Matroids

Michel Habib
MPRI, M2 2016

21 septembre 2016

Schedule

Rappels sur les arbres

Matroids

Lemme d'échange sur les arbres

On considère un graphe connexe $G = (X, E)$ et deux arbres recouvrants sur G : $T = (X, F)$ et $T' = (X, F')$ (F, F' sont inclus dans E).

Echange entre arbres

$\forall e \in F\text{-}F'$, $\exists f \in F'\text{-}F$ such that $T'' = (X, F\text{-}e + f)$ est un arbre recouvrant de G.

Conséquences

Etant donné deux arbres recouvrants T, T' de G, il existe une suite d'échanges de 2 arêtes qui permet de passer de T à T'.

Mais aussi

Une structure d'espace vectoriel sur le corps $\{0, 1\}$ pour les cycles et les cocyles.

Based on an abstract idea of independence, matroids structures generalize both vector spaces and forest of trees in a graph.

A matroid by its independents

Withney [1935] : Let $M = (X, \mathcal{I})$ be a set family where X is a finite ground set and $\mathcal{I}$ a family of subsets of X.
M is a matroid if it satisfies the 3 following axioms :

1. $\emptyset \in \mathcal{I}$
2. If $I \in \mathcal{I}$ alors $\forall J \subseteq I$, $J \in \mathcal{I}$
3. If $I, J \in \mathcal{I}$ et $|I| = |J| + 1$ then $\exists x \in I$-J such that : $J \cup \{x\} \in \mathcal{I}$.

Variations around axiom 3

- 3' If $I, J \in \mathcal{I}$ and $|I| > |J|$ then $\exists x \in I$-J such that : $J \cup \{x\} \in \mathcal{I}$.
- 3'' $\forall A \subseteq X$ and $\forall I, J \subseteq A$, $I, J \in \mathcal{I}$ and maximal then $|I| = |J|$
- It is easy to be convinced that 3, 3' and 3'' are equivalent.

1) and 2) just define an hereditary family
The basis are the independent sets of maximum size and they all have the same cardinality (from axiom 3).

A matroid defined by its basis

1. $\mathcal{B} \neq \emptyset$
2. $\forall B, B' \in \mathcal{B}$, $\forall x \in B \setminus B'$ and $\exists y \in B' \setminus B$ such that : $B\text{-}x\text{+}y \in \mathcal{B}$.

For spanning trees

Let T, T' be two spanning trees of a connected graph G : $\forall e \in F\text{-}F'$, $\exists f \in F'\text{-}F$ such that $T'' = (X, F\text{-}e + f)$ is a spanning tree of G.

Graphic matroids

A partir d'un graphe non orienté G on peut définir un matroide comme suit :
$M = (E(G), \mathcal{I})$ où $\mathcal{I}$ est constitué des ensembles d'arêtes des forêts de G.

- $\mathcal{I}$ est trivialement une famille héréditaire.
- Concernant l'axiome 3. Considérons deux forêts $I, J \in \mathcal{I}$ telles que $|I| = |J| + 1$.
  Soient $X_1, \ldots, X_k$ les ensembles de sommets des composantes connexes de J, n.b. ces ccx contiennent les sommets isolés du graphe. S'il existe une arête e de I, entre deux de ces composantes connexes, alors $J \cup \{e\}$ n'a pas de cycle et donc $J \cup \{e\} \in \mathcal{I}$.
  Sinon toutes les arêtes de I sont internes aux composantes connexes, mais alors $|I| \leq |J|$, contradiction.

A matroid $M = (X, \mathcal{I})$ is called graphic, if there exists a graph G such that
$E(G) = X$ et $\mathcal{I}$ is exactly the forest of G (as edge sets).
Not all matroids are graphic.

Théorème principal Edmonds, Rado

1. Soit $M = (E, \mathcal{I})$ un matroide et $\omega : E \to \mathcal{R}^+$ une fonction de poids, alors :
   un élément $J \in \mathcal{I}$ lexicographiquement maximum de M est de poids maximum (pour la somme).
2. Sorte de réciproque
   On considère une famille de parties J sur E qui vérifie les axiomes 1 et 2. Si pour toute fonction de poids ω un élément J lexicographiquement maximum de M est de poids maximum (pour la somme), alors $M = (E, \mathcal{J})$ est un matroide.

Preuve du 1

Soit $A = (a_1, \dots a_k)$ t.q. $\omega(a_1) \geq \omega(a_2) \dots \omega(a_k)$, un indépendant lexicographiquement maximal. Vu que la valuation ω est positive, A est nécessairement une base de M. De même si l'on considère un indépendant de poids maximal $B = (b_1, \dots b_h)$ t.q. $\omega(b_1) \geq \omega(b_2) \dots \omega(b_h)$ c'est aussi une base. Donc posons h=k=n. Par définition de l'ordre lexicographique (qui est l'ordre de la première différence) et donc $A \gg B$ implique :
$\exists k \in [1, n]$ t.q. pour $i \leq k$, $\omega(a_i) = \omega(b_i)$ and $\omega(a_{k+1}) > \omega(b_{k+1})$. (On supposera ici $A \neq B$).

