J. Appl. Math. & Informatics Vol. 29(2011), No. 5 - 6, pp. 1395 - 1407
Website: http://www.kcam.biz

ON THE GLOBAL CONVERGENCE OF A MODIFIED SEQUENTIAL QUADRATIC PROGRAMMING ALGORITHM FOR NONLINEAR PROGRAMMING PROBLEMS WITH INEQUALITY CONSTRAINTS†

BINGZHUANG LIU

Abstract. When a Sequential Quadratic Programming (SQP) method is used to solve the nonlinear programming problems, one of the main difficulties is that the Quadratic Programming (QP) subproblem may be incompatible. In this paper, an SQP algorithm is given by modifying the traditional QP subproblem and applying a class of l_∞ penalty function whose penalty parameters can be adjusted automatically. The new QP subproblem is compatible. Under the extended Mangasarian-Fromovitz constraint qualification condition and the boundedness of the iterates, the algorithm is showed to be globally convergent to a KKT point of the non-linear programming problem.

AMS Mathematics Subject Classification : 90C30, 90C33, 65K05.
Key words and phrases : Nonlinear programming, sequential quadratic programming, l_∞ penalty function, global convergence.

1. Introduction

Consider the following nonlinear programming problem:

$$(P) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0, \quad i = 1, \cdots, m, \end{array} \tag{1}$$

where $f : R^n \to R$ and $g_i : R^n \to R$, $i = 1, 2, \cdots, m$, are all continuously differentiable functions in R^n.

It is well known that SQP method is a very popular and important method among all the effective methods for (P). The procedure of the traditional SQP algorithm is as follows.

Received December 3, 2010. Revised February 7, 2011. Accepted March 22, 2011. †This work was supported by the National Natural Science Foundation (No. 10971118 and 10901096) and by the research grant of Shandong University Technology.

© 2011 Korean SIGCAM and KSCAM.

At a current iteration point x^k, we usually solve the following QP subproblem:

$$\begin{array}{ll} \min & \langle f'(x^k), d\rangle + \frac{1}{2}\langle H^k d, d\rangle \\ \text{s.t.} & g_i(x^k) + \langle g_i'(x^k), d\rangle \le 0, \quad i = 1, \cdots, m, \end{array} \tag{2}$$

where H^k is a symmetric positive definite matrix. Then the next iterate has the form

$$x^{k+1} = x^k + \gamma_k d^k,$$

where d^k is a solution of the problem (2), γ_k is a step length which is chosen to make the value of some penalty function descend.

Except for the specific feasible descent SQP algorithm such as [**?**], most of SQP methods do not need that the approximate solution obtained at each iteration is feasible for (P). However, this is possible to make the QP subproblem (2) incompatible, that is, the feasible set of (2) is possible to be an empty set. By using some techniques, Burke and Han [**?**] modify the QP subproblem of their SQP method and make the feasible set of the QP subproblem nonempty for each $x \in R^n$, and get the global convergence of their SQP algorithm. Jiang and Ralph [**?**] also make the QP subproblem compatible by adding a variable to the subproblem. Similar situation can be found in [**?**, **?**, **?**, **?**, **?**, **?**, **?**, **?**]. In[**?**], Solodov propose an SQCQP method that the QCQP subproblems are compatible and have quadratic approximations to the constraint functions that are assumed to be convex. In[**?**], the SQP algorithm enjoys the global convergence and the boundedness of the primal iterates based on the strong conditions for the constraints, including convexity and the boundedness of the level set. Here we consider the QP subproblem with linear constraints without convexity of constraint functions in the problem (1). The QP subproblem of our paper is as follows.

At a current iterate x^k, we consider the following QP subproblem:

$$\begin{array}{lll} \min & \langle f'(x^k), d\rangle + \frac{1}{2}\langle H^k d, d\rangle + \beta_k t & \\ \text{s.t.} & g_i(x^k) + \langle g_i'(x^k), d\rangle \le t, & t \ge 0, \quad i \in I_k, \end{array} \tag{3}$$

where β_k is a penalty parameter, H^k is a positive definite matrix, $t \in R$ is a nonnegative variable added to make the QP subproblem compatible (In practical calculation, $t \ge 0$ is sometimes taken as $(t_1, \cdots, t_m)^T$ to get better results.), and I_k is an index set that satisfies that

$$\begin{array}{l} I(x^k) \subset I_k \subset \{1, \cdots, m\}, \\ I(x^k) = \{i \mid g_i(x^k) = p(x^k)\}, \quad p(x^k) = \max\limits_{i=i,\cdots,m} [g_i(x^k)]_+, \end{array} \tag{4}$$

where $[z]_+ = \max\{0, z\}$. It is useful that the index set $I_k \subset \{1, \cdots, m\}$, since it decreases the number of the constraints of QP subproblem. In [**?**], Zheng et.al give a QP subproblem without penalty parameter β_k in the objective function.

