Теорія оптимальних рішень

Рассматривается декомпозиция блочных задач выпуклого программирования со связывающими переменными. Исследуются свойства подзадач, полезные при вычислении ε-субградиентов целевых функций по связывающим переменным. Предлагаются некоторые алгоритмы решения рассматриваемой задачи.

УДК 519.8

Ю.П. ЛАПТИН

ε-СУБГРАДИЕНТЫ В МЕТОДАХ ДЕКОМПОЗИЦИИ ПО ПЕРЕМЕННЫМ ДЛЯ НЕКОТОРЫХ ЗАДАЧ ОПТИМИЗАЦИИ

При решении оптимизационных задач, возникающих в ходе моделирования сложных технических объектов, обычно приходится учитывать их особенности и структуру. Рассматриваемые объекты, как правило, состоят из совокупности связанных между собой блоков и подсистем, что отражается в структуре оптимизационных моделей [1]. В некоторых случаях такие модели сводятся к блочным задачам выпуклого программирования со связывающими переменными [2]. Эффективным подходом к решению таких задач в случае линейного программирования является применение схем декомпозиции по переменным. В случае нелинейного программирования использование схем декомпозиции сопряжено с определенными трудностями. Преодолению некоторых из них посвящена настоящая работа.

1. Рассматривается блочная задача математического программирования со связывающими переменными:

$$\min_{y,\,X} F(y,\ X) = \min_{y,X} \sum_{q=1}^{Q} f_q^0(y,\ x^q) \quad (1)$$

$$f_q^i(y,\, x^q) \le 0,\, i = 1,\dots,I_q,\ q = 1,\dots,Q, \quad (2)$$

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,…, I_q, q = 1,…, Q, $X = (x^1,\dots,\, x^Q)$.

Пусть переменные y зафиксированы, $y=\bar{y}$, тогда задача (1), (2) распадается на подзадачи $P_q(\bar{y})$ следующего вида:

$$\min\left\{f_q^0(\bar{y},\ x^q) : x^q \in E^{N_q},\ f_q^i(\bar{y},\ x^q) \le 0,\ i=1,\dots,I_q\right\},\ q=1,\dots,Q \tag{3}$$

Обозначим W_q множество тех значений вектора y, для которых решение задачи $P_q(y)$ существует, и определим функцию $\Phi^q(y)$:

$$\Phi^q(y)=\begin{cases}\min\left\{f_q^0(y,\ x^q) : x^q \in D_q(y)\right\},\ y\in W_q,\\ +\infty,\ y\notin W_q,\end{cases} \tag{4}$$

где $D_q(y)$ - множество всех x^q, удовлетворяющих ограничениям задачи $P(y)$.

Используя (4), задачу (1), (2) можно представить в виде следующей (координирующей): найти

$$\min\left\{\sum_{q=1}^{Q}\Phi^q(y) : y\in E^L\right\}, \tag{5}$$

В работе [2] показано, что если задача (1), (2) имеет решение, то $\Phi^q(y),\ q=1,\dots,Q$ являются собственными выпуклыми функциями, там же определены правила вычисления субградиентов $g_\Phi^q(y)$ функций $\Phi^q(y)$. При этом каждая задача (3) должна решаться точно, субградиенты $g_\Phi^q(y)$ выражаются через субградиенты функций $f_q^i(y,x)$ в оптимальной точке $x^{*q}(y)$. Для недифференцируемых функций эти субградиенты должны выбираться специальным образом. Поскольку для нелинейных задач алгоритмы оптимизации сходятся к оптимуму как к предельной точке, возникают значительные трудности в определении субградиентов $g_\Phi(y)$ и применении схем декомпозиции к нелинейным блочным задачам со связывающими переменными. В случае недифференцируемых функций возникают дополнительные трудности, поскольку отсутствуют процедуры формирования субдифференциалов функций (в точке может быть вычислен только некоторый субградиент функции). Преодолеть эти трудности можно переходя к использованию ε-субградиентов $g_\Phi^q(y)$ функций $\Phi^q(y)$. Способам вычисления таких ε-субградиентов посвящена основная часть настоящей работы.

