An Interior-point Method with Polynomial Complexity and Superlinear Convergence for Linear Complementarity Problems

Jun Ji
F. Potra
R.A Tapia
Y. Zhang

CRPC-TR91170
July, 1991

An Interior-Point Method with Polynomial Complexity and Superlinear Convergence for Linear Complementarity Problems

Jun Ji [*], Florian Potra [†], Richard Tapia [‡] and Yin Zhang [§]

July, 1991

Key Words: Linear programming, quadratic programming, linear complementarity problems, primal-dual interior-point algorithms, polynomiality, superlinear convergence.

Abbreviated Title: Interior-Point Methods for Linear Complementarity Problems

[*]Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242

[†]Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242 and visiting member of the Center for Research on Parallel Computation, Rice University, Houston, Texas, 77251-1892. Research supported in part by NSF Coop. Agr. No. CCR-8809615.

[‡]Department of Mathematical Sciences and the Center for Research on Parallel Computation, Rice University, Houston, Texas, 77251-1892. Research supported in part by NSF Coop. Agr. No. CCR-8809615, AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.

[§]Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21228 and visiting member of the Center for Research on Parallel Computation, Rice University, Houston, Texas, 77251-1892. Research supported in part by NSF Grant DMS-9102761 and NSF Coop. Agr. No. CCR-8809615.

1

Abstract

For linear programming, a primal-dual interior-point algorithm was recently constructed by Zhang and Tapia that achieves both polynomial complexity and Q-superlinear convergence (Q-quadratic in the nondegenerate case). In this paper, we extend their results to quadratic programming and linear complementarity problems.

1 Introduction

Primal-dual interior-point algorithms, originally introduced by Megiddo [10] in 1986 for linear programming, have a close connection to Newton's method. Exploiting this connection, Zhang, Tapia and Dennis [16] established sufficient conditions for primal-dual algorithms to achieve superlinear convergence for linear programming (LP). Zhang, Tapia and Potra [17] extended these results to quadratic programming (QP) and linear complementarity problems (LCP). Recently, Zhang and Tapia [18] constructed a primal-dual algorithm that exhibits both polynomial complexity and Q-superlinear convergence (Q-quadratic in the nondegenerate case). More recently, Kojima, Kurita and Mizuno [2] proposed rules for choosing parameters in primal-dual algorithms, and showed that for LCP one set of rules leads to both global convergence (but not necessarily polynomial) and superlinear convergence.

In this paper, we extend the Zhang-Tapia algorithm [18] from LP to QP and LCP, and prove the same theoretical properties, i.e., polynomial complexity and superlinear convergence, for the extended algorithm.

We take a unified approach towards linear programming quadratic programming and linear complementarity problems. The basis of this unified approach is the following nonlinear system with nonnegativity restrictions on the variables

$$F(x,y) = \begin{pmatrix} Mx + Ny - h \\ XYe \end{pmatrix} = 0, \quad (x,y) \geq 0, \tag{1.1}$$

2

where $x, y, h, e \in \mathbf{R}^n$, $M, N \in \mathbf{R}^{n \times n}$, $X = diag(x)$, $Y = diag(y)$ and e has all components equal to one.

Problem (1.1) is sufficiently general to include linear complementarity problems, quadratic programming problems and linear programming problems. To begin with, observe that if $N = -I$, then this problem is the standard linear complementarity problem. Consider the following equality constrained quadratic programming problem

$$
\begin{aligned}
\text{minimize} \quad & c^T x + \tfrac{1}{2} x^T Q x \\
\text{subject to} \quad & Ax = b, \\
& x \geq 0,
\end{aligned}
\tag{1.2}
$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n} (m < n)$ and has full row rank, and $Q \in \mathbf{R}^{n \times n}$. If Q is symmetric and positive semi-definite on the null space of A; then QP (1.2) is a convex program and the first order optimality conditions are both necessary and sufficient for optimality. The first-order conditions for (1.2) can be transformed into the form of (1.1), as was shown in [17], with

$$
M = \begin{bmatrix} A \\ -BQ \end{bmatrix}, \quad N = \begin{bmatrix} 0 \\ B \end{bmatrix} \quad \text{and} \quad h = \begin{pmatrix} b \\ Bc \end{pmatrix},
\tag{1.3}
$$

where $B \in \mathbf{R}^{(n-m) \times n}$ is any matrix such that the columns of B^T form a basis for the null space of A. When $Q = 0$, the quadratic programming problem (1.2) reduces to the standard form linear programming problem.

A direct calculation shows that

$$
F'(x, y) = \begin{bmatrix} M & N \\ Y & X \end{bmatrix}.
\tag{1.4}
$$

The feasibility set of problem (1.1) is:

$$
\mathcal{F} = \{(x, y) : x, y \in \mathbf{R}^n, Mx + Ny = h, (x, y) \geq 0\}.
$$

Moreover, a feasible pair $(x, y) \in \mathcal{F}$ is said to be strictly feasible if it is positive.

Throughout this paper, whenever we consider LCP, we assume that the matrix M is positive semi-definite; and whenever we consider QP, we assume that the matrix Q is symmetric and positive semi-definite on the null space of A. Also, we will assume that a strictly feasible point exists.

Subscripts will be used to distinguish values of quantities at a particular iteration and superscripts will indicate components of vectors. We also use the notation:

$$\min(v) \equiv \min(v^i) \equiv \min_{1 \leq i \leq n} v^i \quad \text{and} \quad \max(v) \equiv \max(v^i) \equiv \max_{1 \leq i \leq n} v^i$$

for a vector $v \in \mathbf{R}^n$. The symbol $\| \cdot \|$ denotes the ℓ_2 norm unless otherwise stated.

The paper is organized as follows. In Section 2, we describe the basic interior-point algorithm for our general model (1.1). In Sections 3, we specify our choice of the step-length and the centering parameter. Polynomiality and superlinear convergence are established in Sections 4 and 5, respectively. Concluding remarks are given in Section 6.

