15 Lecture 15, November 24, 2014

15.1 The Lovász local lemma

We now introduce the Lovász local lemma. The local lemma is one of the few widely applicable tools we have for controlling the interactions among a large collection of random variables.

Consider a probability space in which there is a long (possibly infinite) list of "bad" events $B_1, \ldots$ that might occur. We wish to show that the union of the bad events is not the entire space, or in other words, it is *possible* for every event to come out "good", simultaneously. That is, we wish to show (with ^{c} denoting complementation): $\bigcap B_i^c \neq \emptyset$.

There are, in the probabilistic method, two elementary tools to show this kind of statement:

1. The union bound. If $\sum P(B_i) < 1$ then $(\bigcup B_i)^c \neq \emptyset$.
2. Independence. If $P(B_i) < 1$ for all i, and the B_i are mutually independent, then for any finite n, $P\left(\bigcap B_i^c\right) = \prod_1^n (1 - P(B_i)) > 0$.

Let's consider two toy examples of using just the second tool, independence.

First Toy Example: no matter how many fair coins you toss, it is *possible* for all to come up Heads.

Second Toy Example: Show that any finite (or even locally-finite) tree has a valid 2-coloring. Of course this is trivial, but can you show it by just coloring the vertices uniformly at random? Suppose the tree has $n+1$ vertices. There are n "bad" events, each corresponding to a particular edge being monochromatic; these are mutually independent (exercise). So the probability that the tree is properly colored is $2^{-n} > 0$. This shows that there is a valid coloring of the tree, even though the probability that a random coloring is valid is vanishing in n. (For an infinite, locally finite tree, extend the argument by compactness.)

The second toy example illustrates that the use of independence is very fragile. If you insert into the tree just one extra edge closing an odd-length cycle—no matter how long that cycle is—the dependence induced among distant events is enough to ruin the 2-colorability. So even an assumption of $(n-1)$-way independence among n variables is not sufficient to imply that all good events may intersect!

What Lovász did was to create an argument, somewhat like the independence argument we set out above, but more robust, which still works in situations where the bad events are not entirely independent. His argument is a wonderful combination of tools (1) and (2). We present here one form of Lovász's bound.

Definition 62 (Dependency graph) *Let $B_1, \ldots$ be a finite set of events labeled by a set $\mathcal{S}$. A "dependency graph" for the events is a directed graph whose vertices are the set $\mathcal{S}$, with the following property. Let D_i be the set of in-neighbors of i. (We do not include an event among its own in-neighbors.) Then for all i, B_i is independent of the product random variable $\prod_{j \in \mathcal{S}-\{i\}-D_i} B_j$.*

Observe that the same set of events may be assigned many different dependency graphs. In particular, any edges may be added; more significantly, there can be incomparable minimal dependency graphs. (Exercise: give an example.)

Lemma 63 (Lovász) *Suppose that for all i, $|D_i| \leq \Delta$ and $P(B_i) \leq \frac{1}{e(\Delta+1)}$. Then $\bigcap_{i \in \mathcal{S}} B_i^c \neq \emptyset$. In other words, it is possible for all good events to coincide.*

The factor of e here is best possible; this was shown by Shearer [64].

An application: k-SAT with restricted intersections. This is a canonical application of the Lovasz lemma.

Let $H = (V, E)$ be a SAT instance in conjunctive normal form (CNF).

That is, V is a set of boolean variables; a *literal* is a variable $v \in V$ or its negation. E is a collection of *clauses*, each $T \in E$ being a set of literals, which is satisfied if at least one of them is satisfied. H is satisfied if *all* $T \in E$ are satisfied.

We say that two clauses are neighbors if they share any variable (not necessarily literal).

