В. Н. Буров, Некоторые эффективные способы решения задачи П. Л. Чебышева о наилучшем приближении, *Изв. вузов. Матем.*, 1957, номер 1, 67–79

ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ
1957 МАТЕМАТИКА № 1

В. Н. Буров

НЕКОТОРЫЕ ЭФФЕКТИВНЫЕ СПОСОБЫ РЕШЕНИЯ ЗАДАЧИ П. Л. ЧЕБЫШЕВА О НАИЛУЧШЕМ ПРИБЛИЖЕНИИ

§ 1. Постановка задачи. Основные теоремы

Основное содержание настоящей работы составляют методы фактического решения задачи П. Л. Чебышева о наилучшем равномерном приближении в весьма общей постановке.

Пусть D — точечное множество в пространстве $2n+2$ измерений R_{2n+2}, удовлетворяющее условиям: 1) если точка $P\equiv(x_0, x_1, \dots, x_n, y_0, y_1, \dots, y_n)$ принадлежит D, то необходимо $a\leqslant x_0<x_1<\dots<x_n\leqslant b$, где $a<b$ — заданные вещественные числа; 2) при любых фиксированных значениях $a\leqslant \overset{*}{x}_0<\overset{*}{x}_1<\dots<\overset{*}{x}_n\leqslant b$ все точки D вида $(\overset{*}{x}_0, \overset{*}{x}_1, \dots, \overset{*}{x}_n, y_0, y_1, \dots, y_n)$ образуют в $n+1$-мерном подпространстве Y_{n+1} пространства R_{2n+2} *непустой открытый параллелепипед* D_{X^*} (не обязательно ограниченный).

Определение 1. Назовем множество K непрерывных на сегменте $[a, b]$ функций $F(x)$ аргумента x *интерполяционным классом* (относительно множества D), если 1) для любой функции $F(x)\in K$ и любых значений $a\leqslant x_0<x_1<\dots<x_n\leqslant b$ точка $(x_0, x_1, \dots, x_n, F(x_0), F(x_1), \dots F(x_n))$ принадлежит множеству D, и 2) какова бы ни была точка $P\equiv(x_0, x_1, \dots, x_n, y_0, y_1, \dots, y_n)\in D$, существует одна и только одна функция $F(x)\in K$, для которой

$$F(x_i)=y_i \quad (i=0, 1, \dots, n). \tag{1,1}$$

Как частные случаи, данное определение охватывает собой интерполяционные классы функций, рассматривавшиеся в работах Т. Моцкина [11], М. И. Морозова [2], Е. П. Новодворского и И. Ш. Пинскера [5].

Определение 2. Условимся говорить, что функция $f(x)$ аргумента x принадлежит классу $C_D([a, b])$, если она определена и непрерывна на всем сегменте $[a, b]$, и при любых значениях $a\leqslant x_0<x_1<\dots<x_n\leqslant b$ точка $(x_0, x_1, \dots, x_n, f(x_0), f(x_1), \dots, f(x_n))$ принадлежит множеству D.

Пусть $f(x)\in C_D([a, b])$, и пусть S — какое-либо *замкнутое* множество точек сегмента $[a, b]$, содержащее не менее $n+2$ точек. Задача состоит в отыскании функции $F^*(x)$ класса K, *наименее уклоняющейся* от $f(x)$ на S, т. е. такой функции $F^*(x)\in K$, для которой

$$\max_{x\in S}|f(x)-F^*_{(x)}|=\inf_{F(x)\in K}\left\{\max_{x\in S}|f(x)-F(x)|\right\}\equiv E[f; S]. \tag{2,1}$$

Существование и единственность такой функции во всяком интерполяционном классе легко устанавливаются с помощью рассуждений работы М. И. Морозова [2]. Именно, можно доказать следующие фундаментальные теоремы, обобщающие хорошо известные в теории полиномиальных приближений теоремы Э. Бореля [9], Ш.-Ж. Валле Пуссена [12, 13] и П. Л. Чебышева [8].

Теорема 1. *Для каждой функции* $f(x) \in C_D([a, b])$ *существует одна и только одна функция* $F^*(x) \in K$, *наименее уклоняющаяся от* $f(x)$ *на замкнутом множестве* $S \subset [a, b]$, *содержащем не менее* $n+2$ *различных точек.*

Теорема 2. *Наилучшее приближение функции* $f(x) \in C_D([a, b])$ *посредством функции интерполяционного класса* K *на замкнутом множестве* $S \subset [a, b]$, *содержащем более чем* $n+2$ *точки, равно ее наилучшему приближению на множестве* X_{n+2}, *состоящем из* $n+2$ *точек* S, *выбранных таким образом, чтобы это наилучшее приближение было наибольшим.*

Теорема 3. *Пусть* $\overline{F}(x)$ — *произвольная функция класса* K. *Если разность* $f(x) - \overline{F}(x)$ *приобретает в* $n+2$ *последовательных точках* $x_1 < x_2 < \ldots < x_{n+2}$ *множества* S *чередующиеся по знаку значения* $\overline{r}_1, \overline{r}_2, \ldots, \overline{r}_{n+2}$ и $r = \min(|\overline{r}_1|, |\overline{r}_2|, \ldots, |\overline{r}_{n+2}|)$, *то*

$$E[f; S] \geqslant r, \tag{3,1}$$

причем знак равенства имеет место лишь при условии, что все

$$|\overline{r}_i| = r = \max_{x \in S} |f(x) - \overline{F}(x)|.$$

Теорема 4. *Для того, чтобы функция* $F^*(x) \in K$ *наименее уклонялась от функции* $f(x) \in C_D([a, b])$ *на замкнутом множестве* $S \subset [a, b]$, *содержашем не менее* $n+2$ *различных точек, необходимо и достаточно, чтобы абсолютный максимум функции* $|f(x) - F^*(x)|$ *достигался не менее, чем в* $n+2$ *последовательных точках* $x_1 < x_2 < \ldots < x_{n+2}$ *множества* S, *в которых знаки разности* $f(x) - F^*(x)$ *последовательно противоположны.*

Таким образом, как видно из теорем 2 и 4, функция $F^*(x)$, наименее уклоняющаяся от $f(x)$ на S, оказывается *равноосциллирующей* относительно $f(x)$ на некотором $n+2$-точечном множестве $X_{n+2} \subset S$. Это множество X_{n+2} мы, следуя И. П. Натансону [3], будем называть *$n+2$-членным чебышевским альтернансом*. Нетрудно показать, что величина наилучшего приближения $E[f; S]$ непрерывным образом зависит от точек альтернанса (при переменном множестве S).