2 Algorithm

The basic algorithm we will study is the following.

Algorithm 1

Given a strictly feasible pair (x_0, y_0). For $k = 0, 1, 2, \ldots,$ do

Step 1 *Compute the Newton step*

$$\begin{pmatrix} \Delta x_k^N \\ \Delta y_k^N \end{pmatrix} = -[F'(x_k, y_k)]^{-1} F(x_k, y_k)$$

and the centering step

$$\begin{pmatrix} \Delta x_k^C \\ \Delta y_k^C \end{pmatrix} = \frac{1}{n} x_k^T y_k [F'(x_k, y_k)]^{-1} \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

4

Step 2 *Choose* $\sigma_k \in (0,1)$ *and form the combined step*

$$\begin{pmatrix} \Delta x_k \\ \Delta y_k \end{pmatrix} = \begin{pmatrix} \Delta x_k^N \\ \Delta y_k^N \end{pmatrix} + \sigma_k \begin{pmatrix} \Delta x_k^C \\ \Delta y_k^C \end{pmatrix}.$$

Step 3 *Choose* $\alpha_k = \tau_k \hat{\alpha}_k$, *where* $\tau_k \in (0,1)$ *and*

$$\hat{\alpha}_k = \frac{-1}{\min(X_k^{-1}\Delta x_k, Y_k^{-1}\Delta y_k)}.$$

Step 4 *Compute the new iterate*

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} + \alpha_k \begin{pmatrix} \Delta x_k \\ \Delta y_k \end{pmatrix}.$$

Notice that since (x_0, y_0) is strictly feasible, all the subsequent iterates $\{(x_k, y_k)\}$ will also be strictly feasible. Also for QP and LCP, the matrix $F'(x, y)$ is nonsingular for all $(x, y) > 0$; hence Algorithm 1 is well-defined.

For $(x, y) \in \mathcal{F}$, we use the ℓ_1-norm of the residual of (1.1), i.e.,

$$\|F(x, y)\|_1 = x^T y,$$

as our measure of convergence. This measure can be shown to be the duality gap in linear and quadratic programming.

Observe that all three pairs $(\Delta x_k, \Delta y_k)$, $(\Delta x_k^N, \Delta y_k^N)$ and $(\Delta x_k^C, \Delta y_k^C)$ satisfy the equation $Mu + Nv = 0$. We also have the following useful relationships:

$$Y_k \Delta x_k + X_k \Delta y_k = -X_k Y_k e + \sigma_k \frac{1}{n} x_k^T y_k e \tag{2.1}$$

and

$$x_{k+1}^T y_{k+1} = x_k^T y_k (1 - (1 - \sigma_k)\alpha_k) + \Delta x_k^T \Delta y_k \alpha_k^2. \tag{2.2}$$

Algorithm 1 covers, or is closely related to, a wide range of existing interior-point algorithms for linear programming, quadratic programming and linear complementarity

problems. In particular, it covers most of the existing primal-dual interior-point algorithms for linear programming as well as quadratic programming, including Kojima, Mizuno and Yoshise [8], Todd and Ye [13], Monteiro and Adler [11, 12], Lustig [9], to name a few. Algorithms for linear complementarity problems that are covered by Algorithm 1 include Kojima, Mizuno and Yoshise [6, 7], Kojima, Megiddo and Noma [3], Kojima, Megiddo and Ye [4], Kojima, Mizuno and Noma [5]. For more references, see two recent review papers by Ye [14, 15]. Although these algorithms have been motivated and presented in various ways including path-following (homotopy or continuation), potential reduction or affine scaling algorithms, most of them fit into the framework of Algorithm 1.

3   Choices of Parameters

There are two control parameters in Algorithm 1. One is the so-called centering parameter σ_k and the other is the step-length parameter α_k. In [17], Zhang, Tapia and Potra established sufficient conditions for choosing the two parameters for Algorithm 1 to achieve superlinear convergence. Our concern in this paper is both global and local behavior of the algorithm appled to QP and LCP. For LP, this issue has been studied by Zhang and Tapia [18]. They gave rules for choosing the centering and the step-length parameters so that both polynomial complexity and superlinear convergence are obtained. In this section, we will modify their rules for choosing the step-length parameter for LP. The modified rules will fit not only LP but also QP and LCP.

We adopt the following notation:

$$
\begin{aligned}
x_k(\alpha) &= x_k + \alpha\Delta x_k, \qquad y_k(\alpha) = y_k + \alpha\Delta y_k, \\
f_k(\alpha) &= X_k(\alpha)Y_k(\alpha)e, \quad f_k^{\mathrm{ave}}(\alpha) = \tfrac{1}{n}x_k(\alpha)^T y_k(\alpha), \\
f_k^{\mathrm{min}}(\alpha) &= \min(f_k(\alpha)), \quad f_k^{\mathrm{max}}(\alpha) = \max(f_k(\alpha)).
\end{aligned} \tag{3.1}
$$

Whenever $\alpha = 0$, we will drop the argument from the above functions. For example, $x_k \equiv$

$x_k(0)$, $f_k^{\text{ave}} \equiv f_k^{\text{ave}}(0)$ and so on. Clearly, we also have $x_{k+1} = x_k(\alpha_k)$, $f_{k+1}^{\text{ave}} = f_k^{\text{ave}}(\alpha_k)$ and so on.

It is important not to allow any $x_k^i y_k^i$ to become too small compared with the average value $\frac{1}{n} x_k^T y_k$ (see [17], for example). Hence we require that $\alpha = \alpha_k$ satisfies

$$\frac{f_k^{\min}(\alpha)}{f_k^{\text{ave}}(\alpha)} \geq \gamma_k, \quad \alpha > 0, \tag{3.2}$$

where

$$\gamma_k \in [\gamma, f_k^{\min}/f_k^{\text{ave}}] \quad \text{and} \quad 0 < \gamma \leq f_0^{\min}/f_0^{\text{ave}} \leq 1. \tag{3.3}$$

Note that in the case $f_0^{\min}/f_0^{\text{ave}} > \gamma$, we allow γ_k to decrease monotonically as long as $\gamma_k > \gamma$.

