

# Un théorème de Birkhoff-von Neumann

Jules Nodé-Langlois

## I. Le théorème

**Théorème de Birkhoff-von Neumann :** Les points extrémaux des matrices bistochastiques sont les matrices de permutations.

## II. Matrices magiques

On commence par montrer quelques propriétés élémentaires sur les matrices dites magiques. On fixe un entier  $n \in \mathbb{N}^*$ .

### II.1) Définitions

Définition. Une matrice  $M \in \mathcal{M}_n(\mathbb{R})$  est dite magique s'il existe  $\lambda \in \mathbb{R}$  tel que :

$$\forall i, j \in \{1, \dots, n\}, \sum_{k=1}^n m_{kj} = \sum_{k=1}^n m_{ik} = \lambda$$

Le cas échéant, on note  $\sigma(M) = \lambda$ . De plus, on note  $\mathfrak{M}$  l'ensemble des matrices magiques.

Notons  $\mathcal{C} = (e_1, \dots, e_n)$  la base canonique de  $\mathbb{R}^n$ ,  $H = \ker(\sum_{i=1}^n e_i^*)$  et  $X_0 = \sum_{i=1}^n e_i$ . On remarque tout d'abord que :

$$\mathbb{R}^n = H \oplus \mathbb{R}X_0$$

**Propriété :** Une matrice  $M \in \mathcal{M}_n(\mathbb{R})$  est magique si et seulement si  $H$  est stable par  $M$  et  $X_0$  est un vecteur propre pour  $M$ .

*Preuve.*

Soit  $M \in \mathcal{M}_n(\mathbb{R})$  magique. Alors  $X_0$  est vecteur propre pour  $M$  associé à la valeur propre  $\sigma(M)$ . Tandis que pour  $x \in H$ , on a :

$$\sum_{i=1}^n (MX)_i = \sum_{i=1}^n \sum_{k=1}^n m_{i,k} x_k = \sum_{k=1}^n x_k \sum_{i=1}^n m_{i,k} = \sum_{k=1}^n x_k \sigma(M) = 0$$

Donc  $H$  est stable par  $M$ . Réciproquement si  $M \in \mathcal{M}_n(\mathbb{R})$  stabilise  $H$  et admet  $X_0$  comme vecteur propre,

il existe  $\lambda \in \mathbb{R}$  :

$$MX_0 = \lambda X_0$$

Soit  $\varphi = \sum_{k=1}^n e_k^*$ ,  $H$  est stable par  $M$  donc en notant  $m$  l'application linéaire canoniquement associée à  $M$ , on obtient que  $\ker \varphi \subset \ker \varphi \circ m$ . En particulier, les formes linéaires sont colinéaires :

$$\exists \mu \in \mathbb{R}, \mu \varphi = \varphi \circ m$$

Donc :

$$\mu n = \lambda n$$

Donc  $\mu = \lambda$ . Et pour tout  $x \in \mathbb{R}^n$  :

$$\sum_{i=1}^n \sum_{k=1}^n m_{i,k} x_k = \lambda \sum_{k=1}^n x_k$$

Avec cette égalité sur la base canonique on obtient :

$$\forall i \in \{1, \dots, n\}, \sum_{k=1}^n m_{k,i} = \lambda$$

Finalement,  $M \in \mathfrak{M}$ .

**Théorème :**  $\mathfrak{M}$  est une sous-algèbre de  $\mathcal{M}_n(\mathbb{R})$  et l'application  $\sigma : \mathfrak{M} \rightarrow \mathbb{R}$  est un morphisme d'algèbres.

**Propriété :** Si  $M \in \mathfrak{M} \cap GL_n(\mathbb{R})$ , alors  $M^{-1} \in \mathfrak{M}$  et  $\sigma(M^{-1}) = \frac{1}{\sigma(M)}$ .

## II.2) Calcul de la dimension par dualité

On calcule ici la dimension de l'algèbre des matrices magiques. On définit pour tout  $i \in \{1, \dots, n\}$ ,  $f_i \in (\mathcal{M}_n(\mathbb{R}))^*$ ,  $g_i \in (\mathcal{M}_n(\mathbb{R}))^*$  par :

$$f_i : M \mapsto \sum_{k=1}^n m_{i,k} \text{ et } g_i : M \mapsto \sum_{k=1}^n m_{k,i}$$

On remarque dans un premier temps que  $\mathbb{R}I_n$  est un supplémentaire de  $\ker(\sigma)$  dans  $\mathfrak{M}$ .

