

A modified SQP algorithm for minimax problems[☆]Qing-jie Hu^{a,*}, Yu Chen^a, Nei-ping Chen^a, Xue-quan Li^b^a Department of Information, Hunan Business College, 410205, Changsha, PR China^b Department of Mathematics, Hunan First Normal College, 410205, Changsha, PR China

ARTICLE INFO

Article history:

Received 27 August 2007

Available online 10 June 2009

Submitted by J.A. Filar

Keywords:

Minimax problems

SQP algorithm

Nonmonotone line search

Global convergence

Superlinear convergence

ABSTRACT

In this paper, a modified nonmonotone line search SQP algorithm for nonlinear minimax problems is presented. During each iteration of the proposed algorithm, a main search direction is obtained by solving a reduced quadratic program (QP). In order to avoid the Maratos effect, a correction direction is generated by solving the reduced system of linear equations. Under mild conditions, the global and superlinear convergence can be achieved. Finally, some preliminary numerical results are reported.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

We consider the following minimax problem:

$$(P) \quad \min_{x \in R^n} F(x), \quad (1.1)$$

where $F(x) = \max\{f_j(x), j \in I\}$, $I = \{1, 2, \dots, m\}$, and $f_j(x) : R^n \rightarrow R$ is continuously differentiable. Problem (1.1) has strong practical background. It arises in many engineering design problems (see Refs. [1–4]).

Since the objective function $F(x)$ is nondifferentiable even when the $f_i(x)$, $i \in I$, are all differentiable, the classical methods for smooth optimization problems may fail to reach an optimum if they are applied directly to the nonlinear minimax problem. To overcome this difficulty, many of the methods that have been proposed for solving minimax problems are based on the following equivalent translation of the original problem (1.1):

$$(P') \quad \begin{aligned} & \min_{(x,z) \in R^{n+1}} z, \\ & \text{s.t. } f_j(x) \leq z, \quad j \in I. \end{aligned} \quad (1.2)$$

Obviously, the KKT conditions of (1.2) can be stated as follows:

$$\begin{aligned} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \sum_{j \in I} \lambda_j \begin{pmatrix} \nabla f_j(x) \\ -1 \end{pmatrix} = 0, \\ & \lambda_j \geq 0, \quad f_j(x) - z \leq 0, \quad \lambda_j(f_j(x) - z) = 0, \quad j \in I, \end{aligned}$$

[☆] Project supported by Hunan Provincial Natural Science Foundation of China (05JJ40103), Soft Science Research Fund of Hunan Province of China (2006ZK3028) and Scientific Research Fund of Hunan Provincial Education Department of China (105B0707, 08C470).

* Corresponding author.

E-mail address: hqj0525@126.com (Q.-j. Hu).

and these relationships are equivalent to

$$\sum_{j \in I} \lambda_j \nabla f_j(x) = 0, \quad \sum_{j \in I} \lambda_j = 1, \quad \lambda_j (f_j(x) - F(x)) = 0, \quad \lambda_j \geq 0, \quad j \in I, \quad (1.3)$$

where a point $x \in R^n$ is called as a stationary point of (P) (Ref. [5]) and λ is said to be a multiplier vector.

It is well known that the Sequential Quadratic Programming (SQP) method has satisfactory convergence, and it is one of the most effective algorithms for solving nonlinearly constrained optimization problems (see Refs. [6,8,10–13]). So several authors have extended the popular SQP technique to the minimax problems (see Refs. [14–21]). Among them, Zhou and Tits [19] proposed an algorithm: the search direction is generated by solving two quadratic programs, and avoiding the Maratos effect by means of nonmonotone line search. However, it obtains only two-step superlinear convergence. Recently, some SQP algorithms are also proposed to overcome the shortcoming of the two-step superlinear convergence, such as [20], but their assumptions are a little strong: (i) the algorithm is assumed to be strongly convergent; (ii) the step size is supposed to always equal to one after finite iterations.

In this paper, we present a modified SQP algorithm for the minimax problem (1.1). In this algorithm, a main search direction is obtained by solving a quadratic program (QP). In order to avoid the Maratos effect, unlike [19], a correction direction is generated by solving the system of linear equations. Under mild conditions, the global and superlinear convergence can be obtained. Finally, some preliminary numerical results are reported.

The rest of this paper is organized as follows. The algorithm and its properties are presented in Section 2. Global and superlinear convergences are analyzed in Section 3 and Section 4, respectively. Numerical results are reported in Section 5. Section 6 is devoted to final remarks.

2. Algorithm

For convenience of presentation, for a given $x \in R^n$, we use the following notation throughout this paper

$$\begin{aligned} f(x) &= (f_j(x), j \in I)^T, \quad I(x) = \{j \in I: f_j(x) = F(x)\}, \\ g_j(x) &= \nabla f_j(x), \quad j \in I, \quad g(x) = (\nabla f_j(x), j \in I). \end{aligned} \quad (2.1)$$

We suppose that the following assumptions hold in this paper.

(H1) Functions f_j ($j \in I$) are all first order continuously differentiable.

(H2) Vectors $\left\{ \begin{pmatrix} g_j(x^k) \\ -1 \end{pmatrix}, j \in I(x^k) \right\}$ in R^{n+1} are linearly independent.

Let $x^k \in R^n$ be a given iteration point, based on (H2), we use the following technique to generate an ε -active constraint subset $I_k \supseteq I(x^k)$ such that the matrix $A_k \triangleq \left(\begin{pmatrix} g_j(x^k) \\ -1 \end{pmatrix}, j \in I_k \right)$ is full of column rank.

First, we give the following notations:

$$\begin{aligned} y &= (x, z), \quad L(y, \lambda) = z + \sum_{j \in I} \lambda_j (f_j(x) - z), \quad \nabla_y L(y, \lambda) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \sum_{j \in I} \lambda_j \begin{pmatrix} \nabla f_j(x) \\ -1 \end{pmatrix}, \\ G(x) &= \text{diag}(f_j(x) - F(x)), \quad \nabla c_j(x) = \begin{pmatrix} \nabla f_j(x) \\ -1 \end{pmatrix}, \quad \nabla c(x) = (\nabla c_j(x), j \in I), \\ M(x) &= \nabla c(x)^T \nabla c(x) + G^2(x), \quad \lambda(x) = -M^{-1}(x) \nabla c(x)^T \nabla f_0(x), \quad \nabla f_0(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \Phi(x, \lambda) &= \begin{pmatrix} \nabla_y L(y, \lambda) \\ \min\{f(x) - F(x)e, \lambda\} \end{pmatrix}, \quad e = (1, 1, \dots, 1)^T \in R^{|I|}, \quad \rho(x, \lambda) = \sqrt{\|\Phi(x, \lambda)\|}. \end{aligned}$$

Now, we define the following “guessing” of the active set $I(x)$:

$$I(x, \varepsilon) = \{i: f_i(x) - F(x) + \varepsilon \rho(x, \lambda(x)) \geq 0\},$$

where ε is a nonnegative parameter. It is obvious that (x^*, λ^*) is a KKT pair of problem (P) if and only if $\Phi(x^*, \lambda^*) = 0$ or $\rho(x^*, \lambda^*) = 0$. Facchinei et al. [22] showed that if the second order sufficient condition and the Mangasarian–Fromovitz constraint qualification [26] hold, then for any $\varepsilon > 0$, when x is sufficiently close to x^* , the $I(x, \varepsilon)$ is an exact identification of $I(x)$.

