Properties of a New Semidefinite Relaxation for the Max-Cut Problem

Miguel F. Anjos
(joint work with Prof. Henry Wolkowicz)

Department of Combinatorics & Optimization
University of Waterloo
Waterloo, Canada

Workshop on Novel Approaches to Hard Discrete Optimization

University of Waterloo, April 2001

The MAX-CUT problem (MC)

Given a graph $G = (V, E)$ and weights w_{ij} for all edges $(i, j) \in E$, find an edge-cut of maximum weight, i.e. find a set $S \subseteq V$ s.t. the sum of the weights of the edges with one end in S and the other in $V \setminus S$ is maximum.

We assume wlog that $w_{ii} = 0$ for all $i \in V$,

and that G is complete (assign $w_{ij} = 0$ if edge $ij \notin E$).

We present some theoretical and computational results for a new semidefinite programming (SDP) relaxation of MC.

Formulation of MC in $\Re^n$

Let the vector $v \in \{\pm 1\}^n$, $n = |V|$, represent a cut in the graph G via $S = \{i : v_i = +1\}$ and $V \backslash S = \{i : v_i = -1\}$.

Then we can formulate MC as:

$$\begin{array}{rll} \mu^* = & \max & \sum_{1 \le i < j \le n} w_{ij} \left(\frac{1 - v_i v_j}{2} \right) \\ & \text{s.t.} & v \in \{\pm 1\}^n. \end{array}$$

Equivalently,

$$(\text{MC1}) \qquad \begin{array}{rlll} \mu^* = & \max & v^T Q\, v & \\ & \text{s.t.} & v_i^2 = 1, & i = 1, \ldots, n, \end{array}$$

where $Q = \frac{1}{4}(\text{Diag}\,(A\, e) - A)$, and $A = (w_{ij})$ is the weighted adjacency matrix of G.

First lifting

Taking the Lagrangian dual MC1, and then the dual of the dual, yields SDP1, the well-known SDP relaxation for Max-Cut:

$$
\text{(SDP1)} \qquad \begin{array}{rll} \nu_1^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \succeq 0 \\ & & X \in \mathcal{S}^n , \end{array}
$$

where $\mathcal{S}^n$ denotes the space of $n \times n$ symmetric matrices.

Formulation of MC in $\mathcal{S}^n$

Consider the change of variable $X = vv^T, v \in \{\pm 1\}^n$. Then $X_{ij} = v_i v_j$ and:

- X is symmetric and rank-one;
- $X \succeq 0 : y^T X y = y^T v v^T y = (v^T y)^2 \geq 0$;
- and if $v_i^2 = 1$, then $\operatorname{diag}(X) = e$ (all ones);
- obj. function:
  $v^T Q v = \operatorname{trace}(v^T Q v) = \operatorname{trace}(Q v v^T) = \operatorname{trace} QX$.

In fact, MC is equivalent to

$$
(\mathrm{MC2}) \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & \operatorname{rank}(X) = 1 \\ & & X \succeq 0,\ X \in \mathcal{S}^n\,. \end{array}
$$

Question: How can we handle the rank constraint?

First we apply the following result:

Lemma 1 *For $X \in \mathcal{S}^n$*

$$X \succeq 0, X \circ X = E \text{ (the matrix of all ones)}$$

if and only if

$$X = vv^T \text{ for some } v \in \{\pm 1\}^n.$$

Hence Max-Cut can be formulated as

$$\text{(MC3)} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \circ X = E \\ & & X \succeq 0,\ X \in \mathcal{S}^n\,. \end{array}$$

Furthermore,

$$X^2 = nX \quad \Rightarrow \quad X \succeq 0$$

(because X and X^2 are simultaneously diagonalizable)

therefore another formulation of MC is:

$$\begin{array}{ll} \max & \operatorname{trace} QX \\ \text{s.t.} & \operatorname{diag}(X) = e \\ & X^2 - nX = 0 \\ & X \circ X = E. \end{array}$$

But let's look more closely at the $X^2 - nX = 0$ constraint...

Consider again $X = vv^T, v \in \{\pm 1\}^n$ and recall that all the entries $X_{ij} = v_i v_j$ equal ± 1.

Since the equation $X^2 - nX = 0$ is componentwise equivalent to

$$X_{ij} = \frac{1}{n} \sum_{k=1}^{n} X_{ik}\, X_{kj}$$

we see that each term $X_{ik}\, X_{kj}$ must in fact equal X_{ij}.

