ОБ ОЦЕНКЕ ГРАДИЕНТНОГО ЭКСТРЕМУМА С ПОМОЩЬЮ ПАРАМЕТРИЗАЦИИ ГРАДИЕНТНОГО АЛГОРИТМА*

А.Б. Рамазанов[1]

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

Резюме. В настоящей работе рассматривается один из вариантов градиентного алгоритма параметризованного с помощью меры строгой выпуклости целевой функции на конечном подмножестве координатных решеток. При этом за счет введения параметра в градиентный алгоритм число шагов градиентного алгоритма существенно уменьшается.

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

AMS Subject Classification: 74P10.

1. Введение

Одним из возможных вариантов улучшения качества точности и уменьшения число шагов градиентного алгоритма в задачах выпуклой дискретной оптимизации является параметризация алгоритма покоординатного подъема (см., например,[3, 5] . С помощью различных способов параметризации градиентного алгоритма получены оценки точности градиентного экстремума или доказана оптимальность градиентного решения в таких известных задачах, как задача коммивояжера, задачах распределения ресурсов и др. (см., напр., [3]).

В настоящей работе рассматривается один из вариантов градиентного алгоритма параметризованного с помощью меры строгой выпуклости целевой функции на конечном подмножестве координатных решеток. При этом за счет введения параметра в градиентный алгоритм число шагов градиентного алгоритма существенно уменьшается. Отметим, что таких вариантов применения градиентного алгоритма в задачах дискретной оптимизации многочисленно (см., напр., [3, 5], там же ссылки). Если в упомянутых и в других работах существенно используются свойства допустимой области, то в настоящей работе основной априорной информацией является параметр строгой выпуклости, который до начала работы градиентного алгоритма априори известен.

Пусть Z_+^n (R_+^n) - множество $n-$мерных неотрицательных целочисленных (действительных) векторов. Для $x=(x_1,\dots,x_n), y=(y_1,\dots,y_n)\in Z_+^n$ будем

* Работа была представлена на семинаре Института Прикладной Математики 30.06.2015

писать $x=(x_1,...,x_n)\le y=(y_1,...,y_n)$, если $x_i \le y_i$, $i=1,2,...,n$. Если среди последних неравенств есть хотя бы одно строгое, то будем писать $x<y$. Для $x=(x_1,...,x_n), y=(y_1,...,y_n)\in Z_+^n$ будем писать $x=(x_1,...,x_n)\le y=(y_1,...,y_n)$, если $x_i \le y_i$, $i=1,2,...,n$. Если среди последних неравенств есть хотя бы одно строгое, то будем писать $x<y$. Пусть $P\subseteq Z_+^n$. 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 [1, 4-5], множество P, обладающее свойствами 1)-3), будем называть конечным порядково-выпуклым множеством с нулем.

Введем следующие обозначения:

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

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

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

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

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

где $e^i - i-$й единичный n-мерный орт.

Для функции $f: Z_+^n\to R$ (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 $\rho=(\rho_1,...,\rho_n)\in R_+^n$, $\Re_\rho(Z_+^n)$ - класс ρ-координатно-выпуклых функций на Z_+^n [3-5], т.е. таких функций $f: Z_+^n\to R$, что для любого $x\in Z_+^n$

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

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

Рассмотрим задачу A выпуклой дискретной оптимизации: найти

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

где $f(x)$ - $\rho-$координатно-выпуклая функция, P - порядково - выпуклое множество в Z_+^n.

Градиентным экстремумом (максимумом) x^g функции $f(x)\in\Re_\rho(Z_+^n)$ на множестве $P \subseteq Z_+^n$ назовем точку, построенную с помощью следующего градиентного алгоритма $G(\rho)$ покоординатного подъема.

Алгоритм $G(\rho)$.

1. $x^0 = 0 = (0,\ldots,0), \quad t = 0$,
2. $x^{t+1} = \pi_{i(t)}^{+}(x^t), \quad i(t) = \arg\max_i \{\Delta_i f(x^t) - \rho_i : i \in fes(x^t, P)\}$

2. Если $fes(x^t, P) = \varnothing$ или $\Delta_{i(t)} f(x^t) - \rho_{i(t)} \le 0$, то конец. Иначе полагаем $t \leftarrow t+1$ и повторяем пункт 1.

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

Пусть $x^* = (x_1^*,\ldots,x_n^*)$ - оптимальное решение задачи A, т.е.

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

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

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

Как обычно, функции $f : Z_+^n \to R$ будем называть не убывающей, если $\Delta_i f(x) \ge 0, \forall i \in fes(x, P), \forall x \in P$.

