Slide 1

Detecting Infeasibility in Infeasible-Interior-Point Methods for Optimization

Michael J. Todd,

School of Operations Research and Industrial Engineering,
Cornell University,

miketodd@cs.cornell.edu,
http://www.orie.cornell.edu/ miketodd/todd.html

FOCM '02, Minneapolis

Slide 2

Linear Programming

For most of the talk, we confine ourselves to linear programming. Consider the standard-form primal problem

$$(P) \quad \min_x \quad c^T x, \qquad Ax = b, \quad x \geq 0,$$

together with its dual

$$(D) \quad \max_{y,s} \quad b^T y, \qquad A^T y + s = c, \quad s \geq 0.$$

Here A is an $m \times n$ matrix, wlog of rank m, and the vectors are of appropriate sizes.

Slide 3

Optimality conditions

If x is feasible in (P) and (y, s) in (D), then (weak duality)

$$c^T x - b^T y = (A^T y + s)^T x - (Ax)^T y = s^T x \geq 0.$$

Hence if we have feasible solutions with equal objective values, or equivalently with $s^T x = 0$, these solutions are optimal. We therefore have the following optimality conditions:

$$(OC) \quad \begin{array}{lclcll} & A^T y & + & s & = & c, & s \geq 0, \\ Ax & & & & = & b, & x \geq 0, \\ & XSe & & & = & 0, & \end{array}$$

where $X = \mathrm{Diag}(x)$, $S = \mathrm{Diag}(s)$, and $e \in \mathbf{R}^n$ denotes the vector of ones. These conditions are in fact necessary as well as sufficient for optimality (strong duality).

Slide 4

Central path equations

Path-following interior-point methods iterate approximate solutions to

$$(CPE_\nu) \quad \begin{array}{lclcll} & A^T y & + & s & = & c, & (s > 0) \\ Ax & & & & = & b, & (x > 0) \\ & XSe & & & = & \nu e, & \end{array}$$

for $\nu > 0$. The perturbation of the complementary slackness conditions $XSe = 0$ is designed to make the inequality (hard) constraints secondary. $(n + m + n) \times (n + m + n)$ system.

Slide 5

Central path theorem

Theorem 1 *Suppose (P) and (D) have strictly feasible solutions ($x > 0$, $s > 0$). Then, for every positive ν, there is a unique solution $(x(\nu), y(\nu), s(\nu))$ to (CPE_ν). These solutions, for all $\nu > 0$, form a smooth path, and as ν approaches 0, $x(\nu)$ and $(y(\nu), s(\nu))$ converge to optimal solutions to (P) and (D) respectively. Moreover, for every $\nu > 0$, $x(\nu)$ is the unique solution to the primal barrier problem*

$$\min \quad c^T x - \nu \sum_j \ln x_j, \qquad Ax = b, \quad x > 0,$$

and $(y(\nu), s(\nu))$ the unique solution to the dual barrier problem

$$\max \quad b^T y + \nu \sum_j \ln s_j, \qquad A^T y + s = c, \quad s > 0.$$

Slide 6

Infeasible case

The theorem leads to nice algorithms ($O(\sqrt{n} \ln(1/\epsilon))$ iterations) *in the strictly feasible case*. What if (P) or (D) infeasible? (Note: if (P) feasible and (D) infeasible, then (P) is unbounded.)

Then we want (approximate) *certificates of infeasibility*. These are guaranteed by the Farkas Lemma.

Slide 7

Farkas Lemma

Lemma 1 *(i)* (P) *(*$Ax = b$, $x \geq 0$*) is infeasible iff* $\exists\,(\bar{y}, \bar{s})$ *with*

$$A^T\bar{y} + \bar{s} = 0, \qquad \bar{s} \geq 0, \qquad b^T\bar{y} > 0.$$

(ii) (D) *(*$A^Ty + s = c$, $s \geq 0$*) is infeasible iff* $\exists\,\bar{x}$ *with*

$$A\bar{x} = 0, \qquad \bar{x} \geq 0, \qquad c^T\bar{x} < 0.$$

Goal

We want an algorithm that will produce *either* (approximately) optimal solutions to (P) and (D) *or* an (approximate) certificate of infeasibility for (P) or (D).

Slide 8

1st approach: homogenization

Consider the Goldman-Tucker system:

$$\begin{array}{rcrcrcrcl} s & = & & - & A^Ty & + & c\tau & \geq & 0, \\ & & Ax & & & - & b\tau & = & 0, \\ \kappa & = & -c^Tx & + & b^Ty & & & \geq & 0, \\ & & x \geq 0, & & y \text{ free} & & \tau \geq 0. & & \end{array}$$

A solution with $\tau > 0$, $\kappa = 0$ gives optimal solutions.

