An Efficient Kriging Modeling Method Based on Multidimensional Scaling for High-Dimensional Problems

Yu Ge 1, Junjun Shi 1,2, Yaohui Li 1,2,* and Jingfang Shen 2

1 College of Mechanical and Electrical Engineering, Xuchang University, Xuchang 461000, China; 11996010@xcu.edu.cn (Y.G.); sjj_516@webmail.hzau.edu.cn (J.S.)
2 College of Science, Huazhong Agricultural University, Wuhan 430070, China; sjf_712@mail.hzau.edu.cn
* Correspondence: lyh@xcu.edu.cn

Abstract: Kriging-based modeling has been widely used in computationally intensive simulations. However, the Kriging modeling of high-dimensional problems not only takes more time, but also leads to the failure of model construction. To this end, a Kriging modeling method based on multidimensional scaling (KMDS) is presented to avoid the “dimensional disaster”. Under the condition of keeping the distance between the sample points before and after the dimensionality reduction unchanged, the KMDS method, which mainly calculates each element in the inner product matrix due to the mapping relationship between the distance matrix and the inner product matrix, completes the conversion of design data from high dimensional to low dimensional. For three benchmark functions with different dimensions and the aviation field problem of aircraft longitudinal flight control, the proposed method is compared with other dimensionality reduction methods. The KMDS method has better modeling efficiency while meeting certain accuracy requirements.

Keywords: surrogate model; Kriging method; multidimensional scaling; high-dimensional problems

1. Introduction

With the improvement of design levels and application requirements [1–3], the optimal performance of products has attracted much attention. Although the computing power of computers continues to increase, both the computer simulation and analysis of complex problems are still very time consuming. The traditional modeling method requires more expensive simulation evaluation, so it is not suitable for solving black-box problems. Therefore, using Kriging surrogate models with higher approximation accuracy instead of complex simulation sequential modeling methods has become a research hotspot in the field of product design and optimization [4–9]. Approximate construction and space exploration are mainly completed through experimental design in Kriging models, given their potential to estimate information. The Kriging model, which could also help to shorten the design cycle, reduce research and development costs, and improve design accuracy, is suitable for solving complex black-box problems.

Although Kriging has demonstrated its great advantages in many fields [10–14], there are still bottlenecks that need to be overcome when using Kriging to solve high-dimensional modeling problems. There are two main solutions. One is the use of parallel computing [15–17]; the other is to use dimensionality reduction methods to reduce the dimensionality of Kriging modeling, thereby improving the efficiency of Kriging modeling. In this work, we prefer the latter.

Depending on the different commonly used benchmark functions, it takes about 2000 s to 3500 s to complete the modeling of the 80-dimensional problem by sampling 200 training design points through sequential Kriging optimization according to our test. Therefore, the construction of a Kriging surrogate under high-dimensional problems is time consuming. Most of the time is spent on the inversion of the covariance correlation matrix and the solving of the Kriging correlation parameter (also known as the hyper-parameter).
Kriging modeling [18–20] not only considers the relationship between the location of the estimated point and the known data location, but also considers the spatial correlation of the variables. The relationship between positions is mainly coordinated and improved by the distance between data points. Moreover, the spatial correlation of variables [21,22], which is the key to constructing the covariance correlation matrix, is manifested in the Kriging correlation parameters obtained by maximizing the likelihood function through the pattern search algorithm. For Kriging sequential modeling with an increasing number of data points, the dimensions of the correlation parameter vectors are consistent with the dimensions of the design variables without considering dimensionality reduction, while the optimal correlation parameters are obtained by maximizing the complex multimodal likelihood functions. The estimation of the likelihood function in each loop demands the inversion of the covariance matrix. In addition, the dimensionality of the covariance matrix increases with the increase in the amount of sample data, which will cause the complexity of the inversion operation of the covariance matrix to also increase, so that the time spent obtaining the optimal correlation parameters will also increase. The above observations may explain why the Kriging modeling of high-dimensional problems is time consuming.

If the dimensionality of the design variables is high, the optimization of the correlation parameters becomes a complex and multimodal high-dimensional optimization problem, which will exceed our preset calculation amount. Many of the latest works have been focused on solving the “curse of dimensionality” problems that affect Kriging modeling. There are usually three ways to deal with such high-dimensional problems: one is to simplify the calculation of the inverse of the correlation matrix by dividing the high-dimensional space into several subspaces, and then build a high-dimensional model through integration; one is to improve the efficiency of Kriging modeling by reducing the calculation time of hyper-parameters under the condition that the dimensionality does not change; the other way is to establish a Kriging approximation model after dimensionality reduction.

Applying gradient enhancement Kriging (GEK) [23–26] using cheap gradient information to solve high-dimensional problems can effectively enhance the accuracy of the Kriging model. To this end, solving the large correlation matrices problem that affects the GEK models applied to high-dimensional problems enables GEK to be efficient and accurate. For example, the weighted gradient enhancement Kriging method [27] reduces the time spent decomposing large correlation matrices in high-dimensional gradient enhancement Kriging by establishing a series of sub-models with smaller correlation matrices and integrating into these sub-models appropriate weight coefficients. Chen et al. proposed a screening-based gradient-enhanced Kriging modeling method [28], which uses feature selection technology to evaluate and rank the impact of each input variable on the output, and proposes empirical screening criteria to select the gradient, and finally uses the selected partial gradient to construct the Kriging model, thereby improving the modeling efficiency of high-dimensional problems. By combining recursive decomposition and sequential sampling for the establishment of accurate Kriging problems in high-dimensional spaces, Kang et al. [29] perform variable decomposition using an interactive estimation of the full-dimensional Kriging meta-model, and divide the whole space into several subspaces according to the result of variable decomposition. The conversion from high-dimensional space sampling to low-dimensional space sampling is realized in each subspace.

Bouhlel et al. [30,31] proposed a method (KPLS) whereby the partial least squares (PLS) technique is combined with the Kriging model. This method uses the least squares dimensionality reduction method in the process of establishing the Kriging model, which reduces the number of hyper-parameter calculations of the model to be consistent with the number of principal components retained by the PLS, thereby accelerating the construction of the Kriging model. Bouhlel et al. [32] combined KPLS and GEK to propose a new GE-KPLS method. It adds gradient information to the correlation matrix. Although the size of the correlation matrix is slightly increased, it greatly reduces the number of hyper-parameter calculations. Zhao et al. [33] proposed a new Kriging modeling method (KMIC)
that combines Kriging and the maximum information coefficient (MIC). This method uses
the MIC to estimate the relative size of hyper-parameters according to the characteristics
of the hyper-parameters after optimization. Then, the maximum likelihood estimation is
reconstructed to reduce the calculation amount of hyper-parameter estimation. A distance
correlation-based Kriging modeling method (KDIC) [34], which is somewhat similar to the
KMIC method, uses distance correlation (DIC) to estimate the relative size of the hyper-
parameters in the new correlation function, and transforms the hyper-parameter tuning
process into a one-dimensional optimization problem.

For the dimensionality reduction methods used to improve modeling efficiency,
Liu et al. [35] used Sammon mapping [36] to transform the training data from the original
space to the lower dimensional space by minimizing the distance error between the sample
points in the two data sets before and after dimensionality reduction to achieve the purpose
of dimensionality reduction. However, this mapping method does not consider the relation-
ship between input and response when reducing dimensionality. Zhou et al. [37] proposed
a modeling method that combines Kriging and sliced inverse regression (SIR). The method
uses SIR to construct a new projection vector to achieve dimensionality reduction. The
projection vector reduces the dimensionality of the original input vector without losing
the basic information of the model response. In low-dimensional space, a new Kriging
correlation function is constructed using the tensor product of several correlation functions
in each projection direction.

The above-mentioned dimensionality reduction method reduces the modeling time
while ensuring that certain model accuracy requirements are met. After all, things have
two sides. The improvement of modeling efficiency will lead to the loss of accuracy to a
certain extent. Therefore, how to improve the modeling efficiency while minimizing the
loss of precision still needs further research.

