+1)}^{(1)} \cdot Y^{(0),k}, \end{aligned}$$

$$\text{где } MF_1 = \begin{pmatrix} F_1'(S_1^{(1),k}) & 0 & \dots & 0 \\ 0 & F_1'(S_2^{(1),k}) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F_1'(S_{m_1}^{(1),k}) \end{pmatrix}$$

— матрица размерности $m_1 \times m_1$, а матрица $M_{j_1 j_0}^{(1)}$ размерности $m_1 \times (m_0 + 1)$ состоит из числа на позиции $j_1 j_0$ и нулей в качестве остальных элементов матрицы.

Формулы для изменения весов и порогов первого слоя в матричной форме примут вид:

$$\begin{aligned} w_{j_0 j_1}^{(1)}(t+1) &= w_{j_0 j_1}^{(1)}(t) - \\ &- \alpha^{(1)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L \varepsilon^k \cdot MF_2 \cdot W^{(2)} \cdot MF_1 \cdot M_{j_1 j_0}^{(1)} \cdot Y^{(0),k}, \\ T_{j_1}^{(1)}(t+1) &= T_{j_1}^{(1)}(t) - \\ &- \alpha^{(1)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L \varepsilon^k \cdot MF_2 \cdot W^{(2)} \cdot MF_1 \cdot M_{j_1(m_0+1)}^{(1)} \cdot Y^{(0),k}, \end{aligned}$$

$$\text{где } j_0 = \overline{1, m_0}, \quad j_1 = \overline{1, m_1}.$$

Заметим, что аналогичным образом можно получить соответствующие соотношения и для многослойных гетерогенных нейронных сетей.

3. АЛГОРИТМЫ ОБУЧЕНИЯ МНОГОСЛОЙНОЙ НЕЙРОННОЙ СЕТИ

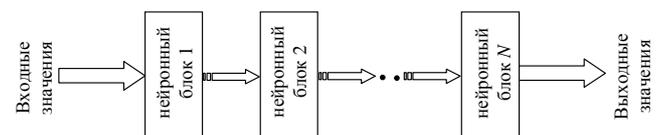


Рис.2. Блочное представление многослойной нейронной сети.

Рассмотрим многослойную гетерогенную нейронную сеть, состоящую из N нейронных блоков (рис.2), каждый из которых имеет структуру, представленную на рис. 3.

Входными значениями для каждого нейронного блока являются выходы предыдущего; для первого – последовательность входных образов $\bar{x}^k = (x_1^k, \dots, x_{m_0}^k)$, ($k = \overline{1, L}$).

Выходное значение i_n -ого нейрона n -ого блока сети для k -ого образа определяется рекуррентным соотношением:

$$y_{i_n}^{(n),k} = F_n(S_{i_n}^{(n),k}),$$

где

$$S_{i_n}^{(n),k} = \sum_{i_{n-1}=1}^{m_{n-1}} w_{i_{n-1}i_n}^{(n)} y_{i_{n-1}}^{(n-1),k} - T_{i_n}^{(n)}, \quad i_n = \overline{1, m_n}, \quad k = \overline{1, L}.$$

При этом формируется вектор $Y^{(n),k} = (y_1^{(n),k}, y_2^{(n),k}, \dots, y_{m_n}^{(n),k}, -1)^T$.

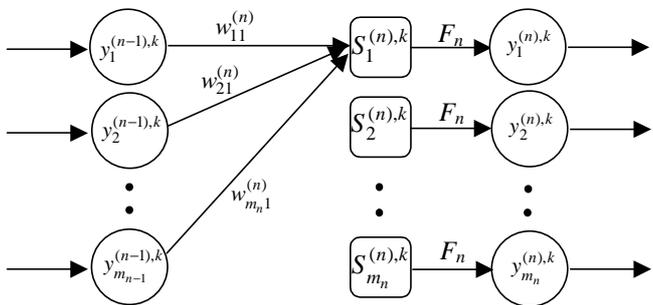


Рис.3 – Архитектура n -ого блока многослойной нейронной сети.

