Proceedings of IAM, V.9, N.2, 2020, pp.126-131

ПОГРЕШНОСТЬ ГРАДИЕНТНОГО АЛГОРИТМА НА ОСНОВЕ РАСШИРЕНИЕ ОКРЕСТНОСТИ ПОИСКА РЕШЕНИЙ ДЛЯ ВЫПУКЛОЙ ФУНКЦИИ ДИСКРЕТНОГО АРГУМЕНТА

А.Б.Рамазанов[1,2], С.Ф.Ахундов[1], Р.Т.Зульфугарова[3]

[1] Факультет прикладной математики и кибернетики,при БГУ, Баку. Азербайджан
[2] Институт Прикладной математики при БГУ, Баку, Азербайджан
[3] Факультет международных отношений и экономики , при БГУ, Баку. Азербайджан
e-mail: ram-bsu@mail.ru

Резюме. В настоящем работе предложена одна из модификаций градиентного алгоритма покоординатного подъема, основанная на выборе направления с помощью разности двух градиентов, вычисленных в двух последовательных элементах (другими словами, расширяется окрестность поиска градиентного решения).

Ключевые слова: градиент, выпуклый, алгоритм, точность, дискрет.

AMC Subject Classification: 74P10

1. Введение.

Одним из возможных вариантов улучшения качества точности градиентных алгоритмов в задачах выпуклой дискретной оптимизации является различные модификации алгоритма покоординатного подъема (см., например, [1, 2, 4, 6, 7]. С помощью различных способов выбора направления получены оценки точности градиентного алгоритма или доказана оптимальность градиентного решения в таких известных задачах, как задача коммивояжера, задача о ранце, задачах распределения ресурсов и др. (см., напр., [2, 6]).

В настоящем работе предложена одна из модификаций градиентного алгоритма покоординатного подъема, основанная на выборе направления с помощью разности двух градиентов, вычисленных в двух последовательных элементах (другими словами, расширяется окрестность поиска градиентного решения). При этом предложенная модификация градиентного алгоритма позволяет найти новые, априорные и апостериорные оценки точности и упростить процедуры нахождения гарантированных оценок точности для выпуклой функции дискретного аргумента по сравнению с ранее предложенной методикой нахождения таких оценок (ср. с [1, 2, 6, 7]). Также сформулированы условия совпадения глобального и градиентного экстремума. Отметим, что хотя основные результаты излагаются на координатной решетке, но учитывая, что целочисленные и координатные решетки изоморфны, то все результаты полученные на координатной решетке остаются справедливыми и на целочисленной решетке и обратно.

2. Определение обозначения.

Пусть $H = (H, \prec)$ - линейно упорядоченное дискретное множество (цепь), на котором задано отношение порядка $\prec$ и

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

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 $x \in P$.

Следуя [2,6,7], множество *P*, обладающее свойствами 1)-3), будем называть конечным порядково-выпуклым множеством с нулем.

Введем следующие обозначения:

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

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

$$h(x) = h(0, x), h = h(P) = \max\{h(0, x) : x \in P\},$$

$$fes(x, P) = \{1 \le i \le n : \pi_i^+(x) \in P,\ x \in P\},\ \pi_i^+(x) = (x_1, \ldots, x_{i-1}, x_i^+, x_{i+1}, \ldots, x_n),\ h(x_i, x_i^+) = 1$$

Следуя [2,6,7], 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

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

Пусть $\rho = (\rho_1, \ldots, \rho_n) \in R_+^n$ (R_+^n - множество n- мерных действитель-ных векторов), $\Re_\rho(H^n)$ - класс ρ-координатно-выпуклых функций на H^n [см., напр., 2,7], т.е. таких функций $F : H^n \to R$, что для любого $x \in H^n$

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

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

3. Постановка задачи. Рассмотрим задачу A выпуклой дискретной оптимизации: найти

$$\max\{F(x) : x \in P\},$$

где $F(x)$ - $\rho-$координатно-выпуклая функция.

Пусть $x^* = (x_1^*,\dots,x_n^*)$ - оптимальное, а $x^g = (x_1^g,\dots,x_n^g)$ - градиентное решение (т.е. точка, полученная с помощью градиентного алгоритма покоординатного подъема) задачи A.

