1 Lovász Local Lemma

Usually, when we use the probabilistic method, the desired event holds with very high probability. On the other hand, suppose we have n mutually independent events, each of which holds with probability p, then the conjunction of all of them holds with probability p^n. This is exponentially small but strictly positive.

In most applications, the desired event cannot be decomposed into n mutually independent ones. The Lovász Local Lemma provides a way to deal with dependencies. It was first found by Erdős and Lovász in 1975 and is extremely powerful, especially if the dependencies are rare.

Let A_1, A_2, ..., A_n be events in an arbitrary probability space. A graph $D = (V, E)$ on the set of vertices $V = [n]$ is called *the dependency graph* for the events $A_1, \ldots, A_n$ if for each $i \in [n]$, the event A_i is mutually independent of all events that are *not* in $\Gamma^+(i) := \Gamma(i) \cup \{i\}$, where $\Gamma(i) = \{j : (i, j) \in E\}$.

Lemma 1 (The Local Lemma). *Suppose $D = ([n], E)$ is a dependency graph for events $A_1, \cdots, A_n$ and there are real numbers $x_1, \cdots, x_n$ such that $0 \le x_i < 1$ and for all $i \in [n]$*

$$\Pr(A_i) \le x_i \prod_{j \in \Gamma(i)} (1 - x_j). \tag{1}$$

Then

$$\Pr\left(\bigwedge_{i=1}^{n} \overline{A_i}\right) \ge \prod_{i=1}^{n} (1 - x_i) > 0.$$

One should think A_i's as "bad" events, and what we are after is some "perfect" object that avoids all undesired events. In particular, when the condition of Lemma 1 holds, with positive probability, no event A_i holds and thus a perfect object exists.

Proof of Lemma 1. We first claim that for any $S \subseteq \{1, \cdots, n\}$ such that $i \notin S$,

$$\Pr\left(A_i \,\middle|\, \bigwedge_{j \in S} \overline{A_j}\right) \le \Pr(A_i) \prod_{j \in \Gamma(i)} (1 - x_j)^{-1}. \tag{2}$$

We prove the inequality by induction on $s = |S|$. The base case is when S is empty and the claim holds trivially.

For the induction step, let $S_1 = S \cap \Gamma(i)$ and $S_2 = S \setminus S_1$. If $S_1 = \emptyset$, then the lemma holds trivially as A_i is independent from S in this case. Otherwise S_2 is a proper subset of S. For the set S, define a new event $B(S) := \bigwedge_{i \in S} \overline{A_i}$, which is the event that none of S occurs. We have that

$$
\begin{aligned}
\Pr(A_i \mid B(S)) &= \frac{\Pr(A_i \wedge B(S_1) \mid B(S_2))}{\Pr(B(S_1) \mid B(S_2))} \\
&\leq \frac{\Pr(A_i \mid B(S_2))}{\Pr(B(S_1) \mid B(S_2))} \\
&= \frac{\Pr(A_i)}{\Pr(B(S_1) \mid B(S_2))},
\end{aligned}
\tag{3}
$$

where the last line is because A_i is independent from $B(S_2)$. We then use the induction hypothesis to bound the denominator. Suppose $S_1 = \{j_1, j_2, \ldots, j_r\}$ for some $r > 0$. Then,

$$
\begin{aligned}
\Pr(B(S_1) \mid B(S_2)) &= \Pr\left(\bigwedge_{j \in S_1} \overline{A_j} \;\middle|\; \bigwedge_{j \in S_2} \overline{A_j} \right) \\
&= \prod_{t=1}^{r} \Pr\left(\overline{A_{j_t}} \;\middle|\; \bigwedge_{s=1}^{t-1} \overline{A_{j_s}} \wedge \bigwedge_{j \in S_2} \overline{A_j} \right) \\
&= \prod_{t=1}^{r} \left(1 - \Pr\left(A_{j_t} \;\middle|\; \bigwedge_{s=1}^{t-1} \overline{A_{j_s}} \wedge \bigwedge_{j \in S_2} \overline{A_j} \right) \right).
\end{aligned}
$$

