УДК 519.676

А.С.Тихомиров

О БЫСТРЫХ АЛГОРИТМАХ ОДНОРОДНОГО МАРКОВСКОГО МОНОТОННОГО ПОИСКА ЭКСТРЕМУМА

Институт электронных и информационных систем НовГУ, Alexey.Tikhomirov@novsu.ru

An estimate of the convergence rate of some homogeneous Markov monotone random search optimization algorithms is obtained. This estimate is used to construct a class of fast optimization methods. It is shown that the number of evaluations of the objective function required for achieving a given accuracy ε increases slowly (logarithmically) as ε tends to zero.

Ключевые слова: гарантирующее число шагов, пространство оптимизации, скорость сходимости

1. Введение

Работа посвящена теоретическому исследованию алгоритмов случайного поиска экстремума функции (см. [1-6]). В ней представлена оценка скорости сходимости однородного марковского монотонного случайного поиска, и с ее помощью построены быстрые алгоритмы оптимизации невырожденных целевых функций. Данная работа является продолжением работы [12] и дополняет результаты, изложенные в [3,7-12]. Доказательства всех утверждений приведены в [13].

В качестве *пространства оптимизации* рассмотрим пространство $\mathbf{R}^d$ с какой-либо обычной метрикой ρ (см. [12]) и *d*-мерной мерой Лебега mes. Замкнутый шар радиуса *r* с центром в точке *x* обозначим как $S_r(x)=\{y\in\mathbf{R}^d:\rho(x,y)\le r\}$ и положим $\varphi(r)=\mathrm{mes}(S_r(x))$.

Пусть *целевая функция* $f:\mathbf{R}^d\mapsto\mathbf{R}$ принимает максимальное значение в единственной точке $x_0=\arg\max\{f(x):x\in\mathbf{R}^d\}$, а нашей целью является отыскание точки x_0 с заданной точностью $\varepsilon>0$. Для поиска точки максимума воспользуемся *однородным марковским монотонным* случайным поиском (см. [12]), описанным далее с помощью алгоритма моделирования.

Алгоритм 1

Шаг 1. $\xi_0\leftarrow x$, $i\leftarrow 1$.

Шаг 2. $\varsigma\leftarrow P(\xi_{i-1},\cdot)$.

Шаг 3. Если $f(\varsigma)\ge f(\xi_{i-1})$, то $\xi_i\leftarrow\varsigma$, иначе $\xi_i\leftarrow\xi_{i-1}$.

Шаг 4. $i\leftarrow i+1$ и перейти к шагу 2.

Здесь *x* — начальная точка поиска, а $P(x,\cdot)$ — марковские переходные функции (см. [3]), называемые *пробными переходными функциями*. Отметим, что введенный случайный поиск является *монотонным* в том смысле, что неравенства $f(\xi_i)\ge f(\xi_{i-1})$ выполняются при всех $i>0$. Ниже для вероятностей событий и математических ожиданий случайных величин, связанных со случайным поиском алгоритма 1, начинающимся в точке $x\in\mathbf{R}^d$, используются обозначения $\mathbf{P}_x$ и $\mathbf{E}_x$.

Особое внимание будет уделено однородному марковскому монотонному случайному поиску, пробные переходные функции $P(x,\cdot)$ которого обладают *симметричной* плотностью вида $p(x,y)=g(\rho(x,y))$, где ρ — метрика, а g — невозрастающая функция, определенная на полуоси $(0,+\infty)$. Функцию *g* назовем *формой поиска*.

При отыскании точки максимума x_0 с заданной точностью $\varepsilon>0$ нас будет интересовать попадание поиска в множество

$$M_\varepsilon=M(\varepsilon)=\{x\in S_\varepsilon(x_0):f(x)>f(y)$$
$$\text{для любого } y\notin S_\varepsilon(x_0)\}.$$

Монотонный поиск, попав в множество M_ε, из него больше не выйдет. Поэтому мы будем изучать момент попадания поиска в множество M_ε. Обозначим $\tau_\varepsilon=\min\{i\ge 0:\xi_i\in M_\varepsilon\}$ — момент первого попадания поиска в множество M_ε. *Трудоемкость* случайного поиска определяется как $\mathbf{E}_x\tau_\varepsilon$ и имеет смысл среднего числа шагов поиска до достижения им множества M_ε. В [12] мы ограничились изучением трудоемкости случайного поиска. В этой работе мы исследуем другую характеристику τ_ε. *Гарантирующее число шагов* определяется как такое минимальное число $N=N(x,f,\varepsilon,\gamma)$ шагов поиска, при котором достижение множества M_ε гарантировано с вероятностью не меньшей γ. Иначе говоря,

$$N(x,f,\varepsilon,\gamma)=\min\{i\ge 0:\mathbf{P}_x(\xi_i\in M_\varepsilon)\ge\gamma\}=$$
$$=\min\{i\ge 0:\mathbf{P}_x(\tau_\varepsilon\le i)\ge\gamma\}.$$

Если целочисленная функция $N_1(x,f,\varepsilon,\gamma)$ обладает тем свойством, что для $\gamma\in(0,1)$ выполнено $\liminf_{\varepsilon\to 0}\mathbf{P}_x(\xi_{N_1}\in M_\varepsilon)\ge\gamma$, то N_1 называется *асимптотически гарантирующим числом шагов* поиска с надежностью γ.

