Рассматриваются оптимизационные задачи, возникающие в ходе моделирования сложных технических объектов. Такие задачи включают в себя хорошо структурированные подсистемы нелинейных ограничений-равенств. Предлагается подход, позволяющий осуществить декомпозицию рассматриваемых задач.

УДК 519.8

Ю.П. ЛАПТИН

РЕШЕНИЕ ОДНОГО КЛАССА ЗАДАЧ ОПТИМИЗАЦИИ С НЕЛИНЕЙНЫМИ ОГРАНИЧЕНИЯМИ

Оптимизационные задачи, возникающие в ходе моделирования сложных технических объектов, обычно имеют специальную структуру, которую необходимо учитывать при разработке программных средств их решения. Одной из особенностей таких задач является наличие структурированных нелинейных ограничений-равенств. Такие ограничения часто распадаются на небольшие, слабо связанные, зависящие от параметров подсистемы нелинейных уравнений, которые в отдельных случаях удается эффективно решать на каждой итерации оптимизационных алгоритмов [1, 2]. В результате значительно понижается размерность оптимизационной задачи. В общем случае возникают проблемы, обусловленные тем, что нелинейные функции обычно определены на ограниченных областях, а подсистемы нелинейных уравнений имеют решения не при всех возможных значениях параметров.

В работе вводится понятие регуляризованного решения системы нелинейных уравнений, зависящих от параметров. Рассматриваются системы уравнений, состоящие из совокупности взаимосвязанных блоков, случай, когда решение системы может быть получено в результате последовательного решения подсистем уравнений блоков. Для построения регуляризованного решения блока необходимо решать гладкую оптимизационную задачу. Формулируется редуцированная задача, эквивалентная исходной, исследуются свойства введенных функций. Редуцированная задача оказывается негладкой невыпуклой задачей. Обсуждается возможность использования некоторых алгоритмов для решения сформулированных задач.

1. Рассмотрим задачу математического программирования: найти

$$\min f_0(x, y) \tag{1}$$

при ограничениях

$$f_k(x, y) \le 0,\ k = 1,\dots, K. \tag{2}$$

$$\chi_i(x, y) = 0,\ i = 1,\dots, m, \tag{3}$$

где $x \in R^n$, $y \in R^m$, $f_k(x, y)$, $\chi_i(x, y)$ - непрерывно дифференцируемые функции $k = 0,\dots, K$, $i = 1,\dots, m$.

Предположим, что система (3) разрешима относительно y при любых значениях вектора x. Если система (3) имеет специальную структуру, это может быть использовано для построения эффективных алгоритмов ее решения. В этом случае для текущих значений вектора x целесообразно найти решение $y(x)$ системы (3) и вместо задачи (1) - (3) рассмотреть следующую: найти

$$\min\left\{ \varphi_0(x) \colon \varphi_k(x) \le 0,\ x \in R^n,\ k = 1,\dots, K\right\},$$

где $\varphi_k(x) = f_k(x,\ y(x)),\ k = 0,\dots, K$.

Если система (3) линейная, т.е. имеет вид

$$Ax + By + b = 0, \tag{3'}$$

то специальная структура может выражаться в том, что матрица B имеет блочный, треугольный или блочно-треугольный вид.

Рассмотрим нелинейный случай. Будем называть блоком B^q совокупность уравнений вида

$$\chi_m(z^q, y^q) = 0,\ m \in M^q, \tag{3''}$$

если система (3'') при фиксированных значениях z^q может быть разрешена относительно y^q, где $z^q = (z_j,\ j \in J^q)$, $y^q = (y_i,\ i \in I^q)$, $\left|I^q\right| = \left|M^q\right|$. Решение системы (3'') обозначим $y^q(z^q)$. Вектор z^q назовем входом блока B^q, $y^q = y^q(z^q)$ - выходом блока B^q, J^q - множество входных переменных, I^q - множество выходных переменных.

Обозначим I совокупность индексов переменных y исходной задачи. Будем говорить, что система (3) представима некоторой сетью блоков, если совокупность уравнений (3) может быть разбита на непересекающиеся подмножества, каждое из которых является блоком. При этом для множества выходных переменных каждого блока B^q имеет место $I^q \subseteq I$, множества выходных переменных разных блоков непересекаются. Компонентами входов каждого блока могут быть как компоненты вектора x исходной задачи, так и компоненты выходов других блоков. Блоки B^q, B^p связаны дугой $(q,\ p)$, если некоторые выходы блока B^q являются входами блока B^p.

