УДК 519.676

А.С.Тихомиров

О БЫСТРЫХ АЛГОРИТМАХ СЛУЧАЙНОГО ПОИСКА

Новгородский государственный университет им. Ярослава Мудрого

A class of random search methods for a global maximum of an objective function is investigated. It is shown that the number of an objective function evaluation required to reach a given accuracy for methods from this class has a slow (logarithmic) order of growth as the accuracy tends to zero.

Введение

Рассмотрим семейство хорошо известных [1, 2] методов случайного поиска экстремума функции. Такие методы успешно используются при решении сложных задач оптимизации. Однако сравнительно мало теоретических работ о скорости сходимости этих алгоритмов. В статье продолжено исследование скорости сходимости «быстрых» алгоритмов случайного поиска, построенных в работе [3] и принадлежащих рассматриваемому классу методов оптимизации.

Метод поиска будем называть «быстрым», если число вычислений целевой функции, требуемое для достижения заданной точности ε, имеет медленный (логарифмический) порядок роста при стремлении ε к нулю. Примеры быстрых алгоритмов случайного поиска можно найти в [3-5]. Данная работа улучшает результаты [4] и дополняет результаты [3, 5].

1. Постановка задачи

1.1. Пространство оптимизации

Назовем *пространством оптимизации* множество оптимизации X, снабженное метрикой ρ. Мы ограничимся случаем $X = \mathbf{R}^d$, d-мерной мерой Лебега mes, и следующими вариантами метрик $\rho(x, y)$ для $\mathbf{R}^d$:

$$\rho_\nu(x,y) = \left(\sum_{i=1}^{d} |x_i - y_i|^\nu \right)^{1/\nu}, \quad \rho_\infty(x,y) = \max_{1 \le i \le d} |x_i - y_i|,$$

где $\nu \ge 1$ — любое фиксированное число, $x = (x_1, ..., x_d)$ и $y = (y_1, ..., y_d)$. Замкнутый шар радиуса r с центром в x обозначим как $S_r(x) = \{y \in \mathbf{R}^d : \rho(x,y) \le r\}$.

1.2. Целевая функция

В дальнейшем всегда будет предполагаться, 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 $r > 0$.

Ввиду условия 3 из сходимости $f(x_n) \to f(x_0)$ следует, что $\rho(x_n, x_0) \to 0$. Отметим, что функции указанного класса могут быть многоэкстремальными в любой окрестности глобального максимума. Еще одно условие на целевую функцию будет введено ниже.

1.3. Случайный поиск

Случайным поиском называется произвольная (конечная или бесконечная) последовательность случайных величин $\{\xi_i\}_{i \ge 0}$ со значениями в $\mathbf{R}^d$. Применим случайный поиск для оценки точки максимума x_0 с заданной положительной точностью $\varepsilon > 0$ и опишем исследуемый поиск с помощью алгоритма моделирования. Обозначение «$\eta \leftarrow P(\cdot)$» читается как «получить реализацию случайного вектора η с распределением P».

Алгоритм 1

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

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

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

Шаг 4. Если $i < n$, то $i \leftarrow i+1$ и перейти к шагу 2, иначе — STOP.

Здесь x — начальная точка поиска, n — число шагов поиска, а $P_i(x, \cdot)$ — вспомогательные переходные функции. Будем полагать, что переходные функции $P = P_i$ не зависят от номера шага i. Таким образом, исследуемый случайный поиск является *однородным*. Кроме того, введенный случайный поиск является *монотонным* в том смысле, что неравенства $f(\xi_i) \ge f(\xi_{i-1})$ выполняются при всех $i > 0$.

Далее будем полагать, что распределение вероятностей $P(x, dy)$ обладает симметричной плотностью вида $p(x, y) = g(\rho(x, y))$, где ρ — метрика, а g — невозрастающая неотрицательная функция, определенная на полуоси $(0,+\infty)$. Легко видеть, что тогда $p(x,x+y)=p(0,y)$ при всех $y\neq 0$, $x\in\mathbf{R}^d$. Функцию g будем называть *формой поиска*. Не умаляя общности будем считать, что функция g непрерывна слева.

Описанный поиск будем называть *однородным марковским монотонным симметричным* случайным поиском.

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 $\mathbf{E}_x$.

