A Norm-Relaxed Algorithm with Identification Function for General Constrained Optimization*

JIAN Jinbao[1†] WEI Xiaopeng[1] ZENG Hanjun[1] PAN Huaqin[1]

Abstract Based on a semi-penalty function and an identification function used to yield a "working set", as well as the norm-relaxed SQP idea, a new algorithm for solving a kind of optimization problems with nonlinear equality and inequality constraints is proposed. At each iteration, to yield the search directions the algorithm solves only one reduced quadratic program (QP) subproblem and a reduced system of linear equations. The proposed algorithm possesses global convergence and superlinear convergence under some mild assumptions without the strictly complementarity. Finally, some elementary numerical experiments are reported.

Keywords Operations research, general constraints, optimization, norm-relaxed algorithm, identification function, global convergence, superlinear convergence

Chinese Library Classification O22

一般约束优化基于识别函数的模松弛算法

简金宝[1†] 韦小鹏[1] 曾汉君[1] 潘华琴[1]

摘要 借助于半罚函数和产生工作集的识别函数以及模松弛 SQP 算法思想，建立了求解带等式及不等式约束优化的一个新算法. 每次迭代中，算法的搜索方向由一个简化的二次规划子问题及一个简化的线性方程组产生. 算法在不包含严格互补性的温和条件下具有全局收敛性和超线性收敛性. 最后给出了算法初步的数值试验报告.

关键词 运筹学，一般约束，最优化，模松弛算法，识别函数，全局收敛性，超线性收敛性

中图分类号 O22

0 Introduction

In this paper, we consider the general constrained optimization problems as follows

$$
(\mathrm{P}) \quad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_j(x) \leqslant 0, \quad j \in I_1 = \{1, \dots, m'\}, \\ & g_j(x) = 0, \quad j \in I_2 = \{m'+1, \dots, m\}, \end{array} \tag{0.1}
$$

收稿日期：2009 年 1 月 6 日.

∗ Supported by the National Natural Science Foundation of China (Nos. 71061002, 10771040), and the Natural Science Foundation of Guangxi Province (No. 2011GXNSFD018022)

1. College of Mathematics and Information Science, Guangxi University, 530004 Nanning, China; 广西大学数学与信息科学学院，南宁 530004.

† Corresponding author 通讯作者

where $x \in R^n, f$ and $g_j \ (j=1,\ldots,m) : R^n \to R$ are smooth functions. Denote the feasible set of the problem (0.1) as

$$X = \{x \in R^n : \ g_j(x) \leqslant 0, j \in I_1; \ g_j(x) = 0, j \in I_2\}.$$

It is well known that the method of feasible directions (MFD) is one of the important methods to deal with the optimization problems with inequality constraints. The MFD possesses several good properties such as the feasibility of the iterative points and the approximate optimal solutions as well as computational efficiency and so on, see Refs. [1-4]. To use the idea of the MFD to study the general constrained optimization problem (0.1), In [5], Mayne and Polak converted the problem (P) to the following semi-penalty optimization problem with only inequality constraints

$$(\mathrm{P}_c) \quad \begin{array}{ll} \min & F_c(x) \overset{\mathrm{Def}}{=} f(x) - c \sum\limits_{j \in I_2} g_j(x) \\ \text{s.t.} & g_j(x) \leqslant 0, \quad j \in I \overset{\mathrm{Def}}{=} I_1 \cup I_2. \end{array} \qquad (0.2)$$

where the penalty parameter $c > 0$, which is updated by a suitable rule. Then, based on the auxiliary problem (P_c), Mayne and Polak presented a MFD associated to (P_c) for the original problem (P). Further research based on this idea can be seen in Refs. [6]-[13].

In 1994, Cawood and Kostreva[3] generalized the idea of Pironneau-Polak's MFD[14] and proposed a norm-relaxed MFD algorithm for the problem (P) with $I_2 = \emptyset$, i.e., with only inequality constraints. At each iteration, the feasible direction of descent in [3] is generated by solving a direction finding subproblem (DFS) as follows

$$\begin{array}{ll} \min & z + \dfrac{r}{2} d^{\mathrm{T}} B_k d \\ \text{s.t.} & \nabla f(x^k)^{\mathrm{T}} d \leqslant z, \\ & g_j(x^k) + \nabla g_j(x^k)^{\mathrm{T}} d \leqslant z, \quad j \in I. \end{array}$$

where B_k is a positive definite matrix, r is a constant and x^k is a current feasible iteration point.

Then in 1999, by introducing some parameters, Chen and Kostreva[4] proposed a so-called generalized norm-relaxed MFD, which can improve the numerical effect. With positive constants $\gamma_0, \gamma_j \ (j \in I)$, they considered the following DFS

$$\begin{array}{ll} \min & z + \dfrac{1}{2} d^{\mathrm{T}} B_k d \\ \text{s.t.} & \nabla f(x^k)^{\mathrm{T}} d \leqslant \gamma_0 z, \\ & g_j(x^k) + \nabla g_j(x^k)^{\mathrm{T}} d \leqslant \gamma_j z, \ j \in I. \end{array}$$

The normal-relaxed method is further improved and extended by Jian et al in [12, 15]. To obtain superlinear convergence, Kostreva and Chen[16], Lawrence and Tits[17], Zhu and

Zhang[18] further studied the norm-relaxed MFD. However, to obtain superlinear convergence, these methods depend on the strict complementarity assumption, additionally, these algorithms can't solve directly the optimization problem (P) with $I_2 \neq \emptyset$.

Recently, with the help of the idea of the strongly sub-feasible direction and some suitable technique yielding a high-order correction direction used to avoid the Maratos effect[19], the norm-relaxed method is further researched by Jian et al, in [20-21], all of them deal with the optimization problem (P) with only inequality constraints. By using an ε-active constraint set technique, Ref. [21] constructs the following DFS to generate a master search direction:

$$\text{DFS}\qquad \begin{array}{ll} \min\limits_{(z,d)\in R^{n+1}} & \gamma_0 z+\dfrac{1}{2}d^{\mathrm{T}}B_k d \\ \text{s.t.} & \nabla f(x^k)^{\mathrm{T}}d\leqslant \gamma_0 z+\gamma\varphi(x^k)^{\sigma}, \\ & g_j(x^k)+\nabla g_j(x^k)^{\mathrm{T}}d\leqslant \gamma_j\eta_k z,\quad j\in I_k^-\overset{\text{Def}}{=}I^-(x^k,\varepsilon), \\ & g_j(x^k)+\nabla g_j(x^k)^{\mathrm{T}}d\leqslant \gamma_j\eta_k z+\varphi(x^k),\quad j\in I_k^+\overset{\text{Def}}{=}I^+(x^k,\varepsilon), \end{array}$$

where η_k is a positive parameter associated with x^k, γ, σ, γ_j are all positive constant parameters, and ε-active constraint sets are defined by

$$I_k^-=\{j\in I_1:\ -\varepsilon\leqslant g_j(x^k)\leqslant 0\}$$
$$I_k^+=\{j\in I_1:\ 0<g_j(x^k),\ 0\leqslant g_j(x^k)-\max_{i\in I_1}\{0,g_i(x^k)\}\}.$$

Then a system of linear equations (SLE) as

$$V_k\begin{pmatrix} d\\ h\end{pmatrix}=\begin{pmatrix} B_k & N_k\\ N_k^T & -G^k\end{pmatrix}\begin{pmatrix} d\\ h\end{pmatrix}=\begin{pmatrix} 0\\ -\max\{\|d^k\|^{\tau},|\ \eta_k^{\nu}z_k\ |\ \|d^k\|\}e_{I_k}-\widetilde{g}^k\end{pmatrix}$$

is solved to yield a high-order updated direction where $N_k=(\nabla g_j(x^k),\ j\in I_k^-\cup I_k^+)$, G^k is a suitable diagonal matrix and $\tilde{g}^k$ is a suitable vector. The high-order correction technique here is much different from the ones used in Refs. [16-18,21], and the numerical effect is further improved.

