J Glob Optim (2007) 37:85–94
DOI 10.1007/s10898-006-9038-8

ORIGINAL PAPER

Exact penalty functions for constrained minimization problems via regularized gap function for variational inequalities

Wu Li · Jiming Peng

Received: 21 October 2004 / Accepted: 27 April 2006 /
Published online: 30 June 2006

Abstract By using the regularized gap function for variational inequalities, we introduce a new penalty function $P_\alpha(x)$ for the problem of minimizing a twice continuously differentiable function in a closed convex subset of the n-dimensional space $\mathbb{R}^n$. Under certain assumptions, it is shown that any stationary point of the penalty function $P_\alpha(x)$ satisfies the first-order optimality condition of the original constrained minimization problem, and any local (or global) minimizer of $P_\alpha(x)$ on $\mathbb{R}^n$ is a locally (or globally) optimal solution of the original optimization problem.

1 Introduction

Consider the following constrained minimization problem:

$$\min_{x\in X} \ f(x), \tag{1}$$

where $f(x)$ is a twice continuously differentiable function on $\mathbb{R}^n$ and X is a closed convex subset of $\mathbb{R}^n$. Related to (1) is the following variational inequality problem of finding x in X such that

$$(y-x)^T\nabla f(x) \geq 0 \quad \text{for } y \in X, \tag{2}$$

where z^T denotes the transpose of a vector z, and $\nabla f(x)$ is the gradient of $f(x)$. Note that $\bar{x}$ is called a stationary point for the constrained minimization problem (1) if $\bar{x}$ satisfies (2).

W. Li (✉)
Senior Research Engineer, Aeronautics Systems Analysis Branch,
NASA Langley Research Center Mail Stop 442, Hampton, VA 23681, USA
e-mail: w.li@nasa.gov

J. Peng
Department of Computing and Software, McMaster University 1280 Main Street West, Hamilton Ont., Canada L8S 4L7
e-mail: pengj@mcmaster.ca

One approach for solving the constrained optimization problem (1) is by using exact penalty functions, which tries to attack problem (1) by solving the following reformulation of (1):

$$\min_{x\in\mathbb{R}^n} f_\alpha(x), \tag{3}$$

where $f_\alpha(x)$ is a penalty function depending on the penalty parameter α, usually constructed by using the objective $f(x)$ and the constraint functions. Nondifferentiable exact penalty functions are easier to construct than continuously differentiable exact penalty functions. However, we then have to solve a nonsmooth unconstrained minimization problem instead of a smooth constrained one. Continuously differentiable exact penalty functions can be obtained under some constraint qualifications such as the Mangasarian–Fromovitz constraint qualification. Many exact penalty functions have been introduced in the literature [1–5, 7, 10, 15]. Typically, these exact penalty functions are much more complicated than $f(x)$ even when X is defined by simple bound constraints [1, 4]. Moreover, though there have been several results regarding the existence of a penalty parameter α such that a local (or global) minimizer of $f_\alpha(x)$ in $\mathbb{R}^n$ is a local (or global) solution of (1), there is no simple way to estimate the penalty parameter α.

Different from the above-mentioned approaches, in this paper we propose to construct exact penalty functions for (1) by using the projection onto the feasible set X. Our method is closely associated with the so-called regularized gap function for (2), introduced by Fukushima [6],

$$\begin{aligned} G_\alpha(x) &:= \max_{y\in X}\left\{(x-y)^T\nabla f(x) - \frac{1}{2\alpha}\|x-y\|^2\right\} \\ &= (x-H_\alpha(x))^T\nabla f(x) - \frac{1}{2\alpha}\|x-H_\alpha(x)\|^2, \end{aligned} \tag{4}$$

where $H_\alpha(x) = \Pi_X(x-\alpha\nabla f(x))$ and $\Pi_X(y)$ denotes the projection of y onto X. It has been observed that several exact penalty functions scattered in the literature [11, 13, 14] can be rewritten as [12, Eq. 16]

$$P_\alpha(x) = f(x) - G_\alpha(x) = f(x) + (H_\alpha(x)-x)^T\nabla f(x) + \frac{1}{2\alpha}\|x-H_\alpha(x)\|^2. \tag{5}$$

Our main purpose here is to propose a unified way of constructing exact penalty functions which not only extends, but also improves these scattered results in the literature [11, 13, 14, 16] under different settings. In particular, we give conditions under which a local (global) minimizer of $P_\alpha(x)$ is a local (global) minimizer of (1) so that the function $P_\alpha(x)$ can act as an exact penalty function for (1). In special cases where X is defined by some simple constraints, we give explicit formulas for $P_\alpha(x)$ in terms of $f(x)$ and the constraints (see Propositions 13, 16, and 17). From a theoretical perspective, our approach can also be applied to minimization problems with general convex feasible set, while its practical efficiency still remains a challenge.

