A Genetic Algorithm for Pointwise Source Reconstruction by the Method of Fundamental Solutions

J. ROCHA DE FARIA

Received on May 14, 2020 / Accepted on January 30, 2022

ABSTRACT. Inverse source reconstruction problems offer great potential for applications of interest to engineering, such as the identification of polluting sources, and to medicine, such as electroencephalography, to cite at least two relevant examples. From a mathematical point of view, the identification of a concentrated source (intensity and location) corresponds to the identification of the centroid (location) and size (intensity) of a distributed source. On the other hand, from a numerical point of view, it is observed that the use of domain discretization methods is intrinsically associated with the introduction of numerical noise in reconstruction algorithms, which is strongly inadvisable since inverse problems are reckoned to be ill-posed. The objective of this work is to explore, in the context of a Poisson problem and taking into account a numerical point of view, a new reconstruction algorithm based on the method of fundamental solutions, where a source point adequately represents the pointwise source within the domain. The inverse problem is reformulated as an optimization problem solved through a genetic algorithm. Finally, numerical examples are performed to analyze the accuracy of the proposed algorithm for two and three dimensions.

Keywords: inverse problems, method of fundamental solutions, genetic algorithms, source reconstruction.

1 INTRODUCTION AND RELATED WORK

Inverse source identification problems have several applications of interest to engineering, such as diffusion or groundwater flow processes, convection-diffusion processes, acoustic problems, indoor and outdoor air pollution, detecting and monitoring underground water pollution, for example [4, 14]. Several works of literature start from an initial guess for the reconstruction of a source represented by a characteristic function.

Motivated by the possible real applications, the objective of the present work is to extend the algorithm proposed in the related conference paper [22]. More specifically, we apply the methodology for choosing the initial guess (centroid and size) in source reconstruction algorithms in two- and three-dimensional cases. In particular, the method of fundamental solutions (MFS) for direct problem solving is adopted, and a genetic algorithm (GA) is used to minimize the elected cost function to approach the inverse problem as an optimization problem.

Universidade Federal da Paraíba, Department of Computer Systems, Av. dos Escoteiros, s/n, Mangabeira, 58058-600, João Pessoa, PB, Brazil – E-mail: jairo@ci.ufpb.br https://orcid.org/0000-0001-6954-4071
The main contribution of this work is in the modeling of the pointwise source through a source point of the method of fundamental solutions positioned within the domain. As far as we know, this procedure is original, considering that the essence behind the method of fundamental solutions (MFS) [18] and other analogous methods, such as auxiliary sources (MAS) [16, 19], of fictitious fonts (MFS) [27, 31], and the source model technique (SMT) [11, 28] concern representation of the approximate solution of a boundary value problem as a finite superposition of fictitious source fields located outside the problem domain [17, 29].

In related work, Machado et al. [20] propose a non-iterative second-order reconstruction algorithm, which performs exhaustive research about \( n \) points to reconstruct intensities and locations for \( m \) pointwise sources. Because of the combinatorial nature of the problem, this exhaustive search becomes unfeasible for \( n \) much greater than \( m \), as \( m \) increases. Then, the authors discuss two other algorithms. Firstly, considering a “Multi-grid approach” and, secondly, considering a “Metaheuristic approach”. Although the exhaustive research and the multi-grid approach showed good results, they were able to reconstruct up to \( m = 4 \) inclusions. On the other hand, the metaheuristic procedure did not present this limitation. The numerical experiments are carried out while taking into account the FEM [13, 30], considering only the two-dimensional case and assuming that the locations belong to the mesh (sub-grid).

This paper is arranged as follows: Section 2 presents the mathematical formulation of the problem, the following Section 3 presents the method of the fundamental solutions in the context of the problem under analysis. In Section 4 some inspirational ideas for genetic algorithms are explored. Section 5 is intended for the presentation of numerical experiments and, finally, Section 6 presents the conclusions of the present study.

