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

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

together with its dual

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

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

\[
\begin{align*}
A^T y & + s = c, \quad s \geq 0, \\
(OC) \quad Ax & = b, \quad x \geq 0, \\
XSe & = 0,
\end{align*}
\]

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

Central path equations

Path-following interior-point methods iterate approximate solutions to

\[
\begin{align*}
A^T y & + s = c, \quad (s > 0) \\
(CPE_\nu) \quad Ax & = b, \quad (x > 0) \\
XSe & = \nu e,
\end{align*}
\]

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.
Central path theorem

**Theorem 1** Suppose \((P)\) and \((D)\) have strictly feasible solutions \((x > 0, s > 0)\). Then, for every positive \(\nu\), there is a unique solution \((x(\nu), y(\nu), s(\nu))\) to \((\text{CP}_E^\nu)\). These solutions, for all \(\nu > 0\), form a smooth path, and as \(\nu\) 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
\[
\min \quad c^T x - \nu \sum_j \ln x_j, \quad Ax = b, \quad x > 0,
\]
and \((y(\nu), s(\nu))\) the unique solution to the dual barrier problem
\[
\max \quad b^T y + \nu \sum_j \ln s_j, \quad A^T y + s = c, \quad s > 0.
\]

Infeasible case

The theorem leads to nice algorithms \((O(\sqrt{n} \ln(1/\epsilon)) \text{ iterations})\) 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.
Farkas Lemma

**Lemma 1**

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

\[ A^T \bar{y} + \bar{s} = 0, \quad \bar{s} \geq 0, \quad b^T \bar{y} > 0. \]

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

\[ A \bar{x} = 0, \quad \bar{x} \geq 0, \quad c^T \bar{x} < 0. \]

**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 \((D)\).

1st approach: homogenization

Consider the Goldman-Tucker system:

\[
\begin{align*}
    s &= -A^T y + c\tau \geq 0, \\
    Ax &= -b\tau = 0, \\
    \kappa &= -c^T x + b^T y \geq 0, \\
    x &\geq 0, \quad y \text{ free} \quad \tau \geq 0.
\end{align*}
\]

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.
Ye-Todd-Mizuno self-dual problem

(HLP):
\[
\begin{align*}
\text{min} & \quad \bar{h} \theta \\
\text{s.t.} & \quad s = -A^T y + c \tau - \bar{c} \theta \geq 0, \\
& \quad Ax - b \tau + \bar{b} \theta = 0, \\
& \quad \kappa = -c^T x + b^T y + \bar{g} \theta \geq 0, \\
& \quad \bar{c}^T x - \bar{b}^T y - \bar{g} \tau = -\bar{h}, \\
x \geq 0, & \quad y \text{ free, } \quad \tau \geq 0, & \quad \theta \text{ free},
\end{align*}
\]

where
\[
\bar{b} := b_0 - Ax_0, \ldots
\]

Self-dual. Have strictly feasible initial solution.
Apply favorite feasible interior-point method.

2nd approach: Infeasible-interior-point method

Try to approximate solution to \((CPE')\) directly (even if there is none!) by applying a damped Newton method from infeasible interior point (IIP) \((x, y, s)\) \((x > 0, s > 0, \text{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
\[
\begin{align*}
x_+ := x + \alpha_P \Delta x, & \quad y_+ := y + \alpha_D \Delta y, & \quad s_+ := s + \alpha_D \Delta s,
\end{align*}
\]
for some \(\alpha_P > 0 \text{ and } \alpha_D > 0\).
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 \((\bar{D})\) similar). Start with 
\((x_0, y_0, s_0), \ x_0 > 0, \ s_0 > 0.\)

Current iterate \((x, y, s)\).

**Assumption 1**

\[
Ax = \phi Ax_0 + (1 - \phi)b, \ x > 0, \ \phi > 0, \\
A^T y + s = c, \ s > 0, \ \beta := b^T y > 0.
\]

---

**The Farkas optimization problems**

We formulate the optimization problem

\[
\begin{align*}
\bar{D} & \quad \max \quad (Ax_0)^T \bar{y} \\
A^T \bar{y} + \bar{s} & = 0, \\
b^T \bar{y} & = 1, \quad \bar{s} \geq 0.
\end{align*}
\]

This is strictly feasible. Its dual is

\[
\begin{align*}
\bar{P} & \quad \min \quad \bar{\zeta} \\
A\bar{x} + b\bar{\zeta} & = Ax_0, \\
\bar{x} & \geq 0.
\end{align*}
\]

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

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

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 \((CP E_{\sigma\mu})\).

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

\[
(x_+, \zeta_+) = (\bar{x}, \bar{\zeta}) \pm \bar{\alpha}_P (\Delta \bar{x}, \Delta \bar{\zeta}),
\]

\[
(\bar{y}_+, \bar{s}_+) = (\bar{y}, \bar{s}) \pm \bar{\alpha}_D (\Delta \bar{y}, \Delta \bar{s}),\]

where
Slide 17

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

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

Slide 18

**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.
**Implications**

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

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

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

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

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

**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.