Applied Mathematics Letters 22 (2009) 464–469

Contents lists available at ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

An improved SQP algorithm for solving minimax problems☆

Zhibin Zhu [a,*], Xiang Cai [a], Jinbao Jian [b]

[a] College of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, 541004, PR China

[b] College of Mathematics and Informational Science, Guangxi University, Nanning, 510004, PR China

ARTICLE INFO

Article history:
Received 10 December 2006
Received in revised form 28 April 2008
Accepted 3 June 2008

Keywords:
Minimax problem
SQP algorithm
Global convergence
Superlinear convergence rate

ABSTRACT

In this work, an improved SQP method is proposed for solving minimax problems, and a new method with small computational cost is proposed to avoid the Maratos effect. In addition, its global and superlinear convergence are obtained under some suitable conditions.

1. Introduction

Consider the following minimax optimization problems:

$$\min \quad f(x) = \max_{1\le j\le m} f_j(x), \quad \text{s.t.}\, x \in R^n, \tag{1.1}$$

where f_j $(j = 1 \sim m)$ are continuously differentiable real-valued functions defined on R^n.

Due to the non-differentiability of the object function $f(x)$, we cannot use the classical gradient methods directly to solve such optimization problems [1–3,6].

Many schemes have been proposed for solving the problem (1.1), by converting it to a smooth constrained optimization problem in R^{n+1} as follows:

$$\begin{array}{ll} \min & z \\ \text{s.t.} & f_j(x) - z \le 0, j \in I = \{1, \dots, m\}. \end{array} \tag{1.2}$$

Obviously, from the problem (1.2), the K–T condition of (1.1) is defined as follows:

$$\begin{array}{l} \sum_{j=1}^{m} \lambda_j \nabla f_j(x) = 0, \qquad \sum_{j=1} \lambda_j = 1, \\ \lambda_j(f_j(x) - f(x)) = 0, \lambda_j \ge 0, j = 1, \dots, m, \end{array} \tag{1.3}$$

It is well known that, because of its superlinear convergence rate, the SQP method is one of the most effective methods for solving nonlinear programming problems. In [4], an SQP algorithm is proposed for solving the problem (1.1). There, the

☆ This work was supported in part by NNSF (Nos 10501009, 10771040) of China and Guangxi Province Science Foundation (Nos 0728206, 0640001).

* Corresponding author.
E-mail address: zhuzbma@hotmail.com (Z. Zhu).

0893-9659/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2008.06.017

main search direction d and its Maratos revised direction $\widetilde{d}$ are computed by solving two QP subproblems. In order to obtain the global convergence as well as the superlinear convergence rate, it is necessary to perform a nonmonotone line search along the arc $td + t^2\widetilde{d}$.

Recently, an SQP type algorithm was proposed in [5] for solving the problem (1.1). In contrast with the case for [4], the monotone line search assures acceptance of the unit step size. However, in order to avoid the Maratos effect, it is still necessary to solve two QP subproblems to obtain the search direction d^k and its high-order revised version $\widetilde{d}^k$. Moreover, in order to obtain the superlinear convergence rate, an additional posteriori assumption is made: $\sum_{j=1}^{m}(\lambda_j^k - u_j^*)\nabla f_j(x^k) = o(\|d^k\|)$, which involves the behavior of the gradient of the functions $f_j(x^k)$, the direction d^k, the approximate multiplier λ^k, and the K–T multiplier u^*.

In this work, an improved SQP algorithm for solving the problem (1.1) is proposed. This algorithm overcomes the shortcomings just pointed out. It takes advantage of the active set of the QP subproblem, and constructs a suitable nonsingular matrix. With this matrix, the new revised method is proposed to overcome the Maratos effect problem, that is to say, the Maratos revised direction is obtained by solving a system of linear equations, instead of a QP subproblem which is equivalent to a nonlinear system. Unlike in [5], under some general conditions, the global convergence is obtained as well as the superlinear convergence rate.

