ORF 522: Lecture 15

The Homogeneous Self-Dual Method

Robert J. Vanderbei

November 14, 2013

Slides last edited at 9:08am on Friday 15th November, 2013

Operations Research and Financial Engineering, Princeton University
http://www.princeton.edu/∼rvdb

The Homogeneous Self-Dual Problem

Primal-Dual Pair

$$\begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b \\ & x \ge 0 \end{array} \qquad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \ge c \\ & y \ge 0 \end{array}$$

Homogeneous Self-Dual Problem

$$\begin{array}{lrrrl} \text{maximize} & 0 & & & \\ \text{subject to} & & -A^T y & +c\phi & \le 0 \\ & Ax & & -b\phi & \le 0 \\ & -c^T x & +b^T y & & \le 0 \\ & x, & y, & \phi & \ge 0 \end{array}$$

In Matrix Notation

$$
\begin{array}{ll}
\text{maximize} & 0 \\
\text{subject to} & \begin{bmatrix} 0 & -A^T & c \\ A & 0 & -b \\ -c^T & b^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ \phi \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\
& x, y, \phi \geq 0.
\end{array}
$$

HSD is self-dual (constraint matrix is skew symmetric).

HSD is feasible ($x = 0$, $y = 0$, $\phi = 0$).

HSD is homogeneous—i.e., multiplying a feasible solution by a positive constant yields a new feasible solution.

Any feasible solution is optimal.

If ϕ is a null variable, then either primal or dual is infeasible (see text).

Theorem 1

Let (x, y, ϕ) *be a solution to HSD. If* $\phi > 0$*, then*

- $x^* = x/\phi$ *is optimal for primal, and*
- $y^* = y/\phi$ *is optimal for dual.*

Proof.

x^* is primal feasible—obvious.

y^* is dual feasible—obvious.

Weak duality theorem implies that $c^T x^* \leq b^T y^*$.

3rd HSD constraint implies reverse inequality.

Primal feasibility, plus dual feasibility, plus no gap implies optimality.

Change of Notation

$$\begin{bmatrix} 0 & -A^T & c \\ A & 0 & -b \\ -c^T & b^T & 0 \end{bmatrix} \longrightarrow A \qquad \begin{bmatrix} x \\ y \\ \phi \end{bmatrix} \longrightarrow x \qquad n + m + 1 \longrightarrow n$$

In New Notation:

$$\begin{array}{ll} \text{maximize} & 0 \\ \text{subject to} & Ax + z = 0 \\ & \quad x, z \geq 0 \end{array}$$

More Notation

Infeasibility:	$\rho(x,z) = Ax + z$
Complementarity:	$\mu(x,z) = \frac{1}{n}x^T z$

Nonlinear System

$$
\begin{aligned}
A(x + \Delta x) + (z + \Delta z) &= \delta\rho(x,z) \\
(X + \Delta X)(Z + \Delta Z)e &= \delta\mu(x,z)e
\end{aligned}
$$

(Here, δ is some positive parameter.)

Linearized System

$$
\begin{aligned}
A\Delta x + \Delta z &= -(1-\delta)\rho(x,z) \\
Z\Delta x + X\Delta z &= \delta\mu(x,z)e - XZe
\end{aligned}
$$

Algorithm

Solve linearized system for $(\Delta x, \Delta z)$.

Pick step length θ.

Step to a new point:

$$\bar{x} = x + \theta \Delta x, \qquad \bar{z} = z + \theta \Delta z.$$

Even More Notation

$$\bar{\rho} = \rho(\bar{x}, \bar{z}), \qquad \bar{\mu} = \mu(\bar{x}, \bar{z})$$

Theorem 2

1. $\Delta z^T \Delta x = 0$.
2. $\bar{\rho} = (1 - \theta + \theta\delta)\rho$.
3. $\bar{\mu} = (1 - \theta + \theta\delta)\mu$.
4. $\bar{X}\bar{Z}e - \bar{\mu}e = (1 - \theta)(XZe - \mu e) + \theta^2 \Delta X \Delta Z e$.

Proof.

1. Tedious but not hard (see text).

2.

$$\begin{aligned}
\bar{\rho} &= A(x + \theta\Delta x) + (z + \theta\Delta z) \\
&= Ax + z + \theta(A\Delta x + \Delta z) \\
&= \rho - \theta(1 - \delta)\rho \\
&= (1 - \theta + \theta\delta)\rho.
\end{aligned}$$

3.

$$
\begin{aligned}
\bar{x}^T \bar{z} &= (x + \theta \Delta x)^T (z + \theta \Delta z) \\
&= x^T z + \theta (z^T \Delta x + x^T \Delta z) + \theta^2 \Delta x^T \Delta z \\
&= x^T z + \theta e^T (\delta \mu e - X Z e) \\
&= (1 - \theta + \theta \delta) x^T z.
\end{aligned}
$$

Now, just divide by n.