1.4. Цель поиска

При изучении случайного поиска нам придется анализировать попадание поиска в окрестность $S_r(x_0)$ точки максимума x_0. Может, однако, случиться так, что поиск, оказавшись в $S_r(x_0)$ на шаге *i*, выйдет из $S_r(x_0)$ на одном из следующих шагов. Чтобы избежать анализа таких эффектов, введем множества

$$M_r=M(r)=\{x\in S_r(x_0): f(x)>f(y) \text{ для любого } y\notin S_r(x_0)\}.$$

Легко видеть, что множества M_r обладают следующими свойствами: а) если $r<z$, то $M_r\subset M_z$, б) если $x\in M_r$ и $y\notin M_r$, то $f(x)>f(y)$. В силу своей монотонности поиск, попав во множество M_r, из него больше не выйдет. Поэтому мы будем изучать момент попадания поиска во множество M_r, а не в шар $S_r(x_0)$ (где *r* сохраняет смысл достигнутой точности поиска). Соответственно мерой близости точки x к x_0 оказывается не расстояние $\rho(x,x_0)$, а число $\delta(x)=\inf\{r\geq 0: x\in M_r\}$.

1.5. Информация о целевой функции

Нам потребуется еще одно ограничение на поведение целевой функции *f*.

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

Это условие гарантирует попадание любой начальной точки поиска во множество M_r при некотором *r*. В силу своей монотонности поиск не может покинуть множество M_r и, значит, не может неограниченно удаляться от точки x_0.

Далее всегда будем полагать, что целевая функция удовлетворяет условиям 1-4.

В задачах оптимизации сведения о целевой функции *f* присутствуют в двух видах. Во-первых, от свойств функции *f* зависит скорость сходимости случайного поиска к точке экстремума (и оценки этой скорости). Во-вторых, некоторые заранее известные характеристики целевой функции могут использоваться в качестве априорной информации при построении поиска. Ниже информация о целевой функции *f* будет содержатся 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 *h*, для которой $F_h\equiv 1$. 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*. Иногда вместо $F_f(r)$ будет удобно иметь дело с функцией $m(r)=m_f(r)=\mathrm{mes}(M_r)$. Функция $m_f(r)$ называется *функцией асимметрии* целевой функции. Подробнее свойства множеств M_r и функций m_f, F_f и δ обсуждаются в [3-5].

1.6. Характеристики случайного поиска

Положим $n=+\infty$ в алгоритме 1 и обозначим $\tau_\varepsilon=\min\{i\geq 0:\xi_i\in M_\varepsilon\}$ — момент первого попадания поиска в множество M_ε. Мы всегда будем предполагать, что для моделирования распределений P_i в алгоритме 1 не требуется вычислений функции *f*. Тем самым при каждой итерации $\xi_{i-1}\mapsto\xi_i$ алгоритма 1 происходит ровно одно вычисление целевой функции, и распределение случайной величины τ_ε дает нам достаточно полную информацию о качестве случайного поиска. Действительно, при выполнении τ_ε итераций алгоритма значения функции *f* вычисляются $\tau_\varepsilon+1$ раз.

В [3] мы ограничились изучением одной характеристики случайной величины τ_ε — трудоемкости. *Трудоемкость* случайного поиска определяется как $\mathbf{E}_x\tau_\varepsilon$ и имеет смысл среднего числа шагов поиска до достижения им множества M_ε.

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

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

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 $a_i=r_i+r_{i-1}$. Положим $t(x)=\sup\{i:x\in M(r_i)\}$ и $\kappa_\varepsilon=\min\{i:r_i\leq\varepsilon\}$. Отметим, что при $x\neq x_0$ верны неравенства: $r_{t(x)+1}<\delta(x)\leq r_{t(x)}$.

При $t<\kappa$ введем величины

$$J(t,\kappa,f,g)=\frac{1}{m(r_{t+1})g(a_{t+1})}+\sum_{i=t+2}^{\kappa}\frac{1}{g(a_i)}\left(\frac{1}{m(r_i)}-\frac{1}{m(r_{i-1})}\right),$$

$$D(t,\kappa,f,g)=\frac{1}{m^2(r_{t+1})g^2(a_{t+1})}+$$

$$+\sum_{i=t+2}^{\kappa}\frac{1}{g^2(a_i)}\left(\frac{1}{m^2(r_i)}-\frac{1}{m^2(r_{i-1})}\right)-J(t,\kappa,f,g),$$

$$K(t,\kappa,f,g)=\frac{1}{m^3(r_{t+1})g^3(a_{t+1})}+$$

$$+\sum_{i=t+2}^{\kappa}\frac{1}{g^3(a_i)}\left(\frac{1}{m^3(r_i)}-\frac{1}{m^3(r_{i-1})}\right).$$

Здесь m — функция асимметрии, а g — форма поиска.

Полезную информацию о зависимости J, D и K от свойств целевой функции и начальной точки поиска (при использовании $t=t(x)$) дает следующее утверждение.

