Matrix relaxations for optimization problems on graphs

Franz Rendl

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

Alpen-Adria-Universität Klagenfurt

Austria

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Abstract combinatorial optimization

$E \ldots$ finite ground set (e.g. edges of graph)
$F \in \mathcal{F}$ feasible solutions $F \subseteq E$ (e.g. spanning trees)
$c_e : \; e \in E$ cost elements, $c(F) := \sum_{e \in F} c_e$
Combinatorial optimization problem (COP)

$$(COP) \quad \min\{c(F) : \; F \in \mathcal{F}\}$$

$x_F \in \{0,1\}^n$ characteristic vector of F
$\mathcal{P} := conv\{x_F : \; F \in \mathcal{F}\}$ convex hull of feasible solutions.

Abstract combinatorial optimization

$E \ldots$ finite ground set (e.g. edges of graph)
$F \in \mathcal{F}$ feasible solutions $F \subseteq E$ (e.g. spanning trees)
$c_e: \; e \in E$ cost elements, $c(F) := \sum_{e \in F} c_e$
Combinatorial optimization problem (COP)

$$(COP) \quad \min\{c(F) : \; F \in \mathcal{F}\}$$

$x_F \in \{0,1\}^n$ characteristic vector of F
$\mathcal{P} := conv\{x_F : \; F \in \mathcal{F}\}$ convex hull of feasible solutions.

Classical Polyhedral approach to (COP): use (partial) description of $\mathcal{P}$ in combination with linear optimization

$$z_{cop} = \min\{c(F) : \; F \in \mathcal{F}\} = \min\{c^T x : x \in \mathcal{P}\}$$

First problem is min over finite set, last problem is LP.

Polyhedral approach for (COP)

Classical example: Assignment Problem:

$$z_{AP} = \min\{\sum_i c_{i,\phi(i)} : \ \phi \in \Pi\}$$

Π set of permutations. $X_\phi \ldots$ permutation matrix

Polyhedral approach for (COP)

Classical example: Assignment Problem:

$$z_{AP} = \min\{\sum_i c_{i,\phi(i)} : \ \phi \in \Pi\}$$

Π set of permutations. $X_\phi \dots$ permutation matrix

Theorem (Birkhoff): $conv\{X_\phi : \phi \in \Pi\} = \Omega$
$\Omega = \{X : Xe = X'e = e, \ X \geq 0\}$ doubly stochastic matrices

$$z_{AP} = \min\{\langle C, X\rangle : \ X \in \Omega\}$$

In general $\mathcal{P}$ is not easily available.

Why nonpolyhedral relaxations?

Some graph optimization problems have natural formulations using quadratic functions. Example Max-Cut:

Given a graph G with edge weights $c_{ij}, ij \in E(G)$, find vertex bisection $(X, V(G) \setminus X)$ of maximum edge weight.

Why nonpolyhedral relaxations?

Some graph optimization problems have natural formulations using quadratic functions. Example Max-Cut:

Given a graph G with edge weights $c_{ij}, ij \in E(G)$, find vertex bisection $(X, V(G) \setminus X)$ of maximum edge weight.

Linear model: introduce edge variables $y_{ij} \in \{0, 1\}$ for $ij \in E$. Then $\sum_{ij \in E} c_{ij} y_{ij}$ is linear in y but how describe edge vectors y corresponding to edge sets of cuts ?

Why nonpolyhedral relaxations?

Some graph optimization problems have natural formulations using quadratic functions. Example Max-Cut:

Given a graph G with edge weights $c_{ij}, ij \in E(G)$, find vertex bisection $(X, V(G) \setminus X)$ of maximum edge weight.

Linear model: introduce edge variables $y_{ij} \in \{0, 1\}$ for $ij \in E$. Then $\sum_{ij \in E} c_{ij} y_{ij}$ is linear in y but how describe edge vectors y corresponding to edge sets of cuts ?

Quadratic model: introduce node variables $x_i \in \{-1, 1\}$ for $i \in V$. Then cut edges ij are characterized by $x_i x_j = -1$, hence we get unconstrained quadratic optimization in $-1, 1$ variables.

Quadratic structure of problem leads to matrix relaxations.

Matrix relaxations of (COP)

Consider $\mathcal{M} := conv(x_F x_F^T : F \in \mathcal{F})$.
Generalizes polyhedral approach as $diag(x_F x_F^T) = x_F$.

- quadratic constraints in x_F are linear for $X \in \mathcal{M}$.
- $\mathcal{M}$ contained also in nonpolyhedral matrix cones.