In this paper, based on the auxiliary problem (P_c), we further extend and improve the norm-relaxed algorithm in [21] such that it can not only deal with general constrained optimization problems, but also improve some characters of the algorithm in [21], as a result, we propose a new norm-relaxed algorithm for general constrained optimization problem (P). To reduce the cost of computation, we use the technique of identification function and working set to construct the DFS and the high-order correction direction. The working set technique can be seen in [22-24], which has been proved to be effectively. Similar to the rules in [23], in this paper, combining with the information used in updating the penalty parameter c, we derive a simple form of working set, and we will show it is equivalent to the active set of (P).

The main features of the algorithm can be summarized as follows:

• the penalty parameter c is adjusted automatically only finite number of times;

• the cost of computation is reduced by using the technique of working set;

• an improved direction for the problem (0.2) is obtained by solving only one norm-relaxed QP subproblem, and a high-order correction direction avoiding the Maratos effect is obtained by solving one SLE;

• possesses global and superlinear convergence under some suitable assumptions without the strictly complementarity.

1 Description of algorithm

For the sake of simplicity, we denote and use the following notations

$$\begin{gathered} X^+=\{x\in R^n:\ g_i(x)\leqslant 0,\ i\in I\},\\ I_1(x)=\{j\in I_1:\ g_j(x)=0\},\ I(x)=I_1(x)\cup I_2,\ I_0(x)=\{j\in I:\ g_j(x)=0\}. \end{gathered} \tag{1.1}$$

First, assume that the following basic assumptions hold in the paper:

Assumption A1 The functions f and g_j $(j\in I)$ are all continuously differentiable in X^+

Assumption A2 The gradient vectors $\{\nabla g_j(x), j\in I(x)\}$ are linearly independent for any $x\in X^+$.

Lemma 1.1 Suppose that Assumptions A1 and A2 hold. Then for any $x^k\in X^+$, the matrix $(\widetilde{N}_k^{\mathrm{T}}\widetilde{N}_k+D(x^k))$ is nonsingular and positive definite, where

$$\widetilde{N}_k=(\nabla g_j(x^k), j\in I),\ D(x^k)=\mathrm{diag}(D_j^k, j\in I),\ D_j^k=\begin{cases} |g_j(x^k)|, & \text{if } j\in I_1;\\ 0, & \text{if } j\in I_2. \end{cases} \tag{1.2}$$

Using Assumption A2, the proof is elementary and is omitted here.

According to the lemma above, for a current iteration point $x^k\in X^+$, we use multiplier vector

$$\pi(x^k)=(\pi_j(x^k), j\in I)=-(\widetilde{N}_k^{\mathrm{T}}\widetilde{N}_k+D(x^k))^{-1}\widetilde{N}_k^{\mathrm{T}}\nabla f(x^k) \tag{1.3}$$

to update the penalty parameter c in (0.2) (the detail can be seen in Step 2 of the algorithm below). Obviously, (1.3) is equivalent to the following SLE in variable d:

$$(\widetilde{N}_k^{\mathrm{T}}\widetilde{N}_k+D(x^k))d=-\widetilde{N}_k^{\mathrm{T}}\nabla f(x^k). \tag{1.4}$$

From Lemma 2.2 in [12], we have the following conclusion.

Lemma 1.2 (i) Let $x^k\in R^n$. If parameter $c>|\ \pi_j(x^k)\ |,\ \forall j\in I_2$, then (x^k,λ^k) is a KKT pair of the original problem (0.1) if and only if $(x^k,\bar{\mu}^k)$ is a KKT pair of the problem

(0.2), where λ^k and $\bar{\mu}^k$ satisfy

$$\lambda_j^k = \bar{\mu}_j^k,\ j \in I_1;\ \lambda_j^k = \bar{\mu}_j^k - c,\ j \in I_2. \tag{1.5}$$

(ii) If x^k is a KKT point of (0.1), then $\pi(x^k)$ is the unique corresponding KKT multiplier.

Now, define function $\Phi : R^{n+m} \to R^{n+m}$:

$$\Phi(x,\lambda) = \begin{pmatrix} \nabla_x L(x,\lambda) \\ \min\{-g_{I_1}(x),\ \lambda_{I_1}\} \\ g_{I_2}(x) \end{pmatrix}$$

with Lagrange function $L(x,\lambda) = f(x) + \sum\limits_{i\in I} \lambda_i g_i(x)$, then from Theorem 4.3 of [23], we know function $\rho : R^{m+n} \to R$:

$$\rho(x,\lambda) = \|\Phi(x,\lambda)\|^{\frac{1}{2}} \tag{1.6}$$

is an optimal identification function of (P), i.e., (x,λ) is a KKT pair of (P) if and only if $\rho(x,\lambda) = 0$. So for the current iteration point x^k, if one denotes

$$I(x,\lambda) = \{i \in I_1 :\ g_i(x) + \rho(x,\lambda) \geqslant 0\},$$

then $I(x,\lambda) \equiv I_1(x^*)$ when (x,λ) is sufficiently close to a KKT pair (x^*,λ^*) of (P) if the second-order sufficient conditions and the Mangasarian-Fromovitz constraint qualification (MFCQ) hold at (x^*,λ^*) (see [22]). Based on Lemma 1.2 and the construction of $\pi(x)$, which is an estimate of the KKT multiplier vector λ, we introduce and use the following working set

$$I_{1k} = \{i \in I_1 : g_j(x^k) + \rho(x^k, \pi(x^k)) \geqslant 0\},\ I_k = I_{1k} \cup I_2. \tag{1.7}$$

We can also prove that $I_k \equiv I(x^*)$ when x^k is sufficiently close to a KKT point x^* (see Lemma 3.1).

From Lemma 1.2 above, we can see that solving the original problem (0.1) can be transformed to solve a sequence optimization (0.2) of problems with inequality constraints. Motivated by this and the important property of working set, we introduce an effective feasible direction method for the problems (0.2), then for the original problem (0.1) indirectly.

For a given c_k and iteration point $x^k \in X^+$ as well as a symmetric positive definite matrix B_k, we use the following DFS to yield our master search direction d^k

$$\text{DFS}_{c_k} \qquad \begin{array}{ll} \min\limits_{(z,d)\in R^{n+1}} & \gamma_0 z + \dfrac{1}{2} d^{\mathrm{T}} B_k d \\ \text{s.t.} & \nabla F_{c_k}(x^k)^{\mathrm{T}} d \leqslant \gamma_0 z, \\ & g_j(x^k) + \nabla g_j(x^k)^{\mathrm{T}} d \leqslant \gamma_j \eta_k z, \quad j \in I_k, \end{array} \tag{1.8}$$

where η_k is a positive parameter associated with x^k, and γ_j $(j \in \{0\} \cup I_k)$ are all positive constant parameters. The parameter η_k accelerates the convergence rate and attaches much importance in the proof of global and superlinear convergence of our algorithm.

