Дискретная оптимизация

МФТИ, весна 2016

Александр Дайняк

www.dainiak.com

Задача оптимизации аддитивной функции на семействе подмножеств конечного множества (DLS problem, Discrete Linear Subset problem)

- Е конечное множество
- Каждому $e \in E$ приписан вес $w(e) \ge 0$
- $\mathcal{F} \subset 2^E$ «допустимые» подмножества
- Для каждого $A \subseteq E$ полагаем

$$w(A) \coloneqq \sum_{e \in A} w(e)$$

• Требуется найти $A \in \mathcal{F}$, такое, что $w(A) \to \max/\min$

Задача DLS

Частные случаи задачи DLS:

- TSP (DLS-задача минимизации)
- MST (DLS-задача минимизации)
- 0,1-рюкзак (DLS-задача максимизации)

TSP и MST как DLS-задачи максимизации

$$E \coloneqq$$
 все рёбра графа G

- В задаче TSP
 - $w(e) \coloneqq (bigconst исходный вес ребра e в графе)$
 - $\mathcal{F} \coloneqq \{E' \subseteq E \mid E'$ образует г. ц. в $G\}$
- В задаче об остовном дереве:
 - $\mathcal{F} \coloneqq \{E' \subseteq E \mid E'$ образует о. д. в $G\}$
 - $w(e) \coloneqq (bigconst исходный вес ребра е в графе)$

Трюк с изменением весов работает только потому, что количество рёбер во всех г.ц./о.д. одинаковое!

Наследственные системы

Наследственная система (система независимости, independence system)

— это пара
$$(E,\mathcal{F})$$
, где $\emptyset \neq \mathcal{F} \subseteq 2^E$, для которой

$$\forall A \in \mathcal{F} \quad \forall A' \subset A \quad A' \in \mathcal{F}$$

Множества $I \in \mathcal{F}$ называются *независимыми*.

Базой множества $X \subseteq E$ называется любое такое $X' \in 2^X \cap \mathcal{F}$, что $\forall x \in X \setminus X' \quad X' \cup \{x\} \notin \mathcal{F}$

Базы множества E называются просто базами или базами системы (E,\mathcal{F}) .

Циклы — это множества $C \notin \mathcal{F}$, такие, что $C' \in \mathcal{F}$ для любого $C' \subset C$.

Ранг

Пусть (E, \mathcal{F}) — наследственная система.

Для множества $X \subset E$ определим

• ранг *X*

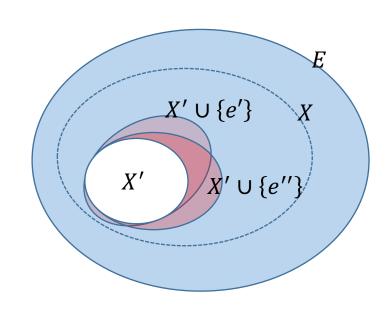
$$\operatorname{rk} X \coloneqq \max_{X' - \operatorname{fasa} X} |X'|,$$

ullet нижний ранг X

$$\widetilde{\operatorname{rk}} X \coloneqq \min_{X' - \operatorname{fasa} X} |X'|.$$

Ранговый разброс (rank quotient):

$$q(E,\mathcal{F}) \coloneqq \min_{X \subseteq E} \frac{\widetilde{\operatorname{rk}} X}{\operatorname{rk} X}.$$



Матроиды

Наследственная система (E, \mathcal{F}) называется матроидом, если $\forall X, Y \in \mathcal{F} \quad (|X| > |Y| \implies \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F}).$

Утверждение.