Let $(d^k, t^k) \in R^n \times R$ be the solution of (3), then the next iterate is given as

$$x^{k+1} = x^k + \gamma_k d^k,$$

where γ_k is obtained by using Armijo-type linear search to an l_∞ penalty function

$$\psi_{\beta_k}(x) = f(x) + \beta_k p(x).$$

For the general SQP method, it is difficult to choose the penalty parameter β_k when we add a variable to ensure the feasibility of the subproblem. [**?**] gives a very simple update formula to adjust the size of β_k. In this paper, we still use this formula.

Throughout in this paper, we suppose that *the extended Mangasarian-Fromovitz constraint qualification (EMFCQ) holds for (1) at any $x \in R^n$, that is, there exists a vector $p \in R^n$, such that*

$$\langle g_i'(x), p\rangle < 0,\ i \in I_+(x) = \{i \mid g_i(x) \geq 0\}. \tag{5}$$

It is obvious that EMFCQ implies the more familiar MFCQ condition, that is, there exists a $p \in R^n$ such that

$$\langle g_i'(x), p\rangle < 0,\ i \in I_0(x) = \{i \mid g_i(x) = 0\}.$$

The remaining of this paper is as follows. In Section 2 we establish the model of our SQP algorithm, and get the global convergence of our algorithm under very mild conditions. We get some computational results of the algorithm. Some conclusion Remarks for this paper are given finally.

In the following we give some denotations in this paper.

— For a directionally differentiable function $\psi : R^n \to R$, let $\psi'(x; d)$ denote the directional derivative of ψ in $d \in R^n$ at the point $x \in R^n$;

— Let $\|\cdot\|$ denote 2-norm, $\|\cdot\|_1$ denote l_1-norm, and $\|\cdot\|_\infty$ denote l_∞-norm in Euclidean space R^n, respectively;

— For two positive semi-definite matrices A and B, if $A - B$ is positive semi-definite, we denote that $A \succeq B$; and we let E be the identity matrix;

— $t/0 = +\infty$ is considered to be well defined, where $t > 0$;

— $a = o(b)$ means that for $b \to 0$, $a/b \to 0$.

2. Main results

2.1. Model of an Algorithm. In this section we give the model of our SQP algorithm. First we note that the problem (3) is always feasible and has unique solution. In fact, it is easy to see that for each fixed $d \in R^n$, the minimum with respect to t in (3) is attained at

$$t^k(d) = \max_{i \in I_k}[g_i(x^k) + \langle g_i'(x^k), d\rangle]_+. \tag{6}$$

So (3) is equivalent to

$$\min_{d \in R^n}\ \langle f'(x^k), d\rangle + 1/2\langle H^k d, d\rangle + \beta_k t^k(d). \tag{7}$$

Since H^k is positive definite, the objective function of (7) is strongly convex and has the unique minimizer d^k. It follows that (d^k, t^k) with $t^k = t^k(d)$ is the unique solution of (3). Furthermore, since the constraints in (3) satisfy the

EMFCQ condition (see (5)), the KKT optimality conditions hold, that is, there exist some $\mu_i^k \in R$, $i \in I_k$, and $\nu^k \in R$ such that

$$f'(x^k) + H^k d^k + \sum_{i \in I_k} \mu_i^k g_i'(x^k) = 0, \tag{8}$$

$$\beta_k - \sum_{i \in I_k} \mu_i^k - \nu^k = 0, \tag{9}$$

$$g_i(x^k) + \langle g_i'(x^k), d^k \rangle \leq t^k, \quad \mu_i^k \geq 0, \ i \in I_k, \tag{10}$$

$$\mu_i^k (g_i(x^k) + \langle g_i'(x^k), d^k \rangle - t^k) = 0, \ i \in I_k, \tag{11}$$

$$t^k \geq 0, \ \nu^k \geq 0, \ t^k \nu^k = 0. \tag{12}$$

In the following we state our algorithm.

Algorithm 1

Initial. Choose $x^0 \in R^n$, β_0, δ_1, $\delta_2 \in (0, +\infty)$, $\sigma \in (0, 1/2)$, *and* $\theta \in (0, 1)$, *set* $k := 0$.

Step 1. Choose an index set I_k *according to (4). Choose a* $n \times n$ *matrix* H_k *(positive definite). Compute* (d^k, t^k) *as the solution of (3), and the associated Lagrange multipliers* (μ^k, ν^k).

Step 2. If $d^k = 0$ *and* $t^k = 0$, *stop.*

Step 3. If $d^k = 0$ *and* $t^k > 0$, *let* $j_k = 0$, *and go to Step 4. Otherwise, find* j_k, *the smallest nonnegative integer* j, *such that*

$$\psi_{\beta_k}(x^k + \theta^j d^k) \leq \psi_{\beta_k}(x^k) + \sigma \theta^j \Delta_k, \tag{13}$$

where $\psi_{\beta_k}(x)$ *is the* l_∞ *penalty function*

$$\psi_{\beta_k}(x) = f(x) + \beta_k p(x)$$

given in section 1, and

$$\Delta_k = \langle f'(x^k), d^k \rangle + 1/2 \langle H^k d^k, d^k \rangle + \beta_k (t^k - p(x^k)). \tag{14}$$

Step 4. Let $\gamma_k = \theta^{j_k}$, $x^{k+1} = x^k + \gamma_k d^k$.