Obviously, DFSc_k is equivalent to the following unconstrained strictly convex program

$$\min_{d\in R^n}\left\{\frac{1}{2}d^{\mathrm T}B_kd+\max_{j\in I_k}\left\{\frac{1}{\gamma_0}\nabla F_{c_k}(x^k)^{\mathrm T}d;\frac{1}{\gamma_j\eta_k}(g_j(x^k)+\nabla g_j(x^k)^{\mathrm T}d)\right\}\right\}.$$

Therefore, it has an unique optimal solution d^k. Moreover, since (1.8) is a convex program with linear constraints, (z_k, d^k) is an optimal solution of (1.8) if and only if it is a KKT point of (1.8) (the detail can be seen in Lemma 2.1 of [20]).

Suppose that $(z_k, d^k, \mu_0^k, \mu_{I_k}^k)$ is a KKT pair of DFSc_k. Then the corresponding KKT conditions of (1.8) can be expressed as

$$\gamma_0\mu_0^k+\eta_k\sum_{j\in I_k}\gamma_j\mu_j^k=\gamma_0, \tag{1.9}$$

$$B_kd^k+\mu_0^k\nabla F_{c_k}(x^k)+\sum_{j\in I_k}\mu_j^k\nabla g_j(x^k)=0, \tag{1.10}$$

$$0\leqslant\mu_0^k\bot\left(-\nabla F_{c_k}(x^k)^{\mathrm T}d^k+\gamma_0z_k\right)\geqslant 0, \tag{1.11}$$

$$0\leqslant\mu_j^k\bot\left(-g_j(x^k)-\nabla g_j(x^k)^{\mathrm T}d^k+\gamma_j\eta_kz_k\right)\geqslant 0,\ j\in I_k, \tag{1.12}$$

where the symbol $x \perp y$ means $x^T y = 0$. In the case of $\mu_0^k \neq 0$, we define multiplier

$$\bar\mu_{I_k}^k=(\bar\mu_j^k=\mu_j^k/\mu_0^k, j\in I_k),\ \bar\mu^k=(\bar\mu_{I_k}^k, 0_{I\setminus I_k}). \tag{1.13}$$

Lemma 1.3 Let (z_k, d^k) be an optimal solution to the DFS (1.8) and suppose that Assumptions A1-A2 hold as well as B_k is a symmetric positive definite matrix. Then

(i) $\gamma_0 z_k + \frac{1}{2}(d^k)^{\mathrm T}B_k d^k \leqslant 0,\ z_k \leqslant 0$;

(ii) $z_k = 0 \Longleftrightarrow d^k = 0 \Longleftrightarrow x^k$ is a KKT point for (0.2), and $\bar\mu^k$ defined by (1.13) is the associated KKT multiplier;

(iii) if $d^k \neq 0$, then $\nabla g_j(x^k)^{\mathrm T}d^k \leqslant \gamma_j\eta_k z_k < 0,\ \forall j \in I_0(x^k)$, and $\nabla F_{c_k}(x^k)^{\mathrm T}d^k \leqslant \gamma_0 z_k < 0$. So d^k is a feasible direction of decent of (P_{c_k}) at x^k.

The proof of the results above is similar to the one of Lemma 2.3 in [12] associated with the case of $x^k \in X^+$, i.e., $\varphi(x^k) \stackrel{\mathrm{Def}}{=} \max\limits_{j\in I}\{0, g_j(x^k)\} = 0$.

Remark 1 If parameter $c_k > |\pi_j(x^k)|, j \in I_2$, then from Lemma 1.2 and Lemma 1.3 (ii), we know that $d^k = 0$ if and only if (x^k, λ^k) is a KKT pair of the original problem (0.1), where λ^k satisfies (1.5) and $\bar\mu^k$ is defined by (1.13).

It is well known that to overcome the Maratos effect, a suitable high-order updated direction of d^k must be adopted. In this paper, based on the technique of working set I_k, similar to [21], we introduce the following SLE

$$V_k\begin{pmatrix} d \\ h \end{pmatrix}=\begin{pmatrix} 0 \\ -\max\{\|d^k\|^\tau, |\ \eta_k^\nu z_k\ |\ \|d^k\|\}\varpi_{I_k}-\widetilde{g}^k \end{pmatrix} \tag{1.14}$$

to yield a high-order updated direction, where $\varpi_{I_k}=(1,\dots,1)^T\in R^{|I_k|}$, $\tau\in(2,3)$, $\nu\in(0,1)$ and

$$V_k=\begin{pmatrix} B_k & N_k \\ N_k^T & -G_k \end{pmatrix},\ N_k=(\nabla g_j(x^k),j\in I_k),\ G_k=\mathrm{diag}(G_j^k,j\in I_k),\ \widetilde{g}^k=(\widetilde{g}_j^k,j\in I_k), \tag{1.15}$$

$$G_j^k=|\ g_j(x^k)\ |\ (|\ g_j(x^k)+\nabla g_j(x^k)^{\mathrm T}d^k-\gamma_j\eta_k z_k\ |+|\ \eta_k z_k\ |+\|d^k\|), \tag{1.16}$$

$$\widetilde{g}_j^k=g_j(x^k+d^k)-g_j(x^k)-\nabla g_j(x^k)^{\mathrm T}d^k+\gamma_j\eta_k z_k. \tag{1.17}$$

Lemma 1.4 Suppose that Assumptions A1–A2 hold and matrix B_k is positive definite. Then the matrix V_k defined by (1.15) is nonsingular.

Proof Obviously, from the definition of G_k, we know that $G_j^k\geqslant 0$, $\forall j\in I_k$, and column vectors

$$\{\nabla g_j(x^k)|\ j:\ G_j^k=0\}=\{\nabla g_j(x^k)|\ j\in I_0(x^k)\subseteq I(x^k)\}$$

are linearly independent, So, by Corollary 1.1.9(3) in [25], one knows the conclusion follows.

Based on the master direction d^k and the high-order correction $\widetilde{d}^k$ yielded by (1.14), we state our algorithm as follows.

Algorithm

Step 0. (*Initialization*) Let parameters $a_0>0$, $\alpha\in(0,0.5)$, $\beta,\nu\in(0,1)$, $\tau\in(2,3)$, $\xi,\zeta,\eta_0,c_{-1},\iota,r,\gamma,\gamma_j>0$, $j=0,1,\dots,m$, $\delta_1,\delta_2\geqslant 0,\delta_1+\delta_2>0$, and choose a starting point $x^0\in X^+$ and an initial symmetric positive definite matrix B_0, set $\eta_0=a_0$ and $k:=0$.

Step 1. (*Generating the working set*) Compute $\rho(x^k,\pi(x^k))$ by (1.3) and (1.6) and the working set I_k by (1.7).

