L9 The Lovász Local Lemma

$n \in \mathbb{N}$

$[n] = \{1, 2, \ldots, n\}$, $A_i \in \Sigma$ for $i \in [n]$ – events in a prob. space. $D = ([n], E)$ ($E \subset [n] \times [n]$) is a <u>dependency digraph</u> for the ~~ev~~ events $A_1, \ldots, A_n$ if:

$\forall i \in [n]: J \subset \{j \in [n] \mid (i,j) \notin E\}$ (i.e., some non-neighbours of i) $\Rightarrow \Pr(A_i \cap \bigcap_{j \in J} A_j) = \Pr(A_i) \cdot \Pr(\bigcap_{j \in J} A_j)$ $\cancel{\prod_{j \in J} \Pr(A_j)}$

– A_i is ~~mutually~~ independent of all the events $\{A_j \mid (i,j) \notin E\}$, of all "non-neighbours" of A_i.

also means in terms of $\Pr(.)$

Here one should add that there is a lemma analogous to that of the previous lecture:

Lemma A is ~~mutually~~ independent of $A_1, \ldots, A_n$ $\Rightarrow$ A is ~~not~~ independent of $A'_1, A'_2, \ldots, A'_n$ where each $A'_i = A_i$ or $\overline{A_i}$.

<u>Proof.</u> Like that of the previous lemma $\Pr(A \cap B) = \Pr(A)\Pr(B) \Rightarrow \Pr(A \cap \overline{B}) =$
$= \Pr(A) - \Pr(A \cap B) = \Pr(A) - \Pr(A)\Pr(B) = \Pr(A)(1 - \Pr(B)) =$
$= \Pr(A)\Pr(\overline{B})$, … – exercise for you. $\boxtimes$

Now back to the statement of the LLL.

$\boxed{\text{Thm. (LLL-general } \cancel{\text{case}} \text{ form)}}$ $n \in \mathbb{N}$, $A_1, \dots, A_n \in \Sigma$ are events in a pr. space, $D = ([n], E)$ is a dep. digraph of $A_1, \dots, A_n$, $x_1, \dots, x_n \in [0,1)$, $\forall i \in [n]: \Pr(A_i) \leq x_i \prod_{\substack{j \\ (i,j) \in E}} (1 - x_j)$.

Then $\Pr\left(\bigcap_{i=1}^{n} \bar{A}_i\right) \geq \prod_{i=1}^{n} (1 - x_i) > 0$.

$\boxed{\text{Thm. (LLL-symmetric form)}}$ $p \in [0,1]$

$n \in \mathbb{N} (= \{1, 2, \dots\})$, $A_1, \dots, A_n \in \Sigma$ are events in a pr. space, $d \in$ ~~and~~ $\forall i \in [n]: A_i$ is ~~not~~ independent of ~~of~~ all of $A_1, \dots, A_n$ but at most d events, and $\Pr(A_i) \leq p$ for every $i \in [n]$. Then

$$\underbrace{e}_{2.71828\dots} p(d+1) \leq 1 \Rightarrow \Pr\left(\bigcap_{i=1}^{n} \bar{A}_i\right) > 0.$$

Proof of the general form: see the Pr. Method of N. Alon & J. H. Spencer, pp. 54–55 (1992 edition); not good, they write things like $\Pr(A_i | \bigwedge_{j \in S} \bar{A}_j)$ without justifying that $\Pr(\bigwedge_{j \in S} \bar{A}_j) > 0$.

Instead, we prove a version of the symm. LLL from Mitzenmacher and Upfal, with weaker condition

$\boxed{4dp \leq 1.}$ **Proof.** Let $S \subset [n]$, By induction on $s = 0, 1, \dots, n-1$:

(*) $|S| \leq s$, $k \notin S \Rightarrow \Pr(A_k \mid \underbrace{\bigcap_{j \in S} \bar{A}_j}_{:= \Omega \text{ if } S = \emptyset}) \leq 2p$. For $S \neq \emptyset$ we need to show that also $\Pr(\bigcap_{j \in S} \bar{A}_j) > 0$.

If $s = 0$ ✓ $(\Pr(A_k) \leq p)$. For the ind. step we first prove – true for $s = 1$, then $\Pr(\bar{A}_j) \geq 1 - p > 0$. For $s > 1$, wlog $S = [s]$ and we have

$$\Pr\Big(\bigcap_{i=1}^{s} \bar{A}_i\Big) = \prod_{i=1}^{s} \frac{\Pr\big(\bigcap_{j=1}^{i} \bar{A}_j\big)}{\underbrace{\Pr\big(\bigcap_{j=1}^{i-1} \bar{A}_j\big)}_{>0 \text{ by induction}}} = \quad (\Pr(\bar{X}) = 1 - \Pr(X))$$

