A Primal-Infeasible Interior Point Algorithm For Linearly Constrained Convex Programming *

Yan-jin Wang †, Pu-sheng Fei ‡

Abstract

In the paper a primal-infeasible interior point algorithm is proposed for linearly constrained convex programming. The starting point is any positive primal-infeasible dual-feasible point in a large region. The method maintains positivity of the iterates which point satisfies primal-infeasible dual-feasible point. At each iterates it requires to solve approximately a nonlinear system. It is shown that, after polynomial iterations a sufficiently good approximation to the optimal point is found, or there is no optimal point in a large nonnegative region.

Keywords: Linearly constrained convex programming; Primal-infeasible interior point algorithm; Polynomial complexity

1. Introduction

In this paper we consider the linearly constrained convex programming problem:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b,\ x \geq 0 \end{array} \tag{1}$$

where $x \in R^n$, $A \in R^{m\times n}$, $b \in R^m$, $m \leq n$ and $f : R^n \to R$ is a sufficiently smooth convex function.

The dual problem for (1) can be put in the form

$$\begin{array}{ll} \text{maximize} & b^T y - (x^T \nabla f(x) - f(x)) \\ \text{subject to} & \nabla f(x) - (A^T y + s) = 0,\ s \geq 0 \end{array} \tag{2}$$

We let Ω denote the feasible point set of dual-primal problem (1) and (2).

$$\Omega = \{(x, y, s) : \ Ax = b, (x, s) \geq 0, \nabla f(x) - A^T y - s = 0\}$$

It is well-known that $(x, y, s) \in R^n \times R^m \times R^n$ is an primal-dual optimal solution of (1) and (2) if and only if the point (x, y, s) satisfies the following first-order optimality conditions (called KKT-Condition) for (1) and (2).

$$\begin{aligned} A^T y + s &= \nabla f(x) \\ Ax &= b \\ (x, s) &\geq 0 \\ x^T s &= 0 \end{aligned}$$

*This work was supported by Foundation for doctors(20020486035),China.

†Corresponding author. PhD candidate, School of Math. and Statistics, Wuhan University, Wuhan, Hubei, 430072, China **E-mail:** wang_jasonyj2002@yahoo.com

‡Professor, School of Math. and Statistics, Wuhan University, Wuhan, Hubei, 430072, China

If we relax the fourth condition of KKT-condition as follows

$$Xs - \beta_1 \mu e = 0$$

where $\mu \geq 0$ and $\beta_1 \in [0, 1]$ is a constant, then KKT-condition is called the perturbed. Clearly, when $\mu \to 0$, then the point (x, y, s), which satisfies the perturbed KKT-condition, will converge to an optimal point of problem (1) and (2).

In recent years various feasible interior point methods have been developed for solving convex programming and nonlinear complementarity problems based on the idea of either reducing the primal-dual complementarity gap $x^T s$ or reducing the value of some primal-dual potential functions, to readers we here refer books Refs. 1 - 4, papers Refs. 5 - 8 and references of these books for convex programming , Refs. 9 - 11 and references of these books for nonlinear complementarity problems. Most of these methods have achieved globally linear convergence with polynomial complexity. In the framework of infeasible-interior-point algorithms (cf. Refs. 12 - 14), it is also considered to solve convex programming Refs. 15 and 16, nonlinear complementarity problems Refs. 17 and 18, these algorithms also have properties of global convergence. But as far as the authors know, there is no result of polynomial complexity. In this paper we are interested in a polynomial infeasible interior point method for convex programming. The infeasible interior point method for problem (1) is more intricate compared to the analysis of the infeasible interior point algorithm for linear programming because of the nonlinear term of the convex object function for problem (1). Thus, based on the interior point algorithms Ref. 5, to avoid the nonlinear term, here we present a primal-infeasible dual-feasible interior point algorithm. The starting point requires any positive primal-infeasible dual-feasible point at each iterate. It asks to find an approximate solution of a nonlinear system. And under some conditions the paper analyzes the complexity, we prove that after polynomial iterations we get an approximate optimal point, or show that there is no optimal point in a large given region.

Throughout the paper, the following notations are used. All vectors are column vectors. We frequently use (x, y) as shorthand for the vector $(x^T, y^T)^T$. R^m denotes the m-dimentional Euclidean space. The set of all $m \times n$ matrices with real entries is denoted by $R^{m \times n}$. The diagonal matrix corresponding to a vector x is denoted by X, i.e., $X = \text{diag(x)}$, and $e = (1, 1, \cdots, 1)^T \in R^n$ We also denote by $\|x\|_1$, $\|x\|_2$ and $\|x\|_\infty$ the 1-, 2- and ∞-norm of x, that is to say, $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, $\|x\|_2 = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$ and $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$. The superscript ^{T} denotes transpose.

The rest of the paper is organized as follows. In Section 2, we describe the primal-infeasible dual-feasible interior point algorithm for convex programming; Section 3 analyze the polynomial convergence of our algorithm; At last section, we make some concluding remarks.

2. Algorithm

In this section, before we state our algorithm, several assumptions are introduced, and we also give out a definition of a neighborhood set of the central path for the algorithm.

Assumption 1 Without loss of the generalization we let $\text{Rank}(A) = m$;
This assumption is quite standard for convex programming.

Assumption 2 $f(x)$ is continuously differentiable and convex. $\nabla f(x)$ satisfies the Lipschitz condition with the Lipschitz index L, i.e.,

$$\|\nabla f(x') - \nabla f(x'')\| \leq L\|x' - x''\|$$

where $L > 0$.

Under the Assumption 3, at Assumption 2 the Lipschitz condition is weaker than the twice continuous differentiablity.

Assumption 3 We want to find an optimal point of convex programming, if it exists, in the following region.

$$\Phi = \{(x, y, s) \in R^n \times R^m \times R^n | (x, s) \geq 0, \|(x, s)\|_\infty \leq \rho\}$$

Note that Assumption 3 is a frequently used assumption for infeasible interior point methods to obtain polynomial complexity bounds.

Assumption 4 Let $\Omega_1 = \{(x, y, s) : (x, s) > 0 \text{ and } s = \nabla f(x) - A^T y\}$, we here suppose that $\Omega_1 \bigcap \Phi \neq \emptyset$.

