Реализация смешанно-целочисленных отсечений Гомори

А. О. Махорин*

Декабрь 2007 г.

1 Математическое обоснование

Рассмотрим смешанно-целочисленную задачу:

$$\min\{c^T x : x \in X\}, \tag{1}$$

$$X = \{x \in \mathbf{Z}_+^p \times \mathbf{R}_+^{n-p} : Ax = b\}, \tag{2}$$

и допустим, что получено базисное оптимальное решение ЛП-релаксации данной задачи, т. е. для этого решения указано разбиение переменных на базисные и небазисные $x = (x_B, x_N)$, а также соответствующее разбиение $A = (B \,|\, N)$, где B — невырожденная матрица, называемая базисом. Указанное разбиение позволяет явно выразить базисные переменные через небазисные посредством надлежащего эквивалентного преобразования исходной системы ограничений-равенств:

$$Ax = b \quad \Leftrightarrow \quad Bx_B + Nx_N = b \quad \Leftrightarrow \quad x_B = -B^{-1}Nx_N + B^{-1}b.$$

Таким образом:

$$x_B = -\widetilde{A}x_N + \beta, \tag{3}$$

где матрица $\widetilde{A} = (\alpha_{ij}) = B^{-1}N$ — это так называемая симплекс-таблица, вектор $\beta = (\beta_i) = B^{-1}b$ — преобразованный вектор правых частей ограничений. Поскольку в базисном решении все небазисные переменные находятся на своих нулевых границах, то такое решение имеет вид:

$$x^* = (x_B^*, x_N^*) = (\beta, 0), \tag{4}$$

т. е. значения базисных переменных равны соответствующим компонентам вектора β (при этом $\beta \geq 0$, поскольку базисное решение предполагается оптимальным, а значит, допустимым для ЛП-релаксации).

*Кафедра прикладной информатики, Московский авиационный институт, Москва, Россия. E-mail: <mao@mai2.rcnet.ru>, <mao@gnu.org>.

Если x^* является целочисленно допустимым решением, т. е. оптимальные значения всех целочисленных переменных являются целыми числами, то x^* одновременно является оптимальным решением задачи (1)—(2), поэтому процесс решения заканчивается. В противном случае найдется хотя бы одна целочисленная переменная, оптимальное значение которой (для ЛП-релаксации) является дробным. Понятно, что такой переменной может быть только базисная переменная, поскольку оптимальные значения всех небазисных переменных равны нулю, а значит, являются целыми числами. Итак, пусть такой переменной является целочисленная переменная $(x_B)_i$, причем $(x_B^*)_i = \beta_i \notin \mathbf{Z}$, т. е. требование целочисленности нарушено. Строка симплекс-таблицы (3), соответствующая выбранной базисной переменной, имеет вид:

$$(x_B)_i = \beta_i - \sum_{j \in N} \alpha_{ij} (x_N)_j. \tag{5}$$

Поскольку симплекс-таблица (3) эквивалентна исходной системе ограничений-равенств $Ax = b$, то равенство (3), а значит, и равенство (5) должны выполняться для любого целочисленно допустимого решения $x \in X$. Отсюда следует, что условие:

$$\beta_i - \sum_{j \in N} \alpha_{ij} (x_N)_j \in \mathbf{Z} \tag{6}$$

является правильным (valid) условием для X. Заметим, что β_i можно представить в виде суммы $\beta_i = \lfloor \beta_i \rfloor + f(\beta_i)$, где $\lfloor \cdot \rfloor$ — целая часть числа, $f(\cdot)$ — дробная часть числа. Поэтому условие (6) можно заменить следующим эквивалентным условием:

$$f(\beta_i) - \sum_{j \in N} \alpha_{ij} (x_N)_j \in \mathbf{Z}, \tag{7}$$

которое также является правильным для X.

