Research Article
New Exact Penalty Functions for Nonlinear Constrained Optimization Problems

Bingzhuang Liu and Wenling Zhao

School of Science, Shandong University of Technology, Zibo, Shandong 255049, China

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

For two kinds of nonlinear constrained optimization problems, we propose two simple penalty functions, respectively, by augmenting the dimension of the primal problem with a variable that controls the weight of the penalty terms. Both of the penalty functions enjoy improved smoothness. Under mild conditions, it can be proved that our penalty functions are both exact in the sense that local minimizers of the associated penalty problem are precisely the local minimizers of the original constrained problem.

1. Introduction

Penalty function has always taken an important role in solving many constrained optimization problems in the fields such as industry design and management science. It is traditionally constructed to solve nonlinear programs by adding some penalty or barrier terms with respect to the constraints to the objective function or a corresponding Lagrange function. Then it can be optimized by some unconstrained or bounded constrained optimization software or sequential quadratic programming (SQP) techniques. No matter what kind of techniques are employed, the penalty function always depends on a small parameter ε or large parameter $\rho = \varepsilon^{-1}$. As $\varepsilon \to 0$; the minimizer of the penalty function, such as a barrier function or the quadratic penalty function [1], converges to a minimizer of the original problem. By using some exact penalty function such as l_1 penalty function (see [2–7]), the minimizer of the corresponding penalty problem must be a minimizer of the original problem when ε is sufficiently small. There are some nonsmooth penalty functions for nonsmooth optimization problems, such as the exact penalty function using the distance function for the nonsmooth variational inequality problem in Hilbert spaces [8]. In [9], the convergence of lower-order exact penalization for a constrained scalar set-valued optimization problem is given under sufficient conditions which are easy to verify.

The traditional exact penalty functions [3] are always nonsmooth. When it is used as a merit function to accept a new iterate in an SQP method, it may cause the Maratos effect [10]. On the other hand, a traditional smooth penalty function such as the quadratic penalty function cannot be an exact one. So we must compute a sequence of minimization subproblem as $\varepsilon \to 0$. At that time, ill-conditioning may occur when the penalty parameter is too large or small, which also brings difficulty of computation. In [11, 12], some kinds of augmented Lagrangian penalty functions have been proposed with improved exactness under strong conditions. In [13], exact penalty functions via regularized gap function for variational inequalities have also been given. In [14], the authors study exactness and algorithm of an objective penalty function for inequality constrained optimization. All these functions enjoy some smoothness, but at the very beginning, to use this smoothness we need second-order or third-order derivative information of the problem function that is difficult to estimate in practice. Besides, all the above kinds of penalty functions (see [11, 15–18] for summary) may be unbounded below even when the constrained problem is bounded, which may make it difficult to locate a minimizer.

Most results in the literature of exact penalization are mainly concerned with finding conditions under which a solution of the constrained optimization problem is a solution of an unconstrained penalized optimization problem, and

the reverse property is rarely studied. In [19], the author studies the reverse property. In this paper, two modified simple exact penalty functions are proposed for two kinds of constrained nonlinear programming problem, where the term *simple* means that the penalty function constructed in the primal variable space only contains the original information of the objective function and the constraint functions in the constrained optimization problem but does not contain the information of their differentials and multipliers. This kind of traditional exact penalty function can be expressed into the following form:

$$p(x; \varepsilon) = f(x) + \frac{1}{\varepsilon} p(x), \quad (1)$$

where $p(\cdot)$ satisfies that $p(x) = 0$, if x is feasible, and $p(x) > 0$ otherwise. A simple exact penalty function of this kind is the l_1 penalty function, and it is known that this penalty function is nonsmooth. The penalty functions without the multipliers have been given in [12, 17, 20, 21]. Under mild conditions, these penalty functions have been proved exact and smooth; however, since they include the information of differentials of the objective function and the constraint functions in the constrained optimization problem, they are not simple ones by our definition. In [22], a new exact penalty function is constructed by adding a new finite-dimensional or even one-dimensional decision variable to control the penalty terms. Under mild conditions, it is proved that for sufficiently large penalty parameter $\sigma > 0$, every local minimizer (x^*, ε_*) of the above penalty problem with finite penalty function value $f_\sigma(x^*, \varepsilon_*)$ has the form $\varepsilon_* = 0$, where x^* is a local minimizer of the original problem.

