

# Многомерная интерполяция и аппроксимация на основе теории случайных функций

Материал из MachineLearning.

## Содержание

- 1 Аннотация
- 2 Вступление
- 3 Подход к многомерной интерполяции и аппроксимации на основе теории случайных функций.
  - 3.1 Введение.
  - 3.2 Многомерная интерполяция.
  - 3.3 Дисперсия случайной функции. Множество реализаций, удовлетворяющих узлам интерполяции.
  - 3.4 Многомерная аппроксимация.
  - 3.5 Выводы.
- 4 Метод машинного обучения.
- 5 Демонстрация в среде Matlab.
- 6 Литература

## Аннотация

В статье исследуется задача многомерной интерполяции и аппроксимации как вариант машинного обучения. Показано, что эти задачи могут быть решены на основе теории случайных функций. Предложен метод машинного обучения, гарантирующий получение оптимального результата с точки зрения рассматриваемого математического аппарата случайных функций. Метод прост в реализации и позволяет сводить разнообразные задачи машинного обучения к решению систем линейных уравнений.

## Вступление

Данную статью условно можно поделить на две части.

- Первая часть посвящена использованию основ теории случайных функций применительно к задачам многомерной интерполяции и аппроксимации, а также машинному обучению и их теоретическому обоснованию. Цель теоретической части показать, что машинное обучение в его парадигме “обучения с учителем”, задачи многомерной интерполяции и аппроксимации, могут быть обобщены на основе теории случайных функций.
- Во второй части изложены практические выводы из положений первой части в виде законченного метода машинного обучения (метода многомерной интерполяции и аппроксимации). Если читателя интересуют сугубо практические вопросы, Вы можете перейти сразу к изложению метода.
- Предложенный метод позволяет получить точное решение задачи многомерной интерполяции или аппроксимации (“обучение с учителем”) гарантирующее оптимальность (при определенных минимальных допущениях, указанных в теоретической части). Метод достаточно прост и сводится к системе линейных уравнений, что может у читателя не знакомого с вопросом без изучения теоретической части вызвать подозрения в работоспособности или обобщающей способности метода. Но смею Вас уверить, что столь простая математическая конструкция в данном случае никак не ограничивает возможности. Если постараться объяснить кратко, то способности метода обеспечивает “специальная функция”, полученная в теоретической части, применение которой в системе линейных уравнений гарантированно обеспечивает оптимальное, с точки зрения аппарата случайных функций, решение. Использование данной функции позволяет сводить задачи многомерной интерполяции и аппроксимации или самые разнообразные задачи машинного обучения (с определенными допущениями) к решению системы линейных уравнений, гарантируя оптимальность полученного решения, отсутствие переобучения, осциллирующий интерполянта и других нежелательных эффектов (если говорить более строго, то в реальных вычислениях используется лишь приближение теоретической “идеальной” функции в используемой части спектра, функция, которую можно записать алгебраически).

## Подход к многомерной интерполяции и аппроксимации на основе теории случайных функций.

### Введение.

Цель данной статьи показать, что машинное обучение в его парадигме “обучения с учителем”, задачи многомерной интерполяции и аппроксимации, могут быть обобщены на основе теории случайных функций.

Использование математического аппарата случайных функций удобно для теоретического обоснования методов машинного обучения, а понятие “вероятность реализации” может служить критерием для обоснования обобщающих способностей метода обучения.

Ниже будет показано, что использование понятия случайной функции и связанных с ней теоретических построений позволяет получать точные решения задач машинного обучения как задач многомерной интерполяции (или аппроксимации) гарантируя оптимальность решения с точки зрения критерия вероятности. Существует способ при минимальных допущениях получить оптимальное решение задачи машинного обучения путем решения системы линейных уравнений (с доказательством его единственности и оптимальности). Это позволяет получить точное решение, что снимает проблему поиска глобального оптимума и выбора топологии сети, которые имеют место во многих существующих итеративных алгоритмах “обучения с учителем” искусственных нейронных сетей.

Теорией случайных процессов называется математическая наука, изучающая закономерности случайных явлений в динамике их развития. Аналогично тому, как понятие случайного процесса является обобщением понятия случайной величины, случайная функция является обобщением понятия случайного процесса, поскольку функции могут зависеть не от одного, а от нескольких аргументов.

Реализациями случайной функции являются функции, вид которых она принимает при наблюдении или, например, при проведении опыта. Сама же случайная функция представляет собой множество своих реализаций, с соответствующим распределением вероятностей. Для любой функции, проявления которой мы можем наблюдать, любую функцию которую мы можем в каком-либо смысле рассматривать, всегда можно поставить в соответствие случайную функцию, реализацией которой мы можем ее считать без каких-либо дополнительных предположений или допущений. Даже если мы рассматриваем какую-то конкретную функцию, например, вида  $y=2x$ , то и в этом случае мы можем ее считать частным случаем случайной функции, множество реализаций которой состоит из одного элемента с вероятностью реализации равной единице.

При решении задачи многомерной интерполяции или аппроксимации или при машинном обучении по набору примеров требуется связать с помощью некоторой функции значения входных и выходных переменных. Решая подобные задачи, требуется найти нужные функции, удовлетворяющие обучающим

векторам, если это машинное обучение, либо узлам интерполяции в случае многомерной интерполяции, притом что таких функций, удовлетворяющих данным, может быть бесконечное множество.

При создании любого метода решения, выбор всегда будет осуществляться на основе определенных положений. Рассмотрим задачу машинного обучения либо многомерную интерполяцию или аппроксимацию. Предположим, что искомую нами функцию можно рассматривать как реализацию некоторой случайной функции, о которой (о реализации) несет информацию рассматриваемый набор обучающих векторов (либо узлов интерполяции).

При таком подходе естественным критерием для поиска требуемой функции будет вероятность ее как реализации. Поиск нужной функции в этом случае превращается в поиск наиболее вероятной из реализаций случайной функции, которые удовлетворяют обучающим примерам либо узлам интерполяции в случае интерполяции.

Эта идея является основой предлагаемого на рассмотрение в данной работе метода. Для любой наблюдаемой в окружающем мире функции всегда можно поставить в соответствие случайную функцию, реализацией которой она будет являться, поэтому считая искомую функцию реализацией случайной функции, не уточняя ее характеристик, мы пока еще не делаем никаких дополнительных предположений или допущений, но в то же время можем использовать математический аппарат теории случайных функций.

Адекватность такого подхода, его область применимости, будет целиком определяться тем, какими конкретными характеристиками мы наделим априорную случайную функцию, используя которую мы будем решать ту или иную задачу.

## Многомерная интерполяция.

Рассмотрим задачу многомерной интерполяции, предполагая, что узлы интерполяции принадлежат реализации некоторой случайной функции.

- Предположим, что функцию (1) можно считать реализацией некоторой случайной функции  $F(x)$  которую можно записать как (2):

$$y = f(x), \quad x \in R^n, \quad y \in R \quad (1)$$

$$F(x) = \sum_{i=1}^m V_i \varphi_i(x) + f_m(x), \quad (2)$$

где  $\varphi_i(x)$  - координатные функции,  $V_i$  - случайные величины,  $f_m(x)$  - функция математического ожидания

- Под координатными функциями в (2) мы можем понимать любые произвольные функции. Условие заключается в том, что существуют какие-то координатные функции, такие, что мы можем через них записать случайную функцию вида (2). Количество слагаемых в (2) может быть произвольным, т.е. полагаем, что  $m$  может быть любым, в том числе и равным бесконечности (как вариант, сумма в (2) может быть обобщена до интеграла). В качестве частного случая (2) можно рассматривать, например, разложение в бесконечный ряд синусов и косинусов, когда реализациями данной случайной функции может быть множество всех непрерывных вещественных функций.
- Предполагаем, что в (2) случайные величины  $V_i$  независимы, распределены по нормальному закону, с математическими ожиданиями равными нулю и дисперсией равной единице.
- Потенциально требование именно к нормальному закону на данном необязательно (но может быть использовано в дальнейшем, например, при анализе и отображении подмножеств реализаций). Для выполнения следующих нескольких преобразований достаточно было бы потребовать, чтобы совместный закон распределения вероятностей для случайных величин из (2) был симметричным относительно начала координат и монотонно убывал с увеличением расстояний от начала координат.
- Требования в (2) для случайных величин к математическому ожиданию равному нулю и дисперсии равной единице сделаны для удобства дальнейших преобразований. Если они не выполняются, можно сделать преобразование, получив снова выражение (2), но чтобы эти требования были выполнены. Предположим, что для какой-либо случайной величины дисперсия не равна единице, тогда мы можем заменить ее и одновременно заменить соответствующую ей координатную функцию (умножив ее на коэффициент), снова получив выражение вида (2). Если же математическое ожидание какой-либо случайной величины в (2) не равно нулю, тогда можно выполнить преобразование, изменив функцию математического ожидания таким образом, чтобы опять свести к виду (2).
- Рассмотрим подробнее функцию  $f_m(x)$ . В частном случае она может быть равна нулю или константе.
- Если эта функция заранее известна, то мы можем исключить ее из дальнейших преобразований, осуществив вычитание ее значений из узлов интерполяции и одновременно удалив из выражения (2). После того как интерполяция по модифицированным узлам будет выполнена мы можем снова прибавить ее к полученному решению для получения окончательного результата.
- Если эта функция заранее не известна, тогда мы можем принять ее за некую реализацию еще одной случайной функции которую можно снова выразить через выражение (2), и дальше повторять эту процедуру в случае необходимости.

Исключив функцию математического ожидания, получим выражение случайной функции:

$$F(x) = \sum_{i=1}^m V_i \varphi_i(x) \quad (3)$$

- Обозначим через  $v_1, v_2, \dots, v_m$  конкретные значения случайных величин из (3), которые они приняли для какой-либо реализации. Тогда плотность распределения вероятностей реализаций в пространстве параметров  $v_1, v_2, \dots, v_m$  (т.е. вероятность реализации случайной функции вида (3), когда случайные величины  $V_1, V_2, \dots, V_m$  принимают некие конкретные значения  $v_1, v_2, \dots, v_m$ ) запишется в виде (4):

$$P(v_1, v_2, \dots, v_m) = \frac{1}{(2\pi)^{\frac{m}{2}}} e^{-\frac{v_1^2 + v_2^2 + \dots + v_m^2}{2}} \quad (4)$$

Пусть последовательности

$$\begin{matrix} x_1, x_2, \dots, x_k & (x_i \in R^n) \\ y_1, y_2, \dots, y_k & (y_i \in R) \end{matrix} \quad (5)$$

представляют собой координаты и соответствующие им значения узлов интерполяции.

- Считая, что узлы интерполяции принадлежат одной из реализаций случайной функции (3), которой соответствует вектор  $v_1, v_2, \dots, v_m$ , их можно выразить в виде системы уравнений (6):

$$\begin{cases} \sum_{i=1}^m v_i \varphi_i(x_1) = y_1 \\ \sum_{i=1}^m v_i \varphi_i(x_2) = y_2 \\ \vdots \\ \sum_{i=1}^m v_i \varphi_i(x_k) = y_k \end{cases} \quad (6)$$

- Используя приведенные выше выражения, можно сказать, что наиболее вероятной реализацией случайной функции вида (3) соответствующей узлам интерполяции будет та, при которой функция плотности вероятности (4) принимает максимальное значение с учетом системы ограничений в виде равенств (6).

Поиск максимума функции (4), равносильно поиску минимума функции (7):

$$v_1^2 + v_2^2 + \dots + v_m^2 \rightarrow \min \quad (7)$$

- Отсюда можно сделать вывод, что решение задачи многомерной интерполяции как поиск наиболее вероятной реализации функции (2) или (3) удовлетворяющей узлам интерполяции, сводится к задаче квадратичного программирования с ограничениями в виде равенств.

Для поиска минимума (7) при системе ограничений (6), воспользуемся методом множителей Лагранжа.

Найдем Функцию Лагранжа:

$$L(v_1, \dots, v_m, \lambda_1, \dots, \lambda_k) = v_1^2 + v_2^2 + \dots + v_m^2 + \sum_{j=1}^k \lambda_j \left( \sum_{i=1}^m v_i \varphi_i(x_j) - y_j \right) \quad (8)$$

Обозначим вектором  $V^* (v_1^*, v_2^*, \dots, v_m^*)$  решение задачи.

Из условия  $\frac{dL}{d\lambda_j} = 0$  опять получим уравнения (6).

