ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—февраль 2008. Том 15, № 1. 44–57

УДК 519.8

ОЦЕНКИ ПОГРЕШНОСТИ ЖАДНЫХ АЛГОРИТМОВ ДЛЯ ЗАДАЧ НА НАСЛЕДСТВЕННЫХ СИСТЕМАХ

В. П. Ильев

Исследуются задачи максимизации и минимизации аддитивных функций на наследственных системах, которые являются обобщениями многих сложных в вычислительном отношении задач комбинаторной оптимизации. Доказана оценка погрешности жадного алгоритма в терминах параметров допустимой области и целевой функции задачи максимизации, уточняющая известную оценку Дженкинса–Корте–Хаусмана. Аналогичный результат получен для задачи минимизации аддитивной функции на наследственной системе.

Введение

Пусть U — конечное множество и $\mathcal{A} \subset 2^U$ — непустое семейство его подмножеств, удовлетворяющее следующей *аксиоме наследственности*:

$$(A \in \mathcal{A},\ A' \subseteq A) \Rightarrow A' \in \mathcal{A}.$$

Пару $\mathcal{S} = (U, \mathcal{A})$ будем называть *наследственной системой* на U. Семейство $\mathcal{A}$ называется также *наследственным семейством* или *системой независимости* [2]. Множества семейства $\mathcal{A}$ называются *независимыми*, а все остальные подмножества множества U — *зависимыми*. Семейство всех зависимых множеств обозначим через $\mathcal{D}$. Очевидно, что $\mathcal{D}$ обладает свойством *наследственности «вверх»*:

$$(D \in \mathcal{D},\ D \subseteq D') \Rightarrow D' \in \mathcal{D}.$$

Поскольку каждое из семейств $\mathcal{A}$, $\mathcal{D}$ однозначно определяет наследственную систему $\mathcal{S}$, будем записывать $\mathcal{S} = (U, \mathcal{A})$ или $\mathcal{S} = (U, \mathcal{D})$ в зависимости от того, какая сторона наследственной системы будет нас интересовать.

Базами системы $\mathcal{S}$ называются максимальные по включению независимые множества, а *циклами* — минимальные по включению зависимые множества.

Объектом нашего исследования будут алгоритмы приближённого решения следующих двух задач комбинаторной оптимизации: найти

$$\max\{f(X) \mid X \in \mathcal{B}\}, \tag{1}$$

$$\min\{f(X) \mid X \in \mathcal{C}\}, \tag{2}$$

где $\mathcal{B}$ — семейство всех баз, $\mathcal{C}$ — семейство всех циклов некоторой наследственной системы $\mathcal{S} = (U, \mathcal{A}) = (U, \mathcal{D})$, а $f : 2^U \to R_+$ — неотрицательная аддитивная функция. Всюду далее для $u \in U$ вместо $f(\{u\})$ будем записывать $f(u)$.

Задачи комбинаторной оптимизации на наследственных системах и их частных случаях — матроидах и коматроидах — являются обобщениями очень многих сложных в вычислительном отношении практически важных задач, таких как задача о рюкзаке, задача о максимальном независимом множестве вершин графа, задача о покрытии множества, задача о минимальном k-связном остовном подграфе и другие. В большинстве случаев оптимизационные задачи на наследственных системах являются NP-трудными.

В качестве приближённого метода решения задачи (1) рассмотрим следующий простой алгоритм.

Алгоритм GA (*жадный алгоритм*)

Шаг 0. Упорядочить множество $U = \{u_1, \dots, u_n\}$ по невозрастанию весов, $X \leftarrow \varnothing$, перейти на шаг 1.

Шаг i ($i = 1, \dots, n$). Если $X \cup \{u_i\} \in \mathcal{A}$, то $X \leftarrow X \cup \{u_i\}$. Если $i < n$, то перейти на шаг $i + 1$, иначе $S_{GA} \leftarrow X$.

Конец.

Для приближённого решения задачи (2) будем применять следующий «обратный» аналог жадного алгоритма.

Алгоритм GR (*обратный жадный алгоритм*)

Шаг 0. Упорядочить множество $U = \{u_1, \dots, u_n\}$ по невозрастанию весов, $X \leftarrow U$, перейти на шаг 1.

Шаг i ($i = 1, \dots, n$). Если $X \setminus \{u_i\} \in \mathcal{D}$, то $X \leftarrow X \setminus \{u_i\}$. Если $i < n$, то перейти на шаг $i + 1$, иначе $S_{GR} \leftarrow X$.

Конец.

Заметим, что алгоритм GA всегда находит базу, а алгоритм GR — цикл наследственной системы, т. е. множество S_{GA} является допустимым решением задачи (1), а S_{GR} — допустимым решением задачи (2).

В настоящей статье получены оценки погрешности алгоритмов GA и GR в терминах параметров допустимых областей и целевых функций задач (1) и (2).

1. Известные результаты

Рассмотрим два важных частных случая наследственных систем.