2 PROBLEM FORMULATION

Let \( \Omega \subset \mathbb{R}^n, n = 2, 3 \), be an open and bounded domain with Lipschitz boundary. Further, \( b_i(\alpha, \delta_{x_i}) \) denotes a source of intensity \( \alpha_i \) concentrated at the point \( x_i \in \Omega \), where \( \delta_{x_i} = \delta(x-x_i) \) represents the Dirac delta function. Considering a problem modeled by the Poisson equation, the inverse problem under investigation is the reconstruction of a pointwise source (intensity and location) taking into account measurements on the boundary \( \partial \Omega \). In particular, you have the following overdetermined problem:

\[
\begin{align*}
-\Delta u &= b^* \quad \text{in} \quad \Omega \\
 u &= u^* \quad \text{on} \quad \partial \Omega \\
-\partial_n u &= q^* \quad \text{on} \quad \partial \Omega
\end{align*}
\]

(2.1)

where \( b^* = b(\alpha_i^*, \delta_{x_i^*}) \). More specifically, the inverse problem is the reconstruction of intensity \( \alpha_i^* \) and source location \( x_i^* \) taking one boundary condition as data and the other as the corresponding measurement. In the present work, in particular, the data is \( u^*|_{\partial \Omega} \) (Dirichlet) and \( q^*|_{\partial \Omega} \) (Neumann) is the corresponding measurement.
Since the inverse problem (2.1) is ill-posed, the usual strategy of reformulating it as an optimization problem is adopted. More precisely, the minimization of the following cost function is considered:

$$ F(u) = \frac{1}{2} \int_{\partial \Omega} (q^* - \partial_n u)^2 dS, \quad (2.2) $$

where $u$ satisfies the problem (2.3)

$$ \begin{cases} -\Delta u = b_0 & \text{in} \quad \Omega \\ u = u^* & \text{on} \quad \partial \Omega \end{cases} \quad (2.3) $$

for an initial guess $b_0$ given by:

$$ b_0 = \alpha_0 \delta(x - x_0). \quad (2.4) $$

Finally, to solve the inverse problem, a genetic algorithm is applied to minimize the cost function (2.2).

In addition to the advantages of meshless methods over domain discretization strategies, such as ease of implementation in any dimension and low computational cost, for example, it is essential to emphasize that adopting the method of fundamental solutions in the present context allows an adequate representation of the pointwise source by an MFS source point. In fact, contrary to the methodology praxis, this point is allocated within the domain, as will be clarified in the next Section.

3 THE METHOD OF FUNDAMENTAL SOLUTIONS

The method of fundamental solutions, introduced by Kupradze and Aleksidze [18] in 1964, is a meshless method that has received much attention from the scientific community in recent years, especially in reconstruction algorithms, where dependence on discretization can generate artificially accurate results known as inverse crimes [5]. Also, MFS stands out for ease of deployment, computational speed, low storage requirements, and exponential convergence [1, 3, 8, 18]. Thus, in the context of inverse problems, where iterative algorithms are often used, and the associated direct problem needs to be solved repeatedly, these advantages are amplified.

Given the problem (2.3) and the definition of fundamental solution, it can be inferred that identifying a concentrated $b_i$ source is the same as identifying an MFS source point located in $\xi = x_i \in \Omega$ and with intensity $\alpha_i$, the Poisson problem being therefore solved through the Laplace equation, except at the points where there are concentrated sources.

For the sake of simplicity, we introduce the MFS for the two-dimensional case. The fundamental solution, in this case, is given by:

$$ G(\|x - \xi\|) = -\frac{1}{2\pi} \log \|x - \xi\|, \quad (3.1) $$

and the MFS solution is given by the linear combination

$$ u(x) = \sum_{j=1}^{M} a_j G(\|x - \xi_j\|) + \alpha_0 G(\|x - x_0\|), \quad \text{for} \quad x \in \Omega \cup \partial \Omega, \quad (3.2) $$

where $\xi^j$ are the source points (singularities) allocated on a pseudo-boundary outside the domain $\Omega$ and $x_0$ is allocated within the domain to represent the point source (Fig. 1). To determine the $a_j$ coefficients, that is, to solve the direct problem (2.3), it is enough to impose the boundary condition

$$u^*(x_l) = \sum_{j=1}^{M+1} a_j G(||x_l - \xi^j||), \text{ for } x_l \in \partial\Omega, \quad l = 1, M,$$ (3.3)

where $x_l$ are called collocation points (to simplify the notation $a_{M+1} = \alpha_0$ and $\xi_{M+1} = x_0$). Fig. 1 presents a scheme for applying MFS to the direct problem.

