ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

2005 МАТЕМАТИКА № 12 (523)

КРАТКИЕ СООБЩЕНИЯ

УДК 519.853

Ю.А. ЧЕРНЯЕВ

СХОДИМОСТЬ МЕТОДА ПРОЕКЦИИ ГРАДИЕНТА ДЛЯ ОДНОГО КЛАССА НЕВЫПУКЛЫХ ЗАДАЧ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Работа посвящена обобщению метода проекции градиента, применяемого для оптимизации гладких функций, на случай невыпуклых допустимых множеств. В статье [1] введено понятие предвыпуклого множества как множества, дополнение которого до его выпуклой оболочки выпукло, и показано, что такое множество всегда представляется в виде теоретико-множественной разности двух выпуклых множеств. В [2]–[4] предложено обобщение метода проекции градиента на случай предвыпуклых допустимых множеств с непустой внутренностью. В данном сообщении полученные результаты обобщаются на случай теоретико-множественной разности произвольного выпуклого множества и объединения нескольких выпуклых множеств. Получены необходимые условия экстремума и сформулированы предложения о сходимости предлагаемого метода.

1. Постановка задачи и алгоритм

Рассматривается задача нахождения точки, удовлетворяющей необходимому условию локального минимума функции $\varphi(x)$ на множестве X в n-мерном евклидовом пространстве R^n, где $\varphi(x) \in C^1(X)$, а X является теоретико-множественной разностью некоторых множеств F и $\bigcup\limits_{i=1}^{l} \operatorname{int} G_i$, причем F и G_i выпуклы и замкнуты, множества внутренних точек X и G_i, $i=\overline{1,l}$, непусты. Пусть каждое из множеств G_i, $i=\overline{1,l}$, в любой своей граничной точке x имеет единственную опорную гиперплоскость, нормаль $N^i(x)$ которой считается внешней, т.е. для всех $y \in G_i$ имеет место $\langle N^i(x), y-x\rangle \le 0$. Будем полагать, что для каждого i орт $n^i(x)$ нормали $N^i(x)$ является непрерывной вектор-функцией на границе G_i, т.е. при любом $\varepsilon > 0$ существует такое $\delta(\varepsilon) > 0$, что из выполнения условия $\|x-y\| < \delta(\varepsilon)$ для произвольных x и y, лежащих на границе G_i, следует справедливость неравенства $\|n^i(x) - n^i(y)\| < \varepsilon$. Необходимое условие локального минимума $\varphi(x)$ на множестве X указанного вида приведено ниже.

Введем обозначения: $s^i(x)$ — проекция точки x на множество G_i, $n^i(x)$ — орт нормали опорной гиперплоскости к G_i в точке $s^i(x)$, $\Gamma^i(x) = \{e \in R^n : \langle n^i(x), e - s^i(x)\rangle \ge 0\}$, $P(x) = F \cap \Gamma^1(x) \cap \Gamma^2(x) \cap \dots \cap \Gamma^l(x)$. Проекции $s^i(x)$ определяются однозначно, т.к. G_i, $i=\overline{1,l}$, являются выпуклыми множествами евклидова пространства R^n. Поскольку каждое из G_i, $i=\overline{1,l}$, в любой своей граничной точке x имеет только одну опорную гиперплоскость, то векторы $n^i(x)$, а значит, и полупространства $\Gamma^i(x)$, $i=1,2,\dots,l$, определяются единственным образом для любого $x \in X$. Если при некотором i точки x и $s^i(x)$ не совпадают, то векторы $n^i(x)$ и $x - s^i(x)$ имеют одно направление, а значит, $\langle n^i(x), x - s^i(x)\rangle > 0$, т.е. $x \in \operatorname{int}\Gamma^i(x)$. Если же x и $s^i(x)$ совпадают, то x лежит на границе $\Gamma^i(x)$. Поскольку $x \in X \subset F$ и при каждом $i=1,2,\dots,l$ имеет место $x \in \Gamma^i(x)$, то всегда $x \in P(x)$.

Предлагается следующий алгоритм построения последовательных приближений.

Шаг 0. Положим $k=0$.

