

# Аппроксимация нелинейных характеристик

## Постановка задачи аппроксимации.

В общем случае использование методов моделирования важно, когда мы имеем сложные объекты или реальные системы, на поведение которых влияет множество различных взаимосвязанных факторов.

С точки зрения математического моделирования при решении многих практических задач часто приходится вычислять значения каких-то функциональных зависимостей  $y = f(x)$ . При этом, как правило, имеют преобладающее место две ситуации.

1. Явная зависимость между  $x$  и  $y$  на интервале  $[a, b]$  отсутствует, а имеется только таблица экспериментальных данных  $\{x_i, y_i\}, i = 1, \dots, n$  и возникает необходимость определения функции  $y = f(x)$  на любом интервале  $\{x_i, x_{i+1}\} \in [a, b]$ . К этой задаче относится также уточнение таблиц экспериментальных данных.

Кроме того, при изучении поведения того или иного объекта реальной системы, описываемой непрерывными макроскопическими параметрами иногда удобно воспользоваться их дискретной математической моделью либо набор входных параметров (вектора состояния) объекта бывает задан лишь в табличном виде. **Дискретизация** – это переход от непрерывной к дискретной функции.

2. Зависимость  $y = f(x)$ . известна и непрерывна, но настолько сложна, что не пригодна для практических расчетов. Стоит задача упрощения вычисления значений функции  $y = f(x)$  и ее характеристик (производной, интеграла и т.д.). Поэтому, с точки зрения экономии времени, вычислительных и материальных ресурсов, приходят к необходимости построения какой-то другой функциональной зависимости  $Y = F(x)$ , которая была бы близка к  $y = f(x)$  по основным ее параметрам, но более проста и удобна в реализации при последующих расчетах.

В общем случае, например, анализ процесса преобразования сигналов в нелинейных цепях радиоэлектронных систем является весьма сложной задачей, что связано с проблемой решения нелинейных дифференциальных уравнений. При этом неприменим принцип суперпозиции, так как параметры нелинейной цепи при воздействии одного источника входного сигнала отличаются от её параметров при подключении нескольких источников. Однако исследование нелинейных

цепей удаётся осуществить сравнительно простыми методами, если нелинейный элемент (НЭ) отвечает условиям безынерционности.

Физически безынерционность НЭ означает мгновенное установление отклика на его выходе вслед за изменением входного воздействия. Строго говоря, безынерционных (резистивных, или омических, т.е. только поглощающих энергию входного сигнала) практически не существует. Все нелинейные элементы – диоды, транзисторы, аналоговые и цифровые микросхемы, – обладают инерционными свойствами. Но в то же время современные полупроводниковые приборы достаточно совершенны по своим частотным параметрам и их удаётся идеализировать с точки зрения безынерционности.

Так что, задача приближенного описания подобных нелинейных функций представляет большой практический интерес не только при синтезе линейных электрических цепей, но и при моделировании работы различных полупроводниковых и иных приборов и систем.

Следует отметить, что оба эти преобразования существуют не только в моделях. Так, в радиотехнике и радиоэлектронике для сигналов их выполняют специальные устройства, которые называются ЦАП (цифро-аналоговый преобразователь) и АЦП (аналого-цифровой преобразователь). ЦАП осуществляет переход к аналоговому описанию процесса, а АЦП - дискретизацию.

Так что, способы представления характеристик относительно простыми функциями, лишь приближенно отображающими истинные характеристики, получили широкое распространение.

Суть задачи в одномерном случае состоит в замене приближенным аналитическим выражением (моделью)  $F(x(t))$  исходной функции  $f(x(t))$ ,  $x(t) = x_1(t), \dots, x_m(t)$ , заданной на замкнутом интервале  $[a, b]$  аналитически, графиком или таблицей. **Замена истинной характеристики приближенно представляющей ее функцией называется аппроксимацией характеристики.**

Функцию  $f(x)$  называют **аппроксимируемой** функцией, а  $F(x)$  — **аппроксимирующей**. С латинского *ap-proximo*- означает «почти близкий». Аппроксимирующая функция должна упрощать решение конкретной задачи.

Для аппроксимации (приближения) используются функции  $F(x)$ , принадлежащие некоторому определённом классу. Существует много разных вариантов задачи о приближении функции, зависящих от того, какие функции  $f(x)$  аппроксимируются, какие функции  $y = F(x)$ , используются для аппроксимации, как понимается близость функций  $f(x)$ ,

и  $F(x)$ , то есть какой критерий эффективности  $W$  используется и какие ограничения  $q(x, y)$  накладываются при решении задачи, и т.д.

