УДК 519.85

Верхняя оценка числа локальных максимумов в задаче максимизации нормы вектора на выпуклом множестве

А. И. Косолап, Ю.В. Черноусова
Украинский государственный химико-технологическийуниверситет, Украина

В работе рассмотрены многоэкстремальные задачи нелинейной оптимизации. Эти задачи преобразуются методом точной квадратичной регуляризации к одноэкстремальным либо к максимуму нормы вектора на выпуклом множестве. При таком преобразовании число локальных экстремумов значительно сокращается, что упрощает решение исходной задачи. Найдена оценка сверху числу различных локальных экстремумов преобразованной многоэкстремальной задачи.

Ключевые слова: *многоэкстремальные задачи, точная квадратичная регуляризация.*

В роботі розглянуті багатоекстремальні задачі нелінійної оптимізації. Ці задачі перетворюються методом точної квадратичної регуляризації до одноекстремальних або до максимуму норми вектора на опуклій множині. При такому перетворенні число локальних екстремумів значно скорочується, що спрощує розв'язок початкової задачі. Знайдена оцінка зверху числа різних локальних екстремумів перетвореної багатоекстремальної задачі.

Ключові слова*: багатоекстремальні задачі, точна квадратична регуляризація*

In this paper we consider the multi-extremal problems of nonlinear optimization. By the exact quadratic regularization, these problems are converted to those that are one-extremal or the problem of vector norm maximum on a convex set. Under this transformation, the number of local extremais greatly reducedthat simplifies the original problem solution. We have found an upper bound for the number of different local extrema in the transformed multi-extremal problem.

Key words*: multi-extremal problems, the exact quadratic regularization*

1. Введение

Моделирование сложных систем приводит к большому количеству многоэкстремальных задач, которые относятся к классу NP-полных. Число локальных экстремумов в таких задачах часто равно 2^n, $n!$ и более [1]. В настоящее время разработаны детерминированные и стохастические методы для решения многоэкстремальных задач [2–4]. Детерминированные методы позволяют находить решения только для задач малой размерности. С увеличением размерности время решения задачи растет экспоненциально. Стохастические методы позволяют находить решение многоэкстремальной задачи только с некоторой вероятностью. Как показывают многочисленные эксперименты на тестовых задачах, стохастические методы иногда позволяют находить решения близкие к оптимальным, однако в большинстве случав эти решения далеки от оптимальных [5].

Для решения многоэкстремальных задач эффективным является поиск преобразований, при которых преобразованная задача будет иметь только один или небольшое число локальных экстремумов. С помощью полуопределенной релаксации общие квадратичные и полиномиальные задачи преобразуются к одноэкстремальным. Однако, в общем случае, полуопределенная релаксация позволяет находить только оценки оптимальных решений[1].

В работе используется точная квадратичная регуляризация, которая позволяет преобразовывать многоэкстремальные задачи к одноэкстремальным либо к задаче максимизации нормы вектора на выпуклом множестве [3].

2. Метод точной квадратичной регуляризации

Рассмотрим задачу глобальной оптимизации

$$\min\left\{f_0(x)\middle| f_i(x) \le 0, i = 1,..., m, x \in E^n\right\} \tag{1}$$

где все $f_i(x)$ – дважды дифференцированные функции, E^n –евклидово пространство. Допустим, что решение задачи (1) существует. Для этого достаточно чтобы функция $f_0(x)$ была непрерывна, а допустимое множество задачи (1) было компактно. Введем новую переменную x_{n+1} и сведем задачу (1) к виду

$$\min\left\{x_{n+1}\middle| f_0(x) + s \le x_{n+1}, f_i(x) \le 0, i = 1,.., m, x \in E^n\right\}, \tag{2}$$

где значение параметра s выбираем таким, чтобы

$$f_0(x^*) + s \ge \left\|x^*\right\|^2 \tag{3}$$

x^* – решение задачи (1).

Если глобальный минимум задачи (1) достигается на множестве точек, то достаточно найти только одну точку глобального минимума. Поэтому для определения параметра s можно выбрать минимальное значение $\left\|x^*\right\|^2$. Например, в задаче $\left\{-x_1\middle|2 \le x_1 \le 6\right\}$ имеем $s \ge 42$.

