A New Integral Function Algorithm for Global Optimization and Its Application to the Data Clustering Problem

Ridwan Pandiya[1,✉], **Atina Ahdika**[2], **Siti Khomsah**[3], **Rima Dias Ramadhani**[3]

[1]Department of Informatics, Institut Teknologi Telkom Purwokerto, Indonesia
[2]Department of Statistics, Universitas Islam Indonesia, Indonesia
[3]Department of Data Sciences, Institut Teknologi Telkom Purwokerto, Indonesia

ridwanpandiya@ittelkom-pwt.ac.id[✉], atina.a@uii.ac.id, siti@ittelkom-pwt.ac.id, rima@ittelkom-pwt.ac.id

Abstract

The filled function method is an approach to finding global minimum points of multidimensional unconstrained global optimization problems. The conventional parametric filled functions have computational weaknesses when they are employed in some benchmark optimization functions. This paper proposes a new integral function algorithm based on the auxiliary function approach. The proposed method can successfully be used to find the global minimum point of a function of several variables. Some testing global optimization problems have been used to show the ability of this recommended method. The integral function algorithm is then implemented to solve the center-based data clustering problem. The results show that the proposed algorithm can solve the problem successfully.

Keywords: *Global Optimization, Global Minimum Point, Integral Function Algorithm, Data Clustering.*

Received: 23 August 2023
Accepted: 2 October 2023
Online: 2 October 2023
Published: 20 December 2023

1 Introduction

Many problems in real life require optimal solutions. It encourages scholars working in numerical optimization to create the most efficient method. Bio-inspired algorithms and surrogate-assisted methods (see [20, 14, 10, 18, 12, 27]) can solve the black-box optimization problem. Flexibility is the main characteristic of these methods. However, the convergence is based on probability. Therefore, those algorithms are expensive in CPU time. Deterministic approaches promise guaranteed convergence, although the problems solved are usually less extensive than those that stochastic approaches can solve. The filled function method [7, 6, 8] and the DIRECT method [11, 31] are examples of methods that fall into the deterministic approach. This research is focused on the filled function method.

There has been a recent growth in parametric-filled function methods in global optimization problems. This is due to the development of nonlinear function optimization problems, which have many local minimum points in the search domains. These problems are widespread in applications such as engineering, finance, social science, biology, etc. The methods such as Newton's steepest descent, quasi-Newton, trust region, and conjugate gradient methods are well executed for obtaining a local solution and globally determined when solving convex optimization problems [25]. Nevertheless, if the problems contain a non-convexity structure, those methods fail to find a global solution.

Consequently, the global optimization problem is a challenging computational job for researchers [32]. The parametric filled function method is an auxiliary function method that uses a point in an objective function, usually a minimum point, to construct a function that satisfies the definition of the filled function. The parametric filled function, which was first presented by Ge [6, 7], gives us an excellent idea of using the local minimum point to find the global minimum point. Furthermore, the basic outline of the function is explained as follows.

a. Minimize the objective function by employing a local minimization method; a minimum point of the objective function is attained.

b. The filled function is constructed, and the point close to the present minimum point is used as an initial point to further minimize the filled function.

c. The minimum point of the filled function obtained in b is used as an initial point for minimizing the objective function, and the better minimum point of the objective function will be obtained.

d. The global minimum point will be found if phases b and c are repeated until the stopping criterion is satisfied.

The filled function proposed in [7] has some unexpected attributes. The intended method strongly depends on two parameters that are difficult to adjust to satisfy several conditions. Because of this problem, Ge and Qin in [8] proposed filled function that has only one parameter and is twice continuously differentiable everywhere. However, the filled function proposed in [8]

and its derivative are still influenced by the exponential term, which will be very large when the iterative points are far from the current local minimum point.

Subsequently, several two-parameter and one-parameter filled functions have been proposed, as presented in [4, 19, 3, 21, 23, 22], to solve global optimization problems. However, many filled function methods proposed by various authors in the literature have some drawbacks since their filled functions still include exponential terms, logarithmic terms, weight factors, and norms. In addition, selecting an appropriate method for searching minimum points, particularly for filled functions, involves numerous variables because they must contain the Hessian matrix and the inverse of a matrix. Several methods do not include inverse or Hessian matrix, but it is still believed that the gradient-based methods are more effective in determining the minimum point of the function.