Наследственная система (E,\mathcal{F}) — матроид т. и т.т., когда выполнено любое из условий:

- 1. $\forall X, Y \in \mathcal{F} \quad (|X| > |Y| \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$
- 2. $\forall X, Y \in \mathcal{F}$ $(|X| = |Y| + 1 \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$
- 3. Для любого $X \subseteq E$ все базы X равномощны.
- 4. $q(E,\mathcal{F})=1$

Эквивалентность определений матроида

- 1. $\forall X, Y \in \mathcal{F} \quad (|X| > |Y| \Rightarrow \exists x \in X \setminus Y \text{ T.ч. } Y \cup \{x\} \in \mathcal{F})$
- 2. $\forall X, Y \in \mathcal{F} \quad (|X| = |Y| + 1 \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$
- 3. Для любого $X \subseteq E$ все базы X равномощны.
- 4. $q(E, \mathcal{F}) := \min_{X \subseteq E} \frac{\min\{|X'| \mid X' \text{база } X\}}{\max\{|X'| \mid X' \text{база } X\}} = 1$

Импликации $(1) \Leftrightarrow (2)$, $(1) \Rightarrow (3)$, $(3) \Leftrightarrow (4)$ очевидны.

Остаётся доказать, например, импликацию $(3) \Rightarrow (1)$.

Пусть система (E,\mathcal{F}) удовлетворяет (3) и пусть |X|>|Y|.

Тогда Y не может быть базисом $X \cup Y$.

Тогда найдётся $x \in (X \cup Y) \setminus Y = X \setminus Y$, такой, что $Y \cup \{x\} \in \mathcal{F}$.

Примеры матроидов

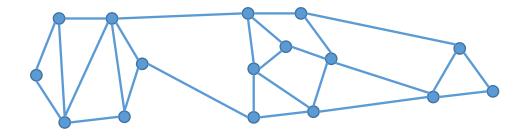
- 1. $\forall X, Y \in \mathcal{F}$ $(|X| > |Y| \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$ 2. $\forall X, Y \in \mathcal{F}$ $(|X| = |Y| + 1 \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$ 3. Для любого $X \subseteq E$ все базы X равномощны.
 4. $q(E, \mathcal{F}) \coloneqq \min_{X \subseteq E} \frac{\min\{|X'| \mid X' 6a3a X\}}{\max\{|X'| \mid X' 6a3a X\}} = 1$
- E строки матрицы, \mathcal{F} всевозможные л.н.з. наборы строк (матричный матроид)
- E конечное множество, $\mathcal{F} \coloneqq \{E' \subseteq E \mid |E'| \le k\}$ (однородный матроид)

Примеры матроидов

- 1. $\forall X, Y \in \mathcal{F} \quad (|X| > |Y| \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$
- 2. $\forall X, Y \in \mathcal{F} \quad (|X| = |Y| + 1 \Rightarrow \exists x \in X \setminus Y \text{ т.ч. } Y \cup \{x\} \in \mathcal{F})$
- 3. Для любого $X \subseteq E$ все базы X равномощны.

4.
$$q(E, \mathcal{F}) := \min_{X \subseteq E} \frac{\min\{|X'| \mid X' - \text{база } X\}}{\max\{|X'| \mid X' - \text{база } X\}} = 1$$

• E — рёбра графа, $\mathcal F$ — наборы рёбер, образующие леса



Жадный алгоритм для задачи DLS

Жадный алгоритм решения DLS-задачи максимизации для наследственной системы (E,\mathcal{F}) с неотрицательными весами:

```
1. A \coloneqq \emptyset

2. if \exists s \in E : A \cup \{s\} \in \mathcal{F} then A := A \cup \{argmax w(s)\} goto 2 else: return A
```

Жадный алгоритм для задачи DLS

Какие требования наложить на структуру (E,\mathcal{F}) , чтобы жадный алгоритм давал оптимальное решение задачи оптимизации? Ответ: пара (E,\mathcal{F}) должна быть матроидом.

Теорема. (T.A. Jenkyns '1976, B. Korte и D. Hausmann '1978)

(Обобщение теоремы R. Rado '1957 и J. Edmonds '1971)

Пусть в задаче максимизации на наследственной системе (E,\mathcal{F})

- w^* оптимальное значение,
- $w_{\rm жал}$ значение, полученное жадным алгоритмом.

Тогда

$$q(E,\mathcal{F}) \leq \frac{w_{\text{жад}}}{w^*} \leq 1,$$

причём для любой системы (E,\mathcal{F}) можно указать веса, на которых нижняя граница достигается.

Пусть $E = \{e_1, e_2, \dots, e_n\}$, причём б.о.о. $w(e_1) \ge \dots \ge w(e_n)$.

