Поиск оптимальной вероятности мутации для решения задачи XdivK

Антипов Д.С., аспирант кафедры компьютерных технологий Университета ИТМО, antipovden@yandex.ru

Буздалов М. В., к.т.н., доцент кафедры компьютерных технологий Университета ИТМО, mbuzdalov@gmail.com

Аннотация

Модельные задачи часто помогают изучить поведение вероятностных алгоритмов при различных условиях. Задача XdivK позволяет понять поведение алгоритма на различных плато, а также найти оптимальные параметры алгоритма для скорейшего прохождения плато.

В данной работе рассматривается теоретическая оценка оптимальной вероятности мутации для (1+1)-стратегии при решении задачи XdivK.

Введение

Теоретический анализ времени работы эволюционных алгоритмов [1] часто помогает найти слабые места эволюционных алгоритмов, или, наоборот, найти оптимальный набор параметров алгоритма, при которых он демонстрирует максимальную производительность [5, 6]. Особый интерес при теоретическом анализе эволюционных алгоритмов представляет их способность проходить плато и локальные оптимумы целевых функций [2, 8, 4].

В данной работе рассматривается оптимизация функции XdivK, особеностью которой являются множественные плато в пространстве поиска [3]. В качестве алгоритма оптимизации в данной работе рассмотрена $(1 + 1)$-стратегия [7]. Данный алгоритм имеет один параметр – вероятность мутации. Целью данной работы является нахождение оптимальной вероятности мутации в следующих смыслах:

- Оптимальная вероятность мутации, минимизирующая ожидаемое время прохождения одного фиксированного плато.
- Оптимальная вероятность мутации, минимизирующая полное ожидаемое время работы алгоритма.

Постановка задачи

В данной работе рассматривется оптимизация функции XdivK $(1+1)$-стратегией.

Функция XdivK задана над пространством двоичных векторов фиксированной длины n. Определить ее проще всего через другую функцию, заданную над тем же пространством поиска – OneMax(x), которая возвращает число единиц в аргументе [3]. Тогда XdivK$(x) = \lfloor$OneMax$(x)/K\rfloor$, где K – это параметр алгоритма, характеризующий размер плато в пространстве поиска.

В качестве алгоритма оптимизации рассматривается $(1 + 1)$-стратегия. Данный алгоритм оптимизации работает следующим образом. Сначала генерируется случайный двоичный вектор. Затем на каждой итерации алгоритма создается новый двоичный вектор, который является мутацией текущего двоичного вектора. Если значение целевой функции на вновь созданном двоичном векторе не меньше, чем на текущем, тогда новый вектор заменят текущий.

Используется следующий оператор мутации: каждый бит двоичного вектора меняет свое значение с вероятностью $\frac{\lambda}{n}$, где λ – это некоторый параметр данного алгоритма.

В данной работе поставлена задача нахождения оптимального значения λ в зависимости от величины плато K

Решение для последнего уровня

Для последнего уровня (то есть для той области пространства поиска, в которой все векторы имеют хотя бы $n-k$ единиц) было выдвинуто предположение, что асимптотически оптимальная вероятность мутации $\lambda = \sqrt[k]{k!}$.

Данное предположение доказано для $k = 2$ и $k = 3$. Рассмотрим доказательство для $k = 2$. Для этого составим марковскую цепь, в которой будет три состояния: состояние 0 – конечное и состояния 1 и 2, в которых текущая особь имеет ровно один и ровно 2 нуля соответственно. Пусть p_i^j – вероятность совершить переход из состояния i в состояние j. В случае , если $i > j$, то эта вероятнсоть складывается из вероятности переключить $i - j$ бит, находящихся в сотоянии “ноль” и, если возможно, вероятности переключить $i - j + 1$ “нулевых” бит и 1 “единичный” бит. Для $i < j$ можно привести аналогичное рассуждение. Таким образом:

$$p_2^0 = \left(\frac{\lambda}{n}\right)^2 \left(1 - \frac{\lambda}{n}\right)^{n-2}$$
$$p_2^1 = 2\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)^{n-1} + (n-2)\left(\frac{\lambda}{n}\right)^3 \left(1 - \frac{\lambda}{n}\right)^{n-3}$$
$$p_2^2 = 1 - p_2^1 - p_2^0$$
$$p_1^0 = \frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)^{n-1}$$
$$p_1^1 = 1 - p_1^2 - p_1^0$$
$$p_1^2 = (n-1)\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)^{n-1} + \frac{(n-1)(n-2)}{2}\left(\frac{\lambda}{n}\right)^3 \left(1 - \frac{\lambda}{n}\right)^{n-3}$$