The paper is organized as follows, In Sect. 2, we mainly study the stationary points of $P_\alpha(x)$. A result about the convexity of $P_\alpha(x)$ is also given. In Sect. 3, we consider local (or global) minimizers of the merit function $P_\alpha(x)$ and show that, for some carefully chosen parameter α (depending only on the Hessian of $f(x)$), any local (or global) minimizer of $P_\alpha(x)$ is a local (or global) solution of (1). In Sect. 4, we give explicit forms of $P_\alpha(x)$ when X is defined by either simple bound constraints, or a spherical constraint, or an ice-cream cone. Conclusions are given in Sect. 5.

2 Gradient, stationary point, and convexity

We first consider the gradient of the function $P_\alpha(x)$. From Theorem 3.2 in [6], we have

Lemma 1 *Suppose that $P_\alpha(x)$ is defined as in* (5) *and $f(x)$ is twice continuously differentiable. Then*

$$\nabla P_\alpha(x) = \frac{1}{\alpha}(I - \alpha\nabla^2 f(x))(x - H_\alpha(x)).$$

Since the gradient of $P_\alpha(x)$ is a linear transform of $(x - H_\alpha(x))$, a fixed point of $H_\alpha(x)$ must be a stationary point of $P_\alpha(x)$. We note that many iterative algorithms have been proposed for the variational inequality problem (2) based on the following characterization of $H_\alpha(x)$[6, 9].

Lemma 2 *A vector $\bar{x} \in \mathbb{R}^n$ is a solution of* (2) *if and only if $\bar{x} - H_\alpha(\bar{x}) = 0$.*

We point out that if $f(x)$ is convex, then solving (1) is equivalent to finding a fixed point of $H_\alpha(x)$. This is exactly what Tao and Hoai An did in [16] for minimizing a quadratic function $f(x) = \frac{1}{2}x^T Qx + b^T x$ on a Euclidean ball. In their work, Tao and Hoai An studied a special DCA method where the iterates are generated by $x^{k+1} := \Pi_X(x^k - \alpha(Qx^k + b))$ with $\alpha^{-1}I - Q$ being positive definite. It is easy to verify that the special DCA method in [16] is nothing more than a fixed-point iteration method for $H_\alpha(x)$

It follows from Lemmas 1 and 2 that if $(I - \alpha\nabla^2 f(x))$ is nonsingular for every x, then any stationary point of $P_\alpha(x)$ is a solution of (2) and vice versa.

Theorem 3 *Suppose that $(I - \alpha\nabla^2 f(\bar{x}))$ is nonsingular. Then $\bar{x}$ is a solution of* (2) *if and only if $\bar{x}$ is a stationary point of $P_\alpha(x)$(i.e.,$\nabla P_\alpha(\bar{x}) = 0$). In particular, if $f(x) = \frac{1}{2}x^T Qx + q^T x$ with a symmetric matrix $Q \in \mathbb{R}^{n\times n}$ and $\alpha\|Q\| < 1$, then $\bar{x}$ is a solution of* (2) *if and only if $\bar{x}$ is a stationary point of $P_\alpha(x)$.*

Next we show that the convexity of $f(x)$ implies the convexity of $P_\alpha(x)$ when $f(x)$ is a quadratic function and α is sufficiently small.

