УДК 519.10

УСТОЙЧИВОСТЬ ГРАДИЕНТНОГО АЛГОРИТМА В ОДНОЙ ЗАДАЧЕ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ

А.Б.РАМАЗАНОВ

Бакинский Государственный Университет

ram-bsu@mail.ru

На основе меры изменения градиентов целевой функции показано, что в нкоторых задачах выпуклой дискретной оптимизации на целочисленной решетке градиентный алгоритм устойчив при "малых" возмущениях параметра меры изменения целевой функции. Как следствие, получены улучшенные гарантированные оценки точности градиентного алгоритма, а также новые достаточные условия, когда значения целевой функции рассматриваемой задачи в глобальном и градиентном экстремумах совпадают.

Ключевые слова: градиент, крутизна, устойчивость, выпуклость, дискрет

Исходные данные многих задач дискретной оптимизации носят приближенный характер. Поэтому актуальным является анализ устойчивости решений при колебаниях параметров задачи. Исследованию различных аспектов устойчивости скалярных и векторных задач дискретной оптимизации посвящены многочисленные публикации (см., напр., [1, 2]). Актуальными являются вопросы устойчивости не только решений задач дискретной оптимизации, но и алгоритмов их решения (см., напр., [3-5]). Одним из возможных вариантов исследования устойчивости локальных (градиентных) алгоритмов является нахождение изменения гарантиированных (относительных) оценок при "малых" возмущениях параметров задачи. Если гарантированные оценки локальных (градиентных) алгоритмов возмущенной задачи не хуже, чем гарантированные оценки исходной задачи, то естественно такой локальный (градиентный) алгоритм можно называть устойчивым [5]. Другими словами, под устойчивостью градиентного алгоритма понимают некоторые свойства инвариантности гарантированных оценок точности градиентного алгоритма при "малых" возмущениях параметров, входящие в эту оценку. Например, в работах [4, 5] в терминах гарантированных оценок установлено, что при "малых" возмущениях чебышевской нормы матрицы входящие в условие ограничения, градиентный алгоритм покоординатного подъема устойчив в некоторых задачах выпуклой дискретной оптимизации.

В данной работе на основе введенного в [6] меры изменения градиентов целевой функции показано, что в некоторых задачах выпуклой дискретной оптимизации на целочисленной решетке градиентный алгоритм устойчив при “малых” возмущениях параметра меры изменения целевой функции. Как следствия получены улучшенные гарантированные оценки точности градиентного алгоритма, а также новые достаточные условия, когда значения целевой функции рассматриваемой задачи в глобальном и градиентном экстремумах совпадают.

Определения и обозначения

Пусть Z_+^n (R_+^n) - множество $n-$мерных неотрицательных целочисленных (действительных) векторов и $P \subseteq Z_+^n$. Будем в дальнейшем считать, что множество P обладает свойствами:

(i) $|P| < \infty$;

(ii) $0 = (0,...,0) \in P$;

(iii) $[0, x] = \{z \in Z_+^n |\ 0 \le z \le x\} \subseteq P$ для любого $x \in P$.

Введем следующие обозначения:

$N(x, y) = \{i|\ x_i < y_i, 1 \le i \le n\}$, $h(x, y) = \sum\limits_{i \in N(x,y)} h(x_i, y_i)$,

$h(x_i, y_i) = \left|\{z_i|\ x_i \le z_i \le y_i\}\right| - 1, 1 \le i \le n$, $h(x) = h(0, x)$,

$h = h(P) = \max\{h(0, x)|\ x \in P\}$, $fes(x, P) = \{1 \le i \le n|\ \pi_i(x) \in P, x \in P\}$,

$\pi_i(x) = (x_1,..., x_{i-1}, x_i + 1, x_{i+1},..., x_n), r = r(P) = \min\{h(x) - 1|\ x \in Z_+^n \setminus P\}$

