Optimisation Combinatoire Avancée (MPRO)
Notes de cours

Frédéric Meunier

CHAPTER 1

Matroïdes

Les matroïdes ont été introduits par Whitney en 1935. On s'est très vite rendu compte qu'ils unifiaient des propriétés de domaines variées comme les graphes, l'algèbre linéaire ou la géométrie. Plus tard, dans les années 60, on a réalisé qu'ils jouaient également un rôle central en optimisation combinatoire.

1 Définitions

Soit V un ensemble fini. Soit $\mathcal{I}$ une collection non-vide de parties de V. La paire $M = (V, \mathcal{I})$ est un *matroïde* si les deux conditions suivantes sont satisfaites :

(i) Pour tout $S, T \subseteq V$: si $S \in \mathcal{I}$ et $T \subseteq S$, alors $T \in \mathcal{I}$.

(ii) Pour tous $S, T \in \mathcal{I}$ tels que $|S| < |T|$, il existe $x \in T \setminus S$ tel que $S \cup \{x\} \in \mathcal{I}$.

Les parties de V dans $\mathcal{I}$ sont les indépendants du matroïde M. Noter que la propriété (i) implique en particulier que $\varnothing$ est toujours indépendant. Une partie S de V qui n'est pas dans $\mathcal{I}$ est un *dépendant*.

Un indépendant maximal pour l'inclusion est une *base*. Une conséquence de la propriété (ii) est que toutes les bases ont le même cardinal (exercice). Le cardinal commune des bases est le *rang du matroïde* M. Pour $S \subseteq V$, on définit le *rang de la partie* S (respectivement à M), noté $r_M(S)$, par

$$r_M(S) = \max\{|I| \colon I \in \mathcal{I}, I \subseteq S\}.$$

(L'indice M sera omis en général, s'il n'y a pas d'ambiguïté.) Noter que le rang du matroïde M n'est rien d'autre que $r_M(V)$.

Un dépendant minimal pour l'inclusion est un *circuit* :

$$C \text{ est un circuit} \qquad \Longleftrightarrow \qquad C \notin \mathcal{I} \text{ et on a } C \setminus \{x\} \in \mathcal{I} \text{ pour tout } x \in C.$$

Le *span* d'une partie S de V, noté $\mathrm{span}(S)$, est défini par

$$\mathrm{span}(S) = \{x \in V \colon r(S \cup \{x\}) = r(S)\}.$$

Noter que si $S \in \mathcal{I}$, on a aussi $\mathrm{span}(S) = \{x \in V \colon S \cup \{x\} \notin \mathcal{I}\} \cup S$ (exercice). Un élément $x \in V$ tel que $\{x\}$ est un circuit est appelé *boucle*.

Enfin, on utilisera parfois la notation $S + x$ à la place de $S \cup \{x\}$ et $S - x$ à la place de $S \setminus \{x\}$.

2 Quelques exemples

2.1 Exemples de matroïdes Dans cette section, des matroïdes classiques sont décrits. Pour chacun d'eux, on pourra vérifier en exercice qu'ils satisfont bien les axiomes des matroïdes. De même, on pourra déterminer les bases, les circuits, la fonction de rang, le rang du matroïde, les spans.

2.1.1 Matroïde uniforme. Pour k un entier positif et V un ensemble fini, posons $\mathcal{I} = \{I \subseteq V : |I| \leqslant k\}$. La paire $M = (V, \mathcal{I})$ est un matroïde, appelé *matroïde uniforme*.

2.1.2 Matroïde de partition. Soient $k_1, \ldots, k_m$ des entiers positifs et $U_1, \ldots, U_m$ une partition d'un ensemble fini V. Posons $\mathcal{I} = \{I \subseteq V : |I \cap U_i| \leqslant k_i \text{ pour tout } i \in [m]\}$. La paire $M = (V, \mathcal{I})$ est un matroïde, appelé *matroïde de partition*.

2.1.3 Matroïde linéaire. Soit A une matrice $m \times n$ sur un corps $\mathbb{F}$. Posons

$$\mathcal{I} = \{I \subseteq [n] : \text{les colonnes de } A \text{ indicées par les éléments dans } I \text{ forment un système libre}\}.$$

La paire $M = ([n], \mathcal{I})$ est un matroïde, appelé *matroïde linéaire*. On dit alors que A *représente* M.

Un matroïde est *représentable* sur $\mathbb{F}$ s'il est isomorphe à un matroïde linéaire représenté par une matrice à coefficients dans $\mathbb{F}$.

2.1.4 Matroïde algébrique. Soit $\mathbb{F}$ un corps et $\mathbb{E}$ une extension de $\mathbb{F}$ (penser $\mathbb{F}$ comme un sous-corps de $\mathbb{E}$). Des éléments $a_1, \ldots, a_t \in \mathbb{E}$ sont *algébriquement indépendants* sur $\mathbb{F}$ s'il n'existe pas de polynôme non-nul $P \in \mathbb{F}[X_1, \ldots, X_t]$ tel que $P(a_1, \ldots, a_t) = 0$.

Soit $V \subseteq \mathbb{E}$. Posons $\mathcal{I} = \{I \subseteq V : \text{les éléments de } I \text{ sont algébriquement indépendants}\}$. La paire $M = (V, \mathcal{I})$ est un matroïde, appelé *matroïde algébrique*.

2.1.5 Matroïde graphique. Soit $G = (V, E)$ un graphe (non-orienté). Posons

$$\mathcal{I} = \{F \subseteq E : (V, F) \text{ est une forêt}\}.$$

La paire $M = (E, \mathcal{I})$ est un matroïde, appelé *matroïde graphique*.

