Strong Duality and Stability in Conic Convex Optimization

Henry Wolkowicz
University of Waterloo

Joint work with: Simon Schurr and Levent Tunçel

Monday, January 21, 2008
at: AdvOL McMaster University

Outline

Cone Optimization

Primal-Dual Pair of Optimization Problems in Conic Form

$$(\textit{finite}) \quad v_P = \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_K c\}, \qquad (\mathcal{P})$$

$$v_D = \inf_x \{\langle c, x\rangle \ : \ \mathcal{A}x = b, \ x \succeq_{K^*} 0\}. \qquad (\mathcal{D})$$

where

- $\mathcal{A}$ - an onto linear transformation; adjoint is $\mathcal{A}^*$
- K - a proper convex cone with dual/polar cone $K^* = \{x : \langle x, z\rangle \geq 0, \ \forall z \in K\}$.
- $z' \preceq_K z''(z' \prec_K z'')$ - partial order, $z'' - z' \in K(\in \mathrm{int} K)$

Cone Optimization

Primal-Dual Pair of Optimization Problems in Conic Form

$$(\textit{finite}) \qquad v_P = \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_K c\}, \qquad (\mathcal{P})$$

$$v_D = \inf_x \{\langle c, x\rangle \ : \ \mathcal{A} x = b, \ x \succeq_{K^*} 0\}. \qquad (\mathcal{D})$$

where

- $\mathcal{A}$ - an onto linear transformation; adjoint is $\mathcal{A}^*$
- K - a proper convex cone with dual/polar cone $K^* = \{x : \langle x, z\rangle \geq 0, \ \forall z \in K\}$.
- $z' \preceq_K z''(z' \prec_K z'')$ - partial order, $z'' - z' \in K(\in \mathrm{int} K)$

Semidefinite Programming, SDP

SDP

$$v_P = \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_K c\}, \qquad (\mathcal{P})$$

$$v_D = \inf_x \{\langle c, x\rangle \ : \ \mathcal{A}x = b, \ x \succeq_{K^*} 0\}. \qquad (\mathcal{D})$$

For SDP, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$, $b \in \mathbb{R}^m$, $c \in \mathbb{S}^n$, and $K = K^* = \mathbb{S}^n_+ := \{X \in \mathbb{S}^n : X \text{ is PSD}\}$.

Faces of Cones

Face

A convex cone F is a face of K, denoted $F \trianglelefteq K$, if

$$x, y \in K \text{ and } x + y \in F \implies x, y \in F.$$

If $F \trianglelefteq K$ and $F \neq K$, write $F \triangleleft K$.

Conjugate Face

If $F \trianglelefteq K$, the conjugate face (or complementary face) of F is

$$F^c := F^\perp \cap K^* \trianglelefteq K^*.$$

If $x \in \mathrm{ri}(F)$, then $F^c = \{x\}^\perp \cap K^*$.

Minimal Face (Minimal Cone)

Feasible set of ($\mathcal{P}$)

Let $\mathcal{F}_P := \{y : c - \mathcal{A}^* y \succeq_K 0\}$

Minimal Face

Assuming that $\mathcal{F}_P$ is nonempty, the minimal face (or minimal cone) of ($\mathcal{P}$) is

$$f_P := \bigcap \{F \trianglelefteq K : c - \mathcal{A}^*(\mathcal{F}_P) \subset F\}.$$

i.e., the minimal face that contains all the feasible slacks.

Minimal Face (Minimal Cone)

Feasible set of $(\mathcal{P})$

Let $\mathcal{F}_P := \{y : c - \mathcal{A}^* y \succeq_K 0\}$

Minimal Face

Assuming that $\mathcal{F}_P$ is nonempty, the minimal face (or minimal cone) of $(\mathcal{P})$ is

$$f_P := \bigcap \{F \trianglelefteq K : c - \mathcal{A}^*(\mathcal{F}_P) \subset F\}.$$

i.e., the minimal face that contains all the feasible slacks.

