

Верхняя оценка сложности одного из вариантов метода ветвей и границ для задачи о сумме подмножеств

Р.М. Колпаков, М.А. Посыпкин, Си Ту Тант Син

Аннотация— Получена верхняя оценка сложности решения задачи о сумме подмножеств, являющейся частным случаем задачи о ранце, одним из вариантов метода ветвей и границ с дополнительным условием отсева подзадач, основанным на сравнении максимального и минимального количества предметов, которые можно положить в ранец. В качестве вспомогательных результатов установлены различные комбинаторные свойства подзадач, получаемых в процессе решения задачи о сумме подмножеств рассматриваемым вариантом метода ветвей и границ.

Ключевые слова — задача о сумме подмножеств, оценки сложности, комбинаторный анализ, булев куб.

I. ВВЕДЕНИЕ

В данной работе рассматривается задача о сумме подмножеств [1] и один из методов ее решения – метод ветвей и границ (МВГ). Задача о сумме подмножеств является частным случаем задачи о ранце [1,2], в которой вес каждого предмета совпадает с его стоимостью. Помимо традиционных способов сокращения перебора при решении данной задачи методом ветвей и границ, рассматривается дополнительное условие исключения из дальнейшего рассмотрения подзадач, возникающих в процессе решения, основанное на сравнении максимального и минимального числа предметов, которые можно положить в ранец.

Получена верхняя оценка сложности решения задачи о сумме подмножеств рассматриваемым методом в «наихудшем» случае, зависящая только от числа переменных. Показано, что введение дополнительного условия исключения подзадач сокращает сложность решения задачи о сумме подмножеств приблизительно вдвое.

II. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Рассматривается задача о сумме подмножеств в следующей постановке:

$$\begin{aligned} f(x) &= \sum_{i=1}^n w_i x_i \rightarrow \max, \\ \sum_{i=1}^n w_i x_i &\leq C, \\ x_i &\in \{0,1\}, i = 1, \dots, n. \end{aligned} \quad (1)$$

Предполагается, что коэффициенты $C, w_i, i = 1, \dots, n$, данной задачи являются натуральными числами. Для решения этой задачи методом ветвей и границ [1,2] используются понятия рекордного решения и подзадачи.

Любой набор значений переменных $x = (x_1, \dots, x_n)$ задачи (1), удовлетворяющий ее ограничениям, называется *допустимым решением*. *Оптимальным* называется допустимое решение, на котором достигается максимум целевой функции f задачи (1). Значение целевой функции на оптимальном решении называется *оптимумом*. *Рекордным* называется некоторое найденное допустимое решение задачи (1), на котором функция f достигает максимального значения среди всех найденных на данный момент времени решений. Значение целевой функции на рекордном решении называется *рекордом*. Из определения рекорда очевидно, что рекорд не уменьшается в процессе решения задачи и не превосходит оптимума.

Подзадачей задачи (1) называется задача, получаемая из задачи (1) фиксацией значений переменных x_1, \dots, x_τ :

$$\begin{aligned} f(x) &= \sum_{i=1}^n w_i x_i \rightarrow \max, \\ \sum_{i=1}^n w_i x_i &\leq C, \\ x_i &= \theta_i, 1 \leq i \leq \tau, \\ x_i &\in \{0,1\}, i = \tau + 1, \dots, n. \end{aligned} \quad (2)$$

где θ_i - заданные значения фиксированных переменных, $\theta_i \in \{0,1\}, i = 1, \dots, n$. Очевидно, что любое допустимое решение задачи (2) является допустимым решением задачи (1). Для подзадачи P вида (2) будем использовать следующие обозначения:

Работа выполнена при поддержке гранта РФФИ № 15-07-03102 А.

