The Homogeneous Self-Dual Method

Robert J. Vanderbei

December 14, 2005
ORF 522

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

The Homogeneous Self-Dual Problem

Primal-Dual Pair

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

Homogeneous Self-Dual Problem

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

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. 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:

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

More Notation

Infeasibility: $\rho(x, z) = Ax + z$

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

Nonlinear System

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

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.

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

Now, just divide by n.

4.

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

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

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

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 $\delta = 1$.

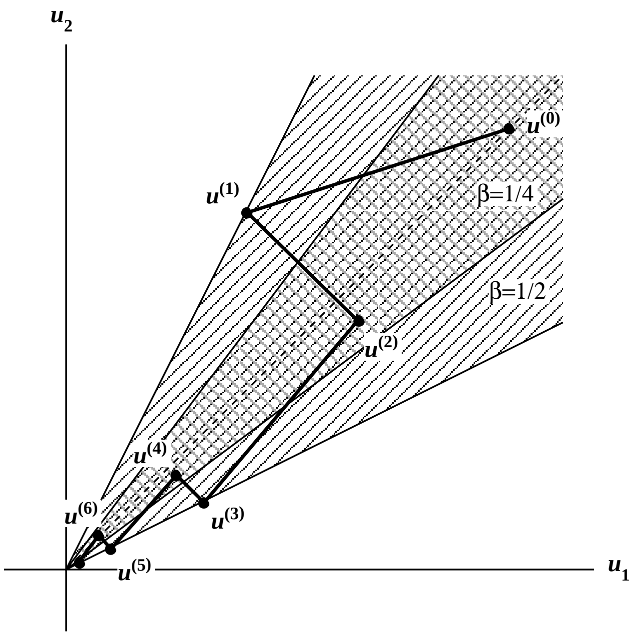
Put $\theta = 1$.

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.

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. 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.

Back to Original Primal-Dual Setting

Just a final remark: If primal and dual problems are feasible, then algorithm will produce a solution to HSD with $\phi > 0$ from which a solution to original problem can be extracted.
See text for details.