В.Б. Бериков, В.М. Неделько

Доктора технических наук

Институт математики СО РАН, г. Новосибирск, Россия

Исследование эффективности адаптивных алгоритмов многоэкстремальной оптимизации*

1. Введение

Задача поиска глобального экстремума функций имеет большую практическую важность. В настоящее время существует ряд методов решения указанной задачи, основанных на различных подходах: эволюционном, «мультистарт», алгоритмы «муравьиной колонии» и т.д. [1,2]. Пионерскими работами, в которых были заложены основные идеи метода случайного поиска с адаптацией (СПА), являются работы Г.С. Лбова [3,4]. Несколько позднее, с использованием аналогичного подхода, были разработаны «генетические» алгоритмы многоэкстремальной оптимизации.

При дальнейшем развитии методов случайного адаптивного поиска возникли сложные теоретические вопросы, касающиеся обоснования их эффективности [5,6]. Одной из наиболее важных задач является установление взаимосвязи между сложностью класса многоэкстремальных функций и свойствами алгоритма оптимизации, которыми он должен обладать для достижения решения с приемлемым качеством, при заданных ограничениях на используемые ресурсы (число испытаний и т.п.).

В работе предлагается подход к решению поставленной задачи, основанный на задании специального вида меры сложности функции.

2. Задача поиска глобального экстремума

Сформулируем задачу оптимизации [7].

Пусть Ψ – некоторый класс функций с общими областью определения Z и областью значений R, т.е. $\psi : Z \to R$, $\psi \in \Psi$. Множество R упорядочено.

Алгоритмом (детерминированным алгоритмом) поиска экстремума будем называть множество функций

$$A = \{ A_i | \ i = 1,..., L \},$$

где $A_i : (Z \times R)^{i-1} \to Z$, при этом A_1 – константа (из Z). Параметр L определяет число точек, которые можно использовать при поиске экстремума.

Результатом применения алгоритма A к функции ψ будет последовательность $A|\psi = \left((z_i, r_i) | \ i = 1,..., L \right)$, где $z_i = A_i(z_1, r_1; z_2, r_2; ...; z_{i-1}, r_{i-1};)$, $r_i = \psi(z_i)$.

Таким образом, первая точка фиксирована, а остальные ставятся в зависимости от полученных значений оптимизируемой функции.

Стохастическим (случайным) алгоритмом поиска экстремума будем называть совокупность семейств вероятностных мер

$$P = \{ P_i[Z/\alpha_i] | \ i = 1,..., L \}, \ \alpha_i \in (Z \times R)^{i-1}.$$

Все меры заданы на некоторой σ-алгебре подмножеств пространства Z.

Результатом применения алгоритма P к функции ψ будет вероятностная мера $P|\psi = P_L^{\psi}[Z^L]$, которую можно определить рекуррентным соотношением

$$P_i^{\psi}[Z^i] = P_{i-1}^{\psi}[Z^{i-1}] \times P_i[Z/\alpha_i], \ \alpha_i = (z_1, r_1; ...; z_i, r_i;), \ r_i = \psi(z_i).$$

Таким образом, первая точка выбирается в соответствии с фиксированным вероятностным распределением, для остальных вероятностная мера корректируется в зависимости от полученных значений оптимизируемой функции.

Чтобы задача поиска глобального экстремума была поставленной, нужно задать критерий качества алгоритма оптимизации, как функцию $K:\Psi \times \Lambda \to [1,\infty)$, где Λ – множество всех детерминированных и стохастических алгоритмов. В настоящее время такая задача решена только для некоторых частных постановок. Задача сравнения эффективности алгоритмов оптимизации и выбора алгоритма, наиболее подходящего для заданного класса задач, остается открытой.

3. Алгоритм СПА

Алгоритм СПА является представителем стохастических алгоритмов оптимизации. В этом алгоритме все меры $P_i[Z/\alpha_i]$, $i = 2,..., L$, являются дифференцируемыми по мере $P_1[Z]$, поэтому их можно задавать через плотности $\varphi_i(z/\alpha_i) = \frac{dP[Z/\alpha_i]}{dP_1[Z]}$. Здесь $\frac{dP[Z/\alpha_i]}{dP_1[Z]} \equiv \frac{P[dz/\alpha_i]}{P_1[dz]}$ обозначает производную по мере.

