JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 119, No. 2, pp. 357–377, November 2003 (© 2003)

Globally and Quadratically Convergent Algorithm for Minimizing the Sum of Euclidean Norms[1]

G. ZHOU,[2] K. C. TOH,[3] AND D. SUN[4]

Communicated by O. L. Mangasarian

Abstract. For the problem of minimizing the sum of Euclidean norms (MSN), most existing quadratically convergent algorithms require a strict complementarity assumption. However, this assumption is not satisfied for a number of MSN problems. In this paper, we present a globally and quadratically convergent algorithm for the MSN problem. In particular, the quadratic convergence result is obtained without assuming strict complementarity. Examples without strictly complementary solutions are given to show that our algorithm can indeed achieve quadratic convergence. Preliminary numerical results are reported.

Key Words. Sum of norms, strict complementarity, quadratic convergence.

1. Introduction

Consider the problem of minimizing a sum of Euclidean norms (MSN):

$$\min_{x \in \Re^n} \sum_{i=1}^{m} \|a_i - A_i^T x\|, \tag{1}$$

where $a_1, a_2, \ldots, a_m \in \Re^d$ are column vectors and $A_1, A_2, \ldots, A_m$ are $n \times d$ matrices. Let

$$f(x) = \sum_{i=1}^{m} f_i(x), \tag{2}$$

[1]This research was supported in part by the Singapore-MIT Alliance.
[2]Research Associate, Department of Mathematics and Statistics, Curtin University of Technology, WA 6102, Australia.
[3]Associate Professor, Department of Mathematics, National University of Singapore, Singapore.
[4]Assistant Professor, Department of Mathematics, National University of Singapore, Singapore.

where

$$f_i(x) = \|a_i - A_i^T x\|, \qquad i = 1, 2, \ldots, m.$$

The problem (1) arises in many applications, such as VLSL design, Euclidean facilities location, and Steiner minimal tree with a given topology; see e.g. Refs. 1–3 for more details.

Many algorithms have been proposed for the problem (1); see e.g. Refs. 1–4. Under a strict complementarity assumption, quadratic convergence results have been obtained in Refs. 1, 2, 4. However, the assumption is not satisfied for some MSN problems; even simple ones such as Euclidean single-facility location problems.

In this paper, we reformulate the problem (1) as a monotone variational inequality problem (MVIP for short). Then, we present an algorithm for (1) by solving the MVIP. The algorithm is globally and quadratically convergent. In particular, unlike current results obtained in Refs. 1, 2, 4, the quadratic convergence is obtained without assuming strict complementarity.

The rest of the paper is organized as follows. In Section 2, we reformulate the problem (1) as an MVIP. In Section 3, we propose an algorithm for the MSN problem by solving the MVIP. In Section 4, we show that the algorithm is quadratically convergent without assuming the strict complementarity. A number of examples are given in Section 5 to show that conditions used in this paper are satisfied, but that strict complementarity does not hold. In Section 6, we report some preliminary numerical results. We conclude the paper in Section 7.

Concerning notation, we let

$$\Re^n_+ = \{x \in \Re^n : x \geq 0\}$$

and

$$\Re^n_{++} = \{x \in \Re : x > 0\}.$$

If $n = 1$, then $\Re^n_+$ and $\Re^n_{++}$ are denoted by $\Re_+$ and $\Re_{++}$ respectively. In this paper, unless otherwise stated, all vectors are column vectors. We denote the cardinality of a set $\mathscr{A}$ by $|\mathscr{A}|$. For a closed convex set $\Omega \subseteq \Re^n$ and $x \in \Re^n$, we let $\Pi_\Omega(x)$ be the Euclidean projection of x onto Ω. If $\Omega = \Re^n_+$, then we denote $\Pi_\Omega(x)$ by x_+. We use $\epsilon \downarrow 0^+$ to denote the limit of a positive scalar ϵ which tends to 0.

We let I_d denote the $d \times d$ identity matrix, $0_{n \times m}$ the $n \times m$ zero matrix, $0_n \in \Re^n$ the zero column vector, and $e_n \in \Re^n$ the column vector of ones. To represent a large matrix with several smaller matrices, we use semicolons for column concatenation and commas for row concatenation. These notations also applies to vectors. Given a finite number of square matrices

$Q_1, \dots, Q_n$, we denote the block diagonal matrix with these matrices as blocks by either $\mathrm{diag}(Q_1, \dots, Q_n)$ or $\mathrm{diag}(Q_i, i = 1, \dots, n)$. Let

$$M = \{1, 2, \dots, m\}.$$

Given vectors $y_i \in \Re^d$, $i \in M$, we let

$$\mathrm{D}(y_i, i \in M) := \begin{bmatrix} y_1 & 0_d & \cdots & 0_d \\ 0_d & y_2 & \cdots & 0_d \\ \vdots & \vdots & \ddots & \vdots \\ 0_d & 0_d & \cdots & y_m \end{bmatrix}.$$

Let

$$Q = \mathrm{D}(y_i, i \in M), \qquad \Lambda = \mathrm{diag}(\lambda_i I_d, i \in M),$$

$$R = \mathrm{diag}(\lambda_i, i \in M), \qquad X = \mathrm{diag}(\lambda_i y_i y_i^T, i \in M).$$

If

$$\|y_i\| = 1, \qquad i \in M,$$

then it is readily shown that

$$Q^T \Lambda Q = R, \tag{3}$$

$$Q R Q^T = X, \tag{4}$$

$$R Q^T \Lambda^{-1} = Q^T. \tag{5}$$

2. Reformulation

In this section, we reformulate the problem (1) as a monotone variational inequality problem. This reformulation is important to our design of quadratically convergent algorithms. Firstly, we give the following lemma. We omit the proof of the lemma as it is easy.

