C. R. Acad. Sci. Paris, Ser. I 334 (2002) 793–795

Probabilités/*Probability Theory*

Problème de Monge pour *n* probabilités

Henri Heinich

INSA de Rouen, Département de génie mathématique, place E. Blondel, 76131 Mont-Saint-Aignan cedex, France

Reçu le 15 juin 2001 ; accepté après révision le 19 février 2002

Note présentée par Marc Yor.

Résumé Dans cette Note nous généralisons un théorème de Gangbo et Swiech, sur une solution au problème de Monge pour n probabilités avec la distance de Wasserstein. Dans le cadre des espaces d'Orlicz et plus généralement celui des espaces de Köthe, nous étudions ce problème pour une fonction $c(x_1,\ldots,x_n)=h(\sum x_i)$, h convexe sur $\mathbb{R}^d$. ***Pour citer cet article : H. Heinich, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 793–795.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Monge problem for *n* probabilities

Abstract In this Note, we generalize Gangbo–Swiech theorem for the Monge–Kantorovich problem. We study this problem for Orlicz and Köthe spaces when the function c has the form $c(x_1,\ldots,x_n)=h(\sum x_i)$, h convex on $\mathbb{R}^d$. ***To cite this article: H. Heinich, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 793–795.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Toutes les variables aléatoires (v.a.) considérées sont définies sur l'espace canonique $([0,1],\mathcal{B},\lambda)$, on note $\mathcal{L}(X)$ la loi d'une v.a. X, et $E[X]$ son espérance. Dans la suite $\Gamma=(P_1|\cdots|P_n)$ est l'ensemble des probabilités sur $(\mathbb{R}^d)^n$ dont les marges sont $P_1,\ldots,P_n$, probabilités sur $\mathbb{R}^d$. Si $|\cdot|$ est la norme euclidienne de $\mathbb{R}^d$, sur Γ on considère la fonctionnelle de Wasserstein :

$$l_2(\Gamma):=\inf\bigl\{E\bigl[c(X)\bigr],\ \mathcal{L}(X)\in\Gamma\bigr\},\qquad \text{où } c(x_1,\ldots,x_n)=\sum_1^n|x_i-x_{i+1}|^2,\ x_{n+1}:=x_1.$$

Le problème de Monge est celui de l'existence de n fonctions de $\mathbb{R}^d\to\mathbb{R}^d$, ϕ_i, $i=1,\ldots,n$, $\phi_1(x)=x$, $\mathcal{L}(\phi(X_1):=(\phi_1(X_1),\ldots,\phi_n(X_1))\in\Gamma$ et $l_2(\Gamma)=E[c(\phi(X_1))]$. Lorsque $n=2$, l'existence d'une telle fonction, dite de Monge, résulte de [2,9]. Le cas $n\geqslant 3$, est plus complexe. Pour des lois gaussiennes, Knott et Smith montrent dans [7] l'existence d'une variable gaussienne X optimale. Gangbo et Swiech dans [5] donnent la solution au problème de Monge pour des probabilités absolument continues par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Rüschendorf et Uckelmann simplifient ce résultat dans [10]. Toutes ces preuves utilisent la bilinéarité du produit scalaire et la linéarité de l'espérance. Ainsi la preuve de [10]

Adresse e-mail : heinich@insa-rouen.fr (H. Heinich).

S1631-073X(02)02341-5/FLA

consiste à voir que le problème de Monge, pour la fonction c précédente, est équivalent au problème pour la fonctionnelle $\sup E[|\sum_1^n X_i|^2]$, puis que $(X_i, \sum_{j\neq i} X_j)$ est un couple optimal. Le problème pour n probabilités est ainsi ramené à celui des couples.

Nous étendons ce résultat dans deux directions, en remplaçant la fonctionnelle de Wasserstein par des fonctionnelles plus générales et la fonction de coût $c(x_1,\dots,x_n) = |\sum x_i|^2$ par $c(x_1,\dots,x_n) = h(\sum x_i)$, h convexe, $h(0) = 0$.

