Probabilistic Method and Random Graphs

Lecture 7. Second Moment Method and Lovász Local Lemma

Xingwu Liu

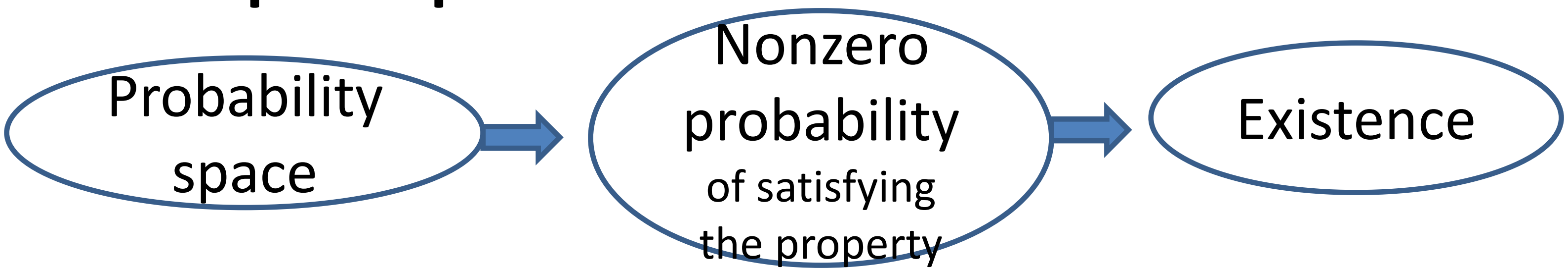
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

[1]The slides are mainly based on Chapter 5 of Probability and Computing.

Comments, questions, or suggestions?

A Review of Lecture 6

- **Principle of probabilistic method**



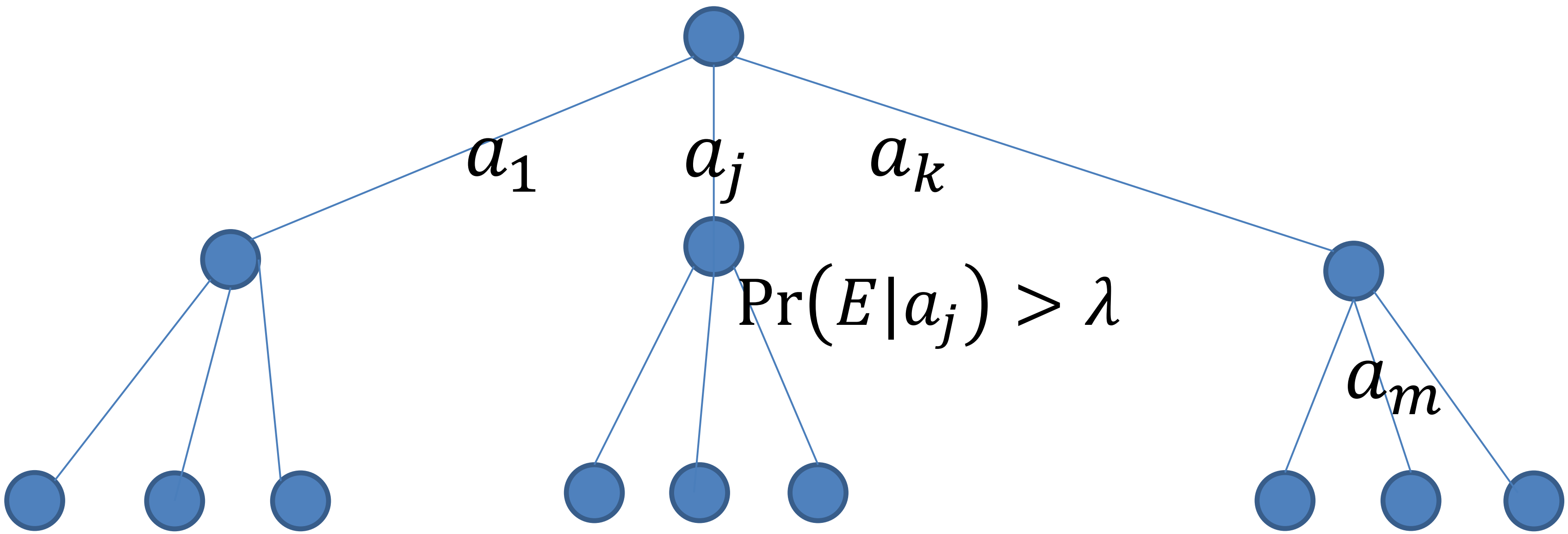
- Naturally lead to (randomized) algorithms
- First Moment method
  - $\Pr(X \geq E[X]) > 0, \Pr(X \leq E[X]) > 0$
    - Trick: estimate $\Pr(X \geq E[X])$ by $E[X]$
  - Markov's inequality: $\mathrm{Prob}[X \geq a] \leq \mathrm{E}[X]/a$

