УДК 519.8

Ю.П. ЛАПТИН, Н.Г. ЖУРБЕНКО , В.Н. КУЗЬМЕНКО

РЕШЕНИЕ БЛОЧНЫХ НЕЛИНЕЙНЫХ ЗАДАЧ ОПТИМИЗАЦИИ СО СВЯЗЫВАЮЩИМИ ПЕРЕМЕННЫМИ[1]

Рассматривается использование схем декомпозиции по переменным для решения блочных нелинейных задач выпуклого программирования со связывающими переменными. Описываются процедуры вычисления ε-субградиентов функций координирующей задачи на основании приближенных решений подзадач отдельных блоков. Приводятся описания модификаций известных методов. Рассматриваются процедуры оценивания величины ε. Обсуждаются результаты вычислительных экспериментов.

К блочным нелинейным задачам выпуклого программирования со связывающими переменными сводятся многие задачи, в частности, возникающие в ходе моделирования сложных технических объектов [1]. Эффективным подходом к решению таких задач является использование схем декомпозиции по переменным – применение некоторых методов оптимизации к специально сформулированной координирующей задаче, на каждой итерации которых решаются подзадачи для каждого блока. Свойства функций, входящих в координирующую задачу, исследовались в разных работах, в частности в [2]. При этом предполагалось, что подзадачи для каждого блока решаются точно. Вопросам вычисления ε-субградиентов функций координирующей задачи на основании приближенных решений подзадач посвящены работы [3, 4]. При вычислении ε-субградиентов для решения каждой подзадачи должны использоваться методы, генерирующие подпоследовательность допустимых точек, сходящуюся к решению подзадачи.

Функции, входящие в координирующую задачу, в общем случае определены на ограниченных множествах, которые описываются неявно. С этим связаны сложности при использовании методов оптимизации.

В работе рассматривается регуляризация исходной задачи, при которой функции

[1] Работа выполнена при частичной финансовой поддержке Украинского научно-технологического центра (грант № 1625).

координирующей задачи определены при любых значениях связывающих переменных, приводятся описания модификаций известных методов [7, 8], генерирующих требуемую подпоследовательность допустимых точек, а также результаты вычислительных экспериментов. Рассматриваются процедуры оценивания величины ε.

1. Рассмотрим блочную задачу математического программирования со связывающими переменными:
найти

$$\min_{y,x}\left\{\sum_{q=1}^{Q} f_q^0(y,\, x^q) : f_q^i(y,\, x^q) \le 0,\ i=1,\dots,I_q,\ q=1,\dots,Q\right\}, \tag{1}$$

где $f_q^i(y,\, x^q)$ - выпуклые собственные функции $(L+N_q)$–размерного вектора (y, x^q), $y\in E^L$, $x^q \in E^{N_q}$, $i = 0,\dots, I_q$, $q = 1,\dots, Q$.

Пусть связывающие переменные y зафиксированы, $y=\bar{y}$. Обозначим $D_q(y)=\left\{x^q \in E^{N_q} : f_q^i(y,\ x^q)\le 0,\ i=1,\dots,I_q\right\}$и определим функцию $\Phi^q(y)$:

$$\Phi^q(y)=\begin{cases}\min\left\{f_q^0(y,\, x^q) : x^q\in D_q(y)\right\},\ y\in W_q,\\ +\infty,\ y\notin W_q,\end{cases} \tag{2}$$

где W_q – множество тех значений вектора *y*, для которых решение оптимизационной задачи в (2) существует.

В схемах декомпозиции решается следующая задача, которая эквивалентна исходной задаче (1) : найти

$$\min\left\{\sum_{q=1}^{Q}\Phi^q(y) : y\in E^L\right\}. \tag{3}$$

Свойства функций $\Phi^q(y)$ исследовались в [2]. Процедуры, позволяющие вычислять ε-субградиенты функций $\Phi^q(y)$ на основании приближенного решения оптимизационной задачи в (2), рассматривались в [3, 4]. В работе [3] предложена регуляризация задачи (1), позволяющая определить функции, аналогичные $\Phi^q(y)$, принимающие конечные значения на всем пространстве E^L.

