Semidefinite Relaxations for Hard Combinatorial Problems

Henry Wolkowicz
University of Waterloo
Dept of Combinatorics and Optimization

OUTLINE

- Introduction to SDP
(geometry, duality, algorithms)

- Max-Cut Problem instance
(quadratic boolean programming)
tractable relaxations versus Lagrangian duality

- Other instances, e.g. QAP, GP, ...

- Theme: - Tractable relaxation implies Lagrangian relaxation which can be solved efficiently using SDP.
- Advantages of Lagrangian approach (recipe for SDP relaxation).

INTRODUCTION

Semidefinite programming (denoted SDP) is an extension of linear programming (LP), e.g. an (linear) SDP is:

$$(P) \qquad \begin{array}{rll} \mu^* := \min & C \bullet X \\ \text{s.t.} & \mathcal{A}X = b \\ & X \succeq 0, \end{array}$$

$C, X \in \mathcal{S}^n$, space of symmetric $n \times n$ matrices

---

inner product $C \bullet X = \text{trace}\, CX$

$A \succeq B$ if $A - B \succeq 0$ positive semidefinite

$$\mathcal{A} : \mathcal{S}^n \to \Re^m$$

linear operator: $(\mathcal{A}X)_i = \text{trace}\,(A_i X)$
$A_i \in \mathcal{S}^n,\ i = 1, \ldots m.$

Why use SDP?

Many applications: engineering, combinatorial optimization, statistics, matrix completions, approximation theory, nonlinear programming, ...

Many computationally hard problems can be modelled as quadratic programs.

These quadratic programs are themselves hard to solve numerically.

But, the Lagrangian relaxation can be solved efficiently using SDP.

Properties of SDP

SDP is a special case of the cone programming problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \succeq_K 0, \end{array}$$

K is a convex cone
$g(x) \succeq_K 0$ is cone partial order, $g(x) \in K$.

very general mathematical program

e.g. standard equality and inequality constraints when $K = \Re^n_+ \otimes \mathcal{S}^n_+ \otimes \{0\}$.

GEOMETRY

Much of the elegant geometry of polyhedral sets developed for LP can be extended to SDP. (e.g. Bohnenblust 1948, Barker-Carlson 1975 and more recently Lewis/98 and Pataki/98)

similarities to LP:
self-polar

$$\mathcal{P} = \mathcal{P}^+ := \{Y : X \bullet Y \geq 0, \forall X \in \mathcal{P}\};$$

homogeneous
for any $X, Y \in \mathrm{int}(\mathcal{P})$, there exists an invertible linear operator $\mathcal{A}$ that leaves $\mathcal{P}$ invariant and $\mathcal{A}(X) = Y$.

faces

$$\mathcal{F} = \{Y : \mathcal{N}(Y) \supset \mathcal{N}(X)\}, \quad X \in \text{ relint } \mathcal{F}$$

faces are *exposed*
$\mathcal{F} = \mathcal{P} \cap \phi^{\perp}$, where $\phi \in \mathcal{P} \cap \mathcal{F}^{\perp}$ *conjugate face*,
(Here $\cdot^{\perp}$ denote orthogonal complement.)

But difficulties in duality theory can arise due to:

$\mathcal{P} + \mathcal{F}^{\perp}$ is always closed

$\mathcal{P} + \text{span}(\mathcal{F})$ is never closed.

Duality Theory, Optimality Conditions

Extentions from LP to SDP, e.g. Bellman and Fan 1963, but strong duality theorems require a Slater-type constraint qualification (strict feasibility).

(Modified optimality conditions without CQ exit Borwein-Wolkowicz/81 and Ramana/98.)

$$(P) \qquad \begin{array}{rl} \mu^* := \min & C \bullet X \\ \text{s.t.} & \mathcal{A}X = b \\ & X \succeq 0, \end{array}$$

---

weak duality (using hidden constraints)

$$\begin{array}{rcl} \mu^* & = & \min_{X \succeq 0} \max_y C \bullet X + y^T(b - \mathcal{A}X) \\ & \geq & \max_y \min_{X \succeq 0} y^T b + (C - \mathcal{A}^* y) \bullet X \\ & = & \nu^*, \end{array}$$

$\mathcal{A}^*$ adjoint operator of $\mathcal{A}$

$$\mathcal{A}(X) \bullet y = X \bullet \mathcal{A}^*(y), \quad \forall X, \forall y$$

---

dual program

$$(D) \qquad \begin{array}{rl} \nu^* := \max & b^T y \\ \text{s.t.} & \mathcal{A}^* y \preceq C \end{array}$$

The duality theory gives rise to the characterization of optimality

$$
\begin{array}{rcll}
\mathcal{A}^* y + Z - C & = & 0 & \text{dual feasibility} \\
b - \mathcal{A}(X) & = & 0 & \text{primal feasibility} \\
ZX & = & 0 & \text{complementary slackness} \\
Z, X \succeq 0. & & &
\end{array}
\tag{1}
$$

$X, (y, Z)$ a primal-dual optimal pair. Z is the (dual) slack variable.

$$ZX = \mu I, \qquad \mu > 0$$

perturbed compl. slack. used for interior-point methods

First and second order optimality conditions e.g. Shapiro/94,99

Nondegeneracy and strict complementarity (Theorem of Goldman and Tucker) do not directly follow through from LP to SDP though generic, e.g. Shapiro/99, Pataki-Tuncel/98, Alizadeh-Haeberly-Overton/98.

