Math-Net.Ru

Общероссийский математический портал

Е. С. Горская, Об одном алгоритме линеаризации выпуклых экстремальных задач, *Матем. сб.*, 2010, том 201, номер 4, 3–24

DOI: <https://doi.org/10.4213/sm7594>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением <http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 34.229.108.108

26 апреля 2023 г., 06:55:21

УДК 519.853.3+ 517.518.8+514.172.45

Е. С. Горская

Об одном алгоритме линеаризации выпуклых экстремальных задач

Предложен метод приближенного решения задач минимизации выпуклых функций многих переменных при выпуклых ограничениях. Основная идея этого метода состоит в аппроксимации выпуклой функции кусочно линейной функцией, в результате чего задача выпуклого программирования сводится к задаче линейного программирования. Для осуществления подобной аппроксимации надграфик выпуклой функции приближается проекцией многогранника большей размерности. В первой части статьи эта задача рассматривается для функций одной переменной. Для этого случая представлен алгоритм аппроксимации надграфика выпуклой функции многоугольником, доказана его оптимальность по числу вершин многоугольника и получены точные оценки для этого числа. Затем с помощью индуктивной процедуры алгоритм обобщен на определенные классы функций многих переменных. На основе предложенного метода получены полиномиальные алгоритмы приближенного вычисления L_p -нормы матрицы и минимума функции энтропии на многограннике.

Библиография: 19 названий.

Ключевые слова: выпуклые задачи, кусочно линейные функции, приближение функций, вычисление нормы операторов.

§ 1. Введение

Как известно, основная проблема, возникающая при численном решении задач выпуклого программирования вида: $f_0(x) \rightarrow \min, x \in G$, где G – выпуклое множество, заданное системой неравенств $f_i(x) \leq 0, i = 1, \dots, m$, f_i – выпуклые функции, $i = 0, \dots, m$, заключается в том, что для каждого метода существует класс функций, на котором он плохо сходится. Исключение представляют задачи линейного программирования (целевая функция и функции ограничения – линейные), которые даже при очень большом числе переменных (сотни тысяч) эффективно решаются симплекс-методом и методом внутренней точки (см., например, [1]–[3]). Поэтому вполне естественно попытаться свести задачу выпуклого программирования к задаче линейного программирования. При этом решение задачи линейного программирования может иметь большую размерность и давать не точное, а приближенное решение исходной задачи. Такое сведение может быть сделано путем приближения области определения G и надграфика целевой функции f_0 многогранниками. Задача приближения выпуклых множеств многогранниками исследовалась в литературе (см. [4]–[6]). Показано, в частности, что выпуклое тело в $\mathbb{R}^n$ можно всегда приблизить (в метрике Хаусдорфа) многогранником с точностью $\varepsilon > 0$, при этом

число граней (т.е. линейных неравенств, задающих многогранник) не превосходит $C\epsilon^{(n-1)/2}$, где константа C зависит от тела. Эта оценка неулучшаема уже для евклидова шара (см. [7], [8]). Таким образом, число граней сильно растет с ростом размерности n . Методы поиска приближающего многогранника достаточно сложны и, как правило, приводят к большим значениям константы C . Кроме того, для наших целей нужно приближение не в метрике Хаусдорфа, а в метрике пространства $\mathbb{C}(\mathbb{R}^{n-1})$, т.е. в смысле расстояния между графиками функций.

Как мы увидим, целесообразно искать приближающий многогранник как проекцию некоторого многогранника большей размерности. Насколько нам известно, эта идея принадлежит А. С. Немировскому.

В частных случаях приближение выпуклых множеств проекциями многогранников большей размерности встречалось в литературе ранее. Так, например, специальные классы выпуклых множеств (так называемые зоноиды) приближались зонотопами, которые, как известно, являются проекциями параллелепипедов большей размерности (см. [9], [10]).

В настоящей работе мы начинаем со случая функции одной переменной и представляем достаточно естественный алгоритм приближения выпуклой функции кусочно линейной, согласующийся с известными алгоритмами приближения выпуклых фигур на плоскости многоугольниками (см., например, [11]–[15]). При этом мы доказываем оптимальность этого алгоритма и получаем асимптотически точную оценку на число узлов приближающей ломаной. Затем показываем, как с помощью этой конструкции можно приближать надграфики функций многих переменных. Речь идет не обо всех функциях, а о некоторых классах функций, которые, однако, содержат достаточно много важных в приложениях примеров. Это позволяет сводить многие задачи выпуклого программирования к задачам линейного программирования.

Напомним, что *надграфиком выпуклой функции* $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ называется множество

$$\text{epi } f = \{(x, t), \text{ где } x \in \mathbb{R}^n, t \in \mathbb{R} \mid t \geq f(x)\}.$$

Обозначим через $\bar{x}$ набор координат $(x_1, \dots, x_n)$. Будем говорить, что проекция многогранника $G \subset \mathbb{R}^k$, $k > n$, приближает надграфик $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ на выпуклом множестве $V \subset \mathbb{R}^n$ снизу с точностью ϵ , если выполняются следующие два условия:

- 1) для любой точки $(\bar{x}, y) \in \text{epi } f$ такой, что $\bar{x} \in V$, существует набор $\bar{z} = (z_1, \dots, z_{k-n-1})$, для которого точка с координатами $(\bar{x}, y, \bar{z})$ принадлежит G ;
- 2) для любой точки $(x_1, \dots, x_k) \in G$ ее первые $n+1$ координат таковы, что $x_{n+1} \geq f(\bar{x}) - \epsilon$.

В этом случае рассмотрим следующий способ перехода от задачи выпуклого программирования к задаче линейного программирования: приблизим надграфик каждой из функций снизу с точностью ϵ многогранником $G_i \subset \mathbb{R}^k$ ($A_i x \leq b_i$, $i = 0, \dots, m$). Для приближения каждой функции f_i нам понадобится ввести $k - n$ дополнительных переменных $(t_1^i, \dots, t_{k-n}^i)$, где $(\bar{x}, t_i) \in \text{epi } f$.

Затем рассмотрим следующую задачу линейного программирования:

$$\begin{cases} t_0 \rightarrow \min, \\ t_1^i \leq 0, \\ A_i x \leq b_i, \quad i = 1, \dots, m, \\ x \in \mathbb{R}^k. \end{cases}$$

Главный вопрос: как построить приближающий многогранник? Напомним, что если приближать n -мерное выпуклое множество многогранником с точностью ε , то число граней такого многогранника асимптотически равно $C\varepsilon^{(1-n)/2}$, что очень велико при больших значениях n , и эта оценка неувлучшаема уже для единичного шара (см., например, [7], [8]). Поэтому если представлять многогранник в виде решения системы линейных неравенств от $n + 1$ переменной, то количество неравенств будет расти как $\varepsilon^{(1-n)/2}$ (ε – точность приближения); кроме того, задача нахождения такого многогранника алгоритмически сложна.

Как будет показано далее, для некоторых частных случаев оказывается возможным свести задачу приближения надграфика функции из $\mathbb{R}^n$ к задаче приближения функции от одной переменной.

Статья имеет следующую структуру. Сначала дано описание алгоритма для приближения надграфика выпуклой функции одной переменной и оценена его сходимость. Затем с его помощью получен алгоритм линеаризации задач выпуклого программирования с функциями f_i , которые являются симметричными функциями координат или которые раскладывается в прямую сумму выпуклых функций одной переменной (в частности, для L_p -нормы, энтропии и т.д.).

Итак, в § 2 подробно рассмотрена задача линеаризации выпуклых функций одной переменной. В этом случае приближающий многогранник является многоугольником и для того, чтобы его построить, достаточно построить ломаную, приближающую график функции с данной точностью.

В работе описан естественный алгоритм, позволяющий приблизить график вогнутой возрастающей функции $f(x)$ (не обязательно дифференцируемой) с точностью ε на данном отрезке $[0, b]$ кусочно линейной функцией ($f(0) = 0$). Доказано, что приведенный алгоритм является оптимальным. Из более общего результата А. Лигуна и В. Сторчая [16] следует, что для числа шагов N_ε рассмотренного алгоритма верна оценка

$$N_\varepsilon \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\int_0^b \sqrt{-f''(x)} dx \right)^{1/4} + C\varepsilon^{1/4}, \quad (1)$$

где C – абсолютная константа. При этом проблема состоит в том, чтобы оценить константу

$$\left(\int_0^b \sqrt{-f''(x)} dx \right)^{1/4}$$

для данной конкретной функции. Кроме того, остается открытым вопрос, как получить эффективную оценку в случае, когда функция недифференцируема.

В теореме 3 получена оценка сверху на N_ε , в которой не используется интеграл от $f''(x)$. Фактически найдена асимптотически наилучшая константа

в неравенстве (1) при $\varepsilon \rightarrow 0$. Для этого сформулирована и решена с помощью принципа максимума Понтрягина соответствующая задача оптимального управления. В результате получена оценка:

$$N_\varepsilon \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\sqrt{f(b)} J(\alpha) \right)^{1/4} + C\varepsilon^{1/4}, \quad (2)$$

где $\alpha = f'(0)b/f(b)$, p – корень уравнения $\alpha p \ln p = p^2 - 2p + 1 + \alpha p - \alpha$ в случае, когда $\alpha \neq 2$, и

$$J(\alpha) = \begin{cases} \frac{\sqrt{\alpha p}}{p-1} \ln p, & \text{если } \alpha \neq 2, \\ \sqrt{2}, & \text{если } \alpha = 2. \end{cases}$$

Затем доказано, что величина N_ε может быть оценена сверху следующим образом:

$$N_\varepsilon \leq \frac{1}{\sqrt{\varepsilon}} \cdot 2^{-5/2} \left(f(b) G^2 \left[\frac{f'(0)}{f(b)} b \right] \right)^{1/8} + C\varepsilon^{1/4}, \quad (3)$$

$$G(\alpha) = \begin{cases} 2\sqrt{\ln \alpha}, & \text{если } \alpha > 2; \\ \sqrt{2}, & \text{если } 1 \leq \alpha \leq 2. \end{cases}$$

Далее показано, как оценивается число шагов алгоритма в случае, когда $f'(0) = \infty$. Эта ситуация характерна для многих часто используемых функций, например, $f(x) = -x \ln x$, $f(x) = x^p$, $0 < p < 1$.

