ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—июнь 2005. Серия 2. Том 12, № 1, 3–11

УДК 519.87

ЭФФЕКТИВНЫЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ МИНИМИЗАЦИИ ПОЛИНОМОВ ОТ БУЛЕВЫХ ПЕРЕМЕННЫХ, ОБЛАДАЮЩИХ СВОЙСТВОМ СВЯЗНОСТИ*)

В. Л. Береснев

Рассматривается задача отыскания минимума функции с переменными, принимающими значения 0 и 1, представленной в виде полинома степени 1 по каждой переменной. Для решения данной задачи известны эффективные алгоритмы в случае, когда характеристическая матрица полинома обладает свойством связности, а коэффициенты при нелинейных членах положительны. В статье предлагается полиномиальный алгоритм решения задачи минимизации полиномов с характеристической матрицей, обладающей свойством связности, с коэффициентами произвольных знаков. Он построен на основе метода рекуррентных соотношений динамического программирования.

Введение

В задаче минимизации полиномов от булевых переменных [1] рассматривается функция $p(y_1, \ldots, y_m)$, в которой переменные принимают значения 0 и 1, а функция представлена в виде

$$p(y_1, \ldots, y_m) = c_0 + \sum_{i \in I} f_i(1 - y_i) + \sum_{t=1}^{T} c_t \prod_{i \in \gamma_t} y_i,$$

где $I = \{1, \ldots, m\}; \gamma_t \subset I, c_t \neq 0, t = 1, \ldots T; f_i \geqslant 0, i \in I$. Предполагается, что в данном представлении полинома $p(y_1, \ldots, y_m)$ отсутствуют подобные члены, т. е. множества $\gamma_t, t = 1, \ldots, T$, попарно различны и если $\gamma_t = \{i\}$ для некоторого $t, 1 \leqslant t \leqslant T$, то $c_t > 0$ и $f_i = 0$. Требуется

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 03–01–00455) и гранта по президентской программе поддержки ведущих научных школ РФ (проект НШ–313.2003.1).

найти такой неединичный (0,1)-вектор $(y_1^*, \dots, y_m^*)$, на котором функция $p(y_1, \dots, y_m)$ принимает наименьшее значение.

Матрицу $D = (d_{it})(i \in I, t = 1, \dots, T)$ такую, что

$$d_{it} = \begin{cases} 0, & \text{если } i \in \gamma_t, \\ 1 & \text{в противном случае,} \end{cases}$$

называют *характеристической матрицей* полинома $p(y_1, \dots, y_m)$.

Рассмотрим также задачу, эквивалентную задаче минимизации полиномов от булевых переменных, которая более удобна для дальнейших исследований. Эта задача, для которой используем обозначение MINF, состоит в отыскании (0,1)-вектора $(z_1^*, \dots, z_m^*)$, доставляющего минимум функции

$$f(z_1, \dots, z_m) = \sum_{i \in I} f_i z_i + \sum_{t=1}^{T} c_t \min_{i | z_i = 1} d_{it}.$$

Вектор $(z_1^*, \dots, z_m^*)$ является оптимальным решением задачи MINF тогда и только тогда, когда вектор $(y_1^*, \dots, y_m^*)$, где $y_i^* = 1 - z_i^*, i \in I$, — оптимальное решение задачи минимизации полинома $p(y_1, \dots, y_m)$.

В случае, когда коэффициенты $c_t, t = 1, \dots, T$, положительны, задача MINF эквивалентна задаче размещения предприятий с неограниченными мощностями [1, 3]. Поэтому задача MINF относится к числу NP-трудных. Одним из направлений исследования этой задачи, как и любой NP-трудной задачи дискретной оптимизации, является исследование её частных случаев, специфика которых позволяет строить для их решения полиномиальные алгоритмы.

