WSAA 8

Jiří Navrátil

The Kantorovič-Rubinstein distance

In: Zdeněk Frolík (ed.): Abstracta. 8th Winter School on Abstract Analysis. Czechoslovak Academy of Sciences, Praha, 1980. pp. 117–120.

Persistent URL: http://dml.cz/dmlcz/701191

Terms of use:

© Institute of Mathematics of the Academy of Sciences of the Czech Republic, 1980

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

The Kantorovič-Rubinstein distance

J. Navrátil, Praha

If (X,ρ) is a pseudometric space then the pseudometric ρ naturally induces some distance between probability measures. There are the following natural possibilities how to define such a distance :

$$\rho_1(\mu,\nu)=\sup\{\mu(f)-\nu(f)\mid \operatorname{Lip}(f)\leq 1,\ f \text{ is bounded}\},$$
$$\rho_2(\mu,\nu)=\inf\{\eta(\rho)\mid \pi_1\eta-\pi_2\eta=\mu-\nu\ ,\ \eta \text{ is a measure on } \Sigma\otimes\Sigma\},$$
$$\rho_3(\mu,\nu)=\inf\{\eta(\rho)\mid \pi_1\eta=\mu,\ \pi_2\eta=\nu\ ,\ \eta \text{ is a probability measure on } \Sigma\otimes\Sigma\},$$

where μ,ν are probability measures on a σ-algebra Σ on X containing all Borel sets ($\mu(f)=\int f\,d\mu$, $\pi_1\eta(A)=\eta(A\times X)$ etc.). The last two metrics are usually called the Kantorovič-Rubinstein distances.

Under certain conditions all metrics given above are equal. That's the reason why it is convenient to work with the Kantorovič-Rubinstein distance.

Kantorovič has shown in [1] that $\rho_1=\rho_2$ if X is a compact metric space. In [2] and [3] Kantorović and Rubinstein proved (essentially) that $\rho_1=\rho_2$ if X is a compact metric space. We shall show that $\rho_1=\rho_2=\rho_3$ if X is an arbitrary separable pseudometric space.

We shall use the following theorem on a non-negative extension of a linear functional.

<u>Theorem 1.</u> Let E be an ordered vector space, let I be a non-negative linear functional on a subspace F of E. Let

$$(\forall y\in E)(\exists z\in F)\quad y\leq z$$

(i.e. F is a majorizing subspace of E).

Then there is a non-negative linear extension of I to E.

.or th ro se [4 . 3

Now nu m

Theorem 2. t (,ρ b a e abl p eu o tric space, let μ, ν be two probability measures on a σ-algebra Σ on X containing all Borel sets. Let ρ_1, ρ_2, ρ_3 be the distances defined above.

Then $\rho_1 = \rho_2 = \rho_3$.

Proof: (1) Let η be a measure on $\Sigma \otimes \Sigma$ such that $\pi_1\eta - \pi_2\eta = \mu - \nu$. Then it holds

$\mu(f) - \nu(f) = \int f.\, d\pi_1\eta - \int f.\, d\pi_2\eta = \int f(x)\, d\eta(x,y) - \int f(y)\, d\eta(x,y) \leq \int \rho\, d\eta = \eta(\rho)$ for each bounded function f with $\mathrm{Lip}(f) \leq 1$.

Thus we have $\rho_1 \leq \rho_2$. Obviously $\rho_2 \leq \rho_3$.

If $\rho_1 = \infty$, then $\rho_1 = \rho_2 = \rho_3 = \infty$. Thus let us consider the case $\rho_1 < \infty$.

(2) Lemma. Let φ, ψ be bounded measurable functions, $a \in R$ and

$\varphi(x) + \psi(y) + a\rho(x,y) \geq 0$ for all $x, y \in X$.

Then it holds $\mu(\varphi) + \nu(\psi) + a\rho_1 \geq 0$.

Proof of the lemma:

a) For $a \leq 0$ we have

$\mu(\varphi) + \nu(\psi) = \int \varphi(x)\, d\eta(x,y) + \int \psi(y)\, d\eta(x,y) \geq (-a)\int \rho(x,y)\, d\eta(x,y) \geq -a\rho_3 \geq -a\rho_1$, where $\eta = \mu \otimes \nu$ (the last inequality is valid by virtue of (1)).

b) Let $a > 0$. Put $h(x) = \inf \{\psi(y) + a\rho(x,y) \mid y \in X\}$.

Then substituting y=x we get $h(x) \leq \psi(x)$ and by the assumption $h(x) \geq -\varphi(x)$, hence h is a bounded function.

For a fixed $y \in X$

$$\psi(y) + a\rho(x,y)$$

is a Lipschitz function with the constant a, thus $\mathrm{Lip}(h) \leq a$ as well. Hence we have

$$a\rho_1 \geq \mu(h) - \nu(h) \geq -\mu(\varphi) - \nu(\psi).$$

(3) By means of the lemma one can easily show that the formula

$$\tilde{\eta}(f) = \mu(\varphi) + \nu(\psi) + a\rho_1$$

gives a sound definition of a non-negative linear functional for all $f(x,y)=\varphi(x)+\psi(y)+a\rho(x,y)$, where φ,ψ are bounded measurable functions and $a\in\mathbf{R}$.