Step 5. Compute $r_k = \min\{\|d^k\|^{-1}, \|\mu^k\|_1 + \delta_1\}$. *Set*

$$\beta_{k+1} := \begin{cases} \beta_k, & \textit{if } \beta_k \geq r_k; \\ \beta_k + \delta_2, & \textit{otherwise.} \end{cases} \tag{15}$$

Step 6. Let $k := k + 1$, *Go to Step 1.*

From the above algorithm it is clear that if $\{\beta^k\}$ is unbounded, then it must have $\|\mu^k\| \to +\infty$ and $d^k \to 0$. We will show that this situation cannot occur, thus lead to the boundedness of $\{\beta_k\}$. In [**?**], the penalty parameter $\{\beta^k\}$ with another update rule is not guaranteed to be bounded.

2.2. Global Convergence. In this section, we get the global convergence of the algorithm given in Section 2. We suppose in this section that $f : R^n \to R$ and $g_i : R^n \to R$, $i = 1, 2, \cdots, m$ in Problem (P), are all continuously differentiable functions in R^n and the constraint qualification condition EMFCQ holds for (P).

First we give a useful lemma that is given in [**?**].

Lemma 2.1. *Let* $\{x|g_i(x) \leq 0, i = 1, 2, \cdots, m\} \neq \emptyset$, *then for any* $x \in R^n$ *satisfying*

$$I_{++}(x) := \{i|g_i(x) > 0\} \neq \emptyset,$$

we have that

$$\sum_{i \in I_{++}(x)} \mu_i g_i'(x) = 0, \ \ \mu_i \geq 0, \ \ i \in I_{++}(x)$$

if and only if

$$\mu_i = 0, \ \ i \in I_{++}(x).$$

In the following we show that Algorithm 1 is well-defined.

Lemma 2.2. *Let* $\{x|g_i(x) \leq 0, i = 1, 2, \cdots, m\} \neq \emptyset$,

(i) if we get that $d^k = 0$, *and* $t^k = 0$ *in Algorithm 1, then* (x^k, μ^k) *is a KKT point of (1);*

(ii) if $d^k = 0$ *and* $t^k > 0$, *then there exists a finite integer* q *such that Algorithm 1 generates* $d^{k+q} \neq 0$.

The proof of this lemma is similar with Proposition 1 in [**?**], so we omit it.

The following result shows that whenever $d^k \neq 0$, it is a descent direction for ψ_{β_k} at x^k, which in turn implies that the line-search step is well-defined. Combining this fact with the above lemma, it follows that Algorithm 1 is well-defined.

Lemma 2.3. *In Algorithm 1 it holds that*

$$\begin{aligned} \psi'_{\beta_k}(x^k; d^k) &\leq \Delta_k - 1/2\langle H^k d^k, d^k\rangle \\ &\leq -\langle H^k d^k, d^k\rangle - \nu^k p(x^k), \end{aligned} \tag{16}$$

therefore Step 3 of Algorithm 1 is well-defined and terminates with some finite integer j_k.

Proof. It is easy to see that

$$\psi'_{\beta_k}(x^k; d^k) = \langle f'(x^k), d^k\rangle + \beta_k \begin{cases} 0, & \text{if } I(x^k) = \emptyset; \\ \max\limits_{i \in I(x^k)} \langle g_i'(x^k), d^k\rangle, & \text{if } p(x^k) > 0; \\ \max\limits_{i \in I(x^k)} [\langle g_i'(x^k), d^k\rangle]_+, & \text{if } p(x^k) = 0, I(x^k) \neq \emptyset. \end{cases} \tag{17}$$

Then we consider the three possible cases in (17).

If $I(x^k) = \emptyset$, then $p(x^k) = 0$ (i.e. $g_i(x^k) < 0$, for any i). Therefore,

$$0 \leq t^k = t^k - p(x^k).$$

If $I(x^k) \neq \emptyset$, from (10), for any $i \in I(x^k) \subset I_k$, we have

$$\langle g_i'(x^k), d^k \rangle \leq t^k - g_i(x^k) = t^k - p(x^k),$$

where the equality follows from that $g_i(x^k) = p(x^k),\ i \in I(x^k)$.

Furthermore, if $p(x^k) = 0$, then $\langle g_i'(x^k), d^k \rangle \leq t^k$. From the monotonicity of $[\cdot]_+$, it follows that

$$[\langle g_i'(x^k), d^k \rangle]_+ \leq [t^k]_+ = t^k - p(x^k).$$

Combining the above three cases, we have from (17) that

$$\begin{aligned} \psi_{\beta_k}'(x^k; d^k) &\leq \langle f'(x^k), d^k \rangle + \beta_k \langle t^k - p(x^k) \rangle \\ &= \Delta_k - 1/2 \langle H^k d^k, d^k \rangle, \end{aligned} \tag{18}$$

then the first inequality in (16) holds.

