Semismooth Newton Methods for Minimizing a Sum of Euclidean Norms with Linear Constraints

Guanglu Zhou* Kim-Chuan Toh† Defeng Sun‡

August 1, 2001

Abstract. We consider the problem of minimizing a sum of Euclidean norms with linear constraints. In this paper we establish the optimality conditions, duality and uniqueness results for the problem. It is shown that the problem has a dual problem with duality theory very similar to that in linear programming. We then present two semismooth Newton methods for the problem by solving the strongly semismooth equations derived from the optimality conditions. The methods are globally and quadratically convergent. In particular, the quadratic convergence of the methods are proved without requiring strict complementarity. Preliminary numerical results are also reported.

Key Words. sum of norms, nonsmooth optimization, duality, uniqueness, semismooth equation, convergence

AMS(MOS) subject classifications. 90C33, 90C30, 65H10.

1 Introduction

Consider the problem of minimizing a sum of Euclidean norms with constraints:

$$\begin{array}{ll} \min & \sum_{i=1}^{m} \|a_i - A_i^T x\| \\ \text{s.t.} & x \in X, \end{array} \tag{1.1}$$

where $a_1, a_2, \cdots, a_m \in \Re^d$, $A_1, A_2, \cdots, A_m \in \Re^{n\times d}$, and $X \subseteq \Re^n$.

When $X = \Re^n$, (1.1) is the problem of minimizing a sum of Euclidean norms (the MSN problem for abbreviation) studied in [1, 3, 19, 24, 30]. If $X := \{x \in \Re^n : B_e^T x = b_e\}$, then (1.1) is the problem of minimizing a sum of Euclidean norms subject to linear equality

*The Singapore-MIT Alliance, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Republic of Singapore. E-mail: smazgl@nus.edu.sg.

†Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Republic of Singapore. E-mail: mattohkc@nus.edu.sg. Research supported in part by the Singapore-MIT alliance.

‡Department of Mathematics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Republic of Singapore. E-mail: matsundf@nus.edu.sg.

constraints (the EMSN problem for abbreviation) [2]. In this paper, unless otherwise stated, we assume that

$$X := \{x \in \Re^n : B_e^T x = b_e,\ B^T x \geq b\} \tag{1.2}$$

is a nonempty set, where $B_e \in \Re^{n\times l}$, $b_e \in \Re^l$, $B \in \Re^{n\times \nu}$ and $b \in \Re^\nu$. Let $A = [A_1, A_2, \cdots, A_m]$, $a = (a_1; a_2; \cdots; a_m)$, $M = \{1, 2, \cdots, m\}$ and $V = \{1, 2, \cdots, \nu\}$. In this paper, we make the following assumptions.

(A1) A has full row rank n;

(A2) Let $\mathcal{S}$ be the solution set of problem (1.1). For any $x \in \mathcal{S}$, define

$$I(x) = \{i : B_i^T x = b_i, i \in V\} \quad \text{and} \quad B_{I(x)} = [B_i, i \in I(x)].$$

The matrix $[B_e, B_{I(x)}]$ has full column rank.

Let

$$f(x) = \sum_{i=1}^{m} f_i(x), \tag{1.3}$$

where $f_i(x) = \|a_i - A_i^T x\|$ for $i = 1, 2, \cdots, m$. If $d = 1$, then the problem (1.1) is equivalent to a linear program [11]. The problem (1.1) is a convex programming problem, but its objective function f is not differentiable at any point x satisfying $a_i - A_i^T x = 0$ for some i. It arises in many applications, such as the VLSL design, the Euclidean facilities location problem and the Steiner minimal tree problem under a given topology; e.g., see [3, 30, 31] for details.

Many algorithms have been designed to solve the MSN problem and the EMSN problem. For the MSN problem, Weiszfeld [29] gave a simple iterative algorithm in 1937. Later, a number of important results were obtained along this line; see [6, 7, 12, 16, 18, 26, 27]. Practical algorithms for solving these problems began with the work of Calamai and Conn [4, 5], and Overton [19], where they proposed projected Newton algorithms with the quadratic rate of convergence. More recently, Andersen [1] presented a Newton barrier method for solving this problem. In [2], Andersen and Christiansen developed an algorithm for the EMSN problem by transforming it into a sequence of unconstrained MSN problems based on l_1 penalty function approach. Andersen, Christiansen, Conn and Overton [3] proposed a primal-dual interior-point method based on smoothing $\|x\|$ to $\sqrt{\|x\|^2 + \varepsilon^2}$ and presented impressive computational results. Xue and Ye [30, 31] presented polynomial time primal-dual potential reduction algorithms by transforming this problem into a standard second-order cone problem. Qi and Zhou [24] presented a globally and quadratically convergent method by solving a system of strongly semismooth equations derived from the problem.

In [5, 19, 24], the global and quadratic convergence results were obtained under a strict complementarity assumption. However, this condition is not satisfied for many MSN problems. Therefore, one question arises: can one design a globally and quadratically convergent method for the problem without assuming strict complementarity ? The main goal of this paper is to answer this question.

In this paper we establish the optimality conditions, duality and uniqueness results for the problem (1.1). The problem is shown to have a dual problem with duality theory very similar to that in linear programming. The optimality conditions of the problem are reformulated as strongly semismooth equations. We then present two semismooth Newton methods for the problem by solving these strongly semismooth equations. The methods are globally and quadratically convergent. In particular, the quadratic convergence of the methods are proved without requiring strict complementarity.

The rest of the paper is organized as follows. In section 2, we present the optimality conditions, duality theory and uniqueness results for (1.1). In section 3, we reformulate the optimality conditions as two systems of strongly semismooth equations based on the Robinson's normal map [25], and the Fisher-Burmeister function [14]. We then present a smoothing Newton method to solve the semismooth equation derived from the Robinson's map in section 4 and a generalized Newton method to solve the other semismooth equation in section 5. In section 6, we report our preliminary numerical results. We conclude the paper in section 7.

1.1 Notation and Terminology

For a continuously differentiable function $F : \Re^n \to \Re^m$, we denote the Jacobian of F at $x \in \Re^n$ by $F'(x)$, whereas the transposed Jacobian as $\nabla F(x)$. In particular, if $m = 1$, the gradient $\nabla F(x)$ is viewed as a column vector.

Let $F : \Re^n \to \Re^m$ be a locally Lipschitzian vector function. By Rademacher's theorem, F is differentiable almost everywhere. Let Ω_F be the set of points where F is differentiable. Then the B-subdifferential of F at $x \in \Re^n$ is defined to be

$$\partial_B F(x) = \{ \lim_{\substack{x^k \to x \\ x^k \in \Omega_F}} \nabla F(x^k)^T \}, \tag{1.4}$$

while Clarke's generalized Jacobian of F at x is defined to be

$$\partial F(x) = \text{conv}\partial_B F(x); \tag{1.5}$$

see [10, 20, 23] for more details. The function F is said to be *semismooth* at x if F is directionally differentiable at x and for all $V \in \partial F(x+h)$ and $h \to 0$,

$$F(x+h) - F(x) = Vh + o(\|h\|). \tag{1.6}$$

F is said to be *p-order semismooth*, $p \in (0, 1]$, at x if F is semismooth at x and for all $V \in \partial F(x+h)$ and $h \to 0$,

$$F(x+h) - F(x) = Vh + O(\|h\|^{1+p}). \tag{1.7}$$

F is said to be *strongly semismooth* at x if F is 1-order semismooth at x. The function F is termed a (strongly) semismooth function if it is (strongly) semismooth everywhere (see [20, 23]). Here, $o(\|h\|)$ stands for a vector function $e : \Re^n \to \Re^m$, satisfying

$$\lim_{h \to 0} \frac{e(h)}{\|h\|} = 0;$$

while $O(\|h\|^2)$ stands for a vector function $e : \Re^n \to \Re^m$ such that for some constants $M > 0$ and $\delta > 0$,

$$\|e(h)\| \leq M\|h\|^2$$

for all h satisfying $\|h\| \leq \delta$.

Let $\Phi : \Re^n \to \Re^n$ be a locally Lipschitzian vector function. We say that Φ is *BD-regular* at a point x if all $V \in \partial_B \Phi(x)$ is nonsingular and Φ is *CD-regular* at a point x if all $V \in \partial \Phi(x)$ is nonsingular.

We denote the cardinality of a set $\mathcal{A}$ by $|\mathcal{A}|$. 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_+. To represent a large matrix with several small matrices, we use semicolons ";" for column concatenation and commas "," for row concatenation. This notation also applies to vectors. We let I_d denote the $d \times d$ identity matrix. Given a finite number of square matrices $Q_1, \cdots, Q_n$, we denote the block diagonal matrix with these matrices as block diagonals by $\mathrm{diag}(Q_1, \cdots, Q_n)$, or by $\mathrm{diag}(Q_i, i = 1, \cdots, n)$. We let $\Re_+ = \{\varepsilon \in \Re : \varepsilon \geq 0\}$ and $\Re_{++} = \{\varepsilon \in \Re : \varepsilon > 0\}$. Finally, we use $\varepsilon \downarrow 0^+$ to denote the case that a positive scalar ε tends to 0.

2 Optimality Conditions, Duality and Uniqueness Theorems

Duality for the MSN and EMSN problems has been studied in [1, 3, 19]. However, there are few results on duality for (1.1). In this section we will establish the optimality conditions, duality and uniqueness results of for (1.1). The following lemma will be used later. We shall omit its proof since it is easy.