(strict complementarity for SDP translates to $Z + X \succ 0$)

Complexity/algorithms: SDP are convex programs and fall into the class of problems that can be approximately solved in polynomial time by int-pt algorithms, Nesterov-Nemirovski/93.

Tractable Relaxations of Max-Cut

(quadratic boolean programming)

$$(MCQ) \quad \mu^* := \max_{x \in \mathcal{F}} q_0(x) \quad (:= x^TQx - 2c^Tx).$$

where $\mathcal{F} = \{\pm 1\}^n$

---

perturbing diagonal of Q on $\mathcal{F}$:

$$\begin{array}{rcl} q_u(x) & := & x^T(Q + \mathrm{Diag}\,(u))x - 2c^Tx - u^Te \\ & = & q_0(x), \quad \forall x \in \mathcal{F}, \end{array}$$

where e vector of ones.

Bound 0: trivial bound from diagonal perturbations:

$$\mu^* \le f_0(u) := \max_x q_u(x).$$

The function f_0 can take on the value $+\infty$. Let

$$S := \left\{ u : u^T e = 0, Q + \mathrm{Diag}\,(\mathrm{u}) \preceq 0 \right\}.$$

Then:

$$\boxed{\begin{array}{c} \mu^* \le B_0 := \min_u f_0(u) \\ \left(= \min_{u^T e = 0} f_0(u), \text{ if } S \ne \emptyset \right). \end{array}}$$

Using the hidden semidefinite constraint:

$$\mu^* \le B_0 = \min_{Q + \mathrm{Diag}\,(\mathrm{u}) \preceq 0} f_0(u).$$

Bound 1: relax the feasible set to the sphere of radius $\sqrt{n}$ (tractable trust region subproblem, TRS):

$$\mu^* \leq f_1(u) := \max_{||x||^2=n} \; q_u(x)$$

and

$$\boxed{\mu^* \leq B_1 := \min_u f_1(u).}$$

The inner maximization problem is called a trust region subproblem and is tractable. This bound provides the central tool in the proofs of equivalence.

Bound 2: box constraint:

$$\mu^* \leq f_2(u) := \max_{|x_i| \leq 1} \; q_u(x).$$

add the semidefinite constraint to make bound tractable.

$$\mu^* \leq \min_u f_2(u)$$

and

$$\boxed{\mu^* \leq B_2 := \min_{Q + \mathrm{Diag}\,(u) \preceq 0} f_2(u).}$$

Bound B_1^c: lift to eigenvalue bound:

$$Q^c := \begin{bmatrix} 0 & -c^T \\ -c & Q \end{bmatrix}$$

$$q_u^c(y) := y^T(Q^c + \mathrm{diag}\,(\mathrm{u}))\mathrm{y} - \mathrm{u}^{\mathrm{T}}\mathrm{e}$$

$$\mu^* \leq f_1^c(u) := \max_{||y||^2=n+1} q_u^c(y)$$

where

$$\max_{||y||^2=n+1} q_u^c(y) = (n+1)\lambda_{\max}(Q^c+\mathrm{diag}\,(\mathrm{u}))-\mathrm{u}^{\mathrm{T}}\mathrm{e}$$

$$\boxed{\mu^* \leq B_1^c := \min_u f_1^c(u).}$$

Similarly, we get equivalent bounds B_0^c and homogenized bounds for the other models.

Bound B_3: SDP bound:

After homogenization ($c = 0$), use

$$x^T Q x = \operatorname{trace} \mathrm{x^T Q x} = \operatorname{trace} \mathrm{Q x x^T}$$

and, for $x \in \mathcal{F}$, $y_{ij} = x_i x_j$ defines a symmetric, rank one, positive semidefinite matrix Y with diagonal elements 1. Relax the rank one condition.

$$\boxed{\begin{array}{lcc} B_3 := & \max & \operatorname{trace} \mathrm{QY} \\ & \text{subject to} & \operatorname{diag}(\mathrm{Y}) = \mathrm{e} \\ & & Y \succeq 0. \end{array}}$$

Summary: ref. Polyak-Rendl-Wolkowicz (PRW) 1995. (without restrictions e.g. $u^T e = 0$)

$$\boxed{\begin{array}{rcl} B_0 & = & \min\limits_u \max\limits_x q_u(x) \\ B_1 & = & \min\limits_u \max\limits_{x^T x = n} q_u(x) \\ B_2 & = & \min\limits_u \max\limits_{-1 \le x_i \le 1} q_u(x) \\ B_3 & = & \max\{\text{trace}\, \mathrm{Q}^{\mathrm{c}}\mathrm{Y} : \text{diag}\,(\mathrm{Y}) = \mathrm{e},\ \mathrm{Y} \succeq 0.\} \\ B_1^c & = & \min\limits_u \max\limits_{y^T y = n+1} q_u^c(y) \end{array}}$$

Now replace ± 1 constraints with $x_i^2 = 1, \forall i$.

$$
\begin{array}{lll}
(P_E) & \text{max} & q_0(x) = x^T Q x - 2c^T x \\
& \text{subject to} & x_i^2 = 1, \quad i = 1, \cdots, n.
\end{array}
$$

B_L denotes Lagrangian relaxation bound.

Following our theme:

Theorem B_L equals all above bounds.

ref PR, PRW/95.
(The proofs come from exploiting strong Lagrangian duality of TRS. More details below.)