In 2010, [26] introduced the parameter-free filled function. However, the function is discontinuous at some points. Besides, the filled function proposed by [26] has the *arctan* term that makes this function very flat if the points are far from the incumbent local minimum point. This nature is the same as Ge's filled function [7]. Several continuously differentiable filled functions were proposed in [24, 28, 1]. The local minimum point of those filled functions does, however, occasionally exist. According to this fact, this study introduces an integral function algorithm. Our new integral function method is made to eliminate the limitation of previous methods because our study has successfully developed such an auxiliary function that can solve multidimensional optimization problems. The advantage of our filled function lies in the non-involvement of parameters, exponential, logarithmic, and weight factors.

This paper is organized as follows: In section 2, our new integral function, including its definitions and theorems, is presented. Section 3 describes the algorithm that includes an integral function in its phase. Numerical results are presented in Section 4. Application of the integral function to the data clustering problem is provided in Section 5, and the conclusion will be presented in Section 6.

2 A New Integral Function and its Properties

This study solves the following unconstrained global optimization problems.

$$\begin{cases} \min \omega(t) \\ t \in R^n \end{cases}. \tag{1}$$

Throughout this paper, we require the following assumptions.

Assumption 1: $\omega(t)$ is continuously differentiable.

Assumption 2: $\omega(t) \to \infty$ as $\|t\| \to \infty$.

Assumption 3: The set $\mathcal{G}$, defined by

$$\mathcal{G} = \{\omega(t) \,|\, t \in \mathcal{F}\},$$

is finite, with $\mathcal{F}$ representing the set of all local minimum points of Problem (1)

Since the number of local minimum points can be infinite, Assumption 3 requires that the number of local minimum values of Problem (1) is finite. Assumption 2 implies the existence of a compact set $\xi \in R^n$ whose interior contains all minimum points of $\omega(t)$. We assume that the value of $\omega(t)$ for t on the boundary of ξ is greater than the value of $\omega(t)$ for any t in the interior of ξ. Then, Problem (1) is equivalent to Problem (2).

$$\begin{cases} \min \omega(t) \\ t \in \xi \end{cases} \tag{2}$$

In order to develop an auxiliary function approach that can address the deficiencies experienced by the filled function method, we offer the following integral function.

$$\psi_i(t, t_m^*) = \begin{cases} -\int_{t_{m(i)}^*}^{t_i} (\omega(t) - \omega(t_m^*))\, dt_i & \left(t_i \geq t_{m(i)}^*\right) \\ -\int_{t_i}^{t_{m(i)}^*} (\omega(t) - \omega(t_m^*))\, dt_i & \left(t_i < t_{m(i)}^*\right) \end{cases}, \tag{3}$$

where

$$t_m^* = \left(t_{m(1)}^*, \ldots, t_{m(n)}^*\right)$$

is a local minimum point of $\omega(t)$, $i = 1, \ldots, n$, with n is the number of variables of $\omega(t)$, and

$$\begin{aligned} \omega(t) - \omega(t_m^*) &= \omega\left(t_{m(1)}^*, \ldots, t_{m(i)}, \ldots, t_{m(n)}^*\right) \\ &\quad -\omega\left(t_{m(1)}^*, \ldots, t_{m(n)}^*\right), \end{aligned} \tag{4}$$

for $m = 1, \ldots, q$, with q is the number of local minimum point of $\omega(t)$. Based on the form shown in Eq. (3), it is clear that $\psi_i(t, t_m^*)$ is a one-dimensional function.

It is known that all filled functions given in the literature still need to select the directions e_i $(i = 1, \ldots, m_0)$ and $m_0 \geq 2n$. In contrast, we do not need to select the direction in this paper since every function in Eq.(3) has included the directions. In other words, each direction has a different function. Based on Eq.(3), $\psi_1(t, t_m^*)$ and $\psi_2(t, t_m^*)$ are integral functions for searching the global minimum point of $\omega(t)$ in t_1-axis (right side and left side of t_i respectively), while $\psi_3(t, t_m^*)$ and $\psi_4(t, t_m^*)$ are integral functions for searching the global minimum point of $\omega(t)$ in t_2-axis, and so forth. Therefore, we will have $2n$ of integral function where each function indicates the directions.