Следуя [5, 7], для функции $f|\ Z_+^n \to R$ (R - множество действиительных чисел) введем понятия i - градиента

$$\Delta_i f(x) = f(\pi_i(x)) - f(x),$$

и (i, j) - градиента

$$\Delta_{ij} f(x) = \Delta_j f(\pi_i(x)) - \Delta_j f(x).$$

Пусть $\rho = (\rho_1,..., \rho_n) \in R_+^n$, $\Re_\rho(Z_+^n)$ - класс ρ-координатно-выпуклых функций на Z_+^n [5, 7], т.е. таких функций $f|\ Z_+^n \to R$, что для любого $x \in Z_+^n$

$$\Delta_{ij} f(x) \le 0, i \ne j, 1 \le i, j \le n,$$

$$\Delta_{ii} f(x) \le -\rho_i, 1 \le i \le n.$$

Определение. Пусть $f(x)$ - неубывающая ρ-координатно-выпуклая функция на множестве $P \subseteq Z_+^n$. Крутизной функции $f(x)$ на множестве $P \subseteq Z_+^n$ называется величина [6]

$$c = c(f) = \min\left\{\frac{\Delta_i f(x) - \Delta_j f(y)}{\Delta_i f(x)} : (i, j, x, y) \in I^+(f) = I^+ \right\},$$

где

$$I^+ = I^+(f) = \{(i, j, x, y) : \Delta_i f(x) > \Delta_j f(y) \geq 0, i \in fes(x, P), j \in fes(y, P), x < y\}$$

Постановка задачи

Рассматривается следующая задача A выпуклой дискретной оптимизации: найти

$$\max\{f(x)|\ x = (x_1,..., x_n) \in P \subseteq Z_+^n\},$$

где $f(x) \in \Re_\rho(Z_+^n)$, $f(x) -$ неубывающая функция на множестве P.

Пусть $x^* = (x_1^*,..., x_n^*)$ - оптимальное решение задачи A, т.е. $f(x^*) = \max\{f(x)|\ x \in P\}$.

Через $x^g = (x_1^g,..., x_n^g)$ обозначим градиентное решение (градиентный максимум функции $f(x)$ на множестве P) задачи A, т.е. x^g точка из P, полученная с помощью градиентного алгоритма покоординатного подъема (см., напр., [5, 7, 8]).

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

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(0)} \leq \varepsilon .$$

Возмущением задачи A будем называть следующие задачи A^δ: найти

$$\max\{f^\delta(x)|\ x \in P \subseteq Z_+^n\},$$

где $f^\delta(x) \in \Re_q(Z_+^n)$ - неубывающая функция на множестве $P, c(f^\delta) = c(f) + \delta, \delta \in R_+^1, q = (q_1,..., q_n)$.

Класс задач типа A^δ не пуст. Действительно, рассмотрим следующие задачи A_1^δ: найти

$$\max\{f^\delta(x) = f(x) - (b, x)|\ x \in P \subseteq Z_+^n\},$$

где $f^\delta(x), f(x)$ - неубывающие функции на множестве P, (b, x) - скалярное произведение векторов $x, b \in Z_+^n$, $I^+(f^\delta) \neq \varnothing$, $f(x) \in \Re_\rho(Z_+^n)$. Тогда очевидно, что $f^\delta(x) \in \Re_\rho(Z_+^n), c(f) < c(f^\delta)$, т.е. $\exists \delta \in R_+^1 \Rightarrow c(f^\delta) = c(f) + \delta$.

Пусть $\varepsilon(\delta)$ и ε, соответственно, гарантированные оценки возмущенной задачи A^{δ} и исходной задачи A. Градиентный алгоритм покоординатного подъема назовем устойчивым [4, 5] для задачи A, если

$$\varepsilon(\delta) \le K(\delta)\varepsilon,$$

где $K(\delta) \to 1$ при $\delta \to 0$.

По существу устойчивость градиентного алгоритма в терминах гарантированных оценок означает выделение класса задач, для которого при “малых” возмущениях параметров задачи (в частности крутизны целевой функции), гарантированные оценки для возмущенных задач не ухудшаются.