Multiplying both sides of (8) by d^k, we have that

$$\langle f'(x^k), d^k \rangle = -\langle H^k d^k, d^k \rangle - \sum_{i \in I(x^k)} \mu_i^k \langle g_i'(x^k), d^k \rangle. \tag{19}$$

Furthermore, we have that

$$\begin{aligned} -\sum_{i \in I(x^k)} \mu_i^k \langle g_i'(x^k), d^k \rangle &= \sum_{i \in I(x^k)} \mu_i^k (g_i(x^k) - t^k) \\ &\leq (p(x^k) - t^k) \sum_{i \in I(x^k)} \mu_i^k \\ &= (\beta_k - \nu^k)(p(x^k) - t^k) \\ &= \beta_k (p(x^k) - t^k) - \nu^k p(x^k), \end{aligned} \tag{20}$$

where (11) was used in the first equality, (9) was used in the second equality, and (12) in the last. Combining (20) with (18) and (19), we get (16).

If $d^k = 0$, by Step 3 of Algorithm 1, we have that $j_k = 0$. If $d^k \neq 0$, then for $\gamma \in [0, 1]$, we have that

$$\begin{aligned} \psi_{\beta_k}(x^k + \gamma d^k) &= \psi_{\beta_k}(x^k) + \gamma \psi_{\beta_k}'(x^k; d^k) + o(\gamma) \\ &\leq \psi_{\beta_k}(x^k) + \gamma \Delta_k + o(\gamma), \end{aligned} \tag{21}$$

where the inequality was by (16). It follows that (13) is guaranteed to hold whenever $\gamma = \theta^j > 0$ satisfies that

$$(1 - \sigma)\gamma \Delta_k \leq o(\gamma).$$

Since $d^k \neq 0$ when $\Delta_k < 0$, the above inequality holds for all $\gamma = \theta^j$ sufficiently small. Hence, Step 3 of Algorithm 1 terminates with finite integer j_k. □

We next show that when close to the feasible region of (1), the solution of the subproblem (3) is given by the solution of the subproblem without the slack variable:

$$\begin{aligned} \min \quad & \langle f'(x^k), d \rangle + \tfrac{1}{2} \langle H^k d, d \rangle \\ \text{s.t.} \quad & g_i(x^k) + \langle g_i'(x^k), d \rangle \leq 0, \quad i \in A_k, \end{aligned} \tag{22}$$

where $A_k \subset \{1, 2, \cdots, m\}$. This fact will be used latter to establish that the penalty parameters β_k stay fixed from some point on. We first make the following assumption:

(**A**) *there exist ρ_1, ρ_2 such that the matrices in (22) satisfy $\rho_2 E \succeq H^k \succeq \rho_1 E$, for all k, where $\rho_2 \geq \rho_1 > 0$.*

Lemma 2.4. *Let $\{x^k\}$ be any sequence converging to some $\bar{x} \in R^n$ such that $p(\bar{x}) = 0$ and the assumption (A) hold. Then the problem (22) is feasible for all sufficiently large k and any $A_k \subset \{1, 2, \cdots, m\}$. Moreover, if (s^k, λ^k) is a KKT point of (22), then for any $\beta_k \geq \|\lambda^k\|_1$, $(s^k, 0)$ is the unique solution of (3) $(I_k = A_k)$. Conversely, if $(d^k, 0, \mu^k, \nu^k)$ is a KKT point of (3), then (d^k, μ^k) is a KKT point of (22). Furthermore, the sequences $\{s^k\}$, $\{\lambda^k\}$ are both bounded.*

Proof. Since $p(\bar{x}) = 0$, we have that $g_i(\bar{x}) = 0$, for $i \in I(\bar{x})$. By the EMFCQ conditon (5), there exists a vector $p \in R^n$ such that

$$\langle g_i'(\bar{x}), p\rangle < 0, \; i \in I(\bar{x}). \tag{23}$$

For each $i \in \{1, 2, \cdots, m\}$, there exists $\eta_i > 0$ and $c_i > 0$ such that

$$\eta_i \langle g_i'(\bar{x}), p\rangle \leq \begin{cases} -2c_i, & \text{if } \; i \in I(\bar{x}); \\ -g_i(\bar{x})/4, & \text{otherwise,} \end{cases} \tag{24}$$

where (23) is used for $i \in I(\bar{x})$. Then for each $i \in \{1, 2, \cdots, m\}$, there exists an index k_i such that for all $k \geq k_i$,

$$\varepsilon_i^k := g_i(x^k) + \eta_i \langle g_i'(x^k) - g_i'(\bar{x}), p\rangle \leq \begin{cases} c_i & \text{if } i \in I(\bar{x}); \\ g_i(\bar{x})/2 & \text{otherwise.} \end{cases}$$

Denoting $\bar{k} = \max_{i=1,2,\cdots,m} k_i$, $\eta = \min_{i=1,2,\cdots,m} \eta_i$, $c = \min\{\min_{i \in I(x)} c_i, -\max_{i \notin I(x)} g_i(\bar{x})/4\}$, and using the two relations above, we have that $d = \eta p$ satisfies

$$g_i(x^k) + \langle g_i'(x^k), d\rangle \leq -c < 0, \; \forall k \geq \bar{k}, \; i = 1, 2, \cdots, m. \tag{25}$$

It follows that d is (strictly) feasible in (22), for any choice of $A_k \subset \{1, 2, \cdots, m\}$ and $k \geq \bar{k}$.

