The Lovász Local lemma

Definition (6.1)

A dependency graph for a set of events $E_1, \ldots, E_N$ is a graph $G = (V, E)$ such that $V = \{1, \ldots, N\}$ and for each $i = 1, \ldots, N$, the event $i$ is mutually independent with the events $\{E_j | (i, j) \notin E\}$. The degree of the dependency graph is the max degree vertex of $G$.

Theorem (6.11, Lovász Local Lemma)

Let $E_1, \ldots, E_N$ be a set of events and assume we know $p \in (0, 1), d \in \mathbb{N}$ such that all the following conditions hold:

1. For all $i$, $\Pr[E_i] \leq p$;
2. The degree of the dependency graph on $\{E_1, \ldots, E_N\}$ is $\leq d$;
3. $4dp \leq 1$

Then

$$\Pr[\bigcap_{i=1}^{N} E_i] > 0.$$ 

Proof of Lovász Local Lemma

The proof depends on showing the following claim by induction:

For $s = 0, 1, \ldots, n - 1$, if $|S| \leq s$, $\Pr[\bigcap_{j \in S} E_j] > 0$, and for every $k \in [n] \setminus S$,

$$\Pr[E_k \cap \bigcap_{j \in S} E_j] \leq 2p.$$ 

After the claim is shown, it is not hard to obtain our result as follows:

$$\Pr[\bigcap_{j \in S} E_j] = \prod_{i=1}^{n} \Pr[E_i | \bigcap_{j \in S} E_j] = \prod_{i=1}^{n} \left(1 - \Pr[E_i | \bigcap_{j \in S} E_j]\right) = \prod_{i=1}^{n} (1 - 2p) > 0.$$

The last step uses the fact that $4dp \leq 1$ (hence certainly $2p < 1$).
LLL: proof of Key Claim

Claim: Under the conditions of the LLL, if we take any $s = 0, \ldots, n - 1$, and any $S \subseteq [n], |S| = s$, then for any $k \not\in S$ we have

$$\Pr \left[ E_k \mid \bigcap_{j \in S} \bar{E}_j \right] \leq 2p.$$

Base case $(s = 0)$: Follows from assumption that $\Pr[E_j] \leq p$ for all $j$.

Induction step: Assume true for $0, 1, \ldots, s - 1$. Next show for $s$.

First need to show $\Pr[\bigcap_{j \in S} \bar{E}_j] > 0$ (to be able to evaluate conditional expectation).

If $s = 1$ follows from assumptions ($\Pr[E_j] > 1 - p$).

For $s > 1$, then we can use claim for $(s - 1), \ldots, 0$ to show $\Pr[\bigcap_{j \in S} \bar{E}_j] \geq (1 - 2p)^s > 0$.

---

LLL: proof of Key Claim (cont'd.)

Induction step (cont'd): (|$S_2| < s$ case) Our target quantity $\Pr[ E_k \mid \bigcap_{j \in S} \bar{E}_j ]$ is equal to

$$\Pr[ E_k \mid F_{S_1} \cap F_{S_2} ] = \frac{\Pr[ E_k \cap F_{S_1} \cap F_{S_2} ]}{\Pr[ F_{S_1} \cap F_{S_2} ]}.$$

We can use a trick to re-write top/bottom as

$$\Pr[ E_k \cap F_{S_1} \cap F_{S_2} ] = \Pr[ E_k \cap F_{S_1} ] \cdot \Pr[ F_{S_2} ]$$

We can cancel the common factor (well-defined - see end slide 5).

We will then argue about $\Pr[ E_k \cap F_{S_1} ]$ ("top") and $\Pr[ F_{S_1} ]$ ("bottom") separately.

---

LLL: proof of Key Claim (cont'd.)

Induction step (cont'd): (|$S_2| < s$ case) We want to show $\Pr[ E_k \mid \bigcap_{j \in S} \bar{E}_j ] \leq 2p$.

