Численные методы

УДК 512.25/26

ВАРИАНТ МЕТОДА УСЛОВНОГО ГРАДИЕНТА ПРИ ДВУСТОРОННЕЙ ДЕКОМПОЗИЦИИ

Н.В.Шмырева

Иногда при построении декомпозиционных алгоритмов для решения задач линейного программирования большой размерности фиксируется часть переменных (так называемых управляющих) прямой или двойственной задачи, и исходная задача сводится к минимизации некоторой кусочно-линейной выпуклой функции от управляющих переменных. В [I] для минимизации функций такого вида рассматривался вариант метода возможных направлений с квадратичной нормализацией.

В настоящей работе этот процесс обобщается на случай функций вида $F(z)=\max\limits_i F^i(z)$, где каждая из функций $F^i(z)$ связана с некоторой задачей линейного программирования в упомянутом выше смысле. В качестве примера ситуации, в которой возникают функции такого вида, здесь рассматривается декомпозиционная схема для задачи линейного программирования с окаймлением, основанная на сведении исходной задачи к поиску седловой точки некоторой функции.

I. Пусть имеется k линейных систем вида

$$A^s x^s + B^s z = b^s, \qquad x^s \geq 0,\ z \geq 0, \tag{I}$$

где $x^s \in R^{n_s}$, $z \in R^n$, $b^s \in R^{m_s}$, A^s и B^s - матрицы соответствующих размерностей. Каждая их этих систем определяет множества $X_s(z)$ и Z_s, представляющие собой проекции множества решений системы на соответствующие подпространства.

Пусть, кроме того, заданы векторы $c^s \in R^{n_s}$ и $d^s \in R^n$. Каждая из систем (I) порождает кусочно-линейную выпуклую функцию $F^s(z)$, заданную на множестве Z_s и имеющую вид

$$F^s(z) = (d^s, z) + \min_{x^s \in X_s(z)} (c^s, x^s).$$

Ставится задача минимизации функции

$$F(z) = \max_s F^s(z)$$

на множестве $Z = \bigcap_s Z_s$.

Функция $F(z)$ является выпуклой кусочно-линейной, и для ее минимизации в принципе применима та же схема метода возможных направлений, что и в [I]. Особенности конкретизации этой схемы для данного случая порождаются неявным характером задания как самой функции, так и области ее определения Z, и проявляются, главным образом, при поиске подходящего направления. Поэтому, не останавливаясь на полном описании метода, напомним итеративную процедуру поиска подходящего направления, использованную в [I], для минимизации функции $F^s(z)$ на множестве Z_s.

Пусть $\bar{z} \in Z_s$. Для вычисления значения функции F^s в точке $\bar{z}$ нужно решить задачу линейного программирования

$$\begin{aligned} &(c^s, x^s) \to \min! \\ &A^s x^s = b^s - B^s \bar{z}, \qquad (2) \\ &x^s \geq 0. \end{aligned}$$

Как отмечалось в [I], вектор $d^s - y^s B^s$, где y^s - любое оптимальное решение задачи, двойственной к (2), является субградиентом функции F^s в точке $\bar{z}$. Необходимым и достаточным условием того, чтобы точка $\bar{z}$ была точкой минимума функции F^s, служит условие $0 \in \partial F^s(\bar{z})$. Субдифференциал $\partial F^s(\bar{z})$ в общем случае представляет собой сумму ограниченного выпуклого многогранника с вершинами v^α, $\alpha = 1, \dots, V$, и конечно-порожденного конуса с образующими w^β, $\beta = 1, \dots, W$, которые являются нормалями граничных гиперплоскостей множества Z_s, проходящих через точку $\bar{z}$. Если $0 \notin \partial F^s(\bar{z})$, то напраление, задаваемое вектором $(-h)$, где h - наименьший по евклидовой норме вектор из множества $\partial F^s(\bar{z})$, будет допусти-.

