Digital Object Identifier (DOI) 10.1007/s101079900098

Thomas F. Coleman · Jianguo Liu

An interior Newton method for quadratic programming*

Received October 11, 1993 / Revised version received February 20, 1996
Published online July 19, 1999

Abstract. We propose a new (interior) approach for the general quadratic programming problem. We establish that the new method has strong convergence properties: the generated sequence converges globally to a point satisfying the second-order necessary optimality conditions, and the rate of convergence is 2-step quadratic if the limit point is a strong local minimizer. Published alternative interior approaches do not share such strong convergence properties for the nonconvex case. We also report on the results of preliminary numerical experiments: the results indicate that the proposed method has considerable practical potential.

Key words. nonconvex quadratic programming – interior method – Newton method – trust-region method – dogleg method – quadratic convergence

1. Introduction

Consider a quadratic program in the standard form:

$$\min_x \left\{ q(x) = \frac{1}{2} x^T Hx + c^T x \right\} \qquad (1)$$
$$\text{subject to } Ax = b \text{ and } x \geq 0,$$

where $H \in \Re^{n\times n}$ is symmetric and, in general, indefinite; $c \in \Re^n$, $A \in \Re^{m\times n}$, and $b \in \Re^m$. When H is indefinite, we are interested in locating a local minimizer and call such a minimizer a solution to (1).

Problem (1) has been studied intensively due to its importance in optimization: many real world problems are posed in the form (1). In addition, many algorithms for general nonlinear programming require successively solving subproblems of this form (e.g., "SQP" methods).

There are basically two kinds of existing approaches for solving (1). The most popular strategies follow an active-set or gradient projection philosophy: a piecewise

T.F. Coleman: Computer Science Department and Center for Applied Mathematics, Cornell University, Ithaca, NY 14853, USA

J. Liu: Center for Applied Mathematics, Cornell University, Ithaca, NY 14853, USA
Current address: Department of Mathematics, University of North Texas, Denton, TX 75067, USA

Mathematics Subject Classification (1991): 65K05, 90C20, 90C06

* Research partially supported by the Applied Mathematical Sciences Research Program (KC-04-02) of the Office of Energy Research of the U.S. Department of Energy under grant DE-FG02-86ER25013.A000, and by NSF, AFOSR, and ONR through grant DMS-8920550, and by the Advanced Computing Research Institute, a unit of the Cornell Theory Center which receives major funding from the National Science Foundation and IBM Corporation, with additional support from New York State and members of its Corporate Research Institute.

linear path is generated, following faces of the polytope defined by the constraints in (1), e.g., [3], [5], [6], [11], [14], [16], [20], [29]. These methods usually generate a finite sequence of intermediate "approximations". An alternative philosophy is to generate an infinite sequence of strictly feasible (or interior) points, converging in the limit to a local solution. One such method, the affine-scaling method, was proposed over twenty years ago [13]. Following the publication of Karmarkar's interior-point method [23], there has been a resurgence of interest in such methods for linear programming, linear complementarity problem, and to some degree quadratic programming, especially the convex case: see for example, [1], [2], [8], [21], [22], [24], [26], [27], [30], [31], [35], [36], [38], and [39].

Most recently, Ye [37] proposes an interior algorithm for problem (1). Ye's method uses the solution of a projected trust region subproblem as the search direction in each iteration and generates a sequence of points converging, in the limit, to a point satisfying the second-order necessary conditions. The convergence rate of Ye's algorithm is believed to be linear. (For references on trust region methods, see for example, [15], [28], [33], and [34].)

We propose a new interior Newton method for problem (1). In our algorithm, each step involves two directions: the solution to a projected trust region subproblem and a projected steepest descent direction. We show that our new method has strong convergence properties: the generated sequence converges globally to a point satisfying the second-order necessary conditions, and the rate of convergence is 2-step quadratic if the limit point is a strong local minimizer. Published alternative interior approaches do not share such strong convergence properties for the nonconvex case. As a byproduct, we show that for general QP, the projected steepest descent directions yield a procedure converging globally to a point satisfying the first-order necessary conditions (KKT conditions). Finally, we describe a modification of this algorithm using a new scaling strategy which dramatically improves computational performance while maintaining strong convergence properties.

Another feature of our method is that the underlying linear system in the limit is well-conditioned, in contrast to certain formulations of alternative interior methods. Our preliminary numerical experiments indicate that the new algorithm is promising. A related work for box-constrained minimization problems is given in [8].

This paper is organized as follows. In Sect. 2, we motivate and present the basic algorithm. In Sects. 3 and 4, we discuss its strong convergence properties. In Sect. 5, we describe modifications of the basic method, and discuss results of preliminary numerical experiments. (In the Appendix we show that an efficient modified version of the basic algorithm maintains the theoretical convergence properties of the basic method.) Finally, in Sect. 6, we have concluding remarks and observations.

We use superscripts to denote the iteration counts and use subscripts to indicate the indices of vector components. We occasionally drop the superscripts when there is no confusion. We use ":=" to denote the phrase "is defined to be". The norm $\|\cdot\|$ used in this paper is the l_2 norm unless otherwise specified. Sets will be denoted by calligraphic capital letters. Given a vector $x \in \Re^n$, the notation $x > 0$ means $x_i > 0$ for every $1 \leq i \leq n$ and $x \geq 0$ means $x_i \geq 0$ for every $1 \leq i \leq n$. We call x *a feasible point* if $Ax = b$ and $x \geq 0$. We call x *an interior point* if x is feasible and $x > 0$, and we call x *a boundary point* if x is feasible and $x_i = 0$ for some $1 \leq i \leq n$. When $M \in \Re^{n \times n}$ is a square matrix, the notation $M > 0$ indicates that M is positive definite and the notation $M \geq 0$ indicates that M is positive semidefinite. If x denotes a vector, $X = diag(x)$ will denote the diagonal matrix whose entries are the components of x. Finally, if $x = [x_1, x_2, \cdots, x_n]^T$ and $M = (m_{ij}) \in \Re^{n \times n}$, then $|x| = [|x_1|, |x_2|, \cdots, |x_n|]^T$ and $|M| = (|m_{ij}|) \in \Re^{n \times n}$.

2. The basic algorithm

In this section, we first introduce the optimality conditions for (1). Then we describe how to compute the search directions and the motivation. Finally, we give the basic algorithm.

Conditions for a point $x \in \Re^n$ to be a local minimizer of (1) are well-known and a set of numerically verifiable conditions can be phrased as follows (see for example [19]): if x^* is a solution to (1), then there exists $w^* \in \Re^m$ such that

$$\text{feasibility: } Ax^* = b \text{ and } x^* \geq 0, \tag{2}$$

$$\text{complementarity: } X^*(Hx^* + c + A^T w^*) = 0, \tag{3}$$

$$\text{sign condition: } x_i^* = 0 \Longrightarrow (Hx^* + c + A^T w^*)_i \geq 0 \quad (i = 1, 2, \cdots, n), \tag{4}$$

$$\text{positive semidefiniteness: } p^T H p \geq 0 \text{ for every } p \in \mathcal{N}(x^*), \tag{5}$$

where for a given feasible point x,

$$\mathcal{N}(x) := \{\, p \in \Re^n \ : \ Ap = 0; \ p_i = 0 \text{ for every } i \in \mathcal{A}(x) \,\},$$

and

$$\mathcal{A}(x) := \{\, i \ : \ x_i = 0 \,\}.$$

Conditions (2) – (4) are first-order necessary conditions and are known as the Karush-Kuhn-Tucker (KKT) conditions. Conditions (2) – (5) are called the second-order necessary conditions. Second-order sufficiency conditions are obtained by replacing "$\geq$" with "$>$" in (4) and (5). A point x^* is called a strong local minimizer if x^* satisfies these second-order sufficiency conditions.

The two equations in the KKT conditions form a nonlinear system

$$F(x, w) := \begin{bmatrix} X(Hx + c + A^T w) \\ Ax - b \end{bmatrix} = 0 \tag{6}$$

which will be useful in the proof of quadratic convergence.

:.1. Projected trust region subproblem and projected steepest descent direction

(e's method [37] uses the solution of the following projected trust region subproblem .s the search direction:

$$\min_{\Delta x} \left\{ \frac{1}{2} \Delta x^T H \Delta x + \Delta x^T (Hx + c) \; : \; A\Delta x = 0, \; \|X^{-1}\Delta x\| \le \delta \right\}.$$

In our algorithm, each step involves two directions: the solution of a projected trust egion subproblem and a projected steepest descent direction. The projected trust region ubproblem is motivated by the Newton system of (6): its solution is ultimately the Newton direction (with respect to x) of (6) (see Sect. 4). More specifically, for given x .nd w, we define

$$g = g(x, w) := Hx + c + A^T w, \quad M = M(x, w) := H + X^{-1}|G|,$$

ınd solve

$$\min_{\Delta x} \left\{ \psi(\Delta x) := \frac{1}{2} \Delta x^T M \Delta x + \Delta x^T (Hx + c) \; : \; A\Delta x = 0, \; \|X^{-\frac{1}{2}}\Delta x\| \le \delta \right\}, \tag{7}$$

vhere $\delta \in [\delta_l, \delta_u]$ for some given $0 < \delta_l < \delta_u$.

Let Δx_{tr} denote the solution of (7) (the subscript *tr* stands for *trust region*). Due to he choices of M and the scaling in (7), the solution Δx_{tr} has two important properties. First, it is a feasible descent direction for q subject to the constraints (i.e., $A\Delta x_{tr} = 0$ ınd $\nabla q(x)^T \Delta x_{tr} < 0$, see next section for the proof). This is important for global :onvergence. Second, Δx_{tr} is ultimately the Newton direction (with respect to x) of ;ystem (6). Therefore, local quadratic convergence can be expected.

We emphasize that Δx_{tr} is used as part of the search direction. A line search is ›erformed to maintain feasibility (see next subsection). Hence the parameter δ in (7) ıeed not be updated in theory for convergence (this will be made clear in the next ;ection), though it may affect computational performance.

Problem (7) can be described in a different way. Let

$$D := X^{\frac{1}{2}}, \quad \bar{A} := AD, \quad \bar{M} := DMD, \quad \text{and} \quad \bar{g} := Dg.$$

We may write Δx_{tr} as

$$\Delta x_{tr} = D\bar{Z}\Delta\bar{x}_{tr}, \tag{8}$$

vhere $\bar{Z} = \bar{Z}(x)$ is a matrix whose columns form an orthonormal basis for the null ;pace of $\bar{A}$ and $\Delta\bar{x}_{tr}$ denotes the solution to

$$\min_{\Delta\bar{x}} \left\{ \bar{\psi}(\Delta\bar{x}) := \frac{1}{2} \Delta\bar{x}^T \bar{Z}^T \bar{M} \bar{Z} \Delta\bar{x} + \Delta\bar{x}^T \bar{Z}^T \bar{g} \; : \; \|\Delta\bar{x}\| \le \delta \right\}. \tag{9}$$

Clearly,

$$\psi(\Delta x_{tr}) = \bar{\psi}(\Delta\bar{x}_{tr}). \tag{10}$$

From the results in [15] and [34], there exists $\lambda_{tr} \geq 0$ such that

$$(\bar{Z}^T \bar{M} \bar{Z} + \lambda_{tr} I)\Delta \bar{x}_{tr} = -\bar{Z}^T \bar{g}, \tag{11}$$

$$\bar{Z}^T \bar{M} \bar{Z} + \lambda_{tr} I \geq 0, \tag{12}$$

$$\lambda_{tr}(\delta - \|\Delta \bar{x}_{tr}\|) = 0, \tag{13}$$

and

$$\lambda_{tr} \leq \frac{\|\bar{g}\|}{\delta} + \|\bar{M}\|. \tag{14}$$

By the definition of $\bar{Z}$, there exists Δw_{tr} such that

$$(\bar{M} + \lambda_{tr} I)\bar{Z}\Delta \bar{x}_{tr} + \bar{g} + \bar{A}^T \Delta w_{tr} = 0. \tag{15}$$

Or equivalently,

$$(M + \lambda_{tr} X^{-1})\Delta x_{tr} + g + A^T \Delta w_{tr} = 0 \text{ and } A\Delta x_{tr} = 0. \tag{16}$$

When $\bar{A}\bar{A}^T$ is nonsingular,

$$\Delta w_{tr} = -(\bar{A}\bar{A}^T)^{-1}\bar{A}\,((\bar{M} + \lambda_{tr} I)\bar{Z}\Delta \bar{x}_{tr} + \bar{g}). \tag{17}$$

In our algorithm, Δw_{tr} will not be computed but it will be useful in the convergence analysis.

