COMPSCI 614: Randomized Algorithms with Applications to Data Science

Prof. Cameron Musco

University of Massachusetts Amherst. Spring 2024.

Lecture 23

Logistics

- Optional Problem Set 5 due 5/13 at 11:59pm.
- Final exam will be **Tuesday 5/14, 10:30-12:30pm in the classroom**. See Piazza post for info on study materials.
- I will hold additional final review office hours **Monday 5/13 from 3-4:30pm**.
- Final project due the last day of finals: Friday 5/17 – if you have questions as you come into the last couple of weeks of the project feel free to reach out.

Summary

Last Time:

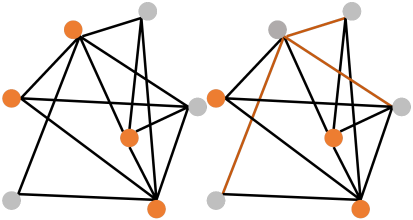
- Finish Markov chain unit.
- Analysis of Metropolis Hastings algorithm
- Example sampling to counting reduction for independent sets.

Today:

- Convex relaxation + randomized rounding for NP-Hard problems.
- Example application to vertex cover and set cover.

Many NP-hard optimization problems can be formulated as convex optimization problems subject to integral constraints.

Example 1: Vertex cover – find a minimum set of vertices such that any edge in a graph is covered by at least one vertex.



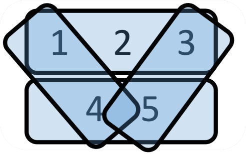
$$\min \sum_{i=1}^{n} x_v \quad s.t. \quad x_u + x_v \geq 1 \text{ for all } (u, v) \in E$$

$$x_i \in \{0, 1\} \text{ for all } i \in [n].$$

Combinatorial Optimization

Many NP-hard optimization problems can be formulated as convex optimization problems subject to integral constraints.

Example 2: Set cover – given a universe of elements $[n]$ and a collection of sets $S_1, S_2, \ldots, S_m \subseteq [n]$, find the minimum number of sets that cover all items in $[n]$.



$$\min \sum_{i=1}^{m} x_i \quad s.t. \quad \sum_{i:j\in S_i} x_i \geq 1 \text{ for all } j \in [n]$$
$$x_i \in \{0,1\} \text{ for all } i \in [m].$$

Applications Beyond Theory

Convex optimization problems with non-convex constraints arise all over the place outside of algorithms textbooks.

- Sparse linear regression: $\min_{x:\|x\|_0 \le k} \|Ax - b\|_2^2$.
- Low-rank matrix completion: $\min_{M:\mathrm{rank}(M) \le k} \sum_{(i,j)\in\Omega} [B_{i,j} - M_{i,j}]^2$.
- Matching matrices with permutations: $\min_{\text{permutation matrices } P_1, P_2} \|A - P_1 B P_2\|_F^2$. Recently, these types of problems are very relevant e.g. in identifying permutation invariances in neural networks.

Convex Relaxation

- **Step 1:** ‘Relax’ the non-convex constraint to be a related (and weaker) convex constraint.
- **Step 2:** Solve the resulting convex problem in polynomial time.
- **Step 3:** Map the relaxed solution back to a solution to the original problem. For integral constraints this is called ‘rounding’.

Key Challenge: Need to argue that the rounding step both gives a feasible solution and does not increase the cost of the relaxed solution too much.

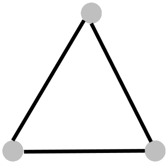
Applications: This very general approach yields the best known approximation algorithms for a huge range of problems: set cover, vertex cover, max-cut (Goemans-Williamson SDP), etc. In many cases, the approximation ratios obtained are known to be optimal under complexity theoretic assumptions.

Vertex Cover Relaxation

$$\min \sum_{i=1}^{n} x_v \quad s.t. \quad x_u + x_v \geq 1 \text{ for all } (u,v) \in E$$

$$x_i \in \{0,1\}[0,1] \text{ for all } i \in [n].$$

- This is now a linear program. It can be solved in polynomial time.
- A solution may no longer be a valid vertex cover.



- How should be round to solution to obtain a true vertex cover?

