



Math-Net.Ru

All Russian mathematical portal

N. S. Uzentsova, S. P. Sidorov, Simultaneous approximation of polynomial functions and its derivatives by feedforward artificial neural networks with one hidden layer, *Izv. Saratov Univ. Math. Mech. Inform.*, 2013, Volume 13, Issue 4, 78–82

DOI: 10.18500/1816-9791-2013-13-2-2-78-82

Use of the all-Russian mathematical portal Math-Net.Ru implies that you have read and agreed to these terms of use

<http://www.mathnet.ru/eng/agreement>

Download details:

IP: 18.97.9.172

February 14, 2025, 07:17:00





Библиографический список

1. Тьяев Л. Б. Геометрическая модель поведения автоматов и их неотличимость // Математика, механика, математическая кибернетика : сб. науч. тр. Саратов : Изд-во Саратов. ун-та, 1999. С. 139–143.
2. Тьяев Л. Б. Решение некоторых задач для конечных автоматов на основе анализа их поведения // Изв. Саратов. ун-та. Нов. сер. Сер. Математика. Механика. Информатика. 2006. Т. 6, вып. 1/2. С. 121–133.
3. Тьяев Л. Б. Геометрические образы автоматов и динамические системы // Дискретная математика и ее приложения : материалы X междунар. семинара / под ред. О. М. Касим-Заде. М. : Изд-во мех.-мат. ф-та Моск. ун-та, 2010. С. 510–513.
4. Матов Д. О. Аффинные преобразования геометрических образов конечных автоматов // Проблемы теоретической кибернетики : материалы XVI междунар. конф. / под ред. Ю. И. Журавлева. Нижний Новгород : Изд-во Нижегородского госун-та, 2011. С. 303–306.

Discrete Dynamical Systems Defined Geometrical Images of Automata

L. B. Tyapayev, D. V. Vasilenko, M. V. Karandashov

Saratov State University, Russia, 410012, Saratov, Astrahanskaya st., 83, tyapayevlb@info.sgu.ru, egoschauer@mail.ru, norg113@gmail.com

The object of study is the dynamic system defined by geometrical images of automata. The phase space of the system is determined by orthogonal and affine transformations of geometric images. Compositions of dynamical systems of a given type and their characteristics are studied.

Key words: automata, discrete dynamical systems, geometric images of automata.

References

1. Tyapayev L. B. The geometric model of the behavior of automata and their indistinguishability. *Matematika, Mekhanika, Matematicheskaya kibernetika: Sb. nauch. tr.* Saratov, Saratov Univ. Press, 1999, pp. 139–143 (in Russian).
2. Tyapayev L. B. Solving Some Problems of Automata Behaviour Analysis. *Izv. Sarat. Univ. N. S. Ser. Math. Mech. Inform.*, 2006, vol. 6, no. 1/2, pp. 121–133 (in Russian).
3. Tyapayev L. B. Geometric images of automata and dynamical systems. *Discretnaya matematika i eyo prilozheniya. Materialy X Mezhd. seminar.* Ed. O. M. Kasim-Zade. Moscow, 2010, pp. 510–513 (in Russian).
4. Matov D. O. Affine transformations of geometric images of finite automata. *Problemy teoreticheskoy kibernetiki : Materialy XVI Mezhdunar. konf.* Ed. Yu. I. Zhuravlyova. Nizhni Novgorod, 2011, pp. 303–306 (in Russian).

УДК 519.71, 519.651

ОБ ОДНОВРЕМЕННОМ ПРИБЛИЖЕНИИ АЛГЕБРАИЧЕСКИХ МНОГОЧЛЕНОВ И ИХ ПРОИЗВОДНЫХ НЕЙРОННЫМИ СЕТЯМИ ПРЯМОГО РАСПРОСТРАНЕНИЯ СИГНАЛА С ОДНИМ СКРЫТЫМ СЛОЕМ

Н. С. Узенцова¹, С. П. Сидоров²

¹ Аспирантка кафедры математической экономики, Саратовский государственный университет им. Н. Г. Чернышевского, Uzentsovans@gmail.com

² Кандидат физико-математических наук, доцент кафедры математической экономики, Саратовский государственный университет им. Н. Г. Чернышевского, sidorovsp@info.sgu.ru

В данной статье приводится алгоритм нахождения весов нейронной сети прямого распространения сигнала с одним основным слоем. Алгоритм используется для решения задачи равномерного приближения алгебраического многочлена совместно с его производными с наперед заданной точностью. В качестве функции активации используется рациональная сигмоида.

