Hanson-Wright inequality in Banach spaces

Rafał Meller (based on joint work with R. Adamczak and R. Latała)

University of Warsaw

Probability and Analysis, Będlewo May 2019

Notation and convention

In this talk, we use the letter C to denote universal, nonnegative constant which may differ at each occurrence. So using this convention we may write

$$2C \leq C \text{ or } \mathbb{P}(|X| \geq Ct) \leq e^{C^{-1}t^2}.$$

We write $C(\alpha)$ if the constant may depend on some parameter α. We write $a \sim b$ ($a \sim^{\alpha} b$) if there exists C ($C(\alpha)$) such that $a/C \leq b \leq aC$ ($a/C(\alpha) \leq b \leq aC(\alpha)$). For example

$$1 \sim 2,\ t^2 \sim 2t^2,\ e^{x^2} \sim e^{x^2} + x^8.$$

Classical Hanson-Wright inequality

Definition

We say that a random variable X is α-subgaussian if for every $t > 0$, $\mathbb{P}(|X| \geq t) \leq 2\exp\left(-t^2/(2\alpha^2)\right)$.

Let us consider a sequence $X_1, X_2, \ldots$ of independent, mean zero and α-subgaussian random variables. The classical Hanson-Wright inequality states that for any real valued matrix $A = (a_{ij})_{ij \leq n}$

$$\mathbb{P}\left(\left|\sum_{ij} a_{ij}(X_iX_j - \mathbb{E}X_iX_j)\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{C\alpha^4 \|A\|_{HS}^2} - \frac{t}{C\alpha^2 \|A\|_{op}}\right)$$

where $\|A\|_{HS}^2 = \sum_{ij} a_{ij}^2$, $\|A\|_{op} = \sup_{x,y \in B_2^n} \sum_{ij} a_{ij} x_i y_j$.

Problems with Classical Hanson-Wright inequality

In many problems one need to analyze not a single quadratic form but a supremum of a collection of them i.e. expression of the form

$$\mathbb{P}\left(\sup_{k\leq n}\left|\sum_{ij} a_{ij}^k(X_iX_j - \mathbb{E}X_iX_j)\right| \geq t\right) \tag{1}$$

where $A^1 = (a_{ij}^1)_{ij}, A^2 = (a_{ij}^2)_{ij}, \ldots$ is a sequence of real-valued matrices. Equivalently one may need to estimate from the above the expression

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_iX_j - \mathbb{E}X_iX_j)\right\| \geq t\right), \tag{2}$$

where $A = (a_{ij})_{ij\leq n}$ is a matrix with values in a Banach space $(F, \|\cdot\|)$.

Moment estimates imply tail estimates

We want to find an upper bound for

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_iX_j - \mathbb{E}X_iX_j)\right\| \geq t\right) = \mathbb{P}\left(S \geq t\right),$$

where $A = (a_{ij})_{ij\leq n}$ is a matrix with values in a Banach space $(F, \|\cdot\|)$. A naive idea (which luckily is enough) is to use Chebyshev's inequality:

$$\mathbb{P}(S \geq t) \leq (\|S\|_p / t)^p \text{ for any } p \geq 1.$$

So we need to estimate from the above $\|S\|_p$. Standard arguments (decoupling, symmetrization and the contraction principle) yield

$$\|S\|_p \leq C\alpha^2 \left\|\sum_{ij} a_{ij}(g_ig_j - \delta_{ij})\right\|_p .$$

Moments of Gaussian quadratic forms

Our goal is to find upper bounds (and preferably two-sided bounds) for moments of $\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|$ (recall that $(a_{ij})_{ij}$ are from Banach space). Some results exist in the literature.

Theorem (C. Borell; M. A. Arcones and E. Giné ; M. Ledoux and M. Talagrand)

Let $(F, \|\cdot\|)$ be a Banach space and A be a symmetric, F-valued matrix. Then, for any $p \geq 1$ we have

$$\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_p \sim \mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\| + \sqrt{p}\,\mathbb{E}\sup_{x\in B_2^n}\left\|\sum_{ij} a_{ij} g_i x_j\right\| + p\sup_{x,y\in B_2^n}\left\|\sum_{ij} a_{ij} x_i y_j\right\|.$$

Problems in L_q spaces

The previous Theorems yields (for $t > C\mathbb{E}\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\|$)

$$\mathbb{P}(\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\| \geq C\alpha^2 t)$$

$$\leq 2\exp\left(-\frac{t^2}{\left(\mathbb{E}\sup_{x\in B_2^n}\left\|\sum_{ij} a_{ij} g_i x_j\right\|\right)^2} - \frac{t}{\sup_{x,y\in B_2^n}\left\|\sum_{ij} a_{ij} x_i y_j\right\|}\right)$$

Consider $(F, \|\cdot\|) = (l_q, \|\cdot\|_q)$. Then $a_{ij} = (a_{ij}^k)_{k\geq 1}$ and

$$\mathbb{E}\sup_{x\in B_2^n}\left\|\sum_{ij} a_{ij} g_i x_j\right\| = \mathbb{E}\sup_{x\in B_2^n}\sqrt[q]{\sum_k \left|\sum_{ij} a_{ij}^k g_i x_j\right|^q}$$

It is nontrivial to estimate the last expression (even in the case $q = 2$).

