Concentration and Convexity

Petros Valettas

Department of Mathematics
University of Missouri

EIMI: Asymptotic Geometric Analysis IV

St. Petersburg, July 2, 2019

Gaussian concentration

Theorem (Sudakov-Tsirel'son '74, Borell '75)

Let f be a Lipschitz map, i.e. $|f(x) - f(y)| \leq L\|x - y\|_2,\ x, y \in \mathbb{R}^n$. Then, for the standard Gaussian random vector G,

$$\mathbb{P}\left(f(G) < \mathrm{med}(f) - tL\right) \leq \Phi(-t) \leq \frac{1}{2}\exp\left(-t^2/2\right), \quad t > 0. \tag{1}$$

In particular,

$$\mathbb{P}(|f(G) - \mathrm{med}(f)| > tL) \leq \exp(-t^2/2),\ t > 0. \tag{2}$$

Gaussian concentration

Theorem (Sudakov-Tsirel'son '74, Borell '75)

Let f be a Lipschitz map, i.e. $|f(x) - f(y)| \leq L\|x - y\|_2,\ x, y \in \mathbb{R}^n$. Then, for the standard Gaussian random vector G,

$$\mathbb{P}\left(f(G) < \mathrm{med}(f) - tL\right) \leq \Phi(-t) \leq \frac{1}{2}\exp\left(-t^2/2\right), \quad t > 0. \tag{1}$$

In particular,

$$\mathbb{P}(|f(G) - \mathrm{med}(f)| > tL) \leq \exp(-t^2/2),\ t > 0. \tag{2}$$

- The probabilistic oscillations (fluctuations) are controlled by the metric oscillations (Lipschitz constant). That is $\mathrm{Var}[f(G)] \leq L^2$.

Gaussian concentration

Theorem (Sudakov-Tsirel'son '74, Borell '75)

Let f be a Lipschitz map, i.e. $|f(x) - f(y)| \leq L\|x - y\|_2$, $x, y \in \mathbb{R}^n$. Then, for the standard Gaussian random vector G,

$$\mathbb{P}\left(f(G) < \mathrm{med}(f) - tL\right) \leq \Phi(-t) \leq \frac{1}{2}\exp\left(-t^2/2\right), \quad t > 0. \tag{1}$$

In particular,

$$\mathbb{P}(|f(G) - \mathrm{med}(f)| > tL) \leq \exp(-t^2/2), \; t > 0. \tag{2}$$

- The probabilistic oscillations (fluctuations) are controlled by the metric oscillations (Lipschitz constant). That is $\mathrm{Var}[f(G)] \leq L^2$.
- Sharp for linear functionals.

Gaussian concentration

Theorem (Sudakov-Tsirel'son '74, Borell '75)

Let f be a Lipschitz map, i.e. $|f(x) - f(y)| \leq L\|x - y\|_2$, $x, y \in \mathbb{R}^n$. Then, for the standard Gaussian random vector G,

$$\mathbb{P}\left(f(G) < \mathrm{med}(f) - tL\right) \leq \Phi(-t) \leq \frac{1}{2}\exp\left(-t^2/2\right), \quad t > 0. \tag{1}$$

In particular,

$$\mathbb{P}(|f(G) - \mathrm{med}(f)| > tL) \leq \exp(-t^2/2), \; t > 0. \tag{2}$$

- The probabilistic oscillations (fluctuations) are controlled by the metric oscillations (Lipschitz constant). That is $\mathrm{Var}[f(G)] \leq L^2$.
- Sharp for linear functionals.
- Consequence of the isoperimetric principle. "isoperimetry + continuity".

