

ОБУЧЕНИЕ МНОГОСЛОЙНЫХ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА УСЛОВНОЙ МИНИМИЗАЦИИ ОШИБКИ КАЖДОГО СЛОЯ

А.В. Дворниченко
(БГТУ, Брест)

На основе матричной алгоритмизации процесса обучения многослойных гетерогенных нейронных сетей прямого распространения без обратных связей [1] предлагается метод условной минимизации ошибки. Изменение синаптических связей и весов каждого слоя, а также изменение «целевых ошибок» производится в соответствии с нижеследующей теоремой, которая также предлагает формулы для вычисления адаптивного шага.

Теорема. При использовании алгоритма условной минимизации ошибки вычисление «эталонных» значений выходов, при некотором заданном параметре β , производится в соответствии с формулой:

$$t_{j_n}^{(n),k} = y_{j_n}^{(n),k} - \alpha \cdot C^{(n+1)} \cdot W^{(n+1)} \cdot \Delta_{j_n}^n, \quad j_n = \overline{1, m_n},$$

с последующей коррекцией:

$$t_{j_n}^{(n),k} := \begin{cases} a + \beta, & \text{if } t_{j_n}^{(n),k} < a + \beta \\ t_{j_n}^{(n),k}, & \text{if } t_{j_n}^{(n),k} \in [a + \beta, b - \beta] \\ b - \beta, & \text{if } t_{j_n}^{(n),k} > b - \beta \end{cases}$$

адаптивный шаг при этом вычисляется в соответствии с равенством:

$$\alpha = \frac{\sum_{j_n=1}^{m_n} (C^{(n+1)} \cdot P_{j_n}^n)^2}{\sum_{j_n=1}^{m_n} \sum_{i_n=1}^{m_n} (C^{(n+1)} \cdot P_{j_n}^n) \cdot \left((P_{i_n}^n)^T \cdot U^{(n+1),k} \cdot (P_{i_n}^n) \right) \cdot (C^{(n+1)} \cdot P_{i_n}^n)}$$

При этом настройка весов и порогов сети производится по формулам:

$$w_{j_{n-1}j_n}^{(n)}(t+1) = w_{j_{n-1}j_n}^{(n)}(t) - \alpha^{(n)} \cdot G_{j_{n-1}j_n}^{(n),layer}, \quad j_{n-1} = \overline{1, m_{n-1}}, \quad j_n = \overline{1, m_n}$$

$$T_{j_n}^{(n)}(t+1) = T_{j_n}^{(n)}(t) - \alpha^{(n)} \cdot G_{j_n}^{(n),layer}, \quad j_n = \overline{1, m_n},$$

с адаптивным шагом, вычисляемым в соответствии с

$$\alpha^{(n)} = \frac{\sum_{j_{n-1}=1}^{m_{n-1}+1} \sum_{j_n=1}^{m_n} (G_{j_{n-1}j_n}^{(n),layer})^2}{\sum_{j_{n-1}, j_{n-1}=1}^{m_{n-1}+1} \sum_{j_n, j_n=1}^{m_n} G_{j_{n-1}j_n}^{(n),layer} \cdot (S_{(n)})_{j_{n-1}j_n} \cdot G_{j_{n-1}j_n}^{(n),layer}}$$

ЛИТЕРАТУРА

1. Маньяков Н.В., Махнист Л.П. Матричная алгоритмизация обучения многослойных нейронных сетей с использованием градиентных методов // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2002. - №5 (17): Физика, математика, химия. – С. 60-64.