Inspired by this idea, in this paper, by augmenting the dimension of the program with a variable, we propose a simple exact penalty function for the equality constrained mathematical program and a simple exact barrier-penalty function for the inequality constrained mathematical program, respectively. Our new penalty function for equality constrained mathematical program is different from the one in [22] since that in [22], as the variable ε is controlled by the function β that has the properties $\beta(0) = 0$, and $\beta'(0) \geq \beta_1 > 0$, which are not needed for our function. In [22], to construct the penalty function for the inequality constrained mathematical program, the original optimization problem must be converted to be an equality constrained problem. For the inequality constrained mathematical program, we propose a new simple exact log-type barrier-penalty function, which is different from the classical log-barrier function and has broader feasible region.

2. A Modified Simple Exact Penalty Function for Equality Constrained Optimization Problems

We are now ready to propose a simple exact penalty function for equality constrained mathematical programs.

We consider the following problem:

$$\left(P^1\right) \quad \begin{matrix} \min & f(x) \\ \text{s.t.} & x \in D, \ F(x) = 0, \end{matrix} \quad (2)$$

where D is a bounded open set in the n-dimensional Euclidean space $\mathfrak{R}^n$ and $f : D \to \mathfrak{R}$ and $F : D \to \mathfrak{R}^m$ are all continuously differentiable in D. We assume that f is bounded below in D.

We then consider a new penalty function as follows:

$$\overline{f}_\sigma(x, \varepsilon)$$
$$= \begin{cases} f(x), & \text{if } \varepsilon = 0, \ \Delta(x, \varepsilon) = 0, \\ f(x) + \frac{1}{2}\varepsilon^{-\alpha}\Delta(x, \varepsilon) + \sigma\varepsilon^\beta, & \text{if } \varepsilon \neq 0, \\ +\infty, & \text{otherwise (i.e., } \varepsilon = 0, \\ & \Delta(x, \varepsilon) \neq 0), \end{cases} \quad (3)$$

where ε is a new one-dimensional variable, $\Delta(x, \varepsilon) = \|F(x) - \varepsilon^\gamma w\|^2$ is the constraint violation measure, $\gamma > \alpha \geq \beta \geq 2$ are three integers, α, β are both even, and $\sigma > 0$ is a penalty parameter. In particular, we can set $\gamma = 3$, and $\alpha = \beta = 2$. $w \in \mathfrak{R}^m$ is a preset variable, for example, we can set $w = (1, 1, \ldots, 1)$. Compared with the paper [22, 23], here we get rid of the restriction that ε is bounded and positive.

Based on the function (3), we establish the following penalty problem:

$$\left(\overline{P}_\sigma\right) \quad \min_{(x, \varepsilon) \in D \times \mathfrak{R}} \overline{f}_\sigma(x, \varepsilon). \quad (4)$$

Let $\nabla_{(x, \varepsilon)}\overline{f}_\sigma(x, \varepsilon)$ denote the gradient of $\overline{f}_\sigma(x, \varepsilon)$ in (x, ε), then we have

$$\nabla_{(x, \varepsilon)}\overline{f}_\sigma(x, \varepsilon)$$
$$= \left(\frac{\partial}{\partial x_1}\overline{f}_\sigma(x, \varepsilon), \ldots, \frac{\partial}{\partial x_n}\overline{f}_\sigma(x, \varepsilon), \frac{\partial}{\partial \varepsilon}\overline{f}_\sigma(x, \varepsilon)\right)^T. \quad (5)$$

In the following we consider the smoothness of the penalty function $\overline{f}_\sigma$.

For $(x, \varepsilon) \in D \times \mathfrak{R}$, if $\varepsilon = 0, \Delta(x, \varepsilon) = 0$, then

$$\frac{\partial}{\partial x_i}\overline{f}_\sigma(x, \varepsilon) = \frac{\partial}{\partial x_i}f(x), \quad i = 1, \ldots, n,$$
$$\frac{\partial}{\partial \varepsilon}\overline{f}_\sigma(x, \varepsilon) = \frac{\partial}{\partial \varepsilon}f(x) = 0; \quad (6)$$

if $\varepsilon \neq 0$, then

$$\frac{\partial}{\partial x_i}\overline{f}_\sigma(x, \varepsilon) = \frac{\partial}{\partial x_i}f(x) + \frac{1}{\varepsilon^\alpha}\sum_{j=1}^m \left(F_j(x) - \varepsilon^\gamma w_j\right)\frac{\partial}{\partial x_i}F_j(x),$$
$$i = 1, \ldots, n,$$

$$\frac{\partial}{\partial \varepsilon} \overline{f}_\sigma (x, \varepsilon) = -\frac{\alpha}{2} \varepsilon^{\alpha-1} \sum_{j=1}^m \left(F_j (x) - \varepsilon^\gamma w_j \right)^2$$

$$- \varepsilon^{\gamma-\alpha-1} \sum_{j=1}^m \left(F_j (x) - \varepsilon^\gamma w_j \right) + \beta \sigma \varepsilon^{\beta-1}. \tag{7}$$

Obviously, $\overline{f}_\sigma (x, \varepsilon)$ is continuously differentiable in the set $D \times \{\Re \setminus \{0\}\}$.