For this purpose, a new Kriging modeling method (KMDS) based on multidimen-
sional scaling (MDS) is proposed in this work. The basic idea is as follows: firstly, high-
dimensional design data are converted to low-dimensional data through MDS; then, the
reduced dimensional data are used to establish a low-dimensional Kriging model. After
that, new high-dimensional design points can be obtained via Latin hypercube sampling
and estimated using expensive functions, and then the sampling points can be added to the
high-dimensional data set. Finally, we can go back to the first step and repeat the process
until the number of sample points meets the stopping criterion.

The remaining arrangements for this work are as follows. The background of the
Kriging model is introduced Section 2. The third states the key technology and outlines
the realization process of the proposed method. The benchmark functions are tested and
analyzed in Section 4. The fifth section is the study of the simulation examples, which are
based on aircraft longitudinal flight control data. Conclusions are given in Section 6.

The KMDS algorithm addresses the difficulties of achieving high-dimensional mod-
eling based on the Kriging model, and greatly improves the modeling efficiency on the
basis of meeting certain accuracy requirements. It guides a new direction for the modeling
of high-dimensional engineering problems. Therefore, it is suitable for publication in the
Algorithms journal.

2. Kriging Model

Kriging is widely used to build surrogate models in applied mathematics, statistics,
and machine learning. The introduction of its concise theory is given below.

For d-dimensional design data, \( X = [x_1, \ldots, x_m]^T \in \mathbb{R}^{m \times d} \) and objective response
\( Y = [y_1, \ldots, y_m]^T, Y \in \mathbb{R}^{m \times 1} \), the Kriging model is expressed as follows:

\[
y(x) = \mu(x) + z(x)
\]

The formula consists of two parts. The first part \( \mu(x) \) is the trend function (also called
the expectation function). When \( \mu(x) \) is equal to 0, \( \mu \), and \( \sum_{i=1}^{d} \beta_i f_i(x) \), the models are
called simple Kriging, ordinary Kriging, and standard Kriging models, respectively, where
the function \( f_i(x) \) is the regression model, and the parameter \( \beta_i \) is the coefficient of the \( i \)-th
regression model. Standard Kriging uses the regression function to determine the change
in the process mean \([38]\).

The second part in Formula (1) is a stochastic process model established by observing
the data and quantifying the correlation of the data. It is often considered to be the
realization of the stochastic process \( Z(x) \) with a mean of 0 and a variance of \( \sigma^2 \). The
covariance between design points is shown in Formula (2):

\[
\text{Cov}(Z(x_i)Z(x_j)) = k(x_i, x_j) = \sigma^2 R(\theta, x_i, x_j) \tag{2}
\]

where process variance \( \sigma^2 \) is a scalar and \( \theta \) is a relative parameter (or called hyper-
parameter) of the kernel function. By optimizing \( \theta \), the correlation between the design
points can be adjusted adaptively. The expression of the spatial correlation function
\( R(\theta, x_i, x_j) \) is shown in Equation (3):

\[
R(\theta, x_i, x_j) = \prod_{k=1}^{d} R_k(\theta_k, x_i^k - x_j^k) \tag{3}
\]

There are many options for spatial correlation functions. However, the most widely
used correlation function is the Gaussian function. The expression of the Gaussian model
is shown in Equation (4):

\[
R_k(\theta_k, x_i^k - x_j^k) = \exp\left(-\theta_k |x_i^k - x_j^k|^2\right) \tag{4}
\]

Under the above assumptions, the best linear unbiased estimate of \( y(x) \) is:

\[
\hat{y}(x) = f^T(x)\hat{\beta} + r^T(x)R^{-1}(Y - F\hat{\beta}) \tag{5}
\]

Parameter \( r \) and \( R \) in Formula (5) are:

\[
r(x) = \{R(x, x_1), R(x, x_2), \ldots , R(x, x_m)\}
R_{m \times m} = \begin{bmatrix}
R(x_1, x_1) & \cdots & R(x_1, x_m) \\
\vdots & \ddots & \vdots \\
R(x_m, x_1) & \cdots & R(x_m, x_m)
\end{bmatrix} \tag{6}
\]

Among them, the least square estimation \( \hat{\beta} \) of \( \beta \) can be calculated using Equation (7):

\[
\hat{\beta} = \left( F^TR^{-1}F \right)^{-1}F^TR^{-1}Y \tag{7}
\]

In addition, the predicted mean square error \( \hat{s}(x) \) at any point \( x \) can be obtained by
calculating the formula in Equation (8):

\[
\hat{s}(x) = \text{MSE}[\hat{y}(x)] = \hat{\sigma}^2 \left\{ 1 - \left[ f(x)^T r(x) \right] \begin{bmatrix} 0 & F^T \end{bmatrix} \begin{bmatrix} f(x) \\ r(x) \end{bmatrix} \right\} \tag{8}
\]

where,

\[
\hat{\sigma}^2 = \frac{1}{m} \left( Y - F\hat{\beta} \right)^T R^{-1} \left( Y - F\hat{\beta} \right) \tag{9}
\]

It can be seen from Equations (3), (7) and (8) that the matrix \( R \), vector \( \hat{\beta} \), and the
estimated variance \( \hat{s}(x) \) of any point \( x \) all depend on the value of the parameter \( \theta \). Based
on the maximum likelihood estimation theory, the unconstrained optimization algorithm is used to maximize Equation (10) to obtain the optimal $\theta$ value [39]:

$$\theta = \text{argmax} \left(-\frac{m}{2} \ln \hat{\sigma}^2 - \frac{1}{2} \ln |R| \right)$$ (10)

In fact, the best approximation curve is not necessarily obtained from the optimal value. As long as the obtained theta value is close to the optimal value, the Kriging model will have a good approximation result [39].

3. KMDS Method

3.1. Multidimensional Scaling (MDS) Method

In most cases, although the sample data that people observe or collect are high-dimensional, they may only be low-dimensionally distributed and closely related to the learning task. Generally speaking, a good dimensionality reduction method can usually greatly reduce the computational cost while preserving most of the function information. Dimensionality reduction means transforming the original high-dimensional attribute space into a low-dimensional “subspace” through a certain mathematical transformation. Multidimensional scaling (MDS) [40–43] can alleviate the problems associated with sparse sample data and the difficulty of distance calculations in high-dimensional situations. Different from the idea of linear dimensionality reduction, multidimensional scaling pays more attention to the internal characteristics of high-dimensional data. It focuses on preserving the “similarity” of information in the high-dimensional space, and the “similarity” is usually defined by the Euclidean distance. That is to say, the main idea of this work is to keep the “similarity” of information in the original space unchanged in the low-dimensional space.

The goal of dimensionality reduction is to obtain the representation $Z \in \mathbb{R}^{d' \times m}$ of the sample in the $d'(d' \leq d)$ dimensional space, and the Euclidean distance of any two samples in the $d'$ dimensional space is equal to the distance in the original space; that is, $\|z_i - z_j\| = t_{ij}$.

Let $B = Z^T Z \in \mathbb{R}^{m \times m}$, where $B$ is the inner product matrix of the sample data after dimensionality reduction, and the calculation formula of each element in $B$ is $b_{ij} = z_i^T z_j$.

