An Improved Kriging Model based on Differential Evolution

Xiaobing Shang, Ping Ma*, Ming Yang
Control and Simulation Center, Harbin Institute of Technology, P.R. China, shangxiaobing1992@126.com, pingma@hit.edu.cn, myang@hit.edu.cn

Abstract
Kriging model is a commonly interpolate approximation method which is widely used in the computer simulation in the past decade. The fitting accuracy is one of the fundamental problems in the research of kriging model, which can be summarized in two aspects, the accurate estimation of model’s parameters and approximate form selection of kriging model. In order to solve the existed problems, an improved parameter estimation method of kriging model base on differential evolution (DE) algorithm is set out in the present paper. Firstly, establish the objective function of DE algorithm depends on the estimation of the model’s accuracy, and get the optimum solution of model’s parameters under the initial condition. Then, a variety of regression function and correlation function in kriging models are selected to compare the fitting accuracy. Finally, the simulation case for outer ballistic data on electromagnetic railgun is examined to determine whether the improved method has priority over traditional one in the approximation accuracy.

Keywords: Kriging model, DE algorithm, approximation accuracy, EM railgun

1 Introduction
Kriging interpolation model, which is an optimal linear unbiased estimate model, is a commonly spatial interpolation model based on the geostatistical variogram theory. In 1989, this theory is introduced to the computer simulation field, and then gradually become a frequently-used interpolation method, which is used in a variety of applications including mechanics engineering, structural optimization and sequential experimental design (Volpi et al, 2015). The approximation accuracy is one of the key problems of current research which including how to effectively estimate the kriging model’s parameters and select the approximate form of model. These two aspects determine the fitting accuracy, which is the significant approach of model investigation.

A mounts of research is taken to solve the problems of improving the model’s accuracy. Jay D. Martin proposed an estimate method of model’s parameters by using maximum likelihood estimate (MLE) and cross validation (CV) methods. The analysis results showed that the MLE is prior to CV method, which are applied in three different dimensional fitting problems (Martin et al, 2005; Martin et al, 2004). Jack P.C. summarized currently research achievements and pointed out that the approximation accuracy is determined by the form and parameters of correlation function (Kleijnen et al, 2009). Søren N. Lophaven developed a matlab tool DACE, which used to compute the kriging model effectively. The model’s parameters are made easy through DACE. Unfortunately, this tool need to limit the minimum and maximum value of parameters, and cannot find the optimal results, which determine the narrow application fields of this method (Lophaven et al, 2002; Lophaven et al, 2002). LIU Xiaolu considers the model’s accuracy is not only determined by parameters estimation, but also by the sampling points. The improved general pattern search (IGPS) algorithm was used to get these points. And a satellite orbit parameter optimization problem is formulated, which showed that kriging models based on global approximations are more accurate than Analyzer. However, this improved method is too complexity and time-consuming and can’t applied in the engineering (Hui et al, 2016; Huang et al, 2011; Liu et al, 2013).

Differential Evolution algorithm (DE) is a novel swarm intelligence method to search for the optimal result based on the cooperation and competition between different individualities. The DE algorithm is regarded as one of the best optimization method. Some experiments over several numerical benchmarks show that DE performs better than the Genetic algorithm (GA) or the Particle Swarm Optimization (Steentjes et al, 2016; Civicioglu et al, 2013). In order to improve the DE’s performance and solve the problems such as convergence speed and time-consuming, some improved DE algorithm is proposed (Sharma et al, 2014; Padhye et al, 2015). Considering the robustness and briefness, the DE method gradually gets more and more concentration among pattern recognition, nonlinear optimize control, mechanical engineering and so on (Neri et al, 2010; Weber et al, 2010).

Considering current situation, how to establish a more accuracy model effectively and quickly is the main problem. Due to the disadvantage of existed approaches, an improved DE based method to estimate
kriging model’s parameters is presented in this paper. DE algorithm is an outstanding method to search for the global optimization parameters with high speed. The approximate optimal parameters can be obtained by this algorithm and be used to build an optimal model. At the end of paper, an engineering example of electromagnetic railgun exterior ballistic data is examined to demonstrate the superiority of kriging model based on DE algorithm.

