Journal of Global Optimization **20:** 49–65, 2001.

Filled functions for unconstrained global optimization

ZHENG XU[1,*], HONG-XUAN HUANG[2,†], PANOS M. PARDALOS[1,*] and CHENG-XIAN XU[3,‡]
[1]*Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA;*
[2]*Department of Mathematical Sciences, Tsinghua University, Beijing 100084, P. R. China;*
[3]*Faculty of Science, Xi'an Jiaotong University, Xi'an 710049, P. R. China*

(Received 28 August 1998; accepted in revised form 23 February 2001)

Abstract. This paper is concerned with filled function techniques for unconstrained global minimization of a continuous function of several variables. More general forms of filled functions are presented for smooth and non-smooth optimization problems. These functions have either one or two adjustable parameters. Conditions on functions and on the values of parameters are given so that the constructed functions have the desired properties of filled functions.

Key words: Global optimization, Local minimizer, Filled function, Basin, Hill

1. Introduction

Global optimization problems arise in many disperse fields of science and technology. Many methods have been proposed to search for a globally optimal solution of a given function (Ge, 1990; Horst and Pardalos, 1995; Horst et al., 2000; Levy and Gómez, 1985; Wales and Scheraga, 1999). Many deterministic methods, including Filled Function (Ge, 1990), Tunneling (Levy and Gómez, 1985) and Basin-Hopping (Wales and Scheraga, 1999), use a transformed objective function strategy to construct a path from one of the local minimizers of a given function to another local minimizer with lower function value (if the objective function has many minimizers).

This paper is concerned with filled functions for unconstrained global minimization of a continuous function $F(x)$, $x \in R^n$. Let $\mathcal{X}$ be a closed and bounded nonempty set which contains a finite number of minimizers of the function $F(x)$, and $x_k^* \in \mathcal{X}$ be a known local minimizer of $F(x)$ with $F(x_k^*) > F^* = \min\{F(x) | x \in \mathcal{X}\}$. The basic idea of the filled function methods is to construct an auxiliary

* The author was partially supported by NSF under Grants DBI 9808210 and EIA 9872509.
† The author was partially supported by the China Scholarship Council while the author was visiting the Department of Industrial and Systems Engineering, University of Florida, Gainesville, USA.
‡ The author was supported by Natural Science Foundation of China.

function, called *filled function* of $F(x)$, such that minimizing the filled function will generate a point x_{k+1} in a basin (a particular connected domain around a local minimizer, see Definition 2.2 in Section 2) of $F(x)$ lower than the basin B_k^* of $F(x)$ at x_k^*. Then the minimization of the function $F(x)$ can be restarted at the point x_{k+1} to generate a new minimizer x_{k+1}^* of $F(x)$ with $F(x_{k+1}^*) < F(x_k^*)$. Repeat the process until a global minimizer of $F(x)$ is found. The filled function is updated at successively local minimizers of $F(x)$. The filled function at a local minimizer x_k^* of $F(x)$ is required to reach its maximum at x_k^*, to have neither a minimizer nor a saddle point in the basin B_k^* and in any basin of $F(x)$ higher than B_k^*, and to have minimizers or saddle points in basins of $F(x)$ lower than B_k^*.

The first filled function with two adjustable parameters was proposed for smooth optimization by Ge in 1983 and finally published in 1990. Theoretical analyses and numerical experiments show that the filled function method is promising. However, the filled function with two parameters have some disadvantages, especially, the excessive restriction on the choices of the parameter values. Modifications to the filled function are made to avoid the restricted choices of the parameter values (Ge and Qin, 1987), and to extend to non-smooth optimization (Ge, 1987; Kong and Zhuang, 1996). In this paper we will present extensions of Ge's type filled functions in (Ge, 1987, 1990; Ge and Qin, 1987, 1990; Kong and Zhuang, 1996; Zhuang, 1994) to more general forms for smooth optimization, and present some filled functions for non-smooth optimization.

The paper is organized as follows: In Section 2, some definitions and assumptions related to the filled functions are presented. Sections 3 and 4 are devoted to general forms of the filled functions with two parameters and with one parameter, respectively, for smooth optimization. In Section 5, we will present some filled functions for non-smooth optimization.

2. Definitions and assumptions

In this section we give some definitions and assumptions that will be used in the paper. It is assumed that the function $F(x)$ is *globally convex* (or *coercive*) in the sense $F(x) \to +\infty$ as $\|x\| \to +\infty$. Global convexity implies that there exists a closed and bounded set $\mathcal{X} \subset R^n$ such that $\mathcal{X}$ contains all global minimizers of $F(x)$, and that all minimizers (local or global) of $F(x)$ in $\mathcal{X}$ are interior points of $\mathcal{X}$. It is also assumed that $F(x)$ has a finite number of minimizers in $\mathcal{X}$ and hence, each minimizer is isolated.

DEFINITION 2.1. *Let $x_1, x_2 \in R^n$, $x_1 \neq x_2$. $x_1 - x_2$ is called a descent (ascent) segment of $F(x)$ if the function $f(\lambda) = F(x_2 + \lambda(x_1 - x_2))$ is monotonically decreasing (increasing) for $\lambda \in [0, 1]$. $x_1 - x_2$ is called a strong descent (ascent) segment of $F(x)$ if one of the following conditions holds:*

(i) *When* $F(x)$ *is continuously differentiable,*

$$f'(\lambda) = \nabla F(x_2 + \lambda(x_1 - x_2))^T (x_1 - x_2) < 0 \ (> 0), \qquad \forall \lambda \in [0, 1].$$

(ii) *When* $F(x)$ *is directional differentiable,*

$$\max\{(x_1 - x_2)^T \xi \mid \xi \in \partial F(x(\lambda))\} < 0 \ (> 0),$$

where $\partial F(x)$ *is the sub-gradient of* $F(x)$ *at* x *and* $x(\lambda) = x_2 + \lambda(x_1 - x_2),\ \lambda \in [0, 1]$.

DEFINITION 2.2. *The basin* B_k^* *at a local minimizer* x_k^* *of* $F(x)$ *is a connected domain with the following properties:*

(i) $x_k^* \in B_k^*$;
(ii) *for any* $x \in B_k^*$ *such that* $x \neq x_k^*$ *and* $F(x) > F(x_k^*)$, *there exists a descent trajectory from* x *to* x_k^*.

The hill of $F(x)$ *at its maximizer* $\hat{x}_k^*$ *is the basin of* $-F(x)$ *at* $\hat{x}_k^*$.

