УДК 519.8

Ю.П. ЛАПТИН, Д.Л. КРОШКО

# НЕКОТОРЫЕ НЕЛИНЕЙНЫЕ ОПТИМИЗАЦИОННЫЕ ЗАДАЧИ СЕТЕВОЙ СТРУКТУРЫ

*Рассматриваются задачи, которые могут быть описаны совокупностью взаимосвязанных блоков. Некоторые выходы одних блоков могут быть входами для других блоков. Функциональные зависимости блоков определены на ограниченных областях. Разработан подход, основанный на доопределении рассматриваемых функций на все пространство.*



В работе рассматриваются задачи, часто возникающие при проектировании сложных технических объектов [1], в частности, энергетических котлоагрегатов [2–3]. К таким относятся задачи, которые могут быть описаны совокупностью взаимосвязанных блоков. Каждый блок характеризуется наборами входов и выходов. При фиксированных входах однозначно определяются выходы. Некоторые выходы одних блоков могут быть входами для других блоков. Особенностью рассматриваемых задач является ограниченность областей определения функциональных зависимостей блоков, что затрудняет применение существующих методов. Разработан подход, основанный на доопределении рассматриваемых функций на все пространство. Результирующая задача является негладкой, и для ее решения используются методы негладкой оптимизации. Результаты вычислительных экспериментов на задачах этого класса показали определенные преимущества методов негладкой оптимизации по сравнению с существующим программным обеспечением.

**Постановка задачи.** Задана совокупность $V$ непрерывно дифференцируемых вектор-функций $f^q(x^q)$, $x^q \in R^{n^q}$, $q \in V$. Обозначим $I^q$ – множество индексов переменных, $|I^q| = n^q$, $J^q$ – множество индексов функций для вектор-функции $f^q$. Предполагается, что область определения вектор-функции $f^q$ задается множеством $S^q$ простой структуры. Например,

$$S^q = \left\{ x^q : A^q x^q \le b^q \right\}, \tag{1}$$

где $b^q \in E^{m^q}$, $A^q$ – матрица соответствующей размерности, или

$$S^q = \left\{ x^q : (x^q, A^q x^q) + (b^q, x^q) \le r^q \right\}, \tag{2}$$

где $A^q$ – положительно определенная квадратная матрица, $r^q$ – положительное число.

Простота структуры множеств $S^q$ подразумевает, что существуют эффективные алгоритмы решения задачи проектирования произвольной точки на множество $S^q$.

Вектор-функцию $f^q$ вместе с областью определения $S^q$ назовем блоком $B^q$, переменные $x_i^q$, $i \in I^q$ будем называть входами, функции $f_j^q$, $j \in J^q$ – выходами блока $B^q$.

Совокупность блоков является связанной, т.е. выходы некоторых блоков могут быть входами для других блоков. Связи для каждой пары блоков $B^p$ и $B^q$ будем описывать множествами $M(p,q) = \left\{ (i,j) : i \in I^q,\ j \in J^p \right\}$. Множество $M(p,q)$ определяет выходы блока $B^p$, которые являются входами блока $B^q$:

$$x_i^q = f_j^p(x^p), \quad (i,j) \in M(p,q). \tag{3}$$

Будем говорить, что блоки $p$, $q$ связаны дугой $(p,q)$, если $M(p,q) \ne \varnothing$, множество всех дуг обозначим $E$.

Предполагается, что в совокупности функций $f_i^q$, $j \in J^q$, $q \in V$ выделено подмножество $T \subseteq \left\{ (j,q) : j \in J^q,\ q \in V \right\}$, которое определяет ограничения, и выделена функция $f_{j^0}^{q^0}$, которая является целевой функцией оптимизационной задачи.