By the induction hypothesis and (1), we have that for any $1 \leq t \leq r$,

$$
\begin{aligned}
\Pr\left(A_{j_t} \;\middle|\; \bigwedge_{s=1}^{t-1} \overline{A_{j_s}} \wedge \bigwedge_{i \in S_2} \overline{A_j} \right) &\leq \Pr(A_{j_t}) \prod_{j \in \Gamma(j_t)} (1 - x_j)^{-1} \\
&\leq x_{j_t} \prod_{j \in \Gamma(j_t)} (1 - x_j) \prod_{j \in \Gamma(j_t)} (1 - x_j)^{-1} \\
&= x_{j_t}.
\end{aligned}
$$

Thus,

$$
\Pr(B(S_1) \mid B(S_2)) \geq \prod_{j \in S_1} (1 - x_j) \geq \prod_{j \in \Gamma(i)} (1 - x_j).
$$

This together with (3) proves the claim (2).

The lemma follows easily. First, by (2) and (1),

$$
\Pr\left(A_i \;\middle|\; \bigwedge_{j \in S} \overline{A_j} \right) \leq \Pr(A_i) \prod_{j \in \Gamma(i)} (1 - x_j)^{-1} \leq x_i.
$$

2

Therefore,

$$\Pr\left(\bigwedge_{i=1}^{n}\overline{A_i}\right) = \left(1 - \Pr(A_1)\right)\left(1 - \Pr\left(A_1 \mid \overline{A_2}\right)\right)\cdots\left(1 - \Pr\left(A_n \mid \bigwedge_{i=1}^{n-1}\overline{A_i}\right)\right)$$

$$\geq \prod_{i=1}^{n}(1 - x_i). \qquad \square$$

The inequality (2) actually holds for any event E (not necessarily one of A_i) if we define the "neighbourhood" of E appropriately. This observation is useful in some applications, but we will not use it.

The symmetric case of Lemma 1 is often useful.

Corollary 2. *Suppose D is the dependency graph and $|\Gamma(i)| \leq d$ for any $i \in [n]$. If*

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

then

$$\Pr\left(\bigwedge_{i=1}^{n}\overline{A_i}\right) > 0.$$

Proof. If $d = 0$ then the corollary is trivial. Otherwise, take $x_i = \frac{1}{d+1} < 1$. The corollary follows trivially since for any $d \geq 1$, $\left(1 - \frac{1}{d+1}\right)^d > \frac{1}{e}$. $\qquad \square$

As shown by Shearer in 1985, the constant e in Corollary 2 is the best possible. Indeed, Shearer 1985 gives the optimal condition for the Local Lemma with a fixed dependency graph and the probability vector of the events. Using more complicated arguments about zeros of the independence polynomial, Scott and Sokal 2005 improved the condition of Corollary 2 into $p \leq \frac{(d-1)^{d-1}}{d^d} = \frac{1}{d-1} \cdot \left(1 - \frac{1}{d}\right)^d \leq \frac{1}{e(d-1)}$.

2 Hypergraph 2-colourability

As the first example, let us see a straightforward application of Corollary 2. Recall that a hypergraph $H = (V, E)$ is 2-colourable if there is a 2-vertex-colouring such that no edge $e \in E$ is monochromatic.

Theorem 3. *Let $H = (V, E)$ be a k-uniform hypergraph with maximum degree d. $(k \geq 2)$ If $edk \leq 2^{k-1}$, then H is 2-colourable.*

Proof. We still colour every vertex uniformly at random. The bad events are A_e, which denotes that an edge $e \in E$ is monochromatic. Thus,

$$\Pr(A_e) = \frac{2}{2^k} = 2^{1-k}.$$

This is p in Corollary 2.