Figure 1: MFS scheme for pointwise source reconstruction.

4 GENETIC ALGORITHMS

Motivated by Darwin’s theory of the natural evolution of the species [6], Holland [12] proposed a stochastic algorithm based on a process of natural selection to find the solution of optimization problems with or without restrictions.

Starting from an initial population, the algorithm uses an adaptability function, which depends on the problem under analysis, as a measure of how much more fit one member of the population is than another. This measure is then used to transform the population through three evolutionary mechanisms: selection, crossover, and mutation, in an elitist process that aims at a new population with a greater adaptability average than that of the previous population. In Fig. 2 we present the flowchart of a basic genetic algorithm.

4.1 Genetic Algorithms and the Method of Fundamental Solutions

The use of GAs associated with the MFS has relevant applications in the scientific computing literature. To cite some relevant examples, Jopek and Kołodziej [15] applied GA to the optimal...
Figure 2: Flowchart for a genetic algorithm.

location of source points in MFS. In turn, Gorzelanczyk [10] used both methodologies in the context of twisting bars with multiple connected cross-sections. Santos et al. [24] performed numerical simulations of cathodic protection systems combining MFS and GA. More recently, Antunes et al. [2] use GA as a preprocessor of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm and the MFS as a solver for the associated direct problem to reconstruct the shape and location of holes in an elastic body.

4.2 Genetic Algorithms and Inverse Problems

In this section, we consider how GAs can be applied to the solution of ill-posed IPs. Initially, we consider an associated adaptability function \( F_{\text{adap}} \), to measure the fittest members of the population. In this work, we adopt the most immediate choice consisting of the opposite of the cost function \( F \), given by the equation (2.2)

\[
F_{\text{adap}} = -F = -\frac{1}{2} \int_{\partial\Omega} (q^* - \partial_n u)^2 dS.
\]  (4.1)

Thus, the individuals with the greatest adaptability are those for whom the cost function is closest to zero. We must emphasize that effectiveness and efficiency of the genetic algorithm are strongly related to the appropriate choice of the adaptability function [21].

The first step in the genetic algorithm is to define an initial population of solutions that, in general (and in this work, in particular), is taken randomly. However, the use of supplementary information can be used to accelerate convergence [7].

After the definition of the fitness function, it is necessary to define how the selection is implemented. There are different schemes proposed in the literature, but the most usual are the roulette-wheel selection, the ranking selection, and the stochastic binary tournament selection [9]. In a roulette wheel selection, some chromosomes will have very few chances to be selected if the fitnesses differ greatly. On the other hand, rank selection can lead to a slower convergence, if the
best chromosomes do not differ as much from other ones. In this work, we adopted the stochastic binary tournament selection, where pairs of individuals are randomly chosen from the population, and the fittest individual in the pair is selected. This scheme, in general, converges faster than both the roulette-wheel and ranking schemes, as related in the literature.

The next step in a genetic algorithm is named crossover and it is the most significant phase in the search for the optimal solutions. The idea is randomly choosing two parents who have a relatively high degree of fitness and substituting them with their children obtained by crossing the genes of the parents, taking into account a probability $p_c$, defined a priori. In our numerical experiments, we adopted the single-point crossover, where a point on both parent’s chromosomes is picked randomly at a specific point termed locus. Then, bits to the right of that point are swapped between the two parents to generate children.

Finally, to ensure genetic diversity in the genetic heritage of the population and avoid early convergence, the following procedure is the mutation, which implies that some of the bits in the bit string can be flipped with a probability $p_m$, also predefined.

The genetic algorithm terminates if any stopping criteria are met. The convergence occurs when the population does not produce offspring which significantly differs from the previous generation.

Readers interested in applying genetic algorithms in the context of inverse problems can consult [23, 25] and the references therein.

5 NUMERICAL EXPERIMENTS

In this section, we carry out numerical experiments taking into account the Matlab optimization toolbox Genetic Algorithm function, to illustrate the accuracy of the proposed algorithm. All experiments were performed in a workstation running Windows 10 with Intel Core i7 and 8 GB RAM.

