

Сравнительный анализ методов аппроксимационного моделирования при наличии выборок разной точности

Н. В. Капкаев^{1,2}

¹ ФГАОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского», г. Нижний Новгород, Российская Федерация

Адрес: 603022, Российская Федерация, г. Нижний Новгород, пр. Гагарина, д. 23

² ООО «Сириус», г. Нижний Новгород, Российская Федерация

Адрес: 603034, Российская Федерация, г. Нижний Новгород, ул. Кисловодская, д. 1/6
kapkaevn@gmail.com

Аннотация

В процессе технологического развития и наращивания вычислительных мощностей качество математических, имитационных моделей увеличилось, появилась возможность реализации сложных многопараметрических систем. Однако, несмотря на высокую точность получаемых результатов, важным является количество времени, требуемое для вычислений. Например, при решении задач оптимизации необходимо производить множество вычислений значений функции, и в случае когда подсчет одного значения может занимать несколько часов, провести поиск оптимального набора параметров не представляется возможным (за приемлемое время, меньшее чем недели и месяцы). В таких ситуациях прибегают к использованию аппроксимационных (суррогатных) моделей (surrogate model), которые ускоряют процесс получения значения функции. Целью данной статьи является представление обзора по существующим методам построения аппроксимационных моделей по одной и двум разноточным выборкам, а также предлагается новый метод для аппроксимации по разноточным выборкам с использованием уточнения при помощи градиента или его оценки, для MF моделей семейства kriging. Сформулирована и доказана теорема о виде аппроксимирующей функции при наличии разноточных выборок и знаний о градиенте на базе модели семейства kriging, продемонстрировано применение на тестовых функциях. Для построения аппроксимации по одной выборке рассматриваются модели PRS (Polynomial Response Surface), IDW (Inverse Distance Weighting), RBF (Radial Basis Function), Kriging. Для случая двух разноточных выборок MFMS (Multi-Fidelity Multiplicative Surrogate), MFAS (Multi-Fidelity Additive Surrogate), MFHS (Multi-Fidelity Hybrid Surrogate), Co-kriging модели, MFG (Multi-Fidelity Gradient) модель.

Ключевые слова: суррогатное моделирование, разноточные выборки, kriging, регрессия гауссовских процессов

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Для цитирования: Капкаев Н. В. Сравнительный анализ методов аппроксимационного моделирования при наличии выборок разной точности // Современные информационные технологии и ИТ-образование. 2024. Т. 20, № 3. С. 554-562. <https://doi.org/10.25559/SITITO.020.202403.554-562>

© Капкаев Н. В., 2024



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.



Comparative Analysis of Surrogate Modeling Methods in the Presence of Multi-Precision Samples

N. V. Kapkaev^{a,b}

^a National Research Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russian Federation

Address: 23 Gagarin Ave., Nizhny Novgorod 603022, Russian Federation

^b Sirius LLC, Nizhny Novgorod, Russian Federation

Address: 1/6 Kislovodskaya St., Nizhny Novgorod 603034, Russian Federation

kapkaevn@gmail.com

Abstract

In the process of technological development and increasing computing power, the quality of mathematical and simulation models has increased, and it has become possible to implement complex multiparametric systems. However, despite the high accuracy of the results obtained, the amount of time required for calculations is important. For example, when solving optimization problems, it is necessary to perform many calculations of function values, and in the case when calculating one value may take several hours, it is not possible to search for the optimal set of parameters (in an acceptable time less than weeks and months). In such situations, they resort to using approximation (surrogate) models, which speed up the process of obtaining the value of the function.

The purpose of this article is to provide an overview of existing methods for constructing approximation models for one and two multi-point samples, and also a new method for approximation for multi-point samples using gradient refinement or its estimation for MF models of the kriging family is proposed. A theorem on the form of an approximating function is formulated and proved in the presence of multi-precision samples and knowledge about the gradient based on the kriging model, and application to test functions is demonstrated. To build an approximation for a single sample, the models PRS (Polynomial Response Surface), IDW (Inverse Distance Weighing), RBF (Radial Basis Function), Kriging are considered. For the case of two different-precision samples, MMS (Multi-Fidelity Multiplicative Surrogate), MFA (Multi-Fidelity Additive Surrogate), MFHS (Multi-Fidelity Hybrid Surrogate), Co-kriging models, MFG (Multi-Fidelity Gradient) model.

Keywords: surrogate modeling, multi-fidelity, kriging, regression of Gaussian processes

Conflict of interests: The author declares no conflict of interests.

For citation: Kapkaev N.V. Comparative Analysis of Surrogate Modeling Methods in the Presence of Multi-Precision Samples. *Modern Information Technologies and IT-Education*. 2024;20(3):554-562. <https://doi.org/10.25559/SITITO.020.202403.554-562>



Введение

Все современные естественные и общественные науки используют для решения своих задач математический аппарат и необходимые математические модели. Однако, составленная математическая модель представляет собой некую идеализацию реальной сложной системы. Для математической модели важной составляющей является адекватность и согласованность с реальным объектом, процессом или явлением. С развитием компьютерных технологий идеализированные модели уточнялись дополнительными сведениями об исследуемой системе, появилась возможность реализации усложненных моделей реальных систем, которые обладают большой точностью, и позволяют в некоторых случаях заменять дорогостоящие натурные эксперименты [1-4]. Главным недостатком, полученной высокоточной модели, является то, что ей требуется большое количество времени и вычислительных мощностей на предоставление необходимых результатов – от нескольких часов до нескольких суток на мощных суперкомпьютерах [5]. Таким образом, в силу ограниченности временных, вычислительных и материальных ресурсов, выборка, получаемая с высокоточной модели или в результате натурного эксперимента имеет достаточно небольшой размер (здесь под выборкой подразумевается некоторый набор пар (X_i, Y_i) $i = 1, \dots, k$, $X_i \in R^n$, $Y_i \in R$, где X_i – входной набор параметров, а Y_i – отклик на заданный входной набор X_i – будем называть такой исходный набор тренировочной (обучающей) выборкой или тренировочным вектором). В связи с этим возникает идея построения аппроксимационной (суррогатной модели). Особенно это актуально для задач оптимизации сложных многопараметрических систем, где применение классических методов влечет огромные затраты по времени, в силу необходимости вычисления значений функционала и его градиента в большом количестве точек. Эта тема, так называемой суррогатной оптимизации широко развита в англоязычной литературе [6-10]. Отдельно можно выделить суррогатное моделирование по разноточным выборкам [11-15], когда имеется высокоточный набор данных (долгое время получения) и набор данных более низкой точности, который может быть получен в гораздо большем размере с упрощенной модели с меньшими вычислительными затратами. При этом упрощенная модель также будет отражать основные тенденции исследуемой системы. Цель данной статьи заключается в представлении обзора по основным существующим суррогатным моделям (PRS, IDW, RBF, семейство Kriging), а также в представлении нового метода для аппроксимации по разноточным данным (Multi-Fidelity) с учетом информации о градиенте.