Lemma 2.1. Let $d(x) = \|x\|$, $x \in \Re^n$. Then, $y \in \partial d(x)$ if and only if there exist $g \geq 0$ and $h \geq 0$ such that

$$x - gy = 0, \qquad 1/2 - (1/2)\|y\|^2 = h, \qquad gh = 0. \tag{6}$$

Let $\mathcal{S}$ be the solution set of problem (1). It is known that $x \in \mathcal{S}$ if and only if

$$0 \in \sum_{i=1}^{m} \partial f_i(x). \tag{7}$$

From Lemma 2.1 and Theorem 4.2.1 in Ref. 5, Chapter VI, for $i \in M$,

$$\partial f_i(x) = \begin{Bmatrix} -A_i y_i : y_i \in \Re^d, & A_i^T x - a_i + \lambda_i y_i = 0, \\ 1/2 - (1/2)\|y_i\|^2 = h_i, & \lambda_i h_i = 0, \lambda_i \geq 0, h_i \geq 0 \end{Bmatrix}. \quad (8)$$

Thus, (7) is equivalent to the following system:

$$-Ay = 0, \qquad A_i^T x - a_i + \lambda_i y_i = 0, \qquad i \in M, \quad (9)$$

$$1/2 - (1/2)\|y_i\|^2 = h_i, \qquad \lambda_i h_i = 0, \qquad h_i \geq 0, \qquad \lambda_i \geq 0, \qquad i \in M, \quad (10)$$

where

$$A = [A_1, A_2, \ldots, A_m] \quad \text{and} \quad y = [y_1; y_2; \ldots; y_m].$$

Let

$$\lambda = [\lambda_1; \ldots; \lambda_m] \in \Re^m \quad \text{and} \quad h = [h_1; \ldots; h_m] \in \Re^m.$$

Let $[x^*; y^*; \lambda^*; h^*]$ be a solution of the system (9)–(10). We say that strict complementarity holds at $[x^*; y^*; \lambda^*; h^*]$ if

$$\lambda_i^* + h_i^* > 0, \qquad \forall i \in M. \quad (11)$$

Note that (11) is equivalent to the following:

$$\|y_i^*\| < 1 \text{ whenever } a_i - A_i^T x^* = 0, \qquad \forall i \in M. \quad (12)$$

Let

$$\Lambda = \text{diag}(\lambda_i I_d, i \in M), \qquad Y = [\|y_1\|^2; \ldots; \|y_m\|^2] \in \Re^m. \quad (13)$$

Let

$$u = [x; y; \lambda] \in \Re^q, \qquad \text{where } q = n + md + m.$$

Define $F: \Re^q \to \Re^q$ by

$$F(u) = \begin{bmatrix} -Ay \\ A^T x - a + \Lambda y \\ (1/2)e_m - (1/2)Y \end{bmatrix}, \quad (14)$$

and the set $\Omega \subset \Re^q$ by

$$\Omega = \{u := [x; y; \lambda] \in \Re^q : x \in \Re^n, y \in \Re^{md}, \lambda \in \Re_+^m\}. \quad (15)$$

Then, it is readily shown that the system (9)–(10) is equivalent to the following variational inequality problem: find a vector $u^* = [x^*; y^*; \lambda^*] \in \Omega$ such that

$$F(u^*)^T (u - u^*) \geq 0, \qquad \forall u \in \Omega. \quad (16)$$

Lemma 2.2. The function F is a smooth monotone mapping in Ω. Moreover, if A has rank n, then the solution set of (16) is nonempty and bounded.

Proof. Let

$$J(y) = \mathrm{D}(y_i, i \in M), \tag{17}$$

where $\mathrm{D}(y_i, i \in M)$ is defined in (3). The Jacobian matrix of F is given by

$$F'(u) = \begin{bmatrix} 0_{n\times n} & -A & 0_{n\times m} \\ A^T & \Lambda & J(y) \\ 0_{m\times n} & -J(y)^T & 0_{m\times m} \end{bmatrix}. \tag{18}$$

Since $F'(u)$ is the sum of a skew-symmetric matrix and a diagonal matrix with nonnegative diagonal elements, $F'(u)$ is positive semidefinite in Ω. Therefore, F is a smooth monotone mapping in Ω. It follows from Lemma 2.1 in Ref. 4 that the solution set $\mathcal{S}$ of (1) is nonempty and bounded if A has rank n. Thus, it is readily proven that the solution set of (9)–(10) is nonempty and bounded. Therefore, the solution set of (16) is nonempty and bounded. □

Let $z = [x; y; s] \in \Re^q$ and let $\Pi_\Omega(z)$ be the Euclidean projection of z onto Ω. It is well known that solving (16) is equivalent to solving the Robinson normal equation,

$$E(z) := F(\Pi_\Omega(z)) + z - \Pi_\Omega(z) = 0, \tag{19}$$

in the following sense: if z^* is a solution of (19), then $\Pi_\Omega(z^*) = [x^*; y^*; s_+^*]$ is a solution of (16); conversely, if u^* is a solution of (16), then $z^* = u^* - F(u^*)$ is a solution of (19); see Ref. 6. Let

$$\Lambda = \mathrm{diag}((s_i)_+ I_d, i \in M). \tag{20}$$

The function $E(z)$ can be rewritten as follows:

$$E(z) = \begin{bmatrix} -Ay \\ A^T x - a + \Lambda y \\ (1/2)e_m - (1/2)Y + s - s_+ \end{bmatrix}. \tag{21}$$

Lemma 2.3. The function E has the following properties:

(i) E is strongly semismooth.
(ii) If A has rank n, then the solution set of (19), i.e., $E^{-1}(0)$, is nonempty and bounded.

Proof.

(i) It is readily shown that the function s_+ is strongly semismooth; see Ref. 7 for the definition of semismoothness. By using Theorem 19 in Ref. 8 and (21), we can conclude that E is strongly semismooth.

