Lecture 21: Lovász Local Lemma

November 12, 2025

1 Lovász Local Lemma

LLL is a powerful tool in the probabilistic method and has found several highly non-trivial applications in algorithms.

In the probabilistic method we prove the existence of some object by running a probabilistic experiment and arguing that the object property holds with non-zero probability.

1.1 First Moment Method

We use expectation analysis. As an example, we showed that in any graph $G = (V, E)$ there exists a max cut of value $\frac{|E|}{2}$ by picking a random cut whose expectation is $\frac{|E|}{2}$.

1.2 Second Moment Method

Here we use variance analysis plus something like the Chebyshev bound.

Example 1. Let $G(n, p)$ be a random graph on n vertices where each edge is chosen independently with probability p. At what value of p will $G(n, p)$ have a clique of size 4?

Clearly when $p = 0$, the graph will be very sparse and there will not be a 4-clique. When $p = 1$ there will be a 4-clique with high probability since the graph becomes dense.

Turns out that $p = \frac{1}{n^3}$ is the threshold.

To see one direction we use first moment method. Let X be the number of 4-cliques.

$$\mathbb{E}[X] = \binom{n}{4} p^6$$

since if we fix a set of 4 vertices, it will be a clique iff all six edges are chosen.

If $p \leq \frac{c}{n^3}$ for sufficiently small constant c, then $\mathbb{E}[X] \leq 0.1$.

For a non-negative integer random variable:

$$\Pr[X \geq 1] \leq \mathbb{E}[X]$$

So $\Pr[X = 0] \geq 1 - \mathbb{E}[X] \geq 0.9$.

We would like to compute the variance of X. Note that $X = \sum_{S \in \mathcal{S}} X_S$ where S ranges over all $\binom{n}{4}$ subsets of four vertices and X_S is an indicator for S being a clique. $\Pr[X_S = 1] = p^6$.

If $S_i, S_j \in \mathcal{S}$, then X_{S_i} and X_{S_j} are independent if S_i and S_j do not share any edges. Otherwise they are dependent.

To estimate $\mathrm{Var}(X)$ we write:

$$\mathbb{E}[X^2] = \mathbb{E}\left[\left(\sum_S X_S\right)^2\right] = \sum_S \mathbb{E}[X_S^2] + 2\sum_{S\sim S'} \mathbb{E}[X_S \cdot X_{S'}]$$

where $S \sim S'$ means S and S' are dependent and $S \perp S'$ means independent.

Suppose we pretend all S and S' are independent. Then there are roughly n^8 such pairs and in that case:

$$\mathrm{Var}(X) = \sum_{S\in\mathcal{S}} \mathrm{Var}(X_S) = n^4p^6(1-p^6)$$

Then by Chebyshev:

$$\Pr[X=0] \le \Pr[|X-\mathbb{E}[X]| \ge \mathbb{E}[X]] \le \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} \le \frac{n^4p^6}{(n^4p^6)^2} = \frac{1}{n^4p^6}$$

One can check that if $p \ge \frac{C}{n^{3/2}}$ for sufficiently large C, then $\Pr[X=0] \le 0.1$.

However, X_S and $X_{S'}$ are not independent for all S_i and $S_j \in \mathcal{S}$. But if we calculate $\mathrm{Var}(X)$ more carefully we see that:

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_{S_i\in\mathcal{S}} \mathbb{E}[X_{S_i}^2] + \sum_{S_i\sim S_j} \mathbb{E}[X_{S_i}X_{S_j}]\\ &= \sum_{S\in\mathcal{S}} \mathbb{E}[X_S] + \sum_{S\perp S'} \mathbb{E}[X_S]\mathbb{E}[X_{S'}]\end{aligned}$$

There are $\binom{n}{4}^2 = O(n^8)$ total pairs. How many are not independent? $S \sim S'$ if they share at least one edge. But the number of dependent pairs is $O(n^7)$ while the number of independent pairs is $\Theta(n^8)$, which is close to all pairs. So $\mathrm{Var}(X)$ still behaves as if not all pairs are independent and one can show that $p \ge \frac{C}{n^{3/2}}$ for sufficiently large constant ensures $G(n,p)$ has a 4-clique with probability ≥ 0.9.

1.3 Concentration plus Union Bound

We saw several examples of using concentration bounds plus union bound. The general strategy is to show that for some events $A_1, A_2, \ldots, A_n$:

$$\Pr[A_i] \le \frac{1}{n}$$

where A_i is a bad event.

Then by the union bound:

$$\Pr[A_1 \cup A_2 \cup \cdots \cup A_n] \le n \cdot \frac{1}{n} = 1$$

$$\Pr[\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}] > 0$$

Thus we have all good events happening with non-zero probability, and if we can express the property we want in those terms then we are done.