Точка, в которой сумма $\sum_{j \in N} \alpha_{ij} (x_N)_j$ равна нулю, не может принадлежать множеству X, поскольку значение β_i предполагается дробным. Это позволяет представить X в виде объединения двух непересекающихся множеств:

$$X = X^+ \cup X^-, \tag{8}$$

где

$$X^+ = \{x \in X : \sum_{j \in N} \alpha_{ij} (x_N)_j > 0\}, \tag{9}$$

$$X^- = \{x \in X : \sum_{j \in N} \alpha_{ij} (x_N)_j < 0\}. \tag{10}$$

Поскольку $f(\beta_i) < 1$, то неравенство:

$$f(\beta_i) - \sum_{j \in N} \alpha_{ij} (x_N)_j < 1 \tag{11}$$

является правильным для X^+. Но тогда из (11) с учетом (7) следует, что неравенство:

$$f(\beta_i) - \sum_{j \in N} \alpha_{ij}(x_N)_j \le 0 \quad \Leftrightarrow \quad \sum_{j \in N} \alpha_{ij}(x_N)_j \ge f(\beta_i) \tag{12}$$

также является правильным для X^+. Аналогично, так как β_i есть дробная величина, то $f(\beta_i) > 0$. Поэтому неравенство:

$$f(\beta_i) - \sum_{j \in N} \alpha_{ij}(x_N)_j > 0 \tag{13}$$

является правильным для X^-, откуда с учетом (7) следует, что тогда неравенство:

$$f(\beta_i) - \sum_{j \in N} \alpha_{ij}(x_N)_j \ge 1 \quad \Leftrightarrow \quad \frac{f(\beta_i)}{1 - f(\beta_i)} \sum_{j \in N} (-\alpha_{ij})(x_N)_j \ge f(\beta_i) \tag{14}$$

также является правильным для X^-.

Далее нам потребуется следующая

Т е о р е м а. Пусть неравенство $a^T x \ge b$ является правильным для множества P, а неравенство $c^T x \ge d$ — правильным для множества Q. Тогда неравенство:

$$\sum_j \max(a_j, c_j) x_j \ge \min(b, d) \tag{15}$$

является правильным для множеств $P \cup Q$ и $\mathrm{conv}(P \cup Q)$.

Неравенство (12) является правильным для множества X^+, а неравенство (14) — правильным для множества X^-. Поэтому в соответствии с приведенной теоремой неравенство:

$$\sum_{j \in N} \max[\alpha_{ij}, \frac{f(\beta_i)}{1 - f(\beta_i)}(-\alpha_{ij})](x_N)_j \ge f(\beta_i) \tag{16}$$

является правильным для исходного смешанно-целочисленного множества $X = X^+ \cup X^-$. Так как $0 < f(\beta_i) < 1$, то $f(\beta_i)/[1 - f(\beta_i)] > 0$. Поэтому:

$$\max[\alpha_{ij}, \frac{f(\beta_i)}{1 - f(\beta_i)}(-\alpha_{ij})] = \begin{cases} \alpha_{ij}, & \text{если } \alpha_{ij} \ge 0 \\ \dfrac{f(\beta_i)}{1 - f(\beta_i)}(-\alpha_{ij}), & \text{если } \alpha_{ij} < 0 \end{cases} \tag{17}$$

что позволяет записать неравенство (16) в следующем виде:

$$\sum_{j \in N^+} |\alpha_{ij}|(x_N)_j + \frac{f(\beta_i)}{1 - f(\beta_i)} \sum_{j \in N^-} |\alpha_{ij}|(x_N)_j \ge f(\beta_i), \tag{18}$$

где $N^+ = \{j \in N : \alpha_{ij} \geq 0\}$, $N^- = \{j \in N : \alpha_{ij} < 0\}$.

В оптимальной точке x^* (см. (4)) неравенство (18) нарушается, так как в этой точке $x^*_N = 0$, а $f(\beta_i) > 0$, поскольку значение базисной переменной $(x^*_B)_i = \beta_i$ предполагается дробным. С другой стороны, неравенство (18) является правильным для множества допустимых решений X задачи (1)—(2). Следовательно, преобразуя указанное неравенство к виду равенства (введением переменной избытка) и добавляя его к исходной системе ограничений-равенств $Ax = b$, мы можем отсечь дробное решение x^* от смешанно-целочисленного множества X.