Задача обучения данной многослойной гетерогенной нейронной сети состоит в нахождении матриц весовых коэффициентов

$$W^{(n)} = \begin{pmatrix} w_{11}^{(n)} & w_{21}^{(n)} & \dots & w_{m_{n-1}1}^{(n)} \\ w_{12}^{(n)} & w_{22}^{(n)} & \dots & w_{m_{n-1}2}^{(n)} \\ \dots & \dots & \dots & \dots \\ w_{1m_n}^{(n)} & w_{2m_n}^{(n)} & \dots & w_{m_{n-1}m_n}^{(n)} \end{pmatrix}_{m_n \times m_{n-1}}$$

и столбцов порогов $\overline{T^{(n)}} = (T_1^{(n)}, T_2^{(n)}, \dots, T_{m_n}^{(n)})^T$,

$n = \overline{1, N}$, которые минимизируют некоторую ошибку сети E_s , как отклонение выходных значений сети $y_{i_n}^{(N),k}$ от эталонных $t_{i_n}^k$ – i_n -ого нейрона сети для k -ого образа. В качестве ошибки рассматривается усредненное по количеству образов «квадратичное отклонение»

$$E_s = \frac{1}{2L} \sum_{k=1}^L \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k)^2.$$

Теорема. Модификация синаптических связей и порогов многослойной гетерогенной нейронной сети (рис.2) производится в соответствии с формулами:

$$w_{j_{n-1}j_n}^{(n)}(t+1) = w_{j_{n-1}j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k}$$

$$T_{j_n}^{(n)}(t+1) = T_{j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n (m_{n-1}+1)}^{(n)} \cdot Y^{(n-1),k},$$

где $C^{(n)}$ вычисляется рекуррентно:

$$C^{(n)} = C^{(n+1)} \cdot W^{(n+1)} \cdot MF_n, \quad C^{(N)} = \epsilon^k \cdot MF_N,$$

$$\epsilon^k = \left((y_1^{(2),k} - t_1^k) \quad (y_2^{(2),k} - t_2^k) \quad \dots \quad (y_{m_2}^{(2),k} - t_{m_2}^k) \right),$$

$$а \quad MF_n = \begin{pmatrix} F_n'(S_1^{(n),k}) & 0 & \dots & 0 \\ 0 & F_n'(S_2^{(n),k}) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & F_n'(S_{m_n}^{(n),k}) \end{pmatrix}$$

– матрица размерности $m_n \times m_n$, а матрица $M_{j_n j_{n-1}}^{(n)}$ размерности $m_n \times (m_{n-1} + 1)$, в которой на пересечении j_n -ой строки и j_{n-1} -ого столбца расположено число 1, а остальные элементы матрицы равны нулю.

Изменение синаптических связей и порогов сети производится, начиная от последнего N -ого до первого блока сети.

Доказательство: Вычислим градиент ошибки для k -ого образа:

$$\begin{aligned} \frac{\partial E_s^{(k)}}{\partial w_{j_{n-1}j_n}^{(n)}} &= \frac{\partial \left(\sum_{i_n=1}^{m_n} \frac{1}{2} (y_{i_n}^{(N),k} - t_{i_n}^k)^2 \right)}{\partial w_{j_{n-1}j_n}^{(n)}} = \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k) \cdot \frac{\partial y_{i_n}^{(N),k}}{\partial w_{j_{n-1}j_n}^{(n)}} = \\ &= \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k) \cdot F_n'(S_{i_n}^{(N),k}) \cdot \frac{\partial S_{i_n}^{(N),k}}{\partial w_{j_{n-1}j_n}^{(n)}} = \\ &= \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k) \cdot F_n'(S_{i_n}^{(N),k}) \cdot \sum_{i_{n-1}=1}^{m_{n-1}} w_{i_{n-1}i_n}^{(N)} \cdot \frac{\partial y_{i_{n-1}}^{(N-1),k}}{\partial w_{j_{n-1}j_n}^{(n)}} = \\ &= \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k) \cdot F_n'(S_{i_n}^{(N),k}) \times \\ &\times \sum_{i_{n-1}=1}^{m_{n-1}} w_{i_{n-1}i_n}^{(N)} \cdot F_{n-1}'(S_{i_{n-1}}^{(N-1),k}) \cdot \frac{\partial S_{i_{n-1}}^{(N-1),k}}{\partial w_{j_{n-1}j_n}^{(n)}} = \\ &= \dots = \\ &= \sum_{i_n=1}^{m_n} (y_{i_n}^{(N),k} - t_{i_n}^k) \cdot F_n'(S_{i_n}^{(N),k}) \cdot \sum_{i_{n-1}=1}^{m_{n-1}} w_{i_{n-1}i_n}^{(N)} \cdot F_{n-1}'(S_{i_{n-1}}^{(N-1),k}) \cdot \dots \cdot \\ &\dots \cdot \sum_{i_1=1}^{m_1} w_{i_1 i_2}^{(n+1)} \cdot F_1'(S_{i_1}^{(n),k}) \cdot y_{j_{n-1}}^{(n-1),k} \cdot \delta_{j_n}^{i_n} = \\ &= \epsilon^k \cdot MF_N \cdot W^{(N)} \cdot MF_{N-1} \cdot \dots \cdot W^{(n+1)} \cdot MF_n \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k} = \\ &= C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k}, \end{aligned}$$

где

$$C^{(n)} = C^{(n+1)} \cdot W^{(n+1)} \cdot MF_n, \quad C^{(N)} = \epsilon^k \cdot MF_N.$$

Т.к. $\frac{\partial E_s}{\partial z} = \frac{\partial \left(\frac{1}{L} \sum_{k=1}^L E_s^{(k)} \right)}{\partial z} = \frac{1}{L} \sum_{k=1}^L \frac{\partial E_s^{(k)}}{\partial z}$, то формулы настройки синаптических связей

$$w_{j_{n-1}j_n}^{(n)}(t+1) = w_{j_{n-1}j_n}^{(n)}(t) - \alpha^{(n)} \frac{\partial E_s}{\partial w_{j_{n-1}j_n}^{(n)}}$$

$j_{n-1} = \overline{1, m_{n-1}}, j_n = \overline{1, m_n}$, принимают вид:

$$w_{j_{n-1}j_n}^{(n)}(t+1) = w_{j_{n-1}j_n}^{(n)}(t) - \alpha^{(n)} \cdot \frac{1}{L} \cdot \sum_{k=1}^L C^{(n)} \cdot M_{j_n j_{n-1}}^{(n)} \cdot Y^{(n-1),k}$$

Аналогичным образом можно получить формулы для настройки порогов $T_{j_n}^{(n)}$ нейронной сети.

Теорема доказана.

УДК 681.3.005.23

Ракецкий В.М., Прожерин И.Г., Павлюкович С.В.

ОПТИМИЗАЦИЯ ДВИЖЕНИЯ КОЗЛОВЫХ КРАНОВ НА ПОГРУЗОЧНО-РАЗГРУЗОЧНОЙ ПЛОЩАДКЕ БРЕСТСКОГО УЗЛА ЖЕЛЕЗНОЙ ДОРОГИ

1. ПОСТАНОВКА ЗАДАЧИ

На брестском узле железной дороги существует проблема несоответствия стандартов ширины колеи между западной и белорусской веткой. В связи с этим возникает задача перегрузки контейнеров с западных вагонов на белорусские.

Для перегрузки контейнеров с платформ западного образца на белорусские необходимо контейнер выгрузить на площадку для таможенного осмотра и декларирования грузов. Затем контейнеры загружаются на белорусский состав.

Контейнеры могут быть двух типов 20 футовые и 40 футовые. Вагоны могут быть трех типов 2-х местных, 3-х местных и полувагон.

Соответствие погрузки контейнеров на вагоны приведено в табл.1.