Example 2 (Routing paths)**.** We saw that we can convert fractional solution to integral solution by randomized rounding. We get $O(\log n/\log\log n)$ congestion by using Chernoff bounds on each single edge and then union bound over all edges.

1.4 Local Phenomena

There are many situations where we cannot use union bound because the individual bad event probability is not that small.

If A_i are independent this does not matter because we have:

$$\Pr[\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}] = \prod_{i=1}^{n} (1 - \Pr[A_i])$$

and hence all we need is $\Pr[A_i] < 1$.

However, independence is rarely possible in complex events.

LLL considers a local setting where the events $A_1, \ldots, A_n$ are not completely independent but there is some limited dependence.

How can one capture such a scenario? For that we use a **dependency graph** on the events. The vertices are the events and we have an edge (A_i, A_j) if A_i and A_j are dependent. No edge means that A_i and A_j are conditionally independent. That is, $\Pr[A_i \mid A_j] = \Pr[A_i]$.

Note that A_i is conditionally independent of all events that it has no edges to. We define it formally:

Definition 1. An event A is conditionally independent with respect to $B_1, B_2, \ldots, B_k$ if:

$$\forall S \subseteq \{1, 2, \ldots, k\} : \Pr\left[A \mid \bigcap_{i \in S} B_i\right] = \Pr[A]$$

When can we easily identify conditional independence?

Claim 1. *Suppose $X_1, X_2, \ldots, X_n$ are independent random variables. Suppose event A_i is completely determined by a subset $S_i \subseteq \{X_1, X_2, \ldots, X_n\}$. If $S_i \cap S_j = \emptyset$ for $j \in \{j_1, j_2, \ldots, j_k\}$ then A_i is mutually independent of $A_{j_1}, A_{j_2}, \ldots, A_{j_k}$.*

With this in place we state the Symmetric version of the LLL:

Theorem 1 (Symmetric LLL)**.** *Suppose $A_1, A_2, \ldots, A_n$ are events in an underlying probability space and let d be the max degree of the dependency graph and $\Pr[A_i] \leq p$ for all i. Then if $ep(d+1) \leq 1$, i.e., if $pd \leq \frac{1}{e}$, then:*

$$\Pr\left[\bigcap_{i=1}^{n} \overline{A_i}\right] \geq (1 - 2p)^n > 0$$

so in particular if $pd \leq \frac{1}{e}$ then $\Pr[\bigcap_i \overline{A_i}] > 0$.

A more general version of the LLL called asymmetric or lopsided LLL is the following:

Theorem 2 (LLL)**.** *Suppose $A_1, \ldots, A_n$ are events in a probability space and let H be the dependency graph. Suppose there exist numbers $x_1, x_2, \ldots, x_n \in (0, 1)$ such that:*

$$\Pr[A_i] \leq x_i \prod_{j \in N(i)} (1 - x_j)$$

Then:

$$\Pr\left[\bigcap_i \overline{A_i}\right] \geq \prod_{i=1}^{n} (1 - x_i)$$

Here $N(i)$ is the set of dependent neighbors of A_i.

1.5 Proof of Symmetric Version

The heart of the proof is the following lemma:

Lemma 1. *For any* $S \subseteq \{1, 2, \ldots, n\}$ *and* $i \notin S$*:*

$$\Pr\left[A_i \mid \bigcap_{j \in S} \overline{A_j}\right] \leq 2p$$

Assuming the lemma above, the symmetric LLL follows as below:

$$\begin{aligned}
\Pr[\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n}] &= \Pr[\overline{A_1}] \cdot \Pr[\overline{A_2} \mid \overline{A_1}] \cdot \Pr[\overline{A_3} \mid \overline{A_1} \cap \overline{A_2}] \cdots \\
&\quad \cdot \Pr[\overline{A_n} \mid \overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_{n-1}}] \\
&= (1 - \Pr[A_1]) \cdot (1 - \Pr[A_2 \mid \overline{A_1}]) \\
&\quad \cdot (1 - \Pr[A_3 \mid \overline{A_1} \cap \overline{A_2}]) \cdots \\
&\geq (1 - 2p)^n > 0
\end{aligned}$$

Now we prove the lemma by induction on $|S|$.

Suppose $|S| = 0$. Then $\Pr[A_i] \leq p \leq 2p$.

Assume true for $|S| = k$. Consider $|S| = k+1$, $i \notin S$. Want to prove $\Pr[A_i \mid \bigcap_{j \in S} \overline{A_j}] \leq 2p$.

Let $S_{dep} = S \cap N(i)$ be the set of events in S that A_i is dependent on, and $S_{ind} = S \setminus S_{dep}$. Thus $S = S_{dep} \cup S_{ind}$.