Зачастую можно предположить, что в задачах аппроксимации функций нормированное пространство, например,  $L_2(x, y, q)$  - соответствует *среднеквадратическому* приближению, а пространство  $C(x, y, q)$  - *равномерному* приближению. И тогда наиболее часто при решении практических задач модель  $Y = F(X)$  выбирается зависящей от каких-то свободных варьируемых параметров эксперимента, т.е.  $F(X) = \varphi(x, \alpha_1, \dots, \alpha_n) = \varphi(x, \vec{\alpha})$ . Значения вектора варьируемых параметров  $\vec{\alpha}$  как раз и выбираются так, чтобы обеспечить выполнение критерия  $W$  и ограничений  $q(x, y)$ , т.е. каких-то заданных условий близости для  $f(x)$  и  $F(x)$ .

В зависимости от способа подбора вектора  $\vec{\alpha}$ , получают различные виды аппроксимации.

Если приближение строится на каком-то дискретном множестве аргументов  $\{x_i\}, i = 1, \dots, n$  то аппроксимация называется точечной. К ней относится интерполирование, среднеквадратичное приближение (МНК). Если множество  $\{x_i\}$  непрерывно и задано например, в виде отрезка  $[a, b]$ , аппроксимация называется непрерывной или интегральной (полиномы Чебышева). В настоящее время на практике хорошо изучена и широко применяется линейная аппроксимация, при которой  $\varphi(x, \vec{\alpha})$  выбирается линейно зависящий от параметров  $\vec{\alpha}$  в виде так называемого обобщенного многочлена:

$$F_n(x, \alpha_n) = \sum_{j=0}^n \alpha_j \varphi_j(x), \quad (1)$$

где  $\vec{\alpha}$  — вектор варьируемых параметров;  $\{\varphi_j(x)\}$  — некоторая выбранная линейно-независимая система базисных функций. В качестве их могут быть, например:

- алгебраическая:  $1, x, x^2, \dots, x^n, \dots$ ;
- тригонометрическая:  $1, \sin(x), \cos(x), \dots, \sin(nx), \cos(nx), \dots$ ;
- экспоненциальная:  $e^{a_0x}, e^{a_1x} \dots e^{a_nx}, \dots$ ;

Здесь коэффициенты  $\{a_i\}$  представляют собой некоторую числовую последовательность попарно различных действительных чисел.

Важным является, чтобы эта система была полной, т.е. обеспечивающей аппроксимацию посредством  $Y = F(X)$  с заданной точностью на всех интервалах  $[a, b]$  определения  $f(x)$ .

Наибольший практический интерес в пространстве  $L_2$  представляют ортогональные функции, а в пространстве  $C$  — степенные.

Аппроксимация (1) степенным полиномом является одним из наиболее распространенных способов. Например, характеристика  $F(x)$

нелинейного элемента в виде степенного полинома упрощает проведение спектрального анализа при полигармоническом воздействии и, кроме того, позволяет получить точные результаты. Но оптимальный выбор способа аппроксимации зависит от вида нелинейной характеристики, а также от режима работы нелинейного элемента в цепи.

### **Интерполирование непрерывной функции полиномами**

В вычислительной математике существенную роль играет интерполяция функций, т.е. построение по заданной функции другой (как правило, более простой). Она по определению предполагает нахождение промежуточных неизвестных значений величины заданной таблицей или графиком по некоторым ее опорным значениям. Относительно функциональных зависимостей она является одним из основных видов точечной аппроксимации.

**Интерполяция** - это вычисление значений модели  $Y = F(X)$  во всей области определения аргумента (входного параметра  $X$ ) по заданному дискретному множеству точек, т.е. переход от дискретной функции к непрерывной. От латинского: *inter* – «внутри», *pole* – «узел».

При интерполировании требуется, чтобы в определённых точках (узлах интерполяции) совпадали значения исходной функции  $f$  и аппроксимирующей модельной функции  $F$ . В более общем случае должны совпадать значения производных в заданных точках.

Иногда интерполяцию и аппроксимацию рассматривают как синонимы. Однако все же они различаются. **Интерполяция**, в отличие от аппроксимации, - это переход от дискретной к непрерывной функции, проходящей точно через заданные точки.

Причем интерполяция имеет как практическое, так и теоретическое значение. Так теория интерполирования используется при построении и исследовании квадратурных формул для численного интегрирования, для получения методов решения дифференциальных и интегральных уравнений.

