

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.651

АППРОКСИМАЦИЯ ТАБЛИЧНО ЗАДАНЫХ ФУНКЦИЙ: МНОГОКРИТЕРИАЛЬНЫЙ ПОДХОД¹⁾

© 2023 г. А. П. Нелюбин^{1,*}, В. В. Подиновский^{2,**}

¹ 101990 Москва, М. Харитоньевский пер., 4, ИМАШ РАН, Россия

² 101000 Москва, ул. Мясницкая, 20, НИУ ВШЭ, Россия

*e-mail: nelubin@gmail.com;

**e-mail: podinovski@mail.ru

Поступила в редакцию 08.09.2022 г.
Переработанный вариант 29.09.2022 г.
Принята к публикации 02.02.2023 г.

Развит новый подход к оцениванию параметров аппроксимации, при котором удаленность аппроксимирующей функции от заданного конечного множества точек оценивается векторным критерием, компонентами которого являются модули невязок во всех точках. При помощи этого критерия задается отношение предпочтения в удаленности и лучшей считается аппроксимирующая функция, недоминируемая по такому отношению. Изучена аппроксимация для нескольких отношений предпочтения, в том числе для отношения Парето и отношения, порожденного информацией о равноважности критериев. Рассмотрены вычислительные вопросы и исследованы взаимоотношения введенных аппроксимирующих функций с классическими (получаемыми методами наименьших квадратов, наименьших модулей и наименьшего максимального модуля отклонений). Библ. 15. Фиг. 8.

Ключевые слова: аппроксимация функций, регрессионный анализ, многокритериальный анализ, теория важности критериев.

DOI: 10.31857/S0044466923050174, **EDN:** GFZWfZ

ВВЕДЕНИЕ

Аппроксимация (приближение), т.е. замена одних математических объектов другими, в том или ином смысле близкими к исходным, широко применяется в математике и ее приложениях. В частности, она является основой для аналитического представления табличных данных функцией заданного вида с неизвестными параметрами, значения которых определяются так, чтобы график функций был по возможности максимально близок к заданным точкам согласно выбранной числовой мере близости. Задачи такого оценивания являются базовыми в теории аппроксимации функций и регрессионном анализе [1], [2].

Меры близости в регрессионном анализе достаточно обоснованно (в частности, с привлечением принципа максимума правдоподобия) выбираются в тех случаях, когда распределение вероятностей ошибок измерения подчинено некоторому известному закону, чаще всего нормальному закону, или закону Гаусса (при этом мерой служит сумма квадратов отклонений рассчитанных значений аппроксимирующей функции от табличных), реже — закону Лапласа (и тогда мерой служит сумма модулей таких отклонений).

В общем же случае при решении задач аппроксимации проблема обоснованного выбора меры близости остается нерешенной, на практике ей обычно уделяется недостаточно внимания (нередко о ней вообще не вспоминают) и в качестве меры берут сумму квадратов отклонений, так как это — “общепринятый подход”, для него имеются давно и хорошо разработанный математический аппарат аппроксимации и созданы компьютерные программы. А ведь выбор меры близости самым существенным образом влияет на результат аппроксимации.

В настоящей статье проблема выбора меры близости решается за счет того, что близость оценивается при помощи векторного критерия, состоящего из модулей отклонений значений ап-

¹⁾ Работа выполнена при поддержке Международного центра анализа и выбора решений.

проксимирующей функции от заданных для всех табличных значений аргумента. Ценой общности такого подхода является многозначность получающейся аппроксимирующей функции.

1. МНОГОКРИТЕРИАЛЬНАЯ ЗАДАЧА АППРОКСИМАЦИИ

Исходными данными в задаче аппроксимации являются пары значений независимой переменной x и зависимой переменной y :

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n). \quad (1.1)$$

В регрессионном анализе независимая переменная x именуется фактором, а зависимая переменная y — откликом (наблюдённым значением).

Далее будем полагать, что аппроксимирующая функция (функция регрессии) является линейной:

$$y = ax + b, \quad (1.2)$$

где a и b — коэффициенты аппроксимации, которые нужно оценить с использованием исходных данных (1.1). Для такого оценивания вводятся в рассмотрение модули невязок δ_i , или уклонений теоретических значений зависимой переменной — рассчитанных по формуле (1.2) — от заданных:

$$\delta_1 = ax_1 + b - y_1, \quad \delta_2 = ax_2 + b - y_2, \quad \dots, \quad \delta_n = ax_n + b - y_n. \quad (1.3)$$

Общепринятым при аппроксимации является путь, основанный на введении числовой меры близости аппроксимирующей функции к исходным данным, такой, как сумма квадратов невязок, сумма модулей невязок или максимальный модуль невязок. Наиболее распространённым методом оценивания является метод наименьших квадратов (МНК) (the least squares method — LS), согласно которому значения коэффициентов регрессии определяются из условия:

$$\sum_{i=1}^n \delta_i^2 \rightarrow \min_{a,b}. \quad (1.4)$$

Для расчета значений коэффициентов согласно (1.4) имеются простые формулы [1], [2].

Достаточно распространённым, особенно в последние годы, стал метод наименьших модулей (МНМ) (the least modulus method — LM), согласно которому значения коэффициентов аппроксимации определяются из условия:

$$\sum_{i=1}^n |\delta_i| \rightarrow \min_{a,b}. \quad (1.5)$$

Разработаны специальные алгоритмы нахождения значений коэффициентов согласно (1.5), см. [3].

Метод, основанный на использовании наибольшего модуля невязок, называемый далее *методом наименьшего максимального модуля* (МНММ) (the least maximum modulus method — LMM), по сути, использует меру Чебышёва и предполагает получение значений коэффициентов регрессии из условия

$$\max_{i \in N} |\delta_i| \rightarrow \min_{a,b}, \quad (1.6)$$

где $N = \{1, 2, \dots, n\}$. Отметим, что нелинейные оптимизационные задачи (1.5) и (1.6) за счет введения дополнительных переменных можно свести к эквивалентным задачам линейного программирования [4].

Далее рассматривается альтернативный метод оценивания, учитывающий, что по своей сути задача оценивания параметров регрессии изначально является многокритериальной: критериями f_i , подлежащими минимизации ($i \in N$), служат модули невязок δ_i из (1.3), составляющие векторный критерий $f(a, b) = (f_1(a, b), f_2(a, b), \dots, f_n(a, b))$, где $f_i(a, b) = z_i = |\delta_i|$. Значение векторного критерия для пары (a, b) называется ее векторной оценкой: $z = f(a, b)$.

Пусть на множестве векторных оценок $Z = \text{Re}_+^n = [0, +\infty)^n$ задан строгий частичный порядок P^Γ , где Γ — информация, касающаяся оценивания удаленности: если верно $z' P^\Gamma z$, то при модулях невязок $z'_1, z'_2, \dots, z'_n$ аппроксимирующая функция считается ближе к заданным точкам (1.1), чем

при модулях невязок $z_1'', z_2'', \dots, z_n''$. Отношение P^Γ индуцирует на плоскости Re^2 строгий частичный порядок P_Γ :

$$(a', b') P_\Gamma (a'', b'') \Leftrightarrow f(a', b') P^\Gamma f(a'', b''),$$

имеющий аналогичный смысл: если верно $(a', b') P_\Gamma (a'', b'')$, то при значениях параметров a', b' аппроксимирующая функция считается ближе к заданным точкам (1.1), чем при значениях параметров a'', b'' . В качестве решения задачи оценивания параметров аппроксимации рассматривается множество точек (векторов) $(a, b) \in \text{Re}^2$, которые недоминируемы по отношению P_Γ . (Точка (a^*, b^*) называется недоминируемой по P_Γ , если не существует точки (a, b) такой, что верно $(a, b) P_\Gamma (a^*, b^*)$).