Определение 3. Интерполяционный класс K назовем *ограниченно компактным*, если из всякого бесконечного ограниченного семейства функций данного класса можно выделить последовательность, равномерно сходящуюся к некоторой функции этого же класса.

Примером интерполяционного класса, не являющегося ограниченно компактным, служит двупараметрическое семейство функций $F(x; t_0, t_1) = t_0 e^{t_1 x}$ при $0 \leqslant x \leqslant 1$;

$$0 < t_0 < +\infty, \quad -\infty < t_1 < +\infty.$$

На случай ограниченно компактных интерполяционных классов удается обобщить важную для теории равномерного приближения теорему Е. Я. Ремеза (см. [6], стр. 93).

Теорема 5. *Пусть интерполяционный класс* K *является ограниченно компактным, и пусть* $\overline{F}(x)$ — *произвольная функцик этого класса. Тогда для любого числа* $\varepsilon > 0$ *можно определить число* $\eta > 0$ *таким образом, чтобы неравенство*

$$\max_{a \leqslant x \leqslant b} |f(x) - \overline{F}(x)| < E[f; [a, b]] + \eta \tag{4,1}$$

необходимо влекло за собой неравенство

$$\max_{a \leqslant x \leqslant b} |\overline{F}(x) - F^*(x)| < \varepsilon, \tag{5,1}$$

где $F^*(x)$ — функция, наименее уклоняющаяся от $f(x)$ на всем сегменте $[a, b]$.

Эта теорема легко доказывается от противного. Из нее вытекает

Следствие. Если последовательность функций $\{F_k(x)\}$ ограниченно компактного интерполяционного класса K такова, что для нее

$$\lim_{k \to \infty} L_k[f] = E[f;\ [a, b]], \tag{6,1}$$

где

$$L_k[f] = \max_{a \leqslant x \leqslant b} |f(x) - F_k(x)|,$$

то эта последовательность равномерно сходится к функции $F^*(x)$, наименее уклоняющейся от $f(x)$ на всем сегменте $[a, b]$.

Определение 4. Интерполяционный класс K назовем *линейным*, если с каждыми двумя его функциями $F_1(x)$ и $F_2(x)$ ему принадлежит также их любая линейная комбинация $c_1F_1(x) + c_2F_2(x)$ при вещественных c_1, c_2.

Аналогичным образом можно рассматривать и просто *однородный* интерполяционный класс K, не требуя его *аддитивности*. Нетрудно видеть, что функция однородного интерполяционного класса определяется своими n нулями с точностью до постоянного множителя.

Пусть $\{F(x)\}$ — некоторое семейство функций *однородного ограниченно компактного* интерполяционного класса K, каждая из которых имеет на сегменте $[a, b]$ n различных нулей, не будучи равной нулю тождественно, причем расстояние между двумя различными нулями каждой из этих функций остается не меньше некоторого фиксированного числа $\sigma > 0$. Обозначив через $x_1 < x_2 < \ldots < x_n$ нули функции $F(x)$, положим

$$X_F^\sigma = \mathop{\mathrm{E}}_{x \in [a, b]} (|x - x_i| \geqslant \sigma;\quad i = 1, 2, \ldots, n). \tag{7,1}$$

При $0 < \sigma < \dfrac{b-a}{2n}$ будет $X_F^\sigma \neq \Lambda$.

Положим, далее,

$$\varphi_\sigma(F) = \min_{x \in X_F^\sigma} |F(x)|. \tag{8,1}$$

Тогда имеет место

Теорема 6. *Существует единая для всего семейства функций $\{F(x)\}$ постоянная μ_σ такая, что*

$$\frac{\varphi_\sigma(F)}{\max\limits_{a \leqslant x \leqslant b} |F(x)|} \geqslant \mu_\sigma > 0. \tag{9,1}$$

Доказательство. Допустим противное. Тогда из данного семейства можно выделить последовательность функций $\{F_k(x)\}$ такую, что для нее

$$\lim_{k \to \infty} \frac{\varphi_\sigma(F_k)}{\max\limits_{a \leqslant x \leqslant b} |F_k(x)|} = 0. \tag{10,1}$$

Пронормируем каждую из функций последовательности, полагая, что

$$\overline{F}_k(x) - \frac{F_k(x)}{\max\limits_{a \leqslant x \leqslant b} |F_k(x)|}. \tag{11,1}$$

Тогда получим

$$\lim_{k \to \infty} \varphi_\sigma(\overline{F}_k) = 0. \tag{12,1}$$

Поскольку $|\overline{F}_k(x)| \leqslant 1$, а класс K предположен ограниченно компактным, мы вправе считать последовательность функций $\{\overline{F}_k(x)\}$ равномерно сходящейся к некоторой функции $\overline{F}(x) \in K$.

Очевидно,

$$\max_{a \leqslant x \leqslant b} |\overline{F}(x)| = 1. \tag{13,1}$$

Пусть $x_1^{(k)} < x_2^{(k)} < \ldots < x_n^{(k)}$ — нули функции $\overline{F}_k(x)$ (равно как и $F_k(x)$). Последовательность индексов k может быть выбрана так, чтобы

$$\lim_{k \to \infty} x_i^{(k)} = \overline{x}_i \quad (i = 1, 2, \ldots, n). \tag{14,1}$$

При этом все предельные точки $\overline{x}_i$ будут различными, ибо и для них $|\overline{x}_{i_1} - \overline{x}_{i_2}| \geqslant \sigma$ при $i_1 \neq i_2$. Легко убедиться, что предельная функция $\overline{F}(x)$ будет иметь точки $\overline{x}_j$ своими нулями.