To apply Corollary 2, we need to upper bound the number of events A_f that are correlated with A_e. If e and f do not share any vertex, then A_e and A_f are independent. Thus, the maximum degree Δ of the dependency graph is the maximum number of edges that e intersect with for any e. Every vertex has maximum degree d, which implies that there are at most $k(d-1)$ many edges intersecting with e. Hence, $\Delta \leq k(d-1)$ and $\Delta + 1 \leq kd$. The lemma follows from Corollary 2 as $ep(\Delta + 1) \leq e2^{1-k}dk \leq 1$. □

3 Ramsey numbers: re-re-visited

We can use the Local Lemma to obtain yet another lower bound for Ramsey numbers $R(k, k)$. Consider a uniformly at random 2-colouring of the edges of K_n. For each subset $S \subset [n]$ of vertices of size k, let A_S be the "bad" events where S is monochromatic. Then we see that

$$\Pr(A_S) = 2^{1-\binom{k}{2}}.$$

On the other hand, A_S and A_T are dependent if and only if $|S \cap T| \geq 2$, as this is the only case when S and T share at least one edge. For a fixed S, there are at most $\binom{k}{2}\binom{n-2}{k-2}$ T such that $|S \cap T| \geq 2$.[1] Thus, to apply Corollary 2, we have that $p = 2^{1-\binom{k}{2}}$ and $\Delta \leq \binom{k}{2}\binom{n-2}{k-2}$.

Theorem 4. *If* $e\binom{k}{2}\binom{n-2}{k-2}2^{1-\binom{k}{2}} \leq 1$, *then* $R(k, k) > n$.

Calculation yields that $R(k, k) > \frac{\sqrt{2}}{e}(1 + o(1))k2^{k/2}$. This is yet another factor $\sqrt{2}$ improvement upon the alteration method, and a factor 2 improvement of the basic method. Why the improvement is small? The reason is that the local lemma works best when the dependency is rare. On the other hand, in this application, the dependency is rather high.

Following this intuition, we should expect much better improvement for off diagonal Ramsey numbers $R(\ell, k)$ when ℓ is small. Let us do $\ell = 3$ (the rest of this section follows from Spencer 1977). Of course, we will have to apply the asymmetric version Lemma 1. In particular, there are two types of "bad" events and we need to find a real number x_i for each type.

Colour every edge blue with probability p and red $1-p$. Let A_S (or B_T) be the event that S is blue (or T is red). If $|S| = 3$, then $\Pr(A_S) = p^3$. If $|T| = k$, then $\Pr(B_T) = (1-p)^{\binom{k}{2}}$. Construct the dependency graph for events A_S and B_T by joining two vertices if the two subset of vertices share more than 2 vertices (A_S with $A_{S'}$, A_S with B_T, or B_T with $B_{T'}$).

Each A_S vertex is adjacent to $3(n-3) < 3n$ $A_{S'}$ vertices, and at most $\binom{n}{k}$ B_T vertices. Similarly, each B_T vertex is adjacent to at most $\binom{k}{2}(n-2) < \frac{k^2 n}{2}$ A_S vertices and at most $\binom{n}{k}$ $B_{T'}$ vertices. To apply Lemma 1, we need to find the probability $0 < p < 1$ and two real numbers $0 < x < 1$ and $0 < y < 1$ such that

$$p^3 \leq x(1-x)^{3n}(1-y)^{\binom{n}{k}},$$

---

[1] This is an over counting! If $|S \cap T| \geq 3$, then T is counted more than once.

and

$$(1-p)^{\binom{k}{2}} \le y(1-x)^{k^2 n/2}(1-y)^{\binom{n}{k}}.$$

If so, $R(3,k) \ge n$.

The calculation is elementary but tedious (for details, see Spencer 1977). The best choices are $p = \Theta(n^{-1/2})$, $k = \Theta(n^{1/2} \log n)$, $x = \Theta(n^{-3/2})$, and $y = \Theta(1/\binom{n}{k})$. It yields that $R(3,k) \ge \Omega\left(\left(\frac{k}{\log k}\right)^2\right)$. A similar argument yields that $R(4,k) \ge k^{5/2+o(1)}$, improving the bound of $k^{2-o(1)}$ by the alteration method. In fact, the bound on $R(4,k)$ is the best we know and is better than any other bound without using the Local Lemma.

