Semidefinite Relaxations for Hard Combinatorial Problems

Henry Wolkowicz

Department of Combinatorics & Optimization
University of Waterloo

OUTLINE

- Introduction (short) to SDP

(geometry, duality, algorithms)

- Max-Cut Problem instance

(quadratic boolean programming)
tractable relaxations versus Lagrangian duality
equivalence of various bounds

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

SDPs equivalent to min-max eigenvalue problems

- Theme: - Tractable relaxation implies Lagrangian relaxation which can be solved efficiently using SDP.

- Advantages of Lagrangian approach (recipe for SDP relaxation).

INTRODUCTION

Semidefinite programming – SDP
extension of linear programming – LP

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

$C, X \in \mathcal{S}^n$, space of symmetric $n \times n$ matrices
inner product $C \bullet X = \mathrm{trace}\, \mathrm{CX}$
$A \succeq B$ if $A - B \succeq 0$ positive semidefinite
$\mathcal{A} : \mathcal{S}^n \to \Re^m$, linear operator: $(\mathcal{A}X)_i = \mathrm{trace}\,(\mathrm{A_i X})$
$A_i \in \mathcal{S}^n, \; i = 1, \ldots m.$

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 $X, Y \in \operatorname{int}(\mathcal{P})$, 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

Extensions 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.)

primal program

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

dual program

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

weak duality (using hidden constraints)

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

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

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

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.

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

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)

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

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

---

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

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

where e vector of ones.

Bound 0: trivial bound from diagonal perturbations:

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

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:

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

Using the hidden semidefinite constraint:

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

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

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

(adding tractable constraint – equivalence to other bounds)

Bound B_1^c: lift to eigenvalue bound:

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

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

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

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

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

Similarly, equivalence to bounds B_0^c and other homogenized bounds.

Bound B_3: SDP bound:

After homogenization ($c = 0$), use

$$x^T Qx = \mathrm{trace\, x^T Qx} = \mathrm{trace\, Qxx^T}$$

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.

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

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

$$\begin{aligned}
B_0 &= \min_u \max_x q_u(x) \\
B_1 &= \min_u \max_{x^T x = n} q_u(x) \\
B_2 &= \min_u \max_{-1 \le x_i \le 1} q_u(x) \\
B_3 &= \max\{\mathrm{trace}\, \mathrm{Q^c Y} : \mathrm{diag}\,(\mathrm{Y}) = \mathrm{e},\ \ \mathrm{Y} \succeq 0.\} \\
B_1^c &= \min_u (n+1)\lambda_{\max}(Q^c + \mathrm{Diag}\,(\mathrm{u})) - \mathrm{u^T e}
\end{aligned}$$

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

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

B_L denotes Lagrangian relaxation bound. Following theme:

Theorem (PRW) B_L equals all above bounds.

(proofs from – strong Lagrangian duality of TRS.)

(question: which relaxation is best numerically, e.g. SDP in dense case, eigenvalue in large sparse case)

Strengthened SDP Bounds for MC

(Anjos-Wolkowicz/99)

Illustration of recipe for SDP relaxation.

(some parts similar to S-procedure Yakubovich/71, Shor relaxation/87)

- replace $Ax = b$ by $||Ax - b||^2 = 0$ and homogenize
- add redundant quadratic constraints
- take Lagrangian dual
- take dual of Lagrangian dual (use hidden SDP constraint)
- eliminate redundant constraints

motivation for $||Ax - b||^2 = 0$:

$0 = \max\{x^2 : x = 0\}$ primal optimal value

$0 < \infty = \inf_{\lambda} \max_{x} x^2 + \lambda x$ Lagrangian dual optimal value

BUT:

replace linear constraint by $x^2 = 0$

Then: Lagrangian dual is

$$0 = \inf_{\lambda} \max_{x} x^2 + \lambda x^2$$

the trick

$$\text{replace } Ax = b \text{ by } ||Ax - b||^2 = 0$$

works in general for most applications

Theorem (PRW) Let $K \subset \Re^n$ be a **finite set.** Let $q(x) = x^T Qx - 2c^T x$, $A \in \Re^{m \times n}$ and $b \in \Re^n$. Then there exists $\bar{\lambda} \in \Re$ such that $\forall\ \lambda \geq \bar{\lambda}$

$$\max_{x \in K} \{q(x) : ||Ax - b||^2 = 0\} = \max_{x \in K} \{q(x) - \lambda ||Ax - b||^2\}.$$

$\blacksquare$

i.e. quadratic penalty function is exact in the case that K is a finite set (no need for e.g. augmented Lagrangian) first lifting procedure

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

implies second lifting procedure

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

and provides an equivalent quadratic matrix model for MC

MCQ is equivalent to:

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

which is equivalent to

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

Since X and X^2 can be mutually diagonalized.