The following are details of this technique.

Algorithm A.

Step (i) Select an initial parameter $\varepsilon = \varepsilon_{k-1} > 0$.

Step (ii) Generate the ε -active constraint subset $I(x^k, \varepsilon)$ and matrix N_k , where

$$N_k = \left(\begin{pmatrix} g_j(x^k) \\ -1 \end{pmatrix}, j \in I(x^k, \varepsilon) \right). \quad (2.2)$$

Step (iii) If $\det(N_k^T N_k) \geq \varepsilon$, set $I_k = I(x^k, \varepsilon)$, $A_k := N_k$ and $\varepsilon_k = \varepsilon$, stop; otherwise set $\varepsilon := \frac{1}{2}\varepsilon$ and repeat Step (ii).

Similar to Lemma 1.1 and Lemma 2.8 in Ref. [23], we present the following lemma, and its proof is omitted here.

Lemma 2.1. Suppose that (H1) and (H2) hold, and let $x^k \in R^n$. Then

- (i) Algorithm A can be terminated in a finite number of computations, i.e., there is no infinite times of loop between Step (ii) and Step (iii);
- (ii) if a sequence $\{x^k\}$ has an accumulation point, then there exists an $\bar{\varepsilon} > 0$ such that the sequence $\{\varepsilon_k\}$ of parameters generated by Algorithm A satisfies $\varepsilon_k \geq \bar{\varepsilon}$ for all k .

For a given iteration point $x^k \in R^n$ and a symmetric positive matrix $H_k = H(x^k)$ (the problem of how H_k is chosen will be discussed much later), we introduce a new quadratic program as follows:

$$\begin{aligned} \text{(QP)} \quad & \min \quad z + \frac{1}{2} d^T H_k d, \\ & \text{s.t.} \quad f_j(x^k) + g_j(x^k)^T d - F(x^k) \leq z, \quad j \in I_k. \end{aligned} \quad (2.3)$$

To describe the main characters of the (QP) (2.3), we give two lemmas as follows.

Lemma 2.2. Suppose that the matrix H_k is symmetric positive definite. Then

- (i) the (QP) (2.3) has a unique optimal solution;
- (ii) (z_k, d^k) is an optimal solution of (2.3) if and only if it is a KKT point of (2.3).

It is not difficult to finish this proof, so it is omitted.

Lemma 2.3. Suppose that (H1) and (H2) hold, and (z_k, d^k) is an optimal solution of (QP) (2.3). Then

- (i) $z_k + \frac{1}{2} (d^k)^T H_k d^k \leq 0$, $z_k \leq 0$; $d^k = 0 \Leftrightarrow z_k = 0$;
- (ii) $d^k = 0 \Leftrightarrow x^k$ is a stationary point of (P);
- (iii) if $d^k \neq 0$, then $z_k < 0$, moreover, d^k is a descent direction of $F(x)$ at point x^k .

Proof. (i) From the fact that $(0, 0)$ is a feasible solution of (QP) (2.3) and H_k is positive definite, one has

$$z_k + \frac{1}{2} (d^k)^T H_k d^k \leq 0, \quad z_k \leq -\frac{1}{2} (d^k)^T H_k d^k \leq 0.$$

If $d^k = 0$, then from the constraints of (2.3) we have

$$F(x^k) - f_j(x^k) + z_k \geq 0, \quad j \in I_k.$$

In view of $\phi \neq I(x^k) \subseteq I_k$, one has $z_k \geq 0$. Combining that $z_k \leq 0$, we have $z_k = 0$.

Conversely, if $z_k = 0$, then $\frac{1}{2} (d^k)^T H_k d^k = \frac{1}{2} (d^k)^T H_k d^k + z_k \leq 0$, taking into account the positive definite property of H_k , one has $d^k = 0$.

(ii) In view of Lemma 2.2(ii), we know that the optimal solution (z_k, d^k) of (2.3) is a KKT point of (QP) (2.3), then there exists a corresponding KKT multiplier vector $\lambda^k = (\lambda_j^k, j \in I_k, 0_{I \setminus I_k})$ such that

$$\begin{aligned} & \begin{pmatrix} 1 \\ H_k d^k \end{pmatrix} + \sum_{j=1}^m \lambda_j^k \begin{pmatrix} -1 \\ g_j(x^k) \end{pmatrix} = 0, \\ & f_j(x^k) + g_j(x^k)^T d^k - F(x^k) - z_k \leq 0, \quad j \in I_k, \\ & (f_j(x^k) + g_j(x^k)^T d^k - F(x^k) - z_k) \lambda_j^k = 0, \quad j \in I_k, \\ & \lambda_j^k \geq 0, \quad j \in I_k; \quad \lambda_j^k = 0, \quad j \in I \setminus I_k. \end{aligned} \quad (2.4)$$

If $d^k = 0$, then we get $z_k = 0$ from Lemma 2.3(i), and we further have from (2.4)

$$\begin{aligned} \sum_{j=1}^m \lambda_j^k g_j(x^k) &= 0, \quad \sum_{j=1}^m \lambda_j^k = 1, \\ f_j(x^k) - F(x^k) &\leq 0, \quad j \in I, \\ (f_j(x^k) - F(x^k))\lambda_j^k &= 0, \quad \lambda_j^k \geq 0, \quad j \in I. \end{aligned} \quad (2.5)$$

Hence x^k is a stationary point of (P) from (1.3).

Conversely, if x^k is a stationary point of (P), then $z_k = 0$ and $d^k = 0$ satisfy (2.4), so $(0, 0)$ is the unique optimal solution of (QP) (2.3) from Lemma 2.2. Therefore $d^k = 0$.

(iii) Using $z_k + \frac{1}{2}(d^k)^T H_k d^k \leq 0$, $d^k \neq 0$, and the positive definite property of the matrix H_k , we know that $z_k < 0$ holds. Furthermore, in view of the constraints of (QP) (2.3), one gets

$$g_j(x^k)^T d^k \leq z_k + F(x^k) - f_j(x^k) = z_k < 0, \quad j \in I(x^k).$$

On the other hand, it is easy to show that the directional derivative $F'(x; d)$ of $F(x)$ at point x along direction d can be expressed as

$$F'(x; d) = \lim_{\lambda \rightarrow 0^+} \frac{F(x + \lambda d) - F(x)}{\lambda} = \max\{g_j(x)^T d, \quad j \in I(x)\}. \quad (2.6)$$

Thus

$$F'(x^k; d^k) \leq z_k < 0, \quad (2.7)$$

and d^k is a descent direction of $F(x)$ at point x^k . The whole proof is completed. \square

Now we give the details of our algorithm as follows.

