J. Appl. Math. & Informatics Vol. **28***(2010), No. 3 - 4, pp. 823 - 835*
Website: http://www.kcam.biz

A GLOBALLY AND SUPERLIEARLY CONVERGENT FEASIBLE SQP ALGORITHM FOR DEGENERATE CONSTRAINED OPTIMIZATION

YU CHEN* AND XIAO-LIANG XIE

ABSTRACT. In this paper, A FSQP algorithm for degenerate inequality constraints optimization problems is proposed. At each iteration of the proposed algorithm, a feasible direction of descent is obtained by solving a quadratic programming subproblem. To overcome the Maratos effect, a higher-order correction direction is obtained by solving another quadratic programming subproblem. The algorithm is proved to be globally convergent and superlinearly convergent under some mild conditions. Finally, some preliminary numerical results are reported.

AMS Mathematics Subject Classification : 65K05,90C30
Key word and phrases : Degenerate inequality constrained optimization, quadratic programming, method of feasible directions, global convergence, convergence rate.

1. Introduction

In this paper, we consider the following nonlinear inequality constrained optimization:

$$(P) \qquad \begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_j(x) \le 0, \quad j \in I = \{1, 2, ..., m\}, \end{array} \tag{1.1}$$

where the functions f_0, $f_j(j \in I) : R^n \to \mathrm{R}$ are all continuously differentiable.

It is well known that sequential quadratic programming (SQP) algorithms are widely acknowledged to be among the most successful algorithms for solving (P)(See[5]–[10], [12], [14]–[16] and [17]). A good survey of SQP algorithms by Boggs and Toll can be found in [4]. Many existing SQP algorithms for handling

Received August 29, 2009. Revised September 8, 2009. Accepted October 7, 2009
*Corresponding author. Project supported by Scientific Research Fund of Hunan Provincial Science technology Department(2008FJ3142) and Scientific Research Fund of Hunan Provincial Education Department(08C470)
© 2010 Korean SIGCAM and KSCAM.

constrained optimization problems focus on using penalty functions, while a feasible sequence of iterates is very important for many practical problems, such as engineering design, real-time applications and that problems whose objective functions are not well defined outside the feasible set. To overcome this shortcoming, In [12], variations on the standard SQP algorithms for solving (P) are proposed which generate iterations lying within the feasible set of (P), which is called as feasible sequential quadratic programming (FSQP) algorithm. FSQP is proved to be globally convergent and superlinearly convergent under some mild assumptions including the strict complementary condition. However, at each iteration, these algorithms require to solve two QP subproblems and a linear least squares problem, or two linear systems of equations and a linear least squares problem. Clearly, their computational cost per single iteration is relatively high. Recently, Another type of FSQP algorithm in [13] is proposed. In this algorithm, the following QP subproblem is considered, for an iteration point x^k:

$$\text{(QP)} \qquad \begin{array}{ll} \min\limits_{(z,d)} & z + \frac{1}{2} d^T H_k d \\ \text{s.t.} & \nabla f_0(x^k)^T d \leq z, \\ & f_j(x^k) + \nabla f_j(x^k)^T d \leq \sigma_k z, \quad j \in I, \end{array} \tag{1.2}$$

where H_k is a symmetric positive definite matrix and an approximation of the Lagrangian Hessian matrix for (P), and σ_k is a positive parameter. In [13], it is necessary to solve an equality constrained QP subproblem to update the parameter σ_k such that $\sigma_k = O(\|d_0^{k-1}\|^2)$. On the other hand, in order to accept the unit step size, a correction direction is obtained by solving another equality constrained QP subproblem. Furthermore, the algorithm is proved to be locally two-step superlinearly convergent under certain conditions including the strict complementary condition. Ref [11] proposed a similar algorithm to solve the problem (P) too, it needs to solve two QP subproblems with inequality constraints, and like [13], it is proved to be locally two-step superlinearly convergent under certain conditions including the strict complementary condition, furthermore, it is required that σ_k approaches to zero fast enough as $d^k \to 0$, i.e., $\sigma_k = o(\|d^k\|)$. In [18], Zhu proposed a algorithm, in his algorithm, a feasible direction of descent is obtained by solving the QP subproblem (1.2), in order to avoid the Marotos effect, a high-order revised direction is computed by solving a reduced linear system. Furthermore, it is proved to be globally convergent and superlinearly convergent under some certain conditions including the strict complementary condition. Different from [11], [13], it needn't compute any auxiliary problem to update σ_k. In [7, 8], Wright presented an infeasible SQP algorithm for degenerate inequality constraints optimization problems(i.e.,The strict complementary condition or linear independence condition at the solution is not assumed) by modifying the QP subproblem, the proposed algorithm is proved to be local superlinear convergent under some weaker conditions. However, little attention has been given to the FSQP algorithm for degenerate inequality constraints optimization problems.

In this paper, we have proposed a FSQP algorithm for degenerate inequality constraints optimization problems. At each iteration of the proposed algorithm, a feasible direction of descent is obtained by solving a quadratic programming subproblem. To overcome the Maratos effect, a higher-order correction direction is obtained by solving another quadratic programming problem. The algorithm is proved to be globally convergent and superlinearly convergent under some mild conditions.