We do not specify the matrix $\bar{Z}$ in (9). It turns out that global convergence to a point satisfying the KKT conditions holds for any choice of $\bar{Z}$. But continuity of $\bar{Z}$ is required for convergence to a point which satisfies the second-order conditions. See Sect. 3 for more discussion on this.

We prefer to compute Δx_{tr} by (8) and (9) instead of (7) since some entries of the matrix M may approach infinity as some diagonal components of X go to zero. The matrix $\bar{M}$ does not have this disturbing property.

In order to ensure that Δx_{tr} converges to the Newton step (with respect to x) of system (6), we need to update w appropriately. A reasonable way to update w is based on condition (3), i.e., at a local minimizer

$$X\,(Hx + c + A^T w) = 0.$$

Therefore, if AXA^T is nonsingular, we may compute w by

$$w = -(AXA^T)^{-1}AX(Hx + c) \tag{18}$$

for every given x, which solves the least square problem

$$\min_w \|(AD)^T w + D(Hx + c)\|. \tag{19}$$

If w is so chosen, then g and M will only depend on x.

Trust region methods have strong convergence properties and have exhibited robust performance in unconstrained minimization (see [15], [28], [33], and [34]). In our case, however, Δx_{tr} may not always be a good choice for the search direction due to the constraints. Similar to the *dogleg* method [32] and the algorithm in [8], we follow a hybrid strategy. We choose the step by combining Δx_{tr} with a projected steepest descent direction

$$\Delta x_g := \mu_g D \frac{\bar{g}}{\|\bar{g}\|}, \tag{20}$$

where μ_g is defined as the solution to the following problem:

$$\min_{\mu} \left\{ \psi_g(\mu) := \frac{\mu^2}{2} \frac{\bar{g}^T}{\|\bar{g}\|} \bar{M} \frac{\bar{g}}{\|\bar{g}\|} + \mu \frac{\bar{g}^T}{\|\bar{g}\|} \bar{g} \ : \ |\mu| \leq \delta \right\}. \tag{21}$$

It is easy to verify that if w is defined by (18) then

$$\psi(\Delta x_g) = \psi_g(\mu_g), \quad A \Delta x_g = 0, \tag{22}$$

and Δx_g is a projection of $\nabla q(x) = Hx + c$ onto the null space of A. Moreover, Δx_g is a feasible descent direction for q subject to the constraints (i.e., $A\Delta x_g = 0$ and $\nabla q(x)^T \Delta x_g < 0$), and, similar to (11) – (13), there exists $\lambda_g \geq 0$ such that

$$\left(\frac{\bar{g}^T}{\|\bar{g}\|} \bar{M} \frac{\bar{g}}{\|\bar{g}\|} + \lambda_g I \right) \mu_g = - \frac{\bar{g}^T}{\|\bar{g}\|} \bar{g}, \tag{23}$$

$$\frac{\bar{g}^T}{\|\bar{g}\|} \bar{M} \frac{\bar{g}}{\|\bar{g}\|} + \lambda_g \geq 0, \tag{24}$$

$$\lambda_g (\delta - |\mu_g|) = 0, \tag{25}$$

and

$$\lambda_g \leq \frac{\|\bar{g}\|}{\delta} + \|\bar{M}\|. \tag{26}$$

The direction Δx_g is similar to the affine-scaling direction introduced in [1], [13], and [35] but differs in that the affine-scaling direction is more tangential to the near active constraints.

We describe in Sect. 2.3 how to combine Δx_{tr} and Δx_g to form the updating step.

2.2. *Maintaining feasibility*

Let x^k be the current iterate, an interior point, and let Δx be a feasible descent direction of q at x^k subject to the constraints, i.e., $A\Delta x = 0$ and $\nabla q(x^k)^T \Delta x < 0$. When moving in the direction Δx starting from x^k, a variable may reach a bound, i.e., $(x^k + \alpha \Delta x)_i = 0$ for some $1 \leq i \leq n$ and for some $\alpha > 0$. Therefore, to maintain strict feasibility and yet allow the solution (which may have some of the variables at their bounds) to be approached asymptotically and sufficiently fast, we define the step length as follows.

For each k, let

$$\theta^k := \frac{\|\tilde{X}^k g^k\| + |\psi^k(\Delta x_{tr}^k)|}{1 + \|\tilde{X}^k g^k\| + |\psi^k(\Delta x_{tr}^k)|}, \tag{27}$$

where for given (x, w), $\tilde{x} = \tilde{x}(x, w) \in \Re^n$ is a function of (x, w) defined as

$$\tilde{x}_i := \begin{cases} x_i & \text{if } g_i := (Hx + c + A^T w)_i \geq 0, \\ -1 & \text{otherwise} \end{cases} \quad (1 \leq i \leq n). \tag{28}$$

It is clear that $\tilde{X}g = 0$ if and only if the KKT conditions are satisfied with (x, w). Even though $\tilde{x}_i$ might be discontinuous when $g_i = 0$, the product $\tilde{X}g$ will be continuous at a solution of (1).

Let

$$\beta_{tr}^k = \min_{1 \leq i \leq n} \left\{ -\frac{x_i^k}{(\Delta x_{tr}^k)_i} > 0 \right\},$$

$$\text{(so that } (x^k + \beta_{tr}^k \Delta x_{tr}^k)_i = 0 \text{ for some } 1 \leq i \leq n) \tag{29}$$

$$\rho_{tr}^k = \max(\tau_\rho, 1 - \theta^k) \text{ for some given } \tau_\rho \in (0, 1), \tag{30}$$

$$\alpha_{tr}^k = \min(\tau_\alpha, \rho_{tr}^k \beta_{tr}^k) \text{ for some given } \tau_\alpha \in (1, 2), \tag{31}$$

and

$$\beta_g^k = \min_{1 \leq i \leq n} \left\{ -\frac{x_i^k}{(\Delta x_g^k)_i} > 0 \right\}, \tag{32}$$

$$\rho_g^k = \max(\tau_\rho, 1 - \theta^{k-1}), \text{ (with } \rho_g^0 = 1) \tag{33}$$

$$\alpha_g^k = \min(\tau_\alpha, \rho_g^k \beta_g^k). \tag{34}$$

The step lengths for Δx_{tr}^k and Δx_g^k will be α_{tr}^k and α_g^k, respectively. In Theorem 5, we show that $\theta^k > 0$ unless x^k satisfies the second-order necessary conditions (2) – (5). Therefore, if x^k is an interior feasible point not satisfying (2) – (5), then both $x^k + \alpha_{tr}^k \Delta x_{tr}^k$ and $x^k + \alpha_g^k \Delta x_g^k$ are interior feasible points. In addition, Theorem 5 shows that θ^k will not be very small unless x^k is sufficiently close to optimality. Hence the iterates x^k will be prevented from prematurely getting too close to the lower bounds. Moreover, we will show that $\theta^k \longrightarrow 0$ (Theorem 5) and $\alpha_{tr}^k \longrightarrow 1$ (Lemma 11) as x^k converges to a solution of (1), thus the iterates x^k will be allowed to approach a solution sufficiently fast.

In (33), θ^{k-1} is used (instead of θ^k) because Δx_{tr}^k may not be available when computing Δx_g^k, e.g., in the algorithm presented in Sect. 5.

2.3. The algorithm

We compute the updating step s^k as follows:

$$s^k = \sigma^k \alpha_{tr}^k \Delta x_{tr}^k + (1 - \sigma^k) \alpha_g^k \Delta x_g^k, \tag{35}$$

where for some given $\gamma \in (0, 1)$,

$$\sigma^k = \begin{cases} 1, & \text{if } \psi^k(\alpha_{tr}^k \Delta x_{tr}^k) \leq \gamma\, \psi^k(\alpha_g^k \Delta x_g^k), \\ 0, & \text{otherwise.} \end{cases} \tag{36}$$

We can now state our basic algorithm.

Algorithm Interior-Newton

Let x^0 be an interior feasible point.

For $k = 0, 1, 2, \cdots$

1. Solve (19) for w^k and let $g^k = Hx^k + c + A^T w^k$;

   Compute Δx_g^k and α_g^k by (20) and (34).

2. Solve (9) for $\Delta \bar{x}_{tr}^k$; Compute Δx_{tr}^k and α_{tr}^k by (8) and (31).

3. Determine s^k by (35); Update $x^{k+1} = x^k + s^k$.

Note that because of the choice of w in the algorithm, g and M will be functions of only x (instead of functions of both x and w). In the next two sections, we consider the convergence properties of Algorithm Interior-Newton.

3. Convergence of Algorithm Interior-Newton

In this section we establish global convergence properties of Algorithm Interior-Newton. We first show that every limit point of the sequence $\{x^k\}$ generated by Algorithm Interior-Newton satisfies the complementarity condition (3). We then prove that $\{x^k\}$ converges. Finally, we show that the limit point of $\{x^k\}$ satisfies the second-order necessary conditions (2) – (5).

These convergence properties are proved under the following assumptions:

(A1) The level set $\mathcal{L} := \{x : x \text{ is feasible and } q(x) \le q(x^0)\}$ is compact.

(A2) AXA^T is nonsingular for every $x \in \mathcal{L}$.

(A3) For every feasible point x satisfying $X(Hx + c + A^T w) = 0$ for some $w \in \Re^m$,
$x_i = 0 \Longrightarrow (Hx + c + A^T w)_i \ne 0 \;\; (1 \le i \le n)$.

Assumption **(A2)** is known as primal nondegeneracy. Assumption **(A3)** says that for every feasible point satisfying the complementarity condition, strict complementarity holds. Note that assumptions **(A2)** and **(A3)** are different from the primal and dual nondegeneracy assumptions for linear programming in that **(A2)** and **(A3)** do not restrict a feasible point x satisfying the complementarity condition to be a vertex. That is important since for a QP problem, solutions are not necessarily vertices.

By **(A1)** and **(A2)**, there exists $C_1 > 0$ such that

$$\|(AXA^T)^{-1}\| \le C_1 \quad \text{for every } \; x \in \mathcal{L}. \tag{37}$$

Recall that in each iteration,

$$\delta^k \in [\delta_l, \, \delta_u], \tag{38}$$

where $0 < \delta_l < \delta_u < \infty$ are two given scalars.

The following lemma defines a few basic equalities used throughout the remainder of this paper.

Lemma 1. *Let each s^k, Δx_{tr}^k and Δx_g^k be defined by (35), (8) and (20). Then*

$$q(x^k) - q(x^k + s^k) = -\psi^k(s^k) + \frac{1}{2}(s^k)^T (X^k)^{-1} |G^k| s^k; \tag{39}$$

$$\psi^k(t\,\Delta x_{tr}^k) = \bar{\psi}^k(t\,\Delta \bar{x}_{tr}^k) \tag{40}$$

$$= -t\,(1 - \frac{t}{2})\,(\Delta \bar{x}_{tr}^k)^T ((\bar{Z}^k)^T \bar{M}^k \bar{Z}^k + \lambda_{tr}^k I)\Delta \bar{x}_{tr}^k - \frac{t^2}{2}\,\lambda_{tr}^k\,\|\Delta \bar{x}_{tr}^k\|^2 \leq 0$$

$$\forall t \in [0,\ \min(2,\ \beta_{tr}^k)];$$

$$\psi(t\,\Delta x_g^k) = \psi_g^k(t\,\mu_g^k) \tag{41}$$

$$= -t\left(1 - \frac{t}{2}\right)\left(\frac{(\bar{g}^k)^T}{\|\bar{g}^k\|}\bar{M}^k \frac{\bar{g}^k}{\|\bar{g}^k\|} + \lambda_g^k\right)(\mu_g^k)^2 - \frac{t^2}{2}\,\lambda_g^k\,(\mu_g^k)^2 \leq 0$$

$$\forall t \in [0,\ \min(2,\ \beta_g^k)].$$

Proof. The relation (39) is a direct consequence of the definitions of ψ^k and $q(x)$, while (40) and (41) are true by (11), (12), (23), and (24).

