Серия «Математика»
2015. Т. 11. С. 54—68

Онлайн-доступ к журналу:
http://isu.ru/izvestia

ИЗВЕСТИЯ
Иркутского
государственного
университета

УДК 519.853.3

Применение модифицированного метода симплексных погружений для решения специального класса задач выпуклой недифференцируемой оптимизации *

А. В. Колосницын
Институт систем энергетики им. Л. А. Мелентьева СО РАН

Аннотация. Рассматривается модифицированный метод симплексных погружений, который относится к классу методов центрированных сечений. Особенностью метода является оценка скорости сходимости, которая зависит только от числа отсеченных вершин симплекса построенной секущей плоскостью. Чем больше вершин отсекает секущая плоскость, тем выше скорость сходимости метода. Модифицированный метод симплексных погружений, снабженный данным критерием выбора секущей плоскости, используется для решения специального класса задач выпуклой недифференцируемой оптимизации, который состоит из двух типов функций. Для возможности формировать секущую плоскость, отсекающую наибольшее число вершин симплекса, возникает необходимость в описании субдифференциала функции, зависящего от одного или нескольких параметров, по которым можно провести оптимизацию. С этой целью приводится описание субдифференциалов функций из введенного класса задач в параметрическом виде, что позволяет формировать вспомогательные минимаксные задачи для поиска результирующих секущих плоскостей, отсекающих наибольшее число вершин симплекса, и сокращает количество итераций метода симплексных погружений. Приводятся результаты численного эксперимента.

Ключевые слова: модифицированных метод симплексных погружений, субдифференциал выпуклой функции, результирующая секущая плоскость.

1. Введение

Для решения задач выпуклой недифференцируемой оптимизации нам доступен достаточно обширный арсенал различных методов, информацию о которых можно найти, например, в [3; 4; 5; 6; 7; 8; 9].

* Работа выполнена при частичной финансовой поддержке РФФИ, грант 15-07-08986.

Исторически первыми можно выделить субградиентные методы, которые представляют собой прямой аналог градиентных методов с заменой градиента функции на произвольный субградиент. При этом не удается достичь монотонного убывания значений минимизируемой функции, однако важно, что убывает другая функция — расстояние до точки минимума. Еще одна особенность метода заключается в правиле выбора длины шага. Его нужно уменьшать, например, задавая априори стремящуюся к нулю последовательность. Главным достоинством таких методов можно считать их чрезвычайную простоту, если известен простой способ вычисления субградиента. Однако, на практике они чаще всего оказываются бесполезными, обладая низкой сублинейной скоростью сходимости и не имея достаточно надежного критерия остановки. Тем не менее субградиентные методы используются чаще всего тогда, когда достаточно найти очень грубое приближение к решению.

Возможность ввести разумное правило остановки дает группа методов кусочно-линейной аппроксимации. Их базовая идея заключается в использовании информации, накапливаемой по ходу итераций, для построения все более точной кусочно-линейной аппроксимации снизу целевой функции. К недостаткам данных методов можно отнести необходимость решать задачу линейного программирования, в которой с ростом числа итераций возрастает и число ограничений. Кроме того, за исключением некоторых классов задач, например, задач с острым минимумом (острый минимум здесь понимается в смысле определения, данного в [7], стр. 127), в общем случае методы кусочно-линейной аппроксимации могут сходиться очень медленно. Еще один существенный недостаток данных методов связан с их неустойчивостью, что выражается в хаотичном поведении последовательностей значений целевой функции $\{f(x^k)\}$ и последовательности приближений $\{x^k\}$.

Эта проблема может быть решена в многошаговых методах с квадратичными подзадачами, которые основаны на методах кусочно-линейной аппроксимации и дополнены необходимым стабилизирующим механизмом. Стабилизация осуществляется за счет введения в модель квадратичного члена, который призван обеспечить близость получаемого решения x^{k+1} к предыдущему приближению, избегая чрезмерных скачков в последовательности $\{x^k\}$. Не вдаваясь в излишние детали, отметим, что для некоторых важных классов негладких выпуклых функций в настоящее время вполне возможно получать сверхлинейно сходящиеся методы [3].

И, наконец, выделим еще одну группу методов, основанных на понятии «центра» допустимого множества. Суть методов так называемых центрированных сечений заключается в определении центра допустимого множества и последующем проведении через него отсекающей гиперплоскости, после чего часть множества, содержащая решение, вновь подвергается операции нахождения центра и проведения через него отсечения. Повторение таких шагов метода осуществляется до тех пор, пока мы не найдем решение с заданной точностью.