The remainder of this paper is organized as follows. The proposed algorithm is stated in Section 2. In Section 3 and Section 4, under some mild assumptions, we show that this algorithm is globally convergent and locally superlinear convergent, respectively. In section 5, some preliminary numerical results are reported. Finally, we give concluding remarks about the proposed algorithm.

2. Description of algorithm

We denote the feasible set X of (P) by

$$X = \{x \in R^n : f_i(x) \le 0, i \in I\},$$

and define the active set by

$$I(x) = \{i \in I : f_i(x) = 0\}.$$

In this paper, we suppose the feasible set X is not empty and the following basic hypothesis holds.

Assumption A_1. *We assume that the Mangasarian-Fromovitz constraint qualification (**MFCQ**) holds at $x \in X$. That is, there is some $d \in R^n$ such that $\nabla f_j(x)^T d < 0, \forall\, j \in I(x)$.*

Remark 1. This constraint qualification is weaker than Linear Independent constraint qualification which is a common assumption for global convergence analysis of many kinds of SQP methods.

The following algorithm is proposed for solving (P).

ALGORITHM

Parameters $\tau \in (2,3), \beta \in (0,1),\ \alpha \in (0,\frac{1}{2}),\ \nu \in (0,1),\ \sigma_1 > 0$.

Data Choose an initial feasible point $x^1 \in X$, a symmetric positive matrix H_1. Set $k = 1$.

Step 1 (Compute the search direction). For the current iteration point x^k, solve

$$(QP) \quad \begin{array}{ll} \min & z + \frac{1}{2} d^T H_k d \\ \text{s.t.} & \nabla f_0(x^k)^T d \le z, \\ & f_j(x^k) + \nabla f_j(x^k)^T d \le \sigma_k z, \quad j \in I, \end{array} \tag{2.1}$$

to obtain an optimal solution (z_k, d^k), let (u_0^k, u_I^k) be corresponding KKT multipliers. If $d^k = 0$, then x^k is a KKT point for (P) and stop; otherwise go to Step 3.

Step 2. Compute the higher-order direction $\widetilde{d^k}$ by solving the following quadratic programming subproblem:

$$(\widetilde{QP}) \quad \begin{array}{ll} \min & \frac{1}{2}d^T H_k d + \nabla f_0(x^k)^T(d - d^k) \\ \text{s.t.} & f_j(x^k + d^k) + \nabla f_j(x^k)^T(d - d^k) \leq -\|d^k\|^\tau \quad j \in I. \end{array} \tag{2.2}$$

If there exists no solution or $\|\widetilde{d^k} - d^k\| > \|d^k\|$, set $\widetilde{d^k} = d^k$.

Step 3 (Do curve search). Compute the step size λ_k, which is the first number λ of the sequence $\{1, \beta, \beta^2, ...\}$ satisfying

$$f_0(x^k + \lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) \leq f_0(x^k) + \alpha\lambda\nabla f_0(x^k)^T d^k, \tag{2.3}$$

$$f_j(x^k + \lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) \leq 0, \quad \forall j \in I. \tag{2.4}$$

Step 4. Set a new iteration point by $x^{k+1} = x^k + \lambda_k d^k + \lambda_k^2(\widetilde{d}^k - d^k)$, $\sigma_{k+1} = min\{\sigma_1, \|d^k\|^\nu\}$. Compute a new symmetric positive definite matrix H_{k+1}, set $k := k + 1$, and go back to Step 1.

We now show that the proposed algorithm is well defined.

Lemma 2.1. *Suppose that H_k is symmetric positive definite. Then (QP) always has a unique optimal solution.*

The proof is similar to the one of Lemma 1 in [18].

Lemma 2.2. *Suppose that the conditions in the above lemma are satisfied and (z_k, d^k) is an optimal solution of (2.1). Then*

(i) $z_k + \frac{1}{2}(d^k)^T H_k d^k \leq 0$ and $z_k \leq 0$;

(ii) $z_k = 0 \Longleftrightarrow d^k = 0 \Longleftrightarrow x^k$ is a KKT point for (P);

(iii) $z_k < 0 \Longrightarrow d^k$ is a feasible direction of descent for (P) at point x^k.

The proof is similar to the one of Lemma 2 in [18].

Lemma 2.3. *The line search in Step 3 of the proposed algorithm yields a stepsize $\lambda_k = \beta^j$ for some finite $j = j(k)$.*

It is not difficult to finish the proof of this lemma.

3. Global Convergence

In this section, we analyze the global convergence of the proposed algorithm. The following assumptions are necessary.

Assumption A_2. *The sequence $\{x^k\}$, which is generated by the proposed algorithm, is bounded.*

Assumption A_3. *There exist a, $b > 0$ such that $a\|d\|^2 \leq d^T H_k d \leq b\|d\|^2$ for all k and all $d \in R^n$.*

We suppose that x^* is a given accumulation point of $\{x^k\}$. In view of J_k being a subset of the finite and fixed set I, there exist an infinite index set K such that

$$\lim_{k \in K} x^k = x^*, \quad J_k \equiv J, \quad \forall k \in K, \tag{3.1}$$

where

$$J_k = \{j \in I : f_j(x^k) + \nabla f_j(x^k)^T d^k = \sigma_k z_k\}.$$

Lemma 3.1. *Suppose that Assumptions A_2 and A_3 hold. Then the sequences $\{d^k : k \in K\}, \{z_k : k \in K\}$ and $\{\widetilde{d}^k : k \in K\}$ are all bounded.*