Р.М. Колпаков работает в МГУ имени М.В. Ломоносова и в ВЦ ФИЦ ИУ РАН (e-mail: foroman@mail.ru)

М.А.Посыпкин работает в ВЦ ФИЦ ИУ РАН (e-mail: mposypkin@gmail.com)

Си Ту Тант Син является аспирантом МИЭТ (e-mail: sithuthantsin86@gmail.com)

$$\tau(P) = \tau, C(P) = C - \sum_{i=1}^{\tau(P)} \theta_i w_i,$$

$$I_0 = \{i : 1 \leq i \leq \tau, \theta_i = 0\}, I_1 = \{i : 1 \leq i \leq \tau, \theta_i = 1\},$$

$$\underline{k}(P) = \left\{ \max k : 0 \leq k \leq n - \tau(P), \sum_{\tau(P)+1}^{\tau(P)+k} w_i \leq C(P) \right\},$$

$$\bar{k}(P) = \left\{ \max k : 0 \leq k \leq n - \tau(P), \sum_{i=n-k+1}^n w_i \leq C(P) \right\},$$

$$\Delta(P) = \bar{k}(P) - \underline{k}(P).$$

Здесь и далее сумма по пустому множеству индексов считается равной нулю. В рассматриваемом в данной работе варианте метода ветвей и границ используется *правило декомпозиции*, сопоставляющее подзадаче (2) две новых подзадачи, получаемые присваиванием переменной $x_{\tau+1}$ значений 0 и 1 соответственно. Подзадачу (2) будем называть *родительской* подзадачей для полученных в результате декомпозиции подзадач. Переменная $x_{\tau+1}$ называется *переменной ветвления* для подзадачи (2). Предполагается, что переменные ветвления выбираются в порядке убывания их весов, т.е. справедливо соотношение $w_1 \geq \dots \geq w_n$.

В целях сокращения перебора в процессе решения задачи (1) методом ветвей и границ для подзадач этой задачи, возникающих в процессе решения, осуществляется проверка определенных условий, при выполнении которых подзадачи могут быть исключены из дальнейшего рассмотрения. Такие условия называются *условиями отсева*. Условие отсева для подзадач задачи (1) называется *условием отрицательного отсева*, если подзадача, удовлетворяющая данному условию, не имеет допустимых решений, и называется *условием положительного отсева*, если для подзадачи, удовлетворяющей данному условию, оптимальное решение может быть найдено без дальнейшей декомпозиции этой подзадачи.

Пусть S_0 — некоторое условие отрицательного отсева, а S_1 — некоторое условие положительного отсева для подзадач задачи (1). Рассмотрим вариант метода ветвей и границ для задачи (1), который состоит в последовательном выполнении итераций следующего цикла (в процессе выполнения данной процедуры поддерживаются список подзадач, подлежащих дальнейшему рассмотрению, и текущее рекордное решение задачи).

1. В список подзадач помещается исходная задача (1).
2. Если список подзадач пуст, то алгоритм завершает свою работу, в противном случае выбирается и удаляется из списка одна из подзадач.
3. Если выбранная подзадача удовлетворяет условию S_0 , выполняется переход к шагу 2.
4. Если выбранная подзадача удовлетворяет условию S_1 , вычисляется оптимальное решение z для данной

подзадачи. Если значение целевой функции на решении z больше текущего рекорда, текущее рекордное решение заменяется на z . Затем выполняется переход к шагу 2.

5. Производится декомпозиция выбранной подзадачи на две подзадачи, которые добавляются в список, и выполняется переход к шагу 2.

В качестве оптимального решения задачи (1) берется рекордное решение, полученное в результате выполнения описанной процедуры. Для отличия от других вариантов метода ветвей и границ для решения задачи (1), данный алгоритм будем называть *мажоритарным* с условиями отрицательного отсева S_0 и положительного отсева S_1 .

Утверждение 1. Мажоритарный метод ветвей и границ с условиями отрицательного и положительного отсева находит оптимальное решение для любой задачи (1).

Доказательство. Прежде всего, заметим, что множество допустимых решений подзадачи совпадает с объединением множеств допустимых решений подзадач, полученных из нее в результате декомпозиции. Так как выполнение условия отрицательного отсева означает, что подзадача не имеет решения, то исключение из рассмотрения подзадач, удовлетворяющих условию отрицательного отсева, не приводит к потере допустимых решений. Таким образом, любое оптимальное решение задачи (1) является также оптимальным решением некоторой подзадачи, удовлетворяющей условию положительного отсева. Поскольку для всех подзадач, удовлетворяющих условию положительного отсева, вычисляются оптимальные решения, исключение их из рассмотрения не приводит к потере всех оптимальных решений задачи (1). □

Сложностью решения задачи (1) мажоритарным методом ветвей и границ называется число итераций цикла 2-5, выполняемых при решении задачи данным методом. Отметим, что это число не зависит от способа выбора подзадач из списка, поэтому данное понятие сложности является корректно определенным.

