ТОЧНЫЙ ПСЕВДОПОЛИНОМИАЛЬНЫЙ АЛГОРИТМ ДЛЯ ОДНОЙ ЗАДАЧИ ДВУХКЛАСТЕРНОГО РАЗБИЕНИЯ МНОЖЕСТВА ВЕКТОРОВ[1]

А.В. Кельманов, В.И. Хандеев

Институт математики им. С.Л. Соболева Сибирского отделения РАН,
Новосибирский государственный университет, Новосибирск
e-mail: kelm@math.nsc.ru, vladimir.handeev@gmail.com

В докладе рассматривается следующая NP-трудная [1] в сильном смысле

Задача. *Дано: множество $\mathcal{Y} = \{y_1, \ldots, y_N\}$ векторов из $\mathbb{R}^q$ и натуральное число M. Найти: разбиение множества $\mathcal{Y}$ на два кластера $\mathcal{C}$ и $\mathcal{Y} \setminus \mathcal{C}$ такое, что*

$$\sum_{y \in \mathcal{C}} \|y - \overline{y}(\mathcal{C})\|^2 + \sum_{y \in \mathcal{Y} \setminus \mathcal{C}} \|y\|^2 \to \min,$$

где $\overline{y}(\mathcal{C}) = \frac{1}{|\mathcal{C}|} \sum\limits_{y \in \mathcal{C}} y$ — центр кластера $\mathcal{C}$, при ограничении $|\mathcal{C}| = M$.

В [2] построен 2-приближённый полиномиальный алгоритм, временная сложность которого есть величина $\mathcal{O}(qN^2)$. В [3] обоснована полиномиальная приближённая схема (PTAS) с временной сложностью $\mathcal{O}(qN^{2/\varepsilon+1}(9/\varepsilon)^{3/\varepsilon})$, где ε — гарантированная оценка относительной погрешности алгоритма. В [4] предложен алгоритм, который при заданных относительной ошибке и вероятности несрабатывания для установленного значения параметра k позволяет находить приближенное решение задачи за время $\mathcal{O}(2^k q(k+N))$, а также найдены условия, при которых алгоритм асимптотически точен и имеет трудоёмкость $\mathcal{O}(qN^2)$.

Основным результатом настоящей работы является псевдополиномиальный алгоритм с временной сложностью $\mathcal{O}(qN(2MD+1)^q)$, где D — максимальное абсолютное значение координат векторов входного множества, гарантирующий оптимальное решение задачи в случае, когда компоненты векторов имеют целочисленные значения, а размерность пространства фиксирована.

ЛИТЕРАТУРА

1. А.В. Кельманов, А.В. Пяткин *О сложности некоторых задач поиска подмножеств векторов и кластерного анализа.* — Журн. вычисл. математики и мат. физики. — 2009, т.49, №11, с. 2059-2067.
2. А.В. Долгушев, А.В. Кельманов *Приближённый алгоритм решения одной задачи кластерного анализа.* — Дискретный анализ и исследование операций. — 2011, т.18, №2, с. 29-40.
3. А.В. Долгушев, А.В. Кельманов, В.В. Шенмайер *Приближенная полиномиальная схема для одной задачи кластерного анализа.* Интеллектуализация обработки информации: 9-я международная конференция. Республика Черногория. Сб. докл. М.: Торус Пресс, 2012, с. 242-244.
4. Alexander Kel'manov, Vladimir Khandeev *A randomized algorithm for a clustering problem.* Proceedings of IV International Conference «Optimization and applications» (OPTIMA-2013), Petrovac, Montenegro, September 22–28, 2013, p. 86.

[1]Работа выполнена при поддержке РФФИ (проекты № 12-01-00090 и № 13-07-00070)