Le résultat principal de cette Note est :

THÉORÈME. – *Supposons les probabilités P_i absolument continues par rapport à la mesure de Lebesgue de $\mathbb{R}^d$. Si $(\mathbb{E}, \|\cdot\|)$ est un espace d'Orlicz modéré avec sa norme d'Orlicz ou celle de Luxemburg, il existe une fonction de Monge, $\phi : \mathbb{R}^d \to (\mathbb{R}^d)^n$, telle que*

$$\left\|c\big(\phi(X_1)\big)\right\| = \operatorname{Inf}\big\{\left\|c(X)\right\|,\ \mathcal{L}(X)\in\Gamma\big\},\quad \mathcal{L}(X_1) = P_1.$$

Si $(\mathbb{E}, \|\cdot\|)$ est un espace de Köthe Id-*convexe, il existe une fonction de Monge ϕ telle que*

$$\left\|c\big(\phi(X_1)\big)\right\| = \operatorname{Sup}\big\{\left\|c(X)\right\|,\ \mathcal{L}(X)\in\Gamma\big\},\quad \mathcal{L}(X_1) = P_1.$$

2. Rappels et résultats

Dans la suite $\mathbb{E}$ est un espace de Köthe de v.a. réelles, [8], que nous supposons séquentiellement complet, et invariant par réarrangement : $\mathcal{L}(X) = \mathcal{L}(Y)$, $X \in \mathbb{E}$ implique $Y \in \mathbb{E}$ et $\|X\| = \|Y\|$. Un tel espace est dit *lisse* si, pour tout X, $\|X\| = 1$, on écrit $X \in S(\mathbb{E})$, il existe une unique application de dualité $X \to D_X$, $D_X \in S(\mathbb{E}')$ vérifiant $\|X\| = E[D_X \cdot X]$.

L'espace $\mathbb{E}$ est Id-*convexe* s'il est lisse et si l'application de dualité s'écrit $D_X = f_X(X)$ où f_X ne dépend que de la loi de X et $x \to x f_X(x)$ est convexe.

Çitons comme exemples les espaces $\mathbb{L}^p$, $1 \leqslant p < \infty$, et plus généralement les espaces d'Orlicz avec la norme de Luxemburg ou celle d'Orlicz, dès que U est modérée et $xu(x)$ convexe, cf. [3].

Pour $c(x = (x_1,\dots,x_n)) = h(\sum x_i)$, $h(0) = 0$, h continue convexe, on définit les deux fonctionnelles : $I(\Gamma) := \inf\{\|c(X)\|,\ \mathcal{L}(X)\in\Gamma\}$ et $S(\Gamma) := \sup\{\|c(X)\|,\ \mathcal{L}(X)\in\Gamma\}$. La première est adaptée lorsque c représente un coût, la seconde lorsque c est un profit.

On sait que ces deux fonctionnelles sont atteintes, i.e. si ∂ désigne une des deux fonctionnelles I ou S, il existe une v.a. X, dite ∂-optimale, telle que $\partial(\Gamma) = \|c(X)\|$, [6].

Montrons que l'étude de la fonctionnelle I pour un espace d'Orlicz se ramène au cadre traditionnel de $\mathbb{L}^1$. À cette fin rappelons que la norme de *Luxemburg* d'une v.a.r. X, soit $\|X\|_{(U)}$, est définie par $E_\lambda\big[U\big(\frac{X}{\|X\|_{(U)}}\big)\big] = \int_0^1 U\big(\big|\frac{X(t)}{\|X\|_{(U)}}\big|\big)\,\mathrm{d}t = 1$ et sa norme d'*Orlicz* (par dualité) est : $\|X\|_U := \sup\{E[XY],\ \|Y\|_{(V)} \leqslant 1\}$. Si $k = I_{c,\mathbb{L}^{(U)}}(P,Q)$, et μ est la loi d'un couple optimal (X,Y), on a l'inégalité, cf. [3] :

$$E_\lambda\left[U\left(\frac{c(X,Y)}{k}\right)\right] = E_\mu\left[U\left(\frac{c(x,y)}{k}\right)\right] = 1 \leqslant E_\nu\left[U\left(\frac{c(x,y)}{k}\right)\right],\quad \forall \nu \in (P|Q).$$

En posant $c^k(\cdot) = U(\frac{c(\cdot)}{k})$, cette relation s'interpète de la manière suivante :

(∗) μ est optimale pour $(c, \mathbb{L}^{(U)}, P, Q)$ et $k = I_{c,\mathbb{L}^{(U)}}(P,Q)$, si et seulement si, μ est optimale pour $(c^k, \mathbb{L}^1, P, Q)$ et $I_{c^k,\mathbb{L}^1}(P,Q) = 1$.

