Matroids and combinatorial optimization

Combinatoria 2022/2023

Lorenzo Gregoris (1867373)

January 10, 2024

SAPIENZA
UNIVERSITÀ DI ROMA

Table of Contents

Combinatorial optimization

Many combinatorial optimization problems can be formulated as follows:

- a set system $(E, \mathcal{F})$, i.e. a finite set E and some $\mathcal{F} \subseteq 2^E$;
- a cost function $c : \mathcal{F} \mapsto \mathbb{R}$;

Find an element of $\mathcal{F}$ whose cost is minimum or maximum.

Goal

We will introduce **independence systems** and **matroids** and show that many combinatorial optimization problems can be described in this context.
The main reason why matroids are important is that a **simple greedy algorithm can be used for optimization over matroids**.

Table of Contents

2 Indipendence systems and matroids

Definition (Indipendence system)

A set system $(E, \mathcal{F})$ is an **indipendence system** if and only if:

- $(M1)$ $\emptyset \in \mathcal{F}$;
- $(M2)$ if $X \subseteq Y$, $Y \in \mathcal{F}$ then $X \in \mathcal{F}$.

Definition (Indipendence system)

A set system $(E, \mathcal{F})$ is an **indipendence system** if and only if:

- $(M1)$ $\emptyset \in \mathcal{F}$;
- $(M2)$ if $X \subseteq Y$, $Y \in \mathcal{F}$ then $X \in \mathcal{F}$.

The elements of $\mathcal{F}$ are called *independent*, the elements of $2^E \setminus \mathcal{F}$ *dependent*. Maximal independent sets are called **bases**. For $X \subseteq E$, the maximal independent subsets of X are called bases of X.

Example

Consider any finite subset E of a vector space V. We can define an independence system by choosing $\mathcal{F} = \{Y \in 2^E \ : \ Y \text{ are linearly independent } \}$.
This set system clearly satisfies $(M1)$ (there is no way to obatin $\vec{0}$ from a linear combination of an empty set of vectors) and $(M2)$.
Keeping in mind this example, we introduce some definitions.

Definition (Rank)

Let $(E, \mathcal{F})$ be an independence system. For $X \subseteq E$ we define the rank of X by:

$$r(X) := \max\{|Y| : Y \subseteq X,\ Y \in \mathcal{F}\}.$$

The rank of a finite collection of vectors X is just the size of any basis for X. Does this hold for any indipendence system? No.

Counter example

Consider the indipendence set system $(E, \mathcal{F})$ where $E = \{1, 2, 3\}$ and
$\mathcal{F} = \Big\{\emptyset, \{1, 2\}, \{1\}, \{2\}, \{3\}\Big\}$
We have two basis for E with different cardinality!
This suggests that subsets of linearly indipendent vectors have another property, in addition to $(M1)$ and $(M2)$. This is the *augmentation property*: Given two finite subsets of linearly indipendent vectors X and Y such that $|X| < |Y|$, we can find some vectors of Y to add to X such that they have the same rank.

Definition (Matroid)

An indipendence system is a **matroid** if and only if:

- $(M3)$ if $X, Y \in \mathcal{F}$, $|X| < |Y|$ then $\exists y \in Y \setminus X$ such that $X \cup \{y\} \in \mathcal{F}$.

Introduced in a paper called *"On the Abstract Properties of Linear Dependence"* by Hassler Whitney in 1935.

Equivalent axiom

This new axiom captures the propriety we were looking for:

Theorem

$(M3) \iff$ *For each $X \subseteq E$, all bases of X have the same cardinality.*

Equivalent axiom

This new axiom captures the propriety we were looking for:

Theorem

$(M3) \iff$ *For each $X \subseteq E$, all bases of X have the same cardinality.*

Proof

$\implies$ Suppose a subset $X \subseteq E$ has two bases of different size $|B_1| > |B_2|$, we can add elements of the bigger base to the small and still be an indipendent set, but B_2 was choosen to be maximal.

