A Full-Newton-Step Infeasible Interior-Point Algorithm for Linear Optimization

Kees Roos

e-mail: C.Roos@tudelft.nl
URL: http://www.isa.ewi.tudelft.nl/~roos

TU Delft

Workshop on Advances in Continuous Optimization
Reykjavik, Iceland
Friday June 30 and Saturday July 1, A.D. 2006
Outline

- State-of-the-art in IIPMs
- Usual search direction
- Complexity results for feasible IPM
- Definition of perturbed problems
- Idea underlying the algorithm
- Complexity analysis
- Conjecture
- Some references
State-of-the-art in IIPMs

\[(P) \quad \min \{c^T x : Ax = b, \, x \geq 0\}\]
\[(D) \quad \max \{b^T y : A^T y + s = c, \, s \geq 0\}\]

$A$ is $m \times n$ and $\text{rank}(A) = m$.

The best known iteration bound for IIPMs is

\[
O \left( n \log \frac{\max \{(x^0)^T s^0, \|b - Ax^0\|, \|c - A^T y^0 - s^0\|\}}{\epsilon} \right).
\]

$x^0 > 0$, $y^0 > 0$ and $s^0 > 0$ denote the starting points, and $b - Ax^0$ and $c - A^T y^0 - s^0$ are the initial primal and dual residue vectors, respectively.

Assumption: the initial iterates are $(x^0, y^0, s^0) = \zeta(e, 0, e)$ where $\zeta$ satisfies $\zeta \geq \|x^*, s^*\|_\infty$, for some optimal triple $(x^*, y^*, s^*)$. 
Usual search directions

\[ A\Delta x = b - Ax \]
\[ A^T\Delta y + \Delta s = c - A^T y - s \]
\[ s\Delta x + x\Delta s = \mu e - xs. \]

When moving from \( x \) to \( x^+ := x + \Delta x \) the new iterate \( x^+ \) satisfies \( Ax^+ = b \), but not necessarily \( x^+ \geq 0 \). To keep the iterate positive, one uses a damped step \( x^+ := x + \alpha \Delta x \), where \( \alpha < 1 \) denotes the step size.

Our search directions are designed in such a way that full Newton steps can be used; each iteration reduces the sizes of the residual vectors with exactly the same speed as the duality gap.

Our aim is to show that this is a promising strategy.
Feasible full-Newton step IPMs

Central path

\[ Ax = b, \quad x \geq 0 \]
\[ A^T y + s = c, \quad s \geq 0 \]
\[ xs = \mu e. \]

If this system has a solution, for some \( \mu > 0 \), then a solution exists for every \( \mu > 0 \), and this solution is unique.

This happens if and only if \( (P) \) and \( (D) \) satisfy the interior-point condition (IPC), i.e., if \( (P) \) has a feasible solution \( x > 0 \) and \( (D) \) has a solution \( (y, s) \) with \( s > 0 \).

If the IPC is satisfied, then the solution is denoted as \( (x(\mu), y(\mu), s(\mu)) \), and called the \( \mu \)-center of \( (P) \) and \( (D) \). The set of all \( \mu \)-centers is called the central path of \( (P) \) and \( (D) \). As \( \mu \) goes to zero, \( x(\mu), y(\mu) \) and \( s(\mu) \) converge to optimal solution of \( (P) \) and \( (D) \).

The system is hard to solve, but by applying Newton’s method one can easily find approximate solutions.
Definition of the Newton step

Given a primal feasible \( x > 0 \), and dual feasible \( y \) and \( s > 0 \), we want to find displacements \( \Delta x \), \( \Delta y \) and \( \Delta s \) such that

\[
A(x + \Delta x) = b, \\
A^T(y + \Delta y) + s + \Delta s = c, \\
(x + \Delta x)(s + \Delta s) = \mu e.
\]

Neglecting the quadratic term \( \Delta x \Delta s \) in the third equation, and using \( b - Ax = 0 \) and \( c - A^Ty - s = 0 \), we obtain the following linear system of equations in the search directions \( \Delta x \), \( \Delta y \) and \( \Delta s \).

\[
A\Delta x = 0, \\
A^T\Delta y + \Delta s = 0, \\
s\Delta x + x\Delta s = \mu e - xs.
\]

Since \( A \) has full rank, and the vectors \( x \) and \( s \) are positive, the coefficient matrix in this linear system is nonsingular. Hence the system uniquely defines the search directions \( \Delta x \), \( \Delta y \) and \( \Delta s \). These search directions are used in all existing primal-dual (feasible) IPMs and called after Newton.
Properties of the Newton step

\[
\begin{align*}
A \Delta x &= 0, \\
A^T \Delta y + \Delta s &= 0, \\
\Delta x + x \Delta s &= \mu e - xs.
\end{align*}
\]

The new iterates are given by

\[
\begin{align*}
x^+ &= x + \Delta x, \\
y^+ &= y + \Delta y, \\
s^+ &= s + \Delta s.
\end{align*}
\]

An important observation is that \( \Delta x \) lies in the null space of \( A \), whereas \( \Delta s \) belongs to the row space of \( A \). This implies that \( \Delta x \) and \( \Delta s \) are orthogonal, i.e.,

\[
(\Delta x)^T \Delta s = 0.
\]

As a consequence we have the important property that after a full Newton step the duality gap assumes the same value as at the \( \mu \)-centers, namely \( n\mu \).