Therefore, the relationship between $T$ and $B$ can be stated by Equation (12):

$$t_{ij}^2 = \|z_i - z_j\|^2 = \|z_i\|^2 + \|z_j\|^2 - 2z_i^T z_j$$

$$= b_{ii} + b_{jj} - 2b_{ij}$$ (12)

Center the sample $Z$ after dimensionality reduction; that is, $\sum_{i=1}^{m} z_i = 0$ ($0 \in \mathbb{R}^{d'}$ is a vector of all zeros). Obviously, the sum of all rows and the sum of all columns in matrix $B$ are both zero, i.e., $\sum_{i=1}^{m} b_{ij} = \sum_{j=1}^{m} b_{ij} = 0$. Under this condition, Equations (13)–(15) can be generated from Equation (12):

$$\sum_{i=1}^{m} t_{ij}^2 = \sum_{i=1}^{m} b_{ii} + \sum_{i=1}^{m} b_{jj} - 2 \sum_{i=1}^{m} b_{ij}$$

$$= tr(B) + mb_{jj}$$ (13)
where the symbol \( tr(\cdot) \) refers to the trace of the matrix, and the trace of the inner product matrix \( B \) is \( tr(B) = \sum_{i=1}^{m} \| z_i \|^2 \). Assume that the following Formulas (16)–(18) hold, Equation (19) can be obtained due to Equations (12)–(18):

\[
\begin{aligned}
    h_1 &= \frac{1}{m} \sum_{j=1}^{m} t_{ij}^2 \\
    h_2 &= \frac{1}{m} \sum_{i=1}^{m} t_{ij}^2 \\
    h_3 &= \frac{1}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} t_{ij}^2
\end{aligned}
\]  

(14)

(15)

(16)

(17)

(18)

The inner product matrix \( B \) can be obtained through the derivation of Equation (19) and the distance matrix \( T \), which remains unchanged before and after dimensionality reduction.

\[
    b_{ij} = -\frac{1}{2} \left( t_{ij}^2 - b_{ii} - b_{jj} \right)
    \begin{align*}
    &= -\frac{1}{2} \left( t_{ij}^2 - \frac{1}{m} \left( \sum_{i=1}^{m} t_{ij}^2 - tr(B) \right) - \frac{1}{m} \left( \sum_{i=1}^{m} t_{ij}^2 - tr(B) \right) \right) \\
    &= -\frac{1}{2} \left( t_{ij}^2 - h_1 - h_2 + \frac{2}{m} tr(B) \right) \\
    &= -\frac{1}{2} \left( t_{ij}^2 - h_1 - h_2 + h_3 \right)
    \end{align*}
\]  

(19)

The eigenvalue decomposition of matrix \( B \) is performed in Equation (20) to complete the dimensionality reduction:

\[
    B = V \Lambda V^T
\]  

(20)

where parameter \( \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_d) \) is the diagonal matrix composed of eigenvalues and \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d \), \( V \) is the eigenvector matrix, since \( \Lambda \) is a diagonal matrix, \( \Lambda = \Lambda^T \). According to Equation (12) and \( B = Z^T Z \), the Equations (21) and (22) can be formed:

\[
    B = Z^T Z = V \Lambda V^T = V \sqrt{\Lambda} \sqrt{\Lambda} V^T
\]  

(21)

\[
    Z = \sqrt{\Lambda} V^T
\]  

(22)

Assuming that there are \( d^* \) non-zero eigenvalues, they form a diagonal matrix \( \Lambda_* = diag(\lambda_1, \lambda_2, \ldots, \lambda_{d^*}) \). Let \( V_* \) denote the corresponding eigenvector matrix, then \( Z \) can be expressed as follows:

\[
    Z = \Lambda_*^{1/2} V_*^T \in \mathbb{R}^{d^* \times m}
\]  

(23)

In practical applications, in order to reduce the dimensionality effectively, it is often only necessary that the distance in the space after dimensionality reduction is as close as possible to the distance in the original space, and does not have to be strictly equal. Next, arrange the \( d' \) non-zero eigenvalues in order from largest to smallest (it is worth noting that the order has been arranged during eigenvalue decomposition in MATLAB), and take the first \( d' \) (\( d' << d \)) eigenvalues as a low-dimensional diagonal matrix \( \tilde{\Lambda} = \)
At this time, let \( \tilde{V} \) denote the corresponding eigenvector matrix, then \( Z \) can be expressed by Equation (24):

\[
Z = \tilde{\Lambda}^{1/2} \tilde{V}^T \in \mathbb{R}^{d' \times m}
\]

(24)

Based on the above analysis, the MDS method can be simply described in Algorithm 1.

**Algorithm 1 Multidimensional scaling (MDS) method**

**Input:** Distance matrix \( T \in \mathbb{R}^{m \times m} \); low-dimensional space dimension \( d' \).

**Output:** Matrix \( \tilde{\Lambda}^{1/2} \tilde{V} \in \mathbb{R}^{m \times d'} \).

**Process:**
1. Calculate \( h_1, h_2 \) and \( h_3 \) according to Equations (16)–(18);
2. Calculate the inner product matrix \( B \) according to Equation (19);
3. Perform eigenvalue decomposition on inner product matrix \( B \);
4. Take \( \tilde{\Lambda} \) as the diagonal matrix composed of \( d' \) maximum eigenvalues, and \( \tilde{V} \) as the corresponding eigenvector matrix.

### 3.2. Determination of the Space Dimension \( d' \) after Dimensionality Reduction

Multidimensional scaling can freely determine the low-dimensional space dimension \( d' \) when performing dimensionality reduction on high-dimensional data. In this work, the dimension \( d' = 3 \) is selected. For specific reasons, we use the 50-dimensional Griewank function as an example to explain. The expression of the Griewank function is shown in Equation (25):

\[
y(x) = \sum_{i=1}^{50} \frac{x_i^2}{4000} - \prod_{i=1}^{50} \cos \left( \frac{x_i}{\sqrt{i}} \right) + 1 \quad -600 \leq x_i \leq 600 \quad i = 1, \ldots, 50
\]

(25)

Table 1 is the modeling result of the 50-D Griewank function through Kriging and KMDS. Among them, KMDS-1, KMDS-2, KMDS-3, KMDS-4 and KMDS-5 refer to reducing the high-dimensional data to 1-D (one dimension), 2-D, 3-D, 4-D and 5-D data, respectively. During the modeling process, the Latin hypercube sampling (LHS) method is used to generate \( N \) (\( N = 10 \) or \( N = 20 \)) initial sample points. The modeling time and accuracy of Kriging and KMDS methods when adding 10\( N \) sample points are compared. Each method was run 10 times separately, and finally the average value was taken as the result so as to ensure the accuracy of the results.

**Table 1.** Modeling time and accuracy of Kriging and KMDS methods.

| Griewank (50D) | Kriging | KMDS-1 | KMDS-2 | KMDS-3 | KMDS-4 | KMDS-5 |
|---------------|---------|--------|--------|--------|--------|--------|
| 110 t (s)     | 29.2118 | 0.5247 | 0.5648 | 1.1065 | 2.0409 | 3.0003 |
| RMSE          | 17.9327 | 187.3951 | 85.5546 | 18.0510 | 17.9595 | 17.8875 |
| 220 t (s)     | 279.4006 | 2.9883 | 3.1055 | 5.7695 | 11.6390 | 15.5642 |
| RMSE          | 12.8756 | 120.5884 | 73.9394 | 12.9491 | 12.7546 | 12.6127 |

The RMSE (root mean square error) values in Table 1 are obtained by using leave-one-out cross-validation [44]. The concrete expression for finding the RMSE is as follows:

\[
\text{RMSE} = \frac{1}{k} \sqrt{\frac{1}{k} \sum_{i=1}^{k} \delta_i^2}
\]

(26)

The parameter \( k \) here represents the number of samples in the current data. If the Kriging model is used to estimate the variance of point \( x_i \), we first need to reconstruct the Kriging model with the remaining \( k - 1 \) sampling points, except point \( x_i \). Then, we
calculate the estimated variance \( \hat{s}_i^2 \) of point \( x_i \) by using the newly built Kriging model and Formula (8). After repeating \( k \) times to complete the variance estimation of these \( k \) sampling points, the average value can be calculated to obtain the RMSE following Equation (26).

To visually prove the stability and effectiveness of KMDS, the Griewank function was tested 10 times with different sample points, and the results are shown in the box plots in Figures 1 and 2. The middle line (solid red line) of each box plot shows the median value of the 10 data. The blue solid line above the box is the upper quartile line, and the blue solid line below the box is the lower quartile line. The top horizontal line is drawn on the basis of the upper quartile line at a distance of 1.5 times the interquartile range (the upper quartile minus the lower quartile) or the limit of the data (if the limit of the data is within 1.5 times the interquartile range). The bottom horizontal line is drawn in the same way. Abnormal data points outside the horizontal line are represented by “+”. Judging from the median time (red solid line) and the size of the box (the area formed by the upper quartile and the lower quartile), the median value of the KMDS method is very low, and the box area is also very small. It can be seen from the RMSE box plot that the RMSE of the other four methods are ideal except for the larger median value and frame area of the KMDS-1 and KMDS-2 methods. This proves that the KMDS method can guarantee accuracy and has a greater advantage with regard to speed. In addition, this method has very little fluctuation in terms of the time spent on these ten modeling scenarios, and the error convergence of KMDS-3, KMDS-4 and KMDS-5 is relatively stable.

