Modified lower order penalty functions based on quadratic smoothing approximation

BAI Fusheng[1†] LUO Xiaoyan[1]

Abstract In this paper, two function forms of quadratic smoothing approximation to the lower order exact penalty function are proposed to generate modified smooth penalty functions for inequality-constrained optimization problems. It is shown that under certain conditions, any global minimizer of the modified smooth penalty problem is a global minimizer to the original constrained optimization problem when the penalty parameter is sufficiently large. Two numerical examples are given to show the effectiveness of the present smoothing scheme.

Keywords modified penalty function, smoothing approximation, lower order penalty function, inequality-constrained optimization problem

Chinese Library Classification O221.2

2010 Mathematics Subject Classification 90C30

基于二次函数光滑化逼近的修正低阶罚函数

白富生 [1†] 罗晓艳 [1]

摘要 针对不等式约束优化问题，给出了通过二次函数对低阶精确罚函数进行光滑化逼近的两种函数形式，得到修正的光滑罚函数．证明了在一定条件下，当罚参数充分大时，修正的光滑罚问题的全局最优解是原优化问题的全局最优解．给出的两个数值例子说明了所提出的光滑化方法的有效性．

关键词 修正罚函数，光滑化逼近，低阶罚函数，不等式约束优化问题

中图分类号 O221.2

数学分类号 90C30

0 Introduction

Consider the following global optimization problem:

$$
\begin{aligned}
[P]\quad \min\quad & f(x),\\
\text{s.t.}\quad & g_i(x)\leqslant 0, i=1,2,\ldots,m,\\
& x\in\mathbb{R}^n,
\end{aligned}
$$

where $f:\mathbb{R}^n\to\mathbb{R}, g_i:\mathbb{R}^n\to\mathbb{R}, i=1,2,\ldots,m$ are twice continuously differentiable. In the last fifty years, a significant amount of investigations have been devoted to exact penalty

收稿日期： 2010 年 6 月 28 日．

1. School of Mathematics, Chongqing Normal University, Chongqing 401331, China; 重庆师范大学数学学院，重庆， 401331

† 通讯作者 Corresponding author

functions.

Let $p_k(u) = (\max\{0, u\})^k$, that is,

$$p_k(u) = \begin{cases} u^k, & \text{if} \quad u > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{0.1}$$

Denote

$$\varphi_{q,k}(x) := f(x) + q\sum_{i=1}^{m} p_k(g_i(x)). \tag{0.2}$$

When $k = 1$, the function $\varphi_{q,k}(x)$ is l_1 penalty function of problem $[P]$; when $k \in (0, 1)$, the function is a lower order penalty function of problem $[P]$ (see [1-8]). It is shown in [8] that the second-order sufficient condition implies local exact penalty property for the lower order penalty function with any positive penalty parameter.

Since $p_k(u)$ is not differentiable, in general $\varphi_{q,k}(x)$ $(k \in (0, 1])$ is a non-differentiable function. However, most powerful methods in optimization require a differentiable cost function. This motivates the smoothing of $\varphi_{q,k}(x)$ via the smoothing of $p_k(u)$. The case with $k = 1$ has been investigated in e.g. [6, 9]. The case with $k = \frac{1}{2}$ has been investigated in e.g. [5, 8]. References [4] and [10] investigated the general cases for $k \in (0, 1)$ and $k \in (\frac{1}{2}, 1]$ respectively.

In this paper, we propose a quadratic smoothing approximation to $p_k(u)$ with $k \in (0, 1)$. The approximation takes two similar function forms. Unlike the smoothing approximations in [4, 10], where the constructed smooth functions satisfying traditional definition of penalty function, the auxiliary function constructed on the smoothing approximation in this paper is a modified penalty function which does not satisfy the traditional definition of penalty function. However, the present modified penalty function can be used to implement penalty to infeasible points effectively, as indicated by Theorems 1.1 and 1.2. It should be noted that modified smooth penalty functions have been proposed in [8, 9] based on quadratic smoothing approximation to deal with the case of a single k value.

The rest of this paper is organized as follows. In Section 1, we introduce the smoothing function to $p_k(u)$, and give some fundamental properties about the constructed modified penalty function based on the smoothing function. In Section 2, a simple algorithm is proposed to obtain an appropriate global optimal solution to the original optimization problem. Two numerical examples are given in this section to show the effectiveness of the present smoothing scheme.

1 Smoothing approximation

To begin with, we introduce the concept of second-order sufficient condition (see [2], p. 169). Let

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

We say that the pair (x^*, λ^*) satisfies the second-order sufficient condition, if

$$\begin{aligned} \nabla_x L(x^*, \lambda^*) = 0 & \\ g_i(x^*) \leqslant 0, & \quad i \in \{1, \ldots, m\} \\ \lambda_i^* \geqslant 0, & \quad i \in \{1, \ldots, m\} \\ \lambda_i^* g_i(x^*) = 0, & \quad i \in \{1, \ldots, m\} \\ y^T \nabla^2 L(x^*, \lambda^*) y > 0, & \quad \text{for any } y \in V(x^*), \end{aligned} \tag{1.1}$$

where

$$V(x^*) = \left\{ y \in R^n \,\middle|\, \begin{array}{ll} \nabla^T g_i(x^*) y = 0, & i \in A(x^*) \\ \nabla^T g_i(x^*) y \leqslant 0, & i \in B(x^*) \end{array} \right\},$$

$$A(x^*) = \{i \in \{1, \ldots, m\} | g_i(x^*) = 0,\ \lambda_i^* > 0\},$$

$$B(x^*) = \{i \in \{1, \ldots, m\} | g_i(x^*) = 0,\ \lambda_i^* = 0\}.$$

Assumption 1 $f(x)$ satisfied the following coercive condition:

$$\lim_{\|x\| \to \infty} f(x) = +\infty.$$

By Assumption 1, there exists a box X such that $G[P] \subset \operatorname{int}(X)$, where $G[P]$ is the set of global minimizers of problem $[P]$, $\operatorname{int}(X)$ denotes the interior of the set X. Then, problem $[P]$ is equivalent to the following problem $[P']$:

$$\begin{aligned} [P'] \quad \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leqslant 0, i = 1, 2, \ldots, m, \\ & x \in X, \end{aligned}$$

in the sense of $G[P] = G[P']$, where $G[P]$ is the set of global minimizers of problem $[P]$.