Обозначим
$$\Delta_k\coloneqq w(e_k)-w(e_{k+1}).$$
 $(\Delta_n\coloneqq w(e_n))$

Обозначим $E_k \coloneqq \{e_1, e_2, \dots, e_k\}$, и, по определению, $E_0 \coloneqq \emptyset$.

Пусть G — допустимое множество, выданное ж.а., и пусть $G_k \coloneqq G \cap E_k$.

Пусть O — оптимальное допустимое множество, и пусть $O_k \coloneqq O \cap E_k$.

Имеем

$$w(G) = \sum_{k=1}^{n} (|G_k| - |G_{k-1}|) \cdot w(e_k) = \sum_{k=1}^{n} |G_k| \cdot \Delta_k.$$

•
$$w(e_1) \ge \cdots \ge w(e_n)$$
, $\Delta_k := w(e_k) - w(e_{k+1})$,
• $E_k := \{e_1, e_2, \dots, e_k\}$, $G_k := G \cap E_k$, $O_k := 0 \cap E_k$.
• $r\bar{k}X := min\{|X'| \mid X' - 6a3aX\}$
• $rkX := max\{|X'| \mid X' - 6a3aX\}$
• $q(E, \mathcal{F}) := \min_{X \subseteq E} \frac{r\bar{k}X}{rkX}$
• $w(G) = \sum_{k=1}^{n} |G_k| \cdot \Delta_k \ge \sum_{k=1}^{n} r\bar{k}E_k \cdot \Delta_k \ge \sum_{k=1}^{n} q(E, \mathcal{F}) \cdot r\bar{k}E_k \cdot \Delta_k = q(E, \mathcal{F}) \cdot w(O)$

Осталось привести пример, когда оценка достигается.

Пусть $X\subseteq E$, и пусть \tilde{B} и B — такие базы X , для которых

$$\frac{\left|\tilde{B}\right|}{\left|B\right|}=q(E,\mathcal{F}).$$

Положим

$$w(e) \coloneqq \begin{cases} 1 & \text{при } e \in X \\ 0 & \text{иначе} \end{cases}$$

Если ж.а. не повезёт, и он выберет вначале все элементы \tilde{B} , то не сможет к ним ничего положительного добавить.

Жадность vs. локальность

Утверждение (упражнение).

На матроиде корректно будет работать такой алгоритм локального поиска:

- 1. *А* ≔ ∀база
- 2. $\mathcal{N}(A) \coloneqq \{A' \in \mathcal{F} \mid A' = (A \setminus \{s\}) \cup \{s'\}, \ s' \notin A\}$
- 3. if $\exists A' \in \mathcal{N}(A)$: w(A') > w(A) then $A \coloneqq A'$ goto 2

Жадный алгоритм в задаче MST

- Дан граф с весами на рёбрах.
- Требуется выбрать дерево, покрывающее все вершины и имеющее как можно меньший вес
- Формальная постановка:
 - $E \coloneqq E(G)$
 - $w(s) \coloneqq (bigconst вес ребра s в исходном графе)$
 - $\mathcal{F} = \{E' \subseteq E \mid E'$ образует остовное дерево в $G\}$
 - Трюк с изменением весов работает потому, что количество рёбер во всех остовных деревьях одинаковое!

Жадный алгоритм для MST

• Дан граф с весами на рёбрах

• Требуется выбрать дерево, покрывающее все вершины

и имеющее как можно меньший вес

- Матроидная постановка:
 - *S* множество всех рёбер графа
 - \mathcal{F} содержит все ациклические подмножества рёбер

Теорема (D. Hausmann, T.A. Jenkyns, B. Korte '1980).

Пусть (E,\mathcal{F}) — наследственная система.

Если для любого $A \in \mathcal{F}$ и любого $e \in E$ во множестве $A \cup \{e\}$ не более t циклов, то $q(E,\mathcal{F}) \geq \frac{1}{t}$.

Пусть $X \subseteq E$ и пусть \tilde{B} и B — наименьшая и наибольшая по мощности базы X.