Equivalent axiom

This new axiom captures the propriety we were looking for:

Theorem

$(M3) \iff$ *For each $X \subseteq E$, all bases of X have the same cardinality.*

Proof

$\implies$ Suppose a subset $X \subseteq E$ has two bases of different size $|B_1| > |B_2|$, we can add elements of the bigger base to the small and still be an indipendent set, but B_2 was choosen to be maximal.

$\impliedby$ Let $X, Y \in \mathcal{F}$ such that $|X| > |Y|$. This means Y can't be a basis for $Y \cup X$. This means there must be an $x \in X \setminus Y$ such that $Y \cup \{x\} \in \mathcal{F}$. $\square$

Bases

Given a set of bases we can construct the remaining indipendent sets by adding all of their subsets.

Bases

Given a set of bases we can construct the remaining indipendent sets by adding all of their subsets.
Given a set of bases, when they generate a matroid?

Bases

Given a set of bases we can construct the remaining indipendent sets by adding all of their subsets.
Given a set of bases, when they generate a matroid?

Theorem (Bases characterization)

Let E be a finite set and $\mathcal{B} \subseteq 2^E$. $\mathcal{B}$ is the set of bases of some matroid $(E, \mathcal{F})$ if and only if:

- $(B1)$ $\mathcal{B} \neq \emptyset$.
- $(B2)$ *For any* $A, B \in \mathcal{B}$ *and* $a \in A \setminus B$ *there exists* $b \in B \setminus A$ *such that* $(A \setminus \{a\}) \cup \{b\} \in \mathcal{B}$.

Given a set of bases we can construct the remaining indipendent sets by adding all of their subsets.
Given a set of bases, when they generate a matroid?

Theorem (Bases characterization)

Let E be a finite set and $\mathcal{B} \subseteq 2^E$. $\mathcal{B}$ is the set of bases of some matroid $(E, \mathcal{F})$ if and only if:

- $(B1)$ $\mathcal{B} \neq \emptyset$.
- $(B2)$ *For any $A, B \in \mathcal{B}$ and $a \in A \setminus B$ there exists $b \in B \setminus A$ such that $(A \setminus \{a\}) \cup \{b\} \in \mathcal{B}$.*

$(B2)$ is called the *base exchange property*.

Proof $\implies$ The set of bases of a matroid satisfies $(B1)$ by $(M1)$, and $(B2)$: the set $A \setminus \{a\}$ is still indipendent for $(M2)$, so by $(M3)$ we can find an element in $B \setminus A$ to create a new basis.
$\impliedby$ Combinatorial Optimization. Theory and Algorithms (Bernhard Korte, Jens Vygen), theorem 13.9, page 337. $\square$

Properties of the rank function

Theorem

Let $\mathcal{M}$ be a matroid on a finite set E. The rank function r of $\mathcal{M}$ has the following properties:

- $(R1)$ $0 \leq r(X) \leq |X|$ *for all* $X \subseteq E$.
- $(R2)$ *If* $X \subseteq Y \subseteq E$ *then* $r(X) \leq r(Y)$.
- $(R3)$ *For every* $X, Y \subseteq E$ $r(X \cup Y) + r(X \cap Y \leq r(X) + r(Y)$.

Theorem

Let $\mathcal{M}$ be a matroid on a finite set E. The rank function r of $\mathcal{M}$ has the following properties:

- $(R1)$ $0 \leq r(X) \leq |X|$ *for all* $X \subseteq E$.
- $(R2)$ *If* $X \subseteq Y \subseteq E$ *then* $r(X) \leq r(Y)$.
- $(R3)$ *For every* $X, Y \subseteq E$ $r(X \cup Y) + r(X \cap Y \leq r(X) + r(Y)$.

Proof $(R1)$ and $(R2)$ follow from the definition, they are satisfied in any indipendence system.

Given a finite set E and a function $r : 2^E \mapsto \mathbb{N}$, we can define a matroid in the following way:

$$\mathcal{F} := \{F \subseteq E \ : \ r(F) = |F|\}$$

Corollary

Let E be a finite set. A function $r : 2^E \mapsto \mathbb{N}$ is the rank function of a matroid on E if and only if r satisfies properties $(R1)$, $(R2)$ and $(R3)$.

