

Théorie du transport dans le cas discret

Maël Colin

2025

Introduction :

Ce document a pour but d'expliquer toutes les bases du transport optimal. Il n'est pas donc nécessaire de s'y connaître en transport optimal pour le comprendre, cependant, un tel document nécessite tout de même un certain niveau en mathématique (plutôt master 1 puisque la majorité des démonstrations sont des démonstrations de l'agrégation externe). Ce document fait suite à mon 1^{er} document « *l'histoire du transport optimal* ». Il est donc recommandé de l'avoir lu pour comprendre le fil du document. De plus, ce document s'inscrit également dans les travaux de vulgarisation du mathématicien Gabriel Peyré, je recommande donc également de lire son document « *Le transport optimal numérique et ses applications* » ([TransportFR.pdf](#)) pour bien appréhender le miens. Pour ceux qui ont du mal sur certaine notion, même après avoir lu le document, je vous recommande la conférence du même nom ([Gabriel Peyre - Le transport optimal numérique et ses applications](#)).

Sommaire :

- 1 - Définitions
- 2 - Problèmes de Monge et de Kantorovitch
- 3 - Preuve de l'existence d'une solution du problème de Monge
- 4 – Méthode pour trouver une solution au problème de Monge

1 – Définitions

Dans la suite du document, on note $n \in \mathbf{N}$, le nombre de sites de production/d'utilisation de terre, $(.|.)$ le produit scalaire usuelle des matrices et $N(.)$ la norme associée.

De plus, pour toutes matrices carrées de A de taille n , $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2$, on note $a_{i,j}$ le coefficient de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne.

On appelle matrice de permutation toutes les matrices dont les coefficients sont soit 0, soit 1 et dont la somme des coefficients de chaque ligne et de chaque colonne vaut 1. On note alors P_n l'ensemble des matrices de permutation. C'est-à-dire, $\forall A \in M_n(\mathbf{R})$, $A \in P_n$ si et seulement si :

$$\forall (i, j) \in \left(\mathbf{N} \cap [1, n] \right)^2, a_{i,j} \in \{0, 1\}$$

$$\forall i \in \mathbf{N} \cap [1, n] \sum_{j=1}^n a_{i,j} = 1$$

$$\forall j \in \mathbf{N} \cap [1, n] \sum_{i=1}^n a_{i,j} = 1$$

On appelle matrice bistochastique toutes les matrices dont les coefficients sont positifs et dont la somme des coefficients de chaque ligne et de chaque colonne vaut 1. On note alors B_n l'ensemble des matrices bistochastiques. C'est-à-dire, $\forall A \in M_n(\mathbf{R})$, $A \in B_n$ si et seulement si :

$$\forall (i, j) \in \left(\mathbf{N} \cap [1, n] \right)^2, a_{i,j} \geq 0$$

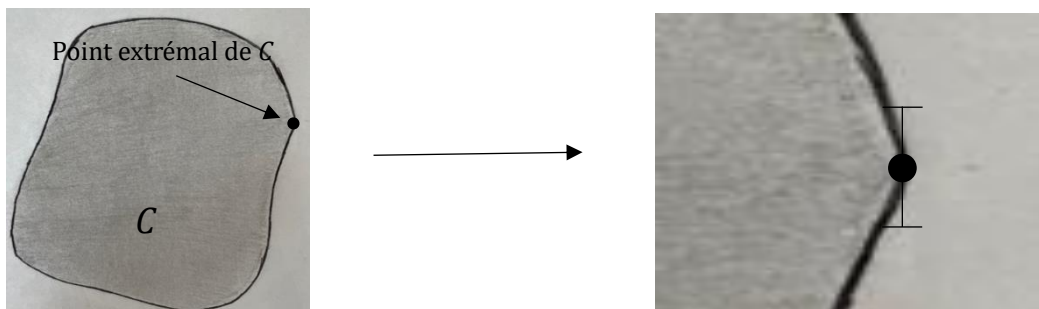
$$\forall i \in \mathbb{N} \bigcap [1, n] \sum_{j=1}^n a_{i,j} = 1$$

$$\forall j \in \mathbb{N} \bigcap [1, n] \sum_{i=1}^n a_{i,j} = 1$$

En particulier, l'ensemble des matrices de permutation est inclus dans l'ensemble des matrices bistochastiques.

Soit C un convexe d'un espace vectoriel E . On dit qu'un point x de C est un point extrémal de C si $C \setminus \{x\}$ est convexe. On note alors $\mathcal{E}(C)$ l'ensemble des points extrémaux de C . C'est-à-dire, $\forall x \in C, x \in \mathcal{E}(C)$ si et seulement si :

$$\forall (a, b) \in C^2, \forall \lambda \in [0, 1], x = \lambda a + (1 - \lambda)b \rightarrow a = b \text{ ou } \lambda \in \{0, 1\}$$



On appelle fonction coût, noté c , la fonction modélisant le coût total des déplacements de matière entre des sites d'extractions et d'utilisation de matières. On a alors $c : P \in B_n \mapsto \sum_{i=1}^n \sum_{j=1}^n p_{i,j} c_{i,j}$ où $p_{i,j}$ est la quantité de matière déplacée du site d'extraction i au site d'utilisation j et $c_{i,j}$ est le coût de transport du site de production i au site d'utilisation j

On appelle polyèdre l'intersection d'un nombre fini de demi-espaces fermés. Un polyèdre d'un espace vectoriel E est donc défini par l'ensemble des solutions du système d'inégalités :

$$f_i(x) \leq \alpha_i, i \in I$$

Où I est un ensemble d'indices de cardinal fini, $\forall i \in I, \alpha_i$ est une constante réelle et f_i est une forme linéaire de E .