В § 3 мы переходим к многомерному случаю: функции $f_i(x)$ зависят от n переменных, $x = (x_1, \dots, x_n)$. Показано, что если функция f есть симметричная функция координат или раскладывается в прямую сумму выпуклых функций от координат, то можно эффективно осуществить ее линейризацию. В частности, можно линейризовать L_p -норму $f(x) = (x_1^p + \dots + x_n^p)^{1/p}$, функцию энтропии $f(x) = \sum_{i=1}^n x_i \ln x_i$ и ей сопряженную $f(x) = \ln(\sum_{i=1}^n e^{x_i})$ в $\mathbb{R}^n$. Таким образом, мы получаем возможность решать задачи выпуклого программирования, в которых используются такие функции, сводя их к задачам линейного программирования.

Параграф 3 завершается рассмотрением задач вычисления L_p -нормы оператора в $\mathbb{R}^n$ и оценки числа обусловленности $(n \times n)$ -матрицы в норме L_p . Обе задачи, как известно, NP-сложные, однако их можно приближенно решить предложенным методом.

§ 2. Линейризация функций одной переменной

В этом параграфе мы рассмотрим алгоритм для приближения надграфика выпуклой функции одной переменной и оценим его сходимость. Также докажем асимптотически точную оценку для количества узлов приближающей ломаной.

Итак, пусть дана выпуклая функция $f: \mathbb{R} \rightarrow \mathbb{R}$. Необходимо приблизить ее надграфик с точностью ε при помощи системы линейных неравенств, т.е. найти A и b , задающие многогранник $G = \{(x, y) \in \mathbb{R}^2 : A \begin{pmatrix} x \\ y \end{pmatrix} \leq b\}$ такой, что для всех x имеем $(x, f(x)) \in G$ и для всех пар $(x, y) \in G$ выполнено $f(x) - y \leq \varepsilon$.

Можно считать, что мы приближаем не надграфик системой линейных неравенств, а график функции $f(x)$ кусочно линейной функцией $\tilde{f}_\varepsilon$ с точностью ε (т.е. для любого x выполнено $|f(x) - \tilde{f}_\varepsilon(x)| \leq \varepsilon$). Действительно, построив ломаную $\tilde{f}_{\varepsilon/2}(x)$ и сдвинув ее на вектор $(0, -\varepsilon/2)$, получим границу искомого многогранника.

Также заметим, что вместо выпуклой функции мы можем рассматривать вогнутую функцию. Для оценки количества линейных неравенств будем оценивать количество вершин ломаной.

Итак, имеется функция $f(x)$ на отрезке $[0, b]$, причем $f(0) = 0$, $f(b) = m$, $f'(0) = \alpha$, $f'(b) = \beta$. Нужно приблизить график этой функции ломаной с точностью ε . Можно также считать, что $f(x)$ – возрастающая, в противном случае отрезок $[0, b]$ разбивается на две части, на одной функция возрастает, а на другой – убывает (алгоритм построения ломаной на втором участке аналогичен первому).

Алгоритм. Зададим индуктивно последовательность абсцисс узлов ломаной $\{x_n\}$ и для каждого узла положим $y_k = f(x_k) + \varepsilon$. Пусть $x_1 = 0$, $y_1 = \varepsilon$. Предположим, что мы уже построили узлы с номерами от 1 до k . Рассмотрим прямую $y = \alpha(x - x_k) + y_k$ такую, что $\max_{x \geq x_k} (f(x) - \alpha(x - x_k) - y_k) = \varepsilon$. Затем найдем x_{k+1} из уравнения $f(x) - \alpha(x - x_k) - y_k = -\varepsilon$. В случае, когда α или x_{k+1} не существуют, полагаем $x_{k+1} = b$ и алгоритм заканчивает свою работу. Ломаная, последовательно соединяющая точки (x_k, y_k) , является искомой кусочно линейной функцией $\tilde{f}$, приближающей график $f(x)$ с точностью ε .

Докажем, что алгоритм является оптимальным. Далее будем обозначать через $N(\tilde{f})$ количество узлов ломаной (кусочно линейной функции) $\tilde{f}$.

ТЕОРЕМА 1. Для любой кусочно линейной функции $\hat{f}(x)$ такой, что $\|\hat{f}(x) - f(x)\|_{C[0,b]} \leq \varepsilon$ имеем $N(\hat{f}) \geq N(\tilde{f})$.

Рис. 1

ДОКАЗАТЕЛЬСТВО. Предположим, что это не так, т.е. существует ломаная g такая, что $\sup_{[0,b]} |f(x) - g(x)| \leq \varepsilon$, с меньшим количеством узлов, чем $\tilde{f}$. Рассмотрим три первых узла ломаной $\tilde{f}$ (обозначим их через A_1, A_2 и A_3) и три первых узла ломаной g (обозначим их через A'_1, A'_2, A'_3), (см. рис. 1). Пусть B_i и C_i – точки, принадлежащие f и $\tilde{f}$ соответственно, в которых достигается равенство $\max(f(x) - \tilde{f}(x)) = \varepsilon$, а D_i – точки, принадлежащие f , в которых достигается равенство $\max(\tilde{f}(x) - f(x)) = \varepsilon$.

Рассмотрим точки $A_i(x_i, f(x_i))$, $B_i(a_i, f(a_i))$, $A'_i(x'_i, g(x'_i))$. Заметим, что A'_1 принадлежит отрезку с координатами $[0, \varepsilon]$ и $[0, -\varepsilon]$. Предположим, что A'_1 отлична от A_1 . Тогда если первое звено $A'_1 A'_2$ ломаной g не пересекает отрезка $A_1 C_1$, то $f(a_1) - g(a_1) > \varepsilon$, что невозможно. Следовательно, $A'_1 A'_2$ пересекает отрезок $A_1 C_1$, но тогда $g(x_2) - f(x_2) > \varepsilon$, т.е. узел A'_2 лежит левее, чем A_2 .

Далее, если второе звено $A'_2 A'_3$ ломаной g не пересекает $A_2 C_2$, то либо A'_3 лежит левее, чем A_3 , либо приходим к противоречию ($f(a_2) - g(a_2) > \varepsilon$, что невозможно). Если же $A'_2 A'_3$ пересекает отрезок $A_2 C_2$, то $g(x_3) - f(x_3) > \varepsilon$, т.е. в любом случае $x'_3 < x_3$. Далее, рассуждая аналогичным образом, получаем, что $x'_{n-1} < x_{n-1}$, т.е. количество узлов ломаной g не меньше, чем количество узлов ломаной $\tilde{f}$.

ЗАМЕЧАНИЕ 1. Рассмотрим задачу в другой постановке: дана вогнутая возрастающая функция $f(x)$ на отрезке $[0, 1]$ и N – количество узлов приближающей ломаной. Нужно найти наименьшее ε и кусочно линейную функцию $\tilde{f}_\varepsilon$ такую, что $N(\tilde{f}_\varepsilon) \leq N$ и

$$\|f(x) - \tilde{f}_\varepsilon(x)\|_{C[0,1]} \leq \varepsilon.$$

Используем так называемый “метод стрельбы”.

Пусть искомое значение $\varepsilon \leq 1$. Проведем процесс поиска в два этапа.

1) На каждом шаге полагаем $\varepsilon_m = (\varepsilon_l + \varepsilon_r)/2$, строим ломаную $\tilde{f}_{\varepsilon_m}$ согласно алгоритму, находим $N(\tilde{f}_{\varepsilon_m})$. Если $N(\tilde{f}_{\varepsilon_m}) > N$, то полагаем $\varepsilon_l = \varepsilon_m$; если $N(\tilde{f}_{\varepsilon_m}) < N$, то полагаем $\varepsilon_r = \varepsilon_m$. Затем переходим к следующему шагу (в случае равенства первый этап алгоритма завершен). На первом шаге $\varepsilon_l = 0$, $\varepsilon_r = \varepsilon_m = 1$.

После окончания первого этапа алгоритма найдено некоторое ε_m такое, что $N(\tilde{f}_{\varepsilon_m}) = N$. Кроме того, мы запоминаем последнее значение ε_l . Пусть δ – точность, с которой мы хотим определить ε . Далее поступаем следующим образом. Если $\tilde{f}_{\varepsilon_m}(1) - f(1) \geq \varepsilon_m - \delta$, то считаем, что искомое значение $\varepsilon = \varepsilon_m$ и искомая ломаная – $\tilde{f}_{\varepsilon_m}$. В противном случае переходим ко второму этапу.

2) На каждом шаге вычисляем $\varepsilon_m = (\varepsilon_l + \varepsilon_m)/2$, согласно алгоритму находим $\tilde{f}_{\varepsilon_m}$ и вычисляем $\tilde{f}_{\varepsilon_m}(1) - f(1)$. Если $\tilde{f}_{\varepsilon_m}(1) - f(1) < \varepsilon_m - \delta$, то ε_m можно еще немного уменьшить, поэтому переходим к следующему шагу. В противном случае алгоритм заканчивает свою работу и мы считаем, что искомое значение $\varepsilon = \varepsilon_m$ и искомая ломаная – $\tilde{f}_{\varepsilon_m}$.

