

удалена строки, соответствующие источнику и стоку. Можно показать, что каждой базисной матрице соответствуют дуги, которые вместе с узлами сети образуют две компоненты связности, одна из которых содержит источник, вторая – сток. Каждая компонента является деревом. Верно и обратное: если множество узлов разбить на 2 части, одна из которых содержит источник, а вторая – сток, а на соответствующих фрагментах сети построить деревья, то дугам сети, вошедшим в дерево, будет соответствовать базисная матрица.

4. *Алгоритм решения систем вида (3)*. Ассоциация базисной матрицы с деревьями на сети позволяет предложить весьма простой алгоритм решения систем типа (3) (для определенности – первой):

- 1) подсчитаем и запомним для каждого уравнения системы (3) количество входящих в него (неизвестных) переменных;
- 2) последовательно анализируем уравнения системы (3) на предмет количества входящих в них неизвестных переменных. Если для всех уравнений это количество равно 0, то система решена. Работа алгоритма окончена;
- 3) для всех уравнений, в которых не известна только одна переменная,
 - а) найдем значение неизвестной переменной;
 - б) уменьшим количество неизвестных переменных на 1 для всех уравнений системы, в которые входит найденная переменная;
- 4) перейдем к п. 2.

5. *Заключение*. Поскольку системы уравнений типа (3) решаются без использования обратной базисной матрицы или какого-либо ее разложения, а элементы матрицы A задачи (2) легко моделируются по описанию графа задачи, предложенный алгоритм а) нетребователен к памяти компьютера, б) численно устойчив, в) эффективен с вычислительной точки зрения.

СПИСОК ЛИТЕРАТУРЫ

1. Ракецкий, В. М. Прямой опорный метод решения сетевой задачи квадратичного программирования / В. М. Ракецкий // Динамические системы: устойчивость, управление, оптимизация : тез. докл. Междунар. конф., Минск, 29 сент. – 4 окт. 2008 г. – Минск : Ин-т математики НАН Беларуси, 2008. – С. 139–140.

А.Ю. САВИЦКИЙ

БрГУ (г. Брест, Беларусь)

ОБУЧЕНИЕ СИГМОИДАЛЬНЫХ НЕЙРОНОВ В АРХИТЕКТУРЕ МНОГОСЛОЙНОЙ НЕЙРОННОЙ СЕТИ

В последнее время в мире активизировались исследования в области глубокого обучения многослойных нейронных сетей. Это связано с определенными успехами в данной области, достигнутыми рядом исследователей [1; 2], а также высокой практической значимостью сильно-многослойных нейронных сетей (СМНС). При этом перспективным считается подход к предобучению (pre-training) СМНС не только с помощью ограниченной машины Больцмана (RBM), но и с применением нейросетевых автоэнкодеров (AutoEncoder, AE). Каждый такой AE представляет собой трехслойный персептрон архитектуры $N \rightarrow M \rightarrow N$, где

параметр N соответствует количеству входов текущего предобучаемого слоя, M – количеству нейронов указанного слоя СМНС. Последовательное (начиная с входного слоя СМНС) обучение совокупности таких АЕ позволяет получить наборы весовых коэффициентов для финальной настройки синаптических связей всей СМНС (fine-tuning). При этом для обучения как АЕ, так и СМНС, как правило, применяется алгоритм обратного распространения ошибки (Back Propagation Error, BPE) [3]. Очевидно, что эффективность алгоритма BPE напрямую определяет эффективность (точность обучения, обобщающие свойства) результирующей модели СМНС в целом.

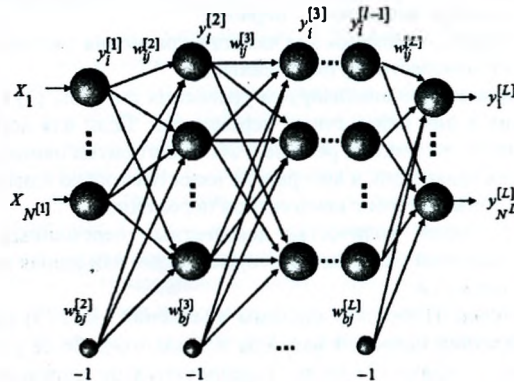


Рисунок. – Архитектура многослойной нейронной сети

На рисунке приведена обобщенная архитектура многослойной нейронной сети, структура искусственного НЭ, а также введены обозначения параметров сети.