Proof Combinatorial Optimization. Theory and Algorithms (Bernhard Korte, Jens Vygen), Page 338. □

Rank quotient

We introduce another definition of rank of subsets for indipendent set system, the size of the smallest base contained:

Definition (Lower rank)

Let $(E, \mathcal{F})$ be an independence system. For $X \subseteq E$ we define the lower rank by:

$$\rho(X) := \min\{|Y| : Y \subseteq X,\ Y \in \mathcal{F} \text{ and } Y \cup \{x\} \notin \mathcal{F}\ \forall x \in X \setminus Y\}.$$

Rank quotient

We introduce another definition of rank of subsets for indipendent set system, the size of the smallest base contained:

Definition (Lower rank)

Let $(E, \mathcal{F})$ be an independence system. For $X \subseteq E$ we define the lower rank by:

$$\rho(X) := \min\{|Y| : Y \subseteq X,\ Y \in \mathcal{F} \text{ and } Y \cup \{x\} \notin \mathcal{F}\ \forall x \in X \setminus Y\}.$$

Definition (Rank quotient)

The rank quotient of $(E, \mathcal{F})$ is defined by:

$$q(E, \mathcal{F}) := \min_{X \subseteq E} \frac{\rho(X)}{r(X)}$$

It follows from the definition that $q(E, \mathcal{F}) \leq 1$, and an indipendence system is a matroid if and only if $q(E, \mathcal{F}) = 1$, since this is equivalent to the statment that every bases for any $X \subseteq E$ has the same cardinality.

$(E, \mathcal{F})$ **is a Matroid** $\iff q(E, \mathcal{F}) = 1$

Table of Contents

Maximization and minimization problems for IS

Definition (Maximization problem)

Istance: A set system $(E, \mathcal{F})$ and a cost function $c : E \mapsto \mathbb{R}$.
Task: Find $X \in \mathcal{F}$ such that $c(X) := \sum_{e \in X} c(e)$ is maximum.

Definition (Maximization problem)

Istance: A set system $(E, \mathcal{F})$ and a cost function $c : E \mapsto \mathbb{R}$.
Task: Find $X \in \mathcal{F}$ such that $c(X) := \sum_{e \in X} c(e)$ is maximum.

Definition (Minimization problem)

Istance: A set system $(E, \mathcal{F})$ and a cost function $c : E \mapsto \mathbb{R}$.
Task: Find a basis B such that $c(B) := \sum_{e \in B} c(e)$ is minimum.

Note: $c(\emptyset) = 0$.

There is a natural duality between minimization and maximization problems on indipendence systems: finding a minimum cost basis B in $(E, \mathcal{F})$ is the same as finding the maximum cost complement $E \setminus B$. This suggests the following definition:

There is a natural duality between minimization and maximization problems on indipendence systems: finding a minimum cost basis B in $(E, \mathcal{F})$ is the same as finding the maximum cost complement $E \setminus B$. This suggests the following definition:

Definition (Dual IS)

Let $(E, \mathcal{F})$ be an indipendence system. We define its **dual** $(E, \mathcal{F}^*)$ so that the bases in $\mathcal{F}^*$ are exaclty the complements of the bases of $\mathcal{F}$.

There is a natural duality between minimization and maximization problems on indipendence systems: finding a minimum cost basis B in $(E, \mathcal{F})$ is the same as finding the maximum cost complement $E \setminus B$. This suggests the following definition:

Definition (Dual IS)

Let $(E, \mathcal{F})$ be an indipendence system. We define its **dual** $(E, \mathcal{F}^*)$ so that the bases in $\mathcal{F}^*$ are exaclty the complements of the bases of $\mathcal{F}$.

It's clear that $(E, \mathcal{F}^{**}) = (E, \mathcal{F})$.

So which sets are in $\mathcal{F}^*$?

