Weak Optimal Transport

Nathaël Gozlan

Université de Paris
Kantorovich Initiative Seminar
27 january 2022

Outline

Works in collaboration with P. Choné, M. Fathi, N. Juillet, F. Kramarz, M. Prod'homme, C. Roberto, P-M Samson, Y. Shu and P. Tetali

I - Weak Optimal Transport : examples and general results

II - WOT and contraction properties of Brenier map

III - WOT and concentration of measure

IV - One word on WOT with unnormalized kernels

Weak Optimal Transport : examples and general results

Optimal Transport - classical definition

Let $\omega : E \times E \to \mathbb{R}^+$ be a measurable function on a Polish space (E, d).

Definition

The optimal transport cost between two probability measures μ and ν is given by

$$\mathcal{T}_\omega(\nu, \mu) = \inf_{\pi \in \Pi(\mu,\nu)} \iint_{E \times E} \omega(x, y) d\pi(x, y),$$

where $\Pi(\mu, \nu)$ denotes the set of probability measures π on $E \times E$ having μ and ν as marginals (called 'transport plans between μ and ν').

Equivalently

$$\mathcal{T}_\omega(\nu, \mu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[\omega(X, Y)]$$

Classical Examples : Kantorovich distances of order $p \geq 1$

$$W_p^p(\nu, \mu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[d^p(X, Y)].$$

Weak Optimal Transport

G.-Roberto-Samson-Tetali (2017), Alibert-Bouchitté-Champion (2018)

Let $\pi \in \Pi(\mu, \nu)$ be a transport plan between μ and ν written in disintegrated form

$$d\pi(x, y) = d\mu(x)dp_x(y),$$

with $x \mapsto p_x$ a transition kernel (μ a.s unique).

If $\omega : E \times E \to \mathbb{R}^+$ is a cost function then

$$\iint \omega(x, y)\, d\pi(x, y) = \int \left(\int \omega(x, y)\, dp_x(y) \right) d\mu(x).$$

In other words, transports of mass coming from x are penalized through their mean cost : $\int \omega(x, y)\, dp_x(y)$.

Idea of WOT :introduce more general penalizations.

Weak Optimal Transport

Let $\mathcal{P}(E)$ denote the set of all probability measures on E.

Definition

Let $c : E \times \mathcal{P}(E) \to \mathbb{R}^+ \cup \{+\infty\}$; the weak optimal transport cost $\mathcal{T}_c(\nu|\mu)$ is defined by

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

where $\mathcal{P}(\mu, \nu)$ is the set of all probability kernels p such that $\mu p = \nu$.

Classical transport :

$$c(x, p) = \int \omega(x, y)\, dp(y).$$

In all useful examples, the function c is convex in p.

Comments

- First examples go back to the works of K. Marton (1996) on concentration of measure.
- The framework of weak transport contains many variants of the transport problem : Schrödinger transport problem, martingale transport problem, semi-martingale transport problem,. . .
- General tools (duality, cyclical monotonicity) have been developed to study weak transport problems. See Backhoff-Veraguas, Beiglböck, Pammer (2019).

  Nice survey paper by Backhoff-Veraguas and Pammer (2020).

Examples

(1) **Barycentric transport** : $E = \mathbb{R}^n$ and

$$c(x,p) = \theta\left(x - \int y\,dp(y)\right),$$

where $\theta : \mathbb{R}^n \to \mathbb{R}^+$ (convex).
We will denote by $\overline{\mathcal{T}}_\theta(\nu|\mu)$ the corresponding weak optimal transport cost.

(2) **Transport with martingale constraints** : $E = \mathbb{R}^n$ and

$$c(x,p) = \begin{cases} \int \omega(x,y)\,dp(y) & \text{if } \int y\,dp(y) = x \\ +\infty & \text{otherwise} \end{cases}$$

Examples

(3) **Entropic regularized transport / Schrödinger bridges** :
Let R be a reference probability measure on $E \times E$

$$\mathcal{T}_H(\nu|\mu) = \inf_{\pi \in \Pi(\mu,\nu)} H(\pi|R),$$

where H is the relative entropy defined by

$$H(\pi|R) = \int \log \frac{d\pi}{dR}\, d\pi$$

if $\pi \ll R$ (and $+\infty$ otherwise).