A Strengthened SDP Bound for MC

(ref Anjos-Wolkowicz/99 -Illustration of the recipe for SDP relaxation. based on a second lifting - motivated by strong duality results that follow from adding redundant constraints of type XX^T, ref. Anstreicher-Wolkowicz/98 and below)

first lifting procedure

$$X = xx^T, \quad \text{Diag}\, \mathrm{X} = \mathrm{e}$$

implies second lifting procedure

$$X^2 = xx^T xx^T = nX.$$

and provides an equivalent quadratic matrix model for MC

MCQ is equivalent to:

$$\begin{array}{rll} \mu^* := & \max & \text{trace}\, QX \\ & \text{s.t.} & \text{diag}\,(X) = e \\ & & X \succeq 0 \\ & & X \text{ is rank } 1. \end{array}$$

which is equivalent to

$$\begin{array}{rll} \mu^* := & \max & \text{trace}\, QX \\ & \text{s.t.} & \text{diag}\,(X) = e \\ & & X^2 - nX = 0 \end{array}$$

Since X and X^2 can be mutually diagonalized.

Now add redundant constraints and get:

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

where E is the matrix of ones.

Note: $X^2 = nX$, $\text{trace}\,\mathrm{X} = \mathrm{n}$ implies X is rank one. But, we cannot solve this nonconvex problem in general, i.e. we have to look at the Lagrangian relaxation with vectors dimension $t(n)$.

Recipe for SDP relaxations:

1. add redundant constraints

2. take Lagrangian dual

3. homogenize

4. use hidden semidefinite constraint to obtain SDP equivalent *(check Slater's constraint qualification - strict feasibility)*

5. take Lagrangian dual again

6. check Slater's CQ again - project if it fails

7. delete redundant constraints

Notation:
(Interesting operators and adjoints –fun for me, if for no one else)

- $S \in \mathcal{S}^n, \quad t(n) = \frac{n(n+1)}{2}$

- $s = \operatorname{svec}(S) \in \Re^{t(n)}$ vector formed (columnwise) from S ignoring strictly lower triang.

- $S = \operatorname{sMat}(s)$ inverse of svec

- $\operatorname{hMat}(v)$ is adjoint of svec, off-diagonal terms are multiplied by a half

- $\operatorname{dsvec}(S)$ is adjoint of sMat, operator like svec but off diagonal elements are multiplied by 2

- $\operatorname{sdiag}(s) := \operatorname{diag}(\operatorname{sMat}(s))$

- $\operatorname{vsMat}(s) := \operatorname{vec}(\operatorname{sMat}(s))$
- $\operatorname{vsMat}^*(s) = \operatorname{dsvec}\left(\left(\operatorname{Mat}(v) + \operatorname{Mat}(v)^T\right)/2\right).$

start with equivalent program to MCQ:

$$
\text{MC2} \qquad \begin{array}{rll} \mu^* = & \max & \operatorname{trace} QX \\ & \text{s.t.} & \operatorname{diag}(X) = e \\ & & X \circ X = E \\ & & X^2 - nX = 0, \end{array} \qquad (2)
$$

to efficiently apply Lagrangian relaxation and not lose information from the linear constraint, replace the constraint with the norm constraint $||\operatorname{diag}(X) - e||^2 = 0$. Add $1 - y_0^2 = 0$.

We use: $X = \operatorname{sMat}(x)$ is a symmetric matrix.

$$
\begin{array}{ll} \max & \operatorname{trace}(Q \operatorname{sMat}(x))\, y_0 \\ \text{s.t.} & \operatorname{sdiag}(x)^T \operatorname{sdiag}(x) - 2e^T \operatorname{sdiag}(x) y_0 + n = 0 \\ & \operatorname{sMat}(x) \circ \operatorname{sMat}(x) = E \\ & \operatorname{sMat}(x)^2 - n \operatorname{sMat}(x) y_0 = 0 \\ & 1 - y_0^2 = 0. \end{array}
$$

take the Lagrangian dual; Lagrange multipliers w, T, S

$$
\begin{aligned}
\mu^* &\leq \\
\nu_2^* &:= \min_{w,T,S} \max_{x, y_0^2=1} \operatorname{trace}\left(\mathrm{Q}\,\mathrm{sMat}\,(\mathrm{x})\right) \mathrm{y}_0 \\
&\quad + w(\mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n) \\
&\qquad + \operatorname{trace} \mathrm{T}(\mathrm{E} - \mathrm{sMat}\,(\mathrm{x}) \circ \mathrm{sMat}\,(\mathrm{x})) \\
&\qquad + \operatorname{trace} \mathrm{S}((\mathrm{sMat}\,(\mathrm{x}))^2 - \mathrm{n}\,\mathrm{sMat}\,(\mathrm{x}) \mathrm{y}_0).
\end{aligned}
$$

move variable y_0 into the Lagrangian without increasing the duality gap since this is a trust region subproblem

$$
\begin{aligned}
\nu_2^* = \; & \min_{t,w,S} \max_{x,y_0} \operatorname{trace}\left(\mathrm{Q}\,\mathrm{sMat}\,(\mathrm{x})\right) \mathrm{y}_0 \\
& + w(\mathrm{sdiag}\,(x)^T \mathrm{sdiag}\,(x) - 2e^T \mathrm{sdiag}\,(x) y_0 + n) \\
& + \operatorname{trace} \mathrm{T}(\mathrm{sMat}\,(\mathrm{x}) \circ \mathrm{sMat}\,(\mathrm{x}) - \mathrm{E}) \\
& + \operatorname{trace} \mathrm{S}((\mathrm{sMat}\,(\mathrm{x}))^2 - \mathrm{n}\,\mathrm{sMat}\,(\mathrm{x}) \mathrm{y}_0) \\
& + t(1 - y_0^2).
\end{aligned}
$$

The inner maximization of the above relaxation is an unconstrained pure quadratic maximization, i.e. the optimal value is infinity unless the Hessian is negative semidefinite **(hidden constraint)** is which case $x = 0$ is optimal.

