Mathematical and Computational Applications

MDPI

Article

A New Smoothing Nonlinear Penalty Function for Constrained Optimization

Touna Yang [1], Nguyen Thanh Binh [2,3,*], Tran Manh Thang [4] and Duong Thi Hoa [3]

[1] School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China; touna1988@yahoo.com
[2] Department of Mathematics, Shanghai University, Shanghai 200444, China
[3] Yen Bai Teacher's Training College, Yen Bai City 320000, Vietnam; dhoayb@gmail.com
[4] Department of Education and Training Yen Bai, Yen Bai City 320000, Vietnam; tmthang@yenbai.edu.vn
* Correspondence: bingbongyb@gmail.com; Tel.: +86-150-2134-6869

Academic Editor: Fazal M. Mahomed
Received: 5 November 2016; Accepted: 4 April 2017; Published: 12 April 2017

Abstract: In this study, a new smoothing nonlinear penalty function for constrained optimization problems is presented. It is proved that the optimal solution of the smoothed penalty problem is an approximate optimal solution of the original problem. Based on the smoothed penalty function, we develop an algorithm for finding an optimal solution of the optimization problems with inequality constraints. We further discuss the convergence of this algorithm and test this algorithm with three numerical examples. The numerical examples show that the proposed algorithm is feasible and effective for solving some nonlinear constrained optimization problems.

Keywords: penalty function; smoothing method; constrained optimization

1. Introduction

Consider the following constrained optimization problem:

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

where the functions $f,\ g_i : \mathbb{R}^n \to \mathbb{R},\ i \in I = \{1, 2, \ldots, m\}$, are continuously differentiable functions.

Let $X_0 = \{x \in \mathbb{R}^n \mid g_i(x) \le 0,\ i \in I\}$ be the feasible solution set and we assume that X_0 is not empty.

For a general constrained optimization problem, the penalty function method has attracted many researchers in both theoretical and practical aspects. However, to obtain an optimal solution for the original problem, the conventional quadratic penalty function method usually requires that the penalty parameter tends to infinity, which is undesirable in practical computation. In order to overcome the drawbacks of the quadratic penalty function method, exact penalty functions were proposed to solve problem (P). Zangwill [1] first proposed the l_1 exact penalty function

$$F^1(x, \rho) = f(x) + \rho \sum_{i=1}^{m} g_i^+(x), \tag{1}$$

where $\rho > 0$ is a penalty parameter, and $g_i^+(x) = \max\{0, g_i(x)\},\ i = 1, 2, \ldots, m$. It was proved that there exists a fixed constant $\rho_0 > 0$, for any $\rho > \rho_0$, and any global solution of the exact penalty problem is also a global solution of the original problem. Therefore, the exact penalty function methods have been widely used for solving constrained optimization problems (see, e.g., [2–9]).

Recently, the nonlinear penalty function of the following form has been investigated in [10–13]:

$$F^k(x,\rho) = \left[f(x)^k + \rho \sum_{i=1}^{m} (g_i^+(x))^k \right]^{\frac{1}{k}}, \tag{2}$$

where $f(x)$ is assumed to be positive and $k \in (0, +\infty)$. It is called the k-th power penalty function in [14,15]. Obviously, if $k = 1$, the nonlinear penalty function $F^k(x,\rho)$ is reduced to the l_1 exact penalty function. In [12], it was shown that the exact penalty parameter corresponding to $k \in (0, 1]$ is substantially smaller than that of the l_1 exact penalty function. Rubinov and Yang [13] also studied a penalty function as follows:

$$\psi_\rho^k(x,c) = [f(x) - c]^k + \rho \sum_{i=1}^{m} g_i^+(x), \tag{3}$$

where $c \in \mathbb{R}$ such that $f(x) - c > 0$ for any $x \in \mathbb{R}^n$, and $k \in (0, \ +\infty)$. The corresponding penalty problem of (P) is defined as

$$(P_\rho) \qquad \min \ \psi_\rho^k(x,c) \qquad \text{s.t. } x \in \mathbb{R}^n.$$

In fact, the original problem (P) is equivalent to the problem as follows:

$$\begin{aligned} (P') \qquad & \min \ [f(x) - c]^k \\ & \text{s.t. } g_i(x) \le 0, \ i = 1, 2, \ldots, m, \\ & \quad\ x \in \mathbb{R}^n. \end{aligned}$$

Obviously, the penalty problem (P_ρ) is the l_1 exact penalty problem of problem (P') defined as (1).

It is noted that these penalty functions $F^1(x,\rho)$, $F^k(x,\rho)$ $(0 < k \le 1)$ and $\psi_\rho^k(x,c)$ are not differentiable at x such that $g_i(x) = 0$ for some $i \in I$, which prevents the use of gradient-based methods and causes some numerical instability problems in its implementation, when the value of the penalty parameter becomes large [3,5,6,8]. In order to use existing gradient-based algorithms, such as a Newton method, it is necessary to smooth the exact penalty function. Thus, the smoothing of the exact penalty function attracts much attention [16–24]. Pinar and Zenios [21] and Wu et al. [22] discussed a quadratic smoothing approximation to nondifferentiable exact penalty functions for constrained optimization. Binh [17] and Xu et al. [23] proposed a second-order differentiability technique to the l_1 exact penalty function. It is shown that the optimal solution of the smoothed penalty problem is an approximate optimal solution of the original optimization problem. Zenios et al. [24] discussed an algorithm for the solution of large-scale optimization problems.

In this study, we aim to develop the smoothing technique for the nonlinear penalty function (3). First, we define the following smoothing function $q_{\epsilon,\rho}^k(t)$ by

$$q_{\epsilon,\rho}^k(t) = \begin{cases} 0 & \text{if } \ t \le 0, \\ \dfrac{2m^2\rho^2}{9\epsilon^2} t^{3k} & \text{if } \ 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ t^k + \dfrac{\epsilon}{3m\rho} e^{-\frac{m\rho}{\epsilon} t^k + 1} - \dfrac{10\epsilon}{9m\rho} & \text{if } \ t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

where $0 < k < +\infty$, $\rho > 0$ and $\epsilon > 0$. By considering this smoothing function, a new smoothing nonlinear penalty function is obtained. We use this smoothing nonlinear penalty function that is able to convert a constrained optimization problem into minimizations of a sequence of continuously differentiable functions and propose a corresponding algorithm for solving constrained optimization problems.

