ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—июнь 2007. Серия 2. Том 14, № 1. 32–42

УДК 519.854

ЗАДАЧА ОТЫСКАНИЯ ПОДМНОЖЕСТВА ВЕКТОРОВ С МАКСИМАЛЬНЫМ СУММАРНЫМ ВЕСОМ*)

А. Е. Бабурин, Э. Х. Гимади, Н. И. Глебов, А. В. Пяткин

Доказана NP-трудность дискретных оптимизационных задач, связанных с выбором из конечного семейства векторов в евклидовом пространстве подмножества векторов с максимальной нормой суммы. Предложены приближённые алгоритмы и получены оценки для относительной погрешности и временно́й сложности. В случае фиксированной размерности пространства построена полиномиальная аппроксимационная схема. Выделен подкласс задач, для которых за псевдополиномиальное время отыскивается точное решение. Полученные результаты могут быть использованы для решения задачи выбора фиксированного числа фрагментов в числовой последовательности квазипериодически повторяющегося фрагмента при заданном числе повторов.

Введение

В статье приводится обоснование NP-трудности и построение приближённых алгоритмов решения дискретных оптимизационных задач, связанных с выбором из конечного семейства векторов в евклидовом пространстве $\mathbb{R}^k$ подмножества векторов с максимальной нормой суммы. Под нормой понимается евклидова норма, т. е. $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_k^2}$. В качестве основных задач в настоящей статье рассматриваются следующие две.

Задача 1. *Задано конечное семейство векторов $V = \{v_1, v_2, \ldots, v_n\}$ в евклидовом пространстве $\mathbb{R}^k$ и натуральное число $m < n$. Требуется из V выбрать m векторов, норма суммы которых максимальна.*

Вторая основная задача связана с целесообразным поиском сигналов в зашумлённой аддитивным шумом последовательности импульсов. Согласно [2, 4] эта задача поиска сводится к следующей.

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проекты 05–01–00395 и 07–07–00022) и INTAS (проект 04–77–7173).

Задача 2. *Задано конечное семейство векторов* $V = \{v_1, v_2, \ldots, v_n\}$ *в евклидовом пространстве* $\mathbb{R}^k$ *и натуральные числа* m *и* l, *удовлетворяющие условию* $l(m-1) < n$. *Требуется выделить в* V *подсемейство векторов* $X = \{v_{a_1}, v_{a_2}, \ldots, v_{a_m}\}$, *обладающее максимальной нормой суммы при условии* $a_{i+1} - a_i \geqslant l,\ i = 1, 2, \ldots, m-1$.

Для решения указанных задач предложены приближённые алгоритмы и оценены относительная погрешность и временна́я сложность. В случае фиксированной размерности пространства построена полиномиальная аппроксимационная схема. Выделен подкласс задач, для которых точное решение отыскивается за псевдополиномиальное время.

Полученные результаты могут быть использованы для решения задачи выбора фиксированного числа фрагментов в числовой последовательности квазипериодически повторяющегося фрагмента при заданном числе повторов. Как указывается в [2, 4], эта задача типична для таких приложений как электронная разведка, радиолокация, телекоммуникация, геофизика, обработка речевых сигналов, медицинская и техническая диагностика и др.

Для доказательства NP-трудности основных задач рассматриваются также некоторые вспомогательные задачи.

В первом разделе статьи доказывается, что NP-полна следующая

Задача Π'. *Дано* N' *булевых* k-*мерных векторов* $u^1, u^2, \ldots, u^{N'}$ *и натуральные константы* $K' > 0, m < N'$. *Можно ли среди этих векторов выбрать такие* m *векторов, что квадрат евклидовой нормы их суммы не меньше* K'?

При доказательстве этого факта используется сводимость к этой задаче известной NP-полной задачи КЛИКА [3].

NP-трудность задачи 1 следует из NP-полноты задачи Π' и сводимости последней к задаче 1.

Во втором разделе доказывается NP-полнота следующей вспомогательной задачи.

