

1 Задача минимизации полиномов от булевых переменных

Вещественная функция $p(y_1, \dots, y_m)$ от переменных, принимающих значение 0 и 1, представленная в виде

$$p(y_1, \dots, y_m) = \sum_{\alpha \subset I} c_\alpha \prod_{i \in \alpha} y_i,$$

где $I = \{1, \dots, m\}$, а c_α — действительные числа, как сказано ранее, называется *полиномом от булевых переменных*.

Будем говорить, что полином от булевых переменных $p(y_1, \dots, y_m)$ задан в *канонической форме*, если он представлен в виде

$$p(y_1, \dots, y_m) = a_0 + \sum_{r=1}^R a_r (1 - \prod_{i \in \alpha_r} y_i) + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i,$$

где $a_r > 0$, $r = 1, \dots, R$; $b_s > 0$, $s = 1, \dots, S$; α_r , $r = 1, \dots, R$, — попарно различные подмножества множества I и β_s , $s = 1, \dots, S$, — также попарно различные подмножества множества I . Если множества α_r , $r = 1, \dots, R$, и β_s , $s = 1, \dots, S$, совместно попарно различны, то каноническую форму будем называть *формой без подобных членов*.

Пусть полином $p(y_1, \dots, y_m)$ задан в канонической форме без подобных членов. Пару $(0,1)$ -матриц (G, H) , $G = (g_{ir})$ ($i \in I$, $r = 1, \dots, R$), $H = (h_{is})$, ($i \in I$, $s = 1, \dots, S$) таких, что

$$g_{ir} = \begin{cases} 0, & \text{если } i \in \alpha_r, \\ 1 & \text{иначе,} \end{cases}$$

$$h_{is} = \begin{cases} 0, & \text{если } i \in \beta_s, \\ 1 & \text{иначе,} \end{cases}$$

назовем *парой характеристических матриц* полинома. При этом коэффициенты a_1, \dots, a_R будем называть *весами столбцов* матрицы G , а коэффициенты b_1, \dots, b_S — *весами столбцов* матрицы H .

Задача минимизации полинома $p(y_1, \dots, y_m)$, которая ранее обозначена как *MINP*, имеет вид

$$\min \{ p(y_1, \dots, y_m) \mid y_i \in \{0, 1\}, i \in I \}.$$

Исходными данными для задачи *MINP* считаем пару матриц (\tilde{F}, \tilde{C}) , $\tilde{F} = (a_r(1 - g_{ir}))$ ($i \in I$, $r = 1, \dots, R$), $\tilde{C} = (b_s h_{is})$ ($i \in I$, $s = 1, \dots, S$), где a_r , $r = 1, \dots, R$, и b_s , $s = 1, \dots, S$, — положительные величины, а $G = (g_{ir})$ ($i \in I$, $r = 1, \dots, R$) и $H = (h_{is})$, ($i \in I$, $s = 1, \dots, S$) — $(0,1)$ -матрицы. Полином $p(y_1, \dots, y_m)$, задаваемый парой матриц (\tilde{F}, \tilde{C}) , записывается следующим образом:

$$p(y_1, \dots, y_m) = a_0 + \sum_{r=1}^R a_r (1 - \prod_{i | g_{ir}=0} y_i) + \sum_{s=1}^S b_s \prod_{i | h_{is}=0} y_i.$$

Если \tilde{F} — диагональная матрица с диагональными элементами $f_i \geq 0$, $i \in I$, то полином $p(y_1, \dots, y_m)$ имеет вид

$$p(y_1, \dots, y_m) = a_0 + \sum_{i \in I} f_i (1 - y_i) + \sum_{s=1}^S b_s \prod_{i|h_{is}=0} y_i$$

и является тем самым полиномом с неотрицательными коэффициентами, задаваемым парой (F_0, \tilde{C}) , где $F_0 = (f_i)$ ($i \in I$).

Возвращаясь к задаче *MINP*, отметим, что поскольку каждая из величин R и S может быть сравнима с числом $(0,1)$ -векторов (y_1, \dots, y_m) , то длина записи исходных данных задачи *MINP* сравнима с числом допустимых решений этой задачи. Так что для решения задачи *MINP* существует эффективный и простой алгоритм решения — перебор допустимых решений. В этом смысле задача *MINP* в общем случае представляется неинтересной. Проблема становится нетривиальной для классов полиномов, у которых величины R и S ограничены полиномами от числа переменных m . Поэтому, говоря о задаче *MINP*, будем подразумевать именно такие подзадачи задачи *MINP*.

Задача *MINP*₀ является частным случаем задачи *MINP*. Это означает, что задача *MINP* также как и задача *MINP*₀ является NP-трудной. Вместе с тем задача *MINP* не является более трудной, чем задача *MINP*₀.

Теорема 4.1.1. Задача *MINP* сводится к задаче *MINP*₀.

Доказательство. Исходному полиному

$$p(y_1, \dots, y_m) = a_0 + \sum_{r=1}^R a_r (1 - \prod_{i \in \alpha_r} y_i) + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i$$

поставим в соответствие полином от булевых переменных $\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ вида

$$\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R) = a_0 + \sum_{r=1}^R a_r \tilde{y}_r + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i + \Phi \sum_{r=1}^R \sum_{i \in \alpha_r} (1 - \tilde{y}_r)(1 - y_i),$$

где Φ — достаточно большая положительная величина такая, что $\Phi > \sum_{r=1}^R a_r$.

Заметим, что коэффициенты при нелинейных членах у данного полинома неотрицательные. В этом нетрудно убедиться, если переписать рассматриваемый полином следующим образом:

$$\begin{aligned} \tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R) &= a_0 + \sum_{r=1}^R a_r \tilde{y}_r + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i + \\ &+ \sum_{r=1}^R \Phi |\alpha_r| - \sum_{r=1}^R \Phi |\alpha_r| \tilde{y}_r - \sum_{i=1}^m \Phi |\{r | i \in \alpha_r\}| \cdot y_i + \sum_{r=1}^R \sum_{i \in \alpha_r} \Phi \tilde{y}_r y_i = \\ &= A_0 + \sum_{r=1}^R A_r (1 - \tilde{y}_r) + \sum_{i=1}^m B_i (1 - y_i) + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i + \sum_{r=1}^R \sum_{i \in \alpha_r} \Phi \tilde{y}_r y_i, \end{aligned}$$

где

$$\begin{aligned} A_0 &= a_0 + \sum_{r=1}^R a_r - \Phi \sum_{i=1}^m |\{r \mid i \in \alpha_r\}|; \\ A_r &= \Phi |\alpha_r| - a_r, \quad r = 1, \dots, R; \\ B_i &= \Phi |\{r \mid i \in \alpha_r\}|, \quad i = 1, \dots, m. \end{aligned}$$

Покажем, что если $(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ — оптимальное решение задачи $MINP$ для полинома $\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$, то (y_1, \dots, y_m) — оптимальное решение задачи $MINP_0$ полинома $p(y_1, \dots, y_m)$.

Пусть $(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ — оптимальное решение. В силу выбора величины Φ имеем для всякого $r = 1, \dots, R$

$$(1 - \tilde{y}_r)(1 - y_i) = 0, \quad i \in \alpha_r,$$

и, следовательно,

$$\tilde{y}_r = 1 - \prod_{i \in \alpha_r} y_i.$$

Отсюда получаем, что для решения (y_1, \dots, y_m) , построенного по оптимальному решению $(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ имеет место равенство

$$p(y_1, \dots, y_m) = \tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R).$$

Пусть теперь (y_1, \dots, y_m) — оптимальное решение задачи $MINP$ полинома $p(y_1, \dots, y_m)$. Рассмотрим решение $(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ задачи $MINP_0$ полинома $\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$, определенное следующим образом:

$$\tilde{y}_r = 1 - \prod_{i \in \alpha_r} y_i, \quad r = 1, \dots, R.$$

Заметим, что для всякого $r = 1, \dots, R$ и любого $i \in \alpha_r$ выполняются равенства

$$(1 - y_i)(1 - \tilde{y}_r) = (1 - y_i) \prod_{i \in \alpha_r} y_i = 0,$$

из которых следует, что

$$\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R) = p(y_1, \dots, y_m).$$

Таким образом, в силу принципа оптимальности (лемма 1.1.1) получаем, что если $(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$ — оптимальное решение задачи $MINP_0$ полинома $\tilde{P}_0(y_1, \dots, y_m, \tilde{y}_1, \dots, \tilde{y}_R)$, то (y_1, \dots, y_m) — оптимальное решение задачи $MINP$ исходного полинома $p(y_1, \dots, y_m)$, что доказывает требуемую сводимость.

Таким образом, полиномы с неотрицательными коэффициентами при нелинейных членах — это класс полиномов, для которых задача минимизации содержит все трудности задачи минимизации полиномов общего вида. В следующем параграфе будет построен подкласс полиномов с неотрицательными коэффициентами, для которых задача миними-

зации эффективно разрешима, и указан «симметричный» класс полиномов, задача минимизации которых также является полиномиально разрешимой.

2 Полиномы с разделяющимися переменными

В настоящем параграфе рассматривается класс полиномов, задача минимизации которых эквивалентна задаче о максимальном потоке в сети и двойственной ей задаче о минимальном разрезе в сети []. Эти задачи хорошо изучены и для их решения построены различные эффективные алгоритмы. Более того, задачу о максимальном потоке можно считать одной из достаточно общих экстремальных задач, для которых найдены эффективные алгоритмы.

2.1. Задача о максимальном потоке в сети

Рассмотрим конечную ориентированную двухполюсную сеть $G = (V, E)$ с множеством вершин V , множеством дуг E , источником $s \in V$ и стоком $t \in V$. Для каждого ребра $e \in E$ обозначим через $v_1(e) \in V$ начальную вершину, а через $v_2(e) \in V$ — конечную вершину. Для каждой вершины $v \in V$ обозначим через $A(v)$ множество $\{e \in E \mid v_1(e) = v\}$, то есть множество ребер, для которых вершина v — начальная, а через $B(v)$ множество $\{e \in E \mid v_2(e) = v\}$, то есть множество ребер, для которых вершина v — конечная.

Считаем, что для каждой дуги $e \in E$ задана величина a_e , называемая *пропускной способностью* дуги. *Разрезом* (V_1, V_2) называется разбиение множества вершин сети V на два подмножества V_1 и V_2 такие, что $s \in V_1$ и $t \in V_2$. *Пропускной способностью* $\gamma(V_1, V_2)$ разреза (V_1, V_2) называется сумма пропускных способностей всех дуг сети, начальные вершины которых лежат в V_1 , а конечные в V_2 . Пропускная способность разреза равняется

$$\gamma(V_1, V_2) = \sum_{\substack{e=(v_1, v_2) \\ v_1 \in V_1, v_2 \in V_2}} a_e.$$

Разрез в сети G с наименьшей пропускной способностью называется *минимальным*, а задача поиска минимального разреза в сети G — *задачей о минимальном разрезе*.

Двойственной к задаче о минимальном разрезе является задача о максимальном потоке в сети G []. Для всякого $e \in E$ рассмотрим величину f_e , обозначаемую *величину потока* по дуге e . С использованием переменных f_e , $e \in E$, задача о максимальном потоке в сети G записывается следующим образом:

$$\begin{aligned} & \max_{(f_e)} \sum_{e \in A(s)} f_e; \\ & \sum_{e \in B(i)} f_e - \sum_{e \in A(i)} f_e = 0, \quad v \in V \setminus \{s, t\}; \\ & 0 \leq f_e \leq a_e, \quad e \in E. \end{aligned}$$

Данная задача является эффективно разрешимой и для ее решения построен целый ряд полиномиальных алгоритмов. Лучшим является алгоритм из [], для которого оценка

трудоемкости равна $O(|V|^3)$. В случае, когда сеть $G = (V, E)$ — двудольная, а V_1 и V_2 — соответствующие подмножества вершин, то оценка равна $O(|V_1|^3 + |V_2|^3)$.

Рассмотрим квадратичный полином $q_1(y_1, \dots, y_m)$ от $(0,1)$ -переменных вида

$$q_1(y_1, \dots, y_m) = a_0 + \sum_{i=1}^{m-1} \sum_{k=i+1}^m c_{ik} (1 - y_i y_k) + \sum_{i=1}^m b_i y_i, \quad (4.2.1)$$

где $b_i \geq 0$, $i = 1, \dots, m$; $c_{ik} \geq 0$, $i = 1, \dots, m-1$, $k = i+1, \dots, m$.

Это есть квадратичный полином с неположительными коэффициентами при нелинейных членах общего вида. Следующая теорема показывает, что задача *MINP* для таких полиномов эквивалентна задаче о минимальном разрезе. Тем самым дается формулировка задачи о максимальном потоке в терминах полиномов от булевых переменных.

Теорема 4.2.1. Задача о минимальном разрезе и задача *MINP* для квадратичного полинома с неположительными коэффициентами эквивалентны.

Доказательство. Пусть $G = (V, E)$ — двухполюсная сеть. Поставим ей в соответствие полином $q((y_v))$ от булевых переменных $y_v, v \in V$, вида

$$q((y_v)) = \sum_{e=(v_1, v_2)} a_e y_{v_2} (1 - y_{v_1})$$

и рассмотрим задачу минимизации этого полинома при дополнительном условии $y_s = 0$, $y_t = 1$. Легко видеть, что данный полином есть полином вида (4.2.1). Разрез (V_1, V_2) сети G и допустимое решение (y_v) задачи минимизации полинома $q((y_v))$ при условии $y_s = 0$, $y_t = 1$ назовем соответствующим, если

$$y_v = \begin{cases} 0, & \text{если } v \in V_1, \\ 1, & \text{если } v \in V_2. \end{cases}$$

Понятно, что по всякому разрезу легко строится допустимое решение и, наоборот, по произвольному допустимому решению легко строится соответствующий ему разрез.

Заметим, что для соответствующих разреза (V_1, V_2) и допустимого решения (y_v) равенство $y_{v_2} (1 - y_{v_1}) = 1$ возможно тогда и только тогда, когда $v_1 \in V_1$ и $v_2 \in V_2$. Следовательно, значения целевых функций рассматриваемых задач на соответствующих разрезе и допустимом решении совпадают. Отсюда по признаку оптимальности следует, что если (y_v) — оптимальное решение, то соответствующий разрез (V_1, V_2) является минимальным и наоборот. Таким образом, задача о минимальном разрезе сводится к задаче *MINP* полинома (4.2.1).

Покажем обратную сводимость. Полиному $q_1(y_1, \dots, y_m)$ вида (4.2.1) поставим в соответствие двухполюсную сеть $G = (V, E)$ (см. рис. 4.2.1) с множеством вершин

$$V = \{1, \dots, m\} \cup \{s, t\},$$

множеством дуг

$$E = \{(i, k), 1 \leq i < k \leq m\} \cup \{(s, i), 1 \leq i \leq m\} \cup \{(k, t), 1 \leq k \leq m\}$$

и пропускными способностями дуг, определяемыми следующим образом:

$$a_{(i,k)} = c_{ik}, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m;$$

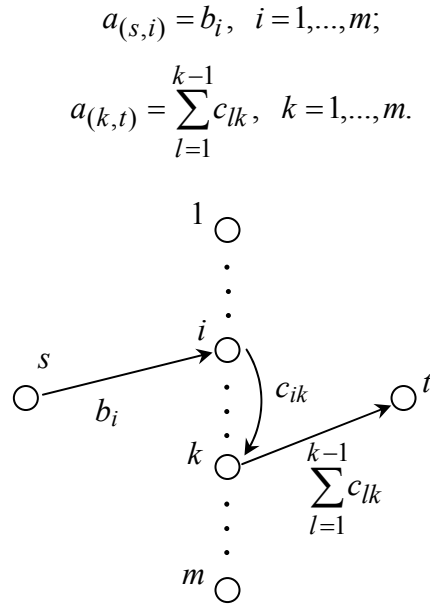


Рис. 4.2.1.

Для построенной сети G рассмотрим полином $q((y_v))$ и покажем, что этот полином при значениях переменных y_s и y_t , равных соответственно 0 и 1, совпадает с полиномом (4.2.1). Действительно, справедливы следующие равенства:

$$\begin{aligned} \sum_{e=(v_1, v_2)} a_e y_{v_2} (1 - y_{v_1}) &= \sum_{i=1}^{m-1} \sum_{k=i+1}^m c_{ik} y_k (1 - y_i) + \sum_{i=1}^m b_i y_i + \sum_{k=1}^m \sum_{l=1}^{k-1} c_{lk} (1 - y_k) = \\ &= \sum_{i=1}^{m-1} \sum_{k=i+1}^m c_{ik} y_k (1 - y_i) + \sum_{i=1}^m b_i y_i + \sum_{l=1}^{m-1} \sum_{k=l+1}^m c_{lk} (1 - y_k) = \sum_{i=1}^m b_i y_i + \sum_{i=1}^{m-1} \sum_{k=i+1}^m c_{ik} (1 - y_i y_k). \end{aligned}$$

Отсюда получаем, что если (V_1, V_2) наименьший разрез сети G , то соответствующее ему допустимое решение (y_v) задачи минимизации полинома $q((y_v))$ при условии $y_s = 0$, $y_t = 1$ является оптимальным и, следовательно, полученный из этого решения вектор (y_1, \dots, y_m) является оптимальным решением задачи $MINP$ полинома (4.2.1). Таким образом, задача $MINP$ полинома (4.2.1) сводится к задаче о минимальном разрезе и, следовательно, эти задачи эквивалентны. Теорема доказана.

