Problema de Monge-Kantorovich

Glenier L. Bello (UZ)
María de Gádor Cabrera (UAL)
Ángela Capel (UGR)
Antonio J. Molero (US)
Leticia Pardo (UM)
Alejandro Poveda (UV)
Profesor: Rafael Espínola

IV Escuela-Taller de Análisis Funcional
7 de marzo de 2014

1. Introducción
2. Medidas imagen
3. Representación de ciertos funcionales
4. Ejemplos
   - Masa de Dirac
   - Hipercubo discreto
5. Principio de Dualidad de Kantorovich
   - Dualidad Monge-Kantorovich en espacios completamente regulares
   - Teorema de Kantorovich en funciones superiormente semicontinuas negativas
   - Teorema de Kantorovich: demostración general
   - Teorema de Kantorovich en espacios métricos
   - Teorema de Kantorovich-Rubinstein

Problema: ¿Cómo realizar el transporte a mínimo coste?

Antes de dar respuesta a esta cuestión, tenemos que tener claro qué es un camino de transporte o **plan de transferencia**. Modelaremos los planes de transferencia con la medida de probabilidad π sobre el espacio producto $X \times Y$.
Decir que un plan de transferencia $\pi \in P(X \times Y)$ es **admisible** (**coupling**) significa que

$$\int_Y d\pi(x,y) = d\mu(x), \qquad \int_X d\pi(x,y) = d\nu(y)$$

Más rigurosamente, exigimos que

$$\pi[A \times Y] = \mu[A], \qquad \pi[X \times B] = \nu[B], \tag{1}$$

para todo subconjunto A de X y B de Y.

Esto es equivalente a declarar que para toda función φ, ψ en una clase adecuada de funciones,

$$\int_{X\times Y}[\varphi(x)+\psi(y)]d\pi(x,y) = \int_X \varphi(x)d\mu(x) + \int_Y \psi(y)d\nu(y) \quad (2)$$

Las medidas de probabilidad π que satisfacen (1) se dice que tienen **marginales** μ y ν, y serán planes de transferencia admisibles. El conjunto de todas estas medidas de probabilidad lo denotaremos por

$$\Pi(\mu,\nu) = \{\pi \in P(X\times Y) : (1)\,\text{se cumple para todo medible}\,A, B\}$$

Este conjunto es siempre no vacío, ya que el tensor producto $\mu \otimes \nu$ pertenece a $\Pi(\mu,\nu)$.

Problema de transporte óptimo de Kantorovich.

Definición

$$Minimizar \quad I[\pi] = \int_{X\times Y} c(x,y)d\pi(x,y) \qquad para \quad \pi \in \Pi(\mu,\nu)$$

Dado un plan de transferencia π, la cantidad no negativa $I[\pi]$ se llama **coste de transporte total** asociado a π. El **coste de transporte óptimo** para este problema es el valor

$$T_c(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi]$$

Si el ínfimo es un mínimo π, i.e., $I[\pi] = T_c(\mu,\nu)$, se denominará **plan de transferencia óptimo**.

Medidas imagen / Problema de Monge

- El problema de Monge es, en esencia, el de Kantorovich pero con una condición adicional: **las masas que se transportan no se pueden dividir**, esto es, a cada origen x le corresponde un único destino y.
- Viéndolo en el caso de planes de transporte:

$$d\pi(x,y) = d\pi_T(x,y) \equiv d\mu(x)\delta[y = T(x)],$$

siendo $T : X \to Y$ una función medible.

Definición

$(\mathbf{Id} \times \mathbf{T})\#\mu$ es la medida de probabilidad en $Z = X \times Y$ que satisface la siguiente propiedad: para cualquier función medible no negativa ζ en Z, se tiene que:

$$\int_{X\times Y} \zeta(x,y)d\pi_T(x,y) = \int_X \zeta(x,T(x))d\mu(x).$$

Medidas imagen / Problema de Monge

- El problema de Monge es, en esencia, el de Kantorovich pero con una condición adicional: **las masas que se transportan no se pueden dividir**, esto es, a cada origen x le corresponde un único destino y.
- Viéndolo en el caso de planes de transporte:

$$d\pi(x,y) = d\pi_T(x,y) \equiv d\mu(x)\delta[y = T(x)],$$

siendo $T : X \to Y$ una función medible.

Definición

$(\mathbf{Id} \times \mathbf{T})\#\mu$ es la medida de probabilidad en $Z = X \times Y$ que satisface la siguiente propiedad: para cualquier función medible no negativa ζ en Z, se tiene que:

$$\int_{X\times Y} \zeta(x,y)d\pi_T(x,y) = \int_X \zeta(x,T(x))d\mu(x).$$

Medidas imagen / Problema de Monge

- El problema de Monge es, en esencia, el de Kantorovich pero con una condición adicional: **las masas que se transportan no se pueden dividir**, esto es, a cada origen x le corresponde un único destino y.
- Viéndolo en el caso de planes de transporte:
$$d\pi(x, y) = d\pi_T(x, y) \equiv d\mu(x)\delta[y = T(x)],$$
siendo $T : X \to Y$ una función medible.

Definición

$(\mathbf{Id} \times \mathbf{T})\#\mu$ es la medida de probabilidad en $Z = X \times Y$ que satisface la siguiente propiedad: para cualquier función medible no negativa ζ en Z, se tiene que:

$$\int_{X\times Y} \zeta(x, y) d\pi_T(x, y) = \int_X \zeta(x, T(x)) d\mu(x).$$

Medidas imagen / Problema de Monge

- En particular, el coste total de transporte asociado es:

$$I[\pi_T] = \int_X c(x, T(x))d\mu(x)$$

- **¿Cuál es la condición que debe satisfacer T para que π_T pertenezca a $\Pi(\mu, \nu)$?**

Usando la definición anterior, y la expresión (2), se tiene que:

$$\int_X [\varphi(x) + \psi \circ T(x)]d\mu(x) = \int_X \varphi(x)d\mu(x) + \int_Y \psi(y)d\nu(y).$$

$$\int_X (\psi \circ T)d\mu = \int_Y \psi d\nu.$$

- Para toda $\psi \in \mathcal{L}^1(d\nu)$, la función medible $\psi \circ T$ ha de pertenecer a $\mathcal{L}^1(d\mu)$ y los valores de las integrales han de coincidir.

Medidas imagen / Problema de Monge

- En particular, el coste total de transporte asociado es:

$$I[\pi_T] = \int_X c(x, T(x)) d\mu(x)$$

- **¿Cuál es la condición que debe satisfacer T para que π_T pertenezca a $\Pi(\mu, \nu)$?**

Usando la definición anterior, y la expresión (2), se tiene que:

$$\int_X [\varphi(x) + \psi \circ T(x)] d\mu(x) = \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y).$$

$$\int_X (\psi \circ T) d\mu = \int_Y \psi d\nu.$$

- Para toda $\psi \in \mathcal{L}^1(d\nu)$, la función medible $\psi \circ T$ ha de pertenecer a $\mathcal{L}^1(d\mu)$ y los valores de las integrales han de coincidir.