5.1 The two-dimensional case

In this case we adopt $\Omega = B_1(0) = \{(x, y) \in \mathbb{R}^2; x^2 + y^2 < 1\}$, the pseudo-boundary was chosen $\partial B_2(0)$, that is, the boundary of the open ball, centered at the origin, with radius $R = 2$. In addition, we adopt $M = 40$. To avoid inverse crimes, related to the use of the same numerical scheme for the simulation of synthetic data and the solution of the direct problem [5], the MFS scheme for the direct problem has been modified using $M = 60$ and $R = 3.0$. Since the number of source points $(M + 1)$ is greater than the number of collocation points $(M)$, the system associated with MFS was solved by the least squares method.

The parameters of the genetic algorithm are: size population $m = 50$, crossing probability $p_c = 0.5$, and mutation probability $p_m = 0.01$. The stopping criteria are given by the $L = 300$ generations limit and the $tol = 10^{-6}$ tolerance function.
5.1.1 Example 1

Consider a source point located at \( x^* = (-0.4, -0.14) \) with intensity \( \alpha^* = 7.1 \). Computed values are denoted with the subscript \( \text{comp} \): \( (x_{\text{comp}}, y_{\text{comp}}) \) and \( \alpha_{\text{comp}} \). Using MFS to solve the direct problem and get the simulated data and GA to minimize the cost function, you get \( x_{\text{comp}} = (-0.4, -0.14) \) and \( \alpha_{\text{comp}} = 7.1 \). After 5 interactions the value of the adaptability function reaches \( F_{\text{adap}} = -4.9 \times 10^{-6} \). This result is shown in Fig. 3, where the center of the circle represents the position and the radius is proportional to the intensity \( \alpha \) of the reconstructed source. The average computational time in ten runs is about 12 seconds.

![Figure 3: Example 1: a) target: \( b^* \) b) reconstructed source](image)

To mimic the random errors inherent in the experimental measurements, the next example will use additive white Gaussian noise (WGN). More specifically, we will consider an additive noise, with uniform power across the frequency band (the name white noise comes from the analogy with white light). Furthermore, white noise has the property of being noise with the Gaussian distribution, with zero mean. For a justification of using WGN to represent some experimental uncertainties, see [26].

In the next example, the WGN was added to the calculated flow, taking into account the reconstruction of the same synthetic data as the previous example. Results are presented in Table (1) taking into account 1%, 2% and 5% WGN. These results are also presented in Fig. 4. In all cases, the average computational time in ten runs is less than 14 seconds.

5.1.2 Example 2

The following results were obtained:

Observing the results from Table 1, even for polluted data with 5% white Gaussian noise, it was possible to reconstruct, with excellent accuracy, the position, and intensity of a concentrated source after about 250 cost function evaluations, i.e., after 5 generations of the genetic algorithm. More specifically, in all of the above cases, the algorithm converges after 5 interactions.
Table 1: Example 2 for $x^* = (-0.4, -0.14)$ and $\alpha^* = 7.1$.

<table>
<thead>
<tr>
<th>WGN</th>
<th>$x_{\text{comp}}$</th>
<th>$y_{\text{comp}}$</th>
<th>$\alpha_{\text{comp}}$</th>
<th>iterations</th>
<th>$F_{\text{adap}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>-0.406</td>
<td>-0.136</td>
<td>7.111</td>
<td>5</td>
<td>-0.0489</td>
</tr>
<tr>
<td>2%</td>
<td>-0.390</td>
<td>-0.153</td>
<td>7.160</td>
<td>5</td>
<td>-0.0820</td>
</tr>
<tr>
<td>5%</td>
<td>-0.382</td>
<td>-0.154</td>
<td>7.210</td>
<td>5</td>
<td>-0.157</td>
</tr>
</tbody>
</table>

Figure 4: Example 2: Results from noisy data with (a) 1%, (b) 2% and (c) 5% of WGN.