2.2. Полиномы, задача минимизации которых сводится к задаче о минимальном разрезе

Рассмотрим квадратичный полином $q_2(y_1, \dots, y_m, z_1, \dots, z_n)$ от $(0,1)$ -переменных вида:

$$q_2(y_1, \dots, y_m, z_1, \dots, z_n) = \sum_{i=1}^m a_i (1 - y_i) + \sum_{k=1}^n b_k (1 - z_k) + \sum_{i=1}^m \sum_{k=1}^n c_{ik} y_i z_k, \quad (4.2.2)$$

где $a_i \geq 0$, $b_k \geq 0$, $c_{ik} \geq 0$, $i = 1, \dots, m$, $k = 1, \dots, n$.

Такой полином назовем полиномом с разделяющимися переменными.

Покажем, что задача $NIMP_0$ для полинома с разделяющимися переменными эффективно разрешима.

Теорема 4.2.2. Задача $MINP$ для полинома с разделяющимися переменными сводится к задаче о минимальном разрезе

Доказательство. Покажем, что задача $NIMP_0$ полинома (4.2.2) сводится к задаче $MINP$ полинома (4.2.1). Для этого заметим, прежде всего, что для коэффициентов полинома (4.2.2) можно считать справедливыми следующие соотношения

$$b_k \leq \sum_{i=1}^m c_{ik}, \quad k = 1, \dots, n.$$

В самом деле, если при некотором k , $1 \leq k \leq n$, данное неравенство не выполняется, то в оптимальном решении задачи $MINP$ полинома (4.2.2) значение переменной z_k равняется единице. Заметим далее, что заменой переменных $y_i = 1 - y'_i$, $i = 1, \dots, m$, в полиноме (4.2.2) позволяет переписать его следующим образом:

$$a_0 + \sum_{i=1}^m a_i y'_i + \sum_{k=1}^n b'_k z_k + \sum_{i=1}^m \sum_{k=1}^n c_{ik} (1 - y'_i z_k),$$

где

$$b'_i = \sum_{k=1}^n c_{ik} - b_i, \quad i = 1, \dots, m.$$

Этот полином является, очевидно, полиномом вида (4.2.1), что с учетом теоремы 4.2.1 доказывает требуемое. Однако для получения в явном виде сети, минимальный разрез, которой дает оптимальное решение задачи $MINP$ полинома с разделяющимися переменными, продолжим доказательство и построим требуемую сеть.

Полиному (4.2.2) поставим в соответствие двухполюсную сеть $G = (V, E)$ (см. рис. 4.2.2) с множеством вершин

$$V = \{1, \dots, m\} \cup \{1, \dots, n\} \cup \{s, t\},$$

множеством дуг

$$E = \{(i, k), 1 \leq i \leq m, 1 \leq k \leq n\} \cup \{(s, i), 1 \leq i \leq m\} \cup \{(k, t), 1 \leq k \leq n\}$$

и пропускной способностью дуг, определяемой следующим образом:

$$a_{(i,k)} = c_{ik}, \quad i = 1, \dots, m, \quad k = 1, \dots, n;$$

$$a_{(s,i)} = a_i, \quad i = 1, \dots, m;$$

$$a_{(k,t)} = b_k, \quad k = 1, \dots, n.$$

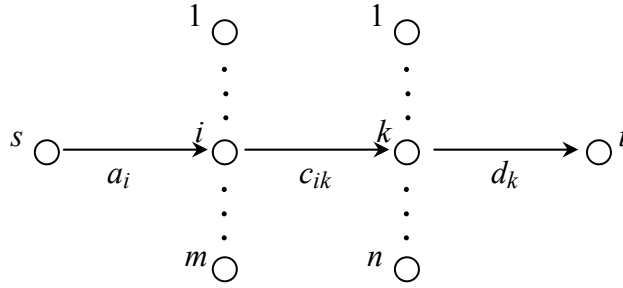


Рис. 4.2.2.

Пусть (V_1, V_2) — разрез для рассматриваемой сети G . Разрез (V_1, V_2) и допустимое решение $((y_i), (z_k))$ задачи $MINP_0$ полинома (4.2.2) назовем соответствующими, если

$$y_i = \begin{cases} 1, & \text{если } i \in V_1, \\ 0, & \text{если } i \in V_2; \end{cases}$$

$$z_k = \begin{cases} 0, & \text{если } k \in V_1, \\ 1, & \text{если } k \in V_2. \end{cases}$$

Покажем, что если разрез (V_1, V_2) и решение $((y_i), (z_k))$ соответствующие, то величина пропускной способности разреза равняется значению полинома $q_2(y_1, \dots, y_m, z_1, \dots, z_n)$ на данном решении. Действительно, можно написать

$$\gamma(V_1, V_2) = \sum_{i|i \in V_2} a_i + \sum_{i|i \in V_1} \sum_{k|k \in V_2} c_{ik} + \sum_{k|k \in V_1} b_k =$$

$$\sum_{i=1}^m a_i(1 - y_i) + \sum_{i=1}^m \sum_{k=1}^n c_{ik} y_i z_k + \sum_{k=1}^n d_k(1 - z_k) = q_2(y_1, \dots, y_m, z_1, \dots, z_n).$$

Таким образом, в силу признака оптимальности (лемма 1.1.1) построенное по минимальному разрезу соответствующее ему решение является оптимальным и требуемая сводимость показана.

Из доказательства теоремы следует, что для решения задачи $NIMP_0$ для полинома с разделяющимися переменными (4.2.2) существует эффективный алгоритм трудоемкости $O(m^3 + n^3)$.

Рассмотрим полином от булевых переменных вида

$$s_3(y_1, \dots, y_m) = \sum_{r=1}^R a_r (1 - \prod_{i \in \alpha_r} y_i) + \sum_{t=1}^T d_t (1 - \prod_{i \in \gamma_t} (1 - y_i)), \quad (4.2.3)$$

где $\alpha_r \subset \{1, \dots, m\}$, $r = 1, \dots, R$; $a_r \geq 0$, $r = 1, \dots, R$; $\gamma_t \subset \{1, \dots, m\}$, $t = 1, \dots, T$; $d_t > 0$, $t = 1, \dots, T$.

Теорема 4.2.3. Задача $MINP$ для полиномов (4.2.3) сводится к задаче $MINP_0$ для полиномов с разделяющимися переменными.

Доказательство. Полиному (4.2.3) поставим в соответствие следующий полином вида (4.2.2)

$$\sum_{r=1}^R a_r (1 - \tilde{y}_r) + \sum_{t=1}^T d_t (1 - \tilde{z}_t) + \sum_{r=1}^R \sum_{t=1}^T c_{rt} \tilde{y}_r \tilde{z}_t,$$

где

$$c_{rt} = \begin{cases} 0, & \text{если } \alpha_r \cap \gamma_t = \emptyset, \\ \Phi, & \text{если } \alpha_r \cap \gamma_t \neq \emptyset, \end{cases}$$

а Φ — некоторая величина такая, что $\Phi > \sum_{r=1}^R a_r$.

Пусть $((y_r), (z_t))$ — оптимальное решение задачи $MINP_0$ для рассматриваемого полинома с разделяющимися переменными. В силу выбора коэффициентов этого полинома для данного решения $((y_r), (z_t))$ при любых $r = 1, \dots, R$ и $t = 1, \dots, T$ справедливо равенство

$$\tilde{y}_r \cdot \tilde{z}_t = 0, \text{ если } \alpha_r \cap \gamma_t \neq \emptyset.$$

Рассматриваемому оптимальному решению поставим в соответствие допустимое решение (y_1, \dots, y_m) задачи $MINP$ для исходного полинома (4.2.3), определяемое следующим образом:

$$y_i = \begin{cases} 1, & \text{если } \tilde{y}_r = 1 \text{ для некоторого } r = 1, \dots, R, \text{ для которого } i \in \alpha_r; \\ 0, & \text{если } \tilde{z}_t = 1 \text{ для некоторого } t = 1, \dots, T, \text{ для которого } i \in \gamma_t; \\ 0, & \text{если } \tilde{y}_r = \tilde{z}_t = 0 \text{ для любых } r = 1, \dots, R, t = 1, \dots, T, \text{ для которых } i \in \alpha_r \cap \gamma_t. \end{cases}$$

Корректность этого определения следует из соотношения $\tilde{y}_r \cdot \tilde{z}_t = 0$, если $\alpha_r \cap \gamma_t \neq \emptyset$, выполняющегося для любых $r = 1, \dots, R$ и $t = 1, \dots, T$ в силу оптимальности решения $((\tilde{y}_r), (\tilde{z}_t))$.

Покажем, что таким образом определенное допустимое решение (y_1, \dots, y_m) является оптимальным. Для этого покажем, что выполняется признак оптимальности. Действительно, поскольку

$$\begin{aligned} \tilde{y}_r &\leq \prod_{i \in \alpha_r} y_i, \quad r = 1, \dots, R; \\ \tilde{z}_t &\leq \prod_{i \in \gamma_t} (1 - y_i), \quad t = 1, \dots, T, \end{aligned}$$

то значение полинома $s_3(y_1, \dots, y_m)$ на построенном решении (y_1, \dots, y_m) не больше оптимального значения рассматриваемого полинома с разделяющимися переменными. С другой стороны, пусть (y_1, \dots, y_m) — оптимальное решение задачи $MINP$ полинома $s_3(y_1, \dots, y_m)$. Построим по данному оптимальному решению решение $((\tilde{y}_r), (\tilde{z}_t))$ следующим образом:

$$\begin{aligned} \tilde{y}_r &\leq \prod_{i \in \alpha_r} y_i, \quad r = 1, \dots, R; \\ \tilde{z}_t &\leq \prod_{i \in \gamma_t} (1 - y_i), \quad t = 1, \dots, T. \end{aligned}$$

Поскольку для построенного решения равенство

$$\tilde{y}_r \cdot \tilde{z}_t = 0, \text{ если } \alpha_r \cap \gamma_t \neq \emptyset$$

выполняется для любых $r = 1, \dots, R$ и $t = 1, \dots, T$, то это решение дает значение рассматриваемого полинома с разделяющимися переменными, равное оптимальному значению по-

линома $s_3(y_1, \dots, y_m)$. Отсюда следует, что найденное по оптимальному решению $((\tilde{y}_r), (\tilde{z}_t))$ задачи $MINP_0$ для построенного полинома с разделяющимися переменными решение (y_1, \dots, y_m) является оптимальным решением задачи $MINP$ для исходного полинома $s_3(y_1, \dots, y_m)$ и, тем самым, необходимая сводимость доказана.

Из доказанного с учетом теоремы 4.2.2 получаем, что задача $MINP$ для полиномов вида (4.2.3) сводится к задаче о минимальном разрезе и, следовательно, может быть эффективно решена алгоритмом с оценкой трудоемкости $O(R^3 + T^3)$.

Полиномы $s_3(y_1, \dots, y_m)$ образуют наиболее широкий класс полиномов от булевых переменных, для которых задача $MINP$ сводится к задаче о минимальном разрезе. С другой стороны, как несложно видеть, квадратичный полином с неположительными коэффициентами $q_1(y_1, \dots, y_m)$ является полиномом вида $s_3(y_1, \dots, y_m)$. Поэтому задача о минимальном разрезе сводится к задаче $MINP$ для полинома $s_3(y_1, \dots, y_m)$. Это означает, что задачи минимизации каждого из полиномов $q_1(y_1, \dots, y_m)$, $s_2(y_1, \dots, y_m)$ и $s_3(y_1, \dots, y_m)$ эквивалентны и каждая из них эквивалентна задаче о минимальном разрезе. В частности, имеет место следующая

Теорема 4.2.4. Задача $MINP$ для полиномов с разделяющимися переменными и задача о минимальном разрезе эквивалентны.

В заключение сделаем еще одно важное замечание. Легко видеть, что в семействе полиномов $s_3(y_1, \dots, y_m)$ содержатся полиномы, имеющие неположительные коэффициенты при нелинейных членах, которые в канонической форме записываются следующим образом:

$$a_0 + \sum_{r=1}^R a_r (1 - \prod_{i \in \alpha_r} y_i) + \sum_{i \in I'} b_i y_i,$$

где $a_r > 0$, $r = 1, \dots, R$; $b_i > 0$, $i \in I' \subset I$.

Задача $MINP$ для таких полиномов, согласно сказанному выше, сводится к задаче о минимальном разрезе и, следовательно, для ее решения существует эффективный алгоритм с оценкой трудоемкости $O(R^3 + m^3)$. Заметим, что полином с неположительными коэффициентами при нелинейных членах при изменении знаков коэффициентов на противоположные становится полиномом вида $p_0(y_1, \dots, y_m)$. Следовательно, задача максимизации полиномов с неотрицательными коэффициентами при нелинейных членах может быть эффективно решена алгоритмом с оценкой трудоемкости $O(m^3 + S^3)$. Таким образом, полином $p_0(y_1, \dots, y_m)$ обладает тем свойством, что задача его минимизации является трудной, а задача максимизации — относительно легкой. Эти свойства полинома $p_0(y_1, \dots, y_m)$ аналогичны свойству вогнутых функций, задача минимизации которых, как известно, является сложной, а задача максимизации — относительно простой.

3 Эквивалентные формулировки задачи минимизации полиномов от булевых переменных

В главе 1 установлена тесная взаимосвязь между задачей $MINP_0$, задачей FL и задачей $MINF_0$. Эта взаимосвязь заключается не только в том, что данные задачи эквивалентны, но и в том, что эти задачи имеют одни и те же исходные данные, что позволяет говорить не о трех различных задачах, а о разных формах записи с использованием разных переменных одной и той же задачи. Помимо названных задач в главе 1 рассматривается задача SC , которая также эквивалентна задаче FL . Однако эта задача не может рассматриваться как еще одна форма записи задачи FL и в некотором смысле является менее сложной, чем задача FL .

В настоящем параграфе исследуются задачи, обобщающие указанные выше четыре задачи. К ним относится задача минимизации полиномов от булевых переменных, которая уже рассмотрена в предыдущем параграфе, а также так называемая двухстадийная задача размещения, обобщенная задача выбора множества строк пары матриц и обобщенная задача о покрытии множества системой подмножеств. Эти четыре задачи обладают свойствами, аналогичными свойствам задач $MINP_0$, FL , $MINF_0$ и SC . Во-первых, они эквивалентны, во-вторых, первые три из них можно рассматривать как разные формы записи одной и той же задачи и, наконец, в-третьих, четвертую задачу в некотором смысле можно считать менее сложной, чем первые три.

Из указанных четырех задач, как уже отмечалось ранее, только двухстадийная задача размещения средств обслуживания является достаточно известной и исследованной. Это связано с многочисленными ее приложениями в различных областях от размещения складских систем до выбора системы изделий, в составе которых используются унифицированные составные части или узлы.

3.1. Обобщенная задача выбора множества строк пары матриц

Обозначим через I множество $\{1, \dots, m\}$, а через J' и J соответственно множества $\{1, \dots, n'\}$ и $\{1, \dots, n\}$. Пусть $F = (f_{il})$ и $C = (c_{ij})$ — матрицы размера соответственно $m \times n'$ и $m \times n$. Обобщенной задачей выбора множества строк пары матриц (F, C) назовем задачу отыскания множества $X \subset I$, доставляющего минимум функции

$$f(X) = \sum_{l \in J'} \max_{i \in X} f_{il} + \sum_{j \in J} \min_{i \in X} c_{ij};$$

то есть задачу

$$\min_X \{f(X) \mid X \subset I\}.$$

При этом считаем, что если $X = \emptyset$, то

$$\begin{aligned} \max_{i \in X} f_{il} &= \min_{i \in I} f_{il}, \quad l \in J'; \\ \min_{i \in X} c_{ij} &= \max_{i \in I} c_{ij}, \quad j \in J; \end{aligned}$$

Такую задачу будем обозначать $MINF$.

Понятно, что если F — диагональная матрица, то функция $F(X)$ превращается в функцию $f_0(X)$, рассмотренную в главе 1. Поэтому сформулированная задача $MINF$ является обобщением задачи $MINF_0$, также как задача $MINP$ является обобщением задачи $MINP_0$.

Нашей целью будет сейчас доказательство того, что задача $MINF$ и задача $MINP$ эквивалентны. Для этого по аналогии со случаем задач $MINF_0$ и $MINP_0$ нам потребуется понятие пары характеристических матриц для пары матриц (F, C) .

Рассмотрим пару матриц (F, C) и для каждого столбца $l \in J'$ матрицы F и каждого столбца $j \in J$ матрицы C построим перестановки (i_1^l, \dots, i_m^l) и (i_1^j, \dots, i_m^j) множества I такие, что

$$f_{i_1^l l} \leq f_{i_2^l l} \leq \dots \leq f_{i_m^l l},$$

$$c_{i_1^j j} \leq c_{i_2^j j} \leq \dots \leq c_{i_m^j j},$$

и определим неотрицательные коэффициенты следующими равенствами:

$$\Delta f_k^l = f_{i_{k+1}^l l} - f_{i_k^l l}, \quad k = 1, \dots, m-1;$$

$$\Delta c_k^j = c_{i_{k+1}^j j} - c_{i_k^j j}, \quad k = 1, \dots, m-1.$$

Указанные перестановки назовем *порожденными* соответственно l -м столбцом матрицы F и j -м столбцом матрицы C , а введенные коэффициенты — *коэффициентами роста* элементов соответственно l -го столбца матрицы F и j -го столбца матрицы C .