Ключевые слова: нейронные сети, аппроксимация функций.



ВВЕДЕНИЕ

Определение и детальное описание основных понятий теории нейронных сетей (а именно понятие многослойной нейронной сети прямого распространения с различными функциями активации), а также подробное изложение используемых понятий и архитектур сетей, можно найти, например, в [1, 2].

Для решения многих прикладных задач используется способность нейронных сетей аппроксимировать полиномиальную функцию. В статье [3] полиномиальная функция используется как мера емкости искусственной нейронной сети прямого распространения с конечным числом скрытых узлов, при этом под емкостью сети понимается наибольшая степень полиномов, которые данная нейронная сеть приближает с наперед заданной точностью.

Полиномиальная функция $f : \mathbb{R}^m \rightarrow \mathbb{R}$ степени r с m переменными может быть записана в следующем виде:

$$f(x) = \left(a_0 + \sum_{i=1}^m a_{i1}x_i \right) + \left(\sum_{i=1}^m a_{i2}x_i \right)^2 + \dots + \left(\sum_{i=1}^m a_{ir}x_i \right)^r, \quad (1)$$

где $x = (x_1, x_2, \dots, x_m)^T \in \mathbb{R}^m$, $a_0, a_{ij} \in \mathbb{R}$, $j = 1, \dots, r$, $i = 1, \dots, m$.

Рассмотрим архитектуру нейронной сети F с одним скрытым слоем. Пусть вектор $x = (x_1, x_2, \dots, x_m)^T \in \mathbb{R}^m$ есть вектор входа нейронной сети, m есть количество узлов входного слоя сети F . Обозначим через N количество узлов скрытого слоя сети. Количество входов в скрытый слой равно числу компонент вектора x . Обозначим через y выход сети F . Вес связи между i -м узлом входного слоя и j -м скрытым узлом обозначим w_{ji} , $i = 1, \dots, m$, $j = 1, \dots, n$. Вес связи между j -м узлом скрытого слоя и (единственным) нейроном выходного слоя обозначим c_j . Выход скрытого узла j определяется следующим образом:

$$\sigma(\theta_j) = \frac{\theta_j}{\alpha + |\theta_j|}, \quad \theta_j = \sum_{i=1}^m w_{ji}x_i + b_j, \quad (2)$$

где $j = 1, \dots, N$, α — положительная константа, m — количество входов, b_j — пороговое значение j -го узла скрытого слоя.

Выходное значение $y \in \mathbb{R}$ нейронной сети F будет равно

$$y = \sum_{j=1}^N c_j \sigma(\theta_j).$$

В настоящей статье, следуя идеям работы [3], мы находим алгоритм настройки весов сети для одновременного приближения в равномерной метрике с заданной ошибкой полиномиальной функции фиксированной степени и ее производных. В отличие от работы [3], мы рассматриваем нейронную сеть прямого распространения сигнала с рациональной сигмоидой в качестве функции активации, в то время как в работе [3] изучалась нейронная сеть с экспоненциальной сигмоидальной функцией активации. Мы покажем, что результаты, аналогичные результатам работы [3], будут справедливы и для задачи равномерного приближения алгебраического многочлена совместно с его производными.

ИСКУССТВЕННАЯ НЕЙРОННАЯ СЕТЬ ПРЯМОГО РАСПРОСТРАНЕНИЯ С ОДНИМ ВХОДОМ И ОДНИМ ВЫХОДОМ

Рассмотрим нейронную сеть $F : \mathbb{R} \rightarrow \mathbb{R}$ прямого распространения с одним входом и одним выходом:

$$F(x) = \sum_{j=1}^N c_j \sigma(w_j x + b_j),$$

где $c_j, w_j, b_j \in \mathbb{R}$, $j = 1, \dots, N$, есть параметры сети. Обозначим множество всех таких отображений как $\mathcal{F}_N(\sigma)$.

Полиномиальная функция $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ степени r имеет вид

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_r x^r,$$

где $a_i \in \mathbb{R}$, $i = 0, 1, \dots, r$.



