Copositive versus Semidefinite Relaxations for some combinatorial optimization problems

Franz Rendl

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

Alpen-Adria-Universität Klagenfurt

Austria

Overview

- Theta function of Lovasz and variations
- Quadratic Assignment Problem
- Graph Coloring
- The Parrilo-DeKlerk-Pasechnik hierarchy
- Theta function revisited

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

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

Copositive Relaxation

DeKlerk, Pasechnik consider strengthening towards $\alpha(G)$ by asking that X is completely positive:

$$X \in C^* := \{\sum_i y_i y_i^k : y_i \geq 0\}$$

The cone of completely positive matrices is dual to the cone of copositive matrices C:

$$C := \{M : x^T M x \geq 0 \;\; \forall x \geq 0\}.$$

It is however co-NP-complete to test whether a matrix is co-positive.

Copositive relaxation (2)

Theorem (DeKlerk, Pasechnik (2003))

$$\alpha(G) := \max\langle J, X\rangle : X \in C^*, \text{ tr}(X) = 1, \;\; x_{ij} = 0 \; (ij) \in E$$

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 nonzeros of x equal to one another.

We will show similar results for QAP and copositive approximations to Coloring.

Overview

- Theta function of Lovasz and variations
- Quadratic Assignment Problem
- Graph Coloring
- The Parrilo-DeKlerk-Pasechnik hierarchy
- Theta function revisited

Quadratic Assignment Problem (QAP)

(QAP) $\min\langle AXB + C, X\rangle$ such that X is permutation matrix

Using $x =$ vec(X), $x \circ x = x$ we get

$$\langle AXB + C, X\rangle = \langle B \otimes A + \text{Diag}(vec(C)), xx^T\rangle$$

Now linearize $Y = xx^T$ to get SDP or copositive relaxations.

A technical problem:
How translate permutation properties from x to Y?

$$X = (x_1, \ldots, x_n), Y = \begin{pmatrix} Y^{11} & \ldots & Y^{1n} \\ \vdots & & \vdots \\ Y^{n1} & \ldots & Y^{nn} \end{pmatrix}, Y^{ij} = x_i x_j^T$$

QAP (2)

$$\sum_i Y^{ii} = \sum_i x_i x_i^T = I, \quad \text{tr}(Y^{ij}) = x_i^T x_j = \delta_{ij}$$

$$\langle J, Y \rangle = (e^T x)(x^T e) = n^2$$

X is orthogonal, sums of all elements =n.

$$\mathcal{F} := \{Y \in C^*, \sum_i Y^{ii} = I, \text{ tr}(Y^{ij}) = \delta_{ij}, \langle J, Y \rangle = n^2\}$$

Theorem (J. Povh, F. Rendl 2005)

$\mathcal{F} = conv\{xx^T : x = vec(X), X \text{ permutation matrix}\} = \Pi$

Proof

$\Pi \subseteq \mathcal{F}$, because each $X \in \Pi$ is by construction feasible for $\mathcal{F}$.

Now let $Y \in \mathcal{F}$, hence $Y = \sum_k y_k y_k^T = \sum_k Z_k$ and $Z_k = y_k y_k^T, \; y_k \geq 0$.

Let Y_k be $n \times n$ matrix formed from $y_k \in \mathbb{R}^{n^2}$.
We need to show that each $Y_k \in \Pi$.

The proof is based on the following facts:
(a) each main diagonal block Z_k^{ii} is diagonal
(b) each off diagonal block has diag$(Z_k^{ij}) = 0 \;\; \forall i \neq j$
(c) each Y_k has at most one nonzero in each row / column.
(d) Each Y_k is multiple of permutation matrix.

The last step makes use of the Cauchy-Schwary inequality.

Copositive relaxation of QAP

$$L := B \otimes A + \text{Diag}(vec(C)).$$

As a consequence, QAP is equivalent to the copositive program

$$\min\langle L, Y\rangle : \sum_i Y^{ii} = I,\ \text{tr}(Y^{ij}) = \delta_{ij},\ \langle J, Y\rangle = n^2,\ Y \in C^*.$$

Replacing $Y \in C^*$ by $Y \succeq 0$ gives SDP relaxation investigated by Zhao, Karisch, Wolkowicz, Sotirov, Rendl.

Further constraints could be added, like

$$Y_{ij,ik} = 0, \quad Y \geq 0.$$

SDP relaxations of QAP

The SDP relaxation

$$\min\langle L,Y\rangle : \sum_i Y^{ii} = I,\ \mathrm{tr}(Y^{ij}) = \delta_{ij},\ \langle J,Y\rangle = n^2,$$

$$Y \succeq 0,\ Y_{ij,ik} = Y_{ij,kj} = 0,\ Y \geq 0$$

provides currently the strongest bounds, which are manageable for interesting sizes.

Recently Burer and Vandenbussche (2004) investigated the Lovasz-Schriver lifting of QAP. It can be shown (see dissertation Povh, Klagenfurt (2006)), that this is actually equivalent to the above SDP model.
The computational effort to solve this SDP for $n = 30$ is still a serious computational challenge.

Overview

- Theta function of Lovasz and variations
- Quadratic Assignment Problem
- Graph Coloring
- The Parrilo-DeKlerk-Pasechnik hierarchy
- Theta function revisited

Coloring and dual theta function

We now consider Graph Coloring and recall Theta function:

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

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

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

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: $\alpha(G) \leq \vartheta(G) \leq \chi(\bar{G})$.