Obviously this assumption is also weaker than one which the feasible interior point set of problem (1) and (2) is not empty.

The central path of problems (1) and (2) is defined as follows

$$S = \left\{ (x, y, s) : x > 0, s > 0, Ax - b = 0, A^T y + s - \nabla f(x) = 0, Xs = \frac{x^T s}{n} e \right\}$$

in primal-dual form.

Without considering the feasibility, we let a neighborhood set of the central path as follows.

$$\mathcal{N} = \left\{ (x, y, s) > 0 : \ (x, s) > 0, \|Xs - \frac{x^T s}{n} e\| \leq \sigma \frac{x^T s}{n}, \sigma \in (0, 1) \right\} \tag{3}$$

We note that for an infeasible interior point algorithm it always defines a neighborhood set of the central path just like the set above.

From the aboving definition, we can easily get the following lemma.

Lemma 2.1 For $(x, y, s) \in \mathcal{N}$, then

$$\begin{aligned} \max\{x_i s_i\} &\leq (1 + \sigma) \frac{x^T s}{n} \\ \min\{x_i s_i\} &\geq (1 - \sigma) \frac{x^T s}{n}. \end{aligned}$$

The primal-dual affine scaling search direction $(\Delta x, \Delta y, \Delta s)$ at a given infeasible interior point $(x^k, y^k, s^k) \in R^n \times R^m \times R^n$ is computed by applying one step of Newton's method to the perturbed KKT-condition. Hence,

$$\left\{ \begin{array}{rcl} A\Delta x &=& -(Ax^k - b) \\ \nabla^2 f(x^k)\Delta x - A^T \Delta y - \Delta s &=& -(\nabla f(x^k) - A^T y^k - s^k) \\ X^k \Delta s + S^k \Delta x &=& -(X^k s^k - \beta_1 \frac{(x^k)^T s^k}{n}) \end{array} \right. \tag{4}$$

According to the idea of the infeasible interior point algorithm Refs. 12 - 14, the search direction of an infeasible interior point algorithm is the solution of the system above. But for convex problem (1) and (2) because of the nonlinearity of $f(x)$, the second equation of the system (4) will reduce to difficulty when one analyzes the polynomial complexity of the primal-dual infeasible interior point algorithm. Therefore, to avoid the difficulty, here we will consider the dual-feasible point, thus we can cancel to solve the second equation of the system (4). That is, we consider take the solution of the following system as the iterative direction at a point $(x^k, y^k, s^k) \in R^n \times R^m \times R^n$

$$\left\{ \begin{array}{rcl} X^k[\nabla f(x^k + \alpha^k \Delta x) - \nabla f(x^k) - \alpha^k A^T \Delta y] + \alpha^k S^k \Delta x &=& -\alpha^k (X^k s^k - \beta_1 \mu e) \\ A\Delta x &=& -(Ax^k - b) \end{array} \right. \tag{5}$$

where α^k and μ is some positive real numbers.

As is known, it is difficulty to find an exact solution of the nonlinear system above. Thus we induce an inexact technic to the system. To remain the property

$$A(x^k + \alpha^k \Delta x) - b = (1 - \alpha^k)(Ax^k - b) \tag{6}$$

it requires to solve the linear equation exactly, then the system can be rewritten in the form

$$\begin{cases} X^k[\nabla f(x^k + \alpha^k \Delta x) - \nabla f(x^k) - \alpha^k A^T \Delta y] + \alpha^k S^k \Delta x &= -\alpha^k (X^k s^k - \beta_1 \mu e) + \alpha^k r^k \\ A\Delta x &= -(Ax^k - b) \end{cases} \tag{7}$$

where r^k satisfies that $\|r^k\|_1 \leq \nu\mu$ and ν is some small positive constant. Obviously, it holds that $\|r^k\| \leq \nu\mu$.

Now we state the primal-infeasible dual-feasible interior point algorithm as follows.

Algorithm 2.1

Step 0. Find an initial starting point (x^0, y^0, s^0) in $\Omega_1 \bigcap \Phi$. We here put

$$\rho_0 = \min\{x_1^0, x_2^0, \cdots, x_n^0, s_1^0, s_2^0, \cdots, s_n^0\}.$$

Let $0 < \nu \leq 1/2 \leq \beta_1 < \beta_2 \leq 1$, $\theta^0 = 1$ satisfying $\sigma\beta_1 > 2\nu$ and $\beta_2 > \beta_1 + \nu$. Set $k = 0$.

Step 1. If the point (x^k, y^k, s^k) satisfies

$$\|Ax^k - b\| \leq \epsilon_p \text{ and } (x^k)^T s^k \leq \epsilon \tag{8}$$

then the algorithm is terminated.

Step 2. Compute α^k as follows.

$$\alpha^k = \left(\frac{\rho_0}{\rho}\right)^2 \frac{\min\{\sigma\beta_1 - 2\nu, \beta_2 - \beta_1 - \nu/n\} \min_i (x_i^k s_i^k)}{2n(2+\tau)^2 \left(1 + (1+L)n^{\frac{1}{2}}\right)^2 (x^k)^T s^k} \tag{9}$$

where

$$\tau = \frac{\rho\|(s^0, x^0)\|_1}{(x^0)^T s^0}$$

Step 3. Let $\mu = \dfrac{(x^k)^T s^k}{n}$, then find $(\Delta x, \Delta y) \in R^n \times R^m$ satisfying th equations (7).

Step 4. Set $x^{k+1} := x^k + \alpha^k \Delta x, y^{k+1} := y^k + \alpha^k \Delta y, s^{k+1} := \nabla f(x^k + \alpha^k \Delta x) - A^T(y^k + \alpha^k \Delta y)$, $\theta^{k+1} := \theta^k (1 - \alpha^k)$ and $k := k + 1$, then go to step 1.

