STRONG DUALITY FOR SEMIDEFINITE PROGRAMMING

(Linear Programming for the 90's and 00's)

Henry Wolkowicz
University of Waterloo

(work with Levent Tuncel and Motakuri Ramana)

Linear Programming ($A : \Re^m \to \Re^n$)

$\Re^n_+$ is a closed convex cone

$$(\mathbf{P}) \quad \begin{array}{rll} p^* = & \text{sup} & c^t x \\ & \text{s.t.} & Ax \preceq b \quad (b - Ax \in \Re^n_+) \\ & & x \in \Re^m \end{array}$$

Lagrangian (payoff function):
$L(x, U) = c^t x + \langle U, b - Ax \rangle$

$$p^* = \max_x \min_{U \succeq 0} L(x, U)$$

(the constraint $U \succeq 0$ is needed to recover the hidden constraint $Ax \preceq b$.)

The dual is obtained from the optimal strategy of the competing player

$$p^* \leq d^* = \min_{U \succeq 0} \max_x L(x, U) = \langle U, b \rangle + x^t (c - A^* U)$$

The hidden constraint $c - A^* U = 0$ yields the dual

$$(\mathbf{D}) \qquad \begin{array}{rcc} d^* = & \inf & \text{trace}\, bU \\ & \text{s.t.} & A^* U = c \\ & & U \succeq 0. \end{array}$$

for the primal

$$(\mathbf{P}) \qquad \begin{array}{rcc} p^* = & \sup & c^t x \\ & \text{s.t.} & Ax \preceq b \\ & & x \in \Re^m \end{array}$$

If Slater's condition fails for the primal LP, then there are an infinite number of different dual programs.

The implicit equality constraints are:

$$A_e x = b_e$$

where $A = \begin{bmatrix} A_l \\ A_e \end{bmatrix}$

$$\textbf{(D)} \qquad \begin{array}{rll} d^* = & \inf & \text{trace}\, bU \\ & \text{s.t.} & A_l^* U_l + A_e^* U_e = c \\ & & U \in \mathcal{U} \\ & & \{U : U \succeq 0\} \subset \mathcal{U} \\ & & \mathcal{U} \subset \{U : U_l \succeq 0, U_e \text{ free}\}. \end{array}$$

for the equivalent primal program

$$\textbf{(P)} \qquad \begin{array}{rll} p^* = & \sup & c^t x \\ & \text{s.t.} & A_l x \preceq b_l \\ & & A_e x = b_e \\ & & x \in \Re^m \end{array}$$

DUALITY THEOREM

1. If one of the problems is inconsistent, then the other is inconsistent or unbounded.

2. **WEAK DUALITY**
   Let the two problems be consistent, and let x^0 be a feasible solution for P and U^0 be a feasible solution for D. Then

$$c^t x^0 \leq \left\langle b, U^0 \right\rangle .$$

3. **STRONG DUALITY**
   If both P and D are consistent, then they have optimal solutions and their optimal values are equal.

4. **COMPLEMENTARY SLACKNESS**
Let x^0 and U^0 be feasible solutions of P and D, respectively. Then x^0 and U^0 are optimal if and only if

$$\left\langle U^0, (b - Ax^0) \right\rangle = 0.$$

if and only if

$$U^0 \circ (b - Ax^0) = 0.$$

5. **SADDLE POINT**
The vectors x^0, U^0 are optimal solutions of P and D, respectively, if and only if (x^0, U^0) is a saddle point of the Lagrangian $L(x, U)$ for all (x,U),

$$L(x, U^0) \leq L(x^0, U^0) \leq L(x^0, U)$$

and then

$$L(x^0, U^0) = c^t x^0 = \left\langle b, U^0 \right\rangle.$$

Characterization of optimality for the dual pair x, U

$$Ax \preceq b \qquad \text{primal feasibility}$$

$$A^*U = c \qquad \text{dual feasibility}$$

$$U \circ (Ax - b) = 0e \quad \text{complementary slackness}$$

$$U \circ (Ax - b) = \mu e \qquad \text{perturbed}$$

Forms the basis for:

primal simplex method
dual simplex method
interior point methods

What is SEMIDEFINITE PROGRAMMING?

Why use it?

Quadratic approximations are better than linear approximations. And, we can solve relaxations of quadratic approximations efficiently using semidefinite programming.