Medidas imagen / Problema de Monge

- En particular, el coste total de transporte asociado es:

$$I[\pi_T] = \int_X c(x, T(x))d\mu(x)$$

- **¿Cuál es la condición que debe satisfacer T para que π_T pertenezca a $\Pi(\mu, \nu)$?**

Usando la definición anterior, y la expresión (2), se tiene que:

$$\int_X [\varphi(x) + \psi \circ T(x)]d\mu(x) = \int_X \varphi(x)d\mu(x) + \int_Y \psi(y)d\nu(y).$$

$$\int_X (\psi \circ T)d\mu = \int_Y \psi d\nu.$$

- Para toda $\psi \in \mathcal{L}^1(d\nu)$, la función medible $\psi \circ T$ ha de pertenecer a $\mathcal{L}^1(d\mu)$ y los valores de las integrales han de coincidir.

Medidas imagen / Problema de Monge

- En particular, el coste total de transporte asociado es:

$$I[\pi_T] = \int_X c(x, T(x))d\mu(x)$$

- **¿Cuál es la condición que debe satisfacer T para que π_T pertenezca a $\Pi(\mu, \nu)$?**

Usando la definición anterior, y la expresión (2), se tiene que:

$$\int_X [\varphi(x) + \psi \circ T(x)]d\mu(x) = \int_X \varphi(x)d\mu(x) + \int_Y \psi(y)d\nu(y).$$

$$\int_X (\psi \circ T)d\mu = \int_Y \psi d\nu.$$

- Para toda $\psi \in \mathcal{L}^1(d\nu)$, la función medible $\psi \circ T$ ha de pertenecer a $\mathcal{L}^1(d\mu)$ y los valores de las integrales han de coincidir.

Medidas imagen / Problema de Monge

- Equivalentemente, en términos de conjuntos medibles, la condición de que π_T pertenezca a $\Pi(\mu, \nu)$ es la siguiente:

Para cualquier conjunto medible $B \subset Y$, $\nu[B] = \mu[T^{-1}(B)]$.

Definición

Si se verifican cualesquiera de las dos condiciones equivalentes, escribiremos:

$$\nu = \mathbf{T}\#\mu$$

y diremos que ν es el **pushforward** *o la* **medida imagen** *de μ por T. También se puede decir que T* **transporta** *a μ hasta ν.*

Medidas imagen / Problema de Monge

Proposición

Sean X e Y espacios completamente regulares, $T : X \to Y$ una aplicación sobreyectiva continua y sean $\mu \in \mathcal{P}(X)$ y $\nu \in \mathcal{P}(Y)$. Entonces, las siguientes afirmaciones son equivalentes:

1. $\nu = T\#\mu$.
2. $\int_Y f d\nu = \int_X f \circ T d\mu$, *para toda función* $f \in \mathcal{C}_b(Y)$.
3. *Para toda función de Borel* $f : Y \to [-\infty, \infty]$ *en* $\mathcal{L}^1(d\nu)$, *tenemos que* $f \circ T \in \mathcal{L}^1(d\mu)$ *y que* $\int_X f \circ T d\mu = \int_Y f d\nu$.

Demostración

- $1. \Rightarrow 3.$ *Es consecuencia elemental de la definición de* $T\#\mu$.
- $3. \Rightarrow 2.$ *Trivial.*

Medidas imagen / Problema de Monge

Demostración

- $2. \Rightarrow 1.$ *Sea* G *un abierto de* Y. *Entonces, su función indicatriz,* 1_G *es inferiormente semicontinua. Al ser* Y *completamente regular, existe una red creciente* (f_α) *en* $\mathcal{C}_b^+(Y)$ *que converge puntualmente a* 1_G. *Pasando al límite la expresión*

$$\int_Y f_\alpha d\nu = \int_X f_\alpha \circ T d\mu$$

y recordando que $1_G \circ T = 1_{T^{-1}(G)}$, *obtenemos:*

$$\nu(G) = \mu(T^{-1}(G)) = T\#\mu(G).$$

Esto prueba que ν *y* $T\#\mu$ *coinciden en conjuntos abiertos y, así, en conjuntos de Borel, luego, son iguales.*

Medidas imagen / Problema de Monge

- Ahora, vamos a considerar X e Y dos espacios completamente regulares. pr_X y pr_Y serán las proyecciones naturales de Z a X e Y y dadas dos funciones $u : X \to (-\infty, \infty]$, $v : Y \to (-\infty, \infty]$, $u \oplus v$ será la función $(x, y) \mapsto u(x) + v(y)$.

Corolario

Supongamos que $\pi \in \mathcal{P}(Z)$. Son equivalentes:

1. $\pi \in \Pi(\mu, \nu)$.
2. *Para todo $A \in \mathcal{B}(X)$ y $B \in \mathcal{B}(Y)$ tenemos que $\pi(A \times Y) = \mu(A)$ y $\pi(X \times B) = \nu(B)$.*
3. *Para todo $u \in \mathcal{C}_b(Y)$, se tiene que $\pi(u \oplus v) = \mu(u) + \nu(v)$.*

Medidas imagen / Problema de Monge

- Ahora, vamos a considerar X e Y dos espacios completamente regulares. pr_X y pr_Y serán las proyecciones naturales de Z a X e Y y dadas dos funciones $u : X \to (-\infty, \infty]$, $v : Y \to (-\infty, \infty]$, $u \oplus v$ será la función $(x, y) \mapsto u(x) + v(y)$.

Corolario

Supongamos que $\pi \in \mathcal{P}(Z)$. Son equivalentes:

1. $\pi \in \Pi(\mu, \nu)$.
2. *Para todo $A \in \mathcal{B}(X)$ y $B \in \mathcal{B}(Y)$ tenemos que $\pi(A \times Y) = \mu(A)$ y $\pi(X \times B) = \nu(B)$.*
3. *Para todo $u \in \mathcal{C}_b(Y)$, se tiene que $\pi(u \oplus v) = \mu(u) + \nu(v)$.*

Medidas imagen / Problema de Monge

Teorema

El conjunto $\Pi(\mu,\nu)$ es un subconjunto de $\mathcal{P}(Z)$ no vacío, convexo y $\sigma(\mathcal{M}_b(Z),\mathcal{C}_b(Z))$-compacto.

