Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2012, Article ID 620949, 12 pages
doi:10.1155/2012/620949

Research Article

On the Convergence of a Smooth Penalty Algorithm without Computing Global Solutions

Bingzhuang Liu,[1] Changyu Wang,[2] and Wenling Zhao[1]

[1] *School of Science, Shandong University of Technology, Zibo 255049, China*
[2] *Institute of Operations Research, Qufu Normal University, Qufu 273165, China*

Correspondence should be addressed to Bingzhuang Liu, lbzlyj@126.com

Received 18 September 2011; Accepted 9 November 2011

Academic Editor: Yeong-Cheng Liou

We consider a smooth penalty algorithm to solve nonconvex optimization problem based on a family of smooth functions that approximate the usual exact penalty function. At each iteration in the algorithm we only need to find a stationary point of the smooth penalty function, so the difficulty of computing the global solution can be avoided. Under a generalized Mangasarian-Fromovitz constraint qualification condition (GMFCQ) that is weaker and more comprehensive than the traditional MFCQ, we prove that the sequence generated by this algorithm will enter the feasible solution set of the primal problem after finite times of iteration, and if the sequence of iteration points has an accumulation point, then it must be a Karush-Kuhn-Tucker (KKT) point. Furthermore, we obtain better convergence for convex optimization problem.

1. Introduction

Consider the following nonconvex optimization problem

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

where $f, g_i : R^n \to R,\ i = 1, \ldots, m$, are all continuously differentiable functions. Without loss of generality, we suppose throughout this paper that $\inf_{x\in R^n} f(x) \ge 0$, because otherwise we can substitute $f(x)$ by $\exp(f(x))$. Let $\Omega_\varepsilon = \{x \in R^n \mid g_i(x) \le \varepsilon,\ i = 1, \ldots, m\}$ be the relax feasible set for $\varepsilon > 0$. Then Ω_0 is the feasible set of (NP).

The classical l_1 exact penalty function [1] is

$$f_\beta(x) = f(x) + \beta \sum_{i=1}^{m} (g_i(x))^+, \tag{1.1}$$

where $\beta > 0$ is a penalty parameter, and

$$(g_i(x))^+ = \max\{0, g_i(x)\}, \quad i = 1, \ldots, m. \tag{1.2}$$

The obvious advantage of the traditional exact penalty functions such as the l_1 exact penalty function is that when the penalty parameter is sufficiently large, their global optimal solutions exist and are optimal solutions of (NP). But they also have obvious disadvantage, that is, their nonsmoothness, which prevent the use of many efficient unconstrained optimization algorithms (such as Gradient-type or Newton-type algorithm). Therefore the study on the smooth approximation of exact penalty functions has attracted broad interests in scholars [2–8]. In recent years based on the smooth approximation of the exact penalty function, several smooth penalty methods are given to solve (NP). For example, [9] gives a smooth penalty method based on approximating the l_1 exact penalty function. Under the assumptions that the optimal solution satisfies MFCQ and the iterate sequence is bounded, it is proved that the iterative sequence will enter the feasible set and every accumulation point is the optimal solution of (NP). In [10, 11], smooth penalty methods are considered based on approximating low-order exact penalty functions. Reference [10] proves the similar results as [9] under very strict conditions (some of them are uneasy to check). The conditions for convergence of the smooth penalty algorithm in [11] are weaker than that in [10], but in [11] it is only proved that the accumulation point of the iterate sequence is a Fritz-John (FJ) point of (NP).

In the algorithms given by [9–11], at each iteration a global optimal solution of the smooth penalty problem is needed. As we all know, it is very difficult to find a global optimal point of a nonconvex function. To avoid this difficulty, in this paper we give a smooth penalty algorithm based on the smooth approximation of the l_1 exact penalty function. The feature of this algorithm lies in that only a stationary point of the penalty function is needed to compute at each iteration. To prove the convergence of this algorithm, we first establish a generalized Mangasarian-Fromovitz constraint qualification condition (GMFCQ) weaker and more comprehensive than the traditional MFCQ. Under this condition, we prove that the iterative sequence of the algorithm will enter the feasible set of (NP). Moreover, we prove that if the iterative sequence has accumulation points, then each of them is a KKT point of (NP). Finally, we apply this algorithm to solve convex optimization and get better convergence results.

