La brouette de Monge ou le transport optimal

Yann Brenier, *directeur de recherche CNRS à l'École polytechnique*

Né d'un problème concret – comment déplacer au mieux un tas de sable – le transport optimal est un outil qui trouve des applications aussi bien à l'intérieur des mathématiques (de la géométrie à l'analyse fonctionnelle) que dans d'autres domaines, comme la gestion de ressources, par exemple.

Dès le XVIII^{e} siècle, le mathématicien français Gaspard Monge, dans son *Mémoire sur la théorie des déblais et des remblais*, étudiait un problème des plus concrets (déplacer au mieux un tas de sable !) en lui appliquant une méthode rigoureuse, « optimale » dirions-nous aujourd'hui. Ce problème est l'un des premiers exemples de *recherche opérationnelle*, la branche des mathématiques qui s'intéresse aux méthodes pour traiter de manière efficace des problèmes combinatoires (voir l'encadré *La combinatoire*) et qui est une théorie encore très vivace aujourd'hui.

En quoi ce problème (qu'on pourrait appeler « Brouette de Monge », le but étant de déplacer du sable) est-il un problème combinatoire ? Nous pouvons le lire de la façon suivante. Nous avons un tas de sable, composé de petits éléments (un élément étant un grain ou une poignée). Nous avons aussi un trou à combler, dont on connaît la forme et dont le volume est supposé coïncider avec celui du tas. On peut penser le trou comme composé de petites cases, chacune pouvant contenir un élément de sable. La question est : où envoyer chaque élément pour minimiser les chemins parcourus au total ? On peut essayer de répondre à partir d'une description combinatoire du problème, mais dans certains cas il peut être plus utile d'en donner une description différente, basée sur les formes du profil du tas et du trou, considérées comme des fonctions. Et nous allons voir aussi que ce problème d'optimisation peut tout à fait être utilisé pour parler d'autres choses, et pas seulement de sable à déplacer.

Des applications diverses

La question du transport a été réexaminée dans les années 1940 par Leonid Kantorovich (un des quelques mathématiciens lauréats du prix Nobel d'économie) à propos d'allocation optimale des ressources. Dans son œuvre, il cherche à résoudre des problèmes de ce genre : comment utiliser les rares facteurs de production (travail, capital…) afin de réaliser une satisfaction maximale des besoins ? Par exemple, étant données des mines produisant du métal et des usines le travaillant, on peut essayer de déterminer quelle mine doit approvisionner quelle usine pour minimiser les dépenses dues au transport du matériel : cette question est finalement la même que celle du tas de sable.

Bien d'autres questions (files d'attente, gestion de stock, etc.), que nous ne détaillerons pas ici, font partie de ce domaine. Il s'agit toujours de trouver comment passer de la situation initiale à la situation finale de la meilleure manière possible, c'est-à-dire par le meilleur chemin dans l'espace compliqué de tous les choix possibles.

Déplacer un nuage de points

Le point clé de la théorie est probablement le concept de *géodésique*, mot qui désigne le plus court chemin entre deux points. Prenons deux points dans le plan : quel est le meilleur chemin qui les relie ? La réponse est bien connue : c'est la ligne droite. Et si on se place sur une sphère ? Ce sont les arcs de cercle maximal : l'équateur, les méridiens (mais pas les parallèles) ; on le voit d'ailleurs si l'on pense à la trajectoire d'un vol transatlantique.

Ce qui est vrai pour des points peut être vrai pour des « nuages de points » ou d'autres objets : les manipuler au mieux revient à trouver le meilleur chemin possible pour amener ces objets tous ensemble de leur état initial à l'état final désiré.

On va considérer tout d'abord un exemple : le cas d'un nuage fini constitué de quatre points dans le plan (voir la figure 1).

On se donne la configuration du nuage à deux instants donnés, et on cherche comment le transporter de façon optimale entre les deux configurations. À l'instant de départ, le nuage est donc défini par une liste de quatre points A_1, A_2, A_3, A_4, et à l'instant final par quatre autres points B_1, B_2, B_3, B_4.

