

Министерство науки и высшего образования Российской Федерации

Амурский государственный университет

А.Г. Масловская

ЧИСЛЕННАЯ ОБРАБОТКА
ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ.
ФУРЬЕ- И ВЕЙВЛЕТ-АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Практикум

Благовещенск
2020

*Рекомендовано
учебно-методическим советом университета*

Рецензент:

Л.В. Никифорова, старший научный сотрудник лаборатории моделирования, обработки информации и управления АмГУ, канд. техн. наук, доцент

Масловская А.Г. Численная обработка экспериментальных данных. Фурье- и вейвлет-анализ временных рядов. Практикум / А.Г. Масловская – Благовещенск: Амурский гос. ун-т, 2020. – 55 с.

Практикум состоит из двух тематических разделов. В первом – рассматриваются вопросы обработки экспериментальных данных на основе численных методов – полиномиальной интерполяции, сплайн-интерполяции и метода наименьших квадратов. Во втором разделе представлены подходы к частотному (методом Фурье) и частотно-временному анализу (на основе вейвлет-преобразований) динамических данных, заданных временными рядами.

Приводятся краткие теоретические сведения о методах решения задач, освещаются практические аспекты обработки и анализа данных с использованием инструментария пакета прикладных программ Matlab. Практикум содержит вопросы для самопроверки и варианты индивидуальных заданий к лабораторным работам.

Учебное издание предназначено для студентов бакалавриата, обучающихся по направлению подготовки 01.03.02 – «Прикладная математика и информатика», а также может быть использовано студентами других направлений подготовки и специальностей, применяющих методы и средства математической и компьютерной обработки данных при решении прикладных задач.

ББК 22.193я73

ВВЕДЕНИЕ

В прикладной математической подготовке студентов, обучающихся по направлению «Прикладная математика и информатика», важное место занимает формирование устойчивых навыков и умений, позволяющих проводить обработку данных натурального или имитационного экспериментов с использованием методов численного анализа и инструментальных средств современных пакетов математических программ.

Первая часть практикума «Численная обработка экспериментальных данных. Фурье- и вейвлет-анализ временных рядов» посвящена теоретическим и практическим аспектам обработки экспериментальных данных на основе методов вычислительной математики – теории интерполирования и метода наименьших квадратов. Вторая часть практикума включает подходы к анализу временных рядов – частотному на основе Фурье-преобразования и частотно-временному анализу с использованием теории вейвлетов. Методическое содержание данного раздела разработано студентом магистратуры Афанасовым Л.С. в рамках педагогической практики.

Материал, содержащийся в данном издании, необходим для организации лабораторного практикума и самостоятельной работы студентов, изучающих методы и средства математической обработки данных. В практикуме приводятся базовые теоретические сведения по каждой из рассматриваемых тем, примеры решения практических задач с использованием инструментария ППП Matlab, контрольные вопросы для самопроверки, а также индивидуальные задания к лабораторным работам. Материал пособия ориентирован на студентов бакалавриата, обучающихся по направлению подготовки 01.03.02 – «Прикладная математика и информатика», а также может быть использован студентами других направлений подготовки и специальностей, применяющих методы и средства математической и компьютерной обработки данных для решения прикладных задач.

1 ПРИМЕНЕНИЕ МЕТОДОВ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ В ЗАДАЧАХ ОБРАБОТКИ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

В основе многих практических задач требуется выполнить замену одной функции $f(x)$ другой функцией $\varphi(x)$, близкой к $f(x)$ и обладающей рядом свойств, позволяющими легко производить над ней аналитические или вычислительные операции. Такая замена называется *аппроксимацией* или *приближением функции* $f(x)$ функцией $\varphi(x)$. Задача аппроксимации функции возникает, в частности, если используется табличный способ задания функции. Наиболее распространенным способом решения задачи аппроксимации основан на использовании многочленов. Наряду с многочленами применяют ряды Фурье, экспоненциальные, логарифмические и другие элементарные функции. Для оценки «близости» функций выбирают тот или иной критерий согласия. Эти критерии основаны на способах измерения расстояния между функциями.

Геометрический смысл процедуры интерполирования функции состоит в том, чтобы найти кривую $y = \varphi(x)$ некоторого определенного типа, проходящую через заданную систему точек (x_i, y_i) , $i = \overline{0, n}$.

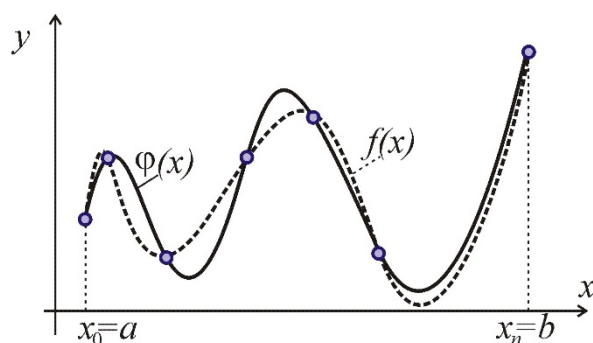


Рис. 1.1. Геометрическая интерпретация интерполирования функции.

Пусть функция $f(x)$ задана множеством своих значений для дискретного набора точек:

x_0	x_1	...	x_n
f_0	f_1	...	f_n

Здесь $f_i = f(x_i)$. Задача полиномиальной интерполяции имеет следующую формулировку: требуется найти интерполяционный многочлен $P(x) = P_n(x)$ степени не выше n , значения которого в узлах интерполяции x_j совпадают со значениями данной функции $P(x_j) = f_j$.

1.1 Интерполяционные полиномы

Одним из фундаментальных подходов, предназначенных для решения задачи интерполирования, является построение полинома Лагранжа. Для функции $f(x)$, заданной таблицей, *интерполяционный многочлен Лагранжа* имеет вид:

$$L_n(x) = \sum_{i=1}^n f_i \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)}. \quad (1.1)$$

Формула полинома Лагранжа (1.1) хотя и имеет аналитическую ценность, на практике достаточно неудобна в применении, поскольку имеет достаточно громоздкий вид, особенно при большом числе узлов. Альтернативные подходы позволяют строить интерполяционные полиномы в виде, приближенном в ряду Тейлора, в котором члены записаны в порядке убывания их значимости.

Рассмотрим формулы Ньютона для таблиц с равноотстоящими узлами. Будем считать, интерполируемая функция $f(x)$ задана своими значениями f_0, f_1, \dots, f_n а системе *равноотстоящих узлов* x_0, x_1, \dots, x_n т.е. таких, что любой узел x_i , этой *сетки* можно представить в виде:

$$x_i = x_0 + ih,$$

где $i = \overline{0, n}$, $h > 0$ – некоторая постоянная величина, называемая *шагом сетки* (таблицы).

Введем понятие конечной разности, для этого определим разности между значениями функции в узлах интерполяции. Конечная разность первого порядка имеет вид: $\Delta f_j = f_{j+1} - f_j$.

Конечная разность второго порядка задается выражением:

$$\Delta^2 f_j = \Delta f_{j+1} - \Delta f_j = f_{j+2} - 2f_{j+1} + f_j.$$

Конечная разность n -го порядка вычисляется по формуле:

$$\Delta^k f_i = f_{i+k} - k \cdot f_{i+k-1} + \frac{k(k-1)}{2!} f_{i+k-2} - \dots + (-1)^k f_i.$$

Для функции, заданной таблицей своих значений, конечные разности разных порядков удобно помещать в одну общую таблицу с узлами и значениями функции (последние можно интерпретировать как конечные разности нулевого порядка). Эту общую таблицу называют *таблицей конечных разностей*.

Пример 1. Составить таблицу конечных разностей для табулированной функции $f(x)$, отражающую зависимость спроса от цены на некоторый вид товара.

x_i , отн. ед.	0.1	0.2	0.3	0.4	0.5
f_i , отн. ед.	8.52	6.32	5.14	2.31	1.05

Используя общую концепцию построения конечных разностей, запишем итоговую таблицу:

x_i	f_i	Δf_i	$\Delta^2 f_i$	$\Delta^3 f_i$	$\Delta^4 f_i$
0.1	8.52	- 2.2	1.02	- 2.67	5.89
0.2	6.32	- 1.18	- 1.65	3.22	
0.3	5.14	- 2.83	1.57		
0.4	2.31	- 1.26			
0.5	1.05				

Первая интерполяционная формула Ньютона используется для интерполирования и экстраполирования в точках x , близких к началу таблицы – узлу x_0 , и имеет следующий вид:

$$P_n(x) = f_0 + q \cdot \Delta f_0 + \frac{q(q-1)}{2!} \Delta^2 f_0 + \dots + \frac{q(q-1) \cdot \dots \cdot (q-n+1)}{n!} \Delta^n f_0, \quad (1.2)$$

где $q = \frac{x - x_0}{h}$, число n выбирают так, чтобы конечные разности n -го порядка были практически постоянными.

Когда значение аргумент находится ближе к концу отрезка интерполяции x_n , используется *вторая интерполяционная формула Ньютона*:

$$P_n(x) = f_n + q \cdot \Delta f_{n-1} + \frac{q(q+1)}{2!} \Delta^2 f_{n-2} + \dots + \frac{q(q+1) \cdot \dots \cdot (q+n-1)}{n!} \Delta^n f_0, \quad (1.3)$$

где $q = \frac{x - x_n}{h}$.

Наряду с выведенными специально для начала и конца таблицы первой и второй интерполяционными формулами Ньютона, имеется еще несколько формул, рассчитанных на их применение в центральной части таблицы и потому называемых *центральными интерполяционными формулами*. Прежде, чем определять эти формулы, введем понятие центральных разностей.

Будем считать, что узел x_0 расположен в середине таблицы, и нумерация остальных узлов производится, начиная с x_0 с использованием как положительных, так и отрицательных индексов, т.е. считаем $x_i = x_0 + ih$, где $i = 0, \pm 1, \pm 2, \dots$

Первая интерполяционная формула Гаусса, предназначенная для интерполирования в середине таблицы представима в виде:

$$\begin{aligned} f(x) \approx \bar{P}(x_0 + qh) = & y_0 + q \cdot \Delta y_0 + \frac{q(q-1)}{2!} \Delta^2 y_{-1} + \frac{(q+1)q(q-1)}{3!} \Delta^3 y_{-1} + \\ & + \frac{(q+1)q(q-1)(q-2)}{4!} \Delta^4 y_{-2} + \dots \end{aligned} \quad (1.4)$$

где $q = \frac{x - x_0}{h}$.

Записанные слагаемые легко дополнить следующими, если знать, что в этой формуле используются нижние центральные разности все возрастающих порядков.

Вторая интерполяционная формула Гаусса, использующая верхние центральные разности имеет вид:

$$\begin{aligned} f(x) \approx \bar{P}(x_0 + qh) = & y_0 + q \cdot \Delta y_{-1} + \frac{q(q+1)}{2!} \Delta^2 y_{-1} + \frac{(q+1)q(q-1)}{3!} \Delta^3 y_{-2} + \\ & + \frac{(q-1)q(q+1)(q+2)}{4!} \Delta^4 y_{-2} + \dots \end{aligned} \quad (1.5)$$

Интерполяционные формулы Гаусса служат полуфабрикатами для получения более симметричных, использующих все центральные разности интерполяционных формул. Так, полусумма первого и второго интерполяционных многочленов Гаусса после преобразований приводит к формуле

$$f(x) \approx P_S(x_0 + qh) = y_0 + q \frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{q^2}{2!} \Delta^2 y_{-1} + \frac{q(q^2 - 1)}{3!} \frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2} + \frac{q(q^2 - 1)}{4!} \Delta^4 y_{-2} + \dots, \quad (1.6)$$

называемой *интерполяционной формулой Стирлинга*.