SDP Example from Ramana, 1995

Primal SDP

$$0 = v_P = \sup_y \left\{ y_2 : \begin{pmatrix} y_2 & 0 & 0 \\ 0 & y_1 & y_2 \\ 0 & y_2 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

$$y^* = \begin{pmatrix} y_1^* & 0 \end{pmatrix}^T, \quad y_1^* \leq 0, \quad Z^* = c - \mathcal{A}^* y^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -y_1^* & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Slater's CQ fails for primal

Dual of SDP Example

Dual Program

$$1 = v_D = \inf_X \{X_{11} \ : \ X_{22} = 0, \ X_{11} + 2X_{23} = 1, \ X \succeq 0\}$$

$$X^* = \begin{pmatrix} 1 & 0 & X_{13} \\ 0 & 0 & 0 \\ X_{13} & 0 & X_{33} \end{pmatrix}, \quad X_{33} \geq (X_{13}^2)$$

Slater's CQ for (primal) dual & complementarity fails

duality gap $v_D - v_P = 1 - 0 = 1$,

$$\text{trace}\, X^* Z^* = \text{trace} \begin{pmatrix} 1 & 0 & X_{13} \\ 0 & 0 & 0 \\ X_{13} & 0 & X_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -y_1^* & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1 > 0$$

Minimal Face for Ramana Example

Feasible Set/Minimal Face

$\mathcal{F}_P = \{y \in \mathbb{R}^2 : y_1 \leq 0,\ y_2 = 0\}$

$$\begin{array}{rcl} f_P & = & \bigcap\{F \trianglelefteq K : c - \mathcal{A}^*(\mathcal{F}_P) \subset F\} \\ & = & \begin{pmatrix} \mathbb{S}^2_+ & 0 \\ 0 & 0 \end{pmatrix} \\ & \lhd & \mathbb{S}^3_+ \end{array}$$

Slater CQ and Minimal Face

If $(\mathcal{P})$ is feasible, then

$$c - \mathcal{A}^* y \not\succ_K 0\ \forall y \text{ (Slater's CQ fails for } (\mathcal{P}) \text{) } \iff f_P \lhd K$$

Regularization of $(\mathcal{P})$

Borwein-W (1981)

If v_P is finite, then $(\mathcal{P})$ is equivalent to regularized $(\mathcal{P})$

$$v_{RP} = \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_{f_P} c\}. \qquad \text{(RP)}$$

Lagrangian Dual DRP Satisfies Strong Duality:

$$v_P = v_{RP} = v_{DRP} = \inf_x \{\langle c, x\rangle \ : \ \mathcal{A}x = b, \ x \succeq_{f_P^*} 0\} \qquad \text{(DRP)}$$

and v_{DRP} is attained

Implementation Problems with Regularization

Difficulties

Borwein and W. also gave an algorithm to compute f_P.
But Difficulties:

1. The algorithm requires the solution of several (homogeneous) cone programs (constraints are: $\mathcal{A}x = 0, \langle c, x\rangle = 0, 0 \neq x \succeq_K 0$)
2. If Slater's CQ fails for $(\mathcal{D})$, then it also fails for each of these cone programs.

Ramana's Strong Dual for SDP

Ramana '95: Extended Lagrange-Slater dual (ELSD) for $(\mathcal{P})$

Construction of this dual takes advantage of the well understood facial structure of $\mathbb{S}^n_+$.

Advantages:

1. ELSD is explicit in terms of original data $(\mathcal{A}, b, c)$
2. ELSD is poly. size (# vrbles is (kn^2), $k \leq \min\{m, n\}$)

Disdvantages:

1. Slater's CQ may fail for ELSD and its Lagrangian dual.
2. ELSD can potentially be very large.

Our Goals:

Equivalence in the case of SDP

Ramana, Tunçel, and W. ’97: Ramana’s ELSD is equivalent to (DRP) (dual of regularized primal of Borwein and W.)
(Both approaches may require solution of potentially large SDPs that need not satisfy Slater’s CQ.)

Goals: Derive an Algorithm that Satisfies

1. recognizes if Slater’s CQ holds and if $(\mathcal{P})$–$(\mathcal{D})$ has a zero duality gap
2. size of any intermediate cone program solved does not exceed that of $(\mathcal{P})$ or $(\mathcal{D})$
3. intermediate cone programs to be solved are well behaved

Our Goals:

Equivalence in the case of SDP

Ramana, Tunçel, and W. '97: Ramana's ELSD is equivalent to (DRP) (dual of regularized primal of Borwein and W.)
(Both approaches may require solution of potentially large SDPs that need not satisfy Slater's CQ.)

Goals: Derive an Algorithm that Satisfies

1. recognizes if Slater's CQ holds and if $(\mathcal{P})$–$(\mathcal{D})$ has a zero duality gap
2. size of any intermediate cone program solved does not exceed that of $(\mathcal{P})$ or $(\mathcal{D})$
3. intermediate cone programs to be solved are well behaved

Theorem of the Alternative for Slater's CQ

THEOREM

Suppose that $(\mathcal{P})$ is feasible. Then exactly one of the following two systems is consistent:

(1) $\mathcal{A}x = 0,\ \langle c, x\rangle = 0,$ and $0 \neq x \succeq_{K^*} 0$

(2) $\mathcal{A}^* y \prec_K c$ (Slater's CQ holds for $(\mathcal{P})$)

Difficult?

