ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

1998 МАТЕМАТИКА № 5 (432)

УДК 517.518:519.713

А.Н. ГОРБАНЬ

ОБОБЩЕННАЯ АППРОКСИМАЦИОННАЯ ТЕОРЕМА И ТОЧНОЕ ПРЕДСТАВЛЕНИЕ МНОГОЧЛЕНОВ ОТ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ СУПЕРПОЗИЦИЯМИ МНОГОЧЛЕНОВ ОТ ОДНОГО ПЕРЕМЕННОГО

1. Введение

Вопрос о представлении непрерывных функций многих переменных с помощью суперпозиций непрерывных функций малого числа переменных составлял ядро 13-й проблемы Гильберта. В.И. Арнольдом и А.Н. Колмогоровым был получен ответ [1], [2]: любую непрерывную функцию многих переменных можно представить с помощью операции суперпозиции с использованием только функции двух переменных.

В работе А.Н. Колмогорова [3] доказана также возможность представления непрерывных функций многих переменных с помощью непрерывных функций одного переменного.

Возврат к вопросу о представлении функций многих переменных с помощью суперпозиций и сумм функций одного переменного связан с исследованиями нейронных сетей. В прикладных задачах аппроксимации в последнее время значительное внимание уделяется устройствам, вычисляющим суперпозиции простых функций одного переменного и их линейных комбинаций. Эти устройства получили название "искусственные нейронные сети" (*neural networks*) [4], [5]. Актуальным становится вопрос о том, какие именно функции могут быть аппроксимированы. Именно поэтому вновь оживился интерес к классическому вопросу, у которого имеются два варианта: точный и аппроксимационный.

1. Можно ли получить *точное* представление функции многих переменных с помощью суперпозиции функций меньшего числа переменных?
2. Можно ли получить *сколь угодно точную аппроксимацию* функции многих переменных с помощью некоторых более простых функций и операций?

Теоремы Колмогорова-Арнольда [1]–[3] посвящены первому варианту вопроса. Наиболее известным результатом, касающимся второго варианта вопроса, является теорема Стоуна-Вейерштрасса [6].

В рамках первого варианта вопроса особый интерес представляют конструкции, в которых для точного представления всех функций многих переменных используется один и тот же набор функций одного переменного. Традиционно считается, что эти функции должны иметь весьма специальный и довольно экзотический вид (напр., как в работе А.Н. Колмогорова [3], где использовались существенно негладкие функции).

Напротив, свобода в выборе функций одного переменного для решения второго варианта вопроса при том же самоограничении (один набор функций одного переменного — для приближенного представления всех функций многих переменных) очень велика. Для этого можно использовать практически любую нелинейную функцию и достаточно всего одной, что было доказано при исследовании нейронных сетей [7] (см. также [8]–[11]).

Работа выполнена при поддержке Красноярского краевого фонда науки, грант 6F0124.

В данной статье доказываются теоремы, относящиеся к первому варианту вопроса (точное представление). В частности, показано, что можно точно представить любой многочлен от многих переменных с помощью суперпозиций произвольного нелинейного многочлена от одного переменного и линейных функций. Эти теоремы позволили доказать обобщенную аппроксимационную теорему Стоуна-Вейерштрасса.

2. Полугруппы многочленов от одного переменного

Пусть $\mathbf{R}[X]$ — кольцо многочленов от одного переменного над полем $\mathbf{R}$ характеристики 0, $\mathbf{E} \subset \mathbf{R}[X]$ – линейное пространство многочленов над $\mathbf{R}$.

Теорема 1. *Если $\mathbf{E}$ замкнуто относительно суперпозиции многочленов, содержит все многочлены первой степени и хотя бы один многочлен $p(x)$ степени $m > 1$, то $\mathbf{E} = \mathbf{R}[X]$.*