Процесс решения задачи (1) мажоритарным методом ветвей и границ удобно представить в виде *дерева ветвления*, вершинам которого соответствуют рассматриваемые в процессе решения подзадачи. Вершина, соответствующая подзадаче, соединяется дугами с вершинами, соответствующими подзадачам, полученным из этой подзадачи в результате декомпозиции. Дуга, ведущая в вершину, соответствующую подзадаче, полученной присваиванием переменной ветвления значения 0 (значения 1), помечается символом 0 (символом 1). Корнем дерева ветвления полагается вершина, соответствующая исходной задаче. Концевым вершинам дерева ветвления соответствуют подзадачи, удовлетворяющие условиям отсева. Отметим, что число вершин в дереве ветвления совпадает со сложностью

решения задачи методом ветвей и границ. Поскольку дерево ветвления является бинарным, число вершин в этом дереве равно $2K - 1$, где K – число конечных вершин дерева. Таким образом, сложность решения задачи (1) мажоритарным МВГ равна $2K - 1$, где K – число рассматриваемых в процессе решения задачи подзадач, удовлетворяющих условиям отсева.

В данной работе рассматривается вариант мажоритарного метода ветвей и границ, в котором применяются следующие условия отсева C_0, C_1 : подзадача P вида (2) удовлетворяет условию C_0 , если $C(P) < 0$, и удовлетворяет условию C_1 , если $\bar{k}(P) = \underline{k}(P)$. Очевидно, что подзадача, удовлетворяющая условию C_0 , не имеет допустимых решений, тем самым C_0 является условием отрицательного отсева. Для подзадачи P , удовлетворяющей условию C_1 , имеется допустимое решение z вида

$$z_i = \begin{cases} \theta_i, & \text{если } 1 \leq i \leq \tau(P), \\ 1, & \text{если } \tau(P) + 1 \leq i \leq \tau + \underline{k}(P), \\ 0, & \text{если } \tau + \underline{k}(P) + 1 \leq i \leq n. \end{cases}$$

Учитывая соотношения $w_1 \geq \dots \geq w_n$, нетрудно заметить, что данное решение является оптимальным для P , тем самым C_1 является условием положительного отсева. В работе получена верхняя оценка сложности решения задачи (1) мажоритарным МВГ с условиями отрицательного отсева C_0 и положительного отсева C_1 в зависимости от числа n переменных задачи.

III. ОЦЕНКА СЛОЖНОСТИ МВГ

Обозначим через MP множество всех подзадач задачи (1), получаемых в процессе решения задачи рассматриваемым вариантом мажоритарного МВГ.

Утверждение 2. Пусть подзадачи P_0 и P_1 получены в результате декомпозиции из подзадачи P присваиванием переменной ветвления $x_{\tau(P)+1}$ значений 0 и 1 соответственно. Тогда $\Delta(P) \geq \Delta(P_0), \Delta(P) \geq \Delta(P_1)$.

Доказательство. Если $\Delta(P) = 0$, то утверждение очевидно. Пусть $\Delta(P) > 0$, тогда $\bar{k}(P) < n - \tau(P)$.

Справедливо равенство $C(P_0) = C(P)$. Из свойства $w_1 \geq \dots \geq w_n$ следует, что $\underline{k}(P_0) \geq \underline{k}(P), \bar{k}(P_0) = \bar{k}(P)$. Следовательно, $\Delta(P_0) \leq \Delta(P)$.

