Modeling with semidefinite and copositive matrices

Franz Rendl

http://www.math.uni-klu.ac.at

Alpen-Adria-Universität Klagenfurt

Austria

Overview

- Node and Edge relaxations for Max-Cut
- Stable-Set Problem and Theta Function
- Graph Coloring and dual Theta Function
- Theta function for sparse and dense problems
- Copositive relaxations

Max-Cut (1)

Unconstrained quadratic 1/-1 optimization:

$$\max \ x^T L x \text{ such that } x \in \{-1, 1\}^n$$

Linearize (and simplify) to get tractable relaxation
$x^T L x = \langle L, xx^T \rangle$, New variable is X.
Basic SDP relaxation:

$$\max \langle L, X \rangle : \text{diag}(X) = e, \ X \succeq 0$$

See Poljak, Rendl (1995) primal-dual formulation, Goemans, Williams (1995) worst-case error analysis (at most 14 % above optimum if weights nonnegative)

Max-Cut (2)

This SDP relaxation can be further tightened by including Combinatorial Cutting Planes: A simple observation:
Barahona, Mahjoub (1986): Cut Polytope,
Deza, Laurent (1997): Hypermetric Inequalities

$$x \in \{-1, 1\}^n, \quad f = (1, 1, 1, 0, \dots, 0)^T \quad \Rightarrow |f^T x| \geq 1.$$

Results in $x^T f \; f^T x = \langle (xx^T), (ff^T) \rangle = \langle \mathbf{X}, \mathbf{ff^T} \rangle \geq \mathbf{1}$.
Can be applied to any **triangle** $i < j < k$.
Nonzeros of f can also be -1.
There are $4\binom{n}{3}$ such triangle inequality constraints.

Direct application of standard methods not possible for $n \approx 100$.

SDP edge-model for Max-Cut (1)

The previous SDP (implicitely) assumes that the graph is dense. If the number m of edges is small, say $O(n)$, then the following model provides a stronger relaxation, see Dissertation Wiegele, Klagenfurt, 2006.

Using $x \in \{1, -1\}^n$ we form an edge vector $y = (y_{ij})$ indexed by 0 and $[ij] \in E(G)$ as follows

$$y_0 = 1, \; y_{ij} = x_i x_j \text{ for } [ij] \in E(G).$$

Forming $Y = yy^T$ we get the SDP edge-relaxation for Max-Cut by putting the cost coefficients in the row and column corresponding to y_0, yielding L_E.

Note that Y is now of order $m + 1$ instead of n before.

SDP edge-model for Max-Cut (2)

Constraints on Y:

$$\text{diag}(Y) = e, \;\; Y \succeq 0 \quad \text{like in node model}$$

If i, j, k is a triangle in G:

$$y_{ij,ik} = y_{0,jk} \quad \text{because } y_{ij,ik} = (x_i x_j)(x_i x_k) = x_j x_k = y_{0,jk}$$

If i, j, k, l is 4-cycle in G:

$$y_{ij,kl} = x_i x_j x_k x_l = x_i x_l x_j x_k = y_{il,jk}$$

$$y_{ij,jk} = y_{il,lk}$$

Similar to second lifting of Anjos, Wolkowicz (2002), and Lasserre (2002) in case of complete graphs.

SDP edge-model for Max-Cut (3)

The second lifting of Anjos, Wolkowicz and Lasserre goes from matrices of order $\mathbf{n}$ to matrices of order $\binom{\mathbf{n}}{\mathbf{2}} + \mathbf{1}$, independent of the number m of edges.

It is computationally intractable once $n \approx 100$.

The present model can handle graphs with up to 2000 edges (number of vertices is irrelevant). Computational results in the forthcoming dissertation of Wiegele (Klagenfurt, 2006).

This model assumes that the graph contains a star. If not, add edges of weight 0 from node 1 to all other nodes.

The resulting SDP is too expensive for standard methods, once number of triangle and 4-cycles gets large.

Overview

- Node and Edge relaxations for Max-Cut
- Stable-Set Problem and Theta Function
- Graph Coloring and dual Theta Function
- Theta function for sparse and dense problems
- Copositive relaxations

Stable sets and theta function

$G = (V, E) \ldots$ Graph on n vertices.
$x_i = 1$ if i in some stable set, otherwise $x_i = 0$.

$$\max \sum_i x_i \text{ such that } x_i x_j = 0 \; ij \in E, \; x_i \in \{0, 1\}$$

Linearization trick: Consider $X = \frac{1}{x^T x} x x^T$.
X satisfies:

$$X \succeq 0, \quad \text{tr}(X) = 1, \; x_{ij} = 0 \forall ij \in E, \; rank(X) = 1$$

Note also: $e^T x = x^T x$, so $e^T x = \langle J, X \rangle$. Here $J = ee^T$.

