Interior Point Algorithms II: Potential Reduction

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye (LY, Chapter 5.6-5.7)
Next week: Chapters Chapters 5.7, 14.5-14.6

Primal-Dual Potential Function for LP

Typically, a single merit-function driven algorithm is preferred since it can adaptively take large step sizes as long as the merit value is sufficiently reduced, comparing to check and balance of hyper-parameters/measures of the path-following type of algorithms.

For $\mathbf{x} \in \text{int}\,\mathcal{F}_p$ and $(\mathbf{y}, \mathbf{s}) \in \text{int}\,\mathcal{F}_d$, the joint primal-dual potential function is defined by

$$\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) := (n+\rho)\log(\mathbf{x}^T\mathbf{s}) - \sum_{j=1}^{n} \log(x_j s_j),$$

where $\rho \geq 0$ and it is fixed.

$$\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) = \rho\log(\mathbf{x}^T\mathbf{s}) + \psi_n(\mathbf{x}, \mathbf{s}) \geq \rho\log(\mathbf{x}^T\mathbf{s}) + n\log n,$$

then, for $\rho > 0$, $\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) \rightarrow -\infty$ implies that $\mathbf{x}^T\mathbf{s} \rightarrow 0$. More precisely, we have

$$\mathbf{x}^T\mathbf{s} \leq \exp(\frac{\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) - n\log n}{\rho}).$$

Primal-Dual Potential Reduction Algorithm for LP

Once have a pair $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k) \in \operatorname{int} \mathcal{F}$, we compute direction vectors $\mathbf{d}_x$, $\mathbf{d}_y$ and $\mathbf{d}_s$ from the system equations:

$$\begin{aligned} S^k\mathbf{d}_x + X^k\mathbf{d}_s &= \frac{(\mathbf{x}^k)^T\mathbf{s}^k}{n+\rho}\mathbf{e} - X^kS^k\mathbf{e}, \\ A\mathbf{d}_x &= \mathbf{0}, \\ -A^T\mathbf{d}_y - \mathbf{d}_s &= \mathbf{0}. \end{aligned} \tag{1}$$

Note that $\mathbf{d}_x^T\mathbf{d}_s = -\mathbf{d}_x^T A^T \mathbf{d}_y = 0$ here. Then choose a step-size scalar $\theta (> 0)$ and assign

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \theta\mathbf{d}_x > \mathbf{0},\ \mathbf{y}^{k+1} = \mathbf{y}^k + \theta\mathbf{d}_y,\ \mathbf{s}^{k+1} = \mathbf{s}^k + \theta\mathbf{d}_s > \mathbf{0}.$$

This is the Newton method for the optimality conditions/equations of the potential minimization problem:

$$\begin{aligned} XS\mathbf{e} &= \frac{(\mathbf{x}^k)^T\mathbf{s}^k}{n+\rho}\mathbf{e}, \\ A\mathbf{x} &= \mathbf{b}, \\ -A^T\mathbf{y} - \mathbf{s} &= -\mathbf{c}. \end{aligned} \tag{2}$$

To simplify rotations, let

$$\begin{aligned} \mathbf{d}_{x'} + \mathbf{d}_{s'} &= \mathbf{r}' := (XS)^{-0.5}(\frac{\mathbf{x}^T\mathbf{s}}{n+\rho}\mathbf{e} - XS\mathbf{e}), \\ A'\mathbf{d}_{x'} &= \mathbf{0}, \\ -(A')^T\mathbf{d}_y - \mathbf{d}_{s'} &= \mathbf{0}. \end{aligned}$$

where

$$D = X^{0.5}S^{-0.5},\ A' = AD,\ \mathbf{d}_{x'} = D^{-1}\mathbf{d}_x,\ \mathbf{d}_{s'} = D\mathbf{d}_s.$$

Again, we maintain $\mathbf{d}_{x'}^T\mathbf{d}_{s'} = 0$.

Unlike in the path-following algorithm, $\|\mathbf{r}'\|^2$ may be too big to make $\mathbf{x}+\mathbf{d}_x$ or $\mathbf{s}+\mathbf{d}_s$ positive. So that we need to add a step size θ to scale $\mathbf{r}'$ such that it makes new iterate feasible.