Algorithm B.

Parameters: $\varepsilon_{-1} > 0$, $\tau \in (2, 3)$, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$.

Step 0. Initialization: $x^0 \in R^n$, a symmetric positive definite matrix $H_0 \in R^{n \times n}$ (usually, H_0 is chosen as a unitary matrix). Let $k := 0$.

Step 1. Generate an ε -active set I_k : Set parameter $\varepsilon = \varepsilon_{k-1}$, generate an active constraint set I_k by Algorithm A and let ε_k be the corresponding termination parameter.

Step 2. Generate a main search direction d^k : Solve (QP) (2.3) to get a solution (z_k, d^k) with the corresponding KKT multiplier vector $\lambda_{I_k}^k = (\lambda_j^k, \quad j \in I_k)$. If $d^k = 0$, then x^k is a stationary point of problem (P) and stop; otherwise, go to Step 3.

Step 3. Generate a correction direction \tilde{d}^k : Compute direction \tilde{d}^k by solving the following system of linear equations

$$\begin{bmatrix} \tilde{H}_k & A_k \\ A_k^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{t} \\ \tilde{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ -\|d^k\|^\tau e - \tilde{f}_k \end{bmatrix}, \quad (2.8)$$

where $\tilde{H}_k = \begin{bmatrix} H_k & 0 \\ 0 & 1 \end{bmatrix}$, $\tilde{t} = \begin{bmatrix} \tilde{d} \\ \tilde{z} \end{bmatrix}$, $e = (1, 1, \dots, 1)^T \in R^{|I_k|}$ and $\tilde{f}_k = (\tilde{f}_j^k, \quad j \in I_k)$, $\tilde{f}_j^k = f_j(x^k + d^k) - F(x^k) - z_k$, $j \in I_k$. If $\|\tilde{d}^k\| \geq \|d^k\|$, set $\tilde{d}^k = 0$.

Step 4. Perform line search: Compute the step size t_k , the first number t of the sequence $\{1, \beta, \beta^2, \dots\}$ satisfying

$$F(x^k + td^k + t^2 \tilde{d}^k) \leq \max_{l=0,1,2} F(x^{k-l}) - \alpha t (d^k)^T H_k d^k. \quad (2.9)$$

Step 5. Update: Generate a new symmetric positive definite matrix H_{k+1} using the damped BFGS formula proposed by [24], set $x^{k+1} = x^k + t_k d^k + t_k^2 \tilde{d}^k$ and $k := k + 1$, go to Step 1.

To explain that the algorithm is well defined, we present the following lemma.

Lemma 2.4. The line search in Step 4 can be carried out if $d^k \neq 0$, that is, there exists $\bar{t}_k > 0$ such that (2.9) holds.

Proof. We assume by contradiction that the conclusion is not correct, that is, (2.9) does not hold for all $\lambda = \beta^j$, $j = 1, 2, \dots$, then from (2.6), (2.7), $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$ and Lemma 2.3(i), we have

$$\begin{aligned} z_k &\geq F'(x^k; d^k) = \lim_{j \rightarrow \infty} \frac{F(x^k + \beta^j d^k) - F(x^k)}{\beta^j} = \lim_{j \rightarrow \infty} \frac{F(x^k + \beta^j d^k + (\beta^j)^2 \tilde{d}^k) - F(x^k)}{\beta^j} \\ &\geq \lim_{j \rightarrow \infty} \frac{F(x^k + \beta^j d^k + (\beta^j)^2 \tilde{d}^k) - \max_{l=0,1,2} F(x^{k-l})}{\beta^j} \\ &\geq - \lim_{j \rightarrow \infty} \alpha (d^k)^T H_k d^k > -\frac{1}{2} (d^k)^T H_k d^k \geq z_k, \end{aligned}$$

which is a contradiction. The proof is completed. \square

3. Global convergence

In this section, we will establish the global convergence of the proposed algorithm. If the solution d^k generated at Step 2 equals to zero, then Algorithm B stops at x^k , moreover, from Lemma 2.3(ii) we know that x^k is a stationary point of the problem (P). And if $d^k \neq 0$, one knows from Lemma 2.3(iii) that d^k is a descent direction of $F(x)$ at point x^k .

We further assume that an infinite sequence $\{x^k\}$ is generated by Algorithm B, and the next object is to show that every accumulation point x^* of $\{x^k\}$ is a stationary point of problem (P).

Firstly, the following assumption is necessary in the rest of this paper.

(H3) The sequence $\{H_k\}$ of matrices is uniformly positive definite, i.e., there exist two positive constants a and b such that

$$a \|d\|^2 \leq d^T H_k d \leq b \|d\|^2, \quad \forall d \in \mathbb{R}^n, \quad \forall k.$$

(H4) For any $x^0 \in \mathbb{R}^n$, the set $\Omega = \{x \in \mathbb{R}^n: f(x) \leq f(x^0)\}$ is compact.

In the rest of this paper, we suppose that x^* is a given accumulation point of $\{x^k\}$. In view of I_k being the subset of the fixed and finite set I and Lemma 2.1, we may assume without loss of generality that there exists an infinite index set K such that

$$x^k \rightarrow x^*, \quad k \rightarrow \infty \quad (k \in K); \quad I_k \equiv I', \quad \forall k \in K. \quad (3.1)$$

Lemma 3.1. (See [19].) The sequence $\{x^k\}$ is bounded and the sequences $\{t_k d^k\}$ and $\{x^{k+1} - x^k\}$ both converge to zero.

Lemma 3.2. Suppose that (H1)–(H3) hold. Then

- (i) the sequences $\{z_k, k \in K\}$, $\{d^k, k \in K\}$ and $\{\tilde{d}^k, k \in K\}$ are all bounded;
- (ii) $\lim_{k \in K} d^k = \lim_{k \in K} \tilde{d}^k = 0$, $\lim_{k \in K} z_k = 0$.

