|  |  |
| --- | --- |
|  | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет**  **имени Н.Э. Баумана (национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

|  |  |
| --- | --- |
| ФАКУЛЬТЕТ | *Робототехника и комплексная автоматизация* |
|  |  |
| КАФЕДРА | *Системы автоматизированного проектирования (РК-6)* |

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

***К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ***

***НА ТЕМУ***

***«Задача линейного упорядочивания: области применения и способы решения»***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Студент | РК6-41М | |  |  | А. С. Антонов |
|  | (Группа) | |  | (подпись, дата) | (инициалы и фамилия) |
|  | |  |  |  |  |
| Руководитель | | |  |  | А. Н. Божко |
|  | |  |  | (подпись, дата) | (инициалы и фамилия) |
|  | |  |  |  |  |
|  | |  |  |  |  |

*Москва, 2025 г*

Оглавление

[1. ВВЕДЕНИЕ 3](#_Toc197774703)

[2. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ 4](#_Toc197774704)

[3. МЕТОДЫ РЕШЕНИЯ 9](#_Toc197774705)

[4. ОБЛАСТИ ПРИМЕНЕНИЯ 9](#_Toc197774706)

[5. ЗАКЛЮЧЕНИЕ 9](#_Toc197774707)

[6. СПИСОК ЛИТЕРАТУРЫ 10](#_Toc197774708)

1. ВВЕДЕНИЕ

Задача линейного упорядочения (Linear Ordering Problem, LOP) занимает центральное место в комбинаторной оптимизации благодаря своей универсальности и широкому спектру приложений. В условиях роста объема данных [1] и потребности в эффективном управлении сложными системами LOP становится инструментом для решения задач ранжирования, планирования и анализа зависимостей.

Гэри и Джонсон (1979) продемонстрировали, что LOP является NP-сложной задачей. Однако, благодаря её многочисленным применениям в различных областях, таких как археология (Гловер и др., 1972), экономика (Леонтьев, 2008), теория графов (Харон и Худри, 2007), машинный перевод (Тромб и Эйснер, 2009) мы можем найти большое количество работ, в которых бы LOP решался с помощью точных, эвристических и метаэвристических методов.

В экономике LOP используется для ранжирования инвестиционных проектов на основе их потенциальной доходности и рисков. В машинном обучении задача помогает строить консенсусные рейтинги в рекомендательных системах [2]. В спортивной аналитике с её помощью определяют рейтинги команд, учитывая исторические результаты матчей [3]. В биоинформатике LOP применяется для упорядочения геномных последовательностей, что критично для понимания эволюционных процессов [4]. В археологии LOP нашёл применение в задаче стратификации для определения наиболее вероятного хронологического порядка образцов, найденных в разных местах. Матрица, описывающая эту проблему, известна как матрица Харриса.

Рост интереса к Big Data и искусственному интеллекту усиливает потребность в алгоритмах, способных работать с высокоразмерными данными, что делает исследование LOP особенно актуальным.

Цель работы состоит в изучении теоретических основ задачи линейного упорядочения, исследовании областей применения и проведении сравнительного анализа методов её решения.

Работа включает в себя пять разделов. В теоретической части раскрывается постановка задачи и её связь с другими проблемами оптимизации. Далее анализируются методы решения, а в практическом разделе рассматриваются области применения данной задачи.

1. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Для заданной матрицы размера задача линейного упорядочения (LOP), формулируется, как задача нахождения перестановки строк и столбцов , которая бы максимизировала функцию в формуле 1.

где обозначает индекс строки (и столбца), занимающей позицию в решении . Иными словами, цель заключается в нахождении такой перестановки строк и столбцов матрицы , которая бы максимизировала сумму элементов, находящихся выше главной диагонали. В данной постановке задача линейного упорядочивания известна, как triangulation problem of input-output matrices. Тем не менее существуют и альтернативные варианты представления LOP. Так Марти и Рейнельт в своей книге 2011-го года «The linear ordering problem: exact and heuristic methods in combinatorial optimization» дают интерпретацию в терминах теории графов с поиском ациклического турнира (полный ориентированный ациклический подграф), максимизирующего сумму весов дуг полного ориентированного графа.