**Lemma 1** After a (feasible!) primal-dual Newton step one has \( (x^+)^T s^+ = n\mu \).
Fundamental results

We use the quantity $\delta(x, s; \mu)$ to measure proximity of a feasible triple $(x, y, s)$ to the $\mu$-center $(x(\mu), y(\mu), s(\mu))$.

$$\delta(x, s; \mu) := \delta(v) := \frac{1}{2} \|v - v^{-1}\|$$

where $v := \sqrt{\frac{xs}{\mu}}$.

Lemma 2 Let $(x, s)$ be a positive primal-dual pair and $\mu > 0$ such that $x^Ts = n\mu$. Moreover, let $\delta := \delta(x, s; \mu)$ and let $\mu^+ = (1 - \theta)\mu$. Then

$$\delta(x, s; \mu^+)^2 = (1 - \theta)\delta^2 + \frac{\theta^2n}{4(1 - \theta)}.$$  

Lemma 3 If $\delta := \delta(x, s; \mu) \leq 1$, then the primal-dual Newton step is feasible, i.e., $x^+$ and $s^+$ are nonnegative. Moreover, if $\delta < 1$, then $x^+$ and $s^+$ are positive and

$$\delta(x^+, s^+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^2)}}.$$  

Corollary 1 If $\delta := \delta(x, s; \mu) \leq \frac{1}{\sqrt{2}}$, then $\delta(x^+, s^+; \mu) \leq \delta^2$.  


Primal-Dual full-Newton step feasible IPM

Input:

Accuracy parameter $\varepsilon > 0$;
barrier update parameter $\theta$, $0 < \theta < 1$;
feasible $(x^0, y^0, s^0)$ with $(x^0)^T s^0 = n\mu^0$, $\delta(x^0, s^0, \mu^0) \leq 1/2$.

begin
$x := x^0$; $y := y^0$; $s := s^0$; $\mu := \mu^0$;
while $x^T s \geq \varepsilon$ do
begin
$\mu$-update:
$\mu := (1 - \theta)\mu$;
centering step:
$(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$;
end
end

Theorem 1 If $\theta = 1/\sqrt{2n}$, then the algorithm requires at most $\sqrt{2n} \log \frac{n \sqrt{\mu^0}}{\epsilon}$ iterations. The output is a primal-dual pair $(x, s)$ such that $x^T s \leq \varepsilon$. 
Graphical illustration primal-dual full-Newton-step path-following method

One iteration.

\[(1 - \theta) \mu e, \bar{z}^1\]
Definition of perturbed problems

We start with choosing arbitrarily \( x^0 > 0 \) and \( y^0, s^0 > 0 \) such that \( x^0 s^0 = \mu_0 e \) for some (positive) number \( \mu_0 \). For any \( \nu \) with \( 0 < \nu \leq 1 \) we use the perturbed problem \( (P_\nu) \), defined by

\[
(P_\nu) \quad \min \left\{ \left( c - \nu \left( c - y^0 - s^0 \right) \right)^T x : Ax = b - \nu \left( b - Ax^0 \right), \ x \geq 0 \right\},
\]

and its dual problem \( (D_\nu) \), which is given by

\[
(D_\nu) \quad \min \left\{ \left( b - \nu \left( b - Ax^0 \right) \right)^T y : A^T y + s = c - \nu \left( c - y^0 - s^0 \right), \ s \geq 0 \right\}.
\]

Note that if \( \nu = 1 \) then
\[
x = x^0 \text{ yields a strictly feasible solution of } (P_\nu), \text{ and }
\]
\[
(y, s) = (y^0, s^0) \text{ a strictly feasible solution of } (D_\nu).
\]

Conclusion: if \( \nu = 1 \) then \( (P_\nu) \) and \( (D_\nu) \) satisfy the IPC.
Perturbed problems satisfy IPC

**Lemma 4** The original problems, \((P)\) and \((D)\), are feasible if and only if for each \(\nu\) satisfying \(0 < \nu \leq 1\) the perturbed problems \((P_\nu)\) and \((D_\nu)\) satisfy the IPC.

**Proof:** Let \(x^*\) be feasible for \((P)\) and \((y^*, s^*)\) for \((D)\). Then \(Ax^* = b\) and \(A^T y^* + s^* = c\), with \(x^* \geq 0\) and \(s^* \geq 0\). Now let \(0 < \nu \leq 1\), and consider

\[
    x = (1 - \nu) x^* + \nu x^0, \quad y = (1 - \nu) y^* + \nu y^0, \quad s = (1 - \nu) s^* + \nu s^0.
\]

One has

\[
    Ax = A \left( (1 - \nu) x^* + \nu x^0 \right) = (1 - \nu) b + \nu Ax^0 = b - \nu \left( b - Ax^0 \right),
\]

showing that \(x\) is feasible for \((P_\nu)\). Similarly,

\[
    A^T y + s = (1 - \nu) \left( A^T y^* + s^* \right) + \nu \left( A^T y^0 + s^0 \right) = c - \nu \left( c - A^T y^0 - s^0 \right),
\]

showing that \((y, s)\) is feasible for \((D_\nu)\). Since \(\nu > 0\), \(x\) and \(s\) are positive, thus proving that \((P_\nu)\) and \((D_\nu)\) satisfy the IPC.