3. Convergence

In this section we will analyze the polynomial convergence of Algorithm 2.1. Firstly, before we state the complexity result of algorithm 2.1, we will describe several lemmas. And throughout this section we will denote $D = (X^k)^{\frac{1}{2}}(S^k)^{-\frac{1}{2}}$ for a point $(x^k, s^k) > 0$

Lemma 3.1 Assume the problems (1) and (2) have an optimal point$(x^*, y^*, s^*) \in \Phi$, and for $k \geq 1$ the point (x^j, y^j, s^j) is generated by Algorithm 2.1 with $j = 1, 2, \cdots, k$ and satisfies

$$(x^j, s^j) > 0; \tag{10}$$

$$(x^j)^T s^j \geq (1 - \alpha^{(j-1)})(x^{j-1})^T s^{j-1}. \tag{11}$$

Then for $k \geq 1$

$$\|(x^k, s^k)\|_1 \leq \frac{2+\tau}{\theta^k \rho_0}(x^k)^T s^k \tag{12}$$

where τ is defined in Algorithm 2.1 depending on the starting point and the large region.

Proof. From Algorithm 2.1, we have

$$A(x^k - \theta^k x^0 - (1-\theta^k)x^*) = 0$$

And

$$\begin{aligned}(s^k - \theta^k s^0 - (1-\theta^k)s^*) \quad &= \quad (\nabla f(x^k) - A^T y^k) - \theta^k(\nabla f(x^0) - A^T y^0) \\ &\quad\; -(1-\theta^k)(\nabla f(x^*) - A^T y^*)\end{aligned}$$

Noting $0 < \theta^k < 1$, from the above equations and the fact that $(\dot{x} - \ddot{x})^T(\nabla f(\dot{x}) - \nabla f(\ddot{x})) \geq 0$ we get

$$\begin{aligned}
&(x^k - \theta^k x^0 - (1-\theta^k)x^*)^T(s^k - \theta^k s^0 - (1-\theta^k)s^*) \\
&= (x^k - \theta^k x^0 - (1-\theta^k)x^*)^T(\nabla f(x^k) - \theta^k \nabla f(x^0) - (1-\theta^k)\nabla f(x^*))^T \\
&= [\theta^k(x^k - x^0) + (1-\theta^k)(x^k - x^*)]^T[\theta^k(\nabla f(x^k) - \nabla f(x^0)) + (1-\theta^k)(\nabla f(x^k) - \nabla f(x^*))] \\
&\geq \theta^k(1-\theta^k)(x^k - x^0)^T(\nabla f(x^k) - \nabla f(x^*)) + \theta^k(1-\theta^k)(x^k - x^*)]^T(\nabla f(x^k) - \nabla f(x^0)) \\
&= \theta^k(1-\theta^k)(x^k - x^* + x^* - x^0)^T(\nabla f(x^k) - \nabla f(x^*)) \\
&\quad + \theta^k(1-\theta^k)(x^k - x^0 + x^0 - x^*)]^T(\nabla f(x^k) - \nabla f(x^0)) \\
&\geq \theta^k(1-\theta^k)(x^* - x^0)^T(\nabla f(x^*) - \nabla f(x^0)) \\
&\geq 0
\end{aligned}$$

Therefore, $(\theta^k x^0 + (1-\theta^k)x^* - x^k)^T(\theta^k s^0 + (1-\theta^k)s^* - s^k) \geq 0$ which imples that $(\theta^k x^0 + (1-\theta^k)x^*)^T s^k + \left(x^k\right)^T(\theta^k s^0 + (1-\theta^k)s^*) \leq (\theta^k x^0 + (1-\theta^k)x^*)^T(\theta^k s^0 + (1-\theta^k)s^*) + \left(x^k\right)^T s^k$.

Thus, according to the definition of the starting point $()x^0, y^0, s^0$ of our algorithm, Assumption 3 and (10) we know

$$\begin{aligned}
\theta^k \rho_0 \|(x^k, s^k)\|_1 \quad &\leq \quad \theta^k\left[\left(x^0\right)^T s^k + \left(s^0\right)^T x^k\right] \\
&\leq \quad (\theta^k x^0 + (1-\theta^k)x^*)^T s^k + \left(x^k\right)^T(\theta^k s^0 + (1-\theta^k)s^*) \\
&\leq \quad (\theta^k x^0 + (1-\theta^k)x^*)^T(\theta^k s^0 + (1-\theta^k)s^*) + \left(x^k\right)^T s^k \\
&= \quad (\theta^k (x^0)^T s^0 + \theta^k(1-\theta^k)((x^*)^T s^0 + (s^*)^T x^0)) + \left(x^k\right)^T s^k \\
&\leq \quad \theta^k(1+\zeta)(x^0)^T s^0 + \left(x^k\right)^T s^k
\end{aligned}$$

where

$$\zeta = \frac{(x^*)^T s^0 + (s^*)^T x^0}{(x^0)^T s^0} > 1$$

We have by (11)

$$(x^k)^T s^k \geq (1 - \alpha^{k-1})(x^{k-1})^T s^{k-1} \geq \theta^k (x^0)^T s^0$$

And from Assumption 3 it follows that $\zeta \leq \tau$. Thus we have

$$\|(x^k, s^k)\|_1 \leq \frac{2+\tau}{\theta^k \rho_0}(x^k)^T s^k$$

Heretofore we have completed the proof of this lemma.

□

Lemma 3.2 Assume that problems (1) and (2) satisfy Assumption 1–3 stated in the section 2, and the problems have an optimal point$(x^*, y^*, s^*) \in \Phi$. Let the points (x^j, y^j, s^j) with $j = 1, 2, \cdots, k$ generated by Algorithm 2.1 and satisfying the relations (10) and (11). Then there are the following estimates

$$\|D^{-1}\Delta x\| \leq \left(1 + (2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0} + \beta_1 + \nu/n\right)\frac{(x^k)^T s^k}{\min_i (x_i s_i)^{\frac{1}{2}}}, \tag{13}$$

$$\|D\Delta s(\alpha^k)\| \leq \alpha^k \left(1 + (2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0} + \beta_1 + \nu/n\right)\frac{(x^k)^T s^k}{\min_i (x_i s_i)^{\frac{1}{2}}}, \tag{14}$$

where $\Delta s(\alpha^k) = \nabla f(x^k + \alpha^k \Delta x) - \nabla f(x^k) - \alpha^k A^T \Delta y$.
Proof. In the proof, sometimes we will omit the superscript k of x^k and s^k.