The idea about the basin appeared in the 1970s (Dixon et al., 1976). If the function $F(x)$ is continuously differentiable, the descent trajectory is a smooth curve, for example, the steepest descent trajectory. If the function $F(x)$ is non-smooth, the trajectory may consist of a finite number of descent segments of $F(x)$.

DEFINITION 2.3. *Let* x_k^* *and* x_{k+1}^* *be two distinct minimizers of* $F(x)$. *If* $F(x_k^*) > F(x_{k+1}^*)$, *we say that the basin* B_{k+1}^* *at* x_{k+1}^* *is lower than the basin* B_k^* *at* x_k^* *or the basin* B_k^* *is higher than the basin* B_{k+1}^*.

Let U_k^* denote the union of all basins higher than B_k^*. It is clear that $F(x) > F(x_k^*)$ holds for any point $x \in U_k^*$.

DEFINITION 2.4. *The simple basin* S_k^* *at a local minimizer* x_k^* *of* $F(x)$ *is a connected domain with the following properties:*

(i) $x_k^* \in S_k^* \subset B_k^*$;
(ii) $(x - x_k^*)$ *is a strong ascent segment of* $F(x)$ *for any* $x \in S_k^*$, $x \neq x_k^*$.

By the Definition 2.1, we know that

$$(x - x_k^*)^T \nabla F(x) > 0, \quad \forall x \in S_k^*,\ x \neq x_k^*,$$

if $F(x)$ is continuously differentiable. In addition, if x_k^* is an isolated minimizer of $F(x)$, then

$$D_k = \min\{\|x - x_k^*\| \mid x \notin S_k^*\} > 0, \tag{2.1}$$

that is, the minimal radius of the simple basin S_k^* is not zero. If $F(x)$ is twice continuously differentiable and $\nabla^2 F(x)$ is positive definite at x_k^*, then x_k^* is an isolated minimizer.

DEFINITION 2.5. *$P(x)$ is called a filled function of the function $F(x)$ at a local minimizer x_k^* if $P(x)$ has the following properties:*

(i) *x_k^* is a local maximizer of $P(x)$;*
(ii) *$P(x)$ has neither minimizer nor saddle point in $W_k^* \backslash \{x_k^*\}$ and the set W_k^* becomes a part of a hill of $P(x)$ at x_k^*, where $W_k^* = B_k^* \cup U_k^*$;*
(iii) *If $F(x)$ has a basin, e.g., B_{k+1}^*, lower than B_k^*, then there exists a point $x' \in B_{k+1}^*$ that minimizes $P(x)$ along the ray $x_k^* + \lambda(x' - x_k^*),\ \lambda > 0$.*

These properties of a filled function ensure that when a descent method, for example a quasi-Newton method, is employed to minimize the constructed filled function, the sequence of iterate points will not terminate at any point in W_k^*, and when there exist basins of $F(x)$ lower than B_k^*, the sequence will either terminate at a point in a basin lower than B_k^* or generate a point such that the value of $F(x)$ is less than $F(x_k^*)$.

In the following sections, it is assumed that x_k^* is an available local minimizer of $F(x)$ and F^* is the global minimum of $F(x)$.

3. Filled functions with two parameters

In this section we study filled functions with two parameters for smooth optimization. It is assumed that the function $F(x)$ is twice continuously differentiable and has a finite number of minimizers in the compact set $\mathcal{X}$. Furthermore, suppose that $\nabla^2 F(x)$ is positive definite at every minimizer. In this case, each minimizer of $F(x)$ is an isolated minimizer and is contained in a certain simple basin.

The first filled function proposed by Ge (1990) has the form

$$p(x, r, \rho) = \frac{1}{r + F(x)} \exp(-\frac{\|x - x_k^*\|^2}{\rho^2}),$$

where r and ρ are two adjustable parameters. Under some conditions on the function $F(x)$ and on the values of the parameters r and ρ, the function $p(x, r, \rho)$ is a filled function of $F(x)$.

A more general form of filled functions with two parameters can be expressed as

$$P(x, r, A) = \psi(r + F(x)) \exp(-A w(\|x - x_k^*\|^\beta)), \quad \beta \geqslant 1,\ A > 0, \tag{3.2}$$

where the value of r is chosen so that $r + F(x) > 0$ for all $x \in \mathcal{X}$, and the functions $\psi(t),\ w(t)$ have the following properties:

(i) $\psi(t)$ and $w(t)$ are continuously differentiable for $t \in (0, +\infty)$;
(ii) for $t \in (0, +\infty)$, $\psi(t) > 0$, $\psi'(t) < 0$ and $\psi'(t)/\psi(t)$ is monotonically increasing;
(iii) $w(0) = 0$ and for any $t \in (0, +\infty)$, $w(t) > 0$, $w'(t) \geqslant c > 0$.

Choices for these two functions can be $1/t^a (a > 0)$, $\mathrm{csch}(t)$, $\exp(1/t) - 1, \cdots$ for $\psi(t)$ and t, $\sinh(t)$, $e^t - 1, \cdots$ for $w(t)$.

The following theorems give the conditions on the values of the parameters r and A so that $P(x, r, A)$ is a filled function of $F(x)$.

THEOREM 3.1. *x_k^* is a local maximizer of $P(x, r, A)$ for any $A > 0$.*

Proof. Since x_k^* is an isolated local minimizer of $F(x)$ and is contained in a simple basin S_k^*, $F(x) > F(x_k^*)$ for all $x \in S_k^*$, $x \neq x_k^*$. From the properties of the functions ψ and w, it follows that

$$\begin{aligned} P(x_k^*, r, A) &= \psi(r + F(x_k^*)) > \psi(r + F(x)) \\ &> \psi(r + F(x))\exp(-Aw(\|x - x_k^*\|^\beta)) = P(x, r, A), \end{aligned}$$

for all $x \in S_k^*$, $x \neq x_k^*$. Thus, x_k^* is a local maximizer of $P(x, r, A)$. □

THEOREM 3.2. *If r and A satisfy the inequality*

$$A(r + F(x_k^*)) \geqslant \frac{L}{c\alpha\beta D_k^{\beta-1}}, \tag{3.3}$$

then any $x \in W_k^* \backslash \{x_k^*\}$ is not a stationary point of $P(x, r, A)$, where

$$\begin{aligned} L &= \max\{\|\nabla F(x)\| \mid x \in \mathcal{X}\}, \\ \theta(t) &= -\frac{\psi(t)}{t\psi'(t)}, \quad t \in (0, +\infty), \\ \alpha &= \min\{\theta(r + F(x)) \mid x \in \mathcal{X}\}, \end{aligned}$$

D_k is defined in (2.1) *and W_k^* has the same meaning as one in Definition* 2.5.