2. Description of the algorithm

For the sake of convenience, define

$$I = \{1, \dots, m\}, I(x) = \{j \mid f_j(x) = f(x)\}.$$

The following general assumptions are true throughout the work.

H 2.1. f_j $(j = 1, \dots, m)$ are continuously differentiable.

H 2.2. For all $x \in R^n$, vectors $\left\{\begin{pmatrix}\nabla f_j(x)\\ -1\end{pmatrix}, j \in I(x)\right\}$ are linearly independent.

Now, the algorithm for the solution of the problem (1.1) can be stated as follows.

Algorithm A.

Step 0 *Initialization and data:*
Given $x^0 \in R^n$, $H_0 \in R^{n\times n}$, a symmetric positive definite matrix, parameters $\alpha \in (0, \frac{1}{2})$, $\tau \in (2, 3)$, set $k = 0$.

*Step*1 *Computation of the search direction:*

1.1 Compute (z_k, d^k) by solving the following quadratic problem at x^k:

$$\begin{aligned} \min \quad & z + \frac{1}{2}d^T H_k d \\ \text{s.t.} \quad & f_j(x^k) - f(x^k) + \nabla f_j(x^k)^T d \le z, j \in I. \end{aligned} \tag{2.1}$$

Let λ^k be the corresponding multiplier vector. If $(z_k, d^k) = (0, 0)$, stop.

1.2 Let

$$J_k = \{j \mid f_j(x^k) + \nabla f_j(x^k)^{\mathrm{T}} d^k - f(x^k) - z_k = 0\}, j_k = \min\{j : j \in I(x^k)\}, \tag{2.2}$$

and compute

$$\begin{aligned} &\overline{f}_{j_k}(x^k) = \left(\overline{f}_{j,j_k}(x^k), j \in J_k \setminus \{j_k\}\right) \in R^{|J_k\setminus\{j_k\}|}, \\ &\overline{f}_{j,j_k}(x^k) = f_j(x^k) - f_{j_k}(x^k), j \in J_k \setminus \{j_k\}, \\ &A_k = \nabla\overline{f}_{j_k}(x^k) = \left(\nabla f_j(x^k) - \nabla f_{j_k}(x^k), j \in J_k \setminus \{j_k\}\right) \in R^{n\times(|J_k\setminus\{j_k\}|)}. \end{aligned} \tag{2.3}$$

1.3 If the matrix A_k is of full rank, obtain $\widetilde{s}^k$ by solving the following system of linear equations:

$$A_k^{\mathrm{T}} s = -\|d^k\|^\tau e - \overline{f}_{j_k}(x^k + d^k), \tag{2.4}$$

where

$$e = (1, \dots, 1)^{\mathrm{T}} \in R^{|J_k\setminus\{j_k\}|}, \overline{f}_{j_k}(x^k + d^k) = \left(\overline{f}_{j,j_k}(x^k + d^k), j \in J_k \setminus \{j_k\}\right). \tag{2.5}$$

1.4 If $\|\widetilde{s}^k\| > \|d^k\|$ or the matrix A_k is not of full rank, set $\widetilde{d}^k = 0$; otherwise, set $\widetilde{d}^k = \widetilde{s}^k$.

*Step*2 *The line search:*
Compute t_k, the first number t of the sequence $\{1, \frac{1}{2}, \frac{1}{4}, \dots\}$ satisfying

$$f(x^k + td^k + t^2\widetilde{d}^k) \le f(x^k) - \alpha t (d^k)^{\mathrm{T}} H_k d^k. \tag{2.6}$$

*Step*3 *Updates*:
Compute a new symmetric definite positive matrix H_{k+1}. Let $x^{k+1} = x^k + t_k d^k + t_k^2\widetilde{d}^k$, $k = k + 1$. Go back to step 1.

3. Global convergence of the algorithm

In this section, we first show that Algorithm A given in Section 2 is globally convergent. For this reason, we make another assumption and let it hold for the remainder of this work.

H 3.1. There exist two constants $0 < a \le b$ such that $a\|d\|^2 \le d^{\mathrm{T}} H_k d \le b\|d\|^2$, for all k, for all $d \in R^n$.

