A New One-parameter Filled Function Method and Its Application in Pathological Analysis

Qiyue He [1], Ying Zhang [1], and Shenggang Wang [2,*]

[1] Department of Mathematics, Zhejiang Normal University, Jinhua 321004, P R China

[2] College of Agriculture and Biological Engineering, Jinhua Polytechnic, Jinhua 321000, P R China

* Correspondence: 736853363@qq.com

Abstract: For the unconstrained global optimization problems, a new one-parameter filled function is constructed with an assumption that the objective function needs to be continuous and differentiable but do not following the Lipschitz condition, it overcomes the problem that the numerical experiment of the filled function with exponential terms may decrease the efficiency of filled function algorithm. The related properties of the filled function were studied and the feasibility of this method was verified by numerical analysis. Applying the filled function method to the study of pathological factors in renal cell carcinoma metastasis, in the process of treating the renal cell carcinoma metastasis by the two classifications, the filled function method is used to minimize the cross entropy loss function, and the global optimal solution is obtained.

Keywords: Global optimal; Unconstraint; One parameter; Filled function method; Renal cell carcinoma metastasis

1. Introduction

Nowadays the global optimization problems have attracted more and more scholars' attention. Many problems encountered in the fields of finance, image and engineering can be abstracted into global optimization problems. The filled function method used in this paper is a deterministic algorithm for solving nonlinear global optimization problems. It is convenient that this method would not stop after finding a local minimizer and keep going to figure out whether there are better local minimizers or not. Especially in the unconstrained global optimization problem, it is more effective. The filled function method was originally proposed by Ge [1]. While, in that reference the constructed filled functions made the effect in numerical calculations unsatisfactory and the filled function given in Ref.[1,2] cannot guarantee that there must be local minimizers in the valleys lower than the current valley. In response to the above problems, the forms of filled functions were modified and more complete definitions of filled functions were given by some later researchers [3-8]. The implementation of the filled function algorithm is generally divided into two phases: the minimization phase and the filled phase. The main idea is firstly to obtain a local minimizer through the local optimization algorithm, and then minimize the filled function constructed at that point. This two-phase repeated loop iteration makes it possible to jump out of the valley where the current local minimizer is located and find the global minimizer.

There are many scholars having constructed many different types of filled functions but these different functions always come with problems. For example, two-parameters filled functions are difficult in tuning and have low computational efficiency [3,4]; Although one-parameter filled functions are simpler, the index items and logarithmic items are prone to false stable points which makes the calculation effect poor [9-11]. In this paper, we propose a continuous and differentiable one-parameter filled function which is simple and overcomes the deficiencies mentioned above. Most of the previous filled functions study only carried out theoretical research, and did not give the application of the filled function algorithm in practical problems. We attempt to study the pathological factors of renal cell carcinoma metastasis. The filled function method is used to minimize the loss function so that the optimal parameters affecting the factors of renal cell carcinoma metastasis can be obtained in the process of renal cell carcinoma metastasis with two classification treatment.

2. A New Filled Function and Its Properties

Consider the following unconstrained global optimization problem:

$$\begin{aligned} &min && f(\boldsymbol{x}) \\ &s.t. && \boldsymbol{x} \in R^n. \end{aligned}$$

Throughout this paper we make the following assumptions:

2.1. Assumptions and Definitions

Assumption 2.1.1. $f(\boldsymbol{x})$ is Lipchitz continuous on R^n.

Assumption 2.1.2. $f(\boldsymbol{x})$ is coercive, i.e. $f(\boldsymbol{x}) \to +\infty$, as $\| \boldsymbol{x} \| \to +\infty$.

According to Assumption 2.1.2, there exist a robust compact set $\Omega \in R^n$ such that all global minimizers of

$f(\boldsymbol{x})$ can be covered by Ω.

Assumption 2.1.3. $f(\boldsymbol{x})$ has only a finite number of local minimum, but there can be an infinite local minimizers. This paper adopts the definition of the filled function in Ref. [3].

Definition 2.1.1. [3] $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ is called a filled function of $f(\boldsymbol{x})$ at a local minimizer $\boldsymbol{x}^*$ if $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ have the following properties:

(1) $\boldsymbol{x}^* \in \Omega$ is a local maximizer of $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$.

