Generalized transport costs and applications to concentration of measure

Nathaël Gozlan*

*Université Paris Est – Marne-la-Vallée

Convexity, probability and discrete structures, a geometric view point
October 26 2015

Introduction

Goal of the talk is to revisit the well known chain of implications

Curvature $\Rightarrow$ Log-Sobolev ineq. $\Rightarrow$ Talagrand transport ineq. $\Rightarrow$ Conc. of measure

in a discrete setting.

Based on joints works with Cyril Roberto, Paul-Marie Samson, Yan Shu and Prasad Tetali.

Outline

Curvature $\Rightarrow$ Log-Sobolev ineq. $\Rightarrow$ Talagrand ineq. $\Rightarrow$ Concentration of measure

I - Generalized transport costs and Kantorovich duality

II - Generalized transport cost inequalities and concentration of measure

III - Examples

IV - Links with Log-Sobolev inequalities

V - Links with discrete curvature

I – Generalized transport costs and Kantorovich duality

Notation

$(\mathcal{X}, d)$ is a polish space

$\mathcal{P}(\mathcal{X})$ denotes the set of all Borel probability measures on $\mathcal{X}$

Let $k : \mathcal{X} \times \mathcal{X} \to \mathbf{R}^+$ be a cost function.

$\rightsquigarrow$ $k(x, y)$ represents the cost of moving one unit of mass from x to y.

Optimal transport cost in the sense of Kantorovich

Let $k : \mathcal{X} \times \mathcal{X} \to \mathbf{R}^+$ be a cost function.

$\rightsquigarrow$ $k(x,y)$ represents the cost of moving one unit of mass from x to y.

Definition

If μ, ν are probability measures on $\mathcal{X}$,

$$\mathcal{T}_k(\mu,\nu) := \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{X}\times\mathcal{X}} k(x,y)\,\pi(dxdy),$$

Optimal transport cost in the sense of Kantorovich

Let $k : \mathcal{X} \times \mathcal{X} \to \mathbf{R}^+$ be a cost function.

$\rightsquigarrow$ $k(x, y)$ represents the cost of moving one unit of mass from x to y.

Definition

If μ, ν are probability measures on $\mathcal{X}$,

$$\mathcal{T}_k(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{X} \times \mathcal{X}} k(x, y)\, \pi(dxdy),$$

where $\Pi(\mu, \nu) = \{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X}) \text{ such that } \pi_1 = \mu \text{ and } \pi_2 = \nu\}$.

Optimal transport cost in the sense of Kantorovich

Let $k : \mathcal{X} \times \mathcal{X} \to \mathbf{R}^+$ be a cost function.

$\rightsquigarrow$ $k(x, y)$ represents the cost of moving one unit of mass from x to y.

Definition

If μ, ν are probability measures on $\mathcal{X}$,

$$\mathcal{T}_k(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \iint_{\mathcal{X} \times \mathcal{X}} k(x, y)\, \pi(dxdy),$$

where $\Pi(\mu, \nu) = \{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X}) \text{ such that } \pi_1 = \mu \text{ and } \pi_2 = \nu\}$.

Example : When $k(x, y) = d(x, y)^p$, with $p \geq 1$, then

$$W_p(\mu, \nu) := \left(\mathcal{T}_{d^p}\right)^{1/p}(\mu, \nu)$$

defines the Kantorovich-Rubinstein metric of order p.

Generalized optimal transport costs

Notation :if $\pi \in \Pi(\mu, \nu)$ is a coupling between μ and ν, one writes

$$\pi(dxdy) = \mu(dx)p_x(dy)$$

its disintegration with respect to its first marginal.

$\rightsquigarrow$ $p_x(dy)$ tells where the mass taken at x is re-allocated.

Generalized optimal transport costs

Notation : if $\pi \in \Pi(\mu, \nu)$ is a coupling between μ and ν, one writes

$$\pi(dxdy) = \mu(dx)p_x(dy)$$

its disintegration with respect to its first marginal.

$\rightsquigarrow$ $p_x(dy)$ tells where the mass taken at x is re-allocated.

Definition

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}^+$ be a cost function ; for all $\mu, \nu \in \mathcal{P}(\mathcal{X})$, one sets

$$\mathcal{T}_c(\nu|\mu) = \inf_{p \in P(\mu,\nu)} \int c(x, p_x)\, \mu(dx),$$

where $P(\mu, \nu)$ is the set of probability kernels $p_x(dy)$ such that

$$\mu p = \nu \qquad i.e. \int \mu(dx)p_x(A) = \nu(A), \forall A.$$

Examples

- Usual transport costs.
  If $c(x,p) = \int k(x,y)\, p(dy)$, then

$$\mathcal{T}_c(\nu|\mu) = \inf_p \iint k(x,y)\, p_x(dy)\mu(dx) = \inf_\pi \iint k(x,y)\, \pi(dxdy) = \mathcal{T}_k(\mu,\nu).$$

Examples

- Usual transport costs.
  If $c(x,p) = \int k(x,y)\, p(dy)$, then

$$\mathcal{T}_c(\nu|\mu) = \inf_p \iint k(x,y)\, p_x(dy)\mu(dx) = \inf_\pi \iint k(x,y)\, \pi(dxdy) = \mathcal{T}_k(\mu,\nu).$$

- Transport costs $\widetilde{\mathcal{T}}$.
  If $c(x,p) = \alpha\left(\int d(x,y)\, p(dy)\right)$, with α a convex function, one denotes

$$\widetilde{\mathcal{T}}_\alpha(\nu|\mu) := \inf_p \int \alpha\left(\int d(x,y)\, p_x(dy)\right)\, \mu(dx).$$

Examples

- Usual transport costs.
  If $c(x,p) = \int k(x,y)\, p(dy)$, then
  $$\mathcal{T}_c(\nu|\mu) = \inf_p \iint k(x,y)\, p_x(dy)\mu(dx) = \inf_\pi \iint k(x,y)\, \pi(dxdy) = \mathcal{T}_k(\mu,\nu).$$
- Transport costs $\widetilde{\mathcal{T}}$.
  If $c(x,p) = \alpha\left(\int d(x,y)\, p(dy)\right)$, with α a convex function, one denotes
  $$\widetilde{\mathcal{T}}_\alpha(\nu|\mu) := \inf_p \int \alpha\left(\int d(x,y)\, p_x(dy)\right)\, \mu(dx).$$
- Transport costs $\overline{\mathcal{T}}$.
  If $c(x,p) = \theta(x - \int y\, p(dy))$, with $\theta : \mathbf{R}^d \to \mathbf{R}$ a convex function, one denotes
  $$\overline{\mathcal{T}}_\theta(\nu|\mu) := \inf_p \int \theta\left(x - \int y\, p_x(dy)\right)\, \mu(dx).$$