Proof. Firstly, from $\nabla f_0(x^k) \to \nabla f_0(x^*)$, $k \in K$, there exists a constant $c_0 > 0$ such that $\| \nabla f_0(x^k) \| \leq c_0$, $\forall k \in K$. Furthermore, from Lemma 2.2, formulas (2.1) and Assumption A_3, one has

$$\begin{aligned} 0 \geq z_k + \tfrac{1}{2}(d^k)^T H_k d^k &\geq \nabla f_0(x^k)^T d^k + \tfrac{a}{2}\|d^k\|^2 \\ &\geq -\|\nabla f_0(x^k)\| \cdot \|d^k\| + \tfrac{a}{2}\|d^k\|^2 \\ &\geq -c_0\|d^k\| + \tfrac{a}{2}\|d^k\|^2, \qquad k \in K. \end{aligned}$$

This shows that $\{d^k : k \in K\}$ is bounded.

Secondly, the boundedness of $\{z_k : k \in K\}$ follows from the boundedness of $\{d^k : k \in K\}$ as well as the following inequalities

$$0 \geq z_k \geq \nabla f_0(x^k)^T d^k \geq -\|\nabla f_0(x^k)\| \cdot \|d^k\| \geq -c_0\|d^k\|, \quad k \in K. \tag{3.2}$$

Lastly, the boundedness of $\{\widetilde{d}^k : k \in K\}$ follows immediately from the boundedness of $\{d^k : k \in K\}$. □

We know that the KKT conditions of (QP) can be formulated as follows:

$$H_k d^k + u_0^k \nabla f(x^k) + \sum_{j \in I} u_j^k \nabla f_j(x^k) = 0, \tag{3.3}$$

$$1 = u_0^k + \sum_{j \in I} \sigma_k u_j^k, \tag{3.4}$$

$$0 \leq u_0^k \perp (z_k - \nabla f_0(x^k)^T d^k) \geq 0, \tag{3.5}$$

$$0 \leq u_j^k \perp (\sigma_k z_k - f_j(x^k) - \nabla f_j(x^k)^T d^k) \geq 0, \quad j \in I, \tag{3.6}$$

where the notation $x \perp y$ means $x^T y = 0$.

Lemma 3.2. (*i*) *The multiplier sequence $\{u_0^k\}_{k=0}^{\infty}$ is bounded.*

(*ii*) *Let multiplier vector $u^k = (u_{J_k}^k, 0_{I \setminus J_k})$. If $\lim_{k \in K} x^k = x^*$ and $\lim_{k \in K} d^k = 0$, then $\{u^k : k \in K\}$ is bounded under Assumptions A_1, A_2 and A_3.*

Proof. (i) From the KKT condition (3.4), one has

$$1 = u_0^k + \sum_{j \in J_k} \sigma_k u_j^k \geq u_0^k, \quad 0 \leq u_0^k \leq 1.$$

(ii) Suppose by contradiction that the given statement is not true, then there exists an infinite index $K' \subseteq K$ such that $\| u^k \| = \| u_J^k \| \to \infty$, $k \in K'$. Therefore, dividing (3.3) by $\|u_J^k\|$ to yield

$$\frac{1}{\|u_J^k\|} H_k d^k + \frac{u_0^k}{\|u_J^k\|} \nabla f_0(x^k) + \sum_{j \in J} \frac{u_j^k}{\|u_J^k\|} \nabla f_j(x^k) = 0. \tag{3.7}$$

Noting that the sequence $\{\frac{u_j^k}{\|u_j^k\|} : k \in K^{'}\}$ is bounded with norm one, we can assume without loss of generality that

$$\frac{u_j^k}{\|u_j^k\|} \to \overline{u}_j, \quad k \in K^{'}, \quad j \in J, \quad 0 \le (\overline{u}_j,\ j \in J) \ne 0. \tag{3.8}$$

Thus, passing to the limit $k \in K^{'}$ and $k \to \infty$ in (3.7), and taking into account Assumption A_3 as well as the given conditions, we have

$$\sum_{j\in J} \overline{u}_j \nabla f_j(x^*) = 0. \tag{3.9}$$

On the other hand, from the given conditions, one has $J \subseteq I(x^*)$, so we can bring a contradiction from (3.8), (3.9) and Assumption A_1, therefore the boundedness of $\{u^k : k \in K\}$ is at hand. □

Lemma 3.3. *Suppose that $\{x^k\}$ is a sequence generated by the proposed algorithm, $\lim_{k\in K} x^k = x^*$ and $\lim_{k\in K} d^k = 0$ hold. Then x^* is a KKT point of (P).*

Proof. Taking into account the boundedness of $\{u_0^k : k \in K\}$, $\{u^k : k \in K\}$ and $\{\sigma_k\}$, we can assume without loss of generality that

$$u^k = (u_j^k, j \in I) \to u^* = (u_j^*, j \in I), \quad u_0^k \to u_0^*, \quad \sigma_k \to \sigma_* \quad k \in K.$$

