



Departamento de Matemáticas, Facultad de Ciencias
Universidad Autónoma de Madrid

Introducción al transporte óptimo

TRABAJO DE FIN DE GRADO

Grado en Matemáticas

Autor: Judith Manchón

Tutor: Daniel Faraco

Curso 2019-2020

Resumen

Abstract

Índice general

1	Problema de Monge	1
2	Problema de Kantorovich	3
3	Problema Dual	5
3.1	Caso Discreto	7
3.2	Caso General	8
4	Existencia para coste cuadrático	11
4.1	Geometría del transporte óptimo	11
4.2	Solución del problema dual	15
4.3	Existencia de solución del problema de transporte óptimo	17
5	Aplicación	21
5.1	Definiciones previas	21
5.2	Existencia y unicidad de un mínimo	22
5.3	Ecuación de Euler-Lagrange	24
5.4	Convergencia	24

CAPÍTULO 1

Problema de Monge

El problema original propuesto por Monge buscaba transportar de manera óptima una cantidad de arena a un hueco del mismo volumen.

En el lenguaje moderno diríamos que, tenemos dos medidas de Radón no negativas μ y ν , y los espacios de medias (X, μ) y (Y, ν) en \mathbb{R}^n , que cumplen la siguiente condición:

$$(1.1) \quad \mu(\mathbb{R}^n) = \nu(\mathbb{R}^n) < \infty$$

y consideramos que existe la función $s : \mathbb{R}^n \rightarrow \mathbb{R}^n$ que lleva μ a ν del siguiente modo:

$$(1.2) \quad s_{\#}(\mu) = (\nu)$$

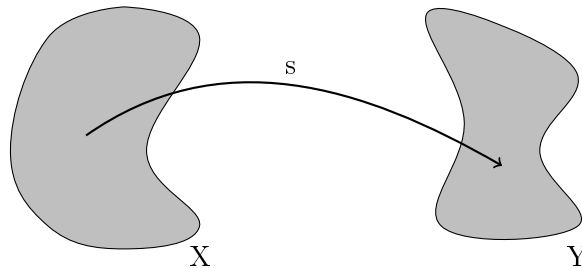
Es decir,

$$(1.3) \quad \int_X h(s(x))d\mu(x) = \int_Y h(y)d\nu(y)$$

para toda función continua h . Llamamos \mathcal{A} al conjunto de todas las funciones s que cumplen (1.2) y (1.3).

Además hay que tener en cuenta el coste del transporte desde el punto $x \in X$ al punto $y \in Y$, para ello tenemos la función de coste

$$c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty)$$



El coste total de transportar el conjunto X dada la función $s \in \mathcal{A}$ es

$$(1.4) \quad I[s] := \int_X c(x, s(x))d\mu(x)$$

Por último, solo nos queda encontrar función s^* que minimiza dicho coste

$$(1.5) \quad I[s^*] := \min_{s \in \mathcal{A}} M[s]$$

Supongamos ahora que nuestras medidas μ y ν son absolutamente continuas respecto a la medida de Lebesgue, entonces

$$d\mu(x) = f(x)dx$$

$$d\nu(y) = g(y)dy$$

si aplicamos esto a la condición (1.3) obtenemos la siguiente relación

$$(1.6) \quad f(x) = g(s(x))\det(D_s(x))$$

donde D_s es la matriz jacobiana de s , definida del siguiente modo:
sea $s = (s^1, s^2, \dots, s^n)$, s^i las componentes de s ,

$$D_s = \begin{pmatrix} s_{x_1}^1 & \cdots & s_{x_n}^1 \\ \vdots & \ddots & \vdots \\ s_{x_1}^n & \cdots & s_{x_n}^n \end{pmatrix}_{n \times n}$$

CAPÍTULO 2

Problema de Kantorovich

Kantorovich introduce una versión *relajada* del problema original de Monge, con la intención de convertirlo en un problema lineal. Para ello introduce la clase de medidas en el espacio producto $\mathbb{R}^n \times \mathbb{R}^n$

$$(2.1) \quad \Pi(\mu, \nu) = \{ \text{Medidas de probabilidad de Radón } \gamma \in \mathbb{R}^n \times \mathbb{R}^n \} \\ (\text{proy}_x \gamma = \mu, \text{ proy}_y = \nu)$$

es decir, las proyecciones de las primeras n coordenadas y las últimas n son, respectivamente μ y ν .

Dado $\gamma \in \Pi(\mu, \nu)$, definimos el nuevo coste

$$(2.2) \quad J[\gamma] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\gamma(x, y)$$

Por tanto, si tenemos una función $s \in \mathcal{A}$, entonces definimos la siguiente medida

$$(2.3) \quad \gamma(E) = \mu\{x \in \mathbb{R}^n \mid (x, s(x)) \in E \subset \mathbb{R}^n \times \mathbb{R}^n\} \quad (E \text{ Borel})$$

que pertenece a $\Pi(\mu, \nu)$. Además, el nuevo funcional (2.2) es lineal respecto a γ . Por tanto, bajo las condiciones apropiadas sobre el coste, ciertos argumentos de compacidad confirman la existencia de al menos una medida $\gamma^* \in \Pi(\mu, \nu)$ que satisface

$$(2.4) \quad J[\gamma^*] = \min_{\gamma \in \Pi(\mu, \nu)} J[\gamma]$$

Sin embargo γ^* no está generada necesariamente por una función $s \in \mathcal{A}$ y por ello consideramos esta solución como una *solución débil* del problema original de Monge.

CAPÍTULO 3

Problema Dual

Comencemos conociendo que es el problema dual y algunos de sus teoremas más importantes.

Asumimos que conocemos $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ y $A_{(m \times n)}$. El problema lineal primal consiste en encontrar $x^* \in \mathbb{R}^n$ que cumpla:

$$(3.1) \quad (P) \quad \begin{cases} \text{máx } c^\top x \\ \text{s.a. } Ax \leq b \end{cases}$$

El problema lineal dual consiste en encontrar $y^* \in \mathbb{R}^m$ que cumpla:

$$(3.2) \quad (D) \quad \begin{cases} \text{mín } y^\top b \\ \text{s.a. } A^\top y = c, y \geq 0 \end{cases}$$

Teorema 3.1 (Dualidad débil). *Sea \bar{x} una solución factible de (P) y sea \bar{y} una solución factible de (D) $\Rightarrow c^\top \bar{x} \leq \bar{y}^\top b$*

Demostración. Partiendo de las hipótesis de nuestro teorema, observamos por un lado, que $A^\top \bar{y} = c \Rightarrow \bar{x}^\top A^\top \bar{y} = \bar{x}^\top c$. Por otro lado, $A\bar{x} \leq b$ e $\bar{y} \geq 0 \Rightarrow \bar{y}^\top A\bar{x} \leq \bar{y}^\top b$.

