Journal of Computational Mathematics, Vol.13, No.2, 1995, 161-171.

PREDICTOR-CORRECTOR ALGORITHM FOR CONVEX QUADRATIC PROGRAMMING WITH UPPER BOUNDS*

Guo Tian-de
(*Mathematics Department of Teacher's College, Qingdao University, Shandong, China*)
Wu Shi-quan
(*Institute of Applied Mathematics, Academia Sinica, Beijing, China*)

Abstract

Predictor-corrector algorithm for linear programming, proposed by Mizuno et al.[1], becomes the best well known in the interior point methods. The purpose of this paper is to extend these results in two directions. First, we modify the algorithm in order to solve convex quadratic programming with upper bounds. Second, we replace the corrector step with an iteration of Monteiro and Adler's algorithm[2]. With these modifications, the duality gap is reduced by a constant factor after each corrector step for convex quadratic programming. It is shown that the new algorithm has a $O(\sqrt{n}L)$-iteration complexity.

1. Introduction

The predictor-corrector method for linear programming is a well known interior point method developed by Mizuno et al.[1], due to its quadratically convergent analysis. This kind of analysis usually contains two steps, i.e., predictor step and corrector step as one iteration. The corrector step is used only to ensure that the iterates stay close to the central path so that large step can be taken during the predictor step. The duality gap remains unchanged at corrector step for linear programming, but in case of convex quadratic programming, as shown later of this paper, this gap even increases. In this paper, we extend these results in order to solve convex quadratic programming with upper bounds. The predictor directions generated by our algorithm are similar to those generated by the algorithm presented in [1]. However, the corrector directions are replaced by the Monteiro and Adler's algorithm[2]. With these modifications, the duality gap is reduced by a constant factor after each corrector step. Therefore, we obtain a faster algorithm for convex quadratic programming.

The paper is organized as follows. In section 2, we outline the procedure of a predictor-corrector method. In section 3, we present convergence results for the algorithm. Final section contains further discussions.

* Received March 12, 1994.

2. The Algorithm

We consider the following quadratic programming problem in standard form

$$(\mathrm{QP})\qquad \begin{aligned} \min\quad & c^Tx+\tfrac{1}{2}x^TQx \\ \text{s.t.}\quad & \mathrm{Ax}=\mathrm{b}, \\ & x+z=d, \\ & x\geq 0,\quad z\geq 0, \end{aligned}$$

where $c\in R^n, A\in R^{m\times n}$, $b\in R^m$, $d\in R^n$, and $Q\in R^{n\times n}$ are given, and Q is positive semi-definite, $x\in R^n$, $z\in R^n$, and the superscript T denotes the transpose. The standard logarithmic barrier interior point method is to incorporate the inequalities into a logarithmic barrier term and then append it to the objective function to obtain the following problem

$$(\mathrm{QP}_\mu)\qquad \begin{aligned} \min\quad & c^Tx+\tfrac{1}{2}x^TQx-\mu\textstyle\sum_{i=1}^n \ln x_i-\mu\sum_{i=1}^n \ln z_i \\ \text{s.t.}\quad & \mathrm{Ax}=\mathrm{b}, \\ & x+z=d. \end{aligned}$$

The first order optimality conditions for (QP) are

$$Ax-b=0, \tag{1}$$
$$x+z-d=0, \tag{2}$$
$$A^Ty+s-w-Qx-c=0, \tag{3}$$
$$XSe=0, \tag{4}$$
$$ZWe=0, \tag{5}$$
$$x,z,s,w\geq 0, \tag{6}$$

where X,S,Z and W are diagonal matrices with the elements x_i,s_i,z_i, and w_i respectively, y,w and s are dual variables and e denotes the n dimensional vector of all 1's Similarly, the first order optimality conditions for (QP_μ) are

$$Ax-b=0, \tag{7}$$
$$x+z-d=0, \tag{8}$$
$$A^Ty+s-w-Qx-c=0, \tag{9}$$
$$XSe-\mu e=0, \tag{10}$$
$$ZWe-\mu e=0. \tag{11}$$