Note that since both $\theta(t)$ and $F(x)$ are continuous, $\mathcal{X}$ is compact and $r + F(x) > 0$ for all $x \in \mathcal{X}$, the minimum of the function $\theta(r + F(x))$ exists on the set $\mathcal{X}$.
Proof. The gradient of the function $P(x, r, A)$ with respect to x is given by

$$\begin{aligned} \nabla P(x, r, A) &= -\exp(-Aw(\|x - x_k^*\|^\beta))[-\psi'(r + F(x))\nabla F(x) + \\ &\quad A\beta\psi(r + F(x))w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^{\beta-2}(x - x_k^*)]. \end{aligned} \tag{3.4}$$

For $x \in S_k^*$, $x \neq x_k^*$, it follows from $(x - x_k^*)^T \nabla F(x) > 0$ that $(x - x_k^*)^T \nabla P(x, r, A) < 0$. Hence, $\nabla P(x, r, A) \neq 0$, x is not a stationary point of $P(x, r, A)$ and $x - x_k^*$ is a descent direction of the function $P(x, r, A)$ at x.

For any $x \in V_k^* = W_k^* \backslash S_k^*$, it follows from (3.4) that if x is a stationary point of $P(x, r, A)$, then the equation

$$\frac{-\psi'(r + F(x))}{A\psi(r + F(x))} \nabla F(x) + \beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^{\beta-2}(x - x_k^*) = 0$$

must hold. Hence, a necessary condition for x to be a stationary point of $P(x, r, A)$ is

$$\frac{1}{A(r + F(x))} = \frac{\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^{\beta-1}}{\|\nabla F(x)\|} \theta(r + F(x)).$$

On one hand, from the property of the function $w(t)$ and definitions of D_k, α and L, we have

$$\frac{\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^{\beta-1}}{\|\nabla F(x)\|} \theta(r + F(x)) \geqslant \frac{c\alpha\beta D_k^{\beta-1}}{L}. \tag{3.5}$$

On the other hand, by using the hypothesis (3.3) and the fact $x \in V_k^*$, we have

$$\frac{1}{A(r + F(x))} < \frac{1}{A(r + F(x_k^*))} \leqslant \frac{c\alpha\beta D_k^{\beta-1}}{L}. \tag{3.6}$$

Therefore, (3.5), (3.6) and the above necessary condition show that when the values of r and A satisfy (3.3), $x \in V_k^*$ can not be a stationary point of $P(x, r, A)$. □

From Theorem 3.2 and its proof, we can obtain the following conclusions:

(i) for any $x \in S_k^*$, $x \neq x_k^*$, $x - x_k^*$ is a strong descent segment of $P(x, r, A)$ at x_k^*;

(ii) for $x \in V_k^*$, appropriate choices for the values of r and A can make $x - x_k^*$ a descent direction of $P(x, r, A)$ at x. Thus, W_k^* becomes a part of a hill of $P(x, r, A)$ with top x_k^*. In fact, it follows from (3.4) that $(x - x_k^*)^T \nabla P(x, r, A) < 0$ is equivalent to the inequality

$$\frac{-\psi'(r + F(x))}{A\psi(r + F(x))}(x - x_k^*)^T \nabla F(x) + \beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta > 0. \tag{3.7}$$

If $(x - x_k^*)^T \nabla F(x) \geqslant 0$, the inequality (3.7) certainly holds. If $(x - x_k^*)^T \nabla F(x) < 0$, then the inequality (3.7) also holds when

$$\frac{1}{A(r + F(x))} < \frac{1}{A(r + F(x_k^*))} \leqslant -\frac{\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta}{(x - x_k^*)^T \nabla F(x)} \theta(r + F(x)). \tag{3.8}$$

It can be seen that condition (3.8) is implied by condition (3.3).

The following two theorems have similar proofs to that of Theorem 3.2, and hence their proofs are omitted.

THEOREM 3.3. *If the values of r and A satisfy the following inequality*

$$A(r + F^*) \geqslant \frac{L}{c\alpha\beta D_k^{\beta-1}}$$

then $x - x_k^*$ is a descent direction of $P(x, r, A)$ for any $x \in \mathcal{X}$, $x \neq x_k^*$, and $P(x, r, A)$ has neither minimizer nor saddle point in $\mathcal{X}$.

THEOREM 3.4. *Let $x \notin W_k^*$. If $(x - x_k^*)^T \nabla F(x) \geqslant 0$, $x - x_k^*$ is a descent direction of $P(x, r, A)$ at x. If $(x - x_k^*)^T \nabla F(x) < 0$, then $x - x_k^*$ is an ascent direction of $P(x, r, A)$ at x when*

$$\frac{1}{A(r + F(x))} > -\frac{\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta}{(x - x_k^*)^T \nabla F(x)} \theta(r + F(x)). \tag{3.9}$$

Theorem 3.4 indicates that if $F(x)$ has basins lower than B_k^*, the function $P(x, r, A)$ may have either minimizers or saddle points in a basin of $F(x)$ lower than B_k^* when the values of r, A and $F(x)$ satisfy condition (3.9). If x is a point in such a basin with $(x - x_k^*)^T \nabla F(x) < 0$, then $P(x, r, A)$ has a minimizer in the line $x(\lambda) = x_k^* + \lambda(x - x_k^*)$, $\lambda \in [0, 1]$. This is because $P(x(\lambda), r, A)$ is descent for small λ and is ascent for λ close 1.

The following theorem shows that for appropriately chosen values of r and A, either minimizers or saddle points, $\hat{x}$ say, of $P(x, r, A)$ are really in basins of $F(x)$ lower than the basin B_k^*, and $\hat{x} - x_k^*$ is a descent direction of $F(x)$ at $\hat{x}$.

