Universidade Federal do Rio Grande do Sul

Instituto de Matemática
Programa de Pós-Graduação em Matemática

O Teorema da Dualidade de Kantorovich para o Transporte Ótimo

Dissertação de Mestrado

Aline Duarte de Oliveira

Porto Alegre, julho de 2011

Dissertação submetida por Aline Duarte de Oliveira[1] como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande de Sul.

Professor Orientador:

Dr. Rafael Rigão Souza

Banca Examinadora:

Dr. Artur Oscar Lopes

Dr. Jairo Krás Mengue

Dra. Joana Mohr

Dr. Rogério Ricardo Steffenon

Data da Defesa: 15 de julho de 2011

[1]bolsista da Capes

Ao Guilherme,

e a minha família

Agradecimentos

Agradeço primeiramente ao meu orientador, Rafael Rigão Souza, por toda orientação que me deu e pelos conselhos acadêmicos, que foram de suma importância em minha formação.

Aos meus pais e minha irmã que sempre me apoiaram e acreditaram em mim.

Ao Guilherme por todo carinho e pelas inúmeras horas de estudo durante esses dois anos e meio.

Aos meus amigos pelos momentos de alegria e pela compreensão nos momentos em que não estive presente.

A Capes pelo apoio financeiro.

Resumo

Abordaremos a teoria do transporte ótimo demonstrando o teorema da dualidade de Kantorovich para uma classe ampla de funções custo. Tal resultado desempenha um papel de suma importância na teoria do transporte ótimo. Uma ferramenta importante utilizada é o teorema da dualidade de Fenchel-Rockafellar, aqui enunciado e demonstrado em bastante generalidade.

Demonstramos também o teorema da dualidade de Kantorovich-Rubinstein, que trata do caso particular da função custo distância.

Palavras-chave: Kantorovich, dualidade, transporte ótimo.

Abstract

We analyze the optimal transport theory proving the Kantorovich duality theorem for a wide class of cost functions. Such result plays an extremely important role in the optimal transport theory. An important tool used here is the Fenchel-Rockafellar duality theorem, which we state and prove in a general case.

We also prove the Kantorovich-Rubinstein duality theorem, which deals with the particular case of cost function given by the distance.

Keywords: Kantorovich, duality, optimal transportation.

Sumário

Introdução

Ainda que a origem da área de transporte ótimo se deva ao problema de Gaspard Monge, em 1781, foi a partir de 1942, com a contribuição de Leonid Kantorovich no assunto, que o tema passou a despertar interesse em áreas como equações diferenciais parciais, mecânica de fluidos, geometria, teoria da probabilidade e análise funcional.

Intuitivamente o problema de Kantorovich de transporte ótimo consiste em deslocar uma quantidade de massa ou de energia de um estado inicial x, a um estado final y, com um *custo de transporte* dado por uma função mensurável $c(x, y)$ em $X \times Y$. Procura-se realizar o transporte com o menor custo possível.

Além da função custo c, levaremos em conta a distribuição da massa que pode ser transportada de X, modelada por uma probabilidade μ em X, e a distribuição desejada de massa em Y, modelada por uma probabilidade ν em Y. A quantidade de massa que efetivamente é transferida de um estado para o outro é modelada por uma probabilidade π em $X \times Y$, contendo projeções na primeira e segunda coordenadas dadas respectivamente por μ e ν. π é denominada um *plano de transporte* ou um *acoplamento.*

O problema original, chamado problema de Monge, é um pouco mais simples. A diferença entre o problema de Kantorovich e o de Monge é que o de Monge admite que a massa não possa ser divida, isto é, que toda a massa de um estado inicial é transportada para um mesmo estado final, enquanto que no problema de Kantorovich não existe restrição quanto a divisão da massa no transporte. Isto explica porque, atualmente, o problema de Monge é visto como um caso particular do problema de Kantorovich. De fato o problema de Kantorovich é considerado a versão relaxada do problema de Monge.

Em termos mais rigorosos, o problema de Kantorovich é interpretado da seguinte forma: dados dois espaços de probabilidade $(X, \mathcal{A}^X, \mu)$ e $(Y, \mathcal{A}^Y, \nu)$, e uma função mensurável e não-negativa $c : X \times\ Y \to \mathbb{R}$, considere o conjunto $\Pi(\mu, \nu)$ dado por todas as probabilidades π em $\mathcal{A}^X \times \mathcal{A}^Y$ que satisfazem

$$\pi[A \times Y] = \mu(A) \quad \text{e} \quad \pi[X \times B] = \nu(B)\,,$$

para quaisquer conjuntos mensuráveis $A \subset X$ e $B \subset Y$.

O problema de Kantorovich consiste em minimizar a *função custo de transporte total*, dada por

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

entre todos os planos de transporte $\pi \in \Pi(\mu, \nu)$.

O *custo de transporte ótimo* é dado por

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

Quando existir uma probabilidade em $\Pi(\mu, \nu)$ que satisfaça $I[\pi] = \mathcal{T}_c(\mu, \nu)$, tal probabilidade será chamada um *plano de transporte ótimo.*

Como Monge considera apenas o caso particular em que a massa não é divida no transporte, matematicamente podemos considerar o estado y em função do estado x através de uma função mensurável $T : X \to Y$. Neste caso, o plano de transporte pode ser reescrito como

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

Equivalentemente, $\Pi(\mu, \nu)$ passa a ser caracterizado pela condição

$$\forall B \in \mathcal{A}^Y, \quad \nu(B) = \mu[T^{-1}(B)]$$

e neste caso diremos que ν é a *imagem mensurável* de μ por T, ou que T *transporta* μ para ν e escreveremos $T\#\mu = \nu$.

Finalmente, o problema de Monge será minimizar a função

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

sobre o conjunto de todas as funções mensuráveis T tal que $T\#\mu = \nu$.

Note que os problemas de Monge e de Kantorovich são distintos, uma vez que no problema de Kantorovich procuramos um plano de transporte que minimize o custo do transporte total, enquanto que o problema de Monge trata de obter a minimização a partir de uma função de transporte. De fato, como já dito, o problema de Monge é um caso particular do problema de Kantorovich.

A diferença entre o problema de Monge e o de Kantorovich fica evidente com o seguinte exemplo:

Exemplo 0.0.1. Sejam $X = Y = [0,1]$, $\mu(x) = \delta_{\{1/2\}}(x)$, $\nu(y) = \frac{1}{2}\delta_{\{1/3\}}(y) + \frac{1}{2}\delta_{\{2/3\}}(y)$ e c uma função custo semicontínua inferiormente qualquer. Afirmamos que sob essas condições não existe solução para o problema de Monge.

De fato, suponha por absurdo que exista $T : [0,1] \to [0,1]$ que transporta μ para ν. Sem perda de generalidade, suponha que $T(1/2) = 1/3$ e $T(a) = 2/3$ para algum $a \in [0,1]\backslash\{\frac{1}{2}\}$. Logo, $1/2 = \nu(2/3) = \mu[T^{-1}(2/3)] = \mu(a) = 0$, uma contradição.

Por outro lado, pelo teorema 1.3.2 que demonstraremos durante a dissertação, existe um plano ótimo para o exemplo anterior que minimiza o problema no sentido de Kantorovich.

Em nossa dissertação de mestrado pretendemos abordar a teoria do transporte provando um teorema essencial: o teorema de dualidade de Kantorovich, que mostra que o ínfimo da esperança de custo entre todos os planos de transporte é atingido por ao menos um plano e coincide com um problema de minimização de uma soma de integrais (em relação às medidas de partida e de chegada) de certas funções cuja soma é majorada pela própria função custo. Antes de enunciar e provar tal resultado, e após introduzir os conceitos básicos da teoria de transporte e certos exemplos selecionados, pretendemos introduzir os conceitos básicos de análise necessários ao referido teorema. Pretendemos dar a demonstração do teorema de dualidade para uma classe ampla de funções custo.

Esta dissertação é baseada no livro de Cédric Villani, *Topics in Optimal Transportation*.

Capítulo 1

A Dualidade de Kantorovich

1.1 Preliminares

Em toda a dissertação, qualquer que seja o espaço métrico (X, d), que denotaremos por simplicidade por X, vamos equipá-lo com a topologia induzida pela distância, e denotaremos por $B(x, r)$ a bola de centro x e raio r. Além disso, dados dois espaços métricos (X, d) e $(Y, \tilde{d})$, consideraremos o espaço métrico produto $(X \times Y, [d + \tilde{d}])$ onde $[d + \tilde{d}]((x_1, y_1), (x_2, y_2)) = d(x_1, x_2) + \tilde{d}(y_1, y_2)$, $\forall x_1, x_2 \in X$ e $\forall y_1, y_2 \in Y$.

Dado $A \subset X$, diremos que $x \in A$ é um ponto interior se existe $r > 0$ tal que a bola aberta $B(x, r) \subset A$. Denotaremos por A^c o complementar do conjunto A, por $\mathrm{Int}(A)$ o conjunto dos pontos interiores de A e escreveremos $\bar{A}$ para significar o fecho de A.

Considere o espaço mensurável $(X, \mathcal{B}(X))$, onde $\mathcal{B}(X)$ corresponde a σ-álgebra de Borel de X. Denotaremos por $M(X)$ o conjunto das medidas com sinal σ-finitas em $\mathcal{B}(X)$, $M_+(X)$ o conjunto das medidas não-negativas σ-finitas em $\mathcal{B}(X)$ e por $P(X)$ o conjunto de todas as probabilidades definidas em $\mathcal{B}(X)$. O espaço $M(X)$ está equipado com a norma da variação total, $\|\mu\|_{TV} = \inf\{\mu_+[X] + \mu_-[X]\}$, onde o ínfimo é tomado sobre todas as medidas não-negativas μ_+, μ_- tais que $\mu = \mu_+ - \mu_-$.

Dada uma medida μ em $\mathcal{B}(X)$, para $p \in [1, +\infty)$ o conjunto $\mathcal{L}^p(d\mu) = \{f : X \to \mathbb{R} \cup \{+\infty\}; \int |f|^p d\mu < \infty\}$ representará o espaço de Lebesgue de ordem p em relação a μ, com a usual identificação das funções que coincidem em quase toda parte.

O espaço das funções contínuas e limitadas definidas no espaço métrico X e tomando valores em $\mathbb{R}$ será denotado por $C_b(X)$ e equipado com a norma do supremo, $\|u\|_\infty = \sup_{x\in X}[u(x)]$. O conjunto de todas as funções Lipschitz definidas no espaço métrico X e tomando valores em $\mathbb{R}$ será denotado por $\mathrm{Lip}(X)$ e equipado com a pseudo-norma $\|\varphi\|_{\mathrm{Lip}} = \sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{d(x,y)}$. Note que $\|\varphi\|_{\mathrm{Lip}}$ é a menor dentre as possíveis constantes de Lipschitz para φ.

Definição 1.1.1. Sejam $(X, \mathcal{A}^X, \mu)$ e $(Y, \mathcal{A}^Y, \nu)$ dois espaços de probabilidade. Definiremos $\Pi(\mu, \nu)$ como sendo o conjunto de todas as medidas não-negativas π em $\mathcal{A}^X \times \mathcal{A}^Y$ que satisfazem $\forall A \in \mathcal{A}^X$, $B \in \mathcal{A}^Y$

$$\pi(A \times Y) = \mu(A) \quad \text{e} \quad \pi(X \times B) = \nu(B). \tag{1.1}$$

Observe que, como consequência imediata de (1.1), temos que $\pi(X \times Y) = \mu(X) = \nu(Y) = 1$ e portanto π é uma probabilidade em $\mathcal{A}^X \times \mathcal{A}^Y$.