(ii) Let $\mathcal{S}_F$ be the solution set of (16) and let $\mathcal{S}_E$ be the solution set of (19). By Lemma 2, $\mathcal{S}_F$ is bounded. For any $z = [x; y; s] \in \mathcal{S}_E$,

$$u := \Pi_\Omega(z) = [x; y; s_+] \in \mathcal{S}_F.$$

By (19),

$$z = u - F(u).$$

Since $\mathcal{S}_F$ is bounded, this implies that $\|F(u)\|$ is bounded for all $u \in \mathcal{S}_F$. Thus, there exists a $c > 0$ such that

$$\|z\| \leq c, \qquad \text{for all } z \in \mathcal{S}_E. \qquad \square$$

Now, we will give a smooth approximation to the function E defined in (19). In Ref. 9, Chen and Mangasarian presented a class of smooth approximations to the function

$$r_+ = \max\{0, r\}, \qquad r \in \Re.$$

Among these smooth approximations, the Chen–Harker–Kanzow–Smale smooth function is the most commonly used. It is defined by

$$\phi(t, r) = 1/2(r + \sqrt{r^2 + 4t^2}), \qquad (t, r) \in \Re_{++} \times \Re. \tag{22}$$

Let $p: \Re^2 \to \Re$ be defined by

$$p(t, r) = \begin{cases} \phi(|t|, r), & \text{if } t \neq 0, \\ r_+, & \text{if } t = 0. \end{cases} \tag{23}$$

The properties of the above functions ϕ and p can be found in Refs. 4 and 10–12.

Let $v := [t; z] = [t; x; y; s]$. Define $D: \Re^{m+1} \to \Re^m$ by

$$D(t, s) = [p(t, s_1); \ldots; p(t, s_m)], \tag{24}$$

define $P: \Re^{q+1} \to \Re^q$ by

$$P(v) = [x; y; D(t, s)], \tag{25}$$

define $K: \Re^{m+1} \to \Re^{md \times md}$ by

$$K(t, s) = \operatorname{diag}(p(t, s_i) I_d, i \in M), \tag{26}$$

and define $H: \Re^{q+1} \to \Re^{q+1}$ by

$$H(v) = \begin{bmatrix} t \\ F(P(v)) + (1+t)z - P(v) \end{bmatrix} = \begin{bmatrix} t \\ -Ay + tx \\ A^T x - a + K(t,s)\,y + ty \\ (1/2)e_m - (1/2)Y + (1+t)s - D(t,s) \end{bmatrix}, \tag{27}$$

where F is defined in (14). Note that

$$F(P(v)) + (1+t)z - P(v)$$

is the Tikhonov regularization of $F(P(v)) + z - P(v)$. The Tikhonov regularization was used to study variational inequalities and complementarity problems; see e.g. Refs. 11 and 12.

Lemma 2.4. The function H has the following properties:

(i) H is continuously differentiable on $(\Re \backslash \{0\}) \times \Re^q$ and strongly semismooth on $\Re \times \Re^q$.

(ii) For any $z \in \Re^q$, $\lim_{t\to 0} H(v) = [0; E(z)]$.

Proof. The results follow readily from (21) and (27). □

Let

$$\begin{aligned} D_t'(t,s) &= [\partial p(t,s_1)/\partial t; \ldots; \partial p(t,s_m)/\partial t], \\ D_s'(t,s) &= \mathrm{diag}(\partial p(t,s_i)/\partial s_i,\ i \in M). \end{aligned} \tag{28}$$

Then, we have the following lemma.

Lemma 2.5. For any $v = [t; z] \in \Re_{++} \times \Re^q$, the Jacobian of H is given by

$$H'(v) := \begin{bmatrix} 1 & 0_n^T & 0_{md}^T & 0_m^T \\ x & tI_n & -A & 0_{n\times m} \\ y + JD_t'(t,s) & A^T & K(t,s) + tI_{md} & JD_s'(t,s) \\ s - D_t'(t,s) & 0_{m\times n} & -J^T & (1+t)I_m - D_s'(t,s) \end{bmatrix}, \tag{29}$$

where $J = J(y)$ is defined as in (17) and $H'(v)$ is nonsingular.

Proof. It is readily shown that (29) holds by simple computation. For any $v = [t; z] \in \Re_{++} \times \Re^q$, in order to prove that $H'(v)$ is nonsingular, we need only to prove that the submatrix

$$W := \begin{bmatrix} tI_n & -A & 0_{n\times m} \\ A^T & K(t,s) + tI_{md} & JD'_s(t,s) \\ 0_{m\times n} & -J^T & (1+t)I_m - D'_s(t,s) \end{bmatrix} \tag{30}$$

is nonsingular. Let

$$N = \begin{bmatrix} 0_n & -A & 0_{n\times m} \\ A^T & 0_{md} & J \\ 0_{m\times n} & -J^T & 0_m \end{bmatrix},$$

$$P = \mathrm{diag}(I_n, I_{md}, D'_s(t,s)),$$

$$Q = \mathrm{diag}(tI_n, K(t,s) + tI_{md}, (1+t)I_m - D'_s(t,s)).$$

Then,

$$W = NP + Q.$$

Since N is positive semidefinite and since P, Q are positive-definite diagonal matrices, by Theorem 3.3 in Ref. 13, W is nonsingular. Therefore, $H'(v)$ is nonsingular. □

3. Algorithm

Let $\gamma \in (0, 1)$. Define $\Psi: \Re^{q+1} \to \Re_+$ by

$$\Psi(v) := \|H(v)\|^2 \tag{31}$$

and define $\beta: \Re^{q+1} \to \Re_+$ by

$$\beta(v) := \gamma \min(\sqrt{\Psi(v)}, \Psi(v)). \tag{32}$$

Now, we will describe an algorithm for finding a solution of $H(v) = 0$. The algorithm is a modified version of the smoothing Newton algorithm proposed in Ref. 10.

Algorithm 3.1.

Step 0. Choose $\bar{t} \in \Re_{++}, \delta \in (0, 1)$, and $\sigma \in (0, 1/2)$. Let $\bar{v} := [\bar{t}; 0_q] \in \Re \times \Re^q$ and $v^0 = [t^0; z^0]$, where $t^0 := \bar{t}$ and z^0 is an arbitrary initial point in $\Re^q$. Choose $\gamma \in (0, 1)$ such that

$$\gamma \bar{t} \leq \gamma \sqrt{\Psi(v^0)} < 1. \tag{33}$$

Set $k := 0$.

Step 1. If $H(v^k) = 0$, then stop. Otherwise, let $\beta_k := \beta(v^k)$.

Step 2. Compute $\Delta v^k := [\Delta t^k; \Delta z^k]$ by solving the linear system

$$H(v^k) + H'(v^k)\Delta v^k = \beta_k \bar{v}. \tag{34}$$

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$\Psi(v^k + \delta^l \Delta v^k) \leq [1 - 2\sigma(1 - \gamma \bar{t})\delta^l]\Psi(v^k). \tag{35}$$

Define $v^{k+1} := v^k + \delta^{l_k}\Delta v^k$.

