MATH 547: HOMEWORK 4
DUE ON: MONDAY, NOVEMBER 4, 9AM.

Problem 1: more on the symmetrization inequality, 30 points:

(1) (De-symmetrization inequality, 15pts) Assume that $X_1, \ldots, X_n \in S$ are i.i.d. random variables. Let $\mathcal{F}$ be a class of functions $\mathcal{F} = \{f : S \mapsto \mathbb{R}\}$, and set $Pf := \mathbb{E}f(X)$. Let $\varepsilon_1, \ldots, \varepsilon_n$ be i.i.d. Rademacher random variables (i.e., random signs). Following the same steps as we used in the proof of the symmetrization inequality, prove that

$$\mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{j=1}^{n} (f(X_i) - Pf) \right| \geq \frac{1}{2} \mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j (f(X_j) - Pf) \right|$$

$$\geq \frac{1}{2} \mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{j=1}^{n} \varepsilon_j f(X_j) \right| - \frac{1}{2\sqrt{n}} \sup_{f \in \mathcal{F}} Pf.$$

(2) (Symmetrization with Gaussian weights, 15pts) Let $g_1, \ldots, g_n$ be i.i.d. $N(0,1)$ random variables. Prove the following version of the symmetrization inequality:

$$\mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{j=1}^{n} (f(X_i) - Pf) \right| \leq 2\sqrt{\frac{\pi}{2}} \mathbb{E} \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{j=1}^{n} g_i f(X_i) \right|.$$

[Hint: represent $g_i = \varepsilon_i |g_i|$, where ε_i is independent of $|g_i|$, and use Jensen's inequality]

Problem 2: facts about VC dimension, 30 points:

(a) (10pts) Let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ be k classes of sets such that $3 \leq d = \max_{j=1,\ldots,k} \mathrm{VC}(\mathcal{C}_j) < \infty$. Prove that

$$\mathrm{VC}\left(\bigcup_{j=1}^{k} \mathcal{C}_j \right) \leq 4d \log(2d) + 2\log(k)$$

If you prove the bound with different constants in place of 2 and 4, you will also receive full credit.

(b) (10pts) Prove directly, without using Dudley's theorem, that the VC dimension of the set of all halfspaces in $\mathbb{R}^d$ is equal to $d+1$.

[Hint: use Radon's theorem which states that any collection of $d+2$ points in $\mathbb{R}^d$ can be partitioned into 2 (disjoint) subsets such that the convex hulls of these subsets intersect]

(c) (10pts) Use part (b) to show that the VC dimension of balls in $\mathbb{R}^d$ is also $d+1$. *[Hint: suppose that the set of points is shattered by the spheres. Then for any partition of this set into 2 disjoint subsets, there exist two spheres such that each contains only the points only from one of these disjoint subsets. Show that there must also be a hyperplane such that the corresponding half-spaces have the same property. This will give you an upper bound for the VC dimension]*

(d) (*bonus + 20pts)

(1) (10 bonus pts) Find an upper and lower bounds for VC dimension of ellipsoids (an ellipsoid is a set of points $x \in \mathbb{R}^d$ that satisfy $(x - x_0)^T A (x - x_0) \leq 1$, where $x_0 \in \mathbb{R}^d$ is fixed and $A \in \mathbb{R}^{d \times d}$ is a positive definite matrix.

(2) (10 bonus pts) Find an upper and lower bounds for VC dimension spherical cones in $\mathbb{R}^d$. A spherical cone is a set

$$C_u(t;b) = \Big\{x \in \mathbb{R}^d : \ \langle x-b, u\rangle \geq t\|x-b\|_2\},\ \|u\|_2 = 1,\ b \in \mathbb{R}^d,\ t \in (0,1]\Big\}.$$

Problem 3: Rademacher complexity of a ball, 15 points:

Let $\mathcal{F}$ be a d-dimensional space of functions, and $B(r) = \{f \in \mathcal{F} : \|f\|_{L_2(P)} = \mathbb{E}f^2(X) \leq r\}$. Show that

$$\mathbb{E} \sup_{f \in B(r)} |R_n(f)| \leq r\sqrt{\frac{d}{n}},$$

where $R_n(f) = \frac{1}{n}\sum_{j=1}^n \varepsilon_j f(X_j)$ is the Rademacher process.

[Hint: use linearity of the Rademacher process together with the fact that any $f \in \mathcal{F}$ can be written as $f = \sum_{j=1}^d \alpha_j(f)\phi_j$ where $\phi_1, \ldots, \phi_d$ are the basis functions; you do not need to use generic chaining or entropy integral bounds in this problem!]