Нейронный элемент слоя l осуществляет функцию преобразования некоторого вектора входных сигналов $y^{[l-1]}$ в выходную активность $y^{[l]}$ по следующему правилу:

$$S_j^{[l]} = \sum_{i=1}^{N^{[l-1]}} y_i^{[l-1]} w_{ij}^{[l]} - w_{ij}^{[l]}, \quad y_j^{[l]} = g^{[l]}(S_j^{[l]}), \quad j = 1, \dots, N^{[l]}, \quad (1)$$

где $S_j^{[l]}$ – взвешенная сумма входных активностей НЭ j ; $w_{ij}^{[l]}$, $w_{ij}^{[l]}$ – весовые коэффициенты; $g^{[l]}(S_j^{[l]})$ – активационная функция НЭ.

В рамках данной архитектуры сформулирована и доказана теорема, позволяющая осуществлять точную настройку весовых коэффициентов нейроэлементов с сигмоидной функцией активации.

Теорема. Правила модификации синаптических связей НЭ j , находящегося в слое L , с сигмоидной функцией активации $g^{[L]}$, минимизирующие среднеквадратичную ошибку $E_j^p(t) = 1/2 (y_j^{[L]k,p}(t) - D_j^p)^2$ данного НЭ для эталона p на итерации обучения t , определяются следующим образом:

$$\left\{ \begin{aligned} w_{ij}^{[L]}(t+1) &= w_{ij}^{[L]}(t) - \frac{S_j^{[L],p}(t) - \ln\left(\frac{D_j^p}{1-D_j^p}\right)}{1 + \sum_{k=1}^{N^{[L-1]}} (y_k^{[L-1],p})^2} y_i^{[L-1],p}(t), \\ w_{hj}^{[L]}(t+1) &= w_{hj}^{[L]}(t) + \frac{S_j^{[L],p}(t) - \ln\left(\frac{D_j^p}{1-D_j^p}\right)}{1 + \sum_{k=1}^{N^{[L-1]}} (y_k^{[L-1],p})^2}, \end{aligned} \right. \quad (2)$$

где D_j^p – эталонное выходное значение j -го НЭ.

Автором также получены аналогичные правила модификации синаптических связей для НЭ с функциями активации биполярная сигмоидная и гиперболический тангенс. Следует также отметить, что результаты теоремы (с учетом алгоритма ВРЕ) можно обобщить на НЭ последующих слоев многослойной нейронной сети.

Ряд вычислительных экспериментов был проведен при обучении нейросетевых энкодеров архитектуры $900 \rightarrow 50 \rightarrow 900$. С этой целью для обучения были сгенерированы битовые образы печатных символов размером 30×30 и количеством 100 экземпляров. При выполнении 100 итераций обучения среднеквадратичная ошибка обучения (в расчете на один эталон) достигла значения 0.000945, чего не удавалось достичь методом эмпирического подбора шага. Таким образом, полученные правила (2) можно применять как на этапе предобучения в нейросетевых энкодерах, так на этапе финальной настройки целостной архитектуры СМНС.

СПИСОК ЛИТЕРАТУРЫ

1. Hinton, G. E. A fast learning algorithm for deep belief networks / G. E. Hinton, S. Osindero, Y.-W. Teh // *Neural Computation*. – 2006. – Vol. 18, no. 7. – P. 1527–1554.
2. Bengio, Y. Learning deep architectures for AI / Y. Bengio // *Foundations and Trends in Machine Learning*. – 2009. – Vol. 2, no. 1. – P. 1–127.
3. Rumelhart, D. E. Learning representations by back-propagating errors / D.E. Rumelhart, G. E. Hinton, R. J. Williams // *Nature*. – 1986. – Vol. 323. – P. 533–536.

Н.Л. САКОВИЧ, А.В. БИЛЕВИЧ

БрГТУ (г. Брест, Беларусь), филиал ГУО «ИПК технологий информатизации и управления» БГУ (г. Брест, Беларусь)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБЕСПЕЧЕНИИ УЧЕБНО-ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ УЧРЕЖДЕНИЙ ДОПОЛНИТЕЛЬНОГО ОБРАЗОВАНИЯ ВЗРОСЛЫХ

В современном мире, характеризующемся высоким динамизмом изменений во внешней среде, конкурентоспособной становится только экономика, основанная на потоке инноваций, постоянном технологическом совершенствовании