2 Theory of kriging model

Considering a simulation system, a set of m design sites and the output response sites be expressed as

\[ S = [s_1, s_2, ..., s_m]^T, s_i \in \mathbb{R}^n \]

\[ Y = [y_1, y_2, ..., y_n]^T, y_i \in \mathbb{R}^t \]

Where \( s_i = [s_{i1}, s_{i2}, ..., s_{in}] \) and \( y_i = [y_{i1}, y_{i2}, ..., y_{it}] \) is the \( i \)th order in the experiments. The kriging model regards the deterministic response \( y(x) \) for an \( n \) dimensional input \( x \in \mathbb{R}^n \) as a combination of a regression model and a random function

\[ y(x) = f(x)^T \beta + z(x) \]

Where, \( f(x) \) is a vector component of 1, \( x \) and the other high order items, \( \beta \) is the regression coefficient, \( z(x) \) is the random function, which is assumed to have a zero mean and \( \sigma^2 \) variance. The covariance between \( n \) dimensional inputs \( x_i \) and \( x_j \) be expressed as

\[ Cov(z(x_i), z(x_j)) = \sigma^2 R(\theta, x_i, x_j) \]

Where, \( R(\theta, x, x) \) is the correlation function with specified parameters \( \theta \).

For the set \( S \) of design sites, a \( m \times p \) design matrix \( F \) is constructed with \( F_\theta = f_j(s_i) \)

\[ F = [f(x_1), f(x_2), ..., f(x_n)]^T \]

Furthermore, a correlation matrix \( R \) between the design sites be defined as \( R_\theta = R(\theta, s_i, s_j) \)

\[ r(x) = [R(\theta, s_1, x), R(\theta, s_2, x), ..., R(\theta, s_n, x)]^T \]

The estimation value of output response \( \hat{y}(x) \) is a linear combination of response in the design sites

\[ \hat{y}(x) = \lambda(x)^T Y \]

The kriging model regards \( \hat{y}(x) \) as an optimal linear unbiased estimation of output response \( y(x) \). Then, the problem of kriging interpolation can be transformed to an optimization problem

\[ \min \ E\{\hat{y}(x) - \lambda(x)^T Y\}^2 \]

\[ s.t. \ E\{\hat{y}(x) - \lambda(x)^T Y\} = 0 \]

The solution to the optimization problem above is

\[ \hat{y}(x) = f^T(x) \hat{\beta} + r^T(x) R^{-1} (Y - F \hat{\beta}) \]

Where, the least squares solution of parameter \( \hat{\beta} \) is

\[ \hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y \]

The Mean Square Error (MSE) of \( \hat{y}(x) \) is

\[ \varphi(x) = \sigma^2 \left[ 1 - \left( f^T(x) r(x) \right) \right] \left( \frac{0}{F} \right) \left( \frac{F}{R} \right) \left[ \frac{f(x)}{r(x)} \right] \]

3 Kriging model based on DE algorithm

In this section, the improved method based on DE algorithm is described in detail. Firstly, the basic theory and operations are introduced to support the application in the kriging model. Then, the second part is to assess the kriging model’s performance, and lead to an optimization problem. Finally, DE method is applied to solve the optimization problem, and the flowchart of DE based kriging model is used to show the process.

