программирования и их модификаций, возможно решение комплексных и изменяемых во времени задач, которые решить стандартными линейными математическими методами практически невозможно.

Список литературы:

1. Жолобов, Д.А. Введение в математическое программирование: Учебное пособие. – М.:МИФИ, 2008. – 376 с.

2. Таха, Хэмди А. Введение в исследование операций, 7-е издание.: Пер. с англ. – М. : Издательский дом “Вильямс”, 2007. – 912 с.

УДК 004.021:032.26

ОБУЧЕНИЕ ДВУХКОМПОНЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ

В. М. Жежеров, Т. А. Лузько
Брестский государственный технический университет, г. Брест
Научные руководители: Л. П. Махнист, канд. техн. наук, доцент,
Т. И. Каримова, канд. физ.-мат. наук, доцент

В работе рассматривается обучение искусственных нейронных сетей прямого распространения с использованием метода обратного распространения ошибки.

Предполагалось, что все слои сети содержат два нейронных элемента.

Обучение такой нейронной сети с использованием метода обратного распространения ошибки состояло в нахождении весовых коэффициентов w_{ij} и порогов T_j нейронной сети, которые минимизируют функцию ошибки сети, построенную по методу наименьших квадратов

$$E(w_{11}, w_{21}, T_1, w_{12}, w_{22}, T_2) = \frac{1}{2}\left((y_1 - t_1)^2 + (y_2 - t_2)^2\right),$$

где $y_j = F(S_j)$ – значение гладкой функции активации j-го выходного нейрона сети, $S_j = w_{1j}x_1 + w_{2j}x_2 - T_j$, x_i – выходное значение i-го нейрона предыдущего слоя,

t_j – ожидаемый выход j-го выходного нейрона ($i = \overline{1,2}$, $j = \overline{1,2}$).

Используем обозначения:

$\overline{W} = (w_{11}, w_{21}, T_1, w_{12}, w_{22}, T_2)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порогов T_j нейронной сети, а $\overline{W}_j = (w_{1j}, w_{2j}, T_j)^T$ – вектор-столбец весовых коэффициентов w_{ij} и порога T_j, связанных с j-ым выходным нейроном сети,

$E\left(\overline{W}\right)=E\left(\overline{W}_1\right)+E\left(\overline{W}_2\right)$ – 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.

Изменение весовых коэффициентов w_{ij} и порогов T_j нейронной сети на каждом шаге обучения $(t+1)$ ($t=1,2,\ldots$) производился по формулам:

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

где $\nabla E\left(\overline{W}_j(t)\right)$ – градиент функции ошибки сети j-го выходного нейрона $E\left(\overline{W}_j(t)\right)$ (например, в [1–2]).

Заметим, что в [3] рассматривались различные подходы к выбору шага обучения нейронной сети $\alpha(t)$, производился их сравнительный анализ с точки зрения сходимости алгоритма обучения с использованием метода наискорейшего спуска.

После вычисления весовых коэффициентов w_{ij} и порогов T_j последнего слоя вычисляли ожидаемые выходы нейронов предыдущего слоя $t_i=x_i$, которые минимизируют ошибку сети, как функцию, зависящую от значений x_i:

$$E\left(\overline{X}\right)=E(x_1,x_2)=\frac{1}{2}\left((y_1-t_1)^2+(y_2-t_2)^2\right).$$

Пусть $\overline{X}=(x_1,x_2)^T$ – вектор-столбец выходных значений нейронов предыдущего слоя.

Ожидаемые выходы нейронов предыдущего слоя $t_i=x_i$ производились с использованием метода наискорейшего спуска на каждом шаге $(t+1)$ $(t=1,2,\ldots)$ по формуле:

$$\overline{X}(t+1)=\overline{X}(t)-\beta(t)\nabla E\left(\overline{X}(t)\right), \tag{2}$$

где $\nabla E\left(\overline{X}(t)\right)$ – градиент функции ошибки сети $E\left(\overline{X}(t)\right)$.

Соблюдая условия, приведенные, например, в [4], шаг $\beta(t)$ вычислялся по формуле:

$$\beta(t)=\frac{\left(\nabla E\left(\overline{X}(t)\right),\nabla E\left(\overline{X}(t)\right)\right)}{\left(\nabla^2 E\left(\overline{X}(t)\right)\cdot\nabla E\left(\overline{X}(t)\right),\nabla E\left(\overline{X}(t)\right)\right)}, \tag{3}$$

где $\nabla^2 E\left(\overline{X}(t)\right)\cdot\nabla E\left(\overline{X}(t)\right)$ – произведение матрицы Гессе $\nabla^2 E\left(\overline{X}(t)\right)$ функции $E\left(\overline{X}(t)\right)$ и вектора градиента $\nabla E\left(\overline{X}(t)\right)$, а $\left(\nabla E\left(\overline{X}(t)\right),\nabla E\left(\overline{X}(t)\right)\right)$ и $\left(\nabla^2 E\left(\overline{X}(t)\right)\cdot\nabla E\left(\overline{X}(t)\right),\nabla E\left(\overline{X}(t)\right)\right)$ – скалярные произведения соответствующих векторов.

Формулы (2) и (3) дают возможность определить ожидаемые выходы нейронов предыдущего слоя $t_i = x_i$. По формуле (1) производится изменение весовых коэффициентов и порогов предыдущего слоя.

Получены простые формулы (1)–(3) для рассматриваемой архитектуры сети.

Список литературы

1. Maniakov, N. Traing algorithm for forecasting multilayer neural network / N. Maniakov, L. Makhnist, V. Rubanov // Pattern Recognition and Information Processing : Proc. of The Seventh Intern. Conf. (PRIP'2003), Minsk, Republic of Belarus, 21–23 May 2003 : in 2 vol. – Minsk, 2003. – Vol. 1. – P. 26–30.

2. Makhnist, L. Some Methods of Adaptive Multilayer Neural Network Training / L. Makhnist, N. Maniakov // International Journal of Computing. – 2004. – Vol. 3. – P. 99–106.

3. 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) : Proc. of the Ninth Intern. Conf., Minsk, Republic of Belarus, 22–24 May 2007 : in 2 vol. – Minsk, 2007. – Vol. 1. – P. 285–289.

4. Махнист, Л. П. Использование алгоритмов обучения однослойной сети для многослойных нейронных сетей прямого распространения / Л. П. Махнист, В. А. Головко, И. И. Гладкий // Вестн. Брест. гос. техн. ун-та. Сер.: Физика, математика, информатика. – 2020. – № 5. – С. 32–37.

УДК 519.6+517.983

ВЫБОР МОМЕНТА ОСТАНОВА В ИТЕРАЦИОННОМ МЕТОДЕ ЯВНОГО ТИПА РЕШЕНИЯ НЕКОРРЕКТНЫХ УРАВНЕНИЙ ПЕРВОГО РОДА

А. В. Жорох,
БрГУ имени А. С. Пушкина, Брест
Научный руководитель:
О. В. Матысик, кандидат физико-математических наук, доцент

В действительном гильбертовом пространстве H решается линейное уравнение первого рода $Ax = y$, где $A: H \to H$ — положительный ограниченный самосопряженный оператор ($0 \in SpA$, и, следовательно, рассматриваемая задача некорректна). Решать данную задачу будем при помощи *итерационного метода явного типа*

$$x_{n+1} = x_n + \alpha A(y - Ax_n), \quad x_0 = 0. \tag{1}$$