Step 4. Set $k := k + 1$ and go to Step 1.

Remark 3.1. Define

$$h_1 = -Ay + tx + \Delta t x,$$

$$h_2 = A^T x - a + K(t, s)y + ty + \Delta t(y + JD'_t(t, s)),$$

$$h_3 = (1/2)e_m - (1/2)Y + (1 + t)s - D(t, s) + \Delta t(s - D'_t(t, s)),$$

$$N_1 = (1 + t)I_m - D'_s(t, s), \qquad N_2 = D'_s(t, s),$$

$$N_3 = K(t, s) + tI_{md} + JN_2 N_1^{-1} J^T,$$

where J, $K(t, s)$, $D(t, s)$, $D'_t(t, s)$, $D'_s(t, s)$ are defined as in (17), (26), (24), (28). Let

$$d_i = \partial p(t, s_i)/\partial s_i, \qquad i \in M.$$

Then,

$$N_1^{-1} = \mathrm{diag}(1/(1 + t - d_i), i \in M),$$

$$N_3 = \mathrm{diag}([p(t, s_i) + t]I_d + [d_i/(1 + t - d_i)]y_i y_i^T, i \in M).$$

For $i \in M$, let

$$b_i = d_i/[(1 + t - d_i)(p(t, s_i) + t)^2 + d_i\|y_i\|^2(p(t, s_i) + t)].$$

By simple computation, we have

$$N_3^{-1} = \mathrm{diag}([1/(p(t, s_i) + t)]I_d - b_i y_i y_i^T, i \in M).$$

From (27) and (29), we can solve (34) by the following procedure. For simplicity, we omit k in (34).

Procedure 3.1.

(i) Compute $\Delta t = -t + \beta(v)\bar{v}$.
(ii) Compute N_1^{-1} and N_3^{-1}.
(iii) Compute Δx by

$$(tI_n + AN_3^{-1}A^T)\,\Delta x = -h_1 + AN_3^{-1}h_4, \tag{36}$$

where $h_4 = -h_2 + JN_2N_1^{-1}h_3$.
(iv) Compute $\Delta y = -N_3^{-1}A^T\Delta x + N_3^{-1}h_4$.
(v) Compute $\Delta s = -N_1^{-1}h_3 + N_1^{-1}J^T\Delta y$.

The system (36) is an n-dimensional symmetric positive-definite linear system, which can be solved by a direct method such as the Cholesky factorization method. The above procedure for a search direction is similar to the one used in the interior-point method in Ref. 3.

Theorem 3.1. Algorithm 3.1 is well defined. Let $\{v^k = [t^k; z^k]\}$ be an infinite sequence generated by Algorithm 3.1. Then,

$$\lim_{k\to+\infty} H(v^k) = 0 \quad \text{and} \quad \lim_{k\to+\infty} t^k = 0. \tag{37}$$

Moreover, the sequence $\{v^k\}$ is bounded if A has rank n.

Proof. We omit the proof as it is similar to that of Theorem 4.5 in Ref. 11. □

4. Quadratic Convergence

Let $v^* = [t^*; z^*]$ be a limit point of the sequence $\{v^k\}$ generated by Algorithm 3.1. By Theorem 3.1, $t^* = 0$ and v^* is a solution of $H(v) = 0$. Define

$$\mathscr{A}(v^*) = \{\lim H'(v^k)\colon v^k = [t^k; z^k],\ t^k \downarrow 0^+ \text{ and } z^k \to z^*\}. \tag{38}$$

Theorem 4.1. Suppose that $v^* = [t^*; z^*]$ is an accumulation point of an infinite sequence $\{v^k\}$ generated by Algorithm 3.1. Assume that all $W \in \mathscr{A}(v^*)$ are nonsingular. Then, the whole sequence $\{v^k\}$ converges to v^* quadratically, i.e.,

$$\|v^{k+1} - v^*\| = O(\|v^k - v^*\|^2) \tag{39}$$

and

$$\Psi(v^{k+1}) = O([\Psi(v^k)]^2). \tag{40}$$

Proof. See Theorem 4.8 in Ref. 4. □

Let $v^* = [t^*; x^*; y^*; s^*]$ be a solution of $H(v) = 0$. Then, x^* is an optimal solution to problem (1) and $[x^*; y^*; s^*]$ is a solution to (19). Let

$$M_0(x^*) = \{i \in M: \|a_i - A_i^T x^*\| = 0\}.$$

Define

$$A_0 = [A_i, i \in M_0(x^*)]$$

and

$$G(x^*) = \sum_{i \in M \setminus M_0(x^*)} \nabla^2 f_i(x^*), \tag{41}$$

where, for $i \in M \setminus M_0(x^*)$,

$$\nabla^2 f_i(x^*) = [1/\|a_i - A_i^T x^*\|] A_i A_i^T \\ - [1/\|a_i - A_i^T x^*\|^3] A_i (a_i - A_i^T x^*) (a_i - A_i^T x^*)^T A_i^T.$$

To prove a quadratic convergence result for Algorithm 3.1, we made the following assumptions:

(A1) The matrix $G(x^*)$ is positive definite.
(A2) The matrix A_0 has full column rank.

Without loss of generality, we suppose that

$$\|a_i - A_i^T x^*\| = 0, \qquad i = 1, \ldots, j,$$

where $j = |M_0(x^*)|$ and

$$\|a_i - A_i^T x^*\| > 0, \qquad i = j+1, \ldots, m.$$

From (21), for $i \in M$, we have

$$A_i^T x^* - a_i + (s_i^*)_+ y_i^* = 0, \tag{42}$$

$$(1/2)(1 - \|y_i^*\|^2) + s_i^* - (s_i^*)_+ = 0. \tag{43}$$

We claim that

$$s_i^* \leq 0, \qquad i = 1, \ldots, j. \tag{44}$$

By (43),

$$\|y_i^*\| \leq 1, \qquad i \in M,$$

since

$$s_i^* - (s_i^*)_+ \leq 0.$$

Suppose that there exists $i \in \{1, \ldots, j\}$ such that $s_i^* > 0$. By (43), $\|y_i^*\| = 1$ and hence $(s_i^*)_+ y_i^* \neq 0$, but this contradicts (42). Therefore, we proved our claim.