2. При описании свойств функций $\Phi^q(y)$, $f_q^i(y,\, x^q)$, множеств W_q, $D_q(y)$ индекс q, если это не вызывает неоднозначности, опускается.

Пусть $D=\Big\{(y,x): f^i(y,x)\le 0,\ i=1,...,I\Big\}$ - непустое замкнутое множество. Предполагается, что $D\subset \operatorname{int}\operatorname{dom} f^i,\ i=0,...,I$.

Теорема 1 [3]. Пусть $(\bar y,\bar x)\in \operatorname{int}\operatorname{dom} f^i,\ i=0,...,I,\ \bar y\in \operatorname{int} W,\ \bar x\in D(\bar y)$, для множества $D(\bar y)$ выполняется условие Слейтера, в точке $(\bar y,\bar x)$ вычислены ε_i-субградиенты $g^i=(g_y^i,\ g_x^i)$ функций $f^i,\ i=0,1,...,I$ такие, что для некоторых чисел $\bar u_i,\ i=1,...,I$, выполняется условие

$$g_x^0+\sum_{i=1}^{I}\bar u_i g_x^i=0,\qquad \bar u_i\ge 0,\ i=1,...,I. \tag{4}$$

Тогда $\Phi(\bar y)\ge f^0(\bar y,\bar x)-\bar\varepsilon$, а $\bar\varepsilon$-субградиент функции $\Phi(y)$ в точке $y=\bar y$ равен $g_\Phi^{\bar\varepsilon}(\bar y)=g_y^0+\sum\limits_{i=1}^{I}\bar u_i g_y^i$, где $\bar\varepsilon=\varepsilon_0+\sum\limits_{i=1}^{I}\bar u_i(\varepsilon_i-f^i(\bar y,\bar x))$.

Пусть $\bar x\in D(\bar y)$, задана совокупность $\{x_1,\ x_2,...,\ x_m\}$ точек из E^N, в каждой точке $(\bar y,x_k)$ вычислены значения $f^{ik}=f^i(\bar y,x_k)$ и субградиенты $g^{ik}=(g_y^{ik},\ g_x^{ik})$ функций $f^i,\ i=0,...,I,\quad k=1,...,m$.

Рассмотрим линейную аппроксимацию задачи (2) при фиксированном $y=\bar y$: найти

$$\min_{x,\xi}\xi, \tag{5}$$

при ограничениях

$$f^{ik}+(g_x^{ik},x-x_k)\le\xi,\quad k=1,...,m,\ i=0, \tag{6}$$

$$f^{ik}+(g_x^{ik},x-x_k)\le 0,\quad k=1,...,m,\ i=1,...,I. \tag{7}$$

Теорема 2 [4]. Пусть ξ^* - оптимальное значение, $\bar u_{ik}$ - оптимальные значения двойственных переменных, $k=1,...,m,\ i=0,...,I$, задачи (5) - (7). Тогда $g_\Phi^{\bar\varepsilon}(\bar y)=\sum\limits_{k=1}^{m}\sum\limits_{i=0}^{I}\bar u_{ik}g_y^{ik}$, где $\bar\varepsilon=f^0(\bar y,\bar x)-\xi^*$.

Приведенная теорема позволяет строить правила остановки в алгоритмах приближенного решения задачи (2). Используемые алгоритмы должны формировать подпоследовательность допустимых точек, сходящуюся к решению задачи (2). Совокупность $\{x_1,\ x_2,...,\ x_m\}$ может формироваться из точек, генерируемых используемым алгоритмом. Часто возникает необходимость добавления дополнительных точек, улучшающих аппроксимацию задачи (2).