Therefore we need to evaluate the Hessian of the Lagrangian.

Using $Q\mathrm{sMat}\,(x) = x^T \mathrm{dsvec}\,(Q)$, and adding a 2 for convenience, we get the constant part (no Lagrange multipliers) of the Hessian:

$$2H_c := 2 \begin{pmatrix} 0 & \frac{1}{2}\mathrm{dsvec}\,(Q)^T \\ \frac{1}{2}\mathrm{dsvec}\,(Q) & 0 \end{pmatrix}.$$

nonconstant part:

Use:

$$\text{dsvec}\,\text{Diag}\,\text{diag}\,\text{sMat} = \text{sdiag}^*\text{sdiag} = \text{Diag}\,\text{svec}\,(\text{I})$$

rewrite the quadratic forms as follows:

$$\text{sdiag}\,(x)^T\text{sdiag}\,(x) = x^T\,(\text{dsvec}\,\text{Diag}\,\text{diag}\,\text{sMat}\,)\,x;$$
$$e^T\text{sdiag}\,(x) = (\text{dsvec}\,\text{Diag}\,\text{e})\,x;$$

$$\begin{aligned}\text{trace}\,\text{S}(\text{sMat}\,(\text{x}))^2 &= \text{trace}\,\text{sMat}\,(\text{x})\,\text{S}\,\text{sMat}\,(\text{x}) \\ &= x^T\text{dsvec}\,(S\text{sMat}\,(x)) \\ &= x^T\,(\text{dsvec}\,S\text{sMat}\,)\,x;\end{aligned}$$

$$\begin{aligned}&\text{trace}\,\text{T}(\text{sMat}\,(\text{x}) \circ \text{sMat}\,(\text{x})) \\ &\quad = x^T\,\{\text{dsvec}\,(T \circ \text{sMat}\,(x))\} \\ &\quad = x^T\,(\text{dsvec}\,(T \circ \text{sMat}\,))\,x.\end{aligned}$$

use the *negative* of the Hessian and split it into four linear operators with the factor 2:

$$
\begin{aligned}
2\mathcal{H} &:= 2\mathcal{H}_1(w) + 2\mathcal{H}_2(T) + 2\mathcal{H}_3(S) + 2\mathcal{H}_4(t) \\
&:= 2w \begin{pmatrix} 0 & (\text{dsvec}\,\text{Diag}\,\text{e})^{\text{T}} \\ (\text{dsvec}\,\text{Diag}\,\text{e}) & -\text{sdiag}\,{}^{*}\text{sdiag} \end{pmatrix} \\
&\quad + 2 \begin{pmatrix} 0 & 0 \\ 0 & \text{dsvec}\,(T \circ \text{sMat}\,) \end{pmatrix} \\
&\quad + 2 \begin{pmatrix} 0 & \frac{1}{n}\text{dsvec}\,(S)^T \\ \frac{1}{n}\text{dsvec}\,(S) & -\text{dsvec}\,S\text{sMat} \end{pmatrix} \\
&\quad + 2t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}
$$

The matrix sdiag *sdiag $\in \mathcal{S}^{t(n)}$ is diagonal with elements determined using

$$\begin{array}{rcl} e_i^T \left(\text{sdiag}\,^*\text{sdiag}\,\right) e_j & = & \text{sdiag}\,(e_i)^T \text{sdiag}\,(e_j) \\ & = & \left\{ \begin{array}{ll} 1 & \text{if i=j=t(k)} \\ 0 & \text{otherwise.} \end{array} \right. \end{array}$$

Similarly, we find that, for $T = \sum_{ij} t_{ij} E_{ij}$, where the matrices E_{ij} are the elementary matrices $e_i e_j^T + e_j e_i^T$, we have

$$\text{dsvec}\,(T \circ \text{sMat}\,) = \sum_{ij} t_{ij} \text{dsvec}\,\left(E_{ij} \circ \text{sMat}\right).$$

cancel the 2 and get the (equivalent to the Lagrangian dual) semidefinite program MCDSDP2

$$\begin{array}{lll} \nu_2^* = & \min & nw + \text{trace}\,\mathrm{E}\mathrm{T} + \text{trace}\,0\mathrm{S} + \mathrm{t} \\ & \text{s.t.} & \mathcal{H}(w, T, S, t) \succeq H_c \end{array}$$

take T sufficiently positive definite and t sufficiently large, then we can guarantee Slater's constraint qualification.

the dual of this SDP is the strengthened SDP relaxation of MC:

$$\text{MCPSDP2} \qquad \begin{array}{rll} \nu_2^* = & \max & \text{trace}\,\mathrm{H_c}\mathrm{Y} \\ & \text{s.t.} & \mathcal{H}_1^*(Y) = n \\ & & \mathcal{H}_2^*(Y) = E \\ & & \mathcal{H}_3^*(Y) = 0 \\ & & \mathcal{H}_4^*(Y) = 1 \\ & & Y \succeq 0. \end{array}$$

we need to calculate the adjoint operators and remove redundant constraints in MCDSDP2.