On note C_j, L_i et $M_{i,j}$ les applications définies par :

$$\forall j \in \mathbf{N} \cap [1, n], C_j : A \in B_n \mapsto \sum_{i=1}^n a_{i,j}$$

$$\forall i \in \mathbf{N} \cap [1, n], L_i : A \in B_n \mapsto \sum_{j=1}^n a_{i,j}$$

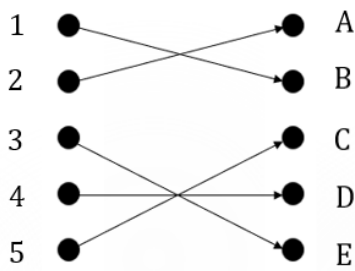
$$\forall (i, j) \in \left(\mathbf{N} \cap [1, n] \right)^2, M_{i,j} : A \in B_n \mapsto a_{i,j}$$

2 - Problèmes de Monge et de Kantorovitch

Le problème de Monge naît suite à la publication de l'ouvrage *Mémoire sur la théorie des déblais et des remblais*. Dans l'ouvrage, Monge associe, 1 à 1, des lieux de production de terre, à des lieux d'utilisation de terre en essayant de minimiser le coût total des déplacements des tas de terre entre les sites. Le problème reposant sur des permutations, le calcul de l'ensemble des coûts de déplacement devient impossible lorsque le nombre de sites devient très important. Monge a alors cherché une méthode mathématique pour trouver, ou même prouver l'existence, du minimum du coût total, chose qu'il n'arrivera pas à faire.

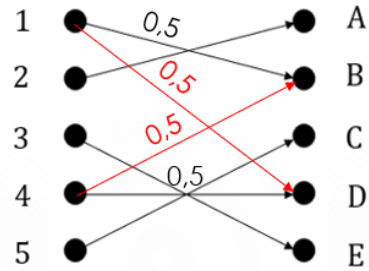
Kantorovitch a alors l'idée 150 ans plus tard d'au lieu d'envoyer chaque tas de terre à un unique site d'utilisation, de diviser le tas de terre en parties pour pouvoir envoyer chaque partie à un site d'utilisation. Ainsi, la nouvelle formulation du problème, dite de Kantorovitch, est modélisable par des matrices bistochastiques contrairement au problème de Monge qui est modélisable par des matrices de permutation. Soit pour illustrer :

Formulation de Monge



	A	B	C	D	E
1	0	1	0	0	0
2	1	0	0	0	0
3	0	0	0	0	1
4	0	0	0	1	0
5	0	0	1	0	0

Formulation de Kantorovitch



	A	B	C	D	E
1	0	0,5	0	0,5	0
2	1	0	0	0	0
3	0	0	0	0	1
4	0	0,5	0	0,5	0
5	0	0	1	0	0

Ainsi avec le problème de Monge, le site de production 1 envoie entièrement le tas de terre sur le site d'utilisation B tandis qu'avec la formulation de Kantorovitch, le site de production 1 envoie la moitié du tas de terre sur le site d'utilisation B et l'autre moitié sur le site d'utilisation D. Ainsi, la formulation de Kantorovitch et le problème de Monge sont similaires, seul leurs ensembles de solutions changent. Ce changement d'ensemble de solution permet alors de prouver l'existence d'une solution au problème de Monge. La suite de ce document consiste donc à la preuve de l'existence d'une telle solution.

3 - Preuve de l'existence d'une solution du problème de Monge

Lemme 1 : Les coefficients des matrices bistochastiques appartiennent à l'intervalle $[0,1]$.

Démonstration : Procédons par l'absurde

Soit $(A, i, j) \in B_n \times (\mathbf{N} \cap [1, n])^2$

Supposons qu'il existe au moins un coefficient qui n'appartient pas à l'intervalle $[0,1]$. Sans perte de généralité on peut supposer qu'il s'agit du coefficient $a_{i,j}$

Alors puisque $\forall k \in \mathbf{N} \cap [1, n] \quad a_{i,k} \geq 0, \quad a_{i,j} \in]1, +\infty[$ et :

$$\sum_{k=1}^n a_{i,k} = \sum_{\substack{k=1 \\ k \neq j}}^n a_{i,k} + a_{i,j} \geq a_{i,j} > 1$$

Ainsi :

$$\sum_{k=1}^n a_{i,k} \neq 1$$

On obtient donc $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2, \quad a_{i,j} \in [0,1]$

Théorème 1 : B_n est un compact convexe de $M_n(\mathbf{R})$.

Démonstration :

Compacité :

On va montrer que l'ensemble des matrices bistochastiques est un fermé borné en dimension finie.