The primal-dual method, proposed by Monteiro and Adler[2], Carpenter [3], applies Newton's method directly to (7)-(11). Denote by F the set of all (x,z) and (y,s,w) that are feasible for the primal and dual, respectively. Denote by F^0 the set of all points with $(x, z, s, w) > 0$ in F. We note that (7)-(11) define a path in F^0:

$$C = \left\{ (x, z, y, s, w) \in \mathrm{F}^0 : \begin{pmatrix} Xs \\ Zw \end{pmatrix} = \frac{\mathrm{x^T s + z^T w}}{\mathrm{2n}} \mathrm{e} \right\}, \tag{12}$$

which is usually called central path for (QP). We define the neighbourhood of the path as

$$N(\alpha) = \left\{ (x, z, y, s, w) \in \mathrm{F}^0 : \left\| \begin{pmatrix} Xs \\ Zw \end{pmatrix} - \mu \mathrm{e} \right\| \le \frac{\alpha}{2}\mu, \text{ where } \mu = \frac{\mathrm{x^T s + z^T w}}{\mathrm{2n}} \right\}, \tag{13}$$

for $\alpha \in (0, 1)$. We obtain the predictor direction of predictor-corrector method by applying Newton's method to (1)-(6). This entails solving the system

$$Ad_{xp} = 0, \tag{14}$$
$$d_{xp} + d_{zp} = 0, \tag{15}$$
$$A^T d_{yp} + d_{sp} - d_{wp} - Qd_{xp} = 0, \tag{16}$$
$$Sd_{xp} + Xd_{sp} = -Xs, \tag{17}$$
$$Wd_{zp} + Zd_{wp} = -Zw, \tag{18}$$

its solution is the predictor direction $(d_{xp}, d_{zp}, d_{yp}, d_{sp}, d_{wp})$. As the predictor step may stray from the central path, we need to pull back a little bit along corrector direction $(d_{xc}, d_{zc}, d_{yc}, d_{sc}, d_{wc})$ which is obtained by solving the system

$$Ad_{xc} = 0, \tag{19}$$
$$d_{xc} + d_{zc} = 0, \tag{20}$$
$$A^T d_{yc} + d_{sc} - d_{wc} - Qd_{xc} = 0, \tag{21}$$
$$Sd_{xc} + Xd_{sc} = \bar{\mu}e - Xs, \tag{22}$$
$$Wd_{zc} + Zd_{wc} = \bar{\mu}e - Zw. \tag{23}$$

Algorithm

Initialization

Let $\alpha \in (0, 1)$ and $\tau > 0$ be such that

$$\frac{\alpha^2 + \tau^2}{1 - \alpha} \le \alpha \left(1 - \frac{\tau}{\sqrt{2n}}\right), \quad \delta = \frac{\tau}{\sqrt{2n}} - \frac{\alpha^2 + \tau^2}{8n(1 - \alpha)} > 0. \tag{24}$$

Set $k = 0$.

One can check that (24) is satisfied if $\alpha = \tau = 0.25$. Take a strictly feasible point $(x^0, z^0, y^0, s^0, w^0)$, such that

$$\left\| \begin{pmatrix} X^0 s^0 \\ Z^0 w^0 \end{pmatrix} - \mu^0 e \right\| \le \frac{\alpha}{2}\mu^0.$$

Predictor step

For each even integer $2k$ ($k = 0, 1, 2, \cdots$), we have $(x^{2k}, z^{2k}, y^{2k}, s^{2k}, w^{2k}) \in F^0$ satisfying

$$\left\| \begin{pmatrix} X^{2k}s^{2k} \\ Z^{2k}w^{2k} \end{pmatrix} - \mu^{2k}e \right\| \le \frac{\alpha}{2}\mu^{2k}.$$

Replace (x, z, y, s, w) by $(x^{2k}, z^{2k}, y^{2k}, s^{2k}, w^{2k})$ and solve the linear system (14)-(18) in $(d_{xp}, d_{zp}, d_{yp}, d_{sp}, d_{wp})$. Let $(d_{xp}^{2k}, d_{zp}^{2k}, d_{yp}^{2k}, d_{sp}^{2k}, d_{wp}^{2k})$ be a solution and let

$$\begin{aligned} x(\theta) &= x^{2k} + \theta d_{xp}^{2k}, \\ z(\theta) &= z^{2k} + \theta d_{zp}^{2k}, \\ y(\theta) &= y^{2k} + \theta d_{yp}^{2k}, \\ s(\theta) &= s^{2k} + \theta d_{sp}^{2k}, \\ w(\theta) &= w^{2k} + \theta d_{wp}^{2k}. \end{aligned}$$

Calculate

$$\theta^{2k} = \max \{\theta : (x(\theta), z(\theta), y(\theta), s(\theta), w(\theta)) \in N(2\alpha)\}.$$