Let $G[P']$ denote the set of global minimizers of problem $[P']$, then $G[P] = G[P']$.

Assumption 2 The set $G[P']$ is a finite set.

For the following penalty problem:

$$[LOP]_k \quad \min_{x \in X} \varphi_{q,k}(x),$$

where $\varphi_{q,k}(x)$ is given in (0.2), we have the following lemma.

Lemma 1.1 *(See [8]) Suppose that Assumptions 1 and 2 hold, and furthermore, that for any $x^* \in G[P]$, there exists $\lambda^* \in R_+^m$ such that the pair (x^*, λ^*) satisfies the second-order sufficient condition (1.1). Then, for any $k \in (0, 1)$, there exists $q^* > 0$, such that when $q > q^*$, $G[P] = G([LOP]_k)$, where $G([LOP]_k)$ is the set of global minimizers of problem $[LOP]_k$.*

Now we consider the smoothing approximation to the lower order penalty function $\varphi_{q,k}(x)$. We use a quadratic function $l(x) = au^2 + bu + c$ to approximate $p_k(u)$ for $u \in [\delta, 0)$, where $\delta < 0$, and a k-order power function $r(u) = (u + \gamma)^k$ to approximate $p_k(u)$ for $u \geqslant 0$; that is, we use the following piecewise function to approximate $p_k(u)$:

$$s(u)=\begin{cases}0, & \text{if} \quad u<\delta;\\ au^2+bu+c, & \text{if } \delta\leqslant u<0;\\ (u+\gamma)^k, & \text{if} \quad u\geqslant 0.\end{cases} \tag{1.2}$$

To determine the unknowns in (1.2), we can follow two different paths. The first one is to set the difference between $l(0)$ and $p_u(0)$ as $\frac{\varepsilon}{qm}$, i.e. $l(0)-p_k(0)=c-0=c=\frac{\varepsilon}{qm}$, and find the unknowns accordingly. Since the function used for approximation should be continuously differentiable, we have $l(0)=r(0)$, i.e. $c=(0+\gamma)^k$ which yields $\gamma=(\frac{\varepsilon}{qm})^{\frac{1}{k}}$, and $l'(0)=r'(0)$, i.e. $b=k\gamma^{k-1}=k(\frac{\varepsilon}{qm})^{1-\frac{1}{k}}$. To find a and δ, we need to solve the following system of linear equations:

$$\begin{cases} l(\delta)=a\delta^2+b\delta+c=a\delta^2+k\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{1}{k}}\delta+\dfrac{\varepsilon}{qm}=p_k(0)=0;\\ l'(\delta)=2a\delta+b=2a\delta+k\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{1}{k}}=p_k'(0)=0.\end{cases}$$

The solution of the above system is

$$a=\frac{k^2}{4}\left(\frac{\varepsilon}{qm}\right)^{1-\frac{2}{k}},\qquad \delta=-\frac{2}{k}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}.$$

Thus

$$s(u)=\begin{cases}0, & \text{if} \quad u<-\dfrac{2}{k}\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}};\\ \dfrac{k^2}{4}\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{2}{k}}u^2+k\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{1}{k}}u+\dfrac{\varepsilon}{qm}, & \text{if } -\dfrac{2}{k}\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\leqslant u<0;\\ \left(u+\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k, & \text{if} \quad u\geqslant 0.\end{cases} \tag{1.3}$$

The second path is to set the interval $[\delta,0]$ as $[-(\frac{\varepsilon}{qm})^{\frac{1}{k}},0]$ first, then find all other unknowns accordingly. Again, since the function used for approximation should be continuously differentiable, we have

$$\begin{cases} l\left(-\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)=a\left(\dfrac{\varepsilon}{qm}\right)^{\frac{2}{k}}-b\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}+c=p_k(-\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}})=0;\\ l'\left(-\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)=-2a\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}+b=p_k'\left(-\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)=0;\\ l(0)=c=r(0)=\gamma^k;\\ l'(0)=b=r'(0)=k\gamma^{k-1}.\end{cases}$$

The solution of the above system is

$$a=\left(\frac{k}{2}\right)^k\left(\frac{\varepsilon}{qm}\right)^{1-\frac{2}{k}},\quad b=k\left(\frac{k}{2}\right)^{k-1}\left(\frac{\varepsilon}{qm}\right)^{1-\frac{1}{k}},\quad c=\left(\frac{k}{2}\right)^k\frac{\varepsilon}{qm},\quad \gamma=\frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}.$$

Thus we have

$$s(u)=\begin{cases}0, & \text{if} \quad u<-\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}};\\ \left(\dfrac{k}{2}\right)^{k}\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{2}{k}}u^2+k\left(\dfrac{k}{2}\right)^{k-1}\left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{1}{k}}u+\left(\dfrac{k}{2}\right)^{k}\dfrac{\varepsilon}{qm}, & \text{if} \ -\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\leqslant u<0;\\ \left(u+\dfrac{k}{2}\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^{k}, & \text{if} \quad u\geqslant 0.\end{cases} \tag{1.4}$$

Note that the two $s(u)$s given by (1.3) and (1.4) are very similar to each other. In the rest of this paper we take the $s(u)$ given by (1.4) as the smoothing function. To indicate the links with q,ε,k, we use $p_{q,\varepsilon,k}(u)$ to denote this function.