Proof. (i) Due to the fact that $(0, 0)$ is a feasible solution of (QP) (2.3), combining (H3) and the constraints of (QP) (2.3), we have

$$\begin{aligned} 0 &\geq z_k + \frac{1}{2} (d^k)^T H_k d^k \geq f_j(x^k) + g_j(x^k)^T d^k - F(x^k) + \frac{1}{2} (d^k)^T H_k d^k \\ &= g_j(x^k)^T d^k + \frac{1}{2} (d^k)^T H_k d^k \geq -\|g_j(x^k)\| \cdot \|d^k\| + \frac{1}{2} a \|d^k\|^2, \quad \forall j \in I(x^k), \quad \forall k. \end{aligned}$$

These inequalities show that $\{z_k, k \in K\}$ and $\{d^k, k \in K\}$ are all bounded. Taking into account the definition of \tilde{d}^k at Step 3 of Algorithm B, we can conclude that $\{\tilde{d}^k, k \in K\}$ is bounded.

(ii) Similar to the proof of Theorem 3.1 in [19], we can prove $\lim_{k \in K} d^k = 0$, this shows conclusion (ii) holds. \square

Lemma 3.3. The whole multiplier sequence $\{\lambda^k = (\lambda_{I_k}^k, 0_{I \setminus I_k})\}$ is bounded.

Proof. From (2.4), we get $\sum_{j=1}^m \lambda_j^k = 1$ and $\lambda_j^k \geq 0$, $j \in I$. Thus sequence $\{\lambda^k\}$ is bounded. \square

Now, we give the following globally convergent theorem for the proposed algorithm.

Theorem 3.1. Suppose that (H1)–(H4) hold, then the proposed Algorithm B either stops at a stationary point of problem (P) in a finite number of iterations, or generates an infinite sequence $\{x^k\}$ such that each accumulation x^* of $\{x^k\}$ is a stationary point of (P).

The proof is similar to the one of Theorem 3.1 in [19].

4. Rate of convergence

In this section, firstly, we give a proposition as follows, which is useful in the next discussions.

Proposition 4.1. Suppose that (H1) and (H2) hold. Then the multiplier vector corresponding to a stationary point \tilde{x} of (P) is unique.

Proof. Suppose that $\tilde{\lambda}, \tilde{\mu}$ are two multiplier vectors corresponding to the same stationary point \tilde{x} . Then we have from (1.3)

$$\sum_{j \in I(\tilde{x})} \tilde{\lambda}_j \begin{pmatrix} g_j(\tilde{x}) \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \sum_{j \in I(\tilde{x})} \tilde{\mu}_j \begin{pmatrix} g_j(\tilde{x}) \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \tilde{\lambda}_{I \setminus I(\tilde{x})} = \tilde{\mu}_{I \setminus I(\tilde{x})} = 0.$$

The above equations give $\sum_{j \in I(\tilde{x})} (\tilde{\mu}_j - \tilde{\lambda}_j) \begin{pmatrix} g_j(\tilde{x}) \\ -1 \end{pmatrix} = 0$. Therefore, $\tilde{\lambda} = \tilde{\mu}$ holds according to (H2). \square

Theorem 4.1. Suppose that (H1)–(H4) hold. Then

$$\lim_{k \rightarrow \infty} d^k = \lim_{k \rightarrow \infty} \tilde{d}^k = 0, \quad \lim_{k \rightarrow \infty} z_k = 0.$$

The proof is similar to the one of Theorem 3.1 in [19].

In order to obtain the superlinearly convergent rate of the proposed algorithm, we further suppose that the following assumption holds.

- (H5) (i) The functions $f_j(x)$ ($j \in I$) are all twice continuously differentiable for any $x \in \mathbb{R}^n$;
(ii) The sequence $\{x^k\}$ generated by Algorithm B possesses an accumulation point x^* with the corresponding unique multipliers μ^* (by Theorem 3.1, x^* is a stationary point of problem (P)), such that the stationary point pair (x^*, μ^*) of problem (P) satisfies the following second order sufficiency conditions for some index $t_0 \in I(x^*)$

$$d^T \nabla_{xx}^2 L(x^*, \mu^*) d > 0, \quad \forall d \in \{d \in \mathbb{R}^n: d \neq 0, (g_j(x^*) - g_{t_0}(x^*))^T d = 0, j \in I(x^*)\},$$

where

$$\nabla_{xx}^2 L(x^*, \mu^*) = \sum_{j \in I} \mu_j^* \nabla^2 f_j(x^*) = \sum_{j \in I(x^*)} \mu_j^* \nabla^2 f_j(x^*).$$

- (iii) The strict complementarity condition holds at (x^*, μ^*) , that is, $\mu_j^* > 0, \forall j \in I(x^*)$.

Now we prove that x^* is an isolated stationary point of (P) under certain conditions.

Lemma 4.1. Suppose that (H2) and (H5) hold. Then x^* is an isolated stationary point of (P).

The proof is similar to the one of Proposition 4.1 in [7].

Theorem 4.2. Suppose that (H2)–(H5) hold. Then $\lim_{k \rightarrow \infty} x^k = x^*$.

Proof. From Lemma 4.1, we know that x^* is an isolated stationary point of (P). Furthermore, one can conclude x^* is an isolated limit point of $\{x^k\}$ and this together with Theorem 4.1 implies $\lim_{k \rightarrow \infty} x^k = x^*$ (see Proposition 4.1 in [7]). \square

The following lemma indicates that the active constraints can be accurately identified when it is close to the solution even if the strict complementarity condition does not hold at x^* .

Lemma 4.2. Let x^* be a stationary point of problem (P) and assume that (H2), (H5)(i) and (ii) hold. Then there exists a neighborhood of x^* such that, for each x in this neighborhood,

$$I(x, \varepsilon) = I(x).$$

The sketch of the proof is as follows:

- (i) First, by using Lemma 3.3, Lemma 3.4 and Theorem 3.5, Theorem 3.6, Theorem 3.7 in [22], we know that $\rho(x, \lambda)$ is an identification function (see Definition 2.1 in [22]).
- (ii) Second, parallelling to the proof of Theorems 2.2, 2.3 in [22], it isn't difficult to show that there exists a neighborhood of x^* such that, for each x in this neighborhood, $I(x, \varepsilon) = I(x)$.

Lemma 4.3. *Let (H1), (H2) and (H5)(iii) hold. Then, when k is sufficiently large,*

$$J_k = I(x^*),$$

where $J_k = \{j \mid f_j(x^k) + g_j(x^k)^T d^k - F(x^k) = z_k\}$.

Proof. For any $j \in J_k$, we have

$$f_j(x^k) + g_j(x^k)^T d^k - F(x^k) = z_k.$$

Taking into account Theorems 4.1 and 4.2, and by taking the limit in the above equation, we have $J_k \subseteq I(x^*)$. Conversely, for any $j \in I(x^*)$, we have $\lambda_j^k > 0$ for k large enough from (H5)(iii). In view of (2.4), one gets $I(x^*) \subseteq J_k$. The proof is complete. \square

Lemma 4.4. *If (H2)–(H5) hold, then, for all k , the matrix*

$$M_k \stackrel{\text{def}}{=} \begin{bmatrix} \tilde{H}_k & A_k \\ A_k^T & 0 \end{bmatrix}$$

is nonsingular, furthermore, there exists a constant $C > 0$ such that $\|M_k^{-1}\| \leq C$.