    $\mathrm{Prob}[X \neq 0] = \mathrm{Prob}[X > 0] = \mathrm{Prob}[X \geq 1] \leq \mathrm{E}[X]$

A Review of Lecture 5

- De-randomize using conditional probability

$$\lambda < \Pr(E) = \sum \Pr(E|a_i)\Pr(a_i)$$



- Sample and modify

Second moment argument

- Chebyshev inequality

$$\mathrm{Prob}[|X - \mathrm{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$$

- Specifically,

$$\mathrm{Prob}[X = 0] \leq \mathrm{Prob}[|X - \mathrm{E}[X]| \geq \mathrm{E}[X]] \leq \frac{\mathrm{Var}[X]}{\mathrm{E}[X]^2}$$

- Typically works when nearly independent
  - Due to the difficulty in estimating the variance

An improved version by Shepp

- $\text{Prob}[X = 0] \leq \dfrac{\text{Var}[X]}{\text{E}[X^2]}$
- Proof: $$\begin{aligned} \text{E}[X]^2 &= (\text{E}[1_{X \neq 0} \cdot X])^2 \\ &\leq \text{E}[1^2_{X \neq 0}]\,\text{E}[X^2] \\ &= \text{Prob}[X \neq 0]\text{E}[X^2] \\ &= \text{E}[X^2] - \text{Prob}[X = 0]\text{E}[X^2] \end{aligned}$$
  - The inequality is due to $\text{E}[XY]^2 \leq \text{E}[X^2]\text{E}[Y^2]$
- Remark: better but harder

App.: Erdös distinct sum problem

- $A \subset R^+$ has distinct subset sums
  - different subsets have different sums
  - Example: $\{2^0, 2^1, \dots 2^k\}$
- Given $n \in Z^+$, $f(n)$ is the max k such that: $\exists S \subset [n], |S| = k, S$ has distinct subset sums
- $f(n) \geq \lfloor \ln_2 n \rfloor + 1$
- Erdös promised 500\$: $f(n) \leq \lfloor \ln_2 n \rfloor + c$
  - Now offered by Ron Graham

An easy upper bound

- Assume k-set $S \subset [n]$ has distinct subset sums
- There are 2^k subset sums
- Each subset sum $\in [nk]$
- So, $2^k \leq nk$
- $k \leq \ln_2 n + \ln_2 k \leq \ln_2 n + \ln_2(\ln_2 n + \ln_2 k)$
  $\leq \ln_2 n + \ln_2(2\ln_2 n)$
  $= \ln_2 n + \ln_2 \ln_2 n + O(1)$
- Can it be tighter? Yes!