Soit $A \in B_n$

Soit $(A_k) \in B_n^{\mathbf{N}}$ une suite convergeant vers la matrice A

Alors, $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2$:

$$\lim_{k \rightarrow +\infty} a_{k_{i,j}} = a_{i,j}$$

De plus, $\forall k \in \mathbf{N}$, A_k étant bistochastique, par passage à la limite dans une somme finie :

$$\forall i \in \mathbf{N} \cap [1, n] \sum_{j=1}^n a_{i,j} = \sum_{j=1}^n \lim_{k \rightarrow +\infty} a_{k_{i,j}} = \lim_{k \rightarrow +\infty} \sum_{j=1}^n a_{k_{i,j}} = 1$$

$$\forall j \in \mathbf{N} \cap [1, n] \sum_{i=1}^n a_{i,j} = \sum_{i=1}^n \lim_{k \rightarrow +\infty} a_{k_{i,j}} = \lim_{k \rightarrow +\infty} \sum_{i=1}^n a_{k_{i,j}} = 1$$

De plus, $\forall (i, j, k) \in (\mathbf{N} \cap [1, n])^2 \times \mathbf{N}$, $a_{k_{i,j}} \geq 0$ donc $a_{i,j} \geq 0$. Ainsi A est bistochastique, l'ensemble des matrices bistochastiques est un fermé.

Soit $A \in B_n$

D'après le lemme 1 $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2$, $a_{i,j} \in [0, 1]$

En particulier $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2$, $a_{i,j}^2 \leq a_{i,j}$. Ainsi en sommant :

$$\sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2 \leq \sum_{i=1}^n \sum_{j=1}^n a_{i,j}$$

De plus :

$$\sum_{i=1}^n \sum_{j=1}^n a_{i,j} = \sum_{i=1}^n 1 = n$$

Ainsi :

$$N(A)^2 = (A|A) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j}^2 \leq n$$

Ainsi $N(A) \leq \sqrt{n}$, l'ensemble des matrices bistochastiques est donc borné par \sqrt{n} .

L'ensemble des matrices bistochastiques est donc un fermé borné en dimension finie, c'est un compact.

Convexité :

Soit $(D, A, B, \lambda) \in M_n(\mathbf{R}) \times B_n^2 \times [0,1]$, telle que $D = \lambda A + (1 - \lambda)B$:

$$\forall i \in \mathbf{N} \cap [1, n] \quad \sum_{j=1}^n d_{i,j} = \sum_{j=1}^n \lambda a_{i,j} + (1 - \lambda) b_{i,j} = \lambda \sum_{j=1}^n a_{i,j} + (1 - \lambda) \sum_{j=1}^n b_{i,j} = \lambda + (1 - \lambda) = 1$$

$$\forall j \in \mathbf{N} \cap [1, n] \quad \sum_{i=1}^n d_{i,j} = \sum_{i=1}^n \lambda a_{i,j} + (1 - \lambda) b_{i,j} = \lambda \sum_{i=1}^n a_{i,j} + (1 - \lambda) \sum_{i=1}^n b_{i,j} = \lambda + (1 - \lambda) = 1$$

De plus, $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2$, $d_{i,j} = \lambda a_{i,j} + (1 - \lambda) b_{i,j} \geq 0$ car $d_{i,j}$ est une somme de réels positifs. Ainsi D est une matrice bistochastique, l'ensemble des matrices bistochastiques est donc bien convexe.

Ainsi l'ensemble des matrices bistochastiques est un compact convexe.

Remarque : La suite de cette section est la partie la plus difficile du document. La suite du document traite du théorème de Choquet. La démonstration choisie est la démonstration en dimension quelconque. En dimension finie, une démonstration plus simple existe. Ainsi si vous souhaitez uniquement voir la démonstration en dimension finie, regardez uniquement le lemme 2, le théorème de Birkhoff et la remarque page 19 du document.

Théorème de Krein-Milman : Soit K un convexe compact non vide d'un espace vectoriel normé. Alors K est l'enveloppe convexe de ses points extrémaux.

Démonstration : Maxime BOUCHEREAU a déjà réalisé une très bonne démonstration du théorème. Voici le lien de la démonstration : [kreinmilman.pdf](#)

Théorème de Choquet : Soit K un convexe compact non vide d'un espace vectoriel normé, $f : K \rightarrow \mathbf{R}$ une application linéaire continue. Alors il existe un point extrémal de K qui minimise f .

Démonstration :

Soit K un compact convexe non vide d'un espace vectoriel normé et $f : K \rightarrow \mathbf{R}$ une fonction linéaire continue.

K est un compact et f est une application continue, ainsi d'après le théorème des bornes atteintes, f est borné. En particulier, f admet un minimum x_0 sur K . Il reste donc juste à montrer qu'un point extrémal atteint ce minimum.

Procédons par l'absurde : Supposons qu'aucun point extrémal atteint le minimum x_0 .

