

# Моделирование нестационарной ковариационной функции гауссовского процесса на основе разложения по словарю базисных функций

Евгений Бурнаев, Алексей Зайцев, Максим Панов, Павел Приходько и Юрий Янович

1. *Институт Проблем Передачи Информации,*  
127994, г.Москва, ГСП-4, Большой Каретный переулок, 19, стр.1

2. *DATAADVANCE,*

105064, г.Москва, Садово-Черногорская улица, 13/3

3. *МФТИ,*

141700, г.Долгопрудный, Институтский переулок, 9

{evgeny.burnaev, alexey.zaytsev, maxim.panov, pavel.prikhodko, yurij.yanovich}@datadvance.net

## Аннотация

*В работе предложен метод моделирования нестационарной ковариационной функции гауссовского процесса на основе словаря параметрических базисных функций, настроенных по данным. Применение разработанного подхода к задаче восстановления неизвестной зависимости позволило существенно повысить точность аппроксимации. Отдельное внимание уделено аппроксимации функций с многомерным выходом и интерполирующим свойствам аппроксиматоров на основе гауссовских процессов.*

## 1. Введение

Одной из основных задач, которые приходится решать при построении метамоделей (моделей на основе данных) является задача аппроксимации неизвестной зависимости по данным [1, 2]. Наиболее популярная модель для построения аппроксиматоров, основанная на гауссовских процессах [3, 4, 5], используется в большом количестве разнообразных прикладных задач, включая концептуальное проектирование [6], структурную оптимизацию [7], многокритериальную оптимизацию при проектировании [8], конструирование в аэрокосмической [9] и автомобильной отраслях [10].

В своих классических реализациях, таких как DACE [15], аппроксиматоры на основе гауссовских процессов обычно применяются для моделирования пространственно однородных функций, что определяется использованием стационарных ковариационных функций. В данной работе мы предлагаем метод моделирования нестационарной ковариационной

функции на основе словаря параметрических базисных функций, настроенных по данным, который позволяет с высокой степенью точности аппроксимировать пространственно неоднородные и даже разрывные функции.

Статья устроена следующим образом:

- в разделе 2 описывается постановка задачи аппроксимации;
- в разделе 3 рассказывается про построение аппроксиматора на основе стационарных гауссовских процессов;
- в разделе 4 описывается метод построения нестационарной ковариационной функции гауссовского процесса и аппроксиматора на ее основе;
- в разделах 5 и 6 уделено внимание соответственно моделированию гауссовских процессов с многомерным выходом и интерполяционным свойствам аппроксиматоров на основе гауссовских процессов;
- в разделах 7 и 8 рассказывается об экспериментальных результатах работы предложенного метода и делаются выводы об особенностях предложенного метода и качестве его работы.

## 2. Постановка задачи

В наиболее общем виде задача аппроксимации может быть сформулирована следующим образом. Пусть  $y = f(x)$  некоторая неизвестная функция со входом  $x \in \mathbb{X} \subset \mathbb{R}^n$  и выходом  $y \in \mathbb{R}^m$ . Далее будем предполагать, что  $y \in \mathbb{R}^1$ , а случай многомерного выхода будет подробно рассмотрен в разделе 5. Пусть

$D_{learn} = (X, Y) = \{(x_i, y_i = f(x_i)), i = 1, \dots, N\}$  - обучающая выборка. Задача состоит в построении аппроксимации  $\hat{y} = \hat{f}(x) = \hat{f}(x|D_{learn})$  для исходной зависимости  $y = f(x)$  по обучающей выборке  $D_{learn}$ .