Далее, используем преобразование пространства $x = Az$, где матрицаA порядка $(n+1)\times(n+1)$ имеет вид

$$\begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ z_1 & z_2 & ... & z_{n+1} \end{pmatrix},$$

что сводит задачу (2) к следующей:

$$\min\left\{\|z\|^2 \middle| f_0(\bar{z}) + s \le \|z\|^2, f_i(\bar{z}) \le 0, i = 1,..., m, x \in E^n\right\}, \tag{4}$$

где $\bar{z} = (z_1,..., z_n), z = (\bar{z}, z_{n+1})$.

Существует такое значение $r > 0$, что все функции

$$g_0(z) = f_0(\bar{z}) + s + (r-1)\|z\|^2, g_i(z) = f_i(\bar{z}) + r\|z\|^2, i = 1,..., m$$

будут выпуклыми для допустимых значений z. Действительно, при соответствующем выборе параметра $r > 0$, гессианы функций $g_0(z)$ и $g_i(z)$, $i = 1,...,m$ будут положительно определенными матрицами (матрицы с преобладающей главной диагональю). Если среди $f_i(\bar{z}), i = 1,...,m$ есть выпуклые функции, то эти ограничения остаются неизменными.

Таким образом, задача (4) сведена к следующей:

$$\min\left\{\|z\|^2 \middle| g_i(z) \le d, i = 0,..., m, r\|z\|^2 = d\right\} \tag{5}$$

где все $g_i(z)$ выпуклые функции.

Следовательно, задача (1) преобразована к минимизации квадрата нормы вектора z, где переменными задачи (5) есть $(z,d) - (n+2)$-мерный вектор. Докажем следующее утверждение.

Теорема* 1.Пусть* $\left(z^0, d_0\right)$ – *решение задачи* (5) и *для параметра s выполняется условие* (3), *тогда* $x^* = \bar{z}^0$ – *решение задачи* (1).

Доказательство. Имеем

$$f_0(\bar{z}^0) + s + (r-1)\left\|z^0\right\|^2 \le d_0,$$

$$f_i(\bar{z}^0) + r\left\|z^0\right\|^2 \le d_0, i = 1,...,m$$

с учетом $r\left\|z^0\right\|^2 = d_0$, получим

$$f_0(\bar{z}^0) + s \le \left\|z^0\right\|^2, f_i(\bar{z}^0) \le 0, i = 1,...,m$$

Первое ограничение равносильно $f_0(\bar{z}^0) + s \le \left\|z^0\right\|^2 (z_{n+1}^0 = 0)$ или

$f_0(\bar{z}^0) + s = \left\|z^0\right\|^2$. Пусть z^* – решение задачи (1) и $z_{n+1}^* = f_0(\bar{z}^*) + s$, $d^* = r\left\|z^*\right\|^2$

($f_0(\bar{z}^*) + s = \left\|z^*\right\|^2$). Тогда из условий

$$f_0(\bar{z}^0) + s \le \left\|z^0\right\|^2,$$

$$f_0(\bar{z}^*) + s \le \left\|z^*\right\|^2$$

и $\left\|z^0\right\|^2 \le \left\|z^*\right\|^2$ следует $f_0(\bar{z}^*) \ge f_0(\bar{z}^0)$, откуда $f_0(\bar{z}^*) = f_0(\bar{z}^0)$.

Аналогично, из условий

$$f_0(\bar{z}^0)+s=\left\|z^0\right\|^2,$$

$$f_0(\bar{z}^*)+s=\left\|z^*\right\|^2$$

и $\left\|z^0\right\|^2 \le \left\|z^*\right\|^2$ будет $f_0(\bar{z}^*) \ge f_0(\bar{z}^0)$, откуда снова $f_0(\bar{z}^*) = f_0(\bar{z}^0)$.

Теорема доказана.

Следствие.*Первое ограничение задачи* (4) *активно в точке глобального минимума.*

Из $f_0(\bar{z}^*)+s \le \left\|z^*\right\|^2$ и (3) имеем $f_0(\bar{z}^*)+s = \left\|z^*\right\|^2$.