Vertex Cover Relaxation

Deterministic Rounding for Vertex Cover: Given a fractional solution $\tilde{x}_1, \ldots, \tilde{x}_n$, obtain integral solution $x_1, \ldots, x_n$ by applying the rule: if $\tilde{x}_u \geq 1/2$, set $x_u = 1$. if $\tilde{x}_u < 1/2$, set $x_u = 0$.

Claim 1: The rounded solution is feasible.

Proof: For any $(u, v) \in E$, we must have $x_u + x_v \geq 1$, and thus at least one of x_u or $x_v \geq 1/2$. So all edges are covered in the rounded solution.

Claim 2: The rounded solution is within a 2-factor of optimal.

Proof: $\sum_{i=1}^{n} x_i \leq 2 \sum_{i=1}^{n} \tilde{x}_i = 2 \cdot OPT_{relax} \leq 2 \cdot OPT$.

Vertex Cover Integrality Gap

Could we do any better than a 2-approximation for vertex cover via this approach?

- There exist graphs for which $OPT_{relax} \leq OPT/2$. I.e., this relaxation has an integrality gap of 2.
- So any rounding scheme must at least double OPT_{relax} in the worst case, or would have to be infeasible on such graphs.
- Since there also exist solutions where $OPT_{relax} = OPT$, this makes it unlikely to get an approximation factor better than 2 for this problem.
- Assuming the unique games conjecture, vertex cover is hard to approximate to a factor better than 2 in general [Khot, Regev '08]]. Assuming $P \neq NP$ it cannot be approximated to a factor better than ≈ 1.36 [Dinur, Safra '05].

Set Cover Relaxation

$$\min \sum_{i=1}^{m} x_i \quad s.t. \quad \sum_{i:j \in S_i} x_i \geq 1 \text{ for all } j \in [n]$$

$$x_i \in \{0,1\}[0,1] \text{ for all } i \in [m].$$

Will deterministic rounding work here?

1 2 3	1 2 4
1 3 4	2 3 4

Randomized Rounding for Set Cover

Naive Randomized Rounding: Given a fractional set cover solution $\tilde{x}_1, \ldots, \tilde{x}_m$, obtain integral solution $x_1, \ldots, x_m$ by independently setting $x_i = 1$ with probability $\tilde{x}_i$ and 0 otherwise.

- What is the expected cost $\mathbb{E}[\sum_{i=1}^{m} x_i]$?
- Is the rounded solution feasible?
- No with pretty good probability. Consider an item that is covered by t sets, each with weight $1/t$.
  $\Pr[\text{not feasible}] = (1 - 1/t)^t \approx 1/e$.
- How could we fix this issue?

Randomized Rounding for Set Cover

Scaled Randomized Rounding: Given a fractional set cover solution $\tilde{x}_1, \ldots, \tilde{x}_m$, obtain integral solution $x_1, \ldots, x_m$ by independently setting $x_i = 1$ with probability $\min(1, \alpha \cdot \tilde{x}_i)$ and 0 otherwise.

- **Expected cost:** $\mathbb{E}[\sum_{i=1}^m x_i] = \sum_{i=1}^m \min(1, \alpha \tilde{x}_i) \leq \alpha \sum_{i=1}^m \tilde{x}_i \leq \alpha \cdot OPT$.
- **Feasibility:** For any given item j, if there is some $S_i \ni j$ with $\tilde{x}_j = 1$, and so j is covered.
- Otherwise, $\mathbb{E}[\sum_{i:j \in S_i} x_i] = \alpha \cdot \sum_{i:j \in S_i} \tilde{x}_i \geq \alpha$.
- How big must we set α such that, with probability at least $1 - 1/n^c$, $\sum_{i:j \in S_i} x_i \geq 1$? $\alpha = O(\log n)$ suffices via a Chernoff bound
- By a union bound over all n items, the solution will be feasible with probability at least $1 - 1/n^{c-1}$.

Set Cover Approximation Via Randomized Rounding

Upshot: We obtain a $O(\log n)$ approximation algorithm for Set Cover via relaxation + randomized rounding.