Theorem 4 *Suppose that $f(x) = \frac{1}{2}x^T Qx + q^T x$ is convex. If $\alpha\|Q\| \le 1$, then $P_\alpha(x)$ is convex.*

Proof It suffices to show that

$$(y - x)^T(\nabla P_\alpha(y) - \nabla P_\alpha(x)) \ge 0 \quad \text{for } x, y \in \mathbb{R}^n. \tag{6}$$

Since Q is symmetric positive semidefinite, we know that $B := (I - \alpha Q)$ is positive semidefinite and $\|B\| \le 1$ when $\alpha\|Q\| \le 1$. Denote $y = x + d$, we have

$$\begin{aligned}
&\alpha(y - x)^T(\nabla P_\alpha(y) - \nabla P_\alpha(x)) \\
&\quad = d^T Bd - d^T B(\Pi_X(B(x + d) - \alpha q) - \Pi_X(Bx - \alpha q)) \\
&\quad \ge d^T Bd - \|Bd\| \cdot \|\Pi_X(B(x + d) - \alpha q) - \Pi_X(Bx - \alpha q)\| \\
&\quad \ge d^T Bd - \|Bd\|^2 = d^T(B - B^2)d \ge 0,
\end{aligned}$$

Where the first equality follows from Lemma 1, the first inequality is the Cauchy–Schwarz inequality, the second inequality is the nonexpansive property of the projection $\Pi_X(x) : \|\Pi_X(x) - \Pi_X(y)\| \le \|x - y\|$ [8 (3.1.6)], and the last inequality is by the positive semidefiniteness of $(B - B^2)$. This proves the theorem. □

3 Global and local minimizers

The following lemma about the equivalence of (2) and the (global) constrained minimization of $G_\alpha(x)$ is due to Fukushima [6].

Lemma 5 *Let the function* $G_\alpha(x)$ *be defined by* (4). *Then* $G_\alpha(x) \geq 0$ *for all* $x \in X$. *Moreover,* $G_\alpha(\bar{x}) = 0$ *with* $\bar{x} \in X$ *if and only if* $\bar{x}$ *solves the variational inequality problem* (2).

The next theorem follows directly from Lemma 5 and Theorem 3.

Theorem 6 *Let* $P_\alpha(x)$ *be defined by* (5). *Then*

$$P_\alpha(x) \leq f(x) \quad \text{for}\, x \in X \tag{7}$$

and the equality holds if and only if x solves (2). *Moreover, for any* $x \in \mathbb{R}^n$, *if* $\nabla P_\alpha(x) = 0$ *and* $(I - \alpha\nabla^2 f(x))$ *is nonsingular, then* $P_\alpha(x) = f(x)$.

The above theorem indicates that $P_\alpha(x)$ is indeed a penalty function of (1) while $G_\alpha(x)$ acts as a penalty term.

Now we can prove that local (or global) minimizers of $P_\alpha(x)$ are local (or global) solutions of (1) for some penalty parameter α.

Theorem 7 *Suppose* $\bar{x} \in \mathbb{R}^n$ *and* $(I - \alpha\nabla^2 f(\bar{x}))$ *is nonsingular. If* $\bar{x}$ *is a global (or local) minimizer of* $P_\alpha(x)$ *in* $\mathbb{R}^n$, *then* $\bar{x}$ *is a global (or local) minimizer of* $f(x)$ *in* X.

Proof Since $\nabla P_\alpha(\bar{x}) = 0$ and $(I - \alpha\nabla^2 f(\bar{x}))$ is nonsingular, by Theorem 3, $\bar{x}$ is a solution of (2), so $\bar{x} \in X$. By Theorem 6, $P_\alpha(\bar{x}) = f(\bar{x})$. Thus for any $x \in X$ (or $x \in X$ and in a neighborhood of $\bar{x}$), it follows from (7) that

$$f(x) \geq P_\alpha(x) \geq P_\alpha(\bar{x}) = f(\bar{x}),$$

which shows that $\bar{x}$ is also a global (ox local) minimizer of $f(x)$ in X. □

If there is a penalty parameter $\alpha > 0$ such that $(I - \alpha\nabla^2 f(x))$ is nonsingular for every $x \in \mathbb{R}^n$, then we can solve the constrained minimization problem (1) by finding local or global minimizers of $P_\alpha(x)$ in $\mathbb{R}^n$. However, one open question is whether or not $P_\alpha(x)$ has a local or global minimizer when (1) has local or global solutions. That is, we have to study the converse of Theorem 7 in order to establish a complete equivalence of (1) with the following unconstrained minimization problem:

$$\min_{x\in\mathbb{R}^n} P_\alpha(x). \tag{8}$$

For this purpose, we study the following family of functions involved in definition of $P_\alpha(x)$:

$$h_\alpha(x,y) := f(x) + (y-x)^T\nabla f(x) + \frac{1}{2\alpha}\|y-x\|^2. \tag{9}$$