Application to ℓ_p norms

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G) - m| \geq tm\right) \leq \exp(-ct^2 k), \quad t > 0, \quad k \equiv k(f) := (m/L)^2.$$

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G) - m| \geq tm\right) \leq \exp(-ct^2 k), \quad t > 0, \quad k \equiv k(f) := (m/L)^2.$$

In the *large deviation regime* the latter is optimal, that is

$$\mathbb{P}(f(G) \geq (1+t)m) \geq ce^{-Ct^2 k}, \quad t > 1.$$

Let's focus on the *small deviation regime*:

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G)-m| \geq tm\right) \leq \exp(-ct^2k), \quad t>0, \quad k \equiv k(f) := (m/L)^2.$$

In the *large deviation regime* the latter is optimal, that is

$$\mathbb{P}(f(G) \geq (1+t)m) \geq ce^{-Ct^2k}, \quad t>1.$$

Let's focus on the *small deviation regime*:

- The ℓ_p norm for $1 \leq p \leq 2$: $f(x) = \|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. *Optimal.*

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G) - m| \geq tm\right) \leq \exp(-ct^2 k), \quad t > 0, \quad k \equiv k(f) := (m/L)^2.$$

In the *large deviation regime* the latter is optimal, that is

$$\mathbb{P}(f(G) \geq (1+t)m) \geq ce^{-Ct^2 k}, \quad t > 1.$$

Let's focus on the *small deviation regime*:

- The ℓ_p norm for $1 \leq p \leq 2$: $f(x) = \|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$. *Optimal.*
- The ℓ_q norm for $2 < q \leq \infty$. *Sub-optimal.*
  $\mathrm{Lip}(\|\cdot\|_q) = 1$, but $\mathrm{Var}(\|G\|_q) = o_n(1)$. E.g. $\mathrm{Var}(\|G\|_4) \asymp n^{-1/2}$.

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G) - m| \geq tm\right) \leq \exp(-ct^2 k), \quad t > 0, \quad k \equiv k(f) := (m/L)^2.$$

In the *large deviation regime* the latter is optimal, that is

$$\mathbb{P}(f(G) \geq (1+t)m) \geq ce^{-Ct^2 k}, \quad t > 1.$$

Let's focus on the *small deviation regime*:

- The ℓ_p norm for $1 \leq p \leq 2$: $f(x) = \|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$. *Optimal.*
- The ℓ_q norm for $2 < q \leq \infty$. *Sub-optimal.*
$\mathrm{Lip}(\|\cdot\|_q) = 1$, but $\mathrm{Var}(\|G\|_q) = o_n(1)$. E.g. $\mathrm{Var}(\|G\|_4) \asymp n^{-1/2}$.
For improving the variance we may employ the Poincaré inequality:

$$\mathrm{Var}[f(G)] \leq \mathbb{E}\|\nabla f(G)\|_2^2 \leq \mathrm{Lip}(f)^2.$$

Application to ℓ_p norms

- For f being a norm we obtain:

$$\mathbb{P}\left(|f(G)-m|\geq tm\right)\leq \exp(-ct^2k), \quad t>0, \quad k\equiv k(f):=(m/L)^2.$$

In the *large deviation regime* the latter is optimal, that is

$$\mathbb{P}(f(G)\geq (1+t)m)\geq ce^{-Ct^2k}, \quad t>1.$$

Let's focus on the *small deviation regime*:

- The ℓ_p norm for $1\leq p\leq 2$: $f(x)=\|x\|_p:=(\sum_{i=1}^n |x_i|^p)^{1/p}$. *Optimal.*
- The ℓ_q norm for $2<q\leq\infty$. *Sub-optimal.*
  $\mathrm{Lip}(\|\cdot\|_q)=1$, but $\mathrm{Var}(\|G\|_q)=o_n(1)$. E.g. $\mathrm{Var}(\|G\|_4)\asymp n^{-1/2}$.
  For improving the variance we may employ the Poincaré inequality:

$$\mathrm{Var}[f(G)]\leq \mathbb{E}\|\nabla f(G)\|_2^2 \leq \mathrm{Lip}(f)^2.$$

Indeed; we have

$$\mathbb{E}\|\nabla\|G\|_4\|_2^2\asymp n^{-1/2}.$$

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal*.

