XAVIER FERNIQUE

Sur le théorème de Kantorovitch-Rubinstein dans les espaces polonais

Séminaire de probabilités (Strasbourg), tome 15 (1981), p. 6-10
<http://www.numdam.org/item?id=SPS_1981__15__6_0>

Novembre 1980

SUR LE THEOREME DE KANTOROVITCH-RUBINSTEIN DANS LES ESPACES POLONAIS

par

X. FERNIQUE

Le théorème de Kantorovitch-Rubinstein dans les espaces polonais est connu et utile ; nous rappelons son énoncé :

THEOREME K.R. : *Soient* (S,d) *un espace polonais,* μ *et* ν *deux probabilités sur* S *vérifiant* :

$$\iint d(x,y)d\mu(x)d\mu(y) < \infty \ , \ \iint d(x,y)d\nu(x)\,d\nu(y) < \infty \ ;$$

il existe alors une probabilité π *sur* $S \times S$ *ayant pour marges* μ *et* ν *et vérifiant* :

$$\iint d(x,y)d\pi(x,y) = \sup\{\int fd\mu - \int fd\nu \ , \ f \in \mathfrak{F}_S\} \ ,$$

$$\mathfrak{F}_S = \{f : \forall\,(x,y) \in S \times S \ , f(x) - f(y) \leq d(x,y)\} \ .$$

Ce théorème est en général prouvé ([1],[2]) à partir de constructions explicites de mesures ; nous en donnons ci-dessous une preuve différente et peut-être plus élémentaire qu'on peut utiliser dans tous les problèmes de ce type.

(a) Supposons pour commencer que μ et ν aient un support fini commun $A = \{a_i, 1 \leq i \leq n\}$ et posons pour tout $i \in [1,n]$, $\mu_i = \mu(a_i) > 0$, $\nu_i = \nu(a_i) > 0$. Nous allons comparer 3 problèmes :

P_1 : maximiser $\sum_1^n f(a_i)(\mu_i - \nu_i)$ sous les conditions : $\forall\, k,\ell \in [1,n]$,
$$f(a_k) - f(a_\ell) \leq d(a_k, a_\ell) \ .$$

P_2 : maximiser $\sum_1^n f(a_i)\mu_i + g(a_i)\nu_i$ sous les conditions : $\forall\, k,\ell \in [1,n]$,
$$f(a_k) + g(a_\ell) \leq d(a_k, a_\ell) \ .$$

P_2' : minimiser $\sum_{1,1}^{n,n} d(a_k,a_\ell)\pi_{k,\ell}$ sous les conditions : $\forall\, k,\ell \in [1,n]$, $\pi_{k,\ell} \geq 0$,

$$\sum_{i=1}^{n} \pi_{i,\ell} = \nu_\ell \ , \ \sum_{j=1}^{n} \pi_{k,j} = \mu_k \ .$$

Remarquons que les problèmes P_2 et P_2' sont des problèmes de transport duaux ; on sait qu'ils ont l'un et l'autre des solutions et la même valeur ; ceci montre que le théorème sera établi dans ce premier cas si on prouve le lemme :

LEMME 1. Dans les conditions ci-dessus, les problèmes P_1 et P_2 ont la même valeur.

Démonstration du lemme : Le problème P_1 ayant évidemment une valeur inférieure ou égale à celle de P_2 , l'inégalité inverse sera prouvée et le lemme établi si nous montrons que tout couple (f,g) maximal pour P_2 a une somme nulle. Or un tel couple vérifie :

$$\forall\, k,\ell \in [1,n] \ , \qquad f(a_k) \leq \inf_{1\leq j\leq n} \{d(a_k,a_j) - g(a_j)\} \ ,$$

$$g(a_\ell) \leq \inf_{1\leq j\leq n} \{d(a_i,a_\ell) - f(a_i)\} \ ,$$

$$f(a_k) + g(a_k) \leq 0 \ ,$$

et nous allons montrer qu'aucune de ces inégalités n'est stricte ; en première et deuxième lignes, c'est immédiat ; supposons en effet par exemple qu'au rang k_o , on ait :

$$f(a_{k_o}) < f'(a_{k_o}) = \inf_{1\leq j\leq n} \{d(a_{k_o},a_j) - g(a_j)\}$$

on en déduirait immédiatement, puisque μ_{k_o} est strictement positif, que (f,g) n'est pas maximal pour P_2 , c'est absurde. Supposons maintenant qu'au rang k_o , on ait :

$$f(a_{k_o}) + g(a_{k_o}) < 0 \ ,$$

la preuve précédente montrerait l'existence de deux rangs i et j tels que :

$$f(a_{k_o}) + g(a_j) = d(a_{k_o}, a_j) \ , \ f(a_i) + g(a_{k_o}) = d(a_i, a_{k_o})$$

on aurait alors :

$$f(a_i)+g(a_j) \le d(a_i,a_j) \le d(a_i,a_{k_o})+d(a_{k_o},a_j) \le f(a_i)+g(a_j)+[f(a_{k_o})+g(a_{k_o})]$$
$$< f(a_i)+g(a_j) \ ,$$

c'est encore absurde ; le lemme est établi et le théorème prouvé dans ce cas (a).

