ИСПОЛЬЗОВАНИЕ ЗНАЧЕНИЙ ПРОИЗВОДНЫХ В ПАРАЛЛЕЛЬНЫХ АЛГОРИТМАХ УСЛОВНОЙ ГЛОБАЛЬНОЙ ОПТИМИЗАЦИИ

К.А. Баркалов

Нижегородский госуниверситет им. Н.И. Лобачевского

Предложен алгоритм решения многоэкстремальных задач с невыпуклыми ограничениями, использующий значения первых производных функций задачи. Предполагается, что производные удовлетворяют условию Липшица. Для редукции условной задачи к безусловной используется индексная схема, основанная на раздельном учете каждого ограничения задачи; идеи метода штрафных функций не используются. Проведены численные эксперименты, подтверждающие сходимость метода и его эффективность по сравнению с алгоритмами без учета значений производных.

Введение

Оригинальный подход к минимизации многоэкстремальных функций при невыпуклых ограничениях – индексный метод – основан на раздельном учете каждого ограничения задачи и не связан с использованием штрафных функций [1, 2]. В соответствии с правилами индексного метода каждая итерация включает последовательную проверку выполнимости ограничений задачи в этой точке. При этом обнаружение первого нарушенного ограничения прерывает испытание и инициирует переход к точке следующей итерации.

Алгоритмы из работ [1, 2] основаны на предположении о липшицевости всех функционалов задачи, что типично и для многих других подходов (см., например, [3, 4]). Но если для функций, входящих в постановку задачи оптимизации, известны значения производных, их использование может существенно ускорить сходимость. Так, в [5-7] предложены эффективные алгоритмы, использующие значения производной целевой функции в задачах глобальной оптимизации без ограничений. В то же время остается открытым вопрос использования значений производных в алгоритмах условной глобальной оптимизации.

1. Постановка задачи

Рассматривается одномерная задача глобальной оптимизации

$$\varphi(x^*) = \min\{\varphi(x) : x \in [a,b],\ g_j(x) \le 0,\ 1 \le j \le m\}, \tag{1}$$

когда целевая функция $\varphi(x)$ (в дальнейшем обозначаемая также g_{m+1}) и левые части ограничений $g_j(x)$, $1 \le j \le m$, определены и вычислимы лишь в соответствующих подобластях $Q_j \subset [a,b]$, где

$$Q_1 = [a,b],\ Q_{j+1} = \{x \in Q_j : g_j(x) \le 0\},\ 1 \le j \le m. \tag{2}$$

С учетом условий (2), исходная задача (1) может быть представлена в виде

$$\varphi(x^*) = \min\{g_{m+1}(x) : x \in Q_{m+1}\}.$$

Вводится классификация точек x из области поиска $[a,b]$ с помощью *индекса* $\nu = \nu(x)$. Указанный индекс ν определяется условиями

$$g_j(x) \le 0,\ 1 \le j \le \nu - 1,\ g_\nu(x) > 0,$$

где последнее неравенство несущественно, если $\nu = m+1$, и удовлетворяет неравенствам

$$1 \le \nu = \nu(x) \le m+1.$$

Данная классификация порождает «индексную» функцию

$$f(x) = g_\nu(x),\ \nu = \nu(x),$$

определенную и вычислимую всюду на $[a,b]$. Ее значение в точке x есть либо значение левой части ограничения, нарушенного в этой точке (случай, когда $\nu \le m$), либо значение минимизируемой функции (случай, когда $\nu = m+1$).

Основная идея индексного подхода состоит в редукции условной задачи (1) к безусловной задаче

$$\psi(x^*) = \min\{\psi(x):\ x \in [a,b]\},$$

где

$$\psi(x) = \begin{cases} g_\nu(x), & \nu < m, \\ g_m - g_m^*, & \nu = m. \end{cases}$$

Как результат, функция $\psi(x)$ будет дифференцируемой на каждом множестве Q_ν, $1 \le \nu \le m$. Однако эта новая функция будет иметь разрывы первого рода на граничных точках множеств Q_ν из (2), поэтому применять для ее минимизации известные алгоритмы [6-8] непосредственно нельзя. Требуются дополнить алгоритм новыми правилами учета разрывов функции $\psi(x)$.

2. Оценка минимума миноранты для «индексной» функции

Основой рассматриваемого в работе алгоритма является предположение о липшицевости первых производных функций задачи, т.е. выполнение условия

$$\left|g_j'(x_1) - g_j'(x_2)\right| \le L_j\left|x_1 - x_2\right|,\ 1 \le j \le m+1,\ [x_1, x_2] \subset [a,b].$$

В работе [7] приведен вывод нижней оценки значения минимизируемой функции на отрезке $[x_1, x_2]$ в задачах безусловной оптимизации (т.е. при m=0, j=1)

$$g_j(x) \ge g_j(x_1) + g_j'(x_1)(\hat{x} - x_1) - 0.5L_j(\hat{x} - x_1)^2,$$

где

$$\hat{x} = \frac{-(g_j(x_2) - g_j(x_1)) + (g_j'(x_2)x_2 - g_j'(x_1)x_1) + 0.5L_j(x_2^2 - x_1^2)}{L_j(x_2 - x_1) + (g_j'(x_2) - g_j'(x_1))}.$$

