STOR 655 Homework 12

1. Read the statement and proof of the basic Gaussian comparison lemma in the online notes. Fill in the necessary details for equation (1.3), which makes use of Gaussian integration-by-parts. Write out a proof of the Gaussian comparison lemma in the case $n = 1$, following the proof of the general result. As $n = 1$, you will not need the conditioning argument, but you will need to exchange the operations of expectation and differentiation. Provide sufficient conditions on G and its derivatives to justify this exchange of limit operations, and show as carefully as you can why these conditions are sufficient. (You need not worry about finding the most general sufficient conditions; any reasonable conditions will do.)

2. Carefully verify that the Gaussian comparison lemma holds for the quadratic function $G(x) = x^t A x$, where A is a symmetric matrix.

3. Let $V \subseteq \mathbb{R}^n$ be a finite set of vectors $v = (v_1, \ldots, v_n)^t$ with $L = \max_{v \in V} ||v||_2$, and let $\varepsilon_1, \ldots, \varepsilon_n$ be independent Rademacher (sign) variables.

   (a) Use Hoeffding's MGF inequality to bound the moment generating functions of the random variables $\sum_{i=1}^n \varepsilon_i v_i$ in terms of the constant L.

   (b) Show that

$$\mathbb{E}\left[\max_{v \in V} \sum_{i=1}^n \varepsilon_i v_i\right] \;\leq\; \sqrt{2L^2 \log |V|}$$

4. Let $\mathcal{X}$ be a set, and let $\mathcal{C} \subseteq 2^{\mathcal{X}}$ be a (possibly infinite) family of sets $C \subseteq \mathcal{X}$. Let $X_1, \ldots, X_n \in \mathcal{X}$ be i.i.d. with distribution μ and define

$$\Delta(X_1^n) \;=\; \sup_{C \in \mathcal{C}} \left| n^{-1} \sum_{i=1}^n I(X_i \in C) \;-\; \mu(C) \right|$$

   (a) Use the Symmetrization inequality to establish that

$$\mathbb{E}\Delta(X_1^n) \;\leq\; 2\,\mathbb{E} \sup_{C \in \mathcal{C}} \left| n^{-1} \sum_{i=1}^n \varepsilon_i \, I(X_i \in C) \right| \qquad (1)$$

   where $\varepsilon_1, \ldots, \varepsilon_n$ are independent Rademacher (sign) variables. Carefully justify your work. The quantity on the right, without the leading factor of two, is sometimes called the expected *Rademacher complexity of* $\mathcal{C}$ with respect to $X_1, \ldots, X_n$.

(b) Show that the Rademacher complexity can be bounded as follows

$$\mathbb{E} \sup_{C \in \mathcal{C}} \left| \sum_{i=1}^{n} \varepsilon_i \, I(X_i \in C) \right| \;\leq\; \sqrt{2n \log \mathbb{E} S(X_1^n : \mathcal{C})}$$

where $S(x_1^n : \mathcal{C}) = |\{C \cap \{x_1, \ldots, x_n\} : C \in \mathcal{C}\}|$ is the shatter coefficient of the family $\mathcal{C}$. [Hint: First condition on the observations $X_1, \ldots, X_n$.]

(c) Combine the bounds above with the bounded difference inequality to get a high probability bound on $\Delta(X_1^n)$.

5. Let (S, d) be a metric space, and let $N(S, \epsilon)$ be the covering number of S under the metric $d(.,.)$ at radius ϵ.

(a) What can you say about the limit of $N(S, \epsilon)$ as $\epsilon \to 0$? [Consider the case where S is finite and S is infinite.]

(b) Now let $S_0 \subseteq S$ be a subset of S. By definition, an ϵ-cover of S_0 contains of balls of radius ϵ centered at points in S_0, and $N(S_0, \epsilon)$ is the size of the smallest such cover. Consider instead general ϵ-covers of S_0 that are centered at points in S, so that centers need *not* be in S_0. Let $\tilde{N}(S_0, \epsilon)$ be the smallest such general cover. Find a simple relationship between $N(S_0, \epsilon)$ and $\tilde{N}(S_0, \epsilon)$.