Un point clé à comprendre est qu'on ne se soucie aucunement de l'individualité des « particules » et que leur numérotation est arbitraire. Ainsi, il est possible de transporter, par exemple, A_1 vers B_1, A_2 vers B_3, A_3 vers B_4 et A_4 vers B_2. (En revanche, on n'est pas autorisé à déplacer A_1 et A_2 vers B_1 en laissant B_4 vacant, par exemple.) Il faudra donc chercher une solution optimale parmi tous les arrangements possibles.

Un arrangement correspond à une manière d'associer à chaque point A un point B, de manière à ce que chaque point de départ soit connecté à un et un seul point à l'arrivée. Le nombre d'arrangements possibles des N premiers entiers s'appelle la factorielle de N et s'écrit $N!$ Il vaut 24 pour $N = 4$ (et plus de 20 milliards pour $N = 15$).

L'association σ correspondant à notre exemple (figure du milieu) s'écrit : $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 4$, $\sigma(4) = 2$, alors que l'association optimale serait $\sigma(1) = 4$, $\sigma(2) = 2$, $\sigma(3) = 3$, $\sigma(4) = 1$.

Trouver la solution optimale revient à minimiser un *coût*, qui est choisi au cas par cas d'après les applications envisagés. Le coût de Monge était donné par la somme de toutes les distances entre les points de départ et les points d'arrivée correspondants. D'autres coûts sont possibles. Par exemple, on pourrait considérer la somme des carrés de ces distances (c'est le coût *quadratique*).

Un problème d'optimisation combinatoire

Tel que nous l'avons formulé, notre problème de transport optimal appartient à une branche importante des mathématiques : l'optimisation combinatoire. Dans cette discipline, notre problème est considéré comme « facile ». En effet, il existe des algorithmes qui permettent de trouver l'arrangement optimal en un nombre d'opérations proportionnel au cube du nombre de points (cela veut dire que, sur un ordinateur où les opérations s'effectuent séquentiellement, sans parallélisme, le temps de calcul pour N points vaudra plus ou moins une constante fois N^3).

Rappelons que, comme le nombre d'arrangements est $N! = N(N-1) \times \ldots \times 3 \times 2 \times 1$, il est

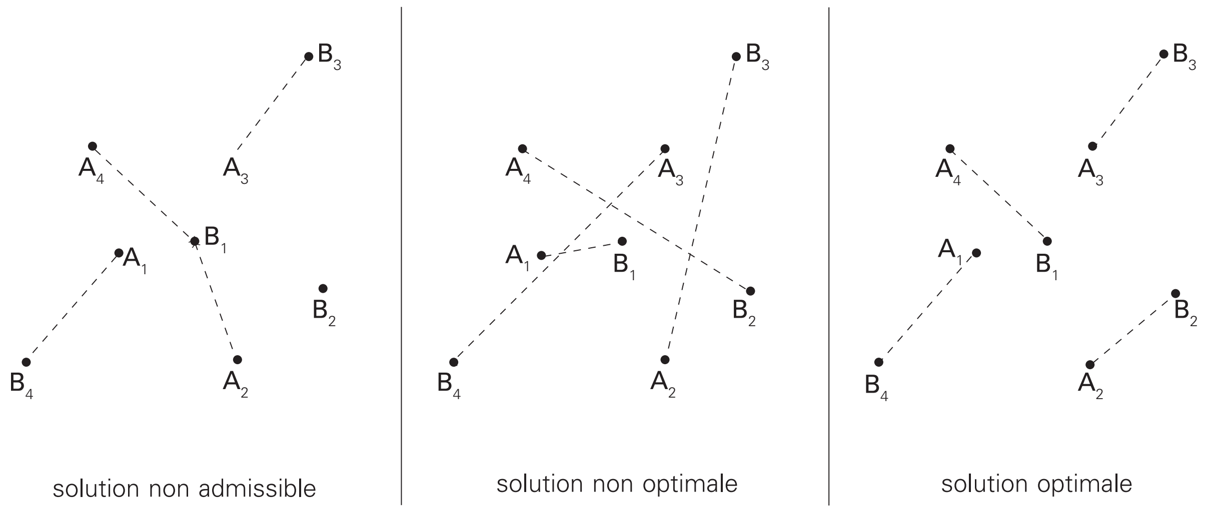


Figure 1. Transport d'un nuage de quatre points.

donc déjà remarquable de rendre le temps de calcul cubique par rapport à N (en effet, si pour $N = 4$ la factorielle vaut 24 et reste modérée, lorsque N est grand, elle augmente presque comme N^N, bien plus vite que N^3 en tous cas, et devient rapidement astronomique).