In theory, we can solve $\min\{0 : x \text{ satisfies } (1)\}$ to determine if Slater's CQ fails for $(\mathcal{P})$.

But this problem need not satisfy the generalized Slater CQ.

So how can we solve (1)?

Theorem of the Alternative for Slater's CQ

THEOREM

Suppose that $(\mathcal{P})$ is feasible. Then exactly one of the following two systems is consistent:

(1) $\mathcal{A}x = 0,\ \langle c, x\rangle = 0,$ and $0 \neq x \succeq_{K^*} 0$
(2) $\mathcal{A}^* y \prec_K c$ (Slater's CQ holds for $(\mathcal{P})$)

Difficult?

In theory, we can solve $\min\{0 : x \text{ satisfies } (1)\}$ to determine if Slater's CQ fails for $(\mathcal{P})$.
But this problem need not satisfy the generalized Slater CQ.
So how can we solve (1)?

Stable Theorem of the Alternative

Stable Auxiliary Problem

Let $e \in \mathrm{int}(K) \cap \mathrm{int}(K^*)$; define $\mathcal{A}_c x := \begin{pmatrix} \mathcal{A}x \\ \langle c, x \rangle \end{pmatrix}$

$$\alpha^* := \left\{ \inf_{x,\alpha} \alpha : \mathcal{A}_c x = 0, x + \alpha e \succeq_{K^*} 0, \langle e, x \rangle \leq 1 \right\} \qquad (\mathcal{A})$$

Properties/Advantages

- size of $(\mathcal{A})$ essentially that of $(\mathcal{D})$
- A strictly feasible primal-dual point is easily found.
- Apply primal-dual IPM; assume a barrier for K^* such that the central path defined by it converges to a point in the relative interior of the optimal face; follow central path closely at end of algorithm.

Slater's Condition and the Auxiliary problem

Solution to $(\mathcal{A})$ yields info on $(\mathcal{P})$–$(\mathcal{D})$

Theorem: The x component of the central path for $(\mathcal{A})$ converges to a point in $\mathrm{ri}(\mathrm{face}(G_P))$, where

$$G_P := \{x \;:\; Ax = 0,\ \langle c, x\rangle = 0,\ x \succeq_{K^*} 0\}.$$

Moreover, since $f_P \subset \{x^*\}^\perp \cap K = [\mathrm{face}(G_P)]^c \trianglelefteq K$, one of the following holds:

1. $\alpha^* = 0$ and $x^* = 0$, so Slater's CQ holds for $(\mathcal{P})$, or
2. $\alpha^* = 0$ and $0 \neq x^* \succeq_{K^*} 0$, so $f_P \subset \{x^*\}^\perp \cap K \lhd K$, or
3. $\alpha^* < 0$ and $x^* \succ_{K^*} 0$, so the generalized Slater CQ holds for $(\mathcal{D})$.

Our Algorithm

Input: $\mathcal{A}, b, c, K$, and $\varepsilon > 0$

Compute an optimal α^* and $x^* \in \mathrm{ri}(\mathrm{face}(G_P))$ from $(\mathcal{A})$ using a primal-dual IPM.

While $\|x^*\| > \varepsilon$

If *$\alpha^* < 0$, then $x^* \succ_{K^*} 0$, $f_P = \{0\}$. Hence optimal y for $(\mathcal{P})$ satisfies $A^* y = c$; exit algorithm.*

Else

1. *$y \in \mathcal{F}_P$ implies $c - \mathcal{A}^* y \in [\mathrm{face}(G_P)]^c \lhd K$, get reduced primal with cone $K' = [\mathrm{face}(G_P)]^c$.*
2. *Replace (update) primal by reduced primal.*

End

Conclusion of Algorithm

Finish

Finally, solve reduced primal problem for which Slater's CQ holds.
(This provides a certificate of optimality.)

$\rightarrow$ For cones such as $\mathbb{S}^n_+$, auxiliary problems get progressively smaller.

Regularization for SDP

The (Conjugate) Faces, $\mathcal{F} \trianglelefteq \mathbb{S}^n_+$ are of the Form

$$\begin{aligned}\mathcal{F} &= \begin{pmatrix} P & Q \end{pmatrix} \begin{pmatrix} \mathbb{S}^r_+ & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \end{pmatrix}^T = P\mathbb{S}^r_+ P^T \\ \mathcal{F}^c &= \begin{pmatrix} P & Q \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{S}^{n-r}_+ \end{pmatrix} \begin{pmatrix} P & Q \end{pmatrix}^T = Q\mathbb{S}^{n-r}_+ Q^T\end{aligned}$$

where matrix $\begin{pmatrix} P & Q \end{pmatrix}$ is orthogonal.