Remark 1.1 If the $s(u)$ given by (1.3) is adopted as the smoothing function, all the following results still hold with minor modification.

Figure 1 shows the behavior of $p_{q,\varepsilon,k}(u)$ and $p_k(u)$ with $m=2,q=5,\varepsilon=0.2,k=\frac{1}{3}$.

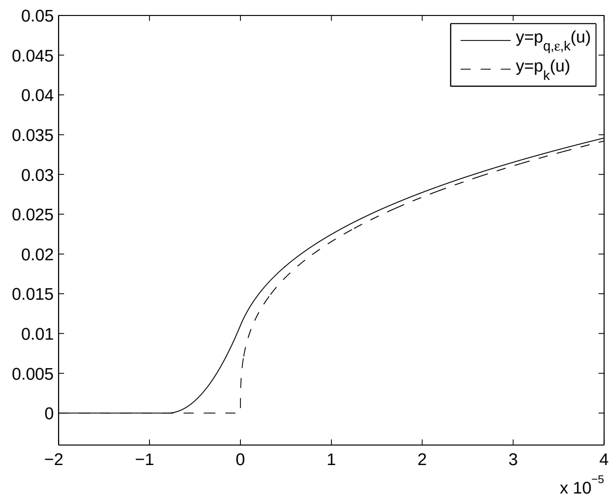

Figure 1: The behavior of $p_{q,\varepsilon,k}(u)$ and $p_k(u)$ with $m=2,q=5,\varepsilon=0.2,k=\frac{1}{3}$

Let

$$\varphi_{q,\varepsilon,k}(x)=f(x)+q\sum_{i=1}^{m}p_{q,\varepsilon,k}(g_i(x)). \tag{1.5}$$

Then it is easy to see that $\varphi_{q,\varepsilon,k}(x)$ is continuously differentiable on $\mathbb{R}^n$. It should be noted that $\varphi_{q,\varepsilon,k}(x)$ does not satisfy the traditional definition of penalty function, as the penalty term function $\sum\limits_{i=1}^{m}p_{q,\varepsilon,k}(g_i(x))$ may take positive value on a feasible point. Thus we call $\varphi_{q,\varepsilon,k}(x)$ a *modified penalty function*.

Consider the following *modified penalty problem*:

$$[SP]_k \quad \min_{x \in X} \varphi_{q,\varepsilon,k}(x)$$

Proposition 1.1 *For any $x \in R^n$, $q > 0$ and $\varepsilon > 0$, we have*

$$0 \leqslant \varphi_{q,\varepsilon,k}(x) - \varphi_{q,k}(x) \leqslant \left(\frac{k}{2}\right)^k \varepsilon, \tag{1.6}$$

where $\varphi_{q,\varepsilon,k}(x)$ and $\varphi_{q,k}(x)$ are given in (1.5) and (0.2) respectively.

Proof

Let

$$\Delta p_i = p_{q,\varepsilon,k}(g_i(x)) - p_k(g_i(x)).$$

Note that

$$\Delta p_i = \begin{cases} 0, & \text{if } g_i(x) < -\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}; \\ \left(\dfrac{k}{2}\right)^k \left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{2}{k}} (g_i(x))^2 + k\left(\dfrac{k}{2}\right)^{k-1} \left(\dfrac{\varepsilon}{qm}\right)^{1-\frac{1}{k}} & \\ \quad \cdot (g_i(x)) + \left(\dfrac{k}{2}\right)^k \dfrac{\varepsilon}{qm}, & \text{if } -\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}} \leqslant g_i(x) < 0; \\ \left(g_i(x) + \dfrac{k}{2}\left(\dfrac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k - (g_i(x))^k, & \text{if } g_i(x) \geqslant 0. \end{cases}$$

and for any $k \in (0,1), q > 0, x \in X$,

$$0 \leqslant \left(\frac{k}{2}\right)^k \left(\frac{\varepsilon}{qm}\right)^{1-\frac{2}{k}} (g_i(x))^2 + k\left(\frac{k}{2}\right)^{k-1} \left(\frac{\varepsilon}{qm}\right)^{1-\frac{1}{k}} (g_i(x)) + \left(\frac{k}{2}\right)^k \frac{\varepsilon}{qm} \leqslant \left(\frac{k}{2}\right)^k \frac{\varepsilon}{qm}$$

and

$$0 \leqslant \left(g_i(x) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k - (g_i(x))^k \leqslant \left(\frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k = \left(\frac{k}{2}\right)^k \frac{\varepsilon}{qm},$$

we have

$$0 \leqslant \varphi_{q,\varepsilon,k}(x) - \varphi_{q,k}(x) = q\sum_{i=1}^m \Delta p_i \leqslant q\sum_{i=1}^m \left(\frac{k}{2}\right)^k \frac{\varepsilon}{qm} = \left(\frac{k}{2}\right)^k \varepsilon.$$

Proposition 1.2 *Let $x^*_{q,k} \in X$ be a global minimizer of problem $[LOP]_k$ and $\overline{x}_{q,\varepsilon,k} \in X$ be a global minimizer of problem $[SP]_k$ for some $q > 0$, $k \in (0,1)$ and $\varepsilon > 0$. Then we have*

$$0 \leqslant \varphi_(\overline{x}_{q,\varepsilon,k}) - \varphi_{q,k}(x^*_{q,k}) \leqslant \left(\frac{k}{2}\right)^k \varepsilon. \tag{1.7}$$

Proof By Proposition 1.1

$$\begin{aligned} 0 \leqslant \varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) - \varphi_{q,k}(\overline{x}_{q,\varepsilon,k}) &\leqslant \varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) - \varphi_{q,k}(x^*_{q,k}) \\ &\leqslant \varphi_{q,\varepsilon,k}(x^*_{q,k}) - \varphi_{q,k}(x^*_{q,k}) \\ &\leqslant \left(\frac{k}{2}\right)^k \varepsilon. \end{aligned}$$