A solution with $\tau = 0$, $\kappa > 0$ gives an infeasibility certificate.

Not clear how to find an approximate solution.

Slide 9

Ye-Todd-Mizuno self-dual problem

(HLP):

$$
\begin{array}{lrcrcrcrcr}
\min & & & & & & & \bar{h}\theta & & \\
s = & & - & A^T y & + & c\tau & - & \bar{c}\theta & \geq & 0, \\
 & Ax & & & - & b\tau & + & \bar{b}\theta & = & 0, \\
\kappa = & -c^T x & + & b^T y & & & + & \bar{g}\theta & \geq & 0, \\
 & \bar{c}^T x & - & \bar{b}^T y & - & \bar{g}\tau & & & = & -\bar{h}, \\
 & x \geq 0, & & y \text{ free}, & & \tau \geq 0, & & \theta \text{ free}, & &
\end{array}
$$

where

$$\bar{b} := b\tau_0 - Ax_0, \ldots.$$

Self-dual. Have strictly feasible initial solution.

Apply favorite feasible interior-point method.

Slide 10

2nd approach: Infeasible-interior-point method

Try to approximate solution to (CPE_ν) directly (*even if there is none!*) by applying a damped Newton method from infeasible interior point (IIP) (x, y, s) ($x > 0$, $s > 0$, eq'ns not satisfied).

Set $\nu := \sigma\mu$, $\mu := s^T x/n$, $\sigma \in [0, 1]$ and get search direction $(\Delta x, \Delta y, \Delta s)$. Then set

$$x_+ := x + \alpha_P \Delta x, \quad y_+ := y + \alpha_D \Delta y, \quad s_+ := s + \alpha_D \Delta s,$$

for some $\alpha_P > 0$ and $\alpha_D > 0$.

Slide 11

This works very well if the problems are strictly feasible ($O(n^2 \ln(1/\epsilon))$ iterations).

But aiming for a non-existent central path if not!

We show that, *implicitly*, the IIP method is trying to find an infeasibility certificate.

Suppose that (P) is strictly infeasible ((D) similar). Start with (x_0, y_0, s_0), $\quad x_0 > 0, \quad s_0 > 0$.

Current iterate (x, y, s).

Assumption 1

$$Ax = \phi A x_0 + (1-\phi)b, \quad x > 0, \quad \phi > 0,$$

$$A^T y + s = c, \quad s > 0, \quad \beta := b^T y > 0.$$

Slide 12

The Farkas optimization problems

We formulate the optimization problem

$$\begin{array}{llrcrcl} (\bar{D}) & \max & (Ax_0)^T \bar{y} & & & & \\ & & A^T \bar{y} & + & \bar{s} & = & 0, \\ & & b^T \bar{y} & & & = & 1, \\ & & & & \bar{s} \geq 0. & & \end{array}$$

This is strictly feasible. Its dual is

$$\begin{array}{llrcrcl} (\bar{P}) & \min & & & \bar{\zeta} & & \\ & & A\bar{x} & + & b\bar{\zeta} & = & Ax_0, \\ & & \bar{x} \geq 0. & & & & \end{array}$$

We use bars to indicate the variables of $(\bar{D})$ and $(\bar{P})$.

Slide 13

Shadow iterates

Note that, from our assumption, $(x/\phi, -(1-\phi)/\phi)$ is a strictly feasible solution to $(\bar{P})$. Also, $(y/\beta, s/\beta)$ is an approximate solution for $(\bar{D})$.

Definition 1 *The shadow iterate corresponding to* (x, y, s) *is given by*

$$(\bar{x}, \bar{\zeta}) := (\frac{x}{\phi}, -\frac{1-\phi}{\phi}), \quad (\bar{y}, \bar{s}) := (\frac{y}{\beta}, \frac{s}{\beta}).$$

Slide 14

Real and shadow iterations

We now wish to compare the results of applying one iteration of the IIP method from (x, y, s) for (P) and (D), and from $(\bar{x}, \bar{\zeta}, \bar{y}, \bar{s})$ for $(\bar{P})$ and $(\bar{D})$.

The idea is shown in the figure below. While the step from (x, y, s) to (x_+, y_+, s_+) is in some sense "following a nonexistent central path," the shadow iterates follow the central path for the strictly feasible pair $(\bar{P})$ and $(\bar{D})$.