3.1 Theory of DE algorithm

Differential Evolution (DE) is a well-known and simple approach for global optimization, which consists of three basic operations: mutation, crossover and selection. The compute process of DE algorithm is familiar with GA method and can be summarized in the following steps:

- Initialization: set the algorithm’s parameters including population members \( N_P \), variable number \( D \), mutagenic factor \( F \) and crossover probability \( CR \), then the initial population \( \{x_{ij}(0)| x_{ij} \leq x_{ij}(0) \leq x_{ij}, i=1,2,...,N_P, j=1,2,...,D\} \) can be generated by

\[ x_{ij}(0) = x_{ij}^l + \text{rand}(0,1) \times (x_{ij}^u - x_{ij}^l) \]

Where \( x_{ij}(0) \) is the \( j \)th variable among the \( i \)th individuality in the \( 0 \)th generation, \( x_{ij}^l \) and \( x_{ij}^u \) are the upper and lower bound of \( x_{ij} \), respectively. The variable \( \text{rand}(0,1) \) is the random number of uniform distribution in the interval \((0,1)\).

- Mutation: mutation is a basic operation in the DE algorithm. Which is the largest difference compared with the GA method. The mutation operation can be described as

\[ v_{i}(t+1) = x_{i}(t) + F \times (x_{r_1}(t) - x_{r_2}(t)) \]

Where, \( r_1, r_2, r_3, i \) are not equal with each other and \( r_1, r_2, r_3 \in [1,2,...,N_P] \). \( v_{i}(t+1) \) is the \((t+1)\)th generation mutation individuality, \( x_{i}(t) \) is the \( i \)th individuality in the \( t \)th generation.
Crossover: for the given individuality $x_i(t)$, it’s necessary to use the operation of crossover to generate new experiment individuality $u_j(t)$. The equation is

$$u_j(t+1) = \begin{cases} v_{ij}(t+1), \text{rand}(j) \leq CR \text{ or } j = \text{randn}(i) \\ x_i(t), \text{rand}(j) > CR \text{ or } j \neq \text{randn}(i) \end{cases}$$

(13)

Where $\text{rand}(j) \in [0,1]$ is the random number of uniform distribution, $j$ is the $j$th variable of the individuality, $\text{randn}(i) \in \{1,2,...,D\}$.

Selection: the greedy strategy is applied in the DE algorithm to search for the best individuality in a population between $u_j(t)$ and $x_i(t)$, the operation can be expressed as

$$x_i(t+1) = \begin{cases} u_j(t+1), f(u_j(t+1)) < f(x_i(t)) \\ x_i(t), f(u_j(t+1)) \geq f(x_i(t)) \end{cases}$$

(14)

After iterating the operations above for many times, it is easy to get the optimal vector in the solving space, which will be used in the optimization of kriging model’s parameters in the next step.

### 3.2 Kriging interpolation based on DE algorithm

In general, the quality of a model can be measured by two aspects: 1) the accuracy of data in the design sites. 2) Accuracy in predicting the output response at the estimating points. However, due to the unbiased predictor of kriging model, which means the estimation in the design sites is exactly, the first measurements don’t need to consider in this work. And the second aspect to calculate the accuracy at the estimating points is very significant for kriging model.

For a simulation system, the input set of predictor is expressed as

$$P = \{p_1, p_2,...,p_n\}^T, p_i \in \mathbb{R}^n$$

and the system real output response and kriging predictor can be expressed as

$$Y_p = \{y_{p_1}, y_{p_2},...,y_{p_n}\}^T, y_i \in \mathbb{R}^q$$

$$\hat{Y}_p = \{\hat{y}_{p_1}, \hat{y}_{p_2},...,\hat{y}_{p_n}\}^T, \hat{y}_i \in \mathbb{R}^q$$

(15)

The sum of squared prediction errors is

$$SS_P = \sum_{i=1}^{n}(y_{p_i} - \hat{y}_{p_i})^2$$

$$= \sum_{i=1}^{n}(y_{p_i} - \bar{y}_p + \bar{y}_p - \hat{y}_{p_i})^2 = \sum_{i=1}^{n}(y_{p_i} - y_p - \bar{y}_p + \bar{y}_p - \hat{y}_{p_i})^2 = SS_F + SS_e$$

The accuracy of kriging model is defined as

$$R^2_\beta = 1 - \frac{SS_F}{SS_T}, R^2_\beta > 0$$

(17)

