Программа. Орг моменты

1. Матроиды.
2. Быстрое преобразование Фурье.
3. ρ-метод Полладра.
4. Автоматы. Регулярные языки.
5. Классы алгоритмов.
6. Численные методы.
7. Симплекс метод.
8. Венгерский алгоритм.
9. Local sensitive hashing.

Формула такая же, как и в прошлом году:

$0.3 \cdot O_{\text{контесты}} + 0.25 \cdot O_{\text{семинарские листки}} + 0.15 \cdot O_{\text{кр}} + 0.3 \cdot O_{\text{экзамен}} + \text{Б}$.

Округление вверх.

Лекция 01 от 02.09.2016. Матроиды.

Пока чуть отдаленно от матроидов.

У нас есть конечное множество A, которое в будущем мы будем называть *носителем*. Пусть $F \subset 2^A$, и F мы будем называть *допустимыми* множествами.

Также у нас есть весовая функция $c(w)$ $\forall w \in A$. Для каждого $B \in F$ мы определим *стоимость* множетсва, как $\sum\limits_{w \in B} c(w)$. Наша задача заключается в том, чтобы найти максимальный вес из всех допустимых множеств.

Пример 1 (Задача о рюкзаке)**.** *У каждого предмета есть вес и стоимость. Мы хотим унести как можно больше вещей максимальной стоимости с весом не более k.*

Вес не более k нам задает ограничение, то есть множество F. А максимизация унесенной суммы нам и задаёт задачу.

Матроид

Множество F теперь будет всегда обозначаться как I.

Матроидом называется множество подмножеств множества A таких, что выполняются следующие 3 свойства:

1. $\varnothing \in I$

2. $B \in I \Rightarrow \ \forall D \subset B \Rightarrow D \in I$

3. Если $B, D \in I$ и $|B| < |D| \Rightarrow \exists w \in D \setminus B$ такой, что $B \cup w \in I$

Дальнейшее обозначение матроидов — $\langle A, I \rangle$.

Определение 1. *Базой матроида называют множество всех таких элементов $B \in I$, что* **не** *существует B', что $B \subset B', |B'| > |B|$ и $B' \in I$. Обозначение β.*

Свойство 1. *Все элементы из базы имеют одну и ту же мощность. И все элементы из I, имеющие эту мощность, будут в базе.*

Доказательство очевидно из определения.

Пример 2 (Универсальный матроид)**.** *Это все подмножества B множества A такие, что $|B| \leqslant k$ при $k \geqslant 0$. Все свойства проверяются непосредственно.*

База такого матроида — все множества размера k.

Пример 3 (Цветной матроид)**.** *У элементов множества A имеются цвета. Тогда $B \in I$, если все элементы множества B имеют разные цвета. Свойства проверяются непосредственно, в 3 свойстве надо воспользоваться принципом Дирихле.*

База такого матроида — множества, где присутствуют все цвета.

Пример 4 (Графовый матроид на n вершинах)**.** *$\langle E, I \rangle$. Множество ребер $T \in I$, если T не содержит циклов.*

Докажем 3 свойство:

Доказательство. Пусть у нас есть T_1 и T_2 такие, что $|T_1| < |T_2|$. Разобьём граф, построенный на T_1 на компоненты связности. Так как ребер ровно $|T_1|$ на n вершинах, то компонент связности будет $n - |T_1|$. В другом случае компонент связности будет $n - |T_2| < n - |T_1|$. То есть во 2-ом графе будет меньше компонент связности, а значит по принципу Дирихле найдётся ребро, которое соединяет 2 компоненты связности в 1-ом графе.

Этот алгоритм чем-то отдаленно напоминает алгоритм Краскала. □

Базой в таком матроиде являются все остовные деревья.

Пример 5 (Матричный матроид)**.** *Носителем здесь будут столбцы любой фиксированной матрицы. I — множество всех подмножеств из линейно независимых столбцов. Все свойства выводятся из линейной алгебры (3-е из метода Гаусса, если быть точным).*

Пример 6 (Трансверсальный матроид)**.** *$G = \langle X, Y, E \rangle$ — двудольный граф с долями X, Y. Матроид будет $\langle X, I \rangle$ такой, что $B \in I$, если существует паросочетание такое, что множество левых концов этого паросочетания совпадает с B.*

Докажем 3 свойство:

Доказательство. Пусть есть 2 паросочетания на $|B_1|$ и $|B_2|$ ($|B_1| < |B_2|$) вершин левой доли. Тогда рассмотрим симметрическую разность этих паросочетаний. Так как во 2-ом паросочетании ребер больше, то существует чередующаяся цепь, а значит при замене ребер на этой чередующейся цепи с новой добавленной вершиной (а она найдётся по принципу Дирихле) получим паросочетание с ещё 1 добавленной вершиной. □

Базой в таком матроиде будут вершины левой доли максимального паросочетания.