В [7] показано, что наиболее эффективным с точки зрения отсечения части множества, не содержащей решения, является метод центров тяжести, однако нахождение центра тяжести множества представляет собой весьма трудную задачу и зачастую не может быть реализовано. Поэтому задача поиска центра множества может быть сведена к нахождению центра более простых аппроксимирующих множеств. К наиболее важным методам таких типов относятся методы вписанных и описанных эллипсоидов, метод аналитического центра, метод объемного центра. Данные методы обладают полиномиальной вычислительной сложностью [5].

В данной статье предлагается модификация метода симплексных погружений, относящегося к классу методов центрированных сечений. Его главная особенность и исследовательский интерес автора связаны с полученной в [1] оценкой скорости сходимости, которая открывает достаточно богатый выбор для создания различных модификаций по ускорению работы метода. В статье рассматривается специальный класс задач, состоящий из комбинаций двух типов выпуклых недифференцируемых функций, для которых приводится описание субдифференциалов в параметрическом виде. Показано, как с помощью вспомогательных минимаксных задач, основанных на параметрическом описании субдифференциалов, формируются результирующие секущие плоскости, отсекающие наибольшее число вершин симплекса, что позволяет ускорить поиск решения рассматриваемой задачи.

2. Постановка задачи и основные определения

Рассмотрим следующую задачу выпуклой оптимизации

$$\begin{aligned} &f_0(x) \to \min, \\ &f_i(x) \le 0, \end{aligned} \tag{2.1}$$

где $x \in R^n$, $f_i(x)$, $i = 1, ..., m$ — выпуклые, не обязательно гладкие функции.

Основная идея метода симплексных погружений заключается в следующем. На k-м шаге допустимое множество решений поставленной задачи погружается в симплекс S_k. Находится центр данного симплекса $x^{c,\,k}$, через который проводится отсекающая плоскость вида $L = g^T(x - x^{c,\,k}) = 0$. Затем часть симплекса, содержащая решение, погружается в новый симплекс минимального объема S_{k+1}. Повторяя такую процедуру, мы строим новые симплексы меньшего объема, последовательно локализуя решение, и останавливаемся, когда объем симплекса становится достаточно малым.

Введем ряд важных определений, которые нам необходимы для более детального изучения метода симплексных погружений.

Определение 1. *Симплексом $S \subset R^n$ с вершиной в точке x^0 и ребрами $(x^1 - x^0, ..., x^n - x^0)$, образующими базис в R^n, называется множество*

$$S = \left\{ x \in R^n : x = x^0 + \sum_{i=1}^{n} \sigma_i \left(x^i - x^0 \right), \ \sigma_i \geq 0, \ \sum_{i=1}^{n} \sigma_i \leq 1 \right\}.$$

Определение 2. *Центром симплекса S называется точка x^c, которая находится по следующей формуле*

$$x^c = \frac{1}{n+1} \left(x^1 + ... + x^n \right).$$

Определение 3. *Объем симплекса S определяется формулой*

$$V(S) = \frac{1}{n!} \left| det \left(\bar{X} \right) \right|,$$

где $\bar{X}$ — матрица размеров $n \times n$, столбцы которой есть суть векторы $x^1, x^2, ..., x^n$.

Определение 4. *Любую плоскость вида*

$$L = \left\{ x : g^T (x - x^c) = 0 \right\},$$

проходящую через центр x^c симплекса S будем называть секущей плоскостью.

Определение 5. *Будем говорить, что вершина x^i симплекса S не отсекается плоскостью L, если*

$$\alpha_i = g^T \left(x^i - x^c \right) < 0,$$

и отсекается, если

$$\alpha_i \geq 0.$$

3. Построение симплекса минимального объема и оценка скорости сходимости

Отметим несколько ключевых особенностей метода симплексных погружений. Важным принципом его работы является процедура погружения усеченного симплекса в новый симплекс минимального объема, что обеспечивает сходимость метода к решению задачи. Исчерпывающую информацию о построении симплекса минимального объема, содержащего заданный усеченный симплекс, можно найти в [1]. Мы же подробно остановимся на оценке скорости сходимости метода, которая также сформулирована в [1] в виде следующей теоремы.

Теорема. *Пусть $S \subset R^n$ – есть n-мерный симплекс, x^c – его центр, $S_G = \left\{x \in S,\ g^T(x - x^c) \le 0\right\}$ – усеченный симплекс. Симплекс S_G всегда можно погрузить в симплекс S^* такой, что для объемов $V(S)$ и $V(S^*)$ симплексов S и S^* будет выполняться неравенство*

$$q_k^* = \frac{V(S^*)}{V(S)} \le \begin{cases} \frac{1}{2} & k = 1; \\ \left(\frac{k}{k+1}\right)^k \left(\frac{k}{k-1}\right)^{k-1}, & 2 \le k \le n, \end{cases} \tag{3.1}$$

где k – число сохраненных при отсечении вершин симплекса.