Además, sabemos que $\bar{x}^\top A^\top \bar{y}$ es una matriz de dimensión 1×1 , por tanto $\bar{x}^\top A^\top \bar{y} = (\bar{x}^\top A^\top \bar{y})^\top = \bar{y}^\top A\bar{x}$. Por el mismo motivo $\bar{x}^\top c = c^\top \bar{x}$. De este modo obtenemos:

$$c^\top \bar{x} = \bar{x}^\top c = \bar{x}^\top A^\top \bar{y} = \bar{y}^\top A\bar{x} \leq \bar{y}^\top b$$

□

Teorema 3.2 (Dualidad Fuerte). *Existen cuatro posibilidades para los problemas primal (P) y dual (D):*

1. *El primal y el dual no son factibles.*
2. *El primal no es factible y el dual no está acotado.*
3. *El dual no es factible y el primal no está acotado.*
4. *Tanto el primal como el dual son factibles y tienen la misma solución.*

Demostración. Para esta demostración damos por conocidos los lemas de Farkas que podemos encontrar en el Apéndice I.

Probamos los cuatro casos por separado:

Caso 1: Un ejemplo en el que tanto el primal como el dual no son factibles:

$$\begin{cases} c = (1) \\ b = (-1) \\ A = (0) \end{cases}$$

Caso 2: Supongamos que \bar{y} es una solución factible del problema dual y asumimos que el primal no es factible. Aplicamos el segundo Lema de Farkas, (1') no se cumple, por lo tanto (2') debe cumplirse, es decir, existe \hat{y} tal que $A^T \hat{y} = 0$, $\hat{y}^T b < 0$, $\hat{y} \geq 0$. Consideramos la recta $\bar{y} + \lambda \hat{y}$, $\lambda \geq 0$. Entonces,

$$A^T(\bar{y} + \lambda \hat{y}) = c + \lambda \cdot 0 = c$$

y

$$\bar{y} + \lambda \hat{y} \geq 0$$

por tanto, para todo $\lambda \geq 0$, $\bar{y} + \lambda \hat{y}$ es una solución factible del dual. Además,

$$(\bar{y} + \lambda \hat{y})^T b = \bar{y}^T b + \lambda \hat{y}^T b$$

Sabemos que $\hat{y}^T b < 0 \Rightarrow$ cuando $\lambda \rightarrow \infty$, $(\bar{y} + \lambda \hat{y})^T b \rightarrow -\infty$. Esto prueba que, en este caso, el problema dual no está acotado.

Caso 3: Sea \bar{x} una solución factible para el problema primal y asumiendo que el dual no tiene solución factible $\Rightarrow \nexists y$ tal que $A^T y = c$, $y \geq 0$. Empleamos el primer Lema de Farkas, (si renombramos $A \rightarrow A^T$, $x \rightarrow y$, $b \rightarrow c$) (1) no se cumple, por tanto (2) debe cumplirse, es decir, existe \hat{x} tal que $A \hat{x} \geq 0$, $\hat{x}^T c = c^T \hat{x} < 0$. Consideramos la recta $\bar{x} - \lambda \hat{x}$, $\lambda \geq 0$. Entonces,

$$A(\bar{x} - \lambda \hat{x}) \leq b - \lambda A \hat{x} \leq b$$

por tanto, para todo $\lambda \geq 0$, $\bar{x} - \lambda \hat{x}$ es una solución factible del primal. Además,

$$c^T(\bar{x} - \lambda \hat{x}) = c^T \bar{x} - \lambda c^T \hat{x}$$

Sabemos que $c^T \hat{x} < 0 \Rightarrow$ cuando $\lambda \rightarrow \infty$, $c^T(\bar{x} - \lambda \hat{x}) \rightarrow -\infty$. Esto prueba que, en este caso, el problema primal no está acotado.

Caso 4: Sea \bar{x} y \bar{y} soluciones factibles de los problemas primal y dual, respectivamente. Por el teorema de Dualidad Débil, $c^T \bar{x} \leq \bar{y}^T b$, por tanto ambos problemas están acotados. Sea γ el valor óptimo para el dual. Supongamos que el valor óptimo del problema primal es menor que γ , es decir, supongamos que $\nexists x$ tal que

$$Ax \leq b, cx \geq \gamma$$

es decir, $\nexists x$ tal que

$$\begin{bmatrix} A \\ -c^T \end{bmatrix} x \leq \begin{bmatrix} b \\ -\gamma \end{bmatrix}$$

Esto implica que la condición (1') del segundo lema de Farkas no se cumple, por tanto (2') debe hacerlo. Entonces, existe un vector $\begin{bmatrix} y^T \\ \lambda \end{bmatrix}$, donde $\lambda \in \mathbb{R}$, que satisface

$$\begin{bmatrix} A \\ -c^T \end{bmatrix}^T \begin{bmatrix} y \\ \lambda \end{bmatrix} = 0, \quad \begin{bmatrix} b \\ -\gamma \end{bmatrix}^T \begin{bmatrix} y \\ \lambda \end{bmatrix} < 0, \quad \begin{bmatrix} y \\ \lambda \end{bmatrix} \geq 0$$

Supongamos $\lambda = 0$. Esto implica que $A^T y = 0$, $b^T < 0$, $y \geq 0$, es decir, la condición (2') del segundo lema de Farkas se cumple. Por este motivo, la condición (1') no se cumple, es decir, $\nexists x$ tal que $Ax \leq b$. Esto contradice la hipótesis, ya que según ésta el problema primal es factible. Por tanto, $\lambda > 0$.

Si extendemos la ecuación matricial anterior:

$$A^T y - \lambda c = 0 \Rightarrow A^T \left(\frac{y}{\lambda} \right) = c$$

Además, $\frac{y}{\lambda} \geq 0 \Rightarrow \frac{y}{\lambda}$ es una solución factible del problema dual. Sin embargo,

$$b^T y - \lambda \gamma < 0 \Rightarrow b^T \left(\frac{y}{\lambda} \right) = \gamma$$

lo cual contradice que γ es el valor óptimo del dual. Por tanto, si tanto el problema primal como el dual son factibles, sus valores óptimos son iguales.

□

TODO: el teorema y la prueba de que el dual de el dual es el primal, porque "partimos de un mínimo y queremos llegar a un máximo"

3.1. Caso Discreto

Ahora que conocemos estos teoremas los aplicamos al problema de Kantorovich. Para ello, observamos primero el caso discreto. Supongamos entonces que c_{ij}, μ_i, ν_j ($i = 1, \dots, n; j = 1, \dots, m$) son conocidos y no negativos, satisfacen la condición de equilibrio

$$(3.3) \quad \sum_{i=1}^n \mu_i = \sum_{j=1}^m \nu_j$$

Queremos encontrar γ_{ij}^* que

$$(3.4) \quad \min \sum_{i=1}^n \sum_{j=1}^m c_{ij} \gamma_{ij}$$

y esté sujeto a las siguientes condiciones

$$(3.5) \quad \sum_{j=1}^m \gamma_{ij} = \mu_i, \sum_{i=1}^n \gamma_{ij} = \nu_j, \gamma_{ij} \geq 0$$

Como podemos observar (3.3), (3.4) y (3.5) son las equivalencias para el caso discreto de (1.1), (2.2) y (2.1) respectivamente.