Для матрицы F рассмотрим множество $\alpha \subset I$, $|\alpha| = k$, $1 \leq k \leq m-1$, такое, что множество

$$J'(\alpha) = \{l \in J' \mid \{i_{m-k+1}^l, \dots, i_m^l\} = \alpha, \Delta f_{m-k}^l > 0\}$$

не пусто. Множество α назовем *характеристическим множеством* длины k матрицы F , а величину $a(\alpha) = \sum_{l \in J'(\alpha)} \Delta f_{m-k}^l$ — *коэффициентом* или *весом* характеристического множества α матрицы F .

Аналогично, для матрицы C рассмотрим множество $\beta \subset I$, $|\beta| = k$, $1 \leq k \leq m-1$, такое, что множество

$$J(\beta) = \{j \in J \mid \{i_1^j, \dots, i_k^j\} = \beta, \Delta c_k^j > 0\}$$

не пусто. Множество β назовем *характеристическим множеством* длины k матрицы C , а величину $b(\beta) = \sum_{j \in J(\beta)} \Delta c_k^j$ — *коэффициентом* или *весом* характеристического множества β матрицы C .

Пусть $\alpha_1, \dots, \alpha_R$ — характеристические множества матрицы F с весами равными a_1, \dots, a_R . Матрицу $G = (g_{ij})$ ($i \in I$, $r = 1, \dots, R$) с $(0,1)$ -элементами

$$g_{ir} = \begin{cases} 0, & \text{если } i \in \alpha_r, \\ 1 & \text{иначе} \end{cases}$$

назовем *характеристической* для матрицы F , а вес a_r характеристического множества α_r назовем *весом* r -го столбца матрицы G . При этом для определенности будем считать, что столбцы матрицы G лексикографически упорядочены по возрастанию. Определим также матрицу $\tilde{F} = (a_r(1 - g_{ir}))(i \in I, r = 1, \dots, R)$, которую назовем *канонической формой* матрицы F .

Аналогично, пусть β_1, \dots, β_S — все характеристические множества матрицы C с весами равными b_1, \dots, b_S . Матрицу $H = (h_{is}) (i \in I, s = 1, \dots, S)$ с $(0,1)$ -элементами

$$h_{is} = \begin{cases} 0, & \text{если } i \in \beta_s, \\ 1 & \text{иначе} \end{cases}$$

назовем *характеристической* для матрицы C , а вес b_s характеристического множества β_s — *весом* s -го столбца матрицы H . Столбцы матрицы H считаем лексикографически упорядоченными по возрастанию. Матрицу $\tilde{C} = (b_s h_{is})(i \in I, s = 1, \dots, S)$ назовем *канонической формой* матрицы C .

С использованием характеристических матриц G и H определим пару (G', H') характеристических матриц для исходной пары (F, C) , а также каноническую форму (\tilde{F}', \tilde{C}') исходной пары (F, C) . Для этого исключим из пары (G, H) одинаковые столбцы аналогично приведению подобных членов в полиноме.

Пусть r -й столбец матрицы G и s -й столбец матрицы H совпадают. Если при этом $a_r = b_s$, то исключим из матриц G и H оба эти столбца. Если $a_r > b_s$, то исключим из матрицы H s -й столбец и одновременно пересчитаем коэффициент a_r , положив $a'_r = a_r - b_s$. Если $a_r < b_s$, то исключим из матрицы G r -й столбец и положим $b'_s = b_s - a_r$. Пусть в результате указанных действий построены матрица $G' = (g'_{ir})(i \in I, r = 1, \dots, R')$, $R' \leq R$, и матрица $H' = (h'_{is})(i \in I, s = 1, \dots, S')$, $S' \leq S$, не имеющие одинаковых столбцов, и пусть $a'_1, \dots, a'_{R'}$ и $b'_1, \dots, b'_{S'}$ — веса столбцов этих матриц. Пару матриц (G', H') назовем *парой характеристических матриц* для исходной пары (F, C) , а пару матриц (\tilde{F}', \tilde{C}') , где $\tilde{F}' = (a'_r(1 - g'_{ir}))(i \in I, r = 1, \dots, R')$ и $\tilde{C}' = (b'_s h'_{is})(i \in I, s = 1, \dots, S')$ — *канонической формой* пары матриц (F, C) .

В качестве примера построим для следующей пары матриц (F, C) :

$$F = \begin{pmatrix} 2 & 10 & 1 \\ 4 & 1 & 10 \\ 4 & 10 & 4 \\ 10 & 2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 6 & 3 & 3 \\ 4 & 2 & 4 \\ 3 & 3 & 1 \\ 1 & 6 & 2 \end{pmatrix}$$

характеристические матрицы G и H , канонические формы \tilde{F} и \tilde{C} для матриц F и C , пару характеристических матриц (G', H') и каноническую форму (\tilde{F}', \tilde{C}') для пары (F, C) .

Характеристическими множествами 1-го столбца матрицы F являются множества $\{4\}$ и $\{2, 3, 4\}$ с весами равными соответственно 6 и 2.

Характеристическими множествами 2-го столбца будут множества $\{1,3\}$ и $\{1,3,4\}$ с весами 8 и 1.

Наконец, характеристическими множествами 3-го столбца будут множества $\{2\}$, $\{2,3\}$ и $\{2,3,4\}$ с весами равными 6, 2 и 1.

Характеристическими множествами матрицы F в целом будут следующие 6 множеств:

$$\alpha_1 = \{4\}, \alpha_2 = \{2\}, \alpha_3 = \{2,3\}, \alpha_4 = \{2,3,4\}, \alpha_5 = \{1,3\}, \alpha_6 = \{1,3,4\}$$

со значениями весов равными

$$a_1 = 6, a_2 = 6, a_3 = 2, a_4 = 3, a_5 = 8, a_6 = 1.$$

Характеристическими множествами матрицы C и их веса вычислены в главе 1 и равняются:

$$\beta_1 = \{4\}, \beta_2 = \{3\}, \beta_3 = \{3,4\}, \beta_4 = \{2\}, \beta_5 = \{2,3,4\}, \beta_6 = \{1,3,4\}, \beta_7 = \{1,2,3\}$$

$$b_1 = 2, b_2 = 1, b_3 = 2, b_4 = 1, b_5 = 2, b_6 = 1, b_7 = 3.$$

Поэтому характеристическими матрицами G и H для матриц F и C будут (0,1)-матрицы

$$G = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix},$$

а каноническими формами \tilde{F} и \tilde{C} матриц F и C — матрицы

$$\tilde{F} = \begin{pmatrix} 1 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 2 & 6 & 0 \\ 1 & 8 & 3 & 2 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 & 6 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} 0 & 0 & 2 & 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 \\ 3 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

Пару характеристических матриц (G', H') для пары матриц (F, C) будут образовывать (0,1)-матрицы

$$G' = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad H' = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

а каноническую форму (\tilde{F}', \tilde{C}') пары (F, C) — матрицы

$$\tilde{F}' = \begin{pmatrix} 8 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 5 & 0 \\ 8 & 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 4 \end{pmatrix}, \quad \tilde{C}' = \begin{pmatrix} 0 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

Справедлива следующая лемма, дополняющая лемму 1.3.1 и позволяющая переписать задачу $MINF$ в виде задачи $MINP$.

Лемма 4.3.1. Пусть (i_1^l, \dots, i_m^l) и (i_1^j, \dots, i_m^j) — перестановки, порожденные соответственно l -м столбцом матрицы $F = (f_{il})$ и j -м столбцом матрицы $C = (c_{ij})$, а Δf_k^l , $k = 1, \dots, m-1$, и Δc_k^j , $k = 1, \dots, m-1$, коэффициенты роста элементов соответственно l -го столбца матрицы F и j -го столбца матрицы C . Тогда для любого $(0,1)$ -вектора (y_1, \dots, y_m) имеют место равенства

$$\begin{aligned} \max_{i|y_i=0} f_{il} &= - \sum_{k=1}^{m-1} \Delta f_k^l y_{i_{k+1}^l} \cdot \dots \cdot y_{i_m^l} + f_{i_m^l}^l, \\ \min_{i|y_i=0} c_{ij} &= c_{i_1^j} + \sum_{k=1}^{m-1} \Delta c_k^j y_{i_1^j} \cdot \dots \cdot y_{i_k^j}. \end{aligned}$$

Доказательство. Покажем выполнение первого равенства. Справедливость второго установлена при доказательстве леммы 1.3.1.

Если (y_1, \dots, y_m) — единичный вектор, то равенство, очевидно, выполняется. Пусть p — наибольший номер k , $1 \leq k \leq m$, для которого $y_{i_k^l} = 0$. Тогда, с одной стороны, имеем

$$\max_{i|y_i=0} f_{il} = f_{i_p^l}^l,$$

а с другой стороны, справедливы равенства

$$\begin{aligned} & - \sum_{k=1}^{m-1} \Delta f_k^l y_{i_{k+1}^l} \cdot \dots \cdot y_{i_m^l} + f_{i_m^l}^l = - \sum_{k=p}^{m-1} \Delta f_k^l + f_{i_m^l}^l = \\ & = -(f_{i_{p+1}^l}^l - f_{i_p^l}^l) - \dots - (f_{i_m^l}^l - f_{i_{m-1}^l}^l) + f_{i_m^l}^l = f_{i_p^l}^l. \end{aligned}$$

Это доказывает требуемое.

Теорема 4.3.1. Задача $MINP$ и задача $MINF$ эквивалентны.

Доказательство. Используя равенства из леммы 4.3.1, паре матриц (F, C) поставим в соответствие полином $p(y_1, \dots, y_m)$ от булевых переменных, определяемый следующим образом:

$$\begin{aligned} p(y_1, \dots, y_m) &= \sum_{l \in J'} \max_{i|y_i=0} f_{il} + \sum_{l \in J} \min_{i|y_i=0} c_{ij} = \\ &= \sum_{l \in J'} \left(\sum_{k=1}^{m-1} \Delta f_k^l y_{i_{k+1}^l} \cdot \dots \cdot y_{i_m^l} + f_{i_m^l}^l \right) + \sum_{j \in J} \left(c_{i_1^j} + \sum_{k=1}^{m-1} \Delta c_k^j y_{i_1^j} \cdot \dots \cdot y_{i_k^j} \right) = \\ &= \sum_{l \in J'} \left(f_{i_1^l}^l + \sum_{k=1}^{m-1} \Delta f_k^l - \sum_{k=1}^{m-1} \Delta f_k^l y_{i_{k+1}^l} \cdot \dots \cdot y_{i_m^l} \right) + \sum_{j \in J} \left(c_{i_1^j} + \sum_{k=1}^{m-1} \Delta c_k^j y_{i_1^j} \cdot \dots \cdot y_{i_k^j} \right) = \\ &= \sum_{l \in J'} f_{i_1^l}^l + \sum_{j \in J} c_{i_1^j} + \sum_{l \in J'} \sum_{k=1}^{m-1} \Delta f_k^l (1 - y_{i_{k+1}^l} \cdot \dots \cdot y_{i_m^l}) + \sum_{j \in J} \sum_{k=1}^{m-1} \Delta c_k^j y_{i_1^j} \cdot \dots \cdot y_{i_k^j}. \end{aligned}$$

Используя обозначения α_r , $r = 1, \dots, R$ и β_s , $s = 1, \dots, S$, для характеристических множеств соответственно матрицы F и матрицы C , а также обозначения a_r , $r = 1, \dots, R$ и b_s ,

$s=1, \dots, S$, для весов характеристических множеств соответственно матрицы F и матрицы C , перепишем полученный полином $p(y_1, \dots, y_m)$ следующим образом:

$$p(y_1, \dots, y_m) = \sum_{l \in J'} f_{i_1^l l} + \sum_{j \in J} c_{i_1^j j} + \sum_{r=1}^R a_r (1 - \prod_{i \in \alpha_r} y_i) + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i.$$

Полином $p(y_1, \dots, y_m)$, построенный указанным выше способом по паре матриц (F, C) , назовем *соответствующим* паре матриц (F, C) . Заметим, что исходными данными для этого полинома $p(y_1, \dots, y_m)$ является пара матриц (\tilde{F}, \tilde{C}) , где \tilde{F} — каноническая форма матрицы F , а \tilde{C} — каноническая форма матрицы C . Отметим также, что в силу построения полинома $p(y_1, \dots, y_m)$, соответствующего паре матриц (F, C) , решение (y_1, \dots, y_m) является оптимальным решением задачи *MINP* для полинома $p(y_1, \dots, y_m)$ тогда и только тогда, когда множество $X = \{i \in I \mid y_i = 0\}$ является оптимальным решением исходной задачи *MINF* с парой матриц (F, C) . Отсюда вытекает, что задача *MINF* сводится к задаче *MINP*.

Покажем обратную сводимость. Пусть задан полином $p(y_1, \dots, y_m)$ с исходными данными (\tilde{F}, \tilde{C}) , где $\tilde{F} = (a_r (1 - g_{ir})) (i \in I, r = 1, \dots, R)$, $\tilde{C} = (b_s h_{is}) (i \in I, s = 1, \dots, S)$. Рассмотрим задачу *MINF* с парой (\tilde{F}, \tilde{C}) . Полином, соответствующий паре (\tilde{F}, \tilde{C}) , может отличаться от исходного полинома $p(y_1, \dots, y_m)$ только на константу. Поэтому если X — оптимальное решение задачи *MINF* с парой (\tilde{F}, \tilde{C}) , то вектор (y_1, \dots, y_m) , где

$$y_i = \begin{cases} 0, & \text{если } i \in X, \\ 1 & \text{иначе} \end{cases}$$

является оптимальным решением исходной задачи *MINP*. Это доказывает требуемую сводимость и завершает доказательство теоремы.

Из доказательств теоремы ясно, как устроен полином $p(y_1, \dots, y_m)$, соответствующий паре матриц (F, C) . Член полинома $a_r (1 - \prod_{i \in \alpha_r} y_i)$ порождается характеристическим

множеством α_r матрицы F и при этом коэффициент a_r равняется весу характеристического множества α_r . Аналогично, член полинома $b_s \prod_{i \in \beta_s} y_i$ порождается характеристическим

множеством β_s матрицы C , а коэффициент b_s совпадает с весом характеристического множества β_s . Наконец, свободный член a_0 есть сумма наименьших элементов столбцов матриц F и C .

Отсюда понятно, какими свойствами должны обладать пары матриц, чтобы соответствующие им полиномы различались не более чем на константу. Пусть пары матриц (F, C) и (F_1, C_1) таковы, что, во-первых, совокупности характеристических множеств у матриц F и F_1 , а также у матриц C и C_1 совпадают и, во-вторых, веса соответствующих множеств у матриц F и F_1 , а также у матриц C и C_1 равны. Тогда соответствующие парам (F, C) и (F_1, C_1) полиномы будут отличаться разве что на константу.

Пары матриц (F, C) и (F_1, C_1) будем называть *эквивалентными*, если соответствующие им полиномы отличаются только на константу.

Из сказанного выше ясно, что пары матриц (F, C) и (F_1, C_1) будут эквивалентными, если канонические формы \tilde{F} и \tilde{F}_1 матриц F и F_1 , а также канонические формы \tilde{C} и \tilde{C}_1 матриц C и C_1 совпадают. Обратно, как легко понять, неверно и эквивалентными могут быть пары матриц, этим свойством не обладающие. Необходимым достаточным условием эквивалентности пары матриц (F, C) и (F_1, C_1) является совпадение их канонических форм (\tilde{F}', \tilde{C}') и $(\tilde{F}_1', \tilde{C}_1')$.

Таким образом, всякий полином

$$f(y_1, \dots, y_m) = a_0 + \sum_{r=1}^R a_r (1 - \prod_{i|g_{ir}=0} y_i) + \sum_{s=1}^S b_s \prod_{i|h_{is}=0} y_i$$

с парой характеристических матриц (G, H) , весами столбцов a_1, \dots, a_R матрицы G и b_1, \dots, b_S матрицы H и свободным членом a_0 , принимающим произвольные значения, является соответствующим для некоторого класса эквивалентных пар матриц. Характеристическая пара (G', H') у всякой пары матриц (F, C) из этого класса совпадает с парой характеристических матриц (G, H) полинома, а веса соответствующих столбцов у матриц G и G' , а также у матриц H и H' равны. В этом классе можно выделить пару матриц (\tilde{F}', \tilde{C}') , $\tilde{F}' = (a_r(1 - g_{ir}))$, $\tilde{C}' = (b_s h_{is})$, являющуюся канонической формой любой пары матриц (F, C) из данного класса эквивалентности. Эту пару (\tilde{F}', \tilde{C}') естественно считать представителем класса эквивалентных пар матриц. Матрицы \tilde{F}' и \tilde{C}' обладают тем свойством, что в каждом столбце этих матриц наименьшие элементы равны нулю, а ненулевые элементы каждого столбца одинаковые.

Поскольку эквивалентным парам матриц (F, C) и (F_1, C_1) соответствуют полиномы, отличающиеся разве что на константу, то согласно теореме 4.3.1 оптимальные решения задач $MINF$ с парой (F, C) и с парой (F_1, C_1) совпадают. Таким образом, класс эквивалентных пар матриц обладает тем свойством, что задачи $MINF$ с такими парами матриц имеют одинаковые решения. Следовательно, поиск оптимального решения задачи $MINF$ с исходной парой (F, C) может быть сведен к решению задачи $MINF$ с некоторой парой матриц, эквивалентной исходной, например, с парой (\tilde{F}', \tilde{C}') .