So which sets are in $\mathcal{F}^*$? Every subset of the dual basis $\mathcal{B}^*$, this means:

$$\mathcal{F}^* = \{F \subseteq E \ : \ \exists B \in \mathcal{B} \text{ s.t. } B \cap F = \emptyset\}$$

Properties of the dual

Theorem

Let $(E, \mathcal{F})$ be an indipendence system, $(E, \mathcal{F}^*)$ its dual, and let r and r^* be the corresponding rank functions.

1. *$(E, \mathcal{F})$ is a matroid if and only if $(E, \mathcal{F}^*)$ is a matroid.*
2. *If $(E, \mathcal{F})$ is a matroid, then $r^*(E) = |F| + r(E \setminus F) - r(E)$ for $F \subseteq E$.*

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.
Let $(E, \mathcal{F})$ be a matroid with bases $\mathcal{B}$ and $(E, \mathcal{F}^*)$ its dual.

$$r^*(F) = \max_{I \subseteq F,\, I \in \mathcal{F}^*} |I|$$

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.
Let $(E, \mathcal{F})$ be a matroid with bases $\mathcal{B}$ and $(E, \mathcal{F}^*)$ its dual.

$$\begin{aligned} r^*(F) &= \max_{I \subseteq F,\, I \in \mathcal{F}^*} |I| \\ &= \max_{B \in \mathcal{B}} |F \cap (E \setminus B)| \end{aligned}$$

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.
Let $(E, \mathcal{F})$ be a matroid with bases $\mathcal{B}$ and $(E, \mathcal{F}^*)$ its dual.

$$\begin{aligned} r^*(F) &= \max_{I \subseteq F, I \in \mathcal{F}^*} |I| \\ &= \max_{B \in \mathcal{B}} |F \cap (E \setminus B)| \\ &= \max_{B \in \mathcal{B}} |F \setminus B| \end{aligned}$$

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.
Let $(E, \mathcal{F})$ be a matroid with bases $\mathcal{B}$ and $(E, \mathcal{F}^*)$ its dual.

$$\begin{aligned} r^*(F) &= \max_{I \subseteq F,\, I \in \mathcal{F}^*} |I| \\ &= \max_{B \in \mathcal{B}} |F \cap (E \setminus B)| \\ &= \max_{B \in \mathcal{B}} |F \setminus B| \\ &= |F| - \min_{B \in \mathcal{B}} |F \cap B| \end{aligned}$$

Proof We only give a proof of 2. The idea to prove 1. is via the rank function characterization: verify that r^* satisfies $(R1)$, $(R2)$ and $(R3)$.
Let $(E, \mathcal{F})$ be a matroid with bases $\mathcal{B}$ and $(E, \mathcal{F}^*)$ its dual.

$$\begin{aligned} r^*(F) &= \max_{I \subseteq F,\, I \in \mathcal{F}^*} |I| \\ &= \max_{B \in \mathcal{B}} |F \cap (E \setminus B)| \\ &= \max_{B \in \mathcal{B}} |F \setminus B| \\ &= |F| - \min_{B \in \mathcal{B}} |F \cap B| \\ &= |F| - |B| + \max_{B \in \mathcal{B}} |(E \setminus F) \cap B| \\ &= |F| - r(E) + r(E \setminus F) \qquad \square \end{aligned}$$

- *Maximum weight stable set problem:* given a graph G and weights $c : V(G) \mapsto \mathbb{R}$, find a stable set $X \subseteq V(G)$ of maximum weight. Here $E = V(G)$ and $\mathcal{F} = \{F \subseteq E : F \text{ is stable in } G\}$.

Examples

- *Maximum weight stable set problem:* given a graph G and weights $c : V(G) \mapsto \mathbb{R}$, find a stable set $X \subseteq V(G)$ of maximum weight. Here $E = V(G)$ and $\mathcal{F} = \{F \subseteq E : F \text{ is stable in } G\}$.
- *TSP:* given a complete undirected graph G and weights $c : E(G) \mapsto \mathbb{R}^+$, find a minimum weight Hamiltonian circuit in G. Here $\mathcal{F} = \{F \subseteq E : F \text{ is a subset of edges in an Hamiltonian circuit in } G\}$.