We are now to discuss the exactness of the penalty function $\overline{f}_\sigma (x, \varepsilon)$.

Theorem 1. *Suppose that $\{(x^k, \varepsilon_k, \sigma_k)\}$ satisfies that (x^k, ε_k) is the local minimizer of the penalty problem $(\overline{P}_{\sigma_k})$ with finite $\overline{f}_{\sigma_k}(x^k, \varepsilon_k)$, for any k, $\varepsilon_k \neq 0$, when $k \to \infty$, one has $\sigma_k \to +\infty$ and $(x^k, \varepsilon_k) \to (x^*, \varepsilon_*)$, then if $\nabla F(x^*)$ have full rank, then $\varepsilon_* = 0$, $\Delta(x^*, \varepsilon_*) = 0$, $F(x^*) = 0$.*

Proof. Since for any k, (x^k, ε_k) is the local minimizer of $(\overline{P}_{\sigma_k})$ with finite $\overline{f}_\sigma(x^k, \varepsilon_k)$, and $(x^k, \varepsilon_k) \in D \times \Re$, then we have $\nabla_{(x,\varepsilon)} \overline{f}_{\sigma_k}(x^k, \varepsilon_k) = 0$, that is,

$$\frac{\partial}{\partial x_i} \overline{f}_{\sigma_k} \left(x^k, \varepsilon_k \right) = \frac{\partial}{\partial x_i} f \left(x^k \right)$$

$$+ \varepsilon_k^{-\alpha} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right) \frac{\partial}{\partial x_i} F_j \left(x^k \right) = 0,$$

$$i = 1, \dots, n, \tag{8}$$

$$\frac{\partial}{\partial \varepsilon} \overline{f}_{\sigma_k} \left(x^k, \varepsilon_k \right) = -\gamma \varepsilon_k^{\gamma-\alpha-1} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right) w_j$$

$$- \frac{1}{2} \alpha \varepsilon_k^{-\alpha-1} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right)^2 \tag{9}$$

$$+ \beta \sigma_k \varepsilon_k^{\beta-1} = 0.$$

Equation (9) is equivalent to

$$- 2\gamma \varepsilon_k^{\gamma-\alpha-\beta} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right) w_j$$

$$- \frac{1}{2} \alpha \varepsilon_k^{-\alpha-\beta} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right)^2 + \beta \sigma_k = 0. \tag{10}$$

Let $k \to \infty$, from the assumption we know $x^k \to x^* \in D$, $\varepsilon_k \to \varepsilon_*$, if $\varepsilon_* \neq 0$, then by the above equality, we know that the first and second term on the left side are all finite value, and the third term tends to $+\infty$. This is impossible, so $\varepsilon_* = 0$.

Moreover, by (8) we know that

$$\varepsilon_k^\alpha \frac{\partial}{\partial x_i} f \left(x^k \right) + \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right) \frac{\partial}{\partial x_i} F_j \left(x^k \right) = 0, \tag{11}$$

$$i = 1, \dots, n.$$

Let $k \to \infty$, from $x^k \to x^* \in D$ and $\varepsilon_k \to \varepsilon_* = 0$, it follows that

$$\nabla F(x^*)^T F(x^*) = 0. \tag{12}$$

Since $\nabla F(x^*)$ has full rank, it follows that $F(x^*) = 0$ and $\Delta(x^*, \varepsilon_*) = 0$; this completes the proof. □

Theorem 2. *Suppose that for $(\overline{P})$, if there exists a sequence $\{(x^k, \varepsilon_k, \sigma_k)\}$ satisfying that $(x^k, \varepsilon_k) \in L(\overline{P}_{\sigma_k})$, where $L(\overline{P}_{\sigma_k})$ is the set of local minimizers of the penalty problem $(\overline{P}_{\sigma_k})$ and $\nabla F(x^k)$ have full rank, $\{\sigma_k\}$ is a strictly increasing sequence and $\overline{f}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, then there exists $k_0 > 0$, such that when $k \geq k_0$,*

$$\varepsilon_k = 0, \quad x^k \in L(P), \tag{13}$$

where $L(P)$ is the set of local minimizers of (2).

