CSE 5401 Lec. 19

Randomized Rounding of Semidefinite Programs

1. Random hyperplane rounding
2. Approximation guarantee for Max Cut
3. Guarantee for positive semidefinite objectives

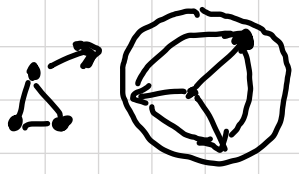
We were motivated to consider semidefinite programs as a family of quadratic optimization problems that we can solve efficiently. In general, we noticed that the maximum cut problem posed a barrier: given an undirected graph $G=(V,E)$, find $S \subseteq V$ that maximizes $|E(S)| = |\{(u,v) \in E : u \in S, v \notin S\}|$ (edges crossing S). This NP-complete problem was captured precisely by

$$\text{maximize } \frac{1}{4}\sum_{(u,v)\in E} (s_u - s_v)^2 \text{ s.t. } s_v^2 = 1 \text{ for all } v \in V.$$

We noticed that we could relax the problem to the semidefinite program maximize $\frac{1}{4}\sum_{(i,j)\in E} \|\vec{v}_i - \vec{v}_j\|_2^2$ s.t. $\|\vec{v}_i\|_2^2 = 1$ for all i. where we can solve the semidefinite program in polynomial time (up to a small additive error). We now turn to the question of how we can use the vector solutions to the relaxation to obtain approximate solutions to the original quadratic program. In particular, for Max Cut, we want to recover a good S, corresponding to an assignment of values in $\{1,-1\}$ to the s_v vars. We refer to this mapping from $\mathbb{R}^n$ to $\{1,-1\}$ as "rounding".

To motivate the method, let's interpret the semidefinite program relaxation: each vertex is assigned a vector $\vec{v}$ with length $\|\vec{v}\|_2 = 1$, so it lies on the unit ball in $\mathbb{R}^n$. The objective maximizes the sum of the distances between the vectors

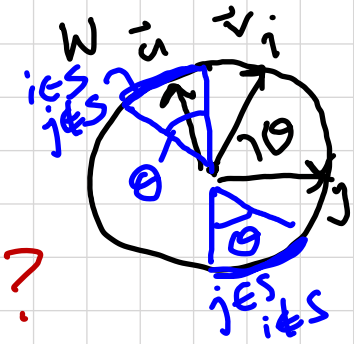
corresponding to vertices joined by edges — so these are pushed as far apart on average as possible. So, cuts that correspond to vertices that have been pushed close together should contain few edges (have many leaving the cut).



The method we use is to choose S containing the vertices on half of the sphere chosen uniformly at random. In other words, we pick $\vec{r} \in \mathbb{R}^n$ with $\|\vec{r}\|_2 = 1$ uniformly at random. Then we put $S = \{i : \langle \vec{v}_i, r \rangle \geq 0\}$.

Approximation guarantee for Max Cut

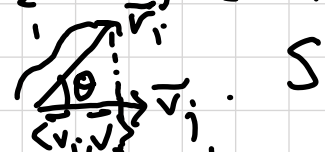
How many edges are cut by S in expectation? To analyze this, we'll consider a single edge $(i,j) \in E$. Observe $(i,j) \in E(S)$ if $\langle \vec{v}_i, \vec{r} \rangle \geq 0$ and $\langle \vec{v}_j, \vec{r} \rangle < 0$ or $\langle \vec{v}_i, \vec{r} \rangle < 0$ and $\langle \vec{v}_j, \vec{r} \rangle \geq 0$. To find the probability of this event, it's enough for us to consider the projection of $\vec{r}$ on to the plane $W = \text{span}\{\vec{v}_i, \vec{v}_j\}$ since the component of $\vec{r}$ in the orthogonal subspace $W^\perp$ is orthogonal to $\vec{v}_i$ and $\vec{v}_j$ — i.e., $\vec{r} = \text{proj}_W(\vec{r}) + \text{proj}_{W^\perp}(\vec{r})$ where $\langle \vec{r}, \vec{v}_i \rangle = \langle \text{proj}_W(\vec{r}), \vec{v}_i \rangle$, $\langle \vec{r}, \vec{v}_j \rangle = \langle \text{proj}_W(\vec{r}), \vec{v}_j \rangle$. Furthermore, since membership in S is determined by the sign of $\langle \vec{r}, \vec{v}_i \rangle$ and $\langle \vec{r}, \vec{v}_j \rangle$, we can rescale these, and just consider $\vec{u} = \frac{\text{proj}_W(\vec{r})}{\|\text{proj}_W(\vec{r})\|_2}$. Notice: $\vec{u}$ is distributed uniformly at random on the unit circle in W, where $\vec{v}_i$ and $\vec{v}_j$ are points on the circle.



