



A full-Newton step interior-point algorithm based on modified Newton direction

Lipu Zhang^{a,*}, Yinghong Xu^b

^a Department of Mathematics, Zhejiang A&F University, Zhejiang 311300, China

^b Department of Mathematics, Zhejiang Sci-Tech University, Zhejiang 310018, China

ARTICLE INFO

Article history:

Received 23 September 2010

Accepted 7 May 2011

Available online 13 June 2011

Keywords:

Interior-point algorithm

Full-Newton step

Central path

Complexity analysis

ABSTRACT

The central path plays a very important role in interior-point methods. By an equivalent reformulation of the central path, we obtain a new search direction which targets at a small neighborhood of the central path. For a full-Newton step interior-point algorithm based on this search direction, the complexity bound of the algorithm is the best known for linear optimization.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Consider the linear optimization problem (LO)

$$(P) \min\{c^T x : Ax = b, x \geq 0\},$$

where $A \in R^{m \times n}$, $\text{rank}(A) = m$, $b \in R^m$, $c \in R^n$, and its dual problem

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

Without loss of generality, we assume that (P) and (D) satisfy the interior-point condition (IPC) [6], i.e., there exist x^0 , y^0 , and s^0 such that

$$Ax^0 = b, \quad x^0 > 0, \quad A^T y^0 + s^0 = c, \quad s^0 > 0. \quad (1)$$

It is well-known that finding an optimal solution of (P) and (D) is equivalent to solving the following nonlinear system

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= 0, \end{aligned} \quad (2)$$

where xs denotes the coordinatewise product of the vectors x and s . The third equation in system (2) is called the complementarity condition.

The basic idea underlying primal–dual interior-point methods (IPMs) is to replace the complementarity condition by the

nonlinear equation $xs = \mu e$, with parameter $\mu > 0$ and with $e = (1, \dots, 1)^T$. The system (2) now becomes

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= \mu e. \end{aligned} \quad (3)$$

Surprisingly enough, if the IPC is satisfied, then a solution exists, for each $\mu > 0$, and this solution is unique. It is denoted as $(x(\mu), y(\mu), s(\mu))$ and we call $x(\mu)$ the μ -center of (P) and $(y(\mu), s(\mu))$ the μ -center of (D). The set of μ -centers (with μ running through all positive real numbers) gives a homotopy path, which is called the central path of (P) and (D) [3]. If $\mu \rightarrow 0$, then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D).

Many IPMs use the central path. Some algorithms explicitly use the central path as they force the iterates to follow the central path. Even for many algorithms that do not use the central path directly in the algorithm statements, the central path is used for convergence analysis [1,2,4–6].

Through an equivalent reformulation of equation $xs = \mu e$, in Section 2, we derive a new search direction. In Section 3, we present a full-Newton step IPM based on the new search direction. The complexity analysis for the algorithm and the implication for the new search direction are given in Section 4. Finally, we end the paper by Section 5.

2. A modified Newton direction

We derive an equivalent reformulation for the equation $xs = \mu e$.

* Corresponding author. Tel.: +86 571 63732773; fax: +86 571 63732772.
E-mail addresses: zhanglipu@shu.edu.cn (L. Zhang), xyh7913@126.com (Y. Xu).

2.1. Equivalent reformulation

We define

$$v = \sqrt{\frac{xs}{\mu}}. \quad (4)$$

The equation $xs = \mu e$ can be rewritten as $v^2 = e$. Remember that $v \geq 0$, one has

$$v^2 = e \Leftrightarrow v = e \Leftrightarrow v^2 = v.$$

Transforming the left-hand side of the equation $v^2 = v$ to the xs space, we obtain

$$xs = \mu v. \quad (5)$$

Note: The implication for this equivalent reformulation will be given at the end of Section 4.

It will turn out that this equivalence of v can be used to design a new search direction for (P) and (D). The full-Newton step IPM based on this search direction still enjoys the best-known complexity.

