ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И РАСПРОСТРАНЕНИЕ ОГРАНИЧЕНИЙ
МКВМ-2004 РАБОЧИЕ СОВЕЩАНИЯ С. 184–192

Метод глобальной оптимизации на основе метода ветвей и границ

Ю.Г. Долгов

1. Введение

Решение задач глобальной оптимизации является, несомненно, актуальной проблемой. Подобные задачи возникают в самых различных областях человеческой деятельности (экономике, геологии, химии, биологии, физике и т. д.) Для гарантированного поиска глобального оптимума обычно пользуются интервальными методами, однако, подобные методы имеют, как правило низкую скорость сходимости.

Алгоритм, предложенный в данной статье позволяет избежать некоторых недостатков известных методов глобальной оптимизации и существенно уменьшить время решения.

Текст данной работы построен следующим образом. В п. 2 приведены определения, использованные в рамках данной работы и описываются методы, использованные в данной работе. В п. 3 предлагается описание предложенного метода и некоторые пояснения. В п. 4 представлены результаты работы алгоритма в сравнении с пакетом Numerica.

2. Основные определения, понятия и используемые методы

Приведем формальные определения ограничения и других терминов, используемых в дальнейшем.

Определение 2.1 (ограничение). *Ограничением* на множестве переменных $\{x_1, \dots, x_m\}$ будем называть подмножество S декартова произведения: $D_{x_1} \times \dots \times D_{x_m}$, где D_{x_i} – область определения переменной x_i.

Определение 2.2 (CSP – Constraint Satisfaction Problem). *CSP или задача удовлетворения ограничениям* – это тройка $P = (X, D, C)$, где $X = \{x_1, \dots, x_m\}$ множество переменных, $D = \{D_{x_1}, \dots, D_{x_n}\}$ – множество областей определения переменных (D_{x_i} – множество всевозможных значений переменной x_i), $C = \{c_1, \dots, c_k\}$ – это множество ограничений.

Определение 2.3 (решение CSP). *Решение задачи удовлетворения ограничениям* $P = (X, D, C)$ – это вектор $a = (a_1, \dots a_n) \in D = \{D_{x_1}, \dots, D_{x_n}\}$, для которого выполняются все ограничения $c_j \in C$.

Определение 2.4 (аппроксимация). Пусть $S \subseteq \Re$, *аппроксимацией* S будем называть наименьший интервальный вектор I, такой, что $S \subseteq I$ и обозначать $\operatorname{approx}(S)$.

Определение 2.5 (2B-совместность ограничения). Пусть c – это некоторое ограничение, определенное на множестве переменных $\{x_1, \ldots, x_m\}$. Будем говорить, что c 2B-*совместно*, если $\forall i \in \{1, \ldots m\} D_{x_i} = \operatorname{approx}\{\nu_i \in D_{x_i} | \exists \nu_1 \in D_{x_1}, \ldots, \nu_{i-1} \in D_{x_{j-1}}, \nu_{i+1} \in D_{x_{j+1}}, \ldots, \nu_m \in D_{x_m}$, такое что $c(\nu_1, \ldots, \nu_{i-1}, \nu_i, \nu_{i+1}, \ldots, \nu_m)\}$ – совместно.

Определение 2.6 (2B-совместность CSP). Задача удовлетворения ограничениям $P = (X, D, C)$ называется 2B-*совместной*, если все ее ограничения $c_j \in C$ 2B-совместны.

Определение 2.7 (оператор сужения). Пусть $P = (X, D, C)$ – некоторая задача удовлетворения ограничениям. С каждым ограничением $c_j \in C$и каждой переменной $x_j \in X$ сопоставим оператор сужения $\operatorname{red}_{c_i}^{x_j}$, который действует на множестве областей определений переменных, используемых в c_i, и результатом его применения является множество $S_j \subseteq D_j$, с удаленными из него несовместными значениями переменной x_j.