2.1.6 Matroïde transversal. Soit $G = (V, E)$ un graphe biparti, dont la bipartition des sommets est donnée par $X \sqcup Y = V$. Posons

$$\mathcal{I} = \{I \subseteq X : \text{il existe un couplage de } G \text{ couvrant } I\}.$$

La paire $M = (X, \mathcal{I})$ est un matroïde, appelé *matroïde transversal*.

2.2 Exemples de non-matroïdes

2.2.1 Un peu comme la matroïde de partition. Soient $k_1, \ldots, k_m$ des entiers positifs et $U_1, \ldots, U_m$ des parties d'un ensemble fini V. Posons $\mathcal{S} = \{I \subseteq V : |I \cap U_i| \leqslant k_i \text{ pour tout } i \in [m]\}$. La paire $(V, \mathcal{S})$ n'est en général pas un matroïde (exercice : donner un exemple concret qui ne soit pas un matroïde). La différence avec la matroïde de partition est que pour ce dernier, les U_i sont 2 à 2 disjoints. La condition (i) est bien satisfaite, mais pas la condition (ii).

2.2.2 Indépendants de graphe. Soit $G = (V, E)$ un graphe. Un sous-ensemble S de V est un *indépendant* s'il n'induit aucune arête (i.e., aucune arête de G a ses deux extrémités dans S). En général, les indépendants d'un graphe ne forme pas matroïde. La condition (i) est bien satisfaite, mais pas la condition (ii).

3 Premières propriétés

Soit $M = (V, \mathcal{I})$ un matroïde.

3.1 Rang et span. Une conséquence immédiate de la définition du rang est sa monotonie : si $S \subseteq T$, on a $r(S) \leqslant r(T)$. La fonction rang a également une autre propriété.

THÉORÈME 1. *La fonction de rang r est sous-modulaire, i.e.*

$$r(S) + r(T) \geqslant r(S \cup T) + r(S \cap T) \quad \forall S, T \subseteq V.$$

DÉMONSTRATION. Soit I un indépendant inclus dans $S \cap T$ tel que $r(S \cap T) = |I|$. Il existe alors un indépendant J inclus dans $S \cup T$ tel que $r(S \cup T) = |J|$ et $I \subseteq J$: il suffit en effet de prendre un indépendant $J' \subseteq S \cup T$ tel que $r(S \cup T) = |J'|$, et d'ajouter le plus d'éléments de J' à I qui maintiennent l'indépendance. On a alors

$$r(S) + r(T) \geqslant |S \cap J| + |T \cap J| = |(S \cup T) \cap J| + |(S \cap T) \cap J| \geqslant |J| + |I| = r(S \cup T) + r(S \cap T).$$

□

PROPOSITION 2. *Si* $r(S \cup X) = r(S)$ *et* $r(S \cup Y) = r(S)$*, alors* $r(S \cup X \cup Y) = r(S)$*.*

DÉMONSTRATION. On a

$$2r(S) = r(S \cup X) + r(S \cup Y) \geqslant r(S \cup X \cup Y) + r((S \cup X) \cap (S \cup Y)) \geqslant r(S \cup X \cup Y) + r(S).$$

La première inégalité est une conséquence de la sous-modularité de la fonction rang. Donc, on a $r(S) \geqslant r(S \cup X \cup Y)$. L'inégalité opposée est immédiate. □

PROPOSITION 3. *Si* $S \subseteq \mathrm{span}(T)$*, alors* $\mathrm{span}(S) \subseteq \mathrm{span}(T)$*.*

DÉMONSTRATION. Soit $x \in \mathrm{span}(S)$. On veut montrer que $x \in \mathrm{span}(T)$. Si $x \in T$, cette dernière relation est immédiate. Supposons donc que $x \notin T$. On a

$$r(S + x) + r(S \cup T) \geqslant r(S \cup T \cup \{x\}) + r(S) \geqslant r(T + x) + r(S).$$

La première inégalité est une conséquence de la sous-modularité de la fonction rang. Comme $r(S + x) = r(S)$ et $r(S \cup T) = r(T)$ (cette seconde égalité s'obtient par répétition de la proposition 2), on a $r(T) \geqslant r(T + x)$. On a donc $r(T + x) = r(T)$, i.e. $x \in \mathrm{span}(T)$. □

3.2 Quelques opérations. Soit $X \subseteq V$. La *restriction* de M à X, notée $M|_X$, est le matroïde $(X, \mathcal{I}')$ où $\mathcal{I}' = \{I \in \mathcal{I} \colon I \subseteq X\}$. (Exercice : Vérifier que c'est bien un matroïde.) Dans le cas où X est de la forme $V \setminus \{x\}$ avec $x \in V$, le matroïde $M|_X$ est aussi noté $M \setminus x$ (on dit que le matroïde est obtenu par *suppression* de x).

Soit k un entier positif. La *troncature* de M, notée $M^{\leqslant k}$, est le matroïde $(V, \mathcal{I}')$ où $\mathcal{I}' = \{I \in \mathcal{I} \colon |I| \leqslant k\}$. (Exercice : Vérifier que c'est bien un matroïde.)

Il existe aussi des notions de *contraction* et de *dualité*, mais elles ne sont pas abordées dans ce cours.