Therefore, (22) is a feasible strongly convex program. Hence, it is uniquely solvable. Thus, there exist some $\lambda_i^k \in R, \; i \in A_k$ such that

$$f'(x^k) + H^k s^k + \sum_{i \in A_k} \lambda_i^k g_i'(x^k) = 0, \tag{26}$$

$$g_i(x^k) + \langle g'(x^k), s^k\rangle \leq 0, \; \lambda_i^k \geq 0, \; i \in A_k, \tag{27}$$

$$\lambda_i^k (g_i'(x^k) + \langle g'(x^k), s^k\rangle) = 0, \; i \in A_k. \tag{28}$$

If $\beta_k \geq \|\lambda^k\|_1$, then the conditions (26)-(28) imply that $d^k = s^k, \; t^k = 0, \; \nu^k = \beta_k - \|\lambda^k\|_1$ and $\mu_i^k = \lambda_i^k, \; i \in A_k = I_k$, satisfy KKT conditions (8)-(12). Therefore, $(s^k, 0)$ is a solution of (3). Conversely, if $(d^k, 0, \mu^k, \nu^k)$ is a KKT point of (3), then for $A_k = I_k$, (d^k, μ^k) is a KKT point of (22).

Next, note that since $d = \eta p$ is feasible in (22) for all $k \geq \bar{k}$, we have that

$$\begin{aligned} & \langle f'(x^k), d\rangle + 1/2\langle H^k d, d\rangle \\ \geq\ & \langle f'(x^k), s^k\rangle + 1/2\langle H^k s^k, s^k\rangle \\ \geq\ & \|s^k\|(\rho_1\|s^k\|^2/2 - \|f'(x^k)\|), \end{aligned}$$

where the second inequality is by $H^k \succeq \rho_1 E$. Since $\{f'(x^k)\}$ is bounded and $\rho_2 E \succeq H^k$, the above relation implies that $\{s^k\}$ must be bounded.

Suppose now that $\{\lambda^k\}$ is unbounded. Let $\lambda_i^k = 0$, for $i \in \{1, 2, \cdots, m\} \setminus A_k$. Passing onto a subsequence if necessary, we can assume that $\|\lambda^k\| \to \infty$, and $\{s^k\} \to \bar{s}, \{H^k\} \to \bar{H}$. Dividing both sides of (26) by $\|\lambda^k\|$ and passing onto the limit as $k \to \infty$, we obtain that

$$0 = \sum_{i=1}^{m} \bar{\lambda}_i g_i(\bar{x}),\ \ \bar{\lambda} \geq 0,\ \ \|\bar{\lambda}\| = 1.$$

The latter is equivalent to that there doesn't exist $s \in R^n$, such that

$$\langle g_i'(\bar{x}), s\rangle < 0,\quad i \in \{i|\bar{\lambda}_i > 0\}. \tag{29}$$

Clearly, $\bar{\lambda}_i > 0$ means that there exists an infinite subsequence of indices $\{k_j\}$ such that $i \in A_{k_j}$ and $\lambda_i^{k_j} > 0$. For such i, (28) implies that

$$g_i(x^{k_j}) + \langle g_i'(x^{k_j}), s^{k_j}\rangle = 0,\ \forall j.$$

Passing onto the limit as $j \to \infty$, we obtain that

$$g_i(\bar{x}) + \langle g_i'(\bar{x}), \bar{s}\rangle = 0,\ i \in \{i|\bar{\lambda}_i > 0\}. \tag{30}$$

Passing onto the limit in (25) as $k \to \infty$, we also have that

$$g_i(\bar{x}) + \langle g_i'(\bar{x}), d\rangle \leq -c < 0.$$

Subtracting (30) from the above inequality, we have that

$$\langle g_i'(\bar{x}), d - \bar{s}\rangle < 0,\ i \in \{i|\bar{\lambda}_i > 0\},$$

which is a contradiction. We conclude that $\{\lambda^k\}$ is bounded. □

By (15), we can obtain that either β_k is a constant starting from some iteration index k_0, or it diverges to $+\infty$. We next show that the latter case cannot occur.

Lemma 2.5. *Let the sequence $\{x^k\}$ generated by Algorithm 1 be bounded and the assumption (A) hold. Then there exists some iteration index k_0 such that*

$$\beta_k = \beta_{k_0},\ \forall k \geq k_0.$$

Proof. Suppose the opposite, i.e., $\beta_k \to +\infty$, as $k \to \infty$. Then from that

$$\beta_k < r_k = \min\{\|d^k\|^{-1},\ \|\mu^k\|_1 + \delta_1\},$$

it follows that there exists a subsequence of iteration indices $\{k_j\}$ such that

$$\lim_{j\to\infty} d^{k_j} = 0 \text{ and } \lim_{j\to\infty} \mu^{k_j} = +\infty. \tag{31}$$

Taking a further subsequence, if necessary, we can assume that $\{x^{k_j}\} \to \tilde{x}$, as $j \to \infty$. We next consider the two possible cases: $p(\tilde{x}) > 0$ or $p(\tilde{x}) = 0$.