Demostración

Sea $\epsilon > 0$. Como μ y ν son medidas de Radon, son apretadas (tight), es decir, existen $K \in \mathcal{K}(X)$ y $L \in \mathcal{K}(Y)$ tales que

$$\mu(X \setminus K) < \tfrac{\epsilon}{2}; \;\; \nu(Y \setminus L) < \tfrac{\epsilon}{2}.$$

Entonces, para todo $\pi \in \Pi(\mu,\nu)$, tenemos:
$\pi(Z \setminus (K \times L)) \leq \pi((X \times K) \times Y) + \pi(X \times (Y \setminus L)) =$
$\mu(X \setminus K) + \nu(Y \setminus L) < \epsilon$
Luego, $\Pi(\mu,\nu)$ es uniformemente apretado y, por el Teorema de Prokhorov, es relativamente $\sigma(\mathcal{M}_b(Z),\mathcal{C}_b(Z))$-compacto.

Medidas imagen / Problema de Monge

Demostración

Probamos ahora que $\Pi(\mu,\nu)$ *es cerrado con la convergencia débil:*

Sea ρ *un elemento de la clausura débil de* $\Pi(\mu,\nu)$ *y sea* (ρ_α) *una red en* $\Pi(\mu,\nu)$ *que converge débilmente a* ρ. *Entonces:*

$$\rho \in \mathcal{P}(Z); \quad lim_\alpha \rho_\alpha(f) = \rho(f), \forall f \in \mathcal{C}_b(Z)$$

En particular, $lim_\alpha \rho_\alpha(u \circ pr_X) = \rho(u \circ pr_X)$, *para todo* $u \in \mathcal{C}_b(X)$. *Pero, para todo* $u \in \mathcal{C}_b(X)$ *y todo* α*:*
$\rho_\alpha(u \circ pr_X) = \mu(u)$ *y, por tanto,* $\rho(u \circ pr_X) = \mu(u)$.
Así, $pr_X \# \rho = \mu$ *y* $pr_Y \# \rho = \nu$, *luego,* $\rho \in \Pi(\mu,\nu)$ *y* $\Pi(\mu,\nu)$ *es débilmente cerrado.*
Con todo esto, hemos demostrado que $\Pi(\mu,\nu)$ *es débilmente compacto.*

Problema de transporte óptimo de Monge

Problema de Monge

$$Minimizar \quad I[T] = \int_X c(x, T(x))d\mu(x)$$

sobre el conjunto de aplicaciones medibles T *tal que* $T\#\mu = \nu$.

Problema de Kantorovich

$$Minimizar \quad I[\pi] = \int_{X\times Y} c(x,y)d\pi(x,y) \qquad para \quad \pi \in \Pi(\mu,\nu)$$

Reformulación de Kantorovich del problema de Monge

Reemplazar $T : X \to Y$ por la medida $0 \leq \pi$ sobre $X \times Y$ cuyas marginales son μ y ν, respectivamente, y entre tales medidas elegimos π que minimice el funcional

$$coste(\pi) = \int_{X \times Y} c(x, y) d\pi(x, y)$$

Existen dos razones por las que este problema es matemáticamente más sencillo que el de Monge:

- El funcional a minimizar no depende linealmente de π.
- El conjunto $\Pi(\mu, \nu)$ es un subconjunto convexo de un espacio de Banach.

En este contexto, se garantiza la existencia de un mínimo π bajo ciertas hipótesis generales sobre c, μ y ν.

Definiciones y resultados básicos

Consideramos los siguientes conjuntos:
$\mathcal{F}(Z)$ ={f: $Z \to [-\infty, \infty)$ superiormente semicontinuas}
$\Phi(f) = \{(u, v)$ funciones Borel tal que $u : X \to (-\infty, \infty]$,
$v : Y \to (-\infty, \infty]$ con $u \in \mathcal{L}^1(\mu)$ y $v \in \mathcal{L}^1(\nu)$ que verifican
$f(x, y) \leq u(x) + v(y)\}$

Definición

$$p(f) = \begin{cases} \inf\{\mu(u) + \nu(v) : (u, v) \in \Phi(f)\} & si \quad \Phi(f) \neq \emptyset \\ \infty & si \quad \Phi(f) = \emptyset \end{cases}$$

- La aplicación $f \mapsto p(f)$ es isótona y p(c)=c para cualquier función c constante.
- $\pi(f) \leq p(f)$ para toda $\pi \in \Pi(\mu, \nu)$ y $f \in \mathcal{F}(Z)$.

Propiedades

Proposición

En $\mathcal{C}_b(Z)$ el funcional $p(\cdot)$ tiene las propiedades:

1. $-\|f\|_\infty \leq inf f \leq p(f) \leq sup f \leq \|f\|_\infty$ *para toda* $f \in \mathcal{C}_b(Z)$.
2. *La aplicación $f \mapsto p(f)$ para $f \in \mathcal{C}_b(Z)$ es sublineal.*
3. $p(u \oplus v) = \mu(u) + \nu(v)$ *para todo* $u \in \mathcal{C}_b(X)$ *y* $v \in \mathcal{C}_b(Y)$.

Demostración:

1. $p(f) \leq sup f$ tomando u(x)= sup f y v(y)=0. El resto obvias.
2. $p(\lambda f) = \lambda p(f)$. $p(f_1 + f_2) \leq \mu(u_1 + u_2) + \nu(v_1 + v_2) = \mu(u_1) + \nu(v_1) + \mu(u_2) + \nu(v_2) = p(f_1) + p(f_2)$.
3. $\mu(u) + \nu(v) = \theta(u \oplus v) \leq p(u \oplus v) \leq \mu(u) + \nu(v)$.

Propiedades

Proposición

En $\mathcal{C}_b(Z)$ el funcional $p(\cdot)$ tiene las propiedades:

1. $-\|f\|_\infty \leq inf f \leq p(f) \leq sup f \leq \|f\|_\infty$ *para toda* $f \in \mathcal{C}_b(Z)$.
2. *La aplicación* $f \mapsto p(f)$ *para* $f \in \mathcal{C}_b(Z)$ *es sublineal.*
3. $p(u \oplus v) = \mu(u) + \nu(v)$ *para todo* $u \in C_b(X)$ *y* $v \in C_b(Y)$.

Demostración:

1. $p(f) \leq sup f$ tomando u(x)= sup f y v(y)=0. El resto obvias.
2. $p(\lambda f) = \lambda p(f)$. $p(f_1 + f_2) \leq \mu(u_1 + u_2) + \nu(v_1 + v_2) = \mu(u_1) + \nu(v_1) + \mu(u_2) + \nu(v_2) = p(f_1) + p(f_2)$.
3. $\mu(u) + \nu(v) = \theta(u \oplus v) \leq p(u \oplus v) \leq \mu(u) + \nu(v)$.

Propiedades

Proposición

En $\mathcal{C}_b(Z)$ el funcional $p(\cdot)$ tiene las propiedades:

1. $-\|f\|_\infty \leq inf f \leq p(f) \leq sup f \leq \|f\|_\infty$ *para toda* $f \in \mathcal{C}_b(Z)$.
2. *La aplicación* $f \mapsto p(f)$ *para* $f \in \mathcal{C}_b(Z)$ *es sublineal.*
3. $p(u \oplus v) = \mu(u) + \nu(v)$ *para todo* $u \in \mathcal{C}_b(X)$ *y* $v \in C_b(Y)$.

