Un problème de transport optimal discret

Référence : Un sujet de devoir de prépa de N. Tosel et mon stage de Master 1 avec F. Bolley.

Leçons : 105, 106, 151, 159, 181, 219, 253.

1 Cadre du problème

On considère n tas de sable de même volume disons 1 par commodité et n fossés tous de volume 1 également. Le but étant de déplacer de manière optimale, c'est-à-dire en minimisant un certain coût, l'ensemble des n tas de sable dans les fossés. **A priori on s'autorise à repartir un même tas de sable dans plusieurs fossés.** On peut aussi imaginer le transport de marchandises dans des entrepôts sachant que chaque stock contient la même quantité de marchandises et chaque dépôt la même capacité de réception. Le but étant de minimiser le coût de transport qu'on peut donner sous la forme d'un coût par kilogramme de marchandises transportées d'un stock donné à un entrepôt donné. Modélisons mathématiquement le problème.

Modélisation d'un plan de déplacement (stratégie) :

Une stratégie de déplacement est donnée par une matrice de $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ où $m_{i,j}$ représente la quantité de sable déplacée du tas i au fossé j. Il est naturel de supposer qu'en plus, on a :

- $m_{i,j} \geq 0$
- $\forall i,\ \sum\limits_{j=1}^{n} m_{i,j} = 1$: il ne reste pas de sable dans chaque tas, tout le sable a été déplacé lorsque de l'opération puisque le volume de sable à déplacer a le même volume que l'ensemble des fossés.
- $\forall j,\ \sum\limits_{i=1}^{n} m_{i,j} = 1$ pour la même raison, chaque fossé est rempli entièrement.

On modélise donc une stratégie de déplacement par une matrice bistochastique (i.e. une matrice vérifiant les 3 conditions précédentes) : on note $\mathcal{B}_n$ leur ensemble.

Modélisation du coût de transport :

On introduit la matrice $C \in \mathcal{M}_n(\mathbb{R})$ dont le coefficient $c_{i,j}$ représente le coup volumique de transport du tas i au trou j. Ainsi le coût total de transport, qui dépend de C et de la stratégie M est :

$$C(M) = \sum_{i,j \leq n} c_{i,j} m_{i,j} = \mathrm{Tr}(C^T M) := C.M$$

Objectif :

Existe-t-il une stratégie qui réalise (et si oui laquelle) : $\inf_{M \in \mathcal{B}_n} C.M \quad ?$.

2 Le théorème de Krein-Milman

On commence par définir la notion de point extrémal d'un ensemble convexe.

Définition 2.1

Si K est un convexe, on dit que $M \in K$ est extrémal si $M = \frac{C_1+C_2}{2}$, avec C_1 et C_2 dans $K \Rightarrow C_1 = C_2 = M$.

On note $Extr(K)$ leur ensemble.

Passons maintenant au fameux théorème de Krein-Milman, énoncé ici en dimension finie seulement car la preuve est bien plus simple.

Théorème 2.2

Soit K un convexe compact non vide d'un $\mathbb{R}$-espace vectoriel de dimension finie. Alors K est l'enveloppe convexe de ses points extrémaux.

On commence par démontrer l'existence d'un hyperplan d'appui.

Théorème 2.3

Soit $K \subset \mathbb{R}^n$ un compact convexe et soit $c \in \partial K$. Alors il existe $\phi \in (\mathbb{R}^n)^* \backslash \{0\}$ telle que :

$$\forall x \in K, \, \phi(x) \leq \phi(c).$$

C'est-à-dire que l'hyperplan affine $c + \ker(\phi)$ sépare c et K au sens large.

Démonstration : Quitte à se placer dans l'espace affine engendré par K, on se ramène au cas où $\mathring{K} \neq \emptyset$. Il suffit alors d'appliquer le théorème de Hahn-Banach géométrique (cas ouvert) aux convexes $\mathring{K}$ et $\{c\}$. En fait si K est d'intérieur vide, il est contenu dans un hyperplan. En effet, l'espace affine engendré par K ne peut pas être de dimension n sinon K contiendrait un simplexe engendré par une famille génératrice et ne serait donc pas d'intérieur vide.

