Théorème de Birkhoff et lien avec le transport optimal discret

__Définitions :__

(1) Soit $M \in \mathcal{M}_n([0,1])$. Elle est dite bistochastique si la somme des éléments sur les lignes et les colonnes fait 1, ie

$$\sum_i m_{ij} = \sum_j m_{ij} = 1$$

(2) Soit E evn et K convexe compact de E. L'ensemble des points extrémaux de K est

$$\mathcal{E}(K) = \left\{ x \in K, \left(\forall a,b \in K, \forall \lambda \in [0,1], \; x = \lambda a + (1-\lambda) b \right) \Rightarrow \left(a = b \text{ ou } \lambda \in \{0,1\} \right) \right\}$$

Ce sont les pts de K qui ne peuvent pas s'écrire comme combinaison linéaire non triviale de points de K.

__Théorème (Birkhoff - von Neumann) :__

B_n = ensemble des matrices bistochastiques de taille $n \times n$.

Alors $\mathcal{E}(B_n) = \{ \text{matrices de permutation} \}$.

__Preuve :__

- __B_n est convexe et compact :__

↳ si $A, B \in B_n$, $\sum_i (\lambda a_{ij} + (1-\lambda) b_{ij}) = \lambda + (1-\lambda) = 1 \quad \forall \lambda \in [0,1]$

↳ $B_n \subset \mathcal{M}_n([0,1])$ compact et B_n est fermé, donc compact.

• $\underline{\underline{\{\text{matrices de permutation}\} \subset \mathcal{E}(B_n)}}$:

Soit $\sigma \in S_n$ et $M_\sigma = \text{Mat}(\sigma) \in \{\text{matrices permutation}\}$.

$M_\sigma \in M_n(\{0,1\})$.

Supposons que M_σ s'écrive comme combinaison linéaire non triviale d'éléments de B_n :

$$\exists \lambda \in]0,1[, \ \exists A, B \in B_n \text{ tq } M_\sigma = \lambda A + (1-\lambda) B.$$

En chaque position (i,j), on a

$$\{0,1\} \ni m_{ij} = \lambda a_{ij} + (1-\lambda) b_{ij}.$$

Deux cas :

(1) $0 = \lambda a_{ij} + (1-\lambda) b_{ij} \Rightarrow a_{ij} = b_{ij} = 0$ car tout positif.

$\Rightarrow a_{ij} = b_{ij} = m_{ij}$ dans ce cas.

(2) $1 = \lambda a_{ij} + (1-\lambda) b_{ij}$.

Supposons par l'absurde que $a_{ij} < 1$, $b_{ij} \leq 1$. Alors

$1 = \lambda a_{ij} + (1-\lambda) b_{ij} < \lambda + (1-\lambda) = 1$ $\lightning$.

Donc $a_{ij} = b_{ij} = 1$

$\Rightarrow a_{ij} = b_{ij} = m_{ij}$ dans ce cas.

On a donc montré que $A = B = M_\sigma$. <u>Donc M_σ est extrémale</u>.

• $\underline{\underline{\mathcal{E}(B_n) \subset \{\text{Matrices de permutation}\}}}$

On va montrer que tout point extrémal de B_n est une matrice à coefs dans $\{0,1\}$. C'est suffisant, car la matrice étant bistochastique, on aura au plus un 1 par ligne et par colonne. Ce sera donc bien une matrice de permutation.