It is easy to verify that

$$f_k^i(\alpha) = f_k^i - (f_k^i - \sigma_k f_k^{\text{ave}})\alpha + \Delta x_k^i \Delta y_k^i \alpha^2. \tag{3.4}$$

Hence, $f_k^i(\alpha)$ is a quadratic for every i. As a result, $f_k^{\min}(\alpha)$ and $f_k^{\max}(\alpha)$ are piecewise quadratics. Moreover, $f_k^{\text{ave}}(\alpha)$ is also a quadratic and

$$f_k^{\text{ave}}(\alpha) = f_k^{\text{ave}}[1 - (1 - \sigma_k)\alpha] + \frac{\Delta x_k^T \Delta y_k}{n} \alpha^2. \tag{3.5}$$

For notational convenience, let us introduce the piecewise quadratic function

$$h_k(\alpha) = f_k^{\min}(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha). \tag{3.6}$$

It follows that condition (3.2) is equivalent to

$$h_k(\alpha) \geq 0, \quad \alpha > 0. \tag{3.7}$$

In determining α_k we will use the following quantity:

$$\alpha_k^\gamma = \min\{\alpha > 0 : h_k(\alpha) = 0\}. \tag{3.8}$$

The following lemma is given in Zhang and Tapia [18].

Lemma 3.1 *The quantity α_k^γ is well defined and $\alpha_k^\gamma \in (0, \hat{\alpha}_k)$. Moreover, condition (3.2) is satisfied for all $\alpha \in (0, \alpha_k^\gamma]$.*

An equivalent expression for α_k^γ is

$$\alpha_k^\gamma = \min\{\alpha > 0 : f_k^i(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha) = 0, \ i = 1, 2, \ldots, n\}.$$

The computation of α_k^γ involves calculating the roots of at most n quadratics and therefore requires $O(n)$ operations.

In addition to a lower bound for $\{f_k^i(\alpha_k)/f_k^{\text{ave}}(\alpha_k)\}$ (i.e., condition (3.2)), we also impose an upper bound on these quantities; namely, we require α_k to satisfy

$$\frac{f_k^{\text{max}}(\alpha)}{f_k^{\text{ave}}(\alpha)} \leq \Gamma_k, \quad \alpha > 0, \tag{3.9}$$

where

$$\Gamma_k \in [f_k^{\text{max}}/f_k^{\text{ave}}, \Gamma] \quad \text{and} \quad 1 \leq f_0^{\text{max}}/f_0^{\text{ave}} \leq \Gamma \leq n. \tag{3.10}$$

Since $f_k^i(\alpha)/f_k^{\text{ave}}(\alpha) < n$ for all i, condition (3.9) will be redundant if $\Gamma_k = n$. Our reason for introducing condition (3.9) is to improve our complexity bound.

Similarly, we introduce the piecewise quadratic function

$$H_k(\alpha) = f_k^{\text{max}}(\alpha) - \Gamma_k f_k^{\text{ave}}(\alpha). \tag{3.11}$$

It is easy to see that condition (3.9) is equivalent to

$$H_k(\alpha) \leq 0, \quad \alpha > 0. \tag{3.12}$$

We define

$$\alpha_k^\Gamma = \begin{cases} +\infty, & \text{if } \{\alpha > 0 : H_k(\alpha) = 0\} = \emptyset, \\ \min\{\alpha > 0 : H_k(\alpha) = 0\}, & \text{otherwise.} \end{cases} \tag{3.13}$$

Analogous to Lemma 3.1 for condition (3.2), we have the following lemma for condition (3.9).

Lemma 3.2 *The quantity α_k^Γ is well-defined and $\alpha_k^\Gamma \in (0, \hat{\alpha}_k)$. Moreover, condition (3.9) is satisfied by all $\alpha \in (0, \alpha_k^\Gamma)$ and also by $\alpha = \alpha_k^\Gamma$ when the latter is finite.*

8

Proof: The proof is similar to that for Lemma 3.1, so we omit it. □

Similarly, α_k^Γ can be found by calculating the roots of at most n quadratics and requires $O(n)$ operations.

The duality gap

$$x_k(\alpha)^T y_k(\alpha) = x_k^T y_k - x_k^T y_k (1 - \sigma_k)\alpha + \Delta x_k^T \Delta y_k \alpha^2 \qquad (3.14)$$

is a quadratic function of α with the second derivative $\Delta x_k^T \Delta y_k$. If $\Delta x_k^T \Delta y_k > 0$, it will reach its minimum at

$$\alpha_k^\nu = \frac{(1 - \sigma_k)x_k^T y_k}{2\Delta x_k^T \Delta y_k}. \qquad (3.15)$$

If $\Delta x_k^T \Delta y_k = 0$, the duality gap is a decreasing function of α; so it is always desirable to take the largest step-length possible. For LCP and QP, one always has $\Delta x_k^T \Delta y_k \geq 0$ (see Lemma 3.3 below). Therefore our choice of the step-length α_k is

$$\alpha_k = \begin{cases} \min(1, \alpha_k^\gamma, \alpha_k^\Gamma, \alpha_k^\nu), & \text{if } \Delta x_k^T \Delta y_k > 0, \\ \min(1, \alpha_k^\gamma, \alpha_k^\Gamma), & \text{otherwise,} \end{cases} \qquad (3.16)$$

where α_k^γ is defined by (3.8), α_k^Γ by (3.13) and α_k^ν by (3.15). This choice of step-length will guarantee (3.2), (3.9) and $\alpha_k \in (0, \hat{\alpha}_k)$.