Consider $E_k$’s node in the dependency graph, and let

$S_1 = \{ j \in S : E_j \text{ is a neighbour of } E_k \}$,

$S_2 = S \setminus S_1$.

If $|S_2| = S$, then $E_k$ is mutually independent of all events in $S$. Done!

Else $|S_2| < s$, and there are some mutually dependent events (those in $S_1$) with $E_k$ in the "conditional".

We re-write the conditional $\bigcap_{j \in S} \bar{E}_j$ as $F_{S_1} \cap F_{S_2}$, with

$$F_{S_1} = \bigcap_{j \in S_1} \bar{E}_j$$

$$F_{S_2} = \bigcap_{j \in S_2} \bar{E}_j$$

---

LLL: proof of Key Claim (cont'd.)

Induction step (cont'd): (|$S_2| < s$ case) We have re-written our target quantity as

$$\frac{\Pr[ E_k \cap F_{S_1} \cap F_{S_2} ]}{\Pr[ F_{S_1} \cap F_{S_2} ]}.$$ 

"top": For sure, we have $\Pr[ E_k \cap F_{S_1} ] \leq \Pr[ E_k ]$.

But $E_k$ is mutually independent of all of $F_{S_2}$ events (by definition of $S_2$). Hence $\Pr[ E_k ] = \Pr[ E_k ]$, which by assumption is $< p$.

"bottom": We have $\Pr[ F_{S_1} ]$, where $|S_2| < s$. So ... can use our I.H. with any individual events of $F_{S_1}$.

$$\Pr[ F_{S_1} ] = \Pr \left[ \bigcap_{j \in S_1} \bar{E}_j \right] \geq 1 - \sum_{i \in S_1} \Pr[ E_i \mid \bigcap_{j \in S_2} \bar{E}_j ]$$
### LLL: proof of Key Claim (cont’d.)

**Induction step (cont’d):** \(|S_2| < s\) case We have re-written our target quantity as
\[
\frac{\Pr[E_k \cap F_{S_i}, F_{S_j}]}{\Pr[F_{S_i}, F_{S_j}]} \leq \frac{p}{\Pr[F_{S_i}, F_{S_j}]}.
\]

“top”: For sure, we have \(\Pr[E_k \cap F_{S_i}, F_{S_j}] \leq \Pr[E_k | F_{S_i}].\)
We have shown
\[
\Pr[F_{S_i}, F_{S_j}] \geq 1 - \sum_{i \in S_i} \Pr[E_i | \bigcap_{j \in S_j} \bar{E_j}]
\]

Now (because \(|S_2| < s\), and fits our I.H.), we have that
\[
\Pr[E_i | \bigcap_{j \in S_j} \bar{E_j}] \leq 2p \text{ for all } i \in S_i.
\]
We also know that \(S_t\) (items dependent with \(E_k\)) has \(\leq d\) items.
Hence we have
\[
\Pr[F_{S_i}, F_{S_j}] \geq 1 - 2p \cdot d, \text{ which is } \geq \frac{1}{2}.
\]

### LLL: proof of Key Claim (wrapping-up.)

**Induction step (cont’d):** \(|S_2| < s\) case We have re-written our target quantity as
\[
\frac{\Pr[E_k \cap F_{S_i}, F_{S_j}]}{\Pr[F_{S_i}, F_{S_j}]} \leq \frac{p}{\Pr[F_{S_i}, F_{S_j}]}.
\]
We now have \(\Pr[F_{S_i}, F_{S_j}] \geq \frac{1}{2}\).
Hence we get
\[
\frac{\Pr[E_k \cap F_{S_i}, F_{S_j}]}{\Pr[F_{S_i}, F_{S_j}]} \leq \frac{p}{2} = 2p,
\]
as required.
We had already proved that this “claim” gives us the LLL.