Теорема 1. 1. Пусть $t<\kappa$, g — форма поиска и $g(a_{t+1})>0$, а функции f и h таковы, что $F_h(r_i)\le F_f(r_i)$ при всех $t+1\le i\le\kappa$. Тогда $J(t,\kappa,f,g)\le J(t,\kappa,h,g)$, $D(t,\kappa,f,g)\le D(t,\kappa,h,g)$ и $K(t,\kappa,f,g)\le K(t,\kappa,h,g)$.

2. Если $\theta\le t<\kappa$, g — форма поиска и $g(a_{\theta+1})>0$, то $J(t,\kappa,f,g)\le J(\theta,\kappa,f,g)$, $D(t,\kappa,f,g)\le$ $\le D(\theta,\kappa,f,g)$ и $K(t,\kappa,f,g)\le K(\theta,\kappa,f,g)$.

Из теоремы 1 следует, что в приводимых далее оценках трудоемкости и гарантирующего числа шагов случайного поиска вместо коэффициента асимметрии $F_f(r)$ и величины $t(x)$, точные значения которых могут быть неизвестны, можно использовать оценки снизу и коэффициента асимметрии $F_f(r)$ и величины $t(x)$.

Приведем вначале оценку трудоемкости случайного поиска из работы [3].

Теорема 2. Для любой целевой функции f, удовлетворяющей условиям 1-4, и любого однородного марковского монотонного симметричного случайного поиска, начинающегося в точке x, при $0<\varepsilon<\delta(x)$ и $g(a_{t(x)+1})>0$ верно неравенство $\mathbf{E}_x\tau_\varepsilon\le J(t(x),\kappa_\varepsilon,f,g)$.

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 которой в силу неравенства Маркова имеет место неравенство $\mathbf{P}_x(\tau_\varepsilon\le J(t(x),\kappa_\varepsilon,f,g)/(1-\gamma))\ge\gamma$. Значит величина

$$N_M(t(x),\kappa_\varepsilon,f,g,\gamma)=[J(t(x),\kappa_\varepsilon,f,g)/(1-\gamma)],\quad(1)$$

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

Для невырожденных целевых функций получены более точные оценки асимптотически гарантирующего числа шагов и гарантирующего числа шагов поиска.

Теорема 3. Пусть для $x\ne x_0$, функции f, удовлетворяющей условиям 1-4, и семейства форм поиска g_ε выполнены соотношения $g_\varepsilon(a_{t(x)+1})>0$ и

$$\lim_{\varepsilon\to0}K(t(x),\kappa_\varepsilon,f,g_\varepsilon)/(D(t(x),\kappa_\varepsilon,f,g_\varepsilon))^{3/2}=0.\quad(2)$$

Тогда для любого $v\in\mathbf{R}$

$$\limsup_{\varepsilon\to0}\mathbf{P}_x(\tau_\varepsilon>J(t(x),\kappa_\varepsilon,f,g_\varepsilon)+$$
$$+v\sqrt{D(t(x),\kappa_\varepsilon,f,g_\varepsilon)})\le1-\Phi(v),\quad(3)$$

где Φ — функция распределения стандартного нормального закона.

Ясно, что неравенство (3) позволяет получить асимптотически гарантирующее число шагов случайного поиска, причем поведение этой величины при $\varepsilon\to0$ определяется порядками стремления к бесконечности $J(t(x),\kappa_\varepsilon,f,g_\varepsilon)$ и $D(t(x),\kappa_\varepsilon,f,g_\varepsilon)$.

Теорема 4. Пусть в условиях теоремы 3 $v=\Phi^{-1}(\gamma)$ для $\gamma\in(0,1)$. Обозначим

$$N_0(t(x),\kappa_\varepsilon,f,g_\varepsilon,\gamma)=[J(t(x),\kappa_\varepsilon,f,g_\varepsilon)+$$
$$+\Phi^{-1}(\gamma)\sqrt{D(t(x),\kappa_\varepsilon,f,g_\varepsilon)}].\quad(4)$$

Тогда

$$\mathbf{P}_x(\xi_{N_0}\in M_\varepsilon)\ge\gamma-16c_0K(t(x),\kappa_\varepsilon,f,g_\varepsilon)/$$
$$/(D(t(x),\kappa_\varepsilon,f,g_\varepsilon))^{3/2}\xrightarrow[\varepsilon\to0]{}\gamma,\quad(5)$$

где c_0 — абсолютная константа неравенства Эссеена.

Неравенство (5) показывает, что асимптотически гарантирующее число шагов $N_0(t(x),\kappa_\varepsilon,f,g_\varepsilon,\gamma)$ является оценкой сверху гарантирующего числа шагов случайного поиска с надежностью $\gamma_0=$ $=\gamma-16c_0K(t(x),\kappa_\varepsilon,f,g_\varepsilon)/(D(t(x),\kappa_\varepsilon,f,g_\varepsilon))^{3/2}$. Таким образом, при малых ε неравенство (5) позволяет получить оценку сверху гарантирующего числа шагов случайного поиска.