Неравенство (18) можно усилить следующим образом. Заметим, что это неравенство является следствием условия (7), которое не нарушится, если любой коэффициент α_{ij} при *целочисленной* переменной $(x_N)_j$ заменить коэффициентом $\alpha_{ij} - \delta$, где δ — произвольное *целое* число. Если $\alpha_{ij} - \delta \geq 0$, то $j \in N^+$, и поэтому соответствующий коэффициент при $(x_N)_j$ в неравенстве (18) будет равен $|\alpha_{ij} - \delta|$. Если же $\alpha_{ij} - \delta < 0$, то $j \in N^-$, и поэтому соответствующий коэффициент при $(x_N)_j$ в неравенстве (18) будет равен $f(\beta_i)/[1-f(\beta_i)] \cdot |\alpha_{ij} - \delta|$. Поскольку все коэффициенты в неравенстве (18) являются положительными, то для получения возможно более сильного отсечения необходимо сделать эти коэффициенты как можно меньше. Отсюда следует, что если $\alpha_{ij} - \delta \geq 0$, то наименьшее значение $|\alpha_{ij} - \delta|$ получается при выборе $\delta = \lfloor \alpha_{ij} \rfloor$ и равно $f(\alpha_{ij})$, а если $\alpha_{ij} - \delta < 0$, то наименьшее значение $|\alpha_{ij} - \delta|$ получается при выборе $\delta = \lceil \alpha_{ij} \rceil = \lfloor \alpha_{ij} \rfloor + 1$ и равно $1 - f(\alpha_{ij})$ (в последнем случае α_{ij} предполагается дробным). Таким образом, в качестве *наиболее подходящего* коэффициента при целочисленной переменной $(x_N)_j$ в неравенстве (18) следует взять величину:

$$\min\{f(\alpha_{ij}), \frac{f(\beta_i)}{1 - f(\beta_i)}[1 - f(\alpha_{ij})]\}. \tag{19}$$

Чтобы придать формуле (19) более удобный вид, рассмотрим случай:

$$f(\alpha_{ij}) \leq \frac{f(\beta_i)}{1 - f(\beta_i)}[1 - f(\alpha_{ij})].$$

Поскольку $f(\beta_i) < 1$, то $1 - f(\beta_i) > 0$. Следовательно:

$$\begin{aligned} & f(\alpha_{ij})[1 - f(\beta_i)] \leq f(\beta_i)[1 - f(\alpha_{ij})] \quad \Leftrightarrow \\ & f(\alpha_{ij}) - f(\alpha_{ij})f(\beta_i) \leq f(\beta_i) - f(\beta_i)f(\alpha_{ij}) \quad \Leftrightarrow \\ & f(\alpha_{ij}) \leq f(\beta_i). \end{aligned}$$

Таким образом, наиболее подходящий коэффициент при целочисленной переменной $(x_N)_j$ равен:

$$\begin{cases} f(\alpha_{ij}), & \text{если } f(\alpha_{ij}) \leq f(\beta_i) \\ \dfrac{f(\beta_i)}{1 - f(\beta_i)}[1 - f(\alpha_{ij})], & \text{если } f(\alpha_{ij}) > f(\beta_i) \end{cases} \tag{20}$$

(Заметим, в частности, что если исходный коэффициент α_{ij} является целочисленным, то наиболее подходящий коэффициент будет равен нулю.)