Si comparamos este problema con el dual (3.2) mediante las equivalencias

$$\begin{aligned} y^\top &= (\gamma_{11}, \gamma_{12}, \dots, \gamma_{1m}, \gamma_{21}, \dots, \gamma_{nm}), \\ b^\top &= (c_{11}, \dots, c_{1m}, c_{21}, \dots, c_{nm}), \\ c^\top &= (\mu_1, \dots, \mu_n, \nu_1, \dots, \nu_m), \\ x^\top &= (u_1, \dots, u_n, v_1, \dots, v_m) \end{aligned}$$

siendo A^\top una matriz de dimensión $(n+m) \times nm$ de la forma

$$\begin{pmatrix} \mathbf{1} & 0 & \cdots & 0 \\ 0 & \mathbf{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1} \\ e_1 & e_1 & \cdots & e_1 \\ e_2 & e_2 & \cdots & e_2 \\ \vdots & \vdots & \ddots & \vdots \\ e_m & e_m & \cdots & e_m \end{pmatrix}$$

Las primeras n filas están compuestas por los elementos $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^m$ y $\mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^m$. Las siguientes m filas contienen los vectores $e_i = (0, 0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^m$ de modo que el 1 se encuentra en la posición i del vector. Con esto obtenemos el problema dual (del dual)

$$(3.6) \quad \text{máx} \sum_{i=1}^n u_i \mu_i + \sum_{j=1}^m v_j \nu_j$$

sujeto a

$$(3.7) \quad u_i + v_j \leq c_{ij} (i = 1, \dots, n; j = 1, \dots, m)$$

3.2. Caso General

Como sabemos, no todas las medidas son discretas, por lo que debemos ver el problema dual análogo para todo tipo de medidas.

Introduzcamos primero como sería la forma correspondiente a las condiciones (3.7) a las que está sujeto el problema dual

$$(3.8) \quad \mathcal{L} := \{(u, v) \mid u, v : \mathbb{R}^n \rightarrow \mathbb{R} \text{ continuas}, u(x) + v(y) \leq c(x, y)\}$$

para $x, y \in \mathbb{R}^n$.

Del mismo modo enunciamos la ecuación equivalente a (3.6) para el caso general

$$(3.9) \quad K[u, v] := \int_{\mathbb{R}^n} u(x) d\mu(x) + \int_{\mathbb{R}^n} d\nu(y)$$

De modo que el problema dual del problema de Kantorovich (2.4), para cualquier medida, consistiría en encontrar el par $(u^*, v^*) \in \mathcal{L}$ tal que

$$(3.10) \quad K[u^*, v^*] = \max_{(u, v) \in \mathcal{L}} K[u, v]$$

Este nuevo problema nos muestra una importante diferencia respecto al problema original (1.5). En el problema de Monge necesitábamos encontrar una función óptima $s^* \in \mathcal{A}$ que satisfaga restricciones no lineales; en cambio, en el problema dual de Kantorovich buscamos un par óptimo $(u^*, v^*) \in \mathcal{L}$. Esto, como veremos más adelante es más sencillo debido a su estructura.

Nota. No hemos demostrado este resultado debido a que hemos preferido enfocarnos en otros temas que nos resultan más interesantes, pero podemos encontrar una prueba para el problema dual general en el apartado 1.6.3 de Santambrogio [1].

CAPÍTULO 4

Existencia para coste cuadrático

En esta sección vamos a considerar un caso en el que el coste es uniformemente convexo, un coste cuadrático

$$(4.1) \quad c(x, y) := \frac{1}{2}|x - y|^2 \quad (x, y \in \mathbb{R}^n),$$

donde $|\cdot|$ es la norma Euclídea. Con este nuevo coste nos preguntamos si podemos construir un $s^* \in \mathcal{A}$ que resuelva el problema de Monge (1.5) dado

$$(4.2) \quad I[s] := \frac{1}{2} \int_{\mathbb{R}^n} |x - s(x)|^2 d\mu(x) \quad (s \in \mathcal{A})$$

Pero antes de responder a esta pregunta primero vamos a observar algunas de las propiedades geométricas de s^* en caso de que exista.

4.1. Geometría del transporte óptimo

Supongamos entonces que existe un $s^* \in \mathcal{A}$, entre todas las funciones $s \in \mathcal{A}$, que minimiza (4.2). Fijamos un m entero positivo y tomamos distintos puntos $\{x_k\}_{k=1}^m \subset X$, asumimos que podemos hallar bolas disjuntas

$$(4.3) \quad E_k := B(x_k, r_k) \quad (k = 1, \dots, m)$$

además tomamos los radios $\{r_k\}_{k=1}^m$ de forma que la medida en estas bolas sea la misma

$$(4.4) \quad \mu(E_1) = \dots = \mu(E_m) = \varepsilon$$

Ahora aplicamos s^* de modo que

$$(4.5) \quad \begin{cases} y_k := s^*(x_k) \\ F_k := s^*(E_k) \end{cases}$$

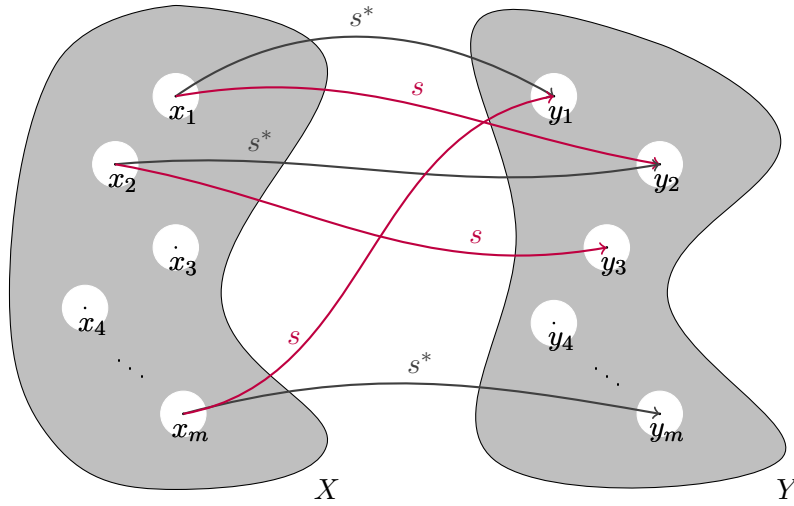
y si recordamos, $\nu = s^*_{\#}(\mu)$ y por ello

$$(4.6) \quad \nu(F_1) = \dots = \nu(F_m) = \varepsilon$$

Elegimos ahora una función $s \in \mathcal{A}$, la cual es una variación de s^* que se obtiene al permutar (de forma cíclica) las bolas imágenes $\{E_k\}_{k=1}^m$. Es decir, obtenemos $s \in \mathcal{A}$ tan que

$$(4.7) \quad \begin{cases} s(x_k) = y_{k+1} \\ s(E_k) = F_{k+1} \\ s \equiv s^* \text{ en } X - \bigcup_{k=1}^m E_k \end{cases}$$

para $k = 1, \dots, m$, donde $y_{m+1} := y_1$, $F_{m+1} := F_1$.