Suppose the angle between $\vec{v}_i$ and $\vec{v}_j$ is θ.
What is the probability $(i,j) \in E(S)$ in terms of θ?

$\frac{2\theta}{2\pi}$ (in radians): $i \in S$ and $j \notin S$ if $\vec{u}$ falls in the sector of angle θ between the vector $\frac{\pi}{2}$ counter-clockwise from $\vec{v}_i$ and $\frac{\pi}{2}$ counter-clockwise from $\vec{v}_j$. Similarly, $i \notin S$ and $j \in S$ if $\vec{u}$ falls in the sector between the vectors $\frac{\pi}{2}$ clockwise from $\vec{v}_i + \vec{v}_j$. Each has prob. $\frac{\theta}{2\pi}$ of being hit since $\vec{u}$ is uniformly distributed.

Thus, the expected number of edges cut is $\sum_{(i,j)\in E} \frac{1}{\pi}\theta(\vec{v}_i, \vec{v}_j)$.

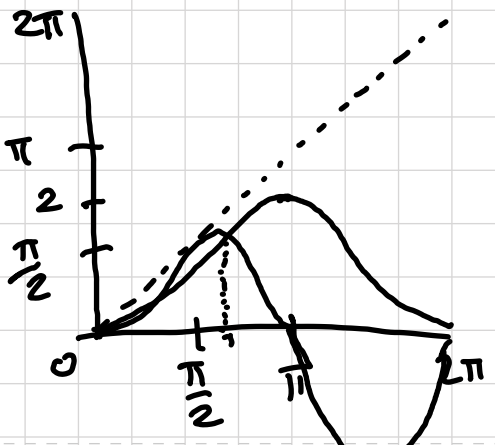
To relate this to the objective of the semidefinite program notice that since $\|\vec{v}_i\|_2 = \|\vec{v}_j\|_2 = 1$, $\|\vec{v}_i - \vec{v}_j\|_2^2 = 2 - 2\langle \vec{v}_i, \vec{v}_j \rangle$ where $\langle \vec{v}_i, \vec{v}_j \rangle = \cos\theta(\vec{v}_i, \vec{v}_j)$ 

. So, the objective may be written $\frac{1}{2}\sum_{(i,j)\in E}(1 - \cos\theta(\vec{v}_i, \vec{v}_j))$. It is therefore enough to bound the ratio $\frac{\frac{1}{\pi}\theta(\vec{v}_i, \vec{v}_j)}{\frac{1}{2}(1-\cos\theta(\vec{v}_i, \vec{v}_j))}$: if this is at least α for all θ, we have term-by-term that $\frac{1}{\pi}\theta(\vec{v}_i, \vec{v}_j) \geq \alpha \cdot \frac{1}{2}(1 - \cos\theta(\vec{v}_i, \vec{v}_j))$, and then summing over $(i,j) \in E$, the expected size of $E(S)$ is at least $\alpha \cdot \frac{1}{2}\sum_{(i,j)\in E} 1 - \cos\theta(\vec{v}_i, \vec{v}_j)$ (our objective value). Since we argued previously that the semidefinite program is a relaxation of the Max Cut problem, $\frac{1}{2}\sum_{(i,j)\in E} 1 - \cos\theta(\vec{v}_i, \vec{v}_j) \geq \max_{S'} |E(S')|$. So in expectation, we get a cut S with $\alpha \cdot \max_{S'} |E(S')|$ edges.

So, we bound $\frac{2}{\pi}\cdot\frac{\theta}{1-\cos\theta}$. Recalling $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$, $1 - \cos x$ is $\frac{x^2}{2} + O(x^4)$, so $\frac{\theta}{1-\cos\theta} \to \infty$ as $\theta \to 0$. (It clearly also diverges at $\theta = 2\pi$).

$\frac{d}{d\theta}\left[\frac{\theta}{1-\cos\theta}\right] = \frac{1}{1-\cos\theta} - \frac{\theta\sin\theta}{(1-\cos\theta)^2}$, so apart from the singularities at 0 & 2π, the only critical point(s) are $1 - \cos\theta = \theta\sin\theta$.