Сами значения отображают силу предпочтения -го элемента -ому.

Рассмотрим пример для , который будет использоваться в дальнейшем. На рисунке 1 представлены три различные решения: и . Исходная матрица представлена перестановкой (рис. 1а) и значением её фитнес функции , равным 138. Решение (рис. 1b) иллюстрирует другой вариант решения, имеющий более оптимальное значение фитнеса . Наилучшее возможное решение представлено перестановкой (рис. 1c). с фитнесом .

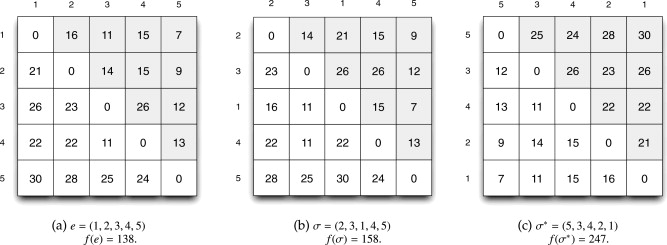


Рис. 1 – Три различных решения для матрицы : (а) исходная матрица; (b) неоптимальное решение; (c) оптимальное решение.

В ходе анализа проблемы было выявлено, что для любой перестановки индексов в матрице размером справедливо следующее:

* С каждым индексом связано записей из матрицы : из строки и из столбца .
* Набор связанных записей каждого индекса может быть сгруппирован в пары. Например, каждой записи из строки , , соответствует точка из столбца , , симметрично расположенная относительно главной диагонали.
* Все пары записей, связанные с индексом , остаются связанными с ним даже после перестановок.
* Каждая запись связана с двумя индексами, .
* Для каждой пары одна из записей всегда расположена выше главной диагонали, а другая ниже.

Утверждения проиллюстрированы примером на рисунке 2. В этом примере представлены два различных решения: на рис. 2a и на рис. 2b. На обоих рисунках записи, связанные с индексом 2, выделены жирным шрифтом. Мы видим, что, несмотря на разное упорядочивание, в обоих решениях набор записей, связанных с индексом 2, одинаков, то есть (21, 14, 15, 9, 16, 23, 22, 28). Несмотря на то, что позиция индекса 2 в различна , попарное отношение связанных с ним записей остается неизменным (см. обведенные индексы на рис. 2).

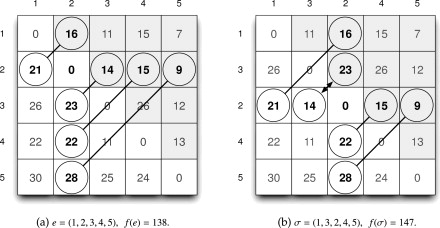


Рис. 2 – Два различных решения для экземпляра . Обведенные записи, соединенные ребрами, обозначают пары записей, связанные с индексом 2: (a) исходное состояние; (b) после перестановки.

Проверяя расположение связанных записей индекса 2 в , мы замечаем, что пара (14, 23) поменяла свои позиции в , 14 теперь ниже главной диагонали, а 23 выше неё. Введём новое понятие: вклад индекса в фитнес-функцию.

Когда индекс занимает позицию в , то есть , вклад индекса в функцию определяется суммой записей столбца в строках и суммой записей строки в столбцах . Иными словами, предыдущие индексы и последующие индексы определяют вклад индекса k в фитнес-функцию. Формальное описание представлено формулой 2.

Вернемся к примеру на рисунке 2, благодаря обмену позициями в пару (14, 23) вклад индекса 2 изменился с 54 (16 + 14 + 15 + 9) в (рис. 2a) до 63 (16 + 23 + 15 + 9) в (рис. 2b). В случае индекса 3 его вклад также увеличился, поскольку пара (14, 23) связана с обоими индексами, 2 и 3. И наоборот, в случае индексов 1, 4 и 5 их вклад не меняется от к .

Если мы внимательно посмотрим на уравнение (2), то поймем, что вклад индекса на самом деле не определяется конкретным упорядочиванием индексов в предыдущей и последующей позициях , но их группировкой в этих двух наборах позиций. Как показано в примере 2, вклад индексов 1, 4 и 5 не меняется от к , поскольку группировка остальных индексов в предыдущем и последующем наборах позиций, связанных с индексами 1, 4 и 5, была одинаковой.