Set

$$\begin{aligned} x^{2k+1} &= x(\theta^{2k}), \\ z^{2k+1} &= z(\theta^{2k}), \\ y^{2k+1} &= y(\theta^{2k}), \\ s^{2k+1} &= s(\theta^{2k}), \\ w^{2k+1} &= w(\theta^{2k}), \\ \mu^{2k+1} &= \frac{(x^{2k+1})^T s^{2k+1} + (z^{2k+1})^T w^{2k+1}}{2n}. \end{aligned}$$

Corrector Step

For each odd integer $2k+1$ ($k = 0, 1, \cdots$), we have $(x^{2k+1}, z^{2k+1}, y^{2k+1}, s^{2k+1}, w^{2k+1}) \in N(2\alpha)$, in other words

$$\left\| \begin{pmatrix} X^{2k+1}s^{2k+1} \\ Z^{2k+1}w^{2k+1} \end{pmatrix} - \mu^{2k+1}e \right\| \le \alpha\mu^{2k+1}.$$

Replace (x, z, y, s, w) and $\bar{\mu}$ by $(x^{2k+1}, z^{2k+1}, y^{2k+1}, s^{2k+1}, w^{2k+1})$ and $\bar{\mu}^{2k+1}$ respectively and solve the linear system (19)-(23) in $(d_{xc}, d_{zc}, d_{yc}, d_{sc}, d_{wc})$, where $\bar{\mu}^{2k+1} = (1 - \tau/\sqrt{2n})\mu^{2k+1}$. Let $(d_{xc}^{2k+1}, d_{zc}^{2k+1}, d_{yc}^{2k+1}, d_{sc}^{2k+1}, d_{wc}^{2k+1})$ be a solution and let

$$\begin{aligned} x^{2k+2} &= x^{2k+1} + d_{xc}^{2k+1}, \\ z^{2k+2} &= z^{2k+1} + d_{zc}^{2k+1}, \\ y^{2k+2} &= y^{2k+1} + d_{yc}^{2k+1}, \\ s^{2k+2} &= s^{2k+1} + d_{sc}^{2k+1}, \\ w^{2k+2} &= w^{2k+1} + d_{wc}^{2k+1}, \\ \mu^{2k+2} &= \frac{(x^{2k+2})^T s^{2k+2} + (z^{2k+2})^T w^{2k+2}}{2n}. \end{aligned}$$

3. Convergence Results

In this section, we present convergence results for the new algorithm. Firstly, to implement the corrector step, it is necessary that $(x^{2k+1}, z^{2k+1}, y^{2k+1}, s^{2k+1}, w^{2k+1})$ is in $N(2\alpha)$. In the linear case, this condition is satisfied[1]. It is extended to quadratic objective in this section. Secondly, to ensure the predictor step in the next iteration is implementable, it is necessary that $(x^{2k+2}, z^{2k+2}, y^{2k+2}, s^{2k+2}, w^{2k+2})$ obtained from the current corrector step is in $N(\alpha)$.

Clearly the direction $(d_x, d_z, d_y, d_s, d_w)$ satisfies

$$Ad_x = 0, \tag{25}$$

$$d_x + d_z = 0, \tag{26}$$

$$A^T d_y + d_s - d_w - Qd_x = 0, \tag{27}$$

then we can get the following lemma.

Lemma 3.1. *For any $(d_x, d_z, d_y, d_s, d_w)$ satisfying* (25)–(27),

$$d_x^T d_s + d_z^T d_w = d_x^T Q d_x \geq 0. \tag{28}$$

Proof. Using (25)–(27), $d_s = Qd_x + d_w - A^T d_y$ and $d_x = -d_z$, we have

$$\begin{aligned} d_x^T d_s + d_z^T d_w &= d_x^T (Qd_x + d_w - A^T d_y) + d_z^T d_w \\ &= d_x^T Q d_x - d_z^T d_w - (Ad_x)^T d_y + d_z^T d_w \\ &= d_x^T Q d_x \geq 0. \end{aligned}$$

The following lemma is a direct extension of a result in Mizuno et al.[1].

Lemma 3.2. *Suppose $(x, z, y, s, w) \in \mathrm{F}^0$ and $(d_x, d_z, d_y, d_s, d_w)$ satisfies (25)-(27). Let $(x(\theta), z(\theta), y(\theta), s(\theta), w(\theta)) = (x, z, y, s, w) + \theta(d_x, d_z, d_y, d_s, d_w)$. If there exists some $\bar{\theta} < 1$ such that*

$$\left\| \begin{pmatrix} X(\theta)s(\theta) \\ Z(\theta)w(\theta) \end{pmatrix} - \mu(\theta)e \right\| \leq \alpha\mu(\theta), \text{ for all } \theta \in [0, \bar{\theta}],$$

then $(x(\theta), z(\theta), y(\theta), s(\theta), w(\theta)) \in \mathrm{F}^0$, where $\mu(\theta) = (x(\theta)^T s(\theta) + z(\theta)^T w(\theta))/(2n)$.