Debido a que hemos asumido que s^* minimiza (4.2) sabemos que

$$(4.8) \quad I[s^*] \leq I[s]$$

es decir,

$$(4.9) \quad \frac{1}{2} \int_X |x - s^*(x)|^2 d\mu(x) \leq \frac{1}{2} \int_X |x - s(x)|^2 d\mu(x)$$

dado que $X = (\bigcup_{k=1}^m E_k) \cup (X - \bigcup_{k=1}^m E_k)$ podemos separar las integrales y simplificar las partes en las que $s \equiv s^*$, es decir, en $X - \bigcup_{k=1}^m E_k$

$$\sum_{k=1}^m \int_{E_k} |x - s^*(x)|^2 d\mu(x) \leq \sum_{k=1}^m \int_{E_k} |x - s(x)|^2 d\mu(x)$$

si desarrollamos los cuadrados llegamos a la siguiente inecuación

$$\sum_{k=1}^m \int_{E_k} 2\langle x, s(x) - s^*(x) \rangle d\mu(x) \leq \int_{\bigcup_{k=1}^m E_k} |s(x)|^2 d\mu(x) - \int_{\bigcup_{k=1}^m E_k} |s^*(x)|^2 d\mu(x)$$

fijémonos ahora únicamente en la parte derecha de esta desigualdad. Si tenemos en cuenta que $s_{\#}(\mu) = s_{\#}^*(\mu)$, puesto que cumplen (1.3), vemos que la parte derecha se anula

$$\int_{\bigcup_{k=1}^m F_k} |y|^2 d\nu(y) - \int_{\bigcup_{k=1}^m F_k} |y|^2 d\nu(y) = 0$$

Por último dividimos por ε y forzamos que $\varepsilon \rightarrow 0$ de modo que obtenemos

$$\sum_{k=1}^m \langle x_k, y_{k+1} - y_k \rangle \leq 0$$

que puede ser reordenada de la siguiente forma

$$(4.10) \quad \sum_{k=1}^m \langle x_{k-1}, y_k \rangle \leq \sum_{k=1}^m \langle x_k, y_k \rangle$$

donde $x_{-1} := x_m$.

Antes de continuar debemos definir algunos conceptos.

Definición 4.1 (Relación). Se dice que ρ es una relación en $E \times E^*$ si ρ es una función multivaluada, es decir $\rho(x) = X^* \subseteq E^*$.

Definición 4.2 (Cíclicamente monótono). Sea ρ una relación en $\mathbb{R}^n \times \mathbb{R}^n$ tal que $x_i^* \in \rho(x_i)$, se dice que ρ es cíclicamente monótona si y solo si

$$\sum \langle x_i, x_i^* \rangle \geq \sum \langle x_{\sigma(i)}, x_i^* \rangle$$

para todo conjunto finito de puntos y toda permutación σ .

Definición 4.3 (Función propia). La función f en X se dice propia si existe $x_0 \in X$ tal que $f(x_0) < +\infty$ y $f(x) > -\infty$ para todo $x \in X$.

Definición 4.4 (Función convexa). una función f en X es convexa si para todo $x, y \in X$ se cumple

$$(4.11) \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

para $\lambda \in (0, 1)$.

Definición 4.5 (Subgradiente y subdiferencial). un subgradiente de f en $x \in E$ es un $x^* \in E^*$ tal que

$$f(y) \geq f(x) + \langle y - x, x^* \rangle$$

para todo $y \in E$. Además, el subdiferencial de f en $x \in E$ ($\partial f(x)$) es el conjunto de todos los posibles subgradientes de f en x .

Teorema 4.6. *Sea ρ una relación en $X \times Y \subset \mathbb{R}^n \times \mathbb{R}^n$. ρ es cíclicamente monótona \Leftrightarrow existe una función convexa propia f en X tal que $\rho \subseteq \partial f$.*

Demostración.

\Leftarrow) Sea ρ una relación en $E \times E^* \subset \mathbb{R}^n \times \mathbb{R}^n$, y dado un conjunto de pares $\{(x_i, x_i^*), i \in I\} \subset E \times E^*$ (denominado *el grafo de ρ*), elegimos una función convexa propia f , que por ello debe cumplir

$$f(y) \geq f(x_i) + \langle y - x_i, x_i^* \rangle,$$

para todo $i \in I, y \in E$.

Por tanto si f es una función convexa propia en E y si

$$x_i^* \in \rho(x_i) \subseteq \partial f$$

donde $i = 0, 1, \dots, n$, entonces

$$+\infty > f(x_j) \geq f(x_i) + \langle x_j - x_i, x_i^* \rangle$$

para todo i, j . Si aplicamos la desigualdad varias veces para los distintos x_i

$$0 \geq \langle x_0 - x_n, x_n^* \rangle + \dots + \langle x_2 - x_1, x_1^* \rangle + \langle x_1 - x_0, x_0^* \rangle$$

o lo que es lo mismo

$$\sum_{i=1}^n \langle x_i, x_i^* \rangle \geq \sum_{i=1}^n \langle x_{i+1}, x_i \rangle$$

Pero esta desigualdad no es la única a la que podríamos llegar, esto se cumpliría para todo conjunto de $n+1$ pares, es decir,

$$\sum_{i=1}^n \langle x_i, x_i^* \rangle \geq \sum_{i=1}^n \langle x_{\sigma(i)}, x_i \rangle$$

para toda permutación σ , y por ello ρ es cíclicamente monótono.

\Rightarrow) Suponemos que ρ es una relación cíclicamente monótona en $E \times E^*$. Sin pérdida de generalidad podemos suponer que ρ es no vacía y fija algún $x_0 \in E$ y $x_0^* \in E^*$ tal que $x_0^* \in \rho(x_0)$. Definimos ahora f de modo que para cada $x \in E$

$$f(x) = \sup\{\langle x - x_n, x_n^* \rangle + \dots + \langle x_1 - x_0, x_0^* \rangle\}$$

donde $x_i^* \in \rho(x_i), i = 1, \dots, n$ y el supremo se toma de entre todos los posibles conjuntos finitos de pares (x_i, x_i^*) .

Dado que f es el supremo de un conjunto no vacío de funciones afines, $f(x) < +\infty$ y cumple la definición de convexidad (4.11). Además,

$$0 \leq \langle x_0 - x_n, x_n^* \rangle + \dots + \langle x_1 - x_0, x_0^* \rangle$$

ya que ρ es cíclicamente monótona. Por tanto $f(x_0) < +\infty$, lo que implica que f es una función propia convexa.