Простейшая задача интерполирования заключается в следующем. Для некоторой функции  $y = f(x)$  на отрезке  $[a, b]$  заданы  $n + 1$  значений  $y_0, y_1, \dots, y_n$  в точках  $x_0, x_1, \dots, x_n$ , которые называются узлами интерполяции. При этом:

$$a \leq x_0 < x_1 < \dots < x_n \leq b$$

Требуется построить интерполирующую функцию  $F(X)$ , принимающую в узлах интерполяции те же значения, что и  $f(x)$ :

$$F(x_0) = f(x_0), F(x_1) = f(x_1), \dots, F(x_n) = f(x_n).$$

Геометрически это означает нахождение кривой определённого типа, проходящей через заданную систему точек  $\{x_i, y_i\}, i = 0, 1, \dots, n$ .

Если значения аргумента выходят за область  $[a, b]$ , то говорят об экстраполировании – продолжении функции за область её определения. Происхождение этого термина также связано с латинскими словами: *extra* – вне, *pole* – узел.

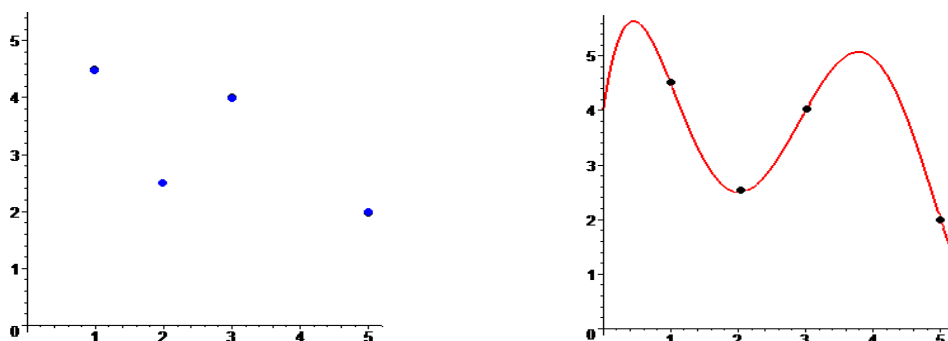


Рис.1

### Интерполяция степенным многочленом (полиномом)

**Известно, что через  $(n + 1)$  точку на плоскости можно провести кривую, являющуюся графиком степенного многочлена (полинома) степени  $n$ , причем такой полином единственный.**

Например, через две точки на плоскости можно провести только одну единственную прямую, т.е. полином первой степени, через три точки – параболу – полином второй степени и т.д. Этот факт и лежит в основе полиномиальной интерполяции, при которой функцию  $F(X)$  строят в виде канонического полинома степени  $n$  :

$$P_n(x) = a_0 + a_1x + a_2x^2 \dots + a_nx^n. \quad (1)$$

Степень интерполяционного полинома, проходящего через  $(n + 1)$  узлов интерполяции, равна  $n$ .

### Виды локальной интерполяции.

И если на всем интервале интерполяции  $[a, b]$ , содержащем  $(n + 1)$  узлов, строят один полином степени  $n$ , то говорят о *глобальной интерполяции*.

Если интервал интерполяции  $[x_0, x_n]$ , разбивают на меньшие отрезки, содержащие два или  $m, m < n$  узлов, и на каждом из отрезков строят свой

(локальный) интерполяционный полином соответствующей степени, то говорят о *локальной* или *многоинтервальной интерполяции*.

Существует несколько представлений алгебраических интерполяционных полиномов.

### Кусочно-линейная интерполяция.

Линейная интерполяция состоит в том, что заданные точки таблицы  $\{x_i, y_i\}, i = 0, 1, \dots, n$  соединяются прямыми линиями и исходная функция  $y = f(x)$  на интервале  $[a, b] = [x_0, x_n]$  заменяется ломаной линией с конечным числом прямолинейных отрезков и вершинами в узлах интерполяции  $\{x_i, y_i\} \in [a, b]$ .

Такая интерполяция называется **кусочной**. Примером может служить кусочно-линейная интерполяция по двум соседним точкам на интервале  $[x_i, x_{i+1}]$ .

Для каждого участка ломаной линии определяются эквивалентные линейные параметры наклона прямой. Области изменения входных параметров задаются таким образом, чтобы в их пределах выходная характеристика изменялась плавно. Получившаяся линейная задача для каждого отрезка решается в отдельности.

В общем случае частичные интервалы  $\{x_i, y_i\}$  различны. Для каждого отрезка ломаной можно написать уравнение прямой, проходящей через точки:

$$\frac{y - y_{i-1}}{y_i - y_{i-1}} = \frac{x - x_{i-1}}{x_i - x_{i-1}}$$

Тогда рабочую формулу можно записать как:

$$y = a_i x + b_i, \tag{2}$$

где  $a_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}, b_i = y_{i-1} - a_i x_{i-1}$ .