According to (1.3) and (2.1), it is easy to obtain the following result.

Lemma 3.1. *Let (z_k, d^k) be the solution of the QP (2.1) at x^k. If $d^k = 0$, then x^k is a K–T point of (1.1); otherwise, it holds that*

$$\nabla f_j(x^k)^{\mathrm{T}} d^k \le z_k \le -\frac{1}{2}(d^k)^{\mathrm{T}} H_k d^k < 0, \quad j \in I(x^k).$$

From Lemma 3.1, it is obvious, if $d^k \neq 0$, that the line search in step 2 yields a step size $t_k = (\frac{1}{2})^j$ for some finite $j = j(k)$, that is to say, the line search step 2 is always completed.

Lemma 3.2. *If $d^k \neq 0$, then the line search in step 2 yields a step size $t_k = (\frac{1}{2})^j$ for some finite $j = j(k)$.*

Proof. For (2.1), it obvious that

$$z_k + \frac{1}{2}(d^k)^{\mathrm{T}} H_k d^k \le 0, \qquad z_k \le -\frac{1}{2}(d^k)^{\mathrm{T}} H_k d^k < 0.$$

For $j \in I$, define

$$\begin{aligned} a_k &\stackrel{\triangle}{=} f_j(x^k + td^k + t^2\widetilde{d}^k) - f(x^k) + \alpha t (d^k)^{\mathrm{T}} H_k d^k \\ &= f_j(x^k) - f(x^k) + t\nabla f_j(x^k)^{\mathrm{T}}(d^k + t\widetilde{d}^k) + \alpha t(d^k)^{\mathrm{T}} H_k d^k + o(t) \\ &\le (1-t)(f_j(x^k) - f(x^k)) + tz_k + \alpha t (d^k)^{\mathrm{T}} H_k d^k + o(t) \\ &\le \left(\alpha - \frac{1}{2}\right) t (d^k)^{\mathrm{T}} H_k d^k + o(t). \end{aligned}$$

This implies that there exists some $\bar{t}_j > 0$ such that $a_k \le 0$. Define $\bar{t} = \min\{\bar{t}_j, j \in I\}$. It is clear that the line search condition (2.6) is satisfied for all t in $[0, \bar{t}]$. ∎

Lemma 3.3. *$\forall x^k \in R^n$, $i_k \in I(x^k)$, the matrix*

$$C_k = \nabla \overline{f}_{i_k}(x^k) = \left(\nabla f_j(x^k) - \nabla f_{i_k}(x^k), j \in I(x^k) \setminus \{i_k\}\right)$$

is always of full rank.

Proof. It is only necessary to prove that the vectors $\{\nabla f_j(x^k) - \nabla f_{i_k}(x^k), j \in I(x^k) \setminus \{i_k\}\}$ are linearly independent. Suppose that there exist $\lambda_j, j \in I(x^k) \setminus \{i_k\}$ such that

$$\sum_{j \in I(x^k) \setminus \{i_k\}} \lambda_j (\nabla f_j(x^k) - \nabla f_{i_k}(x^k)) = 0.$$

Define $\lambda_{i_k} = -\sum_{j \in I(x^k)\setminus\{i_k\}} \lambda_j$; then it holds that

$$\sum_{j \in I(x)} \lambda_j \nabla f_j(x) = 0, \sum_{j \in I(x)} \lambda_j = 0,$$

and thereby,

$$\sum_{j \in I(x)} \lambda_j \begin{pmatrix} \nabla f_j(x) \\ -1 \end{pmatrix} = 0.$$

According to H 2.2, it is easy to see that $\lambda_j = 0, \forall j \in I(x^k) \setminus \{i_k\}$. So, the matrix C_k is of full rank. ∎

Theorem 3.4. *Algorithm A either stops at the K–T point x^k of (1.1) in finite iterations, or generates an infinite sequence $\{x^k\}$ of which any accumulation point is a K–T point of (1.1).*