мым направлением убывания функции $F^s(\bar{x})$. Однако этим свойством, вообще говоря, не обладают близкие к h векторы из $\partial F^s(\bar{x})$. Следовательно, лишь точное определение h позволяет указать подходящее направление, что не дает возможности применять итеративные методы для его нахождения. В связи с этим вместо множества $\partial F^s(\bar{x})$ рассматривается многогранник $P^s(\bar{x})$, являющийся выпуклой оболочкой всех точек v^α и w^β. В предположении телесности множества Z_s условие $0 \in \partial F^s(\bar{x})$ эквивалентно условию $0 \in P^s(\bar{x})$. Отметим, что введение многогранника $P^s(\bar{x})$ соответствует схеме метода возможных направлений, предложенной для дискретной задачи минимакса в [2].

Для того чтобы вектор $(-h)$ задавал подходящее направление для минимизации функции $F^s(\bar{x})$, достаточно, чтобы выполнялись следующие неравенства:

$$(h, v^\alpha) > 0, \ \alpha = 1, \dots, V, \tag{3}$$

$$(h, w^\beta) > 0, \ \beta = 1, \dots, W. \tag{4}$$

Если $0 \notin P^s(\bar{x})$, то минимальный по евклидовой норме элемент из $P^s(\bar{x})$ удовлетворяет этим условиям. Однако поиск такого элемента весьма затруднителен ввиду неявного задания вершин многогранника $P^s(\bar{x})$. В [I] для поиска требуемого вектора h была предложена процедура, являющаяся модификацией метода условного градиента и состоящая в следующем.

Пусть $\theta \in (0,1)$ и $h^\tau \in P^s(\bar{x})$ – вектор, полученный после τ шагов процесса.

I. Для h^τ проверяется выполнение неравенства

$$h_i^\tau \leq 0, \quad i \in I(\bar{x}) = \{i / \bar{x}_i = 0\}. \tag{5}$$

Если для некоторого $i_0 \in I(\bar{x})$ оказалось $h_{i_0}^\tau > 0$, то это означает, что i_0-й орт $e_{i_0} \in R^n$, взятый со знаком " – ", является одним из векторов w^β, на котором нарушается условие (4). В этом случае принимаем $z = -e_{i_0}$ и переходим к п.4.

2. Решаем задачу линейного программирования

$$\begin{gathered} (d^s - y^s B_1^s h^\tau) \to min! \\ y^s A_j^s \leq c_j^s, \ j \in J(\bar{x}^s), \end{gathered} \tag{6}$$

здесь $\bar{x}^s$ – какое-либо из оптимальных решений задачи (2), а $J(\bar{x}^s)=\{j/\bar{x}^s_j=0\}$. Существование оптимального решения в задаче (6) вместе с выполнением условий (5) означает, что все неравенства (4) для $h=h^\tau$ выполняются. Если же окажется, что целевая функция этой задачи на множестве допустимых решений не ограничена снизу, то процедура модифицированного симплекс-метода позволяет получить один из векторов w^β, для которого $(h^\tau, w^\beta)<0$. В этом случае принимаем $z=w^\beta$ и переходим к п.4.

Пусть задача (6) разрешима и $\bar{y}^s$ – ее оптимальное решение. Полагаем $z=d^s-\bar{y}^s d^s$ и переходим к п.3.

3. Проверяем выполнение условия

$$(h^\tau, z)\geqslant\theta\|h^\tau\|^2. \qquad (7)$$

Если это неравенство выполняется, процесс окончен, в противном случае переходим к п.4.

4. В качестве $h^{\tau+1}$ принимаем точку отрезка $[h^\tau, z]$, ближайшую к нулю, и возвращаемся к п.I.

Для описанной процедуры в [I] было показано, что если $0\notin P^s(\bar{x})$, то она заканчивается через конечное число шагов, и полученный вектор $(-h^\tau)$ задает подходящее направление. Если же $0\in P^s(\bar{x})$ и X_s телесно, то $\|h^\tau\|\xrightarrow[\tau\to\infty]{}0$.