The proof of Lemma 4.2 is similar to the one of Lemma 3.1 in [9], and it is omitted.

Lemma 4.5. *Suppose that (H2), (H3) and (H5)(iii) hold. Then $\|\tilde{d}^k\| = O(\|d^k\|^2)$.*

Proof. Taking into account Taylor expansion, the definition of \tilde{f}_j^k and Lemma 4.3, we get

$$\tilde{f}_j^k = f_j(x^k + d^k) - F(x^k) - z_k = f_j(x^k) + g_j(x^k)^T d^k + O(\|d^k\|^2) - F(x^k) - z_k = O(\|d^k\|^2).$$

So, by using Lemma 4.4, $\tau \in (2, 3)$ and (2.8), we have $\|\tilde{d}^k\| = O(\|d^k\|^2)$. \square

Lemma 4.6. *If (H2)–(H5) are all satisfied, then the KKT multiplier $\lambda_{I_k}^k$ of (2.3) corresponding to (z_k, d^k) satisfies $\lim_{k \rightarrow \infty} \lambda^k = \mu^*$ with $\lambda^k = (\lambda_{I_k}^k, 0_{I \setminus I_k})$.*

Proof. We assume by contradiction that $\lim_{k \rightarrow \infty} \lambda^k \neq \mu^*$, then there exists an infinite subset K and a constant $\bar{a} > 0$ such that

$$\|\lambda^k - \mu^*\| \geq \bar{a}, \quad k \in K.$$

In view of $\lim_{k \rightarrow \infty} x^k = x^*$ and the boundedness of $\{\lambda^k\}$, there exists another infinite set $K' \subseteq K$ such that

$$x^k \rightarrow x^*, \quad \|\lambda^k - \mu^*\| \geq \bar{a}, \quad \lambda^k \rightarrow \lambda^*, \quad k \in K' \subseteq K. \quad (4.1)$$

Taking into account of Theorem 4.1 and passing to the limit $k \in K'$ and $k \rightarrow \infty$ in (2.4), we have

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j=1}^m \lambda_j^* \begin{pmatrix} -1 \\ g_j(x^*) \end{pmatrix} = 0,$$

$$f_j(x^*) - F(x^*) \leq 0, \quad (f_j(x^*) - F(x^*))\lambda_j^* = 0, \quad \lambda_j^* \geq 0, \quad j \in I.$$

From the above equations, we know that (x^*, λ^*) is a stationary point pair of (P), thus $\lambda^* = \mu^*$ (since the multiplier vector is unique), which contradicts (4.1). So the whole proof is finished. \square

In order to obtain the superlinearly convergent rate of the proposed algorithm, we should guarantee that the unit step size is accepted by the line search for k large enough. For this, the following assumptions are necessary.

(H6) Suppose that $\|(\nabla_{xx}^2 L(x^k, \lambda^k) - H_k)d^k\| = o(\|d^k\|)$.

Theorem 4.3. If (H2)–(H6) hold, then the step size in Algorithm B always equals to one, i.e., $t_k \equiv 1$, when k is sufficiently large.

Proof. It is sufficient to show that (2.9) holds under $t = 1$ and k large enough. Firstly, in view of Taylor expansion and Lemma 4.5, we get

$$\begin{aligned} f_i(x^k + d^k + \tilde{d}^k) &= f_i(x^k + d^k) + g_i(x^k + d^k)^T \tilde{d}^k + O(\|\tilde{d}^k\|^2) \\ &= f_i(x^k + d^k) + g_i(x^k)^T \tilde{d}^k + O(\|d^k\|^3), \quad i \in I(x^*). \end{aligned} \quad (4.2)$$

From (2.8), we also have

$$A_k^T \begin{pmatrix} \tilde{d}^k \\ \tilde{z}_k \end{pmatrix} = -\|d^k\|^\tau e - \tilde{f}^k, \quad g_i(x^k)^T \tilde{d}^k = \tilde{z}_k - \|d^k\|^\tau - \tilde{f}_i^k, \quad i \in I(x^*). \quad (4.3)$$

Combining (4.2), (4.3) and the definition of \tilde{f}_i^k , we have

$$\begin{aligned} f_i(x^k + d^k + \tilde{d}^k) &= f_i(x^k + d^k) + \tilde{z}_k - \|d^k\|^\tau - \tilde{f}_i^k + O(\|d^k\|^3) \\ &= F(x^k) + z_k + \tilde{z}_k - \|d^k\|^\tau + O(\|d^k\|^3), \quad i \in I(x^*). \end{aligned}$$

So

$$f_j(x^k + d^k + \tilde{d}^k) = F(x^k) + z_k + \tilde{z}_k - \|d^k\|^\tau + O(\|d^k\|^3), \quad j \in I(x^*).$$

From the two above equations, we obtain

$$f_i(x^k + d^k + \tilde{d}^k) = f_j(x^k + d^k + \tilde{d}^k) + O(\|d^k\|^3), \quad \forall i, j \in I(x^*). \quad (4.4)$$

Taking into account $I(x^k + d^k + \tilde{d}^k) \subseteq I(x^*)$ for k large enough. So, for $\forall j_k \in I(x^k + d^k + \tilde{d}^k) \subseteq I(x^*)$, one has

$$F(x^k + d^k + \tilde{d}^k) = f_{j_k}(x^k + d^k + \tilde{d}^k) = f_{j_k}(x^k + d^k + \tilde{d}^k) + O(\|d^k\|^3), \quad \forall j \in I(x^*).$$

On the other hand, from (2.4), Taylor expansion and Lemma 4.5, we get

$$\begin{aligned} \sum_{j \in I(x^*)} \lambda_j^k &= 1, \quad \lambda_j^k F(x^k + d^k + \tilde{d}^k) = \lambda_j^k f_j(x^k + d^k + \tilde{d}^k) + O(\|d^k\|^3), \quad j \in I(x^*), \\ F(x^k + d^k + \tilde{d}^k) &= \sum_{j \in I(x^*)} \lambda_j^k F(x^k + d^k + \tilde{d}^k) = \sum_{j \in I(x^*)} \lambda_j^k f_j(x^k + d^k + \tilde{d}^k) + O(\|d^k\|^3) \\ &= \sum_{j \in I(x^*)} \lambda_j^k \left(f_j(x^k) + g_j(x^k)^T (d^k + \tilde{d}^k) + \frac{1}{2} (d^k)^T \nabla^2 f_j(x^k) d^k \right) + o(\|d^k\|^2). \end{aligned} \quad (4.5)$$