Оценка скорости сходимости (3.1) зависит только от числа отсеченных вершин симплекса, причем, при отсечении n вершин симплекса мы получаем аналог метода дихотомии. На основе данной оценки была разработана модификация метода, представим ее подробное описание в следующей главе.

4. Модифицированный метод симплексных погружений

Сделаем небольшое замечание. В дальнейшем под базовым методом симплексных погружений будем понимать метод и соответствующий алгоритм, описанный в работе [1], не использующий формирование вспомогательных задач для построения результирующей секущей плоскости.

Идея модификации метода симплексных погружений, описанная в [2], заключается в учете сразу нескольких секущих плоскостей. Пусть центр симплекса, полученный на данной итерации, не является допустимым решением задачи (2.1) и в нем нарушаются $l \le m$ ограничений. Тогда можно построить l секущих плоскостей, соответствующих нарушаемым ограничениям с целью отсечь как можно больше вершин симплекса.

Сложность такого подхода заключается в том, что на данный момент нет метода построения симплекса минимального объема, содержащего усеченный симплекс, полученный в результате введения l секущих плоскостей ($l \ge 2$). Однако, следующие результаты из [2] дают нам возможность рассматривать линейную комбинацию из нормалей таких секущих плоскостей и получать одну результирующую секущую плоскость.

Предположим, что на i-й итерации метода сиплексных погружений можно провести l секущих плоскостей с нормалями $a^i, i = 1, \ldots, l$. Будем искать нормаль результирующей секущей плоскости b в виде линейной комбинации $\sum_{i=1}^{l} \lambda_i a^i$ нормалей a^i. Тогда, если существует вектор

$\bar{\lambda} = \left(\bar{\lambda_1}, ..., \bar{\lambda}_l\right)$, который является решением системы

$$\begin{gathered} \textstyle\sum_{i=1}^{l} \lambda_i \left(a^i\right)^T \left(h^j - x^c\right) \geq 0, \quad j = 1, ..., r, \\ \textstyle\sum_{i=1}^{l} \lambda_i = 1, \quad \lambda_i \geq 0, \end{gathered} \tag{4.1}$$

где x^c – центр текущего симплекса а r – число отсеченных вершин, то отсекающее полупространство

$$b^T(x - x^c) \leq 0, \quad b = \sum_{i=1}^{l} \overline{\lambda}_i a^i$$

отсекает то же число вершин, что и совокупность отсекающих полупространств.

Стоит отметить, что не всегда бывает возможно найти такой вектор $\overline{\lambda}$, который являлся бы решением системы (4.1). Поэтому в общем случае для построения результирующей секущей плоскости формируется следующая минимаксная задача:

$$\min_{\lambda} \max_{1 \leq j \leq r} \sum_{i=1}^{l} \lambda_i \left(a^i\right)^T \left(h^j - x^c\right), \quad \lambda_i \geq 0, \quad \sum_{i=1}^{l} \lambda_i = 1, \tag{4.2}$$

из решения которой и находится вектор $\overline{\lambda}$. Данная техника построения результирующей секущей плоскости справедлива и в том случае, когда центр симплекса попадает в точку, где целевая функция или функция, задающая ограничение задачи (2.1), является недифференцируемой. Тогда при условии, что мы можем параметрически описать субдифференциал соответствующей функции в данной точке, у нас есть возможность с помощью постановки задачи (4.2) построить результирующую секущую плоскость, которая отсекает наибольшее число вершин симплекса.

Приведенная идея построения результирующей секущей плоскости является ключевым элементом исследования автора данной статьи и положена в основу модифицированного метода симплексных погружений. Следующая глава содержит описание класса задач выпуклой недифференцируемой оптимизации, для которого можно построить субдифференциал в параметрическом виде в заданной точке.

5. Описание класса задач выпуклой недифференцируемой оптимизации

Рассмотрим два типа функций:

$$f^I(x) = \sum_{i=1}^{m} \alpha_i \left|\left(a^i\right)^T x - b_i\right|, \tag{5.1}$$

$$f^{II}(x) = \max_{1 \le i \le m} \left\{ \alpha_i \left| \left(a^i\right)^T x - b_i \right| \right\}, \tag{5.2}$$

где $\alpha \in R^m_+$, $a^i \in R^n$, $i = 1, ..., m$, $x \in R^n$, $b \in R^m$.

