Probabilistic Method: Lovasz Local Lemma

April 28, 2011

Motivation. In the basic setup of the probabilistic method to prove the existence of a certain object, we have to prove that the probability of a certain event is greater than zero. In many cases involving the first- or second moment method, we actually proved something stronger and showed that the probability of the event is not only greater than zero but also very high (examples include the high girth/high chromatic number result and the result about the number of prime factors). We are now going to focus on the case where the probability of the desired event will hold with positive, albeit very small, probability. A trivial example is when we have n mutually independent events and each of them holds with probability at least $p > 0$. Then the probability that all these events occur is at least $p^n > 0$. The Lovasz Local Lemma provides a way to obtain a similar statement, provided that the events are *not* mutually independent but the dependencies are not "too strong".

1 Recap: Pairwise Independence and Mutual Independence

Definition 1. *A collection of events* $\{\mathcal{E}_i \colon i \in \mathcal{I}\}$ *is* pairwise independent *if for every pair* $i \neq j \in \mathcal{I}$,

$$\mathbf{Pr}\left[\mathcal{E}_i \wedge \mathcal{E}_j\right] = \mathbf{Pr}\left[\mathcal{E}_i\right] \cdot \mathbf{Pr}\left[\mathcal{E}_j\right].$$

(Note that if two events $\mathcal{E}_1$ *and* $\mathcal{E}_2$ *are disjoint, then they are dependent!) A collection of random variables* $\{X_i : \Omega \to \mathbb{R}, : i \in \mathcal{I}\}$ $X, Y : \Omega \to \mathbb{R}$ *is* pairwise independent *if for every values* $\{x_i \in \mathbb{R}\}$,

$$\mathbf{Pr}\left[\wedge_{i \in \mathcal{I}} (X_i \leq x_i)\right] = \prod_{i \in \mathcal{I}} \mathbf{Pr}\left[X_i \leq x_i\right].$$

Remark: for random variables whose range is a finite or countably infinite set (which will be the common setting in this course) one can replace in the above definition"$\leq x_i$" by "$= x_i$".

Note that pairwise independence is equivalent to $\mathbf{Pr}\left[\mathcal{E}_i \mid \mathcal{E}_j\right] = \mathbf{Pr}\left[\mathcal{E}_i\right]$ for all $j \neq i$ (assuming that $\mathcal{E}_j \neq \emptyset$). A useful fact about pairwise independence is linearity of the variance, i.e., when $X_1, \ldots, X_n$ are pairwise independent random variables

$$\mathbf{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{Var}\left[X_i\right].$$

Definition 2. *A collection of events* $\{\mathcal{E}_i \colon i \in \mathcal{I}\}$ *is* (mutually) independent *if for all subsets* $\mathcal{S} \subseteq \mathcal{I}$,

$$\mathbf{Pr}\left[\wedge_{i \in \mathcal{S}} \mathcal{E}_i\right] = \prod_{i \in \mathcal{S}} \mathbf{Pr}\left[\mathcal{E}_i\right].$$

(Mutual) independence of random variables is defined in the same way as pairwise independence.

Similar to pairwise independence, (mutually) independence is equivalent to

$$\forall i \in \mathcal{I}, \forall S \subseteq \mathcal{I} \setminus \{i\}: \quad \mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j \in S} \mathcal{E}_j\right] = \mathbf{Pr}\left[\mathcal{E}_i\right],$$

again assuming that the event we condition on has non-zero probability. Further, it is easy to prove by induction that if a collection $\{\mathcal{E}_i : i \in \mathcal{I}\}$ is *(mutually) independent*, then for any subset $S \subseteq \mathcal{I}$, the collection $\{\mathcal{E}_i : i \in \mathcal{S}\} \cup \{\overline{\mathcal{E}_j} : j \in \mathcal{I} \setminus \mathcal{S}\}$ is also (mutually) independent.

Example: Three random variables which are pairwise independent, but *not* mutually independent. Let $X = \{0, 1\}$ and $Y = \{0, 1\}$ be (pairwise) independent random variables taking 0/1 with probability 1/2. Then let $Z = X \otimes Y$ (i.e., X "xor" Y). Then, the random variables $\{X, Y, Z\}$ are pairwise independent, but not mutually independent (knowing the outcome of any two random variables determines the third one).

