УДК 519.8

Ю.П. ЛАПТИН

ОПТИМИЗАЦИОННЫЕ ЗАДАЧИ С БЛОЧНОЙ СИСТЕМОЙ НЕЛИНЕЙНЫХ ОГРАНИЧЕНИЙ-РАВЕНСТВ

Рассматриваются задачи оптимизации с блочной системой нелинейных ограничений-равенств. Вводится понятие псевдорешения системы уравнений блока, что позволяет свести исходную оптимизационную задачу к редуцированной оптимизационной задаче меньшей размерности. Для ее решения предлагается использовать методы негладких штрафных функций и методы негладкой оптимизации. Исследуются свойства функций редуцированной задачи, определены правила вычисления градиентов в точках, в которых функции дифференцируемы.

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. Подсистемы нелинейных уравнений при фиксированных значениях параметров могут иметь множество решений, из которых необходимо найти одно, принадлежащее достаточно узкой области. В этой области должны выполняться условия сходимости используемых методов.

Вводится понятие псевдорешения системы уравнений блока, что позволяет свести исходную оптимизационную задачу к редуцированной оптимизационной задаче меньшей размерности. Для ее решения предлагается использовать методы негладких штрафных функций и методы негладкой оптимизации [3, 4].

Исследуются свойства функций редуцированной задачи. Эти функции оказываются дифференцируемыми по направлению, негладкими и невыпуклыми. Получены условия регулярности, при выполнении которых функции дифференцируемы в точке, определены правила вычисления градиентов в точках, в которых функции дифференцируемы.

1. Рассматривается задача математического программирования следующего вида: найти

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

при ограничениях

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

$$g^q(x, y^q) = 0,\ q = 1,..., Q, \tag{3}$$

$$A_x^q x + A_y^q y^q \le b^q,\ q = 1,..., Q, \tag{4}$$

где $x \in E^L$, $y^q \in E^{N_q}$, функции $g^q(x, y^q): E^L \times E^{N_q} \to E^{N_q}$, $f_k(x, y^1, y^2,..., y^Q)$: $E^L \times E^{N_1} \times ... \times E^{N_Q} \to E^1$, предполагаются непрерывными и непрерывно дифференцируемыми; $b^q \in E^{m_q}$, матрицы A_x^q, A_y^q имеют соответствующие размерности.

Предполагается, что функции f_k определены при любых значениях аргументов, допустимые точки системы неравенств (4) при фиксированном q принадлежат области определения функции g^q, которая может не совпадать с пространством $E^L \times E^{N_q}$.

Более того, предположим, что для любого фиксированного q, если при некотором x подсистема (3) имеет решение, удовлетворяющее ограничениям (4), то такое решение единственное.

Систему уравнений (3) и ограничений (4) для каждого q будем называть блоком B^q. Обозначим S^q множество точек x, для которых система (3), (4) имеет решение, $y^q(x)$ – решение системы (3) в точке $x \in S^q$.

Очевидно, что следующая задача эквивалентна исходной (1)–(4): найти

$$f_0^* = \min f_0(x, y^1(x),..., y^Q(x)) \tag{5}$$

при ограничениях

$$f_k(x, y^1(x),..., y^Q(x)) \le 0,\ k = 1,..., K, \tag{6}$$

$$x \in S^q,\ q = 1,..., Q. \tag{7}$$

Для решения задачи (5) – (7) будем использовать метод негладких штрафных функций. Обозначим $p_{S^q}(x)$ – проекция точки $x \in E^L$ на множество S^q, $q = 1,..., Q$, $d_{S^q}(x) = \left\| x - p_{S^q}(x) \right\|$ – расстояние от точки x до множества S^q.

Псевдорешением системы (3), (4) в произвольной точке x назовем отображение $\tilde{y}^q(x) = y^q(p_{S^q}(x))$. Пусть заданы неотрицательные векторы $\alpha = (\alpha_1, ..., \alpha_K)$, $\beta = (\beta_1,..., \beta_Q)$. Обозначим $\varphi_k(x) = f_k(x, \tilde{y}^1(x),..., \tilde{y}^Q(x))$, $k = 0, ..., K$ и рассмотрим задачу: найти

$$\varphi_0^*(\alpha,\beta) = \min_x \left\{ \varphi_0(x) + \sum_{k=1}^{K} \alpha_k \left[\varphi_k(x)\right]^+ + \sum_{q=1}^{Q} \beta_q d_{S^q}(x) \right\}. \tag{8}$$

Здесь $\left[\varphi_k(x)\right]^+ = \begin{cases} \varphi_k(x), & \text{если } \varphi_k(x) \ge 0, \\ 0, & \text{если } \varphi_k(x) < 0. \end{cases}$

Метод негладких штрафных функций основан на том, что при достаточно больших, но конечных, значениях коэффициентов α, β решения задач (5)–(7) и (8) совпадают. Условия, при которых решения этих задач совпадают, следует рассматривать отдельно.