Lemma 8 *For any* $x, y \in \mathbb{R}^n$, *there exists* θ *such that* $0 < \theta < 1$ *and*

$$f(y) \leq h_\alpha(x,y) - \frac{1-\alpha\|\nabla^2 f(y_\theta)\|}{2\alpha}\|y-x\|^2, \tag{10}$$

where $y_\theta = x + \theta(y-x)$.

Proof By the Taylor expansion, there is θ such that $0 < \theta < 1$ and

$$f(y) = f(x) + (y-x)^T \nabla f(x) + \frac{1}{2}(y-x)^T \nabla^2 f(y_\theta)(y-x), \tag{11}$$

where $y_\theta = x + \theta(y-x)$. Thus,

$$\begin{aligned} f(y) &\le f(x) + (y-x)^T \nabla f(x) + \tfrac{\|\nabla^2 f(y_\theta)\|}{2} \|y-x\|^2 \\ &= h_\alpha(x,y) - \tfrac{1-\alpha\|\nabla^2 f(y_\theta)\|}{2\alpha} \|y-x\|^2, \end{aligned} \tag{12}$$

where the inequality follows from the Cauchy–Schwarz inequality. □

A direct consequence of the above lemma is the following corollary.

Corollary 9 *For any $x \in \mathbb{R}^n$ there exists $x_\theta := x + \theta(H_\alpha(x) - x)$ with $0 < \theta < 1$ such that*

$$P_\alpha(x) \ge f(H_\alpha(x)) + \frac{1-\alpha\|\nabla^2 f(x_\theta)\|}{2\alpha} \|H_\alpha(x) - x\|^2. \tag{13}$$

If $\alpha\|\nabla^2 f(x)\| \le 1$ for all $x \in \mathbb{R}^n$, then

$$P_\alpha(x) \ge f(H_\alpha(x)) \quad \text{for } x \in \mathbb{R}^n. \tag{14}$$

Proof By (5), $P_\alpha(x) = h_\alpha(x, H_\alpha(x))$. Replacing y by $H_\alpha(x)$ in (10) we obtain (13). The inequality (14) follows from (13) and the assumption that $\alpha\|\nabla^2 f(x_\theta)\| \le 1$. □

Now we can prove the following converse of Theorem 7.

Theorem 10 *For any $\bar{x} \in \mathbb{R}^n$, the following statements hold.*

(1) *If $\alpha\|\nabla^2 f(\bar{x})\| < 1$ and $\bar{x}$ is a local minimizer of $f(x)$ in X, then $\bar{x}$ is a local minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.*
(2) *If $\alpha\|\nabla^2 f(x)\| \le 1$ for $x \in \mathbb{R}^n$ and $\bar{x}$ is a global minimizer of $f(x)$ in X, then $\bar{x}$ is also a global minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.*

Proof First we assume that $\bar{x}$ is a local minimizer of $f(x)$ in X and $\alpha\|\nabla^2 f(\bar{x})\| < 1$ Since $\bar{x}$ is a local minimizer of $f(x)$ in X,there exists a positive scalar ϵ_1 such that

$$f(\bar{x}) \le f(y) \quad \text{for } y \in X \quad \text{with } \|y - \bar{x}\| \le \epsilon_1. \tag{15}$$

By the continuity of $\nabla^2 f(x)$ and $\alpha\|\nabla^2 f(\bar{x})\| < 1$, there is a positive constant ϵ_2 such that

$$\alpha\|\nabla^2 f(x)\| < 1 \quad \text{for } \|x - \bar{x}\| \le \epsilon_2.$$

Since $\bar{x}$ is also a solution of (2), by Lemma 2, $H_\alpha(\bar{x}) = \bar{x}$. By continuity of the projection mapping $\Pi_X(\cdot)$, for $\epsilon = \min\{\epsilon_1, \epsilon_2\}$, there exists a positive scalar δ such that $\|H_\alpha(x) - \bar{x}\| = \|H_\alpha(x) - H_\alpha(\bar{x})\| \le \epsilon$ if $\|(x) - \bar{x}\| \le \delta$.