THEOREM 3.5. *Let the values of r and A satisfy*

$$\frac{1}{A(r + F(x_k^*))} < \frac{c\beta D_k^{\beta-1}\theta(r + F(x_k^*))}{L}. \tag{3.10}$$

If $\hat{x}$ is either a minimizer or a saddle point of $P(x, r, A)$, then

$$(\hat{x} - x_k^*)^T \nabla F(\hat{x}) < 0 \quad \text{and} \quad F(\hat{x}) < F(x_k^*).$$

Proof. It follows from $\nabla P(\hat{x}, r, A) = 0$ and (3.4) that

$$\frac{-\psi'(r + F(\hat{x}))}{A\psi(r + F(\hat{x}))}(\hat{x} - x_k^*)^T \nabla F(\hat{x}) + \beta w'(\|\hat{x} - x_k^*\|^\beta)\|\hat{x} - x_k^*\|^\beta = 0. \tag{3.11}$$

Thus we have

$$(\hat{x} - x_k^*)^T \nabla F(\hat{x}) < 0.$$

Since $(x - x_k^*)^T \nabla F(x) > 0$ for any $x \in S_k^* \backslash \{x_k^*\}$, $\hat{x}$ cannot be a point in S_k^*.

Suppose $F(\hat{x}) \geqslant F(x_k^*)$. It follows from (3.11) and the monotonic property of the function $\psi'(t)/\psi(t)$ that

$$\frac{c\beta D_k^{\beta-1}}{L} \leqslant \frac{\beta\|\hat{x}-x_k^*\|^{\beta-1}w'(\|\hat{x}-x_k^*\|^{\beta})}{\|\nabla F(\hat{x})\|}$$
$$\leqslant -\frac{\psi'(r+F(\hat{x}))}{A\psi(r+F(\hat{x}))} \leqslant -\frac{\psi'(r+F(x_k^*))}{A\psi(r+F(x_k^*))} = \frac{1}{A(r+F(x_k^*))\theta(r+F(x_k^*))}.$$

This contradicts with (3.10). Thus $F(\hat{x}) < F(x_k^*)$ must hold. □

The above analyses indicate that the function $P(x, r, A)$ is a desired filled function of the function $F(x)$ provided that the values of r and A are properly chosen. Hence, $P(x, r, A)$ can be used in the process of finding global minimizers of $F(x)$. However, there are some disadvantages for the filled function $P(x, r, A)$:

(i) The evaluations of the function $P(x, r, A)$ and the gradient $\nabla P(x, r, A)$ are easily affected by the factor $\exp(-Aw(\|x-x_k\|^{\beta})$. Large values of $Aw(\|x-x_k^*\|^{\beta})$ will either lead to an overflow of calculations or make very small changes in $P(x, r, A)$ and $\nabla P(x, r, A)$.

(ii) The choices for the values of the parameters r and A are very restricted. It follows from Theorems 3.2 and 3.3 that r and A must satisfy the following inequalities

$$A(r+F(x_k^*)) \geqslant L/(c\alpha\beta D_k^{\beta-1}) > A(r+F^*),$$

when $r + F^* > 0$, in order that $P(x, r, A)$ becomes a desired filled function.

(iii) An undesirable property of the filled function $P(x, r, A)$ is indicated by Theorem 3.3, that is, global minimizers of the function $F(x)$ may be lost if the values of r and A are not properly chosen.

4. Filled functions with one parameter

Attempts have been made to improve the properties of the filled functions. The basic idea for the modification is to cancel the parameter r by taking $r = -F(x_k^*)$. The first modification is proposed by Ge and Qin (1987) and has the form

$$q(x, A) = -[F(x) - F(x_k^*)]\exp(A\|x-x_k^*\|^2), \quad A > 0.$$

A more general form for filled functions with one parameter can be expressed as

$$Q(x, A) = -\phi(F(x) - F(x_k^*))\exp(Aw(\|x-x_k^*\|^{\beta})), \;\; \beta \geqslant 1, \quad A > 0,$$

where the function $w(t)$ has the same properties as one in Section 3 and $\phi(t)$ has the following properties:

(i) $\phi(t)$ is continuously differentiable for $t \geqslant 0$;
(ii) $\phi(0) = 0, \phi'(t) > 0, \ \forall t \geqslant 0$;
(iii) $\phi'(t)/\phi(t)$ is monotonically decreasing for $t \in (0, +\infty)$.

Choices for the function $\phi(t)$ can be t, $a^t - 1$ $(a > 1)$, $\sinh(t)$ and so on. The filled function properties of $Q(x, A)$ will be proved in the following theorems.

THEOREM 4.1. *x_k^* is a local maximizer of $Q(x, A)$ for any $A > 0$.*

Proof. It follows from the properties of the function $\phi(t)$ that

$$Q(x_k^*, A) = 0 > Q(x, A), \ \forall x \in S_k^*, \ x \neq x_k^*.$$

Thus, x_k^* is a local maximizer of $Q(x, A)$. □

THEOREM 4.2. *For any $x \in W_k^*$, $x \neq x_k^*$, if*

$$A > \max\{0, \frac{-\phi'(F(x) - F(x_k^*))\nabla F(x)^T(x - x_k^*)}{\phi(F(x) - F(x_k^*))\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta}\}, \tag{4.12}$$

then $x - x_k^*$ is a descent direction of $Q(x, A)$ at x and x is not a stationary point of $Q(x, A)$.

Proof. The gradient of the function $Q(x, A)$ with respect to x is given by

$$\begin{aligned}\nabla Q(x, A) &= -\exp(Aw(\|x - x_k^*\|^\beta))[\phi'(F(x) - F(x_k^*))\nabla F(x) \\ &\quad + A\beta\phi(F(x) - F(x_k^*))w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^{\beta-2}(x - x_k^*)].\end{aligned} \tag{4.13}$$

If $x \in S_k^*$, $x \neq x_k^*$, then it follows from $(x - x_k^*)^T\nabla F(x) > 0$ that $(x - x_k^*)^T\nabla Q(x, A) < 0$. Therefore, $x - x_k^*$ is a descent direction of $Q(x, A)$ at x for any $A > 0$ and x is not a stationary point of $Q(x, A)$ since $\nabla Q(x, A) \neq 0$.