4 Optimiser sur les matroïdes

Étant donnés un matroïde M et une fonction de poids $w \colon V \to \mathbb{R}$, un problème naturel, possédant de nombreuses applications, est celui de la base de poids maximal : en notant $\mathcal{B}$ l'ensemble des bases de M, résoudre $\max_{B \in \mathcal{B}} w(B)$.

L'algorithme 1, appelé *algorithme glouton*, est un algorithme naturel et “naïf” : je regarde les éléments de mon matroïde l'un après l'autre, dans l'ordre décroissant de leurs poids ; je garde un élément dès qu'il n'est pas dans le span des éléments déjà acceptés. Le petit miracle, c'est que cet algorithme résout bien le problème de la base de poids maximal. Noter qu'en général, l'algorithme glouton est sous-optimal (couplage de poids maximal, sac-à-dos, coloration propre, etc.).

Données : Matroïde $M = (V, \mathcal{I})$, poids $w\colon V \to \mathbb{R}$
Sortie **:** Base B^* résolvant $\max_{B \in \mathcal{B}} w(B)$

```
Trier les éléments de V : les écrire v_1, ..., v_n avec w(v_1) ⩾ w(v_2) ⩾ ··· ⩾ w(v_n);
B ← ∅;
for i = 1, ..., n do
    if B + v_i ∈ I then
        B ← B + v_i;
    end
end
return B
```

Algorithme 1 : Algorithme glouton pour la base de poids maximal

Théorème 4. *L'algorithme glouton permet de calculer une base de poids maximal.*

Démonstration. □

Dans le cas particulier du matroïde graphique, l'algorithme glouton est l'algorithme de Kruskal calculant l'arbre couvrant de poids maximal.

En se restreignant aux éléments de poids positifs ou en considérant la troncature, on obtient les deux corollaires suivants.

Corollaire 5. *L'algorithme glouton permet de calculer un indépendant de poids maximal.*

Corollaire 6. *L'algorithme glouton permet de calculer un indépendant de cardinal k de poids maximal.*

Pour les algorithmes, lorsqu'on a un matroïde $M = (V, \mathcal{I})$ en données, on suppose souvent que l'on dispose de l'*oracle d'indépendance* : pour $S \subseteq V$, on sait décider en temps constant si S est un indépendant de M. Dans ce modèle de complexité, les problèmes de la base de poids maximal, de l'indépendant de poids maximal et de l'indépendant de cardinal fixé et de poids maximal sont polynomiaux.

On a une "réciproque" du théorème 4.

Théorème 7. *Soit V un ensemble fini et soit $\mathcal{I} \subseteq 2^V$ vérifiant (i) de la section 1. Si, pour toute fonction de poids $w\colon V \to \mathbb{R}$, l'algorithme glouton calcule un élément de $\mathcal{I}$ de poids maximal, alors $(V, \mathcal{I})$ est un matroïde.*

Démonstration. □

5 Intersection de matroïdes

5.1 Théorème et applications. Considérons deux matroïdes ayant le même ensemble d'éléments $M_1 = (V, \mathcal{I}_1)$ et $M_2 = (V, \mathcal{I}_2)$. La paire $(V, \mathcal{I}_1 \cap \mathcal{I}_2)$ n'est en général pas un matroïde. Cependant, on a le théorème remarquable suivant, trouvé par Edmonds en 1970.

Théorème 8. *On a l'égalité suivante*

$$\max_{I \in \mathcal{I}_1 \cap \mathcal{I}_2} |I| = \min_{S \subseteq V} \big(r_{M_1}(S) + r_{M_2}(V \setminus S) \big),$$

et dans le modèle de l'oracle d'indépendance, les ensembles optimaux pour les membres gauche et droit se calculent en temps polynomiaux.

L'inégalité $\leqslant$ entre le membre de gauche et le membre de droite est une inégalité de dualité faible, facile : pour $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ et $S \subseteq V$, on a évidemment

$$|I| = |I \cap S| + |I \setminus S| \leqslant r_{M_1}(S) + r_{M_2}(V \setminus S).$$

La partie difficile du théorème est l'inégalité inverse.

5.1.1 Théorème de Kőnig Soit $G = (V, E)$ un graphe biparti, dont la bipartition des sommets est donnée par $X \sqcup Y = V$. Posons

$$\mathcal{I}^X = \{F \subseteq E \colon |F \cap \delta(v)| \leqslant 1 \text{ pour tout } v \in X\}.$$
$$\mathcal{I}^Y = \{F \subseteq E \colon |F \cap \delta(v)| \leqslant 1 \text{ pour tout } v \in Y\}.$$

Les paires $M^X = (E, \mathcal{I}^X)$ et $M^Y = (E, \mathcal{I}^Y)$ sont des matroïdes de partition. L'ensemble $\mathcal{I}^X \cap \mathcal{I}^Y$ est l'ensemble des couplages de G et le théorème 8 est dans ce cas le théorème de Kőnig (exercice).

5.1.2 Arbre arc-en-ciel Soit $G = (V, E)$ un graphe, et supposons qu'on ait une partition $E = E_1 \sqcup E_2 \sqcup \cdots \sqcup E_k$ des arêtes (on peut voir cette partition comme une coloration des arêtes). Décider s'il existe un arbre couvrant arc-en-ciel (avec les arêtes de couleurs deux à deux distinctes) peut donc se faire en temps polynomial : on applique le théorème 8 à l'intersection d'un matroïde graphique et d'un matroïde de partition.