One occurs at $\theta \approx 2.331$, where $\frac{2}{\pi}\frac{\theta}{1-\cos\theta} = 0.8786\cdots$ This is the minimum.



(Near 0, $x\sin x = x^2 - \frac{x^4}{3!} + \cdots > \frac{x^2}{2} - \frac{x^4}{4!} + \cdots = 1 - \cos x$)

Thus, we find that random hyperplane rounding obtains a $0.8786\cdots$ approximation to the maximum cut. This is the state-of-the-art for the Max Cut problem and is conjectured to be the best approximation attainable by polynomial-time algorithms. Actually, for a large family of problems where we are maximizing the number of satisfied constraints, each of which only depend on a bounded number of variables, a similar algorithm based on rounding a semidefinite program relaxation is conjectured to give the optimal polynomial-time approximation

Random hyperplane rounding for positive semidefinite objectives

What about other quadratic optimization problems? In particular, we can consider integer optimization problems of the form $\underset{\vec{x}\in\{-1,1\}^n}{\text{maximize}}\ \vec{x}^T A\vec{x}$. In general, the objective could be negative, which makes the analysis of approximation challenging (sometimes $\alpha\,\vec{x}^TA\vec{x} > \vec{x}^TA\vec{x}$ for $\alpha<1$ if $\vec{x}^TA\vec{x}<0$). We focus on $A\succeq 0$, therefore, since then $\vec{x}^TA\vec{x}\geq 0$ always. We still need to address the issue that A_{ij} may be negative, which prevents us from simply bounding the individual terms and summing over all (i,j). We will still use the random hyperplane method, obtaining $x_i=-1$ if $\langle\vec{r}_i,\vec{v}_i\rangle<0$ and $x_i=1$ if $\langle\vec{r}_i,\vec{v}_i\rangle\geq 0$. Then

Lemma $\mathbb{E}[x_ix_j] = \frac{2}{\pi}\arcsin(\langle\vec{v}_i,\vec{v}_j\rangle)$.

Proof: What is the probability $x_ix_j=1$? $x_ix_j=1$ if $x_i=x_j$, if $i\neq j$ would lie on the same side of the cut. Similarly, $x_ix_j=-1$ if $i\neq j$ are on different sides. We saw this was with probability $\frac{\theta}{\pi}$ where $\theta=\arccos(\langle\vec{v}_i,\vec{v}_j\rangle)$. Thus, $\Pr[x_ix_j=1] = 1-\frac{\theta}{\pi} = 1-\frac{\arccos(\langle\vec{v}_i,\vec{v}_j\rangle)}{\pi}$

Now, $\arcsin(z) + \arccos(z) = \frac{\pi}{2}$ (since $\cos(\frac{\pi}{2} - \theta) = \sin\theta$ so $\frac{\pi}{2} - \arccos(z) = \arcsin(z)$)

so $\Pr[X_i X_j = 1] = 1 - \frac{1}{\pi}\left[\frac{\pi}{2} - \arcsin(\langle \vec{v}_i, \vec{v}_j \rangle)\right]$

$= \frac{1}{2} + \frac{\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle)}{\pi}$ and $\Pr[X_i X_j = -1] = \frac{1}{\pi}\left[\frac{\pi}{2} - \arcsin(\langle \vec{v}_i, \vec{v}_j \rangle)\right]$

giving $\mathbb{E}[X_i X_j] = \Pr[X_i X_j = 1] - \Pr[X_i X_j = -1] = \frac{2\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle)}{\pi}$ ▨

To bound the approximation, we use the following inequalities for the

Schur product: for matrices M and N, $M \circ N = [M_{ij} \cdot N_{ij}]$ (entrywise)

Schur product Theorem If $M \succeq 0$ and $N \succeq 0$, $M \circ N \succeq 0$.

Corollary: If $M \succeq 0$ and $N \succeq 0$, $\sum_{i,j} M_{ij} N_{ij} \geq 0$.