Пусть $\mathcal{S} = (U, \mathcal{A}) = (U, \mathcal{D})$ — произвольная наследственная система и $W \subseteq U$. *Базой множества* W называется любое максимальное по включению независимое множество, содержащееся в W. *Циклом множества* W назовём любое минимальное по включению зависимое множество, содержащее W.

Наследственная система $\mathcal{S}$ называется *матроидом*, если все базы любого множества $W \subseteq U$ имеют одинаковую мощность. В частности, все базы множества U равномощны, их общая мощность r называется *рангом* матроида.

С каждой наследственной системой $\mathcal{S} = (U, \mathcal{A})$ тесно связана *дополнительная система* или *косистема* $\mathcal{S}' = (U, \mathcal{D}')$, семейство зависимых множеств которой определяется как $\mathcal{D}' = \{U \setminus A \mid \ A \in \mathcal{A}\}$. Очевидно, семейства независимых множеств, баз и циклов системы $\mathcal{S}'$ могут быть заданы как $\mathcal{A}' = \{U \setminus D \mid \ D \in \mathcal{D}\}$, $\mathcal{B}' = \{U \setminus C \mid \ C \in \mathcal{C}\}$, $\mathcal{C}' = \{U \setminus B \mid \ B \in \mathcal{B}\}$ соответственно. Ясно, что $(\mathcal{S}')' = \mathcal{S}$, т. е. наследственные системы $\mathcal{S}$ и $\mathcal{S}'$ взаимно дополнительны.

Наследственную систему, дополнительную к матроиду, будем называть *коматроидом*. Несложно показать,что система $\mathcal{S}$ является коматроидом, если и только если все циклы любого $W \subseteq U$ имеют одинаковую мощность. В частности, все циклы пустого множества равномощны, их общая мощность p называется *обхватом* коматроида.

Пример 1. Наследственная система графа и дополнительная к ней.

Дан граф $G = (V, E)$. Множество вершин $A \subseteq V$ называется *независимым*, если $uv \notin E$ для любых $u, v \in A$. Семейство всех независимых множеств вершин графа G обозначим через $\mathcal{A}_G$. Тогда $\mathcal{S}_G = (V, \mathcal{A}_G)$ — наследственная система, она называется *наследственной системой графа* G. Циклы этой системы взаимно однозначно соответствуют рёбрам графа G.

Несложно понять, что в системе $\mathcal{S}'_G = (V, \mathcal{D}'_G)$, дополнительной к наследственной системе графа, $\mathcal{D}'_G = \{V \setminus A \mid A \in \mathcal{A}_G\}$ — семейство всех вершинных покрытий в графе G.

Одним из центральных результатов теории матроидов является известная теорема Радо–Эдмонса [1, 6], в которой утверждается, что алгоритм GA находит оптимальное решение задачи (1) на произвольной наследственной системе $\mathcal{S}$ для любой аддитивной целевой функции тогда и только тогда, когда $\mathcal{S}$ — матроид.

Для задачи (2) справедлива теорема, аналогичная теореме Радо–Эд-

мондса. Как показано в [3], алгоритм GR находит оптимальное решение задачи (2) на произвольной наследственонй системе $\mathcal{S}$ для любой аддитивной целевой функции тогда и только тогда, когда $\mathcal{S}$ — коматроид.

Как следует из теоремы Радо–Эдмондса, если наследственная система отлична от матроида, то жадный алгоритм в общем случае может не найти оптимальное решение задачи (1) с аддитивной целевой функцией. В [4, 5] получена оценка погрешности алгоритма GA для задачи максимизации аддитивной функции на наследственной системе. Сформулируем этот результат.

Для любой наследственной системы $\mathcal{S} = (U, \mathcal{A}) = (U, \mathcal{D})$ рассмотрим величины

$$c(\mathcal{A}) = \min_{\substack{W \subseteq U, \\ W \notin \mathcal{A}}} \frac{r_{\min}(W)}{r_{\max}(W)}, \quad c(\mathcal{D}) = \max_{\substack{W \subseteq U, \\ W \notin \mathcal{D}}} \frac{g_{\max}(W) - |W|}{g_{\min}(W) - |W|},$$

где $r_{\min}(W)$ и $r_{\max}(W)$ — минимальная и максимальная мощности баз множества W, а $g_{\min}(W)$ и $g_{\max}(W)$ — минимальная и максимальная мощности циклов множества W соответственно.

Параметры $c(\mathcal{A})$ и $c(\mathcal{D})$ характеризуют близость наследственной системы $\mathcal{S} = (U, \mathcal{A}) = (U, \mathcal{D})$ к матроиду и коматроиду соответственно. Очевидно, что $c(\mathcal{A}) \leqslant 1 \leqslant c(\mathcal{D})$ для любой системы, причём $c(\mathcal{A}) = 1$ тогда и только тогда, когда система $\mathcal{S}$ является матроидом, а $c(\mathcal{D}) = 1$ тогда и только тогда, когда система $\mathcal{S}$ является коматроидом.