Proof. We first show that there exists a sufficiently large $k_0 > 0$, such that when $k \geq k_0$, $\varepsilon_k = 0$. If it is not the case, then there exists a subsequence of $\{(x^k, \varepsilon_k, \sigma_k)\}$, which can be still as $\{(x^k, \varepsilon_k, \sigma_k)\}$ without loss of generality, such that for each k, (x^k, ε_k) is the local minimizer of $(\overline{P}_{\sigma_k})$ with finite $\overline{f}_{\sigma_k}(x^k, \varepsilon_k)$, and $\varepsilon_k \neq 0$, when $k \to \infty$, we have $\sigma_k \to +\infty$ and $(x^k, \varepsilon_k) \to (x^*, \varepsilon_*)$. By (9) we know that

$$\frac{\partial}{\partial \varepsilon} \overline{f}_{\sigma_k} \left(x^k, \varepsilon_k \right) = -\gamma \varepsilon_k^{\gamma-\alpha-1} \sum_{j=1}^m \left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right) w_j$$

$$- \frac{1}{2} \alpha \varepsilon_k^{-\alpha-1} \Delta \left(x^k, \varepsilon_k \right) + \beta \sigma_k \varepsilon_k^{\beta-1} = 0, \tag{14}$$

then we have that

$$\frac{1}{2} \gamma \varepsilon_k^{-\alpha-1} \sum_{j=1}^m \left[\left(F_j \left(x^k \right) - \varepsilon_k^\gamma w_j \right)^2 - F_j^2 \left(x^k \right) - \varepsilon_k^{2\gamma} w_j^2 \right]$$

$$- \frac{1}{2} \alpha \varepsilon_k^{-\alpha-1} \Delta \left(x^k, \varepsilon_k \right) + \beta \sigma_k \varepsilon_k^{\beta-1} = 0, \tag{15}$$

thus,

$$(\gamma - \alpha) \Delta \left(x^k, \varepsilon_k \right) - \gamma \varepsilon_k^{2\gamma} \|w\|^2 + 2\beta \varepsilon_k^{\alpha+\beta} \sigma_k = \gamma \left\| F(x^k) \right\|^2, \tag{16}$$

then we can get that

$$(\gamma - \alpha) \varepsilon_k^{-2\alpha} \Delta \left(x^k, \varepsilon_k \right) - \gamma \varepsilon_k^{2(\gamma-\alpha)} \|w\|^2 + 2\beta \sigma_k \varepsilon_k^{\beta-\alpha}$$

$$= \gamma \varepsilon_k^{-2\alpha} \left\| F(x^k) \right\|^2. \tag{17}$$

By Theorem 1, we know that when $k \to \infty$, $\varepsilon_k \to \varepsilon_* = 0$, $\Delta(x^k, \varepsilon_k) \to \Delta(x^*, \varepsilon_*) = 0$. From the assumption that $\gamma > \alpha \geq \beta \geq 2$, we know that the third term on the left side of (17) tends to ∞, and the second term tends to zero, thus we have

$$\varepsilon_k^{-2\alpha} \left\| F(x^k) \right\|^2 \longrightarrow +\infty. \tag{18}$$

Let $y^k = F(x^k)/\varepsilon_k^\alpha$, $z^k = y^k/\|y^k\|$, it is obvious that $\|y^k\| \to +\infty$, and $\|z^k\| = 1$; without loss of generality, we suppose $z^k \to z^*$, then $\|z^*\| = 1$.

Besides, by (8) we know

$$u_i^k = \frac{1}{\|y^k\|} \frac{\partial}{\partial x_i} f\left(x^k\right) + \left(\nabla F\left(x^k\right)^T \left(F\left(x^k\right) - \varepsilon_k^{\gamma-\alpha} \frac{w}{\|y^k\|}\right)\right)_i$$

$$= 0, \quad i = 1, \dots, n. \tag{19}$$

Let $k \to \infty$, then by $\varepsilon_k \to 0$, $\|y^k\| \to +\infty$, and $x^k \to x^*$ we obtain that

$$\nabla F(x^*)^T z^* = 0. \tag{20}$$

Since $\nabla F(x^*)$ is full-rank, then $z^* = 0$, and this leads a contradiction with $\|z^*\| = 1$, so there exists a k_0, such that when $k \geq k_0$, $\varepsilon_k = 0$. Since $\overline{f}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, then by the definition of $\overline{f}_{\sigma_k}$ we have