Под гарантированной оценкой погрешности градиентного алгоритма решения задачи A, как обычно [1-7], понимают такое число $\varepsilon \geq 0$, что

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

Естественно, чем меньше величина ε, тем лучше погрешность. Поэтому, для уменьшение величину ε применяют различные варианты градиентного алгоритма. Ниже описываемая алгоритм основана с расширением поиска окрестности нахождение градиентного решения. Градиентным решением (максимумом) x^g задачи А (функции $f(x) \in \Re_\rho(H^n)$) на множестве $P \in H^n$ назовем точку, построенную с помощью следующего градиентного алгоритма G покоординатного подъема.

Алгоритм G.

1. Принимаем $x^0 = 0, t = 0$. Если $fes(x^t, P) = \varnothing$, то конец. Иначе полагаем

$$x^{t+1} = \pi^+_{i(t)}(x^t), \quad i(t) = \arg\max_i \{\Delta_i f(x^t) : i \in fes(x^t, P)\}$$

и переходим к пункту 2.

2. Полагаем $t \leftarrow t+1$ и находим

$$i(t) = \arg\max_i \{\Delta_i f(x^t) - \Delta_i f(x^{t-1}) : i \in fes(x^t, P)\},$$

$$x^{t+1} = \pi^+_{i(t)}(x^t)$$

Если $fes(x^t, P) = \varnothing$, то конец. Иначе полагаем $t \leftarrow t+1$ и повторяем пункт 2.

Пусть k - число шагов алгоритма G. Тогда полученное решение $x^k = (x_1^k,\dots,x_n^k)$ обозначим через - $x^g = (x_1^g,\dots,x_n^g)$ и будем называть градиентным максимумом функции $f(x)$ на множестве P.

Как обычно, функцию $f : H^n \to R$ будем называть неубывающей на множестве $P \subseteq H^n$, если $\Delta_i f(x) \leq 0, \ \forall i \in fes(x, P), \ x \in P.$

Пусть x^* - глобальный максимум (оптимальное решение задачи А) функции $f(x) \in \Re_\rho(H^n)$ на множестве P.

Теорема. Пусть $f(x) \in \Re_\rho(H^n)$ неубывающая функция на множестве P и $\gamma(k,\rho) > 0$. Тогда глобальный максимум x^* и градиентный максимум x^g функции $f(x)$ на множестве P связаны соотношением

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(0)} \le \frac{(h-k)L}{\gamma(k,\rho)} - \frac{(h-k)^2\omega(\rho)}{2h\gamma(k,\rho)} - \frac{(h-k)Q(k,\rho)}{\gamma(k,\rho)},$$

где

$$L = \max_i \{\Delta_i f(\theta) : i \in fes(\theta, P)\},\ \gamma(k,\rho) = \sum_{s=0}^{k-1} \rho_{i(s)},$$

$$Q(k,\rho) = \sum_{j=1}^{k} e_{i(j)i(j-1)} \rho_{i(j-1)},\ \ e_{i(j)i(j-1)} = \begin{cases} 0, & если \quad i(j) \ne i(j-1), \\ 1, & если \quad i(j) = i(j-1), \end{cases}$$

$$\omega(\rho) = \left(\sum_{i \in N_\rho^+} \frac{1}{\rho_i} \right)^{-1},\ N_\rho^+ = \{i : \rho_i > 0, i = \overline{1,n}\},\ h = \max\{h(\theta, x) : x \in P\},$$

множество индексов $\{i(0), i(1), \ldots, i(k)\}$ определяется в алгоритме G, k - число шагов алгоритма G.

Доказательство теоремы вытекает из п.3 теоремы 2 [1] при $x = x^g, y = x^*$ и из описания алгоритма G с учетом соотношений

$$\Delta_i f(x^g) \le \Delta_i f(x^0) - \sum_{j=1}^{k} e_{i(j)i(j-1)} \rho_{i(j-1)} = \Delta_i f(\theta) - Q(k,\rho),$$

$$f(x^*) - f(\theta) \ge \sum_{s=0}^{k-1} \rho_{i(s)} = \gamma(k,\rho).$$

Следствие. В условиях теоремы 1, если $h = k$, то $f(x^*) = f(x^g)$.

Если $h > k$, то качество точности решения, построенного с алгоритмом G, зависит от поведения величины

$$(h/k - 1)\frac{L}{\gamma(k,\rho)}$$

в наихудшем случае.

