

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

ИСПОЛЬЗОВАНИЕ ОВРАЖНОГО СПУСКА В БЛОЧНОМ ПРОГРАММИРОВАНИИ

А. И. ВЕРЕСКОВ

(Москва)

1. Рассматривается блочная задача линейного программирования

$$A^t x + D^t y^t = b^t, \quad (1)$$

$$x \geq 0, y^t \geq 0, t = 1, \dots, T, \quad (2)$$

$$(c, x) + \sum_{t=1}^T (f^t, y^t) \rightarrow \min. \quad (3)$$

Здесь A^t — матрицы размера $m \times n$; D^t — матрицы $m \times n_t$; D_j^t , $j = 1, \dots, n_t$ — столбцы матриц D^t . Заданные векторы b^t , c , f^t и неизвестные x , y^t — соответственно m , n , n_t и n_t -мерные векторы.

Для простоты изложения предположим, что при любом $x \geq 0$ задачи

$$[(f^t, y^t) \rightarrow \min | D^t y^t = b^t - A^t x, y^t \geq 0], t = 1, \dots, T \quad (4)$$

обладают планами.

Введем в рассмотрение функции $\varphi^t(x) = \{\min (f^t, y^t) | D^t y^t = b^t - A^t x, y^t \geq 0\}$, $t = 1, \dots, T$, определенные на неотрицательном октанте евклидова пространства E^n . Задача (1)–(3) эквивалентна задаче минимизации вы-

пуклой кусочно-линейной функции $\Phi(x) = (c, x) + \sum_{t=1}^T \varphi^t(x)$ при условии $x \geq 0$. Множество векторов, на котором $\Phi(x)$ достигает минимума, обозначим через M . Будем считать, что задача (1)–(3) разрешима, т. е. что $M \neq \emptyset$. В [1–3] рассматриваются различные варианты метода обобщенного градиента, который можно использовать для минимизации функции $\Phi(x)$. Если вектор $x^0 \geq 0$ не оптимален, то обобщенный градиентный спуск с достаточно малым постоянным шагом (на порядок меньшим, чем $\min_{x \in M} \|x^0 - x\|$) за конечное число шагов приводит нас в точку $x^1 \geq 0$, для которой $\Phi(x^1) < \Phi(x^0)$. Пусть $x^k \notin M$, и направление $\delta^k = \delta(x^k)$ (построенное при помощи обобщенного градиентного спуска или другим методом) является возможным, т. е. обеспечивает убывание функции $\Phi(x)$. Найдя $\min_{\lambda \geq 0} \Phi(x^k + \lambda \delta^k)$, получим новую точку x^{k+1} , для которой можно определить направление $\delta^{k+1} = \delta(x^{k+1})$ и т. д. Последовательность $\{x^k\}$, построенная таким полношаговым методом, вообще говоря, не сходится к множеству M из-за возможных зигзагообразных движений; для обеспечения сходимости пришлось бы использовать довольно сложные правила регулировки шага.

Если для любой точки $x \in M$ правило $\delta(x)$ позволяет получить возможное направление, то описанное ниже сочетание полношагового метода с движением по оврагу функции $\Phi(x)$ гарантирует достижение $\min \Phi(x)$ за конечное число итераций. Однако по существу предлагаемый метод является лишь приближенным, так как возможные направления $\delta(x)$ обычно удается получить только для точек, достаточно удаленных от множества M .

2. Пусть для фиксированного вектора $x \geq 0$ получены опорные решения задач (4): $y^t(x) = (y_1^t(x), \dots, y_{p_t}^t(x), 0, \dots, 0)$, $t = 1, \dots, T$, где $y_j^t(x) > 0$, $p_t \leq m$. Базис плана $y^t(x)$ обозначим через $B^t(x)$. Известно [4], что если базис задачи $[(f, y) \rightarrow \min | Dy = \tau, y \geq 0]$, оптимальный для $\tau = \tau^1$, допустим при $\tau = \tau^2$, то он оптимален также и для $\tau = \tau^2$. Если при $x = \bar{x}$ в задаче (4) $p_t = m$, то $b^t - A^t \bar{x} = \sum_{j=1}^m y_j^t(\bar{x}) D_j^t$, $p_j^t(x) > 0$ и векторы $D_1^t, \dots, D_{p_t}^t$ образуют базис пространства E^m . Поэтому в некоторой окрестности точки $\bar{x}$ базис $B^t(\bar{x})$ остается допустимым и сохраняет оптимальность, а значит, функция $\varphi^t(x)$ линейна в этой окрестности. При $p^t < m$ план $y^t(\bar{x})$ вырожден и функция $\varphi^t(x)$, вообще говоря, недифференцируема в точке $\bar{x}$. Однако если вектор z таков, что

