SUT Journal of Mathematics
Vol. 39, No. 1 (2003), 85–104

Local convergence properties of primal-dual interior point methods based on the shifted barrier KKT conditions for nonlinear optimization

Hiroshi Yabe and Hiroshi Yamashita

(Received April 18, 2003; Revised June 8, 2003)

Abstract. In this paper, we consider the shifted barrier KKT conditions for nonlinear optimization. We propose a primal-dual interior point method based on these conditions. By choosing suitable parameters used in our method, we prove local and q-quadratic convergence of the Newton interior point method, and local and q-superlinear convergence of the quasi-Newton interior point method.

AMS 2000 *Mathematics Subject Classification.* 90C30, 90C51, 90C53.

Key words and phrases. Constrained optimization, primal-dual interior point method, local convergence property.

§1. Introduction

In this paper, we consider the following constrained optimization problem:

$$\text{(1.1)} \qquad \begin{array}{lll} \text{minimize} & f(x), & x \in \mathbf{R}^n \\ \text{subject to} & g(x) = 0, & h(x) \geq 0, \end{array}$$

where we assume that the functions $f : \mathbf{R}^n \to \mathbf{R}$, $g : \mathbf{R}^n \to \mathbf{R}^m$ and $h : \mathbf{R}^n \to \mathbf{R}^l$ are twice continuously differentiable. By introducing slack variables $s_i \geq 0,\ i = 1, \ldots, l$, problem (1.1) is written as:

$$\text{(1.2)} \qquad \begin{array}{lll} \text{minimize} & f(x), & x \in \mathbf{R}^n \\ \text{subject to} & g(x) = 0, & h(x) - s = 0,\ s \geq 0. \end{array}$$

Define the Lagrangian function of the above problem by

$$L(x, y, u, s, z) = f(x) - y^T g(x) - u^T (h(x) - s) - z^T s,$$

where $y \in \mathbf{R}^m, u \in \mathbf{R}^l$ are Lagrange multiplier vectors corresponding to the equality constraints, and $z \in \mathbf{R}^l$ is a Lagrange multiplier vector corresponding to the inequality constraint. Then Karush-Kuhn-Tucker (KKT) conditions for optimality of the above problem are given by

$$\begin{pmatrix} \nabla f(x) - A(x)^T y - B(x)^T u \\ g(x) \\ h(x) - s \\ u - z \\ SZe \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad s \geq 0, \quad z \geq 0,$$

where

$$\begin{aligned} A(x) &= (\nabla g_1(x), \ldots, \nabla g_m(x))^T, \\ B(x) &= (\nabla h_1(x), \ldots, \nabla h_l(x))^T, \\ S &= \operatorname{diag}(s_1, \ldots, s_l), \\ Z &= \operatorname{diag}(z_1, \ldots, z_l), \\ e &= (1, \ldots, 1)^T \in \mathbf{R}^l. \end{aligned}$$

Since the fourth equation of the above conditions implies $u = z$, the Lagrangian function can be rewritten as

$$L(w) = f(x) - y^T g(x) - z^T h(x)$$

and the KKT conditions reduce to

$$r_0(w) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ h(x) - s \\ SZe \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad s \geq 0, \quad z \geq 0,$$

where

$$w = (x, y, z, s)^T$$

and

$$\nabla_x L(w) = \nabla f(x) - A(x)^T y - B(x)^T z.$$

We note that the Jacobian matrix of $r_0(w)$ is represented by

$$r_0'(w) = \begin{pmatrix} \nabla_x^2 L(w) & -A(x)^T & -B(x)^T & 0 \\ A(x) & 0 & 0 & 0 \\ B(x) & 0 & 0 & -I \\ 0 & 0 & S & Z \end{pmatrix}. \tag{1.3}$$

To solve problem (1.2) by a primal-dual interior point method, some researchers have considered the barrier function minimization problem:

$$\begin{array}{ll} \text{minimize} & f(x) - \mu \sum_{i=1}^{l} \log s_i, \qquad (x,s) \in \mathbf{R}^n \times \mathbf{R}_+^l \\ \text{subject to} & g(x) = 0, \quad h(x) - s = 0, \end{array}$$

where $\mu > 0$ is a barrier parameter and

$$\mathbf{R}_+^l = \{ v \in \mathbf{R}^l \mid \ v_i > 0, \ i = 1, \ldots, l \}.$$

The first order necessary conditions for optimality of this minimization problem are given by the following equations:

$$\begin{pmatrix} \nabla f(x) - A(x)^T y - B(x)^T z \\ g(x) \\ h(x) - s \\ z - \mu S^{-1} e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad s \in \mathbf{R}_+^l.$$

By noting $z = \mu S^{-1} e$ (> 0), these equations are written as

$$(1.4) \qquad r_1(w;\mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ h(x) - s \\ SZe - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (s,z) \in \mathbf{R}_+^l \times \mathbf{R}_+^l.$$

We call these conditions as the barrier KKT conditions. When we apply the Newton method to the nonlinear equations, the Newton step $\Delta w = (\Delta x, \Delta y, \Delta z, \Delta s)^T$ is defined by a solution to the Newton equation

$$r_1'(w;\mu) \Delta w = -r_1(w;\mu),$$

where $r_1'(w;\mu)$ coincides with $r_0'(w)$ in (1.3).

To globalize Newton-like methods, Yamashita [13] introduced the following barrier penalty function as a merit function:

$$(1.5) \quad F_1(x,s;\mu,\sigma,\rho) = f(x) - \mu \sum_{i=1}^{l} \log s_i + \sigma \sum_{i=1}^{m} |g_i(x)| + \rho \sum_{i=1}^{l} |h_i(x) - s_i|,$$

where $\mu > 0$ is a barrier parameter, and $\sigma > 0$ and $\rho > 0$ are penalty parameters. The above function is called the l_1-type barrier penalty function. We should note that this function is nondifferentiable. Yamashita showed that if σ and ρ are sufficiently large, the necessary conditions for optimality of the l_1-type barrier penalty function minimization problem for a given $\mu > 0$ can be represented by the barrier KKT conditions (1.4). Convergence properties of primal-dual interior point methods based on (1.4) have been studied by many authors. Byrd, Liu and Nocedal [4], El-Bakry, Tapia, Tsuchiya and Zhang [7], Martinez, Parada and Tapia [11], Yabe and Yamashita [12], and Yamashita and Yabe [14] analyzed rate of convergence of these methods, for example. Global convergence properties were also studied by Byrd, Gilbert and Nocedal [2], Byrd, Hribar and Nocedal [3], El-Bakry, Tapia, Tsuchiya and Zhang [7], Yamashita [13], and Yamashita, Yabe and Tanabe [16], for example. See also Forsgren, Gill and Wright [10] as a comprehensive review of recent studies of interior point methods for nonlinear optimization.