Рассмотрим множество:

$$\overline{X} = \mathop{\mathrm{E}}_{x \in [a,b]} \left(|x - \overline{x}_i| \geqslant \frac{\sigma}{2};\ i = 1, 2, \ldots, n \right).$$

Ввиду (14,1) при достаточно больших k будет иметь место включение

$$X_{\overline{F}_k}^{\sigma} \subset \overline{X}$$

и, значит, неравенство

$$0 \leqslant \min_{x \in \overline{X}} |\overline{F}_k(x)| \leqslant \varphi_\sigma(\overline{F}_k).$$

Отсюда, в силу (12,1), вытекает:

$$\min_{x \in \overline{X}} |\overline{F}(x) = 0.$$

Следовательно, предельная функция $\overline{F}(x)$ имеет, помимо n нулей $\overline{x}_i$, по крайней мере, еще один нуль на множестве $\overline{X}$. А тогда необходимо $\overline{F}(x) \equiv 0$, что однако невозможно ввиду (13,1). Полученное противоречие и доказывает теорему.

Несложной модификацией рассуждений П. Лаасонена [10] может быть доказана также следующая, весьма полезная для практических построений.

Теорема 7. ***Всякий линейный ограниченно компактный интерполяционный класс K содержит функции, имеющие на всем сегменте $[a, b]$ только m наперед заданных нулей***

$$a < x_1 < x_2 < \ldots < x_m < b \quad (0 \leqslant m \leqslant n) \tag{15,1}$$

и изменяющие знак при переходе через каждый из этих нулей.

§ 2. Однопараметрическая чебышевская задача

В основу вычислительных схем предлагаемых ниже методов положено решение следующей вспомогательной *однопараметрической* задачи: для ограниченного и замкнутого плоского точечного множества Q требуется определить значение вещественного параметра $\lambda = \lambda^*$ из условия

$$\max_{(u,v) \in Q} |v - \lambda u| \equiv L(\lambda) = \min. \tag{1,2}$$

Простое графо-аналитическое решение этой задачи предложено Е. Я. Ремезом [7]. Именно, искомое решение $\lambda = \lambda^*$ дается угловым коэффициентом любой опорной прямой t множества $Q_0 = Q \cup (-Q)$, пересекающей положительное направление оси O_v и удовлетворяющей по крайней мере одному из двух условий: либо прямая t проходит через две точки $P_1 \equiv (u_1, v_1)$, $P_2 \equiv (u_2, v_2)$ множества Q_0, расположенные по обе стороны от оси ординат (и тогда решение заведомо однозначно), либо прямая t проходит через некоторую точку $(0, \overline{v}) \in Q_0$, лежащую на самой оси Ov.

Нетрудно видеть, что в первом случае решение λ^* может быть получено как тангенс такого угла поворота α^* $\left(0 \leqslant |\alpha^*| < \frac{\pi}{2}\right)$ координатных осей Ouv вокруг начала O, при котором ординаты упомянутых точек P_1, P_2 в новой системе координат $Ou'v'$ окажутся равными между собой и максимальными для всего множества Q_0 (аналог основной теоремы П. Л. Чебышева).

Условимся считать, что начало координат O всегда принадлежит данному множеству Q. Достаточно ограничиться случаем

$$l_0 \equiv \max_{\substack{u \leqslant 0 \\ (u,v) \in Q_0}} v < \max_{(u,v) \in Q_0} v \equiv L_0, \tag{2,2}$$

поскольку случай $\max\limits_{u \geqslant 0} v < L_0$ симметричен данному относительно u и λ, а случай $\max\limits_{u \leqslant 0} v = \max\limits_{u \geqslant 0} v = L_0$ тривиален ($\lambda^* = 0$). Если положить

$$u_{\max} \equiv \max_{(u,v) \in Q_0} u = \max_{(u,v) \in Q} |u|,$$

то для решения λ^* сразу получаем положительную оценку снизу

$$\lambda^* \geqslant \frac{L_0 - l_0}{2u_{\max}} \equiv \lambda_0 > 0. \tag{3,2}$$

В самом деле, неотрицательность λ^* при условии (2,2) очевидна. Возьмем, далее, произвольное значение $\lambda \in (0, \lambda_0)$ и выполним поворот координатных осей на угол $\alpha = \operatorname{arctg} \lambda$. Как известно, ординаты всех точек Q_0 преобразуются при этом по формуле

$$v' = -u \sin \alpha + v \cos \alpha = \cos \alpha \, [v - \lambda u].$$

Отсюда видно, что при выбранном λ будет выполняться соотношение

$$\max_{\substack{u \leqslant 0 \\ (u,v) \in Q_0}} v' < \max_{(u', x') \in Q_0} v' \equiv L'(\lambda), \tag{4,2}$$

а тогда достаточно малый дополнительный поворот координатных осей на положительный угол $\Delta \alpha$ позволит нам еще более понизить величину отклонения $L(\lambda)$. Стало быть, ни одно из значений $\lambda < \lambda_0$ не может решать задачу (1,2), откуда и следует (3,2).

В частности, при любом $\lambda \in (0, \lambda_0)$ оказывается

$$u_\lambda \equiv \min_{v' = L'(\lambda)} u > 0, \tag{5,2}$$

что вытекает из неравенства (4,2) и замкнутости множества Q_0. Это позволяет оценить сверху величину L^* наилучшего приближения задачи (1,2)

$$L^* \leqslant L_0 - \sup_{0 < \lambda \leqslant \lambda_0} \{\lambda u_\lambda\}. \tag{6,2}$$

Однако на практике удобнее пользоваться более грубой оценкой

$$L^* \leqslant L_0 - \bar{\lambda} u_{\bar{\lambda}}, \tag{7,2}$$

где $\bar{\lambda}$ — произвольное значение из $(0, \lambda_0)$.