Theorems 1 - 5 show that $\psi_i(t, t_m^*)$ satisfies the specification to be qualified as a filled function, or in other

words, Eq.(3) satisfies three conditions demanded by the Definition 1.

Definition 1. [33] A function $\psi(t, t_m^*)$ is categorized as a filled function constructed at t_m^* if $\psi(t, t_m^*)$ meets the following three conditions.

- t_m^* is a strict local maximum point of $\psi(t, t_m^*)$;
- $\psi(t, t_m^*)$ does not own local minimum and saddle points in $\hat{\xi}$;
- If $\bar{\xi} \neq \emptyset$, then $\bar{\xi}$ contains local minimum point of $\psi(t, t_m^*)$,

where

$$\hat{\xi} = \{t \in \xi | \omega(t) > \omega(t_m^*)\} \setminus \{t_m^*\}$$

and

$$\bar{\xi} = \{t \in \xi | \omega(t) \leq \omega(t_m^*)\}.$$

For the simplicity, integral function (3) can be written as

$$\psi(t, t_m^*) = -\int_{t_m^*}^{t} \omega(t) - \omega(t_m^*)\, dt,$$

because either

$$-\int_{t_{m(i)}^*}^{t_i} \omega(t) - \omega(t_m^*)\, dt_i \quad \left(t_i \geq t_{m(i)}^*\right)$$

or

$$-\int_{t_i}^{t_{m(i)}^*} \omega(t) - \omega(t_m^*)\, dt_i \quad \left(t_i < t_{m(i)}^*\right)$$

has the same behavior.

Theorem 1. *If (1) $t_m^* \in \mathcal{F}$, (2) t_m is a point such that $t_m \neq t_m^*$, where $\omega(t_m) \geq \omega(t_m^*)$. Then,*

$$\psi(t, t_m^*) \leq 0 = \psi(t_m^*, t_m^*).$$

Proof. From Eq. (3), we have

$$\psi(t_m, t_m^*) = -\int_{t_m^*}^{t_m} \omega(t_m) - \omega(t_m^*)\, dt_m.$$

It follows from the assumption of the theorem,

$$\omega(t_m) \geq \omega(t_m^*).$$

This implies

$$\omega(t_m) - \omega(t_m^*) \geq 0.$$

Since $t_m \neq t_m^*$, we have

$$\int_{t_m^*}^{t_m} \omega(t_m) - \omega(t_m^*)\, dt_m > 0.$$

Therefore,

$$\psi(t_m, t_m^*) < 0 = \psi(t_m^*, t_m^*).$$

□

Theorem 2. *If $t_m^* \in \mathcal{F}$, then t_m^* is a strict local maximum point of $\psi(t, t_m^*)$.*

Proof. t_m^* is the element of $\mathcal{F}$. This implies there exists $\mathcal{N}(t_m^*, \zeta)$, which is the neighborhood of $\mathcal{F}$, such that

$$\omega(t) \geq \omega(t_m^*),$$

$\forall t \in \mathcal{N}(t_m^*, \zeta)$. It follows from Theorem 1, for all

$$t \in \mathcal{N}(t_m^*, \zeta),$$

$t \neq t_m^*$,

$$\psi(t, t_m^*) < 0 = \psi(t_m^*, t_m^*).$$

Thus, t_m^* is a strict local maximum point of $\psi(t, t_m^*)$. □

Theorem 3. *If (1) $t_m^* \in \mathcal{F}$, (2) t_1 and t_2 are two points such that $\|t_1 - t_m^*\| < \|t_2 - t_m^*\|$ and*

$$\omega(t_m^*) < \omega(t_1) < \omega(t_2).$$

Then,

$$\psi(t_2, t_m^*) < \psi(t_1, t_m^*) < 0 = \psi(t_m^*, t_m^*).$$

Proof. Consider

$$\begin{aligned}\psi(t_2, t_m^*) - \psi(t_1, t_m^*) &= -\int_{t_m^*}^{t_2} (\omega(t_2) - \omega(t_m^*))\, dt_2 \\ &\quad + \int_{t_m^*}^{t_1} (\omega(t_1) - \omega(t_m^*))\, dt_1.\end{aligned}$$