$PSD = \{X : a^T X a \geq 0 \ \forall a\}$ semidefinite matrices
$C = \{X : a^T X a \geq 0 \ \forall a \geq 0\}$ copositive matrices

Matrix relaxations of (COP)

Consider $\mathcal{M} := conv(x_F x_F^T : F \in \mathcal{F})$.
Generalizes polyhedral approach as $diag(x_F x_F^T) = x_F$.

- quadratic constraints in x_F are linear for $X \in \mathcal{M}$.
- $\mathcal{M}$ contained also in nonpolyhedral matrix cones.

$PSD = \{X : a^T X a \geq 0 \ \forall a\}$ semidefinite matrices
$C = \{X : a^T X a \geq 0 \ \forall a \geq 0\}$ copositive matrices

dual cone of PSD is again PSD
dual cone of C:
$C^* = conv\{aa^T : a \geq 0\}$ completely positive matrices

Bad news: $X \notin C$ NP-hard to decide.

Semidefinite and Copositive Programs

Problems of the form

$$\max\langle C, X\rangle \text{ s.t. } A(X) = b,\ \ X \in PSD$$

are called Semidefinite Programs.

Semidefinite and Copositive Programs

Problems of the form

$$\max\langle C, X\rangle \text{ s.t. } A(X) = b,\ X \in PSD$$

are called Semidefinite Programs.

Problems of the form

$$\max\langle C, X\rangle \text{ s.t. } A(X) = b,\ X \in C$$

or

$$\max\langle C, X\rangle \text{ s.t. } A(X) = b,\ X \in C^*$$

are called Copositive Programs, because the primal or the dual involves copositive matrices.

Summary: LP versus SDP

Linear Optimization :

- Simplex method efficient for polyhedral approach,
- natural dimension: $\mathbb{R}^n$

Summary: LP versus SDP

Linear Optimization :

- Simplex method efficient for polyhedral approach,
- natural dimension: $\mathbb{R}^n$

Semidefinite Optimization :

$$\min\langle C, X\rangle \text{ subject to } A(X) = b, \quad X \succeq 0.$$

- Can be solved in polynomial time (for given precision).
- Strong duality: holds under some technical condition.
- natural dimension: $n \times n$ symmetric matrices.

Summary: LP versus SDP

Linear Optimization :

- Simplex method efficient for polyhedral approach,
- natural dimension: $\mathbb{R}^n$

Semidefinite Optimization :

$$\min \langle C, X \rangle \text{ subject to } A(X) = b, \quad X \succeq 0.$$

- Can be solved in polynomial time (for given precision).
- Strong duality: holds under some technical condition.
- natural dimension: $n \times n$ symmetric matrices.

SDP should be considered (only) if

- original formulation has something quadratic
- other (easier) approaches fail

Semidefinite Relaxations (SDP)

- First results by Lovasz (1979) for coloring, max-clique
- Goemans-Williamson rounding for Max-Cut (1994)
- Interior-Point methods generalized to SDP

- Approximation results for:

- Max-k-Cut (Frieze, Jerrum, DeKlerk, Pasechnik, etc),
- Coloring (Karger, Motwani, Sudan, Arora, Chlamtac),
- Max-2-Sat (Goemans, Williamson 1995)
- Max-Sat (Asano, Williamson 2002)
- Bandwidth minimization (Blum et al 2000)
- Vertex Cover (Goemans, Kleinberg, Halperin 1998,2002)
- and many others

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Text book example: Max Cut

Max-Cut as a binary quadratic problem.

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

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

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

This model goes back to A. Schrijver.
See also Poljak, R. (1995) primal-dual formulation, and Goemans, Williamson (1995) for the hyperplane rounding analysis.

A fundamental SDP Problem

The SDP relaxation of Max-Cut

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

appears in many other matrix relaxations of graph optimization problems.

For instance:

- max-k-cut
- coloring
- ordering problems

Basic SDP Relaxation of Max-Cut

We solve $\max\langle L, X\rangle : \quad diag(X) = e, \quad X \succeq 0$.
Matrices of order n, and n simple equations $x_{ii} = 1$

n	seconds
1000	12
2000	102
3000	340
4000	782
5000	1570

Seconds on a PC. Implementation of primal-dual interior-point method in MATLAB, 30 lines of source code

