Probabilistic Method and Random Graphs

Lecture 9. Second Moment Method and Lovász Local Lemma

Xingwu Liu

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

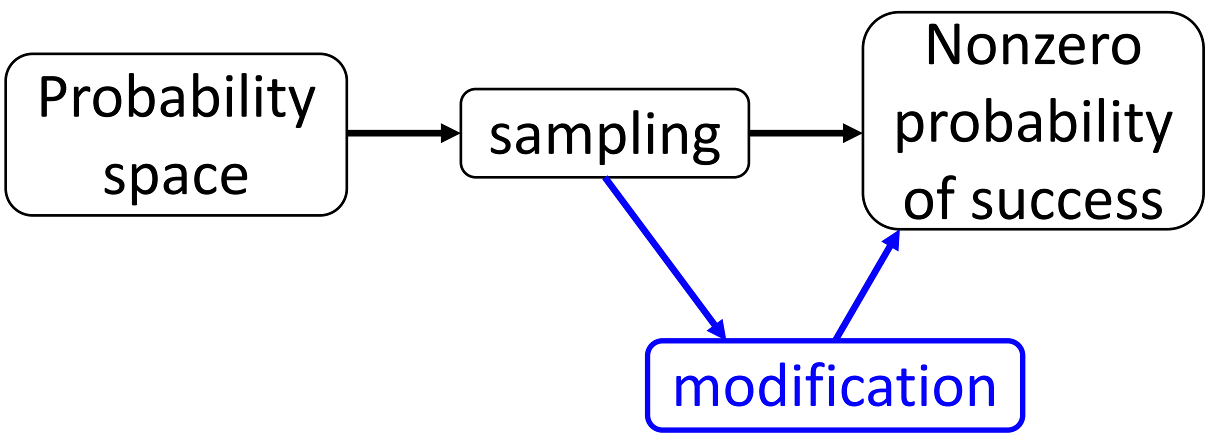
[1]The slides are mainly based on Chapter 6 of Probability and Computing.

Comments, questions, or suggestions?

A Review of Lecture 9

- Derive a deterministic algorithm from expectation argument
- Markov's Ine.: graphs with arbitrarily big girth and chro. number

First Moment method



Main Probabilistic Methods

- Counting argument
- First-moment method
- Second-moment method
- Lovasz local lemma

Second moment argument

- Chebyshev Ineq.: $\Pr(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}$
- A special case:

$$\Pr(X = 0) \leq \Pr(|X - \mathbb{E}[X]| \geq \mathbb{E}[X]) \leq \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2}$$

  - Compare with $\Pr(X \neq 0) \leq \mathbb{E}[X]$ for integer r.v. X
- Typically works when nearly independent
  - Due to the difficulty in computing the variance

An improved version by Shepp

- $\Pr(X = 0) \leq \frac{\mathrm{Var}[X]}{\mathbb{E}[X^2]}$
- Proof: $(\mathbb{E}[X])^2 = (\mathbb{E}[1_{X\neq 0} \cdot X])^2$

$$\leq \mathbb{E}[1_{X\neq 0}^2]\mathbb{E}[X^2]$$
$$= \Pr(X \neq 0)\mathbb{E}[X^2]$$
$$= \mathbb{E}[X^2] - \Pr(X = 0)\mathbb{E}[X^2]$$

  - The inequality is due to $\left(\int fg\right)^2 \leq \int f^2 \int g^2$
- When $X \geq 0$, $\Pr(X > 0) > \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}$

Generalizing Shepp's Theorem

- $\Pr(X > \theta\mathbb{E}[X]) \geq (1-\theta)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}, \theta \in (0,1)$
- Paley&Zygmund, 1932
- Proof:

$$\mathbb{E}[X] = \mathbb{E}\left[X 1_{X \leq \theta\mathbb{E}[X]}\right] + \mathbb{E}\left[X 1_{X > \theta\mathbb{E}[X]}\right]$$
$$\leq \theta\mathbb{E}[X] + \left(\mathbb{E}[X^2]\Pr(X > \theta\mathbb{E}[X])\right)^{\frac{1}{2}}$$