Second lifting

So for the second lifting, our starting formulation of MC is

$$
\begin{array}{ll}
\max & \operatorname{trace} QX \\
\text{s.t.} & \operatorname{diag}(X) = e \\
 & X_{ij} = X_{ik} X_{kj}, \quad k = 1, \ldots, n, \quad 1 \le i < j \le n \\
 & X \circ X = E.
\end{array}
$$

Taking the Lagrangian dual of this problem,

and then the dual of the dual, yields SDP2.

Define (by example) the linear operators

$$\operatorname{dsvec}\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_4 \\ \alpha_2 & \alpha_3 & \alpha_5 \\ \alpha_4 & \alpha_5 & \alpha_6 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ \alpha_3 \\ 2\alpha_4 \\ 2\alpha_5 \\ \alpha_6 \end{pmatrix} \qquad \text{and}$$

$$\operatorname{sMat}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_6 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_4 \\ \alpha_2 & \alpha_3 & \alpha_5 \\ \alpha_4 & \alpha_5 & \alpha_6 \end{pmatrix}.$$

These are defined so that, for the appropriate inner products,

$$< X, \operatorname{sMat}(y) > = < \operatorname{dsvec}(X), y > \text{ for all } X \in \mathcal{S}^n, y.$$

Then SDP2 may be written as:

$$
\begin{array}{rll}
\nu_2^* = & \max & \operatorname{trace} H_Q Y \qquad\qquad \left[\text{Partition } \ Y = \begin{pmatrix} 1 & x^T \\ x & \bar{Y} \end{pmatrix}\right] \\
& \text{s.t.} & \operatorname{diag}(Y) = e \\
& & Y_{t(i),0} = 1, \quad i = 1, \ldots, n \\
& & Y_{T(i,j),0} = Y_{T(i,k),T(k,j)}, \ k = 1, \ldots, n, \ 1 \leq i < j \leq n \\
& & Y \succeq 0, Y \in \mathcal{S}^{t(n)+1},
\end{array}
$$

where $H_Q = \begin{pmatrix} 0 & \frac{1}{2}\operatorname{dsvec}(Q)^T \\ \frac{1}{2}\operatorname{dsvec}(Q) & 0 \end{pmatrix}$,

$$
T(i,j) := \begin{cases} t(j-1) + i, \text{ if } i \leq j \\ t(i-1) + j, \ \text{otherwise,} \end{cases} \quad \text{and } t(i) = \frac{i(i+1)}{2} = \binom{i+1}{2}.
$$

Projection back to $\mathcal{S}^n$

The linear operator sMat maps the feasible set of SDP2 into $\mathcal{S}^n$ by defining

$$F_n := \{X \in \mathcal{S}^n \, : X = \operatorname{sMat}(Y_{1:t(n),0}), Y \text{ feasible for SDP2}\}.$$

The set F_n is convex and compact.

What can we say about F_n?

Convex Relaxations of MC in $\mathcal{S}^n$

The smallest convex set containing all the rank-one matrices X corresponding to cuts is the *cut polytope*:

$$C_n := \mathrm{Conv}\{X : X = vv^T, v \in \{\pm 1\}^n\}.$$

Define the *elliptope* $\mathcal{E}_n$ to be the feasible set of SDP1:

$$\mathcal{E}_n := \{X \in \mathcal{S}^n \; : \mathrm{diag}\,(X) = e, X \succeq 0\}.$$

It is straightforward to show that

$$C_n \subseteq F_n \subseteq \mathcal{E}_n.$$

Another convex set containing all the rank-one matrices is the *metric polytope* M_n, defined by

$$\begin{aligned} M_n := \{X \in \mathcal{S}^n \mid \operatorname{diag}(X) = e, \text{ and } X_{ij} + X_{ik} + X_{jk} \geq -1, \\ X_{ij} - X_{ik} - X_{jk} \geq -1, -X_{ij} + X_{ik} - X_{jk} \geq -1, \\ - X_{ij} - X_{ik} + X_{jk} \geq -1, \forall\, 1 \leq i < j < k \leq n\}. \end{aligned}$$

These inequalities model the fact that for any three mutually connected vertices in the graph, it is only possible to cut either zero or two of the edges joining them.

It is known that $C_n = M_n$ for $n \leq 4$, but $C_n \subsetneq M_n$ for $n \geq 5$.