Задача Π. *Даны слово длины* $N > qM$ *над алфавитом* $\{0, 1\}$ *и натуральная константа* $K > 0$. *Можно ли в этом слове выбрать* M *таких непересекающихся отрезков длины* q, *что квадрат евклидовой нормы суммы выбранных* q-*мерных векторов не меньше* K?

Для этого доказывается сводимость к ней задачи Π'. А из NP-полноты вспомогательной задачи Π и её сводимости к задаче 2 вытекает NP-трудность задачи 2.

В разделе 3 приводится алгоритм приближённого решения задачи 1. При фиксированной размерности пространства построена полиномиальная аппроксимационная схема. Кроме того, обосновываются условия, при которых точное решение задачи 1 отыскивается за псевдополиномиальное время.

В разделе 4 описывается алгоритм приближённого решения задачи 2, временна́я сложность которого несущественно увеличивается по сравнению с алгоритмом решения задачи 1 при той же точности.

1. NP-полнота задачи Π' и её сводимость к задаче 1

Для доказательства NP-полноты вспомогательной задачи Π' воспользуемся известной [3] NP-полной задачей КЛИКА. *Дан n-вершинный граф G и натуральное число r. Есть ли в графе G r взаимно смежных вершин*?

Нам потребуется следующее очевидное

Утверждение. *Если $a \leqslant b$, то $(a-1)^2 + (b+1)^2 > a^2 + b^2$.*

Теорема 1. *Задача Π' является NP-полной.*

Доказательство. Сведём задачу КЛИКА к задаче Π'. Выберем числа p и s такими, чтобы выполнялись неравенства

$$(s+r-1)^2 > (s+r-2)^2 + (n-r)r^2; \quad p > n(s+n-1)^2 + 1.$$

Положим $N' = sn, k = pn, m = sr$ и $K' = r(s+r-1)^2 + (p-1)rs^2$.

Каждой вершине v_i в графе G поставим в соответствие s векторов $u^{(i-1)s+1}, u^{(i-1)s+2}, \ldots, u^{is}$ размерности k. Положим $u^{(i-1)s+t}_{(i-1)p+l} = 1$ для всех $t = 1, 2, \ldots, s$ и $l = 1, 2, \ldots, p$. Кроме того, для каждой пары смежных вершин v_i и v_j будем полагать, что $u^{(i-1)s+1}_{(j-1)p+1} = u^{(j-1)s+1}_{(i-1)p+1} = 1$. Остальные координаты всех векторов положим равными 0.

Назовем j-ю координату вектора *значимой*, если $j \equiv 1 \pmod p$, и *незначимой* в противном случае.

Покажем, что клика размера r в графе G существует тогда и только тогда, когда среди данных N' векторов можно выбрать m таких векторов $a_1, a_2, \ldots, a_m$, что $\|a_1 + a_2 + \ldots + a_m\|^2 \geqslant K'$.

Если в графе G есть клика размера r, то в качестве $a_1, a_2, \ldots, a_m$ выберем векторы, соответствующие вершинам этой клики. Тогда в векторе $a = a_1 + a_2 + \ldots + a_m$ окажется $(p-1)r$ незначимых координат, равных s, и r значимых координат, равных $(s+r-1)$. Следовательно, $\|a\|^2 \geqslant K'$.

Для доказательства обратного утверждения придётся оценить квадрат нормы суммы произвольных m векторов $a_1, a_2, \ldots, a_m$ из рассматриваемого семейства. Среди выбранных m векторов обозначим через x_i число векторов, соответствующих вершине v_i. Тогда вектор $a = a_1 +$

$a_2 + \ldots + a_m$ имеет по $p-1$ незначимых координат, равных x_i для каждого $i = 1, 2, \ldots, n$, и n значимых координат, каждая из которых не превосходит $s + n - 1$. Ясно, что $0 \leqslant x_i \leqslant s$ и $x_1 + x_2 + \ldots + x_n = sr$. Можно считать, что $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n$. Из утверждения следует, что $x_1^2 + x_2^2 + \ldots + x_n^2 \leqslant rs^2$, причём равенство достигается тогда и только тогда, когда $x_1 = x_2 = \ldots = x_r = s$, а $x_{r+1} = x_{r+2} = \ldots = x_n = 0$.