Введем обозначения

$$S_1(d)=\{z|g_i(z)\le d, i=0,...,m\} \text{ и } S_2(d)=\left\{z\left|r\|z\|^2 \le d\right.\right\}.$$

Множества $S_1(d)$ и $S_2(d)$ будут выпуклыми и определяют допустимое множество задачи (5). Справедливо следующее включение $S_1(d)\subset S_1(d+\Delta)$, $S_1(d)\subset S_2(d+\Delta)$ для любого $\Delta>0$, что следует из выпуклости множеств $S_1(d)$ и $S_2(d)$. Анализ расположения этих множеств в пространстве позволяет разбить исходную задачу (1) на два класса сложности.Задачипервогокласса будут одноэкстремальными, а второго, в общем случае, многоэкстремальными.

Пусть d_0– минимальное значение d, для которого множество $S_1(d_0)\ne\varnothing$. Нахождениеd_0 равнозначно решению выпуклой задачи

$$\min\{d|g_i(z)\le d, i=1,...,m\}.$$

Найдем $d_m=\max\{0,d_0\}$.

Покажем, что сложность решения задачи (5) зависит от взаимного расположения множеств $S_1(d_m)$ и $S_2(d_m)$ в пространстве.

1. Если $S_1(d_m)\bigcap \operatorname{int} S_2(d_m)=\otimes$ или $S_1(0)\ne\otimes$, то задача (5) эквивалентна выпуклой задаче

$$\min\left\{d\left|g_i(z)\le d, i=0,...,m, r\|z\|^2\le d\right.\right\}, \tag{6}$$

которая эффективно решается методом локальной оптимизации (прямодвойственным методом внутренней точки PDIPM) [6]. Если $\left(z^*,d^*\right)$ – решение задачи (6), то $x^*=\bar{z}^*$ – точка глобального минимума задачи (1) (при выполнении условия $S_1(0)\ne\otimes$ решение задачи (1) – тривиально $x^*=0$). Решение задачи (6) эквивалентно нахождению точки соприкосновения двух выпуклых множеств при минимальном значении d. Очевидно, что точка соприкосновения будет допустимой для задачи (5) и в этой точке достигается минимальное значение $\|z\|^2$.

2. Если $S_1(d_m)\bigcap \operatorname{int} S_2(d_m) \neq \otimes$ то задача (5) эквивалентна задаче максимизации квадрата нормы вектора

$$\max\left\{\|z\|^2 \middle| g_i(z) \le d, i=0,...,m, r\|z\|^2 \le d\right\}. \qquad (7)$$

В задаче (7) необходимо найти минимальное значение *d*, при котором множество $S_1(d)$ касается границы множеств $S_2(d)$ изнутри. При меньших значениях *d* допустимое множество задачи (5) будет пустым.

Таким образом, если в точке локального минимума (z^*, d^*) задачи (6) $r\|z^*\|^2 = d^*$, то соответствующая задача (1) относится к первому классу сложности, иначе – ко второму.

Теорема 2. *Пусть* (z^*, d^*)*–решение задачи* (6) *и* $r\|z^*\|^2 = d^*$ *тогда* $x^* = \bar{z}^*$ *– точка глобального минимума задачи* (1).

Доказательство. Задача (1) эквивалентна задаче (5). Покажем, что (z^*, d^*) – решение задачи (5). Допустим противное, что существует (z^0, d_0) с меньшим значением $\|z^0\|^2 < \|z^*\|^2$. Если $r\|z^0\|^2 = d_0$, то $d_0 < d^*$, но это невозможно, так как задача (6) выпуклая. Тогда $r\|z^0\|^2 < d_0$, но это противоречит условию теоремы.

Теорема показывает, что задачи второго класса сложности также могут быть решением выпуклой задачи (6), если выполняются условия теоремы 9.