Lemma 2

$$F_n \subseteq M_n.$$

And furthermore,

Theorem 1 *(A. & Wolkowicz)*

$$C_n \subsetneq F_n \subsetneq \mathcal{E}_n \cap M_n \text{ for } n = 5, \text{ and hence for } n \geq 5.$$

Further Strengthening Results for SDP2

It is easy to check that

X is a vertex of C_n if and only if $X = v\,v^T$ for some $v \in \{\pm 1\}^n$.

It is known that the elliptope reflects this part of the structure of C_n quite well:

Theorem 2 *(Laurent & Poljak) The elliptope $\mathcal{E}_n$ has precisely 2^{n-1} vertices, each of the form $v\,v^T$ for some $v \in \{\pm 1\}^n$.*

The same holds for SDP2:

Theorem 3 *(A. & Wolkowicz)*

$$\operatorname{rank} Y = 1 \quad \Leftrightarrow \quad \operatorname{rank} X = 1 \quad \Leftrightarrow \quad X \text{ is a vertex of } C_n.$$

Theorem 4 *For the case $n = 3$,*

- $\operatorname{rank} Y = 1$ *iff X is a vertex of C_3;*
- $\operatorname{rank} Y = 2$ *iff X lies in the relative interior of a face of C_3 of dimension 1;*
- $\operatorname{rank} Y = 3$ *iff X lies in the relative interior of a facet (face of dimension 2) of C_3;*
- $\operatorname{rank} Y = 4$ *iff X lies in the relative interior of C_3.*

A useful result is:

Lemma 3 *If* $X \in F_n$ *and* $\operatorname{rank} X = 2$, *then for every triple* (i, j, k) *of distinct indices from* $\{1, \ldots, n\}$, *at least one of* X_{ij}, X_{ik}, X_{jk} *has absolute value equal to 1.*

Theorem 5 *(A. & Wolkowicz)*

For $n \geq 3$,

$$\operatorname{rank} Y = 2 \quad \Leftrightarrow \quad \operatorname{rank} X = 2$$

$\Leftrightarrow$ X *is in the relative interior of a 1-dimensional face of* C_n.

And, given $X \in F_n$ such that $\operatorname{rank} X = 2$, a simple algorithm yields the two cuts whose convex combination equals X.

Selected problems (toroidal graphs with ±1 weights)

	Number of nodes	SDP1 $\frac{\mu^*}{\nu_1^*}$	SDP2 $\frac{\mu^*}{\nu_2^*}$	SDP1 found (rank-1) optimal cut	SDP2 found (rank-1) optimal cut
2-D	16	0.9859	1.0000	No	Yes
	16	0.7379	1.0000	No	No
	20	0.9432	1.0000	No	Yes
	20	0.8863	1.0000	No	No
	20	0.8151	0.9852	No	No
	28	0.8704	1.0000	No	No
3-D	18	0.9470	1.0000	No	Yes
	18	0.8711	1.0000	No	No
	18	0.9075	1.0000	No	No
	18	0.9642	1.0000	No	Yes
	18	0.8428	1.0000	No	Yes

Selected problems (toroidal graphs with Gaussian weights)

	Number of nodes	SDP1 $\frac{\mu^*}{\nu_1^*}$	SDP2 $\frac{\mu^*}{\nu_2^*}$	SDP1 found (rank-1) optimal cut	SDP2 found (rank-1) optimal cut
2-D	16	0.9985	1.0000	No	Yes
	20	0.9775	1.0000	No	Yes
	20	0.9666	1.0000	No	Yes
	20	0.9157	1.0000	No	Yes
	24	0.9288	1.0000	No	Yes
3-D	16	0.9999	1.0000	No	Yes
	16	0.9192	1.0000	No	Yes
	18	0.9720	1.0000	No	Yes
	20	0.9942	1.0000	No	Yes
	27	0.8731	0.9999	No	Yes
	27	0.9220	1.0000	No	Yes

Empirical observations

1. SDP2 found the weight of the optimal cut
   - *always* for the problems with Gaussian distributed weights.
   - almost always for the problems with $\pm$ 1 weights.
2. Furthermore, SDP2 seems to often find an optimal rank-one matrix and hence an optimal cut:
   - it found an optimal cut for *all* problems with Gaussian distributed weights,
   - and also for several of the problems with $\pm$ 1 weights.