А.С.Тихомиров

О СКОРОСТИ СХОДИМОСТИ ОДНОРОДНОГО МАРКОВСКОГО ПОИСКА

The multidimensional problem of optimization is simulated by means of the uniform Markov monotonic random search. The intensiveness characteristics and guaranteed step number are investigated. Earlier the author has obtained the upper estimates for them. In this paper, they are made more precise by the below estimates.

Для широкого класса локальных методов оптимизации и регулярных целевых функций (см., например, [1, 2]) число шагов поиска до достижения точности ε имеет вид $O(|\ln \varepsilon|)$. Для методов стохастической глобальной оптимизации (см. [3]) типичным результатом является гораздо более худшая — степенная зависимость ($O(1/\varepsilon^t)$ при $t > 0$). В работах [4, 5] представлены такие варианты однородного марковского монотонного симметричного случайного поиска, для которых оценки сверху величин трудоемкости и гарантирующего числа шагов имеют вид $O(\ln^2 \varepsilon)$.

В представленной работе продолжено исследование описанного в [4, 5] однородного марковского монотонного симметричного случайного поиска.

Случайный поиск

Назовем *пространством оптимизации* множество оптимизации X , снабженное метрикой ρ . Мы ограничимся случаем $X = \mathbf{R}^k$ и метрикой $\rho(x, y) = \rho_\infty(x, y) = \max_{1 \leq i \leq k} |x_i - y_i|$, где $x = (x_1, \dots, x_k)$ и $y = (y_1, \dots, y_k)$. Замкнутый шар обозначим как $S_r(x) = \{y \in \mathbf{R}^k : \rho(x, y) \leq r\}$.

В дальнейшем всегда будет предполагаться, что *целевая функция* $f: \mathbf{R}^k \rightarrow \mathbf{R}$ ограничена сверху, измерима и принимает максимальное значение в единственной точке $x_0 = \arg \max \{f(x) : x \in \mathbf{R}^k\}$.

Случайным поиском называется произвольная (конечная или бесконечная) последовательность случайных величин $\{\xi_i\}_{i \geq 0}$ со значениями в $\mathbf{R}^k$. Если последовательность $\{\xi_i\}_{i \geq 0}$ образует марковскую цепь относительно потока σ -алгебр $\sigma(\xi_0, \dots, \xi_i)$, то поиск называется *марковским*, а если для любого $i > 0$ неравенство $f(\xi_i) \geq f(\xi_{i-1})$ выполняется с вероятностью 1, то поиск является *монотонным*. Далее будут изучаться марковские монотонные случайные поиски с переходными функциями специального вида. Опишем класс исследуемых поисков. Обозначив $B_x = \{y \in \mathbf{R}^k : f(y) \geq f(x)\}$, рассмотрим (вообще говоря, неоднородную) марковскую цепь $\{\xi_i\}_{i \geq 0}$ с начальной точкой $\xi_0 = x$ и переходными функциями

$$R_i(x, \cdot) = \delta_x(\cdot)P_i(x, B_x^c) + P_i(x, \cdot \cap B_x^c), \quad (1)$$

где через δ_x обозначено распределение, сосредоточенное в точке x , а B_x^c — дополнение множества B_x . Как обычно, $P_i(x, \cdot)$ при любых i и $x \in \mathbf{R}^k$ является вероятностной мерой, а $P_i(\cdot, A)$ для всех i и любого борелевского множества A является борелевской функцией в $\mathbf{R}^k$. Очевидно, что $R_i(x, B_x) = 1$, и, значит, неравенства $f(\xi_i) \geq f(\xi_{i-1})$ выполняются с вероятностью 1 при всех $i > 0$.

Запишем алгоритм моделирования n шагов описанного поиска. Обозначение « $\xi \leftarrow P(\cdot)$ » читается как «получить реализацию случайной величины ξ с распределением P ».

Алгоритм

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

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

Шаг 3. Если $f(\eta) \geq f(\xi_{i-1})$, 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.

Ниже для вероятностей событий и математических ожиданий случайных величин, связанных со случайным поиском алгоритма, начинающимся в точке $x \in \mathbf{R}^k$, используются обозначения $\mathbf{P}_x$ и $\mathbf{E}_x$.

Далее будем рассматривать *однородный* марковский монотонный случайный поиск, переходные функции которого определяются формулой (1), причем вероятностные меры $P(x, dy)$ не зависят от i и обладают *симметричной* плотностью $p(x, y)$ вида

$$p(x, y) = g(p(x, y)), \tag{2}$$

где g — невозрастающая неотрицательная функция, определенная на полуоси $(0, +\infty)$. Легко видеть, что тогда $p(x, x+y) = p(0, y)$ при всех $x, y \in \mathbf{R}^k$. Функцию g будем называть *формой поиска*, а также формой переходной плотности p . Чтобы функция $p(x, y)$, определенная в (2), была плотностью, форма поиска g должна удовлетворять условию нормировки

$$2^k k \int_{(0, +\infty)} g(r)r^{k-1} dr = 1. \tag{3}$$

Не умаляя общности будем считать, что функция g непрерывна слева.

Далее мы ограничимся формой поиска следующего вида. Пусть $h(r)$ — монотонно невозрастающая строго положительная функция, определенная на полуоси $(0, +\infty)$ и удовлетворяющая следующим двум условиям: а) $h(r)r^k \rightarrow 1$ при $r \rightarrow 0$, б) $h(r)r^{k-1}$ суммируема на промежутке $[1, +\infty)$. Не умаляя общности будем считать, что функция h непрерывна слева.