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal.*
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have

$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla \|G\|_\infty\|_2^2.$$

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal*.
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have
$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla \|G\|_\infty\|_2^2.$$

This is a typical example of *superconcentration*. Following Chatterjee, superconcentration occurs when the classical methods fail to capture the sharp estimates.

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal.*
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have

$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla\|G\|_\infty\|_2^2.$$

This is a typical example of *superconcentration*. Following Chatterjee, superconcentration occurs when the classical methods fail to capture the sharp estimates.

Definition (Chatterjee)

A smooth function f on $\mathbb{R}^n$ is said to be ε_n-superconcentrated if

$$\mathrm{Var}[f(G)] \leq \varepsilon_n \mathbb{E}\|\nabla f(G)\|_2^2.$$

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal.*
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have

$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla \|G\|_\infty\|_2^2.$$

This is a typical example of *superconcentration*. Following Chatterjee, superconcentration occurs when the classical methods fail to capture the sharp estimates.

Definition (Chatterjee)

A smooth function f on $\mathbb{R}^n$ is said to be ε_n-superconcentrated if

$$\mathrm{Var}[f(G)] \leq \varepsilon_n \mathbb{E}\|\nabla f(G)\|_2^2.$$

Inadequacy of the classical concentration in direct-problem solving

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal.*
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have

$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla \|G\|_\infty\|_2^2.$$

This is a typical example of *superconcentration*. Following Chatterjee, superconcentration occurs when the classical methods fail to capture the sharp estimates.

Definition (Chatterjee)

A smooth function f on $\mathbb{R}^n$ is said to be ε_n-superconcentrated if

$$\mathrm{Var}[f(G)] \leq \varepsilon_n \mathbb{E}\|\nabla f(G)\|_2^2.$$

Inadequacy of the classical concentration in direct-problem solving
Q.1.Under what conditions is the classical concentration inequality tight?

- The ℓ_∞ norm: $f(x) = \|x\|_\infty$. *Sub-optimal.*
  $\|\nabla \|G\|_\infty\|_2 = 1$ a.s. However, we have

$$\mathrm{Var}(\|G\|_\infty) \asymp (\log n)^{-1} \ll \mathbb{E}\|\nabla \|G\|_\infty\|_2^2.$$

This is a typical example of *superconcentration*. Following Chatterjee, superconcentration occurs when the classical methods fail to capture the sharp estimates.

Definition (Chatterjee)

A smooth function f on $\mathbb{R}^n$ is said to be ε_n-superconcentrated if

$$\mathrm{Var}[f(G)] \leq \varepsilon_n \mathbb{E}\|\nabla f(G)\|_2^2.$$

Inadequacy of the classical concentration in direct-problem solving

Q.1.Under what conditions is the classical concentration inequality tight?

Q.2. What assumptions ensure sharper concentration bounds?

An anticoncentration inequality

An anticoncentration inequality

If (2) is optimal, then $\mathrm{Var}[f(G)] \gtrsim L^2$ (necessary condition).

An anticoncentration inequality

If (2) is optimal, then $\mathrm{Var}[f(G)] \gtrsim L^2$ (necessary condition).
It turns out that for convex, Lipschitz maps it is also a sufficient condition.

An anticoncentration inequality

If (2) is optimal, then $\mathrm{Var}[f(G)] \gtrsim L^2$ (necessary condition).
It turns out that for convex, Lipschitz maps it is also a sufficient condition.

Theorem (V. '17)

Let $\alpha \in (0,1)$ and let f be a convex and Lipschitz map with $\mathrm{Var}[f(G)] \geq \alpha L^2$. Then, for all $t > 0$ we have

$$\mathbb{P}(|f(G) - m| \geq tL) \geq ce^{-Ct^2},$$

where $c, C > 0$ depend solely on $\alpha > 0$.

An anticoncentration inequality

If (2) is optimal, then $\mathrm{Var}[f(G)] \gtrsim L^2$ (necessary condition).
It turns out that for convex, Lipschitz maps it is also a sufficient condition.