Lemma 2.1 *Let* $f(x) = \|x\|,\ x \in \Re^n$. *Then*

(i) $y \in \partial f(x)$ *if and only if* $\|y\| \leq 1$ and $\|x\| - y^T x = 0$;

(ii) If $x \neq 0$, *then* $\partial f(x) = \left\{ \frac{x}{\|x\|} \right\}$.

Lemma 2.2 *Assuming that* **(A1)** *and* **(A2)** *hold, then the solution set* $\mathcal{S}$ *of the problem (1.1) is bounded, and* $x \in \mathcal{S}$ *if and only if there exist* $g \in \Re^l$ *and* $h \in \Re^\nu$ *such that*

$$\begin{cases} 0 \in \displaystyle\sum_{i=1}^{m} \partial f_i(x) - B_e g - Bh, \\ B_e^T x = b_e, \\ (B^T x - b)^T h = 0,\ h \geq 0,\ B^T x \geq b. \end{cases} \tag{2.1}$$

The solution set of (2.1) is also bounded.

Proof. It follows from $(\mathbf{A1})$ that

$$\min_{\|x\|=1} \|A^T x\| = \tau > 0. \tag{2.2}$$

From (2.2) we obtain

$$\|A^T x\| \geq \tau \|x\|. \tag{2.3}$$

Now

$$\sum_{i=1}^{m} \|a_i - A_i^T x\| \;\geq\; \left(\sum_{i=1}^{m} \|a_i - A_i^T x\|\right)^{1/2} = \|a - A^T x\|$$

$$\geq\; \|A^T x\| - \|a\| \geq \tau\|x\| - \|a\|.$$

The above inequality shows that the solution set of (1.1) is bounded. It follows from (**A1**), (**A2**) and Theorems 2.1.4, 2.2.5 and 2.3.2 in [15] that $x \in \mathcal{S}$ if and only if there exist $g \in \Re^l$ and $h \in \Re^\nu$ such that system (2.1) holds and the solution set of (2.1) is bounded. ∎

From Lemma 2.1, we have for $i \in M$,

$$\partial f_i(x) = \{-A_i y_i : y_i \in \Re^d, \|y_i\| \leq 1, \|a_i - A_i^T x\| - (a_i - A_i^T x)^T y_i = 0\}, \tag{2.4}$$

and if $a_i - A_i^T x \neq 0$, then $y_i = (a_i - A_i^T x)/\|a_i - A_i^T x\|$. Thus (2.1) is equivalent to the following system.

$$\begin{cases} -Ay - B_e g - Bh = 0, \\ (B^T x - b)^T h = 0, \quad h \geq 0, \quad B^T x \geq b, \\ B_e^T x = b_e, \\ \|y_i\| \leq 1, \quad i \in M, \\ \|a_i - A_i^T x\| - (a_i - A_i^T x)^T y_i = 0, \quad i \in M, \end{cases} \tag{2.5}$$

where $y = (y_1; y_2; \cdots; y_m)$. The dual of problem (1.1) is as follow:

$$\begin{array}{ll} \max & a^T y + b_e^T g + b^T h \\ \text{s.t.} & (y; g; h) \in X_D, \end{array} \tag{2.6}$$

where

$$X_D := \{(y; g; h) \in \Re^{md+l+\nu} : Ay + B_e g + Bh = 0, h \geq 0, \|y_i\| \leq 1, i \in M\}.$$

Theorem 2.1 *Let $x \in X$, $(y; g; h) \in X_D$ and let x^*, $(y^*; g^*; h^*)$ be optimal solutions to problems (1.1) and (2.6), respectively. Then*

$$a^T y + b_e^T g + b^T h \leq f(x) \quad \text{(weak duality)}$$

and

$$a^T y^* + b_e^T g^* + b^T h^* = f(x^*) \quad \text{(strong duality)}.$$

Proof. Let $x \in X$ and $(y; g; h) \in X_D$. Then

$$x^T A y + x^T B_e g + x^T B h = 0.$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}
a^T y + b_e^T g + b^T h &= (a - A^T x)^T y + (b_e - B_e^T x)^T g + (b - B^T x)^T h \\
&\le \sum_{i=1}^{m} (a_i - A_i^T x)^T y_i \\
&\le \sum_{i=1}^{m} \|a_i - A_i^T x\| = f(x).
\end{aligned}$$

This establishes the weak duality.

Let x^* be an optimal solution to problem (1.1). It follows Lemma 2.2 and (2.5) that there exist $y \in \Re^{md}$, $g \in \Re^l$ and $h \in \Re^\nu$ such that

$$\begin{cases}
-Ay - B_e g - Bh = 0, \\
(B^T x^* - b)^T h = 0, \quad h \ge 0, \quad B^T x^* \ge b, \\
B_e^T x^* = b_e, \\
\|y_i\| \le 1, \quad i \in M, \\
\|a_i - A_i^T x^*\| - (a_i - A_i^T x^*)^T y_i = 0, \quad i \in M.
\end{cases}$$

Thus $(y; g; h) \in X_D$, and

$$\begin{aligned}
a^T y + b_e^T g + b^T h &= (a - A^T x^*)^T y + (b_e - B_e^T x^*)^T g + (b - B^T x^*)^T h \\
&= \sum_{i=1}^{m} (a_i - A_i^T x^*)^T y_i \\
&= \sum_{i=1}^{m} \|a_i - A_i^T x^*\| = f(x^*),
\end{aligned}$$

which proves that the strong duality holds. ∎

Let x^* and $(y^*; g^*; h^*)$ be optimal solutions to problems (1.1) and (2.6), respectively. We call x^* and $(y^*; g^*; h^*)$ a pair of optimal solutions to problems (1.1) and (2.6) if x^* and $(y^*; g^*; h^*)$ satisfy system in (2.5). We said that *strict complementarity* holds at $(x^*; y^*; g^*; h^*)$ if

(i) for $\forall\ i \in M$, $\|y_i^*\| < 1$ whenever $b_i - A_i^T x^* = 0$;

(ii) $(B^T x^* - b)_j + h_j^* \neq 0$ for $j \in V$.

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

$$G(x^*) = \sum_{i \in M \setminus M_0(x^*)} \nabla^2 f_i(x^*), \tag{2.7}$$

where

$$\nabla^2 f_i(x^*) = \frac{1}{\|a_i - A_i^T x^*\|} A_i A_i^T - \frac{1}{\|a_i - A_i^T x^*\|^3} A_i (a_i - A_i^T x^*)(a_i - A_i^T x^*)^T A_i^T.$$

Let x^* and $(y^*; g^*; h^*)$ be a pair of optimal solutions to problems (1.1) and (2.6). Define

$$\mathcal{I} = \{i : B_i^T x^* > b_i, i \in V\}, \tag{2.8}$$

$$\mathcal{J} = \{i : B_i^T x^* = b_i \text{ and } h_i^* = 0, i \in V\}, \tag{2.9}$$

$$\mathcal{K} = \{i : B_i^T x^* = b_i \text{ and } h_i^* > 0, i \in V\}, \tag{2.10}$$

and

$$D(x^*) = \{d \in \Re^n : B_e^T d = 0, B_i^T d = 0 \text{ for } i \in \mathcal{K}\}. \tag{2.11}$$

We assume that

(A3) For all $d \in D(x^*)\backslash\{0\}$, $d^T G(x^*) d > 0$;

(A4) The matrix $[A_0, B_e, B_i, i \in \mathcal{J} \cup \mathcal{K}]$ has full column rank.

Theorem 2.2 *Let x^* and $(y^*; g^*; h^*)$ be a pair of optimal solutions to problems (1.1) and (2.6). If x^* and $(y^*; g^*; h^*)$ satisfy* **(A3)** *and* **(A4)***, then x^* is a unique solution to problem (1.1).*

Proof. Suppose that $x' \neq x^*$ is another solution to problem (1.1). Let $d = x' - x^*$. Then it is readily shown that for all $t \in [0, 1]$, $x(t) = x^* + td$ is a solution of the problem by using the convexity of f (thus, $\sum_{i=1}^m \|a_i - A_i^T x^*\| = \sum_{i=1}^m \|a_i - A_i^T x(t)\|$), and $B_e^T d = 0$ and $B_i^T d \geq 0$ for $i \in \mathcal{J} \cup \mathcal{K}$. From Lemma 2.2 and (2.5), there exist $y(t) \in \Re^{md}$, $g(t) \in \Re^l$ and $h(t) \in \Re^\nu$ such that

$$\begin{cases} -Ay(t) - B_e g(t) - Bh(t) = 0, \\ (B^T x(t) - b)^T h(t) = 0, \quad h(t) \geq 0, \quad B^T x(t) \geq b, \\ B_e^T x(t) = b_e, \\ \|y_i(t)\| \leq 1, \quad i \in M, \\ \|a_i - A_i^T x(t)\| - (a_i - A_i^T x(t))^T y_i(t) = 0, \quad i \in M. \end{cases} \tag{2.12}$$

Next we prove by contradiction that $(y(t), g(t), h(t)) \to (y^*, g^*, h^*)$ as $t \to 0$. There exists $\delta \in (0, 1)$ such that for $t \in [0, \delta]$, $B_i^T x(t) > b_i$ for $i \in \mathcal{I}$ and $\|a_i - A_i^T x(t)\| \neq 0$ for $i \in M\backslash M_0(x^*)$. Therefore, when $t \in [0, \delta]$, $h_i(t) = 0$ for $i \in \mathcal{I}$ and $y_i(t) = (a_i - A_i^T x(t))/\|a_i - A_i^T x(t)\|$ for $i \in M\backslash M_0(x^*)$. Thus, $h_i(t) \to h_i^*$ for $i \in \mathcal{I}$ and $y_i(t) \to y_i^*$ for $i \in M\backslash M_0(x^*)$ as $t \to 0$. Hence, from **(A4)** and the first equation in (2.12), $\{(y(t), g(t), h(t)), t \in [0, \delta]\}$ is bounded. Suppose that there exists a sequence $\{t^k\}$ in $[0, \delta]$ such that $t^k \to 0$ and

