

A TRUST REGION INTERIOR POINT ALGORITHM FOR LINEARLY CONSTRAINED OPTIMIZATION*

J. FRÉDÉRIC BONNANS[†] AND CECILIA POLA[‡]

Abstract. We present an extension, for nonlinear optimization under linear constraints, of an algorithm for quadratic programming using a trust region idea introduced by Ye and Tse [Math. Programming, 44 (1989), pp. 157–179] and extended by Bonnans and Bouhtou [RAIRO Rech. Opér., 29 (1995), pp. 195–217]. Due to the nonlinearity of the cost, we use a linesearch in order to reduce the step if necessary. We prove that, under suitable hypotheses, the algorithm converges to a point satisfying the first-order optimality system, and we analyze under which conditions the unit stepsize will be asymptotically accepted.

Key words. trust region, quadratic model, linesearch, interior points

AMS subject classifications. 90C30, 65K05, 49M40

PII. S1052623493250639

1. Introduction. In this paper, we study an algorithm for minimizing a nonlinear cost under linear constraints. We consider problems with linear equality constraints and nonnegative variables. At each step, a direction is computed by minimizing a convex quadratic model over an ellipsoidal trust region, and then a linesearch of Armijo type is performed in this direction. At each iteration, the ellipsoid of the quadratic problem is so small that it forces the nonnegativity constraints to be satisfied. However, the ellipsoid is not necessarily contained in the set of feasible points.

In the case of linear programming (LP) or convex quadratic programming (QP), we may assume the quadratic model to be equal to the cost function. Then the unit step will be accepted by the linesearch. In the case of LP, the algorithm is then reduced to the celebrated Dikin's algorithm [10] (see also Tsuchiya [26]). Ye and Tse [27] have extended this algorithm to convex quadratic programming using the trust region idea. This problem was also considered by Sun [25]. Bonnans and Bouhtou [2] studied such methods for nonconvex quadratic problems by taking a variable size for the trust region. An early extension of trust region algorithms to nonlinear costs is done in Dikin and Zorkalcev [11]. Among the related work, we quote Gonzaga and Carlos [13]. Interior point algorithms for the solution of constrained convex optimization problems have been studied by many other researchers; see, for instance, Den Hertog, Roos, and Terlaky [8], Jarre [15], Mehrotra and Sun [19], McCormick [18], Monteiro and Adler [20], Dennis, Heinkenschloss, and Vicente [9], and Coleman and Li [7]. Gonzaga [14] explores the shape of the trust regions to generate ellipsoidal regions adapted to the shape of the feasible set. The resulting algorithm generates sequences of points in the interior of the feasible set.

In this paper, we obtain some results of global convergence, comparable to those obtained in [2] for QP; by global convergence we only mean that the limit points of the sequence generated by the algorithm satisfy the first-order optimality system. The main novelty of the paper, however, is in the local analysis in the vicinity of

* Received by the editors June 17, 1993; accepted for publication (in revised form) January 19, 1996. This research was partially supported by DGICYT (Madrid).

<http://www.siam.org/journals/siopt/7-3/25063.html>

[†] INRIA, Domaine de Voluceau, 78153 Rocquencourt, France (frederic.bonnans@inria.fr).

[‡] Dpto Mat., Estad. y Comp.-Facultad de Ciencias, 39005 Santander, Spain (pola@matsun3.unican.es).

a local solution satisfying some strong second-order sufficient conditions. We check that if such a point is a limit point of the sequence computed by the algorithm and is under a “sufficient curvature” condition satisfied by the Hessian of the quadratic approximation, then the sequence actually converges to this point and the unit step is asymptotically accepted. Unfortunately, the acceptance of the unit step is not by itself a guarantee of a rapid convergence (the convergence might be linear at a very poor rate). The interest of the result lies in the fact that in the case of convex QP, this type of algorithm converges reasonably well in practice, although the convergence rate is only linear (see, e.g., the numerical results reported in Bonnans and Bouhtou [2] and Bouhtou [5]). Therefore, the question is to know to which extent the features of Dikin’s type algorithms may be kept when dealing with nonlinear cost functions. In particular, we do not expect the rate of convergence of the cost to be superlinear, as this is not the case for quadratic programs.

The paper is organized as follows. In section 2 we present the algorithm and give a result of global convergence in the sense that under some convenient hypotheses, the sequence computed by the algorithm converges towards a point satisfying the first-order optimality system. Then, in section 3 we perform the local analysis: we check that if the sequence computed by the algorithm has some regular limit point \bar{x} and if a condition of “sufficient curvature” holds, then the sequence converges to this point and the unit step is asymptotically accepted.

2. The algorithm. We consider the following problem:

$$(P) \quad \min f(x); Ax = b; x \geq 0,$$

where f is a smooth mapping from \mathbb{R}^n in \mathbb{R} , not necessarily convex; A is a $p \times n$ matrix; and $b \in \mathbb{R}^p$. We define the following sets:

$$F := \{x \in \mathbb{R}^n; Ax = b, x \geq 0\},$$

$$\overset{\circ}{F} := \{x \in \mathbb{R}^n; Ax = b; x > 0\},$$

so that F is the set of feasible points and $\overset{\circ}{F}$ is the set of “strictly feasible” points. In the sequel, we assume that F is bounded and $\overset{\circ}{F}$ is nonempty.

The algorithm will use two matrices at each iteration. The first is $X_k := \text{diag}(x^k)$, where $\{x^k\}$ is the current feasible point. This is a scaling matrix that takes care of the positivity constraints. The second matrix is M_k , a symmetric approximation of the Hessian of the cost function. We assume M_k to be positive semidefinite (i.e., $d^t M_k d \geq 0$ for all d in \mathbb{R}^n). We consider the following algorithm.

ALGORITHM 1.

- 0) Choose $x^\circ \in \overset{\circ}{F}$, $\delta \in (0, 1)$, $\beta \in (0, 1)$, $\gamma \in (0, 1)$; $k \leftarrow 0$.
- 1) Choose an $n \times n$ symmetric matrix M_k . Compute δ_k in $(\delta, 1/\delta)$ such that the point d^k that solves

$$(SP) \quad \min_d \varphi_k(d) := f(x^k) + \nabla f(x^k)^t d + \frac{1}{2} d^t M_k d; \quad Ad = 0; \quad d^t X_k^{-2} d \leq \delta_k^2$$

satisfies $x^k + d^k > 0$.