Из условия  $\frac{dL}{dv_i} = 0$  получим:

$$v_i^* = -0.5 \sum_{j=1}^k \lambda_j \varphi_i(x_j), \quad i = 1, \dots, m \quad (9)$$

Обозначим последовательности значений координатных функций  $(\varphi_1(x_i), \varphi_2(x_i), \dots, \varphi_m(x_i))$  как вектора  $L_i$ ,  $i = 1, \dots, k$ ,  $L_i \in R^m$ .

Введем коэффициенты  $q_1, q_2, \dots, q_k$ , обозначив  $q_i = -0.5 \lambda_i$ ,  $i = 1, \dots, k$ . Тогда (9) можно записать через новые обозначения в виде линейной комбинации векторов (10):

$$V^* = q_1 L_1 + q_2 L_2 + \dots + q_k L_k \quad (10)$$

- Выражение (10) можно интерпретировать с геометрической точки зрения. Поиск минимума функции (7), с системой ограничений (6), равносильно опусканию перпендикуляра из начала координат на гиперплоскость, образуемую пересечением гиперплоскостей, заданных в виде уравнений (6) (подставив в (6) значения координатных функций, мы можем понимать их как задание гиперплоскостей и их пересечения в  $n$ -мерном пространстве), следовательно,  $V^*$  будет являться линейной комбинацией их перпендикуляров.

Систему ограничений (6) можно записать через скалярные произведения векторов (11):

$$\begin{cases} L_1 V^* = y_1 \\ L_2 V^* = y_2 \\ \vdots \\ L_k V^* = y_k \end{cases} \quad (11)$$

Заменяя  $V^*$  используя (10) получаем систему уравнений:

$$\begin{cases} q_1 L_1 L_1 + q_2 L_1 L_2 + \dots + q_k L_1 L_k = y_1 \\ q_1 L_2 L_1 + q_2 L_2 L_2 + \dots + q_k L_2 L_k = y_2 \\ \vdots \\ q_1 L_k L_1 + q_2 L_k L_2 + \dots + q_k L_k L_k = y_k \end{cases} \quad (12)$$

Обозначим как вектор  $L_*$  последовательность значений  $(\varphi_1(x), \varphi_2(x), \dots, \varphi_m(x))$ ,  $x \in R^n$ .

Тогда наиболее вероятную реализацию  $y = f^*(x)$  случайной функции (3) можно выразить как выражение (13):

$$f^*(x) = \sum_{i=1}^m v_i^* \varphi_i(x) = L_* V^* = q_1 L_* L_1 + q_2 L_* L_2 + \dots + q_k L_* L_k \quad (13)$$

Рассмотрим более подробно скалярное произведение  $L_i L_j$ :

$$(14)$$

$$L_i L_j = \sum_{t=1}^m \varphi_t(x_i) \varphi_t(x_j) = K f(x_i, x_j), \quad x_i, x_j \in R^n$$

что является не чем иным как каноническим разложением корреляционной функции.

Тогда систему уравнений (12) можно записать в виде (15):

$$\begin{cases} q_1 K f(x_1, x_1) + q_2 K f(x_1, x_2) + \dots + q_k K f(x_1, x_k) = y_1 \\ q_1 K f(x_2, x_1) + q_2 K f(x_2, x_2) + \dots + q_k K f(x_2, x_k) = y_2 \\ \vdots \\ q_1 K f(x_k, x_1) + q_2 K f(x_k, x_2) + \dots + q_k K f(x_k, x_k) = y_k \end{cases} \quad (15)$$

а выражение (13), для наиболее вероятной реализации, в виде (16):

$$f^*(x) = q_1 K f(x, x_1) + q_2 K f(x, x_2) + \dots + q_k K f(x, x_k) \quad (16)$$

- Таким образом, можно сделать вывод, что наиболее вероятной реализацией случайной функции (3), удовлетворяющей узлам интерполяции, будет линейная комбинация сечений ее корреляционных функций.
- Следовательно, зная корреляционную функцию и решив систему уравнений (15), мы получим наиболее вероятную реализацию (16) в рамках данной случайной функции.
- Если же теперь рассмотреть интерполяцию как вариант метода машинного обучения то полученный результат можно легко обобщить на вариант с несколькими выходами, поскольку левая часть уравнений (15) для разных выходных переменных будет одинаковой (если конечно зависимость между входами и разными выходами можно рассматривать, пользуясь одной случайной функцией). В этом случае можно вычислить единую обратную матрицу уравнений (15), которую можно использовать для интерполяции при любых возможных вариантах значений выходных переменных.
- Однако чтобы осуществить подобные расчеты, необходимо знать, какой корреляционной функцией пользоваться (или знать все координатные функции, что будет достаточно для вычисления корреляционной функции), а также знать функцию математического ожидания в (2) либо корректно преобразовать в (3).

Рассмотрим случай, когда нет априорной информации о том, в рамках какой случайной функции можно решать задачу интерполяции. Устранение неопределенности осуществляется выбором корреляционной функции, сделав определенные предположения.

- Предположим, что изначально потенциальными вариантами интерполанта (искомой функции) могут быть любые непрерывные вещественные функции и об их априорном распределении вероятностей между собой не известно. Предположим, что в этом случае, априорно любые точки, а также направления в пространстве интерполяции равнозначны между собой. Поэтому потребуем, чтобы интерполяция не зависела от выбора системы отсчета узлов интерполяции, от операций ее поворота, положения начала координат и масштаба. Т.е. если мы в пространстве интерполяции введем новую систему отсчета (оси координат, путем операций поворота и сдвига исходных), найдем координаты узлов интерполяции в этой новой системе отсчета и выполним интерполяцию, то полученная функция, при переводе ее в исходную систему отсчета должна совпасть с функцией, полученной по исходным узлам. Потребуем выполнения того же свойства и при переходе к новой системе отсчета путем одновременного изменения масштаба по всем координатам (умножив на единый коэффициент значения всех входных и выходных переменных).
- Эти требования можно обеспечить, введя набор положений об априорных вероятностях существования той или иной реализации рассматриваемой случайной функции. Примем априорно равновероятными реализации преобразующиеся друг в друга с помощью операций сдвига поворота или являющиеся подобными друг другу при изменении масштаба. Это допущение обеспечивает выполнение свойств интерполяции при переходе к другой системе отсчета (описанные выше). Действительно, если это допущение выполняется, то при переходе к другой системе отсчета (любым описанным выше способом) появляется возможность отобразить множество реализаций на самого себя таким образом, что мы получим эквивалентную задачу.
- Обозначенных выше требований, не смотря на то, что они и так выглядят естественными в условиях априорной неопределенности, достаточно, чтобы однозначно вычислить корреляционную функцию, необходимую для выражений (15) и (16).

Обозначим сформулированные требования в виде набора положений:

1. Плотность вероятности реализаций, преобразующихся друг в друга параллельным переносом системы координат должна быть одинаковой.

Предположим, что мы имеем две реализации  $f_1(x)$  и  $f_2(x)$  такие, что:

$$\forall x \quad f_1(x) = f_2(x+t), \quad x, t \in R^n, \quad t = const \quad (17)$$

Тогда плотность вероятности реализаций  $f_1(x)$  и  $f_2(x)$  должна быть равна.

2. Плотность вероятности реализаций, преобразующихся друг в друга одинаковым изменением масштаба по всем направлениям должна быть одинаковой.

Предположим, что мы имеем две реализации  $f_1(x)$  и  $f_2(x)$  такие, что:

$$\forall x \quad f_1(x) = k f_2(x/k), \quad x \in R^n$$

где  $k$  - некоторый коэффициент

(18)

Тогда плотность вероятности реализаций  $f_1(x)$  и  $f_2(x)$  должна быть равна.

3. Плотность вероятности реализаций, преобразующихся друг в друга поворотом системы координат должна быть одинаковой. Предположим, что мы имеем две реализации  $f_1(x)$  и  $f_2(x)$  такие, что:

$$\forall x_1 \quad f_1(x_1) = f_2(x_2), \quad x_2 = A x_1, \quad x_1, x_2 \in R^n$$

где  $A$  - некоторая матрица поворота

(19)

Тогда плотность вероятности реализаций  $f_1(x)$  и  $f_2(x)$  должна быть равна.

- Выполнение первого пункта означает, что случайная функция будет стационарной. В этом случае ее можно записать в виде многомерного спектрального разложения. Обозначим  $S(\omega)$  ее спектральную плотность (неотрицательная вещественная функция, симметричная относительно любых частот  $\omega$  и  $-\omega$ ).

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{R^n} S^*(\omega) \sqrt{S(\omega)} e^{ix\omega} d\omega$$

$x, \omega \in R^n$ ,  $x\omega$  - их скалярное произведение

- В выражении (20)  $S^*(\omega)$  - реализация комплексной случайной функции, каждое значение которой является независимой случайной величиной (кроме дополнительного условия, о котором ниже) имеющей нормальное распределение, единичную дисперсию и математическое ожидание равно нулю, как для вещественной, так и мнимой части. Выражение (20) можно считать частным случаем (3), в котором  $m = \infty$ , роль координатных функций выполняет  $\sqrt{S(\omega)} e^{ix\omega}$ , а аналогом значений случайных величин выступает  $S^*(\omega)$ . Функцию можно считать разложением реализации по базису функций пропорциональному  $\sqrt{S(\omega)} e^{ix\omega}$ .
- Наложим дополнительное условие на  $S^*(\omega)$  (а соответственно, на случайную функцию), поскольку мы рассматриваем в данный момент интерполяцию вещественных функций, то ее значения для любых значений частот  $\omega$  и  $-\omega$  должны быть комплексно-сопряженными. Чтобы отразить это условие поделим пространство частот  $R^n$  на две части  $R_1^n$  и  $R_2^n$ . Осуществить это разделение нужно таким образом, чтобы любые частоты  $\omega$  и  $-\omega$  лежали разных частях (например, любой гиперплоскостью, проходящей через начало координат). Тогда  $S^*(\omega)$  можно записать как:

$$S^*(\omega) = \begin{cases} R^*(\omega) + iI^*(\omega), & \omega \in R_1^n \\ R^*(\omega) - iI^*(\omega), & \omega \in R_2^n \end{cases} \quad (21)$$

- Где  $R^*(\omega)$  и  $I^*(\omega)$  можно рассматривать как реализации вещественных случайных функций, каждое значение которых одинаково для частот  $\omega$  и  $-\omega$ , является независимой случайной величиной, имеющей нормальное распределение, математическое ожидание равно нулю и дисперсию равную единице.

При выполнении условия на вещественность реализаций случайную функцию (20) можно записать как (22):

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \left( \int_{R_1^n} (R^*(\omega) + iI^*(\omega)) \sqrt{S(\omega)} e^{ix\omega} d\omega + \int_{R_2^n} (R^*(\omega) - iI^*(\omega)) \sqrt{S(\omega)} e^{ix\omega} d\omega \right) \quad (22)$$

Поскольку  $R_1^n$  и  $R_2^n$  симметричны относительно  $\omega$  и  $-\omega$ , то во втором интеграле в (22) можно перейти к области интегрирования  $R_1^n$  заменив знак перед частотой. В этом случае получим:

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \left( \int_{R_1^n} (R^*(\omega) + iI^*(\omega)) \sqrt{S(\omega)} e^{ix\omega} d\omega + \int_{R_1^n} (R^*(-\omega) - iI^*(-\omega)) \sqrt{S(-\omega)} e^{-ix\omega} d\omega \right) \quad (23)$$

Выполнив далее преобразования, получим:

$$\begin{aligned} f(x) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) (e^{i\omega x} + e^{-i\omega x}) + iI^*(\omega) (e^{i\omega x} - e^{-i\omega x})) d\omega = \\ &= \frac{2}{(2\pi)^{\frac{n}{2}}} \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) \cos(\omega x) - I^*(\omega) \sin(\omega x)) d\omega \end{aligned} \quad (24)$$

- Поскольку каждое значение  $S^*(\omega)$  (для пар частот  $\omega$  и  $-\omega$ ) является независимой комплексной случайной величиной, с математическим ожиданием равным нулю и дисперсией равной единице, мы можем получить (25) - аналог выражения (7), минимизация которого при выполнении ограничений на узлы интерполяции обеспечит наиболее вероятную реализацию.

$$\int_{R_1^n} |S^*(\omega)|^2 d\omega \rightarrow \min \quad (25)$$

Для двух любых реализаций, имеющих одинаковую априорную вероятность значение (25) должно иметь одинаковое значение.