Докажем, что $|B| \leq t \cdot |\tilde{B}|$.

Пусть $B_{\text{обш}} \coloneqq B \cap \tilde{B}$.

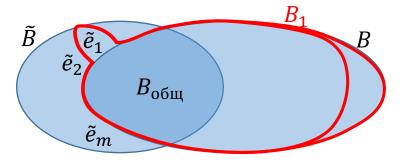
Пусть $\tilde{B} \setminus B_{\text{общ}} = \{\tilde{e}_1, \dots, \tilde{e}_m\}.$

Положим $B_0 \coloneqq B$.

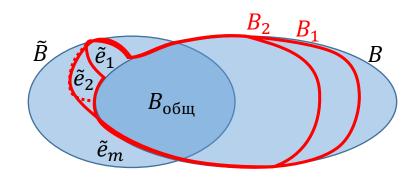
По условию, в $B_0 \cup \{\tilde{e}_1\}$ не более t циклов. У каждого из них найдётся общий элемент с $B_0 \setminus \left(B_{\text{общ}} \cup \{\tilde{e}_1\}\right)$. Удалив не более t таких элементов, получим множество B_1 ,

такое, что

- B_1 независимо,
- $B_{\text{общ}} \cup \{\tilde{e}_1\} \subseteq B_1 \subseteq X$,
- $|B_1 \setminus (B_{\text{общ}} \cup \{\tilde{e}_1\})| \ge |B \setminus B_{\text{общ}}| t$.



- B_1 независимо,
- $B_1 \supseteq B_{\text{общ}} \cup \{\tilde{e}_1\},$
- $|B_1 \setminus B_{\text{общ}}| \ge |B_0 \setminus B_{\text{общ}}| t$.



В B_1 U $\{ ilde{e}_2\}$ не более t циклов. У каждого из них найдётся общий элемент с $B_1\setminus \left(B_{\mathrm{общ}}\cup\{ ilde{e}_1, ilde{e}_2\}\right)$. Удалив не более t таких элементов, получим множество B_2 , такое, что

- B_2 независимо,
- $B_{\text{общ}} \cup \{\tilde{e}_1, \tilde{e}_2\} \subseteq B_2 \subseteq X$,
- $|B_2 \setminus (B_{\text{общ}} \cup \{\tilde{e}_1, \tilde{e}_2\})| \ge |B_1 \setminus (B_{\text{общ}} \cup \{\tilde{e}_1\})| t \ge |B \setminus B_{\text{общ}}| 2t$.

Получим в итоге независимое множество B_m , такое, что

- B_m независимо,
- $\tilde{B} \subseteq B_m \subseteq X$,
- $|B_m \setminus \tilde{B}| \ge |B \setminus B_{\text{общ}}| t \cdot m$.

Так как по условию, \tilde{B} — база X, то $B_m=\tilde{B}$, следовательно $\left|B\setminus B_{\mathrm{общ}}\right|\leq t\cdot m=t\cdot \left|\tilde{B}\setminus B_{\mathrm{общ}}\right|.$

Следовательно,

$$|B| - |B_{\text{общ}}| \le t \cdot (|\tilde{B}| - |B_{\text{общ}}|)$$

И

$$|B| \le t \cdot |\tilde{B}| + (1-t) \cdot |B_{\text{общ}}| \le t \cdot |\tilde{B}|.$$

Субмодулярность ранговой функции

Утверждение.

Если (E,\mathcal{F}) — матроид, то для любых $X,Y\subseteq E$ выполнено $\operatorname{rk} X + \operatorname{rk} Y \geq \operatorname{rk}(X\cup Y) + \operatorname{rk}(X\cap Y)$.

Доказательство:

Пусть A — база $X \cap Y$.

По свойству матроида, можно расширить A до базы $(A \sqcup B)$ множества X, а затем до базы $(A \sqcup B \sqcup C)$ множества $X \cup Y$.

Тогда $A \sqcup C$ — независимое подмножество в Y, поэтому $\operatorname{rk} X + \operatorname{rk} Y \geq |A \sqcup B| + |A \sqcup C| = |A \sqcup B \sqcup C| + |A| = \operatorname{rk}(X \cup Y) + \operatorname{rk}(X \cap Y)$.