Moreover, the fact $\lim_{k\in K} x^k = x^*$ and $\lim_{k\in K} d^k = 0$ implies $\lim_{k\in K} z^k = 0$. Thus, passing to the limit $k \in K$ and $k \to \infty$ in (3.3)—(3.6) and the given conditions, we obtain

$$\begin{aligned} &u_0^* \nabla f_0(x^*) + \sum_{j\in J} u_j^* \nabla f_j(x^*) = 0, \\ &u_j^* f_j(x^*) = 0, \quad u_j^* \ge 0, \quad f_j(x^*) \le 0, \quad j \in J, \\ &1 = u_0^* + \sigma_* \sum_{j\in J} u_j^*, \quad u_0^* \ge 0. \end{aligned} \tag{3.10}$$

From the third formula of (3.10), we know that $(u_0^*, u_J^*) \ne 0$, furthermore, $u_0^* > 0$ from Assumption A_1, which together with (3.10) shows that $(x^*, \dfrac{u^*}{u_0^*})$ is a KKT pair of (P). The proof is complete. □

Based on Lemma 3.1, Lemma 3.2 and Lemma 3.3, we now can present the global convergence theorem of the proposed algorithm as follows.

Theorem 3.1. *Suppose that Assumptions A_1, A_2 and A_3 hold, then the proposed algorithm either stops at a KKT point x^k for problem (P) in a finite number of steps or generates an infinite sequence $\{x^k\}$ of points such that each accumulation point x^* is a KKT point for problem (P). Furthermore, there exists an infinite index K such that the sequence $\{\frac{u^k}{u_0^k} : k \in K\}$ converges to a KKT multiplier associated with x^* and $\lim_{k\in K} u_0^k > 0$.*

Proof. The first statement is obvious. Thus, assume that the proposed algorithm generated an infinite sequence $\{x^k\}$ and (3.1)holds. The cases $\sigma_* = 0$ and $\sigma_* > 0$ are considered, separately.

A. $\sigma_* = 0$. From step 4, there exists an infinite index set $K_1 \subseteq K$ such that $\lim\limits_{k\in K_1} d^{k-1} = 0$. By step 4, it holds that

$$\| x^k - x^{k-1} \| \leq t_k \| d^{k-1} \| + t_k^2 \| \widetilde{d}^{k-1} - d^{k-1} \| \leq 2t^k \| d^{k-1} \| \to 0, \quad k \in K_1.$$

So, the fact that $\lim\limits_{k\in K_1} x^k = x^*$ implies that $\lim\limits_{k\in K_1} x^{k-1} = x^*$. Moreover, we know that x^* is a KKT point for problem (P) from Lemma 3.3.

B. $\sigma_* > 0$. Obviously, it is sufficient to show $\lim\limits_{k\in K} d^k = 0$. For this, we suppose by contradiction that $\lim\limits_{k\in K} d^k \neq 0$, then there exist an infinite subset $K' \subseteq K$ and a constant $\Delta > 0$ such that $\| d^k \| \geq \Delta$ holds for all $k \in K'$. The remainder proof is divided into two steps as follows, and we always assume that $k \in K'$ is sufficient large and $\lambda > 0$ is sufficient small.

a. Show that there exists a constant $\overline{\lambda} > 0$ such that the step size $\lambda_k \geq \overline{\lambda}$ for $k \in K'$.

$$\begin{aligned}
&f_0(x^k + \lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) - f_0(x^k) - \alpha\lambda\nabla f_0(x^k)^T d^k \\
=&\nabla f_0(x^k)^T(\lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) - \alpha\lambda\nabla f_0(x^k)^T d^k + o(\lambda) \\
=&\lambda(1-\alpha)\nabla f_0(x_k)^T d^k + o(\lambda) \\
\leq&\lambda(1-\alpha)z_k + o(\lambda) \\
\leq&-\frac{1}{2}\lambda(1-\alpha)(d^k)^T H_k d^k + o(\lambda) \\
\leq&-\frac{1}{2}a\lambda(1-\alpha)\|d^k\|^2 + o(\lambda) \\
\leq&-\frac{1}{2}a\lambda(1-\alpha)\Delta^2 + o(\lambda).
\end{aligned}$$

The last inequality above shows that (2.3) holds for $k \in K'$ and $\lambda > 0$ small enough.

Analyze (2.4): if $j \notin I(x^*)$, i.e., $f_j(x^*) < 0$. from the continuity of $f_j(x)$ and the boundedness of $\{d^k : k \in K\}$ and $\{\widetilde{d}^k : k \in K\}$, we know $f_j(x^k + \lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) \leq 0$ holds for $k \in K'$ large enough and $\lambda > 0$ small enough.

Let $j \in I(x^*)$, i.e., $f_j(x^*) = 0$. By using Taylor expansion and (2.1), we have

$$\begin{aligned}
f_j(x^k + \lambda d^k + \lambda^2(\widetilde{d}^k - d^k)) &= f_j(x^k) + \lambda\nabla f_j(x^k)^T d^k + o(\lambda) \\
&\leq f_j(x^k) + \lambda(\sigma_k z_k - f_j(x^k)) + o(\lambda) \\
&= (1-\lambda)f_j(x^k) + \lambda\sigma_k z_k + o(\lambda) \\
&\leq \lambda\sigma_k z_k + o(\lambda).
\end{aligned}$$

Therefore, from (2.1) and Assumption A_3, we have

$$\begin{array}{rcl} f_j(x^k+\lambda d^k+\lambda^2(\widetilde{d}^k-d^k)) & \leq & -\frac{\lambda}{2}\sigma_k(d^k)^T H_k d^k+o(\lambda) \\ & \leq & -\frac{\lambda}{2}\sigma_k a \parallel d^k \parallel^2 +o(\lambda) \\ & \leq & -\frac{\lambda}{2}\sigma_* a\Delta^2+o(\lambda). \end{array}$$

Thus, from the inequality above, we can conclude the search inequality (2.4) holds for $k\in K^{'}$ large enough and $\lambda>0$ small enough.