1.2 Resultados

A seguir, provaremos um lema que será utilizado no decorrer da dissertação.

Lema 1.2.1. *Sejam $(X, \mathcal{A}^X, \mu)$ e $(Y, \mathcal{A}^Y, \nu)$ dois espaços de probabilidade. $\pi \in \Pi(\mu, \nu)$ se, e somente se, é uma medida não-negativa em $\mathcal{A}^X \times \mathcal{A}^Y$ tal que $\forall(\varphi, \psi) \in \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$, π satisfaz*

$$\int_{X\times Y} [\varphi(x) + \psi(y)] d\pi = \int_X \varphi d\mu + \int_Y \psi d\nu. \tag{1.2}$$

Demonstração. ($\Longrightarrow$) Pela teoria da medida é suficiente mostrar para o caso em que φ e ψ são funções indicadoras de conjuntos mensuráveis. Se $\pi \in \Pi(\mu, \nu)$, então $\forall A \in \mathcal{A}^X$ e

$\forall B \in \mathcal{A}^Y$

$$\begin{aligned}
\int_{X\times Y}[1_A(x)+1_B(y)]d\pi(x,y) &= \int_{X\times Y}[1_{A\times Y}(x,y)+1_{X\times B}(x,y)]d\pi(x,y)\\
&= \pi(A\times Y)+\pi(X\times B)\\
&= \mu(A)+\nu(B)\\
&= \int_X 1_A d\mu + \int_X 1_B d\nu,
\end{aligned}$$

onde na terceira igualdade foi utilizado a definição 1.1.1.

($\Longleftarrow$) Se π é uma medida não-negativa que satisfaz (1.2), então $\forall A \in \mathcal{A}^X$ e $\forall B \in \mathcal{A}^Y$

$$\begin{aligned}
\pi(A\times Y)+\pi(X\times B) &= \int_{X\times Y}[1_{A\times Y}(x,y)+1_{X\times B}(x,y)]d\pi(x,y)\\
&= \int_{X\times Y}[1_A(x)+1_B(y)]d\pi(x,y)\\
&= \int_X 1_A d\mu + \int_X 1_B d\nu\\
&= \mu(A)+\nu(B),
\end{aligned}$$

onde a terceira igualdade é justificada pela identidade (1.2). Em particular, escolhendo $B = \emptyset$ segue que

$$\pi(A\times Y) = \mu(A),\ \forall A \in \mathcal{A}^X.$$

Analogamente, $\pi(X\times B) = \nu(B),\ \forall B \in \mathcal{A}^Y$, e portanto $\pi \in \Pi(\mu,\nu)$. □

Definição 1.2.2. Um espaço métrico completo e separável é dito um espaço *polonês*. Um espaço topológico é dito de *Hausdorff* se para qualquer par de pontos distintos p_1 e p_2 existem dois abertos disjuntos U_1 e U_2 tais que $p_1 \in U_1$ e $p_2 \in U_2$.

Em particular espaços métricos são espaços de Hausdorff.

Definição 1.2.3. Seja μ uma medida na σ-álgebra de Borel de um espaço topológico de Hausdorff X.

(i) μ é dita *regular interior* se $\mu(B) = \sup\{\mu(K);\ K \text{ compacto},\ K \subset B\}$ para todo boreliano B.

(ii) μ é dita *regular exterior* se $\mu(B) = \inf\{\mu(V);\ V \text{ aberto},\ V \supset B\}$ para todo boreliano B.

(iii) μ é dita *regular* se é regular interior e exterior.

Segundo **Billigsley** [3], teorema 1.1 da página 7, toda probabilidade $\mu \in P(X)$, onde X é um espaço métrico, é regular.

Definição 1.2.4. Sejam X um espaço métrico e $\mathcal{P}$ uma classe de subconjuntos de X.

(i) A *σ-álgebra gerada por* $\mathcal{P}$, denotada por $\sigma(\mathcal{P})$, é definida como a intersecção de todas as σ-álgebras que contém $\mathcal{P}$.

(ii) A classe $\mathcal{P}$ é um *π-system* se é fechada para intersecções finitas, isto é, se para quaisquer $A, B \in \mathcal{P}$, $A \cap B \in \mathcal{P}$.

Enunciaremos agora uma proposição de teoria da medida que não será demonstrada, para maiores detalhes ver **Billingsley** [4], teorema 3.3 da página 38, segunda edição.

Proposição 1.2.5. *Seja X um espaço métrico e $\mathcal{P}$ um π-system. Se μ_1 e μ_2 são probabilidades em $\sigma(\mathcal{P})$, tais que $\mu_1(A) = \mu_2(A)$ para todo $A \in \mathcal{P}$, então $\mu_1(A) = \mu_2(A)$ para todo $A \in \sigma(\mathcal{P})$.*

Em seguida, provaremos uma outra versão do lema 1.2.1.

Lema 1.2.6. *Sejam (X, d) e $(Y, \tilde{d})$ espaços métricos, $\mu \in P(X)$ e $\nu \in P(Y)$. Então $\pi \in \Pi(\mu, \nu)$ se, e somente se, $\pi \in P(X \times Y)$ e $\forall(\varphi, \psi) \in C_b(X) \times C_b(Y)$, π satisfaz a igualdade (1.2).*

Demonstração. ($\Longrightarrow$) Basta observar que $C_b(X) \times C_b(Y) \subset \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$, $\mathcal{B}(X) \times \mathcal{B}(Y) = \mathcal{B}(X \times Y)$ e aplicar o lema 1.2.1.

($\Longleftarrow$) Faremos a prova em 4 etapas.

Afirmação 1. Sejam (X, d) um espaço métrico e $A \subset X$ um aberto. Então existe uma sequência $(f_n)_{n\geq 1}$ de funções contínuas não-negativas, e uniformemente limitadas por 1, tal que $\lim_{n\to\infty} f_n(x) = 1_A(x), \forall x \in X$.

Prova da afirmação 1. De fato, defina $A_n = \{x \in A : d(x, \partial A) \geq 1/n\}$ e

$$f_n(x) = \max\{1 - nd(x, A_n), 0\}.$$

Observe que, para todo n natural, f_n é uma função contínua, já que a função $1 - nd(x, A_n)$ é contínua e a parte positiva de uma função contínua é contínua. Além disso,

$$0 \leq f_n(x) = \max\{1 - nd(x, A_n), 0\} \leq 1,$$

o que mostra que f_n é não-negativa e limitada por 1 para todo n natural.

Note que se $x \notin A$, então $d(x, A_n) \geq 1/n$, o que implica $f_n(x) = 0$. Logo, $f_n(x) \leq 1_A(x)$. Por outro lado, se $x \in A_n$ é imediato que $f_n(x) = 1$. Assim segue que

$$1_{A_n}(x) \leq f_n(x) \leq 1_A(x), \ \forall x \in X.$$

Afirmamos que

$$\lim_{n\to\infty} 1_{A_n}(x) = 1_A(x). \tag{1.3}$$

Com efeito, se $x \notin A$ então $x \notin A_n$ para todo n. Assim,

$$\forall n \in \mathbb{N}, \ 1_{A_n}(x) = 0 = 1_A(x)$$

e vale (1.3). Por outro lado, se $x \in A$, seja $r > 0$ tal que $d(x, \partial A) = r$. Tomando n_0 suficientemente grande tal que $1/n_0 < r$, temos que $\forall n \geq n_0, \ d(x, \partial A) > 1/n$. Portanto, $x \in A_n, \ \forall n \geq n_0$, o que implica

$$1_{A_n}(x) = 1 = 1_A(x), \ \forall n \geq n_0.$$

Neste caso, vemos novamente que a igualdade (1.3) é verdadeira.

Afirmação 2. Sejam (X, d) e $(Y, \tilde{d})$ dois espaços métricos, e $A \subset X$ e $B \subset Y$ abertos. Mostraremos que existe uma sequência de funções $(\varphi_n)_{n\geq 1}$ contínuas e uniformemente limitadas por 2, definidas em $X \times Y$ tal que $\lim_{n\to\infty} \varphi_n(x, y) = 1_A(x) + 1_B(y), \ \forall (x, y) \in X \times Y$.

Prova da afirmação 2. Pela afirmação 1, existem sequências $(f_n)_{n\geq 1}$ e $(g_n)_{n\geq 1}$ de funções contínuas uniformemente limitadas por 1 tais que

$$\lim_{n\to\infty} f_n(x) = 1_A(x) \quad \text{e} \quad \lim_{n\to\infty} g_n(y) = 1_B(y).$$

Defina $\varphi_n : X \times Y \longrightarrow \mathbb{R}$ por $\varphi_n(x,y) = f_n(x) + g_n(y)$. Dados $(a,b) \in X \times Y$ e $\varepsilon > 0$, pela continuidade de f_n e g_n existe um $\delta > 0$ tal que

$$d(x,a) < \delta \Longrightarrow |f_n(x) - f_n(a)| < \varepsilon/2$$

e

$$\tilde{d}(y,b) < \delta \Longrightarrow |g_n(y) - g_n(b)| < \varepsilon/2.$$

Assim,

$$\begin{aligned} [d+\tilde{d}]((x,y),(a,b)) &< \delta \Longrightarrow d(x,a) + \tilde{d}(y,b) < \delta \\ \Longrightarrow |\varphi_n(x,y) - \varphi_n(a,b)| &\leq |f_n(x) - f_n(a)| + |g_n(y) - g_n(b)| < \varepsilon. \end{aligned}$$

Logo, φ_n é contínua. Além disso, para todo n, $|\varphi_n(x,y)| \leq |f_n(x)| + |g_n(y)| \leq 2$ e $\lim_{n\to\infty} \varphi_n(x,y) = 1_A(x) + 1_B(y)$.

Afirmação 3. Sejam X e Y espaços métricos. Para todo $A \subset X$ e $B \subset Y$ abertos

$$\pi(A \times Y) = \mu(A) \quad \text{e} \quad \pi[X \times B] = \nu(B).$$

.

Prova da afirmação 3. De fato,

$$\begin{aligned} \pi(A \times Y) + \pi(X \times B) &= \int [1_A(x) + 1_B(y)] d\pi(x,y) \\ &= \int \lim_{n\to\infty} \varphi_n d\pi(x,y) \\ &= \lim_{n\to\infty} \int \varphi_n d\pi(x,y) \\ &= \lim_{n\to\infty} \int f_n d\mu + \lim_{n\to\infty} \int g_n d\nu \\ &= \int 1_A d\mu + \int 1_B d\nu = \mu(A) + \nu(B), \end{aligned}$$

onde a segunda igualdade vem da afirmação 2, a terceira e quinta igualdades do Teorema da Convergência Dominada, e a quarta igualdade vem da demonstração da afirmação 2 e da hipótese de π satisfazer (1.2.1).

Portanto, como π é uma probabilidade, fazendo $B = Y$ obtemos $\pi(A \times Y) = \mu(A)$, $\forall A \subset X$ aberto, da mesma forma, fazendo $A = X$ obtemos $\pi(X \times B) = \nu(B)$, $\forall B \subset Y$ aberto.

Afirmação 4. Concluiremos o lema.

Prova da afirmação 4. Defina $\pi_X : \mathcal{B}(X) \to [0, 1]$ dada por

$$\pi_X(A) = \pi(A \times Y),$$

π_X é claramente uma probabilidade. Além disso, dados quaisquer subconjuntos abertos A_1 e A_2 de X, $A_1 \cap A_2$ é um subconjunto aberto de X, portanto a classe $\mathcal{P}$ dos subconjuntos abertos de X é um π-system.