Corollary 1.1 *Let $x^*_{q,k} \in X$ be a global minimizer of problem $[LOP]_k$ and $\overline{x}_{q,\varepsilon,k} \in X$ be a global minimizer of problem $[SP]_k$ for some $q > 0$, $k \in (0,1)$ and $\varepsilon > 0$. If $x^*_{q,k}$ and $\overline{x}_{q,\varepsilon,k}$ are feasible to problem $[P]$, then we have*

$$0 \leqslant f(\overline{x}_{q,\varepsilon,k}) - f(x^*_{q,k}) \leqslant \left(\frac{k}{2}\right)^k \varepsilon. \tag{1.8}$$

Proof As $x^*_{q,k} \in X$ is a global minimizer of problem $[P]$, we have

$$\sum_{i=1}^m p_k(g_i(x_{q,k})) = 0.$$

Thus

$$\begin{aligned} f(\overline{x}_{q,\varepsilon,k}) - f(x^*_{q,k}) &= f(\overline{x}_{q,\varepsilon,k}) + q\sum_{i=1}^m p_{q,\varepsilon,k}(g_i(\overline{x}_{q,\varepsilon,k})) - f(x^*_{q,k}) \\ &\quad + q\sum_{i=1}^m p_k(g_i(x^*_{q,k})) - q\sum_{i=1}^m p_{q,\varepsilon,k}(g_i(\overline{x}_{q,\varepsilon,k})) \\ &= \varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) - \varphi_{q,k}(x^*_{q,k}) - q\sum_{i=1}^m p_{q,\varepsilon,k}(g_i(\overline{x}_{q,\varepsilon,k})). \end{aligned}$$

By Proposition 1.2, we have

$$\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) - \varphi_{q,k}(x^*_{q,k}) \leqslant \left(\frac{k}{2}\right)^k \varepsilon.$$

By the nonnegativity of $p_{q,\varepsilon,k}(u)$, we then have

$$f(\overline{x}_{q,\varepsilon,k}) - f(x^*_{q,k}) \leqslant \left(\frac{k}{2}\right)^k \varepsilon.$$

Since $x^*_{q,k}$ is feasible to problem $[P']$, it is a global minimizer of problem $[P']$. Note that $\overline{x}_{q,\varepsilon,k}$ is also feasible to problem $[P']$, so it holds $f(\overline{x}_{q,\varepsilon,k}) - f(x^*_{q,k}) \geqslant 0$. Then we have

$$0 \leqslant f(\overline{x}_{q,\varepsilon,k}) - f(x^*_{q,k}) \leqslant \left(\frac{k}{2}\right)^k \varepsilon.$$

From the above result, if the global minimizer $x^*_{q,k}$ of the non-smooth penalty problem $[LOP]_k$ and the global minimizer $\overline{x}_{q,\varepsilon,k}$ of the modified smooth penalty problem $[SP]_k$ are feasible to problem $[P]$, then the difference between the objective function values on $\overline{x}_{q,\varepsilon,k}$ and $x^*_{q,k}$ can be controlled through the smoothing parameter ε.

Let

$$S = \{x \in X | g_i(x) \leqslant 0, i = 1, \ldots, m\}, \tag{1.9}$$

$$\overline{S} = \{x \in X | g_i(x) < 0, i = 1, \ldots, m\}. \tag{1.10}$$

and $G[SP]_k$ be the set of all the global minimizers of problem $[SP]_k$.

Theorem 1.1 *Suppose that Assumptions 1 and 2 hold, and there exists $x^* \in G[P] \bigcap \bar{S}$. Then, for any given $\varepsilon > 0$, there exists $q^* > 0$ such that any global minimizer of the modified smooth penalty problem $[SP]_k$ is a global minimizer of the original constrained optimization problem $[P]$, i.e., $G[SP]_k \subset G[P]$ when $q > q^*$.*

Proof For any given $\varepsilon > 0$, $k \in (0,1)$, $q > 0$ and any $\overline{x}_{q,\varepsilon,k}$, let

$$A_{q,\varepsilon,k} = \left\{ i | -\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}} \leqslant g_i(\overline{x}_{q,\varepsilon,k}) \leqslant 0, i = 1, \ldots, m \right\}, \tag{1.11}$$

$$B_{q,\varepsilon,k} = \{ i | g_i(\overline{x}_{q,\varepsilon,k}) > 0, i = 1, \ldots, m\}. \tag{1.12}$$

If $B_{q,\varepsilon,k} \neq \emptyset$, we have

$$\begin{aligned}
\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) &= f(\overline{x}_{q,\varepsilon,k}) + q \sum_{i \in B_{q,\varepsilon,k}} \left(g_i(\overline{x}_{q,\varepsilon,k}) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}} \right)^k \\
&\quad + q \sum_{i \in A_{q,\varepsilon,k}} \left(\left(\frac{k}{2}\right)^k \left(\frac{\varepsilon}{qm}\right)^{1-\frac{2}{k}} (g_i(\overline{x}_{q,\varepsilon,k}))^2 \right. \\
&\quad \left. + k\left(\frac{k}{2}\right)^{k-1} \left(\frac{\varepsilon}{qm}\right)^{1-\frac{1}{k}} g_i(\overline{x}_{q,\varepsilon,k}) + \left(\frac{k}{2}\right)^k \frac{\varepsilon}{qm} \right) \\
&\geqslant f(\overline{x}_{q,\varepsilon,k}) + q \sum_{i \in B_{q,\varepsilon,k}} \left(g_i(\overline{x}_{q,\varepsilon,k}) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}} \right)^k \\
&= f(\overline{x}_{q,\varepsilon,k}) + q \sum_{i \in B_{q,\varepsilon,k}} (g_i(\overline{x}_{q,\varepsilon,k}))^k \\
&\quad + q \sum_{i \in B_{q,\varepsilon,k}} \left(\left(g_i(\overline{x}_{q,\varepsilon,k}) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k - (g_i(\overline{x}_{q,\varepsilon,k}))^k \right) \\
&= \varphi_{q,k}(\overline{x}_{q,\varepsilon,k}) + q \sum_{i \in B_{q,\varepsilon,k}} \left(\left(g_i(\overline{x}_{q,\varepsilon,k}) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k - (g_i(\overline{x}_{q,\varepsilon,k}))^k \right).
\end{aligned}$$