$$F\left(x^k\right) = 0. \tag{21}$$

Again by the definition of $\overline{f}_{\sigma_k}$, there exists a neighborhood $N(x^k, r)$ of x^k, with $r > 0$ sufficiently small, such that for all x in $N(x^k, r)$ satisfying $F(x) = 0$, we have

$$\overline{f}_{\sigma_k}\left(x^k, 0\right) = f\left(x^k\right) \leq \overline{f}_{\sigma_k}(x, 0) = f(x), \tag{22}$$

thus $x^k \in L(P)$. $\qquad\square$

It is shown from Theorems 1 and 2 that under some constraint qualification condition, the local minimizer of the penalty problem corresponds to a local minimizer of the original problem, thus our penalty function for equality constrained mathematical program enjoys exactness.

3. A New Simple Exact Barrier-Penalty Function for Inequality Constrained Optimization Problem

We are now to construct a class of simple smooth exact penalty function for inequality constrained optimization problem:

$$\min_{x \in \mathscr{F}} f(x), \tag{23}$$

where $\mathscr{F} = \{x \in \mathfrak{R}^n \mid g_j(x) \leq 0, j = 1, \dots, m\}$, and f, $g_j : \mathfrak{R}^n \to \mathfrak{R}$ are all continuously differentiable functions. Throughout this section, we assume that $\mathscr{F}$ is a nonempty and bounded set. In [22], the authors transform the inequality constrained problem into a kind of equality constrained optimization problem by adding some parameters to control the constraints. In this section, we will give a new smooth and exact barrier-penalty function.

For problem (23), the classical barrier function is

$$p(x; \varepsilon) = f(x) - \varepsilon \sum_{j=1}^m \ln\left(-g_j(x)\right), \tag{24}$$

where $\varepsilon > 0$ is a parameter. The corresponding problem is

$$\min_{x \in \mathfrak{R}} p(x; \varepsilon). \tag{25}$$

Because $p(x; \varepsilon)$ constructs a barrier wall at the boundary points which satisfy $g_j(x) = 0$, the above problem is equivalent to the following problem:

$$\begin{aligned} \min \quad & p(x; \varepsilon) \\ \text{s.t.} \quad & g_j(x) < 0, \quad j = 1, 2, \dots, m. \end{aligned} \tag{26}$$

So the operation set is the interior $\mathscr{F}^0 := \{x \in \mathfrak{R}^n \mid g_j(x) < 0, j = 1, \dots, m\}$ of $\mathscr{F}$, this implies that the interior point method will need a strict interior point as an original point.

In this section, our penalty function is constructed by augmenting a variable ε. Problem (23) is equivalent to

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_j(x) \leq \varepsilon, \quad j = 1, 2, \dots, m, \\ & \varepsilon = 0. \end{aligned} \tag{27}$$

Consider the following penalty function:

$$\widetilde{p}_\sigma(x, \varepsilon)$$
$$= \begin{cases} f(x), & \text{if } \varepsilon = 0, \ x \in \mathscr{F}, \\ f(x) - \varepsilon \sum_{j=1}^m \ln\left(\varepsilon - g_j(x)\right) + \sigma\varepsilon, & \text{if } \varepsilon > 0, \ x \in \mathscr{F}_\varepsilon, \\ +\infty, & \text{otherwise}, \end{cases} \tag{28}$$

where $\mathscr{F}_\varepsilon = \{x \in \mathfrak{R}^n \mid g_j(x) < \varepsilon, j = 1, \dots, m\}$, and $\sigma > 0$ is a penalty parameter. Penalty function $\widetilde{p}_\sigma(x, \varepsilon)$ is a class of logarithmic barrier-penalty function, and the operation set can be enlarged as a set that contains the feasible region of the original problem.

If $x \in \mathscr{F}$, $\varepsilon = 0$, then

$$\frac{\partial}{\partial x_i} \widetilde{p}_\sigma(x, \varepsilon) = \frac{\partial}{\partial x_i} f(x), \quad i = 1, \dots, n,$$
$$\frac{\partial}{\partial \varepsilon} \widetilde{p}_\sigma(x, \varepsilon) = 0; \tag{29}$$

if $x \in \mathscr{F}_\varepsilon$, $\varepsilon > 0$, then

$$\frac{\partial}{\partial x_i} \widetilde{p}_\sigma(x, \varepsilon) = \frac{\partial}{\partial x_i} f(x) + \sum_{j=1}^m \frac{\varepsilon}{\varepsilon - g_j(x)} \frac{\partial}{\partial x_i} g_j(x),$$
$$i = 1, \dots, n, \tag{30}$$

$$\frac{\partial}{\partial \varepsilon} \widetilde{p}_\sigma(x, \varepsilon) = -\sum_{j=1}^m \left(\ln\left(\varepsilon - g_j(x)\right) + \frac{\varepsilon}{\varepsilon - g_j(x)}\right) + \sigma.$$