Puisque K est un compact convexe, d'après le théorème de Krein-Milman, K est l'enveloppe convexe de ses points extrémaux. En particulier, tout point de K est combinaison convexe de points extrémaux K , c'est-à-dire :

$$\forall x \in K, \exists \rho_x \in \mathbf{P}(K), \quad x = \int_{\mathcal{E}(K)} y \rho_x(y), \rho_x \text{ étant une mesure de probabilité}$$

En particulier, il existe une mesure de probabilité $\rho_{x_0} \in \mathbf{P}(K)$ tel que

$$x_0 = \int_{\mathcal{E}(K)} y \rho_{x_0}(y),$$

Or x_0 est un minimum sur K et par hypothèse, aucun point extrémal n'atteint ce minimum ainsi on a :

$$\int_{\mathcal{E}(K)} y \rho_{x_0}(y) < \int_{\mathcal{E}(K)} x_0 \rho_{x_0}(y) < x_0$$

On obtient $x_0 < x_0$ ce qui est absurde

Ainsi un point extrémal atteint bien le minimum.

Lemme 2 : La fonction coût c est une application linéaire et continue.

Démonstration :

Linéarité : Soit $(A, B, \lambda, \mu) \in B_n^2 \times \mathbf{R}^2$ on a alors par linéarité de la somme :

$$\begin{aligned}
c(\lambda A + \mu B) &= \sum_{i=1}^n \sum_{j=1}^n (\lambda A_{i,j} + \mu B_{i,j}) c_{i,j} \\
&= \sum_{i=1}^n \sum_{j=1}^n \lambda A_{i,j} c_{i,j} + \sum_{i=1}^n \sum_{j=1}^n \mu B_{i,j} c_{i,j} \\
&= \lambda \sum_{i=1}^n \sum_{j=1}^n A_{i,j} c_{i,j} + \mu \sum_{i=1}^n \sum_{j=1}^n B_{i,j} c_{i,j} \\
&= \lambda c(A) + \mu c(B)
\end{aligned}$$

Continuité : La fonction coût c est une application linéaire en dimension, par théorème, c est donc continue.

Corollaire 1 : Il existe une solution à la formulation de Kantorovitch.

Démonstration :

D'après le Théorème 1, B_n est un compact convexe. De plus, $M_n(\mathbf{R})$ est un espace vectoriel normé et d'après le lemme 2 la fonction coût c est une application linéaire continue sur B_n . Ainsi d'après le théorème de Choquet, il existe un point extrémal de B_n minimisant c , la formulation de Kantorovitch admet donc une solution.

Théorème de Birkhoff : L'ensemble des points extrémaux de B_n est l'ensemble des matrices de permutation.

Démonstration : La démonstration consiste à montrer que toutes les matrices bistochastiques qui sont des matrices de permutation sont des points extrémaux de B_n et que à l'inverse, toutes les autres matrices bistochastiques ne sont pas des points extrémaux.

Commençons par montrer que toutes les matrices de permutation sont des points extrémaux de B_n :

Soit $P \in P_n$

Soit $(A, B, \lambda) \in (B_n)^2 \times [0,1]$, tel que $P = \lambda A + (1 - \lambda)B$

P est une matrice de permutation ainsi :

$$\forall (i, j) \in \left(\mathbf{N} \cap [1, n] \right)^2, p_{i,j} \in \{0,1\}$$

Donc , soit $(i, j) \in (\mathbf{N} \cap [1, n])^2$, on a :

$$1 = \lambda a_{i,j} + (1 - \lambda)b_{i,j} \text{ ou } 0 = \lambda a_{i,j} + (1 - \lambda)b_{i,j}$$

Procédons par disjonction de cas :

-Dans le cas où $1 = \lambda a_{i,j} + (1 - \lambda)b_{i,j}$, soit $a_{i,j} = b_{i,j} = 1$, soit $\lambda \in \{0,1\}$

En effet, si $\lambda \in]0,1[$, puisque $a_{i,j}$ et $b_{i,j}$ appartiennent à l'intervalle $[0,1]$, on aurait si $a_{i,j} \neq 1$ et $b_{i,j} \neq 1$: $1 = \lambda a_{i,j} + (1 - \lambda)b_{i,j} < \lambda + (1 - \lambda) = 1$ ce qui est absurde.

-Dans le cas où $0 = \lambda a_{i,j} + (1 - \lambda)b_{i,j}$, soit $a_{i,j} = b_{i,j} = 0$, soit $\lambda \in \{0,1\}$

En effet si $\lambda \in]0,1[$, puisque $a_{i,j}$ et $b_{i,j}$ appartiennent à l'intervalle $[0,1]$, $\lambda a_{i,j} + (1 - \lambda)b_{i,j}$ est la somme de deux termes positifs. On a donc nécessairement $a_{i,j} = b_{i,j} = 0$.

On obtient bien $a_{i,j} = b_{i,j}$ ou $\lambda \in \{0,1\}$.

Ainsi dans tous les cas, $\forall (i,j) \in (\mathbf{N} \cap [1,n])^2$, $a_{i,j} = b_{i,j}$ ou $\lambda \in \{0,1\}$ donc $\forall (P,A,B,\lambda) \in P_n \times B_n^2 \times [0,1]$ tel que $P = \lambda A + (1 - \lambda)B$, $A = B$ ou $\lambda \in \{0,1\}$. On obtient donc que toutes les matrices de permutations sont des points extrémaux de l'ensemble des matrices bistochastiques.