Writing $d\pi(x,y) = d\mu(x)dp_x(y)$ and $dR(x,y) = dm(x)dr_x(y)$, one gets

$$H(\pi|R) = H(\mu|m) + \int H(p_x|r_x)\, d\mu(x) := H(\mu|m) + \int c(x,p_x)\, d\mu(x)$$

'Zero noise limit' : Mikami, Thieullen, Léonard, Carlier-Duval-Peyré,. . .
Applications : Cutturi, Peyré,. . .
Functional inequalities : Gentil-Léonard-Ripani, Conforti-Ripani, Gigli-Tamanini,. . .

(4) . . .

A general duality result

Theorem (Backhoff-Veraguas - Beiglboeck - Pammer (2018))

If $c : E \times \mathcal{P}(E) \to \mathbb{R} \cup \{+\infty\}$ is jointly lower semi-continuous, lower bounded and convex in p, then

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

with

$$R_c f(x) = \inf_{p \in \mathcal{P}(E)} \left\{ \int f \, dp + c(x, p) \right\}, \qquad x \in E.$$

Improves G.-Roberto-Samson-Tetali (2017) and Alibert-Bouchitté-Champion (2018).
Links with backward linear mass transfers Bowles-Ghoussoub (2019).

Duality holds under more general conditions on the cost function : μ, ν have finite k-th moment and c is lower semicontinuous w.r.t W_k topology, $k \geq 1$.

Duality for barycentric transport costs

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

Duality for barycentric transport costs

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

Corollary

Let $\theta : \mathbb{R}^n \to \mathbb{R}$ 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

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

A proof of Strassen theorem using barycentric L_1 cost

Notation : $\mathcal{P}_1(\mathbb{R}^n)$ the set of probability measures with a finite first moment.

Definition

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^n)$; μ is dominated by ν in the convex order, denoted by $\mu \leq_c \nu$, if

$$\int f \, d\mu \leq \int f \, d\nu, \qquad \text{for all convex function } f : \mathbb{R}^n \to \mathbb{R}.$$

Theorem (Strassen (1965))

Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^n)$; the following propositions are equivalent

(1) $\mu \leq_c \nu$,

(2) there exists a martingale (X_0, X_1) such that $X_0 \sim \mu$ and $X_1 \sim \nu$.

The implication (2) $\Rightarrow$ (1) comes from Jensen inequality.

A proof of Strassen theorem using barycentric L_1 cost

Let $\|\cdot\|$ be some norm on $\mathbb{R}^n$; consider

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

A proof of Strassen theorem using barycentric L_1 cost

Let $\| \cdot \|$ be some norm on $\mathbb{R}^n$; consider

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

A proof of Strassen theorem using barycentric L_1 cost

Let $\|\cdot\|$ be some norm on $\mathbb{R}^n$; consider

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

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.

A proof of Strassen theorem using barycentric L_1 cost

Let $\|\cdot\|$ be some norm on $\mathbb{R}^n$; consider

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

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 proof of Strassen theorem using barycentric L_1 cost

Let $\|\cdot\|$ be some norm on $\mathbb{R}^n$; consider

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

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.

Thus, if $\mu \leq_c \nu$, then

$$\overline{\mathcal{T}}_1(\nu|\mu) = \sup_{\varphi} \left\{ \int \varphi \, d\mu - \int \varphi \, d\nu \right\} = 0$$

and so there exists a martingale (X_0, X_1) with marginals μ and ν.

II - WOT and contraction properties of the Brenier map

Quadratic barycentric cost

$E = \mathbb{R}^n$ equipped with the Euclidean norm $|\cdot|$.

Consider

$$\overline{\mathcal{T}}_2(\nu|\mu) = \inf_{p \in \mathcal{P}(\mu,\nu)} \int \left| x - \int y \, dp_x(y) \right|^2 d\mu(x)$$
$$= \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[|X - \mathbb{E}[Y|X]|^2],$$

the weak transport cost associated to the cost function

$$c(x,p) = \left| x - \int y \, dp(y) \right|^2.$$

By Jensen,

$$\overline{\mathcal{T}}_2(\nu|\mu) \leq W_2^2(\nu,\mu).$$

Goal : characterize optimal transport plan for $\overline{\mathcal{T}}_2$.