Доказательство того, что не существует другой функции с таким же количеством узлов, приближающей с большей точностью, аналогично доказательству теоремы 1. Пусть существует кусочно линейная функция g с N узлами, приближающая f с точностью $\varepsilon < \varepsilon_m$. Обозначим через $x_1, \dots, x_N$

абсциссы узлов ломаной f_{ε_m} , а через $x'_1, \dots, x'_N$ – абсциссы узлов ломаной g и рассмотрим последовательно несколько звеньев этих ломаных. Получим, что $x'_1 < x_1 < \dots < x'_N < x_N$, т.е. ломаная g имеет больше, чем N узлов.

Устремив в изложенном рассуждении δ к нулю, получим следствие 1, которое мы приведем без доказательства.

СЛЕДСТВИЕ 1. Пусть $f(x)$ – вогнутая возрастающая функция на отрезке $[0, b]$ такая, что $f(0) = 0$. Пусть также K_N – множество кусочно линейных функций, имеющих не более N узлов. Тогда задача

$$\begin{cases} \|\tilde{f} - f\|_{C[0,b]} \rightarrow \min, \\ \tilde{f} \in K_N, \end{cases}$$

имеет единственное решение, которое образует альтернанс с функцией f .

ЗАМЕЧАНИЕ 2. Классическая задача теории приближений заключается в аппроксимации данной гладкой функции функциями из некоторого класса F . При этом, как известно, если F является конечномерным линейным пространством, то элемент наилучшего приближения в равномерной метрике на отрезке образует альтернанс с функцией f (теорема Чебышёва, см., например, [17]).

В нашем случае мы приближаем выпуклую функцию (не обязательно гладкую!) функциями из класса кусочно линейных функций, т.е. элементами бесконечномерного пространства. Тем не менее, как мы видим (теорема 1), элемент наилучшего приближения при данном числе узлов также образует альтернанс с функцией f .

Оценка количества узлов ломаной. Оценим количество шагов алгоритма, или, что то же самое, количество N_ε узлов ломаной $\tilde{f}_\varepsilon$.

ТЕОРЕМА 2. Для общего количества узлов ломаной $\tilde{f}_\varepsilon$, порожденной алгоритмом, справедлива следующая оценка:

$$N_\varepsilon \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\int_0^b \sqrt{-f''(x)} dx \right)^{1/4} + C\varepsilon^{1/4}, \quad (4)$$

где C – абсолютная константа.

ДОКАЗАТЕЛЬСТВО. Теорема 2 следует из более общего результата, изложенного в работе [16], где рассматривалась задача приближения гладкой функции кусочно полиномиальными функциями (сплайнами) заданной степени в пространстве L_p . Теорема 2 является частным случаем результата [16], если положить в нем $p = \infty$, а степень считать равной единице.

Возникает естественный вопрос: можно ли оценить число узлов кусочно линейного приближения, не прибегая к вычислению интеграла, приведенного в теореме 2, которое может быть технически сложным для многих функций? Кроме того, интеграл

$$\int_0^b \sqrt{-f''(x)} dx$$

может расходиться (в этом случае оценка (4) ничего не дает). Наконец, как оценить число узлов приближения для негладких функций?

Пусть $\alpha = f'(0)b/f(b)$, p – корень уравнения $\alpha p \ln p = p^2 - 2p + 1 + \alpha p - \alpha$ в случае, когда $\alpha \neq 2$, и

$$J(\alpha) = \begin{cases} \frac{\sqrt{\alpha p}}{p-1} \ln p, & \text{если } \alpha \neq 2, \\ \sqrt{2}, & \text{если } \alpha = 2. \end{cases}$$

ТЕОРЕМА 3. Для любой вогнутой функции f для общего количества узлов ломаной $\tilde{f}_\varepsilon$, порожденной алгоритмом, справедлива следующая оценка:

$$N_\varepsilon \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} (\sqrt{f(b)} J(\alpha))^{1/4} + C\varepsilon^{1/4}, \quad (5)$$

где C – абсолютная константа.

ДОКАЗАТЕЛЬСТВО. 1) Пусть f – вогнутая, дважды дифференцируемая на отрезке $[0, b]$ функция. Обозначим через $J(f)$ решение следующей задачи:

$$\int_0^b \sqrt{-f''(t)} dt \rightarrow \max,$$

$$f(0) = 0, \quad f(b) = m, \quad f'(0) = \alpha_1, \quad f'(b) = \beta_1, \quad f''(t) \leq 0 \text{ для любого } t \in [0, b].$$

Сведем эту задачу к случаю, когда $f'(b) = 0$.

Для этого выберем некоторое значение $\delta > 0$ и доопределим $f(t)$ на отрезке $[b, b + \delta]$ при помощи квадратичной функции так, чтобы значение первой производной новой функции в точке $b + \delta$ равнялось нулю, т.е. возьмем $g(t) = -c(t - b - \delta)^2 + d$ и

$$f_1(t) = \begin{cases} f(t), & \text{если } t \in [0, b], \\ g(t), & \text{если } t \in [b, b + \delta], \end{cases}$$

где c и d определяются из условий $f(b) = g(b)$ и $f'(b) = g'(b)$. Тогда f_1 вогнута и непрерывна на отрезке $[0, b + \delta]$; кроме того, $f'_1(b + \delta) = 0$.

Пусть $J_1(f)$ – решение задачи:

$$\int_0^{b+\delta} \sqrt{-f''_1(t)} dt \rightarrow \max,$$

$$f_1(0) = 0, \quad f_1(b + \delta) = d, \quad f'_1(0) = \alpha_1, \quad f'_1(b + \delta) = 0, \quad f''_1(t) \leq 0, \quad t \in [0, b].$$

Вычислим $c = f'(b)/(2\delta)$ и $d = f(b) + f'(b)\delta/2$. Рассмотрим разность

$$\int_0^{b+\delta} \sqrt{-f''_1(t)} dt - \int_0^b \sqrt{-f''(t)} dt = \int_b^{b+\delta} \sqrt{-g''(t)} dt = \int_b^{b+\delta} \sqrt{2c} dt = \sqrt{f'(b)} \cdot \sqrt{\delta}.$$

Следовательно, $J_1(f) - J(f) = \sqrt{f'(b)} \cdot \sqrt{\delta} \rightarrow 0, \delta \rightarrow 0$. Поэтому нам достаточно решить задачу, в которой $f'(b) = 0$.

Для того чтобы оценивать интеграл на отрезке $[0, 1]$, сделаем замену $f_1(t) = f(bt)/f(b)$.

Таким образом, необходимо решить следующую задачу:

$$\begin{cases} J = \int_0^1 \sqrt{-\ddot{x}(t)} dt \rightarrow \max, \\ x(0) = 0, \quad x(1) = 1, \\ \dot{x}(0) = \alpha, \quad \dot{x}(1) = 0, \\ \ddot{x} \leq 0. \end{cases}$$

Сделаем замену $x_1 = x$, $x_2 = \dot{x}_1$, $u = \dot{x}_2$. Получим

$$\begin{cases} \int_0^1 \sqrt{-u} dt \rightarrow \max, \\ x_1(0) = 0, \quad x_1(1) = 1, \\ x_2(0) = \alpha, \quad x_2(1) = 0, \\ u \leq 0. \end{cases}$$

Применив принцип максимума Понтрягина, выпишем интегрант и терминант:

$$\begin{aligned} L &= \lambda_0 \sqrt{-u} + p_1(t)(\dot{x}_1(t) - x_2(t)) + p_2(t)(\dot{x}_2(t) - u), \\ l &= \lambda_1 x_1(0) + \lambda_2(x_1(1) - 1) + \lambda_3(x_2(0) - \alpha) + \lambda_4 x_2(1). \end{aligned}$$

Уравнения Эйлера–Лагранжа дают следующие условия:

$$\begin{cases} \dot{p}_1(t) = 0, \\ \dot{p}_2(t) = -p_1(t). \end{cases}$$

Условия трансверсальности таковы: $p_1(0) = \lambda_1$, $p_1(1) = -\lambda_2$, $p_2(0) = \lambda_3$, $p_2(1) = -\lambda_4$. Следовательно, $p_1(t) = \lambda_1 = -\lambda_2$, $p_2(t) = \lambda_2 t + \lambda_3$ и $\lambda_2 + \lambda_3 = -\lambda_4$.

Условие стационарности по управлению формулируется так: в любой точке t непрерывности управления $u(t)$ значение $u(t)$ есть точка минимума функции $L(u) = \lambda_0 \sqrt{-u} - p_2(t)u$. Рассмотрим два случая.

а) $\lambda_0 = 0$. Тогда $u \in \operatorname{argmin} L(u) = -p_2(t)u$. Если $p_2(t) > 0$ на $[0, 1]$, то $u \equiv 0$; если $p_2(t) < 0$ на $[0, 1]$, то $u = -\infty$; случай, когда $p_2(t)$ меняет знак на $[0, 1]$, также не годится.

б) Таким образом, $\lambda_0 \neq 0$. Без ограничения общности можем считать, что $\lambda_0 = 2$. Заметим, что

$$u \in \operatorname{argmin} L(u) \iff \sqrt{-u} \in \operatorname{argmax} L(\sqrt{-u}) = p_2(t)(\sqrt{-u})^2 + 2\sqrt{-u}.$$

Последнее уравнение задает параболу, проходящую через начало координат. Если $p_2(t) > 0$ на $[0, 1]$, то ее ветви направлены вверх и $\sqrt{-u} = +\infty$ (следовательно, случай, когда $p_2(t)$ меняет знак на $[0, 1]$, не годится). Если же $p_2(t) < 0$ на $[0, 1]$, то ее ветви направлены вниз и максимум достигается в вершине параболы: $\sqrt{-u} = -1/p_2(t)$.