- 2) If $\varphi_k(d^k) = f(x^k)$, stop.
 3) *Linesearch*: Compute $\rho^k = \beta^{\ell_k}$, with ℓ_k the smallest nonnegative integer such that

$$(1) \quad f(x^k) - f(x^k + \beta^{\ell_k} d^k) \geq \gamma \beta^{\ell_k} (f(x^k) - \varphi_k(d^k)).$$

- 4) $x^{k+1} = x^k + \rho^k d^k$; $k \leftarrow k + 1$. Go to 1.

Some comments are needed to clarify the description of the algorithm. First, let us note that the stopping criterion of step 2 is, of course, unrealistic. The algorithm will typically never stop. This is convenient for studying the asymptotic properties of the sequence generated by the algorithm. A practical stopping criterion might require that we stop when $\varphi_k(d^k)$ is close enough to $f(x^k)$. Because the cost function may be nonconvex, there is, of course, no guarantee that the limit points are close to a global or even local solution (our results below deal with the optimality system at the limit points).

Our second comment deals with the fact that we allow δ_k to be greater or equal to 1. If we specify a value of δ_k smaller than 1, then we automatically have $x^k + d^k > 0$. What is the meaning of allowing $\delta_k \geq 1$? In order to understand that, let us observe that the trust region problem (SP) cannot be solved directly because of the nonlinear constraint (see, e.g., Moré [21], Sorensen [24]). Instead, one typically solves a sequence of equality constrained quadratic problems of type

$$\min_d f(x^k) + \nabla f(x^k)^t d + \frac{1}{2} d^t M_k d + \frac{\nu}{2} d^t X_k^{-2} d; \quad Ad = 0,$$

where $\nu \geq 0$ is an estimate of ν_k (the Lagrange multiplier associated with the nonlinear constraint). As M_k is semidefinite positive for any $\nu > 0$, this problem has a unique solution $d = d(\nu)$, and the mapping $\nu \rightarrow d^t(\nu) X_k^{-2} d(\nu)$ is strictly decreasing. Let us say that ν is *too small* if either $d^t(\nu) X_k^{-2} d(\nu) > 1/\delta$ or $\min_i x_i^k + d_i(\nu) < 0$ and *too large* if $d^t(\nu) X_k^{-2} d(\nu) < \delta$ (as $\delta < 1$, it follows in that case that $\min_i x_i^k + d_i(\nu) > 0$). A dichotomic procedure based on these notions of “too large” and “too small” allows us to compute a solution of (SP) with $\delta_k \in (\delta, 1/\delta)$ in a finite number of steps; this is associated with a value of δ_k that may be greater than 1. It was observed already in [2] that to allow the possibility that $\delta_k \geq 1$ may speed up the convergence, and therefore it is worth taking this possibility into account in the analysis.

We note that if the algorithm stops at iteration k , then x^k satisfies the first-order optimality condition of (P). To see this, we need the following lemma, which states the optimality system of (SP). This is a simple extension of the known result for problems without equality constraints; see [6].

LEMMA 2.1. *The point d^k that solves (SP) is characterized by the existence of λ^{k+1} in \mathbb{R}^p , $\nu_k \geq 0$ such that*

$$(2) \quad \nabla f(x^k) + M_k d^k + A^t \lambda^{k+1} + \nu_k X_k^{-2} d^k = 0,$$

$$(3) \quad Ad^k = 0,$$

$$(4) \quad \nu_k \geq 0, \quad (d^k)^t X_k^{-2} d^k \leq \delta_k^2, \quad \nu_k [(d^k)^t X_k^{-2} d^k - \delta_k^2] = 0.$$

We now come back to the discussion of step 2 of the algorithm. Using (2), we deduce that

$$\begin{aligned} f(x^k) - \varphi_k(d^k) &= -\nabla f(x^k)^t d^k - \frac{1}{2}(d^k)^t M_k d^k, \\ &= (\lambda^{k+1})^t A d^k + \nu_k (d^k)^t X_k^{-2} d^k + \frac{1}{2}(d^k)^t M_k d^k. \end{aligned}$$

Using (3) and (4), we get

$$(5) \quad f(x^k) - \varphi_k(d^k) = \nu_k \delta_k^2 + \frac{1}{2}(d^k)^t M_k d^k.$$

So, if $f(x^k) = \varphi_k(d^k)$, as M_k is a positive semidefinite matrix, then each of the nonnegative terms on the right-hand side is equal to 0. We deduce that $\nu_k = 0$ and $M_k^{1/2} d^k = 0$, so $M_k d^k = 0$, where $(M_k)^{1/2}$ is the square root of the symmetric positive semidefinite matrix M_k . That is,

$$(M_k)^{1/2} = \sum_{i=1}^n (\lambda_i)^{1/2} u^i (u^i)^t,$$

where $\{\lambda_i, u^i\}$, $i = 1$ to n , are the eigenvalues and associated orthonormal eigenvectors of M_k . Hence, again using (2), we get

$$\begin{aligned} \nabla f(x^k) + A^t \lambda^{k+1} &= 0, \\ Ax^k &= b, \quad x^k > 0. \end{aligned}$$

So, x^k satisfies the first-order optimality condition of (P).

In the sequel, when studying the convergence of the algorithm, we will assume that it generates an infinite sequence of iterates.

Remark 2.1. From Lemma 2.1 it follows that the convex quadratic function

$$\psi_k(x) := \varphi_k(x - x^k) + \nu_k (x - x^k)^t X_k^{-2} (x - x^k)/2$$

attains its minimum on $\{x \in \mathbb{R}^n; Ax = b\}$ at $x^k + d^k$.

In step 3, we see that the linesearch is of Armijo type [1], i.e., it consists simply of testing the unit step, then reducing the step by a factor $\beta < 1$ until a convenient point is found. We note that this linesearch is well defined because, as M_k is positive semidefinite, the function φ_k is convex. It follows that

$$\nabla f(x^k)^t d^k = \nabla \varphi_k(0)^t d^k \leq \varphi_k(d^k) - \varphi_k(0) = \varphi_k(d^k) - f(x^k),$$

hence, for $\rho > 0$ small enough,

$$\begin{aligned} f(x^k) - f(x^k + \rho d^k) &= -\rho \nabla f(x^k)^t d^k + o(\rho), \\ &\geq \rho [f(x^k) - \varphi_k(d^k)] + o(\rho). \end{aligned}$$

As $\gamma \in (0, 1)$ and $f(x^k) > \varphi_k(d^k)$, condition (1) is satisfied whenever ℓ_k is large enough.