Theorem (C. Borell; M. A. Arcones and E. Giné ; M. Ledoux and M. Talagrand)

Let $(F, \|\cdot\|)$ be a Banach space and A be a symmetric, F-valued matrix. Then, for any $p \geq 1$ we have

$$\left\| \left\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij}) \right\| \right\|_p \sim \mathbb{E} \left\| \sum_{ij} a_{ij}(g_i g_j - \delta_{ij}) \right\| + \sqrt{p}\,\mathbb{E} \sup_{x \in B_2^n} \left\| \sum_{ij} a_{ij} g_i x_j \right\| + p \sup_{x,y \in B_2^n} \left\| \sum_{ij} a_{ij} x_i y_j \right\| .$$

Moments of Gaussian quadratic forms

Theorem (R. Adamczak, R. Latała, R. Meller)

Under the assumption of the previous theorem we have

$$\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_p \lesssim \mathbb{E}\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\| + \mathbb{E}\left\|\sum_{\mathbf{i}\neq\mathbf{j}} \mathbf{a_{ij} g_{ij}}\right\|$$

$$+\sqrt{p}\sup_{\mathbf{x}\in\mathbf{B_2^n}} \mathbb{E}\left\|\sum_{ij} a_{ij} g_i x_j\right\| + \sqrt{\mathbf{p}}\sup_{\mathbf{x}\in\mathbf{B_2^{n^2}}}\left\|\sum_{\mathbf{ij}} \mathbf{a_{ij} x_{ij}}\right\|$$

$$+ p \sup_{x,y\in B_2^n}\left\|\sum_{ij} a_{ij} x_i y_j\right\|.$$

This inequality cannot be reversed. To see this, consider $p = 1$ and the Banach space $(M_{n\times n}(\mathbb{R}), \|\cdot\|_*)$, where $\|A\|_* = \sup_{\|T\|_{op}=1, T\in M_{n\times n}} \sum a_{ij} t_{ij}$.

Hanson-Wright inequality in Banach spaces

Theorem

Let $X_1, X_2, \ldots$ be independent, mean-zero, α-subgaussian random variables. Then for any matrix $A = (a_{ij})_{ij}$ with values in $(F, \|\cdot\|)$ and any $t \geq C\alpha^2(\mathbb{E}\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\| + \mathbb{E}\|\sum_{i\neq j} a_{ij} g_{ij}\|)$ we have

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_i X_j - \mathbb{E}X_i X_j)\right\| \geq t\right) \leq 2\exp\left(-\frac{t^2}{C\alpha^4 U^2} - \frac{t}{C\alpha^2 V}\right),$$

$$U = \sup_{x\in B_2^n} \mathbb{E}\left\|\sum_{ij} a_{ij} g_i x_j\right\| + \sup_{x\in B_2^{n^2}} \left\|\sum_{ij} a_{ij} x_{ij}\right\|$$

$$V = \sup_{x,y\in B_2^n} \left\|\sum_{ij} a_{ij} x_i y_j\right\|.$$

Gaussian quadratic forms in L_q spaces.

Theorem

In the L_q spaces the following holds

$$\left\|\sum_{ij} a_{ij}(g_i g_j - \delta_{ij})\right\|_p \sim^q \left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q} + \sqrt{p} \sup_{x \in B_2^{n^2}} \left\|\sum_{ij} a_{ij} x_{ij}\right\|_{L_q}$$
$$+ \sqrt{p} \sup_{x \in B_2^n} \left\|\sqrt{\sum_i \left(\sum_j a_{ij} x_j\right)^2}\right\|_{L_q} + p \sup_{x,y \in B_2^n} \left\|\sum_{ij} a_{ij} x_i y_j\right\|_{L_q}.$$

The reason why in L_q space we have such an simplification is the following

$$\mathbb{E}\left\|\sum_{ij} a_{ij} g_{ij}\right\|_{L_q} \leq Cq\mathbb{E}\left\|\sum_{ij} a_{ij} g_i g_j\right\|_{L_q}.$$

Hanson-Wright inequality in L_q spaces

Theorem

Let $X_1, X_2, \ldots$ be independent, mean-zero, α-subgaussian random variables. Then for any matrix $A = (a_{ij})_{ij}$ with values in $(L_q(T), \|\cdot\|_{L_q})$ and any $t \geq C\alpha^2 q \left\|\sqrt{\sum_{ij} a_{ij}^2}\right\|_{L_q}$ we have

$$\mathbb{P}\left(\left\|\sum_{ij} a_{ij}(X_i X_j - \mathbb{E}X_i X_j)\right\|_{L_q} \geq t\right) \leq 2\exp\left(-\frac{t^2}{C\alpha^4 q U^2} - \frac{t}{C\alpha^2 V}\right),$$

$$U = \sup_{x \in B_2^{n^2}} \left\|\sum_{ij} a_{ij} x_{ij}\right\|_{L_q} + \sup_{x \in B_2^n} \left\|\sqrt{\sum_i \left(\sum_j a_{ij} x_j\right)^2}\right\|_{L_q}$$

$$V = \sup_{x,y \in B_2^n} \left\|\sum_{ij} a_{ij} x_i y_j\right\|_{L_q}.$$