Одновыборочные модели

В данном разделе представим обзор некоторых классических аппроксимационных моделей, построенных по одной выборке. 1) PRS – модель (Polynomial Response Surface) [16, 17]. Это регрессионная полиномиальная модель общий вид которой можно представить следующим образом:

$$f(x) = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{j=1}^n \sum_{j \geq i} \beta_j x_i x_j + \dots + \varepsilon \quad (1)$$

где ε – статистическая ошибка, $\beta_0, \beta_i, \beta_j$ – коэффициенты, которые необходимо оценить для построения модели.

Несложно заметить, что для поверхности второго порядка необходимо оценить $(n+1)(n+2)/2$ параметров, а значит для построения качественной модели количество точек будет пропорционально $O(n^2)$.

Для удобства оценки необходимых коэффициентов формула (1) может быть представлена в обобщенном матричном виде:

$$f(x^o) = \tilde{X}_o \cdot \beta + \varepsilon, \quad (k = 1, 2, \dots, n)$$

где $\tilde{X}_o$ – это вектор, образованный компонентами $x^o = (x_1^o, x_2^o, \dots, x_n^o)$ в порядке соответствующем коэффициентам вектора β . Опуская индекс o -ой точки, получаем, что общий вид будет следующим: $\tilde{X} = (1, x_1, \dots, x_n, x_1^2, x_1 x_2, \dots, x_{n-1} x_n, x_n^2)^T \in R^{(n+1)(n+2)/2}$. По имеющимся данным, и исходя из точной модели $Y = X \cdot \beta$, где $Y^T = (y_1, \dots, y_{(n+1)(n+2)/2})$, $X^T = (\tilde{X}_1^T, \dots, \tilde{X}_{(n+1)(n+2)/2}^T)$, используя оценки наименьших квадратов вектор оцененных коэффициентов $\hat{\beta}$ будет иметь вид:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \quad (2)$$

Подставляя выражение (2) в (1), получим вид модели PRS, коэффициенты которой оценены по известному тренировочному набору точек (x^o, y^o) , $(o = 1, 2, \dots, k)$, $x^o = (x_1^o, x_2^o, \dots, x_n^o)$:

$$\hat{f}(x) = \hat{\beta}_0 + \sum_{i=1}^n \hat{\beta}_i x_i + \sum_{i=1}^n \sum_{j \geq i}^n \hat{\beta}_j x_i x_j + \dots$$

2) IDW – модель (Inverse Distance Weighting) [18]. Это интерполяционный метод, использующий в качестве весовой функции обратные расстояния между точками. Выражение для вычисления значения функции можно представить следующим образом:

$$f(x) = \begin{cases} \frac{\sum_{i=1}^k \beta(x, x^{(i)}) \cdot y^{(i)}}{\sum_{i=1}^k \beta(x, x^{(i)})}, & \text{если } x \neq x^{(i)} \\ y^{(i)}, & \text{если } x = x^{(i)} \end{cases}$$

где $x^{(i)} \in R^n$; $y^{(i)} \in R$ – входной тренировочный вектор и точное значение в i -ой тренировочной (обучающей) точке, k – количество точек тренировочной выборки. В простых случаях, могут использоваться все обучающие точки, однако в многомерных пространствах может оказаться более эффективным использование подвыборки (например, k – nearest neighbors). Весовая функция β определяется как:

$$\beta(x^{(i)}, x^{(j)}) = \|x^{(i)} - x^{(j)}\|_2^{-s},$$

и для непрерывности производных значение параметра s должно быть больше единицы.

3) RBF – модель (Radial Basis Function) [19]. Данная аппроксимационная модель основана на выражении искомой зависимости в виде линейной комбинации радиальных базисных функций. В качестве радиальной базисной функции может выступать любая вещественная функция, значения которой вычисляются от передаваемого в виде аргумента расстояния между точками. Выделяется несколько классических типов



используемых базисных функций ($r = \|x - x_i\|$):

- Гауссова функция: $\phi = \exp(-(\varepsilon r)^2)$ $\varepsilon > 0$
- Мультикватричная: $\phi(r) = \sqrt{1 + (\varepsilon r)^2}$
- Обратная квадратичная: $\phi(r) = \frac{1}{1 + (\varepsilon r)^2}$
- Обратная мультикватричная: $\phi(r) = \frac{1}{\sqrt{1 + (\varepsilon r)^2}}$
- Полигармонический сплайн: $\phi(r) = \begin{cases} r^s, & s = 1, 3, 5, \dots \\ r^s \ln(r), & s = 2, 4, 6, \dots \end{cases}$

Кроме этого модель RBF может быть дополнена глобальными полиномами, чтобы отразить общие тенденции. Тогда формула для вычислений примет вид:

$$f(x) = p(x) \cdot w_p + \sum_{i=1}^k w_i \phi(\|x - x_i\|),$$

где $p(x)$ – глобальный полином; $\phi(\|x - x_i\|)$ – базисная функция, принимающая на вход расстояние между точками; x_i – i -ая из k тренировочных точек; w_i, w_p – весовые коэффициенты для базисных и глобальной функций соответственно.

Для составления модели RBF необходимо найти вектор коэффициентов w , который может быть получен с помощью метода наименьших квадратов. Пусть $f(x_i) = y_i$, тогда получим:

$$\begin{pmatrix} \phi(\|x_1 - x_1\|) \cdots \phi(\|x_n - x_1\|) & p(x_1) \\ \vdots & \vdots \\ \phi(\|x_1 - x_n\|) \cdots \phi(\|x_n - x_n\|) & p(x_n) \\ p(x_1) & \cdots & p(x_n) & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \\ w_p \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ 0 \end{pmatrix}.$$

4) Kriging – модель.

Kriging – это классический подход используемый в геостатистике¹ [20] и основанный на регрессии гауссовских процессов. В общем случае любая модель кригинга может быть представлена в виде (3)

$$f(x) = m(x) + Z(x) \quad (3)$$

где $Z(x)$ – стационарный гауссовский процесс, а $m(x)$ – тренд. Выделяется три основных типа кригинга:

- Простой кригинг. Предполагается, что $m(x) = m = \text{Const}$ и данное значение известно во всей исследуемой области.
- Ординарный кригинг. Значение $m(x)$ в области неизвестно, однако является постоянным в исследуемой области, поэтому строится оценка $m(x)$.
- Универсальный кригинг. В данном случае предполагается, что тренд $m(x)$ является гладкой функцией на исследуемой области, обычно полиномом или комбинацией известных функций.