Пусть зафиксировано некоторое $q \in \{1,..., Q\}$. Для блока B^q введем вспомогательные переменные $z^q \in E^L$ и рассмотрим задачу: найти

$$d^q(x) = \min_{z,y} \left\| x - z^q \right\|, \tag{9}$$

$$g^q(z^q, y^q) = 0 , \tag{10}$$

$$A_x^q z^q + A_y^q y^q \le b^q . \tag{11}$$

Очевидно, что если при некотором $\bar{x}$, решение системы (3), (4) существует, то задача (9)–(11) имеет решения при любых x, функция $d^q(x)$ определена для всех x. Если при заданном x система (3), (4) разрешима, то $d^q(x) = 0$, иначе – $d^q(x) > 0$.

Пусть (z^{*q}, y^{*q}) – решение задачи (9)–(11) при заданном x. Легко видеть, что $d_{S^q}(x) = d^q(x)$, $S^q = \left\{x : d^q(x) = 0\right\}$, $p_{S^q}(x) = z^{*q}$.

Очевидно, что множество S^q в общем случае не выпукло, задача (9)–(11) – многоэкстремальна.

Невыпуклость множеств S^q порождает существенные проблемы как при решении задач (9)–(11), так и при решении задачи (5)–(7).

Подход, позволяющий несколько облегчить возникающие проблемы, основан на том, что при решении многих практических задач приемлемым является некоторое сужение областей локализации, путем введения дополнительных линейных ограничений в (11) , не зависящие от переменных y^q и обеспечивающие выпуклость множеств S^q .

Задача построения вспомогательных ограничений требует отдельного рассмотрения. Такие ограничения должны вводиться при участии специалистов в конкретных прикладных областях. Для выполнения этих построений следует разрабатывать специальное программное обеспечение. При малых размерностях блоков q (систем ограничений (10), (11)) задача построения вспомогательных ограничений может иметь приемлемую трудоемкость.

Для разработки алгоритмов решения задачи (8) необходимо изучить свойства функций $\varphi_k(x)$, $k = 0,..., K$, $d_{S^q}(x)$ и получить соотношения, позволяющие вычислять градиенты этих функций.

2. Рассмотрим задачу: найти

$$d_S(x) = \min_{z\in S} \| x - z \| , \tag{12}$$

где $x, z \in E^L$, $S \subset E^L$ – замкнутое множество (не обязательно выпуклое).

Функция $d_S(x)$ – непрерывна. Если $x \in \operatorname{int} S$, то $\nabla d_S(x) = 0$. В граничных точках множества S функция $d_S(x)$ недифференцируема.

Имеет место следующая лемма (см., например, [6])

Лемма 1. Пусть $x \notin S$, минимум в задаче (12) достигается в единственной точке $\bar{x}$. Тогда функция $d_S(x)$ дифференцируема в точке x и

$$\nabla d_S(x) = (x - \bar{x}) \big/ d_S(x) . \tag{13}$$

Если S – выпуклое множество, то $d_S(x)$ – выпуклая функция.

Пусть множество S описывается следующим образом:

$$S = \left\{ z \in E^L : \ h_i(z) \le 0, \ i = 1,..., m \right\} , \tag{14}$$

где $h_i(z)$ – непрерывно дифференцируемые функции.

Для $z \in S$ обозначим:

$I(z)$ – множество индексов ограничений, активных в точке z, т.е. $I(z) = \left\{ i \in \{1,..., m\} : h_i(z) = 0 \right\}$,

$\Gamma(z),\ z\in S$ – конус возможных в широком смысле [6] направлений множества S в точке z.