Для формирования дополнительных точек модифицируем метод Келли [5]. Рассмотрим функцию

$$\varphi(x) = \max\left\{ f^0(\bar{y}, x) - \xi^*,\ f^i(\bar{y}, x),\ i = 1,..., I \right\}. \quad (8)$$

Обозначим x^* решение задачи (5) - (7), $g_\varphi(x)$ - субградиент функции φ в точке x. Для простоты будем предполагать, что функции f^i, $i = 0,..., I$, непрерывно дифференцируемы.

Лемма 1. $(g_\varphi(x_k), x^* - x_k) \le 0,\quad k \in \{1,..., m\}$.

Доказательство. Очевидно, что $\varphi(x_k) \ge 0$. Более того, если $\varphi(x_k) = 0$, то x_k - решение задачи (2). Пусть максимум в (8) достигается на $f^{i^*}(\bar{y}, x_k)$, $i^* \in \{1,..., I\}$, т.е. $f^{i^*}(\bar{y}, x_k) > 0$. Утверждение леммы следует из (7), поскольку x^* - допустимая точка. Аналогично для случая, когда максимум в (8) достигается на $f^0(\bar{y}, x_k) - \xi^*$. Лемма доказана.

Лемма 2. Пусть $(g_\varphi(x^*), p) < 0,\ x^p_{\min} = \arg\min\left\{ \varphi(x) : x = x^* + tp,\ t \ge 0\right\}$, $x^\sim \in [x^*, x^p_{\min}]$, где $p \in E^N$, $\|p\| = 1$.

Тогда ограничения

$$f^0(\bar{y}, x^\sim) + (g_x^0(\bar{y}, x^\sim), x - x^\sim) \le \xi,$$

$$f^i(\bar{y}, x^\sim) + (g_x^i(\bar{y}, x^\sim), x - x^\sim) \le 0, \quad i = 1,..., I,$$

отсекают точку (x^*, ξ^*).

Доказательство. Пусть максимум функции (8) в точке $x^\sim$ достигается на $i^* \in \{1,..., I\}$, т.е. $f^{i^*}(\bar{y}, x^\sim) > 0$. Очевидно, что $(g_\varphi(x^\sim), p) \le 0$, откуда $(g_x^{i^*}(\bar{y}, x^\sim), x^\sim - x^*) \le 0$, $f^{i^*}(\bar{y}, x^\sim) + (g_x^{i^*}(\bar{y}, x^\sim), x^* - x^\sim) > 0$. Аналогично рассматривается случай, когда максимум в (8) достигается на $f^0(\bar{y}, x^\sim) - \xi^*$. Лемма доказана.

Пусть для точки x_m выполняется условие $(g_\varphi(x^*), x_m - x^*) \le 0$. Тогда отрезок $[x_m, x^*]$ содержит минимум функции φ по направлению, определяемому этим отрезком. В модифицированном методе Келли на каждой итерации в совокупность $\{x_1, x_2,..., x_m\}$ вместо точки x^* добавляется точка $x^\sim$, удовлетворяющая условиям леммы 2. Для выбора точки $x^\sim$ осуществляется поиск на отрезке $[x_m, x^*]$ для минимизации величины $\|x^\sim - x_m\|$.

При практическом формировании линейной аппроксимации в задачу (5)-(7) включаются только наиболее существенные ограничения. Правила, которые при этом используются, являются эвристическими, однако позволяют существенно понизить размерность генерируемой задачи.

3. Рассмотрим модификацию [7] метода линеаризации, формирующую сходящуюся подпоследовательность допустимых точек задачи (2).

Точки, генерируемые обычным методом линеаризации, являются допустимыми для линейной аппроксимации задачи (2). Поскольку допустимая область, задаваемая ограничениями (7), «шире», чем исходная допустимая область, задаваемая функциями $f^i(\overline{y},x)$, $i=1,...,I$, то генерируемая последовательность точек $\{ x_1, x_2,..., x_m \}$ лежит, как правило, вне допустимой области нелинеаризированной задачи. Чтобы получить подпоследовательность точек, лежащую в допустимой области и сходящуюся к решению задачи, применяется комбинированный метод [6], включающий как составную часть метод линеаризации.