For the statement of the result of global convergence, we need some definitions. Given $x \in F$, we denote the set of active constraints by

$$I(x) := \{i \in \{1, \dots, n\}; x_i = 0\}.$$

To any $I \subset \{1, \dots, n\}$ we associate the optimization problem

$$(P)_I \quad \min f(x); Ax = b; x_I = 0.$$

The first-order optimality system associated to $(P)_I$ is

$$(OS)_I \quad \begin{cases} \nabla f(x) + A^t \lambda - \mu = 0, \\ Ax = b, \\ x_I = 0; \mu_i = 0, i \notin I. \end{cases}$$

We will use the following hypotheses:

(H1) For all $I \subset \{1, \dots, n\}$, system $(OS)_I$ has no nonisolated solutions.

(H1)' For all $I \subset \{1, \dots, n\}$, system $(OS)_I$ has at most one solution.

(H2) There exists $\alpha > 0$; $(d^k)^t (M_k + 2\nu_k X_k^{-2}) d^k \geq \alpha \|d^k\|^2$.

(H3) $\begin{cases} \text{The constraints of (P) are qualified in the sense that} \\ (A^t \lambda)_i = 0, \forall i \notin I(\bar{x}) \text{ implies that } \lambda = 0. \end{cases}$

We briefly discuss these hypotheses. If f is strictly convex, then the optimality system $(OS)_I$, which characterizes the minima of F over the feasible set of $(P)_I$, has at most one primal solution; therefore, if (H3) is satisfied in addition, then (H1)' will be satisfied. (H1) is a weaker condition that may be useful especially for nonconvex problems. Hypothesis (H2) is a means that allows control of the decrease of the cost function at each iteration. Indeed, from (5) it follows easily that (H2) is equivalent to

$$\text{there exists } \alpha > 0; f(x^k) - \varphi_k(d^k) \geq \frac{\alpha}{2} \|d^k\|^2.$$

We have no control on the value of ν_k , except that it is nonnegative. Still, we may observe that (H2) will be satisfied if M_k is uniformly positive definite in the sense that

$$\text{there exists } \alpha > 0; (d^k)^t M_k d^k \geq \alpha \|d^k\|^2.$$

In particular, (H2) is satisfied if M_k is close to the Hessian of f and f satisfies a strong convexity condition of the type

$$\forall x \in F, \exists \alpha > 0; (d^k)^t \nabla^2 f(x) d^k \geq \alpha \|d^k\|^2 \forall d \in \mathbb{R}^n; Ad = 0.$$

Also, (H3) is no more than the hypothesis of linear independence of the gradients of active constraints.

THEOREM 2.2. *Let $\{x^k\}$ be computed by Algorithm 1. We assume that $\{M_k\}$ is bounded. Then,*

- (i) *any limit point \bar{x} of $\{x^k\}$ is a solution of $(OS)_{I(\bar{x})}$;*
- (ii) *if either (H1)' or (H1) and (H2) hold, then $\{x^k\}$ converges. If, in addition, (H3) holds then \bar{x} satisfies the first-order optimality system of (P); i.e.,*

$$(OS) \quad \begin{cases} \nabla f(\bar{x}) + A^t \bar{\lambda} - \bar{\mu} = 0, \\ A\bar{x} = b, \\ \bar{x} \geq 0, \bar{\mu} \geq 0, \bar{x}^t \bar{\mu} = 0. \end{cases}$$

The proof of the theorem uses the following lemma.

LEMMA 2.3. *The sequence $\{x^k\}$ generated by Algorithm 1 satisfies the following conditions:*

- (i) $\sum_k (f(x^k) - \varphi_k(d^k))^2 < \infty$.
- (ii) $\nu_k \rightarrow 0$.
- (iii) $(M_k)^{1/2} d^k \rightarrow 0$.
- (iv) *If, in addition, $\{M_k\}$ is bounded, then*

$$X_k[\nabla f(x^k) + A^t \lambda^{k+1}] \rightarrow 0.$$

Proof. (i) As F is bounded, $\{x^k\}$ and $\{d^k\}$ are bounded too. We deduce that for some $c_1 > 0$,

$$f(x^k) - f(x^k + \rho d^k) \geq -\rho \nabla f(x^k)^t d^k - c_1 \rho^2.$$

Using the convexity of φ_k , we get

$$-\nabla f(x^k)^t d^k \geq f(x^k) - \varphi_k(d^k),$$

so that

$$f(x^k) - f(x^k + \rho d^k) \geq \rho[f(x^k) - \varphi_k(d^k)] - c_1 \rho^2.$$

It follows after some algebra that the linesearch test is satisfied whenever

$$\rho \leq \hat{\rho}_k := \min \left\{ 1, \frac{1-\gamma}{c_1} [f(x^k) - \varphi_k(d^k)] \right\}.$$

This implies that $\rho^k \geq \beta \hat{\rho}_k$. Plugging this in the linesearch test and using the fact that as F is bounded, $\{f(x^k)\}$ is bounded from below, we deduce that necessarily $(f(x^k) - \varphi_k(d^k))$ vanishes and, for k large enough,

$$f(x^k) - f(x^{k+1}) \geq \gamma \beta \frac{1-\gamma}{c_1} (f(x^k) - \varphi_k(d^k))^2.$$

Relation (i) follows.

(ii), (iii) By (i), we get that the left-hand side of (5) goes to 0. Then each of the nonnegative terms on the right-hand side must go to 0, and that proves (ii) and (iii).

(iv) From (2) we deduce

$$(6) \quad X_k[\nabla f(x^k) + A^t \lambda^{k+1}] = -\nu_k X_k^{-1} d^k - X_k M_k d^k.$$

From (4) we have that $\|\nu_k X_k^{-1} d^k\|_2 = \nu_k \delta_k$. So, using Lemma 2.3 (ii) and the boundedness of $\{\delta_k\}$, it follows that $\|\nu_k X_k^{-1} d^k\| \rightarrow 0$. If, in addition, $\{M_k\}$ is bounded, we get that $X_k M_k d^k = X_k (M_k)^{1/2} (M_k)^{1/2} d^k \rightarrow 0$ by using the boundedness of $\{X_k\}$ and Lemma 2.3 (iii). Henceforth, the left-hand side of (6) goes to 0. \square

Proof of Theorem 2.1. (i) Let us denote by $R(\cdot)$ the range of an operator. Define

$$\bar{I} := \{1, \dots, n\} - I(\bar{x}).$$

From point (iv) of Lemma 2.3 it follows that

$$[\nabla f(x^k) + A^t \lambda^{k+1}]_{\bar{I}} \rightarrow 0.$$

Since $R(A^t)_{\bar{I}}$ is closed, we deduce that $\nabla f(\bar{x})_{\bar{I}} \in R(A^t)_{\bar{I}}$, i.e., $(\nabla f(\bar{x}) + A^t \bar{\lambda})_{\bar{I}} = 0$ for some $\bar{\lambda} \in \mathbb{R}^p$; system (OS) $_{I(\bar{x})}$ follows.