Lovasz (1979): relax the (diffcult) rank constraint

Stable sets and theta function (2)

$$\vartheta(G) := \max\{\langle J, X\rangle : X \succeq 0,\ \text{tr}(X) = 1,\ \ x_{ij} = 0\ (ij) \in E\}$$

This SDP has $m + 1$ equations, if $|E| = m$.

Can be solved by interior point methods if $n \approx 500$ and $m \approx 5000$.

Notation: We write $A_G(X) = 0$ for $x_{ij} = 0,\ (ij) \in E(G)$.
Hence $A_G(X)_{ij} = \langle E_{ij}, X\rangle$ with $E_{ij} = e_i e_j^T + e_j e_i^T$.
Any symmetric matrix M can therefore be written as

$$M = \text{Diag}(m) + A_G(u) + A_{\bar{G}}(v).$$

Overview

- Node and Edge relaxations for Max-Cut
- Stable-Set Problem and Theta Function
- Graph Coloring and dual Theta Function
- Theta function for sparse and dense problems
- Copositive relaxations

Coloring and dual theta function

We now consider Graph Coloring and recall the Theta function:

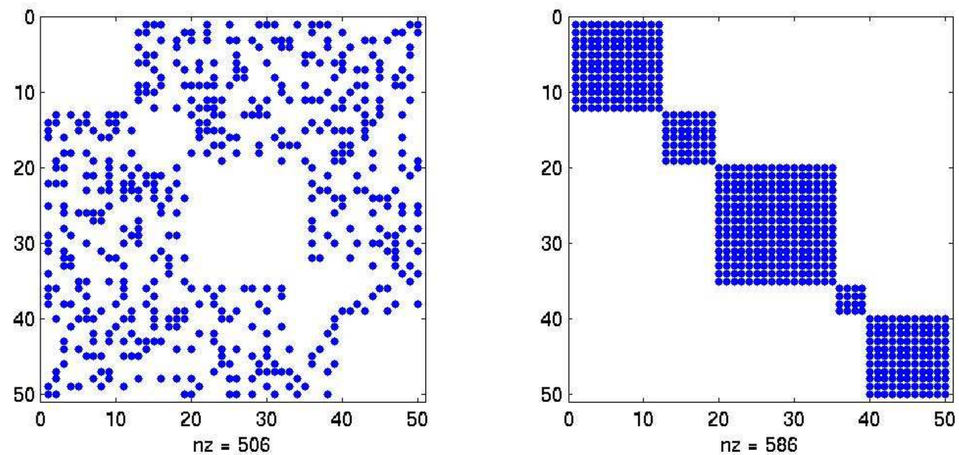
$$\vartheta(G) := \{\max\langle J, X\rangle : X \succeq 0,\ \text{tr}(X) = 1,\ \ A_G(X) = 0\}$$

$$= \min t \text{ such that } tI + A_G^T(y) \succeq J.$$

Here $A_G^T(y) = \sum_{ij} y_{ij} E_{ij}$. Coloring viewpoint: Consider complement graph $\bar{G}$ and partition V into stable sets $s_1, \ldots, s_r$ in $\bar{G}$, where $\chi(\bar{G}) = r$.

Let $M = \sum_i^r s_i s_i^T$ where s_i is characteristic vector of stable set in $\bar{G}$. M is called coloring matrix.

Coloring Matrices



Adjacency matrix A of a graph (left), associated Coloring Matrix (right). The graph can be colored with 5 colors.

Coloring Matrices (2)

Note: A 0-1 matrix M is coloring matrix if and only if

$m_{ij} = 0 \; (ij) \in E$, diag$(M) = e$, $(tM - J \succeq 0 \;\Leftrightarrow\; t \geq rank(M))$

Hence

$$\chi(\bar{G}) = \min t \text{ such that}$$

$$tM - J \succeq 0, \;\; \text{diag}(M) = e, \; m_{ij} = 0 \; \forall (ij) \in \bar{E}, \;\; m_{ij} \in \{0, 1\}$$

Setting $Y = tM$ we get $Y = tI + \sum_{ij \in E} y_{ij} E_{ij} = tI + A_G(y)$.
Leaving out $m_{ij} \in \{0, 1\}$ gives dual of theta function.

$$\vartheta(G) = \min t : \text{ such that } tI + A_G(y) - J \succeq 0.$$

This gives Lovasz sandwich theorem: $\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G})$.

Overview

- Node and Edge relaxations for Max-Cut
- Stable-Set Problem and Theta Function
- Graph Coloring and dual Theta Function
- Theta function for sparse and dense problems
- Copositive relaxations

Sparse and dense Graphs

Since $m \leq \binom{n}{2}$, we say that G is sparse if $m < \frac{1}{2}\binom{n}{2}$ and call it dense otherwise.

$$\vartheta(G) := \max\langle J, X\rangle \text{ such that } X \succeq 0,\ \text{tr}(X) = 1,\ \ A_G(X) = 0$$

$$= \min t \text{ such that } tI + A_G^T(y) - J \succeq 0.$$

There are $m + 1$ equations in the primal, so this can be handled by interior-point methods if m is not too large. For dense graphs, we can use the following reformulation. Let $Y = tI + A_G^T(y)$ and set $Z = Y - J$.