Examples

3 Maximization and minimization problems

- *Maximum weight stable set problem:* given a graph G and weights $c : V(G) \mapsto \mathbb{R}$, find a stable set $X \subseteq V(G)$ of maximum weight. Here $E = V(G)$ and $\mathcal{F} = \{F \subseteq E : F \text{ is stable in } G\}$.
- *TSP:* given a complete undirected graph G and weights $c : E(G) \mapsto \mathbb{R}^+$, find a minimum weight Hamiltonian circuit in G. Here $\mathcal{F} = \{F \subseteq E : F \text{ is a subset of edges in an Hamiltonian circuit in } G\}$.
- *MST:* given a connected undirected graph G and weights $c : E(G) \mapsto \mathbb{R}$, find a minimum weight spanning tree in G. Here $E = E(G)$ and the elements of $\mathcal{F}$ are the edge sets of the forests in G.

The first two problems are NP-hard, while the last one is in P. The indipendence sytems of the first two problems are not matroids, while the last one is! This isn't a coincidence, we will show that simple (polynomial) algorithms solve optimization problems on matroids.

The IS of MST is a matroid

Proposition

The set system $(E, \mathcal{F})$, where E is the set of edges of some undirected graph G and $\mathcal{F} := \{F \subseteq E \ : \ (V(G), F) \text{ is a forest } \}$ is a matroid.

Proposition

The set system $(E, \mathcal{F})$, where E is the set of edges of some undirected graph G and $\mathcal{F} := \{F \subseteq E \ : \ (V(G), F)$ is a forest $\}$ is a matroid.

Proof It is obvious that is an independence sytem, it remains to show that $(M3)$ holds. We will prove its contrapositive.

The IS of MST is a matroid

Proposition

The set system $(E, \mathcal{F})$, where E is the set of edges of some undirected graph G and $\mathcal{F} := \{F \subseteq E \,:\, (V(G), F) \text{ is a forest }\}$ is a matroid.

Proof It is obvious that is an independence sytem, it remains to show that $(M3)$ holds. We will prove its contrapositive.Let $X, Y \in \mathcal{F}$, and suppose $Y \cup \{x\} \notin \mathcal{F}$ for all $x \in X \setminus Y$. We show $|X| \leq |Y|$. For each $x = \{u, v\} \in X$, u and v are in the same connected component of $(V(G), Y)$. Hence each connected component $Z \subseteq (V(G), X)$ is a subset of a connected component of $(V(G), Y)$. So the number of connected component p of the forest $(V(G), X)$ is greater or equal to the number q of connected components of the forest $(V(G), Y)$. But $|V(G)| - |X| = p \geq q = |V(G)| - |Y|$, implying $|X| \leq |Y|$. $\square$

Table of Contents

4 Greedy algorithms

Again, let $(E, \mathcal{F})$ be an independence system and $c : E \mapsto \mathbb{R}^+$ (non-negative weights). We consider the **Maximization problem** for $(E, \mathcal{F}, c)$ and formulate two “greedy” algorithms.

Again, let $(E, \mathcal{F})$ be an independence system and $c : E \mapsto \mathbb{R}^+$ (non-negative weights).
We consider the **Maximization problem** for $(E, \mathcal{F}, c)$ and formulate two "greedy" algorithms.
The first one, called **Best-In** builds the solution choosing at each step the best possible element.

Greedy algorithms

Again, let $(E, \mathcal{F})$ be an independence system and $c : E \mapsto \mathbb{R}^+$ (non-negative weights).
We consider the **Maximization problem** for $(E, \mathcal{F}, c)$ and formulate two “greedy” algorithms.
The first one, called **Best-In** builds the solution choosing at each step the best possible element.
The second, called **Worst-Out** at each step removes the worst element, until a feasible solution is atteined.

Again, let $(E, \mathcal{F})$ be an independence system and $c : E \mapsto \mathbb{R}^+$ (non-negative weights).
We consider the **Maximization problem** for $(E, \mathcal{F}, c)$ and formulate two "greedy" algorithms.
The first one, called **Best-In** builds the solution choosing at each step the best possible element.
The second, called **Worst-Out** at each step removes the worst element, until a feasible solution is atteined.
They boath employ the use of an oracle to decide if a given subset $X \subseteq E$ is feasible.