При указанном многокритериальном подходе в качестве решения задачи аппроксимации выделяется не одно оптимальное значение коэффициентов (a^*, b^*) , а множество недоминируемых значений $U^\Gamma \subset \text{Re}^2$. Каждой паре $(a, b) \in U^\Gamma$ соответствует “своя” функция $y = f(x, a, b) = ax + b$ и ее график — прямая (1.2), а множеству U^Γ — множество функций $F^\Gamma = \{f(x|a, b) | (a, b) \in U^\Gamma\}$ и множество их графиков — связка (совокупность) прямых. При этом фиксированному значению аргумента x ставится в соответствие множество чисел $Y^\Gamma(x) = \{y \in \text{Re} | y = ax + b, (a, b) \in U^\Gamma\}$. Иными словами, полученная указанным путем аппроксимирующая функция F^Γ оказывается многозначной. Такие аппроксимирующие функции будем называть *недоминируемыми*, или, более конкретно, *Γ -недоминируемыми*.

Многозначность аппроксимации (множественность выделяемых значений параметров) — плата за то, что при многокритериальной постановке задачи их оценивания не делается никаких дополнительных ограничительных допущений о характере (свойствах) исходных данных (1.1) типа фиксирования вида закона распределения вероятностей ошибок измерений и т.п.

Для снятия неопределенности из-за многозначности аппроксимации (за счет снижения информативности) можно множество $Y^\Gamma(x)$ представлять его центром тяжести (центроидом) $c^\Gamma(x)$. В частности, если $Y^\Gamma(x)$ — отрезок $[\alpha(x), \beta(x)]$, то $c^\Gamma(x)$ — его середина: $c^\Gamma(x) = 1/2(\alpha(x) + \beta(x))$. Если функция $y = c^\Gamma(x)$ оказывается линейной (на интересующем интервале оси Ox), то ее можно принять за однозначное упрощенное решение задачи линейной аппроксимации.

2. О ПОСТРОЕНИИ МНОЖЕСТВ U^Γ И $Y^\Gamma(x)$

Для приближенного построения (конечной аппроксимации) множества U^Γ можно построить “достаточно мелкую” сетку, покрывающую “подходящую” область в Re^2 (она определяется с учетом специфики задачи, т.е. расположения точек из (1.1) на плоскости), рассматривать эту сетку вместо множества Re^2 и выделять те из ее узлов, которые являются недоминируемыми по P_Γ . Подобный подход широко используется в многокритериальной оптимизации, когда множество допустимых вариантов решения является областью в Re^n (см., например, [5]). Он применялся также при построении множества ПН-средних, определяемого в рамках многокритериального подхода как совокупности точек прямой, каждая из которых недоминируема по отношению предпочтения на множестве векторов, компонентами которых служат расстояния от нее до заданных точек [6], [7].

В качестве ориентира при первоначальном определении области поиска в Re^2 можно использовать точки $(a^\bullet, b^\bullet)$, соответствующие прямым, проходящим через 2 или более точек множества (1.1). Авторы работы [8] называют такие точки “узловыми” и доказывают, что линию, определяемую по МНМ, следует искать среди $y = a^\bullet x + b^\bullet$. Также полезными могут оказаться точки $(a^\blacksquare, b^\blacksquare)$, соответствующие прямым, проходящим через 2 или более середин отрезков, соединяющих пары точек множества (1.1). В последующих примерах линии, определяемые по МНММ, оказываются среди $y = a^\blacksquare x + b^\blacksquare$.

Для сокращения области поиска множества U^Γ можно воспользоваться следующим замечанием. Зафиксируем значение параметра $a = a^0$ и рассмотрим локальную задачу поиска точек $(a^0, b) \in U^\Gamma(a^0)$, недоминируемых по P_Γ на множестве $\{(a, b) | a = a^0, b \in \text{Re}\}$. Обозначим $b_i = y_i - a^0 x_i$ для всех $i \in N$ и преобразуем выражение для уклонений (1.3):

$$\delta_i = a^0 x_i + b - y_i = b - b_i.$$

В результате получаем задачу определения множества ПН-средних $G^\Gamma(B(a^0))$ среди множества вещественных чисел $B(a^0) = \{b_1, b_2, \dots, b_n\}$ (см. [6], [7]). Если точка b не входит в множество средних $G^\Gamma(B(a^0))$, то это означает, что точка (a^0, b) оказывается доминируемой по P_Γ на рассматриваемом множестве $\{(a, b) | a = a^0, b \in \text{Re}\}$, а значит, и на всем множестве Re^2 . Следовательно, должно выполняться включение

$$\{(a, b) \in U^\Gamma | a = a^0\} \subseteq \{(a, b) | a = a^0, b \in G^\Gamma(B(a^0))\}.$$

В ходе перебора точек методом построения сетки, это замечание можно использовать следующим образом. Для каждого перебираемого значения a сначала можно определить ПН-средние $G^\Gamma(B(a))$, а затем перебирать только значения b из этого множества. Это может быть удобно, так как для некоторых типов информации Γ существуют быстрые аналитические методы определения ПН-средних.

Для произвольного фиксированного значения x^0 множество $Y^\Gamma(x^0)$ графически представляет сечение вертикальной прямой $x = x^0$ связки прямых F^Γ . Для построения множеств $Y^\Gamma(x)$ можно использовать также графоаналитический метод, не требующий построения связки прямых и основанный на следующих соображениях. Произвольной точке (x^0, y^0) пространства Re^2 данных (x, y) соответствует прямая $b = -x^0a + y^0$ в пространстве Re^2 параметров (a, b) . Если эта прямая содержит хотя бы одну точку $(a^*, b^*) \in U^\Gamma$, то точке (x^0, y^0) соответствует, по крайней мере, одна прямая из связки аппроксимирующих прямых F^Γ . Фиксируем $x = x^0$ и будем менять y . Тогда соответствующая прямая $b = -x^0a + y^0$ будет сдвигаться параллельно самой себе: $b = -x^0a + y$, т.е. вертикальной прямой $x = x^0$ в пространстве Re^2 данных (x, y) соответствует множество прямых $b = -x^0a + y$ в пространстве Re^2 параметров (a, b) с одинаковым угловым коэффициентом $-x^0$.