(ii) We first discuss the convergence of $\{x^k\}$. Note that $x_i^{k+1} = x_i^k(1 + \rho^k d_i^k / x_i^k)$; hence,

$$x_i^{k+1} \leq (1 + 1/\delta)x_i^k.$$

It follows that if $(x^k, x^{k+1}) \rightarrow (\bar{x}, \hat{x})$ for a subsequence, then $I(\bar{x}) \subset I(\hat{x})$.

If (H1)' holds, using point (i) we deduce that $\bar{x} = \hat{x}$ and, in particular, $\|x^{k+1} - x^k\| \rightarrow 0$; hence, the set of limit points of $\{x^k\}$ is connected. Using (H1)' again, it follows that the set of limit points is finite. Hence, the entire sequence converges towards the same point.

Now let us analyze the case when (H1) and (H2) hold. We know by Lemma 2.3 (i) that $f(x^k) - \varphi_k(d^k) \rightarrow 0$. With (5) and (H2), this implies that $d^k \rightarrow 0$. As $\|x^{k+1} - x^k\| = \rho^k \|d^k\|$ and $\rho_k \leq 1$, the set of limit points of $\{x^k\}$ is connected. By (i) and (H1) each of them is isolated. It follows that the sequence converges.

We now prove that (OS) is satisfied under the additional assumption (H3). If $x^k \rightarrow \bar{x}$ then there exists $(\bar{\lambda}, \bar{\mu})$ such that $(\bar{x}, \bar{\lambda}, \bar{\mu})$ verifies the first-order optimality system of (P) $_{I(\bar{x})}$ by (i). We have to show that $\bar{\mu}_{I(\bar{x})} \geq 0$. With Lemma 2.3 (iv) and (H3), we deduce that $\{\lambda^k\}$ converges to $\bar{\lambda}$; hence, by (2) we have $\mu^{k+1} := -\nu_k X_k^{-2} d^k$ converges to $\bar{\mu}$. Let $i \in I(\bar{x})$ be such that $\bar{\mu}_i < 0$; then $d_i^k = -(x_i^k)^2 \mu_i^{k+1} / \nu_k > 0$ for k large enough, and this contradicts the fact that $x_i^k \rightarrow \bar{x}_i = 0$. \square

3. Acceptance of the unit stepsize. In this section we perform a local analysis around some point \bar{x} , local solution of (P). We seek conditions implying that if \bar{x} is a limit point of $\{x^k\}$, the sequence $\{x^k\}$ converges to \bar{x} and $\rho^k = 1$ is accepted. We note that the rate of convergence of the cost will not be better than linear, as this is the case in LP. Hence, the interest in obtaining a unit stepsize might be questionable. Our motivation is the following. We know that for QP problems, the solution can be computed with a good precision in a small number of iterates by using the exact Hessian for M_k (see [2] and [5]). Hence, we try to reproduce, for problems with a nonquadratic cost, this behavior. What we may prove, by a theoretical study, is that provided that M_k approximates the Hessian of the cost in a certain sense, the stepsize 1 is accepted; we then may hope that the contribution of the “nonquadratic part” of the cost is asymptotically negligible so that the rapid (although linear) convergence still occurs.

It might be argued that the need for M_k to be both positive semidefinite and an approximation of the Hessian in a certain sense makes the theory applicable only in the case of a convex f . This is not so. The situation is comparable to the one for sequential QP algorithms that use a positive definite approximation of the Hessian. The key property is that the Hessian of the cost is positive definite in the tangent space under some natural second-order assumptions, whereas the approximation in the normal space plays no role. This allows approximation in an effective way of a possibly indefinite Hessian by a positive semidefinite matrix.

We need a few definitions. Assuming that \bar{x} satisfies (H3), it follows that \bar{x} is associated with a unique pair $(\bar{\lambda}, \bar{\mu})$ such that (OS) holds. Define the set of strictly

active constraints as

$$J(\bar{x}) := \{i \in \{1, \dots, n\} ; \bar{\mu}_i > 0\}$$

and the extended critical cone as

$$T := \{d \in \mathbb{R}^n ; Ad = 0 ; d_i = 0, i \in J(\bar{x})\}.$$

We say that \bar{x} satisfies the strong second-order condition (see Robinson [23]) whenever

$$(SSOC) \quad \exists \alpha_1 > 0 ; d^t \nabla^2 f(\bar{x}) d \geq \alpha_1 \|d\|^2 \forall d \in T.$$

This is a sufficient condition for the strong regularity, as defined in [23], of the associated optimality system. It has proven useful in sensitivity analysis as well as in the study of convergence properties of algorithms (see, e.g., [16], [4], and [3]).

Given d in $\mathcal{N}(A)$, the null space of A , we now define d_T , d_N as the orthogonal projection (in $\mathcal{N}(A)$) of d onto T and N , where N is the orthogonal complement of T in $\mathcal{N}(A)$, i.e.,

$$N = \{z \in \mathcal{N}(A) ; z^t d = 0 \forall d \in T\},$$

of course, $d = d_T + d_N$ and $\|d\|^2 = \|d_T\|^2 + \|d_N\|^2$. Similarly, we associate d^k with d_T^k and d_N^k . Last, but not least, we define the sufficient curvature condition as

$$(SCC) \quad \begin{cases} \exists \varepsilon_0 > 0, \text{ if } \|x^k - \bar{x}\| \leq \varepsilon_0 \text{ then} \\ (d_T^k)^t M_k d_T^k \geq \frac{1}{2-\gamma} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k + \varepsilon_0 \|d_T^k\|^2. \end{cases}$$

We briefly discuss this condition. Specifically, we check that if M_k satisfies the inequality below and condition (SSOC) holds, then (SCC) is satisfied. We consider the following condition:

$$(7) \quad (d_T^k)^t M_k d_T^k \geq (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k + o(\|d_T^k\|^2).$$

To see that (7) implies (SCC), note that $1/(2-\gamma) \in (0, 1)$ and $(d_T^k)^t \nabla^2 f(\bar{x}) d_T^k \geq \alpha_1 \|d_T^k\|^2$ by (SSOC). This and (7) imply that

$$(d_T^k)^t M_k d_T^k \geq \frac{1}{2-\gamma} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k + \alpha_1 \left(1 - \frac{1}{2-\gamma}\right) \|d_T^k\|^2 + o(\|d_T^k\|^2),$$

from which (SCC) follows. In particular, (SCC) is satisfied if (SSOC) holds and $M_k = \nabla^2 f(x^k)$ (which, of course, is possible only if f is convex).