Следующее утверждение уточняет результаты для поисков теорем 3 и 4 из [3].

Теорема 5. 1. Для поисков теорем 3 и 4 из [3] из невырожденности целевой функции f следует выполнение условия (2). Поэтому выполняется неравенство (3).

2. Для поисков теорем 3 и 4 из [3] и невырожденной функции f имеют место соотношения $J(t(x),\kappa_\varepsilon,f,g_\varepsilon)=h_1(t(x),\kappa_\varepsilon,f,g_\varepsilon)\kappa_\varepsilon^2$ и $D(t(x),\kappa_\varepsilon,f,g_\varepsilon)=$ $=h_2(t(x),\kappa_\varepsilon,f,g_\varepsilon)\kappa_\varepsilon^3$, где $\kappa_\varepsilon=O(|\ln\varepsilon|)$ задается формулой (6) из [3], а функции h_1 и h_2 ограничены. Величины $J(t(x),\kappa_\varepsilon,f,g_\varepsilon)$ и $N_0(t(x),\kappa_\varepsilon,f,g_\varepsilon,\gamma)$ асимптотически эквивалентны при $\varepsilon\to0$. Кроме того, существуют такие ограниченные функции h, что

$$\mathbf{P}_x(\xi_{N_0}\in M_\varepsilon)\ge\gamma-h(t(x),\kappa_\varepsilon,f,g_\varepsilon)/\sqrt{\kappa_\varepsilon}\xrightarrow[\varepsilon\to0]{}\gamma\,.$$

Таким образом, случайные поиски теорем 3 и 4 из [3] являются быстрыми. Их трудоемкость и гарантирующее число шагов имеют медленный (логарифмический) порядок роста при стремлении ε к нулю. Кроме того, оценки их трудоемкости и гарантирующего числа шагов асимптотически эквивалентны. Для сравнения отметим, что для методов стохастической глобальной оптимизации (см., например, [1]) типичным результатом является гораздо более худшая — степенная (т. е. $O(1/\varepsilon^\alpha)$ при $\alpha>0$) зависимость требуемого числа вычислений целевой функции от ε.

Оценки гарантирующего числа шагов

Надежность γ	0,9	0,95	0,99	0,995	0,999	0,9995	0,9999
N_*	176	211	294	330	412	448	531
N_0 / N_*	2,22	2,05	1,75	1,65	1,47	1,41	1,28
N_M / N_*	14	23	81	144	577	1061	4476

В завершение для поиска теоремы 4 из [3] сравним величину N_0 (см. (4)) с оценкой N_M (см. (1)), полученной с помощью неравенства Маркова. Продолжим рассмотрение числового примера из [3]. Возьмем пространство $(\mathbf{R}^2, \rho_\infty)$, $\varepsilon = 0{,}01$, $\delta(x) = R = 1$, $\rho_2(x, x_0) = \sqrt{2}$, $F_f \equiv 1$. Тогда при $q = 0{,}3981$ имеем $J = 238$, $D = 14229$, $\sqrt{D} = 119$. В качестве статистической оценки гарантирующего числа шагов используем выборочные квантили $N_*(x, f, \varepsilon, \gamma)$, для вычисления которых поиск повторялся 10^7 раз. В результате численных экспериментов и расчетов при различных значениях надежности γ получим результаты, представленные в таблице (см.).

Отметим, во-первых, что величина N_* и ее оценка N_0 достаточно медленно растут с увеличением надежности γ. Медленный рост гарантирующего числа шагов при увеличении γ — это важное достоинство рассматриваемого семейства методов случайного поиска. Кроме того, полученная оценка N_0 существенно (во много раз) лучше «простой» оценки N_M, которая получается при использовании результатов [3].

1. Жиглявский А.А., Жилинскас А.Г. Методы поиска глобального экстремума. М.: Наука, 1991. 248 с.
2. 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.
3. Тихомиров А.С. // Вестник НовГУ. Сер.: Техн. науки. 2005. №34. С.90-95.
4. Тихомиров А.С., Некруткин В.В. Марковский монотонный поиск экстремума. Обзор некоторых теоретических результатов // Математические модели. Теория и приложения / Под ред. М.К.Чиркова. Вып. 4. СПб.: ВВМ, 2004. С.3-47.
5. Тихомиров А.С. // Журнал вычислительной математики и математической физики. 2006. Т.46. №3. С.379-394.