Let $x \in \mathbb{R}^n$ be such that $\|x - \bar{x}\| \le \min\{\epsilon, \delta\}$. Then $\|H_\alpha(x) - \bar{x}\| \le \epsilon$. Choose θ between 0 and 1 such that (13) holds. Since $\|x_\theta - \bar{x}\| \le \epsilon$, we have $\alpha\|\nabla f(x_\theta)\| < 1$. Thus, it follows from (14) and (15) that

$$P_\alpha(x) \ge f(H_\alpha(x)) \ge f(\bar{x}) = P_\alpha(\bar{x}) \quad \text{for } x \in \mathbb{R}^n \quad \text{with } \|x - \bar{x}\| \le \min\{\epsilon, \delta\}, \tag{16}$$

where the last equality follows from Theorem 6. Thus, $\bar{x}$ is a local minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.

Next assume that $\bar{x}$ is a global minimizer of $f(x)$ in X and $\alpha\|\nabla^2 f(x)\| \leq 1$ for $x \in \mathbb{R}^n$. Then it follows from (14) and (7) that

$$P_\alpha(x) \geq f(H_\alpha(x)) \geq f(\bar{x}) \geq P\alpha(\bar{x}) \quad \text{for } x \in \mathbb{R}^n,$$

which implies that $\bar{x}$ is also a global minimizer of $P_\alpha(x)$. □

Note that the penalty parameter α can be explicitly determined by using the 2-norm of the Hessian of $f(x)$. However, when $\|\nabla^2 f(x)\|$ is unbounded on $\mathbb{R}^n$, we do not know whether or not $P_\alpha(x)$ has any global minimizer in $\mathbb{R}^n$. The significance of (1) of Theorem 10 in nonlinear programming is that it ensures the existence of a local minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$ when $f(x)$ has a local minimizer in X and α is sufficiently small. Heuristically, it means that by choosing a sufficiently small α, we can obtain a local minimizer of $f(x)$ in X by finding a local minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.

Combining Theorems 7 and 10, we obtain the following equivalence of the original constrained minimization problem (1) and the unconstrained minimization of $P_\alpha(x)$ on $\mathbb{R}^n$, which is the main conclusion in this section.

Theorem 11 *For any $\bar{x} \in \mathbb{R}^n$, the following statements hold.*

(1) *Suppose $\alpha\|\nabla^2 f(\bar{x})\| < 1$. Then $\bar{x}$ is a local minimizer of $f(x)$ in X if and only if $\bar{x}$ is a local minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.*
(2) *Suppose $\alpha\|\nabla^2 f(x)\| < 1$ for all $x \in \mathbb{R}^n$. Then $\bar{x}$ is a global minimizer of $f(x)$ in X if and only if $\bar{x}$ is a global minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.*

Proof If $\alpha\|\nabla^2 f(\bar{x})\| < 1$, then $(I - \alpha\nabla^2 f(\bar{x}))$ is nonsingular. Hence, (1) follows from Theorem 7 and (1) of Theorem 10, If $\alpha\|\nabla^2 f(x)\| < 1$ for $x \in \mathbb{R}$, then $(I - \alpha\nabla^2 f(x))$ is nonsingular for $x \in \mathbb{R}$. Thus, (2) follows from Theorem 7 and (2) of Theorem 10. □

One important application of Theorem 11 is when $f(x)$ is quadratic, as shown in the following corollary that improves the main result in [11].