Condition (SCC) is similar to a condition recently used in the analysis of successive QP algorithms [3]. It is checked in [3] that in the case of unconstrained optimization (then actually d_T^k and d^k coincide), this condition is very weak in the following sense: assuming that the second-order sufficient optimality condition hold for $(\nabla^2 f(\bar{x}) > 0)$, a necessary condition for the acceptance of the unit step for x^k close to \bar{x} is

$$(d_T^k)^t M_k d_T^k \geq \frac{1}{2-\gamma} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k + o(\|d_T^k\|^2).$$

THEOREM 3.1. Assume that $\{M_k\}$ is bounded, \bar{x} satisfies (H3) and (SSOC), and (SCC) is satisfied for x^k close enough to \bar{x} . Then, there exists $\varepsilon > 0$; if, for some k_0 , $\|x^{k_0} - \bar{x}\| < \varepsilon$ then $d^k \rightarrow 0$, $\rho^k = 1$ for all $k \geq k_0$, and $x^k \rightarrow \bar{x}$.

We need a few lemmas (Lemma 3.2 is stated in [3]; we give its proof for the reader's convenience).

LEMMA 3.2. Given $\varepsilon > 0$ and an $n \times n$ symmetric matrix M , define

$$K(\varepsilon, M) := \|M\|(1 + \|M\|/\varepsilon).$$

The two inequalities below then hold:

$$(8) \quad d_T^t M d_T \geq d^t M d - \varepsilon \|d_T\|^2 - K(\varepsilon, M) \|d_N\|^2,$$

$$(9) \quad d^t M d \geq d_T^t M d_T - \varepsilon \|d_T\|^2 - K(\varepsilon, M) \|d_N\|^2.$$

Proof. Since $d = d_T + d_N$, it follows that

$$d^t M d = d_T^t M d_T + 2d_T^t M d_N + d_N^t M d_N.$$

Hence,

$$|d^t M d - d_T^t M d_T| = |2d_T^t M d_N + d_N^t M d_N| \leq \|M\|(2\|d_T\| \cdot \|d_N\| + \|d_N\|^2).$$

Using the inequality $2ab \leq a^2 + b^2$ with $a = \sqrt{\varepsilon}\|d_T\|$ and $b = \|M\|\|d_N\|/\sqrt{\varepsilon}$, we get

$$|d^t M d - d_T^t M d_T| \leq \varepsilon \|d_T\|^2 + \|M\|(1 + \|M\|/\varepsilon) \|d_N\|^2,$$

from which the conclusion follows. \square

LEMMA 3.3. There exists $c_1 > 0$ such that

$$\|z_N\| \leq c_1 \sum_{i \in J(\bar{x})} |z_i| \quad \forall z \in \ker A.$$

Proof. We have $z_N = z - z_T$ and $(z_T)_i = 0, i \in J(\bar{x})$. Henceforth, $z_i = (z_N)_i, i \in J(\bar{x})$, and it suffices to prove that

$$\|z\| \leq c_1 \sum_{i \in J(\bar{x})} |z_i| \quad \forall z \in N.$$

Since both sides are positively homogeneous, it suffices to establish the inequality when $\|z\| = 1$. Then, the existence of c_1 amounts to saying that the problem

$$\min \sum_{i \in J(\bar{x})} |z_i| ; z \in N, \|z\| = 1$$

has a positive infimum. If this were not the case, there would exist $z \in N, \|z\| = 1$, with $z_i = 0, i \in J(\bar{x})$ because this problem has a solution by compactness arguments; hence, $z \in T$ (by definition of T), i.e., $z \in T \cap N = \{0\}$, a contradiction. \square

LEMMA 3.4. Assume that $\{M_k\}$ is bounded and \bar{x} satisfies (H3). Given $K \geq 0$, if x^k is sufficiently close to \bar{x} , the following relation holds:

$$(10) \quad \nu_k (d^k)^t X_k^{-2} d^k > K \|d_N^k\|^2.$$

Proof. Denote $\mu^{k+1} := -\nu_k X_k^{-2} d^k$. From (H3) and Lemma 2.3 (iv), we deduce that for any subsequence of $\{x^k\}$ converging to \bar{x} , the associated subsequence λ^k converges. Combining this with (2), the boundedness of M_k , and Lemma 2.3 (iii), we deduce that the associated subsequence of $\{\mu^k\}$ converges to $\bar{\mu}$. Hence, if x^k is close enough to \bar{x} , one has $d_i^k < 0$ and $\mu_i^k > \bar{\mu}_i/2$, $i \in J(\bar{x})$. Denote

$$\theta := \min\{\bar{\mu}_i/2, i \in J(\bar{x})\}.$$

It follows that

$$(11) \quad \nu_k (d^k)^t X_k^{-2} d^k \geq \nu_k \sum_{i \in J(\bar{x})} (d_i^k / x_i^k)^2 \geq \frac{1}{2} \sum_{i \in J(\bar{x})} -\bar{\mu}_i d_i^k \geq \theta \sum_{i \in J(\bar{x})} |d_i^k|.$$

Also, since $|d_i^k| \leq |x_i^k|/\delta$, it follows that $|d_i^k|$, $i \in J(\bar{x})$ can be made arbitrarily small by taking x^k close to \bar{x} . It follows with (11) that

$$(12) \quad \nu_k (d^k)^t X_k^{-2} d^k / \left(\sum_{i \in J(\bar{x})} |d_i^k| \right)^2 \rightarrow \infty.$$

We conclude with Lemma 3.3. \square

LEMMA 3.5. *Let $\alpha_1 > 0$ be given by (SSOC). Given $K > 0$, under the hypotheses of Theorem 3.1, if x^k is sufficiently close to \bar{x} then*

$$(13) \quad (d^k)^t (M_k + 2\nu_k X_k^{-2}) d^k \geq \frac{\alpha_1}{2} \|d^k\|^2 + K \|d_N^k\|^2.$$

