ЗАДАЧИ КОМБИНАТОРНОЙ ОПТИМИЗАЦИИ НА НАСЛЕДСТВЕННЫХ СИСТЕМАХ

В. П. Ильев

1. Наследственные системы и жадные алгоритмы

Пусть U — конечное множество и $\mathcal{A} \subseteq 2^U$ — непустое семейство его подмножеств, удовлетворяющее следующей *аксиоме наследственности*: $A \in \mathcal{A},\ A' \subseteq A \Rightarrow A' \in \mathcal{A}$. Пара $\mathcal{S} = (U, \mathcal{A})$ называется *системой независимости* или *наследственной системой* на U. Множества семейства $\mathcal{A}$ называются *независимыми*, все остальные подмножества U — *зависимыми*. Семейство всех зависимых множеств обозначим $\mathcal{D}$. 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 оптимизационные задачи вида:

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

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

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 функция множеств.

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 p-медиане, задача о покрытии множества, задача о минимальном k-связном остовном подграфе и другие. Как правило, оптимизационные задачи на наследственных системах являются NP-трудными.

В большинстве задач, математическими моделями которых являются задачи (1) и (2), целевые функции являются аддитивными, субмодулярными или супермодулярными. Напомним, что функция множеств $f : 2^U \to R_+$ называется *субмодулярной*, если для любых $X, Y \subseteq U$ выполняется неравенство $f(X \cup Y) + f(X \cap Y) \leq f(X) + f(Y)$, и *супермодулярной*, если имеет место обратное неравенство. В случае равенства функцию f будем называть *модулярной*. Несложно показать, что неубывающая функция множеств с условием $f(\emptyset) = 0$ модулярна тогда и только тогда, когда она аддитивна.

В качестве приближенного метода решения задачи (1) рассмотрим следующий

Алгоритм GA (*жадный алгоритм*).

Шаг 0. $X_0 \leftarrow \emptyset$, перейти на шаг 1.

Шаг i ($i \geq 1$). Выбрать такой $x_i \notin X_{i-1}$, что $f(X_{i-1} \cup \{x_i\}) = \max\limits_{\substack{x \notin X_{i-1},\\ X_{i-1} \cup \{x\} \in \mathcal{A}}} f(X_{i-1} \cup \{x\})$.

$X_i \leftarrow X_{i-1} \cup \{x_i\}$, перейти на шаг $i+1$. Если такого $x_i \notin X_{i-1}$ нет, то $GA \leftarrow X_{i-1}$.

Конец.

Для приближенного решения задачи (2) будем применять следующий "обратный" аналог жадного алгоритма.

Алгоритм GR.
Шаг 0. $X_0 \leftarrow U$, перейти на шаг 1.
Шаг i ($i \geq 1$). Выбрать такой $x_i \in X_{i-1}$, что $f(X_{i-1} \setminus \{x_i\}) = \min\limits_{\substack{x \in X_{i-1}, \\ X_{i-1} \setminus \{x\} \in \mathcal{D}}} f(X_{i-1} \setminus \{x\})$.
$X_i \leftarrow X_{i-1} \setminus \{x_i\}$, перейти на шаг $i+1$. Если такого $x_i \in X_{i-1}$ нет, то $GR \leftarrow X_{i-1}$.
Конец.

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

2. Задачи с аддитивными и модулярными целевыми функциями

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 назовем любое минимальное по включению зависимое множество, содержащее W. Определим величины

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

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. Очевидно, что $c_{\mathcal{A}}(\mathcal{S}) \leq 1 \leq c_{\mathcal{D}}(\mathcal{S}) \leq n-1$ для любой системы $\mathcal{S}$.

Наследственная система $\mathcal{S}$ называется *матроидом*, если $c_{\mathcal{A}}(\mathcal{S}) = 1$, и *коматроидом*, если $c_{\mathcal{D}}(\mathcal{S}) = 1$. Число $r_{\max}(U)$ называются *рангом* матроида, а $g_{\max}(\emptyset)$ — *обхватом* коматроида. Примерами могут служить *p-униформный* матроид $\mathcal{S} = (U, \mathcal{A})$ и *p-униформный* коматроид $\mathcal{S}' = (U, \mathcal{D}')$, где $\mathcal{A} = \{A \subseteq U : |A| \leq p\}$, $\mathcal{D}' = \{D \subseteq U : |D| \geq p\}$, p — натуральное число, $p < |U|$.

