

---



---

**МАТЕМАТИЧЕСКИЕ МЕТОДЫ  
И МОДЕЛИРОВАНИЕ В ПРИБОРОСТРОЕНИИ**

---



---

УДК 537.213, 537.612, 517.5

© А. С. Бердников, С. В. Масюкевич, 2023

**ЧИСЛЕННЫЙ АЛГОРИТМ ДЛЯ МИНИМАКСНОЙ  
ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ ФУНКЦИЙ  
С ЗАДАНЫМ ВЕСОМ**

В статье рассматривается быстросходящийся численный алгоритм для определения полиномов заданной степени, который обеспечивает на заданном интервале оптимальное приближение заданной функции в минимаксной норме с заданным весом при условии, что весовая функция не обращается в ноль на рассматриваемом интервале, за исключением, быть может, начальной и/или конечной точек интервала.

*Кл. сл.:* минимаксная норма, полиномы Чебышева, оптимальная аппроксимация, интерполяция

**ВВЕДЕНИЕ**

Применение аппроксимаций функций, оптимальных по минимаксной (равномерной) норме, имеет существенные преимущества по сравнению с более простой аппроксимацией функций по методу наименьших квадратов [1, 2]. В качестве полезного инструмента при конструировании таких аппроксимаций используются полиномы Чебышева первого рода, наименее отклоняющиеся от нуля с весовой функцией, равной единице [3, 4]. Разложение функции в усеченный ряд, состоящий из полиномов, наименее отклоняющихся от нуля (в случае единичной весовой функции — полиномов Чебышева), является одним из способов приближенного конструирования полиномиальных аппроксимаций, оптимальных по минимаксной норме [4, 5]. Также для конструирования полиномиальных аппроксимаций, которые являются хорошим приближением к оптимальной минимаксной аппроксимации, используют аппроксимацию по дискретному набору коллокационных точек, соответствующих нулям полинома Чебышева со степенью на единицу больше, чем степень аппроксимирующего полинома [1, 2, 4, 5].

Такие способы характеризуются удобством и легкостью конструирования, а также определенными эмпирическими предпосылками получить на выходе достаточно хорошую точность. Однако упрощенные способы аппроксимации отличаются от истинной аппроксимации, оптимальной по минимаксной норме, и, следовательно, обеспечивают неконтролируемо большую ошибку аппроксимации, чем точное решение соответствующей минимаксной оптимизационной задачи. Использование аппроксимаций, обеспечивающих для заданной функции точное решение минимаксной оптимизационной задачи (когда это возможно), определено более предпочтительно.

Множество функций, для которых имеется явное алгебраическое (аналитическое) представление для полиномиальной аппроксимации, являющейся оптимальной по минимаксной норме, не слишком велико (практически все такие случаи приводятся в [6–8]). В данной работе исследуется уточненный вариант быстро сходящегося численного алгоритма для определения коэффициентов полиномов, наименее отклоняющихся от заданной функции на заданном интервале с заданным весом, который ранее был рассмотрен в [9].

Множество функций, для которых имеется явное алгебраическое (аналитическое) представление для полиномиальной аппроксимации, являющейся оптимальной по минимаксной норме, не слишком велико (практически все такие случаи приводятся в [6–8]). В данной работе исследуется уточненный вариант быстро сходящегося численного алгоритма для определения коэффициентов полиномов, наименее отклоняющихся от заданной функции на заданном интервале с заданным весом, который ранее был рассмотрен в [9].

**МИНИМАКСНАЯ АППРОКСИМАЦИЯ  
ФУНКЦИЙ И ТЕОРИЯ ЧЕБЫШЕВА**

**Постановка задачи**

Пусть имеется непрерывная функция  $f(x)$  и конечный интервал  $[a, b]$ , для которого задана непрерывная весовая функция  $\omega(x)$ , являющаяся строго положительной на этом интервале, за исключением, быть может, концов интервала, для которых она может обращаться в ноль.<sup>1</sup> Требуется найти полином  $p(x)$  степени не выше  $n$  с заранее не-

---

<sup>1</sup> Для весовой функции с нулями внутри интервала  $[a, b]$  потребуется, чтобы максимумы и минимумы Чебышева (и пробные точки для численного алгоритма) не совпадали с нулями весовой функции, а знаки попеременных максимумов и минимумов изменялись в соответствии со знаком весовой функции. Также интервал  $[a, b]$  может быть бесконечным справа и/или слева, но при этом и аппроксимируемая функция  $f(x)$ , и весовая функция  $q(x)$  должны стремиться к нулю на бесконечности не медленнее, чем степенные функции  $1/x^k$ , причем с одинаковой скоростью [6–8].

определенными коэффициентами  $a_0, a_1, a_2, \dots, a_n$ :

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n, \quad (1)$$

который является решением оптимизационной задачи

$$\max |\omega(x) (f(x) - p(x))| \rightarrow \min, \quad (2)$$

где максимизация выполняется по переменной  $x \in [a, b]$ , а минимизация выполняется по коэффициентам  $a_0, a_1, a_2, \dots, a_n$ .

Без ограничения общности можно считать, что функция  $f(x)$  не является полиномом степени  $n$  или ниже, — в противном случае оптимальное решение для задачи (2) заранее известно и совпадает с функцией  $f(x)$ .

### Критерий Чебышева для минимаксной полиномиальной аппроксимации

Для того чтобы полином  $p(x)$  степени  $n$  был решением задачи (2), необходимо и достаточно, чтобы существовали такой набор из  $n + 2$  точек  $x_0 < x_1 < x_2 < \dots < x_{n+1}$ , принадлежащих интервалу  $[a, b]$ , и такое число  $\varepsilon$  (положительное либо отрицательное), чтобы были выполнены условия:

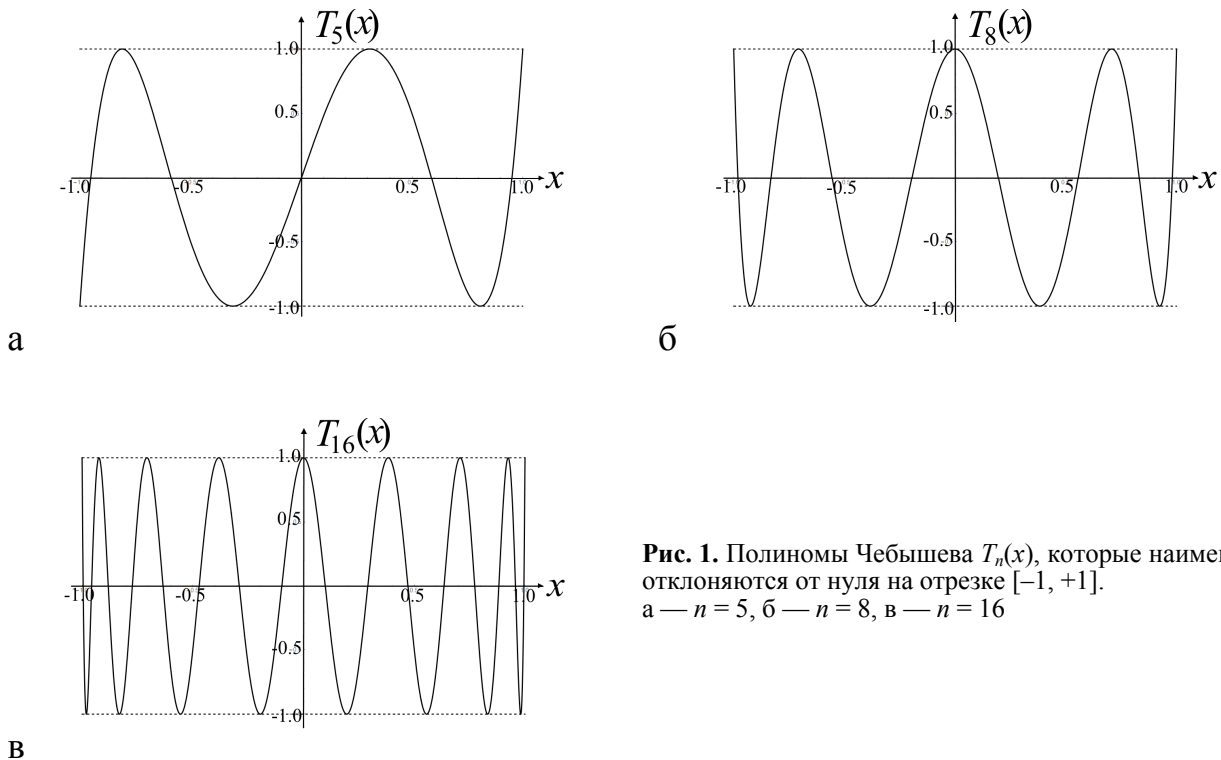
$$-\varepsilon \leq \omega(x) (f(x) - p(x)) \leq \varepsilon \text{ при } x \in [a, b], \quad (3)$$

$$\omega(x_k) (f(x_k) - p(x_k)) = (-1)^k \varepsilon \text{ при } k = 0, 1, \dots, n + 1. \quad (4)$$

Точки  $x_0 < x_1 < x_2 < \dots < x_{n+1}$  с чередующимися положительными минимумами и отрицательными максимумами, равными по абсолютной величине максимуму модуля исследуемой функции, называются альтернансом Чебышева (см. [7, 8, 10–14]). Графики полиномов Чебышева первого рода [3], демонстрирующие выполнение данного условия на интервале  $[-1, +1]$  с весом  $\omega(x) = 1$ , приводятся на рис. 1.

Смысл условий (3), (4) состоит в том, что для полинома  $p(x)$  имеется набор из  $n + 2$  пробных точек  $x_k$ , в которых невязка  $\omega(x) (f(x) - p(x))$  имеет локальные чередующиеся отрицательные минимумы и положительные максимумы, равные  $\pm\varepsilon$ , причем значения этих минимумов и максимумов являются глобальными на интервале  $[a, b]$ . Этот критерий является частным случаем теории Чебышева минимаксной аппроксимации функций с помощью дробно-рациональных функций [7, 8], однако в случае аппроксимации с помощью полиномов доказательство соответствующих утверждений упрощается.

**Утверждение 1 (теорема Валле – Пуссена [7, 8, 15–17]).** Если рассматриваемая функция  $\omega(x) (f(x) - p(x))$  для некоторого многочлена  $p(x)$  степени  $n$  принимает в  $N$  последовательных точках  $x_0 < x_1 < \dots < x_{N-1}$  интервала  $[a, b]$  отличные от нуля значения  $\lambda_0, \lambda_1, \dots, \lambda_{N-1}$  с чередующимися



**Рис. 1.** Полиномы Чебышева  $T_n(x)$ , которые наименее отклоняются от нуля на отрезке  $[-1, +1]$ . а —  $n = 5$ , б —  $n = 8$ , в —  $n = 16$

знаками, где  $N \geq n + 2$ , то для любого другого многочлена  $r(x)$  степени  $n$  справедливо

$$\max |\omega(x) (f(x) - r(x))| \geq \min \{|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|\}.$$

*Доказательство.* Пусть имеется многочлен  $r(x)$ , для которого выполняется условие  $\max |\omega(x) (f(x) - r(x))| < \min (|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|)$ . Будем считать, что в последовательности  $x_k$  для четных точек значения  $\omega(x_k) (f(x_k) - r(x_k))$  строго положительные, а для нечетных точек строго отрицательные. (Когда для нечетных точек значения положительные, а для четных точек отрицательные, рассуждения аналогичны.) Это означает, что в четных точках  $x_k$  выполнено условие  $\omega(x_k) (f(x_k) - r(x_k)) < |\lambda_k| = \omega(x_k) (f(x_k) - p(x_k))$ , а в нечетных точках  $x_k$  выполнено условие  $\omega(x_k) (f(x_k) - r(x_k)) > -|\lambda_k| = \omega(x_k) (f(x_k) - r(x_k))$ . В результате величина  $\omega(x) (p(x) - r(x))$  строго отрицательна в четных точках  $x_k$  и строго положительна в нечетных точках  $x_k$ . Значения весовой функции  $\omega(x_k)$  в точках  $x_k$ , для которых функция  $\omega(x) (f(x) - r(x))$  принимает строго положительные или строго отрицательные значения, не обращаются в ноль и, следовательно, строго положительны. Следовательно, не равный тождественно нулю многочлен  $p(x) - r(x)$  степени  $n$  попеременно принимает строго положительные и строго отрицательные значения как минимум в  $n + 2$  разных точках и в силу этого имеет не менее  $n + 1$  нулей кратности 1. Значит, многочлен  $r(x)$ , для которого выполняется условие  $\max |\omega(x_k) (f(x_k) - r(x_k))| < \min (|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|)$ , не существует.

