

УДК 519.626

О ПЕРЕХОДНЫХ ФУНКЦИЯХ АЛГОРИТМА ОТЖИГА

А.С.Тихомиров

ON THE TRANSITION FUNCTIONS OF THE SIMULATED ANNEALING ALGORITHM

A.S.Tikhomirov

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

Рассмотрена задача построения оптимальных алгоритмов отжига и показано, что такие поиски, оптимальные в достаточно широком классе методов оптимизации, имеют простую структуру переходных функций.

Ключевые слова: случайный поиск, алгоритм отжига, глобальная оптимизация, стохастическая оптимизация

The problem of constructing optimal simulated annealing algorithms is considered. It is shown that such methods being optimal in a very wide sense have a simple structure of the transition functions.

Keywords: random search, simulated annealing algorithm, global optimization, stochastic optimization

1. Введение

Пусть целевая функция $f: \mathbf{R}^d \mapsto \mathbf{R}$ ограничена снизу и измерима. Рассмотрим задачу оценки минимального значения целевой функции f с заданной точностью ε (аппроксимация «по функции»). Один из способов решения этой задачи состоит в применении алгоритмов отжига (см. [1-14]). Алгоритмы отжига (simulated annealing) давно и успешно используются при решении сложных задач оптимизации и являются одними из самых знаменитых алгоритмов стохастической глобальной оптимизации. Тем не менее, существует мало теоретических результатов о скорости сходимости этих алгоритмов (см. [3-6]). Теоретическое исследование некоторых алгоритмов марковского случайного поиска экстремума выполнено в работах [11-16]. Данная работа является продолжением работ [15,16] и посвящена теоретическому исследованию свойств алгоритма отжига.

Построение конкретного варианта алгоритма отжига заключается в выборе переходных функций для получения новых точек в пространстве оптимизации

и задании вероятностей перехода поиска в эти новые точки. Данная работа посвящена обсуждению выбора переходных функций. Показано, что можно сузить семейство используемых переходных функций, не теряя свойства оптимальности в смысле минимальности числа шагов поиска, при котором достижение искомого множества гарантировано с заданной надежностью.

Результаты данной работы распространяют результаты статьи [15], полученные для монотонного марковского поиска, на другой класс методов случайного поиска — алгоритмы отжига. Кроме того, здесь рассмотрен другой, широко используемый на практике класс переходных функций, применяемый, в частности, Л.Ингбером в методе сверхбыстрого отжига (very fast annealing) см. [7,8].

2. Постановка задачи

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

$$\rho_\gamma(x, y) = \left(\sum_{n=1}^d |x_n - y_n|^\gamma \right)^{1/\gamma}, \quad \rho_\infty(x, y) = \max_{1 \leq n \leq d} |x_n - y_n|,$$

где $\gamma \geq 1$ — любое фиксированное число, $x = (x_1, \dots, x_d)$ и $y = (y_1, \dots, y_d)$.

Случайным поиском называется произвольная последовательность случайных величин $\{\xi_n\}_{n \geq 0}$ со значениями в $\mathbf{R}^d$. Следуя [17] приведем общую схему моделирования m шагов алгоритма отжига $\xi_0, \xi_1, \dots, \xi_m$.

Алгоритм 1

Шаг 1. $\xi_0 \leftarrow \pi(\cdot)$, $n \leftarrow 1$.

Шаг 2. $\eta_n \leftarrow P_n(\xi_{n-1}, \cdot)$.

Шаг 3. $\xi_n \leftarrow \begin{cases} \eta_n, & \text{с вероятностью } Q_n, \\ \xi_{n-1}, & \text{с вероятностью } 1 - Q_n, \end{cases}$

где

$$Q_n = \begin{cases} 1, & \text{если } \Delta_n \leq 0, \\ \exp(-\beta_n \Delta_n), & \text{если } \Delta_n > 0, \end{cases}$$

$\Delta_n = f(\eta_n) - f(\xi_{n-1})$, а величины $\beta_n \geq 0$ являются параметрами алгоритма.

