Day 5 Talk 2

Radosław Adamczak

"On the convex Poincaré inequality and weak transportation inequality"

joint w Michał Strzelecki (Ph.D student) Warsaw

Dimension-free concentration

μ on (X,d) satisfies dimension-free concentration ineq (Cl_2^∞) iff: $\forall_{\forall N} \forall_{f: X^N \to \mathbb{R}}$ $\mu^{\otimes}(|f - \mathrm{Med}_{\mu}\, f| \geq t) \leq \alpha(t)$

for fixed $\alpha : [0,\infty) \to [0,1]$

$\lim_{t\to\infty} \alpha(t) = 0$

Talagrand $\rightsquigarrow$ always $\alpha(t) \leq e^{-ct}$ suffices

Cl_2^∞ with

Thm Gromov + V. Milman '83 Poincaré $\Rightarrow$ $\alpha(t) = 2e^{-ct}$

actually $\Leftrightarrow$!

Recall: $\mathrm{Ent}\, X = \mathbb{E}(X \log X) - \mathbb{E} X \log \mathbb{E} X$

Equivalent conditions, classical

CI_c^∞ $\quad \mu^{\otimes}(|f - \mathrm{Med}_{\mu^{\otimes N}} f| > t) \le \alpha(t)$ ~~$\forall N$ $\forall$ 1-Lip f~~

$\Updownarrow$

Poincaré $\quad \lambda \mathrm{Var}_\mu(f) \le \mathbb{E}_\mu |\nabla f|^2$

$\Updownarrow$

mod-log Sobolev $\quad |\nabla f| \le c \Rightarrow \mathrm{Ent}_\mu e^f \le C \mathbb{E}_\mu |\nabla f|^2 e^f$

$\Updownarrow$

transportation ineq $\quad T_\Theta \quad : \forall \nu \quad T_\Theta(\nu, \mu) \le H(\nu|\mu) = \mathbb{E}_\nu \log \frac{d\nu}{d\mu}$

(Θ ← quadratic-linear function; H ← relative entropy)

$$T_\Theta = \inf_\pi \iint \Theta(x-y)\, \pi(dx\,dy)$$

$$= \inf_{\substack{x\sim\mu \\ y\sim\nu}} \mathbb{E}(\Theta(x-y))$$

(couplings)

Restrict concentration to convex functions f. (Conv CI_2^∞)

Annoying complication: must deal with upper/lower tails separately

- Versions of Poincaré, log-Sobolev, T_Θ for convex?
- for convex conditions see slides

⇝ same picture as above for convex conditions

$\text{conv } CI_c^\infty$ $\mu^{\otimes N}(|f - \mathrm{Med}_{\mu^{\otimes N}} f| \geq t) \leq \alpha(t)$
$\forall N$ $\forall$ 1-Lip
<u>convex</u> f

$\Updownarrow$

convex Poincaré

$\lambda \, \mathrm{Var}_\mu f \leq \mathbb{E}_\mu |\nabla f|^2$ $\forall f$ convex

$\Updownarrow$

subexp.-$\text{conv } CI_c^\infty$ $\alpha(t)$ subexponential

$\Updownarrow$

modified log-Sobolev ineq for convex & concave

$\Updownarrow$

$\bar{T}_\theta$: $\bar{T}_\theta(\nu|\mu), \bar{T}_\theta(\mu|\nu) \leq H(\nu|\mu)$ $\forall \nu$

$$\bar{T}_\theta(\nu|\mu) = \inf_\pi \int \theta\left(x - \int y p_x(dy)\right) d\mu(x)$$

$\mu = \underbrace{p_x \, dy}_{\text{cond. exp}} \, d\mu$

$$= \inf_{\substack{X\sim\mu \\ Y\sim\nu}} \mathbb{E}\left(\theta\left(X - \mathbb{E}(Y|X)\right)\right)$$

Note: bounds not dimension free

Q: μ Prob meas on $\mathbb{R}^n$ satisfying convex-Poincaré with constant λ. Does there exist a constant $c(\lambda)$ ... see slides

$\alpha(t) < e^{-c(\lambda) t}$ ↑ dim indep.

