$\{$ Csc 80030 | Lecture 8 $\}$

Probabilistic Analysis & Randomized Algorithms

Hoeteck Wee · hoeteck@cs.qc.edu

PART 0

- HOMEWORK: HW4 due next week, HW5 out by sun
- PROGRAMMING ASSIGNMENT: due tonight
- TODAY: random-walk algorithms

PART 1 | undirected s-t connectivity

Undirected S-T Connectivity

- S-T CONNECTIVITY: Given a directed graph G and two vertices s and t, is there a path from s to t in G?
- can be solved in linear time and space using BFS or DFS.
- UNDIRECTED S-T CONNECTIVITY (USTCON): Given an undirected graph G and two vertices s and t, is there a path from s to t in G?
- TODAY: randomized algorithm for USTCON running in space $O(\log n)$.

USTCON algorithm

USTCON via random walks

Input: (G, s, t), where $G = (V, E)$ has n vertices

1. Let $v = s$. Repeat up to n^4 times:
   1.1 If $v = t$, halt and accept.
   1.2 Else update v to be a random neighbor of v.
2. Reject (if we haven't visited t yet).

- space $O(\log n)$ to store current vertex v and a counter for # steps
- never accepts when there isn't a path from s to t.
- if G is a connected d-regular graph, then a random walk of length $\tilde{O}(d^2 n^3)$ from s will hit t w.h.p.
- replace each vertex v with a cycle of length $\deg(v)$.

random-walk matrix

- Define *random-walk matrix* of an n-vertex digraph G to be the $n \times n$ matrix M where $M_{i,j}$ is the prob. of going from vertex i to j in a single step.
- If G is d-regular, then M is $\frac{1}{d}$ times adjacency matrix of G.
- For every probability distribution $\pi \in \mathbb{R}^n$ on vertices of G, the vector πM is the probability distribution obtained after taking one step of the random walk.
- Start with π concentrated at vertex s, interested in πM^k after taking k steps on the graph.
- We will show that random walk converges to the uniform distribution $u = \begin{pmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$ after $\mathrm{poly}(n)$ steps.

Basic linear algebra

- A non-zero vector $v \in \mathbb{R}^n$ is an *eigenvector* of a $n \times n$ matrix M if $vM = \lambda v$ for some $\lambda \in \mathbb{R}$, the corresponding *eigenvalue*.
- EXAMPLE. the graph K_3

$$M = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

$$v_1 = (\tfrac{1}{3} \quad \tfrac{1}{3} \quad \tfrac{1}{3}), \quad \lambda_1 = 1;$$

$$v_2 = (1 \quad -1 \quad 0), \quad \lambda_2 = -1/2;$$

$$v_3 = (1 \quad 0 \quad -1), \quad \lambda_3 = -1/2$$

Basic linear algebra

Theorem

If M is a symmetric $n \times n$ real matrix, then M has n eigenvectors which form an orthogonal basis for $\mathbb{R}^n$.

Henceforth, M is random-walk matrix for a d-regular undirected graph.

- M is symmetric, with eigenvectors $v_1, \ldots, v_n$.
- Since $uM = u$, $v_1 = u$ is an eigenvector of eigenvalue $\lambda_1 = 1$.
- Assume $|\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$.
- Note $\langle \pi - u, u \rangle = 0$, so we can write $\pi = u + c_2 v_2 + \cdots c_n v_n$.
- After k steps, the distribution is $\pi M^k = u + \lambda_2^k c_2 v_2 + \cdots \lambda_n^k c_n v_n$, so $\|\pi M^k - u\| \leq |\lambda_2|^k \|\pi - u\| \leq |\lambda_2|^k$.

Random walk analysis

Will show on homework:

- All eigenvalues of M have absolute value at most 1.
- If G is connected and non-bibartite, then all eigenvalues (other than 1) have absolute value at most $1 - \Omega(1/(dn)^2)$.

From before,

- $\|\pi M^k - u\| \leq 2(1 - \Omega(1/(dn)^2))^k$.
- Take $k = O(d^2 n^2 \log n)$, every entry of πM^k is at least $1/2n$.
- Take $O(n)$ such walks, we will hit vertex t with probability at least $1/2$.

PART 2 | randomized 2-SAT

Satisfiability problems

- general satisfiability problem (SAT):
  - input: Boolean formula given as AND of OR's: e.g.