(2) $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ has no stationary point in the region $U_1 = \{\boldsymbol{x} \mid f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*), \boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}\}$

(3) If $\boldsymbol{x}^*$ is not a global minimizer of $f(\boldsymbol{x})$, then $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ will have a minimizer in the region

$$U_2 = \{\boldsymbol{x} \mid f(\boldsymbol{x}) < f(\boldsymbol{x}^*), \boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}\}.$$

The Definition 2.1.1 indicates that if $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ is a filled function that satisfies the definition and $\boldsymbol{x}^*$ is not a global minimizer, then by introducing a filled function, the next local minimizer obtained is guaranteed to be only in the function value ratio in the process of minimizing $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$. A small area of $f(\boldsymbol{x})$ exists such that the current local minimizer is jumped out, and the iteration is repeated until the termination condition is satisfied, thereby obtaining a global minimizer lower than the current local minimizer.

2.2. A New One-parameter Filled Function

The new one-parameter filled function at the local minimizer $\boldsymbol{x}^*$ has the following form:

$$P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = \frac{1}{1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2} \varphi_\gamma(f(\boldsymbol{x}) - f(\boldsymbol{x}^*) + \gamma). \quad (1)$$

Where

$$\varphi_\gamma(t) = \begin{cases} 0, & t \leq 0, \\ -\dfrac{(t-\gamma)^3}{\gamma^3} - \dfrac{2t}{\gamma} + 1, & 0 < t \leq \gamma, \\ 1 & t > \gamma. \end{cases} \quad (2)$$

Next we will show that the function $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ is a filled function satisfying Definition 2.1.1.

Theorem 2.2.1. If $\boldsymbol{x}^*$ is a local minimizer of $f(\boldsymbol{x})$, for any $\gamma > 0$, $\boldsymbol{x}^*$ must be a local maximizer of $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$.

Proof. $\boldsymbol{x}^*$ is a local minimizer of $f(\boldsymbol{x})$. From Assumption 2.1.1, $f(\boldsymbol{x})$ is continuously differentiable on R^n, then there exists a neighborhood $N(\boldsymbol{x}^*, \delta)$ of $\boldsymbol{x}^*$ with $\delta > 0$, such that $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$. Namely for $\forall \gamma > 0$, then

$$f(\boldsymbol{x}) - f(\boldsymbol{x}^*) + \gamma \geq \gamma,$$

since that $\varphi_\gamma(t) = 1$. Then for all $\boldsymbol{x} \in N(\boldsymbol{x}^*, \delta)$ and $\boldsymbol{x} \neq \boldsymbol{x}^*$, we have

$$P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = \frac{1}{1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2} \cdot 1 < \frac{1}{1 + \| \boldsymbol{x}^* - \boldsymbol{x}^* \|^2} = P(\boldsymbol{x}^*, \boldsymbol{x}^*, \gamma),$$

Hence, $\boldsymbol{x}^*$ must be a local maximizer of $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$.

Theorem 2.2.2. $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ has no stationary point in the region $U_1 = \{\boldsymbol{x} \mid f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*), \boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}\}$, and for any $\boldsymbol{x} \in \Omega$ and $\boldsymbol{x} \neq \boldsymbol{x}^*$, let $d(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x}^*$, then $\nabla^T P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) d(\boldsymbol{x}) < 0$.

Proof. For $\forall \boldsymbol{x} \in U_1$, if $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$ and $\boldsymbol{x} \neq \boldsymbol{x}^*$, we can have $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = \dfrac{1}{1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2}$, then

$$\nabla P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = -\frac{2(\boldsymbol{x} - \boldsymbol{x}^*)}{(1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2)^2},$$

for $\forall \boldsymbol{x} \in U_1$, $\nabla P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) \neq 0$, then $\nabla P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) \neq 0$ is not a stationary point of $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$. For $\boldsymbol{x}$ that meets the above conditions, we have

$$\nabla^T P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) d(\boldsymbol{x}) = -\frac{2(\boldsymbol{x} - \boldsymbol{x}^*)}{(1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2)^2} (\boldsymbol{x} - \boldsymbol{x}^*) = -\frac{2(\boldsymbol{x} - \boldsymbol{x}^*)^2}{(1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2)^2} < 0.$$

Therefore, $d(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x}^*$ is the downward direction of $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$. If choose the direction $d(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x}^*$, we will find a point smaller than $P(\boldsymbol{x}^*, \boldsymbol{x}^*, \gamma)$.