Следующая теорема — один из центральных результатов теории матроидов.

Теорема 1 (Радо-Эдмондс) [5, 12]. *Пусть $\mathcal{S} = (U, \mathcal{A})$ — наследственная система на конечном множестве U. Алгоритм* **GA** *находит оптимальное решение задачи (1) на системе $\mathcal{S}$ для любой аддитивной целевой функции $f : U \to R_+$ тогда и только тогда, когда $\mathcal{S}$ — матроид.*

Этот результат может быть обобщен на случай произвольной модулярной функции. Для задачи (2) справедлива теорема, аналогичная теореме Радо-Эдмондса.

Теорема 2. *Пусть $\mathcal{S} = (U, \mathcal{D})$ — наследственная система. Алгоритм* **GR** *находит оптимальное решение задачи (2) на системе $\mathcal{S}$ для любой модулярной целевой функции тогда и только тогда, когда $\mathcal{S}$ — коматроид.*

Для задачи минимизации аддитивной функции подобный результат доказан в [6].

Как следует из теоремы 1, если наследственная система отлична от матроида, то жадный алгоритм может не найти оптимальное решение задачи (1) с аддитивной целевой функцией. В работах [8, 9] получена оценка погрешности алгоритма **GA** для задачи максимизации аддитивной функции на наследственной системе:

$$\frac{f(GA)}{f(OPT)} \geq c_{\mathcal{A}}(\mathcal{S}), \tag{3}$$

где OPT — оптимальное решение задачи (1).

В работе [6] доказана аналогичная оценка погрешности алгоритма **GR** для задачи (2) минимизации аддитивной функции на наследственной системе:

$$\frac{f(GR)}{f(OPT)} \leq c_{\mathcal{D}}(\mathcal{S}). \tag{4}$$

Привлечение дополнительной информации о целевой функции позволяет получить более точные, чем (3) и (4), оценки погрешности алгоритмов **GA** и **GR** для задач (1) и (2) с аддитивными целевыми функциями.

Кроме того, в работе [2] доказано, что задача (2) с аддитивной целевой функцией эквивалентна задаче о покрытии множества, что дает возможность применять известные результаты для задачи о покрытии к конкретным оптимизационным задачам, являющимся частными случаями (2).

3. Задачи с последовательными целевыми функциями

Если целевая функция задачи (1) на матроиде не является модулярной, то жадный алгоритм может не найти оптимальное решение. В то же время существуют задачи на матроидах с немодулярными целевыми функциями, в которых жадный алгоритм гарантированно находит оптимальное решение. В этом разделе охарактеризован класс целевых функций задач (1) на матроидах, разрешимых жадным алгоритмом, и приведено обобщение теоремы Радо-Эдмондса.

Пусть U конечное множество, $f : 2^U \to R_+$. Рассмотрим множество $V \subseteq U$ мощности $|V| \geq 3$ и натуральное число $k \leq |V| - 2$. Упорядоченное множество $X = \{x_1, \ldots, x_k\} \subset V$ будем называть *GA-подмножеством* множества V, если

$f(\{x_1\}) \geq f(\{x\})$ для любого $x \in V$,

$f(\{x_1, x_2\}) \geq f(\{x_1, x\})$ для любого $x \in V \setminus x_1$,

. . . .

$f(X) \geq f(\{x_1, \ldots, x_{k-1}, x\})$ для любого $x \in V \setminus \{x_1, \ldots, x_{k-1}\}$.

Функция множеств $f : 2^U \to R_+$ называется *GA-последовательной*, если для любого множества $V \subseteq U$ мощности $|V| \geq 3$ и любого его GA-подмножества $X = \{x_1, \ldots, x_k\}$ имеют место неравенства $f(V \setminus x_k) \leq f(V \setminus x)$ для всех $x \in V \setminus X$.