Если же взять полусумму второго интерполяционного многочлена Гаусса и такого же многочлена, но с нижними индексами, увеличенными на единицу (т.е. с базовой точкой x_1 вместо x_0), можно получить *интерполяционную формулу Бесселя*:

$$f(x) \approx P_B(x_0 + qh) = \frac{y_0 + y_1}{2} + \left(q - \frac{1}{2}\right) \Delta y_0 + \frac{q(q-1)}{2!} \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{q(q-1)\left(q - \frac{1}{2}\right)}{3!} \Delta^3 y_{-1} + \frac{q(q-1)(q+1)(q-2)}{4!} \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \dots \quad (1.7)$$

1.2 Сплайн-интерполяция

В тех случаях, когда промежуток, на котором нужно подменить функцию $f(x)$ функцией $\varphi(x)$, велик, и отсутствуют основания считать данную функцию $f(x)$ достаточно гладкой, нет смысла пытаться повышать качество ее полиномиальной аппроксимации за счет использования в роли $\varphi(x)$ многочленов высоких степеней. Более перспективным в этих условиях является применение *кусочно-полиномиальной аппроксимации*, предполагающей, что аппроксимирующая функция $\varphi(x)$ составляется из отдельных многочленов, как правило, одинаковой небольшой степени, определенных каждый на своей части отрезка $[a, b]$.

Использование низких степеней многочленов, составляющих $\varphi(x)$, позволяет легко находить их коэффициенты как из интерполяционных, так и из иных

условий. Так, если заданы значения y_i функции $y = f(x)$ на системе узлов x_i таких, что $a \leq x_0 < x_1 < \dots < x_n \leq b$ и требуется аппроксимировать $f(x)$ кусочно-линейной функцией $\varphi(x)$, исходя из условий интерполяции $\varphi(x_i) = y_i, i = 0, 1, \dots, n$, то, выбирая функцию $\varphi(x)$ в виде

$$\varphi(x) = \begin{cases} a_1x + b_1 & \text{при } x \in [x_0, x_1], \\ a_2x + b_2 & \text{при } x \in [x_1, x_2], \\ \dots & \dots \\ a_nx + b_n & \text{при } x \in [x_{n-1}, x_n] \end{cases}$$

для нахождения n пар ее коэффициентов a_k, b_k имеем систему из $2n$ линейных уравнений:

$$\begin{cases} \begin{cases} a_1x_0 + b_1 = y_0 \\ a_1x_1 + b_1 = y_1 \end{cases} \\ \begin{cases} a_2x_1 + b_2 = y_1 \\ a_2x_2 + b_2 = y_2 \end{cases} \\ \dots \\ \begin{cases} a_nx_{n-1} + b_n = y_{n-1} \\ a_nx_n + b_n = y_n \end{cases} \end{cases} \quad (1.8)$$

Причем, каждая пара соседних уравнений системы (5.10), имеющих коэффициенты с одинаковыми индексами, не связана с остальными и может решаться отдельно.

Аналогично, каждое звено кусочно-квадратичной функции (при $n = 2m$) определяется тройкой коэффициентов a_k, b_k, c_k которые могут быть найдены последовательным решением (при $k = \overline{1, m}$) трехмерных линейных систем, соответствующим выставленным интерполяционным условиям. Фактически, в рассмотренных случаях речь идет о последовательной линейной интерполяции по перемещаемым вдоль отрезка парам соседних точек разбиения и о последовательной квадратичной интерполяции по тройкам таких точек.

При большом количестве узлов интерполяции приходится использовать

полиномы высокой степени. Можно этого избежать, разбив отрезок интерполяции на несколько частей и построив на каждой части свой интерполяционный многочлен. Существенный недостаток такого интерполирования состоит в том, что в точках сшивки разных интерполяционных полиномов их первая производная будет разрывной, поэтому для решения задачи кусочно-линейной интерполяции используют особый вид кусочно-полиномиальной интерполяции – сплайн-интерполяцию.

Сплайн – это функция, которая на каждом частичном отрезке интерполирования является алгебраическим многочленом, а на заданном отрезке непрерывна вместе с несколькими своими производными.

Пусть на отрезке $[a, b]$ задана упорядоченная система несовпадающих точек x_k , $k = \overline{0, n}$.

Определение. Сплайном $S_m(x)$ называется определенная на $[a, b]$ функция, принадлежащая классу $C_{[a, b]}^l$ l раз непрерывно дифференцируемых функций, такая, что на каждом промежутке $[x_{k-1}, x_k]$, $k = \overline{2, n}$ – это многочлен n -й степени. Разность $d = m - l$ между степенью сплайна m и показателем его гладкости l называется *дефектом* сплайна.

Если сплайн $S_m(x)$ строится по некоторой функции $f(x)$ так, чтобы выполнялись условия $S_m(x_i) = f(x_i)$, то такой сплайн называется *интерполяционным сплайном* для функции $f(x)$ при этом *узлы сплайна* x_k , вообще говоря, могут не совпадать с узлами интерполяции x_i .

Тривиальные примеры интерполяционных сплайнов: кусочно-линейная функция $\varphi(x)$, определенная с параметрами a_k, b_k очевидно, является интерполяционным сплайном степени 1 дефекта 1, а кусочно-квадратичная функция есть интерполяционный сплайн степени 2 дефекта 2.

Совпадение дефекта сплайна с его степенью обеспечивает просто непрерывность сплайна. Интерес представляет построение сплайнов с большей гладкостью, т. е. с малым дефектом. Такие сплайны являются дальнейшим совершенствованием идеи кусочно-полиномиальной аппроксимации.

Наиболее известным и широко применяемым интерполяционным сплайном является сплайн степени 3 дефекта 1. Положим, что узлы сплайна $a = x_0, x_1, \dots, x_n = b$ одновременно служат узлами интерполяции, т.е. в них известны значения функции $f_k = f(x_k)$, $k = 0, 1, \dots, n$.

Определение. Кубическим сплайном дефекта 1, интерполирующим на отрезке $[a, b]$ данную функцию $f(x)$, называется функция

$$g(x) := \left\{ g_k(x) = a_k + b_k(x - x_k) + c_k(x - x_k)^2 + d_k(x - x_k)^3, x \in [x_{k-1}, x_k] \right\}_{k=1}^n, \quad (1.9)$$

удовлетворяющая совокупности условий: а) $g(x_k) = f_k$ – условие интерполяции в узлах сплайна; б) $g(x) \in C_{[a,b]}^2$ – двойная непрерывная дифференцируемость; в) $g''(a) = g''(b) = 0$ краевые условия.

Определенный таким образом сплайн называют еще *естественным* или *чертежным сплайном*. Несложно убедиться, что определяемая условиями а) – в) функция (11), представляющая собой кубический n -звенник с гладким сопряжением звеньев.

Замечание. Краевые условия в определении могут быть заменены на другие. Например, можно наложить дополнительные условия на первую производную функции $g(x)$ в точках a и b . В таком случае кубический сплайн, оставаясь интерполяционным дефекта 1, утрачивает свойство быть естественным. Для построения по данной функции $f(x)$ интерполирующего ее сплайна нужно найти $4n$ его коэффициентов a_k, b_k, c_k, d_k ($k = 1, 2, \dots, n$).

1.3 Графики данных

Графики данных строят в зависимости от их номера поступления в процессе наблюдения или от их номера в сортированном списке данных. Двумерная

графическая визуализация данных в среде Matlab обеспечивается следующими функциями: `plot(x,y)` – для построения точечного графика, полигонов и графиков функций; `stem(x,y)` – для построения точечного графика с визуализацией ординат значения функции, где x – одномерный массив значений абсцисс, представляющих номера наблюдаемой или измеряемой в опыте величины или значения входных данных, y – одномерный массив значений ординат, представляющих значения выходных данных или реализаций случайной переменной.

Наглядным способом представления векторных и матричных данных являются диаграммы и гистограммы. Значение элемента вектора пропорционально высоте столбика диаграммы или площади сектора диаграммы. Гистограммы используются для получения информации о распределении данных по заданным интервалам. Отображение вектора в виде столбчатой диаграммы осуществляется функцией `bar(x,y,h)`. Разметку горизонтальной оси можно задать вектором с возрастающими значениями, что учитывается в первом аргументе функции `bar`. Выбор ширины столбцов осуществляется заданием третьего дополнительного аргумента. Для учета вклада каждого из элементов вектора в общую сумму его элементов оказываются полезными круговые диаграммы `pie(x)`. Визуализация векторных данных может быть осуществлена при помощи `pie3` и `bar3`, которые строят трехмерные круговые и столбчатые диаграммы.

Для получения наглядного представления о распределении данных служит функция `hist(x,n)`. Данная последовательность команд заполняет вектор x числами, распределенными по указанному закону, разбивает интервал, которому они принадлежат, на n равных частей и строят гистограмму попадания в каждый из интервалов.

Для отображения функции двух переменных следует: сгенерировать матрицы с координатами узлов сетки на прямоугольной области определения функции; вычислить функцию в узлах сетки и записать полученные значения в матрицу; использовать одну из встроенных функций пакета. Сетка генерируется функцией `meshgrid`, вызываемой с двумя входными аргументами (диапазонами

изменения значений соответствующих координат). Функции `mesh(x,y,z)` и `surf(x,y,z)` строят каркасную поверхность графика функции, при этом вызов `surf(x,y,z)` приводит к заливке каждой клетки определенным цветом, зависящим от значения функции в точках, соответствующим углам клетки. Команда `shading flat` позволяет убрать каркасные линии, а команда `shading interp` позволяет получить поверхность плавно-залитую цветом. Команда `colorbar` выводит рядом с графиком цветовую шкалу. Линии уровня можно построить с помощью команд `meshc`, `surfc`, которые имеют те же входные аргументы, что и `mesh`, `surf`.

1.4 Численные примеры. Реализация в пакете Matlab

Пример 2. Решить задачу интерполяции с помощью многочлена Лагранжа для функциональной зависимости спроса от цены на некоторую группу товара $f(x)$, заданную таблицей 1.