For the sake of simplicity, we will enforce the conditions

$$\gamma_k \leq 1/2 \quad \text{and} \quad \Gamma_k \geq 2. \qquad (3.17)$$

The specific values in (3.17) do not constitute a loss of generality because they will only affect expressions for some constants in our analysis. These values of γ_k and Γ_k will lead to much simplified expressions for those constants.

Procedure 1

Given

$$0 < \gamma \leq \min(1/2, f_0^{\min}/f_0^{\text{ave}}), \quad \max(2, f_0^{\max}/f_0^{\text{ave}}) \leq \Gamma < n. \qquad (3.18)$$

Step 1 *Choose $\gamma_k \in [\gamma, \min(1/2, f_k^{\min}/f_k^{\text{ave}})]$ and $\Gamma_k \in [\max(2, f_k^{\max}/f_k^{\text{ave}}), \Gamma]$.*

9

Step 2 *Compute* α_k^γ *by (3.8),* α_k^Γ *by (3.13), and if necessary,* α_k^ν *by (3.15).*

Step 3 *Compute* α_k *by (3.16).*

When $\Delta x_k^T \Delta y_k = 0$, the above procedure basically reduces to Procedure 1 in [18], which was designed for LP where $\Delta x_k^T \Delta y_k \equiv 0$.

Now we shift our attention to the choice of the centering parameter. We will use exactly the same strategy to choose σ_k as that used in [18]. However, for the sake of completeness we still describe the choice in detail. Let

$$
\begin{aligned}
p_k &= X_k^{-1} \Delta x_k , & q_k &= Y_k^{-1} \Delta y_k , \\
p_k^N &= X_k^{-1} \Delta x_k^N, & q_k^N &= Y_k^{-1} \Delta y_k^N, \\
p_k^C &= X_k^{-1} \Delta x_k^C, & q_k^C &= Y_k^{-1} \Delta y_k^C,
\end{aligned}
\tag{3.19}
$$

and

$$
\omega_k = \max_{1 \le i \le n} (|(p_k^N)^i (q_k^N)^i|, |(p_k^N)^i (q_k^C)^i|, |(p_k^C)^i (q_k^N)^i|, |(p_k^C)^i (q_k^C)^i|).
\tag{3.20}
$$

Lemma 3.3 *For QP and LCP,*

(i) $\Delta x_k^T \Delta y_k \ge 0$, $(\Delta x_k^N)^T (\Delta y_k^N) \ge 0$ *and* $(\Delta x_k^N)^T (\Delta y_k^N) \ge 0$;

(ii) *if* $f_k^{\min} / f_k^{\text{ave}} \ge \gamma$, *then* $\omega_k \le n/\gamma^2$.

Proof: (i) Since all three pairs $(\Delta x_k, \Delta y_k)$, $(\Delta x_k^N, \Delta y_k^N)$ and $(\Delta x_k^C, \Delta y_k^C)$ satisfy the equation $Mu + Nv = 0$. It suffices to prove that $u^T v \ge 0$ for all $u, v \in \mathbf{R}^n$ satisfying $Mu + Nv = 0$.

In the case of LCP $(N = -I)$, $Mu + Nv = 0$ is equivalent to $v = Mu$. Hence $u^T v = u^T Mu \ge 0$ because M is positive semi-definite.

In the case of QP, $Mu + Nv = 0$ is equivalent to $Au = 0$ and $BQu = Bv$ (see (1.3)). Using the representations $u = B^T u_2$ and $v = A^T v_1 + B^T v_2$, where $v_1 \in \mathbf{R}^m$ and $u_2, v_2 \in \mathbf{R}^{n-m}$, and noticing that $A^T \perp B^T$, we have $u^T v = u_2^T BB^T v_2$. Moreover,

10

$BQu = Bv$ is equivalent to $BQB^T u_2 = BB^T v_2$. Hence, if Q is positive semi-definite on the null space of A, then

$$u^T v = u_2^T BB^T v_2 = u_2^T (BQB^T) u_2 \geq 0.$$

(ii) Multiply both sides of $Y_k \Delta x_k^N + X_k \Delta y_k^N = -X_k Y_k e$ by $(X_k Y_k)^{-\frac{1}{2}}$ and consider the square of the ℓ_2-norm of both sides. Using (i) of Lemma 3.3, we have

$$\|(X_k Y_k)^{\frac{1}{2}} p_k^N\|^2 + \|(X_k Y_k)^{\frac{1}{2}} q_k^N\|^2 \leq x_k^T y_k;$$

or equivalently after dividing both sides by $\frac{1}{n} x_k^T y_k$,

$$\|T_k^{-\frac{1}{2}} p_k^N\|^2 + \|T_k^{-\frac{1}{2}} q_k^N\|^2 \leq n, \tag{3.21}$$

where $T_k = \frac{1}{n} x_k^T y_k (X_k Y_k)^{-1}$ is a diagonal matrix. Since $f_k^{\min}/f_k^{\mathrm{ave}} \geq \gamma$, it is easy to see that

$$|(p_k^N)^i| \leq \sqrt{n/\gamma} \leq \sqrt{n}/\gamma \quad \text{and} \quad |(q_k^N)^i| \leq \sqrt{n/\gamma} \leq \sqrt{n}/\gamma.$$

Similarly, we can prove

$$|(p_k^C)^i| \leq \sqrt{n}/\gamma \quad \text{and} \quad |(q_k^C)^i| \leq \sqrt{n}/\gamma.$$

The result in (ii) of Lemma 3.3 follows directly from the definition of ω_k and the above estimates. □

We now state our procedure for choosing the centering parameter σ_k.

Procedure 2

Given

$$\sigma \in (0,1), \quad \rho^l = \frac{\gamma^2 \sigma}{2n}, \quad \rho^u \geq \frac{\gamma^2 \sigma}{n}. \tag{3.22}$$

Step 1 *Compute ω_k from (3.20).*

Step 2 *Compute $\rho_k^u = \min(\rho^u, \sigma/\omega_k)$.*

Step 3 *Choose $\rho_k \in [(\rho^l + \rho_k^u)/2, \rho_k^u].$*

11

Step 4 *Let $\sigma_k = \rho_k \omega_k$.*

It is easy to verify, using (ii) of Lemma 3.3, that $\rho_k^u > \rho^l$; thus σ_k is well-defined. Moreover, we have $\sigma_k \leq \rho_k^u \omega_k \leq \sigma < 1$.

