Concentration inequalities for non-Lipschitz functions

Paweł Wolff

University of Warsaw

Berkeley, October 1, 2013

joint work with Radosław Adamczak (University of Warsaw)

Gaussian concentration (Sudakov-Tsirelson, Borell, 1970s)

$f\colon \mathbb{R}^n \to \mathbb{R}$ — L-Lipschitz

G — standard Gaussian random vector in $\mathbb{R}^n$

$$\mathbb{P}(|f(G) - \mathbb{E}f(G)| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

Gaussian concentration (Sudakov-Tsirelson, Borell, 1970s)

$f\colon \mathbb{R}^n \to \mathbb{R}$ — L-Lipschitz

G — standard Gaussian random vector in $\mathbb{R}^n$

$$\mathbb{P}(|f(G) - \mathbb{E}f(G)| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

Examples:

- $f(x) = \langle a, x\rangle$ with $|a|_2 = 1$; then $f(G) \sim \mathcal{N}(0, 1)$

Gaussian concentration (Sudakov-Tsirelson, Borell, 1970s)

$f\colon \mathbb{R}^n \to \mathbb{R}$ — L-Lipschitz

G — standard Gaussian random vector in $\mathbb{R}^n$

$$\mathbb{P}(|f(G) - \mathbb{E}f(G)| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

Examples:

- $f(x) = \langle a, x \rangle$ with $|a|_2 = 1$; then $f(G) \sim \mathcal{N}(0,1)$
- $f(x) = |Ax|_2$ is Lipschitz with constant $\|A\|_{\mathrm{op}}$

Gaussian concentration (Sudakov-Tsirelson, Borell, 1970s)

$f\colon \mathbb{R}^n \to \mathbb{R}$ — L-Lipschitz

G — standard Gaussian random vector in $\mathbb{R}^n$

$$\mathbb{P}(|f(G) - \mathbb{E}f(G)| \geq t) \leq 2\exp\left(-\frac{t^2}{2L^2}\right)$$

Examples:

- $f(x) = \langle a, x\rangle$ with $|a|_2 = 1$; then $f(G) \sim \mathcal{N}(0,1)$
- $f(x) = |Ax|_2$ is Lipschitz with constant $\|A\|_{\mathrm{op}}$

Proofs:

- Gaussian isoperimetric inequality
- stochastic calculus (Maurey)
- functional inequalities, semigroup methods (Gross-Herbst, Bobkov, Ledoux, ...)

Functional inequalities approach

$$Y \geq 0; \qquad \operatorname{Ent} Y = \mathbb{E} Y \log Y - \mathbb{E} Y \log \mathbb{E} Y$$

$Y \geq 0; \qquad \mathrm{Ent}\, Y = \mathbb{E} Y \log Y - \mathbb{E} Y \log \mathbb{E} Y$

Logarithmic Sobolev inequality

X random vector in $\mathbb{R}^n$ satisfies (LSI) with constant L if

$$\forall_{f\colon \mathbb{R}^n \to \mathbb{R}} \quad \mathrm{Ent} f^2(X) \leq 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

Functional inequalities approach

$Y \geq 0; \qquad \mathrm{Ent}\, Y = \mathbb{E} Y \log Y - \mathbb{E} Y \log \mathbb{E} Y$

Logarithmic Sobolev inequality

X random vector in $\mathbb{R}^n$ satisfies (LSI) with constant L if

$$\forall_{f\colon \mathbb{R}^n \to \mathbb{R}} \quad \mathrm{Ent} f^2(X) \leq 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

Classical facts:

- standard Gaussian in $\mathbb{R}^n$ satisfies (LSI) with $L = 1$ (Gross)

$Y \geq 0; \qquad \operatorname{Ent} Y = \mathbb{E} Y \log Y - \mathbb{E} Y \log \mathbb{E} Y$

Logarithmic Sobolev inequality

X random vector in $\mathbb{R}^n$ satisfies (LSI) with constant L if

$$\forall_{f\colon \mathbb{R}^n \to \mathbb{R}} \quad \operatorname{Ent} f^2(X) \leq 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

