

УДК 519.8

**МЕТОД ЭЛЛИпсоИДОВ ДЛЯ НАХОЖДЕНИЯ
 L_p -РЕШЕНИЯ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ**

П. И. Стецюк, д.ф.-м.н., с.н.с.

Институт кибернетики им. В.М. Глушкова НАН Украины
stetsyukp@gmail.com

Г. Д. Била, к.ф.-м.н

Институт кибернетики им. В.М. Глушкова НАН Украины
bila.galyna@gmail.com

В. А. Стовба, аспирант

Институт кибернетики им. В.М. Глушкова НАН Украины
stovbaviktor@yandex.ua

Предложена модификация метода эллипсоидов для нахождения L_p -решения переопределенной системы линейных уравнений при двусторонних ограничениях на компоненты решения.

Stetsyuk P.I., Bila H.D., Stovba V.O. Ellipsoid method for finding L_p -solution of linear equations system. Modification of ellipsoid method is proposed to find L_p -solutions of overdetermined system of linear equations with two-sided constraints on the components of the solution.

Ключевые слова: ДВУСТОРОННИЕ ОГРАНИЧЕНИЯ, ЗАДАЧА ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ, МЕТОД ЭЛЛИпсоИДОВ.

Keywords: TWO-SIDED CONSTRAINTS, CONVEX PROGRAMMING PROBLEM, ELLIPSOID METHOD.

Введение. Пусть имеется переопределенная система линейных алгебраических уравнений:

$$Ax \approx b \quad \text{при условиях} \quad l \leq x \leq u, \quad (1)$$

где A – $m \times n$ -матрица, такая что $m > n$, b – m -мерный вектор;

$l \in R^n$, $u \in R^n$ – n -мерные векторы, такие, что для всех $i = 1, \dots, n$, $u_i > l_i$; $x \in R^n$ – n -мерный вектор неизвестных параметров. Требуется найти такой вектор x^* , который в (1) "наилучшим образом" выполняет соотношение $Ax \approx b$.

Под термином "наилучшим образом" будем понимать "наилучшее" решение системы (1) в так называемой L_p -норме, т. е. когда норма вектора невязок $y = Ax - b = (y_1, \dots, y_m)^T$ определена следующим образом: $\|y\|_p = \left(\sum_{i=1}^m |y_i|^p\right)^{1/p}$, где $p \geq 1$. Случай $p = \infty$ определяется как $\|y\|_\infty = \max_{i=1, \dots, m} |y_i|$. Случай $p = 2$ соответствует стандартной евклидовой норме $\|y\|$ для вектора невязок.

Нахождению "наилучшего" решения системы (1) поставим в соответствие следующую задачу выпуклого программирования: найти

$$f_p^* = f_p(x_p^*) = \min_{x \in R^n} \left\{ f_p(x) = \|Ax - b\|_p \right\}, \quad (2)$$

при ограничениях

$$l \leq x \leq u, \quad (3)$$

где $p \in R$ – скалярный параметр, такой что $p \geq 1$, который гарантирует выпуклость функции $f_p(x)$. Здесь x_p^* – решение задачи (2)–(3), для удобства будем считать, что оно единственное.

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 других), при поиске компромиссных решений в моделях с противоречивыми данными, и т.д. При нахождении параметров линейной регрессии решение задачи (2)–(3) будет соответствовать методу наименьших квадратов ($p = 2$), методу наименьших модулей ($p = 1$), минимаксному (чебышевскому) методу ($p = \infty$).

1. Задача (2)-(3) и метод эллипсоидов [1, 2]. Чтобы для решения задачи (2)–(3) применить метод эллипсоидов требуется: во первых – определить градиентное поле $g(x)$ (способ построения в точке $x \in R^n$ гиперплоскости, которая локализует точку x_p^* в одном из полупространств пространства R^n); во-вторых – выбрать начальный радиус области локализации оптимального решения x_p^* . Удовлетворить эти требования для задачи (2)–(3) не представляет особых проблем. Так, первую часть этих требований можно удовлетворить, используя следующую лемму.

