ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2001 МАТЕМАТИКА № 12 (475)

УДК 517.977

Я.И. ЗАБОТИН, А.А. АНДРИАНОВА

АЛГОРИТМЫ В МЕТОДЕ ЦЕНТРОВ С АППРОКСИМАЦИЕЙ МНОЖЕСТВА ДОПУСТИМЫХ РЕШЕНИЙ

При применении итерационных методов решения задач математического программирования приходится, как правило, останавливать вычислительный процесс на основании эвристических соображений о достижении заданной точности решения, т. к. теоретические критерии оптимальности выполняются лишь в искомой точке оптимума. Поэтому для вычислителя при применении им того или иного метода оптимизации важно иметь не только критерии оптимальности, но и легко проверяемые условия, при выполнении которых гарантируется достижение за конечное число итераций наперед заданной точности приближенного решения. Разработке алгоритмов, вырабатывающих такие условия, и посвящена данная статья. Эти условия получены для методов внутренних и внешних центров ([1], с. 83–90) за счет использования в алгоритмах окрестности и соответственно подмножества множества допустимых решений. Предполагается, что вспомогательные задачи безусловной минимизации решаются точно.

Задаче прерывания итерационного процесса в методе центров были посвящены, например, работы [2]–[6]. В [2] была предложена параметризация функции расстояния и было показано, что в принципе за счет выбора значения параметра можно достичь заданной точности решения за один итерационный процесс минимизации вспомогательной функции максимума. Но так как указать такие значения параметров не удалось, в [3]–[5] проводилась процедура адаптации параметра, которая, однако, имела итерационный характер. Наконец, в отличие от [6] в данной статье вырабатываются критерии, при которых заданная точность достигается за конечное число итераций с помощью одностороннего, а не двустороннего процесса приближений.

1. Свойства функции максимума

В данном разделе приводятся вспомогательные утверждения для аддитивно параметризованной функции максимума в той форме, в которой они понадобятся в дальнейшем.

Пусть функции $\varphi(x)$, $\psi(x)$ определены и непрерывны в n-мерном евклидовом пространстве R_n, $G = \{x : x \in R_n,\ \psi(x) \le 0\}$, $G \ne \emptyset$, $x^* \in \operatorname{Arg\,min}\{\varphi(x),\ x \in G\}$, $\varphi^* = \varphi(x^*)$.

Как и в [6], через M обозначается множество функций, определенных в R_n, для которых каждый локальный минимум является абсолютным. Считается, что $\psi \in M$.

Положим

$$\Phi(x,t,\gamma) = \max\{\varphi(x) - t, \psi(x) - \gamma\}, \quad t,\gamma \text{ — константы},$$
$$Z = \operatorname{Arg\,min}\{\Phi(x,t,\gamma),\ x \in R_n\}.$$

Лемма 1. *Пусть параметры t, γ зафиксированы так, что существует точка $z \in Z$, $z \notin G$. Тогда*

1. $\varphi(x) - t > \psi(x) - \gamma\ \ \forall x \in G$;
2. $\varphi(z) - t \ge \psi(z) - \gamma$;
3. $\varphi(z) \le \varphi^*$;

4. *если при этом для какой-либо точки $y \in G$ выполняется неравенство $\varphi(y) \le t$, то*

$$0 \le \varphi(y) - \varphi^* < \gamma. \tag{1}$$

Доказательство. Допустим, что существует такая точка $\overline{x} \in G$, что $\varphi(\overline{x}) - t \le \psi(\overline{x}) - \gamma$. Тогда $\Phi(\overline{x}, t, \gamma) = \psi(\overline{x}) - \gamma$, а т. к. $\psi(\overline{x}) \le 0$, то $\Phi(\overline{x}, t, \gamma) \le -\gamma$. С другой стороны, из определения функции $\Phi(x, t, \gamma)$ следует неравенство $\Phi(z, t, \gamma) \ge \psi(z) - \gamma$, а т. к. по условию леммы $\psi(z) > 0$, то $\Phi(z, t, \gamma) > -\gamma$. Кроме того, $\Phi(z, t, \gamma) \le \Phi(x, t, \gamma)$ для любого $x \in R_n$, в том числе для точки $\overline{x}$. Из указанных неравенств вытекает противоречивое неравенство $-\gamma < \Phi(z, t, \gamma) \le \Phi(\overline{x}, t, \gamma) \le -\gamma$, что доказывает утверждение 1.

Предположим теперь, что справедливо неравенство $\varphi(z) - t < \psi(z) - \gamma$. Известно ([6], лемма 2), что тогда точка z является точкой локального, а т. к. $\psi \in M$, то и абсолютного минимума функции $\psi(x)$. Но $\psi(z) > 0$, тогда как $G \ne \emptyset$ и, следовательно, существует точка x, для которой $\psi(x) \le 0$. Это противоречие доказывает утверждение 2.