Now, for $i = 1, \ldots, j$, by (43),

$$\|y_i^*\| < 1 \qquad \text{if and only if } s_i^* < 0.$$

Let

$$\mathscr{I} := \{i: s_i^* < 0, i = 1, \ldots, j\}. \tag{45}$$

For $i = j+1, \ldots, m$, from (42),

$$(s_i^*)_+ > 0.$$

Thus,

$$s_i^* > 0, \qquad i = j+1, \ldots, m,$$

and by (43),

$$\|y_i^*\| = 1, \qquad i = j+1, \ldots, m.$$

Using (42), we have

$$s_i^* = \|a_i - A_i^T x^*\|, \tag{46a}$$

$$y_i^* = (a_i - A_i^T x^*)/\|a_i - A_i^T x^*\|, \qquad i = j+1, \ldots, m. \tag{46b}$$

Define

$$A_0 = [A_1, \ldots, A_j], \qquad \bar{A} = [A_{j+1}, \ldots, A_m].$$

By (42) and (46), the matrix $G(x^*)$ can be rewritten as follows:

$$\begin{aligned} G(x^*) &= \sum_{i \in M \setminus M_0(x^*)} (1/s_i^*) A_i (I_d - y_i^* (y_i^*)^T) A_i^T \\ &= \bar{A} P_1^{-1} (I_{(m-j)d} - J_1 J_1^T) \bar{A}^T, \end{aligned} \tag{47}$$

where

$$P_1 = \mathrm{diag}(s_i^* I_d, i = j+1, \ldots, m), \tag{48}$$

$$J_1 = \mathrm{D}(y_i^*, i = j+1, \ldots, m). \tag{49}$$

Proposition 4.1. Suppose that $v^* = [t^*; x^*; y^*; s^*]$ is a solution of $H(v) = 0$ and that x^* satisfies (A1) and (A2). Then, all $W \in \mathscr{A}(v^*)$ are nonsingular.

Proof. For any $W \in \mathscr{A}(v^*)$, there exists a sequence $\{v^k = [t^k; x^k; y^k; s^k]\}$ such that

$$W = \lim_{k \to +\infty} H'(v^k) = \begin{bmatrix} 1 & 0_n^T & 0_{md}^T & 0_m^T \\ x^* & 0_{n \times n} & -A & 0_{n \times m} \\ y^* + J^* D_t^* & A^T & K^* & J^* D_s^* \\ s^* - D_t^* & 0_{m \times n} & -(J^*)^T & I_m - D_s^* \end{bmatrix}.$$

Here,

$$J^* = J(y^*), \qquad K^* = K(0, s^*)$$

are defined as in (17)–(26) and

$$D_t^* = \lim_{\substack{t^k \downarrow 0^+ \\ s^k \to s^*}} D_t'(t^k, s^k), \qquad D_s^* = \lim_{\substack{t^k \downarrow 0^+ \\ s^k \to s^*}} D_s'(t^k, s^k),$$

where $D_t'(t^k, s^k)$ and $D_s'(t^k, s^k)$ are defined as in (28). From (23), (44), and (46),

$$p(0, s_i^*) = 0, \qquad i = 1, \dots, j,$$

and

$$p(0, s_i^*) = s^*, \qquad i = j+1, \dots, m.$$

Thus,

$$K^* = \operatorname{diag}(0_{jd \times jd}, P_1), \tag{50}$$

where P_1 is defined in (48). By (44), (46), and simple computation, we have

$$\lim_{\substack{t^k \downarrow 0^+ \\ s_i^k \to s_i^*}} \partial p(t^k, s_i^k)/\partial s_i^k \in [0, 1], \qquad i = 1, \dots, j, \tag{51}$$

$$\lim_{\substack{t^k \downarrow 0^+ \\ s_i^k \to s_i^*}} \partial p(t^k, s_i^k)/\partial s_i^k = 0, \qquad i \in \mathscr{I}, \tag{52}$$

$$\lim_{\substack{t^k \downarrow 0^+ \\ s_i^k \to s_i^*}} \partial p(t^k, s_i^k)/\partial s_i^k = 1, \qquad i = j+1, \dots, m. \tag{53}$$

Hence,

$$D_s^* = \operatorname{diag}(D_1, I_{m-j}), \tag{54}$$

where

$$D_1 = \operatorname{diag}(d_i, i = 1, \dots, j), \qquad d_i \in [0, 1]. \tag{55}$$

We suppose that

$$d_i = 0, \qquad i = 1, \ldots, n_1,$$
$$d_i \in (0, 1), \qquad i = n_1 + 1, \ldots, n_2$$
$$d_i = 1, \qquad i = n_2 + 1, \ldots, j.$$

Then,

$$\mathscr{I} \subseteq \{1, 2, \ldots, n_1\}, \tag{56}$$
$$\|y_i\| = 1, \qquad i = n_1 + 1, \ldots, j. \tag{57}$$

Let

$$N = \mathrm{diag}(d_i, i = n_1 + 1, \ldots, n_2),$$
$$\bar{N} = \mathrm{diag}(1 - d_i, i = n_1 + 1, \ldots, n_2).$$

Then,

$$D_1 = \mathrm{diag}(0, N, I_{j-n_2}).$$

Define

$$J_{01} = \mathrm{D}(y_i^*, i = 1, \ldots, n_1), \qquad J_{02} = \mathrm{D}(y_i^*, i = n_1 + 1, \ldots, n_2), \tag{58}$$
$$J_{03} = \mathrm{D}(y_i^*, i = n_2 + 1, \ldots, j). \tag{59}$$

Let

$$A_{01} = [A_1, \ldots, A_{n_1}], \qquad A_{02} = [A_{n_1+1}, \ldots, A_{n_2}],$$
$$A_{03} = [A_{n_2+1}, \ldots, A_j].$$

