Modified log-Sobolev inequalities for convex functions

(based on joint work with Radosław Adamczak)

Michał Strzelecki

University of Warsaw

Warsaw, June 1, 2015

Introduction: concentration of measure phenomenon

Theorem

- $X \sim \mathrm{Unif}(S^{n-1})$,
- $f : \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz.

Then

$$\mathbb{P}(|f(X) - \mathrm{Med}\, f(X)| \geq t) \leq 2\exp(-(n-2)t^2/(2L^2)).$$

Introduction: concentration of measure phenomenon

Theorem

- $X \sim \mathrm{Unif}(S^{n-1})$,
- $f : \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz.

Then

$$\mathbb{P}(|f(X) - \mathrm{Med}\, f(X)| \geq t) \leq 2\exp(-(n-2)t^2/(2L^2)).$$

Theorem (Talagrand, 1995)

- $X_1, \ldots, X_n$ – independent r.v., $|X_i| \leq 1$,
- $X = (X_1, \ldots, X_n)$,
- $\varphi : \mathbb{R}^n \to \mathbb{R}$ is convex, L-Lipschitz.

Then

$$\mathbb{P}(|\varphi(X) - \mathrm{Med}\, \varphi(X)| \geq t) \leq 4\exp(-t^2/(16L^2)).$$

A class of (symmetric) probability measures on $\mathbb{R}$

Definition

$\beta \in [0, 1]$, $m > 0$, $\sigma \geq 0$. We say that $\mu \in \mathcal{M}_\beta(m, \sigma^{\beta+1})$ if

$$\int_{x \vee m}^{\infty} y^\beta \mu([y, \infty)) dy \leq \sigma^{\beta+1} \mu([x, \infty)). \tag{1}$$

Equivalently, for some $h > 0$, $\alpha < 1$ and all $x > m$,

$$\mu([x + h/x^\beta, \infty)) \leq \alpha \mu([x, \infty)). \tag{2}$$

A class of (symmetric) probability measures on $\mathbb{R}$

Definition

$\beta \in [0,1]$, $m > 0$, $\sigma \geq 0$. We say that $\mu \in \mathcal{M}_\beta(m, \sigma^{\beta+1})$ if

$$\int_{x \vee m}^{\infty} y^\beta \mu([y,\infty))dy \leq \sigma^{\beta+1}\mu([x,\infty)). \tag{1}$$

Equivalently, for some $h > 0$, $\alpha < 1$ and all $x > m$,

$$\mu([x + h/x^\beta, \infty)) \leq \alpha\mu([x,\infty)). \tag{2}$$

„$\Downarrow$": For $x \geq m$,

$$\sigma^{\beta+1}\mu([x,\infty)) \geq \int_x^{x+\frac{2\sigma^{\beta+1}}{x^\beta}} y^\beta \mu([y,\infty))dy \geq \frac{2\sigma^{\beta+1}}{x^\beta} \cdot x^\beta \mu([x + \frac{2\sigma^{\beta+1}}{x^\beta}, \infty)),$$

so we can take $h = 2\sigma^{\beta+1}$ and $\alpha = 1/2$.

A class of (symmetric) probability measures on $\mathbb{R}$

Definition

$\beta \in [0,1]$, $m > 0$, $\sigma \geq 0$. We say that $\mu \in \mathcal{M}_\beta(m, \sigma^{\beta+1})$ if

$$\int_{x \vee m}^{\infty} y^\beta \mu([y,\infty))dy \leq \sigma^{\beta+1}\mu([x,\infty)). \tag{1}$$

Equivalently, for some $h > 0$, $\alpha < 1$ and all $x > m$,

$$\mu([x + h/x^\beta, \infty)) \leq \alpha\mu([x,\infty)). \tag{2}$$

„$\Uparrow$": For $x \geq m$ define $a_0 = x$, $a_{n+1} = a_n + h/a_n^\beta$. Then $a_n \nearrow \infty$ and

$$\int_x^\infty y^\beta \mu([y,\infty))dy \leq \sum_{n=0}^{\infty}(a_{n+1} - a_n)a_{n+1}^\beta \mu([a_n,\infty)) \leq K\sum_{n=0}^{\infty}\alpha^n\mu([a_0,\infty)).$$