*Замечание.* Теорема Валле – Пуссена позволяет получить для решения оптимизационной задачи (2) оценку снизу. Если  $\lambda_0, \lambda_1, \dots, \lambda_{N-1}$  — это локальные максимумы и минимумы функции  $\omega(x) (f(x) - r(x))$  с чередующимися знаками, то такая конструкция называется альтернансом Валле – Пуссена, и для решения оптимизационной задачи (2) она дает не только оценку снизу, но и оценку сверху, равную  $\max \{|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|\}$ .

**Утверждение 2.** Если для полинома  $p(x)$  выполнены условия (3), (4) критерия Чебышева, то оптимальное значение для правой части оптимизационной задачи (2) равно  $|\varepsilon|$ , а полином  $p(x)$  является ее решением (возможно, одним из многих).

*Доказательство.* Альтернанс Чебышева (3), (4) является частным случаем альтернанса Валле – Пуссена, поэтому, согласно Утверждению 1, правая часть оптимизационной задачи (2) не меньше  $|\varepsilon|$ . С другой стороны,  $\max |\omega(x) (f(x) - p(x))| = |\varepsilon|$ . Следовательно, минимум оптимизационной задачи (2) равен  $|\varepsilon|$ , а полином  $p(x)$  является одним из ее решений.

**Утверждение 3.** Существует полином  $p(x)$ , который обеспечивает решение оптимизационной задачи (2). (Это исключает вариант, когда есть полиномы  $p_k(x)$  со все более и более уменьшающимися значениями для  $P_k = \max |\omega(x) (f(x) - p_k(x))|$ , в то время как минимум этой величины никогда не достигается на множестве полиномов заданной степени.)

*Доказательство.* Значения величин  $P_k = \max |\omega(x) (f(x) - p_k(x))|$  ограничены снизу нулем, поэтому для значений  $P_k$  на множестве полиномов существует точная нижняя грань  $P$ . В соответствии с определением точной нижней грани существует последовательность полиномов  $p_k(x)$  степени  $n$ :

$$p_k(x) = a_{0,k} + a_{1,k}x + a_{2,k}x^2 + \dots + a_{n-1,k}x^{n-1} + a_{n,k}x^n,$$

для которой  $\lim P_k = P$  при  $k \rightarrow \infty$ . Следовательно, начиная с некоторого номера  $k$  значения многочленов  $p_k(x)$  на интервале  $[a + \delta, b - \delta]$ , где  $\delta$  достаточно мало, будут ограничены сверху и снизу (от концов интервала пришлось сделать отступ, чтобы учесть случай, когда весовая функция на концах интервала может быть равна нулю). Из формулы Лагранжа [1, 2] для полинома степени  $n$ , который в  $n + 1$  заданных точках  $x_k$  принимает значения  $y_k$

$$p(x) = \sum_{j=1, n+1} \left( y_j \prod_{i=1, n+1, i \neq j} \frac{(x - x_i)}{(x_j - x_i)} \right),$$

следует, что когда в  $n + 1$  фиксированных точках  $x_k$  значения полинома  $y_k$  ограничены сверху и снизу, то каждый из коэффициентов полинома также ограничен сверху и снизу. Следовательно, в соответствии с теоремой Больцано – Вейерштрасса (иначе называемой леммой Больцано – Вейерштрасса) о предельной точке, согласно которой из всякой бесконечной ограниченной последовательности точек пространства  $R^n$  можно выделить сходящуюся подпоследовательность, в последовательности полиномов  $p_k(x)$  можно выбрать подпоследовательность полиномов, в которой для каждого из коэффициентов полинома имеется предел. То есть имеется такая последовательность полиномов  $p_k(x)$ , для которой  $\lim P_k = P$  при  $k \rightarrow \infty$ , а у коэффициентов полиномов имеются предельные значения  $b_j = \lim a_{j,k}$  при  $k \rightarrow \infty$ .

Рассмотрим полином  $r(x)$  с коэффициентами  $b_j$ :

$$r(x) = b_0 + b_1x + b_2x^2 + \dots + b_{n-1}x^{n-1} + b_nx^n.$$

Поскольку  $b_j = \lim a_{j,k}$ , то для любого фиксированного значения  $x$  выполнится условие  $\lim p_k(x) = r(x)$  при  $k \rightarrow \infty$ , причем стремление к пределу является равномерным на рассматри-



Новый интервал обладает тем свойством, что из функции  $F(x)$  можно вычесть такую положительную константу, чтобы положительный максимум функции на интервале уменьшился, но при этом абсолютное значение отрицательного минимума функции на интервале не настолько увеличилось, чтобы превзойти новый максимум функции.

Точно так же если на каком-то интервале минимум равен  $-\lambda$ , а на соседнем интервале максимум меньше  $\lambda$ , то этот интервал вместе с последующим интервалом, содержащим отрицательный минимум, присоединяется к текущему интервалу (см. рис. 2, б). В этом случае на полученном интервале к функции  $F(x)$  можно прибавить такую положительную константу, чтобы уменьшилось абсолютное значение отрицательного минимума на интервале, но при этом не настолько увеличился положительный максимум функции на интервале, чтобы он стал больше абсолютного значения нового минимума.

В конечном итоге на интервале  $[a, b]$  для заданного полинома  $p(x)$  будет получен набор из  $N$  интервалов и принадлежащих им точек  $x_0, x_1, x_2, \dots, x_{N-1}$  с чередующимися положительными максимумами и отрицательными минимумами функции  $F(x)$ , равными  $\pm\lambda$ , где  $\lambda = \max |\omega(x) (f(x) - p(x))|$ . Если  $N$  больше или равно  $n + 2$ , то такой полином  $p(x)$  удовлетворяет критерию Чебышева и, в соответствии с Утверждением 1, будет решением оптимизационной задачи (2).

Пусть  $N$  меньше или равно  $n + 1$ . На этапе объединения интервалов был получен набор точек  $y_1, y_2, \dots, y_{N-1}$  (границы интервалов), в которых функция  $F(x)$  обращается в ноль и меняет свой знак. Эти точки разбивают интервал  $[a, b]$  на интервалы меньшего размера, на каждом из которых имеется точка с чебышевским минимумом или чебышевским максимумом. Также имеются константы  $\delta_k > 0$ , которые можно вычитать из функции  $F(x)$  на интервалах с положительными максимумами или прибавлять к функции  $F(x)$  на интервалах с отрицательными минимумами, уменьшая при этом минимаксную норму на соответствующем интервале.

Без ограничения общности можно считать, что на первом интервале функция  $F(x)$  принимает положительные значения, — когда это не так, можно изменить знак одновременно у функции  $f(x)$  и у полинома  $p(x)$ , не меняя оптимизационную задачу (2).

Рассмотрим функцию  $\Omega(x)$ , которая представляет собой произведение полинома степени не выше  $n$  и весовой функции  $\omega(x)$ :

$$\Omega(x) = \omega(x) (y_1 - x) (y_2 - x) \dots (y_{N-1} - x).$$

Функция  $\Omega(x)$  строго положительна на тех интер-

валах, на которых имеется чебышевский максимум функции  $F(x)$ , и строго отрицательна на тех интервалах, на которых имеется чебышевский минимум функции  $F(x)$ . Если вычесть из функции  $F(x)$  функцию  $s\Omega(x)$  с достаточно малым положительным множителем  $s$  (например, можно выбрать  $s = (\delta / R)$ , где  $R = \max |\Omega(x)|$ , а  $\delta = \min \delta_k > 0$ ), это безопасным образом уменьшит положительные максимумы функции  $F(x)$  и уменьшит отрицательные минимумы функции  $F(x)$ , тем самым уменьшив величину  $\max |F(x)|$ . Следовательно, такой полином  $p(x)$  не может быть решением для оптимизационной задачи (2), поскольку можно получить еще меньшее значение для  $\max |F(x)|$  при замене текущего полинома  $p(x)$  степени  $n$  на новый полином  $r(x)$  степени  $n$ :

$$\begin{aligned} G(x) &= F(x) - (\delta / R) \Omega(x) = \\ &= \omega(x) (f(x) - p(x)) - (\delta / R) \omega(x) (y_1 - x) \dots (y_{N-1} - x) = \\ &= \omega(x) (f(x) - p(x) - (\delta / R) (y_1 - x) \dots (y_{N-1} - x)) = \\ &= \omega(x) (f(x) - r(x)), \end{aligned}$$

$$\max |G(x)| < \max |F(x)|.$$

Вывод: оптимальный полином степени  $n$ , рассмотренный в Утверждении 2, обязан удовлетворять критерию Чебышева.

**Утверждение 5.** Полином  $p(x)$ , удовлетворяющий критерию Чебышева и обеспечивающий решение оптимизационной задачи (2), является единственным.

*Доказательство.* Идея доказательства заимствована из [7, 8]. Пусть имеются два полинома  $p(x)$  и  $r(x)$  степени  $n$ , которые являются решением оптимизационной задачи (2). Для вырожденного случая, для которого  $\max |\omega(x) (f(x) - p(x))| = 0$  и  $\max |\omega(x) (f(x) - r(x))| = 0$ , должны выполняться условия  $f(x) - p(x) \equiv 0$  и  $f(x) - r(x) \equiv 0$ , — что означает, что  $p(x) \equiv r(x)$ . Поэтому будем считать, что

$$\max |\omega(x) (f(x) - p(x))| = \max |\omega(x) (f(x) - r(x))| > 0,$$

так что у выражений  $\omega(x) (f(x) - p(x))$  и  $\omega(x) (f(x) - r(x))$  на интервале  $[a, b]$  имеются не равные нулю максимумы и минимумы.

В соответствии с Утверждением 3, каждый из полиномов  $p(x)$  и  $r(x)$  удовлетворяет критерию Чебышева, причем значение  $\varepsilon \neq 0$  для них одно и то же. Пусть  $x_1, x_2, \dots, x_{n+2}$  — это чебышевские точки уклонения с чередующимися максимумами и минимумами, равными  $\pm\varepsilon$ , для функции  $F(x) = \omega(x) (f(x) - p(x))$ . Если для полинома  $r(x)$  функция  $G(x) = \omega(x) (f(x) - r(x))$  принимает в точ-

ках  $x_1, x_2, \dots, x_{n+2}$  те же самые значения  $\pm\epsilon$ , кроме, может быть, одной точки из этого списка, то полиномы  $p(x)$  и  $r(x)$  тождественно равны друг другу. Действительно, в этом случае разность

$$\begin{aligned} d(x) &= F(x) - G(x) = \\ &= \omega(x) (f(x) - p(x)) - \omega(x) (f(x) - r(x)) = \\ &= \omega(x) (r(x) - p(x)), \end{aligned} \quad (5)$$

которая представляет собой произведение положительной весовой функции  $\omega(x)$  и полинома степени не выше  $n$ , обращается в ноль в  $n + 2$  точках или, в крайнем случае, в  $n + 1$  точке, — и поэтому полином  $r(x) - p(x)$  обязан быть тождественным нулем, поскольку иначе не объяснить наличия у него такого количества нулей (весовая функция  $\omega(x)$ , даже если она обращается в ноль на концах интервала  $[a, b]$ , не равна нулю в точках минимумов и максимумов функции  $F(x)$ ).