Since

$$\omega(t_m^*) < \omega(t_1) < \omega(t_2),$$

then

$$\omega(t_2) - \omega(t_m^*) > \omega(t_1) - \omega(t_m^*) > 0.$$

Therefore,

$$\int_{t_m^*}^{t_1} (\omega(t_1) - \omega(t_m^*))\, dt_1 - \int_{t_m^*}^{t_2} (\omega(t_2) - \omega(t_m^*))\, dt_2 < 0.$$

Consequently, $\psi(t_2, t_m^*) < \psi(t_1, t_m^*) < 0 = \psi(t_m^*, t_m^*)$. □

Theorem 4. *If (1) $t_m^* \in \mathcal{F}$, (2) t is a point such that $\omega(t) > \omega(t_m^*)$. Then t is not a stationary point of $\psi(t, t_m^*)$.*

Proof. For a given t_m with $\omega(t_m) > \omega(t_m^*)$, we have

$$\nabla\psi(t_m, t_m^*) = D_{t_m}\left(-\int_{t_m^*}^{t_m} (\omega(t_m) - \omega(t_m^*))\, dt_m\right).$$

From the fundamental theorem of Calculus,

$$\nabla\psi(t_m, t_m^*) = -(\omega(t_m) - \omega(t_m^*)).$$

Since $\omega(t_m) > \omega(t_m^*)$, then $\nabla\psi(t_m, t_m^*) \neq 0$ □

Theorem 5. *If $t_m^* \in \mathcal{F}$, then any local minimum point or saddle point of $\psi(t, t_m^*)$ must belong to the set*

$$\bar{\xi} = \{t \in R^n | \omega(t) \leq \omega(t_m^*)\}.$$

Proof. Assume that Theorem 5 is not true. Then, there is a local minimum point or saddle point of $\psi(t, t_m^*)$, denoted by $\tilde{t}$, such that $\tilde{t} \notin \bar{\xi}$ and $\omega(\tilde{t}) > \omega(t_m^*)$. It follows Theorem 2 t_m^* is a strict local maximum point of $\psi(t, t_m^*)$. Since $\tilde{t}$ is a local minimum or saddle point of $\psi(t, t_m^*)$, thus $t_m^* \neq \tilde{t}$. If $\tilde{t}$ is a local minimum point of $\psi(t, t_m^*)$, it contradicts Theorem 3, while if $\tilde{t}$ is a saddle point of $\psi(t, t_m^*)$, it contradicts Theorem 4. As a result, Theorem 5 is true. □

Theorem 6. *If t_m^* is a global minimum point of $\omega(t)$, then $\psi(t, t_m^*) < 0, \forall t \in \xi$.*

Proof. Since t_m^* is a global minimum point of $\omega(t)$, $\omega(t) \geq \omega(t_m^*)$ for all $t \in \xi$. Thus, it follows from Theorem 1, $\psi(t, t_m^*) < 0$ for all $t \in \xi$. □

3 The Algorithm

Derived from the theoretical properties of $\psi_i(t, t_m^*)$, in this section, we describe our integral function algorithm, which applies to an unconstrained global optimization problem. The general iterative of this integral function is as follows: assume that t_1^* is not a global minimum point of $\omega(t)$, then we can create an integral function $\psi(t, t_m^*)$ at t_1^* to acquire such a point that can be applied as a new initial point in order to bring the minimization of $\omega(t)$ to the better minimum point than t_1^*. The process is conducted repeatedly until the stopping criterion is met.

Input: t_0, the initial point
Output: t_I^*, global minimum point of $\omega(t)$
Initialization:
a. Choose a constant υ, e.g., set $\upsilon := 0.1$
b. Set $i = 1, \ldots, i_0$, where $i_0 = 2n$
c. Set $i := 1$.
Main step
1. Minimize $\omega(t)$ by using t_0 to obtain t_I^*.
2. Built Eq.(3)
3. **if** $i > i_0$
   then go to 7
   else go to the next step
4. Find $\bar{t}_i$, the stationary point of $\psi_i(t, t_I^*)$
5. Minimize $\omega(t)$ by using $\bar{t}_i$ to obtain t_{II}^*
6. **if** $\omega(t_{II}^*) < \omega(t_I^*)$ and $t_{II}^* \in \xi$
   then set $i := 1$, $t_I^* = t_{II}^*$, and go to 2
   else set $i := i + 1$ and go to 3
7. Algorithm stops and t_I^* is taken as a global minimum.