Частные случаи экстремальных задач выделяются посредством наложения дополнительных условий на исходные данные. В случае задачи MINF такие условия накладываются на характеристическую матрицу D. Для задачи минимизации полиномов с положительными коэффициентами $c_t, t = 1, \dots, T$, известны два основных свойства матрицы D, при выполнении которых удается построить эффективные алгоритмы. Такими свойствами являются гриди–свойство [2,5] и свойство связности [3,4] характеристической матрицы D. Эти свойства позволяют строить алгоритмы решения задачи MINF с оценками временно́й сложности, равными соответственно $O(m(m + T))$ и $O(mT)$. Если матрица D обладает гриди–свойством, то знаки коэффициентов $c_t, t = 1, \dots, T$, не являются существенными и эффективный алгоритм решения задачи MINF в случае коэффициентов разных знаков построен в [2]. Оценка временно́й сложности этого алгоритма остается той же, что и для алгоритма в случае неотрицательных коэффициентов, и равна $O(m(m+T))$. В настоящей статье предлагается алгоритм решения задачи MINF с коэффициентами $c_t, t = 1, \ldots, T$, произвольных знаков в случае характеристической матрицы D, обладающей свойством связности. Алгоритм построен с использованием метода рекуррентных соотношений динамического программирования [3], его временна́я сложность не превосходит величины $O(m^2T)$.

1. Свойство связности

Матрица $D = (d_{it})(i \in I, t = 1, \ldots, T)$ обладает *свойством связности* (*является связной*), если для любых k и i из I разность $d_{kt} - d_{it}$ меняет знак при монотонном изменении t не более одного раза.

Пусть матрица $D = (d_{it})(i \in I, t = 1, \ldots, T)$ обладает свойством связности и в ней нет одинаковых строк. Заметим, что существует перестановка $\{i_1, \ldots, i_m\}$ множества I такая, что для любых l и $k, 1 \leqslant l < k \leqslant m$, разность $d_{i_l t} - d_{i_k t}$ либо не положительна для любого $t = 1, \ldots, T$, либо меняет знак с минуса на плюс при монотонно возрастающем изменении t.

Действительно, определим на множестве I бинарное отношение следования $\prec$ следующим образом: для любых k, i из I положим $k \prec i$, если первая ненулевая компонента вектора $(d_{kt} - d_{it})(t = 1, \ldots, T)$ отрицательна. Поскольку в матрице D нет одинаковых строк, для любых k, i из I выполняется одно из двух условий: либо $k \prec i$, либо $i \prec k$. Заметим также, что данное отношение является транзитивным, т. е. если $l \prec k$ и $k \prec i$, то $l \prec i$. В самом деле, пусть p и q — наименьшие номера такие, что $d_{lp} - d_{kp} < 0$ и $d_{kq} - d_{iq} < 0$. Пусть $t_0 = \min\{p, q\}$. Тогда $d_{lt_0} - d_{it_0} < 0$ и $d_{lt} - d_{it} = 0$ для $t = 1, \ldots, t_0 - 1$. Таким образом, множество I можно линейно упорядочить отношением $\prec$ и перестановка строк, полученная в результате такого упорядочения, будет искомой.

Пользуясь сказанным, при рассматрении (0,1)-матрицы $D = (d_{it})$ $(i \in I,\ t = 1, \ldots, T)$, обладающей свойством связности, будем считать, что в D нет одинаковых строк и подматрицы $D_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Пусть $1 \leqslant k < i \leqslant m$ и $1 \leqslant p < q \leqslant T$. Через $D(\{k, i\}, \{p, q\})$ будем обозначать подматрицу вида $\begin{pmatrix} d_{kp} & d_{kq} \\ d_{ip} & d_{iq} \end{pmatrix}$ матрицы D. Понятно, что если (0,1)-матрица D обладает свойством связности, то при любых k, i, p и q, таких, что $k < i$ и $p < q$, имеем $D(\{k, i\}, \{p, q\}) \neq D_0$.

Рассмотрим задачу MINF в случае, когда (0,1)-матрица $D = (d_{it})$ $(i \in I, t = 1, \ldots, T)$ обладает свойством связности. Для любых k, i из $I, k < i$, обозначим через $p(k, i)$ наибольший элемент $p, 1 \leqslant p \leqslant T$, такой, что $d_{kt} \leqslant d_{it}$ для любого $t \leqslant p$. Для любого $i \in I$ по определению положим $p(0, i) = 0$.