![Box Plot](https://via.placeholder.com/150)

**Figure 1.** (a) Modeling time for 110 points. (b) Modeling error (RMSE) for 110 points.

![Box Plot](https://via.placeholder.com/150)

**Figure 2.** (a) Modeling time for 220 points. (b) Modeling error (RMSE) for 220 points.
The following characteristics can be further explained on the basis of the data shown in Algorithm 1 and Figures 1 and 2. (1) The running time of the KMDS method is greatly shortened compared to Kriging modeling. When adding 100 sample points, its modeling time is already 10–60 times faster than the Kriging method. When 200 sample points are added, the time spent during the KMDS method is about 18–94 times shorter than that of Kriging method. (2) In the low-dimensional space after dimensionality reduction, the higher the dimensionality, the longer the modeling time required by the KMDS method. For example, KMDS-5 took 15.5642 s to model the 220 sampling points, while KMDS-1 only took 2.9883 s. However, compared with the 279.4006 s required by the Kriging method, the actual time they took is not worth mentioning. (3) In terms of error, it is obvious that as the dimensionality of the low-dimensional space increases, the number of errors produced by the KMDS model becomes lower and lower. The error score of KMDS-1 and KMDS-2 is very large compared with the Kriging model, but when \( d' = 3 \), the error score of KMDS-3 is very close to that of the Kriging model. Even though the RMSE will become lower and lower as the dimension \( d' \) increases, the rate of decline is almost invisible. In addition, these figures also show that KMDS-1 and KMDS-2 not only have large number of errors, but also have poor stability. However, KMDS-3, KMDS-4 and KMDS-5 not only produce fewer errors, but they are also more stable in comparison to KMDS-1 and KMDS-2.

Based on the above examples and analysis, and balancing the requirements of accuracy and efficiency, it is reasonable to reduce the high-dimensional problem to a three-dimensional problem. The reason is that when \( d' = 3 \), the model error score of KMDS-3 is very close to that of the Kriging model, which can meet the accuracy requirements, and the time taken is dozens of times faster than the Kriging model. When \( d' > 3 \), the KMDS-4 and KMDS-5 methods do have fewer errors, but compared to KMDS-3, the error score is only reduced a little, while the time has doubled.

In addition, Sammon mapping [36] was used to further verify the dimensionality reduction effect of the KMDS method in this work. Sammon mapping achieves dimensionality reduction by minimizing the distance difference between corresponding points in the two spaces of the original space and the low-dimensional space; that is, by minimizing the error function of Equation (27):

\[
E = \frac{1}{\sum_{i<j} \text{dist}(x_i, x_j)} \times \sum_{i<j} \left[ \text{dist}(x_i, x_j) - \text{dist}(\bar{x}_i, \bar{x}_j) \right]^2 \text{dist}(x_i, x_j)
\]  

(27)

The above error is calculated for the original data and low-dimensional data of each modeling process, and the average error is taken as the final error. Accordingly, Table 2 shows the results of the error given by Equation (27). It can be seen from Table 2 that when \( d' = 3 \), the range of the distance error score between the two spaces is 0.4 to 0.5, which is very small. It also proves the effectiveness of the KMDS dimensionality reduction method.

| Griewank (50D) | KMDS-1 | KMDS-2 | KMDS-3 | KMDS-4 | KMDS-5 |
|----------------|--------|--------|--------|--------|--------|
| 110            | E      | 0.6818 | 0.5288 | 0.4315 | 0.3601 | 0.3008 |
| 220            | E      | 0.7148 | 0.5711 | 0.4753 | 0.4055 | 0.3509 |

3.3. Flow Chart and Specific Steps of the KMDS Method

The flowchart of the KMDS method is shown in Figure 3, and the specific process is completed in the following seven steps.

**Step 1: Initial experimental design.** Set the sets \( X, X_1 \) and \( Y \) to be empty sets. The \( N \) initial sample points \( x_i (i = 1, \cdots, N) \) are generated using the Latin hypercube sampling (LHS) method, and they are added to the set \( X \). Estimate the expensive function of the initial sample points to obtain \( y_i (i = 1, \cdots, N) \), which is added to the set \( Y \).
Step 2: Dimensionality reduction. The multidimensional scaling method is used to reduce the dimension of data set \(X\). Section 3.1 shows that the ideal low-dimensional space dimension \(d'\) after dimensionality reduction is 3; that is, the original high-dimensional data are reduced to three-dimensional data. See Section 3.1 for the specific dimensionality reduction process. The low-dimensional data are stored in the set \(X_1\).

Step 3: Build the Kriging model. The Kriging model is constructed by using new sample data sets \(X_1\) and \(Y\). At this time, the Kriging model is three-dimensional.

Step 4: Obtain a new sample point. Obtain a new sample point \(x_i\) through the LHS, and calculate its expensive function value \(y_i\).

Step 5: Determine whether the stopping criterion is met. In this work, the estimated number of sampling points is used as the stop criterion, and the maximum number of evaluations is set to \(N_{\text{max}}\). If the number of evaluations is less than \(N_{\text{max}}\), proceed to Step 6; if the number of evaluations is greater than or equal to \(N_{\text{max}}\), skip to Step 7.

Step 6: Update the data set. The new sample point \(x_i\) is added to the sample set \(X\), and \(y_i\) is added to the sample set \(Y\). Go back to Step 2.

Step 7: End. Output the Kriging model.

---

Figure 3. Flowchart of the KMDS method.

4. Numerical Test

The KMDS method is compared with the ordinary Kriging model and the KPLS method to prove the effectiveness of the KMDS method. The Kriging model has been explained in detail in the second section. The KPLS method proposed by Bouhlel in 2016 [30,31] is a combination of Kriging and the partial least square method. The core idea is to improve the modeling speed by reducing the complexity of hyper-parameter calculations. Therefore, the comparison between the KPLS method and the KMDS method will further
demonstrate and verify the effectiveness of the KMDS method from the perspective of dimensionality reduction.

We use three different types and different dimensions of benchmark functions to test the effectiveness of the KMDS method. Different test conditions are used for each benchmark function to meet most of the situations encountered by the Kriging surrogate model in the modeling field. All tests are performed in MATLAB 2019a using a RedmiBook Pro machine with an AMD Ryzen 7700U 1.8 GHz CPU and 16 GB RAM.

4.1. Griewank Function Test

The first test function used is the Griewank with greater complexity, and its specific expression is shown in Equation (28):

\[
y(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad -600 \leq x_i \leq 600 \quad i = 1, \ldots, d
\]

For the case where the dimensions of the function are 30 and 70, 110 and 220 sampling points are used for testing in order to effectively verify the effect and stability of the KMDS method in different dimensions and with different sample sizes. The Kriging method, KMDS, KPLS-1, KPLS-2 and KPLS-3 were used to test the modeling time and convergence accuracy of the Griewank function. Table 3 shows the average modeling time and its standard deviation (STD). Furthermore, the average modeling error and its standard deviation were obtained by testing the Griewank function in five different modeling methods, each of which represents the average of 10 test results.