Если для всех  $x \in \mathbb{X}$  (не только для  $x \in D_{learn}$ ) имеет место примерное равенство  $\hat{f}(x) \approx f(x)$ , то считается, что аппроксиматор хорошо воспроизводит исходную зависимость. Это факт проверяется на независимой тестовой выборке  $D_{test} = (X_*, Y_*) = \{(x_j, y_j = f(x_j)), j = 1, \dots, N_*\}$ . Мерой качества аппроксимации является среднеквадратичная ошибка на контрольной выборке:

$$\varepsilon(\hat{f}|D_{test}) = \sqrt{\frac{1}{N_*} \sum_{j=1}^{N_*} (y_j - \hat{f}(x_j))^2}. \quad (1)$$

### 3. Аппроксиматор на основе стационарного гауссовского процесса

#### 3.1. Гауссовские процессы

Гауссовский процесс является одним из возможных способов задания распределения на пространстве функций. Гауссовский процесс  $f(x)$  полностью определяется своей функцией среднего  $m(x) = \mathbb{E}[f(x)]$  и ковариационной функцией  $cov(y, y') = k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]$ . Если положить функцию среднего нулевой:  $m(x) = \mathbb{E}[f(x)] = 0$ , а ковариационную функцию считать известной, то функция апостериорного (для заданной обучающей выборки) среднего значения гауссовского процесса в точках контрольной выборки  $X_*$  выглядит следующим образом [12]:

$$\hat{f}(X_*) = K_* K^{-1} Y, \quad (2)$$

где  $K_* = K(X_*, X) = [k(x_i, x_j), i = 1, \dots, N_*; j = 1, \dots, N]$ ,  $K = K(X, X) = [k(x_i, x_j), i, j = 1, \dots, N]$ .

В типичных, более реалистичных ситуациях при моделировании мы не имеем доступа непосредственно к значениям функции, а наблюдаем их только в зашумленном виде:

$$y(x) = f(x) + \varepsilon(x), \quad (3)$$

где шум  $\varepsilon(x)$  моделируется независимыми одинаково распределенными нормальными случайными величинами с нулевым средним и дисперсией  $\tilde{\sigma}^2$ . В таком случае наблюдения  $y(x)$  будут гауссовским процессом с нулевым средним и ковариационной функцией  $cov(y(x), y(x')) = k(x, x') + \tilde{\sigma}^2$ . Таким образом, функция апостериорного (для заданной обучающей выборки) среднего значения гауссовского процесса  $f(x)$  в точках контрольной выборки  $X_*$  принимает вид:

$$\hat{f}(X_*) = K_* (K + \tilde{\sigma}^2 I)^{-1} Y, \quad (4)$$

где  $I$  - единичная матрица размера  $N \times N$ .

Заметим, что наличие в формуле (4) дисперсии шума  $\tilde{\sigma}^2$  фактически приводит к регуляризации, что позволяет улучшить обобщающую способность аппроксиматора. При этом апостериорная ковариационная функция гауссовского процесса в точках контрольной выборки имеет вид:

$$\mathbb{V}[\hat{f}(X_*)] = K(X_*, X_*) + \tilde{\sigma}^2 I_* - K_* (K + \tilde{\sigma}^2 I)^{-1} K_*^T \quad (5)$$

где  $K(X_*, X_*) = [k(x_i, x_j), i, j = 1, \dots, N_*]$ ,  $I_*$  - единичная матрица размера  $N_* \times N_*$ .

Дисперсии гауссовского процесса в точках контрольной выборки могут быть использованы как оценки ожидаемой ошибки аппроксимации в этих точках. Заметим, что для этого нет необходимости вычислять по формуле (5) всю матрицу  $\mathbb{V}[\hat{f}(X_*)]$ , а достаточно вычислить только элементы ее главной диагонали, которые и являются искомыми дисперсиями.

При работе с реальными данными ковариационная функция породившего их гауссовского процесса, как правило, не известна, поэтому необходимо уметь идентифицировать ее по данным.