Lemma 1 *Let the direction vector* $\mathbf{d} = (\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_s)$ *be generated by equation (2), and let*

$$\theta = \frac{\alpha\sqrt{\min(XS\mathbf{e})}}{\|\mathbf{r}'\|}\ , \tag{3}$$

where α *is a positive constant less than 1. Let*

$$\mathbf{x}^+ = \mathbf{x} + \theta\mathbf{d}_x, \quad \mathbf{y}^+ = \mathbf{y} + \theta\mathbf{d}_y, \quad \textit{and} \quad \mathbf{s}^+ = \mathbf{s} + \theta\mathbf{d}_s.$$

Then, we have $(\mathbf{x}^+, \mathbf{y}^+, \mathbf{s}^+) \in \mathrm{int}\,\mathcal{F}$ *and*

$$\psi_{n+\rho}(\mathbf{x}^+, \mathbf{s}^+) - \psi_{n+\rho}(\mathbf{x}, \mathbf{s})$$

$$\leq -\alpha\sqrt{\min(XS\mathbf{e})}\|(XS)^{-1/2}(\mathbf{e} - \frac{(n+\rho)}{\mathbf{x}^T\mathbf{s}}X\mathbf{s})\| + \frac{\alpha^2}{2(1-\alpha)}\ .$$

Logarithmic Approximation Lemma

We first present a technical lemma:

Lemma 2 *If* $\mathbf{d} \in \mathcal{R}^n$ *such that* $\|\mathbf{d}\|_\infty < 1$ *then*

$$\mathbf{e}^T\mathbf{d} \geq \sum_{i=1}^{n} \log(1+d_i) \geq \mathbf{e}^T\mathbf{d} - \frac{\|\mathbf{d}\|^2}{2(1-\|\mathbf{d}\|_\infty)}.$$

The proof is based on the Taylor expansion of $\ln(1+d_i)$ for $-1 < d_i < 1$.

