Feasible generalized monotone line search SQP algorithm for nonlinear minimax problems with inequality constraints

Jin-bao Jian*, Ran Quan, Xue-lu Zhang

College of Mathematics and Information Science, Guangxi University, 530004, Nanning, P.R. China

Received 7 March 2006; received in revised form 15 May 2006

Abstract

In this paper, the nonlinear minimax problems with inequality constraints are discussed, and a sequential quadratic programming (SQP) algorithm with a generalized monotone line search is presented. At each iteration, a feasible direction of descent is obtained by solving a quadratic programming (QP). To avoid the Maratos effect, a high order correction direction is achieved by solving another QP. As a result, the proposed algorithm has global and superlinear convergence. Especially, the global convergence is obtained under a weak Mangasarian–Fromovitz constraint qualification (MFCQ) instead of the linearly independent constraint qualification (LICQ). At last, its numerical effectiveness is demonstrated with test examples.

© 2006 Elsevier B.V. All rights reserved.

MSC: 90C30; 65K05

Keywords: Inequality constraints; Minimax problems; Generalized monotone line search; Feasible SQP algorithm; Global convergence; Superlinear convergence

1. Introduction

It is very convenient to express some engineering design problems in a minimax form as follows:

\begin{equation}
\min_{x} \quad F(x) \\
\text{s.t.} \quad g_j(x) \leq 0, \quad j \in J,
\end{equation}

where \( F(x) = \max\{ f_i(x), \ i \in I \} \) with \( I = \{1, 2, \ldots, m\} \), \( J = \{1, 2, \ldots, m'\} \). Denote the feasible set of (P) as \( X = \{ x : g_j(x) \leq 0, \ j \in J \} \).

Since the objective function \( F(x) \) is continuous but nondifferentiable even if the functions \( f_i \ (i \in I) \) are all differentiable, the classical methods for smooth optimization problems can not be used directly to solve this kind of constrained minimax problems.
It is well known that the sequential quadratic programming (SQP) method is one of the efficient methods for solving smooth constrained optimization problems, because it has fast convergence rate. Thus, the SQP method plays an important role in highly time-consuming simulations. So, many authors have applied the idea of SQP method to present effective algorithms for solving the minimax problems, such as Refs. [5,11,20,15,16,19]. In Ref. [20], the minimax problems without constraints are discussed with “nonmonotone” line search. Refs. [15] and [16] use appropriate penalty function and augmented Lagrangian formulation to solve directly minimax problems with equality and inequality constraints, respectively. In Refs. [15,16], the penalty parameter selection rule is very important. In [15], the penalty parameter is adaptively adjusted. However, in the inequality-constrained case [16], a further adjustment factor is introduced for each constraint. Ref. [19] includes general constraints and is a extension of [20] with the following form:

\[
\begin{align*}
\min_{F(x)} & \quad F(x) \\
\text{s.t.} & \quad g_j(x) \leq 0, \quad j \in J, \\
& \quad h_i(x) = 0, \quad i \in E.
\end{align*}
\]

where \(E = \{1, \ldots, m_e\}\). Similar to the general constrained optimization problem, the direction finding subproblem (DFS) in [19] at iteration point \(x^k\) has the following form:

\[
\begin{align*}
\min_{(z,d)} & \quad z \\
\text{s.t.} & \quad \nabla f_1(x^k)^T d + \max\{\nabla f_i(x^k)^T d + f_i(x^k) : i \in I\} \\
& \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq 0, \; j \in J, \\
& \quad h_i(x^k) + \nabla h_i(x^k)^T d = 0, \; i \in E.
\end{align*}
\]

The feasible set of subproblem above may be empty, thus [19] gave a modified subproblem, but the analysis in [19] is still corresponding to (1.2). What’s more, Liu and Yuan [9] gave a counterexample to show that the modified subproblem may be incompatible. Superlinear convergence is obtained in [20,15,16,19] with different techniques to overcome the Maratos effect [10]. But [15,16,19] all use penalty functions and the algorithms are not the method of feasible direction (MFD) since the iterations may not be feasible though they move towards feasible set, which is a drawback in some strictly feasible cases, especially in the engineering designs.

Since the idea of this paper has something to do with the MFD, especially the idea of the norm-relaxed MFD, we first remind the idea and its progress of the MFD.

MFD is also an important class of method for solving problem (P) with single level objective i.e., \(m = 1\) (suppose that \(m = 1\) in this paragraph). MFD has been studied widely by a number of authors, but almost each improvement in MFD occurred in the modifying of DFS. MFD was originally proposed by Zoutendijk [21]. A good property of the Zoutendijk’s MFD is that a feasible direction of descent can be obtained by solving one linear program. However, Wolfe [18] proved that Zoutendijk’s MFD does not possess global convergence with a counterexample. Later, Topkis–Veinott [17] modified Zoutendijk’s MFD by modifying the DFS. The modified DFS assures that the sequence of points generated by the algorithm converges to a Fritz John point and has the following form:

\[
\begin{align*}
\min_{(z,d)} & \quad z \\
\text{s.t.} & \quad \nabla f_1(x^k)^T d \leq z, \\
& \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq z, \quad j \in J, \\
& \quad d \in S,
\end{align*}
\]

where \(S\) is a special set bounding the direction \(d\) and \(z\) is an auxiliary variable. However, the Topkis–Veinott’s MFD does not converge linearly even under certain convexity assumptions [12]. Based on Topkis–Veinott’s MFD, Pironneau–Polak’s MFD [13] further improved the MFD [17] such that the DFS has the following form:

\[
\begin{align*}
\min_{(z,d)} & \quad z + \frac{1}{2} \|d\|^2 \\
\text{s.t.} & \quad \nabla f_1(x^k)^T d \leq z, \\
& \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq z, \quad j \in J.
\end{align*}
\]
But, the Pironneau–Polak’s MFD only has linear convergence. Motivated by the Pironneau–Polak’s MFD, Cawood and Kostreva [1] proposed a norm-relaxed MFD. And their DFS is as follows:

\[
\min_{(z, d)} z + \frac{\nu}{2} d^T B_k d
\]

\[
\text{s.t. } \nabla f_1(x^k)^T d \leq z, \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq z, \quad j \in J,
\]

where \(\nu\) is a positive constant. Under certain assumptions, the Cawood and Kostreva’s method has global convergence.

To speed up the convergence rate, Chen and Kostreva [2] proposed another MFD in 1999 by reforming the DFS above as follows:

\[
\min_{(z, d)} z + \frac{1}{2} d^T B_k d
\]

\[
\text{s.t. } \nabla f_1(x^k)^T d \leq \gamma_0 z, \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq \gamma_j z, \quad j \in J,
\]

where \(\gamma_0, \gamma_j (j \in J)\) are all positive constants. Basing on the DFS above, Jian et al. [8] proposed a norm-relaxed strongly sub-feasible direction method for solving (P) with an arbitrary initial iteration point. In [8], the DFS has the following form:

\[
\min_{(z, d)} z + \frac{1}{2} d^T B_k d
\]

\[
\text{s.t. } \nabla f_1(x^k)^T d \leq \gamma_0 z + \alpha(x^k), \quad g_j(x^k) + \nabla g_j(x^k)^T d \leq \gamma_j \eta_k z, \quad j \in J^-(x^k),
\]

\[
\nabla g_j(x^k)^T d \leq \gamma_j \eta_k z + \varphi(x^k), \quad j \in J^+(x^k),
\]

where \(\alpha(x^k)\) is a penalty function, \(\eta_k\) is a positive parameter associated with \(x^k\), \(J^-(x) = \{j \in J : g_j(x) \leq 0\}\), \(J^+(x) = \{j \in J : g_j(x) > 0\}\), \(\varphi(x) = \max\{0; g_j(x), j \in J\}\). But the superlinear convergence is not discussed in [8].

In 1986, Grippo et al. [4] proposed a “nonmonotone” line search according to which the objective function is not forced to decrease at every iteration but merely every \(L\) iterations, where \(L\) is a freely selected positive integer. They showed that, with such a line search, global convergence is still guaranteed, and they also pointed out that, as the full Newton step can then be taken earlier, convergence may often be speeded up.

In this paper, motivated by the techniques of the norm-relaxed MFD in [2,8] and the “non-monotone” line search in [20,19,4], we present a new SQP algorithm for minimax problem with inequality constraints. To get a feasible direction of descent and reduce the computational cost, basing on an \(\varepsilon\)-active objective subset and a \(\theta\)-active constraint subset, we construct a new quadratic programming (QP) subproblem, which always has a feasible solution and possesses small size. By solving the QP subproblem, we get a feasible direction of descent. To avoid the Maratos effect, we construct another QP subproblem corresponding to the first one to get a height-order correction direction. Then, we present our “generalized monotone” line search algorithm, i.e., the merit function is forced to decrease at every \(r + 1\) iterations, where \(r\) is a nonnegative integer. If \(r = 0\), then our algorithm is a usual monotone algorithm, else, we name it \(r\)-monotone algorithm. At every iteration, we only need to solve two QPs with small size. Under mild conditions, the global and superlinear convergence can be obtained. Especially, the global convergence can be obtained only under a weak Mangasarian–Fromovitz constraint qualification (MFCQ).

The rest of this paper is organized as follows. The next section gives our algorithm and some properties of it. In Section 3 and Section 4, we discuss the global and superlinear convergence, respectively. Some preliminary numerical results are reported in Section 5. Finally, we give some remarks about our algorithm.

2. Algorithm

First, for a feasible point \(x \in X\), we denote sets \(I(x)\) and \(J(x)\) by

\[
I(x) = \{i \in I : f_i(x) = F(x)\}, \quad J(x) = \{j \in J : g_j(x) = 0\}.
\]
To yield a feasible direction of descent at point \( x^k \in X \), we construct a quadratic program as follows:

\[
QP(x^k, H_k) \min_{(z,d)} z + \frac{1}{2}d^T H_k d
\]

s.t. \( f_i(x^k) + \nabla f_i(x^k)^T d - F(x^k) \leq z, \quad i \in t^k_e, \)
\( g_j(x^k) + \nabla g_j(x^k)^T d \leq c_k z, \quad j \in J^k_h, \)

(2.1)

where \( c_k, \varepsilon_k, 0_k > 0 \), \( I^k_e = \{ i : F(x^k) - f_i(x^k) \leq \varepsilon_k \} \), \( J^k_h = \{ j : -0_k \leq g_j(x^k) \} \), and the symmetric positive definite matrix \( H_k \) is an approximation to the Lagrangian Hessian matrix of (P).

**Lemma 2.1.** Suppose that the matrix \( H_k \) is symmetric positive definite. Then

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

The proof of this lemma is similar to the one of Lemma 2.1 in [8], thus it is omitted here.

Suppose that \( (z_k, d^k) \) is the solution of (2.1). Then there exist multiplier vectors \( \lambda^k_{I^k_e} \) and \( \mu^k_{J^k_h} \) such that

\[
\left\{
\begin{aligned}
H_k d^k &+ \sum_{i \in I^k_e} \lambda^k_i \left( \begin{array}{c}
-1 \\
\nabla f_i(x^k)
\end{array} \right) + \sum_{j \in J^k_h} \mu^k_j \left( \begin{array}{c}
-c_k \\
\nabla g_j(x^k)
\end{array} \right) = 0, \\
0 \leq \lambda^k_i \perp (f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) - z_k) \leq 0, & i \in I^k_e, \\
0 \leq \mu^k_j \perp (g_j(x^k) + \nabla g_j(x^k)^T d^k - c_k z_k) \leq 0, & j \in J^k_h,
\end{aligned}
\right.
\]

(2.2)

where \( w \perp y \) indicates orthogonality of any vectors \( w \) and \( y \).

A point \( x^k \in X \) is said to be a stationary point of problem (P) if there exist multiplier vectors \( \pi^k = (\pi^k_i, i \in I) \) and \( v^k = (v^k_j, j \in J) \) such that

\[
\left\{
\begin{aligned}
\sum_{i \in I} \pi^k_i \nabla f_i(x^k) + \sum_{j \in J} v^k_j \nabla g_j(x^k) & = 0; \\
\sum_{i \in I} \pi^k_i = 1, \\
0 \leq \pi^k_i \perp (f_i(x^k) - F(x^k)) \leq 0, & i \in I; \\
0 \leq v^k_j \perp g_j(x^k) \leq 0, & j \in J.
\end{aligned}
\right.
\]

(2.3)

First, the two following hypotheses are necessary in this paper.