Slide 15

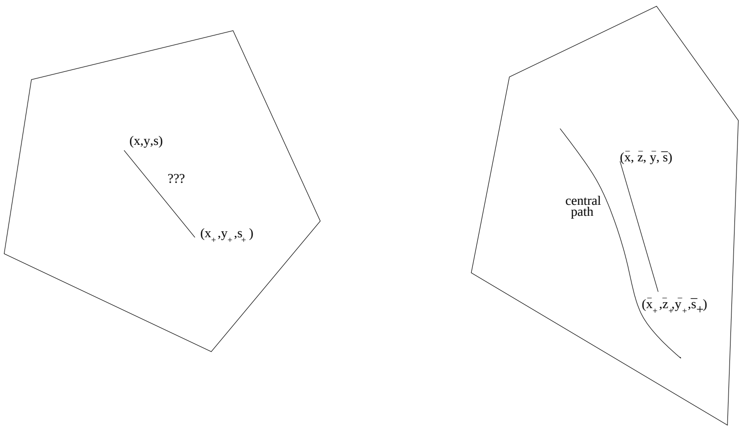

Figure 1. Comparing the real and shadow iterations: a "commutative diagram."

Slide 16

The new iterate (x_+, y_+, s_+) comes from a step (step sizes $\alpha_P > 0$, $\alpha_D > 0$) in direction $(\Delta x, \Delta y, \Delta s)$, which is the solution to the Newton step for $(CPE_{\sigma\mu})$.

Let $(\bar{x}_+, \bar{\zeta}_+, \bar{y}_+, \bar{s}_+)$ be the corresponding shadow iterate. Some algebra shows that, given

Assumption 2 $\Delta\beta := b^T \Delta y$ *is positive,*

then

$(\bar{x}_+, \bar{\zeta}_+) = (\bar{x}, \bar{\zeta}) + \bar{\alpha}_P(\Delta\bar{x}, \Delta\bar{\zeta})$,

$(\bar{y}_+, \bar{s}_+) = (\bar{y}, \bar{s}) + \bar{\alpha}_D(\Delta\bar{y}, \Delta\bar{s})$, where

Slide 17

$$\bar{\alpha}_P := \frac{\alpha_P \Delta\beta}{(1-\alpha_P)\beta},\ \Delta\bar{x} := \frac{\beta}{\phi\Delta\beta}(\Delta x + x),\ \Delta\bar{\zeta} := -\frac{\beta}{\phi\Delta\beta},$$

$$\bar{\alpha}_D := \frac{\alpha_D \Delta\beta}{\beta + \alpha_D\Delta\beta},\ \Delta\bar{y} := \frac{\Delta y}{\Delta\beta} - \bar{y},\ \Delta\bar{s} := \frac{\Delta s}{\Delta\beta} - \bar{s}.$$

Slide 18

Equivalence

Theorem 2 *The directions* $(\Delta\bar{x}, \Delta\bar{\zeta}, \Delta\bar{y}, \Delta\bar{s})$ *defined above solve the Newton step for the central path equations for* $(\bar{P})$ *and* $(\bar{D})$ *for* $\bar{\nu} := \bar{\sigma}\bar{\mu}$*, where* $\bar{\sigma} := \frac{\beta}{\Delta\beta}\sigma$ *and* $\bar{\mu} := \bar{s}^T\bar{x}/n$*.*

This theorem substantiates our main claim. Indeed, the shadow iterates are being generated by damped Newton steps for the problems $(\bar{P})$ and $(\bar{D})$, for which the central path exists.

Further, the argument can be reversed, giving $(\Delta x, \Delta y, \Delta s)$ in terms of $(\Delta\bar{x}, \Delta\bar{\zeta}, \Delta\bar{y}, \Delta\bar{s})$.

Analogous statements hold in the dual infeasible case.

Slide 19

Implications

We assumed $\Delta\beta > 0$. This always holds for sufficiently small σ, but in practice usually holds for all $\sigma \in [0,1]$.

Even if σ is close to one, $\bar{\sigma} := \frac{\beta}{\Delta\beta}\sigma$ is usually close to zero.

Even if α_P close to zero, $\bar{\alpha}_P := \frac{\alpha_P \Delta\beta}{(1-\alpha_P)\beta}$ can be close to one.

Since $\bar{\alpha}_D := \frac{\alpha_D \Delta\beta}{\beta + \alpha_D \Delta\beta}$, to get $\bar{\alpha}_D$ close to one want α_D to approach $+\infty$.

These observations suggest modifications of choices of parameters to more easily detect infeasibility.

Slide 20

Extensions

Development so far only for LP. But all arguments extend directly to more general conic programming problems ($x \in K$, $s \in K^*$) as long as Newton system equations "scale correctly."

E.g., self-scaled cones with Nesterov-Todd direction;

AHO direction for SOCP or SDP;

HKM direction for SOCP or SDP;

Dual HKM direction for SDP.