

УДК 519.10

АНАЛИЗ ТОЧНОСТИ ГРАДИЕНТНОГО АЛГОРИТМА
НА СТРУКТУРАХ ЖОРДАНА-ДЕДЕКИНДА

А.Б.РАМАЗАНОВ

Бакинский Государственный Университет
ram-bsu@mail.ru

В работе рассматривается задача максимизации порядково-выпуклой функции на структурах Жордана-Дедекинда. Для этой задачи получена апостериорная гарантированная оценка точности градиентного алгоритма, которая уточняет и дополняет ранее известные оценки. Кроме того, приведены достаточные условия оптимальности градиентного решения.

Ключевые слова: выпуклость, точность, градиент, алгоритм, крутизна.

Исследование дискретных экстремальных задач на структурах Жордана-Дедекинда является актуальным как для общего развития теории дискретных экстремальных задач, так и для получения более точных оценок точности градиентных (локальных) алгоритмов [1, 2]. Хорошо известно, что градиентный алгоритм не всегда гарантирует получение оптимального решения в соответствующей задаче дискретной оптимизации. Поэтому, естественно, возникает проблема оценки точности градиентного алгоритма. Одна из естественных формализаций градиентных алгоритмов дискретной оптимизации и методик оценки их погрешности предложена в модели порядковой выпуклости [1] (др. ссылки см., напр. [3]). В работе [2] показано, что приведенные в [1] оценки можно улучшить, используя информацию о крутизне целевой функции. В данной работе рассматривается задача максимизации порядково-выпуклой функции на конечных структурах Жордана-Дедекинда. В терминах крутизны целевой функции получена апостериорная гарантированная оценка точности жадного (градиентного) алгоритма, которая уточняет и дополняет ранее известные оценки [1-3]. Кроме того, приведены достаточные условия оптимальности градиентного решения в терминах крутизны целевой функции. Отметим, что данная работа является развитием и обобщением работы [2].

Определения и обозначения

Пусть $H = (H, \prec)$ - множество, на котором задано отношение частичного порядка \prec .

Функция $f : H \rightarrow R$ называется ρ -порядково-выпуклой [1], если

$$2f(y) - f(x) - f(z) \geq \rho, \forall x \triangleright y \triangleright z,$$

где ρ -фиксированное неотрицательное действительное число; $x \triangleright y$ означает, что y непосредственно следует за x в H . Через $\mathfrak{F}_\rho = \mathfrak{F}_\rho(H)$ обозначим класс всех ρ -порядково-выпуклых функций, заданных на множестве H .

Напомним, что множество элементов x^0, x^1, \dots, x^k из H , обладающих свойством $x = x^0 \prec x^1 \prec \dots \prec x^k = y$ называется цепью между x и y . Число k называется длиной цепи. Цепь $x = x^0 \triangleright x^1 \triangleright \dots \triangleright x^k = y$ называется максимальной цепью между x и y . Будем предполагать, что частично упорядоченное множество H удовлетворяет условию Жордана-Дедекинда [4]: все максимальные цепи между произвольными сравнимыми элементами x и y имеют одинаковую длину, которая обозначается через $h(x, y)$. Кроме того, предполагаем, что множество H обладает единственным минимальным элементом (нулем), который будем обозначать через θ . Таким образом, $\theta \prec x \quad \forall x \in H, x \neq \theta$. Будем также пользоваться обозначением $h(x) = h(\theta, x)$.

Введем обозначение

$$h = h(H) = \max\{h(x) \mid x \in H\}, r = \min\{h(x) \mid x \in H^{\max}\},$$

где H^{\max} - множество максимальных элементов в H .

Функция $f \in \mathfrak{F}_\rho(H)$ называется неубывающей, если из $x \prec y$ следует $f(x) \leq f(y)$.

Правым градиентом функции $f(x)$, как обычно [1, 2], называем функции

$$\Delta^+ f(x) = \max\{f(y) - f(x) \mid x \triangleright y, x, y \in H\}.$$

Крутизной функции $f(x)$ будем называть величину $c = c(f)$ [2, 3]:

$$c = c(f) = \begin{cases} \min\{(\Delta^+ f(x) - \Delta^+ f(y)) / \Delta^+ f(x) \mid (x, y) \in I\}, & \text{если } I \neq \emptyset, \\ 0, & \text{если } I = \emptyset, \end{cases}$$

где $I = \{(x, y) \mid \Delta^+ f(x) > \Delta^+ f(y) \geq 0, x \triangleright y, x, y \in H\}$.