В частности, если $f(x)$ строго неубывающая функция $(x \prec y, x \ne y \Rightarrow f(x) < f(y)),\ \ k = h$, то $x^* = x^g$.

Если, кроме того, H^n и функция $f(x)$ имеет вид:

$$f(x)=\sum_{i=1}^{n}c_i h(0,x_i)-\frac{1}{2}\sum_{i=1}^{n}\rho_i h(0,x_i^2),$$

где

$$c=(c_1,...,c_n),\rho=(\rho_1,...,\rho_n)\in R_+^n,\quad c_i-\rho_i x_i-\rho_i/2\geq 0,\forall i\in fes(x,P),\forall x\in P,$$

то с учетом, что

$$L=\max_i\{c_i-\rho_i/2:i\in I_n\},$$

оценка из теоремы может быть конкретизирована.

Следствие дополняет результаты о совпадении глобального и градиентного экстремума, ранее полученного для сепарабельной координатно-выпуклой функции (см., напр., [2,6,7]).

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. Ramazanov A.B. New of Accuracy of Gradient Algorithm in the Jordan-Dedekinds Structure // Applied and Computational Mathematics, 2018, V. 17, N 1, pp. 109-103.
3. Ramazanov A.B. On stability of the gradient algorithm in convex discrete optimisation problems and related questions // Discrete Mathematics and Applications, 2011, volume 21, Issue 4, Pages 465-476.
4. Srivastova V.K., Fahim A. A two – phase optimization for integer programming problems // Comput. and Math. Appl., 2001, 42, N 12, pp. 1585-1595.
5. Емеличев В.А., Овчинников В.Г. К теории экстремума на координатных решетках // Докл. АН БССР, 1983, т. 27, № 7, с.581-583.
6. Ковалев М.М. Матроиды в дискретной оптимизации. Изд-во Университетское, Минск, 1987, 222 с.
7. Рамазанов А.Б. Об оценке градиентного экстремума с помощъю параметризации градиентного алгоритма // Proceedings of IAM , 2015, V. 4, N 2, pp. 214-220.

ERROR OF GRADIENT ALGORITHM BASED ON EXPANSION OF NEIGHBORHOOD OF SOLUTION SEARCH FOR CONVEX FUNCTION OF DISCRETE ARGUMENT

A.B.Ramazanov, S.F.Akhundov, R.T.Zulfugarova

ABSTRACT

In the present work, one of the modifications of the gradient algorithm of obordinate lifting is proposed, based on the choice of direction using the difference of two gradients calculated in two successive elements (in other words, the vicinity of the search for a gradient solution is expanded).

Keywords: gradient, convex, algorithm, errors, discret.

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. Ramazanov A.B. New of Accuracy of Gradient Algorithm in the Jordan-Dedekinds Structure // Applied and Computational Mathematics, 2018, V. 17, N 1, pp. 109-103.
3. Ramazanov A.B. On stability of the gradient algorithm in convex discrete optimisation problems and related questions // Discrete Mathematics and Applications, 2011, volume 21, Issue 4, Pages 465-476.
4. Srivastova V.K., Fahim A. A two – phase optimization for integer programming problems // Comput. and Math. Appl., 2001, 42, N 12, pp. 1585-1595.
5. Emelichev V.A., Ovchinnikov V.G. K teorii ekstremuma na koordinatnykh reshetkakh // Dokl. AN BSSR, 1983, t. 27, № 7, s.581-583 (Emelichev V.A., Ovchinnikov V.G. On the theory of an extremum on coordinate lattices // Dokl. AN BSSR in Russian) .
6. Kovalev M.M. Matroidy v diskretnoy optimizatsii. Izd-vo Universitetskoe, Minsk, 1987, 222 s.(Kovalev M.M. Matroids in discrete optimization. Universitetskoe Publishing House, Minsk, 1987, 222 p. in Russian)
7. Ramazanov A.B. Ob otsenke gradientnogo ekstremuma s pomoshch"yu parametrizatsii gradientnogo algoritma // Proceedings of IAM , 2015, V. 4, N 2, pp. 214-220 (Ramazanov A.B. Estimation of the gradient extremum using the parameterization of the gradient algorithm // Proceedings of IAM, 2015, V. 4, N 2, pp. 214-220 in Russian)