Как мы показали выше, $p_2(t) = \lambda_2 t + \lambda_3$. Далее необходимо рассмотреть два случая.

Начнем со случая $\lambda_2 = 0$. Получаем, что $p_1(t) = \lambda_1 = \lambda_2 = 0$, $\lambda_3 = -\lambda_4$, $p_2(t) = \lambda_3$, $\sqrt{-u} = -1/\lambda_3$. Тогда $u = -1/\lambda_3^2$ и $x_2(t) = -t/\lambda_3^2 + C_1$, $x_1(t) = -t^2/(2\lambda_3^2) + C_1 t + C_2$. Запишем граничные условия на $x_1(t)$ и $x_2(t)$:

$$\begin{cases} C_2 = 0, & C_1 = \alpha, \\ C_1 - \frac{1}{2\lambda_3^2} = 1, \\ C_1 - \frac{1}{\lambda_3^2} = 0. \end{cases}$$

Решив эту систему, получим, что $\lambda_3 = -1/\sqrt{2}$, $\alpha = C_1 = 2$. Таким образом, рассмотренный случай дает решение задачи, только если $\alpha = 2$. Искомая экстремаль задается так: $x_1(t) = -t^2 + 2t$, $\int_0^1 \sqrt{-\ddot{x}(t)} dt = \sqrt{2}$.

Рассмотрим случай, когда $\lambda_2 \neq 0$. Тогда $p_2(t) = \lambda_2 t + \lambda_3$, где $\lambda_2 \neq 0$. Следовательно,

$$u = -\frac{1}{(\lambda_2 t + \lambda_3)^2},$$

$$x_2(t) = \frac{1}{\lambda_2} \frac{1}{\lambda_2 t + \lambda_3} + C, \quad x_1(t) = \frac{1}{\lambda_2^2} \ln(\lambda_2 t + \lambda_3) + Ct + D,$$

где C и D – некоторые константы.

Преобразуем $x_1(t) = (1/\lambda_2^2) \ln(\lambda_2 t + \lambda_3) + Ct + D + (\ln \lambda_3)/\lambda_2^2$. Поскольку $x_1(0) = 0$, то $D + (\ln \lambda_3)/\lambda_2^2 = 0$. Итак,

$$x_1(t) = \frac{1}{\lambda_2^2} \ln\left(\frac{\lambda_2}{\lambda_3} t + 1\right) + Ct, \quad x_2(t) = \frac{1}{\lambda_2^2} \frac{1}{\lambda_2 t + \lambda_3} + C,$$

$$u(t) = -\frac{1}{(\lambda_2 t + \lambda_3)^2}, \quad J = \int_0^1 \sqrt{-u} dt = \frac{1}{\lambda_2} \ln\left(\frac{\lambda_2}{\lambda_3} + 1\right).$$

Сделав замену, получим $x_1(t) = c_1 \ln(c_2 t + 1) + c_3 t$, $x_2(t) = c_1 c_2 / (c_2 t + 1) + c_3$, $J = \sqrt{c_1} \ln(c_2 + 1)$ и граничные условия принимают вид

$$\begin{cases} c_1 \ln(c_2 + 1) + c_3 = 1, \\ c_1 c_2 + c_3 = \alpha, \\ \frac{c_1 c_2}{c_2 + 1} + c_3 = 0. \end{cases} \iff \begin{cases} \frac{\alpha(c_2 + 1)}{c_2^2} \ln(c_2 + 1) - \frac{\alpha}{c_2} = 1, \\ c_1 = \frac{\alpha(c_2 + 1)}{c_2^2}, \\ c_3 = -\frac{\alpha}{c_2}. \end{cases}$$

Таким образом, c_2 – корень уравнения $\alpha(c_2 + 1) \ln(c_2 + 1) - \alpha c_2 = c_2^2$. Пусть $p = c_2 + 1$. Тогда $p > 0$, $p \neq 1$ и, кроме того, p – корень уравнения

$$\alpha p \ln p = p^2 - 2p + 1 + \alpha p - \alpha. \quad (6)$$

ЛЕММА 1. При любом значении α одним из корней уравнения (6) является $p = 1$. При $\alpha = 2$ оно имеет единственный корень $p = 1$. Если же $\alpha > 2$ или $1 < \alpha < 2$, то это уравнение всегда имеет второй корень, отличный от единицы и положительный. В первом случае $p > 1$, а во втором случае $1 > p > 0$.

Доказательство леммы элементарно и мы оставляем его читателю.

ЗАМЕЧАНИЕ 3. Случаю $\alpha = 2$ соответствует экстремаль $x_1(t) = -t^2 + 2t$ (когда $p_2(t) = \text{const}$, см. выше).

Продолжим доказательство теоремы. Выразим J через p и α . Имеем

$$J = \sqrt{c_1} \ln(c_2 + 1) = \sqrt{\frac{\alpha(c_2 + 1)}{c_2^2}} \ln(c_2 + 1) = \frac{\sqrt{\alpha p}}{p - 1} \ln p. \quad (7)$$

Сделав обратные замены, получим утверждение теоремы для дважды дифференцируемой функции.

2) Поскольку любую непрерывную вогнутую функцию можно сколь угодно точно приблизить гладкой функцией, то утверждение теоремы верно и для любых вогнутых функций.

ЗАМЕЧАНИЕ 4. Пусть K_N – класс кусочно линейных функций на отрезке $[0, b]$ с не более чем N узлами. Тогда для любой вогнутой возрастающей функции f на этом отрезке и $\hat{f} \in K_N$ справедливо неравенство

$$\inf \|f - \hat{f}\| \leq G(f) \cdot N^{-2} + o(N^{-2}). \quad (8)$$

Действительно, умножим обе части неравенства (5) на $\sqrt{\varepsilon}$ и N_ε^{-1} и затем возведем в квадрат. Заметим, что мы нашли (теорема 3) точную константу $G(f)$ в этом неравенстве:

$$G(f) = \frac{\|f\|_{C[0,b]}^{1/4}}{32} J(\alpha)^{1/2}, \quad \alpha = \frac{\|f'\|_{C[0,b]} b}{\|f\|_{C[0,b]}}$$

p – корень уравнения $\alpha p \ln p = p^2 - 2p + 1 + \alpha p - \alpha$ в случае, если $\alpha \neq 2$,

$$J(\alpha) = \begin{cases} \frac{\sqrt{\alpha p}}{p - 1} \ln p, & \text{если } \alpha \neq 2, \\ \sqrt{2}, & \text{если } \alpha = 2. \end{cases}$$

Неравенство (8) может быть интерпретировано как классическое неравенство Джексона, связывающее расстояние от функции до пространства алгебраических (либо тригонометрических) полиномов с гладкостью функции f . В нашем случае приближаемая функция f вогнута, а пространство полиномов заменяется на класс кусочно линейных функций с данным числом узлов. Таким образом, неравенство (8) дает точную оценку в этой задаче.

Итак, для любого значения $\alpha \neq 2$, $\alpha > 1$, мы можем численно решить уравнение (6), подставить его корень (отличный от единицы) в (7) и получить точное

значение $J(\alpha)$. При $\alpha = 2$ значение $J = \sqrt{2}$, при $\alpha \rightarrow 1$ значение $J \rightarrow 0$, так как экстремаль стремится к прямой $x(t) = t$.

Таким образом, мы нашли точную константу в оценке (4) и экстремаль, на которой она достигается. Оценим ее сверху более простой функцией для того, чтобы получить представление о том, насколько велика может быть эта константа для произвольной $f(x)$. Найдем функцию $G(\alpha)$ такую, что $J(\alpha) \leq G(\alpha)$ для всех допустимых значений α . Пусть

$$G(\alpha) = \begin{cases} 2\sqrt{\ln \alpha}, & \text{если } \alpha > 2, \\ \sqrt{2}, & \text{если } 1 \leq \alpha \leq 2. \end{cases}$$

Тогда верна

ТЕОРЕМА 4. $J(\alpha) \leq G(\alpha)$ для любого $\alpha \geq 1$.

ДОКАЗАТЕЛЬСТВО. Выразим α через p . Имеем $\alpha = (p-1)^2/(p \ln p - p + 1)$.

1) Рассмотрим случай, когда $1 < \alpha < 2$. Согласно лемме 1 в этом случае $0 < p < 1$. Тогда $J(\alpha(p)) = -\ln p \sqrt{p/(p \ln p - p + 1)}$. Докажем, что $J(p)$ возрастает по p при $0 < p < 1$. Найдем

$$J'(p) = \frac{2p - 2 - (p+1) \ln p}{2(p \ln p - p + 1)^2 \sqrt{p/(p \ln p - p + 1)}}.$$

Знаменатель больше нуля для любого $p > 0$, числитель $h(p) = 2p - 2 - (p+1) \times \ln p > 0$, так как $h(1) = 0$, а $h'(p) = 1 - \ln p - 1/p < 0$, $0 < p < 1$. Следовательно, $J(\alpha(p)) \leq J(2) = \sqrt{2}$ при $0 < p < 1$.