Преобразуем задачу глобальной оптимизации (1) так, чтобы ее переменные принимали значения $x \ge 0$. Если ограничения на переменные задачи (1) заданы двухсторонними ограничениями $a_i \le x_i \le b_i$, то замена $x_i = x_i - a_i$ переводит поиск глобального минимума в положительный ортант. Можно также представить переменные в виде разности двух положительных переменных $x_i = x_i^+ - x_i^-$, где $x_i^+ \ge 0, x_i^- \ge 0$. После выполнения вышеуказанных преобразований, задача (7) будет иметь вид

$$\max\left\{\|x\|^2 \middle| g_i(x) \le d, i=0,...,m, x \ge 0, r\|z\|^2 = d\right\} \qquad (8)$$

Задачу (8) будем решать следующим образом. Фиксируем значение переменной *d* и находим решение x^* задачи

$$\max\left\{\|x\|^2 \middle| g_i(x) \le d, i = 0,...,m, x \ge 0\right\}. \tag{9}$$

Если $r\left\|x^*\right\|^2 = d$, то задача (8) решена, и x^* – ее решение, иначе найдем отрезок $[d_{\min}, d_{\max}]$ для переменной d. Достаточно взять $d_{\min} = d_m$, а $d_{\max}$ определить, решая последовательность задач (9) методом локальной максимизации PDIPM [6] для $d = d_m + kh$, где h – величина шага, k=1,…Тогда $d_{\max}$ находим на интервале $[d_m + (k_0 - 1)h, d_m + k_0 h]$ методом дихотомии, решая задачу (9) до достижения равенства $r\|x\|^2 = d_{\max}$. Заметим, что при увеличении dзначение $\|x\|^2$ монотонно возрастает, что упрощает поиск $d_{\max}$.

3. Верхняя оценка числа локальных максимумов

Задача (9) многоэкстремальна, но в некоторых частных случаях ее решение сводится к одноэкстремальной. Это будет тогда, когда допустимое выпуклое множество является вписанным в шар (с центром в точке c) выпуклым многогранником. Многогранник вписан в шар, если все его вершины принадлежат шару. В таком случае, замена целевой функции в задаче (5) на линейную $c^T x$ преобразует ее к одноэкстремальной задаче линейного программирования. Если решение соответствующей задачи линейного программирования достигается на границе шара, то это решение совпадает с решением задачи (5).Это утверждение справедливо и для произвольного выпуклого множества.

В общем случае, найдем верхнюю оценку числа локальных максимумов преобразованной задачи (5). Будем различать два локальных максимума x^1 и x^2 задачи (5), если

$$\left|\left\|x^1\right\|^2 - \left\|x^2\right\|^2\right| \ge \mu. \tag{10}$$

Таким образом, нас интересуют локальные максимумы с различными значениеми целевой функции. Справедливо следующее утверждение.

Лемма. *Пусть x^1и x^2 – два соседних локальных максимума задачи* (5), *для которых условие* (10) *выполняется, тогда треугольник Ox^1x^2 не содержит точек локального максимума задачи* (5).

Доказательство. Точки x^1 и x^2 принадлежат поверхностям шаров с радиусами d/r и d/r-μ соответственно (см. рис. 1). Допустим противное, что треугольник Ox^1x^2 содержит точку максимума x^0. Тогда прямая, проходящая через точки x^1 и x^0 будет секущей шара радиуса d/r–μ. Так как точки x^1 и x^0 допустимы для задачи (5), то допустимым будут и точки отрезка $[x^1, x^2]$ (допустимое множество задачи (5) выпуклое). Но тогда вдоль этого отрезка значение целевой функции задачи (5) возрастает и точка x^0 не может быть точкой локального максимума задачи (5). Лемма доказана.

Теорема 3. *Количество локальных максимумов в задаче максимизации нормы на выпуклом множестве* (5) *не превосходит числа k, которое удовлетворяет следующему неравенству*

$$\sum_{i=1}^{k} \arccos\left(\frac{d/r - i\mu}{d/r}\right) \le \frac{\pi}{2}. \qquad (11)$$

Доказательство. Сначала проведем доказательство для двумерного случая (рис. 1).