Для ненулевого (0,1)-вектора $(z_1, \ldots, z_m)$ введем следующие обозначения:

$$I_0 = \{i \in I \mid z_i = 1\},$$
$$I_t = \{i \in I_0 \mid d_{it} = \min_{k \in I_0} d_{kt}\}, t = 1, \ldots, T.$$

Наряду с исходной задачей MINF рассмотрим семейство задач MINF(q), $q = 1, \ldots, T$, где MINF(q) — задача минимизации на множестве ненулевых (0,1)-векторов $(z_1, \ldots, z_m)$ функции $f_q(z_1, \ldots, z_m)$ вида

$$f_q(z_1, \ldots, z_m) = \sum_{i \in I} f_i z_i + \sum_{t=1}^{q} c_t \min_{i | z_i = 1} d_{it}.$$

Пусть $(z_1, \ldots, z_m)$ — оптимальное решение задачи MINF(q). Заметим, что относительно этого решения можно считать, что для любого $i \in I_0$ найдется $t, 1 \leqslant t \leqslant q$, такое, что $I_t = \{i\}$. Действительно, пусть для некоторого $i_0 \in I_0$ при любом $t, 1 \leqslant t \leqslant q$, либо $i_0 \notin I_t$, либо $|I_t| > 1$. Рассмотрим решение $(z'_1, \ldots, z'_m)$, отличающееся от исходного решения $(z_1, \ldots, z_m)$ лишь тем, что $z'_{i_0} = 0$. Поскольку при замене решения $(z_1, \ldots, z_m)$ на решение $(z'_1, \ldots, z'_m)$ значение функции $f_q(z_1, \ldots, z_m)$ может только уменьшиться, новое решение $(z'_1, \ldots, z'_m)$ также будет оптимальным. Если и это решение не обладает нужным свойством, то продолжим процесс построения оптимального решения с меньшим числом единичных компонент.

При фиксированном $i_0 \in I$ рассмотрим такой (0,1)-вектор $(z_1, \ldots, z_m)$, что $z_{i_0} = 1, z_i = 0$ при $i > i_0$ и существует такое $t, 1 \leqslant t \leqslant q$, что $I_t = \{i_0\}$. Вектор $(z_1, \ldots, z_m)$, обладающий указанными свойствами, назовем i_0*-допустимым решением* задачи MINF(q). Нашей целью является отыскание i_0-допустимого решения, доставляющего минимум функции $f_q(z_1, \ldots, z_m)$. Такое решение назовем i_0*-оптимальным решением* задачи MINF(q). Понятно, что если для каждого $i_0 \in I$ найдено i_0-оптимальное решение задачи MINF(q), то лучшее из них будет оптимальным решением задачи MINF(q). Понятно также, что если $(z_1, \ldots, z_m)$ является i_0-оптимальным решением задачи MINF(q), то можно считать, что для любого $i \in I_0$ найдется такое $t, 1 \leqslant t \leqslant q$, что $I_t = \{i\}$.

Обозначим через $F_q(i_0)$, где $i_0 \in I$ и $1 \leqslant q \leqslant T$, значение функции $f_q(z_1, \ldots, z_m)$ на i_0-оптимальном решении задачи MINF(q). По определению положим $F_0(0) = 0$.

В следующих леммах говорится об основных свойствах, которыми обладают i_0-допустимые решения в случае, когда матрица D является связной, и которые используются при построении i_0-оптимальных решений.

Лемма 1. *Пусть* i_0, q *и* k *таковы, что* $i_0 \in I, 1 \leqslant q \leqslant T$, $1 \leqslant k < i_0$ *и* $p(k, i_0) < q$. *Тогда если* $(z'_1, \ldots, z'_m)$ — k*-допустимое решение задачи* $\mathrm{MINF}(p(k, i_0))$, *то для любого* $l, l < k$, $z'_l = 1$ *и каждого* $t, t > p(k, i_0))$, *выполняются неравенства* $d_{lt} \geqslant d_{kt} \geqslant d_{i_0 t}$.

