Non-traditional costs and set dualities

Katarzyna Wyczesany

Tel Aviv University

joint work with S. Artstein-Avidan and S. Sadovsky

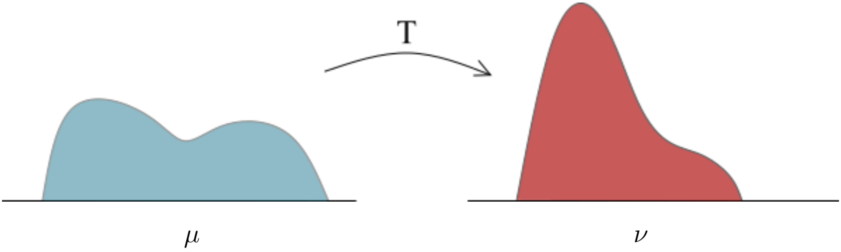
October 18, 2021

Outline

1. Introduction to optimal transport problem
   - Cost induced transforms
   - c-subgradient
   - Geometric notion of optimality
2. Existence of a potential
3. Transporting measures
   - Compatibility
4. Set dualities

Transport of measures

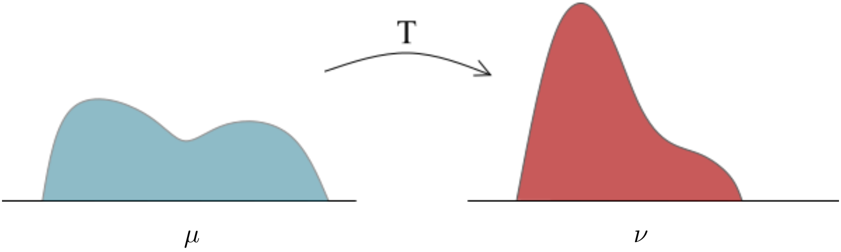
Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ be two probability measures.



T transports μ to ν if $\mu(T^{-1}(A)) = \nu(A)$ for all ν-measurable sets $A \subset Y$.

Transport of measures

Let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ be two probability measures.



T transports μ to ν if $\mu(T^{-1}(A)) = \nu(A)$ for all ν-measurable sets $A \subset Y$.

We say that $\pi \in \mathcal{P}(X \times Y)$ is a **transport plan**, $\pi \in \Pi(\mu, \nu)$, if for any measurable sets $A \subset X$ and $B \subset Y$ we have

$$\pi(A \times Y) = \mu(A) \quad \text{and} \quad \pi(X \times B) = \nu(B)$$

Kantorovich duality

Given a cost function $c : X \times Y \to (-\infty, \infty]$ one is interested in finding an **optimal plan**, that is the plan $\pi \in \Pi(\mu, \nu)$ with infimal **total cost**

$$C(\mu, \nu) = \inf \int_{X \times Y} c(x, y) d\pi(x, y)$$

Kantorovich duality

Given a cost function $c : X \times Y \to (-\infty, \infty]$ one is interested in finding an **optimal plan**, that is the plan $\pi \in \Pi(\mu, \nu)$ with infimal **total cost**

$$C(\mu, \nu) = \inf \int_{X \times Y} c(x, y) d\pi(x, y)$$

Theorem

Theorem (Kantorovich) *For a lower semicontinuous cost function* $c : X \times Y \to (-\infty, \infty]$ *and probability measures* μ, ν *we have*

$$C(\mu, \nu) = \sup \left\{ \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y) : \varphi(x) + \psi(y) \leq c(x, y) \right\},$$

where $\varphi \in L^1(X, \mu)$, $\psi \in L^1(Y, \nu)$.

Fix a cost function $c : X \times Y \to (-\infty, +\infty]$

Definition. A pair of functions φ, ψ is called c-**admissible** if for all x, y

$$\varphi(x) + \psi(y) \leq c(x, y)$$

Fix a cost function $c : X \times Y \to (-\infty, +\infty]$

Definition. A pair of functions φ, ψ is called c-**admissible** if for all x, y

$$\varphi(x) + \psi(y) \leq c(x, y)$$

Definition. The c-**transform** of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x (c(x, y) - \varphi(x))$$

Fix a cost function $c : X \times Y \to (-\infty, +\infty]$

Definition. A pair of functions φ, ψ is called c-**admissible** if for all x, y

$$\varphi(x) + \psi(y) \leq c(x, y)$$

Definition. The c-**transform** of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x \left(c(x, y) - \varphi(x)\right)$$