- Further improvement, tight when X is constant

$$\Pr(X > \theta\mathbb{E}[X]) \geq \frac{(1-\theta)^2(\mathbb{E}[X])^2}{\mathrm{Var}[X] + (1-\theta)^2(\mathbb{E}[X])^2}$$

due to $\mathbb{E}\left[X - \theta\mathbb{E}[X]\right] \leq \mathbb{E}\left[(X - \theta\mathbb{E}[X]) 1_{X > \theta\mathbb{E}[X]}\right]$

App.: Erdős distinct sum problem

- $A \subset \mathbb{R}^+$ has distinct subset sums
  - different subsets have different sums
  - Example: $A = \{2^0, 2^1, \dots 2^k\}$
- Fix $n \in \mathbb{Z}^+$. Consider $S \subset [n]$ having distinct subset sums. $f(n)$ is the max size of such S
- Easy lower bound: $f(n) \geq \lfloor \ln_2 n \rfloor + 1$
- Erdős promised 500$: $f(n) \leq \lfloor \ln_2 n \rfloor + c$
  - Now offered by Ron Graham

An easy upper bound

- Assume k-set $S \subseteq [n]$ has distinct subset sums
- There are 2^k subset sums
- Each subset sum $\in [nk]$
- So, $2^k \leq nk$
- $k \leq \ln_2 n + \ln_2 k \leq \ln_2 n + \ln_2(\ln_2 n + \ln_2 k)$

$$\leq \ln_2 n + \ln_2(2\ln_2 n)$$

$$= \ln_2 n + \ln_2 \ln_2 n + 1$$

- Can it be tighter? Yes!

A tighter upper bound

- Intuition underlying the proof:
  - A small interval ($[nk]$) has many (2^k) distinct sums
- If the sums are not distributed uniformly
  - *Most* of the sums lie in a *much smaller* interval
  - k must be smaller
  - 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 a k-set $S \subset [n]$ with distinct subset sums
- X: the sum of a random subset of S
  - $\mu = \mathbb{E}[X], \sigma^2 = Var[X]$
- $\Pr(|X-\mu| \geq \alpha\sigma) \leq \frac{1}{\alpha^2} \Rightarrow$

$$1 - \frac{1}{\alpha^2} \leq \Pr(|X-\mu| < \alpha\sigma) \Rightarrow$$

$$1 - \frac{1}{\alpha^2} \leq \sum_{i=\mu-\alpha\sigma}^{\mu+\alpha\sigma} \Pr(X = i) \leq \frac{2\alpha\sigma+1}{2^k}$$

Since $\Pr(X = i)$ is either 0 or 2^{-k}

Proof (continued)

- Estimating σ (assume $S = \{a_1, \dots, a_k\}$):

$$\sigma^2 = \frac{a_1^2 + \cdots + a_k^2}{4} \leq \frac{n^2 k}{4} \Rightarrow \sigma \leq \frac{n\sqrt{k}}{2}$$

$$\Rightarrow 1 - \frac{1}{\alpha^2} \leq \frac{1}{2^k}(\alpha n\sqrt{k} + 1)$$

$$\Rightarrow n \geq \frac{2^k\left(1 - \frac{1}{\alpha^2}\right) - 1}{\alpha\sqrt{k}}$$

- This holds for any $\alpha > 1$. Let $\alpha = \sqrt{3}$
- $n \geq \frac{2}{3\sqrt{3}} \frac{2^k}{\sqrt{k}} \Rightarrow k \leq \ln_2 n + \frac{1}{2} \ln_2 \ln_2 n + O(1)$

Application: threshold function

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

$$\lim_{n\to\infty} \Pr\left(G_{n,p} \text{ has } P\right) = \begin{cases} 0 \text{ if } p = o(t(n)) \\ 1 \text{ if } p = \omega(t(n)) \end{cases}$$

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

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
- $\mathbb{E}[X] = \binom{n}{4} p^6 < \frac{n^4 p^6}{24}$
- 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 $\Pr(X > 0) \to 1$
  - By Chebychev's Ineq.: $\Pr(X = 0) \leq \frac{\mathrm{Var}[X]}{(\mathbb{E}[X])^2}$
  - Try to show $\mathrm{Var}[X] = o(\mathbb{E}[X])^2$
- Recall $\mathrm{Var}[X] = \sum \mathrm{Var}[X_S] + \sum_{S \neq T} \mathrm{Cov}(X_S, X_T)$
- X_S is an indicator $\Rightarrow \mathrm{Var}[X_S] \leq \mathbb{E}[X_S]$
- $\mathrm{Cov}(X_S, X_T) \leq \mathbb{E}[X_S X_T] = \Pr(X_S = 1, X_T = 1)$