Classical facts:

- standard Gaussian in $\mathbb{R}^n$ satisfies (LSI) with $L = 1$ (Gross)
- (LSI) for $X \implies$ Gaussian concentration for X (Herbst):
  $f\colon \mathbb{R}^n \to \mathbb{R}$ is 1-Lipschitz

(LSI) for $\exp(\lambda f(\cdot)/2) \implies \mathbb{P}(|f(X) - \mathbb{E} f(X)| \geq t) \leq 2 \exp\left(-\frac{t^2}{2L}\right)$

Functional inequalities approach

$$\forall_{f\colon \mathbb{R}^n\to\mathbb{R}} \quad \mathrm{Ent} f^2(X) \le 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

(Aida-Stroock, 1994) $\Downarrow$

$$\forall_{f\colon \mathbb{R}^n\to\mathbb{R}} \quad \|f(X) - \mathbb{E}f(X)\|_p \le \sqrt{L}\sqrt{p}\,\||\nabla f(X)|\|_p \qquad (*)$$

Functional inequalities approach

$$\forall_{f\colon \mathbb{R}^n\to\mathbb{R}} \quad \mathrm{Ent} f^2(X) \le 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

(Aida-Stroock, 1994) $\Downarrow$

$$\forall_{f\colon \mathbb{R}^n\to\mathbb{R}} \quad \|f(X)-\mathbb{E}f(X)\|_p \le \sqrt{L}\sqrt{p}\,\||\nabla f(X)|\|_p \qquad (*)$$

For Gaussian measure — ineq. (*) due to Pisier (1986)

Functional inequalities approach

$$\forall_{f\colon \mathbb{R}^n \to \mathbb{R}} \quad \mathrm{Ent} f^2(X) \le 2L\mathbb{E}|\nabla f(X)|^2 \qquad \text{(LSI)}$$

(Aida-Stroock, 1994) $\Downarrow$

$$\forall_{f\colon \mathbb{R}^n \to \mathbb{R}} \quad \|f(X) - \mathbb{E}f(X)\|_p \le \sqrt{L}\sqrt{p}\,\||\nabla f(X)|\|_p \qquad (*)$$

For Gaussian measure — ineq. (*) due to Pisier (1986)

Equivalent form for (*): $G \sim \mathcal{N}(0, I_n)$, independent of X

$$\|f(X) - \mathbb{E}f(X)\|_p \le C\sqrt{L}\,\|\langle \nabla f(X), G\rangle\|_p$$

Tails vs moments

Let $Y \geq 0$ (e.g. $Y = |f(X) - \mathbb{E}f(X)|$). By **Chebyshev's** inequality

$$\mathbb{P}(Y \geq e\|Y\|_p) \leq \exp(-p).$$

Tails vs moments

Let $Y \geq 0$ (e.g. $Y = |f(X) - \mathbb{E}f(X)|$). By **Chebyshev's** inequality

$$\mathbb{P}(Y \geq e\|Y\|_p) \leq \exp(-p).$$

Examples:

- $\|Y\|_p \leq C\sqrt{p} \quad \Longrightarrow \quad \mathbb{P}(Y \geq t) \leq \exp(-t^2/(eC)^2)$
- $\|Y\|_p \leq Cp \quad \Longrightarrow \quad \mathbb{P}(Y \geq t) \leq \exp(-t/(eC))$

Tails vs moments

Let $Y \geq 0$ (e.g. $Y = |f(X) - \mathbb{E}f(X)|$). By **Chebyshev's** inequality

$$\mathbb{P}(Y \geq e\|Y\|_p) \leq \exp(-p).$$

Examples:

- $\|Y\|_p \leq C\sqrt{p} \quad \Longrightarrow \quad \mathbb{P}(Y \geq t) \leq \exp(-t^2/(eC)^2)$
- $\|Y\|_p \leq Cp \quad \Longrightarrow \quad \mathbb{P}(Y \geq t) \leq \exp(-t/(eC))$