Note that:

- c-transform is order reversing
- For any admissible pair (φ, ψ) we have $\psi \leq \varphi^c$
- Further, we have that $\varphi \leq \varphi^{cc}$ and hence $\varphi^{ccc} = \varphi^c$

Fix a cost function $c : X \times Y \to (-\infty, +\infty]$

Definition. A pair of functions φ, ψ is called c-**admissible** if for all x, y

$$\varphi(x) + \psi(y) \leq c(x, y)$$

Definition. The c-**transform** of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x (c(x, y) - \varphi(x))$$

Note that:

- c-transform is order reversing
- For any admissible pair (φ, ψ) we have $\psi \leq \varphi^c$
- Further, we have that $\varphi \leq \varphi^{cc}$ and hence $\varphi^{ccc} = \varphi^c$

The c-**class** associated to a cost function c is the image of the c-transform.

Examples

The *c-transform* of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x \left(c(x, y) - \varphi(x)\right)$$

For $c(x, y) = |x - y|^2/2$ we have

$$\begin{aligned}\varphi^c(y) &= \inf_x(|x|^2/2 - \langle x, y\rangle + |y|^2/2 - \varphi(x)) \\ &= |y|^2/2 - \sup_x(\langle x, y\rangle - (|x|^2/2 - \varphi(x)))\end{aligned}$$

Examples

The *c-transform* of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x \left(c(x,y) - \varphi(x)\right)$$

For $c(x,y) = |x-y|^2/2$ we have

$$\begin{aligned}\varphi^c(y) &= \inf_x(|x|^2/2 - \langle x, y\rangle + |y|^2/2 - \varphi(x)) \\ &= |y|^2/2 - \sup_x(\langle x, y\rangle - (|x|^2/2 - \varphi(x)))\end{aligned}$$

Hence,

$$|y|^2/2 - \varphi^c(y) = \mathcal{L}\left(|x|^2/2 - \varphi(x)\right),$$

where we recall

$$\mathcal{L}\varphi(y) = \sup_{x\in\mathbb{R}^n}\left(\langle x, y\rangle - \varphi(x)\right).$$

Examples

The c-*transform* of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x \left(c(x,y) - \varphi(x)\right)$$

For $p(x,y) = -\ln(\langle x,y\rangle - 1)_+$ we have

$$\begin{aligned}\varphi^p(y) &= \inf_x(-\ln(\langle x,y\rangle - 1)_+ - \varphi(x)) \\ &= -\ln\left(\sup_x \frac{(\langle x,y\rangle - 1)_+}{e^{-\varphi(x)}}\right)\end{aligned}$$

Examples

The c-*transform* of a function $\varphi : X \to [-\infty, +\infty]$ is given by

$$\varphi^c(y) = \inf_x \left(c(x,y) - \varphi(x)\right)$$

For $p(x,y) = -\ln(\langle x,y\rangle - 1)_+$ we have

$$\begin{aligned}\varphi^p(y) &= \inf_x(-\ln(\langle x,y\rangle - 1)_+ - \varphi(x))\\ &= -\ln\left(\sup_x \frac{(\langle x,y\rangle - 1)_+}{e^{-\varphi(x)}}\right)\end{aligned}$$

Which we rewrite as

$$e^{-\varphi^p(y)} = \mathcal{A}\left(e^{-\varphi(\cdot)}\right)(y),$$

where

$$\mathcal{A}\varphi(y) = \sup_x \frac{(\langle x,y\rangle - 1)_+}{\varphi(x)}.$$

c-subgradient

For any admissible pair (φ, φ^c) and any transport plan π we obviously have

$$\sup\left(\int \varphi(x)d\mu(x) + \int \varphi^c(y)d\nu(y)\right) \leq \inf \int c(x,y)d\pi(x,y)$$

To find and optimal plan we need "$=$".

c-subgradient

For any admissible pair (φ, φ^c) and any transport plan π we obviously have

$$\sup \left(\int \varphi(x) d\mu(x) + \int \varphi^c(y) d\nu(y) \right) \leq \inf \int c(x,y) d\pi(x,y)$$

To find and optimal plan we need "$=$".