$(y(t^k), g(t^k), h(t^k)) \to (\bar{y}, \bar{g}, \bar{h}) \neq (y^*, g^*, h^*)$ as $k \to +\infty$. From the first equation in (2.12) and the fact that (y^*, g^*, h^*) satisfies the first equation in (2.5), we have

$$\begin{aligned} &\sum_{i \in M_0(x^*)} A_i(y_i(t^k) - y_i^*) + \sum_{i \in M \backslash M_0(x^*)} A_i(y_i(t^k) - y_i^*) + B_e(g(t^k) - g^*) \\ &+ \sum_{i \in \mathcal{I}} B_i(h_i(t^k) - h_i^*) + \sum_{i \in \mathcal{J} \cup \mathcal{K}} B_i(h_i(t^k) - h_i^*) = 0. \end{aligned} \tag{2.13}$$

Letting $k \to +\infty$ and noting that $y_i(t^k) \to y_i^*$ for $i \in M \backslash M_0(x^*)$ and $h_i(t^k) = h_i^* = 0$ for $i \in \mathcal{I}$, we have

$$\sum_{i \in M_0(x^*)} A_i(\bar{y}_i - y_i^*) + B_e(\bar{g} - g^*) + \sum_{i \in \mathcal{J} \cup \mathcal{K}} B_i(\bar{h}_i - h_i^*) = 0.$$

This contradicts **(A4)**. Thus, $(y(t), g(t), h(t)) \to (y^*, g^*, h^*)$ as $t \to 0$. Note that from (2.12), $B_i^T d = 0$ for $i \in \mathcal{K}$. Hence, $d \in D(x^*)$. For $i \in M \backslash M_0(x^*)$, using Taylor expansion,

$$\begin{aligned} &\|a_i - A_i^T(x^* + td)\| - \|a_i - A_i^T x^*\| \\ = \;& -t\left[(a_i - A_i^T x^*)^T / \|a_i - A_i^T x^*\|\right](A_i^T d) + \tfrac{1}{2} t^2 d^T \nabla^2 f_i(x^*) d + o(t^2) \\ = \;& -t(y_i^*)^T(A_i^T d) + \tfrac{1}{2} t^2 d^T \nabla^2 f_i(x^*) d + o(t^2). \end{aligned}$$

Now for a sufficiently small $t > 0$,

$$\begin{aligned} & f(x^* + td) - f(x^*) \\ = \;& \sum_{i=1}^{m} \left(\|a_i - A_i^T(x^* + td)\| - \|a_i - A_i^T x^*\|\right) \\ = \;& \sum_{i \in M_0(x^*)} t\|A_i^T d\| + \sum_{i \in M \backslash M_0(x^*)} \left(-t(y_i^*)^T(A_i^T d) + \frac{1}{2} t^2 d^T \nabla^2 f_i(x^*) d + o(t^2)\right) \\ = \;& \sum_{i \in M_0(x^*)} t\|A_i^T d\| - \sum_{i \in M \backslash M_0(x^*)} t(y_i^*)^T(A_i^T d) \;+\; \frac{1}{2} t^2 d^T G(x^*) d + o(t^2) \\ = \;& \sum_{i \in M_0(x^*)} t\|A_i^T d\| - \sum_{i \in M \backslash M_0(x^*)} t(A_i y_i^*)^T d \;+\; \frac{1}{2} t^2 d^T G(x^*) d + o(t^2) \\ = \;& \sum_{i \in M_0(x^*)} t\left(\|A_i^T d\| + (y_i^*)^T(A_i^T d)\right) + t d^T(B_e g^* + B h^*) + \frac{1}{2} t^2 d^T G(x^*) d + o(t^2), \\ \geq \;& \tfrac{1}{2} t^2 d^T G(x^*) d + o(t^2) \\ > \;& 0, \text{ by } \textbf{(A3)}. \end{aligned}$$

The fifth equation above follows from the fact that (y^*, g^*, h^*) satisfies the first equation in (2.5). However, $f(x^* + td) = f(x^*)$ for all $t \in [0, 1]$. This contradiction shows that x^* is a unique solution to (1.1). ∎

Remark 1. For MSN problems studied in [5, 19], assumptions **(A3)** and **(A4)** reduce to the following conditions.

(i) The matrix $G(x^*)$ is nonsingular;

(ii) A_0 has full column rank.

The conditions (i) and (ii) are weaker than the conditions of Theorem 8.2 in [5] and Theorem 3 in [19]. Moreover, in [5, 19], to guarantee a solution of the problem to be a strong (isolated) minimizer, the strict complementarity assumption is needed.

Remark 2. For the Euclidean single facility location (ESFL) problem [17], conditions (i) and (ii) in Remark 1 hold. Let $a_1, a_2, \cdots, a_m$ be m $(m > 2)$ distinct points in $\Re^2$. Let $\omega_1, \omega_2, \cdots, \omega_m$ be m positive weights. Find a point $x \in \Re^2$ that minimizes

$$f(x) = \sum_{i=1}^{m} \omega_i \|x - a_i\|. \tag{2.14}$$

This is called the ESFL problem. This problem can be easily transformed into a special case of problem (1.1) where $b_i = \omega_i a_i$, $A_i^T = \omega_i I_2$, $i = 1, 2, \cdots, m$, and $X = \Re^2$. Clearly, if a_i, $i = 1, \cdots, m$ are collinear, then $\min_{x \in X} f(x) = \min\{f(a_i) : 1 \le i \le m\}$. In what follows, we always assume that a_i, $i = 1, \cdots, m$ are not collinear. For any $x \in \Re^2$, let $M_0(x) = \{i \in M : \|x - a_i\| = 0\}$. Clearly, $M_0(x) = \emptyset$ or $M_0(x)$ only has one element. This means (ii) holds. Define $f_i(x) = \omega_i \|x - a_i\|$, $i = 1, 2, \cdots, m$. It is easy to prove for any $x \in \Re^2$, $G(x) = \sum_{i \in M \setminus M_0(x)} \nabla^2 f_i(x)$ is positive definite. This shows (i) holds.

3 Reformulation

In this section we reformulate the optimality conditions (2.5) as two systems of strongly semismooth equations based on the Robinson's normal map [25] and the Fisher-Burmeister function [14]. These reformulations are important to our design of quadratically convergent algorithms.

For $i \in M$, it is readily shown that $\|a_i - A_i^T x\| - (a_i - A_i^T x)^T y_i = 0$ and $\|y_i\| \le 1$ if and only if y_i is the solution to the following problem.

$$\begin{array}{ll} \text{maximize} & (a_i - A_i^T x)^T s \\ \text{subject to} & \|s\| \le 1, \\ & s \in \Re^d. \end{array} \tag{3.1}$$

Thus there exist $\lambda_i \in \Re$ and $t_i \in \Re$ such that

$$\begin{array}{rcll} A_i^T x - a_i + \lambda_i y_i & = & 0, & \lambda_i \ge 0, \\ \frac{1}{2} - \frac{1}{2}\|y_i\|^2 & = & t_i, & t_i \ge 0, \\ \lambda_i t_i & = & 0. & \end{array} \tag{3.2}$$

From (3.2), if $A_i^T x - a_i \neq 0$ then $\lambda_i = \|a_i - A_i^T x\|$ and $y_i = (a_i - A_i^T x)/\|a_i - A_i^T x\|$. Therefore (2.5) is equivalent to the following system:

$$\begin{cases} -Ay - B_e g - Bh = 0, \\ (B^T x - b)^T h = 0, \quad B^T x \geq b, \quad h \geq 0, \\ B_e^T x = b_e, \\ A_i^T x - a_i + \lambda_i y_i = 0, \quad i \in M, \\ \frac{1}{2} - \frac{1}{2}\|y_i\|^2 = t_i, \quad i \in M, \\ \lambda_i t_i = 0, \quad t_i \geq 0, \quad \lambda_i \geq 0, \quad i \in M. \end{cases} \tag{3.3}$$

Let

$$\Lambda = \mathrm{diag}(\lambda_i I_d, i \in M), \tag{3.4}$$

$\lambda = (\lambda_1; \cdots; \lambda_m)$, $e_m = (1; \cdots; 1) \in \Re^m$, and $S = (\|y_1\|^2; \cdots; \|y_m\|^2)$. Define $F : \Re^q \to \Re^q$, where $q = n + md + l + \nu + m$, by

$$F(x, y, g, h, \lambda) \;=\; \begin{pmatrix} -Ay - B_e g - Bh \\ A^T x - a + \Lambda y \\ B_e^T x - b_e \\ B^T x - b \\ \frac{1}{2} e_m - \frac{1}{2} S \end{pmatrix}, \tag{3.5}$$

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

$$\Omega = \{(x; y; g; h; \lambda) \in \Re^q : h \geq 0, \ \lambda \geq 0\}. \tag{3.6}$$