In this paper, we consider the following differentiable barrier penalty function instead of (1.5):

$$(1.6)\quad F_2(x,s;\mu,\sigma,\rho) = f(x) - \mu\sum_{i=1}^{l}\log s_i + \frac{1}{2\sigma}\sum_{i=1}^{m}(g_i(x))^2 + \frac{1}{2\rho}\sum_{i=1}^{l}(h_i(x)-s_i)^2$$

which is extensively described in the book by Fiacco and McCormick [8]. We call this function the quadratic barrier penalty function. The necessary conditions for optimality of the minimization problem

$$\text{minimize}\quad F_2(x,s;\mu,\sigma,\rho), \quad (x,s)\in\mathbf{R}^n\times\mathbf{R}^l_+$$

are given by the following:

$$\nabla F_2 = \begin{pmatrix} \nabla f(x) + \frac{1}{\sigma}\sum_{i=1}^{m} g_i(x)\nabla g_i(x) + \frac{1}{\rho}\sum_{i=1}^{l}(h_i(x)-s_i)\nabla h_i(x) \\ -\mu S^{-1}e + \frac{1}{\rho}(s - h(x)) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and $s\in\mathbf{R}^l_+$. As in [8, 9, 15], we introduce the variables y and z by

$$y = -\frac{1}{\sigma}g(x) \quad \text{and} \quad z = -\frac{1}{\rho}(h(x)-s).$$

Since $\nabla_s F_2 = 0$ implies $z = \mu S^{-1}e$, the above conditions are written as

$$(1.7)\quad r_2(w;\mu,\sigma,\rho) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x)+\sigma y \\ h(x)-s+\rho z \\ SZe-\mu e \end{pmatrix} = \begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}, \quad (s,z)\in\mathbf{R}^l_+\times\mathbf{R}^l_+.$$

We call these conditions the shifted barrier KKT conditions. It should be noted that we treat x, y, z and s as independent variables. These conditions are also considered by Forsgren and Gill [9], and Yamashita and Yabe [15]. Based on these conditions, they proposed a differentiable primal-dual merit function in order to obtain global convergence properties.

We are interested in condition (1.7), because the parameters σ and ρ stabilize the Jacobian matrix $r_2'(w;\mu,\sigma,\rho)$ defined below. In fact, the regularity condition is necessary for the Jacobian matrix $r_1'(w;\mu)$ to be nonsingular at the solution, while the Jacobian matrix $r_2'(w;\mu,\sigma,\rho)$ becomes nonsingular at the solution by means of the existence of the fixed positive parameters σ and ρ even if the rank of $A(x)$ or $B(x)$ is deficient. This property is important in the global convergence analysis for fixed positive parameters μ, σ and ρ.

In this paper, we will analyze local behavior of primal-dual interior point methods based on (1.7) instead of (1.4). Yamashita and Yabe [15] showed q-superlinear convergence property of the method in the case where the iterates move along the central path near a solution. On the other hand, this paper shows the fast rate of convergence in the case where the iterates are in the neighborhood of a solution without considering central paths. Convergence results of this paper are closely related with those given by Yamashita and Yabe [14] for the barrier KKT conditions (1.4).

This paper is organized as follows. Section 2 will describe an algorithm of our method. In Section 3, we will present some useful lemmas in proving convergence properties. In Section 4, we will show local and q-quadratic convergence of the primal-dual interior point method based on the Newton method. In Sections 5, we will show local and q-superlinear convergence of the primal-dual interior point method based on the quasi-Newton method. Finally, Section 6 will give concluding remarks.

Throughout this paper, we call w satisfying $s > 0$ and $z > 0$ an interior point. The algorithm in this paper will generate such interior points. In what follows, the subscript k denotes an iteration count. Let $(w_k)_i$ be the ith element of the kth iterate w_k.

§2. Algorithm of primal-dual interior point methods

We consider the shifted barrier KKT conditions (1.7). Then the Jacobian matrix of r_2 is represented by

$$r_2'(w;\mu,\sigma,\rho) = \begin{pmatrix} \nabla_x^2 L(w) & -A(x)^T & -B(x)^T & 0 \\ A(x) & \sigma I & 0 & 0 \\ B(x) & 0 & \rho I & -I \\ 0 & 0 & S & Z \end{pmatrix}.$$

We note that

$$r_2(w;\mu,\sigma,\rho) = r_0(w) + \begin{pmatrix} 0 \\ \sigma y \\ \rho z \\ -\mu e \end{pmatrix} = r_0(w) - \mu\hat{e} + \sigma\hat{y} + \rho\hat{z},$$

where

$$\hat{e} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ e \end{pmatrix}, \quad \hat{y} = \begin{pmatrix} 0 \\ y \\ 0 \\ 0 \end{pmatrix}, \quad \hat{z} = \begin{pmatrix} 0 \\ 0 \\ z \\ 0 \end{pmatrix}.$$

Now we give an algorithm of our method as follows.

Algorithm IP

Given an initial point $w_0 = (x_0, y_0, z_0, s_0)$ with $s_0 > 0$ and $z_0 > 0$ and an initial matrix G_0, for $k = 0, 1, 2, \ldots$, do

(1) Choose the parameter $\mu_k > 0$, $\sigma_k > 0$, $\rho_k > 0$ and $\gamma_k \in (0, 1)$.

(2) Solve the following system for $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k, \Delta s_k)^T$:

$$J_k \Delta w_k = -r_2(w_k; \mu_k, \sigma_k, \rho_k), \tag{2.1}$$

where

$$J_k = \begin{pmatrix} G_k & -A(x_k)^T & -B(x_k)^T & 0 \\ A(x_k) & \sigma_k I & 0 & 0 \\ B(x_k) & 0 & \rho_k I & -I \\ 0 & 0 & S_k & Z_k \end{pmatrix} \tag{2.2}$$

and G_k is the Hessian matrix $\nabla_x^2 L(w_k)$ of the Lagrangian function or its approximation.

