Proceedings of IAM, V.3, N.2, 2014, pp.242-248

АНАЛИЗ ТОЧНОСТИ И УСТОЙЧИВОСТИ ГРАДИЕНТНОГО АЛГОРИТМА НА СТРУКТУРАХ ЖОРДАНА-ДЕДЕКИНДА*

А.Б. Рамазанов [1]

[1]Бакинский Государственный Университет, Баку, Азербайджан
e-mail: ram-bsu@mail.ru

Резюме. 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. Установлена устойчивость градиентного алгоритма в терминах гарантированных оценок. Кроме того, приведены достаточные условия оптимальности градиентного решения.

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

AMS Subject Classification: 74P10.

1. Введение

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 [1,5]. 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., [2]). В работе [5] показано, что приведенные в [1] оценки можно улучшить, используя информацию о крутизны 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 [1,2,5]. Установлена устойчивость градиентного алгоритма в терминах гарантированных оценок. Кроме того, приведены достаточные условия оптимальности градиентного решения в терминах крутизны целевой

---

* Работа была представлена на семинаре Института Прикладной Математики 09.12.2014

функции. Отметим, что данная работа является развитием и обобщением работы [5].

2. Определения и обозначения

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

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

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

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

Напомним, что множество элементов $x^0, x^1, ..., x^k$ из H, обладающих свойством $x = x^0 \prec x^1 \prec ... \prec x^k = y$ называется цепью между x и y. Число k называется длиной цепи. Цепь $x = x^0 \rhd x^1 \rhd ... \rhd x^k = y$ называется максимальной цепью между x и y. Будем предполагать, что частично упорядоченное множество H удовлетворяет условию Жордана-Дедекинда [3]: 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

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

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

Функция $f \in \Im_\rho(H)$ называется неубывающей, если из $x \prec y$ следует $f(x) \leq f(y)$.

Правым градиентом функции $f(x)$, как обычно [1, 5], называем функции

$$\Delta^+ f(x) = \max\{ f(y) - f(x) | x \rhd y, \; x, y \in H\} .$$

Крутизной функции $f(x)$ будем называть величину $c = c(f)$ [2, 5]:

$$c = c(f) = \begin{cases} \min\{(\Delta^+ f(x - \Delta^+ f(y)) / \Delta^+ f(x | (x, y) \in I\} , \text{\textit{если}} \quad I \neq \varnothing, \\ 0, \text{\textit{если}} \; I = \varnothing, \end{cases}$$

где $I = \{(x, y) \big| \Delta^+ f(x) > \Delta^+ f(y) \ge 0,\ x \rhd y,\ x, y \in H\}$.

Очевидно, что если $f(x) \in \Im_\rho(H),\ \rho > 0$, то $I \ne \varnothing$.

3. Гарантированная оценка и устойчивость градиентного алгоритма

Рассматривается следующая задача A: найти

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

где $f(x)$ - неубывающая функция из класса $\Im_\rho = \Im_\rho(H)$.

Градиентным (локальным) решением (максимумом) x^g задачи A (функции $f(x)$) на множестве H, назовем точку, построенную с помощью следующей итерационной процедуры [1, 5]:

$$x^{t+1} = \arg\max\{ f(y) - f(x^t) \big| x^t \rhd y, y \in H\}, t = 0,1,..., x^0 = \theta,$$

заканчивающей работы на шаге τ, когда либо $\Delta^+ f(x^\tau) \le 0$, 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 $\varepsilon \ge 0$, что

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

При возмущение крутизны целевой функции задачи A полученную задачу обозначим через - $A(\delta)$: найти

$$\max\{ f_\delta(x) \big| x \in H\},$$

где $f_\delta(x)$ - неубывающая функция из класса $\Im_q = \Im_q(H)$.

Пусть ε и $\varepsilon(\delta)$ соответственно, гарантированные оценки для задачи A и $A(\delta)$. Градиентный алгоритм называется устойчивым для задачи A, если $\varepsilon(\delta) \le K(\delta)\varepsilon$, где $K(\delta) \to 1$ при $\delta \to 0$ [2, 5]. Другими словами, устойчивость означает инвариантность грантированных оценок при “малых” возмущениях параметров исходной задачи.

Теорема 1. Для задачи A справедлива следующая гарантированная оценка точности градиентного алгоритма

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

где $c = c(f)$.

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

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

Отсюда при $y=x^{*}$, $x=x^{t}$, $t=1,...,r$, получаем

$$\begin{gathered}f(x^{*})-f(x^{t})\le h(x^{t},x^{*})\Delta^{+}f(x^{t})-\rho(h(x^{t},x^{*})-1)\le \\ (h-1)\Delta^{+}f(x^{t}),\ t=1,...,r\end{gathered}\ . \tag{1}$$

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

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

Учитывая (2) в (1), выводим

$$f(x^{*})-f(x^{t})\le(1-c)(h-1)\Delta^{+}f(x^{t-1}),\ t=1,2,...,r \tag{3}$$

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

Теорема 1 доказана.