Вернемся теперь к исходной задаче минимизации функции $F(x)=\max_s F^s(x)$ на множестве $X=\bigcap_s X_s$. Субдифференциал $\partial F(\bar{x})$ этой функции в точке $\bar{x}$ является выпуклой оболочкой субдифференциалов $\partial F^s(\bar{x})$ для $s\in S(\bar{x})=\{s/F^s(\bar{x})=F(\bar{x})\}$ [3]. В соответствии со схемой описанного процесса множеству $\partial F(\bar{x})$ сопоставляется многогранник $P(\bar{x})$, который будет выпуклой оболочкой многогранников $P^s(\bar{x})$, $s\in S(\bar{x})$. Непосредственное применение изложенной выше процедуры поиска подходящего направления к многограннику $P(\bar{x})$ в принципе возможно, но требует на каждом шаге этой процедуры решения, вообще говоря, всех задач (6) при $s\in S(\bar{x})$. Это вызвано тем, что при выполнении условий (4) в качестве z используется та из вершин v^α, на которой достигается минимум (h^τ, v^α). Оказывается, однако, что требование минимума здесь не является необходимым. Нужными свойствами будет обладать и следующая модификация описанной процедуры. При решении каждой задачи (6) выполнение неравенства (7) будем проверять на каждом шаге симплекс-

метода для очередного допустимого базисного вектора y^s. Решение задачи (6) следует прекратить как только нарушается неравенство (7). В этом случае полагаем $z = d^s - y^s B^s$ и переходим к п.4. Весь процесс поиска подходящего направления оканчивается, когда задача (6) доведена до оптимального решения, и неравенство (7) выполнено.

Обоснование этого варианта процедуры не многим отличается от приведенного в [I]. Без каких-либо изменений остается доказательство того факта, что если $0 \notin P^s(\bar{x})$, то процесс заканчивается через конечное число шагов, и последний вектор $(-h^\tau)$ определяет подходящее направление. Если же $0 \in P^s(\bar{x})$, то, как и раньше, показывается, что процесс будет бесконечным и для достаточно больших τ будет выполняться равенство

$$\|h^\tau\|^2 - \|h^{\tau+1}\|^2 = \frac{(h^\tau, h^\tau - z)^2}{\|h^\tau - z\|^2}.$$

Ввиду ограниченности многогранника $P^s(\bar{x})$ для достаточно больших значений K имеем $\|h^\tau - z\| \leq K$. Учитывая, что на каждом шаге процесса $(h^\tau, z) < \theta \|h^\tau\|^2$, получаем

$$\|h^\tau\|^2 - \|h^{\tau+1}\|^2 \geq \frac{(1-\theta)\|h^\tau\|^4}{K^2}.$$

Левая часть этого неравенства стремится к нулю ввиду сходимости последовательности $\|h^\tau\|$ (она убывает и ограничена снизу). А тогда и правая часть стремится к нулю, в результате получаем $\|h^\tau\| \xrightarrow[\tau \to \infty]{} 0$, что и требовалось показать.

Описанное видоизменение процесса поиска подходящего направления в применении к многограннику $P(\bar{x})$ позволяет отказаться от решения всех задач (6) на одном шаге процесса и делает возможным организовать процедуру по так называемому "барьерному" принципу. Один из вариантов такого процесса состоит в следующем.

Зафиксируем некоторый "барьер" $\rho \in [0, \theta]$ и будем при реализации процесса проверять неравенство (7), заменив в нем θ на ρ. При этом задачи (6) при $s \in S(\bar{x})$ будем перебирать в каком-либо заранее установленном порядке по циклу: если на шаге τ проверяемое неравенство нарушалось в задаче s_0, то

на шаге $\tau+1$ мы начинаем с этой же задачи. Если на полном просмотре всех задач нарушений указанного неравенства не обнаружено, то "барьер" увеличивается и перебор возобновляется. Так поступаем до тех пор, пока ρ не достигнет значения θ.

П. Описанной схемой возможных направлений можно воспользоваться, например, при построении итеративных алгоритмов для решения задач линейного программирования с окаймлением.