Доказательство. Заметим, что степень многочлена $p'(x)=p(x+1) - p(x)$ равна $m - 1$, и $p'(x) \in \mathbf{E}$, т. к. $\mathbf{E}$ содержит многочлены первой степени (поэтому $x + 1 \in \mathbf{E}$), замкнуто относительно суперпозиции (поэтому $p(x + 1) \in \mathbf{E}$) и линейных операций (поэтому $p'(x) \in \mathbf{E}$). Если $m > 2$, то понижаем степень с помощью конечных разностей (переходим к p', p'' и т. д.), пока не получим многочлен второй степени. Вычитая из него линейную часть и умножая на константу, получаем $x^2 \in \mathbf{E}$. Поэтому для любого $f \in \mathbf{E}$ имеем $f^2 \in \mathbf{E}$ (т. к. $\mathbf{E}$ — полугруппа). Дальнейшее очевидно в силу тождества

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2],$$

а с помощью умножения и сложения многочленов первой степени порождается все кольцо $\mathbf{R}[X]$. $\square$

3. Представление многочленов от многих переменных

Обозначим через $\mathbf{R}[X_1, \dots, X_n]$ кольцо многочленов от n переменных над полем $\mathbf{R}$ характеристики 0. Пусть p — многочлен от одного переменного, $\mathbf{E}_p[X_1, \dots, X_n]$ — множество многочленов от n переменных, которое можно получить из p и многочленов первой степени, принадлежащих $\mathbf{R}[X_1, \dots, X_n]$, с помощью операций суперпозиции, сложения и умножения на число.

Следующие два предложения дают удобную для дальнейшего характеризацию $\mathbf{E}_p[X_1, \dots, X_n]$ и следуют непосредственно из определений.

Предложение 1. *Множество $\mathbf{E}_p[X_1, \dots, X_n]$ является линейным пространством над $\mathbf{R}$ и для любого многочлена $g(x_1, \dots, x_n)$ из $\mathbf{E}_p[X_1, \dots, X_n]$*

$$p(g(x_1, \dots, x_n)) \in \mathbf{E}_p[X_1, \dots, X_n].$$

Предложение 2. *Для данного p семейство линейных подпространств $\mathbf{L} \subseteq \mathbf{R}[X_1, \dots, X_n]$, содержащих все многочлены первой степени и удовлетворяющих условию:*

$$\textit{если} \quad g(x_1, \dots, x_n) \in \mathbf{L}, \quad \textit{то} \quad p(g(x_1, \dots, x_n)) \in \mathbf{L},$$

замкнуто относительно пересечений. Минимальным по включению элементом этого семейства является $\mathbf{E}_p[X_1, \dots, X_n]$.

Для любого линейного подпространства $\mathbf{E} \subseteq \mathbf{R}[X_1, \dots, X_n]$ рассмотрим множество алгебраических унарных операций, которые переводят элементы $\mathbf{E}$ в элементы $\mathbf{E}$,

$$\mathbf{P_E} = \{p \in \mathbf{R}[X] \mid p(g(x_1, \dots, x_n)) \in \mathbf{E} \quad \text{для любого} \quad g(x_1, \dots, x_n) \in \mathbf{E}\}. \tag{1}$$

Предложение 3. *Для любого линейного подпространства $\mathbf{E} \subseteq \mathbf{R}[X_1, \dots, X_n]$ множество полиномов $\mathbf{P_E}$ является линейным пространством над $\mathbf{R}$, замкнуто относительно суперпозиции и содержит все однородные многочлены первой степени. Если линейное пространство $\mathbf{E}$ содержит 1, а $\mathbf{P_E}$ включает хотя бы один многочлен степени $m > 1$ (т. е. нелинейный), то $\mathbf{P_E} = \mathbf{R}[X]$.*

Доказательство. Замкнутость $\mathbf{P_E}$ относительно суперпозиции следует из определения, все однородные полиномы первой степени входят в $\mathbf{P_E}$, поскольку $\mathbf{E}$ является линейным пространством, отсюда также следует, что $\mathbf{P_E}$ является линейным пространством. Наконец, если $1 \in \mathbf{E}$ и $\mathbf{P_E}$ содержит многочлен степени $m > 1$, то $1 \in \mathbf{P_E}$, тогда $\mathbf{P_E} = \mathbf{R}[X]$ по теореме 1. $\square$

Теорема 2. *Пусть* $\mathbf{E} \subseteq \mathbf{R}[X_1, \ldots, X_n]$ *— линейное подпространство,* $\mathbf{P_E}$ *содержит хотя бы один многочлен степени* $m > 1$ *и* $1 \in \mathbf{E}$*, тогда* $\mathbf{E}$ *является кольцом (с единицей).*