Soit $M \in \mathcal{E}(B_n)$ une matrice extrémale. Supposons par l'absurde qu'un de ses coefficients soit strictement compris entre 0 et 1, ie

$$\exists\ i_0, j_0 \ \text{tq} \quad m_{i_0, j_0} \in \left]0, 1\right[.$$

Puisque M est bistochastique, il y a nécessairement un autre coef non nul sur la même ligne : $\exists\ m_{i_0, j_1} \in \left]0,1\right[$, $j_1 \neq j_0$. (la Σ doit $= 1$).
De même, la somme sur la colonne j_1 devant être 1, il y a un coef $m_{i_1, j_1} \in \left]0,1\right[$. Etc...

$$\begin{bmatrix} m_{i_0 j_0} & m_{i_0 j_1} \\ & m_{i_1 j_1} \end{bmatrix}$$

On construit ainsi deux suites d'indices (i_s) et (j_s) tels que

$$\begin{cases} i_s \neq i_{s+1} \\ j_s \neq j_{s+1} \\ m_{i_s, j_s} \in \left]0,1\right[\quad \text{et} \quad m_{i_s, j_{s+1}} \in \left]0,1\right[\end{cases}$$

Ces deux suites étant à valeurs dans un ensemble fini, elles vont se répéter : $\exists\ k_0, k_1, l_0, l_1$ tq $\begin{cases} i_{k_0} = i_{k_1} \quad \text{et} \quad j_{l_0} = j_{l_1}, \\ m_{i_{k_0}, j_{l_0}} = m_{i_{k_1}, j_{l_1}}. \end{cases}$

L'idée est de perturber la matrice M de deux manières différentes et d'écrire M comme combinaison linéaire non triviale des deux matrices perturbées ainsi obtenues. Difficulté : les matrices perturbées restent-elles bistochastiques ?

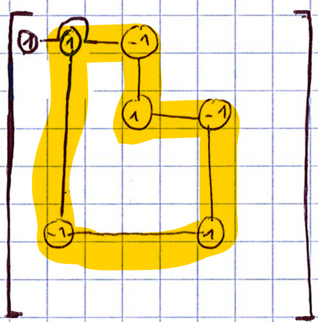
On a deux cas à voir :

1:) Notons k le premier rang d'apparition du terme qui se répète, et l le second. On suppose dans ce cas que les indices $\{j_k, \dots, j_l\}$ sont distincts 2 à 2.

On définit alors une matrice N de la manière suivante :

$$\begin{cases} \forall s \in [\![k+1, l]\!], & n_{i_s, j_s} = 1 \\ \forall s \in [\![k, l-1]\!], & n_{i_s, j_{s+1}} = -1 \\ \text{sinon}, & n_{i,j} = 0 \end{cases}$$

Ex :



Sur ce chemin ci-dessus, on ne garde que les termes non nuls orange, et on met des 0 partout ailleurs.

La somme sur les lignes et les colonnes fait toujours 0.

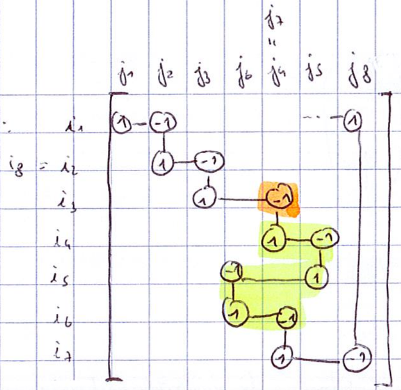
2:) Cas pathologique : on a $\{j_k, \dots, j_l\}$ pas distincts.

On prend alors $k', l' \in [\![k, l]\!]$ tq

$$\begin{cases} j_{k'}, \dots, j_{l'-1} \text{ distincts } 2 \text{ à } 2 \\ j_{k'} = j_{l'} \end{cases}$$

On définit alors N par
$$\begin{cases} \forall s \in [\![k', l'-1]\!], & n_{i_s, j_s} = 1 \\ \forall s \in [\![k', l'-1]\!], & n_{i_s, j_{s+1}} = -1 \\ \text{sinon}, & n_{i,j} = 0 \end{cases}$$

Ex :



$k' = 4, \quad l' = 7$

On ne met des ± 1 que sur la partie jaune, et des 0 partout ailleurs.

La matrice N ainsi obtenue vérifie la m̂ propriété que dans le cas 1 :

la somme sur les lignes et les colonnes est toujours 0.