Como, pela afirmação 3, para todo $A \subset X$ aberto $\pi_X(A) = \pi(A \times Y) = \mu(A)$ o lema 1.2.5 acima, garante que

$$\pi_X(A) = \mu(A), \ \forall A \in \sigma(\mathcal{P}).$$

Mas, $\sigma(\mathcal{P}) = \mathcal{B}(X)$, então $\mu(A) = \pi_X(A) = \pi(A \times Y)$ para todo $A \in \mathcal{B}(X)$.

Analogamente, definindo $\pi_Y : \mathcal{B}(Y) \to [0, 1]$ dada por

$$\pi_Y(B) = \pi(X \times B),$$

concluiremos que $\nu(B) = \pi_Y(B) = \pi(X \times B)$ para todo $B \in \mathcal{B}(Y)$. □

1.3 Dualidade Geral

Nesta seção faremos a prova do Teorema da Dualidade de Kantorovich.

Definição 1.3.1. Uma função φ definida em um espaço métrico X e tomando valores em $\mathbb{R} \cup \{+\infty\}$ é dita *semicontínua inferiormente* se para toda sequência (x_n) em X tal que $x_n \to x$ quando $n \to \infty$, temos

$$\liminf_{n \to \infty} \varphi(x_n) \geq \varphi(x).$$

Um dos objetivos principais nesta dissertação é a demonstração do seguinte teorema, no caso X e Y compactos e c contínua:

Teorema 1.3.2 (Dualidade de Kantorovich). *Sejam X e Y espaços poloneses, $\mu \in P(X)$ e $\nu \in P(Y)$ e seja $c : X \times Y \to \mathbb{R}^+ \cup \{+\infty\}$ uma função custo semicontínua inferiormente.*

Sempre que $\pi \in P(X \times Y)$ e $(\varphi, \psi) \in \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$, defina

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

Defina $\Pi(\mu,\nu)$ como o conjunto de todas as probabilidades de Borel π em $X \times Y$ que satisfazem $\forall A \in \mathcal{B}(X)$, $B \in \mathcal{B}(Y)$,

$$\pi[A \times Y] = \mu(A) \quad e \quad \pi[X \times B] = \nu(B)$$

e defina Φ_c como o conjunto de todas as funções mensuráveis $(\varphi, \psi) \in \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$ satisfazendo

$$\varphi(x) + \psi(y) \leq c(x,y) \tag{1.4}$$

para $d\mu$-quase todo ponto em X e $d\nu$-quase todo ponto em Y.

Então,

$$\inf_{\Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi). \tag{1.5}$$

Ainda, o ínfimo do lado esquerdo de (1.5) é atingido, e o valor do supremo não muda se restringirmos a definição de Φ_c às funções contínuas e limitadas.

Observação 1.3.3. A priori não é claro que o valor do supremo em J não mude com a restrição na definição de Φ_c às funções contínuas e limitadas, uma vez que não é claro que os pares de funções em $\mathcal{L}^1$ satisfazendo (1.4) possam ser aproximadas por pares de funções contínuas limitadas satisfazendo a mesma desigualdade. Contudo, a partir da proposição 1.3.5 abaixo, tal afirmação ficará mais clara. Quando quisermos enfatizar a distinção entre essas definições, escreveremos informalmente $\Phi_c \cap C_b$ e $\Phi_c \cap \mathcal{L}^1$.

A seguir, introduziremos um exemplo que servirá como uma interpretação informal para o teorema 1.3.2.

Exemplo 1.3.4 (O problema do carregador)**.** Suponha que você seja empresário, e que precise transferir uma grande quantidade de carvão das suas minas para as suas fábricas. Você pode contratar caminhões para fazer o transporte, mas precisa pagá-los $c(x, y)$ por cada tonelada de carvão que é transportada de um lugar x para um lugar y. A quantidade de carvão que você pode extrair de cada mina e a quantidade que cada fábrica precisa receber estão fixadas.

Como você está tentando resolver o problema de Monge-Kantorovich associado, a fim de minimizar o preço que você tem que pagar, um matemático vem até você e diz "Meu amigo, deixe-me cuidar disso para você: eu enviarei todo o seu carvão com meus próprios caminhões e você não terá que se preocupar com o que vai aonde. Vou apenas fixar um preço $\varphi(x)$ para carregar uma tonelada de carvão de um lugar x e um preço $\phi(y)$ para descarregar em um destino y. Vou definir os preços de tal maneira que você deixará que eu faça todo o transporte necessário! De fato, você pode checar que para qualquer x, y a soma $\varphi(x) + \phi(y)$ será sempre menor ou igual que o custo $c(x, y)$ (para atingir esse objetivo, estou disposto a dar compensações financeiras para alguns lugares, na forma de preços negativos!)".

É claro que você aceita a proposta. Agora, o que a dualidade de Kantorovich afirma é que, se este carregador for inteligente o suficiente, ele pode organizar os preços de tal forma que você pagará a ele tanto quanto você iria pagar com o método de transporte inicial.

Como primeiro passo para a demonstração do teorema 1.3.2 provaremos a parte fácil da dualidade de Kantorovich

Proposição 1.3.5. *Sob as mesmas hipóteses do teorema 1.3.2,*

$$\sup_{\Phi_c \cap C_b} J(\varphi, \psi) \leq \sup_{\Phi_c \cap \mathcal{L}^1} J(\varphi, \psi) \leq \inf_{\Pi(\mu,\nu)} I[\pi]. \tag{1.6}$$

Demonstração. Como $C_b(X) \times C_b(Y) \subset \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$, a desigualdade da esquerda em (1.6) é trivial. Portanto basta mostrar a desigualdade da direita.

Sejam $(\varphi, \psi) \in \Phi_c \cap \mathcal{L}^1$ e $\pi \in \Pi(\mu, \nu)$. Então pelo lema 1.2.1,

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

Afirmamos que a desigualdade (1.4) vale para $d\pi$-quase toda parte: de fato, sejam $N_X \subset X$ e $N_Y \subset Y$ conjuntos mensuráveis tais que $\mu[N_X] = 0$, $\nu[N_Y] = 0$ e a desigualdade (1.4) valha para todo $(x, y) \in N_X^c \times N_Y^c$. Uma vez que π tem marginais μ e ν, podemos escrever $\pi[N_X \times Y] = \nu[N_X] = 0,\ \pi[X \times N_Y] = \nu[N_Y] = 0$. Portanto

$$\pi[(N_X^c \times N_Y^c)^c] \leq \pi[N_X \times Y] + \pi[X \times N_Y] = 0.$$

Logo $\varphi(x) + \psi(y) \leq c(x, y)$ em dπ-quase toda parte e como consequência

$$J(\varphi, \psi) = \int_{X\times Y} [\varphi(x) + \psi(y)] d\pi \leq \int_{X\times Y} c(x, y) d\pi(x, y) = I[\pi]. \tag{1.7}$$

Assim,

$$\sup_{\Phi_c \cap \mathcal{L}^1} J(\varphi, \psi) \leq I[\pi] \quad \forall \pi \in \Pi(\mu, \nu).$$

O que implica que

$$\sup_{\Phi_c \cap \mathcal{L}^1} J(\varphi, \psi) \leq \inf_{\Pi(\mu,\nu)} I[\pi].$$

□

Agora introduziremos algumas ferramentas necessárias para a demonstração do teorema 1.3.2.

Definição 1.3.6. Seja E um espaço vetorial normado, um conjunto $C \subset E$ é dito *convexo* quando para quaisquer $x, y \in C$ e $t \in [0, 1]$

$$tx + (1 - t)y \in C.$$

Precisamos do seguinte resultado para demonstrar o teorema 1.3.10 abaixo, cuja demonstração encontra-se no apêndice.

Lema 1.3.7. *Seja E um espaço vetorial normado. Se $C \subset E$ um conjunto convexo, então $Int(C)$ é convexo. Se além disso $Int(C) \neq \emptyset$, então $\overline{C} = \overline{Int(C)}$.*

Definição 1.3.8. Dada uma função φ definida em um espaço métrico X e tomando valores em $\mathbb{R} \cup \{+\infty\}$.

(i) O *domínio* de φ, $D(\varphi)$, é dado por

$$D(\varphi) = \{x \in X; \varphi(x) < +\infty\}.$$

(ii) O *epigráfico* de φ é o conjunto

$$\text{epi}\varphi = \{(x, \lambda) \in X \times \mathbb{R}; \varphi(x) \leq \lambda\}.$$

Definição 1.3.9. Seja E um espaço vetorial normado.

(i) O *dual topológico* de E, denotado por E^* é o conjunto

$$E^* = \{f : E \to \mathbb{R}; f \text{ é contínua e linear}\}.$$

(ii) $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ é dita *convexa* se para quaisquer $(x, y, \lambda) \in E \times E \times [0, 1]$

$$\varphi(\lambda x + (1 - \lambda)y) \leq \lambda\varphi(x) + (1 - \lambda)\varphi(y).$$

(iii) Dada uma função Θ convexa em E tomando valores em $\mathbb{R} \cup \{+\infty\}$, a*transformada de Legendre-Fenchel* de Θ é a função Θ^* definida no dual topológico E^* de E, dada por

$$\Theta^*(f) = \sup_{x \in E}[f(x) - \Theta(x)]. \tag{1.8}$$

(iv) Um *hiperplano* $H \subset E$ é um conjunto da forma

$$H = \{x \in E;\ f(x) = \alpha\}$$

onde f é um funcional linear não identicamente nulo, e $\alpha \in \mathbb{R}$.

(v) Dados A e B subconjuntos de E, o hiperplano $H = [f = \alpha]$ *separa* A e B quando

$$f(x) \leq \alpha,\ \forall x \in A \quad \text{e} \quad f(x) \geq \alpha,\ \forall x \in B$$

A seguir provaremos o resultado central para a demostração do teorema 1.3.2. Usaremos para isso, alguns resultados conhecidos, dos quais o segundo é um resultado clássico de Análise Funcional. Sua respectivas demonstrações podem ser encontradas em **Rockafellar** [8], teorema 4.1 da página 25, e **Brezis** [2], teorema 1.6 da página 5.

- O epigráfico de uma função convexa é um conjunto convexo;

- **A primeira forma geométrica de Hahn-Banach:** Sejam dois conjuntos convexos e não-vazios A e B contidos no espaço vetorial E, tais que $A \cap B = \emptyset$. Suponha A aberto, então existe um hiperplano fechado que separa A e B.

Teorema 1.3.10 (**Dualidade de Fenchel-Rockafellar**). *Seja E um espaço vetorial normado, E^* o seu espaço dual topológico, e Θ, Ξ duas funções convexas definidas em E tomando valores em $\mathbb{R} \cup \{+\infty\}$. Sejam Θ^*, Ξ^* as transformadas de Legendre-Fenchel de Θ, Ξ respectivamente. Suponha que exista um $x_0 \in E$ tal que $x_0 \in D(\Theta) \cap D(\Xi)$ e Θ é contínua em x_0.*

Então,

$$\inf_{x \in E}[\Theta(x) + \Xi(x)] = \sup_{f \in E^*}[-\Theta^*(-f) - \Xi^*(f)] \tag{1.9}$$

Além disso, o supremo em (1.9) é atingido por, ao menos, um elemento de E^*.

Note que as hipóteses do teorema não excluem o caso em que

$$\inf_{x \in E}[\Theta(x) + \Xi(x)] = -\infty\,.$$

Demonstração. Definamos

$$\begin{aligned} a &= \inf_{x \in E}[\Theta(x) + \Xi(x)]\,, \\ b &= \sup_{f \in E^*}[-\Theta^*(-f) - \Xi^*(f)]. \end{aligned}$$

Afirmação 1. $b \leq a$.