□

Theorem 1. *Let $\{x^k\}$ be generated by Algorithm Interior-Newton. Then $\{q(x^k)\}$ converges and*

$$\bar{g}^k = D^k(Hx^k + c + A^T w^k) \longrightarrow 0. \tag{42}$$

Proof. Since $x^{k+1} = x^k + s^k$, by (39), (35), and (36),

$$q(x^k) - q(x^{k+1}) \geq -\gamma\,\psi^k(\alpha_g^k \Delta x_g^k) \geq 0. \tag{43}$$

So $\{q(x^k)\}$ is monotonically decreasing. Therefore, $\{q(x^k)\}$ converges by **(A1)**. To show (42), using (43), (41), and letting $y^k = \frac{\bar{g}^k}{\|\bar{g}^k\|}$, we have

$$\alpha_g^k \left(1 - \frac{\alpha_g^k}{2}\right)((y^k)^T \bar{M}^k y^k + \lambda_g^k)(\mu_g^k)^2 \longrightarrow 0. \tag{44}$$

From (23),

$$((y^k)^T \bar{M}^k y^k + \lambda_g^k)\mu_g^k = -\|\bar{g}^k\|. \tag{45}$$

So if (42) is false, then there exists $\epsilon > 0$ and a subsequence $\{k_j\}$ such that

$$\|\bar{g}^{k_j}\| \geq \epsilon. \tag{46}$$

Applying (45), we see that $((y^{k_j})^T \bar{M}^{k_j} y^{k_j} + \lambda_g^{k_j})(\mu^{k_j})^2$ is bounded away from zero. Hence (44) and (34) imply (remembering $\tau_\alpha < 2$)

$$\alpha_g^{k_j} \longrightarrow 0. \tag{47}$$

herefore, (34) yields $\beta_g^{k_j} \longrightarrow 0$. Since n is finite, from the definition of β_g^k, we may vithout loss of generality) assume that $\beta_g^{k_j} = \frac{-x_1^{k_j}}{(\Delta x_g^{k_j})_1}$. Then by (20),

$$\frac{\|\bar{g}^{k_j}\|}{\mu_g^{k_j} g_1^{k_j}} = \frac{x_1^{k_j}}{(\Delta x_g^{k_j})_1} \longrightarrow 0. \tag{48}$$

herefore, we must have $\|\bar{g}^{k_j}\| \longrightarrow 0$ since $|\mu_g^{k_j}| \leq \delta_u$ for every k_j and by **(A1)** $\{|g_1^{k_j}|\}$ bounded above. This contradicts (46).

□

'orollary 1. *Let $\{x^k\}$ be generated by Algorithm Interior-Newton. Suppose x^* is any mit point of $\{x^k\}$. Then*

$$X^* g^* = 0, \tag{49}$$

'here

$$g^* = Hx^* + c + A^T w^* \quad \textit{and} \quad w^* = -(AX^*A^T)^{-1}AX^*(Hx^* + c). \tag{50}$$

ı other words, every limit point x^* satisfies the complementarity condition.

Next we show that $\{x^k\}$ actually converges. We first state the following two technical ›sults without proofs. Readers may find the proofs in [9].

,emma 2. *Suppose that $x^* \in \mathcal{L}$ satisfies the complementarity condition. If $x \in \mathcal{L}$ is ny point satisfying the complementarity condition and*

$$x_i = 0 \;\; \textit{if and only if} \;\; i \in \mathcal{A}(x^*) \;\; (1 \leq i \leq n),$$

ıen $x = x^*$.

,emma 3. *The number of limit points of $\{x^k\}$ is finite. If x^* is a limit point of $\{x^k\}$ ıtisfying*

$$|\mathcal{A}(x^*)| = \max\{|\mathcal{A}(x)| \; : \; x \textit{ is any limit point of } \{x^k\}\}, \tag{51}$$

'here $|\mathcal{A}(x)|$ denotes the number of elements in the set $\mathcal{A}(x)$, then there is no other limit oint x of $\{x^k\}$ satisfying $x_i = 0$ for every $i \in \mathcal{A}(x^*)$.

'heorem 2. *The sequence $\{x^k\}$ generated by Algorithm Interior-Newton converges to * satisfying (51).*

Proof. If $\mathcal{A}(x^*)$ is an empty set, then any limit point of $\{x^k\}$ would be interior. By Lemma 2, x^* is the only limit point and the theorem is clearly true. Now we assume that $\mathcal{A}(x^*)$ is not empty.

If $\{x^k\}$ does not converge to x^*, then there exists another limit point, say, $\bar{x} \neq x^*$.

By Lemma 3, there exists $\epsilon \in (0,\ 1)$ sufficiently small such that for every limit point $x \neq x^*$, there is an index $i \in \mathcal{A}(x^*)$ such that $x_i > \epsilon$. In particular,

$$\bar{x}_i > \epsilon \quad \text{for some} \quad i \in \mathcal{A}(x^*). \tag{52}$$

Therefore, there exists $k_1 > 0$ such that for every $k \geq k_1$,

$$\text{either } (x_i^k)^{\frac{1}{2}} < \tfrac{\epsilon}{3(1+2\delta_u)} \text{ for every } i \in \mathcal{A}(x^*), \text{ or } x_i^k > \tfrac{2}{3}\epsilon \text{ for some } i \in \mathcal{A}(x^*). \tag{53}$$

(In other words, x^k will be either sufficiently close to x^*, or sufficiently close to another limit point.) Noting that for every k, $\|(D^k)^{-1}\Delta x_{tr}^k\| \leq \delta_u$, $\|(D^k)^{-1}\Delta x_g^k\| \leq \delta_u$, $\alpha_{tr}^k < 2$, and $\alpha_g^k < 2$, we have

$$\|(D^k)^{-1}s^k\| \leq 2\,\delta_u \quad \text{for every } k, \tag{54}$$

which implies that

$$|s_i^k| \leq 2\,\delta_u\,(x_i^k)^{\frac{1}{2}} \quad \text{for every } i \text{ and for every } k. \tag{55}$$

Since x^* is a limit point, there exists $k_2 \geq k_1$ such that

$$(x_i^{k_2})^{\frac{1}{2}} < \frac{\epsilon}{3(1+2\delta_u)} \quad \text{for every} \quad i \in \mathcal{A}(x^*). \tag{56}$$

Then by (55) and (56),

$$x_i^{k_2+1} \leq (1+2\,\delta_u)(x_i^{k_2})^{\frac{1}{2}} \leq \frac{\epsilon}{3} \quad \text{for every } i \in \mathcal{A}(x^*). \tag{57}$$

Therefore, by (53),

$$(x_i^{k_2+1})^{\frac{1}{2}} < \frac{\epsilon}{3(1+2\delta_u)} \quad \text{for every } i \in \mathcal{A}(x^*). \tag{58}$$

By induction, we have that for every $k \geq k_2$,

$$x_i^k \leq (x_i^k)^{\frac{1}{2}} < \frac{\epsilon}{3(1+2\delta_u)} \leq \frac{\epsilon}{3} \quad \text{for every } i \in \mathcal{A}(x^*). \tag{59}$$

This is a contradiction to (52).

□

The next result gives a sufficient condition for x^* to satisfy the KKT conditions.

Lemma 4. *The limit point x^* of $\{x^k\}$ satisfies the KKT conditions (2) – (4) if for every k sufficiently large and for every $i \in \mathcal{A}(x^*)$, s_i^k and g_i^* have opposite sign.*

'roof. We need only to show that $g_i^* \geq 0$ for every $i \in \mathcal{A}(x^*)$. In fact, if $g_i^* < 0$ for ome $i \in \mathcal{A}(x^*)$, we would have $s_i^k > 0$ for every k sufficiently large. Then $x_i^{k+1} > x_i^k$ nd $\lim x_i^k > 0$, which is a contradiction to $\lim x_i^k = x_i^* = 0$.

□

The condition in Lemma 4 is crucial to the convergence to a point satisfying the :KT conditions. Therefore, it is important to realize that this condition can be satisfied 'ith $\{s^k\}$. This we prove next by several lemmas.

,emma 5. *Let x^* be the limit point of $\{x^k\}$. Then for every k sufficiently large and for very $i \in \mathcal{A}(x^*)$, $(\Delta x_g^k)_i$ and g_i^* have opposite sign.*

'roof. Let $i \in \mathcal{A}(x^*)$. By **(A3)**, $g_i^* \neq 0$. So when k is sufficiently large, g_i^k and g_i^* have ie same sign, and, by (24) and (45), $\mu_g^k < 0$. Therefore, by (20), $(\Delta x_g^k)_i$ and g_i^k have pposite sign, and the lemma follows.

□

'emark. One implication of the previous two results is that Algorithm Interior-Newton uarantees convergence to a point satisfying the KKT conditions if s^k is replaced with ie truncated projected steepest descent step $\alpha_g^k \Delta x_g^k$ for every k sufficiently large. In ther words, the new projected steepest descent directions yield a procedure converging lobally to a point satisfying the KKT conditions, even in the nonconvex case.

,emma 6. *There exists $\bar{\alpha} > 0$ such that $\alpha_{tr}^k \geq \bar{\alpha}$ for every k.*

'roof. If the lemma is false, then there exists a subsequence $\{k_j\}$ such that

$$\alpha_{tr}^{k_j} \longrightarrow 0.$$

'herefore, by (31), $\beta_{tr}^{k_j} \longrightarrow 0$. Since n is finite, from the definition of β_{tr}^k, we may without loss of generality) assume that $\beta_{tr}^{k_j} = \frac{-x_1^{k_j}}{(\Delta x_{tr}^{k_j})_1}$. Since $\|\Delta x_{tr}^k\| \leq \delta_u \|D^k\|$ by (8), 'e see by **(A1)** that $\{\|\Delta x_{tr}^k\|\}$ is bounded above. So $x_1^{k_j} \longrightarrow 0$. Hence, $x_1^* = 0$ and by **A3)**, $|g_1^*| > 0$. But multiplying the first equation of (16) with X^k gives

$$X^k H \Delta x_{tr}^k + |G^k| \Delta x_{tr}^k + \lambda_{tr}^k \Delta x_{tr}^k + X^k g^k + X^k A^T \Delta w_{tr}^k = 0, \tag{60}$$

'hich yields

$$\frac{|g_1^{k_j}| + \lambda_{tr}^{k_j}}{(H\Delta x_{tr}^{k_j})_1 + g_1^{k_j} + (A^T \Delta w_{tr}^{k_j})_1} = \frac{-x_1^{k_j}}{(\Delta x_{tr}^{k_j})_1} = \beta_{tr}^{k_j} \longrightarrow 0.$$

'herefore, by the fact that $g_1^{k_j} \longrightarrow g_1^* > 0$, we must have

$$\left|(H\Delta x_{tr}^{k_j})_1 + g_1^{k_j} + (A^T \Delta w_{tr}^{k_j})_1\right| \longrightarrow \infty,$$

'hich is impossible since by **(A1)**, (14), (17) and (37), $\{\|H\Delta x_{tr}^k + g^k + A^T \Delta w_{tr}^k\|\}$ is ounded. Therefore, the lemma is true.

□

Lemma 7. *Let $\{\lambda_{tr}^k\}$ be given in (11) – (13). Then*

$$\lambda_{tr}^k \longrightarrow 0. \tag{61}$$

Proof. By (39), (40) and the convergence of $\{q(x^k)\}$,

$$q(x^k) - q(x^{k+1}) \geq -\psi^k\,(\alpha_{tr}^k \Delta x_{tr}^k) \geq \frac{(\alpha_{tr}^k)^2}{2} \lambda_{tr}^k \|\Delta \bar{x}_{tr}^k\|^2 \longrightarrow 0.$$

So by Lemma 6, we have $\lambda_{tr}^k \|\Delta \bar{x}_{tr}^k\|^2 \longrightarrow 0$. Then (61) follows by (13).