Далее нам понадобится следующая лемма.

Лемма 1. [4]. Если функция $f(x) \in \Re_\rho(Z_+^n)$ неубывающая, то справедливо неравенство

$$\sum_{i \in N(x^t, x^*)} \rho_i h(x_i^t, x_i^*)(h(x_i^t, x_i^*) - 1) \ge (h-t)^2 \omega(\rho)/h \ge (h-k)^2 \omega(\rho)/h ,\ t = 0,1,\ldots,k,$$

где

$$\omega(\rho) = \begin{cases} 0, \textit{если} \quad N_\rho^+ = \varnothing, \\ \left(\sum_{i \in N_\rho^+} \dfrac{1}{\rho_i} \right)^{-1}, \textit{если}\, N_\rho^+ \ne \varnothing, \end{cases}$$

$$N_\rho^+ = \{i : \rho_i > 0, i = \overline{1,n}\},\ h = \max\{h(0,x) : x \in P\}$$

Теорема 1. Пусть $f(x) \in \Re_\rho(Z_+^n)$ изотонная функция на множестве P. Тогда глобальный максимум x^* и градиентный экстремум x^g функции $f(x)$ на множестве P связаны соотношением

$$f(x^*) \le A(k,h) f(x^g) + (1 - A(k,h)) f(\theta) - B(k,h,\rho), \qquad (1)$$

Если, кроме того справедливы неравенства $\omega_1(k,h,\delta_f) > 0, \quad \gamma(k,\rho) > 0$, то справедлива следующая гарантированная оценка точности градиентного алгоритма $G(\rho)$ в задаче максимизации функции $f(x)$ на множестве P

$$\frac{f(x^*)-f(x^g)}{f(x^*)-f(\theta)} \le B_1(k,h,\rho,\delta_f)\ , \tag{2}$$

где

$$A(k,h)=(1-(1-1/h)^k)^{-1},$$

$$B(k,h,\rho)=A(k,h)\sum_{s=0}^{k-1}(1-1/h)^{k-1-s}\rho_{i(s)}+$$

$$(h-k)^2\omega(\rho)/2h-h\Omega(\rho),\Omega(\rho)=\sum_{i=1}^{n}\rho_i,$$

$$B_1(k,h,\rho,\delta_f)=(1-1/h)^k-\sum_{s=0}^{k-1}((1-1/h)^{k-1-s}\rho_{i(s)})/\omega_1(\rho,h,\delta_f)-$$

$$(h-k)^3\omega(\rho)/2h\omega_1(\rho,h,\delta_f)+$$

$$h\Omega(\rho)/\gamma(k,\rho),\gamma(k,\rho)=\sum_{s=0}^{k-1}\rho_{i(s)},\omega_1(\rho,h,\delta_f)=h\Omega(\delta_f)-h\omega(\rho)/2,$$

$$\delta_f=(\delta_f^1,...,\delta_f^n),\delta_f^i=\Delta_i f(\theta),\ i=1,2,...,n,\Omega(\delta_f)=\sum_{i=1}^{n}\delta_f^i$$

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

Замечание 1. Отметим, что алгоритм $G(0,...,0)$ совпадает с обычным алгоритмом покоординатного подъема (см., напр., [3]) для которого ранее в [1,3,5] найдены оценки качества точности этого алгоритма в задачах максимизации ρ – координатно-выпуклых функций.

Доказательство теоремы 1. Из п.3 теоремы 6 [1] при $x=x^g, y=x^*$ и из леммы с учетом соотношений

$$\Delta_i f(x^t)-\rho_i \le \Delta_{i(t)} f(x^t)-\rho_{i(t)}, \quad f(x^t)=f(0)+\sum_{s=0}^{t-1}\Delta_{i(s)} f(x^s),$$

имеем

$$f(x^*)-f(0)\le\sum_{s=0}^{t-1}\Delta_{i(s)}f(x^s)+h\Delta_{i(t)}f(x^t)-h\rho_{i(t)}+h\Omega(\rho)-(h-k)^2\omega(\rho)/2h,$$

$$t=0,1,...,k$$

Отсюда, принимая обозначения

$$a=f(x^*)-f(0)-h\Omega(\rho)+(h-k)^2\omega(\rho)/2h,$$

$$\Delta_s^1=\Delta_{i(s)}f(x^s)/a,\ b(t)=1+h\rho_{i(t)}/a$$

$$\alpha_t = \sum_{s=0}^{t-1} \Delta_s^1$$

(здесь и далее, при $t=0$ считаем, что $\alpha_0 = 0$), получаем

$$\alpha_{t+1} \ge (1-1/h)\alpha_t + b(t)/h\ , t = 0,1,...,k.$$

Из этого рекуррентного соотношения получаем

$$\alpha_t \ge (1-1/h)^t \alpha_0 + \sum_{s=0}^{t-1}(1-1/h)^{t-1-s} b(s)/h\ , t = \overline{0,k}\ .$$