Очевидно, что если $f(x) \in \mathfrak{F}_\rho(H)$, $\rho > 0$, то $I \neq \emptyset$.

Гарантированная оценка

Рассматривается следующая задача A : найти

$$\max\{f(x) \mid x \in H\},$$

где $f(x)$ - неубывающая функция из класса $\mathfrak{F}_\rho = \mathfrak{F}_\rho(H)$.

Градиентным (локальным) решением (максимумом) x^g задачи A (функции $f(x)$) на множестве H , назовем точку, построенную с помощью следующей итерационной процедуры [1, 2]:

$$x^{t+1} = \arg \max \{f(y) - f(x^t) \mid x^t \triangleright y, y \in H\}, t = 0, 1, \dots, x^0 = \theta,$$

заканчивающей работы на шаге τ , когда либо $\Delta^+ f(x^\tau) \leq 0$, либо x^τ - максимальный элемент множества H .

Через x^* - обозначим оптимальное решение задачи A .

Под гарантированной (относительной) оценкой погрешности градиентного алгоритма решения задачи A , как обычно, понимают такое число $\varepsilon \geq 0$, что

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} \leq \varepsilon.$$

Теорема. Для задачи A справедлива следующая гарантированная оценка точности градиентного алгоритма

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} \leq \left(1 - \frac{1}{1 + (1-c)(h-1)} \right)^r = B_1,$$

где $c = c(f)$.

Доказательство. Из п.2 теоремы 4 [1], имеем

$$f(y) - f(x) \leq h(x, y) \Delta^+ f(x) - \rho(h(x, y) - 1), x \prec y, x, y \in H.$$

Отсюда при $y = x^*$, $x = x^t$, $t = 1, \dots, r$, получаем

$$\begin{aligned} f(x^*) - f(x^t) &\leq h(x^t, x^*) \Delta^+ f(x^t) - \rho(h(x^t, x^*) - 1) \leq \\ &(h-1) \Delta^+ f(x^t), t = 1, \dots, r \end{aligned} \quad (1)$$

Из определения величины $c = c(f)$ при $y = x^t$, $x = x^{t-1}$, $t = 1, 2, \dots, r$, имеем

$$\Delta^+ f(x^t) \leq (1-c) \Delta^+ f(x^{t-1}), t = 1, 2, \dots, r. \quad (2)$$

Учитывая (2) в (1), выводим

$$f(x^*) - f(x^t) \leq (1-c)(h-1) \Delta^+ f(x^{t-1}), t = 1, 2, \dots, r \quad (3)$$

Далее, повторяя схему доказательства теоремы 4 [1], из неравенства (3), выведем оценку из теоремы.

Теорема доказана.

Следствия и пример

В случае, когда $\rho = 0$ в [2] для задачи A приведена следующая оценка

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} \leq \left(1 - \frac{1}{h} \right)^r = B_2. \quad (4)$$

Следующее следствие показывает, что оценка из теоремы лучше, чем оценка (4).

Следствие 1. Справедливы соотношения

$$B_1 \begin{cases} = B_2, \text{ если } (c = 0) \text{ или } (h = 1), \\ < B_2, \text{ если } (c > 0) \text{ и } (h > 1). \end{cases}$$

Пункт 1 очевиден. Пункт 2 следствия 1 из следующей цепочки неравенств

$$(1-c)(h-1) < h-1, \frac{1}{1+(1-c)(h-1)} > \frac{1}{h}, 1 - \frac{1}{1+(1-c)(h-1)} < 1 - \frac{1}{h},$$

т.е.

$$B_1 = \left(1 - \frac{1}{1+(1-c)(h-1)} \right)^r < \left(1 - \frac{1}{h} \right)^r = B_2.$$

Следствие 2. Если в задаче A , $c = c(f) = 1$, то $f(x^*) = f(x^g)$.