5.2 The three-dimensional case

Similar to the two-dimensional case, $\Omega$ is an open and limited domain with Lipschitz boundary $\partial \Omega$. The pseudo-boundary $\partial \Omega'$ is a surface linearly homotopic to $\partial \Omega$. Finally, $M$ is the number of collocation points. The fundamental solution of the Laplace operator, in this case, is given by:

$$G(||x - \xi||) = \frac{1}{4\pi ||x - \xi||}.$$  \hfill (5.1)

5.2.1 Example 3

In this example, the domain $\Omega$ is also a unitary ball centered on the origin $B(0,1)$. Assuming $M = 53$ equally spaced collocation points and the same amount of source points ($N = 53$) uniformly distributed on $B(0,2)$. Fig. 5 illustrates a simulation of the distribution of source and collocation points.

To avoid inverse crimes, the radius of the pseudo-boundary was changed from $R = 2.0$ to $R = 3.0$ in the MFS scheme for the simulation of $q^*$. Again, the system associated with MFS was solved by the least squares method. Dirichlet data prescribed in the boundary is null, that is, $u^*(x) = 0, \forall x \in \partial \Omega$.

The parameters of the genetic algorithm were $P = 500$, $p_c = 0.5$ and $p_m = 0.01$ and the stopping criteria $tol = 10^{-6}$ and $L = 300$;

The Table (2) below shows the results for noisy data at the 1%, 2% and 5% levels for the reconstruction of $x^* = (-0.4, -0.14, 0.5)$ and $\alpha^* = 7.1$. 

Figure 5: MFS source and collocations points in Example 3 ($R = 2$ for the pseudo-boundary, $M = N = 53$ points on each boundary).

Table 2: Example 3 for $x^* = (-0.4, -0.14, 0.5)$ and $\alpha^* = 7.1$.

<table>
<thead>
<tr>
<th>WGN</th>
<th>$x_{\text{comp}}$</th>
<th>$y_{\text{comp}}$</th>
<th>$z_{\text{comp}}$</th>
<th>$\alpha_{\text{comp}}$</th>
<th>iterations</th>
<th>$\mathcal{F}_{\text{adap}}$</th>
<th>$t(s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>-0.412</td>
<td>-0.149</td>
<td>0.523</td>
<td>7.16</td>
<td>5</td>
<td>-0.0598</td>
<td>110</td>
</tr>
<tr>
<td>2%</td>
<td>-0.448</td>
<td>-0.147</td>
<td>0.541</td>
<td>7.15</td>
<td>7</td>
<td>-0.0881</td>
<td>160</td>
</tr>
<tr>
<td>5%</td>
<td>-0.506</td>
<td>-0.152</td>
<td>0.545</td>
<td>7.21</td>
<td>10</td>
<td>-0.1830</td>
<td>230</td>
</tr>
</tbody>
</table>

Where $t(s)$ is the average computational time from ten runs. We should note that the algorithm converged after 2500, 3500 and 5000 evaluations of the adaptability function when the noises were 1%, 2% and 5%, respectively.

6 CONCLUSIONS

The focus of the present work was the reconstruction of a source concentrated in a Poisson problem by applying the method of fundamental solutions to solve the direct problem and a genetic algorithm to minimize the cost function that associates the inverse problem with an optimization problem. We must emphasize the present work is a significant extension of the conference paper [22] for the three-dimensional case, where, in addition to changing the fundamental solution (3.1) by (5.1), the numerical method becomes more sophisticated.

It should be stressed that the identification of a singularity within the domain constitutes an additional difficulty for the numerical solution of the problem. In fact, the Motz problem, which has a singularity on the boundary, is a benchmark for boundary element methods and MFS [15].
In this work, a modification in the classic MFS was considered, allowing the choice of a source point within the domain, which provided an adequate representation of the pointwise source. Also, this change allowed us to use the fundamental solution of the homogeneous problem, simplifying the algorithm. This original innovation in using the meshless method method allowed the accurate reconstruction of the source location without additional hypotheses, as occurs in domain discretization methods [20].

In both cases studied, in two and three dimensions, the use of Matlab Optimization Toolbox’s “Genetic Algorithm function” led to very accurate results for reconstructing a single source, even when noisy data were considered.

Future work will address the identification of several pointwise sources with a partial reading of the data, considering Helmholtz-type equations, in addition to the identification of the shape and position of a distributed source using the algorithm now proposed for choosing the initial guess in composition with a shape optimization algorithm.

REFERENCES


---