The rest of this paper is organized as follows. In the next section, we give a family of smooth penalty functions. In Section 3 based on the smooth penalty functions given in Section 2, we propose an algorithm for (NP) and analyze its convergence under the GMFCQ condition. We give an example that satisfies GMFCQ at last in this section.

2. Smooth Approximation to l_1 Exact Penalty Function

In this section we give a family of penalty functions, which decreasingly approximate the l_1 exact penalty function. At first we consider a class of smooth function $\phi : R \to R_+$ with the following properties:

(I) $\phi(\cdot)$ is a continuously differentiable convex function with $\phi'(0) > 0$;

(II) $\lim_{t\to-\infty}\phi(t) = a$, where a is a nonnegative constant;

(III) $\phi(t) \geq t$, for any $t > 0$;

(IV) $\lim_{t\to+\infty}(\phi(t))/t = 1$.

From (I)–(IV), it follows that ϕ satisfies

(V) $0 \leq \phi'(t) \leq 1$, for any $t \in R$, and $\lim_{t\to-\infty}\phi'(t) = 0$, $\lim_{t\to+\infty}\phi'(t) = 1$;

(VI) $r\phi(t/r)$ increases with respect to $r > 0$, for any $t \in R$;

(VII) $r\phi(t/r) \downarrow t^+ (r \downarrow 0)$, for any $t \in R$.

The following functions are often used in the smooth approximation of the l_1 exact penalty function and satisfy properties (I)–(IV).

(1) $\phi(t) = \log(1 + e^t)$.

(2) $\phi(t) = (t + \sqrt{t^2 + 4})/2$.

(3) $\phi(t) = \begin{cases} e^t, & t\leq 0; \\ t+1, & t>0. \end{cases}$

We now use $\phi(\cdot)$ to construct the smooth penalty function

$$f_{\beta,r}(x) = f(x) + r\sum_{i=1}^{m}\phi\left(\frac{\beta g_i(x)}{r}\right), \tag{2.1}$$

where $\beta \geq 1$ is a penalty parameter.

By (VII), we easily know when $r \to 0^+$, $f_{\beta,r}(x)$ decreasingly converges to $f_\beta(x)$, that is,

$$f_{\beta,r}(x) = f(x) + r\sum_{i=1}^{m}\phi\left(\frac{\beta g_i(x)}{r}\right) \downarrow f(x) + \beta\sum_{i=1}^{m}\left(g_i(x)\right)^+. \tag{2.2}$$

Therefore $f_{\beta,r}(x)$ smoothly approximates the l_1 exact penalty function, where r decreases to improve the precision of the approximation. It is worth noting that the smooth function $\phi(\cdot)$ and penalty function $f_{\beta,r}(\cdot)$ given in this paper make substantive improvement of the corresponding functions given in [9]. This gives $f_{\beta,r}(\cdot)$ better convergence properties (refer to (2.2) and Theorem 3.9).

3. The Algorithm and Its Convergence

We propose a penalty algorithm for (NP) in this section based on computing the stationary point of $f_{\beta,r}(\cdot)$. We assume that for any $\beta \geq 1$ and $0 < r \leq 1$, $f_{\beta,r}(\cdot)$ always has stationary point.

Algorithm

Step 0. Given $x^0 \in R^n,\ \beta_1 = 1,\ r_1 = 1,\ 0 < \eta_1 < 1$, and $\eta_2 > 1$. Let $k = 1$.

Step 1. Find x^k such that

$$\nabla f_{\beta_k, r_k}\left(x^k\right) = 0. \tag{3.1}$$

Step 2. Put $r_{k+1} = \eta_1 r_k$,

$$\beta_{k+1} = \begin{cases} \beta_k & \text{if } x^k \in \Omega_0, \\ \eta_2 \beta_k & \text{otherwise.} \end{cases} \tag{3.2}$$

Step 3. Let $k = k + 1$ and return to Step 1.

Let $\{x^k\}$ be the iterative sequence generated by the algorithm. We shall use the following assumption:

(A_1) the penalty function value sequence $\{f_{\beta_k, r_k}(x^k)\}$ is bounded.

