Probabilistic Method and Random Graphs

Lecture 9. De-randomization and Second Moment Method

Xingwu Liu

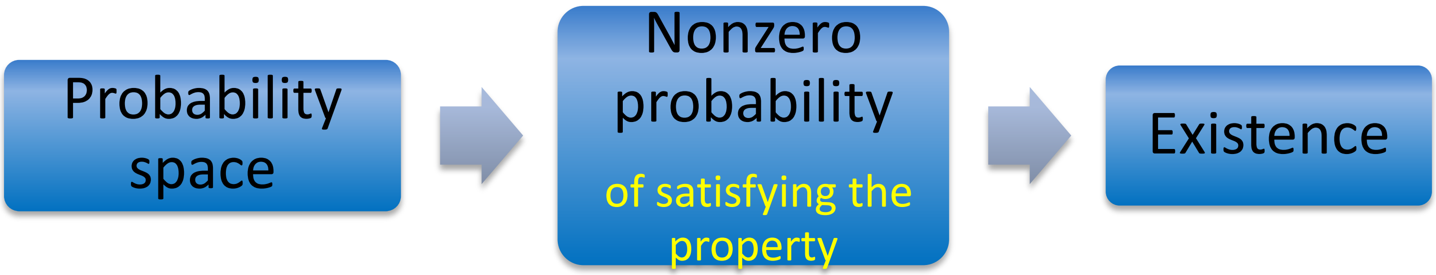
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

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

Comments, questions, or suggestions?

A Review of Lecture 8

- Principle of probabilistic method

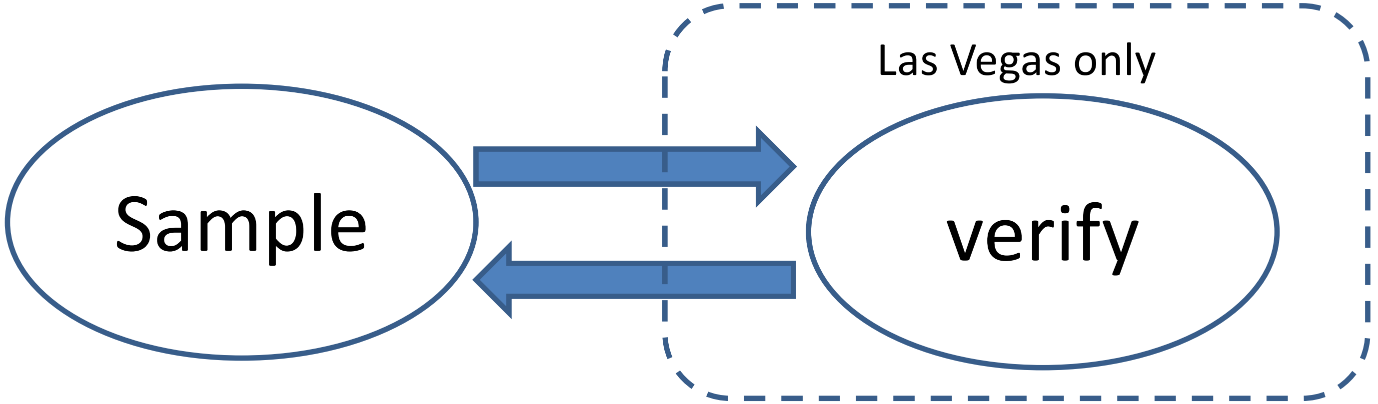


- Counting: Tournament, Ramsey number
- First moment method: Max-3SAT, MIS
  - Expectation argument: $\Pr(X \geq \mathbb{E}[X]) > 0, \Pr(X \leq \mathbb{E}[X]) > 0$
  - Markov's inequality: $\Pr(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$

    $\Pr(X \neq 0) = \Pr(X > 0) = \Pr(X \geq 1) \leq \mathbb{E}[X]$

A Review of Lecture 8

- How to find an desirable object? By sampling!
- Algorithmic paradigm



- First moment method guarantees efficiency

- Cool to get an efficient randomized algorithm
- Can we derive a deterministic one?
- Yes, if expectation argument is used

De-randomization: an example

- **MAX-3SAT**: Given a 3-CNF Boolean formula, find a truth assignment satisfying the maximum number of clauses
  - E.g.: $(x_1 \vee x_2 \vee x_3) \wedge \ldots \wedge (\overline{x_1} \vee \overline{x_3} \vee x_4)$
- Known: at least $\frac{7}{8}n$ clauses can be satisfied
- Randomized algo. to find a good assignment
  - Independently, randomly assign values
  - Succeed if lucky
    - Can we make good choice, rather than pray for luck?