By the assumption that there exists $x^* \in G[P] \bigcap \bar{S}$, there must exist $q_1 > 0$, such that $g_i(x^*) < -(\frac{\varepsilon}{qm})^{\frac{1}{k}}$ for any $i = 1, \ldots, m$ when $q > q_1$. Hence, when $q > q_1$, we have

$$\varphi_{q,\varepsilon,k}(x^*) = f(x^*). \tag{1.13}$$

By Lemma 1.1, there exists q_0 such that $G[P] = G[LOP]_k$ for any $q > q_0$. Let $q^* = \max\{q_1, q_0\}$. Then for any $q > q^*$, we have

$$\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) \leqslant \varphi_{q,\varepsilon,k}(x^*) = f(x^*) = \varphi_{q,k}(x^*) \leqslant \varphi_{q,k}(\overline{x}_{\varepsilon,q,k}). \tag{1.14}$$

Then, we claim that $\overline{x}_{q,\varepsilon,k}$ is a feasible solution to the original problem $[P]$ when $q > q^*$. In fact, if $q > q^*$ and $\overline{x}_{\varepsilon,q,k}$ is not feasible to problem $[P]$, i.e., $B_{q,\varepsilon,k} \neq \emptyset$, then we have

$$\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) \geqslant \varphi_{q,k}(\overline{x}_{q,\varepsilon,k}) + q \sum_{i \in B_{q,\varepsilon,k}} \left(\left(g_i(\overline{x}_{q,\varepsilon,k}) + \frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k - (g_i(\overline{x}_{q,\varepsilon,k}))^k \right) > \varphi_{q,k}(\overline{x}_{q,\varepsilon,k}),$$

which contradicts (1.14).

By (1.13) and

$$f(x^*) \leqslant f(\overline{x}_{q,\varepsilon,k}) \leqslant f(\overline{x}_{q,\varepsilon,k}) + q\sum_{i=1}^{m} p_{q,\varepsilon,k}(g_i(\overline{x}_{q,\varepsilon,k})) = \varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) \leqslant \varphi_{q,\varepsilon,k}(x^*),$$

we have

$$f(x^*) = f(\overline{x}_{q,\varepsilon,k}).$$

Therefore, $\overline{x}_{q,\varepsilon,k}$ is a global minimizer of the original problem $[P]$, i.e, $\overline{x}_{q,\varepsilon,k} \in G[P]$. Thus it follows $G[SP]_k \subset G[P]$ when $q > q^*$.

Now we give the definition of ε-approximate global minimizer.

Definition 1.1 *we say that $\overline{x}$ is an ε-approximate global minimizer of problem $[P]$ if $\overline{x} \in X$ is feasible to problem $[P]$, and $|f(\overline{x}) - f^*| < \varepsilon$, where f^* is the global minimal value of problem $[P]$.*

The following result is needed to establish the result on achieving approximate global optimality of the original problem by the global minimizer of the modified penalty problem $[SP]_k$.

Lemma 1.2 *Let $k \in (0,1)$. Function $\frac{1+x^k}{(1+x)^k}$ attains its maximum on $(0,+\infty)$ at $x=1$ with the associated function value 2^{1-k}.*

Proof Let $v(x) = \frac{1+x^k}{(1+x)^k}, x \in (0,+\infty)$. Then

$$\begin{aligned} v'(x) &= \frac{kx^{k-1}(1+x)^k - k(1+x)^{k-1}(1+x^k)}{(1+x)^{2k}} \\ &= \frac{k(1+k)^{k-1}(x^{k-1}(1+x) - (1+x^k))}{(1+x)^{2k}} \\ &= \frac{k(x^{k-1}-1)}{(1+x)^{k+1}}. \end{aligned}$$

When $0 < x < 1$, it holds $x^{1-k} < 1$, i.e. $x^{k-1} > 1$ or $x^{k-1} - 1 > 0$, thus $v'(x) > 0$. On the other hand, when $x > 1$, it holds $x^{1-k} > 1$, i.e. $x^{k-1} < 1$ or $x^{k-1} - 1 < 0$, thus $v'(x) < 0$. Therefore, $v(x)$ attains its maximum on $(0,+\infty)$ at $x=1$, and $v(1) = \frac{1+1^k}{(1+1)^k} = 2^{1-k}$.

Theorem 1.2 *Suppose that Assumptions 1 and 2 hold, furthermore, $cl(\overline{S})=S$, where $\overline{S}$ and S are given in (1.10) and (1.9) respectively. Then, for any given $\varepsilon>0$, there exists $q_1^*>0$ such that when $q>q_1^*$, any global minimizer $\overline{x}_{q,\varepsilon,k}$ of problem $[SP]_k$ is a $(\frac{k}{2})^k\varepsilon$-approximate global minimizer of problem $[P]$, and satisfies $f(\overline{x}_{q,\varepsilon,k})-f^*\geqslant 0$.*