Из рис. 1 видно, что для реализации алгоритма (2) сначала нужно определить интервал, в который попадает значение  $x_T$ , а затем воспользоваться его границами. Например, так:

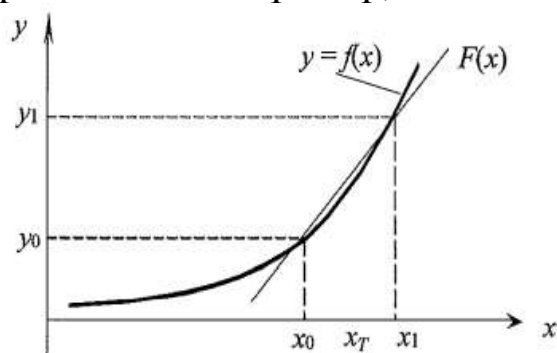


Рис.1

Ошибка приближения функции  $f(x)$  интерполяционным полиномом  $n$  – ой степени  $P_n(x)$  в точке  $x$  определяется разностью:

$$R_n(x) = f(x) - P_n(x). \quad (3)$$

Теоретическая погрешность приближения функции  $f(x)$  кусочно-линейной интерполяцией  $F(x)$  в точках  $x$ , отличных от узлов интерполяции определяется разностью:

$$R_1(x) = f(x) - P_1(x)$$

и равна

$$R_1(x) = \frac{M_2}{8} h^2, M_2 = \max |f''(x)|, h = (x_i - x_{i-1}). \quad (4)$$

### Квадратичная (кусочно-параболическая) интерполяция

При необходимости можно воспользоваться **кусочно-параболической** аппроксимацией по трем соседним точкам  $[x_{i-1}, x_i, x_{i+1}]$  и в качестве интерполяционного многочлена использовать квадратный трехчлен на отрезке  $[x_{i-1}, x_{i+1}]$ .

$$y = a_i x^2 + b_i x + c_i,$$

Для определения коэффициентов  $a_i, b_i, c_i$ , составляется и решается система из трех уравнений согласно условиям совпадения в узлах, а именно:

$$\begin{cases} a_i x_{i-1}^2 + b_i x_{i-1} + c_i = y_{i-1} \\ a_i x_i^2 + b_i x_i + c_i = y_i \\ a_i x_{i+1}^2 + b_i x_{i+1} + c_i = y_{i+1} \end{cases}. \quad (5)$$

Чтобы определить значение функции при определенном  $x$  при квадратичной интерполяции нужно провести следующие вычисления:

- определить три ближайших к  $x$  узла интерполяции;
- составить и решить систему уравнений (5);
- по полученным коэффициентам вычислить значение функции в точке  $x$ .

Следует отметить, что результаты могут отличаться в зависимости от того,  $x \in (x_{i-1}, x_i)$  или же  $x \in (x_i, x_{i+1})$ . Для примера:

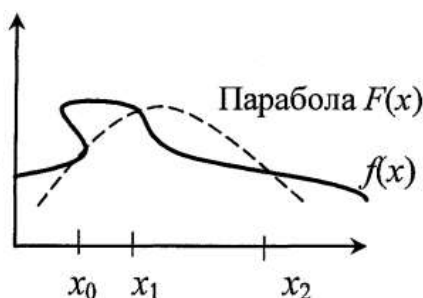


Рис.2

Теоретическая погрешность вне узлов интерполяции для квадратичной интерполяции равна

$$R_2(x) = (x - x_0)(x - x_1)(x - x_2) \frac{M_3}{6}, M_3 = \max |f'''(x)|. \quad (6)$$

В подобных ситуации численные методы требуют определения производной, однако производные в вышеприведенных методах будут терпеть разрыв в узлах кусочной интерполяции.

### Виды глобальной интерполяции

В данном случае интерполяционный многочлен вида  $P_n(x)$  ищется для всего интервала области определения  $[a, b] = [x_0, x_n]$ .

Для получения коэффициентов составляется система уравнений

$$\begin{cases} a_0 + a_1x_0 + a_2x_0^2 \dots + a_nx_0^n = y_0 \\ a_0 + a_1x_1 + a_2x_1^2 \dots + a_nx_1^n = y_1 \\ \dots \\ a_0 + a_1x_n + a_2x_n^2 \dots + a_nx_n^n = y_n \end{cases} \quad (7)$$

**Система таких линейных алгебраических уравнений относительно неизвестных коэффициентов  $a_i$  будет иметь решение, если определитель системы отличен от нуля.**

Определитель этой матрицы  $A$ , известный в алгебре как *определитель Вандермонда*, имеет аналитическое выражение следующего вида:

$$\det A = \prod_{i,j=0, i \neq j}^n (x_i - x_j). \quad (8)$$

Из этого выражения видно, что условие существования и единственности решения, выполнено, если среди всех  $(n + 1)$  узлов  $x_i$  нет совпадающих.