Prova da afirmação 1. Segue da definição de transformada de Legendre-Fenchel que

$$\begin{aligned} b &= \sup_{f \in E^*} \left\{ -\sup_{x \in E}[-f(x) - \Theta(x)] - \sup_{y \in E}[f(y) - \Xi(y)] \right\} \\ &= \sup_{f \in E^*} \left\{ \inf_{x \in E}[f(x) + \Theta(x)] + \inf_{y \in E}[-f(y) + \Xi(y)] \right\} \\ &= \sup_{f \in E^*} \left\{ \inf_{x,y \in E}[f(x-y) + \Theta(x) + \Xi(y)] \right\}. \end{aligned}$$

Como

$$\inf_{x,y \in E}[f(x-y) + \Theta(x) + \Xi(y)] \leq \inf_{x \in E}[\Theta(x) + \Xi(x)] = a, \forall f \in E^*,$$

tem-se $b \leq a$, o que prova a afirmação 1.

Se $a = -\infty$, então $-\infty = a \leq b$ para todo $b \in \mathbb{R}$. Logo $a = b$ e vale (1.9). Assim, no que segue, podemos assumir que $a \in \mathbb{R}$.

Seja $C = \mathrm{epi}\Theta$.

Afirmação 2. O ponto $(x_0, \Theta(x_0) + 1) \in \mathrm{Int}(C)$.

Prova da afirmação 2. Suponha que d é a distância em E e defina a distância

$$\tilde{d}((x_1, y_1), (x_2, y_2)) = d(x_1, x_2) + |y_1 - y_2|$$

em $E \times \mathbb{R}$.

Seja um $\varepsilon > 0$ pequeno. Pela continuidade de Θ em x_0 existe um $\delta > 0$ pequeno (de forma a termos $1 - \delta > \varepsilon$) tal que

$$|\Theta(x) - \Theta(x_0)| < \varepsilon, \forall x \in B(x_0, \delta) = \{x \in E;\ d(x, x_0) < \delta\}.$$

Seja agora

$$\tilde{B} = \{(x, y) \in E \times \mathbb{R};\ \tilde{d}\left((x, y), (x_0, \Theta(x_0) + 1)\right) < \delta\}.$$

Se $(x, y) \in \tilde{B}$ então

$$y > \Theta(x_0) + 1 - \delta > \Theta(x_0) + \varepsilon > \Theta(x),$$

ou seja, temos $(x, y) \in C$. Logo o aberto $\tilde{B}$, que contém $(x_0, \Theta(x_0) + 1)$, está contido em C, e portanto finalizamos a demonstração da afirmação 2.

Agora sejam $A = \mathrm{Int}(C)$ e

$$B = \{(x, \lambda) \in E \times \mathbb{R} : \lambda \leq a - \Xi(x)\} = -\mathrm{epi}(\Xi - a).$$

Pelo lema 1.3.7 segue que A é convexo, e pela convexidade de Ξ que B é convexo. É claro que B é não-vazio, uma vez que $\Xi(x_0) < +\infty$. Ainda, A é não-vazio como consequência da afirmação 2.

Provaremos que $A \cap B = \emptyset$. Com efeito, se $(x, \lambda) \in A$ então $\lambda > \Theta(x)$. Por outro lado, pela definição de a temos que

$$\Theta(x) \geq a - \Xi(x), \forall x \in E$$

de forma que $(x, \lambda) \notin B$.

Assim, pela primeira forma geométrica de Hahn-Banach, existe uma hiperplano fechado H que separa A e B, isto é, existem $\alpha \in \mathbb{R}$ e $\Phi \in (E \times \mathbb{R})^*$ tais que

$$\Phi((x, \lambda)) \geq \alpha, \ \forall (x, \lambda) \in A, \tag{1.10}$$

$$\Phi((x, \lambda)) \leq \alpha, \ \forall (x, \lambda) \in B. \tag{1.11}$$

Usando a continuidade de Φ podemos substituir em (1.10), A por $\overline{A}$. Mas pelo lema 1.3.7 $\overline{A} = \overline{C}$. Note que a função $x \longmapsto \Phi(x, 0)$ é um funcional linear contínuo em E, e assim $\Phi(x, 0) = f(x)$ para algum $f \in E^*$. Definindo $k = \Phi(0, 1)$ segue que

$$\Phi(x, \lambda) = f(x) + k\lambda \ \ \forall (x, \lambda) \in E \times \mathbb{R}. \tag{1.12}$$

Assim, temos

$$f(x) + k\lambda \geq \alpha, \ \forall (x, \lambda) \in C, \tag{1.13}$$

$$f(x) + k\lambda \leq \alpha, \ \forall (x, \lambda) \in B. \tag{1.14}$$

Escolhendo $x = x_0$ e fazendo $\lambda \to +\infty$ em (1.13), vemos que $k \geq 0$.

Afirmação 3. $k > 0$.

Prova da afirmação 3. Assuma por contradição que $k = 0$, como $\Phi \not\equiv 0$, segue de (1.12) que $||f||_\infty \neq 0$. Além disso, por (1.13) e (1.14) segue

$$f(x) \geq \alpha, \ \forall x \in D(\Theta), \tag{1.15}$$

$$f(x) \leq \alpha, \ \forall x \in D(\Xi). \tag{1.16}$$

Mas, pela continuidade de Θ em x_0, existe ε_0 suficientemente pequeno tal que $B(x_0, \varepsilon_0) \subset \text{Dom}(\Theta)$, e assim de (1.15) temos

$$f(x_0 - \varepsilon_0 w) \geq \alpha \ \ \forall w \in B(0,1).$$

Note que da última desigualdade temos que

$$f(x_0) \geq \alpha + \varepsilon_0 f(w), \ \forall w \in B(0,1),$$

o que implica que $f(x_0) \geq \alpha + \varepsilon_0 ||f||$. Por outro lado, de (1.16), temos que $f(x_0) \leq \alpha$, já que $x_0 \in D(\Xi)$. Portanto obtemos $||f|| = 0$, o que é uma contradição e prova a afirmação 3.

Escolha agora $\lambda = \Theta(x)$ em (1.13), e $\lambda = a - \Xi(x)$ em (1.14), então obtemos

$$-\frac{f(x)}{k} - \Theta(x) \leq -\frac{\alpha}{k}, \quad \forall x \in E$$

e

$$\frac{f(x)}{k} - \Xi(x) \leq \frac{\alpha}{k} - a, \quad \forall x \in E,$$

de forma que

$$-\Theta^* \left(\frac{-f}{k} \right) - \Xi^* \left(\frac{f}{k} \right) \geq a.$$

Por outro lado, da definição de b, temos

$$-\Theta^* \left(\frac{-f}{k} \right) - \Xi^* \left(\frac{f}{k} \right) \leq b,$$

o que implica $a \leq b$.

Finalmente, com a afirmação 1, concluímos que

$$a = b = -\Theta^* \left(\frac{-f}{k} \right) - \Xi^* \left(\frac{f}{k} \right),$$

o que nos demonstra (1.9) e que o supremo é atingido. $\square$

Definição 1.3.11. Seja μ uma medida na σ-álgebra de Borel de um espaço topológico de Hausdorff X.

(i) μ é dita *localmente finita* se $\forall x \in X$ existe V vizinhança de x tal que $\mu(V) < \infty$.

(ii) μ é dita *medida de Radon* se é regular interior e localmente finita.

Denotaremos o conjunto das medidas de Radon de um espaço topológico de Hausdorff X como $M_R(X)$, e o conjunto das medidas de Radon não-negativas por $M_R^+(X)$.

Para a prova do teorema 1.3.2 admitiremos o seguinte resultado, que não será provado, mas que pode ser encontrado em **Fitzpatrick** [6], página 164.

Teorema 1.3.12 (**Teorema de Representação de Riesz-Markov**). *Seja X um espaço de Hausdorff compacto e $C(X)$ o espaço das funções contínuas em X tomando valores em $\mathbb{R}$, equipado com a norma do supremo e considere $M_R(X)$ equipado com a norma da variação total. Defina o funcional linear $T : M_R(X) \to [C(X)]^*$ dado por*

$$T_\mu(f) = \int_X f d\mu, \ \forall f \in C(X).$$

Então T é um isomorfismo isométrico de $M_R(X)$ em $[C(X)]^*$. Além disso, T é um funcional linear positivo se, e somente se, $\mu \in M_R^+(X)$.

Faremos a prova do teorema 1.3.2 para o caso particular em que X e Y são compactos e c é continua. Este é o primeiro passo para a demostração do teorema no caso geral, que pode ser encontrada em **Villani** [11], página 26.

Nesta demonstração, e no que segue nesta dissertação, convencionaremos que $0 \cdot (+\infty) = 0$.

Demonstração do teorema 1.3.2. Suponhamos adicionalmente que X e Y são compactos e c é contínua em $X \times Y$. Defina

$$E = C_b(X \times Y)$$

o conjunto de todas as funções contínuas, e consequentemente limitadas, em $X \times Y$ tomando valores em $\mathbb{R}$, equipado com a norma usual do supremo $\| \, . \, \|_\infty$. Note que X e

Y compactos implica $X \times Y$ compacto. Logo, pelo teorema de representação de Riesz-Markov acima, o dual topológico de E pode ser identificado com o espaço das medidas de Radon

$$E^* = M_R(X \times Y)$$

normado pela variação total, isto é, $\| \pi \|_{VT} = \inf\{\pi_+(X) + \pi_-(X)\}$.

Defina $\Theta : C_b(X \times Y) \longrightarrow \mathbb{R} \cup \{+\infty\}$ por

$$\Theta(u) = \begin{cases} 0 \ \text{ se } u(x,y) \geq -c(x,y), \ \forall (x,y) \in X \times Y, \\ \\ +\infty \ \text{ caso contrário,} \end{cases}$$

onde c é a função custo, e $\Xi : C_b(X \times Y) \longrightarrow \mathbb{R} \cup \{+\infty\}$ por

$$\Xi(u) = \begin{cases} \int_X \varphi d\mu + \int_Y \psi d\nu \ \text{ se } u(x,y) = \varphi(x) + \psi(y), \text{ com } (\varphi, \psi) \in C_b \times C_b, \\ \\ +\infty \ \text{ caso contrário.} \end{cases}$$

Note que Ξ está bem definida, já que, se

$$\varphi(x) + \psi(y) = u(x,y) = \tilde{\varphi}(x) + \tilde{\psi}(y), \forall (x,y) \in X \times Y,$$

então

$$\varphi(x) - \tilde{\varphi}(x) = \tilde{\psi}(y) - \psi(y).$$

Logo, existe $s \in \mathbb{R}$ tal que

$$\varphi(x) = \tilde{\varphi}(x) + s \quad \text{e} \quad \psi(y) = \tilde{\psi}(y) - s.$$

Assim,

$$\begin{aligned} \int \varphi d\mu + \int \psi d\nu &= \int [\tilde{\varphi} + s] d\mu + \int [\tilde{\psi} - s] d\nu \\ &= \int \tilde{\varphi} d\mu + s + \int \tilde{\psi} d\nu - s \\ &= \int \tilde{\varphi} d\mu + \int \tilde{\psi} d\nu. \end{aligned}$$

Afirmamos que Θ e Ξ são funções convexas. De fato, observe que para quaisquer $(u,v,\lambda) \in E \times E \times [0,1]$ temos

- se $u(x,y) \geq -c(x,y)$ e $v(x,y) \geq -c(x,y)$ para todo $(x,y) \in X \times Y$, então $\lambda u(x,y) + (1-\lambda)v(x,y) \geq -c(x,y)$, portanto $\Theta(\lambda u + (1-\lambda)v) = 0 = \lambda\Theta(u) + (1-\lambda)\Theta(v)$.

