Parallel Multi-objective Optimization on CPU Using Information Framework for Constructing Global Optimization Algorithms

Vladislav V. Sovrasov

State University of Nizhny Novgorod, Nizhny Novgorod, Russia
sovrasov.vlad@gmail.com

Abstract. В данной работе рассматривается параллельный алгоритм многокритериальной оптимизации. Рассматриваемый подход основан на применении информационно-статистического алгоритма к некоторой редуцированной однокритериальной задаче, множество глобальных оптимумов в которой совпадает с множеством слабоэффективных решений в исходной многокритериальной задаче. Последовательная версия данного метода была рассмотрена ранее. В данной работе к последовательному алгоритму многокритариальной оптимизации применяется схема распараллеливания по характеристикам, общая для всех информационно-статистических алгоритмов глобальной оптимизации. Также в работе впервые для многокритериального метода рассматривается одна из техник учёта локальных свойств оптимизируемой функции, позволяющая существенно ускорить сходимость.

Keywords: deterministi global optimization, multi-objective optimization, parallel numerical methods, derivative-free algorithms

1 Introduction

2 Problem Statement and Dimension Reduction

Задача многокритериальной оптимизации ставится следующим образом:

$$\min\{f(y) : y \in D\}, D = \{y \in \mathbb{R}^n : a_i \leqslant y_i \leqslant b_i, i = \overline{1, n}\} \tag{1}$$

Будем считать, что компоненты вектор-функции (частные критерии) $f_i(y), 1 \leqslant i \leqslant m$, удовлетоворяют в D условию Липшица с константами L_i:

$$|f_i(y_1) - f_i(y_2)| \leqslant L_i \|y_1 - y_2\|, y_1, y_2 \in D, 0 < L_i < \infty, i = \overline{1, m}$$

As the solution to the problem (1) usually accepted the set $S(D) \in D$ of strictly non-dominated points from the range of search, i. e.,

$$S(D) = \{y \in D : \nexists z \in D, f_i(z) < f_i(y), 1 \leqslant i \leqslant m\} \tag{2}$$

which is usually referred as the set of semi-effective (or weakly effective) solutions. The conditions in the right-hand side of the definition (2) are known as the principle of weak Pareto-optimality (or Slater's optimality principle).

The use of the evolvents $y(x)$ i.e. the curves filling the space are a classic dimension-reduction scheme for global optimization algorithms [?].

$$\{y \in R^N : -2^{-1} \leqslant y_i \leqslant 2^{-1}, 1 \leqslant i \leqslant N\} = \{y(x) : 0 \leqslant x \leqslant 1\}$$

Such a mapping allows the reduction of a problem (1) stated in a multidimensional space to solving a one-dimensional problem at the expense of worsening its properties. In particular, the one-dimensional functions $f_i(y(x))$ are not Lipschitzian but a Hölderian functions:

$$|f_i(y(x_1)) - f_i(y(x_2))| \leqslant H_i|x_1 - x_2|^{\frac{1}{N}}, x_1, x_2 \in [0, 1] \tag{3}$$

where the Hölder constants H_i are related to the Lipschitz constant L_i by the relation

$$H_i = 4L_i d\sqrt{N}, d = \max\{b_i - a_i : 1 \leqslant i \leqslant n\}$$

Therefore, not limiting the generality, one can consider the solving of the one-dimensional problem $\min\{f(y(x)) : x \in [0; 1]\}$, satisfying Hölder condition. The issues of numerically building the mapping like a Peano curve and the corresponding theory have been considered in detail in [?]. Here we would note that an evolvent built numerically is an approximation to the theoretical Peano curve with a precision of the order 2^{-m} where m is the building parameter of the evolvent.

3 Description of the Parallel Algorithm With Local Refinement

Рассмотрим схему скаляризации редуцированной задачи (1), представленную в []. Пусть

$$\varphi(x) = \max\{h(x, y) : y \in [0; 1]\}, x \in [0; 1]. \tag{4}$$

Рассмотрим скалярную задачу

$$\varphi^* = \min\{\varphi(x) : x \in [0; 1]\}. \tag{5}$$

Как показано в [], множество слабо-эффективных решений редуцированной задачи (1) совпадает множеством глобально оптимальных решений задачи (5), т.е.

$$S([0; 1]) = \{x \in [0; 1] : \varphi(x) = \varphi^*\} \tag{6}$$

Также в [] показано, что функция $\varphi(x)$ удовлетворяет условию Гёльдера при выполнении требований (**??**). Таким образом, к функции $\varphi(x)$ можно применить информационно-статистический алгоритм глобального поиска, чтобы решить задачу (5). Однако, $\varphi(x)$ задаётся через оператор $\max\{...\}$, поэтому непосредственно вычислить её затруднительно. В [] приведена модификация классического информационно-статистического алгоритма [], в которой значения $\varphi(x)$ вычисляются приближённо. Далее приведём модифицированную версию указанного алгоритма. Модификация заключается в использовании техники local refinement, описанной в [], а также в распараллеливании по характеристикам [].