Для использования алгоритма СПА необходимо чтобы на точках из Z была определена мера сходства $\rho : Z^2 \to [0,\infty)$, которую будем для удобства называть расстоянием, при этом не подразумевая обязательного выполнения свойств метрики.

Идея алгоритма СПА состоит в том, что плотности задаются так, чтобы их значения были больше в окрестностях тех из уже поставленных точек, в которых значения целевой функции ближе к оптимальному, а также в тех областях, где плотность точек (экспериментов) ниже (в «малоизученных областях»). При этом данные плотности зависят от числа уже проведенных экспериментов.

Эта идея может реализовываться (конкретизироваться) различными способами. В настоящей работе используется следующая реализация алгоритма СПА.

Пусть на $(i+1)$–м шаге уже выбрано i точек $Z_i = \{z_1,..., z_i\}$ и получено i значений оптимизируемой функции $r_1,...,r_i$, то есть определен набор параметров $\alpha_i = (z_1, r_1;...; z_i, r_i)$. Требуется задать плотность вероятности $\varphi_i(z/\alpha_i)$, в соответствии с которой будет поставлена следующая точка z_{i+1}.

Обозначим через $\upsilon(z) = \arg\min\limits_{z' \in Z_i} \rho(z', z)$ – функцию, сопоставляющую точке $z \in Z$ ближайшую точку из Z_i – выбранных на предыдущих шагах алгоритма. Если минимум расстояния достигается для более, чем одной, точки, то значением функции $\upsilon(z)$ полагаем точку с меньшим индексом.

Для произвольной точки $z \in Z$ определим величины:

$$R_\psi(z) = \tfrac{1}{i} \cdot \left| \{ z' \in Z_i \mid \psi(z) \geq \psi(\upsilon(z')) \} \right|$$

– относительный ранг ближайшего известного значения функции (доля точек из Z_i, где значение функции ψ «хуже», чем в ближайшей к z) и

$$R_\rho(z) = P_1(\rho(z, \upsilon(z)) \geq \rho(z', \upsilon(z')))$$

– относительный «ранг» расстояния до ближайшей точки, который равен вероятности того, что случайно выбранная точка из Z окажется ближе к множеству Z_i, чем заданная точка.

Теперь, чтобы выразить идею адаптации, достаточно задать плотность $\varphi_i(z/\alpha_i)$ как монотонно возрастающую по обоим аргументам функцию

$$\varphi_i(z/\alpha_i) = \zeta(R_\psi(z), R_\rho(z)).$$

Простейшим вариантом такой функции может выступать

$$\zeta(R_\psi(z), R_\rho(z)) = \begin{cases} 0, & R_\psi(z) + R_\rho(z) < \beta \\ b, & \text{else}, \end{cases}$$

β – заданное пороговое значение, а константа b определяется условием нормировки вероятностной меры. В этом случае на каждом шаге алгоритма выделяется область, куда текущая точка поставлена быть не может, а вовне этой области точка ставиться в соответствии с равномерной плотностью вероятности.

Открытым является вопрос формулирования критерия, определяющего, насколько та или иная функция (класс функций) подходит для оптимизации методом СПА. В частности, ограничение на константу Липшица не является необходимым условием применимости алгоритма.

В качестве модельного примера для тестирования приведенной реализации алгоритма СПА использована задача поиска глобального экстремума функции

$$\psi_1(z)=\psi_1(x_1,x_2)=e^{-15(x_1-0,4)^2-12(x_2-0,7)^2}+0,7e^{-12(x_1-0,85)^2-7(x_2-0,3)^2}$$

на интервале $Z=X_1\times X_2$, $X_1=[0,1]$, $X_2=[0,1]$.