Допустим, что $x_r < s$. Тогда имеем

$$\|a\|^2 \leqslant (p-1)\sum_{j=1}^{n} x_j^2 + n(s+n-1)^2 \\ \leqslant (p-1)rs^2 - (p-1) + n(s+n-1)^2 < (p-1)rs^2 < K'.$$

Теперь предположим, что $x_1 = x_2 = \ldots = x_r = s$ и $x_{r+1} = x_{r+2} = \ldots = x_n = 0$. Тогда ровно $(p-1)r$ незначимых координат вектора a равны s, а остальные незначимые координаты равны 0. Кроме того, среди значимых координат лишь $a_1, a_{p+1}, \ldots, a_{(r-1)p+1}$ могут быть больше r. Если вершины $v_1, v_2, \ldots, v_r$ не образуют клику, то не более $r-1$ из этих координат могут быть равны $s + r - 1$. Таким образом, имеем

$$\|a\|^2 \leqslant (p-1)rs^2 + (r-1)(s+r-1)^2 + (s+r-2)^2 + (n-r)r^2 \\ < (p-1)rs^2 + r(s+r-1)^2 = K'.$$

Значит, $\|a\|^2 \geqslant K'$ тогда и только тогда, когда вершины $v_1, v_2, \ldots, v_r$ образуют клику. Теорема 1 доказана.

Таким образом, вспомогательная задача Π' является NP-полной.

Из NP-полноты задачи Π' и её сводимости к задаче 1 следует

Теорема 2. *Задача* 1 *является* NP-*трудной.*

2. NP-полнота задачи Π и её сводимость к задаче 2

Теорема 3. *Задача* Π *является* NP-*полной.*

Доказательство. Сведём задачу Π' к задаче Π. По данному входу задачи Π' строим вход задачи Π следующим образом. Положим $q = 2k+4$. Каждому вектору u^i задачи Π' ставим в соответствие подслово A^i длины q следующего вида. Положим $A^i_1 = A^i_2 = \ldots = A^i_{k+2} = 0$, $A^i_{k+3} = A^i_q = 1$ и $A^i_{k+3+j} = u^i_j$ при $j = 1, 2, \ldots, k$. Эти подслова назовём *основными*. Пусть $B_1 = B_{k+3} = B_q = 1$ и $B_j = 0$ при $j \in \{2, 3, \ldots, k+2, k+4, k+5, \ldots, q-1\}$. Выберем число $p > (N'(k+2))^2$ и возьмём p копий подслова B (эти подслова назовём *вспомогательными*). Полученные $p + N'$ подслов запишем в произвольном порядке, отделяя соседние подслова друг от друга группами из q нулей. Получим слово длины $N = q(2N' + 2p - 1)$. Пусть $M = p + m$ и $K = 2(p+m)^2 + p^2 + K'$.

Будем говорить, что отрезок b_j в решении задачи Π *порождён* данным подсловом (B или A^i), если b_j с ним пересекается, и *правильно порождён* данным подсловом, если он с ним совпадает (т. е. $b_j = B$ или $b_j = A^i$). Ясно, что каждый отрезок порождён ровно одним словом и никакое решение задачи Π не может содержать более двух отрезков, порождённых одним подсловом. Для удобства будем отождествлять понятия «отрезок b_j длины q» и «q-мерный вектор b_j».

Покажем, что в задаче Π можно выбрать M непересекающихся отрезков $b_1, b_2, \ldots, b_M$ таких, что $\|b_1 + b_2 + \ldots + b_M\|^2 \geqslant K$ тогда и только тогда, когда в задаче Π' можно выбрать m векторов, квадрат евклидовой нормы суммы которых не меньше K'.