Examples

- Usual transport costs.
  If $c(x,p) = \int k(x,y)\,p(dy)$, then
  $$\mathcal{T}_c(\nu|\mu) = \inf_p \iint k(x,y)\,p_x(dy)\mu(dx) = \inf_\pi \iint k(x,y)\,\pi(dxdy) = \mathcal{T}_k(\mu,\nu).$$
- Transport costs $\widetilde{\mathcal{T}}$.
  If $c(x,p) = \alpha\left(\int d(x,y)\,p(dy)\right)$, with α a convex function, one denotes
  $$\widetilde{\mathcal{T}}_\alpha(\nu|\mu) := \inf_p \int \alpha\left(\int d(x,y)\,p_x(dy)\right)\,\mu(dx).$$
- Transport costs $\overline{\mathcal{T}}$.
  If $c(x,p) = \theta(x - \int y\,p(dy))$, with $\theta : \mathbf{R}^d \to \mathbf{R}$ a convex function, one denotes
  $$\overline{\mathcal{T}}_\theta(\nu|\mu) := \inf_p \int \theta\left(x - \int y\,p_x(dy)\right)\,\mu(dx).$$

Notation : When
$$c(x,p) = \int d^s(x,y)\,p(dy) \text{ or } \left(\int d(x,y)\,p(dy)\right)^s \text{ or } \left\|x - \int y\,p(dy)\right\|^s,$$
one uses the notations $\mathcal{T}_s$, $\widetilde{\mathcal{T}}_s$ and $\overline{\mathcal{T}}_s$.

Comparisons

For a given $s \geq 1$, and $\| \cdot \|$ a given norm on $\mathbf{R}^d$, applying Jensen yields to

$$\left\| x - \int y \, p(dy) \right\|^s \leq \left(\int \|x - y\| \, p(dy) \right)^s \leq \int \|x - y\|^s \, p(dy)$$

Comparisons

For a given $s \geq 1$, and $\|\cdot\|$ a given norm on $\mathbf{R}^d$, applying Jensen yields to

$$\left\| x - \int y \, p(dy) \right\|^s \leq \left(\int \|x - y\| \, p(dy) \right)^s \leq \int \|x - y\|^s \, p(dy)$$

and so

$$\overline{\mathcal{T}}_s(\nu|\mu) \leq \widetilde{\mathcal{T}}_s(\nu|\mu) \leq \mathcal{T}_s(\nu, \mu).$$

For a given $s \geq 1$, and $\|\cdot\|$ a given norm on $\mathbf{R}^d$, applying Jensen yields to

$$\left\| x - \int y\, p(dy) \right\|^s \leq \left(\int \|x-y\|\, p(dy) \right)^s \leq \int \|x-y\|^s\, p(dy)$$

and so

$$\overline{\mathcal{T}}_s(\nu|\mu) \leq \widetilde{\mathcal{T}}_s(\nu|\mu) \leq \mathcal{T}_s(\nu,\mu).$$

Similarly,

$$(\mathcal{T}_1(\nu,\mu))^s \leq \widetilde{\mathcal{T}}_s(\nu|\mu).$$

Duality

Recall Kantorovich's duality theorem

Theorem

If $k : \mathcal{X} \times \mathcal{X} \to \mathbf{R}^+$ is lower semicontinuous, then

$$\mathcal{T}_k(\mu,\nu) = \sup_{f\in\mathcal{C}_b(\mathcal{X})} \left\{ \int Qf\, d\mu - \int f\, d\nu \right\},$$

where

$$Qf(x) = \inf_{y\in X}\{f(y) + k(x,y)\}, \qquad x \in \mathcal{X}.$$

Duality for generalized transport costs

Kantorovich's duality theorem extends as follows

Theorem [GRST '15]

Assume that $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}^+$ is convex with respect to p, then under some mild regularity conditions, it holds

$$\mathcal{T}_c(\nu|\mu) = \sup_{f \in \mathcal{C}_b(\mathcal{X})} \left\{ \int Rf \, d\mu - \int f \, d\nu \right\},$$

where

$$Rf(x) = \inf_{p \in \mathcal{P}(\mathcal{X})} \left\{ \int f(y) \, p(dy) + c(x,p) \right\}, \qquad x \in \mathcal{X}.$$

Duality holds in particular for the costs families $\widetilde{\mathcal{T}}$ and $\overline{\mathcal{T}}$.

Duality for $\overline{\mathcal{T}}$ transport costs

Transport costs of the form $\overline{\mathcal{T}}$ are naturally related to convex functions :

Duality for $\overline{\mathcal{T}}$ transport costs

Transport costs of the form $\overline{\mathcal{T}}$ are naturally related to convex functions :

Corollary

Let $\theta : \mathbf{R}^d \to \mathbf{R}$ be a convex function and $c(x,p) = \theta(x - \int y\, p(dy))$; then

$$\overline{\mathcal{T}}_\theta(\nu|\mu) = \sup_\varphi \left\{ \int Q\varphi\, d\mu - \int \varphi\, d\nu \right\},$$

where the supremum runs over the set of all convex functions bounded from below and

$$Q\varphi(x) = \inf_{y\in\mathbf{R}^d} \{\varphi(y) + \theta(x-y)\}, \qquad x \in \mathbf{R}^d.$$

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\overline{\mathcal{T}}_1(\nu|\mu) = \inf_{p\in P(\mu,\nu)} \int \left\| x - \int y \, p_x(dy) \right\| \mu(dx)$$

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\begin{aligned}\overline{\mathcal{T}}_1(\nu|\mu) &= \inf_{p\in P(\mu,\nu)} \int \left\| x - \int y\, p_x(dy) \right\| \mu(dx)\\ &= \inf_{(X_0,X_1),X_0\sim\mu,X_1\sim\nu} \mathbb{E}\left[\|X_0 - \mathbb{E}[X_1|X_0]\|\right].\end{aligned}$$

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\begin{aligned}\overline{\mathcal{T}}_1(\nu|\mu) &= \inf_{p\in P(\mu,\nu)} \int \left\| x - \int y\, p_x(dy)\right\| \mu(dx)\\ &= \inf_{(X_0,X_1), X_0\sim\mu, X_1\sim\nu} \mathbb{E}\left[\|X_0 - \mathbb{E}[X_1|X_0]\|\right].\end{aligned}$$

Therefore, $\overline{\mathcal{T}}_1(\nu|\mu) = 0$ if and only if there exists a martingale $(X_i)_{i\in\{0,1\}}$ with marginals μ and ν.

