Кучумов Е.В. Применение инвариантных (автоморфных) функций в качестве функций активации искусственных нейронных сетей. // Проблемы информатики в образовании, управлении, экономике и технике: Сб. статей XIV Междунар. научно-техн. конф. – Пенза: ПДЗ, 2014. – С. 137-140.

УДК 519.673:517.442:517.443

ПРИМЕНЕНИЕ ИНВАРИАНТНЫХ (АВТОМОРФНЫХ) ФУНКЦИЙ В КАЧЕСТВЕ ФУНКЦИЙ АКТИВАЦИИ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Е.В. Кучумов

APPLICATION OF INVARIANT (AUTOMORPHIC) FUNCTIONS AS ARTIFICIAL NEURAL NETWORKS ACTIVATION FUNCTIONS

E.V. Kuchumov

Аннотация. В данной работе исследована возможность применения инвариантных и, в частности, автоморфных функций в качестве функций активации для искусственных нейронных сетей с целью их упрощения, а также повышения скорости, эффективности и устойчивости решения с их помощью некоторых некорректных математических задач, представимых в виде интегральных уравнений.

Ключевые слова: искусственные нейронные сети, инвариантные функции, автоморфные функции, интегральные преобразования, преобразование Лапласа, преобразование Фурье.

Abstract. This article discusses a probable using of invariant and participle automorphic functions as activation functions for artificial neural networks with purpose of it simplifying and building up speed, efficiency and stability of its solving some incorrect mathematical problems presented by integral equations.

Keywords: artificial neural networks, invariant functions, automorphic functions, integral transformations, Laplace transform, Fourier transform.

Искусственные нейронные сети (ИНС) являются эффективными средствами аппроксимации, что позволяет использовать их при решении многих задач математической физики [1]. В применении интегральных преобразований для их решения показали свою перспективность радиально-базисные сети [2]. Поэтому представляет интерес рассмотреть функции, используемые в качестве функций активации ИНС, с особенными свойствами, дающими преимущество при решении этих задач. В данной статье выбран класс инвариантных функций.

Инвариантной, в общем смысле, называется функция $f(x) \in F$, которая удовлетворяет условию

$$f(\mathrm{T}(x)) = a(x) f(x) + b(x), \; x \in X, \; \mathrm{T} \in L, \tag{1}$$

где $\mathrm{T}(x)$ – инварианты множества X, образующих группу L, $a(x)$ и $b(x)$ – некоторые заданные функции. Если функция $f(x)$ дифференцируема, то она называется автоморфной [3,4]. Простейшим примером автоморфных функций, удовлетворяющих соотношению (1) с множителями $a(x) \equiv 1$ и $b(x) \equiv 0$, являются периодические функции на действительной оси, с группой преобразований $\mathrm{T}_n(x) = x + n \cdot p$, где $n \in Z$ и p – период функции.

Рассмотрим операторное уравнение с правой частью выражения (1)

$$Af(x) = a(x)f(x) + b(x), \tag{2}$$

где A – заданный оператор, отображающий на функциональное подпространство F с тем же носителем X. В случае $a(x) = Const$ и $b(x) \equiv 0$ соотношение (2) является классической задачей на собственные значения, а тождество $f(\mathrm{T}(x)) = Af(x)$ означает, что автоморфная (инвариантная) функция является собственной функцией оператора A. Далее будем рассматривать соотношения (1)-(2) при условии $b(x) \equiv 0$.

Таким образом, если отталкиваться от тождества $f(\mathrm{T}(x)) = Af(x)$, то наличие инвариантов группы L преобразований $\mathrm{T}(x)$ для функции $f(x)$, соответствующие симметриям оператора A, обозначает, что эта функция является собственной функции у данного оператора. Однако инварианты оператора A не обязаны быть симметриями собственной функции $f(x)$.