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

Здесь $\pi(\cdot)$ — начальное распределение, m — число шагов поиска, а n — номер итерации алгоритма. Обозначение $\eta_n \leftarrow P_n(\xi_{n-1}, \cdot)$ читается как «получить реализацию случайной величины η_n с распределением $P_n(\xi_{n-1}, \cdot)$ ». Распределение $P_n(\xi_{n-1}, \cdot)$ зависит от номера шага n и «старой» точки поиска ξ_{n-1} . В соответствии со структурой алгоритма 1, распределения $P_n(\xi_{n-1}, \cdot)$ будем называть *пробными переходными функциями*, а случайные величины η_n — *пробными точками*.

После получения новой пробной точки η_n (на втором шаге алгоритма) на третьем шаге поиск или переходит в эту точку η_n с вероятностью Q_n , или остается в старой точке поиска ξ_{n-1} .

Случайный поиск используем для оценки минимального значения целевой функции f с заданной точностью ε (аппроксимация «по функции»). При аппроксимации по функции нас будет интересовать попадание поиска в множество

$$A_\varepsilon = \{x \in \mathbf{R}^d : f(x) \leq \inf f + \varepsilon\}.$$

Может, однако, случиться так, что поиск $\{\xi_n\}_{n \geq 0}$, оказавшись в множестве A_ε на шаге n , выйдет из A_ε на одном из последующих шагов. Чтобы избежать анализа таких эффектов, введем величины $\xi_n^* = \operatorname{argmin} \{f(\xi_0), \dots, f(\xi_n)\}$. Будем считать, что $\operatorname{argmin} \{f(\xi_0), \dots, f(\xi_n)\} = \xi_j$, где $j = \max \{k \in \{0, \dots, n\} : f(\xi_k) = \min \{f(\xi_0), \dots, f(\xi_n)\}\}$. Случайная величина ξ_n^* , попав в множество A_ε , из него больше не выйдет.

Мы рассмотрим одну характеристику скорости сходимости случайного поиска. *Гарантирующее число шагов* $N(f, \varepsilon, \gamma)$ определяется как такое минимальное число шагов поиска, при котором достижение множества A_ε гарантировано с вероятностью большей, чем γ . Иначе говоря,

$$N(f, \varepsilon, \gamma) = \min \{n \geq 0 : \mathbf{P}(\xi_n^* \in A_\varepsilon) > \gamma\}.$$

Полагаем $N(f, \varepsilon, \gamma) = +\infty$ в случае, когда $\mathbf{P}(\xi_n^* \in A_\varepsilon) \leq \gamma$ при всех $n \geq 0$. Вероятность $\gamma \in (0, 1)$ будем называть *надежностью*.

Мы рассмотрим марковские алгоритмы случайного поиска, переходные функции $P_n(x, \cdot)$ которых обладают плотностями $p_n(x, y)$ вида

$$p_n(x, y) = \prod_{k=1}^d p_{n,k}(x_k, y_k) = \prod_{k=1}^d g_{n,k}(|x_k - y_k|), \quad (1)$$

где $x = (x_1, \dots, x_d)$ и $y = (y_1, \dots, y_d)$, $p_{n,k}$ — плотности в одномерном пространстве $\mathbf{R}$, а $g_{n,k}$ — невозрастающие неотрицательные функции, определенные на множестве $(0, +\infty)$. Не умаляя общности будем считать, что функции $g_{n,k}$ непрерывны слева. Функцию $g_{n,k}$ будем называть *формой плотности* $p_{n,k}$.

Пусть $\mathbf{D}$ — множество всех переходных функций с плотностями вида (1). Переходные функции такого вида широко используются на практике (см. [7, 8, 12]), и применяются, в частности, в методе сверхбыстрого отжига Ингбера.