Proof: $\sum_{ij} M_{ij} N_{ij} = T^T M \circ N T \geq 0$ by the Schur product Theorem. ▨

Lemma If $Z \succeq 0$, $|z_{ij}| \leq 1$ for all i,j. then $[\arcsin(z_{ij}) - z_{ij}] \succeq 0$

Proof: The Taylor expansion for $\arcsin(z) = z + \frac{1}{2 \cdot 3} z^3 + \frac{3}{2 \cdot 4 \cdot 5} z^5 + \cdots \frac{3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdots 2n(2n+1)} z^{2n+1}$

so $[\arcsin(z_{ij}) - z_{ij}] = \frac{1}{2 \cdot 3} Z \circ Z \circ Z + \frac{3}{2 \cdot 4 \cdot 5} Z \circ Z \circ Z \circ Z \circ Z + \cdots$ where $Z \circ \cdots \circ Z \succeq 0$ and any positive linear combination of psd matrices is also psd. ▨

Theorem Random hyperplane rounding of the semidefinite program relaxation of $\underset{\vec{x} \in \{-1,1\}^n}{\text{maximize}} \vec{x}^T A \vec{x}$ for $A \succeq 0$ gives $\mathbb{E}[\vec{x}^T A \vec{x}] \geq \frac{2}{\pi} \max_{\vec{x} \in \{-1,1\}^n} \vec{x}^T A \vec{x}$.

Proof: $\mathbb{E}[\vec{x}^T A \vec{x}] = \sum_{i,j} A_{ij} \frac{2\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle)}{\pi}$ where the semidefinite program objective is $\sum_{i,j} A_{ij} \langle \vec{v}_i, \vec{v}_j \rangle \geq \max_{\vec{x} \in \{-1,1\}^n} \vec{x}^T A \vec{x}$. Now, notice, $\mathbb{E}[\vec{x}^T A \vec{x}] - \frac{2}{\pi} \sum_{i,j} A_{ij} \langle \vec{v}_i, \vec{v}_j \rangle = \frac{2}{\pi} \sum_{i,j} A_{ij} (\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle) - \langle \vec{v}_i, \vec{v}_j \rangle)$ where since $VV^T \succeq 0$, our lemma gives $[\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle) - \langle \vec{v}_i, \vec{v}_j \rangle] \succeq 0$, and since $A \succeq 0$ also, our corollary to the Schur product Theorem gives $\sum_{i,j} A_{ij} [\arcsin(\langle \vec{v}_i, \vec{v}_j \rangle) - \langle \vec{v}_i, \vec{v}_j \rangle) \geq 0$. Thus we find $\mathbb{E}[\vec{x}^T A \vec{x}] \geq \frac{2}{\pi} \sum_{i,j} A_{ij} \langle \vec{v}_i, \vec{v}_j \rangle \geq \frac{2}{\pi} \max_{\vec{x} \in \{-1,1\}^n} \vec{x}^T A \vec{x}$. ▨

Note: proof of Schur product Theorem:

We can write $M = U\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}U^T$, $N = V\begin{bmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{bmatrix}V^T$ where, since $M, N \succeq 0$, $\lambda_1, \dots, \lambda_n$ and $\mu_1, \dots, \mu_n \geq 0$. Equiv, since $\langle \vec{u}_i, \vec{u}_j \rangle = 0$ for $i \neq j$, $M = \sum_{i=1}^{n} \lambda_i \vec{u}^{(i)}\vec{u}^{(i)T}$ and $N = \sum_{i=1}^{n} \mu_i \vec{v}^{(i)}\vec{v}^{(i)T}$

Thus $(M \circ N)_{ij} = \left(\sum_{k=1}^{n} \lambda_k u_i^{(k)} u_j^{(k)}\right)\left(\sum_{l=1}^{n} \mu_l v_i^{(l)} v_j^{(l)}\right) =$

$$\sum_{k=1}^{n}\sum_{l=1}^{n}\left(\sqrt{\lambda_k}u_i^{(k)}\sqrt{\mu_l}v_i^{(l)}\right)\left(\sqrt{\lambda_k}u_j^{(k)}\sqrt{\mu_l}v_j^{(l)}\right)$$

So: $M \circ N = \sum_{k=1}^{n}\sum_{l=1}^{n}\left(\sqrt{\lambda_k}\vec{u}^{(k)} \circ \sqrt{\mu_l}\vec{v}^{(l)}\right)\left(\sqrt{\lambda_k}\vec{u}^{(k)} \circ \sqrt{\mu_l}\vec{v}^{(l)}\right)^T$

which is a sum of psd matrices, hence $M \circ N \succeq 0$ also. ☒