Now add redundant constraints;
and replace $\mathrm{diag}\,(\mathrm{X}) = \mathrm{e}$ with $||\mathrm{diag}\,(\mathrm{X}) - \mathrm{e}||_{\mathrm{F}}^{2} = 0$

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

where E is the matrix of ones.

Notation:

(Interesting operators and adjoints)

- $S \in \mathcal{S}^n, \quad t(n) = \frac{n(n+1)}{2}$
- $s = \operatorname{svec}(S) \in \Re^{t(n)}$ vector formed (columnwise) from S

(ignoring strictly lower triang.)

- $S = \operatorname{sMat}(s)$ inverse of svec
- $\operatorname{hMat}(v)$ adjoint of svec , off-diagonal terms *divided* by 2
- $\operatorname{dsvec}(S)$ adjoint of sMat , off diagonal elements *multiplied* by 2
- $\operatorname{sdiag}(s) := \operatorname{diag}(\operatorname{sMat}(\mathrm{s}))$
- $\operatorname{vsMat}(s) := \operatorname{vec}(\operatorname{sMat}(s))$

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

Summary:

$$\begin{aligned}
\operatorname{diag}^* &= \operatorname{Diag} \\
\operatorname{svec}^* &= \operatorname{hMat} \\
\operatorname{svec}^{-1} &= \operatorname{sMat} \\
\operatorname{dsvec}^* &= \operatorname{sMat} \\
\operatorname{vsMat}^* &= \operatorname{dsvec}\left[\tfrac{1}{2}\left(\operatorname{Mat}(\cdot) + \operatorname{Mat}(\cdot)^T\right)\right].
\end{aligned}$$

homogenize $1 - y_0^2 = 0$.

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

$$\begin{array}{rc} \max & \mathrm{trace}\ (\mathrm{Q}\,\mathrm{sMat}\,(\mathrm{x}))\,\mathrm{y}_0 \\ \text{s.t.} & \mathrm{sdiag}\,(x)^T\mathrm{sdiag}\,(x) - 2e^T\mathrm{sdiag}\,(x)y_0 + n = 0 \\ & \mathrm{sMat}\,(x) \circ \mathrm{sMat}\,(x) = E \\ & \mathrm{sMat}\,(x)^2 - n\,\mathrm{sMat}\,(x)y_0 = 0 \\ & 1 - y_0^2 = 0. \end{array}$$

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

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

constant part (no Lagrange multipliers) of Lagrangian Hessian:

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

negative nonconstant part of Lagrangian Hessian

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

Use **hidden SDP constraint**

Lagrangian dual:

$$\text{MCDSDP2} \qquad \begin{array}{rll} \nu_2^* = & \min & nw + \operatorname{trace} \mathrm{ET} + \operatorname{trace} 0\mathrm{S} + \mathrm{t} \\ & \text{s.t.} & \mathcal{H}(w, T, S, t) \succeq H_c \end{array}$$

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

take dual of this SDP – strengthened SDP relaxation of MC:

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

we need to calculate the adjoint operators and remove redundant constraints in MCPSDP2.

we use

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

simplified SDP relaxation MCPSDP2

$$
\begin{array}{lll}
\nu_2^* = & & \\
\max & \operatorname{trace} \mathrm{H_c Y} & \\
\text{s.t.} & \operatorname{diag}(\mathrm{Y}) = \mathrm{e} & \\
& Y_{0,t(i)} = 1, i = 1, \ldots, n & \\
& Y_{0,T(i,j)} = \frac{1}{n} \sum_{k=1}^{n} Y_{T(i,k),T(k,j)}, & \forall\, i, j \text{ s.t. } 1 \leq i < j \leq n \\
& Y \succeq 0, Y \in \mathcal{S}^{t(n)+1}, &
\end{array}
$$

where

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

This problem has $2t(n) - 1$ constraints; but very sparse.

if dual program is

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

Helmberg-Rendl/2000: if $\operatorname{trace} \mathrm{X} = \text{const}$ (is constant) for all primal feasible X,

$$\mathcal{A}X = b \quad \textbf{implies} \quad \operatorname{trace} \mathrm{X} = \text{const} \qquad (*)$$

then

$$\nu^* = \min b^T y + \text{const} \lambda_{\max}(C - \mathcal{A}^* y)$$

Moreover, (*) holds iff $I = \mathcal{A}^* w$, for some w. But, this holds if original QQP is bounded; add redundant constraint $||x||^2 \leq K$. So – apply min-max eigenvalue approach in large sparse case.