Definition 3. *Let $\{\mathcal{E}_i : i \in \mathcal{I}\}$ be a collection of events. An event $\mathcal{E}_i$ is* mutually independent of *a set $\cup_{j \in \mathcal{S}} \mathcal{E}_j$ of events if for every two disjoint subsets $\mathcal{T}, \overline{\mathcal{T}} \subseteq \mathcal{S}$,*

$$\mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j \in \mathcal{T}} (\mathcal{E}_j) \wedge (\wedge_{k \in \overline{\mathcal{T}}} \overline{\mathcal{E}_k})\right] = \mathbf{Pr}\left[\mathcal{E}_i\right].$$

Note that if $\{\mathcal{E}_i : i \in \mathcal{I}\}$ is a mutually independent collection of events, then any event $\mathcal{E}_i$ is mutually independent of $\cup_{j \in \mathcal{I} \setminus \{i\}} \mathcal{E}_j$. However, the converse is not true, as for instance, $\mathcal{E}_1$ could be mutually independent of $\cup_{j=2}^{n} \mathcal{E}_j$ while $\mathcal{E}_2$ and $\mathcal{E}_3$ are dependent, i.e., not pairwise independent.

2 The Lovasz Local Lemma

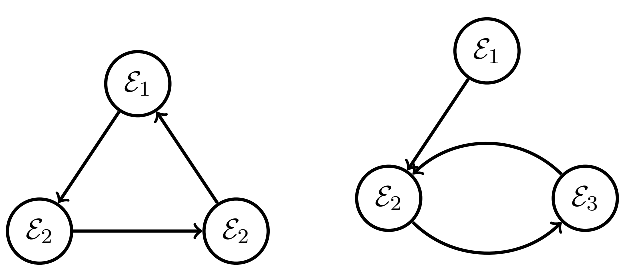
As indicated earlier, the Lovasz Local Lemma (LLL) deals with events that may be dependent, but their dependencies should be rather small. To this end, we make the following definition.

Definition 4. *Let $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ be n events on a probability space Ω. The* dependency graph *is a directed graph $D = (V, E)$ on the set of vertices $V = \{1, \ldots, n\}$ (corresponding to $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$) if for each $1 \leq i \leq n$, $\mathcal{E}_i$ is mutually independent of all the events $\{\mathcal{E}_j : (i, j) \notin E\}$.*

Note that the dependency graph is *not* unique. For instance, consider two independent coin flips with outcome H or T. Consider the events

$$\begin{aligned}
\mathcal{E}_1 &:= \{(H, H), (H, T)\} && \text{(first coin flip is } H\text{)}, \\
\mathcal{E}_2 &:= \{(H, H), (T, H)\} && \text{(second coin flip is } H\text{)}, \\
\mathcal{E}_3 &:= \{(H, H), (T, T)\} && \text{(both coin flips are the same)}.
\end{aligned}$$

Then, two possible dependency graphs are as follows:



More generally, any directed graph with minimum outdegree 1 is a dependency graph.
If all events $\mathcal{E}_1, \dots, \mathcal{E}_n$ are mutually independent, then the empty graph is a dependency graph.

Theorem 5 (Lovasz Local Lemma (LLL), general case)**.** *Let $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$ be n "bad" events on a probability space Ω with a dependency graph D. Suppose that there are real numbers $x_1, x_2, \dots, x_n \in [0, 1)$ such that*

$$\forall 1 \leq i \leq n\colon \quad \mathbf{Pr}\left[\mathcal{E}_i\right] \leq \underbrace{x_i}_{\text{"ideal probability"}} \cdot \underbrace{\prod_{j\colon (i,j) \in E} (1 - x_j)}_{\text{"penalty due to dependencies"}} .$$

Then,

$$\mathbf{Pr}\left[\wedge_{i=1}^n \overline{\mathcal{E}_i}\right] \geq \prod_{i=1}^n (1 - x_i) > 0.$$

Before proving the LLL, let us look at the case where the dependency graph D is emtpy, i.e., the events $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$ are mutually independent. Then,

$$\mathbf{Pr}\left[\wedge_{i=1}^n \overline{\mathcal{E}_i}\right] = \prod_{i=1}^n \mathbf{Pr}\left[\overline{\mathcal{E}_i}\right],$$

so we can set $x_i := \mathbf{Pr}\left[\mathcal{E}_i\right]$ and the statement of the LLL is tight. Hence we can think of the term $\prod_{i=1}^n (1 - x_i)$ as a "penalty factor" for dependencies among $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$; the larger the dependencies among these events, the smaller the individual probabilites for $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_n$ have to be in order for the LLL to apply.

