PROCEEDINGS OF IAM, V.7, N.2, 2018, pp.227-233

ТОЧНОСТЬ ГРАДИЕНТНОГО АЛГОРИТМА В ЗАДАЧАХ С (ρ, q) - КООРДИНАТНО-ВЫПУКЛЫМИ ЦЕЛЕВЫМИ ФУНКЦИЯМИ*

А.Б. Рамазанов [1,2]

1 Факультет Прикладной математики и кибернетики при БГУ, Баку, Азербайджан
2 Институт прикладной математики при БГУ, Баку, Азербайджан

e-mail: ram-bsu@mail.ru

Резюме. В настоящей работе градиентный алгоритм применяется для задач с (ρ, q)-координатно-выпуклыми целевыми функциями. Получено новое гарантированная оценка для рассмотренных задач. Найденная оценка обобщает и развивает ранее полученные оценки.

Ключевые слова: градиент, выпуклость, алгоритм, оценка, решетка.

AMC Subject Classification: 74P10.

Введение.

Известна роль координатно-выпуклых функций в задачах выпуклой дискретной оптимизации (см., напр., [1, 2, 4-8]). С помощью этих понятий построена достаточно универсальная методика оценки качества точности градиентного алгоритма [2, 4-8]. В силу потребности развития теории выпуклых дискретных экстремальных задач, а также с целью получить более точные оценки точности градиентного алгоритма выводятся новые классы выпуклых функции дискретного аргумента [2, 5-8].

В настоящем работе введены новые классы координатно-выпуклых функции дискретного аргумента на координатных решетках. Исследованы свойства этих функций. Исходя из определения этого класса, предложен градиентный алгоритм покоординатного подъема. Найдена гарантированная оценка точности предложенного градиентного алгоритма.

Обозначения и основной результат. Пусть $H = (H, \prec)$ - линейно упорядоченное дискретное множество (цепь), на котором задано отношение порядка $\prec$. 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. Длина цепи между x и y обозначается через $h(x, y)$. Минимальный элемент множества H будем обозначать через 0.

---

* Работа была представлена на семинаре Института Прикладной Математики 06.03.2018

Будем также пользоваться обозначением $h(x) = h(0, x)$. Тогда на координатной решетке

$$H^n = \underbrace{H \times H \times ... \times H}_{n} = \{x = (x_1, ..., x_n) : x_i \in H, 1 \le i \le n\}$$

задается отношение порядка $x \prec y \Leftrightarrow x_i \prec y_i, 1 \le i \le n.$ Очевидно, что минимальным элементом множества H^n является n-мерный вектор $0 = (0,...,0)$. Для всякого элемента $x_i \in H$ через x_i^+ будем обозначать такой элемент из H, что в H нет такого элемента x, что $x_i \prec x \prec x_i^+$. Пусть $P \subseteq H^n$. Будем в дальнейшем считать, что множество P обладает свойствами:

1) $|P| < +\infty$;
2) $0 \in P$;
3) $[0, x] = \{z \in H^n : 0 \prec z \prec x\} \subseteq P$ для любого $x \in P$.

Множество обладающие свойством 1)-3) называют порядково-выпуклой [2, 8]. Введем следующие обозначения:

$$N(x, y) = \{i : x = (x_1, ..., x_n) \prec (y_1, ..., y_n) = y, x_i \prec y_i, x_i \ne y_i, 1 \le i \le n\},$$

$$h(x, y) = \sum_{i \in N(x,y)} h(x_i, y_i), h(x_i, y_i) = |\{z_i : x_i \prec z_i \prec y_i\}| - 1, 1 \le i \le n,$$

$$h(x) = h(0, x), h = h(P) = \max\{h(0, x) : x \in P\}, r = \min\{h(x) - 1 : x \in H^n \setminus P\},$$

$$fes(x, P) = \{1 \le i \le n : \pi_i^+(x) \in P, x \in P\},$$

$$\pi_i^+(x) = (x_1, ..., x_{i-1}, x_i^+, x_{i+1}, ..., x_n), I_n = \{1,2,...,n\}, h(x_i, x_i^+) = 1$$