The best-in greedy algorithm for maximization

Best-In-Greedy algorithm

Input: An independence system $(E, \mathcal{F})$, given by an independence oracle, weights $c : E \mapsto \mathbb{R}^+$.

Output: A set $F \in \mathcal{F}$.

1. Sort E such that $c(e_1) \geq c(e_2) \geq \cdots \geq c(e_n)$.
2. Set $F \leftarrow \emptyset$.
3. **For** i = 1 to n **do**:
   If $F \cup \{e_i\} \in \mathcal{F}$ **then** set $F \leftarrow F \cup \{e_i\}$

Best-In-Greedy algorithm

Input: An independence system $(E, \mathcal{F})$, given by an independence oracle, weights $c : E \mapsto \mathbb{R}^+$.

Output: A set $F \in \mathcal{F}$.

1. Sort E such that $c(e_1) \geq c(e_2) \geq \cdots \geq c(e_n)$.
2. Set $F \leftarrow \emptyset$.
3. **For** i = 1 to n **do**:
   If $F \cup \{e_i\} \in \mathcal{F}$ **then** set $F \leftarrow F \cup \{e_i\}$

The output F of Best-In-Greedy algorithm will be a basis for E.

The second algorithm requires a more complicated oracle. Given a set $F \subseteq E$, this oracle decides whether F contains a basis. Let us call such an oracle a basis-superset oracle.

Worst-out-Greedy algorithm

Input: An independence system $(E, \mathcal{F})$, given by a basis-superset oracle, weights $c : E \mapsto \mathbb{R}^+$.

Output: A basis $F \in \mathcal{F}$.

1. Sort E such that $c(e_1) \leq c(e_2) \leq \cdots \leq c(e_n)$.
2. Set $F \leftarrow E$.
3. **For** i = 1 to n **do**:
   If $F \setminus \{e_i\}$ contains a basis **then** set $F \leftarrow F \setminus \{e_i\}$

Greedy algorithms for minimization

One can analogously formulate both greedy algorithms for the Minimization problem.

Greedy algorithms for minimization

One can analogously formulate both greedy algorithms for the Minimization problem. It is easy to see that the Best-In for the Maximization problem for $(E, \mathcal{F}, c)$ corresponds to Worst-Out for the Minimization problem for $(E, \mathcal{F}^*, c)$: adding an element to F in Best-In corresponds to removing an element from F in Worst-out.

Greedy algorithms for minimization

One can analogously formulate both greedy algorithms for the Minimization problem. It is easy to see that the Best-In for the Maximization problem for $(E, \mathcal{F}, c)$ corresponds to Worst-Out for the Minimization problem for $(E, \mathcal{F}^*, c)$: adding an element to F in Best-In corresponds to removing an element from F in Worst-out. This is the dual relationship discussed before.

Quality of the solution

Theorem (Jenkyns 1976, Korte and Hausmann 1978)

Let $(E, \mathcal{F})$ be an independence system. For $c : E \mapsto \mathbb{R}^+$ we denote by $G(E, \mathcal{F}, c)$ the cost of some solution found by the Best-In-Greedy for the Minimization problem, and by $OPT(E, \mathcal{F}, c)$ the cost of an optimum solution. Then

$$q(E, \mathcal{F}) \leq \frac{c(G)}{c(OPT)} \leq 1$$

for all $c : E \mapsto \mathbb{R}^+$. There is a cost function where the lower bound is attained.

Theorem (Jenkyns 1976, Korte and Hausmann 1978)

Let $(E, \mathcal{F})$ be an independence system. For $c : E \mapsto \mathbb{R}^+$ we denote by $G(E, \mathcal{F}, c)$ the cost of some solution found by the Best-In-Greedy for the Minimization problem, and by $OPT(E, \mathcal{F}, c)$ the cost of an optimum solution. Then

$$q(E, \mathcal{F}) \leq \frac{c(G)}{c(OPT)} \leq 1$$

for all $c : E \mapsto \mathbb{R}^+$. There is a cost function where the lower bound is attained.