Пусть среди чебышевских точек уклонения  $x_1, x_2, \dots, x_{n+2}$  для функции  $F(x)$  имеются хотя бы две точки  $x_k$  и  $x_{k+m}$ , в которых значение функции  $G(x)$  не совпадает со значением функции  $F(x)$  в этих точках (минимумом либо максимумом функции  $F(x)$ ). Требуется доказать, что в таком случае внутри интервала  $[x_k, x_{k+m}]$  найдется по крайней мере еще одна точка, не совпадающая с точками  $x_1, x_2, \dots, x_{n+2}$ , в которой значения функций  $F(x)$  и  $G(x)$  совпадают (хотя, возможно, в этой точке общее значение функций  $F(x)$  и  $G(x)$  и не равно  $\pm\epsilon$ ), — а следовательно, у функции  $d(x)$ , заданной формулой (5), имеется дополнительный неучтенный ноль.

Если в какой-то точке  $x_j$  выполнено условие  $d(x_j) \neq 0$ , то знак разности  $d(x_j)$ , заданной формулой (5), совпадает со знаком  $F(x_j)$ . Действительно, если  $x_j$  это точка отрицательного минимума функции  $F(x)$ , то  $G(x_j) \geq F(x_j)$ , а если  $x_j$  это точка положительного максимума функции  $F(x)$ , то  $G(x_j) \leq F(x_j)$ . Если

$$\begin{aligned} d(x_k) \neq 0, d(x_{k+1}) = 0, d(x_{k+2}) = 0, \dots, d(x_{k+m-1}) = 0, \\ d(x_{k+m}) \neq 0, \end{aligned} \quad (6)$$

то выражения  $(-1)^k F(x_k)$  и  $(-1)^{k+m} F(x_{k+m})$  имеют одинаковые знаки и, следовательно, одинаковые знаки имеют выражения  $d(x_k)$  и  $(-1)^m d(x_{k+m})$ . Следовательно, четность числа перемен знака функции  $d(x)$  на интервале  $[x_k, x_{k+m}]$  совпадает с четностью числа  $m$ , а тем самым и четность числа корней (с учетом их кратности) для функции  $d(x)$  на интервале  $[x_k, x_{k+m}]$  совпадает с четностью числа  $m$ . Согласно (6) у функции  $d(x)$  на интервале  $[x_k, x_{k+m}]$  имеется как минимум  $m - 1$  не равных друг другу корней и, следовательно, либо должен существовать еще один корень, либо один из корней  $x_{k+1}, x_{k+2}, \dots, x_{k+m-1}$  имеет четную кратность, как

минимум равную двум. Прделав эту операцию для всех интервалов, ограниченных точками  $x_k$  с ненулевыми значениями  $d(x_k)$ , получим, что число корней функции  $d(x)$  на интервале  $[a, b]$  с учетом кратности корней будет не меньше, чем  $n + 1$ . Следовательно, в формуле  $d(x) = \omega(x) (r(x) - p(x))$  полином  $r(x) - p(x)$ , имеющий степень не выше  $n$ , обязан быть тождественным нулем.

### АЛГОРИТМ КОНСТРУИРОВАНИЯ МИНИМАКСНОЙ ПОЛИНОМИАЛЬНОЙ АППРОКСИМАЦИИ ФУНКЦИИ

Из критерия Чебышева для минимаксной аппроксимации функции выводится алгоритм численного построения аппроксимирующих полиномов, который представляет собой измененный и оптимизированный алгоритм Е.Я. Ремеза [18–22]:

1. Произвольным образом выбирается начальный набор точек  $x_0 < x_1 < x_2 < \dots < x_{n+1}$ , принадлежащих интервалу  $[a, b]$ . Если весовая функция  $\omega(x)$  обращается в ноль в точке  $x = a$ , необходимо обеспечить выполнение условия  $a < x_0$ . Если весовая функция  $\omega(x)$  обращается в ноль в точке  $x = b$ , необходимо обеспечить выполнение условия  $x_0 < b$ .

2. Для текущего набора точек  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$  конструируется вспомогательная функция  $g(x) = a f(x) - p(x)$ , где  $a$  это некоторая константа, а  $p(x)$  — полином степени  $n$  с некоторыми коэффициентами. Свободные коэффициенты функции  $g(x)$ , т.е. множитель  $a$  и коэффициенты полинома  $p(x)$ , выбираются так, чтобы были выполнены условия  $\omega(x_k)g(x_k) = (-1)^k$  при  $k = 0, 1, \dots, n + 1$ . Такая функция существует и определяется единственным образом (см. далее).

3. Теперь надо определить корни алгебраического уравнения  $g(x) = 0$ , которые находятся на интервале  $[a, b]$ . Поскольку на краях интервалов  $[x_k, x_{k+1}]$  значения функции  $g(x)$  имеют разные знаки, то внутри каждого такого интервала существует по крайней мере один корень уравнения  $g(x) = 0$ . Если функция  $f(x)$  меняется достаточно медленно, то на каждом из интервалов  $[x_k, x_{k+1}]$  имеется ровно один корень кратности единицы, а численное определение этих корней не представляет трудности. Если же это условие нарушается и на рассматриваемом интервале имеется несколько корней, все эти корни добавляются к списку (кратные корни с четной кратностью, для которых значения функции  $g(x)$  по обе стороны от корня имеют один и тот же знак, в список не включаются). По завершении работы по поиску корней к списку значений  $a < y_1 < y_2 < \dots < y_m < b$ , являющихся корнями нечетной кратности функ-

ции  $g(x)$  на интервале  $[a, b]$ , добавляются начало и конец интервала  $y_0 = a$  и  $y_{m+1} = b$ .

4. Теперь интервал  $[a, b]$  представляет собой совокупность не менее чем  $n + 2$  интервалов  $[y_k, y_{k+1}]$ , которые соприкасаются своими конечными точками, причем на каждом интервале функция  $g(x)$  сохраняет знак, а при пересечении границы между интервалами знак функции меняется. Рассмотрим функцию  $G(x) = \omega(x) g(x)$ , которая на интервале  $[a, b]$  принимает значения того же знака, что и функция  $g(x)$ . Найдем для функции  $G(x)$  на каждом интервале  $x \in [y_k, y_{k+1}]$  глобальный (для рассматриваемого интервала) экстремум  $G_k$  и точку  $z_k \in [y_k, y_{k+1}]$ , в которой расположен этот экстремум. Для интервалов с положительными значениями функции  $G(x)$  надо искать положительный максимум, а для интервалов  $[y_k, y_{k+1}]$  с отрицательными значениями функции — отрицательный минимум.

5. Если число найденных точек  $G_k = G(z_k)$ ,  $z_k \in [y_k, y_{k+1}]$  с чередующимися знаками больше  $n + 2$ , необходимо исключить лишние точки. Для этого выбирается наименьшее по модулю значение  $G_k$ . Если интервал с этим значением первый либо последний в списке, значение  $G_k$  и соответствующая этому значению точка  $z_k$  отбрасываются. Если интервал с этим значением находится в середине списка, надо проверить предшествующий и последующий интервалы и отбросить не только рассматриваемое значение  $G_k$ , но также и то из двух соседних значений  $G_{k-1}$  и  $G_{k+1}$ , которое будет меньше по модулю. Действительно, знаки значений  $G_{k-1}$  и  $G_{k+1}$  противоположны по знаку значению  $G_k$ , а т.к. из-за удаления из списка значения  $G_k$  два значения одного знака окажутся рядом, то из двух значений  $G_{k+1}$  и  $G_{k-1}$  одно надо отбросить. Важно: поскольку на каждом шаге удалится одна либо две точки, то есть риск из  $n + 3$  точек  $z_k$  получить сразу  $n + 1$  точек, вместо требуемых  $n + 2$  точек, — в этом случае следует оставить точку  $G_k$  в списке, а вместо нее удалить ту из начальной и конечной точек, которая меньше по модулю. Процесс продолжается, пока в списке не окажется  $n + 2$  экстремумов  $G_k = G(z_k)$ , которым соответствуют  $n + 2$  точек  $z_k$  с чередующимися положительными максимумами и отрицательными минимумами функции  $G(x) = \omega(x) g(x)$ .

6. Для имеющегося списка точек  $a \leq z_0 < z_1 < z_2 < \dots < z_{n+1} \leq b$  рассмотрим более внимательно значения экстремумов  $G_k = \omega(z_k) g(z_k) = (-1)^k \varepsilon_k$ . Если в пределах заданной точности  $\varepsilon_k \approx 1$ , то полином, удовлетворяющий требованиям критерия Чебышева, найден. При этом из теоремы Валле – Пуссена о наилучшем приближении функции полиномами (Утверждение 1) следует, что точный оптимум для задачи (2) лежит в диапазоне между  $\min |\varepsilon_k / a|$  и  $\max |\varepsilon_k / a|$  (где  $a$  это множитель в вы-

ражении  $g(x) = a f(x) - p(x)$ ). Если же условие  $\varepsilon_k \approx 1$  не выполнено, заменяем пробные точки  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$  на точки  $a \leq z_0 < z_1 < z_2 < \dots < z_{n+1} \leq b$  и возвращаемся к шагу 2.

7. Остается нормировать функцию  $g(x) = a f(x) - p(x)$  так, чтобы коэффициент при функции  $f(x)$  стал равен единице. Такая нормировка всегда осуществима, поскольку коэффициент  $a$  не может быть нулем, — иначе окажется, что имеется отличный от тождественного нуля полином  $p(x)$  степени  $n$ , у которого есть  $n + 1$  корней, находящихся между соседними максимумами и минимумами разных знаков в количестве  $n + 2$  штук.

Очевидно, что если этот алгоритм сходится, то он сходится к полиному, удовлетворяющему критерию Чебышева, и тем самым результат работы алгоритма является решением оптимизационной задачи (2), причем единственным таким решением. Теоретически алгоритм сходится всегда, причем со скоростью геометрической прогрессии: для любой непрерывной весовой функции и для любой степени  $n$  найдутся такие числа  $C > 0$  и  $0 < \lambda < 1$ , для которых

$$\max |\omega(x) (a_i f(x) - p_i(x))| \leq 1 + C \lambda^i,$$

где  $i$  это номер итерации. Доказательство аналогично доказательству сходимости алгоритма Е.Я. Ремеза в монографии [20]. К сожалению, из-за ошибок округления алгоритм может расходиться для "плохих" функций  $f(x)$  и  $\omega(x)$ , — а именно для функций, которые не меняются медленно на интервале  $[a, b]$ , а вместо этого демонстрируют многочисленные осцилляции, так что для них имеется более одного корня для интервалов  $[x_k, x_{k+1}]$  или более одного локального максимума или минимума для интервалов  $[y_k, y_{k+1}]$ . В таких случаях алгоритм может выполнять бесконечный цикл, фактически переключаясь между одним и тем же списком тестируемых полиномов. Такое поведение алгоритма следует контролировать вручную, задавая максимальное количество циклов, которое разрешено выполнить. Другим полезным критерием для прекращения работы алгоритма, попавшего в бесконечный цикл, может быть проверка возрастания на очередном шаге разброса между максимумами и минимумами  $\pm \varepsilon_k$ , найденными на интервалах  $[y_k, y_{k+1}]$ , по сравнению с этой величиной для предыдущего шага. Также подозрительным признаком, сигнализирующим о заикливании алгоритма, может быть регулярное возникновение ситуации, когда на шаге 5 при "чистке" точек  $G_k$  на каком-то шаге в списке оказывается меньше чем  $n + 2$  точки и возникает необходимость сделать шаг назад, удалив первую или последнюю точку, — ведь даже если оптимальным многочленом  $p(x)$  является тождест-

венный ноль (например, такая ситуация возникает при попытке аппроксимации на интервале  $[0, 1]$  многочленом конечной степени функции  $\sin(1/x)$ ), то, согласно Утверждению 4, все равно должен существовать набор из  $n + 2$  точек  $z_k$ , которые формируют альтернанс Чебышева.

