Math-Net.Ru

Общероссийский математический портал

А. Б. Рамазанов, Об устойчивости градиентного алгоритма в задачах выпуклой дискретной оптимизации и некоторые смежные вопросы, *Дискрет. матем.*, 2011, том 23, выпуск 3, 82–92

DOI: <https://doi.org/10.4213/dm1154>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.80.73.141

26 апреля 2023 г., 14:42:15

Об устойчивости градиентного алгоритма в задачах выпуклой дискретной оптимизации и некоторые смежные вопросы

© 2011 г. А. Б. Рамазанов

В работе введено понятие крутизны координатно-выпуклой функции дискретного аргумента на порядково-выпуклом множестве. В терминах гарантированных оценок показано, что в задачах максимизации координатно-выпуклых функций на порядково-выпуклом множестве градиентный алгоритм покоординатного подъема устойчив при малых возмущениях крутизны 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.

1. Введение

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., например, [1–3]). Актуальными являются вопросы устойчивости не только решений задач дискретной оптимизации, но и алгоритмов их решения (см., например, [4]). Одним из возможных вариантов исследования устойчивости локальных (градиентных) алгоритмов является нахождение изменения гарантированных (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 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. Так например, в работах [6, 7] в терминах гарантированных оценок установлено, что при малых возмущениях чебышёвской нормы матрицы, входящей в условие ограничения, градиентный алгоритм покоординатного подъема устойчив в некоторых задачах выпуклой дискретной оптимизации.

В данной работе введено понятие крутизны координатно-выпуклой функции дискретного аргумента на порядково-выпуклом множестве. В терминах гарантированных оценок показано, что в задачах максимизации координатно-выпуклых функций на порядково-выпуклом множестве градиентный алгоритм покоординатного подъема устойчив при малых возмущениях крутизны целевой функции. Как следствия, получены улучшенные

гарантированные оценки точности градиентного алгоритма, а также новые достаточные условия того, что значения целевой функции рассматриваемой задачи в глобальном и градиентном экстремумах совпадают.

Отметим, что первоначально понятие крутизны для субмодулярных функций введено и использовано в [8] для улучшения оценки погрешности градиентного алгоритма.

2. Определения и обозначения

Пусть $\mathbf{Z}_+^n$ (соответственно $\mathbf{R}_+^n$) — множество n -мерных неотрицательных целочисленных (действительных) векторов. Для $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbf{Z}_+^n$ будем писать $x = (x_1, \dots, x_n) \leq y = (y_1, \dots, y_n)$, если $x_i \leq y_i$, $i = 1, 2, \dots, n$. Если среди последних неравенств есть хотя бы одно строгое, то будем писать $x < y$. Пусть $P \subseteq \mathbf{Z}_+^n$. Будем в дальнейшем считать, что множество P обладает следующими свойствами:

- (i) $|P| < \infty$,
- (ii) $0 = (0, \dots, 0) \in P$,
- (iii) $[0, x] = \{z \in \mathbf{Z}_+^n \mid 0 \leq z \leq x\} \subseteq P$ для любого $x \in P$.

Множество P , обладающее свойствами (i)–(iii), будем называть конечным порядково-выпуклым множеством с нулем (см., например, [7]).

Введем следующие обозначения:

$$\begin{aligned}
 N(x, y) &= \{i \mid x < y, 1 \leq i \leq n\}, \\
 h(x, y) &= \sum_{i \in N(x, y)} h(x_i, y_i), \\
 h(x_i, y_i) &= |\{z_i \mid x_i \leq z_i \leq y_i\}| - 1, \quad 1 \leq i \leq n, \\
 h(x) &= h(0, x), \\
 h &= h(P) = \max\{h(0, x) \mid x \in P\}, \\
 \text{fes}(x, P) &= \{1 \leq i \leq n \mid \pi_i(x) \in P, x \in P\}, \\
 \pi_i(x) &= (x_1, \dots, x_{i-1}, x_i + 1, x_{i+1}, \dots, x_n), \\
 r &= r(P) = \min\{h(x) - 1 \mid x \in \mathbf{Z}_+^n \setminus P\}.
 \end{aligned}$$