The rest of this paper is organized as follows. In Section 2, we propose a new smoothing penalty function for inequality constrained optimization problems, and some fundamental properties of its are proved. In Section 3, an algorithm based on the smoothed penalty function is presented and its global convergence is proved. In Section 4, we report results on application of this algorithm to three test problems and compare the results obtained with other similar algorithms. Finally, conclusions are discussed in Section 5.

2. Smoothing Nonlinear Penalty Functions

In this section, we first construct a new smoothing function. Then, we introduce our smoothing nonlinear penalty function and discuss its properties.

Let $q^k(t) : \mathbb{R} \to \mathbb{R}$ be as follows:

$$q^k(t) = \begin{cases} 0 & \text{if} \quad t \leq 0, \\ t^k & \text{if} \quad t \geq 0, \end{cases}$$

where $0 < k < +\infty$. Obviously, the function $q^k(t)$ is C^1 on $\mathbb{R}$ for $k > 1$, but it is not C^1 for $0 < k \leq 1$. It is useful in defining exact penalty functions for constrained optimization problems (see, e.g., [14,15,21]). Consider the nonlinear penalty function

$$F^k_\rho(x, c) = [f(x) - c]^k + \rho \sum_{i=1}^{m} q^k(g_i(x)), \tag{4}$$

and the corresponding penalty problem

$$(NP_\rho) \qquad \min \ F^k_\rho(x, c) \qquad \text{s.t.} \ x \in \mathbb{R}^n.$$

As previously mentioned, for any $\epsilon > 0$ and $\rho > 0$, the function $q^k_{\epsilon,\rho}(t)$ is defined as:

$$q^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \leq 0, \\ \dfrac{2m^2\rho^2}{9\epsilon^2} t^{3k} & \text{if} \quad 0 \leq t \leq \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ t^k + \dfrac{\epsilon}{3m\rho} e^{-\frac{m\rho}{\epsilon} t^k + 1} - \dfrac{10\epsilon}{9m\rho} & \text{if} \quad t \geq \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

where $0 < k < +\infty$.

Figure 1 shows the behavior of $q^k(t)$ and $q^k_{\epsilon,\rho}(t)$.

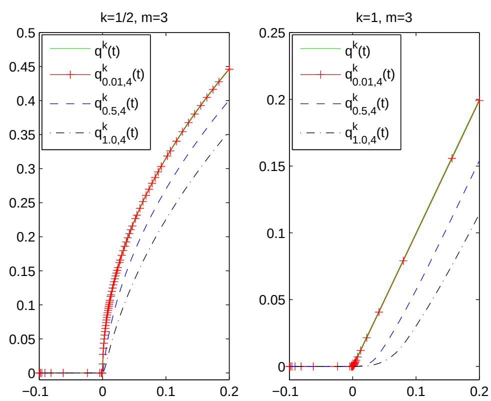

Figure 1. The behavior of $q^k(t)$ and $q^k_{\epsilon,\rho}(t)$.

In the following, we discuss the properties of $q^k_{\epsilon,\rho}(t)$.

Lemma 1. *For $0 < k < +\infty$ and any $\epsilon > 0$, we have*

(i) $q^k_{\epsilon,\rho}(t)$ *is continuously differentiable for* $k > \frac{1}{3}$ *on* $\mathbb{R}$, *where*

$$[q^k_{\epsilon,\rho}(t)]' = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \dfrac{2km^2\rho^2}{3\epsilon^2} t^{3k-1} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ kt^{k-1} - \dfrac{1}{3}kt^{k-1}e^{-\frac{m\rho}{\epsilon}t^k+1} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

(ii) $\lim_{\epsilon \to 0} q^k_{\epsilon,\rho}(t) = q^k(t)$.

(iii) $q^k(t) \ge q^k_{\epsilon,\rho}(t), \ \forall t \in \mathbb{R}$.

Proof. (i) First, we prove that $q^k_{\epsilon,\rho}(t)$ is continuous. Obviously, the function $q^k_{\epsilon,\rho}(t)$ is continuous at any $t \in \mathbb{R} \backslash \left\{0, \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right\}$. We only need to prove that $q^k_{\epsilon,\rho}(t)$ continuous at the separating points: 0 and $\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$.

(1) For $t = 0$, we have

$$\lim_{t \to 0^-} q^k_{\epsilon,\rho}(t) = \lim_{t \to 0^-} 0 = 0, \quad \lim_{t \to 0^+} q^k_{\epsilon,\rho}(t) = \lim_{t \to 0^+} \frac{2m^2\rho^2}{9\epsilon^2} t^{3k} = 0,$$

which implies

$$\lim_{t \to 0^-} q^k_{\epsilon,\rho}(t) = \lim_{t \to 0^+} q^k_{\epsilon,\rho}(t) = 0 = q^k_{\epsilon,\rho}(0).$$

Thus, $q^k_{\epsilon,\rho}(t)$ is continuous at $t = 0$.

(2) For $t = \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$, we have

$$\lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^-} q^k_{\epsilon,\rho}(t) = \lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^-} \frac{2m^2\rho^2}{9\epsilon^2} t^{3k} = \frac{2\epsilon}{9m\rho},$$

$$\lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^+} q^k_{\epsilon,\rho}(t) = \lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^+} \left[t^k + \frac{\epsilon}{3m\rho}e^{-\frac{m\rho}{\epsilon}t^k+1} - \frac{10\epsilon}{9m\rho}\right] = \frac{2\epsilon}{9m\rho},$$

which implies

$$\lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^-} q^k_{\epsilon,\rho}(t) = \lim_{t \to \left[\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right]^+} q^k_{\epsilon,\rho}(t) = \frac{2\epsilon}{9m\rho} = q^k_{\epsilon,\rho}\left(\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}\right).$$

Thus, $q^k_{\epsilon,\rho}(t)$ is continuous at $t = \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$.