Лемма 1. Производная порядка s функции F существует в каждой точке $x \in \mathbb{R}$, при этом справедливо равенство

$$\frac{d^s F(x)}{dx^s} = \sum_{j=1}^N c_j w_j^s Q_s(\sigma(\theta_j)), \quad (3)$$

где $Q_s(\sigma(\theta_j)) = (-1)^{s+1} s! \theta_j^{-s} \sigma^s(\theta_j) (1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j)$, $\theta_j = \theta_j(x) = w_j x + b_j$.

Предположим, что справедливы следующие равенства:

$$\frac{d^s F(0)}{dx^s} = \frac{d^s f(0)}{dx^s} = (s!) a_s, \quad s = 0, 1, 2, \dots, r,$$

т. е. имеет место следующее матричное равенство:

$$\begin{pmatrix} Q_0(\sigma_1^0) & Q_0(\sigma_2^0) & \dots & Q_0(\sigma_N^0) \\ Q_1(\sigma_1^0)w_1 & Q_1(\sigma_2^0)w_2 & \dots & Q_1(\sigma_N^0)w_N \\ \dots & \dots & \dots & \dots \\ Q_r(\sigma_1^0)w_1^r & Q_r(\sigma_2^0)w_2^r & \dots & Q_r(\sigma_N^0)w_N^r \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_N \end{pmatrix} = \begin{pmatrix} a_0 \\ (1!)a_1 \\ \dots \\ (r!)a_r \end{pmatrix}, \quad (4)$$

где $Q_0(\sigma(\theta)) = \sigma(\theta)$, $\sigma_j^0 := \sigma(\theta_j(0)) = \frac{\theta_j(0)}{\alpha + |\theta_j(0)|}$, $j = 1, \dots, N$.

Обозначим

$$\Omega = \begin{pmatrix} Q_0(\sigma_1^0) & Q_0(\sigma_2^0) & \dots & Q_0(\sigma_N^0) \\ Q_1(\sigma_1^0)w_1 & Q_1(\sigma_2^0)w_2 & \dots & Q_1(\sigma_N^0)w_N \\ \dots & \dots & \dots & \dots \\ Q_r(\sigma_1^0)w_1^r & Q_r(\sigma_2^0)w_2^r & \dots & Q_r(\sigma_N^0)w_N^r \end{pmatrix}. \quad (5)$$

Тогда матричное равенство (4) можно записать в виде $\Omega c = a$, где $c = (c_1, c_2, \dots, c_N)^T$, $a = (a_0, (1!)a_1, \dots, (r!)a_r)^T$.

Лемма 2. Пусть $\delta > 0$ и $w_i = (i - 1)\delta$, $i = 1, \dots, N$. Если c_1, \dots, c_N есть решение системы (4), то справедлива система неравенств

$$|c_j| \leq \psi_{\max} q_{\max} \sum_{i=0}^r |i! a_i| \delta^{-i}, \quad j = 1, \dots, N,$$

где ψ_{\max} есть положительное число, зависящее от N и r , q_{\max} есть положительное число, зависящее от порогового значения в скрытом слое.

В следующей теореме показывается, что сеть с одним скрытым слоем, содержащим N узлов, может аппроксимировать любой алгебраический многочлен степени $N - 1$ совместно с его производными с любой степенью точности на произвольном замкнутом множестве числовой прямой.

Теорема 1. Пусть $x_{\max} > 0$ — произвольное число. Тогда для произвольной функции $f \in \mathcal{P}_{N-1}$ и для всякого $\varepsilon > 0$ найдется функция $F(x) \in \mathcal{F}_N(\sigma)$ такая, что

$$\sup_{x \in [-x_{\max}, x_{\max}]} \sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| < \varepsilon,$$

где ε является ошибкой равномерного приближения на $[-x_{\max}, x_{\max}]$.

