Об одном алгебраическом подходе к обучению в теоретической нейроинформатике

Шибзухов З. М.
szport@gmail.com
НИИ прикладной математики и автоматизации КБНЦ РАН, Нальчик

В теоретической нейроинформатике известно, что любое непрерывное преобразование, определенное на компактном множестве, можно аппроксимировать при помощи сети из аддитивных сумматоров и функциональных преобразователей, реализующих произвольно заданную непрерывную нелинейную скалярную функцию [1, 2]. Из него следует, что любое непрерывное преобразование, определенное на компактном множестве, можно аппроксимировать при помощи сети из элементов, реализующих операции сложения, умножения, и произвольно выбранные нелинейные скалярные функции. Такие сети называют *полиномиальными сетями.* Оказывается, что можно построить *прямую алгебраическую процедуру* [3] для построения некоторых достаточно представительных классов полиномиальных сетей, обрабатывающих сигналы, кодируемые в произвольно заданном *коммутативном кольце*, не содержащем делителей нуля [4]. С ее помощью можно эффективно конструировать искусственные нейронные сети для решения задач распознавания и прогнозирования. Ниже рассматривается один класс полиномиальных нейронных сетей, содержащих специальный слой алгебраических *мультиплицирующих нейронов* и слой алгебраических *суммирующих нейронов.* Такая организация сети делает возможным построение прямой алгебраической процедуры обучения с учителем.

Алгебраические мультиплицирующие функции

Пусть $\mathbb{K}$ — это коммутативное кольцо, не содержащее делителей нуля, $\mathbb{X}$ — скалярная область значений входов, $\mathbb{Y}$ — скалярная область значений выходов. Пусть $\mathcal{F}_{0,1}(\mathbb{K})$ — некоторый класс функций $\eta\colon \mathbb{K} \to \mathbb{K}$, таких, что $\eta(p) = 0 \Leftrightarrow p = 0$ и $\eta(1) = 1$. Рассмотрим класс $\mathfrak{P}[x_1, \ldots, x_n]$ всех алгебраических мультиплицирующих функций, которые можно построить из тождественных функций $\mathrm{id}(x_1) = x_1, \ldots, \mathrm{id}(x_n) = x_n$, применяя операции умножения и композиции функций из $\mathcal{F}_{0,1}(\mathbb{K})$. Например, функции вида $\eta\bigl(\prod_{i=1}^n \varphi_i(x_i)\bigr)$, где $\eta \in \mathcal{F}_{0,1}(\mathbb{K})$ и $\varphi_i \in \mathcal{F}_{0,1}(\mathbb{K})$, принадлежат $\mathfrak{P}^n[x_1, \ldots, x_n]$.

Пусть $\Pi(\boldsymbol{x}) \in \mathfrak{P}[x_1, \ldots, x_n]$ — функция n переменных $\boldsymbol{x} = (x_1, \ldots, x_n)$. Для любого мультииндекса $\mathbf{i} \subset \{1, \ldots, n\}$ определена функция $\Pi(\boldsymbol{x}, \mathbf{i})$ вида $\mathbb{K}^r \to \mathbb{K}$, зависящая от набора переменных $\{x_i : i \in \mathbf{i}\}$, $r = |\,\mathbf{i}\,|$, и получающаяся из $\Pi(\boldsymbol{x})$ в результате «исключения переменных» с индексами из $\{i\colon i \notin \mathbf{i}\}$.

Множество функций $\mathcal{P}[\boldsymbol{x}] = \bigl\{\Pi(\boldsymbol{x}, \mathbf{i})\colon \mathbf{i} \subseteq \{1, \dots, n\}\bigr\}$ удовлетворяет следующим условиям:

1) если $\Pi(\boldsymbol{x}, \mathbf{i}) = 0$, то $\Pi(\boldsymbol{x}, \mathbf{i}') = 0$ для любого $\mathbf{i}' \supset \mathbf{i}$;
2) если $\Pi(\boldsymbol{x}, \mathbf{i}) \neq 0$, то $\Pi(\boldsymbol{x}, \mathbf{i}') \neq 0$ для любого $\mathbf{i}' \subset \mathbf{i}$.

Алгебраические суммирующие функции

Пусть $\mathbb{A}$ — некоторый $\mathbb{K}$–модуль. Пусть $\mathcal{F}_0(\mathbb{A})$ — некоторый класс отображений $\sigma\colon \mathbb{A} \to \mathbb{A}$ таких, что $\sigma(s) = 0 \Leftrightarrow s = 0$. Рассмотрим класс $\mathfrak{S}[s_1, \dots, s_N]$ всех алгебраических суммирующих функций, которые можно построить из тождественных функций $\mathrm{id}(s_1) = s_1, \dots, \mathrm{id}(s_n) = s_n$, применяя операции сложения, умножения на скаляры из $\mathbb{K}$ и композиции функций из $\mathcal{F}_0(\mathbb{A})$. Например, множества функций, определяемые индуктивно:

1) $\Sigma(s_1) = \sigma_1(s_1)$;
2) $\Sigma(s_1, \dots, s_k) = \sigma_k\bigl(\Sigma(s_1, \dots, s_{k-1}) + w_k s_k\bigr)$ при $k > 1$;

где $\sigma_k \in \mathcal{F}_0(\mathbb{A})$, $k = 1, \dots, N$, принадлежат $\mathfrak{S}[s_1, \dots, s_N]$.