Theorem (V. '17)

Let $\alpha \in (0,1)$ and let f be a convex and Lipschitz map with $\mathrm{Var}[f(G)] \geq \alpha L^2$. Then, for all $t > 0$ we have

$$\mathbb{P}(|f(G) - m| \geq tL) \geq ce^{-Ct^2},$$

where $c, C > 0$ depend solely on $\alpha > 0$.

- The Gaussian convexity is used in a crucial way. Ehrhard's work on Gaussian rearrangements ('84): $f^*(s) = \inf\{t : \Phi^{-1} \circ P(f(G) \leq t) \geq s\}$.

An anticoncentration inequality

If (2) is optimal, then $\mathrm{Var}[f(G)] \gtrsim L^2$ (necessary condition).
It turns out that for convex, Lipschitz maps it is also a sufficient condition.

Theorem (V. '17)

Let $\alpha \in (0,1)$ and let f be a convex and Lipschitz map with $\mathrm{Var}[f(G)] \geq \alpha L^2$. Then, for all $t > 0$ we have

$$\mathbb{P}(|f(G) - m| \geq tL) \geq ce^{-Ct^2},$$

where $c, C > 0$ depend solely on $\alpha > 0$.

- The Gaussian convexity is used in a crucial way. Ehrhard's work on Gaussian rearrangements ('84): $f^*(s) = \inf\{t : \Phi^{-1} \circ P(f(G) \leq t) \geq s\}$.
- An improved Poincaré inequality due to Bobkov and Houdré ('99).

Theorem (V. '17)

Let f be a convex map. Then, for the standard Gaussian vector G we have

$$\mathbb{P}(f(G) \leq m - t) \leq \mathbb{P}(f(G) \geq m + t), \; t > 0.$$

Theorem (V. '17)

Let f be a convex map. Then, for the standard Gaussian vector G we have

$$\mathbb{P}(f(G) \leq m - t) \leq \mathbb{P}(f(G) \geq m + t), \ t > 0.$$

- The r.v. $Z = f(G)$ is skew to the left. Hence, convex functions concentrate more below the mean.

Theorem (V. '17)

Let f be a convex map. Then, for the standard Gaussian vector G we have

$$\mathbb{P}(f(G) \leq m - t) \leq \mathbb{P}(f(G) \geq m + t), \ t > 0.$$

- The r.v. $Z = f(G)$ is skew to the left. Hence, convex functions concentrate more below the mean.
- Upper deviation estimate suffices to obtain two-sided results.

Theorem (V. '17)

Let f be a convex map. Then, for the standard Gaussian vector G we have

$$\mathbb{P}(f(G) \leq m - t) \leq \mathbb{P}(f(G) \geq m + t), \ t > 0.$$

- The r.v. $Z = f(G)$ is skew to the left. Hence, convex functions concentrate more below the mean.
- Upper deviation estimate suffices to obtain two-sided results.
- Proof uses "Gaussian convexity": Ehrhard's inequality (1983). For any A, B convex sets in $\mathbb{R}^n$ and $0 < \lambda < 1$ one has

$$\Phi^{-1} \circ \gamma_n((1-\lambda)A + \lambda B) \geq (1-\lambda)\Phi^{-1} \circ \gamma_n(A) + \lambda \Phi^{-1} \circ \gamma_n(B)$$

"Isoperimetry + Convexity"

"Isoperimetry + Convexity"

Theorem (Bobkov-Götze '99, Samson '03)

Let f be a convex map and let μ be a Borel probability measure on $\mathbb{R}^n$ which satisfies a transportation cost inequality with constant $A > 0$ (e.g. γ_n does with $A = 1$), i.e. $W_2(\mu, \nu) \leq \sqrt{2AD(\nu||\mu)}$ for any probability measure ν. Then, for any convex map f one has for all $t > 0$,

$$\mu(x : f(x) \leq \mathbb{E}_\mu f - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2}) \leq e^{-t^2/A}. \tag{3}$$

"Isoperimetry + Convexity"