Доказательство. Пусть $p_0 = p(k, i_0)$. Покажем, что $d_{kt} \geqslant d_{i_0 t}$, если $t > p_0$. В силу определения величины p_0 имеем $d_{kp_0+1} > d_{i_0 p_0+1}$. Поэтому если предположить, что $d_{kt} < d_{i_0 t}$ для некоторого $t, t > p_0+1$, то подматрица $D(\{k, i_0\}, \{p_0+1, t\})$ будет подматрицей вида D_0, что противоречит свойству связности матрицы D. Покажем, что $d_{lt} \geqslant d_{kt}$, если $t > p_0$. Так как $(z'_1, \ldots, z'_m)$ является k-допустимым решением задачи $\mathrm{MINF}(p_0)$, то для некоторого $p, p \leqslant p_0$, имеем $d_{lp} > d_{kp}$. Если же $d_{lt} < d_{kt}$ для некоторого $t, t > p_0$, то подматрица $D(\{l, k\}, \{p, t\})$ будет подматрицей вида D_0, что противоречит свойству связности матрицы D. Лемма 1 доказана.

Лемма 2. *Пусть* i_0, q *и* k *таковы, что* $i_0 \in I, 1 \leqslant q \leqslant T$, $1 \leqslant k < i_0$ *и* $p(k, i_0) < q$. *Тогда если* $(z'_1, \ldots, z'_m)$ — k*-допустимое решение задачи* $\mathrm{MINF}(p(k, i_0))$, *то вектор* $(z_1, \ldots, z_m)$, *где* $z_{i_0} = 1$ *и* $z_i = z'_i$ *при* $i \neq i_0$, *является* i_0*-допустимым решением задачи* $\mathrm{MINF}(q)$; *при этом*

$$f_q(z_1, \ldots, z_m) = f_{p(k,i_0)}(z'_1, \ldots, z'_m) + f_i + \sum_{t=p(k,i_0)+1}^{q} c_t d_{i_0 t}.$$

Доказательство. Пусть, как и выше, $p_0 = p(k, i_0)$. По лемме 1 при любом l таком, что $z_l = 1$, имеем $d_{lp_0+1} \geqslant d_{kp_0+1} > d_{i_0 p_0+1}$. Поэтому вектор $(z_1, \ldots, z_m)$ является i_0-допустимым решением задачи $\mathrm{MINF}(q)$. По определению p_0 для всякого $t \leqslant p_0$ имеем $d_{kt} \leqslant d_{i_0 t}$. Поэтому

$$f_{p_0}(z_1, \ldots, z_m) = f_{p_0}(z'_1, \ldots, z'_m) + f_{i_0}.$$

Из леммы 1 следует, что если $d_{lt} = 0$ для некоторых l и t таких, что $z_l = 1$ и $t > p_0$, то $d_{i_0 t} = 0$. Отсюда получаем, что $\min\limits_{i|z_i=1} d_{it} = d_{i_0 t}$. С учетом сказанного можно написать

$$\begin{aligned} f_q(z_1, \ldots, z_m) = f_{p_0}(z_1, \ldots, z_m) + \sum_{t=p_0+1}^{q} c_t \min_{i|z_i=1} d_{it} \\ = f_{p_0}(z'_1, \ldots, z'_m) + f_{i_0} + \sum_{t=p_0+1}^{q} c_t d_{i_0 t}. \end{aligned}$$

Лемма 2 доказана.

2. Рекуррентные соотношения

В следующей теореме устанавливается справедливость рекуррентных соотношений, которые позволяют вычислять i_0-оптимальные и, следовательно, оптимальные решения задач MINF(q).

Теорема. *При любых $i_0 \in I$ и $q, 1 \leqslant q \leqslant T$, справедливо равенство*

$$F_q(i_0) = \min_{\substack{0 \leqslant k < i_0 \\ p(k,i_0) < q}} \left\{ F_{p(k,i_0)}(k) + f_i + \sum_{t=p(k,i_0)+1}^{q} c_t d_{i_0 t} \right\}.$$

Доказательство. Сначала покажем, что если k таково, что $0 \leqslant k < i_0$ и $p(k, i_0) < q$, то справедливо неравенство

$$F_q(i_0) \leqslant F_{p(k,i_0)}(k) + f_i + \sum_{t=p(k,i_0)+1}^{q} c_t d_{i_0 t}.$$