Si on n'a pas envie d'utiliser le théorème de Hahn-Banach, on peut ruser en utilisant le théorème de projection. En effet, puisque $c \in \partial K$, on dispose d'une suite $(x_n)_n$ d'éléments qui ne sont pas dans K et qui converge vers c (puisque $\partial K \subset (\mathring{K})^C = \overline{K^C}$). On note y_n la projection de x_n sur K. Par propriété des angles obtus, on a :

$$\forall n \in \mathbb{N}, \forall z \in K, \, \langle x_n - y_n, z - y_n \rangle \leq 0. \quad (*)$$

Posons alors pour tout n :

$$u_n = \frac{x_n - y_n}{\|x_n - y_n\|}.$$

Par compacité de la sphère unité en dimension finie, il existe $u \in \mathbb{R}^n$ de norme 1 telle qu'une sous-suite de $(u_n)_n$ converge vers u. On vérifie que $H := c + u^\perp$ est un hyperplan d'appui. L'inégalité $(*)$ assure que :

$$\forall z \in K, \, \langle u_n, z - y_n \rangle \leq 0.$$

Par continuité de l'application projection, on a $y_n \to p(c) = c$. On déduit que :

$$\forall z \in K,\ \langle u, z - c\rangle \leq 0.$$

La forme linéaire $\langle u, .\rangle$ convient. $\square$

Énonçons et démontrons un lemme avant de passer à la preuve du théorème de Krein-Milman.

Lemme 2.4

Soit $H = \phi^{-1}(\lambda)$ un hyperplan d'appui à K en $c \in \partial K$. $H \cap K$ est convexe compact et un point $d \in H \cap K$ est extrémal dans K si et seulement si il l'est dans $H \cap K$.

Démonstration : On raisonne par contraposée. Il est clair que si d n'est pas extrémal dans $H \cap K$, alors a fortiori il ne l'est pas dans K. Supposons que $d = \frac{x+y}{2}$ avec $x, y \in K$ et $x \neq y$, n'est pas extrémal dans K. On sait déjà que :

$$\phi(x) \leq \lambda \quad \text{et} \quad \phi(y) \leq \lambda.$$

On déduit par linéarité que :

$$\phi(d) = \phi\left(\frac{x+y}{2}\right) \leq \lambda = \phi(d) \quad (\text{car}\, d \in H).$$

D'où $\phi(x) = \phi(y) = \lambda$ donc $x, y \in H \cap K$. Cela montre exactement que d n'est pas extrémal dans $H \cap K$. $\square$

On peut enfin passer à la preuve théorème de Krein-Milman !

Démonstration : On suppose K non vide sinon il n'y a pas beaucoup d'intérêt. On raisonne par récurrence sur d la dimension du sous-espace affine engendré par K (c'est-à-dire si $x_0 \in K$: $\dim(\mathrm{Vect}(K - x_0))$).

- Si $d = 0$, alors $K = \{x_0\}$ et le résultat est trivial.
- Si le résultat est vrai pour les dimensions inférieures à $d-1$ et K " de dimension " d. Soit $x \in \partial K$, montrons que x s'écrit comme combinaison convexe de points extrémaux. Soit $H = \phi^{-1}(\lambda)$ un hyperplan d'appui à K en x. Si x est extrémal dans K il n'y a rien à montrer. Sinon x n'est pas extrémal dans $H \cap K$ d'après le lemme précédent. On peut appliquer l'hypothèse de récurrence à $H \cap K$ qui est bien de dimension inférieure à $d-1$. Ainsi x s'écrit comme combinaison convexe de point extrémaux dans $H \cap K$ qui sont des point extrémaux de K. Cela montre que :

  $$\partial K \subset \mathrm{Conv}(\mathrm{Extr}(K)).$$

  Si maintenant $x \in \mathring{K}$, alors une droite issue de x rencontre la frontière de K en deux points. Ainsi x est une combinaison convexe de deux points de la frontière qui sont eux même combinaison de points extrémaux. D'où :

  $$K \subset \mathrm{Conv}(\mathrm{Extr}(K)),$$

  ce qui achève la preuve puisque l'autre inclusion est immédiate.

$\square$

3 Propriétés de l'ensemble des matrices bistochastiques

3.1 Compacité

Théorème 3.1

$\mathcal{B}_n$ est un ensemble compact de $\mathcal{M}_n(\mathbb{R})$.