Приводимость одной базы к другой.

Лемма 1. *Пусть $B, D \in \beta$. Тогда существует последовательность $B = B_0, B_1, \ldots, B_k = D$ такие, что $|B_i \triangle B_{i+1}| = 2$.*

Доказательство. Будем действовать по шагам. Если текущее $B_i \neq D$, тогда возьмём произвольный элемент w из $B_i \setminus D$. Тогда по 2-ому пункту определения матроида следует, что $B_i \setminus w \in I$. Так как $|B_i \setminus w| < |D|$, то существует $u \in D$ такой, что $(B_i \setminus w) \cup u \in I$. И теперь $B_{i+1} \leftarrow (B_i \setminus w) \cup u$. Мы сократили количество несовпадающих элементов с D на 1, симметрическая разность B_i и B_{i+1} состоит из 2 элементов — w и u. □

Наконец, мы подошли к основной теореме лекции — жадный алгоритм или теорема Радо-Эдмондса.

Жадный алгоритм на матроиде.

Доказательство будет в несколько этапов.

Для начала определимся с обозначениями. $M = \langle A, I \rangle, n = |A|, w_i$ — элементы множестваA. Решаем обычную задачу на максимизацию необходимого множества.

Теорема 1. *Если отсортировать все элементы A по невозрастанию стоимостей весовой функции: $c_1 \geqslant c_2 \geqslant \ldots \geqslant c_n$, то такой алгоритм решает исходную задачу о нахождении самого дорогого подмножества:*

$B = \varnothing$
for c_i **do**
    if $B \cup w_i \in I$ **then**
        $B = B \cup w_i$
    end
end

Доказательство. Теперь поймём, что наш алгоритм в итоге получит какой-то элемент из базы. Пусть B_i — множество, которое мы получим после i шагов цикла нашего алгоритма. Действительно, если это не так, что существует множество из базы, которое его накрывает: формально $\exists D \in I : B_n \subset D$ и $|B_n| < |D|$, так как можно взять любой элемент из базы и добавлять в B_n по 1 элементу из пункта 3 определения матроида. Тогда у нас существует элемент w_i, который мы не взяли нашим алгоритмом, но $B_{i-1} \cup w_i \in I$, так как $B_{i-1} \cup w_i \subset B_n \cup w_i \subset D$, то есть это лежит в I по пункту 2 определения матроида. Значит мы должны были взять w_i, противоречие.

Рассмотрим последовательность d_i из 0 и 1 длины n такую, что $d_i = 1$ только в том случае, если мы взяли алгоритмом i-ый элемент. А оптимальное решение задачи пусть будет e_i — тоже последовательность из 0 и 1. Последовательности будут обозначаться d и e соответственно.

Если на каком-то префиксе последовательности d единиц стало меньше, чем в e, то возьмём все элементы, которые помечены последовательностью e единицами. Пусть это множество будет E. Аналогично на этом префиксе последовательности d определим множество D. $|D| < |E|, D \in I, E \in I$, поэтому мы можем дополнить D каким-то элементом из E, которого не было в D. То есть на этом префиксе у d стоит 0 (пусть это будет место i), но заметим, что на i-ом шаге мы обязаны были брать этот элемент, из-за рассуждений аналогичным рассуждению про базу (2 абзаца вверх).

Получаем, что на каждом префиксе d единиц не меньше, чем на этом же префиксе последовательности e. Значит 1-ая единица в d встретится не позже, чем в e, 2-ая единица в d не позже, чем 2-ая в e и т.д. по рассуждениям по индукции.

$\square$

На лекции была теория про ранги. В доказательстве можно обойтись без неё, просто приложу то, что сказал Глеб. Может быть понадобиться в задачах.

Рангом множества $B \subset A$ (обозн. $r(B)$) называют максимальное число k такое, что $\exists C \subset B$ такое, что $|C| = k, C \in I$.

Эта функция обладает таким свойством: для любого элемента $w \in A$ следует, что $r(B \cup w) \leqslant r(B) + r(w)$. Давайте поймём, почему так:

Если $r(B \cup w) = r(B)$, то всё хорошо, так как $r(w) \geqslant 0$. Если $r(B \cup w) = r(B) + 1$ (других вариантов не бывает из определения), то тогда $w \in I$, так как в $B \cup w$ найдётся такое $C \subset (B \cup w)$, что $|C| = r(B \cup w), w \in C$ (иначе C годилось бы для B и $r(B \cup w) = r(B)$), значит $r(w) = 1$, так как $C \in I$, а $\{w\} \subset C$.