Definition. Given a c-class function $\varphi : X \to [-\infty, \infty]$ consider the set

$$\partial^c \varphi = \{(x,y) : \varphi(x) + \varphi^c(y) = c(x,y) < \infty\} \subset X \times Y$$

We call the section $\partial^c \varphi(x)$ the c-**subgradient** of φ at $x \in X$.

Analogously, $\partial^c \varphi^c(y)$ denotes the c-subgradient of φ^c at $y \in Y$.

c-cyclic monotonicity

Definition. The set $G \subseteq X \times Y$ is called *c-cyclically monotone* if for any $(x, y) \in G$ we have that $c(x, y) < \infty$ and for any $m \in \mathbb{N}$ and any $\{(x_i, y_i)\}_{i=1}^m \subset G$ we have that

$$\sum_{i=1}^{m} (c(x_i, y_i) - c(x_{i+1}, y_i)) \leq 0,$$

where we identify $x_{m+1} = x_1$.

c-cyclic monotonicity

Definition. The set $G \subseteq X \times Y$ is called *c-cyclically monotone* if for any $(x, y) \in G$ we have that $c(x, y) < \infty$ and for any $m \in \mathbb{N}$ and any $\{(x_i, y_i)\}_{i=1}^m \subset G$ we have that

$$\sum_{i=1}^{m} (c(x_i, y_i) - c(x_{i+1}, y_i)) \leq 0,$$

where we identify $x_{m+1} = x_1$.

Fact. For any cost function c and any function φ in the c-class the set

$$\partial^c \varphi = \{(x, \partial^c \varphi(x)) : \ x \in X\} \subset X \times Y$$

is c-cyclically monotone.

Rockafellar-Rochet-Rüschendorf theorem

Theorem

Let X, Y be two arbitrary sets, $c : X \times Y \to \mathbb{R}$ a **real-valued** *cost function and fix a set $G \subset X \times Y$. Then G is c-cyclically monotone if and only if there exists a c-class function $\varphi : X \to [-\infty, +\infty]$ such that $G \subset \partial^c \varphi$.*

Fix some element $(x_0, y_0) \in G$ and define

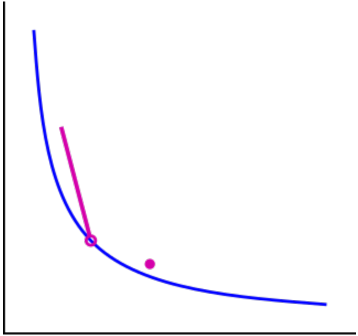
$$\varphi(x) = \inf\{c(x, y_m) - c(x_0, y_0) + \sum_{i=1}^{m} (c(x_i, y_{i-1}) - c(x_i, y_i))\}.$$

Here the infimum runs over all $m \in \mathbb{N}$ and all m-tuples $(x_i, y_i) \in G,\ i = 1, \dots, m$.

Let $p(x, y) = -\ln(\langle x, y\rangle - 1)_+$. Then

$$G = \{(x, y) : \tfrac{3}{4} \le x < 1,\ y = 3 - 2x\} \cup \{(\tfrac{3}{2}, \tfrac{3}{4})\}$$

is p-cyclically monotone but does not have a potential.



S. Artstein-Avidan, S. Sadovsky, K. W.

Theorem

Let X, Y *be two arbitrary sets and* $c : X \times Y \to (-\infty, \infty]$ *an arbitrary cost function. For a given subset* $G \subset X \times Y$ *there exists a* c*-class function* $\varphi : X \to [-\infty, \infty]$ *such that* $G \subset \partial^c \varphi$ *if and only if* G *is* c*-path-bounded.*

S. Artstein-Avidan, S. Sadovsky, K. W.

Theorem

Let X, Y be two arbitrary sets and $c : X \times Y \to (-\infty, \infty]$ an arbitrary cost function. For a given subset $G \subset X \times Y$ there exists a c-class function $\varphi : X \to [-\infty, \infty]$ such that $G \subset \partial^c \varphi$ if and only if G is c-path-bounded.

Definition. The set $G \subset X \times Y$ is called c-**path-bounded** if it satisfies that for any $(x_i, y_i), (x_j, y_j) \in G$ there exists some constant $M = M(i, j)$ such that for any $m \in \mathbb{N}$ and any $\{(x_i, y_i) : 2 \leq i \leq m-1\} \subset G$, letting $(x_i, y_i) = (x_1, y_1)$ and $(x_j, y_j) = (x_m, y_m)$, we have

$$\sum_{i=1}^{m-1} \big(c(x_i, y_i) - c(x_{i+1}, y_i)\big) \leq M.$$

Note that the c-path-boundedness implies c-cyclic monotonicity. Indeed, if $(x_i, y_i) = (x_j, y_j)$, then if there is some path for which the sum is positive, one can duplicate it many times to get paths with arbitrarily large sums.