#### 3.2. Нахождение параметров гауссовского процесса

Предположим, что ковариационная функция гауссовского процесса является членом некоторого параметрического семейства  $k(x, x') = k(x, x'|a)$ , где  $a \in \mathbb{R}^K$  - вектор параметров ковариационной функции. Семейство  $k(x, x'|a)$  обычно берется из класса так называемых стационарных ковариационных функций, т.е. функций значение которых зависит только от разности значений аргументов  $k(x, x'|a) = k(x - x'|a)$ . Значение параметра  $a$  предлагается восстанавливать по обучающей выборке  $D_{learn}$ , исходя из принципа максимума правдоподобия. Для этого выпишем логарифм правдоподобия гауссовского процесса в точках обучающей выборки [12]:

$$\log p(Y|X, a, \tilde{\sigma}) = -\frac{1}{2} Y^T (K + \tilde{\sigma}^2 I_N)^{-1} Y - \frac{1}{2} \log |K + \tilde{\sigma}^2 I| - \frac{n}{2} \log 2\pi, \quad (6)$$

где  $|K + \tilde{\sigma}^2 I|$  - детерминант матрицы  $K + \tilde{\sigma}^2 I$ .

Кроме параметров  $a$  ковариационной функции параметром функционала (6) является также значение дисперсии шума наблюдений  $\tilde{\sigma}^2$ , которое также можно настраивать по обучающей выборке. Таким образом, нахождение оптимальных значений параметров сводится к отысканию максимума правдоподобия по параметрам:

$$\log p(Y|X, a, \tilde{\sigma}) \rightarrow \max_{a, \tilde{\sigma}} \quad (7)$$

Выбор конкретного семейства ковариационных функций  $k(x, x'|a)$  обычно продиктован соображениями удобства, а также априорными представлениями о свойствах аппроксимируемой зависимости. В данной работе мы используем ковариационные функции вида  $k(x - \tilde{x}|a) = \hat{\sigma}^2 \exp(-\{\sum_{i=1}^n \theta_i^2 |x_i - \tilde{x}_i|^p\}^q)$ , где  $p \in (0, 2]$ ,  $q \in (0, 1]$  являются задаваемыми априори параметрами, а параметры  $a = \{\theta_i, i = 1, \dots, n; \hat{\sigma}\}$  настраиваются по обучающей выборке при решении задачи (7).

Заметим, что использование стационарной ковариационной функции не позволяет получать хорошие результаты в том случае, когда свойства гауссовского процесса меняются в зависимости от точки пространства. Это соответствует случаю, когда аппроксимируемая функция является пространственно неоднородной. В следующем разделе будет предложен метод построения ковариационной функции для нестационарного гауссовского процесса.

## 4. Аппроксиматор на основе нестационарного гауссовского процесса

### 4.1. Модель нестационарного гауссовского процесса

Рассмотрим нестационарный гауссовский процесс следующего вида:

$$y = f(x) + \varepsilon(x), \quad (8)$$

где  $f(x) = \sum_{i=1}^Q \alpha_i \psi_i(x) + \tilde{f}(x)$ ,  $\tilde{f}(x)$  - гауссовский процесс с нулевым средним и стационарной ковариационной функцией  $k(x, x') = k(x - x'|a)$ , а  $\alpha_i$ ,  $i = 1, \dots, Q$  — независимые одинаково распределенные нормальные случайные величины с нулевым средним и дисперсией  $\frac{\sigma^2}{Q}$ , а  $\{\psi_i(x), i = 1, \dots, Q\}$  — набор фиксированных функций. Для такого процесса ковариационная функция принимает вид:

$$\text{cov}(y, y') = k(x - x'|a) + \hat{\sigma}^2 \tilde{\psi}(x)^T \tilde{\psi}(x') + \hat{\sigma}^2, \quad (9)$$

где  $\tilde{\psi}(x) = \{\psi_i(x), i = 1, \dots, Q\}$ .