There are two phases in the algorithm. First, minimizing the objective function $\omega(t)$ using a local search procedure (we implement the steepest descent method). In this phase, t_I^* is obtained. In the second phase, integral function $\psi_i(t, t_I^*)$ is constructed. Integral function $\psi_i(t, t_I^*)$ is minimized, then we will yield $\bar{t}_i$, and phase 2 ends. The algorithm will reenter phase 1, with $\bar{t}_i$ as the starting point to find t_{II}^*, a new minimum point of $\omega(t)$ if such one exists. The iteration process is repeated until a stopping criterion is met. The last local minimum point to be found is assumed to be the global minimum point of $\omega(t)$. The explanation of our integral function in the algorithm is given in the following paragraphs.

A set of $2n$ integral functions is chosen in step 2, where each integral function represents the direction. For example, assume that we solve the two-variable function $\omega(t)$, with $t_I^* = (t_1^*, t_2^*)$ as the first local minimum point of $\omega(t)$, which is gained in step 1. Thus, we have the following four integral functions.

$$\psi_1(t, t_I^*) = -\int_{t_1^*}^{t_1} (\omega(t_1, t_2^*) - \omega(t_1^*, t_2^*))dt_1,$$

for $(t_1 \geq t_1^*, t_2 = t_2^*)$.

$$\psi_2(t, t_I^*) = -\int_{t_1}^{t_1^*} (\omega(t_1, t_2^*) - \omega(t_1^*, t_2^*))dt_1,$$

for $(t_1 < t_1^*, t_2 = t_2^*)$.

$$\psi_3(t, t_I^*) = -\int_{t_2^*}^{t_2} (\omega(t_1^*, t_2) - \omega(t_1^*, t_2^*))dt_2,$$

for $(t_2 \geq t_2^*, t_1 = t_1^*)$.

$$\psi_4(t, t_I^*) = -\int_{t_2}^{t_2^*} (\omega(t_1^*, t_2) - \omega(t_1^*, t_2^*))dt_2,$$

for $(t_2 < t_2^*, t_1 = t_1^*)$.

It can be seen clearly that ψ_1 - ψ_4 are one-variable functions. ψ_1 is one-variable function for which t_2 is fixed for the positive t_1-axis direction, ψ_2 is one-variable function for which t_2 is fixed, for the negative t_1-axis direction, ψ_3 is one-variable function which t_1 is fixed for the positive t_2-axis direction, and ψ_4 is one-variable function for which t_1 is fixed for the negative t_2-axis direction.

Each stationary point of $\psi(t, t_m^*)$ will be paired with the number that has been fixed. For example, if we obtain the local minimum point or inflection point $\bar{t}_1$ of ψ_1 or ψ_2, then $(\bar{t}_1, t_2^*)$ will become the initial point for minimizing $\omega(t)$ in step 5. Otherwise, if the stationary point $\bar{t}_2$ is obtained from ψ_3 or ψ_4, $(t_1^*, \bar{t}_2)$ will be taken as the initial point.

4 Numerical Results

In this section, we apply our integral function algorithm to some standard testing problems derived from the literature [26]. The proposed algorithm, and the algorithm given in [26] for comparison purposes, is programmed by Microsoft Visual C++ 6.0 for working on the Windows 10 system with Intel(R) Core(TM) i3 Processor 2.3G CPU and 4G RAM. Numerical results illustrate that the method is competent. The following notations are used in Tables 1 - 3

1. n is the number of variables of a testing functions,
2. F_t is the total number of function evaluations of $\omega(t)$ and $\psi_i(t, t_m^*)$ before termination by our algorithm,
3. G_t is the total number of gradient evaluations of $\omega(t)$ before termination by our algorithm,
4. F_{tl} is the total number of function evaluations of $\omega(t)$ and $\psi_i(t, t_m^*)$ before termination by the algorithm proposed by [26],
5. G_{tl} is the total number of gradient evaluations of $\omega(t)$ before termination by the algorithm proposed by [26]
6. υ is a small positive number used as an addition to the minimum point when it will be used to minimize the integral function.