1. Создадим функцию `bas_fun.m`, возвращающую значение многочлена $l_i(x)$:

```
function l=bas_fun(x,i,X,Y)
%x - абсцисса точек интерполяции, i - номер узла
%X - вектор значений абсцисс узлов интерполирования
%Y - вектор значений ординат узлов интерполирования
N=length(X);
l=1;
for j=1:N
    if not(i==j)
        l=l*(x-X(j))/(X(i)-X(j));
    end
end
l=l*Y(i);
```

2. Создадим файл-функцию `Lagr.m`, возвращающий значение полинома Лагранжа:

```
function L=Lagr(x,X,Y)
N=length(X);
```

```
L=0;
for i=1:N
    L=L+bas_fun(x,i,X,Y);
end
```

3. В скрипт-файле зададим табулированную функцию и вычислим значения полинома Лагранжа:

```
x=[0.1 0.2 0.3 0.4 0.5];
y=[8.52 6.32 5.14 2.31 1.05];
M=length(x);
N=21; %задаем число точек интерполирования
h=(x(M)-x(1))./(N-1); % рассчитываем шаг интерполирования
%вычисляем координаты промежуточных точек
for j=1:N
    X(j)=x(1)+(j-1)*h;
end
% значение полинома Лагранжа в промежуточных точках
for j=1:N
    Yp(j)=Lagr(X(j),x,y);
end
plot(x,y,'ro',X,Yp,'b-'); grid on;
```

4. Значения табличной функции и интерполированные значения функции представлены на графике:

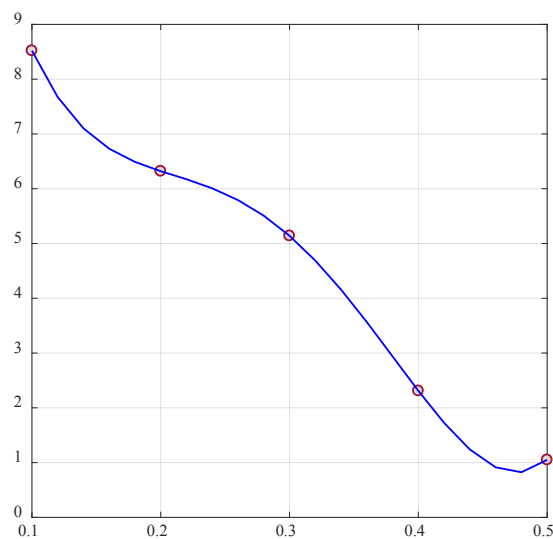


Рис. 1.2. Результат интерполирования функции с помощью полинома Лагранжа

Пример 3. Решить задачу интерполяции для функции, заданной таблично в примере 2, средствами пакета Matlab.

Для решения задачи одномерной интерполяции в пакете Matlab используется функция `interp1()`, которая реализует один из способов интерполирования: `'nearest'` – интерполяция по соседним элементам; `'linear'` – линейная интерполяция (применяется по умолчанию, если способ интерполирования не задан); `'spline'` – интерполяция кубическими сплайнами; `'pchip'` – интерполяция кубическими эрмитовыми сплайнами.

Для демонстрации использования функционала Matlab для решения задачи интерполирования продолжим запись предыдущего скрипт-файла:

```
Ylin=interp1(x,y,X,'linear');  
Ysp=interp1(x,y,X,'spline');  
Yne=interp1(x,y,X,'nearest');  
Ypc=interp1(x,y,X,'pchip');  
figure (2)  
plot(x,y,'o',X,Ylin,X,Ysp,X,Yne,X,Ypc); grid on;
```

Исходные данные и результат всех видов интерполяции представлены на рисунке 1.3.

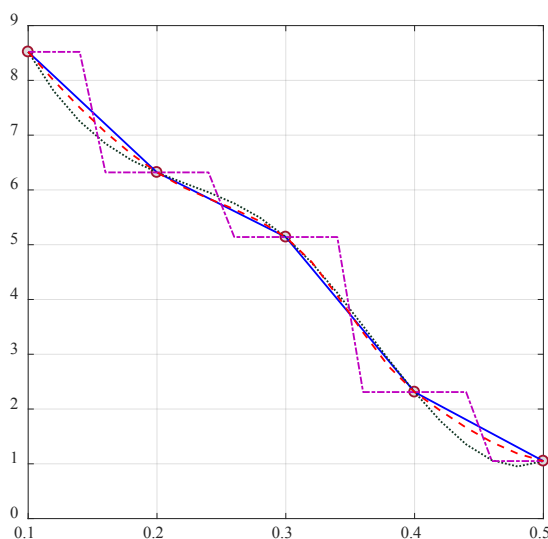


Рис. 1.3. Результат интерполирования функции с помощью встроенной функции пакета Matlab `interp1`.

Альтернативным подходом может оказаться использование функции пакета Matlab `spline()`. Покажем решение задачи с помощью кубических сплайнов:

```
figure (3); Yspline=spline(x,y,X);  
plot(x,y,'o',X,Yspline); grid on;
```

Визуализация исходных данных и результатов кубической сплайн-интерполяции показана на графике (рисунок 1.4).

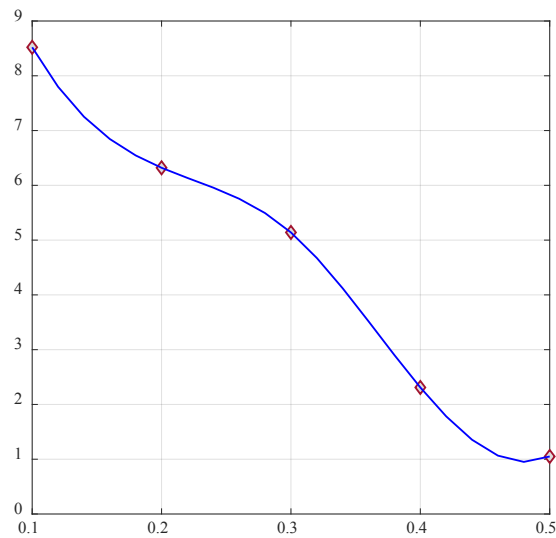


Рис. 1.4. Результат интерполирования функции с помощью встроенной функции пакета Matlab `spline`.

Пример 4. В Matlab можно также осуществить интерполяцию многомерных данных. Для интерполирования двумерных данных следует задать промежуточные узлы командой `meshgrid` и воспользоваться функцией `interp2`, которая реализует один из способов интерполирования:

- ‘nearest’ – интерполяция по соседним элементам,
- ‘bilinear’ или ‘linear’ – билинейная интерполяция,
- ‘bicubic’ или ‘cubic’ – интерполяция бикубическими сплайнами,
- ‘spline’ – интерполяция кубическими сплайнами.

Пример использования функций, реализующих двумерную интерполяцию, показан на следующем рисунке 1.5.

```
X,Y]=meshgrid(0:0.2:1);
```



```

Z=sin(3*pi*X).*sin(3*pi*Y).*exp(-X.^2-Y.^2);
subplot(2,2,1)
surf(X,Y,Z)
title('tabular function')
[Xi,Yi]=meshgrid(0:0.02:1);
ZNear=interp2(X,Y,Z,Xi,Yi,'nearest');
ZBiLin=interp2(X,Y,Z,Xi,Yi,'bilinear');
ZBiCub=interp2(X,Y,Z,Xi,Yi,'bicubic');
subplot(2,2,2)
surf(Xi,Yi,ZNear);
title('near')
subplot(2,2,3)
surf(Xi,Yi,ZBiLin);
title('bilinear')
subplot(2,2,4)
surf(Xi,Yi,ZBiCub);
title('cubic')

```

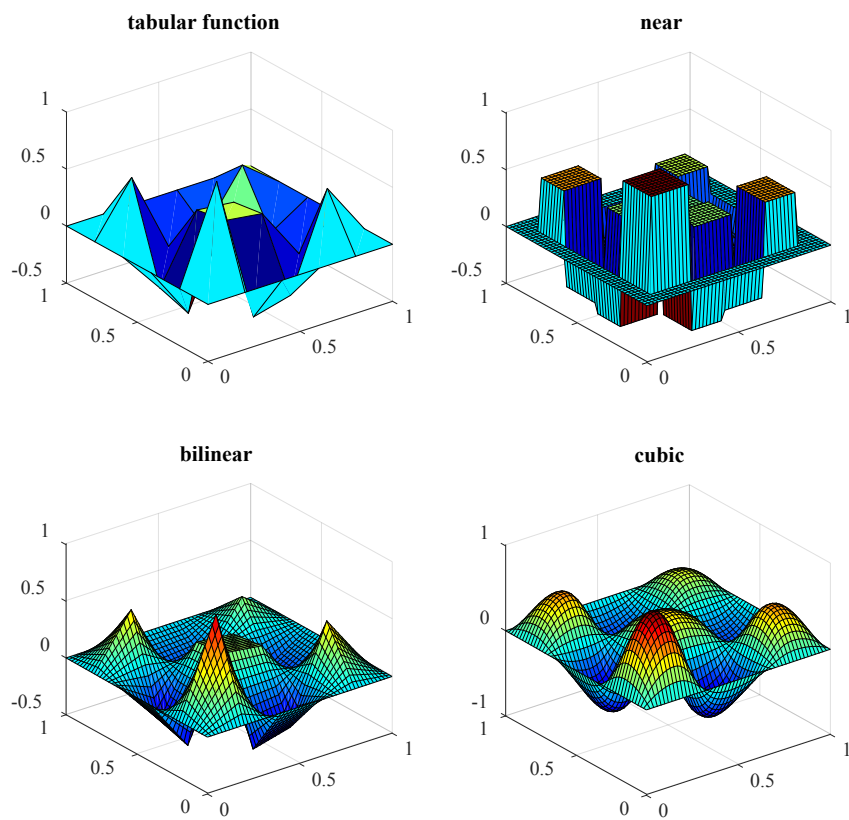


Рис. 1.5. Интерполяция многомерных данных.

Контрольные вопросы

1. В каких практических случаях может потребоваться аппроксимация функции?
2. Какие точки называются узлами интерполяции? Какие узлы называются равноотстоящими?
3. При решении каких задач используются интерполяционные формулы Ньютона, Гаусса, Бесселя, Стирлинга?
4. В каких случаях применяется сплайн-интерполяция? Какой недостаток «кусочного» интерполирования с помощью многочленов Лагранжа или Ньютона устраняется при интерполяции сплайнами?

Индивидуальные задания

1. Дать геометрическое представление табулированной функции. Двумерную геометрическую реализацию обеспечить следующими функциями:

plot(x, y) – для построения точечного графика,

bar(x, y) и **line(x, y)** – для построения столбчатой диаграммы,

stem(x, y) – для построения точечного графика с визуализацией ординат значения функции

2. Применяя методы интерполяции решить прикладную задачу, соответствующую варианту задания. Использовать следующие способы аппроксимации:

а) полиномиальная интерполяция (многочлен Лагранжа или Ньютона),

б) встроенная функция пакета Matlab **interp1()** с различными способами интерполяции (по соседним элементам, линейную интерполяцию, кубическими сплайнами, кубическими эрмитовыми сплайнами).

Исходя из эмпирического анализа данных и сравнительного анализа результатов всех видов интерполяции, сделать выводы.

Вариант 1

1. Найти непрерывный график зависимости изменения температуры в городе Благовещенске по заданным табличным данным и вычислить примерное значение температуры в 14 час. 30 мин.:

t , час.	13-00	14-00	15-00	16-00	17-00	18-00
T , °C	22	23.5	25	23	21	19

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности двумерной пластины представлены на рисунке (симметричное распределение). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 1 см по абсциссам и ординатам графика.

144	103	75	24	116	174
118	94	56	42	101	145
102	91	49	38	98	119
101	90	45	34	95	102
118	94	56	42	101	145
144	103	75	24	116	174

Вариант 2

1. Построить график зависимости значений удельной магнитной восприимчивости χ_e от температуры, используя различные подходы. Установить численно приближенное значение удельной магнитной восприимчивости при температуре 1200 К, сравнить с экспериментальным – -0.1055.

T , К	100	200	400	600	700	800	1000
χ , 10^{-6} см ³ /Г	-0.110	-0.107	-0.102	-0.099	-0.097	-0.095	-0.096

2. При решении задачи упругого кручения призматических стержней получили следующий набор дискретных значений U – функции напряжений (симметричный набор данных). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 0.2 см по абсциссам и ординатам графика.

26	24	18	11	19	28
24	23	15	9	14	25
15	19	16	9	13	18
16	18	13.2	11	10.1	17.5
23	19	15.5	10.5	8	15.1
25.5	18.1	13	10.3	9	12

Вариант 3

1. Построить график изменения ежемесячных средних значений курса акций компании ГАЗПРОМ за некоторый календарный период, используя различные подходы. Численно оценить значение курса в ноябре 2019 года.