Overall, we have the following algorithm.

Algorithm 2 *Given a strictly feasible pair (x_0, y_0). Choose (recall (3.18))*

$$0 < \gamma \leq \min(1/2, f_0^{\min}/f_0^{\text{ave}}), \quad \max(2, f_0^{\max}/f_0^{\text{ave}}) \leq \Gamma < n,$$

and $\sigma \in (0, 1)$. Set $\rho^l = \gamma^2 \sigma / 2n$ and $\rho^u \geq \gamma^2 \sigma / n$. For $k = 0, 1, 2, \ldots$, do

Step 1 *Compute $(\Delta x_k^N, \Delta y_k^N)$ and $(\Delta x_k^C, \Delta y_k^C)$ from Algorithm 1.*

Step 2 *Choose σ_k by Procedure 2 and form $(\Delta x_k, \Delta y_k)$ from Algorithm 1.*

Step 3 *Choose α_k by Procedure 1.*

Step 4 *Form (x_{k+1}, y_{k+1}) from Algorithm 1.*

4 Polynomiality

In this section, we will establish the global behavior of Algorithm 2. We first define

$$\begin{aligned}
\xi_k &= \tfrac{\Gamma_k}{n} \Delta x_k^T \Delta y_k - \max(\Delta x_k^i \Delta y_k^i) \\
\eta_k &= \tfrac{\gamma_k}{n} \Delta x_k^T \Delta y_k - \min(\Delta x_k^i \Delta y_k^i).
\end{aligned} \tag{4.1}$$

Lemma 4.1 *Let α_k^γ be given by (3.8). If $\eta_k \leq 0$, then $\alpha_k^\gamma \geq 1$; otherwise,*

$$\alpha_k^\gamma \geq \frac{(1 - \gamma_k)\sigma_k x_k^T y_k}{n \eta_k} = \frac{(1 - \gamma_k)\sigma_k f_k^{\text{ave}}}{\eta_k}. \tag{4.2}$$

Proof: First we notice that for $\alpha \in [0, 1]$,

$$f_k^i(\alpha) = f_k^i - (f_k^i - \sigma_k f_k^{\text{ave}})\alpha + \Delta x_k^i \Delta y_k^i \alpha^2 \geq f_k^{\min} - (f_k^{\min} - \sigma_k f_k^{\text{ave}})\alpha + \min(\Delta x_k^i \Delta y_k^i)\alpha^2.$$

12

Therefore,

$$
\begin{aligned}
f_k^i(\alpha) - \gamma_k f_k^{\text{ave}}(\alpha) &\geq f_k^{\min} - (f_k^{\min} - \sigma_k f_k^{\text{ave}})\alpha + \min(\Delta x_k^i \Delta y_k^i)\alpha^2 - \gamma_k f_k^{\text{ave}}(\alpha) \\
&= (f_k^{\min} - \gamma_k f_k^{\text{ave}})(1 - \alpha) + (1 - \gamma_k)\sigma_k f_k^{\text{ave}}\alpha - \eta_k \alpha^2 \qquad (4.3) \\
&\geq (1 - \gamma_k)\sigma_k f_k^{\text{ave}}\alpha - \eta_k \alpha^2.
\end{aligned}
$$

If $\eta_k \leq 0$, then it follows that $h_k(\alpha)$ has no root in $(0,1)$; hence $\alpha_k^\gamma \geq 1$. Otherwise, the quadratic in the right-hand side of the last inequality has a unique positive root

$$
\alpha = \frac{(1 - \gamma_k)\sigma_k x_k^T y_k}{n\eta_k} = \frac{(1 - \gamma_k)\sigma_k f_k^{\text{ave}}}{\eta_k}.
$$

Therefore, if $\alpha_k^\gamma < 1$, from (4.3) α_k^γ must be greater than or equal to the above positive number. This completes the proof. $\qquad \square$

Lemma 4.2 *Let α_k^Γ be given by (3.13). If $\xi_k \geq 0$, then $\alpha_k^\Gamma \geq 1$; otherwise,*

$$
\alpha_k^\Gamma \geq \frac{(1 - \Gamma_k)\sigma_k x_k^T y_k}{n\xi_k} = \frac{(1 - \Gamma_k)\sigma_k f_k^{\text{ave}}}{\xi_k}. \qquad (4.4)
$$

Proof: The proof is similar to that of Lemma 4.1. So we omit it. $\qquad \square$

Lemma 4.3 *Let ξ_k and η_k be given by (4.1).*

(i) $\xi_k \geq -4\omega_k f_k^{\max}$,

(ii) $\eta_k \leq 6\omega_k f_k^{\max}$.

Proof: (i) Note that

$$
\begin{aligned}
|\Delta x_k^i \Delta y_k^i| &= |(x_k^i y_k^i)(p_k^i q_k^i)| \\
&\leq \max(X_k Y_k e)\|diag(p_k)q_k\|_\infty \\
&= f_k^{\max}\|diag(p_k^N + \sigma_k p_k^C)(q_k^N + \sigma_k p_k^C)\|_\infty \\
&\leq f_k^{\max}\omega_k(1 + \sigma_k)^2 \\
&\leq 4f_k^{\max}\omega_k.
\end{aligned}
$$