В заключение параграфа приведем без доказательства некоторые дополнительные сведения о решении однопараметрической задачи (1,2).

Для *единственности* решения задачи (1,2) необходимо и достаточно выполнение условия

$$\lambda^- \equiv \sup_{\substack{u \neq 0 \\ (u, v) \in Q}} \left\{\frac{v \operatorname{sign} u - \tilde{v}}{|u|}\right\} \geqslant \inf_{\substack{u \neq 0 \\ (u, v) \in Q}} \left\{\frac{v \operatorname{sign} u + \tilde{v}}{|u|}\right\} \equiv \lambda^+, \tag{8,2}$$

где

$$\tilde{v} \equiv \max_{\substack{u = 0 \\ (u, v) \in Q}} |v|.$$

Если условие (8,2) нарушено, то значения

$$\lambda^- \leqslant \lambda \leqslant \lambda^+ \tag{9,2}$$

образуют совокупность всех решений задачи (1,2) для данного случая.

Наконец, одним из решений задачи (1,2) всегда оказывается значение

$$\lambda^* = \inf_{\substack{u_1 < 0 \\ (u_1, v_1) \in Q_0}} \left\{ \max_{\substack{u_2 \geqslant 0 \\ (u_2, v_2) \in Q_0}} \frac{v_2 - v_1}{u_2 - u_1} \right\}. \tag{10,2}$$

§ 3. Метод возрастающих отклонений Его монотонизация для случая линейных классов

В связи с проблемой приближенного определения функции наилучшего приближения большую практическую важность приобретает тот частный случай, когда множество $S = S_m$ состоит из конечного числа m последовательных точек сегмента $[a, b[$

$$a \leqslant x_1 < x_2 < \ldots < x_m \leqslant b \quad (m > n + 2). \tag{1,3}$$

Для нахождения функции $F^*(x) \in K$, наименее уклоняющейся от функции $f(x) \in C_D([a, b])$, может быть применен метод *возрастающих отклонений*, являющийся конечным аналогом второго алгорифма Е. Я. Ремеза [6]. Он состоит в следующем. Выберем исходное множество $X^0_{n+2} = \{x^0_1 < x^0_2 < \ldots < x^0_{n+2}\}$, состоящее из $n + 2$ различных точек S_m, и построим функцию $F_0(x) \in K$, наименее уклоняющуюся от $f(x)$ на этом множестве. Ее можно определить как равноосциллирующую относительно функции $f(x)$ на множестве X^0_{n+2}, т. е. из условий

$$(-1)^{i-1} \varepsilon E[f;\ X^0_{n+2}] + F_0(x^0_i) = f(x^0_i) \tag{2,3}$$
$$(i = 1, 2, \ldots, n + 2;\ \varepsilon = \pm 1).$$

Для очередного шага образуем множество X^1_{n+2} из множества X^0_{n+2} путем исключения некоторых точек $x^0_{i_1}$, $x^0_{i_2}, \dots, x^0_{i_\nu}$ $(\nu \leqslant n+2)$ множества X^0_{n+2} и замены их точками $x^1_{i_1}$, $x^1_{i_2}, \dots, x^1_{i_\nu}$ множества S_m таким образом, чтобы оказались выполненными два условия: 1) в точках множества X^1_{n+2} сохраняется чередование знаков разности $f(x) - F_0(x)$; 2) во вновь включенных точках

$$|f(x^1_{i_p}) - F_0(x^1_{i_p})| \geqslant E[f;\ X^0_{n+2}] \quad (p = 1, 2, \dots, \nu), \tag{3,3}$$

причем, хотя бы в одной из них имеет место строгое неравенство. Такое построение всегда возможно, если только $F_0(x) \not\equiv F^*(x)$. Построим теперь функцию $F_1(x) \in K$, наименее уклоняющуюся от $f(x)$ на X^1_{n+2}. Тогда согласно теореме 3 будет $E[f;\ X^1_{n+2}] > E[f;\ X^0_{n+2}]$. Отсюда, в силу теоремы 2, вытекает, что повторив указанный процесс конечное число раз (очевидно, не превышающее C_m^{n+2}), мы необходимо придем к искомой функции $F^*(x)$, дающей решение нашей задачи на всем множестве S_m.

Изложенный метод имеет один существенный недостаток — отсутствие монотонности для последовательно получаемых величин

$$L_k \equiv \max_{x \in S_m} |f(x) - F_k(x)|,$$

вследствие чего последующее приближение может, вообще говоря, оказаться во много раз хуже предыдущего (в смысле абсолютной погрешности). Однако для случая линейных интерполяционных классов этот недостаток можно легко устранить, если использовать в вычислительной схеме решение вспомогательной однопараметрической задачи (1,2). Действительно, для *линейного* интерполяционного класса K чебышевская задача (2,1) оказывается инвариантной относительно любой аддитивной интерполяционной поправки $\bar{F}(x) \in K$ в том смысле, что добавление к приближаемой функции $f(x)$ произвольной функции $\bar{F}(x) \in K$ не меняет ни расположения точек альтернаса, ни самой величины наилучшего приближения $E[f;\ S]$. Поэтому оказывается возможным с помощью поправок такого рода перейти в конечное число шагов от функции $f(x)$ к функции $\Delta^*(x) = f(x) - F^*(x)$, наименее уклоняющейся от *нуля* на всем множестве S_m, причем этот переход может быть осуществлен *монотонным* образом.