Then,

$$A_0 = [A_{01}, A_{02}, A_{03}].$$

Let

$$U = \begin{bmatrix} 0_{n\times n} & -A & 0_{n\times m} \\ A^T & K^* & J^* D_s^* \\ 0_{m\times n} & -(J^*)^T & I_m - D_s^* \end{bmatrix}$$
$$= \begin{bmatrix} 0 & -A_{01} & -A_{02} & -A_{03} & -\bar{A} & 0 & 0 & 0 & 0 \\ A_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{02}^T & 0 & 0 & 0 & 0 & 0 & J_{02} N & 0 & 0 \\ A_{03}^T & 0 & 0 & 0 & 0 & 0 & 0 & J_{03} & 0 \\ \bar{A}^T & 0 & 0 & 0 & P_1 & 0 & 0 & 0 & J_1 \\ 0 & -J_{01}^T & 0 & 0 & 0 & I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & -J_{02}^T & 0 & 0 & 0 & \bar{N} & 0 & 0 \\ 0 & 0 & 0 & -J_{03}^T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -J_1^T & 0 & 0 & 0 & 0 \end{bmatrix}.$$

To prove that W is nonsingular, it suffices to show that U is nonsingular. Let

$$U_q = 0, \qquad \text{where } q = (q_1; \dots; q_9).$$

Here, $q_1 \in \Re^n$, $q_2 \in \Re^{n_1 d}$, $q_3 \in \Re^{(n_2 - n_1)d}$, $q_4 \in \Re^{(j - n_2)d}$, $q_5 \in \Re^{(m-j)d}$, $q_6 \in \Re^{n_1}$, $q_7 \in \Re^{n_2 - n_1}$, $q_8 \in \Re^{j - n_2}$, $q_9 \in \Re^{m-j}$. Then, we have

$$A_{01} q_2 + A_{02} q_3 + A_{03} q_4 + \bar{A} q_5 = 0, \tag{60}$$

$$A_{01}^T q_1 = 0, \tag{61}$$

$$A_{02}^T q_1 + J_{02} N q_7 = 0, \tag{62}$$

$$A_{03}^T q_1 + J_{03} q_8 = 0, \tag{63}$$

$$\bar{A}^T q_1 + P_1 q_5 + J_1 q_9 = 0, \tag{64}$$

$$-J_{01}^T q_2 + q_6 = 0, \tag{65}$$

$$-J_{02}^T q_3 + \bar{N} q_7 = 0, \tag{66}$$

$$-J_{03}^T q_4 = 0, \tag{67}$$

$$J_1^T q_5 = 0. \tag{68}$$

It follows from (64) that

$$q_5 = -P_1^{-1} \bar{A}^T q_1 - P_1^{-1} J_1 q_9. \tag{69}$$

Premultiplying (69) by J_1^T and using (68) and (3), we get

$$J_1^T P_1^{-1} \bar{A}^T q_1 = -J_1^T P_1^{-1} J_1 q_9 = -\Lambda_1^{-1} q_9,$$

where

$$\Lambda_1 = \operatorname{diag}(s_i^*, i = j+1, \dots, m).$$

Thus, we have

$$q_9 = -\Lambda_1 J_1^T P_1^{-1} \bar{A}^T q_1 = -J_1^T \bar{A}^T q_1. \tag{70}$$

Note that the second equation in (70) is obtained by using (5). Premultiplying (69) by $\bar{A}$ and applying (70), we get

$$\bar{A} q_5 = -\bar{A} P_1^{-1} (I_{(m-j)d} - J_1 J_1^T) \bar{A}^T q_1 = -G(x^*) q_1,$$

i.e.,

$$\bar{A} q_5 + G(x^*)\, q_1 = 0. \tag{71}$$

Premultiplying (60) by q_1^T and applying (61), we have

$$q_1^T A_{02} q_3 + q_1^T A_{03} q_4 + q_1^T \bar{A} q_5 = 0. \tag{72}$$

Premultiplying (63) by q_4^T and applying (67), we obtain

$$q_4^T A_{03}^T q_1 = -q_4^T J_{03} q_8 = 0. \tag{73}$$

From (66), we have

$$q_7 = \bar{N}^{-1} J_{02}^T q_3. \tag{74}$$

Let

$$\tilde{N} = \text{diag}([d_i/(1-d_i)]\, y_i y_i^T,\ i = n_1+1, \ldots, n_2).$$

By using (4), we obtain

$$J_{02} N \bar{N}^{-1} J_{02}^T = \tilde{N}. \tag{75}$$

Premultiplying (62) by q_3^T and applying (74) and (75), we have

$$q_3^T A_{02}^T q_1 = -q_3^T J_{02} N q_7 = -q_3^T J_{02} N \bar{N}^{-1} J_{02}^T q_3 = -q_3^T \tilde{N} q_3. \tag{76}$$

From (71), (72), (73), (76), we get

$$q_1^T G(x^*)\, q_1^T + q_3^T \tilde{N} q_3 = 0. \tag{77}$$

Since $\tilde{N}$ is positive semidefinite, by (77) and (A1), $q_1 = 0$. Hence, $q_8 = 0$ from (63), $q_9 = 0$ from (70), and then $q_5 = 0$ from (69). Since $q_5 = 0$, by (60) and (A2), $q_2 = 0$, $q_3 = 0$ and $q_4 = 0$. Thus $q_7 = 0$ from (74) and $q_6 = 0$ from (65). Therefore, $q = 0$. This implies that the matrix U is nonsingular and the proof is completed. □

By combining Theorem 4.1 and Proposition 4.1, we obtain the main result of the paper as follows.

Theorem 4.2. Suppose that $v^* = [t^*; x^*; y^*; s^*]$ is an accumulation point of the infinite sequence $\{v^k\}$ generated by Algorithm 3.1 and that x^* satisfies (A1) and (A2). Then, the whole sequence $\{v^k\}$ converges to v^* quadratically.

Remark 4.1. Theorem 4.2 shows that, under conditions (A1) and (A2), Algorithm 3.1 has quadratic convergence. In Refs. 1–2, in order to get quadratic convergence results, besides assumptions (A1) and (A2), a strict complementarity assumption is needed.

Remark 4.2. In Ref. 3, by reformulating the MSN problem as a second-order cone programming problem, an interior-point algorithm with polynomial complexity is presented. However, most existing interior-point algorithms with quadratic convergence require a strict complementarity assumption.