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\overline{\mathcal{T}}_1(\nu|\mu) = \inf_{p\in P(\mu,\nu)} \int \left\| x - \int y\, p_x(dy) \right\| \mu(dx)$$
$$= \inf_{(X_0,X_1), X_0\sim\mu, X_1\sim\nu} \mathbb{E}\left[\|X_0 - \mathbb{E}[X_1|X_0]\|\right].$$

Therefore, $\overline{\mathcal{T}}_1(\nu|\mu) = 0$ if and only if there exists a martingale $(X_i)_{i\in\{0,1\}}$ with marginals μ and ν.

For the cost $\overline{\mathcal{T}}_1$ the duality specializes to

$$\overline{\mathcal{T}}_1(\nu|\mu) = \sup_{\varphi} \left\{ \int \varphi\, d\mu - \int \varphi\, d\nu \right\},$$

where the supremum runs over the set of all 1-Lipschitz and convex functions.

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\overline{\mathcal{T}}_1(\nu|\mu) = \inf_{p \in P(\mu,\nu)} \int \left\| x - \int y\, p_x(dy) \right\| \mu(dx)$$
$$= \inf_{(X_0,X_1), X_0\sim\mu, X_1\sim\nu} \mathbb{E}\left[\|X_0 - \mathbb{E}[X_1|X_0]\|\right].$$

Therefore, $\overline{\mathcal{T}}_1(\nu|\mu) = 0$ if and only if there exists a martingale $(X_i)_{i\in\{0,1\}}$ with marginals μ and ν.

For the cost $\overline{\mathcal{T}}_1$ the duality specializes to

$$\overline{\mathcal{T}}_1(\nu|\mu) = \sup_{\varphi} \left\{ \int \varphi\, d\mu - \int \varphi\, d\nu \right\},$$

where the supremum runs over the set of all 1-Lipschitz and convex functions.

Corollary [Strassen]

There exists a martingale $(X_i)_{i\in\{0,1\}}$ such that $X_o \sim \mu$ and $X_1 \sim \nu$ if and only if μ is majorized by ν in the convex order

.

A quick application to martingale optimal transport

Let $\|\cdot\|$ be some norm on $\mathbf{R}^d$; recall

$$\overline{\mathcal{T}}_1(\nu|\mu) = \inf_{p \in P(\mu,\nu)} \int \left\| x - \int y \, p_x(dy) \right\| \mu(dx)$$
$$= \inf_{(X_0,X_1),X_0\sim\mu,X_1\sim\nu} \mathbb{E}\left[\|X_0 - \mathbb{E}[X_1|X_0]\|\right].$$

Therefore, $\overline{\mathcal{T}}_1(\nu|\mu) = 0$ if and only if there exists a martingale $(X_i)_{i\in\{0,1\}}$ with marginals μ and ν.

For the cost $\overline{\mathcal{T}}_1$ the duality specializes to

$$\overline{\mathcal{T}}_1(\nu|\mu) = \sup_{\varphi} \left\{ \int \varphi \, d\mu - \int \varphi \, d\nu \right\},$$

where the supremum runs over the set of all 1-Lipschitz and convex functions.

Corollary [Strassen]

There exists a martingale $(X_i)_{i\in\{0,1\}}$ such that $X_o \sim \mu$ and $X_1 \sim \nu$ if and only if μ is majorized by ν in the convex order i.e. $\int \varphi \, d\mu \leq \int \varphi \, d\nu$ for all convex function φ.

II- Generalized transport cost inequalities and concentration of measure

Definitions

Definition

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}$ be a cost function ; a probability measure μ is said to satisfy the inequality $\mathbf{T}_c(a_1, a_2)$ if

$$\mathcal{T}_c(\nu_1|\nu_2) \leq a_1 H(\nu_1|\mu) + a_2 H(\nu_2|\mu), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\mathcal{X}),$$

where the relative entropy of ν with respect to μ is defined by

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu}\, d\nu$$

if $\nu \ll \mu$ (and $+\infty$ otherwise).

Definitions

Definition

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}$ be a cost function ; a probability measure μ is said to satisfy the inequality $\mathbf{T}_c(a_1, a_2)$ if

$$\mathcal{T}_c(\nu_1|\nu_2) \leq a_1 H(\nu_1|\mu) + a_2 H(\nu_2|\mu), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\mathcal{X}),$$

where the relative entropy of ν with respect to μ is defined by

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu}\, d\nu$$

if $\nu \ll \mu$ (and $+\infty$ otherwise).

Notation :
When $\mathcal{T}_c = \widetilde{\mathcal{T}}_\alpha$, i.e. $c(x,p) = \alpha\left(\int d(x,y)\, p(dy)\right)$, one writes $\widetilde{\mathbf{T}}_\alpha(a_1, a_2)$.

Definitions

Definition

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}$ be a cost function ; a probability measure μ is said to satisfy the inequality $\mathbf{T}_c(a_1, a_2)$ if

$$\mathcal{T}_c(\nu_1|\nu_2) \leq a_1 H(\nu_1|\mu) + a_2 H(\nu_2|\mu), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\mathcal{X}),$$

where the relative entropy of ν with respect to μ is defined by

$$H(\nu|\mu) = \int \log \frac{d\nu}{d\mu}\, d\nu$$

if $\nu \ll \mu$ (and $+\infty$ otherwise).