A tighter upper bound

- Intuition underlying the proof:
  - A small interval ($[nk]$) has many (2^k) sums
- If the sums are not distributed uniformly
  - *most* of the sums lie in a *much smaller* interval
  - Better estimation of k can be achieved
  - It is the case by Chebyshev's inequality

Proof: $f(n) = \ln_2 n + \frac{1}{2}\ln_2 \ln_2 n + O(1)$

- Fix n, k-set $S \subset [n]$ with distinct subset sums
- X: the sum of a random subset of S
  - $\mu = E[X], \sigma^2 = Var[X]$
- $\text{Prob}[|X-\mu| \geq \alpha\sigma] \leq \frac{1}{\alpha^2} \Rightarrow \text{Prob}[|X-\mu| < \alpha\sigma] \geq 1 - \frac{1}{\alpha^2}$
- So, $1 - \frac{1}{\alpha^2} \leq \frac{1}{2^k}(2\alpha\sigma + 1)$
  - $\Pr(X = i)$ is either 0 or 2^{-k}
  - There are $\leq 2\alpha\sigma + 1$ integers in $\alpha\sigma$-neigh. of μ

Proof (continued)

- Estimating σ (assume $S = \{a_1, \ldots a_k\}$)
  - $\sigma^2 = \frac{a_1^2 + \cdots + a_k^2}{4} \leq \frac{n^2 k}{4} \Rightarrow \sigma \leq n\sqrt{k}/2$
- $\Rightarrow 1 - \frac{1}{\alpha^2} \leq \frac{1}{2^k}(\alpha n\sqrt{k} + 1)$
- $\Rightarrow n \geq \frac{2^k\left(1 - \frac{1}{\alpha^2}\right) - 1}{\alpha\sqrt{k}}$
- This holds for any $\alpha > 1$
- Let $\alpha = \sqrt{3}$, we have $n \geq \frac{2}{3\sqrt{3}} \frac{2^k}{\sqrt{k}}$.

Application: threshold function

- Consider a property P of random graph $G_{n,p}$
- Threshold function $t(n)$ for P is such that

$$\lim_{n\to\infty} \operatorname*{Prob}_{G\in\mathcal{G}(n,p)} [G \text{ has property } P] = \begin{cases} 0 & \text{if } p = o(t(n)) \\ 1 & \text{if } p = \omega(t(n)) \end{cases}$$

- Example (clique number $c(G)$: max clique size)
  - $P: c(G) \geq 4$
  - $t(n) = n^{-\frac{2}{3}}$ is its threshold function

$n^{-\frac{2}{3}}$ is the threshold function of $c(G) \geq 4$

- Proof: when $p = o(n^{-\frac{2}{3}})$
  - S: a 4-subset of the n vertices
  - X_S: indicator of whether S spans a clique
  - $X = \sum_S X_S$: the number of 4-cliques
  - $\mathrm{E}[X] = \binom{n}{4} p^6 < \frac{n^4 p^6}{24}$
  - By Markov's inequality

$$\Pr(c(G) \geq 4) = \Pr(X > 0)$$

$$\leq E[X] < \frac{n^4 p^6}{24} = o(1)$$

Proof: when $p = \omega(n^{-\frac{2}{3}})$

- To derive $\Pr(X > 0) \to 1$, use the Chebychev's ineq. $\Pr(X = 0) \leq Var[X]/E[X]^2$
  - Try to show $Var[X] = o(E[X]^2)$
- Recall $Var[X] = \sum Var[X_S] + \sum_{S \neq T} \mathrm{Cov}(X_S, X_T)$
- X_S is an indicator $\Rightarrow Var[X_S] \leq E[X_S]$
- $\mathrm{Cov}(X_S, X_T) \leq E[X_S X_T] = \Pr(X_S = 1, X_T = 1)$

$$= E[X_S] \Pr(X_T = 1 | X_S = 1)$$

Proof: estimating the variance

- $Var[X] = \sum E[X_S] \sum_{T \sim S} \Pr(X_T = 1 | X_S = 1)$
  $= \sum E[X_S] \Delta_S$
- $\Delta_S = 1 + \sum_{|T \cap S| = 2} \Pr(X_T = 1 | X_S = 1)$
  $+ \sum_{|T \cap S| = 3} \Pr(X_T = 1 | X_S = 1)$
  $= 1 + \binom{n-4}{2}\binom{4}{2} p^5 + \binom{n-4}{1}\binom{4}{3} p^3$
  $= o(n^4 p^6) = o(E[X])$
- $Var[X] = o(E[X]^2) \Rightarrow \Pr(X = 0) = o(1)$
  $\Rightarrow \Pr(X > 0) \rightarrow 1$