| Griewank Statistic | Kriging | KMDS | KPLS-1 | KPLS-2 | KPLS-3 |
|--------------------|---------|------|--------|--------|--------|
| 30D 110 t(s)       | 21.3954 | 1.1498 | 1.2686 | 2.5323 | 5.7447 |
| STD               | 0.7999  | 0.0578 | 0.2175 | 0.1594 | 0.1042 |
| RMSE              | 14.0585 | 13.3769 | 12.6699 | 12.7925 | 12.6464 |
| STD               | 0.8021  | 0.9089  | 0.4453 | 0.1568 | 0.4736 |
| 220 t(s)          | 169.0297 | 6.8024 | 10.8794 | 25.7107 | 58.5037 |
| STD               | 1.8902  | 0.2296  | 0.5263 | 0.3441 | 1.2995 |
| RMSE              | 10.0340 | 9.6798  | 8.4632 | 7.9802 | 8.3273 |
| STD               | 0.5234  | 0.2891  | 0.4772 | 0.2215 | 0.5698 |
| 70D 110 t(s)       | 55.8389 | 1.1410 | 2.6329 | 6.3931 | 15.2219 |
| STD               | 1.1604  | 0.0359  | 0.1836 | 0.4135 | 0.5666 |
| RMSE              | 21.2036 | 20.1981 | 20.0653 | 20.0093 | 20.0010 |
| STD               | 1.3604  | 0.7127  | 0.7714 | 1.0447 | 1.2915 |
| 220 t(s)          | 504.3929 | 5.7970 | 21.3492 | 59.6396 | 138.2662 |
| STD               | 1.4190  | 0.1934  | 1.0168 | 2.1632 | 2.9013 |
| RMSE              | 14.9673 | 14.5079 | 13.5973 | 13.5935 | 13.5839 |
| STD               | 0.5226  | 0.3136  | 0.3038 | 0.5506 | 0.6101 |

The test results show that the KMDS method always takes the shortest time to build the Kriging model under the conditions of different dimensions and different samples. With the increase in the number of dimensions and the number of sample points, the efficiency advantage of the KMDS method becomes more obvious. The main reason is because the KMDS method transforms high-dimensional modeling into three-dimensional modeling, so the amount of time spent calculating the correlation matrix and hyper-parameters in the modeling process is greatly reduced, thereby improving the modeling efficiency of Kriging. In terms of accuracy, the accuracy of the KMDS method is very close to that of ordinary Kriging. Compared with the KPLS method, when the condition of the test function is 30D and 110 samples, the test time of the KMDS method is only 0.1188 s faster.
than that of the KPLS-1 method, but the error score of KPLS-1 is 0.707 lower than that of KMDS. The KPLS-1 method is better at this time. The reason may be that the KPLS method considers the relationship between design variables and response variables to make the accuracy higher. In addition, retaining a principal component greatly reduces the amount of calculation of hyper-parameters, thereby improving the efficiency of Kriging modeling. However, when the Griewank function is 30D and 220 sample points, KMDS is faster than KPLS-1 by 4.077 s, but the difference in the error score is only 1.2166. For the Griewank function of 70D and 110 sample points, KMDS is 2.3 times faster than KPLS-1, and the difference in the error score is only 0.1328. Therefore, with the increase in the number of sample points or the increase in the number of test dimensions, the test time of the KPLS-1 method is much longer than that of the KMDS method. At this time, the advantage of the KMDS method in terms of time is far greater than the disadvantage in terms of accuracy. For KPLS-2 and KPLS-3, although the accuracy of these two methods is a bit higher than that of KMDS, the modeling efficiency cannot match the KMDS method. The reason may be that the calculation burden of the correlation matrix and hyper-parameters in the KMDS method is greatly reduced during the modeling process, while the KPLS method mainly reduces the calculation cost of hyper-parameters to improve efficiency. Therefore, the efficiency of the KPLS method is lower than that of KMDS. Therefore, the KMDS method can effectively improve modeling efficiency while maintaining a small loss of accuracy. In addition, the standard deviation also reflects the stability of the KMDS method, especially in terms of modeling efficiency.

Figure 4 shows the average change trend of the modeling time and modeling error in ten tests of five different modeling methods for the 70D Griewank function. For each test, the Kriging model is firstly established through 10 initial sample points, and the corresponding modeling time is used as the first time. In the next iteration process, whenever a new sampling point is added, the Kriging model is updated again, and the total time spent executing the algorithm is recorded. When the number of newly added sampling points reaches 100, the entire iterative process is stopped. Figure 4 clearly shows how as the number of sample points increases, the modeling error scores of the five modeling methods gradually decrease, and the time it takes for KMDS is much less than that of the other four methods. In addition, when the modeling accuracy of the five different methods gradually stabilized, their RMSE values differed slightly, while the RMSE obtained by the KMDS method was generally small at first. Therefore, the KMDS method is superior to the other four methods in terms of efficiency under the premise of ensuring accuracy for high-dimensional problems.

![Figure 4](image_url)
The red and blue curves in Figures 5 and 6 are error bar graphs for the ten test results. The point on the solid line is the average of the test results of each method. The upper end of the error bar is the maximum value of the test result of each method. Similarly, the lower end of the error bar is minimum of the test result of each method. The green and purple curves are the standard deviations of ten tests with different methods. Compared with Kriging, KPLS-1, KPLS-2 and KPLS-3, the proposed method uses the least modeling time, and the error accuracy obtained is not much different from that of KPLS-1, KPLS-2 and KPLS-3, which is exactly what we expect. Moreover, the standard deviation chart also reflects the stability of the proposed method.

Figure 5. (a) Modeling time and its standard deviation for the 30D Griewank function. (b) RMSE and its standard deviation for the 30D Griewank function.

Figure 6. (a) Modeling time and its standard deviation for the 70D Griewank function. (b) RMSE and its standard deviation for the 70D Griewank function.

4.2. Rothyp Function Test

The second test function is the Rothyp function, whose expression is shown in Equation (29):

$$y(x) = \sum_{i=1}^{d} \sum_{j=1}^{i} x_{ij}^2 - 65.536 \leq x_{i} \leq 65.536 \quad i = 1, \cdots, d$$ (29)

For the 50-dimensional and 100-dimensional problems of this function, 110 and 220 sampling points were used for testing. Its testing process is consistent with that of the first function.
of the first function. Table 4 shows the average modeling time and its standard deviation and the average modeling error score and its standard deviation, which were obtained by testing the Rothyp function under different modeling methods. Each result is the average of 10 experiments. Figure 7 shows the average change trend for the modeling time and modeling error in the ten-time modeling process of the five different modeling methods for 100D Rothyp function. Figures 8 and 9 are the error bar graphs and standard deviation curve graphs of the ten test results.

Table 4. Time and RMSE for Rothyp function.

|    | Rothyp | Kriging | KMDS | KPLS-1 | KPLS-2 | KPLS-3 |
|----|--------|---------|------|--------|--------|--------|
| 50D| 110    | 30.455  | 12.036| 1.8799 | 4.3277 | 10.8365|
|    | STD    | 0.5218  | 0.0859| 0.1281 | 0.2367 | 0.8983 |
|    | RMSE   | 2.55 × 10^4 | 2.29 × 10^4 | 2.31 × 10^4 | 2.36 × 10^4 | 2.41 × 10^4 |
|    | STD    | 1.17 × 10^3 | 9.64 × 10^2 | 1.19 × 10^3 | 1.63 × 10^3 | 1.22 × 10^3 |
| 220|       | 280.586 | 5.7309| 15.8398| 41.9871| 94.4945|
|    | STD    | 3.6827  | 0.1871| 1.0820 | 1.2273 | 4.5775 |
|    | RMSE   | 1.78 × 10^4 | 1.81 × 10^4 | 1.66 × 10^4 | 1.64 × 10^4 | 1.59 × 10^4 |
|    | STD    | 6.32 × 10^2 | 9.65 × 10^2 | 7.17 × 10^2 | 7.19 × 10^2 | 5.77 × 10^2 |
| 100D| 110    | 112.640 | 1.1639| 3.8012 | 9.4307 | 22.8246|
|    | STD    | 3.0488  | 0.0426| 0.1356 | 0.5199 | 1.4751 |
|    | RMSE   | 7.21 × 10^4 | 7.16 × 10^4 | 6.88 × 10^4 | 6.63 × 10^4 | 6.92 × 10^4 |
|    | STD    | 5.04 × 10^3 | 5.45 × 10^3 | 4.80 × 10^3 | 1.64 × 10^3 | 4.46 × 10^3 |
| 220|       | 1033.631| 5.9439| 30.8228| 80.9683| 197.8211|
|    | STD    | 6.3241  | 0.2184| 1.7778 | 3.8911 | 7.5289 |
|    | RMSE   | 5.06 × 10^4 | 5.13 × 10^4 | 4.82 × 10^4 | 4.76 × 10^4 | 4.75 × 10^4 |
|    | STD    | 1.80 × 10^3 | 2.62 × 10^3 | 1.55 × 10^3 | 1.19 × 10^3 | 1.34 × 10^3 |