Основной целью настоящей работы является установление устойчивости градиентного алгоритма покоординатного подъема для задачи A в терминах крутизны целевой функции.

Основной результат

Сформулируем основное утверждение данной работы.

Теорема. Градиентный алгоритм покоординатного подъема устойчив в задаче A при “малых” возмущениях крутизны целевой функции.

Для доказательства теоремы нам понадобятся следующие леммы.

Лемма 1 [7]. Если функция $f(x) \in \Re_{\rho}(Z_{+}^{n})$ - неубывающая, то справедливы неравенства

$$f(y) - f(x) \le \sum_{i \in N(x,y)} h(x_i, y_i)\Delta_i f(x) - \frac{1}{2}\sum_{i \in N(x,y)} \rho_i h(x_i, y_i)(h(x_i, y_i) - 1),$$

$$\forall x \le y, x, y \in Z_{+}^{n}$$

Лемма 2. Пусть функция $f(x) \in \Re_{\rho}(Z_{+}^{n})$ - неубывающая на множестве P. Тогда для любых $x \in P$ и $i \in fes(x,P), j \in fes(\pi_i(x),P)$ справедливы неравенства

$$\Delta_j f(\pi_i(x)) \le (1-c)\Delta_i f(x),$$

где $c = c(f)$.

Доказательство леммы 2 вытекает непосредственно из определения крутизны.

Лемма 3. Пусть функция $f(x) \in \Re_{\rho}(Z_{+}^{n})$ - неубывающая на множестве P. Тогда глобальный максимум x^{*} и градиентный максимум x^{g} функции $f(x)$ на множестве $P \subseteq Z_{+}^{n}$ связаны соотношением

$$\frac{f(x^{*}) - f(x^{g})}{f(x^{*}) - f(0)} \le B\ (c,h,r) = B,$$

где

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

Доказательство леммы 3. Из леммы 1, имеем

$$f(y) - f(x) \le \sum_{i \in N(x,y)} h(x_i, y_i)\Delta_i f(x), \ \forall x \le y, x, y \in Z_+^n$$

Отсюда, полагая $y = x^*, x = x^t, \ t = 1,...,r$ с учетом

$$\sum_{i \in N(x^t, x^*)} h(x_i^t, x_i^*) \le h-1, \ \Delta_i f(x^t) \le \Delta_{i(t)} f(x^t), \ f(x^t) = f(0) + \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s),$$

$$i(s) = \arg\max_i \{\Delta_i f(x^s) \mid i \in fes(x^s, P)\}, 0 \le s \le t-1, 1 \le t \le r, \quad f(x^0) = f(0),$$

имеем

$$f(x^*) - f(0) - \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s) \le (h-1)\Delta_{i(t)} f(x^t), t = 1,...,r$$

Из леммы 2, при $j = i(t), i = i(t-1), x = x^{t-1}$, получаем

$$\Delta_{i(t)} f(x^t) \le (1-c)\Delta_{i(t-1)} f(x^{t-1}), 1 \le t \le r.$$

Поэтому

$$f(x^*) - f(0) - \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s) \le (h-1)(1-c)\Delta_{i(t-1)} f(x^{t-1}), t = 1,...,r.$$

Далее, принимая обозначения

$$\alpha_t = \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s), t = 1,...,r, \ \alpha_0 = 0,$$

и повторяя процедуры аналогичные как в [7] с учетом обозначений доказываем леммы 3.

Лемма 3 доказана.

Доказательство теоремы. Пусть ε и $\varepsilon(\delta)$, соответственно, гарантированные оценки для задачи A и A^δ. Тогда из леммы 3 имеем

$$\varepsilon = B(c(f), h, r), \ \varepsilon(\delta) = B(c(f^\delta), h, r).$$

Поэтому для доказательства теоремы достаточно показать, что

$$B(c(f^\delta), h, r) \le B(c(f), h, r)$$

Возмущение крутизны функции $f(x)$ обозначим через δ.