Следуя [9–12], для функции $f: \mathbf{Z}_+^n \rightarrow \mathbf{R}$ ($\mathbf{R}$ — множество действительных 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 $\rho = (\rho_1, \dots, \rho_n) \in \mathbf{R}_+^n$, $\mathfrak{R}_\rho(\mathbf{Z}_+^n)$ — класс ρ -координатно-выпуклых функций на $\mathbf{Z}_+^n$ [10, 12], то есть таких функций $f: \mathbf{Z}_+^n \rightarrow \mathbf{R}$, что для любого $x \in \mathbf{Z}_+^n$

$$\begin{aligned}
 \Delta_{ij} f(x) &\leq 0, \quad i \neq j, \quad 1 \leq i, j \leq n, \\
 \Delta_{ii} f(x) &\leq -\rho_i, \quad 1 \leq i \leq n.
 \end{aligned}$$

Если $f(x) \in \mathfrak{R}_0(\mathbf{Z}_+^n)$ (то есть, если $\rho = (0, \dots, 0)$), то функция $f(x)$ называется координатно-выпуклой [11]. Кроме того, очевидно, что если $\rho \leq \mu$, то $\mathfrak{R}_\mu(\mathbf{Z}_+^n) \subseteq \mathfrak{R}_\rho(\mathbf{Z}_+^n)$.

Функции $f: \mathbf{Z}_+^n \rightarrow \mathbf{R}$ будем называть неубывающими на множестве $P \subseteq \mathbf{Z}_+^n$, если для любых $i \in \text{fes}(x, P)$ и $x \in P$ имеет место неравенство $\Delta_i f(x) \geq 0$ (то есть неубывание функции $f(x)$ требуется вдоль координатных цепей).

Определение 1. Пусть $f(x)$ — неубывающая ρ -координатно-выпуклая функция на множестве $P \subseteq \mathbf{Z}_+^n$. Крутизной функции $f(x)$ на множестве $P \subseteq \mathbf{Z}_+^n$ называется величина

$$c = c(f) = \begin{cases} \min(\rho_i / \Delta_i f(0) \mid \Delta_i f(0) > 0, i \in \text{fes}(0, P)), & I^+(f) \neq \emptyset, \\ 1, & I^+(f) = \emptyset, \end{cases}$$

где

$$I^+ = I^+(f) = \{i \mid \Delta_i f(\pi_i(0)) \geq 0, i \in \text{fes}(0, P)\}.$$

Отметим, что если $f(x)$ — неубывающая ρ -координатно-выпуклая функция на $\mathbf{Z}_+^n$, то очевидно, что всегда $I^+ \neq \emptyset$. Однако, если функция $f(x)$ неубывающая на множестве $P \subseteq \mathbf{Z}_+^n$, то может быть $I^+ = \emptyset$. Действительно, пусть

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

Очевидно, что функция $f(x)$ — неубывающая на множестве P и

$$\begin{aligned} \Delta_1 f(x) &= -x_1 + 1/2, & \Delta_2 f(x) &= -x_2 + 1/2, \\ \Delta_{11} f(x) &= \Delta_{22} f(x) = -1, & \Delta_1 f(0) &= 1/2 > 0, \\ \Delta_2 f(0) &= 1/2 > 0, & \Delta_1 f(\pi_1(0)) &= \Delta_1 f(1, 0) = -1/2, \\ \Delta_2 f(\pi_2(0)) &= \Delta_2 f(0, 1) = -1/2, \end{aligned}$$

то есть $f(x) \in \mathfrak{R}_{(1,1)}(\mathbf{Z}_+^n)$, $I^+ = \emptyset$.

Кроме того, следует отметить, что $0 \leq c \leq 1$. Неравенство $c \geq 0$ очевидно. Если $I^+ \neq \emptyset$ (случай $I^+ = \emptyset$ очевиден), то в силу неубывания функции $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$ на множестве P для любых $i \in \text{fes}(0, P)$

$$\Delta_i f(0) \geq \Delta_i f(\pi_i(0)) + \rho_i \geq \rho_i,$$

то есть $c \leq 1$.