Определим класс задач выпуклой недифференцируемой оптимизации, который описывается функциями вида (5.1), (5.2):

$$\begin{cases} \phi_0(x) = \sum\limits_{i=1}^{m_0} f_i^I(x) + \sum\limits_{i=1}^{n_0} f_i^{II}(x) \to \min, \\ \\ \phi_k(x) = \sum\limits_{i=1}^{m_k} f_{ki}^I(x) + \sum\limits_{i=1}^{n_k} f_{ki}^{II}(x) \le 0, \quad k = 1, ..., N, \\ \\ x \in X, \end{cases} \tag{5.3}$$

где X — множество простой структуры.

Для возможности использовать модифицированный метод симплексных погружений при решении задачи вида (5.3) нам потребуется описание субдифференциалов функций (5.1) – (5.2). Напомним основные формулы для вычисления субдифференциала выпуклой функции [7].

Если $f_i(x)$, $i = 1, ..., m$ – выпуклые функции и $\alpha_i \ge 0$, тогда

$$\partial \left(\sum_{i=1}^{m} \alpha_i f_i(x) \right) = \sum_{i=1}^{m} \alpha_i \partial f_i(x),\ \alpha_i \ge 0.$$

Пусть $f(x) = \max\limits_{1 \le i \le m} f_i(x)$, $f_i(x)$, $i = 1, ..., m$ – выпуклые функции, тогда

$$\partial f(x) = co \left\{ \bigcup_{i \in I(x)} \partial f_i(x) \right\} =$$

$$\left\{ q : q = \sum_{i \in I(x)} \lambda_i g_i(x),\ \ g_i(x) \in \partial f_i(x),\ \ \lambda_i \ge 0,\ \ i \in I(x), \sum_{i \in I(x)} \lambda_i = 1 \right\},$$

$$I(x) = \left\{ i : f_i(x) = f(x) \right\}.$$

Отметим один важный принцип, которым мы будем руководствоваться при описании субдифференциалов функций (5.1) и (5.2). Класс задач выпуклой недифференцируемой оптимизации (5.3) описывается с помощью различных комбинаций функций, которые представляют собой модуль выпуклых функций и максимум из выпуклых функций. В качестве операций, производимых над данными функциями, используются их сложение и умножение на неотрицательные константы. Субдифференциал функции из такого класса укладывается в параметрическое описание путем его представления как выпуклой оболочки соответствующих субградиентов.

При вычислении субдифференциала функции $f^I(x)$ в заданной точке $\hat{x}$ первым действием, очевидно, будет представление модуля выпуклой функции как максимума из двух выпуклых функций:

$$\begin{gathered}\partial\left(\alpha_i\left|(a^i)^T\hat{x}-b_i\right|\right)=\\ \partial\left(\max\left\{-\alpha_i\left((a^i)^T\hat{x}-b_i\right),\alpha_i\left((a^i)^T\hat{x}-b_i\right)\right\}\right),\ \ i=1,...,m.\end{gathered}\tag{5.4}$$

Субдифференциал максимума выпуклых функций представляет собой выпуклую оболочку из объединения субдифференциалов активных функций:

$$\partial\left(\max\left\{-\alpha_i\left((a^i)^T\hat{x}-b_i\right),\alpha_i\left((a^i)^T\hat{x}-b_i\right)\right\}\right)=$$

$$=co\left\{\bigcup_{l\in L_i^1(\hat{x})}\alpha_{il}(a^i),\ \bigcup_{l\in L_i^2(\hat{x})}\alpha_{il}(a^i)\right\},\ \ i=1,...,m.$$

Как можно заметить, в последней записи используются объединения по множествам L_i^1, L_i^2, $i=1,...,m$. Это множества, которые либо пусты, либо состоят всего из одного элемента и показывают, какая из выпуклых функций определенного ранее максимума (5.4) из данных функций является активной. Определим структуру этих множеств:

$$L_i^1(\hat{x})=\left\{l:\max_{1\le l\le 2}\left\{\alpha_{il}\left(a^T\hat{x}-b_i\right)\right\}=\alpha_{i1}\left(a^T\hat{x}-b_i\right)=-\alpha_i\left(a^T\hat{x}-b_i\right)\right\},$$

$$L_i^2(\hat{x})=\left\{l:\max_{1\le l\le 2}\left\{\alpha_{il}\left(a^T\hat{x}-b_i\right)\right\}=\alpha_{i2}\left(a^T\hat{x}-b_i\right)=\alpha_i\left(a^T\hat{x}-b_i\right)\right\},$$

$$i=1,...,m.$$

Следующим шагом в описании субдифференциала функции $f^I(\hat{x})$ является переход от выпуклой оболочки из объединения субдифференциалов активных функций к всевозможным выпуклым комбинациям субградиентов из данных субдифференциалов:

$$co\left\{\bigcup_{l\in L_i^1(\hat{x})}\alpha_{il}(a^i),\ \bigcup_{l\in L_i^2(\hat{x})}\alpha_{il}(a^i)\right\}=$$

$$=\left\{q:q=\left(-\sum_{l\in L_i^1(\hat{x})}\lambda_{il}\alpha_i+\sum_{l\in L_i^2(\hat{x})}\lambda_{il}\alpha_i\right)a^i,\right.$$

$$\left.\lambda_{il}\ge 0,\ \ l\in L_i^1(\hat{x})\bigcup L_i^2(\hat{x}),\ \ \sum_{l\in L_i^1(\hat{x})\bigcup L_i^2(\hat{x})}\lambda_{il}=1\right\},\ \ i=1,...,m.$$

Чтобы иметь параметрическое представление субдифференциала функции $f^I(\hat{x})$, нам остается просуммировать по всем $i = 1, ..., m$ полученные субдифференциалы модуля выпуклых функций:

$$\partial f^I(\hat{x}) = \partial \left(\sum_{i=1}^m \alpha_i \left| (a^i)^T \hat{x} - b_i \right| \right) =$$

$$= \Big\{ q : q = \sum_{i=1}^m \Big(- \sum_{l \in L_i^1(\hat{x})} \lambda_{il} \alpha_i + \sum_{l \in L_i^2(\hat{x})} \lambda_{il} \alpha_i \Big) a^i,$$
$$\lambda_{il} \geq 0, \ \ l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x}), \ \ \sum_{l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x})} \lambda_{il} = 1, \tag{5.5}$$
$$i = 1, ..., m \Big\}.$$

Множество (5.5) определяет параметрическое описание субдифференциала функции $f^I(\hat{x})$.

Математические выкладки при описании субдифференциала функции $f^{II}(\hat{x})$ аналогичны тем, что мы использовали при описании субдифференциала функции $f^I(\hat{x})$ за исключением последней части. Вместо суммирования по всем $i = 1, ..., m$ мы еще раз берем выпуклую оболочку субдифференциалов активных функций. Распишем все преобразования поэтапно:

$$\partial f^{II}(\hat{x}) = \partial \left(\max_{1 \leq i \leq m} \left\{ \alpha_i \left| (a^i)^T \hat{x} - b_i \right| \right\} \right) =$$

$$= co \left\{ \bigcup_{i \in I(\hat{x})} \partial \left(\max \left\{ -\alpha_i((a^i)^T \hat{x} - b_i), \ \alpha_i((a^i)^T \hat{x} - b_i) \right\} \right) \right\} =$$

$$= co \left\{ \bigcup_{i \in I(\hat{x})} co \left\{ \bigcup_{l \in L_i^1(\hat{x})} \alpha_{li} a^i, \ \bigcup_{l \in L_i^2(\hat{x})} \alpha_{li} a^i \right\} \right\}.$$

Последовательно раскрывая выпуклые оболочки, получаем следующее множество:

$$\Big\{ q : q = \sum_{i \in I(\hat{x})} \mu_i \Big(- \sum_{l \in L_i^1(\hat{x})} \lambda_{il} \alpha_i + \sum_{l \in L_i^2(\hat{x})} \lambda_{il} \alpha_i \Big) a^i,$$
$$\mu_i \geq 0, \ \ \lambda_{il} \geq 0, \ \ l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x}), \tag{5.6}$$
$$\sum_{i \in I(\hat{x})} \mu_i = 1, \ \ \sum_{l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x})} \lambda_{il} = 1, \ \ i \in I(\hat{x}) \Big\},$$

где множество $I(\hat{x}) = \left\{ i : \alpha_i \left| (a^i)^T \hat{x} - b_i \right| = f^{II}(\hat{x}) \right\}.$

Множество (5.6) определяет субдифференциал функции $f^{II}(\hat{x})$.

6. Постановки минимаксных задач

Используя описание субдифференциалов функций $f^I(\hat{x})$ и $f^{II}(\hat{x})$, запишем минимаксные задачи для определения результирующей секущей плоскости в методе симплексных погружений. Для функции $f^I(\hat{x})$ получаем:

$$\min_{\lambda} \max_{1 \leq k \leq n+1} w_k^I(\lambda), \tag{6.1}$$

где

$$w_k^I(\lambda) = \left[\sum_{i=1}^{m} \left(- \sum_{l \in L_i^1(\hat{x})} \lambda_{il} \alpha_i + \sum_{l \in L_i^2(\hat{x})} \lambda_{il} \alpha_i \right) a^i \right]^T \left[v^k - \hat{x} \right],$$

$$\lambda_{il} \geq 0, \;\; l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x}), \qquad \sum_{l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x})} \lambda_{il} = 1,$$

v^k – k-ая вершина симплекса.