Предлагаемый монотонный вариант метода возрастающих отклонений характеризуется на каждом шаге следующими двумя переходами: 1) переходом от функции $\Delta_{k-1}(x)$ к функции $\Delta_k(x)$ при условии

$$L_k \equiv \max_{x \in S_m} |\Delta_k(x)| < \max_{x \in S_m} |\Delta_{k-1}(x)| \equiv L_{k-1}, \tag{4,3}$$

причем $[\Delta_k(x) - \Delta_{k-1}(x)] \in K$, и 2) переходом от множества X^k_{n+2} к множеству X^{k+1}_{n+2} при условии

$$E[f;\ X^{k+1}_{n+2}] > E[f;\ X^k_{n+2}].$$

Определение 5. Будем называть $n+2$-точечное множество X_{n+2} *альтернирующим* относительно функции $\Delta(x)$, если последняя принимает в его точках ненулевые значения с последовательно противоположными знаками.

Не нарушая общности, можно считать, что существует по крайней мере одно $n+2$-точечное множество $X^0_{n+2} \subset S_m$, альтернирующее относительно приближаемой функции $f(x) \equiv \Delta_0(x) \in C_D([a,\ b])$.

Перейдем к изложению самого алгорифма. Условимся говорить, что на данном k-ом шаге мы имеем *нормальный* случай, если существует $n+2$-точечное множество $X^k_{n+2} \subset S_m$, альтернирующее относительно функции $\Delta_{k-1}(x)$ и содержащее *все* точки абсолютного максимума $|\Delta_{k-1}(x)|$ на S_m, и *особый* случай, если такого множества заведомо не существует. (Выделенный здесь особый случай аналогичен случаю стационарной точки в алгорифме С. И. Зуховицкого [1]).

Начнем с рассмотрения *нормального* случая. Построим функцию $F_k(x)$, наименее уклоняющуюся от $\Delta_{k-1}(x)$ на X^k_{n+2}, и решим однопараметрическую задачу (1,2) для множества Q_k пар чисел $(F_k(x), \Delta_{k-1}(x))$, получаемого при $x \in S_m$. Пусть ее решение дается значением $\lambda = \lambda_k$. (Если оно не единственно, то в качестве λ_k будем брать *наименьшее* из всех решений, т. е. λ^- в обозначениях (8,2), (9,2)). В соответствии с выбором множества X^k_{n+2} соотношение (2,2) будет выполнено, а тогда необходимо $\lambda_k > 0$.

Если окажется $0 < \lambda_k \leqslant 1$, то положим $\Delta_k(x) = \Delta_{k-1}(x) - \lambda_k F_k(x)$. Нетрудно показать, что в этом случае возможно выделить для следующего шага множество $X^{k+1}_{n+2} \subset S_m$, альтернирующее относительно функции $\Delta_k(x)$, и такое, что $E[\Delta_{k-1}; X^{k+1}_{n+2}] > E[\Delta_{k-1}; X^k_{n+2}]$. С этой целью достаточно лишь включить в X^{k+1}_{n+2} хотя бы одну из точек абсолютного максимума $|\Delta_k(x)|$ на S_m, не входящую в X^k_{n+2}.

Если же $\lambda_k > 1$, то обычный шаг по методу возрастающих отклонений сам по себе уже является монотонным.

Однако можно еще более понизить величину максимального отклонения

$$L_k(\lambda) \equiv \max_{x \in S_m} |\Delta_{k-1}(x) - \lambda F_k(x)|,$$

сохраняя при этом возможность выбора $n+2$-точечного множества $X^{k+1}_{n+2} \subset S_m$, альтернирующего относительно получаемой функции $\Delta_k(x)$. Для этого нужно положить $\Delta_k(x) = \Delta_{k-1}(x) - \bar{\lambda}_k F_k(x)$, взяв $\bar{\lambda}_k > 1$ достаточно близким к $\min(\lambda_k, \gamma_k)$, где γ_k — точная верхняя грань всех тех значений λ, при которых множество X^k_{n+2} остается альтернирующим относительно разности $\tau_k(x; \lambda) \equiv \Delta_{k-1}(x) - \lambda F_k(x)$. Обозначим через X^k_0 множество всех точек абсолютного максимума $|\Delta_k(x)|$ на S_m. Тогда для всякого $n+2$-точечного множества

$$X_{n+2} \subset X^k_{n+2} \cup X^k_0 \left(X_{n+2} \cap X^k_0 \neq \Delta\right),$$

альтернирующего относительно функции $\Delta_k(x)$, непременно окажется

$$E[\Delta_{k-1}; X_{n+2}] > E[\Delta_{k-1}; X^k_{n+2}].$$

Этим обстоятельством следует воспользоваться для выбора очередного множества X^{k+1}_{n+2}.

Прежде чем перейти к рассмотрению особого случая, покажем, что на любом k-ом шаге всегда можно выделить $n+2$-точечное множество $X^k_{n+2} \subset S_m$ такое, что во всех точках абсолютного максимума $|\Delta_{k-1}(x)|$ на S_m будет выполняться соотношение

$$\operatorname{sign} F_k(x) = \operatorname{sign} \Delta_{k-1}(x), \tag{5,3}$$

где $F_k(x)$ — функция, наименее уклоняющаяся от $\Delta_{k-1}(x)$ на X^k_{n+2}.

С этой целью рассмотрим всевозможные $n+2$-точечные множества $X_{n+2} \subset S_m$, альтернирующие относительно $\Delta_{k-1}(x)$, и выберем из них то множество X^k_{n+2}, для которого величина $E[\Delta_{k-1}; X_{n+2}]$ имеет наибольшее значение. (Если таких множеств несколько, то возьмем любое из них.) Утверждается, что оно и будет искомым. Действительно, если допустить противное, то в некоторой точке x' максимума $|\Delta_{k-1}(x)|$ на S_m заведомо окажется

$$|\Delta_{k-1}(x') - F_k(x')| > E[\Delta_{k-1};\ X_{n+2}^k], \tag{6,3}$$

причем

$$\operatorname{sign}[\Delta_{k-1}(x') - F_k(x')] = \operatorname{sign}\Delta_{k-1}(x').$$

Образуем из X_{n+2}^k множество $\overline{X}_{n+2}$, заменяя точкой x' одну из точек X_{n+2}^k таким образом, чтобы в точках $\overline{X}_{n+2}$ сохранилось чередование знаков разности $\Delta_{k-1}(x) - F_k(x)$. Легко видеть, что множество $\overline{X}_{n+2}$ также будет альтернирующим относительно функции $\Delta_{k-1}(x)$. Однако для него, в силу (6,3) и теоремы 3, должно быть

$$E[\Delta_{k-1};\ \overline{X}_{n+2}] > E[\Delta_{k-1};\ X_{n+2}^k],$$

что противоречит выбору множества X_{n+2}^k.