Corollary 64 *Suppose every clause in $T \in E$ has size $\geq k$ and has at most d neighbors. If $d + 1 \leq 2^k/e$ then H is satisfiable.*

Two cases of this corollary are easy: if the total number of clauses is small, or if the clauses are all disjoint (share no variables). The local lemma uses both effects:

$$\begin{cases} \text{union bound: } \textit{local} \text{ (in the dependency graph)} \\ \text{independence: } \textit{global} \end{cases}$$

Proof: Make a random, uniform assignment to the variables V. For each T there is a "bad event" $B_T =$ no literal in T is satisfied. $\Pr(B_T) = 2^{-k}$. After excluding the edges intersecting T, B_T is independent of the collection of all remaining events, because even finding out the exact coloring of V within those edges (not merely which events occurred) does not affect the probability of event B_T. Now apply Lemma (63). □

An application: Property B (or equivalently, NAE-SAT): Let $H = (V, E)$ be *hypergraph* (a *set system* whose elements we call edges). Specifically V is finite and $E \subseteq 2^V$. H has *Property B* if V can be two-colored so that no edge is monochromatic.

Corollary 65 *Suppose every edge $T \in E$ has size $\geq k$ and intersects at most d other edges. If $d + 1 \leq 2^{k-1}/e$ then H has Property B.*

Property B is just like SAT except that two assignments, rather than one assignment, are ruled out per clause.

Proof: of the local lemma:

We demonstrate more concretely that for any finite subset of $\mathcal{S}$, which we relabel for convenience $B_1, \ldots, B_m$,

$$P(\bigcap_{i \in \{B_1,\ldots,B_m\}} B_i^c) \geq \left(\frac{\Delta}{\Delta+1}\right)^m.$$

More specifically, we show by induction on m that for any set of $m-1$ events $B_1, \ldots, B_{m-1}$ and any event B_m,

$$\underbrace{P(B_m^c | \bigcap_{1 \leq j \leq m-1} B_j^c)}_{(I)} \geq \frac{\Delta}{\Delta+1}. \tag{34}$$

The case $m = 1$ is trivial; now suppose the claim is true for values smaller than m. Partition $\{B_1, \ldots, B_{m-1}\}$ into $\mathcal{D} = D_m \cap \{B_1, \ldots, B_{m-1}\}$ and $\mathcal{D}' = \{B_1, \ldots, B_{m-1}\} - \mathcal{D}$. Restating, the goal is to show that

$$P(B_m | \bigcap_{j \in \mathcal{D} \cup \mathcal{D}'} B_j^c) \leq \frac{1}{\Delta+1}.$$

If $\mathcal{D}$ is empty this is immediate. Otherwise, for $\Delta \geq d \geq 1$ relabel the events as $\mathcal{D} = \{m-d, \ldots, m-1\}$ and $\mathcal{D}' = \{1, \ldots, m-d-1\}$, and write

$$\underbrace{P(B_m \cap \bigcap_{\mathcal{D}} B_j^c | \bigcap_{\mathcal{D}'} B_j^c)}_{(II)} = \underbrace{P(B_m | \bigcap_{\mathcal{D} \cup \mathcal{D}'} B_j^c)}_{(I)} \underbrace{P(\bigcap_{\mathcal{D}} B_j^c | \bigcap_{\mathcal{D}'} B_j^c)}_{(III)} \tag{35}$$

We're going to upper bound (I) by expressing it in the form

$$(I) = \frac{(II)}{(III)}.$$

Term (II) is the application of independence at the global level. We use a simple upper bound for it: $B_m \cap \bigcap_{\mathcal{D}} B_j^c \subseteq B_m$, so

$$\underbrace{P(B_m \cap \bigcap_{\mathcal{D}} B_j^c | \bigcap_{\mathcal{D}'} B_j^c)}_{(II)} \leq P(B_m | \bigcap_{\mathcal{D}'} B_j^c) = P(B_m) \leq \frac{1}{e(\Delta+1)}.$$

Term (III) is the union bound at the local level. We could in fact write it explicitly as a union bound but the lemma would suffer the slightly inferior factor of 4 in place of e, so we use the following slightly slicker derivation.

$$\begin{aligned}
\underbrace{P(\bigcap_{\mathcal{D}} B_j^c | \bigcap_{\mathcal{D}'} B_j^c)}_{(III)} &= \prod_{m-d \leq j \leq m-1} P(B_j^c | \bigcap_{j'<j} B_{j'}^c) \\
&\geq \left(\frac{\Delta}{\Delta+1}\right)^d \\
&\geq \left(\frac{\Delta}{\Delta+1}\right)^\Delta \\
&> 1/e
\end{aligned}$$

Where the first inequality is by induction because every conditional probability on the right-hand side is of the form (34) and involves at most $m-1$ sets.

