Problem Set 5

Sublinear Algorithms

Due Tuesday, November 25

1. In this problem we look at suprema of Gaussian processes, Dudley's entropy integral, and chaining in more detail. Consider the Gaussian process where we choose $g \sim N(0, I_n)$ and for all $x \in B_1$ we define

$$G_x = \langle g, x \rangle.$$

Suppose we want to bound

$$\gamma = \mathrm{E}_g \sup_{x \in B_1} G_x.$$

   (a) First, let's solve this from first principles. For a fixed g, express $\sup_{x \in B_1} G_x$ in terms of g. Therefore, what is γ up to constant factors?

   (b) Now, let's look at using Dudley's entropy integral. Recall that this says that

$$\mathrm{E} \sup_{x \in T} G_x \lesssim \int_0^\infty \sqrt{\log N(T, d, u)} du$$

   where N denotes the covering number and $d(x, y) = \sqrt{\mathrm{E}(G_x - G_y)^2}$.

      i. What is a nicer expression for $d(x, y)$ in this case?

      ii. Using Maurey's empirical method or otherwise, give a bound on $N(T, d, u)$ for the relevant T and d in this problem.

      iii. Therefore, show that $\gamma \lesssim (\log n)^{3/2}$ using Dudley's entropy integral. Hint: you may need to use two different bounds on N, depending on whether u is "large" or "small".

      iv. How does this compare to your bound from part 1a?

(c) OK, let's introduce Talagrand's *generic chaining*. This says the following:

Theorem 1 (Generic Chaining)**.** *Let $A_1, A_2, \ldots$ be subsets of T with $|A_i| \leq 2^{2^i}$. Define $d(x, A_i)$ to be $\min_{y \in A_i} d(x, y)$. Then*

$$\operatorname{E} \sup_{x \in T} G_x \lesssim \sup_{x \in T} \sum_{i=1}^{\infty} d(x, A_i) \sqrt{\log |A_i|} = \sup_{x \in T} \sum_{i=1}^{\infty} d(x, A_i) 2^{i/2}$$

Moreover, there exists a choice of the A_i with

$$\operatorname{E} \sup_{x \in T} G_x \eqsim \sup_{x \in T} \sum_{i=1}^{\infty} d(x, A_i) 2^{i/2}$$

[Notice that this theorem is very similar to how Dudley's entropy integral looked before we converted it from a sum into an integral. The difference is that we choose the A_i to have a fixed size, rather than a fixed $\sup_{x \in T} d(x, A_i)$, and the the sup is outside the sum.]

i. Consider the set $T_i \subset B_1$ consisting of all 2^i-sparse vectors with coordinates that are multiples of $1/2^{2i}$. Show that $|T_i| \leq (cn)^{2^i}$ for some constant c.
   Conclude that we can set $A_{i+\log\log(cn)} = T_i$.

ii. We will show that for any $x \in B_1$,

$$\sum_{i=1}^{\infty} d(x, T_i) 2^{i/2} \lesssim 1. \tag{1}$$

   What would this give as a bound on γ?

iii. Consider any $x \in B_1$ that is k-sparse and fairly uniform: i.e., either $x_i = 0$ or $x_i \in [1/2^j, 2/2^j]$ for some j. Show that

$$\sum_{i=1}^{\infty} d(x, T_i) 2^{i/2} \lesssim \|x\|_2 \sqrt{k} + 1/2^{j/2} \lesssim \|x\|_1 + 1/2^{j/2}.$$

iv. Now, split any $x \in B_1$ into "level sets" of fairly uniform magnitude, and show that

$$\sum_{i=1}^{\infty} d(x, T_i) 2^{i/2} \lesssim \|x\|_1 + 1 \lesssim 1.$$

   Conclude that generic chaining gives the right bound on γ for this example.