Теорема 1 [4, 5]. *Пусть $\mathcal{S} = (U, \mathcal{A})$ — произвольная наследственная система. Тогда для любой аддитивной целевой функции задачи максимизации* (1) *имеет место оценка*

$$\frac{f(S_{GA})}{f(S_O)} \geqslant c(\mathcal{A}), \tag{3}$$

где S_O — оптимальное решение задачи (1).

Кроме того, в [4, 5] показано, что оценка (3) точна в следующем смысле: для любой наследственной системы $\mathcal{S} = (U, \mathcal{A})$ существует такая аддитивная функция $f : 2^U \to R_+$, что

$$\frac{f(S_{GA})}{f(S_O)} = c(\mathcal{A}).$$

В [3] доказана аналогичная оценка погрешности алгоритма GR для задачи (2) минимизации аддитивной функции на наследственной системе.

Теорема 2 [3]. *Пусть $\mathcal{S} = (U, \mathcal{D})$ — произвольная наследственная система. Тогда для любой аддитивной целевой функции задачи минимизации* (2) *имеет место оценка*

$$\frac{f(S_{GR})}{f(S_O)} \leqslant c(\mathcal{D}), \tag{4}$$

где S_O — оптимальное решение задачи (2).

В [3] также доказано, что оценка (4) точна в том смысле, что для любой наследственной системы $\mathcal{S} = (U, \mathcal{D})$ существует такая аддитивная функция $f : 2^U \to R_+$, что

$$\frac{f(S_{GR})}{f(S_O)} = c(\mathcal{D}).$$

2. Уточнение оценки Дженкинса–Корте–Хаусмана для задачи о максимальном независимом множестве

Привлечение дополнительной информации о целевой функции позволяет получить оценку погрешности алгоритма GA для задачи (1), которая в ряде случаев оказывается существенно более точной, чем (3).

Обозначим через $\mathcal{F}(a, b)$ класс всех аддитивных функций $f : 2^U \to R_+$ таких, что $f(u) \in [a, b]$ для любого $u \in U$, где $a \geqslant 0$.

Введём следующую характеристику задачи (1) с целевой функцией $f \in \mathcal{F}(a, b)$:

$$c(\mathcal{A}, a, b) = \min_{\substack{W \subseteq U, \\ W \notin \mathcal{A}}} \frac{(b-a) r_{\min}(W) + a r_{\min}(U)}{(b-a) r_{\max}(W) + a r_{\max}(U)} .$$

Нетрудно видеть, что при $a = 0$ новая характеристика совпадает с $c(\mathcal{A})$.

В этом разделе будет доказано следующее уточнение оценки (3).

Теорема 3. *Пусть $\mathcal{S} = (U, \mathcal{A})$ — произвольная наследственная система. Тогда для любой аддитивной целевой функции $f \in \mathcal{F}(a, b)$ задачи максимизации* (1) *имеет место оценка*

$$\frac{f(S_{GA})}{f(S_O)} \geqslant c(\mathcal{A}, a, b), \tag{5}$$

где S_O — оптимальное решение задачи (1).

Для доказательства теоремы 3 потребуется вспомогательное утверждение и следующие обозначения. Пусть $S_{GA} = \{v_1, \ldots, v_m\}$,

$f(v_i) = w_i \in [a,b]$, причём $w_1 \geqslant \ldots \geqslant w_m$. Положим $U_0 = \varnothing$, а при $i = 1, \ldots, m$ положим

$$U_i = \{v_1, \ldots, v_i\} \cup \{u \in U \mid \ \{v_1, \ldots, v_i, u\} \in \mathcal{D}\}.$$

Тогда $S_{GA} \cap U_i = \{v_1, \ldots, v_i\}$. Заметим, что $U_m = U$. Обозначим $d_i = |S_O \cap (U_i \setminus U_{i-1})|$, $i = 1, \ldots, m$. Следовательно, $|S_O| = d_1 + \ldots + d_m$.

Лемма 1. *Пусть* $b \geqslant w_1 \geqslant \ldots \geqslant w_m \geqslant a$ *и*

$$\frac{bk + a(m-k)}{b\sum\limits_{j=1}^{k} d_j + a\sum\limits_{j=k+1}^{m} d_j} = \min_{0 \leqslant i \leqslant m} \frac{bi + a(m-i)}{b\sum\limits_{j=1}^{i} d_j + a\sum\limits_{j=i+1}^{m} d_j} \tag{6}$$

(при $i = 0$ *и* $i = m$ *дробь в правой части условия* (6) *равна* $\frac{m}{d_1+\ldots+d_m}$*). Тогда*

$$\frac{w_1 + \ldots + w_m}{w_1 d_1 + \ldots + w_m d_m} - \frac{bk + a(m-k)}{b\sum\limits_{j=1}^{k} d_j + a\sum\limits_{j=k+1}^{m} d_j} \geqslant 0. \tag{7}$$

