УДК 519.8

П.И. СТЕЦЮК, Ю.С. КОЛЕСНИК, О.А.. БЕРЕЗОВСКИЙ

ОБ ОДНОМ МЕТОДЕ НАХОЖДЕНИЯ L_p-РЕШЕНИЯ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассматривается задача нахождения L_p–решений переопределенной системы линейных алгебраических уравнений при интервальных ограничениях. Для ее решения предложен метод, основанный на использовании модификации метода эллипсоидов. Показано, что этот метод находит L_p–решение системы линейных алгебраических уравнений за конечное число итераций, зависящее от числа переменных.

Постановка задачи. Пусть имеется система линейных алгебраических уравнений:

$$Ax \approx b, \tag{1}$$

$$l \le x \le u\,, \tag{2}$$

где A – $n \times m$ вещественная матрица; b – m–мерный вещественный вектор; l, u – n-мерные векторы, такие, что для всех $i = 1,\ldots n$ $u_i \ge l_i$; x – n-мерный вектор неизвестных параметров. Требуется найти вектор x^*, который удовлетворяет ограничениям (2) и "наилучшим образом" (обозначено знаком "$\approx$") выполняет соотношение (1).

Понятие "наилучшим образом", как правило, принято понимать как наилучшее решение системы (1)–(2) в так называемой L_p–норме, т.е. когда норма некоторого вектора $y = (y_1, \ldots, y_n)$ определена следующим образом: $\|y\|_p = \left(\sum_{i=1}^{n} |y_i|^p \right)^{1/p}$, где $p \ge 1$.

Случай $p = \infty$ определяется как $\|y\|_\infty = \max_{i=1,\ldots,n} |y_i|$. Случай $p = 2$ соответствует стандартной евклидовой норме.

Нахождению наилучшего L_p–решения системы (1)-(2) может быть поставлена в соответствие следующая задача выпуклого программирования: найти

$$\min_{x \in R^n} \left\{ f_p(x) = \|Ax - b\|_p \right\} \tag{3}$$

при ограничениях:

$$l \le x \le u\,, \tag{4}$$

где $p \in R^1$ – скалярный параметр, такой что $p \geq 1$, который гарантирует выпуклость функции $f_p(x)$.

Подобные задачи часто встречаются в самых разных областях прикладной математики, например: при обработке результатов наблюдений, построении и анализе различного рода моделей (физических, биологических, экономических, социальных и др.), при поиске компромиссных решений в моделях с противоречивыми данными и т.д.

Когда для переопределенной системы линейных уравнений ранг матрицы A равен n, то задача (3)-(4) имеет единственное решение (даже, когда ни одно из ограничений (4) не активно). В общем случае, сложно что-то утверждать о единственности x_p^* (все зависит как от границ на переменные $x_i, i = 1, \ldots n$, так и от свойств матрицы A). Однако при создании метода для нахождения x_p^* будем использовать только тот факт, что задача (3)-(4) всегда имеет решение, и то, что $f_p^* = f_p(x_p^*) \geq 0$. В случае неоднозначности x_p^* будем находить одно из решений, соответствующее минимальному значению функции f_p^*.

О методах решения задачи. В частных случаях (т.е. при некоторых конкретных значениях p) задача (3)-(4) может быть решена достаточно эффективно посредством стандартных численных процедур оптимизации. Так, например, когда $p = 1$, либо $p = +\infty$, задача может быть сведена к задаче линейного программирования и решена симплекс-методом, либо методом внутренних точек. Когда $p = 2$, то она может быть сведена к задаче квадратичного программирования и решена посредством численных методов, например из [1].