□

Lemma 8. *Let x^* be the limit point of $\{x^k\}$. Let $\{\Delta x_{tr}^{k_j}\}$ be any subsequence satisfying $\Delta x_{tr}^{k_j} \longrightarrow 0$. Then for every k_j sufficiently large and for every $i \in \mathcal{A}(x^*)$, $(\Delta x_{tr}^{k_j})_i$ and g_i^* have opposite sign.*

Proof. Let $\epsilon^* := \min\{|g_i^*| \,:\, i \in \mathcal{A}(x^*)\}$. Then by **(A3)**, $\epsilon^* > 0$ and

$$|g_i^k| > \frac{\epsilon^*}{2} > 0 \text{ for every } k \text{ sufficiently large and for every } i \in \mathcal{A}(x^*). \tag{62}$$

By **(A1)**, (14), (17), (37) and the convergence of $\{x^k\}$, $\{\|H\Delta x_{tr}^k + g^k + A^T \Delta w_{tr}^k\|\}$ is bounded. So by (60), there exists $C_2 > 0$ such that for every k sufficiently large,

$$\left|\,(\Delta x_{tr}^k)_i\right| \leq x_i^k \frac{\left|(H\Delta x_{tr}^k)_i + g_i^k + (A^T \Delta w_{tr}^k)_i\right|}{|g_i^k| + \lambda_{tr}^k} \leq C_2 x_i^k \;\text{ for every }\; i \in \mathcal{A}(x^*).$$

Consequently, by (8),

$$\left|\,(\bar{Z}^k \Delta \bar{x}_{tr}^k)_i\right| = \frac{\left|\,(\Delta x_{tr}^k)_i\right|}{(x_i^k)^{\frac{1}{2}}} \leq C_2 (x_i^k)^{\frac{1}{2}} \longrightarrow 0 \;\text{ for every }\; i \in \mathcal{A}(x^*).$$

So,

$$|G^k| \bar{Z}^k \Delta \bar{x}_{tr}^k \longrightarrow 0 \;\text{ since }\; g_i^k \longrightarrow 0 \;\text{ for every }\; i \notin \mathcal{A}(x^*).$$

Hence by the assumption that $D^{k_j} \bar{Z}^{k_j} \Delta \bar{x}_{tr}^{k_j} = \Delta x_{tr}^{k_j} \longrightarrow 0$, we have

$$\bar{M}^{k_j} \bar{Z}^{k_j} \Delta \bar{x}_{tr}^{k_j} = D^{k_j} H D^{k_j} \bar{Z}^{k_j} \Delta \bar{x}_{tr}^{k_j} + |G^{k_j}| \bar{Z}^{k_j} \Delta \bar{x}_{tr}^{k_j} \longrightarrow 0.$$

Therefore, by **(A1)**, (17), (37), (61), and the fact that $\bar{g}^k \longrightarrow 0$, we have

$$\Delta w_{tr}^{k_j} \longrightarrow 0. \tag{63}$$

From (60),

$$(\Delta x_{tr}^{k_j})_i = -x_i^{k_j} \frac{(H\Delta x_{tr}^{k_j})_i + g_i^{k_j} + (A^T \Delta w_{tr}^{k_j})_i}{|g_i^{k_j}| + \lambda_{tr}^{k_j}}.$$

By (62), (63) and the fact that $\Delta x_{tr}^{k_j} \longrightarrow 0$, we see that for every k_j sufficiently large and for every $i \in \mathcal{A}(x^*)$, $(\Delta x_{tr}^{k_j})_i$ and $g_i^{k_j}$ have opposite sign. Then the lemma follows.

□

We can now state the following convergence result.

heorem 3. *Let x^* be the limit point of $\{x^k\}$. Then for every k sufficiently large and for very $i \in \mathcal{A}(x^*)$, s_i^k and g_i^* have opposite sign. Therefore, the sequence $\{x^k\}$ generated y Algorithm Interior-Newton converges (globally) to x^* satisfying the KKT conditions.*

roof. The convergence of $\{x^k\}$ implies that $s^k = x^{k+1} - x^k \longrightarrow 0$. By Lemma 6, the ıbsequence

$$\left\{\Delta x_{tr}^k \,:\, \sigma^k = 1\right\} = \left\{\frac{s^k}{\alpha_{tr}^k} \,:\, \sigma^k = 1\right\}$$

onverges to zero. So by Lemma 8, s_i^k (with $\sigma^k = 1$) and g_i^* have opposite sign for very k sufficiently large and for every $i \in \mathcal{A}(x^*)$. On the other hand, by Lemma 5, $^k = \alpha_g^k\,(\Delta x_g^k)_i$ (with $\sigma^k = 0$) and g_i^* have opposite sign for every k sufficiently large nd for every $i \in \mathcal{A}(x^*)$. Then the lemma follows.

□

Next, we turn to the second-order optimality conditions. To simplify the notation, we efine $g^* = g(x^*)$, $\bar{M}^* = \bar{M}(x^*)$, $\bar{Z}^* = \bar{Z}(x^*)$, and $\mathcal{N}^* = \mathcal{N}(x^*)$ for a given feasible oint x^*. The following lemma reveals the relation between the positive definiteness of I in $\mathcal{N}^*$ and that of $(\bar{Z}^*)^T\bar{M}^*\bar{Z}^*$. The proof can be found in [9].

emma 9. *Let x^* be any feasible point satisfying the complementarity condition. Then*

) $(\bar{Z}^*)^T\bar{M}^*\bar{Z}^* > 0$ *if and only if* $p^THp > 0$ *for every* $p \in \mathcal{N}^*$ *and* $p \neq 0$.

i) $(\bar{Z}^*)^T\bar{M}^*\bar{Z}^* \geq 0$ *if and only if* $p^THp \geq 0$ *for every* $p \in \mathcal{N}^*$.

All the previous results hold as long as the columns of $\bar{Z}^k$ form an orthonormal basis ər the null space of $\bar{A}^k = AD^k$ and we do not specify $\bar{Z}^k$ in Algorithm Interior-Newton. 'or convergence to a point satisfying the second-order conditions, however, continuity eeds to be imposed on $\bar{Z}^k$ and we will assume here that $\bar{Z}(x)$ is continuous in the ppropriate region. This may not be true for an arbitrary choice of $\bar{Z}$. For discussions n this regard see [4], [10], and [17].

heorem 4. *The sequence $\{x^k\}$ generated by Algorithm Interior-Newton converges globally) to x^* satisfying the second-order necessary conditions (2) – (5).*

roof. By Theorem 3, we need only to show that the limit point x^* satisfies condition (5). n fact, by (12), (61) and the convergence of $\{x^k\}$ to x^*, we have $(\bar{Z}^*)^T\bar{M}^*\bar{Z}^* \geq 0$. hen condition (5) follows by Lemma 9.

□

To conclude this section, we justify the statement about θ^k given at the end of ect. 2.2.

heorem 5. *Let θ^k be defined by (27). Let ν^k denote the least eigenvalue of $(\bar{Z}^k)^T\bar{M}^k\bar{Z}^k$. hen*

(i) $\theta^k \longrightarrow 0$.
(ii) $\|\bar{X}^k g^k\| \leq 2\,\theta^k$ *and* $-\frac{4\theta^k}{\delta_l^2} \leq \nu^k$ *for every k such that* $\theta^k \leq \frac{1}{2}$*, where* $\delta_l > 0$ *is the constant used in (38).*
(iii) $\theta^k = 0$ *if and only if* x^k *satisfies the second-order necessary conditions (2) – (5).*

Proof. Proof of (i). Since x^k converges to a point satisfying the KKT conditions, we have $\|\bar{X}^k g^k\| \longrightarrow 0$. So it suffices to show that

$$\psi^k\,(\Delta x_{tr}^k) = \bar{\psi}^k\,(\Delta \bar{x}_{tr}^k) \longrightarrow 0. \tag{64}$$

By (11) and the fact that $\bar{g} \longrightarrow 0$, we have

$$\lim\,\left((\Delta \bar{x}_{tr}^k)^T\,(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}_{tr}^k + \lambda_{tr}^k\,\|\Delta \bar{x}_{tr}^k\|^2\right) = 0$$

which implies that

$$\limsup\,(\Delta \bar{x}_{tr}^k)^T\,(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}_{tr}^k \leq 0.$$

On the other hand, $\liminf\,(\Delta \bar{x}_{tr}^k)^T\,(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}_{tr}^k \geq 0$ since $(\bar{Z}^*)^T \bar{M}^* \bar{Z}^* \geq 0$. Therefore,

$$\lim\,(\Delta \bar{x}_{tr}^k)^T\,(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}_{tr}^k = 0.$$

Then (64) follows from (40) and (61).

Proof of (ii). By (27), it is clear that $\|\bar{X}^k g^k\| \leq 2\,\theta^k$. In addition, by (27) and (64),

$$\left|\bar{\psi}^k\,(\Delta \bar{x}_{tr}^k)\right| = \left|\psi^k\,(\Delta x_{tr}^k)\right| \leq 2\,\theta^k,$$

which by (40) implies that

$$\frac{1}{2}\lambda_{tr}^k\,\|\Delta \bar{x}_{tr}^k\|^2 \leq 2\,\theta^k.$$

If $\|\Delta \bar{x}_{tr}^k\| < \delta^k$, then (13) yields $\lambda_{tr}^k = 0$ which implies that $\nu^k \geq 0 \geq -\frac{4\theta^k}{\delta_l^2}$. Otherwise, $\|\Delta \bar{x}_{tr}^k\| = \delta^k \geq \delta_l$ and

$$\lambda_{tr}^k \leq \frac{4\theta^k}{\delta_l^2}.$$

Therefore, by (12),

$$\nu^k \geq -\lambda_{tr}^k \geq -\frac{4\theta^k}{\delta_l^2}.$$

Proof of (iii). Assume first that $\theta^k = 0$. Then it follows from Part (i) that x^k satisfies the second-order conditions (2) – (5).

Now we assume that x^k satisfies conditions (2) – (5). Then $\bar{X}^k g^k = 0$ and $\bar{g}^k = D^k g^k = 0$. So (11) implies

$$(\Delta \bar{x}_{tr}^k)^T\,(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}_{tr}^k + \lambda_{tr}^k\|\Delta \bar{x}_{tr}^k\|^2 = 0. \tag{65}$$

n the other hand, by Lemma 9, condition (5) implies that $(\Delta \bar{x}^k_{tr})^T (\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}^k_{tr} \geq 0$. nen, by (65) and (13), $\lambda^k_{tr} = 0$. Therefore, using (11) and the fact that $\bar{g}^k = 0$, we have

$$\psi^k (\Delta x^k_{tr}) = \bar{\psi}^k (\Delta \bar{x}^k_{tr}) = \frac{1}{2} (\Delta \bar{x}^k_{tr})^T (\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \Delta \bar{x}^k_{tr} = 0.$$

hen $\theta^k = 0$.

□

Quadratic convergence of Algorithm Interior-Newton

ı this section we consider the local convergence rate properties of Algorithm Interior-ewton: we establish superlinear and quadratic convergence results.

In addition to **(A1)**, **(A2)**, and **(A3)**, we assume the following:

4) $p^T H p > 0$ for every $p \in \mathcal{N}(x^*)$ and $p \neq 0$, where x^* is the limit point of $\{x^k\}$.

That is, we assume the limit point x^* is a strong local minimizer.

emma 10. *Let* $\{\Delta x^k_{tr}\}$, $\{\Delta w^k_{tr}\}$, *and* $\{\lambda^k_{tr}\}$ *be given by (8), (17), and (11). Then*

$$\Delta x^k_{tr} \longrightarrow 0, \quad \Delta w^k_{tr} \longrightarrow 0, \tag{66}$$

nd

$$\lambda^k_{tr} = 0, \quad (\bar{Z}^k)^T \bar{M}^k \bar{Z}^k > 0 \ \textit{ for every } k \textit{ sufficiently large.} \tag{67}$$

roof. By **(A4)** and Lemma 9, $(\bar{Z}^*)^T \bar{M}^* \bar{Z}^* > 0$. Then by (61), there exists $C_3 > 0$ uch that for every k sufficiently large, $(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k + \lambda^k_{tr} I > 0$ and

$$\| ((\bar{Z}^k)^T \bar{M}^k \bar{Z}^k + \lambda^k_{tr} I)^{-1} \| \leq C_3.$$

o using (11) and the fact that $\bar{g}^k \longrightarrow 0$, we have

$$\Delta \bar{x}^k_{tr} \longrightarrow 0.$$

herefore,

$$\Delta x^k_{tr} = D^k \bar{Z}^k \Delta \bar{x}^k_{tr} \longrightarrow 0, \tag{68}$$

nd by (13),

$$\lambda^k_{tr} = 0 \ \text{ for every } k \text{ sufficiently large.}$$

onsequently,

$$(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k > 0 \ \text{ for every } k \text{ sufficiently large.}$$

inally, similar to (63), we have

$$\Delta w^k_{tr} \longrightarrow 0.$$

□

Lemma 11. *Let $\{\alpha_{tr}^k\}$ be given by (31). Then $\alpha_{tr}^k \longrightarrow 1$.*

Proof. By definition, $x_i^* > 0$ for every $i \notin \mathcal{A}(x^*)$. So by (68),

$$-\frac{x_i^k}{(\Delta x_{tr}^k)_i} \longrightarrow \infty \text{ for every } i \notin \mathcal{A}(x^*).$$

By **(A3)**, $g_i^* \neq 0$ for every $i \in \mathcal{A}(x^*)$. Using (60), (66), and (67), we have

$$-\frac{x_i^k}{(\Delta x_{tr}^k)_i} = \frac{|g_i^k| + \lambda_{tr}^k}{(H\Delta x_{tr}^k)_i + g_i^k + (A^T \Delta w_{tr}^k)_i} \longrightarrow 1 \text{ for every } i \in \mathcal{A}(x^*). \tag{69}$$

Therefore, by (29), (30), (31), and the fact that $\theta^k \longrightarrow 0$, we have $\beta_{tr}^k \longrightarrow 1$, $\rho_{tr}^k \longrightarrow 1$, and $\alpha_{tr}^k \longrightarrow 1$.