Поскольку в задаче может отсутствовать априорная информация о константном значении $m(x)$ на всем пространстве оценивания, а также о возможности выражения $m(x)$ через

гладкую функцию, то остается только один способ построения – это ординарный кригинг. Тогда в формуле (3) $m(x) = m$ – подлежащая оценке константа, а $Z(x)$ – стационарный гауссовский процесс для которого будем считать, что заданы нулевое математическое ожидание $E[Z(x)] = 0$ и ковариационная функция $\text{Cov}(Z(x), Z(x')) = \sigma^2 r(x, x') = \sigma^2 r_{xx'}$, где σ^2 – дисперсия гауссовского процесса, $r(x, x') = r_{xx'}$ – корреляционная функция для точек x и x' , называемая также ядром. По исходному тренировочному набору (начальная выборка) $X = (X_1, \dots, X_k)$, $Y = (Y_1, \dots, Y_k)^T$, $X_i \in R^n$, $Y_i = f(X_i) \in R^1$ составляется корреляционная матрица $R_\Theta = (r_{x^{(1)}x}, \dots, r_{x^{(k)}x})$, где $r_{xx'} = (r_{xx^{(1)}}, \dots, r_{xx^{(k)}})^T$ – корреляционный вектор, но для составления необходимо выбрать тип корреляционной функции. Чаще всего используется гауссова корреляционная функция $r(x, x') = \exp(-\sum_{i=1}^n \theta_i (x_i - x'_i)^2)$ [21].

Корреляционная функция зависит от вектора параметров $\Theta = (\theta_1, \dots, \theta_n)$. Таким образом, для построения предикативной модели нужно оценить значения параметров m, σ^2, Θ . Используем для оценки метод максимального правдоподобия:

$$L(Y | m, \sigma^2, \Theta) = \frac{1}{(2\pi\sigma^2)^{k/2} |R_\Theta|^{1/2}} \cdot \exp\left(-\frac{(Y - m \cdot I)^T R_\Theta^{-1} (Y - m \cdot I)}{2\sigma^2}\right),$$

где m, σ^2, Θ – среднее, дисперсия и параметры ядра, соответственно; $I = (1, \dots, 1)^T$ – единичный вектор размерности Y ; $|R_\Theta|$ – определитель матрицы R_Θ .

Учитывая, что $L(Y | m, \sigma^2, \Theta)$ – это функционал который необходимо максимизировать и $L(Y | m, \sigma^2, \Theta) > 0$, обычно используется логарифм оценки максимального правдоподобия:

$$l = \log L = -\frac{k}{2} \log(2\pi\sigma^2) - \frac{1}{2} \log |R_\Theta| - \frac{(Y - mI)^T R_\Theta^{-1} (Y - mI)}{2\sigma^2}$$

Чтобы найти максимум приравняем к нулю частные производные и выразим оценки m, σ^2, θ_i .

$$\begin{cases} \frac{\partial l}{\partial m} = 0 \\ \frac{\partial l}{\partial(\sigma^2)} = 0 \Rightarrow \\ \frac{\partial l}{\partial \theta_i} = 0 \end{cases} \Rightarrow \begin{cases} m = \frac{I^T R_\Theta^{-1} Y}{I^T R_\Theta^{-1} I} \\ \sigma^2 = \frac{(Y - mI)^T R_\Theta^{-1} (Y - mI)}{k} \\ \frac{k}{2} \cdot \frac{(Y - mI)^T (R_\Theta^{-1} \frac{\partial R_\Theta}{\partial \theta_i} R_\Theta^{-1}) (Y - mI)}{(Y - mI)^T R_\Theta^{-1} (Y - mI)} - \frac{1}{2} \text{tr}(R_\Theta^{-1} \frac{\partial R_\Theta}{\partial \theta_i}) = 0 \end{cases}$$

После того как все необходимые параметры оценены можно построить предикативную формулу для вычисления значений, а также для оценки дисперсии в неизвестных точках:

$$\tilde{f}(x) = \hat{m} + r_{xx}^T R_\Theta^{-1} (Y - mI); \quad (4)$$

$$s^2(x) = \hat{\sigma}^2 (1 - r_{xx}^T R_\Theta^{-1} r_{xx}). \quad (5)$$

Подробный вывод формул (4), (5) можно найти в [22].

¹ Krige D. G. A statistical approach to some mine valuation and allied problems on the Witwatersrand : Thesis M.Sc.(Engineering). University of the Witwatersrand, Faculty of Engineering, 1951. 52 p. [Electronic resource]. URL: <https://wiredspace.wits.ac.za/items/ae034a42-dd51-44c9-b405-d14a37f76472> (дата обращения: 12.07.2024).



Двухвыборочные модели

Пусть $X_{LF} = \{x_{1,l}, \dots, x_{v,l}\}$, $f_{LF} = \{f_{1,l}, \dots, f_{v,l}\}$ и $X_{HF} = \{x_{1,h}, \dots, x_{k,h}\}$ – это выборки LF (*Low-Fidelity*) и HF (*High-Fidelity*) моделей, соответственно. При этом будем полагать, что $X_{HF} \subset X_{LF}$, поскольку выборка для LF модели может быть отобрана с учетом имеющихся высокоточных данных, или легко дополнена недостающими элементами.

Рассмотрим в данном пункте четыре базовые модели, которые, концептуально, лежат в основе различных MF методов [23].

1) Мультипликативная модель [24].

В данном случае строится MFMS (*Multi-Fidelity Multiplicative Surrogate*) модель, которая может быть представлена следующим образом:

$$\hat{f}_{MF}(x) = \hat{\beta}(x) \cdot \hat{f}_{LF}(x)$$

где $\hat{\beta}(x)$ – это аппроксимационная модель весового коэффициента, $\hat{f}_{LF}(x)$ – аппроксимационная LF модель или известная LF функция. Аппроксимация $\hat{\beta}(x)$ строится по набору точек $B = \{(x_{1,h}, \beta(x_{1,h})), \dots, (x_{k,h}, \beta(x_{k,h}))\}$,

где $\beta(x_{i,h}) = \frac{f_{HF}(x_{i,h})}{f_{LF}(x_{i,h})}$, для X_{HF} точек.

2) Аддитивная модель [25].