Notation :

When $\mathcal{T}_c = \widetilde{\mathcal{T}}_\alpha$, i.e. $c(x,p) = \alpha\left(\int d(x,y)\, p(dy)\right)$, one writes $\widetilde{\mathbf{T}}_\alpha(a_1, a_2)$.

When $\mathcal{T}_c = \overline{\mathcal{T}}_\alpha$, i.e. $c(x,p) = \theta\left(x - \int y\, p(dy)\right)$, one writes $\overline{\mathbf{T}}_\theta(a_1, a_2)$.

Two classical examples

(1) Talagrand's inequality
A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu,\mu) \leq CH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Two classical examples

(1) Talagrand's inequality

A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu,\mu) \leq CH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Equivalently,

$$\mathcal{T}_2(\nu_1,\nu_2) \leq \frac{C}{s}H(\nu_1|\mu) + \frac{C}{1-s}H(\nu_2|\mu), \qquad \forall s \in (0,1), \forall \nu_1,\nu_2 \in \mathcal{P}(\mathcal{X}).$$

Two classical examples

(1) Talagrand's inequality

A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu,\mu) \leq CH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Equivalently,

$$\mathcal{T}_2(\nu_1,\nu_2) \leq \frac{C}{s}H(\nu_1|\mu) + \frac{C}{1-s}H(\nu_2|\mu), \qquad \forall s \in (0,1), \forall \nu_1,\nu_2 \in \mathcal{P}(\mathcal{X}).$$

$\mathbf{T}_2$ is true for instance for Gaussian measures (Talagrand '96), or more generally for uniformly log-concave probability measures.

Two classical examples

(1) Talagrand's inequality

A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu,\mu) \leq CH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Equivalently,

$$\mathcal{T}_2(\nu_1,\nu_2) \leq \frac{C}{s}H(\nu_1|\mu) + \frac{C}{1-s}H(\nu_2|\mu), \qquad \forall s \in (0,1), \forall \nu_1,\nu_2 \in \mathcal{P}(\mathcal{X}).$$

$\mathbf{T}_2$ is true for instance for Gaussian measures (Talagrand '96), or more generally for uniformly log-concave probability measures.

$\mathbf{T}_2$ is never satisfied by discrete probability measures.

Two classical examples

(1) Talagrand's inequality

A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu, \mu) \leq CH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Equivalently,

$$\mathcal{T}_2(\nu_1, \nu_2) \leq \frac{C}{s} H(\nu_1|\mu) + \frac{C}{1-s} H(\nu_2|\mu), \qquad \forall s \in (0,1), \forall \nu_1, \nu_2 \in \mathcal{P}(\mathcal{X}).$$

$\mathbf{T}_2$ is true for instance for Gaussian measures (Talagrand '96), or more generally for uniformly log-concave probability measures.

$\mathbf{T}_2$ is never satisfied by discrete probability measures.

(2) Marton's inequality

Theorem [Marton '96]

Any probability measure μ on $\mathcal{X}$ satisfies the inequality $\mathbf{T}_c(4,4)$, with

$$c(x,p) = \left(\int \mathbf{1}_{x \neq y} \, p(dy) \right)^2 = p(\mathcal{X} \setminus \{x\})^2$$

Two classical examples

(1) Talagrand's inequality

A probability measure μ is said to satisfy the inequality $\mathbf{T}_2(C)$ if

$$\mathcal{T}_2(\nu,\mu) \leq C H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Equivalently,

$$\mathcal{T}_2(\nu_1,\nu_2) \leq \frac{C}{s} H(\nu_1|\mu) + \frac{C}{1-s} H(\nu_2|\mu), \qquad \forall s \in (0,1), \forall \nu_1,\nu_2 \in \mathcal{P}(\mathcal{X}).$$

$\mathbf{T}_2$ is true for instance for Gaussian measures (Talagrand '96), or more generally for uniformly log-concave probability measures.

$\mathbf{T}_2$ is never satisfied by discrete probability measures.

(2) Marton's inequality

Theorem [Marton '96]

Any probability measure μ on $\mathcal{X}$ satisfies the inequality $\mathbf{T}_c(4,4)$, with

$$c(x,p) = \left(\int \mathbf{1}_{x\neq y}\, p(dy) \right)^2 = p(\mathcal{X} \setminus \{x\})^2$$

This refinement of Pinsker's inequality was introduced by Marton to recover Talagrand's convex-hull distance concentration results for product probability measures.

Equivalent dual formulation

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}^+$ be a cost function for which our Kantorovich type duality result holds.

Theorem [GRST'15]

The following are equivalent :

(1) μ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) μ satisfies

$$\left(\int e^{Rf/a_2}\, d\mu\right)^{a_2} \left(\int e^{-f/a_1}\, d\mu\right)^{a_1} \leq 1, \qquad \forall f \geq 0,$$

where $Rf(x) = \inf_{p \in \mathcal{P}(\mathcal{X})}\{\int f(y)\, p(dy) + c(x,p)\}$.

Equivalent dual formulation

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}^+$ be a cost function for which our Kantorovich type duality result holds.

Theorem [GRST'15]

The following are equivalent :

(1) μ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) μ satisfies

$$\left(\int e^{Rf/a_2}\, d\mu\right)^{a_2} \left(\int e^{-f/a_1}\, d\mu\right)^{a_1} \leq 1, \qquad \forall f \geq 0,$$

where $Rf(x) = \inf_{p \in \mathcal{P}(\mathcal{X})}\{\int f(y)\, p(dy) + c(x,p)\}$.

This is an easy extension of a result by Bobkov and Götze ('99) for usual transport inequalities.

Equivalent dual formulation

Let $c : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbf{R}^+$ be a cost function for which our Kantorovich type duality result holds.

Theorem [GRST'15]

The following are equivalent :

(1) μ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) μ satisfies

$$\left(\int e^{Rf/a_2}\,d\mu\right)^{a_2}\left(\int e^{-f/a_1}\,d\mu\right)^{a_1} \leq 1, \qquad \forall f \geq 0,$$

where $Rf(x) = \inf_{p\in\mathcal{P}(\mathcal{X})}\{\int f(y)\,p(dy) + c(x,p)\}$.

This is an easy extension of a result by Bobkov and Götze ('99) for usual transport inequalities.

