Тогда существует $f \in V_p$, для которой $E_n(f)_{V_p} \leqslant C\varepsilon_n$, $n \in \mathbb{N}$, и ряд $\sum\limits_{n=1}^{\infty} n^\delta(\varphi(|a_n|) + \varphi(|b_n|))$ *расходится.*

Теоремы 1–4 обобщают результаты из [3].

СПИСОК ЛИТЕРАТУРЫ

1. *Терехин А. П.* Приближение функций ограниченной p-вариации // Изв. вузов. Математика., 1965, № 2, С. 171-187.

2. *Gogoladze L., Meskhia R.*, On the absolute convergence of trigonometric Fourier series // Proc. Razmadze Math. Inst., 2006, V. 141, P. 29-40.

3. *Volosivets S. S.* Convergence of series of Fourier coefficients of p-absolutely continuous functions // Analysis Math., 2000, V. 26, № 1, P. 63-80.

УДК 519.651, 519.652, 517.518.85, 519.852

АППРОКСИМАЦИЯ ПОЛИНОМИАЛЬНОЙ ФУНКЦИИ И ЕЕ ПРОИЗВОДНЫХ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТЬЮ ПРЯМОГО РАСПРОСТРАНЕНИЯ В ПРОСТРАНСТВЕ L^p [1]

Н. С. Узенцова, С. П. Сидоров (Саратов, РФ)

uzentsovans@gmail.com

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.

Рассмотрим архитектуру нейронной сети F с одним скрытым слоем. Пусть вектор $x = (x_1, x_2, \ldots, x_m)^T \in \mathbb{R}^m$ есть вектор входа нейронной сети, m есть количество узлов входного слоя сети F. Обозначим N количество узлов скрытого слоя сети. Количество входов в скрытый слой равно числу компонент вектора x. Обозначим $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ выходной вектор сети F. Вес связи между i-м узлом входного слоя и j-м

[1]Работа выполнена при финансовой поддержке РФФИ (проект 13-01-00238).

скрытым узлом обозначим w_{ji}, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Вес связи между j-м узлом скрытого слоя и k-м узлом выходного слоя обозначим c_{kj}. Выход скрытого узла j определяется следующим образом:

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

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

Компонентами выходного вектора $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ нейронной сети F являются числа

$$y_k = \sum_{j}^{N} c_{kj} \sigma(\theta_j), \quad k = 1, \ldots, n.$$

В настоящей статье, следуя идеям работы [3], мы находим алгоритм настройки весов сети для одновременного приближения в метрике L^p 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 приближения алгебраического многочлена совместно с его производными в метрике L^p.

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

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

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

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

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_r x^r,$$

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

В следующей теореме показывается, что сеть с одним скрытым слоем, содержащем N узлов, может аппроксимировать любой алгебраический многочлен степени $N-1$ совместно с его производными с любой степенью точности в метрике L^p.

Теорема 1. Пусть $x_{max} > 0$ — произвольное число. Тогда $\forall f \in \mathcal{P}_{N-1}$, $\forall \varepsilon > 0$, $\exists F(x) \in \mathcal{F}_N(\sigma)$ такая, что:

$$\sum_{s=0}^{N-1} \left(\int_{-x_{max}}^{x_{max}} (F^{(s)}(x) - f^{(s)}(x))^p dx \right)^{1/p} < \varepsilon,$$

где ε является ошибкой равномерного приближения на $[-x_{max}, x_{max}]$ в метрике L^p.

Доказательство. Доказательство данной теоремы следует из основного результата работы [5], где показано, что для произвольного ε_0 найдется $F(x) \in \mathcal{F}_N(\sigma)$ такая, что

$$\sup_{x \in [-x_{max}, x_{max}]} \sum_{s=0}^{N-1} |F^{(s)}(x) - f^{(s)}(x)| < \varepsilon_0.$$

Имеем

$$\sum_{s=0}^{N-1} \left(\int_{-x_{max}}^{x_{max}} (F^{(s)}(x) - f^{(s)}(x))^p \, dx \right)^{1/p} \leqslant \sum_{s=0}^{N-1} \left(\int_{-x_{max}}^{x_{max}} \varepsilon_0^p dx \right)^{1/p}.$$

Взяв число ε_0 таким образом, чтобы $\varepsilon_0 (2x_{max})^{1/p} N < \varepsilon$, получаем утверждение теоремы.

Теорема доказана.

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

СПИСОК ЛИТЕРАТУРЫ

1. *Галушкин А.* Нейронные сети. Основы теории. М.: Горячая Линия – Телеком, 2012.

2. *Haykin S.* Neural Networks: A Comprehensive Foundation. Second ed. NJ : Prentice Hall, Upper Saddle River, 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. 11–17.

5. *Узенцова Н. С., Сидоров С. П.* Об одновременном приближении алгебраических многочленов и их производных нейронными сетями прямого распространиия сигнала с одним скрытым слоем. Известия Саратовского университета. Новая серия. Серия Математика. Механика. Информатика. 2013. Т. 13, вып. 2, ч. 2. С. 78–82.

УДК 517.984

УПРАВЛЕНИЕ ПРОЦЕССОМ ИЗМЕРЕНИЯ В ДИНАМИЧЕСКИХ СИСТЕМАХ[1]

А. В. Фоминых (Санкт-Петербург, РФ)
alexfomster@mail.ru

Доклад посвящён нелинейной системе, характерезующей непрерывный закон изменения ковариационной матрицы наблюдаемого объекта [1]

$$\dot{D}(t) = A(t)D(t) + D(t)A^T(t) - D(t)V(t)D(t) + K(t),$$
$$D(0) = D_0, \quad t \in [0, T], \tag{1}$$

Здесь $K(t)$ характеризует возмущения, действующие на систему (если они отсутствуют, то $K(t) \equiv 0$). $V(t)$ характеризует процесс наблюдения и зависит только от состояния измерителей. Если можно изменять выбор наблюдаемых параметров или точность измерения, то $V(t)$ играет роль управления.

Рассмотрим функционал $I = \int_0^T f_0(V,t)dt$, где $f_0(V,t) \equiv 1$. Тогда $I = T$ представляет собой длительность наблюдения. Теперь можно сформулировать задачу оптимального по быстродействию управления наблюдением V

$$\dot{D}(t) = A(t)D(t) + D(t)A^T(t) - D(t)V(t)D(t) + K(t),$$
$$D(0) = D_0, \quad D(T) = D_T, \quad I = T \to \min_V, \quad V \in U, \tag{2}$$

где D_0 — начальное условие, D_T — значение матрицы D в конечный момент времени T, U — класс допустимых управлений.

Сведём задачу быстродействия (2) к задаче оптимального управления в смысле демпфирования функции W, которая характеризует расстояние от текущего значения $D(t)$ до заданного D_T. Тогда для решения

[1]Работа выполнена при финансовой поддержке РФФИ (проект 12-01-00752).