Figure 7. (a) The average change trend for the modeling time for the 100D Rothyp. (b) The average change trend for RMSE for the 100D Rothyp.
From Table 4 and Figures 7–9, the following characteristics can be observed. (1) The advantage of the modeling time for the KMDS method is particularly obvious. For example, in the case of 100 dimensions, the modeling time of KPLS-3 is approximately 33.28 times that of KMDS. Even the KPLS-1 method that retains only one principal component takes 5.19 times as much time as the proposed method, which fully illustrates the advantages of the KMDS method in terms of modeling efficiency. (2) Although the accuracy of the KMDS method is not as good as that of the KPLS methods, it is already very close to them. After all, any method has two sides. Increased modeling speed usually leads to lower accuracy. Significantly, this work has greatly improved the modeling efficiency under the condition of close accuracy, which is what we expect. (3) From the perspective of the error bar graph for modeling time and modeling accuracy and the trend chart for the standard deviation, the KMDS method also has good stability and effectiveness compared to the other four methods. This shows that the proposed dimensionality reduction method is effective in actual tests.
4.3. Michalewicz Function Test

The third test function is the Michalewicz function, which is expressed in Equation (30). This function is tested in 40 and 80 dimensions with 110 and 220 data points. The testing process of the third function is consistent with the testing process of the first function. Table 5 shows the average modeling time and its standard deviation and the average modeling error and its standard deviation, which were obtained by testing the Michalewicz function under different modeling methods. Each result is the average of 10 experiments. Figure 10 shows the average change trend for the modeling time and modeling error in the ten-time modeling process for the different modeling methods for the 40D Michalewicz function. Figures 11 and 12 are error bar graphs and standard deviation curve graphs of the ten test results.

\[
y(x) = -\sum_{i=1}^{d} \sin(x_i) \cdot \left(\sin\left(i \cdot x_i^2 / \pi\right)\right)^{160} \quad 0 \leq x_i \leq \pi \quad i = 1, \cdots, d \quad (30)
\]

Table 5. Time and RMSE for Michalewicz function.

|          | Michalewicz | Kriging | KMDS  | KPLS-1 | KPLS-2 | KPLS-3 |
|----------|-------------|---------|-------|--------|--------|--------|
| 40D      |             |         |       |        |        |        |
| 110      | t(s)        | 27.7459 | 1.2073| 1.5069 | 3.1195 | 6.8043 |
|          | STD         | 0.6586  | 0.0650| 0.1738 | 0.3594 | 0.7782 |
|          | RMSE        | 0.0860  | 0.0869| 0.0873 | 0.0869 | 0.0870 |
|          | STD         | 0.0054  | 0.0111| 0.0053 | 0.0090 | 0.0033 |
| 220      | t(s)        | 227.0483| 6.6940| 11.6145| 27.3576| 59.8047|
|          | STD         | 1.2774  | 0.1752| 2.1331 | 2.1720 | 5.4872 |
|          | RMSE        | 0.0613  | 0.0656| 0.0604 | 0.0598 | 0.0602 |
|          | STD         | 0.0021  | 0.0041| 0.0023 | 0.0032 | 0.0023 |
| 80D      |             |         |       |        |        |        |
| 110      | t(s)        | 74.3425 | 1.1404| 2.5859 | 7.5471 | 15.3412|
|          | STD         | 1.219   | 0.0181| 0.1840 | 0.4382 | 1.4657 |
|          | RMSE        | 0.1260  | 0.1254| 0.1264 | 0.1209 | 0.1198 |
|          | STD         | 0.0115  | 0.0072| 0.0058 | 0.0133 | 0.0055 |
| 220      | t(s)        | 685.0288| 6.0802| 20.8315| 65.1704| 138.4195|
|          | STD         | 3.5189  | 0.2259| 1.5820 | 6.2537 | 14.2178|
|          | RMSE        | 0.0869  | 0.0906| 0.0882 | 0.0864 | 0.0864 |
|          | STD         | 0.0058  | 0.0051| 0.0041 | 0.0043 | 0.0023 |

Figure 10. (a) The average change trend for the modeling time for the 40D Michalewicz function. (b) The average change trend for RMSE for the 40D Michalewicz function.
The aircraft longitudinal flight control system [45,46] can stabilize and control the aircraft according to the signal law, so as to realize the control and stability of the various functions of the aircraft. What the aircraft longitudinal flight control system needs to aid their design. This paper uses the Simulink simulation software in MATLAB to simulate the aircraft longitudinal flight control system. What the aircraft longitudinal flight control system needs to aid their design. This paper uses the Simulink simulation software in MATLAB to simulate the aircraft longitudinal flight control system. Through the control system, the aircraft is deflected, and all the design goals of aircraft must obey this highest goal. A good flight control system should be able to ensure the safe flight of the aircraft. The aircraft functions and flight control requirements that modern aircraft need to meet are becoming more and more complex, resulting in more and more complex flight control systems. Therefore, the flight control system is used to transmit control signals from the pilot. Through the control system, the control panel of the aircraft is deflected, and all the design goals of aircraft must obey this highest goal. A good flight control system should be able to ensure the safe flight of the aircraft.

From Table 5, it is easy to see that the test results for the Michalewicz function are similar to those of the Griewank function and the Rothyp function. However, in the case of 40D and 110 sampling points, the test results of the Michalewicz function are slightly different: the modeling time of KMDS is still the shortest, while the modeling error of the ordinary Kriging method is slightly smaller than that of the KMDS method. The complexity of the function itself and the uniformity of the proportions of each dimension may cause this problem. Figures 10–12 also show the advantages of the KMDS method. Comparing KMDS and KPLS, it can be seen that KMDS is better than KPLS in terms of modeling time and its standard deviation. In some cases, the model accuracy or the average accuracy of the proposed method is better than the KPLS method, which further illustrates the effectiveness and stability of the KMDS method.

In summary, it can be seen from the test results of all benchmark functions that the modeling efficiency of the KMDS method is the highest. In terms of modeling accuracy, test results for the KMDS method are superior to those of Kriging method in most cases. Compared with the KPLS methods, the accuracy of KMDS method is slightly worse, but it is also relatively close, and even better than them in some cases. Moreover, with the increase in the dimensionality and sample size, the KMDS method can show its absolute advantage in terms of time spent with a small loss of precision. Therefore, the KMDS method is the best choice for high-dimensional problems.
advantage in terms of time spent with a small loss of precision. Therefore, the KMDS method is effective at solving high-dimensional problems.

5. Test on Aircraft Longitudinal Flight Control Problem

Since the advent of the aircraft, flight safety has always been the highest goal of aircraft design, and all the design goals of aircraft must obey this highest goal. A good flight control system should be able to ensure the safe flight of the aircraft. The aircraft functions and flight control requirements that modern aircraft need to meet are becoming more and more complex, resulting in more and more complex flight control systems. Therefore, flight simulation systems, as an important tool for aircraft design and simulated flight training, have received more and more attention. As the core component of modern aircraft and flight simulators, the flight control system is used to transmit control signals from the pilot. Through the control system, the control panel of the aircraft is deflected according to the signal law, so as to realize the control and stability of the various functions of the aircraft.

The aircraft longitudinal flight control system [45,46] can stabilize and control the aircraft’s pitch angle, altitude, speed, etc. When pulling the control stick backwards, the elevator deflects upwards by an angle, which produces a downward pitching force on the horizontal tail, forming a pitching operating moment on the center of gravity of the aircraft, forcing the nose to tilt up, and the angle of attack increases. Similarly, when pulling the control stick forward, the elevator deflects an angle downward, generating upward momentum on the horizontal tail, forming a pitching operating moment on the center of gravity of the aircraft, forcing the nose to descend, and the angle of attack increases. Since the flight process of the aircraft is affected by the pitch angle, angle of attack, lift, drag, acceleration of gravity, sideslip angle, aircraft gravity, etc., it is necessary to continuously adjust the pitch angle of the aircraft to keep the aircraft in balance.