Theorem (Bobkov-Götze '99, Samson '03)

Let f be a convex map and let μ be a Borel probability measure on $\mathbb{R}^n$ which satisfies a transportation cost inequality with constant $A > 0$ (e.g. γ_n does with $A = 1$), i.e. $W_2(\mu, \nu) \leq \sqrt{2AD(\nu||\mu)}$ for any probability measure ν. Then, for any convex map f one has for all $t > 0$,

$$\mu(x : f(x) \leq \mathbb{E}_\mu f - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2}) \leq e^{-t^2/A}. \quad (3)$$

- Improves upon (1), since $\mathbb{E}\|\nabla f\|_2^2 \leq \mathrm{Lip}(f)^2$.

"Isoperimetry + Convexity"

Theorem (Bobkov-Götze '99, Samson '03)

Let f be a convex map and let μ be a Borel probability measure on $\mathbb{R}^n$ which satisfies a transportation cost inequality with constant $A > 0$ (e.g. γ_n does with $A = 1$), i.e. $W_2(\mu, \nu) \leq \sqrt{2AD(\nu||\mu)}$ for any probability measure ν. Then, for any convex map f one has for all $t > 0$,

$$\mu(x : f(x) \leq \mathbb{E}_\mu f - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2}) \leq e^{-t^2/A}. \tag{3}$$

- Improves upon (1), since $\mathbb{E}\|\nabla f\|_2^2 \leq \mathrm{Lip}(f)^2$.
- *An observation.* For a smooth function f with $\nabla^2 f \succ -aI$, one has

$$\mu(x : f(x) \leq m - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2} - at^2) \leq \alpha_\mu(t), \ t > 0,$$

"Isoperimetry + Convexity"

Theorem (Bobkov-Götze '99, Samson '03)

Let f be a convex map and let μ be a Borel probability measure on $\mathbb{R}^n$ which satisfies a transportation cost inequality with constant $A > 0$ (e.g. γ_n does with $A = 1$), i.e. $W_2(\mu, \nu) \leq \sqrt{2AD(\nu||\mu)}$ for any probability measure ν. Then, for any convex map f one has for all $t > 0$,

$$\mu(x : f(x) \leq \mathbb{E}_\mu f - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2}) \leq e^{-t^2/A}. \tag{3}$$

- Improves upon (1), since $\mathbb{E}\|\nabla f\|_2^2 \leq \mathrm{Lip}(f)^2$.
- *An observation.* For a smooth function f with $\nabla^2 f \succ -aI$, one has

$$\mu(x : f(x) \leq m - t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2 - at^2}) \leq \alpha_\mu(t), \ t > 0,$$

- In particular, for smooth f with $\|\nabla^2 f\|_{\mathrm{op}} \leq K$, we obtain *Hanson-Wright* type bounds:

$$\mu(|f - m| \geq t\sqrt{\mathbb{E}_\mu \|\nabla f\|_2^2} + t^2 K) \leq 2\alpha_\mu(t).$$

"Convexity + Convexity"

"Convexity + Convexity"

Theorem (Paouris, V., '16)

Let f be a convex function on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$. Then, we have

$$\mathbb{P}\left(f(G) - \mathrm{med}(f) \leq -t\sqrt{\mathrm{Var}[f(G)]}\right) \leq \Phi\left(-\frac{t}{\sqrt{2\pi}}\right) \leq e^{-ct^2} \quad t > 0, \tag{4}$$

where G is the standard Gaussian vector.

"Convexity + Convexity"

Theorem (Paouris, V., '16)

Let f be a convex function on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$. Then, we have

$$\mathbb{P}\left(f(G) - \mathrm{med}(f) \leq -t\sqrt{\mathrm{Var}[f(G)]}\right) \leq \Phi\left(-\frac{t}{\sqrt{2\pi}}\right) \leq e^{-ct^2} \quad t > 0, \qquad (4)$$

where G is the standard Gaussian vector.

- The function is not necessarily Lipschitz.