Montrons maintenant que les matrices bistochastiques qui ne sont pas des matrices de permutation ne sont pas des points extrémaux de l'ensemble des matrices bistochastiques. Autrement dit, montrons que les matrices de bistochastiques dont au moins l'un coefficient appartient à l'intervalle $]0,1[$ ne sont pas des points extrémaux de l'ensemble des matrices bistochastiques.

La démonstration consiste à perturber une matrice bistochastique quelconque afin deux matrices bistochastiques différentes de tel sorte que la première matrice soit combinaison linéaire des deux autres.

Pour bien comprendre la démonstration, on peut réaliser un exemple avec les matrices bistochastiques de taille 3.

Pour commencer, toute matrice bistochastique peut être construit en suivant une certaine méthode :

Soit $A \in B_3$

Par hypothèse, A possède au moins un coefficient non nul. Dans la suite de la démonstration, chaque coefficient sera choisi sans perte de généralité. Nous choisissons donc sans perte de généralité, que le coefficient non nul est le coefficient $a_{1,1} \in]0,1[$.

On a donc $A =$

$$\begin{pmatrix} a_{1,1} & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{pmatrix}$$

De plus $a_{1,1} \in]0,1[$ et la somme des coefficients de la première ligne vaut 1. Il existe donc au maximum deux coefficients $a_{1,2}$ et $a_{1,3}$ appartenant respectivement aux intervalles $]0,1[$ et $[0,1[$ tel que $a_{1,1} + a_{1,2} + a_{1,3} = 1$.

De la même manière, la somme des coefficients de la première colonne vaut 1. Ainsi Il existe donc au maximum deux coefficients $a_{2,1}$ et $a_{3,1}$ appartenant respectivement aux intervalles $]0,1[$ et $[0,1[$.

On obtient donc $A =$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & ? & ? \\ a_{3,1} & ? & ? \end{pmatrix}$$

De la même manière les sommes des coefficients de la deuxième ligne /troisième ligne et de la deuxième colonne /troisième colonne valent 1. Il existe donc au maximum quatre autres coefficients $a_{2,2}$, $a_{2,3}$, $a_{3,2}$ et $a_{3,3}$ appartenant respectivement aux intervalles $]0,1[$, $[0,1[$, $[0,1[$ et $[0,1[$ tel que

$$a_{1,2} + a_{2,2} + a_{3,2} = 1, a_{2,1} + a_{2,2} + a_{2,3} = 1, a_{1,3} + a_{2,3} + a_{3,3} \text{ et } a_{3,1} + a_{3,2} + a_{3,3} = 1.$$

La matrice devient ainsi $A =$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

A possédant un nombre fini de coefficients, on peut créer deux suites finies, v_3 et w_3 , de coefficients appartenant à l'intervalle $]0,1[$ de tel manière que pour tout coefficient de la suite v_3 , il existe deux coefficients de la suite w_3 , l'un se trouvant sur la même ligne que le coefficient de v_3 , l'autre se trouvant sur la même colonne que le coefficient de v_3 . En particulier si la matrice bistochastique est d'ordre n pair, les vecteurs contiennent n éléments, sinon, ils contiennent $n - 1$ éléments.

Par exemple : Pour la matrice

$$\begin{pmatrix} 0,3 & 0,5 & 0,2 \\ 0,2 & 0,3 & 0,5 \\ 0,5 & 0,2 & 0,3 \end{pmatrix}$$

On peut construire les suite $v_3 = (0,3 ; 0,3)$

Et $w_3 = (0,2 ; 0,2)$

Pour une matrice avec des coefficients valant 0 ou 1 :

$$\begin{pmatrix} 0,3 & 0,7 & 0 \\ 0 & 0 & 1 \\ 0,7 & 0,3 & 0 \end{pmatrix}$$

On peut construire les suites $v_3 = (0,3 ; 0,3)$

Et $w_3 = (0,7 ; 0,7)$

Enfin on pose ε la moitié du minimum des distances des coefficients de A avec 0 ou 1 de) de tel sorte que :

$$\forall (i, j) \in (\mathbf{N} \cap [1, 3])^2, a_{i,j} \in]0, 1[\rightarrow a_{i,j} \pm \varepsilon \in]0, 1[$$

On construit la matrice A^+ en augmentant ε tous les coefficients de A présent dans la suite v_3 et en diminuant de ε tous les coefficients de A présent dans la suite w_3 .

De même, on construit la matrice A^- en diminuant de ε tous les coefficients de A présent dans la suite v_3 et en augmentant ε tous les coefficients de A présent dans la suite w_3 .

Ainsi pour la matrice bistochastique précédente on obtient :

$$A^+ = \begin{pmatrix} 0,3+\varepsilon & 0,7-\varepsilon & 0 \\ 0 & 0 & 1 \\ 0,7-\varepsilon & 0,3+\varepsilon & 0 \end{pmatrix} \quad A^- = \begin{pmatrix} 0,3-\varepsilon & 0,7+\varepsilon & 0 \\ 0 & 0 & 1 \\ 0,7+\varepsilon & 0,3-\varepsilon & 0 \end{pmatrix}$$

Les $\pm\varepsilon$ s'éliminant deux à deux, par construction les matrices A^+ et A^- sont également bistochastiques. De plus, $A = \frac{1}{2} A^+ + \frac{1}{2} A^- = \lambda A^+ + (1 - \lambda) A^-$ avec $\lambda = \frac{1}{2}$. Ainsi A n'est pas un point extrême de B_3 .