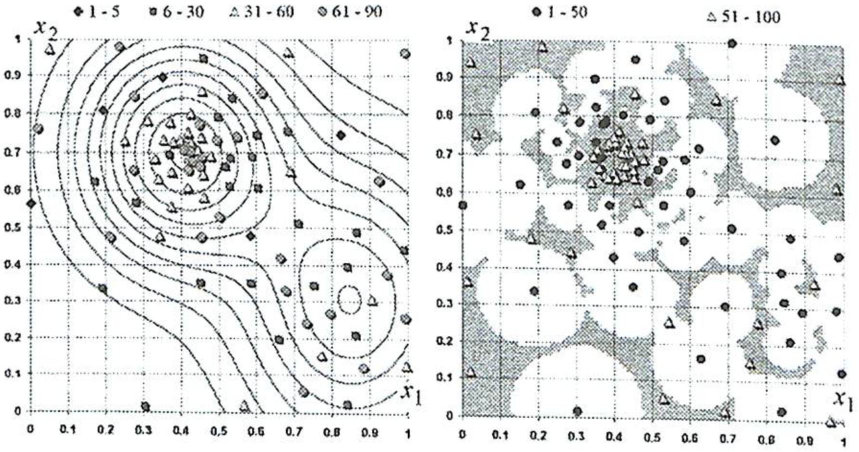

Рис. 1. Последовательность случайной адаптивной расстановки точек для тестовой функции.

На рис. 1 приведена последовательность расстановки точек описанной реализацией (при $\beta=1$) алгоритма СПА для данной тестовой функции. Левая диаграмма отражает последовательность случайной адаптивной расстановки точек. Оптимизируемая функция изображена линиями уровней. На правой диаграмме темным тоном показана область разрешенных положений следующей точки после постановки серии из 50-ти точек. Как можно видеть, расстановка точек достаточно быстро концентрируется вблизи глобального экстремума функции.

4. О связи множества функций и эффективности алгоритма

Под эффективностью метода (алгоритма) поиска глобального экстремума будем понимать некоторый функционал $K(f)$, характеризующий точность нахождения экстремума.

Из NFL-теоремы можно сделать практический вывод о том, что не существует безусловно наилучшего метода многоэкстремальной оптимизации, и обосновать полезность метода можно только обосновав разумность класса функций, на которые он ориентирован.

Если задан класс функций, можно ставить задачу поиска в некотором смысле (например, минимаксном или в среднем, если на функциях задана вероятностная мера) оптимального метода поиска экстремума. В этом случае алгоритм АДАПТ (или СПА) может использоваться для метаоптимизации, т.е. для поиска наилучших значений параметров алгоритма поиска экстремума. Заметим, что в случае поиска экстремума дискретной функции в дискретном пространстве переменных вообще все алгоритмы оптимизации (при заданном числе испытаний) могут быть заданы конечным числом параметров. В этом случае СПА может использоваться для поиска оптимального в среднем алгоритма оптимизации.

Описание класса функций, на которых заданный алгоритм достигает заданного качества оптимизации, является нетривиальной задачей, решённой только для простых частных случаев.

Для описания класса функций, эффективно оптимизируемых алгоритмом АДАПТ, Г.С. Лбовым был предложен подход, основанный на введении меры сложности функции.

Мерой сложности $C(f)$ функции f будем называть некоторую числовую характеристику функции, которая монотонно связана с эффективностью (точностью нахождения экстремума) алгоритма на этой функции, т.е. $C(f_1)\geq C(f_2)\Leftrightarrow K(f_1)\geq K(f_2)$ для любых f_1 и f_2.

На рис. 2 приведены примеры «простой» и «сложной» для поиска экстремума функций.

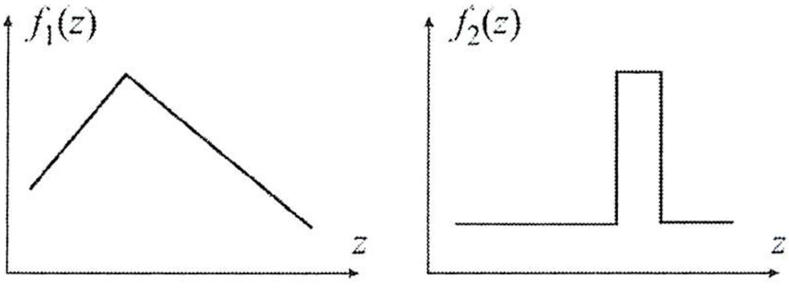

Рис. 2. Примеры «простой» и «сложной» для поиска экстремума функций.

Мера сложности функции может подбираться экспериментально на основе моделирования. В одномерном случае один из очевидных вариантов — функционал вида

$$C(f)=\mathrm{E}\left(f(z)-f^*\right)\cdot\left|z-z^*\right|-\mathrm{E}\left(f(z)-f^*\right)\cdot\mathrm{E}\left|z-z^*\right|,$$

где f^* – значение глобального экстремума.