Следуя [2, 5, 8], для функции $f : H^n \to R$ (R - множество действительных чисел) введем понятия i - градиента

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

и (i, j) – градиента

$$\Delta_{ij} f(x) = \Delta_j f(\pi_i^+(x)) - \Delta_j f(x),$$

где $\pi_i^+(x) = (x_1, ..., x_{i-1}, x_i^+, x_{i+1}, ..., x_n)$.

Пусть $\rho = (\rho_1, ..., \rho_n), q = (q_1, ..., q_n) \in R_+^n$, где R_+^n - множество n-мерных неотрицательных действительных векторов. Через $\Re_{\rho,q}(H^n)$ - обозначим класс всех функций $f : H^n \to R$, обладающих свойством

$$\Delta_{ij} f(x) \le 0, \ \forall x \in H^n, \forall i, j \in I_n = \{1,2,...,n\}, i \ne j,$$

$$-q_i \le \Delta_{ii} f(x) \le -\rho_i, \forall x \in H^n, i = \overline{1,n}.$$

То есть числа $-q_1,...,-q_n$ и $-\rho_1,...,-\rho_n$ являются соответственно нижними и верхними границами для диагональных элементов матрицы $\left\|\Delta_{ij} f(x)\right\|$, представляющей собой дискретный аналог гессиана.

Отметим, что в непрерывном случае аналогичные функции подробно исследованы и названы (l,Q)-вогнутый (см., например, [3]).

Следующая теорема позволяет дать несколько эквивалентных определений класса $\Re_{\rho,q}(H^n)$.

Теорема 1. Следующие утверждения эквивалентны:

1). $f(x) \in \Re_{\rho,q}(H^n)$;

2). $$\sum_{i\in N(x,y)} h(x_i,y_i)\Delta_i f(x) + \sum_{i\in N(y,x)} h(y_i,x_i)\Delta_i^- f(x\vee y) + \sum_{i\in N(x,y)} q_i h(x_i,y_i)h(x_i) -$$

$$\sum_{i\in N(y,x)} q_i h(y_i,x_i)h((\pi_i^-(x\vee y)_i) + \frac{1}{2}\sum_{i\in N(x,y)} q_i h(x_i,y_i) - \frac{1}{2}\sum_{i\in N(y,x)} q_i h(y_i,x_i) -$$

$$\frac{1}{2}\sum_{i=1}^{n} q_i(h^2(y_i) - h^2(x_i)) \le f(y) - f(x) \le \sum_{i\in N(x,y)} h(x_i,y_i)\Delta_i^+ f(x) + \sum_{i\in N(y,x)} h(y_i,x_i)\Delta_i^- f(x\vee y) +$$

$$\sum_{i\in N(x,y)} \rho_i h(x_i,y_i)h(x_i) - \sum_{i\in N(y,x)} \rho_i h(y_i,x_i)h((\pi_i^-(x\vee y))_i) + \frac{1}{2}\sum_{i\in N(x,y)} \rho_i h(x_i,y_i) -$$

$$\frac{1}{2}\sum_{i\in N(y,x)} \rho_i h(y_i,x_i) - \frac{1}{2}\sum_{i=1}^{n} \rho_i(h^2(y_i) - h^2(x_i)),\ \forall x,y \in H^n;$$

3). $$\sum_{i\in N(x,y)} h(x_i,y_i)\Delta_i f(x) + \frac{1}{2}\sum_{i\in N(x,y)} q_i\ {}_i\ h(x_i,y_i) -$$

$$\sum_{i\in N(y,x)} q_i h(y_i,x_i)h((\pi_i^-(x\vee y)_i) + \frac{1}{2}\sum_{i\in N(x,y)} q_i h(x_i,y_i) - \frac{1}{2}\sum_{i\in N(x,y)} q_i h^2(x_i,y_i) \le f(y) - f(x) \le$$

$$\sum_{i\in N(x,y)} h(x_i,y_i)\Delta_i f(x) + \frac{1}{2}\sum_{i\in N(x,y)} \rho_i h(x_i,y_i) - \frac{1}{2}\sum_{i=1}^{n} \rho_i h^2(x_i,y_i),\ \forall x,y \in H^n,\ x \prec y.$$