H1. Functions \( f_i (i \in I) \) and \( g_j (j \in J) \) are all first order continuously differentiable.

H2. The weak MFCQ holds at each \( x \in X \), i.e., there exists a vector \( d \) such that \( \nabla g_j(x)^T d < 0 \) for all \( j \in J(x) \).

**Remark 2.1.** Corresponding to the assumption H2+ in Section 4, the MFCQ holding at each \( x \in X \) implies that there exists a vector \( d \) and for some \( i_x \in I(x) \) such that \( (\nabla f_i(x) - \nabla f_{i_x}(x))^T d < 0 \) for all \( i \in I(x) \) \( \setminus \{i_x\} \) and \( \nabla g_j(x)^T d < 0 \) for all \( j \in J(x) \). Thus, H2 is called to be a weak MFCQ assumption.

**Lemma 2.2.** Suppose that H1 and H2 hold, matrix \( H_k \) is symmetric positive definite and \( (z_k, d^k) \) is an optimal solution of \( QP(x^k, H_k) \). Then

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

**Proof.** (i) From the fact that \( (0, 0) \in \mathbb{R}^{n+1} \) is a feasible solution of \( QP(x^k, H_k) \) and \( H_k \) is positive definite, one has

\[
z_k + \frac{1}{2}(d^k)^T H_k d^k \leq 0, \quad z_k \leq -\frac{1}{2}(d^k)^T H_k d^k \leq 0.
\]
(ii) Firstly, if $d^k = 0$, then from the constraints $f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) \leq z_k$ for $i \in I(x^k)$, we have $z_k \geq 0$. Combining $z_k \leq 0$, 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 $d^k = 0$.

Secondly, if $d^k = 0$, then $z_k = 0$ follows from the conclusion above. In view of Lemma 2.1(ii), we know that the optimal solution $(z_k, d^k) = (0, 0)$ of QP($x^k, H_k$) is also a KKT point of (2.1), so we have from (2.2)

$$
\begin{align*}
\sum_{i \in I^k} \lambda_i^k \nabla f_i(x^k) + \sum_{j \in J^k} \mu_j^k \nabla g_j(x^k) = 0; \quad \sum_{i \in I^k} \lambda_i^k + c_k \sum_{j \in J^k} \mu_j^k = 1, \\
0 \leq \nabla f_i(x^k) - (f_i(x^k) - F(x^k)) \leq 0, \quad i \in I^k; \\
0 \leq \nabla f_j(x^k) - (F(x^k) + g_j(x^k)) \leq 0, \quad j \in J^k.
\end{align*}
$$

(2.4)

On the other hand, we get easily that $\sum_{i \in I^k} \lambda_i^k > 0$ from H2 and (2.4). Thus, (2.4) implies

$$
\begin{align*}
\sum_{i \in I} \lambda_i^k \nabla f_i(x^k) + \sum_{j \in J} \mu_j^k \nabla g_j(x^k) = 0; \quad \sum_{i \in I} \lambda_i^k = 1, \\
0 \leq \nabla f_i(x^k) - (f_i(x^k) - F(x^k)) \leq 0, \quad i \in I; \\
0 \leq \nabla f_j(x^k) - (F(x^k) + g_j(x^k)) \leq 0, \quad j \in J,
\end{align*}
$$

(2.5)

with $\lambda_i^k = \lambda_i^k / (\sum_{i \in I^k} \lambda_i^k)$ for $i \in I^k$, $\mu_j^k = \mu_j^k / (\sum_{i \in I^k} \lambda_i^k)$ for $j \in J^k$ and $\lambda_i^k = 0$ for $i \in I \setminus I^k$, $\mu_j^k = 0$ for $j \in J \setminus J^k$.

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

Conversely, if $x^k$ is a stationary point of (P) with multiplier vectors $\lambda^k$ and $\mu^k$, then $z_k := 0$ together with $d^k := 0$ satisfies (2.2) with multiplier vectors $\lambda_i^k = \lambda_i^k / (1 + c_k \sum_{j \in J} v_j^k)$ for $i \in I^k$ and $\mu_j^k = v_j^k / (1 + c_k \sum_{j \in J} v_j^k)$ for $j \in J^k$, that is, $(0, 0) \in \mathbb{R}^{n+1}$ is a KKT point of QP($x^k, H_k$). So $d^k = 0$ follows from the uniqueness of the KKT point of QP($x^k, H_k$).

(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 $z_k < 0$. Furthermore, in view of the constraints of QP($x^k, H_k$), one gets

$$
\nabla f_i(x^k)^T d^k \leq z_k + F(x^k) - f_i(x^k) = z_k, \quad i \in I(x^k) \subseteq I^k.
$$

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

$$
F'(x; d) = \lim_{\lambda \to 0^+} \frac{F(x + \lambda d) - F(x)}{\lambda} = \max \{\nabla f_i(x)^T d, \quad i \in I(x)\}.
$$

(2.6)

Thus,

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

(2.7)

and $d^k$ is a descent direction of $F(x)$ at point $x^k$. On the other hand, for any $j \in J(x^k) \subseteq J^k$, it follows that

$$
g_j(x^k + \lambda d^k) = g_j(x^k) + \lambda \nabla g_j(x^k)^T d^k + o(\lambda) = \lambda (g_j(x^k) + \nabla g_j(x^k)^T d^k) + o(\lambda) \leq \lambda c_k z_k + o(\lambda) \leq 0
$$

for $\lambda > 0$ small enough. The whole proof is completed. \[\square\]

We know that, to overcome the Maratos effect, a suitable correction direction $\tilde{d}^k$ must be adopted. For this purpose, we yield the correction direction $d^k$ by solving another QP subproblem as follows:

$$
\text{QP}(x^k, d^k, H_k) \min_{(y, d)} y + \frac{1}{2} \left( d + \frac{1}{\sigma_k} d^k \right)^T H_k \left( d + \frac{1}{\sigma_k} d^k \right) \\
\text{s.t.} \quad f_i(x^k + d^k) + \nabla f_i(x^k)^T d - F(x^k + d^k) \leq y, \quad i \in I^k, \\
g_j(x^k + d^k) + \nabla g_j(x^k)^T d \leq c_k z_k \|d^k\| - \|d^k\|^\tau, \quad j \in J^k.
$$

(2.8)
Remark 2.2. Basing on QP($x_k$, $H_k$) and the idea of the norm-relaxed MFD, we construct our QP($x_k$, $d_k$, $H_k$), which is different from the QP($x_k$, $d_k$, $H_k$) in [20,19].

Since QP($x_k$, $d_k$, $H_k$) is a convex quadratic program with linear constraints, the following lemma is at hand.

**Lemma 2.3.** $(y_k, \tilde{d})$ is an optimal solution of QP($x_k$, $d_k$, $H_k$) if and only if it is a KKT point of QP($x_k$, $d_k$, $H_k$). Furthermore, the solution of QP($x_k$, $d_k$, $H_k$) is unique if it exists.

Now, we give the details of our algorithm as follows.

**Algorithm A.** Parameters: nonnegative integer $r$; $v$, $\beta \in (0, 1)$, $\tau \in (2, 3)$, $\varepsilon \in (0, 0.5)$; positive constants $\xi$, $\zeta$, $\tilde{c}$, $M$, $c_0$, $\varepsilon_0$, $\theta_0$, where $M$ is a positive suitably large constant.

Data: an initial point $x^{-r} = \cdots = x^{-2} = x^{-1} = x^0 \in X$, a symmetric positive definite matrix $H_0 \in \mathbb{R}^{n \times n}$.

Step 0: Initialization: Let $k := 0$.

Step 1: Solve QP: Solve QP($x_k$, $H_k$) to get a (unique) solution $(z_k, d_k)$ with corresponding KKT multiplier vectors $\lambda^k$, $\mu^k$, set $\lambda^k = (x^k, 0)_{i \in \mathcal{I}}$ and $\mu^k = (\mu^k, 0)_{j \in \mathcal{J}}$. If $d_k = 0$, stop.

Step 2: Trial of step length unit: If

$$F(x_k + d_k) \leq F_k - \varepsilon |d_k|^T H_k d_k,$$

$$g_j(x_k + d_k) \leq 0, \quad \forall j \in J,$$

set $t_k = 1$ and $\tilde{d}_k = 0$, enter Step 5, where $F_k = \max\{F(x_k + t): t = 0, 1, 2, \ldots, r\}$.

Step 3: Generate a correction direction $\tilde{d}$: If $\sigma_k = \sum_{i \in \mathcal{I}} \lambda^k_i = 0$, set $\tilde{d}_k = 0$, go to Step 4; else solve QP($x_k$, $d_k$, $H_k$), if there is no solution or its solution $(y_k, \tilde{d})$ satisfies $\|\tilde{d}\| > M \cdot \|d_k\|$, set $\tilde{d}_k = 0$.

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

$$F(x_k + t d_k + t^2 \tilde{d}_k) \leq F_k - \varepsilon (d_k)^T H_k d_k,$$

$$g_j(x_k + t d_k + t^2 \tilde{d}_k) \leq 0, \quad \forall j \in J.$$

(2.9)

(2.10)

Step 5: Update: Compute a new symmetric positive definite matrix $H_{k+1}$, set $c_{k+1} = \min\{\tilde{c}, \xi \|d_k\| \tilde{c}\}$, $x_{k+1} = x_k + t_k d_k + t_k^2 \tilde{d}_k$. Choose new positive parameters $\varepsilon_{k+1}$, $\theta_{k+1}$. Let $k := k + 1$, go back to Step 1.

Remark 2.3. Line search (2.9) is our generalized monotone line search. If $r = 0$, then (2.9) is the usual monotone line search, else if $r > 0$, it is a nonmonotone line search, and in this case, the step length may be enlarged. In Section 5 of this paper, we will compare the numerical efficiency corresponding to different values of $r$.

The following lemma shows the proposed algorithm is well defined.

**Lemma 2.4.** The generalized monotone line search at Step 4 can be carried out if $d_k \neq 0$, that is, there exists $\tilde{t}_k > 0$ such that (2.9) and (2.10) hold.

**Proof.** Similar to the proof of Lemma 2.2(iii), we get that $g_j(x_k + t d_k + t^2 \tilde{d}_k) \leq 0$, $j \in J$ for $t > 0$ small enough. Analyze (2.9), by contradiction, we assume that (2.9) does not hold for $t = \beta^j$, $j = 0, 1, 2, \ldots$. Then from (2.7), (2.6), $\varepsilon \in (0, 0.5)$, $\beta \in (0, 1)$ and Lemma 2.2(i), we have

$$z_k \geq F'(x_k; d_k) = \lim_{j \to \infty} \frac{F(x_k + \beta^j d_k) - F(x_k)}{\beta^j} = \lim_{j \to \infty} \frac{F(x_k + \beta^j d_k + \beta^{2j} \tilde{d}_k) - F(x_k)}{\beta^j} \\
\geq \lim_{j \to \infty} \frac{F(x_k + \beta^j d_k + \beta^{2j} \tilde{d}_k) - F_k}{\beta^j} \geq - \lim_{j \to \infty} \varepsilon (d_k)^T H_k d_k > - \frac{1}{2} (d_k)^T H_k d_k \geq z_k,$$

which is a contradiction, thus (2.9) holds and the proof is completed. \(\Box\)
3. Global convergence analysis

In this section, we will discuss the global convergence of the proposed algorithm. If the solution \( d^k \) generated at Step 1 equals zero, then Algorithm A stops at \( x^k \), and from Lemma 2.2(ii) we know that \( x^k \) is a stationary point of problem (P). Thus, we assume that an infinite sequence \( \{x^k\} \) of points is generated by Algorithm A, and the consequent task is to show that every accumulation point \( x^* \) of \( \{x^k\} \) is a stationary point of problem (P). First of all, the following three assumptions are assumed to be held 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

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

H4. For any \( x^0 \in X \), the level set \( \Omega = \{ x \in X : F(x) \leq F(x^0) \} \) is compact. The role of this hypothesis is to ensure that the yielded sequence \( \{x^k\} \) is bounded.

H5. \( \inf_k \{\varepsilon_k\} \geq 0, \quad \inf_k \{\theta_k\} \geq 0 \).