Легко проверяется справедливость соотношения $C(P_1) = C(P) - w_{\tau(P)+1}$. Следовательно

$\underline{k}(P_1) = \underline{k}(P) - 1, \bar{k}(P_1) \geq \bar{k}(P) - 1$. Отсюда получаем

$$\Delta(P) \geq \Delta(P_1) . \square$$

Сопоставим каждой подзадаче P из MP , удовлетворяющей условию C_0 , бинарный набор $\sigma(P)$ длины n , определяемый следующим образом:

$$\sigma_i(P) = \begin{cases} \theta_i, & \text{если } 1 \leq i \leq \tau(P), \\ 0, & \text{если } \tau(P) + 1 \leq i \leq n. \end{cases}$$

Справедливо следующее утверждение.

Утверждение 3. Пусть подзадача P из MP удовлетворяет условию C_0 . Тогда для любого i такого, что $\sigma_i(P) = 1$, справедливо $f(\sigma(P)) - w_i \leq C$.

Доказательство. Рассмотрим подзадачу P' , из которой подзадача P получена декомпозицией. Так как P' не удовлетворяет условию C_0 , то $f(\sigma(P')) \leq C$. Наборы $\sigma(P)$ и $\sigma(P')$ отличаются только в компоненте с номером $\tau(P)$: $\sigma_{\tau(P)}(P) = 1, \sigma'_{\tau(P)}(P) = 0$. Следовательно $f(\sigma(P)) - w_{\tau(P)} = f(\sigma(P')) \leq C$. Так как выполнено $w_1 \geq \dots \geq w_{\tau(P)}$, то $f(\sigma(P)) - w_i \leq f(\sigma(P)) - w_{\tau(P)} \leq C$. \square

Из доказанного утверждения и того, что для любой подзадачи P , удовлетворяющей условию C_0 , справедливо $f(\sigma(P)) > C$, вытекает следующее утверждение.

Следствие 1. Для любых двух различных подзадач P_1 и P_2 из MP , отсеянных по условию C_0 , наборы $\sigma(P_1)$ и $\sigma(P_2)$ несравнимы.

Сопоставим каждой подзадаче P из MP , отсеянной по условию C_1 , набор $\sigma(P)$ длины n , определяемый следующим образом:

$$\sigma_i(P) = \begin{cases} \theta_i, & \text{если } 1 \leq i \leq \tau(P), \\ 1, & \text{если } \tau(P) + 1 \leq i \leq \tau + \underline{k}(P), \\ 0, & \text{если } \tau + \underline{k}(P) + 1 \leq i \leq n. \end{cases}$$

Утверждение 4. Пусть P – подзадача из MP , отсеянная по условию C_1 . Тогда для любого i такого, что $\sigma_i(P) = 0$, выполняется $C < f(\sigma(P)) + w_i$.

Доказательство. Рассмотрим два возможных случая: $\underline{k}(P) = n - \tau(P)$ и $\underline{k}(P) < n - \tau(P)$. Пусть $\underline{k}(P) = n - \tau(P)$. Тогда $i \leq \tau(P)$. Рассмотрим родительскую подзадачу P' для подзадачи P . Так как подзадача P' не удовлетворяет условию отсева C_1 , то

$$\theta_{\tau(P)} = 0 \quad \text{и} \quad \sum_{j=1}^{\tau(P)-1} \theta_j w_j + \sum_{j=\tau(P)}^n w_j > C. \quad \text{Так как}$$

$$\sum_{j=1}^{\tau(P)-1} \theta_j w_j + \sum_{j=\tau(P)}^n w_j = f(\sigma(P)) + w_{\tau(P)} \quad \text{и} \quad w_i \geq w_{\tau(P)},$$

то $f(\sigma(P)) + w_i > C$.

Пусть $\underline{k}(P) < n - \tau(P)$. Из определения $\underline{k}(P)$ следует, что $f(\sigma(P)) + w_{\tau(P)+\underline{k}(P)+1} > C$. В силу

неравенства $w_1 \geq \dots \geq w_n$ получаем $f(\sigma(P)) + w_i > C$ для любого $i \leq \tau(P) + \underline{k}(P) + 1$.