Let $u = (x; y; g; h; \lambda)$. Then the system (3.3) is equivalent to the following box variational inequality problem: find a vector $u^* = (x^*; y^*; g^*; h^*; \lambda^*) \in \Omega$ such that

$$F(u^*)^T (u - u^*) \geq 0, \ \forall \ u \in \Omega. \tag{3.7}$$

Lemma 3.1 *F is a smooth monotone mapping in Ω. Assuming that (A1) and (A2) hold, then the set of solutions to (3.7) is bounded.*

Proof. Let

$$J(y) = \mathrm{diag}(y_i, i \in M). \tag{3.8}$$

The Jacobian matrix of F is given by

$$F'(x, y, g, h, \lambda) = \begin{pmatrix} 0 & -A & -B_e & -B & 0 \\ A^T & \Lambda & 0 & 0 & J(y) \\ B_e^T & 0 & 0 & 0 & 0 \\ B^T & 0 & 0 & 0 & 0 \\ 0 & -J(y)^T & 0 & 0 & 0 \end{pmatrix}. \tag{3.9}$$

Clearly, F' is positive semi-definite in Ω as F' is the sum of a skew-symmetric matrix and a diagonal matrix with nonnegative diagonals. Therefore, F is a smooth monotone mapping in Ω. Moreover, from Lemma 2.2 and (3.3) the set of solutions to (3.7) is bounded. ∎

Let $z = (x; y; g; r; s) \in \Re^q$ and $\Pi_\Omega(z)$ be the Euclidean projection of z onto Ω. It is well known that solving (3.7) is equivalent to solving the following Robinson's normal equation

$$E(z) := F(\Pi_\Omega(z)) + z - \Pi_\Omega(z) = 0 \tag{3.10}$$

in the sense that if z^* is a solution of (3.10), then $\Pi_\Omega(z^*) = (x^*; y^*; g^*; r^*_+; s^*_+)$ is a solution of (3.7), and conversely if u^* is a solution of (3.7) then $z^* = u^* - F(u^*)$ is a solution of (3.10); see [25].

Lemma 3.2 *The function E is strongly semismooth. Assuming that (A1) and (A2) hold, then the set of zeros of E is bounded.*

Proof. It is readily shown that the function $\Pi_\Omega(\cdot)$ is strongly semismooth. It follows from Theorem 19 in [14] that E is strongly semismooth. From Lemma 3.1 the set of zeros of E is bounded. ∎

The system (3.3) can be reformulated into another equivalent system via the Fischer-Burmeister function [14]

$$\psi(a, b) = \sqrt{a^2 + b^2} - a - b. \tag{3.11}$$

The main property of this function is the following characterization of its zeros:

$$\psi(a, b) = 0 \iff a \geq 0, \ b \geq 0, \ ab = 0.$$

The system that is equivalent to (3.3) is as follows:

$$\Phi(u) \ := \ \Phi(x; y; g; h; \lambda) \ = \ 0, \tag{3.12}$$

where Φ is defined by

$$\Phi(u) = \Phi(x; y; g; h; \lambda) := \begin{pmatrix} Ay + B_e g + Bh \\ \\ A^T x - a + \Lambda y \\ \\ B_e^T x - b_e, \\ \\ \psi(h, B^T x - b) \\ \\ \psi(\lambda, e_m - S) \end{pmatrix}, \tag{3.13}$$

and

$$\begin{aligned} \psi(h, B^T x - b) &= (\psi(h_1, B_1^T x - b_1); \cdots; \psi(h_\nu, B_\nu^T x - b_\nu)) \in \Re^\nu, \\ \psi(\lambda, e_m - S) &= (\psi(\lambda_1, 1 - \|y_1\|^2); \cdots; \psi(\lambda_m, 1 - \|y_m\|^2)) \in \Re^m. \end{aligned}$$

Note that the function ψ is not differentiable at the origin, so Φ is nonsmooth. However, from the fact that ψ is strongly semismooth, we have that Φ is strongly semismooth.

In recent years, globally and superlinearly (or quadratically) convergent Newton methods have been developed for solving the semismooth equations; see [9, 13, 21, 20, 23, 22, 28, 32]. In the next two sections we will apply the smoothing Newton method proposed by Qi, Sun and Zhou in [22] to solving the semismooth equation (3.10) and use the generalized damped Newton method in [13] to solve (3.12).

4 The Smoothing Newton Method

In [8], Chen and Mangasarian presented a class of smooth approximations to the function $s_+ = \max\{0, s\}$. Among these smooth approximations, the Chen-Harker-Kanzow-Smale smooth function is most commonly used. It is defined by

$$\phi(t,s) = \frac{s + \sqrt{s^2 + 4t^2}}{2}, \ (t,s) \in \Re_{++} \times \Re. \tag{4.1}$$

Proposition 4.1 *The function $\phi(t,s)$ has the following properties:*

(i) For any given $t > 0$, $\phi(t,s)$ is continuously differentiable;

(ii) $0 \le \phi(t,s) - s_+ \le t$;

(iii) $0 < \phi'_s(t,s) < 1$.

Proof. See [8]. ■

Define $p : \Re^2 \to \Re$ by

$$p(t,s) = \begin{cases} \phi(|t|, s) & \text{if } t \neq 0, \\ \max\{0, s\} & \text{if } t = 0. \end{cases} \tag{4.2}$$

From Proposition 1 in [22], Theorem 19 in [14] and Proposition 4.1, we have the following proposition.

Proposition 4.2 *p has the following properties:*

(i) p is a strongly semismooth function on $\Re^2$;

(ii) For any given $t \neq 0$, $p(t,\cdot)$ is continuously differentiable;

(iii) $p(t,s) > 0$, for any given $t \neq 0$;

(iv) For any $s \in \Re$ and $t > 0$, $0 < \partial p(t,s)/\partial s < 1$;

(v) If $s^* > 0$, then

$$\lim_{\substack{t^k \downarrow 0^+ \\ s^k \to s^*}} \partial p(t,s)/\partial s = 1;$$

(vi) If $s^* < 0$, then

$$\lim_{\substack{t^k \downarrow 0^+ \\ s^k \to s^*}} \partial p(t,s)/\partial s = 0.$$

Let $v := (t; z) = (t; x; y; g; r; s)$ and

$$P(v) = (x; y; g; p(t, r_1); \cdots; p(t, r_\nu); p(t, s_1); \cdots; p(t, s_m)).$$

Define H by

$$H(v) := \begin{pmatrix} t \\ F(P) + (1+t)z - P \end{pmatrix}, \tag{4.3}$$

where $P = P(v)$ and F is defined as in (3.5). From Propositions 4.2, H is continuously differentiable on $(\Re\setminus\{0\})\times\Re^q$ and strongly semismooth on $\Re\times\Re^q$. Clearly, solving (3.10) is equivalent to finding the zeros of H. Let

$$P'_t = (0_n; 0_{md}; 0_l; C_t; D_t) \quad \text{and} \quad P'_z = \operatorname{diag}(I_n, I_{md}, I_l, C_r, D_s),$$

where

$$C_t = (\partial p(t, r_1)/\partial t; \cdots; \partial p(t, r_\nu)/\partial t), \quad D_t = (\partial p(t, s_1)/\partial t; \cdots; \partial p(t, s_m)/\partial t), \tag{4.4}$$
$$C_r = \operatorname{diag}(\partial p(t, r_i)/\partial r_i, i \in V) \quad \text{and} \quad D_s = \operatorname{diag}(\partial p(t, s_i)/\partial s_i, i \in M).$$

Define

$$P_s = \operatorname{diag}(p(t, s_i)I_d, i \in M).$$

Then we have the following lemma.

Lemma 4.1 *For any $v = (t, z) \in \Re_{++}\times\Re^q$, the Jacobian of H is given by*

$$H'(v) := \begin{pmatrix} 1 & 0 \\ (F'(P) - I)P'_t + z & F'(P)P'_z + (1+t)I_q - P'_z \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -BC_t + tx & tI_n & -A & -B_e & -BC_r & 0 \\ JD_t + ty & A^T & P_s + tI_{md} & 0 & 0 & JD_s \\ tg & B_e^T & 0 & tI_l & 0 & 0 \\ tr - tC_t & B^T & 0 & 0 & (1+t)I_\nu - C_r & 0 \\ ts - tD_t & 0 & -J^T & 0 & 0 & (1+t)I_m - D_s \end{pmatrix}, \tag{4.5}$$

where $J = J(y)$ is defined as in (3.8), and $H'(v)$ is nonsingular.

Proof. It is readily shown that (4.5) holds by simple computation. For any $v = (t, z) \in \Re_{++}\times\Re^q$, in order to prove $H'(v)$ is nonsingular, we only need to prove that

$$W := F'(P)P'_z + (1+t)I_q - P'_z$$

is nonsingular. Let $N = (1+t)I_q - P'_z$. From Proposition 4.2, P'_z and N are diagonal matrices with positive diagonals. Because $F'(P)$ is positive semi-definite, W is nonsingular. Therefore, $H'(v)$ is nonsingular. This completes the proof. ■

Define the merit function Ψ by

$$\Psi(v) := \|H(v)\|^2$$

and let

$$\beta(v) := \gamma \min\{1, \Psi(v)\}.$$

Now we will describe the smoothing Newton method for the zeros of $H(v)$.

Algorithm 4.1

Step 0. *Choose* $\bar{t} \in \Re_{++}$ *and* $\gamma \in (0,1)$ *such that* $\gamma\bar{t} < 1$. *Let* $\bar{v} := (\bar{t}; 0_q) \in \Re \times \Re^q$. *Choose constants* $\delta \in (0,1)$, *and* $\sigma \in (0, \frac{1}{2})$. *Let* $t^0 := \bar{t}$, z^0 *be an arbitrary initial point in* $\Re^q$. *Set* $v^0 = (t^0; z^0)$ *and* $k := 0$.