Firstly, we consider the following equations: $(\Delta x', \Delta s')$ satisfies

$$\begin{pmatrix} A & 0 \\ S & X \end{pmatrix}\begin{pmatrix} \Delta x' \\ \Delta s' \end{pmatrix} = \begin{pmatrix} 0 \\ p \end{pmatrix} \tag{15}$$

Here $\Delta s' = \nabla f(x + \Delta x') - \nabla f(x) - A^T \Delta y'$. Then by the convexity of $f(x)$, $(\Delta x')^T \Delta s' \geq 0$. And we also have

$$D^{-1}\Delta x' + D\Delta s' = (XS)^{-\frac{1}{2}} p$$

here $D = (X')^{1/2}(S')^{-1/2}$, thus we see that the following estimates are valid.

$$\|D^{-1}\Delta x'\| = \|D(\Delta s' - X^{-1}p)\| \leq \|(XS)^{-\frac{1}{2}} p\| \tag{16}$$

$$\|D\Delta s'\| = \|D^{-1}(\Delta x' - S^{-1}p)\| \leq \|(XS)^{-\frac{1}{2}} p\| \tag{17}$$

It follows from the proof of Lemma 3.1

$$A(\Delta x + \theta^k(x^0 - x^*)) = 0$$

Let $\hat{\Delta x} = \alpha^k(\Delta x + \theta^k(x^0 - x^*))$ and $\hat{\Delta s} = \nabla f(x^k + \hat{\Delta x}) - \nabla f(x^k) - \alpha^k A^T \Delta y$ where Δy is the solution of the system (7) at $k-$iteration of algorithm 2.1, then $(\hat{\Delta x}, \hat{\Delta s})$ satisfies the system (15) with

$$p = -\alpha^k(X^k s^k - \beta_1 \mu^k e) + \alpha^k\theta^k S^k(x^0 - x^*) + X^k(\nabla f(x^k + \alpha^k(\Delta x + \theta^k(x^0 - x^*))) - \nabla f(x^k + \alpha^k \Delta x)) + \alpha^k r^k$$

Now let $(\hat{\Delta x}_1, \hat{\Delta s}_1)$,$(\hat{\Delta x}_2, \hat{\Delta s}_2)$, $(\hat{\Delta x}_3, \hat{\Delta s}_3)$ and $(\hat{\Delta x}_4, \hat{\Delta s}_4)$ satisfying the system (15) when p is $-\alpha^k(X^k s^k - \beta_1\mu^k e)$, $\alpha^k\theta^k S^k(x^0 - x^*)$, $X^k(\nabla f(x^k + \alpha^k(\Delta x + \theta^k(x^0 - x^*))) - \nabla f(x^k + \alpha^k \Delta x))$ and $\alpha^k r^k$ respectively. Then from (16) and (17) it follows that

$$\begin{aligned}
\|\alpha^k D^{-1}\Delta x\| &= \|D^{-1}(\hat{\Delta x}_1 + \hat{\Delta x}_2 + \hat{\Delta x}_3 + \hat{\Delta x}_4 - \alpha^k\theta^k(x^0 - x^*))\| \\
&\leq \|D^{-1}\hat{\Delta x}_1\| + \|D^{-1}\hat{\Delta x}_3\| + \|D^{-1}\hat{\Delta x}_4\| + \|D\hat{\Delta s}_2\| \\
&\leq \alpha^k\|(XS)^{-\frac{1}{2}}(Xs - \beta_1\mu e)\| + \alpha^k\|(XS)^{-\frac{1}{2}} r^k\| + \alpha^k\theta^k\|D(s^0 - s^*)\| \\
&\quad + \|D^{-1}[\nabla f(x + \alpha^k(\Delta x + \theta^k(x^0 - x^*))) - \nabla f(x + \alpha^k \Delta x)]\| \\
&\leq \alpha^k\|(XS)^{-\frac{1}{2}}(Xs - \beta_1\mu e)\| + \alpha^k\|(XS)^{-\frac{1}{2}} r^k\| + \alpha^k\theta^k\|D(s^0 - s^*)\| + \alpha^k\theta^k L\|D^{-1}\| \cdot \|x^* - x^0\|
\end{aligned}$$

Using the Lipschitz condition of $\nabla f(x)$, Assumption 3 and (10) we have

$$\|D^{-1}\Delta x\| \leq \|(XS)^{-\frac{1}{2}}\|(1+\beta_1)x^T s + \nu\|(XS)^{-\frac{1}{2}}\|(x^k)^T s^k/n + \theta^k\|(XS)^{-\frac{1}{2}}\| \cdot \|X\| \cdot \|(s^0 - s^*)\|$$

$$
\begin{aligned}
& +\theta^k L\|(XS)^{-\frac{1}{2}}\| \cdot \|S\| \cdot \|(x^0 - x^*)\| \\
\leq\ & \frac{1}{\min\limits_i (x_i s_i)^{\frac{1}{2}}}(1+\beta_1+\nu/n)x^T s + (1+L)\frac{1}{\min\limits_i (x_i s_i)^{\frac{1}{2}}}\theta^k \frac{(2+\tau)(x^k)^T s^k}{\theta^k \rho_0} n^{\frac{1}{2}}\rho \\
\leq\ & \left(1+(2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}+\beta_1+\nu/n\right)\frac{(x^k)^T s^k}{\min\limits_i (x_i s_i)^{\frac{1}{2}}}
\end{aligned}
$$

where the second inequality follows from Lemma 3.1.

Similarly we can also get

$$
\begin{aligned}
\|D\Delta s(\alpha^k)\| &= \|D(\hat{\Delta s_1}+\hat{\Delta s_2}+\hat{\Delta s_3}+\hat{\Delta s_4} - [\nabla f(x+\alpha^k(\Delta x+\theta^k(x-x^0))) - \nabla f(x+\alpha^k \Delta x)])\| \\
&\leq \alpha^k \left(1+(2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}+\beta_1+\nu/n\right)\frac{(x^k)^T s^k}{\min\limits_i (x_i s_i)^{\frac{1}{2}}}
\end{aligned}
$$

□

Next lemma shows that the complementarity gap of primal-problem (1) and dual-problem (2) does not decrease too much at every iteration, and the sequence (x^k, s^k), which is generated by algorithm 2.1, is positive.