Доказательство. По предложению 3 в условиях теоремы $\mathbf{P_E} = \mathbf{R}[X]$. В частности, $x^2 \in \mathbf{P_E}$. Это означает, что для любого $f \in \mathbf{E}$ также и $f^2 \in \mathbf{E}$. Поэтому для любых $f, g \in \mathbf{E}$ аналогично доказательству теоремы 1 получаем $fg \in \mathbf{E}$. $\square$

Теорема 3. *Для любого многочлена* p *степени* $m > 1$

$$\mathbf{E}_p[X_1, \ldots, X_n] = \mathbf{R}[X_1, \ldots, X_n].$$

Доказательство. Заметим, что $\mathbf{E} = \mathbf{E}_p[X_1, \ldots, X_n]$ — линейное подпространство в $\mathbf{R}[X_1, \ldots, X_n]$, $\mathbf{P_E}$ содержит хотя бы один многочлен (p) степени $m > 1$ и $\mathbf{E}$ содержит все многочлены первой степени (и поэтому также 1). В силу теоремы 2 $\mathbf{E}$ является кольцом, а т. к. оно содержит все многочлены первой степени, то совпадает с $\mathbf{R}[X_1, \ldots, X_n]$, поскольку, используя умножение и сложение, из этих многочленов получим любой.

Таким образом, из p и многочленов первой степени с помощью операций суперпозиции, сложения и умножения на число можно получить все элементы $\mathbf{R}[X_1, \ldots, X_n]$. $\square$

4. Обобщенная теорема Стоуна-Вейерштрасса

Проиллюстрируем некоторые приложения полученных теорем. Пусть далее $\mathbf{R}$ — поле вещественных чисел, X — компактное топологическое пространство, $\mathbf{C}(X)$ — пространство непрерывных функций на X со значениями в $\mathbf{R}$. В теореме Стоуна-Вейерштрасса рассматриваются подпространства $\mathbf{C}(X)$, состоящие из функций, разделяющих точки в X: если такое подпространство замкнуто, является кольцом и содержит 1, то оно совпадает со всем $\mathbf{C}(X)$.

Основной смысл следующих двух теорем состоит в том, что они позволяют в теореме Стоуна-Вейерштрасса рассматривать не только кольца (с обычным произведением элементов), но и произвольные алгебры с любыми нелинейными алгебраическими операциями.

Пусть $\mathbf{E} \subseteq \mathbf{C}(X)$ – линейное подпространство. Сохраним обозначение (1). Аналогично предложению 3 и теореме 2 формулируются и доказываются следующие утверждения.

Предложение 4. *Для любого линейного подпространства* $\mathbf{E} \subseteq \mathbf{C}(X)$ *множество полиномов* $\mathbf{P_E}$ *является линейным пространством над* $\mathbf{R}$*, замкнуто относительно суперпозиции и содержит все однородные многочлены первой степени. Если линейное пространство* $\mathbf{E}$ *содержит* 1*, а* $\mathbf{P_E}$ *включает хотя бы один многочлен степени* $m > 1$*, то* $\mathbf{P_E} = \mathbf{R}[X]$.

Теорема 4. *Пусть* $\mathbf{E} \subseteq \mathbf{C}(X)$ *— замкнутое линейное подпространство и* $\mathbf{E}$ *содержит* 1*, а* $\mathbf{P_E}$ *включает хотя бы один многочлен степени* $m > 1$*. Тогда* $\mathbf{E}$ *является кольцом (с единицей).*

Непосредственно из теоремы Стоуна-Вейерштрасса и теоремы 4 получаем следующую *обобщенную аппроксимационную теорему.*

Теорема 5. *Пусть* $\mathbf{E} \subseteq \mathbf{C}(X)$ *— замкнутое линейное подпространство, функции из* $\mathbf{E}$ *разделяют точки в* X*,* $\mathbf{E}$ *содержит* 1*, а* $\mathbf{P_E}$ *включает хотя бы один многочлен степени* $m > 1$*. Тогда* $\mathbf{E} = \mathbf{C}(X)$.