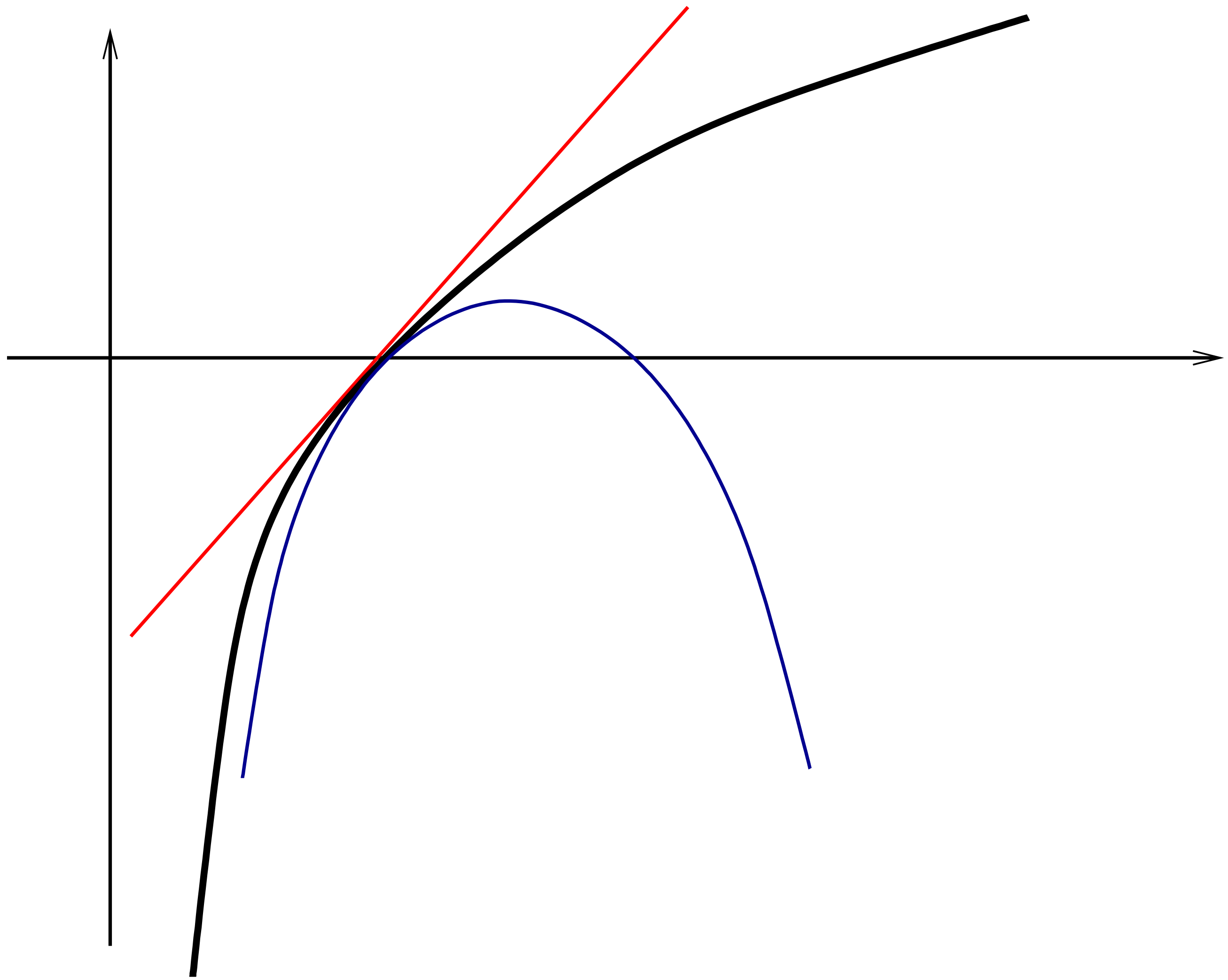

Figure 1: Logarithmic approximation by linear and quadratic functions

Proof Sketch of the Theorem

It is clear that $A\mathbf{x}^+ = \mathbf{b}$ and $A^T\mathbf{y}^+ + \mathbf{s}^+ = \mathbf{c}$. We now show that $\mathbf{x}^+ > \mathbf{0}$ and $\mathbf{s}^+ > \mathbf{0}$. This is similar to the previous proof for the path-following algorithm

$$\|\theta X^{-1}\mathbf{d}_x\|^2 + \|\theta S^{-1}\mathbf{d}_s\|^2 \le \theta^2 \frac{\|\mathbf{r}'\|^2}{\min(XS\mathbf{e})} = \frac{\alpha^2 \min(XS\mathbf{e})}{\|\mathbf{r}'\|^2} \frac{\|\mathbf{r}'\|^2}{\min(XS\mathbf{e})} = \alpha^2 < 1.$$

Therefore,

$$\mathbf{x}^+ = \mathbf{x} + \theta\mathbf{d}_x = X(\mathbf{e} - \theta X^{-1}\mathbf{d}_x) > \mathbf{0}$$

and

$$\mathbf{s}^+ = \mathbf{s} + \theta\mathbf{d}_s = S(\mathbf{e} - \theta S^{-1}\mathbf{d}_s) > \mathbf{0}.$$

Sketch of the proof continued

$$
\begin{aligned}
& \psi(\mathbf{x}^+, \mathbf{s}^+) - \psi(\mathbf{x}, \mathbf{s}) \\
= \; & (n+\rho)\log\left(1 + \frac{\theta \mathbf{d}_s^T\mathbf{x} + \theta \mathbf{d}_x^T\mathbf{s}}{\mathbf{x}^T\mathbf{s}}\right) - \sum_{j=1}^n \left(\log(1 + \frac{\theta d_{s_j}}{s_j}) + \log(1 + \frac{\theta d_{x_j}}{x_j})\right) \\
\le \; & (n+\rho)\left(\frac{\theta \mathbf{d}_s^T\mathbf{x} + \theta \mathbf{d}_x^T\mathbf{s}}{\mathbf{x}^T\mathbf{s}}\right) - \sum_{j=1}^n \left(\log(1 + \frac{\theta d_{s_j}}{s_j}) + \log(1 + \frac{\theta d_{x_j}}{x_j})\right) \\
\le \; & (n+\rho)\left(\frac{\theta \mathbf{d}_s^T\mathbf{x} + \theta \mathbf{d}_x^T\mathbf{s}}{\mathbf{x}^T\mathbf{s}}\right) - \theta \mathbf{e}^T(S^{-1}\mathbf{d}_s + X^{-1}\mathbf{d}_x) + \frac{||\theta S^{-1}\mathbf{d}_s||^2 + ||\theta X^{-1}\mathbf{d}_x||^2}{2(1-\alpha)} \\
\le \; & \frac{n+\rho}{\mathbf{x}^T\mathbf{s}}\theta(\mathbf{d}_s^T\mathbf{x} + \mathbf{d}_x^T\mathbf{s}) - \theta \mathbf{e}^T(S^{-1}\mathbf{d}_s + X^{-1}\mathbf{d}_x) + \frac{\alpha^2}{2(1-\alpha)} \\
= \; & \theta\left(\frac{n+\rho}{\mathbf{x}^T\mathbf{s}}\mathbf{e}^T(X\mathbf{d}_s + S\mathbf{d}_x) - \mathbf{e}^T(S^{-1}\mathbf{d}_s + X^{-1}\mathbf{d}_x)\right) + \frac{\alpha^2}{2(1-\alpha)} \\
= \; & \theta\left(\frac{n+\rho}{\mathbf{x}^T\mathbf{s}}\mathbf{e}^T(X\mathbf{d}_s + S\mathbf{d}_x) - \mathbf{e}^T(XS)^{-1}(X\mathbf{d}_s + S\mathbf{d}_x)\right) + \frac{\alpha^2}{2(1-\alpha)} \\
= \; & \theta\left(\frac{n+\rho}{\mathbf{x}^T\mathbf{s}}XS\mathbf{e} - \mathbf{e}\right)^T (XS)^{-1}(X\mathbf{d}_s + S\mathbf{d}_x) + \frac{\alpha^2}{2(1-\alpha)} \\
= \; & \theta\left(\frac{n+\rho}{\mathbf{x}^T\mathbf{s}}XS\mathbf{e} - \mathbf{e}\right)^T (XS)^{-1}\left(\frac{\mathbf{x}^T\mathbf{s}}{n+\rho}\mathbf{e} - XS\mathbf{e}\right) + \frac{\alpha^2}{2(1-\alpha)} \\
= \; & -\theta \cdot \frac{n+\rho}{\mathbf{x}^T\mathbf{s}} \cdot \|\mathbf{r}'\|^2 + \frac{\alpha^2}{2(1-\alpha)} = -\alpha\sqrt{\min(XS\mathbf{e})} \cdot \frac{n+\rho}{\mathbf{x}^T\mathbf{s}} \cdot \|\mathbf{r}'\| + \frac{\alpha^2}{2(1-\alpha)}.
\end{aligned}
$$

