Тема 3

Обобщенная теорема Стоуна

В предыдущей лекции рассматривалась теорема, дающая ответ на вопрос о возможности точного представления функции многих переменных функциями одного переменного. Оказалось, что в классе непрерывных функций такое представление возможно. Но кроме вопроса о точном представлении существует еще один: возможно ли приблизить функцию многих переменных, используя функции одного переменного. Во многих случаях этот вопрос важнее, ведь вычисление большинства функций производится приближенно даже при наличии «точных» формул. Обобщенная теорема Стоуна дает утвердительный ответ на этот вопрос.

Определение. Подмножество *M* пространства *X* называется компактным, если всякая система открытых множеств пространства *X*, покрывающая *M*, содержит конечную подсистему, которая также покрывает множество *M*.

Определение. Пространство *X* называется нормальным, если для любых непересекающихся замкнутых множеств F_1, F_2 из *X* существуют такие непересекающиеся открытые множества G_1, G_2 что $F_1 \subseteq G_1, F_2 \subseteq G_2$.

Пусть $C(X)$ – множество всех вещественных непрерывных функций, определенных на *X*.

Определение. Говорят, что множество $B \subset C(X)$ разделяет точки пространства *X*, когда для любой пары точек $(s_1, s_2) \in X$ существует функция $x \in B$, удовлетворяющая условию $x(s_1) \neq x(s_2)$.

// Следующая теорема для ознакомления

Сначала докажем теорему.

Теорема 3.1 (Стоуна – Вейерштрасса). Пусть *X* – компактное пространство и $C(X)$ – множество всех вещественных непрерывных функций, определенных на *X*. Через *B* обозначим подмножество множества $C(X)$, удовлетворяющее следующим трем условиям:

1) если $f, g \in B$, то произведение $f \cdot g$ и линейные комбинации $\alpha f + \beta g$ с вещественными коэффициентами α, β также принадлежат *B*;
2) постоянная функция 1 принадлежит *B*;

3) предел f_∞ всякой равномерно сходящейся последовательности $\{f_n\}$ функций, принадлежащих подмножеству *B*, также принадлежит *B*.

Тогда $B = C(X)$ в том и только в том случае, когда B разделяет точки пространства *X*.

Доказательство. Необходимость указанного условия следует из того, что компактное пространство нормально, и поэтому по теореме Урысона [2] существует такая непрерывная вещественная функция $x \in B$, что $x(s_1) \neq x(s_2)$.

Для доказательства достаточности удобно ввести обозначения:

$(f \vee g)(s) = \max(f(s), g(s)),\ \ (f \wedge g)(s) = \min(f(s),\, g(s)),\ \ |f|(s) = |f(s)|$.

Согласно теореме Вейерштрасса об аппроксимации непрерывных функций полиномами, существует последовательность полиномов $\{P_n\}$, такая, что

$$\big||t| - P_n(t)\big| < \frac{1}{n} \qquad \text{при } -n \le t \le n.$$

Следовательно, $\big||f(s)| - P_n(f(s))\big| < 1/n$, если $-n \le f(s) \le n$. Отсюда, согласно условию (3), следует, что $|f| \in B$, если $f \in B$, так как всякая функция $f(s) \in B \subseteq C(X)$ ограничена на компактном пространстве *X*. Поэтому, учитывая соотношения

$$f \vee g = \frac{f+g}{2} + \frac{|f-g|}{2} \quad \text{и} \quad f \wedge g = \frac{f+g}{2} - \frac{|f-g|}{2},$$

мы заключаем, что множество *B* замкнуто по отношению к структурным операциям $\vee$ и $\wedge$.

Пусть элемент $h \in C(X)$ и точки $s_1, s_2 \in X$ выбраны произвольно, но так, что $s_1 \neq s_2$. Тогда мы сможем найти функцию $f_{s_1 s_2} \in B$, такую, что $f_{s_1 s_2}(s_1) = h(s_1)$ и $f_{s_1 s_2}(s_2) = h(s_2)$. В самом деле, выберем функцию $g \in B$ так, что $g(s_1) \neq g(s_2)$, и возьмем такие вещественные числа α и β, что функция $f_{s_1 s_2} = \alpha g + \beta$ удовлетворяет условиям