2) Рассмотрим случай, когда $\alpha > 2$. Согласно лемме 1 в этом случае $p > 1$. Пусть

$$f(p) = G(\alpha(p)) - J(\alpha(p)) = 2\sqrt{\ln \frac{(p-1)^2}{p \ln p - p + 1}} - \sqrt{\frac{p}{p \ln p - p + 1}} \ln p.$$

Докажем, что $f(p) > 0$ при $p > 1$. Для этого достаточно доказать, что

$$h(p) = 4 \ln \frac{(p-1)^2}{p \ln p - p + 1} - \frac{p \ln^2 p}{p \ln p - p + 1} > 0, \quad p > 1.$$

Для этого заметим, что при $p \rightarrow 1$ имеем $h(p) \rightarrow 4 \ln 2 - 2 > 0$. Докажем, что при $p > 1$ функция $h(p)$ возрастает по p . Имеем

$$h'(p) = \frac{8(p-1)^2 + (6 + 4p - 10p^2) \ln p + (3p^2 + 4p + 1) \ln^2 p}{(p-1)(p \ln p - p + 1)^2}.$$

Знаменатель этой дроби больше нуля при $p > 1$. Рассмотрим числитель дроби отдельно:

$$\begin{aligned} & 8(p-1)^2 + (6 + 4p - 10p^2) \ln p + (3p^2 + 4p + 1) \ln^2 p \\ &= [4(p-1) - (3p+1) \ln p] [2(p-1) - (p+1) \ln p]. \end{aligned}$$

Докажем, что оба сомножителя меньше нуля при $p > 1$. Действительно, функции $f_1(p) = 4(p-1) - (3p+1)\ln p$ и $f_2(p) = 2(p-1) - (p+1)\ln p$ вогнуты на $[1, \infty)$ и, кроме того, $f_1(1) = f_2(1) = 0$ и $f_1'(1) = f_2'(1) = 0$. Следовательно, $f_1(p) < 0$ и $f_2(p) < 0$ при $p > 1$. Таким образом, $h'(p) > 0$ при $p > 1$ и утверждение теоремы доказано.

Таким образом, при заданной функции $f(x)$ и $\varepsilon \rightarrow 0$ оценка на количество узлов асимптотически не лучше, чем оценка

$$N_\varepsilon \leq \frac{C_0(f)}{\sqrt{\varepsilon}}, \quad C_0 = 2^{-5/2} \left(f(b) G^2 \left[\frac{f'(0)}{f(b)} b \right] \right)^{1/8}. \quad (9)$$

Пусть теперь дана функция $f(x)$ и фиксированное значение ε . Если при этом $f'(0) = \infty$, то полученная выше оценка означает лишь то, что количество узлов растет быстрее, чем $1/\sqrt{\varepsilon}$.

ПРЕДЛОЖЕНИЕ 1. В случае, когда $f'(0)$ достаточно велико, сделаем первый шаг алгоритма и оценим значение $f'(x_2)$ сверху, а x_2 – снизу. Тогда количество узлов приближающей ломаной не превосходит

$$N_\varepsilon \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4f(b-x_2) \ln \left[\frac{f'(x_2)}{f(b-x_2)} (b-x_2) \right] \right)^{1/8} + C\varepsilon^{1/4}.$$

Покажем, как получить оценки для $f'(x_2)$ и x_2 для произвольной вогнутой функции $f(x)$. Пусть $y = kx + \varepsilon$ – уравнение первого звена ломаной. Тогда прямая $y = kx$ пересекает график ломаной в точках с абсциссами x_1 и x_2 (см. рис. 2). Пусть AB – касательная к графику $f(x)$ во втором узле. Тогда $AO = f(x_2) - x_2 \cdot f'(x_2) \geq 2\varepsilon$. Откуда $f'(x_2) \leq (f(x_2) - 2\varepsilon)/x_2$. Решив для данной функции уравнение $f'(x_2) = (f(x_2) - 2\varepsilon)/x_2$, получим искомую оценку на $f'(x_2)$ и x_2 .

Рис. 2

Частные случаи и приложения. На практике нас обычно интересуют функции, ведущие себя в окрестности нуля как $f(x) = Kx^q$ ($0 < q < 1$, поскольку функция выпуклая). Найдем в этом случае точные значения x_2

и $f'(x_2)$. Пусть $y = kx + \varepsilon$ – уравнение первого звена ломаной, $\bar{x}$ – точка, в которой достигается равенство $f(\bar{x}) - (k\bar{x} + \varepsilon) = \varepsilon$ (см. рис. 3). Поскольку $y = kx + 2\varepsilon$ – касательная к графику функции $f(x)$ в точке $\bar{x}$, то $k = f'(\bar{x})$ и $2\varepsilon = f(\bar{x}) - \bar{x}f'(\bar{x})$. Так как $f'(x) = Kqx^{q-1}$, то $2\varepsilon = K\bar{x}^q - \bar{x} \cdot Kq\bar{x}^{q-1}$. Следовательно, $\bar{x} = (2\varepsilon/(K(1-q)))^{1/q}$, т.е. $k = f'(\bar{x}) = Kq(2\varepsilon/(K(1-q)))^{(q-1)/q}$. Найдем x_2 и $f(x_2)$ из условия $f(x_2) = kx_2$. Получим, что

$$f'(x_2) = K^{1/q} \cdot q^2 \left(\frac{1-q}{2} \right)^{(1-q)/q} \cdot \left(\frac{1}{\varepsilon} \right)^{(1-q)/q},$$

$$x_2 = \left(\frac{1}{K} \right)^{1/q} q^{1/(q-1)} \left(\frac{2}{1-q} \right)^{1/q} \varepsilon^{1/q}.$$

Таким образом, на отрезке $[0, 1]$ имеем

$$N_\varepsilon \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4K(1-x_2)^q \ln \left[\frac{f'(x_2)(1-x_2)}{K(1-x_2)^q} \right] \right)^{1/8} + C\varepsilon^{1/4}.$$

Рис. 3

Поскольку $q^2(1-q)^{(1-q)/q} \leq 1$, то $f'(x_2) \leq K^{1/q}(1/(2\varepsilon))^{(1-q)/q}$. Кроме того, $(1-x_2)^q \leq 1$ и $(1-q) \leq 1$. Следовательно,

$$N_\varepsilon \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4K \ln \left(K^{1/q} \left[\frac{1}{2\varepsilon} \right]^{(1-q)/q} \right) - 4K \ln K(1-x_2)^{q-1} \right)^{1/8} + C\varepsilon^{1/4}$$

$$\leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4K \left[\frac{\ln K}{q} + \frac{1-q}{q} \ln \frac{1}{2\varepsilon} - \ln K + (1-q) \ln(1-x_2) \right] \right)^{1/8} + C\varepsilon^{1/4}$$

$$\leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\frac{4K}{q} \ln \frac{1}{2\varepsilon} - 4K \ln K \right)^{1/8} + C\varepsilon^{1/4}.$$

ПРИМЕР 1. Оценим количество узлов ломаной для функции

$$f(x) = (1 - (1-x)^p)^{1/p}, \quad p > 1, \quad x \in [0, 1].$$

Поскольку $f(x) \leq p^{(p-1)/p} x^{1/p}$ на отрезке $[0, 1]$, то $x_2 = c_1 \varepsilon^p$, $f'(x_2) = c_2 (1/\varepsilon)^{p-1}$, где

$$c_1 = p^{(2p-1)/(p-1)} \left(\frac{2}{p-1} \right)^p, \quad c_2 = \frac{1}{p^2} \left(\frac{p-1}{2} \right)^{p-1}.$$

Обозначим искомое количество узлов через $N_\varepsilon(L_p)$. Тогда

$$N_\varepsilon(L_p) \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4 \left[1 - \varepsilon^{p^2} c_1^p \right]^{1/p} \ln \left[c_2 \left(\frac{1}{\varepsilon} \right)^{p-1} \frac{1 - c_1 \varepsilon^p}{[1 - \varepsilon^{p^2} c_1^p]^{1/p}} \right] \right)^{1/8} + C\varepsilon^{1/4}.$$

Заметим, что по построению $x_2 \leq 1$, поэтому

$$\frac{1 - c_1 \varepsilon^p}{(1 - \varepsilon^{p^2} c_1^p)^{1/p}} \leq 1$$

(для доказательства нужно использовать неравенство $(1 - a)^p + a^p \leq 1$ при $a \in [0, 1]$). Кроме того,

$$(1 - \varepsilon^{p^2} c_1^p)^{1/p} < 1, \quad c_2 = \frac{1}{p^2} \left(\frac{p-1}{2} \right)^{p-1} < \left(\frac{p-1}{2} \right)^{p-1}.$$

Следовательно,

$$N_\varepsilon(L_p) \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4(p-1) \ln \frac{p-1}{2\varepsilon} \right)^{1/8} + C\varepsilon^{1/4}. \quad (10)$$

Рассмотрим еще два примера, когда значения x_2 и $f'(x_2)$ можно вычислить точно.

ПРИМЕР 2. Оценим количество узлов ломаной для функции $f(x) = -x \ln x$ на отрезке $[0, 1]$. На этом отрезке функция имеет экстремум в точке (e^{-1}, e^{-1}) . Согласно алгоритму (теорема 1) мы умеем приближать вогнутые возрастающие функции, поэтому оценим количество узлов ломаной на каждом отрезке монотонности отдельно.