Let $\mathbf{v} = XS\mathbf{e}$. Then, we can prove the following technical lemma:

Lemma 3 *Let* $\mathbf{v} \in \mathcal{R}^n$ *be a positive vector and* $\rho \geq \sqrt{n}$*. Then,*

$$\sqrt{\min(\mathbf{v})}\|V^{-1/2}(\mathbf{e} - \frac{(n+\rho)}{\mathbf{e}^T\mathbf{v}}\mathbf{v})\| \geq \sqrt{3/4}\,.$$

Combining these two lemmas we have

$$\psi_{n+\rho}(\mathbf{x}^+, \mathbf{s}^+) - \psi_{n+\rho}(\mathbf{x}, \mathbf{s})$$

$$\leq -\alpha\sqrt{3/4} + \frac{\alpha^2}{2(1-\alpha)} = -\delta$$

for a constant δ.

Description of Algorithm

Given $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{s}^0) \in \mathrm{int}\,\mathcal{F}$. Set $\rho \geq \sqrt{n}$ and $k := 0$.

While $(\mathbf{x}^k)^T\mathbf{s}^k \geq \epsilon$ **do**

1. Set $(\mathbf{x}, \mathbf{s}) = (\mathbf{x}^k, \mathbf{s}^k)$ and $\gamma = n/(n+\rho)$ and compute $(\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_s)$ from (2).
2. Let $\mathbf{x}^{k+1} = \mathbf{x}^k + \bar{\alpha}\mathbf{d}_x$, $\mathbf{y}^{k+1} = \mathbf{y}^k + \bar{\alpha}\mathbf{d}_y$, and $\mathbf{s}^{k+1} = \mathbf{s}^k + \bar{\alpha}\mathbf{d}_s$ where

$$\bar{\alpha} = \arg\min_{\alpha \geq 0} \psi_{n+\rho}(\mathbf{x}^k + \alpha\mathbf{d}_x, \mathbf{s}^k + \alpha\mathbf{d}_s).$$

3. Let $k := k + 1$ and return to Step 1.

Theorem 1 *Let* $\rho \geq \sqrt{n}$ *and* $\psi_{n+\rho}(\mathbf{x}^0, \mathbf{s}^0) \leq \rho \log((\mathbf{x}^0)^T \mathbf{s}^0) + n \log n$*. Then, the Algorithm terminates in at most* $O(\rho \log((\mathbf{x}^0)^T \mathbf{s}^0 / \epsilon))$ *iterations with*

$$(\mathbf{x}^k)^T \mathbf{s}^k = \mathbf{c}^T \mathbf{x}^k - \mathbf{b}^T \mathbf{y}^k \leq \epsilon.$$

$$\begin{aligned}
(\mathbf{x}^k)^T \mathbf{s}^k &\leq \exp(\tfrac{\psi_{n+\rho}(\mathbf{x}^k, \mathbf{s}^k) - n \log n}{\rho}) \\
&\leq \exp(\tfrac{\psi_{n+\rho}(\mathbf{x}^0, \mathbf{s}^0) - n \log n - \rho \log((\mathbf{x}^0)^T \mathbf{s}^0/\epsilon)}{\rho}) \\
&\leq \exp(\tfrac{\rho \log(\mathbf{x}^0, \mathbf{s}^0) - \rho \log((\mathbf{x}^0)^T \mathbf{s}^0/\epsilon)}{\rho}) \\
&= \exp(\log(\epsilon)) = \epsilon.
\end{aligned}$$

The adaptively search of best ρ?