Пусть $i > \tau(P) + \underline{k}(P) + 1$ и $\sigma_i(P) = 0$. Тогда

$$f(\sigma(P)) + w_i = \sum_{j=1}^{\tau(P)} \sigma_j(P) w_j + \sum_{j=\tau(P)+1}^{\tau(P)+\underline{k}(P)} \sigma_j(P) w_j + w_i.$$

Так как подзадача P удовлетворяет условию C1, то $\bar{k}(P) = \underline{k}(P)$. Поэтому

$$f(\sigma(P)) + w_i = \sum_{j=1}^{\tau(P)} \sigma_j(P) w_j + \sum_{j=\tau(P)+1}^{\tau(P)+\bar{k}(P)} \sigma_j(P) w_j + w_i,$$

Таким образом, из неравенства $w_1 \geq \dots \geq w_n$ следует, что

$$f(\sigma(P)) + w_i \geq \sum_{j=1}^{\tau(P)} \sigma_j(P) w_j + \sum_{j=n-\bar{k}(P)}^n \sigma_j(P) w_j.$$

Согласно определению $\bar{k}(P)$ отсюда получаем, что

$$f(\sigma(P)) + w_i > C. \quad \text{Тем самым утверждение доказано.}$$

□

Из этого утверждения вытекает

Следствие 2. Для любых двух различных подзадач P_1 и P_2 из MP , отсеянных по условию C1, наборы $\sigma(P_1)$ и $\sigma(P_2)$ несравнимы.

Утверждение 5. Пусть подзадачи P_0 и P_1 из MP удовлетворяют условиям C0 и C1 соответственно. Тогда наборы $\sigma(P_0)$ и $\sigma(P_1)$, сопоставленные этим подзадачам, несравнимы.

Доказательство. Так как $f(\sigma(P_0)) > C$ и $f(\sigma(P_1)) \leq C$, то в силу монотонности функции f невозможно выполнение неравенства $\sigma(P_1) \geq \sigma(P_0)$.

Покажем, что неравенство $\sigma(P_0) \geq \sigma(P_1)$ также не верно. Предположим обратное: пусть $\sigma(P_0) \geq \sigma(P_1)$.

Тогда найдется такое i , $1 \leq i \leq \tau(P_0)$, что $\sigma_i(P_0) = 1, \sigma_i(P_1) = 0$. Отсюда следует неравенство

$$f(\sigma(P_1)) \leq f(\sigma(P_0)) - w_i. \quad (3)$$

Рассмотрим родительскую подзадачу P'_0 для подзадачи P_0 . Так как P'_0 не удовлетворяет условию C1, то

$$\bar{k}(P'_0) > \underline{k}(P'_0) \geq 0. \quad \text{Так как} \quad \bar{k}(P'_0) > 0, \quad \text{то}$$

$$f(P'_0) + w_n \leq C \quad \text{и} \quad \tau(P'_0) \leq n - 1. \quad \text{Поэтому}$$

$$\sigma_n(P'_0) = 0 \quad \text{и, следовательно,} \quad \sigma_n(P_1) = 0. \quad \text{Так как}$$

$$f(P'_0) = f(P_0) - w_{\tau(P_0)}, \quad \text{то} \quad f(P_0) - w_{\tau(P_0)} + w_n \leq C.$$

Поскольку $1 \leq i \leq \tau(P_0)$, то $f(P_0) - w_i + w_n \leq f(P_0) - w_{\tau(P_0)} + w_n$. Следовательно $f(\sigma(P_1)) + w_n \leq C$. Применяя (3), получим

$$f(\sigma(P_1)) + w_n \leq C. \quad (4)$$

Так как $\sigma_n(P_1) = 0$, то, согласно утверждению 4, $f(\sigma(P_1)) + w_n > C$, тем самым приходим к противоречию с соотношением (4). □

Следствие 3. Для любых двух различных подзадач P_1 и P_2 из MP , отсеянных по условиям C0 или C1, наборы $\sigma(P_1)$ и $\sigma(P_2)$ несравнимы.

Из Следствия 3 получаем, что, наборы, сопоставленные подзадачам из MP , удовлетворяющим условиям отсева, попарно несравнимы, т.е. образуют антицепь булева куба размерности n . Следовательно, (см., например, [3]), общее число таких наборов не может

превосходить $\binom{n}{\lfloor n/2 \rfloor}$, тем самым число подзадач из

MP , удовлетворяющих условиям отсева, не превосходит $\binom{n}{\lfloor n/2 \rfloor}$. Поэтому сложность решения

задачи (1) рассматриваемым вариантом мажоритарного МВГ не превосходит $2^{\binom{n}{\lfloor n/2 \rfloor}} - 1$. Таким образом

справедлива следующая теорема.