N	Май 2019	Июл 2019	Авг 2019	Окт 2019	Дек 2019	Фев 2020	Март 2020
Стоимость, руб.	162	207	188	224	246	229	186

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности некоторой сплошной среды представлены на рисунке. Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 0.5 см по абсциссам и ординатам графика.

105	89	42	37	69	154
118	97	42	16	35	66
67	52	34	39	75	95
78	56	18	29	38	54
122	98	75	49	55	99
190	95	74	89	152	135

Вариант 4

1. Построить график зависимости стоимости нефти Brent за баррель в течение определенного периода времени (данные зафиксированы 1 марта каждого года). Аппроксимировать численно стоимость в сентябре 2017 г. и сравнить с зафиксированным значением 57.54 \$.

N	2015	2016	2017	2018	2019	2020
$S, \$$	55.11	39.6	53.83	70.27	68.39	32.95

2. При решении задачи определения потенциала U течения грунтовых вод под перемычкой получили следующий набор дискретных значений U (симметричный набор данных). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 1.5 см по абсциссам и ординатам графика.

9.63	7.25	6.58	5.02	6.89	10.2
6.9	5.53	4.21	3.2	6.31	2.69
5.6	4.52	3.09	2.89	1.44	5.2
6.3	5.2	3.1	0.69	1.2	4.2
8.6	5.6	3.2	1.02	0.99	2.69
8.91	1.1	0.23	1.1	2.32	0.16

Вариант 5

1. Зависимость изменения температуры грунта на Марсе, зафиксированная Curiosity (данные получены метеостанцией REMS, установленной на борту марсохода и охватывают 16 по 17 августа 2012 г.), представлена таблицей данных.

<i>Sol</i>	10.2	10.4	10.6	10.8	11.0
<i>T</i> , °C	-90	-20	0	-45	-84

Построить аппроксимацию функциональной зависимости температуры от времени (в солах), используя интерполяционный полином Ньютона. Записать выражение для вычисления приближенного значения температуры в марсианский полдень (*Sol*=10.5).

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности двумерной пластины представлены на рисунке (симметричное распределение). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 2.5 см по абсциссам и ординатам графика.

121	99	52	40	59	123
114	97	49	36	42	78
69	52	34	39	75	80
20	15	12	10	11	15
29	25	20	21	25	29
136	98	71	65	52	38

Вариант 6

1. Построить график зависимости удельного сопротивления медно-никелевых сплавов от содержания компонентов. Установить численно приближенное значение сопротивления при содержании компонентов 70 %.

$N, \%$	0	20	40	53	58	63	100
$\rho, 10^{-6}$ Ом·см	2	27	45	50	52	50	8

2. При решении задачи упругого кручения призматических стержней получили следующий набор дискретных значений U – функции напряжений (симметричный набор данных). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 1 см по абсциссам и ординатам графика.

	15.11	9.21	7.44	2.52	6.22	15.1
4.9	3.53	2.21	1.2	5.36	2.64	
5.6	6.52	5.09	3.89	2.55	5.4	
7.3	5.2	1.1	0.69	1.2	3.2	
5.6	2.6	1.2	0.45	2.03	5.3	
8.23	4.9	0.23	1.1	4.32	5.15	

Вариант 7

1. Нормальная термопара Ле-Шателье (Pt-Pt10%Rh) рекомендована международным комитетом как основной прибор при точных измерениях температур. Однако на практике термопара не применяется и служит для контроля показаний рабочих термопар. Построить график зависимости значений электродвижущей силы E от температуры T , найти значение э.д.с. при значении температуры 950°C .

$T, ^{\circ}\text{C}$	0	200	400	600	800	1000	1200	1400	1600
$E,$ мВ	0	1.44	2.26	5.24	7.34	9.61	11.96	14.36	16.73

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности трубы представлены на рисунке. Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 0.5 см по абсциссам и ординатам графика.

102	98	75	44	56	118
95	44	36	28	53	75
62	56	39	27	44	85
47	35	29	18	17	66
89	75	64	52	69	111
112	95	65	79	85	155

Вариант 8

1. Построить график зависимости скорости звука в углекислом газе (при $t=50.8$ °С) от давления, используя различные подходы. Установить численно приблизительное значение скорости при значении давления 90 атм.

P , атм.	0	20	40	80	100	120	130	140
v , м/с	294	265	259	221	218	262	291	330

2. При решении задачи определения потенциала U течения грунтовых вод под перемычкой получили следующий набор дискретных значений U (симметричный набор данных). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 0.5 см по абсциссам и ординатам графика.

10.44	7.24	6.58	3.18	6.55	15.6
6.9	5.53	4.21	3.2	6.31	2.69
7.6	6.52	5.09	3.89	2.45	5.2
9.3	8.2	2.1	1.69	1.2	3.2
7.6	3.6	2.2	1.45	2.03	5.3
6.91	5.1	1.23	2.1	3.32	5.16

Вариант 9

1. Для анализа покупательской активности в некотором супермаркете зафиксирована численность покупателей в различное время суток. Оценить число покупателей в 15-30, построить непрерывный график зависимости.

t	8-00	9-00	10-00	12-00	14-00	16-00	17-30	18-00	20-00
N	33	52	78	126	115	108	136	158	85

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности двумерной пластины представлены на рисунке (симметричное распределение). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 0.5 см по абсциссам и ординатам графика.

118	94	45	39	62	114
113	97	42	36	35	81
67	52	34	39	75	95
25	13	12	19	28	44
29	25	18	21	25	29
144	95	71	65	52	35

Вариант 10

1. Построить график зависимости среднестатистических значений количества выпадающих осадков на о. Хайнань. Установить приблизительное значение числа выпавших осадков в ноябре, используя различные подходы. Сравнить с экспериментально зафиксированным значением 146.8 мм.

месяц	янв	фев	мар	апр	май	июн	авг	окт	дек
H , мм	56.7	43	79.8	133.3	240.3	179.8	275.9	229.5	62.9

2. Экспериментальные данные распределения температуры (в усл.ед.) в некоторый момент времени на поверхности двумерной пластины представлены на рисунке (симметричное распределение). Используя возможности интерполирования многомерных данных дать графическую интерпретацию этим результатам. Координаты сетки соответствуют размеру единичного отрезка 1.5 см по абсциссам и ординатам графика.

26	25	18	10	19	28
24	21	12	9	11	25
19	19	15	9	13	18
11	15	12.2	10	9.1	18.5
24	23	15	9	14	25
36	34	28	21	29	38

2 ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

В некоторых случаях значения функции $y = f(x)$ заданы с некоторой погрешностью и информацию о ее поведении дает только статистическая совокупность данных. Интерполяционный подход при этом оказывается несостоятельным, поскольку он предполагает точное совпадение исходной и аппроксимирующей функций в узлах интерполяции. Если значений табулированной функции будет достаточно много, интерполяция может привести к появлению большого числа экстремумов, не дающих адекватного представления о поведении анализируемых данных. Альтернативой является использование приближение в смысле метода наименьших квадратов.

Пусть в результате измерений получена таблица некоторой зависимости $y = f(x)$:

x	x_1	x_2	...	x_n
$f(x)$	y_1	y_2	...	y_n

Сформулируем задачу следующим образом – найти функцию вида $\varphi(x)$, которая в точках x_1, x_2, \dots, x_n принимает значения, близкие к табличным значениям y_1, y_2, \dots, y_n . Предположим, что приближающая функция $\varphi(x)$ в точках x_1, x_2, \dots, x_n имеет значения $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n$. Тогда нужно найти функцию $\varphi(x)$ определенного вида так, чтобы сумма квадратов отклонений значений табулированной функции y_i от аппроксимирующей \bar{y}_i

$$(y_1 - \bar{y}_1)^2 + (y_2 - \bar{y}_2)^2 + \dots + (y_n - \bar{y}_n)^2 \quad (2.1)$$

была наименьшей (рисунок 2.1).

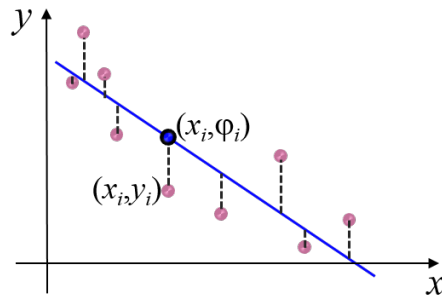


Рис. 2.1. Геометрическая интерпретация метода наименьших квадратов.

2.1 Нахождение приближающей функции в виде линейной

Пусть приближающая функция имеет вид

$$\varphi(x) = ax + b, \quad (2.2)$$

где a, b – параметры.

Составим сумму вида (2.1) для этого случая:

$$\Phi(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2 \rightarrow \min. \quad (2.3)$$

Функция $\Phi(a, b)$ является неотрицательной квадратичной, поэтому в некоторой области она имеет единственную точку минимума (a^*, b^*) , удовлетворяющую условиям:

$$\frac{\partial \Phi}{\partial a} = 0, \quad \frac{\partial \Phi}{\partial b} = 0, \text{ т.е. системе уравнений:}$$

$$\begin{cases} \sum_{i=1}^n (y_i - ax_i - b)x_i = 0, \\ \sum_{i=1}^n (y_i - ax_i - b) = 0. \end{cases} \quad (2.4)$$

Запишем СЛАУ с неизвестными a и b в общепринятом виде:

$$\begin{cases} \left(\sum_{i=1}^n x_i^2 \right) \cdot a + \left(\sum_{i=1}^n x_i \right) \cdot b = \sum_{i=1}^n x_i y_i, \\ \left(\sum_{i=1}^n x_i \right) \cdot a + n \cdot b = \sum_{i=1}^n y_i. \end{cases} \quad (2.5)$$

Вычислив значения параметров a, b , получаем конкретный вид функции (2.2).

В зависимости от характера табличных данных, изучаемого с помощью их изображения в соответствующей системе координат, при обработке экспериментальных данных часто используют иные семейства двухпараметрических функций. Существует возможность с помощью подходящего преобразования переменных получить линейную зависимость и использовать метод (2.5) для следующих семейств функций, показанных в таблице 2.1.

Таблица 2.1 – Замена переменных при линеаризации представителей семейства двухпараметрических функций

Функция $y = f(x)$	Линеаризованная форма	Замена переменных и постоянных для $Y = AX + B$
$y = \frac{a}{x} + b$	$y = a \cdot \frac{1}{x} + b$	$X = \frac{1}{x}, Y = y, A = a, B = b$
$y = \frac{d}{x+c}$	$y = -\frac{1}{c} \cdot xy + \frac{d}{c}$	$X = xy, Y = y,$ $A = -\frac{1}{c}, B = \frac{d}{c}$
$y = \frac{1}{ax+b}$	$\frac{1}{y} = ax + b$	$X = x, Y = \frac{1}{y}, A = a, B = b$
$y = \frac{x}{ax+b}$	$\frac{1}{y} = a + b \frac{1}{x}$	$X = \frac{1}{x}, Y = \frac{1}{y}, A = a, B = b$
$y = a \cdot \ln(x) + b$	$y = a \cdot \ln(x) + b$	$X = \ln(x), Y = y, A = a, B = b$
$y = c \cdot e^{ax}$	$\ln(y) = ax + \ln(c)$	$X = x, Y = \ln(y), A = a, B = \ln(c)$
$y = c \cdot X^a$	$\ln(y) = a \cdot \ln(x) + \ln(c)$	$X = \ln(x), Y = \ln(y), A = a, B = \ln(c)$
$y = \frac{1}{(ax+b)^2}$	$\frac{1}{\sqrt{y}} = ax + b$	$X = x, Y = \frac{1}{\sqrt{y}}, A = a, B = b$
$y = \frac{c \cdot x}{e^{dx}}$	$\ln\left(\frac{y}{x}\right) = -dx + \ln(c)$	$X = x, Y = \ln\left(\frac{y}{x}\right), A = -d, B = \ln(c)$
$y = \frac{1}{1+c \cdot e^{ax}}$	$\ln\left(\frac{1}{y} - 1\right) = ax + \ln(c)$	$X = x, Y = \ln\left(\frac{1}{y} - 1\right), A = a, B = \ln(c)$

2.2 Нахождение приближающей функции в виде квадратного трехчлена

Приближающая функция имеет вид:

$$\varphi(x) = ax^2 + bx + c, \quad (2.6)$$

где a, b, c – параметры.