В случае одномерного пространства $\mathbf{R}$ простейшим из таких распределений является равномерное распределение $U_a(x, \cdot)$ в шаре радиуса $a > 0$ с центром в точке $x \in \mathbf{R}$. Форма $g = g^{(a)}$ плотности такого распределения имеет вид

$$g^{(a)}(r) = \frac{1}{2a} \begin{cases} 1, & \text{если } 0 < r \leq a, \\ 0, & \text{если } r > a. \end{cases} \quad (2)$$

Рассмотрим теперь пространство $\mathbf{R}^d$. Пусть $a = (a_1, \dots, a_d)$ и все $a_k > 0$. Через U_a обозначим переходную функцию с плотностью

$$p^{(a)}(x, y) = \prod_{k=1}^d g^{(a_k)}(|x_k - y_k|), \quad (3)$$

где $x = (x_1, \dots, x_d)$ и $y = (y_1, \dots, y_d)$, а функции $g^{(a_k)}$ задаются формулой (2). Здесь $U_a(x, \cdot)$ — это равномерное распределение в d -мерном прямоугольном параллелепипеде с центром в точке $x \in \mathbf{R}^d$ и сторонами $2a_k$ при $k = 1, \dots, d$. Пусть Π — множество всех переходных функций с плотностями вида (3). Ясно, что $\Pi \subset \mathbf{D}$.

3. Оптимальность простых поисков

Оказалось, что можно сузить семейство используемых переходных функций, не теряя свойства оптимальности в смысле минимальности числа шагов

поиска, при котором достижение множества A_ε гарантировано с заданной надежностью.

Теорема 1. Пусть целевая функция f , точность $\varepsilon > 0$, надежность $\gamma \in (0, 1)$, начальное распределение π и параметры $\beta_n \geq 0$ вероятностей перехода в новые пробные точки фиксированы. Тогда справедливо равенство

$$\min \{N(f, \varepsilon, \gamma) : P_n \in \mathbf{D}\} = \min \{N(f, \varepsilon, \gamma) : P_n \in \mathbf{II}\}. \quad (4)$$

Равенство (4) показывает, что минимальное значение гарантирующего числа шагов при использовании переходных функций из множества $\mathbf{D}$ совпадает с минимальным значением гарантирующего числа шагов при использовании переходных функций из множества $\mathbf{II}$. Таким образом доказано, что переходные функции оптимального алгоритма отжига (т.е. поиска с минимальным значением гарантирующего числа шагов) имеют простую структуру. В качестве таких переходных функций можно использовать равномерные распределения в d -мерных прямоугольных параллелепипедах.

Полученный теоретический результат имеет ясное прикладное значение, так как обосновывает выбор вида поиска, рекомендованного в ряде работ (как правило, прежнее обоснование было либо эмпирическим, либо основывалось на соображениях «простоты»).

1. Ермаков С.М., Жиглявский А.А. О случайном поиске глобального экстремума // Теория вероятностей и ее применения. 1983. №1. С.129-136.
2. Ермаков С.М., Жиглявский А.А., Кондратович М.В. О сравнении некоторых процедур случайного поиска глобального экстремума // Журн. 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. New Jersey: Wiley, 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. Yin G. Rates of convergence for a class of global stochastic optimization algorithms // SIAM Journal on Optimization. 1999. V.10. №1. P.99-120.
7. Ingber L. Very fast simulated re-annealing // Math. Comput. Modelling. 1989. V.12. P.967-973.
8. Лопатин А.С. Метод отжига // Стохастическая оптимизация в информатике. 2005. Вып. 1. С.133-149.
9. Граничин О.Н., Поляк Б.Т. Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003. 291 с.
10. Абакаров А.Ш., Сушков Ю.А. Статистическое исследование случайного поиска // Математические модели. Теория и приложения. Вып.2. СПб.: Изд-во НИИХ СПбГУ, 2002. С.70-86.
11. Тихомиров А.С., Некруткин В.В. Марковский монотонный поиск экстремума. Обзор некоторых теоретических результатов // Математические модели. Теория и приложения. Вып.4. СПб.: БВМ, 2004. С.3-47.
12. Тихомиров А.С. О быстрых вариантах алгоритма отжига (simulated annealing) // Стохастическая оптимизация в информатике. 2009. Вып.5. С.65-90.
13. Тихомиров А.С. О скорости сходимости алгоритма simulated annealing // Журн. вычисл. математики и мат. физики. 2010. Т.50. №1. С.24-37.