Теорема 1. Сложность решения задачи (1) мажоритарным методом ветвей и границ с условиями

отсева C0, C1 не превосходит $2^{\binom{n}{\lfloor n/2 \rfloor}} - 1$.

Пользуясь известной асимптотикой для $\binom{n}{\lfloor n/2 \rfloor}$ (см., например, [5]), получаем, что асимптотика роста рассматриваемой сложности не превосходит

$$2^{\binom{n}{\lfloor n/2 \rfloor}} - 1 \sim \frac{2^{n+1}}{\sqrt{\pi \frac{n}{2}}}.$$

- [7] Лазарев А. А. Графический подход к решению задач комбинаторной оптимизации // Автоматика и телемеханика. – 2007. – №. 4. – С. 13-23.

Ранее в работе [4] рассмотрен вариант мажоритарного МВГ, в котором используется условие положительного отсева, являющееся частным случаем условия С1, а именно, подзадача P задачи (1) удовлетворяет условию

положительного отсева, если $\sum_{i=\tau(P)+1}^n w_i \leq C(P)$.

Показано, что сложность решения задачи (1) таким вариантом МВГ в наихудшем случае равна

$$2^{\binom{n+1}{\lfloor n/2 \rfloor + 1}} - 1.$$

Впервые пример такой задачи

приведен в [6]. Пользуясь асимптотикой для $\binom{n}{\lfloor n/2 \rfloor}$,

получим, что

$$2^{\binom{n+1}{\lfloor n/2 \rfloor + 1}} - 1 \sim \frac{2^{n+2}}{\sqrt{\pi \frac{n}{2}}}.$$

Таким образом, применение введенного нами условия С1 положительного отсева приблизительно в два раза сокращает сложность решения задачи о сумме подмножеств мажоритарным МВГ в наихудшем случае.

IV. ЗАКЛЮЧЕНИЕ

В работе развивается методика оценки сложности метода ветвей и границ для задач булева программирования. Получена верхняя оценка сложности усовершенствованного варианта метода ветвей и границ для решения одного из частных случаев задачи о ранце – задаче о сумме подмножеств. Произведен асимптотический анализ полученной оценки.

В дальнейшем планируется рассмотреть более эффективные варианты метода ветвей и границ, когда для сокращения перебора применяются также оценки, основанные на использовании отношения доминирования подзадач [7].

БИБЛИОГРАФИЯ

- [1] Сигал И.Х., Иванова А.П. Введение в прикладное дискретное программирование. М.: Физматлит, 2003. – 240 с.
- [2] Kellerer H., Pfershy U., Pisinger D. Knapsack Problems. Springer Verlag, 546 p., 2004.
- [3] Таранников Ю.В. Комбинаторные свойства дискретных структур и приложения к криптологии, Москва: Издательство МЦНМО, 2011, — 152 с..
- [4] Колпаков Р. М., Посыпкин М. А. Верхняя и нижняя оценки трудоемкости метода ветвей и границ для задачи о ранце // Дискретная математика. – 2010. – Т. 22. – №. 1. – С. 58-73.
- [5] Яблонский С.В. Введение в дискретную математику. М: Наука, 2003.
- [6] Финкельштейн Ю.Ю. Приближенные методы и прикладные задачи дискретного программирования. М.: Наука, 1976

Upper bound for the complexity of one of the variants of the branch and bound method for the subset sum problem

R.M. Kolpakov, M.A. Posypkin, Si Thu Thant Sin

Abstract – In the paper we obtain an upper bound for the complexity of solving the subset sum problem which is a particular case of the knapsack problem by one of the variants of the branch-and-bound method with an additional pruning rule based on the comparison of the maximum and minimum number of items that can be put into the knapsack. As auxiliary results, we establish various combinatorial properties of subproblems processed for solving the subset sum problem by the considered variant of the branch-and-bound method.

Keywords – the subset sum problem, complexity evaluation, combinatorial analysis, Boolean cube.