(3) Compute the step size

$$\begin{aligned} \alpha_k \equiv \min \Bigg\{ 1, & \gamma_k \min_i \left\{ -\frac{(s_k)_i}{(\Delta s_k)_i} \middle| (\Delta s_k)_i < 0 \right\}, \\ & \gamma_k \min_i \left\{ -\frac{(z_k)_i}{(\Delta z_k)_i} \middle| (\Delta z_k)_i < 0 \right\} \Bigg\}. \end{aligned} \tag{2.3}$$

(4) Update:

$$w_{k+1} = w_k + \alpha_k \Delta w_k.$$

If the matrix G_k is the true Hessian matrix $\nabla_x^2 L(w_k)$ of the Lagrangian function, then Algorithm IP becomes the primal-dual interior point method based on the Newton method, which is called the Newton interior point method. If the matrix G_k is an approximation to the Hessian matrix $\nabla_x^2 L(w_k)$, then Algorithm IP becomes the primal-dual interior point method based on quasi-Newton methods, which is called the quasi-Newton interior point method.

§3. Basic properties

In this section, we analyze the behavior of iteration vectors and step sizes given in Algorithm IP near a solution. Let $w^* = (x^*, y^*, z^*, s^*)^T$ be a KKT point, i.e., $r_0(w^*) = 0$, and let $I(x^*) = \{i \mid h_i(x^*) = 0\}$. We assume the following conditions:

(A1) The second derivatives of the functions f, g and h are Lipschitz continuous at x^*.
(A2) The point x^* satisfies the regularity condition, i.e., the vectors $\nabla g_i(x^*)$, $i = 1, \ldots, m$ and $\nabla h_i(x^*)$, $i \in I(x^*)$ are linearly independent.
(A3) The strict complementarity of w^* is satisfied, i.e., $(z^*)_i > 0$ for $i \in \{ i \mid (s^*)_i = 0\}$.
(A4) The second order sufficiency condition for optimality is satisfied at the point w^*, i.e., for all $v \neq 0$ satisfying $\nabla g_i(x^*)^T v = 0$, $i = 1, \ldots, m$ and $\nabla h_i(x^*)^T v = 0$, $i \in I(x^*)$, $v^T \nabla_x^2 L(w^*) v > 0$ holds.

Let $\|\cdot\|$ denote the l_2 norm for vectors and matrices, and let $\|\cdot\|_M$ and $\|\cdot\|_F$ be a matrix norm and the Frobenius norm for matrices, respectively. Then, by the norm equivalence, there is a positive constant η such that, for any matrix C,

$$\frac{1}{\eta}\|C\|_F \leq \|C\| \leq \eta\|C\|_F \quad \text{and} \quad \|C\|_F \leq \eta\|C\|_M.$$

Under assumption (A1), there exist a positive constant ξ and open convex sets $D_1(\subset \mathbf{R}^n)$ and $D(\subset \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^l \times \mathbf{R}^l)$ such that $x^* \in D_1$ and $w^* \in D$,

$$\|A(x) - A(x^*)\| \leq \xi\|x - x^*\| \quad \text{and} \quad \|B(x) - B(x^*)\| \leq \xi\|x - x^*\|$$

for $\forall x \in D_1$,

$$\|r_0(w) - r_0(w^*)\| \leq \xi\|w - w^*\| \quad \text{and} \quad \|\nabla r_0(w) - \nabla r_0(w^*)\| \leq \xi\|w - w^*\|$$

and

$$\|r_0(w) - r_0(\tilde{w}) - r_0'(w^*)(w - \tilde{w})\| \leq \frac{1}{2}\xi(\|w - w^*\| + \|\tilde{w} - w^*\|)\|w - \tilde{w}\|$$

for $\forall w, \tilde{w} \in D$. The last inequality is given by [6], for example.

In the subsequent sections, we will prove local convergence properties of primal-dual interior point methods that use Newton and quasi-Newton methods. For this purpose, we present some lemmas. The following lemma corresponds to Proposition 4.1 in [7] and guarantees the nonsingularity of the matrix $r_0'(w^*)$. This is an essential result for showing the fast rate of convergence of Newton-like methods.

Lemma 1. *Under assumptions* (A1)–(A4)*, the matrix* $r_0'(w^*)$ *is nonsingular.*

Proof. Though another proof was shown by El-Bakry et al. [7], we give a direct proof.

Let $v = (v_1, v_2, v_3, v_4)^T \in \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^l \times \mathbf{R}^l$. We will show that $r_0'(w^*)v = 0$ implies $v = 0$. Assume that $r_0'(w^*)v = 0$. The equations are represented by

$$\left\{ \begin{array}{c} \nabla_x^2 L(w^*)v_1 - A(x^*)^T v_2 - B(x^*)^T v_3 = 0 \\ A(x^*)v_1 = 0 \\ B(x^*)v_1 - v_4 = 0 \\ S^* v_3 + Z^* v_4 = 0. \end{array} \right. \tag{3.1}$$

With respect to active sets and inactive sets, we define $I^* = \{i \mid (s^*)_i = 0\}$ and $J^* = \{j \mid (s^*)_j > 0\}$, and we denote $s = \begin{pmatrix} s_{I^*} \\ s_{J^*} \end{pmatrix}, B(x^*) = \begin{pmatrix} B_{I^*} \\ B_{J^*} \end{pmatrix}$ without loss of generality. The fourth equation of (3.1) yields

$$(s^*)_i (v_3)_i + (z^*)_i (v_4)_i = 0, \quad i = 1, \ldots, l.$$

By using the strict complementarity condition, we have

$$(v_4)_i = 0, \ i \in I^* \quad \text{and} \quad (v_3)_j = 0, \ j \in J^*,$$

and we have $B_{I^*} v_1 = 0$ by the third equation of (3.1). Thus we have

$$\nabla_x^2 L(w^*)v_1 - A(x^*)^T v_2 - (B_{I^*}^T | B_{J^*}^T) \begin{pmatrix} (v_3)_{I^*} \\ 0 \end{pmatrix} = 0$$

and then

$$v_1^T \nabla_x^2 L(w^*)v_1 - (A(x^*)v_1)^T v_2 - (B_{I^*} v_1)^T (v_3)_{I^*} = 0,$$

which implies $v_1^T \nabla_x^2 L(w^*)v_1 = 0$, because $A(x^*)v_1 = 0$ and $B_{I^*} v_1 = 0$. Since assumption (A4) yields $v_1 = 0$, it follows from the first and third equations of (3.1) that $v_4 = 0$ and

$$A(x^*)^T v_2 + B(x^*)^T v_3 = \sum_{i=1}^m (v_2)_i \nabla g_i(x^*) + \sum_{i \in I^*} (v_3)_i \nabla h_i(x^*) = 0.$$

Furthermore, the regularity condition implies

$$v_2 = 0 \quad \text{and} \quad (v_3)_{I^*} = 0.$$

Therefore we obtain $v = 0$. □

We note that the Newton iteration for the modified complementarity condition yields

(3.2) $$S_k^{-1}\Delta s_k + Z_k^{-1}\Delta z_k = \mu_k (S_k Z_k)^{-1} e - e.$$

The following lemma is very helpful for the convergence analysis and is essentially the same lemma as Lemma 3 in [14].