В общем случае, т.е. при произвольном значении параметра p ($p \geq 1$), задача (3)-(4) есть общей задачей выпуклого программирования при простейших (интервальных) ограничениях на переменные. При этом минимизируемая целевая функция может быть как гладкой, что имеет место при $p = 2$, так и негладкой, что имеет место при $p = 1$ и при $p = \infty$. Поэтому для ее решения требуется разработка численных методов решения, основанных на алгоритмах выпуклой недифференцируемой минимизации. Заметим, что разработка таких методов напрашивается даже тогда, когда речь идет только о наиболее распространенных значениях p (т.е. $p = 1, 2, \infty$). В этом случае одним и тем же методом можно решать задачу для каждого из указанных значений p, что практичнее, чем привлечение для решения этих задач существенно различных методов, какими есть методы линейного программирования и методы квадратичного программирования.

Конкретный метод решения задачи (3)-(4) будет определяться выбранным алгоритмом негладкой оптимизации. Так, например, такой метод можно построить на основе r-алгоритмов [2], как эффективного средства решения задач негладкой оптимизации. Подобная методика применялась в [3] и базировалась на использовании модификации $r(\alpha)$ -алгоритма. Разработанный в [3] метод предназначен для решения задачи безусловной минимизации выпуклой функции $f_p(x)$ и соответствует случаю, когда в задаче (3)-(4) отсутствуют интервальные ограничения вида (4). Однако учет интервальных ограничений для r -алгоритмов не представляет особых проблем. Это можно сделать как посредством использования негладкой штрафной функции максимума из нарушенных ограничений, либо посредством "четного" периодического продолжения целевой функции (заданной на отрезке) на все пространство посредством замены переменных [4].

Однако на практике часто требуется локализовать решение x_p^* с определенной (иногда достаточно «грубой») точностью. Семейство r -алгоритмов таким свойством не обладает. Это наводит на мысль построить более «простой» метод для решения задачи (3)-(4), но с возможностью оценки области локализации x_p^* на каждой итерации этого метода.

Это легко сделать, если в качестве алгоритма негладкой оптимизации использовать модификацию метода эллипсоидов [5]. Заметим, что с таким же успехом можно использовать и классический вариант метода эллипсоидов [2], так как скорость сходимости обоих методов приблизительно одинакова. Однако преимущество модификации в том, что ее итерация единообразно описывается и для одномерного случая (т.е. когда $n=1$), что позволит использовать метод и при решении одномерных задач вида (3)-(4).

При обосновании сходимости такого метода не возникает особых проблем. Тем более, что интервальные ограничения на переменные (4) играют существенную роль при использовании любого метода по типу метода эллипсоидов. Они дают возможность ограничить область поиска точки минимума функции $f_p(x)$ и выполнить начальные условия для применения метода эллипсоидов, которые состоят в том, что точку x_p^* требуется искать в шаре некоторого заданного радиуса. Такой шар легко построить, описав вокруг параллелепипеда, заданного ограничениями (4), либо шар, либо эллипсоид минимального объема. Кроме того, в методах эллипсоидов при построении отсекающих гиперплоскостей важно именно нормированное направление субградиента, что упрощает расчет при произвольном значении параметра p.

Однако, пойдя на использование модификации метода эллипсоидов, мы избавимся от проблем, связанных с обоснованием сходимости метода решения задачи (3)-(4). Но проблема медленной скорости сходимости, сопутствующая методам эллипсоидов, останется, и чем больше n, тем более медленную скорость сходимости будет обеспечивать построенный метод. Поэтому заранее отметим, что рассматриваемый далее метод будет применим для эффективного решения задач (3)-(4) при небольших значениях n (порядка десятка переменных). Величина m роли не играет и метод может быть применим при достаточно больших ее значениях (порядка сотен тысяч).

Метод нахождения x_p^* на основе модификации метода эллипсоидов. Модификация метода эллипсоидов [5] предназначена для решения следующей задачи.

Пусть на E^n задано векторное поле $g(x)$, не обязательно непрерывное, $g(x) \in E^n$, $x \in E^n$, $n \geq 1$. Требуется найти такую точку x^*, что $(g(x), x - x^*) \geq 0$ при всех $x \in E^n$. Предполагается, что $g(x) \neq 0$ при $x \neq x^*$.