Step 2. (*Adjusting parameter c_k*) If $I_2=\emptyset$, go to Step 3; otherwise, compute c_k by

$$c_k=\begin{cases}\max\{s_k,c_{k-1}+r\}, & \text{if } s_k>c_{k-1};\\ c_{k-1}, & \text{if } s_k\leqslant c_{k-1},\end{cases}\quad s_k=\max\{|\pi_j(x^k)|,j\in I_2\}+\iota. \tag{1.18}$$

Step 3. (*Generating master search direction*) Solve the DFS (1.8) to obtain an optimal solution (z_k,d^k) with multiplier $(\mu_0^k,\mu_{I_k}^k)$. If $d^k=0$, then stop.

Step 4. (*Generating correction direction*) Compute the updated direction $\widetilde{d}^k$ by solving the linear system (1.14) with a solution $(\widetilde{d}^k,h^k)$. If $\|\widetilde{d}^k\|>\|d^k\|$, then reset $\widetilde{d}^k=0$.

Step 5. (*Doing arc search*) Compute the step size t_k, which is the first value of t in the sequence $\{1, \beta, \beta^2, \beta^3, \ldots\}$ that satisfies the inequalities:

$$F_{c_k}(x^k + td^k + t^2\widetilde{d}^k) \leqslant F_{c_k}(x^k) + \alpha t \nabla F_{c_k}(x^k)^{\mathrm{T}} d^k, \tag{1.19}$$

$$g_j(x^k + td^k + t^2\widetilde{d}^k) \leqslant 0, \quad j \in I. \tag{1.20}$$

Step 6. (*Updating*) Compute a new symmetric positive definite matrix B_{k+1} by a suitable technique and η_{k+1} by $\eta_{k+1} = \min\{a_0, \delta_1\|d^k\|^\xi + \delta_2 \mid z_k \mid^\varsigma\}$. Set $x^{k+1} = x^k + t_k d^k + t_k^2\widetilde{d}^k$, $k := k+1$, and go back to Step 1.

To show that the proposed algorithm is well defined, we give the following lemma, and taking into account Lemma 1.3 (iii), its proof is similar to the one of Lemma 2.4 in [20] associated with the case of $x^k \in X^+$, i.e., $\varphi(x^k) = 0$.

Lemma 1.5 The inequalities (1.19)–(1.20) hold for $t > 0$ small enough, so the line search in Step 5 is well defined.

Remark 2 The cost of computation in Step 3 and Step 4 is reduced by adopting the technique of working set.

Remark 3 By the construction of c_k, after finite iterations, there exists a constant c such that $c_k \equiv c$ (see Lemma 2.1 below).

2 Global convergence

If the proposed algorithm stops at x^k, then from Step 3 and the definition of c_k in Step 2, we know $d^k = 0$ and $c_k > |\pi_j(x^k|$, $\forall j \in I_2$. According to Lemma 1.2 and Lemma 1.3 (ii), one knows that x^k is a KKT point for (1.1). In this section, we assume that the proposed algorithm yields an infinite iteration sequence $\{x^k\}$, and will show that it is globally convergent, namely, every accumulation x^* of $\{x^k\}$ is a KKT point of (0.1). For this goal, the following basic assumption is necessary.

Assumption A3 The sequence $\{x^k\}$ yielded by the proposed algorithm is bounded and the sequence $\{B_k\}$ of matrices is uniformly positive definite, i.e., there exist two positive constants a and b such that

$$a\|d\|^2 \leqslant d^{\mathrm{T}} B_k d \leqslant b\|d\|^2, \quad \forall d \in R^n, \ \forall k.$$

Define the active constrain set of (1.8) by

$$L_k = \{j \in I_k : g_j(x^k) + \nabla g_j(x^k)^{\mathrm{T}} d^k = \gamma_j \eta_k z_k\}. \tag{2.1}$$

Lemma 2.1 Suppose that Assumptions A1–A2 hold, and the sequence $\{x^k\}$ is bounded. Then there exists a positive integer k_0, such that $c_k = c_{k_0}$, $\forall\ k \geqslant k_0$.

The proof is similar to the fashion of Lemma 2.1 in [12].

According to the lemma above, in the remainder of this paper, we always assume that $c_k \equiv c$ for all k for k large enough .

Lemma 2.2 Suppose that Assumptions A1–A3 hold. Then

(i) the sequences $\{z_k\}$, $\{d^k\}$ and $\{\widetilde{d}^k\}$ are all bounded;

(ii) the KKT multiplier sequences $\{\mu_0^k\}$ and $\{\mu_{I_k}^k\}$ are all bounded; and

(iii) there exists a constant $\bar{c}$ such that $\| V_k^{-1} \| \leqslant \bar{c}$ holds for k large enough.

Proof The proof of conclusions (i) and (ii) are similar to the ones of Lemma 3.2(i) and Lemma 3.3 in [12] associated with the case of $x^k \in X^+$, i.e., $\varphi(x^k) = 0$, respectively. Now we give the proof of conclusion (iii). By contradiction, suppose that there exists an infinite index set K such that $\| V_k^{-1} \| \xrightarrow{K} \infty$. Then, taking into account Assumption A3 and the boundedness of $\{(z_k, d^k, \eta_k)\}$, we can assume without loss of generality, choosing an infinite subset of K if necessary, that

$$B_k \to B_*,\ d^k \to d^*,\ z_k \to z_*,\ \eta_k \to \eta_*\ \ I_k \equiv \widetilde{I},\ V_k \to V_* = \begin{pmatrix} B_* & N_* \\ N_*^T & -G_* \end{pmatrix},\ k \in K,$$

where N_* and G_* is the limits of $\{N_k\}_K$ and $\{G_k\}_K$, respectively, namely,

$$N_* = (\nabla g_j(x^*),\ j \in \widetilde{I}),\quad G_* = \mathrm{diag}(G_j^*,\ j \in \widetilde{I})$$

with

$$G_j^* = | g_j(x^*) | (| g_j(x^*) + \nabla g_j(x^*)^T d^* - \gamma_j \eta_* z_* | + | \eta_* z_* | + \|d^*\|).$$

Obviously, $G_j^* \geqslant 0, \forall j \in \widetilde{I}$ and $G_j^* > 0$ for all $j \in \widetilde{I} \backslash I_0(x^*)$. So, $H := B_*,\ A := N_*$ and $D := G_*$ satisfy the requests in Corollary 1.1.9 in [25], therefore, by this corollary, we know V_* is nonsingular. Thus, $\| V_k^{-1} \| \xrightarrow{K} \| V_*^{-1} \| < \infty$, which contradicts the assumption.

Lemma 2.3 Suppose that Assumptions A1–A3 hold. If an infinite index set K satisfies $\lim\limits_{k\in K} x^k = x^*$ and $\lim\limits_{k\in K} d^k = 0$, then x^* is a KKT point both for the smei-penalty problem (0.2) and the original problem (0.1).