Corollary 12 *Suppose that $f(x) = \frac{1}{2}x^TQx + q^Tx$ is a quadratic function with a symmetric matrix Q and $\alpha\|Q\| < 1$. Then $\bar{x}$ is a local (or global) minimizer of $f(x)$ in X if and only if $\bar{x}$ is a local (or global) minimizer of $P_\alpha(x)$ in $\mathbb{R}^n$.*

4 Explicit forms of the exact penalty function

By Corollary 12, it seems that we can always have an equivalent unconstrained reformulation of a constrained quadratic (or linear) programming problem. The computation of $P_\alpha(x) = f(x) - G_\alpha(x)$ and its gradient $\nabla P_\alpha(x)$ involve $H_\alpha(x)$ and thus the projection mapping $\Pi_X(x - \alpha\nabla f(x))$. Therefore, even though we do not have difficulty in choosing the penalty parameter α, it might be computationally inefficient to calculate $P_\alpha(x)$ and $\nabla P_\alpha(x)$ repeatedly. However, in cases that X is defined by some simple constraints so that an explicit analytical expression of $H_\alpha(x)$ is available, then we can get an explicit form of $P_\alpha(x)$. In this section, we give three explicit forms of $P_\alpha(x)$ when X is defined by simple bound constraints, a spherical constraint, or an ice-cream cone constraint,

Proposition 13 *Suppose that* $X = \{x \in \mathbb{R}^n : l \le x \le u\}$, *Then* $P_\alpha(x)$ *is a continuously differentiable function and*

$$P_\alpha(x) = f(x) + \frac{1}{2\alpha}\|([x - \alpha\nabla f(x)] - u)_+\|^2 + \frac{1}{2\alpha}\|(l - [x - \alpha\nabla f(x)]_+\|^2 \\ -\frac{\alpha}{2}\|\nabla f(x)\|^2, \tag{17}$$

where $t_+ = \max\{0, t\}$.

Proof By algebraic manipulations, we have

$$(y - x)^T\nabla f(x) + \frac{1}{2\alpha}\|x - y\|^2 = \frac{1}{2\alpha}\|[x - \alpha\nabla f(x)] - y\|^2 - \frac{\alpha}{2}\|\nabla f(x)\|^2.$$

From the definition of P_α and the above identity, we get

$$P_\alpha(x) = f(x) + \min_{l \le y \le u}\left\{\frac{1}{2\alpha}\|[x - \alpha\nabla f(x)] - y\|^2 - \frac{\alpha}{2}\|\nabla f(x)\|^2\right\}. \tag{18}$$

On the other hand, it is easy to verify that the orthogonal projection from a vector z to $X = \{x : l \le x \le u\}$ is $(z)_l^u$ whose ith component is $\max\{l_i, \min\{u_i, z_i\}\}$. Let

$$[(z)_l^u]_i = \max\{l_i, \min\{u_i, z_i\}\} \quad \text{for } 1 \le i \le n.$$

Then

$$z - (z)_l^u = (z - u)_+ - (l - z)_+$$

and

$$\min_{l \le y \le u}\|z - y\|^2 = \|z - (z)_l^u\|^2 = \|(z - u)_+ - (l - z)_+\|^2. \tag{19}$$

Since $l_i \le u_i$ for $1 \le i \le n$, we obtain $(z - u)_+^T(l - z)_+ = 0$. Hence (19) implies

$$\min_{l \le y \le u}\|z - y\|^2 = \|(z - u)_+ - (l - z)_+\|^2 = \|(z - u)_+\|^2 + \|(l - z)_+\|^2. \tag{20}$$

Applying (20) with $z = x - \alpha\nabla f(x)$ in (18), we obtain (17). □

As a direct consequence of Corollary 12 and Proposition 13, we have the following slight improvement of the main results given in [11].

Corollary 14 *Suppose that* $X = \{x \in \mathbb{R}^n : l \le x \le u\}$, $f(x) = \frac{1}{2}x^TQx + q^Tx$ *with a symmetric matrix* Q, *and* $\alpha\|Q\| < 1$. *Then a vector* $\bar{x}$ *is a local (or global) minimizer of* $f(x)$ *in* X *if and only if* $\bar{x}$ *is a local (or global) minimizer of the following differentiable piecewise quadratic minimization problem:*

$$\min_{x \in \mathbb{R}^n} f(x) + \frac{1}{2\alpha}\|([x - \alpha\nabla f(x)] - u)_+\|^2 + \frac{1}{2\alpha}\|(l - [x - \alpha\nabla f(x)])_+\|^2 - \frac{\alpha}{2}\|\nabla f(x)\|^2.$$

We point out that the above corollary was proved in [11] under the assumption that $0 < 2\alpha\|Q\| < 1$ with a very complicated analysis.

