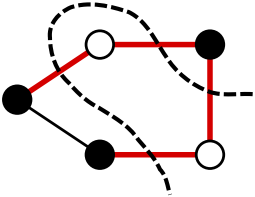
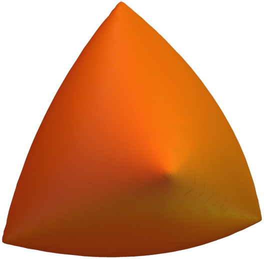
Max-Cut and Goemans-Williamson

Marko Lalovic

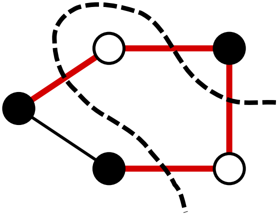




Outline

Max-Cut

- Goal: Given $G = (V, E)$ with $V := \{1, \ldots, n\}$ and $|E| = m$
- Find a subset $S \subseteq V$, such that $f(S) := |cut(S)|$ is maximum



Max-Cut problem is NP-complete [Karp72] [1]. How well can we approximate Max-Cut?

[1] Karp(1972) Reducibility among Combinatorial Problems

Approximation Algorithms

- Denote the optimal value of the Max-Cut problem by $mc(G)$
- And the size of the cut returned by some algorithm by $alg(G)$

Definition

Algorithm $f(S) = alg(G)$ is an α-*approximation* of Max-Cut if

$$f(S) \geq \alpha \cdot mc(G) \tag{1}$$

for all graphs $G = (V, E)$ and some approximation ratio $\alpha \in [0, 1]$.

If algorithm employed is randomized, we say the same, if Inequality (1) holds with an expectation taken on the left-hand side.

Approximation Algorithms

Example

Randomized $\frac{1}{2}-$approximation algorithm for Max-Cut, that assigns each vertex of V to S and $V \setminus S$ independently uniformly at random

$$\mathbb{E}[f(S)] = \sum_{(i,j) \in E} \mathbb{P}[(i,j) \in \text{cut}(S)] = \frac{1}{2} \cdot m \geq \frac{1}{2} \cdot mc(w) \tag{2}$$

[Erd67] [a]

[a] Erdős(1967) On bipartite subgraphs of a graph

- Can we do better?
- Yes: In preliminary version in '94 of [GW95] [2] they improved this by proposing α_{GW}-approximation algorithm with $\alpha_{GW} \geq 0.87$ using semidefinite programming.

[2] Goemans and Williamson(1995) Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming

Preliminaries

A real symmetric matrix X is positive definite, denoted $X \succeq 0$ if the following equivalent conditions hold:

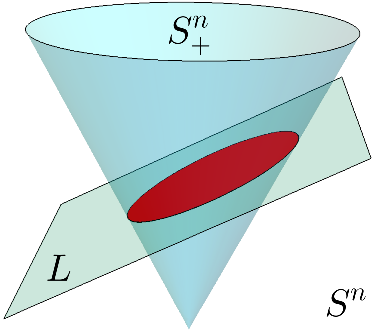
1. All eigenvalues of X are non-negative
2. Quadratic form $v^T X v \geq 0$ for all $v \in \mathbb{R}^n$
3. There exists $Q \in \mathbb{R}^{n \times r}$ with

$$X = QQ^T = \sum_{i=1}^{r} v_i v_i^T$$

   where $v_1, \ldots, v_r$ are columns of Q

Preliminaries

- Let S_+^n denote the convex cone of positive semidefinite matrices in the set of all symmetric matrices S^n
- A *spectrahedron* is the intersection of S_+^n with an affine linear space L



Semidefinite Programming

- *Semidefinite programming (SDP)* solves the following problem: maximize or minimize a linear objective function over the spectrahedron:

$$
\begin{aligned}
&\text{maximize } C \bullet X \\
&\text{subject to:} \\
&\qquad A_i \bullet X = b_i \quad i = 1, \ldots, m \\
&\qquad X \succeq 0
\end{aligned}
\tag{P}
$$

$$
\begin{aligned}
&\text{minimize } b^T y \\
&\text{subject to:} \\
&\qquad \sum_{i=1}^{m} A_i y_i - C \succeq 0
\end{aligned}
\tag{D}
$$

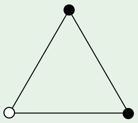
- *Weak duality*: $C \bullet X \geq b^T y$ always holds
- Under primal and dual feasibility, also *strong duality*: $C \bullet X = b^T y$ holds

Reformulation of Max-Cut

- The optimal value of the Max-Cut problem can be expressed by

$$mc(G) = \max_{x \in \{-1,1\}^n} \sum_{i,j} \frac{1 - x_i x_j}{2} \qquad \text{(QP)}$$

Example (K_3)



If: $(x_1, x_2, x_3)^T = (1, -1, -1)^T$

Then:

$$\begin{aligned} mc(G) &= \frac{1 - x_1 x_2}{2} + \frac{1 - x_1 x_3}{2} + \frac{1 - x_2 x_3}{2} \\ &= \frac{1 - 1(-1)}{2} + \frac{1 - 1(-1)}{2} + \frac{1 - (-1)(-1)}{2} \\ &= 1 + 1 + 0 \\ &= 2 \end{aligned}$$

SDP Relaxation Cont.