Lemma 2. *Let assumption* (A3) *hold. Define*

$$\kappa \equiv 2\max\left\{\max_i\left\{\frac{1}{(s^*)_i}\,\middle|\,(s^*)_i > 0\right\}, \max_i\left\{\frac{1}{(z^*)_i}\,\middle|\,(z^*)_i > 0\right\}\right\}.$$

There exists a positive number ε_0 *such that, if*

$$\|w_k - w^*\| \le \varepsilon_0,$$

and if Δw_k *satisfies* (3.2), *then for each* i *such that* $(s^*)_i = 0$,

$$\frac{(\Delta s_k)_i}{(s_k)_i} = -1 + \frac{\mu_k}{(s_k)_i (z_k)_i} + (p_k)_i, \quad |(p_k)_i| \le \kappa \|\Delta w_k\|,$$

$$\left|\frac{(\Delta z_k)_i}{(z_k)_i}\right| \le \kappa \|\Delta w_k\|,$$

and for each i *such that* $(s^*)_i > 0$,

$$\left|\frac{(\Delta s_k)_i}{(s_k)_i}\right| \le \kappa \|\Delta w_k\|,$$

$$\frac{(\Delta z_k)_i}{(z_k)_i} = -1 + \frac{\mu_k}{(s_k)_i (z_k)_i} + (q_k)_i, \quad |(q_k)_i| \le \kappa \|\Delta w_k\|.$$

The next lemma corresponds to Lemma 4 in [14].

Lemma 3. *Let the assumptions of Lemma* 2 *hold. If*

$$\kappa \|\Delta w_k\| \le \gamma_k,$$

then

(3.3) $$1 \ge \alpha_k \ge \gamma_k - \kappa \|\Delta w_k\|.$$

The following lemma estimates the matrix J_k in (2.2) and the step size α_k in (2.3) near the point w^*.

Lemma 4. *Suppose that assumptions* (A1)–(A4) *hold and that the sequence* $\{w_k\}$ *is generated by Algorithm IP. Then there exist* $\varepsilon > 0$, $\delta > 0$, $\bar{\sigma} > 0$ *and* $\bar{\rho} > 0$ *such that, if* $\|w_k - w^*\| \le \varepsilon$, $\left\|G_k - \nabla_x^2 L(w^*)\right\|_M \le \delta$, $0 < \sigma_k \le \bar{\sigma}$ *and* $0 < \rho_k \le \bar{\rho}$, *then*

$$\|J_k - r_0'(w^*)\| \le \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \sigma_k^2 + \rho_k^2} \le \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \bar{\sigma}^2 + \bar{\rho}^2}, \tag{3.4}$$

and $\left\|J_k^{-1}\right\| \le \zeta$ *for some positive constant* ζ.

Furthermore, there exists $\bar{\mu} > 0$ *such that if, in addition,* $0 < \mu_k \le \bar{\mu}$, *then the following holds:*

$$0 \le 1 - \alpha_k \le (1 - \gamma_k) + O(\|r_0(w_k)\|) + O(\mu_k) + O(\sigma_k) + O(\rho_k), \tag{3.5}$$

provided that $0 < \bar{\gamma} \le \gamma_k < 1$ *where* $\bar{\gamma}$ *is a constant.*

Proof. Since

$$\begin{aligned} & J_k - r_0'(w^*) \\ = & \begin{pmatrix} G_k - \nabla_x^2 L(w^*) & A(x^*)^T - A(x_k)^T & B(x^*)^T - B(x_k)^T & 0 \\ A(x_k) - A(x^*) & \sigma_k I & 0 & 0 \\ B(x_k) - B(x^*) & 0 & \rho_k I & 0 \\ 0 & 0 & S_k - S^* & Z_k - Z^* \end{pmatrix}, \end{aligned}$$

we have

$$\begin{aligned} & \|J_k - r_0'(w^*)\|_F^2 \\ & \le \|G_k - \nabla_x^2 L(w^*)\|_F^2 + \sigma_k^2\|I\|_F^2 + \rho_k^2\|I\|_F^2 + \|r_0'(w_k) - r_0'(w^*)\|_F^2 \\ & \le \eta^2\|G_k - \nabla_x^2 L(w^*)\|_M^2 + \sigma_k^2\|I\|_F^2 + \rho_k^2\|I\|_F^2 + \eta^2\|r_0'(w_k) - r_0'(w^*)\|^2 \\ & \le \eta^2\delta^2 + \eta^2\sigma_k^2 + \eta^2\rho_k^2 + \eta^2\xi^2\|w_k - w^*\|^2 \\ & \le \eta^2(\delta^2 + \xi^2\varepsilon^2 + \bar{\sigma}^2 + \bar{\rho}^2). \end{aligned}$$

Thus

$$\|J_k - r_0'(w^*)\| \le \eta\|J_k - r_0'(w^*)\|_F \le \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \bar{\sigma}^2 + \bar{\rho}^2}.$$

This proves inequality (3.4).

By choosing ε, δ, $\bar{\sigma}$ and $\bar{\rho}$ such that

$$\|r_0'(w^*)^{-1}(J_k - r_0'(w^*))\| \le \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \bar{\sigma}^2 + \bar{\rho}^2}\ \|r_0'(w^*)^{-1}\| \le \frac{1}{2},$$

it follows from the Banach perturbation lemma that J_k is nonsingular and

$$\left\|J_k^{-1}\right\| \le \frac{\|r_0'(w^*)^{-1}\|}{1 - \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \bar{\sigma}^2 + \bar{\rho}^2}\,\|r_0'(w^*)^{-1}\|} \le \zeta \equiv 2\left\|r_0'(w^*)^{-1}\right\|.$$

Thus we have

$$\begin{aligned}
(3.6) \qquad \|\Delta w_k\| &= \left\| J_k^{-1} r_2(w_k; \mu_k, \sigma_k, \rho_k) \right\| \\
&\leq \left\| J_k^{-1} \right\| (\|r_0(w_k)\| + \mu_k \|e\| + \sigma_k \|y_k\| + \rho_k \|z_k\|)) \\
&\leq \zeta (\|r_0(w_k)\| + \mu_k \|e\| + \sigma_k \|w_k\| + \rho_k \|w_k\|)) .
\end{aligned}$$