Доказательство. Пусть $f \in \mathcal{P}_{N-1}$ есть произвольный многочлен. Возьмем $\varepsilon > 0$ и покажем, что для некоторой $F(x) \in \mathcal{F}_N(\sigma)$ для всех $x \in [-x_{\max}, x_{\max}]$ будет

$$\sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| < \varepsilon.$$

Потребуем

$$\frac{d^s F(0)}{dx^s} = \frac{d^s f(0)}{dx^s}, \quad s = 0, 1, 2, \dots, N - 1,$$

Пусть $r = N - 1$. Так как $F^{(s)}(x)$ имеет производную любого порядка в каждой точке вещественной оси, то функцию $F^{(s)}(x)$ можно разложить в ряд Тейлора в окрестности точки $x = 0$. Так, для $s = 0$ имеем:

$$F(x) = F(0) + F'(0)x + \frac{1}{2!}F''(0)x^2 + \dots + \frac{1}{r!}F^{(r)}(0)x^r + R_{r,0}(x),$$



где $R_{r,0}(x)$ есть остаточный член форме Лагранжа:

$$R_{r,0}(x) = \frac{F^{(r+1)}(\zeta_0)}{(r+1)!} x^{r+1}.$$

Имеем:

$$F^{(s)}(x) = \sum_{l=s}^r \frac{F^{(l)}(0)x^{l-s}}{(l-s)!} + R_{r,s}(x),$$

где $R_{r,s}(x)$ есть остаточный член в форме Лагранжа,

$$R_{r,s}(x) = \frac{F^{(r+1)}(\zeta_s)}{(r+1-s)!} x^{r+1-s}.$$

Имеем:

$$\sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| = \sum_{s=0}^{N-1} |R_{r,s}(x)|.$$

Согласно лемме 1

$$R_{r-s}(x) = \frac{x^{r+1-s}}{(r+1-s)!} \sum_{j=1}^N c_j w_j^{r+1} Q_{r+1}(\sigma(\theta_j(\zeta_s))),$$

где $-x_{\max} \leq \zeta_s \leq x_{\max}$.

Возьмем $0 < \delta < 1$. Положим $w_j = (j-1)\delta$, $j = 1, 2, \dots, N$. Согласно лемме 1 для $s \geq 1$

$$Q_s(\sigma(\theta_j)) = (-1)^{s+1} s! \theta_j^{-s} \sigma^s(\theta_j) (1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j)$$

и $Q_0(\sigma_j) = \sigma_j$, где $\sigma_j(x) = \sigma(\theta_j(x))$. Так как для всех x имеет место $-1 < \sigma_j < 1$, то $|Q_0(\sigma_j)| \leq 1$ и существует M^0 такое, что $|\sigma^s(\theta_j)(1 - \sigma(\theta_j) \cdot \operatorname{sgn} \theta_j)| < M^0$. Имеем, учитывая лемму 2,

$$\begin{aligned} \sum_{s=0}^{N-1} |R_{r,s}(x)| &\leq \sum_{s=0}^{N-1} \left(\frac{x_{\max}^{r+1-s}}{(r+1-s)!} \sum_{j=1}^N |c_j| |w_j^{r+1}| |Q_{r+1}(\sigma_j(\zeta_s))| \right) \leq \\ &\leq \sum_{s=0}^{N-1} \left(\frac{x_{\max}^{r+1-s}}{(r+1-s)!} \psi_{\max} q_{\max} \sum_{j=1}^N \sum_{p=0}^r |p! a_p| \delta^{-p} (r+1)! j^{r+1} \delta^{r+1} M^0 \right) \leq \\ &\leq \delta \psi_{\max} q_{\max} M^0 \sum_{p=0}^r |p! a_p| \sum_{s=0}^{N-1} \left(\frac{(r+1)!}{(r+1-s)!} x_{\max}^{r+1-s} \sum_{j=1}^N j^{r+1} \right). \end{aligned}$$

Обозначим

$$M = \psi_{\max} q_{\max} M^0 \sum_{p=0}^r |p! a_p| \sum_{s=0}^{N-1} \left(\frac{(r+1)!}{(r+1-s)!} x_{\max}^{r+1-s} \sum_{j=1}^N j^{r+1} \right),$$

т. е. M есть конечное число, не зависящее от x . Имеем:

$$\sum_{s=0}^{N-1} |R_{r,s}(x)| \leq \delta M.$$

Взяв $\delta < \varepsilon/(M+1)$, получаем $\sum_{s=0}^N |R_r(x)| < \varepsilon$, где ε есть ошибка аппроксимации. \square

Заметим, что утверждение, аналогичное теореме, не будет справедливым, если в скрытом слое содержится не $r+1$, а r узлов, т. е. нейронные сети с r узлами в скрытом слое не способны с любой степенью точности приближать полиномы степени r . Это означает, что существует неустранимая ошибка приближения многочленов степени r нейронными сетями с r узлами в скрытом слое. В работе [4] находится оценка этой ошибки.