Пусть $u^{i_1}, u^{i_2}, \ldots, u^{i_m}$ — решение задачи Π', причём $\|u^{i_1} + u^{i_2} + \ldots + u^{i_m}\|^2 \geqslant K'$, $b_j = A^{i_j}$ при $j = 1, 2, \ldots, m$ и $b_j = B$ при $j = m+1, m+2, \ldots, m+p$. Нетрудно убедиться, что

$$\|b_1 + b_2 + \ldots + b_M\|^2 = 2(p+m)^2 + p^2 + \|u^{i_1} + u^{i_2} + \ldots + u^{i_m}\|^2 \geqslant K.$$

Теперь докажем обратное утверждение. Для этого рассмотрим некоторое оптимальное решение задачи Π. Обозначим через $X = (x_1, x_2, \ldots, x_q)$ сумму векторов b_j, соответствующих выбранным отрезкам. По предположению имеем $\|X\|^2 \geqslant K \geqslant 2(p+m)^2 + p^2$.

Назовём j-ю координату вектора X *большой*, если $x_j \geqslant p$. Сначала покажем, что X содержит ровно три больших координаты.

Действительно, общее число единиц в слове не превосходит $3p + N'(q+2) < 3p + \sqrt{p}$ при выбранном p. Следовательно, X содержит не более трёх больших координат. Предположим, что число таких координат меньше трёх. Пусть вектор Y получается из X переобозначением координат в порядке невозрастания, т. е. $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_q$. Тогда справедливы неравенства $y_2 \leqslant y_1 \leqslant p + m$, $\quad y_3 \leqslant p - 1$, $\quad \displaystyle\sum_{i=1}^{q} y_i < 3p + \sqrt{p}$. Чему может быть равен максимальный квадрат евклидовой нормы вектора с такими условиями? Из утверждения следует, что максимум достигается для вектора с координатами $y_2 = y_1 = p+m, y_3 = p-1, y_4 = y < \sqrt{p}$ и $y_j = 0$ при $j = 5, 6, \ldots, q$. Но квадрат нормы такого вектора не превосходит

$$2(p+m)^2 + (p-1)^2 + p = 2(p+m)^2 + p^2 - p + 1 < 2(p+m)^2 + p^2,$$

т. е. имеем противоречие.

Итак, X содержит ровно три большие координаты. Заметим также, что отрезок, порождённый вспомогательным подсловом B, имеет три ненулевые координаты тогда и только тогда, когда он порождён правильно.

Если каждый отрезок, порождённый вспомогательным подсловом, имеет не более двух единиц в больших координатах, то в этом случае может быть не более N' отрезков, имеющих три единицы в больших координатах. Тогда

$$y_1 + y_2 + y_3 \leqslant 2(p + m - N') + 3N' = 2p + N' + 2m < 3p + 2m.$$

Если же найдётся отрезок, порождённый вспомогательным подсловом, в котором имеются три единицы в больших координатах, то он правильно порождён, и номерами больших координат являются $1, k + 3, q$. Учитывая также то, что число отрезков, правильно порождённых вспомогательными словами, не превосходит p, имеем $y_1+y_2+y_3 \leqslant 3p+2m$, так как в каждом отрезке, порождённом основными подсловами, содержится не более двух единиц в координатах $1, k+3, q$. При этом две единицы в указанных координатах будут лишь в случае, когда отрезок порождён правильно (и единицы содержатся в координатах $k + 3$ и q).

Следовательно, в любом случае $y_1+y_2+y_3 \leqslant 3p+2m$ и равенство достигается в том и только том случае, когда решение содержит p отрезков, правильно порождённых вспомогательными подсловами, и m отрезков, правильно порождённых основными подсловами.

Если $y_1 + y_2 + y_3 \leqslant 3p + 2m - 1$, то из утверждения следует, что при имеющихся условиях максимум квадрата нормы достигается при $y_1 = p+m, y_2 = p+m-1$ и $y_3 = p$. При этом сумма квадратов остальных координат не превосходит p. Тогда

$$\|X\|^2 \leqslant (p + m)^2 + (p + m - 1)^2 + p^2 + p = 2(p + m)^2 + p^2 - p - 2m + 1 \\ < 2(p + m)^2 + p^2$$

— противоречие.