Table 1: General computational results by our algorithm

Problem	t_0	υ	n	F_t	F_{tl}
1	(-1.6,0.9)	0.1	2	90	40
2	(-2 , -1)	0.1	2	304	145
3	(2 , 2)	0.1	2	159	74
4	(-1 , -1)	0.1	2	707	349
5	(10 , 10)	0.1	2	425	137
6	(1 , -1)	0.1	2	797	398
7	(1 , 1)	0.1	2	221	94
8	(3 , 3)	0.1	2	788	147
	(0 , 0)	0.1	2	1432	706
	(10 , -10)	0.1	2	2111	1045
9	(2 , 2 , 2)	0.1	3	4447	2208
	(2 , ..., 2)	0.1	5	5087	2525
	(2 , ..., 2)	0.1	7	4354	2156
	(2 , ..., 2)	0.1	10	4692	2316

Table 1 provides the general computational result of the integral function algorithm (Algorithm 3). It indicates that Algorithm 3 successfully obtained the global minimum point of the given examples. Table 2 compares Algorithm 3 and the algorithm offered in [26]. We use the same initial point for our algorithm and algorithm in [26]. In the framework that the comparison is fair enough, we use the same method in minimizing the objective function and integral function (filled function for the method proposed in [26]). We found that in minimizing the filled function in [26], we could not use its gradient for the search direction. Therefore, another search direction stated in [26] was used.

Table 2: Comparison of the results between our algorithm and the algorithm proposed in [26]

Problem	t_0	υ	n	F_t	F_{tl}	F_{tl}	G_{tl}
1	(-1.6,0.9)	0.1	2	90	40	1013	110
2	(-2 , -1)	0.1	2	304	145	1525	168
3	(2 , 2)	0.1	2	159	74	1173	91
4	(-1 , -1)	0.1	2	707	349	1425	775
5	(10 , 10)	0.1	2	425	137	1363	186
6	(1 , -1)	0.1	2	797	398	1775	95
7	(1 , 1)	0.1	2	221	94	1972	231
8	(3 , 3)	0.1	2	788	147	2143	279
	(0 , 0)	0.1	2	1432	706	1723	327
	(10 , -10)	0.1	2	2111	1045	3439	127
9	(2 , 2 , 2)	0.1	3	4447	2208	2467	227
	(2 , ..., 2)	0.1	5	5087	2525	4497	1053
	(2 , ..., 2)	0.1	7	4354	2156	9735	1683
	(2 , ..., 2)	0.1	10	4692	2316	13091	2167

It can be seen in Table 3 that our integral function is more effective. All function evaluations are less than those yielded by the algorithm proposed by [26]. The exception is for the n-dimensional function ($n = 3$). Function evaluation obtained by Algorithm 3 is 4447, while function evaluation gained by the algorithm in [26] is 2467. So as for the aspect gradient evaluation, Algorithm 3 performs better than that algorithm in [26]. Except for gradient evaluations, the Rastrigin and two dimensional functions ($c = 0.5$) have more gradient evaluations. However, it can be understood since our integral function can obtain the entire extreme points of a function. However, in this paper, we are only concerned with obtaining a global minimum point. However, this kind of nature of our integral function has affected function evaluation, especially for functions with many extreme points.

5 Application to Data Clustering Problem

Grouping data sets into some clusters is one of the unsupervised pattern classifications. It has been broadly implemented in various real problems such as pattern recognition, machine learning, text classification, and other problems (see [2, 30]). Some approaches can be employed in clustering data. One of the most popular methods is the K-means algorithm. The algorithm obtains the minimum distance of the data into the center for each partition. The K-means algorithm originally used the Euclidean distance to operate its algorithm. However, according to [29], that distance sometimes is unsuitable. Since clustering is the minimization problem, K-means challenges non-convex and nonlinear optimization. This fact brings many scholars to solve the clustering problem through the optimization point

of view. In the optimization approaches, the data is modeled into the objective function and then solved by implementing minimization procedures. Therefore, the appropriate method chosen is one of the key successes in obtaining the optimal cluster. Since the integral function algorithm in this paper is a global minimization algorithm, in this section, we are willing to examine the proposed algorithm's reliability to solve the clustering problem.