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

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

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

Theorem The optimal values satisfy

$$\nu_2^* \leq \nu^* \quad \text{and} \quad \nu_2^* = \nu^* \Rightarrow \nu_2^* = \mu^*.$$

(Unless MC optimum if sound, the new bound is strictly better.)

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.

From

$$Y_{0,T(i,j)} = \frac{1}{n} \sum_{k=1}^{n} Y_{T(i,k),T(k,j)}, \quad \forall\, 1 \le i < j \le n,$$

average = 1 implies each = 1

$$\text{(MCPSDP3)} \qquad \begin{array}{rl} \nu_3^* = & \\ \max & \operatorname{trace} \mathrm{H_c Z} \\ \text{s.t.} & \operatorname{diag}(\mathrm{Z}) = \mathrm{e} \\ & Z_{0,t(i)} = 1, i = 1, \ldots, n \\ & Z_{0,T(i,j)} = Z_{T(i,k),T(k,j)}, \quad \forall\, k, \forall\, 1 \le i < j \le n \\ & Z \succeq 0, Z \in \mathcal{S}^{t(n)+1}. \end{array}$$

SDP3 has $(n-1) \cdot t(n-1) + 2n + 1$ equality constraints.

n	MC opt value	SDP1 bound	SDP2 bound	SDP1 plus all triangle inequalities	SDP3 bound	Graph
5	4	4.5225 $\rho = 0.8845$ R.E.: 13.06%	4.2889 $\rho = 0.9326$ R.E.: 7.22%	4.0000 $\rho = 1.0000$ R.E.: 0%	4.0000 $\rho = 1.0000$ R.E.: 0%	5-cycle with unit edge weights
5	6	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.1160 $\rho = 0.9810$ R.E.: 1.93%	6.0000 $\rho = 1.0000$ R.E.: 0%	6.0000 $\rho = 1.0000$ R.E.: 0%	$K_5 \backslash e$ with unit edge weights
5	6	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.2500 $\rho = 0.9600$ R.E.: 4.17%	6.2500 $\rho = 0.9600$ R.E.: 4.17%	K_5 with unit edge weights
5	9.28	9.6040 $\rho = 0.9663$ R.E.: 3.49%	9.4056 $\rho = 0.9866$ R.E.: 1.35%	9.2961 $\rho = 0.9983$ R.E.: 0.17%	9.2800 $\rho = 1.0000$ R.E.: 0%	Given by $A(G)$ below
10	12	12.5 $\rho = 0.9600$ R.E.: 4.17%	12.3781 $\rho = 0.9695$ R.E.: 3.15%	12.0000 $\rho = 1.0000$ R.E.: 0%	12.0000 $\rho = 1.0000$ R.E.: 0%	Petersen with unit edge weights
12	88	90.3919 $\rho = 0.9735$ R.E.: 2.72%	89.5733 $\rho = 0.9824$ R.E.: 1.79%	88.0029 $\rho = 1.0000$ R.E.: $3.3e-5$	88.0000 $\rho = 1.0000$ R.E.: $9.9e-7$	Randomly generated

Other illustration of Lagrangian duality strength:

Trust Region Subproblem

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

(more general nonconvex constraints $\alpha \leq q(x) \leq \beta$.)

$$\text{DTRS} \qquad \nu^* := \max_{\lambda \geq 0} \min_x \ q_0(x) + \lambda(x^T x - \delta^2).$$

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

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

Again our theme holds: *tractable implies strong duality* short proof of strong duality (e.g. A. Lewis)

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

$$
\begin{aligned}
\mu^* &= \min_{x^Tx\le\delta^2} && x^T(Q_0-\gamma I)x - 2c_0^Tx + \gamma x^Tx \\
&= \min_{x^Tx=\delta^2} && x^T(Q_0-\gamma I)x - 2c_0^Tx + \gamma x^Tx, && (Q_0 \text{ is indefi} \\
&= \min_{x^Tx=\delta^2} && x^T(Q_0-\gamma I)x - 2c_0^Tx + \gamma\delta^2 \\
&= \min_{x^Tx\le\delta^2} && x^T(Q_0-\gamma I)x - 2c_0^Tx + \gamma\delta^2, && (Q_0 - \gamma I \text{ is} \\
&= \max_{\lambda\ge 0}\min_{x} && x^T(Q_0-\gamma I)x - 2c_0^Tx + \lambda(x^Tx-\delta^2) + \gamma\delta^2 && (\text{convex case} \\
&= \max_{\lambda\ge 0}\min_{x} && x^TQ_0x - 2c_0^Tx + (\lambda-\gamma)(x^Tx-\delta^2) \\
&\le \max_{\lambda\ge \gamma}\min_{x} && x^TQ_0x - 2c_0^Tx + (\lambda-\gamma)(x^Tx-\delta^2) && (\gamma < 0) \\
&= \nu^* \le \mu^*.
\end{aligned}
$$