Запишем более формально: пусть ограничение c_i определенно на множестве переменных $\{x_1, \ldots, x_m\}$, тогда $\operatorname{red}_{c_i}^{x_j} : P(D_{x_1}) \times \ldots \times P(D_{x_m}) \to P\left(D_{x_j}\right)$.

Далее приводится краткое описание методов, используемых для создания метода ветвей и сужений.

2.1. Метод распространения ограничений

Метод распространения ограничений [8] является одним из методов установления 2B-совместности, который аналогичен алгоритму AC-3. Этот алгоритм также называется алгоритмом сужения. В вычислительном плане он может быть представлен как итерационный процесс, последовательно сужающий область, гарантированно содержащую все решения. На каждом шаге итерационного процесса применяется один из операторов сужения, стягивающий область значений переменных, связанных с некоторым ограничением. Как правило, оператор сужения уточняет область определения только одной переменной. Управление итерационным процессом осуществляется по данным, то есть для исполнения на очередном шаге выбирается только тот оператор сужения, аргументы которого были изменены в процессе вычислений на предыдущем шаге.

Операторы сужения могут представляться разными способами. Например, расщеплением исходной системы уравнений на примитивные (унарные и бинарные) соотношения посредством введения дополнительных переменных. В этом случае каждое примитивное соотношение позволяет выразить каждую из входящих в него переменных через остальные. Весь набор полученных таким образом примитивных уравнений и будет задавать множество операторов сужения для исходной системы. Достоинством этого подхода является возможность решения задач с недифференцируемыми функциями.

2.2. Метод ветвей и границ

Метод ветвей и границ [6] – это общий метода поиска решения. Метод начинает работу с определения нижней и верхней границ для исходной задачи. Если верхняя и нижняя границы совпадают, то полученный результат является оптимальным значением, и метод прекращает работу. Иначе, множество переменных разбивается на несколько собственных подмножеств, объединение которых совпадает с исходным множеством. Эти подзадачи становятся потомками исходной. Далее алгоритм применяется рекурсивно к каждой из подзадач, создавая дерево подзадач. Если оптимальное решение найдено для некоторой подзадачи, то оно является достижимым для исходной задачи (не обязательно оптимальным), но так как оно достижимо, его можно использовать для обрезания ветвей у исходного дерева. Процесс поиска продолжается до тех пор, пока каждая из подзадач не будет решена или выкинута или до тех пор, пока не будет достигнут заданный порог между лучшим из найденных решений и нижней границей $f\,(x)$ для всех нерешенных задач.

2.3. Метод Хансена

Метод Хансена был описан в работе [3] в 1992г. Суть работы метода заключается в последовательном удалении из начальной области, подобластей, в которых не содержится глобальный минимум. Удаление происходит одним из ниже перечисленных способов:

- Удаляются подобласти, в которых градиент g функции f отличен от нуля.
- Удаляются подобласти, в которых $f > \overline{f}$, где $\overline{f}$ – верхняя оценка глобального минимума.
- Удаляются подобласти, в которых функция f невыпукла.

Критерии остановки

Критерий A: $\mathrm{wid}\,(\mathbf{x}) < \varepsilon_{\mathbf{x}}$ для некоторого $\varepsilon_{\mathbf{x}} > 0$.
Критерий B: $\overline{f} - \underline{f} < \varepsilon_{\mathbf{f}}$ для некоторого $\varepsilon_{\mathbf{f}} > 0$.

2.4. Метод моделирования обжига

Метод моделирования обжига [5, 1, 4, 7] является представителем семейства методов Монте–Карло. Он основан на физическом процессе замораживания жидкостей или рекристаллизации металлов в процессе прокаливания. В качестве оценки решения может быть выбрана точка не только не уменьшающая значения целевой функции, но и увеличивающая его. Такое поведение поиска оптимального решения помогает избежать попадания в точки локального минимума.