Здесь $x\vee y\,(x\wedge y)$ - объединение (сечение) элементов x и y, т.е. $x\vee y = \sup\{x,y\}, x\wedge y = \inf\{x,y\}$, $\pi_i^-(x) = (x_1,...,x_{i-1},x_i^-,x_{i+1},...,x_n), x_i^- \prec x_i$.

В случае, когда $H^n = Z_+^n$, $x\vee y$ $(x\wedge y)$ имеет координаты $\max(x_1,y_1),...,\max(x_n,y_n)$ $(\min(x_1,y_1),...,\min(x_n,y_n))$.

Доказательство. В силу теоремы 1 [8], имеем

$$f(x) = g(x) + \frac{1}{2}\sum_{i=1}^{n} \rho_i h_i^2(x_i),$$

$$f(x) = \varphi(x) + \frac{1}{2}\sum_{i=1}^{n} q_i h_i^2(x_i),$$

где $g(x) \in \Re_0(H^n)$, $-\varphi(x) \in \Re_0(H^n)$ [2, 8].

Далее следует применить теоремы 4 [8] для функции $g(x), -\varphi(x)$.

Теорема 1 доказана.

Теорема 2. Если $f(x) \in \Re_{\rho,q}(H^n)$ - неубывающая функция, то справедливы неравенства

$$\Delta_i f(\pi_i^+(x)) \le \Delta_i f(0) - (h(0,x)+1)\rho_i, \ \forall i \in fes(\pi_i^+(x), P), \forall x \in P,$$

$$\Delta_i f(x) \le \Delta_i f(0) - (h(0,x)+1)\rho_i + q_i, \ \forall i \in fes(x,P), \ \forall x \in P$$

Доказательство теоремы 2 вытекает из определения класса функции $\Re_{\rho,q}(H^n)$.

Рассмотрим задачу А выпуклой дискретной оптимизации:

Задача А:

$$\max\{ f(x) : x \in P \subseteq H^n \},$$

где $f(x) \in \Re_{\rho,q}(H^n), P -$ порядково-выпуклое множество.

Для решения задачи А предложен следующий алгоритм покоординатного подъема.

Алгоритм G_4 .

1. Полагаем $x^0 = 0 = (0,...,0),\ t = 0$.

2. Если $fes(x^t, P) = \varnothing$, то конец. Иначе переходим п.3.

3. Полагаем

$$x^{t+1} = \pi_{i(t)}^+(x^t),$$

$$i(t) = \arg\max_i \left\{ \frac{q_i}{\rho_i} \Delta_i f(x^t) : i \in fes(x^t, P) \right\}.$$

4. Полагаем $t \leftarrow t+1$ и переходим к п.2.

Пусть число шагов алгоритма G_4 равно k. Тогда полученное решение обозначим через - $x^{G_4} = (x_1^{G_4},...,x_n^{G_4}) = x^k = (x_1^k,...,x_n^k)$. Это решение будем называть градиентным решением (или градиентным максимумом функции $f(x)$) задачи А3 .

Для оценки качества точности градиентного алгоритма G_4 используем методику из [1-5].