Look closer at the randomized algorithm

- In equivalence, choose values sequentially
- Good choices lead to a good final result
  - Which choice is good?
    - Easy to know with hindsight, but how to predict
  - A tentative approach: always make the choice which allows a good final result
    - Fact: a $\frac{7n}{8}$ expect. means the existence of a $\frac{7}{8}$-approx.
    - Make the current choice, keeping the expectation $\geq \frac{7n}{8}$
  - Nice, but does such a choice exist? How to find it?

Conditional expectation says yes!

- The first step
  - $\frac{7n}{8} = \mathbb{E}[X] = \sum_{v_1} \Pr(x_1 = v_1)\, \mathbb{E}[X | x_1 = v_1]$
  - There must be v_1 s.t. $\mathbb{E}[X | x_1 = v_1] \geq \frac{7n}{8}$
- Likewise, if $\mathbb{E}[X | x_1 = v_1, \ldots, x_{k-1} = v_{k-1}] \geq \frac{7n}{8}$, then $\mathbb{E}[X | x_1 = v_1, \ldots, x_k = v_k] \geq \frac{7n}{8}$ for some v_k
- Final correctness
  - $X(x_1 = v_1, \ldots, x_m = v_m) = \mathbb{E}[X | x_1 = v_1, \ldots, x_m = v_m] \geq \frac{7n}{8}$
- Given $v_1, \ldots, v_{k-1}$, what's the v_k?
  - Let v_k s.t. $\mathbb{E}[X | x_1 = v_1, \ldots, x_k = v_k]$ is maximized

Deterministic $\frac{7}{8}$-algorithm for MAX-3SAT

For $k = 1 \cdot\cdot m$ do

$$x_k = \text{argmax}_{v_k} \mathbb{E}[X | x_1 = v_1, \ldots x_{k-1} = v_{k-1}, x_k = v_k]$$

Endfor

- Cool! And this approach can be generalized

De-randomization via conditional expectation

- Expectation argument $\Longrightarrow$ deterministic algorithm
- Basic idea
  - Expectation argument guarantees existence
  - Sequentially make deterministic choices
    - Each choice maintains the expectation, given the past ones

- Only valid for expectation argument where randomness lies in a sequence of random variables
- What if the expectation is hard to compute?

Example: Turán Theorem

- Any graph $G = (V, E)$ contains an independent set of size at least $\frac{|V|}{D+1}$, where $D = \frac{2|E|}{|V|}$
- Expectation argument: the expected size of an independent set S is at least $\frac{|V|}{D+1}$
- Randomly choose vertices into S one by one

- Try the de-randomization routine

Idea of the algorithm (1)

- Choose valid vertices sequentially
- At step $t+1$, find u to maximize $\mathbb{E}\left[Q|S^{(t)},u\right]$
  - $S^{(t)}$: the independent set at step t
  - Q: the size of the final independent set
- Hard to compute the expectation ☹
  - $\mathbb{E}[Q] \geq \sum \frac{1}{d(w)+1} \geq \frac{|V|}{D+1}$
- It suffices to show $\mathbb{E}\left[Q|S^{(t)}\right] \geq \frac{|V|}{D+1}$ for any t

Idea of the algorithm (2)

- Note that $\mathbb{E}\left[Q|S^{(t)}\right] \geq \left|S^{(t)}\right| + \sum_{w \in R^{(t)}} \frac{1}{d(w)+1} \triangleq X^{(t)}$
  - $R^{(t)}$: set of vertices away from $S^{(t)}$ by distance >1
- $X^{(0)} \geq \frac{|V|}{D+1} \Rightarrow$ it's enough if $X^{(t)}$ is non-decreasing
  - Can we achieve this?
- If at step $t+1$, $u \in R^{(t)}$ is chosen,
  $X^{(t+1)} - X^{(t)} = 1 - \sum_{w \in \Gamma^{+}(u)} \frac{1}{d(w)+1}$
- $\mathbb{E}_u\left[X^{(t+1)} - X^{(t)}\right] \geq 1 - \sum_{w \in R^{(t)}} \frac{1}{d(w)+1} \frac{d(w)+1}{|R^{(t)}|} = 0$
- So, there is u s.t. $X^{(t+1)} \geq X^{(t)}$

Can it be non-negative?