For the definition of kriging model accuracy, $R^2_\beta \in (0,1)$, the value of $R^2_\beta$ determine the fitting accuracy. In order to get a higher accuracy, the value of $R^2_\beta$ is as large as it can be. Furthermore, the sum of squared errors of prediction is a fixed value for a specific set of predictors, which means the value of $SS_e$ is as small as it can be. Thus, the model assessment can be transformed to an optimization problem

$$\min \sum_{i=1}^{n}(\hat{y}_{pi} - y_{pi})^2 \quad s.t. \quad \hat{y}_p = f^T(p_i)\hat{\beta} + \sigma^2(p_i)R^{-1}(Y - F\hat{\beta})$$

(18)

For a given form of correlation function, the key factor that has a highest influence on the fitting accuracy is the parameters’ estimation. From the objective function above, it’s obvious that the kriging model accuracy is related with the parameters $\beta$ and $R$. There are two methods to estimate the parameters, which are maximum likelihood estimation (MLE) and cross validation (CV). The MLE method was used in this paper and estimation results of $\beta$ and $\sigma^2$ is

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y$$

$$\sigma^2 = \frac{1}{n}(y - F\hat{\beta})^T R^{-1} (y - F\hat{\beta})$$

(19)

It’s clearly that the parameters $\beta$ and $R$ are related with the estimate value of $\theta = \{\theta_1, \theta_2, ..., \theta_s\}$, which transform the objective function’s optimization to estimate the parameters $\theta$. Two computational problems often exist when estimating the parameters using traditional method: 1) the maximum likelihood estimate of the parameters may be multimodality and hardly to find the optimal result. Therefore, the selection of initial value of parameters has a strong influence on the estimation result. 2) The method above is very suitable for approximating low dimensional model, and has a poor effect of high dimensional one. Considering the robustness and practicability of DE algorithm to solve nonlinear, non-differentiable, multi-extremum and high dimensional problems, an improved kriging model based on DE algorithm is proposed in this paper.

### 3.3 Process of kriging Model based on DE algorithm

The kriging model’s parameters $\theta = \{\theta_1, \theta_2, ..., \theta_s\}$ are the variable need to optimize which is determined by the system input. Due to the optimization’s purpose is to improve the model fitting accuracy, (11) is utilized to act as the objective function of DE algorithm. The process to optimize kriging model’s parameters is presented as show in Figure 1, which can summarize in three aspects:

- Setting DE algorithm parameters: set the objective function with (18) to evaluate the efficiency of DE algorithm. Depend on the research target and compute scope, initialization the DE algorithm’s parameters and...
carry out the operation of crossover, mutation and selection.

- Data preprocessing: generally, the establishment of kriging model is based on the assumption that the system’s input and output data satisfy the normal distribution $N(0,1)$. So it’s significant to verify and normalize data’s normalization before constructing the kriging model. As stated in Sec. II, the system input samples and output response can be expressed as

\[ S = \{s_1, s_2, \ldots, s_n\}^T, s_i \in \mathbb{R}^n \]

\[ Y = \{y_1, y_2, \ldots, y_n\}^T, y_i \in \mathbb{R}^q \]  

(20)

For simple, data normalization can be described as follows:

\[ I_i = \left( s_i - \bar{s}_i \right) / \sigma_{s_i}, i = 1, 2, \ldots, n, \]

\[ O_j = \left( y_j - \bar{y}_j \right) / \sigma_{y_j}, j = 1, 2, \ldots, q \]  

(21)

Where, $\bar{s}_i$ and $\bar{y}_j$ is the mean of ith input and jth output variable, respectively; $\sigma_{s_i}$ and $\sigma_{y_j}$ is the mean variance of ith input and jth output variable, respectively. The calculate equation is

\[ \bar{s}_i = \frac{1}{m} \sum_{i=1}^{m} s_i, \quad \bar{y}_j = \frac{1}{m} \sum_{i=1}^{m} y_j \]

\[ \sigma_{s_i} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (s_i - \bar{s}_i)^2}, \quad \sigma_{y_j} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_j - \bar{y}_j)^2} \]  

(22)

After normalization, the mean and variance of the input and output data are 0 and 1, respectively.