MFAS (*Multi-Fidelity Additive Surrogate*) модель получается после добавления к LF модели аддитивного суррогата:

$$\hat{f}_{MF} = \hat{f}_{LF} + \hat{\gamma}(x)$$

где $\hat{\gamma}(x)$ – аппроксимационная модель, построенная по набору $\Gamma = \{(x_{1,h}, \gamma(x_{1,h})), \dots, (x_{k,h}, \gamma(x_{k,h}))\}$, и значениях в X_{HF} точках $\gamma(x_{i,h}) = f_{HF}(x_{i,h}) - f_{LF}(x_{i,h})$.

3) Гибридная модель [26].

В MFHS (*Multi-Fidelity Hybrid Surrogate*) моделях вводится дополнительный весовой коэффициент с помощью которого строится взвешенная модель на основе MFMS и MFAS:

$$\hat{f}_{MFHS} = w \cdot (\hat{f}_{LF}(x) \cdot \hat{\beta}(x)) + (1 - w) \cdot (\hat{f}_{LF} + \hat{\gamma}(x))$$

Чаще всего параметр w подбирается эмпирически в качестве константы, с опорой на уже имеющиеся данные, однако в некоторых случаях может представлять собой и функцию при наличии соответствующей дополнительной информации [26].

4) Co-kriging модель [27].

Данный метод основан на подходе кригинга, поэтому общая структура аппроксимационной модели будет соответствовать (3), однако, метод имеет свои особенности, поскольку учитывает работу с двумя выборками разной точности.

Зададим два вектора: $X = (X_{LF}, X_{HF})^T$, $Y = (Y_{LF}(X_{LF}), Y_{HF}(X_{HF}))^T$. Как и в случае с кригингом значение в точке обрабатывается так, если бы оно было реализацией стохастического процесса. Будем считать, что HF и LF модели связаны следующим образом:

$$Z_{HF}(x) = \rho \cdot Z_{LF}(x) + Z_d(x),$$

где $Z_{HF}(x)$, $Z_{LF}(x)$ – стационарные гауссовские процессы на основе HF и LF данных соответственно, ρ – постоянный коэффициент, а Z_d – стационарный гауссовский процесс, построен-

ный по набору $(X_{HF}, Y_{HF}(X_{HF}) - \rho \cdot Y_{LF}(X_{HF}))$. Для построения аппроксимации необходимо задать ковариационную матрицу, которая в случае с Co-kriging примет вид:

$$C = \begin{pmatrix} \sigma_{LF}^2 R_{LF}(X_{LF}, X_{LF}) & \rho \sigma_{LF}^2 R_{LF}(X_{LF}, X_{HF}) \\ \rho \sigma_{LF}^2 R_{LF}(X_{HF}, X_{LF}) & \rho^2 \sigma_{LF}^2 R_{LF}(X_{HF}, X_{HF}) + \sigma_d^2 R_d(X_{HF}, X_{HF}) \end{pmatrix}.$$

Здесь R_{LF} и R_d – это корреляционные матрицы, которые могут быть построены как по одному ядру, так и с использованием разных. Для оценки параметров m_{LF} , σ_{LF}^2 , θ_{LF} , m_d , σ_d^2 , θ_d , ρ будут строиться два функционала по методу максимального правдоподобия [28].

В результате будет получена MFC (*Multi-Fidelity Co-kriging*) модель:

$$\hat{f}_{MFC}(x) = \hat{m} + c^T C^{-1} (Y - mI),$$

$$\text{где } c = (\hat{\rho} \hat{\sigma}_{LF}^2 r_{LF}(X_{LF}, x), \hat{\rho}^2 \hat{\sigma}_{LF}^2 r_{LF}(X_{HF}, x) + \hat{\sigma}_d^2 r_d(X_{HF}, x))^T,$$

$$\hat{m} = \frac{I^T C^{-1} Y}{I^T C^{-1} I}, \text{ а оценка выборочной дисперсии будет выра-}$$

$$\text{жаться как: } \hat{s}^2(x) = \hat{\rho}^2 \hat{\sigma}_{LF}^2 + \hat{\sigma}_d^2 - c^T C^{-1} c.$$

В данной статье рассмотрим новый метод, который будет добавлять к построенной MF-модели, являющейся трендом, уточнение с использованием оценки градиента или же его истинного значения, если оно известно. То есть также как и ранее будут известны LF и HF данные, но кроме этого известны либо точные значения градиентов в X_{HF} точках, либо их оценки с помощью LF модели, которые будут использоваться для построения итоговой аппроксимационной модели. При этом в многомерных задачах, вместо k тренировочных точек, в которых определено только значение функции, может быть с большей эффективностью использоваться $k/2$ точек, однако в которых будет вычислено и само значение и значение градиента. Таким образом для построения аппроксимации известны: LF и HF – данные, оценка градиентов в HF – точках или их истинное значение, используемая ковариационная функция. Задача заключается в том, чтобы взяв за основу методы семейства kriging построить аппроксимацию для которой значения параметров будут состоятельными и асимптотически эффективными, что достигается при использовании методов максимального правдоподобия.

Предлагается следующий алгоритм:

1. Строится MF – модель на основе имеющихся LF – и HF – данных. Результирующая функция $y_{MF}(x)$ используется в качестве тренда.

2. Оценивается или используется известная информация о градиенте истинной функции в X_{HF} – точках и в результате составляются вектор

$$y_{HF}^{(g)} = \left(y_{HF}(x_1), \dots, y_{HF}(x_k), \left. \frac{\partial y_{HF}(x)}{\partial x} \right|_{x=x_1}, \dots, \left. \frac{\partial y_{HF}(x)}{\partial x} \right|_{x=x_k} \right)^T \text{ и вектор}$$

$$y_{MF}^{(g)} = \left(y_{MF}(x_1), \dots, y_{MF}(x_k), \left. \frac{\partial y_{MF}(x)}{\partial x} \right|_{x=x_1}, \dots, \left. \frac{\partial y_{MF}(x)}{\partial x} \right|_{x=x_k} \right)^T.$$

3. Дополняем имеющуюся MF модель полученными в пункте (2) векторами с информацией о градиенте, получаем новую уточненную модель (MFG – модель):

$$y_{MFG}(x) = y_{MF}(x) + 0.5 \cdot (c_g^T C_g^{-1} (y_{HF}^{(g)} - y_{MF}^{(g)}) + (y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1} c_g),$$



$$\text{где } c_g = (r_{xx}, \frac{\partial r_{xx}}{\partial x}), C_g = \begin{pmatrix} R & \frac{\partial R}{\partial x^{(i)}} \\ \frac{\partial R}{\partial x^{(j)}} & \frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} \end{pmatrix}$$

Заметим, что в качестве MF модели можно использовать любую из представленных выше, однако первоначальное использование модели семейства kriging позволяет не оценивать дополнительно параметры для новой градиентной корреляционной матрицы.