Для функции $f^{II}(\hat{x})$ минимаксная задача будет иметь следующий вид:

$$\min_{\lambda, \mu} \max_{1 \leq k \leq n+1} w_k^{II}(\lambda, \mu), \tag{6.2}$$

где

$$w_k^{II}(\lambda, \mu) = \left[\sum_{i \in I(\hat{x})} \mu_i \left(- \sum_{l \in L_i^1(\hat{x})} \lambda_{il} \alpha_i + \sum_{l \in L_i^2(\hat{x})} \lambda_{il} \alpha_i \right) a^i \right]^T \left[v^k - \hat{x} \right],$$

$$\mu_i \geq 0, \;\; \lambda_{il} \geq 0, \;\; l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x}),$$

$$\sum_{i \in I(\hat{x})} \mu_i = 1, \qquad \sum_{l \in L_i^1(\hat{x}) \bigcup L_i^2(\hat{x})} \lambda_{il} = 1, \;\; i \in I(\hat{x}),$$

v^k – k-ая вершина симплекса.

В этом параграфе мы описали задачи построения результирующих секущих плоскостей, используя параметрическое описание субдифференциалов каждой из функций $f^I(\hat{x})$ и $f^{II}(\hat{x})$. Общий вывод результирующей секущей плоскости для функции из класса задач (5.3) не представляет труда с учетом приведенных ранее правил вычисления субдифференциалов суммы выпуклых функций, умноженных на неотрицаетльные константы.

7. Алгоритм модифицированного метода симплексных погружений

Используя возможность описания субдифференциалов функций, входящих в класс задач (5.3) и введение минимаксных задач для поиска секущей плоскости, отсекающей наибольшее число вершин, приведем алгоритм модифицированного метода симплексных погружений.

Алгоритм 1. Перед началом l-й итерации, $l = 0, 1, 2, ...$ задаются матрица X^l размеров $(n+1) \times n$, строка i которой соответствует i-й вершине симплекса S_l, содержащего решение x^* задачи (5.3); ρ^l_{max} – длина максимального ребра симплекса S_l т. е.

$$\rho^l_{max} = \max_{1 \leq i \leq n,\, 1 \leq j \leq n} ||x^{l,i} - x^{l,j}||, \;\; i \neq j.$$

При $l = 0$ полагаем $\psi_l = \rho^l_{max} = +\infty$.

На l-й итерации необходимо выполнить следующие действия.

Шаг 1. Найти точку $x^{c,\,l}$ – центр симплекса S_l по формуле

$$x^{c,\,l} = \frac{1}{n+1}\left(x^l_1 + ... + x^l_n\right),$$

где x^l_i есть i-я строка матрицы X^l.

Шаг 2. Вычислить значение максимальной невязки ограничений

$$h_l = \max_{1 \leq k \leq N}\left\{0,\, \phi_k\left(x^{c,\,l}\right)\right\} = \phi_s\left(x^{c,\,l}\right),$$

положить $\psi_{l+1} = \min\{\psi_l, h_l\}$.

Шаг 3. Найти нормаль a_l отсекающей плоскости по следующей формуле:

$$a_l = \begin{cases} \partial\phi_0\left(x^{c,\,l}\right), & \text{если } h_l = 0, \\ \partial\phi_s\left(x^{c,\,l}\right), & \text{если } h_l > 0, \end{cases}$$

При $h_l = 0$ описать субдифференциал функции $\phi_0(x^{c,\,l})$, используя алгоритмы построения субдифференциалов функций $f^I(x)$ и $f^{II}(x)$. Сформировать минимаксную задачу по формулам (6.1), (6.2) и определить нормаль результирующей секущей плоскости. Проделать аналогичные действия и при $h_l > 0$ для функции $\phi_s\left(x^{c,\,l}\right)$.

Шаг 4. Построить усеченный симплекс

$$S^l_G = \left\{x : x \in S_l,\, a^T_l\left(x - x^{c,\,l}\right) \leq 0\right\}.$$

Шаг 5. Найти величины $\alpha_i = a^T_l\left(x^l_i - x^{c,\,l}\right)$, $i = 1, 2, ..., n+1$, определить индекс p из условия

$$\alpha_p = \min_{1 \leq i \leq n+1} \alpha_i.$$

Найти величины $\beta_i = -\alpha_i/\alpha_p$, $i = 1, 2, ..., n$.