$$A^t z = \sum_{j=1}^{p_t} \mu_j D_j^t, \quad \text{то } b^t - A^t(\bar{x} + \lambda z) = \sum_{j=1}^{p_t} (y_j^t(\bar{x}) - \lambda \mu_j) D_j^t,$$

и потому при близких к нулю значениях λ базис $B^t(\bar{x})$ задачи $[(f^t, y^t) \rightarrow \min | D^t y^t = b^t - A^t(\bar{x} + \lambda z), y^t \geq 0]$ допустим и оптимален, а последние $m - p_t$ компонент плана $y^t(\bar{x} + \lambda z)$ остаются равными нулю. Векторы z ,

удовлетворяющие условию $A^t z = \sum_{j=1}^{p_t} \mu_j D_j^t$, образуют подпространство

$Z^t(\bar{x})$ в E^n , и функция от z $\varphi^t(\bar{x} + z)$ линейна при $z \in Z^t(\bar{x})$ в окрестности начала координат. Подпространство $Z^t(\bar{x})$ задается уравнениями

$$(d_i^t, A^t z) = 0, \quad i = 1, \dots, m - p_t, \quad (5)$$

где $d_1^t, \dots, d_{m-p_t}^t$ — базис ортогонального дополнения в E^m подпространства, натянутого на векторы $D_1^t, \dots, D_{p_t}^t$. При $p_t = m$ подпространство $Z^t(\bar{x})$ совпадает с E^n . Если векторы $D_1^t, \dots, D_{p_t}^t$ — первые p_t столбцов базисной матрицы $\bar{D}^t$ задачи (4), то $d_1^t, \dots, d_{m-p_t}^t$ — последние $m - p_t$ строк матрицы $(\bar{D}^t)^{-1}$. Вычислительный процесс решения задачи (4) можно организовать так, чтобы матрица $(\bar{D}^t)^{-1}$ содержалась в симплексной таблице.

3. Выбрав для начального приближения $x = x^0$ возможное направление $\delta^0 = \delta(x^0)$ и решив задачу $[\Phi(x^0 + \lambda \delta^0) \rightarrow \min | \lambda \geq 0, x^0 + \lambda \delta^0 \geq 0]$, придем в точку x^1 , являющуюся точкой излома функции $\Phi(x)$ или граничной точкой октанта $x \geq 0$. Рассмотрим подпространство $Z_h = \bigcap_{t=0}^T Z^t(x^h)$, где

$Z^t(x^h)$, $t = 1, \dots, T$, определены выше, а $Z^0(x^h)$ задается условиями $z_j = 0$ при $x_j^h = 0$.

Если x изменяется в пределах линейного многообразия $x^h + Z_h = \{x | x - x^h \in Z_h\}$, то каждое направление является допустимым и функция $\Phi(x)$ линейна в окрестности точки x^h . Пусть $\Pi_h x$ — проекция вектора $x \in E^n$ на Z_h , $G(x)$ — вектор обобщенного градиента функции $\Phi(x)$; при