Теорема. Пусть известны наборы данных (X_{HF}, y_{HF}) и (X_{LF}, y_{LF}) , $X_{HF} = \{x_{HF}^{(1)}, \dots, x_{HF}^{(k)}\}$, $X_{LF} = \{x_{LF}^{(1)}, \dots, x_{LF}^{(v)}\}$, $x_{HF}^{(i)} \in R^n$, $x_{LF}^{(j)} \in R^n$, $X_{HF} \subset X_{LF}$, $y_{LF}^{(j)} = y_{LF}(x_{LF}^{(j)}) \in R$, $y_{HF}^{(i)} = y_{HF}(x_{HF}^{(i)}) \in R$ по которым построена MF – модель $y_{MF}(x)$, и известна оценка градиентов в точках X_{HF} . Тогда может быть построена функция $y_{MFG}(x) \in C^1$:

$$y_{MFG}(x) = y_{MF}(x) + 0,5 \cdot (c_g^T C_g^{-1} (y_{HF}^{(g)} - y_{MF}^{(g)}) + (y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1} c_g),$$

являющаяся аппроксимацией истинной функции $y(x)$.

Доказательство: Рассмотрим новую точку x_{new} , для которой необходимо вычислить $y_{new} = y(x_{new})$.

Введем обозначения $c_g^{new} = (r_{x_{new}x}, \frac{\partial r_{x_{new}x}}{\partial x})^T$ – корреляционный вектор для множества $\{X_{HF}, x_{new}\}$,

$$\text{а } C_g^{new} = \begin{pmatrix} C_g & c_g^{new} \\ c_g^{newT} & 1 \end{pmatrix}, \text{ где } C_g = \begin{pmatrix} R & \frac{\partial R}{\partial x^{(i)}} \\ \frac{\partial R}{\partial x^{(j)}} & \frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} \end{pmatrix} -$$

корреляционная матрица. Функционал для максимизации по методу максимального правдоподобия будет иметь вид:

$$\ln(L) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln |C_g^{new}| - \frac{(y_{new} - y_{MF}(x_{new}))^T C_g^{new-1} (y_{new} - y_{MF}(x_{new}))}{2\hat{\sigma}^2}.$$

Для невырожденной квадратной блочной матрицы

$$K = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ справедливо следующее представление}$$

для обратной к ней [29]:

$$K^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}.$$

Применяя к матрице C_g^{new} , имеем:

$$C_g^{new-1} = \begin{pmatrix} C_g^{-1} + C_g^{-1}c_g^{new}(1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1}c_g^{newT}C_g^{-1} & -C_g^{-1}c_g^{new}(1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1} \\ -(1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1}c_g^{newT}C_g^{-1} & (1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1} \end{pmatrix},$$

$$\begin{pmatrix} y_{HF}^{(g)} - y_{MF}^{(g)} \\ y_{new} - y_{MF}(x_{new}) \end{pmatrix}^T C_g^{new-1} \begin{pmatrix} y_{HF}^{(g)} - y_{MF}^{(g)} \\ y_{new} - y_{MF}(x_{new}) \end{pmatrix} =$$

$$= (y_{HF}^{(g)} - y_{MF}^{(g)})^T (C_g^{-1} + C_g^{-1}c_g^{new}(1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1}c_g^{newT}C_g^{-1}) (y_{HF}^{(g)} - y_{MF}^{(g)}) -$$

$$- (y_{new} - y_{MF}(x_{new})) (1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1}c_g^{newT}C_g^{-1} (y_{HF}^{(g)} - y_{MF}^{(g)}) -$$

$$- (y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1}c_g^{new}(1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1} (y_{new} - y_{MF}(x_{new})) +$$

$$+ (y_{new} - y_{MF}(x_{new}))^2 (1 - c_g^{newT}C_g^{-1}c_g^{new})^{-1}.$$

Максимизируем $\ln(L)$, находя производную по y_{new} . Имеем:

$$\frac{c_g^{newT}C_g^{-1}(y_{HF}^{(g)} - y_{MF}^{(g)})}{(1 - c_g^{newT}C_g^{-1}c_g^{new})} - \frac{(y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1}c_g^{new}}{(1 - c_g^{newT}C_g^{-1}c_g^{new})} + 2 \frac{(y_{new} - y_{MF}(x_{new}))}{(1 - c_g^{newT}C_g^{-1}c_g^{new})} = 0,$$

$$y_{new} = y_{MF}(x_{new}) + \frac{1}{2} (c_g^{newT}C_g^{-1}(y_{HF}^{(g)} - y_{MF}^{(g)}) + (y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1}c_g^{new}).$$

Отсюда получаем, что в любой точке x оценка истинной функции будет иметь вид:

$$y_{MFG}(x) = y_{MF}(x) + 0,5 \cdot (c_g^T C_g^{-1} (y_{HF}^{(g)} - y_{MF}^{(g)}) + (y_{HF}^{(g)} - y_{MF}^{(g)})^T C_g^{-1} c_g).$$

Заметим, что матрица C_g^{-1} , в свою очередь, может быть представлена следующим образом:

$$C_g^{-1} = \begin{pmatrix} R^{-1} + R^{-1} \frac{\partial R}{\partial x^{(i)}} \left(\frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} - \frac{\partial R}{\partial x^{(j)}} R^{-1} \frac{\partial R}{\partial x^{(i)}} \right)^{-1} \frac{\partial R}{\partial x^{(j)}} R^{-1} & -R^{-1} \frac{\partial R}{\partial x^{(i)}} \left(\frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} - \frac{\partial R}{\partial x^{(j)}} R^{-1} \frac{\partial R}{\partial x^{(i)}} \right)^{-1} \\ - \left(\frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} - \frac{\partial R}{\partial x^{(j)}} R^{-1} \frac{\partial R}{\partial x^{(i)}} \right)^{-1} \frac{\partial R}{\partial x^{(j)}} R^{-1} & \left(\frac{\partial^2 R}{\partial x^{(i)} \partial x^{(j)}} - \frac{\partial R}{\partial x^{(j)}} R^{-1} \frac{\partial R}{\partial x^{(i)}} \right)^{-1} \end{pmatrix}$$

При этом использование методов семейства kriging в качестве MF-модели предполагает, что матрицы R и R^{-1} уже построены и известны. Кроме того, используя для построения корреляционной матрицы гауссово ядро, сразу могут быть вычислены производные, участвующие в разложении:

$$\frac{\partial R^{(i,j)}}{\partial x^{(i)}} = -2\theta(x^{(i)} - x^{(j)})R^{(i,j)}, \frac{\partial R^{(i,j)}}{\partial x^{(j)}} = 2\theta(x^{(i)} - x^{(j)})R^{(i,j)}, \frac{\partial^2 R^{(i,j)}}{\partial x^{(i)} \partial x^{(j)}} = [2\theta - 4\theta^2(x^{(i)} - x^{(j)})^2]R^{(i,j)}.$$