Кроме того, априорная информация, которая требуется модификации метода эллипсоидов, связана с выбором начальной стартовой точки x_0 и такого радиуса шара r_0 с центром в x_0, чтобы в этом шаре находилась искомая точка x^*.

Удовлетворить эти требования для задачи (3)-(4) не представляет особых проблем. Так, первую часть этих требований можно удовлетворить, используя следующую лемму.

Лемма 1. Пусть $t^* = \max\{t_{i^*}, t_{j^*}\}$, где $t_{i^*} = \max\limits_{i=1,\ldots,n} \{x_i - u_i\}$ и $t_{j^*} = \max\limits_{j=1,\ldots,n} \{l_j - x_j\}$. Обозначим: i^* – значение i $(1 \leq i \leq n)$, на котором достигается t_{i^*}; j^* – значение, на котором достигается t_{j^*}; $\partial f_p(x)$ – субградиент функции $f_p(x)$; e_k – k-й орт в E^n, $1 \leq k \leq n$. Тогда вектор

$$g(x) = \left\{ \begin{array}{llll} \partial f_p(x), & \textit{если} & t^* \leq 0, & \\ e_{i^*}, & \textit{если} & t^* > 0 & \textit{и} \quad t^* > t_{i^*}, \\ -e_{j^*}, & \textit{если} & t^* > 0 & \textit{и} \quad t^* \leq t_{j^*} \end{array} \right\}$$

удовлетворяет свойству

$$(g_p(x), x - x^*) \geq 0 \text{ для всех } x \in E^n. \tag{5}$$

Доказательство леммы 1 приводить не будем. Отметим лишь, что она имеет следующий содержательный смысл. Если точка x находится внутри допустимой области, заданной ограничениями (4), то в качестве $g_p(x)$ выбирается субградиент функции $f_p(x)$ в этой точке, а если точка находится вне допустимой области, то выбирается субградиент к максимально нарушенному ограничению вида (4). Учитывая выпуклость функции и выпуклость ограничений (4), очевидно, что такой выбор $g_p(x)$ гарантирует выполнение свойства (5).

В соответствии с правилом вычисления $g_p(x)$ в лемме 1 построим формулу для вычисления "обобщенного" значения функции задачи (3)-(4):

$$F_p(x) = \begin{cases} +\infty, & \textit{если} \quad t^* > 0; \\ f_p(x), & \textit{если} \quad t^* \le 0. \end{cases}$$

Значение $F_p(x)$ будем использовать при построении одного из критериев останова в методе. Этот критерий обусловлен тем, что если $f_p^* = 0$, то вычисление субградиента функции $f_p(x)$ в точке x_p^* некорректно при большинстве значений p, что связано с операцией деления на величину, пропорциональную корню из f_p^* (т.е. будет иметь место деление на нуль).

Вторую часть требований, связанную с априорной информацией о локализации x_p^*, легко обеспечить, выбрав в качестве центра шара центр параллелепипеда, заданного интервальными ограничениями на переменные (4), и установив радиус шара таким, чтобы этот шар содержал параллелепипед и имел минимальный объем. Это обеспечивает следующая лемма.

Лемма 2. Пусть

$$x_0 = \left\{ \frac{u_1 + l_1}{2}, \ldots, \frac{u_n + l_n}{2} \right\} \text{ и } r_0 = \frac{1}{2}\sqrt{\sum_{i=1}^{n}(u_i - l_i)^2}\ .$$

Тогда параллелепипед $P(x) = \{x : l_i \le x_i \le u_i, \quad i = 1, \ldots, n\}$ содержится в шаре $S(x_0, r_0) = \{x : \|x - x_0\| \le r_0\}$.