Выражение (25) можно записать как (26):

$$\int_{R_1^n} (R^*(\omega)^2 + I^*(\omega)^2) d\omega \rightarrow \min \quad (26)$$

Пусть последовательность  $x_1, x_2, \dots, x_k$  ( $x_i \in R^n$ ) и соответствующие им  $y_1, y_2, \dots, y_k$  ( $y_i \in R$ ) - представляют собой узлы интерполяции. Тогда используя полученное выражение (24) ограничения на эти узлы будут представлять собой систему уравнений:

$$\begin{cases} \frac{2}{(2\pi)^{\frac{n}{2}}} \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) \cos(\omega x_1) - I^*(\omega) \sin(\omega x_1)) d\omega = y_1 \\ \frac{2}{(2\pi)^{\frac{n}{2}}} \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) \cos(\omega x_2) - I^*(\omega) \sin(\omega x_2)) d\omega = y_2 \\ \vdots \\ \frac{2}{(2\pi)^{\frac{n}{2}}} \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) \cos(\omega x_k) - I^*(\omega) \sin(\omega x_k)) d\omega = y_k \end{cases} \quad (27)$$

Поиск наиболее вероятной реализации будет соответствовать нахождению функций  $R^*(\omega)$  и  $I^*(\omega)$  таких, при которых выполняются условия (27) и

минимизируется значение выражения (26).

Дальнейшие рассуждения можно выполнить аналогично (8) – (16).  
Функция Лагранжа примет вид:

$$\int_{R_1^n} (R^*(\omega)^2 + I^*(\omega)^2) + \frac{2}{(2\pi)^{\frac{n}{2}}} \sum_{i=1}^k \lambda_i \left( \int_{R_1^n} \sqrt{S(\omega)} (R^*(\omega) \cos(\omega x_i) - I^*(\omega) \sin(\omega x_i)) d\omega - y_i \right) \quad (28)$$

При дифференцировании (28) по  $\lambda_i$  можно снова получить систему (27).

Поскольку любое значение функций  $R^*(\omega)$  и  $I^*(\omega)$  для произвольного  $\omega$  выражения (28) можно считать независимой переменной, то соответственно по каждому из них можно выполнить независимое дифференцирование. В этом случае получим выражения:

$$R^*(\omega) = -\frac{\sqrt{S(\omega)}}{(2\pi)^{\frac{n}{2}}} \sum_{i=1}^k \lambda_i \cos(\omega x_i) \quad (29)$$

$$I^*(\omega) = \frac{\sqrt{S(\omega)}}{(2\pi)^{\frac{n}{2}}} \sum_{i=1}^k \lambda_i \sin(\omega x_i) \quad (30)$$

Подставив решения (29) и (30) в систему (24) получим выражение наиболее вероятной реализации:

$$f(x) = -\frac{2}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R_1^n} S(\omega) (\cos(\omega x_i) \cos(\omega x) + \sin(\omega x_i) \sin(\omega x)) d\omega \quad (31)$$

Выражение (31) преобразуется в (32):

$$f(x) = -\frac{2}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R_1^n} S(\omega) \cos(\omega(x_i - x)) d\omega \quad (32)$$

Поскольку подынтегральное выражение в (32) инвариантно к изменению  $\omega$  на  $-\omega$ , то мы можем перейти целиком в область интегрирования  $R^n$ :

$$f(x) = -\frac{1}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R^n} S(\omega) \cos(\omega(x_i - x)) d\omega \quad (33)$$

Введем коэффициенты:

$$q_i = -\frac{\lambda_i}{(2\pi)^n} \quad (34)$$

Выражение для корреляционной функции (на самом деле, из-за введенных коэффициентов, это будет функция пропорциональная корреляционной):

$$K_f(x_1, x_2) = \int_{R^n} S(\omega) \cos(\omega(x_1 - x_2)) d\omega, \quad x_1, x_2 \in R^n \quad (35)$$

Как видно из (35) спектральная плотность случайной функции играет роль спектра для ее корреляционной функции. Сама же корреляционная функция превращается в автокорреляционную, что и следовало ожидать из условия стационарности:

$$K_f(\tau) = \int_{R^n} S(\omega) \cos(\omega\tau) d\omega, \quad \tau \in R^n \quad (36)$$

- Отсутствие в (35 - 36) синусной составляющей, говорит о том, что автокорреляционная функция должна быть симметрична относительно нуля, что можно было ожидать. Соответственно в случае, если спектральная плотность случайной функции будет известна, появляется потенциальная возможность вычислить автокорреляционную функцию.

Тогда наиболее вероятную реализацию можно выразить:

$$f^*(x) = q_1 K_f(x - x_1) + q_2 K_f(x - x_2) + \dots + q_k K_f(x - x_k) \quad (37)$$

Найдем систему уравнений для вычисления коэффициентов в (37). Подставив (29) и (30) в (27) получим:

$$\begin{cases} -\frac{2}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R_1^n} S(\omega) (\cos(\omega x_i) \cos(\omega x_1) + \sin(\omega x_i) \sin(\omega x_1)) d\omega = y_1 \\ -\frac{2}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R_1^n} S(\omega) (\cos(\omega x_i) \cos(\omega x_2) + \sin(\omega x_i) \sin(\omega x_2)) d\omega = y_2 \\ \vdots \\ -\frac{2}{(2\pi)^n} \sum_{i=1}^k \lambda_i \int_{R_1^n} S(\omega) (\cos(\omega x_i) \cos(\omega x_k) + \sin(\omega x_i) \sin(\omega x_k)) d\omega = y_k \end{cases} \quad (38)$$

Используя (34) и (36) в (38) получим систему уравнений для вычисления коэффициентов для (37) (аналог (15)):

(39)

$$\begin{cases} q_1 K_f(x_1-x_1)+q_2 K_f(x_1-x_2)+...+q_k K_f(x_1-x_k)=y_1 \\ q_1 K_f(x_2-x_1)+q_2 K_f(x_2-x_2)+...+q_k K_f(x_2-x_k)=y_2 \\ \vdots \\ q_1 K_f(x_k-x_1)+q_2 K_f(x_k-x_2)+...+q_k K_f(x_k-x_k)=y_k \end{cases}$$

- Таким образом, мы получили выражения (37) и (39) - аналоги выражение (15) и (16), а также выражение (36) для корреляционной функции. Для раскрытия неопределенности остается определить спектральную плотность  $S(\omega)$  случайной функции.
- Обозначим как  $S_f(\omega)$  спектральное разложение Фурье реализаций рассматриваемой нами случайной функции. Из выражения (20) - спектрального разложения случайной функции видно, что спектр реализаций будет равен:

$$S_f(\omega) = S^*(\omega)\sqrt{S(\omega)} \quad (40)$$

Теперь вернемся ко второму пункту (выражение (18)).

Обозначим за  $S_f^1(\omega)$  спектр некоторой функции  $f_1(x)$ , а  $S_f^2(\omega)$  - спектр некоторой функции  $f_2(x)$ . Будем считать, что

$$f_2(x) = k f_1(x/k), \quad x \in \mathbb{R}^n \quad (41)$$

т.е. функции подобны, но совпадают в масштабе  $k:1$ . Поскольку

$$f_1(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} S_f^1(\omega) e^{ix\omega} d\omega \quad (42)$$

$x, \omega \in \mathbb{R}^n, \quad x\omega$  - их скалярное произведение

Получим

$$\begin{aligned} f_2(x) &= k f_1(x/k) = k \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} S_f^1(\omega) e^{\frac{ix\omega}{k}} d\omega = \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} k^{(n+1)} S_f^1(\omega) e^{\frac{ix\omega}{k}} d\omega = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} k^{(n+1)} S_f^1(\omega k) e^{ix\omega} d\omega \end{aligned} \quad (43)$$

Откуда можно вывести соотношение спектров подобных функций:

$$S_f^2(\omega) = k^{n+1} S_f^1(\omega k) \quad (44)$$

Для обеих рассмотренных функций значение выражения (25) должно быть одинаковым.

Функцию  $S^*(\omega)$  для (25) можно выразить через спектры реализаций и спектральную плотность случайной функции:

$$S^*(\omega) = \frac{S_f(\omega)}{\sqrt{S(\omega)}} \quad (45)$$

Приравнявая выражение (25) для двух рассмотренных выше реализаций (причем, поскольку под знаком интегрирования берется значение по модулю, мы можем интегрировать по  $\mathbb{R}^n$ ), а также используя выражения (44) и (45) получаем:

$$\begin{aligned} \int_{\mathbb{R}^n} \frac{|S_f^1(\omega)|^2}{S(\omega)} d\omega &= \int_{\mathbb{R}^n} \frac{|S_f^2(\omega)|^2}{S(\omega)} d\omega = \int_{\mathbb{R}^n} \frac{k^{2n+2} |S_f^1(\omega k)|^2}{S(\omega)} d\omega = \\ &= \int_{\mathbb{R}^n} \frac{k^{n+2} |S_f^1(\omega k)|^2}{S(\omega)} d(\omega k) = \int_{\mathbb{R}^n} \frac{k^{n+2} |S_f^1(\omega)|^2}{S(\omega/k)} d\omega \end{aligned} \quad (46)$$

Из выражения (46), сравнивая начало и конец полученных равенств можно получить соотношение для спектральной плотности:

$$\frac{S(\omega/k)}{S(\omega)} = k^{n+2} \quad (47)$$

Чтобы выполнялось третье положение (19), автокорреляционная функция должна обладать радиальной симметрией. Соответственно радиальной симметрией должна обладать и спектральная плотность. Тогда из (47) можно получить выражение для спектральной плотности:

$$S(\omega) = \alpha |\omega|^{-(n+2)}, \quad \omega \in \mathbb{R}^n \quad (48)$$

где  $\alpha$  - некоторый коэффициент

или(то же самое):

$$S(\omega_1, \omega_2, \dots, \omega_n) = \alpha (\omega_1^2 + \omega_2^2 + \dots + \omega_n^2)^{-\frac{n+2}{2}} \quad (49)$$

$\omega_1, \omega_2, \dots, \omega_n \in \mathbb{R}$

- Таким образом, в выражении (48-49) мы получили искомую спектральную плотность случайной функции. При использовании корреляционной функции с таким спектром и решении системы линейных уравнений (39) будет гарантированно получена наиболее вероятная реализация случайной функции (если брать в качестве основы положения (17) – (19)).
- Коэффициент  $\alpha$  в (48) – (49) может быть любым, поскольку его изменение одновременно увеличит или уменьшит все коэффициенты левой части системы уравнений (39) и в уравнении (37) в одинаковое количество раз, а соответственно и коэффициенты решения, которые при подстановке в (37) дадут ту же самую функцию.
- Для непосредственно вычисления выражения функции со спектром (49) можно воспользоваться свойством симметричности функции. В этом случае

для упрощения можно сначала рассмотреть одномерный вариант, после чего получить многомерный вариант, воспользовавшись требованием симметричности.