Remark 2.2.1. Theorem 2.2.2 states that the filled function $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ has a downward trend on

$$U_1 = \{\boldsymbol{x} \mid f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*), \boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}\},$$

that is, the current local minimizer is jumped out, and the found point is a local minimizer that is smaller than the current one.

Theorem 2.2.3. If $\boldsymbol{x}^*$ is not the global minimizer of $f(\boldsymbol{x})$, when $0 < \gamma \leq \beta_0 = f(\boldsymbol{x}^*) - f(\bar{\boldsymbol{x}})$, $P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma)$ must have local minimizer $\bar{\boldsymbol{x}}$ on $U_2 = \{\boldsymbol{x} \mid f(\boldsymbol{x}) < f(\boldsymbol{x}^*), \boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}\}$.

Proof. For $\forall \boldsymbol{x} \in U_2$ and $f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$, we have

$$P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = \frac{1}{1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2} \varphi_\gamma(f(\boldsymbol{x}) - f(\boldsymbol{x}^*) + \gamma),$$ then

$$\nabla P(\boldsymbol{x}, \boldsymbol{x}^*, \gamma) = -\frac{2(\boldsymbol{x} - \boldsymbol{x}^*)}{(1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2)^2} \cdot \varphi_\gamma(f(\boldsymbol{x}) - f(\boldsymbol{x}^*) + \gamma) + \frac{1}{1 + \| \boldsymbol{x} - \boldsymbol{x}^* \|^2} \cdot \varphi_\gamma'(f(\boldsymbol{x}) - f(\boldsymbol{x}^*) + \gamma) \cdot \nabla f(\boldsymbol{x}).$$

Next we should proof that there is a point $\bar{\boldsymbol{x}}$ such that

$\nabla P(\bar{x}, x^*, \gamma) = 0$, which means that $\bar{x}$ is a local minimizer on U_2. If x^* is not a global minimizer, there must be $\bar{x} \in U_2$ and $f(\bar{x}) < f(x^*)$. Follow the condition in Theorem 2.2.3, $\beta_0 = f(x^*) - f(\bar{x}) > 0$, when $0 < \gamma \le \beta_0$,we can have $f(\bar{x}) - f(x^*) \le -\gamma$, which can be transformed into $f(\bar{x}) - f(x^*) + \gamma \le 0$. Then $\varphi_\gamma(f(x) - f(x^*) + \gamma) = 0$ can be obtained. So we can draw that $x \in N(\bar{x}, \delta)$, $\nabla P(x, x^*, \gamma) = 0$. Finally, there must be a $\bar{x}$ on U_2 is a local minimizer of $P(x, x^*, \gamma)$.

Remark 2.2.2. It can be seen from Theorem 2.2.3 that if the parameter γ is sufficiently small, the condition $0 < \gamma \le \beta_0 = f(x^*) - f(\bar{x})$ can be guaranteed. The above proofs of Theorem 2.2.1-Theorem 2.2.3 represent that the filled function constructed in this paper satisfies the three conditions of the filled function definition, so the filled function we proposed is feasible.

Theorem 2.2.4. Suppose x^* is a local minimizer of $f(x)$, if $\forall x_1, x_2 \in \Omega$ satisfy $f(x_1) > f(x^*)$ and $f(x_2) > f(x^*)$, then the necessary and sufficient conditions for $\| x_2 - x^* \| > \| x_1 - x^* \|$ are $P(x_2, x^*, \gamma) < P(x_1, x^*, \gamma)$.