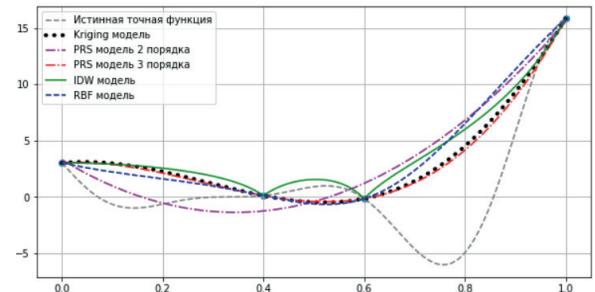
В многомерном случае корреляционная матрица примет вид:

$$C_g = \begin{pmatrix} R & \frac{\partial R}{\partial x_1^{(i)}} & \frac{\partial R}{\partial x_2^{(i)}} & \dots & \frac{\partial R}{\partial x_n^{(i)}} \\ \frac{\partial R}{\partial x_1^{(j)}} & \frac{\partial^2 R}{\partial x_1^{(i)} \partial x_1^{(j)}} & \frac{\partial^2 R}{\partial x_1^{(i)} \partial x_2^{(j)}} & \dots & \frac{\partial^2 R}{\partial x_1^{(i)} \partial x_n^{(j)}} \\ \frac{\partial R}{\partial x_2^{(j)}} & \frac{\partial^2 R}{\partial x_1^{(i)} \partial x_2^{(j)}} & \frac{\partial^2 R}{\partial x_2^{(i)} \partial x_2^{(j)}} & \dots & \frac{\partial^2 R}{\partial x_2^{(i)} \partial x_n^{(j)}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R}{\partial x_n^{(j)}} & \frac{\partial^2 R}{\partial x_1^{(i)} \partial x_n^{(j)}} & \frac{\partial^2 R}{\partial x_2^{(i)} \partial x_n^{(j)}} & \dots & \frac{\partial^2 R}{\partial x_n^{(i)} \partial x_n^{(j)}} \end{pmatrix},$$

$$\text{где } \frac{\partial^2 R^{(i,j)}}{\partial x_1^{(i)} \partial x_m^{(j)}} = -4\theta_i \theta_m (x_1^{(i)} - x_1^{(j)})(x_m^{(i)} - x_m^{(j)})R^{(i,j)}.$$

Сравнение методов на тестовых функциях

Для начала рассмотрим применение классических одновыборочных и двухвыборочных методов на распространенной тестовой задаче [22], а в дальнейшем применим представленный градиентный метод. Пусть $f_{HF}^0 = (6x - 2)^2 \cdot \sin(12x - 4)$ – высокоточная функция для которой имеются данные в точках $X_H = \{0, 0.4, 0.6, 1\}$, и $f_{LF}^0 = 0.5 \cdot (6x - 2)^2 \cdot \sin(12x - 4) + 10x - 5$ – низкоточная функция с известными значениями в точках $X_L = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.



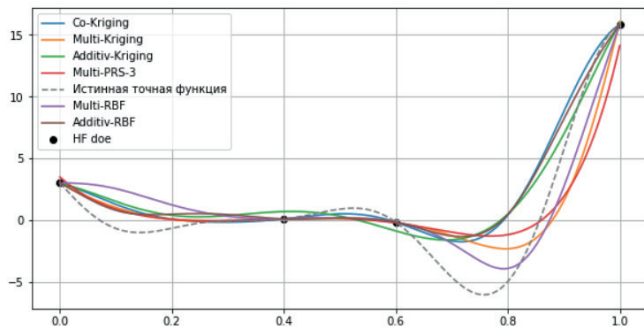
Р и с. 1. Аппроксимационные модели, построенные по выборке высокоточных данных

Fig. 1. Approximation models built from a high-precision data sample

Источник: здесь и далее в статье все таблицы и рисунки составлены автором.
Source: Hereinafter in this article all tables and figures were made by the author.



На рисунке 1 представлены результаты построения аппроксимационных моделей только по высокоточным данным. Несложно заметить, что по набору $(X_H, f_{HF}(X_H))$ исходная зависимость восстанавливается с большой погрешностью. Используя двухвыборочные методы получаем более точную аппроксимацию по фиксируемой среднеквадратической ошибке, что отражено в таблице 1. Также на рисунке 2 можно наблюдать лучшие из полученных аппроксимационных моделей.



Р и с. 2. Аппроксимационные двухвыборочные модели f_{MF}^0
F i g. 2. Approximation two-sample models f_{MF}^0

По размеру получаемой ошибки, наилучший результат из классических аппроксимаций для тестовой функции f^0 показывает Multi-RBF модель, а также семейство kriging моделей (Multi-Kriging, Additiv-Kriging, Co-Kriging). Однако, в случае с kriging моделями мы можем воспользоваться выборочной дисперсией, что может помочь в построении более точной модели. Поскольку kriging – это реализация регрессии гауссовского процесса, то воспользовавшись правилом трех сигм, получим интервальную оценку, которая охватывает 99,7% возможных траекторий, одна из которых, с высокой долей вероятности, является истинной. Используя интервальные оценки, можно указать границы, в которых будет варьироваться градиент в известных HF – точках (т.к. построенная MF модель уже проходит через эти точки).

Т а б л и ц а 1. Среднеквадратические ошибки построенных аппроксимационных моделей
T a b l e 1. Root-mean-square errors of the constructed approximation models

$MSE = \frac{1}{n} \sum_{i=1}^n (f(x_i) - \hat{f}(x_i))^2$	f_{MF}^0	f_{MF}^1	f_{MF}^2	f_{MF}^3	f_{MF}^4	f_{MF}^5
Kriging модель	16.43	0.76	19.89	0.044	2.13	10.25
PRS модель 2 степени	24.17	0.80	17.24	0.051	3.07	20.31
PRS модель 3 степени	15.12	0.83	19.72	0.043	2.27	15.67
IDW модель	22.38	0.66	15.14	0.043	2.78	22.78
RBF модель	22.74	0.62	19.07	0.051	2.51	12.38
Co – Kriging модель	5.10	0.03	2.18	0.010	0.54	5.21
Multi-Kriging модель	4.41	0.77	29.23	0.030	0.67	6.12
Additiv-Kriging модель	4.84	0.0015	2.77	0.041	0.82	4.57
Multi-RBF модель	3.40	1.05	28.21	0.041	1.54	5.65
Additiv-RBF модель	5.16	0.009	2.28	0.049	0.62	7.16
Multi-IDW модель	11.53	0.82	10.91	0.034	1.68	10.70
Additiv- IDW модель	9.18	0.13	3.65	0.043	1.35	11.37
Multi-PRS модель 2	10.36	0.85	13.95	0.031	1.27	9.43
Additiv-PRS модель 2	16.75	0.40	9.34	0.043	2.45	10.39
Multi-PRS модель 3	5.76	0.51	20.66	0.028	3.19	6.12
Additiv-PRS модель 3	9.43	0.37	8.13	0.039	2.37	7.76
MFG-Co-Kriging	1.51	0.0009	1.07	0.0056	0.21	1.23