How does SDP arise from quadratic approximations?

Let

$$q_i(y) = \frac{1}{2} y^t Q_i y + y^t b_i + c_i, \ \ y \in \Re^n$$

$$(\mathbf{QQP}) \quad \begin{array}{rll} q^* = & \text{min} & q_0(y) \\ & \text{s.t.} & q_i(y) = 0 \\ & & i = 1, \ldots m \end{array}$$

Lagrangian:

$$\begin{array}{rcl} L(y,x) & = & \frac{1}{2} y^t (Q_0 - \sum_{i=1}^m x_i Q_i) y \\ & & +y^t (b_0 - \sum_{i=1}^m x_i b_i) \\ & & +(c_0 - \sum_{i=1}^m x_i c_i) \end{array}$$

$$q^* = \min_y \max_x L(y,x) \geq d^* = \max_x \min_y L(y,x).$$

homogenize

$$y_0 y^t (b_0 - \sum_{i=1}^m x_i b_i), \quad y_0^2 = 1.$$

$$
\begin{array}{rcll}
d^* & = & \max_x \min_y & L(y,x) \\
& = & \max_x \min\limits_{y_0^2=1} & \frac{1}{2} y^t (Q_0 - \sum_{i=1}^m x_i Q_i) y \quad (+t y_0^2) \\
& & & + y_0 y^t (b_0 - \sum_{i=1}^m x_i b_i) \\
& & & + (c_0 - \sum_{i=1}^m x_i c_i) \quad (-t)
\end{array}
$$

The hidden semidefinite constraint yields the semidefinite program, i.e. we get
$A : \Re^{m+1} \to \mathcal{S}_{n+1}$

$$
B = \begin{pmatrix} 0 & b_0^t \\ b_0 & Q_0 \end{pmatrix}, A \begin{pmatrix} t \\ x \end{pmatrix} = \begin{bmatrix} -t & \sum_{i=1}^m x_i b_i^t \\ \sum_{i=1}^m x_i b_i & \sum_{i=1}^m x_i Q_i \end{bmatrix}
$$

$$
B - A \begin{pmatrix} t \\ x \end{pmatrix} \succeq 0.
$$

The dual program is equivalent to the SDP (with $c_0 = 0$)

$$
\begin{array}{lrl}
 & d^* = \sup & -\sum_{i=1}^m x_i c_i - t \\
(\mathbf{D}) & \text{s.t.} & A \begin{pmatrix} t \\ x \end{pmatrix} \preceq B \\
 & & x \in \Re^m, t \in \Re
\end{array}
$$

As in linear programming, the dual is obtained from the optimal strategy of the competing player:

$$
\begin{array}{lrl}
 & d^* = \inf & \operatorname{trace} BU \\
(\mathbf{DD}) & \text{s.t.} & A^* U = \begin{pmatrix} -1 \\ -c \end{pmatrix} \\
 & & U \succeq 0.
\end{array}
$$

Example

If the primal is

$$
(\mathbf{P}) \quad
\begin{array}{rll}
p^* = & \sup & x_2 \\
& \text{s. t.} & \begin{bmatrix} x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
\end{array}
$$

then the dual is

$$
(\mathbf{D}) \quad
\begin{array}{rlc}
d^* = & \inf & \operatorname{trace} U_{11} \\
& \text{s. t.} & U_{22} = 0 \\
& & U_{11} + 2U_{23} = 1 \\
& & U \succeq 0.
\end{array}
$$

Then $p^* = 0 < d^* = 1$.

What is a proper duality theory?

Do duality gaps occur in practice?

Are there an infinite number of duals if Slater's condition fails?

the cone $T \subset K$ is a *face* of the cone K, denoted $T \lhd K$, if

$$x, y \in K,\ x + y \in T \Rightarrow x, y \in T.$$

Each face, $K \lhd \mathcal{P}$, is characterized by a subspace, $S \subset \Re^n$.