To prove the inverse implication, suppose that \((P_\nu)\) and \((D_\nu)\) satisfy the IPC for each \(\nu\) satisfying \(0 < \nu \leq 1\). Obviously, then \((P_\nu)\) and \((D_\nu)\) are feasible for these values of \(\nu\). Letting \(\nu\) go to zero it follows that \((P)\) and \((D)\) are feasible. \(\blacksquare\)
Central path of the perturbed problems

We assume that \((P)\) and \((D)\) are feasible. Letting \(0 < \nu \leq 1\), the problems \((P_\nu)\) and \((D_\nu)\) satisfy the IPC, and hence their central paths exist. This means that the system

\[
\begin{align*}
b - Ax &= \nu(b - Ax^0), \quad x \geq 0 \\
c - A^T y - s &= \nu(c - A^T y^0 - s^0), \quad s \geq 0 \\
xs &= \mu e
\end{align*}
\]

has a unique solution, for every \(\mu > 0\). In the sequel this unique solution is denoted as

\[(x(\mu, \nu), y(\mu, \nu), s(\mu, \nu)).\]

These are the \(\mu\)-centers of the perturbed problems \((P_\nu)\) and \((D_\nu)\).

Note that since \(x^0s^0 = \mu^0 e\), \(x^0\) is the \(\mu^0\)-center of the perturbed problem \((P_1)\) and \((y^0, s^0)\) the \(\mu^0\)-center of \((D_1)\). In other words,

\[(x(\mu^0, 1), y(\mu^0, 1), s(\mu^0, 1)) = (x^0, y^0, s^0).\]

In the sequel we will always take \(\mu = \mu^0 \nu\).
Idea underlying the algorithm

We just established that if \( \nu = 1 \) and \( \mu = \mu^0 \), then \( x = x^0 \) is the \( \mu \)-center of the perturbed problem \((P_\nu)\) and \((y, s) = (y^0, s^0)\) the \( \mu \)-center of \((D_\nu)\). These are our initial iterates.

We measure proximity to the \( \mu \)-center of the perturbed problems by the quantity \( \delta(x, s; \mu) \). So, initially, \( \delta(x, s; \mu) = 0 \). We assume that at the start of each of the subsequent iterations \( \delta(x, s; \mu) \leq \tau \), where \( \tau > 0 \). This is certainly true for the first iteration.

Now we describe one iteration of our algorithm.

Suppose that for some \( \mu \in (0, \mu^0] \) we have \( x, y \) and \( s \) that are feasible for \((P_\nu)\) and \((D_\nu)\), with \( \mu = \mu^0 \nu \), and such that \( x^T s = n \mu \) and \( \delta(x, s; \mu) \leq \tau \).

Roughly spoken, one iteration of the algorithm does the following.

Reduce \( \mu \) to \( \mu^+ = (1 - \theta)\mu \) and \( \nu \) to \( \nu^+ = (1 - \theta)\nu \), with \( \theta \in (0, 1) \). Find new iterates \( x^+, y^+ \) and \( s^+ \) that are feasible for \((P_{\nu^+})\) and \((D_{\nu^+})\), and such that \( x^T s = n \mu^+ \) and \( \delta(x^+, s^+; \mu^+) \leq \tau \). Note that \( \mu^+ = \mu^0 \nu^+ \).
Graphical illustration of one iteration

\[ \delta(v) \leq \frac{1}{\sqrt{2}} \]

Feasibility step and subsequent centering steps \((\mu^+ = (1 - \theta)\mu, \nu^+ = (1 - \theta)\nu)\).
We denote the initial values of the primal and dual residuals as $r_b^0$ and $r_c^0$, respectively, as

$$r_b^0 = b - Ax^0, \quad r_c^0 = c - A^T y^0 - s^0.$$ 

Then the feasibility equations for $(P_\nu)$ and $(D_\nu)$ are given by

$$b - Ax = \nu r_b^0, \quad x \geq 0$$

$$c - A^T y - s = \nu r_c^0, \quad s \geq 0.$$ 

and those of $(P_{\nu+})$ and $(D_{\nu+})$ by

$$b - Ax = \nu^+ r_b^0, \quad x \geq 0$$

$$c - A^T y - s = \nu^+ r_c^0, \quad s \geq 0.$$
Feasibility step

Let $x$ be feasible for $(P_\nu)$ and $(y, s)$ for $(D_\nu)$. To get iterates feasible for $(P_{\nu^+})$ and $(D_{\nu^+})$ we need $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$ such that

\[
 b - A(x + \Delta^f x) = \nu + r_0^b \\
 c - A^T(y + \Delta^f y) - (s + \Delta^f s) = \nu + r_0^c.
\]

Since $x$ is feasible for $(P_\nu)$ and $(y, s)$ for $(D_\nu)$, it follows that

\[
 A\Delta^f x = (b - Ax) - \nu + r_0^b = \nu r_0^b - \nu + r_0^b = \theta \nu r_0^b \\
 A^T\Delta^f y + \Delta^f s = (c - A^T y - s) - \nu + r_0^c = \nu r_0^c - \nu + r_0^c = \theta \nu r_0^c.
\]