Отметим, что используя подход, изложенный в работе [3], легко получить следующий результат для нейронных сетей прямого распространения с несколькими входами и одним выходом.



Теорема 2. Пусть $x_{\max} > 0$ есть произвольное действительное число, $I = [-x_{\max}, x_{\max}]$. Если $r \leq \sqrt{2N + 1.5^2} - 1.5$, то для любого $\varepsilon > 0$ и произвольной полиномиальной функции $f : \mathbb{R}^m \rightarrow \mathbb{R}$ степени r с m переменными вида (1) существует сеть прямого распространения $F : \mathbb{R}^m \rightarrow \mathbb{R}$ с одним скрытым слоем, содержащим N узлов и функцией активации (2), которая равномерно аппроксимирует f с любой наперед заданной степенью точности на I^m , т. е.

$$\sup_{x \in I^m} |F(x) - f(x)| < \varepsilon.$$

Библиографический список

1. Галушкин А. Нейронные сети. Основы теории. М. : Горячая Линия – Телеком, 2012.
2. Haykin S. Neural networks : a comprehensive foundation. 2 ed. N. J. : Prentice-Hall, Inc., 1999.
3. Malakooti B., Zhou Y. Q. Approximating polynomial functions by feedforward artificial neural networks : capacity analysis and design // Appl. Math. and Comp. 1998. Vol. 90. P. 27–52.
4. Сидоров С. П. Об ошибке приближения алгебраических полиномов нейронными сетями прямого распространения // Нейрокомпьютеры : разработка, применение. 2005. № 5. С. 13–17.

Simultaneous Approximation of Polynomial Functions and Its Derivatives by Feedforward Artificial Neural Networks with One Hidden Layer

N. S. Uzentsova, S. P. Sidorov

Saratov State University, Russia, 410012, Saratov, Astrahanskaya st., 83, UzentsovaNS@gmail.com, SidorovSP@info.sgu.ru

In this paper we propose the algorithm for finding weights of feedforward artificial neural networks with one hidden layer to approximate polynomial functions and its derivatives with a given error. We use the rational sigmoidal function as a transfer function.

Key words: neural networks, approximation.

References

1. Galyshkin A. *Neironnye seti. Osnovy teorii* [Neural Networks. The foundations of the theory]. Moscow, Goriachaia Liniia – Telekom, 2012 (in Russian).
2. Haykin S. *Neural Networks : A Comprehensive Foundation*. Second ed., New Jersey, Prentice-Hall, Inc., 1999.
3. Malakooti B., Zhou Y. Q. Approximating Polynomial Functions by Feedforward Artificial Neural Networks : Capacity Analysis and Design. *Appl. Math. and Comp.*, 1998, vol. 90, pp. 27–52.
4. Sidorov S. P. Ob oshibke priblizheniia algebraicheskikh polinomov neironnymi setiami priamogo rasprostraneniia [On the error of approximation of algebraical polynomials by means of artificial feedforward neural networks]. *Neirokomp'iutery : razrabotka, primeneniie*, 2005, no. 5, pp. 13–17 (in Russian).

УДК 519.872

МЕТОД УПРАВЛЕНИЯ МАРШРУТИЗАЦИЕЙ В СЕТЯХ МАССОВОГО ОБСЛУЖИВАНИЯ С ПЕРЕМЕННОЙ ТОПОЛОГИЕЙ

Н. П. Фокина¹, И. Е. Тананко²

¹Кандидат физико-математических наук, доцент кафедры системного анализа и автоматического управления, Саратовский государственный университет им. Н. Г. Чернышевского. FokinaNP.sgu@gmail.com

²Кандидат физико-математических наук, доцент кафедры системного анализа и автоматического управления, Саратовский государственный университет им. Н. Г. Чернышевского, TanankoIE@info.sgu.ru

Рассматриваются замкнутые экспоненциальные сети массового обслуживания с изменяющейся топологией. Предложен метод управления маршрутизацией в сетях обслуживания данного типа.

Ключевые слова: сети массового обслуживания, переменная топология, управление маршрутизацией, надежность.