Заметим, что любая аддитивная функция является GA-последовательной.

Доказано следующее обобщение теоремы Радо-Эдмондса для задачи (1).

Теорема 3. *Пусть $\mathcal{S} = (U, \mathcal{A})$ — наследственная система. Алгоритм* **GA** *находит оптимальное решение задачи (1) на системе $\mathcal{S}$ для любой GA-последовательной функции множеств $f : 2^U \to R_+$ тогда и только тогда, когда $\mathcal{S}$ матроид.*

Справедлив также следующий критерий разрешимости жадным алгоритмом задачи (1) на матроиде.

Теорема 4. *Пусть $f : 2^U \to R_+$ — неотрицательная функция множеств. Алгоритм* **GA** *находит оптимальное решение задачи (1) с целевой функцией f на произвольном матроиде $\mathcal{S} = (U, \mathcal{A})$ тогда и только тогда, когда функция f является GA-последовательной.*

Аналогичные результаты получены для задачи (2): охарактеризован класс целевых функций задач (2) на коматроидах, разрешимых алгоритмом **GR** (класс *GR-последовательных* функций), и доказано обобщение теоремы 2 для задачи (2) на наследственной системе с целевой функцией из этого класса.

4. Задачи с субмодулярными и супермодулярными целевыми функциями

Рассмотрим задачу (1), в которой $\mathcal{B}$ — семейство баз матроида ранга p, а $f : 2^U \to R_+$ — неубывающая субмодулярная функция множеств, $f(\emptyset) = 0$. Частным случаем задачи (1) является известная *задача о p-медиане на максимум* с целевой функцией $f(X) = \sum\limits_{j \in J} \max\limits_{i \in X} c_{ij}$, где (c_{ij}) — неотрицательная матрица размера $n \times m$ с множеством индексов строк U и множеством индексов столбцов J.

В работе [4] получена оценка погрешности алгоритма **GA** для задачи о p-медиане на максимум:

$$\frac{f(GA)}{f(OPT)} \geq 1 - \left(\frac{p-1}{p}\right)^p \geq \frac{e-1}{e} \approx 0,63. \tag{5}$$

В [10] оценка (5) была обобщена для задачи максимизации неубывающей субмодулярной функции на p-униформном матроиде. В статье [3] эта оценка была уточнена с учетом дополнительной информации о целевой функции:

$$\frac{f(GA)}{f(OPT)} \geq \frac{1}{c}\left(1 - \left(\frac{p-c}{p}\right)^p\right),$$

где $c \in [0,1]$ — характеристика неубывающей субмодулярной функции f, описывающая замедление ее роста. Величина c определяется следующим образом:

$$c = \max_{\substack{x \in U, \\ f(\{x\}) > f(\emptyset)}} \frac{f(\{x\}) - f(\emptyset) - (f(U) - f(U \setminus \{x\}))}{f(\{x\}) - f(\emptyset)},$$

причем $c = 0$ тогда и только тогда, когда f аддитивна (при условии $f(\emptyset) = 0$).

Рассмотрим теперь вариант задачи (2), в котором $\mathcal{C}$ — семейство циклов коматроида обхвата p, а $f : 2^U \to R_+$ — невозрастающая супермодулярная функция множеств, $f(U) = 0$. Эта задача является обобщением *задачи о p-медиане на минимум* с целевой функцией $f(X) = \sum\limits_{j \in J} \min\limits_{i \in X} c_{ij}$, где (c_{ij}) — неотрицательная матрица размера $n \times m$ с множеством индексов строк U и множеством индексов столбцов J. Легко видеть, что после доопределения $f(\emptyset) = \max\limits_{\substack{X,Y \subseteq U, \\ X \cap Y = \emptyset}} \{f(X) + f(Y) - f(X \cup Y)\}$ невозрастающая целевая функция задачи о p-медиане на минимум становится супермодулярной.