Отметим также, что если $I^+ = \emptyset$ (то есть $c = 1$), то в силу неубывания функции $f(x)$ на множестве $P \subseteq \mathbf{Z}_+^n$ справедливо равенство

$$\max\{f(x) \mid x \in P\} = \max\{f(\pi_i(0)) \mid i \in \text{fes}(0, P)\}$$

(полезно сравнить с описанием градиентного алгоритма и с утверждением следствия 1). Поэтому, не умаляя общности, можно предполагать, что $I^+ \neq \emptyset$. Однако для полноты изложения и для корректности определения крутизны (см. вышеприведенный пример) следует добавить случай $I^+ = \emptyset$, что не является особо стеснительным ограничением.

3. Постановка задачи

Рассматривается следующая задача A выпуклой дискретной оптимизации: найти

$$\max\{f(x) \mid x = (x_1, \dots, x_n) \in P \subseteq \mathbf{Z}_+^n\},$$

где $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$, $f(x)$ — неубывающая функция на множестве P , P — порядково-выпуклое множество.

Пусть $x^* = (x_1^*, \dots, x_n^*)$ — оптимальное решение задачи A , то есть

$$f(x^*) = \max\{f(x) \mid x \in P\}.$$

Пусть $x^g = (x_1^g, \dots, x_n^g)$ — градиентное решение (градиентный максимум функции $f(x)$ на множестве P) задачи A , то есть точка, полученная с помощью следующего градиентного алгоритма покоординатного подъема [9–12]:

$$\begin{aligned} x^{t+1} &= \pi_{i(t)}(x^t), \quad t = 0, 1, \dots, \quad x^0 = 0 = (0, \dots, 0), \\ i(t) &= \arg \max_i \{\Delta_i f(x^t) \mid i \in \text{fes}(x^t, P)\}, \end{aligned}$$

заканчивающего свою работу на шаге τ , если $\Delta_{i(\tau)} f(x^\tau) \leq 0$ или $\text{fes}(x^\tau, P) = \emptyset$.

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

$$\max\{f^\delta(x) \mid x \in P \subseteq \mathbf{Z}_+^n\},$$

где $f^\delta(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$ — неубывающая функция на множестве P , $c(f^\delta) = c(f) + \delta$, $\delta \in \mathbf{R}_+^1$, P — порядково-выпуклое множество.

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

$$\max\{f^\delta(x) = f(x) + g(x) \mid x \in P \subseteq \mathbf{Z}_+^n\},$$

где $f^\delta(x)$, $-g(x)$ — неубывающие функции на множестве P , $I^+(f^\delta) \neq \emptyset$, $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$, $g(x) \in \mathfrak{R}_\mu(\mathbf{Z}_+^n)$.

Очевидно, что $f^\delta(x) \in \mathfrak{R}_{\rho+\mu}(\mathbf{Z}_+^n)$. Поэтому учитывая, что $\mathfrak{R}_{\rho+\mu}(\mathbf{Z}_+^n) \subseteq \mathfrak{R}_\rho(\mathbf{Z}_+^n)$, находим, что

$$f^\delta(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n). \quad (1)$$

В силу неубывания функций $f^\delta(x)$ и $-g(x)$ на множестве P , находим, что

$$0 \leq \Delta_i f^\delta(x) = \Delta_i f(x) + \Delta_i g(x) \leq \Delta_i f(x), \quad (2)$$

то есть $f(x)$ — неубывающая функция на множестве P .

Из (2) следует, что для всех $i \in \text{fes}(0, P)$ справедливо неравенство

$$\Delta_i f^\delta(0) \leq \Delta_i f(0). \quad (3)$$

Так как $I^+(f^\delta) \neq \emptyset$, то из (2) находим, что

$$0 \leq \Delta_i f^\delta(\pi_i(0)) \leq \Delta_i f(\pi_i(0)),$$

то есть

$$I^+(f) \neq \emptyset. \quad (4)$$

Кроме того, в силу (3)

$$\{i \mid \Delta_i f^\delta(0) > 0, i \in \text{fes}(0, P)\} \subseteq \{i \mid \Delta_i f(0) > 0, i \in \text{fes}(0, P)\}. \quad (5)$$