2.2. The new search direction

Substituting the third equation in system (3) with (5), we obtain a new system for LO as follows

$$\begin{aligned} Ax &= b, \quad x \geq 0, \\ A^T y + s &= c, \quad s \geq 0, \\ xs &= \mu v. \end{aligned} \quad (6)$$

In feasible IPM, we are given a positive feasible pair (x, s) , and some $\mu > 0$. Our aim is to define search directions $(\Delta x, \Delta s)$ that move in the direction of the small neighborhood of the μ -center $(x(\mu), s(\mu))$. In fact, we want the new iterates $x + \Delta x, s + \Delta s$ to satisfy system (6) and be positive with respect to μ . After substitution this yields the following conditions on $(\Delta x, \Delta s)$

$$\begin{aligned} A(x + \Delta x) &= b, \quad x + \Delta x > 0, \\ A^T(y + \Delta y) + (s + \Delta s) &= c, \quad s + \Delta s > 0, \\ (x + \Delta x)(s + \Delta s) &= \mu v. \end{aligned}$$

If we neglect for the moment the inequality constraints and the quadratic term $\Delta x \Delta s$, then, since $Ax = b$ and $A^T y + s = c$, this system can be rewritten as follows

$$\begin{aligned} A \Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ x \Delta s + s \Delta x &= \mu v - xs. \end{aligned} \quad (7)$$

Since A has full row rank, the above system uniquely defines a search direction $(\Delta x, \Delta y, \Delta s)$ for any $x > 0$ and $s > 0$ [6]. This new Newton direction will be used in our implementations of IPM.

By taking a full-Newton step along the search direction, one constructs a new triple (x^+, y^+, s^+) , with positive x^+ and s^+ , i.e.,

$$x^+ = x + \Delta x, \quad y^+ = y + \Delta y, \quad s^+ = s + \Delta s.$$

For notational convenience, we define

$$d_x := \frac{v \Delta x}{x} \quad \text{and} \quad d_s := \frac{v \Delta s}{s}. \quad (8)$$

Using this notation, the system (7) can be rewritten as follows

$$\begin{aligned} \bar{A} d_x &= 0, \\ \frac{1}{\mu} \bar{A}^T \Delta y + d_s &= 0, \\ d_x + d_s &= e - v. \end{aligned} \quad (9)$$

where $\bar{A} := AV^{-1}X$ and $V := \text{diag}(v)$, $X := \text{diag}(x)$. Once the search directions d_x and d_s are obtained by solving (9), so Δx and Δs can be computed via (8).

It should be noted that the new search direction in system (9) can also be considered as a special case (when σ goes to one) of the following finite-barrier kernel-function

$$\psi(t) = \frac{t^2 - 1}{2} + \frac{1}{\sigma} (e^{\sigma(1-t)} - 1), \quad \sigma \geq 1.$$

One may refer to the ref. [1] for further details.

In what follows, the 2-norm and the infinity-norm are denoted by $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. Note that since d_x belongs to the null space of the matrix $\bar{A}$ and d_s to its row space, it follows that d_x and d_s are orthogonal vectors, i.e.,

$$d_x^T d_s = 0.$$

Using the third equation in (9) we obtain

$$\|d_x\|^2 + \|d_s\|^2 = \|d_x + d_s\|^2 = \|e - v\|^2.$$

Note that $d_x = d_s = 0$ if and only if $v = e$ and hence x and s satisfy $xs = \mu e$, which implies that (x, s) coincides with the μ -center $(x(\mu), s(\mu))$. Thus, we can use $\|e - v\|$ as a quantity to measure closeness to the pair of μ -centers. We therefore define the proximity measure as follows,

$$\sigma(x, s; \mu) = \sigma(v) = \|e - v\|. \quad (10)$$

The generic primal-dual IPM will be described in the following section.

3. Generic feasible primal-dual IPM for LO

It is assumed that we are given a positive primal-dual pair $(x^0, s^0) > 0$ and $\mu^0 > 0$ such that (x^0, s^0) is close to the μ^0 -center in the sense of the proximity measure $\sigma(x^0, s^0; \mu^0)$. In the algorithm Δx and Δs denote the full-Newton step, as defined before.