5.2 Preuve du théorème.

Lemme 9. *Soit $G = (V, E)$ un graphe biparti, dont la bipartition des sommets est donnée par $X \sqcup Y = V$, avec $|X| = |Y| = k$. Supposons que G possède un unique couplage parfait N. Alors on peut numéroter les sommets dans $X = \{x_1, \ldots, x_k\}$ et ceux dans $Y = \{y_1, \ldots, y_k\}$ de sorte que $N = \{x_i y_i \colon i \in [k]\}$ et $x_i y_j \notin E$ pour tous $1 \leqslant i < j \leqslant k$.*

Démonstration. $\square$

Soit $M = (V, \mathcal{I})$ un matroïde. Soit $I \in \mathcal{I}$. On considère le graphe $G_M(I)$ biparti défini comme suit : la bipartition est donnée par I et $V \setminus I$; les arêtes sont les éléments de l'ensemble $\{xy \colon x \in I, y \in V \setminus I, I - x + y \in \mathcal{I}\}$.

Lemme 10. *Soit $J \subseteq V$ tel que $|J| = |I|$. Si $G_M(I)$ contient un unique couplage couvrant $I \triangle J$, alors $J \in \mathcal{I}$.*

Démonstration. $\square$

Lemme 11. *Soit $J \subseteq V$ tel que $|J| = |I|$ et $J \subseteq \operatorname{span}(I)$. Supposons que $G_M(I)$ contienne un unique couplage couvrant $I \triangle J$. Si $s \in V \setminus (I \cup J)$ est tel que $I + s \in \mathcal{I}$, alors $J + s \in \mathcal{I}$.*

Démonstration. Appliquons le lemme 10. Par conséquent J est indépendant. Si $s \in \operatorname{span}(J)$, alors $s \in \operatorname{span}(I)$ car $J \subseteq \operatorname{span}(I)$. $\square$

Démonstration du théorème 8. Soit $I \in \mathcal{I}_1 \cap \mathcal{I}_2$. On va décrire une procédure qui teste si I est un indépendant commun à $\mathcal{I}_1$ et $\mathcal{I}_2$ de cardinal maximale et si ce n'est pas le cas, explique comment en trouver un plus grand d'une unité.

On considère le graphe orienté $D_{M_1,M_2}(I)$ défini comme suit. C'est un graphe biparti dont la bipartition est donnée par $I \sqcup (V \setminus I)$ et donc les arcs sont les éléments de

$$\big\{(x, y) \in (I, V \setminus I) \colon I - x + y \in \mathcal{I}_1\big\} \cup \big\{(y, x) \in (V \setminus I, I) \colon I - x + y \in \mathcal{I}_2\big\}.$$

On pose $Y_1 = \{y \in V \setminus I \colon I + y \in \mathcal{I}_1\}$ et $Y_2 = \{y \in V \setminus I \colon I + y \in \mathcal{I}_2\}$.

Deux cas sont possibles : soit il existe dans $D_{M_1,M_2}(I)$ un chemin de Y_1 à Y_2, soit il n'en existe pas.

Considérons le premier cas. Notons P un chemin de Y_1 à Y_2 dans $D_{M_1,M_2}(I)$ avec un nombre minimal d'arcs. Posons $I' = I \triangle V(P)$ (où $V(P)$ est l'ensemble des sommets de P). Nous allons montrer que I' appartient simultanément à $\mathcal{I}_1$ et $\mathcal{I}_2$. Soit $z \in Y_1$ l'origine de P. Par optimalité de P, les arcs de P quittant I forment l'unique couplage de $G_{M_1}(I)$ dont l'ensemble de sommets est $V(P) - z$. L'optimalité de P implique aussi que tous les sommets y dans $V(P) - z$ et pas dans I sont tels que $I + y \notin \mathcal{I}_1$. Le lemme 11 montre alors que I' est un indépendant de $\mathcal{I}_1$ (en prenant $J = V(P) \triangle (I + z)$). Un raisonnement semblable montre que I' appartient également à $\mathcal{I}_2$. Comme le cardinal de I' est d'une unité plus grande que celle de I, l'indépendant commun I n'est pas maximum.

Considérons le second cas. Nous allons montrer qu'alors nécessairement I est un indépendant commun à $\mathcal{I}_1$ et $\mathcal{I}_2$ de cardinal maximal. Comme il n'y a pas de chemin allant de Y_1 à Y_2, il existe un ensemble S tel que $Y_2 \subseteq S$ et tel que $\delta^-(S) = \varnothing$ (il suffit de poser S l'ensemble des sommets ne pouvant pas être atteint depuis Y_1). Nous prouvons maintenant que $r_{M_1}(S) \leqslant |I \cap S|$. Supposons que cette inégalité ne soit pas satisfaite. Il existe alors $y \in S \setminus I$ tel que $(I \cap S) + y \in \mathcal{I}_1$. Comme $y \notin Y_1$, on a $I + y \notin \mathcal{I}_1$ et il y a un $x \in I \setminus S$ tel que $I - x + y \in \mathcal{I}_1$ (ajouter des éléments de I à $(I \cap S) + x$ tant qu'on maintient l'indépendance). Mais cela implique l'existence d'un arc (x, y), ce qui contredit la définition de S. On a donc bien $r_{M_1}(S) \leqslant |I \cap S|$. De manière semblable, on montre $r_{M_2}(V \setminus S) \leqslant |I \setminus S|$, ce qui implique que l'on a

$$r_{M_1}(S) + r_{M_2}(V \setminus S) \leqslant |I|.$$

Il a déjà noté que l'égalité opposée était toujours vérifiée. On a donc prouvé l'égalité du théorème, dans laquelle les ensembles I et S ainsi construits réalisent respectivement le minimum et le maximum.