Если $k = 0$, то вектор $(z_1, \ldots, z_m)$, в котором $z_{i_0} = 1$ и $z_i = 0$ при $i \neq i_0$, является i_0-допустимым решением задачи MINF(q) и

$$F_q(i_0) \leqslant f_q(z_1, \ldots, z_m) = F_0(0) + f_i + \sum_{t=1}^{q} c_t d_{i_0 t}.$$

Пусть $k > 0$ и $p_0 = p(k, i_0)$. Рассмотрим k-оптимальное решение $(z'_1, \ldots, z'_m)$ задачи MINF(p_0) и вектор $(z_1, \ldots, z_m)$, в котором $z_{i_0} = 1$ и $z_i = z'_i$ при $i \neq i_0$. По лемме 2 вектор $(z_1, \ldots, z_m)$ является i_0-допустимым решением задачи MINF(q) и для этого решения справедливы соотношения

$$\begin{aligned} F_q(i_0) \leqslant f_q(z_1, \ldots, z_m) = f_{p_0}(z'_1, \ldots, z'_m) + f_{i_0} + \sum_{t=p_0+1}^{q} c_t d_{i_0 t} \\ = F_{p_0}(i_0) + f_{i_0} + \sum_{t=p_0+1}^{q} c_t d_{i_0 t}. \end{aligned}$$

Убедимся в существовании такого k, что $0 \leqslant k < i, p(k, i_0) < q$ и

$$F_q(i_0) = F_{p(k,i_0)}(k) + f_i + \sum_{t=p(k,i_0)+1}^{q} c_t d_{i_0 t}.$$

Для этого рассмотрим i_0-оптимальное решение $(z_1, \ldots, z_m)$ задачи MINF(q). Если для этого решения $I_0 = \{i_0\}$, то при $k = 0$ требуемое равенство выполняется. Пусть $|I_0| \geqslant 2$ и k — наибольший элемент множества $I_0 \setminus \{i_0\}$. Положим $p_0 = p(k, i_0)$. Так как $I_j = \{i_0\}$ для некоторого $j \leqslant q$, то $p_0 < q$. Рассмотрим вектор $(z'_1, \ldots, z'_m)$, отличающийся от $(z_1, \ldots, z_m)$ лишь тем, что $z'_{i_0} = 0$. Так как $(z_1, \ldots, z_m)$ является i_0-оптимальным решением задачи MINF(q), то $(z'_1, \ldots, z'_m)$ — k-допустимое решение задачи MINF(p_0). Покажем, что это решение является k-оптимальным решением задачи MINF(p_0). Предположим противное и пусть $(x'_1, \ldots, x'_m)$ — k-оптимальное решение задачи MINF(p_0). Рассмотрим вектор $(x_1, \ldots, x_m)$, отличающийся от $(x'_1, \ldots, x'_m)$ лишь тем, что $x_{i_0} = 1$. Согласно лемме 2 вектор $(x_1, \ldots, x_m)$ является i_0-допустимым решением. Тогда можно записать

$$\begin{aligned} f_q(x_1, \ldots, x_m) &= f_{p_0}(x'_1, \ldots, x'_m) + f_i + \sum_{t=p_0+1}^{q} c_t d_{i_0 t} \\ &< f_{p_0}(z'_1, \ldots, z'_m) + f_i + \sum_{t=p_0+1}^{q} c_t d_{i_0 t} = f_q(z_1, \ldots, z_m). \end{aligned}$$

Это противоречит тому, что $(z_1, \ldots, z_m)$ — i_0-оптимальное решение задачи MINF(q). Поэтому $(z'_1, \ldots, z'_m)$ является k-оптимальным решением задачи MINF(p_0) и, следовательно, справедливо равенство

$$F_q(i_0) = F_{p_0}(k) + f_{i_0} + \sum_{t=p_0+1}^{q} c_t d_{i_0 t}.$$

Теорема доказана.