Note that the c-path-boundedness implies c-cyclic monotonicity. Indeed, if $(x_i, y_i) = (x_j, y_j)$, then if there is some path for which the sum is positive, one can duplicate it many times to get paths with arbitrarily large sums.

For a real-valued cost we also have that c-cyclic monotonicity implies c-path-boundedness: consider two points $(x_1, y_1), (x_m, y_m) \in G$, then for any path $\{(x_i, y_i)\}_{i=2}^{m-1}$ we have

$$\sum_{i=1}^{m-1} (c(x_i, y_i) - c(x_{i+1}, y_i)) + c(x_m, y_m) - c(x_1, y_m) \leq 0.$$

Note that the c-path-boundedness implies c-cyclic monotonicity. Indeed, if $(x_i, y_i) = (x_j, y_j)$, then if there is some path for which the sum is positive, one can duplicate it many times to get paths with arbitrarily large sums.

For a real-valued cost we also have that c-cyclic monotonicity implies c-path-boundedness: consider two points $(x_1, y_1), (x_m, y_m) \in G$, then for any path $\{(x_i, y_i)\}_{i=2}^{m-1}$ we have

$$\sum_{i=1}^{m-1} (c(x_i, y_i) - c(x_{i+1}, y_i)) + c(x_m, y_m) - c(x_1, y_m) \leq 0.$$

It is important to note that we relied heavily on the fact that $c(x_1, y_m) < \infty$, otherwise this upper bound might be infinite, and therefore meaningless.

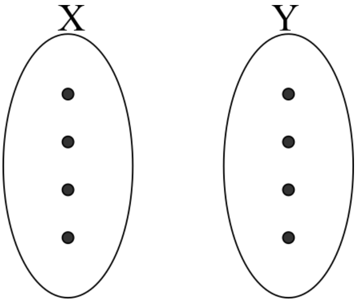
c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

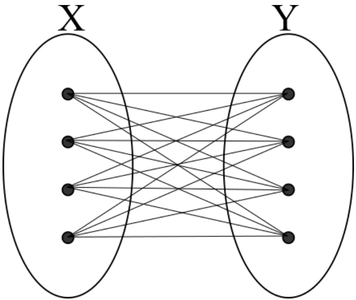
Consider discrete probability measures $\mu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{x_i}$, $\nu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{y_i}$.



c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

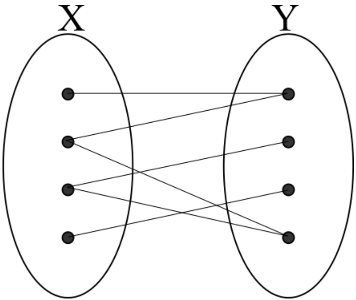
Consider discrete probability measures $\mu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{x_i}$, $\nu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{y_i}$.



c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

Consider discrete probability measures $\mu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{x_i}$, $\nu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{y_i}$.



c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

Consider discrete probability measures $\mu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{x_i}$, $\nu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{y_i}$.

Theorem (Hall's Marriage Theorem)

A bipartite graph G with a vertex set $V_1 \cup V_2$, such that $|V_1| = |V_2|$, contains a complete matching if and only if G satisfies Hall's condition

$$|S| \leq |N_G(S)| \text{ for every } S \subset V_1,$$

where $N_G(S) \subset V_2$ is the set of all neighbors of vertices in S.

c-compatibility

Fix a cost function $c : X \times Y \to (-\infty, \infty]$.

Consider discrete probability measures $\mu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{x_i}$, $\nu = \sum_{i=1}^{m} \frac{1}{m} \mathbb{1}_{y_i}$.

Theorem (Hall's Marriage Theorem)

A bipartite graph G with a vertex set $V_1 \cup V_2$, such that $|V_1| = |V_2|$, contains a complete matching if and only if G satisfies Hall's condition

$$|S| \leq |N_G(S)| \text{ for every } S \subset V_1,$$

where $N_G(S) \subset V_2$ is the set of all neighbors of vertices in S.