Таким образом, $y_1+y_2+y_3 = 3p+2m$ и вектор X порождён некоторыми отрезками $b_j = A^{i_j}$ при $j = 1, 2, \ldots, m$ и $b_j = B$ при $j = m + 1, m + 2, \ldots, m + p$. Тогда из неравенства

$$2(p+m)^2 + p^2 + K' = K \leqslant \|X\|^2 = 2(p+m)^2 + p^2 + \|u^{i_1} + u^{i_2} + \ldots + u^{i_m}\|^2$$

следует, что $\|u^{i_1} + u^{i_2} + \ldots + u^{i_m}\|^2 \geqslant K'$. Теорема 3 доказана.

Заметим, что задача П сводится к задаче 2. Достаточно положить $l = k = q$, $m = M$ и $n = N - q + 1$, а в качестве векторов v_i рассмотреть векторы, порождённые отрезками $[i, i + q - 1]$, $1 \leqslant i \leqslant N - q + 1$, данного слова. Тогда условия $a_{i+1} - a_i \geqslant l$, $1 \leqslant i < m$, гарантируют, что отрезки данного слова в задаче П, соответствующие выбранным векторам v_i, не пересекаются.

Из NP-полноты вспомогательной задачи П и её сводимости к задаче 2 следует

Теорема 4. *Задача* 2 *является* NP-*трудной.*

3. Алгоритм A_1 приближённого решения задачи 1

Для произвольного конечного множества векторов X определим функцию

$$S(X) = \sum_{v \in X} v.$$

Следующий алгоритм A_1 находит приближённое решение задачи 1 в пространстве $\mathbb{R}^k$.

Описание алгоритма

На входе алгоритма A_1 задаётся множество векторов $V = \{v_1, \ldots, v_n\}$ из пространства $\mathbb{R}^k$ и натуральный параметр L. Алгоритм A_1 строит семейство решений. Далее из этих решений выбирается наилучшее. Перейдём к формальному описанию алгоритма A_1.

Положим $W(L) = \{x \in \mathbb{Z}^k \,|\, \max\{|x_1|, |x_2|, \ldots, |x_k|\} = L\}$ (через $\mathbb{Z}$ обозначено множество целых чисел). Пусть множество $W(L)$ упорядочено, например, лексикографически. Обозначим i-й вектор в этом порядке через $W_i(L)$.

Шаг 1. Полагается $i = 1$.

Шаг 2. Выбирается решение — подмножество входных векторов $X^i = \{x_1^i, x_2^i, \ldots, x_m^i\}$, обладающих максимальными значениями скалярного произведения $(x_j^i, W_i(L))$, $j = 1, 2, \ldots, m$.

Шаг 3. Если $i < |W(L)|$, то полагается $i = i + 1$ и выполняется шаг 2. Иначе выполняется шаг 4.

Шаг 4. Из полученных на шаге 2 решений выбирается такое X, что

$$\|S(X)\| = \max_{i=1,\ldots,|W(L)|} \|S(X^i)\|.$$

Множество векторов X является результатом работы алгоритма A_1. Описание алгоритма A_1 закончено.

Теорема 5. *Алгоритм A_1 имеет временну́ю сложность* $O(nk^2(2L +$

$1)^{k-1}$) *и находит решение задачи* 1 *с гарантированной относительной погрешностью, не превышающей* $\frac{1}{8}(k-1)/L^2$.

Доказательство. Определяющий временну́ю сложность алгоритма шаг 2 выполняется $|W(L)|$ раз, а $|W(L)| \leqslant 2k(2L+1)^{k-1}$. Эта оценка следует из неравенства $|W(L)| \leqslant \sum\limits_{l=1}^{k} |W(L;l)|$, где $W(L;l) = \{x \in \mathbb{Z}^k \mid |x_l| = L;\ |x_r| \leqslant L$ при всех $r \neq l,\ 1 \leqslant r \leqslant k\}$.

Сложность выполнения скалярных произведений n входных векторов на вектор $W_i(L)$ равна $O(nk)$. Выбор m максимальных значений из n скаляров можно произвести за линейное по n время. Следовательно, шаг 2 алгоритма выполняется за $O(nk)$ операций и временна́я сложность всего алгоритма не превосходит $O(nk^2(2L+1)^{k-1})$.