Приведем алгоритм рассматриваемого метода.

Выберем начальную точку x_1, зададим $N_1 > 0$, положим $\varepsilon_1 = \gamma/2$, $X_1 = \{x_1\}$, где γ - константа в условии Слейтера.

Точка x_{k+1}, множество X_{k+1} и значения N_{k+1}, ε_{k+1} определяются на k-м шаге по результату решения задачи

$$\min_{x,\,\xi_1,\xi_2} \{ \left\| x-\tilde{x}_k \right\|^2 + \xi_1 + N_k\, \xi_2 \},$$

$$f^{ik} + (g_x^{ik},\, x-x_k) \le \xi_1,\ k=1,...m,\ i=0,$$

$$f^{ik} + (g_x^{ik},\, x-x_k) \le \xi_2,\ k=1,...,m,\ i=1,...,I,$$

$$\xi_2 \ge -\varepsilon_k, \tag{9}$$

где $\tilde{x}_k = \arg\min\{F_k(x) : x \in X_k\}$ является точкой минимума штрафной функции $F_k(x) = f^0(x) + N_k \max\{-\varepsilon_k; f^1(x);...;f^m(x)\}$ на X_k.

Пусть x_{k+1}, ξ_1^k, ξ_2^k – решение задачи (8). Выполним *регулировку 1*:

$X_{k+1} = X_k \cup \{x_{k+1}\}$;

$N_{k+1} = 2N_k$, если $\xi_2^k > -\varepsilon_k$, иначе $N_{k+1} = N_k$;

$\varepsilon_{k+1} = \varepsilon_k/2$, если $\max\{f^1(x);...;f^m(x)\} < 0$, иначе $\varepsilon_{k+1} = \varepsilon_k$.

Лемма 3[6]. Описанный процесс при $\varepsilon_1 > 0$ порождает подпоследовательность точек $x_{p(j)}$, $j = 1,2,\ldots$, сходящуюся к решению задачи и такую, что $f^{ik}(x_{p(j)}) < 0$, для всех $i=1,...,I$.

4. Для регуляризации исходной задачи (1) вводятся вспомогательные переменные $v^q = (v_1^q, ..., v_L^q)$, $q = 1, ..., Q$, и рассматривается следующая задача [3]: найти

$$\min_{x,y,v} \sum_{q=1}^{Q} \left\{ f_q^0(v^q, x^q) + M_q \sum_{l=1}^{L} | y_l - v_l^q | \right\} \tag{10}$$

при ограничениях

$$f_q^i(v^q, x^q) \le 0, i = 1, ..., I_q, q = 1, ..., Q, \tag{11}$$

где $v^q \in E^L$, $q = 1, ..., Q$.

Если задача (1) имеет решение, то при достаточно больших положительных значениях M_q, $q = 1, \ldots, Q$, решения задач (1) и (10), (11) совпадают.

Обозначим $z^q = (v^q, x^q)$, $z^q \in E^L \times E^{N_q}$. При фиксированных значениях y задача (10), (11) распадается на подзадачи

$$\Psi^q(y) = \min_{z^q} \left\{ f_q^0(z^q) + M_q \sum_{l=1}^{L} | y_l - z_l^q | : f_q^i(z^q) \le 0, i = 1, ..., I_q \right\}. \tag{12}$$

Пусть системы ограничений задач (12) совместны и эти задачи имеют решение при $M_q = 0$, тогда функции $\Psi^q(y)$ определены при любых y, а координирующая задача

$$\min \left\{ \sum_{q=1}^{Q} \Psi^q(y) : y \in E^l \right\} \tag{13}$$

- задача выпуклого программирования без ограничений.