$$K = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S\}.$$

Moreover,

$$\operatorname{relint} K = \{X \in \mathcal{P} : \mathcal{N}(X) = S\}.$$

The complementary face of K is $K^c = K^\perp \cap \mathcal{P}$

$$K^c = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S^\perp\}.$$

Moreover,

$$\operatorname{relint} K^c = \{X \in \mathcal{P} : \mathcal{N}(X) = S^\perp\}.$$

the face K (respectively, K^c) is determined by the supporting hyperplane corresponding to any $X \in \operatorname{relint} K^c$ (respectively, $\operatorname{relint} K$);

and

$$XY = 0, \forall X \in K, Y \in K^c$$

The *minimal cone* of P is defined as

$$\mathcal{P}^f = \cap\{\text{faces of } \mathcal{P} \text{ containing } (b - A(F))\}.$$

Therefore, an equivalent program is the *regularized P program*

$$
\textbf{(RP)} \quad
\begin{array}{rll}
p^* = & \max & c^t x \\
& \text{s.t.} & Ax \preceq_{\mathcal{P}^f} b \\
& & x \in \Re^m.
\end{array}
$$

there exists x such that $b - Ax \in \operatorname{relint} \mathcal{P}^f$. (generalized Slater's constraint qualification) strong duality pair is RP and

$$
\textbf{(DRP)} \quad
\begin{array}{rll}
p^* = & \min & \operatorname{trace} bU \\
& \text{s.t.} & A^* U = c \\
& & U \succeq_{(\mathcal{P}^f)^+} 0.
\end{array}
$$

Find $\mathcal{P}^f$? Properties?

LEMMA 1
Suppose $\mathcal{P}^f \lhd K \lhd \mathcal{P}$. Then the system

$$A^*U = 0, U \succeq_{K^+} 0, \operatorname{trace} Ub = 0$$

is consistent only if

$$\text{the minimal cone } \mathcal{P}^f \subset \{U\}^\perp \cap K.$$

PROOF
Since $\operatorname{trace} U(Ax - b) = 0$, for all x, we get $A(F) - b \subset U^\perp$, i.e. $\mathcal{P}^f \subset \{U\}^\perp$.

□

LEMMA 2 (surprising)
Suppose that $K \lhd \mathcal{P}$. Then

$$\mathcal{P}^+ + K^\perp \text{ is closed.}$$

□

Define: $\mathcal{P}_0 := \mathcal{P}$ and

$$\mathcal{U}_1 := \{U \succeq_{(\mathcal{P}_0)^+} 0 : A^*U = 0, \operatorname{trace} Ub = 0\}$$

Choose $U_1 \in \mathcal{U}_1 \cap \operatorname{relint} \mathcal{U}_1^f$ (if 0 - **STOP**)

$$\mathcal{P}_1 := \mathcal{U}_1^c = \{U_1\}^\perp \cap \mathcal{P}_0 \lhd \mathcal{P}_0$$

$$(\mathbf{RP_1}) \quad \begin{array}{rll} p^* = & \max & c^t x \\ & \text{s.t.} & Ax \preceq_{\mathcal{P}_1} b \\ & & x \in \Re^m. \end{array}$$

———————— $p^* \le d_1^* \le d^*$ ————————-

$$(\mathbf{DRP_1}) \quad \begin{array}{rll} d_1^* = & \min & \operatorname{trace} bU \\ & \text{s.t.} & A^*U = c \\ & & U \succeq_{(\mathcal{P}_1)^+} 0. \end{array}$$

——- $(\mathcal{P}_1)^+ = (\mathcal{P} \cap \mathcal{P}_1)^+ = \mathcal{P} + (\mathcal{P}_1)^\perp$ ——-

$$(\mathbf{ELSD_1}) \quad \begin{array}{rll} d_1^* = & \min & \operatorname{trace} b(U + (W + W^t)) \\ & \text{s.t.} & A^*(U + (W + W^t)) = c \\ & & A^*U_1 = 0, \operatorname{trace} U_1 b = 0 \\ & & U \succeq 0, \quad \begin{bmatrix} I & W^t \\ W & U_1 \end{bmatrix} \succeq 0. \end{array}$$