Next, we will show that $q^k_{\epsilon,\rho}(t)$ is continuously differentiable, i.e., $[q^k_{\epsilon,\rho}(t)]'$ is continuous. Actually, we only need to prove that $[q^k_{\epsilon,\rho}(t)]'$ is continuous at the separating points: 0 and $\left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$.

(1) For $t = 0$, we have

$$\lim_{t \to 0^-} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to 0^-} 0 = 0, \quad \lim_{t \to 0^+} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to 0^+} \frac{2km^2\rho^2}{3\epsilon^2} t^{3k-1} = 0,$$

which implies

$$\lim_{t \to 0^-} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to 0^+} [q^k_{\epsilon,\rho}(t)]' = 0 = [q^k_{\epsilon,\rho}(0)]'.$$

Thus, $[q^k_{\epsilon,\rho}(t)]'$ is continuous at $t = 0$.

(2) For $t = (\frac{\epsilon}{m\rho})^{\frac{1}{k}}$, we have

$$\lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-} \frac{2km^2\rho^2}{3\epsilon^2} t^{3k-1} = \frac{2k}{3}\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}},$$

$$\lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+} \left(kt^{k-1} - \frac{1}{3}kt^{k-1}e^{-\frac{m\rho}{\epsilon}t^k+1}\right) = \frac{2k}{3}\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}},$$

which implies

$$\lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^-} [q^k_{\epsilon,\rho}(t)]' = \lim_{t \to \left[(\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right]^+} [q^k_{\epsilon,\rho}(t)]' = \frac{2k}{3}\left(\frac{\epsilon}{m\rho}\right)^{\frac{k-1}{k}} = \left[q^k_{\epsilon,\rho}\left((\frac{\epsilon}{m\rho})^{\frac{1}{k}}\right)\right]'.$$

Thus, $[q^k_{\epsilon,\rho}(t)]'$ is continuous at $t = (\frac{\epsilon}{m\rho})^{\frac{1}{k}}$.

(ii) For $\forall t \in \mathbb{R}$, by the definition of $q^k(t)$ and $q^k_{\epsilon,\rho}(t)$, we have

$$q^k(t) - q^k_{\epsilon,\rho}(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ t^k - \dfrac{2m^2\rho^2}{9\epsilon^2} t^{3k} & \text{if} \quad 0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ \dfrac{10\epsilon}{9m\rho} - \dfrac{\epsilon}{3m\rho} e^{-\frac{m\rho}{\epsilon}t^k+1} & \text{if} \quad t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}. \end{cases}$$

When $0 \le t \le \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$, let $u = t^k$. Then, we have $0 \le u \le \frac{\epsilon}{m\rho}$. Consider the function:

$$G(u) = u - \frac{2m^2\rho^2}{9\epsilon^2} u^3, \quad 0 \le u \le \frac{\epsilon}{m\rho},$$

and we have

$$G'(u) = 1 - \frac{2m^2\rho^2}{3\epsilon^2} u^2, \quad 0 \le u \le \frac{\epsilon}{m\rho}.$$

Obviously, $G'(u) > 0$ for $0 \le u \le \frac{\epsilon}{m\rho}$. Moreover, $G(0) = 0$ and $G(\frac{\epsilon}{m\rho}) = \frac{7\epsilon}{9m\rho}$. Hence, we have

$$0 \le q^k(t) - q^k_{\epsilon,\rho}(t) \le \frac{7\epsilon}{9m\rho}.$$

When $t \ge \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}$, we have

$$0 < q^k(t) - q^k_{\epsilon,\rho}(t) = \frac{10\epsilon}{9m\rho} - \frac{\epsilon}{3m\rho} e^{-\frac{m\rho}{\epsilon}t^k+1} \le \frac{10\epsilon}{9m\rho}.$$

Thus, we have

$$0 \le q^k(t) - q^k_{\epsilon,\rho}(t) \le \frac{10\epsilon}{9m\rho}.$$

That is,

$$\lim_{\epsilon \to 0} q^k_{\epsilon,\rho}(t) = q^k(t).$$

(iii) For $\forall t \in \mathbb{R}$, from (ii), we have

$$q^k(t) - q^k_{\epsilon,\rho}(t) \ge 0,$$

which is $q^k(t) \geq q^k_{\epsilon,\rho}(t)$.
This completes the proof. □

In this study, we always assume that $c < 0$ and $|c|$ is large enough, such that $f(x) - c > 0$ for all $x \in \mathbb{R}^n$. Let

$$F^k_{\epsilon,\rho}(x,c) = [f(x) - c]^k + \rho \sum_{i=1}^{m} q^k_{\epsilon,\rho}(g_i(x)), \quad 0 < k < +\infty. \tag{5}$$

Then, $F^k_{\epsilon,\rho}(x,c)$ is continuously differentiable at any $x \in \mathbb{R}^n$ and is a smooth approximation of $F^k_\rho(x,c)$. We have the following smoothed penalty problem:

$$(NP_{\epsilon,\rho}) \qquad \min \ F^k_{\epsilon,\rho}(x,c) \qquad \text{s.t. } x \in \mathbb{R}^n.$$

Lemma 2. *We have that*

$$0 \leq F^k_\rho(x,c) - F^k_{\epsilon,\rho}(x,c) \leq \frac{10\epsilon}{9}, \quad 0 < k < +\infty \tag{6}$$

for any $x \in \mathbb{R}^n$, $\epsilon > 0$ and $\rho > 0$.