$\left(\frac{A}{D} = \frac{A}{B} \cdot \frac{B}{C} \cdot \frac{C}{D} \quad (D = 1)\right)$

$$= \prod_{i=1}^{s} \frac{\Pr\big(\bigcap_{j=1}^{i-1} \bar{A}_j\big) - \Pr\big(A_i \cap \bigcap_{j=1}^{i-1} \bar{A}_j\big)}{\Pr\big(\bigcap_{j=1}^{i-1} \bar{A}_j\big)} \underset{\text{condit. probab.}}{=}$$

$$= \prod_{i=1}^{s} \Big(1 - \underbrace{\Pr\Big(A_i \,\Big|\, \bigcap_{j=1}^{i-1} \bar{A}_j\Big)}_{\leq 2p \text{ by the inductive assump.}}\Big) \geq \prod_{i=1}^{s} (1 - 2p) > 0.$$

Let $S_1 := \{ j \in S \mid (2, j) \in E \}$, $S_2 := S \setminus S_1$. ④

$S_2 = S \Rightarrow A_2$ is ~~mut.~~ indep. of the $\bar{A}_i$, $i \in S$,

and $Pr(A_2 \mid \bigcap_{j \in S} \bar{A}_j) = \dots = Pr(A_2) \leq p < 2p$.

Let $|S_2| < \Delta$. We introduce the notation

$F_S := \bigcap_{j \in S} \bar{A}_j$ and similarly for F_{S_1} and F_{S_2}. So

$F_S = F_{S_1} \cap F_{S_2}$ and $Pr(A_2 \mid F_S) = \dfrac{\overset{N}{Pr(A_2 \cap F_S)}}{\underset{D}{Pr(F_S)}}$.

$N = Pr(A_2 \cap F_S) = Pr(A_2 \cap F_{S_1} \cap F_{S_2}) =$

$= Pr(A_2 \cap F_{S_1} \mid F_{S_2}) \underbrace{Pr(F_{S_2})}_{>0}$ ~~and by dividing with~~

~~$Pr(A_2 \mid F_S)$~~ $D = Pr(F_S) = \cancel{Pr(F_{S_1})} \cap$

$= Pr(F_{S_1} \cap F_{S_2}) =$

$= Pr(F_{S_1} \mid F_{S_2}) \underbrace{Pr(F_{S_2})}_{>0} \rightsquigarrow Pr(A_2 \mid F_S) =$

$= \dfrac{N}{D} = \dfrac{\overset{N}{Pr(A_2 \cap F_{S_1} \mid F_{S_2})}}{\underset{D}{Pr(F_{S_1} \mid F_{S_2})}}$. (holds for $S_2 = \emptyset$ as well)

5

$$N = \Pr(A_k \cap F_{S_1} \mid F_{S_2}) \underset{\Pr(A\cap B) \le \Pr(A)}{\le} \Pr(A_k \mid F_{S_2}) \underset{S_2 \text{ are non-neighb.}}{=} \Pr(A_k) \le p.$$

Since $|S_2| < |S| = s$, $\underbrace{\text{by induction:}}$

$$D = \Pr(F_{S_1} \mid F_{S_2}) \overset{def}{=} \Pr\Big(\bigcap_{i\in S_1} \bar{A}_i \,\Big|\, \bigcap_{j\in S_2} \bar{A}_j\Big) \ge$$

$$\underset{\substack{\Pr(\bar A \cap \bar B) = \Pr(\overline{A\cup B}),\\ \text{de Morgan}}}{=} \Pr\Big(\overline{\bigcup_{i\in S_1} A_i} \,\Big|\, \bigcap_{j\in S_2} \bar{A}_j\Big) \underset{\substack{\Pr(\bar A) = 1-\Pr(A)\\ \text{and the union bound}}}{\ge} 1 - \sum_{i\in S_1} \Pr\Big(A_i \,\Big|\, \bigcap_{j\in S_2} \bar{A}_j\Big)$$

$$\ge 1 - \sum_{i\in S_1} 2p \underset{|S_1|\le d}{\ge} 1 - 2pd \underset{\text{by } (*)}{\ge} \frac{1}{2}.$$ Thus indeed

$$\Pr(A_k \mid F_S) = \frac{N}{D} \le \frac{p}{1/2} = 2p.$$

$$\Pr(A_k \mid F_S) = \Pr\Big(A_k \,\Big|\, \bigcap_{j\in S} \bar{A}_j\Big) \quad (k \notin S)$$

To conclude, $\Pr\Big(\bigcap_{i=1}^{n} \bar{A}_i\Big) \underset{\text{recall from above}}{=} \prod_{i=1}^{n}\Big(1 - \Pr\Big(A_i \,\Big|\, \bigcap_{j=1}^{i-1} \bar{A}_j\Big)\Big) \ge$

$$\ge \prod_{i=1}^{n} (1-2p) > 0,$$ as $2p < 1$ by $(*)$. $\boxtimes$

The LLL is from the article by • Paul (Pál) Erdős (1913 – 1996) and • László Lovász (1948) in 1975. In 1991, • József Beck (1952) [not to be confused with the Polish politician • Józef Beck (1894 – 1944)] found efficient algorithmization of LLL – typical application is an efficient ($\in \mathbf{P}$) randomized, or even deterministic, algorithm that for a given hypergraph with low degree finds a proper coloring. The algorithmic LLL was much improved in 2010 by R. A. Moser and • Gábor Tardos (1964) [arXiv:0903.0544, see also the Wikipedia article ~~Algo~~ Algorithmic Lovász local lemma] – • Gödel Prize 2020!