Summarizing the analysis above, we conclude that there exists $\overline{\lambda}>0$ such that $\lambda_k\geq\overline{\lambda}$ for all $k\in K^{'}$.

b. Use $\lambda_k\geq\overline{\lambda}>0$ to bring a contradiction. From (2.3), (2.1) and Assumption A_3, we have

$$\begin{array}{rcl} f_0(x^{k+1}) & \leq & f_0(x^k)+\alpha\lambda_k\nabla f_0(x^k)^T d^k \leq f_0(x^k)+\alpha\lambda_k z_k \\ & \leq & f_0(x^k)-\frac{1}{2}\alpha\lambda_k(d^k)^T H_k d^k \leq f_0(x^k)-\frac{1}{2}\alpha\lambda_k a \parallel d^k \parallel^2, \qquad \forall k. \end{array}$$

Therefore the sequence $\{f_0(x^k)\}$ is decreasing, furthermore combining $\lim\limits_{k\in K'} f_0(x^k) = f_0(x^*)$, one knows $\lim\limits_{k\to\infty} f_0(x^k)=f_0(x^*)$. On the other hand, one also has

$$f_0(x^{k+1})\leq f_0(x^k)-\frac{1}{2}a\alpha\overline{\lambda}\Delta^2, \qquad \forall k\in K^{'}.$$

Passing to the limit $k\in K^{'}$ and $k\to\infty$ in the inequality above, we have $-\frac{1}{2}a\alpha\overline{\lambda}\Delta^2\geq 0$, which is a contradiction. So, $d^*=0$. According to Lemma 3.3, x^* is a KKT point for problem (P). □

4. Rate of convergence

In this section, we will analyze the convergent rate of the proposed algorithm, for this, the following further hypothesis is necessary.

Assumption A_4 *(i) The functions $f_j(x)(j\in I)$ are all second-order continuously differentiable.*

(ii) The sequence $\{x^k\}$ generated by the algorithm possesses an accumulation point x^* such that KKT pair (x^*,u^*) satisfies the strong second-order sufficiency conditions, i.e.,

$$d^T\nabla^2_{xx}L(x^*,u^*)d>0, \quad \forall d\in\Omega \stackrel{def}{=} \{d\in R^n : d\neq 0, \nabla f_{I^+}(x^*)^T d=0\}, \qquad (4.1)$$

where

$$L(x,u)=f_0(x)+\sum_{j\in I}u_j f_j(x), \qquad I^+=\{j\in I: u_j^*>0\}. \qquad (4.2)$$

Lemma 4.1. *(i) Suppose that Assumptions A_1, A_2 hold. Then* $\lim\limits_{k\to\infty} d^k = 0$, $\lim\limits_{k\to\infty}\widetilde{d}^k=0$, $\lim\limits_{k\to\infty} z_k=0$, $\lim\limits_{k\to\infty}\sigma_k=0$ *and* $\lim\limits_{k\to\infty}\|x^{k+1}-x^k\|=0$.

(ii) If Assumptions A_1,A_2 and A_3 are satisfied, then $\lim\limits_{k\to\infty} x^k=x^*$.

Proof. (i) Similar to the proof of Lemma 4.2 in [18]. We have that $\lim_{k\to\infty} d^k = 0$, $\lim_{k\to\infty} z^k = 0$. Furthermore, it is easy to conclude that $\lim_{k\to\infty} \widetilde{d}^k = 0$ from Step 2 and $\lim_{k\to\infty} \sigma_k = 0$.

On the other hand, from the conclusion above, one has

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = \lim_{k\to\infty} \|\lambda_k d^k + \lambda_k^2(\widetilde{d}^k - d^k)\| \le \lim_{k\to\infty} (2\|d^k\| + \|\widetilde{d}^k\|) = 0.$$

(ii) Under Assumption A_4 (ii), one can conclude that the given limit point x^* is an isolated KKT point of (1.1)(See Theorem 1.2.5 in [16]), therefore x^* is an isolated accumulation point of $\{x^k\}$ from Theorem 3.1, and this together with $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$ shows that $\lim_{k\to\infty} x^k = x^*$. The proof is finished. □

Lemma 4.2. *Suppose that Assumptions A_1, A_2 and A_3 hold. Then*

$$|z_k| = O(\|d^k\|), \quad \|\widetilde{d}^k - d^k\| = O(\|d^k\|^2), \tag{4.3}$$

$$I^+ \subseteq J_k \subseteq I(x^*). \tag{4.4}$$

Proof. Firstly, from the first inequality constraint of (2.1), we have

$$-\|\nabla f_0(x^k)\| \cdot \|d^k\| \le z_k, \quad |z_k| \le \|\nabla f_0(x^k)\| \cdot \|d^k\|.$$

So it is not difficult to verify that $|z_k| = O(\|d^k\|)$.