Demostración:

1. $p(f) \leq sup f$ tomando u(x)= sup f y v(y)=0. El resto obvias.
2. $p(\lambda f) = \lambda p(f)$. $p(f_1 + f_2) \leq \mu(u_1 + u_2) + \nu(v_1 + v_2) = \mu(u_1) + \nu(v_1) + \mu(u_2) + \nu(v_2) = p(f_1) + p(f_2)$.
3. $\mu(u) + \nu(v) = \theta(u \oplus v) \leq p(u \oplus v) \leq \mu(u) + \nu(v)$.

Resultados

Definición

Sea S un espacio completamente regular, I: $\mathcal{C}_b(Z) \to \mathbb{R}$ un funcional lineal. Se dice que I es **apretado** *si $\forall \epsilon > 0$ existe $Q(\epsilon) \in \mathcal{S}$ tal que $|I(f)| < \epsilon$ para cada $f \in \mathcal{C}_b(Z)$ que satisface:*

- *f se anula en $Q(\epsilon)$.*
- $\|f\|_\infty \leq 1$.

Teorema

Sea S un espacio completamente regular, I: $\mathcal{C}_b(Z) \to \mathbb{R}$ un funcional lineal, apretado y continuo. Entonces existe una única medida $\sigma \in \mathcal{M}_b(S)$ tal que I(f)=$\int_S f d\sigma$ para toda $f \in \mathcal{C}_b(Z)$.

Resultados

Si $\sigma \in \Pi(\mu, \nu)$ y I(f)=$\int_S f d\sigma$ entonces, por lo visto anteriormente, $I(f) \leq p(f)$ para toda $f \in \mathcal{C}_b(Z)$. Existe un resultado recíproco:

Teorema

Sea I: $\mathcal{C}_b(Z) \to \mathbb{R}$ un funcional lineal que verifica $I(f) \leq p(f)$ para toda $f \in \mathcal{C}_b(Z)$. Entonces existe una única medida $\sigma \in \Pi(\mu, \nu)$ tal que I(f)=$\int_S f d\sigma$.

Demostración:

- I es *continua* por ser acotada: $\pm I(f) \leq p(f) \leq \|f\|_\infty$
- I es *apretada*: Tomamos K y L subconjuntos compactos de X e Y respectivamente que cumplan $\mu(X \setminus K) < \epsilon/2$ y $\nu(Y \setminus L) < \epsilon/2$. Sea f tal que se anula en KxL y $\|f\|_\infty \leq 1$, y sean

$$u(x) = \begin{cases} 1 & si \quad x \in X \setminus K \\ 0 & si \quad x \in K \end{cases} \qquad v(y) = \begin{cases} 1 & si \quad y \in Y \setminus L \\ 0 & si \quad y \in L \end{cases}$$

Entonces como $(u, v) \in \Phi(f)$, $I(f) \leq p(f) \leq \mu(u) + \nu(v) < \epsilon$.

- Aplicando el teorema anterior existe $\sigma \in \mathcal{M}_b(S)$ tal que I(f)=$\int_S f d\sigma$.
- $\sigma \in \mathcal{P}(Z)$, y $\sigma \in \Pi(\mu, \nu)$, pues I(f)=$\theta(f)$ para toda f en $\Lambda = \{u \oplus v : u \in \mathcal{C}_b(X), v \in \mathcal{C}_b(Y)\}$

Ejemplos

Masa de Dirac

Supongamos que ν es una masa de Dirac: $\nu = \delta_a$. Entonces hay un único elemento en $\Pi(\mu, \nu)$ (toda la masa debe ser transportada a a) y

$$\mathcal{T}_c(\mu, \delta_a) = \int_X c(x, a) d\mu(x).$$

Además, los problemas de Monge y Kantorovich coinciden, ya que la masa no se divide.

En efecto:

- $\Pi(\mu, \nu) := \{\pi \in P(X \times Y) : \pi(A \times Y) = \mu(A), \pi(X \times B) = \nu(B), \forall A, B$ medibles$\}$
- Si $a \notin B$ entonces

$$\pi(A \times B) \leq \pi(X \times B) = \delta_a(B) = 0.$$

Luego $\pi(A \times B) = 0$.

Ejemplos

Masa de Dirac

Supongamos que ν es una masa de Dirac: $\nu = \delta_a$. Entonces hay un único elemento en $\Pi(\mu, \nu)$ (toda la masa debe ser transportada a a) y

$$\mathcal{T}_c(\mu, \delta_a) = \int_X c(x, a) d\mu(x).$$

Además, los problemas de Monge y Kantorovich coinciden, ya que la masa no se divide.

En efecto:

- $\Pi(\mu, \nu) := \{\pi \in P(X \times Y) : \pi(A \times Y) = \mu(A), \pi(X \times B) = \nu(B), \forall A, B$ medibles$\}$
- Si $a \notin B$ entonces

$$\pi(A \times B) \leq \pi(X \times B) = \delta_a(B) = 0.$$

Luego $\pi(A \times B) = 0$.

Ejemplos

- Si $a \in B$ entonces

$$\pi(A \times B) = \pi(A \times \{a\}) + \pi(A \times (B \setminus \{a\})) = \pi(A \times \{a\}).$$

Luego

$$\pi(A \times B) = \pi(A \times \{a\}) = \pi(A \times Y) = \mu(A).$$

- Por tanto $\pi(A \times B) = \mu(A)\delta_a(B) = (\mu \otimes \delta_a)(A \times B)$.
- Además

$$\begin{aligned} \mathcal{T}_c(\mu, \delta_a) &= \inf_{\pi \in \Pi(\mu, \delta_a)} \mathcal{I}(\pi) = \mathcal{I}(\mu \otimes \delta_a) \\ &= \int_{X \times Y} c(x,y) d(\mu \otimes \delta_a)(x,y) \\ &= \int_X \left(\int_Y c(x,y) d\delta_a(y) \right) d\mu(x) \\ &= \int_X c(x,a) d\mu(x) \end{aligned}$$

Ejemplos

- Si $a \in B$ entonces
$$\pi(A \times B) = \pi(A \times \{a\}) + \pi(A \times (B \setminus \{a\})) = \pi(A \times \{a\}).$$
Luego
$$\pi(A \times B) = \pi(A \times \{a\}) = \pi(A \times Y) = \mu(A).$$
- Por tanto $\pi(A \times B) = \mu(A)\delta_a(B) = (\mu \otimes \delta_a)(A \times B)$.
- Además
$$\begin{aligned}\mathcal{T}_c(\mu, \delta_a) &= \inf_{\pi \in \Pi(\mu,\delta_a)} \mathcal{I}(\pi) = \mathcal{I}(\mu \otimes \delta_a) \\ &= \int_{X \times Y} c(x,y)d(\mu \otimes \delta_a)(x,y) \\ &= \int_X \left(\int_Y c(x,y)d\delta_a(y) \right) d\mu(x) \\ &= \int_X c(x,a)d\mu(x)\end{aligned}$$