Proof. The first statement is obvious, the only stopping point being in step 1.1. Thus, suppose that the algorithm generates an infinite sequence $\{x^k\}$, and we might as well assume that there exists a subsequence K such that

$$x^k \to x^*, \qquad H_k \to H_*, \qquad d^k \to d^*, \qquad z_k \to z_*, \quad k \in K. \tag{3.1}$$

According to Lemma 3.1, in order to finish the proof of this result, it is only necessary to prove that $d^k \to 0$, $k \in K$. In view of (2.6), and Lemma 3.1, it is evident that $\{f(x^k)\}$ is monotonically decreasing. Hence, considering $\{x^k\}_{k\in K} \to x^*$ and the continuity of $f(x)$, it holds that

$$f(x^k) \to f(x^*), \quad k \to \infty. \tag{3.2}$$

Suppose by contradiction that $d^* \neq 0$, then $z^* < 0$. Thereby, according to the conclusion about the successful line search under Lemma 3.1, it is easy to conclude that the step size t_k obtained by the linear search is bounded away from zero on K, i.e.,

$$t_k \geq t_* = \inf\{t_k, k \in K\} > 0, \quad k \in K. \tag{3.3}$$

So, from (2.6), (3.2) and (3.3), we get

$$0 = \lim_{k\in K} \left(f(x^{k+1}) - f(x^k)\right) \leq \lim_{k\in K} -\alpha t_k (d^k)^{\mathrm{T}} H_k d^k \leq -\frac{1}{2}\alpha t_* (d^*)^{\mathrm{T}} H_* d^* < 0.$$

This is a contradiction, which shows that $d^k \to 0$, $k \in K$. The claim holds. ■

4. The rate of convergence

Now we discuss the convergence rate of the algorithm, and prove that the sequence $\{x^k\}$ generated by the algorithm is superlinearly convergent. For this purpose, we add some stronger regularity assumptions.

H 2.1′. The functions $f_j (j \in I)$ are twice continuously differentiable.

H 4.1. The sequence $\{x^k\}$ generated by the algorithm possesses an accumulation point x^* (in view of Theorem 3.4, a K–T point). $H_k \to H_*$, $k \to \infty$.

H 4.2. The matrix $\sum_{j=1}^m u_j^* \nabla^2 f_j(x^*) = \sum_{j\in I(x^*)} u_j^* \nabla^2 f_j(x^*)$ is nonsingular. The second-order sufficiency conditions with strict complementary slackness are satisfied at the K–T point x^* and the corresponding multiplier vector u^*, that is to say, it holds that $u_j^* > 0, j \in I(x^*)$, and

$$d^{\mathrm{T}} \left(\sum_{j=1}^m u_j^* \nabla^2 f_j(x^*) \right) d > 0, \quad \forall 0 \neq d \in Y(x^*, u^*),$$

where

$$Y(x^*, u^*) = \left\{ d \in R^n \mid \nabla f_j(x^*)^{\mathrm{T}} d = \nabla f_{j_k}(x^k)^{\mathrm{T}} d, j \in I(x^*), u_j^* > 0, j_k \in I(x^*) \right\}.$$

According to Theorem 2 and Lemma 6 in [5], we have the following results.

Theorem 4.1. *The K–T point x^* is isolated.*

Lemma 4.2. *For $k \to \infty$, we have*

$$x^k \to x^*, \qquad d^k \to 0, \qquad z_k \to 0, \qquad \lambda^k \to u^*, \quad k \to \infty, \quad J_k \equiv I(x^*).$$

Lemma 4.3. *For k large enough, the matrix A_k is of full rank, and the direction $\widetilde{d}^k$ computed in Step 1.3 satisfies that $\|\widetilde{d}^k\| = O\left(\|d^k\|^2\right)$.*