Paley-Zygmund: l. b. for moments + regularity $\Longrightarrow$ l. b. for tails

$$Z \geq 0 \text{ (take } Z = Y^p), \lambda \in (0,1):$$

$$\mathbb{P}(Z \geq \lambda \mathbb{E}Z) \geq (1-\lambda)^2 \frac{(\mathbb{E}Z)^2}{\mathbb{E}Z^2}.$$

If $\|Y\|_{2p} \leq D\|Y\|_p$ then taking $\lambda = 2^{-p}$,

$$\mathbb{P}(Y \geq \frac{1}{2}\|Y\|_p) \geq \exp(-c_D p).$$

Beyond Lipschitz I

Quadratic form in Gaussian vector:

$$f(G) = G^t A G$$

where A symmetric $n \times n$, zeros on the diagonal

Beyond Lipschitz I

Quadratic form in Gaussian vector:

$$f(G) = G^t A G$$

where A symmetric $n \times n$, zeros on the diagonal

$$G^t A G = G^t (Q^t D Q) G \sim G^t D G = \sum d_i g_i^2 = \sum d_i (g_i^2 - 1)$$

Quadratic form in Gaussian vector:

$$f(G) = G^t A G$$

where A symmetric $n \times n$, zeros on the diagonal

$$G^t A G = G^t (Q^t D Q) G \sim G^t D G = \sum d_i g_i^2 = \sum d_i (g_i^2 - 1)$$

By Bernstein's inequality (for sums of indep. subexp. r.v.'s)

$$\mathbb{P}(|f(G)| \geq t) = \mathbb{P}(|\sum d_i (g_i^2 - 1)| \geq t)$$

$$\leq 2 \exp\left(-c\left(\frac{t^2}{|d|_2^2} \wedge \frac{t}{\|d\|_\infty}\right)\right)$$

$$|d|_2 = \|A\|_{\mathrm{HS}} \qquad \|d\|_\infty = \|A\|_{\mathrm{op}}$$

Beyond Lipschitz I

Quadratic form in Gaussian vector:

$$f(G) = G^t A G$$

where A symmetric $n \times n$, zeros on the diagonal

$$G^t A G = G^t (Q^t D Q) G \sim G^t D G = \sum d_i g_i^2 = \sum d_i (g_i^2 - 1)$$

By Bernstein's inequality (for sums of indep. subexp. r.v.'s)

$$\mathbb{P}(|f(G)| \geq t) = \mathbb{P}(|\sum d_i (g_i^2 - 1)| \geq t)$$

$$\leq 2 \exp\left(-c\left(\frac{t^2}{|d|_2^2} \wedge \frac{t}{\|d\|_\infty}\right)\right)$$

$$|d|_2 = \|A\|_{\mathrm{HS}} \qquad \|d\|_\infty = \|A\|_{\mathrm{op}}$$

Moments: $\|f(G)\|_p \leq C(\sqrt{p}\,\|A\|_{\mathrm{HS}} + p\,\|A\|_{\mathrm{op}})$

Beyond Lipschitz II

$A = (a_{i_1 \ldots i_d})_{i_1,\ldots i_d \le n}$ — symmetric, zeros on all "diagonals"

$$f(G) = A(G, \ldots, G) = \sum_{i_1,\ldots,i_d} a_{i_1 \ldots i_d} g_{i_1} \cdots g_{i_d}$$

Latała (2006): two-sided estimates for moments and tails of $f(G)$

Beyond Lipschitz II

$A = (a_{i_1 \dots i_d})_{i_1, \dots i_d \le n}$ — symmetric, zeros on all "diagonals"

$$f(G) = A(G, \dots, G) = \sum_{i_1, \dots, i_d} a_{i_1 \dots i_d} g_{i_1} \cdots g_{i_d}$$

Latała (2006): two-sided estimates for moments and tails of $f(G)$

For $d = 3$,

$$\|f(G)\|_p \sim \sqrt{p}\|A\|_{\{123\}} + p\|A\|_{\{12\}\{3\}} + p^{3/2}\|A\|_{\{1\}\{2\}\{3\}}$$