Свойство циклов матроида

Утверждение.

Пусть C', C'' — два различных цикла матроида.

Для любого $x \in C' \cap C''$ и любого $y \in C' \setminus C''$ найдётся цикл C, такой, что

$$y \in C \subseteq (C' \cup C'') \setminus \{x\}.$$

Доказательство.

По субмодулярности получаем

$$\operatorname{rk} C' + \operatorname{rk} ((C' \cup C'') \setminus \{x, y\}) + \operatorname{rk} C'' \ge$$

$$\ge \operatorname{rk} C' + \operatorname{rk} ((C' \cup C'') \setminus \{y\}) + \operatorname{rk} (C'' \setminus \{x\})$$

$$\ge \operatorname{rk} (C' \cup C'') + \operatorname{rk} (C' \setminus \{y\}) + \operatorname{rk} (C'' \setminus \{x\})$$

Свойство циклов матроида

Для любого $x \in C' \cap C''$ и любого $y \in C' \setminus C''$ найдётся цикл C, такой, что $y \in C \subseteq (C' \cup C'') \setminus \{x\}$.

• $\operatorname{rk} C' + \operatorname{rk} ((C' \cup C'') \setminus \{x, y\}) + \operatorname{rk} C'' \ge \operatorname{rk} (C' \cup C'') + \operatorname{rk} (C' \setminus \{y\}) + \operatorname{rk} (C'' \setminus \{x\})$

Из соотношений $\operatorname{rk} C' = \operatorname{rk}(C' \setminus \{y\})$ и $\operatorname{rk} C'' = \operatorname{rk}(C'' \setminus \{x\})$ следует

$$\operatorname{rk}((C' \cup C'') \setminus \{x, y\}) = \operatorname{rk}(C' \cup C'').$$

Пусть B — база множества $(C' \cup C'') \setminus \{x, y\}$.

Тогда $B \cup \{y\}$ содержит искомый цикл.

Ещё одно эквивалетное определение матроида

Теорема (D. Hausmann, T.A. Jenkyns, B. Korte '1980).

Пусть (E, \mathcal{F}) — наследственная система.

Если для любого $A \in \mathcal{F}$ и любого $e \in E$ во множестве $A \cup \{e\}$ не более t циклов, то $q(E,\mathcal{F}) \geq \frac{1}{t}$.

Утверждение.

Наследственная система (E,\mathcal{F}) является матроидом т. и т.т., когда для любого $A \in \mathcal{F}$ и любого $e \in E$ множество $A \cup \{e\}$ содержит не более одного цикла.

Доказательство. В одну сторону — прямое следствие теоремы. Нужно доказать в другую сторону.

Ещё одно эквивалетное определение матроида

Утверждение.

Наследственная система (E,\mathcal{F}) является матроидом т. и т.т., когда для любого $A \in \mathcal{F}$ и любого $e \in E$ множество $A \cup \{e\}$ содержит не более одного цикла.

Пусть (E, \mathcal{F}) — матроид, и пусть $A \in \mathcal{F}$.

Допустим, в $A \cup \{e\} \in \mathcal{F}$ есть пара различных циклов: C', C''.

Тогда во множестве $(C' \cup C'') \setminus \{e\}$ есть цикл — но это противоречит тому, что $(C' \cup C'') \setminus \{e\} \subseteq A \in \mathcal{F}$.

Пересечение наследственных систем

Пересечение наследственных систем (E, \mathcal{F}') и (E, \mathcal{F}'') — это наследственная система (E, \mathcal{F}) , в которой

$$\mathcal{F} \coloneqq \mathcal{F}' \cap \mathcal{F}''$$

Утверждение.

Любая наследственная система (E,\mathcal{F}) может быть представлена как пересечение нескольких матроидов.

Доказательство:

Пусть C — произвольный цикл в (E,\mathcal{F}) . Определим $\mathcal{F}_{\mathcal{C}} \coloneqq \{A \subseteq E \mid \mathcal{C} \setminus A \neq \emptyset\}$.