Lovász local lemma: motivation

- Can we avoid all bad events?
- Given bad events $A_1, A_2, \dots A_n$, with $\Pr(A_i) < 1$
  - Is $\Pr(\cap_i \overline{A_i}) > 0$
- Two special cases
  - $\sum_i \Pr(A_i) < 1 \Rightarrow \Pr(\cap_i \overline{A_i}) \geq 1 - \sum_i \Pr(A_i) > 0$
  - Independence $\Rightarrow \Pr(\cap_i \overline{A_i}) = \prod (1 - \Pr(A_i)) > 0$
- What if *almost* independent?

Lovász local lemma: symmetric version

- Given event set $S = \{A_1, A_2, \ldots A_n\}$, $\Gamma(A_i) \subseteq S$ is such that A_i is independent of $S \backslash \Gamma(A_i)$
- **Theorem**: $\Pr(\cap_i \overline{A_i}) > 0$ if
  1. $\forall i, \Pr(A_i) \le p, |\Gamma(A_i)| \le d$ and
  2. $4pd \le 1$
- By Erdös&Lovász in 1975

Lovász local lemma: proof

- Standard trick
  - $\Pr(\cap_i \overline{A_i}) = \prod_{i=1}^{n} \Pr(\overline{A_i} | \cap_{j=1}^{i-1} \overline{A_j})$
  - Valid only if each $\cap_{j=1}^{i-1} \overline{A_j}$ has nonzero probability
  - Hold if each term $\Pr\left(\overline{A_i} \middle| \cap_{j=1}^{i-1} \overline{A_j}\right) > 0$
- **Claim**: for any $t \geq 0$ and $A, B_1, B_2, \ldots B_t \in S$,
  1. $\Pr(\cap_{j=1}^{t} \bar{B}_j) > 0$
  2. $\Pr\left(A \middle| \cap_{j=1}^{t} \bar{B}_j\right) < \frac{1}{2d}$

Inductive proof of the claim

- Basis: $t = 0$. Both 1 and 2 of the claim hold
- Hypothesis: the claim holds for all $t' < t$
- Induction
  - For **1**, $\Pr\left(\bigcap_{j=1}^{t} \overline{B_j}\right)$
    $$= \Pr\left(\overline{B_t} \mid \bigcap_{j=1}^{t-1} \overline{B_j}\right) \Pr\left(\bigcap_{j=1}^{t-1} \overline{B_j}\right) > 0$$
  - For **2**, let $\{C_1, \dots C_x\} = \{B_1, \dots B_t\} \cap \Gamma(A_i)$, and
    $$\left\{D_1, \dots D_y\right\} = \{B_1, \dots B_t\} \backslash \Gamma(A_i)$$
    - $x \leq d, x + y = t$

Proof: induction for 2

- If $x = 0$, A is independent of $\{B_1, \ldots B_t\}$ and $\Pr\left(A \middle| \cap_{j=1}^{t} \overline{B_j}\right) = \Pr(A) < \frac{1}{2d}$
- Assume $x > 0$. Then $y < t$.
  - $\Pr\left(A \middle| \cap_{j=1}^{t} \overline{B_j}\right) = \frac{\Pr(A \cap \cap_{j=1}^{t} \overline{B_j})}{\Pr(\cap_{j=1}^{t} \overline{B_j})}$

$$\leq \frac{\Pr\left(A \cap \cap \overline{D_j}\right)}{\Pr\left(\cap \overline{C_j} \cap \cap \overline{D_j}\right)} = \frac{\Pr\left(A | \cap \overline{D_j}\right)}{\Pr\left(\cap \overline{C_j} | \cap \overline{D_j}\right)}$$

$$= \frac{\Pr(A)}{1 - \Pr\left(\cup C_j | \cap \overline{D_j}\right)} \leq \frac{p}{1 - \frac{d}{2d}} < \frac{1}{2d}$$

Application: edge-disjoint paths

- n pairs of users to communicate

  Pair i can choose a path from m-set F_i

  Can the paths be pair-wise edge-disjoint?