In Section 5, we will show that conditions (A1) and (A2) are satisfied always for the Euclidean single-facility location (ESFL) problem, but that strict complementarity does not hold for a number of problems in this class.

5. ESFL Problem

Let $c_1, c_2, \ldots, c_m$ be m, $m \geq 3$, distinct points in $\Re^d$. Let ω_1, $\omega_2, \ldots,$ ω_m be m positive weights. Find a point $x \in \Re^d$ that minimizes

$$f(x) = \sum_{i=1}^{m} \omega_i \|x - c_i\|. \tag{78}$$

This is called the Euclidean single-facility location (ESFL) problem in Ref. 3. Let

$$a_i = \omega_i c_i \quad \text{and} \quad A_i^T = \omega_i I_d, \qquad i \in M.$$

Then, the ESFL problem can be transformed into a special case of problem (1). In what follows, we assume always that c_i, $i \in M$, are not collinear. The following lemma will be used later. We shall omit its proof, since it is easy.

Lemma 5.1. Let $L = \{1, 2, \ldots, l\}$, $v_i \in \Re^d$, $i \in L$, satisfying $\|v_i\| = 1$, $i \in L$, and let u_i, $i \in L$, be positive numbers. If there exist v_i and v_j, with $i, j \in L$, such that v_i and v_j are linearly independent, then the matrix

$$N = \sum_{i=1}^{l} u_i (I_d - v_i v_i^T)$$

is positive definite.

Let

$$g_i(x) = \omega_i \|x - c_i\|, \qquad i \in M.$$

Then, for $i \in M$,

$$\partial g_i(x) = \begin{cases} \{\omega_i (x - c_i)/\|x - c_i\|\}, & \text{if } x \neq c_i, \\ \{\omega_i y \colon y \in \Re^d, \|y\| \leq 1\}, & \text{if } x = c_i. \end{cases}$$

Proposition 5.1. If the vectors c_i, $i \in M$, are not collinear, then conditions (A1) and (A2) are satisfied at any $x \in \Re^d$.

Proof. For any $x \in \Re^d$, let

$$M_0(x) = \{i \in M \colon g_i(x) = 0\}.$$

Clearly, either $M_0(x) = \varnothing$ or $M_0(x)$ only has one element.

Case 1. $M_0(x) = \{i\}$. In this case,

$$A_0 = [A_i, i \in M_0(x)] = \omega_i I_d,$$

which means that (A2) holds. For $i \in M$ and $x \neq c_i$, let

$$v_i(x) = (x - c_i)/\|x - c_i\|.$$

Then, from (41),

$$G(x) = \sum_{i \in M \setminus M_0(x)} [1/g_i(x)]\,[I_d - v_i(x)\,[v_i(x)]^T].$$

If the vectors $c_i, i \in M$ are not collinear, then there exist $v_i(x)$ and $v_j(x)$, $i, j \in M \setminus M_0(x)$, such that $v_i(x)$ and $v_j(x)$ are linearly independent. By Lemma 5.1, the matrix $G(x)$ is positive definite, i.e., (A1) holds.

Case 2. $M_0(x) = \varnothing$. Similar to the proof of Case 1, we have that

$$G(x) = \sum_{i \in M} [1/g_i(x)]\,(I_d - v_i(x)\,[v_i(x)]^T)$$

is positive definite, which shows that (A1) holds. □

Proposition 5.2. If there exists $i \in M$ such that

$$\omega_i = \left\| \sum_{j \in M, j \neq i} \omega_j (c_i - c_j)/\|c_i - c_j\| \right\|, \tag{79}$$

then c_i is a solution of (78). However, strict complementarity does not holds at the solution point c_i.

Proof. By (79), there exists $y_i \in \Re^d$ satisfying $\|y_i\| = 1$ such that

$$\omega_i y_i + \sum_{j \in M, j \neq i} \omega_j (c_i - c_j)/\|c_i - c_j\| = 0, \tag{80}$$

which means that $0 \in \partial f(c_i)$. Thus, c_i is a solution of (78).

For any $y_i \in \Re^d$ satisfying $\|y_i\| < 1$, by using (79), we have

$$\omega_i y_i + \sum_{j \in M, j \neq 1} \omega_j (c_i - c_j)/\|c_i - c_j\| \neq 0. \tag{81}$$

This shows that the strict complementarity does not holds at the solution point c_i. □

Remark 5.1. (79) holds for a number of ESFL problems. For example, let

$$d \geq 2, \quad c_1 = 0_d, \quad \omega_1 = 1, \quad c_2 = [1 : 0_{d-1}], \quad \omega_2 = 1,$$

$$c_3 = [0; 1; 0_{d-2}], \quad \omega_3 = 1, \quad c_4 = [0; -1; 0_{d-2}], \quad \omega_4 = 1.$$

Then,

$$\omega_1 = \left\| \sum_{j=2}^{4} \omega_j (c_1 - c_j)/\|c_1 - c_j\| \right\|.$$

6. Preliminary Numerical Experiments

To show that the method proposed in the paper has quadratic convergence, we implemented Algorithm 3.1 in MATLAB and tested the following two Euclidean single-facility location (ESFL) problems. Note that, at the solution points of these two examples, strict complementarity does not hold.

Example 6.1. Here

$$d = 2, \quad c_1 = [0; 0], \quad \omega_1 = 1, \quad c_2 = [1; 0], \quad \omega_2 = 1,$$
$$c_3 = [0; 1], \quad \omega_3 = 3, \quad c_4 = [0; -1], \quad \omega_4 = 3.$$

The solution is $x^* = [0; 0]$.

Example 6.2. Here

$$d = 4, \quad c_1 = [0; 0; 0; 0], \quad \omega_1 = 0.5, \quad c_2 = [1; 0; 0; 0], \quad \omega_2 = 0.5,$$
$$c_3 = [0; 1; 0; 0], \quad \omega_3 = 2, \quad c_4 = [0; -1; 0; 0], \quad \omega_4 = 2.$$

The solution is $x^* = [0; 0; 0; 0]$.