В одномерном случае требуемая спектральная плотность будет:

$$S(\omega) = |\omega|^{-3} \quad (50)$$

Тогда выражение для автокорреляционной функции (36) в одномерном случае примет вид:

$$K(\tau) = \int_{-\infty}^{\infty} \frac{\cos(\omega\tau)}{|\omega|^3} d\omega \quad (51)$$

Выражение (51) можно записать как:

$$K(\tau) = \lim_{\omega_0 \rightarrow +0} (2 \int_{\omega_0}^{\infty} \frac{\cos(\omega\tau)}{\omega^3} d\omega) \quad (52)$$

Последовательно интегрируя по частям, получим:

$$K(\tau) = \lim_{\omega_0 \rightarrow +0} ((-\frac{\cos(\omega\tau)}{\omega^2} + \tau^2 \frac{\sin(\omega\tau)}{\omega\tau})|_{\omega_0}^{\infty}) - \tau^2 \int_{\omega_0}^{\infty} \frac{\cos(\omega\tau)}{\omega} d\omega \quad (53)$$

В качестве последнего правого слагаемого в (53) был получен интегральный косинус. Рассмотрим его более подробно:

$$-\int_{\omega_0}^{\infty} \frac{\cos(\omega\tau)}{\omega} d\omega = -\int_{\omega_0\tau}^{\infty} \frac{\cos(\omega\tau)}{\omega\tau} d(\omega\tau) = -\int_{\omega_0|\tau|}^{\infty} \frac{\cos(\omega|\tau|)}{\omega|\tau|} d(\omega|\tau|) \quad (54)$$

Интегральный косинус может быть раскрыт:

$$-\int_{\omega_0|\tau|}^{\infty} \frac{\cos(\omega|\tau|)}{\omega|\tau|} d(\omega|\tau|) = \gamma + \ln(\omega_0|\tau|) + \int_0^{\omega_0|\tau|} \frac{\cos(t)-1}{t} dt \quad (55)$$

где  $\gamma$  - постоянная Эйлера

Используя (53) и (55) получим:

$$\begin{aligned} K(\tau) &= \lim_{\omega_0 \rightarrow +0} (\tau^2 \gamma + \tau^2 \ln(\omega_0|\tau|) + \int_0^{\omega_0|\tau|} \frac{\cos(t)-1}{t} dt + \frac{\cos(\omega_0\tau)}{\omega_0^2} - \tau^2 \frac{\sin(\omega_0\tau)}{\omega_0\tau}) = \\ &= \lim_{\omega_0 \rightarrow +0} (\tau^2 \ln(\omega_0 e^{\gamma-1} |\tau|) + \frac{\cos(\omega_0\tau)}{\omega_0^2}) \end{aligned} \quad (56)$$

Таким образом, получаем результат:

$$K(\tau) = \int_{-\infty}^{\infty} \frac{\cos(\omega\tau)}{|\omega|^3} d\omega = \lim_{\omega_0 \rightarrow +0} (\tau^2 \ln(\omega_0 e^{\gamma-1} |\tau|) + \frac{\cos(\omega_0\tau)}{\omega_0^2}) \quad (57)$$

- Хотя в реальных вычислениях невозможно использовать  $\omega_0 = 0$ , можно взять ее значение сколь угодно близкое к нулю, тем самым приблизив ее спектр к выражению (50). Для вычислений вместо (57) может быть использован более упрощенный вариант (58).
- Учитывая, что коэффициент  $\alpha$  в (48)-(49) можно взять любым (возьмем равным 0.5), а также воспользовавшись свойством симметричности, получаем вариант функции для многомерного случая:

$$K_f(\tau) = \lim_{n, t \rightarrow \infty} (\|\tau\|^2 \ln(\frac{\|\tau\|^2}{t}) + n) \quad (58)$$

где  $\|\tau\|^2$  - норма вектора,  $\tau \in \mathbb{R}^n$

Выпишем еще раз выражения (39) и (37):

$$\begin{cases} q_1 K_f(x_1 - x_1) + q_2 K_f(x_1 - x_2) + \dots + q_k K_f(x_1 - x_k) = y_1 \\ q_1 K_f(x_2 - x_1) + q_2 K_f(x_2 - x_2) + \dots + q_k K_f(x_2 - x_k) = y_2 \\ \vdots \\ q_1 K_f(x_k - x_1) + q_2 K_f(x_k - x_2) + \dots + q_k K_f(x_k - x_k) = y_k \end{cases} \quad (59)$$

$$f^*(x) = q_1 K_f(x - x_1) + q_2 K_f(x - x_2) + \dots + q_k K_f(x - x_k) \quad (60)$$

- Таким образом, в виде (58)-(60) получаем итоговые выражения метода многомерной интерполяции. Метод в пределе функции (58) обеспечивает поиск наиболее вероятной реализации случайной функции, удовлетворяющей положениям (17)-(19) об инвариантности преобразований систем отсчета.



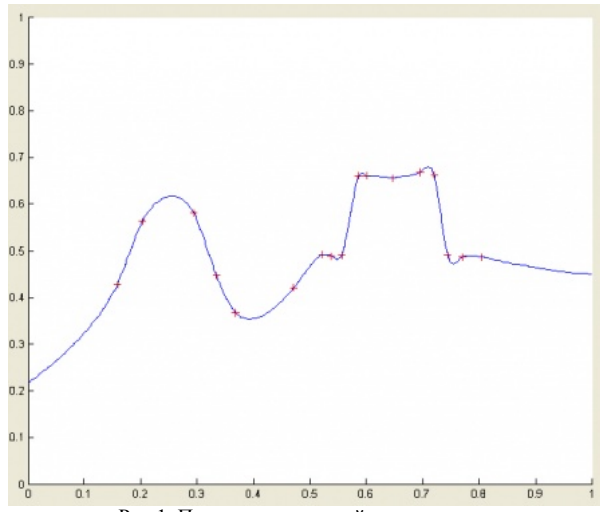


Рис.1. Пример одномерной интерполяции

- При реальных расчетах (как показали вычислительные эксперименты) в качестве  $\tau$  и  $n$  можно брать числа примерно одного порядка и на много порядков больше диапазона изменения  $\tau$ . Ниже еще вернемся к этому вопросу.
- Как видно из рисунка, использование функции (58) позволяет производить качественную интерполяцию нелинейных функций с помощью обычной системы линейных уравнений, что и следует из теории.

#### Дисперсия случайной функции. Множество реализаций, удовлетворяющих узлам интерполяции.

- В рассмотренном выше случае под решением задачи интерполяции (58-60) подразумевался поиск наиболее вероятной реализации случайной функции, удовлетворяющей узлам интерполяции. Однако целиком множество реализаций, удовлетворяющих узлам интерполяции, кроме некоторых частных случаев выражения (3), бесконечно. Было бы полезно иметь характеристики этого множества, кроме того, они необходимы, чтобы перейти от задачи многомерной интерполяции к задаче аппроксимации (как алгоритма полноценного машинного обучения).
- Выразим дисперсию случайной функции (3) для некоторой произвольной точки  $x$ , зная, что дисперсии всех случайных величин из (3) равны единице:

$$D_F(x) = \sum_{i=1}^m \varphi_i^2(x) = K_f(x, x), \quad x \in \mathbb{R}^n \quad (61)$$

- Как видно из (60) дисперсия случайной функции (3) в точке  $x$  равна значению ее корреляционной функции от двух одинаковых значений этой точки. Для случайной функции с автокорреляционной функцией (51), (57) либо (58) дисперсия будет одинаковой во всех точках и равна их значению от нуля. Для (58) значение дисперсии будет равно  $n$ :

$$D_F = K_f(0) = n \quad (62)$$

- Поскольку в (58) значение  $n$  устремлено в бесконечность то в пределе и дисперсия такой случайной функции будет равна бесконечности.

Рассмотрим, чем будет являться значение реализации случайной функции (3) для некоего конкретного значения  $\mathcal{X} \in \mathbb{R}^n$ .

- Обозначим как  $\tilde{L}$  вектор значений  $(\varphi_1(\mathcal{X}), \varphi_2(\mathcal{X}), \dots, \varphi_m(\mathcal{X}))$ .
- Обозначим как  $V$  вектор  $(V_1, V_2, \dots, V_m)$  случайных величин из (3).

Тогда:

$$F(\mathcal{X}) = V\tilde{L} \quad (63)$$

Значение случайной функции (3) или (63) для некоего конкретного значения  $\mathcal{X}$  будет являться случайной величиной распределенной по нормальному закону с дисперсией определяемой формулами (61 - 62).

- Пусть последовательность  $x_1, x_2, \dots, x_k$  ( $x_i \in \mathbb{R}^n$ ) и соответствующие им  $y_1, y_2, \dots, y_k$  ( $y_i \in \mathbb{R}$ ) - представляют собой узлы интерполяции.
- Обозначим векторами  $L_i$  последовательности значений  $(\varphi_1(x_i), \varphi_2(x_i), \dots, \varphi_m(x_i))$ ,  $1 = 1, \dots, k$
- Тогда вектор  $\tilde{L}$  можно будет представить как сумму двух перпендикулярных векторов:

$$\tilde{L} = L_a + L_b \quad (64)$$

где

$$L_a = \sum_{i=1}^k \alpha_i L_i \quad (65)$$

Причем потребуем, чтобы вектор  $L_b$  был перпендикулярен любому из векторов  $L_i$ . Тогда разложение (64) для каждого конкретного  $\mathcal{X}$  будет индивидуальным и единственным. Коэффициенты  $\alpha_i$  в (65) всегда можно подобрать таким образом, чтобы получить разложение (64), поскольку при таком разложении (64-65) количество уравнений равно количеству неизвестных. Вектор  $L_a$ , по сути, является проекцией вектора  $\tilde{L}$  на гиперплоскость, образуемой базисом из векторов  $L_i$ , а вектор  $L_b$  является оставшейся его частью, перпендикулярной данной гиперплоскости.

Тогда (63) можно записать как:

$$F(\mathcal{X}) = V(L_a + L_b) = VL_a + VL_b = F_a(\mathcal{X}) + F_b(\mathcal{X}) \quad (66)$$

Из (66) видно, что при таком представлении случайную величину  $F(\mathcal{X})$  можно считать суммой двух случайных величин  $F_a(\mathcal{X})$  и  $F_b(\mathcal{X})$ , причем они будут независимые. Докажем это.

- Любой набор случайных значений вектора  $V$  также можно разложить на два перпендикулярных вектора:

$$V = V_{a*} + V_{b*} \quad \text{где} \quad (67)$$

$$V_{a*} = \sum_{i=1}^k a_i^* L_i \quad (68)$$

Также потребуем, чтобы вектор  $V_{b*}$  был перпендикулярен любому из векторов  $L_i$ . Тогда получим:

$$F(\mathcal{X}) = V\tilde{L} = (V_{a*} + V_{b*})(L_a + L_b) = V_{a*}L_a + V_{b*}L_b \quad (69)$$

$$F_a(\mathcal{X}) = V_{a*}L_a \quad (70)$$

$$F_b(\mathcal{X}) = V_{b*}L_b \quad (71)$$

Часть слагаемых из (69) сократятся, в силу того, что скалярное произведение перпендикулярных векторов равно нулю.

- Коэффициенты  $a_i^*$  из (68) являются коэффициентами  $q_i$  из (10) или (15).
- Вектора  $V_{a*}$  и  $V_{b*}$  являются перпендикулярными и образуют в сумме вектор  $V$ , имеющий многомерный нормальный закон распределения. Следовательно, в силу свойств нормального закона распределения мы можем рассматривать (66) как разложение исходной случайной функции на сумму  $F_a(\mathcal{X})$  и  $F_b(\mathcal{X})$  двух независимых случайных функций.

Рассмотрим  $F_a(\mathcal{X})$ :

$$F_a(\mathcal{X}) = V_{a*}L_a = (a_1^*L_1 + a_2^*L_2 + \dots + a_k^*L_k)\tilde{L} \quad (72)$$

Что при известных значениях в узлах интерполяции автоматически превращается в наиболее вероятную реализацию, вычисляемую через формулы (15) и (16).

- Таким образом, случайную функцию (3) при известных координатах узлов интерполяции можно представить как сумму двух независимых случайных функций.  $F_a(\mathcal{X})$  - случайную функцию, реализациями которых являются функции, вычисляемые через (15) и (16) как наиболее вероятные реализации исходной случайной функции (3) и  $F_b(\mathcal{X})$  - случайную функцию, представляющую собой разность ("ошибку") между найденной наиболее вероятной реализацией и настоящей реализацией случайной функции о которой нам не известно.
- Если рассматривать  $F_b(\mathcal{X})$  как случайную величину в конкретной точке  $\mathcal{X}$  то она будет распределена по нормальному закону с математическим ожиданием равным нулю.
- Следует отметить тот факт, что поскольку разложение (64)-(65) не зависит от конкретных значений в узлах интерполяции то и функция  $F_b(\mathcal{X})$ , а соответственно и вероятность ошибки интерполяции при любом конкретном  $\mathcal{X}$  не будет зависеть от значений в узлах интерполяции, а только от расположения координат самих узлов.

Предположим, что известны значения искомой функции в узлах интерполяции. Тогда выражение (66) обратится в следующий вид (введем новое обозначение):

$$\tilde{F}(\mathcal{X}) = f^*(\mathcal{X}) + F_b(\mathcal{X}) \quad (73)$$

где  $f^*(\mathcal{X})$  - наиболее вероятная реализация, вычисленная в (15)-(16)

- Выражение (73) можно понимать как случайную функцию выражающую целиком множество реализаций удовлетворяющих узлам интерполяции (с соответствующим распределением вероятностей между ними).