A copositive approximation of Coloring

Since coloring matrices M are in C^*, we consider

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

$$tI + A_G^T(y) \succeq J$$

$$tI + A_G^T(y) \in C^*$$

We clearly have

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

Unlike in the stable set and QAP case, where the copositive model gave the exact problem, we will show now the following.

Theorem (I. Dukanovic, M. Laurent, F. Rendl 2006): $t^* = \chi_f$

Fractional Chromatic Number

$\chi_f(G)$ is defined as follows. Let $S + (S_1, \ldots)$ be the of all characteristic vectors of stable sets S in $\bar{G}$.

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

(χ is obtained by asking $\lambda_i = 0$ or 1.)
Let us collect all characteristic vectors S_i in the matrix A.
(A has n rows, but may have an exponential number of columns).

$$\chi_f := \min e^t \lambda : \; A\lambda = e, \; \lambda \geq 0$$

$$= \max e^t y : \; \sum_{j \in S_i} y_j := y_{S_i} \leq 1, \; \forall S_i \in S, \; y \in \mathbb{R}^n.$$

Auxiliary Lemma

We first show that any feasible solution λ for the primal LP gives a feasible solution t, Y for the copositive program having the same value.
An auxiliary result:
Lemma

Let x_i be 0-1 vectors and $\lambda_i \geq 0$. Let $X_\lambda := \sum_i \lambda_i x_i x_i^T$. 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) We need to show that $y^T M y \geq 0 \;\; \forall y.$

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

(e) $y^T M y = \sum_{ij} \lambda_i \lambda_j a_i^2 - \sum_{ij} \lambda_i \lambda_j a_i a_j =$

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

Claim 1

Claim: $t^* \leq \chi_f$

Proof: Take feasible solution λ of primal LP and define
$Y := \sum_i \lambda_i S_i S_i^T$.
Set $t := \sum \lambda_i$. We need to show:
diag$(Y) = e, \ Y \in C^*, \ tY - J \succeq 0.$
We have diag$(Y) = \sum_i \lambda_i S_i = e$.

$Y \in C^*$ by construction

The Lemma shows that $tY \succeq J$ and so we have feasible solution (with same value t).

Since we minimize, $t^* \leq \chi_f$.

Claim 2

If y feasible for dual LP and t feasible for copositive program, than $e^t y \leq t$.

Consequence: $\chi_f \leq t^*$.

Has recently been pointed out to us by Monique Laurent.

Overview

- Theta function of Lovasz and variations
- Quadratic Assignment Problem
- Graph Coloring
- The Parrilo-DeKlerk-Pasechnik hierarchy
- Theta function revisited

Approximating the copositive cone

The cone dual to C^* is the cone C of copositive matrices:

$$C := \{M : x^T M x \geq 0 \; \forall x \geq 0\}$$

Parrilo (2000) and DeKlerk, Pasechnik (2002) use following idea to approximate C:

$$M \in C \text{ iff } P(x) := (x \circ x)^T M (x \circ x) = \sum_{ij} x_i^2 x_j^2 m_{ij} \geq 0 \;\; \forall x.$$

A sufficient condition for this to hold is that

$P(x)$ has a sum of squares representation.

Parrilo hierarchy

Parrilo proposes to consider

$$P_r(x) := (\sum_i x_i^2)^r P(x)$$

Polya (1928): If M strictly copositive then $P_r(x)$ is SOS for some sufficiently large r.

Parrilo (2000):
$P_0(x)$ is SOS iff $M = P + N$, where $P \succeq 0$ and $N \geq 0$.
$P_1(x)$ is SOS iff $\exists M_1, \ldots, M_n$ such that

$$M - M_i \succeq 0$$

$$(M_i)_{ii} = 0 \; \forall i \quad (M_i)_{jj} + 2(M_j)_{ij} = 0 \; \forall i \neq j$$

$$(M_i)_{jk} + (M_j)_{ik} + (M_k)_{ij} \geq 0 \; \forall i < j < k$$

Overview

- Theta function of Lovasz and variations
- Quadratic Assignment Problem
- Graph Coloring
- The Parrilo-DeKlerk-Pasechnik hierarchy
- Theta function revisited

$r = 1$ case for vertex transitive graphs

The relaxation with $r = 1$ involves n semidefiniteness constraints for matrices of order n and $O(n^3)$ additional linear (in)equalities.
This is computationally intractable even for small problems ($n \approx 100$).

Dukanovic (2005) shows that for vertex transitive graphs, the $r = 1$ approximation of Parrilo applied to stable set relaxation can be done with one additional matrix and one SDP constraint.

Some computational results

We consider Hamming graphs and compare ϑ which is equal to the $r = 0$ relaxation in all these cases with $r = 1$.

graph	n	$\vartheta, (r = 0)$	$r = 1$	χ
H(7,6)	128	53.33	63.9	64
H(8,6)	256	85.33	127.9	128
H(9,4)	512	51.19	53.9	
H(10,8)	1024	383.99	511.9	512
H(12,4)	4096	211.86	255.5	

Further computational simplifications are used by exploiting the automorphism group underlying Hamming graphs. All computations took no more than a few minutes on a PC.

Further details in the forthcoming dissertation of Dukanovic (Klagenfurt 2006).