Proof For any given $\varepsilon>0$, $q>0$, $k\in(0,1)$, and $\overline{x}_{q,\varepsilon,k}\in G[SP]_k$, if there exists a global minimizer x^* of problem $[P]$ lying in $\overline{S}$, then by Theorem 1.1, it is easy to see that the conclusions hold; else take an $x^*\in S\backslash\overline{S}$. By $\mathrm{cl}(\overline{\mathrm{S}})=\mathrm{S}$, there exists a sequence $\{x_n\}\subset\overline{S}$, such that

$$\lim_{k\to\infty} x_n = x^*.$$

Therefore, there exists $n_0>0$ such that

$$f(x_n)<f(x^*)+\frac{k^k\varepsilon}{2m},$$

when $n\geqslant n_0$. Especially, we have

$$f(x_{n_0})<f(x^*)+\frac{k^k\varepsilon}{2m}. \tag{1.15}$$

By $x_{n_0}\in\overline{S}$, for the given $\varepsilon>0$, there exists $q_1>0$, such that $g_i(x_{n_0})<-(\frac{\varepsilon}{qm})^{\frac{1}{k}}$ for any $i=1,\ldots,m$ when $q>q_1$. Thus, when $q>q_1$, we have

$$f(x_{n_0})=\varphi_{q,\varepsilon,k}(x_{n_0})\geqslant\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}). \tag{1.16}$$

Let $q_1^*=2^{1-k}\max\{q_1,q^*\}$, where q^* is given in Lemma 1.1. Then we say that $\overline{x}_{q,\varepsilon,k}$ is a feasible solution of problem (P) when $q>q_1^*$.

In fact, if for $q>q_1^*$, $\overline{x}_{q,\varepsilon,k}$ is not feasible, then there exists at least one index $i_0\in B_{q,\varepsilon,k}$, where $B_{q,\varepsilon,k}$ is given in (1.12). Note that $q>q_1^*$ implies $q>q_1$. Let $q=2^{1-k}\overline{q}$, then for $q>q_1^*$ we have $\overline{q}>q^*$. Thus it holds

$$\begin{aligned}
\varphi_{q,\varepsilon,k}(\overline{x}_{q,\varepsilon,k}) \geqslant\ & f(\overline{x}_{q,\varepsilon,k})+q\sum_{i\in B_{q,\varepsilon,k}}\left(g_i(\overline{x}_{q,\varepsilon,k})+\frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k\\
=\ & f(\overline{x}_{q,\varepsilon,k})+\overline{q}\sum_{i\in B_{q,\varepsilon,k}}(g_i(\overline{x}_{q,\varepsilon,k}))^k\\
& +\overline{q}\sum_{i\in B_{q,\varepsilon,k}}\left(2^{1-k}\left(g_i(\overline{x}_{q,\varepsilon,k})+\frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k-(g_i(\overline{x}_{q,\varepsilon,k}))^k\right)\\
\geqslant\ & f(\overline{x}_{q,\varepsilon,k})+\overline{q}\sum_{i\in B_{q,\varepsilon,k}}(g_i(\overline{x}_{q,\varepsilon,k}))^k\\
& +\overline{q}\left(2^{1-k}\left(g_{i_0}(\overline{x}_{q,\varepsilon,k})+\frac{k}{2}\left(\frac{\varepsilon}{qm}\right)^{\frac{1}{k}}\right)^k-(g_{i_0}(\overline{x}_{q,\varepsilon,k}))^k\right)\\
=\ & f(\overline{x}_{q,\varepsilon,k})+\overline{q}\sum_{i\in B_{q,\varepsilon,k}}(g_i(\overline{x}_{q,\varepsilon,k}))^k\\
& +\overline{q}(g_{i_0}(\overline{x}_{q,\varepsilon,k}))^k\left(2^{1-k}\left(1+\frac{\frac{k}{2}(\frac{\varepsilon}{qm})^{\frac{1}{k}}}{g_{i_0}(\overline{x}_{q,\varepsilon,k})}\right)^k-1\right)
\end{aligned}$$

$$\begin{aligned}
&\geqslant f(\overline{x}_{q,\varepsilon,k})+\overline{q}\sum_{i\in B_{q,\varepsilon,k}}(g_i(\overline{x}_{q,\varepsilon,k}))^k\\
&\quad +\overline{q}(g_{i_0}(\overline{x}_{q,\varepsilon,k}))^k\left(\frac{\frac{k}{2}(\frac{\varepsilon}{qm})^{\frac{1}{k}}}{g_{i_0}(\overline{x}_{q,\varepsilon,k})}\right)^k \qquad \text{(by Lemma 1.2)}\\
&= \varphi_{\overline{q},k}(\overline{x}_{q,\varepsilon,k})+\overline{q}\left(\frac{k}{2}\right)^k\frac{\varepsilon}{qm}\\
&\geqslant \varphi_{\overline{q},k}(x^*)+\frac{k^k\varepsilon}{2m} \qquad (\text{by } \overline{\text{q}} > \text{q}^*)\\
&= f(x^*)+\frac{k^k\varepsilon}{2m}\\
&> f(x_{n_0}) \qquad \text{(by (1.15))}
\end{aligned}$$

which contradicts (1.16).

By Corollary 1.1, we have

$$0\leqslant f(\overline{x}_{q,\varepsilon,k})-f^*\leqslant\left(\frac{k}{2}\right)^k\varepsilon.$$

Thus, $\overline{x}_{q,\varepsilon,k}$ is a $(\frac{k}{2})^k$-approximate global minimizer of problem (P), and further satisfies $f(\overline{x}_{q,\varepsilon,k})-f^*\geqslant 0$.

Remark 1.2 It is not easy to check whether the conditions of Theorems 1.1 or 1.2 hold. However, many practical inequality-constrained optimization problems do satisfy these conditions. For simplicity, in the algorithm and the numerical examples presented in the next section, we assume that when the penalty parameter are appropriately chosen as described in the algorithm, a global minimizer of the modified penalty function can be regarded as an approximate global minimizer of the original optimization problem. The gauge of precision for the approximation is not used in the algorithm and in the numerical examples.