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*

$$H(v^k) + H'(v^k)\Delta v^k = \beta_k \bar{v}. \tag{4.6}$$

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

$$\Psi(v^k + \delta^l \Delta v^k) \ \leq \ [1 - 2\sigma(1 - \gamma\bar{t}\ )\delta^l]\Psi(v^k). \tag{4.7}$$

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

Step 4. *Replace* k *by* $k+1$ *and go to Step 1.*

Let $v^* = (0, z^*)$ and define

$$A(v^*) = \{\lim H'(t^k, z^k) : t^k \downarrow 0^+ \text{ and } z^k \to z^*\}. \tag{4.8}$$

Clearly, $A(v^*) \subseteq \partial_B H(v^*)$. It follows from Lemmas 3.2 and 4.1 that we have the following theorems. The proofs of the theorems are similar to that of Theorems 4.6 and 4.8 in [24], and we shall omit them.

Theorem 4.1 *(i) An infinite sequence* $\{v^k = (t^k, z^k)\}$ *is generated by Algorithm 4.1, and*

$$\lim_{k\to+\infty} H(v^k) = 0, \text{ and } \lim_{k\to+\infty} t^k = 0. \tag{4.9}$$

(ii) The sequence $\{v^k\}$ *is bounded. Hence there exists at least an accumulation point, say* $v^* = (0, z^*)$, *of* $\{v^k\}$ *such that* z^* *is a solution to (3.10).*

Theorem 4.2 *Suppose that* v^* *is an accumulation point of an infinite sequence* $\{v^k\}$ *generated by Algorithm 4.1. Assume that all* $W \in A(v^*)$ *are nonsingular. Then the whole sequence* $\{v^k\}$ *converges to* v^* *quadratically, i.e.,*

$$\|v^{k+1} - v^*\| = O(\|v^k - v^*\|^2), \tag{4.10}$$

and

$$\Psi(v^{k+1}) = O(\Psi(v^k)^2). \tag{4.11}$$

Let $v^* = (0; x^*; y^*; g^*; r^*; s^*)$ be a solution of $H(v) = 0$. Then x^*, $(y^*; g^*; r^*_+)$ are a pair of optimal solutions to problems (1.1) and (2.6). Recall the notation $M_0(x^*)$ used in section 2, without loss of generality, we suppose that $\|a_i - A_i^T x^*\| = 0$ for $i = 1, \cdots, j$, where $j = |M_0(x^*)|$, and $\|a_i - A_i^T x^*\| > 0$ for $i = j+1, \cdots, m$. From (3.5), (3.10) and (4.3), we have for $i \in M$,

$$A_i^T x^* - a_i + (s_i^*)_+ y_i^* = 0,$$

and

$$s_i^* = (s_i^*)_+ - \frac{1}{2}(1 - \|y_i^*\|^2).$$

Then $s_i^* \le 0, i = 1, \cdots, j$ and $s_i^* > 0, i = j+1, \cdots, m$. Furthermore, $s_i^* = \|a_i - A_i^T x^*\|$ and $y_i^* = (a_i - A_i^T x^*)/\|a_i - A_i^T x^*\|$ for $i = j+1, \cdots, m$. If $a_i - A_i^T x^* = 0$ and $\|y_i^*\| < 1$, then $s_i^* < 0$. Define

$$A_0 = [A_1, \cdots, A_j], \quad \text{and} \quad \bar{A} = [A_j, \cdots, A_m].$$

Then the Hessian matrix $G(x^*)$ of the function $\sum_{i=j+1}^m f_i(x)$ at x^* is as follows.

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

where

$$P_1 = \text{diag}(s_i^* I_d, i = j+1, \cdots, m), \tag{4.12}$$

and

$$J_1 = \text{diag}(y_i^*, i = j+1, \cdots, m). \tag{4.13}$$

From (3.5), (3.10) and (4.3),

$$B^T x^* - b = r_+^* - r^*.$$

Then $B_i^T x^* > b_i$ and $r_i^* < 0$ if $i \in \mathcal{I}$, $B_i^T x^* = b_i$ and $r_i^* = 0$ if $i \in \mathcal{J}$, and $B_i^T x^* = b_i$ and $r_i^* > 0$ if $i \in \mathcal{K}$. Here, B_i is the ith row of B and $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{K}$ are defined as in (2.8)-(2.10).

Proposition 4.3 *Suppose that $v^* = (0; x^*; y^*; g^*; r^*; s^*)$ is a solution of $H(v) = 0$ and $(x^*; y^*; g^*; r_+^*)$ satisfies* **(A3)** *and* **(A4)**. *Then all $W \in A(v^*)$ are nonsingular.*

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

$$W = \lim_{k \to +\infty} H'(v^k) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -BC_t^* & 0 & -A & -B_e & -BC_r^* & 0 \\ J^* D_t^* & A^T & P_s^* & 0 & 0 & J^* D_s^* \\ 0 & B_e^T & 0 & 0 & 0 & 0 \\ 0 & B^T & 0 & 0 & I_\nu - C_r^* & 0 \\ 0 & 0 & -(J^*)^T & 0 & 0 & I_m - D_s^* \end{pmatrix},$$

where $J^* = J(y^*)$ is defined as in (3.8), C_t^* and D_t^* are defined as in (4.4), and

$$P_s^* = \begin{pmatrix} 0 & 0 \\ 0 & P_1 \end{pmatrix}, \quad D_s^* = \begin{pmatrix} D_1 & 0 \\ 0 & I_{m-j} \end{pmatrix}, \quad C_r^* = \text{diag}(0_{|\mathcal{I}|}, C_1, I_{|\mathcal{K}|}).$$

Here

$$D_1 = \text{diag}(d_i, i = 1, \cdots, j), \ d_i \in [0, 1], \quad \text{and} \quad C_1 = \text{diag}(c_i, i \in \mathcal{J}), \ c_i \in [0, 1].$$

Without loss of generality, we suppose that $c_i \in (0, 1)$ for all $i \in \mathcal{J}$. Otherwise, let $\mathcal{I} := \mathcal{I} \cup \{i \in \mathcal{J} : c_i = 0\}$, $\mathcal{K} := \mathcal{K} \cup \{i \in \mathcal{J} : c_i = 1\}$, and $\mathcal{J} := \mathcal{J} \backslash \{i \in \mathcal{J} : c_i = 1 \text{ or } c_1 = 0\}$. Let $B = [B_\mathcal{I}, B_\mathcal{J}, B_\mathcal{K}]$ and $\bar{C}_1 = \text{diag}(1 - c_i, i \in \mathcal{J})$. We also suppose that $d_i = 0$ for $i = 1, \cdots, n_1$, $d_i \in (0, 1)$ for $i = n_1 + 1, \cdots, n_2$ and $d_i = 1$ for $i = n_2 + 1, \cdots, j$.

Then we have $\|y_i\| = 1$ for $i = n_1 + 1, \cdots, j$. Let $N = \mathrm{diag}(d_i, i = n_1 + 1, \cdots, n_2)$ and $\bar{N} = \mathrm{diag}(1 - d_i, i = n_1 + 1, \cdots, n_2)$. Then $D_1 = \mathrm{diag}(0, N, I_{j-n_2})$. Define

$$J_{01} = \mathrm{diag}(y_i^*, i = 1, \cdots, n_1), \quad J_{02} = \mathrm{diag}(y_i^*, i = n_1 + 1, \cdots, n_2), \tag{4.14}$$

and

$$J_{03} = \mathrm{diag}(y_i^*, i = n_2 + 1, \cdots, j). \tag{4.15}$$

Let $A_{01} = [A_1, \cdots, A_{n_1}]$, $A_{02} = [A_{n_1+1}, \cdots, A_{n_2}]$ and $A_{03} = [A_{n_2+1}, \cdots, A_j]$. Then $A_0 = [A_{01}, A_{02}, A_{03}]$. Let $\mathcal{W}$ be the following matrix

$$\begin{pmatrix} 0 & -A_{01} & -A_{02} & -A_{03} & -\bar{A} & -B_e & 0 & -B_{\mathcal{J}}C_1 & -B_{\mathcal{K}} & 0 & 0 & 0 & 0 \\ A_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_{02}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & NJ_{02} & 0 & 0 \\ A_{03}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{03} & 0 \\ A_p^T & 0 & 0 & 0 & P_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_1 \\ B_e^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ B_{\mathcal{I}}^T & 0 & 0 & 0 & 0 & 0 & I_{|\mathcal{I}|} & 0 & 0 & 0 & 0 & 0 & 0 \\ B_{\mathcal{J}}^T & 0 & 0 & 0 & 0 & 0 & 0 & \bar{C}_1 & 0 & 0 & 0 & 0 & 0 \\ B_{\mathcal{K}}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -J_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & -J_{02}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{N} & 0 & 0 \\ 0 & 0 & 0 & -J_{03}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -J_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To prove that W is nonsingular, it suffices to show that $\mathcal{W}$ is nonsingular. Let $\mathcal{W}q = 0$, where $q = (q_1; \cdots; q_{13})$. Here, $q_1 \in \Re^n$, $q_2 \in \Re^{n_1 d}$, $q_3 \in \Re^{(n_2-n_1)d}$, $q_4 \in \Re^{(j-n_2)d}$, $q_5 \in \Re^{(m-j)d}$, $q_6 \in \Re^l$, $q_7 \in \Re^{|\mathcal{I}|}$, $q_8 \in \Re^{|\mathcal{J}|}$, $q_9 \in \Re^{|\mathcal{K}|}$, $q_{10} \in \Re^{n_1}$, $q_{11} \in \Re^{n_2-n_1}$, $q_{12} \in \Re^{j-n_2}$ and $q_{13} \in \Re^{m-j}$. Then we have