$A = (a_{i_1 \ldots i_d})_{i_1,\ldots i_d \le n}$ — symmetric, zeros on all "diagonals"

$$f(G) = A(G, \ldots, G) = \sum_{i_1,\ldots,i_d} a_{i_1 \ldots i_d} g_{i_1} \cdots g_{i_d}$$

Latała (2006): two-sided estimates for moments and tails of $f(G)$

For $d = 3$,

$$\|f(G)\|_p \sim \sqrt{p}\|A\|_{\{123\}} + p\|A\|_{\{12\}\{3\}} + p^{3/2}\|A\|_{\{1\}\{2\}\{3\}}$$

where

$$\|A\|_{\{123\}} = \Big(\sum_{ijk} a_{ijk}^2\Big)^{1/2} = \sup\{\sum_{ijk} a_{ijk} x_{ijk} : \sum_{ijk} x_{ijk}^2 \le 1\}$$

$$\|A\|_{\{12\}\{3\}} = \sup\{\sum_{ijk} a_{ijk} x_{ij} y_k : \sum_{ij} x_{ij}^2 \le 1, \sum_{k} y_k^2 \le 1\}$$

$$\|A\|_{\{1\}\{2\}\{3\}} = \sup\{\sum_{ijk} a_{ijk} x_i y_j z_k : \sum_{i} x_i^2 \le 1, \sum_{j} y_j^2 \le 1, \sum_{k} z_k^2 \le 1\}$$

Assume $\forall_{h\colon \mathbb{R}^n \to \mathbb{R} \text{ smooth}} \, \|h(X) - \mathbb{E}h(X)\|_p \le L\sqrt{p}\,\||\nabla h(X)|\|_p$

Theorem (Adamczak, W., 2013)

$f\colon \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^2$ such that for all $x \in \mathbb{R}^n$

$$\left\|D^2 f(x)\right\|_{\mathrm{HS}} \le A_{\mathrm{HS}}, \qquad \left\|D^2 f(x)\right\|_{\mathrm{op}} \le A_{\mathrm{op}}$$

Then for $p \ge 2$ and $t > 0$,

$$\|f(X) - \mathbb{E}f(X)\|_p \le C\left(L\sqrt{p}|\mathbb{E}\nabla f(X)| + L^2\sqrt{p}A_{\mathrm{HS}} + L^2 p A_{\mathrm{op}}\right)$$

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| \ge t) \le 2\exp\left(-c\left(\frac{t^2}{L^2|\mathbb{E}\nabla f(X)|^2 + L^4 A_{\mathrm{HS}}^2} \wedge \frac{t}{L^2 A_{\mathrm{op}}}\right)\right)$$

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Equivalently: $\|h(X) - \mathbb{E}h(X)\|_p \leq CL\,\|\langle \nabla h(X), G\rangle\|_p$ (*)

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Equivalently: $\|h(X) - \mathbb{E}h(X)\|_p \leq CL\,\|\langle \nabla h(X), G\rangle\|_p$ (*)

Apply to f + triangle inequality for L^p-norm:

$$\|f(X) - \mathbb{E}f(X)\|_p$$
$$\leq CL\,\|\langle \nabla f(X), G\rangle - \langle \mathbb{E}\nabla f(X), G\rangle\|_p + CL\,\|\langle \mathbb{E}\nabla f(X), G\rangle\|_p$$

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Equivalently: $\|h(X) - \mathbb{E}h(X)\|_p \leq CL\,\|\langle \nabla h(X), G\rangle\|_p$ (*)

Apply to f + triangle inequality for L^p-norm:

$$\|f(X) - \mathbb{E}f(X)\|_p$$
$$\leq CL\,\|\langle \nabla f(X), G\rangle - \langle \mathbb{E}\nabla f(X), G\rangle\|_p + CL\,\|\langle \mathbb{E}\nabla f(X), G\rangle\|_p$$
$$\approx CL\,\left\|\langle \nabla f(X), \bar{G}\rangle - \mathbb{E}_X\langle \nabla f(X), \bar{G}\rangle\right\|_p + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Equivalently: $\|h(X) - \mathbb{E}h(X)\|_p \leq CL\,\|\langle \nabla h(X), G\rangle\|_p$ (*)