□

The following result shows that $\sigma^k = 1$ will hold for every k sufficiently large, where σ^k is defined in (36). Therefore, the updates will be

$$x^{k+1} = x^k + \alpha_{tr}^k \Delta x_{tr}^k \text{ for every } k \text{ sufficiently large.}$$

Lemma 12. *The equality $\sigma^k = 1$ will hold for every k sufficiently large, where σ^k is defined in (36).*

Proof. By (36), we need to show that

$$\psi^k\,(\alpha_{tr}^k \Delta x_{tr}^k) \leq \gamma\, \psi^k\,(\alpha_g^k \Delta x_g^k) \text{ for every } k \text{ sufficiently large.} \tag{70}$$

In fact, by (11) and (67), we have

$$\psi^k\,(\alpha_{tr}^k \Delta x_{tr}^k) = \alpha_{tr}^k (2 - \alpha_{tr}^k) \bar{\psi}^k\,(\Delta \bar{x}_{tr}^k).$$

On the other hand, it is easy to see that $\lambda_g^k = 0$ for every k sufficiently large. Then, since $\Delta \bar{x}_{tr}^k$ solves (9) and μ_g^k solves (21), we have

$$\psi^k\,(\alpha_g^k \Delta x_g^k) \geq \psi^k\,(\Delta x_g^k) = \psi_g^k\,(\mu_g^k) \geq \bar{\psi}^k\,(\Delta \bar{x}_{tr}^k).$$

Therefore, by Lemma 11 and (40),

$$\frac{\psi^k\,(\alpha_{tr}^k \Delta x_{tr}^k)}{\psi^k\,(\alpha_g^k \Delta x_g^k)} = \frac{\alpha_{tr}^k (2 - \alpha_{tr}^k) \bar{\psi}^k\,(\Delta \bar{x}_{tr}^k)}{\psi^k\,(\alpha_g^k \Delta x_g^k)} \geq \alpha_{tr}^k (2 - \alpha_{tr}^k) \longrightarrow 1.$$

Since $\gamma < 1$, we see that (70) is true.

□

Now we cite a standard result (see, e.g., [12]), used subsequently.

Theorem 6. *Let $\mathcal{D} \subseteq \Re^n$ be an open convex set. Let $y^* \in \mathcal{D}$, $F : \Re^n \longrightarrow \Re^n$, $F(y^*)=0$, $\nabla F(y^*)$ be nonsingular, and ∇F be Lipschitz continuous at y^* in $\mathcal{D}$. Let $\{T_k\}$ be a sequence of nonsingular matrices in $\Re^{n \times n}$. Suppose for some $y^0 \in \mathcal{D}$ that the sequence of points generated by $y^{k+1} = y^k - T_k^{-1} F(y^k)$ remains in $\mathcal{D}$, $y^k \neq y^*$ for every k, and $y^k \longrightarrow y^*$.*

If $\|T_k - \nabla F(y^*)\| \longrightarrow 0$, then $\{y^k\}$ converges superlinearly to y^*.

If $\|T_k - \nabla F(y^*)\| = O\left(\|y^k - y^*\|\right)$, then $\{y^k\}$ converges quadratically to y^*.

We show next that the conditions in Theorem 6 hold with function F defined in (6). We note that

$$\nabla F^* = \nabla F(x^*, w^*) = \begin{bmatrix} G^* + X^* H & X^* A^T \\ A & 0 \end{bmatrix}. \tag{71}$$

Lemma 13. *The matrix ∇F^* is nonsingular.*

Proof. Let $\nabla F^* \begin{bmatrix} \Delta x \\ \Delta w \end{bmatrix} = 0$. Then by (71),

$$A \Delta x = 0 \tag{72}$$

$$(G^* + X^* H)\Delta x + X^* A^T \Delta w = 0. \tag{73}$$

Since $x_i^* = 0$ for every $i \in \mathcal{A}(x^*)$, we have from (73) that $(G^* \Delta x)_i = 0$ for every $i \in \mathcal{A}(x^*)$. But $g_i^* \neq 0$ for every $i \in \mathcal{A}(x^*)$. So $\Delta x_i = 0$ for every $i \in \mathcal{A}(x^*)$ and by (72), we have $\Delta x \in \mathcal{N}(x^*)$. Using (73), and the facts that $g_i^* = 0$ and $x_i^* \neq 0$ for every $i \notin \mathcal{A}(x^*)$, we have $(H\Delta x)_i + (A^T \Delta w)_i = 0$ for every $i \notin \mathcal{A}(x^*)$. Then $\Delta x^T H \Delta x = 0$ which by **(A4)** implies that $\Delta x = 0$. By (73) and **(A2)**, we also have $\Delta w = 0$, which implies that ∇F^* is nonsingular.

□

The next result is obvious so we state it without proof.

Lemma 14. *There exists $\epsilon > 0$ and a neighborhood $B_\epsilon(x^*, w^*) = \{(x, w) : \|(x, w) - (x^*, w^*)\| \leq \epsilon\}$ of (x^*, w^*) such that ∇F is Lipschitz continuous at (x^*, w^*) in $B_\epsilon(x^*, w^*)$, and $\nabla F(x, w)$ is nonsingular for every $(x, w) \in B_\epsilon(x^*, w^*)$.*

It is not clear how to directly apply Theorem 6 to the sequence $\{(x^k, w^k)\}$ since $\{w^k\}$ is not updated in the form $w^{k+1} = w^k + \Delta w^k$. Therefore, to establish the convergence rates of $\{(x^k, w^k)\}$ and $\{x^k\}$, intermediate steps need to be inspected. We shall introduce three auxiliary sequences $\{\hat{w}^k\}$, $\{\Delta \hat{w}^k\}$, and $\{\hat{g}^k\}$ below. Note that these auxiliary sequences are used only for the purpose of analysis. They are not actually computed.

First, we consider the sequence $\{(x^k, \hat{w}^k)\}$ where $\{\hat{w}^k\}$ is defined by

$$\hat{w}^k = w^{k-1} + \Delta w_{tr}^{k-1} \quad (k \geq 1), \tag{74}$$

and $\{\Delta w_{tr}^k\}$ is defined by (17). By (66),

$$\hat{w}^k \longrightarrow w^*.$$

Using (16) and the fact that $\lambda_{tr}^k = 0$ for every k sufficiently large, we have

$$(H + (X^k)^{-1}|G^k|)\Delta x_{tr}^k + A^T \Delta w_{tr}^k = -g^k.$$

So, if for $k \geq 0$ we let

$$\Delta \hat{w}^k = \hat{w}^{k+1} - \hat{w}^k, \tag{75}$$
$$\hat{g}^k = Hx^k + c + A^T \hat{w}^k, \tag{76}$$

then

$$(H + (X^k)^{-1}|G^k|)\Delta x_{tr}^k + A^T \Delta \hat{w}^k = -\hat{g}^k.$$

Define

$$T_\alpha^k := \begin{bmatrix} \frac{1}{\alpha_{tr}^k}(|G^k| + X^k H) & X^k A^T \\ \frac{1}{\alpha_{tr}^k} A & 0 \end{bmatrix}.$$

Since $|G^k| \longrightarrow |G^*| = G^*$, by Lemma 11, we have

$$T_\alpha^k \longrightarrow \nabla F^*. \tag{77}$$

So T_α^k is nonsingular for every k sufficiently large and

$$(\alpha_{tr}^k \Delta x_{tr}^k, \Delta \hat{w}^k) = -(T_\alpha^k)^{-1} F(x^k, \hat{w}^k).$$

Therefore, by Lemma 12

$$(x^{k+1}, \hat{w}^{k+1}) = (x^k, \hat{w}^k) - (T_\alpha^k)^{-1} F(x^k, \hat{w}^k).$$

Using Theorem 6 and (77), we have the following theorem:

Theorem 7. *The sequence $(x^k, \hat{w}^k)$ converges superlinearly to (x^*, w^*).*

Next, we show that $\{(x^k, \hat{w}^k)\}$ converges quadratically to (x^*, w^*).

Lemma 15. *The sequence $\{\alpha_{tr}^k\}$ satisfies $|1 - \alpha_{tr}^k| = O\,(\|x^k - x^*\|)$.*

Proof. First, for every k sufficiently large,

$$w^* - w^k = (AX^*A^T)^{-1}AX^*(Hx^* + c) - (AX^kA^T)^{-1}AX^k(Hx^k + c).$$

By (37), it is easy to verify that

$$\|w^k - w^*\| = O\,(\|x^k - x^*\|). \tag{78}$$

Consequently,

$$\|X^k g^k\| = \|X^k g^k - X^* g^*\| = O\,(\|x^k - x^*\|), \tag{79}$$

and

$$\|\tilde{X}^k g^k\| = O\,(\|x^k - x^*\|).$$

By (16), we have

$$\begin{bmatrix} |G^k| + X^k H & X^k A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{tr}^k \\ \Delta w_{tr}^k \end{bmatrix} = \begin{bmatrix} -X^k g^k \\ 0 \end{bmatrix}.$$

It is easy to see that the sequence of the coefficient matrices converges to ∇F^* since $|G^*| = G^*$. So, by Lemma 13 and (79),

$$\|\Delta x_{tr}^k\| + \|\Delta w_{tr}^k\| = O\,(\|x^k - x^*\|). \tag{80}$$

Also, multiplying the first equation of (16) by $(\Delta x_{tr}^k)^T$ yields

$$|\psi^k\,(\Delta x_{tr}^k)| = O\,(\|x^k - x^*\|).$$

Therefore,

$$\theta^k = O\,(\|x^k - x^*\|), \quad \text{and} \quad |1 - \rho_{tr}^k| = O\,(\|x^k - x^*\|). \tag{81}$$

In addition, similar to (69), for every $i \in \mathcal{A}(x^*)$ and for every k sufficiently large,

$$\left|1 + \frac{x_i^k}{(\Delta x_{tr}^k)_i}\right| = \left|\frac{(H\Delta x_{tr}^k)_i + (A^T \Delta w_{tr}^k)_i}{(H\Delta x_{tr}^k)_i + g_i^k + (A^T \Delta w_{tr}^k)_i}\right| = O\,(\|x^k - x^*\|).$$

So,

$$|1 - \beta_{tr}^k| = O\,(\|x^k - x^*\|),$$

and therefore, by (81),

$$|1 - \alpha_{tr}^k| = O\,(\|x^k - x^*\|). \tag{82}$$

□

Theorem 8. *The sequence $\{(x^k, \hat{w}^k)\}$ converges quadratically to (x^*, w^*).*

Proof. We show that

$$\|T_\alpha^k - \nabla F^*\| = O\,(\|x^k - x^*\|). \tag{83}$$

Then the desired result follows by Theorem 6. In fact, let $\|\cdot\|_F$ denote the Frobenius norm, we have

$$\begin{aligned} \|T_\alpha^k - \nabla F^*\|_F = {} & \left\|\frac{1}{\alpha_{tr}^k}|G^k| - G^*\right\|_F + \left\|\left(\frac{1}{\alpha_{tr}^k}X^k - X^*\right) H\right\|_F \\ & + \|(X^k - X^*)\,A^T\|_F + \left\|\left(\frac{1}{\alpha_{tr}^k} - 1\right) A\right\|_F. \end{aligned}$$

It is not hard to show that

$$\||G^k| - G^*\|_F + \|X^k - X^*\|_F = O\,(\|x^k - x^*\|). \tag{84}$$

So by (82) and the equivalence of the norms, (83) holds.