Лемма 1 [1]. Пусть $t^* = \max\{t_{i^*}, t_{j^*}\}$, где $t_{i^*} = \max_{i=1, \dots, n} \{x_i - u_i\}$ и $t_{j^*} = \max_{j=1, \dots, n} \{l_j - x_j\}$. Обозначим: i^* – значение i ($1 \leq i \leq n$), на котором достигается t_{i^*} ; j^* – значение j ($1 \leq j \leq n$), 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

$$g_p(x) = \begin{cases} \partial f_p(x), & \text{если } t^* \leq 0, \\ e_{i^*}, & \text{если } t^* > 0 \text{ и } t^* > t_{i^*}, \\ -e_{j^*}, & \text{если } t^* > 0 \text{ и } t^* \leq t_{j^*}, \end{cases}$$

удовлетворяет свойству

$$(g_p(x), x - x_p^*) \geq 0 \quad \text{для всех } x \in R^n. \quad (4)$$

Лемма 1 имеет следующий содержательный смысл. Если точка x находится внутри допустимой области, заданной ограничениями (3), то в качестве $g_p(x)$ выбирается субградиент функции $f_p(x)$ в этой точке; если же точка находится вне допустимой области, то выбирается субградиент к максимально нарушенному ограничению вида (3). Выпуклость функции $f_p(x)$ и ограничений (3) для векторного поля $g_p(x)$ гарантирует выполнение свойства (4).

Субградиент $\partial f_p(x)$ в точке x вычисляется по формуле:

$$\partial f_p(x) = \|Ax - b\|_p^{1-p} \sum_{j=1}^m \left(\operatorname{sgn}(a_j x - b_j) \cdot |a_j x - b_j|^{p-1} \right).$$

Вторую часть требований, связанную с априорной информацией о локализации точки x_p^* , легко обеспечить, выбрав за центр шара центр параллелепипеда, заданного двусторонними ограничениями на переменные (3), и установив радиус шара таким, чтобы этот шар содержал параллелепипед и имел минимальный объем. Это обеспечивает следующая лемма.

Лемма 2 [1]. Если $x_0 = \frac{1}{2}(u+l)$ и $r_0 = \frac{1}{2}\|u-l\|$, то тогда параллелепипед $P(x) = \{x: l_i \leq x_i \leq u_i, i=1, \dots, n\}$ содержится в n -мерном шаре $S(x_0, r_0) = \{x: \|x - x_0\| \leq r_0\}$.

Следует отметить, что в лемме 2 используется самый простой и очевидный выбор x_0 и r_0 для начальной локализации точки x_p^* . Более сложный выбор можно осуществить, описав вокруг параллелепипеда эллипсоид минимального объема. Однако, в этом случае усложняется обоснование сходимости метода эллипсоидов, если некоторые переменные фиксированы (т.е. $l_i = u_i$ для некоторых i). Дело в том, что использование эллипсоида минимального объема приводит к проектированию на орты, соответствующие фиксированным переменным, и обратная матрица, задающая этот эллипсоид, является вырожденной. Такая схема выбора x_0 и r_0 возможна, но мы ее рассматривать не будем. Заметим лишь, что она имеет преимущество в сравнении с леммой 2, так как приводит к уменьшению размерности решаемой задачи.

2. Метод эллипсоидов для нахождения x_p^* . В соответствии с правилом вычисления $g_p(x)$ в лемме 1 построим формулу для вычисления "обобщенного" значения функции в задаче (3)–(4):

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

Значение $F_p(x)$ будем использовать при построении критерия останова в методе эллипсоидов для нахождения x_p^* . Параметр p будем считать входным параметром метода ($p \geq 1$), а ε_f – точностью, с которой требуется найти значение $f_p^* = f_p(x_p^*)$.

Учитывая вышеизложенное, метод эллипсоидов для нахождения x_p^* примет следующий вид.

Инициализация. Положим стартовую точку $x_0 = (u+l)/2$ и начальный радиус $r_0 = \|u-l\|/2$. Вычислим $\beta = \sqrt{\frac{n-1}{n+1}}$.

Введем в рассмотрение $n \times n$ -матрицу B и положим $B_0 := I_n$, где I_n – единичная $n \times n$ -матрица. 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: $k^* = k$ и $x_p^* = x_k$ ". Иначе вычислим $g_p(x_k)$. Если $F_p(x_k) < +\infty$ и $\|B_k^T g_p(x_k)\| r_k \leq \varepsilon_f$, то "Останов: $k^* = k$ и $x_p^* = x_k$ ". Иначе переходим к шагу 2.