Рассматриваемую задачу сформулируем следующим образом: найти

$$\min f_{j^0}^{q^0}(x^{q^0}), \tag{4}$$

при ограничениях

$$f_j^q(x^q) \le 0, \quad (j,q) \in T, \tag{5}$$

$$x_i^q = f_j^p(x^p), \quad (i,j) \in M(p,q),\ (p,q) \in E, \tag{6}$$

$$x^q \in S^q,\ q \in V. \tag{7}$$

Совокупность блоков $B^q$, $q \in V$ и связей между ними $E$ будем называть сетью вычислительных блоков $(V, E)$. Выходами сети блоков является объединение выходов всех блоков, входами сети блоков – переменные $x_i^q$, не являющиеся выходами каких-либо блоков. В дальнейшем будем рассматривать ациклические сети блоков.

Задачи вида (4)–(7) являются характерными при проектировании сложных технических объектов [1], в частности, энергетических котлоагрегатов [2–3]. Блоками при этом являются функциональные зависимости $f^q(x^q)$, реализующие инженерные методики расчетов отдельных компонент таких объектов. Ограничениями вида (7) описывается область применения этих методик.

Очевидно, что соотношения (6) при условии ацикличности сети позволяют выразить любой выход сети $(V, E)$ в виде функции от входов сети. При этом размерность задачи (4)–(7) существенно уменьшается, выходы сети являются сложной суперпозицией функций, описывающих отдельные блоки.

Такое представление задачи (4)–(7) с помощью суперпозиций функций, описывающих отдельные блоки, будем называть *прямой редукцией* исходной задачи.

Определенной проблемой при разработке алгоритмов решения редуцированной задачи является учет ограничений (7), которые задают области определения функций блоков и в результате редукции уже не описывают множества простой структуры. Так, в случае (1) левые части ограничений после редукции являются нелинейными функциями от входов сети.

Более того, пусть $x$ – входы сети блоков, $x^q(x)$ – входы блока $B^q$, как функция от входов $x$. Очевидно, что если $x^q(x) \notin S^q$, то выходы блока $B^q$ не определены, а также не определены входы и выходы любого блока $B^p$, для которого существует путь в сети $(V, E)$ из $B^q$ в $B^p$. То есть не определены функции, описывающие область определения (7) для блока $B^p$.

Изложенное подчеркивает сложность описания областей определения функций, полученных в результате прямой редукции исходной задачи.

Для решения редуцированной задачи могут использоваться методы барьерных функций (внутренних штрафов) и методы проектирования на допустимую область. Заметим, что трудоемкость задачи проектирования в данном случае может быть весьма высокой.

В настоящей работе предлагается новый подход, связанный с доопределением используемых функций на все пространство.

**Расширенная редукция исходной задачи.** Рассмотрим сеть вычислительных блоков $(V,E)$. Для блока $B^q$, $q \in V$ обозначим:

$p_{S^q}(x^q)$ – проекция точки $x^q \in R^{n^q}$ на множество $S^q$,

$d_{S^q}(x^q) = \left\| x^q - p_{S^q}(x^q) \right\|$ – расстояние от точки $x^q$ до множества $S^q$.

Расширением вектор-функции $f^q(x^q)$ для произвольной точки $x^q \in R^{n^q}$ назовем вектор-функцию $\overline{f}^q(x^q) = f^q(p_{S^q}(x^q))$. Рассмотрим задачу: найти

$$\min \overline{f}^{q^0}_{j^0}(x^{q^0}) \tag{8}$$

при ограничениях

$$\overline{f}^q_j(x^q) \le 0, \quad (j,q) \in T, \tag{9}$$

$$x_i^q = \overline{f}^p_j(x^p), \quad (i,j) \in M(p,q),\ (p,q) \in E, \tag{10}$$

$$d_{S^q}(x^q) = 0,\ q \in V. \tag{11}$$

Поскольку $\overline{f}^q(x^q) = f^q(x^q)$ при условии $d_{S^q}(x^q) = 0$, $q \in V$, то решения задач (4)–(7) и (8)–(11) совпадают.