We used

$$\mathcal{P}_1^{\perp} = \left\{ (W + W^t) : A^* U_1 = 0, \text{trace}\, U_1 b = 0, \begin{bmatrix} I & W^t \\ W & U_1 \end{bmatrix} \succeq 0 \right\}$$

and

$$U_1 \succeq W W^t \qquad \text{iff} \qquad \begin{bmatrix} I & W^t \\ W & U_1 \end{bmatrix} \succeq 0$$

$$\text{implies} \quad W = U_1 H, \text{for some matrix H}$$

$$
\begin{aligned}
\mathcal{U}_2 &:= \{U \succeq_{(\mathcal{P}_1)^+} 0 : A^*U = 0, \operatorname{trace} Ub = 0\} \\
&= \{U + Z : A^*(U+Z) = 0, \operatorname{trace} Ub = 0, \\
&\qquad U \succeq_{(\mathcal{P}_0)^+}, Z \in (\mathcal{P}_1)^\perp\}
\end{aligned}
$$

Choose $U_2 \in \mathcal{U}_2 \cap \operatorname{relint} \mathcal{U}_2^f$ (if 0 - **STOP**)

$$\mathcal{P}_2 := \mathcal{U}_2^c = \{U_2\}^\perp \cap \mathcal{P}_1 \lhd \mathcal{P}_1$$

$$
(\mathbf{RP_2}) \quad \begin{array}{rcc} p^* = & \max & c^t x \\ & \text{s.t.} & Ax \preceq_{\mathcal{P}_2} b \\ & & x \in \Re^m. \end{array}
$$

————————$p^* \le d_2^* \le d_1^* \le d^*$————————-

$$
(\mathbf{DRP_2}) \quad \begin{array}{rcc} d_2^* = & \min & \operatorname{trace} bU \\ & \text{s.t.} & A^*U = c \\ & & U \succeq_{(\mathcal{P}_2)^+} 0. \end{array}
$$

————$(\mathcal{P}_2)^+ = (\mathcal{P} \cap \mathcal{P}_2)^+ = \mathcal{P} + (\mathcal{P}_2)^\perp$————-

$$
(\mathbf{ELSD_2}) \quad \begin{array}{rcc} d_1^* = & \min & \operatorname{trace} b(U + (W + W^t)) \\ & \text{s.t.} & A^*(U + (W + W^t)) = c \\ & & A^*U_1 = 0, \operatorname{trace} U_1 b = 0 \\ & & A^*(U_2 + (W_1 + W_1^t)) = 0, \\ & & \operatorname{trace}(U_2 + (W_1 + W_1^t))b = 0 \\ & & U \succeq 0, \ \begin{bmatrix} I & W_1^t \\ W_1 & U_1 \end{bmatrix} \succeq 0 \\ & & \begin{bmatrix} I & W^t \\ W & U_2 \end{bmatrix} \succeq 0. \end{array}
$$

HOMOGENIZATION

Alternate view of optimality conditions

$$
\begin{array}{llll}
 & 0 = & \max & c^t x + t(-p^*) \\
(\mathbf{HP}) & & \text{subject to} & Ax + t(-b) + Z = 0 \\
 & & & w \in K = \Re^m \otimes \Re_+ \otimes \mathcal{P}
\end{array}
$$

This defines the objective, constraints, and variables:

$$
\begin{array}{c}
(= \langle a, w \rangle) \\
(Bw = 0) \\
\left(w = \begin{pmatrix} x \\ t \\ Z \end{pmatrix} \right)
\end{array}
$$

the feasible set is

$$F_H = \mathcal{N}(B) \cap K,$$

$$Bw = 0, w \in K \text{ implies } \langle a, w \rangle \leq 0.$$

Optimality conditions:

$$a = \begin{pmatrix} c \\ -p \\ 0 \end{pmatrix} \in -(\mathcal{N}(B) \cap K)^{+}.$$

$$\begin{pmatrix} -c \\ p \\ 0 \end{pmatrix} \in \overline{\mathcal{R}(B^{*}) + K^{+}},$$

WCQ - Weakest Constraint Qualification:

CLOSURE HOLDS

Conditions for closure:

If C, D are closed convex sets and the intersection of their recession cones is empty, then $D - C$ is closed.

$cone(F_H - K)$ is the whole space

(Slater's)

$$\exists \hat{x} \in F \text{ such that } A\hat{x} \prec b.$$

FIX: Find sets, T, to add to attain the closure. Equivalently, find sets, C, $C^+ = T$, to intersect with K to attain the closure since

$$(\mathcal{N}(A) \cap (K \cap C))^+ = \overline{\mathcal{R}(A^*) + K^+ + C^+}.$$