Termination with Exact Optimizers

- The first is a "cross-over" procedure to find a basic feasible solution (BFS, corner point) whose objective value is at least as good as the current interior point. Let $A, \mathbf{b}, \mathbf{c}$ be integers and L be their bit length, and let a second best BFS solution be $\mathbf{x}^{2nd}$ and the optimal objective value be z^*. Then

$$\mathbf{c}^T\mathbf{x}^{2nd} - z^* > 2^{-L}.$$

Thus, one can terminate interior-point algorithm when

$$\mathbf{c}^T\mathbf{x}^k - \mathbf{b}^T\mathbf{y}^k \leq 2^{-L}.$$

- The second approach is to compute a strictly complementary solution pair. The method uses the primal-dual interior-point pair to identify the strict complementarity partition (P^*, Z^*) and then "purify or project" the primal interior solution onto the primal optimal face and the dual interior solution onto the dual optimal face, based on the following theorem:

Theorem 2 *Given an interior solution $\mathbf{x}^k$ and $\mathbf{s}^k$ in the solution sequence generated by an* *interior-point algorithm, define*

$$P^k = \{j : x_j^k \geq s_j^k, \forall j\} \quad \textit{and} \quad Z^k = \{1, ..., n\} \setminus P^k.$$

Then, we have $P^k = P^*$ *whenever*

$$\mathbf{c}^T\mathbf{x}^k - \mathbf{b}^T\mathbf{y}^k \leq 2^{-L}.$$

Thus, the worst-case iteration bound for interior-point algorithms is $O(\sqrt{n}L)$ if the initial point pair $(\mathbf{x}^0)^T\mathbf{s}^0 \leq 2^L$.

Initialization

- Combining the primal and dual into a single linear feasibility problem, then applying LP algorithms to find a feasible point of the problem. Theoretically, this approach can retain the currently best complexity result.
- The big M method, i.e., add one or more artificial column(s) and/or row(s) and a huge penalty parameter M to force solutions to become feasible during the algorithm.
- Phase I-then-Phase II method, i.e., first try to find a feasible point (and possibly one for the dual problem), and then start to look for an optimal solution if the problem is feasible and bounded.
- Combined Phase I-Phase II method, i.e., approach feasibility and optimality simultaneously. To our knowledge, the "best" complexity of this approach is $O(n\log(R/\epsilon))$.

Homogeneous and Self-Dual Algorithm

- It solves the linear programming problem without any regularity assumption concerning the existence of optimal, feasible, or interior feasible solutions, while it retains the currently best complexity result
- It can start at any positive primal-dual pair, feasible or infeasible, near the central ray of the positive orthant (cone), and it does not use any big M penalty parameter or lower bound.
- Each iteration solves a system of linear equations whose dimension is almost the same as that solved in the standard (primal-dual) interior-point algorithms.
- If the LP problem has a solution, the algorithm generates a sequence that approaches feasibility and optimality simultaneously; if the problem is infeasible or unbounded, the algorithm will produce an infeasibility certificate for at least one of the primal and dual problems.

Primal-Dual Alternative Systems

A pair of LP has two alternatives

(Solvable)
$$\begin{array}{rl} A\mathbf{x}-\mathbf{b} & =\mathbf{0} \\ -A^T\mathbf{y}+\mathbf{c} & \geq \mathbf{0}, \\ \mathbf{b}^T\mathbf{y}-\mathbf{c}^T\mathbf{x} & =0, \\ \mathbf{y}\text{ free},\ \mathbf{x} & \geq \mathbf{0} \end{array}$$

or

(Infeasible)
$$\begin{array}{rl} A\mathbf{x} & =\mathbf{0} \\ -A^T\mathbf{y} & \geq \mathbf{0}, \\ \mathbf{b}^T\mathbf{y}-\mathbf{c}^T\mathbf{x} & >0, \\ \mathbf{y}\text{ free},\ \mathbf{x} & \geq \mathbf{0} \end{array}$$

An Integrated Homogeneous System

The two alternative systems can be homogenized as one:

$$
\begin{array}{llr}
(HP) & A\mathbf{x} - \mathbf{b}\tau & = \mathbf{0} \\
 & -A^T\mathbf{y} + \mathbf{c}\tau & = \mathbf{s} \geq \mathbf{0}, \\
 & \mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} & = \kappa \geq 0, \\
 & \mathbf{y} \text{ free}, \ (\mathbf{x}; \tau) & \geq \mathbf{0}
\end{array}
$$

where the two alternatives are

$$
\text{(Solvable)}: \ (\tau > 0, \kappa = 0) \quad \text{or} \quad \text{(Infeasible)}: \ (\tau = 0, \kappa > 0)
$$