- Suponha, sem perda de generalidade, que $u(x,y) \geq -c(x,y)$ para todo $(x,y) \in X \times Y$, e que exista $(x_0, y_0) \in X \times Y$ tal que $v(x_0,y_0) < -c(x_0,y_0)$. Então, $\Theta(v) = +\infty$, por isso, se $\lambda < 1$,

$$\lambda\Theta(u) + (1-\lambda)\Theta(v) = +\infty \geq \Theta(\lambda u + (1-\lambda)v),$$

  onde a desigualdade é justificada pela definição de Θ. Se $\lambda = 1$, como $0 \cdot (+\infty) = 0$, segue que $\Theta(\lambda u + (1-\lambda)v) = \Theta(u) = \lambda\Theta(u) + (1-\lambda)\Theta(v)$.

O que mostra que Θ é convexa. Da mesma forma,

- se $u(x,y) = \varphi_1(x) + \psi_1(y)$ e $v(x,y) = \varphi_2(x) + \psi_2(y)$ com $(\varphi_1, \psi_1), (\varphi_2, \psi_2) \in C_b(X) \times C_b(Y)$ então

$$\begin{aligned} \lambda u(x,y) + (1-\lambda)v(x,y) &= \lambda[\varphi_1(x) + \psi_1(y)] + (1-\lambda)[\varphi_2(x) + \psi_2(y)] \\ &= [\lambda\varphi_1 + (1-\lambda)\varphi_2](x) + [\lambda\psi_1 + (1-\lambda)\psi_2](y). \end{aligned}$$

  Como $(\lambda\varphi_1 + (1-\lambda)\varphi_2, \lambda\psi_1 + (1-\lambda)\psi_2) \in C_b(X) \times C_b(Y)$ temos que

$$\begin{aligned} &\Xi(\lambda u + (1-\lambda)v) = \\ &\quad = \int_X [\lambda\varphi_1 + (1-\lambda)\varphi_2] d\mu + \int_Y [\lambda\psi_1 + (1-\lambda)\psi_2] d\nu \\ &\quad = \lambda \left[\int_X \varphi_1 d\mu + \int_Y \psi_1 d\nu \right] + (1-\lambda) \left[\int_X \varphi_2 d\mu + \int_Y \psi_2 d\nu \right] \\ &\quad = \lambda\Xi(u) + (1-\lambda)\Xi(v). \end{aligned}$$

- Suponha que $u(x,y) = \varphi_0(x) + \psi_0(y)$, $\forall (x,y) \in X \times Y$ com $(\varphi_0, \psi_0) \in C_b(X) \times C_b(Y)$ e que $v \neq \varphi + \psi$, $\forall (\varphi, \psi) \in C_b(X) \times C_b(Y)$, o que implica que $\Xi(v) = +\infty$. Portando, se $\lambda = 1$ então usando novamente que $0 \cdot (+\infty) = 0$, temos $\Xi(\lambda u + (1-\lambda)v) = \Xi(u) = \lambda\Xi(u) + (1-\lambda)\Xi(v)$. Agora, se $\lambda < 1$, então pela definição de Ξ,

$$\lambda\Xi(u) + (1-\lambda)\Xi(v) = +\infty \geq \Xi(\lambda u + (1-\lambda)v).$$

O que demonstra que Ξ também é convexa.

Além disso, se $z_0 \equiv 1$ então $1 = z_0 > -c(x,y)$, pois c é não-negativa, portanto $\Theta(z_0) = 0 < \infty$. Ainda, Θ é contínua em z_0, uma vez que se $\|z_0 - u\| < 1/2$ então $\Theta(u) = 0$.

Portanto as hipóteses do teorema 1.3.10 são satisfeitas, e a identidade (1.9) se verifica.

Observe que

$$\begin{aligned}
&\inf_{u\in E}[\Theta(u) + \Xi(u)] \\
&\quad = \inf\left\{\int_X \varphi d\mu + \int_Y \psi d\nu;\ \varphi(x) + \psi(y) \geq -c(x,y),\ (\varphi,\psi) \in C_b(X) \times C_b(Y)\right\} \\
&\quad = \inf\left\{\int_X \varphi d\mu + \int_Y \psi d\nu;\ -(\varphi(x) + \psi(y)) \leq c(x,y),\ (\varphi,\psi) \in C_b(X) \times C_b(Y)\right\} \\
&\quad = \inf\left\{-\int_X \varphi d\mu - \int_Y \psi d\nu;\ \varphi(x) + \psi(y) \leq c(x,y),\ (\varphi,\psi) \in C_b(X) \times C_b(Y)\right\} \\
&\quad = -\sup\left\{\int_X \varphi d\mu + \int_Y \psi d\nu;\ \varphi(x) + \psi(y) \leq c(x,y),\ (\varphi,\psi) \in C_b(X) \times C_b(Y)\right\} \\
&\quad = -\sup\left\{J(\varphi,\psi); (\varphi,\psi) \in \Phi_c \cap C_b\right\}. \qquad (1.17)
\end{aligned}$$

Calculemos as transformadas de Legendre-Fenchel de Θ e Ξ, inicialmente para qualquer $\pi \in M_R(X \times Y)$: a definição de Θ implica

$$\begin{aligned}
\Theta^*(-\pi) &= \sup_{u\in E}\{-\pi(u) - \Theta(u)\} \\
&= \sup_{u\in E}\left\{-\int u(x,y)d\pi(x,y);\ -u(x,y) \leq c(x,y)\right\} \\
&= \sup_{u\in E}\left\{\int u(x,y)d\pi(x,y);\ u(x,y) \leq c(x,y)\right\}.
\end{aligned}$$

- Se π é uma medida de Radon com sinal com parte negativa não-nula, então existe uma função $v \in C_b(X \times Y)$ tal que $\int v d\pi > 0$ (esse fato é verdadeiro pois, se não fosse, o funcional linear $T : v \in C_b(X \times Y) \mapsto \int v d\pi$ seria um funcional linear positivo e, pelo teorema 1.3.12, a medida que representa tal funcional seria não-negativa). Assim, tomando $u = \lambda v$, e fazendo $\lambda \to +\infty$, vemos que $\sup_{u \in E} \left\{ \int u(x,y) d\pi(x,y);\ u(x,y) \leq c(x,y) \right\} = +\infty$.

- Por outro lado, se $\pi \in M_R^+(X \times Y)$, como $c \in C_b(X \times Y)$ temos que o supremo é, evidentemente, $\int c d\pi = I[\pi]$.

Assim,

$$\Theta^*(-\pi) = \begin{cases} I[\pi] \text{ se } \pi \in M_R^+(X \times Y) \\ \\ +\infty \text{ caso contrário.} \end{cases} \tag{1.18}$$

Analogamente, a definição de Ξ faz com que

$$\begin{aligned} \Xi^*(\pi) &= \sup_{u \in E} \{\pi(u) - \Xi(u)\} \\ &= \sup_{u \in E} \left\{ \begin{array}{r} \int u(x,y) d\pi(x,y) - \int \varphi d\mu - \int \psi d\nu; \\ u(x,y) = \varphi(x) + \psi(y) \text{ e } (\varphi, \psi) \in C_b \times C_b \end{array} \right\} \\ &= \sup_{(\varphi,\psi) \in C_b \times C_b} \left\{ \int [\varphi(x) + \psi(y)] d\pi(x,y) - \int \varphi d\mu - \int \psi d\nu \right\}. \end{aligned}$$

- Se $\int [\varphi(x) + \psi(y)] d\pi(x,y) = \int \varphi d\mu + \int \psi d\nu,\ \forall (\varphi, \psi) \in C_b(X) \times C_b(Y)$ então o supremo é nulo.

- Caso contrário, existe $(\varphi_0, \psi_0) \in C_b(X) \times C_b(Y)$ tal que, sem perda de generalidade, $\int [\varphi_0(x) + \psi_0(y)] d\pi(x,y) > \int \varphi_0 d\mu + \int \psi_0 d\nu$. Então a escolha $(\varphi, \psi) = (\lambda \varphi_0, \lambda \psi_0)$ com $\lambda \to +\infty$, mostra que o supremo acima é $+\infty$.

Afim de simplificar a notação, defina

$$\begin{aligned} \Pi^*(\mu, \nu) &= \left\{ \pi \in M_R(X \times Y);\ \int [\varphi(x) + \psi(y)] d\pi(x,y) = \right. \\ &\quad \left. = \int \varphi d\mu + \int \psi d\nu \quad \forall (\varphi, \psi) \in C_b(X) \times C_b(Y) \right\} \end{aligned}$$

Acabamos de provar que

$$\Xi^*(\pi) = \begin{cases} 0 \text{ se } \pi \in \Pi^*(\mu,\nu), \\ \\ +\infty \text{ caso contrário.} \end{cases} \tag{1.19}$$

Agora, de (1.17), vemos que o lado esquerdo de (1.9) é igual a $-\sup_{\Phi_c \cap C_b} \{J(\varphi,\psi)\}$, enquanto que o lado direito, de (1.18) e (1.19), coincide com

$$\sup_{\pi \in M_R(X)} \left\{ - \begin{Bmatrix} I[\pi] \text{ se } \pi \in M_R^+(X \times Y) \\ \\ +\infty \text{ caso contrário} \end{Bmatrix} - \begin{Bmatrix} 0 \text{ se } \pi \in \Pi^*(\mu,\nu) \\ \\ +\infty \text{ caso contrário.} \end{Bmatrix} \right\}$$

ou equivalentemente,

$$\sup_{\pi \in M_R(X)} \{-I[\pi],\ \pi \in M_R^+(X \times Y) \cap \Pi^*(\mu,\nu)\}.$$

Como as medidas pertencentes a $\Pi^*(\mu,\nu)$ são probabilidades, temos que $M_R^+(X \times Y) \cap \Pi^*(\mu,\nu) = M_+(X \times Y) \cap \Pi^*(\mu,\nu)$. Logo, pelo lema 1.2.6, $M_R^+(X \times Y) \cap \Pi^*(\mu,\nu) = \Pi(\mu,\nu)$. Assim, substituindo em (1.9) o que obtivemos, deduzimos que

$$-\sup_{\Phi_c \cap C_b} J(\varphi,\psi) = -\inf_{\pi \in \Pi(\mu,\nu)} I[\pi]$$

Pela proposição 1.3.5 (lembrando que $\Phi_c \cap \mathcal{L}^1$ é outra notação para Φ_c), segue que

$$\sup_{\Phi_c} J(\varphi,\psi) = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

□

Observe que o teorema 1.3.10 afirma que $\sup\limits_{\pi \in M_R(X \times Y)} [-\Theta^*(-\pi) - \Xi^*(\pi)] = -\inf\limits_{\pi \in \Pi(\mu,\nu)} I[\pi]$ é atingido, por ao menos um elemento de $M_R(X \times Y)$, o que nos garante a existência de um plano de transporte ótimo.

Uma demonstração alternativa, que não será detalhada nesta dissertação, para a existência de um plano de transporte ótimo no caso em que c é continua e X e Y compactos, segue do fato de $\Pi(\mu,\nu)$ ser compacto e $I(\pi)$ ser contínua, portanto $\inf_{\Pi(\mu,\nu)} I(\pi)$

é atingido. Ainda, pode-se mostrar que com pouco trabalho extra adapta-se tal demonstração para o caso c semicontínua inferiormente.

Os próximos resultados mostrarão que podemos restringir o supremo na equação (1.5) a uma classe menor de funções.