Proof. Define

$$K(\varepsilon) := \sup_{k \in N} K(\varepsilon, M_k).$$

Because $\{M_k\}$ is bounded, we have that $K(\varepsilon) < \infty$. Apply Lemma 3.2, with $\varepsilon = \varepsilon_0$, where $\varepsilon_0 > 0$ is such that (SCC) holds. We obtain that if x^k is close to \bar{x} , then

$$(d^k)^t M_k d^k \geq \frac{1}{2-\gamma} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k - K(\varepsilon_0) \|d_N^k\|^2.$$

Since $1/(2-\gamma) \geq 1/2$, by using (SSOC) we get

$$\begin{aligned} (d^k)^t M_k d^k &\geq \frac{\alpha_1}{2} \|d_T^k\|^2 - K(\varepsilon_0) \|d_N^k\|^2, \\ &= \frac{\alpha_1}{2} \|d^k\|^2 - \left(K(\varepsilon_0) + \frac{\alpha_1}{2} \right) \|d_N^k\|^2. \end{aligned}$$

The conclusion is obtained with Lemma 3.4. \square

Proof of Theorem 3.1. (a) We first prove that $x^k \rightarrow \bar{x}$. We use the fact that $\|d^k\|$ is small whenever x^k is close to \bar{x} , k is large enough as a consequence of Lemma 3.5 and (5), and \bar{x} satisfies (SSOC). The last fact implies that \bar{x} is an isolated critical point of (P) (see [23]). As (H3) necessarily holds in a neighborhood of \bar{x} , it follows by Theorem 2.2 that \bar{x} is the only limit point of $\{x^k\}$ in some neighborhood \mathcal{V} of \bar{x} .

We now just have to prove that x^k remains in \mathcal{V} for k large enough. We can take \mathcal{V} of the form

$$\mathcal{V}_\varepsilon := \{x \in F; \|x - \bar{x}\| \leq \varepsilon\}.$$

Note that $\|d^k\| < \varepsilon/2$ whenever $x^k \in \mathcal{V}_{\varepsilon_1}$ for some $\varepsilon_1 > 0$ small enough. We may assume that $\varepsilon_1 < \varepsilon/2$. It follows that if $x^k \in \mathcal{V}_{\varepsilon_1}$, then $\|x^{k+1} - \bar{x}\| \leq \|x^k - \bar{x}\| + \|d^k\| \leq \varepsilon$. In other words, x^{k+1} is in \mathcal{V}_ε whenever x^k is in $\mathcal{V}_{\varepsilon_1}$.

On the other hand, we also know that $f(x^{k+1}) \leq f(x^k)$. So, let us define

$$\hat{f} := \inf\{f(x); x \in \mathcal{V}_\varepsilon - \mathcal{V}_{\varepsilon_1}\}.$$

Because \bar{x} is a strict local minimum of (P), we may assume that $\hat{f} > f(\bar{x})$, reducing ε and ε_1 if necessary. Now, assuming that $f(x^k) \leq \hat{f}$ and $x^k \in \mathcal{V}_{\varepsilon_1}$, it follows that $f(x^{k+1}) < \hat{f}$ and $x^{k+1} \in \mathcal{V}_\varepsilon$; using the definition of \hat{f} , we find that x^{k+1} is in $\mathcal{V}_{\varepsilon_1}$ again. This implies that the sequence $\{x^k\}$ remains in $\mathcal{V}_{\varepsilon_1}$, hence, that $x^k \rightarrow \bar{x}$.

(b) We now check that $\rho^k = 1$ for k large enough. Define

$$H_k := 2 \int_0^1 (1 - \sigma) \nabla^2 f(x^k + \sigma d^k) d\sigma.$$

Then,

$$f(x^k) - f(x^k + d^k) = -\nabla f(x^k)^t d^k - \frac{1}{2} (d^k)^t H_k d^k.$$

If x^k is close enough to \bar{x} , d^k is then close to 0 as was already observed; hence, H_k is close to $\nabla^2 f(\bar{x})$. We deduce that

$$-(d^k)^t H_k d^k \geq -(d^k)^t \nabla^2 f(\bar{x}) d^k - \frac{\varepsilon_0}{2} \|d^k\|^2,$$

with ε_0 given by (SCC). As a consequence,

$$\begin{aligned} f(x^k) - f(x^k + d^k) &\geq -\nabla f(x^k)^t d^k - \frac{1}{2} (d^k)^t \nabla^2 f(\bar{x}) d^k - \frac{\varepsilon_0}{4} \|d^k\|^2, \\ &= f(x^k) - \varphi_k(d^k) + \frac{1}{2} (d^k)^t (M_k - \nabla^2 f(\bar{x})) d^k - \frac{\varepsilon_0}{4} \|d^k\|^2. \end{aligned}$$

So, by (1), the unit step will be accepted if

$$(14) \quad (1 - \gamma)(f(x^k) - \varphi_k(d^k)) + \frac{1}{2} (d^k)^t (M_k - \nabla^2 f(\bar{x})) d^k - \frac{\varepsilon_0}{4} \|d^k\|^2 \geq 0.$$

Using (5), Lemma 3.2 with $\varepsilon = \varepsilon_0/2$ (where ε_0 is given by (SCC)), (SCC), and Lemma 3.4, we get

$$\begin{aligned} f(x^k) - \varphi_k(d^k) &= \frac{1}{2} (d^k)^t M_k d^k + \nu_k \delta_k^2, \\ &\geq \frac{1}{2} (d_T^k)^t M_k d_T^k - \frac{\varepsilon_0}{4} \|d_T^k\|^2 - \frac{K(\varepsilon_0)}{2} \|d_N^k\|^2 + \nu_k \delta_k^2, \\ &\geq \frac{1}{2(2 - \gamma)} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k - \frac{K(\varepsilon_0)}{2} \|d_N^k\|^2 + \frac{\varepsilon_0}{4} \|d_T^k\|^2 + \nu_k \delta_k^2, \\ &\geq \frac{1}{2(2 - \gamma)} (d_T^k)^t \nabla^2 f(\bar{x}) d_T^k + \frac{\nu_k}{2} \delta_k^2. \end{aligned}$$