□

Next we establish the convergence rate of the sequences $\{x^k\}$ and $\{(x^k, w^k)\}$. To do this, we need the following lemma which can be regarded as a complement to Theorem 6. Its proof is similar to that of Theorem 8.2.4 in [12].

Lemma 16. *Suppose in Theorem 6, there is a partition $y = (y_1, y_2)$ and $\|T_k - \nabla F(y^*)\| = O(\|y_1^k - y_1^*\|)$. Then there exists $C > 0$ such that for every k sufficiently large,*

$$\|y_1^{k+1} - y_1^*\| \leq C \, \|y^{k+1} - y^k\| \, \|y_1^k - y_1^*\|.$$

Theorem 9. *There exists $C > 0$ such that for every k sufficiently large,*

$$\|x^{k+1} - x^*\| \leq C \, \|x^{k-1} - x^*\| \, \|x^k - x^*\|, \tag{85}$$

$$\|(x^{k+1}, w^{k+1}) - (x^*, w^*)\| \leq$$
$$C \, \|(x^{k-1}, w^{k-1}) - (x^*, w^*)\| \, \|(x^k, w^k) - (x^*, w^*)\|. \tag{86}$$

Proof. By (83) and Lemma 16, there exists $C > 0$ such that for every k sufficiently large,

$$\|x^{k+1} - x^*\| \leq C \, \|(x^{k+1}, \hat{w}^{k+1}) - (x^k, \hat{w}^k)\| \, \|x^k - x^*\|. \tag{87}$$

Then (85) and (86) follow by (74), (78), and (80).

□

Theorem 9 shows that the sequences $\{x^k\}$ and $\{(x^k, w^k)\}$ have quadratic convergence property: the rate of convergence is at least 2-step quadratic.

5. A modification of Algorithm Interior-Newton and numerical experiments

We have shown above that Algorithm Interior-Newton has strong convergence properties. However, Algorithm Interior-Newton uses the scaling matrix $D^k = (X^k)^{\frac{1}{2}}$ which is based on the complementarity condition (3), but reveals no information about the sign condition (4). On the other hand, conditions (3) and (4) together are equivalent to the following system:

$$\tilde{X}(Hx + c + A^T w) = 0, \tag{88}$$

where $\tilde{X}$ is defined in (28). So system (3) is weaker than (88), and an algorithm with scaling matrices based on (88) would be expected to outperform Algorithm Interior-Newton: those points, satisfying complementary slackness and feasibility but not the KKT conditions, are attractors for (3) but not for (88).

This logic leads to the consideration of $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ in Algorithm Interior-Newton. Unfortunately, the unbridled use of $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ does not appear to yield a convergent process. Nevertheless, computational performance using $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ often yields a significant improvement over the use of $D^k = (X^k)^{\frac{1}{2}}$. Consequently, we develop an algorithm that mixes the use of these two different scaling matrices. We use $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ when progress is good – in most cases this scaling accelerates progress toward a neighborhood of the solution. If progress is weak, $D^k = (X^k)^{\frac{1}{2}}$ is used. In this fashion impressive computational performance is achieved while maintaining the strong global convergence properties of Algorithm Interior-Newton. The second-order convergence rate is maintained. We denote this modified algorithm, Interior-Newton$_2$, which we describe below.

Algorithm Interior-Newton$_2$

Let x^0 be an interior feasible point. Let $0 < \tau_1,\ \tau_2,\ \tau_3 < 1$.

For $k = 0, 1, 2, \cdots$

1. (Select a trial D^k.)

If $k = 0$ or $\dfrac{\psi^{k-1}\ (\alpha_{tr}^{k-1}\Delta x_{tr}^{k-1})}{\psi^{k-1}\ (\Delta x_{tr}^{k-1})} \leq \min\ (\tau_1,\ \tau_2\theta^{k-1})$ or $\dfrac{\psi^{k-1}\ (\alpha_{tr}^{k-1}\Delta x_{tr}^{k-1})}{\psi^{k-1}\ (\alpha_{g}^{k-1}\Delta x_{g}^{k-1})} \leq \tau_3$

$$D^k = (X^k)^{\frac{1}{2}};$$

else

$$D^k = |\tilde{X}^k|^{\frac{1}{2}}, \quad \text{where } \tilde{x}^k = \tilde{x}^k(x^k,\ w^{k-1}) \text{ is defined by (28);}$$

end.

(Compute a trial w^k, g^k, Δx_g^k, α_g^k, and perform an acceptance test to determine D^k.)

If $D^k = |\tilde{X}^k|^{\frac{1}{2}}$

Compute $t^k = \|\tilde{X}^k(Hx^k + c + A^T w^{k-1})\| \,/$
$(1 + \|\tilde{X}^k(Hx^k + c + A^T w^{k-1})\|)$;

Compute $w^k = -(A|\tilde{X}^k|A^T)^{-1}A|\tilde{X}^k|(Hx^k + c)$ and let $g^k = Hx^k + c + A^T w^k$;

Compute Δx_g^k and α_g^k by (20) and (34); (with $D^k = |\tilde{X}^k|^{\frac{1}{2}}$)

If $\psi^k\ (\alpha_g^k \Delta x_g^k) \,/\, \psi^k\ (\Delta x_g^k) \leq \min\ (\tau_1,\ \tau_2\, t^k)$

$$D^k = (X^k)^{\frac{1}{2}};$$

end;

end.

(Compute w^k, g^k, Δx_g^k, and α_g^k if necessary.)

If $D^k = (X^k)^{\frac{1}{2}}$

Solve (19) for w^k and let $g^k = Hx^k + c + A^T w^k$;

Compute Δx_g^k and α_g^k by (20) and (34);

end.

2. Solve (9) for $\Delta \tilde{x}_{tr}^k$; Compute Δx_{tr}^k and α_{tr}^k by (8) and (31)

(Note: if $D^k = |\tilde{X}^k|^{\frac{1}{2}}$, then replace X^k by $|\tilde{X}^k|$ in the computation).

3. Determine s^k by (35); Update $x^{k+1} = x^k + s^k$.

We see that Algorithm Interior-Newton2 differs from Algorithm Interior-Newton only in that procedures to select and determine a trial D^k are added in Algorithm Interior-Newton2. In the Appendix, we verify that Algorithm Interior-Newton2 has the same strong convergence properties as Algorithm Interior-Newton. Here we make a few remarks before discussing our computational experiments.

1. As we show in the Appendix, due to the acceptance test in Step 1 for safeguarding, the convergence of $\{x^k\}$ is independent of the choice between $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ and $D^k = (X^k)^{\frac{1}{2}}$. However, a proper choice is important for improved practical performance: the rule we specify has done well in our numerical experiments.
2. We show in the Appendix that, ultimately, only a single solve of (19) is needed in each iteration; therefore, ultimately there is no extra cost for the acceptance test.
3. With minor modifications, both Algorithm Interior-Newton and Algorithm Interior-Newton2 can handle QP problems of the "general" form:

$$\begin{aligned} &\min \{q(x) = \tfrac{1}{2}x^T Hx + c^T x\} \\ &\text{subject to } \ Ax = b \ \text{ and } \ l \le x \le u, \end{aligned} \tag{89}$$

where $l \in \{\Re \cup \{-\infty\}\}^n$ and $u \in \{\Re \cup \{\infty\}\}^n$. For example, we may replace X^k by $diag\ (\min (x^k - l,\ u - x^k))$. We actually conducted our numerical experiments on problems of this more general form.

We have implemented our algorithms in Matlab and conducted some preliminary testing to investigate the practical viability of our approach. The experiments were performed on a Sun workstation. In the remainder of this section we present and discuss our preliminary numerical results.

Problem Generation: The Moré/Toraldo [29] QP-generator was adapted to generate the test problems. All the problems reported here are dense, and of moderate size. For positive definite problems the single global solution, with prescribed characteristics, was generated. For indefinite problems, the local minimizer determined by the algorithm(s) may have little relationship with the point generated (with prescribed characteristics) by the test problem generator: indefinite problems have many local minimizers.

Starting and Stopping: For each test problem, a feasible starting point was found using a feasible point determination algorithm [25]. This procedure does not involve the matrix H nor the vector c; therefore, the feasible starting point can be considered somewhat arbitrary with regard to problem (1).

As usual, choosing a robust stopping criterion is not easy. We have used stopping criteria based on three computations: the relative difference in the objective function values of two successive iterations, the size of α_{tr}, and the size of θ which is defined by (27). We terminate the iteration if:

$$\left\{ \begin{array}{c} q(x^k) - q(x^{k+1}) \le tol * (1 + |q(x^k)|) \\ \text{and} \\ \alpha_{tr}^k \ge 0.1, \end{array} \right. \tag{90}$$

or

$$\theta^k \le tol, \tag{91}$$

or

$$k = 100. \tag{92}$$

Criterion (91) is reasonable by Theorem 5. Criterion (90) suggests that no significant progress can be made. We introduced $\alpha_{tr}^k \geq 0.1$ since we did not want to stop the iteration if (90) was caused by a very small step length. In our experiments, $tol = 10^{-12}$. This stopping criterion was successful for our experiments since for every problem we tested, the second-order optimality conditions were confirmed.

Computation and Parameter Setting: Our implementation is straightforward: i.e., we solve (19) for w^k, and at the same time, $\bar{Z}^k$ is obtained. We compute Δx_g^k by (20), and compute Δx_{tr}^k by (8), etc..

The parameters are set as follows:

$$\tau_\rho = 0.8; \quad \tau_\alpha = 1.9; \quad \tau_1 = 0.001; \quad \tau_2 = \tau_3 = 0.5, \tag{93}$$

where τ_ρ is used in (30) and (33), τ_α is used in (31) and (34), τ_1, τ_2, and τ_3 are used in Algorithm Interior-Newton$_2$ for selecting the trial D^k and for the acceptance test.

We adjust δ^k as follows:

Updating δ^k

Let $0 < \tau_4 < \tau_5 < 1$ and $0 < \tau_6 < 1 < \tau_7$ be given.

$\delta^0 = 1$;

For $k = 0, 1, 2, \cdots$

$\alpha^k = \sigma^k \alpha_{tr}^k + (1 - \sigma^k)\alpha_g^k$;

If $\alpha^k \leq \tau_4$

$\delta^{k+1} = \max\ (\delta_l,\ \tau_6 \delta^k)$;

end.

If $\alpha^k \geq \tau_5$

$\delta^{k+1} = \min\ (\delta_u,\ \tau_7 \delta^k)$;

end.

The corresponding parameters are set as follows:

$$\tau_4 = 0.5; \quad \tau_5 = 0.9; \quad \tau_6 = 0.75; \quad \tau_7 = 1.25.$$

Experimental Results: We tested four groups of problems using Algorithm Interior-Newton$_2$ and the results are tabulated in Tables 1 – 4. For each case, three problems were attempted. The table entries are the numbers of iterations required to satisfy the stopping criterion. Positive definite problems are reported in Tables 1 and 2; indefinite problems are reported in Tables 3 and 4. For each problem in Tables 3 and 4, about 10% of the eigenvalues of H are negative.

We conducted our experiments on problems of the general form (89). The lower bounds were all set to zero. The upper bounds were all set to unity in Table 1 and 3 while in Table 2 and 4, about 10% were set to infinity (that is what $pctinf = 0.1$ means).

For all the problems, about $0.8 * (n - m)$ components of the generated x were set to the bounds and $cond(H) = cond(A) = cond$.

We did not encounter any instance where $q(x^k) \longrightarrow -\infty$; however, for our generated indefinite problems, with some infinite bounds, there is no guarantee that q is bounded below in the feasible region.

Very few iterations needed more than a single QR factorization (see Theorem 13). Over all 3806 iterations (for the 216 test problems), only 18 extra QR factorizations were used. On average, $D^k = (X^k)^{\frac{1}{2}}$ was used about twice per problem (remember that D^0 is always set to be $(X^0)^{\frac{1}{2}}$). The average cost for finding a feasible starting point was 4.30 linear system solves.