Main result: modified log-Sobolev inequalities

Theorem (Adamczak, St., 2015)

- $X_1, \ldots, X_n$ – independent r.v. with laws in $\mathcal{M}_\beta(m, \sigma^{\beta+1})$, $\beta \in [0,1]$,
- $X = (X_1, \ldots, X_n)$,
- $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is smooth, convex.

If $\beta \in (0,1]$, then

$$\operatorname{Ent} e^{\varphi(X)} \leq C_{\beta,m,\sigma} \mathbb{E}\Big(\|\nabla\varphi(X)\|_2^2 \vee \|\nabla\varphi(X)\|_{(\beta+1)/\beta}^{(\beta+1)/\beta} \Big) e^{\varphi(X)}.$$

If $\beta = 0$ and $|\partial_i \varphi(x)| \leq 1/(2m + 6\sigma)$, then

$$\operatorname{Ent} e^{\varphi(X)} \leq C_{m,\sigma} \mathbb{E} \|\nabla\varphi(X)\|_2^2 e^{\varphi(X)}.$$

$$(\operatorname{Ent} Y = \mathbb{E} Y \log Y - \mathbb{E} Y \log \mathbb{E} Y,\ \|x\|_p^p = |x_1|^p + \ldots + |x_n|^p)$$

Corollaries

- $X_1, \ldots, X_n$ – independent r.v. with laws in $\mathcal{M}_\beta(m, \sigma^{\beta+1})$, $\beta \in [0,1]$,
- $X = (X_1, \ldots, X_n)$.

Concentration (convex functions)

$\varphi \colon \mathbb{R}^n \to \mathbb{R}$ is smooth, Lipschitz, convex. Then for $t \geq 0$,

$$\mathbb{P}(\varphi(X) \geq \mathbb{E}\varphi(X) + t) \leq \exp\Big(-C' \min\Big\{\frac{t^2}{\sup \|\nabla\varphi\|_2^2}, \frac{t^{1+\beta}}{\sup \|\nabla\varphi\|_{(\beta+1)/\beta}^{1+\beta}}\Big\}\Big).$$

Corollaries

- $X_1, \ldots, X_n$ – independent r.v. with laws in $\mathcal{M}_\beta(m, \sigma^{\beta+1})$, $\beta \in [0,1]$,
- $X = (X_1, \ldots, X_n)$.

Concentration (convex functions)

$\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is smooth, Lipschitz, convex. Then for $t \geq 0$,

$$\mathbb{P}(\varphi(X) \geq \mathbb{E}\varphi(X) + t) \leq \exp\Big(- C' \min\Big\{ \frac{t^2}{\sup \|\nabla\varphi\|_2^2}, \frac{t^{1+\beta}}{\sup \|\nabla\varphi\|_{(\beta+1)/\beta}^{1+\beta}} \Big\}\Big).$$

Concentration (enlargements of convex sets)

$A \subset \mathbb{R}^n$ is convex and $\mathbb{P}(X \in A) \geq 1/2$. Then for $r \geq 0$,

$$\mathbb{P}(X \notin A + r^{1/2} B_2 + r^{1/(1+\beta)} B_{1+\beta}) \leq e^{-C'' r}.$$

Proof of the Corollary (the Herbst argument), $\beta \in (0,1]$

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ convex, Lipschitz,
- $A = \sup \|\nabla\varphi(x)\|_2$, $B = \sup \|\nabla\varphi(x)\|_{(\beta+1)/\beta}$,
- For $\lambda > 0$ put $\Phi(\lambda) = \mathbb{E}e^{\lambda\varphi(X)}$.