4.

$$
\begin{aligned}
\bar{X} \bar{Z} e - \bar{\mu} e &= (X + \theta \Delta X)(Z + \theta \Delta Z) e - (1 - \theta + \theta \delta) \mu e \\
&= X Z e - \mu e + \theta (X \Delta z + Z \Delta x + (1 - \delta) \mu e) + \theta^2 \Delta X \Delta Z e \\
&= (1 - \theta)(X Z e - \mu e) + \theta^2 \Delta X \Delta Z e.
\end{aligned}
$$

Neighborhoods of $\{(x,z) > 0 : x_1 z_1 = x_2 z_2 = \cdots = x_n z_n\}$

$$\mathcal{N}(\beta) = \{(x,z) > 0 : \|XZe - \mu(x,z)e\| \leq \beta\mu(x,z)\}$$

Note: $\beta < \beta'$ implies $\mathcal{N}(\beta) \subset \mathcal{N}(\beta')$.

Predictor-Corrector Algorithm

Odd Iterations–Predictor Step

Assume $(x,z) \in \mathcal{N}(1/4)$.

Compute $(\Delta x, \Delta z)$ using $\delta = 0$.

Compute θ so that $(\bar{x}, \bar{z}) \in \mathcal{N}(1/2)$.

Even Iterations–Corrector Step

Assume $(x,z) \in \mathcal{N}(1/2)$.

Compute $(\Delta x, \Delta z)$ using $\delta = 1$.

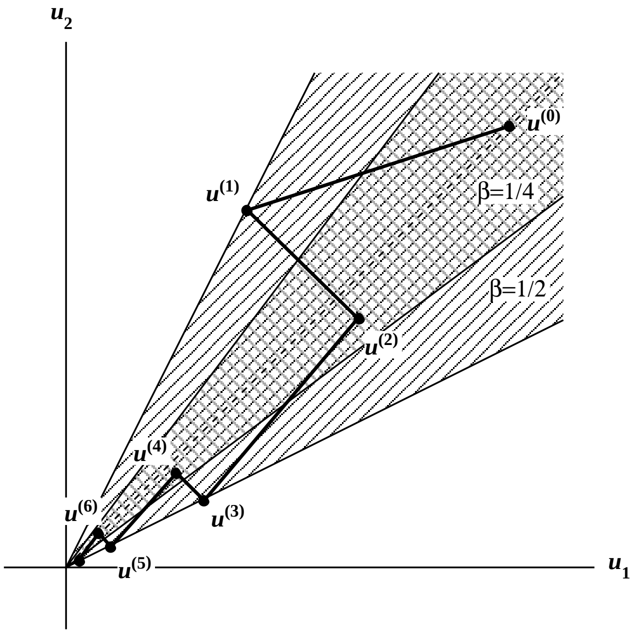
Put $\theta = 1$.

Predictor-Corrector Algorithm

In Complementarity Space

Let

$$u_j = x_j z_j \qquad j = 1, 2, \ldots, n.$$



Well-Definedness of Algorithm

Must check that preconditions for each iteration are met.

Technical Lemma.

1. If $\delta = 0$, then $\|\Delta X \Delta Z e\| \leq \frac{n}{2}\mu$.
2. If $\delta = 1$ and $(x, z) \in \mathcal{N}(\beta)$, then $\|\Delta X \Delta Z e\| \leq \frac{\beta^2}{1-\beta}\mu/2$.

Proof. Tedious **and** tricky. See text.

Theorem 3

1. *After a predictor step,* $(\bar{x}, \bar{z}) \in \mathcal{N}(1/2)$ *and* $\bar{\mu} = (1-\theta)\mu$.
2. *After a corrector step,* $(\bar{x}, \bar{z}) \in \mathcal{N}(1/4)$ *and* $\bar{\mu} = \mu$.

Proof.

1. $(\bar{x}, \bar{z}) \in \mathcal{N}(1/2)$ by definition of θ.

   $\bar{\mu} = (1-\theta)\mu$ since $\delta = 0$.

2. $\theta = 1$ and $\beta = 1/2$. Therefore,

$$\|\bar{X}\bar{Z}e - \bar{\mu}e\| = \|\Delta X \Delta Z e\| \leq \mu/4.$$

   Need to show also that $(\bar{x}, \bar{z}) > 0$. Intuitively clear (see earlier picture) but proof is tedious. See text.

Complexity Analysis

Progress toward optimality is controlled by the stepsize θ.

Theorem 4. *In predictor steps, $\theta \geq \frac{1}{2\sqrt{n}}$.*

Proof.

Consider taking a step with step length $t \leq 1/(2\sqrt{n})$:

$$x(t) = x + t\Delta x, \qquad z(t) = z + t\Delta z.$$

From earlier theorems and lemmas,

$$\begin{aligned}
\|X(t)Z(t)e - \mu(t)e\| &\leq (1-t)\|XZe - \mu e\| + t^2\|\Delta X \Delta Z e\| \\
&\leq (1-t)\frac{\mu}{4} + t^2\frac{n\mu}{2} \\
&\leq (1-t)\frac{\mu}{4} + \frac{\mu}{8} \\
&\leq (1-t)\frac{\mu}{4} + (1-t)\frac{\mu}{4} \\
&= \frac{\mu(t)}{2}.
\end{aligned}$$

Therefore $(x(t), z(t)) \in \mathcal{N}(1/2)$ which implies that $\theta \geq 1/(2\sqrt{n})$.