В многомерном случае (в том числе для разнотипного пространства переменных) аналогичный функционал может быть введён для функций, задаваемых деревом решений, если в качестве «расстояния» между точками пространства взять длину цепи, соединяющей соответствующие вершины.

Заключение

Предложен вариант меры сложности многоэкстремальной функции как величины, характеризующий ожидаемое качество решения оптимизационной задачи методом ADAPT.

Предложенная в работе модификация метода случайного поиска с адаптацией (СПА) позволяет осуществлять поиск глобального экстремума функций, определенных на произвольном метрическом пространстве с вероятностной мерой, способ задания которой допускает эффективную алгоритмическую реализацию процедуры случайного выбора точки. При этом не требуется, чтобы пространство было евклидовым или представлялось в виде декартова произведения множеств значений некоторых характеристик (признаков). Предложенный метод оптимизации полностью сохраняет принципы адаптивной расстановки точек, заложенные в методе СПА.

Проведенное исследование на модельных примерах задач классификации показывает принципиальную пригодность метода СПА для поиска оптимального дерева решений. В рассмотренных примерах метод продемонстрировал существенно большую скорость приближения в оптимальному решению, по сравнению с методом неадаптивного случайного поиска. Дальнейшее улучшение алгоритмов построения логической решающей функции может быть достигнуто за счет комбинирования методов направленного и адаптивного поиска.

Список литературы

1. Weise T. Global Optimization Algorithms – Theory and Application: Ph.D. thesis / University of Kassel.- 2008.
2. C. Voglis, I.E. Lagaris Towards ''Ideal Multistart". A stochastic approach for locating the minima of a continuous function inside a bounded domain
3. Лбов Г.С. Выбор эффективной системы зависимых признаков. // Выч. системы, вып. 19, Новосибирск, 1965, с. 21–34.
4. Лбов Г.С. Методы обработки разнотипных экспериментальных данных. — Новосибирск: Наука, 1981.
5. Лбов Г.С., Неделько В.М., Неделько С.В. Метод адаптивного поиска логической решающей функции. // Сибирский журнал индустриальной математики, 2009, т. XII, № 3 (39). С. 66–74.
6. Неделько С.В., Неделько В.М. Исследование связи характеристик алгоритма поиска глобального экстремума и класса функций в дискретном пространстве малой мощности. // Искусственный интеллект (ISSN 1561-5359), №2, 2006. С. 201-205.
7. Сухарев А.Г. Оптимальный поиск экстремума. М.:Изд-во МГУ. 1975.

Resume

In this paper, we research the efficiency of optimization algorithms, in particular for revealing the relations between algorithm parameters, a class of multi-extreme functions and number of tests. For studying the relation between optimization algorithm and a class of multi-extreme functions, a formal statement of the problem of finding the optimum algorithm for multi-extreme optimization is given, under condition of fixed number of tests and a fixed class of multi-extreme functions. The method for searching the algorithm for extremum finding, optimum on average, is suggested.

Түйін

Мақалада функцияның экстремиумын табуда кездейсоқ адаптивті іздеу әдісіне негізделген тәсіл ұсынылған. Мұнда функция күрделілігі арнайы түрде берілген. Зерттеу нәтижесінде классификациялау есептері үшін тиімді шешімдер ағашын табу мақсатында ұсынылған әдістің түбегейлі жарамды екендігі көрсетілген. Ұсынылған әдістің тиімді шешімге жету жылдамдығы өте жоғары.

Özet

Bu çalışmada, algoritma parametreleri, çok aşırı fonksiyonlar bir sınıf ve test sayısı arasındaki ilişkileri ifşa ettiği için, özellikle verimlilik optimizasyon algoritmaları araştırma. Optimizasyon algoritması ve çok aşırı fonksiyonlar bir sınıf arasındaki ilişkiyi incelemek için, çok aşırı optimizasyonu için en uygun algoritma bulma sorunu biçimsel bir bildirimde testleri sabit sayıda koşul ve multi-aşırı sabit bir sınıf altında verilir fonksiyonları. Algoritması ortalamaoptimum uç bulmak, aramak için bir yöntem önerilmektedir.