Bien maintenant que la méthode de la démonstration est bien comprise, faisons là avec des matrices de taille n .

Soit $A \in B_n$

A étant bistochastique, la matrice admet un coefficient appartenant à l'intervalle $]0,1[$. De la même manière que dans la démonstration avec les matrices de taille 3, on peut construire deux suites finies, v_n et w_n , de coefficients appartenant à l'intervalle $]0,1[$ de telle manière que pour tout coefficient de la suite v_n , il existe deux coefficients de la suite w_n , l'un se trouvant sur la même ligne que le coefficient de v_n , l'autre se trouvant sur la même colonne que le coefficient de v_n .

On pose ε la moitié du minimum des distances des coefficients de A avec 0 ou 1 de telle sorte que :

$$\forall (i, j) \in (\mathbb{N} \cap [1, n])^2, a_{i,j} \in]0,1[\rightarrow a_{i,j} \pm \varepsilon \in]0,1[$$

On construit alors la matrice A^+ en augmentant ε tous les coefficients de A présent dans la suite v_n et en diminuant de ε tous les coefficients de A présent dans la suite w_n .

De même, on construit la matrice A^- en diminuant de ε tous les coefficients de A présent dans la suite v_n et en augmentant ε tous les coefficients de A présent dans la suite w_n .

Les matrices obtenues sont également bistochastiques. En effet, puisque $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2, a_{i,j} \in]0, 1[\rightarrow a_{i,j} \pm \varepsilon \in]0, 1[$ et $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2, a_{i,j} \geq 0$, on a $\forall (i, j) \in (\mathbf{N} \cap [1, n])^2, a^+_{i,j} \geq 0$ et $a^-_{i,j} \geq 0$. De plus :

$$\forall i \in \mathbf{N} \cap [1, n] \sum_{j=1}^n a^+_{i,j} = \sum_{j=1}^n a_{i,j} + \sum_{j=1}^{card(v_n)} +\varepsilon - \varepsilon = 1 + 0 = 1$$

$$\forall i \in \mathbf{N} \cap [1, n] \sum_{j=1}^n a^-_{i,j} = \sum_{j=1}^n a_{i,j} + \sum_{j=1}^{card(w_n)} +\varepsilon - \varepsilon = 1 + 0 = 1$$

$$\forall j \in \mathbf{N} \cap [1, n] \sum_{i=1}^n a^+_{i,j} = \sum_{i=1}^n a_{i,j} + \sum_{i=1}^{card(v_n)} +\varepsilon - \varepsilon = 1 + 0 = 1$$

$$\forall j \in \mathbf{N} \cap [1, n] \sum_{i=1}^n a^-_{i,j} = \sum_{i=1}^n a_{i,j} + \sum_{i=1}^{card(w_n)} +\varepsilon - \varepsilon = 1 + 0 = 1$$

Enfin $A = \frac{1}{2} A^+ + \frac{1}{2} A^- = \lambda A^+ + (1 - \lambda) A^-$ avec $\lambda = \frac{1}{2}$. Ainsi $\forall A \in B_n, \exists (A^+, A^-, \lambda) \in B_n^2 \times [0, 1]$ tel que $A^+ \neq A^-$ et $\lambda \neq 0$ et $\lambda \neq 1$. On

obtient donc que seule les matrices de permutation sont des points extrémaux de B_n .

Ainsi l'ensemble des points extrémaux de B_n est l'ensemble des matrices de permutation.

Corollaire 2 : Il existe une solution au problème de Monge.

Démonstration :

D'après le corollaire 1, il existe un point extrémal de B_n minimisant c . Or d'après le théorème de Birkhoff, l'ensemble des points extrémaux de B_n est l'ensemble des matrices de permutation. Par conséquent, il existe une matrice de permutation de minimisant c , le problème de Monge admet donc une solution.

Remarque : En dimension finie, l'utilisation du théorème de Krein-Milman n'est pas nécessaire pour prouver l'existence d'une solution au problème de Monge/ à la formulation de Kantorovitch. En effet, dans le document d'agrégation *Développement : Théorèmes de Choquet et de Birkhoff* ([Théorèmes de Choquet et de Birkhoff.pdf](#)) l'auteur réalise une démonstration du théorème de Choquet sans le théorème de Krein-Milman. Toutefois, l'auteur rajoute l'hypothèse que l'ensemble des points extrémaux du compact convexe est également un compact. Nous avons vu précédemment que l'ensemble des points extrémaux de B_n est l'ensemble des matrices de permutation. Or en dimension finie, l'ensemble des matrices de permutation est un ensemble fini, c'est donc en particulier un compact.

4 – Méthode pour trouver une solution au problème de Monge

Le problème de Monge reposant sur des permutations, chercher une matrice solution du problème pourrait nécessiter de l'ordre de $n!$ calculs. Il nous faut donc une méthode plus rapide pour trouver une telle solution :

Lemme 3 : Les applications C_j , L_i et $M_{i,j}$ sont des formes linéaires sur B_n .