Проведенное рассуждение показывает, что особый случай не имеет принципиального отличия от нормального случая, поскольку в последнем мы требовали включения в X_{n+2}^k *всех* точек абсолютного максимума $|\Delta_{k-1}(x)|$ на S_m только ради выполнения соотношений (5,3), совпадающих с (2,2) и обеспечивающих положительность соответствующего λ_k. Таким образом, и к *особому* случаю применим тот же самый алгорифм.

Однако практическое нахождение указанного множества X_{n+2}^k может оказаться весьма затруднительным. Поэтому иногда бывает целесообразнее применить в особом случае искусственный прием понижения общего числа экстремальных точек функции $\Delta_{k-1}(x)$, который состоит в следующем. Строится некоторая функция $\overline{F}(x) \in K$, удовлетворяющая условию вида (5,3). Вычитая из функции $\Delta_{k-1}(x)$ интерполяционную поправку вида $\overline{\lambda}\,\overline{F}(x)$ при достаточно малом положительном значении $\overline{\lambda}$, можно понизить общее число экстремальных точек этой функции по меньшей мере на единицу, причем множество X_{n+2}^{k-1}, альтернирующее относительно $\Delta_{k-1}(x)$, будет по-прежнему альтернировать относительно результирующей функции

$$\overline{\Delta}_{k-1}(x) = \Delta_{k-1}(x) - \overline{\lambda}\,\overline{F}(x).$$

При этом окажется

$$\max_{x \in S_m} |\overline{\Delta}_{k-1}(x)| < \max_{x \in S_m} |\Delta_{k-1}(x)|,$$

так что общий характер монотонности алгорифма не будет нарушен. Повторив указанный прием конечное число раз, мы необходимо придем к нормальному случаю.

§ 4. О сходимости методов сеточного типа

Изложенный в предыдущем параграфе метод возрастающих отклонений может также рассматриваться как сеточный (интерполяционный) метод приближенного решения общей чебышевской задачи (2,1) на всем сегменте $[a, b]$ или на бесконечном замкнутом множестве $S \subset [a, b]$. В этом случае за приближенное решение задачи (2,1) принимается точное решение $F_\varepsilon(x)$ аналогичной задачи на конечной ε-сети S_ε исходного множества S, причем $\varepsilon > 0$ предполагается достаточно малым.

При таком подходе к решению задачи (2,1) возникает вопрос о сходимости всего процесса при $\varepsilon \to 0$. Этому вопросу мы и посвятим данный параграф.

Пусть имеется произвольная последовательность положительных чисел $\{\varepsilon_m\}$, монотонно сходящаяся к нулю при $m \to \infty$. Соответствующую ей последовательность конечных ε_m-сетей S_m всего множества S предположим построенной таким образом, что

$$S_1 \subset S_2 \subset \ldots \subset S_m \subset \ldots \subset S. \tag{1,4}$$

Тогда легко доказывается

Теорема 8. *При условии* (1,4)

$$\lim_{n \to \infty} E[f;\ S_m] = E[f;\ S] \tag{2,4}$$

для любой функции $f(x) \in C_D([a,\ b])$.

Таким образом, для всех интерполяционных классов имеет место „слабая“ сходимость вида (2,4). „Сильная“ же сходимость методов сеточного типа может быть доказана в предположении ограниченной компактности интерполяционного класса K. Именно, справедлива

Теорема 9. ***Пусть интерполяционный класс K является ограниченно компактным, и пусть $f(x) \in C_D([a,\ b])$. Обозначим через $F_m(x)$ и $F^*(x)$ функции, наименее уклоняющиеся от $f(x)$ соответственно на множествах S_m и S. Тогда при выполнении условия*** (1,4)

$$\lim_{m \to \infty} F_m(x) = F^*(x) \tag{3,4}$$

равномерно относительно x на всем сегменте $[a,\ b]$.

Доказательство этой теоремы, опирающееся на целый ряд вспомогательных предложений, мы приводить не будем.

Установим теперь некоторые оценки, характеризующие близость отклонений в зависимости от ε. Будем считать $S = [a,\ b]$. Обозначим через $\omega(\delta)$ и $\Omega_m(\delta)$ соответственно модули непрерывности функций $f(x)$ и $F_m(x)$. Положим, далее,

$$L_m[f] = \max_{a \leqslant x \leqslant b} |f(x) - F_m(x)|.$$

Тогда, приближая кривые $y = f(x)$ и $y = F_m(x)$ ломаными с концами звеньев в точках S_m и учитывая, что функция $F_m(x)$ наименее уклоняется от $f(x)$ на S_m, нетрудно установить следующее неравенство:

$$0 \leqslant L_m[f] - E[f;\ S_m] \leqslant \omega(\varepsilon_m) + \Omega_m(\varepsilon_m). \tag{4,4}$$

Если предположить, что интерполяционный класс K совпадает с классом H_n всех алгебраических многочленов степени не выше n, а приближаемая функция $f(x)$ имеет на $[a,\ b]$ ограниченную вторую производную $|f''(x)| \leqslant M_2$, то аналогичным путем получим следующую оценку:

$$L_m[f] - E[f;\ S_m] = 0\left(\varepsilon_m^2\right). \tag{5,4}$$

Последняя оценка, вообще говоря, уже не допускает улучшения порядка за счет дополнительных предположений о структурно-дифференциальных свойствах приближаемой функции $f(x)$. В этом легко убедиться на примере функции $f(x) = x^2$ при $n = 0$.

