Math 608D, Assignment 1: Due Friday, Feb 2

1. We show that Gaussian concentration extends, in an important special case, to sub-Gaussian variables. Let X be a vector with independent random entries $X_1, X_2, \ldots, X_n$. Suppose for each i

$$\boldsymbol{E} X_i = 0, \qquad \boldsymbol{E} X_i^2 = 1, \qquad \|X_i\|_{\Psi_2} \leq 10.$$

Show that the following hold:

   (a) $\| \|X\|_2 \|_{\Psi_2} \leq C\sqrt{n}$. (This part is not so exciting.)

   (b) $\| \|X\|_2 - \sqrt{n} \|_{\Psi_2} \leq C$. (This part matches Gaussian concentration, is tricky to prove, and was the first step in a publication that arose from this class last year.)

   (c) If you replace the assumption $\|X_i\|_{\Psi_2} \leq 10$ with $\|X_i\|_{\Psi_2} \leq \alpha$ for some $\alpha > 0$, then how well can you bound $\| \|X\|_2 - \sqrt{n} \|_{\Psi_2}$? (This part is open-ended, a tight result could possibly lead to a good conference paper.)

2. Prove the following version of Bernstein inequality:

Theorem 0.1 (Bernstein inequality for bounded random variables) *Let $X_1, X_2, \ldots, X_N$ be independent, mean-zero, random variables which are all uniformly bounded by a positive scalar M, i.e., $\|X_i\|_\infty \leq M$. Then for any $t > 0$,*

$$P\left(\left| \sum_{i=1}^{N} X_i \right| \geq t \right) \leq 2 \exp\left(-C \min\left(\frac{t^2}{\sum_i \boldsymbol{E}[X_i^2]}, \frac{t}{M} \right) \right).$$

Remark 0.2 *This version is quite useful for bounded random variables which have standard deviation much lower than M. As seen in class for sums of sub-exponential random variables, there is a combination of sub-Gaussian behavior and sub-exponential behavior in the bound.*

3. As can be seen from the above version of Bernstein inequality, being able to control a random variable in multiple ways can lead to better tail bounds. Now suppose you have a sequence of mean-zero independent random variables $X_1, X_2, \ldots X_N$ having the following properties:

   (a) $\|X_i\|_2 = \sigma$,

   (b) $\|X_i\|_\infty = M$,

   (c) $\|X_i\|_{\Psi_1} = a_1$,

   (d) $\|X_i\|_{\Psi_2} = a_2$.

   (a) What can you say about the ordering of σ, M, a_1, a_2? (Which is largest? etc.)

   (b) In the same spirit as Bernstein inequality above, can you give a tail bound using some subset of the above properties? To be interesting, this bound should be a significant improvement over bounds that can be made using a single property, at least for some values of t. *Note: This is a very open-ended problem. I'm not sure if it is easy, hard, or impossible. We'll see what comes out!*