The second form is a generalization of the infimum-convolution inequalities introduced by Maurey ('91). In particular, $\overline{\mathbf{T}}_\theta(a_1, a_2)$ is equivalent to the so-called convex (τ)-property.

General link to concentration

Theorem [GRST '15]

The following are equivalent :

(1) $\mu \in \mathcal{P}(\mathcal{X})$ satisfies $\mathbf{T}_c(a_1, a_2)$.

General link to concentration

Theorem [GRST '15]

The following are equivalent :

(1) $\mu \in \mathcal{P}(\mathcal{X})$ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) For all integer n, $\mu^n = \mu \otimes \cdots \otimes \mu$ satisfies $\mathbf{T}_{c^n}(a_1, a_2)$, where the cost c^n is defined by

$$c^n(x,p) := \sum_{i=1}^{n} c(x_i, p_i), \qquad \forall x = (x_1, \dots, x_n) \in \mathcal{X}^n, \qquad \forall p \in \mathcal{P}(\mathcal{X}^n),$$

where p_i denotes the i-th marginal of p.

Theorem [GRST '15]

The following are equivalent :

(1) $\mu \in \mathcal{P}(\mathcal{X})$ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) For all integer n, $\mu^n = \mu \otimes \cdots \otimes \mu$ satisfies $\mathbf{T}_{c^n}(a_1, a_2)$, where the cost c^n is defined by

$$c^n(x,p) := \sum_{i=1}^n c(x_i, p_i), \qquad \forall x = (x_1, \ldots, x_n) \in \mathcal{X}^n, \qquad \forall p \in \mathcal{P}(\mathcal{X}^n),$$

where p_i denotes the i-th marginal of p.

(3) Dimension free concentration property : For all integer n, for all $A \subset \mathcal{X}^n$,

$$\mu^n(\mathcal{X}^n \setminus A_t)^{a_2} \mu^n(A)^{a_1} \leq e^{-t}, \qquad \forall t \geq 0,$$

where $A_t = \{x \in \mathcal{X}^n : c^n(x, A) \leq t\}$ and

$$c^n(x, A) = \inf_{p \in \mathcal{P}(\mathcal{X}^n), p(A)=1} c^n(x,p).$$

General link to concentration

Theorem [GRST '15]

The following are equivalent :

(1) $\mu \in \mathcal{P}(\mathcal{X})$ satisfies $\mathbf{T}_c(a_1, a_2)$.

(2) For all integer n, $\mu^n = \mu \otimes \cdots \otimes \mu$ satisfies $\mathbf{T}_{c^n}(a_1, a_2)$, where the cost c^n is defined by

$$c^n(x, p) := \sum_{i=1}^{n} c(x_i, p_i), \qquad \forall x = (x_1, \ldots, x_n) \in \mathcal{X}^n, \qquad \forall p \in \mathcal{P}(\mathcal{X}^n),$$

where p_i denotes the i-th marginal of p.

(3) Dimension free concentration property : For all integer n, for all $A \subset \mathcal{X}^n$,

$$\mu^n(\mathcal{X}^n \setminus A_t)^{a_2} \mu^n(A)^{a_1} \leq e^{-t}, \qquad \forall t \geq 0,$$

where $A_t = \{x \in \mathcal{X}^n : c^n(x, A) \leq t\}$ and

$$c^n(x, A) = \inf_{p \in \mathcal{P}(\mathcal{X}^n), p(A)=1} c^n(x, p).$$

(1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (3) adapts well known arguments due to Marton.

Comparison of enlargements

Let us compare (for $n = 1$) the enlargments of a given set $A \subset \mathbf{R}^d$ for the transport costs $\mathcal{T}_2, \overline{\mathcal{T}}_2$ and $\widetilde{\mathcal{T}}_2$:

Comparison of enlargements

Let us compare (for $n = 1$) the enlargments of a given set $A \subset \mathbf{R}^d$ for the transport costs $\mathcal{T}_2, \overline{\mathcal{T}}_2$ and $\widetilde{\mathcal{T}}_2$:

- When $c(x,p) = \int \|x-y\|^2 \, p(dy)$, then

$$A_t = \{x \in \mathbf{R}^d : \exists y \in A, \|x-y\| \leq t^{1/2}\}$$

Comparison of enlargements

Let us compare (for $n = 1$) the enlargments of a given set $A \subset \mathbf{R}^d$ for the transport costs $\mathcal{T}_2, \overline{\mathcal{T}}_2$ and $\widetilde{\mathcal{T}}_2$:

- When $c(x, p) = \int \|x - y\|^2 \, p(dy)$, then

$$A_t = \{x \in \mathbf{R}^d : \exists y \in A, \|x - y\| \leq t^{1/2}\}$$

- When $c(x, p) = \left(\int \|x - y\| \, p(dy)\right)^2$, then

$$A'_t = \{x \in \mathbf{R}^d : \exists p \text{ with } p(A) = 1 \text{ such that } \int \|x - y\| \, p(dy) \leq t^{1/2}\}$$

Comparison of enlargements

Let us compare (for $n = 1$) the enlargments of a given set $A \subset \mathbf{R}^d$ for the transport costs $\mathcal{T}_2, \overline{\mathcal{T}}_2$ and $\widetilde{\mathcal{T}}_2$:

- When $c(x,p) = \int \|x-y\|^2 \, p(dy)$, then

$$A_t = \{x \in \mathbf{R}^d : \exists y \in A, \|x-y\| \leq t^{1/2}\}$$

- When $c(x,p) = \left(\int \|x-y\| \, p(dy)\right)^2$, then

$$A'_t = \{x \in \mathbf{R}^d : \exists p \text{ with } p(A) = 1 \text{ such that } \int \|x-y\| \, p(dy) \leq t^{1/2}\}$$

- When $c(x,p) = \|x - \int y \, p(dy)\|^2$, then

$$A''_t = \{x \in \mathbf{R}^d : \exists y \in \overline{\mathrm{conv}}(A), \|x-y\| \leq t^{1/2}\}.$$

Comparison of enlargements

Let us compare (for $n = 1$) the enlargments of a given set $A \subset \mathbf{R}^d$ for the transport costs $\mathcal{T}_2, \overline{\mathcal{T}}_2$ and $\widetilde{\mathcal{T}}_2$:

- When $c(x, p) = \int \|x - y\|^2 \, p(dy)$, then
$$A_t = \{x \in \mathbf{R}^d : \exists y \in A, \|x - y\| \leq t^{1/2}\}$$
- When $c(x, p) = \left(\int \|x - y\| \, p(dy)\right)^2$, then
$$A_t' = \{x \in \mathbf{R}^d : \exists p \text{ with } p(A) = 1 \text{ such that } \int \|x - y\| \, p(dy) \leq t^{1/2}\}$$
- When $c(x, p) = \|x - \int y \, p(dy)\|^2$, then
$$A_t'' = \{x \in \mathbf{R}^d : \exists y \in \overline{\mathrm{conv}}(A), \|x - y\| \leq t^{1/2}\}.$$

So
$$A_t \subset A_t' \subset A_t''.$$

Particular case 1 : $\mathbf{T}_2$

Corollary [G. '09 (for (2) ⇒ (1))]

The following are equivalent :

(1) μ satisfies $\mathbf{T}_2(C)$.

(2) There is some $t_o \geq 0$ such that for all $n \geq 1$,

$$\mu^n(f > m(f) + t) \leq e^{-(t-t_o)^2/C}, \qquad \forall t \geq t_o,$$

for all $f : \mathcal{X}^n \to \mathbf{R}$ 1-Lipschitz with respect to the distance d_2 defined by

$$d_2(x,y) = \left[\sum_{i=1}^n d(x_i,y_i)^2\right]^{1/2}, \qquad x,y \in \mathcal{X}^n.$$

Particular case 2 : $\overline{\mathbf{T}}_2$

Corollary

The following are equivalent :

(1) μ satisfies $\overline{\mathbf{T}}_2(C, C)$ (with respect to some norm $\|\cdot\|$ on $\mathbf{R}^d$).

(2) There is some $t_o \geq 0$ and some C' such that for all $n \geq 1$,

$$\mu^n(f > m(f) + t) \leq e^{-(t-t_o)^2/C'}, \qquad \forall t \geq t_o,$$

for all convex or concave $f : (\mathbf{R}^d)^n \to \mathbf{R}$ 1-Lipschitz with respect to the distance defined by

$$d_2(x, y) = \left[\sum_{i=1}^{n} \|x_i - y_i\|^2\right]^{1/2}, \qquad x, y \in \mathcal{X}^n.$$

C and C' are related through universal factors.

Particular case 3 : $\widetilde{\mathbf{T}}_2$

Particular case of Marton's inequality (the general case is still unexplored). Corresponds to

$$d(x,y) = \mathbb{1}_{x\neq y} \text{ and } c(x,p) = \left(\int \mathbb{1}_{x\neq y}\, p(dy)\right)^2.$$

It can be shown that

$$c^n(x,A) = (D_A(x))^2,$$

where $D_A(x)$ is Talagrand's convex-hull distance :

$$D_A(x) = \sup_{|\alpha|_2\leq 1} \inf_{y\in A} \sum_{i=1}^n \alpha_i \mathbb{1}_{x_i\neq y_i}.$$

So, Marton's transport inequality (i.e. $\widetilde{\mathbf{T}}_2(4,4)$) gives back Talagrand's convex hull-distance concentration inequality

Corollary [Talagrand '95]

$$\mu^n(A)\mu^n(\{D_A \geq t\}) \leq e^{-t^2/4}, \qquad \forall t \geq 0.$$

Applications to study of length of the longest increasing sub-sequence, length of the minimal spanning tree, traveling salesman problem, etc . . .

III- Examples

From Bernoulli to Poisson

Let μ_q be the Bernoulli measure on $\{0,1\}$ with parameter $q = \mu_q(1)$.

For any $s \in (0,1)$, let $\theta_{q,s}$ be the best function so that the $\overline{\mathbf{T}}$ inequality

$$\overline{\mathcal{T}}_{\theta_{q,s}}(\nu_1|\nu_2) \leq \frac{1}{1-s} H(\nu_1|\mu_q) + \frac{1}{s} H(\nu_2|\mu_q), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\{0,1\})$$

is satisfied. This function $\theta_{q,s}$ has been identified by P-M Samson ('03).

Tensorizing this inequality yields a transport inequality for μ_q^n on $\{0,1\}^n$.

Then considering the push-forward of μ_q^n under the map $x \mapsto \sum_{i=1}^n x_i$ gives a transport inequality for the binomial $\mu_{n,q}$ with parameters n and q :

$$\overline{\mathcal{T}}_{\theta_{n,q,s}}(\nu_1|\nu_2) \leq \frac{1}{1-s} H(\nu_1|\mu_{n,q}) + \frac{1}{s} H(\nu_2|\mu_{n,q}), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\{0,1\}^n).$$

with

$$\theta_{n,q,s}(u) = n\theta_{q,s}\left(\frac{u}{n}\right), \qquad u \in \mathbf{R}.$$

Finally, if $q = \lambda/n$ then $\mu_{n,q} \to p_\lambda$ Poisson of parameter λ and $\theta_{n,q,s} \to \beta_{\lambda,s}$, as $n \to \infty$.

Transport inequality for the Poisson measure

p_λ : the Poisson measure of parameter $\lambda > 0$,

Transport inequality for the Poisson measure

p_λ : the Poisson measure of parameter $\lambda > 0$, $p_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, $k \in \mathbb{N}$.

Theorem [GRST '15]

The Poisson measure p_λ satisfies $\overline{\mathbf{T}}_{\beta_{\lambda,s}}(1/(1-s), 1/s)$, $s \in (0,1)$, with

$$\beta_{\lambda,s}(h) = \left[\frac{\lambda}{s} w\left(\frac{u_s(h)}{\lambda}\right) + \frac{\lambda}{1-s}\left(\frac{h + u_s(h)}{\lambda}\right)\right] \mathbb{1}_{h \leq 0},$$

where $w(h) = (1-h)\log(1-h) + h$, and $u = u_s(h) \in [0, \lambda)$ is the unique solution of the following equation

$$(\lambda - u)^{1-s}(\lambda - u - h)^s = \lambda, \qquad h \leq 0.$$

Characterization in dimension one

Theorem [GRSST '15]

A probability measure μ on $\mathbf{R}$ satisfies $\overline{\mathbf{T}}_2(C, C)$ for some constant $C > 0$ if and only if the monotone rearrangement map U_μ pushing forward the symmetric exponential measure $\mu_1(dx) = \frac{1}{2}e^{-|x|}\,dx$ on μ is such that

$$(*) \qquad |U_\mu(x) - U_\mu(y)| \leq a(1 + \sqrt{|x-y|}), \qquad \forall x, y \in \mathbf{R}$$

There is a quantitative link between a and C.