- The natural Set Cover LP relaxation has an integrality gap of $\Omega(\log n)$.
- Assuming $P \neq NP$ this approximation factor is optimal up to constants [Raz, Safra '97].
- A simple deterministic greedy algorithm also gives an $O(\log n)$ approximation factor: at each step pick the set that covers the most number of previously uncovered elements.

Bonus Slides: Semidefinite Programming Relaxation of Max-Cut

Max-Cut

Given a graph G output the sets of vertices S such that the number of edges between S and $V \setminus S$ is maximized.

- Decision version is NP-Hard.
- If $P \neq NP$ no algorithm gives better than $16/17$ approximation.
- Best known algorithm is the Goemans-Williamson algorithm, which is based on convex relaxation and randomized rounding. Gives ≈ 0.878 approximation.
- This is optimal assuming the Unique Games Conjecture.

Max-Cut SDP Formulation

$$\max \frac{1}{2} \sum_{(u,v)\in E} (1 - x_u x_v) \quad \text{s.t.} \quad x_v \in \{-1, 1\} \text{ for all } v \in V.$$

- If we just relax $x_v \in [-1, 1]$, this problem is not convex.
- Instead, Goemans and Williamson relax the problem by letting the x_v be unit vectors in $\mathbb{R}^n$:

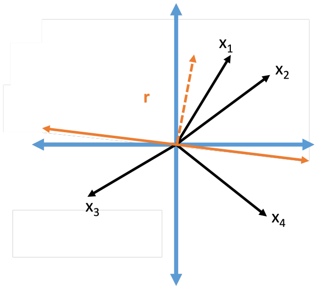
$$\max \frac{1}{2} \sum_{(u,v)\in E} (1 - \langle x_u, x_v \rangle) \quad \text{s.t.} \quad x_v \in \mathbb{R}^n, \|x_v\|_2 = 1 \text{ for all } v \in V.$$

- This is a valid relaxation – given an integral solution could set $\tilde{x}_v = [x_v, 0, 0, 0, \ldots]$ and achieve the same cost.
- Further it can be solved in polynomial time as a semidefinite program (SDP).

Max-Cut Rounding

To round the Max-Cut SDP relaxation, Goemans and Williamson use the following procedure:

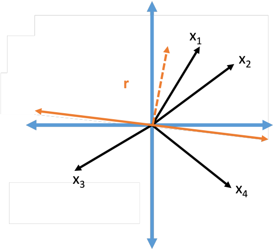
- Let $r \in \mathbb{R}^n$ be a uniform random point with $\|r\|_2 = 1$.
- Let $x_v = 1$ if $\tilde{x}_v : \langle x_v, r \rangle \geq 0$, and $x_v = 0$ otherwise.



Note that the output solution is always a valid cut. So the main challenge is to prove the approximation ratio.

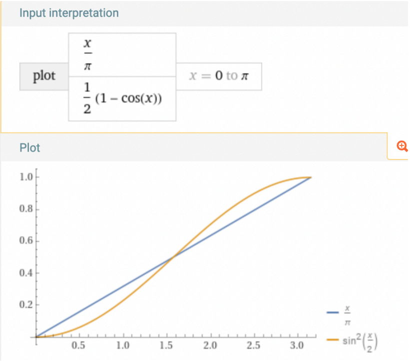
Max-Cut Approximation Ratio

- Focusing on just a single edge (u, v), the relaxed solution gives value $\frac{1-\langle x_u, x_v\rangle}{2} = \frac{1-\cos\theta}{2}$ where θ is the angle between x_u and x_v.
- The rounded solution gives value 1 if x_u and x_v are rounded to different sides of the cut (and value 0 otherwise). What is the probability of this happening? θ/π.



- Thus, summing over all edges, the Goemans Williamson algorithm has expected approximation ratio at least $\min_\theta \frac{\theta/\pi}{\frac{1-\cos\theta}{2}} \approx 0.878$.

Max-Cut Approximation Ratio



- If you took 514 you may recognize that this analysis is very closely related to the SimHash locality sensitive hashing algorithm, and in turn the JL Lemma.
- In fact SimHash, which is used e.g. for high dimensional approximate near neighbor search is exactly the rounding scheme from Goemans Williamson.