Also, from (2.4) and Lemma 4.5, one has

$$\sum_{j \in I(x^*)} \lambda_j^k g_j(x^k)^T (d^k + \tilde{d}^k) = -(d^k)^T H_k d^k + o(\|d^k\|^2), \quad (4.6)$$

and

$$\sum_{j \in I(x^*)} \lambda_j^k f_j(x^k) \leq \sum_{j \in I(x^*)} \lambda_j^k F(x^k) = F(x^k). \quad (4.7)$$

So, from (4.5)–(4.7), (H3) and (H5), we have

$$\begin{aligned} F(x^k + d^k + \tilde{d}^k) &\leq \max_{l=0,1,2} F(x^{k-l}) - (d^k)^T H_k d^k + \frac{1}{2} (d^k)^T \left(\sum_{j \in I(x^*)} \lambda_j^k \nabla^2 f_j(x^k) \right) d^k + o(\|d^k\|^2) \\ &= \max_{l=0,1,2} F(x^{k-l}) - \frac{1}{2} (d^k)^T H_k d^k + \frac{1}{2} (d^k)^T \left(\sum_{j \in I(x^*)} \lambda_j^k \nabla^2 f_j(x^k) - H_k \right) d^k + o(\|d^k\|^2) \\ &= \max_{l=0,1,2} F(x^{k-l}) - \alpha (d^k)^T H_k d^k + \left(\alpha - \frac{1}{2} \right) (d^k)^T H_k d^k + o(\|d^k\|^2) \\ &\leq \max_{l=0,1,2} F(x^{k-l}) - \alpha (d^k)^T H_k d^k + \left(\alpha - \frac{1}{2} \right) a \|d^k\|^2 + o(\|d^k\|^2). \end{aligned}$$

Noting that $\alpha \in (0, \frac{1}{2})$, we have for k large enough

$$F(x^k + d^k + \tilde{d}^k) \leq \max_{l=0,1,2} F(x^{k-l}) - \alpha (d^k)^T H_k d^k,$$

that is, (2.9) holds for $t = 1$ and k large enough. So the whole proof is finished. \square

To analyze the superlinear convergence, we give the following lemma.

Lemma 4.7. Suppose that (H2)–(H5) hold and let

$$R_k = R(x^k) = (g_j(x^k) - g_{t_0}(x^k), \quad j \in I(x^*) \setminus \{t_0\}), \quad P_k = I_n - R_k(R_k^T R_k)^{-1} R_k^T,$$

where I_n denotes a unitary matrix. Then, for all k , the matrix

$$G_k = \begin{pmatrix} P_k \nabla_{xx}^2 L(x^*, \mu^*) & R_k \\ R_k^T & 0 \end{pmatrix} \quad (4.8)$$

is nonsingular and there exists a constant c such that $\|G_k^{-1}\| \leq c$.

The proof of this lemma is similar to the one of Lemma 4.4, and is omitted.

Theorem 4.4. Let (H2)–(H6) be satisfied. Then the proposed algorithm is superlinearly convergent, i.e., the sequence $\{x^k\}$ generated by Algorithm B satisfies

$$\|x^{k+1} - x^*\| = o(\|x^k - x^*\|).$$

The proof is similar to the one of Theorem 2.2.3 in [25].

Proof. In view of the active set J_k being a subset of the fixed and finite set I , one gets, when k is sufficient large, $J_k \equiv J$, where J is some subset of I . So, for $t_0 \in J$, we know that $t_0 \in I(x^*)$ from Lemma 4.3.

For convenience of discussion, we denote

$$J' = J \setminus \{t_0\}, \quad R_k = R(x^k) = (g_j(x^k) - g_{t_0}(x^k), \quad j \in J'), \quad \lambda_{j'}^k = (\lambda_j^k, \quad j \in J'), \quad \mu_{j'}^* = (\mu_j^*, \quad j \in J').$$

From (2.4), we get

$$H_k d^k + \sum_{j \in J} \lambda_j^k g_j(x^k) = 0, \quad \sum_{j \in J} \lambda_j^k = 1, \quad \sum_{j \in J} \lambda_j^k g_{t_0}(x^k) = g_{t_0}(x^k), \quad (4.9)$$

$$g_j(x^k)^T d^k = F(x^k) + z_k - f_j(x^k), \quad j \in J. \quad (4.10)$$

Combining (4.9) and (4.10), one has

$$H_k d^k + \sum_{j \in J'} \lambda_j^k (g_j(x^k) - g_{t_0}(x^k)) = H_k d^k + R_k \lambda_{j'}^k = -g_{t_0}(x^k), \quad (4.11)$$

$$(g_j(x^k) - g_{t_0}(x^k))^T d^k = f_{t_0}(x^k) - f_j(x^k), \quad \forall i, j \in J; \quad R_k^T d^k = (f_{t_0}(x^k) - f_j(x^k), \quad j \in J'). \quad (4.12)$$

Let us define vector-valued function $h(x)$ by

$$h(x) = \sum_{j \in J'} \mu_j^* (g_j(x) - g_{t_0}(x)) = R(x) \mu_{j'}^*.$$

Taking into account Taylor expansion and $\sum_{j \in J} \mu_j^* = 1$, we have

$$\begin{aligned} h(x^k) &= R_k \mu_{j'}^* = h(x^*) + \nabla h(x^*)^T (x^k - x^*) + o(\|x^k - x^*\|) \\ &= \sum_{j \in J'} \mu_j^* (g_j(x^*) - g_{t_0}(x^*)) + \sum_{j \in J'} \mu_j^* (\nabla^2 f_j(x^*) - \nabla^2 f_{t_0}(x^*)) (x^k - x^*) + o(\|x^k - x^*\|) \\ &= \sum_{j \in J'} \mu_j^* (g_j(x^*) - g_{t_0}(x^*)) + \left(\sum_{j \in J} \mu_j^* \nabla^2 f_j(x^*) - \nabla^2 f_{t_0}(x^*) \right) (x^k - x^*) + o(\|x^k - x^*\|) \\ &= \sum_{j \in J'} \mu_j^* (g_j(x^*) - g_{t_0}(x^*)) + \nabla_{xx}^2 L(x^*, \mu^*) (x^k - x^*) - \nabla^2 f_{t_0}(x^*) (x^k - x^*) + o(\|x^k - x^*\|). \end{aligned}$$

By the definition of P_k and the above equations, we have $P_k R_k = 0$ and

$$\begin{aligned} 0 &= P_k R_k \mu_{j'}^* = P_k h(x^k) \\ &= P_k \sum_{j \in J'} \mu_j^* (g_j(x^*) - g_{t_0}(x^*)) + P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^k - x^*) - P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) + o(\|x^k - x^*\|). \end{aligned}$$