Составим сумму вида (2.1) как функцию $\Phi(a, b)$ для этого случая:

$$\Phi(a, b) = \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2 \rightarrow \min. \quad (2.7)$$

Функция $\Phi(a, b)$ является неотрицательной квадратичной, поэтому в некоторой области она имеет единственную точку минимума (a^*, b^*) , удовлетворяющую условиям: $\frac{\partial \Phi}{\partial a} = 0, \quad \frac{\partial \Phi}{\partial b} = 0$, т.е. системе уравнений:

$$\begin{cases} \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)x_i^2 = 0, \\ \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)x_i = 0, \\ \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c) = 0. \end{cases} \quad (2.8)$$

Запишем СЛАУ с неизвестными a, b и c в общепринятом виде:

$$\begin{cases} \left(\sum_{i=1}^n x_i^4 \right) \cdot a + \left(\sum_{i=1}^n x_i^3 \right) \cdot b + \left(\sum_{i=1}^n x_i^2 \right) \cdot c = \sum_{i=1}^n x_i^2 y_i, \\ \left(\sum_{i=1}^n x_i^3 \right) \cdot a + \left(\sum_{i=1}^n x_i^2 \right) \cdot b + \left(\sum_{i=1}^n x_i \right) \cdot c = \sum_{i=1}^n x_i y_i, \\ \left(\sum_{i=1}^n x_i^2 \right) \cdot a + \left(\sum_{i=1}^n x_i \right) \cdot b + n \cdot c = \sum_{i=1}^n y_i. \end{cases} \quad (2.9)$$

Решив систему относительно неизвестных a, b, c , находим значения параметров приближающей функции (2.6).

2.3 Численные примеры. Реализация в пакете Matlab

Пример 1. Даны табличные значения функциональной зависимости прибыли предприятия за некоторый календарный период. Выполнить аппроксимацию данных методом наименьших квадратов, используя квадратичную функцию (2.6).

x	0	1	2	3	4	5	6	7	8	9	10
y	1.2	1.5	2.1	4.8	7.7	12.4	12.1	15.4	24.1	28.9	31.4

Создадим скрипт-файл, реализующий метод наименьших квадратов:

```
x=0:1:10;
n=length(x);
y=[1.2 1.5 2.1 4.8 7.7 12.4 12.1 15.4 24.1 28.9 31.4];
A(1,1)=sum(x.^4);
A(1,2)=sum(x.^3);
A(1,3)=sum(x.^2);
A(2,1)=A(1,2);
A(2,2)=A(1,3);
A(2,3)=sum(x);
A(3,1)=A(1,3);
A(3,2)=A(2,3);
A(3,3)=n;
B(1,1)=dot(x.^2,y);
B(2,1)=dot(x,y);
B(3,1)=sum(y);
C=A^-1*B
xpr=0:0.001:10;
a=C(1,1);
b=C(2,1);
c=C(3,1);
ypr=a*xpr.^2+b*xpr+c;
plot(x,y,'o',xpr,ypr); grid on
```

В результате работы программы получаем значение параметров приближающей функции (2.6):

$$C = \begin{matrix} 0.2585 \\ 0.6167 \end{matrix}$$

На рисунке 2.2 представлены исходные данные и аппроксимирующая функция.

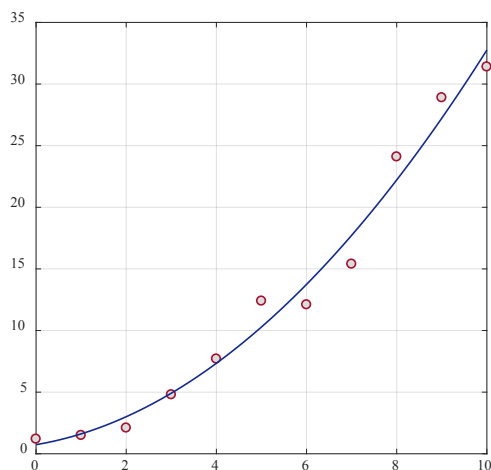


Рис. 2.2. Аппроксимация квадратичной зависимостью табулированной функции.

Пример 2. Используя табличные значения примера 1, найти параметры квадратичной аппроксимирующей функции (2.6) с помощью встроенной функции Matlab.

Для решения задачи обобщенной нелинейной регрессии в пакете Matlab имеется функция `lsqnonlin()`, возвращающая решение задачи нахождения точки минимума функции:

$$\min_x (f(x)) = f_1^2(x) + f_2^2(x) + \dots + f_n^2(x) + L,$$

где $f(x)$ – вектор-функция, x – столбец искомых переменных, L – искомая константа.

Сначала создадим файл-функцию `fun_mnk.m`, которая возвращает значение вектор-функции $f(x)$:

```
function z=fun_mnk(C,x,y)
N=length(x);
i=1:N
z(i)=y(i)-(C(1)*x(i).^2+C(2)*x(i)+C(3));
```


Далее проведем вычисления значения аппроксимирующей функции с помощью функции `lsqnonlin()`, и визуализируем исходные данные и аппроксимирующую функцию:

```
z=[0 0 0];  
C=lsqnonlin('fun_mnk',z,[],[],[],x,y);  
a=C(1);  
b=C(2);  
C=C(3);  
xpr=0:0.001:10;  
ypr=a*xpr.^2+b*xpr+c;  
plot(x,y,'s',xpr,ypr); grid on
```

Параметры приближающей функции:

```
C =    0.2585    0.6167    0.7413
```

График исходных данных и приближающей функции:

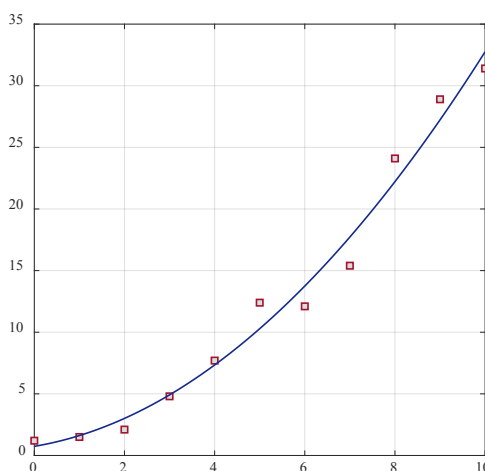


Рис. 2.3. Аппроксимация квадратичной зависимостью табулированной функции с помощью встроенной функции пакета Matlab.

Приближение табличной функции одной переменной полиномом заданного порядка по методу наименьших квадратов может быть реализовано в пакете Matlab с помощью функции `polyfit(x,y,n)`, в которой первые два входных аргумента являются векторами со значениями абсцисс и ординат табличной функции, а третий – требуемой степенью полинома.

Пример 3. Приблизить функцию, заданную таблицей в предыдущих примерах, полиномами четвертой, пятой и шестой степени и вывести график, отражающий характер приближений.

Реализация поставленной задачи приведена в следующем скрипт-файле:

```
xx=0:0.001:10;  
y4=polyfit(x,y,4);  
y5=polyfit(x,y,5);  
y6=polyfit(x,y,6);  
P4=polyval(y4,xx);  
P5=polyval(y5,xx);  
P6=polyval(y6,xx);  
plot(x,y,'d',xx,P4,xx,P5,xx,P6); grid on
```

Результат аппроксимации представлен в графическом окне (рисунок 2.4).

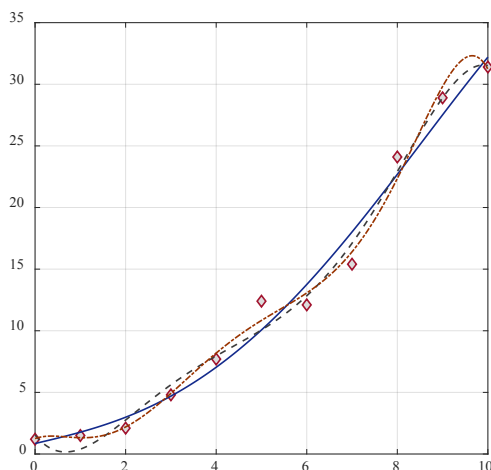


Рис. 2.4. Аппроксимация табулированной функции с помощью встроенной функции пакета Matlab.

Пример 4. Процедуру подбора эмпирических зависимостей продемонстрируем на примере третьего закона Кеплера о движении планет.

Период движения планеты на орбите в днях может быть выражен как функция расстояния (табл. 2.1):

$$T = Cx^{\frac{3}{2}},$$

где C – константа, x – расстояния от Солнца в млн. км.

Таблица 2.1 – Зависимость периода движения некоторых планет Солнечной системы от расстояния

Планета	$x_i, 10^6$ км	T_i , день
Меркурий	57.59	87.99
Венера	108.11	224.77
Земля	149.57	365.26
Марс	227.84	686.98

Выполняя последовательность алгоритма метода наименьших квадратов, можно установить численно значение константы C :

$$\Phi(C) = \sum_{i=1}^n (T_i - Cx_i)^2 \rightarrow \min, \text{ из чего следует:}$$

$$\frac{d\Phi}{dC} = -2 \sum_{i=1}^4 \left(T_i - Cx_i^{\frac{3}{2}} \right) x_i^{\frac{3}{2}} = 0, \text{ откуда } \left[\sum_{i=1}^4 \left(x_i^{\frac{3}{2}} \right)^2 \right] \cdot C = \sum_{i=1}^4 T_i \cdot x_i^{\frac{3}{2}}, \text{ и далее:}$$

$$C = \frac{\sum_{i=1}^4 T_i \cdot x_i^{\frac{3}{2}}}{\left[\sum_{i=1}^4 x_i^3 \right]} = 0.1998.$$

Программная реализация будет представлена листингом:

```
% initial data
x=[57.59 108.11 149.57 227.84];
t=[87.99 224.77 365.26 686.98];

%calculation of parameter C
A=sum(x.^3);
B=sum(t.*(x.^1.5));
C=B/A

% plot the graph
X=57.59:0.01:227.84;
```

```
T=C*(X.^1.5);  
plot(x,t,'rd',X,T,'b-');  
grid on
```

Контрольные вопросы

1. В чем суть приближения таблично заданной функции по методу наименьших квадратов? Чем этот метод отличается от метода интерполяции?
2. Какие элементарные функции чаще всего используются в качестве приближающих?
3. Как можно добиться повышения качества приближения?
4. Какая ошибка является среднеквадратической?
5. Какое из двух приближений одной и той же таблично заданной функции считается лучшим?
6. Каким образом построение приближающих функций в виде различных элементарных функций сводится к случаю линейной функции?
7. За счет чего возникает полиномиальное раскачивание?

