Lecture #19

Last Time: Applications of MWU

(Adaboost, and solving optimization problems via ZSGs)

Today: MAXCUT and Semidefinite Programs

Earlier we saw a framework for designing approx. algorithms:

$$\text{ILP} \longrightarrow \text{LP Relaxation}$$

MAXCUT: Given $G=(V,E)$, $\max_{U \subseteq V} \underbrace{|E(U, V \setminus U)|}_{\text{edges cut}}$

e.g. If G is bipartite, can cut all edges

MAXCUT is NP-hard, let's try to approximate it:

ILP: $\max \sum_{(u,v) \in E} z_{uv}$

$$\text{s.t. } z_{uv} \leq x_u + x_v \qquad \forall u,v$$
$$z_{uv} \leq (1-x_u)+(1-x_v)$$
$$z_{u,v}, x_u \in \{0,1\} \qquad \forall u,v$$

If we drop integer constraint can set $x_u \equiv \frac{1}{2}$ $\forall u$, and $z_{uv} \equiv 1$ $\forall u,v$.

Thus the optimal value of the LP is $|E|$, even though there are graphs where MAXCUT $\sim \frac{|E|}{2}$

At best, rounding the LP leads to a $\frac{1}{2}$-approx.
In fact, there is a trivial way to cut $\frac{1}{2}$ of edges:

Choose to include each u in U independently, with probability $1/2$

Is there a better approx. algorithm?

Today we will introduce a new, powerful family of convex relaxations:

We say $X \in \mathbb{R}^{n \times n}$, symmetric is <u>positive semidefinite</u>, written $X \succeq 0$, if

(1) $a^T X a \geq 0 \quad \forall a \in \mathbb{R}^n$

(2) $X = BB^T$ (Cholesky decomposition)

(3) X's eigenvalues are all nonnegative

Now we can introduce <u>semidefinite programming</u>:

$$(P) \quad \max \langle C, X \rangle := \sum_{ij} C_{ij} X_{ij} \quad \text{(matrix inner-product)}$$

C is $n \times n$

$$\text{s.t. } \langle A_i, X \rangle = b_i \quad \forall i = 1 \text{ to } m$$

$$X \succeq 0$$

Note: degenerates to an LP if X is diagonal

the feasible region $K = \{X \mid \langle A_i, X \rangle = b_i \ \forall i, X \succeq 0\}$

can alternatively be thought of as

$$K = \{X \mid \langle A_i, X \rangle = b_i \ \forall i, \ a^T X a \geq 0 \ \forall a\}$$

Hence it is no longer the intersection of finite number of linear inequalities

Claim: But we can optimize linear functions over K using the Ellipsoid Algorithm

Caveat: Only get optimal value within any desired accuracy $\pm\varepsilon$, cannot "jump" to optimal value as in LPs

---

Is there a notion of duality here?

(D) $\max b^T y$

$$\text{s.t.} \quad \sum y_i A_i + S = C$$

$$S \succeq 0$$

Let's develop some basic facts!

Fact: $\langle A, X\rangle = \text{Trace}(A^T X)$

Fact: $\text{Trace}(AB) = \text{Trace}(BA)$

(Does this mean $\text{Trace}(ABC) = \text{Trace}(CBA)$? No!)

The following lemma will prove very useful:

Lemma 1: Let X be symmetric. Then $X \succeq 0$ iff $\langle A, X\rangle \geq 0 \ \forall A \succeq 0$

Proof: Suppose X is not PSD, let $a^T X a < 0$. then

$$\langle \underbrace{aa^T}_{A}, X \rangle = \text{trace}(aa^T X) = \text{Trace}(a^T X a) = a^T X a < 0$$

Thus there is an $A \succeq 0$ s.t. $\langle A, X \rangle < 0$.

Suppose $X \succeq 0$. Then consider $A \succeq 0$, and by Cholesky

$$A = BB^T = \sum_i b_i b_i^T$$

Then $\langle A, X \rangle = \sum_i b_i^T X b_i \geq 0$ ☒

Now we can prove <u>weak duality</u>:

<u>lemma 2</u>: If X/y are feasible for (P)/(D) then

$$b^T y \leq \langle C, X \rangle$$

<u>Proof</u>: We have:

$$\langle C, X \rangle \underset{\substack{\text{by feasibility} \\ \text{of } y}}{=} \underbrace{\langle \sum y_i A_i, X \rangle}_{\substack{= \sum y_i b_i \\ \text{by feasibility of } X}} + \underbrace{\langle S, X \rangle}_{\substack{\geq 0 \\ \text{by feasibility of } X \\ \text{and lemma 1}}}$$

☒

<u>Warning</u>: Strong duality usually holds (under say Slater's condition) but is more subtle than for LPs

Slater's condition $\equiv$ strong duality holds if feasible region has interior point

Let's return to MAXCUT and try a new relaxation following [Goemans-Williamson]

$$\max \sum_{(u,v)\in E} \left(\tfrac{1}{2} - \tfrac{1}{2} X_{u,v}\right)$$

$$\text{s.t. } X_{uu} = 1 \ \forall u$$

$$X \succeq 0$$

Why is this a relaxation? Let $a_u = \begin{cases} 1 & u \in U \\ -1 & \text{else} \end{cases}$, then

$X = aa^T$ is feasible, and moreover

$$X_{uv} = \begin{cases} -1 & \text{if } u,v \text{ cut by } (U, V\setminus U) \\ 1 & \text{else} \end{cases}$$

Now by Cholesky, given a feasible X we can find a collection of unit vectors $\{y_u\}$ s.t. $X_{uv} = \langle y_u, y_v \rangle$

Theorem [Goemans-Williamson]
Let $\alpha_{GW} := \min_{0 \le \theta \le \pi} \frac{2}{\pi} \frac{\theta}{1-\cos\theta} \approx 0.87856\ldots$

Then there is an α-approx. algorithm for MAXCUT

The algorithm first finds an optimal soln to the SDP, then performs hyperplane rounding:

Choose a uniformly on the sphere

Set $x_u = \text{sgn}(\langle a, y_u \rangle)$

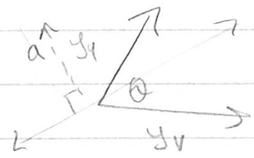
To analyze this approx. algorithm, let's understand the contribution of an edge (u,v)

contribution to SDP:

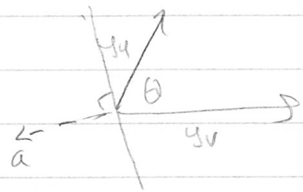
$$\frac{1}{2} - \frac{1}{2}\langle y_u, y_v \rangle$$

$$\underbrace{\qquad\qquad}_{\frac{1}{2}(1-\cos\theta)}$$

contribution to cut (WLOG $0 \le \theta \le \pi$)



an a that cuts the edge (u,v) $\Big\}\ \frac{\theta}{\pi}$



an a that does not cut the edge (u,v)

Key: whether (u,v) is cut or not only depends on the projection of a onto $\mathrm{span}(y_u, y_v)$.

Thus the worst-case ratio of contributions is:

$$\min_{0 \le \theta \le \pi} \frac{2\theta}{\pi(1-\cos\theta)}$$

Theorem [Khot, Kindler, Mossel, O'Donnell] [Mossel, O'Donnel, Oleszkiewicz] It is Unique Games Hard to approximate MAXCUT better than α_{GW}

The UGC is one of the most far-reaching conjectures in complexity. If true, many problems we have matching upper and lower bounds for

(In some cases, we don't know the actual constant)

But maybe it is too good to be true, and it's false?