Cependant, même si on a remarqué que N^3 est beaucoup plus petit que $N!$, en pratique, les algorithmes connus sont peu performants lorsque N devient grand (au delà du millier, disons). On peut rêver d'un algorithme dont le coût de calcul croîtrait à peu près proportionnellement à N. Les applications en seraient spectaculaires. Mais pour approcher ce rêve, il faut changer de point de vue, et considérer non plus un nombre limité de points, mais des points en très grand nombre, voire une infinité de points. On passe alors d'un problème *discret* à un problème *continu*.

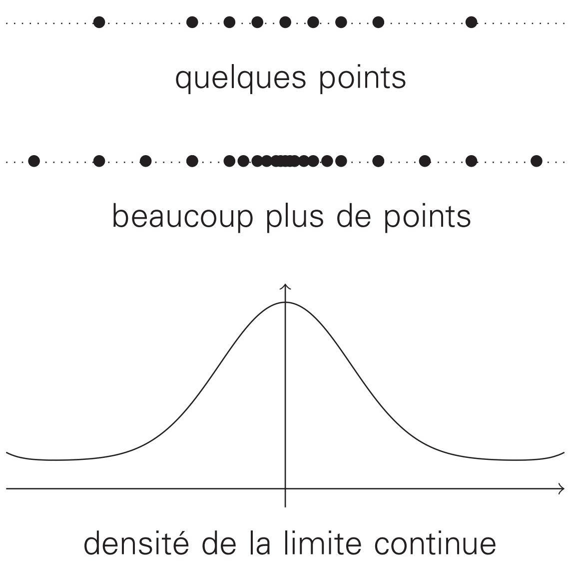


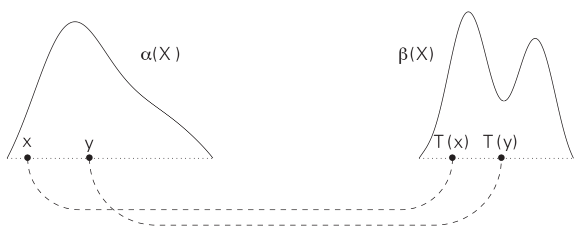
Figure 2. Cas d'un nuage continu de points

On peut rêver d'un algorithme dont le coût de calcul croîtrait à peu près proportionnellement à N. Les applications en seraient spectaculaires.

Transport optimal d'un nuage continu

Comme souvent en mathématiques et en physique, la limite continue du problème discret précédemment discuté permet de mettre en jeu toute la force du calcul différentiel et intégral (en particulier, les sommes que l'on effectue dans le cas discret sont remplacées par des intégrales dans le cas continu).

Dorénavant, notre nuage de points sera décrit, respectivement au départ et à l'arrivée, par deux fonctions (deux *densités de probabilité*) donnant, en tout point X de l'espace, la densité du nuage. On note $\alpha(X)$ la densité de départ et $\beta(X)$ celle d'arrivée. Ce sont deux fonctions réelles positives, dont l'intégrale sur tout l'espace vaut 1 et l'intégrale sur chaque ensemble A vaut la proportion de points du nuage se trouvant dans A. Dans ce cadre continu, on cherche donc à déterminer le transport optimal qui sera une fonction T envoyant les points de départ vers des points d'arrivée.



Dans le cas continu, en considérant le coût quadratique, c'est-à-dire l'intégrale du carré de la distance entre le nuage de départ et celui d'arrivée (au lieu de la somme des carrés des distances entre les points de départ et les points d'arrivée du cas discret), des résultats théoriques obtenus dans la seconde moitié des années 1980 permettent de trouver la solution.