Задача (5) является задачей выпуклого программирования, однако непосредственное ее использование в вычислительных алгоритмах затруднительно в виду сложности описания подмножеств из E^L, на которых функции $\Phi^q(y)$ принимают конечные значения. В работе [2] для преодоления этих трудностей проводилась регуляризация задачи (5), позволяющая определить функции, аналогичные $\Phi^q(y)$, на всем пространстве E^L. При этом предполагается, что функции

$f_q^i(y,\, x^q)$ определены на всем пространстве $E^L \times E^{N_q}$. После регуляризации для решения задачи (5) могут использоваться известные ε- субградиентные методы [2 - 5].

2. В дальнейшем будем рассматривать задачи вида (3), индекс q, если это не вызывает неоднозначности, будем опускать.

Обозначим $D = \left\{ (y,x) : f^i(y,\, x) \le 0,\, i = 1,...,I \right\}$. Предполагается, что

1) D – замкнутое множество, $D \subset \operatorname{int} \operatorname{dom} f^i$, $i = 0,...,I$;

2) множество D непусто и для него выполняется условие Слейтера;
3) в каждой точке области определения функции могут быть вычислены значение и некоторый субградиент этой функции.

Задача $P(y)$ определяется соотношением (3), функция $\Phi(y)$ - соотношением (4) (индекс q опущен), $D(y)$ - допустимое множество задачи $P(y)$.

Теорема 1. Пусть $(\bar{y},\bar{x}) \in \operatorname{int} \operatorname{dom} f^i$, $i = 0,...,I$, $\bar{y} \in \operatorname{int} W$, для допустимого множества $D(\bar{y})$ задачи $P(\bar{y})$ выполняется условие Слейтера, заданы числа $\varepsilon_i \ge 0$, $i = 0, 1,...,I$, в точке $(\bar{y},\bar{x})$ вычислены ε_i - субградиенты $g^i = (g_y^i,\, g_x^i)$ функций f^i такие, что для некоторых чисел $\bar{u}_i$, $i = 1,...,I$, выполняются условия

$$g_x^0 + \sum_{i=1}^{I} \bar{u}_i g_x^i = 0\,, \tag{6}$$

$$\bar{u}_i \ge 0,\; i = 1,...,I\,. \tag{7}$$

Тогда

1) $\Phi(\bar{y}) \ge f^0(\bar{y},\bar{x}) - \bar{\varepsilon}$;

2) если $\bar{x} \in D(\bar{y})$, то $\bar{\varepsilon}$-субградиент функции $\Phi(y)$ в точке $y = \bar{y}$ может быть вычислен по формуле

$$g_\Phi^{\bar{\varepsilon}}(\bar{y}) = g_y^0 + \sum_{i=1}^{I} \bar{u}_i g_y^i, \tag{8}$$

где

$$\bar{\varepsilon} = \varepsilon_0 + \sum_{i=1}^{I} \bar{u}_i (\varepsilon_i - f^i(\bar{y},\bar{x})). \tag{9}$$

Доказательство. Рассмотрим первое утверждение. Используя теорему Куна - Таккера, имеем

$$\Phi(\bar{y}) = \max_{u\ge 0} \min_x \left\{ f^0(\bar{y},\, x) + \sum_{i=1}^{I} u_i f^i(\bar{y}, x) \right\} \ge$$

$$\ge \min_x \left\{ f^0(\bar{y},\, \bar{x}) + (g_x^0, x-\bar{x}) - \varepsilon_0 + \sum_{i=1}^{I} \bar{u}_i (f^i(\bar{y},\bar{x}) + (g_x^i, x-\bar{x}) - \varepsilon_i) \right\}.$$

Откуда, с учетом (6), получаем утверждение теоремы.

Докажем второе утверждение

Пусть в некоторой точке $y \in \operatorname{int} W$ для допустимого множества $D(y)$ задачи $P(y)$ выполняется условие Слейтера. Обозначим *x(y)* – решение задачи $P(y)$, $u_i(y) \ge 0,\ i = 1,...,I$ – значения множителей Лагранжа, таких, что

$$\Phi(y) = L_{u(y)}(y,\, x(y)) = f^0(y,\, x(y)) + \sum_{i=1}^{I} u_i(y) f^i(y, x(y)) =$$

$$= \min_x \left\{ f^0(y,\, x) + \sum_{i=1}^{I} u_i(y) f^i(y, x) \right\} = \max_{u\ge 0} \min_x \left\{ f^0(y,\, x) + \sum_{i=1}^{I} u_i f^i(y, x) \right\}.$$

Такие множители существует в соответствии с теоремой Куна - Таккера.