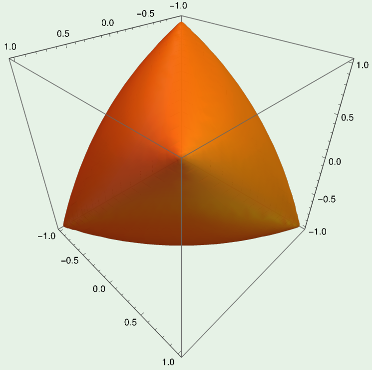
- Cuts in the complete graph K_n can be represented by

$$\{xx^T : x \in \{-1, 1\}^n\}$$

- The *elliptope* $\mathcal{E}_n$ is a set of all $n \times n$ correlation matrices

$$\mathcal{E}_n = \{X \in S_n : X_{ii} = 1 \quad \text{for all } i\} \tag{3}$$

Example (K_3)



SDP Relaxation

- Therefore

$$
\begin{aligned}
mc(G) &= \max_{X \in \mathcal{E}_n, rk(X)=1} \sum_{i,j} \frac{1 - X_{ij}}{2} \\
&\leq \max_{X \in \mathcal{E}_n} \sum_{i,j} \frac{1 - X_{ij}}{2} = sdp(G)
\end{aligned} \tag{4}
$$

- This means that $sdp(G)$ is a *relaxation* of Max-Cut problem.
- Note: Objective function is linear in entries of matrix X

Hyperplane Rounding

Every positive semidefinite matrix X can be decomposed as $X = QQ^T$ and the quantity $sdp(G)$ can be reformulated as

$$sdp(G) = \max_{\|v_i\|_2=1} \sum_{i,j} \frac{1 - v_i^T v_j}{2} \tag{5}$$

Select a random unit vector $r \in \mathbb{R}^n$ and construct the subset

$$S := \{i \in V \mid v_i^T r \geq 0\}$$

This is called *hyperplane rounding*

Hyperplane Rounding

We can show that

$$\mathbb{E}[f(S)] = \sum_{i,j} \mathbb{P}[(i,j) \in cut(S)] \geq \alpha_{GW} \cdot sdp(G) \tag{6}$$

where $\alpha_{GW} = \min_{\theta_{i,j} \in [0,\pi]} \{ \frac{2}{\pi} \frac{\theta_{i,j}}{1-\cos(\theta_{i,j})} \} \geq 0.878$. Combining Inequalities 4 and 6

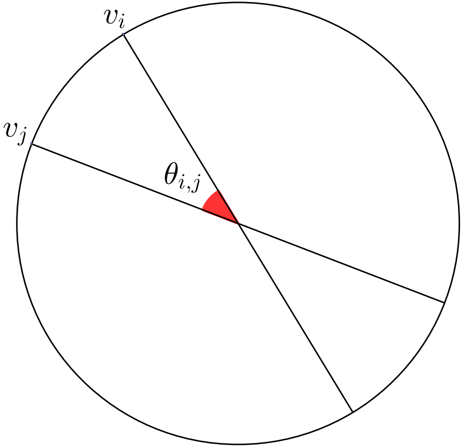
$$\mathbb{E}[f(S)] \geq \alpha_{GW} \cdot sdp(G) \geq \alpha_{GW} \cdot mc(G) \tag{7}$$

and finally conclude

$$\mathbb{E}[f(S)] \geq \alpha_{GW} \cdot mc(G) \tag{8}$$

Sketch of Proof of (7)

$$\mathbb{P}[(i,j) \in cut(S)] = \frac{\theta_{i,j}}{\pi} \tag{9}$$



Sketch of Proof of (7) Cont.