Proof By Lemma 2.2, without loss of generality, we can assume, choosing an infinite subset of K if necessary, that

$$x^k \to x^*,\quad B_k \to B_*,\quad \eta_k \to \eta_*,\quad I_k \equiv \widetilde{I},\quad L_k \equiv L,\ \mu_0^k \to \mu_0^*,\ \mu_{I_k}^k \to \mu_{\widetilde{I}}^*,\ k \in K. \tag{2.2}$$

In view of $\nabla F_c(x^k)^T d^k \leqslant \gamma_0 z_k \leqslant 0$ and $\lim\limits_{k\in K} d^k = 0$ as well as $\lim\limits_{k\in K} x^k = x^*$, it follows that $\lim\limits_{k\in K} z_k = 0$. This together with the definition of L_k in (2.1) shows that $L \subseteq I_0(x^*)$. Therefore, passing to the limit $k \in K$ and $k \to \infty$ in formulas (1.9)–(1.12), one gets

$$\gamma_0 \mu_0^* + \eta_* \sum_{j\in L} \gamma_j \mu_j^* = \gamma_0,\ \mu_0^* \nabla F_c(x^*) + \sum_{j\in L} \mu_j^* \nabla g_j(x^*) = 0,\ \mu_0^* \geqslant 0,\ \mu_j^* \geqslant 0,\ g_j(x^*) = 0,\ j \in L. \tag{2.3}$$

Furthermore, from the relations above and A2, it is easy to see that $\mu_0^* \neq 0$, so $\mu_0^* > 0$. Therefore, we can conclude that x^* is a KKT point of (0.2) with the corresponding multiplier $\bar{\mu}^* \overset{\text{Def}}{=} (\bar{\mu}_L^* = \mu_L^*/\mu_0^*, 0_{I\setminus L})$.

Finally, taking into account the construction of c_k in Step 2 and Lemma 3.1, we obtain

$$c \geqslant \lim_{k\in K} \pi_j(x^k) + \iota = \pi_j(x^*) + \iota > \pi_j(x^*), \forall j \in I_2.$$

Hence, one can conclude that x^* is a KKT point of (1.1) from Lemma 1.2.

Theorem 2.1 Let x^* be an accumulation point of the sequence $\{x^k\}$ yielded by the proposed algorithm. And suppose that Assumptions A1–A3 are satisfied. Then x^* is a KKT point both of the original problem (0.1) and the auxiliary problem (0.2) with c. Therefore, the proposed algorithm is globally convergent.

Proof In view of Lemma 2.2, we can assume that there exists an infinite index set K such that (2.2) holds and $d^k \to d^*,\ \ z_k \to z_*,\ k \in K$. By contradiction, we assume that x^* is not a KKT point of (0.1), so $\rho(x^*, \pi(x^*)) > 0$. Furthermore, by Lemma 2.3, one knows that there exists an infinite subset $K^{'} \subseteq K$ and a constant $\delta > 0$ such that $\|d^k\| \geqslant \delta,\ k \in K^{'}$. Thus, one has the following relations from Lemma 1.3 (i) and Assumption A3,

$$\gamma_0 z_k \leqslant -\frac{1}{2}(d^k)^T B_k d^k \leqslant -\frac{1}{2}a\|d^k\|^2 \leqslant -\frac{1}{2}a\delta^2, \text{ i.e., } z_k \leqslant -\frac{1}{2\gamma_0}a\delta^2, \quad k \in K^{'}. \tag{2.4}$$

Two cases of $\eta_* = 0$ and $\eta_* > 0$ are considered in the remainder discussion.

Case I Suppose that $\eta_* = 0$. Then by the construction of η_k in Step 6, one has

$$\eta_k = \delta_1\|d^{k-1}\|^\xi + \delta_2|z_{k-1}|^\zeta \to 0,\ k \in K^{'}.$$

Therefore, $\delta_1\|d^{k-1}\|^\xi \to 0$ and $\delta_2|z_{k-1}|^\zeta \to 0$. In view of $\delta_2 z_{k-1} \leqslant -\frac{a\delta_2}{2\gamma_0}\|d^{k-1}\|^2$, thus $\delta_2 d^{k-1} \to 0$. So we have $(\delta_1+\delta_2)d^{k-1} \to 0$, this together with $\delta_1+\delta_2 > 0$ shows that $\lim\limits_{k\in K'} \|d^{k-1}\| = 0$. Therefore, according to Steps 4-6, one has

$$\lim_{k\in K'} \|x^k - x^{k-1}\| \leqslant \lim_{k\in K'} (t_{k-1}\|d^{k-1}\| + t_{k-1}^2\|\widetilde{d}^{k-1}\|) \leqslant \lim_{k\in K'} (2\|d^{k-1}\|) = 0.$$

So $\lim\limits_{k\in K'} x^{k-1} = x^*$ follows from $\lim\limits_{k\in K'} x^k = x^*$. Let $\overline{K} = \{k-1,\ k \in K^{'}\}$. Then the discussion above implies that $\lim\limits_{k\in \bar{K}} x^k = x^*$ and $\lim\limits_{k\in \bar{K}} d^k = 0$. Therefore, x^* is a KKT point for (1.1) by Lemma 2.3, which is a contradiction.

Case II Assume that $\eta_* > 0$. Then for k large enough, one has

$$\eta_k \geqslant \frac{\eta_*}{2}, \text{ for } k \in K^{'} \text{ large enough.} \tag{2.5}$$

First, we will show that $\bar{t} \overset{\text{Def}}{=} \inf\{t_k,\ k \in K^{'}\} > 0$, i.e., the relations (1.19) and (1.20) hold for $t > 0$ small enough and $k \in K^{'}$ large enough.

Analyze the inequality (1.19): using Taylor expansion, Lemma 2.2(i) and (1.8) as well as (2.4), one has

$$\begin{aligned} F_c(x^k+td^k+t^2\widetilde{d}^k)-F(x^k)-\alpha t\nabla F_c(x^k)^{\mathrm{T}}d^k &= (1-\alpha)t\nabla F_c(x^k)^{\mathrm{T}}d^k+o(t) \\ &\leqslant (1-\alpha)t\gamma_0 z_k+o(t) \\ &\leqslant -\frac{(1-\alpha)t\delta^2}{2}+o(t). \end{aligned}$$

Therefore, the inequality (1.19) holds for $k\in K^{'}$ large enough and $t>0$ sufficiently small.

Analyze the inequalities (1.20): For $j\notin I_0(x^*)$, i.e., $g_j(x^*)<0$, since g_j is continuous and $\{(d^k,\widetilde{d}^k)\}$ is bounded, we have $g_j(x^k+td^k+t^2\widetilde{d}^k)\leqslant 0$ for $k\in K^{'}$ large enough and $t>0$ small enough.

For $j\in I_0(x^*)$, i.e., $g_j(x^*)=0$. we have $j\in I_k$ since $g_j(x^k)+\rho(x^k,\pi(x^k))\to g_j(x^*)+\rho(x^*,\pi(x^*))>0$. Therefore, in view of (1.8), (2.4) and (2.5) as well as $g_j(x^k)\leqslant 0$, one gets

$$\begin{aligned} g_j(x^k+td^k+t^2\widetilde{d}^k) &= g_j(x^k)+t\nabla g_j(x^k)^{\mathrm{T}}d^k+o(t) \\ &\leqslant g_j(x^k)+t\gamma_j\eta_k z_k-tg_j(x^k)+o(t) \\ &= (1-t)g_j(x^k)+t\gamma_j\eta_k z_k+o(t) \\ &\leqslant t\gamma_j\eta_k z_k+o(t) \\ &\leqslant -\frac{\gamma_j\eta_* a\delta^2}{4\gamma_0}t+o(t) \\ &\leqslant 0. \end{aligned}$$

So, (1.20) holds for $k\in K^{'}$ large enough and $t>0$ small enough.

