

В. Б. Л и л а (Ростов-на-Дону, РГСУ). **Адаптивный метод обучения искусственных нейронных сетей.**

Все градиентные методы обучения нейронных сетей можно представить в виде предлагаемого адаптивного метода.

Общую формулу для изменения весов выразим формулой $\bar{w}_{k+1} = \bar{w}_k + \eta_k \bar{p}_k$, где вектор \bar{p}_k задает направление движения, а η_k — размер шага k -й итерации.

Формулу расчета вектора \bar{p}_k выразим следующим образом: $\bar{p}_k = \bar{g}_k + \sum_{i=1}^{\min\{k-1, m\}} \beta_i^k \bar{g}_{k-i}$, где вектор \bar{p}_k задает направление движения, \bar{g}_i — направление антиградиента на i -й итерации, β_i — коэффициент, определяющий вес i -го градиента, $i = 1, 2, \dots, k$, m — количество запоминаемых градиентов, k — порядковый номер текущей итерации.

Отметим, что методы сопряженных градиентов [1], которые наиболее часто употребляются при обучении нейросетей, являются частным случаем адаптивного метода.

Рассмотрим пример использования такого подхода при оптимизации функции Розенброка [2]:

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2. \quad (1)$$

Проведены эксперименты из одной стартовой точки с координатами $(-1, 5, 3, 5)$ с изменением параметров η, m, β .

Для градиентного спуска ($\beta = 0$) при оптимальном $\eta = 0,005$ значение целевой функции, равное $-0,30103179$, достигается на 121-й итерации.

Для метода тяжелого шарика [1] $\bar{p}_k = \bar{g}_k + \beta \bar{p}_{k-1}$, который также является частным случаем адаптивного метода, при $\beta = 0,9$ и $\eta = 0,005$ значение целевой функции $0,12836612$ достигается на 52-й итерации.

Для адаптивного алгоритма с параметрами $\eta = 0,0004$, $m = 2$, $\beta_1 = 0,9$, $\beta_2 = 0,081$ значение целевой функции $-0,04807101$ получено на 35-й итерации.

Таким образом, для сложной функции Розенброка (1) для адаптивного алгоритма получены лучшие результаты по сравнению с классическими методами градиентного спуска и тяжелого шарика.

Применяя адаптивный алгоритм для обучения на нейронных сетях, мы получаем возможность настройки параметров алгоритма на небольшом количестве примеров, продолжив затем обучение на реальных данных.

СПИСОК ЛИТЕРАТУРЫ

1. *Тархов Д. А.* Нейронные сети. Модели и алгоритмы. Кн. 18. М.: Радиотехника, 2005, 256 с.
2. *Крушель Е. Г., Степанченко О. В.* Методические указания к исследовательской лабораторной работе по дисциплине «Математические основы кибернетики».
3. *Хайкин С.* Нейронные сети. М.: Издательский дом «Вильямс», 2006, 1104 с.