High Anisotropy Space Exploration with Co-Kriging Method

Zebin Zhang[1,a),*], Zhonghua Han[2,†] and Pascal Ferrand[3,**]

[1]*Zhengzhou University, 100, Kexue Av., 450001 Zhengzhou, P.R.China*
[2]*Northwestern Polytechnical University, 710072 Xi'an, P.R.China.*
[3]*Ecole Centrale Lyon, 36, Av. Guy de Collongue, Ecully, France 69134*

[a)]Corresponding author: zebin.zhang@zzu.edu.cn

Abstract. For "black box" data-space exploration, classical DoE method prescribes a cluster of isotropic design points following a certain space-infill criterion, whatever the objective function behaves. Adding new point often disturbs the original spatial-infill properties, which is not practical for dimension-augmentation and reusability of existing data. Surrogate model based sequential design considers not only the spatial features of the design space but also the characteristic of the objective function. This latter may favor a highly anisotropic design for the sake of computational effort effectiveness, especially while high informative samples are used. This work presents a high anisotropy design solution when derivative-assisted co-Kriging method is employed. Adaptation of this method has been proposed for the special case while only one single sample appears at certain dimensions. A 2-dimensional test function with distinctive spatial properties at each dimension is composed and used to assess the validity of the proposed method.

INTRODUCTION

With the improvement of measurement facilities such as high-performance computing, data collected become more and more complex and informative, characterized by increasing dimensionality and larger sample size [1], seriously challenging our ability to keep pace with the need to precisely model the systems we seek to design. In the design optimization, it is a common practice to extract design information through a data modeling process, so called surrogate model or meta-model. Design of Experiments (DoE) is often performed *a priori* of the modeling. Observations are considered following an unknown but modeled functional pattern [2]. To better capture the dominating features of the design space, a well-chosen sample set is critical. However, classical DoE methods consider only the spatial properties of the design variables, the characteristic of the objective function itself has not been taken into account, resulting in an isotropic design where the difference of sensitivities at different dimensions should be but not reflected. What's more, the number of points is fixed once the design is done, in most cases, adding or removing design point would disturbs the original space-infill criterion [3]. The surrogate model based sequential design shows a good compromise between the design space spatial properties and the objective function characteristics. The DoE obtained is not an "one shot" design but a continuously updating one by taking into account the property of objective function at different dimensions. The number of points of the design site is not fixed *a priori*. Instead, a convergence criterion is employed to complete the design. With its inherent benefit of uncertainty prediction, Kriging method and its variant "co-Kriging" method [4, 5], assisted by design site derivatives, are practical sequential design approaches for the exploration of data-space featuring anisotropy properties. For example, one can replace the scalar correlation parameters with a vector of n elements where n is the dimension of the design space [6]. This study particularly draws attention to the modeling of high anisotropic design space, of which the contribution is twofold:

- Integration of derivative information with co-Kriging surrogate model [7]. The sensitivity study manifests a great potential for the modeling of a complex system. One example is the implementation of adjoint method in the large-scale aerodynamic shape optimization [8]. Hessian enhanced co-Kriging method has also been

[*]Lecturer, School of Mechanical Engineering of Zhengzhou University
[†]Professor, National Key Laboratory of Science and Technology on Aerodynamic Design and Research at Northwestern Polytechnical University
[**]Director of research at Laboratoir de Méchanique des Fluides et d'Acoustique (LMFA) of Ecole Centrale Lyon

studied [9]. Comparisons are given in this paper to show the improvements brought by the first and second order derivatives.

- Adaptation of co-Kriging method for a specific case while only one sample point presents at some dimension. Derivative assisted design points are more informative than the derivative free ones, the high is the order of derivatives, the more information it contains for the usage of modeling. Chances are only one single point needed for certain dimensions. Without variation in those dimension, the classical co-Kriging method cannot be applied directly. Adjustment to the method has been proposed in this study. An additional advantage of this "adaptation" resides in the dimensional expansibility of the design space and the reusability of an existing database.