Lemma 3.3 Assume that problems(1) and (2) have an optimal point (x^*, y^*, s^*) satisfying $\|(x^*, s^*)\|_\infty \leq \rho$, $(x^{k+1}, y^{k+1}, s^{k+1})$ is generated at the $(k+1)-$iteration of algorithm 2.1, then for any $k \geq 0$ and α^k in form of (9) we have

$$
\begin{aligned}
(x^{k+1}, s^{k+1}) &> 0, && (18)\\
(x^{k+1})^T s^{k+1} &\geq (1-\alpha^k)(x^k)^T s^k. && (19)
\end{aligned}
$$

Before starting the proof, firstly we introduce two auxiliary functions as following

$$
\varphi_1(\alpha^k) = \sum_{i=1}^n |\frac{\alpha^k \Delta x_i}{x_i^k}|^2,
$$
$$
\varphi_2(\alpha^k) = \sum_{i=1}^n |\frac{\Delta s_i(\alpha^k)}{s_i^k}|^2,
$$

which will play an important role in the proof. Obviously, under the condition $(x^k, s^k) > 0$, if $\varphi_1(\alpha^k) < 1$ and $\varphi_2(\alpha^k) < 1$ are valid, then we see that $(x^{k+1}, s^{k+1}) = (x^k + \alpha^k \Delta x, s^k + \Delta s(\alpha^k)) > 0$ must be satisfied. Moreover, under the same assumption, $(x^k, s^k) > 0$, we have

$$
\varphi_1(\alpha^k) = \sum_{i=1}^n |\frac{\alpha^k \Delta x_i}{x_i^k}|^2 = \sum_{i=1}^n \alpha^2 |\frac{D_{ii}}{x_i^k}|^2 (D_{ii}^{-1}\Delta x_i)^2 \leq \alpha^2 \frac{1}{\min\limits_i \{x_i^k s_i^k\}}\|D^{-1}\Delta x\|^2 \tag{20}
$$

And similarly we have

$$
\varphi_2(\alpha^k) \leq \frac{1}{\min\limits_i \{x_i^k s_i^k\}}\|D\Delta s(\alpha^k)\|^2 \tag{21}
$$

Proof of Lemma 3.3. The proof will be by induction. Firstly, we consider the case: $k = 0$

It is obvious that $(x^0, s^0) > 0$. Then we have by the linear equation of (7)

$$
A(\Delta x + (x^0 - x^*)) = 0
$$

Under the assumption which problems(1) and (2) have an optimal point (x^*, y^*, s^*) satisfying $\|(x^*, s^*)\|_\infty \le \rho$, using similar analyzing method of Lemma 3.2, then we have

$$\begin{aligned}\|D^{-1}\Delta x\| &\le \|(X^0S^0)^{-\frac{1}{2}}\|(1+\beta_1)(x^0)^Ts^0 + \|(X^0S^0)^{-\frac{1}{2}}\|\nu(x^0)^Ts^0/n + \theta^0\|(X^0S^0)^{-\frac{1}{2}}\|\cdot\|X^0\|\cdot\|(s^0-s^*)\| \\ &\quad +\theta^0L\|(X^0S^0)^{-\frac{1}{2}}\|\cdot\|S^0\|\cdot\|(x^0-x^*)\| \\ &\le (1+\beta_1+\nu/n)\frac{(x^0)^Ts^0}{\min\limits_i(x_i^0s_i^0)^{1/2}} + (1+L)\frac{1}{\min\limits_i(x_i^0s_i^0)^{1/2}}\frac{(x^0)^Ts^0}{\rho_0}\rho n^{1/2} \\ &\le \left(1+\beta_1+\nu/n+\frac{\rho}{\rho_0}(1+L)n^{1/2}\right)\frac{(x^0)^Ts^0}{\min\limits_i(x_i^0s_i^0)^{1/2}}\end{aligned}$$

And similarly,

$$\|D\Delta s(\alpha^0)\| \le \alpha^0\left(1+\beta_1+\nu/n+\frac{\rho}{\rho_0}(1+L)n^{1/2}\right)\frac{(x^0)^Ts^0}{\min\limits_i(x_i^0s_i^0)^{1/2}}$$

And by the fact that $\tau > 1$, $\beta_1 + \nu/n < \beta_2 < 1$, $\nu < 1$, (9) with $k = 0$, (20) and (21) we easily see that $\varphi_1(\alpha^0) < 1$ and $\varphi_2(\alpha^0) < 1$ hold.

That is to say, it is proved that

$$(x^1, s^1) = (x^0 + \alpha^0\Delta x, s^0 + \Delta s(\alpha^0)) > 0$$

Furthermore, the inequality follows

$$\begin{aligned}(x^1)^Ts^1 &= (x^0+\alpha^0\Delta x)^T(s^0+\Delta s(\alpha^0)) \\ &= (x^0)^Ts^0 + (\alpha^0(s^0)^T\Delta x + (x^0)^T\Delta s(\alpha^0)) + \alpha^0(\Delta x)^T\Delta s(\alpha^0) \\ &= (x^0)^Ts^0 - \alpha^0(1-\beta_1)(x^0)^Ts^0 + \alpha^k(\sum_{i=1}^n r_i^0) + \alpha^0(\Delta x)^T\Delta s(\alpha^0) \\ &= (1-\alpha^0)(x^0)^Ts^0 + \alpha^0\beta_1(x^0)^Ts^0 + \alpha^k(\sum_{i=1}^n r_i^0) + \alpha^0(\Delta x)^T\Delta s(\alpha^0) \\ &\ge (1-\alpha^0)(x^0)^Ts^0 + \alpha^0(\beta_1-\nu/n)(x^0)^Ts^0 - \alpha^0\|D^{-1}\Delta x\|\cdot\|D\Delta s(\alpha^0)\|\end{aligned}$$

Then by the definition of α^0, the estimates of $\|D^{-1}\Delta x\|$ and $\|D\Delta s(\alpha^0)\|$ above, $0 < \beta_1 < 1$, $0 < \sigma < 1$ and $\sigma\beta_1 - 2\nu < \beta_1 - \nu/n < \beta_2 < 1$ we see that for α^0 with the form of (9) the following is satisfied

$$\alpha^0\beta_1(x^0)^Ts^0 - \alpha^0\|D^{-1}\Delta x\|\cdot\|D\Delta s(\alpha^0)\| \ge 0$$