Proof. Due to $J_k \equiv I(x^*), I(x^k) \subseteq I(x^*)$, imitating the proof of Lemma 3.3, it is obvious that, for k large enough, the matrix A_k is of full rank. So, the direction $\widetilde{d}^k$ is well defined. From H 4.2 and Lemma 4.2, we have, for k large enough, that $\lambda_j^k > 0, j \in I(x^*)$. Thereby, $I(x^k) \subset I(x^*)$, it holds, from (2.2), for $j \in J_k \equiv I(x^*), j_k \in I(x^k) \subseteq I(x^*)$, that

$$\begin{aligned} \bar{f}_{j,j_k}(x^k + d^k) &= f_j(x^k + d^k) - f_{j_k}(x^k + d^k) \\ &= f_j(x^k) + \nabla f_j(x^k)^{\mathrm{T}} d^k + O(\|d^k\|^2) - \left(f_{j_k}(x^k) + \nabla f_{j_k}(x^k)^{\mathrm{T}} d^k + O(\|d^k\|^2)\right) \\ &= O(\|d^k\|^2). \end{aligned}$$

So, from the definition of $\widetilde{d}^k$, it holds that $\|\widetilde{d}^k\| = O\left(\|d^k\|^2\right)$. ■

In order to obtain superlinear convergence, we make the following assumption:

H 4.3. The matrix sequence $\{H_k\}$ satisfies that

$$\left\| P_k \left(H_k - \sum_{j \in I(x^*)} \lambda_j^k \nabla^2 f_j(x^k) \right) d^k \right\| = o\left(\|d^k\|\right),$$

where $P_k = I_n - A_k(A_k^{\mathrm{T}}A_k)^{-1}A_k^{\mathrm{T}}$.

Lemma 4.4. *Define* $\overline{\lambda}^k = \{\lambda_j^k, j \in I(x^*) \setminus \{j_k\}\}$, $\overline{d}^k = -A_k(A_k^{\mathrm{T}}A_k)^{-1}\overline{f}_{j_k}(x^k)$*; then, for* k *large enough, there exists some constant* $c > 0$ *such that*

$$d^k = P_k d^k + \overline{d}^k,\ \|\overline{d}^k\| = O\left(\|\overline{f}_{j_k}(x^k)\|\right), \overline{f}_{j_k}(x^k)^{\mathrm{T}}\overline{\lambda}^k \le -c\|\overline{f}_{j_k}(x^k)\|. \tag{4.1}$$

Proof. According to $J_k \equiv I(x^*), I(x^k) \subset I(x^*)$, it holds, for k large enough, from (2.2), that

$$f_j(x^k) + \nabla f_j(x^k)^{\mathrm{T}} d^k = f_{j_k}(x^k) + \nabla f_{j_k}(x^k)^{\mathrm{T}} d^k, \quad j \in I(x^*) \setminus \{j_k\},$$

so,

$$A_k^{\mathrm{T}} d^k = -\overline{f}_{j_k}(x^k).$$

Thereby,

$$P_k d^k = d^k - A_k(A_k^{\mathrm{T}}A_k)^{-1}A_k^{\mathrm{T}}d^k = d^k + A_k(A_k^{\mathrm{T}}A_k)^{-1}\overline{f}_{j_k}(x^k),$$

i.e.,

$$d^k = P_k d^k + \overline{d}^k.$$

In addition, from the definition of $\overline{d}^k$, it holds that

$$\|\overline{d}^k\| = O\left(\|\overline{f}_{j_k}(x^k)\|\right).$$

In view of $f_{j_k}(x^k) = f(x^k)$, we have that $\overline{f}_{j_k}(x^k) \le 0$. According to Lemma 4.2 and H 4.2, we get that

$$\underline{\lambda} \overset{\triangle}{=} \min\{\lambda_j^k, j \in I(x^*)\} > 0.$$

So, there exists some constant $c > 0$ such that

$$\overline{f}_{j_k}(x^k)^{\mathrm{T}}\overline{\lambda}^k \le \underline{\lambda} \sum_{j \in I(x^*) \setminus \{j_k\}} \overline{f}_{j,j_k}(x^k) \le -c\|\overline{f}_{j_k}(x^k)\|.$$

The claim holds. ■

Lemma 4.5. *Under the above-mentioned assumptions, for* k *large enough,* $t_k \equiv 1$.