3. Метод ветвей и сужений

В рамках данной работы был реализован метод, объединяющий преимущества различных подходов при решении задач глобальной оптимизации.

Описание метода

Данный метод, по сути, является усовершенствованием интервального метода ветвей и границ, на каждой итерации которого применяется метод Хансена для сужения области, содержащей глобальный оптимум. Для улучшения нижней оценки функции используется численный метод симуляции обжига, объединенный с методом мультистарта, который позволяет достаточно быстро найти хорошее приближение глобального оптимума. Для применения метода Хансена используется метод распространения ограничений.

Выбор этих методов не случаен, их комбинация позволяет компенсировать слабые стороны друг друга. Приведем сводную таблицу качественных характеристик используемых методов:

Характеристики	Метод		
	ветвей и границ	Хансена	локальной оптимизации
Гарантированный поиск глобального оптимума	+	–	–
Гарантированность вычислений	+	+	–
Вычисления над интервальными типами данных	+	+	–
Небольшие временные затраты	–	+	+
Небольшое количество используемых ресурсов	–	+	+
Сужение области поиска	+	+	–

Использование комбинации методов позволяет гарантированно находить глобальный оптимум задачи, при этом использование метода локальной оптимизации для улучшения текущего рекорда значительно ускоряет работу метода ветвей и границ. Метод Хансена также позволяет откидывать области, в которых заведомо не содержится решения.

На основе результатов тестов были сделаны некоторые корректировки в используемых методах:

- в методе ветвей и границ производится деление на три части, кроме того, при выборе подзадачи выбирается та, где среднее значение функции наименьшее;
- в методе Хансена используется модернизированный метод распространения ограничений, предложенный в работе [2];
- запуск локального метода происходит только после выполнения некоторого количества итераций (это число может устанавливаться пользователем) метода ветвей и границ, что позволяет уменьшить время решения относительно простых задач. Количество стартовых точек в методе мультистарта вычисляется по формуле: $n = n_0 \cdot \mathrm{dim}$, где dim – размерность задачи, а n_0 – заданное пользователем число.

$\mathrm{GlobOpt}\left(X, F\left(x\right), k_{\max}, \varepsilon, n_{\mathrm{loc}}\right)$
`Псевдокод:`

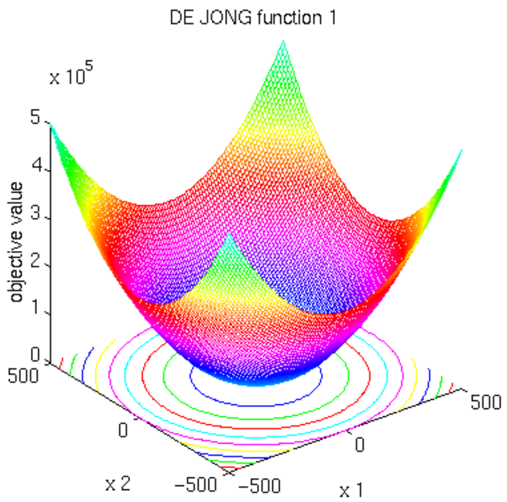
```
1.  BEGIN
2.     k = 1, record = [∞,∞], SolList = null;
3.     T = X_0 = X; – дерево подобластей
```

```
  WHILE T не пусто AND k < k_max DO
    IF k mod n_loc = 0 DO
      opt_record = LocOpt(ε,dim); (dim – размерность задачи)
      IF opt_record < record DO
        Очистить SolList;
      END DO
    END DO
    Взять элемент t_k = X_k из T;
    t_k = HansenMeth(t_k, ε);
    IF F(X_k) ≤ record DO
      IF wid(F(X_k))_l < ε DO
        IF F(X_k) < record DO
          Очистить SolList;
        END DO
        record = [min(record_l, F(X_k)_l), min(record_r, F(X_k)_r)]
        Добавить(X_k, F(X_k)) в SolList;
      ELSE DO
        Поделить X_k на 3 части и добавить их в T;
      END DO
    END DO
    Удалить X_k из T, k = k + 1;
  END DO
  RETURN SolList;
END
```

4. Результаты численных экспериментов

Приведем сравнение результатов решения нескольких задач глобальной оптимизации методом ветвей и сужений и пакетом Numerica.