Легко проверить, что (E, \mathcal{F}_C) — матроид.

Поэтому $(E,\mathcal{F}) = \bigcap_C (E,\mathcal{F}_C)$ — искомое представление.

Утверждение.

Если наследственная система (E,\mathcal{F}) может быть представлена как пересечение t матроидов, то $q(E,\mathcal{F}) \geq \frac{1}{t}$.

Доказательство:

Пусть (E, \mathcal{F}_i) — матроиды, пересечение которых даёт (E, \mathcal{F}) . Пусть $X \in \mathcal{F}$ и пусть $e \in E \setminus X$.

- Чтобы подмножество $C \subseteq X \cup \{e\}$ было циклом в (E, \mathcal{F}) , оно должно быть циклом хотя бы в одном (E, \mathcal{F}_i) .
- Каждое (E, \mathcal{F}_i) матроид, и в нём не более одного цикла в $X \cup \{e\}$.

Единственность решения DLS

• Решаем задачу:

- E конечное множество, для $X \subseteq E$ полагаем $w(X) \coloneqq \sum_{x \in X} w(x)$
- \mathcal{F} семейство «допустимых» подмножеств E (Пара E, \mathcal{F} не обязана быть наследственной системой!)
- Требуется найти $X^* \in \mathcal{F}$, такое, что $w(X^*) \to \max$

• Вопросы:

- Когда решение такой задачи единственное?
- Что будет, если w(x) выбирать для каждого $x \in E$ случайным образом, например, из множества $\{1,2,...,M\}$?

Лемма об изолировании (isolation lemma)

Теорема. (K. Mulmuley, U. Vazirani, V. Vazirani '1987)

При случайном равномерном независимом выборе весов элементов E из M-элементного множества вероятность единственности решения оптимизационной задачи не меньше

$$1-\frac{|E|}{M}.$$

Происхождение названия леммы: набор весов называется uзолирующим для семейства \mathcal{F} , если решение оптимизационной задачи на данном наборе весов единственное.

Доказательство леммы об изолировании

Для $s \in E$ рассмотрим величину

$$\alpha(s) \coloneqq \max_{X \in \mathcal{F}: X \not\ni s} w(X) - \max_{Y \in \mathcal{F}: Y \ni s} w(Y \setminus \{s\})$$

Допустим, есть два различных множества $X',Y'\in\mathcal{F}$, на которых достигается максимум веса.

Тогда рассмотрим произвольный $s \in Y' \setminus X'$. Для такого s выполнено

$$\alpha(s) = \max_{X \in \mathcal{F}: X \not\ni s} w(X) - \max_{Y \in \mathcal{F}: Y \ni s} w(Y \setminus \{s\}) = w(X') - w(Y' \setminus \{s\}) =$$

$$= w(X') - (w(Y') - w(s)) = w(s)$$

Следовательно, если $\forall s \in E \ w(s) \neq \alpha(s)$, то решение оптимизационной задачи единственное.

Доказательство леммы об изолировании

- Если $\forall s \in E$ $w(s) \neq \alpha(s)$, то решение задачи единственное.
- $\alpha(s) \coloneqq \max_{X \in \mathcal{F}: X \not\ni s} w(X) \max_{Y \in \mathcal{F}: Y \ni s} w(Y \setminus \{s\})$

Заметим, что $\alpha(s)$ не зависит от w(s), поэтому

$$\Pr\{w(s) = \alpha(s)\} \le \frac{1}{M}$$

Теперь можно оценить вероятность единственности решения:

$$\Pr\{\forall s \in E \ w(s) \neq \alpha(s)\} =$$

$$= 1 - \Pr\{\exists s \in E: \ w(s) = \alpha(s)\} \ge$$

$$\geq 1 - \sum_{s \in E} \Pr\{w(s) = \alpha(s)\} \geq 1 - \frac{|E|}{M}$$

Резюме

- Жадность простой подход, эффективный в некоторых случаях (например, в оптимизационных задачах на матроидах), но, конечно, далеко не во всех
- Во многих оптимизационных задачах при случайном выборе весов решение с большой вероятностью единственное