Индивидуальные задания

1. Для данного набора экспериментальных данных выполнить аппроксимацию, используя метод наименьших квадратов и, соответственно, следующие приближения:

- а) линейной функцией,
- б) нелинейной функцией, допускающей линеаризацию. Осуществить замену переменных, получить линеаризованную форму.

(Примеры описания таких функций и соответствующих преобразований см. в таблице – описание к лабораторной работе). Построить графики исходных данных и полученных аппроксимирующих функций. Рассчитать значение среднеквадратического отклонения в каждом из случаев. Сделать вывод о качестве и наилучшем виде аппроксимации.

2. Найти аппроксимирующую функцию средствами пакета Matlab с помощью функции `lsqnonlin()` (функцию выбрать самостоятельно исходя из эмпирических сведений об объекте) и приближении полиномами разных степеней `polyfit(x,y,n)`.

3. Сравнить все приближения на основе оценки среднеквадратического отклонения, сделать вывод о наиболее подходящей аппроксимации. Оформить отчет по лабораторной работе.

Вариант 1

В таблице представлены данные зависимости температуры от времени при наблюдении процесса остывания жидкости.

t , мин	0	2	6	10	14	20
T , °C	92	79	64	45	38	22

Вариант 2

В таблице приведены сведения о зависимости предложения при изменении цены на определенные группы товара.

P , отн. ед.	21	23	28	36	38	44
S , отн. ед.	99	105	141	186	193	228

Вариант 3

В таблице представлена временная зависимость зафиксированных значений числа атомов изотопа плутония ^{242}Pu при радиоактивном распаде.

t , сек· 10^5	0	2	4	8	16	20
N , ед· 10^{21}	2.467	1.845	1.400	0.790	0.261	0.131

Вариант 4

Зависимость значений потока энергии радиоизлучения Солнца F_v в зависимости от длины волны λ (данные измерения при спокойном Солнце и максимальном значении пятен) представлена в таблице.

λ , см	600	150	6	0.6
$F_v, \cdot 10^{-22}$, Вт/(м ² ·Гц)	0.49	9.4	155	3400

Вариант 5

В таблице представлены динамические данные о числе заболевших коронавирусом 2019-nCoV за определённый период времени 2020 года.

Дата	20 янв	25 янв	08 фев	25 фев	9 мар
N , чел.	278	1075	34886	80888	111235

Вариант 6

В таблице представлены динамические данные о динамике изменения концентрации патогенных бактерий в некотором объеме.

t , день	0	1	3	5	7	10	15
$C \cdot 10^7$ моль/литр	0.0023	0.0125	3.1244	5.5667	8.5023	9.1451	9.2455

Вариант 7

В таблице представлены данные регистрации процесса зарядки конденсатора: зависимость процента максимального заряда от времени.

t , пост. вр.	0	0.5	1	2	3	4	5
δ , %	0	39	63	86	95	98	99

Вариант 8

В таблице представлена динамика посещения банка клиентами в течение суток.

t , час.	9-00	10-00	12-00	13-00	14-00	18-00	20-00
N , чел.	16	28	96	112	42	86	21

Вариант 9

В таблице приведены данные зависимости числа заболевших ветряной оспой детей некоторого детского сада.

t , день	0	2	5	7	10	15	20
N , чел.	1	3	14	19	23	27	28

Вариант 10

В таблице представлены динамические данные о концентрации сульфата меди в некотором водоеме.

t , день	0	0.2	0.5	0.7	1.5	4
C , мг/литр	0	0.5	2.3	2.5	1.1	0.1

3 ФУРЬЕ И ВЕЙВЛЕТ-АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

3.1 Фурье-анализ временных рядов в ППП Matlab

Для вычисления дискретного преобразования Фурье (ДПФ) или обратного ДПФ последовательности из N элементов требуется выполнить N^2 операций. Если число элементов исходных массивов составляет порядка нескольких тысяч и более, то время, за которое будет выполнено преобразование, резко возрастает и теряется возможность обработки сигналов в реальном масштабе времени.

В 60-е годы XX века Кули и Тьюки предложили метод вычисления коэффициентов Фурье, который позволил снизить объем вычислений до $N \log_2 N$ операций. Данный метод получил название алгоритма быстрого преобразования Фурье (БПФ). В настоящее время процедуры данного алгоритма входят буквально во все математические библиотеки прикладных программ. Рассмотрим основную идею БПФ для случая, когда отсчетов $N = 2^p$, p – целое число.

Разобьем входную последовательность $\{s_k\}$ на две части с нечетными и четными номерами:

$$\{s_k\}_E = \{s_{2k}\}, \{s_k\}_O = \{s_{2k+1}\}, \quad (3.1)$$

где $k = 0, 1, 2, \dots, \frac{N}{2} - 1$.

Данное действие позволит представить n -ый коэффициент ДПФ в виде:

$$\begin{aligned} C_n &= \frac{1}{N} \sum_{k=0}^{\frac{N}{2}-1} \left(s_{2k} e^{-i \frac{2\pi n}{N} 2k} + s_{2k+1} e^{-i \frac{2\pi n}{N} (2k+1)} \right) = \\ &= \frac{1}{N} \left(\sum_{k=0}^{\frac{N}{2}-1} s_{2kE} e^{-i \frac{2\pi n}{N} 2k} + e^{-i \frac{2\pi n}{N}} \sum_{k=0}^{\frac{N}{2}-1} s_{2k+1O} e^{-i \frac{2\pi n}{N} 2k} \right). \end{aligned} \quad (3.2)$$

Из (1) видно, что коэффициенты ДПФ из первой половины исходного сигнала с номерами от 0 до $\frac{N}{2} - 1$ находятся через коэффициенты ДПФ двух частных последовательностей:

$$C_n = C_{nE} + e^{-i\frac{2\pi n}{N}} C_{nO}, \quad (3.3)$$

где $n = 0, 1, 2, \dots, \frac{N}{2} - 1$.

Исходя из того, что последовательности коэффициентов массивов $\{s_k\}_E$ и $\{s_k\}_O$ являются периодическими с периодом $\frac{N}{2}$, то

$$C_{nE} = C_{n+\frac{N}{2}E}, C_{nO} = C_{n+\frac{N}{2}O}. \quad (3.4)$$

Выражение для второй половины множества коэффициентов ДПФ получаем из подстановки (3.4) в (3.3):

$$C_{n+\frac{N}{2}} = C_{nE} - e^{-i\frac{2\pi n}{N}} C_{nO}, \quad (3.5)$$

где $n = 0, 1, 2, \dots, \frac{N}{2} - 1$

Формулы (3.3), (3.5) лежат в основе алгоритма БПФ: последовательности отсчетов с нечетными и четными номерами вновь разбиваются на две части, и продолжается это до тех пор, пока не получится последовательность, состоящая из одного элемента.

В пакете Matlab одномерное преобразование Фурье реализовано двумя функциями, выполняющие прямое и обратное БПФ: `fft` и `ifft` соответственно. Данные функции можно использовать как для действительных, так и для комплексных последовательностей, при этом длина последовательностей может быть любой.

Вызов функций: `fft(v)` – данная функция вернет дискретное преобразование Фурье 2^m -мерного вектора, аргумент которого результат дискретизации через равные промежутки времени некоторой функции. Результат функции – комплексный вектор размерности $2^m + 1$. Элементы данного вектора, возвращаемый функцией `fft`, вычисляется по формуле:

$$c_n = \sum_{k=0}^{N-1} v_k e^{\frac{i2\pi nk}{N}},$$

где N – размерность вектора v ; $fft(v)$ – функция возвращает обратное дискретное преобразование Фурье. Передаваемый в качестве параметра вектор v должен иметь размерность $2^m + 1$ элементов. Результатом функции будет действительный вектор размерностью 2^{m+1} . Вычисление элементов вектора для данной функции производится по формуле:

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} v_k e^{-\frac{i2\pi nk}{N}},$$

где N – число элементов вектора v .

Воспользуемся функциями вычисления БПФ пакета Matlab на примере сигнала, заданного в модельном представлении в виде комбинации синусоид:

$$s(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t), \quad (3.6)$$

где $A_1 = 0,8$, $A_2 = 0,4$ – амплитуды компонент сигнала; $f_1 = 1$ Гц, $f_2 = 10$ Гц – частотные компоненты сигнала.

Шаг 1. Инициализация исходного сигнала в ППП Matlab:

```
N = 256;
i = 1:N;
Tmax = 2; %длительность временного интервала в секундах
t(i) = Tmax / (N-1) * (i-1); %задание временной сетки
f1 = 1; %частота первой составляющей, Гц
f2 = 10; %частота второй составляющей, Гц
%вычисление мгновенных значений функции
f(i) = .8*sin(2*pi*f1*t(i)) + .4*sin(2*pi*f2*t(i));
figure(1);
plot(t, f), grid
```

Полученное графическое представление сигнала показано на рисунке 3.1.

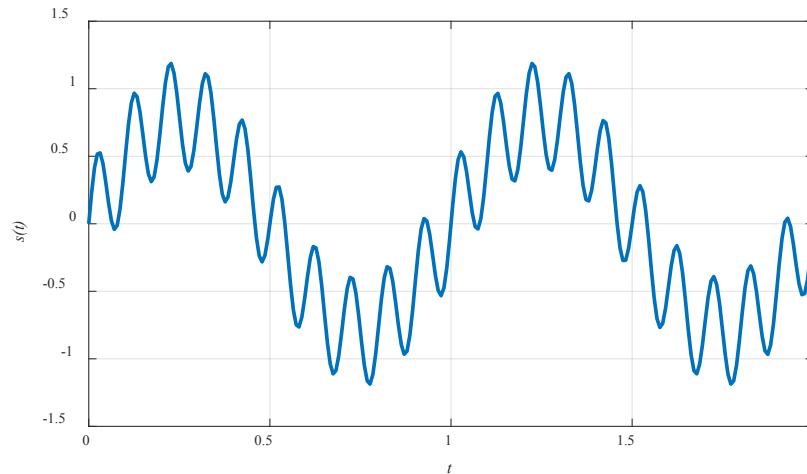


Рис. 3.1. Исходный искусственно созданный временной сигнал с частотами 1 и 10 Гц.

Шаг 2. Проведем Фурье-анализ для данного временного сигнала. Воспользуемся встроенной функцией в Matlab $fft(x)$, осуществляющей быстрое непрерывное преобразование Фурье. Для этого выполним последовательность команд:

```

с = fft(f); % вычисление спектра
j = 2:N/2;
% вычисление нормированного амплитудно-частотного спектра
Cm(j-1) = abs(c(j-1))/(N/2);
%вычисление вектора частот в Гц
Freq(j-1) = (j-1)/Tmax;
figure(2)
plot(Freq, Cm), grid
axis([0 100 0 1]);

```

Полученный в результате Фурье-преобразования график в Matlab, представлен на рисунке 3.2.

Как видно из графика, Фурье-преобразование позволило локализовать частоты сигнала, значение которых соответствует максимумам кривой (1 Гц и 10 Гц).