Утверждение 3 вытекает из утверждения 1 и условий леммы. Действительно, для всех $x \in G$ выполняются соотношения $\varphi(z) - t \le \Phi(z, t, \gamma) \le \Phi(x, t, \gamma) = \varphi(x) - t$. В частности, эти соотношения выполняются для тех $x \in G$, для которых $\varphi(x) = \varphi^*$.

Используя, наконец, утверждения 2 и 3 при $t \ge \varphi(y)$, а также неравенство $\psi(z) > 0$, получаем $0 \ge \varphi^* - \varphi(y) \ge \varphi(z) - t \ge \psi(z) - \gamma > -\gamma$. Отсюда следует (1). □

Замечание. Утверждение 3 леммы 1, конечно, нельзя считать новым, т. к. такая ситуация возникает в методе внешних центров, хотя и для несколько иной функции максимума.

Следующая лемма, как и утверждение 4 леммы 1, носит характер оценки близости значений функции $\varphi(x)$ к ее минимуму на множестве G.

Лемма 2. *Пусть параметры t, γ зафиксированы так, что существует такая точка $z \in Z$, что $z \in G$ и $\varphi(z) > \varphi^*$. Тогда выполняется неравенство*

$$\varphi(x^*) - t < \psi(x^*) - \gamma, \tag{2}$$

а если при этом $t \le \varphi^*$, то

$$0 < \varphi(z) - \varphi^* \le -\gamma. \tag{3}$$

Доказательство. Допустим, что (2) неверно. Тогда в силу определения функции $\Phi(x, t, \gamma)$ и условий леммы выполняются соотношения $\varphi(x^*) - t = \Phi(x^*, t, \gamma) \ge \varphi(z) - t$, что противоречит неравенству $\varphi(z) > \varphi^*$, и неравенство (2) доказано.

Если к тому же $t \le \varphi^*$, то $0 < \varphi(z) - \varphi^* \le \varphi(z) - t \le \Phi(z, t, \gamma) \le \Phi(x^*, t, \gamma) = \psi(x^*) - \gamma$. А так как $x^* \in G$, то $\psi(x^*) \le 0$ и, следовательно, выполняется (3). □

2. Постановка задачи и алгоритмы

Пусть функции $f(x)$, $f_i(x)$ для $i \in H = \{1, 2, \ldots, m\}$ определены и непрерывны в n-мерном евклидовом пространстве R_n, $g(x) = \max\{f_i(x), i \in H\}$. Для $\varepsilon \ge 0$ определим множества $D(\varepsilon) = \{x : x \in R_n, g(x) \le \varepsilon\}$, $D'(\varepsilon) = \{x : x \in R_n, g(x) < \varepsilon\}$.

Требуется найти

$$\min\{f(x),\ x \in D\}, \tag{4}$$

где $D = D(0)$.

Считается, что множество D удовлетворяет условию регулярности по Слейтеру, т. е. $D'(0) \ne \emptyset$.

Таким образом, множество $D(\varepsilon)$ при любом $\varepsilon > 0$ является окрестностью множества допустимых решений задачи (4), а D, в свою очередь, является окрестностью для множества $D(-\varepsilon)$ при некотором $\varepsilon > 0$.

Пусть $f^* = \min\{f(x),\ x \in D\} > -\infty$, $X^* = \{x : x \in D,\ f(x) = f^*\}$.

Будем говорить, что точка $z \in D$ при $\varepsilon > 0$ является ε-решением задачи (4), если выполняется неравенство $f(z) \leq f^* + \varepsilon$.

Первый из алгоритмов основывается на следующей идее. Функция расстояния в форме функции максимума строится не для множества D, а для его замкнутой окрестности $D(\varepsilon)$. Итерационный процесс проводится по методу внутренних центров для задачи минимизации функции $f(x)$ на множестве $D(\varepsilon)$. Попадание точки минимума вспомогательной функции максимума во множество $D(\varepsilon) \setminus D$ является сигналом того, что предыдущая итерационная точка является ε-решением задачи. При этом вырабатывается и критерий оптимальности целевой функции без использования градиента, что позволяет останавливать итерационный процесс при решении задачи (4) для недифференцируемой целевой функции.

Изложим далее алгоритм, построенный по описанной схеме, и его обоснование.

Алгоритм 1. Выбирается точка начального приближения $x_0 \in D$, число $\varepsilon > 0$. Если найдена точка $x_k \in D (k \geq 0)$, то переход к x_{k+1} осуществляется по следующей схеме.