этом [1] $G(x) = c - \sum_{t=1}^T (A^t)' u^t$, где u^t — оптимальный план задачи, двойственной к (4). Вектор $-P_k G(x^k)$ дает направление скорейшего убывания функции $\Phi(x)$ в точке x^k на многообразии $x^k + Z_k$. Если $\gamma^1 = -P_1 G(x^1) \neq 0$, то функция $\Phi(x^1 + \lambda \gamma^1)$ убывает по крайней мере на отрезке $0 \leq \lambda \leq \lambda_1$, где λ_1 — минимальное значение λ , при котором обращается в нуль одна из ненулевых компонент векторов $x^1 + \lambda \gamma^1$, $y^t(x^1 + \lambda \gamma^1)$, $t = 1, \dots, T$. Движение от точки x^1 к точке $x^2 = x^1 + \lambda_1 \gamma^1$ естественно назвать овражным спуском. Эта итерация не является полношаговой, так как функция $\Phi(x^1 + \lambda \gamma^1)$, возможно, продолжала бы убывать и при $\lambda > \lambda_1$. Для точки x^2 построим вектор $\gamma^2 = -P_2 G(x^2)$ и т. д. до тех пор, пока $\gamma^k \neq 0$. Поскольку каждый овражный шаг связан с увеличением числа нулевых компонент вектора $(x^k, y^1(x^k), \dots, y^T(x^k))$, то через конечное число итераций окажется, что $\gamma^k = 0$.

Докажем, что если $\gamma^k = -P_k G(x^k) = 0$, то существует опорный план задачи (1)–(3) $(\tilde{x}, \tilde{y}^1, \dots, \tilde{y}^T)$ такой, что $\Phi(\tilde{x}) = \Phi(x^k)$. Пусть $\tilde{A}_j$, $j = 1, \dots, n$, и $\tilde{D}_j^t$, $j = 1, \dots, n$; $t = 1, \dots, T$, — Tm -мерные векторы условий задачи (1)–(3), $\tilde{b}$ — ее вектор ограничений. Выделим векторы условий $\tilde{A}_j$, $j = 1, \dots, r$; $\tilde{D}_j^t$, $j = 1, \dots, p$; $t = 1, \dots, T$, соответствующие положительным компонентам плана $(x^k, y^1(x^k), \dots, y^T(x^k))$, и рассмотрим задачу

$$\sum_{j=1}^r \tilde{A}_j x_j + \sum_{t=1}^T \sum_{j=1}^{p_t} \tilde{D}_j^t y_j^t = \tilde{b}, \quad (6)$$

$$x_j \geq 0, j = 1, \dots, r; y_j^t \geq 0, j = 1, \dots, p; t = 1, \dots, T, \quad (7)$$

$$\sum_{j=1}^r c_j x_j + \sum_{t=1}^T \sum_{j=1}^{p_t} f_j^t y_j^t \rightarrow \min. \quad (8)$$

Пусть $(\tilde{x}_1, \dots, \tilde{x}_r, \tilde{y}_1^1, \dots, \tilde{y}_{p_1}^1, \dots, \tilde{y}_1^T, \dots, \tilde{y}_{p_T}^T)$ — опорное решение задачи (6)–(8), $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_r, 0, \dots, 0) \in E^n$, $\tilde{y}^t = (\tilde{y}_1^t, \dots, \tilde{y}_{p_t}^t, 0, \dots, 0) \in$

E^{n_t} , $t = 1, \dots, T$. В силу блочной структуры условий (6) $\sum_{j=1}^{p_t} \tilde{D}_j^t \tilde{y}_j^t = b^t -$

$-\sum_{j=1}^r A_j^t \tilde{x}_j$, т. е. $D^t \tilde{y}^t = b^t - A^t \tilde{x}$, и базис $B^t(x^k)$, оптимальный для задачи (4)

при $x = x^k$, оптимален также и при $x = \tilde{x}$. Поэтому

$$\{\min(f^t, y^t) | D^t y^t = b^t - A^t \tilde{x}, y^t \geq 0\} = \sum_{j=1}^{p_t} f_j^t \tilde{y}_j^t \text{ и}$$

$$\Phi(\tilde{x}) = \sum_{j=1}^r c_j \tilde{x}_j + \sum_{t=1}^T \sum_{j=1}^{p_t} f_j^t \tilde{y}_j^t \leq \sum_{j=1}^r c_j x_j^k + \sum_{t=1}^T \sum_{j=1}^{p_t} f_j^t y_j^t(x^k) = (c, x^k) +$$

$$+ \sum_{t=1}^T (f^t, y^t(x^k)) = \Phi(x^k). \text{ Далее, } \sum_{j=1}^r \tilde{A}_j (\tilde{x}_j - x_j^k) = \sum_{t=1}^T \sum_{j=1}^{p_t} \tilde{D}_j^t (y_j^t(x^k) - \tilde{y}_j^t),$$