L'analogue pour la norme d'Orlicz utilise la relation, cf. [3] :

$$\|X\|_U = E\left[Xu\left(\frac{X}{k}\right)\right] = k\left(E\left[U\left(\frac{X}{k}\right)\right] + 1\right),\quad \text{où } E\left[V\circ u\left(\frac{X}{k}\right)\right] = 1.$$

Le minimum de la fonction $x \to f(x,X) = x(E[U(X/x)]+1)$ est atteint pour $x = k_X$ et vaut $f(k_X, X) = \|X\|_U$. Soit (X,Y) un couple optimal

$$\big\|c(X,Y)\big\|_U = I_{c,\mathbb{L}^U}(P,Q),\ \mathcal{L}(X,Y) = \mu \in (P|Q)$$

notons $k = \|c(X,Y)\|_{(V \circ u)}$. Comme $f(k, c(X,Y)) \leqslant f(k, c(U,V))$, si $\mathcal{L}(U,V) = \nu \in (P|Q)$, on obtient :

$$\big\|c(X,Y)\big\|_U = k\left(E_\mu\left[U\left(\frac{c(x,y)}{k}\right)\right]+1\right) \leqslant k\left(E_\nu\left[U\left(\frac{c(x,y)}{k}\right)\right]+1\right).$$

En notant $c^k(\cdot) = k(U(c(\cdot)/k)+1)$, cette équation s'interprète de la manière suivante :

$(**)$ μ est optimale pour $(c, \mathbb{L}^U, P, Q)$ si et seulement si, μ est optimale pour $(c^k, \mathbb{L}^1, P, Q)$ où $k = \|c(X,Y)\|_{(V \circ u)}$, $\mathcal{L}(X,Y) = \mu$.

3. Preuve du théorème

Pour $x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$, notons $x^i = \sum_{j \neq i} x_j$. Pour la fonctionnelle I et les espaces d'Orlicz. Le couple (X_i, X^i) étant optimal pour (P_i, P^i), le Théorème 2–5 de [6] donne l'existence d'une fonction θ_i, différentielle d'une fonction strictement convexe, telle que $X^i = \theta_i(X_i)$. En effet la procédure précédente ramène le problème au cadre classique $\mathbb{L}^1$ où le résultat est connu, [1] ou [4].

Pour la fonctionnelle S, on montre, Théorème 4–2 de [6], que le support d'une probabilité optimale est $(\mathrm{Id} \cdot f_c) \circ c$-cycliquement monotone. Ceci assure aussi de l'existence de θ_i telle que $X^i = \theta_i(X_i)$. Par conséquent, dans les deux cas, $\sum_1^n X_j = X_i + \theta_i(X_i)$, $1 \leqslant i \leqslant n$. Les fonctions $\mathrm{Id} + \theta_i$ sont des différentielles de fonctions strictement convexes, et donc inversibles, [1,2,4] ou [10]. Il suffit de poser, pour $i > 1$, $\phi_i = (\mathrm{Id} + \theta_i)^{-1} \circ \theta_1$, pour obtenir le résultat. $\square$

Références bibliographiques

[1] T. Abdellaoui, H. Heinich, Caractérisation d'une solution optimale au problème de Monge–Kantorovich, Bull. Soc. Math. France 127 (3) (1999) 429–443.
[2] N. Belili, H. Heinich, Transport problem and derivation, Appl. Math. 26 (3) (1999) 299–314.
[3] B. Bru, H. Heinich, Applications de dualité dans les espaces de Köthe, Stud. Math. 43 (1989) 41–69.
[4] W. Gangbo, R.J. McCann, The geometry of optimal transportation, Acta Math. 177 (1996) 113–161.
[5] W. Gangbo, A.J. Świech, Optimal maps for the multidimensional Monge–Kantorovich problem, Comm. Pure Appl. Math. LI (1998) 23–45.
[6] H. Heinich, Monge problem and Köthe functionals, Preprint, 2001.
[7] M. Knott, C.S. Smith, On a generalization of cyclic monotonicity and distances among random vectors, Linear Algebra Appl. 199 (1994) 367–371.
[8] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces, Springer, Berlin, 1996.
[9] S.T. Rachev, L. Rüschendorf, Mass Transportation Problems, Springer, New York, 1998.
[10] L. Rüschendorf, L. Uckelmann, On the n-coupling problem, Preprint of the Institut für Math. Stochastik, University of Freiburg, 1998.