Since $\Delta x_k^T \Delta y_k \geq 0$, we have (see (4.1))

$$\xi_k \geq -\max(\Delta x_k^i \Delta y_k^i) \geq -4 f_k^{\max} \omega_k.$$

(ii) Since $|\Delta x_k^i \Delta y_k^i| \leq 4 f_k^{\max} \omega_k$, $\Delta x_k^T \Delta y_k \leq 4n f_k^{\max} \omega_k$. Therefore,

$$
\begin{aligned}
\eta_k &= \frac{\gamma_k}{n} \Delta x_k^T \Delta y_k - \min(\Delta x_k^i \Delta y_k^i) \\
&\leq 4\gamma_k f_k^{\max} \omega_k + 4 f_k^{\max} \omega_k \\
&\leq 6 f_k^{\max} \omega_k,
\end{aligned}
$$

because $\gamma_k \leq 1/2$. This completes the proof. $\qquad\square$

Lemma 4.4 *Let (x_k, y_k), $(\Delta x_k, \Delta y_k)$ and $(\alpha_k^\gamma, \alpha_k^\Gamma)$ be generated by Algorithm 2. Then there exists a constant $\beta > 0$, such that*

$$\min(1, \alpha_k^\gamma, \alpha_k^\Gamma) \geq \frac{\beta}{n}. \tag{4.5}$$

Moreover,

$$\frac{\Delta x_k^T \Delta y_k}{x_k^T y_k} \leq \frac{1}{4\gamma}. \tag{4.6}$$

Proof: From Lemmas 4.1 and 4.2, it suffices to consider the situation where $\xi_k < 0$ and $\eta_k > 0$. In this case,

$$
\begin{aligned}
\min(1, \alpha_k^\gamma, \alpha_k^\Gamma) &\geq \min\left(1, \frac{(1-\gamma_k)\sigma_k f_k^{\text{ave}}}{\eta_k}, \frac{(1-\Gamma_k)\sigma_k f_k^{\text{ave}}}{\xi_k}\right) & \text{[Lemmas 4.1-4.2]} \\
&\geq \min\left(1, \frac{0.5\sigma_k f_k^{\text{ave}}}{6\omega_k f_k^{\max}}, \frac{\sigma_k f_k^{\text{ave}}}{4\omega_k f_k^{\max}}\right) & \text{[(3.17) and Lemma 4.3]} \\
&\geq \min\left(1, \frac{\rho_k}{12\Gamma}\right) & \text{[}f_k^{\max}/f_k^{\text{ave}} \leq \Gamma \text{ and } \sigma_k = \rho_k \omega_k\text{]} \\
&\geq \min\left(1, \frac{\gamma^2 \sigma}{24\Gamma n}\right) & \text{[}\rho_k \geq \rho^l = \gamma^2\sigma/2n\text{]} \\
&\geq \left(\frac{\gamma^2 \sigma}{24\Gamma}\right)\frac{1}{n} & \text{[}\gamma < 1, \Gamma > 1 \text{ and } \sigma < 1\text{].}
\end{aligned}
$$

Hence,

$$\beta = \frac{\gamma^2 \sigma}{24\Gamma} \tag{4.7}$$

will satisfies (4.5).

Now define $D_k = X_k Y_k^{-1}$ and let

$$r_k = [\sigma_k \tfrac{1}{n} x_k^T y_k (X_k Y_k)^{-1/2} - (X_k Y_k)^{1/2}]e.$$

It follows that

$$r_k^T r_k = x_k^T y_k \left(1 - 2\sigma_k + \sigma_k^2 \tfrac{1}{n} x_k^T y_k \frac{e^T (X_k Y_k)^{-1} e}{n}\right) \le x_k^T y_k (1 - 2\sigma_k + \sigma_k^2/\gamma), \qquad (4.8)$$

since $f_k^{\text{ave}}/f_k^i \le 1/\gamma$.

Multiplying both sides of (2.1) by $(X_k Y_k)^{-1/2}$ gives

$$r_k = D_k^{-1/2} \Delta x_k + D_k^{1/2} \Delta y_k.$$

Therefore,

$$
\begin{aligned}
r_k^T r_k &= \Delta x_k^T D_k^{-1} \Delta x_k + \Delta y_k^T D_k \Delta y_k + 2\Delta x_k^T \Delta y_k \\
&= \|D_k^{-1/2} \Delta x_k - D_k^{1/2} \Delta y_k\|^2 + 4\Delta x_k^T \Delta y_k \qquad (4.9) \\
&\ge 4\Delta x_k^T \Delta y_k.
\end{aligned}
$$

Combining (4.8) and (4.9) and noting $\sigma_k < 1$ leads to

$$
\begin{aligned}
\frac{\Delta x_k^T \Delta y_k}{x_k^T y_k} &\le \tfrac{1}{4}(1 - 2\sigma_k + \sigma_k^2/\gamma) \\
&= \tfrac{1}{4}((1 - \sigma_k)^2 + \sigma_k^2(\tfrac{1}{\gamma} - 1)) \\
&\le \tfrac{1}{4\gamma}.
\end{aligned}
$$

This completes the proof. $\qquad\qquad\square$

Theorem 4.1

Let $\{(x_k, y_k)\}$ be generated by Algorithm 2. Then

$$\frac{x_{k+1}^T y_{k+1}}{x_k^T y_k} \le (1 - \delta_k) < 1,$$

where $\delta_k = \Omega(\tfrac{1}{n})$.

Proof: From (2.2),

$$x_{k+1}^T y_{k+1} = x_k^T y_k \left(1 - (1 - \sigma_k)\alpha_k + \frac{\Delta x_k^T \Delta y_k}{x_k^T y_k}\alpha_k^2\right).$$

15

We must estimate the quantity

$$\delta_k = \left(1 - \sigma_k - \frac{\Delta x_k^T \Delta y_k}{x_k^T y_k} \alpha_k \right) \alpha_k. \tag{4.10}$$

Since $\delta_k < 1$ is guaranteed in Algorithm 2 by the choice of α_k, it is sufficient to verify that $\delta_k = \Omega(\frac{1}{n})$.