Proof of the Corollary (the Herbst argument), $\beta \in (0, 1]$

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ convex, Lipschitz,
- $A = \sup \|\nabla \varphi(x)\|_2$, $B = \sup \|\nabla \varphi(x)\|_{(\beta+1)/\beta}$,
- For $\lambda > 0$ put $\Phi(\lambda) = \mathbb{E} e^{\lambda \varphi(X)}$.

Then $\Phi'(\lambda) = \mathbb{E} \varphi(X) e^{\lambda \varphi(X)}$ and

$$\lambda \Phi'(\lambda) - \Phi(\lambda) \log \Phi(\lambda) = \operatorname{Ent} e^{\lambda \varphi(X)}$$

Proof of the Corollary (the Herbst argument), $\beta \in (0,1]$

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ convex, Lipschitz,
- $A = \sup \|\nabla\varphi(x)\|_2$, $B = \sup \|\nabla\varphi(x)\|_{(\beta+1)/\beta}$,
- For $\lambda > 0$ put $\Phi(\lambda) = \mathbb{E}e^{\lambda\varphi(X)}$.

Then $\Phi'(\lambda) = \mathbb{E}\varphi(X)e^{\lambda\varphi(X)}$ and

$$\begin{aligned}\lambda\Phi'(\lambda) - \Phi(\lambda)\log\Phi(\lambda) &= \operatorname{Ent} e^{\lambda\varphi(X)} \\ &\le C\mathbb{E}\Big((\lambda\|\nabla\varphi(X)\|_2)^2 \vee (\lambda\|\nabla\varphi(X)\|_{\frac{\beta+1}{\beta}})^{\frac{\beta+1}{\beta}}\Big)e^{\lambda\varphi(X)} \\ &\le C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})\Phi(\lambda).\end{aligned}$$

Proof of the Corollary (the Herbst argument), $\beta \in (0,1]$

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ convex, Lipschitz,
- $A = \sup \|\nabla\varphi(x)\|_2$, $B = \sup \|\nabla\varphi(x)\|_{(\beta+1)/\beta}$,
- For $\lambda > 0$ put $\Phi(\lambda) = \mathbb{E}e^{\lambda\varphi(X)}$.

Then $\Phi'(\lambda) = \mathbb{E}\varphi(X)e^{\lambda\varphi(X)}$ and

$$\begin{aligned}\lambda\Phi'(\lambda) - \Phi(\lambda)\log\Phi(\lambda) &= \operatorname{Ent} e^{\lambda\varphi(X)} \\ &\le C\mathbb{E}\Big((\lambda\|\nabla\varphi(X)\|_2)^2 \vee (\lambda\|\nabla\varphi(X)\|_{\frac{\beta+1}{\beta}})^{\frac{\beta+1}{\beta}}\Big)e^{\lambda\varphi(X)} \\ &\le C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})\Phi(\lambda).\end{aligned}$$

Equivalently, $\Big(\frac{1}{\lambda}\log\Phi(\lambda)\Big)' \le C\big((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta}\big)/\lambda^2$.

Proof of the Corollary (the Herbst argument), $\beta \in (0,1]$

- $\varphi : \mathbb{R}^n \to \mathbb{R}$ convex, Lipschitz,
- $A = \sup \|\nabla\varphi(x)\|_2$, $B = \sup \|\nabla\varphi(x)\|_{(\beta+1)/\beta}$,
- For $\lambda > 0$ put $\Phi(\lambda) = \mathbb{E}e^{\lambda\varphi(X)}$.