Следовательно, при заданном решении вклад индекса , в фитнес-функцию не зависит от порядка предыдущих индексов и от порядка последующих индексов .

Пример из рисунка 3 иллюстрирует, как вклад индекса 3 не зависит от упорядочивания предыдущего и последующего наборов индексов. На рис. 3а вклад индекса 3 равен 63, как результат суммы (11 + 14 + 26 + 12). Если мы проверим вклад индекса 3 в (см. рис. 3б), то увидим, что он также равен 63, несмотря на то что индексы {1, 2} и {4, 5} поменялись местами.

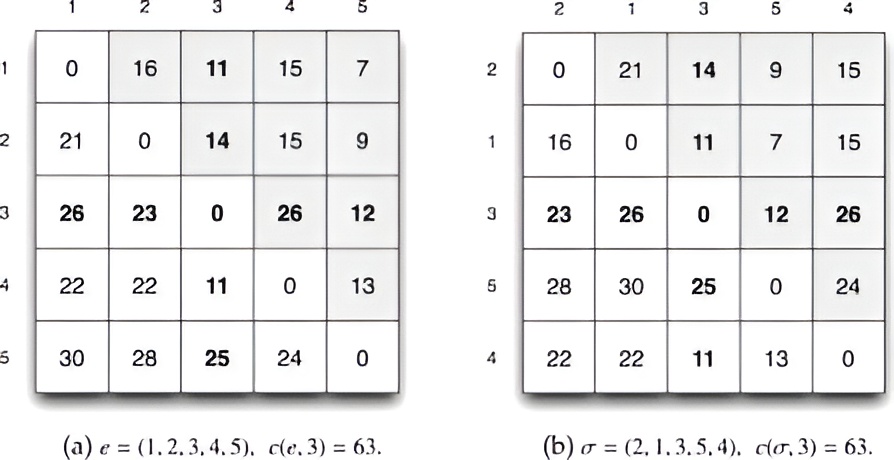


Рис. 3 – Эффект от замены индексов в позициях 1,2 и 4,5 для индекса 3: (a) исходное состояние; (b) после перестановки.

Как было сказано выше, вклад индекса не зависит от порядка следования индексов в предыдущих и последующих множествах. Но что произойдет, если индекс будет перемещен из предыдущего набора индексов в последующий набор индексов? В отличие от предыдущего случая, вклад индекса будет изменён. В этот момент стоит вспомнить, что каждая пара записей в матрице связана с двумя индексами, , а значит, любой обмен местоположением по определению влияет на вклад в функцию приспособленности и . Фактически, перемещение в позицию влияет на вклад всех индексов, расположенных между позициями и . Приведенный ниже пример (рисунок 4) иллюстрирует изменения в фитнес-функции, вызванные перемещением индекса.

На рисунке 4а и b показана матрица в соответствии с решениями . В этом примере мы анализируем последствия перемещения индекса в позицию 4. В результате этой модификации индексы 3 и 4 сдвигаются на одну позицию влево, что изменяет их вклад в фитнес-функцию. В частности, мы видим, что пары {14, 23} и {15, 22}, связанные с индексами 2-3 и 2-4, поменялись местами. Поэтому вклад индекса 3 меняется с 63 (11 + 14 + 26 + 12) на 72 (11 + 26 + 23 + 12). Аналогично, вклад индекса 4 меняется с 69 (15 + 15 + 26 + 13) до 76 (15 + 26 + 22 + 13). Что касается индекса 2, то его вклад также меняется с 54 (16 + 14 + 15 + 9) до 70 (16 + 23 + 22 + 9). Обратите внимание, что вариация фитнес-вклада индекса 2 равна сумме вариаций индексов 3 и 4.