**Propriété :** Le rang de la famille  $(f_1, \dots, f_n, g_1, \dots, g_n)$  est  $2n - 1$ .

On considère alors les espaces vectoriels :

$$F = \bigcap_{k=1}^n \ker f_k, G = \bigcap_{k=1}^n \ker g_k$$

et enfin :

$$W = F \cap G$$

On sait que  $\dim W = n^2 - (2n - 1)$ , or  $W = \ker \sigma$ , en particulier on obtient :

$$\dim \mathfrak{M} = n^2 - 2n + 2$$

### III. Matrices de permutation

Notons  $(E_{i,j})_{i,j \in [1,n]}$  la base canonique de  $\mathcal{M}_n(\mathbb{R})$ . Notons  $J = \sum_{j=1}^n \sum_{i=1}^n E_{i,j}$ , on vérifie aisément que  $J \in \mathfrak{M}$ . On définit pour tout  $s \in \mathfrak{S}_n$ , l'application linéaire  $\rho(s) : \mathbb{R}^n \rightarrow \mathbb{R}^n$  sur la base canonique par  $\rho(s) : e_i \mapsto e_{s(i)}$  pour tout  $i \in \{1, \dots, n\}$ . On rappelle que  $(\mathbb{R}^n, \rho)$  est une représentation de  $\mathfrak{S}_n$  dont  $H$  et  $\mathbb{R}X_0$  sont des sous-représentations non triviales. De plus on note pour tout  $s \in \mathfrak{S}_n$ ,  $P_s = \text{Mat}_{\mathcal{C}}(\rho(s))$ . La caractérisation des matrices magiques montre en fait que pour tout  $s \in \mathfrak{S}_n$ ,  $P_s \in \mathfrak{M}$ .

**Définition :** On définit le commutant de  $\mathfrak{M}$  par  $\mathcal{K}(\mathfrak{M}) = \{M \in \mathfrak{M} | \forall A \in \mathfrak{M}, [A, M] = 0\}$ .

**Théorème.**

$$\mathcal{K}(\mathfrak{M}) = \text{Vect}(I_n, J)$$

*Preuve.*  $I_n$  commute avec toutes la matrices magiques. Soit  $M \in \mathfrak{M}$ , on remarque de  $JM = \sigma(M)J$  et de même que  $MJ = \sigma(M)J$  donc que  $JM = MJ$ ,  $\mathcal{K}(\mathfrak{M})$  étant un sous-espace vectoriel on obtient l'inclusion  $\text{Vect}(I_n, J) \subset \mathcal{K}(\mathfrak{M})$ . Réciproquement, soit  $M \in \mathcal{K}(\mathfrak{M})$ , en particulier  $M$  commute avec toutes les matrices de permutations. En particulier avec les matrices de transpositions. Soient  $i, j \in [1, n], i \neq j$  et  $\tau = (i, j)$ . Pour tout  $k \in [1, n]$  :

$$P_{\tau} M e_k = M P_{\tau} e_k$$

On note  $M$  selon la concaténation de ses colonnes :  $M = (C_1 | \dots | C_n)$ , on a donc :

$$C_k = P_{\tau} C_k, \text{ si } k \neq i, k \neq j$$

et :

$$C_i = P_{\tau} C_j, \text{ sinon}$$

En particulier :

$$m_{i,k} = m_{j,k}, \text{ si } k \neq i, k \neq j$$

et :

$$m_{i,i} = m_{j,j}, \text{ sinon}$$

C'est donc bien que  $M \in \text{Vect}(I_n, J)$  et que  $\mathcal{K}(\mathfrak{M}) \subset \text{Vect}(I_n, J)$ .

### IV. Matrices bistochastiques

**Définition.** Une matrice  $M \in \mathcal{M}_n(\mathbb{R})$  est dite bistochastique si  $M \in \mathfrak{M}$  avec  $\sigma(M) = 1$  et pour tout  $i, j \in \{1; \dots; n\}$ ,  $m_{i,j} \geq 0$ . On note  $\mathfrak{B}$  l'ensemble de ces matrices.

Par exemple, une matrice de permutation est toujours bistochastique.

**Théorème.**  $\mathfrak{B}$  est une partie génératrice de  $\mathfrak{M}$ .

*Preuve.* On pose pour tout  $i, j \in \{1, \dots, n-1\}$ ,  $B_{i,j} = E_{i,j} - E_{i,n} - E_{n,j} + E_{n,n}$  qui est magique avec  $\sigma(B_{i,j}) = 0$  et  $\tilde{B}_{i,j} = \frac{1}{n}(J + B_{i,j})$  qui est bistochastique. Il suffit alors de vérifier que les  $(\tilde{B}_{i,j})_{i,j \in \{1, \dots, n\}}$  forment une famille libre. Et alors par dimension :

$$\boxed{\text{Vect}(\mathfrak{B}) = \mathfrak{M}}$$

## V. Retour au théorème de Birkhoff

Soit  $s \in \mathfrak{S}_n$ , et soient  $A, B \in \mathfrak{B}$ , et  $\lambda \in ]0, 1[$  tels que :  $P_s = \lambda A + (1 - \lambda)B$ . Alors pour tout  $i \in \{1, \dots, n\}$ , et pour tout  $k \in \{1, \dots, n\}, k \neq \sigma^{-1}(i)$  :  $\lambda A_{i,k} + (1 - \lambda)B_{i,k} = 0$ , soit par positivité des coefficients :  $A_{i,k} = 0$  et  $B_{i,k} = 0$ . Finalement c'est que  $A_{i, \sigma^{-1}(i)} = B_{i, \sigma^{-1}(i)} = 1$ . Donc que  $A = B = P_s$ . En particulier, les matrices de permutations sont des points extrémaux de  $\mathfrak{B}$ .