$$Y \cong \begin{pmatrix} y_0 \\ x \end{pmatrix} \begin{pmatrix} y_0 & x^T \end{pmatrix}, \qquad X = \mathrm{sMat}\,(x)$$

simplified SDP relaxation MCPSDP2

$$\begin{array}{ll} \max & \mathrm{trace}\, \mathrm{H_c Y} \\ \text{s.t.} & \mathrm{diag}\,(\mathrm{Y}) = \mathrm{e} \\ & Y_{0,t(i)} = 1, \quad \forall i = 1, \ldots, n \\ & \sum\limits_{k=1}^{i} Y_{t(i-1)+k, t(j-1)+k} \\ & \quad + \sum\limits_{k=i+1}^{j} Y_{t(k-1)+i, t(j-1)+i} \\ & \qquad + \sum_{k=j+1}^{n} Y_{t(i-1)+i, t(k-1)+j} \\ & \qquad\quad - nY_{0,t(j-1)+i} = 0 \\ & \qquad\qquad \forall 1 \le i < j \le n \\ & Y \succeq 0, Y \in S^{t(n)+1}. \end{array}$$

This problem has $2t(n) - 1$ constraints.

surprise result

LEMMA
Suppose that Y is feasible in MCPSDP2. Then the first row

$$\mathrm{sMat}\left(Y_{0,1:t(n)}\right) \succeq 0.$$

Proof: Let $x = Y_{0,2:t(n)}$. $X^2 = nX$ constraint:

$$n\mathrm{sMat}(x) = \mathrm{sMat}(x)\mathrm{sMat}(x) = \mathrm{sMat}(x)x^T\mathrm{sMat}^*.$$

identify xx^T with lower right block of Y; get congruence of a positive semidefinite matrix. Alternatively: using $\mathcal{H}_3^*$, and $\mathrm{dsvec}(\cdot)\mathrm{sMat}$ is self-adjoint operator

$$n\mathrm{dsvec}^*(x) = \mathrm{dsvec}\,\bar{Y}\mathrm{sMat} = (\mathrm{dsvec}\,\bar{Y}\mathrm{sMat})^*,$$

where $\bar{Y}$ is the bottom right block of Y. Again congruence. ∎

n	Wt opt.	MCSDP (% rel. err)	MCPSDP2 (% rel. err)	Num. rank
5	4	4.5225 (13.06%)	4.2890 (7.22%)	2
7	56	56.4055 (0.72%)	56.0954 (0.17%)	3
8	30	30.2015 (0.67%)	30. (e-10%)	1
9	58	58.9361 (1.61%)	58.1182 (0.20%)	3
10	64	64.08 (0.1268%)	64 (e-08%)	3
12	88	90.3919 (2.72%)	89.5733 (1.79%)	4

The first line of results corresponds to solving both MC relaxations for a 5-cycle with unit edge-weights; the others come from randomly generated weighted graphs.

Other instances:

Trust Region Subproblem

$$\begin{array}{rl} \mu^* := \min & q_0(x) \\ \text{s.t.} & x^T x - \delta^2 \leq 0 \text{ (or } = 0). \end{array}$$

(more general nonconvex constraints $\alpha \leq q(x) \leq \beta$.)
for "$\leq$," the Lagrangian dual is:

DTRS $\qquad \nu^* := \max\limits_{\lambda \geq 0} \min\limits_{x} \; q_0(x) + \lambda(x^T x - \delta^2).$

strong duality holds and equivalent to (ref Stern-Wolkowicz/95) the (concave) nonlinear semidefinite program

$$\text{DTRS} \qquad \begin{array}{rl} \nu^* := \max & g_0^T (Q + \lambda I)^\dagger g_0 - \lambda \delta^2 \\ \text{s.t.} & Q + \lambda I \succeq 0 \\ & \lambda \geq 0. \end{array}$$

Again our theme holds: if we have a tractable problem then strong duality holds. (tractable:- ref e.g. More-Sorensen/83) short proof of strong duality (e.g. A. Lewis)

WLOG TRS is nonconvex; smallest eigenvalue of Q_0, denoted γ, is negative:

$$\begin{array}{rcll}
\mu^* & = & \min\limits_{x^Tx\leq\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^t x + \gamma x^T x \\
 & = & \min\limits_{x^Tx=\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^t x + \gamma x^t x, \\
 & = & \min\limits_{x^Tx=\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^t x + \gamma \delta^2 \\
 & = & \min\limits_{x^Tx\leq\delta^2} & x^T(Q_0-\gamma I)x - 2c_0^t x + \gamma \delta^2, \\
 & = & \max\limits_{\lambda\geq 0}\min\limits_{x} & x^T(Q_0-\gamma I)x - 2c_0^t x + \lambda(x^T x - \delta^2) + \\
 & = & \max\limits_{\lambda\geq 0}\min\limits_{x} & x^T Q_0 x - 2c_0^t x + (\lambda-\gamma)(x^T x - \delta^2) \\
 & \leq & \max\limits_{\lambda\geq \gamma}\min\limits_{x} & x^T Q_0 x - 2c_0^t x + (\lambda-\gamma)(x^T x - \delta^2) \\
 & = & \nu^* \leq \mu^*. &
\end{array}$$

Two Trust Region Subproblem TTRS consists in minimizing a (possibly nonconvex) quadratic function subject to a norm and a least squares constraint.

ref in SQP methods by Celis-Dennis-Tapia.

TTRS can have a nonzero duality gap, ref Peng-Yuan.

if objective not convex, then the primal may not be attained, ref Luo-Zhang.

TRS can have at most one local and nonglobal optimum, ref Martinez.

Still an open problem whether TTRS is an NP-hard or a polynomial time problem.

Note: a quadratic constraint can be written as

$$y^t P y \leq \delta$$

after the homogenization. This is lifted to

$$\text{trace}\, \mathsf{PY} \leq \delta.$$

Second lifting:

$$YPY \preceq \delta Y.$$

This strictly strengthens the SDP relaxation.