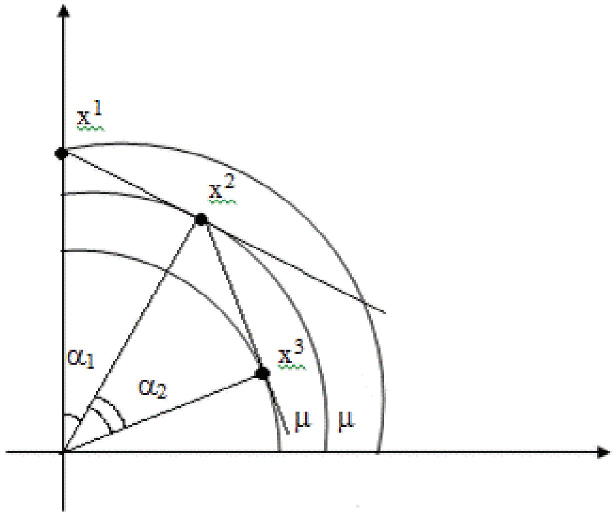

Рис. 1. Точки локальных максимумов

На окружности радиуса d/r берем точку локального максимума x^1 и через нее проведем касательную к окружности радиуса $d/r-\mu$. Касание будет в точке x^2.Два соседних локальных максимума x^1 и x^2 образуют угол α_1, равный

$$\alpha_1 = \arccos\left(\frac{d/r - \mu}{d/r}\right).$$

Следующие два локальных максимума x^2 и x^3 образуют угол α_2

$$\alpha_2 = \arccos\left(\frac{d/r - 2\mu}{d/r}\right).$$

Тогда сумма всех углов, которые образуют все соседние локальные максимумы-задачи (5) удовлетворяет неравенству (11).

Рассмотрим теперь n-мерный случай. Допустим противное, что углы образованные соседними локальными максимумами x^1, z^2 и z^2, z^3 меньше, чем угол, образованный x^1, x^3. Тогда точка z^3 попадает в один из треугольников Ox^1x^2 или Ox^2x^3 (точки x^2, x^3 всегда можно выбрать таким образом, чтобы точки x^1, x^2, x^3, z^3 лежали в одной плоскости). Но, в соответствии с леммой, точка z^3 не может быть локальным максимумом. Полученное противоречие доказывает теорему.

Значение k из неравенства (11) можно определить только численно. Результаты вычислений представлены в табл. 1.

Табл. 1. Количество локальных максимумов задачи (9)

d/r	Δ	*k*
5	0,2	3
10	0,1	6
10	0,001	48
25	0,005	23
100	0,1	13
100	0,5	7
500	0,2	18
1000	0,1	29
10000	0,1	105

Эти результаты показывают, что число локальных максимумов в задаче (5) является небольшим.

4. Выводы и направления дальнейших исследований

Рассмотрена общая задача нелинейной оптимизации. С помощью метода точной квадратичной регуляризации она преобразована к максимизации нормы вектора на выпуклом множестве. Показано, что такое преобразование позволяет существенно сократить число локальных экстремумов исходной задачи, что упрощает решение преобразованной задачи локальными методами. В работе получена верхняя оценка числа максимумов. При определенных условиях число максимумов будет равно единице. Это предмет для дальнейших исследований.

ЛИТЕРАТУРА

1. Косолап А. И. Методы глобальной оптимизации / А. И. Косолап. – Днепропетровск: Наука и образование, 2013. – 316 с.

2. Horst R., Tuy H. Global Optimization: Deterministic Approaches, 3rd ed. – Springer-Verlag, Berlin, 1996. – 726 p.

3. Floudas C. A., Gounaris C. E. A review of recent advances in global optimization // J. Glob. Optim. – 2009. – v. 45, no. 1. – P. 3–38.

4. Kenneth V.P.,Storn R.M., Lampinen J.A. Differential Evolution. A Practical Approach to Global Optimization.– Berlin Heidelberg: Springer-Verlag, 2005.– 542 p.

5. Cagnina L. C. Solving constrained optimization problems with a hybrid particle swarm optimization algorithm / L. C. Cagnina, S. C. Esquivel, C. A Coello //Engineering Optimization, v. 43, No. 8, 2011.– P. 843–866.

6. Nocedal, J., Wright S.J. Numerical optimization. – Springer, 2006. – 685 p.

Надійшла 14.01.2015.