Consider the following two functions of α

$$\delta_k(\alpha) = \left(1 - \sigma_k - \frac{\Delta x_k^T \Delta y_k}{x_k^T y_k} \alpha \right) \alpha \tag{4.11}$$

and

$$\bar{\delta}(\alpha) = \left(1 - \sigma - \frac{1}{4\gamma} \alpha \right) \alpha. \tag{4.12}$$

From (4.6) and $\sigma_k \leq \sigma$, for $\alpha \geq 0$

$$\delta_k(\alpha) \geq \bar{\delta}(\alpha).$$

From the choice of α_k in Procedure 1, it can be verified that

$$\alpha_k = \arg\max\{\delta_k(\alpha) : \alpha \in [0, \min(1, \alpha_k^\gamma, \alpha_k^\Gamma)]\}.$$

Therefore, for all $\alpha \in [0, \min(1, \alpha_k^\gamma, \alpha_k^\Gamma)]$,

$$\delta_k = \delta_k(\alpha_k) \geq \bar{\delta}(\alpha). \tag{4.13}$$

Let β be defined by (4.7). It follows from (4.5) that (4.13) is satisfied by $\alpha = \beta/n$; i.e.,

$$\delta_k \geq \bar{\delta}(\beta/n) = \left(1 - \sigma - \frac{\beta}{4\gamma n}\right) \frac{\beta}{n} = \Omega(\tfrac{1}{n}).$$

This completes the proof. □

The following corollary follows immediately from Theorem 4.1. Its proof is straightforward (see [18]), for example). By a standard argument, it leads to polynomiality assuming integral data.

Corollary 4.1 *Assume* $x_0^T y_0 \leq 2^{\nu L}$, *where* $L > 0$ *and* $\nu > 0$. *Then in at most* $O(nL)$ *iterations, Algorithm 2 will produce* (x_k, y_k) *satisfying* $x_k^T y_k \leq 2^{-L}$.

16

5 Superlinear Convergence

In this section, we will apply the following superlinear convergence result of Zhang, Tapia and Potra [17] to establish Q-superlinear convergence of Algorithm 2 for QP and LCP.

Theorem 5.1 (Zhang, Tapia and Potra)

Let $\{(x_k, y_k)\}$ be generated by Algorithm 1 and (x_*, y_*) be a solution of problem (1.1). Assume

(i) strict complementarity holds at (x_*, y_*),

(ii) $\{f_k^{\text{ave}} / f_k^{\text{min}}\}$ is bounded above,

(iii) $\sigma_k \to 0$ and $\alpha_k \to 1$.

Then if $\{(x_k, y_k)\}$ converges to (x_*, y_*), the sequence $\{X_k Y_k e\}$, component-wise, converges to zero Q-superlinearly. Moreover, under the above conditions, $\alpha_k \to 1$ if and only if $\tau_k \to 1$.

Since τ_k is directly under our control and α_k is not, the more meaningful result is stated in terms of τ_k. However, Lemma 3.1 of Zhang, Tapia and Potra [17] demonstrates that under the conditions of the theorem, $\hat{\alpha}_k$ as defined in Step 3 of Algorithm 1 converges to 1. Hence, $\alpha_k \to 1$ if and only if $\tau_k \to 1$ and we may consider either condition. In the sequel, we choose to verify $\alpha_k \to 1$.

We note that unlike in LP, a strictly complementary solution may not exist in QP or LCP. So the strict complementarity assumption for QP and LCP is a somewhat strong, but often made, assumption.

For the same technical reasons as encountered in Zhang and Tapia [18], we must restrict further the choice of ρ_k. Denote the length of the interval $[\rho^l, \rho_k^u]$ by π_k. It follows from (3.22), Step 2 of Procedure 2, and Lemma 3.3 that

$$\rho_k^u \geq \frac{\gamma^2 \sigma}{n}. \tag{5.1}$$

17

Thus,

$$\pi_k = \rho_k^u - \rho^l \geq \frac{\gamma^2 \sigma}{2n} > 0. \tag{5.2}$$

Let Σ_k be the following set of $2n$ points

$$\Sigma_k = \{-(p_k^N)^i/(p_k^C)^i, \ -(q_k^N)^i/(q_k^C)^i, \ i = 1, 2, \ldots, n\}$$

and define the distance from σ to the set Σ_k as

$$\text{dist}(\sigma, \Sigma_k) = \min\{|\sigma - \varsigma| : \varsigma \in \Sigma_k\}.$$

We choose σ_k according to Procedure 2 with the additional restriction that

$$\text{dist}(\sigma_k, \Sigma_k) \geq \pi_k \omega_k/(8n + 4). \tag{5.3}$$

In other words, we require not only

$$\sigma_k \in [0.5(\rho^l + \rho_k^u)\omega_k, \rho_k^u \omega_k] \tag{5.4}$$

but also that σ_k be bounded away from the set Σ_k by at least the amount $\pi_k \omega_k/(8n+4)$. Since $\{\pi_k\}$ is bounded away from zero, we see from (5.3) that $\{\text{dist}(\sigma_k, \Sigma_k)\}$ is bounded away from zero when $\{\omega_k\}$ is bounded away from zero. It has been shown in Lemma 8.1 of Zhang and Tapia [18] that the set of σ values satisfying (5.3) and (5.4) is nonempty.

Our superlinear convergence theory for Algorithm 2, applied to not only LP, but also to QP and LCP, is the following.

Theorem 5.2

Let (x_*, y_*) be a solution of problem (1.1) and $\{(x_k, y_k)\}$ be generated by Algorithm 2 with the restriction (5.3) on the centering parameter σ_k. Assume

(i) *strict complementarity holds at (x_*, y_*),*

(ii) *ρ^u is sufficiently large, e.g., $\rho^u \geq 24\Gamma$.*

18

For LP, QP and LCP, if $\{(x_k, y_k)\}$ converges to (x_*, y_*), then the sequence $\{X_k Y_k e\}$, component-wise, converges to zero Q-superlinearly.