This along with $\sum_{j \in J} \mu_j^* g_j(x^*) = 0$ and $\sum_{j \in J} \mu_j^* = 1$ implies that

$$\begin{aligned} &P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^k - x^*) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) - P_k \sum_{j \in J'} \mu_j^* (g_j(x^*) - g_{t_0}(x^*)) + o(\|x^k - x^*\|) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) - P_k \left(\sum_{j \in J} \mu_j^* g_j(x^*) - g_{t_0}(x^*) \right) + o(\|x^k - x^*\|) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) + P_k g_{t_0}(x^*) + o(\|x^k - x^*\|). \end{aligned} \quad (4.13)$$

Furthermore, from Theorem 4.3, (4.13) and Lemma 4.5, we have

$$\begin{aligned} &P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^{k+1} - x^*) \\ &= P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^k - x^*) + P_k \nabla_{xx}^2 L(x^*, \mu^*) (d^k + \tilde{d}^k) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) + P_k g_{t_0}(x^*) + P_k \nabla_{xx}^2 L(x^*, \mu^*) d^k + o(\|x^k - x^*\|) + o(\|d^k\|) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) + P_k g_{t_0}(x^*) + P_k (\nabla_{xx}^2 L(x^*, \mu^*) - H_k) d^k + P_k H_k d^k + o(\|x^k - x^*\|) + o(\|d^k\|) \\ &= P_k \nabla^2 f_{t_0}(x^*) (x^k - x^*) + P_k g_{t_0}(x^*) + P_k H_k d^k + o(\|x^k - x^*\|) + o(\|d^k\|). \end{aligned}$$

From (4.11) and the definition of P_k , we obtain $P_k H_k d^k = -P_k g_{t_0}(x^k)$. This along with the above equations and Taylor expansion generates

$$\begin{aligned} &P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^{k+1} - x^*) \\ &= P_k (\nabla^2 f_{t_0}(x^*) (x^k - x^*) + g_{t_0}(x^*) - g_{t_0}(x^k)) + o(\|x^k - x^*\|) + o(\|d^k\|) \\ &= o(\|x^k - x^*\|) + o(\|d^k\|), \end{aligned}$$

that is,

$$P_k \nabla_{xx}^2 L(x^*, \mu^*) (x^{k+1} - x^*) = o(\|x^k - x^*\|) + o(\|d^k\|). \quad (4.14)$$

On the other hand, from Lemma 4.3 and Taylor expansion, we have

$$\begin{aligned} 0 &= f_j(x^*) - f_{t_0}(x^*) = f_j(x^k) - f_{t_0}(x^k) + (g_j(x^k) - g_{t_0}(x^k))^T (x^* - x^k) + o(\|x^k - x^*\|), \quad j \in J_k, \\ 0 &= (f_j(x^k) - f_{t_0}(x^k), j \in J') + R_k^T (x^* - x^k) + o(\|x^k - x^*\|). \end{aligned} \quad (4.15)$$

Moreover, from (4.15) and (4.12), we have

$$\begin{aligned} R_k^T (x^k - x^*) &= (f_j(x^k) - f_{t_0}(x^k), j \in J') + o(\|x^k - x^*\|), \\ R_k^T (x^{k+1} - x^*) &= R_k^T (x^k - x^*) + R_k^T (d^k + \tilde{d}^k) \\ &= (f_j(x^k) - f_{t_0}(x^k), j \in J') + R_k^T d^k + o(\|x^k - x^*\|) + o(\|d^k\|) \\ &= o(\|x^k - x^*\|) + o(\|d^k\|). \end{aligned}$$

That is,

$$R_k^T (x^{k+1} - x^*) = o(\|x^k - x^*\|) + o(\|d^k\|). \quad (4.16)$$

Combining (4.14) and (4.16), we have

$$\begin{pmatrix} P_k \nabla_{xx}^2 L(x^*, \mu^*) & R_k \\ R_k^T & 0 \end{pmatrix} \begin{pmatrix} x^{k+1} - x^* \\ 0 \end{pmatrix} = o(\|d^k\|) + o(\|x^k - x^*\|).$$

Table 1

The detailed information of the solutions to the tested problems.

Algorithm	Prob	Ni	objective	dnorm	eps
Algo 1	1	8	1.9522	4.2478e–006	0.1e–04
Algo 2	–	9	1.9522	9.6073e–006	0.1e–04
Algo 1	2	7	2.0000	1.9868e–014	0.1e–04
Algo 2	–	8	2.0000	5.2721e–006	0.1e–04
Algo 1	3	11	–44.0000	8.2939e–006	0.1e–04
Algo 2	–	11	–43.9900	1.2939e–005	0.1e–04
Algo 1	4	12	0.6164	8.7974e–006	0.1e–04
Algo 2	–	13	0.6164	1.0136e–006	0.1e–04
Algo 1	5	14	3.5997	1.8319e–006	0.1e–04
Algo 2	–	15	3.5997	8.6761e–006	0.1e–04
Algo 1	6	10	0.0508	2.3312e–006	0.1e–04
Algo 2	–	10	0.0508	3.6073e–006	0.1e–04
Algo 1	7	14	2.7545e–006	1.9763e–006	0.1e–04
Algo 2	–	15	2.7546e–006	2.0389e–006	0.1e–04
Algo 1	8	14	680.6301	1.5379e–006	0.1e–04
Algo 2	–	16	680.6380	2.7637e–006	0.1e–04
Algo 1	9	15	24.3012	2.5676e–006	0.1e–04
Algo 2	–	19	24.3062	1.2946e–006	0.1e–04
Algo 1	10	27	1.3261e+002	8.8477e–006	0.1e–04
Algo 2	–	25	1.3261e+002	8.1349e–006	0.1e–04
Algo 1	Vardi-3	10	–48.0158	4.5843e–008	0.1e–04
Algo 2	–	12	–48.0158	6.3061e–008	0.1e–04

This together with Lemmas 4.5 and 4.7 shows that

$$\begin{aligned}
 \|x^{k+1} - x^*\| &= o(\|d^k\|) + o(\|x^k - x^*\|) \\
 &= o(\|d^k + \tilde{d}^k\|) + o(\|x^k - x^*\|) \\
 &= o(\|(x^{k+1} - x^*) - (x^k - x^*)\|) + o(\|x^k - x^*\|) \\
 &\leq o(\|x^{k+1} - x^*\|) + o(\|x^k - x^*\|),
 \end{aligned}$$

i.e.,

$$\frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \left(1 - \frac{o(\|x^{k+1} - x^*\|)}{\|x^{k+1} - x^*\|} \right) \leq \frac{o(\|x^k - x^*\|)}{\|x^k - x^*\|}.$$

This implies

$$\|x^{k+1} - x^*\| = o(\|x^k - x^*\|).$$

The whole proof is completed. \square

5. Numerical experiments

In this section, we test some practical problems based on the proposed algorithm (for the purpose of conveniences, we denote it by **Algo 1**) and the one in [20] (the algorithm is denoted by **Algo 2**). The numerical experiments are implemented on MATLAB 6.5, under Windows XP and 1000 MHz CPU. Eqs. (2.3) and (2.8) are solved by the Optimization Toolbox. To solve (2.3) efficiently, we use the following Hessian approximation of objective function in (2.3):

$$\begin{bmatrix} H_k & 0 \\ 0 & \varepsilon \end{bmatrix},$$

where $\varepsilon = 0.00001$.

