# Seminar 6

High-dimensional Probability and Statistics

*HSE University, spring 2023*

### Talagrand's inequality

**Theorem 1** (Talagrand's inequality, in this form [2]). *Let $X_1, \dots, X_n$ be independent random variables and let $\mathcal{F}$ be a separable family of zero-mean functions, taking their values in $[-1, 1]^n$. Denote*

$$Z = \sup_{f \in \mathcal{F}} \sum_{i=1}^n f_i(X_i), \quad f = (f_1, \dots, f_n).$$

*Then, for any $t > 0$, it holds that*

$$\mathbb{P}\left(Z \geqslant \mathbb{E} Z + t\right) \leqslant e^{-\frac{t^2}{2V + 3t}},$$

*where*

$$V = 2\,\mathbb{E} Z + \sigma_{\mathcal{F}}^2 \quad \text{and} \quad \sigma_{\mathcal{F}}^2 = \sup_{f \in \mathcal{F}} \mathrm{Var}\left(\sum_{i=1}^n f_i(X_i)\right).$$

**Problem 1.** Let a random matrix $X \in \mathbb{R}^{m \times n}$ have independent zero-mean entries, taking their values in $[-1, 1]$. Show that, for any $\delta \in (0, 1)$, with probability at least $1 - \delta$, it holds that

$$\|X\| \leqslant \mathbb{E}\|X\| + \sqrt{(4\mathbb{E}\|X\| + 2)\log(1/\delta)} + 3\log(1/\delta).$$

*Remark* 2. In [1], the author extended the Talagrand concentration inequality to the unbounded case.

### Some properties of Gaussian random variables

**Theorem 3** (Gaussian concentration of Lipschitz functions). *Let $X_1, \dots, X_n$ be i.i.d. random variables with the Gaussian distribution $\mathcal{N}(0, 1)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function (with respect to the Euclidean norm), that is,*

$$|f(x) - f(y)| \leqslant L\|x - y\| \quad \text{for all } x, y \in \mathbb{R}^n.$$

*Then, for any $t > 0$, it holds that*

$$\mathbb{P}\left(f(X_1, \dots, X_n) - \mathbb{E} f(X_1, \dots, X_n) > t\right) \leqslant e^{-\frac{t^2}{2L^2}}.$$

*Remark* 4. The proof follows from the log-Sobolev inequality and the entropy method.

**Theorem 5** (Sudakov-Fernique inequality). *Let* $X = (X_1, \ldots, X_n)^\top$ *and* $Y = (Y_1, \ldots, Y_n)^\top$ *be centered Gaussian vectors, such that*

$$\mathbb{E}(X_i - X_j)^2 \leqslant \mathbb{E}(Y_i - Y_j)^2 \quad \textit{for all } i, j \in \{1, \ldots, n\}.$$

*Then it holds that*

$$\mathbb{E} \max_{1 \leqslant i \leqslant n} X_i \leqslant \mathbb{E} \max_{1 \leqslant i \leqslant n} Y_i.$$

**Problem 2 (Gaussian contraction).** Let $\xi_1, \ldots, \xi_n$ be i.i.d. standard Gaussian random variables and let $\varphi_1, \ldots, \varphi_n$ be a collection of 1-Lipschitz univariate functions, such that $\varphi_i(0) = 0$ for all $i \in \{1, \ldots, n\}$. Prove that

$$\mathbb{E} \sup_{\theta \in \Theta} \sum_{i=1}^n \xi_i \varphi_i(\theta_i) \leqslant \mathbb{E} \sup_{\theta \in \Theta} \sum_{i=1}^n \xi_i \theta_i \quad \text{for any bounded set } \Theta \subset \mathbb{R}^n.$$

**Problem 3 (Sudakov minoration).** Let $X_\theta, \theta \in \Theta$ be a zero-mean Gaussian process defined on the non-empty set $\Theta$. Let $\mathcal{M}(\Theta, \delta)$ be the packing number of $\Theta$ with respect to the metric $\rho(\theta, \theta') = \big(\mathrm{Var}(X_\theta - X_{\theta'})\big)^{1/2}$. Then it holds that

$$\mathbb{E} \sup_{\theta \in \Theta} X_\theta \gtrsim \sup_{\delta > 0} \delta \sqrt{\log \mathcal{M}(\Theta, \delta)}.$$

## References


[1] R. Adamczak. A tail inequality for suprema of unbounded empirical processes with applications to Markov chains. *Electronic Journal of Probability*, 13:no. 34, 1000–1034, 2008.

[2] T. Klein and E. Rio. Concentration around the mean for maxima of empirical processes. *The Annals of Probability*, 33(3):1060 – 1077, 2005.