- Yes, if any path in F_i shares edges with at most $k \leq \frac{m}{8n}$ paths in F_j

Proof

- Each pair randomly choose a path in F_i
- $E_{i,j}$: the paths of pairs i and j share edges
  - $\Pr\left(E_{i,j}\right) \leq \frac{k}{m} =: p$
  - $E_{i,j}$ is independent of $E_{i',j'}$ if $i', j' \notin \{i, j\}$
    - $\left|\Gamma\left(E_{i,j}\right)\right| < 2n =: d$
- $4dp \leq 1$
- $\Pr\left(\cap \overline{E_{i,j}}\right) > 0$ by local lemma

Other forms of Lovász local lemma

- Still true if $4dp \leq 1$ replaced by $ep(d+1) \leq 1$
  - Similar proof, but prove $\Pr\left(A \middle| \cap_{j=1}^{t} \bar{B}_j\right) \leq \frac{1}{d+1}$
  - Example: select colored beads on a circle

- Given events $A_1, A_2, \ldots A_n$, if there are $x_1, x, \ldots$ $x_n \in (0,1)$ s.t. $\Pr(A_i) \leq x_i \prod_{j \in \Gamma(A_i)} (1 - x_j)$, then $\Pr(\cap \overline{A_i}) \geq \prod (1 - x_i)$
  - Similar proof, but prove $\Pr\left(A_i \middle| \cap_{j=1}^{t} \bar{B}_j\right) \leq x_i$
  - Frugal Graph Coloring: HW

Algorithmic aspects

- **Like** other probabilistic methods, LLL proves existence non-constructively
- **Unlike** other probabilistic methods, LLL doesn't lead to efficient algorithms
  - If directly sample, one just knows the *exponentially small* lower bound of success probability
- Is there an efficient, constructive proof?

Constructive proof Lovász Local Lemma

- Breakthrough made by Joszef Beck
  - Under restrictive conditions
  - In terms coloring, SAT …
- Work by Robin Moser and Gabor Tardos in 2009
  - Let Ξ be a finite set of events determined by a finite set Υ of mutually independent random variables. If there exists an assignment of reals $x\colon \Xi \rightarrow (0,1)$ such that $\Pr(A) \leq x(A) \prod_{B \in \Gamma(A)} (1 - x(B))$, then Υ can be assigned s.t. no event in Ξ occurs. The assignment can be algorithmically found in time $\sum \frac{x(A)}{1-x(A)}$

The algorithm finding the assignment

---

function sequential_lll(Υ, Ξ)

  for all $P \in \Upsilon$ **do**

    $v_P \leftarrow$ a random evaluation of P;

  while $\exists A \in \Xi : A$ is violated when $(P = v_P : \forall P \in \Upsilon)$ **do**

    pick an arbitrary violated event $A \in \Xi$;

    for all $P \in \text{vbl}(A)$ **do**

      $v_P \leftarrow$ a new random evaluation of P;

  return $(v_P)_{P \in \Upsilon}$;

- $\text{vbl}(A) \subset \Upsilon$: the set of variables determining A

Basic idea of the proof

Estimate the average number of resampling

⇓

Estimate that of resampling each event A

⇓

Estimate the number of witness trees rooted at A

⇓

Interpret the number in terms of generating probability in a branching process

References

- http://www.cse.buffalo.edu/~hungngo/classes/2011/Spring-694/lectures/sm.pdf
- Robin Moser and Gabor Tardos. A constructive proof of the general Lovasz Local Lemma. 2009
- http://www.openproblemgarden.org/

Thank you!