Remark 3.1. If one chooses \( \varepsilon_k \) and \( \theta_k \) by one of the following fashions, then H5 holds automatically.

**Fashion A:** \( \varepsilon_k \equiv \tilde{\varepsilon} \) and \( \theta_k \equiv \tilde{\theta} \) for all \( k \), where \( \tilde{\varepsilon} \) and \( \tilde{\theta} \) are two positive sufficiently small constants.

**Fashion B:** \( \varepsilon_k = \max \{F(x^k) - f_i(x^k), i \in I\} + \tilde{\varepsilon} \) and \( \theta_k = \max \{-g_j(x^k), j \in J\} + \tilde{\theta} \). Note that in this case, \( I_{\varepsilon_k} \equiv I, \quad J_{\theta_k} \equiv J \).

Lemma 3.1. Suppose that H1, H3 and H4 hold. Then (i) the entire sequence \( \{F(x^k)\} \) is convergent and (ii) the entire sequence \( \{t_k d^k\} \) converges to zero.

**Proof.** (i) Obviously, \( F(x^k) \leq F(x^0), \forall k \), thus H4 implies that \( \{x^k\} \) is bounded. We define index \( l(k) \) as follows:

\[
F(x^{l(k)}) = \max \{F(x^{k-l}) : l = 0, 1, 2, \ldots, r\} = \max \{F(x^l) : l = k-r, k-r+1, \ldots, k\}.
\]

This together with (2.9) shows that

\[
F(x^{l(k+1)}) = \max \{F(x^{k+1-l}) : l = 0, 1, 2, \ldots, r\} \leq \max \{F(x^{k+1-l}) : l = 0, 1, 2, \ldots, r+1\}
= \max \{F(x^{l(k)}), F(x^{k+1})\} \stackrel{(2.9)}{=} F(x^{l(k)}),
\]

which shows that \( \{F(x^{l(k)})\} \) is a monotonely non-increasing sequence. Thus, \( \{F(x^{l(k)})\} \) is convergent since it is bounded. Denote

\[
\lim_{k \to \infty} F(x^{l(k)}) = F_\ast. \quad (3.1)
\]

On the other hand, from Algorithm A, we have

\[
F(x^{l(k)}) \leq \max \{F(x^{l(k)-1-h}) : h = 0, 1, 2, \ldots, r\} - \alpha t_{l(k)-1} (d^{l(k)-1})^T H_{l(k)-1} d^{l(k)-1}
= F(x^{l(k-1)}) - \alpha t_{l(k)-1} (d^{l(k)-1})^T H_{l(k)-1} d^{l(k)-1}.
\]

This relationship along with (3.1) and H3 gives

\[
t_{l(k)-1} d^{l(k)-1} \to 0, \quad t_{l(k)-1} d^{l(k)-1} \to 0, \quad \|x^{l(k)} - x^{l(k)-1}\| \to 0, \quad k \to \infty. \quad (3.2)
\]

Now, we set \( \hat{l}(k) = l(k+ r + 2) \) and show, by induction, that for each \( j \geq 1 \)

\[
\lim_{k \to \infty} t_{l(k)-j} d^{l(k)-j} = 0, \quad \lim_{k \to \infty} F(x^{l(k)-j}) = \lim_{k \to \infty} F(x^{l(k)}) = F_\ast. \quad (3.3)
\]

From H4, (3.1), (3.2) and \( \{\hat{l}(k)\} \subseteq \{l(k)\} \), one has

\[
|F(x^{l(k)-1}) - F(x^{l(k)})| \leq |F(x^{l(k)-1}) - F(x^{l(k)})| + |F(x^{l(k)}) - F(x^{l(k)})| \to 0, \quad k \to \infty,
\]

Hence, \( \lim_{k \to \infty} F(x^{l(k)}) = F_\ast \).
this along with (3.2) shows that (3.3) holds for $j = 1$. Suppose that (3.3) holds for $j = \hat{j}$. Then similar to the proof of (3.2), one has
\[ t_{i(k)-\hat{j}-1}^{-1} d_{\hat{i}(k)-\hat{j}-1} \to 0, \quad t_{i(k)-\hat{j}-1}^{-1} \tilde{d}_{\hat{i}(k)-\hat{j}-1} \to 0, \quad \|x^{\hat{i}(k)-\hat{j}} - x^{\hat{i}(k)-\hat{j}-1}\| \to 0, \quad k \to \infty. \]

Thus, from the property of function $F(x)$ and H4, we obtain
\[ \lim_{k \to \infty} F(x^{\hat{i}(k)-\hat{j}-1}) = \lim_{k \to \infty} F(x^{\hat{i}(k)-\hat{j}}) = \lim_{k \to \infty} F(x^{\hat{i}(k)}). \]

So, (3.3) holds for $j = \hat{j} + 1$.

Now, we turn to complete the rest proof. For each index $k$, one has
\[ k - r \leq l(k) \leq k, \quad k + 2 \leq \hat{i}(k) \leq k + 2 + r, \quad 1 \leq \hat{i}(k) - k - 1 = l(k + 2) - k - 1 \leq r + 1, \]
and
\[ x^{\hat{i}(k)} = x^{k+1} + \sum_{j=1}^{\hat{i}(k)-k-1} (t_{i(k)-j}^{-1} d_{\hat{i}(k)-j} + t_{i(k)-j}^{-1} \tilde{d}_{\hat{i}(k)-j}). \]

From the two former relationships, (3.3) and H4, we get
\[ \|x^{k+1} - x^{i(k)}\| \to 0, \quad |F(x^{k+1}) - F(x^{i(k)})| \to 0, \quad F(x^{k+1}) \to F_*, \quad k \to \infty. \]

(ii) From Step 4 of the proposed algorithm and H3, one has
\[ F(x^{k+1}) \leq F(x^{\hat{i}(k)}) - \alpha_{t_k} (d^k)^T H_k d^k \leq F(x^{\hat{i}(k)}) - \alpha_{t_k} \|d^k\|^2. \]

Thus, from Lemma 3.1(i), one has $\alpha d_k \to 0$, $k \to \infty$. The whole proof is finished. \qed

Denote the active constraint sets of QP($x^k$, $H_k$) by
\[ I_k = \{i \in I^0_{\epsilon_k} : f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) = z_k\}, \quad J_k = \{j \in J^0_{\theta_k} : g_j(x^k) + \nabla f_j(x^k)^T d^k = c_k z_k\}. \]

In the rest of this section, suppose that $x^*$ is a given limit point of the yielded sequence $\{x^k\}$ of points. In view of the approximately active sets $I^k_{\epsilon_k}$, $J^k_{\theta_k}$ and the active sets $I_k$, $J_k$ of (2.1) all being the subsets of the fixed and finite sets $I$ and $J$, and the boundedness of $\{c_k\}$, we can assume without loss of generality that there exist an infinite index set $K$ and a constant $c_*$ such that
\[ x^k \to x^*, \quad I^k_{\epsilon_k} \equiv \bar{I}, \quad J^k_{\theta_k} \equiv \bar{J}, \quad I_k \equiv I', \quad J_k \equiv J', \quad c_k \to c_*, \quad H_k \to H_*, \quad k \in K. \quad (3.4) \]

**Lemma 3.2.** Suppose that H1–H5 hold. Then

(i) the entire sequences $\{z_k\}$, $\{d^k\}$ and $\{\tilde{d}^k\}$ are all bounded;
(ii) the entire multiplier sequences $\{\lambda^k_{\epsilon_k}\}$, $\{\mu^k_{\theta_k}\}$ are both bounded;
(iii) if $\lim_{k \in K} c_k = c_* > 0$, then $\lim_{k \in K} d^k = \lim_{k \in K} \tilde{d}^k = 0$ and $\lim_{k \in K} z_k = 0$.

**Proof.** (i) In view of Lemma 2.2(i), the constraints of QP($x^k$, $H_k$), H1, H3 and H4, there exist two constants $c', \tilde{c} > 0$ such that
\[ 0 \geq z_k + \frac{1}{2} (d^k)^T H_k d^k \geq f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) + \frac{1}{2} (d^k)^T H_k d^k \geq -c' \cdot \|d^k\| - \tilde{c} + \frac{1}{2} \alpha \|d^k\|^2, \quad \forall i \in I^k_{\epsilon_k} \neq \emptyset, \forall k. \]

These inequalities imply that $\{z_k\}$ and $\{d^k\}$ are all bounded. Furthermore, the boundedness of $\{\tilde{d}^k\}$ follows from Step 3 of Algorithm A.
(ii) From (2.2), one gets
\[ 0 \leq \lambda^k_{j_k}, \quad \sum_{i \in I_k} \lambda^k_i = 1 - c_k \sum_{j \in J_k} \mu^k_j \leq 1, \]
so \( \{ \lambda^k_{j_k} \} \) is bounded. Now we show that \( \{ \mu^k_{j_k} \} \) is bounded. By contradiction, without loss of generality, suppose that there exists an infinite index set \( \tilde{K} \) such that \( \| \mu^k_{j_k} \| \to \infty \) and (3.4) holds for \( \tilde{K} \). Then from (2.2), we have
\[
\begin{align*}
H_k d^k + \sum_{i \in I_k} \lambda^k_i f_i(x^k) + \sum_{j \in J_k} \mu^k_j g_j(x^k) &= 0, \\
\sum_{i \in I_k} \lambda^k_i + c_k \sum_{j \in J_k} \mu^k_j = 1; & \quad g_j(x^k) + \nabla g_j(x^k) d^k = c_k z_k, \quad j \in J_k.
\end{align*}
\]
Thus,
\[
- (d^k)^T H_k d^k + \sum_{i \in I_k} \lambda^k_i \nabla f_i(x^k) d^k = \sum_{j \in J_k} \mu^k_j \nabla g_j(x^k) d^k = \sum_{j \in J_k} \mu^k_j (c_k z_k - g_j(x^k))
\]
\[
= z_k \left( 1 - \sum_{i \in I_k} \lambda^k_i \right) - \sum_{j \in J_k} \mu^k_j g_j(x^k).
\]
Hence, from the boundedness of \( \{(x^k, d^k, z_k, \lambda^k_{j_k})\} \), there exists a constant \( \tilde{M} > 0 \) such that \( -\sum_{j \in J_k} \mu^k_j g_j(x^k) \leq \tilde{M} \), that is
\[
\sum_{j \in J \setminus J(x^*)} \mu^k_j (g_j(x^k)) + \sum_{j \in J \setminus J(x^*)} \mu^k_j (g_j(x^k)) \leq \tilde{M}.
\]
If \( J' \cap J(x^*) = \phi \), then the inequality above implies that \( \{ \mu^k_{j_k}, k \in \tilde{K} \} \) is bounded, which brings about a contradiction, else, one gets that \( \sum_{j \in J \setminus J(x^*)} \mu^k_j (g_j(x^k)) \leq \tilde{M} \) and \( \{ \mu^k_{j_k}, J \setminus J(x^*) \} : k \in \tilde{K} \} \) is bounded, thus \( \| \mu^k_{J \setminus J(x^*)} \| \to \infty, k \in \tilde{K} \).

Dividing the first equation of (3.5) by \( \gamma_k \triangleq \| \mu^k_{J \setminus J(x^*)} \| \), we get
\[
\frac{1}{\gamma_k} H_k d^k + \sum_{i \in I'} \frac{\lambda^k_i}{\gamma_k} f_i(x^k) + \sum_{j \in J \setminus J(x^*)} \frac{\mu^k_j}{\gamma_k} \nabla g_j(x^k) + \sum_{j \in J \setminus J(x^*)} \frac{\mu^k_j}{\gamma_k} \nabla g_j(x^k) = 0. \tag{3.6}
\]
Note that \( \{(1/\gamma_k) \mu^k_{J \setminus J(x^*)} : k \in \tilde{K} \} \) is bounded with norm one, thus we can assume without loss of generality that
\[
\frac{\mu^k_j}{\gamma_k} \to \bar{\mu}_j, \quad j \in J' \cap J(x^*), \quad k \in \tilde{K}, \quad 0 \leq \bar{\mu}_{J \setminus J(x^*)} \neq 0.
\]

Passing to the limit \( k \in \tilde{K} \) and \( k \to \infty \) in (3.6), one has \( \sum_{j \in J' \cap J(x^*)} \bar{\mu}_j \nabla g_j(x^*) = 0 \), which contradicts H2 and \( J' \cap J(x^*) \neq \phi \). Thus, \( \{ \mu^k_{j_k} \} \) is bounded.