Указанная оценка может быть использована и для задач с ограничениями, если $j = \nu(x_1) = \nu(x_2)$, т.е. когда на границах отрезка нарушено одно и то же ограничение, или же точки x_1, x_2 являются допустимыми. В случае если $\nu(x_1) \ne \nu(x_2)$, то в качестве нижней оценки можно использовать либо неравенство

$$g_j(x) \ge g_j(x_1) + g_j'(x_1)(x_2 - x_1) - 0.5L_j(x_2 - x_1)^2,\ j = \nu(x_1) > \nu(x_2)$$

(т.е. когда у функции вычислено значение и производная на левой границе отрезка), либо неравенство

$$g_j(x) \ge g_j(x_2) - g_j'(x_2)(x_2 - x_1) - 0.5L_j(x_2 - x_1)^2,\ j = \nu(x_2) > \nu(x_1)$$

(т.е. когда у функции вычислены значение и производная на правой границе отрезка).

Полученные нижние оценки могут быть использованы в алгоритме в качестве характеристик интервалов поиска: очередное испытание нужно проводить в интервале, которому соответствует наименьшее значение миноранты, т.е. в интервале с минимальной характеристикой.

3. Последовательный и параллельный алгоритмы

Первое испытание осуществляется в произвольной внутренней точке $x^1 \in (a,b)$. Выбор точки x^{k+1}, $k \geq 1$, любого последующего испытания определяется следующими правилами.

Правило 1. Перенумеровать точки $x^1, \dots, x^k$ предшествующих испытаний нижними индексами в порядке увеличения значений координаты, т.е.

$$a = x_0 < x_1 < \dots < x_i < \dots < x_k < x_{k+1} = b, \tag{3}$$

и сопоставить им значения $z_i = g_\nu(x_i)$ и $z'_i = g'_\nu(x_i)$, $\nu = \nu(x_i)$, $1 \leq i \leq k$, вычисленные в этих точках; точки $x_0 = a$ и $x_{k+1} = b$ введены дополнительно (значения z_0, z_{k+1}, z'_0, z'_{k+1} не определены) для удобства последующих обозначений.

Правило 2. Провести классификацию номеров $i, 1 \leq i \leq k$, точек из ряда (3) по числу ограничений задачи, выполняющихся в этих точках, путем построения множеств

$$I_\nu = \{ i:\ 1 \leq i \leq k,\ \nu = \nu(x_i) \},\ 1 \leq \nu \leq m+1,$$

содержащих номера всех точек $x_i, 1 \leq i \leq k$, имеющих индексы, равные одному и тому же значению ν. Граничные точки $x_0 = a$ и $x_{k+1} = b$ интерпретируются как имеющие нулевые индексы, и им сопоставляется дополнительное множество $I_0 = \{0, k+1\}$.

Определить максимальное значение индекса

$$M = \max\{ \nu = \nu(x_i), 1 \leq i \leq k \}.$$

Правило 3. Вычислить текущие нижние границы для неизвестных констант Липшица L_ν первой производной функций $g_\nu, 1 \leq \nu \leq m+1$

$$\mu_\nu = \max\left\{ \max\left\{ \begin{array}{c} \left| z'_i - z'_j \right| \big/ (x_i - x_j), \\ 2\left[-(z_i - z_j) + z'_j(x_i - x_j) \right] \big/ (x_i - x_j)^2, \\ 2\left[(z_i - z_j) - z'_i(x_i - x_j) \right] \big/ (x_i - x_j)^2, \end{array} \right\} \quad i, j \in I_\nu,\ i > j \right\} \tag{4}$$

Если множество I_ν содержит менее двух элементов или если μ_ν из (4) оказывается равным нулю, то принять $\mu_\nu = 1$. Из (4) следует, что оценки μ_ν являются неубывающими, начиная с момента, когда (4) порождает первое положительное значение μ_ν.

Правило 4. Для всех непустых множеств I_ν, $1 \leq \nu \leq m+1$, вычислить оценки

$$z_\nu^* = \begin{cases} 0, & \nu < M, \\ \min\{ g_\nu(x_i) : i \in I_\nu \}, & \nu = M. \end{cases}$$

Правило 5. Для каждого интервала (x_{i-1}, x_i), $1 \leq i \leq k+1$, вычислить *характеристику* $R(i)$, при этом если $\nu = \nu(x_{i-1}) = \nu(x_i)$, то

$$R(i) = (z_{i-1} - z_\nu^*) + z'_{i-1}(\hat{x}_i - x_{i-1}) - 0.5 r_\nu \mu_\nu (\hat{x}_i - x_{i-1})^2,$$