Remark : if $\mu \leq_c \nu$, then $\overline{\mathcal{T}}_2(\nu|\mu) = 0$ and any martingale coupling between μ and ν is optimal. What about the general case ?

Brenier Theorem

The following result characterizes optimal transport plans for the cost function

$$\omega(x, y) = |y - x|^2, \qquad x, y \in \mathbb{R}^n.$$

Theorem (Brenier (1991))

If μ is absolutely continuous w.r.t. Lebesgue and if $\int |x|^2 d\mu(x) < +\infty$ and $\int |y|^2 d\nu(y) < +\infty$, then there exists a unique optimal transport plan π°, such that

$$W_2^2(\nu, \mu) = \iint |y - x|^2 \, d\pi^\circ(x, y).$$

Moreover π° has the following structure : there exists a *convex* function $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ such that $\pi^\circ = \mathrm{Law}(X, \nabla\phi(X))$, with $X \sim \mu$ and so

$$W_2^2(\nu, \mu) = \int |\nabla\phi(x) - x|^2 \, d\mu(x).$$

Brenier-Strassen couplings

Elementary remark : it is always possible to compose a deterministic transport and a martingale transport to couple two arbitrary probability measures μ and ν.

Indeed if (X, Y) is an arbitrary coupling then letting $\bar{X} = \mathbb{E}[Y|X]$, the coupling $(X, \bar{X})$ is deterministic and $(\bar{X}, Y)$ is a martingale.

Definition

A coupling (X, Y) between $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^n)$ is of the Brenier-Strassen type if

$$\mathbb{E}[Y|X] = \nabla\phi(X) \qquad \text{a.s}$$

with $\phi : \mathbb{R}^n \to \mathbb{R}$ a convex function of class $\mathcal{C}^1$.

Remark : the independent coupling is of the Brenier-Strassen type.

Main Results

G.-Juillet (2020) / Alfonsi-Corbetta-Jourdain (2020)
Dimension 1 : G.-Roberto-Samson-Shu-Tetali (2018)

Let $\mathcal{P}_2(\mathbb{R}^n)$ denote the set of probability measures with a finite second moment.

Theorem 1 (Interpretation of $\overline{\mathcal{T}}_2$)

Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$; define $B_\nu = \{\eta \in \mathcal{P}_1(\mathbb{R}^n) : \eta \leq_c \nu\}$.
There exists a unique probability measure $\bar{\mu} \in B_\nu$ such that

$$W_2(\bar{\mu}, \mu) = \inf_{\eta \in B_\nu} W_2(\eta, \mu).$$

Moreover

$$\overline{\mathcal{T}}_2(\nu|\mu) = W_2^2(\bar{\mu}, \mu).$$

Remark : It is also possible to define the projection $\bar{\nu}$ of ν onto $\{\eta : \mu \leq_c \eta\}$.
See Alfonsi - Corbetta - Jourdain (2020) and more recently Kim - Ruan (2021)

Main Results

G.-Juillet (2020) / Backhoff-Veraguas - Beiglboeck - Pammer (2019)

Theorem 2 (Optimal plans for $\overline{\mathcal{T}}_2$)

Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$;

(1) There exists a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ such that

$$\bar{\mu} = \nabla\phi_{\#}\mu.$$

Moreover $\nabla\phi$ is 1-Lipschitz.

(2) A coupling (X, Y) between μ and ν is optimal for $\overline{\mathcal{T}}_2(\nu|\mu)$ if and only if $\mathbb{E}[Y|X] = \nabla\phi(X)$ a.s.

Remark

Optimal transport between μ and its projection $\bar{\mu}$ is thus more regular than in the generic case : it is automatically given by a Lipschitz continuous transport map without any particular assumption on μ.

Examples

Theorem

If $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ and $\nu = \sum_{i=0}^k p_i \delta_{y_i}$ with $p_i \geq 0$ and $y_0, \ldots, y_k$ affinely independent points of $\mathbb{R}^n$, then there exists some $c \in \mathbb{R}^n$ such that

$$\bar{\mu} = T_{\#}\mu, \qquad \text{with} \qquad T(x) = \mathrm{Proj}_\Delta(x + c),$$

where Δ is the convex hull of $\{y_0, \ldots, y_k\}$ and Proj_Δ denotes the orthogonal projection on Δ.