DERIVATIVE-ENHANCED CO-KRIGING METHOD

The Kriging method modelizes an unknown function through a stochastic process, represented by a mean function and a covariance function. It is often the best linear unbiased prediction from some given evaluations, i.e., the prediction error is minimized. Two sub-models are needed for this stochastic process:

- A regression model which can be considered as the mean function of all the possible prediction functions subject to the existing evaluations.
- A correlation model which reflects the spatial correlations between the points in the design space. Gaussian model is a most-frequently used model which assumes *a priori* the Weakly Sense Stationary (WSS) property of the design site.

For the regression model, many works presume a constant average of the Kriging process [10, 9]. For a co-Kriging process, the differentiation of the regression model is required, once differentiated, a constant regression model cannot reflect the differences at different design points. L.Zhao has assessed different regression models such as Hermite polynomial, trigonometry function and exponential function and so on [11]. According to his research, the polynomial function is recommended for the universal Kriging and is therefore taken in this study. The regression matrix for a co-Kriging model consists of a regression part and a differentiated part which takes the following form:

$$\mathcal{F}(x) = \begin{bmatrix} \mathcal{F}_0 \\ d(\mathcal{F}) \\ d^2(\mathcal{F}) \\ \dots \end{bmatrix} \tag{1}$$

where $d(\mathcal{F})$ and $d^2(\mathcal{F})$ are respectively the first order and second order derivatives of the regression matrix $\mathcal{F}_0$. Clearly from this representation, if the order of regression model is less than that of the derivatives, the differentiation parts of $\mathcal{F}(x)$ could be all zeros. In this case, the "enhancing" effect of the high order derivatives will not be fully extracted.

The regression function is a mean path of a random process, characterized by an *a priori* chosen covariogram. The interpolation properties primarily depend on the local behavior of the random field. Near to the origin, the Gaussian correlation model shows a quadratic behavior, fitting the data with second order derivatives. For the sake of derivability, Gaussian model is chosen for this study.

The correlation model can be established by evaluating a "hyper-parameter" namely θ. One single θ value could be used all around the design space, where the correlation depends only on the Euclidean distance, this is called isotropic covariogram [12]. However, since the nature of the random field usually differs at different direction, an isotropic model can be inappropriate [7]. The anisotropic model distinguishes the correlation characteristics from one dimension to another by choosing different θ value for different design variables, which yields to a vectorial parameter θ. This latter is applied for the co-Kriging modeling process in this study.

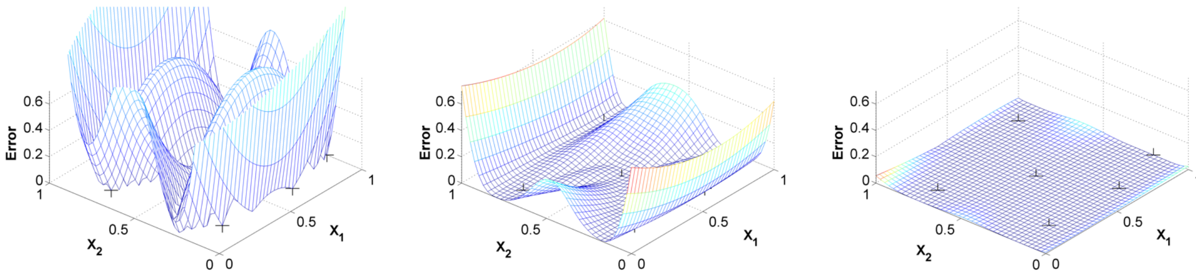
A 2-dimension function is composed to illustrate the derivative-enhanced effect and the dimensional difference in term of complexity and sensitivity:

$$f_{test}(x_1, x_2) = 2 + 5(x_1 - 0.5)^2 + 3x_1x_2 - sin(10x_2) - exp(\frac{x_2}{2}) \tag{2}$$

At the dimension of x_1, the function behaves as a second order polynomial, the x_2 term constitutes of a sinusoidal and an exponential function which results in a function of order of infinite. Interaction effect is added with a cross term

“$3x_1x_2$”. The remarkable difference at 2 dimensions makes this function a fitted choice for the study of the dimension sensitivity in the next section. This function serves as a reference to assess the quality of different models.