Доказательство. Из (6) следует, что при любом $i = 0, 1, \ldots, k-1$

$$\frac{bk + a(m-k)}{b\sum\limits_{j=1}^{k} d_j + a\sum\limits_{j=k+1}^{m} d_j} \leqslant \frac{bk + a(m-k) - b(k-i) + a(k-i)}{b\sum\limits_{j=1}^{k} d_j + a\sum\limits_{j=k+1}^{m} d_j - b\sum\limits_{j=i+1}^{k} d_j + a\sum\limits_{j=i+1}^{k} d_j}\ .$$

Поэтому

$$\begin{gathered}
b^2 k\sum_{j=1}^{k} d_j + abk\sum_{j=k+1}^{m} d_j - b^2 k\sum_{j=i+1}^{k} d_j + abk\sum_{j=i+1}^{k} d_j + ab(m-k)\sum_{j=1}^{k} d_j \\
+ a^2(m-k)\sum_{j=k+1}^{m} d_j - ab(m-k)\sum_{j=i+1}^{k} d_j + a^2(m-k)\sum_{j=i+1}^{k} d_j \\
- b^2 k\sum_{j=1}^{k} d_j - abk\sum_{j=k+1}^{m} d_j - ab(m-k)\sum_{j=1}^{k} d_j - a^2(m-k)\sum_{j=k+1}^{m} d_j \\
+ b^2(k-i)\sum_{j=1}^{k} d_j + ab(k-i)\sum_{j=k+1}^{m} d_j - ab(k-i)\sum_{j=1}^{k} d_j - a^2(k-i)\sum_{j=k+1}^{m} d_j \leqslant 0.
\end{gathered}$$

Следовательно,

$$b^2((k-i)\sum_{j=1}^{k} d_j - k\sum_{j=i+1}^{k} d_j) - ab((k-i)\sum_{j=1}^{k} d_j - k\sum_{j=i+1}^{k} d_j)$$

$$+ ab((k-i)\sum_{j=k+1}^{m} d_j - (m-k)\sum_{j=i+1}^{k} d_j) - a^2((k-i)\sum_{j=k+1}^{m} d_j - (m-k)\sum_{j=i+1}^{k} d_j)$$

$$= (b-a)[b((k-i)\sum_{j=1}^{k} d_j - k\sum_{j=i+1}^{k} d_j) + a((k-i)\sum_{j=k+1}^{m} d_j - (m-k)\sum_{j=i+1}^{k} d_j)] \leqslant 0.$$

Введём обозначение $s_j = b(\sum_{j=1}^{k} d_j - kd_j) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_j)$, $j = 1, \ldots, k$. Тогда последнее неравенство примет вид

$$(b-a)(s_{i+1} + \ldots + s_k) \leqslant 0$$

при любом $i = 0, 1, \ldots, k-1$.

Отсюда с учётом неравенства $b - a \geqslant 0$ получаем

$$s_k \leqslant 0,\ s_{k-1} + s_k \leqslant 0,\ \ldots,\ s_1 + \ldots + s_k \leqslant 0. \tag{8}$$

Далее из (6) следует, что при любом $i = k+1, \ldots, m$

$$\frac{bk + a(m-k)}{b\sum_{j=1}^{k} d_j + a\sum_{j=k+1}^{m} d_j} \leqslant \frac{bk + a(m-k) + b(i-k) - a(i-k)}{b\sum_{j=1}^{k} d_j + a\sum_{j=k+1}^{m} d_j + b\sum_{j=k+1}^{i} d_j - a\sum_{j=k+1}^{i} d_j}.$$

Отсюда путем аналогичных выкладок получаем

$$(a-b)(s_{k+1} + \ldots + s_i) \leqslant 0 \text{ для любого } i = k+1, \ldots, m,$$

где $s_j = b(\sum_{j=1}^{k} d_j - kd_j) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_j)$, $j = k+1, \ldots, m$.

Отсюда и из неравенства $a - b \leqslant 0$ следует, что

$$s_{k+1} \geqslant 0,\ s_{k+1} + s_{k+2} \geqslant 0,\ \ldots,\ s_{k+1} + \ldots + s_m \geqslant 0. \tag{9}$$