Therefore, the following system is used to define $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$:

\[
 A\Delta^f x = \theta \nu r_0^b \\
 A^T\Delta^f y + \Delta^f s = \theta \nu r_0^c \\
 s\Delta^f x + x\Delta^f s = \mu e - xs.
\]

The new iterates are given by

\[
 (x^f, y^f, s^f) = (x + \Delta^f x, y + \Delta^f y, s + \Delta^f s).
\]
Centering steps

After the feasibility step the iterates are feasible for \((P_{\nu^+})\) and \((D_{\nu^+})\) with \(\nu = \nu^+\). The hard part in the analysis will be to guarantee that these iterates lie in the region where Newton’s method is quadratically convergent, i.e., \(\delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2}\).

After the feasibility step we perform centering steps in order to get iterates that moreover satisfy \(x^T s = n\mu^+\) and \(\delta(x, s; \mu^+) \leq \tau\). By using the quadratic convergence of the Newton step the required number of centering steps can be easily be obtained. Because, assuming \(\delta = \delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2}\), after \(k\) centering steps we will have iterates \((x, y, s)\) that are still feasible for \((P_{\nu^+})\) and \((D_{\nu^+})\) and such that

\[
\delta(x, s; \mu^+) \leq \left(\frac{1}{\sqrt{2}}\right)^{2^k}.
\]

So \(k\) must satisfy

\[
\left(\frac{1}{\sqrt{2}}\right)^{2^k} \leq \tau,
\]

which certainly holds if

\[
k \geq \log_2 \left(\log_2 \frac{1}{\tau^2}\right).
\]
Primal-Dual infeasible full-Newton step IPM

Input:
Accuracy parameter \( \varepsilon > 0 \);
barrier update parameter \( \theta, 0 < \theta < 1 \)
threshold parameter \( \tau > 0 \).

\[
\begin{align*}
x & := x^0 > 0; \\
y & := y^0; \\
s & := s^0 > 0; \\
x^0 s^0 & = \mu^0 e; \\
\mu & = \mu^0;
\end{align*}
\]

while \( \max(x^T s, \|b - Ax\|, \|c - A^T y - s\|) \geq \varepsilon \) do

begin
feasibility step: \( (x, y, s) := (x, y, s) + (\Delta^f x, \Delta^f y, \Delta^f s) \);
\( \mu \)-update: \( \mu := (1 - \theta)\mu \);
centering steps:

while \( \delta(x, s; \mu) \geq \tau \) do

\( (x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s) \);

endwhile

end

end
Analysis of the feasibility step

Defining

\[ v = \sqrt{\frac{x_s}{\mu}}, \quad d_x := \frac{v \Delta^f x}{x}, \quad d_s := \frac{v \Delta^f s}{s}, \]

we have

\[ x^f = x + \Delta^f x = x + \frac{xd_x}{v} = \frac{x}{v}(v + d_x) \]
\[ s^f = s + \Delta^f s = s + \frac{sd_s}{v} = \frac{s}{v}(v + d_s). \]

Hence

\[ x^f s^f = xs + \left( s \Delta^f x + x \Delta^f s \right) + \Delta^f x \Delta^f s = \mu e + \Delta^f x \Delta^f s = \mu (e + d_x d_s). \]

Since \(\mu v^2 = xs\), after division of both sides by \(\mu^+\) we get

\[ \left( v^f \right)^2 = \frac{x^f s^f}{\mu^+} = \frac{\mu}{\mu^+} e + \frac{1}{\mu^+} \frac{xd_x}{v} \frac{sd_s}{v} = \frac{e + d_x d_s}{1 - \theta}. \]

**Lemma 5** The iterates \((x^f, y^f, s^f)\) are strictly feasible if and only if \(e + d_x d_s > 0\).

**Corollary 2** The iterates \((x^f, y^f, s^f)\) are certainly strictly feasible if \(\|d_x d_s\|_\infty < 1\).
Proximity after the feasibility step

\[ \delta(x^f, s^f; \mu^+) = \delta(v^f) = \frac{1}{2} \left\| v^f - \frac{e}{v_f^f} \right\|, \quad \text{where} \quad (v^f)^2 = \frac{e + d_x d_s}{1 - \theta}. \]

In the sequel we denote

\[ \omega(v) := \frac{1}{2} \sqrt{\| d_x \|^2 + \| d_s \|^2}. \]

This implies \( \| d_x \| \leq 2 \omega(v) \) and \( \| d_s \| \leq 2 \omega(v) \), and moreover,

\[ d^T_x d_s \leq \| d_x \| \| d_s \| \leq \frac{1}{2} \left( \| d_x \|^2 + \| d_s \|^2 \right) \leq 2 \omega(v)^2 \]

\[ \| d_x d_s \|_\infty \leq \| d_x \| \| d_s \| \leq 2 \omega(v)^2. \]

**Lemma 6** If \( \omega(v) < 1/\sqrt{2} \) then the iterates \( (x^f, y^f, s^f) \) are strictly feasible.

**Lemma 7** Let \( \omega(v) < 1/\sqrt{2} \). Then one has

\[ 4 \delta(v^f)^2 \leq \frac{\theta^2 n}{1 - \theta} + \frac{2 \omega(v)^2}{1 - \theta} + (1 - \theta) \frac{2 \omega(v)^2}{1 - 2 \omega(v)^2}. \]
Choice of $\theta$