Пусть система уравнений (3) представима ациклической сетью блоков. Вектор x будем называть входом, вектор y - выходом сети блоков. Очевидно, что ациклическая сеть определяет некоторое упорядочение блоков (возможно, неединственное). Нетрудно видеть, что последовательное решение (в соответствии с упорядочением сети) систем уравнений блоков дает решение исходной системы уравнений (3).

2. В общем случае система (3") имеет решение не при всех значениях z^q. Существенной особенностью нелинейных функций является то, что они могут быть определены на ограниченных множествах. Предположим, что при формировании каждого блока явно описываются области, в которых определены функции, входящие в систему уравнений блока, и которым должно принадлежать решение, т.е. описание блока B^q имеет вид

$$\chi_m(z^q, y^q) = 0,\ m \in M^q, \tag{4}$$

$$l_s(z^q,\ y^q) \le 0,\ s \in S^q. \tag{5}$$

Обозначим

$$V^q = \left\{ (z^q,\ y^q) : l_s(z^q,\ y^q) \le 0,\ s \in S^q \right\}. \tag{6}$$

Предполагается, что функции $l_s(z^q, y^q)$ - линейны, функции $\chi_m(z^q, y^q)$ определены на множестве V^q, т.е. $V^q \subset \operatorname{int} dom\, \chi_m$, $m \in M^q$, якобиан $\det J_y(z^q, y^q) = \det\left\{\dfrac{\partial \chi_m}{\partial y_i}\right\}$, $m \in M^q$, $i \in I^q$ отличен от нуля для всех точек $(z^q, y^q) \in V^q$.

Введем вспомогательные переменные $v^q = (v_j^q,\ j \in J^q)$ и рассмотрим вспомогательную задачу: найти

$$\delta^q(z^q) = \min_{v,y} \left\| z^q - v^q \right\|, \tag{7}$$

$$\chi_m(v^q, y^q) = 0,\ m \in M^q, \tag{8}$$

$$l_s(v^q,\ y^q) \le 0,\ s \in S^q. \tag{9}$$

Очевидно, что если при некоторых значениях $\bar{z}^q$, решение системы (4) - (5) существует, то задача (7) - (9) имеет решения при любых z^q, функция $\delta^q(z^q)$

определена для всех z^q. Если при заданном $\bar{z}^q$ система (4) - (5) разрешима, то $\delta^q(\bar{z}^q) = 0$, иначе $\delta^q(\bar{z}^q) > 0$.

Пусть (v^{*q}, y^{*q}) – решение задачи (7) - (9) при заданном z^q. Положим $\theta^q(z^q) = y^{*q}$. Вектор-функцию $\theta^q(z^q)$ будем называть регуляризованным решением блока.

Пусть задана ациклическая сеть блоков. При последовательном расчете сети для заданного входа x используем регуляризованные решения блоков.

Обозначим: $\theta(x)$ - «регуляризованное» решение сети блоков (исходной системы уравнений); $\varphi_k(x) = f_k(x, \theta(x)),\ k = 0,..., K$; $z^q(x)$ - зависимость входа z^q блока B^q от входа x сети блоков (полученная в результате последовательного расчета сети); $\mu^q(x) = \delta^q(z^q(x)),\ q \in Q$.

Очевидно, что исходная задача (1) - (3) с учетом дополнительных ограничений (5) эквивалентна следующей: найти

$$\varphi_0^* = \min \varphi_0(x) \tag{10}$$

при ограничениях

$$\varphi_k(x) \le 0,\ k = 1,..., K, \tag{11}$$

$$\mu^q(x) \le 0,\ q \in Q. \tag{12}$$

3. Рассмотрим функцию $\delta^q(z)$, которая непрерывна и определена при всех z. Соотношение (8) определяет зависимость $y^q = y^q(v^q)$. Обозначим $\upsilon_s(v^q) = l_s(v^q,\ y^q(v^q))$, и приведем (7) - (9) к виду: найти

$$\delta^q(z^q) = \min_v \left\| z^q - v^q \right\|, \tag{13}$$

$$\upsilon_s(v^q) \le 0,\ s \in S^q. \tag{14}$$

Обозначим: $v^q(z^q)$ - решение задачи (13) - (14) при заданном z^q; $D^q = \left\{ v^q :\ \upsilon_s(v^q) \le 0,\ s \in S^q \right\}$.

