COMPSCI 690RA: Randomized Algorithms and Probabilistic Data Analysis

Prof. Cameron Musco

University of Massachusetts Amherst. Spring 2022.
Lecture 9

Logistics

- Problem Set 3 is due 4/15 at 8pm.
- Project progress report due this Friday, 4/8. Submit a pdf via email. 1-2 pages.
- Weekly quiz due next Tuesday at 8pm.

Summary

Last Week: **Random sketching and subspace embedding.**

- Subspace embedding via leverage score sampling.
- Analysis via matrix concentration bounds.
- Spectral graph sparsification via leverage score sampling.

Today:

- Finish spectral graph sparsification and physical interpretation
- Start on Markov chains and their analysis
- Markov chain based algorithms for 2-SAT and 3-SAT.
- Gambler's ruin.

Spectral Graph Sparsification

Subpace Embedding via Sampling

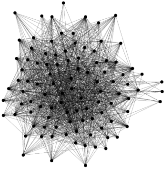
Theorem (Subspace Embedding via Leverage Score Sampling)

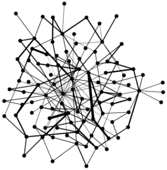
For any $A \in \mathbb{R}^{n \times d}$ with left singular vector matrix U, let $\tau_i = \|U_{i,:}\|_2^2$ and $p_i = \frac{\tau_i}{\sum \tau_i}$. Let $\mathbf{S} \in \mathbb{R}^{m \times n}$ have $\mathbf{S}_{:,j}$ independently set to $\frac{1}{\sqrt{mp_i}} \cdot e_i^T$ with probability p_i.

Then, if $m = O\left(\frac{d \log(d/\delta)}{\epsilon^2}\right)$, with probability $\geq 1 - \delta$, $\mathbf{S}$ is an ϵ-subspace embedding for A.

- Matches oblivious random projection up to the $\log d$ factor.
- Variational characterization: $\tau_i = \max_{x \in \mathbb{R}^d} \frac{[Ax](i)^2}{\|Ax\|_2^2}$.

Spectral Graph Sparsification





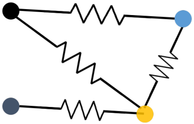
1	-1	0	0
0	1	0	-1
0	0	1	-1
-1	0	1	0
1	0	-1	0
0	1	-1	0
1	0	0	-1
0	0	1	-1

vertex-edge
incidence matrix B

- Given a graph G, find a (weighted) subgraph G' with many fewer edges such that: $(1-\epsilon)L_G \preceq L_{G'} \preceq (1+\epsilon)L_G$.
- Equivilantly, letting $B \in \mathbb{R}^{m \times n}$ be the vertex-edge incidence matrix of G, find a sampling matrix $\mathbf{S}$ that is an ϵ-subspace embedding for B. I.e, $B^T\mathbf{S}^T\mathbf{S}B \approx_\epsilon B^TB$.
- Sampling edges according to their leverage scores in B gives an ϵ-spectral sparsifier with just $O(n \log n/\epsilon^2)$ edges.
- Can be used to approximate many properties of G, including the size of all cuts.

Leverage Scores and Effective Resistance

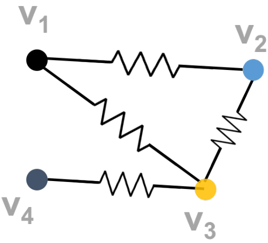
A spectral sparsifier G' of G with $O(n \log n/\epsilon^2)$ edges can be constructed by sampling rows of the vertex-edge incidence matrix via their leverage scores. What are these leverage scores?



- View each edge as a 1-Ohm resistor.
- If we fix a current of 1 between u, v, the voltage drop across the nodes is known as the effective resistance between u and v.
- We will show that the leverage score of each edge is exactly equal to its effective resistance.
- Intuitively, to form a spectral sparsifier, we should sample high resistance edges with high probability, since they are 'bottlenecks'.