The following two lemmas describe the functions of the predictor step and the corrector step respectively.

Lemma 3.3. *Suppose, $(x, z, y, s, w) \in \mathrm{N}(\alpha)$ and $(d_{xp}, d_{zp}, d_{yp}, d_{sp}, d_{wp})$ satisfies* (14)–(18). *Let*

$$(x(\theta), z(\theta), y(\theta), s(\theta), w(\theta)) = (x, z, y, s, w) + \theta(d_{xp}, d_{zp}, d_{yp}, d_{sp}, d_{wp}),$$

$$\bar{\theta} = \frac{\sqrt{\alpha^2 + 8n\alpha} - \alpha}{4n},$$

then for any $\theta \in [0, \bar{\theta}]$, $(x(\theta), z(\theta), y(\theta), s(\theta), w(\theta)) \in \mathrm{N}(2\alpha)$, that is

$$\left\| \begin{pmatrix} X(\theta)s(\theta) \\ Z(\theta)w(\theta) \end{pmatrix} - \mu(\theta)e \right\| \leq \alpha\mu(\theta), \quad \text{for any } \theta \in [0, \bar{\theta}],$$

and $\mu(\theta) \le (1-\theta/2)^2\mu \le (1-\theta/2)\mu$, *where* $\mu(\theta) = (x(\theta)^T s(\theta) + z(\theta)^T w(\theta))/(2n)$.

Proof. To simplify the notations, (x, z) and (s, w) are denoted by $\hat{x}$, and $\hat{s}$, (d_{xp}, d_{zp}) and (d_{sp}, d_{wp}) are denoted by $\hat{d}_{xp}$ and $\hat{d}_{sp}$ respectively, let $\hat{n} = 2n$. Then from Lemma 3.1, we have

$$\hat{d}_{xp}^T \hat{d}_{sp} = d_x^T Q d_x \ge 0. \tag{29}$$

From (17) and (18) we can get

$$\hat{X}\hat{d}_{sp} + \hat{S}\hat{d}_{xp} = -\hat{X}\hat{s}. \tag{30}$$

Multiply both sides of (30) by $(\hat{X}\hat{S})^{-\frac{1}{2}}$, we can obtain

$$\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp} + \hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp} = -(\hat{X}\hat{s})^{\frac{1}{2}}. \tag{31}$$

Then, we get

$$\left\|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}\right\|^2 + \left\|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp}\right\|^2 = \left\|(-\hat{X}\hat{s})^{\frac{1}{2}}\right\|^2 - 2\hat{d}_{sp}^T\hat{d}_{xp} = \hat{x}^T\hat{s} - 2\hat{d}_{xp}^T d_{sp}. \tag{32}$$

From (31) and the properties of diagonal matrices, we have

$$\begin{aligned} 0 \le \hat{d}_{xp}^T\hat{d}_{sp} &= (\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp})^T(\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}) \\ &\le \frac{\|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp}\|^2 + \|(\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}\|^2}{2} \\ &= \frac{\hat{x}^T\hat{s} - 2\hat{d}_{xp}^T\hat{d}_{sp}}{2}, \end{aligned}$$

then

$$0 \le \hat{d}_{xp}^T\hat{d}_{sp} \le \frac{1}{4}\hat{x}^T\hat{s}. \tag{33}$$

From the definition of $\hat{x}(\theta)$ and $\hat{s}(\theta)$ we know that

$$\begin{aligned} \hat{X}(\theta)\hat{s}(\theta) &= (\hat{X} + \theta\hat{D}_{xp})(\hat{s} + \theta\hat{d}_{sp}) \\ &= \hat{X}\hat{s} + \theta(\hat{S}\hat{d}_{xp} + \hat{X}\hat{d}_{sp}) + \theta^2\hat{D}_{xp}\hat{d}_{sp} \\ &= (1-\theta)\hat{X}\hat{s} + \theta^2\hat{D}_{xp}\hat{d}_{sp} \end{aligned} \tag{34}$$

where the last equality is due to (30). Next, we use (34) to estimate $\hat{x}(\theta)^T\hat{s}(\theta)$,