Имеет место (см. [6]) следующая

Теорема. Пусть при заданном z^q выполняется $\delta^q(z^q) > 0$, задача (13) - (14) имеет единственное решение (проекция точки z^q на допустимое множество D^q). Тогда функция δ^q дифференцируема в точке z^q, градиент g^δ этой функции лежит на прямой, порожденной отрезком $[z^q - v^q(z^q)]$, и $\left\| g^\delta \right\| = 1$.

Следствие. Пусть множество D^q выпуклое. Тогда функция δ^q - выпуклая и непрерывно дифференцируема в точках z^q , таких, что $\delta^q(z^q) > 0$.

Будем говорить, что точка z^q регулярна, если для этого z^q задача (13) - (14) имеет единственное решение.

В общем случае множество D^q невыпуклое. Будем говорить, что множество D^q обладает свойством Δ-регулярности, если все точки z^q , такие что $\delta^q(z^q) \le \Delta$, являются регулярными. Т.е. функция δ^q непрерывно дифференцируема в точках z^q , таких, что $0 < \delta^q(z^q) \le \Delta$.

Приведем без доказательства

Утверждение. Пусть функции $\upsilon_s(v^q)$, $s \in S^q$, непрерывно дифференцируемы, градиенты $g^{\upsilon_s}(v^q)$ функций $\upsilon_s(v^q)$ удовлетворяют условию Липшица с некоторой константой L. Существует $\varepsilon > 0$, такое, что для любой точки v^q , в которой $\upsilon_s(v^q) = 0$, выполняется $\left\| g^{\upsilon_s}(v^q) \right\| > \varepsilon$. Тогда существует $\Delta > 0$, такое, что множество D^q обладает свойством Δ - регулярности.

Замечание. В регулярных точках z^q решение $v^q(z^q)$ задачи (13) - (14) непрерывно, но может быть негладким.

4. Предложенная схема решения задач оптимизации при наличии структурированных нелинейных ограничений-равенств позволяет эффективно решать систему нелинейных уравнений при любых значениях параметров x. Вспомогательные задачи (7) - (9) являются гладкими (при условии гладкости исходных функций). Для их решения могут использоваться методы, которые гарантируют выполнение линейных ограничений (9) на каждой итерации (например, некоторые модификации метода линеаризации [5]). Редуцированная задача (10) - (12) является негладкой, размерность ее относительно исходной задачи существенно понижается. В точках x, таких что значения $\mu^q(x)$, $q \in Q$, достаточно малы, функции, входящие в задачу, непрерывны. Если значения $\mu^q(x)$ велики, эти функции могут быть разрывными. Для решения редуцированной задачи могут использоваться алгоритмы с растяжением пространства [3, 4]. Ограничения (11) - (12) при этом должны вводиться в целевой функционал как негладкие штрафы.

Эффективность предложенного подхода и круг задач, для которого применение этого подхода целесообразно, могут быть определены в результате вычислительных экспериментов.

Ю.П. Лаптін

РОЗВ'ЯЗАННЯ ОДНОГО КЛАСУ ЗАДАЧ ОПТИМІЗАЦІЇ З НЕЛІНІЙНИМИ ОБМЕЖЕННЯМИ

Розглядаються оптимізаційні задачі, які виникають при моделюванні складних технічних об'єктів. Такі задачі включають в себе добре структуровані системи нелінійних обмежень-рівнянь. Пропонується підхід, який дозволяє виконувати декомпозицію задач, що розглядаються.

Yu.P. Laptin

SOLVING OF ONE CLASS OF OPTIMIZATION PROBLEMS WHICH CONTAIN NONLINEAR CONSTRAINTS

Optimization problems are discussed which arise when complicated technical objects are modeled. Such problems contain well structured sets of nonlinear equations. Some approach is proposed for decomposition of discussed problems.

1. *Лаптин Ю. П., Журбенко Н. Г.* Разработка программных средств оптимизации сложных технических объектов // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2002. – С. 3 – 12.
2. *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.
3. *Shor N. Z.* Nondifferentiable Optimization and Polynomial Problems. – London: Kluwer Academic Publishers, 1998. – 381 p.
4. *Шор Н.З., Журбенко Н.Г.* Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51 – 59.
5. *Пшеничный Б. Н.* Метод линеаризации. – М.: Наука, 1983. – 136 с.
6. *Ж. - Б. Ириарт – Уррути.* Оптимизация и выпуклый анализ. – Киев: Издательская компания «КИТ», 2004. – 370 с.

Получено 27.03.2005