Пример 1 (Функция де Джонга).



$$f = \sum_{i=1}^{n} x_i^2, \quad x_i = [-10^{30}, 10^{30}].$$

Глобальный минимум:

$$f = 0, \quad x_i = [0, 0],$$

$i \in \{1, \ldots, n\}$.

SibCalc	Numerica	n
0	0.10	5
0.015	0.30	20
0.015	8.00	100
0.064	40.10	200
0.156	Out of Memory	400

Пример 2.

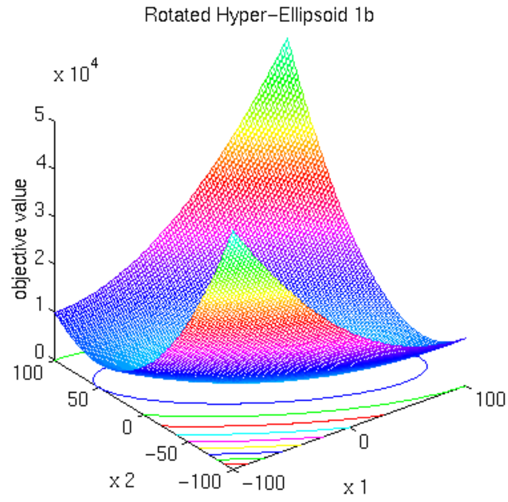
$$f=\sum_{i=1}^{n}\sum_{j=1}^{i}x_j^2,\ \ x_i=[-10^{30},10^{30}].$$

Глобальный минимум:

$$f=0,\quad x_i=[0,0],$$

$i\in\{1,\dots,n\}$.

SibCalc	Numerica	n
0	0.10	5
0.015	3.70	20
0.438	Out of Memory	100
1.813	Out of Memory	200
7.953	Out of Memory	400



Пример 3 (функция Растригина).

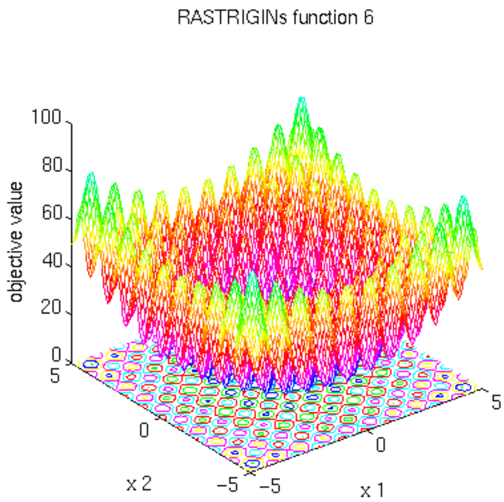
$$f=\sum_{i=1}^{n}(x_i^2-10\cos(2\pi x_i))+10n,$$
$$x_i=[-10^{30},10^{30}].$$

Глобальный минимум:

$$f=0,\quad x_i=[0,0],$$

$i\in\{1,\dots,n\}$.