2 Algorithm and numerical examples

In this section, we propose a simple algorithm to solve problem $[P]$ via solving the modified smooth penalty problem $[SP]_k$. Two numerical examples are provided to show the applicability of the algorithm with $k=\frac{1}{3}$ and $k=\frac{1}{4}$ respectively.

Algorithm (SP):

Step 1. Choose $M>0,\varepsilon>0,k\in(0,1)$. Take an initial point $x_1^0\in X$, and two initial parameters $\varepsilon_1>\varepsilon$, $q_1<M$. Let $n=1$.

Step 2. Solve the following modified penalty problem:

$$[SP]_k \quad \min_{x\in X}\varphi_{q_n,\varepsilon_n,k}(x)$$

with x_n^0 as the starting point. Let x_n^* be a global minimizer of the smooth problem $[SP]_k$.

Step 3. If $q_n \geqslant M$ and $\varepsilon_n = \varepsilon$, then stop. The obtained global minimizer x_n^* can be regarded as an approximate global minimizer of problem $[P]$. Otherwise, let $\varepsilon_{n+1} = \max\{\varepsilon, \frac{\varepsilon_n}{10}\}$, $q_{n+1} = \max\{q_n^2, M\}$, $x_{n+1}^0 = x_n^*$, $n = n+1$, and go to *Step 2*.

In each of the following examples, we take $\varepsilon = 10^{-4}$, $M = 10^8$ and $\varepsilon_1 = 10^{-2}$, $q_1 = 10^2$ and we use

n : the number of iterations to solve the smooth problem $[SP]_k$;

x_n^0 : the initial point to solve the smooth problem $[SP]_k$ at the nth iteration;

x_n^* : the obtained global minimizer of problem $[SP]_k$ at the nth iteration.

In the following tables, floating point format with 6 digits is adopted to record the numerical results except the first columns, where the exact values of initial points are recorded. Also we omit "$e+0$" in the entries of the tables.

Example 2.1

$$\begin{aligned}
\min \quad f(x) &= -x_1 - x_2,\\
\text{s.t.} \quad g_1(x) &= x_1 - 2 - 2x_1^4 + 8x_1^3 - 8x_1^2,\\
g_2(x) &= x_2 - 4x_1^4 + 32x_1^3 - 88x_1^2 + 96x_1 - 36,\\
x &\in \{(x_1, x_2) | 0 \leqslant x_1 \leqslant 3, 0 \leqslant x_2 \leqslant 4\}.
\end{aligned}$$

This example is excerpted from [3] (Test Problem 9 in Section 4.10). Let $X = \{(x_1, x_2) \mid 0 \leqslant x_1 \leqslant 3, 0 \leqslant x_2 \leqslant 4\}$. The corresponding problem $[SP]_{\frac{1}{3}}$ is as follows:

$$\min_{x \in X} \varphi_{q,\varepsilon,\frac{1}{3}}(x),$$

where $\varphi_{q,\varepsilon,\frac{1}{3}}(x) = f(x) + q\sum_{i=1}^{2} p_{q,\varepsilon,k}(g_i(x))$ and for $i = 1, 2$,

$$\begin{aligned}
&p_{q,\varepsilon,\frac{1}{3}}(g_i(x)) \qquad\qquad (2.1)\\
= &\begin{cases}
0, & \text{if} \quad g_i(x) < -(\frac{\varepsilon}{2q})^3;\\
(\frac{1}{6})^{\frac{1}{3}}(\frac{\varepsilon}{2q})^{-5}(g_i(x))^2 + \frac{1}{3}(\frac{1}{6})^{-\frac{2}{3}}(\frac{\varepsilon}{2q})^{-2}g_i(x) + (\frac{1}{6})^{\frac{1}{3}}\frac{\varepsilon}{2q}, & \text{if} \ -(\frac{\varepsilon}{2q})^3 \leqslant g_i(x) < 0;\\
(g_i(x) + \frac{1}{6}(\frac{\varepsilon}{2q})^3)^{\frac{1}{3}}, & \text{if} \quad g_i(x) \geqslant 0.
\end{cases}
\end{aligned}$$

We take the initial point $x_1^0 = (0, 0)$ and let $X = \{(x_1, x_2) | 0 \leqslant x_1 \leqslant 3, 0 \leqslant x_2 \leqslant 4\}$. The global optimizer in each iteration x_n^* is obtained via the quasi-filled function method proposed in [7]. Table 1 gives the numerical results of solving Example 3.1 by the proposed Algorithm (SP).

By Table 1, we know that $x_3^* = (2.329521, 3.178489)$ can be regarded as an approximate global minimizer of Example 3.1 with function value -5.508010. The global minimizer given in [3] is $(2.3295, 3.17846)$ with associated function value -5.50796.

Table 1: Numerical results for Example 2.1

n	x_n^0	x_n^*	q_n	ε_n	$f(x_n^*)$	$\begin{pmatrix} g_1(x_n^*) \\ g_2(x_n^*) \end{pmatrix}$
1	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 2.329521 \\ 3.178489 \end{pmatrix}$	10^2	1×10^{-2}	-5.508010	$\begin{pmatrix} -8.489793e-1 \\ -2.790287e-7 \end{pmatrix}$
2	$\begin{pmatrix} 2.329521 \\ 3.178489 \end{pmatrix}$	$\begin{pmatrix} 2.329521 \\ 3.178489 \end{pmatrix}$	10^4	1×10^{-3}	-5.508010	$\begin{pmatrix} -8.489791e-1 \\ -1.136868e-13 \end{pmatrix}$
3	$\begin{pmatrix} 2.329521 \\ 3.178489 \end{pmatrix}$	$\begin{pmatrix} 2.329521 \\ 3.178489 \end{pmatrix}$	10^8	1×10^{-4}	-5.508010	$\begin{pmatrix} -8.489791e-1 \\ -5.684342e-14 \end{pmatrix}$