A Latin Hypercube Sampling (LHS) design of 6 points has been used to build the Kriging, first order co-Kriging and second order co-Kriging surrogate models. Derivatives at those points are added for co-Kriging models. Absolute error between each model and the reference function is illustrated in Fig. 1.



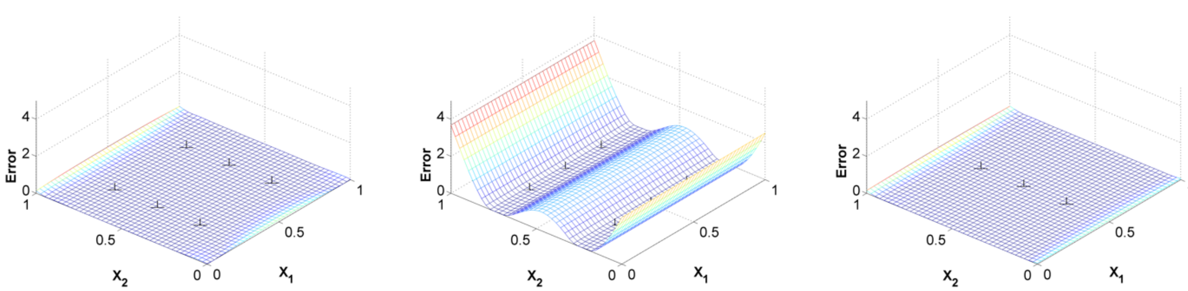
(a) Kriging (b) 1st order co-Kriging (c) 2nd order co-Kriging

FIGURE 1: Absolute error comparison between Kriging, first order co-Kriging and second order co-Kriging models

Figure. 1 (a) shows the absolute error of the prediction by Kriging model. Clearly from the figure, error is very important, especially at the zone of extrapolation, i.e., outside of the range of sampling points. The improvement brought by the first order derivatives is significant, maximum error is not more than 0.7 as shown in Fig. 1 (b). When second order derivatives are employed, the maximum error goes as low as 0.05, illustrated in Fig. 1 (c). Provided the derivatives are easy to obtain, the co-Kriging method can significantly improve the precision of the model economically.

FUNCTION DIMENSIONAL SENSITIVITY STUDY AND ADAPTATION OF CO-KRIGING FOR DIMENSION WITH UNIQUE VALUE

To illustrate the dimensional dependence of the modeling, 6 design points with first and second order derivatives, distributed in 2 different patterns: (a). 2 values (0.25 and 0.75) for x_1, 3 values (0.25, 0.5 and 0.75) for x_2 ; (b). 3 values (0.25, 0.5 and 0.75) for x_1, 2 values (0.25 and 0.75) for x_2, are used to model the reference function with second order co-Kriging method. The hyper-parameter θ obtained are respectively: $[0.10 \quad 0.44]^T$, $[0.10 \quad 29.92]^T$ and $[0.10 \quad 0.52]^T$. The different θ values at 2 dimensions are related to the anisotropic covariogram in the random field, especially for case (b), the θ at dimension 2 appears near to the extreme of the interval which is $[0 \quad 30]^T$. This indicates that the maximum likelihood problem is not well resolved or more samples are need at the corresponding dimension. Absolute errors between each model and the reference function are shown in Fig. 2 (a) and Fig. 2 (b).



(a) (b) (c)

FIGURE 2: Absolute error comparison between 2×3, 3×2 and 1×3 samplings.

The precision of the model differs from each other due to the distribution of sampling points. With the same computational cost, sampling shown in Fig.2 (a) gives obviously more useful information for the design space exploration. This significant difference originates from the difference in term of function complexities for 2 dimensions. At x_1, the function is practically quadratic while a polynomial of much higher order presents at x_2. Less points are necessary for dimension 1 than dimension 2.

Since the second order derivatives are used, one single point should be adequate to capture the quadratic characteristics at dimension x_1 for a given value of x_2, i.e., the unique value for x_1 is allowed theoretically. However, being a stochastic method, the co-Kriging requires a variation at every dimension in order to calculate the corresponding spatial variance σ^2, which is one of the key factors of the modeling process. To cope with this paradox, adaptation of co-Kriging method has been proposed.

The variance is used in the co-Kriging modeling process to calculate the correlation, to resolve the Kriging equation system [6] and to normalize the input data (Eq. 3).