Пусть сеть вычислительных блоков $(V,E)$ ациклическая, $x$ – множество входов сети. Учитывая (10), представим выходы $\overline{f}^q(x^q)$ каждого блока как функцию от входов сети. Такое представление назовем расширенной редукцией сети блоков.

Обозначим $x^q(x)$ значения входов блока $B^q$ для заданного входа $x$ сети $(V,E)$, полученные при расширенной редукции сети блоков. Положим $\varphi^q(x) = \overline{f}^q(x^q(x))$, $\delta^q(x) = d_{S^q}(x^q(x))$. Тогда задачу (8)–(11) можно представить в виде: найти

$$\min \varphi^{q^0}_{j^0}(x), \tag{12}$$

при ограничениях

$$\varphi^q_j(x) \le 0, \ (j,q) \in T, \tag{13}$$

$$\delta^q(x) \le 0, \ q \in V. \tag{14}$$

Задачу (12)–(14) будем называть расширенной редукцией исходной задачи (4)–(7). Функции $\varphi^q(x)$, $\delta^q(x)$ определены при любых $x$, и для решения задачи (12)–(14) может использоваться метод негладких штрафов.

Для разработки алгоритмов решения задачи (12)–(14) необходимо описать свойства функций $\varphi_j^q(x)$, $\delta^q(x)$ и получить соотношения, позволяющие вычислять градиенты этих функций.

Рассмотрим функцию

$$d_S(x) = \min_{z\in S} \| x - z \|, \tag{15}$$

где $x, z \in R^n$, $S \subset R^n$ – замкнутое выпуклое множество. Свойства этой функции исследовались в различных работах (см., например, [5]). Функция $d_S(x)$ выпуклая и непрерывна. Если $x \in \text{int } S$, то $d_S(x) = 0$, $\nabla d_S(x) = 0$. В граничных точках множества $S$ функция $d_S(x)$ недифференцируема. В точках $x \notin S$ функция $d_S(x)$ дифференцируема и

$$\nabla d_S(x) = (x - \bar{x}) / d_S(x) . \tag{16}$$

Рассмотрим отображение

$$p_S(x) = \arg\min_{z\in S} \| x - z \| . \tag{17}$$

Обозначим $\bar{x} = p_S(x)$, $p'_S(x,d)$ – производная отображения $p_S(x)$ по направлению $d$ в точке $x$ (если такая производная существует), $\Gamma(z)$ – конус возможных направлений [5] множества $S$ в точке $z \in S$.

**Теорема 1**. Отображение $p_S(x)$ непрерывно и дифференцируемо по направлениям в точке $x$. При этом

1) если $x \in S$, то $p'_S(x,d) = p_{\Gamma(x)}(d)$;

2) если $x \notin S$, то $p'_S(x,d) = p_{T(x)}(d)$, где $T(x) = \left\{ v : (x - \bar{x},\, v) = 0,\ v \in \Gamma(\bar{x}) \right\}$.

Ввиду ограниченности объема статьи доказательства теорем не приводим.

Пусть множество $S$ описывается следующим образом:

$$S = \left\{ x \in R^n : \ h_i(x) \le 0,\ i = 1, ..., m \right\}, \tag{18}$$

где $h_i(x)$ – выпуклые непрерывно дифференцируемые функции.

Обозначим $I(x)$ множество ограничений, активных в точке $x \in S$, $I(x) = \{ i \in \{1, ..., m\} : h_i(x) = 0 \}$.

Пусть $x \in S$, градиенты $\nabla h_i(x),\ i \in I(x)$ линейно независимы, тогда

$$\Gamma(x)=\left\{v\in R^L : (\nabla h_i(x),\, v)\le 0,\ i\in I(x)\right\}.$$

Точка $x\notin S$ регулярно разложима относительно множества $S$, если градиенты $\nabla h_i(\bar{x}),\ i\in I(\bar{x})$ линейно независимы и в разложении $x-\bar{x}=\sum_{i\in I(\bar{x})}\gamma_i\nabla h_i(\bar{x})$ все коэффициенты отличны от нуля: $\gamma_i>0,\ i\in I(\bar{x})$ .