The bound $R(3,k) \ge \Omega\left(\left(\frac{k}{\log k}\right)^2\right)$ matches a result by Erdős 1961 using a highly compli-cated probabilistic method. It is also close to the correct answer, which is $R(3,k) = \Theta(\frac{k^2}{\log k})$. The upper bound is due to Ajtai, Komlós, Szemerédi 1980, and the lower bound is due to Kim 1995.

To get some intuition of the upper bound, one can show, using induction, that

$$R(\ell, k) \le \binom{k + \ell - 2}{\ell - 1}.$$

In particular, $R(3,k) \le \frac{k(k+1)}{2}$.

4 Lopsided Lovász Local Lemma

In the proof of Lemma 1, we can replace the condition

$$\Pr(A_i) \le x_i \prod_{(i,j) \in E} (1 - x_j),$$

by the weaker assumption that for each i and $S \subseteq [n] \setminus \Gamma^+(i)$,

$$\Pr\left(A_i \,\middle|\, \bigwedge_{j \in S} \overline{A_j}\right) \le x_i \prod_{(i,j) \in E} (1 - x_j).$$

Notice that here S does not intersect the neighbourhoods of i, but we do not require S to be be independent from A_i. Indeed, it is fine if S is positively correlated with A_i.

Formally, we define the negative dependency graph.

Definition 1. D is a negative dependency graph *if for every event A_i and every subset $J \subseteq [n] \setminus \Gamma^+(i)$, we have that*

$$\Pr\left(A_i \wedge \bigvee_{j \in J} A_J\right) \ge \Pr(A_i) \Pr\left(\bigvee_{j \in J} A_J\right). \tag{4}$$

5

Using the negative dependency graph, we have the Lopsided version of Lemma 1.

Lemma 5 (Lopsided Lovász Local Lemma (LLLL)). *Suppose $D = ([n], E)$ is a negative dependency graph for events $A_1, \cdots, A_n$ and there are real numbers $x_1, \cdots, x_n$ such that $0 \leq x_i < 1$ and for all $i \in [n]$*

$$\Pr(A_i) \leq x_i \prod_{j \in \Gamma(i)} (1 - x_j).$$

Then

$$\Pr\left(\bigwedge_{i=1}^{n} \overline{A_i}\right) \geq \prod_{i=1}^{n}(1 - x_i) > 0.$$

The proof of Lemma 5 is exactly the same as Lemma 1.

When the overall dependency is high but negative dependency is rare, Lemma 5 is much more powerful than the vanilla version Lemma 1. It is shown in the next example.

4.1 Latin Transversal

Definition 2. *Let $M \in \mathbb{Z}^{n \times n}$ be a matrix. We call a permutation π a Latin transversal of M if $M_{i,\pi(i)} \neq M_{j,\pi(j)}$ if $i \neq j$.*

For example, the identity permutation is a Latin transversal of the following matrix.

3	2	1
2	1	3
1	3	2

The next theorem shows that if no integer appears too often in M, then there is a Latin transversal. It is due to Erdős and Spencer 1991.

Theorem 6. *If no integer appears in more than $\frac{n}{4e}$ entries of M, then M has a Latin transversal.*

Proof. Let π be a uniform permutation. The "bad" events correspond to pairs that $M_{i,i'} = M_{j,j'}$ where $i \neq j$ and $i' \neq j'$. For each such pair, define the event

$$A_{ii'jj'} := \{\pi \mid \pi(i) = i' \ \wedge \ \pi(j) = j'\}.$$

Then

$$\Pr(A_{ii'jj'}) = \frac{1}{n(n-1)}.$$

6

The negative dependency graph D is defined on these 4-tuples. Two vertices $ii'jj'$ and $kk'\ell\ell'$ are adjacent if and only if

$$\{i, i'\} \cap \{k, k'\} \neq \emptyset \qquad \text{or} \qquad \{j, j'\} \cap \{\ell, \ell'\} \neq \emptyset.$$