Apply to f + triangle inequality for L^p-norm:

$$\|f(X) - \mathbb{E}f(X)\|_p$$

$$\leq CL\,\|\langle \nabla f(X), G\rangle - \langle \mathbb{E}\nabla f(X), G\rangle\|_p + CL\,\|\langle \mathbb{E}\nabla f(X), G\rangle\|_p$$

$$\approx CL\,\left\|\langle \nabla f(X), \bar{G}\rangle - \mathbb{E}_X\langle \nabla f(X), \bar{G}\rangle\right\|_p + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

use (*) conditionally on $\bar{G}$ for $h(x) = \langle \nabla f(x), \bar{G}\rangle$

$$\leq C^2L^2\left\|\langle D^2 f(X)\bar{G}, G\rangle\right\|_p + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

Outline of the proof

Assumption: $\|h(X) - \mathbb{E}h(X)\|_p \leq L\sqrt{p}\,\||\nabla h(X)|\|_p$

Equivalently: $\|h(X) - \mathbb{E}h(X)\|_p \leq CL\,\|\langle\nabla h(X), G\rangle\|_p$ (*)

Apply to f + triangle inequality for L^p-norm:

$$\|f(X) - \mathbb{E}f(X)\|_p$$

$$\leq CL\,\|\langle\nabla f(X), G\rangle - \langle\mathbb{E}\nabla f(X), G\rangle\|_p + CL\,\|\langle\mathbb{E}\nabla f(X), G\rangle\|_p$$

$$\approx CL\,\left\|\langle\nabla f(X), \bar{G}\rangle - \mathbb{E}_X\langle\nabla f(X), \bar{G}\rangle\right\|_p + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

use (*) conditionally on $\bar{G}$ for $h(x) = \langle\nabla f(x), \bar{G}\rangle$

$$\leq C^2L^2\left\|\langle D^2 f(X)\bar{G}, G\rangle\right\|_p + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

condition on X

$$\leq CL^2\left(p\left\|\|D^2 f(X)\|_{op}\right\|_p + \sqrt{p}\left\|\|D^2 f(X)\|_{HS}\right\|_p\right) + CL\sqrt{p}|\mathbb{E}\nabla f(X)|$$

The result — remarks

- **Higher order**: $f\colon \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^D$, iterate D times

$$\|f(X) - \mathbb{E}f(X)\|_p \qquad \text{and} \qquad \mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t)$$

involve tensor-like norms of

$$\mathbb{E}\nabla f(X), \quad \mathbb{E}D^2 f(X), \quad \ldots, \quad \mathbb{E}D^{d-1} f(X)$$

and $\sup_{x\in\mathbb{R}^n}$ of such norms of $D^D f(x)$

The result — remarks

- **Higher order**: $f\colon \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^D$, iterate D times

  $$\|f(X) - \mathbb{E}f(X)\|_p \qquad \text{and} \qquad \mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t)$$

  involve tensor-like norms of

  $$\mathbb{E}\nabla f(X), \quad \mathbb{E}D^2 f(X), \quad \ldots, \quad \mathbb{E}D^{d-1} f(X)$$

  and $\sup_{x\in\mathbb{R}^n}$ of such norms of $D^D f(x)$

  Main ingredient: estimates for moments and tails of $A(G^{(1)}, \ldots, G^{(d)})$ (R. Latała, 2006)