Secondly, we shall show the second equation of (4.3). In view of Assumption A_1 and $d^k \to 0$, and the constraints of $(\widetilde{QP})$ is consistent for large enough k , furthermore, these can be denoted as follows

$$f_j(x^k) + \nabla f_j(x^k)^T \widetilde{d}^k \le -M \parallel d^k \parallel^2, \quad j \in I,$$

for some $M > 0$. Similar to the proof in [1], we obtain $\|\widetilde{d}^k - d^k\| = O(\|d^k\|^2)$.

To show the relationship (4.4), one first gets $J_k \subseteq I(x^*)$ from $\lim_{k\to\infty} (x^k, d^k, z_k, \sigma_k) = (0,0,0,0)$. Furthermore, one has $\lim_{k\to\infty} \lambda_{I^+}^k = \lambda_{I^+}^* > 0$ from Theorem 3.1, so $\lambda_{I^+}^k > 0$ and $I^+ \subseteq J_k$ holds for k large enough. The proof is complete. □

Lemma 4.3. *Suppose that Assumptions $A_1, A_2$and A_3 hold. Then $\{\widetilde{u}^k\}$ converges to the KKT multiplier associated with x^* for (P), where $\{\widetilde{u}^k\}$ is corresponding KKT multipliers for (2.2).*

The proof of the lemma is obviously, and is omitted. To ensure the step size $\lambda_k \equiv 1$ for k large enough, an additional assumption as follows is necessary.

Assumption A_5. *Suppose that $\|(\nabla^2_{xx} L(x^k, \widetilde{u}^k_{\widetilde{J}_k}) - H_k)d^k\| = o(\|d^k\|)$, where*

$$L(x, \widetilde{u}^k_{\widetilde{J}_k}) = f_0(x) + \sum_{j\in\widetilde{J}_k} \widetilde{u}^k_j f_j(x),$$

$$\widetilde{J}_k = \{j \in I : f_j(x^k + d^k) + \nabla f_j(x^k)^T(\widetilde{d}^k - d^k) = -\|d^k\|^\tau\}.$$

Remark 2. This assumption is similar to the well-known Dennis-More Assumption [2] that guarantees superlinear convergence for quasi-Newton methods.

Lemma 4.4. *Suppose that Assumptions A_1, A_2, A_3, A_4 and A_5 hold. Then the step size of the proposed algorithm always equals one, i.e., $\lambda_k \equiv 1$, if k is sufficiently large.*

Proof. We know that it is sufficient to verify (2.3) and (2.4) hold for $\lambda = 1$, and the statement "k large enough" will be omitted in the following discussion.

We first prove (2.4) holds for $\lambda = 1$. For $j \notin I(x^*)$, i.e., $f_j(x^*) < 0$, in view of $(x^k, \widetilde{d}^k) \to (x^*, 0)(k \to \infty)$, we can conclude $f_j(x^k + \widetilde{d}^k) \leq 0$ holds.

For $j \in I(x^*)$, from Taylor expansion, (2.1), (2.2) and formula (4.3), we have

$$\begin{aligned} f_j(x^k + \widetilde{d}^k) &= f_j(x^k + d^k) + \nabla f_j(x^k + d^k)^T(\widetilde{d}^k - d^k) + O(\|(\widetilde{d}^k - d^k)\|^2) \\ &= f_j(x^k + d^k) + \nabla f_j(x^k)^T(\widetilde{d}^k - d^k) + O(\|d^k\|\|(\widetilde{d}^k - d^k)\|) \\ &\quad + O(\|(\widetilde{d}^k - d^k)\|^2) \\ &\leq -\|d^k\|^\tau + O(\|d^k\|^3) \\ &\leq 0. \end{aligned} \tag{4.5}$$

This shows that (2.4) holds for $\lambda = 1$.

The next objective is to show (2.3) holds for $\lambda = 1$. From Taylor expansion and taking into account relationship (4.3), we have

$$\begin{aligned} \omega_k &\stackrel{def}{=} f_0(x^k + \widetilde{d}^k) - f_0(x^k) - \alpha \nabla f_0(x^k)^T d^k \\ &= \nabla f_0(x^k)^T d^k + \nabla f_0(x^k)^T(\widetilde{d}^k - d^k) + \tfrac{1}{2}(d^k)^T \nabla^2_{xx} f_0(x^k) d^k \\ &\quad -\alpha \nabla f_0(x^k)^T d^k + o(\|d^k\|^2). \end{aligned} \tag{4.6}$$

On the other hand, from the KKT condition of (2.2) and formula (4.3), one has

$$\begin{aligned} \nabla f_0(x^k)^T(\widetilde{d}^k - d^k) &= -(\widetilde{d}^k)^T H_k(\widetilde{d}^k - d^k) - \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k \nabla f_j(x^k)^T(\widetilde{d}^k - d^k) \\ &= \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k f_j(x^k + d^k) + o(\|d^k\|^2). \\ \nabla f_0(x^k)^T d^k &= -(\widetilde{d}^k)^T H_k d^k - \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k \nabla f_j(x^k)^T d^k \\ &= -(\widetilde{d}^k)^T H_k d^k + \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k f_j(x^k) - \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k \sigma_k z_k. \end{aligned} \tag{4.7}$$