Aún queda por ver que, si tomamos \bar{x} y \bar{x}^* tal que $\bar{x}^* \in \rho(\bar{x}), \bar{x}^* \in \partial f$. Para ello debemos ver que, para todo $\alpha < f(\bar{x})$

$$f(x) \geq \alpha + \langle x - \bar{x}, \bar{x}^* \rangle \text{ para todo } x \in E$$

Dado que $\alpha < f(\bar{x})$ y podemos elegir un conjunto de pares (x_i, x_i^*) $i = 1, \dots, k$ de modo que

$$\alpha < \langle \bar{x} - x_k, x_k^* \rangle + \dots + \langle x_1 - x_0, x_0^* \rangle$$

Si renombramos $\bar{x} = x_{k+1}$ y $\bar{x}^* = x_{k+1}^*$ y sustituimos en la desigualdad

$$f(x) \geq \langle x - x_{k+1}, x_{k+1}^* \rangle + \dots + \langle x_1 - x_0, x_0^* \rangle$$

y esto se cumple para todo $x \in E$ por definición de f . Por tanto queda probado que $\rho \subseteq \partial f$. \square

Recordemos que hemos llegado a la ecuación (4.10) donde s permutaba de forma cíclica las bolas imágenes $\{E_k\}_{k=1}^m$. Se ha decidido que la permutación sea $\sigma(k) = k - 1$, pero podríamos haber elegido otra cualquiera definiendo una función s distinta. Por tanto, podemos afirmar que dado grafo $\{(x, s^*(x)) \mid x \in X\} \subset \mathbb{R}^n \times \mathbb{R}^n$, s^* es cíclicamente monótona.

Hemos visto que s^* es cíclicamente monótona entonces, por el teorema que acabamos de ver

$$(4.12) \quad s^* \subseteq \partial \phi^*$$

para alguna función convexa $\phi^* : \mathbb{R}^n \rightarrow \mathbb{R}$. Además, por [2] una función convexa es diferenciable en casi todo punto (a.e.), por tanto

$$(4.13) \quad s^* = D\phi^* \quad \text{a.e. en } X$$

donde $D\phi^*$ es el gradiente de ϕ^* .

La conclusión que sacamos entonces, es que, en caso de que exista una función s^* que optimice nuestro problema, ésta será el gradiente de una función convexa ϕ^* .

4.2. Solución del problema dual

En esta sección vamos a probar que existe una solución óptima para el problema de transporte. Lo haremos partiendo del problema dual de Kantorovich. En lo sucesivo supondremos que

$$(4.14) \quad d\mu = f \, dx, \quad d\nu = g \, dy$$

tal que f y g son funciones acotadas, no negativas con soporte compacto. El soporte de estas funciones es X e Y respectivamente. Además, estas dos funciones cumplen la condición de equilibrio

$$(4.15) \quad \int_X f(x) \, dx = \int_Y g(y) \, dy$$

Para este caso el problema es encontrar (u^*, v^*) que maximice

$$(4.16) \quad K[u, v] := \int_X u(x) f(x) \, dx + \int_Y v(y) g(y) \, dy$$

sujeto a

$$(4.17) \quad u(x) + v(y) \leq \frac{1}{2}|x - y|^2 \quad (x \in X, y \in Y)$$

Queremos utilizar las herramientas del análisis convexo, para ello debemos realizar un cambio de variable

$$(4.18) \quad \begin{cases} \phi(x) := \frac{1}{2}|x|^2 - u(x) & (x \in X) \\ \psi(y) := \frac{1}{2}|y|^2 - v(y) & (y \in Y) \end{cases}$$

de modo que las restricciones (4.17) son ahora

$$(4.19) \quad \phi(x) + \psi(y) \geq \langle x, y \rangle$$

y nuestra meta ahora será encontrar ϕ^* y ψ^* que minimice

$$(4.20) \quad L[\phi, \psi] := \int_X \phi(x)f(x)dx + \int_Y \psi(y)g(y)dy$$

Lema 4.7. *Existe (ϕ^*, ψ^*) que resuelve el problema de minimización (4.20). Además, (ϕ^*, ψ^*) son funciones duales convexas del tipo*

$$(4.21) \quad \begin{cases} \phi^*(x) = \max_{y \in Y} (\langle x, y \rangle - \psi^*(y)) & (x \in X) \\ \psi^*(y) = \max_{x \in X} (\langle x, y \rangle - \phi^*(x)) & (y \in Y) \end{cases}$$

Demostración.

Sabemos que las funciones ϕ y ψ deben satisfacer 4.19, entonces

$$\phi(x) \geq \max_{y \in Y} (\langle x, y \rangle - \psi(y)) =: \hat{\phi}(x)$$

por tanto

$$\hat{\phi}(x) + \psi(y) \geq \langle x, y \rangle$$

del mismo modo

$$\psi(y) \geq \max_{x \in X} (\langle x, y \rangle - \hat{\phi}(x)) =: \hat{\psi}(y)$$

entonces

$$(4.22) \quad \hat{\phi}(x) + \hat{\psi} \geq \langle x, y \rangle$$

Dado que $\psi \geq \hat{\psi}$

$$(4.23) \quad \max_{y \in Y} (\langle x, y \rangle - \hat{\psi}(y)) \geq \hat{\phi}(x)$$

Por (4.22) y (4.23) tenemos que

$$\hat{\phi}(x) = \max_{y \in Y} (\langle x, y \rangle - \hat{\psi}(y))$$

Dado que $f, g \geq 0$, $\phi \geq \hat{\phi}$ y $\psi \geq \hat{\psi}$, vemos que $L[\hat{\phi}, \hat{\psi}] \leq L[\phi, \psi]$. Esto quiere decir que para todo ϕ y ψ podemos encontrar un par $\hat{\phi}$ y $\hat{\psi}$ que mejore la solución del problema de transporte óptimo.

Sabemos que X, Y son compactos por tanto tenemos

$$\begin{aligned}\bar{x} &= \sup(X) \\ \bar{y} &= \sup(Y)\end{aligned}$$

Además, si elegimos y_i tal que $\phi(x_i) = \langle x_i, y_i \rangle - \psi(y_i)$ e $y^* = \sup(Y)$

$$\begin{aligned}|\phi(x_1) - \phi(x_2)| &= |\langle x_1, y_1 \rangle - \psi(y_1) - \langle x_2, y_2 \rangle + \psi(y_2)| \leq \\ &\leq |\langle x_1, y_2 \rangle - \langle x_2, y_2 \rangle| \leq \bar{y}_2 |x_1 - x_2| \leq \\ &\leq \bar{y}^* |x_1 - x_2|\end{aligned}$$

tomando $\bar{y}_i := \max_{j \in \{1, \dots, n\}} (y_{ij})$, donde y_{ij} representa la cordenada j -ésima del vector y_i .

Con esto probamos que todo ϕ es Lipschitz continua y de forma similar podemos probar que ψ lo es.