Orthogonally Constrained Programs with Zero Duality Gaps

permutation matrices are a subset of orthogonal matrices

$$\Pi \subset \mathcal{O} := \{X : XX^T = I\}$$

(Stiefel manifold ref e.g. Edelman,Arias,Smith)

A and B $n \times n$ symmetric matrices

$$\mathrm{QQP}_{\mathrm{O}} \qquad \begin{array}{rll} \mu^* := & \min & \operatorname{trace} AXBX^{\mathrm{T}} \\ & \text{s.t.} & XX^T = I. \end{array}$$

Tractable problem; can be solved using classical Hoffman-Wielandt inequality. Provides bounds for e.g. QAP.

Proof using first application of Lagrange multipliers:

$$L(X,S) = \mathrm{trace}\, \mathrm{AXBX}^{\mathrm{T}} + \mathrm{S}(\mathrm{XX}^{\mathrm{T}} - \mathrm{I})$$

$$0 = \langle \nabla L(X,S), h \rangle = 2\mathrm{trace}\, \mathrm{AXBh}^{\mathrm{T}} + \mathrm{SXIh}^{\mathrm{T}} \quad \forall \mathrm{h}.$$

Therefore $AXBX^T = -S = -S^T$, i.e. A and XBX^T are mutually diagonalizable. The optimal value is then

$$\mu^* = \sum_i \lambda_i(A)\lambda_{n-i+1}(B)$$

But, Lagrangian dual can have a duality gap, ref Zhao,Karisch,Rendl,Wolkowicz

Lagrangian of (P)

$L(X,S) = \mathrm{trace}\, AXBX^t + \mathrm{trace}\, SXX^t - \mathrm{trace}\, S.$

Lagrangian dual is

$$\nu^* = \max_S \min_X L(X,S).$$

The hidden constraint (Hessian is psd) yields the dual

$$(D) \quad \begin{array}{rcc} \mu^D = & \max\limits_{\hat{S}=\hat{S}^t} & -\mathrm{trace}\, \hat{S} \\ & \text{subject to} & (B \otimes A + I \otimes \hat{S}) \succeq 0. \end{array}$$

A simple Example

Consider the following 2×2 example

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad B := \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}.$$

$\mu^* = 10$

$$B \otimes A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

But $s_{11} \geq -3$ and $s_{22} \geq -6$; to maximize (D), equality must hold, and therefore $-\text{trace}\, \hat{\mathsf{S}} = 9$ in the optimum.

However, add redundant constraint $X^T X = I$. Get $T \otimes I$ in Hessian and $-\text{trace}\, \mathsf{T}$ in objective function. $s_{11} \geq -3 - t_{11}$ $s_{11} \geq -4 - t_{22}$ $s_{22} \geq -6 - t_{11}$ $s_{22} \geq -8 - t_{22}$ So $t_{22} = -1$. Duality gap is closed.

Summary: constraints $XX^T = I$ and $X^T X = I$ are equivalent. Add redundant constraints.

$$\text{QQP}_{\text{OO}} \qquad \mu^O := \begin{array}{ll} \min & \text{trace}\,\text{AXBX}^{\text{T}} \\ \text{s.t.} & XX^T = I,\ X^T X = I. \end{array}$$

dual problem is

$$\begin{array}{lll} \text{DQQP}_{\text{OO}} & & \\ \mu^O \geq \mu^D := & \max & \text{trace}\,\text{S} + \text{trace}\,\text{T} \\ & \text{s.t.} & (I \otimes S) + (T \otimes I) \preceq (B \otimes A) \\ & & S = S^T,\ T = T^T. \end{array}$$

Theorem (Anstreicher-Wolkowicz/98) Strong duality holds for QQP_{OO} and DQQP_{OO}, i.e. $\mu^D = \mu^O$ and both primal and dual are attained. ∎

Theme again holds.

Other applications:

Weighted Sums of Eigenvalues;

Graph Partitioning Problem;

TRS like constraints $\{X : XX^T \preceq I\}$ (ref Anstreicher-Wolkowicz-Xin-Yuan/99) (extension of Hoffman-Wielandt inequality)

Quadratic Assignment Problem, QAP
(additional constraints)

$$(QAP) \quad \mu^* := \min_{X \in \Pi} \ \mathrm{trace}\, AXBX^T - 2CX^T,$$

A, B real symmetric $n \times n$ matrices, C real $n \times n$ matrix, Π set of permutation matrices.

model for e.g. allocating set of n facilities to set of n locations while minimizing quadratic objective (distance-flow)

QAP is NP-hard and, in practice, problems of moderate sizes, such as $n = 16$, are still considered very hard.

recent surveys/books: Burkard/91, Rendl-Pardalos-Wolkowicz/93, Cellis/98???.

Questions:

1. interesting numerical and theoretical difficulties, e.g. loss of constraint qualification and loss of sparsity in the optimality conditions.

2. Can the new bound compete with other bounding techniques in speed and quality?

3. comparison on the Nugent test problems?

4. add new facet inequalities?