The Minimal Face f_P Using the Auxiliary Problem

With $x^* \in \mathrm{ri}(G_P)$ from Auxiliary Problem

$$x^* = \begin{pmatrix} P & Q \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \end{pmatrix}^T,$$

$$\mathrm{face}(G_P) = \mathrm{face}(x^*) = \begin{pmatrix} P & Q \end{pmatrix} \begin{pmatrix} \mathbb{S}^r_+ & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P & Q \end{pmatrix}^T$$

Then: $f_P \trianglelefteq K' := [\mathrm{face}(G_P)]^c = Q\mathbb{S}^{n-r}_+ Q^T$.

WLOG: shift c and find linear transformation $\mathcal{L}'$

$c' \leftarrow c - \mathcal{A}^* y' \in K' - K'; \quad \mathcal{R}(\mathcal{A}^*\mathcal{L}') \subset K' - K'$

A Reduced Lower Dimensional Primal Problem

equivalent cone constraints: $\mathcal{A}^*\mathcal{L}'y' \preceq_{K'} c'$

$Q^T(A^*\mathcal{L}'y')Q \preceq_{\mathbb{S}^{n-r}_+} Q^T c' Q$

Previous SDP with $K = \mathbb{S}^3_+$ and a Duality Gap of 1

SeDuMi 1.1 Results

$$y^* = \begin{pmatrix} -0.321 \times 10^6 & 0.372 \end{pmatrix}^T$$
$$s^* = \begin{pmatrix} 0.628 \times 10^5 & 0 & 0 \\ 0 & -0.321 \times 10^6 & -0.372 \\ 0 & -0.372 & 0 \end{pmatrix};$$

desired accuracy (10^{-6}) achieved but!!

$\langle c, x^* \rangle - \langle b, y^* \rangle \approx -0.12$! and s^* is not pos. semidef.

After One Step of the Reduction

Our code yields correct primal solution:

$$y^* = \begin{pmatrix} -1.50 \\ 0 \end{pmatrix}, \quad s^* = \begin{pmatrix} 1.00 & 0 & 0 \\ 0 & 1.50 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Higher Dimensional Numerical Experiments

SDP with $m = n \geq 3$, $b = e_2$, $c = 0$

$$\mathcal{A}^* y = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_{n-1} & y_n \\ y_2 & y_3 & & & & \\ \vdots & & & \ddots & & \\ y_{n-1} & & & & y_n & \\ y_n & & & & & 0 \end{pmatrix}$$

SeDuMi/Our Algorithm

SeDuMi gives incorrect primal/dual solution; duality gap of -1; our algorithm gives correct solution

$\mathcal{F}_P = \{y \in \mathbb{R}^n : y_1 \leq 0,\ y_2 = \cdots = y_n = 0\}$

min. face $f_P = \{Z \in \mathbb{S}^n_+ : Z_{11} \geq 0,\ Z_{ij} = 0\ \forall (i,j) \neq (1,1)\}$,

and $(\mathcal{D})$ is infeasible.

$(\mathcal{P})$–$(\mathcal{D})$ in Symmetric Subspace Form

Symmetric Subspace Form

Let: $\bar{s} := c$; $A\bar{x} = b$; $\mathcal{L} = \mathrm{Nullspace}(\mathcal{A})$. Then:

$$v_P = \langle \bar{s}, \bar{x} \rangle \; - \inf_s \{ \langle \bar{x}, s \rangle \; : \; s \in (\bar{s} + \mathcal{L}^\perp) \cap K \}, \qquad (\mathcal{P}')$$

$$v_D = \inf_x \{ \langle \bar{s}, x \rangle \; : \; x \in (\bar{x} + \mathcal{L}) \cap K^* \}. \qquad (\mathcal{D}')$$

Recession Cone Feasibility Problems for $(\mathcal{P}')$ and $(\mathcal{D}')$:

$e \in \mathrm{int}(K) \cap \mathrm{int}(K^*)$; and $(0 \neq x^*, 0 = \alpha^*)$ soln to aux. prob.

$$\mathcal{S} := \{ s \in \mathcal{L}^\perp \cap K : \langle e, s \rangle = 1 \}, \qquad \text{(PRF)}$$

$$x^* / \langle e, x^* \rangle \in \mathcal{X} := \{ x \in \mathcal{L} \cap K^* : \langle e, x \rangle = 1 \}. \qquad \text{(DRF)}$$