Шаг 1. Пусть $x_k \in X$ есть k-е приближение.
Шаг 2. Определяются точки $s^i(x_k)$, $i = \overline{1,l}$.
Шаг 3. Строятся полупространства $\Gamma^i(x_k)$, $i = \overline{1,l}$.
Шаг 4. Строится множество $P(x_k)$.
Шаг 5. Задается величина $\alpha_k > 0$.
Шаг 6. Строится y_k — проекция точки $x_k - \alpha_k\varphi'(x_k)$ на $P(x_k)$.
Шаг 7. Если $x_k = y_k$, то вычисления заканчиваются, иначе осуществляется переход к шагу 8.
Шаг 8. Задается величина $\beta_k \in (0;1]$.
Шаг 9. Пусть $x_{k+1} = x_k + \beta_k(y_k - x_k)$.
Шаг 10. Полагается $k := k+1$ и осуществляется переход к шагу 1.

Множества $P(x_k)$, $k = 0,1,2,\ldots$, выпуклы и замкнуты, т. к. являются пересечением выпуклого замкнутого множества F с замкнутыми полупространствами $\Gamma^i(x_k)$, $i = \overline{1,l}$, и непусты, т. к. по построению $x_k \in P(x_k)$, $k = 0,1,2,\ldots$, поэтому задача проектирования точки $x_k - \alpha_k\varphi'(x_k)$ на множество $P(x_k)$ при $k = 0,1,2,\ldots$ имеет единственное решение. Числа α_k, β_k, $k = 0,1,2,\ldots$, могут выбираться по-разному.

Ниже будет показано, что если точки x_k и y_k совпадают при некотором k, то найденная точка x_k является стационарной, т. е. удовлетворяет необходимому условию локального минимума $\varphi(x)$ на X. Условие $x_k = y_k$ может не выполняться ни при каком k, тогда процесс вычислений становится бесконечным, но при определенных способах выбора чисел α_k, β_k, $k = 0,1,2,\ldots$, и некоторых дополнительных условиях доказывается, что в этом случае любая предельная точка последовательности $\{x_k\}$ будет стационарной. Следует отметить, что найденная с помощью алгоритма стационарная точка, вообще говоря, не будет точкой глобального минимума $\varphi(x)$ на X. Для того чтобы выяснить, доставляет ли точка глобальный минимум, требуются дополнительные исследования функции на множестве.

2. Сходимость алгоритма

В этом разделе формулируется лемма о необходимом условии локального минимума функции $\varphi(x) \in C^1(X)$ на множестве X изучаемого вида, предлагаются различные способы выбора чисел α_k, β_k и приводятся предложения о сходимости алгоритма.

Лемма 1. *Если x_* — точка локального минимума функции $\varphi(x)$ на множестве X указанного вида, то выполняется условие*

$$\forall x \in P(x_*) : \langle \varphi'(x_*), x - x_* \rangle \geq 0. \tag{1}$$

Лемма справедлива, поскольку по построению $x_* \in P(x_*) \subset X$, а условие (1) есть необходимое условие локального минимума $\varphi(x)$ на выпуклом множестве $P(x_*)$.

Из известного свойства проекций на выпуклые множества, а также положительности чисел α_k следует, что совпадение точек x_k и y_k влечет за собой выполнение неравенств

$$\langle -\alpha_k\varphi'(x_k), x - x_k \rangle \leq 0 \quad \text{и} \quad \langle \varphi'(x_k), x - x_k \rangle \geq 0 \quad \text{для всех } x \in P(x_k).$$

Таким образом, совпадение точек x_k и y_k при некотором k означает, что точка x_k удовлетворяет необходимому условию локального минимума из леммы 1.

Приведем три способа выбора параметров α_k, β_k, $k = 0,1,2,\ldots$, и сформулируем предложения о сходимости алгоритма в смысле необходимого условия экстремума из леммы 1.

Способ 1 характеризуется выбором параметров из условий

1) $\beta_k = 1$ при $k = 0,1,2,\ldots$;
2) α_k, $k = 0,1,2,\ldots$, ограничены снизу положительным числом;
3) α_k, $k = 0,1,2,\ldots$, выбираются так, что выполняются неравенства

$$\varphi(x_k) - \varphi(x_{k+1}) \geq \varepsilon\|x_k - x_{k+1}\|^2,$$

где ε — некоторое положительное число.