Так как весовая функция  $\omega(x)$  на каком-то из концов интервала  $[a, b]$  может обращаться в ноль, то для выполнения шага 2 потребуется, чтобы пробные точки, выбираемые на шаге 1, лежали строго внутри интервала  $[a, b]$ :  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$ . Можно убедиться, что если  $\omega(a) = 0$  либо  $\omega(b) = 0$  и при этом  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$ , то условие  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$  сохранится для всех последующих итераций, т.к. точки  $z_k$  на шаге 6 также должны будут удовлетворять условию  $a < z_0 < z_1 < z_2 < \dots < z_{n+1} < b$ .

Для вычисления коэффициентов на шаге 2 нельзя применять явное решение системы линейных уравнений для нахождения коэффициентов функции  $g(x) = a \cdot f(x) - p(x)$ , поскольку обусловленность матрицы линейных уравнений (за исключением первого столбца, совпадающая с матрицей Вандермонда) быстро растет с увеличением степени полинома — в результате малые ошибки округления чисел с плавающей точкой приводят к совершенно неудовлетворительным значениям для неизвестных коэффициентов. Один из работоспособных способов нахождения искомого функции  $g(x) = a \cdot f(x) - p(x)$  на шаге 2 состоит в следующем. Из условий  $p(x_k) = a \cdot f(x_k) - (-1)^k / \omega(x_k)$ , которые равносильны условиям  $\omega(x_k) (a \cdot f(x_k) - p(x_k)) = (-1)^k$ , при  $k = 0, 1, \dots, n$  по формуле Лагранжа<sup>2</sup> [1, 2] находятся коэффициенты полинома  $p(x)$ , который в заданных точках  $x_k$  принимает заданные значения. Эти коэффициенты будут представлять собой линейные выражения со свободным коэффициентом  $a$ . После этого для уравнения  $\omega(x_k) (a \cdot f(x_k) - p(x_k)) = (-1)^k$  при  $k = n + 1$ , которое представляет собой линейное уравнение относительно переменной  $a$ , находится неизвестный коэффициент  $a$ .

Другой, более эффективный способ определения коэффициента  $a$  и коэффициентов полинома  $p(x)$ , для которых на шаге 2 будут выполнены условия

$$\omega(x_k) (a \cdot f(x_k) - p(x_k)) = (-1)^k \text{ при } k = 0, 1, \dots, n+1,$$

состоит в следующем. Для выполнения равенств  $p(x_k) = a \cdot f(x_k) - (-1)^k / \omega(x_k)$  по схеме конечных раз-

ностей Ньютона [1, 2] определяются такой полином  $r(x)$ , для которого выполнены условия  $r(x_k) = (-1)^k / \omega(x_k)$  при  $k = 0, 1, \dots, n$ , и полином  $s(x)$ , для которого выполнены условия  $s(x_k) = f(x_k)$  при  $k = 0, 1, \dots, n$ . Затем из условия

$$\omega(x_{n+1}) (a \cdot f(x_{n+1}) - a \cdot s(x_{n+1}) + r(x_{n+1})) = (-1)^{n+1}$$

определяется константа  $a$ . После этого искомым полином  $p(x)$  определяется в соответствии с формулой  $p(x) = a \cdot s(x) - r(x)$ . Этот способ можно использовать также в следующем виде: вычисляются полином  $s(x)$  степени  $n + 1$ , для которого в точках  $x_k$  при  $k = 0, 1, \dots, n + 1$  выполнены условия  $s(x_k) = f(x_k)$ , и полином  $r(x)$  степени  $n + 1$ , для которого в точках  $x_k$  при  $k = 0, 1, \dots, n + 1$  выполнены условия  $r(x_k) = (-1)^k / \omega(x_k)$ . Остается подобрать коэффициент  $a$  так, чтобы полином  $a \cdot s(x) - r(x)$  имел степень  $n$ , — что всегда возможно, если только функция  $f(x_k)$  не является полиномом степени не выше  $n$ .

Отметим, что на промежуточных этапах нет необходимости определять коэффициенты полинома  $p(x)$  в явном виде, достаточно уметь вычислять его значение в произвольной точке. Поэтому при компьютерной реализации данного алгоритма вплоть до получения итогового результата можно использовать вычисление значений полинома  $p(x)$  по формуле  $p(x) = a \cdot s(x) - r(x)$ , где для вспомогательных полиномов  $s(x)$  и  $r(x)$  используется форма конечных разностей Ньютона.

Для улучшения работы алгоритма на начальном этапе на шаге 1 рекомендуется выбирать для начальных точек  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$  нули полинома Чебышева первого рода степени  $n + 2$ , смещенные и масштабированные с интервала  $[-1, +1]$  к интервалу  $[a, b]$ . На шаге 2 вместо условия  $\omega(x_k) p(x_k) = (-1)^k$  рекомендуется использовать условие  $\omega(x_k) p(x_k) = (-1)^k (4 / (b - a))^{n-1}$ , чтобы избежать неоправданно больших коэффициентов у полинома  $p(x)$  и функции  $f(x)$  (такой выбор масштаба соответствует осцилляциям полинома Чебышева степени  $n + 1$ , пересчитанного от интервала  $[-1, +1]$  к интервалу  $[a, b]$  и масштабированного так, чтобы старший коэффициент был равен единице).

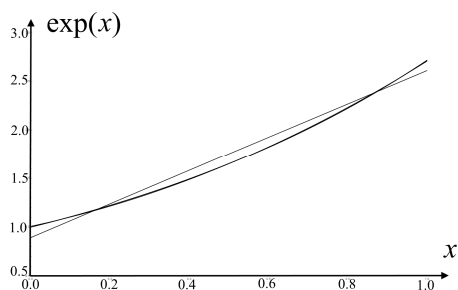
**Пример.** Пусть  $f(x) = \exp(x)$ ,  $a = 0$ ,  $b = 1$ ,  $\omega(x) = 1$ . Для  $n = 1, 2, \dots, 8$  в результате работы описанного алгоритма получаем полиномы, которые обеспечивают оптимальную (по минимаксной норме) аппроксимацию функции  $\exp(x)$  на интервале  $[0, 1]$ . Ошибка аппроксимации в зависимости от степени полинома указана в табл. Зависимость ошибки аппроксимации от степени полинома в графической форме показана на рис. 3.

<sup>2</sup> Вместо явной формулы Лагранжа, особенно при больших степенях полиномов, удобно применять итерационную схему Эйткена [2]. Хотя этот алгоритм требует больше памяти, он обеспечивает существенно большую точность коэффициентов и требует меньше компьютерного времени.

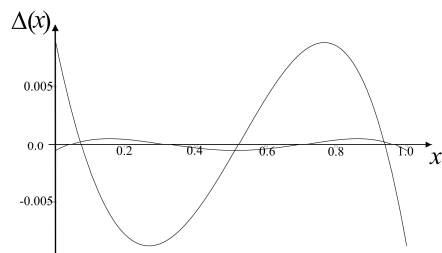


**Табл.** Ошибка аппроксимации функции  $\exp(x)$  на интервале  $[0, 1]$  многочленами степени  $n$ , оптимальными по минимаксной норме с весом  $q(x) = 1$

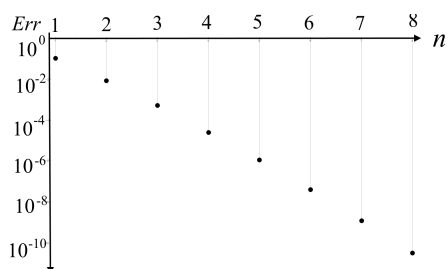
$n$	1	2	3	4	5	6	7	8
$\varepsilon$	0.1	$8 \cdot 10^{-3}$	$5 \cdot 10^{-4}$	$3 \cdot 10^{-5}$	$10^{-6}$	$4 \cdot 10^{-8}$	$10^{-9}$	$3 \cdot 10^{-11}$



а

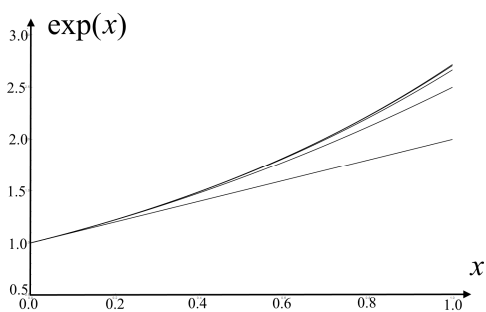


б

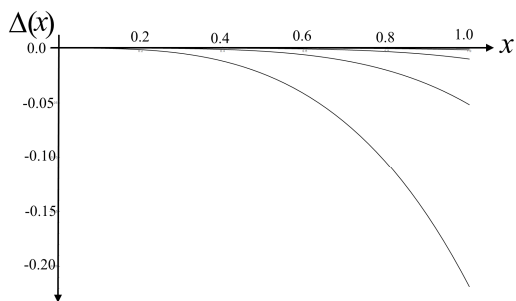


в

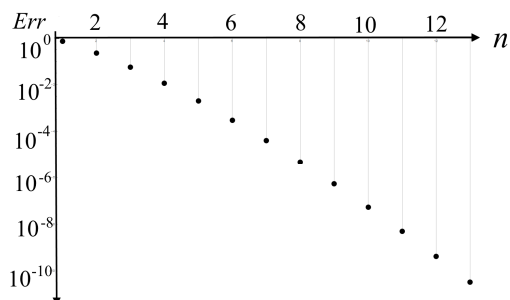
**Рис. 3.** Аппроксимация функции  $\exp(x)$  на отрезке  $[0, 1]$  минимаксными многочленами степени  $n$ . а — экспоненциальная функция и ее аппроксимации, б — разность между экспоненциальной функцией и ее аппроксимациями, в — минимаксная ошибка в логарифмическом масштабе



а



б



в

**Рис. 4.** Аппроксимация функции  $\exp(x)$  на интервале  $[0, 1]$  усеченными рядами Тейлора степени  $n$ . а — экспоненциальная функция и ее аппроксимации, б — разность между экспоненциальной функцией и ее аппроксимациями, в — минимаксная ошибка в логарифмическом масштабе

Для сравнения те же данные для усеченных рядов Тейлора показаны на рис. 4. Следует отметить, что ряд Тейлора для функции  $\exp(x)$  сходится очень быстро (в отличие от ряда Тейлора для функции  $\log(x)$ , например), однако даже в этом случае минимаксное приближение дает существенное преимущество.