A generalization of the above corollary gives a continuously differentiable exact penalty function for nonlinear programming problems with simple bound constraints.

Corollary 15 *Suppose that $X = \{x \in \mathbb{R}^n : l \le x \le u\}$, and $\alpha\|\nabla^2 f(x)\| < 1$ for $x \in \mathbb{R}^n$. Then a vector $\bar{x}$ is a local (or global) minimizer of $f(x)$ in X if and only if $\bar{x}$ is a local (or global) minimizer of the following continuously differentiable unconstrained minimization problem:*

$$\min_{x\in\mathbb{R}^n} f(x) + \frac{1}{2\alpha}\|([x-\alpha\nabla f(x)]-u)_+\|^2 + \frac{1}{2\alpha}\|(l-[x-\alpha\nabla f(x)])_+\|^2 - \frac{\alpha}{2}\|\nabla f(x)\|^2.$$

Proof Corollary 15 follows from Theorem 11 and Proposition 13. □

Next we show that $P_\alpha(x)$ has an explicit form if X is defined by a spherical constraint.

Proposition 16 *Suppose that $\rho > 0$ and $X = \{x \in \mathbb{R}^n : \|x\| \le \rho\}$. Let $p(x) = \min\left\{1, \frac{\rho}{\|x-\alpha\nabla f(x)\|}\right\}$. Then*

$$P_\alpha(x) = f(x) + (p(x)[x-\alpha\nabla f(x)] - x)^T\nabla f(x) + \frac{1}{2\alpha}\|x - p(x)[x-\alpha\nabla f(x)]\|^2, \tag{21}$$

$$\nabla P_\alpha(x) = \frac{1}{\alpha}(I-\alpha\nabla^2 f(x))(p(x)[x-\alpha\nabla f(x)] - x). \tag{22}$$

Moreover, when $f(x) = \frac{1}{2}x^TQx + q^Tx$ with a symmetric matrix Q and $\alpha\|Q\| < 1$, a vector $\bar{x}$ is a local (or global) minimizer of $f(x)$ in X if and only if $\bar{x}$ is a local (or global) minimizer of $P_\alpha(x)$.

Proof It is easy to verify that

$$H_\alpha(x) = \Pi_X(x-\alpha\nabla f(x)) = p(x)(x-\nabla f(x)).$$

Thus, we obtain (21) and (22) from (5) and Lemma 1, respectively. The second part of Proposition 16 follows from Corollary 12. □

Finally we give an explicit form of $P_\alpha(x)$ when X is an ice-cream cone, The following proposition can be proved by calculating $\Pi\hat{X}(y)$ and using Corollary 12, We leave the details to interested readers.

Proposition 17 *Suppose X is an ice-cream defined by $X = \{x \in \mathbb{R}^n : x_n \ge \gamma(x)\}$, where*

$$\gamma(x) = \sqrt{\sum_{i=1}^{n-1} x_i^2}.$$

Then

$$P_\alpha(x) = f(x) + (\varphi(x-\alpha\nabla f(x)) - x)^T\nabla f(x) + \frac{1}{2\alpha}\|x - \varphi(x-\alpha\nabla f(x))\|^2, \tag{23}$$

$$\nabla P_\alpha(x) = \frac{1}{\alpha}(I-\alpha\nabla f(x))(\varphi(x-\alpha\nabla f(x)) - x), \tag{24}$$

where

$$\varphi(x) := \begin{cases} 0, & \text{if } x_n \le -\gamma(x), \\ x, & \text{if } x_n \ge \gamma(x), \\ \left(\frac{\gamma(x)+x_n}{2\gamma(x)}\right)(x_1, x_2, \ldots, x_{n-1}, \gamma(x))^T, & \text{if } -\gamma(x) < x_n < \gamma(x). \end{cases}$$

Moreover, when $f(x) = \frac{1}{2}x^TQx + q^Tx$ *with a symmetric matrix* Q *and* $\alpha\|Q\| < 1$*, a vector* $\bar{x}$ *is a local (or global) minimizer of* $f(x)$ *in* X *if and only if* $\bar{x}$ *is a local (or global) minimizer of* $P_\alpha(x)$.