The purpose of this section is the implementation of Algorithm 3. For this purpose, we use the optimization model initiated by [29], which has the following form.

$$\min_{\tau_1,\ldots,\tau_k \in R} \left(-\varepsilon \sum_{p=1}^{m} \ln \sum_{q=1}^{k} \exp\left(-\frac{\|\tau_q - t_p\|^2}{\varepsilon} \right) \right), \quad (5)$$

where $\tau_1, \ldots, \tau_k \in R$ denotes the centers for each k disjoint partitions of the data set

$$A = \{t_p \in R^n | p = 1, \ldots, m\} \subset R^n.$$

Eq.(5) will be solved by Algorithm 3. We use the example in [13].

Example 1. Let Table 3 is a data set where the elements denoted by t_p, $p = 1, \ldots, 20$. The data will be partitioned into two clusters. In the first phase, all t_p are substituted to Eq.(5). The next step is solving the optimization model using Algorithm 3 to obtain the optimal center for each cluster.

Table 3: Data set

t_1	t_2	t_3	t_4	t_5
0.456535	0.658425	0.868230	0.086283	0.704274
t_6	t_7	t_8	t_9	t_{10}
0.063848	0.795001	0.684515	0.040520	0.824774
t_{11}	t_{12}	t_{13}	t_{14}	t_{15}
0.957827	0.486618	0.008372	0.081007	0.251257
t_{16}	t_{17}	t_{18}	t_{19}	t_{20}
0.907372	0.014313	0.783009	0.743946	0.066294

For the data set given in Table 3, Algorithm 3 success in obtaining centers

$$(\tau_1^*, \tau_2^*) = (0.0765, 0.7392),$$

with the number of iterations and function evaluations are 40 and 1829, respectively. From the results, it can be shown that Algorithm 3 can be used as one of the alternatives when solving center-based data clustering problems.

6 Conclusion and Future work

The more problems in the real world can be modeled into the optimization model, the higher the demand for minimization methods with efficient computational performance. In this regard, many methods have been offered with various advantages. This paper offers a new deterministic approach to solve global optimization problems: the integral function algorithm. Uniquely, even though the optimization problems solved are multi-dimensional, the approach used in this study is a one-variable function approach. With this property, the computation stage becomes simpler. In addition, the integral function initiated in this study does not involve parameters usually included by similar methods. The reliability and efficiency of the proposed method are demonstrated through numerical simulations and numerical comparisons with other similar methods available in the literature. The proposed algorithm is also implemented in the final phase to solve the center-based data clustering problem. The implementation results show that the proposed algorithm can obtain the optimal center of the given data set.

The numerical simulations conducted in this research are limited to several benchmark objective functions. In order to confirm that the proposed method is more reliable in solving global optimization problems, further research will be carried out on non-trivial optimization problems, such as black-box optimization problem [9, 15], GKLS generator [5, 16], and zigzag-based benchmark functions [13, 17].

Acknowledgment

This research is supported by the Ministry of Education, Culture, Research and Technology, Indonesia, through the fundamental research grant 2023.

References

[1] Ahmed, A. A new filled function for global minimization and system of nonlinear equations. *Optimization 71* (2022), 4083–4106.

[2] Dhillon, I., Mallela, S., and Kumar, R. A divisive information-theoretic feature clustering algorithm for text classification. *Journal of Machine Learning Research 3* (2003), 1265–1287.

[3] El-Gindy, T., Salim, M., and Ahmed, A. A new filled function method applied to unconstrained global optimization. *Applied Mathematics and Computation 273* (2016), 1246–1256.

[4] Gao, Y., Yang, Y., and You, M. A new filled function method for global optimization. *Applied Mathematics and Computation 268* (2015), 268–695.

[5] Gaviano, M., Kvasov, D. E., Lera, D., and Sergeyev, Y. D. Algorithm 829: Software for generation of classes of test functions with known local and global minima for global optimization. *ACM Transactions on Mathematical Software 29* (2003), 469–480.

[6] Ge, R. The theory of filled function method for finding global minimizers of nonlinearly constrained minimization problems. *Journal of Computational Mathematics 5* (1987), 1–9.

[7] Ge, R. A filled function method for finding a global minimizer of a function of several variables. *Mathematical Programming 46* (1990), 191–204.

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

[9] Hansen, N., Auger, A., Finck, S., and Ros, R. Real parameter black-box optimization benchmarking 2012: Experimental setup. Technical Report. INRIA 6829, 2012.