*Работа выполнена в Институте аналитического приборостроения Российской академии наук (Санкт-Петербург) в рамках темы FFZM-2022-0009 (номер гос. регистрации 122040600002-3) государственного задания Министерства науки и высшего образования Российской Федерации № 075-01157-23-00 от 29.12.2022.*

*При проведении вычислений использовалась лицензионно чистая программа Wolfram Mathematica версии 11 (Home Edition) [23], лицензированная одному из авторов.*

### СПИСОК ЛИТЕРАТУРЫ

1. Ланцош К. Практические методы прикладного анализа. Справочное руководство. М.: ГИФМЛ, 1961. 524 с.
2. Березин И.С., Жидков Н.П. Методы вычислений, Т. 1. М.: ГИФМЛ, 1962. 464 с.
3. Данилов Ю.А. Многочлены Чебышева. Минск: Вышэйшая школа, 1984. 157 с.
4. Пашковский С. Вычислительные применения многочленов и рядов Чебышева. М.: Наука, 1983. 384 с.
5. Грибкова В.П. Эффективные методы равномерных приближений, основанные на полиномах Чебышева. М.: Изд-во "Спутник", 2017. 193 с.
6. Гончаров В.Л. Теория интерполирования и приближения функций. М.–Л.: ГТТИ, 1934. 316 с.
7. Ахиезер Н.И. Лекции по теории аппроксимации. М.–Л.: ОГИЗ, 1947. 324 с.
8. Ахиезер Н.И. Лекции по теории аппроксимации. 2-е изд. М.: Наука, 1965. 407 с.
9. Berdnikov A., Solovyev K., Krasnova N., Golovitski A., Syasko M. Algorithm for Constructing the Chebyshev-Type Polynomials and the Chebyshev-Type Approximations with a Given Weight // Proceedings of 2022 Int. Conference on Electrical Engineering and Photonics (EEExPolytech). P. 143–145. DOI: 10.1109/EEExPolytech56308.2022
10. Чебышев П.Л. Теория механизмов, известных под именем параллелограммов // П.Л. Чебышев. Избранные труды. М.: Изд-во Академии Наук СССР, 1955. С. 611–648.
11. Чебышев П.Л. Вопросы о наименьших приближениях, связанных с приближенным представлением функций // П.Л. Чебышев. Избранные труды. М.: Изд-во Академии Наук СССР, 1955. С. 462–578.
12. Чебышев П.Л. О функциях, наименее уклоняющихся от нуля // П.Л. Чебышев. Избранные труды. М.: Изд-во Академии Наук СССР, 1955. С. 579–608.
13. Бернштейн С.Н. О наилучшем приближении непрерывных функций посредством многочленов данной степени // Сообщения Харьковского математического общества. Вторая сер. 1912. Т. 13, вып. 2-3. С. 49–144.
14. Бернштейн С.Н. О наилучшем приближении функций нескольких переменных посредством многочленов или тригонометрических сумм // Труды МИАН СССР. 1951. Т. 38. С. 24–29.
15. De la Vallée-Poussin Ch.J. Sur les polynomes d'approximation et la représentation approchée d'un angle // Bulletin de la classe des sciences, Académie royale de Belgique, 1910. No. 12. P. 808–845.
16. De la Vallée-Poussin Ch.J. Leçons sur l'approximation des fonctions d'une variable réelle: professées à la Sorbonne. Paris: Gauthier-Villars, 1919.
17. Butzer P.L., Nessel R.J. Aspects of de La Vallée Poussin's work in approximation and its influence // Archive for History of Exact Sciences. 1993. Vol. 46. P. 67–95. DOI: 10.1007/BF00387727
18. Remez E. Sur un procédé convergent d'approximations successives pour déterminer les polynomes d'approximation // Compt. Rend. Acad. Sci. Paris, 1934. Vol. 198. P. 2063–2065.
19. Remez E.Ya. Sur le calcul effectif des polynomes d'approximation de Tschebyscheff // Compt. Rend. Acad. Sci. Paris, 1934. Vol. 199. P. 337–340.
20. Дзядык В.К. Введение в теорию равномерного приближения функций полиномами. М.: Наука, Главная редакция физико-математической литературы, 1977. 508 с.
21. Рemez Е.Я. Основы численных методов чебышевского приближения. Киев: Наукова думка, 1969. 624 с.
22. Лоран П.-Ж. Аппроксимация и оптимизация. М.: Мир, 1975. 496 с.
23. Wolfram Mathematica: наиболее полная система для математических и технических вычислений. URL: <https://www.wolfram.com/mathematica/> (обращение 21.03.2023).

**Институт аналитического приборостроения РАН, Санкт-Петербург**

Контакты: Бердников Александр Сергеевич, asberd@yandex.ru

Материал поступил в редакцию 28.03.2023

# NUMERICAL ALGORITHM FOR MINIMAX POLYNOMIAL APPROXIMATION OF FUNCTIONS WITH A GIVEN WEIGHT

A. S. Berdnikov, S. V. Masyukevich

*Institute for Analytical Instrumentation of RAS, Saint Petersburg, Russia*

The article discusses a rapidly converging numerical algorithm for determining the coefficients of polynomials. The algorithm provides the optimal approximation of a given function in the minimax norm with a given weight on a given interval. The approximation is made under the condition that the weight function does not turn to zero on the considered interval, except, perhaps, for the initial and/or final points of the interval.

*Keywords:* minimax norm, Chebyshev polynomials, optimal approximation

## INTRODUCTION

The use of function approximations that are optimal in terms of the minimax (uniform) norm has significant advantages over the simpler approximation of functions using the least squares method [1, 2]. As a practical tool in constructing such approximations, the Chebyshev polynomials of the first kind are used, deviating least from zero and having a weight function equal to one [3, 4]. The decomposition of a function into a truncated series consisting of polynomials that deviate least from zero (the Chebyshev polynomials in the case of a unit weight function) is one of the methods for the close construction of polynomial approximations that are optimal in the minimax norm [4, 5]. Also, to construct polynomial approximations that are close to the optimal minimax approximation, one uses an approximation over a discrete set of collocation points corresponding to the zeros of the Chebyshev polynomial with a degree one greater than the degree of the approximating polynomial [1, 2, 4, 5].

Such methods are characterized by convenience and ease of design, as well as having certain empirical prerequisites to obtain quite good accuracy. However, the simplified methods of approximation differ from the true approximation, which is optimal in terms of the minimax norm, and, therefore, the former provide an uncontrollably larger approximation error than the exact solution of the corresponding minimax optimization problem. The use of approximations that provide an exact solution to the minimax optimization problem for a given function (when possible) is definitely preferable.

The set of functions for which there is an explicit algebraic (analytical) representation for a polynomial approximation that is optimal in the minimax norm, is not too large (practically all such cases are given in [6–8]). In this paper, we study an improved version of a rapidly converging numerical algorithm for determining the coefficients of polynomials that deviate

least from a given function on a given interval with a given weight, which was previously considered in [9].

## MINIMAX FUNCTION APPROXIMATION AND THE THEORY OF CHEBYSHEV

### Formulation of the problem

Let there be a continuous function  $f(x)$  and a finite interval  $[a, b]$  for the latter a continuous weight function  $\omega(x)$  is given, which is strictly positive on this interval, except, perhaps, for the ends of the interval, for which it can turn into zero.<sup>1</sup> It is necessary to find a polynomial  $p(x)$  of degree at most  $n$  with previously undefined coefficients  $a_0, a_1, a_2, \dots, a_n$ :

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n, \quad (1)$$

which is the solution to the optimization problem

$$\max |\omega(x) (f(x) - p(x))| \rightarrow \min, \quad (2)$$

where the variable  $x \in [a, b]$  is maximized, and the coefficients  $a_0, a_1, a_2, \dots, a_n$  are minimized.

Without loss of generality, we can assume that the function  $f(x)$  is not a polynomial of degree  $n$  or lower, otherwise the optimal solution for problem (2) is known in advance and coincides with the function  $f(x)$ .

---

<sup>1</sup> For a weight function with zeros inside the interval  $[a, b]$ , it will be required that the Chebyshev maxima and minima (and trial points for the numerical algorithm) do not coincide with the zeros of the weight function, and the signs of the alternating maxima and minima change in accordance with the sign of the weight function. Also, the interval  $[a, b]$  can be infinite on the right and/or left, but both the approximated function  $f(x)$  and the weight function  $q(x)$  must tend to zero at infinity no slower than the power functions  $1/x^k$ , and with the same rate [6–8].

**The Chebyshev criterion for the minimax polynomial approximation**

For a polynomial  $p(x)$  of degree  $n$  to be a solution to problem (2), it is necessary and sufficient that there exist such a set of  $n + 2$  points  $x_0 < x_1 < x_2 < \dots < x_{n+1}$ , belonging to the interval  $[a, b]$ , and such a number  $\varepsilon$  (positive or negative) that the following conditions are met:

$$-\varepsilon \leq \omega(x) (f(x) - p(x)) \leq \varepsilon \text{ при } x \in [a, b], \quad (3)$$

$$\omega(x_k) (f(x_k) - p(x_k)) = (-1)^k \varepsilon \text{ при } k = 0, 1, \dots, n + 1. \quad (4)$$

The points  $x_0 < x_1 < x_2 < \dots < x_{n+1}$  with alternating positive minima and negative maxima equal in absolute value to the maximum modulus of the function under study are called the Chebyshev alternance (see [7, 8, 10–14]). Graphs of Chebyshev polynomials of the first kind [3], demonstrating the fulfillment of this condition on the interval  $[-1, +1]$  with the weight  $\omega(x) = 1$ , are shown in Fig. 1.

**Fig. 1.** Chebyshev polynomials  $T_n(x)$ , which deviate least from zero on the interval  $[-1, +1]$ .  
 a —  $n = 5$ , б —  $n = 8$ , в —  $n = 16$

The meaning of conditions (3), (4) is that for the polynomial  $p(x)$  there is a set of  $n + 2$  test points  $x_k$  at which the discrepancy  $\omega(x) (f(x) - p(x))$  has local alternating negative minima and positive maxima equal to  $\pm|\varepsilon|$ , and the values of these minima and maxima are global on the interval  $[a, b]$ . This criterion is a special case of Chebyshev's theory of minimax approximation of functions using fractional rational functions [7, 8], but in the case of approximation using polynomials, the proof of the appropriate statements is simplified.

**Statement 1 (The de la Vallée – Poussin theorem [7, 8, 15–17]).** If the considered function  $\omega(x) (f(x) - p(x))$  for some polynomial  $p(x)$  of degree  $n$  takes non-zero values  $\lambda_0, \lambda_1, \dots, \lambda_{N-1}$  with alternating signs, where  $N \geq n + 2$ , at  $N$  successive points  $x_0 < x_1 < \dots < x_{N-1}$  in interval  $[a, b]$ , then for any other polynomial  $r(x)$  of degree  $n$ ,

$$\max |\omega(x) (f(x) - r(x))| \geq \min \{|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|\}.$$

*Proof.* Let there be a polynomial  $r(x)$  satisfying the condition  $\max |\omega(x) (f(x) - r(x))| < \min (|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|)$ . We assume that in the sequence  $x_k$  the values  $\omega(x_k) (f(x_k) - r(x_k))$  are strictly positive for even points, and strictly negative for odd points. (When the values are positive for odd points and negative for even points, the reasoning is similar.) This means that the condition  $\omega(x_k) (f(x_k) - r(x_k)) < |\lambda_k| = \omega(x_k) (f(x_k) - p(x_k))$  is satisfied at even points, and the condition  $\omega(x_k) (f(x_k) - r(x_k)) > -|\lambda_k| = \omega(x_k) (f(x_k) - r(x_k))$  is sa-

tisfied at odd points. As a result, the value  $\omega(x) (p(x) - r(x))$  is strictly negative at even points  $x_k$  and strictly positive at odd points  $x_k$ . The values of the weight function  $\omega(x_k)$  at the points  $x_k$ , for which the function  $\omega(x) (f(x) - r(x))$  takes strictly positive or strictly negative values, do not vanish and, therefore, are strictly positive. Therefore, the polynomial  $p(x) - r(x)$  of degree  $n$ , which is not identically zero, alternately takes strictly positive and strictly negative values at least at  $n + 2$  different points and therefore has at least  $n + 1$  zeros of multiplicity 1. Hence, the polynomial  $n + 2$  satisfying the condition  $\max |\omega(x_k) (f(x_k) - r(x_k))| < \min (|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|)$ , does not exist.

*Comment.* The de la Vallée – Poussin theorem allows one to obtain a lower estimate for solving the optimization problem (2). If  $\lambda_0, \lambda_1, \dots, \lambda_{N-1}$  are local maxima and minima of the  $\omega(x) (f(x) - r(x))$  with alternating signs, then this construction is called the de la Vallée – Poussin alternance, and for solving the optimization problem (2) it gives not only a lower bound, but also an upper bound equal to  $\max \{|\lambda_0|, |\lambda_1|, \dots, |\lambda_{N-1}|\}$ .

**Statement 2.** If the conditions (3), (4) of the Chebyshev criterion are satisfied for the polynomial  $p(x)$ , then the optimal value for the right side of the optimization problem (2) is equal to  $|\varepsilon|$ , and the polynomial  $p(x)$  is its solution (possibly, one of many).

*Proof.* The Chebyshev alternance (3), (4) is a special case of the de la Vallée – Poussin alternance, therefore, according to Statement 1, the right side of the optimization problem (2) is not less than  $|\varepsilon|$ . On the other hand,  $\max |\omega(x) (f(x) - p(x))| = |\varepsilon|$ . Therefore, the minimum of the optimization problem (2) is equal to  $|\varepsilon|$ , and the polynomial  $p(x)$  is one of its solutions.