Sejam X e Y espaços métricos, c uma função limitada semicontínua inferiormente definida em $X \times Y$ e φ limitada em X. Defina

$$\varphi^c(y) = \inf_{x \in X} [c(x,y) - \varphi(x)] \quad \text{e} \quad \varphi^{cc}(x) = \inf_{y \in Y} [c(x,y) - \varphi^c(y)].$$

O par $(\varphi^{cc}, \varphi^c)$ é chamado de funções *c-côncavas conjugadas.*

Proposição 1.3.13 (Funções c-côncavas). *Sejam (X, d) e $(Y, \tilde{d})$ espaços métricos, c uma função limitada semicontínua inferiormente em $X \times Y$, e φ limitada em X. Então podemos escrever φ^c como o supremo de uma sequência não-decrescentes ψ_l de funções uniformemente contínuas. Em particular, φ^c é mensurável.*

A demonstração da proposição acima encontra-se no apêndice. Observe que é imediato desta proposição o fato de que φ^{cc} é mensurável.

Lema 1.3.14. *Seja c como no teorema 1.3.2. Suponha também que c seja limitada, então*

$$\sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi) = \sup_{(\varphi,\psi) \in \ \tilde{\phi}_c} J(\varphi, \psi), \tag{1.20}$$

onde $\tilde{\phi}_c$ é o conjunto de todas as funções mensuráveis $(\varphi, \psi) \in \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$ satisfazendo $\varphi(x) + \psi(y) \leq c(x, y), \ \forall (x, y) \in X \times Y$.

Demonstração. Lembre que Φ_c foi definido no enunciado do teorema 1.3.2 como o conjunto de todas as funções mensuráveis $(\varphi, \psi) \in \mathcal{L}^1(d\mu) \times \mathcal{L}^1(d\nu)$ satisfazendo $\varphi(x) + \psi(y) \leq c(x, y)$ para $d\mu$-quase todo ponto em X e $d\nu$-quase todo ponto em Y.

É imediato que $\sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi) \geq \sup_{(\varphi,\psi) \in \ \tilde{\phi}_c} J(\varphi, \psi)$.

Dadas $(\varphi, \psi) \in \Phi_c$ sejam N_X e N_Y tais que $\varphi(x) + \psi(y) \leq c(x, y), \ \forall (x, y) \in N_X^c \times N_Y^c$ e $\mu(N_X) = 0 = \nu(N_Y)$. Em particular, φ e ψ são finitas $\forall (x, y) \in N_X^c \times N_Y^c$.

Definindo,

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x), & \text{se } x \in N_X^c \\ -\infty, & \text{se } x \in N_X \end{cases}$$

e

$$\tilde{\psi}(y) = \begin{cases} \varphi(y), & \text{se } y \in N_Y^c \\ -\infty, & \text{se } y \in N_Y \end{cases}$$

vemos que $\tilde{\varphi}(x)+\tilde{\psi}(y) \leq c(x,y)$ para todo $(x,y) \in X \times Y$. Além disso, $J(\varphi,\psi) = J(\tilde{\varphi},\tilde{\psi})$. O que mostra que $\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) \leq \sup_{(\varphi,\psi)\in\ \tilde{\phi}_c} J(\varphi,\psi)$ e que a igualdade (1.20) é válida. □

Corolário 1.3.15. *Seja c como no teorema 1.3.2. Suponha também que c seja limitada, então*

$$\sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) = \sup\{J(\varphi^{cc},\varphi^c),\ \varphi \textit{ limitada}\}.$$

Demonstração. Do lema anterior, se c é limitada, podemos supor sem perda de generalidade que para qualquer $(\varphi,\psi) \in \Phi_c$ vale que $\varphi(x)+\psi(y) \leq c(x,y)$ para todo $(x,y) \in X\times Y$ e portanto

$$\psi(y) \leq c(x,y) - \varphi(x). \tag{1.21}$$

Além disso, se φ é limitada então

$$\psi(y) \leq \inf_{x\in X}[c(x,y) - \varphi(x)] = \varphi^c(y),\ \forall y \in Y.$$

Ainda, pela proposição 1.3.13, φ^c é mensurável. Da definição de φ^c, vale que, para todo $(x,y) \in X \times Y$, $\varphi(x) + \varphi^c(y) \leq c(x,y)$ o que implica

$$\varphi(x) \leq \varphi^{cc}(x).$$

Novamente, pela proposição 1.3.13, φ^{cc} é mensurável. Logo $J(\varphi^{cc},\varphi^c) \geq J(\varphi^{cc},\psi) \geq J(\varphi,\psi),\ \forall(\varphi,\psi) \in \Phi_c$ com φ limitada. Portanto

$$\begin{aligned} \sup\{J(\varphi^{cc},\varphi^c),\ \varphi \text{ limitada}\} &\geq \sup_{(\varphi,\psi)\in\Phi_c}\{J(\varphi,\psi),\ \varphi \text{ limitada}\} \\ &\geq \sup_{(\varphi,\psi)\in\Phi_c\cap C_b} J(\varphi,\psi) \\ &= \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi). \end{aligned} \tag{1.22}$$

onde a última igualdade é justificada pelo teorema 1.3.2.

Além disso, para toda φ limitada e para todo $(x, y) \in X \times Y$

$$\begin{aligned} \varphi^{cc}(x) = \inf_{x \in X}[c(x,y) - \varphi^c(y)] &\implies \varphi^{cc}(x) \leq c(x,y) - \varphi^c(y) \\ &\implies \varphi^{cc}(x) + \varphi^c(y) \leq c(x,y) \\ &\implies (\varphi^{cc}, \varphi^c) \in \Phi_c. \end{aligned}$$

Por isso

$$\sup \{J(\varphi^{cc}, \varphi^c),\ \varphi \text{ limitada}\} \leq \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi). \tag{1.23}$$

Logo, de (1.22) e (1.23) temos

$$\sup \{J(\varphi^{cc}, \varphi^c), \varphi \text{ limitada}\} = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi).$$

□

Corolário 1.3.16. *Sob as mesmas hipóteses do teorema 1.3.15 vale que*

$$\sup_{(\varphi,\psi) \in \Phi_c} \{J(\varphi, \psi)\} = \sup_{(\varphi,\psi) \in \Phi_c} \{J(\varphi, \psi); 0 \leq \varphi \leq 2\|c\|_\infty, -\|c\|_\infty \leq \psi \leq 0\}.$$

Demonstração. De fato, seja φ limitada, pela definição de φ^c temos

$$\begin{aligned} \varphi^c(y) &= \inf_{x \in X}[c(x,y) - \varphi(x))] \\ &\geq \inf_{x \in X} c(x,y) + \inf_{x \in X}[-\varphi(x)] \\ &= \inf_{x \in X} c(x,y) - \sup_{x \in X} \varphi(x) \\ &\geq -\sup_{x \in X} \varphi(x). \end{aligned}$$

Por outro lado,

$$\begin{aligned} \varphi^c(y) &= \inf_{x \in X}[c(x,y) - \varphi(x))] \leq c(x,y) - \varphi(x) \\ \implies \varphi^c(y) &\leq \|c\|_\infty - \varphi(x) \\ \implies \varphi^c(y) &\leq \|c\|_\infty - \sup_{x \in X} \varphi(x). \end{aligned}$$

Portanto,

$$-\sup_{x\in X}\varphi(x) \leq \varphi^c(y) \leq \|c\|_\infty - \sup_{x\in X}\varphi(x).$$

Raciocinando de maneira análoga para φ^{cc}, temos

$$\begin{cases} -\sup\varphi & \leq & \varphi^c & \leq & \|c\|_\infty - \sup\varphi, \\ -\sup\varphi^c & \leq & \varphi^{cc} & \leq & \|c\|_\infty - \sup\varphi^c. \end{cases} \tag{1.24}$$

Observe que, para todo $s \in \mathbb{R}$,

$$\begin{aligned} (\varphi+s)^c(y) &= \inf_{x\in X}[c(x,y) - (\varphi(x)+s)] \\ &= \inf_{x\in X}[c(x,y) - \varphi(x)] - s \\ &= \varphi^c(y) - s. \end{aligned}$$

Da mesma forma,

$$\begin{aligned} (\varphi+s)^{cc}(x) &= \inf_{y\in Y}[c(x,y) - (\varphi+s)^c(y)] \\ &= \inf_{y\in Y}[c(x,y) - (\varphi^c(y)-s)] \\ &= \inf_{y\in Y}[c(x,y) - \varphi^c(y)] + s \\ &= \varphi^{cc}(x) + s. \end{aligned}$$

Dada φ limitada, lembre que $J((\varphi+s)^{cc}, (\varphi+s)^c) = J(\varphi^{cc}+s, \varphi^c - s) = J(\varphi^{cc}, \varphi^c)$. Assim tomando $s = \|c\|_\infty - \sup_{x\in X}\varphi(x)$, podemos supor, sem perda de generalidade, que $\sup_{x\in X}\varphi(x) = \|c\|_\infty$. Logo por (1.24) temos que

$$\begin{cases} -\|c\|_\infty & \leq & \varphi^c & \leq & 0, \\ 0 & \leq & \varphi^{cc} & \leq & 2\|c\|_\infty, \end{cases} \tag{1.25}$$

o que implica

$$\begin{aligned} \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi) &= \text{pelo corolário 1.3.15} \\ &= \sup\{J(\varphi^{cc},\varphi^c), \varphi \text{ limitada}\} \\ &= \sup\{J(\varphi^{cc},\varphi^c), \varphi \text{ limitada},\ 0 \leq \varphi^{cc} \leq 2\|c\|_\infty,\ -\|c\|_\infty \leq \varphi^c \leq 0\} \\ &\leq \sup_{(\varphi,\psi)\in\Phi_c}\{J(\varphi,\psi),\ 0 \leq \varphi \leq 2\|c\|_\infty,\ -\|c\|_\infty \leq \psi \leq 0\} \\ &\leq \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi). \end{aligned}$$

Portanto

$$\sup_{(\varphi,\psi)\in\Phi_c} \{J(\varphi,\psi),\ 0 \le \varphi \le 2\|c\|_\infty,\ -\|c\|_\infty \le \psi \le 0\} = \sup_{(\varphi,\psi)\in\Phi_c} J(\varphi,\psi).$$

□

Capítulo 2

Função Custo Distância

Quando $X = Y$ e a função custo é uma métrica $c(x,y) = d(x,y)$, o seguinte resultado é válido:

Teorema 2.0.17 (**Teorema de Kantorovich-Rubinstein**). *Sejam $X = Y$ um espaço polonês e d a distância em X. Seja $\mathcal{T}_d$ o custo do transporte ótimo para o custo $c(x,y) = d(x,y)$, isto é*

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

Denote por $\mathrm{Lip}(X)$ o espaço de todas as funções Lipschitz em X tomando valores em $\mathbb{R}$, e

$$\|\varphi\|_{\mathrm{Lip}} \equiv \sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{d(x,y)}$$

a pseudo-norma definida em $\mathrm{Lip}(X)$. Então

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

Além disso, o valor do supremo não muda quando adicionamos a hipótese de que φ é limitada.

Lembre que, uma pseudo-norma mantém quase todas as propriedades da norma, exceto a propriedade que garante que ter norma nula é exclusividade do elemento neutro, isto é, $\|\varphi\|_{\mathrm{Lip}} = 0 \nRightarrow \varphi\equiv 0$ (basta considerar $\varphi\equiv 1$). Uma norma para o espaço $\mathrm{Lip}(X)$ pode ser a definida como $\|.\|_{\mathrm{BL}} = \|.\|_{\mathrm{Lip}} + \|.\|_{\infty}$.