- **Higher order**: $f\colon \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^D$, iterate D times

$$\|f(X) - \mathbb{E}f(X)\|_p \qquad \text{and} \qquad \mathbb{P}(|f(X) - \mathbb{E}f(X)| \geq t)$$

  involve tensor-like norms of

$$\mathbb{E}\nabla f(X), \quad \mathbb{E}D^2 f(X), \quad \ldots, \quad \mathbb{E}D^{d-1} f(X)$$

  and $\sup_{x \in \mathbb{R}^n}$ of such norms of $D^D f(x)$

  Main ingredient: estimates for moments and tails of $A(G^{(1)}, \ldots, G^{(d)})$ (R. Latała, 2006)

- **Optimality**: f polynomial, X Gaussian vector

A — Wigner $n \times n$ matrix, entries satisfy (LSI)

$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ eigenvalues

$f \colon \mathbb{R} \to \mathbb{R}$ smooth function

$Z = \sum f(\lambda_i/\sqrt{n})$ — linear eigenvalue statistic

Theorem (Guionnet-Zeitouni)

$$\mathbb{P}(|Z - \mathbb{E}Z| \geq t) \leq 2\exp(-ct^2/\left\|f'\right\|_\infty^2)$$

A — Wigner $n \times n$ matrix, entries satisfy (LSI)

$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ eigenvalues

$f \colon \mathbb{R} \to \mathbb{R}$ smooth function

$Z = \sum f(\lambda_i/\sqrt{n})$ — linear eigenvalue statistic

Theorem (Guionnet-Zeitouni)

$$\mathbb{P}(|Z - \mathbb{E}Z| \geq t) \leq 2\exp(-ct^2/\left\|f'\right\|_\infty^2)$$

Theorem (Adamczak, W.)

$$\mathbb{P}(|Z - \mathbb{E}Z| \geq t) \leq 2\exp\left(-c\left(\frac{t^2}{\int f'^2\, d\rho + n^{-2/3}\left\|f''\right\|_\infty^2} \wedge \frac{nt}{\left\|f''\right\|_\infty}\right)\right)$$

Semicircle law: $d\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2}1_{(-2,2)}(x)\, dx$.

Applications I: random matrices

$$\mathbb{P}(|Z-\mathbb{E}Z| \geq t) \leq 2\exp\left(-c\left(\frac{t^2}{\int f'^2\,d\rho + n^{-2/3}\,\|f''\|_\infty^2} \wedge \frac{nt}{\|f''\|_\infty}\right)\right)$$

Hoffman-Wielandt: $(A, \|\cdot\|_{\mathrm{HS}}) \mapsto ((\lambda_1,\ldots,\lambda_n), |\cdot|_2)$ is 1-Lipschitz

$$F(\lambda_1,\ldots,\lambda_n) = \sum f(\lambda_i/\sqrt{n})$$

$$|\mathbb{E}\nabla F|^2 = \frac{1}{n}\sum\left(\mathbb{E}f'(\frac{\lambda_i}{\sqrt{n}})\right)^2 \leq \mathbb{E}\left(\frac{1}{n}\sum f'^2(\frac{\lambda_i}{\sqrt{n}})\right) = \int f'^2\,d\mu$$

μ ($\mathbb{E}$ sp. measure of $\frac{1}{\sqrt{n}}A$) $\simeq \rho$ (Bobkov, Götze, Tikhomirov, 2010)

$$\sup_x \|D^2F(x)\|_{\mathrm{HS}} \leq n^{-1/2}\|f''\|_\infty, \qquad \sup_x \|D^2F(x)\|_{\mathrm{op}} \leq n^{-1}\|f''\|_\infty$$

Sub-Gaussian coefficient

$$\|Y\|_{\psi_2} = \inf\{a > 0 \colon \mathbb{E}\exp(Y^2/a^2) \le 2\}$$

Sub-Gaussian coefficient

$$\|Y\|_{\psi_2} = \inf\{a > 0 \colon \mathbb{E}\exp(Y^2/a^2) \le 2\}$$

$X = (X_1, \ldots, X_n)$ independent, $\|X_i\|_{\psi_2} \le L$
$f \colon \mathbb{R}^n \to \mathbb{R}$ multivariate polynomial of degree 2