Characterization in dimension one

Theorem [GRSST '15]

A probability measure μ on $\mathbf{R}$ satisfies $\overline{\mathbf{T}}_2(C, C)$ for some constant $C > 0$ if and only if the monotone rearrangement map U_μ pushing forward the symmetric exponential measure $\mu_1(dx) = \frac{1}{2}e^{-|x|}\,dx$ on μ is such that

$$(*) \qquad |U_\mu(x) - U_\mu(y)| \leq a(1 + \sqrt{|x-y|}), \qquad \forall x, y \in \mathbf{R}$$

There is a quantitative link between a and C.

For instance any probability with bounded support satisfies $\overline{\mathbf{T}}_2$ for some constant (also a consequence of Marton's result).

Characterization in dimension one

Theorem [GRSST '15]

A probability measure μ on $\mathbf{R}$ satisfies $\overline{\mathbf{T}}_2(C,C)$ for some constant $C>0$ if and only if the monotone rearrangement map U_μ pushing forward the symmetric exponential measure $\mu_1(dx)=\frac{1}{2}e^{-|x|}\,dx$ on μ is such that

$$(*) \qquad |U_\mu(x)-U_\mu(y)| \leq a(1+\sqrt{|x-y|}), \qquad \forall x,y\in\mathbf{R}$$

There is a quantitative link between a and C.

For instance any probability with bounded support satisfies $\overline{\mathbf{T}}_2$ for some constant (also a consequence of Marton's result).

In comparison, μ satisfies Talagrand's inequality $\mathbf{T}_2(C)$ if and only if it satisfies $(*)$ and Poincaré inequality (G. '12).

Other examples

Symmetric group, Poisson random measures. Works in progress.

IV- Links with log-Sobolev

Otto-Villani Theorem

Theorem [Otto-Villani '00]

If a probability measure μ on a smooth Riemannian manifold satisfies the logarithmic Sobolev inequality

$$\left(\int f^2 \log\left(\frac{f^2}{\int f^2\,d\mu}\right)\,d\mu\right) := \mathrm{Ent}_\mu(f^2) \leq C \int |\nabla f|^2\,d\mu, \qquad \forall \text{ smooth } f$$

then it satisfies Talagrand's inequality $\mathbf{T}_2(C)$.

Otto-Villani Theorem

Theorem [Otto-Villani '00]

If a probability measure μ on a smooth Riemannian manifold satisfies the logarithmic Sobolev inequality

$$\left(\int f^2 \log\left(\frac{f^2}{\int f^2\,d\mu}\right)\,d\mu\right) := \mathrm{Ent}_\mu(f^2) \leq C\int |\nabla f|^2\,d\mu, \qquad \forall \text{ smooth } f$$

then it satisfies Talagrand's inequality $\mathbf{T}_2(C)$.

Theorem [GRS'11]

A probability measure on $\mathbf{R}^d$ satisfies $\mathbf{T}_2(C)$ if and only if there is $D, K > 0$ such that

$$\mathrm{Ent}_\mu(e^f) \leq D\int |\nabla f|^2 e^f\,d\mu,$$

for all $\mathcal{C}^2$-smooth $f : \mathbf{R}^d \to \mathbf{R}$ such that $\mathrm{Hess}\, f \geq -K$.
The link between C, D and K is quantitative.

$\overline{\mathbf{T}}_2$ and log-Sobolev for log-convex/concave functions

Theorem [GRST'15]

A probability measure μ on $\mathbf{R}^d$ satisfies $\overline{\mathbf{T}}_2(C, C)$ if and only if there is $D, K > 0$ such that

$$\mathrm{Ent}_\mu(e^f) \leq D \int |\nabla f|^2 e^f \, d\mu,$$

for all convex $\mathcal{C}^2$-smooth function $f : \mathbf{R}^d \to \mathbf{R}$

$\overline{\mathbf{T}}_2$ and log-Sobolev for log-convex/concave functions

Theorem [GRST'15]

A probability measure μ on $\mathbf{R}^d$ satisfies $\overline{\mathbf{T}}_2(C, C)$ if and only if there is $D, K > 0$ such that

$$\mathrm{Ent}_\mu(e^f) \leq D \int |\nabla f|^2 e^f \, d\mu,$$

for all convex $\mathcal{C}^2$-smooth function $f : \mathbf{R}^d \to \mathbf{R}$ and for all concave $\mathcal{C}^2$-smooth function $f : \mathbf{R}^d \to \mathbf{R}$ such that $\mathrm{Hess}\, f \geq -K$.

$\overline{\mathbf{T}}_2$ and log-Sobolev for log-convex/concave functions

Theorem [GRST'15]

A probability measure μ on $\mathbf{R}^d$ satisfies $\overline{\mathbf{T}}_2(C, C)$ if and only if there is $D, K > 0$ such that

$$\mathrm{Ent}_\mu(e^f) \leq D \int |\nabla f|^2 e^f \, d\mu,$$

for all convex $\mathcal{C}^2$-smooth function $f : \mathbf{R}^d \to \mathbf{R}$ and for all concave $\mathcal{C}^2$-smooth function $f : \mathbf{R}^d \to \mathbf{R}$ such that $\mathrm{Hess}\, f \geq -K$.
The link between C, D and K is quantitative.