Vamos fazer a demonstração no caso em que d é limitada, o que vale por exemplo se X é compacto. Observamos ainda que

$$\int_X \varphi d(\mu - \nu) = \int_X \varphi d(\mu) - \int_X \varphi d(\nu).$$

Demonstração. Note inicialmente que qualquer f k-Lipschitz é limitada: com efeito, fixado $x_0 \in X$,

$$|f(x)| = |f(x) - f(x_0)| + |f(x_0)| \leq kd(x, x_0) + |f(x_0)| \leq kM + |f(x_0)|, \ \forall x \in X,$$

onde M é escolhido de forma a termos $|d| < M$. Portanto, toda função k-Lipschitz é integrável em relação às probabilidade μ e ν.

Com isso, pelo teorema 1.3.2, basta mostrar que

$$\sup_{(\varphi,\psi)\in\phi_d} J(\varphi, \psi) = \sup \left\{ \int_X \varphi d(\mu - \nu); \ \|\varphi\|_{\text{Lip}} \leq 1 \right\}.$$

Afirmação 1.

$$\sup_{(\varphi,\psi)\in\phi_d} J(\varphi, \psi) = \sup_{\varphi\in\mathcal{L}^1(d\mu)} J(\varphi^{dd}, \varphi^d)$$

onde

$$\varphi^d(y) \equiv \inf_{x\in X}[d(x, y) - \varphi(x)] \quad \text{e} \quad \varphi^{dd}(x) \equiv \inf_{y\in Y}[d(x, y) - \varphi^d(y)].$$

Prova da afirmação 1. De fato, pelo corolário 1.3.15, temos

$$\begin{aligned} \sup_{(\varphi,\psi)\in\phi_d} J(\varphi, \psi) &= \sup \left\{ J(\varphi^{dd}, \varphi^d), \ \varphi \text{ limitada} \right\} \\ &\leq \sup_{\varphi\in\mathcal{L}^1(d\mu)} J(\varphi^{dd}, \varphi^d). \end{aligned}$$

Por outro lado, como d é limitada, e consequentemente integrável, se $\varphi \in \mathcal{L}^1(d\mu)$, então pela definição de φ^d,

$$\begin{aligned} \int \varphi^d d\nu &= \int \varphi^d(y) d\pi(x, y) \leq \int [d(x, y) - \varphi(x)] d\pi(x, y) \\ &= \int d(x, y) d\pi(x, y) - \int \varphi d\mu < \infty, \end{aligned}$$

onde as igualdades são justificadas pelo fato de $\pi \in \Pi(\mu, \nu)$. Portanto $\varphi^d \in \mathcal{L}^1(d\nu)$. Da mesma forma $\varphi^{dd} \in \mathcal{L}^1(d\mu)$. Além disso, para todo $(x, y) \in X \times Y$, pela definição de φ^{dd}, é imediato que $\varphi^{dd}(x) \leq d(x, y) - \varphi^d(y)$. Logo $(\varphi^{dd}, \varphi^d) \in \phi_d$.

Assim,

$$\sup_{(\varphi,\psi)\in\phi_d} J(\varphi, \psi) \quad \geq \quad \sup_{\varphi\in\mathcal{L}^1(d\mu)} J(\varphi^{dd}, \varphi^d),$$

o que finaliza a prova da afirmação 1.

Afirmação 2. φ^d é 1-Lipschitz.

Prova da afirmação 2. Suponha sem perda de generalidade que $\varphi^d(y_1) \geq \varphi^d(y_2)$:

$$\begin{aligned}
|\varphi^d(y_1) - \varphi^d(y_2)| &= \varphi^d(y_1) - \varphi^d(y_2) \\
&= \inf_{x\in X}[d(x, y_1) - \varphi(x)] - \inf_{x\in X}[d(x, y_2) - \varphi(x)] \\
&\leq \inf_{x\in X}[d(x, y_1) - \varphi(x)] - d(x_0, y_2) + \varphi(x_0) + \delta \\
&\leq d(x_0, y_1) - \varphi(x_0) - d(x_0, y_2) + \varphi(x_0) + \delta \\
&\leq d(y_1, y_2) + \delta,
\end{aligned}$$

onde nas duas primeiras desigualdades usamos a definição de ínfimo e na terceira desigualdade usamos a desigualdade triangular. Fazendo δ tender a zero temos $|\varphi^d(y_1) - \varphi^d(y_2)| \leq d(y_1, y_2)$ o que finaliza a demonstração da afirmação 2.

Afirmação 3: $\varphi^{dd} = -\varphi^d$

Prova da afirmação 3. Note que

$$\varphi^{dd}(x) \quad = \quad \inf_{y\in Y}[d(x, y) - \varphi^d(y)] \leq -\varphi^d(x). \tag{2.1}$$

Além disso, como φ^d é 1-Lipschitz, temos que $-\varphi^d(x) \leq d(x, y) - \varphi^d(y)$ e assim

$$\varphi^{dd}(x) = \inf_{y\in Y}[d(x, y) - \varphi^d(y)] \geq -\varphi^d(x). \tag{2.2}$$

De (2.1) e (2.2) temos a demonstração da afirmação 3 e, portanto

$$\begin{aligned}\sup_{(\varphi,\psi)\in\phi_d} J(\varphi,\psi) &= \sup_{\varphi\in\mathcal{L}^1(d\mu)} J(-\varphi^d,\varphi^d)\\ &\leq \sup_{\|\varphi\|_{\mathrm{Lip}}\leq 1} J(-\varphi,\varphi)\\ &= \sup_{\|\varphi\|_{\mathrm{Lip}\leq 1}} J(\varphi,-\varphi)\\ &\leq \sup_{(\varphi,\psi)\in\phi_d} J(\varphi,\psi),\end{aligned}$$

onde a primeira igualdade vem das afirmações 1 e 3, a primeira desigualdade do fato de φ^d ser 1-Lipschitz (afirmação 2), e a última desigualdade vem do fato de φ ser 1-Lipschitz implicar $(\varphi,-\varphi)\in\Phi_d$. Com isso, lembrando que

$$J(\varphi,-\varphi)=\int_X \varphi d(\mu-\nu),$$

e usando a observação imediatamente anterior à afirmação 1, finalizamos a demonstração do teorema de Kantorovich-Rubinstein. □

A seguir faremos algumas considerações, que não serão demonstradas, sobre a topologia dos espaços $P(X)$ e $M(X)$ no caso em que d é limitada. Tais observações podem ser encontradas em **Villani**, [11].

Seja $P_1(X)$ o espaço das probabilidades μ tal que $\int d(x_0,x)d\mu(x)<\infty$ para algum x_0, e seja $E_1(X)$ o espaço vetorial gerado por $P_1(X)$. Note que $P_1(X)=P(X)$ se d é limitada. Em $E_1(X)$ podemos definir a norma

$$\|\mu\|_{\mathrm{KR}}=\sup\left\{\int_X\varphi d\mu;\ \varphi\in\mathrm{Lip}(X)\cap\mathcal{L}^1(d|\mu|);\ \|\varphi\|_{\mathrm{Lip}\leq 1}\right\}.$$

Então o teorema de Kantorovich-Rubinstein implica que $\mathcal{T}_d(\mu,\nu)=\|\mu-\nu\|_{\mathrm{KR}}$ para todas as probabilidades em $\mu,\nu\in P_1(X)$.

Quando d é limitada, $\mathcal{T}_d$ e $\|.\|_{\mathrm{KR}}$ estão bem definidas em $P_1(X)$ e $E_1(X)$ respectivamente, então $\mathcal{T}_d$ algumas vezes chamado é de *distância Lipschitz limitada* e $\|.\|_{\mathrm{KR}}$ de *norma Lipschitz limitada*. Contudo é usual definir a distância Lipschitz limitada de uma maneira um pouco diferente:

$$\|\mu-\nu\|_{\mathrm{BL}^*}\equiv\sup_{\|\varphi\|_{\mathrm{BL}}\leq 1}\int\varphi d(\mu-\nu).$$

Além disso, em $P_1(X)$, também podemos definir a métrica dada por

$$W_1(\mu,\nu) = \sup_{\|\varphi\|_{\mathrm{Lip}\leq 1}} \int \varphi d(\mu - \nu),$$

chamada de *distância de Kantorovich-Rubinstein ou Wassertein-1.* Segundo **Villani** [11] se d é limitada, então W_1 metriza a convergência fraca em $P(X)$, ou seja, dada uma sequência $(\mu_k) \in P(X)$ e $\mu \in P(X)$

$$\left[\lim_{k\to\infty} \int \varphi d\mu_k = \int \varphi d\mu,\ \forall \varphi \in C_b(X)\right] \Leftrightarrow \left[\lim_{k\to\infty} W_1(\mu_k,\mu) = 0\right].$$

Observe ainda que, pelo teorema de Kantorovich-Rubinstein, para todo $\pi \in \Pi(\mu,\nu)$ e toda $\varphi \in \mathrm{Lip}(X)$ com $\|\varphi\|_{\mathrm{Lip}} \leq 1$, cotas inferior e superior para a distância Wasserstein-1 são naturalmente fornecidas através de

$$I(\pi) = \int d(x,y)d\pi(x,y) \geq W_1(\mu,\nu) \geq \int \varphi d(\mu - \nu) = J(\varphi, -\varphi).$$

Apêndice - Algumas Demonstrações

Demonstração do lema 1.3.7: Sejam $x, y \in \mathrm{Int}(C)$, então existe um $r > 0$ tal que $B(x,r) \subset C$ e $B(y,r) \subset C$. Assim

$$tB(x,r) + (1-t)B(y,r) \subset C, \forall t \in [0,1]$$

Afirmamos que

$$tB(x,r) + (1-t)B(y,r) = B(tx + (1-t)y, r), \forall t \in [0,1].$$

De fato, dado $z \in tB(x,r) + (1-t)B(y,r)$, então $z = t\tilde{x} + (1-t)\tilde{y}$, onde $\tilde{x} \in B(x,r)$ e $\tilde{y} \in B(y,r)$. Logo,

$$\begin{aligned} |z - [tx + (1-t)y]| &= |t(\tilde{x} - x) + (1-t)(\tilde{y} - y)| \\ &\leq t|\tilde{x} - x| + (1-t)|\tilde{y} - y| \\ &< tr + (1-t)r = r. \end{aligned}$$

Por outro lado, considere $z \in B(tx+(1-t)y, r)$. Um cálculo imediato mostra que podemos escrever $z = t\tilde{x}+(1-t)\tilde{y}$, onde $\tilde{x} = z+(1-t)(x-y) \in B(x,r)$ e $\tilde{y} = z-t(x-y) \in B(y,r)$. Assim,

$$B(tx + (1-t)y, r) \subset C, \forall t \in [0,1]$$

o que prova que $\mathrm{Int}(C)$ é convexo.

Além disso, se existe $y_0 \in C$ tal que $B(y_0, r) \subset C$, temos que $\forall x \in C$

$$tx + (1-t)B(y_0, r) \subset C, \forall t \in [0,1].$$

Mas,

$$B(tx+(1-t)y_0,(1-t)r) = tx+(1-t)B(y_0,r) \subset C, \forall t \in [0,1]$$

e portanto $tx+(1-t)y_0 \in \mathrm{Int}(C), \forall t \in [0,1]$. Agora, $\forall x \in C$ vale que $x = \lim_{n\to+\infty}[(1-\frac{1}{n})x+\frac{1}{n}y_0]$. Mas, pelo que provamos acima $(1-\frac{1}{n})x+\frac{1}{n}y_0 \in \mathrm{Int}(C)$. Logo, $x \in \overline{\mathrm{Int}(C)}$. Isto prova que $C \subset \overline{\mathrm{Int}(C)}$ e consequentemente $\overline{C} \subset \overline{\mathrm{Int}(C)}$.