Так как, очевидно, $E[f;\ S_m] \leqslant E[f;\ [a,\ b]]$, то оценки (4,4), (5,4) заведомо остаются справедливыми при замене в левой части $E[f;\ S_m]$ на $E[f;\ [a,\ b]]$.

§ 5. Метод убывающих максимумов

Излагаемый в настоящем параграфе метод представляет собой известную модификацию первого алгорифма Е. Я. Ремеза [6] и распространение его на случий *линейно ограниченно компактных* интерполяционных классов. Этот метод можно рекомендовать для приближенного решения чебышевской задачи (2,1) непосредственно на всем сегменте $[a, b]$. Он состоит в следующем. Исходим из любого „нулевого" приближения $F_0(x)$ функции $f(x)$ на $[a, b]$ и полагаем $\Delta_0(x) = f(x) - F_0(x)$. Фиксируем также некоторое положительное число A_0, удовлетворяющее условию $0 < A_0 \leqslant E[f;\ [a, b]]$, взяв его по возможности ближе к $E[f;\ [a, b]]$. (Считаем $E[f;\ [a, b]] > 0$, тогда выбор A_0 может быть осуществлен при помощи теоремы 3.)

Пусть функция $\Delta_k(x)$ уже построена. Чтобы перейти от нее к функции $\Delta_{k+1}(x)$, следует поступать так. Определяем число A_{k+1} из условия

$$E[f;\ [a, b]] \geqslant A_{k+1} \geqslant A_k. \tag{1,5}$$

Тогда A_{k+1} подавно будет удовлетворять неравенству

$$0 < A_{k+1} < \max_{a \leqslant x \leqslant b} |\Delta_k(x)|.$$

Положим,

$$\begin{aligned} \mathsf{E}^+_{k+1} &= \underset{x \in [a, b]}{\mathsf{E}} \ (\Delta_k(x) \geqslant A_{k+1}) \quad \text{и} \\ \mathsf{E}^-_{k+1} &= \underset{x \in [a, b]}{\mathsf{E}} \ (\Delta_k(x) \leqslant -A_{k+1}). \end{aligned} \tag{2,5}$$

Соединение $\mathsf{E}^+_{k+1} \cup \mathsf{E}^-_{k+1}$ будет непустым замкнутым точечным множеством, содержащимся в $[a, b]$. Известно [4], что такое множество может быть получено из сегмента $[a, b]$ путем удаления конечного или счетного числа взамно не налегающих интервалов (с выбрасыванием, если нужно, самих точек a и b). Обозначим через

$$I_1^{k+1}, \quad I_2^{k+1}, \quad I_3^{k+1}, \ldots \tag{3,5}$$

те из этих интервалов, которые одновременно граничат с обоими множествами E^+_{k+1} и E^-_{k+1}.

Мы предполагаем число A_{k+1} выбранным таким образом, чтобы число интервалов типа (3,5) не превышало n. В соответствии с теоремой 3 последнего всегда можно добиться, увеличивая в случае надобности число $A_{k+1} \leqslant E[f;\ [a, b]]$ до тех пор, пока требуемое условие не окажется выполненным, т. к. иначе наша задача была бы уже решена.

Итак, пусть для выбранного значения A_{k+1} мы получили интервалы:

$$I_1^{k+1}, \quad I_2^{k+2}, \ldots, I_\nu^{k+1} \quad (\nu \leqslant n). \tag{4,5}$$

Построим функцию $F_{k+1}(x)$ такую, чтобы на всем соединении $\mathsf{E}^+_{k+1} \cup \mathsf{E}^-_{k+1}$ выполнялось условие

$$\operatorname{sign} F_{k+1}(x) = \operatorname{sign} \Delta_k(x). \tag{5,5}$$

Этого мы добьемся, задавая все нули функции $F_{k+1}(x)$ внутри интервалов (4,5) и опираясь на теорему 7. При этом следует задавать n нулей, если $\nu \equiv n \pmod 2$ и $n-1$ нулей в противном случае, располагая все эти нули так, чтобы при переходе через каждый из интервалов (4,5) функция $F_{k+1}(x)$ изменяла знак в нужную нам сторону. Если следить на каждом шаге за тем, чтобы все расстояния задаваемых нулей друг от друга, равно как и от множества

$$\mathsf{E}^+_{k+1} \cup \mathsf{E}^-_{k+1},$$

были не меньше, чем $\frac{\delta_{k+1}}{n+1}$, где δ_{k+1} — наименьшая из длин интервалов (4,5), то, как нетрудно показать, эти расстояния при всех k будут оставаться выше некоторой фиксированной положительной границы $\sigma > 0$ (зависящей, вообще говоря, от выбора A_0).

Решим теперь однопараметрическую задачу (1,2) для множества Q_{k+1} пар чисел $(F_{k+1}(x), \Delta_k(x))$, где x пробегает весь сегмент $[a, b]$. Пусть это решение дается значением $\lambda = \lambda_{k+1}$. Тогда полагаем, что

$$\Delta_{k+1}(x) = \Delta_k(x) - \lambda_{k+1} F_{k+1}(x) \qquad (6,5)$$

и т. д.

Сходимость данного процесса устанавливает

Теорема 10. *При выполнении указанных выше условий относительно выбора нулей каждой из функции $F_k(x)$ $(k = 1, 2, 3, \ldots)$ оказывается, что*

$$\lim_{k \to \infty} \Delta_k(x) = f(x) - F^*(x) \qquad (7,5)$$

равномерно относительно x на всем сегменте $[a, b]$, где $F^*(x)$ — функция, наименее уклоняющаяся от $f(x)$ на $[a, b]$.