**Statement 3.** There is a polynomial  $p(x)$  that provides a solution to the optimization problem (2). (This excludes the case when there are polynomials  $p_k(x)$  with constantly decreasing values for  $P_k = \max |\omega(x) (f(x) - p_k(x))|$ , while the minimum of this quantity is never reached on the set of polynomials of a given degree.)

*Proof.* The values of  $P_k = \max |\omega(x) (f(x) - p_k(x))|$  are bounded from below by zero; therefore, for the values  $P_k$  on the set of polynomials, there exists an infimum  $P$ . In accordance with the definition of the infimum, there exists a sequence of polynomials  $p_k(x)$  of degree  $n$ :

$$p_k(x) = a_{0,k} + a_{1,k}x + a_{2,k}x^2 + \dots + a_{n-1,k}x^{n-1} + a_{n,k}x^n,$$

for which  $\lim P_k = P$  as  $k \rightarrow \infty$ . Therefore, starting from a certain number  $k$ , the values of the polynomials  $p_k(x)$  on the interval  $[a + \delta, b - \delta]$ , where  $\delta$  is small enough, will be bounded from above and below (we had to make an indent from the ends of the interval to take into account the case when the weight function

may be zero at the ends of the interval). From the Lagrange formula [1, 2] for a polynomial of degree  $n$ , which at  $n + 1$  given points  $x_k$  takes the values  $y_k$

$$p(x) = \sum_{j=1, n+1} \left( y_j \prod_{i=1, n+1, i \neq j} \frac{(x - x_i)}{(x_j - x_i)} \right),$$

it follows that when at  $n + 1$  fixed points  $x_k$  the values of the polynomial  $y_k$  are bounded above and below, then each of the coefficients of the polynomial is also bounded above and below. Therefore, in accordance with the Bolzano – Weierstrass theorem (also called the Bolzano – Weierstrass lemma) on the limit point, according to which from any infinite bounded sequence of points in the space  $R^n$ , one can select a convergent subsequence, in the sequence of polynomials  $p_k(x)$  one can choose a subsequence of polynomials in which there is a limit for each of the coefficients of the polynomial. That is, there is a sequence of polynomials  $p_k(x)$  for which  $\lim p_k = P$  as  $k \rightarrow \infty$  and the polynomial coefficients have limit values  $b_j = \lim a_{j, k}$  as  $k \rightarrow \infty$ .

Consider a polynomial  $r(x)$  with coefficients  $b_j$ :

$$r(x) = b_0 + b_1x + b_2x^2 + \dots + b_{n-1}x^{n-1} + b_nx^n.$$

Since  $b_j = \lim a_{j, k}$ , then for any fixed value  $x$  the condition  $\lim p_k(x) = r(x)$  is satisfied as  $k \rightarrow \infty$ , and the tendency to the limit is uniform on the considered finite interval of values  $x$ . From inequality

$$\max |\omega(x) (f(x) - r(x))| \leq \max |\omega(x) (f(x) - p_k(x))| + \max |\omega(x)| \max |r(x) - p_k(x)|$$

it follows that  $\max |\omega(x) (f(x) - r(x))| = P$ , because this value cannot be less than  $P$ , while as  $k \rightarrow \infty$  the first term on the right side tends to  $P$ , and the second term tends to zero. This means that for the polynomial  $r(x)$  the value (2) reaches its lower bound.

**Statement 4 (The Chebyshev theorem).** The polynomial  $p(x)$ , which provides a solution to the optimization problem (2), satisfies the Chebyshev criterion.

*Proof.* Consider the behavior of the function  $F(x) = \omega(x) (f(x) - p(x))$  on the segment  $[a, b]$ . To simplify the reasoning, we assume that the number of intervals on which the function  $F(x)$  changes sign is finite. This does not mean that the number of zeros of the function  $F(x)$  is also finite: let's assume the variant is when the function  $F(x)$  is equal to zero at any point of some interval inside the segment  $[a, b]$ . Let the points  $y_1, y_2, \dots, y_m$  divide the segment  $[a, b]$  into  $m + 1$  intervals, on each of which the function  $F(x)$  takes alternately positive and negative values, and let us assume the option when on the entire segment  $[a, b]$  the function  $F(x)$  takes only positive or only negative values. If there is continuous interval consisting entirely of zeros of the function  $F(x)$ , this interval is attached to the

neighboring interval with positive or negative values of the function.

For intervals with positive values of  $F(x)$ , we choose a point with the maximum value of the function on this interval (perhaps not the only one on this interval), and for intervals with negative values of  $F(x)$ , we select the point with the minimum value of the function on this interval. We get a set of points  $y_1, y_2, \dots, y_m$ , — function  $F(x)$  zeros dividing the segment  $[a, b]$  into  $m + 1$  intervals, for which the function  $F(x)$  retains its sign, changing it when crossing the boundary of the interval. Also we get a set of points  $x_0, x_1, x_2, \dots, x_m, x_{m+1}$  belonging these intervals and not coinciding with the ends of the intervals. This set consists of alternating positive maxima and negative minima  $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{m+1}$  of the function  $F(x)$  (Vallée – Poussin alternance) and the accompanying partition of the segment  $[a, b]$  into intervals on which the function  $F(x)$  retains its sign).

Let  $\lambda = \max |\lambda_k|$ . If the maximum is equal to  $\lambda$  on some interval, and the modulus of the minimum is less than  $\lambda$  on the neighboring (next or previous) interval, then this interval, together with the subsequent interval with a positive maximum, regardless of its value, merges the current interval (see Fig. 2, a).

**Fig. 2.** Combining segments with alternating maxima and minima.

a — the case of a single minimum or several successively located minima that lie above the minimum level of the function  $F(x)$ ; б — the case of a single maximum or several successively located maxima that lie below the maximum level of the function  $F(x)$ .

The vertical arrows mark the minima and maxima, which correspond to the maximum deviation. Slanted arrows mark minima and maxima, which are less than the maximum deviation in modulus, so that the graph of the function can be shifted down or up, respectively, by a non-zero distance without going beyond the limits of the maximum deviation. Bold dots mark the zeros of the function  $F(x)$ , which are the boundary of the combined interval with strictly positive or strictly negative maximum deviations of the function  $F(x)$

The new interval has the property that such a positive constant can be subtracted from the function  $F(x)$  so that the positive maximum of the function on the interval decreases, but the absolute value of the negative minimum of the function on the interval does not increase so much as to exceed the new maximum of the function.

In the same way, if on some interval the minimum is equal to  $-\lambda$ , and on the neighboring interval the maximum is less than  $\lambda$ , then this interval, together with the subsequent interval containing a negative minimum, merges the current interval (see Fig. 2, б). In this case, on the obtained interval, a positive con-

stant can be added to the function  $F(x)$  so that the absolute value of the negative minimum on the interval decreases, but the positive maximum of the function on the interval does not increase so much that it becomes greater than the absolute value of the new minimum.

Finally, on the interval  $[a, b]$  for a given polynomial  $p(x)$ , a set of  $N$  intervals with points  $x_0, x_1, x_2, \dots, x_{N-1}$  on them, and the function  $F(x)$  alternating positive maxima and negative minima equal to  $\pm\lambda$ , where  $\lambda = \max |\omega(x) (f(x) - p(x))|$ . If  $N$  is greater than or equal to  $n + 2$ , then such a polynomial  $p(x)$  satisfies the Chebyshev criterion and, in accordance with Statement 1, will be a solution to the optimization problem (2).

Let  $N$  be less than or equal to  $n + 1$ . At the stage of combining the intervals, a set of points  $y_1, y_2, \dots, y_{N-1}$  (interval boundaries), at which the function  $F(x)$  vanishes and changes its sign, was obtained. These points divide the interval  $[a, b]$  into smaller intervals, each of which has a point with a Chebyshev minimum or a Chebyshev maximum. There are also constants  $\delta_k > 0$ , which can be subtracted from the function  $F(x)$  on intervals with positive maxima or added to the function  $F(x)$  on intervals with negative minima, while reducing the minimax norm on the appropriate interval.

Without loss of generality, we can assume that the function  $F(x)$  takes positive values on the first interval; otherwise, you can change the sign of the function  $f(x)$  and the polynomial  $p(x)$  simultaneously without changing the optimization problem (2).

Consider the function  $\Omega(x)$ , which is the product of a polynomial of degree at most  $n$  and a weight function  $\omega(x)$ :

$$\Omega(x) = \omega(x) (y_1 - x) (y_2 - x) \dots (y_{N-1} - x).$$

The function  $\Omega(x)$  is strictly positive on those intervals where there is a Chebyshev maximum of the function  $F(x)$ , and strictly negative on those intervals where there is a Chebyshev minimum of the function  $F(x)$ . If we subtract the function  $s\Omega(x)$  with a sufficiently small positive factor  $s$  (for example, we can choose  $s = (\delta / R)$ , where  $R = \max |\Omega(x)|$  and  $\delta = \min \delta_k > 0$ ) from the function  $F(x)$ , this will safely reduce the positive maxima of  $F(x)$  and decrease the negative minima of  $F(x)$ , thereby reducing  $\max |F(x)|$ . Therefore, such a polynomial  $p(x)$  cannot be a solution for the optimization problem (2), since it is possible to obtain an even smaller value for  $\max |F(x)|$  when replacing the current polynomial  $p(x)$  of degree  $n$  with a new polynomial  $r(x)$  of degree  $n$ :

$$\begin{aligned} G(x) &= F(x) - (\delta / R) \Omega(x) = \\ &= \omega(x) (f(x) - p(x)) - (\delta / R) \omega(x) (y_1 - x) \dots (y_{N-1} - x) = \\ &= \omega(x) (f(x) - p(x) - (\delta / R) (y_1 - x) \dots (y_{N-1} - x)) = \\ &= \omega(x) (f(x) - r(x)), \quad \max |G(x)| < \max |F(x)|. \end{aligned}$$

Conclusion: the optimal polynomial of degree  $n$ , considered in Statement 2, must satisfy the Chebyshev criterion.

**Statement 5.** The polynomial  $p(x)$  that satisfies the Chebyshev criterion and provides a solution to the optimization problem (2) is the only one.

*Proof.* The idea of the proof is borrowed from [7, 8]. Let there be two polynomials  $p(x)$  and  $r(x)$  of degree  $n$ , which are the solution to the optimization problem (2). For the degenerate case, for which  $\max |\omega(x) (f(x) - p(x))| = 0$  and  $\max |\omega(x) (f(x) - r(x))| = 0$ , the conditions  $f(x) - p(x) \equiv 0$  и  $f(x) - r(x) \equiv 0$  must be satisfied, which means that  $p(x) \equiv r(x)$ . Therefore, we will assume that  $\max |\omega(x) (f(x) - p(x))| = \max |\omega(x) (f(x) - r(x))| > 0$  so that the expressions  $\omega(x) (f(x) - p(x))$  and  $\omega(x) (f(x) - r(x))$  on the interval  $[a, b]$  have nonzero maxima and minima.

According to Statement 3, each of the polynomials  $p(x)$  and  $r(x)$  satisfies the Chebyshev criterion, and the value  $\varepsilon \neq 0$  for them is the same. Let  $x_1, x_2, \dots, x_{n+2}$  be Chebyshev deviation points with alternating maxima and minima equal to  $\pm\varepsilon$ , for the function  $F(x) = \omega(x) (f(x) - p(x))$ . If for the polynomial  $r(x)$  the function  $G(x) = \omega(x) (f(x) - r(x))$  takes the same values  $\pm\varepsilon$  at the points  $x_1, x_2, \dots, x_{n+2}$ , except, maybe one point from this list, then the polynomials  $p(x)$  and  $r(x)$  are identically equal to each other. Indeed, in this case, the difference

$$\begin{aligned} d(x) &= F(x) - G(x) = \\ &= \omega(x) (f(x) - p(x)) - \omega(x) (f(x) - r(x)) = \\ &= \omega(x) (r(x) - p(x)), \end{aligned} \tag{5}$$

which is the product of a positive weight function  $\omega(x)$  and a polynomial of degree at most  $n$ , vanishes at  $n + 2$  points or, in extremis, at  $n + 1$  points — and therefore the polynomial  $r(x) - p(x)$  must be identically zero, since otherwise it cannot be explained that it has so many zeros (the weight function  $\omega(x)$ , even if it vanishes at the ends of the interval  $[a, b]$ , is not equal to zero at the points of function  $F(x)$  minima and maxima).