For the design variables S and the responses Y, the normalized terms $\tilde{S}$ and $\tilde{Y}$ are:

$$\tilde{S} = \frac{S - \bar{S}}{\sigma_S}, \quad and \quad \tilde{Y} = \frac{Y - \bar{Y}}{\sigma_Y} \tag{3}$$

Without variation, the σ calculated returns zero, presenting at the denominator of the normalization process and also at the denominator of the Lagrangian term in the Kriging equation system, causing some infinite values. The modeling process cannot proceed in this case.

For such case a unique value presenting at some dimension x_i, the authors propose to fix the correlation function by assuming a variance σ^2 equal to 1 at the target dimension x_i so that the random part $z(x_i)$ of the model subjects to the normal distribution $\mathcal{N}(0, 1)$. This can be seen as a second hypothesis of the co-Kriging process, which is natural because data at all dimensions have been normalized so that the errors at design site subject to the normal distribution. This hypothesis respects the originality of the modeling process and has been validated by different tests [13]. According to the tests, at the examined dimensions where unique value presents, the co-Kriging model turns out to be the Taylor expansion of the reference function at the reference point, order of which depending on the order of derivatives used for the modeling.

The adaptation is used to model the reference function (Eq. 2). One single value 0.5 has been taken for the quadratic dimension x_1, and 3 values for x_2. The error is shown in Fig. 2 (c). Comparing to the model obtained with 6 design points, the precision is almost the same but the number of points used is 50% less. The function (Eq. 2) and the different design sites are intentionally composed to show the anisotropic effects but functions of less pronounced dimensional distinctions can be modeled more easily with the methodology. An industrial implementation of this methodology can be found at the thesis work of Zhang [13].

CONCLUSION

It has been observed that the anisotropic sampling should be used for design space exploration. The surrogate model based sequential design is a logical choice for this kind of sampling design. The co-Kriging model based sequential design can make the best of derivatives to improve the quality of the model. Among all the anisotropic designs, the dimension with one unique value has been studied specifically. The classic co-Kriging failed to interpolate this kind of sampling points. Adaptation of the method has been proposed in this paper. The result model shows an accurate Taylor style extrapolation at the questioned dimension. This study is meaningful for the space dimension augmentation and reusability of data, especially when high-informative data are involved.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the technical support from all the staffs of Fluorem S.A.S.

REFERENCES

[1] R. Giraldo and S. Dabo-Niang, Statistics and Probability Letters **In Press** (2018).

[2] J. Ramsay and B. Silverman, *Functional Data Analysis* (Springer Science+Business Media, Inc, 233 Spring St., New York, NY 10013, USA, 2005), pp. 37–57.

[3] Z.Zhang and B.Demory, "Space infill study of kriging meta-model for multi-objective optimization of an engine cooling fan," (Proceedings of ASME Turbo Expo, 2014).
[4] D. Krige, "statistical approach to some mine valuations and allied problems at the witwatersrand," Master's thesis, University of Witwatersrand 1951.
[5] G. Matheron, Economic Geology **58**, 1246–1266 (1963).
[6] S. N. Lophaven and H. B. Nielsen, Dace a matlab kriging toolbox, , technical Report IMM-TR-2002-12.
[7] J. Villemonteix and Vazquez, Journal of Global Optimization **44(4)**, 509–534 (2008).
[8] Z. Han and Zimmermann, AIAA Journal **50(5)**, 1205–1210 (2012).
[9] W. Yamazaki and D. J. Mavriplis, AIAA Journal **51(1)**, 126–137 (2013).
[10] A. March and K. Willcox, AIAA Journal **50(5)**, 1079–1089 (2012).
[11] L. Zhao and K. Choi, AIAA Journal **49(9)**, 2034–2046 (2011).
[12] S. N. Lophaven and H. B. Nielsen, Aspects of the matlab toolbox dace, , technical Report IMM-TR-2002-13.
[13] Z. Zhang, "Intégration des méthodes de sensibilité d'ordre élevé dans un processus de conception optimale des turbomachines. développement de méta-modèles." Ph.D. thesis, Ecole Centrale de Lyon 2014.