Lemma 3.1. *Suppose that the assumption (A_1) holds, then for any $\varepsilon > 0$, there exists $k_0 \in N = \{1, 2, \ldots\}$, such that for $k \geq k_0$,*

$$x^k \in \Omega_\varepsilon. \tag{3.3}$$

Proof. Suppose to the contrary that there exist an $\varepsilon_0 > 0$ and an infinite sequence $K \subseteq N$, such that for any $k \in K$,

$$x^k \notin \Omega_{\varepsilon_0}. \tag{3.4}$$

By the algorithm, we know that

$$\beta_k \longrightarrow +\infty \quad (k \longrightarrow \infty). \tag{3.5}$$

It follows from (3.4) that there exist a subsequence $K_0 \subseteq K$ and an index $i_0 \in I = \{1, \ldots, m\}$, such that for any $k \in K_0$,

$$g_{i_0}\left(x^k\right) > \varepsilon_0. \tag{3.6}$$

Thus, from the assumptions about $f(\cdot)$, the properties about $\phi(\cdot)$, (3.5) and (3.6), it follows that

$$\begin{aligned} f_{\beta_k,r_k}\left(x^k\right) &= f\left(x^k\right) + r_k \sum_{i=1}^{m} \phi\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \\ &\geq f\left(x^k\right) + r_k \phi\left(\frac{\beta_k \varepsilon_0}{r_k}\right) \\ &\geq \left(\frac{\phi(\beta_k \varepsilon_0 / r_k)}{\beta_k \varepsilon_0 / r_k}\right) \beta_k \varepsilon_0 \\ &\longrightarrow +\infty \quad (k \longrightarrow \infty, k \in K_0). \end{aligned} \tag{3.7}$$

This contradicts with (A_1). □

Lemma 3.2. *Suppose that the assumption (A_1) holds, and x^* is any accumulation point of $\{x^k\}$, then $x^* \in \Omega_0$, that is, x^* is a feasible solution of* (NP).

Proof. By Lemma 3.1, we obtain that for any $\varepsilon > 0$ and every sufficiently large k, $x^k \in \Omega_\varepsilon$. Let x^* be an accumulation point of $\{x^k\}$, then there exists a subsequence $\{x^k\}_{k\in K}$ such that $x^k \to x^*(k \in K, k \to \infty)$. Therefore

$$x^* \in \Omega_\varepsilon. \tag{3.8}$$

By the arbitrariness of $\varepsilon > 0$, we have that $x^* \in \Omega_0$. □

Given $\overline{x} \in \Omega_0$, we denote that $I(\overline{x}) = \{i \in I \mid g_i(\overline{x}) = 0\}$.

Definition 3.3 (see [12]). We say that $\overline{x} \in \Omega_0$ satisfies MFCQ, if there exists a $h \in R^n$ such that

$$\nabla g_i(\overline{x})^T h < 0, \quad \text{for any } i \in I(\overline{x}). \tag{3.9}$$

In the following we propose a kind of generalized Mangasarian-Fromovitz constraint qualification (GMFCQ).

Let $K \subseteq N$ be a subsequence, and for sequence $\{z^k\}_{k\in K}$ in R^n denote two index sets as

$$\begin{aligned} I^+(K) &= \left\{ i \in I \limsup_{k\in K, k\to\infty} g_i\left(z^k\right) \geq 0 \right\}, \\ I^-(K) &= \left\{ i \in I \limsup_{k\in K, k\to\infty} g_i\left(z^k\right) < 0 \right\}. \end{aligned} \tag{3.10}$$

Definition 3.4. We say that the sequence $\{z^k\}_{k\in K}$ satisfies GMFCQ, if there exist a subsequence $K_0 \subseteq K$ and a vector $h \in R^n$ such that

$$\limsup_{k\in K_0, k\to\infty} \nabla g_i\left(z^k\right)^T h < 0, \quad \text{for any } i \in I^+(K_0). \tag{3.11}$$

Under some circumstances, the sequence $\{x^k\}$ may satisfy that $\|x^k\| \to +\infty$ $(k \to \infty)$, which can be seen for the example in the last part of this section. At this time MFCQ can not be applied, but GMFCQ can. The following proposition shows that Definition 3.4 is a substantive generalization of Definition 3.3.