Electrical Flows

For a flow $f \in \mathbb{R}^m$, the currents going into each node are given by $B^T f$.



$$\underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}}_{\mathbf{B}^\mathbf{T}} \underbrace{\begin{pmatrix} 3 \\ 0 \\ -1 \\ -1 \end{pmatrix}}_{\mathbf{f}} = \begin{pmatrix} 3 \\ -4 \\ 0 \\ 1 \end{pmatrix}$$

The electrical flow when one unit of current is sent from u to v is:

$$f^e = \underset{f : B^T f = b_{u,v}}{\arg\min} \|f\|_2.$$

Since power (energy/time) is given by $P = I^2 \cdot R$.

$$f^e = \arg\min_{f: B^T f = b_{u,v}} \|f\|_2.$$

By Ohm's law, the voltage drop across (u, v) (i.e., the effective resistance) is simply the entry $f^e_{u,v}$ (since u, v is a unit resistor).

- To solve for f, note that we can assume that f is in the column span of B. Otherwise, it would not have minimal norm. So $f = B\phi$ for some vector $\phi \in \mathbb{R}^n$.
- Then need to solve $B^T B\phi = b_{u,v}$. I.e., $L\phi = b_{u,v}$. ϕ is unique up to its component in the null-space of L.

L

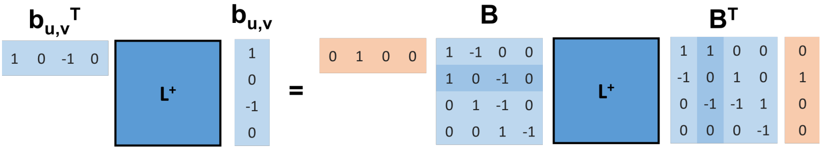
$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

- $\phi = L^+ b_{u,v}$.

Leverage Scores and Effective Resistance

The effective resistance across edge (u, v) is given by

$$b_{u,v}(B^T B)^+ b_{u,v} = e_{u,v}^T B(B^T B)^+ B^T e_{u,v}.$$



Write $B = U\Sigma V^T$ in its SVD.

$$\begin{aligned} e_{u,v}^T B(B^T B)^+ B^T e_{u,v} &= e_{u,v}^T U\Sigma V^T (V\Sigma^{-2} V^T) V\Sigma U^T e_{u,v} \\ &= e_{u,v}^T U U^T e_{u,v} \\ &= U_{u,v}^T U_{u,v} = \|U_{u,v}\|_2^2. \end{aligned}$$

I.e., the effective resistance is exactly the leverage score of the corresponding row in B.

Markov Chains

Markov Chain Definition

- A discrete time stochastic process is a collection of random variables $X_0, X_1, X_2, \ldots,$
- A discrete time stochastic process is a Markov chain if is it memoryless:

$$\Pr(X_t = a_t | X_{t-1} = a_{t-1}, \ldots, X_0 = a_0) = \Pr(X_t = a_t | X_{t-1} = a_{t-1})$$
$$= P_{a_{t-1}, a_t}.$$

Think-Pair-Share: In a Markov chain, is X_t independent of $X_{t-2}, X_{t-3}, \ldots, X_0$?

Transition Matrix

A Markov chain $X_0, X_1, \ldots$ where each X_i can take m possible values, is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

$$P_{j,k} = \Pr(X_{i+1} = k | X_i = j).$$

Let $q_i \in [0,1]^{1 \times m}$ be the distribution of X_i. Then $q_{i+1} = q_i P$.

P

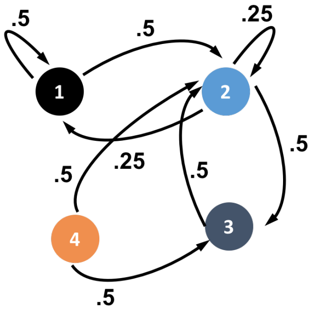
.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

q_0

1	0

Graph View

Often viewed as an underlying state transition graph. Nodes correspond to possible values that each X_i can take.



P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

The Markov chain is irreducible if the underlying graph consists of single strongly connected component.