По условию теоремы в точке $\bar{y}$ для допустимого множества $D(\bar{y})$ условие Слейтера выполняется. Рассмотрим точку $y' \in \operatorname{int} W$, достаточно близкую к $\bar{y}$, в которой для множества $D(y')$ также выполняется условие Слейтера (такая точка существует в силу непрерывности функций $f^i(y,x),\ i = 1,...,\ I$).

Обозначим $x' = x(y')$. Очевидно, $\Phi(y') = L_{u(y')}(y', x') \ge L_{\bar{u}}(y', x')$. Поскольку $\bar{x}$ - допустимая точка задачи $P(\bar{y})$, а $\Phi(\bar{y})$ - значение целевой функции в оптимальной точке, то $\Phi(\bar{y}) \le f^0(\bar{y}, \bar{x})$. Откуда получаем

$$\Phi(y') - \Phi(\bar{y}) \ge L_{\bar{u}}(y', x') - f^0(\bar{y}, \bar{x}) = f^0(y', x') - f^0(\bar{y}, \bar{x}) +$$

$$+ \sum_{i=1}^{I} \bar{u}_i \left[f^i(y', x') - f^i(\bar{y}, \bar{x}) + f^i(\bar{y}, \bar{x}) \right] \ge$$

$$\ge (x' - \bar{x},\ g_x^0) + (y' - \bar{y},\ g_y^0) - \varepsilon_0 + \sum_{i=1}^{I} \bar{u}_i \left[(x' - \bar{x},\ g_x^i) + (y' - \bar{y},\ g_y^i) - \varepsilon_i + \right.$$

$$\left. + f^i(\bar{y}, \bar{x}) \right] = (y' - \bar{y},\ g_y^0 + \sum_{i=1}^{I} \bar{u}_i g_y^i) - \varepsilon_0 + \sum_{i=1}^{I} \bar{u}_i \left[-\varepsilon_i + f^i(\bar{y}, \bar{x}) \right].$$

Последнее равенство выполняется в силу (6), т.е. $g_\Phi^{\bar{\varepsilon}}(\bar{y}) = g_y^0 + \sum_{i=1}^{I} \bar{u}_i g_y^i$. Теорема доказана.

Рассмотрим сформулированный результат применительно к задаче линейного программирования. Пусть $f^i(y,x) = a_x^i x + a_y^i y + b^i$, $i = 0,..., I$, где a_x^i, a_y^i - вектор-строки соответствующих размерностей.

Обозначим $A_x = \begin{Bmatrix} a_x^1 \\ ... \\ a_x^I \end{Bmatrix}$, $A_y = \begin{Bmatrix} a_y^1 \\ ... \\ a_y^I \end{Bmatrix}$, $b = \begin{Bmatrix} b^1 \\ ... \\ b^I \end{Bmatrix}$, $b(y) = b + A_y y$, $b^0(y) = b^0 + a_y^0 y$.

Тогда задачу линейного программирования $P(y)$ и двойственную к ней можно представить в виде

$$\min_x \left\{ a_x^0 x + b^0(y) : A_x x + b(y) \le 0 \right\}, \tag{10}$$

$$\max_u \left\{ b(y)^T u + b^0(y) : A_x^T u + a_x^0 = 0,\ u \ge 0 \right\}. \tag{11}$$

Очевидно, что в условиях теоремы 1 можно положить $\varepsilon_i = 0$, $i = 0, 1,..., I$, а соотношения (6) , (7) есть условия допустимости двойственной задачи (11).

Лемма 1. В условиях теоремы 1 для задачи (10) имеет место $\bar{\varepsilon} = a_x^0 \bar{x} - b(\bar{y})^T \bar{u}$.

Доказательство.

$\bar{\varepsilon} = -\sum_{i=1}^{I} \bar{u}_i f^i(\bar{y}, \bar{x}) = -(A_x \bar{x} + b(\bar{y}))^T u = a_x^0 \bar{x} - a_x^0 \bar{x} - (A_x \bar{x} + b(\bar{y}))^T u = a_x^0 \bar{x} - b(\bar{y})^T \bar{u}$. Лемма доказана.

Сформулированная лемма определяет простое правило остановки при приближенном решении задач линейного программирования и вычислении $\bar{\varepsilon}$-субградиентов функции $\Phi(\bar{y})$. При этом должны использоваться алгоритмы решения задач линейного программирования, генерирующие на каждой итерации допустимые точки прямой и двойственной задач (см., например, [6]). Близкие результаты для задач линейного программирования получены в [5].