Доказательство. Ясно, что условие (5,5) перекрывает собой условие (2,2) для множества Q_{k+1} пар чисел

$$(F_{k+1}(x), \Delta_k(x)), \text{ где } x \in [a, b].$$

Будем предполагать, что на каждом шаге решение λ_{k+1} соответствующей однопараметрической задачи (1,2) вычисляется практически с такой степенью точности, которая позволяет пользоваться для величин

$$L_{k+1} \equiv \max_{a \leqslant x \leqslant b} |\Delta_{k+1}(x)|$$

оценкой вида (7,2) при

$$\overline{\lambda} = \frac{L_k - A_{k+1}}{2 \max\limits_{a \leqslant x \leqslant b} |F_{k+1}(x)|}. \qquad (8,5)$$

Тогда, как нетрудно видеть, при $\lambda = \overline{\lambda}$ все точки абсолютного максимума функции $|\Delta_k(x) - \overline{\lambda} F_{k+1}(x)|$ на $[a, b]$ принадлежат множеству $\mathsf{E}^+_{k+1} \cup \mathsf{E}^-_{k+1}$. Далее, в силу выбора нулей функции $F_{k+1}(x)$, должно иметь место включение

$$\mathsf{E}^+_{k+1} \cup \mathsf{E}^-_{\alpha+1} \subset X^\sigma_{F_{k+1}},$$

где принято обозначение (7,1). А тогда и подавно $\varphi_{\overline{\lambda}} \geqslant \varphi_\sigma(F_{k+1})$, где $\varphi_{\overline{\lambda}}$ соответствует значению $u_{\overline{\lambda}}$ из (5,2).

Применяя с учетом вышесказанного оценку (7,2) при $\lambda = \overline{\lambda}$ из (8,5) и полагая для краткости $E[f; [a, b]] = \rho$, получим

$$L_{k+1} \leqslant L_k - \frac{L_k - A_{k+1}}{2 \max\limits_{a \leqslant x \leqslant b} |F_{k+1}(x)|} \mid \varphi_{\overline{\lambda}} \leqslant L_k - (L_k - \rho) \frac{\varphi_\sigma(F_{k+1})}{2 \max\limits_{a \leqslant x \leqslant b} |F_{k+1}(x)|}. \qquad (9,5)$$

Поскольку выбор нулей каждой из функции $F_{k+1}(x)$ обеспечивает выполнение условий теоремы 6 для всей последовательности функций $\{F_k(x)\}$, неравенство (9,5) можно усилить;

$$L_{k+1} \leqslant L_k \left[1 - \left(1 - \frac{\rho}{L_k}\right) \frac{\mu_\sigma}{2}\right], \tag{10,5}$$

где, очевидно, $0 < \mu_\sigma \leqslant 1$.

Монотонно возрастающая последовательность $\frac{\rho}{L_k}$ ограничена сверху единицей. Поэтому существует

$$\lim_{k \to \infty} \frac{\rho}{L^k} = \vartheta,$$

причем $0 < \vartheta \leqslant 1$. Согласно следствию из теоремы 5 достаточно показать, что $\vartheta = 1$.

Допустим противное: $0 < \vartheta < 1$. В таком случае мы только снова усилим неравенство (10,5), если напишем

$$L_{k+1} \leqslant \alpha L_k, \tag{11,5}$$

где

$$0 < \alpha \equiv \left[1 - (1 - \vartheta) \frac{\mu_\sigma}{2}\right] < 1.$$

Из последнего неравенства непосредственно вытекает, что

$$\lim_{k \to \infty} L_k = 0,$$

поскольку тогда

$$0 < L_{k+1} \leqslant \alpha^{k+1} L_0. \tag{12,5}$$

Но, с другой стороны, при всех k, очевидно, будет $L_k \geqslant \rho > 0$. Мы получили противоречие.

Значит, на самом деле $\vartheta = 1$, т. е. $\lim\limits_{k \to \infty} L_k = \rho$, что и доказывает сходимость метода убывающих максимумов.

Ленинградский государственный университет
имени А. А. Жданова

Поступило
4 X 1957

ЛИТЕРАТУРА

1. С. И. Зуховицкий, О наилучшем в смысле П. Л. Чебышева приближении конечной системы несовместных уравнений, ДАН, LXXIX, 4, стр. 561—564, 1951. 2. М. И. Морозов, О некоторых вопросах равномерного приближения непрерывных функций посредством функций интерполяционных классов, ДАН, LXXVII, 3, стр. 381—383, 1951. 3. И. П. Натансон, Конструктивная теория функций, М.—Л., 1949. 4. И. П. Натансон, Теория функций вещественной переменной, М.—Л., 1950. 5. Е. П. Новодворский, И. Ш. Пинскер, Процесс уравнивания максимумов, УМН, VI, 6 (46), стр. 174—181, 1951. 6. Е. Ремез, Про методи найкращого, в розумінні Чебишова, наближеного представления функцій, Вид. АН УРСР, Київ, 1935. 7. Е. Я. Ремез, Метод графо-аналитического решения некоторых задач чебышевского приближения, ДАН, 101, № 3, стр. 409—412, 1955. 8. П. Л. Чебышев, Теория механизмов, известных под названием параллелограммов, 1854. Избр. тр. Изд. АН СССР, М., стр. 611—648, 1955. 9. E. Borel, Leçons sur les fonctions de variables réelles et les développements en séries de polynomes, Paris, 1905. 10. P. Laasonen, Einige Sätze über Tschebyscheffsche Funktionensysteme, Ann. Acad. Sc. Fenn., 52, p. p. 1—24, 1949. 11. Th. Motzkin, Approximation by curves of a unisolvent family, Bull. Amer. Math. Soc., 55, № 8, p. p. 789—793, 1949. 12. Ch.—J. de la Vallée Poussin, Sur les polynomes d'approximation et la représentation approchée d'un angle, Bull. Acad. de Belgique, p. p. 808—844, 1910. 13. Ch.—J. de la Vallée Poussin. Leçons sur l'approximation des fonctions d'une variable réelle, Paris, 1919.