The Homogeneous System is Self-Dual

$$
\begin{array}{rrll}
(HP) & A\mathbf{x} - \mathbf{b}\tau & = \mathbf{0}, & (\mathbf{y}') \\
& -A^T\mathbf{y} + \mathbf{c}\tau & = \mathbf{s} \geq \mathbf{0}, & (\mathbf{x}') \\
& \mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} & = \kappa \geq 0, & (\tau') \\
& \mathbf{y} \text{ free}, \ (\mathbf{x}; \tau) & \geq \mathbf{0} &
\end{array}
\qquad
\begin{array}{rrl}
(HD) & A\mathbf{x}' - \mathbf{b}\tau' & = \mathbf{0}, \\
& A^T\mathbf{y}' - \mathbf{c}\tau' & \leq \mathbf{0}, \\
& -\mathbf{b}^T\mathbf{y}' + \mathbf{c}^T\mathbf{x}' & \leq 0, \\
& \mathbf{y}' \text{ free}, \ (\mathbf{x}'; \tau') & \geq \mathbf{0}
\end{array}
$$

Theorem 3 *System (HP) is feasible (e.g. all zeros) and any feasible solution* $(\mathbf{y}, \mathbf{x}, \tau, \mathbf{s}, \kappa)$ *is self-complementary:*

$$\mathbf{x}^T\mathbf{s} + \tau\kappa = 0.$$

Furthermore, it has a strictly self-complementary feasible solution

$$\begin{pmatrix} \mathbf{x} + \mathbf{s} \\ \tau + \kappa \end{pmatrix} > \mathbf{0},$$

Let's Find Such a Feasible Solution

Given $\mathbf{x}^0 = \mathbf{e} > \mathbf{0}$, $\mathbf{s}^0 = \mathbf{e} > \mathbf{0}$, and $\mathbf{y}^0 = \mathbf{0}$, we formulate

$$
\begin{array}{lrrrrl}
(HSDP) \quad \min & & & & \theta & \\
\text{s.t.} & & A\mathbf{x} & -\mathbf{b}\tau & +\bar{\mathbf{b}}\theta & = \mathbf{0}, \\
& -A^T\mathbf{y} & & +\mathbf{c}\tau & -\bar{\mathbf{c}}\theta & \geq \mathbf{0}, \\
& \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & +\bar{z}\theta & \geq 0, \\
& \mathbf{y} \text{ free}, & \mathbf{x} \geq \mathbf{0}, & \tau \geq 0, & \theta \text{ free}, &
\end{array}
$$

where

$$
\bar{\mathbf{b}} = \mathbf{b} - A\mathbf{e}, \quad \bar{\mathbf{c}} = \mathbf{c} - \mathbf{e}, \quad \bar{z} = \mathbf{c}^T\mathbf{e} + 1.
$$

But it may just give us the all-zero solution.

A HSD linear program

Let's try to add one more constraint to prevent the all-zero solution

$$
\begin{array}{lrrrrrl}
(HSDP) & \min & & & & (n+1)\theta & \\
& \text{s.t.} & & A\mathbf{x} & -\mathbf{b}\tau & +\bar{\mathbf{b}}\theta & = \mathbf{0}, \\
& & -A^T\mathbf{y} & & +\mathbf{c}\tau & -\bar{\mathbf{c}}\theta & \geq \mathbf{0}, \\
& & \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & +\bar{z}\theta & \geq 0, \\
& & -\bar{\mathbf{b}}^T\mathbf{y} & +\bar{\mathbf{c}}^T\mathbf{x} & -\bar{z}\tau & & = -(n+1), \\
& & \mathbf{y} \text{ free}, & \mathbf{x} \geq \mathbf{0}, & \tau \geq 0, & \theta \text{ free}. &
\end{array}
$$

Note that the constraints of (HSDP) form a skew-symmetric system and the objective coeffcient vector is the negative of the right-hand-side vector, so that it remains a self-dual linear program.