Для построения быстрых алгоритмов случайного поиска на целевую функцию необходимо наложить дополнительные ограничения. Далее будем полагать, что целевая функция $f:\mathbf{R}^d\mapsto\mathbf{R}$ ограничена сверху, измерима и удовлетворяет следующим условиям.

Условие 1. Функция *f* принимает максимальное значение в единственной точке $x_0=\arg\max\{f(x):x\in\mathbf{R}^d\}$.

Условие 2. Функция *f* непрерывна в точке x_0.

Условие 3. Неравенство $\sup\{f(x):x\notin S_r(x_0)\}<$ $<f(x_0)$ выполнено для любого $r>0$.

Условие 4. $\bigcup_{r>0}M_r=\mathbf{R}^d$.

Ниже информация о целевой функции *f* будет выражаться в виде *коэффициента асимметрии* $F_f(r)=\mathrm{mes}(M_r)/\mathrm{mes}(S_r(x))$. Коэффициент асимметрии «сравнивает» поведение *f* с *F-идеальной* одноэкстремальной функцией f_*, для которой $F_{f_*}\equiv 1$. В силу условий, наложенных на целевую функцию, $F_f(r)>0$ при всех $r>0$. Функции, у которых $\liminf F_f(r)>0$ при $r\to 0$, будут называться *невырожденными.* Подробнее наложенные на *f* ограничения и свойства коэффициента асимметрии F_f обсуждаются в [3, 9].

2. Оценки скорости сходимости

Пусть параметры оценки $\{r_i\}_{i=-\infty}^{+\infty}$ таковы, что $0<r_i<r_{i-1}$ при всех *i*, $r_i\to 0$ при $i\to+\infty$ и $r_i\to+\infty$ при $i\to-\infty$. Обозначим $n(\varepsilon)=\min\{i:r_i\le\varepsilon\}$ и $t(x)=\sup\{i:x\in M(r_i)\}$.

Пусть $t<n$ и $0<v_i\le 1$ при $t+1\le i\le n$. Введем величины

$$Y(t,n,v_{t+1},\ldots,v_n)=\sum_{i=t+1}^{n}\frac{1}{v_i},\quad D(t,n,v_{t+1},\ldots,v_n)=\sum_{i=t+1}^{n}\frac{1-v_i}{v_i^2},$$

$$K(t,n,v_{t+1},\ldots,v_n)=\sum_{i=t+1}^{n}\frac{1}{v_i^3}.$$

Приведем вначале оценку трудоемкости случайного поиска работы [12].

Теорема 1. Пусть целевая функция f удовлетворяет условиям 1-4 и однородный марковский монотонный случайный поиск алгоритма 1 начинается в точке $x \notin M_\varepsilon$. Пусть $t = t(x)$, $n = n(\varepsilon)$ и неравенства $0 < v_i \le \inf\{P(y, M(r_i)) : y \in M(r_{i-1})\}$ верны при всех $t+1 \le i \le n$. Тогда трудоемкость случайного поиска удовлетворяет неравенству $\mathbf{E}_x \tau_\varepsilon \le Y(t, n, v_{t+1}, \ldots, v_n)$.

Получим асимптотически гарантирующее число шагов и оценки гарантирующего числа шагов для исследуемого случайного поиска. Отметим, что «простая» оценка гарантирующего числа шагов сразу следует из теоремы 1. В условиях теоремы 1, в силу неравенства Маркова, имеет место неравенство $\mathbf{P}_x(\tau_\varepsilon \le Y(t, n, v_{t+1}, \ldots, v_n)/(1-\gamma)) \ge \gamma$. Значит величина

$$N_{\mathrm{M}}(t, n, v_{t+1}, \ldots, v_n, \gamma) = [Y(t, n, v_{t+1}, \ldots, v_n)/(1-\gamma)], \quad (1)$$

где через $[z]$ обозначена целая часть числа z, служит оценкой сверху гарантирующего числа шагов случайного поиска.