Для любых $i_0 \in I$ и $q, 1 \leqslant q \leqslant T$, положим

$$k_q(i_0) = \arg \min_{\substack{0 \leqslant k < i_0 \\ p(k, i_0) < q}} \left\{ F_{p(k,i_0)}(k) + f_i + \sum_{t=p(k,i_0)+1}^{q} c_t d_{i_0 t} \right\}.$$

Из теоремы при любых $i_0 \in I$ и $q, 1 \leqslant q \leqslant T$, вытекает следующее. Пусть $k = k_q(i_0)$ и $p = p(k, i_0)$. Тогда при $k \neq 0$ i_0-оптимальным решением задачи MINF(p_0) является такой вектор $(z_1, \ldots, z_m)$, что $z_{i_0} = 1$, а вектор $(z_1, \ldots, z_k, 0, \ldots, 0)$ является k-оптимальным решением задачи MINF(p_0). Если $k = 0$, то i_0-оптимальным решением является вектор $(z_1, \ldots, z_m)$, в котором $z_{i_0} = 1$ и $z_i = 0$ при $i \neq i_0$.

3. Алгоритм построения оптимального решения

Алгоритм вычисления оптимального решения задачи MINF в случае матрицы D, обладающей свойством связности, состоит из двух этапов. На первом этапе по полученному рекуррентному соотношению для каждого $t = 1, \ldots, T$ последовательно вычисляются величины $F_t(i)$ и номера $k_t(i)$, $i \in I$. На втором этапе с помощью вычисленных на первом этапе номеров $k_t(i)$ определяются единичные компоненты оптимального решения.

Первый этап включает предварительный шаг и T основных шагов.

На предварительном шаге для любых k, i из $I, k < i$, вычисляются номера $p(k, i)$.

На t-м, $t = 1, \ldots, T$, основном шаге для каждого $i \in I$ с использованием рекуррентного соотношения вычисляются величина $F_t(i)$ и номер $k_t(i)$.

Второй этап состоит из предварительного шага и конечного числа однотипных основных шагов.

На предварительном шаге вычисляется номер $i_1 \in I$ такой, что $F_T(i_1) = \min\limits_{i \in I} F_T(i)$. После этого полагается $p_1 = T$, $z_i = 0$ для $i = i_1 + 1, \ldots, m$ и начинаются основные шаги.

На очередном ℓ-м основном шаге, $\ell = 1, 2, \ldots$, имеются номера i_ℓ, p_ℓ и определяются номера $i_{\ell+1} = k_{p_\ell}(i_\ell)$ и $p_{\ell+1} = p(i_{\ell+1}, i_\ell)$. Далее полагается $z_{i_\ell} = 1$ и $z_i = 0$ для $i = i_{\ell+1} + 1, \ldots, i_\ell - 1$. После этого при $i_{\ell+1} > 0$ начинается следующий шаг; в противном случае второй этап и алгоритм в целом заканчивают работу и $(z_1, \ldots, z_m)$ — искомое оптимальное решение.

Нетрудно видеть, что для реализации каждого шага первого этапа требуется $O(m^2)$ действий. Поэтому временна́я сложность первого этапа не превосходит $O(m^2T)$. Временна́я сложность предварительного шага второго этапа не превосходит величины $O(m)$, а всех основных шагов второго этапа — величины $O(T)$. Таким образом, в целом временна́я сложность рассмотренного алгоритма построения оптимального решения задачи MINF не превосходит $O(m^2T)$.

ЛИТЕРАТУРА

1. **Береснев В. Л.** Алгоритмы минимизации полиномов от булевых переменных // Проблемы кибернетики. Вып. 36. М.: Наука, 1979. С. 225–246.

2. **Береснев В. Л.** Эффективный алгоритм для задачи размещения производства с вполне уравновешенной матрицей // Дискрет. анализ и исслед. операций. Сер. 1. 1998. Т. 5, № 1. С. 20–31.

3. **Береснев В. Л., Гимади Э. Х., Дементьев В. Т.** Экстремальные задачи стандартизации. Новосибирск: Наука, 1978.
4. **Гимади Э. Х.** Выбор оптимальных шкал в одном классе задач типа размещения, унификации и стандартизации // Управляемые системы. Сб. науч. тр. Вып. 6. Новосибирск: Институт математики СО АН СССР, 1970. С. 57–70.
5. **Hoffman A. J., Kolen A. W. G., Sakarovitch M.** Totally-balanced and greedy matrices // SIAM J. Algebraic Discrete Methods. 1985. V. 6, N 4. P. 721–730.

Адрес автора:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск, Россия.
E–mail: beresnev@math.nsc.ru

Статья поступила
5 апреля 2005 г.