$(\mathbf{y} = \mathbf{0},\ \mathbf{x} = \mathbf{e},\ \tau = 1,\ \theta = 1)$ is a strictly feasible point for (HSDP).

$$
\begin{array}{llrrrrl}
(HSDP) & \min & & & & (n+1)\theta & \\
& \text{s.t.} & & A\mathbf{x} & -\mathbf{b}\tau & +\bar{\mathbf{b}}\theta & = \mathbf{0}, \\
& & -A^T\mathbf{y} & & +\mathbf{c}\tau & -\bar{\mathbf{c}}\theta & = \mathbf{s} \geq \mathbf{0}, \\
& & \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & +\bar{z}\theta & = \kappa \geq 0, \\
& & -\bar{\mathbf{b}}^T\mathbf{y} & +\bar{\mathbf{c}}^T\mathbf{x} & -\bar{z}\tau & & = -(n+1), \\
& & \mathbf{y} \text{ free}, & \mathbf{x} \geq \mathbf{0}, & \tau \geq 0, & \theta \text{ free}. &
\end{array}
$$

Denote by $\mathcal{F}_h$ the set of all points $(\mathbf{y}, \mathbf{x}, \tau, \theta, \mathbf{s}, \kappa)$ that are feasible for (HSDP). Denote by $\mathcal{F}_h^0$ the set of interior feasible points with $(\mathbf{x}, \tau, \mathbf{s}, \kappa) > \mathbf{0}$ in $\mathcal{F}_h$. By combining the constraints, we can derive the last (equality) constraint as

$$
\mathbf{e}^T x + \mathbf{e}^T s + \tau + \kappa - (n+1)\theta = (n+1),
$$

which serves indeed as a normalizing constraint for (HSDP) to prevent the all-zero solution.

Theorem 4 *Consider problems (HSDP) and (HSDD).*

i) *(HSDD) has the same form as (HSDP), i.e., (HSDD) is simply (HSDP) with* $(\mathbf{y}, \mathbf{x}, \tau, \theta)$ *being replaced by* $(\mathbf{y}', \mathbf{x}', \tau', \theta')$.

ii) *(HSDP) has a strictly feasible point*

$$\mathbf{y} = \mathbf{0}, \quad x = \mathbf{e} > \mathbf{0}, \quad \tau = 1, \quad \theta = 1, \quad \mathbf{s} = \mathbf{e} > \mathbf{0}, \quad \kappa = 1.$$

iii) *(HSDP) has an optimal solution and its optimal solution set is bounded.*

iv) *The optimal value of (HSDP) is zero, and*

$$(\mathbf{y}, \mathbf{x}, \tau, \theta, \mathbf{s}, \kappa) \in \mathcal{F}_h \quad \textit{implies that} \quad (n+1)\theta = \mathbf{x}^T\mathbf{s} + \tau\kappa.$$

v) *There is an optimal solution* $(\mathbf{y}^*, \mathbf{x}^*, \tau^*, \theta^* = 0, \mathbf{s}^*, \kappa^*) \in \mathcal{F}_h$ *such that*

$$\begin{pmatrix} \mathbf{x}^* + \mathbf{s}^* \\ \tau^* + \kappa^* \end{pmatrix} > \mathbf{0},$$

which we call a strictly self-complementary solution. (Similarly, we sometimes call an optimal solution to (HSDP) a self-complementary solution; the strict inequalities above need not hold.)

Theorem 5 *Let* $(\mathbf{y}^*, \mathbf{x}^*, \tau^*, \theta^* = 0, \mathbf{s}^*, \kappa^*)$ *be a strictly self complementary solution for (HSDP).*

i) *(LP) has a solution (feasible and bounded) if and only if* $\tau^* > 0$. *In this case,* $\mathbf{x}^*/\tau^*$ *is an optimal solution for (LP) and* $(\mathbf{y}^*/\tau^*, \mathbf{s}^*/\tau^*)$ *is an optimal solution for (LD).*

ii) *(LP) has no solution if and only if* $\kappa^* > 0$. *In this case,* $\mathbf{x}^*/\kappa^*$ *or* $\mathbf{s}^*/\kappa^*$ *or both are certificates for proving infeasibility: if* $\mathbf{c}^T\mathbf{x}^* < 0$ *then (LD) is infeasible; if* $-\mathbf{b}^T\mathbf{y}^* < 0$ *then (LP) is infeasible; and if both* $\mathbf{c}^T\mathbf{x}^* < 0$ *and* $-\mathbf{b}^T\mathbf{y}^* < 0$ *then both (LP) and (LD) are infeasible.*

Theorem 6 **i)** *For any* $\mu > 0$*, there is a unique* $(\mathbf{y}, \mathbf{x}, \tau, \theta, \mathbf{s}, \kappa)$ *in* $\mathcal{F}_h^0$*, such that*

$$\begin{pmatrix} X\mathbf{s} \\ \tau\kappa \end{pmatrix} = \mu\mathbf{e}.$$

ii) *Let* $(\mathbf{d}_y, \mathbf{d}_x, d_\tau, d_\theta, \mathbf{d}_s, d_\kappa)$ *be in the null space of the constraint matrix of (HSDP) after adding surplus variables* $\mathbf{s}$ *and* κ*, i.e.,*