Proof: Theorem 5.2 is a generalization of Theorem 8.1 in [18], i.e., the superlinear convergence of Algorithm 2 applied to LP. As for the proof of Theorem 8.1 in [18], we need to show $\sigma_k \to 0$ and $\alpha_k \to 1$.

The proof for $\sigma_k \to 0$ is identical to that in [18], because it only requires that $\{\alpha_k\}$ be bounded below and that $\{p_k^N\}$, $\{q_k^N\}$, $\{p_k^C\}$ and $\{q_k^N\}$ be bounded above, which are also true for QP and LCP (see Lemmas 3.3 and 4.4).

The proof for $\alpha_k \to 1$ is similar to that in [18]. First, the same argument used in [18] gives

$$\alpha_k \leq \hat{\alpha}_k \to 1. \tag{5.5}$$

When $\Delta x_k^T \Delta y_k = 0$, it follows from (3.16) and the proof of Lemma 4.4 that

$$\alpha_k = \min(1, \alpha_k^\gamma, \alpha_k^\Gamma) \geq \min\left(1, \frac{\rho_k}{12\Gamma}\right). \tag{5.6}$$

From the definition of ρ_k and the fact $\omega_k \to 0$, for k sufficiently large

$$\rho_k = \frac{1}{2}(\rho^l + \rho_k^u) \geq \frac{1}{2}\rho_k^u = \frac{1}{2}\rho^u.$$

Therefore, if $\rho^u \geq 24\Gamma$, then from (5.6)

$$\alpha_k \geq 1. \tag{5.7}$$

Now inequalities (5.5) and (5.7) together imply $\alpha_k \to 1$.

When $\Delta x_k^T \Delta y_k > 0$, in order to prove (5.7) we also need to show that $\alpha_k^\nu \geq 1$ for k sufficiently large (see (3.16)). It follows from $\Delta x_k^T \Delta y_k \leq 4n\omega_k f_k^{\max}$ (see the proof of Lemma 4.3 (ii)), $f_k^{\max}/f_k^{\text{ave}} \leq \Gamma$, $\sigma_k \leq \sigma$ and $\omega_k \to 0$ that as $k \to \infty$

$$\alpha_k^\nu = \frac{(1 - \sigma_k)x_k^T y_k}{2\Delta x_k^T \Delta y_k} \geq \frac{(1 - \sigma_k)f_k^{\text{ave}}}{8\omega_k f_k^{\max}} \geq \frac{1 - \sigma}{8\Gamma\omega_k} \to \infty.$$

Hence, $\alpha_k = \min(1, \alpha_k^\gamma, \alpha_k^\Gamma, \alpha_k^\nu) \geq 1$ for k sufficiently large. This completes the proof. $\square$

6 Concluding Remarks

In this paper, we have constructed a primal-dual interior-point method for a general problem which includes linear programming, quadratic programming and linear complementarity problems, and have demonstrated its $O(nL)$-iteration polynomial complexity and its Q-superlinear convergence. We accomplished the above objectives by adopting the rules for choosing step-length in the Zhang-Tapia polynomial and superlinearly convergent algorithm for LP [18] to our general problem, and then extending their theory to our problem.

Acknowledgment

References

[1] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, NJ, 1963.

[2] M. Kojima, Y. Kurita and S. Mizuno. Large-step interior point algorithms for linear complementarity problems. Research Report on Information Sciences B-243, Dept. of Information Sciences, Tokyo Institute of Technology, 1991.

[3] M. Kojima, N. Megiddo, and T. Noma. Homotopy continuation methods for complementarity problems, 1988. manuscript, IBM Almaden Research Center, San Jose, California.

[4] M. Kojima and Y. Ye and N. Megiddo. An interior point potential reduction algorithm for the linear complementarity problems. Technical report, IBM Almaden Research Center, San Jose, California, 1988.

[5] M. Kojima, S. Mizuno, and T. Noma. A new continuation method for complementarity problems with uniform p-functions. *Mathematical Programming*, 43:107–113, 1989.

[6] M. Kojima, S. Mizuno, and A. Yoshise. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Math. Prog.*, 50:331–342, 1991.

[7] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual algorithm for a class of linear complementarity problems. *Mathematical Programming*, 44:1–26, 1989.

[8] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point method for linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 29–47. Springer-Verlag, New York, 1989.

[9] I.J. Lustig, R.E. Marsten, and D.F. Shanno. Computational experience with a primal-dual interior point method for linear programming. Technical Report SOR 89-17, Dept. Civil Eng. and O.R., Princeton University, 1989.

[10] N. Megiddo. Pathways to the optimal set in linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 131–158. Springer-Verlag, New York, 1989.

[11] R.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part I: linear programming. *Math. Prog.*, 44:27–41, 1989.

[12] R.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part II: convex quadratic programming. *Math. Prog.*, 44:43–66, 1989.

[13] M.J. Todd and Yinyu Ye. A centered projective algorithm for linear programming. Technical Report No. 763, School of Operations Research and Industrial Engineering, Cornell University, 1987 (revised 1989). To appear in *Math. Prog.*

[14] Yinyu Ye. Interior point algorithms for quadratic programming. Technical report, Dept. of Management Sciences, The University of Iowa, 1989. Working Paper Series No. 89-29.

[15] Yinyu Ye. Interior point algorithms for global optimization. Technical report, Dept. of Management Sciences, The University of Iowa, 1990. Working Paper Series No. 90-2.

[16] Yin Zhang, R. A. Tapia, and J. E. Dennis. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. Technical Report TR90-6, Dept. Mathematical Sciences, Rice University, 1990. To appear in *SIAM J. on Optimization.*

[17] Y. Zhang, R. A. Tapia and F. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. Technical Report TR90-9, Dept. Mathematical Sciences, Rice University, 1990.

[18] Y. Zhang and R. A. Tapia. A quadratically convergent polynomial primal-dual interior point algorithm for linear programming. Technical Report TR90-40, Dept. Mathematical Sciences, Rice University, Houston, Texas 77251-1892, 1991.