To prove (3.5), we note that if ε, δ, $\bar{\mu}$, $\bar{\sigma}$ and $\bar{\rho}$ are sufficiently small, then from the conditions for the parameters and (3.6), the assumption of Lemma 3 is satisfied. It follows from (3.3) that

$$\begin{aligned}
0 &\leq 1 - \alpha_k \\
&\leq (1 - \gamma_k) + \kappa \|\Delta w_k\| \\
&\leq (1 - \gamma_k) + \kappa \zeta (\|r_0(w_k)\| + \mu_k \|e\| + \sigma_k \|w_k\| + \rho_k \|w_k\|).
\end{aligned}$$

Therefore, from the boundedness of $\{w_k\}$, we obtain (3.5). □

§4. Local and quadratic convergence of the Newton interior point method

In this section, we pay our attention to the local and quadratic convergence property of the Newton interior point method. Letting $G_k = \nabla_x^2 L(w_k)$ in (2.2) of Algorithm IP in Section 2, we have $J_k = r_2'(w_k; \mu_k, \sigma_k, \rho_k)$.

Theorem 1. *Suppose that assumptions* (A1)–(A4) *hold. Let* $G_k = \nabla_x^2 L(w_k)$. *Let* $\{w_k\}$ *be generated by Algorithm IP. Choose the parameters such that*

$$\begin{aligned}
&0 < \mu_k = O(\|r_0(w_k)\|^2), \qquad 0 < \sigma_k = O(\|r_0(w_k)\|^2), \\
&0 < \rho_k = O(\|r_0(w_k)\|^2) \quad \textit{and} \quad 0 < 1 - \gamma_k = O(\|r_0(w_k)\|).
\end{aligned}$$

Then there exists a positive constant ε *such that for*

$$\|w_0 - w^*\| < \varepsilon, \qquad w_0 \in D,$$

then the sequence $\{w_k\}$ *is well defined and converges q-quadratically to* w^*.

Proof. Assume that

$$\|w_k - w^*\| < \varepsilon$$

for ε sufficiently small. Since, by Lemma 4, $r_2'(w_k; \mu_k, \sigma_k, \rho_k)$ is nonsingular and

$$\|r_2'(w_k; \mu_k, \sigma_k, \rho_k)^{-1}\| \leq \zeta,$$

we have

$$
\begin{aligned}
w_{k+1} - w^* &= (w_k - w^*) + \alpha_k \Delta w_k \\
&= (1-\alpha_k)(w_k - w^*) - \alpha_k r_2'(w_k; \mu_k, \sigma_k, \rho_k)^{-1}\{r_0(w_k) - r_0(w^*) \\
&\quad - r_0'(w_k)(w_k - w^*) - \begin{pmatrix} 0 & & & 0 \\ & \sigma_k I & & \\ & & \rho_k I & \\ 0 & & & 0 \end{pmatrix}(w_k - w^*) \\
&\quad - \mu_k \hat{e} + \sigma_k \hat{y}_k + \rho_k \hat{z}_k\},
\end{aligned}
$$

and hence it follows from Lemma 4 that, for ε sufficiently small,

$$
\begin{aligned}
&\|w_{k+1} - w^*\| \\
&\quad \le (1-\alpha_k)\|w_k - w^*\| \\
&\qquad + \alpha_k \|r_2'(w_k; \mu_k, \sigma_k, \rho_k)^{-1}\|(\|r_0(w_k) - r_0(w^*) - r_0'(w^*)(w_k - w^*)\| \\
&\qquad + \|(r_0'(w_k) - r_0'(w^*))(w_k - w^*)\| + (\sigma_k + \rho_k)\|w_k - w^*\| \\
&\qquad + \mu_k\|e\| + \sigma_k\|y_k\| + \rho_k\|z_k\|) \\
&\quad \le \{(1-\gamma_k) + O(\|r_0(w_k)\|) + O(\mu_k) + O(\sigma_k) + O(\rho_k)\}\|w_k - w^*\| \\
&\qquad + O(\|w_k - w^*\|^2) + O(\mu_k) + O(\sigma_k) + O(\rho_k).
\end{aligned}
$$

In the last inequality, the boundedness of $\{y_k\}$ and $\{z_k\}$ are used. Thus there exists a positive constant ν such that

$$\|w_{k+1} - w^*\| \le \nu \|w_k - w^*\|^2 \le \nu \varepsilon^2 < \varepsilon.$$

Thus, by using mathematical induction, it is easy to show that the sequence $\{w_k\}$ converges to w^* and the rate of convergence is quadratic. Therefore the proof is complete. □

§5. Local and superlinear convergence of the quasi-Newton interior point method

By letting the matrix G_k be an approximation to the Hessian matrix $\nabla_x^2 L(w_k)$, Algorithm IP given in Section 2 can be regarded as the quasi-Newton method. In this section, we show the local and superlinear convergence property of the quasi-Newton interior point method. The next theorem gives local and linear convergence of the quasi-Newton method. This theorem corresponds to the bounded deterioration theorem for unconstrained optimization by Broyden, Dennis, and Moré [1].

Theorem 2. *Let $\{w_k\}$ be generated by Algorithm IP. Suppose that assumptions* (A1)–(A4) *hold. Choose the parameters such that*

$$\begin{array}{l} 0 < \mu_k = O(\|r_0(w_k)\|^{1+\tau}), \quad 0 < \sigma_k = O(\|r_0(w_k)\|^{1+\tau}) \\ 0 < \rho_k = O(\|r_0(w_k)\|^{1+\tau}) \quad \text{and} \quad 0 < \hat{\gamma} \le \gamma_k < 1 \end{array}$$

for constants $\tau > 0$ and $\hat{\gamma} \in (0,1)$. Assume that the sequence of matrices $\{G_k\}$ satisfies the bounded deterioration property

$$\|G_{k+1} - \nabla_x^2 L(w^*)\|_M \le (1 + \beta_1 \psi_k)\|G_k - \nabla_x^2 L(w^*)\|_M + \beta_2 \psi_k,$$

where β_1 and β_2 are positive constants, and

$$\psi_k = \max(\|w_{k+1} - w^*\|, \|w_k - w^*\|).$$

Then for each $\nu \in (1 - \hat{\gamma}, 1)$, there exist positive constants $\varepsilon = \varepsilon(\nu)$ and $\delta = \delta(\nu)$ such that if

$$\|w_0 - w^*\| < \varepsilon, \qquad w_0 \in D$$

and

$$\|G_0 - \nabla_x^2 L(w^*)\|_M < \frac{\delta}{2},$$

then the sequence $\{w_k\}$ is well defined and converges to w^*. Furthermore,

$$\|w_{k+1} - w^*\| \le \nu \|w_k - w^*\|$$

for each $k \ge 0$.