Now, we consider points $x \in V_k^* = W_k^* \backslash S_k^*$. If $(x - x_k^*)^T\nabla F(x) \geqslant 0$, we still have $(x - x_k^*)^T\nabla Q(x, A) < 0$ for any $A > 0$. If $(x - x_k^*)^T\nabla F(x) < 0$, then $(x - x_k^*)^T\nabla Q(x, A) < 0$ is equivalent to

$$\begin{aligned}&\phi'(F(x) - F(x_k^*))(x - x_k^*)^T\nabla F(x) \\ &+ A\beta\phi(F(x) - F(x_k^*))w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta > 0.\end{aligned}$$

This inequality will hold when

$$A > \frac{-\phi'(F(x) - F(x_k^*))(x - x_k^*)^T\nabla F(x)}{\beta\phi(F(x) - F(x_k^*))w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta}.$$

This completes the proof. □

Let x_j^*, $j = 1, 2, \cdots, \ell$ be the local minimizers of $F(x)$ with $F(x_j^*) > F(x_k^*)$ and

$$m = \frac{\phi'(d)}{\phi(d)}, \; d = \min\{F(x_j^*) - F(x_k^*) \mid j = 1, 2, \cdots, \ell\} > 0. \tag{4.14}$$

If

$$A > \frac{mL}{c\beta D_k^{\beta-1}},$$

then (4.12) is implied and $x - x_k^*$ is a descent direction of $Q(x, A)$ at any point $x \in W_k^*$, $x \neq x_k^*$. Therefore, for A large enough, the set W_k^* of $F(x)$ becomes a part of a hill of $Q(x, A)$ with top x_k^*.

The following theorem is a direct consequence of $(x - x_k^*)^T \nabla Q(x, A) > 0$ and (4.13), and shows that if $F(x)$ has basins lower than B_k^*, $Q(x, A)$ has either minimizers or saddle points in one of such basins.

THEOREM 4.3. *Let $x \notin W_k^*$, $F(x) > F(x_k^*)$. If $(x - x_k^*)^T \nabla F(x) < 0$ and*

$$A < \frac{-\phi'(F(x) - F(x_k^*))(x - x_k^*)^T \nabla F(x)}{\beta \phi(F(x) - F(x_k^*)) w'(\|x - x_k^*\|^\beta) \|x - x_k^*\|^\beta}, \tag{4.15}$$

then $x - x_k^*$ is an ascent direction of $Q(x, A)$ at x.

When $F(x) - F(x_k^*) \to 0^+$ for x in a basin of $F(x)$ lower than B_k^* (if it exists), $\phi(F(x) - F(x_k^*)) \to 0^+$ and the right hand side of (4.15) tends to infinity. Thus the inequality (4.15) will hold for points in such a basin with $F(x) - F(x_k^*)$ close to 0^+ no matter how large A is. This property of the filled function $Q(x, A)$ ensures that global minimizers of $F(x)$ will not be lost, and that if the minimization of the filled function $Q(x, A)$ finds a point $\hat{x}$ with either $F(\hat{x}) < F(x_k^*)$ or $(\hat{x} - x_k^*)^T \nabla Q(\hat{x}, A) > 0$, then the point $\hat{x}$ is already in a lower basin, hence the minimization of $Q(x, A)$ can be terminated and the minimization of $F(x)$ can be restarted at the point $\hat{x}$.

The filled function $Q(x, A)$ removes the excessive restriction on the choice of the parameter values. However, the factor $\exp(Aw(\|x - x_k^*\|^\beta)$ still exists and influences the evaluation of $Q(x, A)$ and $\nabla Q(x, A)$. A further modification can be motivated from the fact that if $f(x) > 0$, then $f(x)$ and $\ln(f(x))$ have the same extreme points. Thus, a new filled function of $F(x)$ can be obtained from

$$\begin{aligned} &-\ln[(a + \phi(F(x) - F(x_k^*))) \exp(Aw(\|x - x_k^*\|^\beta))] \\ &= -\ln(a + \phi(F(x) - F(x_k^*)) - Aw(\|x - x_k^*\|^\beta), \end{aligned}$$

where $a > 0$ is a constant.

A more general form for this kind of filled functions is

$$U(x, A) = -\eta(F(x) - F(x_k^*)) - Aw(\|x - x_k^*\|^\beta),$$

where the function $w(t)$ has the same properties as one in Section 3 and $\eta(t)$ is required to have the following properties:

(i) $\eta(0) = 0,\ \eta(t) > 0,\ \forall t \in (0, +\infty)$;
(ii) $\eta'(t) > 0$ is monotonically decreasing for $t \in (0, +\infty)$;
(iii) $\lim_{t\to 0^+} \eta'(t) = +\infty$.

It can be similarly proved that when the functions $\eta(t)$ and $w(t)$ have the desired properties, the function $U(x, A)$ is a filled function of $F(x)$ without the factor $\exp(Aw(\|x - x_k^*\|^\beta))$. That is, $U(x, A)$ has the following desired properties:

(i) the local minimizer x_k^* of $F(x)$ is a local maximizer of $U(x, A)$;
(ii) when

$$A > \frac{L\eta'(d)}{c\beta D_k^{\beta-1}},$$

any point $x \in W_k^*,\ x \neq x_k^*$ is not a stationary point of $U(x, A)$, and $x - x_k^*$ is a descent direction of $U(x, A)$ at x, where d is defined in (4.14);
(iii) for a point x in a basin of $F(x)$ lower than B_k^* with $F(x) > F(x_k^*)$, when $(x - x_k^*)^T \nabla F(x) < 0$ and

$$A < \frac{-(x - x_k^*)^T \nabla F(x) \eta'(F(x) - F(x_k^*))}{\beta w'(\|x - x_k^*\|^\beta)\|x - x_k^*\|^\beta}, \tag{4.16}$$

$x - x_k^*$ is an ascent direction of $U(x, A)$ at x. Hence, $U(x, A)$ has either minimizers or saddle points in basins lower than B_k^*.

The third property of the function $\eta(t)$ ensures the satisfaction of (4.16) for points in such a basin with $F(x) - F(x_k^*) \to 0^+$.

5. Filled functions for non-smooth optimization

This section is devoted to the filled functions for non-smooth unconstrained optimization. It is assumed that the function $F(x)$ is a composite function in the form $F(x) = f(x) + h(c(x))$, where $f(x)$ and $c(x)^T = (c_1(x) \cdots c_m(x))$ are smooth functions and $h : R^m \to R$ is convex but non-smooth [3]. Examples of these functions occur when solving a system of nonlinear equations, finding a feasible point of a system of nonlinear inequalities, penalty functions for constrained optimization and so on.

The filled functions for non-smooth optimization can be expressed in the same forms as those for smooth optimization. However, for simplicity of discussion, here we will only consider the filled functions with the function $w(t) = t$ and $\beta = 2$.