1. Строится функция $\Phi(x, f(x_k), \varepsilon) = \max\{f(x) - f(x_k), g(x) - \varepsilon\}$.
2. Находится $z_{k+1} \in \text{Arg}\min\{\Phi(x, f(x_k), \varepsilon),\ x \in R_n\}$.
3. Если $z_{k+1} \notin D$, то вычислительный процесс останавливается, точка x_k принимается за ε-решение задачи (4). Иначе следует переход к п. 4.
4. Если выполняется $f(z_{k+1}) - f(x_k) > g(z_{k+1}) - \varepsilon$, то $z_{k+1} \in \text{Arg}\min\{f(x),\ x \in D\}$. Задача (4) решена. В противном случае переходим к п. 5.
5. За x_{k+1} принимается любая точка из D, для которой $f(x_{k+1}) \leq f(z_{k+1})$, и осуществляется переход к п. 1 при k, замененном на $k+1$.

Пункты 3, 4 и тот факт, что за конечное число шагов, проделанных по алгоритму 1, будет получено решение или ε-решение задачи (4), обосновывает

Теорема 1. *Пусть функции $f(x)$, $g(x)$ определены и непрерывны в R_n, $f, g \in M$ и последовательность $\{x_k\}$ вырабатывается по алгоритму 1. Тогда найдется такой номер m, что для z_{m+1} выполняются условия п. 3 или п. 4 алгоритма и соответственно x_m является ε-решением или z_{m+1} является точным решением задачи (4).*

Доказательство. Докажем сначала, что найдется номер m такой, что при $k = m$ будут выполнены условия п. 3 или п. 4 алгоритма 1. Предположим, что это не так, т. е. что выполняются условия

$$z_{k+1} \in D \text{ и } f(z_{k+1}) - f(x_k) \leq g(z_{k+1}) - \varepsilon \quad \forall k \geq 0. \tag{5}$$

Так как в силу первого из условий (5) выполняется неравенство $g(z_{k+1}) \leq 0$ для всех $k \geq 0$, то $f(z_{k+1}) - f(x_k) \leq -\varepsilon$ в силу второго из этих условий. Отсюда в соответствии с п. 5 алгоритма 1

$$f(x_{k+1}) \leq f(z_{k+1}) \leq f(x_k) - \varepsilon \quad \forall k \geq 0. \tag{6}$$

С одной стороны, из (6) следует, что числовая последовательность $\{f(x_k)\}$ неограниченно убывает. С другой стороны, $x_k \in D$ согласно п. 5 алгоритма 1 и, следовательно, $f(x_k) \geq f^*$ для всех $k \geq 0$. Полученное противоречие показывает, что указанный выше номер m существует.

Если для этого номера m выполнены условия п. 3 алгоритма 1, то $z_{m+1} \notin D$, и согласно утверждению 4 леммы 1 получим $0 \leq f(x_m) - f^* < \varepsilon$, т. е. x_m по определению является ε-решением задачи (4).

Пусть теперь для номера m выполнены условия п. 4 алгоритма 1. Это значит, что процесс построения последовательности $\{x_k\}$ при $k = m$ не был прекращен по условиям п. 3 алгоритма 1. Это означает, что $z_{m+1} \in D$. Кроме того, неравенство в п. 4 является строгим. Следовательно, существует такая окрестность ω_1 точки z_{m+1}, что это неравенство справедливо для всех $x \in \omega_1$. Но тогда $\Phi(x, f(x_m), \varepsilon) = f(x) - f(x_m)$ для всех $x \in \omega_1$. А так как z_{m+1} — точка абсолютного минимума функции $\Phi(x, f(x_m), \varepsilon)$, то $\Phi(z_{m+1}, f(x_m), \varepsilon) \leq \Phi(x, f(x_m), \varepsilon)$, или, что то же самое, $f(z_{m+1}) \leq f(x) \ \ \forall x \in \omega_1$, т. е. z_{m+1} — точка локального минимума функции $f(x)$. А так как $f \in M$, то z_{m+1} является точкой абсолютного минимума функции $f(x)$ и, следовательно, $z_{m+1} \in X^*$. $\square$

Замечание. Пункт 5 алгоритма 1 позволяет выбрать $x_{k+1} = z_{k+1}$ или комбинировать алгоритм 1 с любым другим методом минимизации, позволяющим уменьшить значение целевой функции более простым способом.

Следующий алгоритм строится на идее, аналогичной той, которая использовалась для построения алгоритма 1, и представляет собой метод внешних центров с использованием функции расстояния, построенной для множества $D(-\varepsilon), \varepsilon > 0$.

Алгоритм 2. Выбирается точка начального приближения $x_0 \notin D$, для которой $f(x_0) \leq f^*$, и число $\varepsilon > 0$. Если найдена точка $x_k (k \geq 0)$, то переход к точке x_{k+1} осуществляется по следующей схеме.