Шаг 2. Положим $\xi_k := \frac{B_k^T g_p(x_k)}{\|B_k^T g_p(x_k)\|}$.

Шаг 3. Вычислим очередную точку

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

Шаг 4. Вычислим

$$B_{k+1} := B_k + (\beta - 1)(B_k \xi_k) \xi_k^T \quad \text{и} \quad r_{k+1} := r_k \frac{n}{\sqrt{n^2 - 1}}.$$

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

Сходимость метода обеспечивает следующая теорема.

Теорема. Последовательность точек $\{x_k\}_{k=0}^{k^*}$ удовлетворяет неравенству

$$\|B_k^{-1}(x_k - x_p^*)\| \leq r_k, \quad k = 0, 1, 2, \dots, k^*$$

На каждой итерации k коэффициент q , который определяет величину уменьшения объема эллипсоида, локализирующего x_p^* , есть величина постоянна и равна

$$q = \frac{\text{vol}(E_{k+1})}{\text{vol}(E_k)} = \sqrt{\frac{n-1}{n+1}} \left(\frac{n}{\sqrt{n^2-1}} \right)^n < \exp\left\{-\frac{1}{2n}\right\} < 1.$$

Из теоремы следует, что для уменьшения в 10 раз объема эллипсоида, локализирующего x_p^* , требуется K итераций, где

$$K = -\frac{\ln 10}{\ln q} \approx (2 \ln 10)n \approx 4.6n. \quad \text{Для задачи (2)–(3) это означает}$$

следующее: чтобы на порядок улучшить отклонение найденного рекордного значения функции $f_p(x)$ от ее оптимального значения f_p^* , потребуется сделать $4.6n^2$ итераций. Если в задаче переменных не более десяти, то для нахождения f_p^* с относительной точностью 10^{-10} максимальное количество итераций легко определить из таблицы

n	itn	n	itn	n	itn
2	177	5	1144	8	2940
3	407	6	1651	9	3723
4	730	7	2249	10	4598

где приведены необходимые количества итераций для $n = 2 \div 10$.

Закключение. Метод эллипсоидов можно успешно применять для нахождения x_p^* , если количество переменных небольшое.

На скорость его сходимости величина m не влияет, от нее зависит трудоемкость вычисления значения функции $f_p(x)$ и ее субградиента $\partial f_p(x)$. При $m \sim 1000$ это будет вносить в

трудоемкость метода эллипсоидов более весомый вклад, чем алгоритмические операции (шаги 2–4).

Метод эллипсоидов может быть использован для линейной регрессии с ограничениями на параметры θ [3]. Нахождению θ соответствует задача квадратичного программирования:

$$\sum_{i=1}^m (y_i - x_i^T \theta)^2 \rightarrow \min \quad \text{при ограничениях} \quad b^{low} \leq A\theta \leq b^{up}, \quad (5)$$

которая решается за конечное число итераций средствами многих программных продуктов. Однако это уже будет не так,

если в задаче (5) целевую функцию заменить на $\sum_{i=1}^m \left\| y_i - x_i^T \theta \right\|_p$.

Зато в таком случае применим метод эллипсоидов. Более того, на его основе легко построить метод для нахождения оптимальных значений θ^* , для которого настройка всех его параметров производится автоматически. Кроме того, за конечное число итераций метод гарантирует нахождение θ^* с требуемой точностью, а на каждой итерации можно указать область локализации точки минимума в задаче (5).

Работа выполнена при поддержке НАН Украины (проект № 0116U004558, первый и третий авторы) и Volkswagen Foundation (грант No 90 306, первый и второй авторы).

Литература

1. Стецюк П.И., Колесник Ю.С., Березовский О.А. Об одном методе нахождения L_p -решения системы линейных уравнений // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2003. – С. 83–90.
2. Шор Н.З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. – 1977. – №1. – С. 94-95.
3. Knopov P.S., Korkhin A.S. Regression Analysis Under a Priori Parameter Restrictions. – Springer, 2012. – 234 p.