In order to have $\delta(v^f) \leq 1/\sqrt{2}$, it follows from Lemma 7 that it suffices if

$$\frac{\theta^2 n}{1 - \theta} + \frac{2\omega(v)^2}{1 - \theta} + (1 - \theta) \frac{2\omega(v)^2}{1 - 2\omega(v)^2} \leq 2. \quad (1)$$

Lemma 8 Let $\omega(v) \leq \frac{1}{2}$ and

$$\theta = \frac{\alpha}{\sqrt{2n}}, \quad 0 \leq \alpha \leq \sqrt{\frac{n}{n + 1}}. \quad (2)$$

Then the iterates $(x^f, y^f, s^f)$ are strictly feasible and $\delta(v^f) \leq \frac{1}{\sqrt{2}}$.

At this stage we decide to choose $\theta := \frac{\alpha}{\sqrt{2n}}, \alpha \leq \frac{1}{\sqrt{2}}$. Then we have

$$\omega(v) = \frac{1}{2} \sqrt{||d_x||^2 + ||d_s||^2} \leq \frac{1}{2} \Rightarrow \delta(v^f) \leq \frac{1}{\sqrt{2}}.$$

We proceed by considering the vectors $d_x$ and $d_s$ more in detail.
The scaled search directions $d_x$ and $d_s$

\[
\begin{align*}
A\Delta^f x &= \theta\nu r^0_b \\
A^T \Delta^f y + \Delta^f s &= \theta\nu r^0_c \\
s\Delta^f x + x\Delta^f s &= \mu e - xs.
\end{align*}
\]

\[v = \sqrt{\frac{xs}{\mu}}, \quad d_x := \frac{v\Delta^f x}{x}, \quad d_s := \frac{v\Delta^f s}{s}.
\]

The system defining the search directions $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$, can be expressed in terms of $d_x$ and $d_s$ as follows.

\[
\begin{align*}
\bar{A} d_x &= \theta\nu r^0_b, \\
\bar{A}^T \frac{\Delta^f y}{\mu} + d_s &= \theta\nu vs^{-1} r^0_c, \\
d_x + d_s &= v^{-1} - v,
\end{align*}
\]

where

\[\bar{A} = AV^{-1}X.\]
Geometric interpretation of $\omega(v)$

\[ \{ \xi : A\xi = \theta \nu v_s^0 \} \]

\[ \theta \nu v_s^{-1} r^0 + \{ A^T \eta : \eta \in \mathbb{R}^m \} \]

\[ d_x \]

\[ d_s \]

\[ r \]

\[ 0 \]

\[ 2\omega(v) \]

\[ 2\delta(v) \]

\[ 90^\circ \]
Upper bound for $\omega(v)$

Let us denote the null space of the matrix $\bar{A}$ as $\mathcal{L}$. So,

$$\mathcal{L} := \{ \xi \in \mathbb{R}^n : \bar{A}\xi = 0 \}.$$

Obviously, the affine space $\{ \xi \in \mathbb{R}^n : \bar{A}\xi = \theta \nu r_0^0 \}$ equals $d_x + \mathcal{L}$. The row space of $\bar{A}$ equals the orthogonal complement $\mathcal{L}^\perp$ of $\mathcal{L}$, and $d_s \in \theta \nu \nu s^{-1} r_c^0 + \mathcal{L}^\perp$.

**Lemma 9** Let $q$ be the (unique) point in the intersection of the affine spaces $d_x + \mathcal{L}$ and $d_s + \mathcal{L}^\perp$. Then

$$2\omega(v) \leq \sqrt{\|q\|^2 + (\|q\| + 2\delta(v))^2}.$$

To simplify the presentation we will denote $\delta(x, s; \mu)$ below simply as $\delta$ ($\delta \leq \tau$). Recall that we need $2\omega(v) \leq 1$. Thus we require $\|q\|$ to satisfy

$$\|q\|^2 + (\|q\| + 2\delta)^2 \leq 1.$$
Upper bound for $\|q\|$ (1)

Recall that $q$ is the (unique) solution of the system

\[
\begin{align*}
\bar{A}q &= \theta \nu r_b^0, \\
\bar{A}^T \xi + q &= \theta \nu \nu s^{-1} r_c^0.
\end{align*}
\]

We proceed by deriving an upper bound for $\|q\|$. From the definition of $\bar{A}$ we deduce that $\bar{A} = \sqrt{\mu} AD$, where

\[
D = \text{diag} \left( \frac{x v^{-1}}{\sqrt{\mu}} \right) = \text{diag} \left( \sqrt{\frac{x}{s}} \right) = \text{diag} \left( \sqrt{\mu} \nu s^{-1} \right).
\]

For the moment, let us write

\[
 r_b = \theta \nu r_b^0, \quad r_c = \theta \nu r_c^0
\]

Then the system defining $q$ is equivalent to

\[
\begin{align*}
\sqrt{\mu} AD q &= r_b, \\
\sqrt{\mu} DA^T \xi + q &= \frac{1}{\sqrt{\mu}} Dr_c.
\end{align*}
\]
Upper bound for $\|q\|$ (2)