К сожалению, алгоритм **GR** может давать сколь угодно плохое решение задачи минимизации супермодулярной функции. Кроме того, известно, что существование полиномиального алгоритма, который бы решал задачу о p-медиане на минимум с гарантированной оценкой погрешности, не превосходящей константы, означало бы, что $P = NP$ [11].

Однако, привлечение дополнительной информации о целевой функции задачи (2) делает возможным получение гарантированных оценок погрешности алгоритма **GR**.

Определим *крутизну* невозрастающей супермодулярной функции f как

$$s = \max_{\substack{x \in U, \\ f(\{x\}) < f(\emptyset)}} \frac{f(\emptyset) - f(\{x\}) - (f(U \setminus \{x\}) - f(U))}{f(\emptyset) - f(\{x\})}.$$

Параметр s характеризует замедление убывания функции f. Нетрудно показать, что $s \in [0,1]$, и что $s = 0$ тогда и только тогда, когда f — модулярная функция.

В работе [1] для функций крутизны $s < 1$ получены следующие оценки погрешности алгоритма **GR** минимизации невозрастающей супермодулярной функции на p-униформном коматроиде:

$$\frac{f(GR)}{f(OPT)} \leq \frac{1}{t}\left(\left(\frac{q+t}{q}\right)^q - 1\right) \leq \frac{e^t - 1}{t},$$

где $q = n - p$, а $t = s/(1-s)$. В статье [7] этот результат обобщен для задачи (2) минимизации невозрастающей супермодулярной функции на произвольном коматроиде обхвата p.

Как следствие получены оценки погрешности алгоритма **GR** для общей задачи о p-медиане на минимум в терминах матрицы (c_{ij}).

ЛИТЕРАТУРА

1. В.П. Ильев. Оценка точности алгоритма жадного спуска для задачи минимизации супермодулярной функции // Дискрет. анализ и исслед. операций. Сер. 1. 1998. Т. 5, N 4. С. 45-60.

2. В.П. Ильев, А.С. Талевнин. Две задачи на наследственных системах // Дискрет. анализ и исслед. операций. Сер. 1. 2003. Т. 10, N 3. С. 54-66.

3. M. Conforti, G. Cornuejols. Submodular set functions, matroids and the greedy algorithm: Tight worst-case bounds and some generalizations of the Rado-Edmonds theorem // Discrete Appl. Math. 1984. V. 7, N 3. P. 251-274.

4. G. Cornuejols, M.L. Fisher, G.L. Nemhauser. Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms // Management Science. 1977. V. 23. P. 789-810.

5. J. Edmonds. Matroids and the greedy algorithm // Math. Programming. 1971, V. 1, N 2. P. 127-136.

6. V. Il'ev. Hereditary systems and greedy-type algorithms // Discrete Appl. Math. 2003. V. 132, N 1-3. P. 137-148.

7. V. Il'ev, N. Linker. Performance guarantees of a greedy algorithm for minimizing a supermodular set function // European J. Oper. Res. 2006. V. 171, N 2. P. 648-660.

8. Th.A. Jenkyns. The efficacy of the "greedy" algorithm // Proc. 7th S-E Conf. Combinatorics, Graph Theory and Computing. 1976. P 341-350.

9. B. Korte, D. Hausmann. An analysis of the greedy heuristic for independence systems // Annals of Discrete Mathematics. 1978. V. 2. P. 65-74.

10. G.L. Nemhauser, L.A. Wolsey, M.L. Fisher. An analysis of approximations for maximizing submodular set functions – I // Math. Programming. 1978. V. 14. P. 265-294.

11. G.L. Nemhauser, L.A. Wolsey. Integer and combinatorial optimization. New York.: John Wiley & Sons, Inc., 1988.

12. R. Rado. Note on independence functions // Proc. London. Math. Soc. 1957. V. 7, N 3. P. 300-320.

Ильев Виктор Петрович,

Омский государственный университет, пр. Мира, 55а, Омск, 644077, Россия, тел. (3812) 22-56-96. E-mail: iljev@math.omsu.omskreg.ru