Generic feasible IPM for LO

Input: Accuracy parameter $\epsilon > 0$;

barrier update parameter $\theta, 0 < \theta < 1$;

threshold parameter $\tau, 0 < \tau < 1$;

feasible pair (x^0, y^0, s^0) with $\mu^0 > 0$ such that $\sigma(x^0, s^0; \mu^0) \leq \tau$.

begin:

$x := x^0; y := y^0; s := s^0; \mu := \mu^0$.

while $x^T s \geq \epsilon$

solve (9) and obtain $(\Delta x, \Delta y, \Delta s)$, let

$x^+ = x + \Delta x$;

$y^+ = y + \Delta y$;

$s^+ = s + \Delta s$;

μ -update: $\mu := (1 - \theta)\mu$;

endwhile

end:

4. Complexity analysis

This section describes the effects of a full-Newton step and of a μ -update, and concludes with a complexity result for our algorithm.

4.1. Some basic results

The next lemma gives some upper bounds for the 2-norm and the infinity norm of the componentwise product of d_x and d_s .

Lemma 4.1 ([6, Lemma C.5]). If u and v are orthogonal, then

$$\|uv\|_\infty \leq \frac{1}{4}\|u+v\|^2, \quad \|uv\| \leq \frac{\sqrt{2}}{4}\|u+v\|^2.$$

Since d_x and d_s are orthogonal vectors, it follows from Lemma 4.1, that one has

$$\|d_x d_s\|_\infty \leq \frac{1}{4}\|d_x + d_s\|^2 = \frac{1}{4}\sigma(v)^2 \quad (11)$$

and

$$\|d_x d_s\| \leq \frac{\sqrt{2}}{4}\|d_x + d_s\|^2 = \frac{\sqrt{2}}{4}\sigma(v)^2. \quad (12)$$

We give some basic properties about the proximity measure $\sigma(v)$.

Lemma 4.2. One has

$$1 - \sigma(v) \leq v_i \leq 1 + \sigma(v), \quad 1 \leq i \leq n.$$

Proof. Since

$$|1 - v_i| \leq \|e - v\| = \sigma(v),$$

the result easily follows. $\square$

4.2. Properties of the full-Newton step

By (8) and the third equation of system (9), one has

$$\begin{aligned} x^+ s^+ &= (x + \Delta x)(s + \Delta s) \\ &= xs + (s\Delta x + x\Delta s) + \Delta x \Delta s \\ &= xs + (\mu v - xs) + \Delta x \Delta s \\ &= \mu(v + d_x d_s). \end{aligned} \quad (13)$$

We want the new iterates be strictly positive, so we only have to concentrate on the sign of the coordinates of the vectors x^+ and s^+ . We call the Newton step strictly feasible if x^+ and s^+ are positive. The main aim of this subsection is to find conditions for strict feasibility of the full-Newton step.

Lemma 4.3. The Newton step is strictly feasible if and only if $v + d_x d_s > 0$.

Proof. The “only if” part of both statements in the lemma follows immediately from (13). For the proof of the converse implication we introduce a step length $\alpha \in [0, 1]$, and define

$$x^\alpha = x + \alpha \Delta x \quad \text{and} \quad s^\alpha = s + \alpha \Delta s.$$

We then have $x^0 = x, x^1 = x^+$ and similarly $s^0 = s, s^1 = s^+$. Hence, we have $x^0 s^0 = xs > 0$. The proof uses a continuity argument, namely that x^1 and s^1 are nonnegative if $x^\alpha s^\alpha$ is positive for all α in the open interval $(0, 1)$. We write

$$x^\alpha s^\alpha = (x + \alpha \Delta x)(s + \alpha \Delta s) = xs + \alpha(x\Delta s + s\Delta x) + \alpha^2 \Delta x \Delta s.$$

Using the third equation of system (9), we obtain

$$\begin{aligned} x^\alpha s^\alpha &= xs + \alpha(\mu v - xs) + \alpha^2 \Delta x \Delta s \\ &= \mu[(1 - \alpha)v^2 + \alpha v + \alpha^2 d_x d_s]. \end{aligned}$$