Proposition 3.5. *Suppose that* $\{z^k\}_{k\in K}$ *satisfies*

$$\lim_{k\in K,\, k\to\infty} z_k = z^* \in \Omega_0. \tag{3.12}$$

If z^* *satisfies MFCQ, then* $\{z^k\}_{k\in K}$ *satisfies GMFCQ.*

Proof. By (3.12), we know that $\limsup_{k\in K, k\to\infty} g_i(z^k) \geq 0$ if and only if

$$\lim_{k\in K, k\to\infty} g_i\left(z^k\right) = g_i(z^*) = 0. \tag{3.13}$$

Thus, $I^+(K) = I(z^*)$. By the assumption, there exists a $h \in R^n$ such that

$$\limsup_{k\in K,\, k\to\infty} \nabla g_i\left(z^k\right)^T h < 0, \quad \text{for any } i \in I^+(K). \tag{3.14}$$

□

We need two assumptions in the following:

(A_2) the sequence $\{\nabla f(x^k)\}$ and $\{\nabla g_i(x^k)\}$, $i = 1, \ldots, m$ are both bounded;

(A_3) any subsequence of $\{x^k\}$ satisfies GMFCQ.

Theorem 3.6. *Suppose that the assumptions* (A_1), (A_2), *and* (A_3) *hold, then*

(1) *there exists a* k_0 *such that for any* $k \geq k_0$,

$$x^k \in \Omega_0; \tag{3.15}$$

(2) *any accumulation point of* $\{x^k\}$ *is a KKT point of* (NP).

Proof. If (1) does not hold, that is, there exists a subsequence $K \subseteq N$ such that for any $k \in K$, it holds that

$$x^k \notin \Omega_0. \tag{3.16}$$

By the algorithm, we know that

$$\lim_{k\to\infty} \beta_k = +\infty. \tag{3.17}$$

From the assumption (A_3) and (3.16), it follows that there exist $K_0 \subseteq K$ and $h \in R^n$ such that

$$\limsup_{k\in K_0,\, k\to\infty} \nabla g_i\left(x^k\right)^T h < 0, \quad \text{for any } i \in I^+(K_0), \tag{3.18}$$

$$I^*(K_0) = \left\{i \in I \mid g_i\left(x^k\right) > 0,\ \text{for any } k \in K_0\right\} \neq \emptyset, \qquad I^*(K_0) \subseteq I^+(K_0). \tag{3.19}$$

By (3.18) and the definition of $I^-(K_0)$, there exists a $\delta > 0$, such that for all $k \in K_0$,

$$\nabla g_i\left(x^k\right)^T h \leq -\delta, \quad \text{for any } i \in I^+(K_0), \tag{3.20}$$

$$g_i\left(x^k\right) \leq -\delta, \quad \text{for any } i \in I^-(K_0). \tag{3.21}$$

From the algorithm, we know that x^k satisfies

$$\nabla f\left(x^k\right) + \sum_{i=1}^{m} \beta_k \phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \nabla g_i\left(x^k\right) = 0. \tag{3.22}$$

Let $k \in K_0$, from (3.22) we obtain that

$$\frac{\nabla f(x^k)^T h}{\beta_k} + \sum_{i\in I^-(K_0)} \phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \nabla g_i\left(x^k\right)^T h + \sum_{i\in I^+(K_0)} \phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \nabla g_i\left(x^k\right)^T h = 0. \tag{3.23}$$

We now analyze the three terms on the left side of (3.23).

(a) By (3.17) and (A_2),

$$\lim_{k\in K_0,\, k\to\infty} \frac{\nabla f(x^k)^T h}{\beta_k} = 0. \tag{3.24}$$

(b) By (3.21), for any $i \in I^-(K_0)$, we have

$$\lim_{k\in K_0,\, k\to\infty} \frac{\beta_k g_i(x^k)}{r_k} = -\infty. \tag{3.25}$$

From the properties of $\phi(\cdot)$ and (A_2), we have that the second term satisfies

$$\lim_{k\in K_0,\, k\to\infty} \sum_{i\in I^-(K_0)} \phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \nabla g_i\left(x^k\right)^T h = 0. \tag{3.26}$$

(c) From (3.19), (3.20), and the properties of $\phi(\cdot)$, it follows that

$$\begin{aligned}\sum_{i\in I^{+}(K_0)}\phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right)\nabla g_i\left(x^k\right)^T h &\le -\delta\sum_{i\in I^{+}(K_0)}\phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right)\\ &\le -\delta\sum_{i\in I^{*}(K_0)}\phi'\left(\frac{\beta_k g_i(x^k)}{r_k}\right)\\ &\le -\delta|I^*(K_0)|\phi'(0),\end{aligned}\tag{3.27}$$

where $|I|$ denotes the number of the elements in I.