Polynomials in independent sub-Gaussian r.v.'s

Sub-Gaussian coefficient

$$\|Y\|_{\psi_2} = \inf\{a > 0 \colon \mathbb{E}\exp(Y^2/a^2) \le 2\}$$

$X = (X_1, \ldots, X_n)$ independent, $\|X_i\|_{\psi_2} \le L$

$f \colon \mathbb{R}^n \to \mathbb{R}$ multivariate polynomial of degree 2

$$\|f(X) - \mathbb{E}f(X)\|_p \le C\left(L\sqrt{p}|\mathbb{E}\nabla f(X)| + L^2\sqrt{p}\|D^2 f\|_{\mathrm{HS}} + L^2 p\|D^2 f\|_{\mathrm{op}}\right)$$

Sub-Gaussian coefficient

$$\|Y\|_{\psi_2} = \inf\{a > 0 \colon \mathbb{E}\exp(Y^2/a^2) \le 2\}$$

$X = (X_1, \ldots, X_n)$ independent, $\|X_i\|_{\psi_2} \le L$
$f \colon \mathbb{R}^n \to \mathbb{R}$ multivariate polynomial of degree 2

$$\|f(X) - \mathbb{E}f(X)\|_p \le C\left(L\sqrt{p}|\mathbb{E}\nabla f(X)| + L^2\sqrt{p}\|D^2 f\|_{\mathrm{HS}} + L^2 p\|D^2 f\|_{\mathrm{op}}\right)$$

Higher degree polynomial: estimates involve norms of

$\mathbb{E}D^d f(X)$ for $d = 1, \ldots, D$ with factors $p^{1/2}, p, p^{3/2}, \ldots, p^{D/2}$

Theorem (Adamczak, W.) degree=3

$X = (X_1, \ldots, X_n)$ independent, $\|X_i\|_{\psi_2} \leq L$

$f\colon \mathbb{R}^n \to \mathbb{R}$ polynomial of degree 3

$$\|f(X) - \mathbb{E}f(X)\|_p \leq C\Big(L\sqrt{p}|\mathbb{E}\nabla f(X)| + L^2\sqrt{p}\|\mathbb{E}D^2 f(X)\|_{\mathrm{HS}} + L^2 p\|\mathbb{E}D^2 f(X)\|_{\mathrm{op}} + L^3\sqrt{p}\|D^3 f\|_{\{123\}} + L^3 p\|D^3 f\|_{\{12\}\{3\}} + L^3 p^{3/2}\|D^3 f\|_{\{1\}\{2\}\{3\}}\Big)$$

$$\|A\|_{\{123\}} = \Big(\sum_{ijk} a_{ijk}^2\Big)^{1/2} = \sup\{\sum_{ijk} a_{ijk}x_{ijk} \colon \sum_{ijk} x_{ijk}^2 \leq 1\}$$

$$\|A\|_{\{12\}\{3\}} = \sup\{\sum_{ijk} a_{ijk}x_{ij}y_k \colon \sum_{ij} x_{ij}^2 \leq 1, \sum_k y_k^2 \leq 1\}$$

$$\|A\|_{\{1\}\{2\}\{3\}} = \sup\{\sum_{ijk} a_{ijk}x_i y_j z_k \colon \sum_i x_i^2 \leq 1, \sum_j y_j^2 \leq 1, \sum_k z_k^2 \leq 1\}$$

Outline of the proof — tetrahedral case

$$f(X) = \sum_{d=0}^{D} \sum_{i_1<\ldots<i_d} a^{(d)}_{i_1\ldots i_d} X_{i_1}\cdots X_{i_d} \qquad \text{(tetrahedral)}$$

$$f(X) = \sum_{d=0}^{D} \frac{1}{d!} D^d f(\mathbb{E}X)(X-\mathbb{E}X,\ldots,X-\mathbb{E}X)$$

$$f(\mathbb{E}X) = \mathbb{E}f(X) \quad \text{and} \quad D^d f(\mathbb{E}X) = \mathbb{E}D^d f(X)$$