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

And $\mathrm{Cov}(X_S, X_T)$=0 if independent

Proof: estimating the variance

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

- $\mathrm{Var}[X] = o(\mathrm{E}[X]^2) \Rightarrow \Pr(X = 0) \leq \frac{\mathrm{Var}[X]}{\mathrm{E}[X]^2} = o(1)$
  $\Rightarrow \Pr(X > 0) \to 1$

Lovász local lemma: motivation

- Can we avoid all bad events?
- Given bad events $A_1, A_2, \dots A_n$, is $\Pr(\cap_i \overline{A_i}) > 0$?
  - Applicable to SAT, coloring, Ramsey theory…
- Two special cases
  - $\sum_i \Pr(A_i) < 1 \Rightarrow \Pr(\cap_i \overline{A_i}) \geq 1 - \sum_i \Pr(A_i) > 0$
  - Independent $\Rightarrow \Pr(\cap_i \overline{A_i}) = \prod(1 - \Pr(A_i)) > 0$
- What if *almost* independent?

Lovász local lemma: symmetric version

- Dependency graph
  - Undirected simple graph on $S = \{A_1, A_2, \ldots A_n\}$
  - A_i is independent of its non-neighborhood $S \backslash \Gamma^+(A_i)$
    - $\Gamma(A_i) \triangleq \Gamma^+(A_i) \backslash \{A_i\}$
- **Theorem**: $\Pr(\cap_i \overline{A_i}) > 0$ if
  1. $\forall i, \Pr(A_i) \leq p, |\Gamma(A_i)| \leq d$ and
  2. $4pd \leq 1$
- By Erdös&Lovász in 1973 to Erdös 60th birthday

Lovasz

Erdos

Lovász local lemma: proof

- Standard trick
  - Chain rule: $\Pr(\cap_i \overline{A_i}) = \prod_{i=1}^{n} \Pr(\overline{A_i} | \cap_{j=1}^{i-1} \overline{A_j})$
    - 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 $\{C_1, \ldots C_x\} = \{B_1, \ldots B_t\} \cap \Gamma(A)$, and
    $$\{D_1, \ldots D_y\} = \{B_1, \ldots B_t\} \backslash \Gamma(A)$$
    - $x \leq d, x + y = t$

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 $y < t$.
  - $\Pr\left(A \middle| \cap_{j=1}^{t} \overline{B_j}\right) = \frac{\Pr(A \cap (\cap_{j=1}^{t} \overline{B_j}))}{\Pr(\cap_{j=1}^{t} \overline{B_j})}$

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

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

General case

Application to (k,s)-SAT

- (k,s)-CNF
  - Any clause has k literals
  - Any literal appears in at most s clauses
- Theorem: Any (k,s)-CNF is satisfiable if $s \leq \frac{1}{4}\frac{2^k}{k}$
  - Randomly assign values to the Boolean variables
  - A_i: the event that the ith clause is not satisfied
  - $\Pr(\bigcap \overline{A_i}) > 0 \Leftrightarrow$ satisfiable
  - $p = \Pr(A_i) = 2^{-k}, \;\; d \leq ks$
  - $s \leq \frac{1}{4}\frac{2^k}{k} \Rightarrow 4pd \leq 1 \Rightarrow \Pr(\bigcap \overline{A_i}) > 0 \Rightarrow$ satisfiable

Application to Ramsey Number $R(k)$

- Counting argument: $R(k) \geq k2^{\frac{k}{2}}\left[\frac{1}{e\sqrt{2}} + o(1)\right]$ [1947]
- Best result: $R(k) \geq k2^{\frac{k}{2}}\left[\frac{\sqrt{2}}{e} + o(1)\right]$ [1975, Spencer]
  - Randomly color edges of K_n in red/blue
  - S: a k-subset of the vertices
  - A_S: S is monochromatic
  - $p = \Pr(A_S) = 2^{1-\binom{k}{2}}$, $d \leq \binom{k}{2}\binom{n}{k-2}$
  - By Stirling's formula, $4pd \leq 1$ if $n \leq k2^{\frac{k}{2}}\left[\frac{\sqrt{2}}{e} + o(1)\right]$
- Can we say something about $R(k,t)$?

Thank you!