On the convex Poincaré inequality and weak transportation inequalities

Radosław Adamczak
(joint work with Michał Strzelecki)

Polish Academy of Sciences & University of Warsaw

Geometric functional analysis and applications
Berkeley 2017

Dimension-free concentration

Definition

We will say that a probability measure μ on $(\mathcal{X}, d)$ satisfies a dimension-free concentration inequality (CI_2^∞) iff there exists a function $\alpha\colon [0,\infty) \to [0,1]$, $\lim_{t\to\infty} \alpha(t) = 0$ such that for all N, all 1-Lipschitz functions $f\colon \mathcal{X} \to \mathbb{R}$ and $t > 0$,

$$\mu^{\otimes N}\Big(|f - \mathrm{Med}_{\mu^{\otimes N}} f| \geq t\Big) \leq \alpha(t)$$

(the distance on $\mathcal{X}^N$ is $d(x,y) = \sqrt{\sum_{i=1}^N d(x_i,y_i)^2}$).

- Standard examples: uniform measure on the sphere, Gaussian measure
- By CLT, in non-trivial cases, α cannot decay faster than some Gaussian tail.
- As observed by Talagrand, if μ satisfies CI_2^∞ then one can take $\alpha(t) = 2\exp(-ct)$ for some $c > 0$.

Ways to prove CI_2^∞

- Functional inequalities: Poincaré, modified log-Sobolev
- Transportation cost inequalities
- Infimum convolution inequalities – dual to transportation

Common idea: Tensorization - the ineq. passes from μ to $\mu^{\otimes N}$.

Definition

We will say that μ satisfies the Poincaré inequality iff for some $\lambda > 0$ and all locally Lipschitz functions $f\colon \mathcal{X} \to \mathbb{R}$

$$\lambda \mathrm{Var}_\mu f \leq \mathbb{E}_\mu |\nabla f|^2$$

with the length of gradient

$$|\nabla f|(x) = \limsup_{y \to x} \frac{|f(y) - f(x)|}{d(x,y)}$$

(we set $|\nabla f|(x) = 0$ for isolated points).

Theorem (Gromov-V. Milman '83)

If μ satisfies the Poincaré inequality then it satisfies CI_2^∞ with $\alpha(t) = 2\exp(-ct)$

The standard exponential distribution satisfies Poincaré so this is optimal for large t.

Theorem (Gromov-V. Milman '83)

If μ satisfies the Poincaré inequality then it satisfies CI_2^∞ with $\alpha(t) = 2\exp(-ct)$

The standard exponential distribution satisfies Poincaré so this is optimal for large t.

Theorem (Gozlan-Roberto-Samson '15)

If μ satisfies CI_2^∞ then it satisfies the Poincaré inequality.

As a consequence

$$CI_2^\infty \iff \text{subexp.-}CI_2^\infty \iff \text{Poincaré inequality}$$

Theorem (Bobkov-Ledoux '97, Bobkov-Gentil-Ledoux '01)

If μ on $\mathbb{R}^n$ satisfies the Poincaré inequality then

- *there exist c, C s.t. for every locally Lipschitz function f with $|\nabla f| \leq c$,*
$$\mathrm{Ent}_\mu e^f \leq C\mathbb{E}_\mu |\nabla f|^2 e^f,$$
- *for some C, D, the measure μ satisfies the transportation cost inequality with the quadratic-linear cost*
$$\theta(x) = \begin{cases} \frac{|x|^2}{2C} & \text{for } |x| \leq CD, \\ D|x| - \frac{CD^2}{2} & \text{for } |x| > CD, \end{cases}$$
i.e. for all measures ν on $\mathbb{R}^n$,

$$T_\theta(\nu, \mu) := \inf_\Pi \iint \theta(x-y)\Pi(dx, dy) \leq H(\nu|\mu) := \mathbb{E}_\nu \log(\frac{d\nu}{d\mu}),$$

where the infimum is taken over all couplings Π of μ and ν.