Введем обозначения  $\Psi(X) = \{\tilde{\psi}(x_i), i = 1, \dots, N\}$ ,  $\Psi(X_*) = \{\tilde{\psi}(x_i), i = 1, \dots, N_*\}$ ,  $\bar{K}_* = K_* + \hat{\sigma}^2 \Psi(X_*)^T \Psi(X_*)$ ,  $\bar{K} = K + \hat{\sigma}^2 \Psi(X)^T \Psi(X)$ . Тогда апостериорное математическое ожидание процесса (8) в точках контрольной выборки:

$$\hat{f}(X_*) = \bar{K}_* (\bar{K}_* + \hat{\sigma}^2 I)^{-1} Y. \quad (10)$$

Соответственно, апостериорная ковариация предсказаний в точках контрольной выборки:

$$\begin{aligned} \mathbb{V}[\hat{f}(X_*)] &= K(X_*, X_*) + \hat{\sigma}^2 I_* \\ &+ \hat{\sigma}^2 \Psi(X_*)^T \Psi(X_*) - \bar{K}_* (\bar{K}_* + \hat{\sigma}^2 I)^{-1} \bar{K}_*^T. \end{aligned} \quad (11)$$

Логарифм правдоподобия гауссовского процесса (8) совпадает с выражением (6), в котором матрица  $K$  заменена на матрицу  $\bar{K}$ :

$$\begin{aligned} \log p(Y|X, a, \tilde{\sigma}, \hat{\sigma}) &= -\frac{1}{2} Y^T (\bar{K} + \hat{\sigma}^2 I_N)^{-1} Y - \\ &- \frac{1}{2} \log |\bar{K} + \hat{\sigma}^2 I_N| - \frac{n}{2} \log 2\pi. \end{aligned} \quad (12)$$

Если набор базисных функции задан, то для идентификации аппроксимационной модели правдоподобие (12) необходимо максимизировать по параметрам  $a$  ковариационной функции  $k(x - x'|a)$ , а также по параметрам  $\tilde{\sigma}^2$  и  $\hat{\sigma}^2$ , т.е.

$$\log p(Y|X, a, \tilde{\sigma}, \hat{\sigma}) \rightarrow \max_{a, \tilde{\sigma}, \hat{\sigma}}. \quad (13)$$

Построение набора базисных функций  $\{\psi_i(x), i = 1, \dots, Q\}$  по обучающей выборке представляет собой отдельную задачу, которую мы кратко рассмотрим в следующем разделе.

### 4.2. Построение словаря базисных функций

В статье [13] авторами данной работы был предложен метод построения аппроксимации неизвестной многомерной зависимости на основе данных. Аппроксимация строится в классе линейных разложений по словарю параметрических функций:

1. Сигмоидные базисные функции  $\psi_j(x) = h(\sum_{i=1}^n \beta_{j,i} x^i)$ , где  $x = (x^1, \dots, x^n)$ ,  $h(z) = g(|z|^{g(\gamma_j)+1} \text{sign}(z))$ ,  $g(v) = \frac{e^v - 1}{e^v + 1}$ , а  $\gamma_j$  и  $\beta_{j,i}$  — настраиваемые параметры функции.
2. Радиальные базисные функции  $\psi_j(x) = \exp\left(-\|x - d_j\|^2 / t_j^2\right)$ ,  $d_j$  и  $t_j$  — настраиваемые параметры функции.
3. Линейные базисные функции  $\psi_j(x) = x^j$ ,  $j = 1, \dots, n$ ,  $x = (x^1, \dots, x^n)$ .

Параметры функций из словаря настраиваются с помощью градиентных методов [13].

## 5. Моделирование многомерного выхода

В предыдущих разделах рассматривалось построение аппроксиматора на основе модели гауссовского процесса для случая одномерного выхода  $y \in \mathbb{R}^1$ . В данном разделе мы обобщим эти результаты на случай многомерного выхода  $y = (y_1, \dots, y_m) \in \mathbb{R}^m$ ,  $m > 1$ . Простейшим вариантом является построение такого аппроксиматора для каждого выхода  $y_l$ ,  $l = 1, \dots, m$  в отдельности. Такой подход, однако,

никак не учитывает взаимную корреляцию выходов, учет которой может существенно повысить качество аппроксимации. Рассмотрение не только корреляционных свойств каждого выхода, но и кросскорреляций между выходами, приводит к алгоритму, который известен в геостатистической литературе под названием кокригинг [14].