В общем случае для вычисления $\bar{\varepsilon}$-субградиентов функции $\Phi(\bar{y})$ должны применяться специальные алгоритмы. Результаты, полезные для разработки таких алгоритмов, приводятся ниже.

Пусть задана совокупность $\{ x_1, x_2,..., x_m \}$ точек из E^N, $x_m \in D(\bar{y})$, 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, $i = 0,..., I$, $k = 1,..., m$.

Рассмотрим линейную аппроксимацию задачи (3) при фиксированном $y = \bar{y}$, которая может быть представлена в виде: найти

$$\min_{x,\xi} \xi \tag{12}$$

при ограничениях

$$f^{ik}+(g_x^{ik},x-x_k)\le\xi,\quad k=1,...,m,\ i=0,\tag{13}$$

$$f^{ik}+(g_x^{ik},x-x_k)\le 0,\quad k=1,...,m,\ i=1,...,I,\tag{14}$$

Пусть u_{0k}, u_{ik} - двойственные переменные, соответствующие ограничениям (13), (14), $k=1,...,\ m,\ \ i=1,...,\ I$. Двойственная задача записывается следующим образом: найти

$$\max_u\sum_{i=0}^{I}\sum_{k=1}^{m}\left(f^{ik}-(g_x^{ik},\ x_k)\right)u_{ik},\tag{15}$$

при ограничениях

$$\sum_{k=1}^{m}u_{0k}=1,\tag{16}$$

$$\sum_{i=0}^{I}\sum_{k=1}^{m}u_{ik}g_x^{ik}=0,\tag{17}$$

$$u_{ik}\ge 0,\ i=0,...,I,\ k=1,...,m.\tag{18}$$

Обозначим x_m^*, ξ_m^* - решение задачи (12)-(14), $\bar{u}_{ik}$ - оптимальные значения двойственных переменных, $k=1,...,m,\ i=0,...,I$.

Лемма 2. Пусть для некоторых *i, k* выполняется $\bar{u}_{ik}>0$, тогда g^{ik} есть ε^{ik}-субградиент функции f^i в точке $(\bar{y},x_m)$, где

$$\varepsilon^{ik}=\begin{cases} f^{im}+(g_x^{ik},x_m^*-x_m),\ i>0,\\ f^{0m}-\xi_m^*+(g_x^{0k},x_m^*-x_m),\ i=0.\end{cases}\tag{19}$$

Доказательство. Легко видеть, что

$$\varepsilon^{ik}=f^{im}-f^{ik}-(g_x^{ik},\ x_m-x_k)=f^{im}-f^{ik}-(g_x^{ik},\ x_m-x_k+x_m^*-x_m^*)=$$
$$=f^{im}+(g_x^{ik},\ x_m-x_m^*)-\left[f^{ik}+(g_x^{ik},\ x_m^*-x_k)\right].$$

Поскольку $\bar{u}_{ik}>0$, то соответствующее ограничение выполняется как равенство в оптимальной точке. Откуда следует утверждение леммы. Лемма доказана.

Лемма 3. Пусть $\bar{u}_i=\sum_{k=1}^{m}\bar{u}_{ik}>0$. Тогда вектор $g^i=(\bar{u}_i)^{-1}\left(\sum_{k=1}^{m}\bar{u}_{ik}g^{ik}\right)$ есть ε^i-субградиент функции f^i в точке $(\bar{y},x_m)$, где $\varepsilon^i=(\bar{u}_i)^{-1}\left(\sum_{k=1}^{m}\bar{u}_{ik}\varepsilon^{ik}\right)$, ε^{ik} определено в соответствии с леммой 2.

Доказательство. В соответствии с леммой 2 имеем $f^i(\bar{y},x)\ge f^{im}+(g^{ik},x-x_m)-\varepsilon^{ik}$, $k=1,...,m$. Умножая неравенства на $\bar{u}_{ik}$ и суммируя по *k*, получаем утверждение леммы.