Why do we care about modified log-Sobolev or transportation cost inequalities?

Improved concentration

For $f : (\mathbb{R}^n)^N \to \mathbb{R}$*,*

$$\mu^{\otimes N}\Big(|f - \mathbb{E}_{\mu^{\otimes N}} f| \geq t\Big) \leq 2\exp\Big(-c\min\Big(\frac{t^2}{L_2^2}, \frac{t}{L_1}\Big)\Big)$$

where

$$L_2 = \sup_{x\in\mathbb{R}^{nN}} |\nabla f(x)|, \ L_1 = \max_{i=1,\ldots,N} \sup_{x\in\mathbb{R}^{Nn}} |\nabla_i f(x)|.$$

Why do we care about modified log-Sobolev or transportation cost inequalities?

Improved concentration

For $f\colon (\mathbb{R}^n)^N \to \mathbb{R}$,

$$\mu^{\otimes N}\Big(|f - \mathbb{E}_{\mu^{\otimes N}} f| \geq t\Big) \leq 2\exp\Big(-c\min\Big(\frac{t^2}{L_2^2}, \frac{t}{L_1}\Big)\Big)$$

where

$$L_2 = \sup_{x\in\mathbb{R}^{nN}} |\nabla f(x)|,\ L_1 = \max_{i=1,\ldots,N} \sup_{x\in\mathbb{R}^{Nn}} |\nabla_i f(x)|.$$

Simple example: $n = 1$, $f(x) = (x_1 + \ldots + x_N)/\sqrt{N}$,

- Poincaré:

$$\mu^{\otimes N}(|f - \mathbb{E}f| \geq t) \leq 2\exp(-ct)$$

- modified log-Sobolev:

$$\mu^{\otimes N}(|f - \mathbb{E}f| \geq t) \leq 2\exp\Big(-c\min(t^2, \sqrt{N}t)\Big).$$

dimension-free concentration CI_2^∞

$\Updownarrow$

subexp.-CI_2^∞

$\Updownarrow$

Poincaré inequality

$\Updownarrow$

modified log-Sobolev inequality

$\Updownarrow$

quadratic-linear transportation cost ineq.

$\Updownarrow$

two-level concentration

Getting to the topic of the talk...

Definition

We will say that a probability measure μ on $\mathbb{R}^n$ satisfies **convex** *dimension-free concentration inequality (convCI$_2^\infty$) iff there exists a function $\alpha\colon [0,\infty) \to [0,1]$, $\lim_{t\to\infty} \alpha(t) = 0$ such that for all N, all 1-Lipschitz* **convex** *functions $f\colon \mathcal{X} \to \mathbb{R}$ and $t > 0$,*

$$\mu^{\otimes N}\Big(|f - \mathrm{Med}_{\mu^{\otimes N}} f| \geq t\Big) \leq \alpha(t).$$

Theorem (Talagrand '94)

All measures with bounded support satisfy convCI$_2^\infty$

Questions:

- Do we also have improved concentration?
- Can we get a picture as in the "classical" case?

Why may this be interesting?

- Restricting concentration to convex functions allows for significant weakening of assumptions, while still encompassing many important functions (e.g. norms)
- Relation with concentration for polynomials (Marton, Meckes-Szarek, Vu-Wang, A.)
- New arguments or modifications of existing ones needed, since convexity is not preserved under basic operations
- Investigating convex functions sometimes gives new insight into the classical theory (Gozlan-Roberto-Samson, Gozlan-Roberto-Samson-Shu-Tetali, Shu-Strzelecki).

Questions:

- Do we also have improved concentration?
- Can we get a picture as in the "classical" case?

An annoying complication:
We have to deal with upper and lower tails separately.

Theorem (Bobkov-Götze '99, Gozlan-Roberto-Samson '15)

A measure μ on $\mathbb{R}^n$ satisfies convCI$_2^\infty$ iff it satisfies the Poincaré inequality for all convex functions.

