DOI: 10.17117/na.2016.09.01.475

<http://ucom.ru/doc/na.2016.09.01.475.pdf>

Поступила (Received): 01.09.2016

Полозова Ю.Е.
**Об особенностях применения интервальных алгоритмов
оптимизации при обучении интервальных
нейросетевых моделей**

Polozova Yu.E.
**About peculiarities of application of interval optimization
algorithms for training of interval neural network models**

В работе рассматривается подход к обучению интервальных нейронных сетей на основе интервальных алгоритмов глобальной оптимизации. Рассмотрена постановка оптимизационной задачи обучения интервальной нейронной сети. Предложена модификация алгоритма обучения на основе бисекции вектора весов. Приведены результаты численного моделирования

Ключевые слова: интервальная нейронная сеть, алгоритмы оптимизации

Полозова Юлия Евгеньевна

Аспирант

Липецкий государственный технический университет

г. Липецк, ул. Московская, 30

An approach to training interval neural networks based interval global optimization algorithms is described in this paper. The optimization problem of training interval neural network is formulated. Modification of the learning algorithm based on the bisection vector of weights are discussed. There are results of numerical simulations in conclusion

Key words: interval neural network, optimization algorithms

Polozova Yuliya Evgenyevna

Graduate

Lipetsk state technical university
Lipetsk, Moskovskaya st., 30

Использование алгоритмов глобальной оптимизации на основе методов интервального анализа позволяет гарантировать нахождение глобального оптимума и включение точного решения в выходной интервал. Однако применительно к обучению интервальных нейросетевых моделей были выявлены некоторые особенности использования интервальных алгоритмов глобальной оптимизации. Обсуждению таких особенностей посвящена данная работа.

1. Обзор предыдущих результатов

В качестве исходной предпосылки рассматривается задача нейросетевого прогнозирования временного ряда, заданного интервальными числами. При этом результат прогнозирования должен представлять собой интервал наименьшей ширины, содержащий фактическое интервальное значение. В целях упрощения вычислений рассматривается двухслойная интервальная нейронная сеть (ИНС) с одним выходом.

В ходе ряда проведенных исследований было установлено, что включение фактических значений в прогнозный интервал может быть гарантировано в случае использования функционала качества:

$$J([w]) = \sum_{i=1}^k Q_i([w]) = \sum_{i=1}^k \{p_i(\tilde{y}_i - y_i([w])) + \bar{p}_i(\overline{y_i([w])} - \tilde{y}_i)\} \quad (1)$$

где
$$p_i = \begin{cases} 1, & \tilde{y}_i \geq y_i([w]), \\ N(y_i([w]) - \tilde{y}_i), & \tilde{y}_i < y_i([w]); \end{cases} \quad \bar{p}_i = \begin{cases} 1, & \overline{y_i([w])} \geq \tilde{y}_i, \\ N(\tilde{y}_i - \overline{y_i([w])}), & \overline{y_i([w])} < \tilde{y}_i. \end{cases}$$

Здесь k – число примеров обучающего множества; $[y]([w])$ – вектор выходов сети; $[\tilde{y}]$ – вектор указаний учителя; p_i – некоторый весовой коэффициент (в общем случае интервалом не является); N – уровень толерантности, показывающий порядок отклонения границы интервала, который признается допустимым (например, 10, 100, 1000).

Для обучения полученной модели предлагалось воспользоваться простейшим интервальным адаптивным алгоритмом глобальной оптимизации функций [1] на основе бисекции вектора весов. Однако по результатам вычислительных экспериментов одним из существенных недостатков предложенной модели являлась большая ширина прогнозного интервала.

2. Модификация простейшего интервального адаптивного алгоритма глобальной оптимизации

В ходе проведенных исследований были выявлены следующие особенности обучения ИНС:

1. В случае обучения классической нейронной сети исходные данные рекомендуется нормировать и центрировать в диапазон $[-1;1]$. В случае обучения ИНС центрирование интервала не требуется, исходные данные могут быть нормированы в диапазон $[0;1]$. Эмпирическим путем было выявлено, что повышение точности прогнозирования может быть достигнуто умножением нормированных значений на масштабирующий коэффициент.