SibCalc	Numerica	n
0.328	0.10	5
2.578	0.30	10
11.546	0.70	15
25.141	1.20	20
58.016	1.90	25

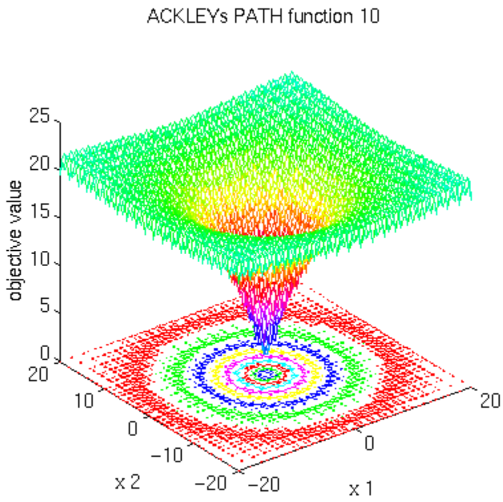


Пример 4 (функция Экли).

$$f=-20\cdot\exp\left(-\frac{1}{5}\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(2\pi x_i\right)\right)+20+\exp(1),$$
$$x_i=[-33,33]\,.$$

Глобальный минимум:

$$f=0,\quad x_i=[0,0],\quad i\in\{1,\dots,n\}.$$



SibCalc	Numerica	n
1.578	0.40	5
15.266	1.20	10
83.546	3.60	15

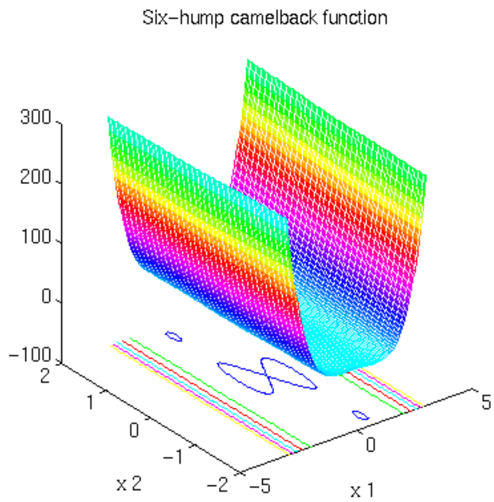
Пример 5 (шестигорбый верблюд).

$$\left(4 - 2.1x_1^2 + \frac{x_1^4}{3}\right)x_1^2 + x_1x_2 + \left(-4 + 4x_2^2\right)x_2^2, \quad x_1 = x_2 = [-10^{30}, 10^{30}].$$

Глобальный минимум:

$$f = [-1.03628453, -1.03628453].$$

Первое решение: $x_1 = [-0.089842, -0.089842], x_2 = [0.7126564, 0.7126564]$.
Второе решение: $x_1 = [0.089842, 0.089842], x_2 = [-0.7126564, -0.7126564]$.



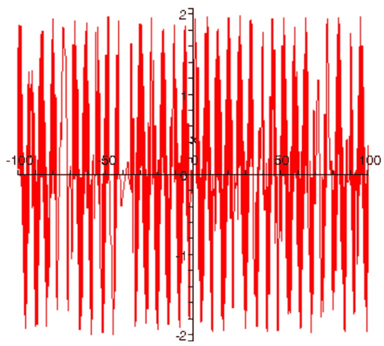
При решении Numerica находит только второе решение.
Время счета: SibCalc – 0.125, Numerica – 0.

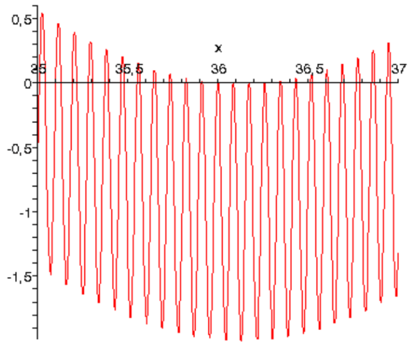
Пример 6.

$$f = \sin x + \sin x^2, \quad x = [-100, 100].$$

Глобальный минимум:

$$f = [-1.999999438712797, -1.999999438712796]\,,$$
$$x = [36.12937493100531, 36.12937493100539]\,.$$





Numerica находит решение:

$$f = [-1.999999438712743, -1.999999438712743]\,,$$
$$x = [99.02767859, 99.02767859]\,.$$

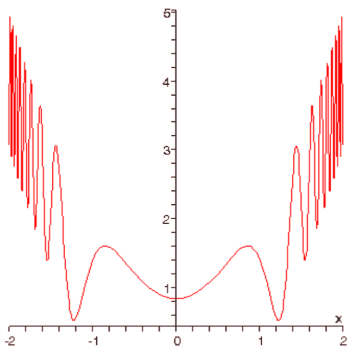
Время счета: SibCalc – 0.015, Numerica – 0.1.