This implies $\sqrt{\mu} q = q_1 + q_2$, where

$$q_1 := \left( I - DAT \left( AD^2 A^T \right)^{-1} AD \right) Dr_c, \quad q_2 := DAT \left( AD^2 A^T \right)^{-1} r_b.$$

Let $(y^*, s^*)$ be such that $A^T y^* + s^* = c$ and $x^*$ such that $Ax^* = b$. Then

$$r_b = \theta \nu r_b^0 = \theta \nu (b - Ax^0) = \theta \nu A(x^* - x^0) = \theta \nu AD \left( D^{-1} \left( x^* - x^0 \right) \right)$$

$$r_c = \theta \nu r_c^0 = \theta \nu \left( c - A^T y^0 - s^0 \right) = \theta \nu \left( A^T (y^* - y^0) + s^* - s^0 \right).$$

Since $DAT(y^* - y^0)$ belongs to the row space of $AD$, we obtain

$$\|q_1\| \leq \theta \nu \|D (s^* - s^0)\|, \quad \|q_2\| \leq \theta \nu \|D^{-1} (x^* - x^0)\|.$$

Since $\sqrt{\mu} q = q_1 + q_2$ and $q_1$ and $q_2$ are orthogonal, we may conclude that

$$\sqrt{\mu} \|q\| \leq \theta \nu \sqrt{\|D (s^* - s^0)\|^2 + \|D^{-1} (x^* - x^0)\|^2},$$

where, as always, $\mu = \mu_0 \nu$.

We are still free to choose $x^*$ and $s^*$, subject to the constraints $Ax^* = b$ and $A^T y^* + s^* = c$. 

Upper bound for $\|q\|$ (3)

$$\sqrt{\mu} \|q\| \leq \theta \nu \sqrt{\|D (s^* - s^0)\|^2 + \|D^{-1} (x^* - x^0)\|^2}$$

Let $x^*$ be an optimal solution of $(P)$ and $(y^*, s^*)$ of $(D)$ (assuming that these exist). Then $x^*$ and $s^*$ are complementary, i.e., $x^* s^* = 0$. Let $\zeta$ be such that $\|x^* + s^*\|_\infty \leq \zeta$. Starting the algorithm with

$$x^0 = s^0 = \zeta e, \quad y^0 = 0, \quad \mu^0 = \zeta^2,$$

the entries of the vectors $x^0 - x^*$ are $s^0 - s^*$ satisfy

$$0 \leq x^0 - x^* \leq \zeta e, \quad 0 \leq s^0 - s^* \leq \zeta e.$$ 

Thus it follows that

$$\sqrt{\|D (s^* - s^0)\|^2 + \|D^{-1} (x^* - x^0)\|^2} \leq \zeta \sqrt{\|De\|^2 + \|D^{-1}e\|^2} = \zeta \sqrt{e^T \left( \frac{x}{s} + \frac{s}{x} \right)}.$$

Substitution gives

$$\sqrt{\mu} \|q\| \leq \theta \nu \zeta \sqrt{e^T \left( \frac{x}{s} + \frac{s}{x} \right)}.$$
Bounds for $x$ and $s$; choice of $\tau$

$$\sqrt{\mu} \|q\| \leq \theta \nu \zeta \sqrt{e^T \left( \frac{x}{s} + \frac{s}{x} \right)}.$$

Recall that $x$ is feasible for $(P_\nu)$ and $(y, s)$ for $(D_\nu)$ and, moreover $\delta(x, s; \mu) \leq \tau$. Based on his information we need to estimate the sizes of the entries of the vectors $x/s$ and $s/x$. Since the concerning result does not belong to the mainstream of the analysis we mention it without proof. If

$$\tau = \frac{1}{8}$$

and $\delta \leq \tau$ then we have

$$\sqrt{\frac{x}{s}} \leq \sqrt{2} \frac{x(\mu, \nu)}{\sqrt{\mu}}, \quad \sqrt{\frac{s}{x}} \leq \sqrt{2} \frac{s(\mu, \nu)}{\sqrt{\mu}}.$$
Upper bound for $\|q\|$ (4)

$$\sqrt{\mu} \|q\| \leq \theta \nu \zeta \sqrt{e^T \left( \frac{x}{s} + \frac{s}{x} \right)}, \quad \sqrt{\frac{x}{s}} \leq \sqrt{2} \frac{x(\mu, \nu)}{\sqrt{\mu}}, \quad \sqrt{\frac{s}{x}} \leq \sqrt{2} \frac{s(\mu, \nu)}{\sqrt{\mu}}.$$ 

Substitution gives

$$\sqrt{\mu} \|q\| \leq \theta \nu \zeta \sqrt{2e^T \left( \frac{x(\mu, \nu)^2}{\mu} + \frac{s(\mu, \nu)^2}{\mu} \right)}.$$ 

This implies

$$\mu \|q\| \leq \theta \nu \zeta \sqrt{2} \sqrt{\|x(\mu, \nu)\|^2 + \|s(\mu, \nu)\|^2}.$$ 