$$A_{01}q_2 + A_{02}q_3 + A_{03}q_4 + \bar{A}q_5 + B_e q_6 + B_{\mathcal{J}}C_1 q_8 + B_{\mathcal{K}}q_9 = 0, \tag{4.16}$$

$$A_{01}^T q_1 = 0, \tag{4.17}$$

$$A_{02}^T q_1 + NJ_{02}q_{11} = 0, \tag{4.18}$$

$$A_{03}^T q_1 + J_{03}q_{12} = 0, \tag{4.19}$$

$$\bar{A}^T q_1 + P_1 q_5 + J_1 q_{13} = 0, \tag{4.20}$$

$$B_e^T q_1 = 0, \tag{4.21}$$

$$B_{\mathcal{I}}^T q_1 + q_7 = 0, \tag{4.22}$$

$$B_{\mathcal{J}}^T q_1 + \bar{C}_1 q_8 = 0, \tag{4.23}$$

$$B_{\mathcal{K}}^T q_1 = 0, \tag{4.24}$$

$$-J_{01}^T q_2 + q_{10} = 0, \tag{4.25}$$

$$-J_{02}^T q_3 + \bar{N} q_{11} = 0, \tag{4.26}$$

$$-J_{03}^T q_4 = 0 \tag{4.27}$$

and

$$J_1^T q_5 = 0. \tag{4.28}$$

It follows from (4.20) that

$$q_5 = -P_1^{-1}\bar{A}^T q_1 - P_1^{-1}J_1 q_{13}. \tag{4.29}$$

From (4.28) and (4.29),

$$J_1^T P_1^{-1}\bar{A}^T q_1 = -J_1^T P_1^{-1} J_1 q_{13} = -\Lambda_1^{-1} q_{13},$$

where $\Lambda_1 = \mathrm{diag}(s_i^*, i = j+1, \cdots, m)$. Then we have

$$q_{13} = -\Lambda_1 J_1^T P_1^{-1}\bar{A}^T q_1 = -J_1^T \bar{A}^T q_1. \tag{4.30}$$

Pre-multiplying (4.29) by $\bar{A}$, and applying (4.30), 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.,

$$\bar{A} q_5 + G(x^*) q_1 = 0. \tag{4.31}$$

Pre-multiplying (4.16) by q_1^T, and applying (4.17), (4.21), (4.23) and (4.24), we have

$$q_1^T A_{02} q_3 + q_1^T A_{03} q_4 + q_1^T \bar{A} q_5 - q_8^T \bar{C}_1 C_1 q_8 = 0. \tag{4.32}$$

Pre-multiplying (4.19) by q_4^T and applying (4.27),

$$q_4^T A_{03}^T q_1 = -q_4^T J_{03} q_{12} = 0. \tag{4.33}$$

From (4.26), we have

$$q_{11} = \bar{N}^{-1} J_{02}^T q_3. \tag{4.34}$$

Pre-multiplying (4.18) by q_3^T and applying (4.34),

$$q_3^T A_{02}^T q_1 = -q_3^T N J_{02} q_{11} = -q_3^T N J_{02}\bar{N}^{-1} J_{02}^T q_3 = -q_3^T N\bar{N}^{-1} q_3. \tag{4.35}$$

From (4.31), (4.33), (4.35) and (4.32),

$$q_1^T G(x^*) q_1^T + q_3^T N\bar{N}^{-1} q_3 + q_8^T \bar{C}_1 C_1 q_8 = 0, \tag{4.36}$$

where

$$\begin{aligned} N\bar{N}^{-1} &= \mathrm{diag}(d_i/(1-d_i), i = n_1+1, \cdots, n_2), \\ \bar{C}_1 C_1 &= \mathrm{diag}(c_i(1-c_i), i \in \mathcal{J}). \end{aligned}$$

By (4.21) and (4.24), $q_1 \in D(x^*)$. From (4.36) and **(A3)**, we have $q_1 = 0$, $q_3 = 0$ and $q_8 = 0$ since $N\bar{N}^{-1}$ and $\bar{C}_1 C_1$ are positive definite. Hence $q_{13} = 0$ from (4.30), $q_5 = 0$ from (4.29), $q_7 = 0$ from (4.22), $q_{11} = 0$ from (4.34) and $q_{12} = 0$ from (4.19). By (4.16) and **(A4)**, $q_2 = 0, q_4 = 0, q_6 = 0$, and $q_9 = 0$. From (4.25), $q_{10} = 0$. Thus $q = 0$. This implies that $\mathcal{W}$ is nonsingular and the proof is completed. ∎

By combining Theorem 4.2 and Proposition 4.3 we can directly obtain the following result.

Theorem 4.3 *Suppose* $v^* = (0; x^*; y^*; g^*; r^*; s^*)$ *is an accumulation point of the infinite sequence* $\{v^k\}$ *generated by Algorithm 4.1 and* $(x^*; y^*; g^*; r^*_+)$ *satisfies* **(A3)** *and* **(A4)**. *Then the whole sequence* $\{v^k\}$ *converges to* v^* *quadratically.*

5 The Generalized Damped Newton Method

In this section we present a generalized damped Newton method for solving $\Phi(u) = 0$, where Φ is defined in (3.13). Let

$$\Theta(u) = \frac{1}{2}\Phi(u)^T\Phi(u). \tag{5.1}$$

By Theorem 2.6.6 and the Corollary to Theorem 2.2.4 in [10], θ is continuously differentiable and

$$\nabla\Theta(u) = W^T\Phi(u), \tag{5.2}$$

for every $W \in \partial\Phi(u)$. For any $u = (x; y; g; h; \lambda) \in \Re^q$, let

$$\begin{aligned} I_0(u) &= \{i : B_i^T x = b_i \text{ and } h_i = 0, i \in V\}, \\ J_0(u) &= \{i : \|y_i\| = 1 \text{ and } \lambda_i = 0, i \in M\}. \end{aligned}$$

Define

$$W = \begin{pmatrix} 0 & A & B_e & B & 0 \\ A^T & \Lambda & 0 & 0 & J \\ B_e^T & 0 & 0 & 0 & 0 \\ D_1B^T & 0 & 0 & D_2 & 0 \\ 0 & -2D_3J^T & 0 & 0 & D_4 \end{pmatrix}, \tag{5.3}$$

where Λ and $J = J(y)$ are defined as in (3.4) and (3.8) respectively. Here $D_1 = \text{diag}(\alpha_1, \cdots, \alpha_\nu)$, $D_2 = \text{diag}(\beta_1, \cdots, \beta_\nu)$, $D_3 = \text{diag}(\eta_1, \cdots, \eta_m)$ and $D_4 = \text{diag}(\theta_1, \cdots, \theta_m)$ are diagonal matrices whose diagonal elements are given by

$$\alpha_i = \frac{B_i^T x - b_i}{\sqrt{h_i^2 + (B_i^T x - b_i)^2}} - 1, \quad \beta_i = \frac{h_i}{\sqrt{h_i^2 + (B_i^T x - b_i)^2}} - 1$$

if $i \in V \backslash I_0(u)$, and $\alpha_i = -1$, $\beta_i = 0$ if $i \in I_0(u)$;

$$\eta_i = \frac{1 - \|y_i\|^2}{\sqrt{\lambda_i^2 + (1 - \|y_i\|^2)^2}} - 1, \quad \theta_i = \frac{\lambda_i}{\sqrt{\lambda_i^2 + (1 - \|y_i\|^2)^2}} - 1$$

if $i \in M \backslash J_0(u)$, and $\eta_i = -1$, $\theta_i = 0$ if $i \in J_0(u)$. We have the following lemma.

Lemma 5.1 *The matrix* W *defined in (5.3) is an element of* $\partial_B\Phi(u)$.

Proof. We shall build a sequence of points $\{u^k = (x^k; y^k; g^k; h^k; \lambda^k)\}$ such that Φ is differentiable at u^k and $\Phi'(u^k)$ converges to W. Let $\{t^k\}$ is a sequence of positive numbers converging to 0. Define $x^k = x$, $y^k = y$, $g^k = g$,

$$h_i^k = \begin{cases} h_i & i \in V \backslash I_0(u), \\ t^k & i \in I_0(u), \end{cases} \quad \text{and} \quad \lambda_i^k = \begin{cases} \lambda_i & i \in M \backslash J_0(u), \\ t^k & i \in J_0(u). \end{cases}$$

Then Φ is differentiable at u^k, and for $i \in I_0(u^k)$, $\alpha_i^k = -1$ and $\beta_i^k = 0$, and for $i \in J_0(u^k)$, $\eta_i^k = -1$ and $\theta_i^k = 0$. Therefore, $\Phi'(u^k)$ tends to W as $k \to +\infty$. ■

Algorithm 5.1

Step 0. *Choose* $u^0 \in \Re^q$, $\sigma \in (0, 1/2)$, $\rho \in (0,1), \eta > 0, a > 2$. *Set* $k = 0$.