Throughout our computational experiments, we used the following parameters:

$$\delta = 0.5, \quad \sigma = 0.0005, \quad \bar{t} = 0.5, \quad z^0 = 0.3e_q, \quad \gamma = 0.5.$$

We terminated the iteration when

$$\|E(z^k)\|_\infty \leq 10^{-12},$$

where E is defined in (19). The outputs of the algorithm for Examples 6.1 and 6.2 are given in Tables 1 and 2, which show the quadratic convergence of this method.

7. Conclusions

In this paper, we proposed a globally and quadratically convergent method for the problem of minimizing a sum of Euclidean norms. In particular, the quadratic convergence of the method was proved without assuming strict complementarity.

Table 1. Output of Algorithm 3.1 for Example 6.1.

k	$f(x^k)$	$\|E(z^k)\|_\infty$	x_1^k	x_2^k	δ^{l_k}
1	7.21E+00	3.67E+00	2.11E–01	1.63E–01	1
2	7.19E+00	2.76E+00	1.66E–01	–1.90E–01	0.5
3	7.12E+00	1.14E+00	1.75E–01	9.06E–02	1
4	7.14E+00	6.26E–01	1.73E–01	1.26E–01	1
5	7.07E+00	2.46E–01	1.34E–01	6.75E–02	1
6	7.01E+00	1.05E–01	5.96E–02	1.26E–02	1
7	7.00E+00	2.73E–02	1.12E–02	3.27E–03	1
8	7.00E+00	1.31E–03	7.67E–04	1.69E–04	1
9	7.00E+00	7.09E–06	2.47E–06	3.25E–07	1
10	7.00E+00	1.29E–10	6.06E–11	1.72E–11	1
11	7.00E+00	6.66E–16	–1.04E–16	2.49E–21	1

Table 2. Output of Algorithm 3.1 for Example 6.2.

k	$f(x^k)$	$\|E(z^k)\|_\infty$	x_1^k	x_2^k	x_3^k	x_4^k	δ^{l_k}
1	5.10E+00	2.51E+00	2.81E–01	2.54E–01	2.54E–01	2.54E–01	1
2	4.82E+00	1.49E+00	2.16E–01	–7.91E–02	1.87E–01	1.87E–01	0.5
3	4.74E+00	8.56E–01	1.94E–01	1.82E–01	1.36E–01	1.36E–01	1
4	4.69E+00	4.19E–01	1.77E–01	1.85E–01	1.12E–01	1.12E–01	1
5	4.61E+00	1.67E–01	1.51E–01	1.38E–01	7.19E–02	7.19E–02	1
6	4.54E+00	5.39E–02	1.01E–01	4.80E–02	3.27E–02	3.27E–02	1
7	4.50E+00	1.50E–01	2.88E–02	2.82E–03	8.29E–03	8.29E–03	1
8	4.50E+00	9.31E–03	1.69E–02	7.11E–03	6.44E–03	6.44E–03	1
9	4.50E+00	6.84E–03	1.44E–03	2.88E–05	4.23E–04	4.23E–04	1
10	4.50E+00	5.22E–05	3.97E–05	1.42E–05	1.47E–05	1.47E–05	1
11	4.50E+00	4.55E–08	9.85E–09	8.97E–10	3.06E–09	3.06E–09	1
12	4.50E+00	2.59E–15	1.61E–15	6.26E–16	6.77E–16	6.77E–16	1

Our numerical implementation of the algorithm is very preliminary. There are numerous computational issues to be investigated in order to make the algorithm practically efficient and robust. The main computational step in each iteration of Algorithm 1 lies in solving the linear system (36). Thus, it is necessary for us to come up with ways to solve it efficiently by exploiting sparsity or special structures present in the linear system. We shall leave these as further research topics.

References

1. Calamai, P. H., and Conn, A. R., *A Projected Newton Method for l_p Norm Location Problems*, Mathematical Programming, Vol. 38, pp. 75–109, 1987.

2. Overton, M. L., *A Quadratically Convergent Method for Minimizing a Sum of Euclidean Norms*, Mathematical Programming, Vol. 27, pp. 34–63, 1983.
3. Xue, G., and Ye, Y., *An Efficient Algorithm for Minimizing a Sum of Euclidean Norms with Applications*, SIAM Journal on Optimization, Vol. 7, pp. 1017–1036, 1997.
4. Qi, L., and Zhou, G., *A Smoothing Newton Method for Minimizing a Sum of Euclidean Norms*, SIAM Journal on Optimization, Vol. 11, pp. 389–410, 2000.
5. Hiriart-Urruty, J. B., and Lemarechal, C., *Convex Analysis and Minimization Algorithms*, Springer Verlag, Berlin, Germany, 1993.
6. Robinson, S. M., *Normal Maps Induced by Linear Transformations*, Mathematics of Operations Research, Vol. 17, pp. 691–714, 1992.
7. Qi, L., and Sun, J., *A Nonsmooth Version of Newton's Method*, Mathematical Programming, Vol. 58, pp. 353–367, 1993.
8. Fischer, A., *Solution of Monotone Complementarity Problems with Locally Lipschitzian Functions*, Mathematical Programming, Vol. 76, pp. 513–532, 1997.
9. Chen, C., and Mangasarian, O. L., *A Class of Smoothing Functions for Nonlinear and Mixed Complementarity Problems*, Computational Optimization and Applications, Vol. 5, pp. 97–138, 1996.
10. Qi, H. D., *A Regularized Smoothing Newton Method for Box-Constrained Variational Inequality Problems with P_0-Functions*, SIAM Journal on Optimization, Vol. 10, pp. 251–264, 2000.
11. Qi, L., Sun, D., and Zhou, G., *A New Look at Smoothing Newton Methods for Nonlinear Complementarity Problems and Box-Constrained Variational Inequalities*, Mathematical Programming, Vol. 87, pp. 1–35, 2000.
12. Sun, D., *A Regularization Newton Method for Solving Nonlinear Complementarity Problems*, Applied Mathematics and Optimization, Vol. 40, pp. 315–339, 1999.
13. Chen, B., and Harker, P. T., *A Noninterior-Point Continuation Method for Linear Complementarity Problems*, SIAM Journal on Matrix Analysis and Applications, Vol. 14, pp. 1168–1190, 1993.