а) Рассмотрим функцию $f(x) = -x \ln x$ на отрезке $[0, e^{-1}]$. Пусть $y = kx + \varepsilon$ — уравнение первого звена ломаной, $\bar{x}$ — точка, в которой достигается равенство $f(\bar{x}) - (k\bar{x} + \varepsilon) = \varepsilon$. Поскольку $y = kx + 2\varepsilon$ — касательная к графику функции $f(x)$ в точке $\bar{x}$, то $k = f'(\bar{x})$ и $2\varepsilon = f(\bar{x}) - \bar{x}f'(\bar{x})$. Так как $f'(x) = -\ln x - 1$, то $2\varepsilon = -\bar{x} \ln \bar{x} + \bar{x} \ln \bar{x} + \bar{x}$, т.е. $\bar{x} = 2\varepsilon$. Тогда $k = -\ln 2e\varepsilon$. Из условия $f(x_2) = kx_2$ определяем, что $x_2 = 2e\varepsilon$. Тогда $f'(x_2) = -\ln 2e^2\varepsilon$. Таким образом,

$$\begin{aligned} N_{1,\varepsilon} &\leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4f(e^{-1} - x_2) \ln \left[\frac{-\ln 2e^2\varepsilon(e^{-1} - 2e\varepsilon)}{-(e^{-1} - 2e\varepsilon) \ln(e^{-1} - 2e\varepsilon)} \right] \right)^{1/8} + C\varepsilon^{1/4} \\ &\leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(2 \ln \left[\frac{\ln 1/(2e^2\varepsilon)}{\ln 1/(e^{-1} - 2e\varepsilon)} \right] \right)^{1/8} + C\varepsilon^{1/4} \\ &\leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(2 \ln \ln \frac{1}{2e^2\varepsilon} \right)^{1/8} + C\varepsilon^{1/4}, \end{aligned}$$

где первое неравенство следует из того, что $2f(e^{-1} - x_2) \leq 2/e < 1$, а второе – из следующих соображений: так как $x_2 = 2e\varepsilon \leq e^{-1}$, то $1/(1 - 2e^2\varepsilon) \geq 1$, откуда

$$\frac{e}{1 - 2e^2\varepsilon} \geq e, \quad \ln \frac{1}{e^{-1} - 2e\varepsilon} = \ln \frac{e}{1 - 2e^2\varepsilon} \geq 1.$$

б) Оценим количество узлов ломаной на отрезке $[e^{-1}, 1]$. Для этого сделаем замену и рассмотрим вогнутую возрастающую функцию $f(x) = -(1-x) \ln(1-x)$ на отрезке $[0, 1 - e^{-1}]$. Тогда $f'(x) = \ln(1-x) + 1$, т.е., $f'(0) = 1$. Следовательно,

$$\begin{aligned} N_{2,\varepsilon} &\leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\frac{4}{e} \ln \frac{1 \cdot (1 - e^{-1})}{e^{-1}} \right)^{1/8} + C\varepsilon^{1/4} \\ &= \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\frac{4}{e} \ln(e - 1) \right)^{1/8} + C\varepsilon^{1/4} \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} + C\varepsilon^{1/4}. \end{aligned}$$

Таким образом, количество узлов ломаной, приближающей $f(x) = -x \ln x$ на отрезке $[0, 1]$, с точностью ε не превосходит

$$N_{en,\varepsilon} \leq 1 + \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(\left(2 \ln \ln \frac{1}{2e^2\varepsilon} \right)^{1/8} + 1 \right) + C\varepsilon^{1/4}. \quad (11)$$

В отличие от двух изложенных выше примеров, в следующем примере $f'(0) \neq 0$. Тем не менее, мы его здесь приведем, чтобы не проводить вычисления в следующем параграфе.

ПРИМЕР 3. Оценим количество узлов ломаной для вогнутой убывающей функции $f(x) = \ln(e - e^x)$ на отрезке $[0, \ln(e - 1)]$. Для этого сделаем замену и рассмотрим вогнутую возрастающую функцию $f(x) = \ln(e - e^{\ln(e-1)-x})$ на отрезке $[0, \ln(e - 1)]$. Тогда $f'(0) = e - 1$, $f(\ln(e - 1)) = \ln(e - 1)$. Следовательно,

$$N_{l,\varepsilon} \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} \left(4 \ln(e - 1) \ln \frac{(e - 1) \ln(e - 1)}{\ln(e - 1)} \right)^{1/8} + C\varepsilon^{1/4} \leq \frac{1}{\sqrt{\varepsilon}} \cdot \frac{1}{4\sqrt{2}} + C\varepsilon^{1/4}. \quad (12)$$

ЗАМЕЧАНИЕ 5. Заметим, что число узлов кусочно линейных функций, приближающих подграфики функций, рассмотренных в примерах 1–3, с точностью ε не превосходит соответственно $N_{\varepsilon/2}(L_p)$, $N_{en,\varepsilon/2}$, $N_{l,\varepsilon/2}$.

§ 3. Линеаризация функций многих переменных

При помощи достаточно естественного алгоритма (теорема 1) мы можем приблизить подграфик вогнутой (или надграфик выпуклой) функции одной переменной решением системы линейных неравенств для двух переменных, а также оценить количество таких неравенств.

Непосредственно обобщить эту задачу для приближения надграфика выпуклой функции в $\mathbb{R}^n$ решением системы линейных неравенств для n переменных не представляется возможным, поскольку число неравенств оценивается как $C\varepsilon^{(1-n)/2}$ (см., например, [7], [8]), что очень велико при больших значениях n ,

и эта оценка неувлучшаема уже для единичного шара. Поэтому если представлять многогранник в виде решения системы линейных неравенств от $n + 1$ переменной, то количество неравенств будет расти как $\varepsilon^{(1-n)/2}$ (ε – точность приближения); кроме того, задача нахождения такого многогранника алгоритмически сложна.

Поэтому предложим другой подход к решению этой задачи. Будем приближать надграфик выпуклой функции проекцией многогранника, принадлежащего пространству большей размерности (т.е. решением системы неравенств для большего, чем n , числа переменных).

Обозначим через $\bar{x}$ набор координат $(x_1, \dots, x_n)$.

ОПРЕДЕЛЕНИЕ. Пусть $f: \mathbb{R}^n \rightarrow \mathbb{R}$ – выпуклая функция. Будем говорить, что проекция многогранника $G \subset \mathbb{R}^k$, $k > n$, приближает надграфик функции $f(x)$ на выпуклом множестве $V \subset \mathbb{R}$ снизу с точностью ε , если выполняются следующие условия:

- 1) для любой точки $(\bar{x}, y) \in \text{epi } f$ такой, что $\bar{x} \in V$, существует набор $\bar{z} = (z_1, \dots, z_{k-n-1})$, для которого точка с координатами $(\bar{x}, y, \bar{z})$ принадлежит G ;
- 2) для любой точки $(x_1, \dots, x_k) \in G$ ее первые $n + 1$ координат таковы, что $x_{n+1} \geq f(\bar{x}) - \varepsilon$.

Далее будем считать, что многогранник определяется решением системы линейных неравенств и, кроме того, он может быть неограниченным.

В некоторых случаях такой подход позволяет свести задачу приближения надграфика выпуклой функции многих переменных к одномерной задаче линеаризации (которую мы умеем эффективно решать). Например, такое возможно в случае, когда рассматриваемая функция есть симметричная функция координат, или когда она раскладывается в прямую сумму выпуклых функций одной переменной.

Аппроксимация L_p -нормы в $\mathbb{R}^n$. Рассмотрим приближение надграфика функции

$$f_p(x) = (x_1^p + \dots + x_n^p)^{1/p}, \quad p > 1,$$

на множестве $V = \{(x_1, \dots, x_n) : x_1^p + \dots + x_n^p \leq 1, x_i \geq 0\}$ проекцией многогранника.

Нужно построить многогранник G , проекция которого приближает надграфик $f_p(x)$. Исследуем сначала случай функции двух переменных: приблизим надграфик функции $f(x) = (x_1^p + x_2^p)^{1/p}$ с точностью ε проекцией многогранника $\tilde{G} \subset \mathbb{R}^3$.

Рассмотрим новые переменные $\tilde{x}_1 = x_1/y$ и $\tilde{x}_2 = x_2/y$, $y \geq 0$, и множество точек $(\tilde{x}_1, \tilde{x}_2)$ на плоскости, удовлетворяющих неравенству $1 \geq (\tilde{x}_1^p + \tilde{x}_2^p)^{1/p}$. Найдем систему линейных уравнений от двух переменных такую, что для любого ее решения $(\tilde{x}_1, \tilde{x}_2)$ выполняется неравенство $1 + \varepsilon \geq (\tilde{x}_1^p + \tilde{x}_2^p)^{1/p}$ и любая точка $(\tilde{x}_1, \tilde{x}_2)$ такая, что $1 \geq (\tilde{x}_1^p + \tilde{x}_2^p)^{1/p}$, есть решение этой системы. Заметим, что при $\tilde{x}_1 \in [0, 1]$ эта задача совпадает с задачей приближения надграфика вогнутой функции $\tilde{x}_2 = (1 - \tilde{x}_1^p)^{1/p}$ (пример 1). Для этого необходимо не более чем $N_{\varepsilon/2}(L_p)$ неравенств, где $N_\varepsilon(L_p)$ удовлетворяет оценке (10).

Возьмем произвольное неравенство $a_{i1}\tilde{x}_1 + a_{i2}\tilde{x}_2 \leq b_i$ найденной системы. Сделав обратную замену, получим, что $a_{i1}(x_1/y) + a_{i2}(x_2/y) \leq b_i$. Тогда $a_{i1}x_1 + a_{i2}x_2 \leq b_i y$. Рассмотрим систему из не более чем $N_{\varepsilon/2}(L_p) + 2$ линейных неравенств:

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 \leq b_i y, \\ y \leq 1, \\ y \geq 0. \end{cases}$$

Она задает многогранник в $\mathbb{R}^3$, проекция которого приближает надграфик $f(x)$ снизу с точностью ε .