В нашей работе мы предлагаем промежуточный подход, который предполагает, что для разных выходов гауссовские процессы  $y_l(x) = \sum_{i=1}^Q \alpha_{i,l} \psi_i(x) + \tilde{f}_l(x) + \varepsilon_l(x)$  независимы, но порождены ковариационными функциями из одного класса. Для этого предположим, что для процесса  $\tilde{f}_l(x)$   $l$ -го выхода ковариационная функция имеет вид

$$k_l(x - \bar{x} | a_l) = \{ \hat{\sigma}_l^2, \theta_i, i = 1, \dots, n \} = \\ = \hat{\sigma}_l^2 \exp \left( - \left\{ \sum_{i=1}^n \theta_i^2 |x_i - \bar{x}_i|^p \right\}^q \right),$$

где масштабирующий коэффициент  $\hat{\sigma}_l^2$  зависит от выхода, параметры  $\{ \theta_i, i = 1, \dots, n \}$  одинаковы для всех выходов, при этом, коэффициент  $p$  может принимать значения в  $(0, 2]$ , а  $q$  в  $(0, 1]$ . Предположим также, что дисперсии коэффициентов  $\{ \alpha_{i,l}, i = 1, \dots, Q \}$  зависят от выхода, т.е.  $\mathbb{E} \alpha_{i,l}^2 = \frac{\tilde{\sigma}_l^2}{Q}, i = 1, \dots, Q, l = 1, \dots, m$ . Компоненты  $\varepsilon_l(x), l = 1, \dots, m$  шума независимы между собой, моделируются независимыми одинаково распределенными нормальными случайными величинами с нулевым средним и одинаковой дисперсией  $\tilde{\sigma}^2$ .

Правдоподобие такой модели равно сумме правдоподобий вида (12) по всем каналам выхода. Оптимальные значения параметров находятся максимизацией правдоподобия:

$$\log p(Y|X, \tilde{\sigma}, a_l, \tilde{\sigma}_l, l = 1, \dots, m) = \\ = \sum_{l=1}^m \log p(Y_l|X, \tilde{\sigma}, a_l, \tilde{\sigma}_l) \rightarrow \max_{\{ \tilde{\sigma}_l, a_l \}_{l=1}^m, \tilde{\sigma}}. \quad (14)$$

Недостатком такого подхода является тот факт, что хотя предсказания среднего значения соответствующих гауссовских процессов будет зависеть от выхода:

$$\hat{f}_l(X_*) = \bar{K}_{l,*} (\bar{K}_l + \tilde{\sigma}^2 I_N)^{-1} Y_l, l = 1, \dots, m,$$

все выходы будут моделироваться на основе гауссовских процессов, ковариационные матрицы которых отличаются только масштабирующим множителем  $\hat{\sigma}_l$ . Это проявляется, например, в том, что для случая интерполяции апостериорные ковариации (11) разных выходов будут отличаться только масштабом.

## 6. Интерполирующие свойства гауссовских процессов

Важным свойством гауссовских процессов является тот факт, что в случае рассмотрения модели наблюдений без шума получаемый аппроксиматор (2)  $\hat{f}(x)$  обладает интерполирующим свойством. Действительно, при  $X_* = X$  мы получаем  $\hat{f}(X) = K(X, X)K(X, X)^{-1}Y = Y$ . Во многих практических задачах интерполирующее поведение аппроксиматора, а именно, выполнение условия  $\hat{f}(X) = Y$  является важным условием его применимости. В моделях с шумом (3) и (8) интерполирующее свойство в силу наличия шума вообще говоря не наблюдается. Тем не менее для модели (8) можно показать, что при достаточно малых значениях шума можно подобрать такие значения параметров  $\{ \hat{\sigma}^2; \theta_i, i = 1, \dots, n; \hat{\sigma}_l, \tilde{\sigma}_l, l = 1, \dots, m \}$ , что поведение аппроксиматора будет интерполирующим с любой наперед заданной точностью. Формально это утверждение дается следующей теоремой:

**Теорема.** Пусть задано  $\varepsilon > 0$ , тогда, если

- $\| \Psi(X)^T \Psi(X) \| < C, C \in \mathbb{R};$
- $\tilde{\sigma}^2 < \frac{1}{8} \min \left( \min_{i=1, N} \lambda_i, \varepsilon \right);$
- $\hat{\sigma}^2 > \frac{1}{2};$
- $\tilde{\sigma}^2 < \frac{1}{8} \cdot \frac{\min_{i=1, N} \lambda_i}{C+1},$

где  $\lambda_i, i = 1, \dots, n$  - это собственные числа матрицы  $K$ , а норма понимается как стандартная евклидова, то:

$$\| Y - \hat{f}(X) \| < \varepsilon \| Y \|.$$

## 7. Экспериментальные результаты

В ходе экспериментов сравнивалась работа трех алгоритмов, основанных на гауссовских процессах, а именно:

- широко известный алгоритм DACE [15], основанный на гауссовских процессах (KR2 на графике);
- алгоритм на основе гауссовских процессов без использования нестационарной ковариационной функции, разработанный авторами статьи (MacrosGP на графике);
- алгоритм на основе гауссовских процессов с использованием нестационарной ковариационной функции, разработанный авторами статьи (MacrosGP\_HDA на графике);

Для демонстрации экспериментальных результатов были использован большой набор тестовых функций, которые применяются для тестирования задач оптимизации [16, 17]. Всего тестирование проводилось на 23 различных функциях, для каждой из которых генерировалось 2 случайных обучающих выборках размером 160 точек. Результаты сравнивались по среднеквадратичной ошибке (1) на больших контрольных выборках из 10000 точек. Для удобства результаты представлены в виде кривых Долан-Мора [18] на рисунке 1. Чем выше кривая находится на графике, тем выше качество работы соответствующего алгоритма. Заметим, что наша реализация аппроксиматора на основе гауссовских процессов заметно превосходит DACE, а добавление нестационарной ковариационной функции позволяет еще значительно повысить точность аппроксимации.

Также предлагается рассмотреть пример, на котором будет видно, как использование нестационарной ковариационной функции позволяет заметно увеличить качество аппроксимации. Для этого рассмотрим аппроксимацию функции Хевисайда на отрезке  $[0, 1]$  по 20 случайным точкам. Графики аппроксимации при использовании стационарной и нестационарной ковариационных функций показаны на рисунке 2. Введение нестационарной ковариационной функции позволило заметно увеличить качество аппроксимации.

## 8. Выводы

В работе предложен метод моделирования нестационарной ковариационной функции гауссовского на основе словаря параметрических базисных функций, настроенного по данным. В отличие от других реализаций аппроксиматоров на основе гауссовских процессов, таких как DACE [15], предложенный метод полностью автоматизирован и не требует от применяющего его инженера каких-либо знаний в области машинного обучения. Экспериментальные результаты показали высокую качество работы предложенного алгоритма по сравнению с обычными методами. Дальнейшие исследования предполагается сосредоточить в области введения дополнительной регуляризации на параметры метода, основанной на вероятностных предположениях. Такая регуляризация позволит увеличить обобщающую способность и надежность алгоритма.

## Список литературы

[1] Бернштейн А.В., Бурнаев Е.В., Кулешов А.П. Интеллектуальный анализ данных в метамоделировании. // Труды 17 Всероссийского Семинара "Нейроинформатика и ее приложения к Анализу Данных Красноярск, 2009 – с. 23-28.