Let $p(\tilde{x}) > 0$, and denote

$$I_{++}(\tilde{x}) := \{i | g_i(\tilde{x}) > 0\} \neq \emptyset.$$

Note that, by the continuity of g, $I_{++}(\tilde{x}) \cap I_{k_j} \neq \emptyset$ for all j large enough. By (10),

$$g_i(x^{k_j}) + \langle g_i'(x^{k_j}), d^{k_j} \rangle \leq t^{k_j}, \ i \in I_{k_j}.$$

For $i \in I_{k_j} \setminus I_{++}(\tilde{x})$, as $j \to \infty$ the left-hand side of the inequality above tends to $g_i(\tilde{x}) \leq 0$ (notice (33)), while the right-hand side tends to $p(\tilde{x}) > 0$ (recall (6)). Hence, such constraints are inactive for all j large enough and, by (11),

$$\mu_i^{k_j} = 0, \ i \in I_{k_j} \setminus I_{++}(\tilde{x}).$$

Formally setting $\mu_i^{k_j} = 0$, $i \in I_{++}(\tilde{x}) \setminus I_{k_j}$, we can write (8) as

$$f'(x^{k_j}) + H^{k_j} d^{k_j} + \sum_{i \in I_{++}(\tilde{x})} \mu_i^{k_j} g_i'(x^{k_j}) = 0.$$

Dividing both sides of this equality by $\|\mu^{k_j}\|$ and let $j \to \infty$, by (31) and $\rho_2 E \succeq H^{k_j}$, we obtain that

$$\sum_{i \in I_{++}(\tilde{x})} \tilde{\mu} g_i'(\tilde{x}) = 0, \ \tilde{\mu}_i \geq 0, \ i \in I_{++}(\tilde{x}), \ \|\tilde{\mu}\| = 1.$$

which contradicts with Lemma 1.

Suppose now that $p(\tilde{x}) = 0$. By Lemma 2.4, the subproblem (22) with $A_k = I_{k_j}$, is solvable for all indices j large enough. Let (s^{k_j}, λ^{k_j}) be a KKT point of (22). In particular, by Lemma 2.4, $\{\lambda^{k_j}\}$ is bounded. Since $\beta_k \to +\infty$, we have that $\beta_{k_j} > \|\lambda^{k_j}\|$, and the unique solution of (3) is $(d^{k_j}, t^{k_j}) = (s^{k_j}, 0)$, with multipliers (μ^{k_j}, ν^{k_j}) generated by Algorithm 1. But then (s^{k_j}, μ^{k_j}) is a KKT point of (22), again by Lemma 2.4. However the unboundedness of $\{\mu^{k_j}\}$ contradicts with Lemma 2.4. Then the proof is completed. □

We are now ready to establish the global convergence of Algorithm 1. First we have an important conclusion as follows.

Lemma 2.6. *Suppose that the conditions assumed in Lemma 2.5, are satisfied, then we have that*

$$\gamma_k d^k \to 0, \ k \to \infty.$$

Proof. First from Algorithm 1, it follows that

$$\begin{aligned} \psi_{\beta_{k+1}}(x^{k+1}) &\leq \psi_{\beta_k}(x^k) + \sigma \gamma_k \Delta_k \\ &\leq \psi_{\beta_k}(x^k) - \sigma/2 \gamma_k \langle H^k d^k, d^k \rangle, \end{aligned} \tag{32}$$

where the second inequality is by (16) ($\nu^k \geq 0$, $p(x^k) \geq 0$). Therefore, $\{\psi_{\beta_k}(x^k)\}$ is a non-increasing sequence. Since $\{x^k\}$ is bounded, $\rho_2 E \succeq H^k \succeq \rho_1 E$, where

$\rho_2 \geq \rho_1 > 0$, from Lemma 2.5, we know that there exists an index k_0, such that $\beta_{k+1} = \beta_k = \beta_{k_0}$, as $k \geq k_0$. Then from (32), we have that for all $k \geq k_0$,

$$\psi_{\beta_{k_0}}(x^k) - \psi_{\beta_{k_0}}(x^{k+1}) \geq \sigma\gamma_k\rho_1\|d^k\|^2/2.$$

Since $\{\psi_{\beta_k}(x^k)\}$ is non-increasing and $\{x^k\}$ is bounded, we have that $\gamma_k d_k \to 0$, as $k \to \infty$. □

In the following we give the main result of the global convergence of our algorithm.