Notation:

$\mathcal{E} := \{X : Xe = X^T e = e\}$
set of matrices satisfying *assignment constraints*

$\mathcal{Z} := \left\{X : X_{ij} \in \{0, 1\}\right\}$
set of (0,1)-matrices

$\mathcal{N} := \left\{X : X_{ij} \geq 0\right\}$
set of *nonnegative matrices*

$\mathcal{O} := \{X : XX^T = X^T X = I\}$
set of orthogonal matrices

$$\Pi = \mathcal{E} \cap \mathcal{Z} = \mathcal{O} \cap \mathcal{Z}$$

rewrite QAP as (add redundant constraints)

$$\begin{array}{ll}
\min & \operatorname{trace} AXBX^T - 2CX^T \\
\text{s.t.} & XX^T = X^T X = I \\
& \|Xe - e\|^2 = 0 \\
& \|X^T e - e\|^2 = 0 \\
& X_{ij}^2 - X_{ij} = 0, \quad \forall i, j. \\
& \operatorname{diag}(X_{:i} X_{:j}^T) = 0, \quad \text{if } i \neq j \\
& X_{:i} X_{:j}^T - \operatorname{Diag}\left(\operatorname{diag}(X_{:i} X_{:j}^T)\right) = 0, \quad \text{if } i = j
\end{array}$$

Lagrangian Relaxation

$$
\begin{aligned}
\mu_{\mathcal{O}} &\geq \mu_{\mathcal{L}} \\
&:= \max_{W,u_0,v_0\cdots} \min_{XX^T=X^TX=I} \{\mathrm{trace}\, AXBX^T - 2CX^T \\
&\qquad + \textstyle\sum_{ij} W_{ij}(X_{ij}^2 - X_{ij}) \\
&\qquad + u_0\|Xe - e\|^2 \\
&\qquad + v_0\|X^Te - e\|^2 \\
&\qquad + \ldots\}.
\end{aligned}
$$

homogenize the Lagrangian using scalar x_0 and constraint $x_0^2 = 1$.

We get the lower bound
(quadratic, linear constant in X)

$$\max_{W,S_b,S_o,u_0 v_0,,w_0} \min_{X,\ x_0} \{\mathrm{trace}\,[\mathrm{AXBX^T} \\ + u_0\|Xe\|^2 + v_0\|X^T e\|^2 \\ + W(X \circ X)^T + w_0 x_0^2 \\ + S_b X X^T + S_o X^T X] \\ - \mathrm{trace}\,\mathrm{x_0(2C+W)X^T} \\ - 2x_0 u_0 e^T (X + X^T) e \\ + \ldots \\ - w_0 - \mathrm{trace}\,\mathrm{S_b} - \mathrm{trace}\,\mathrm{S_o} + 2\mathrm{nu}$$

Apply the hidden semidefinite constraint to get an SDP:

$$\begin{array}{ll} \max & -w_0 - \mathrm{trace}\,\mathrm{S_b} - \mathrm{trace}\,\mathrm{S_o} + \ldots \\ \text{s.t.} & L_Q + \mathrm{Arrow}\,(w) + \mathrm{B^0 Diag}\,(S_b) \\ & \quad + \mathrm{O^0 Diag}\,(S_o) + u_0 D + \ldots \succeq 0. \end{array}$$

check Slater's condition - OK
take Lagrangian dual again

$$
\begin{array}{ll}
\min & \operatorname{trace} L_Q Y \\
\text{s.t.} & b^0\operatorname{diag}(Y) = I, \quad o^0\operatorname{diag}(Y) = I \\
& \operatorname{arrow}(Y) = e_0, \quad \operatorname{trace} DY = 0 \\
& \ldots \\
& Y \succeq 0,
\end{array}
$$

check Slater's condition again - $D \succeq 0$ so it fails!! But we can project onto the minimal face!!

Now remove redundant constraints to get simplified SDP relaxation simple projected relaxation with $n^3 - 2n^2 + 1$ constraints.

$$\begin{array}{rll} \mu_{R2} := & \min & \mathrm{trace}\,(\hat{\mathrm{V}}^{\mathrm{T}}\mathrm{L_Q}\hat{\mathrm{V}})\mathrm{R} \\ & \text{s.t.} & \mathcal{G}_{\bar{J}}(\hat{V}R\hat{V}^T) = E_{00} \\ & & R \succeq 0. \end{array}$$

The dual problem is

$$\begin{array}{rll} \mu_{R2} = & \max & -Y_{00} \\ & \text{s.t.} & \hat{V}^T(L_Q + \mathcal{G}^*_{\bar{J}}(Y))\hat{V} \succeq 0. \end{array}$$

Concluding Remarks:

In each case of a tractable bound for a nonconvex problem, the structure allows for redundant constraints to be added to close the Lagrangian duality gap.

What is the correct question about "best"? Can we close or reduce the duality gap in general?

	Sol.	GLB	ELI	RRD	EVB3	μ_{R1}	μ_{R2}	μ_{R3}
Esc16a	68	38	47	68	50	47	50	47
Esc16b	292	220	250	278	276	250	276	277
Esc16c	160	83	95	118	113	95	113	110
Esc16d	16	3	-19	4	-12	-19	-12	-6
Esc16e	28	12	6	14	13	6	13	9
Esc16g	26	12	9	14	11	9	11	10
Esc16h	996	625	708	704	708	708	909	806
Esc16i	14	0	-25	0	-21	-25	-21	-6
Esc16j	8	1	-6	2	-4	-6	-4	-4
Had12	1652	1536	1573	n.a.	1595	1604	1640	1648
Had14	2724	2492	2609	n.a.	2643	2651	2709	2703
Had16	3720	3358	3560	n.a.	3601	3612	3678	3648
Had18	5358	4776	5104	n.a.	5176	5174	5286	5226
Had20	6922	6166	6625	n.a.	6702	6713	6847	6758
Kra30a	88900	68360	63717	76003	n.a.	69736	–	–
Kra30b	91420	69065	63818	76752	n.a.	70324	–	–
Nug12	578	493	472	523	498	486	530	547
Nug14	1014	852	871	n.a.	898	903	959	967
Nug15	1150	963	973	1041	1001	1009	1060	1075
Nug16a	1610	1314	1403	n.a.	1455	1461	1527	1520
Nug16b	1240	1022	1046	n.a.	1081	1082	1138	1132
Nug17	1732	1388	1487	n.a.	1521	1548	1621	1604
Nug18	1930	1554	1663	n.a.	1707	1723	1801	1776
Nug20	2570	2057	2196	2182	2290	2281	2385	2326
Nug21	2438	1833	1979	n.a.	2116	2090	2252	2157
Nug22	3596	2483	2966	n.a.	3174	3140	3394	3234
Nug24	3488	2676	2960	n.a.	3074	3068	–	–
Nug25	3744	2869	3190	n.a.	3287	3305	–	–
Nug30	6124	4539	5266	4805	5448	5413	–	–