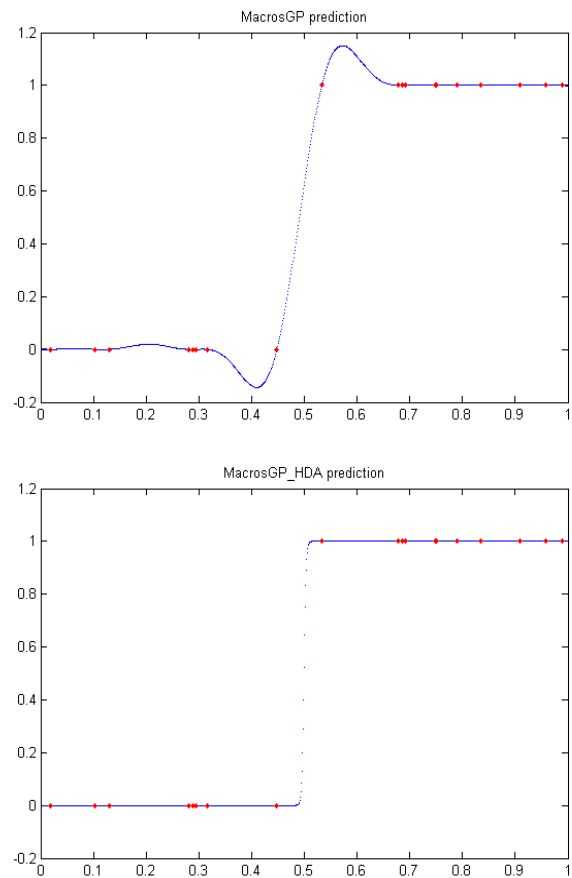


Рис. 2. Аппроксимация функции Хевисайда

- [2] Forrester A., Sobester A., Keane A. Engineering Design via Surrogate Modelling. A Practical Guide. – Wiley, 2008. – 238 p.
- [3] Giunta A., Watson L. T.A Comparison of Approximation Modeling Technique: Polynomial Versus Interpolating Models. // 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Vol. 1, AIAA, Reston, VA, 1998 – pp. 392–404.
- [4] Simpson T. W., Booker A. J., Ghosh S., Giunta A., Koch P. N., Yang, R. J. Approximation Methods in Multidisciplinary Analysis and Optimization: A Panel Discussion. // Structural and Multidisciplinary Optimization, Vol. 27, No. 5, 2004 – pp. 302–313.
- [5] Batill S. M., Renaud J. E., Gu X., Modeling and Simulation Uncertainty in Multidisciplinary Design Optimization // AIAA Paper 2000-4803, Sept. 2000.
- [6] Pacheco J. E., Amon C. H., Finger S. Bayesian Surrogates Applied to Conceptual Stages of the Engineering Design Process // Journal of Mechanical Design, Vol. 125, No. 4, 2003 – pp. 664–672.
- [7] Booker A. J., Dennis J. E., Frank P. D., Serafini D. B. Torczon, V., Trosset, M. A Rigorous Framework for Optimization of Expensive Functions by Surrogates // Structural Optimization, Vol. 17, No. 1, 1999, – pp. 1–13.
- [8] Koch P. N., Wujek B. A., Golovidov O., Simpson, T.