Pour les comprendre, il est utile de voir ce qui se passe dans le cas de l'espace à une seule dimension, c'est-à-dire quand la variable X et les densités α et β vivent sur une droite. Dans ce cas, le résultat nous dit que le transport optimal T est le *transport monotone croissant*, c'est-à-dire celui qui envoie le point le plus à droite de la densité α sur le point le plus à droite de β, et qui respecte l'ordre ($X \leq Y$ donne $T(X) \leq T(Y)$; on peut démontrer qu'il n'y a qu'une seule fonction T avec cette propriété qui respecte les densités α et β).

En dimension supérieure à 1, un problème se pose cependant : qu'est-ce que cela signifie, qu'une fonction de deux variables est « croissante » ? A priori, cela n'a pas vraiment de sens… Pourtant, on peut utiliser une astuce : en dimension 1, T est croissante si et seulement si c'est la dérivée d'une fonction convexe (c'est-à-dire d'une fonction dont le graphe est courbé vers le haut, dont la pente augmente avec l'abscisse, et qui est caractérisée par le fait que la dérivée de sa dérivée est positive). La notion de fonction convexe a aussi un sens en dimension supérieure, et le langage qui a été choisi pour énoncer le théorème concernant le transport optimal en dimension supérieure est bien celui-ci.

Le résultat des années 80 évoqué auparavant fait bien intervenir une fonction convexe $\Phi(x,y)$, que l'on appelle « potentiel ». Il nous dit que le transport optimal T entre les deux fonctions de densité α et β est égal au *gradient* de cette fonction $\Phi(x,y)$, noté $D\Phi(x,y)$, qui est le vecteur dont les composantes sont, en chaque point, la dérivée de Φ par rapport à x et sa dérivée par rapport à y.

Reste cependant à trouver cette fonction « potentiel » Φ qui nous permettra de définir la trajectoire optimale de chacune des particules. En exprimant le fait que T envoie la densité α sur β, on obtient des équations différentielles (faisant intervenir des dérivées). En dimension 1 l'équation impose une égalité sur la dérivée de T, donc la dérivée seconde de Φ. En dimension supérieure, on obtient une formule exprimant la valeur du déterminant de la matrice donnée par les dérivées secondes de Φ (par rapport aux différentes variables), qui doit être égal au ratio entre les densités α et β. Cette équation admet une solution unique Φ et permet donc de trouver le potentiel et, ensuite, la fonction T.

Ainsi, paradoxalement, le fait d'avoir « compliqué » le problème (en passant du discret au continu) nous a permis de le résoudre.

Une application inattendue : modifier les images

Ce problème d'essence mathématique rencontre déjà des applications bien concrètes, comme le montre cet exemple étudié dans les années 2000 par l'équipe de Jean-Michel Morel à l'École Normale Supérieure de Cachan.

On dispose de l'image numérisée d'une peinture dont les couleurs sont défraîchies. Ces couleurs sont numériquement codées selon les couleurs élémentaires bleue, rouge et verte. À chaque pixel est donc associée une certaine proportion de bleu, de rouge ou de vert. En balayant les N pixels de l'image numérique, on obtient un nuage de N points dans l'espace tridimensionnel des couleurs. (Il s'agit d'un espace abstrait, pas de notre espace physique.) Par ailleurs, on a une idée de la palette de l'auteur, obtenue en considérant ses autres œuvres. Si l'on peut en faire une statistique raisonnable, on pourra ainsi définir un nuage « de référence », constitué de N points dans le même espace de couleurs. Une technique de rehaussement des couleurs consiste alors à effectuer le transport optimal, dans l'espace des couleurs, du nuage de points fourni par l'image défraîchie vers le nuage de points de référence. On déplacera alors en conséquence les valeurs des couleurs de chaque pixel.

Notons que dans cet exemple, c'est bien grâce à la formulation continue que le problème a pu être étudié, ceci indépendamment du fait que le calcul numérique est effectué sur un nombre fini de points.

« Jeunes filles au bord de la mer » de P. Puvis de Chavanne.

Une autre application… aux mathématiques elles-mêmes !