Теорема 3. Если в задаче А функция $f(x)$ неубывающая, то для гарантированной оценки погрешности градиентного алгоритма G_4 справедливо неравенство

$$\frac{f(x^*)-f(x^{G_4})}{f(x^*)-f(0)} \le V(h,k,q,\rho)\,, \qquad (1)$$

где

$$V(h,k,\rho,q) = B_4(h,k,\rho,q) - \frac{(h-k)^2}{2h}(V_1(h,\lambda_f,\rho))^{-1} - 1,$$

$$B_4(h,k,\rho,q) = \left(1-\prod_{s=0}^{k-1}\left(1-\frac{1}{\eta_s}\right)\right)^{-1},\ \eta_s = \frac{hq_{i(s)}\Omega(\overline{q})}{\rho_{i(s)}},$$

$$\overline{q} = (\overline{q}_1,...,\overline{q}_n),\ \overline{q_i} = \frac{\rho_i}{q_i},\ i=\overline{1,n},$$

$$V_1(h,\lambda_f,\rho) = h\Omega(\lambda_f) - \frac{1}{2}h\omega(\rho)\,,$$

$$\lambda_f = (\lambda_1^f,...,\lambda_n^f),\ \lambda_i^f = (\Delta_i f(0)),\ i=1,...,n.$$

Доказательство. Полагая в п.3 теоремы 1 при $y=x^*$, $x=x^t$, получаем

$$f(x^*)-f(x^t) \le \sum_{i\in N(x^t,x^*)} h(x_i^t,x_i^*)\Delta_i f(x^t) - \frac{1}{2}\sum_{i=1}^{n}\rho_i h(x_i^t,x_i^*)(h(x_i^t,x_i^*)-1), \quad t=0,1,...,k.$$

Отсюда, учитывая неравенства

$$\Delta_i f(x^t) \le \frac{q_{i(t)}}{\rho_{i(t)}}\frac{\rho_i}{q_i}\Delta_{i(t)} f(x^t)\,,\ t=\overline{0,k}, \quad i\in fes(x^t,P),$$

имеем

$$f(x^*)-f(0)+\frac{1}{2}\frac{(h-k)^2}{h} \le \sum_{s=0}^{t-1}\Delta_s + \frac{q_{i(t)}\Delta_t}{\rho_{i(t)}}\sum_{i=1}^{n}\frac{\rho_i}{q_i},\ t=\overline{0,k}\,.$$

Отсюда повторяя процедуры аналогичные как в [1-5], выводим

$$f(x^*)-f(x^{G_4}) \le (B_4(h,k,\rho,q)-1)(f(x^*)-f(0)) - \frac{(h-k)^2}{2h}, \qquad (2)$$

где

$$B_4(h,k,\rho,q) = \left(1-\prod_{s=0}^{k-1}\left(1-\frac{1}{\eta_s}\right)\right)^{-1},\ \eta_s = \frac{hq_{i(s)}\Omega(\overline{q})}{\rho_{i(s)}},\ \overline{q} = (\overline{q}_1,...,\overline{q}_n),$$

$$\overline{q_i} = \frac{\rho_i}{q_i},\ i=\overline{1,n},\ \ s\in\{0,1,...,k-1\}$$

Из правой части неравенства п.3 теоремы 1 при $y=x^*$, $x=0$, имеем

$$f(x^*)-f(0)\le \sum_{i\in N(0,x^*)} h(0,x_i^*)\Delta_i f(0)-\frac{1}{2}\sum_{i=1}^{n}\rho_i h(0,x_i^*)(h(0,x_i^*)-1).$$

Отсюда с учетом леммы 2 [1], имеем

$$f(x^*)-f(0)\le \sum_{i\in N(0,x^*)} h(0,x_i^*)\Delta_i f(0)-\frac{1}{2}h\omega(\rho)\le$$
$$h\Omega(\lambda_f)-\frac{1}{2}h\omega(\rho)=V_1(h,\lambda_f,\rho)\,, \tag{3}$$

где

$$\lambda_f=(\lambda_1^f,...,\lambda_n^f),\ \lambda_i^f=\Delta_i f(0), i=\overline{1,n}.$$

Учитывая (2) в (3) с учетом обозначений, выводим оценку (1) из теоремы 3.

Теорема 3 доказана.