Вопрос о возможности выбора α_k, $k = 0, 1, 2, \dots$, из условий 2 и 3 рассмотрен в [2].

Предложение 1. *Если* $\varphi(x) \in C^{1,1}(X)$, *множество*

$$M(x_0) = \{x \in X \mid \varphi(x) \le \varphi(x_0)\} \tag{2}$$

ограничено, последовательность $\{x_k\}$ *построена по изложенному алгоритму и числа* α_k, β_k *выбираются согласно способу* 1, *то любая предельная точка* x_* *последовательности* $\{x_k\}$, *для которой* $\operatorname{int} P(x_*) \ne \emptyset$, *удовлетворяет условию* (1).

Способ 2 выбора параметров состоит в следующем:

1) $\alpha_k = \alpha$, $k = 0, 1, \dots$, где α — некоторая положительная константа;
2) $\beta_k \in \operatorname*{Argmin}_{0 \le \beta \le 1} \varphi[x_k + \beta(y_k - x_k)]$, $k = 0, 1, 2, \dots$ (если минимум достигается не при единственном β, то в качестве β_k берется произвольное из соответствующих β).

При использовании способа 2 на каждой итерации нужно решать вспомогательную задачу одномерной минимизации. Справедлива

Лемма 2. *Если точки* x_k *и* y_k *не совпадают, то* $\beta_k > 0$.

На множестве X вводится вспомогательная функция $\xi(x) = \varphi(x) - \min_{\beta \in [0;1]} \varphi[x + \beta(y(x) - x)]$, где $y(x) = \operatorname{Pr}_{P(x)}[x - \alpha\varphi'(x)]$, и оказываются верными следующие две леммы.

Лемма 3. *Точка* $y(x)$ *непрерывно зависит от* x *при всех* $x \in X$, *для которых* $\operatorname{int} P(x) \ne \emptyset$.

Лемма 4. *Функция* $\xi(x)$ *непрерывна в любой точке* $x \in X$, *для которой* $\operatorname{int} P(x) \ne \emptyset$.

С учетом лемм 2, 3 и 4 можно доказать

Предложение 2. *Если множество* (2) *ограничено, последовательность* $\{x_k\}$ *построена по изложенному алгоритму и числа* α_k, β_k *выбираются согласно способу* 2, *то любая предельная точка* x_* *последовательности* $\{x_k\}$, *для которой* $\operatorname{int} P(x_*) \ne \emptyset$, *удовлетворяет условию* (1).

Способ 3 выбора параметров состоит в следующем:

1) $\alpha_k = \alpha$, $k = 0, 1, \dots$, где α — некоторая положительная константа;
2) величина β_k на каждой итерации полагается равной 2^{-j}, где j — первый индекс $i = 0, 1, \dots$, для которого выполняется неравенство

$$\varphi(x_k + 2^{-i}(y_k - x_k)) \le \varphi(x_k) + 0{,}5 \cdot 2^{-i} \cdot \langle \varphi'(x_k), y_k - x_k \rangle.$$

Вопрос о возможности выполнения последнего неравенства изучен в [4].

Предложение 3. *Если* $\varphi(x) \in C^{1,1}(X)$, *множество* (2) *ограничено, последовательность* $\{x_k\}$ *построена по изложенному алгоритму и числа* α_k, β_k *выбираются согласно способу* 3, *то любая предельная точка* x_* *последовательности* $\{x_k\}$, *для которой* $\operatorname{int} P(x_*) \ne \emptyset$, *удовлетворяет условию* (1).

Доказательства предложений 1, 2 и 3 для случая $l > 1$ обобщают доказательства соответствующих предложений для случая $l = 1$, приведенные соответственно в [2], [3] и [4].