Если точка $x\notin S$ регулярно разложима относительно множества $S$ , то

$$T(x)=\left\{v : (\nabla h_i(\bar{x}),\, v)=0,\ i\in I(\bar{x})\right\}. \tag{19}$$

Для $I(x)\neq\varnothing$ обозначим $H(x)$ матрицу, строками которой являются градиенты $\nabla h_i(x),\ i\in I(x)$ .

**Теорема 2**. Пусть точка $x\notin S$ регулярно разложима относительно множества $S$, $\bar{x}=p_S(x)$, $H=H(\bar{x})$. Тогда отображение $p_S(x)$ дифференцируемо в точке $x$ и

$$p'_S(x,d)=\left(I-H^T(HH^T)^{-1}H\right)d\,. \tag{20}$$

Пусть $f(x)$ – непрерывно дифференцируемая функция, $x\in S\subset R^n$ . Рассмотрим расширенную функцию $\bar{f}(x)=f(p_S(x))$ , $\bar{x}=p_S(x)$ .

**Теорема 3**. Функция $\bar{f}(x)$ непрерывна и дифференцируема по направлениям. Если точка $x$ регулярно разложима относительно множества $S$ , то $\bar{f}(x)$ в этой точке дифференцируема и

$$\nabla\bar{f}(x)=\left(I-H^T(HH^T)^{-1}H\right)\nabla f(\bar{x}), \tag{21}$$

**Теорема 4.** Пусть $x\notin S$ , в точке $\bar{x}=p_S(x)$ градиенты $\nabla h_i(\bar{x}),\ i\in I(\bar{x})$ линейно независимы. Тогда для любого $\varepsilon>0$ найдется регулярно разложимая точка $x'$, такая что $\|x'-x\|<\varepsilon$.

Вернемся к задачам (8)–(11) и (12)–(14). Градиенты функций $\bar{f}^q(x^q)$ , $d_{S^q}(x^q)$ в регулярно разложимых точках вычисляются в соответствии с соотношениями (21), (16). Градиенты функций $\varphi^q(x)$ , $\delta^q(x)$ вычисляются по правилам дифференцирования сложных функций.

Следуя [6], определим почти-градиент функции $\varphi(x)$, в точке $x$ как предел некоторой последовательности градиентов $\left\{\nabla\varphi(x^k)\right\}_{k=1}^{\infty}$, где $\left\{x^k\right\}_{k=1}^{\infty}$ – последовательность точек, сходящаяся к $x$ , и такая, что во всех точках этой последовательности функция $\varphi(x)$ дифференцируема. В качестве приближения к почти-градиенту в точке $x$ можно взять градиент $\nabla\varphi(x^k)$ в точке $x^k$, достаточно близкой к $x$.

Для решения задачи (12)–(14) используется метод негладких штрафов, основанный на алгоритме с растяжением пространства, предложенном в [6], для минимизации почти-дифференцируемых функций.

**Программная реализация и вычислительные эксперименты.** Для решения задач (4)–(7) были разработаны программные реализации прямой и расширенной редукции исходной задачи.

Для расширенной редукции исходной задачи реализован случай, когда область определения функций каждого блока $S^q$ есть пересечение шара и множества $\left\{x^q : x_l^q \le x^q \le x_u^q\right\}$, где $x_l^q$, $x_u^q$ – заданные границы.

Обе реализации написаны на языке Python и включены в состав программных средств OpenOpt (http://www.openopt.org), разрабатываемых как свободная альтернатива коммерческим средам AMPL, GAMS, TOMLAB/TOMNET, и др. Также в OpenOpt включены модификация $r$ -алгоритма (http://openopt.org/ralg) и реализация метода негладких штрафов, основанная на $r$ -алгоритме из [4, 6].