Since

$$\mu^{(2k)} = (1 - \theta^{(2k-1)})(1 - \theta^{(2k-3)}) \cdots (1 - \theta^{(1)})\mu^{(0)}$$

and $\mu^{(0)} = 1$, we see from the previous theorem that

$$\mu^{(2k)} \leq \left(1 - \frac{1}{2\sqrt{n}}\right)^k .$$

Hence, to get a small number, say 2^{-L}, as an upper bound for $\mu^{(2k)}$ it suffices to pick k so that:

$$\left(1 - \frac{1}{2\sqrt{n}}\right)^k \leq 2^{-L}.$$

This inequality is implied by the following simpler one:

$$k \geq 2\log(2)L\sqrt{n}.$$

Since the number of iterations is $2k$, we see that $4\log(2)L\sqrt{n}$ iterations will suffice to make the final value of μ be less than 2^{-L}.

Of course,

$$\rho^{(k)} = \mu^{(k)}\rho^{(0)}$$

so the same bounds guarantee that the final infeasibility is small too.

Theorem 5

Fix $\beta > 0$. There exists a constant $c > 0$ such that $x_j + z_j \geq c$ for all j and for all $(x, z) \in \mathcal{N}(\beta)$.

Proof. Let $\mu = \mu(x, z)$ and $\rho = \rho(x, z) = \mu\rho^{(0)}$.

Let (x^*, y^*) be a strictly complementary feasible solution (see next slide for proof of existence).

Consider

$$\begin{aligned} z^T x^* + x^T z^* &= z^T x^* - x^T A x^* && ((x^*, z^*) \text{ is feasible}) \\ &= (-A^T x + z)^T x^* \\ &= (Ax + z)^T x^* && (A \text{ is skew symmetric}) \\ &= \rho^T x^* && (\text{definition of } \rho) \\ &= \mu {\rho^{(0)}}^T x^*. \end{aligned}$$

Let $M = {\rho^{(0)}}^T x^*$. From above, we see that

$$\begin{aligned} z_j x_j^* &\leq \mu M && \text{for all } j \\ x_j z_j^* &\leq \mu M && \text{for all } j \end{aligned}$$

It's easy to check that $(x, z) \in \mathcal{N}(\beta)$ implies that $x_j z_j \geq (1 - \beta)\mu$ for all j.
Hence, $\mu \leq x_j z_j / (1 - \beta)$ for all j.

Putting this altogether, we get...

$$z_j x_j^* \leq x_j z_j \frac{M}{1-\beta} \qquad \text{for all } j$$
$$x_j z_j^* \leq x_j z_j \frac{M}{1-\beta} \qquad \text{for all } j$$

Simplifying,, we get

$$x_j \geq x_j^* \frac{1-\beta}{M}$$
$$z_j \geq z_j^* \frac{1-\beta}{M}$$

Hence,

$$x_j + z_j \geq c,$$

where

$$c = (x_j^* + z_j^*)\frac{1-\beta}{M} > 0.$$

QED

An Aside: Strict Complementarity

Consider a primal-dual pair:

$$(P) \qquad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax + w = b \\ & x, w \geq 0 \end{array} \qquad\qquad (D) \qquad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y - z = c \\ & y, z \geq 0 \end{array}$$

Let (x, w) be primal feasible and (y, z) be dual feasible.

We say that they are *complementary* if $x_j z_j = 0$ for all j and $w_i y_i = 0$ for all i.

We say that they are *strictly complementary* if, in addition, $x + z > 0$ and $y + w > 0$.

Theorem 6. *If the two problems have feasible solutions, then there exists a strictly complementary pair (x, w) and (y, z).*

Proof. We start by constructing $n + m$ feasible pairs.

For each j, we will construct a feasible pair for which $x_j + z_j > 0$.

For each i, we will construct a feasible pair for which $y_i + w_i > 0$.

Then, we average these $n + m$ solutions to get a single strictly complementary pair.

Fix j. If there is a primal feasible solution for which $x_j > 0$, then put that in the collection together with any dual feasible solution.

Now, suppose there does not exist a primal feasible solution for which $x_j > 0$.

Such an x_j is called a *null variable*.

If x_j is a null variable, we must construct a dual feasible solution for which $z_j > 0$.

Consider this primal-dual pair:

$$(P_j) \quad \begin{array}{ll} \text{maximize} & x_j \\ \text{subject to} & Ax + w = b \\ & \quad x, w \geq 0 \end{array} \qquad\qquad (D_j) \quad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y - z = e_j \\ & \quad y, z \geq 0 \end{array}$$

Problem (P_j) is feasible by assumption. It is also bounded. Therefore, its dual has an optimal solution.

Let (y', z') be a feasible solution to (D_j) and let (y'', z'') be a feasible solution to (D).

It is easy to see that $(y, z) = (y' + y'', z' + z'' + e_j)$ is a feasible solution to (D).

And, for this solution, $z_j \geq 1$.

A similar construction works for y_i, w_i pairs. *QED*