Также для сравнения результатов в таблицу 1 добавлены среднеквадратические отклонения при построении различных аппроксимаций для ниже представленных функций.

$$1) f_{HF}^1 = \sin(10x) + 2x + \frac{(10x-5)^2}{16} + 0.5, f_{LF}^0 = \sin(10x) + 2x + 0.5,$$

$$X_H = \{0, 0.4, 0.6, 1\}, X_L = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}, x \in [0, 1].$$

$$2) f_{HF}^2 = e^x \cdot \cos(12x) \cdot (x^2 - 4x) - 5, f_{LF}^2 = 0.5 \cdot e^{1.12x-0.1} \cdot \cos(12x) \cdot ((x-1)^2 - 3x-1) + 5 \cdot (x-0.76),$$

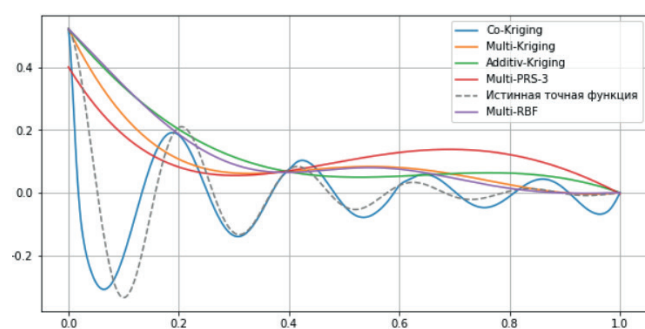
$$X_H = \{0, 0.4, 0.6, 1\}, X_L = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}, x \in [0, 1].$$

3) Функция затухающих колебаний:

$$f_{HF}^3 = 0.523 \cdot e^{-4.378x} \cdot \cos(29.934x),$$

$$f_{LF}^3 = 0.481 \cdot e^{-3.798x-0.12} \cdot \cos(27.212x-11.33), X_H = \{0, 0.4, 0.6, 1\},$$

$$X_L = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}, x \in [0, 1].$$



Р и с. 3. Суррогатные модели для функции затухающих колебаний

F i g. 3. Surrogate models for the damped oscillation function

Для данного примера стоит заметить, что кроме малой величины ошибки, модели Co-kriging и MFG – Co-kriging наиболее точно отражают поведение функции в отличие от других аппроксимаций, что продемонстрировано на рисунке 3. Заметим, что возможность достаточно точного восстановления функций с периодической структурой является важной составляющей для решения обратных задач, в том числе и для дифференциальных уравнений, поскольку функция затухающих колебаний возникает как решение дифференциального уравнения: $f''(x) + 2\beta f'(x) + \omega^2 f(x) = 0$.

4) Функция Розенброка:

$$f_{HF}^4(X) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2, f_{LF}^4(X) = 0.74 \cdot f_{HF}^4(X) + 20, X_{HF} \subset X_{LF} \in [-2; 2]$$

5) Функция Леви:

$$f_{HF}^5(X) = \sin^2(\pi\omega_1) + (\omega_1 - 1)^2 \cdot (1 + 10 \sin^2(\pi\omega_1 + 1)) + (\omega_2 - 1)^2 \cdot (1 + \sin^2(2\pi\omega_2)),$$

$$\omega_i = 1 + \frac{x_i - 1}{4}, X_{HF} \subset X_{LF} \in [-10; 10]$$

Заключение

В данной статье были описаны и применены к конкретным примерам основные методы аппроксимационного моделирования, а также был предложен и применен новый метод восстановления функции по двум разноточным выборкам с использованием значения градиента функции или его оценки. Данный метод, основываясь на методах семейства kriging может быть использован для задач различной размерности. Численная реализация представленных методов и построение необходимых графиков были осуществлены с использованием языка программирования Python 3. В дальнейшем планируется регистрация программы ЭВМ аппроксимации на основе суррогатных моделей.

References

- [1] Kang B.-H., Jo B.-H., Kim B.-G., Hwang J.-H., Choi S.-B. Linear and Nonlinear Models for Drop Simulation of an Aircraft Landing Gear System with MR Dampers. *Actuators*. 2023;12:287. <https://doi.org/10.3390/act12070287>
- [2] Vereschikov D.V., Fedorkevich I.A. The aircraft longitudinal movement dynamics simulation Simulink-model. *Aerospace forces. Theory and practice*. 2021;(17):128-138. (In Russ., abstract in Eng.) EDN: FMILH
- [3] Nguyen T.G. Development and testing simulation model of machine with worm gear. *News of the Tula state university. Technical sciences*. 2022;(3):33-42. (In Russ., abstract in Eng.) <https://doi.org/10.24412/2071-6168-2022-3-33-42>
- [4] Saporito M., Da Ronch A., Bartoli N., Defoort S. Robust multidisciplinary analysis and optimization for conceptual design of flexible aircraft under dynamic aeroelastic constraints. *Aerospace Science and Technology*. 2023;138:108349. <https://doi.org/10.1016/j.ast.2023.108349>
- [5] Zore K., Shah S., Stokes J., Sasanapuri B., Sharkey P. ANSYS CFD for High Lift Aircraft Configurations. In: AIAA AVIATION Forum, 36th AIAA Applied Aerodynamics Conference 2018. Atlanta, Georgia, USA: Curran Associates, Inc.; 2018. <https://doi.org/10.2514/6.2018-2844>
- [6] Han Z.-H., Zhang K.-S. 'Surrogate-Based Optimization'. In: O. Roeva (ed.) Real-World Applications of Genetic Algorithms. InTech; 2012. p. 343-362. <http://dx.doi.org/10.5772/36125>
- [7] Jones D.R., Schonlau M., Welch W.J. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*. 1998;13:455-492. <https://doi.org/10.1023/A:1008306431147>
- [8] Bartoli N., Lefebvre T., Dubreuil S., Olivanti R., Priem R., Bons N., Martins J.R.R.A., Morlier J. Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design. *Aerospace Science and Technology*. 2019;90:85-102. <https://doi.org/10.1016/j.ast.2019.03.041>
- [9] Priem R., Bartoli N., Diouane Y., Sgueglia A. Upper trust bound feasibility criterion for mixed constrained Bayesian optimization with application to aircraft design. *Aerospace Science and Technology*. 2020;105:105980. <https://doi.org/10.1016/j.ast.2020.105980>