Proof. By induction on k, we will prove that

$$\|w_{k+1} - w^*\| \le \nu\|w_k - w^*\| < \varepsilon \quad \text{and} \quad \|G_{k+1} - \nabla_x^2 L(w^*)\|_M < \delta$$

for all $k \ge 0$. For this purpose, we show that if, for $i = 0, 1, \ldots, k$,

$$\|w_i - w^*\| \le \nu\|w_{i-1} - w^*\| < \varepsilon \quad \text{and} \quad \|G_i - \nabla_x^2 L(w^*)\|_M < \delta,$$

then

$$\|w_{k+1} - w^*\| \le \nu\|w_k - w^*\| < \varepsilon \quad \text{and} \quad \|G_{k+1} - \nabla_x^2 L(w^*)\|_M < \delta.$$

If ε and δ are sufficiently small, it follows from Lemma 4 that J_k is nonsingular and $\|J_k^{-1}\| \le \zeta$. ¿From the linear system (2.1), we have

$$\begin{aligned} w_{k+1} - w^* &= (w_k - w^*) + \alpha_k \Delta w_k \\ &= (1 - \alpha_k)(w_k - w^*) - \alpha_k J_k^{-1}\{r_0(w_k) - r_0(w^*) \\ &\qquad -r_0'(w^*)(w_k - w^*) - (J_k - r_0'(w^*))(w_k - w^*) \\ &\qquad -\mu_k \hat{e} + \sigma_k \hat{y}_k + \rho_k \hat{z}_k\} \end{aligned}$$

and hence, for some ζ',

$$\begin{aligned}
\|w_{k+1} - w^*\| &\le (1-\alpha_k)\|w_k - w^*\| \\
&\quad +\alpha_k\|J_k^{-1}\|\{\|r_0(w_k) - r_0(w^*) - r_0'(w^*)(w_k - w^*)\| \\
&\quad +\|J_k - r_0'(w^*)\|\|w_k - w^*\| + \mu_k\|e\| + \sigma_k\|y_k\| + \rho_k\|z_k\|\} \\
&\le \{(1-\gamma_k) + O(\|r_0(w_k)\|)\}\|w_k - w^*\| \\
&\quad +\zeta\{O(\|w_k - w^*\|^2) + O(\mu_k) + O(\sigma_k) \\
&\quad +O(\rho_k) + \eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + \sigma_k^2 + \rho_k^2}\ \|w_k - w^*\|\} \\
&\le \{(1-\gamma_k) + O(\|w_k - w^*\|) + O(\|w_k - w^*\|^\tau) \\
&\quad +\zeta\eta^2\sqrt{\delta^2 + \xi^2\varepsilon^2 + O(\|w_k - w^*\|^{2(1+\tau)})}\}\|w_k - w^*\| \\
&\le \{(1-\hat{\gamma}) + \zeta'(\varepsilon^{\min(1,\tau)} + \sqrt{\delta^2+\varepsilon^2}\ )\}\|w_k - w^*\|.
\end{aligned}$$

Choosing ε and δ such that

$$(1-\hat{\gamma}) + \zeta'(\varepsilon^{\min(1,\tau)} + \sqrt{\delta^2+\varepsilon^2}\ ) < \nu,$$

we obtain

$$\|w_{k+1} - w^*\| \le \nu\|w_k - w^*\| < \varepsilon.$$

Moreover, by using the same technique as in Broyden, Dennis and Moré [1], we can show that

$$\|G_{k+1} - \nabla_x^2 L(w^*)\|_M < \delta.$$

We can prove the case of $k = 0$ in the same way as above. Therefore the proof is complete. □

Now we give necessary and sufficient conditions for superlinear convergence of our method.

Theorem 3. *Suppose that assumptions* (A1)–(A4) *hold and that the sequence* $\{w_k\}$ *generated by Algorithm IP converges linearly to* w^*. *Choose the parameters such that*

$$0 < \mu_k = o(\|r_0(w_k)\|), \quad 0 < \sigma_k = o(\|r_0(w_k)\|),$$

$$0 < \rho_k = o(\|r_0(w_k)\|) \quad \textit{and} \quad 0 < 1 - \gamma_k = o(1).$$

Then the following four conditions are equivalent.
(a) *The sequence* $\{G_k\}$ *satisfies*

$$\lim_{k\to\infty} \frac{\|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\|}{\|w_{k+1} - w_k\|} = 0. \tag{5.1}$$

(b) *The sequence* $\{J_k\}$ *satisfies*

$$\lim_{k\to\infty} \frac{\|(J_k - r_0'(w^*))(w_{k+1} - w_k)\|}{\|w_{k+1} - w_k\|} = 0. \tag{5.2}$$

(c) *The sequence* $\{r_0(w_k)\}$ *satisfies*

$$\lim_{k\to\infty} \frac{\|r_0(w_{k+1})\|}{\|w_{k+1} - w_k\|} = 0. \tag{5.3}$$

(d) *The sequence* $\{w_k\}$ *converges superlinearly to* w^*, *i.e.*,

$$\lim_{k\to\infty} \frac{\|w_{k+1} - w^*\|}{\|w_k - w^*\|} = 0.$$

Proof. First we note that linear convergence implies, for some $\nu \in (0,1)$,

$$\begin{aligned} \|w_k - w^*\| &\le \|w_{k+1} - w^*\| + \|w_{k+1} - w_k\| \\ &\le \nu\|w_k - w^*\| + \|w_{k+1} - w_k\|, \end{aligned}$$