Therefore, also using $\mu = \mu^0 \nu = \zeta^2 \nu$ and $\theta = \frac{\alpha}{\sqrt{2n}}$, we obtain the following upper bound for the norm of $q$:

$$\|q\| \leq \frac{\alpha}{\zeta \sqrt{n}} \sqrt{\|x(\mu, \nu)\|^2 + \|s(\mu, \nu)\|^2}.$$
Choice of $\alpha$

\[ \|q\| \leq \frac{\alpha}{\zeta \sqrt{n}} \sqrt{\|x(\mu, \nu)\|^2 + \|s(\mu, \nu)\|^2}. \]

We define

\[ \kappa(\zeta, \nu) = \frac{\sqrt{\|x(\mu, \nu)\|^2 + \|s(\mu, \nu)\|^2}}{\zeta \sqrt{2n}}, \quad 0 < \nu \leq 1, \mu = \mu^0 \nu \]

and

\[ \bar{\kappa}(\zeta) = \max_{0 < \nu \leq 1} \kappa(\zeta, \nu). \]

Then $\|q\|$ can be bounded above as follows.

\[ \|q\| \leq \alpha \bar{\kappa}(\zeta) \sqrt{2}. \]

We found in (1) that in order to have $\delta(v^f) \leq 1/\sqrt{2}$, we should have $\|q\|^2 + (\|q\| + 2\delta)^2 \leq 1$. Therefore, since $\delta \leq \tau = \frac{1}{8}$, it suffices if $q$ satisfies $\|q\|^2 + (\|q\| + \frac{1}{4})^2 \leq 1$. So we certainly have $\delta(v^f) \leq 1/\sqrt{2}$ if $\|q\| \leq \frac{1}{2}$. Since $\|q\| \leq \alpha \bar{\kappa}(\zeta) \sqrt{2}$, the latter inequality is satisfied if we take

\[ \alpha = \frac{1}{2 \sqrt{2} \bar{\kappa}(\zeta)}. \]

Note that since $x(\zeta^2, 1) = s(\zeta^2, 1) = \zeta e$, we have $\kappa(\zeta, 1) = 1$. As a consequence we obtain that $\bar{\kappa}(\zeta) \geq 1$. We can prove that $\bar{\kappa}(\zeta) \leq \sqrt{2n}$. 

30
Bound for $\bar{\kappa}(\zeta)$ (1)

We have

\[ Ax^* = b, \quad 0 \leq x^* \leq \zeta e \]
\[ A^T y^* + s^* = c, \quad 0 \leq s^* \leq \zeta e \]
\[ x^* s^* = 0. \]

To simplify notation, we denote $x = x(\mu, \nu)$, $y = y(\mu, \nu)$ and $s = s(\mu, \nu)$. Then

\[ b - Ax = \nu(b - A\zeta e), \quad x \geq 0 \]
\[ c - A^T y - s = \nu(c - \zeta e), \quad s \geq 0 \]
\[ xs = \nu\zeta^2 e. \]

This implies

\[ Ax^* - Ax = \nu(Ax^* - A\zeta e), \quad x \geq 0 \]
\[ A^T y^* + s^* - A^T y - s = \nu(A^T y^* + s^* - \zeta e), \quad s \geq 0 \]
\[ xs = \nu\zeta^2 e. \]
Bound for $\bar{\kappa}(\zeta)$ (2)

\[
Ax^* - Ax = \nu(Ax^* - A\zeta), \quad x \geq 0
\]
\[
A^Ty^* + s^* - A^Ty - s = \nu(A^Ty^* + s^* - \zeta), \quad s \geq 0
\]
\[
xs = \nu\zeta^2 e.
\]

We rewrite this system as

\[
A (x^* - x - \nu x^* + \nu \zeta e) = 0, \quad x \geq 0
\]
\[
A^T (y^* - y - \nu y^* + \nu \zeta e) = s - s^* + \nu s^* - \nu \zeta e, \quad s \geq 0
\]
\[
xs = \nu\zeta^2 e.
\]

From this we deduce that

\[
(x^* - x - \nu x^* + \nu \zeta e)^T (s - s^* + \nu s^* - \nu \zeta e) = 0.
\]

Define

\[
a := (1 - \nu)x^* + \nu \zeta e, \quad b := (1 - \nu)s^* + \nu \zeta e.
\]

We then have $a \geq \nu \zeta e$, $b \geq \nu \zeta e$ and $(a - x)^T (b - s) = 0$. The latter gives

\[
a^T b + x^T s = a^T s + b^T x.
\]
Bound for $\bar{\kappa}(\zeta)$ (3)

Since $x^* T s^* = 0$, $x^* + s^* \leq \zeta e$ and $xs = \nu \zeta^2 e$, we may write

$$a^T b + x^T s = ((1 - \nu)x^* + \nu \zeta e)^T ((1 - \nu)s^* + \nu \zeta e) + \nu \zeta^2 n = \nu (1 - \nu) (x^* + s^*)^T \zeta e + \nu^2 \zeta^2 n + \nu \zeta^2 n \leq \nu (1 - \nu) \zeta^2 n + \nu^2 \zeta^2 n + \nu \zeta^2 n = 2 \nu \zeta^2 n.$$ 