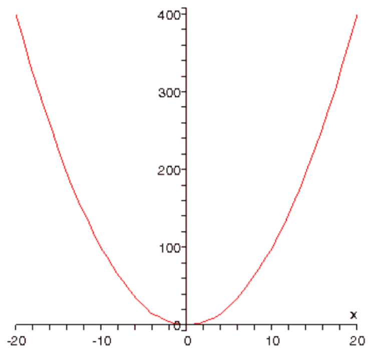
Пример 7.

$$f = x^2 + \sin\left(\exp\left(x^2\right)\right), \quad x = [-100, 100].$$

Глобальный минимум:

$$f = [0.5264768703, 0.5264768703]\,,$$
$$x = \pm[1.22528916901954, 1.225289169019542].$$





Numerica находит решение $x = [0, 0]$.
Время счета: SibCalc – 0.015, Numerica – 0.1.

Результаты экспериментов показали, что метод ветвей и сужений позволяет гарантированно находить глобальный минимум, в отличие от пакета Numerica, который зачастую теряет корни, а в примерах 6 и 7 находит неверные значения. Кроме того, время работы алгортима остается вполне приемлемым, он работает медленнее лишь на задачах с большим количеством локальных экстремумов (примеры 3 и 4).

5. Заключение

В рамках данной работы был проведен анализ существующих алгоритмов поиска глобального оптимума. Разработан и реализован метод глобальной оптимизации. Проведены сравнения с пакетом Numerica. По результатам сравнения можно сделать следующие выводы:

1. Созданный метод позволяет гарантированно находить решения, даже для задач с большим значением локальных минимумов.
2. Метод может реально применяться для задач с размерностями несколько десятков.
3. Скорость работы метода сравнима, а на некоторых задачах превосходит скорость работы существующих методов.

Стоит также отметить, что полученный метод может применяться для решения задач условной оптимизации. Однако, скорость работы метода при этом существенно падает из-за отсутствия возможности применять в этом случае метод Хансена.

В дальнейшем предполагается адаптировать и улучшить данный метод для решения задач условной оптимизации.

Список литературы

[1] Cerny V. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm // J. Opt. Theory Appl. – 1985. – Vol. 45. – P. 41–51.

[2] Dolgov Yu. On strategies of the narrowing operator selection in the constraint propagation method // Lect. Notes Comput. Sci. – 2004. – Vol. 2890. – P. 431–437.

[3] Hansen E., Global Optimization Using Interval Analysis. – New York: Dekker, 1992.

[4] Kirkpatrick, S., Gelatt Jr.C.D., Vecchi M.P. Optimization by simulated annealing // Science. – 1983. – Vol. 220. – P. 671–680.

[5] Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E. Equation of state calculations by fast computing machines // J. Chem. Phys. – 1953. – Vol. 21. – P. 1087–1092.

[6] Nemhauser G.L., Wolsey L.A. Integer combinatorial optimization. – New York: Wiley.

[7] Press Wm. H., Flannery B., Teukolsky S., Vettering Wm. // Numerical Recipes. – New York: Cambridge University Press, 1986. – P. 326–334.

[8] Semenov A., Kashevarova T., Leshchenko A., Petunin D. Combining various techniques with the algorithm of subdefinite calculations // Proc. of the 3rd Intern. Conf. on the Practical Application of Constraint Technology PACT'97. – London (England), April 1997. – P. 287–306.