Следовательно, для построения $Y^\Gamma(x^0)$ можно действовать так. Вначале следует построить множество U^Γ или его конечную аппроксимацию в системе координат параметров aOb . Далее нужно взять (начертить) прямую с наклоном $-x^0$. Параллельно сдвигая эту прямую (т.е. сохраняя угол наклона), нужно определить все те положения, в которых прямая $b = -x^0a + y$ будет пересекать множество U^Γ . Соответствующие этим положениям значения b на оси ординат Ob составят множество $Y^\Gamma(x^0)$. В частности, если U^Γ — компактное (т.е. замкнутое и ограниченное) связное множество, то $Y^\Gamma(x^0)$ — отрезок с концами $y_{\min}$ и $y_{\max}$, соответствующими крайним (верхнему и нижнему) положениям передвигаемой прямой: $y_{\min} = b_{\min}$ и $y_{\max} = b_{\max}$, ибо линейное отображение $b = -x^0a + y^0$ является непрерывным и поэтому, как известно из курса анализа, переводит компактное связное множество в Re^2 в компактное и связное множество в Re .

Примеры построения множеств U^Γ и $Y^\Gamma(x)$ приводятся ниже. В них для конечной аппроксимации множеств U^Γ и $Y^\Gamma(x)$ использовалась специально разработанная компьютерная программа.

3. АППРОКСИМАЦИЯ ПРИ ОТСУТСТВИИ ИНФОРМАЦИИ Γ

Рассмотрим вначале случай, когда никакой информации Γ нет: $\Gamma = \emptyset$. Здесь в роли частично-го порядка P^Γ выступает отношение Парето $P^\emptyset$, которое определяется (с учетом желательности минимизации критериев) так [9]:

$$z' P^\emptyset z'' \Leftrightarrow z'_i \leq z''_i, \quad i = 1, 2, \dots, n, \quad z' \neq z''. \quad (3.1)$$

Структура множества $U^\emptyset$ исследовалась в [10]. Было установлено, что оно компактно (замкнуто и ограничено), связно, возможно, не выпукло и является объединением конечного числа политопов (выпуклых оболочек конечных наборов двумерных векторов). Из этого свойства, как указывалось выше, следует, что множества $Y^\emptyset(x)$ — отрезки.

Изложенные теоретические положения иллюстрирует следующий числовой пример.

Пример 1. Пусть исходные данные (1.1) заданы в табл. 1.

Таблица 1. Исходные данные для задач 1–3

x_i	1	2	3	4	5	7
y_i	3	2	5	5	4	7

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

$$f^{LS}(x) = \frac{47}{70}x + 1\frac{61}{70} \approx 0.671x + 1.871, \quad (3.2)$$

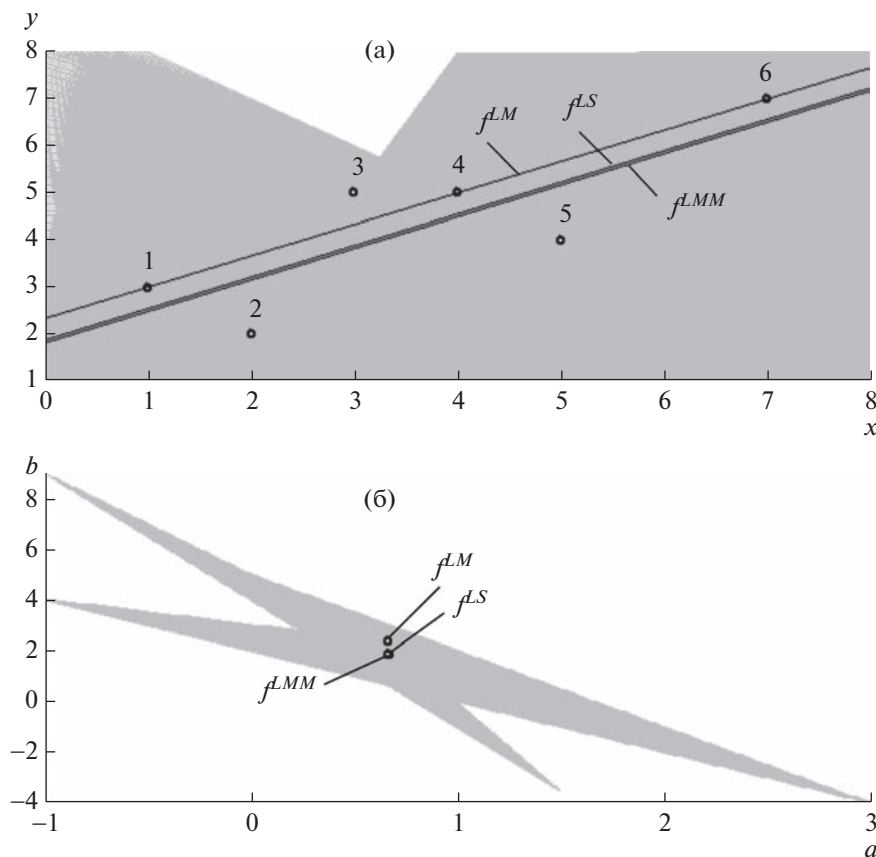
$$f^{LM}(x) = \frac{2}{3}x + 2\frac{1}{3} \approx 0.667x + 2.333, \quad (3.3)$$

$$f^{LMM}(x) = \frac{2}{3}x + 1\frac{5}{6} \approx 0.667x + 1.833. \quad (3.4)$$

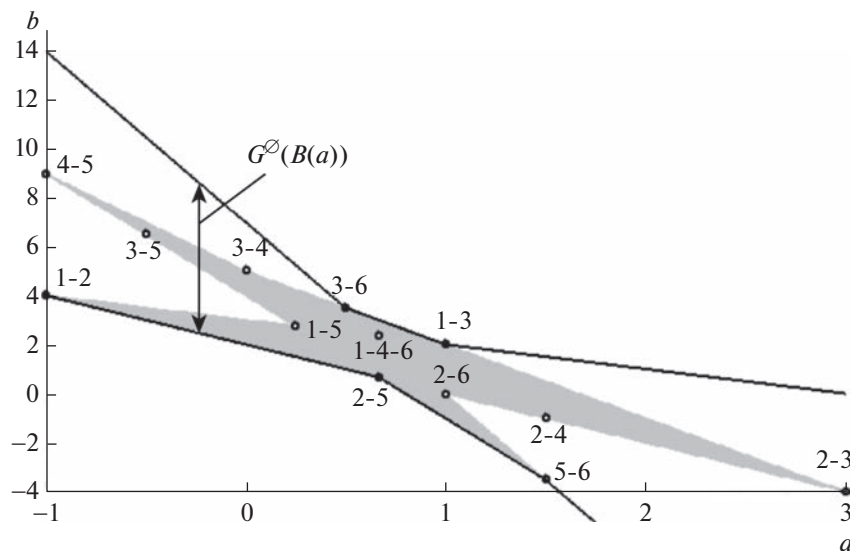
В этом примере при расчетах для аппроксимации множества $U^\varnothing$ использовалась сетка с шагом 0.005 по параметру a и с шагом 0.05 по параметру b , покрывающая область $[-1, 3] \times [-4, 9]$ для параметров (a, b) . Чтобы избежать вычислительных ошибок при сравнении чисел с плавающей точкой, при вычислениях использовались только рациональные числа, представленные целыми числителем и знаменателем. Например, $a = \frac{\hat{a}}{\check{a}}$, где числитель $\hat{a}$ и знаменатель $\check{a} > 0$ — целые числа. В таком представлении числитель и знаменатель уклонений (1.3) рассчитываются по следующим формулам:

$$\hat{\delta}_i = \hat{a}\hat{x}_i\check{b}\check{y}_i + \hat{b}\check{a}\check{x}_i\check{y}_i - \hat{y}_i\check{a}\check{x}_i\check{b}, \quad \check{\delta}_i = \check{a}\check{b}\check{x}_i\check{y}_i.$$

Связка аппроксимирующих функций представлена в области $[0, 8] \times [1, 8]$ для (x, y) на фиг. 1 сверху в виде затемненной области, а множество точек $U^\varnothing$ представлено в области $[-1, 3] \times [-4, 9]$ для (a, b) на фиг. 1б.