Proof. Since $f(x_1) > f(x^*)$, $f(x_2) > f(x^*)$, from (1) and (2), we have

$$P(x, x^*, \gamma) = \frac{1}{1 + \| x - x^* \|^2} \cdot 1,$$

then

$$\begin{aligned} & P(x_1, x^*, \gamma) - P(x_2, x^*, \gamma) \\ &= \frac{1}{1 + \| x_1 - x^* \|^2} - \frac{1}{1 + \| x_2 - x^* \|^2} \\ &= \frac{\| x_2 - x^* \| - \| x_1 - x^* \|}{(\| x_1 - x^* \|^2) \cdot (1 + \| x_2 - x^* \|^2)}, \end{aligned}$$

combine condition $\| x_2 - x^* \| > \| x_1 - x^* \|$, we have $P(x_1, x^*, \gamma) - P(x_2, x^*, \gamma) > 0$. And from Theorem 2.2.1 we have $P(x_1, x^*, \gamma) < P(x^*, x^*, \gamma)$. Furthermore, $P(x_2, x^*, \gamma) < P(x_1, x^*, \gamma) < P(x^*, x^*, \gamma) = 1$.

Conversely, the sufficiency is obvious, if $P(x_2, x^*, \gamma) < P(x_1, x^*, \gamma)$, then

$$\frac{1}{1 + \| x_2 - x^* \|^2} < \frac{1}{1 + \| x_1 - x^* \|^2},$$

hence $\| x_2 - x^* \| > \| x_1 - x^* \|$.

Remark 2.2.3. Theorem 2.2.4 illustrates that in U_1 , the farther away from the local minimizer x^* of $f(x)$, the smaller the value of the function $P(x, x^*, \gamma)$. So in the minimization process, either the point corresponding to $f(x)$ smaller than $f(x^*)$ is found, or it will run all the way to the boundary, thus avoiding the phenomenon that the filled function repeats the round-trip operation during the minimization process.

3. The Filled Function Algorithm and Numerical Experiments

Now, we present algorithm AFF as follows [12]:

3.1. AFF Algorithm

1. Choose $\alpha > 0$ as the search step size, $0 < \gamma \le \beta_0$; K is the upper bound of the number of outer loops k ; m is the upper bound of the number of inner loops i ; $d_i (i = 1, 2, \cdots m; m > 2n)$ is the direction of selection; Set that $i := 1$, $k := 1$.

2. Choose $x_k^0 \in \Omega$ as initial point; Starting from x_k^* ; the local minimizer x_k^0 and the local minimum $f(x_k^*)$ of the objective function $f(x)$ are obtained by using the existing local optimization algorithm.

3. Construct a filled function $P(x, x^*, \gamma)$ at a local minimizer x_k^* .

4. If $i \le m$, the filled function $P(x, x_k^*, \gamma)$ is minimized with $\bar{x}_k := x_k^* + \alpha d_i \in \Omega$ as the starting point, and a local minimizer $x_{\gamma k}^*$ of $P(x, x_k^*, \gamma)$ is obtained, and turn into step 5; otherwise, let $k := k + 1$, and turn into step 7.

5. If $x_{\gamma k}^* \in \Omega$, turn into step 6; otherwise, let $i := i + 1$, turn into step 4.

6. Starting from x_{rk}^* , $f(x)$ is minimized by a local optimization algorithm to obtain a new local minimizer x_{k+1}^* . If $f(x_{k+1}^*) < f(x_k^*)$, let $f(x_k^*) := f(x_{k+1}^*)$, $x_k^* := x_{k+1}^*$, go to step 3; otherwise, go to step 7.

7. If $k \le K$, reduce parameter γ , let $i := 1$, if $\gamma > \beta_0$, let $\gamma = \beta_0$, go to step 4; otherwise stop, output x_k^* and $f(x_k^*)$. That is, the global optimization minimizer or the approximate global optimization minimizer x_k^* is obtained.

The algorithm is mainly divided into two phases for looping. $\gamma = 1$ is selected as the initial value. In the first phase, the objective function $f(x)$ is minimized by a local optimization algorithm to obtain a local minimizer x_k^* .Then the second phase is to minimize the new filled function $P(x, x_k^*, \gamma)$ to get the local minimizer $x_{\gamma k}^*$. If $f(x_{\gamma k}^*) < f(x_k^*)$, then use $x_{\gamma k}^*$ as the new initial point and re-enter the first phase, otherwise, reduce the parameter γ ,re-enter the second phase. Repeat the above process until $k > K$ is satisfied. Then the global minimizer or the approximate global minimizer is obtained.

For the above algorithm AFF, we use Python3.5 to program and verify it through the following numerical experiments.