Other example : In dimension 1, Alfonsi-Corbetta-Jourdain (2020) and Backhoff-Veraguas - Beiglboeck - Pammer (2020) obtained a semi-explicit formula for the transport map T sending μ on $\bar{\mu}$.

Characterization of the contractivity of the Brenier map

The following result is a consequence of our main results :

Corollary 1 (G.-Juillet (2020), Fathi-G.-Prod'Homme (2020))

Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^n)$; the following propositions are equivalent

(1) There exists $\phi : \mathbb{R}^n \to \mathbb{R}$ convex and $\mathcal{C}^1$ such that $\nu = \nabla\phi_{\#}\mu$ with $\nabla\phi$ 1-Lipschitz ;

(2) $\bar{\mu} = \nu$;

(3) $W_2^2(\nu, \mu) = \overline{\mathcal{T}}_2(\nu|\mu)$.

Caffarelli contraction theorem

Theorem (Caffarelli (2000))

If γ is the standard Gaussian measure on $\mathbb{R}^n$ and $d\nu(y) = e^{-V(y)}\,dy$ is a probability measure associated to a $\mathcal{C}^2$ smooth function V on $\mathbb{R}^n$ such that $\mathrm{Hess}\, V \geq \mathrm{Id}$, then there exists a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^1$ such that $\nu = \nabla\phi_{\#}\gamma$ and such that $\nabla\phi$ is 1-Lipschitz.

In other words, the Brenier map from γ to ν is a contraction.

Original proof based on the Monge-Ampère equation satisfied by ϕ.

Generalizations by Kolesnikov ('10), Kim-Milman ('12), Colombo-Figalli-Jhaveri ('17).

Equivalent formulation of Caffarelli theorem

Corollary (Equivalent formulation of Caffarelli theorem)

If γ is the standard gaussian measure on $\mathbb{R}^n$ and $d\nu(y) = e^{-V(y)}\,dy$, with $\mathrm{Hess}\,V \geq \mathrm{Id}$, then

$$\bar{\gamma} = \nu.$$

In a joint paper with M. Fathi and M. Prod'Homme (2020), we obtained a new proof of Caffarelli theorem by directly showing that if $d\nu(y) = e^{-V(y)}\,dy$, with $\mathrm{Hess}\,V \geq \mathrm{Id}$, then

$$W_2^2(\nu, \gamma) \leq W_2^2(\eta, \gamma), \qquad \forall \eta \leq_c \nu.$$

Our proof relies on entropic regularization and Prekopa-Leindler inequality.

Caffarelli contraction theorem - Example of application

The standard Gaussian measure has the following dimension free concentration property :

Theorem

If $X_1, \ldots, X_k$ are independent standard Gaussian random vectors on $\mathbb{R}^n$, then for any function $f : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is L-Lipschitz with respect to the Euclidean norm on $(\mathbb{R}^n)^k$, it holds

$$\mathbb{P}(|f(X_1, \ldots, X_k) - \mathbb{E}[f(X_1, \ldots, X_k)]| \geq t) \leq 2e^{-t^2/(2L^2)}, \qquad \forall t \geq 0.$$

Caffarelli contraction theorem - Example of application

The standard Gaussian measure has the following dimension free concentration property :

Theorem

If $X_1, \ldots, X_k$ are independent standard Gaussian random vectors on $\mathbb{R}^n$, then for any function $f : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is L-Lipschitz with respect to the Euclidean norm on $(\mathbb{R}^n)^k$, it holds

$$\mathbb{P}(|f(X_1, \ldots, X_k) - \mathbb{E}[f(X_1, \ldots, X_k)]| \geq t) \leq 2e^{-t^2/(2L^2)}, \qquad \forall t \geq 0.$$

Corollary

If $Y_1, \ldots, Y_k$ are i.i.d random vectors on $\mathbb{R}^n$ distributed according to a probability ν satisfying the assumptions of Caffarelli theorem, then for any function $g : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is L-Lipschitz with respect to the Euclidean norm, it holds

$$\mathbb{P}(|g(Y_1, \ldots, Y_k) - \mathbb{E}[g(Y_1, \ldots, Y_k)]| \geq t) \leq 2e^{-t^2/(2L^2)}, \qquad \forall t \geq 0.$$