Second, use $t_k\geqslant \overline{t}>0$ $(k\in K^{'})$ to bring a contradiction. From (1.19) and (2.4), it follows that

$$\begin{cases} F_c(x^{k+1})-F_c(x^k)\leqslant \alpha t_k\nabla F_c(x^k)^{\mathrm{T}}d^k\leqslant \alpha\gamma_0 z_k t_k<0, \quad \forall k; \\ F_c(x^{k+1})-F_c(x^k)\leqslant \alpha t_k\nabla F_c(x^k)^{\mathrm{T}}d^k\leqslant \alpha\gamma_0 z_k t_k\leqslant -\dfrac{1}{2}\alpha a\delta^2\overline{t}, \quad \forall k\in K^{'}. \end{cases} \tag{2.6}$$

Therefore, $\{F_c(x^k)\}$ is monotone decreasing, together with $\lim\limits_{k\in K^{'}}F_c(x^k)=F_c(x^*)$, one has $\lim\limits_{k\to\infty}F_c(x^k)=F_c(x^*)$. On the other hand, passing to the limit $k\in K^{'}$and $k\to\infty$ in the second inequality of (2.6), we bring a contradiction. Hence, x^* is a KKT point of (0.1).

Lastly, in view of Lemma 1.2 and $c>|\pi_j(x^*)|$, $\forall j\in I_2$, it further follows that x^* is a KKT point of (0.2).

3 Strong and superlinear convergence

In this section, we first prove the strong convergence of the proposed algorithm. Then under some mild assumptions without the strict complementarity, we discuss the superlinear convergence. In the remainder analysis, the two following additional second-order assumptions are necessary.

Assumption A4 (i) The functions $f(x)$ and $g_j(x)$ $(j \in I)$ are all twice continuously differentiable in X^+; and

(ii) the sequence $\{x^k\}$ generated by the proposed algorithm possesses an accumulation point x^* along with the KKT multiplier λ^* for (P), such that the strongly second-order sufficient condition (SOSC) is satisfied, i.e.,

$$d^T \nabla^2_{xx} L(x^*, \lambda^*) d > 0, \quad \forall d \in \Omega \overset{\text{Def}}{=} \{d \in R^n: \ d \neq 0, \ \nabla g_j(x^*)^T d = 0, \ j \in I_*^+ \cup I_2\},$$

where $I_*^+ = \{j \in I_1 : \lambda_j^* > 0\}$.

Remark 4 From Lemma 1.2, we know that the KKT multiplier $\lambda^* = \pi^* \overset{\text{Def}}{=} \pi(x^*)$, further, $(x^*, \bar{\mu}^*)$ with

$$\bar{\mu}^* = \lambda_j^*, \ j \in I_1; \quad \bar{\mu}_j^* = \lambda_j^* + c, \ j \in I_2 \tag{3.1}$$

is a KKT pair of (0.2). On the other hand, the Lagrange function associated with (P_c) is given by

$$L_c(x, \bar{\mu}^*) = F_c(x) + \sum_{j \in I} \bar{\mu}_j^* g_j(x) = L(x, \lambda^*), \ \nabla^2_{xx} L_c(x^*, \bar{\mu}^*) = \nabla^2_{xx} L(x^*, \lambda^*).$$

So, in view of $\bar{\mu}_j^* > 0, \ \forall j \in I_2$, the SOSC is satisfied for the inequality constrained optimization (P_c) at the KKT pair $(x^*, \bar{\mu}^*)$.

Assumption A5 Suppose that

$$\|(\nabla^2_{xx} L(x^*, \lambda^*) - B_k) d^k\| = \|(\nabla^2_{xx} L_c(x^*, \bar{\mu}^*) - B_k) d^k\| = o(\|d^k\|).$$

Remark 5 By Lemma 3.1 below, one can see that, Assumption A5 holds if and only if

$$\|(\nabla^2_{xx} L_c(x^k, \mu^k/\mu_0^k) - B_k) d^k\| = o(\|d^k\|).$$

Theorem 3.1 Suppose that Assumptions A2–A4 are all satisfied and x^* is the limit point stated in A4. Then

(i) for each index set K such that $x^k \xrightarrow{K} x^*$, it follows that $(z_k, d^k) \xrightarrow{K} (0, 0)$; and

(ii) $\lim\limits_{k \to \infty} x^k = x^*$, $\lim\limits_{k \to \infty} z_k = 0$ and $\lim\limits_{k \to \infty} d^k = \lim\limits_{k \to \infty} \widetilde{d}^k = 0$, so the proposed algorithm is strongly convergent.

Proof (i) Since the whole sequence $\{(z_k, d^k)\}$ is bounded, it is sufficient to show that each limit point (z_*, d^*) of $\{(z_k, d^k)\}_K$ must equal $(0, 0)$. For the given limit point (z_*, d^*), in view of the boundedness of $\{\eta_k\}$, there exists a subset $K' \subseteq K$ such that $(z_k, d^k) \xrightarrow{K'} (z_*, d^*)$ and $\eta_k \xrightarrow{K'} \eta_*$.

Taking into account $(x^k, \pi(x^k)) \xrightarrow{K} (x^*, \lambda^* = \pi(x^*))$, under A2 and A4, from Ref. [22], one has $I_{1k} \equiv I_1(x^*)$ and $I_k = I_{1k} \cup I_2 = I_1(x^*) \cup I_2 = I(x^*)$ for $k \in K$ large enough. Therefore, by passing the limit in the constraints of (1.8) for $k \in K'$, we have

$$\nabla F_c(x^*)^{\mathrm{T}} d^* \leqslant \gamma_0 z_*, \quad \nabla g_j(x^*)^{\mathrm{T}} d^* \leqslant \gamma_j \eta_* z_* \leqslant 0, \quad j \in I_k = I(x^*). \tag{3.2}$$

On the other hand, in view of $(x^*, \bar{\mu}^*)$ is a KKT pair of (0.2), one has

$$\nabla F_c(x^*) + \sum_{j \in I(x^*)} \bar{\mu}_j^* \nabla g_j(x^*) = 0,\ \bar{\mu}_j^* \geqslant 0,\ j \in I(x^*). \tag{3.3}$$

Hence, it is easy to get $z_* = 0$ from (3.2) and (3.3). Furthermore, $d^* = 0$ follows form Lemma 1.3(i), A3 and $z_* = 0$.

(ii) First, Assumptions A2 and A4 ensure that x^* is an isolated KKT point of (P) (see, Corollary 1.4.3 in [25]), which together with Theorem 2.1 shows that x^* is an isolated limit point of $\{x^k\}$. Second, for each index set K such that $x^k \xrightarrow{K} x^*$, from part (i) and Step 4, one has

$$\lim_{k \in K} \|x^{k+1} - x^k\| \leqslant \lim_{k \in K} (t_k \|d^k\| + t_k^2 \|\widetilde{d}^k\|) \leqslant 2 \lim_{k \in K} \|d^k\| = 0.$$

Therefore, by Proposition 7 in [26] or Theorem 1.1.7 in [25], we can conclude $\lim\limits_{k \to \infty} x^k = x^*$.

Finally, the rest conclusions in part (ii) follow immediately from $\lim\limits_{k \to \infty} x^k = x^*$ and part (i) as well as $\|\widetilde{d}^k\| \leqslant \|d^k\|$.