so we have

$$\frac{\|w_k - w^*\|}{\|w_{k+1} - w_k\|} \le \frac{1}{1-\nu} . \tag{5.4}$$

(a) $\Longrightarrow$ (b): Since

$$\begin{aligned} &(J_k - r_0'(w^*))(w_{k+1} - w_k) \\ = &\begin{pmatrix} G_k - \nabla_x^2 L(w^*) & A(x^*)^T - A(x_k)^T & B(x^*)^T - B(x_k)^T & 0 \\ A(x_k) - A(x^*) & \sigma_k I & 0 & 0 \\ B(x_k) - B(x^*) & 0 & \rho_k I & 0 \\ 0 & 0 & S_k - S^* & Z_k - Z^* \end{pmatrix} \bullet \\ &(w_{k+1} - w_k) \\ = &\begin{pmatrix} (G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k) \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &+ \begin{pmatrix} (A(x^*) - A(x_k))^T(y_{k+1} - y_k) \\ (A(x_k) - A(x^*))(x_{k+1} - x_k) \\ (B(x_k) - B(x^*))(x_{k+1} - x_k) \\ (Z_k - Z^*)(s_{k+1} - s_k) \end{pmatrix} + \begin{pmatrix} (B(x^*) - B(x_k))^T(z_{k+1} - z_k) \\ \sigma_k(y_{k+1} - y_k) \\ \rho_k(z_{k+1} - z_k) \\ (S_k - S^*)(z_{k+1} - z_k) \end{pmatrix}, \end{aligned}$$

we have

$$\|(J_k - r_0'(w^*))(w_{k+1} - w_k)\|$$

$$\begin{aligned}
&\leq \|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\| \\
&\quad + \left\| \begin{pmatrix} (A(x^*) - A(x_k))^T (y_{k+1} - y_k) \\ (A(x_k) - A(x^*))(x_{k+1} - x_k) \\ (B(x_k) - B(x^*))(x_{k+1} - x_k) \\ (Z_k - Z^*)(s_{k+1} - s_k) \end{pmatrix} \right\| \\
&\qquad + \left\| \begin{pmatrix} (B(x^*) - B(x_k))^T (z_{k+1} - z_k) \\ \sigma_k (y_{k+1} - y_k) \\ \rho_k (z_{k+1} - z_k) \\ (S_k - S^*)(z_{k+1} - z_k) \end{pmatrix} \right\| \\
&\leq \|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\| + O(\|w_k - w^*\| \|w_{k+1} - w_k\|).
\end{aligned}$$

Thus the following holds

$$\begin{aligned}
&\lim_{k\to\infty} \frac{\|(J_k - r_0'(w^*))(w_{k+1} - w_k)\|}{\|w_{k+1} - w_k\|} \\
&\qquad \leq \lim_{k\to\infty} \frac{\|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\|}{\|w_{k+1} - w_k\|} = 0,
\end{aligned}$$

which implies (b).

(b) $\Longrightarrow$ (a): Since

$$\begin{aligned}
& \|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\| \\
\leq\ & \|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k) + (-A(x_k) + A(x^*))^T (y_{k+1} - y_k) \\
& + (-B(x_k) + B(x^*))^T (z_{k+1} - z_k)\| \\
& + \|(-A(x_k) + A(x^*))^T (y_{k+1} - y_k) + (-B(x_k) + B(x^*))^T (z_{k+1} - z_k)\| \\
\leq\ & \|(J_k - r_0'(w^*))(w_{k+1} - w_k)\| + \|A(x_k) - A(x^*)\| \|y_{k+1} - y_k\| \\
& + \|B(x_k) - B(x^*)\| \|z_{k+1} - z_k\|,
\end{aligned}$$

we have

$$\begin{aligned}
& \frac{\|(G_k - \nabla_x^2 L(w^*))(x_{k+1} - x_k)\|}{\|w_{k+1} - w_k\|} \\
&\quad \leq \frac{\|(J_k - r_0'(w^*))(w_{k+1} - w_k)\|}{\|w_{k+1} - w_k\|} + \|A(x_k) - A(x^*)\| \frac{\|y_{k+1} - y_k\|}{\|w_{k+1} - w_k\|} \\
&\qquad + \|B(x_k) - B(x^*)\| \frac{\|z_{k+1} - z_k\|}{\|w_{k+1} - w_k\|}.
\end{aligned}$$

Thus (b) implies (a).

(b) $\Longrightarrow$ (c): Since

$$\begin{aligned}
r_0(w_{k+1}) &= r_0(w_{k+1}) - J_k(w_{k+1} - w_k) - \alpha_k (r_0(w_k) - \mu_k \hat{e} + \sigma_k \hat{y}_k + \rho_k \hat{z}_k) \\
&= r_0(w_{k+1}) - r_0(w_k) - r_0'(w^*)(w_{k+1} - w_k) \\
&\qquad - (J_k - r_0'(w^*))(w_{k+1} - w_k) + (1 - \alpha_k)(r_0(w_k) - r_0(w^*)) \\
&\qquad - \alpha_k (-\mu_k \hat{e} + \sigma_k \hat{y}_k + \rho_k \hat{z}_k),
\end{aligned} \tag{5.5}$$

we have

$$\begin{aligned}\|r_0(w_{k+1})\| &\leq \|r_0(w_{k+1}) - r_0(w_k) - r_0'(w^*)(w_{k+1} - w_k)\| \\ &\quad +\|(J_k - r_0'(w^*))(w_{k+1} - w_k)\| + (1 - \alpha_k)\|r_0(w_k) - r_0(w^*)\| \\ &\quad +\alpha_k(\mu_k\|e\| + \sigma_k\|y_k\| + \rho_k\|z_k\|) \\ &= O(\|w_k - w^*\|)\|w_{k+1} - w_k\| + \|(J_k - r_0'(w^*))(w_{k+1} - w_k)\| \\ &\quad +\{(1 - \gamma_k) + O(\|r_0(w_k)\|)\}O(\|w_k - w^*\|) \\ &\quad +O(\mu_k) + O(\sigma_k) + O(\rho_k) \\ &= O(\|w_k - w^*\|)\|w_{k+1} - w_k\| + \|(J_k - r_0'(w^*))(w_{k+1} - w_k)\| \\ &\quad +o(1)O(\|w_k - w^*\|) + o(\|w_k - w^*\|).\end{aligned}$$

Therefore the above and expression (5.4) yield (c).
(c) $\Longrightarrow$ (b): Since it follows directly from (5.5) that

$$\begin{aligned}&(J_k - r_0'(w^*))(w_{k+1} - w_k) \\ = \ & r_0(w_{k+1}) - r_0(w_k) - r_0'(w^*)(w_{k+1} - w_k) + (1 - \alpha_k)(r_0(w_k) - r_0(w^*)) \\ &\quad -\alpha_k(-\mu_k\hat{e} + \sigma_k\hat{y}_k + \rho_k\hat{z}_k) - r_0(w_{k+1}),\end{aligned}$$

we can obtain (b) in the same way as above.
(c) $\Longleftrightarrow$ (d): The result follows directly from the same argument as in Dennis and Moré [5].