- Model computation and optimization: select the specific kriging model including type of correlation function and form of regression function, and then optimize the kriging model’s parameters based on DE algorithm, which is stated in Sec. II. As shown in Figure 1, the improved method can be divided into two aspects: the DE algorithm global search and kriging model establishment.

From the process above, it’s obvious that the DE algorithm is very suitable for the kriging model’s optimization. With the DE method, the best parameters can be gained to establish an optimal kriging model.

4 An engineering example

To demonstrate the validity of kriging model based on DE algorithm, EM railgun exterior ballistic simulation data is taken for example. Due to the complexity of railgun ballistic, the flight range of projectile suffers from a variety of factors, which have a different degree of effect, and coupled with each other. Thus, during the research of projectile’s performance, it’s crucial to study the relation between projectile’s flight range and a set of factors. The linear regression model is the ordinary method to solve this problem. However, the different factors may interaction with each other, and hardly to draw an expression with the projectile’s range. Furthermore, the classic regression is the maximum likelihood estimation of sampling points and hardly to solve the problem of multivariable and multimodality. Thus, the kriging model based on DE algorithm is used to establish the model between range and factors.

In this paper, six separate factors including projectile mass, launch velocity, launch angle, deflection angle, wind velocity in x and z direction, are considered in the paper to establish the relationship with flight range using kriging model. Meanwhile, DE algorithm was utilized to optimize the kriging model’s parameters and get a higher accuracy.

4.1 Setting DE algorithm parameters

In order to improve the efficiency of DE algorithm and speed up the algorithm convergence, the parameters of DE algorithm set as follows:

The size of population is 30. Mutagenic factor is 0.5, and crossover probability factor is 0.7. Considering the optimal parameters set $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ represent weight of each dimension, it’s important to restrict the span of $\theta_i, i = 1, 2, \ldots, n$. In general, the range of $\theta$ is set as $[0, d_{min}, 10.0d_{max}]$, where $d_{min}$ and $d_{max}$ represents the minimum and maximum distance of ith input parameter sets, respectively.

4.2 Data preprocessing

Before establishing the kriging model, it’s obvious to preprocess the flight range data. The normalization verification result is demonstrated in Figure 2 and 3, which including two graphs: the left and right one is the bar graph of flight range and result of normality test, respectively. As the verify result show, the disperse position of range data coincide with the standard normalization reference line. Furthermore, it’s necessary to take the range data with normalization operation, which is obtained from (12) and (13). Thus, the railgun data meets the requirement of kriging model construction after preprocessing.

Figure 1. Flow chart of the kriging based on DE method.
4.3 Model computation and optimization

After accomplishing the data preprocessing, the last step is to compare the model accuracy by two aspects: 1) the approximation accuracy of kriging model based on DE algorithm and ordinary one, 2) the accuracy of different forms of correlation function and regression function \( f(x) \). In order to demonstrate the superiority of improved method, three types of correlation function and five forms of regression function \( f(x) \) is adapted in the DE method in this paper. The types of correlation function consist of Exp, Gauss and Linear and the forms of regression function are summarized in Table 1.

As show in Figure 4, the kriging model based on DE algorithm is prior to the traditional one in the model’s accuracy. For the different types of correlation function, the Gauss model has a significant advantage than the other models. Meanwhile, the fourth type of regression function has the highest accuracy in the five forms of regression function. So it’s essential to compare different forms of regression function to select the highest accuracy one.

![Figure 2. Histogram of the range data.](image)

![Figure 3. Normalized test result.](image)

![Figure 4. Histogram of the kriging model error.](image)

Table 1. Forms of the regression function \( f(x) \).