Table 1. Positive Definite Problems, pctinf = 0

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
100	10	18	17	18	16.3	20	17
100	50	16	13	17	15.3	23	17.3
100	90	12	10.3	13	11.3	18	16.7

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
200	20	17	16	21	18.7	23	21
200	100	20	18.3	22	19.3	22	19.3
200	180	15	13.3	16	13.3	17	15.3

Table 2. Positive Definite Problems, pctinf = 0.1

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
100	10	18	16.3	17	16	18	16.3
100	50	15	12.7	18	16	22	16
100	90	13	11	14	11.7	16	14

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
200	20	17	16	22	20	19	18.3
200	100	21	18.3	21	19.7	22	19
200	180	15	13.7	14	13	24	17.3

Observations: On the whole, the experiments indicate that algorithm Interior-Newton$_2$ is efficient. The iteration numbers are insensitive to problem size and condition number. Though our experiments are limited, they clearly indicate that the new algorithm is promising.

Table 3. Indefinite Problems, pctinf = 0

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
100	10	20	18.3	26	20.3	29	23.3
100	50	20	19.3	23	20	17	16
100	90	13	11.3	13	11.7	18	14.7

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
200	20	21	20.3	26	24.7	32	28.7
200	100	25	21.7	28	23.7	23	21.3
200	180	17	15.7	14	12.7	17	16

Table 4. Indefinite Problems, pctinf = 0.1

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
100	10	19	17.7	26	20.3	34	24.3
100	50	20	18.3	24	21.3	19	17.3
100	90	13	11.3	14	12.3	17	14.3

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
200	20	39	26	28	25.3	30	30
200	100	27	22	29	23.3	25	21.7
200	180	16	15.3	19	15.7	25	18

For purposes of comparison, we have used Algorithm Interior-Newton to solve the same set of problems as in Table 1, and the results are presented in the following table. In most cases both algorithms terminated at (almost) the same function values for each problem; however, the required numbers of iterations are quite different. On average, the number of iterations is 16.1 in Table 1 but is 31.0 in Table 1′. The largest number is 23 in Table 1 but is 89 in Table 1′: the modification in Algorithm Interior-Newton$_2$ improves the computational performance significantly.

6. Concluding remarks

We have proposed an interior Newton method with a new scaling strategy for general quadratic programming problems. The algorithm is robust and has stronger convergence properties than existing interior methods for the general quadratic programming problem. Specifically, the main theoretical property of our proposed method can be summarized as follows. Under compactness and nondegeneracy assumptions, i.e., **(A1)**, **(A2)**, and **(A3)**, our proposed algorithm generates a sequence $\{x^k\}$ converging to a point

Table 1′. Positive Definite Problems, pctinf = 0

($D^k = (X^k)^{\frac{1}{2}}$ for every k)

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
100	10	20	18	27	20.7	35	27
100	50	60	31.3	69	36.7	28	19
100	90	53	34.3	21	15.7	89	58

size		cond=10^3		cond=10^6		cond=10^9	
n	m	max	avg	max	avg	max	avg
200	20	20	19	43	27.3	40	31
200	100	59	50.3	56	35.7	59	35
200	180	53	44	56	28.7	40	25.7

x^* satisfying the second-order necessary conditions. Moreover, if x^* satisfies second-order sufficiency conditions, then the local rate of convergence is 2-step quadratic.

It is noted that the strong convergence properties hold under a more general algorithmic framework: we do not have to compute w^k, s^k, etc. exactly as stated. Rather, only the following conditions need be satisfied:

(i) The sequence $\{w^k\}$ satisfies the following: if any subsequence $\{x^{k_j}\}$ converges to a point x^* satisfying the complementarity condition (3), then $w^{k_j} \longrightarrow w^* = -(AX^*A^T)^{-1}AX^*(Hx^*+c)$.

(ii) The sequence $\{s^k\}$ satisfies that for a given $\gamma > 0$ and for every k sufficiently large,
(1) $\psi^k(s^k) \le \gamma \min\left(\psi^k(\alpha_g^k \Delta x_g^k),\ \psi^k(\alpha^k \Delta x_{tr}^k)\right)$.
(2) $\|(D^k)^{-1}s^k\| \le \gamma\,\delta^k$.

Proofs under these more general conditions are provided in [9].

Preliminary numerical experiments suggest that the method is efficient for dense problems of moderate size. Inspection of the conditioning of the underlying linear systems, in the limit, indicates that robust asymptotic behavior is to be expected from this approach, even in the presence of near-degeneracy and ill-conditioning. This is in sharp contrast to certain formulations of interior methods which are inherently afflicted by ill-conditioning. To further explore this remark, consider the following.

Suppose **(A4)** holds. Then for every k sufficiently large, the resulting coefficient matrix of our algorithm can be formulated as

$$(\bar{Z}^k)^T \bar{M}^k \bar{Z}^k = (\bar{Z}^k)^T (D^k H D^k + |G^k|)\bar{Z}^k. \tag{94}$$

Alternatively, the underlying system can be solved using the coefficient matrix

$$\begin{bmatrix} D^k H D^k + |G^k| & (AD^k)^T \\ AD^k & 0 \end{bmatrix}. \tag{95}$$

The limit matrix of (95),

$$\begin{bmatrix} D^* H D^* + |G^*| & (AD^*)^T \\ AD^* & 0 \end{bmatrix} \in \Re^{(m+n)\times(m+n)}, \tag{96}$$

is well-behaved under **(A3)** since there exists a permutation matrix P such that

$$P^T \begin{bmatrix} D^*HD^* + |G^*| & (AD^*)^T \\ AD^* & 0 \end{bmatrix} P = \begin{bmatrix} diag(|g^*_{\mathcal{A}^*}|) & 0 \\ 0 & B \end{bmatrix} \tag{97}$$

where $B \in \Re^{(m+n-|\mathcal{A}^*|)\times(m+n-|\mathcal{A}^*|)}$ is some nonsingular matrix and $g^*_{\mathcal{A}^*}$ is a sub-vector of g^* containing those elements of g^*_i with $i \in \mathcal{A}(x^*)$. The limit matrix of (94) is

$$(\bar{Z}^*)^T(D^*HD^* + |G^*|)\bar{Z}^*. \tag{98}$$

Any possible ill-conditioning of matrix (95) can only come from the ill-conditioning of H, the rank deficiency of A, and the degeneracy of the original problem (1).

The interior approach we have explored in this paper is an exciting new way to solve quadratic programming problems: it is theoretically interesting and practically viable. As it stands it is not suitable for large-scale quadratic programming in that it requires full-dimensional trust region computations and several matrix factorizations in each major iteration. In response to this we are currently investigating a modification of this approach involving iterative and approximate linear solvers. This will be the topic of a future report.

Acknowledgements. We thank our colleague Yuying Li for many helpful remarks on this work. We thank Yinyu Ye for reading a preliminary draft of this report and making many useful comments. We also thank the editor, associate editor, and referees for many valuable suggestions that significantly improved the presentation of this paper.

Appendix: The convergence properties of Algorithm Interior-Newton$_2$

In this appendix, we verify that Algorithm Interior-Newton$_2$ has the same convergence properties as Algorithm Interior-Newton. First, we give a simple general result; the proof is straightforward and so it is omitted.

Lemma 17. *Suppose $H \geq 0$ is any symmetric matrix. Then*

$$\|Hx\|^2 \leq x^T Hx \, \|H\|.$$

The following set will be useful subsequently:

$$\mathcal{K} := \{k : D^k = (X^k)^{\frac{1}{2}} \text{ at the end of Step 1 in Interior-Newton}_2\}.$$

Let $\mathcal{K}^c$ denote the complementary set of $\mathcal{K}$. It is clear that $k \in \mathcal{K}^c$ if and only if $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ at the end of Step 1.

Theorem 10. *Suppose $\{k_j\} \subset \mathcal{K}^c$ and $x^{k_j} \longrightarrow x^*$. Then (49) holds with x^* and*

$$\tilde{x}_i^{k_j} = x_i^{k_j} \text{ for every } i \in \mathcal{A}(x^*) \text{ and for every } k_j \text{ sufficiently large,} \tag{99}$$

where $\tilde{x}^k$ is defined by (28).

Proof. First, assume that zero is a limit point of $\{t^{k_j}\}$, where t^k appears in Step 1 for the acceptance test. Without loss of generality, assume $t^{k_j} \longrightarrow 0$. Then

$$\tilde{X}^{k_j}(Hx^{k_j} + c + A^T w^{k_j-1}) \longrightarrow 0.$$

Since $\{\|\tilde{x}^k\|\}$ and $\{\|w^k\|\}$ are bounded, let $(x^\#,\ w^\#)$ be a limit point of $\{\ (\tilde{x}^{k_j},\ w^{k_j-1})\}$. Then

$$X^\#(Hx^* + c + A^T w^\#) = 0.$$

If any $(Hx^* + c + A^T w^\#)_i \neq 0$, then $x_i^\# = 0$ which implies $x_i^{k_{j_l}} = \tilde{x}_i^{k_{j_l}} \longrightarrow 0$ for some subsequence $\{k_{j_l}\}$ of $\{k_j\}$ and consequently $x_i^* = 0$. Therefore,

$$X^*(Hx^* + c + A^T w^\#) = 0,$$

which implies that $w^\# = w^*$ by **(A2)**, where w^* is defined in (50). So in this case, $w^{k_j-1} \longrightarrow w^*$ and (49) holds.

To show (99), let $i \in \mathcal{A}(x^*)$. By **(A3)**, $(Hx^* + c + A^T w^*)_i \neq 0$. Then $\tilde{x}_i^{k_j} \longrightarrow 0$, which implies that $\tilde{x}_i^{k_j} = x_i^{k_j}$ for every k_j sufficiently large.

Now assume that $\{t^{k_j}\}$ is bounded away from zero. Then for some $\epsilon > 0$,

$$t^{k_j} \geq \epsilon \text{ for every } k_j.$$

It is clear that (43) holds and it implies that $\{q(x^k)\}$ converges. Also, (39) is true and by the acceptance test in Step 1, noting that $\{k_j\} \subset \mathcal{K}^c$, we have

$$\begin{aligned} q(x^{k_j}) - q(x^{k_j+1}) &\geq -\gamma\psi^{k_j}\ (\alpha_g^{k_j}\Delta x_g^{k_j}) \\ &\geq \gamma \min\ (\tau_1, \tau_2\epsilon)\left(-\psi^{k_j}\ (\Delta x_g^{k_j})\right) \longrightarrow 0. \end{aligned}$$

Therefore, by (40) and letting $y^k = \frac{\bar{g}^k}{\|\bar{g}^k\|}$, we have

$$\left((y^{k_j})^T \bar{M}^{k_j} y^{k_j} + \lambda_g^{k_j}\right)\ (\mu^{k_j})^2 \longrightarrow 0.$$

Using (23), we have

$$\bar{g}^{k_j} = D^{k_j}(Hx^{k_j} + c + A^T w^{k_j}) \longrightarrow 0,$$

where $D^{k_j} = |\tilde{X}^{k_j}|^{\frac{1}{2}}$. Repeating the arguments above for the case $t^{k_j} \longrightarrow 0$, replacing $\{w^{k_j-1}\}$ by $\{w^{k_j}\}$, we see the lemma is true.

□

Corollary 2. *Let x^* be any limit point of $\{x^k\}$. Then (49) holds with x^*.*

Proof. Assume $x^{k_j} \longrightarrow x^*$. If $k_j \in \mathcal{K}$, then s^{k_j} is computed based on $D^{k_j} = (X^{k_j})^{\frac{1}{2}}$, and otherwise, based on $D^{k_j} = |\tilde{X}^{k_j}|^{\frac{1}{2}}$. So if there are infinitely many $k_j \in \mathcal{K}^c$, then by Theorem 10, (49) holds with x^*. Otherwise, by Corollary 1, (49) holds with x^* also.

□

Since Lemma 2 and Lemma 3 are independent of the choice of D^k, we have a result similar to Theorem 2 as follows.