(iii) We prove \( \lim_{k \to K} d^k = 0 \) first. By contradiction, we assume that \( \lim_{k \to K} d^k \neq 0 \). Then there exist an infinite index subset \( K_1 \subseteq K \) and a constant \( \sigma > 0 \) such that \( \| d^k \| \geq \sigma, k \in K_1 \). Denote
\[
w_k(t) = F(x^k + td^k + t^2 d^k) - F_k + \alpha t (d^k)^T H_k d^k,
\]
then we have
\[
w_k(t) = \max \{ f_i(x^k + td^k + t^2 d^k) - F_k + \alpha t (d^k)^T H_k d^k, i \in I \}
\]
\[
= \max \{ f_i(x^k) + t \nabla f_i(x^k)^T d^k + o(t) - F_k + \alpha \alpha (d^k)^T H_k d^k, i \in I \}.
\]
We further denote
\[ a_{ki}(t) = f_i(x^k) + t \nabla f_i(x^k)^T d^k + o(t) - F_k + \alpha t(d^k)^T H_k d^k, \quad i \in I. \]

So from the definition of \( F_k \), one has
\[
a_{ki}(t) \leq f_i(x^k) + t \nabla f_i(x^k)^T d^k + o(t) - F(x^k) + \alpha t(d^k)^T H_k d^k \\
= t( f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k)) + (1-t)( f_i(x^k) - F(x^k)) + \alpha t(d^k)^T H_k d^k + o(t), \quad i \in I.
\]

Then from the constraint conditions of (2.1), Lemma 2.2 (i), H3 and \( \alpha \in (0, 0.5) \), we have for \( i \in I_k \)
\[
a_{ki}(t) \leq t z_k + \alpha t(d^k)^T H_k d^k + o(t) \leq (\alpha - \frac{1}{2}) t(d^k)^T H_k d^k + o(t) = (\alpha - \frac{1}{2}) \alpha t \sigma^2 + o(t) \leq 0.
\]

For \( i \in I \setminus I_k \), in view of H5, we get easily that \( a_{ki}(t) \leq 0 \) for \( t > 0 \) small enough. Thus, \( w_k(t) \leq 0 \) and (2.9) holds for \( t > 0 \) small enough and all \( k \in K_1 \).

On the other hand, one also gets
\[
g_j(x^k + t d^k + t^2 \bar{d}^k) = g_j(x^k) + t \nabla g_j(x^k)^T d^k + o(t) \\
= t(g_j(x^k) + \nabla g_j(x^k)^T d^k) + (1-t) g_j(x^k) + o(t) \\
\leq tc_k z_k + o(t) \leq - \frac{1}{2} c_k t(d^k)^T H_k d^k + o(t) \\
\leq - \frac{1}{2} \alpha c_k t \sigma^2 + o(t) \leq - \frac{1}{4} \alpha c_k t \sigma^2 + o(t) \leq 0, \quad j \in J_k.
\]

Also, taking into account H5, we can easily get \( g_j(x^k + t d^k + t^2 \bar{d}^k) \leq 0 \) for \( j \in J \setminus J_k \) and \( t > 0 \) small enough. Thus, (2.10) holds for \( t > 0 \) small enough and all \( k \in K_1 \). Hence there exists a constant \( \bar{t} > 0 \) such that the stepsizes \( t_k \geq \bar{t} \), \( \forall k \in K_1 \), and \( \|h_k d^k\| \geq \bar{t}\|d^k\| \geq \bar{t} \sigma \), \( k \in K_1 \), which contradicts Lemma 3.1(ii). So \( \lim_{k \in K} d^k = 0 \). \( \lim_{k \in K} \bar{d}^k = 0 \) follows from Step 3 of Algorithm A.

Finally, we prove \( \lim_{k \in K} z_k = 0 \). From Lemma 2.2(i) and the constraints of QP(\( x_k, H_k \)), one has
\[
0 \geq z_k \geq f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) = \nabla f_i(x^k)^T d^k, \quad i \in I(x^k) \subseteq I_k,
\]
this along with \( \lim_{k \in K} d^k = 0 \) shows that \( \lim_{k \in K} z_k = 0 \). The whole proof is completed. \( \square \)

**Theorem 3.1.** Suppose that H1–H5 hold. Then Algorithm A either stops at a stationary point of problem (P) in a finite number of iterations, or generates an infinite sequence \( \{x^k\} \) of points such that each accumulation \( x^* \) of \( \{x^k\} \) is a stationary point of (P).

**Proof.** If Algorithm A stops at the \( k \)th iteration, then, from Step 1 of Algorithm A and Lemma 2.2(iii), we know that \( x^k \) is a stationary point of (P). Now we suppose that Algorithm A generates an infinite sequence \( \{x^k\} \). Without loss of generality suppose that the infinite index set \( K \) satisfies (3.4). The rest proof is divided into two cases.

**Case A:** Suppose that \( \lim_{k \in K} c_k = c > 0 \). In view of Lemma 3.2(ii) and (iii), we can assume without loss of generality that the infinite index set \( K \) also satisfies
\[
\bar{\lambda}^*_k \rightarrow \lambda^*, \quad \mu^*_k \rightarrow \mu^*, \quad d^k \rightarrow 0, \quad z_k \rightarrow 0, \quad k \in K.
\]

Passing to the limit \( k \in K \) and \( k \rightarrow \infty \) in (2.2), we have
\[
\begin{align*}
\sum_{i \in I} \lambda^*_i \nabla f_i(x^*) + \sum_{j \in J} \mu^*_j \nabla g_j(x^*) &= 0; \\
\sum_{i \in I} \lambda^*_i + c \sum_{j \in J} \mu^*_j &= 1, \\
0 \leq \lambda^*_i \perp (f_i(x^*) - F(x^*)) &\leq 0, \quad i \in \tilde{I}; \\
0 \leq \mu^*_j \perp g_j(x^*) &\leq 0, \quad j \in \tilde{J}.
\end{align*}
\]

(3.7)

Basing on the relationship (3.7), similar to the proof of Lemma 2.2(ii), we can conclude that \( x^* \) is a stationary point of (P).
Case B: If \( \lim_{k \to K} c_k = c_\ast = 0 \). In view of the definition of \( c_k \) at Step 5, we have \( \lim_{k \to K} d^{k-1} = 0 \), and \( \lim_{k \to K} z_{k-1} = 0 \) follows from the constraints of QP(\( x^k, H_k \)). Thus,

\[
\lim_{k \to K} \|x^k - x^{k-1}\| \leq \lim_{k \to K} (t_{k-1} \|d^{k-1}\| + t_{k-1}^2 \|\tilde{d}^{k-1}\|) \leq \lim_{k \to K} (M + 1) \|d^{k-1}\| = 0.
\]

This along with \( \lim_{k \to K} x^k = x^\ast \) implies that \( \lim_{k \to K} x^{k-1} = x^\ast \). Summarize the discussion above, and let \( \tilde{K} = \{k - 1 : k \in K\} \), then

\[
\lim_{k \to \tilde{K}} x^k = x^\ast, \quad \lim_{k \to \tilde{K}} d^k = 0, \quad \lim_{k \to \tilde{K}} z_k = 0.
\]

Thus, there exists \( K' \subseteq \tilde{K} \) such that \( K' \) satisfies (3.4). In view of Lemma 3.2(ii), \( K' \) also satisfies

\[
\hat{\lambda}^k_{t_k^k} \to \lambda', \quad \mu^k_{j_t^k} \to \mu', \quad d^k \to 0, \quad z_k \to 0, \quad x^k \to x^\ast, \quad k \in K'.
\]

Passing to the limit \( k \to K' \) and \( k \to \infty \) in (2.2), we know that \( x^\ast \) is a stationary point of (P). The proof is completed. \( \square \)

The following results are helpful in analyzing the rate of convergence of the proposed algorithm.

**Lemma 3.3.** Suppose that H1–H5 hold. Then (i) \( \lim_{k \to \infty} d^k = \lim_{k \to \infty} \tilde{d}^k = 0 \), \( \lim_{k \to \infty} z_k = 0 \), and (ii) \( \lim_{k \to \infty} \|x^{k+1} - x^k\| = 0 \).

**Proof.** (i) We assume that \( x^\ast \) is an accumulation point of \( \{x^k\} \), then \( x^\ast \) is a stationary point of (P) from Theorem 3.1. Suppose by contradiction that \( \lim_{k \to \infty} d^k \neq 0 \). Then taking into account the boundedness of \( \{(x^k, c_k, H_k)\} \), there exist an infinite index subset \( \tilde{K}' \) and a constant \( \delta > 0 \) such that \( \|d^k\| \geq \delta \) and (3.4) holds for \( \tilde{K}' \). Furthermore, in view of Lemma 3.2(i), we can assume that there exists another infinite subset \( \tilde{K}' \subseteq \tilde{K} \) such that

\[
x^k \to x^\ast, \quad d^k \to d^\ast \neq 0, \quad z_k \to z^\ast, \quad k \in \tilde{K}'.
\]

Thus, from H3 and Lemma 2.2(i), we have

\[
z^\ast = \lim_{k \in \tilde{K}'} z_k \leq \lim_{k \in \tilde{K}'} -\frac{1}{2} (d^k)^T H_k d^k = -\frac{1}{2} (d^\ast)^T H_\ast d^\ast < 0.
\]

Also from the constraints of QP(\( x^k, H_k \)), one has

\[
\nabla f_i(x^\ast)^T d^\ast \leq z^\ast < 0, \quad i \in I(x^\ast) \subseteq I_{t_k^k}, \quad \nabla g_j(x^\ast)^T d^\ast \leq c_\ast z^\ast \leq 0, \quad j \in J(x^\ast) \subseteq J_{t_k^k}^k.
\]

(3.8)

Note that \( x^\ast \) is a stationary point of (P), let \( \tilde{\lambda}^\ast, \tilde{\mu}^\ast \) be the multiplier vectors corresponding to \( x^\ast \), then we have from (2.3)

\[
\left\{
\begin{aligned}
\sum_{i \in I(x^\ast)} \tilde{\lambda}^*_i \nabla f_i(x^\ast) + \sum_{j \in J(x^\ast)} \tilde{\mu}^*_j \nabla g_j(x^\ast) &= 0, \\
\sum_{i \in I(x^\ast)} \tilde{\lambda}^*_i &= 1, \quad \tilde{\lambda}^*_i \geq 0, \quad \tilde{\mu}^*_j \geq 0, \quad i \in I(x^\ast); \quad j \in J(x^\ast),
\end{aligned}
\right.
\]

which contradicts (3.8). Thus, \( d^k \to 0, k \to \infty \). Further, \( \tilde{d}^k \to 0, k \to \infty \) follows from Step 3 of Algorithm A, and \( z_k \to 0, k \to \infty \) follows from \( 0 \geq z_k \geq \nabla f_i(x^k)^T d^k \) for \( i \in I(x^k) \).

(ii) From (i), we have

\[
\lim_{k \to \infty} \|x^{k+1} - x^k\| = \lim_{k \to \infty} \|t_k d^k + t_k^2 \tilde{d}^k\| \leq \lim_{k \to \infty} (\|d^k\| + \|\tilde{d}^k\|) = 0.
\]

So \( \lim_{k \to \infty} \|x^{k+1} - x^k\| = 0 \). \( \square \)
4. Strong and superlinear convergence

In this section, we discuss the strong and superlinear convergence of the proposed algorithm. For this goal, we should strengthen the assumption H2 as follows.

$H2^+$. For some $i_k \in I(x)$, vectors $\{\nabla f_i(x) - \nabla f_i(x), \ i \in I(x) \setminus \{i_k\}; \nabla g_j(x), \ j \in J(x)\}$ are linearly independent for each $x \in X$.

**Remark 4.1.** Via simple analysis, we know that $H2^+$ and any one of the two following conditions are really equivalent:

1. For each $t \in I(x)$, vectors $\{\nabla f_i(x) - \nabla f_i(x), \ i \in I(x \setminus \{t\}); \nabla g_j(x), \ j \in J(x)\}$ are linearly independent.