Caffarelli contraction theorem - Example of application

The standard Gaussian measure has the following dimension free concentration property :

Theorem

If $X_1, \ldots, X_k$ are independent standard Gaussian random vectors on $\mathbb{R}^n$, then for any function $f : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is L-Lipschitz with respect to the Euclidean norm on $(\mathbb{R}^n)^k$, it holds

$$\mathbb{P}(|f(X_1, \ldots, X_k) - \mathbb{E}[f(X_1, \ldots, X_k)]| \geq t) \leq 2e^{-t^2/(2L^2)}, \qquad \forall t \geq 0.$$

Corollary

If $Y_1, \ldots, Y_k$ are i.i.d random vectors on $\mathbb{R}^n$ distributed according to a probability ν satisfying the assumptions of Caffarelli theorem, then for any function $g : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ which is L-Lipschitz with respect to the Euclidean norm, it holds

$$\mathbb{P}(|g(Y_1, \ldots, Y_k) - \mathbb{E}[g(Y_1, \ldots, Y_k)]| \geq t) \leq 2e^{-t^2/(2L^2)}, \qquad \forall t \geq 0.$$

Proof.

Apply the Gaussian concentration inequality to $f(x_1, \ldots, x_k) = g(\nabla\phi(x_1), \ldots, \nabla\phi(x_k))$ where $\nabla\phi$ is the Brenier map between γ and ν. $\square$

Caffarelli contraction theorem

Numerous consequences in the field of functional inequalities.

Example : the standard Gaussian measure γ satisfies the log-Sobolev inequality (Gross (1975)) :

$$\text{(LSI)} \qquad \mathrm{Ent}_\gamma(f^2) \leq 2 \int |\nabla f|^2 \, d\gamma, \qquad \forall f : \mathbb{R}^n \to \mathbb{R} \ \mathcal{C}^1$$

If $d\nu(y) = e^{-V(y)}\, dy$ with $\mathrm{Hess}\, V \geq \mathrm{Id}$, then according to Caffarelli Theorem $\nu = \nabla\phi_{\#}\gamma$ with $\nabla\phi$ 1-Lispchitz.

Therefore, applying (LSI) to $f = g \circ \nabla\phi$ yields to

$$\begin{aligned} \mathrm{Ent}_\nu(g^2) &\leq 2 \int |\mathrm{Hess}\, \phi(x) \cdot \nabla g(\nabla\phi(x))|^2 \, d\gamma(x), \qquad \forall f : \mathbb{R}^n \to \mathbb{R} \ \mathcal{C}^1 \\ &\leq 2 \int |\nabla g(y)|^2 \, d\nu(y). \end{aligned}$$

So ν satisfies (LSI) : one recovers the Bakry-Emery criterion (with the good constant)

III - WOT and concentration of measure

Concentration of measure

Definition

One says that $\mu \in \mathcal{P}(\mathbb{R}^n)$ satisfies the dimension free gaussian concentration property if there exist $a, b > 0$ such that for all $k \geq 1$ and for all function $f : (\mathbb{R}^n)^k \to \mathbb{R}$ 1-Lipschitz (w.r.t to Euclidean norm), it holds

$$\mathbb{P}(f(X_1, \ldots, X_k) \geq m(f) + t) \leq e^{-b(t-a)^2}, \qquad \forall t \geq a,$$

where $X_1, \ldots, X_k$ are i.i.d of law μ and $m(f)$ is the median of $f(X_1, \ldots, X_k)$.

Examples :

- This property is satisfied by the standard Gaussian measure γ on $\mathbb{R}^n$, with constants $b = 1/2$ and $a = 0$.
  This result goes back to the Borell-Sudakov-Tsirelson isoperimetric theorem for the Gauss space.
- More generally, if $d\mu = e^{-V}\,dx$ with $\mathrm{Hess}\, V \geq c\mathrm{Id}$, with $c > 0$, then μ satisfies the dimension free concentration property with constant $b = c/2$ and $a = 0$.
  This is for instance a consequence of the Caffarelli contraction theorem.
- Many methods were proposed to show this type of concentration inequalities for more general probability measures : logarithmic Sobolev inequality, transport entropy inequality, Brunn-Minkowski inequality, ...