Lemma 3.1 Suppose that Assumptions A2–A4 hold. Then

$$I_2 \cup I_*^+ \subseteq L_k \subseteq I_0(x^*) = I(x^*) \equiv I_k \ (k \text{ large enough}),$$
$$\lim_{k \to \infty} (\eta_k, \mu_0^k, \mu_{I_k}^k) = (0, 1, \bar{\mu}_{I(x^*)}^*).$$

Proof Obviously, $I_0(x^*) = I(x^*)$ follows from $x^* \in X$. We first show $I(x^*) \equiv I_k$ for k large enough. By Theorem 2.1, Lemma 1.2(ii) and Theorem 3.1, we know that $(x^*, \pi(x^*))$ is a KKT pair of (0.1) and $\big(x^k, \pi(x^k)\big) \to (x^*, \pi(x^*))$. So using Assumptions A2 and A4, from [22] one has $I_1(x^*) \equiv I_{1k}$ for k large enough. Therefore, $I_k = I_{1k} \cup I_2 \equiv I_1(x^*) \cup I_2 = I(x^*)$.

By the definition of η_k at Step 6, in view of $\lim\limits_{k \to \infty} d^k = 0$ and $\lim\limits_{k \to \infty} z_k = 0$, we have $\lim\limits_{k \to \infty} \eta_k = 0$. So, taking into account the boundedness of $\{\mu_{I_k}^k\}$, $\lim\limits_{k \to \infty} \mu_0^k = 1$ follows from (1.9). To prove $\lim\limits_{k \to \infty} \mu_{I_k}^k = \bar{\mu}_{I(x^*)}^*$, it is sufficient to show that each limit point $\bar{\mu}_{I(x^*)}$ of $\{\mu_{I_k}^k\}$ equals to $\bar{\mu}_{I(x^*)}^*$, and this can be obtained from the proof of Lemma 2.3 since $(x^k, d^k) \to (x^*, 0)$.

Lastly, $L_k \subseteq I_0(x^*)$ follows from $(d^k, z_k) \to (0, 0)$ and (2.1). For $j \in I_*^+$, $\mu_j^k \to \mu_j^* = \lambda_j^* > 0$; for $j \in I_2$, $\mu_j^k \to \mu_j^* = \lambda_j^* + c = \pi_j(x^*) + c > 0$, so $\mu_j^k > 0$ and $j \in L_k$ if $j \in I_*^+ \cup I_2$ and k is large enough.

Based on Lemma 3.1, corresponding to the special case of $x^k \in X^+$, i.e., $\varphi(x^k) = 0$ of Ref. [21], in a similar fashion to Lemma 4.2 and Theorem 4.2 as well as Theorem 4.3 in Ref. [21], we can prove the following result.

Lemma 3.2 (i) Suppose that Assumptions A2–A4 hold. Then

$$\begin{aligned} \mid z_k \mid &= O(\|d^k\|), \\ \|\widetilde{d}^k\| &= O(\|d^k\|^2) + O(\mid \eta_k z_k \mid) = o(\|d^k\|), \\ \|h^k\| &= O(\|d^k\|^2) + O(\mid \eta_k z_k \mid) = o(\|d^k\|), \\ \|\widetilde{d}^k\|^2 &= O(\|d^k\|^4) + o(\mid \eta_k z_k \mid \|d^k\|), \\ \|d^k\| \ \|\widetilde{d}^k\| &= O(\|d^k\|^3) + O(\mid \eta_k z_k \mid \|d^k\|), \end{aligned}$$

where $\widetilde{d}^k$ is the solution of (1.14). So the correction direction $\widetilde{d}^k$ in the algorithm is always yielded by the solution of (1.14) if k is large enough.

(ii) Suppose that Assumptions A2–A5 hold. Then the step size $t_k = 1$ is always accepted by the arc search (1.19)–(1.20) for k large enough.

At the end of this section, based on Lemma 3.1 and $x^{k+1} = x^k + d^k + \widetilde{d}^k$ as well as $\|\widetilde{d}^k\| = o(\|d^k\|)$, similar to the analysis of Theorem 4.3 in [20], we can present the following superlinear convergence of the algorithm.

Theorem 3.2 Suppose that Assumptions A2–A5 hold. Then the proposed algorithm is superlinearly convergent, i.e., $\|x^{k+1} - x^*\| = o(\|x^k - x^*\|)$.

4 Numerical experiments

In this section, we test some practical problems given in [27, 28]. All numerical tests are implemented on MATLAB 7.1. At each iteration, we use the BFGS formula from Powell [29] to update B_k, and let B_0 be the identity matrix, and use the optimization toolbox to solve the DFS (1.8). The BFGS formula is as follows:

$$B_{k+1} = B_k - \frac{B_k s^k (s^k)^{\mathrm{T}} B_k}{(s^k)^{\mathrm{T}} B_k s^k} + \frac{y^k (y^k)^{\mathrm{T}}}{(s^k)^{\mathrm{T}} y^k} \quad , \quad (k \geqslant 0), \tag{4.1}$$

where

$$\begin{aligned} &s^k = x^{k+1} - x^k, \quad y^k = \theta \hat{y}^k + (1-\theta) B_k s^k, \\ &\hat{y}^k = \nabla F_{c_k}(x^{k+1}) - \nabla F_{c_k}(x^k) + \sum_{j=1}^{m} \bar{\mu}_j^k (\nabla g_j(x^{k+1}) - \nabla g_j(x^k)), \\ &\theta = \begin{cases} 1, & \text{if} \quad (y^k)^{\mathrm{T}} s^k \geqslant 0.2 (s^k)^{\mathrm{T}} B_k s^k; \\ \dfrac{0.8 (s^k)^{\mathrm{T}} B_k s^k}{(s^k)^{\mathrm{T}} B_k s^k - (y^k)^{\mathrm{T}} s^k}, & \text{otherwise}, \end{cases} \end{aligned}$$

and $\bar{\mu}_j^k$ is computed by (1.13). During the numerical experiments, we consider the case that $\gamma_0 = 2.0,\ \gamma_j = 0.5\gamma_0,\ j = 1, 2, \ldots, m$. And the other parameters are selected as follows:

$$\begin{aligned} &\beta = 0.58, \quad \alpha = 0.25,\ \eta_0 = 0.2, \quad a_0 = 0.2,\ \nu = 0.55, \quad \tau = 2.25; \\ &\delta_1 = 0,\ \delta_2 = 1, \quad \xi = \zeta = 0.6. \quad c_{-1} = 1.0,\ \iota = r = 0.01. \end{aligned}$$

The operational process is terminated if one of the two following conditions is satisfied:

(i) $\|\Phi(x^k, \pi(x^k))\| \leqslant 10^{-5}$;

(ii) $\|d^k\| \leqslant 10^{-5}$ or $\|z_k\| \leqslant 10^{-5}$.

The numerical reports are shown in Table 1 below, some numerical results are compared to the ones in [21].