Dans les années 1990, le transport optimal devient un outil de démonstration puissant et élégant pour démontrer de nombreuses « inégalités » de la géométrie et de l'analyse fonctionnelle.

Le cas le plus simple à décrire est la fameuse « inégalité isopérimétrique », solution du problème de Didon déjà mentionné par Virgile, qui peut s'énoncer ainsi : de toutes les courbes fermées de longueur donnée, celle qui entoure l'aire la plus grande est le cercle. Si on veut s'en tenir à la légende, Didon avait obtenu du roi local des terres, mais « autant qu'il en pourrait tenir dans la peau d'un bœuf ». Elle prit alors une peau, la fit découper en ficelles et fit une longue corde pour contourner la plus grande surface possible, en résolvant donc un « problème isopérimétrique » et en fondant la ville de Carthage.

Or, si on prend une figure F dans le plan et on la compare au cercle C de même aire, en considérant le transport optimal entre deux densités uniformes sur F et sur C et en utilisant les propriétés de Φ, on peut démontrer que le périmètre de F est forcément plus grand que celui de C. De même, dans l'espace, le volume maximal entouré par une surface donnée est réalisé par la sphère.

Ces résultats géométriques étaient bien connus avant le transport optimal et cela en donne juste une preuve alternative ; une méthode similaire est néanmoins utilisée pour réaliser la démonstration d'inégalités non démontrées auparavant.

Mathématiques pures et appliquées se répondent.

La voici donc bouclée, cette boucle de deux siècles. Une boucle où mathématiques pures et appliquées se répondent, où l'in-

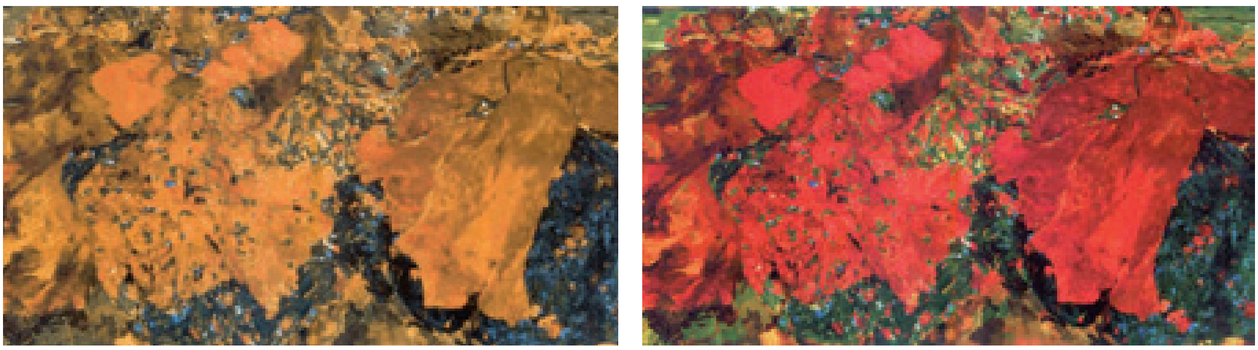


Figure 3. À droite : « Whirlwind » de F. Maliavin. À gauche : le tableau de F. Maliavin repeint avec les teintes de Puvis de Chavanne, par des méthodes de transport optimal. D'après Delon-Salomon-Sobolevski, SIAM J. Appl. Math.

formatique théorique permet de mieux comprendre ce qui peut se calculer efficacement. Et surtout, où c'est en « changeant de point de vue » (ici, en passant d'un problème discret à un problème continu) qu'un problème compliqué peut livrer une solution inattendue.

Pour aller plus loin

Cet article est inspiré de celui du même auteur paru dans *Interstices*.

La combinatoire

La combinatoire est le domaine des problèmes mathématiques où l'on énumère et classifie les manières de relier ou d'arranger un nombre fini, mais souvent élevé, d'objets. Par exemple :

Combien de plaques minéralogiques différentes du type LL CCC LL peuvent être émises ?

Et si l'on veut que deux plaques différentes se distinguent par deux caractères au moins ?

Combien de couleurs sont nécessaires au minimum pour colorier un planisphère de manière à ce que deux états ayant une frontière commune n'aient pas la même couleur ?