Теперь, чтобы узнать ожидамое число итераций, прежде чем алгоритм закончит работу (то есть перейдет в состояние 0), можно составить систему из двух линейных уравнений. E_1 и E_2 – ожидаемые числа итераций, которые потребуются алгоритму, если он находится в состоянии 1 и 2 соответственно:

$$E_1 = 1 + p_1^1 E_1 + p_1^2 E_2$$
$$E_2 = 1 + p_2^1 E_1 + p_2^2 E_2$$

После решения этой системы получаем, асимптотическую формулу, которая одинакова дл E_1 и E_2:

$$E_1, E_2 = \frac{n^2}{e^{-\lambda}(2\lambda + \lambda^2 + o(1))}$$

Для минимизации данной формулы требуется максимизировать знаменатель. Знаменатель является непрерывно дифференцируемой функцией по λ при $\lambda > 0$. Производная знаменателя имеет единственный ноль при $\lambda = \sqrt{2}$. Таким образом, легко найти максимум знаменателя, сравнив значения при $\lambda = 0$, $\sqrt{2}$ и $+\infty$. Максимальное значение будет при $\lambda = \sqrt{2}$, следовательно, это значение λ является оптимальным для $k = 2$.

В случае $k = 3$ можно получить аналогичную систему из трех линейных уравнений, при решении которой E_1, E_2 и E_3 также будут иметь одинаковые асимптотические формулы:

$$E_1, E_2, E_3 = \frac{n^3}{e^{-\lambda}(6\lambda + 3\lambda^2 + \lambda^3 + o(1))}$$

Данная формула минимизируется при $\lambda = \sqrt[3]{6}$, что можно доказать рассуждениями, аналогичными для случая $k = 2$

Полученные результаты наталкивают на следующее предположение. В общем виде все E_i имеют вид $\frac{n^k}{k!e^{-\lambda}(\sum\limits_{i=1}^{k}\frac{\lambda^i}{i!}+o(1))}$. Если данное предположение верно, тогда все E_i минимизируются значением $\lambda = \sqrt[k]{k!}$.

Результаты экспериментов

Для больших значений k было получено численное значение оптимального значения вероятности мутации λ для $n = 100$. Сравнение численных результатов со значением $\sqrt[k]{k!}$ представлено в Таблице 1.

k	Оптимальное значение λ	$\sqrt[k]{k!}$
2	1.419	1.414
3	1.825	1.817
4	2.224	2.213
5	2.620	2.605
6	3.013	2.994
7	3.404	3.380
8	3.798	3.764
9	4.186	4.147
10	4.539	4.529

Таблица 1: Сравнение оптимальных значенией λ со значением $\sqrt[k]{k!}$ для значений k от 2 до 10

Заключение

В данной работе были сделаны первые шаги по нахождению оптимальной вероятности мутации для $(1 + 1)$-стратегии при решении задачи XdivK. Был получен предварительный результат для последнего уровня. В дальнейшем планируется рассмотреть другие плато, находящиеся дальше от оптимума, а также постараться найти вероятность мутации, которая будет оптимальна в глобальном смысле.

Литература

[1] Ю.А. Скобцов. Основы эволюционных вычислений. // ДонНТУ, Донецк, 2008.

[2] Brockhoff, D., Friedrich, T., Hebbinghaus, N., Klein, C., Neumann, F., Zitzler, E.: On the effects of adding objectives to plateau functions. // IEEE Transactions on Evolutionary Computation 13(3), 591–603 (2009)

[3] Buzdalov, M., Buzdalova, A.: OneMax helps optimizing XdivK: Theoretical runtime analysis for RLS and EA+RL. // Proceedings of Genetic and Evolutionary Computation Conference Companion. pp. 201–202. ACM (2014)

[4] Dang, D.C., Friedrich, T., Kötzing, T., Krejca, M.S., Lehre, P.K., Oliveto, P.S., Sudholt, D., Sutton, A.M.: Escaping local optima with diversity mechanisms and crossover. // Proceedings of Genetic and Evolutionary Computation Conference. pp. 645–652 (2016)

[5] Doerr, B., Doerr, C., Ebel, F.: Lessons from the black-box: fast crossover-based genetic algorithms. // Proceedings of Genetic and Evolutionary Computation Conference. pp. 781–788 (2013)

[6] Doerr, B., Doerr, C., Ebel, F.: From black-box complexity to designing new genetic algorithms. // Theoretical Computer Science 567, 87–104 (2015)

[7] Droste, S., Jansen, T., Wegener, I.: On the analysis of the (1+1) evolutionary algorithm. // Theor. Comput. Sci. 276(1-2), 51–81 (2002)

[8] Sudholt, D.: Memetic algorithms with variable-depth search to overcome local optima. // Genetic and Evolutionary Computation Conference, GECCO 2008, Proceedings, Atlanta, GA, USA, July 12-16, 2008. pp. 787–794 (2008)