Proof Let $E = \{e_1, \ldots, e_n\}$ and $c(e_1) \geq c(e_2) \geq \cdots \geq c(e_n)$. Let G_n be the solution found by Best-In-Greedy, while O_n is an optimum solution.

We define $E_j := \{e_1, \dots, e_j\}$, $G_j := G_n \cap E_j$ and $O_j := O_n \cap E_j$. Set $d_n := c(e_n)$ and $d_j := c(e_j) - c(e_{j+1})$ for $j = 1, \dots, n-1$.
Since $O_j \subseteq O_n$, $O_j \in \mathcal{F}$ by $(M2)$. We have $|O_j| \leq r(E_j)$.

We define $E_j := \{e_1, \ldots, e_j\}$, $G_j := G_n \cap E_j$ and $O_j := O_n \cap E_j$. Set $d_n := c(e_n)$ and $d_j := c(e_j) - c(e_{j+1})$ for $j = 1, \ldots, n-1$.
Since $O_j \subseteq O_n$, $O_j \in \mathcal{F}$ by $(M2)$. We have $|O_j| \leq r(E_j)$. Since G_j is a basis for E_j, we have $|G_j| \geq \rho(E_j)$. With these two inequalities we conclude that

We define $E_j := \{e_1, \ldots, e_j\}$, $G_j := G_n \cap E_j$ and $O_j := O_n \cap E_j$. Set $d_n := c(e_n)$ and $d_j := c(e_j) - c(e_{j+1})$ for $j = 1, \ldots, n-1$.
Since $O_j \subseteq O_n$, $O_j \in \mathcal{F}$ by $(M2)$. We have $|O_j| \leq r(E_j)$. Since G_j is a basis for E_j, we have $|G_j| \geq \rho(E_j)$. With these two inequalities we conclude that

$$
\begin{aligned}
c(G_n) &= \sum_{j=1}^{n} (|G_j| - |G_{j-1}|) c(e_j) \\
&= \sum_{j=1}^{n} |G_j| d_j \\
&\geq \sum_{j=1}^{n} \rho(E_j) d_j
\end{aligned}
$$

$$
\begin{aligned}
&\geq q(E, \mathcal{F}) \sum_{j=1}^{n} r(E_j) d_j \\
&\geq q(E, \mathcal{F}) \sum_{j=1}^{n} |O_j| d_j \\
&= q(E, \mathcal{F}) \sum_{j=1}^{n} (|O_j| - |O_{j-1}|) c(e_j) \\
&= q(E, \mathcal{F}) c(O_n)
\end{aligned}
$$

this proves the inequality.

We now prove that the lower bound is tight. Choose $F \subseteq E$ and bases B_1, B_2 of F such that $\frac{|B_1|}{|B_2|} = q(E, \mathcal{F})$, define the cost function:

$$c(e) := \begin{cases} 1 \text{ for } e \in F \\ 0 \text{ for } e \in E \setminus F \end{cases}$$

If we sort $e_1, \ldots, e_n$ such that $B_1 = \{e_1, \ldots e_{|B_1|}\}$ the algorithm will find B_1 as a solution, and the lower bound is attained. $\square$

Rado-Edmonds theorem

Theorem (Rado 1957, Edmonds 1971)

An independence system $(E, \mathcal{F})$ is a matroid if and only if the BEST-IN-GREEDY finds an optimum solution for the MAXIMIZATION PROBLEM for $(E, \mathcal{F}, c)$ for all cost functions $c : E \mapsto \mathbb{R}^+$.

This is one of the rare cases where we can define a structure by its algorithmic behaviour.