That is to say, we finish the proof of the case $k = 0$: $(x^1, s^1) > 0$ and $(x^1)^Ts^1 \ge (1-\alpha^0)(x^0)^Ts^0$
Thus, we can assume that for any positive integer k the following is valid.

$$(x^k, s^k) > 0 \tag{22}$$
$$(x^k)^Ts^k \ge (1-\alpha^{(k-1)})(x^{k-1})^Ts^{k-1} \tag{23}$$

For the case: $k+1$. From hyperthesis of this lemma and (22) - (23) we see that Lemma 3.1 and 3.2 hold.
Thus, by Lemma 3.2, α^k in form of (9) and (20) - (21) we easily obtain that $\varphi_1(\alpha^k) < 1$ and $\varphi_2(\alpha^k) < 1$ hold, it follows that

$$(x^{k+1}, s^{k+1}) > 0 \tag{24}$$

In addition, we have

$$\begin{aligned}
(x^{k+1})^T s^{k+1} &= (x^k + \alpha^k \Delta x)^T (s^k + \Delta s(\alpha^k)) \\
&= (x^k)^T s^k + (\alpha^k (s^k)^T \Delta x + (x^k)^T \Delta s(\alpha^k)) + \alpha^k (\Delta x)^T \Delta s(\alpha^k) \\
&= (x^k)^T s^k - \alpha^k (1-\beta_1)(x^k)^T s^k + \alpha^k (\sum_{i=1}^n r_i^k) + \alpha^k (\Delta x)^T \Delta s(\alpha^k) \\
&= (1-\alpha^k)(x^k)^T s^k + \alpha^k \beta_1 (x^k)^T s^k + \alpha^k (\sum_{i=1}^n r_i^k) + \alpha^k (\Delta x)^T \Delta s(\alpha^k) \\
&\geq (1-\alpha^k)(x^k)^T s^k + \alpha^k (\beta_1 - \nu/n)(x^k)^T s^k - \alpha^k \nu (x^k)^T s^k / n + \alpha^k \|D^{-1}\Delta x\| \cdot \|D \Delta s(\alpha^k)\| \\
&\geq (1-\alpha^k)(x^k)^T s^k \\
&\quad + \alpha^k (\beta_1 - \frac{\nu}{n})(x^k)^T s^k - (\alpha^k)^2 \left(1 + (2+\tau)(1+L)n^{\frac{1}{2}} \frac{\rho}{\rho_0} + \beta_1 + \frac{\nu}{n}\right)^2 \frac{((x^k)^T s^k)^2}{\min_i (x_i s_i)}
\end{aligned}$$

where the equation (7) implies the third equality, and the last inequality follows from Lemma 3.2.

Thus, by the definition (9) of α^k and $\sigma\beta_1 - 2\nu < \beta_1 - \nu/n$, we see that

$$(\beta_1 - \frac{\nu}{n})(x^k)^T s^k - \alpha^k \left(1 + (2+\tau)(1+L)n^{\frac{1}{2}} \frac{\rho}{\rho_0} + \beta_1 + \frac{\mu}{n}\right)^2 \frac{((x^k)^T s^k)^2}{\min_i (x_i s_i)} > 0$$

is valid, therefore, $(x^{k+1})^T s^{k+1} \geq (x^k)^T s^k$ hold.

Hereunto, we have completed the proof of Lemma 3.3

□

From the proof above we note that, for $k = 0$, $(\Delta x, \Delta s(\alpha^k))$ also satisfies (13) and (14).

In next lemma firstly we will show that the complementarity gap of problems (1) and (2) will be reduced at each iteration of algorithm 2.1, and it is also shown that the sequence $\{(x^k, y^k, s^k)\}$ is remained in the neighborhood $\mathcal{N}$ of the central path.

Lemma 3.4 Suppose that $(x^{k+1}, y^{k+1}, s^{k+1})$ is generated by algorithm 2.1, $(x^k, s^k) \in \mathcal{N}$. And assume Lemma 3.1 - 3.3 hold. Then the following statements are satisfied

$$\begin{aligned}
(x^{k+1})^T s^{k+1} &\leq (1 - \alpha^k (1-\beta_2))(x^k)^T s^k \qquad (25) \\
(x^{k+1}, s^{k+1}) &\in \mathcal{N} \qquad (26)
\end{aligned}$$

for

$$\alpha^k = \left(\frac{\rho_0}{\rho}\right)^2 \frac{\min\{\sigma\beta_1 - 2\nu, \beta_2 - \beta_1 - \nu/n\} \min_i (x_i^k s_i^k)}{2n(1+(2+\tau)^2 \left(1+L)n^{\frac{1}{2}}\right)^2 (x^k)^T s^k}$$

Proof. Firstly, we have by (7)

$$\begin{aligned}
X^{k+1} s^{k+1} &= (X^k + \alpha^k \Delta X)(s^k + \Delta s(\alpha^k)) \\
&= X^k s^k + (X^k \Delta s(\alpha^k) + \alpha^k S^k \Delta x) + \alpha^k \Delta X \Delta s(\alpha^k) \\
&= X^k s^k - \alpha^k (X^k s^k - \beta_1 \frac{(x^k)^T s^k}{n} e) + \alpha^k r^k + \alpha^k \Delta X \Delta s(\alpha^k) \qquad (27)
\end{aligned}$$

and

$$(x^{k+1})^T s^{k+1} = (x^k)^T s^k - \alpha^k (1-\beta_1)(x^k)^T s^k + \alpha^k (\sum_{i=1}^n r_i^k) + \alpha^k \Delta x^T \Delta s(\alpha^k) \qquad (28)$$

We here consider the inequality (25).

$$
\begin{aligned}
(1-\alpha^k(1-\beta_2))(x^k)^T s^k - (x^{k+1})^T s^{k+1} &= \alpha^k(\beta_2-\beta_1)(x^k)^T s^k - \alpha^k \Delta x^T \Delta s(\alpha^k) - \alpha^k(\sum_{i=1}^n r_i^k) \\
&\geq \alpha^k(\beta_2-\beta_1)(x^k)^T s^k - \alpha^k\|r^k\|_1 - \alpha^k\|D^{-1}\Delta x\|\cdot\|D\Delta s(\alpha^k)\| \\
&\geq \alpha^k(\beta_2-\beta_1-\nu/n)(x^k)^T s^k \\
&\quad -(\alpha^k)^2\left[\left(1+\beta_1+\nu/n+(2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}\right)^2 \frac{((x^k)^T s^k)^2}{\min_i x_i^k s_i^k}\right]
\end{aligned}
$$

If the righthand of the last inequality above is not smaller than zero, i.e.,

$$
\alpha^k \leq \frac{(\beta_2-\beta_1-\nu/n)\min_i x_i^k s_i^k}{\left(1+\beta_1+\nu/n+(2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}\right)^2 (x^k)^T s^k} \tag{29}
$$

then the inequality (25) must be valid.