MATLAB simulation software is widely used in the simulation of flight control systems to aid their design. This paper uses the Simulink simulation software in MATLAB to simulate the aircraft longitudinal flight control system. What the aircraft longitudinal flight control system needs to complete is the simulation of the pilot stick force, and at the same time calculate the corresponding pitch angle of the aircraft in real time. The pitch angle of the aircraft is closely related to the force of the control stick, the flight state of the aircraft, the control mode and the structural parameters of the control system. These parameters (see Figure 13 for details) are used as the design variables of the KMDS modeling method, and the aircraft longitudinal flight control simulation is thereby combined with KMDS. In the KMDS modeling process, each sample point is used as the parameter value of the aircraft’s longitudinal flight control, so that the simulation result can be obtained with Simulink. The simulation results are time dependent. In order to better illustrate the superiority of the KMDS method, under the same parameter conditions, we take the pitch angle of the aircraft at 3s as the simulation result and output it to the MATLAB workspace. We use the data obtained from the simulation to establish Kriging and record the modeling time and modeling error through the KMDS modeling method. In addition, the Kriging model is directly established with the data obtained from the simulation, and the modeling time and modeling error are also recorded.
The following two conclusions can be drawn from Figure 14. (a) It can be seen from Figure 14 that the time required for KMDS to establish Kriging and to directly establish the Kriging model gradually increases as the number of sample points increases. However, the time required to directly establish the Kriging model is more and more than the time required for KMDS modeling with the increase in the number of sampling points, and it
finally reaches a result where the time difference is more than ten times. (b) The figure shows that the model error is gradually decreasing as the number of sample points increases. However, the modeling error of KMDS to establish Kriging has always been smaller than that of directly establishing Kriging, which shows that the KMDS method has higher modeling accuracy. In summary, for the simulation of an aircraft longitudinal flight control system, the KMDS method can improve the modeling efficiency and modeling accuracy of the Kriging model.

6. Conclusions

The high dimensionality and complexity of engineering design problems leads to expensive calculation costs. The Kriging model can alleviate this bottleneck problem. However, the Kriging model is also very time consuming when approximating high-dimensional problems, solutions to some of which it even fails to construct. The main cost is in the inversion of the covariance correlation matrix and the solution of the hyper-parameters of the Kriging model. To this end, this work proposes a Kriging modeling method based on multidimensional scaling to deal with high-dimensional problems. Firstly, the dimensionality of high-dimensional sample data was reduced via the MDS method to convert high-dimensional problems to low-dimensional problems. Then, a new sample point was added to the high-dimensional sample data \( X \), and the dimensionality of the sample set was reduced through MDS again. Finally, a Kriging model was established using the new sample data \( X_1 \) after dimensionality reduction. The test results from the benchmark functions and simulation problem prove that the proposed method can greatly improve the modeling efficiency of Kriging while ensuring certain accuracy requirements.

Unluckily, the method proposed in this work sometimes loses some model accuracy when solving high-dimensional problems. For further research, we will start with the following three aspects: (1) based on the existing modeling efficiency, we will study how to choose more promising samples to further improve the modeling accuracy of the Kriging model; (2) in order to improve the accuracy of the model after dimensionality reduction, we will study how to combine the characteristic parameters of the Kriging model with mutual entropy or a \( t \)-test to select the dimension with the largest amount of information that is worthy of in-depth study; (3) for non-stationary, high-dimensional engineering problems, we will study how to achieve dimensionality reduction through the non-stationary random Kriging model while ensuring the accuracy requirement is also one of the future research directions.

Author Contributions: Conceptualization, Y.L. and J.S. (Junjun Shi); writing—review and editing, Y.G.; software, J.S. (Junjun Shi); writing—original draft, Y.L.; review and check, J.S. (Junjun Shi), J.S. (Jingfang Shen). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (No. 51775472), Science and Technology Innovation Talents in Universities of Henan Province (No. 21HASTIT027) and Henan Excellent Youth Fund Project (No. 202300410346), Training plan of Young Backbone Teachers in Universities of Henan Province (No. 2020GGJS209).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Ghosh, S.; Ghosh, S.; Chakraborty, S. Seismic fragility analysis in the probabilistic performance-based earthquake engineering framework: An overview. *Int. J. Adv. Eng. Sci. Appl. Math.* **2021**, *13*, 122–135. [CrossRef]
2. Gaxiola-Camacho, J.R.; Azizsoltani, H.; Haldar, A.; Vazirizade, S.M.; Villegas-Mercado, F.J. Chapter 13—Novel concepts for reliability analysis of dynamic structural systems. In Handbook of Probabilistic Models; Samui, P., Tien Bui, D., Chakraborty, S., Deo, R.C., Eds.; Butterworth-Heinemann: Oxford, UK, 2020; pp. 305–346.

3. Ghosh, S.; Roy, A.; Chakraborty, S. Kriging Metamodeling-Based Monte Carlo Simulation for Improved Seismic Fragility Analysis of Structures. *J. Earthq. Eng.* 2021, 25, 1316–1336. [CrossRef]

4. Martin, J.D.; Simpson, T.W. Use of Kriging Models to Approximate Deterministic Computer Models. *AIAA J.* 2005, 43, 853–863. [CrossRef]

5. Kleijnen, J.P.C. Regression and Kriging metamodels with their experimental designs in simulation: A review. *Eur. J. Oper. Res.* 2017, 256, 1–16. [CrossRef]

6. Martínez-Frutos, J.; Herrero-Pérez, D. Kriging-based infill sampling criterion for constraint handling in multi-objective optimization. *J. Glob. Optim.* 2015, 64, 97–115. [CrossRef]

7. Li, Y.; Shi, J.; Shen, J.; Cen, H.; Chao, Y. An adaptive Kriging method with double sampling criteria applied to hydrogen preparation case. *Int. J. Hydrogen Energy* 2020, 45, 31689–31705. [CrossRef]

8. Li, Y.; Shen, J.; Cai, Z.; Wu, Y.; Wang, S. A Kriging-Assisted Multi-Objective Constrained Global Optimization Method for Expensive Black-Box Functions. *Mathematics* 2021, 9, 149. [CrossRef]

9. Dellino, G.; Kleijnen, J.P.C.; Meloni, C. Robust Optimization in Simulation: Taguchi and Kriging Combined. *INFORMS J. Comput.* 2012, 24, 471–484. [CrossRef]

10. Xing, J.; Luo, Y.; Gao, Z. A global optimization strategy based on the Kriging surrogate model and parallel computing. *Struct. Multidiscip. Optim.* 2020, 62, 405–417. [CrossRef]

11. Van Stein, B.; Wang, H.; Kowalczyk, W.; Emmerich, M.; Bäck, T. Cluster-based Kriging approximation algorithms for complexity reduction. *Appl. Intel. 2020*, 50, 778–791. [CrossRef]

12. Kleijnen, J.P.C. Simulation Optimization through Regression or Kriging Metamodels. In *High-Performance Simulation-Based Optimization*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 115–135.

13. Bonnot, J.; Menard, D.; Desnos, K. Fast Kriging-based Error Evaluation for Approximate Computing Systems. In Proceedings of the 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, 9–13 March 2020.

14. Shi, R.; Liu, L.; Long, T.; Wu, Y.; Tang, Y. Filter-based adaptive Kriging method for black-box optimization problems with expensive objective and constraints. *Comput. Methods Appl. Mech. Eng.* 2019, 347, 782–805. [CrossRef]

15. Bae, S.-H.; Choi, J.Y.; Qiu, J.; Fox, G.C. Dimension reduction and visualization of large high-dimensional data via interpolation. In Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, Association for Computing Machinery, Chicago, IL, USA, 21–25 June 2010; pp. 203–214.

16. Mascagni, M. High-dimensional numerical integration and massively parallel computing. *Contemp. Math.* 1991, 115, 53–73.