Combining our two bounds we obtain the following from (35):

$$\underbrace{P(B_m | \bigcap_{j \in \mathcal{D} \cup \mathcal{D}'} B_j^c)}_{(I)} \leq \frac{1/e(\Delta+1)}{1/e} = \frac{1}{\Delta+1}.$$

□

15.2 Ramsey numbers

Theorem 66 (Ramsey) *Fix any nonnegative integers k, ℓ. There is a finite "Ramsey number" $R(k,\ell)$ such that every graph on $R(k,\ell)$ vertices contains either a clique of size k or an independent set of size ℓ.*

This theorem runs in the opposite direction to Property B (although it is more sophisticated, referring to coloring vertices rather than graph edges). Not surprisingly, then, our use of the local lemma will be to provide a *lower bound* on Ramsey numbers. We have already seen a lower bound (Problem Set 1); now we will see how to improve it.

15.3 Digression: The Ramsey Upper Bound

Numerous generalizations of Ramsey's argument have since been developed—see the book by Graham, Rothschild and Spencer (Wiley 1980).

Proof: of Theorem (66): This is outside our main line of development but we include it for completeness. First, $R(k,1) = R(1,k) = 1$. Now consider a graph with $R(k,\ell-1) + R(k-1,\ell) + 1$ vertices and pick any vertex v. Let V_Y denote the vertices connected to v by an edge, and let V_N denote the remaining vertices.

If $|V_N| \geq R(k,\ell-1)$ then the graph spanned by $V_N \cup \{v\}$ contains either a k-clique (not using v) or an independent set of size ℓ (using v).

On the other hand if $|V_Y| \geq R(k-1, \ell)$ then the graph spanned by $V_Y \cup \{v\}$ contains either a k-clique (using v) or an independent set of size ℓ (not using v).

If you work out the numbers, the proof of Ramsey's theorem gives the bound $R(k,k) \leq \binom{2k-2}{k-1} \in O(4^k/\sqrt{k})$.

15.4 Elementary Ramsey lower bound

The classic lower bound, which we saw earlier in the course, relies on (a simplest possible) deviation bound:

Theorem 67 (Erdös) *If $\binom{n}{k} < 2^{\binom{k}{2}-1}$ then $R(k,k) > n$. Thus $R(k,k) \geq \frac{2^{k/2}k}{e\sqrt{2}}(1-o(1))$.*

Proof: (of Theorem (67)) Color the edges uniformly iid. Any particular subgraph of k vertices has probability only $2^{1-\binom{k}{2}}$ of being monochromatic. Take a union bound over all subgraphs. □

Needless to say this leaves an exponential gap to Theorem 66. Actually this gap is small by the standards of Ramsey theory. The local lemma will slightly tighten this gap, as we now show, but the remaining gap is a major open problem in combinatorics.

15.5 Better $R(k,k)$ lower bound through the Local Lemma

Theorem 68 *If $e\binom{k}{2}\binom{n}{k-2} < 2^{\binom{k}{2}-1}$ then $R(k,k) > n$. Thus $R(k,k) \geq \frac{2^{k/2}k\sqrt{2}}{e}(1-o(1))$.*

Proof: As before, color the edges of K_n uniformly iid. For each set of k vertices the "bad event" of a monochromatic clique occurs with probability $2^{1-\binom{k}{2}}$. For the dependency graph, connect two subsets if they share an edge. The degree of this graph is strictly less than $\binom{k}{2}\binom{n-2}{k-2}$, so $\Delta + 1 \leq \binom{k}{2}\binom{n}{k-2}$.

This improves the union bound by a factor of only 2, but is the best lower bound known.

Although the improvement factor is very small, qualitatively it is quite meaningful. It means that a certain negative correlation among edges is possible: you have a graph which is big enough to have on average *many* copies of each graph of size k, yet some kind of negative correlation exists which prevents the occurrence of the extreme graphs (the k-clique and the k-indep-set).