Theorem 11. *The sequence $\{x^k\}$ generated by Algorithm Interior-Newton$_2$ converges to x^* satisfying (51).*

Proof. The proof is almost the same as that of Theorem 2. First, repeat the part from the beginning of the proof for Theorem 2 to (53). Then by Theorem 10, we have $\tilde{x}_i^k = x_i^k$ for every $i \in \mathcal{A}(x^*)$ and for every $k \in \mathcal{K}^c$ sufficiently large. Therefore,

$$|s_i^k| \leq 2\delta_u \, (x_i^k)^{\frac{1}{2}} \text{ for every } k \text{ sufficiently large and for every } i \in \mathcal{A}(x^*).$$

Then we may repeat from the line after (55) to the end of the proof of Theorem 2.

□

Note that the convergence of $\{x^k\}$ does not depend on whether $D^k = (X^k)^{\frac{1}{2}}$ or $D^k = |\tilde{X}^k|^{\frac{1}{2}}$ for the trial D^k in Step 1. This flexibility gives us freedom to choose D^k with performance in mind.

The next result establishes the convergence of $\{w^k\}$.

Lemma 18. *The sequence $\{w^k\}$ converges to w^* defined in (50).*

Proof. Let $w^\#$ be any limit point of $\{w^k\}$ and assume $\{w^{k_j}\} \longrightarrow w^\#$. If there are infinitely many $k_j \in \mathcal{K}$, then clearly $w^\# = w^*$. Otherwise, for every k_j sufficiently large, $k_j \in \mathcal{K}^c$ and

$$w^{k_j} = -(A|\tilde{X}^{k_j}|A^T)^{-1}A|\tilde{X}^{k_j}|(Hx^{k_j} + c), \tag{100}$$

where the existence of $(A|\tilde{X}^{k_j}|A^T)^{-1}$ comes from **(A2)**.

Let $x^\#$ be any limit point of $\{\tilde{x}^{k_j}\}$. Then $w^\# = -(AX^\# A^T)^{-1}AX^\#(Hx^* + c)$. By Theorem 10 and **(A3)**, $\tilde{x}_i^{k_j} = x_i^{k_j}$ for every $i \in \mathcal{A}(x^*)$ and for every $k_j \in \mathcal{K}^c$ sufficiently large. Therefore, $X^\#(Hx^* + c + A^T w^*) = 0$ which implies

$$w^* = -(AX^\# A^T)^{-1}AX^\#(Hx^* + c) = w^\#.$$

□

It is easy to verify that all the remaining results in Sects. 3 and 4 for Interior-Newton also hold for Interior-Newton$_2$. To sum them up, we have the following result:

Theorem 12. *Under assumptions* **(A1)**, **(A2)**, *and* **(A3)**, *the sequence $\{x^k\}$ generated by Algorithm Interior-Newton$_2$ converges (globally) to a point satisfying the second-order necessary conditions (2) – (5). Further more, if assuming* **(A4)**, *then the rate of convergence is at least 2-step quadratic.*

Our final result in this appendix is to show that, ultimately, only one computation is needed for solving (19), etc.. Clearly, this is true if we can establish that in the acceptance test,

$$\psi^k \, (\alpha_g^k \Delta x_g^k) \,/\, \psi^k \, (\Delta x_g^k) > \min \; (\tau_1, \, \tau_2 \, t^k) \text{ for every } k \text{ sufficiently large.} \tag{101}$$

Theorem 13. *The inequality (101) holds under assumption* **(A4)**.

Proof. First, we show the following:

$$\text{there exists } \bar{\alpha} > 0 \text{ such that } \alpha_g^k \geq \bar{\alpha} \text{ for every } k \text{ sufficiently large.} \qquad (102)$$

In fact, by Theorem 10, whether $D^k = (X^k)^{\frac{1}{2}}$ or $D^k = |\tilde{X}^k|^{\frac{1}{2}}$, we always have

$$D_{ii}^k = (x_i^k)^{\frac{1}{2}} \quad \text{for every } k \text{ sufficiently large and for every } i \in \mathcal{A}(x^*).$$

If (102) is false, then repeating the arguments between (47) and (48), noting that $\{\|g^{k_j}\|\}$ is bounded, we would have

$$\frac{\|\bar{g}^{k_j}\|}{|\mu_g^{k_j}|} \longrightarrow 0. \qquad (103)$$

But on the other hand, since $\bar{A}^k \bar{g}^k = 0$, there exists $\xi^k \in \Re^{n-m}$ such that $\bar{g}^k = \bar{Z}^k \xi^k$. Hence

$$\frac{(\bar{g}^k)^T}{\|\bar{g}^k\|} \bar{M}^k \frac{\bar{g}^k}{\|\bar{g}^k\|} = \frac{(\xi^k)^T \, (\bar{Z}^k)^T \bar{M}^k \bar{Z}^k \xi^k}{\|\xi^k\|^2},$$

By Lemma 9, $(\bar{Z}^*)^T \bar{M}^* \bar{Z}^* > 0$. So there exists $\epsilon > 0$ such that

$$\frac{(\bar{g}^k)^T}{\|\bar{g}^k\|} \bar{M}^k \frac{\bar{g}^k}{\|\bar{g}^k\|} \geq \epsilon \quad \text{for every } k \text{ sufficiently large.}$$

Then, using (45), (25) and the fact that $\bar{g}^k \longrightarrow 0$, we have $\mu_g^k \longrightarrow 0$ and $\lambda_g^k \longrightarrow 0$. Using (45) again, we have

$$\frac{\|\bar{g}^{k_j}\|}{|\mu_g^{k_j}|} \geq \frac{\epsilon}{2} \quad \text{for every } k \text{ sufficiently large,}$$

which is a contradiction to (103). Therefore, (102) is true.

Now, to show (101), it suffices to show that $\{\psi^k \, (\alpha_g^k \Delta x_g^k)/\psi^k \, (\Delta x_g^k)\}$ is bounded away from zero since $t^k \longrightarrow 0$. This is true because it is easy to see that $\lambda_g^k = 0$ for every k sufficiently large. Therefore, by (41), (23), and (20),

$$\frac{\psi^k \, (\alpha_g^k \Delta x_g^k)}{\psi^k \, (\Delta x_g^k)} = \alpha_g^k (2 - \alpha_g^k) \quad \text{for every } k \text{ sufficiently large,}$$

which is bounded away from zero by (102) and the fact that $\alpha_g^k \leq \tau_\alpha < 2$.

□

References

1. Barnes, E.R. (1986): A variation on Karmarkar's algorithm for solving linear programming problems. Math. Program. **36**, 174–182
2. Ben Daya, M., Shetty, C.M. (1988): Polynomial barrier function algorithm for convex quadratic programming. Report J 88-5, School of ISE, Georgia Institute of Technology, Atlanta, GA
3. Beale, E.M.L. (1959): On quadratic programming. Naval Research Logistics Quarterly **6**, 227–243
4. Byrd, R.H., Schnabel, R.B. (1986): Continuity of the null space and constrained optimization. Math. Program. **35**, 32–41
5. Calamai, P.H., Moré, J.J. (1987): Projected gradient methods for linearly constrained problems. Math. Program. **39**, 93–116
6. Coleman, T.F., Hulbert, L.A. (1989): A direct active set algorithm for large sparse quadratic programs with simple bounds. Math. Program. **45**, 373–406
7. Coleman, T.F., Li, Y. (1994): On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. Math. Program. **67**, 189–224
8. Coleman, T.F., Li, Y. (1996): An interior trust region approach for nonlinear minimization subject to bounds. SIAM J. Optim. **6**, 418–445
9. Coleman, T.F., Liu, J. (1993): An interior Newton method for quadratic programming. Tech. Rep. 93-1388, Computer Science Department, Cornell University, Ithaca, NY
10. Coleman, T.F., Sorensen, D. (1984): A note on the computation of an orthonormal basis for the null space of a matrix. Math. Program. **29**, 234–242
11. Cottle, R.W., Dantzig, G.B. (1968): Complementary pivot theory of mathematical programming. Linear Algebra Appl. **1**, 103–125
12. Dennis, Jr., J.E., Schnabel, R.B. (1983): Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, NJ
13. Dikin, I.I. (1967): Iterative solution of problems of linear and quadratic programming. Sov. Math. Dokl. **8**, 674–675
14. Fletcher, R. (1971): A general quadratic programming algorithm. J. Inst. Math. Appl. **7**, 76–91
15. Gay, D.M. (1981): Computing optimal locally constrained steps. SIAM J. Sci. Statist. Comput. **2**, 186–197
16. Gill, P.E., Murray, W. (1974): Newton type methods for unconstrained and linearly constrained optimization. Math. Program. **7**, 311–350
17. Gill, P.E., Murray, W., Saunders, M.A., Stewart, G.W., Wright, M.H. (1985): Properties of representation of a basis for the null space. Math. Program. **33**, 172–186
18. Gill, P.E., Murray, W., Saunders, M.A., Tomlin, J.A., Wright, M.H. (1986): On projected Newton methods for linear programming and equivalence to Karmarkar's projective method. Math. Program. **36**, 183–209
19. Gill, P.E., Murray, W., Wright, M.H. (1981): Practical Optimization. Academic Press, London
20. Goldfarb, D. (1972): Extensions of Newton's method and simplex methods for solving quadratic programs. in: Lootsman, F.A., ed., Numerical Methods for nonlinear Optimization. Academic Press, London
21. Goldfarb, D., Liu, S. (1991): An $O(n^3 L)$ primal interior point algorithm for convex quadratic programming. Math. Program. **49**, 325–340
22. Kapoor, S., Vaidya, P. (1986): Fast algorithm for convex quadratic programming and multicommodity flows. Proceedings of the 18th Annual ACM Symposium on Theory Computing 147–159
23. Karmarkar, N. (1984): A new polynomial time algorithm for linear programming. Combinatorica **4**, 373–395
24. Kojima, M., Mizuno, S., Yoshise, A. (1989): A polynomial-time algorithm for a class of linear complementarity problems. Math. Program. **44**, 1–26
25. Liu, J. (1994): Interior and Exterior Newton Methods for Large-scale Quadratic Programming. Ph.D. Thesis, Center for Applied Mathematics, Cornell University, Ithaca, NY
26. Mehrotra, S., Sun, J. (1990): An algorithm for convex quadratic programming that requires $O(n^{3.5}L)$ arithmetic operations. Math. Oper. Res. **15**, 342–363
27. Monteiro, R., Adler, I. (1989): Interior path-following primal-dual algorithms, part II: Convex quadratic programming. Math. Program. **44**, 43–66
28. Moré, J.J., Sorensen, D. (1983): Computing a trust region step. SIAM J. Sci. Statist. Comput. **4**, 553–572
29. Moré, J.J., Toraldo, G. (1989): Algorithms for bound constrained quadratic programming problems. Numer. Math. **55**, 377–400
30. Pardalos, P.M., Ye, Y., Han, C.G. (1990): An interior point algorithm for large–scale quadratic problems with box constraints. In: Bensoussan, A., Lions, J.L., eds., Analysis and Optimization of Systems: Proceedings of the 9th International Conference. Lect. Notes Control Inform. Sci. **144**, pp. 413–422. Springer, Berlin
31. Ponceleón, D.B. (1990): Barrier methods for large-scale quadratic programming. Ph.D. Thesis. Department of Computer Science, Stanford University, Stanford, CA

32. Powell, M.J.D. (1970): A new algorithm for unconstrained optimization. In: Rosen, J.B., Mangasarian, O.L., Ritter, K., eds., Nonlinear Programming, pp. 31–65. Academic Press, New York
33. Schultz, G.A., Schnabel, R.B., Byrd, R.H. (1985): A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties. SIAM J. Numer. Analysis **22**, 47–67
34. Sorensen, D. (1982): Trust region methods for unconstrained optimization. SIAM J. Numer. Analysis **19**, 409–426
35. Vanderbei, R.J., Meketon, M.S., Freedman, B.A. (1986): A modification of Karmarkar's linear programming algorithm. Algorithmica **1**, 395–407
36. Ye, Y. (1989): An extension of Karmarkar's algorithm and the trust region method for quadratic programming. In: Megiddo, N., ed., Progress in Mathametical Programming. Springer, New York
37. Ye, Y. (1992): On affine scaling algorithms for nonconvex quadratic programming. Math. Program. **56**, 285–300
38. Ye, Y., Tse, E. (1989): An extension of Karmarkar's projective algorithm for convex quadratic programming. Math. Program. **44**, 157–179
39. Zhang, Y. (1994): On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. SIAM J. Optim. **4**, 208–227