<table>
<thead>
<tr>
<th>Indication</th>
<th>Forms of ( f(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>([1, x_1, x_2, ..., x_n])</td>
</tr>
<tr>
<td>2</td>
<td>([1, x_1, x_2, ..., x_n, x_1^2, x_2, x_3, x_4, ..., x_n^2])</td>
</tr>
<tr>
<td>3</td>
<td>([1, x_1, x_2, ..., x_n, x_1^2, x_2^2, ..., x_n^2])</td>
</tr>
<tr>
<td>4</td>
<td>([1, x_1, x_2, ..., x_n, x_1x_2, x_1x_3, x_1x_4, ..., x_{n-1}x_n])</td>
</tr>
</tbody>
</table>

Table 2. Sum errors of kriging model based On DE.

<table>
<thead>
<tr>
<th>Indication</th>
<th>EXP</th>
<th>GAUSS</th>
<th>LIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.58e4</td>
<td>4.99e4</td>
<td>4.46e4</td>
</tr>
<tr>
<td>2</td>
<td>3.065e4</td>
<td>1.99e4</td>
<td>2.52e4</td>
</tr>
<tr>
<td>3</td>
<td>1.86e4</td>
<td>5.75e3</td>
<td>1.59e4</td>
</tr>
<tr>
<td>4</td>
<td>1.38e4</td>
<td>6.52e3</td>
<td>1.06e4</td>
</tr>
<tr>
<td>5</td>
<td>3.51e4</td>
<td>2.23e4</td>
<td>3.2e4</td>
</tr>
</tbody>
</table>

Table 3. Sum errors of traditional kriging model.

<table>
<thead>
<tr>
<th>Indication</th>
<th>EXP</th>
<th>GAUSS</th>
<th>LIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.77e4</td>
<td>1.08e5</td>
<td>2.34e5</td>
</tr>
<tr>
<td>2</td>
<td>3.11e4</td>
<td>2.53e4</td>
<td>3.43e4</td>
</tr>
<tr>
<td>3</td>
<td>1.93e4</td>
<td>1.2e4</td>
<td>3.29e4</td>
</tr>
<tr>
<td>4</td>
<td>1.47e4</td>
<td>8.49e3</td>
<td>2.72e4</td>
</tr>
<tr>
<td>5</td>
<td>3.77e4</td>
<td>3.21e4</td>
<td>4.05e4</td>
</tr>
</tbody>
</table>

As show in Figure 4, the kriging model based on DE algorithm is prior to the traditional one in the model’s accuracy. For the different types of correlation function, the Gauss model has a significant advantage than the other models. Meanwhile, the fourth type of regression function has the highest accuracy in the five forms of regression function. So it’s essential to compare different forms of regression function to select the highest accuracy one.
convergence, it’s obvious that the approach could find the optimization result quickly and effectively. Thus, the kriging model based on DE algorithm improves the approximate on accuracy and has a valuable application in engineering.

![Figure 5. Iteration process of the DE optimization.](image)

### 5 Conclusions

The parameters estimation and approximate form selection to obtain a higher accuracy model is a basic existed problem in the research of kriging interpolation. In the traditional method, the model parameters selection always depends on users. In order to solve the problem and get an optimal model, an improved kriging model based on DE algorithm to approximate simulation model was presented in this article. The research of kriging model proves that correlation function’s parameters have a strong influence on the fitting accuracy. Then the DE algorithm establishes the objective function by fitting accuracy to optimize the kriging model’s parameters. The EM railgun exterior ballistic data was taken for instance to demonstrate the priority of improved method. Three forms of correlation function and five types of regression model are used to compare the approximation accuracy of kriging model based on DE algorithm with traditional one. The simulation results show that the kriging model based on DE algorithm has the higher accuracy, and a fine prospect of engineering application.

There is a need for more research on the improved method. Although the DE algorithm perform well in kriging model, it’s essential to have deep research of the DE parameters such as population members NP, variable number D and mutagenic factor F, which determine the search efficiency. So the future work can be concentrated on parameters selection strategy to obtain a higher speed.

### Acknowledgements

This work is supported by National Science Foundation of China (No. 61374164).

### References