Ejemplos

- Si $a \in B$ entonces

$$\pi(A \times B) = \pi(A \times \{a\}) + \pi(A \times (B \setminus \{a\})) = \pi(A \times \{a\}).$$

  Luego

$$\pi(A \times B) = \pi(A \times \{a\}) = \pi(A \times Y) = \mu(A).$$

- Por tanto $\pi(A \times B) = \mu(A)\delta_a(B) = (\mu \otimes \delta_a)(A \times B)$.
- Además

$$\begin{aligned}
\mathcal{T}_c(\mu, \delta_a) &= \inf_{\pi \in \Pi(\mu,\delta_a)} \mathcal{I}(\pi) = \mathcal{I}(\mu \otimes \delta_a) \\
&= \int_{X \times Y} c(x,y) d(\mu \otimes \delta_a)(x,y) \\
&= \int_X \left(\int_Y c(x,y) d\delta_a(y) \right) d\mu(x) \\
&= \int_X c(x,a) d\mu(x)
\end{aligned}$$

Ejemplos

Hipercubo discreto

Sea $C := \{0,1\}^N$ el hipercubo discreto dotado de la métrica de Hamming (cada arista tiene longitud 1). Sean x_0 e y_0 dos vértices contiguos, digamos $x_0 = (0,0,\dots,0)$, $y_0 = (1,0,\dots,0)$. Definimos X e Y como las bolas unidad centradas en x_0 e y_0 respectivamente, y consideramos la masa unidad uniformemente distribuida en X. Nos planteamos el problema de Kantorovich para llevar la masa de X a Y.

Analizaremos dos posibles planes de transporte:

- Medida producto.
- Fijar x_0 e y_0, y mandar el resto de puntos a los vecinos.

Ejemplos

Hipercubo discreto

Sea $C := \{0,1\}^N$ el hipercubo discreto dotado de la métrica de Hamming (cada arista tiene longitud 1). Sean x_0 e y_0 dos vértices contiguos, digamos $x_0 = (0,0,\dots,0)$, $y_0 = (1,0,\dots,0)$. Definimos X e Y como las bolas unidad centradas en x_0 e y_0 respectivamente, y consideramos la masa unidad uniformemente distribuida en X. Nos planteamos el problema de Kantorovich para llevar la masa de X a Y.

Analizaremos dos posibles planes de transporte:

- Medida producto.
- Fijar x_0 e y_0, y mandar el resto de puntos a los vecinos.

Ejemplos

- Medida producto

$$
\begin{aligned}
I[\mu \otimes \nu] &= \int_{X \times Y} d(x,y) d(\mu \otimes \nu)(x,y) \\
&= \sum_{x_i \in X, y_j \in Y} d(x_i, y_j)(\mu \otimes \nu)(\{(x_i, y_j)\}) \\
&= \sum_{x_i \in X, y_j \in Y} d(x_i, y_j)\mu(\{x_i\})\nu(\{y_j\}) \\
&= \frac{1}{(N+1)^2} \sum_{x_i \in X} \sum_{y_j \in Y} d(x_i, y_j) \\
&= \frac{1}{(N+1)^2} \\
&\left([0+1+2(N-1)] + [N] + [(N-1)(2+2+3(N-2))]\right) \\
&= \frac{3N^2 - 2N + 1}{(N+1)^2}
\end{aligned}
$$

Ejemplos

- Fijar x_0 e y_0, y mandar el resto de puntos a los vecinos

Es decir, consideramos la aplicación

$$T : X \to Y$$

dada por

$$T(x_0) = x_0, \quad T(y_0) = y_0, \quad T(x_i) = y_i,$$

donde $x_i = (0, \ldots, 1^{(i}, \ldots, 0)$, $y_i = (1, \ldots, 1^{(i}, \ldots, 0)$, para $i = 2, \ldots, N$.

Así

$$I[\pi] = \int_X d(x, Tx) d\mu(x) = \sum_{i=2}^{N} 1 \cdot \mu(\{x_i\}) = \frac{N-2}{N+1}.$$

Ejemplos

Observaciones

- Notar que $\frac{3N^2-2N+1}{(N+1)^2} > \frac{N-2}{N+1} \quad \forall N \in \mathbb{N}$.
- Veremos más adelante que el último transporte es optimal usando el Teorema de dualidad de Kantorovich en su versión Kantorovich-Rubinstein.

Ejemplos

Observaciones

- Notar que $\frac{3N^2-2N+1}{(N+1)^2} > \frac{N-2}{N+1} \quad \forall N \in \mathbb{N}$.
- Veremos más adelante que el último transporte es optimal usando el Teorema de dualidad de Kantorovich en su versión Kantorovich-Rubinstein.

Principio de Dualidad de Kantorovich

Sean X, Y espacios topológicos completamente regulares.

Teorema

Sea $h : Z \to [-\infty, \infty)$ una función superiormente semicontinua. Supongamos que existen funciones reales inf. semicontinuas $a \in \mathcal{L}^1(\mu)$ y $b \in \mathcal{L}^1(\nu)$ tales que $h(x,y) \leq a(x) + b(y)$ $\forall (x,y) \in Z$. Entonces,

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \inf\{\mu(u) + \nu(v) : (u,v) \in \Phi(h)\}$$

(No se excluye el caso en que ambos términos sean $-\infty$.)

Principio de Dualidad de Kantorovich

Sean X, Y espacios topológicos completamente regulares.

Teorema

Sea $h : Z \to [-\infty, \infty)$ una función superiormente semicontinua. Supongamos que existen funciones reales inf. semicontinuas $a \in \mathcal{L}^1(\mu)$ y $b \in \mathcal{L}^1(\nu)$ tales que $h(x,y) \leq a(x) + b(y)\ \forall (x,y) \in Z$. Entonces,

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \inf\{\mu(u) + \nu(v) : (u,v) \in \Phi(h)\}$$

(No se excluye el caso en que ambos términos sean $-\infty$.)

Principio de Dualidad de Kantorovich

Sean X, Y espacios topológicos completamente regulares.