Theorem 2.7. *Suppose that (A) holds and there exists an index k_1 such that for all $k \geq k_1$, $I_k = \{1, 2, \cdots, m\}$. If the sequence $\{x^k\}$ generated by Algorithm 1 is bounded, then the sequence $\{\mu^k\}$ which is obtained by Algorithm 1 is also bounded, and every accumulation point of the sequence $\{(x^k, \mu^k)\}$ is a KKT point of the problem (1).*

Proof. First from Lemma 2.6, we have that $\gamma_k d^k \to 0$, as $k \to \infty$. Now we prove that $d^k \to 0$, as $k \to \infty$. Suppose it is not so, then there exist a subsequence $\{d^{k_i}\}$ of $\{d^k\}$ and a constant $\varepsilon > 0$, such that

$$\|d^{k_i}\| \geq \varepsilon, \ \forall i. \tag{33}$$

We next prove that there exists γ' such that

$$\gamma_{k_i} \geq \gamma', \ \forall i. \tag{34}$$

Suppose that (34) does not hold, then there exists a subsequence of $\{\gamma_{k_i}\}$, which without loss of generality, is set by $\{\gamma_{k_i}\}$ itself, such that

$$\gamma_{k_i} \to 0, \ i \to \infty.$$

From Step 3, Step 4 of Algorithm 1, and Lemma 2.5, we know that if i is sufficiently large ($i > k_0$), then

$$\psi_{\beta_{k_0}}(x^{k_i} + \frac{\gamma_{k_i}}{\theta}d^{k_i}) > \psi_{\beta_{k_0}}(x^{k_i}) + \sigma\frac{\gamma_{k_i}}{\theta}\Delta_{k_i}..$$

On the other hand, from that $\gamma_{k_i} \to 0, \ i \to \infty$, it follows that

$$\begin{aligned}\psi_{\beta_{k_0}}(x^{k_i} + \frac{\gamma_{k_i}}{\theta}d^{k_i}) &= \psi_{\beta_{k_0}}(x^{k_i}) + \frac{\gamma_{k_i}}{\theta}\psi_{\beta_{k_0}}(x^{k_i}; d^{k_i}) + o(\frac{\gamma_{k_i}}{\theta}) \\ &\leq \psi_{\beta_{k_0}}(x^{k_i}) + \frac{\gamma_{k_i}}{\theta}\Delta_{k_i} + o(\frac{\gamma_{k_i}}{\theta}),\end{aligned}$$

where the inequality above is obtained by (21) in the proof of Lemma 2.3. Therefore, we have that

$$\begin{aligned}\frac{\gamma_{k_i}}{\theta}\Delta_{k_i} + o(\frac{\gamma_{k_i}}{\theta}) &\geq \psi_{\beta_{k_0}}(x^{k_i} + \frac{\gamma_{k_i}}{\theta}d^{k_i}) - \psi_{\beta_{k_0}}(x^{k_i}) \\ &> \sigma\frac{\gamma_{k_i}}{\theta}\Delta_{k_i}.\end{aligned}$$

Thus,

$$(1-\sigma)\Delta_{k_i} + o(\frac{\gamma_{k_i}}{\theta})/\frac{\gamma_{k_i}}{\theta} \geq 0.$$

Since $\Delta_{k_i} \leq -1/2\langle H^{k_i} d^{k_i}, d^{k_i}\rangle$ (by (16)), and from (33), we obtain that

$$-(1-\sigma)\rho_1\varepsilon^2 + o(\frac{\gamma_{k_i}}{\theta})/\frac{\gamma_{k_i}}{\theta} \geq 0.$$

Let $i \to \infty$, noting that $\gamma^{k_i} \to 0$, then we have that

$$1 - \sigma \leq 0,$$

which contradicts with $\sigma \in (0, 1/2)$, therefore (34) holds. From (33) and (34), we know that $\gamma_k d^k \not\to 0$, which contradicts with Lemma 2.6. Therefore,

$$\lim_{k\to\infty} d^k = 0. \tag{35}$$

Furthermore, from Step 5 of Algorithm 1, and Lemma 2.5, we know that

$$\beta_{k_0} \geq r_k = \min\{\|d^k\|^{-1},\ \|\mu^k\|_1 + \delta_1\},$$

then there exists $k_2 \geq \max\{k_0, k_1\}$, such that

$$\beta_{k_0} > \sum_{i\in I_k} \mu^k,\ \forall k \geq k_2,$$

which implies that $\{\mu^k\}$ is bounded. From (9), it follows that

$$\nu^k > 0,\ \forall k \geq k_2.$$

Furthermore, by (12), we obtain that

$$t^k = 0,\ \forall k \geq k_2. \tag{36}$$

Let $k \to \infty$ in (8)-(12), where $I_k = \{1, 2, \cdots, m\}$, we obtain that every accumulation point of $\{(x^k, \mu^k)\}$ satisfies the KKT condition of (1). This completes the proof. □

2.3. Numerical results. In this section, we tested two typical problems taken from [**?**] and [**?**] by using our algorithm. The parameters in the algorithm were selected as $\beta_0 = 10$, $\delta_1 = 1$, $\delta_2 = 1$, $\sigma = 0.1$, and $\theta = 0.5$. And the numerical results are given in Table 1 and Table 2, whose columns have the following meaning:

x^0—— the initial point;
Ni —— the number of iterations;
Nf —— the number of objective function evaluations;
Ng —— the number of constraints evaluations;
obj—— the optimal value of the objective function.