Now, by letting $k\to\infty, k\in K_0$, and taking the limit on both sides of (3.23), we obtain from (a)–(c) that

$$\delta|I^*(K_0)|\phi'(0)\le 0.\tag{3.28}$$

But by (3.19) and the properties of $\phi(\cdot)$, $\delta|I^*(K_0)|\phi'(0)>0$. This contradiction completes the proof of (1).

By (1) we know that there exists a k_0, such that if $k\ge k_0$, then $x^k\in\Omega_0$. Thus by the algorithm, when $k\ge k_0$, we have that

$$\beta_k=\beta_{k_0}.\tag{3.29}$$

Suppose that x^* is an accumulation point of $\{x^k\}$, then there exists a subsequence $\{x^k\}_{k\in K}$, such that

$$\lim_{k\in K,k\to\infty}x^k=x^*.\tag{3.30}$$

By Lemma 3.2, x^* is a feasible point of (NP), that is, $x^*\in\Omega_0$. Thus by (3.22), we obtain that

$$\nabla f\left(x^k\right)+\sum_{i\in I\setminus I(x^*)}\beta_{k_0}\phi'\left(\frac{\beta_{k_0}g_i(x^k)}{r_k}\right)\nabla g_i\left(x^k\right)+\sum_{i\in I(x^*)}\beta_{k_0}\phi'\left(\frac{\beta_{k_0}g_i(x^k)}{r_k}\right)\nabla g_i\left(x^k\right)=0.\tag{3.31}$$

In the second term of (3.31), because $i\in I\setminus I(x^*)$, so by (3.30) and the properties of $\phi'(\cdot)$, we have

$$\lim_{k\in K,k\to\infty}\phi'\left(\frac{\beta_{k_0}g_i(x^k)}{r_k}\right)=0.\tag{3.32}$$

In the third term of (3.31), from the properties of $\phi'(\cdot)$, the sequence $\{\phi'(\beta_{k_0}g_i(x^k)/r_k)\}, i\in I$ is nonnegative and bounded. Thus, there exists a subsequence $K_0\subseteq K$ such that

$$\lim_{k\in K_0,k\to\infty}\beta_{k_0}\phi'\left(\frac{\beta_{k_0}g_i(x^k)}{r_k}\right)=\lambda_i\ge 0,\quad \text{for any } i\in I(x^*).\tag{3.33}$$

At last by letting $k \to \infty, k \in K_0$, and taking the limit on both sides of (3.31), we obtain from (3.30)(3.32) and (3.33) that

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0. \tag{3.34}$$

□

By Lemma 3.2, Proposition 3.5, and Theorem 3.6, we obtain the following conclusion.

Corollary 3.7. *Suppose that* (A_1) *holds,* $\{x^k\}$ *is bounded, and any accumulation point* x^* *of* $\{x^k\}$ *satisfies MFCQ, then*

(1) *there exists a* k_0 *such that for any* $k \geq k_0$

$$x^k \in \Omega_0; \tag{3.35}$$

(2) *any accumulation point of* $\{x^k\}$ *is a KKT point of* (NP).

When (NP) is a convex programming problem, that is, the functions f and $g_i,\ i \in I$ of (NP) are all convex functions, the algorithm has better convergence results.

Theorem 3.8. *Suppose* (NP) *is a convex programming problem, then every accumulation point of* $\{x^k\}$ *is an optimal solution of* (NP).