Поэтому

$$\begin{aligned} c(f^\delta) &= \min \left\{ \frac{\rho_i}{\Delta_i f^\delta(0)} \mid \Delta_i f^\delta(0) > 0, i \in \text{fes}(0, P) \right\} \\ &\geq \min \left\{ \frac{\rho_i}{\Delta_i f(0)} \mid \Delta_i f(0) > 0, i \in \text{fes}(0, P) \right\} = c(f), \end{aligned}$$

то есть существует такое число $\delta \in \mathbf{R}_+^1$, что

$$c(f^\delta) = c(f) + \delta.$$

В частности, если в задаче A_1^δ функция $g(x)$ имеет вид $g(x) = -(b, x)$, где $b = (b_1, \dots, b_n) \in \mathbf{R}_+^n$, $x = (x_1, \dots, x_n) \in \mathbf{Z}_+^n$, (b, x) — скалярное произведение векторов b и x (очевидно, что в этом случае соотношения (1)–(5) выполняются), то получаем другой класс задач типа A^δ .

С учетом соотношений (1)–(5), можно сформулировать следующие свойства крутизны.

Пусть $f(x)$ — неубывающая ρ -координатно-выпуклая функция на множестве $P \subseteq \mathbf{Z}_+^n$. Тогда

- (1) $c(af) = c(f)$ для всех $a \in \mathbf{R}_+^1$;
- (2) $c(f + g) \leq c(f)$, где $g(x) \in \mathfrak{R}_q(\mathbf{Z}_+^n)$ — неубывающая функция на множестве $P \subseteq \mathbf{Z}_+^n$, $I^+(g) \neq \emptyset$;
- (3) $c(f + \varphi) \geq c(f)$, где $\varphi(x) \in \mathfrak{R}_\beta(\mathbf{Z}_+^n)$, $-\varphi(x)$ — неубывающая функция на множестве $P \subseteq \mathbf{Z}_+^n$, $I^+(f + \varphi) \neq \emptyset$.

Справедливость п. 1 вытекает из соотношения

$$af(x) \in \mathfrak{R}_{a\rho}(\mathbf{Z}_+^n), \quad a\rho = (a\rho_1, \dots, a\rho_n), \quad \Delta_i af(0) = a\Delta_i f(0)$$

и из определения крутизны.

П. 3 непосредственно вытекает из соотношений (1)–(5), примененных к функции $f(x) + \varphi(x)$. П. 2 доказывается аналогично, как п. 3.

Тем самым эти свойства, в некотором смысле, позволяют конструировать возмущенные задачи.

Пусть $\varepsilon(\delta)$ и ε соответственно — гарантированные оценки возмущенной задачи A^δ и исходной задачи A . Градиентный алгоритм покоординатного подъема назовем устойчивым [6, 7] для задачи A , если

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

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

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.

4. Основной результат

Сформулируем основное утверждение данной работы.

Теорема 1. *Градиентный алгоритм покоординатного подъема устойчив в задаче A при малых возмущениях крутизны целевой функции.*

Для доказательства теоремы нам понадобятся следующие леммы.

Лемма 1 ([10]). *Если $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$ — неубывающая функция, то справедливо неравенство*

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

для всех $x \leq y$, $x, y \in \mathbf{Z}_+^n$.

Лемма 2. *Пусть $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$ — неубывающая функция на множестве P . Тогда для любых $x \in P$ и $i \in \text{fes}(x, P)$ справедливо неравенство*

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

где $c = c(f)$.

Доказательство. Если $I^+ = I^+(f) = \emptyset$, то неравенство из леммы 2 очевидно. Пусть $I^+ \neq \emptyset$. Тогда из определения класса $\mathfrak{R}_\rho(\mathbf{Z}_+^n)$ следует, что

$$\Delta_i f(\pi_i(x)) - \Delta_i f(x) \leq -\rho_i, \quad (6)$$

$$\Delta_i f(0) \geq \Delta_i f(\pi_i(0)) \geq \dots \geq \Delta_i f(x). \quad (7)$$

Из определения крутизны получаем, что для любых $i \in \text{fes}(0, P)$

$$c \leq \frac{\rho_i}{\Delta_i f(0)}. \quad (8)$$

Поэтому из (6)–(8) следует, что для любых $x \in P$ и $i \in \text{fes}(x, P)$, с учетом того, что $\text{fes}(0, P) \subseteq \text{fes}(x, P)$ (в силу порядковой выпуклости множества P), справедливы неравенства

$$c \leq \frac{\rho_i}{\Delta_i f(0)} \leq \frac{\Delta_i f(x) - \Delta_i f(\pi_i(x))}{\Delta_i f(x)}.$$

Отсюда вытекают неравенства леммы 2. Лемма 2 доказана.