Обозначим через X^* оптимальное решение задачи. Пусть $W_i(L)$ — вектор, максимально близкий по углу к вектору $S(X^*)$ среди $W(L)$, а φ — угол между $S(X^*)$ и $W_i(L)$. Пусть $s = \dfrac{S(X^*)L}{\max\{|S_1(X^*), S_2(X^*)\dots, S_k(X^*)|\}}$.

Рассмотрим вектор $w = (\lfloor s_1\rceil, \lfloor s_2\rceil, \dots, \lfloor s_k\rceil)$, где через $\lfloor x\rceil$ обозначается ближайшее к x целое число. Очевидно, что $w \in W(L)$.

Обозначим через φ угол между $W_i(L)$ и $S(X^*)$, а через ψ угол между w и $S(X^*)$ (между w и s). Ясно, что $\varphi \leqslant \psi$.

Применяя теорему косинусов к треугольнику, образованному векторами w и s, получаем

$$\begin{aligned}\cos\psi = -\frac{|w-s|^2 - |s|^2 - |w|^2}{2|s||w|} &= 1 - \frac{|w-s|^2 - (|s|-|w|)^2}{2|s||w|} \\ &\geqslant 1 - \frac{k-1}{8(\min\{|w|, |s|\})^2} \geqslant 1 - \frac{k-1}{8L^2},\end{aligned}$$

поскольку каждая координата вектора $w-s$ по абсолютной величине не превосходит $1/2$, и хотя бы одна из них равна 0.

Далее для найденного приближённого решения X имеем

$$\begin{aligned}\|S(X)\| \geqslant \|S(X^i)\| \geqslant \frac{(S(X^i), W_i(L))}{\|W_i(L)\|} &= \frac{\sum\limits_{j=1}^{m} (X_j^i, W_i(L))}{\|W_i(L)\|} \\ &\geqslant \frac{\sum\limits_{j=1}^{m} (X_j^*, W_i(L))}{\|W_i(L)\|} = \|S(X^*)\| \cos\varphi.\end{aligned}$$

Отсюда следует оценка относительной погрешности алгоритма A_1:

$$\frac{\|S(X^*)\| - \|S(X)\|}{\|S(X^*)\|} \leqslant 1 - \cos\varphi \leqslant 1 - \cos\psi \leqslant \frac{k-1}{8L^2}.$$

Теорема 5 доказана.

Из теоремы 5 непосредственно вытекает справедливость следующего утверждения об условиях полиномиальности и асимптотической точности предложенного алгоритма.

Теорема 6. *При фиксированной размерности k пространства $\mathbb{R}^k$ алгоритм A_1 решения задачи 1 является асимптотически точным при таком выборе параметра L, что $L(n)$ является произвольной неограниченно растущей функцией от n.*

Тем самым для случая фиксированной размерности k пространства $\mathbb{R}^k$ установлено построение полиномиальной аппроксимационной схемы. Действительно, пусть относительная погрешность равна $\varepsilon = \frac{k-1}{8L^2}$. Тогда $L = \sqrt{\frac{k-1}{8\varepsilon}}$ и для временно́й сложности алгоритма A_1 получим величину

$$O\left(k(kn + m\log n)\left(\sqrt{\frac{k-1}{2\varepsilon}} + 1\right)^{k-1}\right).$$

Теорема 7. *Если размерность k пространства $\mathbb{R}^k$ фиксирована, то алгоритм A_1 с параметром $L = 0{,}5kmb$, где*

$$b = \max_{i=1,\dots,n;\, j=1,\dots,k} |v_{ij}|,$$

является псевдополиномиальным точным алгоритмом решения задачи 1 с целочисленными координатами входных векторов.

Доказательство. В введённых обозначениях для точного решения задачи X^* имеем $\|S(X^*)\|^2 \leqslant km^2b^2$. Для приближённого решения X получается

$$\frac{\|S(X)\|^2}{\|S(X^*)\|^2} \geqslant (1-\varepsilon)^2 = (1 - \frac{k-1}{8L^2})^2 > 1 - \frac{k-1}{4L^2} > 1 - \frac{k-1}{(kmb)^2}$$
$$> 1 - \frac{1}{km^2b^2} \geqslant 1 - \frac{1}{\|S(X^*)\|^2},$$

откуда следует, что $\|S(X)\|^2 > \|S(X^*)\|^2 - 1$.