Teorema

Sea $h : Z \to [-\infty, \infty)$ *una función superiormente semicontinua. Supongamos que existen funciones reales inf. semicontinuas* $a \in \mathcal{L}^1(\mu)$ *y* $b \in \mathcal{L}^1(\nu)$ *tales que* $h(x, y) \leq a(x) + b(y)\ \forall (x, y) \in Z$. *Entonces,*

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = inf\{\mu(u) + \nu(v) : (u, v) \in \Phi(h)\}$$

(No se excluye el caso en que ambos términos sean $-\infty$*.)*

En la literatura de transferencia óptima es usual reformular el teorema anterior en términos de $c(x,y) = -h(x,y)$, la *función coste*.

$$\pi(h) = -\pi(c) \Leftrightarrow \int_Z hd\pi = -\int_Z cd\pi$$

Entonces, el coste total del plan de transferencia π es $\int_Z cd\pi$ y el teorema anterior evalúa el coste total mínimo posible, denotándolo $\underset{\pi\in\Pi(\mu,\nu)}{min} \int_Z cd\pi$.

$$\underset{\pi\in\Pi(\mu,\nu)}{min} \pi(c) = sup\{\mu(u)+\nu(v) : (u,v)\in\Phi(c)\}$$

En la literatura de transferencia óptima es usual reformular el teorema anterior en términos de $c(x,y) = -h(x,y)$, la *función coste*.

$$\pi(h) = -\pi(c) \Leftrightarrow \int_Z h d\pi = -\int_Z c d\pi$$

Entonces, el coste total del plan de transferencia π es $\int_Z c d\pi$ y el teorema anterior evalúa el coste total mínimo posible, denotándolo $\underset{\pi\in\Pi(\mu,\nu)}{min} \int_Z c d\pi$.

$$\underset{\pi\in\Pi(\mu,\nu)}{min} \pi(c) = sup\{\mu(u) + \nu(v) : (u,v) \in \Phi(c)\}$$

En la literatura de transferencia óptima es usual reformular el teorema anterior en términos de $c(x,y) = -h(x,y)$, la *función coste*.

$$\pi(h) = -\pi(c) \Leftrightarrow \int_Z h d\pi = -\int_Z c d\pi$$

Entonces, el coste total del plan de transferencia π es $\int_Z c d\pi$ y el teorema anterior evalúa el coste total mínimo posible, denotándolo $\underset{\pi\in\Pi(\mu,\nu)}{min} \int_Z c d\pi$.

$$\underset{\pi\in\Pi(\mu,\nu)}{min} \pi(c) = sup\{\mu(u) + \nu(v) : (u,v) \in \Phi(c)\}$$

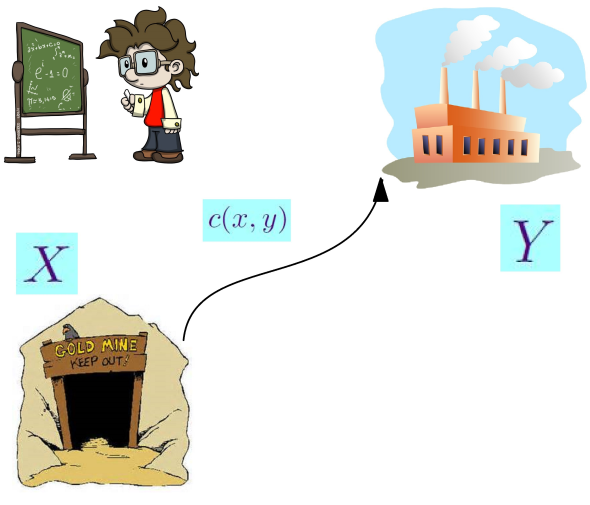
En la literatura de transferencia óptima es usual reformular el teorema anterior en términos de $c(x,y) = -h(x,y)$, la *función coste*.

$$\pi(h) = -\pi(c) \Leftrightarrow \int_Z h d\pi = -\int_Z c d\pi$$

Entonces, el coste total del plan de transferencia π es $\int_Z c d\pi$ y el teorema anterior evalúa el coste total mínimo posible, denotándolo $\underset{\pi \in \Pi(\mu,\nu)}{min} \int_Z c d\pi$.

$$\underset{\pi \in \Pi(\mu,\nu)}{min} \pi(c) = sup\{\mu(u) + \nu(v) : (u,v) \in \Phi(c)\}$$

Problema del Comerciante de Caffarelli



Problema del Comerciante de Caffarelli

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Demostración

Hacemos la demostración en tres pasos.

- *Caso:* $h \in \mathcal{C}_b(Z)$

Consideramos $q : Lin\{h\} \to \mathbb{R}$ *tal que* $q(\lambda h) = \lambda p(h)$

$$\Rightarrow q(f) \leq p(f), \forall f \in Lin\{h\}$$

Por Hahn-Banach, $\exists I : \mathcal{C}_b(Z) \to \mathbb{R}$ *funcional lineal tal que*

$$I(f) \leq p(f) \forall f \in \mathcal{C}_b(Z), I(h) = p(h)$$

Entonces, $\exists \sigma \in \Pi(\mu, \nu)$ *tal que* $I(f) = \sigma(f) \forall f \in \mathcal{C}_b(Z)$.
Expresamos $I(h) = p(h)$ *como* $\sigma(h) = p(h)$.

Principio de Dualidad de Kantorovich

Como $\pi(f) \leq p(f) \quad \forall \pi \in \Pi(\mu,\nu) \quad y \quad \forall f \in \mathcal{F}(Z)$,

$$\pi(h) \leq p(h) \quad \forall \pi \in \Pi(\mu,\nu)$$

Por tanto,

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \sigma(h) = p(h) = inf\{\mu(u) + \nu(v) : (u,v) \in \Phi(h)\}$$

Principio de Dualidad de Kantorovich

Como $\pi(f) \leq p(f) \quad \forall \pi \in \Pi(\mu, \nu) \quad y \quad \forall f \in \mathcal{F}(Z)$,

$$\pi(h) \leq p(h) \quad \forall \pi \in \Pi(\mu, \nu)$$

Por tanto,

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \sigma(h) = p(h) = inf\{\mu(u) + \nu(v) : (u, v) \in \Phi(h)\}$$

- Probado ya el **Teorema de dualidad de Kantorovich** para $f \in \mathcal{C}_b(Z)$ queda generalizarlo para el caso de superiormente semicontinuas.
- Para abordar dicha discusión son necesarios dos Lemas previos...

Lema

Sea K compacto y $\{g_\alpha\}_{\alpha\in A}$ una red decreciente en $\mathcal{C}_b(K)$ con límite puntual $g : K \to [-\infty, \infty[$ entonces se verifica

$$\lim_\alpha \sup_{\omega\in K} g_\alpha(\omega) = \inf_\alpha \sup_{\omega\in K} g_\alpha(\omega) = \sup_{\omega\in K} g(\omega)$$

Solo es necesario probar que $\sup_{\omega\in K} g(\omega) = \inf_\alpha \sup_{\omega\in K} g_\alpha(\omega)$

La prueba se sirve de una red de cerrados $F(\alpha) = \{\omega \in K : g_\alpha(\omega) \geq \Delta\}$ donde Δ es un real verificando $\sup_{\omega\in K} g(\omega) < \Delta$

Si todo $F(\alpha)$ fuese no vacío gastando la caracterización de compacidad de K por la PIF y la cofinalidad del orden del conjunto dirigido llegariamos a que $g_\alpha(\omega_0) \leq \Delta \;\; \forall\alpha \in A$ lo que contradiría la convergencia puntual de la sucesión.