Литература

1. Emelichev V.A., Ramazanov A.B. About the steepness of the function of discrete argument // TWMS Journal of Pure and Applied Mathematics, 2016, V. 7, N 1, pp. 105-111.
2. Ковалев М.М. Матроиды в дискретной оптимизации. Изд-во Университетское, Минск, 1987, 222 с.
3. Немировский А.С., Юдин Д.Б. Сложность задач и эффективность методов оптимизации. М.: Наука, 1979, 384 с.
4. Рамазанов А.Б. Об оценке градиентного экстремума с помощью параметризации градиентного алгоритма // Proceedings of IAM , 2015, V. 4, N 2, pp. 214-220.
5. Рамазанов А.Б. Устойчивость градиентного алгоритма в задачах выпуклой дискретной оптимизации и некоторые смежные вопросы // Дискретная математика, 2011, №3, с. 82-92.
6. Ramazanov A.B. New of Accuracy of Gradient Algorithm in the Jordan-Dedekinds Structure // Applied and Computational Mathematics, 2018, V. 17, N 1, pp. 109-113.
7. Ramazanov A. B. An Estimate for the Curvature of an Order-Convex Set in the Integer Lattice and Related Questions // Math. Notes, 2008, vol 84, N 1, pp. 147-151.
8. Рамазанов А. Б. Оценки точности получаемых алгоритмом покоординатного подъема решений задач дискретной выпуклой оптимизации , Дискретный анализ и исследование операций, 2005, серия 1, том 12, № 4, с. 60-80.

ACCURACY OF THE GRADIENT ALGORITHM IN TASKS WITH

(ρ, q) - COORDINATE-CONVEXITY CRITERION FUNCTIONS

A.B. RAMAZANOV

ABSTRACT

In the work the gradient algorithm is applied to tasks with (ρ, q) -coordinate and convex criterion functions. It is received new the guaranteed assessment for the considered tasks. The assessment found summarizes and develops previously obtained estimating.

Keywords: gradient, convexity, algorithm, errors, lattice.

References

1. Emelichev V.A., Ramazanov A.B. About the steepness of the function of discrete argument , TWMS Journal of Pure and Applied Mathematics, 2016, V. 7, N 1, pp. 105-111.
2. Kovaleb M.M. Matroidi v diskretnoy optimizasii, Izd-vo Universitetskoye , Minsk, 1987, 222 s.
3. Nemirovskiy A.S.,Yudin D.B. Slojnost zadach I effektivnost metodov optimizasii , M.: Nauka, 1979, 384 s.
4. Ramazanov A.B. Ob osenki gradiyentnoqo ekstremuma s pomoshyu parametrizasiyi gradiyentnoqo algoritma , Proceedings of IAM , 2015, V. 4, N 2, pp. 214-220.
5. Ramazanov A.B. Ustoychivost gradiyentnoqo algoritma v zadachax vipukloy diskretnoy optimizasii i nekotoriyi smejniyi voprosi , Diskretnaya matematika , 2011, №3, s. 82-92.
6. Ramazanov A.B. New of Accuracy of Gradient Algorithm in the Jordan-Dedekinds Structure, Applied and Computational Mathematics, 2018, V. 17, N 1, pp. 109-113.
7. Ramazanov A. B. An Estimate for the Curvature of an Order-Convex Set in the Integer Lattice and Related Questions, Math. Notes, 2008, vol 84, N 1, pp. 147-151.
8. Ramazanov A.B. Osenki tochnosti poluchaemix algoritmom pokoordinatnoqo podyema resheniy zadach diskretnoy vipukloy optimizasi , Diskretniy analiz I issledovaniye operasii, 2005, S. 1, T. 12, № 4, s. 60-80.(Ramazanov, A. B., Estimates of the Accuracy Obtained by the Algorithm of Coordinate Recovery of Discrete Convex Optimization Problems, Discrete Analysis and Operations Research, 2005, S. 1, V. 12, N. 4, p. 60-80) (in Russian)