2. Vectors $\left\{ \left( \begin{array}{c} -1 \\ \nabla f_i(x) \end{array} \right), \ i \in I(x); \left( \begin{array}{c} 0 \\ \nabla g_j(x) \end{array} \right), \ j \in J(x) \right\}$ in $\mathbb{R}^{n+1}$ are linearly independent.

**Remark 4.2.** The assumption $H2^+$ is stronger than $H2$ in Section 2, thus we only use a weak assumption $H2$ (MFCQ) in Section 2 to ensure the global convergence of the proposed algorithm.

Also, the following assumption which is used in Refs. [20,19] is necessary.

$H6$. (i) The functions $f_i$ ($i \in I$) and $g_j$ ($j \in J$) are all twice continuously differentiable for any $x \in X$; (ii) The sequence $\{x^k\}$ yielded by the proposed algorithm possesses an accumulation point $x^*$ with the corresponding multipliers $\bar{\lambda}^*, \bar{\mu}^*$ (by Theorem 3.1, $x^*$ is a stationary point of problem (P)) satisfies the following second-order sufficiency conditions with strict complementarity:

$$
\begin{align*}
&d^T \nabla^2_x L(x^*, \bar{\lambda}^*, \bar{\mu}^*) d > 0, \quad \forall d \in \Psi, \\
&\Psi \triangleq \{d \neq 0 : (\nabla f_i(x^*) - \nabla f_j(x^*))^T d = 0, \ \forall i, j \in I(x^*); \nabla g_j(x^*)^T d = 0, \ \forall j \in J(x^*)\}; \\&\bar{\lambda}^*_i > 0, \quad \forall i \in I(x^*); \quad \bar{\mu}^*_j > 0, \quad \forall j \in J(x^*),
\end{align*}
$$

(4.1) (4.2)

where

$$
\nabla^2_x L(x^*, \bar{\lambda}^*, \bar{\mu}^*) = \sum_{i \in I} \bar{\lambda}^*_i \nabla^2 f_i(x^*) + \sum_{j \in J} \bar{\mu}^*_j \nabla^2 g_j(x^*).
$$

First of all, we give a lemma to show that $x^*$ is an isolated stationary point of (P) under certain conditions.

**Lemma 4.1.** Suppose that $x^*$ is a stationary point of (P) and the stated assumptions hold. Then $x^*$ is an isolated stationary point of (P).

The proof of this lemma is similar to the one of Theorem 1.2.5 in [6], thus it is omitted here.

**Theorem 4.1.** Suppose that $H2^+$, $H3$–$H6$ hold. Then $\lim_{k \to \infty} x^k = x^*$, i.e., the proposed algorithm is strongly convergent.

**Proof.** From Lemma 4.1, we know that $x^*$ is an isolated stationary point of (P). Furthermore, in view of Theorem 3.1, one can conclude that $x^*$ is an isolated limit point of $\{x^k\}$, and this together with Lemma 3.3 (ii) implies $\lim_{k \to \infty} x^k = x^*$ (see Theorem 1.1.5 in [6]). □

**Lemma 4.2.** Suppose that $H2^+$, $H3$–$H6$ hold. Then

(i) the multiplier vectors $\bar{\lambda}^k_{I_k}$ and $\bar{\mu}^k_{J_k}$ corresponding to the solution of $QP(x^k, H_k)$ satisfy

$$
\sigma_k \to 1, \quad \bar{\lambda}^k = (\bar{\lambda}^k_{I_k}, 0_{I \setminus I_k}) = (\bar{\lambda}^k_{I(x^*)}, 0_{I \setminus I(x^*)}) \to \bar{\lambda}^*; \quad \bar{\mu}^k = (\bar{\mu}^k_{J_k}, 0_{J \setminus J_k}) = (\bar{\mu}^k_{J(x^*)}, 0_{J \setminus J(x^*)}) \to \bar{\mu}^*,
$$

and for $k$ large enough

$$
I_k = \{i \in I : \bar{\lambda}^k_i > 0\} \equiv I(x^*), \quad J_k = \{j \in J : \bar{\mu}^k_j > 0\} \equiv J(x^*);
$$

(ii) $\lim_{k \to \infty} x^k = x^*$.
(ii) \(\text{QP}(x^k, d^k, H_k)\) (2.8) always has a solution for \(k\) large enough, and its solution \((y_k, d^k)\) satisfies \(\lim_{k \to \infty} (y_k, d^k) = (0, 0)\);

(iii) the multiplier vectors \(\lambda_{ik}^k\) and \(\mu_{jk}^k\) corresponding to the solution \((y_k, d^k)\) of \(\text{QP}(x^k, d^k, H_k)\) satisfy

\[
\{ i \in I_{ik}^k : \lambda_{ik}^k > 0 \} \equiv I(x^*), \quad \{ j \in J_{jk}^k : \mu_{jk}^k > 0 \} \equiv J(x^*).\]

**Proof.**

(i) From formula (2.2), Lemma 3.3(i), H5 and Theorem 4.1 for \(k\) large enough we have \(I_k \subseteq I(x^*) \subseteq I_{ik}^k, J_k \subseteq J(x^*) \subseteq J_{jk}^k\) and

\[
H_k d^k + \sum_{i \in I(x^*)} \lambda_{ik}^k \nabla f_i(x^k) + \sum_{j \in J(x^*)} \mu_j^k \nabla g_j(x^k) = 0,
\]

\[
\sum_{i \in I(x^*)} \lambda_{ik}^k + c_k \sum_{j \in J(x^*)} \mu_j^k = 1. \tag{4.3}
\]

On the other hand, from Lemma 3.3(i) and Step 5 of Algorithm A, one has

\[
c_k \to 0, \quad \sigma_k = \sum_{i \in I_{ik}^k} \lambda_{ik}^k = \sum_{i \in I_k} \lambda_{ik}^k = \sum_{i \in I(x^*)} \lambda_{ik}^k = 1 - c_k \sum_{j \in J_k} \mu_j^k \to 1, \quad k \to \infty.
\]

Thus, dividing the first equality of (4.3) by \(\sigma_k\), we get

\[
\frac{1}{\sigma_k} H_k d^k + \sum_{i \in I(x^*)} \frac{\lambda_{ik}^k}{\sigma_k} \nabla f_i(x^k) + \sum_{j \in J(x^*)} \frac{\mu_j^k}{\sigma_k} \nabla g_j(x^k) = 0,
\]

with

\[
\bar{\lambda}_{ik}^k = \lambda_{ik}^k / \sigma_k, \quad i \in I; \quad \bar{\mu}_{jk}^k = \mu_j^k / \sigma_k, \quad j \in J, \tag{4.5}
\]

and for some \(i_k \in I(x^*) \subseteq I(x^*)\)

\[
\frac{1}{\sigma_k} H_k d^k + \nabla f_{i_k}(x^k) + \sum_{i \in I(x^*) \setminus \{ i_k \}} \bar{\lambda}_{ik}^k (\nabla f_i(x^k) - \nabla f_{i_k}(x^k)) + \sum_{j \in J(x^*)} \bar{\mu}_{jk}^k \nabla g_j(x^k) = 0. \tag{4.6}
\]

In view of Theorem 4.1, Lemma 3.3 (i), H2+ and (4.6), one has

\[
\bar{\lambda}_{ik}^k \to \bar{\lambda}_{ik}^*, \quad i \in I; \quad \bar{\mu}_{jk}^k \to \bar{\mu}_{jk}^*, \quad j \in J, \quad k \to \infty,
\]

this along with (4.5) and \(\sigma_k \to 1\) implies that

\[
\lambda_{ik}^k \to \lambda_{ik}^*, \quad \mu_{jk}^k \to \mu_{jk}^*, \quad k \to \infty.
\]

Taking into account \(I_k \subseteq I(x^*)\) and \(J_k \subseteq J(x^*)\), for \(k\) large enough, we get from (2.2)

\[
\lambda_{ik}^k = 0, \quad i \notin I(x^*), \quad \mu_{jk}^k = 0, \quad j \notin J(x^*).
\]

Furthermore, from the strict complementarity (4.2), we have

\[
I_k = \{ i \in I : \lambda_{ik}^k > 0 \} \equiv I(x^*), \quad J_k = \{ j \in J : \mu_{jk}^k > 0 \} \equiv J(x^*).
\]

(ii) Define matrix \(A_k\) and vector \(b_k\) by

\[
A_k = A(x^k) = \begin{pmatrix} -1 & 0 \\ \nabla f_i(x^k) & \nabla g_j(x^k) \end{pmatrix}, \quad i \in I(x^*); \quad j \in J(x^*),
\]

\[
b_k = \begin{pmatrix} F(x^k + d^k) - f_i(x^k + d^k), \quad i \in I(x^*) \\ c_k^z \| d^k \| - \| d^k \|^\tau - g_j(x^k + d^k), \quad j \in J(x^*) \end{pmatrix}.
\]
From H2+-2, we know that \((A^T A)\) and \((A_k^T A_k)^{-1}\) are well defined for \(k\) large enough. Thus, in view of Lemma 3.3(i) and Lemma 4.2(ii), one obtains \((\bar{y}_k, (d_{0k}^k)^T)\) is a feasible solution of QP
\[
\min_{t, d} \quad t + \frac{1}{2} d^T H_0 d
\]
\[\text{s.t.} \quad f_i(x) + \nabla f_i(x)^T d - F(x) = 0, \quad i \in I,\]
\[g_j(x) + \nabla g_j(x)^T d - c_i^T z_i d = 0, \quad j \in J(x).\]
From the former system, we know that \((y_*, d^k)\) is a KKT point of QP\((x^*, H_*)\). On the other hand, in view of Theorem 3.1 and the structure of QP\((x^*, H^*)\), one can conclude that \((0, 0) \in R^{n+1}\) is the unique KKT point of QP\((x^*, H_*)\). Thus, \((y_*, d^k) = (0, 0)\).

(iii) Finally, based on \((y_k, d^k) \xrightarrow{k \to \infty} (0, 0)\), the proof of conclusion (iii) is similar to the one of conclusion (i), so it is omitted here. The whole proof is finished. □

**Lemma 4.3.** Suppose that \((y_k, d^k)\) is the solution of QP\((x^k, d^k, H_k)\) and the stated assumptions are all satisfied. Then

\[
\|d^k\| = O(\|d^k\|^2) + O(\|c_k^T z_k\|) = O(\|d^k\|) \quad \|\bar{d}^k\| = o(\|d^k\|);
\]

\[
O(\|d^k\| \cdot \|\bar{d}^k\|) = O(\|d^k\|^3) + o(\|c_k^T z_k\| \cdot \|d^k\|) = o(\|d^k\|^2), \quad O(\|\bar{d}^k\|^2) = O(\|d^k\|^3) + o(\|c_k^T z_k\| \cdot \|d^k\|) = o(\|d^k\|^2).
\]

**Proof.** (i) Take into account Lemma 2.3 and Lemma 4.2(iii), we know that \((y_k, d^k)\) is a KKT point of QP\((x^k, d^k, H_k)\) with corresponding multipliers \(\bar{\lambda}^{ik}_{J(x_k)}\) and \(\bar{\mu}^{jk}_{J(x_k)}\) such that

\[
\begin{aligned}
\left( \begin{array}{c}
H_k (d^k + \frac{1}{\sigma_k} \bar{d}^k) + \sum_{i \in I(x^*) \setminus \{i_k\}} \tilde{\lambda}^i (x^k) (\nabla f_i (x^k)) + \sum_{j \in J(x^*)} \tilde{\mu}^j (x^k) (\nabla g_j (x^k)) = 0, \\
f_i (x^k + d^k) + \sum_{j \in J(x^*)} \tilde{\mu}^j (x^k) (\nabla g_j (x^k)) - F(x^k + d^k) = y_k, \\
g_j (x^k + d^k) + \sum_{j \in J(x^*)} \tilde{\mu}^j (x^k) (\nabla g_j (x^k)) = -\|d^k\|, \\
0 \leq \|\bar{d}^k\| \leq \|d^k\|
\end{array} \right)
\end{aligned}
\]