Учитывая целочисленность $\|S(X)|^2$ и $\|S(X^*)\|^2$ при целочисленности координат входных векторов, заключаем, что

$$\|S(X)\| = \|S(X^*)\|.$$

Тем самым при описанных в условии теоремы допущениях алгоритм A_1 находит точное решение задачи. При этом алгоритм использует $O(nk^2(mbk)^{k-1})$ операций, а множитель b^{k-1} определяет псевдополиномиальность предложенного алгоритма. Теорема 7 доказана.

4. Алгоритм A_2 приближённого решения задачи 2

Алгоритм A_2 находит приближённое решение задачи 2 в пространстве $\mathbb{R}^k$. Приведём описание только второго шага этого алгоритма, поскольку именно этим шагом отличаются оба алгоритма.

Шаг 2. Обозначим скалярное произведение векторов v_j и $W_i(L)$ через w_j, т. е. $w_j = (v_j, W_i(L))$, $j = 1, \ldots, n$. Рассматривается ориентированный граф $G_i(V, E)$ со взвешенными вершинами. Вес вершины v_j определяется как w_j для всех $j = 1, 2, \ldots, n$. Положим $E = \{(v_i, v_j) | j - i \geqslant l\}$. Методом динамического программирования в графе G_i находится цепь длины $m-1$ с максимальным суммарным весом входящих в неё вершин. Множество вершин этой цепи обозначается через X^i.

Теорема 8. *Алгоритм A_2 за время $O\Big(nk(k+m)(2L+1)^{k-1}\Big)$ находит решение задачи 2 с относительной погрешностью, не превышающей величины $\frac{1}{8}(k-1)/L^2$.*

Доказательство. Доказательство погрешности алгоритма A_2 аналогично доказательству в теореме 5. Множество вершин цепи, найденной на шаге 2 алгоритма, образует оптимальное решение задачи выбора m скаляров максимального веса при ограничениях, заданных в условии задачи 2. Шаг 2, определяющий временну́ю сложность алгоритма, выполняется $|W(L)|$ раз, причём $|W(L)| \leqslant k(2L+1)^{k-1}$. Сложность расчёта скалярных произведений n входных векторов на вектор $W_i(L)$ не превосходит $O(nk)$. Построение цепи на шаге 2 выполняется за $O(nm)$ операций. Поэтому временна́я сложность шага 2 алгоритма не превосходит $O(nk+nm)$, а всего алгоритма — $O\Big(nk(k+m)(2L+1)^{k-1}\Big)$. Теорема 8 доказана.

Остаётся открытым вопрос о сложностном статусе задач 1 и 2 при фиксированной размерности k пространства $\mathbb{R}^k$.

Выражаем признательность А. А. Кельманову, привлекшему наше внимание к рассмотрению данного класса задач дискретной оптимизации.

ЛИТЕРАТУРА

1. **Ахо А., Хопкрофт Дж., Ульман Дж.** Построение и анализ вычислительных алгоритмов. М.: Мир, 1979.
2. **Гимади Э. Х., Кельманов А. В., Кельманова М. А., Хамидуллин С. А.** Апостериорное обнаружение в числовой последовательности квазипериодически повторяющегося фрагмента при заданном числе повторов // Сиб. журн. индустриальной математики. 2006. Т. IX, № 1(25). С. 55–74.
3. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.
4. **Кельманов А. В., Хамидуллин С. А., Окольнишникова Л. В.** Апостериорное обнаружение одинаковых подпоследовательностей-фрагментов в квазипериодической последовательности // Сиб. журн. индустриальной математики. 2002. Т. 5, № 2(10). С. 94–108.

Адрес авторов:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск,
Россия.
E-mail: gimadi@math.nsc.ru

Статья поступила
7 декабря 2006 г.

Переработанный вариант —
16 мая 2007 г.