Démonstration :

Par combinaison linéaire de coefficients, C_j et L_i sont des applications linéaires. De plus, par somme de coefficient appartenant à l'intervalle $[0,1]$, les applications C_j et L_i sont à valeur dans \mathbf{R} . Les applications C_j et L_i sont donc bien des formes linéaires.

$M_{i,j}$ correspond à une combinaison linéaire d'un seul coefficient appartenant à l'intervalle $[0,1]$. $M_{i,j}$ donc également une forme linéaire.

Théorème 2 : B_n est un polyèdre convexe dans $M_n(\mathbf{R})$.

Démonstration :

Convexité :

D'après le Théorème 1, B_n est un compact convexe $M_n(\mathbf{R})$. En particulier B_n est convexe.

B_n est un polyèdre :

Soit $A \in M_n(\mathbf{R})$

$A \in B_n$ si et seulement si :

$$\forall (i,j) \in \left(\mathbf{N} \bigcap [1,n] \right)^2, a_{i,j} \geq 0$$

$$\forall i \in \mathbf{N} \bigcap [1,n] \sum_{j=1}^n a_{i,j} = 1$$

$$\forall j \in \mathbf{N} \bigcap [1,n] \sum_{i=1}^n a_{i,j} = 1$$

Soit, si et seulement si :

$$\forall (i,j) \in \left(\mathbf{N} \bigcap [1,n] \right)^2, M_{i,j}(A) \geq 0$$

$$\forall i \in \mathbf{N} \bigcap [1,n], L_i(A) = 1$$

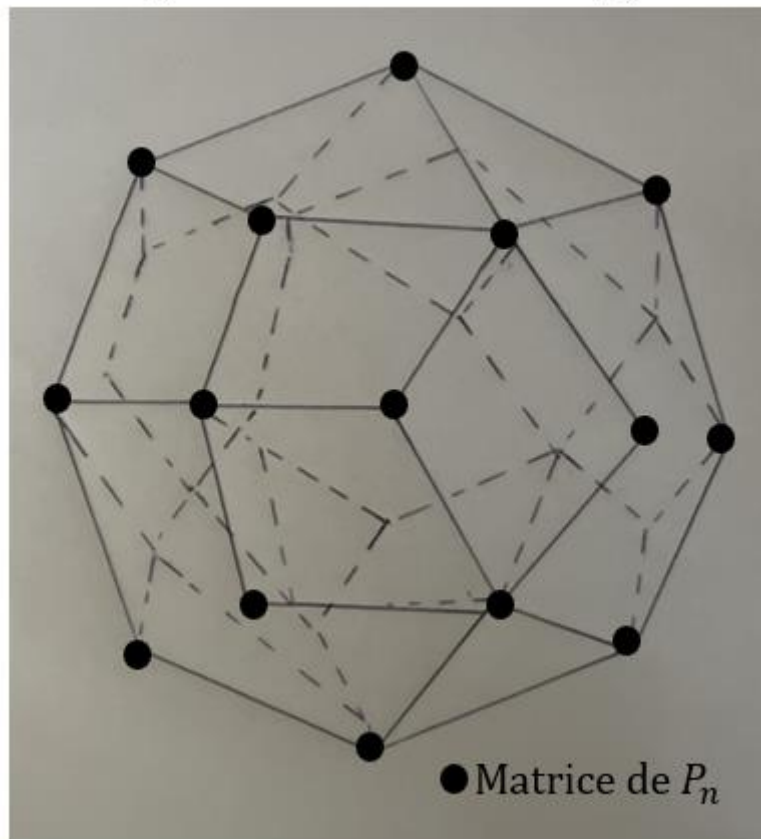
$$\forall j \in \mathbf{N} \bigcap [1,n], C_j(A) = 1$$

On obtient ainsi un système définissant un polyèdre avec comme formes linéaires de $M_n(\mathbf{R})$, $M_{i,j}$, L_i et C_j et comme constantes 0 et 1.

Ainsi B_n est un polyèdre convexe dans $M_n(\mathbf{R})$.

En réalité prouvé que B_n est un polyèdre convexe dans $M_n(\mathbf{R})$ ne change pas le fait qu'il pourrait falloir de l'ordre de $n!$ calculs pour trouver une solution au problème de Monge. Cependant, les matrices de permutations ne sont pas des matrices quelconques de B_n , se sont ses points extrémaux. Les matrices de permutation sont donc les sommets du polyèdre B_n . Par exemple si on représente B_n par un dodécaèdre :