3.2. Двухстадийная неограниченная задача размещения

Рассмотренная в двух первых главах неограниченная задача размещения предприятий (средств обслуживания) моделирует, как известно, ситуацию удовлетворения заданного спроса некоторого множества потребителей в однородном продукте. Этот продукт производится предприятиями, которые могут быть открыты в заданных местах их размещения. В двухстадийной задаче размещения рассматривается та же ситуация удовлетворения спроса потребителей в однородном продукте, но, в отличие от первой задачи, считается, что для открытия предприятия, удовлетворяющего спрос и называемого предпри-

ятием верхнего уровня, необходимо, чтобы были открыты в заданных местах размещения некоторое число связанных с ним вспомогательных предприятий, называемых предприятиями нижнего уровня. Таким образом, в двухстадийной задаче размещения процесс производства потребляемого продукта состоит из двух стадий (этапов). На предприятиях нижнего уровня производится сырье для предприятий верхнего уровня, которые перерабатывают это сырье в продукцию, направляемую потребителям. Следует подчеркнуть, что для каждого открываемого предприятия верхнего уровня число связанных с ним предприятий нижнего уровня и места их размещения заданы. При этом каждое открываемое предприятие нижнего уровня может быть связано ни с одним, а с несколькими предприятиями верхнего уровня.

В двухстадийной задаче размещения требуется определить множество открываемых предприятий верхнего уровня, соответствующее множество открываемых предприятий нижнего уровня и прикрепить каждого потребителя к предприятию верхнего уровня так, чтобы суммарные затраты, включающие фиксированные затраты на открытие предприятий верхнего и нижнего уровней и затраты на удовлетворение спроса потребителей, были наименьшими.

Чтобы сформулировать двухстадийную задачу размещения введем следующие обозначения.

$J = \{1, \dots, n\}$ — множество потребителей с заданным спросом, пронумерованных индексом j ;

$I = \{1, \dots, m\}$ — множество предприятий (мест размещения предприятий) верхнего уровня, пронумерованных индексом i ;

$\{1, \dots, R\}$ — множество предприятий (мест размещения предприятий) нижнего уровня, пронумерованных индексом r ;

g_{ir} — величина, показывающая связаны или нет предприятие i верхнего уровня и предприятие r нижнего уровня; $g_{ir}=1$, если предприятия i и r связаны, $g_{ir}=0$, если нет;

f_r — величина затрат на открытие предприятия r нижнего уровня;

d_i — величина затрат на открытие предприятия i верхнего уровня;

c_{ij} — затрат на удовлетворение предприятием i спроса потребителя j ;

Введем также следующие переменные.

$y_i \in \{0, 1\}$ — переменная, показывающая открывается или нет предприятие i верхнего уровня; $y_i = 1$, если открывается, и $y_i = 0$, если нет;

$z_r \in \{0, 1\}$ — переменная, показывающая открывается или нет предприятие r нижнего уровня; $z_r = 1$, если открывается, и $z_r = 0$, если нет;

$x_{ij} \in \{0, 1\}$ — переменная, показывающая используется или нет предприятие i для удовлетворения спроса потребителя j ; $x_{ij} = 1$, если используется, и $x_{ij} = 0$, если нет.

С использованием указанных обозначений двухстадийная неограниченная задача размещения записывается в виде задачи целочисленного программирования следующим образом:

$$\min_{(x_i), (x_{ij}), (z_r)} \left\{ \sum_{r=1}^R f_r z_r + \sum_{i \in I} d_i x_i + \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \right\} \quad (4.3.1)$$

$$\sum_{i \in I} x_{ij} = 1, \quad j \in J; \quad (4.3.2)$$

$$x_i \geq x_{ij}, \quad i \in I, j \in J; \quad (4.3.3)$$

$$z_r \geq g_{ir} x_i, \quad i \in I, r = 1, \dots, R; \quad (4.3.4)$$

$$x_i, x_{ij}, z_r \in \{0, 1\}, \quad i \in I, j \in J, r = 1, \dots, R; \quad (4.3.5)$$

Эту задачу далее будем обозначать *TFL*. Целевая функция (4.3.1) данной задачи выражает величину суммарных затрат на открытие предприятий и удовлетворения спроса потребителей. Ограничение (4.3.2) означает, что спрос всех потребителей должен быть удовлетворен. Неравенство (4.3.3) показывает, что предприятие верхнего уровня не может удовлетворять спрос потребителей, если оно не открыто, а неравенство (4.3.4) означает, что если открыто предприятие верхнего уровня, то все связанные с ним предприятия нижнего уровня тоже должны быть открыты.

Заметим, что сформулированной задаче условия (4.3.4) могут быть заменены на условия

$$z_r \geq \sum_{i \in I} g_{ir} x_{ij} = 1, \quad r = 1, \dots, R, j \in J; \quad (4.3.4')$$

без изменения множества оптимальных решений задачи.

Действительно, если $((z_r), (x_i), (x_{ij}))$ — оптимальное решение задачи *TFL* с ограничением (4.3.4), то в силу условий (4.3.2) и (4.3.3) выполняются неравенства (4.3.4').

Обратно, пусть $((z_r), (x_i), (x_{ij}))$ — оптимальное решение задачи с условием (4.3.4'). Поскольку это оптимальное решение, то

$$x_i = \max_{j \in J} x_{ij}, \quad i \in I.$$

С учетом этого для любого $i_0 \in I$ получаем

$$z_r \geq \sum_{i \in I} g_{ir} \max_{j \in J} x_{ij} \geq g_{i_0 r} \max_{j \in J} x_{i_0 j} = g_{i_0 r} x_{i_0}$$

и, следовательно, неравенства (4.3.4) выполняются.

Вместе с задачей *TFL* рассмотрим ее частный случай, когда $d_i = 0, i \in I$. Содержательно это означает отсутствие затрат на открытие собственно предприятий верхнего уровня. В этом случае переменные $x_i, i \in I$, могут быть исключены из задачи (4.3.1) – (4.3.5) а сама задача переписана следующим образом:

$$\min_{(x_{ij}), (z_r)} \left\{ \sum_{r=1}^R f_r z_r + \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \right\} \quad (4.3.6)$$

$$\sum_{i \in I} x_{ij} = 1, \quad j \in J; \quad (4.3.7)$$

$$z_r \geq \sum_{i \in I} g_{ir} x_{ij} = 1, \quad j \in J, r = 1, \dots, R; \quad (4.3.8)$$

$$x_{ij}, z_r \in \{0, 1\}, \quad i \in I, j \in J, r = 1, \dots, R. \quad (4.3.9)$$

Эту задачу далее будем обозначать через T_0FL .

Покажем, что в действительности задача T_0FL не является более простой по сравнению с общим случаем. Как следует из приводимой ниже леммы, эти задачи являются эквивалентными и, следовательно, при построении алгоритмов решения задачи TFL можно сосредоточиться на исследовании задачи T_0FL .

Лемма 4.3.1. Задача TFL сводится к задаче T_0FL .

Доказательство. Задаче TFL с множеством предприятий верхнего уровня $I = \{1, \dots, m\}$ и множеством предприятий нижнего уровня $\{1, \dots, R\}$ поставим в соответствие задачу T_0FL с множеством предприятий верхнего уровня $I = \{1, \dots, m\}$ и множеством предприятий нижнего уровня $\{1, \dots, R, R+1, \dots, R+m\}$. В этой задаче к реальным предприятиям нижнего уровня добавлено m фиктивных, таких, что каждое из этих предприятий связано с одним и только с одним соответствующим предприятием верхнего уровня. Положим для рассматриваемой задачи T_0FL

$$f'_r = \begin{cases} f_r, & \text{если } r \leq R, \\ d_i, & \text{если } r = R+i, \end{cases}$$

$$r = 1, \dots, R, R+1, \dots, R+m;$$

$$g'_{ir} = \begin{cases} g_{ir}, & \text{если } r \leq R, \\ 1, & \text{если } r = R+i, \\ 0, & \text{если } r > R \text{ и } r \neq R+i, \end{cases}$$

$$i \in I, r = 1, \dots, R, R+1, \dots, R+m.$$

Несложно увидеть, что эта задача после замены переменных

$$z_r = z'_r, \quad r = 1, \dots, R,$$

$$x_i = z'_{R+i}, \quad i \in I,$$

превращается в исходную задачу TFL . Поэтому если $((z'_r), (x_{ij}))$ — оптимальное решение рассматриваемой задачи T_0FL , то $((z_r), (x_i), (x_{ij}))$ — оптимальное решение исходной задачи TFL , и требуемая сводимость доказана.

Теорема 4.3.1. Задача $MINF$ и задача T_0FL эквивалентны.

Доказательство. Задаче T_0FL поставим в соответствие задачу $MINF$ с парой матриц (F, C) , где $F = (f_r \cdot g_{ir})$ ($i \in I, r = 1, \dots, R$), $C = (c_{ij})$ ($i \in I, j \in J$). Решению X последней задачи поставим в соответствие решение $((z_r), (x_{ij}))$ исходной задачи T_0FL , построенное следующим образом. Для всякого $j \in J$ обозначим через $i(j)$ некоторый номер $i_0 \in X$, для которого $c_{i_0 j} = \min_{i \in X} c_{ij}$. Положим

$$z_r = \begin{cases} 1, & \text{если } \max_{i \in X} g_{ir} = 1, \\ 0 & \text{иначе,} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{если } i = i(j), \\ 0 & \text{иначе.} \end{cases}$$

Заметим, что решение $((z_r), (x_{ij}))$ есть допустимое решение задачи T_0FL . Действительно, равенство (4.3.7), очевидно, выполняется. Неравенство (4.3.8) также выполняется, поскольку, если $z_r = 0$, то $g_{ir} = 0$ для всякого $i \in X$. Кроме того, имеем $x_{ij} = 0$ для любого $i \notin X$. Отсюда получаем $\sum_{i \in I} g_{ir} x_{ij} = 0$.

Допустимые решения X и $((z_r), (x_{ij}))$, связанные указанными выше соотношениями, назовем соответствующими. Заметим, что поскольку для соответствующих решений справедливы равенства

$$\begin{aligned} \max_{i \in X} g_{ir} f_r &= f_r \cdot z_r, \quad r = 1, \dots, R; \\ \min_{i \in X} c_{ij} &= c_{i(j)j} = \sum_{i \in I} c_{ij} x_{ij}, \quad j \in J, \end{aligned}$$

то на таких решениях значения целевых функций рассматриваемых задач равны.

Заметим также, что если $((z_r), (x_{ij}))$ — оптимальное решение задачи T_0FL , то легко строится множество $X \subset I$ такое, что решения X и $((z_r), (x_{ij}))$ являются соответствующими. Действительно, положим $X = \{i \in I \mid \max_{j \in J} x_{ij} = 1\}$. Поскольку $((z_r), (x_{ij}))$ — оптимальное решение, то

$z_r = 1$ тогда и только тогда, когда $\max_{j \in J} \sum_{i \in I} g_{ir} x_{ij} = 1$. Отсюда, с учетом равенств

$$\max_{j \in J} \sum_{i \in I} g_{ir} x_{ij} = \max_{j \in J} \max_{i \in I} g_{ir} x_{ij} = \max_{i \in I} g_{ir} \max_{j \in J} x_{ij} = \max_{i \in X} g_{ir},$$

получаем, что $z_r = 1$ тогда и только тогда, когда $\max_{i \in X} g_{ir} = 1$. Из оптимальности решения

получаем также, что если $x_{i_0 j} = 1$, то $c_{i_0 j} = \min_{i \in X} c_{ij}$. Таким образом, рассматриваемые решения являются соответствующими.

Из сказанного с учетом признака оптимальности (лемма 1.1.1) получаем, что если X — оптимальное решение рассмотренной задачи $MINF$, то построенное по нему решение $((z_r), (x_{ij}))$ задачи T_0FL является оптимальным и, следовательно, задача T_0FL сводится к задаче $MINF$.

Имеет место и обратная сводимость. Рассмотрим задачу $MINF$ с парой матриц (F, C) и задачу $MINF$ с парой матриц (\tilde{F}, C) , где $\tilde{F} = (a_r(1 - g_{ir}))(i \in I, r = 1, \dots, R)$ — каноническая форма матрицы F . Эти задачи, как известно, имеют одни и те же оптимальные решения. Последней задаче поставим в соответствие задачу T_0FL . Из первой части доказательства теоремы ясно, что если $((z_r), (x_{ij}))$ — оптимальное решение построенной задачи T_0FL , то соответствующее ему решение X будет оптимальным решением $MINF$ с парой (\tilde{F}, C) и, следовательно, задачи $MINF$ с исходной парой (F, C) . Таким образом, требуемая сводимость доказана и рассматриваемые задачи $MINF$ и T_0FL являются эквивалентными. Теорема доказана.

3.3. Обобщенная задача о покрытии множества системой подмножеств

Пусть $J = \{1, \dots, n\}$, $I = \{1, \dots, m\}$ — конечные множества и пусть $A = (a_{ij})$ и $G = (g_{ir})$ — $(0,1)$ -матрицы размера соответственно $m \times n$ и $m \times R$, а f_r , $r = 1, \dots, R$, — неотрицательные величины. Обобщенной задачей о покрытии, далее обозначаемой через GSC , назовем следующую задачу:

$$\begin{aligned} \min_{(x_i), (z_r)} \quad & \sum_{r=1}^R f_r z_r; \\ & \sum_{i \in I} a_{ij} x_i \geq 1, \quad j \in J; \\ & z_r \geq g_{ir} x_i, \quad i \in I, r = 1, \dots, R; \\ & x_i, z_r \in \{0, 1\}, \quad i \in I, r = 1, \dots, R. \end{aligned}$$

Эта задача отличается от рассмотренной в главе 1 задачи SC тем, что вес системы подмножеств, покрывающих множество J , определяется не как сумма весов подмножеств, входящих в систему, а более сложным образом через систему подмножеств множества I , задаваемую матрицей $G = (g_{ir})$.

Задача GSC , как следует из приводимой ниже теоремы, тесно связана с задачей $MINF$, а следовательно, и с задачей $MINP$.

Теорема 4.3.2. Задача $MINF$ и задача GSC эквивалентны.

Доказательство. Задаче GSC поставим в соответствие задачу $MINF$ с парой матриц (F, C) , $F = (f_r g_{ir})$, $C = (\Phi(1 - a_{ij}))$, где Φ — некоторая достаточно большая величина такая, что $\Phi > \sum_{r=1}^R f_r$. Пусть X — оптимальное решение рассматриваемой задачи $MINF$. Поставим

ему в соответствие решение $((x_i), (z_r))$ исходной задачи GSC , определив его следующим образом:

$$\begin{aligned} x_i &= \begin{cases} 1, & \text{если } i \in X, \\ 0 & \text{иначе,} \end{cases} \\ z_r &= \begin{cases} 1, & \text{если } \max_{i \in X} g_{ir} = 1, \\ 0 & \text{иначе.} \end{cases} \end{aligned}$$

Заметим, что поскольку X — оптимальное решение, то в силу выбора величины Φ имеем $X \cap \{i \in I \mid a_{ij} = 1\} \neq \emptyset$ для всякого $j \in J$. Следовательно, решение $((x_i), (z_r))$ — допустимое. По этой же причине справедливы равенства

$$\min_{i \in X} \Phi(1 - a_{ij}) = 0, \quad j \in J.$$

Заметим также, что

$$\max_{i \in X} f_r g_{ir} = f_r \max_{i \in X} g_{ir} = f_r z_r, \quad r = 1, \dots, R.$$

Отсюда получаем, что значения целевых функций задач $MINF$ и GSC на рассматриваемых решениях равны.

Кроме того, если $((x_i), (z_r))$ — оптимальное решение исходной задачи GSC , то

$$z_r = \max_{i|x_i=1} g_{ir}, \quad r = 1, \dots, R.$$

Поэтому для решения $X = \{i \in I \mid x_i = 1\}$ построенной задачи $MINF$, получаем, что значение целевой функции равняется значению целевой функции задачи GSC . Отсюда, в силу критерия оптимальности (лемма 1.1.1) следует, что задача GSC сводится к задаче $MINF$.

