COMPSCI 614: Final Review

General Info: The final will be held **Tuesday May 14th from 10:30am-12:30pm in LGRC A203.** The test will be **closed book**, with no cheatsheets or calculators allowed. You must **show your work/derive any answers** as part of the solutions to receive full credit (and partial credit if you make a mistake).

Format: The format will be very similar to the midterm: the test will contain 4-5 questions. The first will be a mix of True/False or Always/Sometimes/Never style questions. The rest will be short answer style questions, like homework questions, but significantly less involved. The exam will be similar in length to the midterm as well, although you will have 2 hours, rather than 1 hour 15 minutes to complete it.

Studying Tips:

- Do as many practice problems as you can – including from the posted practice exam, from this review sheet, the books, the quizzes, the homeworks, and the 'Exercise' or 'Think-Pair-Share' questions given on the slides. For quizzes/homeworks/in class questions – try to re-solve without looking at the answer key or a solution given in the next slide. Then check to see how you did.
- For all practice questions, try to solve (and write down) a solution first without resources and somewhat quickly, as you would on the exam. Then go back and more slowly work through the problem, see if you solution is correct, etc.
- We encourage you to post on Piazza to check answers/discuss approaches.

Pre-Midterm and Last Class Material:

The final **will not** cover material from the last class on the probabilistic method. It will also not focus on material from before the midterm (Lectures 1-12 up to stochastic trace estimation). However, you should be able to use the tools developed in the first half of the course. E.g.,

- Basic probability calculations, applications of linearity of expectation, linearity of variance, concentration bounds, and union bound.
- Do not need to memorize any concentration bounds outside Markov's and Chebyshev's.
- Idea of proving communication lower bounds and other lower bounds via pigeonhole principle style arguments.
- Basics of randomized linear algebra – ability to work with randomly sampled matrices and vectors.

1 Concepts to Study

Randomized Numerical Linear Algebra

- Subspace embedding definition.
- ϵ-net definition, and motivation for why they are used to prove subspace embedding from the Johnson-Lindenstrauss lemma. Should understand high level idea of this proof, but don't need to memorize details.
- The Johnson-Lindenstrauss lemma statement and ability to apply Hanson-Wright inequality. Do not need to memorize Hanson-Wright.
- Application of subspace embeddings to approximate regression
- Definition of leverage scores and high level idea of subspace embedding via leverage score sampling and matrix Chernoff bound. Do not need to memorize matrix Chernoff bound.
- Loewner ordering notation. I.e., for $M, N \in \mathbb{R}^{d \times d}$, $M \preceq N$ if for all $x \in \mathbb{R}^d$, $x^T M x \leq x^T N x$.
- Variational characterization of the leverage scores.
- Spectral sparsifier definition and its connection to cut preservation and subspace embedding.
- Equivalence between leverage scores of the vertex-edge incidence matrix and effective resistances. Don't need to know the proof (we didn't cover it).

Markov Chains

- Definition of a Markov chain and view in terms of state graph and transition matrix.
- High level idea behind Markov chain based 2-SAT and 3-SAT algorithms. Don't need to memorize the analysis but should understand the tools used. E.g., solving for the expected number of steps to reach a satisfying assignment via a linear recurrence.
- Gambler's ruin set up, analysis, and conclusion.
- Definition of a stationary distribution, and conditions for having a stationary distribution and for that distribution being unique (fundamental theorem of Markov chains).
- Related to the above, irreducibility and aperiodicity. You should be able to recognize and prove when a Markov chain is/is not irreducible/aperiodic.
- Fact that symmetric Markov chains have uniform stationary distributions.
- Definition of total variation (TV) distance, and mixing time.
- Min/Max view of the TV distance (Kontorovich-Rubinstein duality) and implication for bounding mixing time via coupling.
- Ability to construct a coupling for a Markov chain and analyze its coupling time.
- Metropolis-hastings algorithm – should understand why it achieves the desired stationary distribution, proportional to the density $p(\cdot)$.

- Ability to design a Markov chain that converges to a given distribution (like the independent set examples using a symmetric Markov chain and Metropolis-Hastings covered in class).
- Definition of a reversible Markov chain and the implication that P has real eigenvalues in this case.
- Do not need to know counting-sampling reductions in detail.

Other

- High level idea of what sort of problems convex relaxation + randomized rounding would be applied to and how to analyze expected solution quality for randomized rounding.