5 Conclusion

By using Fukushima's regularized gap function $G_\alpha(x)$, we obtain an exact penalty function $P_\alpha(x) = f(x) - G_\alpha(x)$ for the constrained minimization problem (1). For a twice continuously differentiable function $f(x), P_\alpha(x)$ is a differentiable function on $\mathbb{R}^n$. In general, under the assumption that all $\alpha\|\nabla^2 f(x)\| < 1$ for all x in $\mathbb{R}^n$ (which implies the uniform boundedness of the Hessian of $f(x)$ on $\mathbb{R}^n$), we have proved that local (or global) minimizers of $f(x)$ in X are local (or global) minimizers of $P_\alpha(x)$ in $\mathbb{R}^n$. This shows that $P_\alpha(x)$ is a continuously differentiable *strongly exact penalty function* on $\mathbb{R}^n$ [3, Definition 3] for the constrained minimization problem (1).

For the special cases where X is defined by simple bound constraints or a spherical constraint, we give the explicit forms of $P_\alpha(x)$. In particular, when $f(x)$ is quadratic, it is very easy to choose α. However, in general the computation of P_α and choice of α are the most difficult issues when using $P_\alpha(x)$ as an exact penalty function since the computation of $P_\alpha(x)$ involves the projection to the feasible set X and it is hard to find a suitable α when the Hessian of $f(x)$ is unbounded on $\mathbb{R}^n$. How to resolve these issues in applications is still a challenging problem.

Acknowledgment The research was done by Wu Li, while he was a professor at Old Dominion University, Norfolk, Virginia, USA

References

1. Contaldi, G., Di Pillo, G., Lucidi, S.: A continuously differentiable exact penalty function for nonlinear programming problems with unbounded feasible set. Oper. Res. Lett. **14**, 153–161 (1993)
2. Pillo, G.D., Grippo, L.: A continuously differentiable exact penalty function for nonlinear programming problems with inequality constraint. Math. Program. **36**, 1–18 (1986)
3. Pillo, G.D., Grippo, L.: Exact penalty functions in constrained optimization. SIAM J. Cont. Optim. **27**, 1333–1360 (1989)
4. Facchinei, F., Lucidi, S.: A class of methods for optimization problems with simple bounds. Optim. **26**, 239–259 (1992)
5. Fletcher, R.E.: A class of methods for nonlinear programming with termination and convergence properties. In Abadie, J. (ed.) Integer and Nonlinear Programming pp. 157–173. North-Holland, Amsterdam (1970)
6. Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**, 99–110 (1992)
7. Glad, T., Polak, E.: A multiplier method with automatic limitation of penalty growth. Math. Program. **17**, 251–296 (1979)
8. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I. Springer-Verlag, New York (1993)
9. Harker, P.T., Pang, J.-S.: Finite dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
10. Huyer, W., Neumaier, A.: A new exact penalty function. SIAM J. Optim. **13**, 1141–1158 (2003)
11. Li, W.: Differentiable piecewise quadratic exact penalty functions for quadratic programs with simple bound constraints. SIAM J. Optim. **6**, 299–315 (1996)
12. Li, W.: A merit function and a Newton-type method for symmetric linear complementarity problems. In: Ferris M., Pang J.-S. (eds) Complementarity and Variational Problems, Proceedings of International Conference on Complementarity Problems, held at the Johns Hopkins University, November 1–4, 1995, SIAM, Philadelphia, 1997, pp. 181–203

13. Li, W., Swetits, J.: A Newton method for convex regression, data smoothing, and quadratic programming with bounded constraints. SIAM J. Optim. **3**, 466–488 (1993)
14. Li, W., Swetits, J.: A new algorithm for strictly convex quadratic programs. SIAM J. Optim. **7**, 595–619 (1997)
15. Lucidi, S.: New results continuously differentiable exact penalty functions. SIAM J. Optim. **2**, 558–574 (1992)
16. Tao P.D., Hoai An, L.T: Difference of convex functions optimization algorithms (DCA) for globally minimizing nonconvex quadratic forms on Euclidean balls and spheres. Oper. Res. Lett. **19**, 207–216 (1996)
17. Williamson, R.-E., Crowell, R.H., Trotter, H.F.: Calculus of Vector Functions, (2nd edn), Prentice-Hall Inc., NJ (1968)