Suppose $v + d_x d_s > 0$, i.e., $d_x d_s > -v$. Substitution gives

$$\begin{aligned} x^\alpha s^\alpha &> \mu[(1 - \alpha)v^2 + \alpha v - \alpha^2 v] \\ &= \mu(1 - \alpha)(v^2 + \alpha v). \end{aligned}$$

Since v^2 and v are positive and $\alpha \in (0, 1)$, it follows that $x^\alpha s^\alpha > 0$ for $0 \leq \alpha < 1$. Hence, none of the entries of x^α and s^α vanish for $0 \leq \alpha < 1$. Since x^0 and s^0 are positive, and x^α and s^α depend linearly on α , this implies that $x^\alpha > 0$ and $s^\alpha > 0$ for $0 \leq \alpha < 1$. Hence, x^1 and s^1 must be positive, proving the if part of the statement in the lemma. $\square$

Corollary 4.4. The new iterates (x^+, s^+) are certainly strictly feasible if

$$\|d_x d_s\|_\infty < \min(v).$$

Proof. By Lemma 4.3, x^+ and s^+ are strictly feasible if and only if $v + d_x d_s > 0$. Since the inequality holds if $\|d_x d_s\|_\infty < \min(v)$, the corollary follows. $\square$

Lemma 4.5. Let $\sigma(v)$ defined as (10), and (x, s) be any positive pair and suppose $\mu > 0$. If $\sigma(v) < 2\sqrt{2} - 2$, then the full-Newton step for LO is strictly feasible, i.e., x^+ and s^+ are positive.

Proof. It follows from (11) and Lemma 4.2, that one has

$$\|d_x d_s\|_\infty \leq \frac{1}{4}\sigma(v)^2 \quad \text{and} \quad 1 - \sigma(v) \leq \min(v).$$

It is easily verified that

$$\|d_x d_s\|_\infty < \min(v)$$

certainly holds for

$$\frac{1}{4}\sigma(v)^2 < 1 - \sigma(v),$$

which is equivalent to $\sigma(v) < 2\sqrt{2} - 2$. By Corollary 4.4, the new iterates after a full-Newton step are certainly strictly feasible, which completes the proof. $\square$

The next lemma gives the effect of full-Newton step on duality gap.

Lemma 4.6. If $\sigma(v) < 2\sqrt{2} - 2$, then

$$(x^+)^T s^+ < (2\sqrt{2} - 1)n\mu.$$

Proof. Using (13), Lemma 4.2 and remembering that the vectors d_x and d_s are orthogonal, one has

$$\begin{aligned} (x^+)^T s^+ &= e^T (x^+ s^+) = \mu e^T (v + d_x d_s) \\ &= \mu(e^T v + (d_x)^T d_s) \leq n\mu \max(v) \\ &\leq n\mu(1 + \sigma(v)). \end{aligned}$$

Since $\sigma(v) < 2\sqrt{2} - 2$, one has

$$(x^+)^T s^+ < (2\sqrt{2} - 1)n\mu,$$

which completes the proof. $\square$

Denote $v^+ = \sqrt{\frac{x^+ s^+}{\mu}}$. It follows from (13) that

$$(v^+)^2 = v + d_x d_s. \quad (14)$$

By (11) and Lemma 4.2, one has

$$\min((v^+)^2) \geq \min(v) - \|d_x d_s\|_\infty \geq 1 - \sigma(v) - \frac{\sigma(v)^2}{4}.$$

Assuming that the iterates x^+ and s^+ are strictly feasible, one can obtain a lower bound for the vector v^+ .

$$\min(v^+) \geq \sqrt{1 - \sigma(v) - \frac{\sigma(v)^2}{4}}. \quad (15)$$

The following theorem gives the effect of the proximity measure after the full-Newton step and a μ -update.