3.2. Numerical Experiments

The meanings of the symbols used are as follows:
k : The number of iterations.
$\boldsymbol{x}_k^0$: The initial point when the original function is locally minimized.
$\boldsymbol{x}_k^*$: The local minimizer of the original function.
$f(\boldsymbol{x}_k^*)$: The local minimum of the original function.
$\bar{\boldsymbol{x}}_k$: The initial point when the filled function is locally minimized, obtained by $\bar{\boldsymbol{x}}_k := \boldsymbol{x}_k^* + \alpha d_i$.
$\boldsymbol{x}_{rk}^*$: The local minimizer point when the filled function is locally minimized.

Problem 1 (Six-Hump Camel-back)[4]

$$\min \quad f(\boldsymbol{x}) = 4\boldsymbol{x}_1^2 - 2.1\boldsymbol{x}_1^4 + \frac{1}{3}\boldsymbol{x}_1^6 - \boldsymbol{x}_1\boldsymbol{x}_2 - \boldsymbol{x}_2^2 + \boldsymbol{x}_2^4$$
$$-3 \le \boldsymbol{x}_i \le 3, \qquad i = 1,2.$$

The results are shown in Table I below:

Table 1. Six-hump camel-back

k	$\boldsymbol{x}_k^0$	$\boldsymbol{x}_k^*$	$f(\boldsymbol{x}_k^*)$	$\bar{\boldsymbol{x}}_k$	$\boldsymbol{x}_{rk}^*$
1	(-2,1)	(-1.60710452, 0.56865148)	2.10425031	(-1.08710452, 0.86865148)	(-0.9999955, 0.99999848)
2	(-0.9999955, 0.99999848)	(0.08984214, 0.7126564)	-1.03162845	\	\

Obviously, there are several local minimizers in the Problem 1. From Table 1, $\boldsymbol{x}_k^* = (0.08984214,\ 0.7126564)^T$ is the global minimizer of the Problem 1, and $f(\boldsymbol{x}_k^*) = -1.03162845$ is the global minimum.

Problem 2 (Rastrigin) [4]

$$\min \quad f(\boldsymbol{x}) = \boldsymbol{x}_1^2 + \boldsymbol{x}_2^2 - \cos(18\boldsymbol{x}_1) - \cos(18\boldsymbol{x}_2)$$
$$-2 \le \boldsymbol{x}_i \le 2, \qquad i = 1,2.$$

The results are shown in Table II below:

Table 2. Rastrigin

k	$\boldsymbol{x}_k^0$	$\boldsymbol{x}_k^*$	$f(\boldsymbol{x}_k^*)$	$\bar{\boldsymbol{x}}_k$	$\boldsymbol{x}_{rk}^*$
1	(1,1)	(1.04075871, 1.04075871)	0.17977497	(1.04075871, 0.74075871)	(1.04075831, 0.61127157)
2	(1.04075831, 0.61127157)	(1.04075871, 0.69384446)	-0.42571623	(0.04075871, -0.24075871)	(0.04075871, -0.24075871)
3	(0.04075871, -0.24075871)	(2.96461066e-11, -3.46923815e-01)	-1.87890065	(2.96461066e-11, -3.46923815e-05)	(1.09370176e-11, 7.03974071e-04)
4	(1.09370176e-11, 7.03974071e-04)	(1.13890947e-13, -3.83195090e-09)	-2	\	\

From Table 2, it can be found that $\boldsymbol{x}_k^* = (1.13890947e{-}13, -3.83195090e{-}09)^T$ is the global minimizer of Problem 2, and the obtained global minimum $f(\boldsymbol{x}_k^*) = -1.9999999999999978$.

Problem 3 (3-Hump back Camel) [4]

$$\min \quad f(\boldsymbol{x}) = 2\boldsymbol{x}_1^2 - 1.05\boldsymbol{x}_1^4 + \frac{1}{6}\boldsymbol{x}_1^6 - \boldsymbol{x}_1\boldsymbol{x}_2 + \boldsymbol{x}_2^2$$
$$-3 \le \boldsymbol{x}_i \le 3, \qquad i = 1,2.$$

The results are shown in Table III below:

Table 3. 3-hump back camel

k	$\boldsymbol{x}_k^0$	$\boldsymbol{x}_k^*$	$f(\boldsymbol{x}_k^*)$	$\bar{\boldsymbol{x}}_k$	$\boldsymbol{x}_{rk}^*$
1	(-3,2)	(-1.74755242, -0.87377621)	0.29863844	(-0.14755242,-0.07377621)	(0.66370003,1.94295319)
2	(0.66370003,1.94295319)	(3.80334484e-10,3.19560893e-10)	2.69887777e-19	\	\

It can be obtained from Table 3 that $\boldsymbol{x}_k^* = (3.80334484e{-}10,\ 3.19560893e{-}10)^T$ is the global minimizer of Problem 3, and $f(\boldsymbol{x}_k^*) = 2.69887777e{-}19$ is the global minimum.

According to the results of the above examples, the results are compared with the previous paper results:

From the Table 4, it is concluded that the number of iterations in Problem 1 and Problem 3 is less than that in Ref. [4], and the results obtained in the numerical Problem 1 and Problem 2 are substantially the same as those in the Ref. [4]. The result of the numerical Problem 3 is less than the result of the Ref. [4], so the filled function we proposed is feasible and more effective.

4. The Application of Filled Function Method in Pathological Analysis of Renal Cell Carcinoma

Metastasis

In order to illustrate the filled function and its algorithm in this paper are feasible and effective, we try to combine it with the actual medical case. The case [13] is as follows: In order to study the clinical and pathological factors related to renal cell carcinoma metastasis, a researcher investigated a group of patients undergoing radical nephrectomy and collected renal cancer specimens. The results of 26 cases are shown in Table 5.

Table 4. Comparison table

Numerical experiment	Comparison	Cycle index	Initial point	The global minimizer	The global minimum
1	Our work	2	(-2,1)	(0.08984214 , 0.7126564)	-1.03162845
	Ref.[4]	3	(-2,1)	(-0.08983722 , -0.712699468)	-1.03162843
2	Our work	4	(1,1)	(1.13890947e-13 , -3.83195090e-09)	-1.9999999999999978
	Ref.[4]	4	(0,0)	(-0.00000002 , -0.00000002)	-2
3	Our work	2	(-3,2)	(3.80334484e-10 , 3.19560893e-10)	2.69887777e-19
	Ref.[4]	3	(-2,-1)	(-0.00001356 , 0.00000492)	4.58551865e-10

Table 5. Renal cell carcinoma metastasis data

x_1	x_2	x_3	x_4	x_5	y	x_1	x_2	x_3	x_4	x_5	y
59	2	43.4	2	1	0	31	1	47	2	1	0
36	1	57.2	1	1	0	36	3	31.6	3	1	1
61	2	190	2	1	0	42	1	66.2	2	1	0
58	3	128	4	3	1	14	3	138.6	3	3	1
55	3	80	3	4	1	32	1	114	2	3	0
61	1	94.4	2	1	0	35	1	40.2	2	1	0
38	1	76	1	1	0	70	3	177.2	4	3	1
42	1	240	3	2	0	65	2	51.6	4	4	1
50	1	74	1	1	0	45	2	124	2	4	0
58	3	68.6	2	2	0	68	3	127.2	3	3	1
68	3	132.8	4	2	0	31	2	124.8	2	3	0
25	2	94.6	4	3	1	58	1	128	4	3	0
52	1	56	1	1	0	60	3	149.8	4	3	1

Where x_1 refers to the age at which the patient was diagnosed; x_2 refers to renal cell carcinoma vascular endothelial growth factor (VEGF), which is positively high to low in three grades; x_3 refers to the number of micro vessels in the renal cell carcinoma tissue (MVC); x_4 refers to the nuclear histology of kidney cancer cells from low to high divided into 4 grades; x_5 refers to renal cell carcinoma from low to high is divided into 4 Staging; y refers to the metastasis of renal cell carcinoma (1 is metastasis; 0 is no metastasis). The case determines whether renal cell carcinoma is metastasized through five influencing factors, and this paper treats it into a two-classification problem due to the quantification criteria are different between different variables, so the data are min-max standardized firstly, and the parameters of the five influencing factors are $\theta_i, i=1,2,3,4,5, \boldsymbol{\Theta}=(\theta_1,\cdots,\theta_1)^T$, $\boldsymbol{X}=(\boldsymbol{x}_{1j},\cdots \boldsymbol{x}_{5j})^T$. Then