$$f_{s_1 s_2}(s_1) = h(s_1) \qquad \text{и} \qquad f_{s_1 s_2}(s_2) = h(s_2).$$

Пусть теперь заданы произвольное $\varepsilon > 0$ и точка $t \in X$. Тогда для всякой точки $s \in X$ существует окрестность *U*(*s*), такая, что $f_{st}(u) > h(u) - \varepsilon$ для всех $u \in U(s)$. Пусть далее множества $U(s_1), U(s_2), \ldots, U(s_n)$ покрывают компактное пространство *X*; положим

$$f_t = f_{s_1 t} \vee \ldots \vee f_{s_n t}.$$

Тогда $f_t \in B$ и $f_t(u) > h(u) - \varepsilon$ для всех точек $u \in X$. Кроме того, $f_t(t) = h(t)$, так как $f_{s_j t}(t) = h(t)$. Следовательно, существует такая окрестность $V(t)$ точки t, что

$f_t(u) < h(u) + \varepsilon$ при всех $u \in V(t)$. Пусть множества $V(t_1), V(t_2), \ldots, V(t_k)$ покрывают компактное пространство *X*. Определим функцию f соотношением

$$f = f_{t_1} \wedge \ldots \wedge f_{t_k} .$$

Тогда $f \in B$ и $f(u) > h(u) - \varepsilon$ для всех точек $u \in X$, ибо $f_{t_j}(u) > h(u) - \varepsilon$ для $u \in X$. Кроме того, для любой точки $u \in X$ и, в частности, для $u \in V(t_i)$ справедливы неравенства $f(u) \le f_{t_j}(u) < h(u) + \varepsilon$.

Таким образом, мы показали, что $|f(u) - h(u)| < \varepsilon$ на множестве *X*, откуда и следует утверждение теоремы.

Кроме аппроксимации функций многочленами, в последнее время все большее внимание уделяется приближению функций многих переменных с помощью линейных операций и суперпозиций функций одного переменного. В данном параграфе мы покажем, что эти функции одного переменного могут быть произвольными и докажем обобщенную теорему Стоуна [1,4], естественным образом охватывающую и классическую теорему Стоуна, и аппроксимацию функций многих переменных суперпозициями и линейными комбинациями функций одного переменного.

Теорема 3.2 (Обобщенная теорема Стоуна). Пусть X - компактное пространство, $C(X)$ - пространство непрерывных функций на X с вещественными значениями, $C(R)$ - пространство непрерывных функций на действительной оси, $E \subseteq C(X)$ - замкнутое линейное подпространство в $C(X)$, $1 \in E$, функции из E разделяют точки в X и E замкнуто относительно нелинейной унарной операции $f \in C(R)$. Тогда $E = C(X)$.

Доказательство. Рассмотрим множество всех таких $p \in C(R)$, что $p(E) \subseteq E$, то есть для любого $g \in E$ выполнено: $p(g) \in E$. Обозначим это множество P_E. Оно обладает следующими свойствами:

1) P_E - полугруппа относительно суперпозиции функций;
2) P_E - замкнутое линейное подпространство в $C(R)$ (в топологии равномерной сходимости на компактах);
3) $1 \in E$ и $id \in E$ $\quad (id(x) \equiv x)$.
4) P_E включает хоть одну непрерывную нелинейную функцию.

Пусть множество $P \subseteq C(R)$ удовлетворяет условиям 1-4. Докажем, что тогда $P = C(R)$.

Доказательство опирается на три леммы.

Лемма 3.1. Для множества P, удовлетворяющего условиям 1-4, существует дважды непрерывно дифференцируемая функция $g \in P$, не являющаяся линейной.

Доказательство. Пусть $v(x) \in C^{\infty}(R)$, $v(x) = 0$ при $|x| > 1, \int_R v(x)dx = 1$.

Рассмотрим оператор осреднения

$$J_{\varepsilon} f(x) = \int_R f(x+y) \frac{1}{\varepsilon} v\left(\frac{y}{\varepsilon}\right) dy, \qquad (3.1)$$

где $f(x)$ - функция из условий теоремы 3.2.

Покажем, что для любого ε>0 выполнено: $J_{\varepsilon} f(x) \in P$.