Max-k-Cut

Max-k-Cut asks to partition the vertices of a graph G into k pieces $S_1, \dots, S_k$, so that the total weight between the pieces (=partition blocks) is maximized.
Partitions modeled by $n \times k$ matrices $S = (s_1, \dots, s_k)$. s_i is characteristic vector of S_i.
Use Laplacian L of G to express edges cut by S as

$$z_{cut} = \frac{1}{2}\langle S, LS\rangle = \langle L, SS^T\rangle.$$

Max-k-Cut

Max-k-Cut asks to partition the vertices of a graph G into k pieces $S_1, \ldots, S_k$, so that the total weight between the pieces (=partition blocks) is maximized.
Partitions modeled by $n \times k$ matrices $S = (s_1, \ldots, s_k)$. s_i is characteristic vector of S_i.
Use Laplacian L of G to express edges cut by S as

$$z_{cut} = \frac{1}{2}\langle S, LS\rangle = \langle L, SS^T\rangle.$$

Matrix lifting: $Y = SS^T$. Then $z_{cut} = \frac{1}{2}\langle L, Y\rangle$.
We also have $diag(Y) = e$, because $Se = e$ and $diag(s_i s_i^T) = s_i$
Asking that $Y \succeq 0$ can be strengthened

Partition Lemma

Let $S = (s_1, \ldots, s_k)$ be 0-1 matrix and $\lambda_i \geq 0$ be such that $S\lambda = \sum_i \lambda_i s_i = e$. Let $t = \sum_i \lambda_i > 0$. Let $Y = \sum_i \lambda_i s_i s_i^T$.

Then $diag(Y) = e$ and $tY - J \succeq 0$.

Partition Lemma

Let $S = (s_1, \ldots, s_k)$ be 0-1 matrix and $\lambda_i \geq 0$ be such that $S\lambda = \sum_i \lambda_i s_i = e$. Let $t = \sum_i \lambda_i > 0$. Let $Y = \sum_i \lambda_i s_i s_i^T$.

Then $diag(Y) = e$ and $tY - J \succeq 0$.

Partition Lemma

Let $S = (s_1, \ldots, s_k)$ be 0-1 matrix and $\lambda_i \geq 0$ be such that $S\lambda = \sum_i \lambda_i s_i = e$. Let $t = \sum_i \lambda_i > 0$. Let $Y = \sum_i \lambda_i s_i s_i^T$.

Then $diag(Y) = e$ and $tY - J \succeq 0$.

Proof:

$$\sum_i \lambda_i \binom{1}{s_i}\binom{1}{s_i}^T = \begin{pmatrix} t & e^T \\ e & Y \end{pmatrix} \succeq 0.$$

Partition Lemma

Let $S = (s_1, \ldots, s_k)$ be 0-1 matrix and $\lambda_i \geq 0$ be such that $S\lambda = \sum_i \lambda_i s_i = e$. Let $t = \sum_i \lambda_i > 0$. Let $Y = \sum_i \lambda_i s_i s_i^T$.

Then $diag(Y) = e$ and $tY - J \succeq 0$.

Proof:

$$\sum_i \lambda_i \begin{pmatrix} 1 \\ s_i \end{pmatrix} \begin{pmatrix} 1 \\ s_i \end{pmatrix}^T = \begin{pmatrix} t & e^T \\ e & Y \end{pmatrix} \succeq 0.$$

Max-k-Cut has $\lambda_i = 1$ and we get stronger condition $kY - J \succeq 0$ (instead of $Y \succeq 0$) and SDP relaxation

$$\max\{\frac{1}{2}\langle L, Y\rangle : diag(Y) = e, \; kY - J \succeq 0, \; Y \geq 0\}$$

Summary Max-Cut and Max-k-Cut

- Convexification of Max-Cut in the node space used by Billionet and Elloumi (2006) in combination with convex QP and Branch and Bound

Summary Max-Cut and Max-k-Cut

- Convexification of Max-Cut in the node space used by Billionet and Elloumi (2006) in combination with convex QP and Branch and Bound
- SDP relaxation with triangle inequalities using interior-point methods in combination with bundle method and Branch and Bound: problems with $n \approx 150$ routinely manageable, see BIQMAC-website (R., Rinaldi, Wiegele (2006)).

Summary Max-Cut and Max-k-Cut

- Convexification of Max-Cut in the node space used by Billionet and Elloumi (2006) in combination with convex QP and Branch and Bound
- SDP relaxation with triangle inequalities using interior-point methods in combination with bundle method and Branch and Bound: problems with $n \approx 150$ routinely manageable, see BIQMAC-website (R., Rinaldi, Wiegele (2006)).
- SDP relaxation for Max-k-Cut used by Frieze, Jerrum (1997) to get approximation results
- Max-k-Cut relaxation used by Ghaddar, Anjos, Liers (2008) in a Branch-and-Cut approach.