В результате усиления неравенства (18) за счет выбора наиболее подходящих коэффициентов при целочисленных переменных в соответствии с формулой (20) мы приходим к окончательному неравенству, которое называется *смешанно-целочисленным отсечением Гомори*:

$$\begin{aligned}\sum_{j\in N^+_{\mathbf{Z}}} f(\alpha_{ij})(x_N)_j + \frac{f(\beta_i)}{1-f(\beta_i)} \sum_{j\in N^-_{\mathbf{Z}}} [1-f(\alpha_{ij})](x_N)_j + \\ \sum_{j\in N^+_{\mathbf{R}}} |\alpha_{ij}|(x_N)_j + \frac{f(\beta_i)}{1-f(\beta_i)} \sum_{j\in N^-_{\mathbf{R}}} |\alpha_{ij}|(x_N)_j \geq f(\beta_i),\end{aligned} \tag{21}$$

где

$$\begin{aligned} N^+_{\mathbf{Z}} &= \{j \in N : (x_N)_j \in \mathbf{Z},\ f(\alpha_{ij}) \leq f(\beta_i)\}, \\ N^-_{\mathbf{Z}} &= \{j \in N : (x_N)_j \in \mathbf{Z},\ f(\alpha_{ij}) > f(\beta_i)\}, \\ N^+_{\mathbf{R}} &= \{j \in N : (x_N)_j \in \mathbf{R},\ \alpha_{ij} \geq 0\}, \\ N^-_{\mathbf{R}} &= \{j \in N : (x_N)_j \in \mathbf{R},\ \alpha_{ij} < 0\}. \end{aligned} \tag{22}$$

(Следует отметить, что хотя неравенство (21) является более сильным, чем неравенство (18), величина нарушения этих неравенств в точке x^* одна и та же и равна $f(\beta_i)$.)

Р. Гомори показал[1], что алгоритм решения смешанно-целочисленной задачи (1)—(2), основанный на последовательном добавлении отсечений вида (21) к исходной системе ограничений-равенств $Ax = b$, определяющих множество X, позволяет найти оптимальное решение указанной задачи за конечное число шагов (при условии, что $c^T x \in \mathbf{Z}$ для всех $x \in X$).

2 Дополнительные преобразования

Для вычисления строки симплекс-таблицы в пакете GLPK предусмотрена подпрограмма `glp_eval_tab_row`, которая возвращает результат в следующем формате:[2]

$$(x_B)_i = \sum_{j\in N} \xi_{ij}(x_N)_j. \tag{23}$$

Данный формат отличается от стандартного формата (5) тем, что здесь небазисные переменные могут иметь ненулевые нижние границы, конечные верхние границы, или иметь фиксированное значение. (Небазисные

[1] R. E. Gomory, An algorithm for the mixed integer problem, Technical Report RM-2597, RAND Corp., 1960.

[2] Подробнее см. справочное руководство по пакету GLPK.

переменные также могут быть свободными, т. е. неограниченными по знаку. Однако, если хотя бы одна такая переменная входит в строку симплекс-таблицы с ненулевым коэффициентом, соответствующая строка не используется для генерации отсечения.)

Переход от формата (23) к стандартному формату (5) выполняется следующим образом. Множество всех небазисных переменных N разбивается на три непересекающихся подмножества:

$$N = N_L \cup N_U \cup N_S, \tag{24}$$

где N_L — множество небазисных переменных, у которых активна нижняя граница, N_U — множество небазисных переменных, у которых активна верхняя граница, N_S — множество небазисных переменных, имеющих фиксированное значение. В зависимости от статуса (24) отдельной исходной небазисной переменной $(x_N)_j$ используется следующая подстановка:

$$(x_N)_j = \begin{cases} (l_N)_j + y_j, & \text{если } j \in N_L \\ (u_N)_j - y_j, & \text{если } j \in N_U \\ (s_N)_j, & \text{если } j \in N_S \end{cases} \tag{25}$$

где $y_j \geq 0$ — преобразованная небазисная переменная, заменяющая исходную небазисную переменную, $(l_N)_j$ — нижняя граница исходной небазисной переменной, $(u_N)_j$ — верхняя граница исходной небазисной переменной, $(s_N)_j$ — фиксированное значение исходной небазисной переменной. Эта подстановка позволяет преобразовать строку симплекс-таблицы (23) к стандартному формату (5):

$$\begin{aligned} (x_B)_i &= \sum_{j \in N_L} \xi_{ij}(x_N)_j + \sum_{j \in N_U} \xi_{ij}(x_N)_j + \sum_{j \in N_S} \xi_{ij}(x_N)_j = \\ &= \sum_{j \in N_L} \xi_{ij}[(l_N)_j + y_j] + \sum_{j \in N_U} \xi_{ij}[(u_N)_j - y_j] + \sum_{j \in N_S} \xi_{ij}(s_N)_j = \\ &= \beta_i - \sum_{j \in N} \alpha_{ij} y_j, \end{aligned}$$