2. В случае обучения классической нейронной сети начальные веса генерируются случайно в малом симметричном относительно нуля диапазоне. В случае обучения ИНС нет необходимости в случайной генерации. Напротив, начальные веса должны представлять собой интервалы достаточно большой ширины, чтобы при первом проходе по сети вектор выходов гарантированно включал вектор указаний учителя.

3. Согласно алгоритму [1] выбор координаты для бисекции необходимо производить в соответствии с наибольшей шириной интервала. Однако в соответствии с п.2 все координаты вектора весов на первой итерации будут иметь одинаковую ширину. На последующих итерациях, по-прежнему, будут оставаться координаты с одинаковой шириной интервала. Таким образом, представляется целесообразным выбирать координату для бисекции простым перебором.

4. В соответствии с алгоритмом [1] элемент списка, подвергнутый бисекции, подлежит удалению. Однако в том случае, если оценка целевой функции

ухудшилась, при удалении будет потеряно наилучшее приближение. Таким образом, предлагается сохранять рассматриваемый элемент списка до тех пор, пока не будет найдена улучшенная оценка целевой функции.

Обобщая вышесказанное, далее представляется модифицированный простейший интервальный адаптивный алгоритм глобальной оптимизации.

Вход: минимальная ширина бруса $\delta > 0$, требуемая величина ошибки обучения $\varepsilon > 0$.

Выход: $[w]$ – вектор весов, доставляющий минимальное значение функционалу (1); J^* – минимальное значение функционала (1).

1. Инициализация начальных весов $[w]$ значениями $[0,0;1]$.
2. Вычисление выхода ИНС y^* и функционала качества J^* .
3. Инициализация списка $L := \{[w], J^*\}$.
4. Инициализация координаты для бисекции $l = 0$, счетчика итераций для выхода из цикла $c = 0$.

5. Цикл: Пока $\min_{i=1}^n \text{wid}[w_i] > \delta$ и $J^* > \varepsilon$.

5.1. Вычисление координаты для бисекции $l = l + 1$. Если l больше размерности вектора весов, то $l = 1$.

5.2. Бисекция $[w]$ по l -й координате на $[w']$ и $[w'']$.

5.3. Вычисление выходов ИНС y', y'' и функционалов качества J', J'' .

5.4. Если $J' > J^*$ и $J'' > J^*$, то переход на шаг 5.5, иначе переход на шаг 5.6.

5.5. Присвоение $c = c + 1$. Если c не равно размерности вектора весов, то переход на шаг 5.1. Иначе выход из цикла.

5.6. Удаление из L элемента $([w], J^*)$.

5.7. Добавление в список L элементов $([w'], J')$ и $([w''], J'')$.

5.8. Сортировка списка L по возрастанию второго поля.

5.9. Обозначение первой записи через $([w], J^*)$.

3. Результаты вычислительного эксперимента

Рис. 1. Прогноз (опыт 1)

Рис. 2. Прогноз (опыт 2)

Для проведения экспериментов используются 2 набора данных, содержащие по 30 интервальных значений. В первом опыте интервальные числа представляют собой частный случай, когда верхняя и нижняя границы интервалов совпадают. По исходным данным производится построение прогноза на 5 значений вперед. Для прогнозирования используется двухслойная ИНС с 5 входными нейронами и 3 нейронами скрытого слоя. Результаты каждого из опытов приведены на рис. 1 и 2.

По результатам эксперимента можно заметить, что несмотря на небольшой объем обучающего множества, достигается гарантированность включения фактических значений в прогнозный интервал. Кроме того, следует отметить, что ширина прогнозных интервалов является приемлемой, хотя и возрастает с увеличением горизонта прогнозирования.

Список используемых источников:

1. Шарый С.П. Рандомизированные алгоритмы в интервальной глобальной оптимизации // Сибирский журнал вычислительной математики. Т. 11. 2008. №4. С. 457-474.

© 2016, Полозова Ю.Е.

Об особенностях применения интервальных алгоритмов оптимизации при обучении интервальных нейросетевых моделей

© 2016, Polozova Yu.E.

About peculiarities of application of interval optimization algorithms for training of interval neural network models