Фиг. 1. Связка прямых и множество $U^\varnothing$.



Фиг. 2. Множество U^O , графики границ отрезков $G^O(B(a))$, а также узловые точки.

Для информации $\Gamma = \emptyset$ множество ПН-средних $G^O(B(a))$ (множество значений функций, реализующих среднее по Коши) определяется легко: это отрезок $[\min_{i \in N} b_i, \max_{i \in N} b_i]$. Графики нижней и верхней границы этого отрезка в зависимости от параметра a представлены на фиг. 2. Также на фиг. 2 изображены “узловые” точки (a^*, b^*) , соответствующие прямым, проходящим через 2 или более точек множества (1.1) (номера $i \in N$ этих точек указаны рядом через дефис, например, точке 1–4–6 соответствует линия, определяемая по МНМ). Очевидно, что эти точки (a^*, b^*) недоминируемы по Парето, так как для каждой такой точки (и только для нее) все соответствующие отклонения $\delta_i = 0$.

4. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ

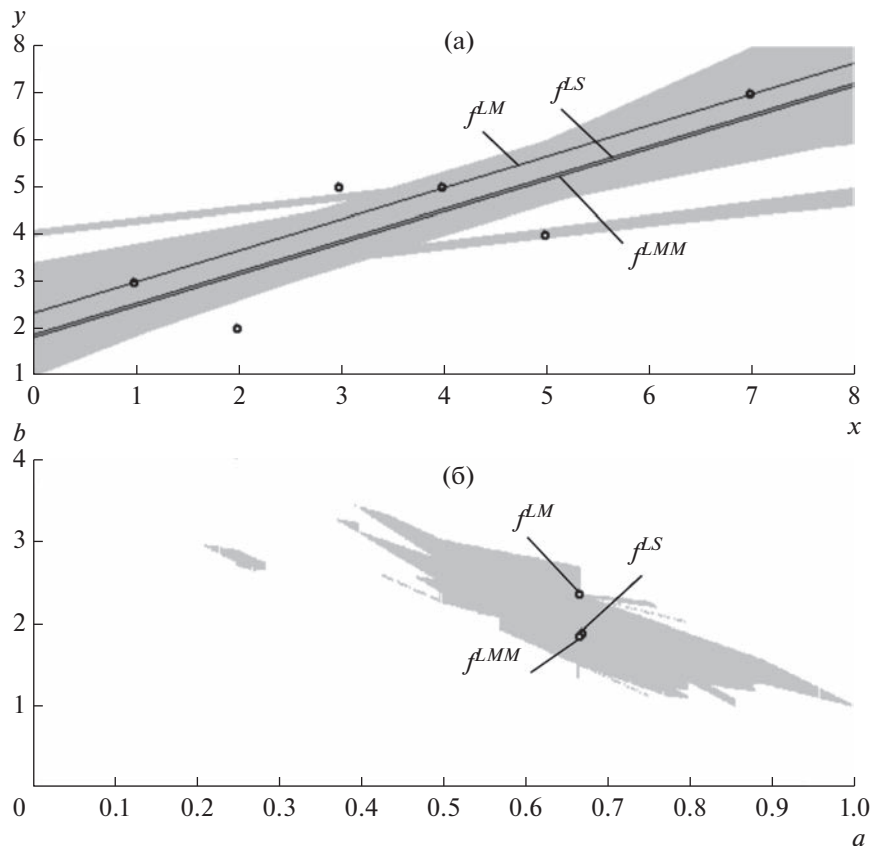
Пусть теперь все критерии имеют равную важность (см. [11], [12]): информация $\Gamma = E$. Содержательно равноважность критериев означает, что после произвольной перестановки компонент векторной оценки z получится векторная оценка, эквивалентная (равная по предпочтительности) исходной. В рассматриваемой постановке задачи аппроксимации все точки (1.1) одинаково существенны (“равноправны”) и поэтому критерии f_i равноважны.

Обозначим через $z_{\setminus} = (z_{[1]}, z_{[2]}, \dots, z_{[n]})$ вектор, полученный упорядочением по невозрастанию компонент вектора z : $z_{[1]} \geq z_{[2]} \geq \dots \geq z_{[n]}$. Отношение P^E определяется следующим аналитическим решающим правилом (см. [11]):

$$z' P^E z'' \Leftrightarrow z'_{[1]} \leq z''_{[1]}, \quad z'_{[2]} \leq z''_{[2]}, \quad \dots, \quad z'_{[n]} \leq z''_{[n]}, \quad z'_{\setminus} \neq z''_{\setminus}.$$

Поскольку $P^E \supset P^O$, то имеет место включение $U^E \subseteq U^O$. Оказывается, что здесь множество U^E может быть несвязным, и тогда связка аппроксимирующих прямых расщепляется на несколько частей. (Несвязность U^E – аналог несвязности средних по P_E [6].) Эту возможность иллюстрирует следующий пример.

Пример 2. Пусть исходные данные (1.1) заданы в табл. 1. При расчетах для аппроксимации множества U^E использовалась сетка с шагом 0.001 по параметру a и с шагом 0.01 по параметру b , покрывающая область $[0, 1.1] \times [0, 5]$ для параметров (a, b) . На фиг. 3 представлены связка графиков аппроксимирующих функций и множество U^E , а также графики функций (3.2)–(3.4) и исходные точки. Фигура 3 показывает, что множество U^E не связно, и из-за этого связка линий аппроксимации расщепилась на несколько частей, а множества $Y^E(x)$ не являются числовыми промежутками.



Фиг. 3. Связка прямых и множество U^E .

5. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ И ОТСУТСТВИИ КОМПЕНСАЦИИ

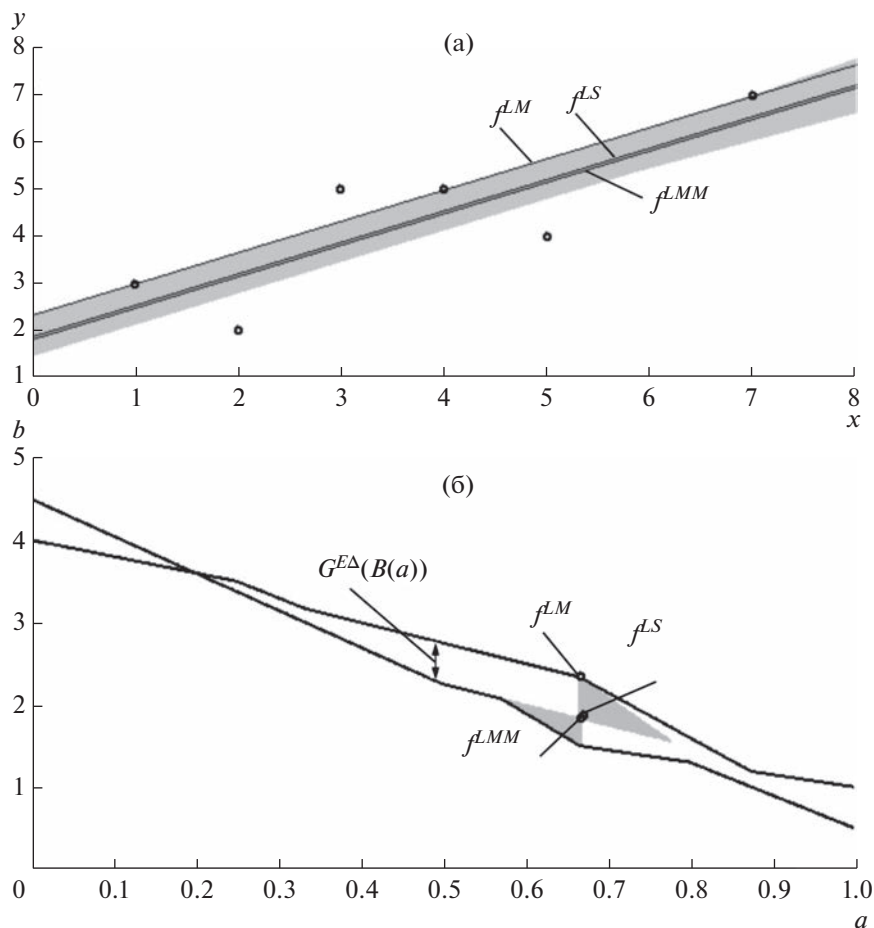
Рассмотрим случай, когда увеличение большего уклонения аппроксимирующей функции в одной из точек x_i не компенсируется уменьшением меньшего уклонения в другой такой точке. Такой подход можно рассматривать как уточнение МНММ (см. (1.6)). Для такой информации EL лексикографического типа (относительно упорядоченных по невозрастанию компонент векторных оценок) отношение P^{EL} определяется следующим решающим правилом:

$$z' P^{EL} z'' \Leftrightarrow [(z'_{[1]} < z''_{[1]}) \vee (z'_{[1]} = z''_{[1]}, z'_{[2]} < z''_{[2]}) \vee \dots \\ \dots \vee (z'_{[i]} = z''_{[i]}, i = 1, 2, \dots, n-1, z'_{[n]} < z''_{[n]})].$$

Поскольку $P^{EL} \supset P^E$, то $U^{EL} \subseteq U^E$.

6. АППРОКСИМАЦИЯ ПРИ РАВНОВАЖНЫХ КРИТЕРИЯХ СО ШКАЛОЙ ПЕРВОЙ ПОРЯДКОВОЙ МЕТРИКИ

Пусть, наконец, все критерии имеют равную важность и шкала критериев является шкалой первой порядковой метрики (см. [12], [13]): информация $\Gamma = E\Delta$. Такая шкала обладает следующим свойством: если в произвольной векторной оценке z , в которой $z_i > z_j$, заменить z_i на $z_i - q$, а z_j — на $z_j + q$, но при этом обеспечить выполнение неравенства $z_i - q \geq z_j + q$, где q — положительное число, то полученная таким образом векторная оценка будет более предпочтительна, чем исходная. Это означает, что при указанном переходе большее отклонение от одной из точек выборки уменьшается за счет равного по величине увеличения меньшего отклонения от другой точки выборки, так что набор отклонений становится более близким к идеальному набору из минимально возможных равных уклонений. Такой подход является аналогом известного в теории обществен-



Фиг. 4. Связка прямых, множество $U^{E\Delta}$ и графики границ отрезков $G^{E\Delta}(B(a))$.

ного благосостояния принципа Пигу-Дальтона: “Общество станет лучше, если единица дохода будет передана от более богатого индивида более бедному” [14].

В рассматриваемом случае имеем частичный порядок $P^{E\Delta}$, задаваемый следующим решающим правилом [12], [15]:

$$\begin{aligned} z' P^{E\Delta} z'' &\Leftrightarrow z'_{[1]} \leq z''_{[1]}, \quad z'_{[1]} + z'_{[2]} \leq z''_{[1]} + z''_{[2]}, \quad \dots \\ \dots, \quad z'_{[i]} + z'_{[2]} + \dots + z'_{[n]} &\leq z''_{[i]} + z''_{[2]} + \dots + z''_{[n]}, \quad z'_\infty \neq z''_\infty. \end{aligned}$$

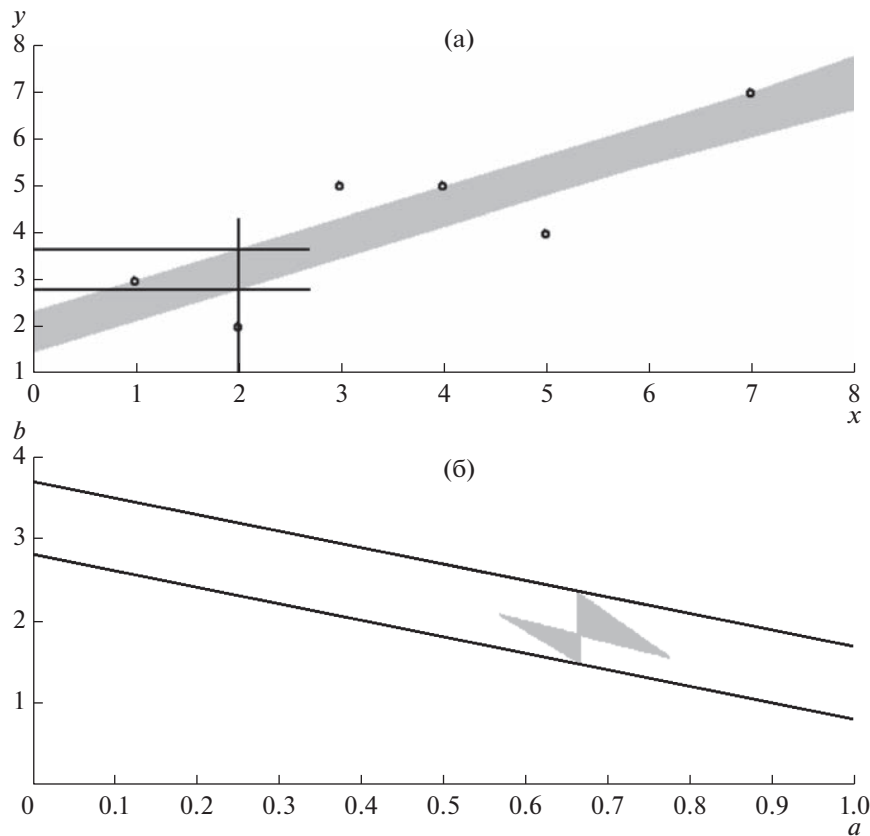
Поскольку $P^{E\Delta} \supset P^E$, имеет место включение $U^{E\Delta} \subseteq U^E$.

Пример 3. Пусть исходные данные (1.1) заданы в табл. 1. На фиг. 4 представлены связки линий аппроксимации и множество $U^{E\Delta}$, а также линии регрессии (3.2)–(3.4) и исходные точки.