Quality of the solution for the minimization problem

Theorem

Korte and Monma 1979 Let $(E, \mathcal{F})$ be an independence system. For $c : E \mapsto \mathbb{R}^+$ let $G(E, \mathcal{F}, c)$ denote a solution found by Worst-Out-Greedy for the Minimization problem. Then

$$1 \leq \frac{G(E, \mathcal{F}, c)}{OPT(E, \mathcal{F}, c)} \leq \max_{F \subseteq E} \frac{|F| - \rho^*(F)}{|F| - r^*(F)}$$

for all $c : E \mapsto \mathbb{R}^+$, where ρ^* and r^* are the rank functions of the dual indipendence system $(E, \mathcal{F}^*)$. There is a cost function where the upper bound is attained.

Proof We use the same notation as in the previous proof. At each step, the partial solution will is $(E \setminus E_j) \cup G_j$, by construction contains a basis of E. Its complement $E_j \setminus G_j$, like G_j in Best-In-Greedy is a basis for E_j in the dual $\mathcal{F}^*$, thus:

$$|E_j| - |G_J| \geq \rho^*(E_j)$$

Likewise, the complement of the solution $E_j \setminus O_j$ is indipendnent in $\mathcal{F}^*$, thus:

$$|E_J| - |O_J| \leq r^*(E_j)$$

We now use the same trick to compute the cost of the solution G_n:

$$
\begin{aligned}
c(G_n) &= \sum_{j=1}^{n} (|G_j| - |G_{j-1}|) c(e_j) \\
&= \sum_{j=1}^{n} |G_j| d_j \\
&\leq \sum_{j=1}^{n} |E_j| - \rho^*(E_j) d_j
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} \left(|E_j| - \rho^*(E_j)\right) \frac{|E_j| - r^*(E_j)}{|E_j| - r^*(E_j)} d_j \\
&\leq \max_{F \subseteq E} \frac{|F| - \rho^*(F)}{|F| - r^*(F)} \sum_{j=1}^{n} |O_j| d_j \qquad \square \\
&= \max_{F \subseteq E} \frac{|F| - \rho^*(F)}{|F| - r^*(F)} \sum_{j=1}^{n} (|O_j| - |O_{j-1}|) c(e_j) \\
&= \max_{F \subseteq E} \frac{|F| - \rho^*(F)}{|F| - r^*(F)} c(O_n)
\end{aligned}
$$

To prove that the upper bound is tight we can use the same strategy as in the Edmons-Rado theorem. $\square$

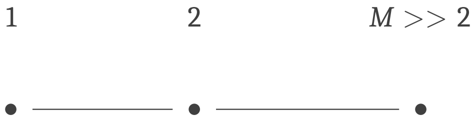
We have shown that using **Best-In-Greedy for maximization problems** on matroids will always find the best solution, likewise **Worst-Out-Greedy for minimization problems**.

Remarks

We have shown that using **Best-In-Greedy for maximization problems** on matroids will always find the best solution, likewise **Worst-Out-Greedy for minimization problems**.
If we apply the **Worst-Out-Greedy** to the maximization problem or the **Best-In-Greedy** to the minimization problems, **there is no positive lower/finite upper bound for** $\frac{G}{OPT}$.

Remarks

We have shown that using **Best-In-Greedy for maximization problems** on matroids will always find the best solution, likewise **Worst-Out-Greedy for minimization problems**.
If we apply the **Worst-Out-Greedy** to the maximization problem or the **Best-In-Greedy** to the minimization problems, **there is no positive lower/finite upper bound for** $\frac{G}{OPT}$.
To see this try finding a minimal vertex cover of maximum weight or a maximal stable set of minimum weight in this simple graph:



However in the case of matroids, it does not matter whether we use the Best-In-Greedy or the Worst-Out-Greedy: since all bases have the same cardinality, the minimization problem for $(E, \mathcal{F}, c)$ is equivalent to the maximization problem for $(E, \mathcal{F}, c')$, where

$$
\begin{aligned}
c'(e) &:= M - c(e) \qquad \forall e \in E \\
M &:= 1 + \max\{c(e) \,:\, e \in E\}
\end{aligned}
$$

Therefore

Corollary

Kruskal's algorithm solves the MST problem optimally.

Matroids and combinatorial optimization

Thank you for listening!