$$\begin{array}{rrrrrrcl} & A\mathbf{d}_x & -\mathbf{b}d_\tau & +\bar{\mathbf{b}}d_\theta & & & = & \mathbf{0}, \\ -A^T\mathbf{d}_y & & +\mathbf{c}d_\tau & -\bar{\mathbf{c}}d_\theta & -\mathbf{d}_s & & = & \mathbf{0}, \\ \mathbf{b}^T\mathbf{d}_y & -\mathbf{c}^T\mathbf{d}_x & & +\bar{z}d_\theta & & -d_\kappa & = & 0, \\ -\bar{\mathbf{b}}^T\mathbf{d}_y & +\bar{\mathbf{c}}^T\mathbf{d}_x & -\bar{z}d_\tau & & & & = & 0. \end{array} \quad (4)$$

$$(\mathbf{d}_x)^T\mathbf{d}_s + d_\tau d_\kappa = 0.$$

Endogenous Potential Function and Central Path

$$\psi_{n+\rho}(\mathbf{x},\mathbf{s},\tau,\kappa) := (n+1+\rho)\log(\mathbf{x}^T\mathbf{s}+\tau\kappa) - \sum_{j=1}^{n}\log(x_j s_j) - \log(\tau\kappa),$$

and

$$\mathcal{C} = \left\{ (\mathbf{y},\mathbf{x},\tau,\theta,\mathbf{s},\kappa) \in \mathcal{F}_h^0 : \begin{pmatrix} X\mathbf{s} \\ \tau\kappa \end{pmatrix} = \frac{\mathbf{x}^T\mathbf{s}+\tau\kappa}{n+1}\mathbf{e} \right\}.$$

Obviously, the initial interior feasible point proposed in Theorem 4 is on the path with $\mu = 1$ or $(\mathbf{x}^0)^T\mathbf{s}^0 + \tau^0\kappa^0 = n+1$.

Solving (HSDP)

Consider solving the following system of linear equations for $(\mathbf{d}_y, \mathbf{d}_x, d_\tau, d_\theta, \mathbf{d}_s, d_\kappa)$ that satisfies (4) and

$$\begin{pmatrix} X\mathbf{d}_s + S\mathbf{d}_x \\ \tau^k d_\kappa + \kappa^k d_\tau \end{pmatrix} = \gamma\mu\mathbf{e} - \begin{pmatrix} X\mathbf{s} \\ \tau\kappa \end{pmatrix}.$$

Theorem 7 *The $O(\sqrt{n}\log((\mathbf{x}^0)^T\mathbf{s}^0/\epsilon))$ interior-point algorithm, coupled with a termination technique described above, generates a strictly self-complementary solution for (HSDP) in $O(\sqrt{n}(\log(c(A,\mathbf{b},\mathbf{c})) + \log n))$ iterations and $O(n^3(\log(c(A,\mathbf{b},\mathbf{c})) + \log n))$ operations, where $c(A,\mathbf{b},\mathbf{c})$ is a positive number depending on the data $(A,\mathbf{b},\mathbf{c})$. If (LP) and (LD) have integer data with bit length L, then by the construction, the data of (HSDP) remains integral and its length is $O(L)$. Moreover, $c(A,\mathbf{b},\mathbf{c}) \le 2^L$. Thus, the algorithm terminates in $O(\sqrt{n}L)$ iterations and $O(n^3L)$ operations.*

Example

Consider the example where

$$A = \begin{pmatrix} -1 & 0 & 0 \end{pmatrix}, \quad b = 1, \qquad \text{and} \quad \mathbf{c} = \begin{pmatrix} 0 & 1 & -1 \end{pmatrix}.$$

Then,

$$y^* = 2, \quad \mathbf{x}^* = (0,2,1)^T, \quad \tau^* = 0, \quad \theta^* = 0, \quad \mathbf{s}^* = (2,0,0)^T, \quad \kappa^* = 1$$

could be a strictly self-complementary solution generated for (HSDP) with

$$\mathbf{c}^T\mathbf{x}^* = 1 > 0, \quad by^* = 2 > 0.$$

Thus $(y^*, \mathbf{s}^*)$ demonstrates the infeasibility of (LP), but $\mathbf{x}^*$ doesn't show the infeasibility of (LD). Of course, if the algorithm generates instead $\mathbf{x}^* = (0,1,2)^T$, then we get demonstrated infeasibility of both.

Software Implementation

Cplex, GUROBI

SEDUMI: `http://sedumi.mcmaster.ca/`

MOSEK: `http://www.mosek.com/products_mosek.html`

IPOPT: `https://projects.coin-or.org/Ipopt`

hsdLPsolver: Sparse Linear Programming Solver (Matlabe .m file).