Далее получена другая оценка гарантирующего числа шагов случайного поиска, и показано ее превосходство над «простой» оценкой N_{M}. Обозначим

$$N_0(t, n, v_{t+1}, \ldots, v_n, \gamma) =$$
$$= [Y(t, n, v_{t+1}, \ldots, v_n) + \Phi^{-1}(\gamma)(D(t, n, v_{t+1}, \ldots, v_n))^{1/2}], \quad (2)$$

где $\gamma \in (0,1)$, Φ — функция распределения стандартного нормального закона, функция Φ^{-1} является обратной к Φ.

Основной результат данной работы представляет следующая теорема.

Теорема 2. В условиях теоремы 1 верно неравенство

$$\mathbf{P}_x(\xi_{N_0} \in M_\varepsilon) \ge$$
$$\ge \gamma - c_0 16 K(t, n, v_{t+1}, \ldots, v_n)/(D(t, n, v_{t+1}, \ldots, v_n))^{3/2},$$

где c_0 — абсолютная константа неравенства Эссеена.

Следствие 1. Пусть в условиях теоремы 2 для семейства пробных переходных функций $P^{(\varepsilon)}$ (зависящих от ε) выполнено соотношение

$$\lim_{\varepsilon \to 0} K(t, n, v_{t+1}, \ldots, v_n)/(D(t, n, v_{t+1}, \ldots, v_n))^{3/2} = 0, \quad (3)$$

где $v_i = v_i^{(\varepsilon)}$, и пусть величина $N_0 = N_0(t, n, v_{t+1}, \ldots, v_n, \gamma)$ задается формулой (2). Тогда

$$\mathbf{P}_x(\xi_{N_0} \in M_\varepsilon) \ge \gamma - c_0 16 K(t, n, v_{t+1}, \ldots, v_n)/$$
$$/(D(t, n, v_{t+1}, \ldots, v_n))^{3/2} \xrightarrow[\varepsilon \to 0]{} \gamma. \quad (4)$$

Соотношение (4) показывает, что величина $N_0(t, n, v_{t+1}, \ldots, v_n, \gamma)$ является асимптотически гарантирующим числом шагов случайного поиска с надежностью γ. Кроме того она является оценкой сверху гарантирующего числа шагов случайного поиска с надежностью $\gamma_0 = \gamma - c_0 16 K(t, n, v_{t+1}, \ldots, v_n)/$ $/(D(t, n, v_{t+1}, \ldots, v_n))^{3/2}$.

3. Быстрые алгоритмы случайного поиска

В этом разделе получим целый класс однородных поисков, дающих для невырожденных целевых функций оценки трудоемкости и гарантирующего числа шагов вида $\mathrm{O}(\ln^2 \varepsilon)$. Мы рассмотрим семейство однородных марковских монотонных случайных поисков алгоритма 1, пробные переходные функции которых зависят от требуемой точности решения задачи ε и обладают симметричными плотностями $p_\varepsilon(x, y) = g_\varepsilon(\rho(x, y))$ с формами g_ε, задаваемыми следующим образом. Пусть $h(r)$ — монотонно невозрастающая строго положительная функция, определенная на полуоси $(0, +\infty)$ и такая, что функция $h(r) r^{d-1}$ суммируема на промежутке $[1, +\infty)$. Кроме того предположим, что $h(r) r^d \to 1$ при $r \to 0$. Не умаляя общности будем считать, что функция h непрерывна слева. Зафиксируем параметр $a > 0$ и положим при $\varepsilon > 0$

$$g_\varepsilon(r) = \lambda(\varepsilon) \begin{cases} h(a\varepsilon) \text{ при } r \le a\varepsilon, \\ h(r) \text{ при } r > a\varepsilon, \end{cases} \quad (5)$$

где множитель $\lambda(\varepsilon)$ обеспечивает условие нормировки (необходимое для плотности).

Определим теперь параметры используемой оценки. Зафиксируем коэффициент сжатия $q \in (0,1)$, радиус $R > 0$ и зададим радиусы окрестностей точки максимума следующим образом: $r_i = R q^i$. Положим $v_i^{(\varepsilon)} = F_f(r_i) \varphi(r_i) g_\varepsilon(r_{i-1} + r_i)$.

Следующее утверждение уточняет результаты теорем 1 и 2 и следствия 1 для поисков с формами (5). Оказывается, что для этих поисков из невырожденности целевой функции f следует выполнение условия (3). Поэтому в этих условиях выполняется соотношение (4).