Let among the Chebyshev deviation points  $x_1, x_2, \dots, x_{n+2}$  for the function  $F(x)$  there be at least two points  $x_k$  and  $x_{k+m}$ , at which the value of the function  $G(x)$  does not coincide with the value of the function  $F(x)$  in the point (minimum or maximum of the function  $F(x)$ ). It is required to prove that in this case inside the interval  $[x_k, x_{k+m}]$  there is at least one more point that does not coincide with the points  $x_1, x_2, \dots, x_{n+2}$ , and at which the values of the functions  $F(x)$  and  $G(x)$  coincide (although, perhaps, at this point the common value of the functions  $F(x)$  and  $G(x)$  is not equal to  $\pm\varepsilon$ ), and, consequently, the function  $d(x)$  given by formula (5) has an additional unaccounted zero.

If at some point  $x_j$  the condition  $d(x_j) \neq 0$  is satis-

fied, then the sign of the difference  $d(x_j)$  given by formula (5) coincides with the sign of  $F(x_j)$ . Indeed, if  $x_j$  is the point of the negative minimum of the function  $F(x)$ , then  $G(x_j) \geq F(x_j)$ , and if  $x_j$  is the point of the positive maximum of the function  $F(x)$ , then  $G(x_j) \leq F(x_j)$ . If

$$d(x_k) \neq 0, d(x_{k+1}) = 0, d(x_{k+2}) = 0, \dots, d(x_{k+m-1}) = 0, \\ d(x_{k+m}) \neq 0, \quad (6)$$

then the expressions  $(-1)^k F(x_k)$  and  $(-1)^{k+m} F(x_{k+m})$  have the same signs and, therefore, the expressions  $d(x_k)$  and  $(-1)^m d(x_{k+m})$  have the same signs. Therefore, the parity of the number of sign changes of the function  $d(x)$  on the interval  $[x_k, x_{k+m}]$  coincides with the parity of the number  $m$ , and thus the parity of the number of roots (taking into account their multiplicity) for the function  $d(x)$  on the interval  $[x_k, x_{k+m}]$  coincides with the parity of the number  $m$ . According to (6), the function  $d(x)$  on the interval  $[x_k, x_{k+m}]$  has at least  $m-1$  unequal roots and, therefore, either one more root must exist, or one of the roots  $x_{k+1}, x_{k+2}, \dots, x_{k+m-1}$  has an even multiplicity of at least 2. After performing this operation for all intervals bounded by points  $x_k$  with non-zero values  $d(x_k)$ , we obtain that the number of roots of the function  $d(x)$  on the interval  $[a, b]$ , taking into account the multiplicity of the roots, will be no less than  $n+1$ . Therefore, the polynomial  $r(x) - p(x)$  of degree at most  $n$  must be identically zero in the formula  $d(x) = \omega(x)(r(x) - p(x))$ .

#### ALGORITHM FOR CONSTRUCTING A MINIMAX POLYNOMIAL APPROXIMATION OF A FUNCTION

An algorithm for the numerical construction of approximating polynomials is derived from the Chebyshev criterion for the minimax approximation of a function. This algorithm is a modified and optimized E.Ya. Remez algorithm [18–22]:

1. An initial set of points  $x_0 < x_1 < x_2 < \dots < x_{n+1}$ , on the interval  $[a, b]$  is chosen arbitrarily. If the weight function  $\omega(x)$  vanishes at the point  $x = a$ , it is necessary to satisfy the condition  $a < x_0$ . If the weight function  $\omega(x)$  vanishes at the point  $x = b$ , then the condition  $x_0 < b$  must be met.

2. For the current set of points  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$ , an auxiliary function  $g(x) = a f(x) - p(x)$  is constructed, where  $a$  is some constant, and  $p(x)$  is a polynomial of degree  $n$  with some coefficients. The free coefficients of the function  $g(x)$ , i.e. multiplier  $a$  and coefficients of polynomial  $p(x)$  are chosen so that the conditions  $\omega(x_k)g(x_k) = (-1)^k$  are satisfied for  $k = 0, 1, \dots, n+1$ . Such a function exists and is uniquely defined (see further).

3. Now we need to determine the algebraic equation  $g(x) = 0$  roots, which are on the interval  $[a, b]$ .

Since the values of the function  $g(x)$  have different signs at the edges of the intervals  $[x_k, x_{k+1}]$ , then inside each such interval there is at least one root of the equation  $g(x) = 0$ . If the function  $f(x)$  changes slowly enough, then on each of the intervals  $[x_k, x_{k+1}]$  there is exactly one root of multiplicity 1, and the numerical determination of these roots is not difficult. If this condition is violated and there are several roots on the interval under consideration, all these roots are added to the list (multiple roots with even multiplicity, for which the values of the function  $g(x)$  on both sides of the root have the same sign, are not included in the list). Upon completion of the search for roots, the beginning and the end of the interval  $y_0 = a$  and  $y_{m+1} = b$  are added to the list of values  $a < y_1 < y_2 < \dots < y_m < b$ , which are roots of odd multiplicity of the function  $g(x)$  on the interval  $[a, b]$ .

4. Now the interval  $[a, b]$  is a set of at least  $n+2$  intervals  $[y_k, y_{k+1}]$ , which are in contact with their end points, and on each interval the function  $g(x)$  retains its sign, and the sign of the function changes when the boundary between the intervals is crossed. Consider the function  $G(x) = \omega(x)g(x)$ , which on the interval  $[a, b]$  takes values of the same sign as the function  $g(x)$ . Let us find the global (for the considered interval) extremum  $G(x)$  and the point  $z_k \in [y_k, y_{k+1}]$ , where this extremum is located, for the function  $G(x)$  on each interval  $x \in [y_k, y_{k+1}]$ . For intervals with positive values of the function  $G(x)$ , one should look for a positive maximum, and for intervals  $[y_k, y_{k+1}]$  with negative values of the function — a negative minimum.

5. It is necessary to exclude extra points if the number of determined points  $G_k = G(z_k)$ ,  $z_k \in [y_k, y_{k+1}]$  with alternating signs is greater than  $n+2$ . To do this, the least modulo value of  $G_k$  is chosen. If the interval with this value is the first or last in the list, the value of  $G_k$  and the corresponding point  $z_k$  are discarded. If the interval with this value is in the middle of the list, check the previous and subsequent intervals and discard not only the considered value of  $G_k$ , but also that of the two neighboring values  $G_{k-1}$  and  $G_{k+1}$ , which will be less in absolute value. Indeed, the signs of the values  $G_{k-1}$  and  $G_{k+1}$  are opposite in sign to the value of  $G_k$ , and since due to the removal of the value  $G_k$  from the list, two values of the same sign will be nearby, then one of the two values  $G_{k+1}$  and  $G_{k-1}$  must be discarded. It's important: there is a risk of getting  $n+1$  points from  $n+3$   $z_k$  points, instead of the required  $n+2$  points at once since one or two points are removed at each step, — in this case, you should keep the point  $G_k$  in the list, and delete instead one of the start or end points, which is less in absolute value. The process continues until the list contains  $n+2$  extrema  $G_k = G(z_k)$ , corresponding to  $n+2$  points  $z_k$  with alternating positive maxima and negative minima of the function  $G(x) = \omega(x)g(x)$ .

6. For the existing list of points  $a \leq z_0 < z_1 < z_2 < \dots < z_{n+1} \leq b$ , let's take a closer look at the values of extrema  $G_k = \omega(z_k) g(z_k) = (-1)^k \varepsilon_k$ . If  $\varepsilon_k \approx 1$  within the given accuracy, then a polynomial that satisfies the requirements of the Chebyshev criterion is determined. Moreover, it follows from the de la Vallée – Poussin theorem on the best approximation of a function by polynomials (Statement 1) that the exact optimum for problem (2) lies in the range between  $\min |\varepsilon_k / a|$  and  $\max |\varepsilon_k / a|$ , (where  $a$  is the factor in the expression  $g(x) = a f(x) - p(x)$ ). If the condition  $\varepsilon_k \approx 1$  is not satisfied, we replace the trial points  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$  with points  $a \leq z_0 < z_1 < z_2 < \dots < z_{n+1} \leq b$  and return to step 2.

7. It remains to normalize the function  $g(x) = a f(x) - p(x)$  so that the coefficient of the function  $f(x)$  becomes equal to 1. Such a normalization is always feasible, since the coefficient  $a$  cannot be zero, otherwise it will turn out that there is a non-identical zero polynomial  $p(x)$  of degree  $n$ , which has  $n + 1$  roots located between neighboring maxima and minima of different signs in the amount of  $n + 2$  pieces.

Obviously, if this algorithm converges, then it converges to a polynomial that satisfies the Chebyshev criterion, and thus the result of the algorithm is a solution to the optimization problem (2), and the only such solution. Theoretically, the algorithm always converges at the rate of a geometric progression: for any continuous weight function and for any degree of  $n$ , there are numbers  $C > 0$  and  $0 < \lambda < 1$  for which

$$\max |\omega(x) (a_i f(x) - p_i(x))| \leq 1 + C \lambda^i,$$

where  $i$  is the iteration number. The proof is similar to the proof of the convergence of the E.Ya. Remez algorithm in the monograph [20]. Unfortunately, due to rounding errors, the algorithm may diverge for "bad" functions  $f(x)$  and  $\omega(x)$ , namely, for functions that do not change slowly on the interval  $[a, b]$ , but instead show multiple oscillations, so that they have more than one root for the intervals  $[x_k, x_{k+1}]$  or more than one local maximum or minimum for the intervals  $[y_k, y_{k+1}]$ . In such cases, the algorithm can perform an infinite loop, in fact switching between the same list of tested polynomials. This behavior of the algorithm should be controlled manually by setting the maximum number of cycles that are allowed to run. Another useful criterion for terminating the operation of an algorithm that has fallen into an infinite loop can be, at the next step, to check the increase of the spread between the maxima and minima  $\pm \varepsilon_k$  determined on the intervals  $[y_k, y_{k+1}]$  compared to this value of the previous step. Also, a suspicious attribute, signaling the looping of the algorithm, may be the regular occurrence of a situation, at Step 5, when "cleaning" points  $G_k$ , at some step there are less than  $n + 2$  points in the list and it becomes necessary to take a step

back, removing the first or the last point, because even if the optimal polynomial  $p(x)$  is the identical zero (for example, such a situation arises when trying to approximate with a polynomial of finite degree of the function  $\sin(1/x)$  on the interval  $[0, 1]$ ), then, according to Statement 4, there must still be a set of  $n + 2$  points  $z_k$  that form the Chebyshev alternance.

Since the weight function  $\omega(x)$  can vanish at an end of the interval  $[a, b]$ , then to perform step 2 it will be required that the test points chosen at step 1 lie strictly inside the interval  $[a, b]$ :  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$ . It can be seen that if  $\omega(a) = 0$  or  $\omega(b) = 0$  and,  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$ , then the condition  $a < x_0 < x_1 < x_2 < \dots < x_{n+1} < b$  will be preserved for all subsequent iterations, because the points  $z_k$ , at step 6, will also have to satisfy the condition  $a < z_0 < z_1 < z_2 < \dots < z_{n+1} < b$ .