Алгебраические ΣΠ–функции

Алгебраическая ΣΠ–функция реализует преобразование

$$\mathrm{spn}(\boldsymbol{x}) = \mathrm{out}\bigl(\mathrm{sp}(\boldsymbol{x})\bigr); \tag{1}$$

$$\mathrm{sp}(\boldsymbol{x}) = \Sigma\bigl(\theta(\boldsymbol{x}), \Pi_1(\boldsymbol{x}, \mathbf{i}_1), \dots, \Pi_N(\boldsymbol{x}, \mathbf{i}_N)\bigr); \tag{2}$$

где $\theta\colon \mathbb{X}^n \to \mathbb{A}$, $w_k \in \mathbb{A}$, $\Sigma(s_0, \dots, s_N) \in \mathfrak{S}[s_0, \dots, s_N]$, $\Pi_k(\boldsymbol{x}) \in \mathfrak{P}[x_1, \dots, x_n]$. Преобразование $\mathrm{out}\colon \mathbb{A} \to \mathbb{Y}$ является *допустимым*, т. е. для любых $s \in \mathbb{A}$, $0 \neq p \in \mathbb{K}$, $y \in \mathbb{Y}$ уравнение $\mathrm{out}(s + wp) = y$ имеет решение относительно w.

При $m = 1$ и $\mathbb{A} = \mathbb{K}$ соотношения (1)–(2) описывают алгебраический ΣΠ-нейрон $\mathrm{spn}(\boldsymbol{x}) = \mathrm{out}\bigl(\mathrm{sf}(\boldsymbol{x})\bigr)$. Если $m = 1$ и $\mathbb{A}$ — $\mathbb{K}$-модуль размерности, большей 1, то соотношения (1)–(2) описывают коллектив ΣΠ-нейронов, принимающих единственное решение, или алгебраический ΣΠ-корректор.

Треугольно упорядоченные последовательности алгебраических мультиплицирующих функций

Пусть $\mathbf{X} = \{\mathbf{x}_k\}$ — некоторая последовательность входов $\mathbf{x}_k \in \mathbb{X}^n$, $\mathbf{P} = \{\Pi_k(\boldsymbol{x}, \mathbf{i}_k)\}$ — последовательность алгебраических мультиплицирующих функций из $\mathfrak{P}[x_1, \dots, x_n]$, треугольно упорядоченная на $\mathbf{X}$, т. е. таких, что

1) $\Pi_k(\mathbf{x}_j, \mathbf{i}_k) = 0$ для любой пары $1 \leqslant j < k$;
2) $\Pi_k(\mathbf{x}_k, \mathbf{i}_k) \neq 0$ для всех k.

Она является минимальной, если каждая последовательность $\left\{\Pi_k(\boldsymbol{x}, \mathbf{i}'_k)\right\}$, где $\mathbf{i}'_k \subseteq \mathbf{i}_k$, не является треугольно упорядоченной на $\mathbf{X}$.

При определенных условиях для некоторых достаточно широких классов последовательностей $\{\mathbf{x}_k\}$ существуют конструктивные минимальные треугольно упорядоченные последовательности $\{\Pi_k(\boldsymbol{x}, \mathbf{i}_k)\}$ [4].

Лемма 1. *(о минимальных треугольно упорядоченных последовательностях алгебраических мультиплицирующих функций). Конструктивно перечисляются все минимальные последовательности* $\left\{\Pi_k(\boldsymbol{x}, \mathbf{i}^\star_k)\right\}$, *треугольно упорядоченные на* $\mathbf{X}$.

Рекуррентные последовательности алгебраических $\Sigma\Pi$–функций

Пусть задана последовательность ожидаемых на выходе значений $\mathbf{Y} = \{y_k\}$, соответствующая $\mathbf{X}$. Определяется последовательность алгебраических $\Sigma\Pi$-функций $\{\mathrm{spn}_k(\boldsymbol{x})\}$, где

$$\mathrm{sp}_k(\boldsymbol{x}) = \sigma_k\big(\mathrm{sp}_{k-1}(\boldsymbol{x}) + w_k\Pi_k(\boldsymbol{x}, \mathbf{i}_k)\big), \tag{3}$$

w_k — решение уравнения $\mathrm{out}\big(\sigma_k\big(\mathrm{sp}_{k-1}(\mathbf{x}_k) + w_k\Pi_k(\mathbf{x}_k, \mathbf{i}_k)\big)\big) = y_k$ относительно w_k, $\mathrm{sp}_0(\boldsymbol{x}) = \sigma_0\big(\theta(\boldsymbol{x})\big)$.

Лемма 2. *(о рекуррентной последовательности алгебраических* **ΣΠ***-функций). Для всех* $j = 1, \ldots, k$ *верно равенство* $\mathrm{spn}_k(\mathbf{x}_j) = y_j$.

Работа выполнена при поддержке ОМН РАН по проекту «Исследование конструктивных последовательностей алгебраических расширений распознающих алгоритмов» в 2007 году.

Литература

[1] *Gilev S. E., Gorban A. N.* On Completness of the Class of Functions Computable by Neural Networks // World Cong. on Neural Networks, 1996. — Pp. 984–991.

[2] *Горбань А. Н.* Возможности нейронных сетей. — Нейроинформатика, под ред. Новикова Е. А., Гл. 1. — Новосибирск: Наука, 1998. — С. 18–46.

[3] *Журавлев Ю. И.* Об алгебраическом подходе к решению задач распознавания и классификации / Избр. науч. тр. — М.: Магистр, 1998. — С. 229–323.

[4] *Шибзухов З. М.* Конструктивные методы обучения ΣΠ-нейронных сетей. — М.: Наука, 2006 — 159 с.