$$
\begin{aligned}
\mathbb{P}[(i,j) \in cut(S)] &= \frac{\theta_{i,j}}{\pi} \\
&= \frac{\arccos(v_i^T v_j)}{\pi} \\
&= \frac{2}{\pi} \frac{\arccos(v_i^T v_j)}{1 - v_i^T v_j} \frac{1 - v_i^T v_j}{2} \\
&\geq \min_{\theta_{i,j} \in [0,\pi]} \left\{ \frac{2}{\pi} \frac{\theta_{i,j}}{1 - \cos(\theta_{i,j})} \right\} \cdot \frac{1 - v_i^T v_j}{2} \\
&= \alpha_{GW} \cdot \frac{1 - v_i^T v_j}{2} \\
&= \alpha_{GW} \cdot sdp(G)
\end{aligned}
$$

Dual Problem

- Warm up: find maximum eigenvalue of a symmetric matrix $X \in S_+^n$
- Suppose X has eigenvalues

$$\lambda_1 \geq \cdots \geq \lambda_n \tag{10}$$

- Then for some $t \in \mathbb{R}$

$$t - \lambda_1 \leq \cdots \leq t - \lambda_n \tag{11}$$

- Note: $tI - X \succeq 0$ if and only if $0 \leq t - \lambda_1$ or equivalently $t \geq \lambda_{\max}(X)$
- This immediately gives us an SDP

$$\lambda_{\max}(X) = \min_t \{tI - X \succeq 0\} \tag{12}$$

Dual Problem

Define the *Laplacian matrix* $L = (L_{i,j})_{i,j}$ of a graph $G = (V, E)$ as

$$L_{i,j} := \begin{cases} deg(v_i) & \text{if } i = j \\ -1 & \text{if } (i,j) \in E(G) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

Dual of Max-Cut SDP relaxed problem can be reformulated as

$$\begin{aligned} &\text{minimize } \frac{n}{4}t \\ &\text{subject to:} \\ &\qquad tI - (L + diag(u)) \succeq 0 \\ &\qquad 1^T u = 0 \end{aligned} \tag{D'}$$

Dual Problem

Dual of Max-Cut SDP relaxed problem can be reformulated as:

$$\frac{n}{4} \min_{u:1^T u=0} \lambda_{\max}(L + diag(u)) \tag{14}$$

By weak duality we get an upper bound given in [MP90] [3]

$$mc(G) \leq sdp(G) \leq \frac{n}{4}\lambda_{\max}(L) \tag{15}$$

For $G = C_5$, $\lambda_{\max} = \frac{1}{5}(5 + \sqrt{5})$, we get the upper bound studied in [DP93] [4]

$$\frac{1}{2}(5 + \sqrt{5})/4 \geq 0.9 \cdot mc(C_n) \tag{16}$$

[3] Mohar and Poljak(1990) Eigenvalues and the max-cut problem

[4] Delorme and Poljak(1993) Laplacian eigenvalues and the maximum cut problem

SoS Relaxation

Find an operator $\tilde{\mathbb{E}}$ that behaves like the expectation over some probability distribution on $x \in \{0,1\}^n$

$$\tilde{\mathbb{E}} : \mathcal{P}_n^{\leq d} \to \mathbb{R} \tag{17}$$

where $P_n^{\leq d}$ represents a set of polynomials $p : \{0,1\}^n \to \mathbb{R}$ of degree at most d in n variables $x_1, \dots x_n$, $x_i \in \{0,1\}$, to get an optimization problem for Max-Cut:

$$\max_{\tilde{\mathbb{E}}} \tilde{\mathbb{E}}[\sum_{(i,j)\in E(G)} (x_i - x_j)^2]$$

$$\text{subject to:} \tag{D''}$$

$$(1)\ \tilde{\mathbb{E}} \text{ is linear} \qquad (2)\ \tilde{\mathbb{E}}[1] = 1$$

$$(3)\ \tilde{\mathbb{E}}[p^2] \geq 0 \qquad (4)\ \tilde{\mathbb{E}}[x_i^2 p] = \tilde{\mathbb{E}}[x_i p]$$

$$\text{for all polynomials } p \text{ with } deg(p) \leq \frac{d}{2}$$

SoS Relaxation

- This is a relaxation of Max-Cut problem
- Given $G = (V, E)$ and a subset S with size of the cut $mc(G)$, there is a feasible solution to (D") with objective value equal to $mc(G)$
  - Denote the indicator vector of S as $a_1, \ldots, a_n$ and let $\tilde{\mathbb{E}}$ be

$$\tilde{\mathbb{E}}[p(x_1, \ldots, x_n)] = p(a_1, \ldots, a_n) \tag{18}$$

  - Then $\tilde{\mathbb{E}}$ satisfies the constraints (1)-(4) and achieves objective value

$$\tilde{\mathbb{E}}[\sum_{(i,j)\in E(G)} (x_i - x_j)^2] = mc(G) \tag{19}$$

- This approach is called *Sum-of-Squares (SoS)* hierarchy and was introduced in [Pa00] [5] and [La01] [6]
- Increasing the degree d, we increase the size of SDP problem. For $d = n$, we get exact relaxation