Тестовые задачи описываются следующим образом. Задана $n$-звенная цепочка. Звенья соединены шарнирами. Трение в шарнирах отсутствует. Массы сосредоточены в шарнирах (масса звена равна 0). Введем обозначения:

$l_k$ – длина звена $k$, $i = k,..., n$, (длина звена не зависит от натяжения);

$m_k$ – масса шарнира $k$, $k = 1,..., n-1$;

$\overline{F}_k$ – прочность звена $i$ (при натяжении более $\overline{F}_k$ звено разрушается);

$z_k = (x_k,\ y_k)$ – координаты $k$ -го шарнира;

$z_0 = (x_0,\ y_0)$ – координаты начальной точки первого звена;

$z_{n+1} = (x_{n+1},\ y_{n+1})$ – координаты конечной точки последнего звена;

Цепочка подвешена за концы в точках $z_0 = (x_0,\ y_0)$, $z_{n+1} = (x_{n+1},\ y_{n+1})$ в однородном поле тяжести. Длины и прочность звеньев считаются заданными.

В задаче предполагается, что точка $z_0 = (x_0,\ y_0)$ зафиксирована. Необходимо найти точку $z_{n+1} = (x_{n+1},\ y_{n+1})$ с максимальным значением $x_{n+1}$ при условии, что цепочка не разрушается и суммарная масса шарниров не меньше заданной величины.

Чтобы представить задачи в виде (4)–(7), поставим в соответствие $k$ -му шарниру (соединяет звенья $k$ и $k+1$) блок $B^k$, входами которого являются координаты $z_k = (x_k,\ y_k)$, масса $m_k$ шарнира $k$ и сила (вектор) $F_k = (F_k^x, F_k^y)$, действующая на шарнир $k$ со стороны звена $k$. Выходами блока $B^k$ будут сила $F_{k+1}=(F_{k+1}^x, F_{k+1}^y)$, действующая на шарнир $k+1$ со стороны звена $k+1$ и координаты $z_{k+1}=(x_{k+1},\ y_{k+1})$ конца звена $k+1$ (координаты шарнира $k+1$).

Очевидно, что выходы блока $B^k$ определяются соотношениями

$$F_k + m_k G - F_{k+1} = 0,$$

$$z_{k+1} = z_k - l_k \frac{F_{k+1}}{\|F_{k+1}\|},$$

где первое соотношение – равенство нулю сил, действующих на шарнир $k$, вектор $G=(0,\ g)$ – ускорение силы тяжести. Выходы блока $B^k$ определены в области

$$S^k = \left\{(F_k^x, F_k^y) : (F_k^x)^2 + (F_k^y)^2 \le \overline{F}_k^{\,2}\right\}.$$

Понятно, что при фиксированных значениях $z_0=(x_0,\ y_0)$, $F_1=(F_1^x, F_1^y)$ и массах $m_k$ $k$, $k=1,\ldots,\ n-1$, координаты всех шарниров и натяжение каждого звена определяются однозначно, если звенья не разрушаются, т.е. число переменных в редуцированных задачах равно $n+1$.

Сравнительный анализ проводился для следующих бесплатных решателей задач нелинейного программирования, подключенных к OpenOpt – IPOPT, COBYLA, ALGENCAN, slsqp

Тестирование этих решателей проводилось на прямой редукции задачи о цепочке, являющейся гладкой. При тестировании расширенной редукции использовался метод негладких штрафов, использующий $r$-алгоритм.

Результаты вычислительных экспериментов приведены в таблице. Поскольку наилучшими решателями оказались IPOPT и метод негладких штрафов для расширенной редукции (ralg в приведенной таблице), то результаты приведены только для этих решателей. Вариант 2 отображает результаты расчетов при введении дополнительных ограничений $y_k \le \overline{y}_k$ в область $S^k$ для каждого шарнира. При этом значения $\overline{y}_k$ подобраны так, что в оптимальном решении активно несколько этих ограничений. В результате IPOPT останавливается практически в стартовой точке. Отметка Fail обозначает, что задача не решена.