Step 1. *If* $\nabla\Theta(u^k) = 0$, *stop. Otherwise, select an element* $W^k \in \partial_B \Phi(u^k)$ *and find a solution* Δu^k *of the linear system*

$$\Phi(u^k) + W^k \Delta u = 0. \tag{5.4}$$

If (5.4) is not solvable, or if $\nabla\Theta(u^k)^T \Delta u^k > -\eta \|\Delta u^k\|^a$, *set* $\Delta u^k = -\nabla\Theta(u^k)$.

Step 2. *Let* $\alpha_k = \rho^{j_k}$, *where* j_k *is the smallest nonnegative integer* j *such that*

$$\Theta(u^k + \rho^j \Delta u^k) - \Theta(u^k) \le \sigma\rho^j \nabla\Theta(u^k)^T \Delta u^k.$$

Step 3. *Set* $u^{k+1} := u^k + \alpha_k \Delta u^k$ *and* $k := k+1$. *Go to Step 1.*

This algorithm is a generalization of the corresponding algorithm for the Fischer-Burmeister equation in [13]. Just like Theorem 11 of [13], we have the following global and superlinear convergence theorem for this algorithm. We omit its proof since it is similar to the one in [13].

Theorem 5.1 *Suppose an infinite sequence of points* $\{u^k\}$ *is generated by Algorithm 5.1. Then each accumulation point of* $\{u^k\}$ *is a stationary point of* Θ. *Moreover, if the sequence* $\{u^k\}$ *has an accumulation point* u^* *which is an isolated solution of* $\Phi(u) = 0$, *then the entire sequence* $\{u^k\}$ *converges to* u^*.

Theorem 5.2 *Suppose* $u^* = (x^*; y^*; g^*; h^*; \lambda^*)$ *is an accumulation point of* $\{u^k\}$ *generated by Algorithm 5.1. If* u^* *is a BD-regular solution of* $\Phi(u) = 0$, *then the entire sequence* $\{u^k\}$ *converges to* u^* *and the rate of convergence is Q-quadratic.*

Proposition 5.1 *Suppose* $u^* = (x^*; y^*; g^*; h^*; \lambda^*)$ *is a solution of* $\Phi(u) = 0$ *and it satisfies* **(A3)** *and* **(A4)**. *Then* Φ *is BD-regular at* u^*.

Proof. Let $u^* = (x^*; y^*; g^*; h^*; \lambda^*)$ is a solution of $\Phi(u) = 0$. Without loss of generality, we suppose that $\|a_i - A_i^T x^*\| = 0$ for $i = 1, \cdots, j$, and $\|a_i - A_i^T x^*\| > 0$ for $i = j+1, \cdots, m$. Then from (3.3), $\lambda_i^* = 0$ for $i = 1, \cdots, j$, and $y_i^* = (a_i - A_i^T x^*)/\|a_i - A_i^T x^*\|$ and $\lambda_i^* = \|a_i - A_i^T x^*\|$ for $i = j+1, \cdots, m$. Let

$$A_0 = [A_1, \cdots, A_j] \quad \text{and} \quad \bar{A} = [A_j, \cdots, A_m].$$

Then the Hessian matrix $G(x^*)$ of the function $\sum_{i=j+1}^m f_i(x)$ at x^* is as follows.

$$G(x^*) = \bar{A}{\Lambda_1}^{-1}(I_{(m-j)d} - J_1 J_1^T)\bar{A}^T,$$

where $\Lambda_1 = \text{diag}(\lambda_i^* I_d, i = j+1, \cdots, m)$ and J_1 is defined as in (4.13).

For any $W \in \partial_B \Phi(u^*)$, W can be written as follows.

$$W = \begin{pmatrix} 0 & A & B_e & B & 0 \\ A^T & \Lambda^* & 0 & 0 & J^* \\ B_e^T & 0 & 0 & 0 & 0 \\ D_1^* B^T & 0 & 0 & D_2^* & 0 \\ 0 & -2D_3^*(J^*)^T & 0 & 0 & D_4^* \end{pmatrix}, \tag{5.5}$$

where $\Lambda^* = \mathrm{diag}(0, \Lambda_1)$, $J^* = J(y^*)$ is defined as in (3.8), and $D_1^* = \mathrm{diag}(\alpha_1^*, \cdots, \alpha_\nu^*)$, $D_2^* = \mathrm{diag}(\beta_1^*, \cdots, \beta_\nu^*)$, $D_3^* = \mathrm{diag}(\eta_1^*, \cdots, \eta_m^*)$ and $D_4^* = \mathrm{diag}(\theta_1^*, \cdots, \theta_m^*)$ are diagonal matrices. Here,

$$\alpha_i^* = \frac{B_i^T x^* - b_i}{\sqrt{(h_i^*)^2 + (B_i^T x^* - b_i)^2}} - 1, \quad \beta_i^* = \frac{h_i^*}{\sqrt{(h_i^*)^2 + (B_i^T x^* - b_i)^2}} - 1$$

if $h_i^* \neq 0$ or $B_i^T x^* > b_i$; α_i^* and β_i^* satisfy $(\alpha_i^* + 1)^2 + (\beta_i^* + 1)^2 \leq 1$ if $h_i^* = 0$ and $B_i^T x^* = b_i$. For $i = 1, 2, \cdots, j$, if $\|y_i^*\| < 1$, then $\eta_i^* = 0$ and $\theta_i^* = -1$, and if $\|y_i^*\| = 1$, then η_i^* and θ_i^* satisfy $(\eta_i^* + 1)^2 + (\theta_i^* + 1)^2 \leq 1$. For $i = j+1, \cdots, m$, $\eta_i^* = -1$ and $\theta_i^* = 0$. Define

$$\begin{aligned} \mathcal{R} &:= \{i : \alpha_i^* = 0, \beta_i^* = -1, i \in V\}, \\ \mathcal{S} &:= \{i : \alpha_i^*, \beta_i^* \in (-1, 0), i \in V\}, \\ \mathcal{T} &:= \{i : \alpha_i^* = -1, \beta_i^* = 0, i \in V\}. \end{aligned}$$

Then $\mathcal{K} \subseteq \mathcal{T} \subseteq \mathcal{J} \cup \mathcal{K}$, where $\mathcal{J}$ and $\mathcal{K}$ are defined in (2.9) and (2.10) respectively. We write B as $B = [B_{\mathcal{R}}, B_{\mathcal{S}}, B_{\mathcal{T}}]$. Let $D_1^* = \mathrm{diag}(0, \mathcal{A}, -I_{|\mathcal{T}|})$, where $\mathcal{A} = \mathrm{diag}(\alpha_i^*, i \in \mathcal{S})$, and $D_2^* = \mathrm{diag}(-I_{|\mathcal{R}|}, \mathcal{B}, 0)$, where $\mathcal{B} = \mathrm{diag}(\beta_i^*, i \in \mathcal{S})$. Without loss of generality, we suppose that $\eta_i^* = 0$ and $\theta_i^* = -1$ for $i = 1, \cdots, n_1$, $\eta_i^*, \theta_i^* \in (-1, 0)$ for $i = n_1 + 1, \cdots, n_2$, and $\eta_i = -1$ and $\theta_i^* = 0$ for $i = n_2 + 1, \cdots, j$. Let $D_3^* = \mathrm{diag}(0, \mathcal{C}, -I_{j-n_2}, -I_{m-j})$, where $\mathcal{C} = \mathrm{diag}(\eta_i^*, i = n_1 + 1, \cdots, n_2)$, and $D_4^* = \mathrm{diag}(-I_{n_1}, \Gamma, 0, 0)$, where $\Gamma = \mathrm{diag}(\theta_i^*, i = n_1 + 1, \cdots, n_2)$. Let J_{01}, J_{02} and J_{03} be defined as in (4.14) and (4.15), $A_{01} = [A_1, \cdots, A_{n_1}]$, $A_{02} = [A_{n_1+1}, \cdots, A_{n_2}]$ and $A_{03} = [A_{n_2+1}, \cdots, A_j]$. Then W can be rewritten as follows:

$$\begin{pmatrix} 0 & A_{01} & A_{02} & A_{03} & \bar{A} & B_e & B_{\mathcal{R}} & B_{\mathcal{S}} & B_{\mathcal{T}} & 0 & 0 & 0 & 0 \\ A_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{01} & 0 & 0 & 0 \\ A_{02}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{02} & 0 & 0 \\ A_{03}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_{03} & 0 \\ A_p^T & 0 & 0 & 0 & \Lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & J_1 \\ B_e^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -I_{|\mathcal{R}|} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{A}B_{\mathcal{S}}^T & 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{B} & 0 & 0 & 0 & 0 & 0 \\ -B_{\mathcal{T}}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & -2\mathcal{C}J_{02}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Gamma & 0 & 0 \\ 0 & 0 & 0 & 2J_{03}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2J_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is now easy to show that W is nonsingular by following the steps in proving $\mathcal{W}$ is nonsingular in Proposition 4.3. ∎

By combining Theorem 5.2 and Proposition 5.1 we can directly obtain the following result.

Theorem 5.3 *Suppose $u^* = (x^*; y^*; g^*; h^*; \lambda^*)$ is an accumulation point of $\{u^k\}$ generated by Algorithm 5.1. If u^* is a solution of $\Phi(u) = 0$ and it satisfies* **(A3)** *and* **(A4)***, then the entire sequence $\{u^k\}$ converges to u^* and the rate of convergence is Q-quadratic.*

6 Preliminary Numerical Experiments

We implemented Algorithm 4.1 in MATLAB and tested the method on the following examples.