Заметим, что это самый простой и очевидный выбор x_0 и r_0 для модификации метода эллипсоидов. Более сложный выбор можно осуществить, описав вокруг параллелепипеда эллипсоид минимального объема. Однако в этом случае усложняется доказательство метода, если некоторые переменные фиксированы (т.е. для некоторых i $l_i = u_i$). Дело в том, что эллипсоид минимального объема приводит к проектированию на орты, соответствующие этим равенствам, и обратная матрица, задающая этот эллипсоид, является вырожденной. Такая схема выбора x_0 и r_0 возможна, хотя мы ее рассматривать не будем. Заметим лишь, что она имеет некоторые преимущества перед леммой 2, так как приводит к уменьшению размерности решаемой задачи.

Учитывая вышеизложенное, метод для нахождения x_p^* примет следующий вид. Здесь p будем считать входным параметром метода ($p \ge 1$), а ε_f определяет абсолютную точность, с которой требуется найти значение f_p^*.

<u>Инициализация</u>. Установим стартовую точку $x_0 = (u + l)/2$ и начальный радиус $r_0 = \frac{1}{2}\sqrt{\sum_{i=1}^{n}(u_i - l_i)^2}$. Вычислим β по формуле $\beta = \left(\sqrt{1 + \frac{1}{n^2}} - \frac{1}{n}\right)$.

Введем в рассмотрение матрицу B размером $n \times n$ и положим $B_0 := I_n$, где I_n – единичная матрица размером $n \times n$. Перейдем к первой итерации со значениями x_0, r_0 и B_0.

Пусть на k-й итерации найдены значения $x_k \in E^n$, r_k, B_k. Переход к ($k+1$)-й итерации состоит в выполнении такой последовательности действий.

Шаг 1. Вычислим $F_p(x_k)$. Если $F_p(x_k) = 0$, то "Останов" $x_p^* = x_k$. Если критерий останова не сработал, то вычислим $g_p(x_k)$. Если $\left\| B_k^T g_p(x_k) \right\| r_k \le \varepsilon_f$, то "Останов" $x_p^* = x_k$. Иначе переходим к шагу 2.

Шаг 2. Положим

$$\xi_k := \frac{B_k^T g_k(x_k)}{\left\| B_k^T g_k(x_k) \right\|}.$$

Шаг 3. Вычислим очередную точку

$$x_{k+1} := x_k - h_k B_k \xi_k \text{, где } h_k = \frac{r_k}{n} \beta .$$

Шаг 4. Вычислим

$$B_{k+1} := B_k R_\beta(\xi) \text{ и } r_{k+1} := r_k \sqrt{1 + \frac{1}{n^2}} ,$$

где $R_\beta(\xi) = I_n + (\beta - 1)\xi\xi^T$ – оператор "сжатия" пространства с коэффициентом β в направлении $\xi \in E^n$ $(\|\xi\| = 1)$, I_n – единичная матрица размером $n \times n$.

Шаг 5. Переходим к ($k+1$)-й итерации со значениями x_{k+1}, r_{k+1}, B_{k+1}.

Факт сходимости вышеприведенного метода обеспечивает следующая теорема, доказательство которой не будем приводить в силу его громоздкости.

Теорема 1. Если не сработал ни один из критериев останова (см. шаг 1), то генерируемая методом последовательность точек $\{x_k\}_{k=o}^{\infty}$ удовлетворяет неравенству

$$\left\| B_k^{-1}(x_k - x_p^*) \right\| \le r_k, \; k = 0,1,2 \ldots .$$

Скорость сходимости метода можно установить посредством следующих рассуждений.

Множество точек x, удовлетворяющих $\left\| B_k^{-1}(x_k - x) \right\| \le r_k$, представляет собой локализующий точку x_p^* эллипсоид Φ_k объемом $vol(\Phi_k) = v_0 r_k^n \big/ \det(B_k^{-1})$. Здесь v_0 – объем единичного n-мерного шара. Следовательно, вопрос о скорости сходимости метода нахождения x_p^* сводится к оценке на каждом шаге этого метода скорости уменьшения объема эллипсоида Φ_k, локализующего x_p^*.