2 Practice Questions

1. Subspace Embeddings

1. Give an upper bound on required the size of an ϵ net over the d-dimensional cube $[-1,1]^d$. Use as volume argument to show that your upper bound is tight up to constant factors.
2. Consider two matrices $A, B \in \mathbb{R}^{n\times d}$ such that $A = BC$ for some invertible $C \in \mathbb{R}^{d\times d}$. If $\mathbf{S}$ is an ϵ-subspace embedding for A, then $\mathbf{S}$ is also an ϵ-subspace embedding for B. ALWAYS SOMETIMES NEVER.
3. True of False: For any matrix $A \in \mathbb{R}^{d\times d}$, $2A \succeq A$. If true, why? If false, what is one assumption you can make on A so that it is true?
4. True or False: For any matrix $A \in \mathbb{R}^{n\times d}$, there is some matrix $S \in \mathbb{R}^{d\times n}$ such that, for all $x \in \mathbb{R}^d$, $\|SAx\|_2 = \|Ax\|_2$.
5. Prove directly, without using an ϵ-net that if $S \in \mathbb{R}^{m\times n}$ is a random ± 1 matrix with $m = O\left(\frac{d+\log(1/\delta)}{\epsilon^2}\right)$, and $A \in \mathbb{R}^{n\times d}$ is any matrix, then with probability $\geq 1-\delta$, for all $x \in \{0,1\}^d$, $(1-\epsilon)\|Ax\|_2 \leq \|SAx\|_2 \leq (1+\epsilon)\|Ax\|_2$.

2. Leverage Scores and Spectral Sparsifiers:

1. There exists a spectral sparsifier of a connected graph G with $< n-1$ edges. ALWAYS SOMETIMES NEVER
2. For a matrix $A \in \mathbb{R}^{n\times d}$ with rows $a_1, \ldots, a_n \in \mathbb{R}^d$, the leverage score of the i^{th} row τ_i, satsifies $\tau_i = \|a_i\|_2^2$ ALWAYS SOMETIMES NEVER.
3. Consider two matrices $A, B \in \mathbb{R}^{n\times d}$ such that $A = BC$ for some invertible $C \in \mathbb{R}^{d\times d}$. How do the leverage scores of A compare to those of B?
4. Let G be the complete graph on n-nodes, and let $\tilde{G}$ be a 1/2-spectral sparsifier of G. Assume that $\tilde{G}$ has $O(n\log n)$ edges. Argue that $\tilde{G}$ has at least one edge in it with weight at least $\Omega(n/\log n)$. **Hint:** Think about how $\tilde{G}$ preserves cuts in G.

3. Markov Chains:

1. Exercises 7.3, 7.6, 7.7, 7.11, 7.20, 7.21 from Mitzenmacher, Upfal.

2. Is a random walk on a connected, undirected graph always irreducible? Is it always aperiodic? What about on a connected undirected graph where one of the nodes has a self loop?

3. Let P be the uniform distribution on the integers $\{1, 2, \ldots, 100\}$. Let Q be the uniform distribution on the even integers $\{2, 4, 6, ..., 100\}$. What is $\|P - Q\|_{TV}$?

4. Prove formally the claim made in class that if $q_{t,i}$ is the state distribution of a Markov chain after taking t steps starting from state i, and π is a stationary distribution of the chain, then $\|q_{t+1,i} - \pi\|_{TV} \leq \|q_{t,i} - \pi\|_{TV}$. Does this fact require that the Markov chain is aperiodic and irreducible? Does it require the Markov chain to have a unique stationary distribution?

5. Consider the 'Glauber dynamics' for sampling an independent set: to generate set $\mathbf{X}_{i+1}$ from set $\mathbf{X}_i$, sample a random vertex v from the graph. Let $\mathbf{X}' = \mathbf{X}_i \cup \{v\}$ with probability $1/2$ and $\mathbf{X}' = \mathbf{X}_i \setminus \{v\}$ with probability $1/2$. If $\mathbf{X}'$ is an independent set, let $\mathbf{X}_{i+1} = \mathbf{X}'$. Else, let $\mathbf{X}_{i+1} = \mathbf{X}_i$. Is this Markov chain irreducible and aperiodic? What is its stationary distribution?

6. Describe a Markov chain whose stationary distribution is the uniform distribution over valid Δ-colorings of a graph. I.e., assignments of each vertex to one of Δ colors, such that no two vertices with the same color are connected by an edge. Assume that there is at least one valid Δ-coloring of the graph.

7. Is the Markov chain you found above irreducible and aperiodic?

8. For a valid coloring X of a graph G, let $c(X)$ be the number of unique colors used in that coloring. Observe that $c(X) \leq n$ where n is the number of nodes and $c(X) \geq \chi(G)$, where $\chi(G)$ is the chromatic number of G. Describe a Markov chain, which is both irreducible and aperiodic, whose stationary distribution samples a valid coloring X with probability $\pi(X) = \frac{\lambda^{c(X)}}{\sum_{\text{valid colorings } Y} \lambda^{c(Y)}}$, for some parameter λ.

9. Consider an irreducible, aperiodic Markov chain, where all states transition to a single 'home state' h with probability $1/c$. I.e., $P_{i,h} = c$ for all i. Give an upper bound on the ϵ-mixing time, $\tau(\epsilon)$ for this chain.

10. Describe a Markov chain for which *any distribution* $\pi \in [0,1]^m$ is a stationary distribution.