Dans le modèle d'oracle d'indépendance, la procédure décrite ci-dessus peut être réalisée en temps polynomial, ce qui montre la seconde partie du théorème. $\square$

CHAPTER 2

Fonctions sous-modulaires

Les fonctions sous-modulaires jouent un rôle central en optimisation combinatoire, en économie et en machine learning.

1 Définitions

Soit V un ensemble fini. Une fonction $f\colon V \to \mathbb{R}$ est *sous-modulaire* si elle vérifie pour tous $S, T \subseteq V$

$$f(S) + f(T) \geqslant f(S \cup T) + f(S \cap T)\,.$$

On a déjà vu un exemple de fonction sous-modulaire, à savoir la fonction rang d'un matroïde.

2 Quelques exemples

2.1 Biens substituables et complémentaires Le lemme suivant fournit une définition équivalente de la sous-modularité. Dans les contextes économiques, cette dernière est particulièrement utile car elle traduit la propriété de l'*utilité marginale décroissante* pour les biens "substituables". (À l'opposé, pour les biens "complémentaires", on a une utilité marginale croissante qui se traduit par une propriété de *supermodularité*.)

Lemme 12. *Une fonction $f\colon 2^V \to \mathbb{R}$ est sous-modulaire si et seulement si pour tous $S \subseteq T \subseteq V$ et $v \in V \setminus T$, on a*

$$f(T \cup \{v\}) - f(T) \leqslant f(S \cup \{v\}) - f(S)\,.$$

Démonstration. Supposons f est sous-modulaire. Alors l'inégalité

$$f(T' \cup \{v\}) - f(T') \leqslant f(S' \cup \{v\}) - f(S') \tag{1}$$

est trivialement satisfaite pour tous $S' \subseteq T' \subseteq V$ et $v \in V \setminus T'$.

Réciproquement, supposons l'inégalité (1) satisfaite pour tous $S' \subseteq T' \subseteq V$ et $v \in V \setminus T'$. Prenons $S, T \subseteq V$, et notons arbitrairement $s_1, s_2, \ldots, s_k$ les éléments de $S \setminus T$. Pour $i \in [k-1]$, (1) pour $S' = (S \cap T) \cup \{s_1, \ldots, s_i\}$, $T' = T \cup \{s_1, \ldots, s_i\}$ et $v = s_{i+1}$ devient

$$f\big((S \cap T) \cup \{s_1, \ldots, s_i\}\big) - f\big((S \cap T) \cup \{s_1, \ldots, s_{i-1}\}\big) \geqslant f(T \cup \{s_1, \ldots, s_i\}) - f(T \cup \{s_1, \ldots, s_{i-1}\})\,.$$

En sommant cette dernière inégalité pour $i = 1, \ldots, k$, on obtient $f(S) - f(S \cap T) \geqslant f(T \cup S) - f(T)$, soit exactement ce qu'on souhaitait obtenir. □

2.2 Rang de matroïde

2.3 Union d'ensembles

2.4 Coupes de graphe orienté

2.5 Probabilités d'événements conjoints On considère un espace de probabilités. Soient $A_1, \dots, A_n$ des événements aléatoires. La fonction $f\colon S \mapsto \mathbb{P}(\bigcap_{i \in S} A_i)$ est une fonction super-modulaire $2^{[n]} \to \mathbb{R}_+$ (son opposée est donc sous-modulaire).

2.6 Entropie et information mutuelle Étant donnée une variable aléatoire X à valeur dans E, la quantité $H(X) \coloneqq -\mathbb{E}[\log_2(p(X))]$, où $p(x)$ est la masse ou la densité de $x \in E$ selon que la variable est discrète ou continue, est l'*entropie* de X.

Considérons n variables aléatoires $X_1, X_2, \dots, X_n$. On considère la fonction $H\colon S \in 2^{[n]} \mapsto H((X_i)_{i \in S}) \in \mathbb{R}$, appelée *entropie jointe* (noter le petit abus de notation).

Théorème 13. *L'entropie jointe H est sous-modulaire.*

Démonstration. On écrit la preuve dans le cas où les variables aléatoires sont discrètes et à support fini. Le support de la variable aléatoire X_i sera noté E_i. De plus, on notera x_B, pour $B \subseteq [n]$, la collection $(x_j)_{j \in B}$. Pour $S \subseteq T \subseteq V \setminus \{i\}$, on considère la quantité

$$\gamma \coloneqq - \sum_{x_{T \cup \{i\}} \in \prod_{j \in T \cup \{i\}} E_j} p(x_{T \cup \{i\}}) \log \frac{p(x_{T \cup \{i\}}) p(x_S)}{p(x_{S \cup \{i\}}) p(x_T)}$$

où $p(x_B)$, pour $B \subseteq [n]$, est la probabilité que $X_j = x_j$ pour tout $j \in B$. Par concavité du logarithme, on a

$$\gamma \leqslant \log \left(\sum_{x_{T \cup \{i\}} \in \prod_{j \in T \cup \{i\}} E_j} \frac{p(x_{S \cup \{i\}}) p(x_T)}{p(x_S)} \right) .$$

On a donc

$$\gamma \leqslant \log \left(\sum_{x_T \in \prod_{j \in T} E_j} \sum_{x_i \in E_i} \frac{p(x_{S \cup \{i\}}) p(x_T)}{p(x_S)} \right) = \log \left(\sum_{x_T \in \prod_{j \in T} E_j} p(x_T) \right) = 0 .$$