Lema

Sea K compacto y $\{g_\alpha\}_{\alpha\in A}$ una red decreciente en $\mathcal{C}_b(K)$ con límite puntual $g : K \to [-\infty, \infty[$ entonces se verifica

$$\lim_{\alpha} \sup_{\omega\in K} g_\alpha(\omega) = \inf_{\alpha} \sup_{\omega\in K} g_\alpha(\omega) = \sup_{\omega\in K} g(\omega)$$

Solo es necesario probar que $\sup_{\omega\in K} g(\omega) = \inf_\alpha \sup_{\omega\in K} g_\alpha(\omega)$
La prueba se sirve de una red de cerrados
$F(\alpha) = \{\omega \in K : g_\alpha(\omega) \geq \Delta\}$ donde Δ es un real verificando $\sup_{\omega\in K} g(\omega) < \Delta$
Si todo $F(\alpha)$ fuese no vacío gastando la caracterización de compacidad de K por la PIF y la cofinalidad del orden del conjunto dirigido llegariamos a que $g_\alpha(\omega_0) \leq \Delta \ \ \forall\alpha \in A$ lo que contradiría la convergencia puntual de la sucesión.

Lema

Sea K compacto y $\{g_\alpha\}_{\alpha \in A}$ una red decreciente en $\mathcal{C}_b(K)$ con límite puntual $g : K \to [-\infty, \infty[$ entonces se verifica

$$\lim_{\alpha} \sup_{\omega \in K} g_\alpha(\omega) = \inf_{\alpha} \sup_{\omega \in K} g_\alpha(\omega) = \sup_{\omega \in K} g(\omega)$$

Solo es necesario probar que $\sup_{\omega \in K} g(\omega) = \inf_\alpha \sup_{\omega \in K} g_\alpha(\omega)$
La prueba se sirve de una red de cerrados
$F(\alpha) = \{\omega \in K \ : g_\alpha(\omega) \geq \Delta\}$ donde Δ es un real verificando $\sup_{\omega \in K} g(\omega) < \Delta$
Si todo $F(\alpha)$ fuese no vacío gastando la caracterización de compacidad de K por la PIF y la cofinalidad del orden del conjunto dirigido llegariamos a que $g_\alpha(\omega_0) \leq \Delta \ \ \forall \alpha \in A$ lo que contradiría la convergencia puntual de la sucesión.

Lema

Sea $z_0 \in Z$ *y* t *una constante real tal que* $h(z_0) < t$. *Entonces existe* $g \in \mathcal{C}_b(Z)$ *verificando* $h \leq g \leq 0$ *y* $g(z_0) < t$

- Notemos que el Lema anterior implica que es posible construir una red $\{h_\alpha\}$ decreciente en $\mathcal{C}_b(Z)$ que converge puntualmente a h verificando $h_\alpha \leq 0 \ \forall \alpha$
- En particular, dada $\pi \in \Pi(\mu, \nu)$ se tiene que $\{\pi(h_\alpha)\}$ es decreciente y convergente a $\pi(h)$
- Hagamos la identificación de h_α y h como elementos del bidual. Es decir consideremos $\hat{h}_\alpha$ y $\hat{h}$ verificando $\hat{h}_\alpha(\pi) = \pi(h_\alpha)$ y $\hat{h}(\pi) = \pi(h)$ para toda $\pi \in \Pi(\mu, \nu)$

- Notemos que el Lema anterior implica que es posible construir una red $\{h_\alpha\}$ decreciente en $\mathcal{C}_b(Z)$ que converge puntualmente a h verificando $h_\alpha \leq 0 \; \forall \alpha$
- En particular, dada $\pi \in \Pi(\mu, \nu)$ se tiene que $\{\pi(h_\alpha)\}$ es decreciente y convergente a $\pi(h)$
- Hagamos la identificación de h_α y h como elementos del bidual. Es decir consideremos $\hat{h}_\alpha$ y $\hat{h}$ verificando $\hat{h}_\alpha(\pi) = \pi(h_\alpha)$ y $\hat{h}(\pi) = \pi(h)$ para toda $\pi \in \Pi(\mu, \nu)$

- Notemos que el Lema anterior implica que es posible construir una red $\{h_\alpha\}$ decreciente en $\mathcal{C}_b(Z)$ que converge puntualmente a h verificando $h_\alpha \leq 0 \; \forall \alpha$
- En particular, dada $\pi \in \Pi(\mu, \nu)$ se tiene que $\{\pi(h_\alpha)\}$ es decreciente y convergente a $\pi(h)$
- Hagamos la identificación de h_α y h como elementos del bidual. Es decir consideremos $\hat{h}_\alpha$ y $\hat{h}$ verificando $\hat{h}_\alpha(\pi) = \pi(h_\alpha)$ y $\hat{h}(\pi) = \pi(h)$ para toda $\pi \in \Pi(\mu, \nu)$

Probemos el teorema de dualidad para funciones superiormente semicontinuas y negativas

- Si recordamos $\hat{h}_\alpha(\pi) = \pi(h_\alpha)$ y $\hat{h}(\pi) = \pi(h) \ \ \forall \pi \in \Pi(\mu, nu)$
- Entonces tenemos que $\hat{h}_\alpha \xrightarrow{\sigma} \hat{h}$ y como $\Pi(\mu, \nu)$ es $\sigma(\mathcal{M}_b(Z), \mathcal{C}_b(Z))$-compacto estamos en condiciones de aplicar el Lema (1).

$$\lim_\alpha \max_{\pi \in \Pi(\mu, nu)} \hat{h}_\alpha(\pi) = \inf \max_{\pi \in \Pi(\mu,\nu)} \hat{h}_\alpha(\pi) = \max_{\pi \in \Pi(\mu,\nu)} \hat{h}(\pi)$$

Ahora $\hat{h}(\pi) \leq p(h) \leq p(h_\alpha) = \max_{\rho\in\Pi(\mu,\nu)} \hat{h}_\alpha(\rho)$[1]

Tomando máximos e ínfimos y aplicando el lema tenemos que

$$\max_{\pi\in\Pi(\mu,\nu)} \hat{h}(\pi) \leq p(h) \leq \inf_\alpha p(h_\alpha) = \inf_\alpha \max_{\rho\in\Pi(\mu,\nu)} \hat{h}_\alpha(\rho) = \max_{\rho\in\Pi(\mu,\nu)} \hat{h}(\rho)$$

Por lo que se da la igualdad y

$$\max_{\pi\in\Pi(\mu,\nu)} \pi(h) = p(h) = \inf \{\mu(u) + \nu(v) \ : \ (u,v) \in \Phi(h)\}$$

[1] En la igualdad se hace uso del caso anterior porque $h_\alpha \in \mathcal{C}_b(Z)$

- Recordar que por ahora tenemos probado que

$$\max_{\pi\in\Pi(\mu,\nu)} \pi(h) = \inf\{\mu(u)+\nu(v) \ : \ (u,v)\in\Phi(h)\}$$

siendo h superiormente semicontinua y negativa.