Similarly, by defining $K'(\varepsilon) := \sup_{k \in \mathbb{N}} K(\varepsilon, M_k - \nabla^2 f(\bar{x}))$, we obtain

$$\begin{aligned} \frac{1}{2}(d^k)^t(M_k - \nabla^2 f(\bar{x}))d^k &\geq \frac{1}{2}(d_T^k)^t(M_k - \nabla^2 f(\bar{x}))d_T^k - \frac{\varepsilon_0}{4}\|d_T^k\|^2 - K'(\varepsilon_0)\|d_N\|^2, \\ &\geq \frac{\gamma - 1}{2(2 - \gamma)}(d_T^k)^t \nabla^2 f(\bar{x}) d_T^k - K'(\varepsilon_0)\|d_N^k\|^2 + \frac{\varepsilon_0}{4}\|d_T^k\|^2. \end{aligned}$$

Combining these inequalities and using Lemma 3.4 again, we get

$$\begin{aligned} (1 - \gamma)(f(x^k) - \varphi_k(d^k)) &+ \frac{1}{2}(d^k)^t(M_k - \nabla^2 f(\bar{x}))d^k - \frac{\varepsilon_0}{4}\|d^k\|^2 \\ &\geq (1 - \gamma)\frac{\nu_k}{2}\delta_k^2 - (K'(\varepsilon_0) + \frac{\varepsilon_0}{4})\|d_N^k\|^2 \geq 0. \end{aligned}$$

We have proven (14) as required. It follows that the unit step is accepted, hence, d^k vanishes, as was to be proved. \square

We now check that if M_k is close to $\nabla^2 f(\bar{x})$ in a very weak sense (see (16) below), then the following holds:

$$(15) \quad \sum_k (\|x^k - \bar{x}\| + \|\lambda^k - \bar{\lambda}\| + \|\mu^k - \bar{\mu}\|) < \infty.$$

THEOREM 3.6. *Assume that the hypotheses of Theorem 3.1 hold and, in addition, that \bar{x} satisfies the strict complementarity condition. If $x^k \rightarrow \bar{x}$ (hence, $\rho^k = 1$ by Theorem 3.1) then there exists $\varepsilon_1 > 0$ such that*

$$(16) \quad \|(M_k - \nabla^2 f(\bar{x}))d_T^k\| \leq \varepsilon_1 \|d^k\|$$

implies (15).

Let us note that Newton's method satisfies (16). Note that if we assume $M_k \rightarrow \nabla^2 f(\bar{x})$, then we may violate the positive definiteness requirement on M_k since $\nabla^2 f(\bar{x})$ need not be positive definite.

Proof. Denote

$$I := I(\bar{x}), \quad \bar{I} := \{1, \dots, n\} - I.$$

The proof is based on the mapping

$$\psi(x, \lambda) := \begin{cases} (\nabla f(x) + A^t \lambda)_{\bar{I}}, \\ Ax - b, \\ x_I. \end{cases}$$

It follows easily from (SSOC) and (H3) that $\psi(x, \lambda)$ has an invertible derivative at $(\bar{x}, \bar{\lambda})$; hence, there exists some $a_1 > 0$ such that

$$(17) \quad \|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\| \leq a_1 \|\psi(x^{k+1}, \lambda^{k+1})\|.$$

(a) Let us prove that

$$(18) \quad \exists K_1, K_4; \|\psi(x^{k+1}, \lambda^{k+1})\| \leq nK_1\nu_{k+1} + nK_4\nu_k + \|x^{k+1} - x^k\|/(4a_1).$$

Indeed, from the convergence of $\{x^k\}$ to \bar{x} and (H3) and by using Lemma 2.3 (iv) and (2), it follows that $(\lambda^k, \mu^k) \rightarrow (\bar{\lambda}, \bar{\mu})$. Now, multiplying (2) by X_k and recalling that $\nu_k \|X_k^{-1} d^k\| = \nu_k \delta_k$, we get

$$\|X_k(\nabla f(x^k) + M_k d^k + A^t \lambda^{k+1})\| = \nu_k \delta_k.$$

Using the strict complementarity hypothesis and the relation $|z_i| \leq \|z\|$, we obtain, for some $K_1 > 0$,

$$(19) \quad x_i^k \leq K_1 \nu_k, \quad i \in I,$$

$$(20) \quad |(\nabla f(x^k) + M_k d^k + A^t \lambda^{k+1})_i| \leq K_1 \nu_k, \quad i \notin I.$$

Now choose ε_1 in (16) as $\varepsilon_1 = 1/(8a_1 n)$.

We have

$$\begin{aligned} \nabla f(x^k) + M_k d^k &= \nabla f(x^k) + \nabla^2 f(\bar{x}) d^k + (M_k - \nabla^2 f(\bar{x})) d^k, \\ (21) \quad &= \nabla f(x^{k+1}) + r^k + (M_k - \nabla^2 f(\bar{x})) d^k, \end{aligned}$$

where the term r^k for x^k close to \bar{x} satisfies

$$(22) \quad \|r^k\| \leq \|d^k\|/(8a_1 n).$$

Also, by (16) and as $\{M_k\}$ is bounded, we get for some $K_2 > 0$

$$\begin{aligned} \|(M_k - \nabla^2 f(\bar{x})) d^k\| &\leq \|(M_k - \nabla^2 f(\bar{x})) d_T^k\| + K_2 \|d_N^k\|, \\ (23) \quad &\leq \|d^k\|/(8a_1 n) + K_2 \|d_N^k\|. \end{aligned}$$

Using (21), (22), (23), and Lemma 3.3, we obtain for some $K_3 > 0$

$$(24) \quad \|\nabla f(x^k) + M_k d^k - \nabla f(x^{k+1})\| \leq \|d^k\|/(4a_1 n) + K_3 \sum_{j \in I} |d_j^k|.$$

Now we prove (18). As $\bar{\mu} = -\lim_{k \rightarrow +\infty} \nu_k X_k^{-2} d^k$, using the strict complementarity hypothesis for k large enough and for all $j \in I$, we get $d_j^k < 0$, hence, $|d_j^k| \leq x_j^k$. So, combining this with (19), (20), and (24), we get for some $K_4 > 0$

$$(25) \quad |(\nabla f(x^{k+1}) + A^t \lambda^{k+1})_i| \leq K_4 \nu_k + \|d^k\|/(4a_1 n), \quad i \notin I.$$

So, by (19) and (25), we get (18).