Имеет место и обратная сводимость. Рассмотрим задачу $MINF$ с парой матриц (F, C) и задачу $MINF$ с парой матриц (\tilde{F}, \tilde{C}) , где $\tilde{F} = (a_r(1 - g_{ir}))(i \in I, r = 1, \dots, R)$ — каноническая форма матрицы F , а $\tilde{C} = (b_s h_{is})(i \in I, s = 1, \dots, S)$ — каноническая форма матрицы C . Эти задачи имеют одни и те же оптимальные решения. Последней задаче поставим в соответствие задачу GSC вида

$$\begin{aligned} \min_{(x_i)(z_r)(w_s)} & \left\{ \sum_{r=1}^R a_r z_r + \sum_{s=1}^S b_s w_s \right\}; \\ \sum_{i \in I} (1 - h_{is}) x_i + w_s & \geq 1, \quad s = 1, \dots, S; \\ z_r & \geq (1 - g_{ir}) x_i, \quad i \in I, r = 1, \dots, R; \\ z_r, x_i, w_s & \in \{0, 1\}, \quad i \in I, r = 1, \dots, R, s = 1, \dots, S. \end{aligned}$$

Пусть $((x_i), (z_r), (w_s))$ — оптимальное решение рассмотренной задачи GSC . Поставим ему в соответствие решение $X = \{i \in I \mid x_i = 1\}$ задачи $MINF$ с парой (\tilde{F}, \tilde{C}) . В силу оптимальности решения $((x_i), (z_r), (w_s))$ имеют место равенства

$$\begin{aligned} z_r &= \max_{i \in I} (1 - g_{ir}) x_i = \max_{i|x_i=1} (1 - g_{ir}) = \max_{i \in X} (1 - g_{ir}), \\ w_s &= 1 - \max_{i \in I} (1 - h_{is}) x_i = 1 - \max_{i|x_i=1} (1 - h_{is}) = \min_{i \in X} h_{is}. \end{aligned}$$

Отсюда получаем, что значения целевых функций задачи $MINF$ с парой (\tilde{F}, \tilde{C}) и приведенной задачи GSC на рассматриваемых решениях равны. Чтобы убедиться в том, что построенное выше решение является оптимальным укажем, согласно принципу оптимальности (лемма 1.1.1), для оптимального решения задачи GSC допустимое решение $((x_i), (z_r), (w_s))$ задачи $MINF$ с парой (\tilde{F}, \tilde{C}) , для которого значение целевой функции совпадает с оптимальным значением целевой функции задачи GSC . Такое решение определяется следующим образом

$$\begin{aligned} x_i &= \begin{cases} 1, & \text{если } i \in X, \\ 0 & \text{иначе,} \end{cases} \\ z_r &= \max_{i \in X} (1 - g_{ir}); \\ w_s &= \min_{i \in X} h_{is}. \end{aligned}$$

Это решение является допустимым поскольку, поскольку

$$z_r = \max_{i \in X} (1 - g_{ir}) = \max_{i \in I} (1 - g_{ir}) x_i \geq (1 - g_{ir}) x_i,$$

$$w_s = \min_{i \in X} h_{is} = 1 - \max_{i \in I} (1 - h_{is})x_i \geq 1 - \sum_{i \in I} (1 - h_{is})x_i.$$

Кроме того, значение целевой функции на этом решении совпадает с оптимальным значением целевой функции задачи *GSC*. Следовательно, задача *MINF* сводится к задаче *GSC* и требуемая эквивалентность рассматриваемых задач доказана.

Таким образом, по аналогии с задачей *FL* и связанных с ней задач *MINF*₀, *MINP*₀ и *SC*, обнаружена тесная взаимосвязь между следующими четырьмя задачами: двухстадийной задачей размещения (*TFL*), задачей выбора набора строк пары матриц (*MINF*), задачей минимизации полиномов от булевых переменных (*MINP*) и обобщенной задачей о покрытии множества системой подмножеств (*GSC*). Эта взаимосвязь не ограничивается лишь тем, что данные задачи взаимно сводимы и построение для каждой задачи соответствующей ей другой задачи осуществляется достаточно просто. Так же как и в случае задачи *FL* и связанных с ней задач, эта взаимосвязь носит более тесный характер, по крайней мере, для трех первых задач. Исходные данные этих задач задаются двумя (0,1)-матрицами и весами столбцов этих матриц. Поэтому можно говорить не о трех разных задачах, а о разных формах записи одной и той же задачи. Это означает, в частности, что если для одной из них выделен некоторый эффективно разрешимый частный случай, то автоматически получены эффективно разрешимые частные случаи и других задач. Таким образом, эти три задачи эквивалентны не только в смысле поиска для них универсальных алгоритмов, но и в смысле поиска эффективно разрешимых частных случаев этих задач.

Несколько менее тесной является связь между указанными тремя задачами и задачей *GSC*. Решение задачи *GSC* с матрицами ограничений $A = (a_{ij})$ и $G = (g_{ir})$, как известно, может быть получено из решения задачи *MINF* с парой матриц (F, C) , имеющих в качестве характеристических соответственно матрицы $\bar{G} = (1 - g_{ir})$ и $\bar{H} = (1 - a_{ij})$. Отсюда следует, что если найдено некоторое полезное свойство характеристических матриц G и H , позволяющее эффективно решать задачу *MINF* с парой (F, C) , то тем самым выделен подкласс эффективно разрешимых задач *GSC*. Этот подкласс определяется матрицами ограничений A и G такими, что матрицы \bar{A} и \bar{G} обладают указанными выше свойствами. При этом отметим, что относительная мощность построенного подкласса эффективно разрешимых задач *GSC* такая же, что и выделенного подкласса эффективно разрешимых задач *MINF*₀. В этом смысле, задачу *GSC* можно считать не сложнее задачи *MINF*.

С другой стороны, решение задачи *MINF* с парой (F, C) , имеющих характеристические матрицы $G = (g_{ir})$ и $H = (h_{ij})$ соответственно, может быть получено из решения задачи

GSC с матрицами ограничений $\tilde{A} = \begin{pmatrix} \bar{H} \\ E \end{pmatrix}$ и $\tilde{G} = \begin{pmatrix} \bar{G} \\ 0 \end{pmatrix}$, где E — диагональная матрица, а 0 —

нулевая матрица. Поэтому, если для матриц ограничений A и G задачи *GSC* найдены свойства, позволяющие строить эффективные алгоритмы решения задачи, то это может не дать никакого результата для задачи *MINF*. Действительно, указанными полезными свойствами

должны обладать матрицы $\tilde{A} = \begin{pmatrix} \bar{H} \\ E \end{pmatrix}$ и $\tilde{G} = \begin{pmatrix} \bar{G} \\ 0 \end{pmatrix}$, чего может не быть в силу специального

вида этих матриц. Поэтому класс пар матриц (F, C) , имеющих характеристические матрицы G и H такие, что матрицы \tilde{A} и \tilde{G} задачи GSC обладают указанными свойствами, может оказаться вырожденным. В этом смысле можно констатировать, что задачи TFL , $MINF$ и $MINP$ более сложные, чем задача GSC .

4 Сложность задачи минимизации специальных классов полиномов от булевых переменных

Задача минимизации квадратичных полиномов с разделяющимися переменными и неотрицательными коэффициентами, как следует из сказанного, в начале настоящей главы, эквивалентна задаче о минимальном разрезе. Это означает, что класс квадратичных полиномов с разделяющимися переменными является одним из достаточно общих классов полиномов, для которого задача минимизации эффективно разрешима. К сожалению, попытки сколь-нибудь значительных расширений данного класса с использованием идеи разделения переменных, сохраняющих возможность эффективно решать задачу минимизации, не приводят к успеху.

В настоящем параграфе исследуется сложность задачи минимизации для трех классов полиномов, являющихся естественными расширениями класса квадратичных полиномов с разделяющимися переменными. Такие расширения образуют квадратичные полиномы от трех групп переменных, полиномы третьей степени от трех групп переменных и квадратичные полиномы с разделяющимися переменными, но с коэффициентами разных знаков. Для таких полиномов близких по строению к квадратичным полиномам с разделяющимися переменными, задача минимизации оказывается существенно более сложной и является NP-трудной. Общий метод доказательства соответствующих утверждений заключается в сведении к каждой из рассматриваемых задач задачи о вершинном покрытии для три раскрашиваемого графа. Доказательство того, что эта задача является NP-трудной вытекает из приводимой в начале параграфа вспомогательной леммы о задаче вершинного покрытия минимальной мощности для графа. В целом изложение данного параграфа следует работам [4, 6].

4.1 Задача о вершинном покрытии для графа

Опираясь на результаты главы 1 о сводимости задачи $MINP_0$ к задаче VC и учитывая то, что задача $VC1$ для произвольного графа является NP-трудной, покажем, что задача $VC1$ для три раскрашиваемого графа также является NP-трудной.

Пусть $G = (V, E)$ — неориентированный граф без петель. Произвольным образом выбрав направление ребер, сделаем данный граф G ориентированным и тем самым для всякого ребра $e \in E$ определим его начальную вершину $v_1(e)$ и конечную вершину $v_2(e)$.

Запишем задачу $VC1$ для графа G следующим образом

$$\min_{(x_v)} \sum_{v \in V} x_v; \quad (4.4.1)$$

$$x_{v_1(e)} + x_{v_2(e)} \geq 1, \quad e \in E; \quad (4.4.2)$$

$$x_v \in \{0, 1\}, \quad v \in V. \quad (4.4.3)$$

Покажем, что эта задача может быть переписана как задача $VC1$ для некоторого более просто упорядоченного графа G' , который оказывается три раскрашиваемым графом.

Лемма 4.4.1. Задача $VC1$ для графа G (4.4.1) – (4.4.3) сводится к задаче $VC1$ следующего вида

$$\min_{(x_v), (t_e^1), (t_e^2)} \left\{ \sum_{v \in V} x_v + \sum_{e \in E} (t_e^1 + t_e^2) \right\}; \quad (4.4.4)$$

$$x_{v_1(e)} + t_e^1 \geq 1, \quad e \in E; \quad (4.4.5)$$

$$x_{v_2(e)} + t_e^2 \geq 1, \quad e \in E; \quad (4.4.6)$$

$$t_e^1 + t_e^2 \geq 1, \quad e \in E; \quad (4.4.7)$$

$$x_v, t_e^1, t_e^2 \in \{0, 1\}, \quad v \in V, e \in E. \quad (4.4.8)$$

Доказательство. Согласно теореме 1.4.2 и следствию 2 из нее задача (4.4.1) – (4.4.3) сводится к задаче VC следующего вида:

$$\min_{(x_v), (t_e^1), (t_e^2)} \left\{ \sum_{v \in V} x_v + \sum_{e \in E} \Phi(t_e^1 + t_e^2) \right\}; \quad (4.4.9)$$

$$x_{v_1(e)} + t_e^1 \geq 1, \quad e \in E; \quad (4.4.10)$$

$$x_{v_2(e)} + t_e^2 \geq 1, \quad e \in E; \quad (4.4.11)$$

$$t_e^1 + t_e^2 \geq 1, \quad e \in E; \quad (4.4.12)$$

$$x_v, t_e^1, t_e^2 \in \{0, 1\}, \quad v \in V, e \in E, \quad (4.4.13)$$

где $\Phi > |I|$.

Утверждение леммы будет доказано, если будет показано, что задача (4.4.9) – (4.4.13) сводится к задаче (4.4.4) – (4.4.8), то есть что оптимальное решение задачи (4.4.4) – (4.4.8) может быть построено по оптимальному решению задачи (4.4.9) – (4.4.13).

Нетрудно видеть, что для любого оптимального решения задачи (4.4.9) – (4.4.13) неравенство (4.4.12) превращается в равенство. Покажем, что существует оптимальное решение задачи (4.4.4) – (4.4.8), для которого неравенство (4.4.7) также выполняется как равенство. Пусть $((x_v), (t_e^1), (t_e^2))$ — некоторое оптимальное решение задачи (4.4.4) – (4.4.8), для которого множество $\tilde{E} = \{e \in E \mid t_e^1 + t_e^2 > 1\}$ не пусто. Возьмем произвольный элемент $\tilde{e} \in \tilde{E}$ и преобразуем исходное оптимальное решение в решение $((\tilde{x}_v), (\tilde{t}_e^1), (\tilde{t}_e^2))$, определяемое следующим образом:

$$\tilde{x}_e = \begin{cases} 1, & \text{если } v = v_1(\tilde{e}), \\ x_e & \text{иначе,} \end{cases}$$

$$\tilde{t}_e^1 = \begin{cases} 0, & \text{если } e = \tilde{e}, \\ t_e^1 & \text{иначе,} \end{cases}$$

$$\tilde{t}_e^2 = t_e^2.$$

Понятно, что построенное решение есть допустимое решение. Более того, это оптимальное решение, поскольку значение целевой функции на этом решении не превосходит значения целевой функции на исходном оптимальном решении. Кроме того, для построенного оптимального решения множество \tilde{E} имеет, очевидно, мощность на единицу меньше, чем множество \tilde{E} для исходного оптимального решения. Поэтому, продолжив при необходимости описанный выше процесс преобразования оптимального решения задачи (4.4.4) – (4.4.8), получим оптимальное решение этой задачи, обладающее необходимым свойством.

Таким образом, получаем, что обе рассматриваемые задачи (4.4.4) – (4.4.8) и (4.4.9) – (4.4.13) эквивалентны одной и той же задаче, имеющей вид:

$$\min_{(x_v), (t_e^1), (t_e^2)_{v \in V}} \sum_{v \in V} x_v;$$

$$x_{v_1(e)} + t_e^1 \geq 1, \quad e \in E;$$

$$x_{v_2(e)} + t_e^2 \geq 1, \quad e \in E;$$

$$t_e^1 + t_e^2 = 1, \quad e \in E.$$

Следовательно, если $((x_v), (t_e^1), (t_e^2))$ — оптимальное решение задачи (4.4.4) – (4.4.8) такое, что для этого решения множество \tilde{E} пусто, то оно является оптимальным решением задачи (4.4.9) – (4.4.13), что доказывает требуемую сводимость.

В качестве следствия из доказанной леммы вытекает, что задача (4.4.4) – (4.4.8) является NP-трудной.

Заметим теперь, что задача (4.4.4) – (4.4.8) является задачей VC1 для графа G' , который получается из графа G путем «деления» каждой дуги $(v_1(e), v_2(e))$ графа G на три дуги $(x_1(e), t_e^1), (t_e^1, t_e^2), (t_e^2, x_2(e))$ посредством добавления двух новых вершин t_e^1 и t_e^2 . Отсюда ясно, что полученный граф G' является три раскрашиваемым, поскольку новые вершины t_e^1 и t_e^2 могут быть выкрашены соответственно первым и вторым цветом, а старые вершины — третьим цветом.

Таким образом, получаем, что NP-трудная задача VC1 для произвольного графа сводится к задаче VC1 для три раскрашиваемого графа G' . Следовательно, последняя задача также является NP-трудной.

4.2 Квадратичные полиномы от трех групп переменных

Рассмотрим квадратичный полином от булевых переменных следующего вида:

$$\begin{aligned}
& \sum_{i=1}^M a_i(1-x_i) + \sum_{k=1}^K b_k(1-y_k) + \sum_{l=1}^L c_l(1-z_l) + \\
& + \sum_{i=1}^M \sum_{k=1}^K d_{ik}x_iy_k + \sum_{i=1}^M \sum_{l=1}^L f_{il}x_iz_l + \sum_{k=1}^K \sum_{l=1}^L h_{kl}y_kz_l,
\end{aligned} \tag{4.4.14}$$

где $a_i \geq 0, b_k \geq 0, c_l \geq 0, d_{ik} \geq 0, f_{il} \geq 0, h_{kl} \geq 0, i = 1, \dots, M, k = 1, \dots, K, l = 1, \dots, L$. Такой полином называется *квадратичным полиномом от трех групп переменных*.

Теорема 4.4.1. Задача минимизации полинома (4.4.14) является NP-трудной.

Доказательство. Рассмотрим три раскрашиваемый граф $G=(V, E)$ и задачу VC1 для этого графа, которая, согласно лемме 4.4.1, является NP-трудной. Разобьем множество E на три подмножества. Поскольку граф G является три раскрашиваемым, то для всякого $e \in E$ вершины $v_1(e)$ и $v_2(e)$ имеют разные цвета. Тогда дуге $e \in E$ припишем цвет, отличный от цветов каждой из вершин $v_1(e)$ и $v_2(e)$. Обозначим через E_s множество дуг цвета $s = 1, 2, 3$. Согласно теоремам 1.3.1 и 1.4.1 задача VC1 сводится к задаче $MINP_0$, где соответствующий полином может быть записан следующим образом:

$$\sum_{v \in V} (1-y_v) + \sum_{s=1}^3 \sum_{e \in E_s} \Phi y_{v_1(e)} \cdot y_{v_2(e)}.$$

Нетрудно увидеть, что это полином вида (4.4.14). Таким образом, NP-трудная задача VC1 для три раскрашиваемого графа сводится к задаче минимизации полинома (4.4.14), что доказывает утверждение теоремы.

4.3 Полиномы третьей степени от трех групп переменных

Рассмотрим полином третьей степени от булевых переменных следующего вида:

$$\sum_{i=1}^M a_i(1-x_i) + \sum_{k=1}^K b_k(1-y_k) + \sum_{l=1}^L c_l(1-z_l) + \sum_{i=1}^M \sum_{k=1}^K \sum_{l=1}^L d_{ikl}x_iy_kz_l, \tag{4.4.15}$$

$a_i \geq 0, b_k \geq 0, c_l \geq 0, d_{ikl} \geq 0, i = 1, \dots, M, k = 1, \dots, K, l = 1, \dots, L$. Такой полином называется *полиномом третьей степени от трех групп переменных*.

Теорема 4.4.2. Задача минимизации полинома (4.4.15) является NP-трудной.