Theorem (Gozlan-Roberto-Samson '15)

Dimension free convex concentration from above

$$\mu^{\otimes N}(f \geq \mathrm{Med}_{\mu^{\otimes N}} f + t) \leq \alpha(t)$$

implies subexponential convex concentration from above and implies the convex Poincaré inequality.

Weak transportation inequalities

Definition (Gozlan-Roberto-Samson-Tetali '14)

For a convex cost function $\theta\colon \mathbb{R}^n \to \mathbb{R}$ and two probability measures μ, ν on $\mathbb{R}^n$ with finite first moments define the weak transportation cost $\overline{T}(\nu|\mu)$ as

$$\overline{T}(\nu|\mu) = \inf_{\Pi} \int \theta(x - \int y p_x(dy)) d\mu(x),$$

where the infimum is taken over all couplings Π of $\mu$and ν and p_x is the conditional distribution given by

$$\Pi(dxdy) = p_x(dy)\mu(dx).$$

In probabilistic notation

$$\overline{T}(\nu|\mu) = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\theta(X - \mathbb{E}(Y|X)).$$

Notation:

$\mathcal{P}_1(\mathbb{R}^n)$ – set of probability measures on $\mathbb{R}^n$ with finite first moment

Definition (Gozlan-Roberto-Samson-Tetali '14)

We will say that $\mu \in \mathcal{P}_1(\mathbb{R}^n)$ *satisfies*

- $\overline{T}_\theta^+$ *if for all* $\nu \in \mathcal{P}_1(\mathbb{R}^n)$,

$$\overline{T}_\theta(\nu|\mu) \leq H(\nu|\mu)$$

- $\overline{T}_\theta^-$ *if for all* $\nu \in \mathcal{P}_1(\mathbb{R}^n)$,

$$\overline{T}_\theta(\mu|\nu) \leq H(\nu|\mu)$$

- $\overline{T}_\theta$ *if it satisfies both* $\overline{T}_\theta^+$ *and* $\overline{T}_\theta^-$.

For measures on **the real line** by combining results by Bobkov-Götze, Gozlan-Roberto-Samson, Gozlan-Roberto-Samson-Tetali, Feldheim-Marsiglietti-Nayar-Wang, Strzelecki-A., Gozlan-Roberto-Samson-Shu-Tetali one obtains

dimension-free $convCI_2^\infty$

$\Updownarrow$

subexp.-$convCI_2^\infty$

$\Updownarrow$

convex Poincaré inequality

$\Updownarrow$

modified log-Sobolev ineq. for convex and concave functions

$\Updownarrow$

$\overline{T}_\theta$ with θ quadratic-linear

$\Updownarrow$

Two-level concentration for convex functions

Remark: Many of the proofs rather indirect, based on the characterization of the convex Poincaré inequality on the line due to Bobkov-Götze dimension-free $convCI_2^\infty$

$\Updownarrow$

subexp.-$convCI_2^\infty$

$\Updownarrow$

convex Poincaré inequality

$\Updownarrow$

modified log-Sobolev ineq. for convex and concave functions

$\Updownarrow$

$\overline{T}_\theta$ with θ quadratic-linear

$\Updownarrow$

Two-level concentration for convex functions

Theorem (Strzelecki-A. '17)

The above picture holds for measures on $\mathbb{R}^n$.

More specifically we proved that the convex Poincaré inequality implies modified log-Sobolev inequalities for convex and concave functions, which in turn imply $\overline{T}_\theta$

Theorem (Strzelecki-A. '17)

Let μ be a probability measure on $\mathbb{R}^n$ which satisfying the convex Poincaré inequality

$$\lambda \mathrm{Var} f \leq \mathbb{E}_\mu |\nabla f|^2$$

for all convex functions $f \colon \mathbb{R}^n \to \mathbb{R}$. Then for some c, C, D,

- *μ satisfies the modified log-Sobolev inequality*

  $$\mathrm{Ent}_\mu e^f \leq C \mathbb{E} |\nabla f|^2 e^f$$

  for all convex or concave functions f with $|\nabla f| \leq c$,
- *μ satisfies the weak transportation inequality $\overline{T}_\theta$:*

  $$\overline{T}_\theta(\nu|\mu), \overline{T}_\theta(\mu|\nu) \leq H(\nu|\mu)$$

  where

  $$\theta(x) = \begin{cases} \frac{|x|^2}{2C} & \text{for } |x| \leq CD, \\ D|x| - \frac{CD^2}{2} & \text{for } |x| > CD. \end{cases}$$

However...