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Graph Coloring

$\{s_1, \ldots, s_k\}$ stable sets (=pairwise non adjacent) with characteristic vectors x_i in G. A Coloring is partition of vertices into stable sets. The chromatic number $\chi(G)$ is the smallest k such that G has k-partition into stable sets.

$$\chi = \min\{\sum \lambda_i : \sum \lambda_i x_i = e,\ \lambda_i \in \{0, 1\}\}$$

Graph Coloring

$\{s_1, \ldots, s_k\}$ stable sets (=pairwise non adjacent) with characteristic vectors x_i in G. A Coloring is partition of vertices into stable sets. The chromatic number $\chi(G)$ is the smallest k such that G has k-partition into stable sets.

$$\chi = \min\{\sum \lambda_i : \sum \lambda_i x_i = e, \ \lambda_i \in \{0,1\}\}$$

The fractional chromatic number χ_f is obtained by replacing $\lambda_i \in \{0,1\}$ by $\lambda_i \geq 0$.
This is LP with exponential number of variables λ_i, and NP-hard.

Graph Coloring

$\{s_1, \ldots, s_k\}$ stable sets (=pairwise non adjacent) with characteristic vectors x_i in G. A Coloring is partition of vertices into stable sets. The chromatic number $\chi(G)$ is the smallest k such that G has k-partition into stable sets.

$$\chi = \min\{\sum \lambda_i : \sum \lambda_i x_i = e, \ \lambda_i \in \{0,1\}\}$$

The fractional chromatic number χ_f is obtained by replacing $\lambda_i \in \{0,1\}$ by $\lambda_i \geq 0$.
This is LP with exponential number of variables λ_i, and NP-hard.

Matrix lifting: $Y = \sum \lambda_i x_i x_i^T$. Partition lemma shows:
$diag(Y) = e, \ \chi_f Y - J \succeq 0$.
partition into stable sets implies $y_{uv} = 0$ if $uv \in E(G)$.

SDP relaxation of Coloring

$$\min\{t:\ diag(Y)=e,\ tY-J\succeq 0,\ y_{uv}=0\ \forall uv\in E(G)\}\leq \chi_f$$

Optimum is called Lovasz theta function (Lovasz 1979). Its dual SDP:

$$\vartheta=\max\{\langle J,X\rangle:\ tr(X)=1,\ x_{uv}=0\ uv\notin E, X\succeq 0\}$$

SDP relaxation of Coloring

$$\min\{t : \ diag(Y) = e, \ tY - J \succeq 0, \ y_{uv} = 0 \ \forall uv \in E(G)\} \leq \chi_f$$

Optimum is called Lovasz theta function (Lovasz 1979). Its dual SDP:

$$\vartheta = \max\{\langle J, X\rangle : \ tr(X) = 1, \ x_{uv} = 0 \ uv \notin E, X \succeq 0\}$$

Can be interpreted as SDP relaxation of Max-Clique ($\omega(G) \ldots$ size of largest clique in G).

Lovasz sandwich theorem: $\omega(G) \leq \vartheta(G) \leq \chi_f(G)$

SDP relaxation of Coloring

$$\min\{t : \ diag(Y) = e, \ tY - J \succeq 0, \ y_{uv} = 0 \ \forall uv \in E(G)\} \leq \chi_f$$

Optimum is called Lovasz theta function (Lovasz 1979). Its dual SDP:

$$\vartheta = \max\{\langle J, X\rangle : \ tr(X) = 1, \ x_{uv} = 0 \ uv \notin E, X \succeq 0\}$$

Can be interpreted as SDP relaxation of Max-Clique ($\omega(G) \ldots$ size of largest clique in G).

Lovasz sandwich theorem: $\omega(G) \leq \vartheta(G) \leq \chi_f(G)$

Many theoretical implications: Grötschel, Lovasz, Schrijver use it to compute $\chi(G)$ and $\alpha(G)$ for perfect graphs G in polynomial time.

Computing the theta function

The number of constraints depends on the edge set $|E|$. If m is small, then the SDP can be solved efficiently using interior point methods.

n	$m = \frac{1}{2}\binom{n}{2}$	time (secs.)	n	$m = 5n$	time (secs.)
100	2475	21	500	2500	28
150	5587	180	1000	5000	200
200	9950	925	1500	7500	618

Times in seconds for computing $\vartheta(G)$ on random graphs with different densities($|E| = \frac{1}{4}n^2,\ 5n$).
In each iteration, a linear equation with $|E|$ variables has to be solved, so no hope if |E| > 10,000.

Theta function: big DIMACS graphs

Boundary point method (Malick, Povh, R., Wiegele (2006))

graph	n	m	ϑ	ω
keller5	776	74.710	31.00	27
keller6	3361	1026.582	63.00	$\geq$59
san1000	1000	249.000	15.00	15
brock800-1	800	112.095	42.22	23
p-hat500-1	500	93.181	13.07	9
p-hat1000-3	1000	127.754	84.80	$\geq$68
p-hat1500-3	1500	227.006	115.44	$\geq$94

The theta number for the bigger instances has not been computed before.