Известно [5], что если градиенты $\nabla h_i(z),\ i\in I(z)$ – линейно независимы, то

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

Пусть $I(z)\ne\varnothing$. Обозначим H матрицу, строками которой являются градиенты $\nabla h_i(z),\ i\in I(z)$.

Будем говорить, что

точка $x\notin S$ регулярна относительно множества S, если минимум в задаче (12) достигается в единственной точке $\bar{x}=p_S(x)$;

точка $x\notin S$ – регулярно разложима относительно множества S, если она регулярна, градиенты $\nabla h_i(\bar{x}),\ i\in I(\bar{x})$ – линейно независимы, в разложении $x-\bar{x}=\sum\limits_{i\in I(\bar{x})}\gamma_i\nabla h_i(\bar{x})$ все коэффициенты не равны нулю: $\gamma_i>0,\ i\in I(\bar{x})$.

Внутренние точки множества S будем считать регулярными и регулярно разложимыми относительно множества S.

Рассмотрим отображение $p_S(x)=\arg\min\limits_{z\in S}\|x-z\|$. 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).

Теорема 1. Пусть точка $x\notin S$ регулярна относительно множества S. Тогда

$p_S(x)$ непрерывно и дифференцируемо по направлениям в точке x;

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

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

Обозначим H матрицу, строками которой являются градиенты $\nabla h_i(\bar{x}),\ i\in I(\bar{x})$.

Лемма 2. Пусть точка $x\notin S$ – регулярно разложима относительно множества S. Тогда отображение $p_S(x)$ дифференцируемо в точке x,

$$p'_S(x,d)=\left(I-H^T(HH^T)^{-1}H\right)d. \qquad (15)$$

3. Рассмотрим задачу (9)–(11) для отдельного блока. Индекс q (номер блока) в пределах данного пункта будем опускать.

Пусть система уравнений

$$g(x,y)=0 \tag{16}$$

при $x=\bar{x}$ имеет решение $\bar{y}=y(\bar{x})$. Обозначим $G_x=\left\|\dfrac{\partial g_i}{\partial x_k}\right\|$, $G_y=\left\|\dfrac{\partial g_i}{\partial y_j}\right\|$, $i,j=1,\ldots,N$, $k=1,\ldots,L$ – матрицы частных производных вектор-функции g в точке $(\bar{x},\bar{y})$. Пусть $\left|G_y\right|\neq 0$, тогда в окрестности точки $\bar{x}$ существует решение $y(x)$, непрерывное и дифференцируемое в точке $\bar{x}$. При этом производная $y'(\bar{x},d)$ по направлению d отображения $y(x)$ имеет вид

$$y'(\bar{x},d)=-G_y^{-1}G_x d\,. \tag{17}$$

Рассмотрим множество S. Учитывая, что $y(z)$ –решение системы (10) для $z\in S$, можно записать

$$S=\left\{z:\ A_x z+A_y y(z)\le b\right\}. \tag{18}$$

Обозначим $h(z)=A_x z+A_y y(z)-b$, $h(z)=(h_1(z),\ldots,h_m(z))$.

Учитывая (17), можно показать, что

$$\nabla h_i(z)=(a_x^i-a_y^i G_y^{-1}G_x)^T, \tag{19}$$

где a_x^i, a_y^i $-i$-е строки матриц A_x, A_y; G_x, G_y сформированы для точки z.

Положим

$$W(z)=\begin{cases}\left(I-H^T(HH^T)^{-1}H\right)(G_y^{-1}G_x)^T, \text{ если } I(z)\neq\varnothing,\\ (G_y^{-1}G_x)^T, \text{ если } I(z)=\varnothing,\end{cases} \tag{20}$$

где матрицы G_x, G_y, H сформированы для точки z.

Пусть $f(x,y)$ – непрерывно дифференцируемая функция, $x\in E^L$, $y\in E^N$. Положим $\varphi(x)=f(x,y(p_S(x)))$, $\bar{x}=p_S(x)$, $\bar{y}=y(\bar{x})$.

Теорема 2. Пусть точка $x\in E^L$ регулярна относительно множества S, тогда функция $\varphi(x)$ в этой точке непрерывна и дифференцируема по направлениям. Если точка x регулярно разложима относительно множества S, то $\varphi(x)$ в этой точке дифференцируема:

$$\nabla\varphi(x)=\nabla f_x(x,\bar{y})-W(\bar{x})\nabla f_y(x,\bar{y}), \tag{21}$$

4. Рассмотрим исходную совокупность блоков.