17. De Luca, P.; Galletti, A.; Marcellino, L. Parallel solvers comparison for an inverse problem in fractional calculus. *J. Earthq. Eng.* 2021, 25, 1316–1336. [CrossRef]

18. Bhosekar, A.; Ierapetritou, M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Comput. Chem. Eng.* 2018, 108, 250–267. [CrossRef]

19. Verdin, A.; Rajagopalan, B.; Kleiber, W.; Funk, C. A Bayesian kriging approach for blending satellite and ground precipitation observations. *Water Resour. Res.* 2015, 51, 908–921. [CrossRef]

20. Doan, V.; Massa, F.; Tison, T.; Naceur, H. Coupling of Homotopy Perturbation Method and Kriging surrogate model for an efficient fuzzy linear buckling analysis: Application to additively manufactured lattice structures. *Appl. Math. Model.* 2021, 97, 602–618. [CrossRef]

21. Schabenberger, O.; Gotway, C.A. Statistical Methods for Spatial Data Analysis; CRC Press: Boca Raton, FL, USA, 2017.

22. Li, Y.; Wu, Y.; Huang, Z. An incremental Kriging method for sequential optimal experimental design. *Comput. Model. Eng. Sci. (CMES)* 2014, 97, 323–357.

23. Laurenst, L.; Le Riche, R.; Soulier, B.; Boucard, P.-A. An Overview of Gradient-Enhanced Metamodels with Applications. *Arch. Comput. Methods Eng.* 2019, 26, 61–106. [CrossRef]

24. Chen, L.; Qiu, H.; Gao, L.; Jiang, C.; Yang, Z. Optimization of expensive black-box problems using Gradient-enhanced Kriging. *Comput. Methods Appl. Mech. Eng.* 2020, 362, 112861. [CrossRef]

25. Liu, W.; Batill, S. Gradient-Enhanced Response Surface Approximations Using Kriging Models. In Proceedings of the 9th International Conference on Theory and Practice in Modern Computing (TPMC 2020), Online, 23–25 July 2020; pp. 197–204.

26. Bhosekar, A.; Ierapetritou, M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. *Comput. Chem. Eng.* 2018, 108, 250–267. [CrossRef]

27. Verdin, A.; Rajagopalan, B.; Kleiber, W.; Funk, C. A Bayesian kriging approach for blending satellite and ground precipitation observations. *Water Resour. Res.* 2015, 51, 908–921. [CrossRef]

28. Doan, V.; Massa, F.; Tison, T.; Naceur, H. Coupling of Homotopy Perturbation Method and Kriging surrogate model for an efficient fuzzy linear buckling analysis: Application to additively manufactured lattice structures. *Appl. Math. Model.* 2021, 97, 602–618. [CrossRef]

29. Schabenberger, O.; Gotway, C.A. Statistical Methods for Spatial Data Analysis; CRC Press: Boca Raton, FL, USA, 2017.

30. Li, Y.; Wu, Y.; Huang, Z. An incremental Kriging method for sequential optimal experimental design. *Comput. Model. Eng. Sci. (CMES)* 2014, 97, 323–357.

31. Laurent, L.; Le Riche, R.; Soulier, B.; Boucard, P.-A. An Overview of Gradient-Enhanced Metamodels with Applications. *Arch. Comput. Methods Eng.* 2019, 26, 61–106. [CrossRef]

32. Chen, L.; Qiu, H.; Gao, L.; Jiang, C.; Yang, Z. Optimization of expensive black-box problems using Gradient-enhanced Kriging. *Comput. Methods Appl. Mech. Eng.* 2020, 362, 112861. [CrossRef]

33. Liu, W.; Batill, S. Gradient-Enhanced Response Surface Approximations Using Kriging Models. In Proceedings of the 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Atlanta, Georgia, 4–6 September 2002.

34. Ulaganathan, S.; Couckuyt, I.; Dhaene, T.; Degroote, J.; Laermans, E. High dimensional Kriging metamodelling utilising gradient information. *Appl. Math. Model.* 2016, 40, 5256–5270. [CrossRef]

35. Han, Z.-H.; Zhang, Y.; Song, C.-X.; Zhang, K.-S. Weighted Gradient-Enhanced Kriging for High-Dimensional Surrogate Modeling and Design Optimization. *AIAA J.* 2017, 55, 4330–4346. [CrossRef]

36. Chen, L.; Qiu, H.; Gao, L.; Jiang, C.; Yang, Z. A screening-based gradient-enhanced Kriging modeling method for high-dimensional problems. *Appl. Math. Model.* 2019, 69, 15–31. [CrossRef]

37. Kang, K.; Lee, I. Efficient high-dimensional metamodeling strategy using recursive decomposition coupled with sequential sampling method. *Struct. Multidiscip. Optim.* 2021, 63, 375–390. [CrossRef]
30. Bouhlel, M.A.; Bartoli, N.; Otsmane, A.; Morlier, J. An Improved Approach for Estimating the Hyperparameters of the Kriging Model for High-Dimensional Problems through the Partial Least Squares Method. *Math. Prob. Eng.* 2016, 2016, 6723410. [CrossRef]

31. Bouhlel, M.A.; Bartoli, N.; Otsmane, A.; Morlier, J. Improving kriging surrogates of high-dimensional design models by Partial Least Squares dimension reduction. *Struct. Multidiscip. Optim.* 2016, 53, 935-952. [CrossRef]

32. Bouhlel, M.A.; Martins, J. Gradient-enhanced kriging for high-dimensional problems. *Eng. Comput.* 2019, 35, 157-173. [CrossRef]

33. Zhao, L.; Wang, P.; Song, B.; Wang, X.; Dong, H. An efficient kriging modeling method for high-dimensional design problems based on maximal information coefficient. *Struct. Multidiscip. Optim.* 2020, 61, 39-57. [CrossRef]

34. Fu, C.; Wang, P.; Zhao, L.; Wang, X. A distance correlation-based Kriging modeling method for high-dimensional problems. *Knowl. Based Syst.* 2020, 206, 106356. [CrossRef]

35. Liu, B.; Zhang, Q.; Gielen, G. A Gaussian Process Surrogate Model Assisted Evolutionary Algorithm for Medium Scale Expensive Optimization Problems. *IEEE Trans. Evol. Comput.* 2014, 18, 180–192. [CrossRef]

36. Sammon, J.W. A Nonlinear Mapping for Data Structure Analysis. *IEEE Trans. Comput.* 1969, 18, 401–409. [CrossRef]

37. Zhou, Y.; Lu, Z. An enhanced Kriging surrogate modeling technique for high-dimensional problems. *Mech. Syst. Signal Process.* 2020, 140, 106687. [CrossRef]

38. Liu, B.; Xie, L. An Improved Structural Reliability Analysis Method Based on Local Approximation and Parallelization. *Mathematics* 2020, 8, 209. [CrossRef]

39. Nielsen, H.B. *Aspects of the Matlab Toolbox DACE*; The Technical University of Denmark: Kongens Lyngby, Denmark, 2002.

40. Groisser, D.; Jung, S.; Schwartzman, A. Geometric foundations for scaling-rotation statistics on symmetric positive definite matrices: Minimal smooth scaling-rotation curves in low dimensions. *Electron. J. Stat.* 2017, 11, 1092-1159. [CrossRef]

41. Entezami, A.; Sarmadi, H.; Behkamal, B.; Mariani, S. Health Monitoring of Large-Scale Civil Structures: An Approach Based on Data Partitioning and Classical Multidimensional Scaling. *Sensors* 2021, 21, 1646. [CrossRef] [PubMed]

42. Erez, P.; Matan, G. Multidimensional scaling of noisy high dimensional data. *Appl. Comput. Harmon. Anal.* 2021, 51, 333–373.

43. Gintautas, D.; Martynas, S. Geometric multidimensional scaling: A new approach for data dimensionality reduction. *Appl. Math. Comput.* 2020, 409, 125561.

44. Vehtari, A.; Gelman, A.; Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 2017, 27, 1413–1432. [CrossRef]

45. Saussié, D.; Saydy, L.; Akhrif, O. Longitudinal flight control design with handling quality requirements. *Aeronaut. J.* 2006, 110, 627–637. [CrossRef]

46. Bahri, S. Longitudinal flight control laws for high aspect ratio light utility aircraft. *J. Phys. Conf. Ser.* 2018, 1130, 012026. [CrossRef]