После приведения к общему знаменателю числитель дроби в левой

части неравенства (7) будет иметь вид

$$
\begin{aligned}
& w_1 b \sum_{j=1}^{k} d_j + \ldots + w_m b \sum_{j=1}^{k} d_j + w_1 a \sum_{j=k+1}^{m} d_j + \ldots + w_m a \sum_{j=k+1}^{m} d_j \\
& - w_1 bkd_1 - \ldots - w_m bkd_m - w_1 a(m-k)d_1 - \ldots - w_m a(m-k)d_m \\
& = w_1 b(\sum_{j=1}^{k} d_j - kd_1) + \ldots + w_m b(\sum_{j=1}^{k} d_j - kd_m) \\
& + w_1 a(\sum_{j=k+1}^{m} d_j - (m-k)d_1) + \ldots + w_m a(\sum_{j=k+1}^{m} d_j - (m-k)d_m) \\
& = w_1(b(\sum_{j=1}^{k} d_j - kd_1) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_1)) + \ldots \\
& + w_m(b(\sum_{j=1}^{k} d_j - kd_m) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_m)) \\
& = w_1 s_1 + \ldots + w_m s_m = w_1 s_1 + \ldots + w_k s_k + w_{k+1} s_{k+1} + \ldots + w_m s_m \\
& = w_1[s_1 + \ldots + s_k - (s_2 + \ldots + s_k)] + \ldots + w_{k-1}(s_{k-1} + s_k - s_k) \\
& + w_k s_k + w_{k+1} s_{k+1} + w_{k+2}(s_{k+1} + s_{k+2} - s_{k+1}) + \ldots \\
& + w_m[s_{k+1} + \ldots + s_m - (s_{k+1} + \ldots + s_{m-1})] \\
& = w_1(s_1 + \ldots + s_k) + (w_2 - w_1)(s_2 + \ldots + s_k) + \ldots + (w_k - w_{k-1})s_k
\end{aligned}
$$

$$
\begin{aligned}
+ (w_{k+1} - w_{k+2})s_{k+1} + \ldots + (w_{m-1} - w_m)(s_{k+1} + \ldots + s_{m-1}) \\
+ w_m(s_{k+1} + \ldots + s_m).
\end{aligned}
$$

Учитывая (8), (9) и то, что $w_1 \leqslant b$, $w_m \geqslant a$, $w_2 - w_1 \leqslant 0, \ldots,$ $w_k - w_{k-1} \leqslant 0$, $w_{k+1} - w_{k+2} \geqslant 0, \ldots, w_{m-1} - w_m \geqslant 0$, последнюю сумму можно оценить снизу величиной $b(s_1 + \ldots + s_k) + a(s_{k+1} + \ldots + s_m)$.

Докажем, что $b(s_1 + \ldots + s_k) + a(s_{k+1} + \ldots + s_m) = 0$.

Действительно,

$$
\begin{aligned}
b[b(\sum_{j=1}^{k} d_j - kd_1) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_1) + \ldots \\
+ b(\sum_{j=1}^{k} d_j - kd_k) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_k)]
\end{aligned}
$$

$$+ a[b(\sum_{j=1}^{k} d_j - kd_{k+1}) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_{k+1}) + \ldots$$

$$+ b(\sum_{j=1}^{k} d_j - kd_m) + a(\sum_{j=k+1}^{m} d_j - (m-k)d_m)]$$

$$= b^2 \sum_{j=1}^{k} d_j - b^2 kd_1 + ab \sum_{j=k+1}^{m} d_j - ab(m-k)d_1 + \ldots$$

$$+ b^2 \sum_{j=1}^{k} d_j - b^2 kd_k + ab \sum_{j=k+1}^{m} d_j - ab(m-k)d_k$$

$$+ ab \sum_{j=1}^{k} d_j - abkd_{k+1} + a^2 \sum_{j=k+1}^{m} d_j - a^2(m-k)d_{k+1} + \ldots$$

$$+ ab \sum_{j=1}^{k} d_j - abkd_m + a^2 \sum_{j=k+1}^{m} d_j - a^2(m-k)d_m$$

$$= b^2(k \sum_{j=1}^{k} d_j - k \sum_{j=1}^{k} d_j) + a^2((m-k) \sum_{j=k+1}^{m} d_j - (m-k) \sum_{j=k+1}^{m} d_j)$$

$$+ ab(k \sum_{j=k+1}^{m} d_j - (m-k) \sum_{j=1}^{k} d_j + (m-k) \sum_{j=1}^{k} d_j - k \sum_{j=k+1}^{m} d_j) = 0.$$

Лемма 1 доказана.

Доказательство теоремы 3. Имеем

$$\frac{f(S_{GA})}{f(S_O)} = \frac{\sum\limits_{i=1}^{m} f(v_i)}{\sum\limits_{i=1}^{m} \left(\sum\limits_{u \in S_O \cap (U_i \setminus U_{i-1})} f(u) \right)} \geqslant \frac{\sum\limits_{i=1}^{m} w_i}{\sum\limits_{i=1}^{m} w_i d_i} \ . \tag{10}$$

Последнее неравенство следует из того, что $f(u) \leqslant f(v_i) = w_i$ для любого элемента $u \in U_i \setminus U_{i-1}$, так как алгоритм GA выбрал v_i.