Покажем это на примере оператора двойного дифференцирования $A = \partial_x^2$. Его собственными функциями являются периодические функции синуса и косинуса, инвариантные относительно сдвига на период $p = 2\pi$. Однако данный оператор инвариантен к любому сдвигу на постоянную величину, а не только на кратный период данной функции. Поэтому группа смещений $\mathrm{T}_n(x)$ с другим периодом p, оставляющий этот оператор неизменным, для данных функций уже не является инвариантом.

Если рассматривать оператор A как некое преобразование, например, интегральное, то выражение (2) так же характеризует существование собственных функций у этого оператора преобразования. Однако под выражение (2) подходят не все важные преобразования, в частности одномерное преобразование Лапласа, т.к. трансформанта определяется на комплексной плоскости. Для выполнения (2) нужно искусственно интерпретировать образ только на действительной положительной полуоси. По соответствующим таблицам трансформант [5] легко проверить, что единственная функция, удовлетворяющая выражению (2) с множителями $a(x) \equiv \sqrt{\pi}$ – это степенная функция x^α с показателем $\alpha = -1/2$, а также производные от неё функции, полученные с помощью сдвига и масштабирующего множителя. Если функция представима в виде ряда (аналитическая), то она не будет удовлетворять (2) для этого преобразования.

Данного недостатка лишено одномерное экспоненциальное преобразования Фурье, т.к. области определения у образа и прообраза эквивалентны. Благодаря чему кроме указанной степенной функции существуют ещё функции, удовлетворяющие условию (2), а именно показательная функция $\exp(-x^2)$, где после преобразования Фурье осуществляется замена $x = \omega/2$, с множителем $a(x) \equiv \sqrt{\pi}$. Выражение x^2 можно интерпретировать как квадрат модуля комплексного числа $p = u + iw$, т.е. $x^2 = pp^*$. Тогда инвариантным преобразованием для данной функции является поворот с произвольным углом по окружности радиуса x на комплексной плоскости вокруг начала координат. Для совпадения соотношений (1) и (2) в этом случае необходимо, чтобы оператор A был экспоненциальным преобразованием Фурье, нормированным на число $\sqrt{\pi}$.

Вообще говоря, можно просто применять явные выражения для выбранного интегрального преобразования конкретной функции. Это позволяет избежать части промежуточных вычислений, тем самым поднять устойчивость решений математических задач с помощью применения выбранного типа преобразования [6]. Удобство от применения инвариантных (автоморфных) функций в качестве функций активации для ИНС заключается в сведении результата преобразования к относительно простой, стандартизированной замене переменных и множителей благодаря их симметриям.

Библиографический список

1. Горбаченко В.И. Нейрокомпьютеры в решении краевых задач теории поля. – М.: Радиотехника, 2003. – 336 с.
2. Бойков И.В., Кучумов Е.В. Об одном итерационном методе решения интегральных уравнений Вольтерра // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. – 2009. – №2 (10). – С. 25–38.
3. Форд Л.Р. Автоморфные функции. – М.: ОНТИ, 1936. – 340 с.
4. Гурвиц А., Курант Р. Теория функций. – М.: Наука, 1968. – 648 с.
5. Бейтмен Г., Эрдейи А. Таблицы интегральных преобразований. Том I. Преобразование Фурье, Лапласа, Меллина. – М.: Наука, 1969. – 344 с.
6. Кучумов Е.В. О прямом и обратном преобразовании Лапласа на радиально-базисных сетях с применением к численному решению обыкновенных дифференциальных уравнений // Проблемы информатики в образовании, управлении, экономике и технике : Сб. статей XII Междунар. науч.-техн. конф. – Пенза: ПДЗ, 2012. – С. 34-40.

Кучумов Евгений Владимирович
ОАО «Научно-исследовательский институт физических измерений», г. Пенза, Россия
E-mail: evgenii_kuchumov@mail.ru

Kuchumov Yevgeniy Vladimirovich
OJSC “Research institute of physical measurements”, Penza, Russia