Proof. Since $f(\cdot), g_i(\cdot),\ i \in I$ are convex, and $\phi(\cdot)$ is increasing, then for any $\beta > 0$ and $r > 0$, $f_{\beta,r}(\cdot)$ is convex. Thus $\nabla f_{\beta_k,r_k}(x^k) = 0$ is equivalent to

$$x^k \in \arg\min_{x \in R^n} f_{\beta_k,r_k}(x). \tag{3.36}$$

Therefore by (3.36) and the properties of $\phi(\cdot)$, we have for any $\overline{x} \in \Omega_0$,

$$\begin{aligned} f_{\beta_k,r_k}\left(x^k\right) &= f\left(x^k\right) + r_k \sum_{i=1}^{m} \phi\left(\frac{\beta_k g_i(x^k)}{r_k}\right) \\ &\leq f(\overline{x}) + r_k \sum_{i=1}^{m} \phi\left(\frac{\beta_k g_i(\overline{x})}{r_k}\right) \\ &\leq f(\overline{x}) + r_k m \phi(0). \end{aligned} \tag{3.37}$$

From (3.37), the arbitrariness of $\overline{x} \in \Omega_0$ and the nonnegativity of $\phi(\cdot)$, it follows that

$$f\left(x^k\right) \leq \inf_{x \in \Omega_0} f(x) + r_k m \phi(0). \tag{3.38}$$

Suppose that x^* is an accumulation point of $\{x^k\}$, there exists a subsequence $K \subseteq N$ such that $\lim_{k\in K, k\to\infty} x^k = x^*$. Thus, by (3.38), we have

$$f(x^*) \leq \inf_{x\in\Omega_0} f(x). \tag{3.39}$$

On the other side, (3.37) implies that (A_1) holds. Then from Lemma 3.2, we know $x^* \in \Omega_0$. □

Theorem 3.9. *Suppose that* (NP) *is a convex programming problem, and the assumptions* (A_2), (A_3) *hold, then*

(1) *there exists a* k_0, *for any* $k \geq k_0$, $\{f_{\beta_k, r_k}(x^k)\}$ *decreases to* $\inf_{x\in\Omega_0} f(x)$.

(2) $\lim_{k\to\infty} f(x^k) = \inf_{x\in\Omega_0} f(x)$.

Proof. Note that for (NP) which is convex, (A_1) holds. By Theorem 3.6 there exists a k_0, such that $x^k \in \Omega_0$ when $k \geq k_0$. Therefore from the algorithm, we have for any $k \geq k_0$, $\beta_k = \beta_{k_0}$. By (3.36) and the property (VI) of $\phi(\cdot)$, when $k \geq k_0$,

$$\begin{aligned} f_{\beta_{k_0}, r_{k+1}}\left(x^{k+1}\right) &\leq f_{\beta_{k_0}, r_{k+1}}\left(x^k\right) \\ &= f\left(x^k\right) + r_{k+1}\sum_{i=1}^m \phi\left(\frac{\beta_{k_0} g_i(x^k)}{r_{k+1}}\right) \\ &\leq f\left(x^k\right) + r_k\sum_{i=1}^m \phi\left(\frac{\beta_{k_0} g_i(x^k)}{r_k}\right) \\ &\leq f_{\beta_{k_0}, r_k}\left(x^k\right). \end{aligned} \tag{3.40}$$

Notice that $x^k \in \Omega_0 (k \geq k_0)$, by (3.37) and the properties of $\phi(\cdot)$, we have for $k \geq k_0$ that

$$\begin{aligned} \inf_{x\in\Omega_0} f(x) &\leq f\left(x^k\right) \\ &\leq f_{\beta_{k_0}, r_k}\left(x^k\right) \\ &\leq \inf_{x\in\Omega_0} f(x) + r_k m\phi(0). \end{aligned} \tag{3.41}$$

Combining (3.40) with (3.41), we obtain the conclusion. □

Example 3.10. Consider that $\min_{x\in\Omega_0} f(x) = (1/4)(x_1 - x_2)^2$, $\Omega_0 = \{x \in R^2 \mid g(x) = x_1 - x_2 \leq 0\}$.

This is a convex case. Denote its optimal solution by $\Omega_0^* = \{x^* \in \Omega_0 \mid x_1^* - x_2^* = 0\}$ and let $\phi(t) = (t + \sqrt{t^2+4})/2$. We consider $f_{\beta,r}(\cdot)$, that is,

$$f_{\beta,r}(x) = \frac{1}{4}(x_1 - x_2)^2 + \frac{r}{2}\left(\sqrt{\frac{\beta^2}{r^2}(x_1 - x_2)^2 + 4} + \frac{\beta}{r}(x_1 - x_2)\right). \tag{3.42}$$