During the numerical experiments, a slight modification of the BFGS formula, which is proposed in [24], is adopted in the algorithm, and we set

$$H_0 = I, \quad \tau = 2.6, \quad \beta = 0.6, \quad \alpha = 0.5, \quad \varepsilon_{-1} = 2,$$

where I is a unitary matrix. The tested problems in Table 1 are selected from [20] and [27]. The initial points for the selected problems are the same as the ones in [20] and [27]. The columns of Table 1 have the following meanings: The **prob** column lists the tested problems taken from [20] and [27]. The columns labelled **Ni** give the number of iterations required to solve the problem. The columns labelled **objective**, **dnorm** and **eps** denote the final objective value, the norm of d^k and the step criterion threshold ϵ , respectively.

The detailed information of the solutions to the tested problems is listed in Table 1. It can be seen from Table 1 that the proposed algorithm may be effective, since it can successfully reach a near-optimal point for all the tested problems. Furthermore, it is easy to see from Table 1 that the two algorithms do not have much difference in the number of iterations. But, we find that the numerical performance is sensitive to the choice of parameters during the numerical experiments. Although special choice of parameters will be better for each problem, we insist on using the same set of parameters for all the tested problems.

6. Concluding remarks

In this paper, we propose a nonmonotone line search SQP algorithm for nonlinear minimax problems. During each iteration, with the solution to a reduced QP subproblem, a main search direction is obtained. Then we correct the main search direction by solving a reduced system of linear equations. Under mild conditions, the global and one-step superlinear convergent properties are obtained. Preliminary numerical results show that the proposed algorithm may be effective.

Acknowledgments

The authors are very grateful to the anonymous referees for their helpful comments and valuable suggestions. The authors also thank Yunhai Xiao of Henan University for his help in numerical experiments, to improve the numerical results in the first version of this paper.

References

- [1] D.Q. Mayne, E. Polak, A. Sangiovanni-Vincentelli, Computer-aided design via optimization: A review, *Auto.* 18 (1982) 147–154.
- [2] E. Polak, D.Q. Mayne, D.M. Stimpler, Control system design via semi-infinite optimization: A review, *Proc. IEEE* 72 (1984) 1777–1794.
- [3] S.E. Salcudean, Algorithms for optimal design of feedback compensators, PhD thesis, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, 1986.
- [4] T.L.S. Wu, Delightmimo: An interactive system for optimization-based multivariable control system design, PhD thesis, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, 1986.
- [5] S.P. Han, Variable metric methods for minimizing a class of nondifferentiable functions, *Math. Program.* 20 (1981) 1–13.
- [6] D.Q. Mayne, E. Polak, A superlinearly convergent algorithm for constrained optimization problems, *Math. Program.* 16 (1982) 45–61.
- [7] E.R. Panier, A.L. Tits, A superlinearly convergent feasible method for the solution of inequality constrained optimization problems, *SIAM J. Control Optim.* 25 (1987) 934–950.
- [8] E.R. Panier, A.L. Tits, On combining feasibility, descent and superlinear convergence in inequality constrained optimization, *Math. Program.* 59 (1993) 261–276.
- [9] H.D. Qi, L.Q. Qi, A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality optimization, *SIAM J. Optim.* 11 (2000) 113–132.
- [10] J.B. Jian, K.C. Zhang, S.J. Xue, A superlinearly and quadratically convergent SQP type feasible method for constrained optimization, *Appl. Math. J. Chinese Univ. Ser. B* 15 (2000) 319–331.
- [11] J.B. Jian, C.M. Tang, An SQP feasible descent algorithm for nonlinear inequality constrained optimization without strict complementarity, *Comput. Math. Appl.* 49 (2005) 223–238.
- [12] M.J.D. Powell, A fast algorithm for nonlinearly constrained optimization calculation, in: *Proceedings of 1977 Dundee Biennial Conference on Numerical Analysis*, Springer, Berlin, 1978, pp. 144–157.
- [13] S.P. Han, A globally convergent method for nonlinear programming, *J. Optim. Theory Appl.* 22 (1977) 297–309.
- [14] A.R. Conn, Y. Li, An efficient algorithm for nonlinear minimax problems, Report CS-88-41, University of Waterloo, Waterloo, Ontario, Canada, 1989.
- [15] L. Luskkan, A compact variable metric algorithm for nonlinear minimax approximation, *Comput.* 36 (1986) 355–373.
- [16] E. Polak, D.Q. Mayne, J.E. Higgins, A superlinearly convergent algorithm for minimax problems, in: *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, Florida, 1989, pp. 894–898.
- [17] E. Polak, D.Q. Mayne, J.E. Higgins, Superlinearly convergent algorithm for min-max problems, *J. Optim. Theory Appl.* 69 (1991) 407–439.
- [18] R.A. Polyak, Smooth optimization methods for minimax problems, *SIAM J. Control Optim.* 26 (1988) 1274–1286.
- [19] J.L. Zhou, A.L. Tits, Nonmonotone line search for minimax problems, *J. Optim. Theory Appl.* 76 (1993) 455–476.
- [20] Y. Xue, The sequential quadratic programming method for solving minimax problem, *J. Systems Sci. Math. Sci.* 22 (2002) 355–364 (in Chinese).
- [21] Y.H. Yu, L. Gao, Nonmonotone line search algorithm for constrained minimax problems, *J. Optim. Theory Appl.* 115 (2002) 419–446.
- [22] F. Facchinei, A. Fischer, C. Kanzow, On the accurate identification of active constraints, *SIAM J. Optim.* 9 (1999) 14–32.
- [23] Z.Y. Gao, G.P. He, F. Wu, A method of sequential systems of linear equations with arbitrary initial point, *Sci. China Ser. A* 27 (1997) 24–33.
- [24] J.F.A. De, O. Pantoja, D.Q. Mayne, Exact penalty function algorithm with simple updating of the penalty parameter, *J. Optim. Theory Appl.* 69 (1991) 441–467.
- [25] J.B. Jian, Researches on superlinearly and quadratically convergent algorithms for nonlinearly constrained optimization, PhD thesis, School of Science, Xi'an Jiaotong University, Xi'an, China, 2000.
- [26] O.L. Mangasarian, S. Fromovitz, The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.* 17 (1967) 37–47.
- [27] A. Vardi, New minimax algorithm, *J. Optim. Theory Appl.* 75 (1992) 613–634.