Proof. For any $x \in \mathbb{R}^n$, $\epsilon > 0$, $\rho > 0$, we have

$$F^k_\rho(x,c) - F^k_{\epsilon,\rho}(x,c) = \rho \sum_{i=1}^{m} \left(q^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x)) \right).$$

Note that

$$q^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x)) = \begin{cases} 0 & \text{if } \ g_i(x) \leq 0, \\ [g_i(x)]^k - \dfrac{2m^2\rho^2}{9\epsilon^2}[g_i(x)]^{3k} & \text{if } \ 0 \leq g_i(x) \leq \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \\ \dfrac{10\epsilon}{9m\rho} - \dfrac{\epsilon}{3m\rho} e^{-\frac{m\rho}{\epsilon}[g_i(x)]^k + 1} & \text{if } \ g_i(x) \geq \left(\frac{\epsilon}{m\rho}\right)^{\frac{1}{k}}, \end{cases}$$

for any $i \in I$.

By Lemma 1, we have

$$0 \leq \sum_{i=1}^{m} (q^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x))) \leq \frac{10\epsilon}{9\rho},$$

which implies

$$0 \leq \rho \sum_{i=1}^{m} (q^k(g_i(x)) - q^k_{\epsilon,\rho}(g_i(x))) \leq \frac{10\epsilon}{9}.$$

Hence,

$$0 \leq F^k_\rho(x,c) - F^k_{\epsilon,\rho}(x,c) \leq \frac{10\epsilon}{9}.$$

This completes the proof. □

Lemma 3. *Let x^* and x^*_ρ be optimal solutions of problem (P) and problem (NP_ρ), respectively. If x^*_ρ is a feasible solution to problem (P), then x^*_ρ is an optimal solution for problem (P).*

Proof. Under the given conditions, we have that

$$0 < [f(x^*_\rho) - c]^k = F^k_\rho(x^*_\rho, c) \leq F^k_\rho(x^*, c) = [f(x^*) - c]^k, \quad 0 < k < +\infty.$$

Therefore, $f(x^*_\rho) - c \le f(x^*) - c$, which is $f(x^*_\rho) \le f(x^*)$.
Since x^* is an optimal solution and x^*_ρ is feasible to problem (P), which is

$$f(x^*_\rho) \ge f(x^*).$$

Therefore, x^*_ρ is an optimal solution for problem (P).
This completes the proof. □

Theorem 1. *Let x^*_ρ and $x^*_{\epsilon,\rho}$ be the optimal solutions of problem (NP_ρ) and problem $(NP_{\epsilon,\rho})$, respectively, for some $\rho > 0$ and $\epsilon > 0$. Then, we have that*

$$0 \le F^k_\rho(x^*_\rho, c) - F^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}, c) \le \frac{10\epsilon}{9}, \quad 0 < k < +\infty. \tag{7}$$

Furthermore, if x^*_ρ satisfies the conditions of Lemma 3 and $x^*_{\epsilon,\rho}$ is feasible to problem (P), then $x^*_{\epsilon,\rho}$ is an optimal solution for problem (P).

Proof. By Lemma 2, for $\rho > 0$ and $\epsilon > 0$, we obtain

$$\begin{aligned} 0 &\le F^k_\rho(x^*_\rho, c) - F^k_{\epsilon,\rho}(x^*_\rho, c) \\ &\le F^k_\rho(x^*_\rho, c) - F^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}, c) \\ &\le F^k_\rho(x^*_{\epsilon,\rho}, c) - F^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}, c) \\ &\le \frac{10\epsilon}{9}. \end{aligned}$$

That is,

$$0 \le F^k_\rho(x^*_\rho, c) - F^k_{\epsilon,\rho}(x^*_{\epsilon,\rho}, c) \le \frac{10\epsilon}{9},$$

and

$$\begin{aligned} 0 \le& \left\{ [f(x^*_\rho) - c]^k + \rho \sum_{i=1}^m q^k(g_i(x^*_\rho)) \right\} - \left\{ [f(x^*_{\epsilon,\rho}) - c]^k + \rho \sum_{i=1}^m q^k_{\epsilon,\rho}(g_i(x^*_{\epsilon,\rho})) \right\} \\ \le& \frac{10\epsilon}{9}. \end{aligned} \tag{8}$$

From the definition of $q^k(t)$, $q^k_{\epsilon,\rho}(t)$ and the fact that x^*_ρ, $x^*_{\epsilon,\rho}$ are feasible for problem (P), we have

$$\sum_{i=1}^m q^k(g_i(x^*_\rho)) = \sum_{i=1}^m q^k_{\epsilon,\rho}(g_i(x^*_{\epsilon,\rho})) = 0.$$

Note that $f(x^*_{\epsilon,\rho}) - c > 0$, and from (8), we have

$$0 < [f(x^*_{\epsilon,\rho}) - c]^k \le [f(x^*_\rho) - c]^k.$$

Therefore, $f(x^*_{\epsilon,\rho}) - c \le f(x^*_\rho) - c$, which is $f(x^*_{\epsilon,\rho}) \le f(x^*_\rho)$.
As $x^*_{\epsilon,\rho}$ is feasible to (P) and by Lemma 3, x^*_ρ is an optimal solution to (P), we have

$$f(x^*_{\epsilon,\rho}) \ge f(x^*_\rho).$$

Thus, $x^*_{\epsilon,\rho}$ is an optimal solution for problem (P).
This completes the proof. □

Definition 1. *A feasible solution x^* of problem (P) is called a KKT point, if there exists a $\mu^* \in \mathbb{R}^m$ such that the solution pair (x^*, μ^*) satisfies the following conditions:*

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*) = 0, \tag{9}$$

$$\mu_i^* g_i(x^*) = 0,\ g_i(x^*) \le 0,\ \mu_i^* \ge 0,\ \ i \in I. \tag{10}$$

Theorem 2. *Suppose the functions f, g_i $(i \in I)$ in problem (P) are convex. Let x^* and $x^*_{\epsilon,\rho}$ be the optimal solutions of problem (P) and problem $(NP_{\epsilon,\rho})$, respectively. If $x^*_{\epsilon,\rho}$ is feasible to problem (P), and there exists a $\mu^* \in \mathbb{R}^m$ such that the pair $(x^*_{\epsilon,\rho}, \mu^*)$ satisfies the conditions in Equations (9) and (10), then we have that*

$$0 \le [f(x^*_{\epsilon,\rho}) - c]^k - [f(x^*) - c]^k \le \frac{10\epsilon}{9}, \quad 0 < k < +\infty. \tag{11}$$