Отсюда с учетом

$$\frac{f(x^g) - f(0)}{a} \ge \sum_{s=0}^{k-1} \Delta_s / a = \alpha_k\ , \alpha_0 = 0$$

и обозначения теоремы выводим оценку (1).

Из оценки (1) с учетом соотношений

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

$$f(x^*) - f(0) \le h\Omega(\delta_f) - h\omega(\rho)/2 = \omega_1(h,\rho,\delta_f),$$

выводим оценку (2).
Теорема 1 доказана.

Повторяя аналогичные, как в теореме 1 вычисления, получаем следующую теорему.

Теорема 2. Пусть неубывающая функция $f(x) \in \Re_\rho(Z_+^n)$. Тогда для градиентного экстремума, построенного с помощью градиентного алгоритма, $G(0,...,0)$ имеет следующие оценки погрешности точности:

1). $f(x^*) \le A(k,h) f(x^g) + (1 - A(k,h)) f(\theta) - B_2(k,h,\rho)$.

2). Если, кроме того выполняется неравенство $\omega_1(k,h,\delta_f) > 0$, то справедлива следующая гарантированная оценка точности градиентного алгоритма $G(0,...,0)$ в задаче максимизации неубывающей функции $f(x)$ $\in \Re_\rho(Z_+^n)$ на множестве P

$$\frac{f(x^*) - f(x^g)}{f(x^*) - f(\theta)} \le B_3(k,h,\rho,\delta_f)\ ,$$

где

$$A(k,h) = (1-(1-1/h)^k)^{-1},$$

$$B_2(k,h,\rho) = B_2 = A(k,h)((h-k)^2 \omega(\rho)/2h,$$

$$B_3(k,h,\rho,\delta_f) = B_3 = (1-1/h)^k -$$

$$A(k,h)(h-k)^3 \omega(\rho)/2h\omega_1(\rho,h,\delta_f),\ \ \omega_1(\rho,h,\delta_f) = h\Omega(\delta_f) - h\omega(\rho)/2$$

$$\delta_f = (\delta_f^1,\ldots,\delta_f^n),\ \delta_f^i = \Delta_i f(\theta),\ \ i = 1,2,\ldots,n,\ \ \Omega(\delta_f) = \sum_{i=1}^{n} \delta_f^i,$$

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

Замечание 2. Хотя алгоритм $G(\rho)$ позволяет найти худшую точность градиентного алгоритма, но преимущество этого алгоритма в том, что число шагов этого алгоритма существенно меньше, чем градиентного алгоритма покоординатного подъема без параметризации.

Литература

1. Emelichev V.A., Kovalev M.M., Ramazanov A.B., Errors of gradient extrema of a strictly convex function of discrete argument, J. Discrete Mathematics and Applications, 1992, Vol.2, No.2, 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, 2011, Vol.21, Issue 4, pp.465-476.
3. Ковалев М.М., Матроиды в дискретной оптимизации, Изд-во БГУ, Минск, 1987, 222 с.
4. 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.
5. Шор Н.З., Методы минимизации недифференцируемых функций и их приложения, Киев, Науково думка, 1979, 200 с.

Qradiyent alqoritmin parametrləşdirilməsi ilə qradiyent ekstremumun qiymətləndirilməsi haqqinda

Ə.B. Ramazanov

XÜLASƏ

İşdə koordinat qəfəslərin sonlu alt çoxluğunda qradiyent alqoritmin ciddi qabarıqlıq ölçüsü ilə parametrləşdirilməsinə baxılır. Belə ki, qradiyent alqoritmə parametrin daxil edilməsi hesabına qradiyent alqoritmin addımlarının sayı ciddi azalır.

Açar sözlər: qradiyent, diskret, alqoritm, xəta, qabarıqlıq.

On a estimation of the gradient extremum by the help of parametrization of the gradient algorithm

A.B. Ramazanov

ABSTRACT

In the present work variant of the gradient method? Parametrized by the help of measure of the strong convexity of the cost function on the finite subset of coordinate grid. Introducing a parameter in the gradient algorithm the number of steps of the algorithm is essentially reduced.

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