2-SAT

Motivating Example: Find a satisfying assignment for a 2-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

A simple 'local search' algorithm:

1. Start with an arbitrary assignment.
2. Repeat $2mn^2$ times, terminating if a satisfying assignment is found:
   - Chose an arbitrary unsatisfied clause.
   - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.
3. If a valid assignment is not found, return that the formula is unsatisfiable.

Claim: If the formula is satisfiable, the algorithm finds a satisfying assignment with probability $\geq 1 - 2^{-m}$.

Randomized 2-SAT Analysis

Fix a satisfying assignment S. Let $X_i \leq n$ be the number of variables that are assigned the same values as in S, at step i.

S

1	0	1	0	0	0	1	1

assignment i

1	**0**	0	1	1	**0**	**1**	**1**

- $X_{i+1} = X_i \pm 1$ since we flip one variable in an unsatisfied clause.
- $\Pr(X_{i+1} = X_i + 1) \geq$
- $\Pr(X_{i+1} = X_i - 1) \leq$

$$(x_1 \vee \bar{x}_2) \wedge (\bar{x}_1 \vee \bar{x}_3) \wedge (x_1 \vee x_2) \wedge (x_4 \vee \bar{x}_3) \wedge (x_4 \vee \bar{x}_1)$$

Coupling to a Markov Chain

The number of correctly assigned variables at step i, X_i, obeys

$$\Pr(X_{i+1} = X_i + 1) \geq \frac{1}{2} \quad \text{and} \quad \Pr(X_{i+1} = X_i - 1) \leq \frac{1}{2}.$$

Is $X_0, X_1, X_2, \ldots$ a Markov chain?

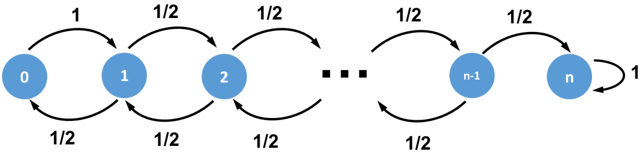
Define a Markov chain $Y_0, Y_1, \ldots$ such that $Y_0 = X_0$ and:

$$\begin{aligned}
\Pr(Y_{i+1} = 1 | Y_i = 0) &= 1 \\
\Pr(Y_{i+1} = j+1 | Y_i = j) &= 1/2 \;\; \text{for } 1 \leq j \leq n-1 \\
\Pr(Y_{i+1} = j-1 | Y_i = j) &= 1/2 \;\; \text{for } 1 \leq j \leq n-1 \\
\Pr(Y_{i+1} = n | Y_i = n) &= 1.
\end{aligned}$$

- Our algorithm terminates as soon as $X_i = n$. We expect to reach this point only more slowly with Y_i. So it suffices to argue that $Y_i = n$ with high probability for large enough i.
- Formally could use a coupling argument (see Chapter 11 of Mitzenmacher Upfal.)

Simple Markov Chain Analysis

Want to bound the expected time required to have $\mathbf{Y}_i = n$.



Let h_j be the expected number of steps to reach n when starting at node j (i.e., the expected termination time when j variables are assigned correctly.)

$$h_n = 0$$
$$h_0 = h_1 + 1$$
$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1 \text{ for } 1 \leq j \leq n-1$$

Simple Markov Chain Analysis

Claim: $h_j = h_{j+1} + 2j + 1$. Can prove via induction on j.

- $h_0 = h_1 + 1$, satisfying the claim in the base case.

$$\begin{aligned} h_j &= \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1 \\ &= \frac{h_j}{2} + (j-1) + \frac{1}{2} + \frac{h_{j+1}}{2} + 1 \\ &= \frac{h_j}{2} + \frac{h_{j+1}}{2} + j + \frac{1}{2}. \end{aligned}$$

- Rearranging gives: $h_j = h_{j+1} + 2j + 1$.

So in total we have:

$$h_0 = h_1 + 1 = h_2 + 3 + 1 = \ldots = \sum_{j=0}^{n-1}(2j+1) = n^2.$$

Simple Markov Chain Analysis

Upshot: Consider the Markov chain $Y_0, Y_1, \ldots$, and let i^* be the minimum i such $Y_{i^*} = n$. Then $\mathbb{E}[i^*] \leq n^2$.