Proof. Since the functions f, g_i $(i \in I)$ are continuously differentiable and convex, we see that

$$f(x^*) \ge f(x^*_{\epsilon,\rho}) + \nabla f(x^*_{\epsilon,\rho})^T (x^* - x^*_{\epsilon,\rho}), \tag{12}$$

$$g_i(x^*) \ge g_i(x^*_{\epsilon,\rho}) + \nabla g_i(x^*_{\epsilon,\rho})^T (x^* - x^*_{\epsilon,\rho}), \quad i = 1, 2, \ldots, m. \tag{13}$$

After applying the conditions given in Equations (9), (10), (12) and (13), we see that

$$\begin{aligned} f(x^*) &\ge f(x^*_{\epsilon,\rho}) + \nabla f(x^*_{\epsilon,\rho})^T (x^* - x^*_{\epsilon,\rho}) \\ &= f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* \nabla g_i(x^*_{\epsilon,\rho})^T (x^* - x^*_{\epsilon,\rho}) \\ &\ge f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* \left[g_i(x^*) - g_i(x^*_{\epsilon,\rho}) \right] \\ &= f(x^*_{\epsilon,\rho}) - \sum_{i=1}^{m} \mu_i^* g_i(x^*) \\ &\ge f(x^*_{\epsilon,\rho}). \end{aligned}$$

Therefore, $f(x^*) - c \ge f(x^*_{\epsilon,\rho}) - c > 0$. Thus,

$$[f(x^*_{\epsilon,\rho}) - c]^k \le [f(x^*) - c]^k \le [f(x^*) - c]^k + \rho \sum_{i=1}^{m} q^k(g_i(x^*)) = F^k_\rho(x^*, c).$$

By Lemma 2, we have

$$F^k_\rho(x^*, c) \le F^k_{\epsilon,\rho}(x^*, c) + \frac{10\epsilon}{9}.$$

It follows that

$$\begin{aligned} [f(x^*_{\epsilon,\rho}) - c]^k &\le F^k_{\epsilon,\rho}(x^*, c) + \frac{10\epsilon}{9} \\ &= [f(x^*) - c]^k + \rho \sum_{i=1}^{m} q^k_{\epsilon,\rho}(g_i(x^*)) + \frac{10\epsilon}{9} \\ &= [f(x^*) - c]^k + \frac{10\epsilon}{9}. \end{aligned} \tag{14}$$

Since $x^*_{\epsilon,\rho}$ is feasible to (P), which is

$$f(x^*) \le f(x^*_{\epsilon,\rho}),$$

then

$$f(x^*) - c \le f(x^*_{\epsilon,\rho}) - c,$$

and, by $f(x^*) - c > 0$, we have

$$0 < [f(x^*) - c]^k \le [f(x^*_{\epsilon,\rho}) - c]^k. \quad (15)$$

Combining Equations (14) and (15), we have that

$$[f(x^*) - c]^k \le [f(x^*_{\epsilon,\rho}) - c]^k \le [f(x^*) - c]^k + \frac{10\epsilon}{9},$$

which is

$$0 \le [f(x^*_{\epsilon,\rho}) - c]^k - [f(x^*) - c]^k \le \frac{10\epsilon}{9}, \quad 0 < k < +\infty.$$

This completes the proof. □

3. Algorithm

In this section, by considering the above smoothed penalty function, we propose an algorithm to find an optimal solution of problem (P), defined as Algorithm 1.

Definition 2. *For $\epsilon > 0$, a point $x^*_\epsilon \in X_0$ is called an ϵ-feasible solution to (P), if it satisfies $g_i(x^*_\epsilon) \le \epsilon, \quad \forall i \in I$.*

Algorithm 1: Algorithm for solving problem (P)

Step 1: Let the initial point x^0_1. Let $\epsilon_1 > 0,\ \rho_1 > 0,\ 0 < \gamma < 1,\ \beta > 1$ and choose a constant $c < 0$ such that $f(x) - c > 0,\ \forall x \in X_0$, let $j = 1$ and go to Step 2.
Step 2: Use x^0_j as the starting point to solve the following problem:

$$(NP_{\epsilon_j,\rho_j}) \qquad \min_{x\in\mathbb{R}^n} F^k_{\epsilon_j,\rho_j}(x,c) = [f(x) - c]^k + \rho_j \sum_{i=1}^{m} q^k_{\epsilon_j,\rho_j}(g_i(x)).$$

Let $x^*_{\epsilon_j,\rho_j}$ be an optimal solution of (NP_{ϵ_j,ρ_j}) (the solution of (NP_{ϵ_j,ρ_j}) we obtained by the BFGS method given in [25]).
Step 3: If $x^*_{\epsilon_j,\rho_j}$ is ϵ-feasible for problem (P), then the algorithm stops and $x^*_{\epsilon_j,\rho_j}$ is an approximate optimal solution of problem (P). Otherwise, let
$\rho_{j+1} = \beta\rho_j,\ \epsilon_{j+1} = \gamma\epsilon_j,\ x^0_{j+1} = x^*_{\epsilon_j,\rho_j}$ and $j = j + 1$. Then, go to Step 2.