Démonstration : En effet, si on définit les applications suivantes continues :

$$p_{.j} : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to \mathbb{R} \\ M = (m_{i,j}) & \mapsto \sum_{i=1}^{n} m_{i,j} \end{cases} .$$

$$p_{i.} : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \to \mathbb{R} \\ M = (m_{i,j}) & \mapsto \sum_{j=1}^{n} m_{i,j} \end{cases} .$$

Alors : $\mathcal{B}_n = \bigcap_{i=1}^{n} p_{i.}^{-1}(\{1\}) \cap \bigcap_{j=1}^{n} p_{.j}^{-1}(\{1\}) \cap \mathcal{M}_n(\mathbb{R}^+)$

Ainsi $\mathcal{B}_n$ est fermé comme intersection de fermés.

De plus, $\mathcal{B}_n$ est borné, par exemple pour la norme du maximum des coefficients par 1. Donc $\mathcal{B}_n$ compact.

$\square$

3.2 Convexité et points extrémaux

Théorème 3.2

$\mathcal{B}_n$ est un ensemble convexe.

Démonstration : En effet, si $\epsilon \in [0,1]$ et $B, B' \in \mathcal{B}_n$, alors $\epsilon B + (1-\epsilon)B' \in \mathcal{B}_n$.

$\square$

En fait on connait exactement les points extrémaux du convexe compact $\mathcal{B}_n$.

Théorème 3.3

Les points extrémaux de $\mathcal{B}_n$ sont les matrices de permutations données par :

$$\forall \sigma \in S_n,\ P_\sigma = (\mathbf{1}_{i=\sigma(j)}).$$

Démonstration : Commençons par montrer que les matrices de permutations sont bien des points extrémaux de $\mathcal{B}_n$. Si $P_\sigma = \frac{C+C'}{2}$, avec $C, C' \in \mathcal{B}_n$. Alors :

$$\forall j \in \{1, \ldots, n\}, \ \frac{c_{\sigma(j),j} + c'_{\sigma(j),j}}{2} = 1$$

$$\forall k \neq j,\ \frac{c_{\sigma(j),k} + c'_{\sigma(j),k}}{2} = 0$$

Donc $C = C' = P_\sigma$. Étudions maintenant la réciproque.

On note J_n l'espace vectoriel des matrices dont la somme de chaque ligne et de chaque colonne est nulle. On exhibe un isomorphisme entre J_n et $\mathcal{M}_{n-1}(\mathbb{R})$.

$$S : \begin{cases} \mathcal{M}_{n-1}(\mathbb{R}) & \to J_n \\ A & \mapsto \begin{pmatrix} A & U \\ V & w \end{pmatrix} \end{cases}$$

où :

1. $V_j = -\sum_{k=1}^{n-1} A_{k,j}$
2. $U_i = -\sum_{l=1}^{n-1} A_{i,l}$
3. $w = -\sum_{j=1}^{n-1} V_j$

Il s'agit bien d'un isomorphisme puisque :

- Injectivité : Si $A \neq A'$, alors $S(A) \neq S(A')$.
- Surjectivité : Si $M \in \mathcal{B}_n$, alors si $A = (m_{i,j})_{i,j \leq n-1}$, alors $M = S(A)$.

Conclusion : $\dim(J_n) = (n-1)^2$.

On note $C_{i,j}$ la forme linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui renvoie le coefficient (i,j).

Définition 3.4

Soit $\alpha, \beta : \{1, \dots, n\}$ deux applications vérifiant : $\forall i,\ \alpha(i) < \beta(i)$.
$C_{\alpha,\beta} = \{M \in \mathcal{M}_n(\mathbb{R}),\ \forall (i,j),\ j \notin \{\alpha(i), \beta(i)\} \Rightarrow C_{i,j}(M) = 0\}$.

Remarquons que :

$$\dim(C_{\alpha,\beta}) \geq 2n.$$

Démontrons le lemme clé suivant.

Lemme 3.5

On considère $B \in \mathcal{B}_n$ telle qu'il existe α, β comme précédemment avec :

$$\forall i,\ C_{i,\alpha(i)}(B), C_{i,\beta(i)}(B) > 0.$$

Alors B n'est pas extrémale dans $\mathcal{B}_n$.