- Thus, by Markov's inequality, with probability $\geq 1/2$, our 2-SAT algorithms finds a satisfying assignment within $2n$ steps.
- Splitting our $2nm$ total steps into m periods of $2n$ steps each, we fail to find a satisfying assignment in all m periods with probability at most $1/2^m$.

Check-in Question: For a fixed i, what roughly is $\mathbb{E}[Y_i]$?

3-SAT

More Challenging Problem: Find a satisfying assignment for a 3-CNF formula with n variables.

$$(x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (\bar{x}_1 \vee \bar{x}_3 \vee x_4) \wedge (x_1 \vee x_2 \vee \bar{x}_3).$$

- 3-SAT is famously NP-hard. What is the naive deterministic runtime required to solve 3-SAT?
- The current best known runtime is $O(1.307^n)$ [Hansen, Kaplan, Zamir, Zwick, 2019].
- Will see that our simple Markov chain approach gives an $O(1.3334^n)$ time algorithm.
- Note that the exponential time hypothesis conjectures that $O(c^n)$ is needed to solve 3-SAT for some constant $c > 1$. The strong exponential time hypothesis conjectures that for $k \to \infty$, solving k-SAT requires $O(2^n)$ time.

Randomized 3-SAT Algorithm

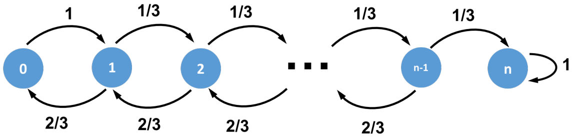
1. Start with an arbitrary assignment.
2. Repeat m times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.
3. If a valid assignment is not found, return that the formula is unsatisfiable.

Randomized 3-SAT Analysis

As in the 2-SAT setting, let X_i be the number of correctly assigned variables at step i. We have:

$$\Pr(X_i = X_{i-1} + 1) \geq$$
$$\Pr(X_i = X_{i-1} - 1) \leq$$

Define the coupled Markov chain $Y_0, Y_1, \ldots$ as before, but with $Y_i = Y_{i-1} + 1$ with probability $1/3$ and $Y_i = Y_{i-1} - 1 = 2/3$.



How many steps do you expect are needed to reach $Y_i = n$?

Randomized 3-SAT Analysis

Letting h_j be the expected number of steps to reach n when starting at node j,

$$h_n = 0$$
$$h_0 = h_1 + 1$$
$$h_j = \frac{2h_{j-1}}{3} + \frac{h_{j+1}}{3} + 1 \text{ for } 1 \leq j \leq n-1$$

- We can prove via induction that $h_j = h_{j+1} + 2^{j+2} - 3$ and in turn, $h_0 = 2^{n+2} - 4 - 3n$.
- Thus, in expectation, our algorithm takes at most $\approx 2^{n+2}$ steps to find a satisfying assignment if there is one.
- Is this an interesting result?

Modified 3-SAT Algorithm

Key Idea: If we pick our initial assignment uniformly at random, we will have $\mathbb{E}[X_0] = n/2$. With very small, but still non-negligible probability, X_0 will be much larger, and our random walk will be more likely to find a satisfying assignment.

Modified Randomized 3-SAT Algorithm:

Repeat m times, terminating if a satisfying assignment is found:

1. Pick a uniform random assignment for the variables.
2. Repeat $3n$ times, terminating if a satisfying assignment is found:
    - Chose an arbitrary unsatisfied clause.
    - Pick one of the variables in the clause uniformly at random, and switch the assignment of the variable.

If a valid assignment is not found, return that the formula is unsatisfiable.

Modified 3-SAT Analysis

Consider a single random assignment with $X_0 = n - j$. I.e., we need to correct j variables to find a satisfying assignment.