[10] Jin, Y., Wang, H., Chugh, T., Guo, D., and Miettinen, K. K data-driven evolutionary optimization: An overview and case studies. *IEEE Transactions on Evolutionary Computation 23* (2018), 442–458.

[11] Jones, D. R., and Martins, J. R. The direct algorithm: 25 years later. *Journal of Global Optimization 79* (2021), 521–566.

[12] Kanagaraj, G., Masthan, S. S., and Yu, V. F. Meta-heuristics based inverse kinematics of robot manipulator's path tracking capability under joint limits. *Mendel 28* (2022), 41–54.

[13] Kudela, J. Novel zigzag-based benchmark functions for bound constrained single objective optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC)* (2021), pp. 857–862.

[14] Kudela, J. A critical problem in benchmarking and analysis of evolutionary computation methods. *Nature Machine Intelligence 4* (2022), 1238–1245.

[15] Kudela, J. Benchmarking state-of-the-art direct-type methods on the bbob noiseless testbed. In *GECCO '23 Companion: Proceedings of the Companion Conference on Genetic and Evolutionary Computation* (2023), pp. 1620–1627.

[16] Kudela, J., and Juricek, M. Computational and exploratory landscape analysis of the gkls generator. In *GECCO '23 Companion: Proceedings of the Companion Conference on Genetic and Evolutionary Computation* (2023), pp. 443–446.

[17] Kudela, J., and Matousek, R. New benchmark functions for single-objective optimization based on a zigzag pattern. *IEEE Access 10* (2022), 8262–8278.

[18] Kudela, J., and Matousek, R. Recent advances and applications of surrogate models for finite element method computations: A review. *Soft Computing 26* (2022), 13709–13733.

[19] Lampariello, F., and Liuzzi, G. A filling function method for unconstrained global optimization. *Computational Optimization and Applications 61* (2015), 713–729.

[20] LaTorre, A., Molina, D., Osaba, E., Poyatos, J., Ser, J. D., and Herrera, F. A prescription of methodological guidelines for comparing bio-inspired optimization algorithms. *Swarm and Evolutionary Computation 67* (2021), 100973.

[21] Li, J., Shang, Y., and Han, P. New tunnel-filled function method for discrete global optimization. *Journal of the Operations Research Society of China 5* (2017), 291–300.

[22] Lin, H., Gao, Y., Wang, X., and Su, S. A filled function which has the same local minimizer of the objective function. *Optimization Letters 13* (2018), 761–776.

[23] Lin, H., Wang, Y., Gao, Y., and Wang, X. A filled function method for global optimization with inequality constraints. *Computational and Applied Mathematics 37* (2018), 1524–1536.

[24] Liu, H., Wang, Y., Guan, S., and Liu, X. A new filled function method for unconstrained global optimization. *International Journal of Computer Mathematics 94* (2017), 2283–2296.

[25] Ma, S., and Gao, Y. Parameter-free filled function method for nonlinear integer program. *Journal of Industrial & Management Optimization 19* (2023), 6835–6855.

[26] Ma, S., Yang, Y., and Liu, H. A parameter free filled function for unconstrained global optimization. *Applied Mathematics and Computation 215* (2010), 3610–3619.

[27] Muller, J. Improving initial aerofoil geometry using aerofoil particle swarm optimisation. *Mendel 28* (2022), 63–67.

[28] Pandiya, R., Salmah, Widodo, W., and Endrayanto, I. A novel parameter-free filled function applied to global optimization. *Engineering Letters 29* (2021), 191–200.

[29] Sabo, K., Scitovski, R., and Vazler, I. One-dimensional center-based l_1-clustering method. *Optimization Letters 7* (2013), 5–22.

[30] Sabo, K., Scitovski, R., Vazler, I., and Zekic-Susac, M. Mathematical models of natural gas consumption. *Energy Conversion and Management 52* (2011), 1721–1727.

[31] Stripinis, L., and Paulavičius, R. Lipschitz-inspired halrect algorithm for derivative-free global optimization. *Journal of Global Optimization* (2023).

[32] Tang, J., Wang, W., and Xu, Y. Lower-order smoothed objective penalty functions based on filling properties for constrained optimization problems. *Optimization 71* (2022), 1601, 2022.

[33] Yang, Y., and Shang, Y. A new filled function method for unconstrained global optimization. *Applied Mathematics and Computation 173* (2006), 501–512.