1. Строится функция $\Phi(x, f(x_k), -\varepsilon) = \max\{f(x) - f(x_k), g(x) + \varepsilon\}$.
2. Находится $x_{k+1} \in \mathrm{Arg}\ \min\{\Phi(x, f(x_k), -\varepsilon),\ x \in R_n\}$.
3. Если $x_{k+1} \in D$, то x_{k+1} принимается за ε-решение задачи (4). Иначе следует переход к п. 1 при k, замененном на $k+1$.

Пункт 3 и тот факт, что за конечное число шагов, проделанных по алгоритму 2, будет получено решение или ε-решение задачи (4), обосновывает

Теорема 2. *Пусть функции $f(x)$, $g(x)$ определены и непрерывны в R_n, $f, g \in M$ и последовательность $\{x_k\}$ вырабатывается по алгоритму 2. Тогда найдется такой номер m, что $x_k \notin D$ при $k \leq m$, а для x_{k+1} выполняются условия п. 3 и x_{m+1} является ε-решением задачи* (4).

Доказательство. Докажем, что найдется номер m такой, что $x_m \notin D$, $x_{m+1} \in D$. Предположим, что $x_k \notin D$ для любого номера $k \geq 0$. Тогда в силу утверждений 2 и 3 леммы 1 для любых $k \geq 0$ справедливы соотношения

$$f(x_{k+1}) - f(x_k) \geq g(x_{k+1}) + \varepsilon, \tag{7}$$

$$f(x_{k+1}) \leq f^*. \tag{8}$$

В силу (8) последовательность $\{f(x_k)\}$ ограничена сверху. Так как $x_{k+1} \notin D$, то $g(x_{k+1}) > 0$ и в силу (7) для любого номера $k \geq 0$ выполняется неравенство $f(x_{k+1}) - f(x_k) > \varepsilon$, что противоречит ограниченности сверху последовательности $\{f(x_k)\}$. Следовательно, существует номер m такой, что $x_{m+1} \in D$.

Так как согласно п. 3 алгоритма 2 вычислительный процесс прекращается, как только $x_{k+1} \in D$ для некоторого k, то из включения $x_{m+1} \in D$ следует, в частности, что $x_m \notin D$. Тогда $f(x_m) \leq f^*$ согласно п. 3 леммы 1. Очевидно, $f(x_{m+1}) \geq f^*$. Если это неравенство выполняется как равенство, точка x_{m+1} является точным решением задачи (4). В противном случае $f(x_{m+1}) > f^*$. Учитывая это и используя лемму 2, получим неравенство $0 < f(x_{m+1}) - f^* \leq \varepsilon$. $\square$

Замечание 1. Требования, налагаемые в алгоритме 2 на выбор точки x_0, можно удовлетворить следующим образом. Построив функцию $F(x) = \max\{f(x) + t, g(x)\}$, выбрать $t \geq 0$ настолько большим, чтобы $f(x) + t > g(x)$ для всех $x \in D$, например, добиться, чтобы $f(x) + t > 0$ для всех $x \in D$. Тогда, как легко убедиться, в качестве x_0 можно выбрать $\mathrm{argmin}\{F(x),\ x \in R_n\}$. В частности, если $f(x)$ — квадратичная положительно определенная форма, то $f(x) > g(x)$ $\forall x \in D$.

Замечание 2. С помощью алгоритма 2 в качестве приближенного с гарантированной точностью решения получается допустимое решение, тогда как в стандартном методе внешних центров любая итерационная точка, принимаемая за приближенное решение, не принадлежит множеству D и останавливать итерационный процесс приходится по эвристическим критериям без гарантии достижения заданной точности приближенного решения.

Литература

1. Гроссман К.,Каплан А.А. *Нелинейное программирование на основе безусловной минимизации.* – Новосибирск: Наука, 1981. – 183 с.
2. Заботин Я.И. *Минимаксный метод решения задачи математического программирования* // Изв.вузов. Математика. – 1975. – № 6. – С. 36–43.
3. Заботин Я.И., Князев Е.А. *Алгоритмы с адаптацией в параметризованном методе центров* // Исследов. по прикл. матем. – Казань, 1987. – № 14. – С. 9–15.
4. Заботин И.Я., Князев Е.А. *Вариант параметризованного метода центров* // Изв. вузов. Математика. – 1995. – № 12. – С. 26–32.
5. Князев Е.А. *Алгоритмы с адаптацией в методе центров*: Дис. … канд. физ.-матем. наук. – Казань, 1987. – 134 с.
6. Заботин Я.И., Даньшин И.Н. *Алгоритмы с комбинированием, параметризацией и двусторонним приближением в методе центров* // Изв. вузов. Математика. – 1998. – № 12. – С. 40–48.

Казанский государственный университет

Поступила 12.04.2001