Теорема 3. Пусть целевая функция f удовлетворяет условиям 1-4 и является невырожденной. Тогда для однородных марковских случайных поисков с формами (5) и начальной точкой $x \ne x_0$ верны соотношение (4) и равенства

$$Y(t, n, v_{t+1}, \ldots, v_n) = \mathrm{O}(\ln^2 \varepsilon),\ D(t, n, v_{t+1}, \ldots, v_n) = \mathrm{O}(|\ln^3 \varepsilon|),$$
$$K(t, n, v_{t+1}, \ldots, v_n)/(D(t, n, v_{t+1}, \ldots, v_n))^{3/2} = \mathrm{O}(|\ln \varepsilon|^{-1/2}),$$

где $t = t(x)$, $n = n(\varepsilon)$, $r_i = R q^i$, $v_i = v_i^{(\varepsilon)} = F_f(r_i) \varphi(r_i) g_\varepsilon(r_{i-1} + r_i)$.

Таким образом, случайные поиски теоремы 3 являются быстрыми, их трудоемкость и гарантирующее число шагов имеют вид $\mathrm{O}(\ln^2 \varepsilon)$. Отметим, что для методов стохастической глобальной оптимизации (см., напр., [3-5]) типичным результатом является гораздо более худшая — степенная (т. е. $\mathrm{O}(1/\varepsilon^\alpha)$ при $\alpha > 0$) зависимость требуемого числа вычислений целевой функции от ε.

В заключение покажем превосходство новой оценки N_0 (см. (2)) над «старой» оценкой N_{M} (см. (1)), полученной с помощью неравенства Маркова. Из определения N_0 видно, что в условиях теоремы 3 и для фиксированных x и γ величина

$N_0(t,n,v_{t+1},\ldots,v_n,\gamma)$ асимптотически эквивалентна величине $Y(t,n,v_{t+1},\ldots,v_n)$ при $\varepsilon \to 0$. Асимптотика оценки $N_{\rm M}$ (см. (1)) будет хуже. Приведем числовой пример для сравнения величины N_0 с оценкой $N_{\rm M}$ при «большем» $\varepsilon = 10^{-2}$. Возьмем пространство оптимизации $\mathbf{R}^2$ с метрикой ρ_∞ (см. [12]), F-идеальную функцию f, $x_0 = (0,0)$ и $x = (1,1)$. Параметры поиска и параметры оценки описаны в [13]. В качестве статистической оценки гарантирующего числа шагов использованы выборочные квантили $N_*(x,f,\varepsilon,\gamma)$, для вычисления которых поиск повторялся 10^7 раз. Результаты статистического моделирования и расчетов представлены в следующей таблице.

Оценки гарантирующего числа шагов

Надежность γ	0,9	0,95	0,99	0,995	0,999	0,9995	0,9999
N_*	196	238	338	381	481	526	627
N_0/N_*	2,14	1,96	1,64	1,54	1,35	1,29	1,17
$N_{\rm M}/N_*$	13	22	76	135	535	978	4101

Видно, что новая оценка N_0 во много раз лучше старой оценки $N_{\rm M}$, а при больших значениях надежности γ преимущество становится огромным. Так как с практической точки зрения интересны как раз большие значения надежности, то превосходство новой оценки очевидно.

1. Ермаков С.М., Жиглявский А.А. // Теория вероятностей и ее применения. 1983. №1. С.129-136.
2. Ермаков С.М., Жиглявский А.А., Кондратович М.В. // Журнал вычисл. математики и мат. физики. 1989. Т.29. №2. С.163-170.
3. Zhigljavsky A., Zilinskas A. Stochastic Global Optimization. Berlin: Springer-Verlag, 2008. 262 p.
4. Spall J.C. Introduction to stochastic search and optimization: estimation, simulation, and control. Wiley, New Jersey, 2003. 618 p.
5. Spall J.C., Hill S.D., Stark D.R. Theoretical framework for comparing several stochastic optimization approaches // Probabilistic and randomized methods for design under uncertainty. L.: Springer, 2006. P.99-117.
6. Абакаров А.Ш., Сушков Ю.А. Статистическое исследование случайного поиска // Математические модели. Теория и приложения / Под ред. М.К.Чиркова. Вып.2. СПб.: Изд-во НИИХ СПбГУ, 2002. С.70-86.
7. Тихомиров А.С. // Вестник НовГУ. Сер.: Техн. науки. 2005. №34. С.90-95.
8. Тихомиров А.С. // Вестник НовГУ. 2006. №39. С.34-37.
9. Тихомиров А.С. // Журнал вычисл. математики и мат. физики. 2006. Т.46. №3. С.379-394.
10. Тихомиров А.С. // Журнал вычисл. математики и мат. физики. 2007. Т.47. №5. С.817-828.
11. Tikhomirov A., Stojunina T., Nekrutkin V. // J. of Statistical Planning and Inference. 2007. Vol.137. Issue 12. P.4031-4047.
12. Тихомиров А.С. // Вестник НовГУ. 2007. №44. С.51-54.
13. Тихомиров А.С. Деп. в ВИНИТИ №68-В2007 от 24.01.2007. 57 с.