Demonstração da proposição 1.3.13. Defina, para cada $l \in \mathbb{N}$

$$c_l(x,y) = \inf_{(\bar{x},\bar{y})\in X\times Y} \left\{ min\,(c(\bar{x},\bar{y}),l) + l[d(x,\bar{x})+\tilde{d}(y,\bar{y})] \right\}$$

e

$$\psi_l(y) = \inf_{x\in X} \left[c_l(x,y)-\varphi(x)\right].$$

Faremos a prova em 4 passos.

Afirmação 1. Afirmamos que c_l é uma função l-Lipschitz. Em particular, c_l é uniformemente contínua.

Prova da afirmação 1. De fato, dados $(x_1,y_1),(x_2,y_2) \in X\times Y$, suponha, sem perda de generalidade, que $c_l(x_1,y_1) \geq c_l(x_2,y_2)$, então

$$\begin{aligned}
&|c_l(x_1,y_1)-c_l(x_2,y_2)| = \\
&\quad = c_l(x_1,y_1)-c_l(x_2,y_2) \\
&\quad = \inf_{(\bar{x}_1,\bar{y}_1)\in X\times Y} \left\{ min\,(c(\bar{x}_1,\bar{y}_1),l) + l[d(x_1,\bar{x}_1)+\tilde{d}(y_1,\bar{y}_1)] \right\} \\
&\qquad - \inf_{(\bar{x}_2,\bar{y}_2)\in X\times Y} \left\{ min\,(c(\bar{x}_2,\bar{y}_2),l) + l[d(x_2,\bar{x}_2)+\tilde{d}(y_2,\bar{y}_2)] \right\}.
\end{aligned}$$

Pela definição de ínfimo, dado $\delta > 0$ existe $(x_0,y_0) \in X\times Y$ tal que

$$\begin{aligned}
&\inf_{(\bar{x}_2,\bar{y}_2)\in X\times Y} \left\{ min\,(c(\bar{x}_2,\bar{y}_2),l) + l[d(x_2,\bar{x}_2)+\tilde{d}(y_2,\bar{y}_2)] \right\} + \delta \\
&\qquad \geq \min\,(c(x_0,y_0),l) + l[d(x_2,x_0)+\tilde{d}(y_2,y_0)].
\end{aligned}$$

Escolhendo $(\bar{x}_1, \bar{y}_1) = (x_0, y_0)$, segue que

$$\begin{aligned}
&|c_l(x_1, y_1) - c_l(x_2, y_2)| \\
&\quad \leq \inf_{(\bar{x}_1, \bar{y}_1) \in X \times Y} \left\{ \min\left(c(\bar{x}_1, \bar{y}_1), l\right) + l[d(x_1, \bar{x}_1) + \tilde{d}(y_1, \bar{y}_1)] \right\} \\
&\qquad\qquad - \min\left(c(x_0, y_0), l\right) - l[d(x_2, x_0) + \tilde{d}(y_2, y_0)] + \delta \\
&\quad \leq \min\left(c(x_0, y_0), l\right) + l[d(x_1, x_0) + \tilde{d}(y_1, y_0)] \\
&\qquad\qquad - \min\left(c(x_0, y_0), l\right) - l[d(x_2, x_0) + \tilde{d}(y_2, y_0)] + \delta \\
&\quad = l[d(x_1, x_0) - d(x_2, x_0) + \tilde{d}(y_1, y_0) - \tilde{d}(y_2, y_0)] + \delta \\
&\quad \leq l[d(x_1, x_2) + \tilde{d}(y_1, y_2)] + \delta \\
&\quad = l[d + \tilde{d}]\left((x_1, y_1), (x_2, y_2)\right) + \delta,
\end{aligned}$$

onde a terceira desigualdade é justificada pela desigualdade triangular. Fazendo δ tender a zero temos que c_l é l-Lipschitz.

Afirmação 2. Mostraremos que $\lim\limits_{l \to \infty} c_l(x, y) = c(x, y)$.

Prova da afirmação 2. De fato,

$$c_l(x, y) \leq \min(c(x, y), l) \leq c(x, y) \Rightarrow \lim c_l(x, y) \leq c(x, y). \tag{2.3}$$

Pela definição de ínfimo, para todo $l \in \mathbb{N}$ existe $(x_l, y_l) \in X \times Y$ tal que

$$c_l(x, y) \leq \min(c(x_l, y_l), l) + l[d(x, x_l) + \tilde{d}(y, y_l)] \leq c_l(x, y) + 1/l. \tag{2.4}$$

Como $c_l(x, y) + 1/l \geq \min(c(x_l, y_l), l) + l[d(x, x_l) + \tilde{d}(y, y_l)]$ e $\min(c(x_l, y_l), l) \geq 0$ temos $c_l(x, y) + 1/l \geq l[d(x, x_l) + \tilde{d}(y, y_l)]$.

Além disso c é limitada, logo existe um $M > 0$ tal que $c(x, y) \leq M$ para todo $(x, y) \in X \times Y$. Portanto

$$l[d(x, x_l) + \tilde{d}(y, y_l)] \leq c_l(x, y) + 1/l \leq c(x, y) + 1/l \leq M + 1/l.$$

Desta maneira $d(x, x_l) + \tilde{d}(y, y_l) \leq M/l + 1/l^2$. Fazendo l tender ao infinito temos que

$$d(x, x_l) + \tilde{d}(y, y_l) \to 0.$$

Portanto

$$x_l \to x \quad \text{e} \quad y_l \to y.$$

Tomando $l_0 \geq M$, para qualquer $l \geq l_0$, $c(x,y) \leq l$ e assim $\min(c(x,y),l) = c(x,y), \forall (x,y) \in X \times Y$.

Então, para qualquer $l \geq l_0$, da desigualdade (2.4) temos

$$\begin{aligned} \min(c(x_l,y_l),l) + l[d(x,x_l) + \tilde{d}(y,y_l)] &= c(x_l,y_l) + l[d(x,x_l) + \tilde{d}(y,y_l)] \\ &\leq c_l(x,y) + 1/l. \end{aligned}$$

Logo

$$c(x_l,y_l) \leq c_l(x,y) + 1/l, \ \forall l \geq l_0.$$

Sendo c uma função semicontínua inferiormente,

$$c(x,y) \leq \liminf_{l\to\infty} c(x_l,y_l) \leq \liminf_{l\to\infty}[c_l(x,y) + 1/l] = \liminf_{l\to\infty} c_l(x,y).$$

Então

$$c(x,y) \leq \lim_{l\to\infty} \inf c_l(x,y) \leq \lim_{l\to\infty} c_l(x,y). \tag{2.5}$$

Juntando (2.3) e (2.5) temos

$$c(x,y) = \lim c_l(x,y). \tag{2.6}$$

Afirmação 3. Afirmamos que ψ_l é uma função l-Lipschitz. Em particular, ψ_l é uniformemente contínua.

Prova da afirmação 3. De fato, dados $y_1, y_2 \in Y$, suponha sem perda de generalidade que $\psi_l(y_1) \leq \psi_l(y_2)$, então

$$\begin{aligned} |\psi_l(y_1) - \psi_l(y_2)| &= \psi_l(y_1) - \psi_l(y_2) \\ &= \inf_{\bar{x}\in X}\{c_l(\bar{x},y_1) + \varphi(\bar{x})\} - \inf_{x\in X}\{c_l(x,y_2) + \varphi(x)\}. \end{aligned}$$

Pela definição de ínfimo, dado $\delta > 0$ existe $x_0 \in X$ tal que

$$\inf_{x\in X}\{c_l(x,y_2) + \varphi(x)\} + \delta \geq c_l(x_0,y_2) + \varphi(x_0). \tag{2.7}$$

Assim, segue que

$$
\begin{aligned}
|\psi_l(y_1) - \psi_l(y_2)| &\leq \inf_{\bar{x} \in X} \{c_l(\bar{x}, y_1) + \varphi(\bar{x})\} - c_l(x_0, y_2) - \varphi(x_0) + \delta \\
&\leq c_l(x_0, y_1) + \varphi(x_0) - c_l(x_0, y_2) - \varphi(x_0) + \delta \\
&= c_l(x_0, y_1) - c_l(x_0, y_2) + \delta \\
&\leq |c_l(x_0, y_1) - c_l(x_0, y_2)| + \delta \\
&\leq l[d(x_0, x_0) + \tilde{d}(y_1, y_2)] + \delta \\
&= l\tilde{d}(y_1, y_2) + \delta,
\end{aligned}
$$

onde a primeira desigualdade usamos (2.7), na segunda escolhemos $\overline{x} = x_0$, na terceira uma propriedade da norma e a desigualdade triangular e na quarta o passo 1.

Fazendo δ tender a zero temos que ψ_l é l-Lipschitz.

Afirmação 4. Provaremos que $\lim_{l\to\infty} \psi_l = \varphi^c$

Prova da afirmação 4. Basicamente precisamos mostrar que para todo $y \in Y$,

$$
\lim_{l\to\infty} \left(\inf_{x \in X} [c_l(x, y) + \varphi(x)] \right) = \inf_{x \in X} [c(x, y) + \varphi(x)] .
$$

Com efeito, pelo passo 2, para qualquer $y \in Y$,

$$
\begin{aligned}
\inf_{x \in X} [c(x, y) + \varphi(x)] &- \lim_{l\to\infty} \left(\inf_{x \in X} [c_l(x, y) + \varphi(x)] \right) \\
&\leq \inf_{x \in X} [c(x, y) + \varphi(x)] - \lim_{l\to\infty} c(x_0, y) - \varphi(x_0) + \delta \\
&\leq c(x_0, y) - \lim_{l\to\infty} c_l(x_0, y) + \delta \\
&= \delta,
\end{aligned}
$$

onde nas desigualdades usamos a definição de ínfimo. Fazendo $\delta \to 0$, segue que

$$
\lim_{l\to\infty} \left(\inf_{x \in X} [c_l(x, y) + \varphi(x)] \right) \geq \inf_{x \in X} [c(x, y) + \varphi(x)] .
$$

Usando (2.3) concluímos que

$$
\lim_{l\to\infty} \left(\inf_{x \in X} [c_l(x, y) + \varphi(x)] \right) \leq \inf_{x \in X} [c(x, y) + \varphi(x)] ,
$$

o que prova a afirmação feita.

Logo, diretamente desta afirmação, para qualquer $y \in Y$,

$$\begin{aligned}
\lim_{l\to\infty} \psi_l(y) &= \lim_{l\to\infty} \left(\inf_{x\in X} [c_l(x,y) + \varphi(x)] \right) \\
&= \inf_{x\in X} [c(x,y) + \varphi(x)] \\
&= \varphi^c(y).
\end{aligned}$$

Referências Bibliográficas

[1] Bartle, R., *The Elements of Integration and Lebesgue Measure*, Wiley, New York, 1995.

[2] Brezis, H., *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2010.

[3] Billingsley, P., *Convergence of Probability Measure*, Wiley, Chicago, 1999.

[4] Billingsley, P., *Probability and Measure*, John Wiley & Sons, 1986.

[5] Fernandez, P., *Medida e Integração*, IMPA, Rio de Janeiro, 1996.

[6] Fitzpatrick, P., *Real Analysis*, China Machine Press, China, 2010.

[7] Lima, E. L., *Espaços métricos*, IMPA, Rio de Janeiro, 2007.

[8] Rockafellar, R., *Convex Analysis*, Princeton University Press, Princeton, 1972.

[9] Souza, E. A., *O problema de Monge-Kantorovich para duas medidas de probabilidade sobre um conjunto finito*, Dissertação de mestrado, USP, São Paulo, 2009

[10] Villani, C., *Optimal transport, old and new*, Springer, 2008.

[11] Villani, C., *Topics in Optimal Transportation*, Board, 2003.