Шаг 6. Найти параметр $t_l \in [0, 1]$, решив задачу одномерной выпуклой минимизации:

$$q_k^* = \min_{0 \le t \le 1} \prod_{i=1}^{n} (1 + \beta_i t)^{-1}.$$

Шаг 7. Найти вектор $\tau^l = \left(\tau_1^l, ..., \tau_{p-1}^l, \tau_{p+1}^l, ..., \tau_n^l\right)$, задающий коэффициенты растяжения ребер симплекса S_l, исходящих из опорной вершины x_p^l: $\tau_i^l = (1 + \beta_i t_l)^{-1}$.

Шаг 8. Перейти от симплекса S_l с матрицей X^l к симплексу S_{l+1} с матрицей X^{l+1}:

$$X_{ij}^{l+1} = \begin{cases} X_{pj}^l + \tau_i^l \left(X_{ij}^l - X_{pj}^l\right), \quad i \ne p, \quad i = 1, 2, ..., n+1, \\ X_{ij}^l, \quad i = p, \quad i = 1, 2, ..., n. \end{cases}$$

Шаг 9. Если

$$\rho_{max}^l \le \varepsilon,$$

тогда $x^{c,\,l}$ – решение задачи (5.3) с заданной точностью ε, в противном случае возвращаемся на шаг 1.

8. Численный эксперимент

Предварительное численное тестирование проводилось на задаче выпуклой безусловной минимизации вида

$$F(x) = \sum_{i=1}^{m} \left(\alpha_i \left|a_i^T x - b_i\right| + r_i\right) \to \min, \;\; x \in R^n$$

с известной заранее точкой минимума x^*. Идея теста заключается в том, что мы изначально задаем точку $x^* \in R^n$, а векторы $a_i \in R^n$, α_i, r_i, $i = 1, ..., m$ генерируются случайным образом. Полагая вектор $b = Ax^*$, получаем $F(x^*) = \sum_{i=1}^{m} r_i$. При этом для остальных точек $F(x) \ge F(x^*)$, поскольку к $F(x^*)$ добавляются неотрицательные в силу использования модуля слагаемые. В итоге получается кусочно-линейная выпуклая функция с минимумом в точке x^*. Цель тестирования состояла в проверке того, насколько эффективным является использование вспомогательной минимаксной задачи.

Вычисления проводились в системе GAMS на компьютере с четырехядерным процессором Intel Core i7/2.3GHz, 8 Gb оперативной памяти. Результаты представлены в таблице 1, в которой приняты следующие обозначения: n — число переменных, m — число слагаемых в целевой функции, k_B — число итераций базового метода, T_B — время работы (в секундах) базового метода, k_M — число итераций модифицированного метода, T_M — время работы (в секундах) модифицированного метода, k_{minmax} — количество минимаксных задач, решенных в ходе работы модифицированного метода. В таблице приведены средние результаты по сериям из пяти задач для каждой размерности. Точность решения ε принималось равной 10^{-5}.

Таблица 1

Результаты численного эксперимента

n	m	k_B	T_B	k_M	T_M	k_{minmax}
5	120	152	2.128	144	2.328	4
10	300	345	4.830	330	6.258	21
20	700	1337	21.392	1298	35.240	216
30	1200	3729	182.732	3582	243.071	866
40	2200	7307	255.745	7169	594.683	2203
50	3000	11250	393.750	11032	815.798	3942

Из проведённого вычислительного эксперимента можно сделать следующие выводы. Для того, чтобы проявился эффект от решаемых минимаксных задач необходимо, чтобы число слагаемых m в целевой функции (или другими словами число линейных кусков целевой функции) существенно превышало число переменных. Чем больше m, тем чаще появляется возможность использовать неоднозначность субградиента. Время работы и базового, и модифицированного методов вполне приемлемо для задач недифференцируемой оптимизации. То, что из-за решения вспомогательных минимаксных задач время работы модифицированного метода превышает время работы базового не должно вводить в заблуждение. Главной целью предлагаемой методики является сокращение количества итераций, что и было достигнуто. Во многих неявных задачах выпуклой недифференцируемой оптимизации (например, при декомпозиции задач линейного или выпуклого пограммирования большой размерности) время вычисления значения целевой функции значительно может превышать время решения вспомогательной минимаксной задачи. Адаптация и тестирование модифицированного метода симплексных погружений на неявных задачах выпуклого программирования — цель дальнейших исследований.

Подводя итог, следует отметить, что в данной статье приведено параметрическое описание субдифференциалов выпуклых недифференцируемых функций специального вида и проведено численное тестирование модифицированного метода симплексных погружений, характеризующее эффективность модификации.