We claim D is a negative dependency graph. Use $a, b, \cdots$ to denote the tuples of the form (i, i', j, j'). We need to verify (4); namely, for any subset $J \subseteq V \setminus \Gamma^+(a)$.

$$\Pr\left(A_a \,\middle|\, \bigvee_{b \in J} A_b\right) \geq \Pr(A_a) = \frac{1}{n(n-1)}. \tag{5}$$

We will show that for $a = (1, 1, 2, 2)$ and a fixed $J \subseteq V \setminus \Gamma^+(a)$. Due to symmetry, the proof for any other tuple is exactly the same. Let P_J be the set of permutations such that $\bigvee_{b \in J} A_b$ holds, and for any $i \neq j$,

$$S_{i,j} := \{\pi \mid \pi \in P_J \text{ s.t. } \pi(1) = i, \ \pi(2) = j\}.$$

Since

$$\Pr\left(A_a \,\middle|\, \bigvee_{b \in J} A_b\right) = \frac{\Pr\left(A_a \wedge \bigvee_{b \in J} A_b\right)}{\Pr\left(\bigvee_{b \in J} A_b\right)} = \frac{|S_{1,2}|}{\sum_{i \neq j} |S_{i,j}|},$$

what we desire, (5), will follow if we show that $|S_{1,2}| \geq |S_{i,j}|$.

To show $|S_{1,2}| \geq |S_{i,j}|$, consider the following mapping $T : S_{i,j} \to S_{1,2}$:

$$T(\pi) = (1, i)(2, j)\pi.$$

In other words, $T(\pi)$ is a permutation that first apply π, and then swap 1 with i and 2 with j. We claim that $T(\pi) \in S_{1,2}$ if $\pi \in S_{i,j}$. This is because:

1. As $\pi \in S_{i,j}$, $\pi(1) = i$ (or $\pi(2) = j$) and $T(\pi)(1) = 1$ (or $T(\pi)(2) = 2$).

2. The image $T(\pi) \in P_J$. Note that $\pi \in P_J$ implies that π satisfies A_b for some tuple $b = (k, k', \ell, \ell') \in J$, which indicates $\pi(k) = k'$ and $\pi(\ell) = \ell'$. Moreover, $1, 2 \notin \{k, k', \ell, \ell'\}$ by the definition of J. It implies $k' \neq i, j$ and $\ell' \neq i, j$ since $\pi^{-1}(i) = 1$ and $\pi^{-1}(j) = 2$. Thus, $T(\pi)(k) = \pi(k) = k'$ and $T(\pi)(\ell) = \pi(\ell) = \ell'$. In other words, $T(\pi)$ satisfies A_b as well and $T(\pi) \in P_J$.

Next we claim that T is injective. Suppose otherwise, then there are two different permutations $\pi, \pi' \in S_{i,j}$, whose images $T(\pi) = T(\pi')$. Since π and π' are distinct, there is some $k \neq 1, 2$ such that $\pi(k) \neq \pi'(k)$. It is easy to see that $T(\pi)(k) \neq T(\pi')(k)$, a contradiction.

We still need to verify the condition of LLLL. This is a symmetric case. For any event $A_{ii'jj'}$, its correlated events $A_{kk'\ell\ell'}$ can be determined by the following:

1. First choose a pair kk' or $\ell\ell'$ such that it lies in the same row/column as either ii' or jj'. There are $n + n + (n - 2) + (n - 2) = 4(n - 1)$ choices.

7

2. Suppose we chose kk'. Since every integer appears at most $\frac{n}{4e}$ times in the matrix, there are at most $\frac{n}{4e} - 1$ entries $\ell\ell'$ such that $M_{kk'} = M_{\ell\ell'}$.

In summary, the number of neighbours of $A_{ii'jj'}$ is at most $4(n-1)\cdot\left(\frac{n}{4e} - 1\right) = \frac{n(n-1)}{e} - 4(n-1)$. Thus, we see that

$$ep(\Delta + 1) \leq e \cdot \frac{1}{n(n-1)} \cdot \frac{n(n-1)}{e} = 1.$$

The theorem follows from Lemma 5. $\qquad\square$