т. е. $A'(\tilde{x} - x^k) = \sum_{j=1}^{p^k} D_j'(y_j^t(x^k) - \tilde{y}_j^t)$, и $\tilde{x}_j - x_j^k = 0$ при $x_j^k = 0$, откуда

$\tilde{x} - x^k \in Z_k$. Из этого следует, что $\Phi(\tilde{x}) = \Phi(x^k)$, так как неравенство $\Phi(\tilde{x}) < \Phi(x^k)$ означало бы, что вдоль направления $\tilde{x} - x^k \in Z_k$ функция $\Phi(x)$ убывает при движении от точки x^k , а это противоречит условию $\gamma^k = 0$. Очевидно, вектор $(\tilde{x}, \tilde{y}^t, \dots, \tilde{y}^T)$ является опорным планом задачи (1) — (3); тем самым требуемое утверждение доказано.

Если $\Phi(x^k) > \min_{x \geq 0} \Phi(x)$, то, выбрав возможное направление $\delta^k = \delta(x^k)$, получим новое приближение $x = x^{k+1}$, $\Phi(x^{k+1}) < \Phi(x^k)$ и снова перейдем к овражному спуску. Таким образом, описанный процесс можно рассматривать как движение по вершинам многогранного множества (1), (2) через какие-то промежуточные точки. Значения функции $\Phi(x)$ при этом монотонно убывают, поэтому решение будет получено за конечное число итераций.

Выше отмечалось, что овражная итерация не является полношаговой: движение от точки x^i продолжается до тех пор, пока векторы $y^t(x^i + \lambda \gamma^t)$ сохраняют прежние базисы $B^t(x^i)$. Возможно, что такие итерации целесообразно чередовать с полношаговыми итерациями в направлении γ^t , обращаясь к описанному методу лишь тогда, когда обнаружено зигзагообразное поведение последовательности $\{x^i\}$.

4. Ниже излагается вариант метода штрафных функций, который может быть использован для решения задач выпуклого программирования общего вида и, в частности, для построения возможных направлений $\delta(x)$ в задаче (1) — (3).

Пусть c, x — n_0 -мерные векторы, h_i, y_i — n_i -мерные векторы, Q_0, R_i, Q_i — выпуклые замкнутые множества соответственно в $E^{n_0}, E^{n_i}, E^{n_0+n_i}, i = 1, \dots, m$. Будем считать, что задача выпуклого программирования

$$x \in Q_0, y_i \in R_i, \begin{pmatrix} x \\ y_i \end{pmatrix} \in Q_i, \quad i = 1, \dots, m, \quad (9)$$

$$(c, x) + \sum_{i=1}^m (h_i, y_i) \rightarrow \min$$

разрешима и множество ее оптимальных планов ограничено.

Рассмотрим задачу выпуклого программирования

$$x_0 \in Q_0, y_i \in R_i, \begin{pmatrix} x_0 \\ y_i \end{pmatrix} \in Q_i, \quad i = 1, \dots, m, \quad (10)$$

$$\frac{1}{m+1} \sum_{i=0}^m (c, x_i) + \sum_{i=1}^m (h_i, y_i) + f(P, x_0, x_1, \dots, x_m) \rightarrow \min.$$

Здесь $x_i \in E^{n_0}, P \in E^1, f(P, x_0, x_1, \dots, x_m)$ — функция, определенная на $E^{m n_0 + 1}$. Приведем без доказательства теорему, устанавливающую связь между задачами (9) и (10).

Теорема. Пусть функция $f(P, x_0, \dots, x_m)$ обладает следующими свойствами: 1) $\forall P > 0$ $f(P, x_0, \dots, x_m)$ выпукла по совокупности переменных $x_0, \dots, x_m$; 2) $\forall P > 0$ $f(P, x_0, \dots, x_m) \geq 0$, причем равенство достигается тогда и только тогда, когда $x_0 = x_1 = \dots = x_m$; 3) для фиксированного вектора $(x_0, \dots, x_m)$ такого, что некоторые $x_i \neq x_0$, $f(P, x_0, \dots, x_m) \rightarrow \infty$

при $P \rightarrow \infty$. Тогда последовательность оптимальных планов задачи (10) при $P \rightarrow \infty$ стремится к вектору $(\bar{x}, \dots, \bar{x}, \bar{y}_1, \dots, \bar{y}_m)$, где $(\bar{x}, \bar{y}_1, \dots, \bar{y}_m)$ — оптимальный план задачи (9).