Из неравенства (10) и леммы 1 следует, что

$$\frac{f(S_{GA})}{f(S_O)} \geqslant \frac{bk + a(m-k)}{b \sum\limits_{j=1}^{k} d_j + a \sum\limits_{j=k+1}^{m} d_j} \ . \tag{11}$$

Поскольку $k = |S_{GA} \cap U_k|$, $m = |S_{GA}|$, $\sum\limits_{j=1}^{k} d_j = |S_O \cap U_k|$, а $\sum\limits_{j=1}^{m} d_j = |S_O|$, неравенство (11) может быть переписано в виде

$$\frac{f(S_{GA})}{f(S_O)} \geqslant \frac{(b-a)|S_{GA} \cap U_k| + a|S_{GA}|}{(b-a)|S_O \cap U_k| + a|S_O|}\ .$$

Осталось заметить, что S_{GA} — база множества U и, как следует из определения множества U_k, $S_{GA} \cap U_k$ — база множества U_k. Следовательно, $|S_{GA}| \geqslant r_{\min}(U)$ и $|S_{GA} \cap U_k| \geqslant r_{\min}(U_k)$. Кроме того, $|S_O| \leqslant r_{\max}(U)$, а поскольку множество $S_O \cap U_k$ независимо, то $|S_O \cap U_k| \leqslant r_{\max}(U_k)$. Поэтому

$$\begin{aligned}\frac{f(S_{GA})}{f(S_O)} &\geqslant \frac{(b-a)r_{\min}(U_k) + ar_{\min}(U)}{(b-a)r_{\max}(U_k) + ar_{\max}(U)}\\ &\geqslant \min_{0 \leqslant i \leqslant m} \frac{(b-a)r_{\min}(U_i) + ar_{\min}(U)}{(b-a)r_{\max}(U_i) + ar_{\max}(U)}\\ &\geqslant \min_{\substack{W \subseteq U,\\ W \notin \mathcal{A}}} \frac{(b-a)r_{\min}(W) + ar_{\min}(U)}{(b-a)r_{\max}(W) + ar_{\max}(U)} = c(\mathcal{A}, a, b).\end{aligned}$$

Теорема 3 доказана.

Замечание 1. Для любой наследственной системы $\mathcal{S} = (U, \mathcal{A})$ и любых a, b $(0 \leqslant a \leqslant b)$ справедливо неравенство $c(\mathcal{A}, a, b) \geqslant c(\mathcal{A})$.

Действительно, по определению величины $c(\mathcal{A})$ для любого множества $W \subseteq U$ $(W \notin \mathcal{A})$ выполняется неравенство $r_{\min}(W) \geqslant c(\mathcal{A})r_{\max}(W)$. Пусть $W^* \notin \mathcal{A}$ — множество, на котором достигает минимума величина

$$\frac{(b-a)r_{\min}(W) + ar_{\min}(U)}{(b-a)r_{\max}(W) + ar_{\max}(U)}.$$

Тогда

$$\begin{aligned}c(\mathcal{A}, a, b) &= \frac{(b-a)r_{\min}(W^*) + ar_{\min}(U)}{(b-a)r_{\max}(W^*) + ar_{\max}(U)}\\ &\geqslant \frac{(b-a)c(\mathcal{A})r_{\max}(W^*) + ac(\mathcal{A})r_{\max}(U)}{(b-a)r_{\max}(W^*) + ar_{\max}(U)} = c(\mathcal{A}).\end{aligned}$$

Замечание 2. Величина $c(\mathcal{A}, a, b)$ может быть сколь угодно близка к 1, в то время как $c(\mathcal{A})$ близка к 0.

Рассмотрим задачу (1) на наследственной системе $\mathcal{S}_G = (V, \mathcal{A}_G)$ графа $G = K_{1,p} \cup O_q$, состоящего из звезды с p лучами и q изолированных

вершин. Пусть каждая изолированная вершина u имеет вес $f(u) = a$, а все остальные вершины графа имеют вес b, где $b \geqslant a > 0$. Множество V вершин графа G упорядочено таким образом, что u_1 — вершина степени p, $u_2, \ldots, u_{p+1}$ — вершины степени 1, $u_{p+2}, \ldots, u_{p+q+1}$ — изолированные вершины.

Здесь $c(\mathcal{A}) = \dfrac{r_{\min}(W^*)}{r_{\max}(W^*)} = \dfrac{1}{p}$, где $W^* = \{u_1, u_2, \ldots, u_{p+1}\}$, а

$$c(\mathcal{A}, a, b) = \frac{(b-a)r_{\min}(W^*) + ar_{\min}(V)}{(b-a)r_{\max}(W^*) + ar_{\max}(V)} = \frac{(b-a) + a(1+q)}{(b-a)p + a(p+q)} = \frac{b + aq}{bp + aq} = \frac{1 + \frac{aq}{b}}{p + \frac{aq}{b}}.$$

Очевидно, что при достаточно большом p величина $c(\mathcal{A})$ близка к 0, а $c(\mathcal{A}, a, b)$ близка к 1 при достаточно большом значении $(aq)/b$.

Замечание 3. Оценка (5) достижима.

Оценка (5) достигается, если наследственная система $\mathcal{S} = (U, \mathcal{A})$ является матроидом, так как в этом случае $c(\mathcal{A}, a, b) = 1$ и алгоритм GA находит оптимальное решение задачи (1).