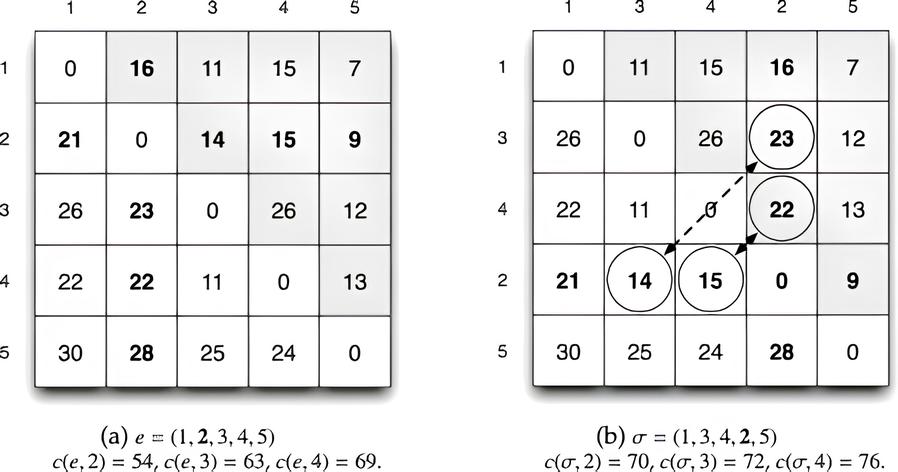


Рис. 4 – Иллюстрация влияния перемещения индекса 2 (из позиции 2 в позицию 4) на вклад индексов 2, 3 и 4 в фитнес-функцию. Числа, выделенные жирным шрифтом, обозначают записи, связанные с индексом 2. Обведенные кружком пары записей выделяют замененные записи: (a) исходное состояние; (b) после перестановки.

1. МЕТОДЫ РЕШЕНИЯ

LOP - это NP-полная задача, то есть мы не можем ожидать алгоритма, который решит ее за полиномиальное время. Однако LOP имеет множеств практических приложений [2, 3, 4], и поэтому алгоритмы для его эффективного решения востребованы. Было предложено несколько точных и эвристических алгоритмов. Точные алгоритмы включают в себя метод границ и ветвей, использующий LP-релаксацию для нижней границы, предложенный Каасом (A branch and bound algorithm for the acyclic subgraph problem.), алгоритм отсечений, предложенный Грёцшелем, Юнгером и Рейнелтом и комбинированный алгоритм Митчелла и Борчерса. Современные точные алгоритмы могут решать достаточно большие задачи из определенных классов задач с числом столбцов и строк до нескольких сотен, в то время как на экземплярах из других классов гораздо меньшего размера они терпят неудачу. Независимо от типа решаемых задача, время вычислений точных алгоритмов сильно увеличивается с ростом размера матрицы.

Проблему LOP также решали с помощью ряда эвристических алгоритмов. К ним относятся конструктивные алгоритмы, такие как жадный алгоритм Беккера [3], алгоритмы локального поиска, такие как эвристика CK от Чанаса и Кобыланьского [8], а также множество метаэвристических подходов, включая элитный поиск с запретами (tabu search), scatter search, представленные в серии работ Марти, Лагуны и Кампоса [17, 7, 6], а также алгоритмы итеративного локального поиска (ILS) [9, 25]. В частности, подходы на основе ILS в настоящее время считаются наиболее успешными метаэвристиками, судя по их результатам на ряде доступных тестовых примеров LOP [9, 25].

Переведено с помощью DeepL.com (бесплатная версия)

1. ОБЛАСТИ ПРИМЕНЕНИЯ
2. ЗАКЛЮЧЕНИЕ
3. СПИСОК ЛИТЕРАТУРЫ
4. Статья с сайта: Fabio Duarte Amount of Data Created Daily (2025) [электронный ресурс] // explodingtopics URL: <https://explodingtopics.com/blog/data-generated-per-day> (дата обращения: 06.05.2025)
5. Aparicio J., Landete M., Monge J. F. A linear ordering problem of sets //Annals of Operations Research. – 2020. – Т. 288. – №. 1. – С. 45-64.
6. Cameron T. R., Charmot S., Pulaj J. On the linear ordering problem and the rankability of data //arXiv preprint arXiv:2104.05816. – 2021.
7. Alcaraz J. et al. The linear ordering problem with clusters: a new partial ranking //Top. – 2020. – Т. 28. – С. 646-671.