Considérons les deux ensembles suivants :
$A_{k+1} = (a_1, \ldots a_{k+1})$ t.q. $\omega(a_1) \geq \omega(a_2) \ldots \omega(a_{k+1})$ et
$B_k = (b_1, \ldots b_k)$ t.q. $\omega(b_1) \geq \omega(b_2) \ldots \omega(b_k)$
Comme un matroide est une famille de parties héréditaires, A_{k+1}, B_k sont aussi indépendants.
D'après l'axiome 3 des matroids, comme $|A_{k+1}| = |B_k| + 1$, il existe $a \in \{a_1, \ldots, a_{k+1}\}$ tel que $B'_{k+1} = B_k + a$ soit un indépendant. En outre $\omega(B'_{k+1}) = \omega(B_k) + \omega(a) > \omega(B_k) + \omega(b_{k+1}) = \omega(B_{k+1})$.
En considérant B'_{k+1} et B on peut compléter en ajoutant que des b_j avec $j > k + 1$ (en utilisant $n - k - 1$ fois l'axiome 3) B'_k en une base B' qui vérifie :
$\omega(B') > \omega(B)$ ce qui contredit la maximalité de B.

Preuve du 2, l'espèce de réciproque

On considère $A \subseteq X$ et $I, J \in \mathcal{I}$ maximaux pour l'inclusion.
Supposons $|I| < |J|$.
Définissons une valuation $\omega : X \rightarrow R^+$ de la manière suivante :
$\omega(x) = 1 + \alpha$ si $x \in I$
$\omega(x) = 1$ si $x \in J$-I
$\omega(x) = 0$ si $x \in X$-I-J
Cela donne $\omega(I) = |I|.(1 + \alpha)$ et $\omega(J) = |I \cap J|.(1 + \alpha) + |J$-$I|$

- I est lexicographiquement plus grand que J.
- $|I| < |J|$ implique que $|I\text{-}J| < |J\text{-}I|$
- $\omega(I) < \omega(J)$ est équivalent à $(1+\alpha) < \frac{|J-I|}{|I-J|}$. Vu la remarque ci-dessus, il existe une telle valeur de $\alpha > 0$, d'où la contradiction.
  Donc si l'axiome 3" n'est pas vérifié, nous avons exhibé une fonction de poids pour laquelle lexicographiquement maximal n'implique pas maximum pour la somme.

The greedy algorithm on a matroid

Standard greedy applied on a matroid
$M = (E, \mathcal{F})$
$X \leftarrow \emptyset$;
while $|A| \neq \emptyset$ **do**
    $A \leftarrow \{x \in E$ such that $X \cup \{x\} \in \mathcal{F}\}$;
    Choose $a \in A$ with maximum weight ;
    $X \leftarrow X \cup \{a\}$;
end

This algorithm computes one basis lexicographically maximum and therefore optimum using Edmonds et Rado's Theorem.
After an element is chosen and added to X usign a local optimal criterium, there is no backtrack. (This is just the greedy algorithmic paradigm).
The complexity highly depends on the computation of the set A and the computation of the maximum weighted element.

Algorithmique sur les matroids

Lié à la représentation du matroide $M = (E, \mathcal{I})$:

- Par un oracle qui répond sur une partie $I \subseteq E$ est un indépendant en $O(1)$
- Par un oracle qui répond sur une partie $I \subseteq E$ est un indépendant en $O(|I|)$
- Mais vérifier que $X \cup \{x\} \in \mathcal{F}$ peut demander beaucoup de ressources pour certains matroids. (i.e. demander de résoudre un problème NP-complet par exemple).

Matroid defined by its circuits

The circuits of the matroid are the minimal dependent subsets, they can be defined with 2 axioms :

1. $\forall C \in \mathcal{C}, \forall C' \subset C,\ C' \notin \mathcal{C}$
2. $\forall C, C' \in \mathcal{C}, \forall a \in C \cap C', \exists C" \in C$ such that : $C" \subseteq C \cup C'$-a

So we can write a matroid version of Kruskal algorithm.

Matroidal rank functions

Let $M = (E, \mathcal{I})$ be a matroid, we define the function r :
$r_M : \mathcal{P}\ (E) \to \mathcal{N}$ such that :
$\forall A \subseteq E$, $r_M(A) = max\{|I|$ avec $I \in \mathcal{I}$ and $I \subseteq A\}$
This function is well defined, since :
$M(A) = (A, \mathcal{J})$ with $\mathcal{J} = \{\mathcal{I} \cap \mathcal{A} | \mathcal{I} \in \mathcal{I}\}$ is also a matroid and $r_M(A)$ is equal to the size of a basis of this matroid.