Пусть задана непрерывно дифференцируемая функция $f\left(x, y^1,..., y^Q\right)$, $x \in E^L$, $y^q \in E^{N_q}$. Для каждого блока q положим $\overline{x}^q = p_{S^q}(x)$, $\overline{y}^q = y^q(\overline{x}^q)$ и вычислим матрицу вида (20) – $W^q(\overline{x}^q)$.

Положим $\varphi(x) = f(x, y^1(p_{S^1}(x)),..., y^Q(p_{S^Q}(x)))$. Функции такого вида входят в задачу (8). Можно показать, что имеет место следующая теорема.

Теорема 3. Пусть точка $x \in E^L$ – регулярна относительно каждого множества S^q, $q = 1,..., Q$. Тогда функция $\varphi(x)$ в этой точке непрерывна и дифференцируема по направлениям.

Если точка x – регулярно разложима относительно каждого множества S^q, $q = 1,..., Q$,, то $\varphi(x)$ в этой точке дифференцируема:

$$\nabla\varphi(x) = \nabla f_x\left(x, \overline{y}^1,..., \overline{y}^Q\right) - \sum_{q=1}^{Q} W^q\left(\overline{x}^q\right)\nabla f_{y^q}\left(x, \overline{y}^1,..., \overline{y}^Q\right). \qquad (22)$$

В случае, если точка $x \in E^L$ нерегулярна относительно какого-либо множества S^q, то функция $\varphi(x)$ в этой точке может быть разрывна, однако в любой окрестности данной точки найдется регулярная и регулярно разложимая точка.

Следуя [7], определяем почти-градиент 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.

5. Для решения задачи (8) будем использовать метод с растяжением пространства, предложенный в [7] для минимизации почти-дифференцируемых функций.

На каждой итерации для вычисления значений функций $\varphi_k(x)$ должны решаться задачи (9)–(11) для каждого блока q. Градиенты функций $\varphi_k(x)$ вычисляются в соответствии с теоремой 3.

Для решения задач (9)–(11) может использоваться модификация метода линеаризации Б.Н. Пшеничного [8], обеспечивающая на каждой итерации выполнение линейных ограничений (11).

Ю.П. Лаптін

ОПТИМІЗАЦІЙНІ ЗАДАЧІ З БЛОКОВОЮ СИСТЕМОЮ НЕЛІНІЙНИХ ОБМЕЖЕНЬ–РІВНЯНЬ

Розглядаються оптимізаційні задачі з блоковою системою нелінійних обмежень-рівнянь. Визначається поняття псевдорішення системи рівнянь блока, що дозволяє звести вихідну оптимізаційну задачу до редукційованої оптимізаційної задачі меншої розмірності. Для її розв'язання пропонується використовувати методи негладких штрафних функцій і методи негладкої оптимізації. Досліджуються властивості функцій редукційованої задачі.

Yu.P. Laptin

OPTIMIZATION PROBLEMS WITH NONLINEAR BLOCK EQUATIONS SYSTEM

Optimization problems with nonlinear block equations system are considered. A pseudo-solution is defined for the equations system of block. Original optimization problem is reduced to the problem of smaller dimension. Nonsmooth penalty method and nonsmooth optimization method are proposed for solving the reduced problem. Properties of the reduced problem are analysed.

1. *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.
2. *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. – № 1. – С. 3 – 12.
3. *Шор Н.З.* Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.
4. *Shor N. Z.* Nondifferentiable Optimization and Polynomial Problems. – London: Kluwer Academic Publishers, 1998. – 381 p.
5. *Фиакко А., Мак-Кормик Г.* Нелинейное программирование. Методы последовательной безусловной оптимизации / Пер.с англ. – М.: Мир, 1972. – 320 с.
6. *Демьянов В.Ф., Васильев Л.В.* Недифференцируемая оптимизация. – М.: Наука, 1981. – 384 с.
7. *Шор Н.З.* О классе почти-дифференцируемых функций и одном методе минимизации функций этого класса // Кибернетика. – 1972. – № 4. – С. 9 – 17.
8. *Пшеничный Б.Н.* Метод линеаризации. – М.: Наука, 1983. – 136 с.

Получено 10.04.2008