Por otro lado,

$$\inf_{\phi \in \Phi, \psi \in \Psi} (L[\phi, \psi]) = m$$

donde Φ y Ψ es el conjunto de todas las funciones de la forma (4.21). Por tanto existen sucesiones ϕ_n y ψ_n tal que $L[\phi_n, \psi_n] \rightarrow m$. Teniendo en cuenta que estas funciones son Lipschitz continuas, por el teorema de Arzelá-Ascoli existen subsucesiones ϕ_{n_k} y ψ_{n_k} tal que $\lim_{k \rightarrow \infty} \phi_{n_k} = \phi^*$ y $\lim_{k \rightarrow \infty} \psi_{n_k} = \psi^*$. Este teorema también nos asegura convergencia uniforme en compactos, por tanto

$$\lim_{k \rightarrow \infty} L[\phi_{n_k}, \psi_{n_k}] = L[\phi^*, \psi^*] = m$$

□

4.3. Existencia de solución del problema de transporte óptimo

Comenzamos esta sección definiendo $\phi^*(x)$ y $\psi^*(y)$ como en el apartado anterior

$$(4.24) \quad \begin{cases} \phi^*(x) = \max_{y \in Y} (\langle x, y \rangle - \psi^*(y)) & (x \in X) \\ \psi^*(y) = \max_{x \in X} (\langle x, y \rangle - \phi^*(x)) & (y \in Y) \end{cases}$$

de modo que, además, estas funciones sean las que minimizan $L[\phi, \psi]$. Sabemos también que ϕ^* y ψ^* son convexas y diferenciables en casi todo punto.

Demostraremos a continuación que s^* es solución del problema de transporte donde

$$(4.25) \quad s^*(x) := D\phi^*(x) \quad a.e. \ x \in X$$

Teorema 4.8. *Sea $s^*(x) := D\phi^*(x)$ a.e $x \in X$ entonces:*

- (1) $s^* : X \rightarrow Y$ es esencialmente biyectiva.
- (2) $\int_X h(s^*(x))d\mu(x) = \int_Y h(y)d\nu(y)$ para toda función $h \in C(Y)$
- (3) Para todo $s : X \rightarrow Y$ que cumpla $s_{\#}(\mu) = \nu$

$$\frac{1}{2} \int_X |x - s^*(x)|^2 d\mu(x) \leq \frac{1}{2} \int_X |x - s(x)|^2 d\mu(x)$$

Demostración.

Comenzamos viendo que, por la definición de ϕ^* dada en (4.24), $s^* = D\phi^* \in Y$ a.e. Tomamos $\tau > 0$ y definimos

$$(4.26) \quad \begin{cases} \phi_\tau(x) := \max_{y \in Y} (\langle x, y \rangle - \psi_\tau(y)) & (x \in X) \\ \psi_\tau(y) := \psi^*(y) + \tau h(y) & (y \in Y) \end{cases}$$

para $h \in C(Y)$.

Estas nuevas funciones cumplen

$$\phi_\tau(x) + \psi_\tau(y) \geq \langle x, y \rangle \quad (x \in X, y \in Y)$$

y teniendo en cuenta que ϕ^* y ψ^* minimizan el problema de transporte

$$L[\phi^*, \psi^*] \leq L[\phi_\tau, \psi_\tau] := i(\tau)$$

La función que lleva $\tau \rightarrow i(\tau)$ tendría por lo tanto un mínimo para $\tau = 0$, ya que $\phi_0 = \phi^*$ y $\psi_0 = \psi^*$, y por ello

$$(4.27) \quad \begin{aligned} 0 &\leq \frac{1}{\tau} (L[\phi_\tau, \psi_\tau] - L[\phi^*, \psi^*]) \\ &= \int_X \frac{\phi_\tau(x) - \phi^*(x)}{\tau} f(x) dx + \int_Y h(y) g(y) dy \end{aligned}$$

Elegimos $y_\tau(x) \in Y$ de modo que, si fijamos x

$$\phi_\tau(x) = \langle x, y_\tau(x) \rangle - \psi_\tau(y_\tau(x))$$

esto quiere decir que, para cada x tenemos un $y_\tau \in Y$ con el que se consigue la igualdad anterior. Podemos observar entonces

$$\begin{aligned} \phi_\tau(x) - \phi^*(x) &= \langle x, y_\tau(x) \rangle - \psi^*(y_\tau(x)) - \tau h(y_\tau(x)) - \phi^*(x) \\ &\leq \langle x, y_\tau(x) \rangle - \psi^*(y_\tau(x)) - \tau h(y_\tau(x)) - \langle x, y_\tau(x) \rangle + \psi^*(y_\tau(x)) \\ &= \tau h(y_\tau(x)) \end{aligned}$$

Por otro lado, si tomamos $y^*(x) \in Y$ del mismo modo que antes, de forma que se cumpla, para un valor de x fijado,

$$(4.28) \quad \phi^*(x) = \langle x, y^*(x) \rangle - \psi^*(y^*(x))$$

entonces

$$\begin{aligned}\phi_\tau(x) - \phi^*(x) &\geq \langle x, y^*(x) \rangle - \psi^*(y^*(x)) - \tau h(y^*(x)) - \phi^*(x) \\ &= \langle x, y^*(x) \rangle - \psi^*(y^*(x)) - \tau h(y^*(x)) - \langle x, y^*(x) \rangle + \psi^*(y^*(x)) \\ &= -\tau h(y^*(x))\end{aligned}$$

Si combinamos los dos resultados anteriores obtenemos

$$(4.29) \quad -h(y^*(x)) \leq \frac{\phi_\tau(x) - \phi^*(x)}{\tau} \leq -h(y_\tau(x))$$

Si tomamos $x \in X$ de modo que exista $s^* := D\phi^*(x)$ entonces $s^*(x) = y^*(x)$

Además, cuando $\tau \rightarrow 0$, $y_\tau \rightarrow y^*(x) = s^*(x)$. En el caso de (4.29), con lo que acabamos de ver, cuando $\tau \rightarrow 0$

$$\frac{\phi_\tau(x) - \phi^*(x)}{\tau} \rightarrow -h(s^*(x))$$

Se puede observar que $\left| \frac{\phi_\tau - \phi^*}{\tau} \right| \leq \|h\|_\infty$. Por el teorema de convergencia dominada junto con (4.27)

$$\int_X -\frac{\phi_\tau(x) - \phi^*(x)}{\tau} f(x) dx \rightarrow \int_X h(s^*(x)) f(x) dx \leq \int_Y h(y) g(y) dy$$

La función h se ha elegido de un modo arbitrario, por lo que si realizásemos el mismo proceso para la función $-h$ obtendríamos

$$\int_X h(s^*(x)) f(x) dx \geq \int_Y h(y) g(y) dy$$

Con lo que conseguimos la igualdad

$$(4.30) \quad \int_X h(s^*(x)) f(x) dx = \int_Y h(y) g(y) dy$$

Esto prueba (2).