Theorem 4.7. Let (x, s) be a positive pair and $\mu > 0$. Moreover, let $\sigma(v) < 2\sqrt{2} - 2$ and $\mu^+ = (1 - \theta)\mu$. Then

$$\sigma(x^+, s^+; \mu^+) \leq \frac{\sigma(v) + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\sigma(v)^2}{1 - \theta + \sqrt{1 - \theta}\sqrt{1 - \sigma(v) - \frac{\sigma(v)^2}{4}}}.$$

Proof. Since

$$\begin{aligned} \sigma(x^+, s^+; \mu^+) &= \left\| e - \sqrt{\frac{x^+ s^+}{\mu^+}} \right\| \\ &= \frac{1}{\sqrt{1 - \theta}} \left\| \sqrt{1 - \theta} e - v^+ \right\|, \end{aligned} \quad (16)$$

and

$$\begin{aligned} \left\| \sqrt{1 - \theta} e - v^+ \right\|^2 &= \sum_{i=1}^n \left(\sqrt{1 - \theta} - v_i^+ \right)^2 \\ &= \sum_{i=1}^n \frac{(1 - \theta - (v_i^+)^2)^2}{(\sqrt{1 - \theta} + v_i^+)^2} \\ &\leq \frac{\sum_{i=1}^n (1 - \theta - (v_i^+)^2)^2}{(\sqrt{1 - \theta} + \min(v^+))^2} \\ &= \frac{\|e - \theta e - v - d_x d_s\|^2}{(\sqrt{1 - \theta} + \min(v^+))^2} \\ &\leq \frac{(\sigma(v) + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\sigma(v)^2)^2}{(\sqrt{1 - \theta} + \min(v^+))^2} \\ &\leq \frac{(\sigma(v) + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\sigma(v)^2)^2}{\left(\sqrt{1 - \theta} + \sqrt{1 - \sigma(v) - \frac{\sigma(v)^2}{4}} \right)^2}, \end{aligned}$$

where the third equation, the second inequality and the last inequality follow from (14), the triangle inequality and (15), respectively. Taking square roots at both sides of the above inequality, and substituting into (16) the result easily follows. $\square$

4.3. Fixing the parameter

We want to find a threshold τ and an update parameter θ , which at the start of the iterate satisfies $\sigma(x, s; \mu) \leq \tau$. After the full-Newton step and a μ -update, the property $\sigma(x^+, s^+; \mu^+) \leq \tau$ should be maintained. In this case, by Theorem 4.7, it suffices if

$$\frac{\sigma(v) + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\sigma(v)^2}{1 - \theta + \sqrt{1 - \theta}\sqrt{1 - \sigma(v) - \frac{\sigma(v)^2}{4}}} \leq \tau.$$

The left-hand side of the above inequality is monotonically increasing with respect to $\sigma(v)$, which implies that

$$\begin{aligned} &\frac{\sigma(v) + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\sigma(v)^2}{1 - \theta + \sqrt{1 - \theta}\sqrt{1 - \sigma(v) - \frac{\sigma(v)^2}{4}}} \\ &\leq \frac{\tau + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\tau^2}{1 - \theta + \sqrt{1 - \theta}\sqrt{1 - \tau - \frac{\tau^2}{4}}}. \end{aligned}$$

Thus the $\sigma(x^+, s^+; \mu^+) \leq \tau$ suffices if

$$\frac{\tau + \theta\sqrt{n} + \frac{\sqrt{2}}{4}\tau^2}{1 - \theta + \sqrt{1 - \theta}\sqrt{1 - \tau - \frac{\tau^2}{4}}} \leq \tau. \quad (17)$$

At this stage, after some simple calculation, if we set $\tau = \frac{1}{2}$ and $\theta = \frac{1}{7\sqrt{n}}$, the inequality (17) certainly holds. Which means that $(x, s) > 0$ and $\sigma(x, s; \mu) \leq \frac{1}{2}$ are maintained during the algorithm. Thus the algorithm is well-defined.

It is easily verified that, if $\theta = \frac{1}{7\sqrt{n}}$ and $\tau = \frac{1}{2}$, after one full-Newton step and a μ -update, the property $\sigma(x, s; \mu) \leq \frac{1}{2}$ is maintained. By Lemma 4.2, one concludes that the scaled central path $xs = \mu v$ satisfies

$$\mu(1 - \sigma(v))e \leq xs \leq \mu(1 + \sigma(v))e,$$

which gives

$$\frac{1}{2}\mu e \leq xs \leq \frac{3}{2}\mu e.$$

This means that our new search keeps the iterates in a small neighborhood of the central path.