Приведем без подробного доказательства *обобщенную аппроксимационную теорему* для произвольных непрерывных (а не только полиномиальных) операций на $\mathbf{C}(X)$. Рассмотрим пространство $\mathbf{C}(\mathbf{R})$ непрерывных функций на вещественной оси $\mathbf{R}$ в топологии равномерной сходимости на компактах. Пространство $\mathbf{C}(\mathbf{R})$ является полугруппой относительно операции суперпозиции функций $(f \circ g)(x) = f(g(x))$. Единицей в этой полугруппе является функция id ($\mathrm{id}\,(x) = x$).

Теорема 5′. *Пусть* $\mathbf{E}$ *— замкнутое подпространство в* $\mathbf{C}(X)$, $1 \in \mathbf{E}$, *элементы* $\mathbf{E}$ *разделяют точки в* X *и для некоторой нелинейной функции* f *из* $\mathbf{C}(\mathbf{R})$ *и любой функции* g *из* $\mathbf{E}$ *функция* $f(g(x))$ *принадлежит* $\mathbf{E}$. *Тогда* $\mathbf{E} = \mathbf{C}(X)$.

Так же, как и для теоремы 5, доказательство базируется на исследовании множества тех f из $\mathbf{C}(\mathbf{R})$, для которых при данном $\mathbf{E}$ выполнены условия теоремы, а также на теореме о полноте полугрупп непрерывных функций, аналогичной теореме 1.

Теорема 1′. *Пусть* $\mathbf{E}$ *— замкнутое подпространство в* $\mathbf{C}(\mathbf{R})$, *являющееся подполугруппой,* $1 \in \mathbf{E}$ *и* $id \in \mathbf{E}$. *Тогда либо* $\mathbf{E} = C(\mathbf{R})$, *либо* $\mathbf{E}$ *— подпространство линейных функций* ($f(x) = ax + b$). *В частности, если* $\mathbf{E}$ *содержит хотя бы одну функцию, не являющуюся линейной, то* $\mathbf{E} = C(\mathbf{R})$.

Автор благодарен С.Е. Гилеву и Э.Э. Шнолю за полезное обсуждение.

Литература

1. Колмогоров А.Н. *О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных* // ДАН СССР. – 1956. – Т. 108. – № 2. – С. 179–182.
2. Арнольд В.И. *О функциях трех переменных* // ДАН СССР. – 1957. – Т. 114. – № 4. – С. 679–681.
3. Колмогоров А.Н. *О представлении непрерывных функций нескольких переменных в виде суперпозиций непрерывных функций одного переменного и сложения* // ДАН СССР. – 1957. – Т. 114. – № 5. – С. 953–956.
4. Горбань А.Н. *Обучение нейронных сетей.* – М.: Изд. СССР-США СП “ПараГраф”, 1990. – 160 с.
5. Zurada J.M. *Introduction to artificial neural systems.* – New York: PWS Publishing Company, 1992. – 785 p.
6. Stone M.N. *The generalized Weierstrass approximation theorem* // Math. Mag. – 1948. – V. 21. – P. 167–183.
7. Горбань А.Н., Россиев Д.А. *Нейронные сети на персональном компьютере.* – Новосибирск: Наука, 1996. – 276 с.
8. Cybenko G. *Approximation by superposition of a sigmoidal function* // Math. Control, Signals and Systems. – 1989. – V. 2. – P. 303–314.
9. Hornik K., Stinchcombe M., White H. *Multilayer feedforward networks are universal approximators* // Neural Networks. – 1989. – V. 2. – P. 359–366.
10. Kochenov D.A., Rossiev D.A. *Approximations of functions of* $C[A, B]$ *class by neural-net predictors* (*architectures and results*) // AMSE Transaction, Scientific Siberian, **A**. – Tassin, France.– 1993. – V. 6. Neurocomputing. – P. 189–203.
11. Gilev S.E., Gorban A.N. *On completness of the class of functions computable by neural networks* // Proc. of the World Congress on Neural Networks (WCNN’96). Sept. 15–18, 1996, San Diego, CA, Lawrens Erlbaum Accociates, 1996. – P. 984–991.

Вычислительный центр СО РАН,
Красноярск

Поступила
25.06.1997