Moreover, also using $a \geq \nu \zeta e$, $b \geq \nu \zeta e$, we get

$$a^T s + b^T x = ((1 - \nu)x^* + \nu \zeta e)^T s + ((1 - \nu)s^* + \nu \zeta e)^T x = (1 - \nu) (s^T x^* + x^T s^*) + \nu \zeta e^T (x + s) \geq \nu \zeta (\|s\|_1 + \|x\|_1).$$

Hence $\|s\|_1 + \|x\|_1 \leq 2 \zeta n$. Since $\|x\|^2 + \|s\|^2 \leq (\|s\|_1 + \|x\|_1)^2$, it follows that

$$\frac{\sqrt{\|x\|^2 + \|s\|^2}}{\zeta \sqrt{2n}} \leq \frac{\|s\|_1 + \|x\|_1}{\zeta \sqrt{2n}} \leq \frac{2 \zeta n}{\zeta \sqrt{2n}} = \sqrt{2n},$$

thus proving $\bar{\kappa}(\zeta) \leq \sqrt{2n}$. 

33
Complexity analysis

We have found values for the parameters in the algorithm such that if at the start of an iteration the iterates satisfy \( \delta(x, s; \mu) \leq \tau = \frac{1}{8} \), then after the feasibility step the iterates satisfy \( \delta(x, s; \mu^+) \leq \frac{1}{\sqrt{2}} \). At most

\[
\log_2 \left( \log_2 \frac{1}{\tau^2} \right) = \log_2 \left( \log_2 64 \right) \leq 3
\]

centering steps suffice to get iterates that satisfy \( \delta(x, s; \mu^+) \leq \tau \). So each (main) iteration consists of at most 4 so-called inner iterations, in each of which we need to compute a new search direction. In each iteration both the duality gap and the residual vectors are reduced by the factor \( 1 - \theta \). Hence, using \( x^0^T s^0 = n\zeta^2 \) and defining \( K := \max \{ n\zeta^2, \| r^0_b \|, \| r^0_c \| \} \), the total number of main iterations is bounded above by

\[
\frac{1}{\theta} \log \frac{K}{\varepsilon} = \frac{\sqrt{2n}}{\alpha} \log \frac{K}{\varepsilon} = 4\bar{\kappa}(\zeta)\sqrt{2n} \log \frac{K}{\varepsilon}.
\]

The total number of inner iterations is therefore bounded above by

\[
16\bar{\kappa}(\zeta)\sqrt{2n} \log \frac{\max \{ n\zeta^2, \| r^0_b \|, \| r^0_c \| \} }{\varepsilon},
\]

where \( \bar{\kappa}(\zeta) \leq \sqrt{2n} \).
Conjecture

**Conjecture 1** If \((P)\) and \((D)\) are feasible and \(\zeta \geq ||x^* + s^*||_\infty\) for some pair of optimal solutions \(x^*\) and \((y^*, s^*)\), then \(\overline{\kappa}(\zeta) = 1\).

Evidence was provided by a simple Matlab implementation of our algorithm. As input we used a primal-dual pair of randomly generated feasible problems with known optimal solutions \(x^*\) and \((y^*, s^*)\), and ran the algorithm with \(\zeta = ||x^* + s^*||_\infty\). This was done for various sizes of the problems and for at least \(10^5\) instances. No counterexample for the conjecture was found. Typically the graph of \(\kappa(\zeta, \nu)\), as a function of \(\nu\) is as depicted below.

![Graph of \(\kappa(\zeta, \nu)\)](image)

Typical behavior of \(\kappa(\zeta, \nu)\) as a function of \(\nu\).

The importance of the conjecture is evident. Its trueness would reduce the currently best iteration bound for IIPMs by a factor \(\sqrt{2n}\).
**Related conjecture**

**Conjecture 2** If $(P)$ and $(D)$ are such that $(x^0, y^0, s^0) = \zeta(e, 0, e)$ is a feasible triple, whereas $\zeta \geq \|x^* + s^*\|_\infty$ for some optimal triple $(x^*, y^*, s^*)$, then

$$\frac{\sqrt{\|x(\mu)\|^2 + \|s(\mu)\|^2}}{\zeta \sqrt{2n}} \leq 1, \quad 0 < \mu \leq \zeta^2.$$

In other words,

$$\|x(\mu)\|^2 + \|s(\mu)\|^2 \leq 2n\zeta^2, \quad 0 < \mu \leq \zeta^2.$$
Concluding remarks

• The techniques that have been developed in the field of feasible IPMs, and which have now been known for almost twenty years, are sufficient to get an IIPM whose theoretical performance is as good as the currently best known theoretical performance of IIPMs. Following a well-known metaphor of Isaac Newton\(^a\), it looks like if a “smooth pebble or pretty shell on the sea-shore of IPMs” has been overlooked for a surprisingly long time.

• The full Newton step method presented in this paper is not efficient from a practical point of view. But just as in the case of feasible IPMs one might expect that large-update methods for IIPMs can be designed whose complexity is not worse than \(\sqrt{n}\) times the iteration bound in this paper. Even better results for large-update methods might be obtained by changing the search direction, by using methods that are based on kernel functions. This requires further investigation. Also extensions to second-order cone optimization, semidefinite optimization, linear complementarity problems, etc. seem to be within reach.

\(^a\)“I do not know what I may appear to the world; but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me” (Westfall).
Some references