Then $\Phi'(\lambda) = \mathbb{E}\varphi(X)e^{\lambda\varphi(X)}$ and

$$\begin{aligned}\lambda\Phi'(\lambda) - \Phi(\lambda)\log\Phi(\lambda) &= \mathrm{Ent}\, e^{\lambda\varphi(X)} \\ &\le C\mathbb{E}\Big((\lambda\|\nabla\varphi(X)\|_2)^2 \vee (\lambda\|\nabla\varphi(X)\|_{\frac{\beta+1}{\beta}})^{\frac{\beta+1}{\beta}}\Big)e^{\lambda\varphi(X)} \\ &\le C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})\Phi(\lambda).\end{aligned}$$

Equivalently, $\left(\frac{1}{\lambda}\log\Phi(\lambda)\right)' \le C\big((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta}\big)/\lambda^2$.

$\lim_{\lambda\to 0} \frac{1}{\lambda}\log\Phi(\lambda) = \mathbb{E}\varphi(X)$, so integration and definition of Φ give:

$$\Phi(\lambda) = \mathbb{E}e^{\lambda\varphi(X)} \le \exp(\lambda\mathbb{E}\varphi(X) + C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})).$$

Proof of the Corollary (the Herbst argument), $\beta \in (0, 1]$

We know that

$$\mathbb{E}e^{\lambda\varphi(X)} \leq \exp(\lambda\mathbb{E}\varphi(X) + C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})).$$

Proof of the Corollary (the Herbst argument), $\beta \in (0, 1]$

We know that

$$\mathbb{E}e^{\lambda\varphi(X)} \leq \exp(\lambda\mathbb{E}\varphi(X) + C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})).$$

Hence

$$\mathbb{P}(\varphi(X) \geq \mathbb{E}\varphi(X)+t) \leq \frac{\mathbb{E}e^{\lambda\varphi(X)}}{e^{\lambda\mathbb{E}\varphi(X)+\lambda t}} \leq \exp(-\lambda t+C((A\lambda)^2\vee(B\lambda)^{(\beta+1)/\beta}))).$$

Proof of the Corollary (the Herbst argument), $\beta \in (0,1]$

We know that

$$\mathbb{E}e^{\lambda\varphi(X)} \le \exp(\lambda\mathbb{E}\varphi(X) + C((A\lambda)^2 \vee (B\lambda)^{(\beta+1)/\beta})).$$

Hence

$$\mathbb{P}(\varphi(X) \ge \mathbb{E}\varphi(X)+t) \le \frac{\mathbb{E}e^{\lambda\varphi(X)}}{e^{\lambda\mathbb{E}\varphi(X)+\lambda t}} \le \exp(-\lambda t+C((A\lambda)^2\vee(B\lambda)^{(\beta+1)/\beta}))).$$

Optimization over λ gives

$$\begin{aligned}\mathbb{P}(\varphi(X) \ge \mathbb{E}\varphi(X) + t) &\le \exp\Big(- C' \min\Big\{\frac{t^2}{A^2}, \frac{t^{1+\beta}}{B^{1+\beta}}\Big\}\Big) \\ &= \exp\Big(- C' \min\Big\{\frac{t^2}{\sup\|\nabla\varphi\|_2^2}, \frac{t^{1+\beta}}{\sup\|\nabla\varphi\|_{(\beta+1)/\beta}^{1+\beta}}\Big\}\Big). \quad \square\end{aligned}$$

Some remarks

- Is our sufficient condition also necessary?

Some remarks

- Is our sufficient condition also necessary?
- Weak transport inequalities (cf. Gozlan, Roberto, Samson, Tetali).

Some remarks

- Is our sufficient condition also necessary?
- Weak transport inequalities (cf. Gozlan, Roberto, Samson, Tetali).
- Lower tail estimates for convex functions?

Some remarks

- Is our sufficient condition also necessary?
- Weak transport inequalities (cf. Gozlan, Roberto, Samson, Tetali).
- Lower tail estimates for convex functions?
- Feldheim, Marsiglietti, Nayar, Wang: stronger result for $\beta = 0$ (property τ).

The end

Thank you for your attention.

Questions?

[1] R.A., M.St. *Modified log-Sobolev inequalities for convex functions. Sufficient conditions*, `http://arxiv.org/abs/1505.05493`