Summary Coloring

- SDP relaxation (ϑ function) currently the basis for the best approximation results for coloring: Arora, Chlamtac (2007) color a three-colorable graph with at most $O(n^{0.211})$ colors.

- No combinatorial algorithm to find coloring (or clique) number for perfect graphs.

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Ordering Problems

Data: Objects $1, \ldots, n$ and numbers c_{ij} for $i \neq j$.
For a given ordering of the objects, say

$$i_1,\ i_2, \ldots, i_n$$

we associate the weight: $c_{i_1,i_2} + c_{i_1,i_3} + \ldots + c_{i_{n-1},i_n}$.

Problem: find order with maximum weight.

Or more general: costs $c_{ij,kl}$ to be gained if i is before j and k before l. Leads to Ordering problem with quadratic objective.

General and powerful modeling tool.

Linear Ordering Problems

Linear Ordering Problem: Given $n \times n$ Matrix $C = (c_{ij})$, find a permutation ϕ maximizing

$$\sum_{i<j} c_{\phi(i)\phi(j)}$$

Equivalently:

Linear Ordering Problems

Linear Ordering Problem: Given $n \times n$ Matrix $C = (c_{ij})$, find a permutation ϕ maximizing

$$\sum_{i<j} c_{\phi(i)\phi(j)}$$

Equivalently:
Find a simultaneous permutation of rows, columns of C, maximizing sum of the entries in the upper triangle.

Equivalently:
Find a complete acyclic subgraph in the complete directed graph with weights given by C.

See recent survey by Marti, Reinelt (2009)

0-1 Linear Integer Formulation

Grötschel, Jünger, Reinelt (1984): $x_{ij} = 1$ if i before j.

$$x_{ij} + x_{ji} = 1 \text{ used to eliminate } x_{ji} \text{ for } j > i.$$

Three-cycle constraints: $\forall$ distinct i, j, k:

$$x_{ij} + x_{jk} + x_{ki} \leq 2.$$

Together with $x_{ij} \in \{0, 1\}$, these constraints describe edge set of complete acyclic graph.

$$(LOR) \ \max \sum_{i<j} c_{ij} x_{ij} \text{ such that}$$

$$x_{ik} \leq x_{ij} + x_{jk} \leq 1 + x_{ik} \ \forall i < j < k$$

$$x_{ij} \in \{0, 1\}$$

Linear relaxation

The 3-cycle relaxation has $2\binom{n}{3} \approx \frac{1}{3}n^3$ inequality constraints. LP can be solved for n up to $n \approx 250$.

Branch and Bound computations: works for smaller values of $n \approx 40$.

Typical gap at root using 3-cycle relaxation: $3 - 10\%$

More facets are known, but difficult to separate.

LOP could also be formulated as Quadratic Assignment Problem, but Branch and Bound computations already hard for $n \approx 30$.

SDP relaxation

Variable transformation: $y_{ij} = 2x_{ij} - 1$ gives $y \in \{-1, 1\}$.
3-cylces become:

$$-1 \leq y_{ij} + y_{jk} - y_{ik} \leq 1 \; \forall i < j < k$$

Matrix lifting idea:

$$\begin{pmatrix} 1 \\ y \end{pmatrix} \begin{pmatrix} 1 \\ y \end{pmatrix}^T = \begin{pmatrix} 1 & y^T \\ y & yy^T \end{pmatrix} = \begin{pmatrix} 1 & y^T \\ y & Y \end{pmatrix}$$

The matrix yy^T is replaced by a new matrix Y. Matrix order is $\binom{n}{2} + 1$. Note that $diag(Y) = e$.

How should the linear inequalities be lifted (=made quadratic)?

Squared 3-cycles

The special form of 3-cycles

$$y_{ij} + y_{jk} - y_{ik} \in \{-1, 1\} \;\; \forall i < j < k$$

suggests squared form:

$$y_{ij,jk} - y_{ij,ik} - y_{jk,ik} = -1 \; \forall i < j < k.$$

Note that the inequalities become equations after lifting.

Other forms, like diagonal lifting are weaker, see Helmberg, R., Weismantel (2000).

Lovasz-Schrijver lifting is stronger (multiply each constraint with each (binary) variable), but has much higher complexity.

Further Cutting Planes

We could also multiply 3-cycles by $(1 - y_{lm})$ or $(1 + y_{lm})$, see Lovasz, Schrijver (1981).

Since Y should be -1,1 matrix, we can also include the triangle inequalities defining the metric polytope.

$$-y_{ij,rs} - y_{rs,uv} - y_{ij,uv} \leq 1,$$

$$y_{ij,rs} + y_{rs,uv} - y_{ij,uv} \leq 1$$

There are roughly $\frac{1}{12}n^6$ such constraints for LOR.