For the filled functions in the form

$$P(x, r, A) = \psi(r + F(x))\exp(-A\|x - x_k^*\|^2),$$

the function $\psi(t)$ is required to have the properties:

(i) $\psi(t) > 0$ for $t \geqslant 0$;
(ii) $\psi(t)$ is monotonically decreasing for $t \geqslant 0$;
(iii) $\psi(t_1) - \psi(t_2) \leqslant c_2(t_2 - t_1)$ for $t_2 > t_1 \geqslant 0$, where $c_2 > 0$ is a constant.

Because of the disadvantages of the filled functions in this form, we just point out here that, if

$$A(r + F(x_k^*)) > L/(2\alpha D_k),$$

then the local minimizer x_k^* of $F(x)$ is a local maximizer of $P(x, r, A)$, and the set W_k^* of $F(x)$ becomes a part of a hill of $P(x, r, A)$ at x_k^*, and $P(x, r, A)$ has either minimizers or saddle points in basins of $F(x)$ lower than B_k^* (if it exists). If

$$A(r + F^*) \geqslant L/(2\alpha D_k),$$

then global minimizers of $F(x)$ can be lost, where

$$\alpha = \min\{ \frac{\psi(r + F(x))}{c_2(r + F(x))} \mid x \in \mathcal{X}\}, \quad \|\xi\| \leqslant L, \ \forall\, \xi \in \partial F(x), \ x \in \mathcal{X}.$$

For the filled functions in the form

$$Q(x, A) = -\phi(F(x) - F(x_k^*))\exp(A\|x - x_k^*\|^2),$$

the function $\phi(t)$ is required to satisfy:

(i) $\phi(0) = 0$, $\phi(t)$ is monotonically increasing for $t \geqslant 0$;
(ii) $c_1(t_2 - t_1) \leqslant \phi(t_2) - \phi(t_1) \leqslant c_2(t_2 - t_1)$ for $t_2 > t_1 \geqslant 0$, where $0 < c_1 \leqslant c_2$ are constants.

THEOREM 5.1. *x_k^* is a local maximizer of $Q(x, A)$.*

Proof. From the fact that x_k^* is a local minimizer of $F(x)$, we have

$$Q(x_k^*, A) = 0 > Q(x, A), \ \forall x \in S_k^*, \ x \neq x_k^*.$$

Therefore, x_k^* is a local maximizer of $Q(x, A)$. □

The following lemmas are required to get the result that the set W_k^* of $F(x)$ becomes a part of a hill of $Q(x, A)$ at x_k^*.

LEMMA 5.2. *Let $x(\lambda) = x_2 + \lambda(x_1 - x_2), \ \lambda \in [0, 1]$. If*

$$(x_1 - x_2)^T(x_2 - x_k^*) \geqslant 0, \tag{5.17}$$

then $\|x(\lambda) - x_k^*\|^2$ and $(x_1 - x_2)^T(x(\lambda) - x_k^*)$ are increasing functions for $\lambda \in [0, 1]$ and hence

$$(x_1 - x_2)^T(x(\lambda) - x_k^*) \geqslant (x_1 - x_2)^T(x_2 - x_k^*) \geqslant 0, \ \forall\lambda \in [0, 1].$$

Proof. The conclusions can directly be obtained from condition (5.17) and the following expansions

$$\begin{aligned}\|x(\lambda)-x_k^*\|^2 &= \|\lambda(x_1-x_2)+(x_2-x_k^*)\|^2=\|x_2-x_k^*\|^2\\ &\quad +\lambda^2\|x_1-x_2\|^2+2\lambda(x_1-x_2)^T(x_2-x_k^*),\\ (x_1-x_2)^T(x(\lambda)-x_k^*) &= (x_1-x_2)^T[\lambda(x_1-x_2)+(x_2-x_k^*)]\\ &= \lambda\|x_1-x_2\|^2+(x_1-x_2)^T(x_2-x_k^*).\end{aligned}$$

□

LEMMA 5.3. *Let x_1-x_2 be an ascent segment of $F(x)$. If (5.17) holds, then x_1-x_2 is a descent segment of $Q(x,A)$ for any $A>0$.*

Proof. Let $f(\lambda)=F(x(\lambda))=F(x_2+\lambda(x_1-x_2))$ and $g(\lambda)=Q(x(\lambda),A)=-\phi(f(\lambda)-F(x_k^*))\exp(A\|x(\lambda)-x_k^*\|^2)$. Then $f(\lambda)$, $\phi(f(\lambda)-F(x_k^*))$, $\exp(\|x(\lambda)-x_k^*\|^2)$ and hence $\phi(f(\lambda)-F(x_k^*))\exp(A\|x(\lambda)-x_k^*\|^2)$ are increasing functions for $\lambda\in[0,1]$. Therefore, $g(\lambda)$ is decreasing for $\lambda\in[0,1]$, and x_1-x_2 is a descent segment of $Q(x,A)$. □

LEMMA 5.4. *Let $x_2\in V_k^*=W_k^*\backslash\{S_k^*\}$, x_1-x_2 be a descent segment of $F(x)$. If*

$$\frac{(x_1-x_2)^T(x_2-x_k^*)}{\|x_1-x_2\|\cdot\|x_2-x_k^*\|}\geqslant\hat{c}>0, \tag{5.18}$$

$$A\geqslant\frac{c_2L}{2c_1\hat{c}dD_k}, \tag{5.19}$$

the x_1-x_2 is also a descent segment of $Q(x,A)$, where d is defined in (4.14), D_k is given in (2.1).

Proof. Since x_1-x_2 is a descent segment of $F(x)$, $f(\lambda)$ is decreasing for $\lambda\in[0,1]$. It is clear that if

$$\begin{aligned}g(\lambda_2) &\stackrel{\triangle}{=} -\phi(f(\lambda_2)-F(x_k^*))\exp(A\|x(\lambda_2)-x_k^*\|^2)\\ &< g(\lambda_1)\stackrel{\triangle}{=}-\phi(f(\lambda_1)-F(x_k^*))\exp(A\|x(\lambda_1)-x_k^*\|^2)\end{aligned} \tag{5.20}$$

holds for any $1\geqslant\lambda_2>\lambda_1\geqslant 0$, then x_1-x_2 is a descent segment of $Q(x,A)$.