*Définition.* Pour  $M \in \mathcal{M}_n(\mathbb{R})$ , et  $I, J$  deux parties de  $\{1; \dots; n\}$ , on note  $M_{I,J}$  la matrice de  $\mathcal{M}_{|I|,|J|}(\mathbb{R})$  construite de la sorte :  $M = (m_{i,j})_{i \in I, j \in J}$ .

*Définition.* Pour  $M \in \mathcal{M}_n(\mathbb{R})$  et  $\sigma \in \mathfrak{S}_n$  on appelle slalom de  $M$  associé à  $\sigma$  la quantité :

$$s_\sigma(M) = \prod_{i=1}^n m_{\sigma(i), i}$$

*Remarque :* Pour  $\sigma \in \mathfrak{S}_n$  et  $M \in \mathcal{M}_n(\mathbb{R})$ , on a :  $P_\sigma M$  est la matrice  $M$  dont a permuté les lignes selon  $\sigma$  et  $M P_\sigma$  la matrice  $M$  dont on a permuté les colonnes selon  $\sigma$ .

**Théorème.** Soit  $M \in \mathcal{M}_n(\mathbb{R})$  telle que pour tout  $\sigma \in \mathfrak{S}_n$ ,  $s_\sigma(M) = 0$  alors il existe  $I, J$  deux parties de  $\{1; \dots; n\}$  telles que  $M_{I,J} = 0$  avec  $|I| + |J| = n + 1$ .

*Preuve.*

Pour tout  $i \in \{1; \dots; n\}$  il existe  $\sigma_i \in \mathfrak{S}_n$  telle que  $m_{\sigma_i(i), i} = 0$  sinon on obtiendrait un slalom non nul. On pose  $I = \{\sigma_i(i) | 1 \leq i \leq n\}$  et  $J = \{i | 1 \leq i \leq n, \exists! j \in \{1; \dots; n\}, \sigma_j(i) = \sigma_i(i)\}$ . On a d'une part par construction :  $M_{I,J} = 0$ . On en déduit alors que si  $M \in \mathcal{M}_n(\mathbb{R})$  est telle que pour tout  $\sigma \in \mathfrak{S}_n$ ,  $s_\sigma(M) = 0$ ,

il existe  $r, s \in \mathbb{N}^*$  tels que  $r + s = n + 1$  et  $\sigma, \sigma' \in \mathfrak{S}_n$  telles que la matrice  $B = P_\sigma M P_{\sigma'}$  vérifie :

$$\forall 1 \leq i \leq r, \forall r \leq j \leq n, b_{i,j} = 0$$

En particulier, une matrice bistochastique admet, nécessairement un slalom non nul.

Soit alors  $A \in \mathcal{M}_n(\mathbb{R})$  une matrice bistochastique et  $\sigma \in \mathfrak{S}_n$  associée a un slalom non nul. On considère  $a = \min_{1 \leq j \leq n} a_{\sigma(j),j}$ .  $A$  étant bistochastique, on a nécessairement :  $a \in [0, 1]$  et le slalom étant non nul,  $a \neq 0$ . Si  $a = 1$ , pour tout  $j \in \{1; \dots; n\}$ ,  $a_{\sigma(j),j} = 1$ , en particulier,  $A = P_\sigma$ . Sinon, si  $a < 1$ , on considère la matrice  $A - aP_\sigma$ , et on remarque qu'il existe  $B \in \mathfrak{B}$  avec strictement plus de coefficients non nuls que  $A$  telle que  $A - aP_\sigma = (1 - a)B$ . Soit  $P$  un point extrémal de  $\mathfrak{B}$ .

## VI. Un corollaire

On suppose ici connus les théorèmes suivants :

**Théorème de Carathéodory** : Dans un espace vectoriel  $E$  réel de dimension  $n \in \mathbb{N}^*$ , l'enveloppe convexe d'un sous-ensemble  $A$  est l'ensemble des barycentres à coefficients positifs ou nuls de familles de  $n+1$  points de  $A$ .

**Théorème de Krein-Milman** : Tout convexe compact d'un espace vectoriel réel de dimension finie est enveloppe convexe de ses points extrémaux.

**Corollaire** :