Therefore the theorem is proved. □

Note that (5.1) or (5.2) corresponds to the Dennis-Moré condition [5] in the case of unconstrained optimization. We also note that condition (5.3) is observable. Thus by observing the sequence $\{\|r_0(w_{k+1})\|/\|w_{k+1} - w_k\|\}$, we can investigate whether the sequence $\{w_k\}$ converges q-superlinearly to a KKT point.

§6. Concluding remarks

In this paper, we have considered the shifted barrier KKT conditions (1.7) that arise from minimizing (1.6), and we have proposed primal-dual interior point methods based on the Newton method and the quasi-Newton method. The shifted barrier KKT conditions are interesting, because the parameters σ and ρ stabilize the Jacobian matrix $r_2'(w;\mu,\sigma,\rho)$ even if the rank of $A(x)$ or $B(x)$ is deficient. Under standard assumptions, we have proved local and quadratic convergence of the Newton interior point method, and local and q-superlinear convergence of the quasi-Newton interior point method. These are closely related with convergence results by Yamashita and Yabe [14] for the barrier KKT conditions.

In [14], they dealt with three kinds of step size rules that include the following two rules in addition to (2.3):

Step size rule A

$$\alpha_{sk} = \min\left\{1, \gamma_k \min_i \left\{ -\frac{(s_k)_i}{(\Delta s_k)_i} \middle| (\Delta s_k)_i < 0 \right\}\right\},$$

and

$$\alpha_{zk} = \min\left\{1, \gamma_k \min_i \left\{ -\frac{(z_k)_i}{(\Delta z_k)_i} \middle| (\Delta z_k)_i < 0 \right\}\right\},$$

where $\gamma_k \in (0,1)$. Step sizes for the other variables are chosen as 1, or α_{sk}, or α_{zk}.

Step size rule B

$$\alpha_{sk} = \min\left\{1, \gamma_k \min_i \left\{ -\frac{(s_k)_i}{(\Delta s_k)_i} \middle| (\Delta s_k)_i < 0 \right\}\right\},$$

where $\gamma_k \in (0,1)$. The step size α_{zk} is the largest step that satisfies

$$\alpha_{zk} \leq 1,$$

$$\begin{aligned} \min\left\{\frac{\mu_k}{M_{Lk}((s_k)_i + \alpha_{sk}(\Delta s_k)_i)}, (z_k)_i\right\} &\leq (z_k)_i + \alpha_{zk}(\Delta z_k)_i \\ &\leq \max\left\{\frac{M_{Uk}\mu_k}{(s_k)_i + \alpha_{sk}(\Delta s'_k)_i}, (z_k)_i\right\} \end{aligned}$$

for $i = 1, \ldots, n$, where $\mu_k > 0$, and where M_{Lk} and M_{Uk} are positive numbers that satisfy

$$M_{Lk} > \max\left\{1, \frac{2\mu_k}{(1-\gamma_k)\min_i \{(s_k)_i (z_k)_i\}}\right\}$$

and

$$M_{Uk} > \max\left\{3, \frac{3\max_i \{(s_k)_i (z_k)_i\}}{\mu_k}\right\}.$$

Step sizes of the other variables are chosen as 1, or α_{sk}, or α_{zk}.

For Algorithm IP with step size rule A or B, similar convergence results to Theorems 1, 2 and 3 of the present paper can be obtained.

Acknowledgments. The authors appreciate the valuable comments from an anonymous referee.

References

[1] C.G. Broyden, J.E. Dennis, Jr. and J.J. Moré, On the local and superlinear convergence of quasi-Newton methods, *Journal of Institute of Mathematics and its Applications*, **12** (1973), 223–245.

[2] R.H. Byrd, J.C. Gilbert and J. Nocedal, A trust region method based on interior point techniques for nonlinear programming, *Mathematical Programming*, **89** (2000), 149–185.

[3] R.H. Byrd, M.E. Hribar and J. Nocedal, An interior point algorithm for large-scale nonlinear programming, *SIAM J. on Optimization*, **9** (1999), 877–900.

[4] R.H. Byrd, G. Liu and J. Nocedal, On the local behaviour of an interior point method for nonlinear programming, in *Numerical analysis 1997*, D.F. Griffiths, D.J. Higham and G.A. Watson eds., Longman, 37–56, 1998.

[5] J.E. Dennis, Jr. and J.J. Moré, A characterization of superlinear convergence and its application to quasi-Newton methods, *Mathematics of Computation*, **28** (1974), 549–560.

[6] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, New Jersey, 1983.

[7] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya and Y. Zhang, On the formulation and theory of the Newton interior-point method for nonlinear programming, *Journal of Optimization Theory and Applications*, **89** (1996), 507–541.

[8] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, 1990 (Reprint of the 1968 original).

[9] A. Forsgren and P.E. Gill, Primal-dual interior methods for nonconvex nonlinear programming, *SIAM J. on Optimization*, **8** (1998), 1132–1152.

[10] A. Forsgren, P.E. Gill and M.H. Wright, Interior methods for nonlinear optimization, *SIAM Review*, **44** (2002), 525–597.

[11] H.J. Martinez, Z. Parada and R.A. Tapia, On the characterization of Q-superlinear convergence of quasi-Newton interior-point methods for nonlinear programming, *Boletín de la Sociedad Matemática Mexicana*, **1** (1995), 137–148.

[12] H. Yabe and H. Yamashita, Q-superlinear convergence of primal-dual interior point quasi-Newton methods for constrained optimization, *Journal of the Operations Research Society of Japan*, **40** (1997), 415–436.

[13] H. Yamashita, A globally convergent primal-dual interior point method for constrained optimization, *Optimization Methods and Software*, **10** (1998), 443–469.

[14] H. Yamashita and H. Yabe, Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization, *Mathematical Programming*, **75** (1996), 377–397.

[15] H. Yamashita and H. Yabe, An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization, *SIAM J. on Optimization* (accepted).

[16] H. Yamashita, H. Yabe and T. Tanabe, *A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization*, Technical Report, July 1997 (revised May 2003).

Hiroshi Yabe
Department of Mathematical Information Science, Tokyo University of Science
1-3, Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
yabe@rs.kagu.tus.ac.jp

Hiroshi Yamashita
Mathematical Systems, Inc.
2-4-3, Shinjuku, Shinjuku-ku, Tokyo 160-0022, Japan
hy@msi.co.jp