ПН-средние $G^{E\Delta}(B(a))$ представляют собой отрезок $[\alpha, \beta]$, границы которого легко вычисляются по формулам (см. [7]):

$$\alpha = \frac{1}{2} \min_{p \in H} (b_{(p)} + b_{(n+1-p)}), \quad \beta = \frac{1}{2} \max_{p \in H} (b_{(p)} + b_{(n+1-p)}),$$

где $H = \{1, 2, \dots, h\}$, а h — целая часть числа $(n+1)/2$. Еще одно полезное свойство ПН-средних при информации $E\Delta$ — устойчивость к изменению положения точек множества $B(a)$. Поскольку при небольшом изменении параметра a на $a + \varepsilon$, положение точек b_i множества $B(a + \varepsilon)$ меняется относительно точек $B(a)$ на величину εx_i , то, согласно [7], максимальное изменение границ множества $G^{E\Delta}(B(a + \varepsilon))$ по сравнению с $G^{E\Delta}(B(a))$ не превышает $\varepsilon \max_{i \in N} |x_i|$. Из этого следует, что



Фиг. 5. Графоаналитический метод определения $Y^{E\Delta}(x)$ при $x = 2$.

функции границ $\alpha(a)$ и $\beta(a)$ должны быть непрерывными. Графики этих функций приведены на фиг. 4.

На фиг. 4 множество $U^{E\Delta}$ связно, и поэтому множества $Y^{E\Delta}(x)$ являются числовыми промежутками.

Рассмотрим графоаналитический метод определения $Y^{E\Delta}(x)$. На фиг. 5б представлено множество $U^{E\Delta}$ и графики прямой $b = -x^0 a + y$ для точки $x^0 = 2$ в двух крайних положениях, которые определяют приближенные значения $b_{\min} = 2.85$ и $b_{\max} = 3.7$, которым соответствует $Y^{E\Delta}(2) = [2.85, 3.7]$. Это согласуется с результатом, который можно получить при помощи фиг. 5а, если построить вертикальную прямую $x = 2$ и рассечь ею связку прямых.

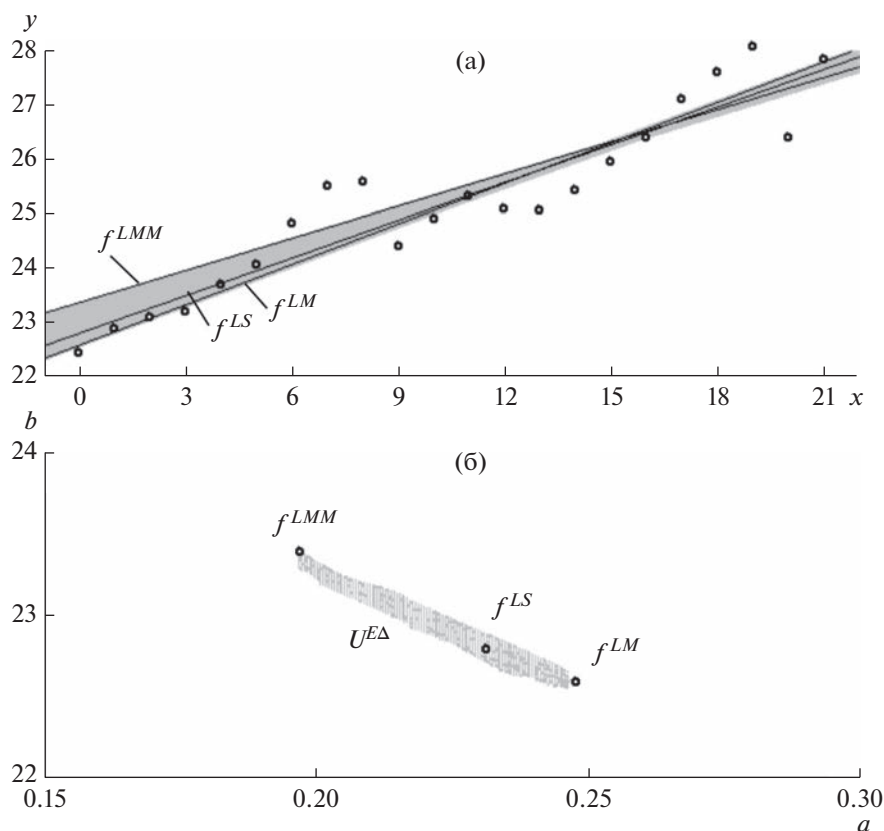
Рассмотрим более сложный пример с реальными данными.

Пример 4. В табл. 2 представлены исходные данные (1.1) по ВВП на душу населения в странах Евросоюза за 2000–2021 гг. [16]. Нумерация годов x_i для удобства начинается с 0, т.е. $x_1 = 0$ соответствует 2000 г. Величины y_i заданы в тысячах евро на человека.

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

Таблица 2. Исходные данные по ВВП на душу населения в Евросоюзе

x_i	0	1	2	3	4	5	6	7	8	9	10
y_i	22.45	22.90	23.09	23.21	23.71	24.06	24.82	25.51	25.58	24.41	24.90
x_i	11	12	13	14	15	16	17	18	19	20	21
y_i	25.32	25.10	25.06	25.42	25.95	26.41	27.11	27.62	28.07	26.39	27.83



Фиг. 6. Связка прямых и множество $U^{E\Delta}$.

$$f^{LS}(x) \approx 0.2313x + 22.7953;$$

$$f^{LM}(x) \approx 0.2478x + 22.5944;$$

$$f^{LMM}(x) \approx 0.197x + 23.3885.$$

При расчетах для аппроксимации множества $U^{E\Delta}$ использовалась сетка с шагом 0.0005 по параметру a и с шагом 0.005 по параметру b , покрывающая область $[0.15, 0.3] \times [22, 24]$ для параметров (a, b) . Связка аппроксимирующих функций представлена в области $[-1, 22] \times [22, 28]$ для (x, y) на фиг. 6а, а множество точек $U^{E\Delta}$ представлено в области $[0.15, 0.3] \times [22, 24]$ для (a, b) на фиг. 6б.

По разбросу значений параметра a примерно от 0.2 до 0.25 можно судить о ежегодном среднем приращении показателя ВВП на данном периоде времени.

7. О ВЗАИМОСВЯЗИ НЕДОМИНИРУЕМЫХ И КЛАССИЧЕСКИХ АППРОКСИМИРУЮЩИХ ФУНКЦИЙ

Поскольку информация $E\Delta$ уточняет информацию E , то, как уже указывалось выше, верно $P^\emptyset \subset P^E \subset P^{E\Delta}$, и поэтому имеют место включения $U^\emptyset \supseteq U^E \supseteq U^{E\Delta}$, а соответствующие связи аппроксимирующих функций последовательно входят друг в друга. Это хорошо видно при сравнении фиг. 1, 3 и 4.

Интересен также вопрос о том, всегда ли графики функций f^{LS} , f^{LM} и f^{LMM} , полученные классическими методами — МНК, МНМ и МНММ — входят в связи прямых для $\Gamma = \emptyset$, $\Gamma = E$ и $\Gamma = E\Delta$, т.е. когда справедливы включения $f^{LS} \in F^\emptyset$, $f^{LM} \in F^E$ и $f^{LMM} \in F^{E\Delta}$.