Obviously, by (9) the inequality (29) is satisfied. The proof of (25) is end.

For the statement (26), from Lemma 3.3, we just need to prove the following inequality

$$
\|X^{k+1}s^{k+1} - \frac{(x^{k+1})^T s^{k+1}}{n} e\| \leq \sigma \frac{(x^{k+1})^T s^{k+1}}{n}
$$

Then, from (27) and (28), we obtain

$$
\begin{aligned}
\|X^{k+1}s^{k+1} - \frac{(x^{k+1})^T s^{k+1}}{n} e\| &\leq (1-\alpha^k)\|X^k s^k - \frac{(x^k)^T s^k}{n} e\| + \alpha^k\|r^k - \frac{\sum_{i=1}^n r_i^k}{n} e\| \\
&\quad +\alpha^k\|\Delta X \Delta s(\alpha^k) - \frac{\Delta x^T \Delta s(\alpha^k)}{n} e\|
\end{aligned}
$$

By the fact that $\frac{\Delta x^T \Delta s(\alpha^k)}{n} e$ and $r^k - \frac{\sum_{i=1}^n r_i^k}{n} e$ are the orthogonal projections of $\Delta X \Delta s(\alpha^k)$ and r^k on the one dimensional subspace by e respectively, we deduce that

$$
\|X^{k+1}s^{k+1} - \frac{(x^{k+1})^T s^{k+1}}{n} e\| \leq (1-\alpha^k)\sigma \frac{(x^k)^T s^k}{n} + \alpha^k\|r^k\| + \alpha^k\|\Delta X \Delta s(\alpha^k)\|
$$

Thus,

$$
\begin{aligned}
\sigma\frac{(x^{k+1})^T s^{k+1}}{n} &- \|X^{k+1}s^{k+1} - \frac{(x^{k+1})^T s^{k+1}}{n} e\| \\
&\geq \sigma\frac{(x^{k+1})^T s^{k+1}}{n} - ((1-\alpha^k)\sigma\frac{(x^k)^T s^k}{n} + \alpha^k\|r^k\| + \alpha^k\|\Delta X \Delta s(\alpha^k)\|) \\
&\geq \alpha^k\sigma\beta_1\frac{(x^k)^T s^k}{n} - \alpha^k(\|r^k\| + |(\sum_{i=1}^n r_i^k)/n|) - \alpha^k(\|\Delta X \Delta s(\alpha^k)\| + |\frac{\Delta x^T \Delta s(\alpha^k)}{n}|) \\
&\geq \alpha^k(\sigma\beta_1 - 2\nu)\frac{(x^k)^T s^k}{n} - 2\alpha^k\|D^{-1}\Delta x\|\cdot\|D\Delta s(\alpha^k)\| \\
&\geq \alpha^k(\sigma\beta_1 - 2\nu)\frac{(x^k)^T s^k}{n} - 2(\alpha^k)^2\left(1+\beta_1+\nu/n+(2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}\right)^2 \frac{((x^k)^T s^k)^2}{\min_i x_i^k s_i^k}
\end{aligned}
$$

By (9), we see that the following inequality holds

$$(\sigma\beta_1 - 2\nu)\frac{(x^k)^T s^k}{n} - 2\alpha^k \left(1 + \beta_1 + \nu/n + (2+\tau)(1+L)n^{\frac{1}{2}}\frac{\rho}{\rho_0}\right)^2 \frac{((x^k)^T s^k)^2}{\min_i x_i^k s_i^k} \geq 0$$

Thus, it is valid that $(x^{k+1}, y^{k+1}, s^{k+1}) \in \mathcal{N}$.

We have gotten all of our desired results.

□

Theorem 3.1 Suppose the problems (1) and (2) has an optimal point in Φ and satisfy Assumption 1-4. Let the starting point (x^0, y^0, s^0) is also taken from the set $\mathcal{N}$. Then Algorithm 2.1 will terminate in at most

$$\left\lceil \frac{\Xi}{-\ln(1-\tilde{\alpha})} \right\rceil \tag{30}$$

iterations, where

$$\tilde{\alpha} = \left(\frac{\rho_0}{\rho}\right)^2 \frac{\min\{(\sigma\beta_1 - 2\nu), (\beta_2 - \beta_1 - \nu/n)\}(1-\sigma)}{2(2+\tau)^2\ n^2\left(1 + (1+L)n^{\frac{1}{2}}\right)^2} \tag{31}$$

$$\Xi = \max\left\{\frac{\ln(x^0)^T s^0}{\varepsilon}, \frac{\ln\|Ax^0 - b\|}{\varepsilon_p}\right\} \tag{32}$$

and $\lceil\xi\rceil$ denotes the smallest integer greater than or equal to ξ.

Proof. By Lemma 2.1, we deduce $\alpha^k \geq \tilde{\alpha}$.

From Lemma 3.4, we have

$$(x^{k+1})^T s^{k+1} \leq (1 - \alpha^k(1-\beta_2))(x^k)^T s^k \leq (1 - \widetilde{\alpha}(1-\beta_2))(x^k)^T s^k$$
$$\|Ax^{k+1} - b\| = (1-\alpha^k)\|Ax^k - b\| \leq (1-\widetilde{\alpha})\|Ax^k - b\|$$

Then easily we obtain the result of Theorem 3.1.

□

Througout the analysis above, it uses the optimal point (x^*, y^*, s^*) in the proof of Lemma 3.1 - 3.3. The proof of Lemma 3.4 and Theorem 3.1 just use the result of Lemma 3.1 - 3.3. So we now give out a theorem with the detection of infeasibility.