QAPLIB instances 1

	Sol.	GLB	ELI	RRD	EVB3	μ_{R1}	
Rou12	235528	202272	200024	224278	201337	208685	2209
Rou15	354210	298548	296705	324869	297958	306833	3231
Rou20	725522	599948	597045	643346	n.a.	615549	6424
Scr12	31410	27858	4727	29872	n.a.	11117	242
Scr15	51140	44737	10355	49264	n.a.	17046	420
Scr20	110030	86766	16113	95113	n.a.	28535	830
Tai12a	224416	195918	193124	n.a.	195673	203595	2153
Tai15a	388214	327501	325019	n.a.	327289	333437	3494
Tai17a	491812	412722	408910	n.a.	410076	419619	4412
Tai20a	703482	580674	575831	n.a.	n.a.	591994	6187
Tai25a	1167256	962417	956657	n.a.	n.a.	974004	
Tai30a	1818146	1504688	1500407	n.a.	n.a.	1529135	
Tho30	149936	90578	119254	100784	n.a.	125972	

QAPLIB instances 2

	Sol.	GLB	ELI	RRD	EVB3	μ_{R1}	μ_{R2}	μ_{R3}
Car10ga	4954	3586	4079	n.a.	4541	4435	4853	4919
Car10gb	8082	6139	7211	n.a.	7617	7600	7960	8035
Car10gc	8649	7030	7837	n.a.	8233	8208	8561	8612
Car10gd	8843	6840	8006	n.a.	8364	8319	8666	8780
Car10ge	9571	7627	8672	n.a.	8987	8910	9349	9473
Car10pa	32835	28722	-4813	n.a.	n.a.	1583	12492	30359
Car10pb	14282	12546	-14944	n.a.	n.a.	-5782	9934	13361
Car10pc	14919	12296	-17140	n.a.	n.a.	-8040	2473	13655
Esc08a	2	0	-2	0	n.a.	-2	0	2
Esc08b	8	1	-2	2	n.a.	-2	3	6
Esc08c	32	13	8	22	n.a.	9	18	30
Esc08d	6	2	-2	2	n.a.	-2	2	6
Esc08e	2	0	-6	0	n.a.	-6	-4	1
Esc08f	18	9	8	18	n.a.	9	13	18
Nug05	50	50	47	50	50	49	50	50
Nug06	86	84	69	86	70	74	85	86
Nug07	148	137	125	148	130	132	144	148
Nug08	214	186	167	204	174	179	197	210

Numerical Results for old instances

	Sol.	μ_{R1}	μ_{R2}	μ_{R3}
Nug12C	578	492	534	545
Nug15C	1150	1012	1075	1081
Nug20C	2570	2292	2396	2335
Nug30C	6124	5424	5648	–

Instances for which symmetry is destroyed

	n	Sol.	GLB	(ir)	μ_{R1}	(ir)	μ_{R2}	(ir)	μ_{R3}
Nug12	12	578	493		486		530		547
Nug12.1	11	586	496	(0.6)	514	(5.4)	551	(3.9)	574
Nug12.2	11	578	495	(0.4)	514	(5.4)	553	(4.3)	571
Nug12.5	11	578	494	(0.2)	524	(7.8)	552	(4.1)	570
Nug12.6	11	586	499	(1.2)	530	(9.0)	561	(5.8)	578
Nug15	15	1150	963		1009		1060		1075
Nug15.1	14	1150	967	(0.4)	1049	(4.0)	1104	(4.2)	1114
Nug15.2	14	1166	974	(1.1)	1076	(6.6)	1124	(6.0)	1135
Nug15.3	14	1200	987	(2.5)	1075	(6.5)	1133	(6.9)	1145
Nug15.6	14	1152	968	(0.5)	1056	(4.7)	1106	(4.3)	1118
Nug15.7	14	1166	979	(1.7)	1052	(4.3)	1112	(5.2)	1125
Nug15.8	14	1168	983	(2.1)	1063	(5.4)	1118	(5.5)	1133
Nug20	20	2570	2057		2281		2385		2326
Nug20.1	19	2628	2082	(1.2)	2358	(3.4)	2449	(2.7)	2414
Nug20.2	19	2600	2130	(3.6)	2401	(5.3)	2489	(4.4)	2464
Nug20.3	19	2588	2067	(0.5)	2331	(2.2)	2428	(1.8)	2392
Nug20.6	19	2570	2064	(0.3)	2341	(2.6)	2431	(1.9)	2403
Nug20.7	19	2634	2105	(2.3)	2387	(4.6)	2479	(3.9)	2436
Nug20.8	19	2636	2127	(3.4)	2381	(4.4)	2485	(4.2)	2438

Results for first level in branching tree