SDP relaxation of LOR

$$\max c^T y \text{ such that } Y - yy^T \succeq 0$$

$$diag(Y) = e$$

Y satisfies squared 3-cycle equations and triangle inequalities.

Matrix order: $\binom{n}{2} + 1$
Equations: $\binom{n}{2} + \binom{n}{3} + 1$
Inequalities: $\frac{1}{12} n^6$

Direct solution using standard interior-point based software not practical.

Dualize inequalities and 3-cycle equations. See Dissertation Hungerländer, Klagenfurt, in preparation

Facets for $n = 7$

Some typical results for facets of linear ordering polytope $n = 7$.

graph	n	opt	LP	3C	3C+M	3C+LSM	all
FC4	7	8	8.5	8.40	8.05	8.13	8
FC9	7	10	10.5	10.42	10.14	10.28	10
FC14	7	10	10.5	10.50	10.35	10.50	10.22
FC26	7	11	11.5	11.46	11.25	11.32	11

Only one out of 27 different classes of facets could not be recovered exactly.

Larger instances

We also consider larger instances from the literature and compare with the 3-cycle LP bound (3C-LP).

name	n	opt	3C-LP	gap	SDP	gap
P50-02	50	43835	44866	2.35	44100	0.60
P50-05	50	42907	44196	3.00	43285	0.82
P50-06	50	42325	43765	3.40	42775	1.06
P50-18	50	46897	48152	2.68	47401	1.07
Pal-31	31	285	300	5.26	297	4.21
Pal-43	43	543	597	9.94	569	4.79
T1d100	100	106852	114468	7.13	109966	2.91

SDP bound much better but also computationally much more expensive

Summary: Ordering Problems

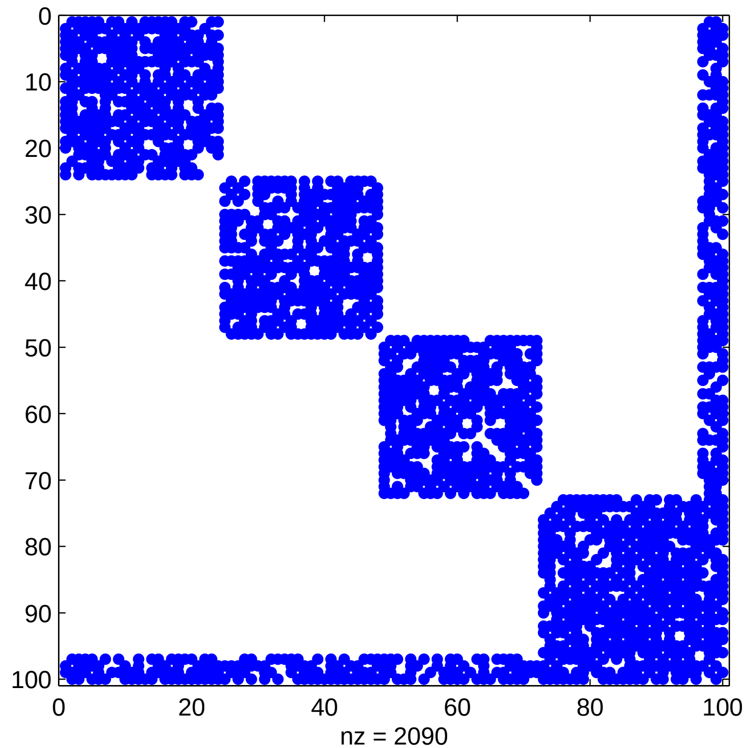
- Grötschel, Jünger, Reinelt (1984): Polyhedral approach, LP-based Branch and Bound
- Anjos, Vanelli (2008): Single-Row Facility Layout (special case of linear ordering with quadratic cost) first SDP based model
- Buchheim, Wiegele, Zheng (2009): linear ordering with quadratic cost: polyhedral investigations and SDP relaxations
- Approximation Results: 2-approximation is trivial, $2-\epsilon$ approximation is open.

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Vertex Separators

Given adjacency matrix A of a graph G. Does G have $S_k \subseteq V(G)$ such that $G \setminus S_k$ decomposes into $k-1$ pieces $S_1, \ldots, S_{k-1}$ of (roughly) equal size?



Here $k = 5$, last block separates the first four.

Partition Model

As before, we model partitions as $n \times k$ 0-1 matrices $X = (x_1, \ldots, x_k)$ such that $Xe = e, X^T e = m$, where $m^T = (\alpha, \ldots, \alpha, \beta)$ and $(k-1)\alpha + \beta = n$.
As a consequence: $X^T X = diag(m) = M$.

$$B = \begin{pmatrix} 0 & 1 & \ldots & 1 & 0 \\ 1 & 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ 1 & \ldots & 1 & 0 & 0 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

The total number of edges between the first $k-1$ blocks is given by $cut(S_1, \ldots, S_{k-1}) = \frac{1}{2}\mathrm{tr}(AXBX^T)$.