Оценка может достигаться также и в тех случаях, когда допустимая область задачи (1) отлична от матроида. Например, в задаче, описанной в замечании 2, $S_{GA} = \{u_1, u_{p+2}, \ldots, u_{p+q+1}\}$ — независимое множество вершин графа G, найденное алгоритмом GA, $S_O = \{u_2, \ldots, u_{p+1}, u_{p+2}, \ldots, u_{p+q+1}\}$ — независимое множество вершин максимального веса, и

$$\frac{f(S_{GA})}{f(S_O)} = \frac{b + aq}{bp + aq} = c(\mathcal{A}, a, b).$$

Следует отметить, что достижимость оценки (5) имеет иной смысл, нежели достижимость оценки (3), поскольку существуют наследственные системы и значения $a, b > 0$ такие, что $\frac{f(S_{GA})}{f(S_O)} > c(\mathcal{A}, a, b)$ для любой целевой функции $f \in \mathcal{F}(a, b)$ задачи (1).

3. Оценка погрешности обратного жадного алгоритма для задачи о минимальном зависимом множестве

Как и ранее, предположим, что аддитивная целевая функция $f : 2^U \to R_+$ задачи (2) принадлежит классу $\mathcal{F}(a, b)$, где $a \geqslant 0$.

Введём следующую характеристику задачи (2) с целевой функцией $f \in \mathcal{F}(a, b)$:

$$c(\mathcal{D}, a, b) = \max_{\substack{W \subseteq U, \\ W \notin \mathcal{D}}} \frac{(b-a)(g_{\max}(W) - |W|) + ag_{\max}(\varnothing)}{(b-a)(g_{\min}(W) - |W|) + ag_{\min}(\varnothing)}.$$

В этом разделе получено следующее уточнение оценки (4).

Теорема 4. *Пусть $\mathcal{S} = (U, \mathcal{D})$ — произвольная наследственная система. Тогда для любой аддитивной целевой функции $f \in \mathcal{F}(a, b)$ задачи минимизации (2) справедлива оценка*

$$\frac{f(S_{GR})}{f(S_O)} \leqslant c(\mathcal{D}, a, b). \tag{12}$$

где S_O — оптимальное решение задачи (2).

Доказательство. Пусть $S_O = \{v_1, \ldots, v_m\}$, $f(v_i) = w_i \in [a, b]$, причём $w_1 \geqslant \ldots \geqslant w_m$. Положим $U_0 = \varnothing$, а при $i = 1, \ldots, m$ положим

$$U_i = \{v_1, \ldots, v_i\} \cup \{u \in U \mid \ U \setminus \{v_1, \ldots, v_i, u\} \in \mathcal{A}\}$$

Пусть $d_i = |S_{GR} \cap (U_i \setminus U_{i-1})|$, $i = 1, \ldots, m$. Следовательно, $|S_{GR}| = d_1 + \ldots + d_m$.

Тогда

$$\frac{f(S_{GR})}{f(S_O)} = \frac{\sum\limits_{i=1}^{m} \left(\sum\limits_{u \in S_{GR} \cap (U_i \setminus U_{i-1})} f(u) \right)}{\sum\limits_{i=1}^{m} f(v_i)} \leqslant \frac{\sum\limits_{i=1}^{m} w_i d_i}{\sum\limits_{i=1}^{m} w_i} \ . \tag{13}$$

Последнее неравенство следует из того, что $f(u) \leqslant f(v_i) = w_i$ для любого элемента $u \in U_i \setminus U_{i-1}$, так как алгоритм GR выбрал u_i.

Пусть

$$\frac{b \sum\limits_{j=1}^{k} d_j + a \sum\limits_{j=k+1}^{m} d_j}{bk + a(m-k)} = \max_{0 \leqslant i \leqslant m} \frac{b \sum\limits_{j=1}^{i} d_j + a \sum\limits_{j=i+1}^{m} d_j}{bi + a(m-i)}. \tag{14}$$

Заметим, что условие (14) равносильно условию (6). Тогда по лемме 1 верно неравенство (7), которое, в свою очередь, равносильно следующему неравенству

$$\frac{w_1 d_1 + \ldots + w_m d_m}{w_1 + \ldots + w_m} - \frac{b \sum\limits_{j=1}^{k} d_j + a \sum\limits_{j=k+1}^{m} d_j}{bk + a(m-k)} \leqslant 0.$$

Отсюда и из неравенства (13) следует, что

$$\frac{f(S_{GR})}{f(S_O)} \leqslant \frac{b \sum\limits_{j=1}^{k} d_j + a \sum\limits_{j=k+1}^{m} d_j}{bk + a(m-k)} \ . \tag{15}$$