$$\hat{x}(\theta)^T\hat{s}(\theta) = e^T\hat{X}(\theta)\hat{s}(\theta) = (1-\theta)\hat{x}^T\hat{s} + \theta^2\hat{d}_{xp}^T\hat{d}_{sp},$$

then

$$\mu(\theta) = \frac{x(\theta)^T s(\theta)}{\hat{n}} = (1-\theta)\mu + \frac{\theta^2}{\hat{n}}\hat{d}_{xp}^T\hat{d}_{sp} \ge (1-\theta)\mu, \tag{35}$$

where the second equality follows from $\mu = \hat{x}^T\hat{s}/\hat{n}$, and the inequality is due to (29). Using the properties of diagonal matrices, we have

$$\begin{aligned} \|\hat{D}_{xp}\hat{d}_{sp}\| &= \|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{D}_{xp}\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}\| \\ &\le \|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp}\| \times \|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}\| \\ &\le \frac{\|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xp}\|^2 + \|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sp}\|^2}{2} \\ &= \frac{\hat{x}^T\hat{s} - 2\hat{d}_{xp}^T\hat{d}_{sp}}{2} = \frac{1}{2}\hat{x}^T\hat{s} - \hat{d}_{xp}^T\hat{d}_{sp}, \end{aligned} \tag{36}$$

where the second equality is due to (32). Combining (34),(35) and (36) yields

$$\begin{aligned}\hat{X}(\theta)\hat{s}(\theta)-\mu(\theta)e &= (1-\theta)\hat{X}\hat{s}+\theta^2\hat{D}_{xp}\hat{d}_{sp}-(1-\theta)\mu e-\frac{\theta^2}{\hat{n}}\hat{d}_{xp}^T\hat{d}_{sp}e\\ &= (1-\theta)(\hat{X}\hat{s}-\mu e)+\theta^2(\hat{D}_{xp}\hat{d}_{sp}-\frac{1}{\hat{n}}\hat{d}_{xp}^T\hat{d}_{sp}e).\end{aligned}$$

Thus

$$\begin{aligned}\left\|\frac{\hat{X}(\theta)\hat{s}(\theta)}{\mu(\theta)}-e\right\| &\le (1-\theta)\frac{\alpha\mu}{2\mu(\theta)}+\theta^2\frac{\|\hat{D}_{xp}\hat{d}_{sp}\|+\hat{d}_{xp}^T\hat{d}_{sp}\|e\|/\hat{n}}{\mu(\theta)}\\ &\le \frac{1}{2}\alpha+\theta^2\frac{\|\hat{D}_{xp}\hat{d}_{sp}\|+\hat{d}_{xp}^T\hat{d}_{sp}/\sqrt{\hat{n}}}{(1-\theta)\mu}\\ &\le \frac{1}{2}\alpha+\theta^2\frac{\hat{x}^T\hat{s}/2-\hat{d}_{xp}^T\hat{d}_{sp}+\hat{d}_{xp}^T\hat{d}_{sp}/\sqrt{\hat{n}}}{(1-\theta)\mu}\\ &\le \frac{1}{2}\alpha+\theta^2\frac{\hat{n}\mu}{2(1-\theta)\mu}\\ &= \frac{1}{2}\alpha+\frac{\hat{n}\theta^2}{2(1-\theta)}\\ &= \frac{1}{2}\alpha+\frac{1}{2}\alpha=\alpha \quad (\text{for } \theta\in[0,\bar{\theta}])\end{aligned} \tag{37}$$

where the first inequality follows from $(x,z,y,s,w)\in N(\alpha)$ and the second inequality is due to (35), the third inequality follows from (36), the fourth inequality is due to the facts that $\sqrt{\hat{n}}\ge 1$ and (29) and the last inequality is due to that the function $f(\theta)=\hat{n}\theta^2/(2(1-\theta)),\ \theta\in[0,1)$ is a monotonic increasing function for $\theta\in[0,1)$, and when $\theta=\bar{\theta}$, $f(\theta)=1/(2\alpha)$. Also from (35) and (33) we get

$$\begin{aligned}\mu(\theta) &\le (1-\theta)\mu+\frac{\theta^2}{4\hat{n}}\hat{x}^T\hat{s}=(1-\theta)\mu+\frac{\theta^2}{4}\mu\\ &= (1-\theta+\frac{\theta^2}{4})\mu=(1-\frac{\theta}{2})^2\mu\\ &\le (1-\frac{\theta}{2})\mu.\end{aligned}$$

This completes the proof.