Sparse and dense Graphs (2)

$Z = Y - J \succeq 0$ has the following properties:
$A_{\bar{G}}(Z) = -2e$, because $z_{ij} = -1$ for $[ij] \notin E$.
$te -$ diag$(Z) = e$, because diag$(Y) = te$. Hence we get the theta function equivalently as

$$\vartheta(G) = \min\{t : te - \textbf{diag}(Z) = e,\ -A_{\bar{G}} = 2e, Z \succeq 0\} =$$

$$\max\{e^T x + 2e^T \xi : \textbf{Diag}(x) + A_{\bar{G}}(\xi) \succeq 0, e^T x = 1\}.$$

Here the dual has $\bar{m} + n$ equations, hence this is good for dense graphs ($\bar{m}$ small in this case).

See Dukanovic and Rendl, working paper, Klagenfurt 2005.

Comparing the two models

The two models have the following running times on graphs with $n = 100$ and various edge densities.

density	0.90	0.75	0.50	0.25	0.10
m	4455	3713	2475	1238	495
dense	1	7	42	130	238
sparse	223	118	34	5	1

Comparison of the computation times (in seconds) for ϑ on five random graphs with 100 vertices and different densities in the dense and the sparse model.

The computation takes no more than half a minute.

Overview

- Node and Edge relaxations for Max-Cut
- Stable-Set Problem and Theta Function
- Graph Coloring and dual Theta Function
- Theta function for sparse and dense problems
- Copositive relaxations

Copositive Relaxation of Stable-Set

DeKlerk, Pasechnik (SIOPT 2002) consider the following copositive relaxation of Stable-Set and show:

$$\alpha(G) := \max\langle J, X\rangle \text{ such that}$$

$$X \in C^*, \ \text{tr}(X) = 1, \ \ A_G(X) = 0.$$

The proof uses
(a) extreme rays are of form xx^T with $x \geq 0$
(b) support of x = some stable set
(c) maximization makes x nonzeros of x equal to one another.

Could also be shown using the Motzkin-Strauss Theorem.

A copositive approximation of Coloring

We have seen that copositive relaxation gives exact value of stable set. Since coloring matrices M are in C^*, we consider

$$t^* := \min t \text{ such that}$$

$$tI + A_{\bar{G}}^T(y) \succeq J, \quad tI + A_{\bar{G}}^T(y) \in C^*$$

We clearly have

$$\vartheta \leq t^* \leq \chi$$

Unlike in the stable set case, where the copositive model gave the exact problem, we will show now the following.
Theorem (I. Dukanovic, F. Rendl 2005): $t^* \leq \chi_f \leq \chi$

χ_f denotes the fractional chromatic number.

Fractional Chromatic Number

$\chi_f(G)$ is defined as follows. Let $s_1, \ldots$ be the characteristic vectors of (all) stable sets in G.

$$\chi_f(G) := \min \sum_i \lambda_i \text{ such that } \sum_i \lambda_i s_i = e, \lambda_i \geq 0.$$

(χ is obtained by asking $\lambda_i = 0$ or 1.)

Lemma Let x_i be 0-1 vectors and $\lambda_i \geq 0$. Let $X_\lambda := \sum_i x_i x_i^T$.

Then $M := (\sum_j \lambda_j) X_\lambda - \text{diag}(X_\lambda) \text{diag}(X_\lambda)^T \succeq 0$.

Proof of Lemma

$M := (\sum_j \lambda_j) X_\lambda - \text{diag}(X_\lambda)\text{diag}(X_\lambda)^T$.

We have

(a) $\text{diag}(x_i x_i^T) = x_i$

(b) $\text{diag}(X_\lambda) = \sum_i \lambda_i x_i$

(c) $M = (\sum_j \lambda_j)(\sum_i \lambda_i x_i x_i^T) - \sum_{ij} \lambda_i \lambda_j x_i x_j^T$

(d) Let y be arbitrary and set $a_i := x_i^T y$.

(e) $y^T M y = \sum_{ij} \lambda_i \lambda_j a_i^2 - \sum_{ij} \lambda_i \lambda_j a_i a_j = \sum_{i<j} \lambda_i \lambda_j (a_i^2 + a_j^2 - 2a_i a_j) \geq 0$.

Proof of Theorem

Let λ_i be feasible for $\chi_f(G)$, hence $\lambda_i \geq 0, \ \sum_i \lambda_i s_i = e$.
Let $X_\lambda := \sum_i \lambda_i s_i s_i^T \in C^*$.
Then diag$(X_\lambda) = \sum_i \lambda_i s_i = e$.
Set $t = \sum_i \lambda_i$.
The Lemma shows that $tX_\lambda \succeq J$ and so we have feasible solution (with same value t).

We do not know whether $t^* = \chi_f$ holds in general, but it is true for vertex-transitive graphs.