5. Предварительные вычислительные эксперименты по использованию предлагаемого подхода проводились при решении тестовых задач, являющихся модификацией известной трудной задачи из [6]. Содержательно рассматриваемые задачи заключаются в поиске минимума потенциальной энергии для цепочки, подвешенной за крайние звенья. В модифицированной задаче рассматривается несколько цепочек, имеющих одно общее звено. Крайние звенья зафиксированы, требуется найти минимум потенциальной энергии. При фиксированном положении общего звена задача распадается на подзадачи для каждой отдельной цепочки. Рассматриваемая задача является задачей квадратичного программирования.

Результаты этих экспериментов показали:

при приближении к оптимуму подзадачи (12) линейные аппроксимации (5)-(7) становятся плохо обусловленными, что приводит к большим погрешностям при использовании программных средств линейного программирования;

для данного класса тестовых задач трудоемкость решения подзадач (12) достаточно высокая и это требует разработки специальных алгоритмов повышенной эффективности.

Для преодоления указанных проблем:

в случае плохой обусловленности линейной аппроксимации (5) - (7) необходимо переходить на другие алгоритмы вычисления ε-субградиентов функций $\Psi^q(y)$;

для повышения эффективности алгоритмов решения подзадач (12) на текущей итерации необходимо использовать аппроксимации, полученные на предыдущих итерациях.

В целом использование предлагаемой схемы декомпозиции целесообразно в случаях, когда исходная задача (1) распадается на подзадачи, большинство из которых имеют специальную структуру и эффективные алгоритмы решения, а подзадач общего вида немного и их размерность невелика.

Ю.П. Лаптін, М.Г. Журбенко, В.М. Кузьменко

РОЗВ'ЯЗУВАННЯ БЛОЧНИХ НЕЛІНІЙНИХ ЗАДАЧ ОПТИМІЗАЦІЇ ЗІ ЗВ'ЯЗУЮЧИМИ ЗМІННИМИ

Розглядається використання схем декомпозиції за змінними для розв'язування блокових нелінійних задач опуклого програмування із зв'язуючими змінними. Описуються процедури обчислення ε-субградієнтів функцій координуючої задачі на підставі наближених розв'язків підзадач окремих блоків. Наводяться описи модифікацій відомих методів, а також результати обчислювальних експериментів

Yu.P.Laptin, N.G. Zhurbenko, V.N. Kuzmenko

SOLVING BLOCK NONLINEAR OPTIMIZATION PROBLEMS WITH CONNECTIVE VARIABLES

Using of decomposition in variables is considered for nonlinear convex block problems with coordinating variables. There are described procedures for calculating of ε-subgradient for functions of master problem. These procedures use approximate solutions of subproblems for individual blocks. Modifications of well-known optimization methods are described. Results of numerical experiments are discussed.

1. Использование средств оптимизации в системе автоматизированного проектирования энергетических котлоагрегатов КРОКУС, Ю.П. Лаптин, М.М. Левин, П.И. Волковицкая и др. // Энергетика и электрификация. – 2003. – № 7. – С. 41 – 51.
2. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – London: Kluwer Academic Publishers, 1998. – 381 p.
3. *Лаптин Ю.П.* Декомпозиция по переменным для некоторых задач оптимизации // Кибернетика и системный анализ. – 2004. – № 1. – С. 98 – 104.

4. *Лаптин Ю.П.* ε-субградиенты в методах декомпозиции по переменным для некоторых задач оптимизации // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2003. – № 2. – С. 75 – 82.
5. *Kelley J.* The cutting plane method for solving convex programs // SIAM J. – 1960 – 8, N 4. – P. 703 –712.
6. *Пшеничный Б.Н., Ненахов Э.И., Кузьменко В.Н.* Комбинированный метод решения общей задачи выпуклого программирования // Кибернетика и системный анализ. – 1998. – № 4. – С. 121 – 134.
7. *Кузьменко В.Н., Бойко В.В.* О применении комбинированного метода выпуклого программирования // Теорія оптимальних рішень. – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2003. – № 2. – С. 19 – 24.
8. *Шор Н.З., Журбенко Н.Г.* Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51 – 59.

Получено 29.04.2004