Доказательство. Пусть $G = (V, E)$ — три раскрашиваемый граф. Преобразуем задачу VC1 для этого графа в задачу SC1 следующего вида:

$$\begin{aligned}
& \min_{(x_v), (t_e), (u_e), (w_e)} \left\{ \sum_{v \in V} x_v + \sum_{e \in E} (t_e + u_e + w_e) \right\}; \\
& x_{v_1(e)} + x_{v_2(e)} + t_e \geq 1, \quad e \in E; \\
& x_{v_1(e)} + x_{v_2(e)} + u_e \geq 1, \quad e \in E; \\
& x_{v_1(e)} + x_{v_2(e)} + w_e \geq 1, \quad e \in E;
\end{aligned}$$

$$x_v, t_e, u_e, w_e \in \{0,1\}, \quad v \in V, \quad e \in E.$$

Нетрудно понять, что если $((x_v), (t_e), (u_e), (w_e))$ — оптимальное решение последней задачи, то $t_e = u_e = w_e = 0, e \in E$. Следовательно, рассматриваемые задачи $VC1$ и $SC1$ эквивалентны. Поскольку граф G является три раскрашиваемым, то для всякого $e \in E$ вершины $v_1(e)$ и $v_2(e)$ имеют разные цвета. Тогда дуге $e \in E$ припишем цвет, отличный от цветов каждой из вершин $v_1(e)$ и $v_2(e)$. Согласно теоремам 1.3.1 и 1.4.1 рассматриваемая задача $SC1$ сводится к задаче $MINP_0$ и соответствующий полином может быть записан следующим образом:

$$\sum_{v \in V} (1 - y_v) + \sum_{e \in E} ((1 - \bar{t}_e) + (1 - \bar{u}_e) + (1 - \bar{w}_e)) + \sum_{e \in E} \Phi y_{v_1(e)} \cdot y_{v_2(e)} \cdot (\bar{t}_e + \bar{u}_e + \bar{w}_e).$$

Поскольку для всякого $e \in E$ элементам тройки $(v_1(e), v_2(e), e)$ приписаны разные цвета, то несложно видеть, что данный полином есть полином вида (4.4.15). Таким образом, NP-трудная задача $VC1$ для три раскрашиваемого графа сводится к задаче минимизации полинома (4.4.15), что доказывает утверждение теоремы.

4.4 Квадратичные полиномы с разделяющимися переменными с коэффициентами разных знаков

Рассмотрим также естественное обобщение полиномов с разделяющимися переменными, каковыми являются квадратичные полиномы с разделяющимися переменными и коэффициентами произвольных знаков. Такие полиномы имеют вид:

$$\sum_{i=1}^m a_i (1 - y_i) + \sum_{k=1}^n b_k (1 - z_k) + \sum_{i=1}^m \sum_{k=1}^n c_{ik} y_i z_k, \quad (4.4.16)$$

Где $a_i, b_i, c_{ik}, i = 1, \dots, m, k = 1, \dots, n$, — коэффициенты разных знаков.

Теорема 4.4.3. Задача $MINP$ для полиномов вида (4.4.6) является NP-трудной.

Доказательство. Покажем, что задача (4.4.4) – (4.4.8), которая, согласно лемме 4.4.1, является NP-трудной, сводится к задаче $MINP$ для полиномов вида (4.4.16). По теоремам 1.3.1 и 1.4.1 задача (4.4.4) – (4.4.8) сводится к задаче $MINP_0$ следующего вида:

$$\sum_{v \in V} (1 - y_v) + \sum_{e \in E} ((1 - q_e^1) + (1 - q_e^2)) + \sum_{e \in E} \Phi (y_{v_1(e)} q_e^1 + y_{v_2(e)} q_e^2 + q_e^1 q_e^2), \quad (4.4.17)$$

где $\Phi > |V| + 2|E|$.

Покажем, что задача $MINP_0$ для полиномов вида (4.4.17) сводится к задаче $MINP_0$ для полиномов следующего вида:

$$\sum_{v \in V} (1 - y_v) + \sum_{e \in E} ((1 - q_e^1) + (1 - q_e^2)) + \sum_{e \in E} y_{v_1(e)} q_e^1 + \sum_{e \in E} \Phi (y_{v_2(e)} + q_e^2) q_e^1. \quad (4.4.18)$$

Действительно, заметим, прежде всего, что оптимальное решение задачи $MINP_0$ для полиномов вида (4.4.17) удовлетворяет условию

$$y_{v_1(e)} q_e^1 = 0, \quad e \in E.$$

Кроме того, заметим, что существует оптимальное решение задачи $MINP_0$ для полиномов вида (4.4.18), которое также удовлетворяет этому свойству. В самом деле, пусть

$(y_v), (q_e^1), (q_e^2)$ — оптимальное решение задачи, для которого множество $\tilde{E} = \{e \in E \mid y_{v_1(e)} \cdot q_e^1 = 1\}$ не пусто. Рассмотрим тогда допустимое решение $(\tilde{y}_v), (\tilde{q}_e^1), (\tilde{q}_e^2)$, определенное следующим образом:

$$\begin{aligned}\tilde{y}_v &= y_v, \\ \tilde{t}_e^1 &= \begin{cases} 0, & \text{если } e \in \tilde{E}, \\ t_e^1 & \text{иначе,} \end{cases} \\ \tilde{t}_e^2 &= t_e^2.\end{aligned}$$

Это решение — оптимальное, поскольку значение полинома на этом решении, очевидно, не превосходит его значения на исходном оптимальном решении. Кроме того, полученное оптимальное решение удовлетворяет требуемому условию. Отсюда получаем, что задача $MINP_0$ для полиномов вида (4.4.17) сводится к задаче $MINP_0$ для полиномов вида (4.4.18).

Заметим теперь, что последняя задача сводится к задаче $MINP_0$ для полиномов вида

$$\sum_{v \in V} (1 - y_v) + \sum_{e \in E} ((1 - q_e^1) + (1 - q_e^2)) + \sum_{e \in E} y_{v_1(e)} q_e^1$$

с дополнительными ограничениями

$$(y_{v_2(e)} + q_e^2) \cdot q_e^1 = 0, \quad e \in E.$$

Но всякое допустимое решение этой задачи удовлетворяет условию

$$q_e^2 = (1 - y_{v_2(e)})(1 - q_e^1), \quad e \in E.$$

Следовательно, последняя задача сводится к задаче $MINP$ для полиномов вида

$$\begin{aligned}\sum_{v \in V} (1 - y_v) + \sum_{e \in E} (1 - q_e^1) + \sum_{e \in E} (y_{v_2(e)} + q_e^1 - y_{v_2(e)} q_e^1) + \sum_{e \in E} y_{v_1(e)} q_e^1 = \\ |E| + \sum_{v \in V} (1 - y_v) + \sum_{e \in E} y_{v_2(e)} + \sum_{e \in E} y_{v_1(e)} q_e^1 + \sum_{e \in E} y_{v_2(e)} q_e^1.\end{aligned}$$

Но этот полином, как несложно увидеть, является полиномом вида (4.4.16). Таким образом, NP-трудная задача (4.4.4) – (4.4.8) сводится к задаче $MINP$ для полиномов вида (4.4.16). Это означает, что последняя задача является NP-трудной. Теорема доказана.

5 Полиномы с p -вхождениями

Полиномы с p -вхождениями образуют класс полиномов, близкий по свойствам к p -связным полиномам. В частности, полиномы с 2-вхождениями являются 3-связными, задача $MINP_0$ для которых, как известно, NP-трудна. Тем не менее, как следует из дальнейшего, задача $MINP_0$ для полиномов с 2-вхождениями эффективно разрешима и тем самым эти полиномы представляют собой пример 3-связных полиномов, для которых задача $MINP_0$ эффективно разрешима.

Рассмотрим полином $p_0(y_1, \dots, y_m)$ с неотрицательными коэффициентами и с характеристической матрицей $H = (h_{is})$ ($i \in I$, $s = 1, \dots, S$), записанный в канонической форме без подобных членов

$$p_0(y_1, \dots, y_m) = a_0 + \sum_{i \in I} f_i(1 - y_i) + \sum_{s=1}^S b_s \prod_{i|h_{is}=0} y_i.$$

Полином $p_0(y_1, \dots, y_m)$ называется *полиномом с p -вхождениями*, $p = 1, 2, 3, \dots$, если для всякого $i \in I$ переменная y_i входит не более чем в p нелинейных членов, то есть в i -й строке матрицы H имеется не более p нулевых элементов. Полином $p_0(y_1, \dots, y_m)$ будем называть полиномом со *строгими p -вхождениями*, $p = 1, 2, 3, \dots$, если для всякого $i \in I$ переменная y_i входит ровно в p нелинейных членов.

Если полином $p_0(y_1, \dots, y_m)$ есть полином с 1-вхождениями, то характеристическая матрица H обладает, очевидно, свойством 1-связности и задача $MINP_0$ для такого полинома эффективно разрешима. Более того, эту задачу можно считать тривиальной, поскольку полином с 1-вхождениями распадается на сумму полиномов с одним нелинейным членом в каждом:

$$p_0(y_1, \dots, y_m) = a_0 + \sum_{s=1}^S \left(\sum_{i|h_{is}=0} y_i + b_s \prod_{i|h_{is}=0} y_i \right).$$

Если $p_0(y_1, \dots, y_m)$ есть полином с 2-вхождениями, то характеристическая матрица H будет 3-всвязной. В общем случае задача $MINP_0$ для таких полиномов является NP-трудной, однако в данном частном случае может быть решена эффективно.

Транспонируем, прежде всего, задачу $MINP_0$ для полиномов с 2-вхождениями к случаю полиномов со строгими 2-вхождениями.

Лемма 4.5.1. Задача $MINP_0$ для полиномов с 2-вхождениями, не являющихся полиномами со строгими 2-вхождениями, сводится к задаче $MINP_0$ для полиномов с 2-вхождениями и числом переменных на единицу меньше.

Доказательство. Пусть у полинома $p_0(y_1, \dots, y_m)$ с числом переменных не меньше двух есть переменная (будем считать, что это переменная y_1), которая либо не входит ни в один нелинейный член, либо входит только в один нелинейный член $b_s \cdot \prod_{i|h_{is}=0} y_i$. В первом

случае сведение очевидно. Во втором случае представим исходный полином $p_0(y_1, \dots, y_m)$ в виде

$$p_0(y_1, \dots, y_m) = a_0 + f_1 + y_1 \left(-f_1 + b_s \cdot \prod_{i>1|h_{is}=0} y_i \right) + p'_0(y_2, \dots, y_m)$$

и рассмотрим функцию

$$y_1(y_2, \dots, y_m) = \begin{cases} 1, & \text{если } b_s \leq f_1, \\ 1 - \prod_{i>1|h_{is}=0} y_i, & \text{если } b_s > f_1. \end{cases}$$

Эта функция, очевидно, обладает свойством

$$y_1(y_2, \dots, y_m) = \begin{cases} 1, & \text{если } -f_1 + b_s \cdot \prod_{i>1|h_{is}=0} y_i \leq 0, \\ 0 & \text{иначе,} \end{cases}$$

и, следовательно, согласно основному принципу псевдобулева программирования, получаем, что задача $MINP_0$ для исходного полинома $p_0(y_1, \dots, y_m)$ сводится к задаче $MINP_0$ для следующего полинома

$$a_0 + f_1 + y_1(y_2, \dots, y_m) \left(-f_1 + b_s \cdot \prod_{i>1|h_{is}=0} y_i \right) + p'_0(y_2, \dots, y_m).$$

Этот полином является полиномом с 2-вхождениями. Действительно, если $y_1(y_2, \dots, y_m) = 1$, то это очевидно, а если $y_1(y_2, \dots, y_m) = 1 - \prod_{i>1|h_{is}=0} y_i$, то справедливость ска-

занного следует из равенства:

$$a_0 + f_1 + \left(1 - \prod_{i>1|h_{is}=0} y_i \right) \left(-f_1 + b_s \cdot \prod_{i>1|h_{is}=0} y_i \right) + p'_0(y_2, \dots, y_m) = a_0 + f_1 \prod_{i>1|h_{is}=0} y_i + p'_0(y_2, \dots, y_m).$$

Таким образом, в обоих случаях приходим к полиному с 2-вхождениями и числом переменных на единицу меньше, что завершает доказательство леммы.

Из леммы следует, что через конечное число шагов исключения переменных по исходному полиному с 2-вхождениями будет построен либо тривиальный полином от одной переменной, либо полином со строгими 2-вхождениями. В любом случае, задача минимизации исходного полинома сводится к задаче минимизации построенного полинома, а полученные при этом функции вида $y_1(y_2, \dots, y_m)$ служат для восстановления оптимальных значений удаленных переменных.

Сказанное означает, что при исследовании задачи $MINP_0$ для полиномов с 2-вхождениями можно ограничиться рассмотрением случая со строгими 2-вхождениями.

Теорема 4.5.1. Задача $MINP_0$ для полиномов со строгими 2-вхождениями сводится к задаче EC .

Доказательство. По теореме 1.4.1 задача $MINP_0$ для полинома $p_0(y_1, \dots, y_m)$ с характеристической матрицей $H = (h_{is})$ ($i \in I, s = 1, \dots, S$) сводится к задаче SC вида

$$\begin{aligned} & \min_{(x_i)(w_s)} \left\{ \sum_{i \in I} f_i x_i + \sum_{s=1}^S b_s w_s \right\}; \\ & \sum_{i=1}^m (1 - h_{is}) x_i + w_s \geq 1, \quad s = 1, \dots, S; \\ & x_i, w_s \in \{0, 1\}, \quad i \in I, \quad s = 1, \dots, S. \end{aligned}$$

Введем дополнительную переменную w_0 и добавим к рассматриваемой задаче SC следующие ограничения:

$$\begin{aligned}\sum_{s=1}^S w_s + w_0 &\geq 1; \\ w_0 &\geq 1; \\ w_0 &\in \{0, 1\}.\end{aligned}$$

Понятно, что рассматриваемая задача SC сводится к этой же задаче с указанными дополнительными ограничениями и, следовательно, исходная задача $MINP_0$ сводится к рассматриваемой задаче SC с дополнительными ограничениями. Но последняя задача, поскольку $\sum_{s=1}^S (1 - h_{is}) = 2$, $i \in I$, является задачей EC , что и доказывает теорему.

Граф G , соответствующий полученной задаче EC , имеет $S + 2$ вершины. Задача EC с S вершинами сводится к задаче о совершенном паросочетании минимального веса [] и может быть решена за полиномиальное время, оцениваемое величиной $O(S^3)$. Таким образом, трудоемкость алгоритма решения задачи $MINP_0$ для полиномов со строгими 2-вхождениями оценивается величиной $O(S^3)$, то есть величиной $O(m^3)$.

Можно выделить подкласс полиномов с 2-вхождениями, задача $MINP_0$ для которых сводится к задаче EC для двудольных графов, то есть к задаче о назначениях, для решения которой существуют значительно менее трудоемкие алгоритмы, чем для задачи о совершенном паросочетании минимального веса.

Рассмотрим полином $p_0(y_1, \dots, y_m)$ следующего вида

$$\sum_{i \in I} f_i(1 - y_i) + \sum_{s=1}^S b_s \prod_{i \in \beta_s} y_i + \sum_{t=1}^T d_t \prod_{i \in \gamma_t} y_i, \quad (4.5.1)$$

где $b_s > 0$, $s = 1, \dots, S$; $d_t > 0$, $t = 1, \dots, T$; β_s , $s = 1, \dots, S$; — попарно непересекающиеся подмножества множества I и γ_t , $t = 1, \dots, T$, — подмножества множества I с такими же свойствами.

Имеет место следующая теорема, позволяющая построить для задачи $MINP_0$ таких полиномов малотрудоемкий алгоритм.

Теорема 4.5.2. Задача $MINP_0$ для полиномов вида (4.5.1) сводится к задаче EC для двудольного графа.

Доказательство. Согласно лемме 4.5.1 полином (4.5.1) можно считать полиномом со строгими 2-вхождениями. Это означает, что

$$\bigcup_{s=1}^S \beta_s = \bigcup_{t=1}^T \gamma_t = I.$$

Пусть $H = (H^1, H^2)$, где $H^1 = (h_{is}^1)(i \in I, s = 1, \dots, S)$, $H^2 = (h_{it}^2)(i \in I, t = 1, \dots, T)$ — характеристическая матрица полинома (4.5.1). По теореме 1.4.1 задача $MINP_0$ для полинома (4.5.1) сводится к задаче SC вида

$$\begin{aligned}
& \min_{(x_i), (w_s^1), (w_t^2)} \left\{ \sum_{i \in I} f_i x_i + \sum_{s=1}^S b_s w_s^1 + \sum_{t=1}^T d_t w_t^2 \right\} \\
& \sum_{i \in I} (1 - h_{is}^1) x_i + w_s^1 \geq 1, \quad s = 1, \dots, S; \\
& \sum_{i \in I} (1 - h_{it}^2) x_i + w_t^2 \geq 1, \quad t = 1, \dots, T; \\
& x_i, w_s^1, w_t^2 \in \{0, 1\}, \quad i \in I, \quad s = 1, \dots, S, \quad t = 1, \dots, T.
\end{aligned}$$

Введем дополнительную переменную w_0 и дополним ограничения рассматриваемой задачи SC следующими соотношениями:

$$\begin{aligned}
& \sum_{s=1}^S w_s + w_0 \geq 1; \\
& \sum_{t=1}^T w_t + w_0 \geq 1; \\
& w_0 \in \{0, 1\}.
\end{aligned}$$

Поскольку рассматриваемая задача SC сводится к этой же задаче с дополнительными ограничениями, то задача $MINP_0$ для полинома (4.5.1) сводится к задаче SC с указанными дополнительными ограничениями. Но последняя задача суть задача EC для графа $G = (V, E)$ с множеством вершин $V = \{1, \dots, S, S+1\} \cup \{1, \dots, T, T+1\}$ и множеством ребер $E = I \cup \{1, \dots, S\} \cup \{1, \dots, T\} \cup \{(S+1, T+1)\}$ (см. рис. 4.5.1)

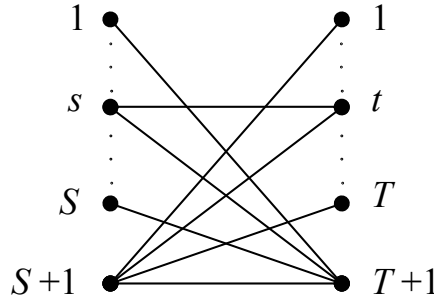


Рис. 4.5.1

Ребро $i, i \in I$, соединяет вершину $s, 1 \leq s \leq S$, и вершину $t, 1 \leq t \leq T$, если $i \in \beta_s \cap \gamma_t$; ребро $s, 1 \leq s \leq S$, соединяет вершины s и $T+1$, а ребро $t, 1 \leq t \leq T$, — вершины t и $S+1$. Поскольку никакое ребро не соединяет две вершины из множества $\{1, \dots, S, S+1\}$ или из множества $\{1, \dots, T, T+1\}$, то рассматриваемый граф G является двудольным, что завершает доказательства теоремы.