Дисперсию (61) случайной функции используя (69) можно записать:

$$D_F(\mathcal{X}) = K_f(\mathcal{X}, \mathcal{X}) = \|L_a\|^2 + \|L_b\|^2 = D_a + D_b \quad (74)$$

Дисперсия случайной функции (73) будет равна:

$$D_{\tilde{F}}(\mathcal{X}) = 0 + \|L_b\|^2 = K_f(\mathcal{X}, \mathcal{X}) - D_a \quad (75)$$

Рассмотрим  $K_f(\mathcal{X}, x_j)$ , где  $x_j$  - координата одного из узлов интерполяции.

$$K_f(\mathcal{X}, x_j) = \tilde{L}L_j = (L_a + L_b)L_j = L_aL_j = \sum_{i=1}^k a_i L_i L_j = \sum_{i=1}^k a_i K_f(x_i, x_j) \quad (76)$$

Получаем систему уравнений:

$$\begin{cases} a_1 K_f(x_1, x_1) + a_2 K_f(x_1, x_2) + \dots + a_k K_f(x_1, x_k) = K_f(\mathcal{X}, x_1) \\ a_1 K_f(x_2, x_1) + a_2 K_f(x_2, x_2) + \dots + a_k K_f(x_2, x_k) = K_f(\mathcal{X}, x_2) \\ \vdots \\ a_1 K_f(x_k, x_1) + a_2 K_f(x_k, x_2) + \dots + a_k K_f(x_k, x_k) = K_f(\mathcal{X}, x_k) \end{cases} \quad (77)$$

Выразим дисперсию  $D_a$ :

$$D_a = L_a L_a = \sum_{i=1}^k a_i L_i L_a = a_1 K_f(\mathcal{X}, x_1) + a_2 K_f(\mathcal{X}, x_2) + \dots + a_k K_f(\mathcal{X}, x_k) \quad (78)$$

Повторим еще раз функцию (58):

$$K_f(\tau) = \lim_{n, t \rightarrow \infty} (\|\tau\|^2 \ln(\frac{\|\tau\|^2}{t}) + n)$$

где  $\|\tau\|^2$  - норма вектора,  $\tau \in R^n$

Если использовать полученную автокорреляционную функцию (58) то получим систему уравнений:

$$\begin{cases} a_1 K_f(x_1 - x_1) + a_2 K_f(x_1 - x_2) + \dots + a_k K_f(x_1 - x_k) = K_f(\mathcal{X} - x_1) \\ a_1 K_f(x_2 - x_1) + a_2 K_f(x_2 - x_2) + \dots + a_k K_f(x_2 - x_k) = K_f(\mathcal{X} - x_2) \\ \vdots \\ a_1 K_f(x_k - x_1) + a_2 K_f(x_k - x_2) + \dots + a_k K_f(x_k - x_k) = K_f(\mathcal{X} - x_k) \end{cases} \quad (79)$$

Как видно из (79) получилась система уравнений, левая часть которой совпадает с (59), но в правой части вместо значений в узлах интерполяции находятся значения автокорреляционной функции. Выражение дисперсии  $D_a$ :

$$D_a = a_1 K_f(\mathcal{X} - x_1) + a_2 K_f(\mathcal{X} - x_2) + \dots + a_k K_f(\mathcal{X} - x_k) \quad (80)$$

Тогда дисперсия множества реализаций, удовлетворяющих узлам:

$$D_{\mathcal{F}}(\mathcal{X}) = n - D_a \quad (81)$$

- Если рассматривать множество реализаций, удовлетворяющих узлам интерполяции с соответствующим им распределением как новую случайную функцию, то среднее квадратическое отклонение такой функции будет:

$$\delta_{\mathcal{F}}(\mathcal{X}) = \sqrt{n - D_a} \quad (82)$$

- Как уже отмечалось ранее, умножение корреляционной или автокорреляционной функции в уравнениях (15) - (16) или (58) - (60) на любой конечный ненулевой коэффициент и использование такой новой функции не изменит найденную в этих уравнениях наиболее вероятную реализацию.
- Однако при работе с дисперсией значения автокорреляционной функции в нуле будет напрямую влиять на ее значение. Кроме того, как будет видно ниже, это влияет на результаты при переходе к задаче многомерной аппроксимации.
- Поскольку умножение автокорреляционной функции на конечный ненулевой коэффициент не влияет на результаты интерполяции, можно выполнить "калибровку" дисперсии автокорреляционной функции в соответствии с априорными предпочтениями (которые могут следовать из условий решаемой задачи).
- Осуществить калибровку можно различными способами. Ниже предложен один из вариантов.
- Калибровку можно осуществить оценкой на основе априорных предпочтений и заданием коэффициента перед автокорреляционной функцией.

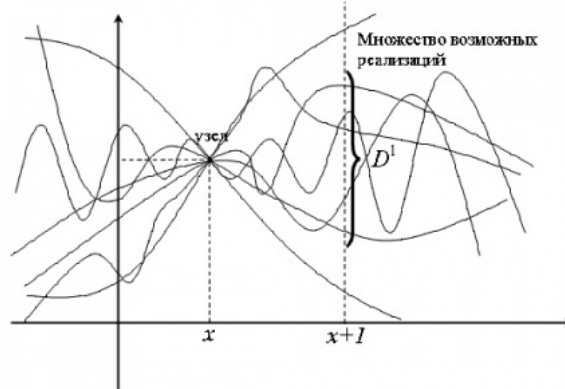


Рис.2. Априорная оценка дисперсии на единичном расстоянии от узла интерполяции.

- Предположим известно значение реализации случайной функции в одном узле интерполяции (Рис.2.). Если существует возможность каким-либо образом оценить или сделать предположение о величине дисперсии разброса реализаций на единичном расстоянии от него, тогда можно выполнить калибровку.
- Допустим, что мы хотим, чтобы используемая нами априорная случайная функция обладала такими свойствами, чтобы на единичном расстоянии от известного узла дисперсия разброса реализаций, удовлетворяющих узлу, равнялась некоторому желаемому значению  $D^1$ .
- Предположим, что используется функция (58) с заданными значениями  $\mathcal{t}$  и  $n$ , используемыми в реальных вычислениях. Введем калибровочный коэффициент:

$$K_f(\tau) = C_K (\|\tau\|^2 \ln(\frac{\|\tau\|^2}{\mathcal{t}}) + n) \quad (83)$$

где  $C_K$  - калибровочный коэффициент

Вычислим значение калибровочного коэффициента при заданных  $\mathcal{t}$  и  $n$  такое, чтобы дисперсия на единичном расстоянии от узла равнялась желаемому значению  $D^1$ . Из (79) получим:

$$a_1 K_f(0) = K_f(1) \quad (84)$$

Подставим (83) в (84):

$$a_1 C_K n = C_K (\ln(\frac{1}{\mathcal{t}}) + n) \quad (85)$$

Выразим  $a_1$ :

$$a_1 = \frac{\ln(\frac{1}{\mathcal{t}}) + n}{n} \quad (86)$$

Тогда:

$$D_a = C_K \frac{(\ln(\frac{1}{\mathcal{t}}) + n)^2}{n} \quad (87)$$

Выразим желаемую дисперсию на единичном расстоянии:

$$D^1 = K_f(0) - D_a \quad (88)$$

Выполнив преобразования, получим:

$$D^1 = C_K \frac{2n - \ln(t)}{n} \ln(t) \quad (89)$$

Откуда можно вычислить коэффициент калибровки:

$$C_K = \frac{D^1 n}{(2n - \ln(t)) \ln(t)} \quad (90)$$

Если вместо единичного расстояния для калибровки удобнее взять другое расстояние равное  $d$  и оценить для него дисперсию  $D^d$ , то по ним также можно выполнить калибровку. Если выполнить преобразования аналогичные (84)-(90) получится выражение для калибровочного коэффициента:

$$C_K = \frac{D^d n}{(2n - d^2 \ln(\frac{1}{d^2})) d^2 \ln(\frac{1}{d^2})} \quad (90.a)$$

Пример:

Рассмотрим случай одномерной интерполяции. Пусть область интерполяции находится в диапазоне от 0 до 10. Будем использовать следующие значения для функции (83):  $t = 10^6$ ,  $n = 10^6$ ,  $D^1 = 1$ .

Вычислим калибровочный коэффициент:

$$C_K = 0.036191456826998$$

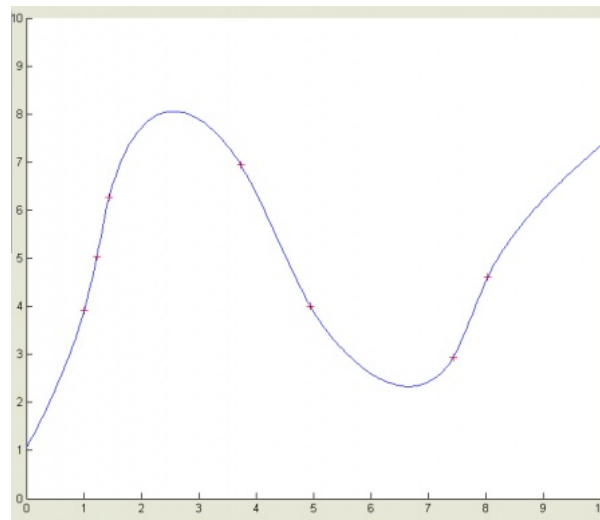


Рис.3. Пример одномерной интерполяции.

- На Рис.3. представлен пример одномерной интерполяции, полученный с использованием функции (83) и уравнений (59) – (60). Однако, как уже было сказано выше, калибровочный коэффициент  $C_K$  не влияет на результат интерполяции (на вычисление наиболее вероятной реализации).
- Найдем среднеквадратическое отклонение случайной функции, представляющей собой подмножество реализаций априорной случайной функции удовлетворяющих узлам интерполяции (выражения (79) – (82)). Результат представлен на Рис.4.

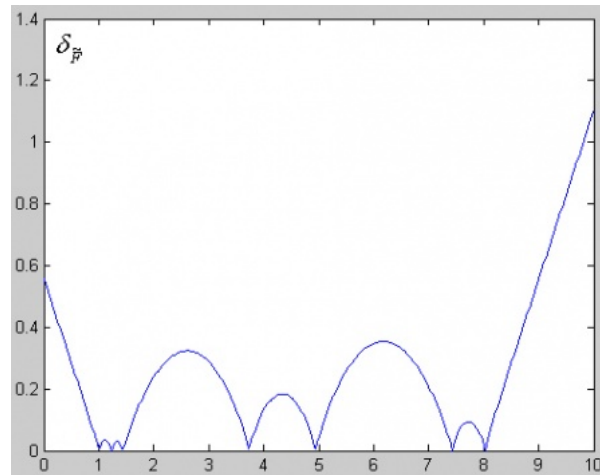


Рис.4. Среднеквадратическое отклонение множества реализаций, удовлетворяющих узлам интерполяции.

- Случайная функция (73) для произвольного  $\mathcal{F}$  представляет собой случайную величину, имеющую нормальное распределение. При известном среднеквадратическом отклонении мы можем отобразить полученное множество реализаций графически (Рис.5).

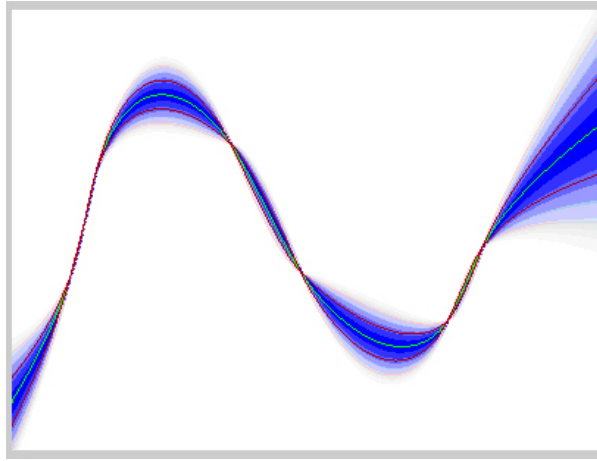


Рис.5. Графическое отображение множества реализаций, удовлетворяющих узлам интерполяции. Интенсивностью синего цвета символизируется вероятность прохождения реализации через данные точки.

- Как видно из Рис.4. среднее квадратическое отклонение в узлах интерполяции равно нулю, что соответствует ожиданиям, поскольку значения искомого функции в этих узлах известны и соответственно никакой неопределенности быть не может.
- Выше был представлен пример одномерной интерполяции и вычисления для данного частного случая среднее квадратическое отклонение по оставшемуся множеству реализаций, удовлетворяющих узлам. Был рассмотрен одномерный случай в силу наглядности его графического отображения. Для представленных преобразований нет ограничения по размерности пространства интерполяции.
- Если рассматривать подобную интерполяцию как метод машинного обучения, то вычисление среднее квадратического отклонения  $\delta_{\mathcal{F}}(\mathcal{X})$  может позволить не только вычислять значения выходных переменных но и оценить их надежность, возможный разброс их значений.
- Несмотря на то, что вычислительные эксперименты убедительно подтверждают работоспособность метода при использовании  $t = 10^6$ ,  $n = 10^6$  (при области интерполяции измеряемой единицами), остается открытым вопрос о влиянии их значений на погрешность интерполяции по сравнению с рассчитанной “идеальной функцией” (57) – (58).
- Очевидно, что такая погрешность связана с отклонением спектра используемой автокорреляционной функции от эталонного спектра (49) – (50), что также может привести к искажению полученного интерполанта в этой же частотной области. Анализируя выражение (52) – поскольку при конечных коэффициентах в (58),  $\omega_0$  в (52) не достигает нуля, можно сделать вывод, что главные искажения сведутся в низкочастотную область в пределах  $\omega \approx \sqrt{n}$  или  $\omega \approx \sqrt{t}$ , что при  $t = 10^6$ ,  $n = 10^6$  и области интерполяции  $[0,10]$  не должно значительно сказываться на результатах.
- Однако вопрос о погрешности остается открытым, чему планируется в дальнейшем посвятить отдельную исследовательскую работу.