Posons maintenant $0 < \varepsilon < \min\limits_{i,j} (m_{ij}, 1 - m_{ij})$

Par les propriétés de N mises en évidence, $M + \varepsilon N$ et $M - \varepsilon N$ sont bistochastiques.

Bien sûr, on a $M = \frac{1}{2}(M + \varepsilon N) + \frac{1}{2}(M - \varepsilon N)$, ce qui est une contradiction puisque M est supposée extrémale.

On a donc montré $M \in \mathcal{E}(B_n) \Rightarrow M \in M_n(\{0,1\}) \cap B_n$

$\Rightarrow$ $\underline{M \text{ matrice de permutation}}$.

<u>Conclusion</u> : $M \in \mathcal{E}(B_n) \Leftrightarrow M$ est une matrice de permutation.

■

<u>Lien avec le transport optimal</u> :

<u>Rappels de l'exposé d'Armand</u> (avant les vacances : 21/10)

- $E = \{x_1, \dots, x_m\}$ $\quad F = \{y_1, \dots, y_n\}$
- $c : E \times F \to [0, +\infty]$
- $\mu = \sum\limits_{i=1}^{m} p_i \delta_{x_i}$; $\quad \nu = \sum\limits_{j=1}^{n} q_j \delta_{y_j}$

- $p = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix}$; $q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$

- $\tilde{c} = \left(c(x_i, y_j)\right)_{i,j} \in \mathcal{M}_{mn}(\mathbb{R})$.

On a vu que se donner un plan de transport $\mu \to \nu$ revient à se donner une matrice.

Aujourd'hui, on va supposer en plus que

(i) $m = n$

(ii) les mesures sont uniformes : $\begin{cases} \mu = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \\ \nu = \frac{1}{n} \sum_{j=1}^{n} \delta_{y_j} \end{cases}$

<u>Ex</u> : n usines produisant chacune une caisse de matériel à distribuer entre n destinations.

On se donne donc une matrice $M = (m_{ij})_{i,j}$. Afin d'assurer la conservation de la masse, cette matrice doit être bistochastique :

$$\sum_i m_{ij} = \sum_j m_{ij} = 1.$$

On veut alors minimiser la fonctionnelle

$$J \quad M \longmapsto \sum_{i,j=1}^{n} m_{ij} \cdot c(x_i, x_j) \quad , \quad M \in \mathcal{B}_n.$$

<u>Théorème</u> (Choquet) :

Soit K un compact d'un evn E, $\ell : K \to \mathbb{R}$ une fonctionnelle linéaire continue. Alors ℓ atteint son minimum sur $\mathcal{E}(K)$.
un ?

↳ dans le cas de la dim finie, ce n'est pas la peine d'utiliser ce théorème : un argument rapide utilisant le théorème de Krein-Milman suffit.

Ainsi notre fonctionnelle J satisfait aux hypothèses de ce théorème. Elle atteint donc son minimum en un point extrémal. Mais par Birkhoff, toute matrice extrémale de B_n est une matrice de permutation.
Un ~~Le~~ plan de transport optimal est donc dans ce cas une permutation.
Pour l'ex des usines, on interprète cela en disant que chaque caisse de matériel doit être expédiée entière. En revanche, on ne sait pas quelle permutation est celle qui réalise ~~la~~ une solution optimale.

Pour le cas de masses rationnelles, on réduit au même dénominateur et on normalise, de manière à pouvoir supposer que les masses sont entières. On se ramène au cas uniforme en scindant un point de masse $m \in \mathbb{N}$ en m points de masse 1 situés au même endroit.