Lemma 3.4. *Suppose, for some $k\ge 0$, $(x^{2k+1},z^{2k+1},y^{2k+1},s^{2k+1},w^{2k+1})$ is generated by the predictor step, so it satisfies $(x^{2k+1},z^{2k+1},y^{2k+1},s^{k+1},w^{2k+1})\in \mathrm{N}(2\alpha)$, then $(x^{2k+2},z^{2k+2},y^{2k+2},s^{k+2},w^{2k+2})\in \mathrm{N}(\alpha)$. Thus the algorithm is well defined. Moreover, we have $\mu^{2k+2}\le(1-\delta)\mu^{2k+1}$, where $\delta=\frac{\tau}{\sqrt{\hat{n}}}-\frac{\alpha^2+\tau^2}{2\hat{n}(1-\alpha)}>0$.*

Proof. To simplify the notations, (x^{2k+1},z^{2k+1}) and (s^{2k+1},w^{2k+1}) are denoted by $\hat{x}$, and $\hat{s}$, $(d_{xc}^{2k+1},d_{zc}^{2k+1})$ and $(d_{sc}^{2k+1},d_{wc}^{2k+1})$ are denoted by $\hat{d}_{xc}$ and $\hat{d}_{sc}$, μ^{2k+1}, $\bar{\mu}^{2k+1}$ are denoted by μ, $\bar{\mu}$, respectively. Then from Lemma 3.2 and the algorithm, we know that $\hat{x}>0$, $\hat{s}>0$ and

$$\hat{X}\hat{d}_{sc}+\hat{S}\hat{d}_{xc}=-(\hat{X}\hat{s}-\bar{\mu}e). \tag{38}$$

Also, we use superscript '+' to denote the superscript $2k+2$, that is, $\hat{x}^+$ and $\hat{s}^+$ denote (x^{2k+2},z^{2k+2}) and (s^{2k+2},w^{2k+2}), respectively, and μ^+ denotes μ^{2k+2}.

Multiplying both sides of (38) by $(\hat{X}\hat{S})^{-\frac{1}{2}}$, one obtains

$$\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc} + \hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xc} = (\hat{X}\hat{S})^{-\frac{1}{2}}(\bar{\mu}e - \hat{X}\hat{s}). \tag{39}$$

Then we get

$$\left\|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}\right\|^2 + \left\|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xc}\right\|^2 = \left\|(\hat{X}\hat{S})^{-\frac{1}{2}}(\bar{\mu}e - \hat{X}\hat{s})\right\|^2 - 2\hat{d}_{sc}^T\hat{d}_{xc}. \tag{40}$$

Note that from (40) and the properties of diagonal matrices, we have

$$\begin{aligned}\|\hat{D}_{xc}\hat{d}_{sc}\| &= \|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{D}_{xc}\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}\| \\ &\le \|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xc}\| \times \|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}\| \\ &\le \frac{\|\hat{S}^{\frac{1}{2}}\hat{X}^{-\frac{1}{2}}\hat{d}_{xc}\|^2 + \|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}\|^2}{2} \\ &= \frac{\|(\hat{X}\hat{S})^{-\frac{1}{2}}(\bar{\mu}e - \hat{X}\hat{s})\|^2 - 2\hat{d}_{sc}^T\hat{d}_{xc}}{2}\end{aligned} \tag{41}$$

Due to the fact that $(x^{2k+1}, z^{2k+1}, y^{2k+1}, s^{2k+1}, w^{2k+1}) \in N(2\alpha)$, in other words,

$$\|\hat{X}\hat{s} - \mu e\| \le \alpha\mu. \tag{42}$$

Using (42) we obtain that $\mu(1-\alpha) \le \hat{x}_i\hat{s}_i \le \mu(1+\alpha)(i = 1, 2, \cdots, \hat{n})$, then $1/(\mu(1+\alpha)) \le 1/(\hat{x}_i\hat{s}_i) \le 1/(\mu(1-\alpha)), (i = 1, 2, \cdots, \hat{n})$, that is

$$\|(\hat{X}\hat{S})^{-\frac{1}{2}}\|^2 = \frac{1}{\|\hat{X}\hat{S}\|} \le \frac{1}{\mu(1-\alpha)}. \tag{43}$$