The condition can be reformulated in terms of the measures, as

$$\mu(A) \leq \nu(\{y : \exists x \in A,\ c(x,y) < \infty\}) \text{ for all } A \subset X.$$

c-compatibility

Definition. Let X, Y be measure spaces and $c : X \times Y \to (-\infty, \infty]$ be a measurable cost function. We say that two probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ are c-**compatible** if for any measurable $A \subset X$ it holds that

$$\mu(A) + \nu(\{y : \forall x \in A,\ c(x, y) = \infty\}) \leq 1.$$

c-compatibility

Definition. Let X, Y be measure spaces and $c : X \times Y \to (-\infty, \infty]$ be a measurable cost function. We say that two probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$ are c-**compatible** if for any measurable $A \subset X$ it holds that

$$\mu(A) + \nu(\{y : \forall x \in A,\ c(x,y) = \infty\}) \leq 1.$$

Lemma

Let X, Y be measure spaces and $c : X \times Y \to (-\infty, \infty]$ be a measurable cost function. Given $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, assume there exists $\pi \in \Pi(\mu, \nu)$ which is concentrated on

$$S = \{(x,y) \in X \times Y :\ c(x,y) < \infty\}.$$

Then μ and ν are c-compatible.

S. Artstein-Avidan, S. Sadovsky, K. W.

Theorem

Let $X = Y$ be a Polish space, let $c : X \times Y \to (-\infty, \infty]$ be a **continuous** *and symmetric cost function, essentially bounded from below with respect to probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. Assume μ and ν are strongly c-compatible, namely satisfy that for any measurable $A \subset X$ we have*

$$\mu(A) + \nu(\{y \in Y : \forall x \in A,\ c(x, y) = \infty\}) < 1.$$

If there exists some finite cost plan transporting μ to ν, then there exists a c-class function φ and an optimal transport plan $\pi \in \Pi(\mu, \nu)$ concentrated on $\partial^c \varphi$.

Cost duality for sets

Definition

*Let X, Y be two sets and let $c : X \times Y \to (-\infty, \infty]$. Fix $t \in (-\infty, \infty]$ (which will be omitted in the notation as it is a fixed parameter). For $K \subset X$ define the c-**dual** set of K as*

$$K^c = \bigcap_{x \in K} \{y \in Y : c(x,y) \geq t\} = \{y \in Y : c(x,y) \geq t, \ \forall x \in K\}.$$

Cost duality for sets

Definition

*Let X, Y be two sets and let $c : X \times Y \to (-\infty, \infty]$. Fix $t \in (-\infty, \infty]$ (which will be omitted in the notation as it is a fixed parameter). For $K \subset X$ define the c-**dual** set of K as*

$$K^c = \bigcap_{x \in K} \{y \in Y : c(x,y) \geq t\} = \{y \in Y : c(x,y) \geq t, \ \forall x \in K\}.$$

Lemma

For every $K, L \subset X$, the following hold

1. $K \subset (K^c)^c = K^{cc}$,
2. *if $L \subset K$ then $K^c \subset L^c$,*
3. $K^c = K^{ccc}$.

Cost duality for sets

Definition

*Let X, Y be two sets and let $c : X \times Y \to (-\infty, \infty]$. Fix $t \in (-\infty, \infty]$ (which will be omitted in the notation as it is a fixed parameter). For $K \subset X$ define the c-**dual** set of K as*

$$K^c = \bigcap_{x \in K} \{y \in Y : c(x,y) \geq t\} = \{y \in Y : c(x,y) \geq t, \ \forall x \in K\}.$$

Definition

*Fix a cost function c. The c-**class** of sets consists of all closed sets $K \subset X$ such that there exists some $L \subset X$ with $K = L^c$.*

*For any set $K \subset X$ we define its c-**envelope** as the set K^{cc}, which is the smallest c-class set containing K.*

S. Artstein-Avidan, S. Sadovsky, K. W.

Theorem

Let $T : \mathcal{P}(X) \to \mathcal{P}(X)$ *be an order-reversing quasi-involution, that is, for every* $K, L \subset X$ *we have:*

(i) $K \subset TTK$
(ii) if $K \subset L$ *then* $TL \subset TK$*.*

Then, there exists a cost function $c : X \times X \to (-\infty, \infty]$ *such that* T *is induced by* c*, that is for all* $K \subset X$ *we have* $TK = K^c$*.*

Thank you for your attention!