Proof. Since the indexes j_k, $\forall k$ have only finite different value, without loss of generality, we assume that $j_k \equiv j_0 \in I(x^*)$. For $s, t \in I(x^*) \setminus \{j_0\}$, according to the definition of $\widetilde{d}^k$, we have that

$$(\nabla f_s(x^k) - \nabla f_{j_0}(x^k))^{\mathrm{T}}\widetilde{d}^k = -\|d^k\|^\tau - (f_s(x^k + d^k) - f_{j_0}(x^k + d^k)).$$

Thereby,

$$\begin{aligned}
f_s(x^k + d^k + \widetilde{d}^k) &= f_s(x^k + d^k) + \nabla f_s(x^k + d^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|\widetilde{d}^k\|^2\right) \\
&= f_s(x^k + d^k) + \nabla f_s(x^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|d^k\|^3\right) \\
&= -\|d^k\|^\tau + f_{j_0}(x^k + d^k) + \nabla f_{j_0}(x^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|d^k\|^3\right), \\
f_t(x^k + d^k + \widetilde{d}^k) &= f_t(x^k + d^k) + \nabla f_t(x^k + d^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|\widetilde{d}^k\|^2\right) \\
&= f_t(x^k + d^k) + \nabla f_t(x^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|d^k\|^3\right) \\
&= -\|d^k\|^\tau + f_{j_0}(x^k + d^k) + \nabla f_{j_0}(x^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|d^k\|^3\right), \\
f_{j_0}(x^k + d^k + \widetilde{d}^k) &= f_{j_0}(x^k + d^k) + \nabla f_{j_0}(x^k + d^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|\widetilde{d}^k\|^2\right) \\
&= f_{j_0}(x^k + d^k) + \nabla f_{j_0}(x^k)^{\mathrm{T}}\widetilde{d}^k + O\left(\|d^k\|^3\right).
\end{aligned}$$

As $\tau \in (2, 3)$, it is easy to see that

$$f_i(x^k + d^k + \widetilde{d}^k) = f_j(x^k + d^k + \widetilde{d}^k) + o\left(\|d^k\|^2\right), \quad \forall i, j \in I(x^*). \tag{4.2}$$

In addition, the facts that $d^k \to 0, \widetilde{d}^k \to 0$ imply, for k large enough, that $I(x^k+d^k+\widetilde{d}^k) \subseteq I(x^*)$. So, for $t \in I(x^k+d^k+\widetilde{d}^k) \subseteq I(x^*)$, we obtain that

$$f(x^k + d^k + \widetilde{d}^k) = f_t(x^k + d^k + \widetilde{d}^k) = f_j(x^k + d^k + \widetilde{d}^k) + o\left(\|d^k\|^2\right) \quad (\forall j \in I(x^*)).$$

Thereby, it holds that

$$\begin{aligned} f(x^k + d^k + \widetilde{d}^k) &= f_j(x^k + d^k + \widetilde{d}^k) + o\left(\|d^k\|^2\right) (\forall j \in I(x^*)) \\ &= \sum_{j\in I(x^*)} \lambda_j^k f_j(x^k + d^k + \widetilde{d}^k) + o\left(\|d^k\|^2\right) \\ &= \sum_{j\in I(x^*)} \lambda_j^k \left(f_j(x^k) + \nabla f_j(x^k)^{\mathrm{T}}(d^k + \widetilde{d}^k) + \frac{1}{2}(d^k)^{\mathrm{T}}\nabla^2 f_j(x^k) d^k\right) + o\left(\|d^k\|^2\right). \end{aligned} \tag{4.3}$$