Proof of Theorem 5. We first prove by induction on s, that for any $S \subseteq \{1, \dots, n\}$ with $|S| = s$,

$$\mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j \in S} \overline{\mathcal{E}_i}\right] \leq x_i. \tag{1}$$

Assuming that (1) holds for all $0 \leq s < n$, the LLL follows, since

$$\begin{aligned}
\mathbf{Pr}\left[\wedge_{i=1}^n \overline{\mathcal{E}_i}\right] &= \prod_{i=1}^n \mathbf{Pr}\left[\overline{\mathcal{E}_i} \mid \wedge_{j=1}^{i-1} \overline{\mathcal{E}_j}\right] \\
&= \prod_{i=1}^n \left(1 - \mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j=1}^{i-1} \overline{\mathcal{E}_j}\right]\right) \geq \prod_{i=1}^n (1 - x_i).
\end{aligned}$$

Hence it remains to prove (1). This is obviously true for $s = 0$. For the induction step, assume that (1) holds for all $s' < s$. Let $S \subseteq \{1, \dots, n\}$ be any set of size $|S| = s$ and $i \notin S$. We partition S into two disjoint sets:

$$\begin{aligned}
S_i &= \{j \in S\colon (i, j) \in E\} && \text{(the dependent part)}, \\
S_{\neg i} &= S \setminus S_i && \text{(the independent part)}.
\end{aligned}$$

By definition of conditional probabilities,

$$\begin{aligned}
\mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j \in S} \overline{\mathcal{E}_j}\right] &= \frac{\mathbf{Pr}\left[\mathcal{E}_i \wedge (\wedge_{j \in S} \overline{\mathcal{E}_j})\right]}{\mathbf{Pr}\left[\wedge_{j \in S} \overline{\mathcal{E}_j}\right]} \\
&= \frac{\mathbf{Pr}\left[\mathcal{E}_i \wedge (\wedge_{j \in S_i} \overline{\mathcal{E}_j}) \mid \wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right] \cdot \mathbf{Pr}\left[\wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right]}{\mathbf{Pr}\left[\wedge_{j \in S_i} \overline{\mathcal{E}_j} \mid \wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right] \cdot \mathbf{Pr}\left[\wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right]} \\
&= \frac{\mathbf{Pr}\left[\mathcal{E}_i \wedge (\wedge_{j \in S_i} \overline{\mathcal{E}_j}) \mid \wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right]}{\mathbf{Pr}\left[\wedge_{j \in S_i} \overline{\mathcal{E}_j} \mid \wedge_{j \in S_{\neg i}} \overline{\mathcal{E}_j}\right]}.
\end{aligned} \tag{2}$$

We first upper bound the numerator of (2):

$$\mathbf{Pr}\left[\mathcal{E}_i \wedge (\wedge_{j\in S_i}\overline{\mathcal{E}_j}) \mid \wedge_{j\in S_{\neg i}}\overline{\mathcal{E}_j}\right] \le \mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j\in S_{\neg i}}\overline{\mathcal{E}_j}\right] = \mathbf{Pr}\left[\mathcal{E}_i\right] \le x_i \cdot \prod_{j\colon (i,j)\in E}(1-x_j),$$

where the equality follows since $\mathcal{E}_i$ is mutually independent of all the events $\{\mathcal{E}_j\colon j\in S_{\neg i}\}$. Let us now lower bound the denominator of (2). To this end, let $S_i=\{j_1,\ldots,j_r\}$. If $r=0$, then the denominator is 1, for $r\ge 1$ we use the induction hypothesis to obtain that