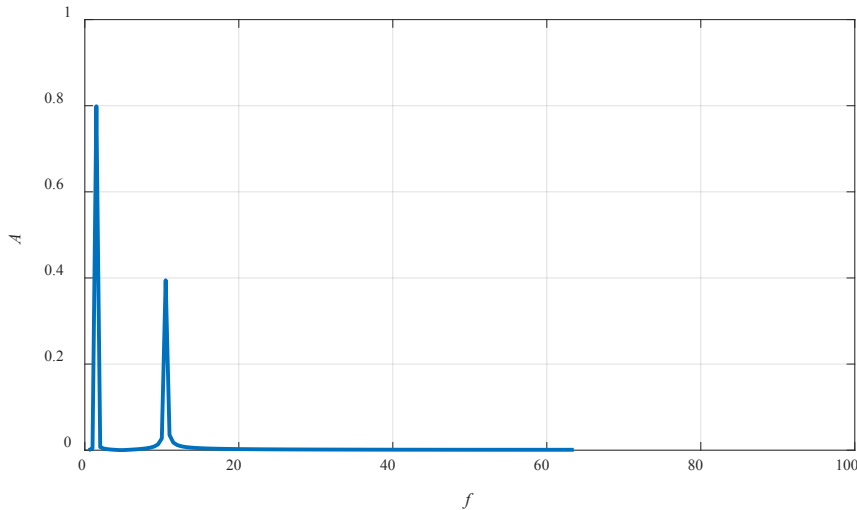


Рис. 3.2. Результат Фурье-преобразования для сигнала (3.6).

3.2 Пример использования вейвлет-преобразования для анализа временных рядов в Matlab

В отличие от преобразования Фурье, локализующего частоты, но не дающего временного разрешения процесса, и от аппарата d -функций, локализующего моменты времени, но не имеющего частотного разрешения, вейвлет-преобразование, обладающее самонастраивающимся подвижным частотно-временным окном, одинаково хорошо выявляет как низкочастотные, так и высокочастотные характеристики сигнала на разных временных масштабах. По этой причине вейвлет-анализ часто сравнивают с «математическим микроскопом», вскрывающим внутреннюю структуру существенно неоднородных объектов.

Непрерывное вейвлет-преобразование строится с помощью непрерывных масштабных преобразований и переносов вейвлета $y(t)$ с произвольными значениями масштабного коэффициента a и параметра сдвига b :

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt. \quad (3.7)$$

Преобразование обратимо:

$$f(t) = \int_{0-\infty}^{\infty} \int_{-\infty}^{\infty} W(a, b) \psi\left(\frac{t-b}{a}\right) \frac{dad b}{a^2}. \quad (3.8)$$

Поскольку вейвлет-преобразование есть скалярное произведение анализирующего вейвлета на заданном масштабе и анализируемого сигнала, коэффициенты $W(a,b)$ содержат комбинированную информацию об анализирующем вейвлете и анализируемом сигнале.

Выбор анализирующего вейвлета, как правило, определяется тем, какую информацию необходимо извлечь из сигнала. Каждый вейвлет имеет характерные особенности во временном и частотном пространстве, поэтому иногда с помощью разных вейвлетов можно выявить и подчеркнуть те или иные свойства анализируемого сигнала.

Примеры различных вейвлетов приведены в таблице 3.1.

Таблица 3.1 – Виды вейвлетов

Наименование	Функция	Изображение
HAAR	$\psi(t) = \begin{cases} 1, 0 \leq t < \frac{1}{2} \\ -1, \frac{1}{2} \leq t < 1 \\ 0, t < 0, t \geq 1 \end{cases}$	
FNAT	$\psi(t) = \begin{cases} 1, t < \frac{1}{3} \\ -\frac{1}{2}, \frac{1}{3} \leq t < 1 \\ 0, t > 1 \end{cases}$	
WAVE	$\psi(t) = t \exp\left(-\frac{t^2}{2}\right)$	
MHAT	$\psi(t) = (1-t^2) \exp\left(-\frac{t^2}{2}\right)$	

Наименование	Функция	Изображение
Вейвлет Морле	$\psi(t) = \cos(5t) \cdot \exp\left(-\frac{t^2}{2}\right)$	

Рассмотрим вейвлет-преобразование для того же сигнала (3.6). Для этого воспользуемся встроенной $cwt(x)$ функцией в ППП Matlab, осуществляющей непрерывное вейвлет-преобразование.

Программный код имеет вид:

```

N = 256;
i = 1:N;
Tmax = 2; %длительность временного интервала в секундах
t(i) = Tmax/(N-1)*(i-1); %задание временной сетки
f1 = 1; %частота первой составляющей, Гц
f2 = 10; %частота второй составляющей, Гц
%вычисление мгновенных значений функции
f(i) = .8*sin(2*pi*f1*t(i)) + .4*sin(2*pi*f2*t(i));
figure(1);
plot(t,f),grid
figure(2)
c2 = cwt(f, 1:64, 'mexh', 'lvl');
title('скелетон, линии локальных экстремумов')

```

В приведенном примере рассматривается вейвлет-преобразование с использованием вейвлета «мексиканская шляпа». Построенный график скелетон сигнала представлен на рисунке 3.3.

Как видно из рисунка 3.3 на скелетоне присутствуют светлые и темные области: максимумы функции коэффициентов вейвлет-преобразования $W(a,b)$ соответствуют светлым областям, а минимумы – темным.

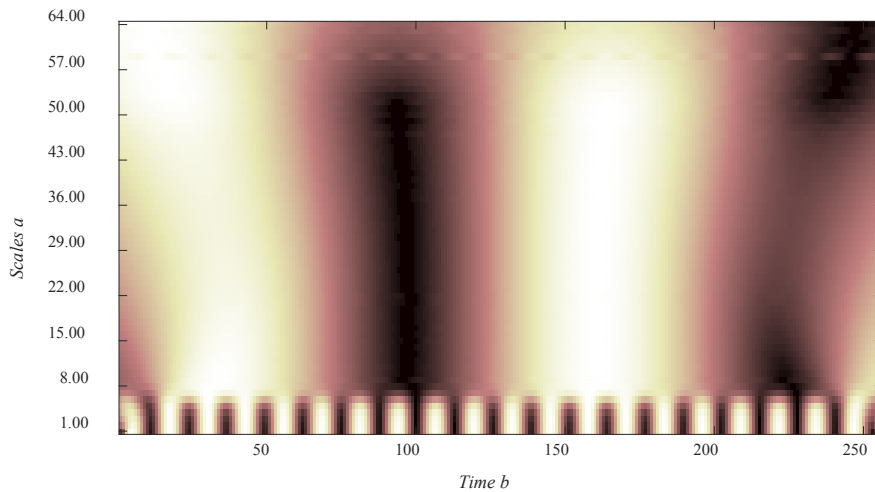


Рис. 3.3. Скелетон (линии локальных экстремумов) для сигнала (3.6).

Построим поверхность $W(a,b)$ и его проекцию для полученного вейвлет-преобразования. Фрагмент программы в Matlab приведен ниже:

```
figure (3)
D = c2;
[X1,Y1] = meshgrid(1:1:N,1:1:64);
%построение поверхности W(a,b)
surf(X1,Y1,D); colormap(jet);
title('Коэффициенты вейвлет-преобразования W(a,b)-частотно-
временная локализация')
xlabel('b'), ylabel('a'), zlabel('W(a,b)')

figure (4)
surf(X1,Y1,D); colormap(jet);
view(2);
title('Коэффициенты вейвлет-преобразования W(a,b) - проекция на
(a,b)')
```

Полученные графики приведены на рисунке 3.4 и рисунке 3.5.

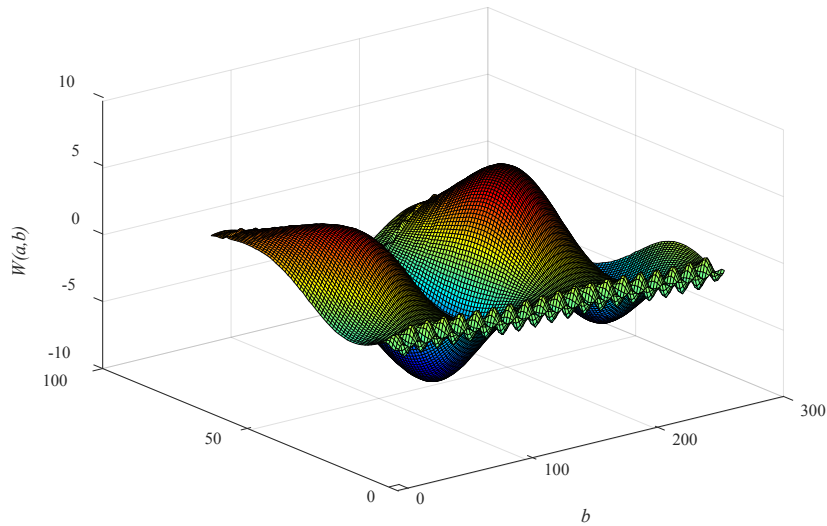


Рис. 3.4. График поверхности коэффициентов вейвлет-преобразования $W(a,b)$

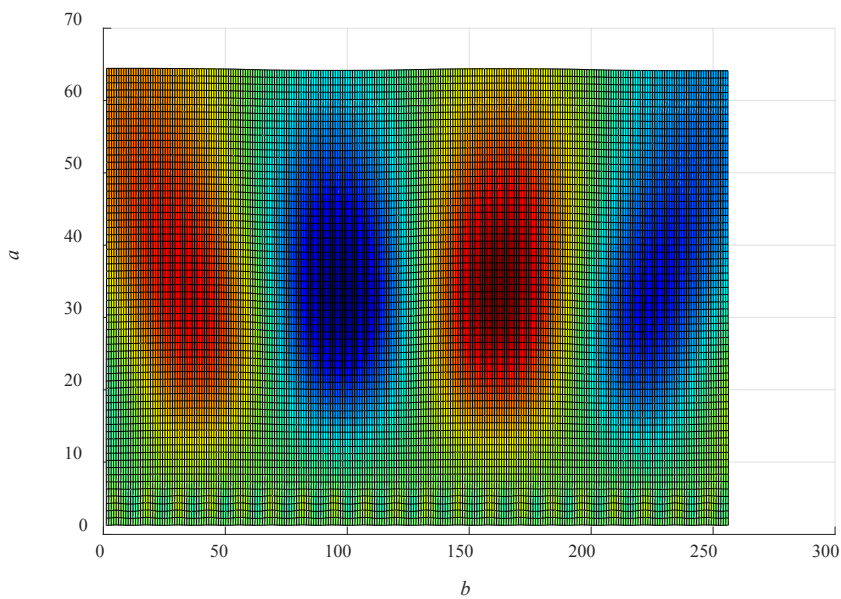


Рис. 3.5. Проекция $W(a,b)$ на плоскость (a,b) для сигнала (3.6).

3.3 Пример применения Фурье- и вейвлет-преобразования для обработки динамических данных

В качестве объекта исследования выбран сигнал, заданный изменением количества солнечных пятен с течением времени. Данные были импортированы с

сайта LISIRD (LASP Interactive Solar Irradiance Datacenter) [5] и представляли собой временной ряд как зависимость количества пятен на Солнце за период с 1974 г. по 2018 г. от времени. График исходных данных представлен на рисунке 3.6.

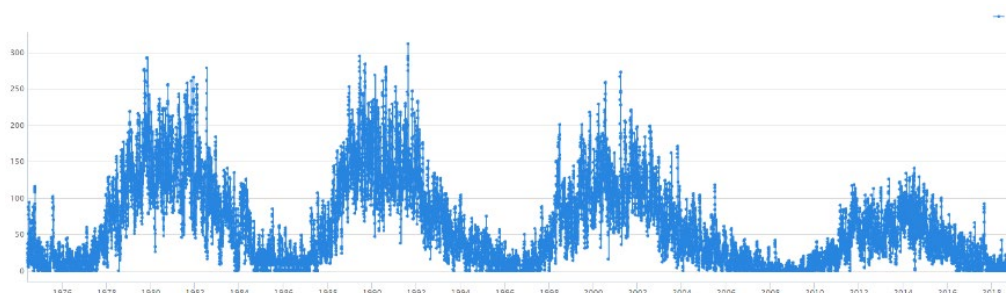


Рис. 3.6. Исходный временный ряд [5].