где

$$\hat{x}_i = \frac{-(z_i - z_{i-1}) + (z'_i x_i - z'_{i-1} x_{i-1}) + 0.5 r_\nu \mu_\nu (x_i^2 - x_{i-1}^2)}{r_\nu \mu_\nu (x_i - x_{i-1}) + (z'_i - z'_{i-1})},$$

если $\nu = \nu(x_i) > \nu(x_{i-1})$, то

$$R(i) = (z_i - z_\nu^*) - z'_i(x_i - x_{i-1}) - 0.5 r_\nu \mu_\nu (x_i - x_{i-1})^2,$$

если $\nu = \nu(x_{i-1}) > \nu(x_i)$, то

$$R(i) = (z_{i-1} - z_\nu^*) + z'_{i-1}(x_i - x_{i-1}) - 0.5 r_\nu \mu_\nu (x_i - x_{i-1})^2.$$

Величины $r_\nu > 1$, $1 \le \nu \le m+1$, являются параметрами алгоритма. Подходящий выбор значений r_ν позволяет использовать произведение $r_\nu \mu_\nu$ как оценку констант Липшица L_ν, $1 \le \nu \le m+1$, для первых производных.

Правило 6. Определить интервал (x_{t-1}, x_t), которому соответствует минимальная характеристика

$$R(t) = \min\{R(i):\ 1 \le i \le k+1\}. \tag{5}$$

Правило 7. Провести очередное испытание в серединной точке интервала (x_{t-1}, x_t), если индексы его концевых точек не совпадают, т.е.

$$x^{k+1} = \frac{x_t + x_{t-1}}{2},\ \ \nu(x_{t-1}) \ne \nu(x_t).$$

В противном случае провести испытание в точке

$$x^{k+1} = \hat{x}_t = \frac{-(z_t - z_{t-1}) + (z'_t x_t - z'_{t-1} x_{t-1}) + 0.5 r_\nu \mu_\nu (x_t^2 - x_{t-1}^2)}{r_\nu \mu_\nu (x_t - x_{t-1}) + (z'_t - z'_{t-1})},$$

$$\nu = \nu(x_{t-1}) = \nu(x_t).$$

Описанные правила можно дополнить условием остановки, прекращающим испытания, если

$$x_t - x_{t-1} \le \varepsilon,$$

где t из (5) и $\varepsilon > 0$ есть заданная точность.

Предложенный алгоритм относится к классу характеристических алгоритмов, поэтому его распараллеливание возможно в рамках синхронной и асинхронной схем, представленных в [8]. При этом алгоритм распараллеливается на единой области оптимизации, что позволяет учитывать результаты испытаний, полученные всеми процессорами.

Проведены численные эксперименты, подтверждающие сходимость метода и его эффективность по сравнению с алгоритмами без учета значений производных. Примененная процедура оценки основана на численном решении сравниваемыми методами всех задач из некоторой случайно генерируемой выборки большого объема и использует аппарат *операционных характеристик, описанный, например, в* [9].

В заключение отметим, что в работе предложен способ использования производных в задачах условной глобальной оптимизации, решение которых осуществляется с помощью индексной схемы. Предложенный алгоритм существенно ускоряет процесс решения задач условной оптимизации. Данное утверждение установлено путем численного решения набора тестовых задач из различных классов.

Работа выполнена при поддержке совета по грантам Президента Российской Федерации (грант № НШ-1960.2012.9), РФФИ (грант № 11-01-00682-а), а также ФЦП «Научные и научно-педагогические кадры инновационной России» (соглашение № 14.B37.21.0878).

Литература

1. Стронгин Р.Г., Баркалов К.А. О сходимости индексного алгоритма в задачах условной оптимизации с ε-резервированными решениями // Математические вопросы кибернетики. М.: Наука, 1999. С. 273–288.
2. Баркалов К.А., Стронгин Р.Г. Метод глобальной оптимизации с адаптивным порядком проверки ограничений // Ж. вычисл. матем. и матем. физ. 2002. Т.42, №9. С. 1338–1350.
3. Евтушенко Ю.Г. Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.

4. Pinter J. Global optimization in action (Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications). Kluwer Acad. Publ., 1996.
5. Гергель В.П. Об одном способе учета значений производных при минимизации многоэкстремальных функций // Ж. вычисл. матем. и матем. физ. 1996. Т.36, №6. С. 51–67.
6. Sergeyev Ya.D. A global optimization algorithm using derivatives and local tuning. ISI-CNR, Report 1, 1994.
7. Gergel V.P., Sergeyev Ya.D. Sequential and parallel algorithms for global minimizing functions with lipschitzian derivatives// Computers & Mathematics with Applications. 1999. V. 37 № 4-5. P. 163–179.
8. Grishagin V.A., Sergeyev Y.D., Strongin R.G. Parallel characteristical algorithms for solving problems of global optimization // Journal of Global Optimization. 1997. V.10. № 2. P. 185–206.
9. Сергеев Я.Д., Квасов Д.Е. Диагональные методы глобальной оптимизации – М.: ФИЗМАТЛИТ, 2008.