In view of Lemma 4.2(i) and Taylor expansion, for some \(i_k \in I(x^k) \subseteq I(x^*)\) the relationships above can be rewritten as follows:

\[
\begin{aligned}
\nabla f_i (x^k) + H_k \bar{d}^k + \frac{1}{\sigma_k} H_k d^k + \sum_{i \in I(x^*) \setminus \{i_k\}} \tilde{\lambda}^i (x^k) (\nabla f_i (x^k)) - \nabla f_{i_k} (x^k) + \sum_{j \in J(x^*)} \tilde{\mu}^j (x^k) (\nabla g_j (x^k)) = 0, \\
(\nabla f_i (x^k) - \nabla f_{i_k} (x^k)) \bar{d}^k = O(\|d^k\|^2), \\
\nabla g_j (x^k) \bar{d}^k = -\|d^k\|^2 + c^T z_k \|d^k\| - c_k z_k + O(\|d^k\|^2),
\end{aligned}
\]

That is

\[
\nabla f_i (x^k) + H_k \bar{d}^k + \frac{1}{\sigma_k} H_k d^k + N_k \bar{u}^k = 0; \quad N_k^T \bar{d}^k = O(\|d^k\|^2) + O(|c_k^T z_k|) + O(|c_k^T z_k| \cdot \|d^k\|),
\]

where vector \(\bar{u}^k = (\tilde{\lambda}^i, i \in I(x^*) \setminus \{i_k\}; \tilde{\mu}^j, j \in J(x^*))\) and matrix

\[
N_k = N(x^k) = (\nabla f_i (x^k) - \nabla f_{i_k} (x^k), \quad i \in I(x^*) \setminus \{i_k\}; \nabla g_j (x^k), \quad j \in J(x^*)),
\]

On the other hand, we also have from (4.6)

\[
\nabla f_i (x^k) + \frac{1}{\sigma_k} H_k d^k + N_k \bar{u}^k = 0,
\]

with \(\bar{u}^k = (\tilde{\lambda}^i, i \in I(x^*) \setminus \{i_k\}; \tilde{\mu}^j, j \in J(x^*))\). Substituting (4.9) into (4.7), one has

\[
\begin{pmatrix}
H_k \\
N_k^T
\end{pmatrix}
\begin{pmatrix}
\bar{d}^k \\
\bar{u}^k - \bar{u}^k
\end{pmatrix}
= O(\|d^k\|^2) + O(|c_k^T z_k|) + O(|c_k^T z_k| \cdot \|d^k\|).
\]

Again, from H2+−1, H3 and Theorem 4.1, it is not difficult to know that the coefficient matrix of (4.10) denoted by \(D_k\) is uniformly nonsingular and there exists a constant \(\rho > 0\) such that \(\|D_k^{-1}\| \leq \rho\). This together (4.10) shows that

\[
\|\bar{d}^k\| = O(\|d^k\|^2) + O(|c_k^T z_k|) + O(|c_k^T z_k| \cdot \|d^k\|).
\]
On the other hand, from the constraints of QP\((x^k, H_k)\), we have
\[
|z_k| \leq F(x^k) - f_i(x^k) - \nabla f_i(x^k)^T d_k \leq F(x^k) - f_i(x^k) + \|\nabla f_i(x^k)\| \cdot \|d_k\|
\]
\[
= \|\nabla f_i(x^k)\| \cdot \|d_k\|, \quad i \in I(x^k),
\]
thus \(|z_k| = O(\|d_k\|)\), which combining \(c_k \to 0\) further shows that \(\bar{d}_k = o(\|d_k\|)\).

(ii) The relationships in part (ii) follow from (i) immediately. \(\square\)

Now, we define the projection matrix
\[
P_k = E_n - N_k(N_k^TN_k)^{-1}N_k^T.
\]

**Lemma 4.4.** Suppose that H2+, H3–H6 hold. Then
\[
d_k = P_kd_k + \bar{d}_k, \quad \bar{d}_k = O(\|\bar{\zeta}(x^k)\|) + o(\|d_k\|),
\]
with
\[
\bar{d}_k = N_k(N_k^TN_k)^{-1}N_k^T d_k, \quad \bar{\zeta}(x^k) = \begin{pmatrix} f_{ik}(x^k) - f_i(x^k), & i \in I(x^*)\backslash\{i_k\} \\ -g_j(x^k), & j \in J(x^*) \end{pmatrix}.
\]

**Proof.** In view of H2+ and Theorem 4.1, we know that \(P_k\) is well defined, thus we get \(d_k = P_kd_k + \bar{d}_k\). On the other hand, from Lemma 4.2(i), one has
\[
f_i(x^k) + \nabla f_i(x^k)^T d_k = F(x^k) + z_k, \quad i \in I(x^*); \quad g_j(x^k) + \nabla g_j(x^k)^T d_k = c_k z_k, j \in J(x^*).
\]
Hence from the equalities above, we obtain
\[
N_k^T d_k = \begin{pmatrix} f_{ik}(x^k) - f_i(x^k), & i \in I(x^*)\backslash\{i_k\} \\ -g_j(x^k), & j \in J(x^*) \end{pmatrix} + \begin{pmatrix} 0_{I(x^*)\backslash\{i_k\}} \\ c_k z_k e_{J(x^*)} \end{pmatrix},
\]
where \(e_{J(x^*)} = (1, \ldots, 1)^T \in R^{\mid J(x^*)\mid}\). Thus, from Lemma 4.3(i), one has
\[
\bar{d}_k = N_k(N_k^TN_k)^{-1} \bar{\zeta}(x^k) + o(\|d_k\|) = O(\|\bar{\zeta}(x^k)\|) + o(\|d_k\|).
\]
The proof is finished. \(\square\)

To ensure the step size unit can be accepted, the following assumption about the matrix \(H_k\) is necessary.

H7. Suppose that \(\|P_k(\nabla^2_{xx} L(x^k, \tilde{x}^k, \tilde{\mu}^k) - H_k)d_k\| = o(\|d_k\|)\), where \(\tilde{x}^k\) and \(\tilde{\mu}^k\) are yielded by (4.5).

**Remark 4.3.** According to Theorem 4.1, Lemma 4.2 and (4.5), it is easy to get
\[
\|P_k(\nabla^2_{xx} L(x^k, \tilde{x}^k, \tilde{\mu}^k) - H_k)d_k\| = o(\|d_k\|) \iff \|P_k(\nabla^2_{xx} L(x^*, \tilde{x}^*, \tilde{\mu}^*) - H_k)d_k\| = o(\|d_k\|).
\]

**Theorem 4.2.** Under all the assumptions H2+ and H3–H7 above, the step size of Algorithm A always equals one, i.e., \(t_k \equiv 1\), if \(k\) is sufficiently large.

**Proof.** It is sufficient to show that the inequalities (2.9) and (2.10) are satisfied with \(t = 1\) and \(k\) large enough. Firstly, in view of Taylor expansion, Lemma 3.3(i), Lemma 4.2(iii) and Lemma 4.3(ii), we have for \(j \in J(x^*) \subseteq \tilde{J}_k^k\)
\[
g_j(x^k + d_k + \bar{d}_k) = g_j(x^k + d_k) + \nabla g_j(x^k + d_k)^T \bar{d}_k + O(\|\bar{d}_k\|^2)
\]
\[
= g_j(x^k + d_k) + \nabla g_j(x^k)^T d_k + O(\|d_k\| \cdot \|\bar{d}_k\|) + O(\|\bar{d}_k\|^2)
\]
\[
= c^*_j z_k \|d_k\| - \|d_k\|^2 + O(\|d_k\|^3) + o(c^*_j z_k \|d_k\|).
\]
(4.11)
Thus, for $k$ large enough, we have $g_j(x^k + d^k + \tilde{d}^k) \leq 0$, $j \in J(x^*)$. For $j \in J \setminus J(x^*)$, we can easily get $g_j(x^k + d^k + \tilde{d}^k) \leq 0$ since $\lim_{k \to \infty} g_j(x^k + d^k + \tilde{d}^k) = g_j(x^*) < 0$. Hence the inequality (2.10) holds for $t = 1$ and $k$ large enough.

Secondly, we prove that $F(x^k + d^k + \tilde{d}^k) \leq F(x^k) - z(d^k)^T H_k d^k$, which implies that the inequality (2.9) holds for $t = 1$. From Lemma 4.2(iii) and Taylor expansion, we have

$$f_i(x^k + d^k + \tilde{d}^k) = f_i(x^k + d^k) + \nabla f_i(x^k)^T \tilde{d}^k + O(\|d^k\| \cdot \|\tilde{d}^k\|) + O(\|\tilde{d}^k\|^2)$$

$$= F(x^k + d^k) + y_k + O(\|d^k\| \cdot \|\tilde{d}^k\|) + O(\|\tilde{d}^k\|^2), \quad i \in I(x^*).$$

Thus, one obtains

$$f_i(x^k + d^k + \tilde{d}^k) = f_i(x^k + d^k + \tilde{d}^k) + O(\|d^k\| \cdot \|\tilde{d}^k\|) + O(\|\tilde{d}^k\|^2), \quad \forall i, j \in I(x^*).$$

(4.12)

For $k$ large enough, we know that there exists an index $j_k \in I(x^k + d^k + \tilde{d}^k) \subseteq I(x^*)$. Thus, combining (4.12) and $\sum_{i \in I(x^*)} \lambda^k_i = 1$ with Lemma 4.3(ii), we have

$$F(x^k + d^k + \tilde{d}^k) = f_{j_k}(x^k + d^k + \tilde{d}^k) = \left( \sum_{i \in I(x^*)} \lambda^k_i \right) f_{j_k}(x^k + d^k + \tilde{d}^k)$$

$$= \sum_{i \in I(x^*)} \lambda^k_i f_{j_k}(x^k + d^k + \tilde{d}^k)$$

$$= \sum_{i \in I(x^*)} \lambda^k_i f_i(x^k + d^k + \tilde{d}^k) + O(\|d^k\| \cdot \|\tilde{d}^k\|) + O(\|\tilde{d}^k\|^2)$$

$$= \sum_{i \in I(x^*)} \lambda^k_i \left( f_i(x^k) + \nabla f_i(x^k)^T (d^k + \tilde{d}^k) + \frac{1}{2} (d^k)^T \nabla^2 f_i(x^k) d^k \right)$$

$$+ \nabla \left( \sum_{i \in I(x^*)} \lambda^k_i \nabla^2 f_i(x^k) d^k \right) + \frac{1}{2} (d^k)^T \left( \sum_{i \in I(x^*)} \lambda^k_i \nabla^2 f_i(x^k) d^k \right) \tilde{d}^k + O(\|d^k\|^2).$$

On the other hand, from (4.4), one has

$$\frac{1}{\sigma_k} (d^k + \tilde{d}^k)^T H_k d^k + \sum_{i \in I(x^*)} \lambda^k_i \nabla f_i(x^k)^T (d^k + \tilde{d}^k) + \sum_{j \in J(x^*)} \tilde{\mu}^k_j \nabla g_j(x^k)^T (d^k + \tilde{d}^k) = 0.$$

Thus, from Lemma 4.3(ii), $\sigma_k \leq 1$ and $\sigma_k \to 1$, we further have

$$F(x^k + d^k + \tilde{d}^k) = \sum_{i \in I(x^*)} \lambda^k_i f_i(x^k) - \frac{1}{\sigma_k} (d^k + \tilde{d}^k)^T H_k d^k - \sum_{j \in J(x^*)} \tilde{\mu}^k_j \nabla g_j(x^k)^T (d^k + \tilde{d}^k)$$

$$+ \frac{1}{2} (d^k)^T \left( \sum_{i \in I(x^*)} \lambda^k_i \nabla^2 f_i(x^k) \right) d^k + O(\|d^k\|^2).$$
Thus, substituting the equality above into (4.13) and combining with Lemma 4.3, we have

\[ F(x^k + d^k + \bar{d}^k) \leq \sum_{i \in I(x^*)} \tilde{\lambda}_i^k f_i(x^k) - (d^k)^T H_k d^k - \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k)^T (d^k + \bar{d}^k) \]

\[ + \frac{1}{2} (d^k)^T \left( \sum_{i \in I(x^*)} \tilde{\lambda}_i^k \nabla^2 f_i(x^k) \right) d^k + o(\|d^k\|^2). \]  

(4.13)