Контейнеры формируются в комплекты по месту назначения. Один комплект устанавливается на один вагон. Контейнеры могут быть беспорядочно расположены на всей площадке.

Таблица 1 – Соответствие погрузки контейнеров на вагоны.

Тип вагона	Варианты расстановки контейнеров
2-х местный	2х20 футовые
3-х местный	3х20 футовых или 20+40 футовый
полувагон	40 футовый

Загрузку и разгрузку контейнеров осуществляют 3 козловых крана. Они находятся на одних рельсах и при движении возникает проблема столкновения. Для работы используются краны разных типов: один кран для 40 футовых контейнеров и 2 крана для 20 футовых.

Вся площадка разбита на 3 зоны (А, В, С), как показано на рис.1.

На площадке выделены позиции для 20 футовых контейнеров. 40 футовые контейнеры занимают 2 позиции. Ряды и столбцы нумеруются с 1. Позиция контейнера на площадке представляется следующим образом, например А/1/2, что означает контейнер находится в зоне А в ряду 1 и столбце 2.

Целью работы является составление плана загрузки состава, с учетом требований предъявляемых заказчиком:

1) время расчета не более 10 мин. (на среднем классе компью-

тер). Полученные в теореме матричные и рекуррентные соотношения для модификации весовых коэффициентов и порогов многослойных гетерогенных нелинейных нейронных сетей для минимизации среднеквадратичной ошибки сети позволяют их эффективно использовать при программной реализации алгоритмов обучения сетей на основе различных градиентных методов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Гладкий И.И., Головкин В.А., Махнист Л.П. Обучение нейронных сетей с использованием метода наискорейшего спуска // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2001. № 5: Физика, математика, химия. – С. 56-61.

тер). Время счета задачи должно быть ограничено, так как некоторые контейнеры могут блокироваться при таможенном осмотре и нужно оперативно получать новый план загрузки в связи с изменением входных данных;

2) минимизация времени загрузки, которая позволяет снизить затраты на электроэнергию, простой составов, износ оборудования и уменьшение количества обслуживающего персонала.

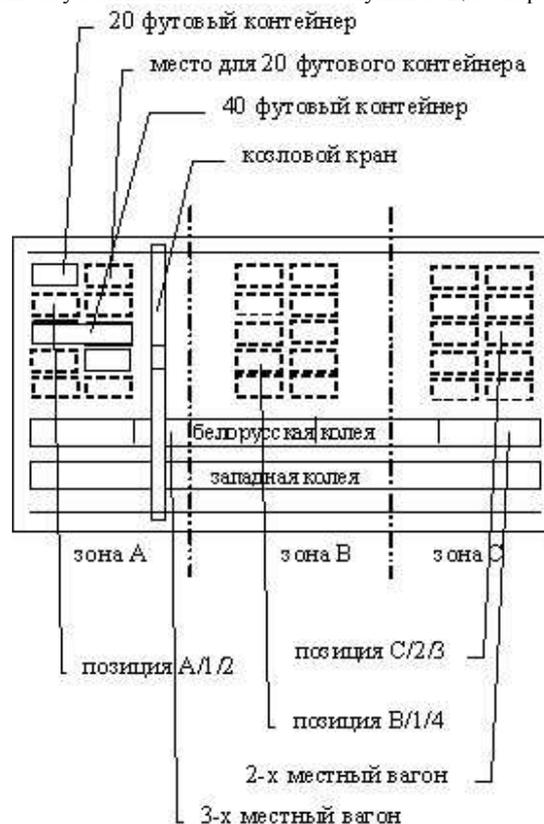


Рис.1 – Погрузочно-разгрузочная площадка на брестском узле железной дороги.

Ракецкий Валерий Михайлович. К. физ.-мат. н., доцент, зав. каф. информатики и прикладной математики Брестского государственного технического университета.

Прожерин Игорь Геннадиевич. Ассистент каф. информатики и прикладной математики Брестского государственного технического университета.

Павлюкович Сергей Вячеславович. Ассистент каф. информатики и прикладной математики Брестского государственного технического университета.

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