Tomamos ahora s cualquier función que cumple (1.3), entonces, si elegimos $h = \psi^*$

$$(4.31) \quad \int_X \psi^*(s(x)) f(x) dx = \int_Y \psi^*(y) g(y) dy$$

Sabemos que $\phi^*(x) + \psi^*(y) \geq \langle x, y \rangle$ con igualdad para $y = s^*$ por (4.28)

$$(4.32) \quad \begin{cases} \psi^*(s(x)) \geq \langle x, s(x) \rangle - \phi^*(x) \\ \psi^*(s^*(x)) = \langle x, s^*(x) \rangle - \phi^*(x) \end{cases}$$

Por otro lado sabemos que

$$(4.33) \quad \int_X \psi^*(s(x)) f(x) dx = \int_Y \psi^*(y) f(y) dy = \int_X \psi^*(s^*(x)) f(x) dx$$

donde la primera igualdad la obtenemos de (4.30) para $h = \psi^*$ y la segunda de (4.31). Ahora, podemos dejar ambas integrales a un lado del igual consiguiendo

$$\begin{aligned} 0 &= \int_X \psi^*(s(x))f(x)dx - \int_X \psi^*(s^*(x))f(x)dx \\ &\geq \int_X (\langle x, s(x) \rangle - \phi^*(x))f(x)dx - \int_X (\langle x, s^*(x) \rangle - \phi^*(x))f(x)dx \\ &= \int_X \langle x, s(x) - s^*(x) \rangle f(x)dx \end{aligned}$$

Esta desigualdad se obtiene al sustituir $\phi^*(s(x))$ y $\phi^*(s^*(x))$ por la desigualdad e igualdad, respectivamente, mostradas en (4.32).

Podemos modificar ligeramente esta desigualdad del siguiente modo

$$\begin{aligned} 0 &\geq 2 \int_X \langle x, s(x) - s^*(x) \rangle f(x)dx \\ &= \int_X 2\langle x, s(x) - s^*(x) \rangle f(x)dx - \int_Y |y|^2 g(y)dy + \int_Y |y|^2 g(y)dy \\ &= \int_X 2\langle x, s(x) - s^*(x) \rangle f(x)dx - \int_X |s(x)|^2 f(x)dx + \int_X |s^*(x)|^2 f(x)dx \\ &= \int_X |x - s^*(x)|^2 f(x)dx - \int_X |x - s(x)|^2 f(x)dx \end{aligned}$$

Aclaro que la segunda igualdad se obtiene sustituyendo h por $|\cdot|^*$ en las ecuaciones (4.30) y (1.3). Así queda probado (3). \square

CAPÍTULO 5

Aplicación

Las ecuaciones de difusión no lineales provienen de distintos fenómenos de difusión que pueden encontrarse en la naturaleza. Se consideran modelos matemáticos de problemas físicos. Para abordar algunas de estas ecuaciones se construye un sistema iterativo basado en la discretización del tiempo. Sin embargo, debemos conocer algunos conceptos y teoremas previos.

5.1. Definiciones previas

Antes de comenzar debemos definir el espacio de funciones de densidad de probabilidades \mathcal{P} con segundo momento finito:

$$\mathcal{P} := \left\{ \rho : \mathbb{R}^n \rightarrow \mathbb{R}^n \mid \rho \geq 0 \text{ a.e.}, \int_{\mathbb{R}^n} \rho(x) dx = 1, \int_{\mathbb{R}^n} |x|^2 \rho(x) dx < \infty \right\}$$

Definición 5.1 (Método directo del cálculo variacional). El método directo del cálculo variacional se emplea típicamente para demostrar la existencia de un mínimo para un funcional $F : V \rightarrow \mathbb{R} \cup \infty$, es decir, un $v \in V$ tal que $F(v) \leq F(u)$ para todo $u \in V$. Aplicar este método cuando V es un subespacio de W espacio de Banach implica realizar los siguientes cinco pasos:

1. Demostrar que el funcional está acotado inferiormente, de modo que

$$\inf_{u \in V} J(u) > -\infty$$

2. Definir una sucesión minimizadora $\{u_n\}$ tal que $u_n \rightarrow \inf_{u \in V} J(u)$.
3. Cualquier sucesión minimizadora $\{u_n\}$ debe estar acotada.
4. Mostrar que existe una subsucesión $\{u_{n_k}\}$ de $\{u_n\}$ que converja a algún $u_0 \in W$ respecto a la topología débil.
5. Demostrar que J es débilmente e inferiormente semi-continua, es decir, para cualquier sucesión $u_n \rightarrow u_0$

$$\liminf_{n \rightarrow \infty} J(u_n) \geq J(u_0)$$

Definición 5.2 (Convergencia débil). Se dice que una subsucesión de medidas $\{\mu_n\} \in \mathcal{P}$ converge débilmente a la medida μ si

$$\int_{\mathbb{R}^n} g \mu_n \rightarrow \int_{\mathbb{R}^n} g \mu$$

para toda función g continua con soporte compacto en \mathbb{R}^n .

Teorema 5.3 (Teorema de Banach-Alaoglu sucesional). *Este teorema afirma que la bola unidad cerrada es sucesionalmente compacta en la topología débil.*

Si aplicamos este teorema a nuestro conjunto \mathcal{P} junto al la definición de convergencia débil obtenemos el siguiente teorema de compacidad, el cual nos garantiza que, sobre una función β , la medida límite es absolutamente continua respecto a la medida de Lebesgue. Además, el hecho de que nuestras medidas tengan momentos finitos nos asegura que la medida será de probabilidad.

Teorema 5.4. *Si existe una función β con crecimiento super lineal tal que $\int_{\mathbb{R}^n} \beta(\rho_{n_k}(x)) dx < \infty$, entonces $\{\rho_{n_k}\}$ converge a $\rho_0 \in \mathcal{P}$.*

Definición 5.5 (Distancia Wasserstein). Para medir la distancia entre dos medidas de probabilidad μ y ν :

$$d(\mu, \nu) = \inf_{\gamma \in P(\mu, \nu)} \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\gamma(x, y)$$

donde, $P(\mu, \nu)$ es el conjunto de medidas de probabilidad en $\mathbb{R}^n \times \mathbb{R}^n$ con primera marginal μ y segunda marginal ν .

Si observamos esta distancia, comprobamos que es la solución al problema de Kantorovich para coste $c = \frac{1}{2}|x - y|^2$

Teorema 5.6 (Teorema de convergencia monótona). *Sea $\{f_n\}$ una sucesión monótona de funciones medibles, si alguna de las funciones de esta sucesión es integrable entonces $\lim \int f_k$ y $\int \lim f_k$ existen y, además*

$$\lim \int f_k = \int \lim f_k$$

Proposición 5.7. *Si β es una función convexa con crecimiento superlineal, entonces es débilmente semicontinua en \mathcal{P} . Es decir, sea $\{u_k\} \in \mathcal{P}$ sucesión que converge a u*

$$\liminf_{k \rightarrow \infty} \beta(u_k) \geq \beta(u)$$

5.2. Existencia y unicidad de un mínimo

Lo primero de todo es describir el funcional que queremos minimizar, utilizando un procedimiento de discretización del tiempo mediante pasos. De modo que tomamos

un tamaño de paso pequeño $h > 0$ y comenzando por un $\rho^0 \in \mathcal{P}$ obtenemos $\rho^{k+1} \in \mathcal{P}$ que minimiza el funcional

$$(5.1) \quad N_k := \frac{d(\rho, \rho^k)^2}{h} + \int \beta(\rho) dx$$

donde $\beta : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función convexa y con crecimiento superlineal. Se supone además que:

$$\int_{\mathbb{R}^n} \beta(\rho^0) < \infty$$

Debemos mostrar que este mínimo existe, para ello utilizaremos el método de cálculo de variaciones definido en la sección previa.