Множество X может задаваться в виде

$$X = \{x \in R^n \mid f_i(x) \le 0,\ i = 1, 2, \dots, m + l\} \tag{3}$$

при выполнении условий

1) $f_i(x) \in C^1(R^n)$, $i = \overline{1, m+l}$;
2) $f_i(x)$, $i = \overline{1, m}$, выпуклые, $f_{m+i}(x)$, $i = \overline{1, l}$, вогнутые в R^n;
3) $f_i(x) < 0$, $i = \overline{1, m+l}$, для некоторой точки $x \in R^n$;
4) $f_{m+i}(x) > 0$, $i = \overline{1, l}$, для некоторой точки $x \in R^n$.

Из этих условий и теорем выпуклого анализа следует, что X — множество вида, указанного в постановке задачи. Из условия 2 следует, что множества $F_i = \{x \in R^n \mid f_i(x) \le 0\}$, $i = 1, 2, \dots, m$, выпуклы, множество $F = \bigcap\limits_{i=1}^{m} F_i$ выпукло как пересечение выпуклых множеств, множества $G_i = \{x \in R^n \mid f_{m+i}(x) \geq 0\}$, $i = \overline{1,l}$, тоже выпуклы. Из условия 3 следует, что $\operatorname{int} X \neq \emptyset$, а из условия 4 следует, что $\operatorname{int} G_i \neq \emptyset$, $i = \overline{1,l}$. Если при некотором i точка $\dot{x}$ лежит на границе G_i, то $n^i(\dot{x}) = -f'_{m+i}(\dot{x})\|f'_{m+i}(\dot{x})\|^{-1}$, а вектор-функция $n^i(x)$ непрерывна на границе G_i, поскольку $f_{m+i}(x) \in C^1(R^n)$. Справедливо

Предложение 4. *Если множество X задано в виде* (3), *справедливы условия* 1)–4) *и выполнены требования одного из предложений* 1, 2 *и* 3, *то любая предельная точка последовательности $\{x_k\}$ стационарна в смысле Лагранжа.*

3. Вычислительные аспекты

В работе предложен алгоритм, обобщающий метод проекции градиента, который применяется для минимизации функций на выпуклых множествах, на один класс невыпуклых допустимых множеств. Идея обобщения состоит в том, что задача проектирования на каждой итерации решается не для самого невыпуклого множества, а для некоторых его выпуклых подмножеств. Эти подмножества являются теоретико-множественным пересечением выпуклого множества F и нескольких полупространств $\Gamma^i(x_k)$. Задача проектирования на множество такого типа решается, как правило, непросто. Кроме того, построению каждого из полупространств предшествует решение задачи проектирования точки на одно из выпуклых множеств G_i, которая тоже не всегда проста. Поэтому для практического применения предлагаемого в этой работе метода, вообще говоря, требуется разработка вычислительных алгоритмов для различных невыпуклых множеств изучаемого вида.

При оптимизации на множествах относительно простой структуры разработанный метод может, однако, работать эффективно. Если, например, допустимое множество задано в виде теоретико-множественной разности выпуклого многогранного множества, представленного линейными неравенствами, и объединения нескольких шаров, то задачи проектирования на каждой итерации сводятся к задачам проектирования на шаровые множества, решаемым элементарно, и к задаче квадратичного программирования, которая всегда может быть решена известными методами. Если же решение задач проектирования на итерациях требует привлечения итерационных процедур, то эффективность метода снижается.

Литература

1. Заботин В.И., Полонский Ю.А. *Предвыпуклые множества, отображения и их приложения к экстремальным задачам* // Кибернетика. – 1981. – № 1. – С. 71–74.
2. Заботин В.И., Черняев Ю.А. *Обобщение метода проекции градиента на экстремальные задачи с предвыпуклыми ограничениями* // Журн. вычисл. матем. и матем. физ. – 2001. – Т. 41. – № 3. – С. 367–373.
3. Черняев Ю.А. *Об одном численном алгоритме решения экстремальных задач с предвыпуклыми ограничениями* // Журн. вычисл. матем. и матем. физ. – 2003. – Т. 43. – № 2. – С. 169–175.
4. Черняев Ю.А. *Два алгоритма решения задачи математического программирования с предвыпуклыми ограничениями* // Журн. вычисл. матем. и матем. физ. – 2004. – Т. 44. – № 7. – С. 1229–1233.

Казанский государственный
технический университет

Поступили
полный текст 22.09.2004
краткое сообщение 27.04.2005