Поскольку $k = |S_O \cap U_k|$, $m = |S_O|$, $\sum\limits_{j=1}^{k} d_j = |S_{GR} \cap U_k|$, а $\sum\limits_{j=1}^{m} d_j = |S_{GR}|$, неравенство (15) может быть переписано в виде

$$\frac{f(S_{GR})}{f(S_O)} \leqslant \frac{(b-a)|S_{GR} \cap U_k| + a|S_{GR}|}{(b-a)|S_O \cap U_k| + a|S_O|} \ .$$

Осталось заметить, что S_{GR} — цикл пустого множества и, как следует из определения множества U_k, $(S_{GR} \cap U_k) \cup \overline{U}_k$ — цикл множества $\overline{U}_k = U \setminus U_k$. Следовательно,

$$|S_{GR}| \leqslant g_{\max}(\varnothing) \text{ и } |S_{GR} \cap U_k| \leqslant g_{\max}(\overline{U}_k) - |\overline{U}_k|.$$

Кроме того, $|S_O| \geqslant g_{\min}(\varnothing)$, а поскольку множество $(S_O \cap U_k) \cup \overline{U}_k$ зависимо, то $|S_O \cap U_k| \geqslant g_{\min}(\overline{U}_k) - |\overline{U}_k|$. Поэтому

$$\begin{aligned} \frac{f(S_{GR})}{f(S_O)} &\leqslant \frac{(b-a)(g_{\max}(\overline{U}_k) - |\overline{U}_k|) + a g_{\max}(\varnothing)}{(b-a)(g_{\min}(\overline{U}_k) - |\overline{U}_k|) + a g_{\min}(\varnothing)} \\ &\leqslant \max_{0 \leqslant i \leqslant m} \frac{(b-a)(g_{\max}(\overline{U}_i) - |\overline{U}_i|) + a g_{\max}(\varnothing)}{(b-a)(g_{\min}(\overline{U}_i) - |\overline{U}_i|) + a g_{\min}(\varnothing)} \\ &\leqslant \max_{\substack{W \subseteq U, \\ W \notin \mathcal{D}}} \frac{(b-a)(g_{\max}(W) - |W|) + a g_{\max}(\varnothing)}{(b-a)(g_{\min}(W) - |W|) + a g_{\min}(\varnothing)} = c(\mathcal{D}, a, b). \end{aligned}$$

Теорема 4 доказана.

Замечание 4. Как и в случае максимизации аддитивной функции на наследственной системе, можно показать, что всегда $c(\mathcal{D}, a, b) \leqslant c(\mathcal{D})$ и разница между этими двумя величинами может быть значительной.

Замечание 5. Оценка (12) является достижимой.

Оценка (12) достижима в более слабом смысле, чем оценка (4). Она достигается, если наследственная система $\mathcal{S} = (U, \mathcal{D})$ является коматроидом, так как в этом случае $c(\mathcal{D}, a, b) = 1$ и алгоритм GR находит оптимальное решение задачи (2).

Оценка может достигаться также и в тех случаях, когда допустимая область задачи (2) отлична от коматроида. Это подтверждается следующим примером. Пусть $\mathcal{S}'_G = (V, \mathcal{D}'_G)$ — система, дополнительная к наследственной системе графа $G = K_{1,p} \cup O_q$ из замечания 2. Напомним, что семейство $\mathcal{D}'_G$ зависимых множеств системы $\mathcal{S}'_G$ состоит из всех вершинных покрытий в графе G. Здесь $S_{GR} = \{u_2, \ldots, u_{p+1}, u_{p+2}, \ldots, u_{p+q+1}\}$ — вершинное покрытие, найденное алгоритмом GR, $S_O = \{u_1, u_{p+2}, \ldots,$

$u_{p+q+1}\}$ — вершинное покрытие минимального веса,

$$\frac{f(S_{GR})}{f(S_O)} = \frac{bp + aq}{b + aq} = c(\mathcal{D}, a, b).$$

Автор признателен рецензенту за полезные замечания.

ЛИТЕРАТУРА

1. **Edmonds J.** Matroids and the greedy algorithm // Math. Programming. 1971. V. 1, N 2. P. 127–136.
2. **Grötschel M., Lovász L.** Combinatorial optimization. Handbook of combinatorics. V. 2. Amsterdam: Elsevier Science B.V., 1995. P. 1541–1598.
3. **Il'ev V.** Hereditary systems and greedy-type algorithms // Discrete Appl. Math. 2003. V. 132, N 1–3. P. 137–148.
4. **Jenkyns Th. A.** The efficacy of the "greedy" algorithm // Congressus Numerantium, N XVIII. Winnipeg, 1976. P. 341–350.
5. **Korte B., Hausmann D.** An analysis of the greedy heuristic for independence systems // Annals Discrete Math. 1978. V. 2. P. 65–74.
6. **Rado R.** Note on independence functions // Proc. London. Math. Soc. 1957. V. 7, N 3. P. 300–320.

Адрес автора:

Омский государственный университет,
пр. Мира, 55а,
644077 Омск,
Россия.
E-mail: iljev@math.omsu.omskreg.ru

Статья поступила
30 октября 2007 г.