If $|S_{ind}| = k + 1$ then $S_{dep} = \emptyset$ and:

$$\Pr[A_i \mid \bigcap_{j \in S} \overline{A_j}] = \Pr[A_i] \leq p \leq 2p$$

Thus we now consider $|S_{ind}| \leq k$.

We now use conditional probability or Bayes theorem. For events X, Y:

$$\Pr[X \mid Y] = \frac{\Pr[X \cap Y]}{\Pr[Y]}$$

and for events X, Y, Z:

$$\Pr[X \mid Y \cap Z] = \frac{\Pr[X \cap Y \mid Z]}{\Pr[Y \mid Z]}$$

Applying with $X = A_i$, $Y = \bigcap_{j \in S_{dep}} \overline{A_j}$ and $Z = \bigcap_{j \in S_{ind}} \overline{A_j}$, we have:

$$\Pr\left[A_i \mid \bigcap_{j \in S} \overline{A_j}\right] = \frac{\Pr\left[A_i \cap \left(\bigcap_{j \in S_{dep}} \overline{A_j}\right) \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right]}{\Pr\left[\bigcap_{j \in S_{dep}} \overline{A_j} \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right]}$$

Consider denominator. Via union bound:

$$\begin{aligned}
\Pr\left[\bigcap_{j \in S_{dep}} \overline{A_j} \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right] &= 1 - \Pr\left[\bigcup_{j \in S_{dep}} A_j \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right] \\
&\geq 1 - \sum_{j \in S_{dep}} \Pr\left[A_j \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right]
\end{aligned}$$

By induction hypothesis, since $|S_{ind}| \leq k$:

$$\Pr\left[A_j \mid \bigcap_{\ell \in S_{ind}} \overline{A_\ell}\right] \leq 2p$$

and $|S_{dep}| \leq d$.

Hence:

$$\Pr\left[\bigcup_{j \in S_{dep}} A_j \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right] \geq 1 - 2pd$$

Numerator is:

$$\begin{aligned}\Pr\left[A_i \cap \left(\bigcap_{j \in S_{dep}} \overline{A_j}\right) \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right] &\leq \Pr\left[A_i \mid \bigcap_{j \in S_{ind}} \overline{A_j}\right] \\ &= \Pr[A_i] \leq p\end{aligned}$$

Hence:

$$\Pr\left[A_i \mid \bigcap_{j \in S} \overline{A_j}\right] \leq \frac{p}{1 - 2pd} \leq 2p$$

2 Application of LLL

2.1 k-SAT

Recall that a k-SAT formula is a Boolean formula in CNF form with each clause having exactly k literals over distinct variables.

Theorem 3. *Let Φ be a k-SAT formula in which each variable occurs in at most $\frac{2^k}{k}$ clauses. Then Φ is satisfiable.*

Note that there is no limitation on number of variables or clauses.

Example 3. If $k = 10$ then it is requiring each variable to be in at most $\frac{2^{10}}{10} \approx 102$ clauses.

The theorem may not be interesting from a SAT perspective but is mainly to showcase the power of LLL and a setting in which it applies.

We prove this by considering a random assignment to the variables. Let A_i be the event that a clause C_i is not satisfiable (bad event), then $\Pr[A_i] = \frac{1}{2^k}$.

What does A_i depend on? C_i has k variables. Each variable that is in C_i is in at most $\frac{2^k}{k} - 1$ other clauses. So C_i shares a variable with at most $k \cdot \left(\frac{2^k}{k} - 1\right) = 2^k - k$ other clauses.

If C_i and C_j do not share variables then A_i and A_j are mutually independent.

Thus we can apply symmetric LLL with $p = \frac{1}{2^k}$ and $d = 2^k - k$.

Since $pd = \frac{2^k - k}{2^k} < \frac{1}{e}$:

$$\Pr\left[\bigcap_i \overline{A_i}\right] > 0$$

where m is the number of clauses.

Therefore $\Pr[\Phi \text{ is satisfiable}] > 0$.

2.2 Routing for Congestion Minimization

Recall that we saw the congestion minimization problem:

- $G = (V, E)$ directed graph
- $(s_1, t_1), \ldots, (s_k, t_k)$: k pairs that we want to connect by paths $P_1, P_2, \ldots, P_k$
- Minimize $\max_{e \in E} |\{i : e \in P_i\}|$ (congestion on e)

We used an LP relaxation and found a fractional routing that minimizes max fractional congestion.