Замечание 1. Переписав неравенства из леммы 2 в виде

$$c \leq \frac{\Delta_i f(x) - \Delta_i f(\pi_i(x))}{\Delta_i f(x)}, \quad \Delta_i f(x) > 0,$$

легко увидеть, что крутизна является нижней границей замедленного убывания градиентов функции $f(x)$ вдоль координатных осей. Возникает вопрос, можно ли определить крутизну функции $f(x)$ по формуле

$$c = \min \left\{ \frac{\Delta_i f(x) - \Delta_i f(\pi_i(x))}{\Delta_i f(x)} \mid i \in \text{fes}(x, P), x \in P, \Delta_i f(x) > 0 \right\}$$

(аналогично тому, как это делается в [8]). Однако вычисление крутизны с помощью этой формулы в вычислительном плане — трудно решаемая задача. Поэтому определение крутизны, данное в предыдущем параграфе, более конструктивно и допускает более простое вычисление.

Лемма 3. Пусть $f(x) \in \mathfrak{R}_\rho(\mathbf{Z}_+^n)$ — неубывающая на множестве P функция. Тогда глобальный максимум x^* и градиентный максимум x^g функции $f(x)$ на множестве $P \subseteq \mathbf{Z}_+^n$ связаны соотношением

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

где

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

Доказательство. Из леммы 1 следует, что для всех $x, y \in \mathbf{Z}_+^n$, $x \leq y$,

$$f(y) - f(x) \leq \sum_{i \in N(x, y)} h(x_i, y_i) \Delta_i f(x).$$

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

$$\begin{aligned} \sum_{i \in N(x^t, x^*)} h(x_i^t, x_i^*) &\leq h, \quad \Delta_i f(x^t) \leq \Delta_{i(t)} f(x^t), \\ f(x^t) &= f(0) + \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s), \quad i(s) = \arg \max_i \{ \Delta_i f(x^s) \mid i \in \text{fes}(x^s, P) \}, \end{aligned}$$

где $0 \leq s \leq t-1$, $1 \leq t \leq r$, и $f(x^0) = f(0)$, находим, что

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

Из леммы 2 получаем, что

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

Поэтому

$$f(x^*) - f(0) - \sum_{s=0}^{t-1} \Delta_{i(s)} f(x^s) \leq h(1 - c) \Delta_{i(t)} f(x^{t-1}), \quad t = 1, \dots, r. \quad (9)$$

Далее, используя обозначения

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

и учитывая, что

$$\Delta_{i(t)} f(x^{t-1}) = f(x^t) - f(x^{t-1}) = \alpha_t - \alpha_{t-1}, \quad t = 1, \dots, r,$$

из (9) получаем, что

$$\beta_t - 1 \geq \left(1 - \frac{1}{1 + (1-c)h}\right) (\beta_{t-1} - 1), \quad t = 1, \dots, r,$$

где

$$\beta_t = \frac{\alpha_t}{f(x^*) - f(0)}.$$

Поэтому при $t = 1, \dots, r$

$$\beta_t - 1 \geq \left(1 - \frac{1}{1 + (1-c)h}\right) (\beta_{t-1} - 1) \geq \dots \geq \left(1 - \frac{1}{1 + (1-c)h}\right)^r (\beta_0 - 1).$$

Отсюда при $t = r$, с учетом соотношений

$$\begin{aligned} f(x^r) &\leq f(x^g), \quad \alpha_0 = 0, \quad \alpha_r = f(x^r) - f(0) \leq f(x^g) - f(0), \\ \beta_r &= \alpha_r / (f(x^*) - f(0)), \quad \beta_0 = \alpha_0 / (f(x^*) - f(0)) = 0, \end{aligned}$$

находим, что

$$\frac{f(x^g) - f(0)}{f(x^*) - f(0)} \geq 1 - \left(1 - \frac{1}{1 + (1-c)h}\right)^r = 1 - B.$$

Учитывая обозначения леммы 3, получаем требуемую оценку

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

Лемма 3 доказана.