Первые две итерации производятся в концевых точках $x^0 = 0$ и $x^1 = 1$ интервала $[0;1]$. Выбор точек $x^{k+j}, 1 \leqslant j \leqslant p$ осуществляется по правилам:

Step 1. Renumber the points in the set $X_k = \{x^1, \dots, x^k\} \cup \{0\} \cup \{1\}$, which includes the boundary points of the interval $[0,1]$ as well as the points of preceding trials, by the lower indices in order of increasing coordinate values i.e.

$$0 = x_0 < x_1 < \ldots < x_{k+1} = 1$$

Step 2. Compute the values

$$\mu_\nu = \max_{1 \leqslant i \leqslant k} \frac{|f_\nu(x_i) - f_\nu(x_{i-1})|}{\Delta_i}, 1 \leqslant \nu \leqslant m \qquad (7)$$

Step 3. Каждой точке x_i, $0 \leqslant i \leqslant k$, сопоставить значение

$$z_i = \max\{h(x_i, x_j) : 0 \leqslant j \leqslant k\}, \qquad (8)$$

где

$$h(x_i, x_j) = \min\{\frac{f_\nu(x_i) - f_\nu(x_j)}{\mu_\nu} : 1 \leqslant \nu \leqslant m\}, 0 \leqslant i, j \leqslant k \qquad (9)$$

Step 4. Для каждого интервала $(x_i, x_{i-1}), 1 \leqslant i \leqslant k$ вычислить величины

$$R(i) = \Delta_i + \frac{(z_i - z_{i-1})^2}{r^2 \Delta_i} - \frac{z_i + z_{i-1}}{2r} \qquad (10)$$

$$R^*(i) = \frac{R(i)}{\sqrt{(z_i - z^*)(z_{i-1} - z^*)} + 1.5^{-\alpha}}, \qquad (11)$$

называемые характеристиками. При этом $\Delta_i = (x_i - x_{i-1})^{\frac{1}{N}}$, $z^* = \min\{z_i : 1 \leqslant i \leqslant k\}$, а $r > 1$ и $\alpha \in [10; 30]$ — параметры метода.

Step 5. Если $q \neq 0$ и $s \mod q \neq 0$, то характеристики $R(i)$, $1 \leqslant i \leqslant k+1$, упорядочить в порядке убывания

$$R(t_1) \geqslant R(t_2) \geqslant \cdots \geqslant R(t_k) \geqslant R(t_{k+1})$$

и выбрать p наибольших характеристик с номерами интервалов t_j, $1 \leqslant j \leqslant p$. Иначе то же самое сделать с характеристиками $R^*(i), 1 \leqslant i \leqslant k+1$. Здесь s — номер текущей итерации. q — параметр метода, отвечающий за степень интенсивности локального уточнения. Чем меньше q, тем чаще используются характеристики R^*, заставляющие метод выбирать следующие точки вблизи текущего найденного минимума.

Step 6. Провести новые испытания в точках x^{k+j}, $1 \leqslant j \leqslant p$:

$$x^{k+j} = \frac{x_{t_j} + x_{t_j - 1}}{2} - \mathrm{sign}(z_{t_j} - z_{t_j - 1}) \frac{|z_{t_j} - z_{t_j - 1}|^n}{2r} \qquad (12)$$

Все p испытаний на этом шаге могут быть произведены параллельно на p вычислительных устройствах.

The algorithm is terminated if the condition $\Delta_{t_j} \leqslant \varepsilon$ is fulfilled at least for one of the numbers t_j, $1 \leqslant j \leqslant p$; here $\varepsilon > 0$ is the predefined accuracy. After the search is terminated, the set $S(\{x^0, \ldots, x^k\})$ of all non-dominated points of the truncated sequence $\{x^0, \ldots, x^k\}$ is accepted as an estimation for S from (6).

The theoretical substantiation of this method when $p = 1$ and $q = 0$ is presented in []. Siffitient condition of convergence is: exists an iteration such that $r\mu_\nu \geqslant 4H_\nu, 1 \leqslant \nu \leqslant m$.

4 Experimental Results

5 Conclusion

References

1. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. 78, 315–333 (1982)
2. Clarke, F., Ekeland, I.: Solutions périodiques, du période donnée, des équations hamiltoniennes. Note CRAS Paris 287, 1013–1015 (1978)
3. Michalek, R., Tarantello, G.: Subharmonic solutions with prescribed minimal period for nonautonomous Hamiltonian systems. J. Diff. Eq. 72, 28–55 (1988)
4. Tarantello, G.: Subharmonic solutions for Hamiltonian systems via a $\mathbb{Z}_p$ pseudoindex theory. Annali di Matematica Pura (to appear)
5. Rabinowitz, P.: On subharmonic solutions of a Hamiltonian system. Comm. Pure Appl. Math. 33, 609–633 (1980)