Поскольку задача EC для двудольного графа сводится к задаче о назначениях [], то из доказанного вытекает, что задача $MINP_0$ для полинома (4.5.1) может быть решена алгоритмом с оценкой трудоемкости $O(S^3 + T^3)$.

Следующая лемма показывает, что наличие у полиномов свойства p -вхождений является полезным для построения полиномиальных алгоритмов только если $p \leq 2$.

Лемма 4.5.2. Задача $VC1$ для кубического графа сводится к задаче $MINP_0$ для полиномов с 3-вхождениями.

Доказательство. Рассмотрим кубический граф $G = (I, E)$, где $|E|=S$ и матрицу инцидентий $H = (h_{is})$ ($i \in I, s = 1, \dots, S$) этого графа. По теоремам 1.3.1 и 1.4.1 задача $VC1$ сводится к задаче $MINP_0$ для полиномов вида

$$\sum_{i \in I} (1 - y_i) + \sum_{s=1}^S \Phi \prod_{i|h_{is}=1} y_i,$$

где $\Phi > t$. По определению кубического графа имеем $\sum_{s=1}^S h_{is} = 3, i \in I$. Это означает, что

приведенный полином есть полином с 3-вхождениями.

Поскольку задача $VC1$ для кубического графа является NP-трудной [], то из доказанной леммы вытекает, что задача $MINP_0$ для полиномов с 3-вхождениями также является NP-трудной. Таким образом, при переходе от класса полиномов с 2-вхождениями к классу полиномов с 3-вхождениями, так же как и случае полиномов со свойством p -связности, происходит качественный скачок в росте сложности задачи $MINP_0$. При $p = 2$ задача эффективно разрешима, а при $p = 3$ — NP-трудна.

В заключение отметим, что полиномы со свойством 2-связности и полиномы с 2-вхождениями образуют несравнимые классы полиномов и что характеристическая матрица полинома с 2-вхождениями не всегда приводима к матрице со свойством 2-связности. Достаточно просто привести пример (0,1)-матрицы, обладающей свойством 2-связности и имеющей более двух нулей в каждой строке. С другой стороны, транспонированная матрица инцидентий контура с двумя непересекающимися хордами не приводима к матрице со свойством 2-связности, но такая матрица имеет только две единицы в каждой строке и, следовательно, матрица \bar{H} является характеристической матрицей полинома с 2-вхождениями.

6 Квадратичные полиномы от булевых переменных

Из доказательства теоремы 4.1.1 о сводимости задачи $MINP$ к задаче $MINP_0$ видно, что задача $MINP$ для квадратичных полиномов с коэффициентами разных знаков сводится к задаче $MINP_0$ для квадратичных полиномов с неотрицательными коэффициентами при квадратичных членах. В силу этого центральным объектом исследования в классе квадратичных полиномов можно считать полиномы с неотрицательными коэффициентами. Задача $MINP_0$ для квадратичных полиномов является NP-трудной, поскольку к ней, согласно теоремам 1.3.1 и 1.4.1 сводится, например, NP-трудная задача VC для произвольного графа. В силу этого попытки построения эффективных универсальных алгоритмов решения данной задачи становятся бесперспективными и для получения точного решения следует

обратиться к алгоритмам, основанным на процедурах неявного перебора и, в частности, на рассмотренной в главе 2 вычислительной схеме ветвей и границ.

Работоспособность алгоритмов, работающих по схеме ветвей и границ, существенно зависит от степени использования полезной специфики рассматриваемой задачи, позволяющей, в частности, построить эффективный алгоритм вычисления достаточно точной нижней границы.

Задача $MINP_0$ для квадратичного полинома, как следует из теорем 1.3.1, 1.4.1 и 1.4.2, сводится к задаче SC и задаче VC . Линейные релаксации этих задач, полученные в результате замены условий целочисленности переменных на условия их неотрицательности, представляют собой задачи линейного программирования, которые могут быть использованы для вычисления нижних границ значений квадратичных полиномов.

Известно, кроме того, что задача VC и ее линейная релаксация обладают рядом замечательных свойств, которые оказываются полезными при построении алгоритмов ветвей и границ. Соответствующие результаты получены в работах [139, 140], касаются так называемой задачи о вершинной упаковке, эквивалентной задаче VC , и могут быть один к одному перенесены на задачу VC . Полученные результаты, во-первых, позволяют ограничить возможные значения компонент оптимального решения релаксированной задачи VC величинами 0, 1, $\frac{1}{2}$ и, во вторых, указывают на тесную связь оптимальных решений задачи VC и релаксированной задачи VC . Эта связь состоит в том, что для оптимального решения релаксированной задачи VC существует оптимальное решение задачи VC , совпадающее с первым решением в части целочисленных компонент оптимального решения релаксированной задачи. Кроме того, указанные результаты дают возможность.....

6.1 Квадратичные полиномы и задача о покрытии множества

Рассмотрим квадратичный полином $q_0(y_1, \dots, y_m)$ с неотрицательными коэффициентами при квадратичных членах, заданный в виде

$$q_0(y_1, \dots, y_m) = \sum_{i \in I} f_i (1 - y_i) + \sum_{i=1}^{m-1} \sum_{k=i+1}^m b_{ik} y_i y_k,$$

где $f_i \geq 0$, $i \in I$; $b_{ik} \geq 0$, $i = 1, \dots, m-1$, $k = i+1, \dots, m$. Задачу минимизации такого полинома далее будем обозначать $MINQ_0$.

Глава 1 содержит ряд важных результатов о сводимости задачи $MINP_0$ к задаче SC (теорема 1.4.1) и задаче VC (теорема 1.4.2). Применительно к задаче $MINQ_0$ эти результаты формулируются следующим образом:

Теорема 4.6.1. Задача $MINQ_0$ сводится к задаче SC вида

$$\min_{(x_i), (w_{lk})} \left\{ \sum_{i \in I} f_i x_i + \sum_{i=1}^{m-1} \sum_{k=i+1}^m b_{ik} w_{ik} \right\} \quad (4.6.1)$$

$$x_i + x_k + w_{ik} \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.2)$$

$$x_i, w_{lk} \geq 0, \quad i \in I, \quad l = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.3)$$

$$x_i, w_{lk} \in \{0, 1\}, \quad i \in I, \quad l = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.4)$$

Теорема 4.6.2. Задача $MINQ_0$ сводится к задаче VC вида

$$\min_{(x_i), (t_{ik}^1), (t_{ik}^2)} \left\{ \sum_{i \in I} f_i x_i + \sum_{i=1}^{m-1} \sum_{k=i+1}^m b_{ik} (t_{ik}^1 + t_{ik}^2) \right\} \quad (4.6.5)$$

$$x_i + t_{ik}^1 \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.6)$$

$$x_i + t_{ik}^2 \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.7)$$

$$t_{ik}^1 + t_{ik}^2 \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.8)$$

$$x_i, t_{lk}^1, t_{lk}^2 \geq 0, \quad i \in I, \quad l = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.9)$$

$$x_i, t_{lk}^1, t_{lk}^2 \in \{0, 1\}, \quad i \in I, \quad l = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.10)$$

Ниже приводятся некоторые полезные свойства оптимальных решений задачи VC и релаксированной задачи VC .

Пусть $G = (V, E)$ — неориентированный граф без петель, ребрам которого произвольным образом приписано направление. Обозначим через $v_1(e)$ и $v_2(e)$ соответственно начальную и конечную вершины образовавшейся дуги $e \in E$. Запишем задачу VC следующим образом:

$$\min_{(x_v)} \sum_{v \in V} c_v x_v \quad (4.6.11)$$

$$x_{v_1(e)} + x_{v_2(e)} \geq 1, \quad e \in E; \quad (4.6.12)$$

$$x_v \geq 0, \quad v \in V; \quad (4.6.13)$$

$$x_v \in \{0, 1\} \quad v \in V. \quad (4.6.14)$$

В уже упоминавшихся работах [139, 140] обнаружены замечательные свойства оптимальных решений так называемой задачи о вершинной упаковке, эквивалентной задаче VC для графа. Применительно к задаче (4.6.11) – (4.6.14) эти свойства формулируются следующим образом.

Теорема 4.6.3. Существует оптимальное решение (x_v) задачи (4.6.11) – (4.6.13) такое, что $x_v \in \{0, \frac{1}{2}, 1\}$, $v \in V$. Если (x_v) — оптимальное решение задачи (4.6.11) – (4.6.13), обладающее указанным свойством, то существует оптимальное решение (x_v^*) задачи (4.6.11) – (4.6.14) такое, что $x_v^* = x_v$, если $x_v \in \{0, 1\}$, $v \in V$.

Теорема 4.6.4. Задача линейного программирования (4.6.11) – (4.6.13) сводится к задаче VC для двудольного графа и может быть решена алгоритмом с оценкой трудоемкости $O(|V|^3)$.

Понятно, что приведенные свойства задачи (4.6.11) – (4.6.14) наряду с наличием относительно быстрого алгоритма решения задачи линейного программирования (4.6.11) – (4.6.13) открывают дополнительные возможности при построении вычислительно эффективных алгоритмов решения задачи VC для графа, основанных на вычислительной схеме ветвей и границ.

Понятно также, что поскольку задача VC , соответствующая задаче $MINQ_0$, обладает такими же замечательными свойствами, что и задача VC для графа, то эти свойства могут быть использованы при построении универсального алгоритма решения задачи $MINQ_0$, основанного на вычислительной схеме ветвей и границ. Работа такого алгоритма предполагает многократное обращение к алгоритму решения задачи (4.6.5) – (4.6.9), имеющему трудоемкость $O(m^6)$. С учетом того, что трудоемкость каждого шага алгоритма ветвей и границ является существенным фактором, влияющим на вычислительную эффективность алгоритма в целом, возникают следующие вопросы. Обладает ли задача (4.6.1) – (4.6.4) свойствами, аналогичными приведенным выше свойствам задачи VC и нельзя ли при построении алгоритма решения задачи $MINQ_0$ перейти от использования задачи (4.6.5) – (4.6.9) к решению задачи (4.6.1) – (4.6.4), имеющей существенно меньшее число переменных и ограничений.

Положительные ответы на эти вопросы базируются на следующей лемме.

Лемма 4.6.1. Задача (4.6.1) – (4.6.4) сводится к задаче (4.6.5) – (4.6.10), а задача (4.6.1) – (4.6.3) сводится к задаче (4.6.5) – (4.6.9).

Доказательство. Справедливость первого утверждения показана при доказательстве теоремы 1.4.2. Здесь отметим лишь то, что если $(x_i), (t_{ik}^1), (t_{ik}^2)$ — оптимальное решение задачи (4.6.5) – (4.6.10), то решение $(x_i), (w_{ik})$, где

$$w_{ik} = t_{ik}^1 + t_{ik}^2 - 1,$$

является оптимальным решением задачи (4.6.1) – (4.6.4).

Покажем справедливость второго утверждения леммы. Пусть $((x_i), (t_{ik}^1), (t_{ik}^2))$ — оптимальное решение задачи (4.6.5) – (4.6.9). Поставим ему в соответствие решение $((x_i), (w_{ik}))$ задачи (4.6.1) – (4.6.3), полагая

$$w_{ik} = t_{ik}^1 + t_{ik}^2 - 1. \quad (4.6.15)$$

Это решение является допустимым. Действительно, из условий (4.6.6) и (4.6.7) имеем

$$x_i + x_k + t_{ik}^1 + t_{ik}^2 \geq 2.$$

Отсюда получаем

$$x_i + x_k + w_{ik} \geq 1.$$

Кроме того, в силу условия (4.6.8) имеем $w_{ik} \geq 0$. Легко видеть также, что значения целевой функции (4.6.5) на исходном оптимальном решении и целевой функции (4.6.1) на построенном допустимом решении совпадают.

Построим теперь по оптимальному решению $(x_i), (w_{ik})$ задачи (4.6.1) – (4.6.3) допустимое решение задачи (4.6.5) – (4.6.9), на котором значение целевой функции (4.6.5) равняется оптимальному значению целевой функции (4.6.1). Тогда в силу признака оптимальности (лемма 1.1.1) решение $(x_i), (w_{ik})$, построенное по оптимальному решению $((x_i), (t_{ik}^1), (t_{ik}^2))$ задачи (4.6.5) – (4.6.9) с использованием равенств (4.6.15), будет оптимальным и, следовательно, задача (4.6.1) – (4.6.3) сводится к задаче (4.6.5) – (4.6.9).

Рассмотрим оптимальное решение $((x_i), (w_{ik}))$ и определим величины t_{ik}^1 и t_{ik}^2 для $i = 1, \dots, m-1, k = i, \dots, m$ следующим образом:

$$\begin{aligned} t_{ik}^1 &= 1 - x_i; \\ t_{ik}^2 &= 1 - x_k + \Delta_{ik}, \end{aligned}$$

где Δ_{ik} – некоторая неотрицательная величина, обеспечивающая выполнение условия (4.6.8). Если $x_i + x_k < 1$, то $\Delta_{ik} = 0$, а если $x_i + x_k \geq 1$, то Δ_{ik} – такая величина, что $t_{ik}^1 + t_{ik}^2 = 1$.

Покажем, что для так определенных величин t_{ik}^1 и t_{ik}^2 выполняется равенство (4.6.15). Действительно, если $w_{ik} > 0$ и, следовательно, $x_i + x_k < 1$, то

$$w_{ik} = 1 - x_i + x_k = 1 - t_{ik}^1 - t_{ik}^2.$$

Если же $w_{ik} = 0$, то $x_i + x_k \geq 1$ и, следовательно, $t_{ik}^1 + t_{ik}^2 = 1$. Таким образом, в обоих случаях равенство (4.6.15) выполняется, поэтому значение целевой функции (4.6.5) на построенном допустимом решении $((x_i), (t_{ik}^1), (t_{ik}^2))$ равняется оптимальному значению целевой функции (4.6.1) и, следовательно, второе утверждение леммы и утверждения леммы в целом доказаны.

Используя доказанную лемму, покажем, что задача (4.6.1) – (4.6.4) обладает теми же замечательными свойствами, что и задача (4.6.5) – (4.6.10).

Теорема 4.6.5. Существует оптимальное решение $((x_i), (w_{ik}))$ задачи (4.6.1) – (4.6.3) такое, что $x_i, w_{lk} \in \{0, \frac{1}{2}, 1\}$, $i \in I, l = 1, \dots, m-1, k = l+1, \dots, m$. Если $((x_i), (w_{ik}))$ — оптимальное решение задачи (4.6.1) – (4.6.3), обладающее указанным свойством, то существует оптимальное решение $((x_i^*), (w_{ik}^*))$ задачи (4.6.1) – (4.6.4) такое, что $x_i^* = x_i$, если $x_i \in \{0, 1\}$, $i \in I$.

Доказательство. Согласно теореме 4.6.3 существует оптимальное решение $((x_i), (t_{ik}^1), (t_{ik}^2))$ задачи (4.6.5) – (4.6.9) такое, что

$$x_i, t_{lk}^1, t_{lk}^2 \in \{0, \frac{1}{2}, 1\}, \quad i \in I, l = 1, \dots, m-1, k = l+1, \dots, m.$$

Но, согласно лемме 4.6.1, задача (4.6.1) – (4.6.3) сводится к задаче (4.6.5) – (4.6.9) и оптимальному решению $((x_i), (t_{ik}^1), (t_{ik}^2))$ последней задачи соответствует оптимальное решение $((x_i), (w_{ik}))$ задачи (4.6.1) – (4.6.3) такое, что $w_{ik} = t_{ik}^1 + t_{ik}^2 - 1$. Отсюда получаем, что

$$w_{ik} \in \{0, \frac{1}{2}, 1\}, \quad i = 1, \dots, m-1, k = i+1, \dots, m$$

и поэтому первое утверждение теоремы справедливо.

Докажем второе утверждение. Пусть $((x_i), (w_{ik}))$ — оптимальное решение задачи (4.6.1) – (4.6.3), обладающее нужным свойством. Рассмотрим соответствующее ему оптимальное решение $((x_i), (t_{ik}^1), (t_{ik}^2))$ задачи (4.6.5) – (4.6.9), построенное в ходе доказательства леммы 4.6.1. Из построения этого решения следует, что

$$t_{ik}^1, t_{ik}^2 \in \{0, \frac{1}{2}, 1\}, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m.$$

В силу теоремы 4.6.3 существует оптимальное решение $(\tilde{x}_i), (\tilde{t}_{ik}^1), (\tilde{t}_{ik}^2)$ задачи (4.6.5) – (4.6.10) такое, что $\tilde{x}_i = x_i$, если $x_i \in \{0, 1\}$, $i \in I$. По лемме 4.6.1 задача (4.6.1) – (4.6.4) сводится к задаче (4.6.5) – (4.6.10), а из доказательства леммы вытекает, что оптимальному решению $(\tilde{x}_i), (\tilde{t}_{ik}^1), (\tilde{t}_{ik}^2)$ задачи (4.6.5) – (4.6.10) соответствует оптимальное решение $((\tilde{x}_i), (\tilde{w}_{ik}))$ задачи (4.6.1) – (4.6.5). Для этого оптимального решения имеем $\tilde{x}_i = x_i$, если $x_i \in \{0, 1\}$, $i \in I$, что доказывает второе утверждение теоремы. Теорема доказана.