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

$$\operatorname{trace} \mathrm{PY} \leq \delta.$$

Second lifting:

$$Y P Y \preceq \delta Y.$$

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

$$\begin{array}{lrll} \mathrm{QQP_O} & \mu^* := & \min & \mathrm{trace\,AXBX^T} \\ & & \text{s.t.} & XX^T = I. \end{array}$$

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

Proof using first application of Lagrange multipliers:

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

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

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

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

A simple 2×2 Example $\mu^* = 10$

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad B := \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \quad B \otimes A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

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

However, add redundant constraint $X^T X = I$. Get $T \otimes I$ in Hessian and $-\mathrm{trace}\,\mathrm{T}$ 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.

Add redundant constraints $X^TX = I$.

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

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

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) add linear constraint to QAP ???? Use equivalence to min-max eig problem after homogenization to get zero duality gap?

Linear Case $A = B = 0$.

$$\begin{array}{lllr} \mathrm{QQPL}_{\mathcal{O}} \qquad \mu_L^{\mathcal{O}} := & \min & \mathrm{trace} - 2\mathrm{CX}^{\mathrm{T}} & \\ & \text{s.t.} & XX^T = I. & \end{array} \tag{2}$$

Characterize the optimal solution using singular values:

Proposition Suppose we have the singular value decomposition $C = U\Sigma V^T$, where the singular values, σ_i, are in the diagonal matrix Σ, and U, V are orthogonal matrices. Then the optimal value of $\mathrm{QQPL}_{\mathcal{O}}$ is $\mu_L^{\mathcal{O}} = -2\mathrm{trace}\,\Sigma = -2\sum_{\mathrm{i}=1}^{\mathrm{n}} \sigma_{\mathrm{i}}$. The optimal solution is obtained using the orthogonal matrices that yield the decomposition, i.e. $X^* = UV^T$. ■

dual of pure linear case: (can assume S, T are diagonal)

$$\begin{array}{lll} \text{LD} & \max & e^T s + e^T t + w \\ & \text{s.t.} & s_j + t_i \leq -\dfrac{1}{|w|}|c_{(j-1)n+i}|^2, \quad i, j = 1, \ldots, n. \end{array}$$

notice independence of the sign of the individual elements of C.

But:

Example

$$\begin{array}{rll} \mu^* := & \min & \text{trace } -2\mathrm{CX}^{\mathrm{T}} \\ & \text{s.t.} & XX^T = I,\; X^TX = I, \end{array} \tag{3}$$

with 2 × 2 matrix

$$C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

We then solve this example with the sign changed on -1, i.e.

$$C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The dual problem in the second case is

$$(D_{\mathcal{O}}) \quad \begin{array}{ll} \max & \mathrm{trace}\,\mathrm{S} + \mathrm{trace}\,\mathrm{T} + \mathrm{w} \\ \text{s.t.} & \mathrm{B}^0\mathrm{Diag}\,(S) + \mathrm{O}^0\mathrm{Diag}\,(T) + wE_{00} \preceq L_Q, \end{array}$$

where

$$L_Q := \begin{bmatrix} 0 & -1 & -1 & -1 & -1 \\ -1 & & & & \\ -1 & & \text{\Huge 0} & & \\ -1 & & & & \\ -1 & & & & \end{bmatrix}. \tag{4}$$

Dual optimal value does not change. But primal does, since the sum of the singular values of the two matrices are: $2\sqrt{2}$ in the first instance and just 2 for the symmetric C, i.e. the optimal values are $-4(\sqrt{2})$ and -4, respectively. (Implies duality gap in second case.)

Add redundant variables:

Theorem Strong duality holds between the pure linear case and its Lagrangian dual if the following equivalent problem for the sum of the singular values is used.

$$\text{SVD} \qquad \begin{array}{rll} \sum_i \sigma_i(C) = \max & & 2\mathrm{trace\,YCX^T} \\ & \text{s.t.} & XX^T + YY^T = I. \end{array}$$