Теорема 2. Коэффициент уменьшения объема эллипсоида, локализующего x_p^*, на каждой итерации k метода есть величина постоянная и равная

$$q = \frac{vol(\Phi_{k+1})}{vol(\Phi_k)} = \left(1 + \frac{1}{n^2}\right)^{n/2} \times \beta < \exp\left\{-\frac{1}{2n} + \frac{1}{2n^2}\right\} < 1 .$$

Из теоремы 2 следует: для того, чтобы уменьшить объем эллипсоида, локализующего x_p^*, в 10 раз, требуется выполнить K итераций, где $K = -\ln 10/\ln q \approx (2\ln 10)n \approx 4.6n$. Так, например, для задачи (3)-(4) такое уменьшение объема означает следующее: чтобы на порядок улучшить отклонение найденного рекордного значения минимизируемой функции $f_p(x)$ от ее оптимального значения f_p^*, потребуется выполнить $4.6n^2$ итераций. Таким образом, когда в задаче (3)-(4) число переменных $n<10$, то для ее решения вышеприведенный метод будет достаточно эффективен. Его эффективность при этих значениях n можно оценить из таблицы, где приведено число итераций *itn* для нахождения x_p^* с точностью 10^{-10} по значению функции.

n	2	3	4	5	6	7	8	9	10
itn	179	408	730	1144	1651	2250	2940	3723	4598

Для современных персональных компьютеров при n~10 и m~3n метод требует небольших вычислительных затрат по времени (порядка секунды). Следовательно, его можно успешно применять для нахождения x_p^*, когда число переменных $n \le 10$. Значение m существенно не влияет на скорость сходимости метода. Однако от него зависит трудоемкость вычислений значения функции $f_p(x)$ и ее субградиента, которые при m~1000 будут вносить более весомый вклад в трудоемкость метода, чем алгоритмические операции (шаги 2–4).

В заключение отметим, что вышеприведенный метод свободен от параметров регулировки шага, настройка всех его параметров делается автоматически. Кроме того, за конечное число итераций он гарантирует нахождение x_p^* с требуемой точностью (на каждой итерации можно указать область локализации точки минимума).

П.І.Стецюк, Ю.С. Колесник, О.А.Березовський

ПРО ОДИН МЕТОД ЗНАХОДЖЕННЯ LP-РОЗВ'ЯЗКУ СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ

Розглядається задача знаходження L_p–розв'язку перевизначеної системи лінійних алгебраїчних рівнянь при інтервальних обмеженнях. Для її розв'язування запропоновано метод, який основано на використанні модифікації метода еліпсоїдів. Показано, що цей метод знаходить L_p–розв'язок системи лінійних алгебраїчних рівнянь за кінцеве число ітерацій, яке залежить від кількості змінних.

P.I.Stetsyuk, Yu.S.Kolesnik, O.A.Berezovski

ON ONE METHOD TO FIND LP-SOLUTION OF LINEAR EQUATIONS SYSTEM

The problem of finding L_p–solution to an overdetermined system of linear algebraic equations with interval restrictions is considered. To solve this problem it's proposed the method which based on using the modification of ellipsoid method. It's shown that the method finds L_p–solution of linear algebraic equations system for finite number of iterations which depends on a number of variables.

1. *Пшеничный Б.Н., Данилин Ю.М.* Численные методы в экстремальных задачах. – М.: Наука, 1975. – 319 с.
2. *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.
3. *Стецюк П.И., Колесник Ю.С.* К вопросу выбора метода аппроксимации результатов измерений // Интеллектуальные информационно-аналитические системы и комплексы. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2000. – С. 62-67.
4. *Шор Н.З., Стеценко С.И.* Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
5. *Стецюк П.И.* Приближенный метод эллипсоидов // Кибернетика и системный анализ. – 2003. – N.3. – С. 141-146.

Получено 01.09.2003