“Convexity + Convexity”

Theorem (Paouris, V., ’16)

Let f be a convex function on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$. Then, we have

$$\mathbb{P}\left(f(G) - \mathrm{med}(f) \leq -t\sqrt{\mathrm{Var}[f(G)]}\right) \leq \Phi\left(-\frac{t}{\sqrt{2\pi}}\right) \leq e^{-ct^2} \quad t > 0, \tag{4}$$

where G is the standard Gaussian vector.

- The function is not necessarily Lipschitz.
- This is a much stronger inequality, inside the class of convex (Lipschitz) functions, since we may have $\mathrm{Var}[f(G)] \ll \mathbb{E}\|\nabla f(G)\|_2^2 \ll L^2$. Takes into account the superconcentration.

"Convexity + Convexity"

Theorem (Paouris, V., '16)

Let f be a convex function on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$. Then, we have

$$\mathbb{P}\left(f(G) - \mathrm{med}(f) \leq -t\sqrt{\mathrm{Var}[f(G)]}\right) \leq \Phi\left(-\frac{t}{\sqrt{2\pi}}\right) \leq e^{-ct^2} \quad t > 0, \qquad (4)$$

where G is the standard Gaussian vector.

- The function is not necessarily Lipschitz.
- This is a much stronger inequality, inside the class of convex (Lipschitz) functions, since we may have $\mathrm{Var}[f(G)] \ll \mathbb{E}\|\nabla f(G)\|_2^2 \ll L^2$. Takes into account the superconcentration.
- New type of concentration, which is not explained by isoperimetry: It exploits the convexity properties of the distribution.

"Convexity + Convexity"

Theorem (Paouris, V., '16)

Let f be a convex function on $\mathbb{R}^n$ with $f \in L_2(\gamma_n)$. Then, we have

$$\mathbb{P}\left(f(G) - \mathrm{med}(f) \leq -t\sqrt{\mathrm{Var}[f(G)]}\right) \leq \Phi\left(-\frac{t}{\sqrt{2\pi}}\right) \leq e^{-ct^2} \quad t > 0, \qquad (4)$$

where G is the standard Gaussian vector.

- The function is not necessarily Lipschitz.
- This is a much stronger inequality, inside the class of convex (Lipschitz) functions, since we may have $\mathrm{Var}[f(G)] \ll \mathbb{E}\|\nabla f(G)\|_2^2 \ll L^2$. Takes into account the superconcentration.
- New type of concentration, which is not explained by isoperimetry: It exploits the convexity properties of the distribution.
- Similar estimates for any log-concave probability measure μ on $\mathbb{R}^n$ and any convex map f.

A (sketch) proof of (4) (with worse constant).

- Let f be convex map and let $F(t) = \mathbb{P}(f(G) \leq t)$. Ehrhard's ineq. implies that $t \mapsto \Phi^{-1} \circ F(t)$ is concave.

A (sketch) proof of (4) (with worse constant).

- Let f be convex map and let $F(t) = \mathbb{P}(f(G) \leq t)$. Ehrhard's ineq. implies that $t \mapsto \Phi^{-1} \circ F(t)$ is concave.
- If m is a median of f, then concavity yields

$$\Phi^{-1} \circ F(m-t) \leq \Phi^{-1} \circ F(m) - t(\Phi^{-1} \circ F)'(m) = -t(\Phi^{-1} \circ F)'(m).$$

A (sketch) proof of (4) (with worse constant).

- Let f be convex map and let $F(t) = \mathbb{P}(f(G) \leq t)$. Ehrhard's ineq. implies that $t \mapsto \Phi^{-1} \circ F(t)$ is concave.
- If m is a median of f, then concavity yields

$$\Phi^{-1} \circ F(m-t) \leq \Phi^{-1} \circ F(m) - t(\Phi^{-1} \circ F)'(m) = -t(\Phi^{-1} \circ F)'(m).$$

- We may compute $(\Phi^{-1} \circ F)'(m) = \sqrt{2\pi} F'(m)$.