Complementarity Partition

Symmetric Subspace Form

$$\mathcal{S} := \{s \in \mathcal{L}^\perp \cap K : \langle e, s\rangle = 1\}, \quad \mathcal{L}^\perp = \mathcal{R}(\mathcal{A}^*) \qquad \text{(PRF)}$$
$$\mathcal{X} := \{x \in \mathcal{L} \cap K^* : \langle e, x\rangle = 1\} \quad \mathcal{L} = \mathcal{N}(\mathcal{A}) \qquad \text{(DRF)}$$

Complementarity Partition for given $\mathcal{F} \unlhd K$:

$(\mathcal{F}, \mathcal{F}^c)$ is a complementarity partition if
$\mathrm{face}(\mathcal{S}) \subset \mathcal{F}$ and $\mathrm{face}(\mathcal{X}) \subset \mathcal{F}^c$;
it is a strict complementarity partition if also
$[\mathrm{face}(\mathcal{S})]^c = \mathrm{face}(\mathcal{X})$ (equiv. $[\mathrm{face}(\mathcal{S})]^c \cap [\mathrm{face}(\mathcal{X})]^c = \{0\}$);
it is proper if $\mathcal{S}$ and $\mathcal{X}$ are both nonempty.

Strict Complementarity and Nonzero Gaps

Theorem: Let K be a proper cone

(1) If (PRF)–(DRF) has a proper complementarity partition but not a strict complementarity partition, then there exists $\bar{s}$ and $\bar{x}$ such that $(\mathcal{P})$–$(\mathcal{D})$ with data $(\mathcal{L}, K, \bar{s}, \bar{x})$ has a finite nonzero duality gap.

(Partial Converse)

(2) If (a) $(\mathcal{P})$–$(\mathcal{D})$ with data $(\mathcal{L}, K, \bar{s}, \bar{x})$ has a finite nonzero duality gap with both optimal values attained, and (b) all feasible solutions of $(\mathcal{P})$ and $(\mathcal{D})$ are optimal, then (PRF)–(DRF) has a proper complementarity partition but not a strict complementarity partition.

Generating SDP Instances with nonzero gaps

$K = \mathbb{S}^n_+$ Instance

Choose positive integers n, p, d with $n > p + d$. Let $e = I_n \in \operatorname{int}(K) \cap \operatorname{int}(K^*)$.

Choose subspace $\mathcal{L}$ and Orthogonal Matrix Q

$$\operatorname{face}(\mathcal{L}^\perp \cap K) = Q \begin{pmatrix} 0 & & \\ & 0 & \\ & & \mathbb{S}^p_+ \end{pmatrix} Q^T, \operatorname{face}(\mathcal{L} \cap K^*) = Q \begin{pmatrix} \mathbb{S}^d_+ & & \\ & 0 & \\ & & 0 \end{pmatrix} Q^T.$$

These faces form a not strict complementarity partition

Choose a nonzero $U \in \mathbb{S}^{n-p-d}_+$

$$\bar{s} := \bar{x} := Q \begin{pmatrix} 0 & & \\ & U & \\ & & 0 \end{pmatrix} Q^T.$$

duality gap is $\langle \bar{s}, \bar{x} \rangle = \|U\|_F^2 > 0$.

Conclusion

Summary:

- presented a stable algorithm to solve (feasible) conic problems for which Slater's CQ fails;
- algorithm requires the solution of problems whose size is the same as that of the original dual; In special cases such as SDP and SOCP, these problems become progressively smaller;
- Failure of strict complementarity for the associated recession problems is closely related to the existence of instances having a finite nonzero duality gap; provides a means of generating instances for testing.

Work in Progress

Future:

- We intend to refine our code and test it on larger SDPs having a finite nonzero duality gap.
- Perform backward error analysis to study how rounding errors and errors in computing approximate solutions to the auxiliary problems affects the number of iterations of our algorithm.
- In particular, we want to reduce the problem when Slater's condition almost fails.

Auxiliary Problem for Distance to Infeasibility

Perturbed Auxiliary Problem

let $\mathcal{Q}$ denote the second order cone, SOC; relax the equality constraints $\mathcal{A}_c x = 0$ to SOC constraint $\|\mathcal{A}_c x\|_2 \leq \delta$.

$$
\begin{array}{rll}
v_P^{aux} := & \inf_{x,\delta} & \delta \\
& \text{s.t.} & \begin{pmatrix} \delta \\ \mathcal{A}_c x \end{pmatrix} \succeq_{\mathcal{Q}} 0 \\
& & \langle x, e \rangle = 1 \\
& & x \succeq_{K^*} 0.
\end{array}
$$

Similar nice properties; and, near failure of Slater's CQ is identified.