A fractional routing for a pair (s_i, t_i) is a probability distribution over paths $p \in \mathcal{P}_i$ where $\mathcal{P}_i$ is the set of all s_i-t_i paths. We let x_p, $p \in \mathcal{P}$, be the amount of flow routed along p. We have $\sum_{p \in \mathcal{P}_i} x_p = 1$.

Suppose:

$$\sum_{i : p \in \mathcal{P}_i} \sum_{p \in \mathcal{P}_i : e \in p} x_p \leq 1$$

i.e., the max fractional congestion is at most 1.

Randomized rounding picks a path for each i independently according to the distribution $\{x_p : p \in \mathcal{P}_i\}$.

Then we used Chernoff bounds to show that $\Pr[\ell_e \geq c \log m]$ is at most $\frac{1}{m^2}$ for some sufficiently large constant c. Here ℓ_e is the load on e (the number of paths that use e).

Then via the union bound we see that with high probability $\max_{e \in E} \ell_e = O(\log m)$.

2.2.1 Better bound when paths are short

Now we will prove a better bound when paths are short.

Suppose $x_p = 0$ or $|p| \leq h$ where h is some parameter. In many applications h is a small constant independent of m, n. This implies locality because even if the graph is large, paths along which flow is routed are short.

Theorem 4. *There exists an integral routing where max congestion is $O(h \log h)$.*

Note: the bound does not depend on graph size.

In order to apply LLL we do some preprocessing. By discretization tricks we will assume that all x_p values that are non-zero have same value $\frac{1}{L}$ for some L. We may duplicate paths to achieve this. Thus each pair now has exactly L paths.

Then we do randomized rounding as before but with the discretized paths. So we pick one of the L paths for each pair.

How can we apply LLL here? Need to set up the events carefully.

Let C be the threshold of congestion we want to avoid. For edge e, let S_e be the set of all paths that use e.

Define $A_{e,S}$ for $e \in E$, $S \subseteq S_e$, $|S| = C$ to be the event that S is the set of paths chosen.

Claim 2. *If S contains two paths from path collection of same pair, then* $\Pr[A_{e,S}] = 0$. *Otherwise it is equal to $\frac{1}{L^C}$.*

We use notation $S = \{(i_1, j_1), (i_2, j_2), \ldots, (i_C, j_C)\}$ where $i_\ell \in \{1, 2, \ldots, K\}$ indicates the pair and $j_\ell \in \{1, 2, \ldots, L\}$ to denote the index of the path in the L paths for the pair. We order the paths in some fashion for each pair. We let P_i^j denote the j-th path for pair i.

Dependencies: Fix two events $A_{e,S}$ and $A_{e',S'}$ where S has paths from distinct pairs and similarly S'.

Suppose $S = \{(i_1, j_1), (i_2, j_2), \ldots, (i_C, j_C)\}$ and $S' = \{(i'_1, j'_1), (i'_2, j'_2), \ldots, (i'_C, j'_C)\}$.

If $\{i_1, i_2, \ldots, i_C\} \cap \{i'_1, i'_2, \ldots, i'_C\} = \emptyset$, i.e., the pairs don't overlap, then $A_{e,S}$ is independent of $A_{e',S'}$.

We need to understand how many other events does $A_{e,S}$ depend on. Suppose $e' \in E$, $S' \subseteq S_{e'}$, and $|S'| = C$ with $\{i_1, i_2, \ldots, i_C\} \cap \{i'_1, i'_2, \ldots, i'_C\} \neq \emptyset$. Say $i_1 = i'_1$.

How many choices of e', $\{i'_2, \ldots, i'_C\}$ and $\{j'_1, j'_2, \ldots, j'_C\}$ do we have?

L choices for j'_1. Fix one such j'_1.

Then we have path $P_{i_1}^{j'_1}$ and this path has at most h edges.

So e' has h choices. For each such edge, at most L paths use the edge (since total flow on each edge ≤ 1). And we can choose any $C-1$ paths from those L.

Thus for fixed choice of j'_1 there are $h \cdot \binom{L}{C-1}$ choices.

Hence total is $L \cdot h \cdot \binom{L}{C-1}$.

There are C choices of pair overlap between S and S'.

Hence the neighborhood size of $A_{e,S}$ in the dependency graph is at most $C \cdot L \cdot h \cdot \binom{L}{C-1}$.

To apply LLL we have $\Pr[A_{e,S}] = \frac{1}{L^C} = p$ and $d = C \cdot L \cdot h \cdot \binom{L}{C-1}$.

To ensure $pd \leq \frac{1}{e}$ we need:

$$\frac{1}{L^C} \cdot C \cdot L \cdot h \cdot \binom{L}{C-1} \leq \frac{1}{e}$$

It suffices to have $C = O(h \log h)$.

Note that if no bad event $A_{e,S}$ happens, then congestion is at most $C-1$.