    $\phi(x_1, x_2, x_3, x_4) = (x_1 \vee \bar{x_2} \vee x_4) \wedge (\bar{x_1} \vee \bar{x_3}) \wedge (x_3 \vee x_4)$
  - n variables x_1, x_2, x_3, x_4; literals e.g. $x_1, \bar{x_2}$; m clauses e.g. $(x_3 \vee x_4)$
  - assignment e.g. (T, F, F, T)
  - ϕ is satisfiable if there exists an assignment for which ϕ evaluates to T
  - SAT problem: decide if the input ϕ is satisfiable
- 2-satisfiability problem (2-SAT):
  - restricted version of SAT where each clause contains exactly 2 literals
  - $\phi(x_1, x_2, x_3, x_4) = (x_1 \vee \bar{x_2}) \wedge (\bar{x_1} \vee \bar{x_3}) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x_3}) \wedge (x_4 \vee \bar{x_1})$

2-satisfiability algorithm RAND2SAT

1. Start with an arbitrary assignment
2. Repeat up to $2mn^2$ times, terminating if all clauses are satisfied:
   2.1 Choose an arbitrary (first) clause that is not satisfied
   2.2 Choose uniformly at random one of the literals in the clause and switch the value of the variable
3. If a satisfying assignment has been found, return it
4. Otherwise, return “unsatisfiable”

2-satisfiability algorithm: example

- $\phi(x_1, x_2, x_3, x_4) = (x_1 \vee \bar{x_2}) \wedge (\bar{x_1} \vee \bar{x_3}) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x_3}) \wedge (x_4 \vee \bar{x_1})$
  1. assignment $A_0 = (\text{F, T, T, F})$

     unsatisfied clauses: $x_1 \vee \bar{x_2},\ x_4 \vee \bar{x_3}$

     switch x_1 in $C_0 = x_1 \vee \bar{x_2}$
  2. assignment $A_1 = (\text{T, T, T, F})$

     unsatisfied clauses: $\bar{x_1} \vee \bar{x_3},\ x_4 \vee \bar{x_3},\ x_4 \vee \bar{x_1}$

     switch x_3 in $C_1 = \bar{x_1} \vee \bar{x_3}$
  3. assignment $A_2 = (\text{T, T, F, F})$

     unsatisfied clauses: $x_4 \vee \bar{x_1}$

     switch x_4 in $C_2 = x_4 \vee \bar{x_1}$
  4. assignment $A_3 = (\text{T, T, F, T})$

     unsatisfied clauses: none

2-satisfiability algorithm: analysis

- suppose ϕ is satisfiable with an assignment S
  - e.g. ϕ as before, S is (T, T, F, T)
- let X_i be the number of variables with same values in A_i and S
  - e.g. $X_0 = 1, X_1 = 2, X_2 = 3, X_3 = 4$
  - if $X_i = n$, then algorithm terminates with satisfying assignment
  - how long does it take for X_i to reach n?
- compute transition probabilities for $X_0, X_1, X_2, \ldots$
  - if we pick a variable where A_i and S agree, $X_{i+1} = X_i - 1$
  - if we pick a variable where A_i and S disagree, $X_{i+1} = X_i + 1$

2-satisfiability algorithm: analysis

- if $X_i = 0$, then $\Pr[X_{i+1} = 1 \mid X_i = 0] = 1$
- if $X_i = j$, where $1 \le j \le n-1$
  - A_i and S disagree on at least one variable in clause C_i
  - $\Pr[\text{pick a variable where } A_i \text{ and } S \text{ disagree}] \ge 1/2$

$$\Pr[X_{i+1} = j+1 \mid X_i = j] \ge 1/2$$

$$\Pr[X_{i+1} = j-1 \mid X_i = j] \le 1/2$$

2-satisfiability algorithm: analysis

- consider Markov chain $Y_0, Y_1, Y_2, \ldots$ with states $0, 1, 2, \ldots, n$

$$
\begin{aligned}
Y_0 &= X_0 \\
\Pr[Y_{i+1} = 1 \mid Y_i = 0] &= 1 \\
\Pr[Y_{i+1} = j+1 \mid Y_i = j] &= 1/2 \\
\Pr[Y_{i+1} = j-1 \mid Y_i = j] &= 1/2
\end{aligned}
$$

  - expected time to reach n in Y is $\geq$ expected time to reach n in X
  - let h_j be expected # of steps it takes to reach n from j

2-satisfiability algorithm: analysis

- computing $h_0, h_1, \ldots, h_n$
  - base case: $h_n = 0$ and $h_0 = h_1 + 1$
  - for $1 \leq j \leq n-1$: $h_j = \frac{1}{2}(h_{j-1}+1) + \frac{1}{2}(h_{j+1}+1)$
  - solving $n+1$ equations in $n+1$ unknowns, we obtain
    $h_j = n^2 - j^2$ and $h_0 = n^2$
  - expected time to reach n in X is at most n^2
- RAND2SAT
  - if ϕ is unsatisfiable, always returns the right answer "unsatisfiable"
  - if ϕ is satisfiable, returns a satisfying assignment with prob $\geq 1 - 2^{-m}$

THE END | next, ...