By the theorem 1 and by the lemma there is a non-negative linear extension of $\tilde{\eta}$ to all functions majorized (in the absolute value) by $a\rho(x,y)+b$ (where a,b are positive constants). We shall denote the extension by $\tilde{\eta}$ as well.

(4) For an arbitrary $\varepsilon>0$ there is a sequence of pairwise disjoint sets $A_n\in\Sigma$ such that diam $A_n<\varepsilon$ and $\bigcup_{n=1}^{\infty}A_n=X$, for (X,ρ) is a separable space.

Put

$$c_{ij}=\begin{cases}0 & \text{for } \mu(A_i).\nu(A_j)=0\\ \dfrac{\tilde{\eta}(A_i\times A_j)}{\mu(A_i)\nu(A_j)} & \text{otherwise,}\end{cases}$$

and

$$\eta(B)=\sum_{i,j=1}^{\infty}c_{ij}\,\mu\otimes\nu\,(B\cap(A_i\times A_j))\ .$$

Then η is a non-negative σ-additive measure on $\Sigma\otimes\Sigma$.

Furthermore we have

$$\eta(A\times X)=\sum_{\mu(A_i)\nu(A_j)\neq 0}\frac{\tilde{\eta}(A_i\times A_j)}{\mu(A_i)\nu(A_j)}\cdot\mu\otimes\nu((A\cap A_i)\times A_j)=$$

$$=\sum_{\mu(A_i)\neq 0}\frac{\tilde{\eta}(A_i\times A_j)}{\mu(A_i)}\cdot\mu(A\cap A_i)$$

for all $A\in\Sigma$ (if $\nu(A_j)=0$ then $\tilde{\eta}(A_i\times A_j)\leq\tilde{\eta}(X\times A_j)=\nu(A_j)=$ $=0$, thus $\tilde{\eta}(A_i\times A_j)=0$).

(5) Denote $B_n=\bigcup_{k=1}^{n}A_k$. Then it holds

$0\leq\tilde{\eta}(A_i\times X)-\sum_{j=1}^{n}\tilde{\eta}(A_i\times A_j)\leq\tilde{\eta}(A_i\times(X-B_n))\leq\tilde{\eta}(X\times(X-B_n))=$ $=\nu(X-B_n)$,

but $(X-B_n)\searrow\emptyset$, hence

$$\sum_{j=1}^{\infty}\tilde{\eta}(A_i\times A_j)=\tilde{\eta}(A_i\times X)=\mu(A_i).$$

Thus we have

$\eta(A\times X)=\sum^{\infty}\mu(A\cap A_i)=\mu(A)$ for all $A\in\Sigma$ (if $\mu(A_i)=0$

then obviously $\tilde{\eta}(A_i \times A_j)=0$),

analogously $\eta(X \times A)=\nu(A)$ for all $A \in \Sigma$, i.e. $\pi_1\eta=\mu$, $\pi_2\eta=\nu$.

(6) Put $\bar{\rho} = \sup \rho \mid A_i \times A_j$ on $A_i \times A_j$ and analogously

$\underline{\rho} = \inf \rho \mid A_i \times A_j$ on $A_i \times A_j$.

Then it holds

$$\eta(\rho) \le \eta(\bar{\rho}) \le \eta(\underline{\rho})+2\varepsilon \le \tilde{\eta}(\rho)+2\varepsilon=\rho_1+2\varepsilon,$$

for $\eta(\underline{\rho}\cdot c_{B_n \times B_n}) \to \eta(\underline{\rho})$ and

$$\eta(\underline{\rho}\cdot c_{B_n \times B_n}) = \sum_{i,j=1}^{n} c_{ij}\,\mu\otimes\nu(A_i \times A_j).\inf \rho \mid A_i \times A_j =$$

$$=\sum_{i,j=1}^{n} \tilde{\eta}(A_i \times A_j).\inf \rho \mid A_i \times A_j = \sum_{i,j=1}^{n} \tilde{\eta}(\underline{\rho}\cdot c_{A_i \times A_j}) =$$

$$= \tilde{\eta}(\underline{\rho}\cdot c_{B_n \times B_n}) \le \tilde{\eta}(\rho).$$

Thus we have $\rho_3 \le \rho_1$.

<u>Remark.</u> The main result (and its proof) remains valid in case that ρ satisfies only the following conditions

$$\rho(x,x)=0, \quad 0 \le \rho(x,y) < \infty, \quad \rho(x,y) \le \rho(x,z)+\rho(z,y)$$

for all $x,y,z \in X$ if we replace $\mathrm{Lip}(f) \le a$ by $f(x)-f(y) \le a\rho(x,y)$ (X is supposed to be separable in the topology defined by the basis $\{y \in X \mid \rho(x,y)<\varepsilon,\ \rho(y,x)<\varepsilon\}$, $x \in X$, $\varepsilon > 0$.

References

[1] L.V.Kantorovič: Dokl. Akad. Nauk SSSR 37 (1942), No 7-8, 227-230

[2] L.V.Kantorovič, G.Š.Rubinstein: Dokl. Akad. Nauk SSSR 115 (1957), No 6, 1058-1061

[3] L.V.Kantorovič, G.Š.Rubinstein: Vestnik Leningrad. Univ., Ser. mat. 1958, No 7, ed.2, 52-59

[4] B.Z.Vulikh, Introduction to the theory of partially ordered vector spaces, Groningen 1967