Тогда из цепочки неравенств

$$1 + (1 - c(f) - \delta)(h-1) \le 1 + (1 - c(f))(h-1),$$

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

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

$$\left(1-\frac{1}{1+(1-c(f))(h-1)}\right)^r=B(c(f),h,r),$$

имеем

$$\varepsilon(\delta)=B(c(f^{\delta}),h,r)=B(c(f)+\delta,h,r)\le B(c(f),h,r)=\varepsilon$$

Теорема доказана.

Следствия. Пример

Приведем ряд следствий непосредственно вытекающих из леммы 3.

Следствие 1. В условиях леммы 3, если $c=c(f)=1$, то $f(x^*)=f(x^g)$.

Покажем, что существуют такие функции из класса $\Re_\rho(Z_+^n)$, что $c=1$.

Действительно, пусть

$$f(x)=-2x_1^2+6x_1-x_2^2+3x_2+2x_1x_2,\, P=\{(0,0),(1,0),(0,1),(0,2)\}.$$

Тогда

$$\Delta_1 f(x)=-4x_1+4,\Delta_2 f(x)=-2x_2+2,\ \rho=(4,2),\Delta_1 f(0,0)=4,\Delta_2 f(0,0)=2.$$

Поэтому $c=1$.

Построим пример, когда $f(x^*)=f(x^g)$, но $c\ne 1$. То есть покажем, что утверждение следствия 1 является достаточным, но не необходимым.

Пусть

$$f(x)=-2x_1^2+7x_1-x_2^2+4x_2,\ P=\{(0,0),(0,1),(0,2),(1,0)\}.$$

Тогда

$$\Delta_1 f(x)=-4x_1+5,\Delta_2 f(x)=-2x_2+3,\rho=(4,2),\Delta_1 f(0,0)=5,$$

$$\Delta_2 f(0,1)=1,x^*=x^g=(1,0),\ f(x^*)=f(x^g)=5,$$

но

$$c(f)=4/5\ne 1$$

Известно (см. теорему 11.1 [8]), что если, в задаче $A\ \ f(x)\in\Re_0(Z_+^n)$, то справедлива оценка

$$\frac{f(x^*)-f(x^g)}{f(x^*)-f(0)}\le\left(1-\frac{1}{h}\right)^r=B_1 \qquad (1)$$

Следующее следствие показывает, что оценка из леммы 3 лучше, чем оценка (1).

Следствие 2. Справедливы соотношения

$$B(c,h,r)\begin{cases} = B_1,\ если\ c=0, \\ =0,\ \ если\ c=1\ или\ h=1, \\ < B_1,\ если\ 0<c<1, h>1, \end{cases}$$

где $c = c(f)$.

Доказательство. Пункты 1 и 2 очевидны. Докажем третий пункт. Так как $h>1, 0<c<1$, то имеем $(1-c)(h-1)<h-1$. Поэтому $1+(1-c)(h-1)<h$. Отсюда находим

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

т.е.

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

Следствие 2 доказано.

Пример. Рассмотрим задачу.

$$\max\{f(x)=-x_1^2+6x_1-x_2^2+5x_2 \mid x=(x_1,x_2)\in P\}, \tag{2}$$

где $P=\{(0,0),(1,0),(0,1),(0,2)\}$.

Очевидно, что $h=h(P)=2,\ r=r(P)=1$. Легко видеть, что

$\Delta_1 f(x)=-2x_1+5, \Delta_2 f(x)=-2x_2+4, \Delta_{11} f(x)=-2, \Delta_{22} f(x)=-2, I^+(f)\neq\varnothing$,

т.е. $f(x)\in\Re_{(2,2)}(Z_+^2)$ и $f(x)$ - неубывающая функция на множестве P. Тогда

$c(f)=3/5$.