Obviously $\widetilde{p}_\sigma(x, \varepsilon)$ is continuously differentiable on the set $S_\varepsilon := \{(x, \varepsilon) \in \mathfrak{R}^n \times (0, \overline{\varepsilon}] \mid x \in \mathscr{F}_\varepsilon\}$, where $\overline{\varepsilon} > 0$ is a constant.

Consider the corresponding penalty function

$$\left(\widetilde{P}_\sigma\right) \quad \min_{x \in \mathfrak{R}^n \times [0,\bar{\varepsilon}]} \widetilde{p}_\sigma\left(x, \varepsilon\right). \tag{31}$$

Assume that $\mathscr{F}_{\bar{\varepsilon}} := \{x \in \mathfrak{R}^n \mid g_j(x) < \bar{\varepsilon}, j = 1, \ldots, m\}$ is a bounded set; problem $(\widetilde{P}_\sigma)$ is equivalent to

$$\min \quad \widetilde{p}_\sigma\left(x, \varepsilon\right)$$

$$\text{s.t.} \quad g_j\left(x\right) < \varepsilon, \quad j = 1, \ldots, m, \tag{32}$$

$$0 \le \varepsilon \le \bar{\varepsilon}.$$

In the following we discuss the exactness of penalty function $\widetilde{p}_\sigma(x, \varepsilon)$.

Theorem 3. *If there exists a sequence $\{(x^k, \varepsilon_k, \sigma_k)\}$ satisfying:*

(1) *$(x^k, \varepsilon_k) \in L(\widetilde{P}_{\sigma_k})$, and $\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, where $L(\widetilde{P}_{\sigma_k})$ is a set of the local minimizers of $(\widetilde{P}_{\sigma_k})$,*

(2) *$\varepsilon_k > 0$, and when $k \to \infty$, $(x^k, \varepsilon_k) \to (x^*, \varepsilon_*)$, $\sigma_k \to +\infty$,*

(3) *EMFCQ condition (extended Mangasarian-Fromovitz constraint qualification) is satisfied at x^*; that is, there exists a vector $z \in \mathfrak{R}^n$ such that*

$$\nabla g_i(x^*)^T z < 0, \quad \forall i \in I_0\left(x^*\right) \cup I_+\left(x^*\right),$$

$$\nabla h_j(x^*)^T z = 0, \quad \forall j = 1, \ldots, l, \tag{33}$$

where

$$I_0\left(x^*\right) := \{i \in \{1, \ldots, m\} \mid g_i\left(x^*\right) = 0\},$$

$$I_+\left(x^*\right) := \{i \in \{1, \ldots, m\} \mid g_i\left(x^*\right) > 0\}. \tag{34}$$

Then we have $\varepsilon_* = 0$, and $x^* \in \mathscr{F}$.

Proof. By the assumptions, we have

$$g_j\left(x^k\right) < \varepsilon_k, \quad j = 1, \ldots, m, \ \varepsilon_k > 0. \tag{35}$$

For any k, (x^k, ε_k) is a local optimizer of $(\widetilde{P}_{\sigma_k})$, and $\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, then for any k, we have that

$$\frac{\partial}{\partial x_i} \widetilde{p}_{\sigma_k}\left(x^k, \varepsilon_k\right) = \frac{\partial}{\partial x_i} f\left(x^k\right)$$

$$+ \sum_{j=1}^{m} \frac{\varepsilon_k}{\varepsilon_k - g_j(x^k)} \frac{\partial}{\partial x_i} g_j\left(x^k\right) = 0, \tag{36}$$

$$i = 1, \ldots, n,$$

$$\frac{\partial}{\partial \varepsilon} \widetilde{p}_{\sigma_k}\left(x^k, \varepsilon_k\right) = \sigma_k - \sum_{j=1}^{m} \left(\ln\left(\varepsilon_k - g_j\left(x^k\right)\right) + \frac{\varepsilon_k}{\varepsilon_k - g_j\left(x^k\right)} \right)$$

$$\le 0, \tag{37}$$

and $(\partial/\partial\varepsilon)\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k) < 0$ holds if and only if $\varepsilon_k = \bar{\varepsilon}$.