Рассмотрим задачу линейного программирования вида

$$
\begin{aligned}
&(c,x)+(d,z)\to min!\\
&Ax+Bz=b,\\
&Gx+Dz=g,\\
&x\geqslant 0,\ z\geqslant 0,
\end{aligned}
$$

здесь x, z – неизвестные векторы c,d,b,g – заданные векторы, A,B,G,D – заданные матрицы соответствующих размерностей. Предполагается, что матрица A имеет простую структуру, позволяющую использовать специальные алгоритмы линейного программирования. Один из возможных способов построения итеративных методов решения таких задач состоит в сведении исходной задачи к задаче поиска седловой точки функции $\varphi(z,w)$ на множестве $Z\times W$:

$$\varphi(z,w)=(d,z)+(w,g)-(w,Dz)+\min_{x\in X(z)}(c-wG,x),$$

где $X(z)=\{x/Ax=b-Bz,\ x\geqslant 0\}$, а множества Z и W определяются условиями непустоты множеств $X(z)$ и $Y(z)=\{y/yA\leqslant c-wG\}$. Аппроксимация множества W его дискретным подмножеством приводит нас к задаче, рассмотренной выше в первом разделе данной статьи; если же сделать дискретным множество Z, то получим аналогичную задачу, только речь будет идти о максимизации вогнутой функции, а не о минимизации выпуклой. Указанная ситуация возникает, в частности, при применении для поиска седловой точки функции $\varphi(z,w)$ алгоритма [4], который состоит в следующем.

Процесс начинается с произвольных $z^1\in Z$ и $w^1\in W$. К началу k-го шага имеются $z^1,z^2,\dots,z^k\in Z$ и $w^1,w^2,\dots,w^k\in W$. На k-м шаге рассматриваются две задачи:

$$M \to \min!$$

$$\varphi(x, w^i) \leq M, \quad i = 1, \ldots, k; \qquad (9)$$

$$x \in Z;$$

$$m \to \max! \qquad (10)$$

$$\varphi(x^j, w) \geq m, \quad j = 1, \ldots, k;$$

$$w \in W.$$

Векторы x и w, решающие эти задачи, принимаются в качестве x^{k+1} и w^{k+1} соответственно. Для такого процесса при условии компактности множеств Z и W легко показать, что последовательность значений $M^k = \max\limits_{1 \leq i \leq k} \varphi(x^{k+1}, w^i)$ не убывает, а последовательность значений $m^k = \min\limits_{1 \leq j \leq k} \varphi(x^j, w^{k+1})$ не возрастает; $M^k \leq m^k$ при всех k, и $\lim\limits_{k \to \infty} M^k = \lim\limits_{k \to \infty} m^k$. Кроме того, если через $\lambda_1^k, \ldots, \lambda_k^k$ и $\mu_1^k, \ldots, \mu_k^k$ обозначить множители Лагранжа ограничений (9) и (10), то любая точка сгущения последовательности $\left\{\left(\sum\limits_{j=1}^{k} \mu_j^k x^j; \sum\limits_{i=1}^{k} \lambda_i^k w^i\right)\right\}_{k=1}^{\infty}$ является седловой точкой функции $\varphi(x, w)$. Ясно, что для задачи получения M^k и m^k являются задачами рассмотренного выше (в п.I) типа. При этом, например, для задачи получения M^k будем иметь:

$$F^s(x) = \varphi(x, w^s), \; c^s = c - w^s G, \; d^s = d - w^s D, \; A^s = A, \; B^s = B, \; b^s = b.$$

То обстоятельство, что матрицы A^s, B^s и векторы b^s не зависят от s (вследствие чего множества Z_s также не будут меняться при изменении s), может быть учтено при разработке конкретных реализаций описанной схемы.

ЛИТЕРАТУРА

1. ШМЫРЕВА Н.В. Об одном декомпозиционном алгоритме линейного программирования.-Оптимизация,1977, вып.19(36),с.100-115.
2. ДЕМЬЯНОВ В.Ф., МАЛОЗЕМОВ В.Н. Введение в минимакс. - М.: Наука, 1972.
3. ПШЕНИЧНЫЙ Б.Н. Выпуклый анализ и экстремальные задачи. - М.: Наука, 1980.
4. STAHL J. Decomposition procedures for convex programs. - INFELOR Kozlemények, N10. Budapest, 1975, p.307-316.

Поступила в ред.-изд. отдел
06.01.1982 г.