De este modo comenzamos probando que $\beta(\rho) > -\infty$. Debido a que tiene un crecimiento superlineal sabemos que

$$\lim_{x \rightarrow \infty} \frac{\beta(x)}{|x|} = \infty$$

y con ello obtenemos la coercividad, que para $C_1 \geq 0$ y $C_2 \geq 0$ constantes, tenemos

$$\beta(x) \geq C_1|x| - C_2$$

como $|x| \geq 0$ queda probado que $\beta(\rho) > -\infty$, $\forall \rho \in \mathcal{P}$. Por ello podemos ver que el funcional (5.1) está acotado inferiormente

$$N_k > -\infty$$

Por tanto podemos definir una sucesión $\{\rho_n\}$ minimizante, es decir, $\lim_{n \rightarrow \infty} N_k(\rho_n) \rightarrow \inf_{\rho \in \mathcal{P}} N_k(\rho)$. Además, por la definición del conjunto \mathcal{P} podemos acotar la sucesión $\{\rho_n\}$

$$0 \leq \int_{\mathbb{R}^n} \rho(x) dx \leq 1$$

El teorema de Banach-Alaoglu nos muestra la existencia de una subsucesión $\{\rho_{n_j}\}$ de $\{\rho_n\}$ que converge débilmente* a $\rho^{k+1} \in \mathcal{P}$.

Por otro lado, dado que $\beta(\rho)$ es convexa $\int_{\mathbb{R}^n} \beta(\rho)$ es débilmente semicontinua respecto a la topología débil*, y por este motivo

$$(5.2) \quad \liminf_{j \rightarrow \infty} \int_{\mathbb{R}^n} \beta(\rho_{n_j}) dx \geq \int_{\mathbb{R}^n} \beta(\rho^{k+1}) dx$$

Nos queda demostrar que

$$(5.3) \quad \liminf_{j \rightarrow \infty} d(\rho_{n_j}, \rho^k)^2 \geq d(\rho^{k+1}, \rho^k)^2$$

Para ello comenzamos eligiendo un $p_j \in \Pi(\rho_{n_j}, \rho^k)$ que satisfaga

$$\frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 dp_j(x, y) \leq d(\rho_{n_j}, \rho^k) + \frac{1}{j}$$

Tomamos ahora un $R > 0$ y la función continua $\eta_R : \mathbb{R}^n \rightarrow [0, 1]$ del siguiente modo:

$$\eta_R = 1 \text{ para } x \in B_R \quad \eta_R = 0 \text{ para } x \notin B_{2R}$$

donde B_R es la bola de centro el origen de \mathbb{R}^n y radio R . Y sea p la medida de probabilidad en $\mathbb{R}^n \times \mathbb{R}^n$ tal que $p_j \rightarrow p \in \Pi(\rho^{k+1}, \rho^k)$ podemos ver que para todo $R < \infty$

$$\begin{aligned} & \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} \eta_R(x)\eta_R(y)|x - y|^2 dp(x, y) \\ &= \lim_{j \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} \eta_R(x)\eta_R(y)|x - y|^2 dp_j(x, y) \\ &\leq \lim_{j \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 dp_j(x, y) \\ &\leq \liminf_{j \rightarrow \infty} d(\rho_{n_j}, \rho^k) \end{aligned}$$

Utilizando el teorema de convergencia monótona obtenemos

$$\begin{aligned} d(\rho^{k+1}, \rho^k) &\leq \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 dp(x, y) \\ &= \lim_{R \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} \eta_R(x)\eta_R(y)|x - y|^2 dp(x, y) \end{aligned}$$

con esto queda probado (5.3).

Por lo tanto se ha demostrado que existe al menos un minimizador $\rho^{k+1} \in \mathcal{P}$ para el funcional (5.1). Pero, aun queda ver que este minimizador es único. Para ello solo es necesario observar tanto β como el funcional $\rho \rightarrow d(\rho, \rho^k)^2$ son convexos.

5.3. Ecuación de Euler-Lagrange

5.4. Convergencia

Apéndice I

Teorema 5.8 (Primer Lema de Farkas). *Sea $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$. Se cumple exactamente una de las siguientes condiciones:*

- (1) $\exists x \in \mathbb{R}^n$ tal que $Ax = b$, $x \geq 0$;
- (2) $\exists y \in \mathbb{R}^m$ tal que $A^T y \geq 0$, $y^T b < 0$.

Demostración. Comenzaremos probando que no se pueden cumplir ambas condiciones simultáneamente. Si tenemos en cuenta $Ax = b$ por (1), y $y^T b < 0$ por (2) obtenemos $y^T Ax = y^T (Ax) = y^T b < 0$. Pero por otro lado, $A^T y \geq 0$ por (2) y $x \geq 0$ por (1), entonces $y^T Ax = (y^T A)x = (A^T y)^T x \geq 0$.

Ahora queremos probar que si (1) no se cumple, entonces se cumple (2). Para ello nombramos v_1, v_2, \dots, v_n a las columnas de A . Definimos

$$Q = \text{cone}(v_1, \dots, v_n) \equiv \left\{ s \in \mathbb{R}^m : s = \sum_{i=1}^n \lambda_i v_i, \lambda_i \geq 0, \right\}$$

Entendida la demo pero tengo que demostrar que es cerrada y convexa, y el teorema de separación por hiperplano

TODO: terminar □

Teorema 5.9 (Segundo Lema de Farkas). *Sea $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$. Se cumple exactamente una de las siguientes condiciones:*

- (1') $\exists x \in \mathbb{R}^n$ tal que $Ax \leq b$;
- (2') $\exists y \in \mathbb{R}^m$ tal que $A^T y = 0$, $y^T b < 0$, $y \geq 0$.

La siguiente condición es equivalente a (2'):

- (2'') $\exists y \in \mathbb{R}^m$ tal que $A^T y = 0$, $y^T b = -1$, $y \geq 0$.

Demostración. **TODO:** leer, entender y escribir □

Bibliografía

- [1] SANTAMBROGIO, F. *Optimal Transport for Applied Mathematicians* Calculus of Variations, PDEs, and Modeling **87**, págs. 35–37 (2015).
- [2] ROCKAFELLAR, R.T. *Convex Analysis* Princeton University Press, Princeton, NJ, pág. 246 (1972).