Because $f_{\beta,r}(\cdot)$ is convex, thus $\nabla f_{\beta,r}(x^*) = 0$ if and only if $x^* \in \arg\min_{x\in R^2} f_{\beta,r}(x)$. By the algorithm, we get stationary points as

$$x^k = \begin{pmatrix} k \\ k+\alpha_k \end{pmatrix}, \quad k = 0, 1, \ldots, \tag{3.43}$$

where $\alpha_k > 0$ and $\lim_{k\to\infty}\alpha_k = 0$. Here $\{x^k\}$ has no accumulation point, that is, $\lim_{k\to\infty}\|x^k\| = +\infty$. Thus in the analysis of convergence, MFCQ may not be appropriate to be applied as a constraint qualification condition for this example. But for any $k \in N$, we have $\nabla f(x^k) = (-(1/2)\alpha_k, (1/2)\alpha_k)^T$, $\nabla g(x^k) = (1, -1)^T$, which implies that assumption (A_2) is satisfied. we can also check that $\{x^k\}$ satisfies GMFCQ. In fact, choose $h = (-1, 1)^T$, then we have

$$\begin{gathered} \lim_{k\to\infty} g\left(x^k\right) = \lim_{k\to\infty}\alpha_k = 0, \\ \lim_{k\to\infty} \nabla g\left(x^k\right)^T h = (1, -1)^T(-1, 1) = -2 < 0. \end{gathered} \tag{3.44}$$

On the other side, by the algorithm, we have $x^k \in \Omega_0$ and $\beta_k = 1$, for all k. By letting $k \to \infty$, we get $f_{\beta_k, r_k}(x^k) \downarrow 0$ and $f(x^k) \to 0$. So by the algorithm we get a feasible solution sequence which is also optimal.

Acknowledgments

This paper was supported by the National Natural Science Foundations of China (10971118, 10701047, and 10901096). The authors would like to give their thanks to the editor and anonymous referees for their valuable suggestions and comments.

References

[1] W. I. Zangwill, "Non-linear programming via penalty functions," *Management Science*, vol. 13, pp. 344–358, 1967.

[2] A. Auslender, R. Cominetti, and M. Haddou, "Asymptotic analysis for penalty and barrier methods in convex and linear programming," *Mathematics of Operations Research*, vol. 22, no. 1, pp. 43–62, 1997.

[3] A. Ben-Tal and M. Teboulle, "A smoothing technique for non-differentiable optimization problems," in *Optimization*, vol. 1405 of *Lecture Notes in Mathematics*, pp. 1–11, Springer, Berlin, Germany, 1989.

[4] C. H. Chen and O. L. Mangasarian, "Smoothing methods for convex inequalities and linear complementarity problems," *Mathematical Programming*, vol. 71, no. 1, pp. 51–69, 1995.

[5] C. Chen and O. L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems," *Computational Optimization and Applications*, vol. 5, no. 2, pp. 97–138, 1996.

[6] M. Herty, A. Klar, A. K. Singh, and P. Spellucci, "Smoothed penalty algorithms for optimization of nonlinear models," *Computational Optimization and Applications*, vol. 37, no. 2, pp. 157–176, 2007.

[7] M. Ç. Pinar and S. A. Zenios, "On smoothing exact penalty functions for convex constrained optimization," *SIAM Journal on Optimization*, vol. 4, no. 3, pp. 486–511, 1994.

[8] I. Zang, "A smoothing-out technique for min-max optimization," *Mathematical Programming*, vol. 19, no. 1, pp. 61–77, 1980.

[9] C. C. Gonzaga and R. A. Castillo, "A nonlinear programming algorithm based on non-coercive penalty functions," *Mathematical Programming*, vol. 96, no. 1, pp. 87–101, 2003.

[10] Z. Y. Wu, F. S. Bai, X. Q. Yang, and L. S. Zhang, "An exact lower order penalty function and its smoothing in nonlinear programming," *Optimization*, vol. 53, no. 1, pp. 51–68, 2004.

[11] Z. Meng, C. Dang, and X. Yang, "On the smoothing of the square-root exact penalty function for inequality constrained optimization," *Computational Optimization and Applications*, vol. 35, no. 3, pp. 375–398, 2006.

[12] O. L. Mangasarian and S. Fromovitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints," *Journal of Mathematical Analysis and Applications*, vol. 17, pp. 37–47, 1967.