Démonstration : On cherche $C \in C_{\alpha,\beta}$ et $\epsilon > 0$ tels que :

- $B = \frac{B-\epsilon C}{2} + \frac{B+\epsilon C}{2}$
- $C \neq 0$

- $B \pm \epsilon C \in \mathcal{B}_n$

Choisissons maintenant grâce à ce qui précède C et ϵ.

1. On a :
$$\dim(J_n \cap C_{\alpha,\beta}) \geq 2n + (n-1)^2 - \dim(J_n + C_{\alpha,\beta}) > 0.$$
Il existe donc $C \in J_n \cap C_{\alpha,\beta} \backslash \{0\}$.
2. Choisissons ϵ.
   - Fixons i. Si $j \notin \{\alpha(i), \beta(i)\}$, alors $C_{i,j}(B \pm \epsilon C) = B_{i,j} \geq 0$.
   - Sinon on peut choisir $\epsilon > 0$ assez petit tel que pour tout i :
$$B_{i,\alpha(i)} \pm \epsilon C_{i,\alpha(i)} \quad \text{et} \quad B_{i,\beta(i)} \pm \epsilon C_{i,\beta(i)} \geq 0$$

Cela assure que si $B \in \mathcal{B}_n$ est telle qu'il existe α et β comme précédemment avec :
$$\forall i,\, C_{i,\alpha(i)}(B), C_{i,\beta(i)}(B) > 0$$
Alors B n'est pas extrémale dans $\mathcal{B}_n$.

□

Achevons maintenant la preuve du fait qu'une matrice extrémale est de permutation. On raisonne par récurrence sur la dimension n. On note :

P_n : "Une matrice extrémale de $\mathcal{B}_n$ est de permutation".

- **Initialisation :** $\mathcal{B}_1 = \{(1)\}$.
- **Hérédité :** Si P_{n-1} est vraie. Soit $B \in \mathcal{B}_n$. D'après ce qui précède il n'existe pas deux applications α et β comme précédemment. On déduit facilement que M est de la forme suivante :
$$\begin{pmatrix} * & * & 0 & * & * & * \\ * & * & \vdots & * & * & * \\ * & * & 0 & * & * & * \\ 0 & \dots & 1 & 0 & \dots & 0 \\ * & * & 0 & * & * & * \\ * & * & \vdots & * & * & * \\ * & * & 0 & * & * & * \end{pmatrix}.$$
La matrice à laquelle on enlève la ligne i et la colonne j est dans $\mathcal{B}_{n-1}$ et reste extrémale sinon B ne le serait pas. Par hypothèse de récurrence, c'est une matrice de permutation. Donc B est elle même une matrice de permutation.

Cela achève la preuve du théorème.

□

4 Résolution du problème initial

Théorème 4.1

Il existe une matrice de permutation P_σ réalisant :
$$\inf_{M \in \mathcal{B}_n} C.M.$$

Démonstration : On veut minimiser la fonction $M \in \mathcal{B}_n \mapsto C.M$. Cette application est continue car c'est la restriction d'une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui est donc continue (dimension finie).

Puisque $\mathcal{B}_n$ est un ensemble compact, on déduit que la fonction atteint bien son infimum C^* en une matrice $M \in \mathcal{B}_n$.

Par l'absurde : si le minimum n'est pas atteint en une matrice de permutation.
Alors d'après le théorème de Krein-Milman :

$$M = \sum_{i=1}^{s} \lambda_i P_{\sigma_i},$$

où $\sigma_i \in S_n$, $\lambda_i \in [0,1]$ avec $\sum \lambda_i = 1$. On a donc :

$$C.M = \sum_{i=1}^{s} \lambda_i C.P_{\sigma_i} > C^*,$$

ce qui est absurde.

□

Conclusion : On a vu que le problème de transport optimal de tas de sable (ou de n'importe quoi d'autre) admet une solution qui est réalisée par le transport de chaque tas de sable entier vers un fossé. Ainsi indépendamment du coût de transport, il est inutile de découper les tas de sable pour espérer optimiser le coût.

Remarque 4.2. *En fait on a résolu un problème de transport optimal discret pour des probabilités qui chargent un ensemble fini de points avec la même masse, précisément on a montré que le problème de Kantorovitch (" s'autoriser à découper les tas de sable ") est équivalent au problème de Monge (" envoyer chaque tas dans un fossé ").*