$$t_j(\boldsymbol{\Theta})=\boldsymbol{X}^T\boldsymbol{\Theta}=\theta_1 x_{1j}+\theta_1 x_{2j}+\theta_3 x_{3j}+\theta_4 x_{4j}+\theta_5 x_{5j}, \tag{3}$$

where x_{ij} is the j component of the i influencing factor. Since $t_j(\boldsymbol{\Theta})$ is a linear combination, the problem of using a linear line to do the two classifications obviously affects straightforward. Affecting the classification effect, in order to convert the linearity into nonlinearity, a nonlinear activation function sigmoid function will be introduced here:

$$p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})=\frac{1}{1+e^{t_j(\boldsymbol{\Theta})}}. \tag{4}$$

The sigmoid function is an S-shaped curve that maps the output to 0-1, which means that the output is a probability value. These data is processed by $t_j(\boldsymbol{\Theta})$, then nonlinearly processed by $p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})$ and the obtained output value is close to 0. It is judged that there is no metastasis of renal cell carcinoma; if the output value obtained is

close to 1, it is judged that renal cell carcinoma metastasizes.

In order to obtain higher accuracy, we will train the optimal parameters by minimizing the loss function. Here we introduce the cross entropy loss function, which is often used in neural networks for multi-classification. Its form is as follows:

$$L(\boldsymbol{\Theta}) = -\frac{1}{n}\sum_{j=1}^{n}[y_j ln(p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})) + (1-y_j) ln(1-p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X}))]. \tag{5}$$

In the process of iterative updating of parameters by the algorithm, a negative gradient direction is needed. Then $L(\boldsymbol{\Theta})$ is used to derive the i parameter θ_i as follows:

$$\begin{aligned}\frac{\partial L(\boldsymbol{\Theta})}{\partial \theta_i} &= -\frac{1}{n}\sum_{j=1}^{n}\left(\frac{y_j}{p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})}\frac{\partial p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})}{\partial \theta_i} + \frac{(1-y_j)}{1-p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})}\frac{\partial(-p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X}))}{\partial \theta_i}\right) \\ &= -\frac{1}{n}\sum_{j=1}^{n}\frac{y_j - p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})}{p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})(1-p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X}))}\frac{\partial p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})}{\partial \theta_i} \\ &= -\frac{1}{n}\sum_{j=1}^{n}(y_j - p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X}))x_{ij}.\end{aligned} \tag{6}$$

where y_j is the j component of the y group data, representing the true value of 0 or 1, when $y_j = 1$, the formula (5) The predicted value of $p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})$ is larger, the closer to the true value, the smaller the loss function; when $y_j = 0$, the predicted value is $p_{\boldsymbol{\Theta}}^{(j)}(\boldsymbol{X})$ is smaller, the closer to the real value, the smaller the loss function. So this paper attempts to minimize the cross entropy loss function (5) by using the filled function algorithm. Overcoming the problem of using local optimization algorithms can only get local minimum.

In this paper, the data of 4 - 23 groups are used as the test group. The data of 1-3 and 24-26 groups are used as the verification group. The optimal parameters obtained by the algorithm AFF training are shown in Table 6:

Table 6. Parameter

k	θ_k	θ_k^*	$L(\theta_k^*)$	$\bar{\theta}_k$	$\theta_{\gamma k}^*$
1	(1,3,2,1,2)	(-7.38054835, 4.33273494, -3.64733445, 4.12222301, 2.11783449)	0.25198217	(-8.38054835,4.83273494, -2.64733445, 4.12222301, 2.11783449)	(-7.57159459,4.28758873, 3.47886350, 3.94119798, 2.01755273)
2	(-7.57159459,4.28758873, -3.4788635, 3.94119798, 2.01755273)	(-9.58074336, 5.25537670, -4.29128969, 5.63852196, 2.52719037)	0.23438636	(-12.58074333, 5.25537672, -4.2912901, 5.63852205, 2.52719047)	(-12.67249218,6.72993542, -5.1179182, 7.4901191, 3.18841998)
3	(-12.67249218,6.72993542, -5.1179182, 7.4901191, 3.18841998)	(-13.40880788,7.09635693, -5.19537445, 7.85842628, 3.32961168)	0.22008213	\	\