(b) On the other hand, by (5), the linesearch rule, and the fact that $\rho^k = 1$, we have

$$f(x^k) - f(x^{k+1}) \geq \gamma(f(x^k) - \varphi_k(d^k)) \geq \gamma \nu_k \delta_k^2.$$

Hence, as $\delta_k \geq \delta > 0$,

$$(26) \quad \nu := \sum_{k=1}^{\infty} \nu_k < \infty.$$

Hence, using (17) and (18), we get

$$\sum_{k=k_0}^{\bar{k}} (\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\|) \leq a_1 n(K_1 + K_4)\nu + \frac{1}{4} \sum_{k=k_0}^{\bar{k}} \|x^{k+1} - x^k\|.$$

Now, by using

$$\frac{1}{4} \sum_{k=k_0}^{\bar{k}} \|x^{k+1} - x^k\| \leq \frac{1}{4} \sum_{k=k_0}^{\bar{k}} (\|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\|) \leq \frac{1}{2} \sum_{k=k_0}^{\bar{k}} \|x^{k+1} - \bar{x}\| + \frac{1}{4} \|x^{k_0} - \bar{x}\|,$$

we deduce that

$$\sum_{k=k_0}^{\bar{k}} (\|x^{k+1} - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\|) \leq 2a_1 n(K_1 + K_4)\nu + \frac{1}{2} \|x^{k_0} - \bar{x}\|.$$

Finally, we obtain (15), noticing that by (2)

$$\begin{aligned} \mu^{k+1} - \bar{\mu} &= O(\|x^k - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\| + \|d^k\|), \\ &= O(\|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\|). \quad \square \end{aligned}$$

Acknowledgment. The authors would like to thank the referees for their useful remarks.

REFERENCES

- [1] L. ARMIJO, *Minimization of functions having Lipschitz continuous first partial derivatives*, Pacific J. Math., 16 (1966), pp. 1–3.
- [2] J. F. BONNANS AND M. BOUHTOU, *The trust region affine interior point algorithm for nonconvex quadratic programming*, RAIRO Rech. Opér., 29 (1995), pp. 195–217.
- [3] J. F. BONNANS AND G. LAUNAY, *An implicit trust region algorithm for constrained optimization*, SIAM J. Optim., 5 (1995), pp. 792–812.
- [4] J. F. BONNANS AND A. SULEM, *Pseudopower expansion of generalized equations and constrained optimization problems*, Math. Programming, 70 (1995), pp. 123–148.
- [5] M. BOUHTOU, *Méthodes de points intérieurs pour l'optimisation des systèmes de grande taille*, Ph.D. dissertation, Université de Paris-IX Dauphine, Paris, France, 1993.
- [6] J. P. BULTEAU AND J. P. VIAL, *A restricted trust region algorithm for unconstrained optimization*, J. Optim. Theory Appl., 47 (1985), pp. 413–435.
- [7] T. F. COLEMAN AND Y. LI, *An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds*, Technical report TR93-1342, Computer Science Department, Cornell University, Ithaca, NY, 1993.
- [8] D. DEN HERTOG, C. ROOS, AND T. TERLAKY, *On the classical logarithmic barrier function method for a class of smooth convex programming problems*, J. Optim. Theory Appl., 73 (1992), pp. 1–25.
- [9] J. E. DENNIS, M. HEINKENSCHLOSS, AND L. N. VICENTE, *Trust-Region Interior Point Algorithms for a Class of Nonlinear Programming Problems*, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1994.
- [10] I. I. DIKIN, *Iterative solutions of problems of linear and quadratic programming*, Soviet Math. Dokl., 8 (1967), pp. 674–675.
- [11] I. I. DIKIN AND V. I. ZORKALCEV, *Iterative Solutions of Mathematical Programming Problems (Algorithms of Interior Point Methods)*, Nauka Publishers, Novosibirsk, 1980 (in Russian).
- [12] R. FLETCHER, *Practical Methods of Optimization*, 2nd ed., John Wiley & Sons, Chichester, New York, 1987.

- [13] C. C. GONZAGA AND L. A. CARLOS, *A primal affine-scaling algorithm for linearly constrained convex programs*, Technical report ES-238/90, Universidade Federal do Rio de Janeiro, Brazil, 1990.
- [14] C. C. GONZAGA, *An interior trust region method for linearly constrained optimization*, COAL Bull., 19 (1991), pp. 55–65.
- [15] F. JARRE, *On the method of analytic centers for solving smooth convex programs*, in Lecture Notes in Mathematics 1405, S. Dolecki, ed., Springer-Verlag, Berlin, New York, 1989, pp. 69–85.
- [16] K. JITTORNTRUM, *Solution point differentiability without strict complementarity in nonlinear programming*, Math. Programming Study, 21 (1984), pp. 127–138.
- [17] N. KARMAKAR, *A new polynomial time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.
- [18] G. MCCORMICK, *The projective SUMT method for convex programming problems*, Math. Oper. Res., 14 (1989), pp. 203–223.
- [19] S. MEHROTRA AND J. SUN, *An interior point algorithm for solving smooth convex programs based on Newton's method*, in Mathematical Developments Arising from Linear Programming, Contemporary Mathematics 114, J. C. Lagarias and M. J. Todd, eds., AMS, Providence, RI, 1990, pp. 265–284.
- [20] R. D. C. MONTEIRO AND I. ADLER, *An extension of Karmarkar-type algorithms to a class of convex separable programming problems with global rate of convergence*, Math. Oper. Res., 15 (1990), pp. 408–422.
- [21] J. J. MORÉ, *Recent developments in algorithms and software for trust region method*, in Mathematical Programming, the State of the Art, A. Bachem, M. Grötschel, and B. Korte, eds., Springer-Verlag, Berlin, New York, 1983, pp. 258–287.
- [22] M. J. D. POWELL, *Algorithms for nonlinear constraints that use Lagrangian functions*, Math. Programming, 15 (1978), pp. 224–248.
- [23] S. M. ROBINSON, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [24] D. C. SORENSEN, *Newton's method with a model trust region modification*, SIAM J. Numer. Anal., 19 (1982), pp. 409–426.
- [25] J. SUN, *A convergence proof for an affine-scaling algorithm for convex quadratic programming without nondegeneracy assumptions*, Math. Programming, 60 (1993), pp. 69–79.
- [26] T. TSUCHIYA, *Global Convergence of the Affine Scaling Algorithm for Primal Degenerate Strictly Convex Quadratic Programming Problems*, Math. Oper. Res., 47 (1993), pp. 509–539.
- [27] Y. YE AND E. TSE, *An extension of Karmarkar's projective algorithm for convex quadratic programming*, Math. Programming, 44 (1989), pp. 157–179.