To calculate the coefficients at step 2, it is impossible to use an explicit solution of the system of linear equations to determine the coefficients of the function  $g(x) = a f(x) - p(x)$ , since the conditionality of the matrix of linear equations (except for the first column, coinciding with the Vandermonde matrix) grows rapidly with the degree of the polynomial: as a result, small errors in rounding floating point numbers lead to completely unsatisfactory values for unknown coefficients. One of the efficient ways to determine the desired  $g$  function  $g(x) = a f(x) - p(x)$  at step 2 is as follows. From the conditions  $p(x_k) = a f(x_k) - (-1)^k / \omega(x_k)$ , which are equivalent to the conditions  $\omega(x_k) (a f(x_k) - p(x_k)) = (-1)^k$ , using the Lagrange formula<sup>2</sup> [1, 2], we determine the coefficients of the polynomial  $p(x)$ , which at given points  $x_k$  takes the given values. These coefficients will be linear expressions with a free coefficient  $a$ . After that, for the equation  $\omega(x_k) (a f(x_k) - p(x_k)) = (-1)^k$ , which is a linear equation with respect to the variable  $a$ , the unknown coefficient  $a$  is determined.

Another, more efficient way to determine the coefficient  $a$  and the coefficients of the polynomial  $p(x)$ , for which the conditions will be satisfied at step 2

$$\omega(x_k) (a f(x_k) - p(x_k)) = (-1)^k \text{ при } k = 0, 1, \dots, n+1,$$

consists of the following. To fulfill the equalities  $p(x_k) = a f(x_k) - (-1)^k / \omega(x_k)$  according to Newton finite difference method [1, 2], we determine such a polynomial  $r(x)$  for which the conditions  $r(x_k) = (-1)^k / \omega(x_k)$  are satisfied if  $k = 0, 1, \dots, n$ , and a polynomial  $s(x)$  for which the conditions  $s(x_k) = f(x_k)$  are satisfied if  $k = 0, 1, \dots, n$ . Then from the condition

<sup>2</sup> Instead of the explicit Lagrange formula, especially for large degrees of polynomials, it is convenient to use the Aitken iterative method [2]. Although this algorithm requires more memory, it provides significantly better coefficient accuracy and requires less computing time.



$$\omega(x_{n+1}) (a \cdot f(x_{n+1}) - a \cdot s(x_{n+1}) + r(x_{n+1})) = (-1)^{n+1}$$

the constant  $a$  is defined. After that, the desired polynomial  $p(x)$  is determined in accordance with the formula  $p(x) = a \cdot s(x) - r(x)$ . This method can also be used in the following form: we calculate a polynomial  $s(x)$  of degree  $n+1$  for which the conditions  $s(x_k) = f(x_k)$  are met at the points  $x_k$  if  $k = 0, 1, \dots, n+1$ , and a polynomial  $r(x)$  of degree  $n+1$  for which the conditions  $r(x_k) = (-1)^k / \omega(x_k)$  are met at the points  $x_k$  if  $k = 0, 1, \dots, n+1$ . It remains to choose the coefficient  $a$  so that the polynomial  $a \cdot s(x) - r(x)$  has degree  $n$ , which is always possible, unless the function  $f(x_k)$  is a polynomial of degree at most  $n$ .

Note that at intermediate stages, there is no need to determine the coefficients of the polynomial  $p(x)$  explicitly, it is enough to be able to calculate its value at an arbitrary point. Therefore, in the case of computer calculation of this algorithm, until the final result is obtained, one can use the calculation of the values of the polynomial  $p(x)$  according to the formula  $p(x) = a \cdot s(x) - r(x)$ , where for the auxiliary polynomials  $s(x)$  and  $r(x)$  the Newton method of finite differences is used.

To improve the operation of the algorithm at the initial stage, at *step 1*, it is recommended to choose the zeros of the Chebyshev polynomial of the first kind of degree  $n+2$ , shifted and scaled from the interval  $[-1, +1]$  to the interval  $[a, b]$  for the initial points  $a \leq x_0 < x_1 < x_2 < \dots < x_{n+1} \leq b$ . At *step 2*, instead of the condition  $\omega(x_k) p(x_k) = (-1)^k$ , it is recommended to use the condition  $\omega(x_k) p(x_k) = (-1)^k (4 / (b-a))^{-n-1}$  to avoid unreasonably large coefficients of the polynomial  $p(x)$  and the function  $f(x)$  (this choice of scale corresponds to the oscillations of the Chebyshev polynomial of degree  $n+1$ , recalculated from the interval  $[-1, +1]$  to the interval  $[a, b]$  and scaled so that the leading coefficient is equal to 1).

**Example.** Let  $f(x) = \exp(x)$ ,  $a = 0$ ,  $b = 1$ ,  $\omega(x) = 1$ . For  $n = 1, 2, \dots, 8$  as a result of the described algorithm, we obtain polynomials that provide the optimal (in terms of minimax norm) approximation of the function  $\exp(x)$  on the interval  $[0, 1]$ . The approximation error depending on the polynomial degree is indicated in Tab. The dependence of the approximation error on the degree of the polynomial is shown in graphical form in Fig. 3.

**Tab.** Error of the function  $\exp(x)$  approximation on the interval  $[0, 1]$  with polynomials of degree  $n$ , optimal in minimax norm with weight  $q(x) = 1$

**Fig. 3.** Approximation of the function  $\exp(x)$  on the interval  $[0, 1]$  with minimax polynomials of degree  $n$ .  $a$  — exponential function and its approximations,  $b$  — difference between the exponential function and its approximations,  $b$  — minimax error on a logarithmic scale

For comparison, the same data for the truncated Taylor series are shown in Fig. 4. It should be noted that the Taylor series for the function  $\exp(x)$  converges very quickly (in contrast to the Taylor series for the function  $\log(x)$ , for example), but even in this case, the minimax approximation gives a significant advantage.

**Fig. 4.** Approximation of the function  $\exp(x)$  on the interval  $[0, 1]$  with a truncated Taylor series of degree  $n$ .  $a$  — exponential function and its approximations,  $b$  — difference between the exponential function and its approximations,  $b$  — minimax error on a logarithmic scale

## REFERENCES

1. Lanczos C. *Applied Analysis*. New York, Prentice Hall Inc., 1956. (Russ. ed.: Lantsosh K. *Prakticheskie metody prikladnogo analiza. Spravochnoe rukovodstvo*. Moscow: GIFML, 1961. 524 p.).
2. Berezin I.S., Zhidkov N.P. *Metody vychislenii, T. 1* [Methods of computing. Vol. 1]. Moscow, GIFML, 1962. 464 p. (In Russ.).
3. Danilov Yu.A. *Mnogochleny Chebysheva* [Chebyshev polynomials]. Minsk, Vyshehishaya shkola, 1984. 157 p. (In Russ.).
4. Paszkowski S. *Numerical Applications of Chebyshev Polynomials*. Warszawa: Państwowe Wydawnictwo Naukowe, 1975. (Russ. ed.: Pashkovskii S. *Vychislitel'nye primeneniya mnogochlenov i ryadov Chebysheva*. Moscow, Nauka, 1983. 384 p.).
5. Gribkova V.P. *Ehffektivnye metody ravnomernykh priblizhenii, osnovannye na polinomakh Chebysheva* [Effective methods of uniform approximation based on the Polynomials of Chebyshev]. Moscow, Sputnik Publ., 2017. 193 p. (In Russ.).
6. Goncharov V.L. *Teoriya interpolirovaniya i priblizheniya funktsii* [Theory of interpolation and approximation of functions]. Moscow, Leningrad, GTTI Publ., 1934. 316 p. (In Russ.).
7. Achiezer N.I. *Lektsii po teorii approksimatsii* [Lectures on approximation theory]. Moscow, Leningrad, OGIZ Publ., 1947. 324 p. (In Russ.).
8. Achiezer N.I. *Lektsii po teorii approksimatsii. 2-e izd* [Lectures on approximation theory, 2nd ed.]. Moscow, Nauka Publ., 1965. 407 p. (In Russ.).

9. Berdnikov A., Solovyev K., Krasnova N., Golovitski A., Syasko M. Algorithm for Constructing the Chebyshev-Type Polynomials and the Chebyshev-Type Approximations with a Given Weight. *Proceedings of 2022 Int. Conference on Electrical Engineering and Photonics (EExPolytech)*, 2022 Oct 21–22, pp. 143–145. DOI: 10.1109/EExPolytech56308.2022
10. Chebyshev P.L. [Theory of mechanisms known by the name of parallelograms]. *P.L. Chebyshev. Izbrannye trudy* [See P.L. Chebyshev. Selected works], Moscow, Izd. Akademii Nauk SSSR, 1955, pp. 611–648. (In Russ.).
11. Chebyshev P.L. [Questions about the best approximations related to approximate representations of functions]. *P.L. Chebyshev. Izbrannye trudy* [See P.L. Chebyshev. Selected works], Moscow, Izd. Akademii Nauk SSSR, 1955, pp. 462–578. (In Russ.).
12. Chebyshev P.L. [On the functions that are the least deviating from zero]. *P.L. Chebyshev. Izbrannye trudy* [See P.L. Chebyshev. Selected works], Moscow, Izd. Akademii Nauk SSSR, 1955, pp. 579–608. (In Russ.).
13. Bernstein S.N. [On the best approximation of continuous functions through polynomials of a given degree]. *Soobshcheniya Khar'kovskogo matematicheskogo obshchestva. Vtoraya ser.* [Transactions of the Kharkov mathematical society. 2 series], 1912, vol. 13, no. 2-3, pp. 49–144. (In Russ.).
14. Bernstein S.N. [On the best approximation of the functions of several variables through polynomials or trigonometric sums]. *Trudy MIAN SSSR* [Proceedings of the Steklov Institute of Mathematics], 1951, vol. 38, pp. 24–29. (In Russ.).
15. De la Vallée-Poussin Ch.J. Sur les polynomes d'approximation et la représentation approchée d'un angle. *Bulletin de la classe des sciences*, Académie royale de Belgique, 1910, no. 12, pp. 808–845.
16. De la Vallée-Poussin Ch.J. *Leçons sur l'approximation des fonctions d'une variable réelle: professées à la Sorbonne*. Paris, Gauthier-Villars, 1919.
17. Butzer P.L., Nessel R.J. Aspects of de La Vallée Poussin's work in approximation and its influence. *Archive for History of Exact Sciences*, 1993, vol. 46, pp. 67–95. DOI: 10.1007/BF00387727
18. Remez E. Sur un procédé convergent d'approximations successives pour déterminer les polynomes d'approximation. *Compt. Rend. Acad. Sci.*, Paris, 1934, vol. 198, pp. 2063–2065.
19. Remez E.Ya. Sur le calcul effectif des polynomes d'approximation de Tschebyscheff. *Compt. Rend. Acad. Sci.*, Paris, 1934, vol. 199, pp. 337–340.
20. Dzyadyk V.K. *Vvedenie v teoriyu ravnomernogo priblizheniya funktsii polinomami* [Introduction to the theory of uniform approximation of functions by polynomials]. Moscow, Nauka Publ., Glavnaya redaktsiya fiziko-matematicheskoi literatury, 1977. 508 p. (In Russ.).
21. Remez E.Ya. *Osnovy chislennykh metodov chebyshevskogo priblizheniya* [Fundamentals of the numerical methods of the Chebyshev approximation]. Kyiv, Naukova Dumka, 1969. 624 p. (In Russ.).
22. Laurent P.-G. *Approximation et optimization*. Paris, Hermann, 1972 (Russ. ed.: Loran P.-Zh. *Approksimatsiya i optimizatsiya*. Moscow: Mir Publ., 1975. 496 p.).
23. *Wolfram Mathematica: naibolee polnaya sistema dlya matematicheskikh i tekhnicheskikh vychislenii* [The world's definitive system for modern technical computing]. URL: <https://www.wolfram.com/mathematica/> (accessed 21.03.2023). (In Russ.).

Contacts: *Berdnikov Aleksandr Sergeevich*,  
asberd@yandex.ru

Article received by the editorial office on 28.03.2023