Доказательство теоремы 1. Пусть ε и $\varepsilon(\delta)$ — гарантированные оценки для задач A и A^δ соответственно. Тогда из леммы 3 следует, что

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

Поэтому для доказательства теоремы достаточно показать, что

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

Возмущение крутизны функции $f(x)$ обозначим через δ . Тогда из цепочки неравенств

$$\begin{aligned} 1 + (1 - c(f) - \delta)h &\leq 1 + (1 - c(f))h, \\ \frac{1}{1 + (1 - c(f) - \delta)h} &\geq \frac{1}{1 + (1 - c(f))h}, \\ B(c(f) + \delta, h, r) &= \left(1 - \frac{1}{1 + (1 - c(f) - \delta)h}\right)^r \\ &\leq \left(1 - \frac{1}{1 + (1 - c(f))h}\right)^r = B(c(f), h, r) \end{aligned}$$

получаем, что

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

Теорема 1 доказана.

5. Следствия и примеры

Приведем ряд следствий, непосредственно вытекающих из леммы 3.

Из леммы 3 получаем следующие достаточные условия того, что значения целевой функции задачи A в глобальном и градиентном экстремумах совпадают.

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

Покажем, что существуют такие функции из класса $\mathfrak{R}_\rho(\mathbf{Z}_+^n)$, что $c = 1$. Действительно, положим

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

Тогда

$$\begin{aligned} \Delta_1 f(x) &= -4x_1 + 4, & \Delta_2 f(x) &= -2x_2 + 2, \\ \rho &= (4, 2), & \Delta_1 f(0, 0) &= 4, & \Delta_2 f(0, 0) &= 2. \end{aligned}$$

Поэтому $c = 1$.

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

$$f(x) = -2x_1^2 + 7x_1 - x_2^2 + 3x_2, \quad P = \{(0, 0), (0, 1), (0, 2), (1, 0)\}.$$

Тогда

$$\begin{aligned} \Delta_1 f(x) &= -4x_1 + 5, & \Delta_2 f(x) &= -2x_2 + 2, & \rho &= (4, 2), \\ \Delta_1 f(0, 0) &= 5, & \Delta_2 f(0, 0) &= 2, & x^* = x^g &= (1, 0), & f(x^*) = f(x^g) &= 5. \end{aligned}$$

С другой стороны,

$$c(f) = \min\{\rho_1/\Delta_1 f(0), \rho_2/\Delta_2 f(0)\} = \min\{4/5, 2/2\} = 4/5 \neq 1.$$

Известно (см. теорему 11.1 в [11]), что если в задаче A $f(x) \in \mathfrak{R}_0(\mathbf{Z}_+^n)$, то справедлива оценка

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(0)} \leq \left(1 - \frac{1}{h}\right)^r = B_1. \quad (10)$$

Следующее следствие показывает, что в некоторых случаях оценка из леммы 3 лучше, чем оценка (10).

Следствие 2. Справедливы соотношения

$$\begin{aligned} B(c, h, r) &> B_1, & \text{если} & \quad c = 0, \\ B(c, h, r) &= 0, & \text{если} & \quad c = 1, \\ B(c, h, r) &< B_1, & \text{если} & \quad 1/h < c < 1, \quad h > 1, \end{aligned}$$

где $c = c(f)$.

Доказательство. Первое и второе соотношения очевидны. Докажем третье соотношение. Так как $h > 1$, $1/h < c < 1$, справедливо неравенство $(1 - c)h < h - 1$. Поэтому $1 + (1 - c)h < h$. Отсюда находим, что

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

то есть

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

Следствие 2 доказано.

Построим простой, но очень наглядный пример, когда градиентный алгоритм устойчив при возмущениях крутизны целевой функции.

Пример 1. Рассмотрим задачу: найти

$$\max\{f(x) = -x_1^2 + 6x_1 - x_2^2 + 5x_2 \mid x = (x_1, x_2) \in P\}, \quad (11)$$

где $P = \{(0, 0), (1, 0), (0, 1), (0, 2)\}$.