Par conséquent,

$$\begin{aligned} &- \sum_{x_{T \cup \{i\}} \in \prod_{j \in T \cup \{i\}} E_j} p(x_{T \cup \{i\}}) \log p(x_{T \cup \{i\}}) + \sum_{x_T \in \prod_{j \in T} E_j} p(x_T) \log p(x_T) - \\ &\leqslant - \sum_{x_{S \cup \{i\}} \in \prod_{j \in S \cup \{i\}} E_j} p(x_{S \cup \{i\}}) \log p(x_{S \cup \{i\}}) + \sum_{x_S \in \prod_{j \in S} E_j} p(x_S) \log p(x_S) . \end{aligned}$$

□

Étant données deux parties S, T de $[n]$, l'*information mutuelle* de $(X_i)_{i \in S}$ et $(X_j)_{j \in T}$ est définie comme la quantité $I(S; T) \coloneqq H(S) + H(T) - H(S, T)$. L'information mutuelle donne lieu à deux fonctions sous-modulaires intéressantes en pratique :

— La fonction $S \in 2^{[p]} \mapsto I(S; [p] \setminus S) \in \mathbb{R}$ est sous-modulaire (et symétrique).

— Soient A, B deux parties de $[p]$ telle que les variables X_i pour $i \in A$ soient indépendantes conditionnellement aux variables X_j pour $j \in B$. Alors la fonction $S \in 2^{[p]} \mapsto I(B; S) \in \mathbb{R}$ est sous-modulaire (et monotone).

2.7 Placement de capteurs Considérons une pièce dont on veut surveiller la température à l'aide de capteurs. Il y a n emplacements potentiels, et l'on veut placer k capteurs. La température X_i à l'emplacement i est une variable aléatoire (on suppose que les capteurs sont de qualité parfaite). Le problème d'identification du "meilleur" ensemble de k emplacements peut se modéliser comme

$$\begin{array}{rl} \text{Max} & H(S) \\ \text{s.c.} & |S| = k \\ & S \subseteq [n] \end{array}$$

ou comme

$$\begin{array}{rl} \text{Max} & I(S; [n] \setminus S) \\ \text{s.c.} & |S| = k \\ & S \subseteq [n]\,. \end{array}$$

Le premier problème cherche à maximiser l'information totale récupérée ; le second cherche à les emplacements fournissant l'information la plus complète sur les emplacements non-couverts.

Tous deux sont des problèmes de maximisation de fonction sous-modulaire sous contraintes de cardinal.

2.8 Clustering Partitionner un ensemble de variables aléatoires $\{X_1, X_2, \ldots, X_n\}$ en deux parties, de manière à minimiser la "longueur de description minimale", revient à chercher S minimisant $I(S; [n] \setminus S)$, ce qui est un problème de minimisation de fonction sous-modulaire.

2.9 Logarithmes de déterminants Soit P une matrice symétrique définie positive, à n lignes et n colonnes. Pour $S \subseteq [n]$, on note P_S la sous-matrice carrée dont les lignes et les colonnes sont indicées par les éléments de S. Noter que P_S est symétrique définie positive pour tout S.

Soit f la fonction $2^{[n]} \to \mathbb{R}$ définie par $f(S) \coloneqq \log\det P_S$.

Proposition 14. *La fonction f est sous-modulaire.*

Démonstration. La matrice P peut s'écrire $M^\top M$, où M est une matrice carrée $n \times n$. Notons $\vec{v}_j$ le vecteur de coordonnées la $j$ème colonne de M. Soient $S \subseteq T \subseteq V \setminus \{i\}$. On peut montrer que l'on a

$$f(S \cup \{i\}) - f(S) = \text{ hauteur } \vec{v}_j \text{ par rapport à l'hyperplan engendré par les } (\vec{v}_j)_{j \in S}$$

et de même,

$$f(T \cup \{i\}) - f(T) = \text{ hauteur } \vec{v}_j \text{ par rapport à l'hyperplan engendré par les } (\vec{v}_j)_{j \in T}\,.$$

(Pour ce faire, on utilise le fait que le volume du parallélépipède des vecteurs $(v_j)_{j \in B}$ est égal à $\sqrt{\det P_B}$, pour tout $B \subseteq [n]$.) La conclusion désirée suit immédiatement. □

3 Polymatroïdes

On se donne V un ensemble fini et $f \colon V \to \mathbb{R}$ une fonction sous-modulaire quelconque. Les polyèdres

$$\begin{array}{rcl} P_f & \coloneqq & \{x \in \mathbb{R}_+^V \colon x(S) \leqslant f(S) \text{ pour tout } S \subseteq V\} \\ EP_f & \coloneqq & \{x \in \mathbb{R}^V \colon x(S) \leqslant f(S) \text{ pour tout } S \subseteq V\} \end{array}$$

sont respectivement le *polymatroïde associé à f* et le *polymatroïde étendu associé à f*. Remarquer que P_f est non vide si et seulement si $f(S) \geqslant 0$ pour tout $S \subseteq V$ et EP_f est non vide si et seulement si $f(\varnothing) \geqslant 0$.

Considérons un vecteur c quelconque de $\mathbb{R}^V$. Si $f(\varnothing) < 0$, alors clairement $\max_{x \in EP_f} c \cdot x = -\infty$. Si $f(\varnothing) \geqslant 0$ et une des composantes de c au moins est strictement négative, alors on peut aisément voir que $\max_{x \in EP_f} c \cdot x = +\infty$. Pour traiter complètement le problème de la maximisation de $x \mapsto c \cdot x$ sur EP_f, il reste donc à considérer le cas où $f(\varnothing) \geqslant 0$ et $c \geqslant \mathbf{0}$.