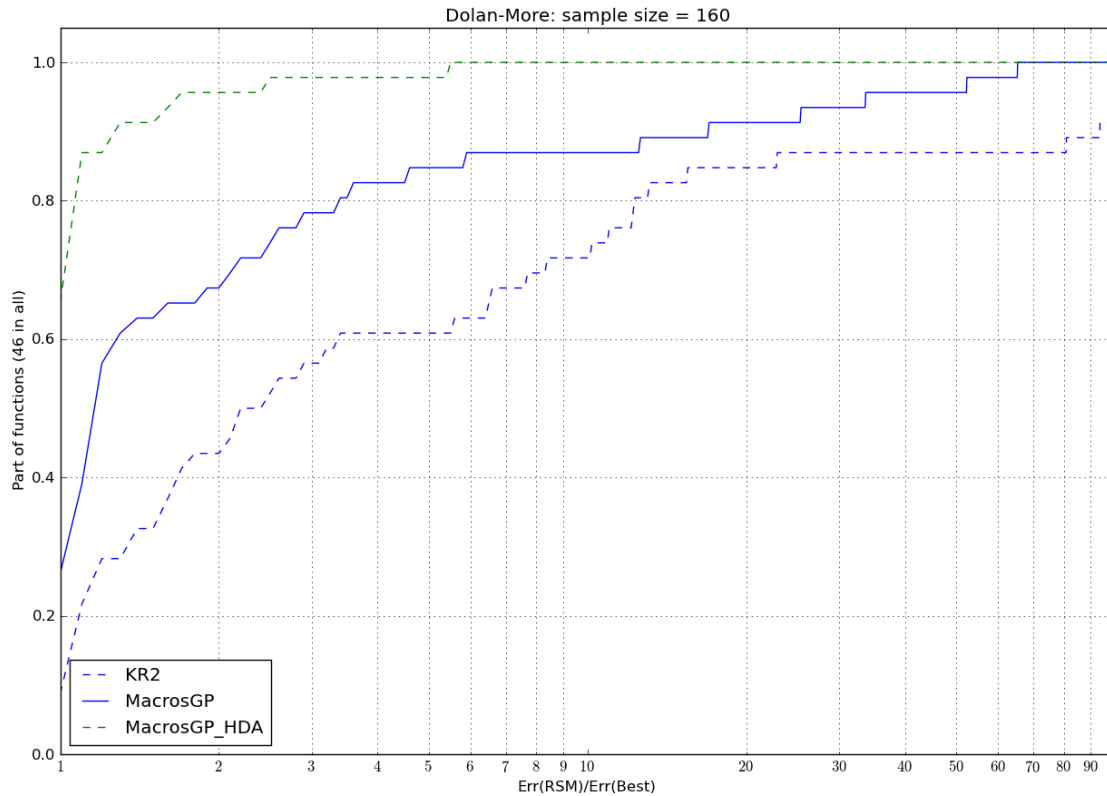


Рис. 1. Кривые Долан-Мора для тестовых функций

- W. Facilitating Probabilistic Multidisciplinary Design Optimization Using Kriging Approximation Models // AIAA Paper 2002-5415, Sept. 2002.
- [9] Simpson T. W., Maurey T. M., Korte J. J., and Mistree F. Kriging Metamodels for Global Approximation in Simulation-Based Multidisciplinary Design Optimization // AIAA Journal, Vol. 39, No. 12, 2001 – pp. 2233–2241.
- [10] Yang R. J., Wang N., Tho C. H., Bobineau J. P., and Wang B. P. Metamodeling Development for Vehicle Frontal Impact Simulation // American Society of Mechanical Engineers, ASME Design Engineering Technical Conf.—Design Automation Conf., DETC2001/DAC-21012, Sept. 2001.
- [11] Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: data mining, inference, and prediction. – Springer, 2008. – 763 p.
- [12] Rasmussen C.E., Williams C.K.I. Gaussian Processes for Machine Learning. – the MIT Press, 2006.
- [13] Burnaev E. V., Belyaev M. G., Prihodko P. V. About hybrid algorithm for tuning of parameters in approximation based on linear expansions in parametric functions. // Proceeding of the 8th International Conference "Intelligent Information Processing Cyprus, 2011 – p.
- [14] Cressie N. A. C. Statistics for Spatial Data. – Wiley, 1993.
- [15] Nielsen H.B. DACE. A Matlab Kriging Toolbox. [http://www2.imm.dtu.dk/~thicksim\\$hbndace/](http://www2.imm.dtu.dk/~thicksim$hbndace/)
- [16] GDR MASCOT-NUM Toy Functions benchmark. <http://gdr-mascotnum.math.cnrs.fr/data2/benchmarks/jm.pdf>
- [17] Lappeenranta University of Technology: evolutionary computation pages - the function testbed. <http://www.it.lut.fi/ip/evo/functions/functions.html>
- [18] Dolan E. D., More J. J. Benchmarking optimization software with performance profiles // Mathematical Programming, Ser. A 91, 2002 – pp. 201–213