Again, from (4.11), Lemma 4.3 and Taylor expansion, we get

\[ o(\|d^k\|^2) = g_j(x^k + d^k + \bar{d}^k), \quad j \in J(x^*), \]

and

\[ o(\|d^k\|^2) = \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k + d^k + \bar{d}^k) \]

\[ = \sum_{j \in J(x^*)} \tilde{\mu}_j^k \left( g_j(x^k) + \nabla g_j(x^k)^T (d^k + \bar{d}^k) + \frac{1}{2} (d^k)^T \nabla^2 g_j(x^k) d^k \right) \]

\[ + o(\|d^k + \bar{d}^k\|^2) + O(\|d^k\| \cdot \|\bar{d}^k\|) + O(\|\bar{d}^k\|^2), \]

\[ - \sum_{j \in J(x^*)} \tilde{\mu}_j^k \left( \nabla g_j(x^k)^T (d^k + \bar{d}^k) \right) = \sum_{j \in J(x^*)} \tilde{\mu}_j^k \left( g_j(x^k) + \frac{1}{2} (d^k)^T \nabla^2 g_j(x^k) d^k \right) + o(\|d^k\|^2). \]

Substituting the equality above into (4.13) and combining with Lemma 4.3, we have

\[ F(x^k + d^k + \bar{d}^k) \leq \sum_{i \in I(x^*)} \tilde{\lambda}_i^k f_i(x^k) - (d^k)^T H_k d^k + \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k) + \frac{1}{2} (d^k)^T \nabla^2 L(x^k, \tilde{\lambda}^k, \tilde{\mu}^k) d^k \]

\[ + o(\|d^k\|^2) \]

\[ = \sum_{i \in I(x^*)} \tilde{\lambda}_i^k f_i(x^k) + \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k) - F(x^k) + \frac{1}{2} (d^k)^T \nabla^2 L(x^k, \tilde{\lambda}^k, \tilde{\mu}^k) - H_k d^k \]

\[ + F(x^k) - \alpha (d^k)^T H_k d^k + \left( \alpha - \frac{1}{2} \right) (d^k)^T H_k d^k + o(\|d^k\|^2). \]

On the other hand, from Lemmas 4.4, 4.2(i), H7, \( f_i(x^k) = F(x^k) \) and the strict complementarity (4.2), there exists a constant \( v > 0 \) such that

\[ \sum_{i \in I(x^*)} \tilde{\lambda}_i^k f_i(x^k) + \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k) - F(x^k) + \frac{1}{2} (d^k)^T \nabla^2 L(x^k, \tilde{\lambda}^k, \tilde{\mu}^k) - H_k d^k \]

\[ = \sum_{i \in I(x^*) \setminus \{i_k\}} \tilde{\lambda}_i^k (f_i(x^k) - f_{i_k}(x^k)) + \sum_{j \in J(x^*)} \tilde{\mu}_j^k g_j(x^k) \]

\[ + \frac{1}{2} (P_k d^k + \bar{d}^k)^T \nabla^2 L(x^k, \tilde{\lambda}^k, \tilde{\mu}^k) - H_k d^k \]

\[ \leq - \sum_{i \in I(x^*) \setminus \{i_k\}} v |f_i(x^k) - f_{i_k}(x^k)| - \sum_{i \in J(x^*)} v |g_j(x^k)| + o(\|d^k\|^2) + o(\|\xi(x^k)\|) \]

\[ \leq - v \|\xi(x^k)\| + o(\|\xi(x^k)\|) + o(\|d^k\|^2) \leq o(\|d^k\|^2). \]

Thus,

\[ F(x^k + d^k + \bar{d}^k) \leq F(x^k) - \alpha (d^k)^T H_k d^k + \left( \alpha - \frac{1}{2} \right) \alpha \|d^k\|^2 + o(\|d^k\|^2). \]
Hence, for \( k \) large enough, we have 
\[
F(x^k + d^k + \tilde{d}^k) \leq F(x^k) - \alpha(d^k)^T H_k d^k \leq F_k - \alpha(d^k)^T H_k d^k.
\]
This completes the whole proof. \( \square \)

To analyze the superlinear convergence, we further give a lemma as follows.

**Lemma 4.5.** Suppose that H2\(^+\) and H3–H6 hold. Then, for \( k \) large enough, the following matrix:
\[
G_k \triangleq \begin{pmatrix}
P_k \nabla^2_{xx} L(x^*, \tilde{x}^*, \tilde{\mu}^*) & N_k \\
N_k^T & 0
\end{pmatrix}
\]

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

**Proof.** It is sufficient to show that any accumulation point \( G_\ast \) of the sequence \( \{G_k\} \) is nonsingular. For a given accumulation point \( G_\ast \), we can assume that there exists an infinite index subset \( K \) such that
\[
\text{if } \equiv i', \quad G_k \rightarrow G_\ast, \quad k \in K.
\]
We denote
\[
N_\ast = (\nabla f_i(x^*) - \nabla f_{i'}(x^*), \quad i \in I(x^*) \setminus \{i'\}; \quad \nabla g_j(x^*), \quad j \in J(x^*)),
\]
\[
P_\ast = E_n - N_\ast (N_\ast^T N_\ast)^{-1} N_\ast^T,
\]
then
\[
G_\ast = \begin{pmatrix}
P_\ast \nabla^2_{xx} L(x^*, \tilde{x}^*, \tilde{\mu}^*) & N_\ast \\
N_\ast^T & 0
\end{pmatrix}.
\]

Now we show that \( G_\ast (y^T, \bar{y}^T)^T = 0 \) has only a solution zero. From \( G_\ast (y^T, \bar{y}^T)^T = 0 \), we get
\[
N_\ast^T y = 0, \quad P_\ast \nabla^2_{xx} L(x^*, \tilde{x}^*, \tilde{\mu}^*) y + N_\ast \bar{y} = 0.
\]

From the first equation of (4.15) and the definition of \( P_\ast \), one gets \( y^T P_\ast = y^T \), thus we further get \( y^T \nabla^2_{xx} L(x^*, \tilde{x}^*, \tilde{\mu}^*) y = 0 \) from the second equation of (4.15). So, combining H6 with (4.15), we get \( y = 0 \), furthermore, \( N_\ast \bar{y} = 0 \) follows from the second equation of (4.15). In view of H2\(^+\), we further get \( \bar{y} = 0 \). Thus, the matrix \( G_\ast \) is nonsingular. \( \square \)

**Theorem 4.3.** Let the assumptions H2\(^+\) and H3–H7 be satisfied. Then the proposed algorithm is superlinearly convergent, i.e., the sequence \( \{x^k\} \) generated by Algorithm A satisfies
\[
\|x^{k+1} - x^\ast\| = o(\|x^k - x^\ast\|).
\]

**Proof.** From Lemma 4.2(i), we have
\[
\nabla f_i(x^k)^T d^k = F(x^k) + z_k - f_i(x^k), \quad i \in I(x^*), \quad i \neq i_k,
\]
\[
\nabla g_j(x^k)^T d^k = c_k z_k - g_j(x^k), \quad j \in J(x^*).
\]

Thus, from (4.9), (4.16) and (4.17), one has
\[
\frac{1}{\sigma_k} H_k d^k + N_k \tilde{u}^k = -\nabla f_{i_k}(x^k),
\]
\[
N_k^T d^k = \begin{pmatrix}
f_{i_k}(x^k) - f_i(x^k), \quad i \in I(x^*) \setminus \{i_k\} \\
c_k z_k - g_j(x^k), \quad j \in J(x^*)
\end{pmatrix}.
\]
Let us define vector-function \( h(x) \) by
\[
h(x) = \sum_{i \in I(x) \setminus \{i_k\}} \tilde{\lambda}_i^* (\nabla f_i(x) - \nabla f_i^*(x)) + \sum_{j \in J(x^*)} \tilde{\mu}_j^* \nabla g_j(x) + N(x) \tilde{u}^*,
\]
with \( \tilde{u}^* = (\tilde{\lambda}_i^*, i \in I(x) \setminus \{i_k\}, \tilde{\mu}_j^*, j \in J(x^*)) \). Then, in view of \( \sum_{i \in I(x) \setminus \{i_k\}} \tilde{\lambda}_i^* = 1 \) and \( \sum_{j \in J(x^*)} \tilde{\mu}_j^* \nabla g_j(x^*) = 0 \), we get \( h(x^k) = -\nabla f_i^*(x^*) \). Therefore, one has by Taylor expansion
\[
\begin{align*}
  h(x^k) &= N_k \tilde{u}^* = h(x^k) + \nabla h(x^k)^T (x^k - x^*) + o(\|x^k - x^*\|) \\
  &= -\nabla f_i^*(x^k) + \sum_{i \in I(x) \setminus \{i_k\}} \tilde{\lambda}_i^* (\nabla^2 f_i(x^k) - \nabla^2 f_i^*(x^k)) (x^k - x^*) \\
 &\quad + \sum_{j \in J(x^*)} \tilde{\mu}_j^* \nabla^2 g_j(x^*) (x^k - x^*) + o(\|x^k - x^*\|) \\
  &= -\nabla f_i^*(x^k) + \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*)(x^k - x^*) - \nabla^2 f_i^*(x^k) (x^k - x^*) + o(\|x^k - x^*\|).
\end{align*}
\]
So from the definition of \( P_k \) and the equalities above, we have \( P_k N_k = 0 \) and
\[
0 = P_k N_k \tilde{u}^* = P_k h(x^k) = \begin{array}{c}
- P_k \nabla f_i^*(x^k) + P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*)(x^k - x^*) \\
- P_k \nabla^2 f_i^*(x^k) (x^k - x^*) + o(\|x^k - x^*\|).
\end{array}
\]
That is
\[
P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*)(x^k - x^*) = P_k \nabla^2 f_i^*(x^k) (x^k - x^*) + P_k \nabla^2 f_i^*(x^k) + o(\|x^k - x^*\|). \tag{4.20}
\]
Furthermore, from Theorem 4.2, (4.20), Lemma 4.3 and Remark 4.3, we have
\[
P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*)(x^k - x^*) = \begin{array}{c}
P_k \nabla^2 f_i^*(x^k) (x^k - x^*) \\
- P_k \nabla f_i^*(x^k) + P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*) (d^k + \tilde{d}^k) \\
- P_k \nabla^2 f_i^*(x^k) (x^k - x^*) + P_k \nabla f_i^*(x^k) + P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*) - H_k) d^k + P_k H_k d^k \\
+ o(\|x^k - x^*\|) + o(\|d^k\|) \\
- P_k \nabla^2 f_i^*(x^k) (x^k - x^*) + P_k \nabla f_i^*(x^k) + P_k H_k d^k + o(\|x^k - x^*\|) + o(\|d^k\|).
\end{array}
\]
On the other hand, from (4.18), we obtain \( (1/\sigma_k) P_k H_k d^k = -P_k \nabla f_i^*(x^k) \). Furthermore, combining the last formula above with \( \sigma_k \rightarrow 1 \) and Taylor expansion, one has
\[
P_k \nabla^2 L(x^*, \tilde{\lambda}^*, \tilde{\mu}^*)(x^k - x^*) = \begin{array}{c}
P_k \left( \nabla^2 f_i^*(x^k) (x^k - x^*) + \nabla f_i^*(x^k) - \sigma_k \nabla f_i^*(x^k) \right) + o(\|x^k - x^*\|) + o(\|d^k\|) \\
- P_k \left( \nabla^2 f_i^*(x^k) (x^k - x^*) + \nabla f_i^*(x^k) - 1 - \sigma_k) \nabla f_i^*(x^k) \right) + o(\|x^k - x^*\|) + o(\|d^k\|) \\
- \sigma_k) \nabla f_i^*(x^k) - (1 - \sigma_k) P_k \nabla f_i^*(x^k) \\
= o(\|x^k - x^*\|) + o(\|d^k\|) - \frac{1 - \sigma_k}{\sigma_k} P_k H_k d^k \\
= o(\|x^k - x^*\|) + o(\|d^k\|),
\end{array}
\]
that is
\[ P_k \nabla^2_{xx} L(x^*, \bar{\lambda}^*, \bar{\mu}^*) (x^{k+1} - x^*) = o(\|x^k - x^*\|) + o(\|d^k\|). \] (4.21)

On the other hand, from \( I(x^k) \subseteq I(x^*) \), \( J(x^k) \subseteq J(x^*) \) and Taylor expansion, one has
\[ 0 = f_i(x^k) - f_i(x^*) = f_i(x^k) - f_i(x^k) + \nabla f_i(x^k) (x^* - x^k) + o(\|x^k - x^*\|), \quad i \in I(x^*), \] (4.22)
\[ 0 = g_j(x^*) = g_j(x^k) + \nabla g_j(x^k) (x^* - x^k) + o(\|x^k - x^*\|), \quad j \in J(x^*). \] (4.23)