Convex concentration of measure

Theorem

Let $\mu \in \mathcal{P}(\mathbb{R}^n)$ and $b > 0$; the following are equivalent

(1) there exists $a \geq 0$ such that μ satisfies the dimension free gaussian concentration property with constants a, b,

(2) μ satisfies the $\mathbb{T}_2$ transport-entropy inequality

$$W_2^2(\nu, \mu) \leq \frac{1}{b} H(\nu|\mu), \qquad \forall \nu \in \mathcal{P}(\mathbb{R}^n).$$

The implication (2) $\Rightarrow$ (1) is due to Marton and Talagrand ($a = \sqrt{(\log 2)/b}$)
The implication (1) $\Rightarrow$ (2) (G. 09) relies on Large deviation theory (Sanov Theorem).

A probability measure satisfying the $\mathbb{T}_2$ transport-entropy inequality has necessarily a connected support. This excludes in particular discrete mesures...

Convex concentration of measure

Definition

One says that $\mu \in \mathcal{P}(\mathbb{R}^n)$ satisfies the dimension free gaussian *convex* concentration property if there exist $b, c > 0$ such that for all $k \geq 1$ and for all function $f : (\mathbb{R}^n)^k \to \mathbb{R}$ 1-Lipschitz convex or concave, it holds

$$\mathbb{P}(f(X_1, \ldots, X_k) \geq m(f) + t) \leq ce^{-bt^2}, \qquad \forall t \geq 0,$$

where $X_1, \ldots, X_k$ are i.i.d of law μ and $m(f)$ is the median of $f(X_1, \ldots, X_k)$.

Remarks :
- Weaker than usual gaussian dimension free concentration property.
- Example (Talagrand, Marton, Maurey) : If μ has a bounded support, then it satisfies this inequality with $c = 2$ and $b = \frac{1}{4\mathrm{Diam}(\mathrm{Supp}(\mu))^2}$.

Convex concentration of measure

Theorem (G.-Roberto-Samson-Tetali (2017))

A probability measure $\mu \in \mathcal{P}(\mathbb{R}^n)$ satisfies the dimension free gaussian *convex* concentration property if and only if there exists $D > 0$ such that

$$\overline{\mathcal{T}}_2(\nu_1|\nu_2) \leq D(H(\nu_1|\mu) + H(\nu_2|\mu)), \qquad \forall \nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^n).$$

+ Quantitative links between constants b, c, D
+ Necessary and sufficient condition in dimension 1 (paper with Y. Shu).

IV - One word on WOT with unnormalized kernels

WOT with unnormalized kernels

Work in progress with P. Choné and F. Kramarz

Let $\mathcal{X}, \mathcal{Y}$ be two compact metric spaces. Denote by $\mathcal{M}(\mathcal{Y})$ the set of all non-negative finite measures on $\mathcal{Y}$.

Definition

Let $c : \mathcal{X} \times \mathcal{M}(\mathcal{Y}) \to \mathbb{R}^+ \cup \{+\infty\}$; the unormalized weak transport cost $\mathcal{I}_c(\mu, \nu)$ between $\mu \in \mathcal{P}(\mathcal{X})$ and $\nu \in \mathcal{P}(\mathcal{Y})$ is defined by

$$\mathcal{I}_c(\mu, \nu) = \inf_{q \in \mathcal{Q}(\mu,\nu)} \int c(x, q^x)\, d\mu(x),$$

where $\mathcal{Q}(\mu, \nu)$ is the set of all non-negative kernels q (i.e $q^x(dy) \in \mathcal{M}(\mathcal{Y})$ for all $x \in \mathcal{X}$) such that $\mu q = \nu$

Economic motivation (Choné - Kramarz 2021) :

- μ represents a distribution of firms (the size of the firms is unknown)
- ν represents a distribution of workers
- q^x represents the workers recruited by the firm x : $q^x(dy) = \sum_{i=1}^k n_i \delta_{y_i}$ means that the firm x has recruited n_i workers with the skill profile y_i.
- $-c(x, m)$ represents the productivity of the firm x when it recruits a distribution of workers m.

Goal : Find the optimal allocation of workers to optimize the total productivity.

ArXiv preprint (soon available) : existence of solutions, duality, study of barycentric costs, generalization of Strassen theorem,. . .

Thank you for your attention !