Возмущение задачи (2) будет следующая задача

$$\max\{f^\delta(x)=-x_1^2+6x_1-x_2^2+5x_2-(x_1+x_2/2) \mid x=(x_1,x_2)\in P\}. \tag{3}$$

Очевидно, что $f^\delta(x)\in\Re_{(2,2)}(Z_+^2)$ и $f^\delta(x)$ - неубывающая функция на множестве P, $I^+(f^\delta)\neq\varnothing$. Тогда $c(f^\delta)=2/3$. Так как $c(f^\delta)=c(f)+\delta,\quad \delta=1/15$, то в силу теоремы градиентный алгоритм устойчив в задаче (2) при возмущении крутизны целевой функции $f(x)$. Действительно, пусть ε_1 (ε_2) гарантированная оценка задачи (2) (задачи (3)). Из леммы 3 находим $\varepsilon_1=2/7$, $\varepsilon_2=1/4$. Тогда $\varepsilon_2<\varepsilon_1$.

ЛИТЕРАТУРА

1. Emelichev V., Podkapaev D. Quantitative stability analysis for vectpr problems of 0-1 programming // Discrete Optimization . 2010, v.7, p. 48-63.
2. Гордеев Э.Н., Леонтьев В.К. Общий подход к исследованию устойчивости решений в задачах дискретной оптимизации // Журн. вычисл. матем. и матем. физики. 1996, т. 36, № 1, с. 66-72.
3. Девятерикова М.В., Колоколов А.А. Анализ устойчивости некоторых алгоритмов дискретной оптимизации // Автоматика и Телемеханика. 2004, №3, с.48-54.

4. Рамазанов А.Б. Об оценке точности градиентного алгоритма в одной задаче выпуклой дискретной оптимизации и некоторые смежные вопросы. В кн. Труды 13-й Байкальской Международной семинар-школы. Иркутск: Инст.-т систем энергетики им. Л.А. Мелентьева СО РАН, 2005, с. 571-576.
5. Ramazanov A.B. An Estimate for the Curvature of an Order-Convex Set in the Integer Lattice and Related Questions // Mathematical Notes. 2008, v. 84, №1, p.147-151.
6. Рамазанов А.Б. О точности жадного алгоритма на структурах Жордана-Дедекинда. В кн. Дискретный анализ и исследование операций. Новосибирск: Инс-т Математики им. С.Л. Соболева СО РАН, 2010, с. 108.
7. Рамазанов А.Б. Оценки точности получаемых алгоритмом покоординатного подъема решений задач дискретной выпуклой оптимизации. // Дискретный анализ и исследование операций. 2005, т. 12, сер. 1, № 4, с. 60-80.
8. Ковалев М.М. Матроиды в дискретной оптимизации. Минск: Бел.гос. унив., 1987, 222 с.

BİR DİSKRET OPTİMALLAŞDIRMA MƏSƏLƏSİNDƏ QRADİYENT ALQORİTMİN DAYANIQLILIĞI

Ə.B. RAMAZANOV

XÜLASƏ

Diskret arqumentli funksiyalar üçün bükülmə anlayışı daxil edilir və həmin parametr əsasında qradiyent alqoritmin bir diskret optimallaşdırma məsələsində qarantiyalı xətaya görə dayanıqlığı isbat edilir. Həmçinin nəticə kimi baxılan məsələdə qradiyent alqoritmin yeni daha dəqiq qarantiyalı xətası və funksional mənada qradiyent həllin optimallığı üçün kafi şərt verilmişdir.

Açar sözlər: qradiyent, bükülmə, dayanıqlıq, qabarıq, diskret

STABILITY OF GRADIENT ALGORITHM FOR A PROBLEM OF DISCRETE OPTIMIZATION

A.B.RAMAZANOV

SUMMARY

The conception of steepness for discretely argumented functions is presented and by means of this parameter the stability of the gradient algorithm in the discrete optimization problem for the guaranteed error is proved. A sufficient condition for a new quite guaranteed error and for the optimization of the gradient solution in the functional meaning in the sought problem is provided.

Key words: gradient, steepness, stability, conves, discrete

Поступила в редакцию: 05.10.2011 г.
Принята к печати: 05.03.2012 г.