Suppose that $\varepsilon_* > 0$, then by (37) and the fact that when $k \to \infty, \sigma_k \to +\infty$, we have that

$$\sum_{j=1}^{m} \left(\ln\left(\varepsilon_k - g_j\left(x^k\right)\right) + \frac{\varepsilon_k}{\varepsilon_k - g_j\left(x^k\right)} \right) \longrightarrow +\infty. \tag{38}$$

Because $\{x^k\}$ is bounded, and $\varepsilon_k \in [0, \bar{\varepsilon}]$, so $\{\varepsilon_k - g_j(x^k)\}$ is bounded. Then when $k \to \infty$, we have

$$\sum_{j=1}^{m} \frac{\varepsilon_k}{\varepsilon_k - g_j\left(x^k\right)} \longrightarrow +\infty. \tag{39}$$

Since $\varepsilon_* > 0$, there exists at least $j_0 \in \{1, \ldots, m\}$ such that when $k \to \infty$,

$$\varepsilon_k - g_{j_0}\left(x^k\right) \longrightarrow 0^+, \tag{40}$$

thus

$$g_{j_0}\left(x^*\right) > 0. \tag{41}$$

Let $J_0 \subseteq \{1, \ldots, m\}$ be the set of such index j_0. Then by (36) and the boundedness of $\{x^k\}$, we have for any $j_0 \in J_0$,

$$\frac{\partial}{\partial x_i} g_{j_0}\left(x^k\right) \longrightarrow 0, \quad i = 1, \ldots, n, \tag{42}$$

when $k \to \infty$; so

$$\frac{\partial}{\partial x_i} g_{j_0}\left(x^*\right) = 0, \quad i = 1, \ldots, n. \tag{43}$$

This contradicts the assumption that EMFCQ is satisfied at x^*. Thus $\varepsilon_* = 0$. By $g_j(x^k) < \varepsilon_k, j = 1, \ldots, m$, then we have

$$g_j\left(x^*\right) \le 0, \quad j = 1, \ldots, m, \tag{44}$$

that is, $x^* \in \mathscr{F}$. $\qquad\square$

Theorem 4. *Assume that $\mathscr{F}_{\bar{\varepsilon}}$ is a bounded set, and there exists a sequence $\{(x^k, \varepsilon_k, \sigma_k)\}$ such that $(x^k, \varepsilon_k) \in L(\widetilde{P}_{\sigma_k})$, where $L(\widetilde{P}_{\sigma_k})$ is the set of local minimizers of $(\widetilde{P}_{\sigma_k})$, EMFCQ is satisfied at x^k, $\{\sigma_k\}$ is an increasing sequence, and $\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, then there exists a sufficiently large $k_0 > 0$, such that when $k \ge k_0$,*

$$\varepsilon_k = 0, \quad x^k \in L\left(P\right), \tag{45}$$

where $L(P)$ is the set of local minimizers of (P).

Proof. We first show that there exists a sufficiently large $k_0 > 0$ such that $k \ge k_0, \varepsilon_k = 0$. If it is not the case, then there exists a subsequence of $\{(x^k, \varepsilon_k, \sigma_k)\}$; here we assume without loss of generality that $\{(x^k, \varepsilon_k, \sigma_k)\}$ is the subsequence, such that $(x^k, \varepsilon_k) \in L(\widetilde{P}_{\sigma_k})$, $\varepsilon_k > 0$, $\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, and when $k \to \infty, \sigma_k \to +\infty$. Then from Theorem 3, it follows that

$$\varepsilon_k \longrightarrow \varepsilon_* = 0, \quad x^k \longrightarrow x^* \in \mathscr{F}. \tag{46}$$

By (37) we have,

$$\sum_{j=1}^{m} \frac{\varepsilon_k}{\varepsilon_k - g_j(x^k)} \longrightarrow +\infty, \quad k \longrightarrow \infty. \tag{47}$$

By the proof of Theorem 3, the above results contradicts with the assumption that EMFCQ condition is satisfied at x^*. Thus there exists a sufficiently large $k_0 > 0$ such that $\varepsilon_k = 0$, for $k \geq k_0$.