$$D^d f(\mathbb{E}X)(X-\mathbb{E}X,\ldots,X-\mathbb{E}X) \quad \text{(tetrahedral)}$$

$$\downarrow \text{(decoupling ineq. + symmetrization)}$$

$$D^d f(\mathbb{E}X)(\varepsilon^{(1)}X^{(1)},\ldots,\varepsilon^{(d)}X^{(d)})$$

$$\downarrow \text{(comparison with Gaussians)}$$

$$L^d D^d f(\mathbb{E}X)(G^{(1)},\ldots,G^{(d)})$$

and use Latała's estimates for Gaussian chaoses

$G(n,p)$ — n vertices, edges independent with probab. p

H — fixed graph

Y_H — # of copies of H in $G(n,p)$

$$\mathbb{P}(Y_H \geq (1+\varepsilon)\mathbb{E}Y_H) \leq ?$$

Application II: random graphs

$G(n,p)$ — n vertices, edges independent with probab. p

H — fixed graph

Y_H — # of copies of H in $G(n,p)$

$$\mathbb{P}(Y_H \geq (1+\varepsilon)\mathbb{E}Y_H) \leq ?$$

$X_1, \ldots X_{\binom{n}{2}}$ independent 0-1 r.v's

$Y_H = \sum X_{i_1} \cdots X_{i_{e(H)}}$ polynomial of degree $e(H)$ in X_i's

$$\|X_i\|_{\psi_2} \leq \frac{1}{\sqrt{\log(1/p)}}$$

Theorem (Janson, Oleszkiewicz, Ruciński, 2004)

For a certain function $M_H^*(n,p)$

$$e^{-C_H(\varepsilon)M_H^*(n,p)\log(1/p)} \le \mathbb{P}(Y_H \ge (1+\varepsilon)\mathbb{E}Y_H) \le e^{-c_H(\varepsilon)M_H^*(n,p)}$$

$$M_{C_k}^*(n,p) = \Theta(n^2p^2), \qquad M_{K_k}^*(n,p) = \Theta(n^2p^{k-1})$$

(whenever $\mathbb{P}(Y_H > 0)$ is not negligible; for $H = C_k$ it is $p \ge 1/n$)

Application II: random graphs

Theorem (Janson, Oleszkiewicz, Ruciński, 2004)

For a certain function $M_H^*(n,p)$

$$e^{-C_H(\varepsilon)M_H^*(n,p)\log(1/p)} \le \mathbb{P}(Y_H \ge (1+\varepsilon)\mathbb{E}Y_H) \le e^{-c_H(\varepsilon)M_H^*(n,p)}$$

$$M_{C_k}^*(n,p) = \Theta(n^2p^2), \qquad M_{K_k}^*(n,p) = \Theta(n^2p^{k-1})$$

(whenever $\mathbb{P}(Y_H > 0)$ is not negligible; for $H = C_k$ it is $p \ge 1/n$)

Theorem (Chatterjee, DeMarco-Kahn, 2012)

For $H = K_k$ and $p \ge n^{-2/(k-1)}\mathrm{poly}(\log n)$,

$$\mathbb{P}(Y_H \ge (1+\varepsilon)\mathbb{E}Y_H) \le e^{-c_k(\varepsilon)n^2p^{k-1}\log(1/p)}$$

Application II: random graphs

Theorem (Adamczak, W.)

For $H = C_k$ and $p \geq n^{-\frac{k-2}{2(k-1)}} \log^{-1/2} n$,

$$\mathbb{P}(Y_H \geq (1+\varepsilon)\mathbb{E} Y_H) \leq e^{-c_k(\varepsilon) n^2 p^2 \log(1/p)}$$

$$Y_H = f(X_1, \ldots, X_{\binom{n}{2}}) = \sum_{\substack{\text{edges } i_1,\ldots,i_k \\ \text{forms a cycle}}} X_{i_1} \cdots X_{i_k}$$

The proof involves estimates for norms of $\mathbb{E} D^d f(X)$, $d = 1, \ldots, k$ (can be done for any graph H)

Thank you!