Теорема 2. Градиентный алгоритм устойчив в задаче A при “малых” возмущениях крутизны функции $f(x)$.

Доказательство. Из теоремы 1 соответственно для задачи A и $A(\delta)$, имеем

$$\varepsilon=\left(1-\frac{1}{1+(1-c(f))(h-1)}\right)^{r},\ \varepsilon(\delta)=\left(1-\frac{1}{1+(1-c(f_{\delta}))(h-1)}\right)^{r}.$$

Для доказательства теоремы 2 достаточно показать, что $\varepsilon(\delta)\le\varepsilon$. Пусть $\exists\delta>0\Rightarrow c(f_{\delta})\ge c(f)+\delta$. Тогда из цепочки неравенств, имеем

$$(1-c(f))(h-1)\ge(1-c(f_{\delta}))(h-1),$$

$$\frac{1}{1+(1-c(f))(h-1)}\le\frac{1}{1+(1-c(f_{\delta}))(h-1)},$$

$$1-\frac{1}{1+(1-c(f))(h-1)}\ge1-\frac{1}{1+(1-c(f_{\delta}))(h-1)},$$

$$\varepsilon=\left(1-\frac{1}{1+(1-c(f))(h-1)}\right)^{r}\ge\left(1-\frac{1}{1+(1-c(f_{\delta}))(h-1)}\right)^{r}=\varepsilon(\delta)\ .$$

То есть $\varepsilon(\delta)\le\varepsilon$.

Теорема 2 доказана.

4. Следствия и примеры

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

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

Следующее следствие показывает, что оценка из теоремы 1 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 $f(x^*)=f(x^g)$.

Пусть выполняются условия: 1) если $c \triangleright a$ и $c \triangleright b$, $a \ne b, a,b,c \in H$, то существует такой элемент $d \in H$, что $a \triangleright d$ и $b \triangleright d$; 2) если $q \triangleright l$ и $q \triangleright m$, $l \ne m, q,l,m \in H$, то существует такой элемент $p \in H$, что $p \triangleright l$ и $p \triangleright m$.

Частично упорядоченное множество H удовлетворяющии условиям 1) и 2) называют полумодулярной [4]. Известно (см., напр., [3, 5]), что в любом полумодулярном множестве выполняется условия Жордана-Дедекинда.

Поэтому справедливо

Следствие 3. Все доказанные утверждения справедливы и для полумодулярных множеств.

Следствие 4. Полученная оценка в теореме 1 достижима.

Для доказательства следствие 4 построим пример.

Пример 1. 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

$$I=\{(a,c)\}, \text{ т.е. } c=1/3.$$

Из теоремы 1 находим $B_1 = 2/5$.
Поэтому

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

Пример 2. Пусть

$$H = \{a,b,c,d,e\}, a \rhd b \rhd e, a \rhd c \rhd d,$$
$$f(a) = 0,\ f(b) = 3,\ f(c) = 2,\ f(e) = 3.5,\ f(d) = 4$$

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 1 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, Vol.2, No.2, 1992, pp.119-131.
2. Ramazanov A.B. On stability of the gradient algorithm in convex discrete optimisation problems and related questions, Discrete Mathematics and Applications, Vol.21, No.4, 2011, pp.465-476.
3. Айгнер М. Комбинаторная теория, М.:Мир, 1982, 558 с.
4. Биркгоф Г. Теория решеток, М.:Мир, 1984, 568 с.
5. Рамазанов А.Б. Устойчивость градиентного алгоритма на структурах Жордана-Дедекинда, В кн. Математическое программирование и

приложение. Изд.-во ИМ и М Уральское Отделение РАН, Екатеринбург, 2011, с.204-208.

Qradiyent alqoritmin xətasinin və dayaniqliğinin Jordan-Dedekind structurlarinda araşdirilmasi

Ə.B. Ramazanov

XÜLASƏ

Məqalədə tərtib-qabarıq funksiyaların Jordan-Dedekind strukturlarında maksimallaşdırma məsələsinə baxılır. Bu məsələ üçün qradiyent alqoritmin əvvəlki xətaları dəqiqləşdirən və inkişaf etdirən zəmanətli xəta alınmışdır. Zəmanətli xəta terminində qradiyent alqoritmin dayanıqlığı isbat edilmişdir. Həmçinin qradiyent həllin optimallığı üçün kafi şərt verilmişdir.

Açar sözlər: qabarıq, xəta, qradient, alqoritm, bükülmə, dayanıqlıq.

Analysis of the error and stability of the gradient algorithm in the Jordan-Dedekinds structures

A.B. Ramazanov

ABSTRACT

In this paper, we investigate maximizing ordered-convex functions in Jordan-Dedekind structure. For this problem, we get a-posterior guaranteed enhencied and progressing error. Stability of the gradient algorithm the terms of guaranteed errors is found. Also sufficient condition of optimality of gradient solution is given.

Keywords: convexity, error, gradient, algorithm, steepness, stability.