## Многомерная аппроксимация.

- При решении задачи многомерной интерполяции предполагалось, что искомая реализация должна точно соответствовать узлам интерполяции. Если переводить в термины машинного обучения, был рассмотрен вариант такого обучения, когда ошибка обучения должна быть равна нулю. Во многих случаях точно воспроизводить обучающую выборку не требуется или это может вредно сказаться на качестве обобщающих способностей. Данные могут быть не точными, содержать погрешности или даже быть противоречивыми. Т.е. требуется решение задачи не интерполяции, а аппроксимации многомерных данных.
- Перейти к задаче аппроксимации можно предположив, что значения узлов интерполяции могут не соответствовать породившей их реализации и могут отличаться от ее значений на некоторую случайную величину.
- Пусть последовательность  $x_1, x_2, \dots, x_k$  и соответствующие им  $y_1, y_2, \dots, y_k$  – представляют собой узлы интерполяции. Предположим, что координаты узлов  $x_1, x_2, \dots, x_k$  представляют собой комплексные числа, к которым были прибавлены бесконечно малые мнимые части, такие, что теперь координаты узлов не могут быть точно равны ни между собой, ни какому либо другому вещественному числу.
- В этом случае выражение (3) мы можем модифицировать следующим образом:

$$F(x) = \sum_{i=1}^m V_i \varphi_i(x) + \sum_{j=1}^k V_{(m+j)} E_j(x) \quad (91)$$

$$E_j(x) = \begin{cases} s, & x = x_j \\ 0, & x \neq x_j \end{cases} \quad \text{где} \quad (92)$$

$x_j \in \mathbb{Z}^n$ ,  $j = 1, \dots, k$  – обновленные координаты узлов интерполяции

- В выражении (91) были добавлены дополнительные случайные величины, а также набор функций  $E_j(x)$ , которые принимают значение равное  $s$  в соответствующих им узлах интерполяции и имеющие нулевое значение во всех остальных точках.
- Поскольку значения узлов интерполяции не могут точно равняться друг другу, ни какому либо другому вещественному значению, функции  $E_j(x)$  принимают ненулевое значение исключительно для соответствующих узлов и, выражения (91) и (92) полностью имитируют наличие погрешностей в значениях узлов интерполяции. Использование такой случайной функции (91) будет эквивалентно наличию в значениях узлов интерполяции шума, с нормальным распределением и среднее квадратическим отклонением равным  $s$ . В то же время мы можем использовать функции  $E_j(x)$  как полноценные дополнительные координатные функции в выражении (3).
- Приняв  $E_j(x)$  как дополнительные координатные функции в (3) можно выполнить без изменений все преобразования (3) – (16). Изменения коснутся лишь корреляционной функции, обозначим ее как  $\widehat{K}_f(x_1, x_2)$ :

$$\widehat{K}_f(x_1, x_2) = \sum_{i=1}^m \varphi_i(x_1) \varphi_i(x_2) + \sum_{j=1}^k E_j(x_1) E_j(x_2) \quad (93)$$

$x_1, x_2 \in \mathbb{Z}^n$  – некоторые произвольные значения

- Тогда корреляционную функцию в  $x_i, x_j$  - узлах интерполяции можно выразить как:

$$\widehat{K_f}(x_1, x_2) = \begin{cases} K_f(x_i, x_j) + s^2, & x_i = x_j \\ K_f(x_i, x_j), & x_i \neq x_j \end{cases} \quad (94)$$

- Для произвольных вещественных значений  $x_1, x_2 \in R^n$  корреляционная функция останется неизменной (14).

$$\widehat{K_f}(x_1, x_2) = K_f(x_1, x_2) \quad (95)$$

- Используя (94) получим, что учет существования шума в значениях узлов интерполяции с нормальным законом распределения и среднеквадратическим отклонением равным  $s$  эквивалентен добавлению в уравнение (15) к главной диагонали значений  $s^2$  (дисперсию шума).
- Для автокорреляционной функции (58) или (83) получаем систему уравнений:

$$\begin{cases} q_1(K_f(x_1-x_1)+s^2)+q_2K_f(x_1-x_2)+\dots+q_kK_f(x_1-x_k)=y_1 \\ q_1K_f(x_2-x_1)+q_2(K_f(x_2-x_2)+s^2)+\dots+q_kK_f(x_2-x_k)=y_2 \\ \vdots \\ q_1K_f(x_k-x_1)+q_2K_f(x_k-x_2)+\dots+q_k(K_f(x_k-x_k)+s^2)=y_k \end{cases} \quad (96)$$

- Уравнение для наиболее вероятной реализации остается неизменным:

$$f^*(x) = q_1K_f(x-x_1)+q_2K_f(x-x_2)+\dots+q_kK_f(x-x_k), \quad x \in R^n \quad (97)$$

Пример аппроксимации:

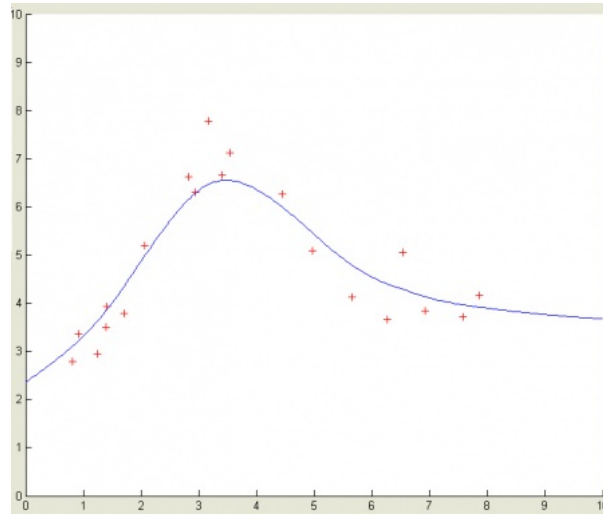


Рис.6. Пример аппроксимации.

Графическое отображение множества реализаций, соответствующее аппроксимации представлено на Рис.7.

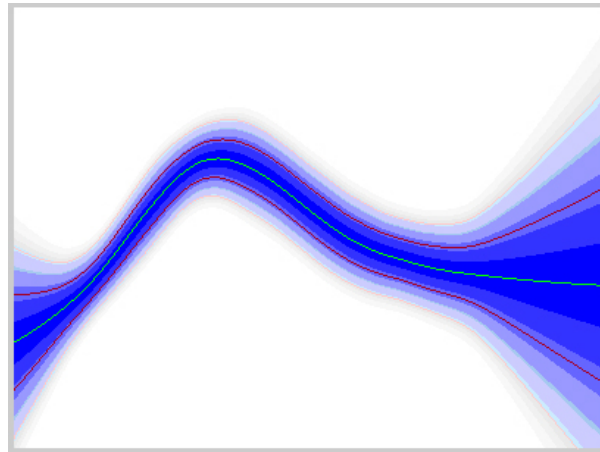


Рис.7. Пример отображения множества реализаций при аппроксимации.

Поскольку мы ввели учет шума с помощью дополнительных координатных функций то все рассуждения, касающиеся дисперсии случайной функции (61) – (62) остаются в силе. Изменения коснутся только системы уравнений (79), которую теперь можно представить как (98):

$$\begin{cases} a_1(K_f(x_1-x_1)+s^2)+a_2K_f(x_1-x_2)+\dots+a_kK_f(x_1-x_k)=K_f(x-x_1) \\ a_1K_f(x_2-x_1)+a_2(K_f(x_2-x_2)+s^2)+\dots+a_kK_f(x_2-x_k)=K_f(x-x_2) \\ \vdots \\ a_1K_f(x_k-x_1)+a_2K_f(x_k-x_2)+\dots+a_k(K_f(x_k-x_k)+s^2)=K_f(x-x_k) \end{cases} \quad (98)$$

- Следует отметить, что в результате вычисления среднеквадратического отклонения реализаций, удовлетворяющих узлам (отображенное на рис.7) –

его значение отражает лишь распределение порождающих реализаций случайной функции без учета возможных погрешностей в их значениях.

- В выражении (92) функции  $E_j(x)$  принимают значение равное 5 либо нулевое значение. Однако если известно распределение шума в области аппроксимации, это распределение можно непосредственно учесть в выражении (92), поскольку для каждого узла может быть задано индивидуальное значение среднеквадратического отклонения шума:

$$E_j(x) = \begin{cases} S_m(x), & x = x_j \\ 0, & x \neq x_j \end{cases} \quad (99)$$

где  $S_m(x)$  - значение среднеквадратического отклонения шума в  $x$

Тогда выражение (96) примет вид:

$$\begin{cases} q_1(K_f(x_1-x_1)+S_m^2(x_1))+q_2K_f(x_1-x_2)+...+q_kK_f(x_1-x_k)=y_1 \\ q_1K_f(x_2-x_1)+q_2(K_f(x_2-x_2)+S_m^2(x_2))+...+q_kK_f(x_2-x_k)=y_2 \\ \vdots \\ q_1K_f(x_k-x_1)+q_2K_f(x_k-x_2)+...+q_k(K_f(x_k-x_k)+S_m^2(x_k))=y_k \end{cases} \quad (100)$$

где  $S_m(x)$  - среднеквадратическое отклонение для шума в точке  $x$

А выражение (98) будет:

$$\begin{cases} a_1(K_f(x_1-x_1)+S_m^2(x_1))+a_2K_f(x_1-x_2)+...+a_kK_f(x_1-x_k)=K_f(x-x_1) \\ a_1K_f(x_2-x_1)+a_2(K_f(x_2-x_2)+S_m^2(x_2))+...+a_kK_f(x_2-x_k)=K_f(x-x_2) \\ \vdots \\ a_1K_f(x_k-x_1)+a_2K_f(x_k-x_2)+...+a_k(K_f(x_k-x_k)+S_m^2(x_k))=K_f(x-x_k) \end{cases} \quad (101)$$

Пример. Рассмотрим одномерный случай.

Предположим, распределение шума (его среднеквадратического отклонения) в области аппроксимации определяется выражением:

$$S_m(x) = \frac{x^2}{40}, \quad x \in R \quad (102)$$

Т.е. в области около нуля шум почти отсутствует и возрастает по мере удаления.

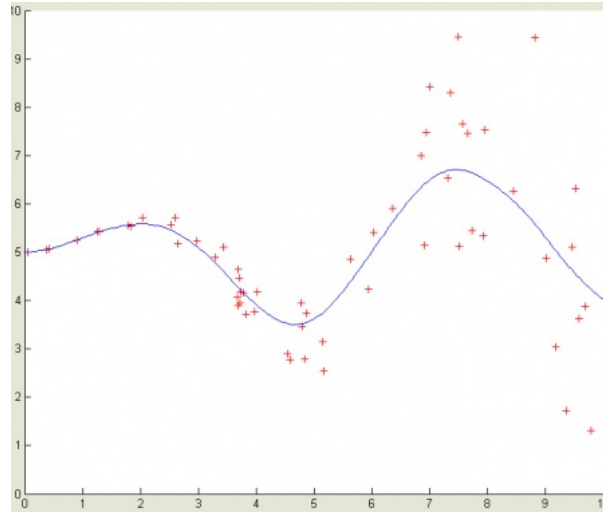


Рис.8. Аппроксимация с учетом распределения шума.