Если в (9) отсутствуют переменные $y_i (n_i = 0)$, то описанный прием делает возможной декомпозицию условий задачи $[(c, x) \rightarrow \min | x \in \bigcap_{i=0}^m Q_i]$,

не обладающей блочной структурой. Штрафную функцию $f(P, x_0, \dots, x_m)$ целесообразно выбирать в зависимости от вида ограничений задачи (9).

Вернемся к задаче (1)–(3). Пусть $\gamma^k = -P_k G(x^k) = 0$ и нужно найти такой вектор x , что $\Phi(x) < \Phi(x^k)$. Нетрудно видеть, что (1) можно разбить на небольшое число групп так, чтобы операторы проектирования P_k^s , построенные для каждой из этих S групп, давали $P_k^s G(x^k) \neq 0$ хотя бы для некоторых $s = 1, \dots, S$. Пусть условия (1) разбиты на две группы: $t = 1, \dots, T_1$ и $t = T_1 + 1, \dots, T$. Рассмотрим задачу

$$\begin{aligned} A^t \xi + D^t y^t &= b^t, \quad t = 1, \dots, T_1, \\ A^t \eta + D^t y^t &= b^t, \quad t = T_1 + 1, \dots, T, \\ \xi &\geq 0, \quad \eta \geq 0, \quad y^t \geq 0, \quad t = 1, \dots, T, \end{aligned} \quad (11)$$

$$\frac{1}{2}(c, \xi + \eta) + \sum_{t=1}^T (f^t, y^t) + P \|\xi - \eta\|^2 \rightarrow \min. \quad (12)$$

Здесь $\xi, \eta \in E^n$. Для любого $P > 0$ можно выбрать малое $\lambda > 0$ такое, что векторы $\xi(\lambda) = x^k - \lambda P_k^1 G(x^k)$, $\eta(\lambda) = x^k - \lambda P_k^2 G(x^k)$ порождают план задачи (11), (12) $(\xi(\lambda), \eta(\lambda), y^1(\lambda), \dots, y^T(\lambda))$, обеспечивающий меньшее значение функции (12), чем план $(\xi(0), \eta(0), y^1(0), \dots, y^T(0))$. Положив $\xi^{(0)} = \xi(\lambda)$, $\eta^{(0)} = \eta(\lambda)$, $y^{t(0)} = y^t(\lambda)$, применим для решения задачи (11), (12) метод типа покомпонентного спуска: при фиксированных значениях $\xi = \xi^{(r)}$, $y^t = y^{t(r)}$, $t = 1, \dots, T_1$, находим векторы $\eta^{(r+1)}$, $y^{t(r+1)}$, $t = T_1 + 1, \dots, T$, удовлетворяющие условиям (11) и минимизирующие функцию (12); затем, зафиксировав найденные векторы, минимизируем (12) по переменным ξ, y^t , $t = 1, \dots, T_1$. Можно показать, что такой итеративный процесс сходится к решению задачи (11), (12), у которого, согласно приведенной выше теореме, ξ мало отличается от η при больших P . Таким образом, оборвав процесс после достаточного числа итераций и положив, например, $x = \frac{1}{2}(\xi^{(r)} + \eta^{(r)})$, получим $\Phi(x) < \Phi(x^k)$.

Автор благодарит Б. Т. Поляка за ряд ценных замечаний по содержанию работы.

ЛИТЕРАТУРА

1. Н. З. Шор, М. Б. Щепаккин. Алгоритм решения двухэтапной задачи стохастического программирования. Кибернетика, 1968, № 3.
2. Н. З. Шор. О скорости сходимости метода обобщенного градиентного спуска с растяжением пространства. Кибернетика, 1970, № 2.
3. Б. Т. Поляк. Минимизация негладких функционалов. Журнал вычислит. матем. и матем. физики, 1969, № 3.
4. Е. Г. Гольштейн, Д. Б. Юдин. Новые направления в линейном программировании. М., «Сов. радио», 1966.

Поступила в редакцию
29 IV 1971