Hoffman-Wielandt relaxation

We use the Hoffman-Wielandt theorem

$$\min_{X^TX=I} \text{tr}(AXBX^T) = \sum_i \lambda_i(A)\lambda_{n+1-i}(B),$$

and get

$$2cut \geq \min_{X\in F} \text{tr}(AXBX^T) = \alpha \sum_{i=2}^{k-1} \lambda_i(L) - \frac{k-2}{n}\alpha\beta\lambda_n(L) = f(L).$$

Here L is the Laplacian associated to A and the set $F = \{X : Xe = e,\ X^Te = m,\ X^TX = \text{diag}(m)\}$. The proof of this result is rather technical, see forthcoming paper with A. Lisser.

Estimates for Separators

We would like to know whether $cut(S_1, \dots, S_{k-1}) > 0$ or not, in which case the desired separator exists.
If $f(L) > 0$ then there can not exist a separator of size β with the required properties, because $cut \geq f(L)$. This involves computing $\lambda_2(L), \dots, \lambda_{k-1}(L)$ and $\lambda_n(L)$. The function $f(L)$ is concave. Now consider

$$A_\epsilon = \{A : \sum_{i<j} a_{ij} = |E|, \ a_{ij} = 0 \text{ if } ij \notin E, \ a_{ij} \geq \epsilon\},$$

so $A \in A_\epsilon$ exactly if $A = \sum_{ij \in E} a_{ij} E_{ij}$ with
$a_{ij} \geq \epsilon, \ \sum_{ij} a_{ij} = |E|$.
See also Helmberg et al (2008), Boyd et al (2004)

Optimizing the bound

We would like to redistribute the edge weights a_{ij} so that the lower bound $f(L)$ is maximized. This leads to

$$\min\{\lambda_{max}(L_A) : \ A \in A_\epsilon\}.$$

Here, L_A denotes the Laplacian of A. This can be solved as SDP, or directly through eigenvalue optimization.

Preliminary computational results are encouraging.

Overview

- Combinatorial optimization and matrix liftings
- Cut Problems
- Coloring Problems
- Ordering Problems
- Graph separators
- Copositive relaxations

Stable-Set and other matrix cones

Let $X = \frac{1}{x^T x} x x^T$ where x is characteristic vector of stable set. Lovasz uses $X \succeq 0$.

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

Schrijver, McEliece et al consider $X \succeq 0, X \geq 0$.

In this case we can add constraints $x_{ij} = 0$ into a single equation (using adjacency matrix A) $\langle A, X\rangle = 0$.
Lovasz-Schrijver number ϑ^+ (1979):

$$\vartheta(G)^+ := \max\{\langle J, X\rangle : \ \langle A, X\rangle = 0, \ \ \text{tr}X = 1, \ \ X \succeq 0, X \geq 0\}.$$

Clearly

$$\alpha(G) \leq \vartheta(G)^+ \leq \vartheta(G).$$

Stable-Set and Copositive Programs

Recall $X = \frac{1}{x^T x} x x^T$. Lovasz uses $X \succeq 0$.
Schrijver considers $X \succeq 0, X \geq 0$.
X is in fact completely positive $X \in C^*$.

Stable-Set and Copositive Programs

Recall $X = \frac{1}{x^T x} x x^T$. Lovasz uses $X \succeq 0$.
Schrijver considers $X \succeq 0, X \geq 0$.
X is in fact completely positive $X \in C^*$.

DeKlerk and Pasechnik (2002)

$$\alpha(G) = \max\{\langle J, X\rangle : \langle A+I, X\rangle = 1, \quad X \in C^*\}$$

$$= \min\{y : y(A+I) - J \in C\}.$$

This is a copositive program with only one equation (in the primal problem).
This is a simple consequence of the Motzkin-Straus Theorem and is implicitly contained in Bomze et al (2000).