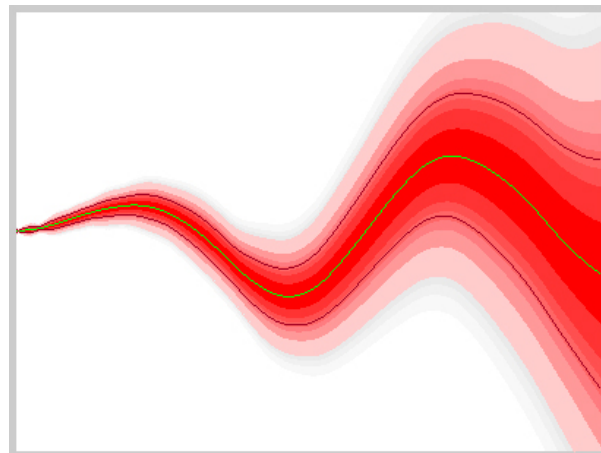


Рис.9. Графическое отображение множества реализаций.

- На рис.8. представлена аппроксимация с использованием системы уравнений (100) и выражением для среднеквадратического значения в распределении шума (102).

- На рис.9. для более наглядного сравнения с рис.8. множество реализаций отображено с учетом шума (102).

## Выводы.

- Проведенные рассуждения и преобразования показывают, что интерполяция и аппроксимация могут быть обобщены на основе теории случайных функций как поиск наиболее вероятных реализаций, удовлетворяющих информации об узлах. В результате был получен метод, которые можно рассматривать как метод “машинного обучения”, имеющий точное решение, обеспечивающий получение оптимального результата по критерию вероятности, при соблюдении введенных условий по преобразованию систем отсчета.

## Метод машинного обучения.

Выпишем отдельно ключевые выражения, полученные в теоретической части в виде законченного метода.

- Пусть последовательность  $x_1, x_2, \dots, x_k$  ( $x_i \in R^n$ ) представляет собой набор входных векторов для обучения. Пусть соответствующие им  $y_1, y_2, \dots, y_k$  ( $y_i \in R$ ) представляют собой набор выходных значений. Если значения на выходе представляют собой не один, а набор значений (вектора), то представленные преобразования можно рассмотреть последовательно для каждого из выходных параметров.
- Метод позволяет определить функцию, связывающую значения на входе и на выходе (которая будет являться наиболее вероятной реализацией случайной функции, о чем шла речь в теоретической части). Метод также позволяет для произвольного входного значения определить среднеквадратическое отклонение ошибки, которую может дать полученная функция.

Функция, связывающая значения обучающей выборки на входе и на выходе будет определяться выражением (103):

$$f^*(x) = q_1 K_f(x-x_1) + q_2 K_f(x-x_2) + \dots + q_k K_f(x-x_k), \quad x \in R^n \quad (103)$$

Функцию  $K_f(\tau)$  из (103) определим как (104):

$$K_f(\tau) = C_K (\|\tau\|^2 \ln(\frac{\|\tau\|^2}{t}) + n) \quad (104)$$

где  $C_K, t, n$  - коэффициенты  
 $\|\tau\|$  - норма (длина) вектора  $\tau$

Коэффициенты  $q_1, q_2, \dots, q_k$  для (103) вычисляются из системы линейных уравнений (105):

$$\begin{cases} q_1(K_f(x_1-x_1) + S_m^2(x_1)) + q_2 K_f(x_1-x_2) + \dots + q_k K_f(x_1-x_k) = y_1 \\ q_1 K_f(x_2-x_1) + q_2(K_f(x_2-x_2) + S_m^2(x_2)) + \dots + q_k K_f(x_2-x_k) = y_2 \\ \vdots \\ q_1 K_f(x_k-x_1) + q_2 K_f(x_k-x_2) + \dots + q_k(K_f(x_k-x_k) + S_m^2(x_k)) = y_k \end{cases} \quad (105)$$

где  $S_m(x)$  - среднеквадратическое отклонение для шума в точке  $x$

- Значения  $S_m(x)$  определяют априорно предполагаемый уровень шума (погрешности) в данных обучения и соответственно степень приближения, с которой функция (103) воспроизведет данные обучения.
- Если  $S_m(x) = 0$ , то это будет соответствовать частному случаю, когда погрешность или любая противоречивость в данных отсутствует и, искомая функция должна точно воспроизвести обучающую выборку. Т.е. задача обучения превращается в задачу многомерной интерполяции.
- Если  $S_m(x) = const > 0$ , то это будет соответствовать случаю, когда предполагается что в данных обучения одинаково на всей выборке содержится шум, имеющий нормальное распределение со среднеквадратическим отклонением равным  $S_m(x) = const$ . Задачу обучения можно рассматривать как многомерную аппроксимацию.
- Если же уровень шума и его распределение в пространстве обучения неравномерно но известно, его наличие может быть задано в виде функции  $S_m(x)$  (или достаточно только ее значений в узлах интерполяции).

Рассмотрим теперь коэффициенты в выражении (104).

- Коэффициенты  $t$  и  $n$  в “идеальном случае” должны быть приблизительно равны и устремлены в бесконечность. В реальных вычислениях в качестве этих коэффициентов можно использовать  $t \approx n \approx 10^5 \div 10^6$  при условии нормирования входных значений в диапазоне [-1,1] (они должны быть на несколько порядков больше диапазона изменения входных значений). С увеличением  $t$  и  $n$  разница от использования функции (104) вместо “идеальной функции” устремляется в область все более низких частот (при разложении искомой функции в спектр) измеряемыми  $\omega \approx \sqrt{t} \approx \sqrt{n}$  и все меньше влияет на результаты обучения в области параметрического пространства, где находится обучающая выборка.
- Коэффициент  $C_K$ , назовем его “калибровочным”, связан со свойствами априорной случайной функции. Хотя непосредственно случайные функции в представленных выражениях (103) - (105) не используются, эти выражения выведены на их основе. (Если  $S_m(x) = 0$  и обучение сводится к многомерной интерполяции, значение  $C_K$  можно взять произвольным, поскольку при решении системы (105) и вычислении функции (103) он сократится.)

Процедура вычисления  $C_K$  (предлагаемый вариант):



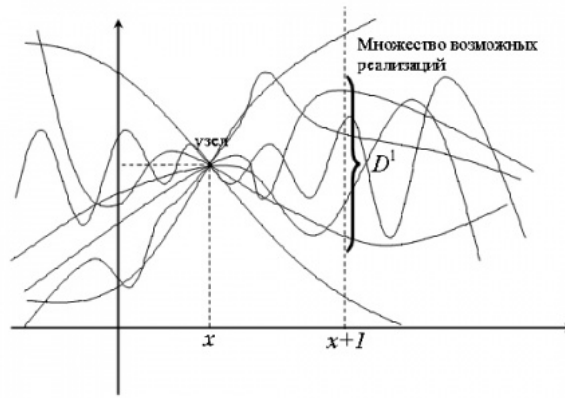


Рис.10. Оценка желаемой дисперсии на единичном расстоянии.

- Калибровка требуется, чтобы найти баланс между возможными погрешностями, неточностями и противоречиями в данных обучения и возможной нелинейностью самой функции, связывающей вход и выход.
- Предположим известно, что искомая функция (103) проходит через некий узел. Для калибровки можно априорно задать желаемый уровень дисперсии множества возможных реализаций  $D^1$  на единичном расстоянии от узла. С увеличением  $D^1$  при вычислениях (103) – (105) все “большая роль” будет отводиться возможной нелинейности самой функции, с уменьшением – все больше “списываться” на наличие неточностей в данных обучения. При известном  $D^1$  можно вычислить требуемый коэффициент  $C_K$ :

$$C_K = \frac{D^1 n}{(2n - \ln(t)) \ln(t)} \quad (106)$$

- Если вместо единичного расстояния для калибровки удобнее взять другое расстояние равное  $d$  и оценить для него дисперсию  $D^d$ , то по ним также можно выполнить калибровку. Выражение для калибровочного коэффициента будет:

$$C_K = \frac{D^d n}{(2n - d^2 \ln(\frac{t}{d^2})) d^2 \ln(\frac{t}{d^2})} \quad (106.a)$$

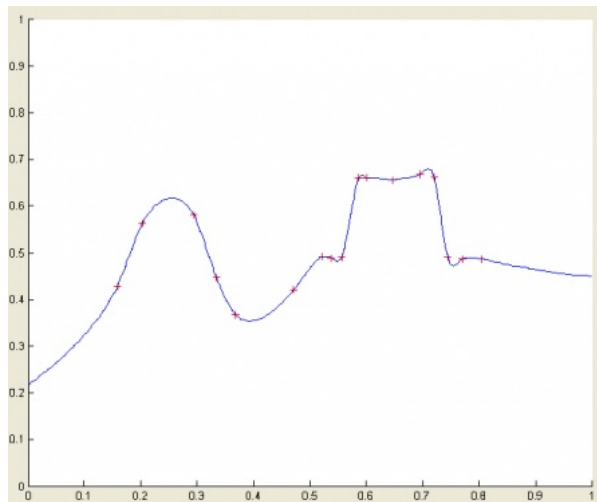


Рис.11. Пример одномерной интерполяции.

- На рис.11. представлен пример одномерной интерполяции (при  $S_m(x) = 0$ ). Как видно из рисунка, интерполяция выполнена качественно (отсутствуют осцилляции, которые, например, часто имеют место при использовании интерполяционного многочлена Лагранжа).
- С вводом  $S_m(x) = \text{const} > 0$ , интерполяция легко превращается в аппроксимацию - Рис.12.

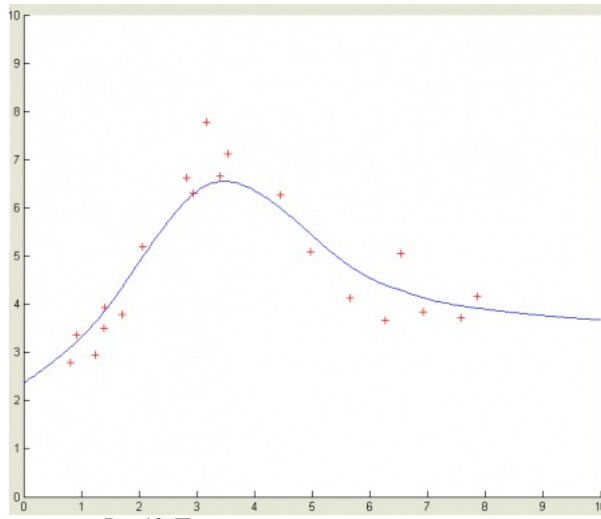


Рис.12. Пример одномерной аппроксимации.

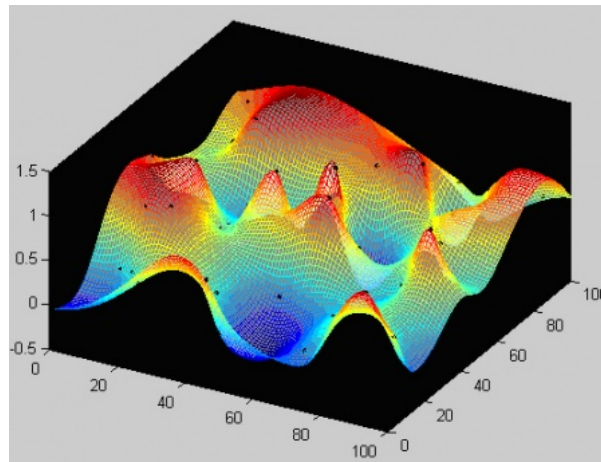


Рис.13. Пример двумерной интерполяции.

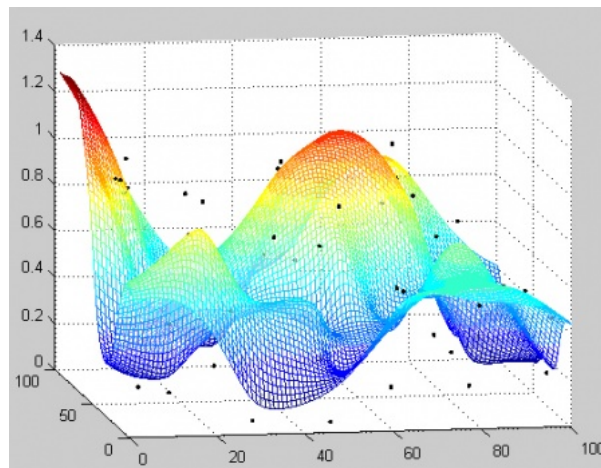


Рис.14. Пример двумерной аппроксимации.