Table 1. Numerical reports

Prob.	(n, m_i, m_e)	Method	Ni	Nf	Ng	I_k	c	f_{final}
012	(2,1,0)	SNQP1	7	4	7	1	0	-2.999 998 9E+01
		SNQP2	7	7	27			-3.000 000 0E+01
029	(3, 1, 0)	SNQP1	10	6	13	1	0	-2.262 741 4E+01
		SNQP2	11	15	42			-2.262 741 7E+01
031	(3,7,0)	SNQP1	14	14	19	1	0	6.000 000 00E+00
		SNQP2	15	21	39			6.000 000 0E+00
035	(3,4,0)	SNQP1	5	0	5	1	0	0.111 169 4E+00
		SNQP2	6	6	0			1.111 111 1E-01
043	(4,3, 0)	SNQP1	31	61	31	2	0	-4.253 242 5E+01
		SNQP2	12	12	77			-4.399 999 9E+01
093	(6, 2, 0)	SNQP1	31	62	120	2	0	1.357 737 36E+02
		SNQP2	16	16	743			1.350 759 4E+02
07	(2,0,1)	SNQP1	59	117	117	1	1	1.778 683 1E+00
032	(3,4,1)	SNQP1	20	410	82	5	2.1052	1.000 000 2E+00
037	(3,8, 0)	SNQP1	31	0	53	1	0	-3.455 876 7E+03
063	(3, 3, 2)	SNQP1	94	1 040	208	4	1.061	9.359 985 9E+02
065	(3,7, 0)	SNQP1	27	129	241	1	0	9.535 295 2E-01
100	(7, 4, 0)	SNQP1	31	58	33	2	0	6.831 124 3E+02
107	(9, 8, 6)	SNQP1	101	7 882	563	14	7.546 679 9E+04	5.709 558 1E+03
252	(3,1, 1)	SNQP1	101	416	208	2	5.806 140 1E+03	1.023 407 5E+04

The columns of Table 1 mean that: *Prob*: the problem number given in [27, 28]; n, m_i, m_e: the number of variables and inequality constraints as well as equality constraints of the test problems; *SNQP1*: our algorithm; *SNQP2*: the algorithm in [21]; *Ni, Nf, Ng*: the number of iterations and objective function evaluations as well as constraint functions evaluations, respectively; I_k: the number of indices in the final working set; c: the number of final value of c_k; f_{final}: the objective function value at the final.

References

[1] Zoutendijk G. Methods of Feasible Directions[M]. Amsterdam: Elsevier , 1960.

[2] Panier E R, Tits A L. On combining feasibility, descent and superlinear convergence in equality constrained optimization[J]. *Mathematical Programming*, 1993, **59**: 261-276.

[3] Cawood M E, Kostreva M M. Norm-relaxed method of feasible direction for solving the nonlinear programming problems[J]. *Jouranl of Optimization Theory Application*, 1994, **83**: 311-320.

[4] Chen X, Kostreva M M. A generalization of the norm-relaxed method of feasible directions[J]. *Applied Mathematics and Computation,* 1999, **102**: 257-273.

[5] Mayne D Q, Polak E. Feasible direction algorithm for optimization problems with equality and inequality constraints[J]. *Mathematical Programming,* 1976, **11**: 67-80.

[6] Lawrence C T, Tits A L. Nonlinear equality constraints in feasible sequential quadratic programming[J]. *Optimization Methods and Software*, 1996, **6**: 252-282.

[7] Herskovits J. A two-stage feasible directions algorithm for nonlinear constrained optimization[J]. *Mathematical Programming*, 1986, **36**: 19-38.

[8] Jian J B. A superlinearly convergent feasible descent algorithm for nonlinear optimization[J]. *Journal of Mathematics (in Chinese),* 1995, **15**(3): 319-326.

[9] Jian J B, Tang Ch M, Hu Q J, et al. A feasible descent SQP algorithm for general constrained optimization without strict complementarity[J]. *Journal of Computational and Applied Mathematics,* 2005, **180**: 391 ＋ 412.

[10] Herskovits J, Potra F A. Feasible direction interior-point technique for nonlinear optimization[J]. *Journal of Optomization Theory and Applications*, 1998, **99**: 121-146.

[11] Tits A L, Wachter A, Bakhtiari S, et al. A primal-dual interior-point method for nonlinear programming with strong global and local convergence properties[J]. *SIAM Journal on Optimization*, 2003, **14**: 173-199.

[12] Jian J B, Xu Q J, Han D L. A Strongly convergent norm-relaxed method of strongly subfeasible direction for optimization with nonlinear equality and inequality constraints[J]. *Applied Mathematics and Computation*, 2006, **182**: 854-870.

[13] Jian J B, Zhu Z B. Algorithm of sequential systems of linear equations with superlinear and quadratical convergence for general constrained optimization[J]. *Journal of Engineering Mathematics (in Chinese),* 2003, **20**: 24-30.

[14] Pironneau O, Polak E. Rate of convergence of a class of methods of feasible directions[J]. *SIAM Journal on Optimization*, 1973, **10**: 161-173.

[15] Jian J B, Zheng H Y, Hu Q J, et al. A new norm-relaxed method of strongly sub-feasible direction for inequality constrained optimization[J]. *Applied Mathematics and Computation*, 2005, **168**: 1-28.

[16] Kostreva M M, Chen X. A superlinearly convergent method of feasible directions[J]. *Applied Mathematics and Computation,* 2000, **116**: 231-244.

[17] Lawrence C T, Tits A L. A computationally efficient feasible sequential quadratic programming algorithm[J]. *SIAM Journal on Optimization*, 2001, **11**: 1092-1118.

[18] Zhu Z B, Zhang K C. A new SQP method of feasible directions for nonlinear programming[J]. *Applied Mathematics and Computation,* 2004, **148**: 121-134.

[19] Maratos N. Exact penalty function algorithm for finite dimensional and control optimization problems[D]. Ph.D. thesis, Imperial College Science, Technology, University of London, 1978.

[20] Jian J B, Zheng H Y, Tang C M, et al. A new superlinearly convergent norm-relaxed method of strongly sub-feasible direction for inequality constrained optimization[J]. *Applied Mathematics and Computation*, 2006, **182**: 955-976.

[21] Jian J B, Ke X Y, Zheng H Y, et al. A method combining norm-relaxed QP subproblem with system of linear equations for constraint optimization[J]. *Journal of Computational and Applied Mathematics*, 2009, **223**: 1013-1027.

[22] Facchinei F, Fischer A, Kanzow C. On the accurate identication of active constraints[J]. *SIAM Journal on Optimization*, 1998, **9**: 14-32.

[23] Wang Y, Chen L, He G. Sequential systems of linear equations method for general constrained optimization without strict complementarity[J]. *Journal of Computational and Applied Mathematics,* 2005, **182**: 447-471.

[24] Wang Y. Study on the QP-free algorithms for solving nonlinear constrained optimization problems[D]. PhD thesis, School of Shang'hai Jiaotong University, Shang'hai, China.

[25] Jian J B. Fast algorithms for smooth constrained optimization–theoretical analysis and numerical experiments[M]. Beijing: Science Press, 2010.

[26] Kanzow C, Qi H D. A QP-free constrained Newton-type method for variational inequality problems[J]. *Mathematical Programming*, 1999, **85**: 81-106.

[27] Hock W, Schittkowski K. Tests Examples for Nonlinear Programming Codes[M]. *Lecture Notes in Economics and Mathematical Systems,* Berlin Heidelberg New York: Springer-Verlag, 1981, **187**.

[28] K. Schittkowski, *More Test Examples for Nonlinear Programming Codes*, Spring Verlag, 1987.

[29] Powell M J D. The convergence of variable metric methods for nonlinearly constrained optimization calculations[J]. *Nonlinear programming*, **3**, Edited by R.R. Meyer and S. M. Robinson, Academic Press, New York, 1978.