В случае $n \geq 3$ переменных будем действовать по индукции. Предположим, что для функции от менее, чем n переменных задача решена. Построим многогранник, приближающий надграфик $y \geq (x_1^p + \dots + x_n^p)^{1/p}$ с точностью ε следующим образом. Введем вспомогательную переменную t . Согласно предположению индукции мы можем построить многогранник $G_1 \subset \mathbb{R}^k$, приближающий надграфик $t \geq (x_1^p + \dots + x_{n-1}^p)^{1/p}$ с точностью ε_1 , и $G_2 \subset \mathbb{R}^3$, приближающий надграфик $y \geq (t^p + x_n^p)^{1/p}$ с точностью ε_2 . Система линейных неравенств, задающих G_1 и G_2 , определяет $G \subset \mathbb{R}^{k+1}$, приближающий надграфик $y \geq (x_1^p + \dots + x_n^p)^{1/p}$ с точностью ε .

Поскольку на каждом шаге мы добавляем одну вспомогательную переменную, то надграфик $y \geq (x_1^p + \dots + x_n^p)^{1/p}$ приближается системой линейных неравенств от $n + 2$ переменных, т.е. $k = n + 2$, $n \geq 3$.

Покажем, что можно выбрать $\varepsilon_1 = \varepsilon(n - 2)/(n - 1)$ и $\varepsilon_2 = \varepsilon/(n - 1)$. Для любой точки с координатами $(x_1, \dots, x_n, y, t)$, принадлежащей G , выполняются неравенства $y \geq (t^p + x_n^p)^{1/p} - \varepsilon_2$ и $t \geq (x_1^p + \dots + x_{n-1}^p)^{1/p} - \varepsilon_1$, откуда

$$y \geq [(x_1^p + \dots + x_{n-1}^p)^{1/p} - \varepsilon_1]^p + x_n^p)^{1/p} - \varepsilon_2.$$

Обозначим $a = (x_1^p + \dots + x_{n-1}^p)^{1/p}$. Тогда выражение $((a - \varepsilon_1)^p + x_n^p)^{1/p}$ есть L_p -норма вектора $(a - \varepsilon_1, x_n)$. Воспользуемся неравенством треугольника:

$$((a - \varepsilon_1)^p + x_n^p)^{1/p} = \|(a - \varepsilon_1, x_n)\|_p \geq \|(a, x_n)\|_p - \|(\varepsilon_1, 0)\|_p = (a^p + x_n^p)^{1/p} - \varepsilon_1.$$

Следовательно, $y \geq (x_1^p + \dots + x_n^p)^{1/p} - \varepsilon_1 - \varepsilon_2$. Таким образом, количество неравенств системы, задающей G , не превосходит $(n - 1)N_{\varepsilon/(2(n-1))}(L_p)$, где $N_{\varepsilon}(L_p)$ оценивается сверху по формуле (10). Теперь L_p -норма на всем пространстве $\mathbb{R}^n$ приближается проекцией многогранника $G' = \{u \in \mathbb{R}^d \mid |U_i| \leq x_i, i = 1, \dots, n; x \in G, x \geq 0\}$.

Таким образом, L_p -норма на множестве V в $\mathbb{R}^n$ приближается данным алгоритмом с точностью ε решением системы не более чем $(n - 1)N_{\varepsilon/(2(n-1))}(L_p)$ линейных неравенств для $n + 2$ переменных. Соответственно, единичный шар L_p -нормы приближается проекцией многогранника из $\mathbb{R}^{n+2}$, заданного не более чем $4(n - 1)N_{\varepsilon/(2(n-1))}(L_p)$ неравенствами, т.е. число неравенств не превосходит

$$C(p) \frac{n^{3/2}}{\sqrt{\varepsilon}} \left(\ln \frac{n}{\varepsilon} \right)^{1/8} + Cn^{3/4}\varepsilon^{1/4},$$

где $C(p)$ – константа, зависящая от p .

Такой подход позволяет свести многие экстремальные задачи, связанные с L_p -нормой, к задачам линейного программирования. Также заметим, что, используя полученное приближение, можно свести задачу поиска минимума выпуклой функции на L_p -шаре к задаче поиска минимума выпуклой функции на многограннике.

Приближенное вычисление L_p -нормы матрицы. Рассмотрим задачу вычисления L_p -нормы матрицы A , определяемой следующим образом:

$$\|A\|_p = \sup_{x \in \mathbb{R}^n} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p \leq 1} \|Ax\|_p.$$

Заметим, что эта задача NP-сложна уже при $p = 4$. Предложим возможный способ приближенного вычисления $\|A\|_p$, использующий приближение единичного шара L_p -нормы проекцией многогранника.

Обозначим через $\mathcal{B}_p = \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$ единичный шар в $\mathbb{R}^n$. Рассмотрим задачу:

$$\begin{cases} r \rightarrow \min \\ A\mathcal{B}_p \subset r\mathcal{B}_p. \end{cases}$$

Решение r_0 этой задачи есть в точности $\|A\|_p$. Решим ее приближенно. Для этого научимся определять, верно ли, что $A\mathcal{B}_p \subset r\mathcal{B}_p$ для произвольного $r > 0$. Тогда методом деления пополам мы сможем с нужной точностью найти r_0 .

Итак, для данного $r > 0$ необходимо определить, существует ли точка $x \in \mathbb{R}^n$ такая, что $x \in A\mathcal{B}_p$, но $x \notin r\mathcal{B}_p$. Пусть G_0 – многогранник, проекция которого приближает единичный шар L_p -нормы с точностью ε , G_1 – многогранник, проекция которого приближает $f_p(x)$ на множестве

$$V = \{(x_1, \dots, x_d) : x_1^p + \dots + x_d^p \leq \|A\|_\infty\}$$

с точностью ε . Тогда условия $x \in A\mathcal{B}_p$ и $x \in r\mathcal{B}_p$ можно заменить такими: $z_x \in G_1$ и $s_x \in rG_0$ для векторов s_x и z_x , соответствующих x . Пусть G_0 задается системой линейных неравенств $(a_i, s_x) \leq b_i, i = 1, \dots, N$. Тогда нужно рассмотреть N систем вида

$$\begin{cases} z_x \in G_1, \\ (a_i, s_x) > rb_i, \end{cases}$$

и для каждой из них определить, имеет ли она решение. Если хотя бы одна система имеет решение, то $x \in A\mathcal{B}_p$ и $x \notin r\mathcal{B}_p$.

Заметим, что рассматриваются системы от $n + 2$ переменных. Количество таких систем не превосходит $4(n - 1)N_{\varepsilon/(2(n-1))}(L_P)$ и количество неравенств в каждой системе не превосходит

$$4(n - 1)N_{\varepsilon/(2(n-1))}(L_P) + 4(n - 1)N_{\varepsilon/(2(n-1))\sqrt{\|A\|_\infty}}(L_P),$$

т.е. число неравенств не превосходит

$$C(\|A\|_\infty, p) \frac{n^{3/2}}{\sqrt{\varepsilon}} \left(\ln \frac{n}{\varepsilon} \right)^{1/8} + Cn^{3/4}\varepsilon^{1/4},$$

где $C(\|A\|_\infty, p)$ – константа, зависящая от A и p .

Используя изложенный алгоритм приближенного вычисления L_p -нормы матрицы, мы можем решить трудную задачу (см., например, [18], [19]) приближенного вычисления числа обусловленности $\sigma_p = \|A\|_p \cdot \|A^{-1}\|_p$ матрицы A в L_p -норме. При этом мы считаем, что A и A^{-1} нам уже даны, т.е. задачу обращения матрицы решать не нужно.

Аппроксимация функции $f(x) = \ln(e^{x_1} + \dots + e^{x_n})$. Приближим надграфик функции $f(x) = \ln(e^{x_1} + \dots + e^{x_n})$ проекцией многогранника на множестве

$$V = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_i, \ln\left(\sum_{i=1}^n e^{x_i}\right) \leq 1 \right\}.$$

Начнем со случая $n = 2$ и приближим надграфик функции $f(x) = \ln(e^{x_1} + e^{x_2})$ с точностью ε проекцией многогранника $\tilde{G} \subset \mathbb{R}^3$.

Рассмотрим новые переменные $\tilde{x}_1 = x_1 - y$ и $\tilde{x}_2 = x_2 - y$, $y \geq 0$, и множество точек $(\tilde{x}_1, \tilde{x}_2)$ на плоскости, удовлетворяющих неравенству $1 \geq \ln(e^{\tilde{x}_1} + e^{\tilde{x}_2})$. Найдем систему линейных уравнений от двух переменных такую, что для любого ее решения $(\tilde{x}_1, \tilde{x}_2)$ выполняется неравенство $1 + \varepsilon \geq \ln(e^{\tilde{x}_1} + e^{\tilde{x}_2})$ и любая точка $(\tilde{x}_1, \tilde{x}_2)$ такая, что $1 \geq \ln(e^{\tilde{x}_1} + e^{\tilde{x}_2})$, есть решение этой системы. Заметим, что при $\tilde{x}_1 \in [0, \ln(e - 1)]$ эта задача совпадает с задачей приближения подграфика вогнутой функции $\tilde{x}_2 = \ln(e - e^{\tilde{x}_1})$ (пример 3); для этого необходимо не более чем $N_{l,\varepsilon/2}$ неравенств, где $N_{l,\varepsilon}$ оценивается сверху по формуле (12).

Рассмотрим произвольное неравенство $a_{i1}\tilde{x}_1 + a_{i2}\tilde{x}_2 \leq b_i$ найденной системы. Сделав обратную замену, получим $a_{i1}x_1 + a_{i2}x_2 \leq b_i + (a_{i1} + a_{i2})y$. Рассмотрим систему из не более чем $N_{l,\varepsilon/2} + 2$ линейных неравенств:

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 \leq b_i + (a_{i1} + a_{i2})y, \\ y \leq 1, \\ y \geq 0. \end{cases}$$

Она задает многогранник в $\mathbb{R}^3$, проекция которого приближает надграфик $f(x)$ снизу с точностью ε .