Ответ на этот вопрос дает

Теорема. Справедливы соотношения $f^{LS} \in F^{E\Delta} \subseteq F^E \subseteq F^\emptyset$, $f^{LM} \in F^E \subseteq F^\emptyset$ и $f^{LMM} \in F^\emptyset$. Если функция f^{LM} (соответственно функция f^{LMM}) единственна, то $f^{LM} \in F^{E\Delta}$ (соответственно $f^{LMM} \in F^E$).

Доказательство. Вначале заметим, что функция $g(z)$ называется убывающей (невозрастающей) по P^Γ (на Z), если из $z' P^\Gamma z''$ следует $f(z') < f(z'')$ (соответственно $f(z') \leq f(z'')$). Известно следующее утверждение о монотонных функциях (см., например, [9, лемма 1.2.2]), которое применительно к рассматриваемому случаю можно сформулировать в следующем виде.

Утверждение. Пусть $Y \subseteq Z$ и невозрастающая по P^Γ функция $g(z)$ достигает минимума на Y в точке z^0 . Для того чтобы точка z^0 была недоминируемой по P^Γ , достаточно выполнения одного из следующих условий: $g(z)$ возрастает по P^Γ ; z^0 – единственная точка минимума $g(z)$ на Y .

Рассмотрим функцию $g(z|s) = \sum_{i=1}^n z_i^s$, где $s > 0$, $z \in \text{Re}_+^n = [0, +\infty)^n$. Очевидно, что из $z' P^\emptyset z''$, как и из $z' P^E z''$, следует $g(z'|s) < g(z''|s)$. Из свойств функций, выпуклых по Шуру (с учетом того, что меньшие значения компонент z_i предпочтительнее больших) [15], следует, что при $s > 1$ из $z' P^E z''$, следует $g(z'|s) < g(z''|s)$, а при $s = 1$ лишь $g(z|s) \leq g(z''|s)$. Следовательно, поскольку минимизация $g(z|1)$ при $z_i = |\delta_i|$ и $g(z|2)$ при $z_i = \delta_i$ лежит в основе МНК и МНМ, то $f^{LS} \in F^{E\Delta}$ и $f^{LM} \in F^E$; а если f^{LM} единственна, то $f^{LM} \in F^{E\Delta}$. Функция $g^-(z) = \max_{i \in N} z_i$ при $z_i = |\delta_i|$ не возрастает по $P^\emptyset$, а потому и по P^E , и по $P^{E\Delta}$. Следовательно, если она единственна, то $f^{LMM} \in F^{E\Delta}$. Теорема доказана.

Теорему иллюстрируют примеры 5–7 с “нетипичными” исходными данными.

Пример 5. Исходные данные (1.1) заданы табл. 3.

Здесь уравнения аппроксимирующих функций, определяемые МНК, МНМ и МНММ, соответственно, таковы:

$$f^{LS}(x) = 1.8x - 2.6; \quad f^{LM}(x) \equiv 1; \quad f^{LMM}(x) = 2.25x - 4.625.$$

Связка прямых для $F^{E\Delta}$, а также три прямые – графики всех аппроксимирующих функций – представлены на фиг. 7. Все функции f^{LS} , f^{LMM} и f^{LM} единственны. Все три соответствующие прямые входят в связку прямых, порождаемых недоминируемыми аппроксимирующими функциями, и соответствуют вершинам множества $U^{E\Delta}$.

Пример 6. Исходные данные (1.1) заданы табл. 4.

Здесь множество $F^{E\Delta}$ одноэлементно: $F^{E\Delta} = \{y = 2\}$, и

$$f^{LM}(x) \equiv l, \text{ где } l \in [1, 3], \text{ а также } f^{LM}(x) = x \text{ и } f^{LM}(x) = 5 - x; f^{LS}(x) \equiv f^{LMM}(x) \equiv 2.$$

Функции f^{LS} и f^{LMM} единственны и равны, а f^{LM} не единственна. Справедливы соотношения $f^{LS} \in F^{E\Delta}$, $f^{LMM} \in F^{E\Delta}$, и в $F^{E\Delta}$ входит только $f^{LM}(x) = l$ при $l = 2$.

Пример 7. Исходные данные (1.1) заданы табл. 5.

Множество $F^{E\Delta}$ одноэлементно: $F^{E\Delta} = \{y = 2\}$, и

$$f^{LS}(x) \equiv 2; f^{LM}(x) = kx + 2(1 - k), \text{ где } k \in [-1, 1]; f^{LMM}(x) = kx + 2 - k, \text{ где } k \in [-1, 1].$$

Таблица 3. Исходные данные для задачи 5

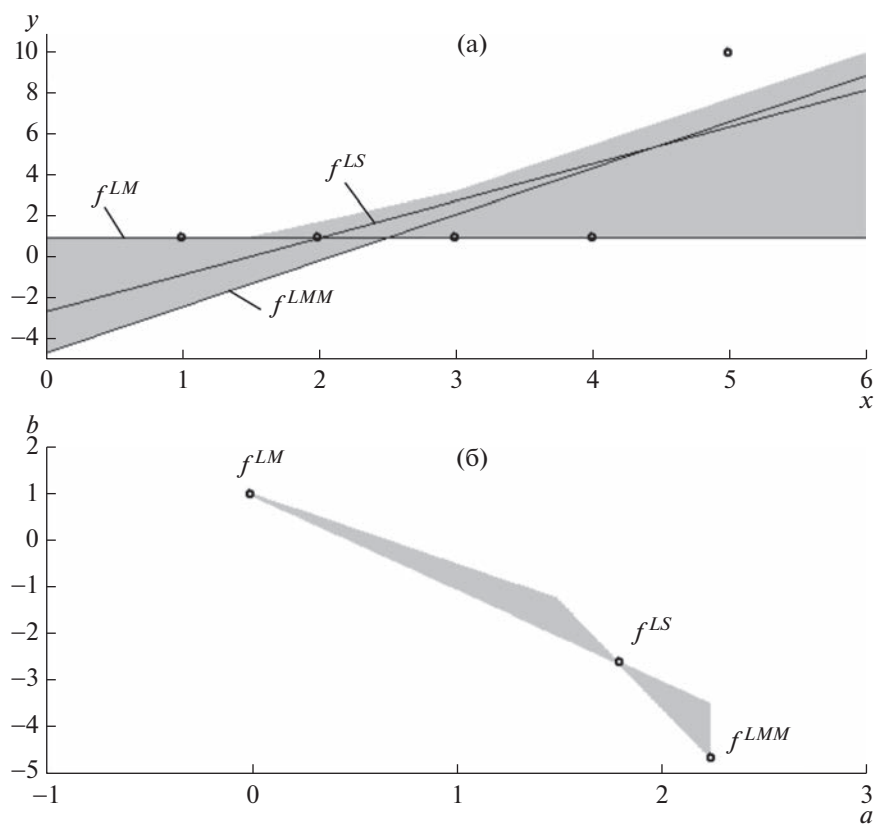
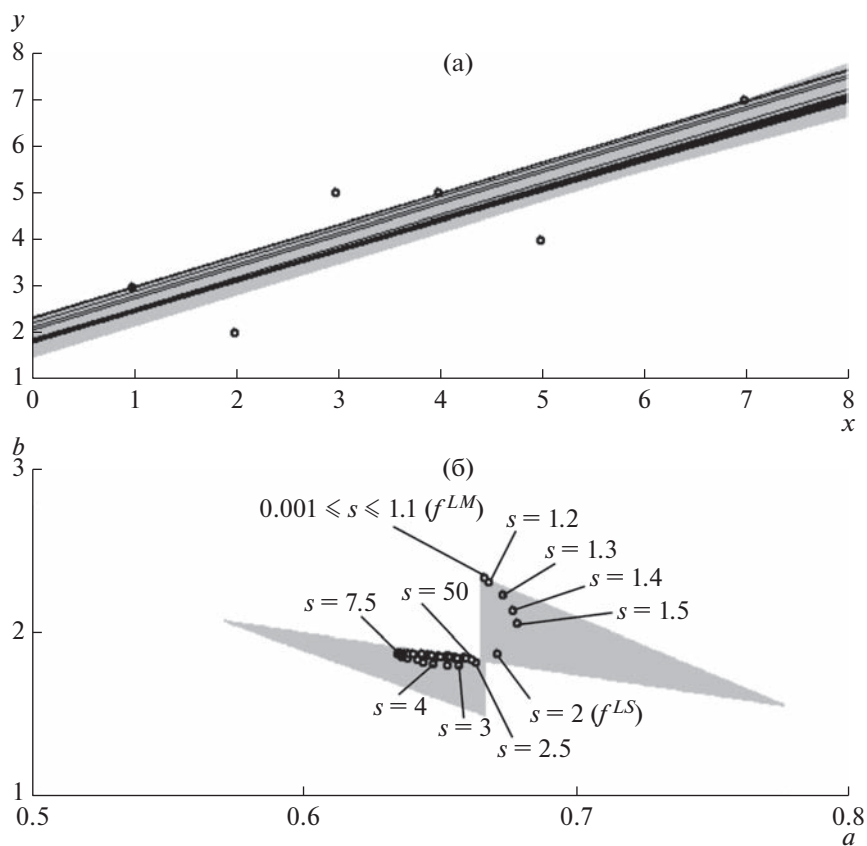
x_i	1	2	3	4	5
y_i	1	1	1	1	10