Since for any k, $\widetilde{p}_{\sigma_k}(x^k, \varepsilon_k)$ is finite, then by the definition of $\widetilde{p}_{\sigma_k}$,

$$x^k \in \mathscr{F}. \tag{48}$$

Then by the definition of $\widetilde{p}_{\sigma_k}$, there exists a neighborhood $N(x^k, r)$ of x^k, where $r > 0$ is sufficiently small, such that for all $x \in N(x^k, r) \cap \mathscr{F}$,

$$\widetilde{p}_{\sigma_k}(x^k, 0) = f(x^k) \leq \widetilde{p}_{\sigma_k}(x, 0) = f(x), \tag{49}$$

thus $x^k \in L(P)$. □

It is shown by Theorem 4 that under some constraint qualification condition, a local minimizer corresponds a local minimizer of the original problem when the penalty parameter is sufficiently large, thus the penalty function (28) is an exact penalty function. Since the penalty function (28) is a penalty function with a barrier, thus for problem (23), we can still apply the interior method. Note that we can use an interior point $(x, \varepsilon) \in \mathscr{F}_\varepsilon \times (0, \varepsilon')$ as the original point, where $\varepsilon' \in (0, \bar{\varepsilon})$.

Conflict of Interests

The authors declared that there is no conflict of interests in their submitted paper.

Acknowledgments

References

[1] R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations," *Bulletin of the American Mathematical Society*, vol. 49, pp. 1–23, 1943.

[2] T. Antczak, "Exact penalty functions method for mathematical programming problems involving invex functions," *European Journal of Operational Research*, vol. 198, no. 1, pp. 29–36, 2009.

[3] S. P. Han and O. L. Mangasarian, "Exact penalty functions in nonlinear programming," *Mathematical Programming*, vol. 17, no. 3, pp. 251–269, 1979.

[4] T. Hoheisel, C. Kanzow, and J. V. Outrata, "Exact penalty results for mathematical programs with vanishing constraints," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 72, no. 5, pp. 2514–2526, 2010.

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

[6] A. J. Zaslavski, "A sufficient condition for exact penalty in constrained optimization," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 250–262, 2005.

[7] A. J. Zaslavski, "Stability of exact penalty for nonconvex inequality-constrained minimization problems," *Taiwanese Journal of Mathematics*, vol. 14, no. 1, pp. 1–19, 2010.

[8] M. Soleimani-damaneh, "Penalization for variational inequalities," *Applied Mathematics Letters*, vol. 22, no. 3, pp. 347–350, 2009.

[9] X. X. Huang, "Convergence of a class of penalty methods for constrained scalar set-valued optimization," *Journal of Global Optimization*, vol. 56, no. 4, pp. 1501–1513, 2013.

[10] N. Maratos, *Exact penalty function algorithms for finite dimensional and control optimization problems [Ph.D. thesis]*, University of London, 1978.

[11] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, NY, USA, 1982.

[12] G. Di Pillo, "Exact penalty methods," in *Algorithms for Continuous Optimization*, E. Spedicato, Ed., vol. 434, pp. 209–253, Kluwer Academic, Dordrecht, The Netherlands, 1994.

[13] W. Li and J. Peng, "Exact penalty functions for constrained minimization problems via regularized gap function for variational inequalities," *Journal of Global Optimization*, vol. 37, no. 1, pp. 85–94, 2007.

[14] Z. Q. Meng, C. Y. Dang, M. Jiang, X. S. Xu, and R. Shen, "Exactness and algorithm of an objective penalty function," *Journal of Global Optimization*, vol. 56, no. 2, pp. 691–711, 2013.

[15] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Optimizaton Theory and Algorithms*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1993.

[16] D. Boukari and A. V. Fiacco, "Survey of penalty, exact-penalty and multiplier methods from 1968 to 1993," *Optimization*, vol. 32, no. 4, pp. 301–334, 1995.

[17] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1987.

[18] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, NY, USA, 1999.

[19] J. J. Ye, "The exact penalty principle," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 75, no. 3, pp. 1642–1654, 2012.

[20] G. Di Pillo and S. Lucidi, "An augmented Lagrangian function with improved exactness properties," *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 376–406, 2001.

[21] G. Di Pillo, G. Liuzzi, and S. Lucidi, "An exact penalty-Lagrangian approach for large-scale nonlinear programming," *Optimization*, vol. 60, no. 1-2, pp. 223–252, 2011.

[22] W. Huyer and A. Neumaier, "A new exact penalty function," *SIAM Journal on Optimization*, vol. 13, no. 4, pp. 1141–1158, 2003.

[23] C. Yu, K. L. Teo, L. Zhang, and Y. Bai, "A new exact penalty function method for continuous inequality constrained optimization problems," *Journal of Industrial and Management Optimization*, vol. 6, no. 4, pp. 895–910, 2010.