4.4. Complexity bound

Lemma 4.8. If the barrier parameter μ has the initial value μ^0 and is repeatedly multiplied by $1 - \theta$, with $0 < \theta < 1$, then after at most

$$\left\lceil \frac{1}{\theta} \log \frac{(2\sqrt{2} - 1)n\mu^0}{\varepsilon} \right\rceil$$

iterations we have $x^T s \leq \varepsilon$.

Proof. At the initial point, one has $(x^0)^T s^0 = n\mu^0$. After one iteration, by Lemma 4.6, the duality-gap equals

$$(x^1)^T s^1 \leq (2\sqrt{2} - 1)(1 - \theta)n\mu^0,$$

thus, after k iterates, the duality-gap satisfies

$$(x^k)^T s^k \leq (2\sqrt{2} - 1)(1 - \theta)^k n\mu^0.$$

So, it suffices if

$$(2\sqrt{2} - 1)(1 - \theta)^k n\mu^0 \leq \varepsilon,$$

which, by taking the logarithm gives

$$k \log(1 - \theta) + \log((2\sqrt{2} - 1)n\mu^0) \leq \log \varepsilon. \quad (18)$$

Since

$$\log(1 - \theta) \leq -\theta,$$

it certainly suffices if

$$-k\theta + \log((2\sqrt{2} - 1)n\mu^0) \leq \log \varepsilon,$$

this gives

$$k \geq \frac{1}{\theta} \log \frac{(2\sqrt{2} - 1)n\mu^0}{\varepsilon}$$

which completes the proof. $\square$

For $\theta = \frac{1}{7\sqrt{n}}$, the following theorem holds trivially.

Theorem 4.9. Setting $\tau = 1/2$ and $\theta = 1/7\sqrt{n}$, the initial duality-gap is $(x^0)^T s^0 = n\mu^0$, the full-Newton step primal-dual IPMs for LO

has the complexity bound

$$\mathcal{O} \left(7\sqrt{n} \log \frac{(2\sqrt{2} - 1)n\mu^0}{\varepsilon} \right).$$

5. Conclusions

We have presented a full-Newton step IPM based on the modified Newton direction and obtained the best-known complexity bound for LO. Our further research may focus on designing the infeasible IPM based on the new search direction and doing numerical tests of the algorithm.

Acknowledgments

The authors kindly acknowledge the help of the associate editor and anonymous referees in improving the readability of the

paper. The relation of the new search direction with the kernel-function was remembered by one of the referees, and is especially acknowledged. The research is supported by National Natural Science Foundation of China, No. 11071221 and Foundation of Zhejiang Educational Committee, No. Y200908058.

References

- [1] Y.Q. Bai, M. El Ghami, C. Roos, A new efficient large-update primal–dual interior-point method based on a finite barrier, *SIAM J. Optim.* 13 (3) (2002) 766–782.
- [2] Y.Q. Bai, M. El Ghami, C. Roos, A comparative study of kernel functions for primal–dual interior-point algorithms in linear optimization, *SIAM J. Optim.* 15 (1) (2004) 101–128.
- [3] N. Megiddo, Pathways to the optimal set in linear programming, in: N. Megiddo (Ed.), *Progress in Mathematical Programming: Interior Point and Related Methods*, Springer Verlag, New York, 1989, pp. 131–158.
- [4] J. Peng, C. Roos, T. Terlaky, New complexity analysis of the primal–dual Newton method for linear optimization, *Ann. Oper. Res.* 99 (1) (2000) 23–39.
- [5] J. Peng, C. Roos, T. Terlaky, *Self-Regularity: A New Paradigm for Primal–Dual Interior-Point Algorithms*, Princeton Univ. Pr., 2002.
- [6] C. Roos, T. Terlaky, J.P. Vial, *Theory and Algorithms for Linear Optimization: An Interior Point Approach*, John Wiley & Son Ltd., 1997.