Очевидно, что P — порядково-выпуклое множество, $h = h(P) = 2$, $r = r(P) = 1$. Легко видеть, что

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

то есть $f(x) \in \mathfrak{R}_{(2,2)}(\mathbf{Z}_+^n)$ и $f(x)$ — неубывающая функция на множестве P . Тогда

$$c(f) = \min\{2/5, 2/4\} = 2/5.$$

Возмущение задачи (11) построим по правилу из свойства, приведенного в разделе 2. Рассмотрим задачу

$$\max f^\delta(x) = -x_1^2 + 6x_1 - x_2^2 + 5x_2 - (x_1 + x_2/2) \mid x = (x_1, x_2) \in P\}. \quad (12)$$

Очевидно, что $f^\delta(x) \in \mathfrak{R}_{(2,2)}(\mathbf{Z}_+^n)$ и $f^\delta(x)$ — неубывающая функция на множестве P , $I^+(f^\delta) \neq \emptyset$. Тогда

$$c(f^\delta) = \min\{2/4, 4/7\} = 1/2.$$

Поскольку $c(f^\delta) = c(f) + \delta$ и $\delta = 0,1$, в силу теоремы градиентный алгоритм устойчив в задаче (11) при возмущении крутизны целевой функции $f(x)$. Действительно, пусть ε_1 (ε_2) — гарантированная оценка задачи (11) (задачи (12)). Из леммы 3 находим, что $\varepsilon_1 = 6/11$, $\varepsilon_2 = 1/2$. Тогда $\varepsilon_2 < \varepsilon_1$.

Список литературы

1. Гордеев Э. Н., Леонтьев В. К., Общий подход к исследованию устойчивости решений в задачах дискретной оптимизации // Журнал вычислительной математики и математической физики (1996) **36**, №1, 66–72.

2. Емеличев В. А., Кричко В. Н., Формула радиуса устойчивости векторной l_∞ -экстремальной траекторной задачи. *Дискретная математика* (2004) **16**, №1, 14–20.
3. Сергиенко И. В., Козерацкая Л. Н., Лебедева Т. Т., *Исследование устойчивости и параметрический анализ дискретных оптимизационных задач*. Наукова думка, Киев, 1995.
4. Девятерикова М. В., Колоколов А. А., Анализ устойчивости некоторых алгоритмов дискретной оптимизации. *Автоматика и телемеханика* (2004), № 3, 48–54.
5. Тихонов А. Н., Арсенин В. Я., *Методы решения некорректных задач*. Наука, Москва, 1986.
6. Рамазанов А. Б., Об оценке точности градиентного алгоритма в одной задаче выпуклой дискретной оптимизации и некоторые смежные вопросы. В сб.: *Труды 13-й Байкальской Международной семинар-школы*. Изд-во Института систем энергетики им. Л. А. Мелентьева СО РАН, Иркутск, 2005, с. 571–576.
7. Рамазанов А. Б., Об оценке кривизны порядково-выпуклого множества на целочисленной решетке и некоторые смежные вопросы. *Математические заметки* (2008) **84**, №1, 153–156.
8. Conforti M., Cornuejols G., Submodular set functions, matroids and the greedy algorithm: Tight worst-case bounds and some generalizations of the Rado–Edmonds theorem. *Discrete Appl. Math.* (1984) **7**, 251–274.
9. Глебов Н. И., О применимости метода покоординатного спуска к некоторым задачам выпуклого целочисленного программирования. *Управляемые системы* (1978) **17**, 52–59.
10. Емеличев В. А., Ковалев М. М., Рамазанов А. Б., Погрешность градиентных экстремумов строго выпуклой функции дискретного аргумента. *Дискретная математика* (1990) **2**, №2, 127–137.
11. Ковалев М. М., *Матроиды в дискретной оптимизации*. Изд-во Университетское, Минск, 1987.
12. Рамазанов А. Б., Оценки точности получаемых алгоритмом покоординатного подъема решений задач дискретной выпуклой оптимизации. *Дискретный анализ и исследование операций* (2005) **12**, №4, 60–80.

Статья поступила 9.07.2009.