From (2.1), we have that

$$\sum_{j\in I(x^*)} \lambda_j^k \nabla f_j(x^k)^{\mathrm{T}}(d^k + \widetilde{d}^k) = -(d^k)^{\mathrm{T}} H_k d^k + o\left(\|d^k\|^2\right), \qquad \sum_{j\in I(x^*)} \lambda_j^k = 1, \tag{4.4}$$

so, for every k, it holds that

$$\begin{aligned} \sum_{j\in I(x^*)} \lambda_j^k f_j(x^k) &= \lambda_{j_0}^k f_{j_0}(x^k) + \sum_{j\in I(x^*)\setminus\{j_0\}} \lambda_j^k f_j(x^k) \\ &= f_{j_0}(x^k) + \sum_{j\in I(x^*)\setminus\{j_0\}} \lambda_j^k \left(f_j(x^k) - f_{j_0}(x^k)\right) \\ &= f_{j_0}(x^k) + \overline{f}_{j_0}(x^k)^{\mathrm{T}}\overline{\lambda}^k \\ &\le f(x^k) - c\|\overline{f}_{j_0}(x^k)\|. \end{aligned} \tag{4.5}$$

From (4.3)–(4.5), we have that

$$\begin{aligned} f(x^k + d^k + \widetilde{d}^k) &\le f(x^k) - c\|\overline{f}_{j_0}(x^k)\| - (d^k)^{\mathrm{T}} H_k d^k + \frac{1}{2}(d^k)^{\mathrm{T}}\left(\sum_{j\in I(x^*)} \lambda_j^k \nabla^2 f_j(x^k)\right) d^k + o\left(\|d^k\|^2\right) \\ &= f(x^k) - c\|\overline{f}_{j_0}(x^k)\| - \frac{1}{2}(d^k)^{\mathrm{T}} H_k d^k + \frac{1}{2}(P_k d^k + \overline{d}^k)^{\mathrm{T}}\left(\sum_{j\in I(x^*)} \lambda_j^k \nabla^2 f_j(x^k) - H_k\right) d^k + o\left(\|d^k\|^2\right). \end{aligned}$$

Thereby, from H 4.3 and (4.1), we have that

$$f(x^k + d^k + \widetilde{d}^k) \le f(x^k) - c\|\overline{f}_{j_0}(x^k)\| - \alpha(d^k)^{\mathrm{T}} H_k d^k + \left(\alpha - \frac{1}{2}\right)(d^k)^{\mathrm{T}} H_k d^k + o\left(\|d^k\|^2\right) + o\left(\|\overline{f}_{j_0}(x^k)\|\right).$$

So, according to $\alpha \in (0, \frac{1}{2})$, it holds, for k large enough, that

$$f(x^k + d^k + \widetilde{d}^k) \le f(x^k) - \alpha(d^k)^{\mathrm{T}} H_k d^k.$$

The claim holds. ■

From Lemma 4.5 and the method of Theorem 5.2 in [7], we can get the following theorem:

Theorem 4.6. *Under the above-mentioned conditions, Algorithm A is superlinearly convergent, that is to say, the sequence* $\{x^k\}$ *satisfies that*

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

References

[1] S.P. Han, Variable metric methods for minimizing a class of nondifferentiable functions, Math. Program. 20 (1981) 1–13.
[2] E. Polak, D.Q. Mayne, J.E. Higgins, Superlinearly convergent algorithm for Min–Max problems, J. Optim. Theory Appl. 69 (1991) 407–439.
[3] A. Vardi, New Minimax algorithm, J. Optim. Theory Appl. 75 (1992) 613–634.
[4] J.L. Zhou, A.L. Tits, Nonmonotone line search for Minimax problems, J. Optim. Theory Appl. 76 (1993) 455–476.
[5] Z.B. Zhu, K.C. Zhang, A superlinearly convergent sequential quadratic programming algorithm for minimax problems, J. Numer. Math. Appl. 27 (4) (2005) 15–32.
[6] I. Husain, Z. Jabeen, Continuous-time fractional minmax programming, Math. Comput. Modelling 42 (2005) 701–710.
[7] F. Facchinei, S. Lucidi, Quadratically and superlinearly convergent for the solution of inequality constrained optimization problem, J. Optim. Theory Appl. 85 (2) (1995) 265–289.