Теорема 2.

$$g_{\Phi}^{\bar{\varepsilon}_m}(\bar{y}) = \sum_{k=1}^{m} \sum_{i=0}^{I} \bar{u}_{ik} g_y^{ik}, \tag{20}$$

$$\bar{\varepsilon}_m = f^{0m} - \xi_m^*. \tag{21}$$

Доказательство. Не трудно видеть, что векторы $g^i = (g_y^i, g_x^i)$ и числа u_i, определенные в соответствии с леммой 3, удовлетворяют соотношениям (6), (7). Подставляя соответствующие выражения в (8), (9) и учитывая, что $\sum_{k=1}^{m} \bar{u}_{0k} = 1$, получаем $\bar{\varepsilon}_m = f^{0m} - \xi_m^* + \left(\sum_{k=1}^{m} \sum_{i=0}^{I} \bar{u}_{ik} g_x^{ik},\ x_m^* - x_m \right) = f^{0m} - \xi_m^*$. Последнее равенство следует из (17). Теорема доказана.

Теорема 3 позволяет строить правила остановки в процессах приближенного решения задачи $P(y)$. Для этого необходимо, чтобы допустимые точки x_m генерировались некоторым процессом, сходящимся к решению задачи $P(y)$, и оптимальные значения линейных аппроксимаций также сходились к оптимальному значению задачи $P(y)$.

Сформулированные условия выполняются, если ограничения задачи $P(y)$ линейны, а для решения используется метод Келли [7]. Тогда все генерируемые точки являются допустимыми и оптимальные значения линейных аппроксимаций сходятся к оптимальному значению задачи $P(y)$. Однако метод Келли имеет медленную сходимость и сложный в программной реализации.

В случае нелинейных ограничений необходимы дополнительные усилия для формирования допустимых точек и специальным образом построенные линейные аппроксимации, обеспечивающие сходимость оптимальных значений.

При практическом формировании линейной аппроксимации в задачу (12)-(14) включаются только наиболее существенные ограничения. Правила, которые при этом используются, являются эвристическими, однако позволяют существенно понизить размерность генерируемой задачи.

3. Представленные результаты могут быть полезны при разработке алгоритмов декомпозиции для решения задач вида (1), (2). Применение таких алгоритмов наиболее эффективно в случаях, когда исходная задача (1), (2) распадается на подзадачи, большинство из которых имеют специальную структуру и эффективные алгоритмы решения, а подзадач общего вида немного и их размерность невелика.

На основе предлагаемого подхода разрабатываются отдельные классы рекурсивной объектно-ориентированной библиотеки RBL [1], предназначенной для оптимизационного моделирования задач, имеющих сложную структуру. При создании приложений прикладные классы должны порождаться от базовых классов библиотеки RBL. Для подзадач, имеющих специальную структуру, могут разрабатываться соответствующие алгоритмы решения.

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

Лаптін Ю.П.

ε-СУБГРАДІЄНТИ В МЕТОДАХ ДЕКОМПОЗИЦІЇ ЗА ЗМІННИМИ ДЛЯ ДЕЯКИХ ЗАДАЧ ОПТИМІЗАЦІЇ

Розглядається декомпозиція блочних задач опуклого програмування зі зв'язуючими змінними. Досліджуються властивості підзадач, корисні при обчисленні ε-субградієнтів цільових функцій за зв'язуючими змінними. Пропонуються деякі алгоритми розв'язування задач, які розглядаються.

Laptin Y.P.

ε -SUBGRADIENT IN DECOMPOSITION ALGORITMS ON VARIABLES FOR SOME OPTIMIZATION PROBLEMS

Decomposition of block convex problems with linking variables is discussed. Block functions are defined on restricted sets. We analyze the subproblem properties which are useful for calculaiting ε -subgradient of the object function by linking variables. Some algorithms are proposed for solving such problems. Computational experiment results are discussed.

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. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – London: Kluwer Academic Publishers, 1998. - 381 p.
3. *Демьянов В.Ф., Васильев Л.В.* Недифференцируемая оптимизация. – М.: Наука, 1981. – 384 с.
4. *Ржевский С.В.* Монотонные методы выпуклого программирования. – Киев: Наук. думка, 1993. – 316 с.
5. *Condzio J., Vial J.-P.* Warm Start and ε-Subgradients in a Cutting Plane Scheme for Block-Angular Linear Programs // Computational Optimization and Applications. – 1999. – N 14. – P. 17–36.
6. *Юдин Д.Б., Гольштейн Е.Г.* Линейное программирование. Теория, методы и приложения. – М.: Наука, 1969. – 424 с.
7. *Kelley J.* The cutting plane method for solving convex programs // SIAM J. – 1960.- 8, N 4. - P.703 –712.

Получено 17.09.2003