Again, from Taylor expansion, we have

$$f_j(x^k + d^k) = f_j(x^k) + \nabla f_j(x^k)^T d^k + \frac{1}{2}(d^k)^T \nabla^2_{xx} f_j(x^k) d^k + o(\|d^k\|^2), \quad j \in \widetilde{J}_k.$$

Thus

$$\sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k f_j(x^k + d^k) = \sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k \sigma_k z_k + \tfrac{1}{2}(d^k)^T (\sum_{j \in \widetilde{J}_k} \widetilde{u}_j^k \nabla^2_{xx} f_j(x^k)) d^k + o(\|d^k\|^2). \tag{4.8}$$

Substituting (4.8) into (4.7), one has

$$\nabla f_0(x^k)^T(\widetilde{d}^k - d^k) = \sum_{j\in\widetilde{J}_k} \widetilde{u}_j^k \sigma_k z_k + \tfrac{1}{2}(d^k)^T(\sum_{j\in\widetilde{J}_k} \widetilde{u}_j^k \nabla^2_{xx} f_j(x^k))d^k + o(\|d^k\|^2). \tag{4.9}$$

Substituting (4.9) and the third equation of (4.7) into (4.6), we obtain

$$\begin{aligned}\omega_k &= (\alpha-1)(\widetilde{d}^k)^T H_k d^k + \tfrac{1}{2}(d^k)^T\nabla^2_{xx}L(x^k,\widetilde{u}^k)d^k \\ &+ (1-\alpha)\sum_{j\in\widetilde{J}_k}\widetilde{u}_j^k f_j(x^k) + \alpha\sum_{j\in\widetilde{J}_k}\widetilde{u}_j^k\sigma_k z_k + o(\|d^k\|^2) \\ &\le ((\alpha-1)+\tfrac{1}{2})a\|d^k\|^2 + \tfrac{1}{2}(d^k)^T(\nabla^2_{xx}L(x^k,\widetilde{u}^k) - H_k)d^k \\ &+ (1-\alpha)\sum_{j\in\widetilde{J}_k}\widetilde{u}_j^k f_j(x^k) + \alpha\sum_{j\in\widetilde{J}_k}\widetilde{u}_j^k\sigma_k z_k + o(\|d^k\|^2).\end{aligned}$$

So, using Assumption A_5 and the given conditions, one has

$$\omega_k \le ((\alpha-1)+\frac{1}{2})a\|d^k\|^2 + o(\|d^k\|^2).$$

Therefore, according to $\alpha \in (0,\frac{1}{2})$, we know (2.3) holds for $\lambda = 1$. The whole proof is finished. □

Theorem 4.1. *Under all above-mentioned assumptions, if* $\|(u_0^{k+1}, u_I^{k+1}) - (u_0^k, u_I^k)\| \to 0 (k\to\infty)$, *the algorithm is superlinearly convergent. i.e., the sequence* $\{x^k\}$ *generated by the algorithm satisfies* $\|x^{k+1} - x^*\| = o(\|x^k - x^*\|)$.

Proof. Similar to the proof of Theorem 5.3 in [8], we can conclude that

$$\|x^k + d^k - x^*\| = o(\|x^k - x^*\|).$$

In view of Lemma 4.2, one gets

$$\begin{aligned}\frac{\|x^{k+1}-x^*\|}{\|x^k - x^*\|} &\le \frac{\|x^k + d^k - x^*\| + \|\widetilde{d}^k - d^k\|}{\|x^k - x^*\|} \\ &\le \frac{\|x^k + d^k - x^*\|}{\|x^k - x^*\|} + \frac{\|\widetilde{d}^k - d^k\|}{\|d^k\|}\frac{\|d^k\|}{\|x^k - x^*\|} \\ &\le \frac{o(\|x^k - x^*\|)}{\|x^k - x^*\|} + \frac{\|\widetilde{d}^k - d^k\|}{\|d^k\|}\frac{\|d^k\|}{\|x^k - x^*\|} \to 0, \quad k\to\infty.\end{aligned}$$

□

5. Numerical experiments

In this section, we test some practical problems based on the proposed algorithm. The numerical experiments are implemented on MATLAB 6.5, under Windows XP and 1000MHZ CPU. The (2.1) and (2.2) are solved by the Optimization Toolbox. The BFGS formula, which is proposed in [3], is adopted in the algorithm.

During the numerical experiments, we set

$$\tau = 2.5, \quad \nu = 0.1, \quad \beta = 0.6, \quad \alpha = 0.3, \quad \sigma_1 = 0.6.$$

The test problem in Table 5.1 are selected from [19] and [20]. The initial points for the selected problems are as same as the ones in [19] and [20]. The columns of Table 5.1 have the following meanings: The (**n**, **m**) is the the number of variable and constraints of the test problems. The **prob** column lists the test problem taken from [19] and [20] in order. The columns labelled **Ni**, **Nf0** and **Nf** give the number of iterations required to solve the problem, objective function evaluations and constraint function evaluations(including linear and nonlinear constraints), respectively. The columns labelled **objective**, **dnorm** and **eps** denote the final objective value, the norm of d^k and the step criterion threshold ϵ, respectively.

The detailed information of the solutions to the test problems is listed in the following Table 5.1.