В случае n переменных будем действовать по индукции. Предположим, что для функции от менее чем n переменных задача решена. Тогда построим многогранник, приближающий надграфик $y \geq \ln(e^{x_1} + \dots + e^{x_n})$ с точностью ε следующим образом. Введем вспомогательную переменную t . Согласно предположению индукции мы можем построить многогранник $G_1 \subset \mathbb{R}^k$, приближающий надграфик $t \geq \ln(e^{x_1} + \dots + e^{x_{n-1}})$ с точностью ε_1 , и $G_2 \subset \mathbb{R}^3$, приближающий надграфик $y \geq \ln(e^t + e^{x_n})$ с точностью ε_2 . Система линейных неравенств, задающих G_1 и G_2 определяет $G \subset \mathbb{R}^{k+1}$, приближающий надграфик $y \geq \ln(e^{x_1} + \dots + e^{x_n})$ с точностью ε . Поскольку на каждом шаге мы добавляем одну вспомогательную переменную, то надграфик $y \geq \ln(e^{x_1} + \dots + e^{x_n})$ приближается системой линейных неравенств от $n + 2$ переменных, т.е. $k = n + 2$, $n \geq 3$.

Покажем, что $\varepsilon = \varepsilon_1 + \varepsilon_2$, т.е. можно выбрать $\varepsilon_1 = \varepsilon(n - 2)/(n - 1)$ и $\varepsilon_2 = \varepsilon/(n - 1)$. Действительно, для любой точки с координатами $(x_1, \dots, x_n, y, t)$,

принадлежащей G , выполняются неравенства $y \geq \ln(e^t + e^{x_n}) - \varepsilon_2$ и $t \geq \ln(e^{x_1} + \dots + e^{x_{n-1}}) - \varepsilon_1$, откуда

$$\begin{aligned} y &\geq \ln\left(\frac{e^{x_1} + \dots + e^{x_{n-1}}}{e^{\varepsilon_1}} + e^{x_n}\right) - \varepsilon_2 = \ln(e^{x_1} + \dots + e^{x_{n-1}} + e^{\varepsilon_1} \cdot e^{x_n}) - \varepsilon_1 - \varepsilon_2 \\ &\geq \ln(e^{x_1} + \dots + e^{x_n}) - \varepsilon_1 - \varepsilon_2. \end{aligned}$$

Таким образом, количество неравенств системы, задающей многогранник G , не превосходит $(n-1)N_{l,\varepsilon/(2(n-1))}$, где $N_{l,\varepsilon}$ оценивается сверху по формуле (12), т.е. число неравенств не превосходит $n^{3/2}/(4\sqrt{\varepsilon}) + Cn^{3/4}\varepsilon^{1/4}$.

Аппроксимация функции энтропии. В следующей многомерной задаче сведение к одномерной задаче осуществляется за счет того, что рассматриваемая функция есть сумма выпуклых функций от координат.

Решим задачу приближения проекцией многогранника надграфика функции $f(x) = x_1 \ln x_1 + \dots + x_n \ln x_n$ с точностью ε на множестве $V = \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_i \leq 1\}$.

Для этого воспользуемся следующим фактом. Пусть $f(x) = f_1(x_1) + \dots + f_n(x_n)$, где $f_i(x)$ – выпуклая функция. При помощи алгоритма (теорема 1) мы можем приблизить надграфик каждой функции $y_i \geq f_i(x_i)$ с точностью ε_i проекцией многогранника G_i (в данном случае G_i – многоугольник). Покажем, как построить многогранник G , проекция которого приближает надграфик функции $f(x) = f_1(x_1) + f_2(x_2)$. Пусть G_i определяется системой линейных неравенств

$$A_i \begin{pmatrix} x_i \\ y_i \end{pmatrix} \leq b_i, \quad i = 1, 2.$$

Введем дополнительную переменную t и рассмотрим систему неравенств

$$\begin{cases} A_i \begin{pmatrix} x_i \\ y_i \end{pmatrix} \leq b_i, & i = 1, 2, \\ t \geq y_1 + y_2. \end{cases}$$

Она определяет многогранник $G \subset \mathbb{R}^5$, проекция которого приближает надграфик $f(x) = f_1(x_1) + f_2(x_2)$ с точностью $\varepsilon = \varepsilon_1 + \varepsilon_2$.

Рассуждая аналогично, по индукции можно построить многогранник $G' \subset \mathbb{R}^{3n-1}$, проекция которого приближает надграфик функции $f(x) = f_1(x_1) + \dots + f_n(x_n)$ с точностью $\varepsilon = \varepsilon_1 + \dots + \varepsilon_n$.

Таким образом, для того чтобы приблизить надграфик функции $f(x) = x_1 \ln x_1 + \dots + x_n \ln x_n$ с точностью ε на множестве $V = \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_i \leq 1\}$, необходимо приблизить надграфик функции $y = x \ln x$ с точностью ε/n . Согласно примеру 2 для этого необходимо не более чем $N_{en,\varepsilon/(2n)}$ неравенств, где $N_{en,\varepsilon}$ оценивается сверху по формуле (11). Тогда надграфик функции $f(x) = \sum_{i=1}^n x_i \ln x_i$ на множестве V приближается с точностью ε проекцией многогранника из пространства $\mathbb{R}^{3n-1}$, заданного системой из не более чем $nN_{en,\varepsilon/(2n)} + n - 1$ неравенств, т.е. число неравенств не превосходит $2n + (n/(4\sqrt{\varepsilon}))(\ln \ln 1/\varepsilon)^{1/8} + Cn\varepsilon^{1/4}$.

В заключение выражаю благодарность профессору В. Ю. Протасову за внимание к работе и ценные замечания.

Список литературы

- [1] S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge Univ. Press, Cambridge, 2004.
- [2] Yu. Nesterov, A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, SIAM Stud. Appl. Math., **13**, SIAM, Philadelphia, PA, 1994.
- [3] Yu. Nesterov, *Introductory lectures on convex optimization*, Appl. Optim., **87**, Kluwer Acad. Publishers, Boston, MA, 2004.
- [4] P. M. Gruber, “Error of asymptotic formulae for volume approximation of convex bodies in $\mathbb{E}^d$ ”, *Monatsh. Math.*, **135**:4 (2002), 279–304.
- [5] C. Schütt, E. Werner, “Random polytopes with vertices on the boundary of a convex body”, *C. R. Acad. Sci. Paris Sér. I Math.*, **331**:9 (2000), 697–701.
- [6] R. Schneider, “Polyhedral approximation of smooth convex bodies”, *J. Math. Anal. Appl.*, **128**:2 (1987), 470–474.
- [7] M. Ludwig, C. Schütt, E. Werner, “Approximation of the Euclidean ball by polytopes”, *Studia Math.*, **173**:1 (2006), 1–18.
- [8] L. Chen, “New analysis of the sphere covering problems and optimal polytope approximation of convex bodies”, *J. Approx. Theory*, **133**:1 (2005), 134–145.
- [9] J. Bourgain, J. Lindenstrauss, V. Milman, “Approximation of zonoids by zonotopes”, *Acta Math.*, **162**:1 (1989), 73–141.
- [10] В. Ю. Протасов, “Обобщенный совместный спектральный радиус. Геометрический подход”, *Изв. РАН. Сер. матем.*, **61**:5 (1997), 99–136; англ. пер.: V. Yu. Protasov, “The generalized joint spectral radius. A geometric approach”, *Izv. Math.*, **61**:5 (1997), 995–1030.
- [11] M. A. Lopez, Sh. Reisner, “Hausdorff approximation of convex polygons”, *Comput. Geom.*, **32**:2 (2005), 139–158.
- [12] J. S. Müller, “Step by step approximation of plane convex bodies”, *Arch. Math. (Basel)*, **58**:6 (1992), 606–610.
- [13] R. Schneider, “Affine-invariant approximation by convex polytopes”, *Studia Sci. Math. Hungar.*, **21**:3–4 (1986), 401–408.
- [14] D. E. McClure, R. A. Vitale, “Polygonal approximation of plane convex bodies”, *J. Math. Anal. Appl.*, **51**:2 (1975), 326–358.
- [15] В. Ю. Протасов, “Совместный спектральный радиус и инвариантные множества линейных операторов”, *Фундамент. и прикл. матем.*, **2**:1 (1996), 205–231.
- [16] А. А. Лигун, В. Ф. Сторчай, “О наилучшем выборе узлов при приближении сплайнами в метрике L_p ”, *Матем. заметки*, **20**:4 (1976), 611–618; англ. пер.: A. A. Ligun, V. F. Storchai, “On the best choice of nodes for approximation by splines in the metric of L_p ”, *Math. Notes*, **20**:4 (1976), 899–903.
- [17] А. Д. Иоффе, В. М. Тихомиров, *Теория экстремальных задач*, Наука, М., 1974; англ. пер.: A. D. Ioffe, V. M. Tihomirov, *Theory of extremal problems*, Stud. Math. Appl., **6**, North-Holland, Amsterdam–New York, 1979.
- [18] D. S. Watkins, *Fundamentals of matrix computations*, Pure Appl. Math. (N. Y.), Wiley, New York, 2002.
- [19] G. H. Golub, Ch. F. Van Loan, *Matrix computations*, Johns Hopkins Stud. Math. Sci., Johns Hopkins Univ. Press, Baltimore, MD, 1996.

Е. С. Горская (E. S. Gorskaya)

Механико-математический факультет
 Московского государственного университета
 им. М. В. Ломоносова
 E-mail: allena@mccme.ru

Поступила в редакцию
 23.06.2009