В целом программы негладкой оптимизации работают более устойчиво, по быстродействию соизмеримы с реализациями гладких методов (бесплатными).

| | Вариант 1 | | | | Вариант 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | ralg | | IPOPT | | ralg | | IPOPT | |
| $N$ var | $f^*$ | $T$ (сек.) | $f^*$ | $T$ (сек.) | $f^*$ | $T$ (сек.) | $f^*$ | $T$ (сек.) |
| 15 | 5,98 | 12,51 | 6,29 | 1,83 | 4,47 | 21,46 | 0,16 | 0,07 |
| 16 | 6,378 | 10,01 | 6,71 | 1,46 | 4,637 | 31,08 | 0,19 | 0,07 |
| 17 | 6,13 | 5,36 | 7,11 | 1,7 | 4,79 | 103,44 | 0,26 | 0,07 |
| 18 | 7,51 | 15,54 | 7,58 | 1,27 | 4,91 | 18,84 | 0,61 | 0,07 |
| 19 | 8,04 | 12,92 | 8,14 | 1,83 | 4,64 | 34,59 | 0,34 | 0,08 |
| 20 | 8,69 | 42,15 | 8,71 | 3,43 | 4,99 | 16,6 | 0,25 | 0,08 |
| 21 | 9,19 | 25,59 | 9,21 | 3,33 | 5,53 | 26,85 | 0,23 | 0,08 |
| 22 | 9,26 | 34,86 | 9,60 | 21,72 | 5,68 | 45,06 | 0,23 | 0,09 |
| 23 | 9,83 | 12,65 | 9,94 | 2,88 | 5,79 | 24,2 | 0,231 | 0,09 |
| 24 | 10,26 | 19,66 | 10,33 | 19,61 | 5,90 | 127,47 | 0,25 | 0,1 |
| 25 | 10,57 | 27,52 | 10,79 | 8,8 | 6,05 | 125,89 | Fail | Fail |

*Ю.П. Лаптін, Д.Л. Крошко*

ДЕЯКІ НЕЛІНІЙНІ ОПТИМІЗАЦІЙНІ ЗАДАЧІ СІТКОВОЇ СТРУКТУРИ

Розглядаються задачі, які описуються сукупністю взаємопов'язаних функціональних блоків. Деякі виходи одних блоків можуть бути входами для інших блоків. Функціональні залежності блоків визначені на обмежених множинах. Розроблено підхід до вирішення задач, заснований на довизначенні функцій на весь простір.

*Yu.P.Laptin, D.L.Kroshko*

SOME NETWORK NONLINEAR OPTIMIZATION PROBLEMS

The problems under consideration are described by a set of interconnected functional blocks. Some block outputs may be used as inputs for other blocks. Functional dependencies of blocks are defined on bounded areas. Problem solving approach is based on function redefining at full space.

1. *Айда-заде К.Р.* Исследование нелинейных оптимизационных задач сетевой структуры // Автоматика и телемеханика – 1990. – № 2. –С. 3 – 14.
2. *Теплосиловые* системы: Оптимизационные исследования. – Новосибирск: Наука, 2005. – 236 с.
3. *Лаптин Ю.П., Журбенко Н.Г., Левин М.М., Волковицкая П.И.* Использование средств оптимизации в системе автоматизированного проектирования энергетических котлоагрегатов КРОКУС // Энергетика и электрификация. – 2003. – № 7. – С. 41 – 51.
4. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – London: Kluwer Academic Publishers, 1998. – 381 p.
5. *Демьянов В.Ф., Васильев Л.В.* Недифференцируемая оптимизация. – М.: Наука, 1981. – 384 с.
6. *Шор Н.З.* О классе почти-дифференцируемых функций и одном методе минимизации функций этого класса // Кибернетика – 1972. – № 4. – С. 9 – 17.