Remark 1. *From $0 < \gamma < 1,\ \beta > 1$, we can easily see that as $j \to +\infty$, the sequence $\{\epsilon_j\} \to 0$ and the sequence $\{\rho_j\} \to +\infty$.*

Theorem 3. *For $k > \frac{1}{3}$, suppose that for $\epsilon \in (0, \epsilon_1]$ and $\rho \in [\rho_1, +\infty)$, the set*

$$\arg\min_{x\in\mathbb{R}^n} F^k_{\epsilon,\rho}(x,c) \ne \emptyset. \quad (16)$$

Let $\{x^*_{\epsilon_j,\rho_j}\}$ be the sequence generated by Algorithm 1. If $\lim_{\|x\|\to+\infty} f(x) = +\infty$ and the sequence $\{F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c)\}$ is bounded, then $\{x^*_{\epsilon_j,\rho_j}\}$ is bounded and the limit point of $\{x^*_{\epsilon_j,\rho_j}\}$ is the solution of (P).

Proof. First, we prove that $\{x^*_{\epsilon_j,\rho_j}\}$ is bounded. Note that

$$F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c) = [f(x^*_{\epsilon_j,\rho_j}) - c]^k + \rho_j \sum_{i=1}^{m} q^k_{\epsilon_j,\rho_j}(g_i(x^*_{\epsilon_j,\rho_j})), \quad j = 0, 1, \ldots. \quad (17)$$

From the definition of $q^k_{\epsilon,\rho}(t)$, we have

$$\rho_j \sum_{i=1}^{m} q^k_{\epsilon_j,\rho_j}(g_i(x^*_{\epsilon_j,\rho_j})) \geq 0. \tag{18}$$

Suppose, on the contrary, that the sequence $\{x^*_{\epsilon_j,\rho_j}\}$ is unbounded and without loss of generality $\|x^*_{\epsilon_j,\rho_j}\| \to +\infty$ as $j \to +\infty$, and $f(x^*_{\epsilon_j,\rho_j}) - c > 0$. Then, $\lim\limits_{j\to+\infty}[f(x^*_{\epsilon_j,\rho_j}) - c]^k = +\infty$, and from Equations (17) and (18), we have

$$F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c) \geq [f(x^*_{\epsilon_j,\rho_j}) - c]^k \to +\infty, \quad j = 0, 1, \ldots,$$

which contradicts with the sequence $\{F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c)\}$ being bounded. Thus, $\{x^*_{\epsilon_j,\rho_j}\}$ is bounded.

Next, we prove that the limit point of $\{x^*_{\epsilon_j,\rho_j}\}$ is the solution of problem (P). Let x^* be a limit point of $\{x^*_{\epsilon_j,\rho_j}\}$. Then, there exists the subset $J \subset \mathbb{N}$ such that $x^*_{\epsilon_j,\rho_j} \to x^*$ for $j \in J$, where $\mathbb{N}$ is the set of natural numbers. We have to show that x^* is an optimal solution of problem (P). Thus, it is sufficient to show (i) $x^* \in X_0$ and (ii) $f(x^*) \leq \inf_{x\in X_0} f(x)$.

(i) Suppose $x^* \notin X_0$. Then, there exists $\theta_0 > 0$ and the subset $J' \subset J$, such that $g_{i'}(x^*_{\epsilon_j,\rho_j}) \geq \theta_0 > 0$ for any $j \in J'$ and some $i' \in I$.

If $\theta_0 \leq g_{i'}(x^*_{\epsilon_j,\rho_j}) < \left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}}$, from the definition of $q^k_{\epsilon,\rho}(t)$ and $x^*_{\epsilon_j,\rho_j}$ is the optimal solution according j-th values of the parameters ϵ_j, ρ_j for any $x \in X_0$, we have

$$\begin{aligned}[f(x^*_{\epsilon_j,\rho_j}) - c]^k + \frac{2m^2\rho_j^3\theta_0^{3k}}{9\epsilon_j^2} &\leq F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c) \\ &\leq F^k_{\epsilon_j,\rho_j}(x, c) = [f(x) - c]^k,\end{aligned}$$

which contradicts with $\rho_j \to +\infty$ and $\epsilon_j \to 0$.

If $g_{i'}(x^*_{\epsilon_j,\rho_j}) \geq \theta_0 \geq \left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}}$ or $g_{i'}(x^*_{\epsilon_j,\rho_j}) \geq \left(\frac{\epsilon_j}{m\rho_j}\right)^{\frac{1}{k}} \geq \theta_0$, from the definition of $q^k_{\epsilon,\rho}(t)$ and $x^*_{\epsilon_j,\rho_j}$ is the optimal solution according j-th values of the parameters ϵ_j, ρ_j for any $x \in X_0$, we have

$$\begin{aligned}[f(x^*_{\epsilon_j,\rho_j}) - c]^k + \rho_j\theta_0^k + \frac{\epsilon_j}{3m}e^{-\frac{m\rho_j}{\epsilon_j}\theta_0^k+1} - \frac{10\epsilon_j}{9m} &\leq F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c) \\ &\leq F^k_{\epsilon_j,\rho_j}(x, c) = [f(x) - c]^k,\end{aligned}$$

which contradicts with $\rho_j \to +\infty$ and $\epsilon_j \to 0$.

Thus, $x^* \in X_0$.

(ii) For any $x \in X_0$, we have

$$[f(x^*_{\epsilon_j,\rho_j}) - c]^k \leq F^k_{\epsilon_j,\rho_j}(x^*_{\epsilon_j,\rho_j}, c) \leq F^k_{\epsilon_j,\rho_j}(x, c) = [f(x) - c]^k.$$

We know that $f(x^*_{\epsilon_j,\rho_j}) - c > 0$, so $f(x^*_{\epsilon_j,\rho_j}) - c \leq f(x) - c$. Therefore, $f(x^*) \leq \inf_{x\in X_0} f(x)$ holds. This completes the proof. □

4. Numerical Examples

In this section, we apply the Algorithm 1 to three test problems. The proposed algorithm is implemented in Matlab (R2011A, The MathWorks Inc., Natick, MA, USA).

In each example, we take $\epsilon = 10^{-6}$. Then, it is expected to get an ϵ-solution to problem (P) with Algorithm 1, and the numerical results are presented in the following tables.