Example 2.2

$$\begin{aligned} \min \quad f(x) &= -2x_1 - 6x_2 + x_1^2 - 2x_1x_2 + 2x_2^2, \\ \text{s.t.} \quad g_1(x) &= x_1 + x_2 - 2 \leqslant 0, \\ g_2(x) &= -x_1 + 2x_2 - 2 \leqslant 0, \\ x &\in \{(x_1, x_2) \mid 0 \leqslant x_i \leqslant 2, i = 1, 2\}. \end{aligned}$$

This example is excerpted from [2] (p. 504). Let $X = \{(x_1, x_2) \mid 0 \leqslant x_i \leqslant 2, i = 1, 2\}$. The corresponding problem $[SP]_{\frac{1}{4}}$ is as follows:

$$\min_{x \in X} \varphi_{q,\varepsilon,\frac{1}{4}}(x),$$

where $\varphi_{q,\varepsilon,\frac{1}{4}}(x) = f(x) + q \sum_{i=1}^{2} p_{q,\varepsilon,\frac{1}{4}}(g_i(x))$ with

$$\begin{aligned} & p_{q,\varepsilon,\frac{1}{4}}(g_i(x)) \\ = & \begin{cases} 0, & \text{if} \quad g_i(x) < -(\frac{\varepsilon}{2q})^4, \\ (\frac{1}{8})^{\frac{1}{4}}(\frac{\varepsilon}{2q})^{-7}(g_i(x))^2 + \frac{1}{4}(\frac{1}{8})^{-\frac{3}{4}}(\frac{\varepsilon}{2q})^{-3}g_i(x) + (\frac{1}{8})^{\frac{1}{4}}\frac{\varepsilon}{2q}, & \text{if} \ -(\frac{\varepsilon}{2q})^4 \leqslant g_i(x) < 0, \\ (g_i(x) + \frac{1}{8}(\frac{\varepsilon}{2q})^4)^{\frac{1}{4}}, & \text{if} \quad g_i(x) \geqslant 0. \end{cases} \end{aligned} \tag{2.2}$$

We take the initial point $x_1^0 = (0, 0)$. The global minimizer x_n^* is obtained by the quasi-filled function method proposed in [7]. Table 2 gives the result of solving Example 3.2 by the proposed Algorithm (SP).

From Table 2, we know that $x_3^* = (8.004453e-1, 1.199555)$ can be regarded as an approximate global minimizer of Example 3.2 with global optimal value $f^* = -7.199999$. The global minimizer of the problem given in [2] is $(0.8, 1.2)$ with the associated function value -7.2.

From the above two examples, we can see that the present method is quite effective.

Table 2: Numerical results for Example 2.2

n	x_n	x_n^*	q_n	ε_n	$f(x_n^*)$	$\begin{pmatrix} g_1(x_n^*) \\ g_2(x_n^*) \end{pmatrix}$
1	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 8.004707e-1 \\ 1.199529 \end{pmatrix}$	10^2	10^{-2}	-7.199999	$\begin{pmatrix} -3.250733e-13 \\ -4.014121e-1 \end{pmatrix}$
2	$\begin{pmatrix} 8.004707e-1 \\ 1.199529 \end{pmatrix}$	$\begin{pmatrix} 8.004710e-1 \\ 1.199529 \end{pmatrix}$	10^4	10^{-3}	-7.199999	$\begin{pmatrix} -1.776357e-15 \\ -4.014130e-1 \end{pmatrix}$
3	$\begin{pmatrix} 8.004710e-1 \\ 1.199529 \end{pmatrix}$	$\begin{pmatrix} 8.004453e-1 \\ 1.199555 \end{pmatrix}$	10^8	10^{-4}	-7.199999	$\begin{pmatrix} -4.440892e-16 \\ -4.013358e-1 \end{pmatrix}$

References

[1] Bai F S, Wu Z Y, Zhu D L. Lower order calmness and exact penalty function [J]. *Optimization Methods and Software,* 2006, **21**(4): 515-525.

[2] Bazaraa M S, Sherali H D, Shetty C M. Nonlinear Programming Theory and Algorithms [M]. New York: John Wiley & Sons Inc, 1993.

[3] Floudas C A, Pardalos M, Adjiman C S, Esposito W R, et al. Handbook of Test Problems in Local and Global Optimization [M]. Dordrecht: Kluwer Academic Publishers, 1999.

[4] He Z H, Bai F S. A smoothing approximation to the lower order exact penalty function [J]. *Operations Research Transactions*, 2010, **14**(2): 11-22.

[5] Meng Z Q, Dang C Y, Yang X Q. On the smoothing of the square-root exact penalty function for inequality constrained optimization [J]. *Computational Optimization and Applications,* 2006, **35**(3): 375-398.

[6] Pinar M C, Zenios S A. On smoothing exact functions for convex constrained optimization [J]. *SIAM Journal on Optimization,* 1994, **4**(3): 486-511.

[7] Wu Z Y, Lee H W J, Zhang L S, et al. A novel filled function method and quasi-filled function method for global optimization [J]. *Computational Optimization and Applications*, 2006, **34**(2): 249-272.

[8] Wu Z Y, Bai F S, Yang X Q, et al. An exact lower order penalty function and its smoothing in nonlinear Programming [J]. *Optimization*, 2004, **53**(1): 51-68.

[9] Wu Z Y, Lee H W J, Bai F S, et al. Quadratic smoothing approximation to l_1 exact penalty function in global optimization [J]. *Journal Of Industral and Management Optimization*, 2005, **1**(4): 533-547.

[10] Yang X Q, Meng Z Q, Huang X X, et al. Smoothing nonlinear penalty functions for constrained optimization [J]. *Numerical Functional Analysis and Optimization*, 2003, **24**(3 & 4): 351-364.