A deterministic algorithm

- Initialize S to be the empty set
- **While** there is a vertex $u \notin \Gamma(S)$
  - Add to S such a vertex u which minimizes $\sum_{w \in \Gamma^+(u)} \frac{1}{d(w)+1}$
- **Return** S

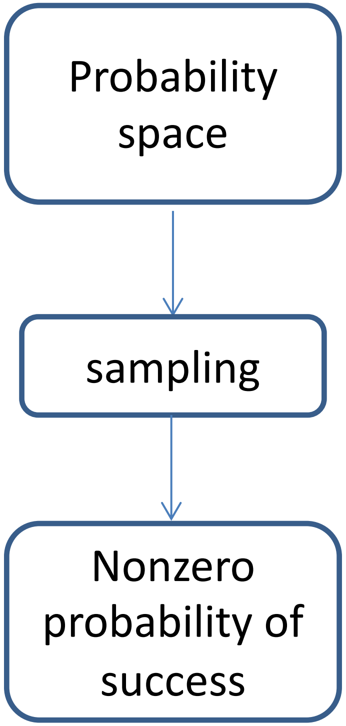
- Paul Turán (1910 –1976)
- Hungarian mathematician
- Founder of

  Probabilistic number theory

  Extremal graph theory

  (in Nazi Camp)

Sample


Probability space
sampling
Nonzero probability of success

Big Chromatic Number and Big Girth

- Chromatic number vs local structure
  – Loose local structure → small chro. number?
  – No! (Erdős 1959)
- One of the first applications of prob. Method
- Theorem: for any integers $g, k > 0$, there is a graph with girth$\geq g$ and chro. number$\geq k$
- We just prove the special case $g = 4$, i.e. triangle-free

Basic Idea of the Proof

- Randomly pick a graph G from $G_{n,p}$
  - $\chi(G)$: the chromatic number of G
  - $\mathbb{I}(G)$: the size of a maximum independent set of G
- With high probability $\mathbb{I}(G)$ is small
  - $\mathbb{I}(G)\chi(G) \geq n$ implies that $\chi(G)$ is big
- With high probability G has few triangles
- Destroy the triangles while keeping $\mathbb{I}(G)$ small

Proof: $\mathbb{I}(G)$ is small w.h.p.

- X: the number of independent sets of size $\frac{n}{2k}$
- $\Pr\left(\mathbb{I}(G) \geq \frac{n}{2k}\right) = \Pr(X \neq 0) \leq \mathbb{E}[X]$

$$= \binom{n}{n/2k} (1-p)^{\binom{n/2k}{2}}$$

$$< 2^n e^{-\frac{pn(n-2k)}{8k^2}}$$

- Small if n is large and $p = \omega(n^{-1})$

Proof: triangles are few w.h.p.

- $\mathcal{T}(G)$: the number of triangles of G
- $\mathbb{E}[\mathcal{T}(G)] = \binom{n}{3}p^3 < \frac{(np)^3}{6} = \frac{n}{6}$ if $p = n^{-2/3}$
- By Markov ineq., $\Pr\left(\mathcal{T}(G) > \frac{n}{2}\right) \le \frac{1}{3}$
- Recall $\Pr\left(\mathbb{I}(G) \ge \frac{n}{2k}\right) < 2^n e^{-\frac{pn(n-2k)}{8k^2}}$

$$< e^n\, e^{-\frac{pn^2}{16k^2}} = e^{n-n^{\frac{4}{3}}/16k^2} \quad \text{if } n > 4k$$

$$< e^{-n} < \frac{1}{6} \quad \text{if } n^{1/3} \ge 32k^2$$

Proof: modification

- $\Pr\left(\mathbb{I}(G) < \frac{n}{2k}, \mathcal{T}(G) \le \frac{n}{2}\right) > \frac{1}{2}$
  - Choose G s.t. $\mathbb{I}(G) < \frac{n}{2k}, \mathcal{T}(G) \le \frac{n}{2}$
- Remove one vertex from each triangle of G, resulting in a graph G' with $n' \ge n - \mathcal{T}(G)$
- $\mathbb{I}(G') \le \mathbb{I}(G) < \frac{n}{2k}$
- $\chi(G') \ge \frac{n'}{\mathbb{I}(G')} \ge \frac{n'}{\mathbb{I}(G)} \ge \frac{n - \mathcal{T}(G)}{\frac{n}{2k}} \ge k$

Algorithm for finding such a graph

- Fix $n^{1/3} \geq 32k^2$ and $p = n^{-2/3}$
- Sample G from $G_{n,p}$
- Destroy the triangles

- Success probability > ½

- Do you have any idea of de-randomizing?

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: $A = \{2^0, 2^1, \ldots 2^k\}$
- 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 $S = \{a_1, \ldots, a_k\}$):

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

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

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

- This holds for any $\alpha > 1$. Let $\alpha = \sqrt{3}$

- $n \ge \frac{2}{3\sqrt{3}} \frac{2^k}{\sqrt{k}} \Rightarrow k \le \ln_2 n + \frac{1}{2} \ln_2 \ln_2 n + O(1)$

References

- http://www.cse.buffalo.edu/~hungngo/classes/2011/Spring-694/lectures/sm.pdf
- http://www.openproblemgarden.org/
- Documentary film of Erdős: N is a Number - A Portrait of Paul Erdős

Thank you!