THÉORÈME 15. *Si $f(\varnothing) \geqslant 0$ et $c \geqslant \mathbf{0}$, alors x^* défini comme suit est un point de EP_f où $c \cdot x$ atteint son maximum :*

— *Écrire $V = \{v_1, \ldots, v_n\}$ de sorte que $c_{v_1} \geqslant \cdots \geqslant c_{v_n}$.*
— *Définir $U_i := \{v_1, \ldots, v_i\}$ pour $i = 1, \ldots, n$.*
— *Définit $x^*_{v_1} := f(U_1)$.*
— *Définir $x^*_{v_i} := f(U_i) - f(U_{i-1})$ pour $i = 2, \ldots, n$.*

DÉMONSTRATION. Vérifions d'abord que x^* est réalisable, c'est-à-dire que $x^*(S) \leqslant f(S)$ pour tout $S \subseteq V$. Cela se fait par récurrence sur $|S|$. L'inégalité souhaitée est trivialement satisfaite quand $S = \varnothing$. Soit S un sous-ensemble quelconque non vide de V. Notons k le plus grand indice tel que $v_k \in S$. Si $k = 1$, alors $S = U_1$ et $x^*(S) \leqslant f(S)$ par définition. Si $k \geqslant 2$, on a alors

$$x^*(S) = x^*(S \setminus \{v_k\}) + x^*_{v_k} \leqslant f(S \setminus \{v_k\}) + f(U_k) - f(U_{k-1}) \leqslant f(S)\,.$$

La première inégalité vient de l'hypothèse de récurrence ; la seconde du lemme 12 appliqué à $S \setminus \{v_k\} \subseteq U_{k-1}$.

Pour montrer l'optimalité de x^*, on va exhiber une solution réalisable du problème dual. Ce dernier s'écrit :

$$\begin{array}{lll} \min & \displaystyle\sum_{S \in 2^V \setminus \{\varnothing\}} f(S) y_S & \\ \text{s.c.} & \displaystyle\sum_{S \subseteq V \,:\, v \in S} y_S = c_v & \forall v \in V \\ & y_S \geqslant 0 & \forall S \in 2^V \setminus \{\varnothing\}\,. \end{array}$$

Définissons alors pour $S \neq \varnothing$

$$y^*_S := \begin{cases} c_{v_i} - c_{v_{i+1}} & \text{si } S = U_i \text{ et } i \in [n-1], \\ c_{v_n} & \text{si } S = V, \\ 0 & \text{sinon.} \end{cases}$$

Ce vecteur $y^* \in \mathbb{R}_+^{2^V \setminus \{\varnothing\}}$ est trivialement solution réalisable du problème dual. Comme on a

$$c \cdot x^* = c_{v_1} f(U_1) + \sum_{i=2}^{n} c_{v_i}(f(U_i) - f(U_{i-1})) = \sum_{i=1}^{n-1} f(U_i)(c_{v_i} - c_{v_{i+1}}) + f(V) c_{v_n} = \sum_{S \in 2^V \setminus \{\varnothing\}} f(S) y^*_S\,,$$

le point x^* maximise $c \cdot x$ sur EP_f (et y^* est solution optimale du problème dual). □

On a le corollaire suivant. Le caractère polynomial de l'optimisation sur P_f est une conséquence du fait que, par les ellipsoïdes, on sait séparer les inégalités de EP_f en temps polynomial. Dans le cas où f est croissante, l'algorithme donné par le théorème 15 donne aussi la solution optimale pour P_f (en notant que l'on peut se restreindre à la face de P_f supportée par les $x_v = 0$ avec $c_v < 0$).

COROLLAIRE 16. *Supposons f donnée sous forme d'un oracle et un vecteur $c \in \mathbb{Q}^V$. Le problème de maximisation de $c \cdot x$ sur EP_f et sur P_f peut être résolu en temps polynomial.*

On a aussi :

COROLLAIRE 17. *Si f est à valeurs entières, alors EP_f et P_f sont des polyèdres entiers.*

COROLLAIRE 18. *On a $f(S) = \max\{x(S) \colon x \in EP_f\}$ pour tout $S \in 2^V \setminus \{\varnothing\}$.*

DÉMONSTRATION. Facile. □

4 Minimiser une fonction sous-modulaire

Soit $f\colon V \to \mathbb{R}$ une fonction sous-modulaire telle que $f(\varnothing) = 0$. Définissons pour tout $S \subseteq V$ la quantité $\bar{f}(S) \coloneqq \min_{T \subseteq S} f(T)$.

LEMME 19. *La fonction $\bar{f}$ est une fonction sous-modulaire et $EP_{\bar{f}} = \{x \in EP_f \colon x \leqslant \mathbf{0}\}$.*

DÉMONSTRATION. □

THÉORÈME 20. *Étant donnée une fonction $f\colon V \to \mathbb{R}$ sous-modulaire, trouver S minimisant $f(S)$ peut se faire en temps polynomial.*

DÉMONSTRATION. Sans perte de généralité, on peut supposer que $f(\varnothing) = 0$. À partir du corollaire 18 appliqué à $\bar{f}$ et du lemme 19, et en utilisant l'équivalence entre optimisation et séparation. □

La preuve donnée ci-dessus s'appuie sur l'algorithme des ellipsoïdes, qui n'est malheureusement pas implémentable. Un algorithme combinatoire, implémentable, a été proposé par Schrijver en 2000.