- Reduzcamos el caso general en que h superiormente semicontinua y $h(x,y)\leq a(x)+b(y) \ \ a\in\mathcal{L}^1(\mu), b\in\mathcal{L}^1(\nu)$ al caso anterior.

Definamos $k = h - a\oplus b$ que es superiormente semicontinua y negativa. Por tanto

$$\max_{\pi\in\Pi(\mu,\nu)} \pi(k) = p(k)$$

Ahora bien, si $\tilde{u},\tilde{v}$ verifican que $k\leq \tilde{u}\oplus\tilde{v}$ entonces

$$p(h) = \inf\{\mu(\tilde{u})+\nu(\tilde{v}) : \ (\tilde{u},\tilde{v})\in\Phi(k)\} + \mu(a)+\nu(b)$$

además

$$\pi(h) = \pi(k) + \pi(a \oplus b) = \pi(k) + \mu(a) + \nu(b)$$

Y finalmente

$$\begin{aligned} \max_{\pi \in \Pi(\mu,\nu)} \pi(h) &= \max_{\pi \in \Pi(\mu,\nu)} \pi(k) + \mu(a) + \nu(b) \\ &= p(k) + \mu(a) + \nu(b) = p(h) \\ &= \inf\{\mu(u) + \nu(v) : \ (u,v) \in \Phi(h)\} \end{aligned}$$

Si consideramos el caso particular en que X, Y son espacios métricos podemos obtener un resultado más fino al anterior.

- El teorema de Kantorovich para espacios $T_{3\frac{1}{2}}$ afirmaba que

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \inf\{\mu(u) + \nu((v)) : \ (u, v) \in \Phi(h)\}$$

  donde la relación $(u, v) \in \Phi(h)$ significaba que $u \in \mathcal{L}^1(\mu), v \in \mathcal{L}^1(\nu)$ y que $h(x, y) \leq u(x) + v(y)$

- El teorema para espacios métricos se enuncia igual imponiendo que la función superiormente semicontinua sea acotada dando como resultado

$$\max_{\pi \in \Pi(\mu,\nu)} \pi(h) = \inf\{\mu(u) + \nu((v)) : \ (u, v) \in \Psi(h)\}$$

  donde $(u, v) \in \Psi(h)$ significa que $u \in \mathcal{C}_{boun}(X), v \in \mathcal{C}_{boun}(Y)$ y $h(x, y) \leq u(x) + v(y)$ donde $\mathcal{C}_{boun}(X)$ simboliza las funciones acotadas y uniformemente continuas en X.

El teorema de Kantorovich-Rubinstein

Teorema

Sea $X = Y$ un espacio polaco, y sea d una métrica inferiormente semicontinua sobre X. Sea $\mathcal{T}_d$ el coste de transporte optimal para el coste $c(x, y) = d(x, y)$,

$$\mathcal{T}_d(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y) d\pi(x, y).$$

Denotemos por Lip(X) el espacio de todas las funciones Lipschitz sobre X y

$$||\varphi||_{\text{Lip}} := \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)}.$$

Entonces

$$\mathcal{T}_d(\mu, \nu) = \sup \left(\int_X \varphi d(\mu - \nu) : \varphi \in L^1(d|\mu - \nu|), ||\varphi||_{\text{Lip}} \leq 1 \right).$$

En el caso del hipercubo, tenemos:

- $X = Y = C$ es el hipercubo
- d es la métrica de Hamming
- μ, ν son las medidas uniformes de antes
- Sea $\varphi \in \mathrm{Lip}(C), ||\varphi||_{\mathrm{Lip}} \leq 1$

$$\begin{aligned}
\int_C \varphi d(\mu - \nu) &= \frac{1}{N+1} \sum_{i=2}^{N} (\varphi(y_i) - \varphi(x_i)) \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} |\varphi(x_i) - \varphi(y_i)| \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} d(x_i, y_i) \\
&= \frac{N-2}{N+1}
\end{aligned}$$

En el caso del hipercubo, tenemos:

- $X = Y = C$ es el hipercubo
- d es la métrica de Hamming
- μ, ν son las medidas uniformes de antes
- Sea $\varphi \in \mathrm{Lip}(C), ||\varphi||_{\mathrm{Lip}} \leq 1$

$$\begin{aligned}
\int_C \varphi d(\mu - \nu) &= \frac{1}{N+1} \sum_{i=2}^{N} (\varphi(y_i) - \varphi(x_i)) \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} |\varphi(x_i) - \varphi(y_i)| \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} d(x_i, y_i) \\
&= \frac{N-2}{N+1}
\end{aligned}$$

En el caso del hipercubo, tenemos:

- $X = Y = C$ es el hipercubo
- d es la métrica de Hamming
- μ, ν son las medidas uniformes de antes
- Sea $\varphi \in \text{Lip}(C), ||\varphi||_{\text{Lip}} \leq 1$

$$
\begin{aligned}
\int_C \varphi d(\mu - \nu) &= \frac{1}{N+1} \sum_{i=2}^{N} (\varphi(y_i) - \varphi(x_i)) \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} |\varphi(x_i) - \varphi(y_i)| \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} d(x_i, y_i) \\
&= \frac{N-2}{N+1}
\end{aligned}
$$

En el caso del hipercubo, tenemos:

- $X = Y = C$ es el hipercubo
- d es la métrica de Hamming
- μ, ν son las medidas uniformes de antes
- Sea $\varphi \in \text{Lip}(C)$, $||\varphi||_{\text{Lip}} \leq 1$

$$
\begin{aligned}
\int_C \varphi d(\mu - \nu) &= \frac{1}{N+1} \sum_{i=2}^{N} (\varphi(y_i) - \varphi(x_i)) \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} |\varphi(x_i) - \varphi(y_i)| \\
&\leq \frac{1}{N+1} \sum_{i=2}^{N} d(x_i, y_i) \\
&= \frac{N-2}{N+1}
\end{aligned}
$$

- Sea $\varphi : C \to C$ la función que manda cada punto a su primera coordenada. Entonces, φ es Lipschitz con $||\varphi||_{\text{Lip}} = 1$ y, además,

$$\int_C \varphi d(\mu - \nu) = \frac{N-2}{N+1}$$

- Luego $\mathcal{T}_d(\mu, \nu) = \frac{N-2}{N+1}$

- Sea $\varphi : C \to C$ la función que manda cada punto a su primera coordenada. Entonces, φ es Lipschitz con $||\varphi||_{\text{Lip}} = 1$ y, además,
$$\int_C \varphi d(\mu - \nu) = \frac{N-2}{N+1}$$
- Luego $\mathcal{T}_d(\mu, \nu) = \frac{N-2}{N+1}$

¡GRACIAS POR SU ATENCIÓN!