Theorem 3.2 Suppose that one lets the following two criticals,

"Step 1$'$ If (x^k, y^k, s^k) does not satisfy one of (12), (13) and (14), then terminated."
in between Step 1 and Step 2, and

"Step 3$'$ If $(\Delta x^k, \Delta s(\alpha^k))$ with α^k does not satisfy one of (17) and (18), then terminated."
in between Step 3 and Step 4, of Algorithm 2.1. Let the starting point (x^0, y^0, s^0) is also taken from the set $\mathcal{N}$. Then Algorithm 2.1 will terminate in at most

$$\left\lceil \frac{\Xi}{-\ln(1-\tilde{\alpha})} \right\rceil \tag{33}$$

iterations.

And if it terminates by Step 1, (x^k, y^k, s^k) is the ε-optimal solution of the primal-dual problem (1) and (2); or else, it terminates by Step 1$'$ or Step 3$'$, there is no optimal $(x^*, y^*, s^*)$of the problems

(1) and (2) satisfying $\|(x^*, s^*)\|_\infty \leq \rho$.

Obviously, according to the proof of Lemma 3.1 -3.3, if Algorithm 2.1 terminates by Step $1'$ or Step $3'$, we can easily get the conclusion above by reduction to absurdity.

We note that, in Theorem 3.1 and 3.2, the complexities dependent no only on the the input data of problems but on the starting point and the region where we want to find an optimal point.

4. Concluding Remark

In this section we come back to consider the approximate solution of the nonlinear system

$$\Psi_k(\Delta x, \Delta y) := X^k(\nabla f(x^k + \alpha^k \Delta x) - \nabla f(x^k) - \alpha^k A^T \Delta y) - S^k \Delta x - \alpha^k (Xs - \frac{x^T s}{n} e) = 0 \quad (34)$$

$$A\Delta x = -(Ax^k - b) \quad (35)$$

Easily, the Jacobi matrix of the system above is the form as

$$\begin{pmatrix} Q_k & \alpha^k A^T \\ A & 0 \end{pmatrix}$$

If we use Newton-like method to solve the nonlinear system, and we firstly put the starting point satisfying $A\Delta x = -(Ax^k - b)$, then we just need to solve the following linear system

$$Q_k \hat{\Delta x} - \alpha^k A^T \hat{\Delta y} = -\Psi_k(\Delta x, \Delta s) \quad (36)$$

$$A\hat{\Delta x} = 0 \quad (37)$$

and let $\tilde{\Delta x} = \Delta x + \hat{\Delta x}, \tilde{\Delta y} = \Delta y + \hat{\Delta y}$. Thus we see that it, $A\tilde{\Delta x} = -(Ax^k - b)$, will be satisfied by all iterates with this method. Other iterative methods may also be used to solve the nonlinear system (34) and (35).

Moreover, if we also induce the inexact technic to the linear equation in (7), then at each iterative, the property (6) cannot be remained. Then the analysis in Section 3 cannot work. it is our future work how to analysis our algorithm in the case. And for this case it has been considered for linear programming, readers may find it in papers Refs. 19 and 20.

References

1. Nesterov, Y., and Nemirovski, A., Interior Point Polynomial Algorithms In Convex Programming. SIAM Publications, Philadelphia, PA. 1994.

2. Hertog, D., Interior point approach to linear, quadratic and convex programming. Mathematics and its applications, 277. Kluwer Academic Publishers Group, 1994.

3. Wright, S., Primal-Dual Interior-Point Methods. SIAM Philadelphia, 1997.

4. Ye, Y., Interior Point Algorithms: Theory and Analysis. New York: Wiley, 1997.

5. Kortanek, K.O., Ptra, F., and Ye, Y., On some efficient interior point methods for nonlinear convex programming. Linear Algebra and its Applications 152(1991), 169-189.

6. Güler, O., Hyperbolic polynomials and interior point methods for convex programming. Math. Oper. Res. 22(1997), 350-377

7. Portra, F.A., and Wright, S., Interior point methods, J. Comput. Appl. Math. 124(2000), no. 1-2, 281-302.

8. Shi, Y., A combination of potential reduction steps and steepest decent steps for solving convex programming prblems. Numerical Linear Algebra with Applications 9(2002), 195-203.

9. Portra, F.A., and Ye, Y., Interior-point methods for nonlinear complementarity problems. J. Optim. Theory Appl. 88(1996), no. 3, 617-642.

10. Ralph, D., and Wright, S., Superlinear convergence of an interior point method for monotone variational inequalities. in Complementarity and Variational Problems: state of the Art. SIAM, 1997.

11. Ralph, D., and Wright, S., Superlinear convergence of an interior-point method depite dependent constrains. Math. Oper. Res. 25(2000), no. 2, 179-194

12. Kojima, M., Megiddo, N. and Mizuo, S., a primal-dual infeasible-interior-point algorithms for linear programming. Math. Programming 61(1993), 263-280

13. Kojima, M., Noma, T., and Yoshise, A., Global convergence of an infeasible-interior-point methods. Math. Programming 65(1994), 43-72.

14. Mizuno, S., Polynmial of infeasible-interior-point algoritms for linear programming. Math. Programming 67(1995), 109-119

15. Monteiro, R.D.C., and Zhou, F., On superlinear convergence of infeasible interior-point algorithms for linearly constrained convex programming. Comput. Optim. Appl. 8(1997), 245-262.

16. Yamashita, N., Kanzow, C., Morimoto, T. and Fukushima, M., An infeasible interior proximal method for convex programming problems with linear constraints. J. Nonlinear Convex Anal. 2(2001), 139-156.

17. Wright, S., and Ralph, D., A superlinear infeasible-interior-point algorithm for monotone complementarity problems. Math. Oper. Res. 21(1996), no. 4, 815-838.

18. Tseng, P., An infeasible path-following method for monotone complementarity problems. SIAM J. Optim. 7(1997), no. 2, 386-402.

19. Mizuno, S., and Jarre, F., Global and polynomial-time convergence of an infeasible-interior-point algorithm using inexact computation. Math. Program. 84(1999), no. 1, 105-122.

20. Korzak, J., Convergence analysis of inexact infeasible-interior-point algorithms for solving linear programming problems. SIAM J. Optim. 11(2000), no. 1, 133-148.