- For concave functions in the modified log-Sobolev inequality and for the inequality $\overline{T}_\theta^+$ the constants we get depend not only on λ from the convex Poincaré inequality, but also on some quantiles of the measure μ, which may be dimension dependent.
- This is not the case for convex functions and the inequality $\overline{T}_\theta^-$.
- One can remove the dependence on quantiles if the following question has affirmative answer

However...

- For concave functions in the modified log-Sobolev inequality and for the inequality $\overline{T}_\theta^+$ the constants we get depend not only on λ from the convex Poincaré inequality, but also on some quantiles of the measure μ, which may be dimension dependent.
- This is not the case for convex functions and the inequality $\overline{T}_\theta^-$.
- One can remove the dependence on quantiles if the following question has affirmative answer

Question

Let μ be a probability measure on $\mathbb{R}^n$ satisfying the convex Poincaré inequality with constant λ. Does there exist a constant $c(\lambda)$ such that for all convex functions $f\colon \mathbb{R}^n \to \mathbb{R}$ and all $t > 0$,

$$\mu(f \le \mathbb{E}_\mu f - t) \le 2\exp(-c(\lambda)t)?$$

A few words about the proof

- To prove modified log-Sobolev inequalities we modify the argument by Bobkov-Ledoux from the classical case.
- WLOG we can assume that $\mathrm{Med}\, f = 0$.

Lemma

If μ satisfies the convex Poincaré ineq. and f is convex, then

$$\mathbb{E}_\mu (f - \mathrm{Med}\, f)^2 \leq \frac{2}{\lambda} \mathbb{E}_\mu |\nabla f|^2.$$

Lemma

Assume that f is convex, $\mathrm{Med}_\mu f = 0$, $|\nabla f| \leq c(\lambda)$. Then

$$\mathbb{E}_\mu f^2 e^f \leq C(\lambda) \mathbb{E} |\nabla f|^2 e^f,$$
$$\mathbb{E}_\mu f^2 \leq C(\lambda) \mathbb{E} |\nabla f|^2 e^f.$$

Lemma

Assume that f is convex, $\operatorname{Med}_\mu f = 0$, $|\nabla f| \le c(\lambda)$. *Then*

$$\mathbb{E}_\mu f^2 e^f \le C(\lambda)\mathbb{E}|\nabla f|^2 e^f,$$
$$\mathbb{E}_\mu f^2 \le C(\lambda)\mathbb{E}|\nabla f|^2 e^f.$$

Denote $F(t) = \mathbb{E} f^2 e^{tf}$.

$$\begin{aligned}\operatorname{Ent} e^f &\le \mathbb{E}(fe^f - e^f + 1)\\ &= \mathbb{E}\int_0^1 tf^2 e^{tf} dt = \int_0^1 tF(t)dt\\ &\le \int_0^1 t(1-t)F(0) + t^2 F(1) dt = \frac{1}{6}F(0) + \frac{1}{3}F(1).\end{aligned}$$

We use the lemma to estimate the right-hand side.

For concave functions one proves

Lemma

Assume that f is concave, $\mathrm{Med}_\mu f = 0$, $|\nabla f| \le c$. *Then*

$$\mathbb{E} f^2 e^f \mathbf{1}_{\{f \ge 0\}} \le C(\lambda, \mu) \mathbb{E} |\nabla f|^2 e^f$$
$$\mathbb{E} f^2 \le C(\lambda) \mathbb{E} |\nabla f|^2 e^f.$$

As before, for $F(t) = \mathbb{E} f^2 e^{tf}$ we have

$$\mathrm{Ent}\, e^f \le \frac{1}{6} F(0) + \frac{1}{3} F(1).$$

Moreover

$$F(1) \le F(0) + \mathbb{E} f^2 e^f \mathbf{1}_{\{f \ge 0\}},$$

so one can use the lemma.

