STA 6448 Homework 2

Due February 13

20 pts for each problem below (in total, 100pts):

1. For any $x_1^n = (x_1, \ldots, x_n)$, let $D(x_1^n) = \sup_{f \in \mathcal{F}} \sqrt{\frac{\sum_{i=1}^n f^2(x_i)}{n}}$ denote the ℓ_2 radius of $\mathcal{F}(x_1^n)/\sqrt{n}$ (under the notation of lecture note 7). Show

$$\mathcal{R}(\mathcal{F}(x_1^n)/n) = \mathbb{E}_\varepsilon \Big[\sup_{f \in \mathcal{F}} \Big| \frac{1}{n} \sum_{i=1}^n \varepsilon_i \, f(x_i) \Big| \Big] \le D(x_1^n) \sqrt{\frac{2 \log(2 \, \Pi_{\mathcal{F}}(n))}{n}}.$$

2. (Gaussian contraction inequality) For each j, let $\phi_j : \mathbb{R} \to \mathbb{R}$ to be a contraction mapping, meaning that ϕ_j is 1-Lipschitz and satisfies $\phi_j(0) = 0$. Given a set $\mathcal{T} \subset \mathbb{R}^d$, consider the set

$$\phi(\mathcal{T}) = \Big\{ \big(\phi_1(\theta_1), \ldots, \phi_d(\theta_d) \big) \, \Big| \, \theta \in \mathcal{T} \Big\}.$$

Prove the Gaussian contraction inequality $\mathcal{G}(\phi(\mathcal{T})) \le \mathcal{G}(\mathcal{T})$.

3. Prove the following relation between Rademacher complexity and Gaussian complexity: for $\mathcal{T} \subset \mathbb{R}^d$,

$$\mathcal{R}(\mathcal{T}) \le \sqrt{\frac{\pi}{2}} \mathcal{R}(\mathcal{G}) \le c \sqrt{\log d} \, \mathcal{R}(\mathcal{T}),$$

where $c > 0$ is some universal constant. (Hint: in proving the second bound, you may assume the Rademacher contraction inequality: $\mathcal{R}(\phi(\mathcal{T})) \le \mathcal{R}(\mathcal{T})$ for any contraction mapping ϕ defined in Problem 2.)

4. Prove the following bounds about Gaussian complexities:
a) Consider the set $\mathcal{S}^d(s) = \{\theta \in \mathbb{R}^d : \|\theta\|_0 \le s, \, \|\theta\|_2 \le 1\}$ corresponding to all s-sparse vectors contained within the Euclidean unit ball. Then there is some universal constant $c > 0$ such that

$$\mathcal{G}(\mathcal{S}^d(s)) \le c \sqrt{s \, \log \frac{ed}{s}}.$$

b) Consider the ℓ_q-ball of unit radius $\mathcal{B}_q^d = \{\theta \in \mathbb{R}^d : \|\theta\|_q = \big(\sum_{j=1}^d |\theta_j|^q \big)^{1/q} \le 1\}$, where $q \in (1, \infty)$. Then there exists some constant c_q only dependent of q such that

$$\sqrt{\frac{2}{\pi}} \le \frac{\mathcal{G}(\mathcal{B}_q^d)}{d^{1-1/q}} \le c_q.$$

5. (Concentration of Gaussian suprema) Let $\{X_\theta : \theta \in \mathcal{T}\}$ be a Gaussian process over a countable set $\mathcal{T}$, and define $Z = \sup_{\theta \in \mathcal{T}} X_\theta$. Prove that

$$\mathbb{P}\big[|Z - \mathbb{E}[Z]| \geq t\big] \leq 2e^{-\frac{t^2}{2\sigma^2}}, \quad \text{for all } t > 0,$$

where $\sigma^2 = \sup_{\theta \in \mathcal{T}} \mathrm{Var}(X_\theta)$ is the maximal variance of the process. (Hint: first prove the inequality for a finite collection of Gaussian variables.)