A (sketch) proof of (4) (with worse constant).

- Let f be convex map and let $F(t) = \mathbb{P}(f(G) \leq t)$. Ehrhard's ineq. implies that $t \mapsto \Phi^{-1} \circ F(t)$ is concave.
- If m is a median of f, then concavity yields

$$\Phi^{-1} \circ F(m-t) \leq \Phi^{-1} \circ F(m) - t(\Phi^{-1} \circ F)'(m) = -t(\Phi^{-1} \circ F)'(m).$$

- We may compute $(\Phi^{-1} \circ F)'(m) = \sqrt{2\pi} F'(m)$.
- A lower bound for $F'(m)$:

A (sketch) proof of (4) (with worse constant).

- Let f be convex map and let $F(t) = \mathbb{P}(f(G) \leq t)$. Ehrhard's ineq. implies that $t \mapsto \Phi^{-1} \circ F(t)$ is concave.
- If m is a median of f, then concavity yields

$$\Phi^{-1} \circ F(m-t) \leq \Phi^{-1} \circ F(m) - t(\Phi^{-1} \circ F)'(m) = -t(\Phi^{-1} \circ F)'(m).$$

- We may compute $(\Phi^{-1} \circ F)'(m) = \sqrt{2\pi} F'(m)$.
- A lower bound for $F'(m)$:

  Note that F is log-concave, hence for any $\delta > 0$ (choose later) we may write

$$\begin{aligned} F'(m) &= F(m)(\log F)'(m) = \tfrac{1}{2}(\log F)'(m) \\ &\geq \frac{\log F(m+\delta) - \log F(m)}{2\delta} \\ &= \dots \\ &= \frac{1}{2\delta} \log\left(1 + 2\mathbb{P}(m \leq f(G) \leq m+\delta)\right) \\ &\geq \frac{1}{2\delta} \mathbb{P}(m \leq f \leq m+\delta). \end{aligned}$$

Proof (cont'd)

Therefore, we obtain

$$F'(m) \geq \frac{1}{2\delta}\left[\frac{1}{2} - \mathbb{P}(f \geq m+\delta)\right] \geq \frac{1}{2\delta}\left(\frac{1}{2} - \frac{\|f-m\|_{L_1}}{\delta}\right).$$

Proof (cont'd)

Therefore, we obtain

$$F'(m) \geq \frac{1}{2\delta}\left[\frac{1}{2} - \mathbb{P}(f \geq m+\delta)\right] \geq \frac{1}{2\delta}\left(\frac{1}{2} - \frac{\|f-m\|_{L_1}}{\delta}\right).$$

To conclude, choose $\delta = 4\|f-m\|_{L_1}$.

Proof (cont'd)

Therefore, we obtain

$$F'(m) \geq \frac{1}{2\delta}\left[\frac{1}{2} - \mathbb{P}(f \geq m+\delta)\right] \geq \frac{1}{2\delta}\left(\frac{1}{2} - \frac{\|f-m\|_{L_1}}{\delta}\right).$$

To conclude, choose $\delta = 4\|f-m\|_{L_1}$.

- Finally, we obtain

$$\Phi^{-1} \circ F(m-t) \leq -\frac{t\sqrt{2\pi}}{32\|f-m\|_{L_1}}$$

Proof (cont'd)

Therefore, we obtain

$$F'(m) \geq \frac{1}{2\delta}\left[\frac{1}{2} - \mathbb{P}(f \geq m+\delta)\right] \geq \frac{1}{2\delta}\left(\frac{1}{2} - \frac{\|f-m\|_{L_1}}{\delta}\right).$$

To conclude, choose $\delta = 4\|f-m\|_{L_1}$.

- Finally, we obtain

$$\Phi^{-1}\circ F(m-t) \leq -\frac{t\sqrt{2\pi}}{32\|f-m\|_{L_1}} \Longrightarrow F(m-t) \leq \Phi\left(-\frac{ct}{\|f-m\|_{L_1}}\right).$$

Thank you for your attention!