[5] Parrilo(2000) Structured semidefinite programs and semi-algebraic geometry methods in robustness and optimization

[6] Lasserre(2001) Global optimization with polynomials and the problem of moments

Gaussian Rounding

Given $\tilde{\mathbb{E}}$ that realizes the maximum in (D"), i.e. SDP solution of SoS with objective value

$$sos(G) := \max_{\tilde{\mathbb{E}}} \tilde{\mathbb{E}}[\sum_{(i,j)\in E(G)} (x_i - x_j)^2] \tag{20}$$

From SoS being a relaxation of Max-Cut it also follows that

$$sos(G) \geq mc(G) \qquad \text{for all } G = (V, E) \tag{21}$$

Gaussian Rounding

- Assume $\tilde{\mathbb{E}}[x_i] = \frac{1}{2}$ for all $i = 1, \ldots n$
- Take y to be a Gaussian vector with the following mean and covariance matrix

$$\mu = \tilde{\mathbb{E}}[x] = \frac{1}{2}1 \qquad \Sigma = \tilde{\mathbb{E}}[(x-\mu)(x-\mu)^T] \tag{22}$$

- Construct the indicator vector a for a subset $S \subseteq V$ as

$$a_i = \begin{cases} 1, & \text{if } y_i \leq \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

Gaussian Rounding

We can show that

$$\mathbb{E}[f(S)] = \sum_{i,j} \mathbb{P}[(i,j) \in cut(S)] \geq \alpha_{GW} \cdot sos(G) \tag{24}$$

where $\alpha_{GW} = \min_{\theta \in [0,\pi]} \{ \frac{2}{\pi} \frac{\theta}{1-\cos(\theta)} \} \geq 0.878$ as before. Combining Inequalities 21 and 24

$$\mathbb{E}[f(S)] \geq \alpha_{GW} \cdot sos(G) \geq \alpha_{GW} \cdot mc(G)$$

we can also in this case conclude that

$$\mathbb{E}[f(S)] \geq \alpha_{GW} \cdot mc(G) \tag{25}$$

Sketch of Proof of (25)

- For each edge $(i,j) \in E(G)$, define $\rho_{i,j} = 4\tilde{\mathbb{E}}[x_i x_j] - 1$
- Given two uncorrelated random variables $(s,t) \overset{\text{i.i.d.}}{\sim} N(0,I)$, we can get $\rho_{i,j}$-correlated random variables s and u, where

$$u = \rho_{i,j} s + \sqrt{1-\rho_{i,j}^2}\, t \tag{26}$$

- Then $\rho_{i,j} = \mathbb{E}[us] = \cos(\theta_{i,j})$, and we can calculate

$$\begin{aligned}
\mathbb{P}[(i,j) \in cut(S)] &= \mathbb{P}[a_i \neq a_j] \\
&= \mathbb{P}[sgn(y_i - \frac{1}{2}) \neq sgn(y_j - \frac{1}{2})] \\
&= \mathbb{P}[sgn(s) \neq sgn(u)] \\
&= \frac{\theta_{i,j}}{\pi}
\end{aligned}$$

Sketch of Proof of (25) Cont.

$$\begin{aligned}
\mathbb{E}[f(S)] &= \sum_{i,j} \mathbb{P}[(i,j) \in cut(S)] \\
&= \sum_{i,j} \frac{\theta_{i,j}}{\pi} \\
&= \sum_{i,j} \frac{\theta_{i,j}}{\pi} \cdot \frac{2}{(1-\rho_{i,j})} \tilde{\mathbb{E}}[(x_i - x_j)^2] \\
&= \sum_{i,j} \frac{2}{\pi} \frac{\theta_{i,j}}{1-\cos(\theta_{i,j})} \tilde{\mathbb{E}}[(x_i - x_j)^2] \\
&\geq \min_{\theta \in [0,\pi]} \{ \frac{2}{\pi} \frac{\theta}{1-\cos(\theta)} \} \cdot \sum_{i,j} \tilde{\mathbb{E}}[(x_i - x_j)^2] \\
&= \alpha_{GW} \cdot \sum_{i,j} \tilde{\mathbb{E}}[(x_i - x_j)^2] \\
&= \alpha_{GW} \cdot \tilde{\mathbb{E}}[\sum_{i,j} (x_i - x_j)^2] \\
&= \alpha_{GW} \cdot sos(G)
\end{aligned}$$