(b) Supposons maintenant que S soit compact ; pour tout entier $p>0$, nous notons $A = A_p$ une partie finie de S (de cardinal n) dont le voisinage d'ordre $\frac{1}{p}$ soit S ; nous notons $h = h_p$ une application mesurable de S dans A vérifiant $d(x,h(x)) \le \frac{1}{p}$; nous choisissons un point a de S et nous posons :

$$\mu_p = (1-\tfrac{1}{p})\mu \circ h^{-1} + \frac{1}{np}\sum_A \delta_a \ , \ \nu_p = (1-\tfrac{1}{p})\nu \circ h^{-1} + \frac{1}{np}\sum_A \delta_a \ ,$$

de sorte que , μ_p et ν_p satisfaisant les hypothèses de (a), il existe une probabilité π_p sur $S \times S$ portée par $A \times A$, ayant μ_p et ν_p pour marges et vérifiant :

$$\iint d(x,y)d\pi_p(x,y) \le \sup\{\int f d\mu_p - \int f d\nu_p \ , \ f \in \mathfrak{F}_A\} \ .$$

Pour transformer ce dernier terme, nous utilisons le lemme suivant :

LEMME 2. <u>Soient</u> (S,d) <u>un espace métrique,</u> A <u>un sous-ensemble de</u> S <u>et</u> f <u>une fonction sur</u> A <u>appartenant à</u> $\mathfrak{F}_A$; <u>alors</u> f <u>est la restriction à</u> A <u>d'une fonction</u> $\bar{f}$ <u>sur</u> S <u>appartenant à</u> $\mathfrak{F}_S$. <u>Si</u> A <u>est mesurable, on a donc pour tout couple</u> (μ,ν) <u>de probabilités sur</u> S <u>portées par</u> A :

$$\sup\{\int f d\mu - \int f d\nu, \ f \in \mathfrak{F}_A\} = \sup\{\int f d\mu - \int f d\nu, \ f \in \mathfrak{F}_S\} \ .$$

La preuve de ce lemme est immédiate ; on peut définir $\bar{f}$ par :

$$\forall \ x \in S \ , \ \bar{f}(x) = \sup\{f(y) - d(x,y) \ , \ y \in A\} \ .$$

L'application du lemme 2 et des définitions de μ_p et ν_p fournit alors :

$$\iint d(x,y)\, d\pi_p(x,y) \leq \sup\{\int f d\mu - \int f d\nu\ ,\ f \in \mathfrak{F}_S\} + \frac{2}{p}\ .$$

On peut alors extraire de la suite π_p une suite convergeant étroitement sur $S \times S$ vers une probabilité π ayant pour marges μ et ν et vérifiant, puisque d est continue bornée :

$$\iint d(x,y) d\pi(x,y) \leq \lim \iint d(x,y) d\pi_p(x,y)$$

c'est le résultat du théorème qui est ainsi prouvé dans le cas (b).

(c) Dans le cas général, nous notons a un élément de S ; les hypothèses du théorème impliquent la convergence de $\int (1+d(x,a))(d\mu(x)+d\nu(x))$; pour tout entier $p>0$, il existe donc une partie compacte $K = K_p$ de S telle que :

$$\int_{S \setminus K_p} [1+d(x,a)](d\mu(x) + d\nu(x)) \leq \frac{1}{p}\ ;$$

nous définissons alors deux probabilités μ_p et ν_p à support compact en posant :

$$\mu_p = I_K \cdot \mu + (1-\mu(K))\, \delta_a\ ,\ \nu_p = I_K \cdot \nu + (1-\nu(K))\, \delta_a\ ;$$

la preuve précédente montre qu'il existe une probabilité π_p sur $S \times S$ de marges μ_p et ν_p vérifiant, d'après les définitions de μ_p et ν_p :

$$\iint d(x,y) d\pi_p(x,y) = \sup\{\int_{K_p} (f(x) - f(a))(d\mu(x) - d\nu(x))\ ,\ f \in \mathfrak{F}_S\}\ ;$$

la définition de K_p majore alors ce dernier membre par :

$$\sup\{\int f(x)(d\mu(x) - d\nu(x)\ ,\ f \in \mathfrak{F}_S\} + \frac{1}{p}\ .$$

Enfin l'ensemble $\{\pi_p\ ,\ p \in \mathbb{N}\}$ ayant des marges étroitement convergentes est relativement compact, on peut en extraire une suite partielle étroitement convergente vers une probabilité π qui a pour marges μ et ν ; on a par ailleurs pour tout nombre $t>0$ et tout entier $p>0$:

$$\begin{aligned}\iint_{d(x,y)>t} d(x,y) d\pi_p(x,y) &\leq \iint_{d(x,a)+d(y,a)>t} [d(x,a)+d(y,a)] d\pi_p(x,y) \\ &\leq 2\int_{d(x,a)>\frac{t}{2}} d(x,a) d\mu(x) + 2\int_{d(y,a)>\frac{t}{2}} d(y,a) d\nu(y)\ .\end{aligned}$$

Ceci montre que d est uniformément intégrable par rapport à l'ensemble $\{\pi_p , p \in \mathbb{N}\}$ et on a donc :

$$\iint d(x,y)d\pi(x,y) \leq \overline{\lim} \iint d(x,y)d\pi_p(x,y) \leq \sup\{ \int fd\mu - \int fd\nu , \ f \in \mathfrak{F}_S\} .$$

Ceci prouve le théorème dans le cas général.

[1] DUDLEY, R.M. Probabilities and metrics, Lecture Notes Series n° 45. Matematik Institut, Aarhus University.

[2] DE ACOSTA, A. Invariance principles in probability for triangular arrays of B - valued random vectors and some applications, to appear.

INSTITUT DE RECHERCHE MATHEMATIQUE AVANCEE
Laboratoire Associé au C.N.R.S.
Université Louis Pasteur
7, rue René Descartes
67084 STRASBOURG CEDEX