To pass from log-Sobolev to transportation one uses the dual form of the latter.

Lemma (Gozlan-Roberto-Samson-Tetali '14)

$$Q_t f(x) := inf_{y \in \mathbb{R}^n}\Big\{f(y) + t\theta\Big(\frac{x-y}{t}\Big)\Big\}.$$

Then

(i) *μ satisfies $\overline{T}_\theta^+$ iff for all convex, Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$, bounded from below,*

$$\exp\Big(\int_{\mathbb{R}^n} Q_1 f d\mu\Big)\int_{\mathbb{R}^n} e^{-f} d\mu \leq 1;$$

(ii) *μ satisfies $\overline{T}_\theta^-$ iff for all convex, Lipschitz $f : \mathbb{R}^n \to \mathbb{R}$, bounded from below,*

$$\int_{\mathbb{R}^n} \exp(Q_1 f) d\mu \exp\Big(-\int_{\mathbb{R}^n} f d\mu\Big) \leq 1.$$

- For $\overline{T}_\theta^-$, following the ideas of Bobkov-Gentil-Ledoux one combines the Hamilton-Jacobi equation

$$\frac{d}{dt} Q_t f(x) + \theta^*(|\nabla_x Q_t f(x)|) = 0$$

with the modified log-Sobolev inequality in order to show that the function

$$F(t) = \frac{1}{t} \log \mathbb{E}_\mu e^{tQ_t f}$$

is non-increasing. Thus

$$\mathbb{E}_\mu e^{Q_1 f} = F(1) \leq \liminf_{t \to 0} F(t) \leq \mathbb{E}_\mu f,$$

which proves the dual form of $\overline{T}^-$.

- For $\overline{T}^+$ – a similar argument

Final remarks

- Weak transportation inequalities $\overline{T}_\theta$ on the line have been characterized by Gozlan-Roberto-Samson-Shu-Tetali. It turns out that μ satisfies the usual strong transportation inequality iff it satisfies the weak one and the Poincaré inequality (for all locally Lipschitz functions)
- Similarly, Shu-Strzelecki showed that the modified log-Sobolev inequality for convex functions on the line is in fact equivalent to $\overline{T}_\theta$. In particular, one has the corollary for a large class of cost functions:

Corollary (Shu-Strzelecki '16)

A probability measure on the line satisfies the strong transportation inequality T_θ iff it satisfies the Poincaré inequality for all functions and the modified log-Sobolev inequality for convex functions.

- **Question:** Does this hold in higher dimensions? Again, the proofs for the line go through explicit characterizations.

Concentration for convex Lipschitz functions can be extended to general convex function. Here is a special case.

Proposition (Strzelecki-A. '17)

Assume that a probability measure μ on $\mathbb{R}$ satisfies the convex Poincaré inequality. Then for any convex function $f\colon \mathbb{R}^N \to \mathbb{R}$ and any $p \geq 1$,

$$\left(\mathbb{E}_{\mu^{\otimes N}}\left|\frac{(f - \mathrm{Med}_{\mu^{\otimes N}} f)_+}{\sqrt{p}|\nabla f|_2 + p|\nabla f|_\infty}\right|^p\right)^{1/p} \leq C(\lambda)$$

and

$$\left(\mathbb{E}_{\mu^{\otimes N}}(f - \mathrm{Med}_{\mu^{\otimes N}} f)_-^p\right)^{1/p} \leq C(\lambda)(\sqrt{p}\mathbb{E}_{\mu^{\otimes N}}|\nabla f|_2 + p\mathbb{E}_{\mu^{\otimes N}}|\nabla_i f|_\infty).$$

In the Gaussian case interesting strengthened estimates for lower tails of convex functions were proved recently by Paouris-Valettas. For self-bounded empirical processes similar ineq. obtained by de la Peña-Klass-Lai. Non-Lipschitz convex functions were also considered by Bobkov-Nayar-Tetali.

Thank you