$$\begin{aligned}
\mathbf{Pr}\left[\wedge_{j\in S_i}\overline{\mathcal{E}_j} \mid \wedge_{j\in S_{\neg i}}\overline{\mathcal{E}_j}\right] &= \prod_{k=1}^{r}\mathbf{Pr}\left[\overline{\mathcal{E}_{j_k}} \mid (\wedge_{l=1}^{k-1}\overline{\mathcal{E}}_{j_l})\wedge(\wedge_{j\in S_{\neg i}}\overline{\mathcal{E}_j})\right]\\
&= \prod_{k=1}^{r}\left(1-\mathbf{Pr}\left[\mathcal{E}_{j_k} \mid \underbrace{(\wedge_{l=1}^{k-1}\overline{\mathcal{E}}_{j_l})\wedge(\wedge_{j\in S_{\neg i}}\overline{\mathcal{E}_j})}_{\le r-1+(s-r)\le s-1\text{ events}}\right]\right)\\
&\overset{\text{IH}}{\ge} \prod_{k=1}^{r}(1-x_{j_k}) \ge \prod_{j\colon (i,j)\in E}(1-x_j).
\end{aligned}$$

Plugging the two bounds into (2),

$$\mathbf{Pr}\left[\mathcal{E}_i \mid \wedge_{j\in S}\overline{\mathcal{E}_j}\right] \le \frac{x_i\cdot\prod_{j\colon (i,j)\in E}(1-x_j)}{\prod_{j\colon (i,j)\in E}(1-x_j)} = x_i,$$

which proves (1) and completes the proof. □

Corollary 6 (Lovasz Local Lemma, Symmetric Version)**.** *Let $\mathcal{E}_1,\ldots,\mathcal{E}_n$ be events and suppose that each event $\mathcal{E}_i$ is mutually independent of a set of all the other events $\mathcal{E}_j$, $j\ne i$, except for at most d, and suppose that $\mathbf{Pr}\left[\mathcal{E}_i\right]\le p$ for all $1\le i\le n$. Then if*

$$ep(d+1)\le 1,$$

then

$$\mathbf{Pr}\left[\wedge_{i=1}^{n}\overline{\mathcal{E}_i}\right] > 0.$$

It was shown by Shearer in 1985 that the constant e in the inequality $ep(d+1)\le 1$ is best possible. One special case of Corollary 6 is when no event $\mathcal{E}_i$ is independent of any other event $\mathcal{E}_j$, $1\le j\le n$, e.g., all events $\mathcal{E}_1,\mathcal{E}_2,\ldots,\mathcal{E}_n$ are disjoint. Assume also that $\mathbf{Pr}\left[\mathcal{E}_i\right]=p$ for each i. Then,

$$\begin{aligned}
\mathbf{Pr}\left[\wedge_{i=1}^{n}\overline{\mathcal{E}_i}\right] &= 1-\mathbf{Pr}\left[\overline{\wedge_{i=1}^{n}\overline{\mathcal{E}_i}}\right]\\
&= 1-\mathbf{Pr}\left[\vee_{i=1}^{n}\mathcal{E}_i\right]\\
&= 1-\sum_{i=1}^{n}\mathbf{Pr}\left[\mathcal{E}_i\right] && \text{(since } \mathcal{E}_i,\ 1\le i\le n, \text{ are disjoint)}\\
&= 1-np,
\end{aligned}$$

which is greater than zero iff $p<1/n$. As here, $d+1=n$, we get a slightly worse result by Corollary 6 which requires $p<1/(en)$.

Proof of Corollary 6. If $d = 0$, then the result is trivial (all the events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ are mutually independent). Otherwise, we have $d > 0$ and then the assumptions imply that there is a dependency graph $D = (V, E)$ for the events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ in which for each $1 \leq i \leq n$, $|\{j \colon (i,j) \in E\}| \leq d$. We choose $x_i := 1/(d+1) < 1$ for all i and we have to check that the condition of Theorem 5 is satisfied. To see this, take any $1 \leq i \leq n$ and note that

$$x_i \cdot \prod_{j:(i,j)\in E} (1 - x_j) \geq \frac{1}{d+1} \cdot \left(1 - \frac{1}{d+1}\right)^d \geq \frac{1}{d+1} \cdot \frac{1}{e} \geq p \geq \mathbf{Pr}\left[\mathcal{E}_i\right].$$

Hence, Corollary 6 follows from Theorem 5. □