Example 1. *Consider the following problem ([20], Example 4.1)*

$$\begin{aligned} \min \;& f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\ s.t. \;& g_1(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1 + x_2 + x_4 - 5 \le 0, \\ & g_2(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 \le 0, \\ & g_3(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \le 0. \end{aligned}$$

For $k = \frac{2}{3}$, let $x_1^0 = (0,\ 0,\ 0,\ 0)$, $\rho_1 = 6$, $\beta = 10$, $\epsilon_1 = 0.01$, $\gamma = 0.01$ and choose $c = -100$. The results are shown in Table 1.

For $k = 1$, let $x_1^0 = (5,\ 5,\ 5,\ 5)$, $\rho_1 = 10$, $\beta = 4$, $\epsilon_1 = 0.01$, $\gamma = 0.1$ and choose $c = -100$. The results are shown in Table 2.

The results in Tables 1 and 2 show that the convergence of Algorithm 1 and the objective function values are almost the same. By Table 1, we obtain that an approximate optimal solution $x^* = (0.166332,\ 0.828748,\ 2.013798,\ -0.959021)$ after two iterations with function value $f(x^*) = -44.233325$. In [20], the obtained approximate optimal solution is $x^* = (0.170056,\ 0.841066,\ 2.004907,\ -0.968785)$ with function value $f(x^*) = -44.225989$. Numerical results obtained by our algorithm are slightly better than the results in [20].

Table 1. Results of Algorithm 1 with $k = \frac{2}{3}$, $x_1^0 = (0,0,0,0)$ for Example 1.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$
1	6	0.01	(0.214257, 0.952510, 1.934008, −1.051001)	−44.266455	0.069500	0.044928	−1.353205
2	60	0.0001	(0.166332, 0.828748, 2.013798, −0.959021)	−44.233325	−0.000071	−0.000006	−1.911178

Table 2. Results of Algorithm 1 with $k = 1$, $x_1^0 = (5,5,5,5)$ for Example 1.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$	$g_3(x^*_{\epsilon_j,\rho_j})$
1	10	0.01	(0.049915, 1.064341, 1.936277, −1.083612)	−44.010514	−0.032469	0.064160	−0.600569
2	40	0.001	(0.161485, 0.837434, 2.012060, −0.962236)	−44.233462	0.000003	0.000003	−1.870398
3	160	0.0001	(0.169837, 0.833494, 2.009512, −0.963710)	−44.233813	0.000000	0.000000	−1.892242

Example 2. *Consider the following problem ([22], Example 3.2)*

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

For $k = \frac{3}{4}$, let $x_1^0 = (3,\ 1)$, $\rho_1 = 5$, $\beta = 10$, $\epsilon_1 = 0.1$, $\gamma = 0.1$, and choose $c = -10$. The results are shown in Table 3.

For $k = 1$, let $x_1^0 = (0,\ 1)$, $\rho_1 = 6$, $\beta = 10$, $\epsilon_1 = 0.02$, $\gamma = 0.01$ and choose $c = -10$. The results are shown in Table 4.

The results in Tables 3 and 4 show that the convergence of Algorithm 1 and the objective function values are almost the same. By Table 3, we obtain an approximate optimal solution is $x^* = (2.112103,\ 3.900086)$ after 2 iterations with function value $f(x^*) = -6.012190$. In [22], the obtained global solution is $x^* = (2.3295,\ 3.1784)$ with function value $f(x^*) = -5.5080$. Numerical results obtained by our algorithm are much better than the results in [22].

Table 3. Results of Algorithm 1 with $k = \frac{3}{4}$, $x_1^0 = (3, 1)$ for Example 2.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	5	0.1	(2.112050, 3.900242)	−6.012292	0.000039	0.000052
2	50	0.01	(2.112103, 3.900086)	−6.012190	−0.000020	−0.000008

Table 4. Results of Algorithm 1 with $k = 1$, $x_1^0 = (0, 1)$ for Example 2.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	6	0.02	(2.111875, 3.900776)	−6.012651	0.000232	0.000278
2	60	0.0002	(2.112763, 3.898814)	−6.011577	−0.000755	−0.000109

Example 3. *Consider the following problem ([26], Example 4.1)*

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2^2 - \cos(17x_1) - \cos(17x_2) + 3 \\ s.t. \quad & g_1(x) = (x_1 - 2)^2 + x_2^2 - 1.6^2 \le 0, \\ & g_2(x) = x_1^2 + (x_2 - 3)^2 - 2.7^2 \le 0, \\ & 0 \le x_1 \le 2, \\ & 0 \le x_2 \le 2. \end{aligned}$$

For $k = \frac{2}{3}$, let $x_1^0 = (0,\ 1)$, $\rho_1 = 1$, $\beta = 3$, $\epsilon_1 = 0.01$, $\gamma = 0.01$ and choose $c = -2$. The results are shown in Table 5.

For $k = \frac{3}{4}$, let $x_1^0 = (0,\ 0)$, $\rho_1 = 1$, $\beta = 9$, $\epsilon_1 = 0.01$, $\gamma = 0.01$, and choose $c = -2$. The results are shown in Table 6.

The results in Tables 5 and 6 show that the convergence of Algorithm 1 and the objective function values are almost the same. By Table 5, we obtain that an approximate optimal solution is $x^* = (0.725355,\ 0.399258)$ after two iterations with function value $f(x^*) = 1.837548$. In [26], the obtained approximate optimal solution is $x^* = (0.7255,\ 0.3993)$ with function value $f(x^*) = 1.8376$. Numerical results obtained by our algorithm are slightly better than the results in [26].