$\widetilde{\mathbf{T}}_2$ and log-Sobolev

Theorem [Shu '14]

A probability measure on a metric space $(\mathcal{X}, d)$ satisfies the following log-Sobolev inequality

$$(\mathbf{LSI}^-) \qquad \mathrm{Ent}_\mu(e^f) \leq C \int |\widetilde{\nabla}^- f|^2 e^f \, d\mu,$$

where

$$|\widetilde{\nabla}^- f|(x) = \sup_{y \in \mathcal{X}} \frac{[f(y) - f(x)]_-}{d(x,y)}$$

if and only if it satisfies

$$(\widetilde{\mathbf{T}}_2^-(D)) \qquad \widetilde{\mathcal{T}}_2(\mu|\nu) \leq DH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Theorem [Shu '14]

A probability measure on a metric space $(\mathcal{X}, d)$ satisfies the following log-Sobolev inequality

$$(\mathbf{LSI}^-) \qquad \mathrm{Ent}_\mu(e^f) \leq C \int |\widetilde{\nabla}^- f|^2 e^f \, d\mu,$$

where

$$|\widetilde{\nabla}^- f|(x) = \sup_{y \in \mathcal{X}} \frac{[f(y) - f(x)]_-}{d(x,y)}$$

if and only if it satisfies

$$(\widetilde{\mathbf{T}}_2^-(D)) \qquad \widetilde{\mathcal{T}}_2(\mu|\nu) \leq DH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

There is a quantitative link between C and D.

Theorem [Shu '14]

A probability measure on a metric space $(\mathcal{X}, d)$ satisfies the following log-Sobolev inequality

$$(\mathbf{LSI}^-) \qquad \mathrm{Ent}_\mu(e^f) \leq C \int |\widetilde{\nabla}^- f|^2 e^f \, d\mu,$$

where

$$|\widetilde{\nabla}^- f|(x) = \sup_{y \in \mathcal{X}} \frac{[f(y) - f(x)]_-}{d(x,y)}$$

if and only if it satisfies

$$(\widetilde{\mathbf{T}}_2^-(D)) \qquad \widetilde{\mathcal{T}}_2(\mu|\nu) \leq DH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

There is a quantitative link between C and D.
Moreover, $\mathbf{LSI}^+(C) \Rightarrow \widetilde{\mathbf{T}}_2^+(D)$.

$\widetilde{\mathbf{T}}_2$ and log-Sobolev

Theorem [Shu '14]

A probability measure on a metric space $(\mathcal{X}, d)$ satisfies the following log-Sobolev inequality

$$(\mathbf{LSI}^-) \qquad \mathrm{Ent}_\mu(e^f) \leq C \int |\widetilde{\nabla}^- f|^2 e^f \, d\mu,$$

where

$$|\widetilde{\nabla}^- f|(x) = \sup_{y \in \mathcal{X}} \frac{[f(y) - f(x)]_-}{d(x,y)}$$

if and only if it satisfies

$$(\widetilde{\mathbf{T}}_2^-(D)) \qquad \widetilde{\mathcal{T}}_2(\mu|\nu) \leq DH(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

There is a quantitative link between C and D.
Moreover, $\mathbf{LSI}^+(C) \Rightarrow \widetilde{\mathbf{T}}_2^+(D)$.

$(\mathbf{LSI}^\pm)$ are implied by usual log-Sobolev or modified log-Sobolev inequalities involving the Dirichlet form of a Markov chain on a graph.

V- Links with discrete curvature

Definition [Lott-Sturm-Villani]

A geodesic probability space (X, d, μ) is said to satisfy the curvature-dimension condition $\mathrm{CD}(K, \infty)$ if for all $\nu_0, \nu_1 \in \mathcal{P}_2(M)$, absolutely continuous with respect to μ, there is a W_2-geodesic $(\nu_t)_{t\in[0,1]}$ connecting ν_0 to ν_1 such that

$$H(\nu_t|\mu) \leq (1-t)H(\nu_0|\mu) + tH(\nu_1|\mu) - K\frac{t(1-t)}{2}W_2^2(\nu_0, \nu_1), \quad \forall t \in [0,1].$$

This definition is motivated by the following result (Mc Cann, Cordero-McCann-Schmuckenschläger, Sturm-Von Renesse, Lott-Villani, Sturm).

LSV

Definition [Lott-Sturm-Villani]

A geodesic probability space (X, d, μ) is said to satisfy the curvature-dimension condition $\mathrm{CD}(K, \infty)$ if for all $\nu_0, \nu_1 \in \mathcal{P}_2(M)$, absolutely continuous with respect to μ, there is a W_2-geodesic $(\nu_t)_{t\in[0,1]}$ connecting ν_0 to ν_1 such that

$$H(\nu_t|\mu) \leq (1-t)H(\nu_0|\mu) + tH(\nu_1|\mu) - K\frac{t(1-t)}{2}W_2^2(\nu_0,\nu_1), \quad \forall t \in [0,1].$$

This definition is motivated by the following result (Mc Cann, Cordero-McCann-Schmuckenschläger, Sturm-Von Renesse, Lott-Villani, Sturm).

Theorem

Let (M, g) be a complete connected Riemannian manifold and suppose that $\mu \in \mathcal{P}(M)$ is absolutely continuous with $\mu(dx) = e^{-V(x)}\,dx$.
The following are equivalent :

1. μ verifies the $\mathrm{CD}(K, \infty)$ condition, for some $K \in \mathbf{R}$
2. $\mathrm{Ric} + \mathrm{Hess}\, V \geq Kg$.

Transport cost and curvature

In a discrete setting, the LSV theory does not work (there is no W_2 geodesics).

Many different approaches were proposed recently by : Ollivier, Erbar-Maas-Mielke, Ollivier-Villani, Bonciocat-Sturm, Lin-Yau, Léonard, Hillion, . . . ,

Question : Is it possible to prove some displacement convexity properties of the relative entropy in a discrete setting using the costs $\widetilde{\mathcal{T}}_2$?

Theorem [GRST '13]

If μ is any probability measure on $\{0,1\}$, then for any integer $n \geq 1$, it holds

$$H(\nu_t|\mu^n) \leq (1-t)H(\nu_0|\mu^n) + H(\nu_1|\mu^n) - \frac{t(1-t)}{2n}\left(\widetilde{\mathcal{T}}_2(\nu_0|\nu_1) + \widetilde{\mathcal{T}}_2(\nu_1|\nu_0)\right), \forall t \in [0,1],$$

where ν_t is a binomial interpolation between ν_0 and ν_1 and $\widetilde{\mathcal{T}}_2$ is defined with respect to the graph distance ($= \ell_1$ distance) on $\{0,1\}^n$.

The same holds for products of complete graphs. Are there other examples ? We have some hope for the symmetric group. . .

Thank you for your attention !