---

### SAT and \(k\)-SAT (standard probabilistic method)

Recall that in propositional logic, a **Boolean variable** \(x_i\) can take on 0 or 1 values, a **literal** is either \(x_i\) or \(\overline{x_i}\), and for the set of variables \(\{x_1, \ldots, x_n\}\) a SAT problem is any conjunction (AND) of a set of clauses, each individual clause being a disjunction (OR) of literals. For example,
\[
\{x_4, \overline{x_7}, \overline{x_9}\}, \{\overline{x_1}, x_3, \overline{x_5}\}, \{x_1, x_2, \overline{x_6}\}, \{x_3, x_9, x_3\}
\]
is an instance of SAT. Since all clauses are of length 3, the one above is also an instance of 3-SAT.

Suppose we have \(m\) clauses, with \(k_i\) literals in the \(i\)th clause, \(1 \leq i \leq m\). Then on a **uniform random** assignment of boolean values to the \(n\) variables, the probability clause \(i\) is satisfied is \((1 - 2^{-k_i})\).

\((2^k\) is the probability we would set all \(k_i\) literals of this clause to be false.)

---

### SAT and \(k\)-SAT (standard probabilistic method)

Suppose we have \(m\) clauses, with \(k_i\) literals in the \(i\)th clause, \(1 \leq i \leq m\). Then on a **uniform random** assignment of boolean values to the \(n\) variables, the expected number of **satisfied** clauses is \(\sum_{i=1}^m (1 - 2^{-k_i})\).

- This is at least \(m \cdot (1 - 2^{-k})\), where \(k = \min_{i=1}^m k_i\).
- If the instance is \(k\)-SAT (all clauses length \(k\)), the expected number of satisfied clauses is **exactly** this.
- Hence (by probabilistic method) there is at least one assignment to \(\{x_1, \ldots, x_n\}\) with at least \(\sum_{i=1}^m (1 - 2^{-k_i})\) satisfied clauses.
- Can’t get any condition guaranteeing “satisfiability” (all \(m\) clauses) even for \(k\)-SAT as \(m \cdot (1 - 2^{-k})\) is strictly less than \(m\). Would need to do a different kind of analysis.
**k-SAT with Lovász Local Lemma**

Now consider the “bad events” $E_i$ to be the event where clause $i$ becomes unsatisfied, and consider the dependency graph.

**Theorem (6.13)**
If we have a k-SAT formula where no variable appears in more than $T = \frac{2^k}{4^k}$ clauses, then that formula has some satisfying assignment.

**Proof**
We assume a uniform random assignment to all the $x_j$ and let $E_i$ be the event that all the $k$ variables get the “wrong” assignment. $\Pr[E_i] \leq 2^{-k}$ for all $i$.

The event $E_i$ is mutually dependent of any $E_{i'}$ such that clause $i'$ shares no logical variables with clause $i$. For each of the variables in clause $i$, they may appear in $T = \frac{2^k}{4^k}$ clauses, so taking all $k$ variables, there are at most $k \cdot T = \frac{2^k}{4}$ clauses which share some variable(s) with clause $i$. So $d \leq \frac{2^k}{4}$.

Then $4dp \leq 4 \cdot \frac{2^k}{4} \cdot 2^{-k} = 1$, and the LLL implies there is some assignment where none of the bad events occur (ie, all clauses are satisfiable).

**Notes**

**Reading**
- Section 6.6 of the book presents the “Conditional Expectation Inequality” and shows how to shorten the proof of the “no clique with 4 vertices” half of Theorem 6.8 - ie, 6.8(b) of Lecture 12.
- Section 6.7 from the book deals with the Lovasz Local Lemma.

**Doing**
- Coursework 2 is due 24th March at 4pm (GMT).
- The final tutorial sheet will be shipped soon, tutorials will take place in week 11 (Tues 31st March, Wed 1st April).

**LLL: can we de-randomize?**

The only negative aspect of using LLL is that we don’t get an explicit randomized process linked to the existence result. So we don’t have a handle on how we might go about finding such a object.

There are ways to convert a LLL result into an explicit construction, but usually you need a lower dependency value.

We won’t cover this (see Sections 6.8, 6.10 of the book)