By combining (43) with (41) we get

$$\begin{aligned}\|\hat{D}_{xc}\hat{d}_{sc}\| &\le \frac{\|(\hat{X}\hat{S})^{-\frac{1}{2}}\|^2 \times \|\bar{\mu}e - \hat{X}\hat{s}\|^2 - 2\hat{d}_{sc}^T\hat{d}_{xc}}{2} \\ &\le \frac{\|\bar{\mu}e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc} \\ &= \frac{\|(\bar{\mu} - \mu)e + \mu e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc} \\ &= \frac{\|(\bar{\mu} - \mu)e\|^2 + \|\mu e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc} \\ &\le \frac{\alpha^2 + \tau^2}{2(1-\alpha)}\mu - \hat{d}_{sc}^T\hat{d}_{xc}.\end{aligned} \tag{44}$$

The second equality is due to the fact that $e^T(\mu e - \hat{X}\hat{s}) = 0$, and the last inequality follows from the definition of $\bar{\mu} = (1-\tau/\sqrt{\hat{n}})\mu$ and (42). Since $\hat{x}^+ = \hat{x}+\hat{d}_{xc}$, $\hat{s}^+ = \hat{s}+\hat{d}_{sc}$, then we get

$$\begin{aligned}\hat{X}^+\hat{s}^+ &= (\hat{X} + \hat{D}_{xc})(\hat{s} + \hat{d}_{sc}) \\ &= \hat{X}\hat{s} + \hat{D}_{xc}\hat{d}_{sc} + \hat{X}\hat{d}_{sc} + \hat{S}\hat{d}_{xc}, \\ &= \bar{\mu}e + \hat{D}_{xc}\hat{d}_{sc}.\end{aligned} \tag{45}$$

where the last equality is due to (38). Now it follows from (45) that

$$\begin{aligned}\mu^+ &= \frac{(\hat{x}^+)^T\hat{s}^+}{\hat{n}} = \frac{e^T(\bar{\mu}e + \hat{D}_{xc}\hat{d}_{sc})}{\hat{n}} \\ &= \frac{\hat{n}\bar{\mu} + \hat{d}_{xc}^T\hat{d}_{sc}}{\hat{n}} = \bar{\mu} + \frac{\hat{d}_{xc}^T\hat{d}_{sc}}{\hat{n}} \\ &\geq \bar{\mu}, \end{aligned} \tag{46}$$

where the last inequality is due to (29). Therefore

$$\begin{aligned}\left\|\frac{\hat{X}^+\hat{s}^+}{\mu^+} - e\right\| &= \left\|\frac{\bar{\mu}e + \hat{D}_{xc}\hat{d}_{sc}}{\mu^+} - e\right\| \\ &= \frac{\|\bar{\mu}e + \hat{D}_{xc}\hat{d}_{sc} - \mu^+ e\|}{\mu^+} \\ &= \frac{\|\bar{\mu}e + \hat{D}_{xc}\hat{d}_{sc} - \bar{\mu}e - (\hat{d}_{xc}^T\hat{d}_{sc}/\hat{n})e\|}{\mu^+} \\ &\leq \frac{\|\hat{D}_{xc}\hat{d}_{sc}\| + (\hat{d}_{xc}^T\hat{d}_{sc}/\hat{n})\|e\|}{\bar{\mu}} \\ &\leq \frac{\frac{\alpha^2+\tau^2}{2(1-\alpha)}\mu - \hat{d}_{sc}^T\hat{d}_{xc} + \hat{d}_{sc}^T\hat{d}_{xc}/\sqrt{\hat{n}}}{\bar{\mu}} \\ &\leq \frac{\alpha^2+\tau^2}{2(1-\alpha)\bar{\mu}}\mu \\ &= \frac{\alpha^2+\tau^2}{2(1-\alpha)(1-\tau/\sqrt{\hat{n}})} \\ &= \frac{\alpha}{2}\frac{\alpha^2+\tau^2}{\alpha(1-\alpha)(1-\tau/\sqrt{\hat{n}})} \leq \frac{\alpha}{2}, \end{aligned} \tag{47}$$

where the third equality is due to (46), the second inequality follows from (44), the third inequality is due to (29) and $\sqrt{\hat{n}} \geq 1$ and the last inequality follows from the first part of (24).

