# S&DS 602: Homework 3

Due Wednesday, September 18 at 2PM, via Gradescope

1. (Vershynin 6.4.6) Let $X_1, \ldots, X_n$ be independent subgaussian random variables with $\mathbb{E}X_i = 0$, and let $\varepsilon_1, \ldots, \varepsilon_n$ be i.i.d. Rademacher variables independent of $X_1, \ldots, X_n$. Show that for universal constants $C, c > 0$,

$$c \left\| \sum_{i=1}^n \varepsilon_i X_i \right\|_{\psi_2} \le \left\| \sum_{i=1}^n X_i \right\|_{\psi_2} \le C \left\| \sum_{i=1}^n \varepsilon_i X_i \right\|_{\psi_2}.$$

2. (Vershynin 6.3.4) Let $E$ be a subspace of $\mathbb{R}^d$ of dimension $k$. Let $X = (X_1, \ldots, X_d) \in \mathbb{R}^d$ have independent $\sigma^2$-subgaussian coordinates satisfying $\mathbb{E}X_i = 0$ and $\operatorname{Var} X_i = 1$, and let $\operatorname{dist}(X, E)$ be the Euclidean distance from $X$ to $E$. Show that $\mathbb{E}\operatorname{dist}(X, E)^2 = d - k$, and for universal constants $C, c > 0$,

$$\mathbb{P}\Big[\big|d(X, E) - \sqrt{d-k}\big| > t\Big] \le Ce^{-ct^2/\sigma^4} \text{ for all } t \ge 0.$$

3. (Vershynin 6.2.6) A random vector $X \in \mathbb{R}^d$ is $\sigma^2$-*subgaussian* if $v^\top X$ is $\sigma^2$-subgaussian for every (deterministic) unit vector $v \in \mathbb{R}^d$.

   Let $X \in \mathbb{R}^d$ be $\sigma^2$-subgaussian with $\mathbb{E}X = 0$, and let $G \in \mathbb{R}^d$ have i.i.d. $\mathcal{N}(0, 1)$ coordinates. For any $B \in \mathbb{R}^{d' \times d}$ and universal constants $C, C', c > 0$, show that

$$\mathbb{E}e^{\lambda^2 \|BX\|_2^2} \le \mathbb{E}e^{C\lambda^2\sigma^2\|BG\|_2^2} \le \mathbb{E}e^{C'\lambda^2\sigma^2\|B\|_F^2} \text{ for all } |\lambda| \le \frac{c}{\sigma\|B\|_{\mathrm{op}}}.$$

4. (Vershynin 6.2.7) Let $X_1, \ldots, X_n$ be independent $\sigma^2$-subgaussian random vectors in $\mathbb{R}^d$ with $\mathbb{E}X_i = 0$. For any $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ and some universal constants $C, c > 0$, show that

$$\mathbb{P}\left[\left|\sum_{1 \le i \ne j \le n} a_{ij} X_i^\top X_j\right| \ge t\right] \le C \exp\left(-c \min\left(\frac{t^2}{\sigma^4 d \|A\|_F^2}, \frac{t}{\sigma^2\|A\|_{\mathrm{op}}}\right)\right).$$