Suite (Guido) :

voir le document de Guido pour plus de détails

P. Hall, 1935

X ensemble, $X_1, X_2, \ldots, X_n$ parties de X (possible que $X_i = X_j$)

Section transversale de la famille : $(X_i, i \in I)$

C'est une suite $(x_i)_i$, $x_i \in X_i$, $x_i \neq x_j$ si $i \neq j$

Théorème :

Une famille $(X_i)_i$ admet une transversale si et seulement si

$$\forall\ 1 \leq k \leq n \quad \text{et} \quad \forall\ 1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq n,$$

$$|X_{i_1} \cup X_{i_2} \cup \ldots \cup X_{i_k}| \geq k$$

Preuve : par récurrence sur n

2 cas : un facile et un difficile

↳ vraie pour toute sous-famille de $(X_i)_i$ avec $n-1$ sous-ensembles des ensembles $\{X_i\}$

il subsiste un petit trou dans le cas $k = n$

voir notes de Guido

① $1 \leq k \leq n$; $1 \leq i_1 < i_2 < \ldots < i_k < n$

$$|X_{i_1} \cup \ldots \cup X_{i_k}| \geq k+1$$

si $x_1 \in X_{i_1}$, $\left(X_i \setminus \{x_1\},\ i \in I \setminus \{1\} \right)$.

② $\exists\ 1 \leq k \leq n$, $|X_{i_1} \cup \ldots \cup X_{i_k}| = k$.

supposons $i_1, \ldots i_k$ consécutifs : $|X_1 \cup \ldots \cup X_k| = k$

vérifie la condition

$\Downarrow$

$\exists$ une transversale $\{x_i,\ 1 \leq i \leq k\}$.

On introduit la nouvelle famille $X_i' = X_i \setminus \{x_1, \ldots, x_k\}$, $i > k$.

Affirmation : cette nouvelle famille vérifie la condition de Hall

$1 \leq \ell \leq n-k$; $1 \leq d_1 < d_2 < \ldots < d_\ell \leq n-k$

On calcule $|X'_{k+d_1} \cup \ldots \cup X'_{k+d_\ell}|$

$$= |X'_{k+d_1} \cup \ldots \cup X'_{k+d_\ell}| + |X_{i_1} \cup \ldots \cup X_{i_k}| - k$$

$$\geq |X_{i_1} \cup \ldots \cup X_{i_k} \cup X'_{k+d_1} \cup \ldots \cup X'_{k+\ell}|$$

$$\geq (k+\ell) - k = \ell .$$

■

Lien avec Birkhoff

Théorème :

$A = (a_{ij})$ tq $\sum_j a_{ij} = t$, $\sum_i a_{ij} = t$, $t > 0$

Alors $\exists\ \alpha_1, \dots, \alpha_s > 0$ tq $\sum \alpha_j = t$
et $\exists\ P_1, \dots, P_s \in \mathcal{P}_n$ permutation
tq $A = \alpha_1 P_1 + \dots + \alpha_s P_s$.

On fait la preuve sur un exemple

$$A = \begin{array}{c} \left[\begin{array}{cc|ccc} 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ \hline 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right] \begin{array}{c} 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{array} \\ \begin{array}{ccccc} 3 & 3 & 3 & 3 & 3 \end{array} \end{array}$$

$w :=$ nbre elts non nuls de A.

ici, $X_1 = \{3, 5\}$ $\quad X_4 = \{2, 4\}$
$X_2 = \{4, 5\}$ $\quad X_5 = \{1, 2, 5\}$
$X_3 = \{1, 3\}$

Cette famille vérifie la condition de Hall.
On prend les premiers termes de chaque X_i jusqu'à $k-1$: $\{3, 4, 1, 2\}$
et je fais une matrice de permutation avec :

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} := P_1$$

(la position du dernier 1 est déterminée par les quatre autres)

$A_1 := A - P_1$ a donc plus de zéros que A et la somme sur les ~~termes~~ lignes et colonnes fait 2.

On continue ... on obtient $A = P_1 + P_2 + P_3$ (exo)

les α_i sont les min des coefficients.

Dans le cas des matrices rectangles, de marges $p = \begin{pmatrix} p_1 \\ \vdots \\ p_m \end{pmatrix}$, $q = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix}$

$U(p,q) = \{ A,\ \text{marges } p \text{ et } q \}$

Qui est $\mathcal{E}(U(p,q))$?

Peut-on décomposer les matrices rectangles comme on a fait avant ?