Chernoff's, or Černov's, bounds

[in fact, after • Herman Chernoff (1923), cf. • Leopold Vietoris (1891 – 2002) – wrote a paper on trig. sums at the age 103]

Recall the notion of a <u>random variable</u>

[for a prob. space (Ω, Σ, Pr), it is a function $X: \Omega \to \mathbb{R}$ that is <u>(Σ-) measurable</u>, which means that $\forall a \in \mathbb{R}: \{\omega \in \Omega \mid X(\omega) \le a\} \in \Sigma$] and of <u>independence of r. variables</u>:

[real r. variables X, Y are <u>independent</u> if $\forall a, b \in \mathbb{R}$: $Pr[X \le a \;\&\; Y \le b] =$

$$= Pr[X \le a]\, Pr[Y \le b]]$$

$= Pr(\{\omega \in \Omega \mid X(\omega) \le a\})$

Similarly for several r. variables $X_1, X_2, \dots, X_n$

$\boxed{\text{Theorem}}$ (essentially Bernstein, 1924)

Бернштейн

[• Sergej N. Bernštejn (1880–1968)]

$X_1, \dots, X_n$ are ± 1-valued r. variables, independent, $Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}$ and $\boxed{X := X_1 + X_2 + \dots + X_n}$; then $\forall$ real $t > 0$:

$$Pr(X \ge t) < e^{-t^2/2\sigma^2} \text{ and } Pr(X \le -t) < \text{ —||— },$$

where $\sigma = \sqrt{n}$.

~~Where~~ ($\sigma = \sqrt{Var[X]} = \sqrt{n}$.)

Proof. Just the 1^{st} ineq., the other follows by symmetry (which means really: ...). We set $Y := e^{uX}$ where $u > 0$ is to be determined. ← u, not μ!

$Pr[X \geq t] = Pr[Y \geq e^{ut}]$. By Markov's ineq.

$$Pr[Y \geq q] \leq \frac{\mathbb{E}Y}{q} \quad (q > 0).$$

But

$$\mathbb{E}Y = \mathbb{E}[\exp(u \cdot (X_1 + X_2 + \dots + X_n))] = \mathbb{E}\prod_{i=1}^{n} e^{uX_i} \overset{\text{independence of } X_i}{=} \prod_{i=1}^{n} \mathbb{E}e^{uX_i} = \left(\frac{e^{u}+e^{-u}}{2}\right)^{n} \leq e^{nu^2/2}.$$

Because $\frac{1}{2}(e^{x}+e^{-x}) = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots \leq 1 + \frac{1}{1!}\frac{x^2}{2} + \frac{1}{2!}\left(\frac{x^2}{2}\right)^2 + \dots = e^{x^2/2}$

Hence

$$Pr[Y \geq e^{ut}] \leq \frac{\mathbb{E}Y}{e^{ut}} \leq e^{nu^2/2 - ut}.$$

The last expression is minimized by setting $u = \frac{t}{n}$, thus

$$Pr(X \geq t) \leq e^{-t^2/2n} \quad \left(= e^{-t^2/2\sigma^2}\right) \qquad \boxtimes$$

An obvious application is, of course, coin flipping. $X_1, \dots, X_n$ tosses of a fair coin ($Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}$). Then for $X := X_1 + \dots + X_n$ (n tosses) we have:

$$\cancel{Pr(|X - \tfrac{n}{2}|} \quad Pr(|X| \geq \tfrac{n}{4}) \leq 2e^{-(\frac{n}{4})^2/2n} =$$

$$= 2e^{-n/32}. \text{ or } Pr(|X| \geq 10\sqrt{n}) \leq$$

$$\leq 2e^{-(10\sqrt{n})^2/2n} = 2e^{-50}$$ – much stronger bounds than those provided by Chebyshev's ineq.

More general $\boxed{\text{Theorem}}$ (from Matoušek and Vondrák) (see also Appendix in Alon & Spencer) $X_1, X_2, \dots, X_n$ indep. r. variables $X_i \in [0,1]$, $X := X_1 + X_2 + \dots + X_n$, $\sigma^2 = \operatorname{Var} X \overset{!}{=} \sum_{i=1}^{n} \operatorname{Var} X_i$. Then $\forall$ real $t \geq 0$:

$$Pr(X \geq \mathbb{E}X + t) < e^{-\frac{t^2}{2(\sigma^2 + t/3)}}$$

and $$Pr(X \leq \mathbb{E}X - t) < \text{— || —}.$$

Thank you.