- [10] Jones D. A Taxonomy of Global Optimization Methods Based on Response Surfaces. *Journal of Global Optimization*. 2001;21:345-383. <https://doi.org/10.1023/A:1012771025575>
- [11] Pehlivan Solak H., et al. Hydrofoil Optimization via Automated Multi-Fidelity Surrogate Models. In: 10th Conference on Computational Methods in Marine Engineering (Marine 2023). Scipedia, S.L.; 2023. <https://doi.org/10.23967/marine.2023.136>
- [12] Mourousias N., Malim A., Marinus B., Runacres M. Multi-fidelity Multi-Objective Optimization of a High-Altitude Propeller. In: 2023 AIAA Aviation and Aeronautics Forum and Exposition (AIAA AVIATION Forum). San Diego, CA, USA: Curran Associates, Inc.; 2023. <https://doi.org/10.2514/6.2023-3590>
- [13] Toal D.J.J. Applications of multi-fidelity multi-output Kriging to engineering design optimization. *Structural and Multidisciplinary Optimization*. 2023;66:125. <https://doi.org/10.1007/s00158-023-03567-z>
- [14] Han Z., et al. Efficient aerodynamic shape optimization using variable-fidelity surrogate models and multilevel computational grids. *Chinese Journal of Aeronautics*. 2020;33(1):31-47. <https://doi.org/10.1016/j.cja.2019.05.001>
- [15] Lili Z., et al. A multi-fidelity surrogate modeling approach for incorporating multiple non-hierarchical low-fidelity data. *Advanced Engineering Informatics*. 2022;51:101430. <https://doi.org/10.1016/j.aei.2021.101430>
- [16] Box G.E.P., Wilson K.B. On the Experimental Attainment of Optimum Condition. *Journal of the Royal Statistical Society, Series B*. 1951;13(1):1-38. <https://doi.org/10.1111/j.2517-6161.1951.tb00067.x>
- [17] Seber G.A.F., Lee A.J. Linear Regression Analysis. John Wiley & Sons, Inc.; 2003. 592 p. <https://doi.org/10.1002/9780471722199>
- [18] Shepard D. A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM National Conference (ACM '68). New York, NY, USA: ACM; 1968. p. 517-524. <https://doi.org/10.1145/800186.810616>
- [19] Powell M.J.D. The Theory of Radial Basis Function Approximation in 1990. In: W. Light (ed.) Advances in Numerical Analysis: Wavelets, Subdivision Algorithms, and Radial Basis Functions. Oxford University Press; 1992. p. 105-210. <https://doi.org/10.1093/oso/9780198534396.003.0003>
- [20] Matheron G. Principles of geostatistics. *Economic geology*. 1963;58(8):1246-1266. <https://doi.org/10.2113/gsecongeo.58.8.1246>
- [21] Rasmussen C.E., Williams C.K.I. Gaussian Processes for Machine Learning. The MIT Press, 2005. 272 p. <https://doi.org/10.7551/mitpress/3206.001.0001>
- [22] Forrester A.I.J., Sóbester A., Keane A.J. Engineering Design via Surrogate Modelling: A Practical Guide. Progress in astronautics and aeronautics. A John Wiley and Sons, Ltd., Publication; 2008. 240 p. <https://doi.org/10.1002/9780470770801>
- [23] Fernández-Godino M.G. Review of Multi-fidelity Models. *Preprints*. 2023:2023041264. <https://doi.org/10.20944/preprints202304.1264.v1>
- [24] Haftka R.T. Combining global and local approximations. *AIAA Journal*. 1991;29(9):1523-1525. <https://doi.org/10.2514/3.10768>
- [25] Lewis R.M., Nash S.G. A multigrid approach to the optimization of systems governed by differential equations. In: 8th Symposium on Multidisciplinary Analysis and Optimization. Long Beach, CA, U.S.A.; 2000. p. 4890. <https://doi.org/10.2514/6.2000-4890>
- [26] Gano S.E., Renaud J.E., Sanders B. Hybrid Variable Fidelity Optimization by Using a Kriging-Based Scaling Function. *AIAA Journal*. 2005;43(11):2422-2433. <https://doi.org/10.2514/1.12466>
- [27] Myers D.E. Co-Kriging – New Developments. In: Verly G., David M., Journel A.G., Marechal A. (eds.) Geostatistics for Natural Resources Characterization. Dordrecht: Springer; 1984. p. 295-305. https://doi.org/10.1007/978-94-009-3699-7_18
- [28] Jiang P., Zhou Q., Shao X. Surrogate Model-Based Engineering Design and Optimization. *Springer Tracts in Mechanical Engineering (STME)*. Singapore: Springer; 2020. 240 p. <https://doi.org/10.1007/978-981-15-0731-1>
- [29] Lu T.-T., Shiou S.-H. Inverses of 2×2 block matrices. *Computers & Mathematics with Applications*. 2002;43(1-2):119-129. [https://doi.org/10.1016/S0898-1221\(01\)00278-4](https://doi.org/10.1016/S0898-1221(01)00278-4)

Поступила 12.07.2024; одобрена после рецензирования 11.09.2024; принята к публикации 26.09.2024.

Submitted 12.07.2024; approved after reviewing 11.09.2024; accepted for publication 26.09.2024.

Об авторе:

Капкаев Никита Викторович, аспирант кафедры дифференциальных уравнений, математического и численного анализа Института информационных технологий, математики и механики, ФГАОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского» (603022, Российская Федерация, г. Нижний Новгород, пр. Гагарина, д. 23); старший программист, ООО «Сириус» (603034, Российская Федерация, г. Нижний Новгород, ул. Кисловодская, д. 1/6), **ORCID:** <https://orcid.org/0009-0004-7400-524X>, kapkaevn@gmail.com

Автор прочитал и одобрил окончательный вариант рукописи.

About the author:

Nikita V. Kapkaev, Postgraduate Student of the Chair of Differential Equations Mathematical and Numerical Analysis, Institute of Information Technology, Mathematics and Mechanics, National Research Lobachevsky State University of Nizhny Novgorod (23 Gagarin Ave., Nizhny Novgorod 603022, Russian Federation); Senior developer, Sirius LLC (1/6 Kislovodskaya St., Nizhny Novgorod 603034, Russian Federation), **ORCID:** <https://orcid.org/0009-0004-7400-524X>, kapkaevn@gmail.com

The author has read and approved the final manuscript.