где

$$\alpha_{ij} = \begin{cases} -\xi_{ij}, & \text{если } j \in N_L \\ +\xi_{ij}, & \text{если } j \in N_U \\ 0, & \text{если } j \in N_S \end{cases} \tag{26}$$

$$\beta_i = \sum_{j \in N_L} \xi_{ij}(l_N)_j + \sum_{j \in N_U} \xi_{ij}(u_N)_j + \sum_{j \in N_S} \xi_{ij}(s_N)_j. \tag{27}$$

(Заметим, что стандартный формат (5) также предполагает неотрицательность базисной переменной $(x_B)_i$. Однако, поскольку в формуле отсечения (21) используется только дробная часть текущего значения этой переменной $f(\beta_i)$, преобразование $(x_B)_i$ не требуется.)

Использование формулы (21) для коэффициентов (26) и (27) дает отсечение, выраженное в терминах преобразованных небазисных переменных:

$$\sum_{j \in N_L \cup N_U} \overline{\varphi}_j y_j \geq \overline{\rho}, \tag{28}$$

где $\overline{\varphi}_j$ — коэффициенты отсечения, $\overline{\rho} = f(\beta_i)$ — правая часть отсечения. Поэтому далее выполняется подстановка, обратная к (25):

$$y_j = \begin{cases} (x_N)_j + (l_N)_j, & \text{если } j \in N_L \\ (u_N)_j - (x_N)_j, & \text{если } j \in N_U \end{cases} \tag{29}$$

чтобы выразить полученное отсечение в терминах исходных небазисных переменных:

$$\sum_{j \in N_L \cup N_U} \overline{\varphi}_j y_j \geq \overline{\rho} \quad \Rightarrow$$
$$\sum_{j \in N_L} \overline{\varphi}_j [(x_N)_j - (l_N)_j] + \sum_{j \in N_U} \overline{\varphi}_j [(u_N)_j - (x_N)_j] \geq \overline{\rho}.$$

Таким образом, отсечение принимает вид:

$$\sum_{j \in N_L \cup N_U} \varphi_j (x_N)_j \geq \rho, \tag{30}$$

где

$$\varphi_j = \begin{cases} +\overline{\varphi}_j, & \text{если } j \in N_L \\ -\overline{\varphi}_j, & \text{если } j \in N_U \end{cases} \tag{31}$$

$$\rho = \overline{\rho} + \sum_{j \in N_L} \overline{\varphi}_j (l_N)_j - \sum_{j \in N_U} \overline{\varphi}_j (u_N)_j. \tag{32}$$

Среди небазисных переменных в неравенстве (30) могут быть как структурные, так и вспомогательные переменные, в то время как пакет GLPK допускает строки (ограничения), содержащие только структурные переменные. Чтобы привести неравенство (30) к формату GLPK, для каждой небазисной вспомогательной переменной $(x_N)_j = (x_R)_i$, входящей в это неравенство, выполняется подстановка:

$$(x_R)_i = \sum_j a_{ij} (x_S)_j, \tag{33}$$

где a_{ij} — коэффициенты i-й строки (ограничения) исходной задачи, $(x_S)_j$ — соответствующие структурные переменные. В процессе выполнения указанной подстановки фиксированные структурные переменные, которые могут присутствовать в (33), не включаются в левую часть неравенства (30), а учитываются соответствующим изменением правой части этого неравенства.

После подстановки вспомогательных переменных отсечение (30) принимает окончательный вид, в котором оно может быть добавлено к формулировке исходной задачи.