5 Intersection de deux polymatroïdes

Tout comme pour les matroïdes, on peut considérer des polymatroïdes qui s'intersectent. Soient f_1 et f_2 deux fonctions sous-modulaires sur un ensemble V, et $a, b \in (\mathbb{Z} \cup \{-\infty, +\infty\})^V$. Considérons le programme linéaire suivant.

$$\text{(P)} \qquad \begin{array}{lll} \max & c \cdot x & \\ \text{s.c.} & x(S) \leqslant f_1(S) & \forall S \subseteq V \\ & x(S) \leqslant f_2(S) & \forall S \subseteq V \\ & a_i \leqslant x_i \leqslant b_i & \forall i \in V\,. \end{array}$$

THÉORÈME 21. *Le dual de (P) admet une solution optimale entière pour tout vecteur entier c pour lequel il est de valeur optimale finie.*

DÉMONSTRATION. On écrit la preuve dans le cas où toutes les composantes de a et b sont finies. Le dual de (P) est

$$\text{(D)} \qquad \begin{array}{lll} \min & \displaystyle\sum_{S \in 2^V \setminus \{\varnothing\}} (f_1(S) y_S^1 + f_2(S) y_S^2) + \sum_{i \in V} (b_i z_i' - a_i z_i) & \\ \text{s.t.} & z_i' - z_i + \displaystyle\sum_{S \subseteq V \colon i \in S} (y_S^1 + y_S^2) = c_i & \forall i \in V \\ & y_S^1, y_S^2 \geqslant 0 & \forall S \in 2^V \setminus \{\varnothing\} \\ & z_i, z_i' \geqslant 0 & \forall i \in V\,. \end{array}$$

Par dualité forte, il admet une solution optimale $(\bar{y}^1, \bar{y}^2, \bar{z}, \bar{z}')$. Posons $c_i^k \coloneqq \bar{z}_i' - \bar{z}_i + \sum_{S \subseteq V \colon i \in S} \bar{y}_S^k$ pour $k \in \{1, 2\}$ et $i \in V$. Remarquer qu'alors $\bar{y}^k$ est solution optimale du programme linéaire

$$\begin{array}{lll} \min & \displaystyle\sum_{S \in 2^V \setminus \{\varnothing\}} f_k(S) y_S^k & \\ \text{s.t.} & \displaystyle\sum_{S \subseteq V \colon i \in S} y_S^k = c_i^k & \forall i \in V \\ & y_S^k \geqslant 0 & \forall S \in 2^V \setminus \{\varnothing\}\,. \end{array}$$

La preuve du théorème 15 montre que l'on peut choisir les $\bar{y}^k$ de sorte que les parties S telles que $\bar{y}_S^k > 0$ forment une chaîne pour l'inclusion. Une matrice dont les colonnes peuvent être partitionnées en deux matrices d'incidence d'un système laminaire est totalement unimodulaire (voir...). Cela implique que la matrice des contraintes de (D) restreinte au support de la solution optimale $(\bar{y}^1, \bar{y}^2, \bar{z}, \bar{z}')$ est totalement unimodulaire. Le résultat souhaité en découle directement. □

Corollaire 22. *Soient f_1 et f_2 deux fonctions sous-modulaires sur V telles que $f_1(\varnothing) = f_2(\varnothing) = 0$. On a l'égalité suivante :*

$$\max\{x(V) \colon x \in EP_{f_1} \cap EP_{f_2}\} = \min\{f_1(S) + f_2(V \setminus S)\}\,.$$

Démonstration. Prendre c le vecteur dont toutes les composantes sont égales à 1. On prend $a_i = -\infty$ et $b_i = +\infty$ pour tout $i \in V$. On applique la dualité forte. Soit $(\bar{y}^1, \bar{y}^2)$ une solution optimale entière du dual. Cette solution est à coefficients dans $\{0,1\}$, et pour tout i, une seule partie S de V contenant i est telle que $\bar{y}^1_S + \bar{y}^2_S = 1$. De plus, si deux parties distinctes S, T sont telles que $\bar{y}^1_S = \bar{y}^1_T = 1$, alors $f_1(S \cup T) \leqslant f_1(S)\bar{y}^1_S + f_1(T)y^1_T$ car $f_1(\varnothing) = 0$, et on peut "fusionner" S et T. □

Corollaire 23. *Soient f_1 et f_2 deux fonctions sous-modulaires croissantes sur V telles que $f_1(\varnothing) = f_2(\varnothing) = 0$. On a l'égalité suivante :*

$$\max\{x(V) \colon x \in P_{f_1} \cap P_{f_2}\} = \min\{f_1(S) + f_2(V \setminus S)\}\,.$$

Démonstration. Même preuve que pour le corollaire 22, à ceci près que l'on prend $a_i = 0$ pour tout $i \in V$. □

Le théorème 8 est alors immédiat. Considérons le cas où $f_1 = r_{M_1}$ et $f_2 = r_{M_2}$. Dans ce cas, par le théorème 15, l'ensemble des solutions réalisables de (P) est un polyèdre entier (valeur optimale entière pour tout coût entier). En prenant c le vecteur dont toutes les composantes sont égales à 1, le corollaire 23 donne alors le résultat souhaité.

6 Maximiser une fonction sous-modulaire

7 Extension de Lovász