Problem 1. Example hs100 [**?**]

$$\begin{aligned}
\min \quad & f(x) = (x_1-10)^2 + 5(x_2-12)^2 + x_3^4 + 3(x_4-11)^2 \\
& \qquad +10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7 \\
\text{s.t.} \quad & g_1(x) = 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 - 127 \leq 0, \\
& g_2(x) = 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 - 282 \leq 0, \\
& g_3(x) = 23x_1 + x_2^2 + 6x_6^2 - 8x_7 - 196 \leq 0,
\end{aligned}$$

$$g_4(x) = 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \leq 0.$$

Table 1. Results for Example hs100 with different initial points.

x^0	Ni	Nf	Ng	obj
(10, 10, 10, 10, 10, 10, 10)	43	89	356	6.806300573632109e+002
(5, 5, 5, 5, 5, 5, 5)	37	91	364	6.806300574685115e+002
(1, 1, 1, 1, 1, 1, 1)	25	66	264	6.806300573898336e+002
(1, 2, 0, 4, 0, 1, 1)	12	31	124	6.806300572078540e+002

Problem 2. Example s264 [?]

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\ \text{s.t.} \quad & g_1(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 - x_3 - x_4 - 8 \leq 0, \\ & g_2(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 9 \leq 0, \\ & g_3(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 \leq 0. \end{aligned}$$

Table 2. Results for Example s264 with different initial points.

x^0	Ni	Nf	Ng	obj
(1, 1, 1, 1)	8	18	54	-44.11340764633566e-002
(0, 0, 0, 0)	8	16	48	-44.11340754856773e-002
(2, 2, 2, 2)	11	18	54	-44.11340879313437e-002
(4, 4, 4, 4)	12	16	48	-44.11340683863549e-002

From Table 1 and Table 2, we can see that the initial points can be chosen arbitrarily, which shows the algorithm is numerically stable for the above two problems.

3. Conclusions remarks

In this paper we propose a modified SQP algorithm with global convergence. The algorithm enjoys well properties. Firstly, the QP subproblems are always compatible. Secondly, only first order derivatives of the problem functions are needed. Thirdly, simple automatic adjustment rules for the penalty function parameters are used. Finally, the line search of the step length is well-defined in Algorithm 1.

References

1. J. V. Burke, *A sequential quadratic programming algorithmfor potentially infeasible mathematical programs*, J. Math. Anal. and Applications **139**(1989), 319-351.
2. J. V. Burke and S. P. Han, *A robust sequential quadratic programming method*, Math. Programming **43**(1989), 277-303.
3. W. Hock and K. Schittkowski, *Test examples for nonlinear programming codes*, lecture notes in economics and mathematical systems **187**, Springer-Verlag, Berlin Heidelberg New York, 1981.
4. A. D. Ioffe, *Necessary and sufficient conditions for a local minimum. III: Second-order conditions and augmented duality*, SIAM J. Optimization **17**(1979), 266-288.

5. J. B. Jian and C. M. Tang, *An SQP feasible descent algorithm for nonlinear inequality constrained optimization without strict complementarity*, Comput. and Math. with Applications **49**(2005), 223-238.
6. H. Y. Jiang and D. Ralph, *Smooth SQP methods for mathematical programs with nonlinear complementarity constraints*, SIAM J. Optimization **10**(2000), 779-808.
7. J. F. A. Pantoja and D. Q, Mayne, *Exact penalty function algorithm with simple updating of the penalty parameter*, J. Optim. Theory and Applications **69**(1991), 441-467.
8. L. Q. Qi and Y. F. Yang, *A Globally and Superlinearly Convergent SQP Algorithm for Nonlinear Constrained Optimization*, J. Global Optimization **21**(2001), 157-184.
9. M. V. Solodov, *On the sequential quadratically constrained quadradic programming methods*, Math. Oper. Research **29**(2004), 64-79.
10. M. V. Solodov, *Global convergence of an SQP method without boundedness assumptions on any of the iterative sequences*, Math. Programming **118**(2009), 1C12.
11. K. Schittkowski, *More test examples for nonlinear programming codes*, Spring-Verlag, 1987.
12. P. Spellucci, *A new technique for inconsistent QP problems in the SQP methods*, Math. Meth. of Oper. Research**47**(1998), 355-400.
13. K. Tone, Revision of constraint approximations in the successive QP-method for nonlinear programming problems. Math. Programming **26**(1983), 144-152.
14. H. Y. Zheng, J. B. Jian, C. M. Tang, and R. Quan, *A new norm-relaxed SQP algorithm with global convergence*, Appl. Math. Letters **23**(2010), 670-675.

Bingzhuang Liu received M.Sc. from Qufu Normal University and Ph.D at Shanghai University. Since 2008 he has been at Shandong University of Technology. His research interests include Mathematical Programming and Nonlinear Equations.

School of Science, Shandong University of Technology, Zibo 255049, P.R. China.
e-mail: lbzlyj@126.com