Зафиксируем $a > 0$ и положим при $\varepsilon > 0$

$$g_\varepsilon(r) = \frac{1}{\lambda(\varepsilon)} \begin{cases} h(a\varepsilon) & \text{при } 0 < r \leq a\varepsilon, \\ h(r) & \text{при } r > a\varepsilon, \end{cases} \tag{4}$$

где множитель $1/\lambda(\varepsilon)$ обеспечивает условие нормировки (3).

Цель и характеристики случайного поиска

Случайный поиск используем для отыскания точки максимума x_0 с заданной точностью $\varepsilon > 0$ (аппроксимация «по аргументу»). В этом случае нас интересует попадание поиска в шар $S_\varepsilon(x_0)$.

Положим в алгоритме $n = +\infty$ и обозначим

$$\tau_\varepsilon = \min \{i \geq 0 : \xi_i \in S_\varepsilon(x_0)\}.$$

Мы всегда будем предполагать, что для моделирования распределений P_i не требуется вычислений функции f . Тем самым на каждом шаге $\xi_{i-1} \mapsto \xi_i$ алгоритма происходит ровно одно вычисление целевой функции f , и распределение случайной величины τ_ε дает нам достаточно полную информацию о качестве случайного поиска. Действительно, при выполнении τ_ε шагов алгоритма значения функции f вычисляются $\tau_\varepsilon + 1$ раз.

Мы будем изучать две характеристики скорости сходимости случайного поиска. *Трудоёмкость* случайного поиска определяется как $E_x \tau_\varepsilon$ и имеет смысл среднего числа шагов поиска до достижения им множества $S_\varepsilon(x_0)$.

Вторая характеристика — *гарантирующее число шагов* — определяется как такое минимальное число $N = N(x, \varepsilon, \gamma)$ шагов поиска, при котором достижение множества $S_\varepsilon(x_0)$ гарантировано с вероятностью, не меньшей γ . Иначе говоря,

$$N(x, \varepsilon, \gamma) = \min\{i \geq 0 : P_x(\xi_i \in S_\varepsilon(x_0)) \geq \gamma\}.$$

Результаты

Приведенные результаты относятся к изучению зависимости величин $E_x \tau_\varepsilon$ и $N(x, \varepsilon, \gamma)$ при фиксированных x и γ от требуемой точности ε , при этом акцент делается на малые значения ε . Тем самым исследуются порядки роста $E_x \tau_\varepsilon$ и $N(x, \varepsilon, \gamma)$ при $\varepsilon \rightarrow 0$.

Конечно, на практике при вычислениях главную роль могут играть константы при соответствующих порядках. Хорошо известно, однако, что теоретическое нахождение этих констант — гораздо более трудная задача, чем определение точных порядков, а даже получение хороших оценок точных порядков в задачах глобальной оптимизации является сложной проблемой.

Полученные в [4, 5] оценки трудоёмкости сверху не могут дать лучшую по порядку зависимость от ε при малых ε , чем $O(\ln^2 \varepsilon)$. Этот наилучший порядок зависимости полученных оценок трудоёмкости от ε достигается для однородного марковского случайного поиска с формой плотности (4) и невырожденной (см. [4, 5]) целевой функции f . Кроме того, порядки зависимости от ε оценок трудоёмкости и гарантирующего числа шагов совпадают. При этом представленные в [4, 5] результаты являются именно оценками сверху, и общие утверждения о точности полученных оценок неизвестны. Ниже приводятся оценки этой точности для однородного марковского монотонного симметричного случайного поиска с формой плотности (4).

Теорема. Рассмотрим однородный марковский монотонный симметричный случайный поиск с формой плотности (4). Тогда существуют такие постоянные $c_1 > 0$ и c_2, c_3, c_4 , что при $\varepsilon < \rho(x, x_0)$ выполняются неравенства

$$E_x \tau_\varepsilon \geq c_1 (|\ln \varepsilon| + c_2) \left(\ln \left(\ln \frac{\rho(x, x_0)}{\varepsilon} + c_3 \right) + c_4 \right), \quad (5)$$

$$N(x, \varepsilon, \gamma) \geq \gamma c_1 (|\ln \varepsilon| + c_2) \left(\ln \left(\ln \frac{\rho(x, x_0)}{\varepsilon} + c_3 \right) + c_4 \right). \quad (6)$$

Неравенства (5) и (6) показывают, что полученные в [4, 5] оценки не далеки по порядку от истинной зависимости $E_x \tau_\varepsilon$ и $N(x, \varepsilon, \gamma)$ от ε при $\varepsilon \rightarrow 0$.

Таким образом, получены оценки трудоёмкости и гарантирующего числа шагов снизу.

-
1. Карманов В.Г. Математическое программирование. М.: Физматлит, 2000. 264 с.
 2. Немировский А.С., Юдин Д.Б. Сложность задач и эффективность методов оптимизации. М.: Наука, 1979. 384 с.
 3. Жиглявский А.А., Жилинскас А.Г. Методы поиска глобального экстремума. М.: Наука, 1991. 248 с.
 4. Nekrutkin V.V., Tikhomirov A.S. // Acta Appl. Math. 1993. V.33. P.89-108.
 1. Tikhomirov A.S. // Model-Oriented Data Analysis. Proceedings of the 3rd International Workshop in Petrodvorets, Russia. 25-30 May 1992. Heidelberg: Physica-Verlag, 1993. P.249-256.