Or equivalently :

$r : \mathcal{P}(\mathcal{X}) \rightarrow \mathcal{N}$ that satisfies :

1. $r(\emptyset) = 0$
2. $\forall A \subseteq X$, $\forall A \subseteq X$, $r(A + x)$-$r(A) \in \{0, 1\}$
3. $\forall A \subseteq X$, $\forall x, y \in X$-A such that : $r(A + x) = r(A + y) = r(A)$ then $r(A + x + y) = r(A)$.

Or else

$r : \mathcal{P}(\mathcal{X}) \to \mathcal{N}$ that satisfies :

1. $\forall A \subseteq X,\ r(A) \leq |A|$
2. $\forall A, B \subseteq X,\ A \subseteq B$ then $r(A) \leq r(B)$
3. submodularité
   $\forall A, B \subseteq X,\ r(A \cap B) + r(A \cup B) \leq r(A) + r(B)$

- independents satisfy $r(A) = |A|$.
  The rank function can be used as an oracle when writing algorithms.
- submodularity $=$ a kind of convexity on subsets
- But matroids can also be defined using a closure operator or via lattices

A duality for matroids

To the matroid $M = (E, \mathcal{I})$ one can associate its dual matroid
$M^* = (E, \mathcal{C}^*)$
$\forall A \subseteq E$, $r^*(A) = |A| + r(E\text{-}A) - r(E)$
B is a basis for M iff E-B is a basis for M^*.

a cocircuit is simply a circuit in the dual but it can also be defined as a cut as a minimal under inclusion subset of E intersecting all independent element in M.
Then we can generalize the framework with blue rule and red rule as used for minimum spanning trees.

Les différentes définitions d'un matroide : par ses indépendants, par ses bases (indépendants maximaux), sa fonction de rang, ses cycles (dépendants minimaux) définissent le même concept mathématique, mais hélas elles ne sont pas du tout équivalentes du point de vue de la complexité algorithmique.
Notion d'oracle pour l'indépendance

Exemple de matroide sur E :
$M_k =$ l'ensemble des parties de E ayant $\leq k$ éléments.
appelé le matroide uniforme.
Les axiomes sont trivialement vérifiés.

Le matroide de partition

Soit $P = \{E_1, \ldots, E_k\}$ une partition de E et k entiers positifs $d_1, \ldots, d_k$ et $\mathcal{I}$=$\{$ $I \subseteq E$ tel que $\forall i \in [1, k]$, $|I \cap E_i| \leq d_i\}$ alors $M = (E, \mathcal{I})$ est un matroide dit de partition.
Les axiomes 1 et 2 sont trivialement vrais.
Considérons 3. Considérons I, J tels que $|I| = |J| + 1$.
Si $\forall i$, $|E_i \cap I| = |E_i \cap J|$ alors $|I| = |J|$.
Donc il existe au moins un entier i tel que $d_i \geq |E_i \cap I| > |E_i \cap J|$ et donc on peut augmenter J d'au moins un élément.

Paradigmes algorithmiques et matroids

- Algorithmes gloutons :
  les prototypes des algorithmes gloutons sont les algorithmes pour l'arbre de poids maximum.
  Indépendant de poids maximum dans un matroide.
- Algorithmes par échanges améliorants (principe des chaînes alternées dans les couplages ou flots)
  Indépendant commun à deux matroids de taille maximum
- Programmation linéaire et algorithmes de type simplexe qui examine tous les sommets d'un polyèdre.
  matroids orientés

Exercices :

1. Comment se calcule la fonction de rang d'un matroide graphique (on suppose disposer du graphe) ?
2. Pour le problème de l'arbre de Steiner avec $|S| = 2$, montrer que la solution est la plus courte chaîne reliant les deux sommets.

$G = (X, U)$ un graphe orienté avec $|X| = n$, une fonction de poids $\omega : U \to R_+$, considérons les 3 matroids suivants définis sur G :

$M_1 = (U, \mathcal{I})$

Le matroide graphique des forêts de G

$M_2 = (U, \mathcal{J})$

avec $\mathcal{J}$=$\{J \subseteq U$ tel que $|J| \leq n-1$ et $\forall x \in X$, J contient au plus un arc sortant de $x\}$

$M_3 = (U, \mathcal{K})$

avec $\mathcal{K} = \{K \subseteq U$ tel que $|K| \leq n-1$ et $\forall x \in X$, K contient au plus un arc entrant en $x\}$

Tout d'abord

M_1, M_2 et M_3 sont bien des matroids.

Intersection de 2

$A \subseteq U$ est un arbre orienté recouvrant de poids maximal
ssi
A est une base de poids maximal commune à M_1 et M_2

Intersection de 3

$A \subseteq U$ est un chemin hamiltonien de poids maximal
ssi
A est une base de poids maximal commune à M_1, M_2 et M_3.