Проведем Фурье-преобразование для локализации частот. Результат, полученный в ППП Matlab, показан на рисунке 3.7.

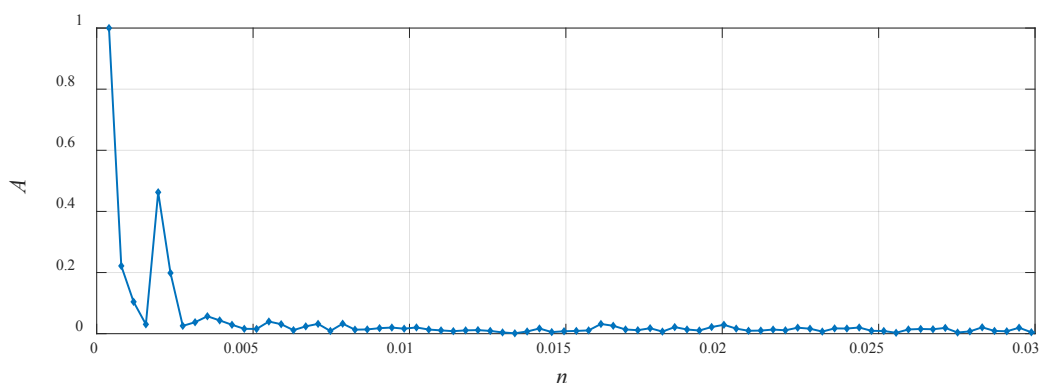


Рис. 3.7. Локализация частот анализируемого сигнала.

Из графика видно, что в данном сигнале присутствует низкочастотные и высокочастотные составляющие. Фурье-преобразование локализует частоты временного сигнала, но не локализует его особенности во времени. Как известно, преобразование Фурье дает хороший результат для линейных и стационарных сигналов. Динамика солнечной активности представляет нестационарный физический процесс. Поэтому для его анализа далее рассмотрим результат вейвлет-преобразования.

Вейвлет-преобразование в отличие от Фурье-преобразования, позволяет проводить частотно-временную локализацию. Результатом вейвлет-преобразования временного ряда является двумерный массив амплитуд – значение коэффициентов $W(a,b)$. Спектр $W(a,b)$ сложного временного сигнала представляет собой поверхность в трехмерном пространстве. Для анализа были использованы два типа вейвлетов: «мексиканская шляпа» (mhat) и «биортогональный вейвлет» (bior). Полученные в ППП Matlab графики поверхности коэффициентов вейвлет-преобразования $W(a,b)$ за последний год представлены на рисунке 8. Также визуализируем проекцию для коэффициентов $W(a,b)$ на плоскость (a,b) для отображения вейвлет-спектров или скелетонов сигнала. Результаты показаны на рисунке 9.

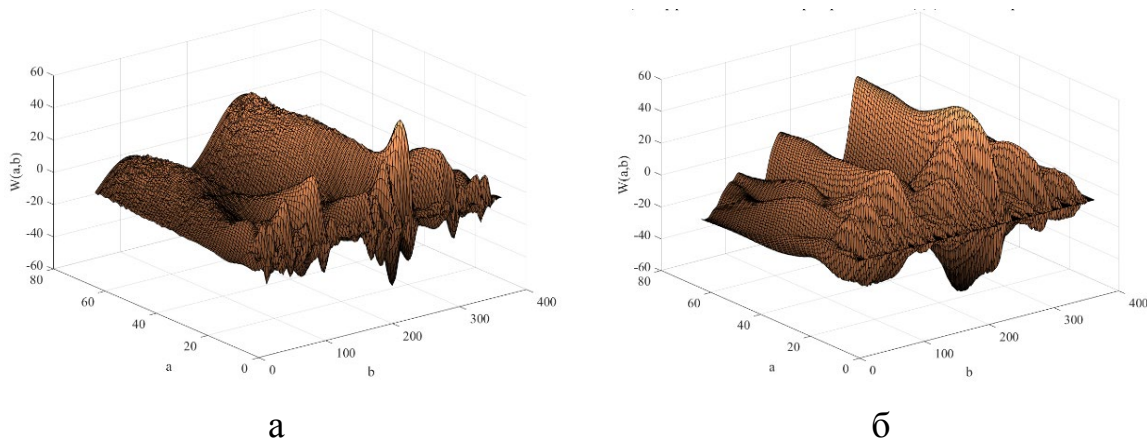


Рис. 3.8. Поверхности коэффициентов вейвлет-преобразования $W(a,b)$ с использованием анализаторов: а) «мексиканская шляпа»; б) «биортогональный вейвлет».

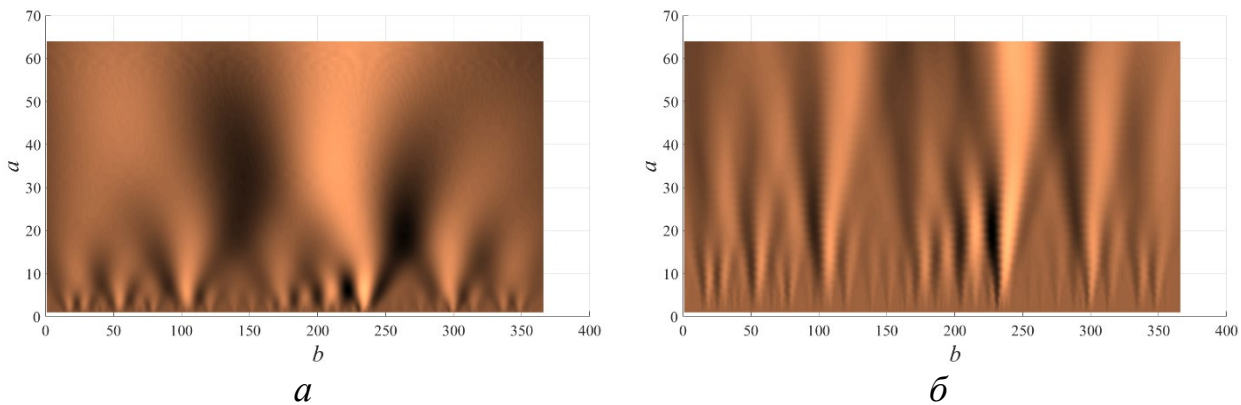


Рис. 3.9. Вейвлет-спектры с использованием анализаторов: а) «мексиканская шляпа»; б) «биортогональный вейвлет».

Скелетон линий локальных экстремумов отображает светлые и темные области, соответствующие максимумам и минимумам соответственно. Во всем временном диапазоне сигнала в большом количестве присутствуют высокочастотные составляющие (на малых масштабах a), в меньшем – низкочастотные (на больших масштабах a). Скелетоны сигнала позволяют наглядно видеть их распределение по частотно-временной области.

Контрольные вопросы

1. Какая встроенная функция в пакете Matlab выполняет прямое быстрое преобразование Фурье?
2. Какая функция в пакете Matlab выполняет обратное быстрое преобразование Фурье?
3. Что позволяет выявить на графике Фурье-преобразование исходного сигнала?
4. Какие виды анализирующих вейвлетов вы знаете?
5. Какая функция в ППП Matlab используется для осуществления вейвлет-преобразования?
6. Что позволяет выявить полученный «скелетон» исходного сигнала?

Индивидуальные задания

1. Визуализировать исходный ряд динамики изменения курса заданной валюты за последний год. Данные инициализировать, используя ресурсы сайта Центрального банка Российской Федерации [6].
2. Построить частотные характеристики с помощью Фурье-преобразования.
3. Построить частотно-временные характеристики с помощью вейвлет-преобразования.
- 4*. Выполнить импорт данных курса заданной валюты в период кризиса (01.08.2014 – 30.12.2015) и провести анализ полученных данных за последний год. Наблюдается ли изменения в спектрах в период кризиса?

Курсы валют для каждого варианта представлены в таблице 3.2.

Таблица 3.2 – Курсы заданных валют по вариантам

Вариант, №	Курс валюты
1	Рубль/Австралийский доллар
2	Рубль/Вона КНДР
3	Рубль/Белорусский рубль
4	Рубль/Датская крона
5	Рубль/Доллар США
6	Рубль/Евро
7	Рубль/Канадский доллар
8	Рубль/Норвежская крона
9	Рубль/Турецкая лира
10	Рубль/Украинская гривна

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Дьяконов, В. П. MATLAB 6.5 SP1/7/7 SP1/7 SP2 + Simulink 5/6. Инструменты искусственного интеллекта и биоинформатики [Электронный ресурс] : монография / В. П. Дьяконов, В. В. Круглов. — Электрон. текстовые данные. — М. : СОЛОН-ПРЕСС, 2009. — 454 с. — 5-98003-255-X. — Режим доступа: <http://www.iprbookshop.ru/8683.html> .
2. Петров, А.В. Моделирование процессов и систем. [Электронный ресурс] – Электрон. дан. – СПб.: Лань, 2015. – 288 с. – Режим доступа: <http://e.lanbook.com/book/68472>.
3. Поршнева, С.В. Компьютерное моделирование физических процессов в пакете MATLAB. + CD. [Электронный ресурс] – Электрон. дан. – СПб.: Лань, 2011. – 736 с. – Режим доступа: <http://e.lanbook.com/book/650> .
4. Самарский, А.А. Математическое моделирование: Идеи, методы, проблемы: моногр. / А.А. Самарский, А.П. Михайлов, – 2-е изд., испр. – М.: Москва: Физматлит, 2005. – 320 с.
5. База данных солнечных пятен [Электронный ресурс] // Lasp Interactive Solar Irrudiance Datacenter.: офиц.сайт. – 2005. – Режим доступа: http://lasp.colorado.edu/lisird/data/american_relative_sunspot_number_daily. – 21.02.2020.
6. Динамика официального курса заданной валюты [Электронный ресурс] // Центральный банк Российской Федерации.: офиц.сайт. – 2000. – Режим доступа: https://www.cbr.ru/currency_base/dynamics/. – 21.02.2020.

СОДЕРЖАНИЕ

Введение	3
1 Применение методов вычислительной математики в задачах обработки экспериментальных данных	4
1.1 Интерполяционные полиномы	5
1.2 Сплайн-интерполяция	8
1.3 Графики данных	11
1.4 Численные примеры. Реализация в пакете Matlab	13
2 Обработка экспериментальных данных методом наименьших квадратов	27
2.1 Нахождение приближающей функции в виде линейной	28
2.2 Нахождение приближающей функции в виде квадратного трехчлена	30
2.3 Численные примеры. Реализация в пакете Matlab	30
3 Фурье и вейвлет-анализ временных рядов	40
3.1 Фурье-анализ временных рядов в ППП Matlab	40
3.2 Пример использования вейвлет-преобразования для анализа временных рядов в Matlab	44
3.3 Пример применения Фурье- и вейвлет-преобразования для обработки динамических данных	48
Библиографический список	53

*Анна Геннадьевна Масловская,
профессор кафедры математического анализа и моделирования АмГУ,
д-р физ.-мат. наук, доцент*

Численная обработка экспериментальных данных. Фурье- и вейвлет-анализ временных рядов. Практикум

Изд-во АмГУ. Подписано к печати 15.03.2020. Формат 60×84/16.

Усл. печ. л. 4. Тираж 100. Заказ ***.

Отпечатано в типографии АмГУ.