14. Тихомиров А.С. Нижние оценки скорости сходимости марковского симметричного случайного поиска // Журн. вычисл. математики и мат. физики. 2011. Т.51. №9. С.1630-1644.
15. Тихомиров А.С. Об оптимальном марковском монотонном симметричном случайном поиске // Журн. вычисл. математики и мат. физики. 1998. Т.38. №12. С.1973-1982.
16. Тихомиров А.С. Об оптимальном марковском случайном поиске // Вестник НовГУ. Сер. Естеств. и техн. науки. 1999. №13. С.109-112.
17. Zhigljavsky A., Zilinskas A. Op. cit. P.118.

References

1. Ermakov S.M., Zhigljavskii A.A. O sluchainom poiske global'nogo ekstremuma [On the Random Search of Global Extremum]. Teoriia veroiatnostei i ee primeneniia – Theory of Probability and its Applications, 1983, no. 1, pp. 129–136.
2. Ermakov S.M., Zhigljavskii A.A., Kondratovich M.V. O sravnenii nekotorykh protsedur sluchainogo poiska global'nogo ekstremuma [Comparison of some random search procedures for a global extremum]. Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki – Computational Mathematics and Mathematical Physics, 1989, vol. 29, no. 2, pp. 112–117.
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. New Jersey, Wiley, 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. London, Springer, 2006, pp. 99–117.
6. Yin G. Rates of convergence for a class of global stochastic optimization algorithms. SIAM Journal on Optimization, 1999, vol. 10, no. 1, pp. 99–120.
7. Ingber L. Very fast simulated re-annealing. Mathl. Comput. Modelling, 1989, v. 12. pp. 967–973.
8. Lopatin A.S. Metod otzhiga [Simulated Annealing Method]. Stokhasticheskaia optimizatsiia v informatike, 2005, no. 1, pp. 133–149.
9. Granichin O.N., Poliak B.T. Randomizirovannye algoritmy otsenivaniia i optimizatsii pri pochte proizvol'nykh pomekakh [Randomized estimation and optimization algorithms in “almost arbitrary” noise]. Moscow, “Nauka” Publ., 2003. 291 p.
10. Abakarov A.Sh., Sushkov Iu.A. Statisticheskoe issledovanie sluchainogo poiska [Statistical Investigation of Random Search]. Matematicheskie modeli. Teoriia i prilozheniia [Mathematical Modeling: Theory and Applications], Research Studies Institute of Chemistry, St. Petersburg State University, St. Petersburg, 2002, no. 2, pp. 70–86.
11. Tikhomirov A.S., Nekrutkin V.V., Markovskii monotonnii poisk ekstremuma. Obzor nekotorykh teoreticheskikh rezul'tatov [Markov Monotone Search for Extrema: Survey of Some Theoretic Results]. Matematicheskie modeli. Teoriia i prilozheniia [Mathematical Modeling: Theory and Applications], VVM Publ., St. Petersburg, 2004, no. 4, pp. 3–47.
12. Tikhomirov A.S. O bystrykh variantakh algoritma otzhiga (simulated annealing) [On some variants of the fast simulated annealing algorithms]. Stokhasticheskaia optimizatsiia v informatike, 2009, no. 5, pp. 65–90.
13. Tikhomirov A.S. Tikhomirov A.S. O skorosti skhodimosti algoritma simulated annealing [On the convergence rate of the simulated annealing algorithm]. Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki – Computational Mathematics and Mathematical Physics, 2010, vol. 50, no. 1, pp. 19–31.
14. Tikhomirov A.S. Nizhnie otsenki skorosti skhodimosti markovskogo simmetrichnogo sluchainogo poiska [Lower bounds on the convergence rate of the Markov symmetric random search]. Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki – Computational Mathematics and Mathematical Physics, 2011, vol. 51, no. 9, pp. 1524–1538.
15. Tikhomirov A.S. Ob optimal'nom markovskom monotonnom simmetrichnom sluchainom poiske [Optimal Markov monotone symmetric random search]. Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki – Computational Mathematics and Mathematical Physics, 1998, vol. 38, no. 12, pp. 1894–1902.
16. Tikhomirov A.S. Ob optimal'nom markovskom sluchainom poiske [On optimal Markov random search]. Vestnik of Novgorod State University. Ser.: Estestv. i tekhn. nauki – Vestnik NovSU. Issue: Natural and Engineering Sciences, 1999, no. 13, pp. 109–112.
17. Zhigljavsky A., Zilinskas A. Op. cit., p.118.