Table 5. Results of Algorithm 1 with $k = \frac{2}{3}$, $x_1^0 = (0, 1)$ for Example 3.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	1	0.01	(0.016363, 1.092060)	2.271514	2.567409	−3.649496
2	3	0.0001	(0.725355, 0.399258)	1.837548	−0.775873	−0.000000

Table 6. Results of Algorithm 1 with $k = \frac{3}{4}$, $x_1^0 = (0, 0)$ for Example 3.

j	ρ_j	ϵ_j	$x^*_{\epsilon_j,\rho_j}$	$f(x^*_{\epsilon_j,\rho_j})$	$g_1(x^*_{\epsilon_j,\rho_j})$	$g_2(x^*_{\epsilon_j,\rho_j})$
1	1	0.01	(0.387905, 1.086812)	2.448684	1.220019	−3.479255
2	9	0.0001	(0.725356, 0.399258)	1.837548	−0.775876	−0.000000

5. Conclusions

In this study, we have proposed a new smoothing approach to the nonsmooth penalty function and developed a corresponding algorithm to solve constrained optimization with inequality constraints. It is shown that any optimal solution of the smoothed penalty problem is shown to be an approximate optimal solution or a global solution of the original optimization problem. Furthermore, the numerical results given in Section 4 show that the Algorithm 1 has a good convergence for an approximate optimal solution.

Acknowledgments: The authors would like to express their gratitude to anonymous referees' detailed comments and remarks that help us improve our presentation of this paper considerably. This work is supported by grants from the National Natural Science Foundation of China (No. 11371242 and 61572099).

Author Contributions: All authors contributed equally to this work.

Conflicts of Interest: The authors declare no competing interests.

References

1. Zangwill, W.I. Nonlinear programming via penalty function. *Manag. Sci.* **1967**, *13*, 334–358.
2. Bazaraa, M.S.; Goode, J.J. Sufficient conditions for a globally exact penalty function without convexity. *Math. Program. Study* **1982**, *19*, 1–15.
3. Di Pillo, G.; Grippo, L. An exact penalty function method with global conergence properties for nonlinear programming problems. *Math. Program.* **1986**, *36*, 1–18.
4. Di Pillo, G.; Grippo, L. Exact penalty functions in constrained optimization. *SIAM J. Control. Optim.* **1989**, *27*, 1333–1360.
5. Han, S.P.; Mangasrian, O.L. Exact penalty function in nonlinear programming. *Math. Program.* **1979**, *17*, 251–269.
6. Lasserre, J.B. A globally convergent algorithm for exact penalty functions. *Eur. J. Oper. Res.* **1981**, *7*, 389–395.
7. Mangasarian, O.L. Sufficiency of exact penalty minimization. *SIAM J. Control. Optim.* **1985**, *23*, 30–37.
8. Rosenberg, E. Exact penalty functions and stability in locally Lipschitz programming. *Math. Program.* **1984**, *30*, 340–356.
9. Yu, C.J.; Teo, K.L.; Zhang, L.S.; Bai, Y.Q. A new exact penalty function method for continuous inequality constrained optimization problems. *J. Indus. Mgmt. Optimiz.* **2010**, *6*, 895–910.
10. Huang, X.X.; Yang, X.Q. Convergence analysis of a class of nonlinear penalization methods for constrained optimization via first-order necessary optimality conditions. *J. Optim. Theory Appl.* **2003**, *116*, 311–332.
11. Rubinov, A.M.; Glover, B.M.; Yang, X.Q. Extended Lagrange and penalty functions in continuous optimization. *Optimization* **1999**, *46*, 327–351.
12. Rubinov, A.M.; Yang, X.Q.; Bagirov, A.M. Penalty functions with a small penalty parameter. *Optim. Methods Softw.* **2002**, *17*, 931–964.
13. Rubinov, A.M.; Yang, X.Q. *Lagrange-Type Functions in Constrained Non-Convex Optimization*; Kluwer Academic Publishers: Dordrecht, Netherlands, 2003.
14. Binh, N.T.; Yan, W.L. Smoothing approximation to the k-th power nonlinear penalty function for constrained optimization problems. *J. Appl. Math. Bioinform.* **2015**, *5*, 1–19.
15. Yang, X.Q.; Meng, Z.Q.; Huang, X.X.; Pong, G.T.Y. Smoothing nonlinear penalty functions for constrained optimization. *Numer. Funct. Anal. Optimiz.* **2003**, *24*, 351–364.
16. Binh, N.T. Smoothing approximation to l_1 exact penalty function for constrained optimization problems. *J. Appl. Math. Inform.* **2015**, *33*, 387–399.
17. Binh, N.T. Second-order smoothing approximation to l_1 exact penalty function for nonlinear constrained optimization problems. *Theor. Math. Appl.* **2015**, *5*, 1–17.
18. Binh, N.T. Smoothed lower order penalty function for constrained optimization problems. *IAENG Int. J. Appl. Math.* **2016**, *46*, 76–81.
19. Chen, C.H.; Mangasarian, O.L. Smoothing methods for convex inequalities and linear complementarity problems. *Math. Program.* **1995**, *71*, 51–69.
20. Meng, Z.Q.; Dang, C.Y.; Jiang, M.; Shen, R. A smoothing objective penalty function algorithm for inequality constrained optimization problems. *Numer. Funct. Anal. Optimiz.* **2011**, *32*, 806–820.
21. Pinar, M.C.; Zenios, S.A. On smoothing exact penalty function for convex constrained optimization. *SIAM J. Optim.* **1994**, *4*, 486–511.
22. Wu, Z.Y.; Lee, H.W.J.; Bai, F.S.; Zhang, L.S. Quadratic smoothing approximation to l_1 exact penalty function in global optimization. *J. Ind. Manag. Optim.* **2005**, *1*, 533–547.
23. Xu, X.S.; Meng, Z.Q.; Sun, J.W.; Huang, L.G.; Shen, R. A second-order smooth penalty function algorithm for constrained optimization problems. *Comput. Optim. Appl.* **2013**, *55*, 155–172.
24. Zenios, S.A.; Pinar, M.C.; Dembo, R.S. A smooth penalty function algorithm for network-structured problems. *Eur. J. Oper. Res.* **1995**, *83*, 220–236.

25. Nocedal, J.; Wright, S.T. *Numerical Optimization*; Springer: New York, NY, USA, 1999.
26. Sun, X.L.; Li, D. Value-estimation function method for constrained global optimization. *J. Optim. Theory Appl.* **1999**, *102*, 385–409.