6.2. Эффективный алгоритм решения релаксированной задачи о покрытии множества

Выше установленные свойства задачи (4.6.1) – (4.6.4) будут полезны при построении для задачи $MINQ_0$ работоспособного алгоритма типа ветвей и границ только при наличии вычислительно эффективного алгоритма решения задачи (4.6.1) – (4.6.3). Приводимая ниже теорема дает основу для построения такого алгоритма.

Теорема 4.6.6. Задача (4.6.1) – (4.6.3) сводится к задаче $MINQ_0$ для полинома с разделяющимися переменными вида

$$\frac{1}{2} \left[\sum_{i \in I} f_i(1 - y_i) + \sum_{i \in I} f_i(1 - z_i) + \sum_{i=1}^{m-1} \sum_{k=i+1}^m b_{ik}(y_i z_k + y_k z_i) \right] \quad (4.6.16)$$

Доказательство. Задаче (4.6.1) – (4.6.3) поставим в соответствие задачу SC вида

$$\min_{(x_i^1), (x_i^2), (w_{ik}^1), (w_{ik}^2)} \frac{1}{2} \left\{ \sum_{i \in I} f_i(x_i^1 + x_i^2) + \sum_{i=1}^{m-1} \sum_{k=i+1}^m b_{ik}(w_{ik}^1 + w_{ik}^2) \right\}; \quad (4.6.17)$$

$$x_i^1 + x_i^2 + w_{ik}^1 \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.18)$$

$$x_i^1 + x_k^2 + w_{ik}^2 \geq 1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.19)$$

$$x_i^1, x_i^2, w_{lk}^1, w_{lk}^2 \in \{0, 1\}, \quad i \in I, \quad l = 1, \dots, m-1, \quad k = l+1, \dots, m; \quad (4.6.20)$$

и покажем, что задача (4.6.1) – (4.6.3) сводится к этой задаче.

Пусть $(x_i^1), (x_i^2), (w_{ik}^1), (w_{ik}^2)$ — оптимальное решение задачи (4.6.17) – (4.6.20). Поставим ему в соответствие решение $(x_i), (w_{ik})$ задачи (4.6.1) – (4.6.3) следующим образом

$$x_i = \frac{1}{2}(x_i^1 + x_i^2), \quad i \in I;$$

$$w_{ik} = \frac{1}{2}(w_{ik}^1 + w_{ik}^2), \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m.$$

Это допустимое решение, поскольку из неравенств (4.6.18) и (4.6.19) получаем

$$x_i + x_k + w_{ik} = \frac{1}{2}(x_i^1 + x_k^2 + w_{ik}^1) + \frac{1}{2}(x_i^2 + x_k^1 + w_{ik}^2) \geq 1.$$

Понятно также, что значения целевых функций задач (4.6.1) – (4.6.3) и (4.6.17) – (4.6.20) на рассматриваемых решениях совпадают.

Пусть теперь $(x_i), (w_{ik})$ — оптимальное решение задачи (4.6.1) – (4.6.3), удовлетворяющее условию

$$x_i, w_{lk} \in \{0, \frac{1}{2}, 1\}, i \in I, l = 1, \dots, m-1, k = l+1, \dots, m.$$

Положим

$$\begin{aligned} x_i^1 &= \begin{cases} 1, & \text{если } x_i = 1 \text{ или } x_i = \frac{1}{2}, \\ 0, & \text{если } x_i = 0; \end{cases} \\ x_i^2 &= \begin{cases} 1, & \text{если } x_i = 1, \\ 0, & \text{если } x_i = \frac{1}{2} \text{ или } x_i = 0; \end{cases} \\ w_{ik}^1 &= \begin{cases} 1, & \text{если } w_{ik} = 1, \\ \max\{0, 1 - x_i^1 - x_k^2\}, & \text{если } w_{ik} = \frac{1}{2}, \\ 0, & \text{если } w_{ik} = 0; \end{cases} \\ w_{ik}^2 &= \begin{cases} 1, & \text{если } w_{ik} = 1, \\ \max\{0, 1 - x_i^2 - x_k^1\}, & \text{если } w_{ik} = \frac{1}{2}, \\ 0, & \text{если } w_{ik} = 0. \end{cases} \end{aligned}$$

Проверим допустимость построенного решения $(x_i^1), (x_i^2), (w_{ik}^1), (w_{ik}^2)$. Ограничения (4.6.20), очевидно, выполняются. Убедимся в справедливости неравенств (4.6.18) и (4.6.19). Рассмотрим следующие возможные случаи.

Пусть $x_i = 1$ или $x_k = 1$. Тогда $x_i^1 = x_i^2 = 1$ или $x_k^1 = x_k^2 = 1$ и, следовательно, неравенства (4.6.18) и (4.6.19) имеют место.

Пусть $x_i = x_k = \frac{1}{2}$. Тогда $x_i^1 = x_k^1 = 1$ и поэтому неравенства (4.6.18) и (4.6.19) выполняются.

Пусть $x_i = \frac{1}{2}, x_k = 0$. Поскольку в этом случае $w_{ik} = \frac{1}{2}$, то имеем $x_i^1 = 1, x_i^2 = 0, x_k^1 = 0, x_k^2 = 0, w_{ik}^1 = 0, w_{ik}^2 = 1$. Отсюда видно, неравенства (4.6.18) и (4.6.19) справедливы.

Заметим также, что в каждом из рассмотренных случаев выполняются равенства

$$\begin{aligned} x_i &= \frac{1}{2}(x_i^1 + x_i^2), \quad i = 1, \dots, m; \\ w_{ik} &= \frac{1}{2}(w_{ik}^1 + w_{ik}^2), \quad i = 1, \dots, m-1, k = i+1, \dots, m. \end{aligned}$$

Отсюда получаем, что построенное решение является допустимым, а значение целевой функции (4.6.17) на этом решении совпадает с оптимальным значением целевой функции (4.6.1). Таким образом, согласно признаку оптимальности (лемма 1.1.1) задача (4.6.1) – (4.6.3) сводится к задаче (4.6.17) – (4.6.20).

Для завершения доказательства остается заметить, что задача (4.6.17) – (4.6.20) сводится к задаче $MINQ_0$ для полинома (4.6.16). Действительно, решения $(x_i^1), (x_i^2), (w_{ik}^1), (w_{ik}^2)$ и $(y_i), (z_i)$ рассматриваемых задач назовем соответствующими, если

$$\begin{aligned} y_i &= 1 - x_i^1, \quad i \in I; \\ z_i &= 1 - x_i^2, \quad i \in I; \\ (1 - x_i^1)(1 - x_k^2) &= w_{ik}^1, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \\ (1 - x_k^1)(1 - x_i^2) &= w_{ik}^2, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m. \end{aligned}$$

Понятно, что если $(x_i^1), (x_i^2), (w_{ik}^1), (w_{ik}^2)$ — оптимальное решение задачи (4.6.17) – (4.6.20), то два последних равенства выполняются. Следовательно, указанные соотношения для оптимального решения каждой из рассматриваемых задач определяют допустимое решение другой задачи. Понятно также, что на соответствующих решениях значения целевых функций рассматриваемых задач равны. Отсюда в силу критерия оптимальности (лемма 1.1.1), получаем сводимость задачи (4.6.17) – (4.6.20) к задаче $MINQ_0$ для полинома (4.6.16) и, следовательно, сводимость задачи (4.6.1) – (4.6.3) к задаче $MINQ_0$ для полинома (4.6.16). Теорема доказана.

Из доказанной теоремы вытекает, что если $((y_i), (z_k))$ — оптимальное решение задачи $MINQ_0$ для полинома (4.6.16), то решение $((x_i), (w_{ik}))$, где

$$\begin{aligned} x_i &= \frac{1}{2}(1 - y_i + 1 - z_i), \quad i \in I; \\ w_{ik} &= \frac{1}{2}(y_i z_k + y_k z_i), \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m, \end{aligned}$$

будет оптимальным решением задачи (4.6.1) – (4.6.3).

Согласно теореме 4.2.2, для получения оптимального решения $(y_i), (z_i)$ задачи $MINQ_0$ для полинома с разделяющимися переменными (4.6.16) достаточно решить задачу о минимальном разрезе соответствующей двухполюсной сети $G=(V, E)$. Эта сеть (см. рис. 4.6.1) имеет множество вершин

$$V = \{v_1^1, \dots, v_m^1\} \cup \{v_1^2, \dots, v_m^2\} \cup \{s, t\},$$

множество дуг

$$E = \{(s, v_i^1), i \in I\} \cup \{(v_i^2, t), i \in I\} \cup \{(v_i^1, v_k^2), (v_k^1, v_i^2), 1 \leq i < k \leq m\}$$

и следующие пропускные способности дуг

$$\begin{aligned} a(s, v_i^1) &= \frac{1}{2} f_i, \quad i \in I; \\ a(v_i^2, t) &= \frac{1}{2} f_i, \quad i \in I; \\ a(v_i^1, v_k^2) &= \frac{1}{2} b_{ik}, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m; \\ a(v_k^1, v_i^2) &= \frac{1}{2} b_{ik}, \quad i = 1, \dots, m-1, \quad k = i+1, \dots, m. \end{aligned}$$

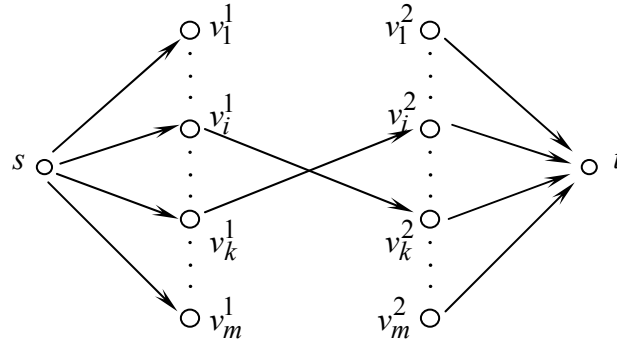


Рис. 4.6.1

Если (V_1, V_2) — оптимальный разрез рассмотренной сети G , то с учетом сказанного выше и теоремы 4.2.2 получаем, что решение $((x_i), (w_{ik}))$, где

$$x_i = \begin{cases} 1, & \text{если } v_i^1 \in V_2, v_i^2 \in V_1; \\ 0, & \text{если } v_i^1 \in V_1, v_i^2 \in V_2; \\ \frac{1}{2} & \text{иначе,} \end{cases}$$

$$w_{ik} = \max \{0, 1 - x_i - x_k\},$$

будет оптимальным решением задачи (4.6.1) – (4.6.3).

6.3 Алгоритм ветвей и границ

Используем рассмотренные выше свойства задачи $MINQ_0$ при построении для данной задачи алгоритма ветвей и границ. Общая схема такого алгоритма описана в главе 2. Для построения на основе этой общей схемы конкретного алгоритма решения той или иной задачи необходимо с учетом специфики и свойств этой задачи конкретизировать и определить следующие составные элементы общей схемы:

- базисные множества решений,
- способ задания подмножеств решений,
- функцию ветвления,
- способ вычисления нижней границы,
- функцию выбора наилучшего решения подмножества решений,
- правило выбора перспективного элемента разбиения.

Множеством допустимых решений задачи $MINQ_0$ для квадратичного полинома является множество B^m $(0,1)$ -векторов длины m . Для такого множества допустимых решений базисные подмножества, способ задания подмножеств решений, функция ветвления, а также правило выбора перспективного элемента разбиения определены в главе 2. Специфика задачи $MINQ_0$ должна быть использована при построении нижней границы и определении функции выбора наилучшего решения (подмножества решений).

С учетом сказанного относительно заданий подмножеств решений рассматриваемой задачи функция нижняя граница H должна быть определена для всякого частичного

решения (x_1, \dots, x_m) или, точнее говоря, для всякого множества $P(x_1, \dots, x_m)$ продолжений частичного решения (x_1, \dots, x_m) . Напомним, что частичным решением мы называем вектор (x_1, \dots, x_m) , элементы которого принимают три значения: 0, 1 и *. При этом считаем, что переменные, значения которых равняются *, — это свободные переменные, значения которых еще не определены и могут быть равными 0 или 1. Для частичного решения (x_1, \dots, x_m) положим $I_0 = \{i \in I \mid x_i = 0\}$, $I_1 = \{i \in I \mid x_i = 1\}$ и $I' = \{i \in I \mid x_i = *\}$. Всякий $(0,1)$ -вектор (y_1, \dots, y_m) называется продолжением частичного решения (x_1, \dots, x_m) , если $I_0 \subset \{i \in I \mid y_i = 0\}$, $I_1 \subset \{i \in I \mid y_i = 1\}$.

Для частичного решения (x_1, \dots, x_m) , у которого $I_0 = \emptyset$ и $I_1 = \emptyset$, значение $H(x_1, \dots, x_m)$ интересующей нас функции нижняя граница определена выше. Это есть оптимальное значение целевой функции задачи (4.6.1) – (4.6.3), которое равняется наименьшему значению полинома с разделяющимися переменными (4.6.16), которое, в свою очередь, равняется величине наименьшего разреза соответствующей двухполюсной сети. Определим значение функции $H(x_1, \dots, x_m)$ на любом частичном решении (x_1, \dots, x_m) . При этом для удобства будем считать, что у данного частичного решения $I' = \{m', m'+1, \dots, m\}$, где $1 < m' \leq m$.

Значения полинома $q_0(y_1, \dots, y_m)$ на решениях из множества $P(x_1, \dots, x_m)$ вычисляются следующим образом:

$$q_0(y_1, \dots, y_m) = \left(\sum_{i \in I_0} f_i + \sum_{i \in I_1} \sum_{k \in I_1 \mid i < k} b_{ik} \right) + \sum_{i \in I'} f_i (1 - y_i) + \sum_{i=m'}^{m-1} \sum_{k=i+1}^m b_{ik} y_i y_k.$$

Задача SC , соответствующая полиному

$$q'_0(y_{m'}, \dots, y_m) = \sum_{i \in I'} f_i (1 - y_i) + \sum_{i=m'}^{m-1} \sum_{k=i+1}^m b_{ik} y_i y_k,$$

имеет вид

$$\min_{(z_i), (w_{ik})} \left\{ \sum_{i \in I'} f_i z_i + \sum_{i=m'}^{m-1} \sum_{k=i+1}^m b_{ik} w_{ik} \right\}; \quad (4.6.21)$$

$$z_i + z_k + w_{ik} \geq 1, \quad i = m', \dots, m-1, \quad k = i+1, \dots, m; \quad (4.6.22)$$

$$z_i, w_{lk} \geq 0, \quad i \in I', \quad l = m', \dots, m-1, \quad k = l+1, \dots, m; \quad (4.6.23)$$

$$z_i, w_{lk} \in \{0,1\}, \quad i \in I', \quad l = m', \dots, m-1, \quad k = l+1, \dots, m. \quad (4.6.24)$$

Тогда значение функции $H(x_1, \dots, x_m)$ на рассматриваемом частичном решении (x_1, \dots, x_m) может быть определено следующим образом:

$$H(x_1, \dots, x_m) = \left(\sum_{i \in I_0} f_i + \sum_{i \in I_1} \sum_{k \in I_1 \mid i < k} b_{ik} \right) + \sum_{i \in I'} f_i z_i^* + \sum_{i=m'}^{m-1} \sum_{k=i+1}^m b_{ik} w_{ik}^*,$$

где $(z_i^*), (w_{ik}^*)$ — оптимальное решение задачи линейного программирования (4.6.21) – (4.6.23). Для вычисления этого оптимального решения может быть использован алгоритм построения минимального разреза соответствующей сети.

Понятно, что если оптимальное решение $((z_i^*), (w_{ik}^*))$ задачи (4.6.21) – (4.6.23) оказывается целочисленным, то тем самым оно является оптимальным решением задачи (4.6.21) – (4.6.24) и, следовательно, в подмножестве $P(x_1, \dots, x_m)$ найдено наилучшее решение. Для таких частичных решений считаем определенной функцию выбора наилучшего решения.

Пусть оптимальное решение $((z_i^*), (w_{ik}^*))$ не является целочисленным, но вектор (z_i^*) имеет целочисленные компоненты и пусть $I'_0 = \{i \in I' \mid z_i^* = 1\}$, $I'_1 = \{i \in I' \mid z_i^* = 0\}$. Тогда согласно теореме 4.6.5 во множестве $P(x_1, \dots, x_m)$ можно выделить несобственное подмножество, содержащее наилучшее решение множества $P(x_1, \dots, x_m)$. Это подмножество задается частичным решением (x'_1, \dots, x'_m) таким, что $\{i \mid x'_i = 0\} = I_0 \cup I'_0$, $\{i \mid x'_i = 1\} = I_1 \cup I'_1$. Для таких частичных решений (x_1, \dots, x_m) считаем определенной функцию выбора наилучшего подмножества.

Указанные особенности процедуры ветвей и границ для задачи $MINQ_0$, в том числе возможность точного решения оценочной задачи линейного программирования за время $O(m^3)$ и, что особенно важно, возможность отыскания для большого числа исследуемых множеств (элементов разбиения множества неотброшенных решений) наилучшего подмножества или даже решения, позволяет надеяться на вычислительную работоспособность данной процедуры, для завершения работы которой может оказаться достаточным относительно небольшого числа шагов.