It can be seen from Table VI that after minimizing $L(\boldsymbol{\Theta})$ by the filled function algorithm, the optimal parameters for the five influencing factors affecting renal cell carcinoma metastasis are

$$\boldsymbol{\Theta} = (-13.40880788, 7.09635693, -5.19537445, 7.85842628, 3.32961168)^T$$

predicting the remaining six sets of data, the probability values are 0.00737699722, 0.00271481598, 0.000119076584, 0.880055784, 0.0541460992, 0.961351203, which representing the predicted group 1, 2, 3, 5 kidney cells no metastasis of cancer is 0, and metastasis of renal cell carcinoma in groups 4 and 6 is 1. Comparing the true values of these six groups 0,0,0,0,0,1, it can be seen that the other five except the fourth groups are correct, the accuracy obtained by the filled function method is 83.33%, so the filled function and its algorithm is considered to be feasible and effective. Since there is less data in the case, if there are more groups of data to be trained, then the accuracy obtained will be higher.

5. Conclusion

In this paper, a new one-parameter filled function is constructed, which overcomes the deficiencies of the previous filled function with exponential term, which decrease the efficiency of filled function algorithm. The feasibility and effectiveness of the method are verified by several examples. The innovation of this paper is that in addition to the experiments with numerical examples, we also tried to combine the filled function algorithm with the actual case, and got good results, which made it further research in application. The disadvantages are that when the algorithm is running, the selection of the initial point and the selection of the parameters need to be adjusted, and these data of the case are not enough, then the accuracy can be better.

Acknowledgments

This work is supported by the Scientific Research Project of Zhejiang Provincial Department of science and technology in China (LGN19C040001, 2017C32034), the Scientific Research Project of Zhejiang Provincial Department of Education in China (Y201636028) and national social science foundation in China (17BTJ028). The authors are highly thankful to the referees for their invaluable suggestions and comments for the improvement of this paper.

References

[1] Ge R.; Qin Y. A class of filled functions for finding global minimizers of a function of several variables. *Journal of Optimization Theory and Applications*, **1987**, (54): pp. 241-252.

[2] Yang Y.; Shang Y. A new filled function method for unconstrained global optimization. *Applied Mathematics and Computation*, **2006**, (173): pp. 501-512.
[3] El-Gindy T. M.; Salim M. S. Ahmed A I. A new filled function method applied to unconstrained global optimization. *Applied Mathematics and computation*, **2016**, (273): pp. 1246-1256.
[4] Wang C.; Yang Y.; Li J. A new filled function method for unconstrained global optimization. *Journal of Computational and Applied Mathematics*, **2009**, (225): pp. 68-79.
[5] Gao C.; Yang Y.; Han B. A new class of filled functions with one parameter for global optimization. *Computers & Mathematics with Applications*, **2011**, (62): pp. 2393-2403.
[6] Wei F.; Wang Y.; Lin H. A new filled function method with two parameters for global optimization. *Journal of Optimization Theory and Applications*, **2014**, (163): pp. 510-527.
[7] Zhang Y.; Zhang L.; Xu Y. A one-parameter filled function for nonsmooth global optimization and its application. *Journal of Systems Science and Complexity*, **2010**(23): pp. 1195-1209.
[8] Yang Y.; Zhang L. Several deterministic algorithms for global optimization. Shanghai University Press, **2010**.
[9] Shang Y.; Pu D., Jiang A. Finding global minimizer with one-parameter filled function on unconstrained global optimization. *Applied Mathematics and Computation*, **2007**, (191): pp. 176-182.
[10] Li B.; Lu D. A new filled function algorithm with one parameter. *Journal of Engineering Mathematics*, **2015**, (32): pp. 269-275.
[11] Wu B.; Gao Y. A filled function method with one parameter for finding unconstrained continuous global optimization problems. *Journal of Ningxia University* (Natural Science Edition), **2017**, pp. 221-223.
[12] He S.; Chen W. A filled function algorithm for solving unconstrained optimization problems. *Journal of Zhejiang University*: Science Edition, **2011**, (38): pp. 144-149.
[13] Ni Z. Health statistics. People's Medical Publishing House, **2001**, pp. 138-142.