From (5.20), we have

$$\frac{\phi(f(\lambda_1)-F(x_k^*))}{\phi(f(\lambda_2)-F(x_k^*))}<\exp(A(\|x(\lambda_2)-x_k^*\|^2-\|x(\lambda_1)-x_k^*\|^2)). \tag{5.21}$$

(5.18) indicates $\|x(\lambda)-x_k^*\|^2$ is increasing for $\lambda\in[0,1]$ (see Lemma 5.2). It follows from the Taylor expansion of e^t that if

$$\begin{aligned}&\frac{\phi(f(\lambda_1)-F(x_k^*))-\phi(f(\lambda_2)-F(x_k^*))}{\phi(f(\lambda_2)-F(x_k^*))}\\ &< A(\|x(\lambda_2)-x_k^*\|^2-\|x(\lambda_1)-x_k^*\|^2)\\ &= A(\lambda_2-\lambda_1)(x_1-x_2)^T[2(x_2-x_k^*)+(\lambda_2+\lambda_1)(x_1-x_2)]\end{aligned} \tag{5.22}$$

holds, then (5.21) will be satisfied. Under the condition (5.18), we know that (5.22) is equivalent to

$$A > \frac{\phi(f(\lambda_1) - F(x_k^*)) - \phi(f(\lambda_2) - F(x_k^*))}{\phi(f(\lambda_2) - F(x_k^*))} \times \frac{1}{(\lambda_2 - \lambda_1)(x_1 - x_2)^T[2(x_2 - x_k^*) + (\lambda_2 + \lambda_1)(x_1 - x_2)]}. \tag{5.23}$$

From the properties of the function $\phi(t)$ and the definition of d, (5.23) is implied by

$$A > \frac{c_2(F(x_2 + \lambda_1(x_1 - x_2)) - F(x_2 + \lambda_2(x_1 - x_2)))}{c_1 d(\lambda_2 - \lambda_1)(x_1 - x_2)^T[2(x_2 - x_k^*) + (\lambda_2 + \lambda_1)(x_1 - x_2)]}. \tag{5.24}$$

Since $x_1 - x_2$ is a descent segment of $F(x)$, it follows from the non-smooth analysis (see [1, 3]) that we have

$$\begin{aligned} &\lim_{\lambda_2 \to \lambda_1} \frac{F(x_2 + \lambda_2(x_1 - x_2)) - F(x_2 + \lambda_1(x_1 - x_2))}{\lambda_2 - \lambda_1} \\ &= \max\{(x_1 - x_2)^T \xi \mid \xi \in \partial F(x(\lambda_1))\} = (x_1 - x_2)^T \hat{\xi} \leqslant 0, \end{aligned}$$

where $\hat{\xi}$ is the element in $\partial F(x(\lambda_1))$ such that $(x_1 - x_2)^T \hat{\xi}$ maximizes the value $(x_1 - x_2)^T \xi$ for all $\xi \in \partial F(x(\lambda_1))$. Taking limit in the right hand side of (5.24) generates

$$A \geqslant \frac{-c_2(x_1 - x_2)^T \hat{\xi}}{2c_1 d(x_1 - x_2)^T (x(\lambda_1) - x_k^*)}. \tag{5.25}$$

Since $(x_1 - x_2)^T (x(\lambda) - x_k^*) \geqslant (x_1 - x_2)^T (x_2 - x_k^*) \geqslant \hat{c}\|x_1 - x_2\| \cdot \|x_2 - x_k^*\| \geqslant \hat{c} D_k \|x_1 - x_2\|$ and $|(x_1 - x_2)^T \hat{\xi}| \leqslant \|x_1 - x_2\| \cdot \|\hat{\xi}\| \leqslant L\|x_1 - x_2\|$, (5.25) is implied by condition (5.19). Therefore, when the value of A satisfies condition (5.19), the inequality (5.20) holds, and $x_1 - x_2$ is a descent segment of $Q(x, A)$. □

THEOREM 5.5. *When the value of A satisfies inequality (5.19), $x - x_k^*$ is a descent segment of $Q(x, A)$ for any $x \in W_k^*$, $x \neq x_k^*$, that is, W_k^* becomes a part of a hill of $Q(x, A)$ at x_k^*. Therefore, there is neither a minimizer nor a saddle point of $Q(x, A)$ in the set $W_k^* \backslash \{x_k^*\}$.*

Proof. If $x \in S_k^*$, $x \neq x_k^*$, then $x - x_k^*$ is an ascent segment of $F(x)$. Using the same method used in the proof of lemma 5.3, we can determine that $x - x_k^*$ is a descent segment of $Q(x, A)$.

If $x \in V_k^* = W_k^* \backslash S_k^*$, then there exists some $\hat{\lambda} \in (0, 1)$ such that $x(\hat{\lambda}) = x_k^* + \hat{\lambda}(x - x_k^*) \in \partial S_k^*$ where ∂S_k^* is the boundary of the set S_k^*. The result in the previous paragraph shows that $x(\hat{\lambda}) - x_k^*$ is a descent segment of $Q(x, A)$. The segment $x - x(\hat{\lambda})$ can be divided into a number of subsegments such that each subsegment is either a descent or an ascent one of $F(x)$. It is clear that for such subsegments, the first inequality of condition (5.18) becomes an equality with $\hat{c} = 1$. Then it follows from lemmas 5.3 and 5.4 that these subsegments are descent segments of $Q(x, A)$, and hence $x - x_k^*$ is a descent segment of $Q(x, A)$. This completes the proof. □

THEOREM 5.6. *Let* $x_2 \notin W_k^*$, $F(x_2) > F(x_k^*)$ *and* $x_1 - x_2$ *be a strong descent segment of* $F(x)$. *If* $(x_1 - x_2)^T(x_2 - x_k^*) > 0$ *and*

$$A < \frac{-c_1(x_1 - x_2)^T\hat{\xi}}{2(x_1 - x_2)^T(x(\lambda) - x_k^*)\phi(F(x(\lambda)) - F(x_k^*))}, \quad \forall\lambda \in [0, 1], \tag{5.26}$$

where $\hat{\xi} \in \partial F(x(\lambda))$ *such that* $(x_1 - x_2)^T\hat{\xi} = \max\{(x_1 - x_2)^T\xi \mid \xi \in \partial F(x(\lambda))\}$, *then* $x_1 - x_2$ *is an ascent segment of* $Q(x, A)$.