- На Рис. 11 – 14 представлены примеры интерполяции и аппроксимации в одномерном и двумерном случае в целях наглядности результатов. Выражения (103) – (105) могут быть использованы без ограничений для пространства любой размерности.
- Для определения среднеквадратического отклонения возможной ошибки были получены выражения:

$$\delta_{\mathcal{F}}(x) = \sqrt{(n - D_a)} \quad (107)$$

- Значение  $D_a$  определяется следующим выражением:

$$D_a = a_1 K_f(x - x_1) + a_2 K_f(x - x_2) + \dots + a_k K_f(x - x_k) \quad (108)$$

- А коэффициенты  $a_1, a_2, \dots, a_k$  из (108) находятся решением системы уравнений (109):

$$\begin{cases} a_1(K_f(x_1 - x_1) + S_m^2(x_1)) + a_2 K_f(x_1 - x_2) + \dots + a_k K_f(x_1 - x_k) = K_f(x - x_1) \\ a_1 K_f(x_2 - x_1) + a_2(K_f(x_2 - x_2) + S_m^2(x_2)) + \dots + a_k K_f(x_2 - x_k) = K_f(x - x_2) \\ \vdots \\ a_1 K_f(x_k - x_1) + a_2 K_f(x_k - x_2) + \dots + a_k(K_f(x_k - x_k) + S_m^2(x_k)) = K_f(x - x_k) \end{cases} \quad (109)$$

- Необходимо отметить, что при расчете среднеквадратического отклонения для каждого значения  $x$ , необходимо рассчитать индивидуальное значение  $D_{\alpha}$  и заново решать систему уравнений (109) (хотя вычисления можно упростить, единожды вычислив обратную матрицу).

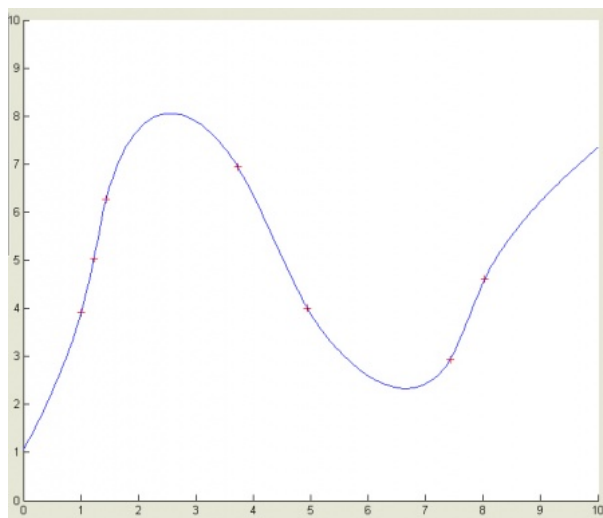


Рис.15. Пример одномерной интерполяции.

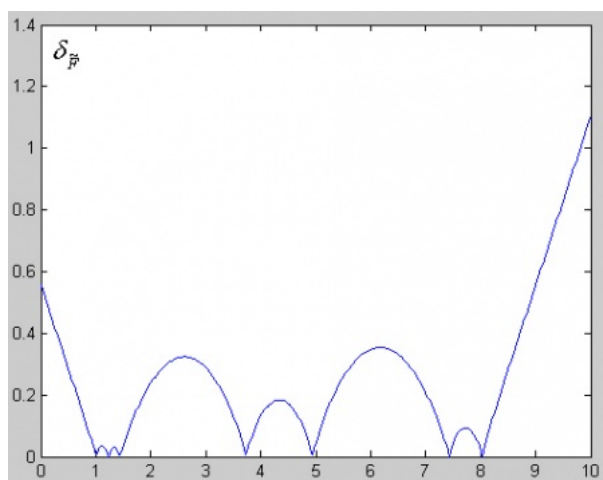


Рис.16. Среднеквадратическое отклонение распределения возможной ошибки как случайной величины для найденной функции на Рис.15

- Ожидаемая ошибка функции (103) как случайная величина (исходя из теоретических положений в первой части) будет иметь нормальное распределение с рассчитанным по (107) среднеквадратическим отклонением и математическим ожиданием равным нулю. Зная это, мы можем ее отобразить графически, как показано на Рис.17.

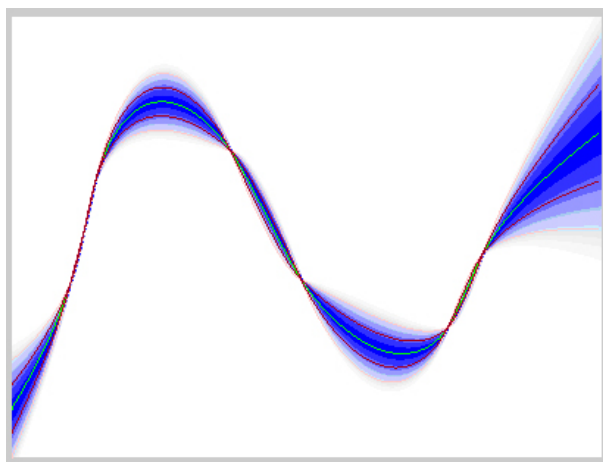


Рис.17. Возможная ожидаемая ошибка, или распределение реализаций случайной функции, удовлетворяющих данным обучения

- С другой стороны, то, что мы можем понимать под возможной ошибкой, представляет собой распределение значений возможных реализаций априорной случайной функции (рассмотренной в теоретической части) которые удовлетворяют данным обучения.

Ниже представлены рисунки, демонстрирующие аппроксимацию.

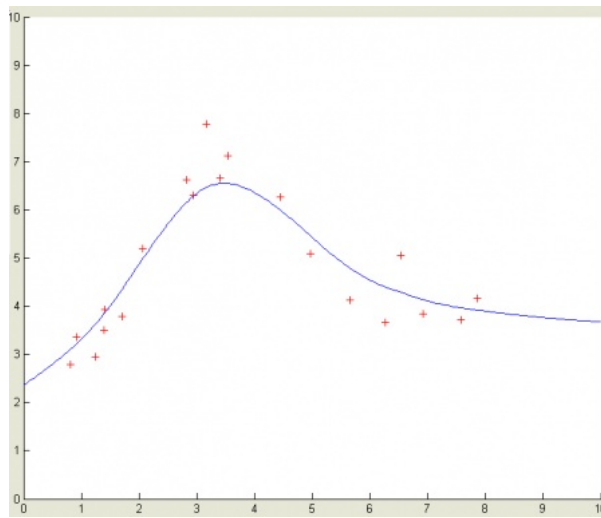


Рис.18. Пример одномерной аппроксимации.

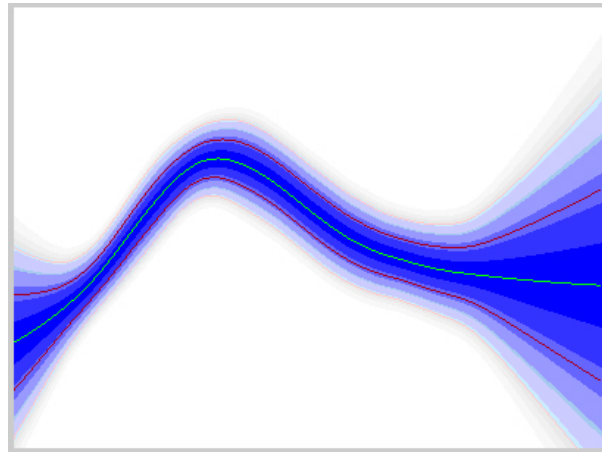


Рис.19. Распределение реализаций или возможная ожидаемая ошибка (для функции на Рис.18).

- Среднеквадратическая ошибка в (107) отражает разброс значений связанный с возможным распределением реализаций. Если требуется учесть возможный шум в самих данных, то для вычисления среднеквадратического отклонения можно воспользоваться выражением:

$$\delta_{\mathcal{F}}(x) = \sqrt{(n - D_a + S_m^2(x))} \quad (110)$$

Пример. Рассмотрим одномерный случай.

- Предположим, распределение шума (его среднеквадратического отклонения) в области аппроксимации определяется выражением:

$$S_m(x) = \frac{x^2}{40}, \quad x \in \mathbb{R}$$

- Т.е. в области около нуля шум почти отсутствует и возрастает по мере удаления.

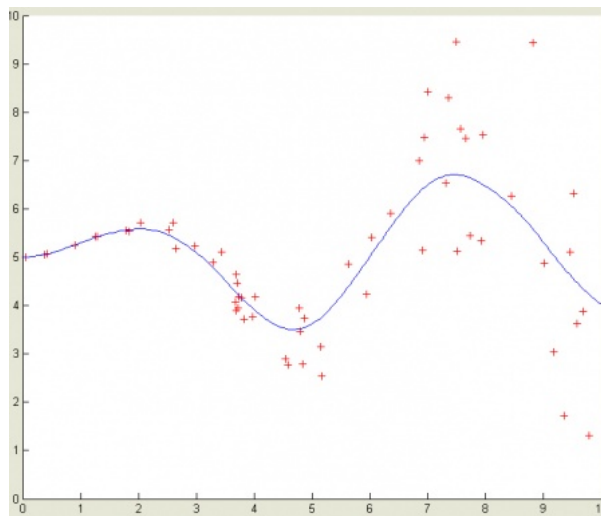


Рис.20. Аппроксимация с учетом распределения шума.

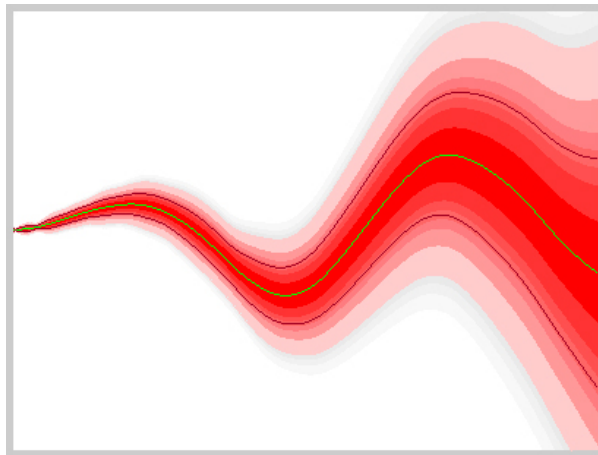


Рис.21. Графическое отображение множества реализаций (с учетом шума).

## Демонстрация в среде Matlab.

Предложенный подход к машинному обучению был реализован в виде набора функций в среде Matlab.

Ссылка [52Kб] на файл-архив с функциями. (Также есть программа на языке Python 3 с демонстрацией одномерной интерполяции/аппроксимации.)

Кроме самих функций в архиве содержатся программы-демонстрации в виде графического интерфейса пользователя в среде Matlab.

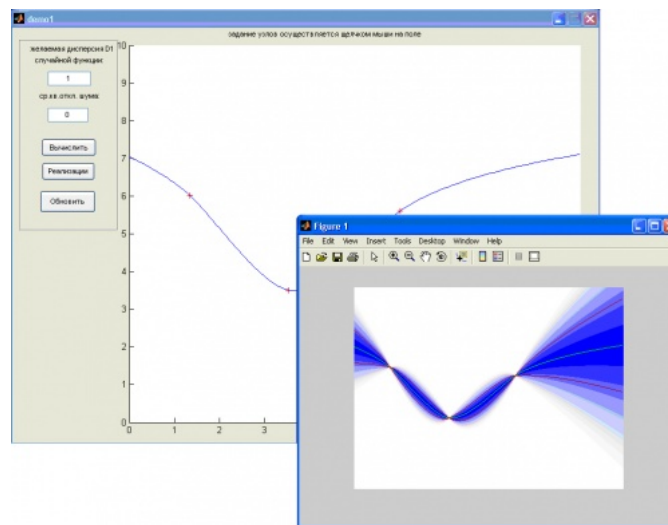


Рис.22. Демонстрация одномерной интерполяции и аппроксимации.

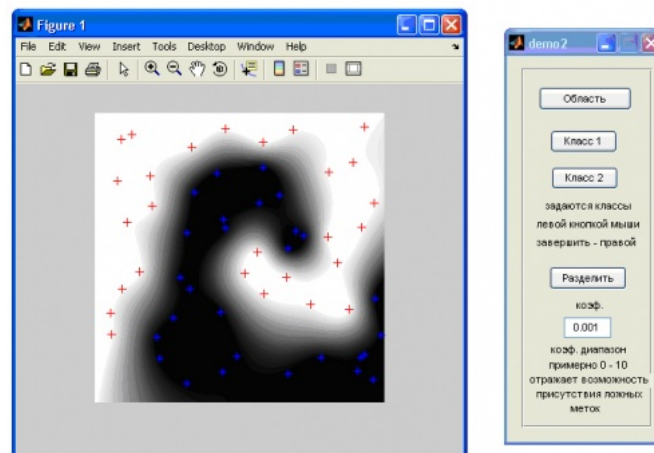


Рис.23. Демонстрация разделения двух классов на плоскости.