Table 5.1 Numerical results

Prob	**(n, m)**	**Ni**	**Nf0**	**Nf**	**objective**	**dnorm**	**eps**
hs001	(2, 2)	20	59	156	-0.0100e+02	4.1528e-012	0.1e-05
hs12	(2, 1)	19	56	74	-0.3000e+02	6.4675e-007	0.1e-05
hs29	(3, 1)	22	63	84	-0.2263e+02	8.1949e-007	0.1e-05
hs35	(3, 4)	11	21	124	0.1111e+00	1.4208e-006	0.1e-05
hs43	(4, 3)	29	88	348	-0.4400e+02	5.5220e-007	0.1e-05
hs100	(7, 4)	25	87	444	0.6806e+03	5.6542e-007	0.1e-05
hs108	(9, 14)	1	1	14	0.0000e+03	1.1124e-016	0.1e-05
s225	(2, 5)	9	17	85	0.0020+03	3.1656e-012	0.1e-05
s264	(4, 3)	28	84	312	0.4411+03	7.6810e-007	0.1e-05
s388	(15, 15)	50	173	3330	-0.5821+03	6.7815e-007	0.1e-05

6. Concluding remarks

In this paper, we have presented a feasible sequential quadratic programming algorithm for degenerate inequality constraints optimization problems. At each iteration of the proposed algorithm, a feasible direction of descent is obtained by solving a quadratic programming subproblem. To overcome the Maratos effect, a higher-order correction direction is obtained by solving another quadratic programms. The algorithm is proved to be globally convergent and superlinearly convergent under some mild conditions. Preliminary numerical results in Section 5 also show that the proposed algorithm is effective.

References

1. Daniel J.W., *Stability of solution of definte quadratic program*, Math. Program.,**5**(1973), 41-53.
2. Broyden, C.G., Dennis, J.E.,and More,J.J., *On the local and superlinear convergence of quasi-Newton methods*, J.Inst. Math. Appl., **12**(1973), 223-245.
3. M.J.D.Powell, *A fast algorithm for nonlinear constrained optimization calculations, in: G.A. Waston(Ed.)*, Numerical Analysis, Springer, Berlin, 1978.
4. P.T. Boggs and J.W. Tolle, *Sequential quadratic programming*, Acta Numerica, **1**(1995), 1-51.

5. L.Qi and Y.F.Yang, *A Globally and superlinearly convergent SQP algorithm for nonlinear constrained optimization,* AMR00/7, Applied Mathematics Report, University of New South Wales, Sydney, March 2000.
6. P.Spellucci, *A new technique for inconsistent QP probrems in the SQP methods, Math. Methods. Oper. Res.,* **47**(1998),355-400.
7. S.J.Wright, *Superlinear convergent of a stabilized SQP to a degenerate solution,* Comput. Optim. Appl., **11**(1998), 253-275.
8. S.J.Wright,*Modifying SQP for degenerate problems*,Preprint ANL/MCS-p699-1097, Mathematics and Computer Science Division, Agronne National Laboratory, Oct,.17,1997.
9. S. P. Han, *A globally convergent method for nonlinear programming,* J.Optim.Theory.Appl., **22**(1977), 297-309.
10. F. Facchinei, *Robust recursive quadratic programming algorithm model with global and superlinear convergence properties*, J.Optim.Theory.Appl., **92**(1997), 543-579.
11. M.M.Kostreva., X.Chen, *A superlinearly convergent method of feasible directions,* Appl.Math.Computation., **116**(2000), 245-255.
12. E.R. Panier and A.L. Tits, *A superlinearly convergent feasible method for the solution of inequality constrained optimization problems*, SIAM Journal on Control. Optim., **25**(1987), 934-950.
13. C.T. Lawrence., and A.L.Tits *A computationally efficient feasible sequential quadratic programming algorithm*, SIAM J.Optim., **11**(2001), 1092-1118.
14. J. B. Jian, K. C. Zhang, S. J. Xue, *A superlinearly and quadratically convergent SQP type feasible method for constrained optimization*, Appl. Math. J. Chinese Univ, Ser. B, **15**(2000), 319-331.
15. J. B. Jian, *Two extension models of SQP and SSLE algorithms for optimization and their superlinear and quadrtical convergence*,Appl. Math. J. Chinese Univ, Ser. A, **16**(2001), 435-444.
16. J.B.Jian., *Researches on superlinearly and quadratically convergent algorithm for nonlinearly constrained optimization*, Ph.D.Thesis. School of Xi'an Jiaotong University. Xi'an, China, 2000.
17. Facchinei.F., Lucidi.S., *Quadratically and superlinearly convergent algorithms for the solution of inequality constrained optimization problems*, J.Optim.Theory.Appl.,**85**(1995),265-289.
18. Zh.b.Zhu., *An efficient sequential quadratic programming algorithm for nonlinear programming*, J.Comput.Appl.Math., **175**(2005),447-464.
19. W. Hock, K. Schittkowski, *Test Examples for Nonlinear Programming Codes, Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin Heidelberg New York, 1981.
20. W. Hock, K. Schittkowski, *More Test Examples for Nonlinear Programming Codes*, Springer-Verlag, Berlin Heidelberg New York, 1987.

Yu Chen received her BS from Guangxi university in 2004. Her research interests focus on nonlinear programming.

School of Information, Hunan University of Commerce, 410205, Changsha, P.R. China
e-mail: chenyu4660@163.com.cn