Таблица 4. Исходные данные для задачи 6

x_i	1	2	3	4
y_i	1	3	3	1

Таблица 5. Исходные данные для задачи 7

x_i	1	2	1
y_i	1	2	3

Фиг. 7. Связка прямых и множество $U^{E\Delta}$.Фиг. 8. Связка прямых, множество $U^{E\Delta}$ и решения, определяемые (7.1).

Функции f^{LM} и f^{LMM} не единственны. Здесь $f^{\text{LS}} \in F^{\text{EA}}$ и лишь $f^{\text{LM}}(x) \equiv 2$ и $f^{\text{LMM}}(x) \equiv 2$ входят в F^{EA} (т.е. при $k = 0$).

Рассмотрим еще один пример, связанный с теоремой и ее доказательством.

Пример 8. Рассмотрим, как расположены решения (a, b) , определяемые условиями

$$\sum_{i=1}^n |\delta_i^s| \rightarrow \min_{a,b} \quad (7.1)$$

при различных значениях показателя $s > 0$ для исходных данных (1.1) из табл. 1. Используем сетку с шагом 0.0005 по параметру a и с шагом 0.005 по параметру b , покрывая область $[0.5, 0.8] \times [1, 3]$. Получившиеся точки оказываются внутри множества U^{EA} , как показано на фиг. 8 (ср. с фиг. 4). Это – иллюстрация соответствующего утверждения из доказательства теоремы.

8. ЗАКЛЮЧЕНИЕ

Множественный подход к определению параметров аппроксимации выделяет не одну прямую линию аппроксимации, а связку таких линий. Множественность выделяемых значений параметров аппроксимации – плата за то, что при многокритериальной постановке задачи их оценивания не делается никаких дополнительных ограничительных допущений о характере (свойствах) исходных данных (например, о законе распределения их вероятностей). Получающуюся при предлагаемом подходе многозначную аппроксимирующую функцию можно интерпретировать двояко: либо как совокупность возможных однозначных функций аппроксимации, либо как множество функций, характеризующих аппроксимацию в целом.

За единственную линию аппроксимации, особенно в случае равноважных критериев со шкалой первой порядковой метрики, можно взять геометрическое место точек – центров тяжести (центроидов) множеств $Y^{\Gamma}(x)$.

Актуальной остается проблема разработки эффективных вычислительных методов построения множеств недоминируемых значений параметров аппроксимации U^{Γ} , а также множеств $Y^{\Gamma}(x)$. Такие методы должны учитывать специфические особенности отношений P_{Γ} для различных видов информации Γ .

Предложенный подход очевидным образом обобщается на случай аппроксимации функций нескольких переменных, но вычислительные проблемы из-за повышения размерности при этом существенно усложняются.

СПИСОК ЛИТЕРАТУРЫ

1. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. Приближение функций, дифференциальные и интегральные уравнения. М.: Наука, 1967.
2. Малов С.В. Регрессионный анализ: теоретические основы и практические рекомендации. СПб.: Изд-во С.-Пб. Ун-та, 2013.
3. Мудров В.И., Кушко В.Л. Метод наименьших модулей. М.: URSS, 2013.
4. Charnes A., Cooper W.W. Management models and industrial applications of linear programming. N.Y.: Wiley, 1961.
5. Статников Р.Б., Матусов И.Б. О недопустимых, допустимых и оптимальных решениях в задачах проектирования // Проблемы машиностр. и надежности машин. 2012. № 4. С. 10–19.
6. Подиновский В.В., Нелюбин А.П. Средние величины: многокритериальный подход // Проблемы управления. 2020. № 5. С. 3–16.
7. Подиновский В.В., Нелюбин А.П. Средние величины: многокритериальный подход. II // Проблемы управления. 2021. № 2. С. 33–41.
8. Тырсин А.Н., Азарян А.А. Точные алгоритмы реализации метода наименьших модулей на основе спуска по узловым прямым // Вычисл. матем. 2017. № 4. С. 21–32.
9. Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. М.: Физматлит, 1982. Изд. второе, испр. и доп., 2007.

10. Зоркальцев В.И. Октаэдрические и евклидовы проекции точки на линейное многообразие // Тр. ин-та матем. и механ. УрО РАН. 2012. № 3. С. 106–118.
11. Подиновский В.В. Многокритериальные задачи с однородными равноценными критериями // Ж. вычисл. матем. и матем. физ. 1975. № 2. С. 330–344.
12. Подиновский В.В. Идеи и методы теории важности критериев в многокритериальных задачах принятия решений. М.: Наука, 2019.
13. Нелюбин А.П., Подиновский В.В. Алгоритмическое решающее правило, использующее ординальные коэффициенты важности критериев со шкалой первой порядковой метрики // Ж. вычисл. матем. и матем. физ. 2012. № 1. С. 43–59.
14. Fishburn P.C., Willig R.D. Transfer principles in income redistribution // J. of Public Economics. 1984. V. 25. P. 323–328.
15. Marshall A.W., Olkin I. Inequalities: Theory of majorization and its applications. N.Y.: Academic press, 1979.
16. Eurostat. Real GDP per capita, 2022. https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table