Birkhoff's theorem - lifted version

Π set of permutations. $X_\phi \ldots$ permutation matrix
Theorem (Birkhoff): $conv\{X_\phi : \phi \in \Pi\} = \Omega$
$\Omega = \{X : Xe = X'e = e,\ X \geq 0\}$ doubly stochastic matrices

We now consider the matrix lifting of this result.
$x_\phi = vec(X_\phi)$, hence $x_\phi \in \mathbb{R}^{n^2}$.

$$\mathcal{P} := conv\{x_\phi x_\phi^T :\ \phi \in \Pi\}$$

Birkhoff's theorem - lifted version

Π set of permutations. $X_\phi \ldots$ permutation matrix
Theorem (Birkhoff): $conv\{X_\phi : \phi \in \Pi\} = \Omega$
$\Omega = \{X : Xe = X'e = e,\ X \geq 0\}$ doubly stochastic matrices

We now consider the matrix lifting of this result.
$x_\phi = vec(X_\phi)$, hence $x_\phi \in I\!R^{n^2}$.

$$\mathcal{P} := conv\{x_\phi x_\phi^T :\ \phi \in \Pi\}$$

This set is contained in the space of $n^2 \times n^2$ matrices.
We describe permutation matrices X_ϕ by the following quadratic constraints:

$$X^T X = X X^T = I,\ \|JX\|^2 = (\sum_{ij} x_{ij})^2 = n^2$$

Birkhoff's theorem - lifted version (2)

Any $Y \in \mathcal{P}$ is completely positive ($Y \in C^*$). The quadratic equations become linear equations on Y which can be partitioned into $n \times n$ blocks Y^{ij}.

$$Y = \begin{pmatrix} Y^{11} & \dots & Y^{1n} \\ \vdots & & \vdots \\ Y^{n1} & \dots & Y^{nn} \end{pmatrix}$$

The quadratic equations in X turn into:

$$(*) \quad \sum Y^{ii} = I, \ \text{tr}(Y^{ij}) = \delta_{ij}, \sum y_{ij} = n^2.$$

Birkhoff's theorem - lifted version (2)

Any $Y \in \mathcal{P}$ is completely positive ($Y \in C^*$). The quadratic equations become linear equations on Y which can be partitioned into $n \times n$ blocks Y^{ij}.

$$Y = \begin{pmatrix} Y^{11} & \dots & Y^{1n} \\ \vdots & & \vdots \\ Y^{n1} & \dots & Y^{nn} \end{pmatrix}$$

The quadratic equations in X turn into:

$$(*) \quad \sum Y^{ii} = I, \ \mathrm{tr}(Y^{ij}) = \delta_{ij}, \ \sum y_{ij} = n^2.$$

Povh and R. (2008): $\mathcal{P} = \{Y : \ Y \in C^*, Y \text{ satisfies } (*)\}$

Application: QAP as copositive program

Quadratic assignment problem:

$$z_{qap} = \min\{\langle AXB, X\rangle : X \text{ permutation matrix}\}.$$

We also get, using previous result

$$z_{qap} = \min\langle B \otimes A, Y\rangle \text{ such that}$$

$$\sum Y^{ii} = I, \ \text{tr}(Y^{ij}) = \delta_{ij}, \ \sum y_{ij} = n^2, \ Y \in C^*.$$

Replacing $Y \in C^*$ by the tractable constraint $Y \succeq 0, \ Y \geq 0$ gives the currently strongest bounds for QAP, see Sotirov, R. (2005), Sun and Toh (2008).

A general copositive modeling theorem

Burer (2007) shows the following general result for the power of copositive programming:
The optimal values of P and C are equal: opt(P) = opt(C)

$$(P) \qquad \min x^T Q x + c^T x$$

$$a_i^T x = b_i,\ x \geq 0,\ x_i \in \{0,1\}\ \forall i \leq m.$$

Here $x \in I\!R^n$ and $m \leq n$.

$$(C) \qquad \min \mathbf{tr}(QX) + c^T x,\ \text{s.t. } a_i^T x = b_i,$$

$$a_i^T X a_i = b_i^2,\ X_{ii} = x_i\ \forall i \leq m,\ \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in C^*$$

Last Slide

- Copositive relaxations are often exact, but intractable: find good tractable approximations

- higher liftings
- optimization with polynomials, using Sum of Squares idea
- local solutions give primal heuristic

Last Slide

- Copositive relaxations are often exact, but intractable: find good tractable approximations

- higher liftings
- optimization with polynomials, using Sum of Squares idea
- local solutions give primal heuristic

- SDP modeling joins nonlinear, integer programming and theoretical computer science:

- reliable large-scale SDP solvers not yet available
- new approximation ideas with hyperplane rounding?

Last Slide

- Copositive relaxations are often exact, but intractable: find good tractable approximations

- higher liftings
- optimization with polynomials, using Sum of Squares idea
- local solutions give primal heuristic

- SDP modeling joins nonlinear, integer programming and theoretical computer science:

- reliable large-scale SDP solvers not yet available
- new approximation ideas with hyperplane rounding?

- Optimize over $X \succeq 0, \; X \geq 0??$

- Basic SDP relaxation of Max-Cut for $n \geq 10,000$??