Similar to (44) we get

$$\begin{aligned}0 \leq \hat{d}_{xc}^T\hat{d}_{sc} &= (\hat{X}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}\hat{d}_{xc})^T(\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}) \\ &\leq \frac{\|\hat{X}^{-\frac{1}{2}}\hat{S}^{\frac{1}{2}}\hat{d}_{xc}\|^2 + \|\hat{X}^{\frac{1}{2}}\hat{S}^{-\frac{1}{2}}\hat{d}_{sc}\|^2}{2} \\ &\leq \frac{\|(\hat{X}\hat{S})^{-\frac{1}{2}}\|^2 \times \|\bar{\mu}e - \hat{X}\hat{s}\|^2 - 2\hat{d}_{sc}^T\hat{d}_{xc}}{2} \\ &\leq \frac{\|\bar{\mu}e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc} \\ &= \frac{\|(\bar{\mu}-\mu)e + \mu e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc} \\ &= \frac{\|(\bar{\mu}-\mu)e\|^2 + \|\mu e - \hat{X}\hat{s}\|^2}{2(1-\alpha)\mu} - \hat{d}_{sc}^T\hat{d}_{xc}\end{aligned}$$

$$\leq \frac{\alpha^2+\tau^2}{2(1-\alpha)}\mu - \hat{d}_{sc}^T \hat{d}_{xc}.$$

Hence we get

$$\hat{d}_{xc}^T \hat{d}_{sc} \leq \frac{\alpha^2+\tau^2}{4(1-\alpha)}\mu. \tag{48}$$

Follows from (46) we have

$$\begin{aligned}\mu^+ = \bar{\mu} + \frac{\hat{d}_{xc}^T \hat{d}_{sc}}{\hat{n}} &\leq \left(1-\frac{\tau}{\sqrt{\hat{n}}}\right)\mu + \frac{\alpha^2+\tau^2}{4\hat{n}(1-\alpha)}\mu \\ &= \left(1-\frac{\tau}{\sqrt{\hat{n}}}+\frac{\alpha^2+\tau^2}{4\hat{n}(1-\alpha)}\right)\mu = (1-\delta)\mu,\end{aligned}$$

where the first inequality is due to (48) and the last equality is follows from the definition of δ. The proof is completed.

The main results is stated as follows.

Theorem 3.1. *Let* $\left\{(x^k, z^k, y^k, s^k, w^k)\right\}$ *be generated by the algorithm, and let*

$$\delta = \frac{\tau}{\sqrt{2n}} - \frac{\alpha^2+\tau^2}{8n(1-\alpha)} > 0,$$

then there holds

(i) $\quad \theta^{2k} \geq \dfrac{\sqrt{\alpha^2+8n\alpha}-\alpha}{4n} = O\left(\dfrac{1}{\sqrt{n}}\right), \ \delta = O\left(\dfrac{1}{\sqrt{n}}\right).$

(ii) $\quad \mu^{2k+2} \leq \mu^0(1-\delta)^{k+1}\displaystyle\prod_{i=0}^{k}\left(1-\frac{\theta^{2k}}{2}\right).$

Thus the algorithm will terminate at an optimal primal-dual pair in $O(\sqrt{n}L)$ *iterations.*

Proof. Part (i) is due to Lemma 3.3. Part (ii) follows from Lemma 3.3 and Lemma 3.4. This ends the proof.

4. Final Remarks

In Mizuno, Todd and Ye's predictor-corrector method for linear programming [1], the corrector steps are used only to ensure that the iterates stay close to the central path so that large steps can be taken during the predictor steps. In fact, they choose $\bar{\mu}^{2k+1} = \mu^{2k+1}$. From (46) we know that if we also choose $\bar{\mu}^{2k+1} = \mu^{2k+1}$, then the duality gap remains unchanged at corrector step for linear programming, but increases for convex quadratic programming. In our modified algorithm, we choose $\bar{\mu}^{2k+1} = (1-\tau/\sqrt{\hat{n}})\mu^{2k+1}$, then the duality gap is reduced by a constant factor after each corrector step for linear and convex quadratic programming.

References

[1] S.Mizuno, M.J. Todd and Y.Ye, On adaptive primal-dual algorithm for linear programming, Technical Report No. 944, School of Operations Research and Industrial Engineering, Cornell University, (Ithaca, New York),1990, to appear in *Mathematics of Operations Research.*

[2] R.C. Monteiro and Adler,Interior path following primal-dual algorithms: Part II, Convex quadratic Programming," *Mathematical Programming*, Vol.44 (1989) 43–66.

[3] T.J.Carpenter, I.J.Lusting, J.M.Mulvey and D.F.Shanno, A primal-dual interior point method for convex separable nonlinear programs, Technical Report SOR-90-2, Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey.

[4] Z.Q.Luo and S.Wu, A modified predictor-corrector method for linear programming, Technical Report, Communications Research Laboratory, McMaster University, Hamilton, Ontario, L8S 4K1, Canada, 1993.

[5] F.Wu, S.Wu and Y.Ye, On quadratic convergence of the $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm, Working paper, The Institute of Applied Mathematics, Academia Sinica,1992.