Hence, from (4.22) and (4.23), we have
\[ N_k^T (x^k - x^*) = \begin{pmatrix} f_i(x^k) - f_i(x^k), & i \in I(x^*) \setminus \{i_k\} \\ g_j(x^k), & j \in J(x^*) \end{pmatrix} + o(\|x^k - x^*\|). \]

This along with Theorem 4.2, (4.19) and Lemma 4.3 implies that
\[ N_k^T (x^{k+1} - x^*) = N_k^T (x^k - x^*) + N_k^T (d^k + \tilde{d}^k) = o(\|x^k - x^*\|) + o(\|d^k\|). \] (4.24)

Therefore, from (4.21) and (4.24), we have
\[ \left( \begin{array}{cc} P_k \nabla^2_{xx} L(x^*, \bar{\lambda}^*, \bar{\mu}^*) & N_k \\ N_k^T & 0 \end{array} \right) \begin{pmatrix} x^k+1 - x^* \\ 0 \end{pmatrix} = o(\|d^k\|) + o(\|x^k - x^*\|). \]

This together with Lemmas 4.5 and 4.3 shows that
\[ \|x^{k+1} - x^*\| = o(\|d^k\|) + o(\|x^k - x^*\|) \]
\[ = o(\|d^k + \tilde{d}^k\|) + o(\|x^k - x^*\|) \]
\[ = o(\|(x^{k+1} - x^*) - (x^k - x^*)\|) + o(\|x^k - x^*\|) \]
\[ \leq o(\|x^{k+1} - x^*\|) + o(\|x^k - x^*\|). \]

Thus,
\[ \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} \left( 1 - \frac{o(\|x^{k+1} - x^*\|)}{\|x^{k+1} - x^*\|} \right) \leq o(\|x^k - x^*\|) \|x^k - x^*\|. \]

This implies that \( \|x^{k+1} - x^*\| = o(\|x^k - x^*\|) \). So the proposed algorithm is superlinearly convergent. The whole proof of Theorem 4.3 is completed. \( \square \)

5. Numerical experiments

In this section, we select several problems to show the efficiency of Algorithm A. The numerical experiments are implemented on MATLAB 6.5 and we use its optimization toolbox to solve the quadratic programings (2.1) and (2.8). All computations are performed on an Intel(R) Celeron(R) CPU 1.80GHz computer. The numerical results show that the proposed algorithm is efficient.

In the implementation, the approximation Hessian matrix \( H_k \) is updated according to the Powell’s modification of BFGS formula [14] as follows:
\[ H_{k+1} = H_k - \frac{H_k s^k (s^k)^T H_k}{(s^k)^T H_k s^k} + \frac{\tilde{y}^k (\tilde{y}^k)^T}{(s^k)^T \tilde{y}^k}, \quad (k \geq 0), \] (5.1)
Table 1
Numerical results of Problems 1–6

| Problem | n, m, m' | IP       | Method | r | Ni | Ns | F(x*) | ||d^k|| |
|---------|----------|----------|--------|---|----|----|-------|--------|
| 1       | 2, 3, 2  | (0, 0)^T | RN     | 0 | 6  |    | 1.95222 | φ      |
| (Problem 1 in [16]) | JQZ     | 0        | 5      | 3 | 1.952224 | 1.661704e − 007 |
|         |          | JQZ     | 1      | 5  | 1.952224 | 9.689462e − 007 |
| 2       | 2, 6, 2  | (1, 3)^T | RN     | 0 | 7  |    | 0.61643 | φ      |
| (Problem 2 in [16]) | JQZ     | 0        | 8      | 2 | 0.616438 | 1.426054e − 008 |
|         |          | JQZ     | 2      | 11 | 5   | 0.616432 | 6.521994e − 008 |
|         |          | JQZ     | 2      | 7  | 1   | 0.616432 | 5.568723e − 008 |
| 3       | 2, 3, 2  | (4, 2)^T | RN     | 0 | 10 |    | 2.25   | φ      |
| (Problem 4 in [16]) | (2.5, -2.5)^T | JQZ | 0 | 10 | 7 | 2.25 | 1.294468e − 006 |
|         |          | JQZ     | 2      | 12 | 9   | 2.25 | 4.866514e − 007 |
| 4       | 4, 4, 3  | (0, 1, 1, 0)^T | RN     | 0 | 11 |    | -44.0  | φ      |
| (Problem 5 in [16]) | JQZ     | 0        | 32     | 32 | 34  | 3.616107e − 005 |
|         |          | JQZ     | 1      | 31 | 31  | -44.0 | 3.145373e − 005 |
|         |          | JQZ     | 2      | 36 | 36  | -44.0 | 3.465435e − 005 |
| 5       | 2, 3, 2  | (0, 1)^T | RN     | 0 | 4  |    | 2.0    | φ      |
| (Problem 6 in [16]) | JQZ     | 0        | 4      | 1 | 2.0 | 5.251334e − 007 |
|         |          | JQZ     | 1      | 4  | 4   | 2.0  | 5.251334e − 007 |
| 6       | 7, 5, 4  | (1, 2, 0, 4, 0, 1, 1)^T | RN     | 0 | 17 |    | 680.6306 | φ      |
| (Problem 7 in [16]) | JQZ     | 0        | 41     | 41 | 680.7535 | 0.095059 |
|         |          | JQZ     | 1      | 37 | 37  | 680.6398 | 0.045855 |
|         |          | JQZ     | 2      | 28 | 28  | 680.6385 | 0.020206 |

where

\[ s^k = x^{k+1} - x^k, \quad y^k = \eta_k y^k + (1 - \eta_k) H_k s^k, \quad y^k = \nabla x L(x^{k+1}, \bar{\lambda}^k, \bar{\mu}^k) - \nabla x L(x^k, \bar{\lambda}^k, \bar{\mu}^k), \]

\[ \bar{\lambda}_i = \sum_{j \in J} \lambda_j \nabla f_j(x) + \sum_{j \in J} \mu_j \nabla g_j(x), \]

\[ \eta_k = \begin{cases} 
1 & \text{if } (s^k)^T y^k > 0.2(s^k)^T H_k s^k, \\
0.8(s^k)^T H_k s^k & \text{otherwise.}
\end{cases} \]

During the numerical experiments, we set parameters \( v = 0.75, \quad \beta = 0.5, \quad \tau = 2.5, \quad \alpha = 0.45, \quad \zeta = 0.5, \quad \bar{c} = 0.1, \quad c_0 = 2.5, \quad \varepsilon_0 = 1, \quad \theta_0 = 1, \quad M = 100 \) and

\[ \varepsilon_{k+1} = \begin{cases} 
\varepsilon_k / 2 & \text{if } \varepsilon_k > \bar{\varepsilon} \leq 0.01, \\
\varepsilon_k & \text{otherwise,}
\end{cases} \]

\[ \theta_{k+1} = \begin{cases} 
\theta_k / 2 & \text{if } \theta_k > \bar{\theta} \leq 0.15, \\
\theta_k & \text{otherwise.}
\end{cases} \]

In formula (5.1), we select \( H_0 = E \), where \( E \in \mathbb{R}^{n \times n} \) is an identity matrix. Execution is terminated if \( ||d^k|| < 10^{-5} \) or \( ||x^{k+1} - x^k|| < 10^{-5} \).

The test problems 1–6 are selected from problems 1, 2, 4–7 in [16], respectively. The following Table 1 gives the numerical results of Algorithm 1 (\( r = 0, \gamma = 0.1 \)) in Ref. [16] and the proposed algorithm in this paper corresponding to parameter \( r = 0, 1, 2 \), respectively.
Table 2
The approximately optimal solutions of the Problems about JQZ.

<table>
<thead>
<tr>
<th>Problem</th>
<th>r</th>
<th>Approximately optimal solution $x^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0, 1, 2</td>
<td>(1.13904, 0.89956)$^T$</td>
</tr>
<tr>
<td>2</td>
<td>0, 1, 2</td>
<td>(−0.45330, 0.90659)$^T$</td>
</tr>
<tr>
<td>3</td>
<td>0, 1, 2</td>
<td>(1.35356, 0.64644)$^T$</td>
</tr>
<tr>
<td>4</td>
<td>0, 1, 2</td>
<td>(0.0, 1.0, 2.0, −1.0)$^T$</td>
</tr>
<tr>
<td>5</td>
<td>0, 1, 2</td>
<td>(1.0, 1.0)$^T$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>(2.27088, 1.96372, −0.48377, 4.34848, −0.61966, 1.04854, 1.52877)$^T$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>(2.32279, 1.94861, −0.52346, 4.37617, −0.62558, 1.03135, 1.59134)$^T$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>(2.34269, 1.95234, −0.47569, 4.35944, −0.622033, 1.02614, 1.60240)$^T$</td>
</tr>
</tbody>
</table>

In Table 1, the following notations mean:

- JQZ: the proposed Algorithm A in this paper;
- RN: the Algorithm 1 ($\gamma = 0.1, r \equiv 0$) in Ref. [16];
- NI: the number of iterations;
- IP: the initial point;
- NS: Number of solutions to QP$(x^k, d^k, H_k)$ in this paper;
- $n$: the dimension of the problem;
- $m$: the number of objective functions;
- $m^i$: the number of constraints;
- $\phi$: the norm of $d^k$ is not given in [16].

From the numerical results for the six problems, we can see that our algorithm is efficient. Firstly, from Table 1, we can see that problem 1 in the cases of $r = 0, 1, 2$ is solved a little faster than that in [16] if initial point $(0, 0)^T$ is chosen. As to problem 2, it is solved a little faster than JQZ if initial point $(1, 3)^T$ is chosen, but we get the same number of iterations as that in [16] if we choose $(1, 2.4)^T$ as initial point. Consider problem 3, because point $(4, 2)^T$ is not a feasible point, JQZ starts with initial feasible point $(2.5, −2.5)^T$ and the number of iterations in the cases of parameter $r = 0, 1$ is as same as that in [16]. To problem 5, RN and JQZ have same number of iterations. On the other hand, considering problems 4 and 6, we see that they are solved faster in [16] than JQZ. But from Table 1, it is obviously that the two algorithms do not have much difference in terms of the number of iterations except problems 4 and 6.

Secondly, we only consider the numerical results of Algorithm A in this paper. From Table 1, it can be seen that the norm of feasible direction of descent $d^k$ converges to zero fast except problem 6. Furthermore, for problem 2 with the same initial point $(1, 3)^T$, the efficiency of Algorithm A in the case of $r = 0$ is better than the other two cases, i.e., the case of monotone line search is better than the other two “nonmonotone” ones. However, the results for problems 4 and 6 show that the “nonmonotone” line search is better than the monotone one. But, to problem 1, problem 2 with initial point $(1, 2.4)^T$ and problem 5, each number of iterations is same for $r = 0, 1, 2$. On the other hand, we can see that the optimal solutions (see Table 2) of each problem are almost same for parameter $r = 0, 1, 2$ except problem 6.

Summarizing, in our opinion, one cannot simply conclude that which case for the parameter $r$ in Algorithm A is better than the others, and it changes with different problems and different initial points.

6. Concluding remarks

In this paper, we propose a feasible “generalized monotone line search” algorithm for nonlinear minimax problems with inequality constraints. With the help of the technique of norm-relaxed MFD, we construct a new QP and by solving this QP subproblem we obtain a feasible direction of descent, then a correction direction is yielded by solving another QP to avoid the Maratos effect and guarantee the superlinear convergence under mild conditions. Then combining the “nonmonotone line search” technique, we propose our generalized monotone line search algorithm and the preliminary numerical results show that the proposed algorithm is efficient.
To further reduce the number of the constraints of QP$(x^k, d^k, H_k)$, we can consider of replacing the index sets $I^k_{\ell_k}$ and $J^k_{\ell_k}$ in QP$(x^k, d^k, H_k)$ subproblem with the active sets $I_k$ and $J_k$ of QP$(x^k, H_k)$, respectively, which does not affect the convergence analysis in this paper. As further work, we can consider of removing the strict complementarity, and we can also obtain the correction direction $\tilde{d}^k$ by other techniques, for example: sequential systems of linear equations technique [3] or generalized projection technique [8,7].

References