Let q_j be a lower bound on the success probability in this case. Since $j \leq n$ and since we run the search process for $3n$ steps,

$$\begin{aligned} q_j &= \Pr[X_{3n} = n] \\ &\geq \Pr[X_{3j} = n] \\ &\geq \Pr[\text{take exactly } 2j \text{ steps forward and } j \text{ steps back in } 3j \text{ steps}] \\ &= \binom{3j}{j} \left(\frac{2}{3}\right)^j \cdot \left(\frac{1}{3}\right)^{2j}. \end{aligned}$$

Via Stirling's approximation, $\binom{3j}{j} \geq \frac{1}{\sqrt{j}} \cdot \frac{3^{3j-2}}{2^{2j-2}}$, giving:

$$q_j \geq \frac{2^2}{3^2\sqrt{j}} \cdot \frac{3^{3j}}{2^{2j}} \cdot \frac{2^j}{3^{3j}} \approx \frac{1}{\sqrt{j} \cdot 2^j} \geq \frac{1}{\sqrt{n} \cdot 2^j}.$$

Modified 3-SAT Analysis

Our overall probability of success in a single trial is then lower bounded by:

$$
\begin{aligned}
q &\geq \sum_{j=0}^{n} \Pr[X_0 = n - j] \cdot q_j \\
&\geq \sum_{j=0}^{n} \binom{n}{j} \cdot \frac{1}{2^n} \cdot \frac{1}{\sqrt{n} \cdot 2^j} \\
&\geq \frac{1}{\sqrt{n} \cdot 2^n} \sum_{j=1}^{n} \binom{n}{j} \cdot \frac{1}{2^j} \\
&= \frac{1}{\sqrt{n} \cdot 2^n} \cdot \left(\frac{3}{2}\right)^n \leq \frac{1}{\sqrt{n}} \cdot \left(\frac{3}{4}\right)^n .
\end{aligned}
$$

Thus, if we repeat for $m = O\left(\sqrt{n} \cdot \left(\frac{4}{3}\right)^n\right) = O(1.33334^n)$ trials, with very high probability, we will find a satisfying assignment if there is one.

Gambler’s Ruin

Gambler's Ruin

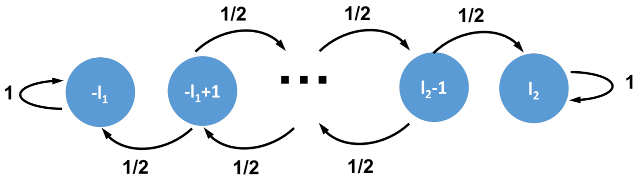
- You and 'a friend' repeatedly toss a fair coin. If it hits heads, you give your friend \$1. If it hits tails, they give you \$1.
- You start with $\$\ell_1$ and your friend starts with $\$\ell_2$. When either of you runs out of money the game terminates.
- What is the probability that you win $\$\ell_2$?

Gambler's Ruin Markov Chain

Let $X_0, X_1, \ldots$ be the Markov chain where X_i is your profit at step i. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{j,j+1} = P_{j,j-1} = 1/2 \text{ for } -\ell_1 < j < \ell_2$$



- ℓ_1 and ℓ_2 are absorbing states.
- All j with $-\ell_1 < j < \ell_2$ are transient states. I.e., $\Pr[X_{i'} = j \text{ for some } i' > i \,|\, X_i = j] < 1$.

Observe that this Markov chain is also a Martingale since $\mathbb{E}[X_{i+1}|X_i] = X_i$.

Gambler's Ruin Analysis

Let $X_0, X_1, \ldots$ be the Markov chain where X_i is your profit at step i. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{j,j+1} = P_{j,j-1} = 1/2 \text{ for } -\ell_1 < j < \ell_2$$

We want to compute $q = \lim_{i\to\infty} \Pr[X_i = \ell_2]$.

By linearity of expectation, for any i, $\mathbb{E}[X_i] = 0.$ Further, for $q = \lim_{i\to\infty} \Pr[X_i = \ell_2]$, since $-\ell_1, \ell_2$ are the only non-transient states,

$$\lim_{i\to\infty} \mathbb{E}[X_i] = \ell_2 q + -\ell_1(1-q) = 0.$$

Solving for q, we have $q = \frac{\ell_1}{\ell_1+\ell_2}$.

Gambler's Ruin Thought Exercise

What if you always walk away as soon as you win just $1. Then what is your probability of winning, and what are your expected winnings?