Список литературы

1. Анциферов Е. Г. Алгоритм симплексных погружений в выпуклом программировании / Е. Г. Анциферов, В. П. Булатов // Журн. вычисл. математики и мат. физики. – 1987. – Т. 27, № 3. – С. 377–384.
2. Апекина Е. В. Модифицированный метод симплексных погружений с одновременным введением нескольких секущих плоскостей, / Е. В. Апекина, О. В. Хамисов // Изв. вузов. Математика. – 1997. – № 3. – С. 16–24.
3. Измаилов А. Ф. Численные методы оптимизации : учеб. пособие / А. Ф. Измаилов, М. В. Солодов. – М. : Физматлит, 2005. — 304 с.
4. Мину М. Математическое программирование. Теория и алгоритмы / М. Мину. – М. : Наука, 1990. – 488 с.
5. Нестеров Ю. Е. Введение в выпуклую оптимизацию / Ю. Е. Нестеров. – М. : МЦНМО, 2010. – 280 с.
6. Нурминский Е.А. Численные методы выпуклой оптимизации / Е. А. Нурминский. – М. Наука, 1991. – 168 с.
7. Поляк Б. Т. Введение в оптимизацию / Б. Т. Поляк. – М. : Наука, 1983. – 384 с.
8. Шор Н. З. Методы минимизации недифференцируемых функций и их приложения / Н. З. Шор. – Киев : Наук думка, 1979. – 200 с.
9. Numerical Optimization. Theoretical and Practical Aspects. Second edition / J. F. Bonnans, J. C. Gilbert, C. Lemarechal, C. A. Sagastizaabal. – Berlin : Springer-Verl., 2006. – 494 p.

Колосницын Антон Васильевич, Институт систем энергетики им. Л. А. Мелентьева СО РАН, 664033, Иркутск, ул. Лермонтова, 130 (e-mail: ankolos25@mail.ru)

A. V. Kolosnitcyn

Using of Modified Simplex Imbeddings Method for Solving Special Class of Convex Non-Differentiable Optimization Problems

Abstract. In this paper it is considered the modified simplex imbeddings method, which is related to the class of cutting plane methods. The main feature of this method is the convergence estimation, which depends only on the quantity of simplex vertices, that are cut off by the cutting plane. The more vertices are cut off by the cutting plane, the higher speed of method convergence. Modified method of simplex imbeddings with such criteria of cutting plane choosing is applied to solving special class of convex non-differentiable problems, which is consists of two types of functions. We come across the necessity to describe the function subdifferential that is depends on one or several parameters, that we can subject to optimization. It is described functions subdifferentials from the introduced class in parametric representation, that let us form auxiliary problems in simplex imbeddings method for searching resulting cutting planes, that cut off as much vertices of simplex as possible. It let us increase the speed of finding optimal solution. The results of numerical experiment are also given in this paper.

Keywords: modified simplex imbeddings method, subdifferential of convex functions, resulting cutting plane.

References

1. Antsiferov E.G., Bulatov V.P. An algorithm of simplex imbeddings in convex programming (in Russian). *Zh. vychisl. Mat. mat. Fiz.*, 1987, vol. 27, no 3, pp. 377-384.
2. Apekina E.V., Khamisov, O.V. A modified simplex immersions method with simultaneous introduction of several intersecting planes (in Russian). *Izv. Vysh. Uchebn Zaved., Mat.*, 1997, no 12, pp. 16-24.
3. Izmailov A.F., Solodov M.V. Numerical optimization methods (in Russian). Moscow, Fizmatlit, 2005. 304 p.
4. Minu M. Mathematical programming. Theory and algorithms (in Russian). Moscow, Nauka, 1990. 488 p.
5. Nesterov Yu.E. Convex optimization methods (in Russian). Moscow, MTSNMO, 2010. 280 p.
6. Nurminski E.A. Numerical methods of convex optimization (in Russian). Moscow, Nauka, 1991. 168 p.
7. Polyak B.T. Introduction to optimization (in Russian). Moscow, Nauka, 1983. 384 p.
8. Shor N.Z. Methods of minimization of non-differentiable functions and their applications (in Russian). Kiev, Nauk. dumka, 1979. 200 p.
9. Bonnans J.F., Gilbert J.C., Lemarechal C., Sagastizaabal C.A. Numerical Optimization. Theoretical and Practical Aspects. Second edition. Berlin : Springer-Verl., 2006. P. 494.

Kolosnitcyn Anton Vasilievich, Melentiev Energy Systems Institute SB RAS, 130, Lermontov st., Irkutsk, 664033 (e-mail: ankolos25@mail.ru)