Proof. If

$$\begin{aligned} &-\phi(f(\lambda_2) - F(x_k^*))\exp(A\|x(\lambda_2) - x_k^*\|^2) \\ &> -\phi(f(\lambda_1) - F(x_k^*))\exp(A\|x(\lambda_1) - x_k^*\|^2) \end{aligned} \tag{5.27}$$

holds for any $1 \geqslant \lambda_2 > \lambda_1 \geqslant 0$, then $x_1 - x_2$ is an ascent segment of $Q(x, A)$. From (5.27) we obtain

$$\begin{aligned} &\frac{\exp(A\|x(\lambda_2) - x_k^*\|^2) - \exp(A\|x(\lambda_1) - x_k^*\|^2)}{\exp(A\|x(\lambda_1) - x_k^*\|^2)} \\ &< \frac{\phi(f(\lambda_1) - F(x_k^*)) - \phi(f(\lambda_2) - F(x_k^*))}{\phi(f(\lambda_2) - F(x_k^*))}. \end{aligned} \tag{5.28}$$

The property of the function $\phi(t)$ implies that if

$$\begin{aligned} &\frac{\exp(A\|x(\lambda_2) - x_k^*\|^2) - \exp(A\|x(\lambda_1) - x_k^*\|^2)}{\exp(A\|x(\lambda_1) - x_k^*\|^2)} \\ &\leqslant \frac{c_1(f(\lambda_1) - f(\lambda_2))}{\phi(f(\lambda_2) - F(x_k^*))} \\ &\quad = \frac{c_1(F(x_2 + \lambda_1(x_1 - x_2)) - F(x_2 + \lambda_2(x_1 - x_2))}{\phi(F(x_2 + \lambda_2(x_1 - x_2)) - F(x_k^*))}, \end{aligned} \tag{5.29}$$

then inequality (5.28) is satisfied. Using $\lambda_2 - \lambda_1$ to divide the both hand sides of (5.29) and then taking limits as $\lambda_2 \to \lambda_1$, we obtain

$$2A(x_1 - x_2)^T(x(\lambda_1) - x_k^*) \leqslant \frac{-c_1(x_1 - x_2)^T\hat{\xi}}{\phi(F(x(\lambda_1)) - F(x_k^*))}.$$

Therefore, when (5.26) holds, $x_1 - x_2$ is an ascent segment of $Q(x, A)$. □

Since $x_2 \notin W_k^*$, i.e. x_2 is in a basin of $F(x)$ lower than B_k^*, the inequality (5.26) will hold when $F(x(\lambda)) - F(x_k^*) \to 0^+$ as λ increases.

Finally, for the filled functions in the form

$$U(x, A) = -\eta(F(x) - F(x_k^*)) - A\|x - x_k^*\|^2.$$

When the function $\eta(t)$ satisfies the properties (i) and (ii) of the function $\phi(t)$, $U(x, A)$ is a desirable filled function and has the following properties:

(i) the local minimizer x_k^* of $F(x)$ is a local maximizer of $U(x, A)$;

(ii) if $A \geqslant c_2 L/(2\hat{c} D_k)$, then the set W_k^* of the function $F(x)$ becomes a part of a hill of the function $U(x, A)$ with peak x_k^*, where $\hat{c}$ is defined in (5.18);

(iii) if $x_1 - x_2$ is a descent segment of $F(x)$ in a basin lower than B_k^* and $(x_1 - x_2)^T(x_2 - x_k^*) > 0$, and if

$$A < \frac{c_1(F(x_2 + \lambda_1(x_1 - x_2)) - F(x_2 + \lambda_2(x_1 - x_2)))}{(\lambda_2 - \lambda_1)(x_1 - x_2)^T[2(x_2 - x_k^*) + (\lambda_2 + \lambda_1)(x_1 - x_2)]}$$

holds for all $1 \geqslant \lambda_2 > \lambda_1 \geqslant 0$, then $x_1 - x_2$ is an ascent segment of $U(x, A)$.

6. Concluding remarks

In the paper, we are concerned with some general forms of filled functions used for unconstrained global minimization of a continuous function (smooth or non-smooth) of several variables. These filled functions have either one or two adjustable parameters. Conditions on the objective function and on the values of parameters are given so that the constructed functions have the desired properties of filled functions. Note that the forms of filled functions considered only use two static parameters or one static parameter. It will be useful from the practical point of view to extend the filled functions to other types, where the parameters are dynamically adjusted.

References

1. Clarke, F.H. (1983), *Optimization and Non-smooth Analysis*, John Wiley & Sons, New York.
2. Dixon, L.C.W., Gomulka, J. and Herson, S.E. (1976), Reflections on Global Optimization Problem, in L.C.W. Dixon (ed), *Optimization in Action*, Academic Press, New York, pp. 398–435.
3. Fletcher, R. (1981), *Practical Methods of Optimization, Vol. 2, Constrained Optimization*, John Wiley & Sons, New York.
4. Ge, R.P. (1987), The Theory of Filled Function Methods for Finding Global Minimizers of Nonlinearly Constrained Minimization Problems, *Journal of Computational Mathematics* 5(1): 1–9.
5. Ge, R.P. (1990), A Filled Function Method for Finding a Global Minimizer of a Function of Several Variables, *Mathematical Programming* 46: 191–204.

6. Ge, R.P. and Qin, Y.F. (1987), A Class of Filled Functions for Finding a Global Minimizer of a Function of Several Variables, *Journal of Optimization Theory and Applications* 54(2): 241–252.
7. Ge, R.P. and Qin, Y.F. (1990), The Globally Convexized Filled Functions for Globally Optimization, *Applied Mathematics and Computations* 35: 131–158.
8. Horst, R. and Pardalos, P.M. (eds) (1995), *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
9. Horst, R., Pardalos, P.M. and Thoai, N.V. (2000), *Introduction to Global Optimization* (2nd edition), Kluwer Academic Publishers, Dordrecht, The Netherlands.
10. Kong, M. and Zhuang, J.N. (1996), A Modified Filled Function Method for Finding a Global Minimizer of a Non-smooth Function of Several Variables, *Numerical Mathematics—A Journal of Chinese Universities* 18(2): 165–174.
11. Levy, A.V. and Gómez, S. (1985), The Tunneling Method Applied to Global Optimization, in: P.T. Boggs, R.H. Byrd and R.B. Schnabel (eds), *Numerical Optimization*, SIAM, pp. 213–244.
12. Wales, D.J. and Scheraga, H.A. (1999), Global Optimization of Clusters, Crystals and Biomolecules, *Science*, 285: 1368–1372.
13. Zhuang, J.N. (1994), A Generalized Filled Function Method for Finding the Global Minimizer of a Function of Several Variables, *Numerical Mathematics—A Journal of Chinese Universities* 16(3): 279–287.