Следствие 3. Полученная оценка в теореме достижима.

Действительно, пусть

$$H = \{a, b, c, d, e, q\}, a \triangleright b, a \triangleright c \triangleright d, a \triangleright e \triangleright q, f(a) = 0,$$

$$f(b) = 3, f(c) = f(e) = 2, f(q) = 5, f(d) = 4.$$

Очевидно, что

$$\theta = a, x^g = b, x^* = q, h = 2, r = 1, \Delta^+ f(a) = 3, \Delta^+ f(c) = 2,$$

$$\Delta^+ f(e) = 3, f(x^*) = 5, f(x^g) = 3.$$

Поэтому $I = \{(a, c)\}$, т.е. $c = 1/3$.

Из теоремы находим $B_1 = 2/5$.

Поэтому

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} = \frac{2}{5} = B_1.$$

Пример. Пусть

$$H = \{a, b, c, d, e\}, a \triangleright b \triangleright e, a \triangleright c \triangleright d,$$

$$f(a) = 0, f(b) = 3, f(c) = 2, f(e) = 3.5, f(d) = 4$$

Тогда очевидно, что

$$a = \theta, h = 2, r = 2, \Delta^+ f(a) = 3, \Delta^+ f(b) = 1.5, \Delta^+ f(c) = 2.$$

То есть

$$I = \{(a, b), (a, c)\}.$$

Тогда

$$c = c(f) = \min \left\{ \frac{\Delta^+ f(a) - \Delta^+ f(b)}{\Delta^+ f(a)}, \frac{\Delta^+ f(a) - \Delta^+ f(c)}{\Delta^+ f(a)} \right\} = \frac{1}{3}.$$

Очевидно, что

$$x^* = d, f(x^*) = 4, x^g = e, f(x^g) = 3.5.$$

Из теоремы и из (4) находим $B_1 = 4/25$, $B_2 = 1/4$, т.е. $B_1 < B_2$.

Отметим, что для этого примера

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} = \frac{1}{8} < B_1.$$

ЛИТЕРАТУРА

1. Emelichev V.A., Kovalev M.M., Ramazanov A.B. Errors of Gradient Extrema of a Strictly Convex Function of Discrete Argument // Discrete Mathematics and Applications, 1992, v. 2, Issue 2, p. 119-131.
2. Рамазанов А.Б. Устойчивость градиентного алгоритма на структурах Жордана-Дедекинда // В кн. Математическое программирование и приложение. ИМ и М Уральское Отделение РАН, Екатеринбург, 2011, с. 204-208.
3. Ramazanov A.B. On Stability of the Gradient Algorithm in Convex Discrete Optimisation Problems and Related Questions // Discrete Mathematics and Applications, 2011, v. 21, Issue 4, p. 465-476.
4. Айгнер М. Комбинаторная теория. М.: Мир, 1982, 558 с.

QRADİYENT ALQORİTMİN XƏTASININ JORDAN-DEDEKİND STRUKTURLARINDA ARAŞDIRILMASI

Ə.B. RAMAZANOV

XÜLASƏ

Məqalədə tərtib-qabarıq funksiyaların Jordan-Dedekind strukturlarında maksimallaşdırma məsələsinə baxılır. Bu məsələ üçün qradiyent alqoritmin əvvəlki xətalrı dəqiqləşdirən və inkişaf etdirən aposterior qarantiyalı xətası alınmışdır. Həmçinin qradiyent həllin optimallığı üçün kafi şərt verilmişdir.

Açar sözlər: qabarıq, xəta, qradiyent, alqoritm, bükülmə.

ANALYSIS OF ACCURACY OF GRADIENT ALGORITHM IN THE JORDAN-DEDEKINDS STRUCTURE

A.B.RAMAZANOV

SUMMARY

In this article, we investigate maximizing ordered-convex functions in Jordan-Dedekind structure. For this problem, we get a-posterior guaranteed enhanced and progressing error. Also, sufficient conditions of optimization of the gradient solution is given.

Key words: convex, errors, gradient, algorithm, steepness.

Поступила в редакцию: 09.10.2014 г.

Подписано к печати: 13.02.2015 г.