Действительно, $f(x+y) \in P$ для каждого фиксированного y (т.к. константы принадлежат P и P замкнуто относительно линейных операций и суперпозиции функций). Интеграл $J_{\varepsilon} f(x)$ принадлежит P, так как P является замкнутым линейным подпространством в $C(X)$, а этот интеграл - пределом конечных сумм.

Функция $J_{\varepsilon} f(x)$ принадлежит $C^{\infty}(R)$ так как

$$J_{\varepsilon} f(x) = \int_R f(x+y) \frac{1}{\varepsilon} v\left(\frac{y}{\varepsilon}\right) dy = \int_R f(z) \frac{1}{\varepsilon} v\left(\frac{z-x}{\varepsilon}\right) dz. \qquad (3.2)$$

(напомним, что v – функция с компактным носителем).

Покажем, что $J_{\varepsilon} f \to f$ (в топологии равномерной сходимости на компактах):

$$J_{\varepsilon} f(x) = \int_R f(x+y) \frac{1}{\varepsilon} v\left(\frac{y}{\varepsilon}\right) dy$$

сделаем замену переменных: $\frac{y}{\varepsilon} = w$, тогда интеграл примет вид

$$J_{\varepsilon} f(x) = \int_R f(w\varepsilon + x) \frac{1}{\varepsilon} d(w\varepsilon + x) = \int_R f(w\varepsilon + x) v(w) dw \qquad (3.3)$$

$$\lim_{\varepsilon \to 0} \int_R f(w\varepsilon + x) v(w) dw = f(x) \int_R v(w) dw = f(x).$$

Таким образом $J_{\varepsilon} f \to f$, при $\varepsilon \to 0$, следовательно, существует такое ε, что функция $g = J_{\varepsilon} f$ не является линейной, поскольку пространство линейных функций замкнуто, а f не является линейной функцией. Таким образом, в предположениях леммы существует нелинейная функция $g \in P \bigcap C^{\infty}(R)$, которую можно выбрать в виде $g = J_{\varepsilon} f$. Лемма доказана.

Лемма 3.2. Пусть существует дважды непрерывно дифференцируемая функция $g \in P$, не являющаяся линейной. Тогда функция $q(x) = x^2$ принадлежит P.

Доказательство. Существует точка x_0, для которой $g''(x_0) \neq 0$. Обозначим $r(x) = 2\dfrac{g(x + x_0) - g(x_0) - xg'(x_0)}{g''(x_0)}$.

Очевидно, что $r \in P$, $r(0) = 0$, $r'(0) = 0$, $r''(0) = 2$, $r(x) = x^2 + o(x^2)$. Поэтому

$$\lim_{\varepsilon \to 0} \frac{r(\varepsilon x)}{\varepsilon^2} = x^2. \qquad (3.4)$$

Поскольку P замкнуто, получаем: функция $q(x) = x^2$ принадлежит P. Лемма доказана.

Лемма 3.3. Пусть в условиях леммы 3.2 функция $q(x) = x^2$ принадлежит P. Тогда P является кольцом – для любых $f, g \in P$ их произведение $fg \in P$.

Доказательство. Действительно, $fg = \frac{1}{2}\left[(f + g)^2 - f^2 - g^2\right]$ и, так как P замкнуто относительно суперпозиции и линейных операций, то $fg \in P$. Лемма доказана.

Доказательство теоремы 3.2 заканчивается обращением к классической теореме Вейерштрасса о приближении функций многочленами: из лемм 3.1-3.3 следует, что в условиях теоремы 3.2 P является кольцом и, в частности, содержит все многочлены (которые получаются из 1 и id с помощью умножения и линейных операций). По теореме Вейерштрасса отсюда следует, что $P = C(X)$. Теорема доказана.

Следствие. Следствием этой теоремы является то, что с помощью операции суперпозиции и линейных комбинаций из любых нелинейных элементов можно получить любую функцию многих переменных с любой заданной точностью.

Литература

1. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. «Наука», Новосибирск, 1996.
2. К. Иосида «Функциональный анализ», «Мир», М., 1967, с. 17.
3. Stone M.N. The generalized Weierstrass approximation theorem. Math. Mag., 1948. V.21. PP. 167-183, 237-254. Cybenko G. Approximation by superposition of a sigmoidal function. Mathematics of Control, Signals, and Systems, 1989. Vol. 2. PP. 303 - 314.