Example 1. This is a SMT problem [30]. The starting point is $z^0 = \mathbf{0}$, where $\mathbf{0}$ is the vector whose elements are all zero.

Example 2. This is a multi-facility location problem [19]. The starting point is $z^0 = \mathbf{0}$.

The following two examples are generated randomly. We use the following pseudo-random sequence:

$$\psi_0 = 7, \psi_{i+1} = (445\psi_i + 1) \bmod 4096, i = 1, 2, \cdots,$$

$$\bar{\psi}_i = \frac{\psi_i}{4096}, i = 1, 2, \cdots.$$

Example 3. The elements of $A_i, i = 1, 2, \cdots, m$, and those of $a_i, i = 1, 2, \cdots, m$, are successively set to $\bar{\psi}_1, \bar{\psi}_2, \cdots$, in the order:

$$(A_1)_{11}, (A_1)_{21}, \cdots, (A_1)_{n1}, (A_1)_{12}, \cdots, (A_1)_{nd}, \cdots, (A_m)_{nd},$$

$$(a_1)_1, \cdots, (a_1)_d, (a_2)_1, \cdots, (a_m)_d,$$

except that the appropriate random number is multiplied by 100 to give $(A_i)_{jk}$ or $(a_i)_j$ if $i \bmod 10 = 1$. The starting point is $z^0 = \mathbf{0}$.

Example 4. Same as Example 3 except with $X = \{x \in \Re^n : x \geq 0\}$. The starting point is $z^0 = \mathbf{0}$.

Throughout our computational experiments, we used the following parameters:

$$\delta = 0.5, \ \sigma = 0.0005, \ \bar{t} = 0.001, \ \text{ and } \ \gamma = 0.5.$$

We terminated the iteration when $\|E(z^k)\|_\infty \leq$ 1e-6, where E is defined in (3.10). The numerical results we obtained are summarized in Table 1. In this table, **n**, **d** and **m** specify the problem dimensions, **It** denotes the number of iterations, which is also equal to the number of Jacobian evaluations for the function H, **NH** denotes the number of function

Example	n	d	m	It	NH	$f(x^k)$	$\|E(z^k)\|_\infty$	CPU
1	16	2	17	9	17	2.54e+01	1.12e-10	0.54
2	10	2	55	13	25	2.26e+02	9.77e-11	3.34
3	10	2	100	7	17	6.86e+01	7.55e-07	6.54
	10	2	200	9	18	1.46e+02	4.55e-08	23.92
	10	2	400	9	19	2.83e+02	1.75e-07	89.34
	10	2	600	10	19	4.38e+02	8.06e-12	200.46
	10	2	800	10	20	5.74e+02	1.22e-07	365.22
	10	2	1000	10	20	7.17e+02	1.87e-08	564.95
4	10	2	100	30	97	6.89e+01	1.47e-08	36.98
	10	2	200	43	118	1.47e+02	2.35e-08	151.00
	10	2	400	27	68	2.83e+02	1.10e-09	315.06
	10	2	600	20	41	4.38e+02	4.28e-07	426.64
	10	2	800	26	40	5.75e+02	4.42e-09	769.52
	10	2	1000	12	22	7.17e+02	3.12e-07	657.52

Table 1: Numerical results for Algorithm 4.1

evaluations for the function H, $f(x^k)$ denotes the value of $f(x)$ at the final iteration, $\|E(z^k)\|_\infty$ denotes the value of $\|E(z)\|_\infty$ at the final iteration, and **CPU** denotes the CPU time in second for solving the problem.

The results reported in Table 1 show that the algorithm performs well. The algorithm was able to solve all examples after a small number of iterations. Moreover, from the results of Examples 3 and 4, the number of iterations does not seem to increase with the number of constraints.

7 Conclusion

In this paper we established the optimality conditions, duality and uniqueness results for the problem of minimizing a sum of Euclidean norms with linear constraints and presented two semismooth Newton methods for the problem by solving the strongly semismooth equations derived from the optimality conditions. The methods were globally and quadratically convergent. In particular, the quadratic convergence of the methods were proved without requiring strict complementarity.

We would point out that if specialized to MSN problems, Algorithm 4.1 appears to have some theoretical advantages over some existing methods in [5], [19] and [24]. In particular, we can prove the quadratic convergence under conditions **(A3)** and **(A4)** which is weaker than the conditions used in [5] and [19] to prove quadratic convergence. For Algorithm 5.1, there remains some issues to be studied, for example, under what conditions is the sequence generated by the algorithm bounded and when is a stationary point of Θ a zero of $\Phi(u)$? 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 4.1 lies in solving the linear system (4.6). Thus, it is necessary for us to come up with ways to solve it efficiently by exploiting sparsity or special structures present in the problem data. However, we shall leave these as further research topics.

Acknowledgements

The first author would like to thank Professor Liqun Qi for his encouragement, K.D. Andersen and E. Christiansen for providing references [2] and [3].

References

[1] K.D. Andersen, An efficient Newton barrier method for minimizing a sum of Euclidean norms, *SIAM Journal on Optimization* 6 (1996) 74–95.

[2] K.D. Andersen and E. Christiansen, Minimizing a sum of norms subject to linear equality constraints, *Computational Optimization and Applications* 11 (1998) 65–79.

[3] K.D. Andersen, E. Christiansen, A.R. Conn and M.L. Overton, An efficient primal-dual interior-point method for minimizing a sum of Euclidean norms, *SIAM Journal on Scientific Computing* 22 (2000) 243–262.

[4] P.H. Calamai and A.R. Conn, A stable algorithm for solving the multifacility location problem involving Euclidean distances, *SIAM Journal on Scientific and Statistical Computing* 1 (1980) 512–526.

[5] P.H. Calamai and A.R. Conn, A projected Newton method for l_p norm location problems, *Mathematical Programming* 38 (1987) 75–109.

[6] R. Chandrasekaran and A. Tamir, Open questions concerning Weiszfeld's algorithm for the Fermat-Weber location problem, *Mathematical Programming* 44 (1989) 293–295.

[7] R. Chandrasekaran and A. Tamir, Algebraic optimization: the Fermat-Weber location problem, *Mathematical Programming* 46 (1990) 219–224.

[8] C. Chen and O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Computational Optimization and Applications* 5 (1996) 97–138.

[9] X. Chen, L. Qi, and D. Sun, Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities, *Mathematics of Computation* 67 (1998) 519-540.

[10] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[11] T.F. Coleman and Y. Li, A globally and quadratically convergent affine scaling method for linear ℓ_1 problems, *Mathematical Programming* 56 (1992) 189–222.

[12] J.W. Eyster, J.A. White and W.W. Wierwille, On solving multifacility location problems using a hyperboloid approximation procedure, *AIIE Transactions* 5 (1973) 1–6.

[13] T. De Luca, F. Facchinei and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, *Mathematical Programming* 75 (1996) 407–439.

[14] A. Fischer, Solution of monotone complementarity problems with locally Lipschitzian functions, *Mathematical Programming* 76 (1997) 513–532.

[15] J.B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithm*, Springer-Verlag Berlin Heidelberg, 1993.

[16] H.W. Kuhn, A note on Fermat's problem, *Mathematical Programming* 4 (1973) 98–107.

[17] R.F. Love, J.G. Morris and G.O. Wesolowsky, Facility Location: Models and Methods, North-Holland, Amsterdam, 1988.

[18] L.M.M. Ostresh, The multifacility location problem: applications and decent theorems, *Journal of Regional Science* 17 (1977) 409–419.

[19] M.L. Overton, A quadratically convergent method for minimizing a sum of Euclidean norms, *Mathematical Programming* 27 (1983) 34–63.

[20] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research* 18 (1993) 227–244.

[21] L. Qi and H. Jiang, Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton methods and quasi-Newton methods for solving these equations, *Mathematics of Operations Research* 22 (1997) 301-325.

[22] L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, *Mathematical Programming* 87(2000) 1–35.

[23] L. Qi and J. Sun, A nonsmooth version of Newton's method, *Mathematical Programming* 58 (1993) 353–367.

[24] L. Qi and G. Zhou, A smoothing Newton method for minimizing a sum of Euclidean norms, *SIAM Journal on Optimization* 11 (2000) 389–410.

[25] S.M. Robinson, Normal maps induced by linear transformation, *Mathematics of Operations Research* 17 (1992) 691–714.

[26] J.B. Rosen and G.L. Xue, On the convergence of Miehle's algorithm for the Euclidean multifacility location problem, *Operations Research* 40 (1992) 188–191.

[27] J.B. Rosen and G.L. Xue, On the convergence of a hyperboloid approximation procedure for solving the perturbed Euclidean multifacility location problem, *Operations Research* 41 (1993) 1164–1171.

[28] D. Sun, A regularization Newton method for solving nonlinear complementarity problems, *Applied Mathematics and Optimization* 40 (1999) 315–339.

[29] E. Weiszfeld, Sur le point par lequel la somme des distances de n points donnes est minimum, *Tohoku Mathematical Journal* 43 (1937) 355–386.

[30] G. Xue and Y. Ye, An efficient algorithm for minimizing a sum of Euclidean norms with applications, *SIAM Journal on Optimization* 7 (1997) 1017–1036.

[31] G. Xue and Y. Ye, An efficient algorithm for minimizing a sum of P-norms, *SIAM Journal on Optimization* 10 (2000) 551–579.

[32] N. Yamashita and M. Fukushima, Modified Newton methods for solving semismooth reformulations of monotone complementarity problems, *Mathematical Programming* 76 (1997) 273–284.