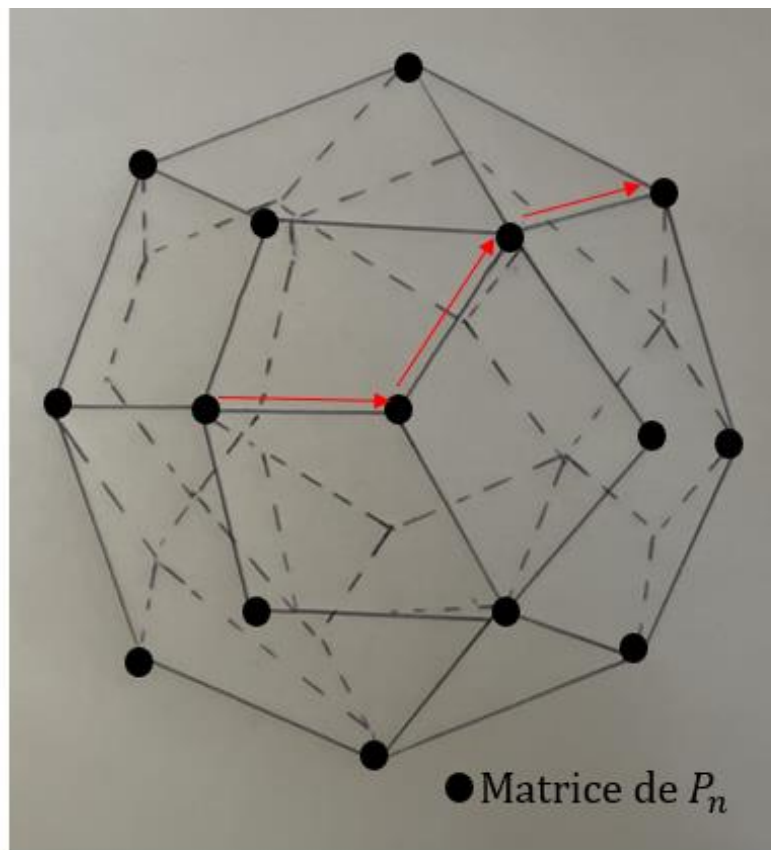
Polyèdre convexe B_n



Les points extrémaux à l'arrière du polyèdre ne sont pas représentés pour éviter de surcharger le dessin.

Il existe alors plusieurs algorithmes d'optimisation linéaire pour obtenir une matrice de permutation solution, l'algorithme du simplexe et l'algorithme des points intérieurs. Ces deux algorithmes fonctionnent uniquement sur des polyèdres convexes.

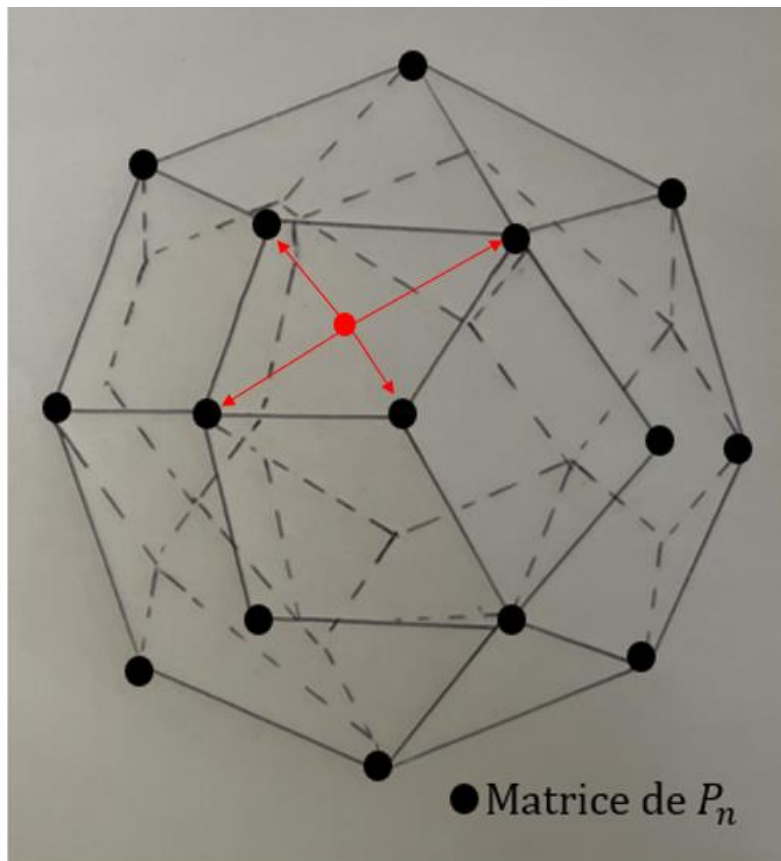
L'algorithme du simplexe consiste à passer d'un sommet du polyèdre à un autre sommet adjacent par une arête jusqu'à trouver un point extrémal solution. Soit pour illustrer :



L'algorithme des points intérieurs se répartissent en plusieurs familles utilisant différentes méthodes d'optimisation :

- méthodes du chemin central
- méthodes de réduction de potentiel
- méthodes dites « affine scaling »

Contrairement au chemin de l'algorithme du simplexe, le chemin de l'algorithme des points intérieurs reste à l'intérieur du polyèdre convexe (d'où son nom). Ainsi pour illustrer :



Bien que l'algorithme des points intérieurs soit théoriquement plus rapide, l'algorithme du simplexe reste plus rapide dans certains cas. Les algorithmes modernes utilisent donc un mixte des deux algorithmes. C'est notamment le cas du package d'optimisation linéaire de python GLPK (GNU Linear Programming Kit) .

Si vous voyez une coquille, merci de me prévenir au mail :
mcolin.maths@gmail.com

Sources :

- Tous les documents cités précédemment
- *Théorie générale du transport et applications* : Cécile CARRERE, Didier LESESVRE, Paul PEGON
- *Méthodes polyédriques et algorithmes* : Alain FAYE
- *Matrices bistochastiques* : Jean Etienne ROMBALD