Таким образом, если  $x_i \neq x_j$  при  $i \neq j$ , то система имеет единственное решение. Для решения (7) можно использовать методы решения для систем линейных алгебраических уравнений.

Прямое решение системы (7) и получение  $F(x)$  явно в виде полинома (1)  $P_n(x) = a_0 + a_1x + a_2x^2 \dots + a_nx^n$  выгодно, когда производится много вычислений по одной и той же таблице.

Для разового вычисления предложены другие алгоритмы, при которых не нужно находить параметры вектора  $\vec{a}$ , а интерполяционные многочлены сразу записываются через значения таблиц  $\{x_i, y_i\}$ ,  $i = 0, 1, \dots, n$ . Это интерполяционные многочлены Лагранжа и Ньютона.



## Построение полинома Лагранжа.

Один из методов интерполирования функции, которая принимает значения  $y_0, y_1, \dots, y_n$  в точках  $x_0, x_1, \dots, x_n$ , - это **построение полинома Лагранжа**.

Многочлен Лагранжа ищется в виде линейной комбинации из значений  $y = f(x)$  в  $i$ -ых узлах интерполяции и специально построенных из системы узлов интерполяции многочленов  $L_n(x)$   $n$ -ой степени. Сам полином имеет следующий вид:

$$L_n(x) = y_0 \frac{(x-x_1)(x-x_2)\dots(x-x_n)}{(x_0-x_1)(x_0-x_2)\dots(x_0-x_n)} + y_1 \frac{(x-x_0)(x-x_2)\dots(x-x_n)}{(x_1-x_0)(x_1-x_2)\dots(x_1-x_n)} + \dots + y_n \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\dots(x_n-x_{n-1})}. \quad (9)$$

Интерполяционный полином Лагранжа является «глобальным», т.к. использует все узлы интерполяции, и при этом для него не требуется предварительного определения в (1) всех коэффициентов. Однако, из его вида (9) следует, что добавление новой узловой точки приводит к изменению всех членов полинома.

В этом состоит неудобство формулы Лагранжа. Зато метод Лагранжа содержит минимальное количество арифметических действий.

На практике, если необходим повторный расчет при различных  $x_i$  в большем количестве, то схема (1) будет предпочтительнее. Однако полином Лагранжа широко используется при реализации других численных методов. Следует подчеркнуть, что при  $n = 1$  - это линейная, а при  $n = 2$  - квадратичная интерполяция.

При интерполяции функции  $f(x)$  по ее  $n$  значениям  $y_i = f(x)$  в узлах  $a \leq x_0 < x_1 < \dots < x_n \leq b$  полиномы Лагранжа высокого порядка применяются сравнительно редко. Гораздо чаще функция  $f(x)$  интерполируется в интервалах между несколькими соседними опорными точками полиномами низких степеней.

## Погрешность глобальной интерполяции.

Ошибка приближения функции  $F(x)$  интерполяционным полиномом  $n$ -ой степени  $P_n(x)$  в точке  $x$  определяется разностью:

$$R_n(x) = F(x) - P_n(x).$$

И для полинома Лагранжа в том числе:  $R_n(x) = F(x) - L_n(x)$ .

Можно показать, что погрешность вычисления  $R_n(x)$  определяется следующим выражением:

$$R_n(x) = \frac{F^{(n+1)}(\theta)}{(n+1)!} (x - x_0)(x - x_1) \dots (x - x_n). \quad (10)$$

Здесь  $F^{(n+1)}(\theta)$  – производная  $(n + 1)$  порядка функции  $F(X)$  в некоторой точке  $\theta \in [x_0, x_n]$ . Конкретная величина погрешности в точке  $x$  зависит, очевидно, от значения функции в этой точке. Однако общим является то, что погрешность интерполяции тем выше, чем ближе точка  $x$  лежит к концам отрезка  $[a, b]$ . За пределами отрезка интерполяции (т.е. при *экстраполяции*) погрешность возрастает существенно.

Вследствие этого глобальная интерполяция в некоторых случаях может давать совершенно неудовлетворительный результат. И, несмотря на выполнение условий Лагранжа в узлах, интерполяционная функция может иметь значительное отклонение от аппроксимируемой кривой и ее поведением между узлами.

С увеличением количества узлов возрастает и степень интерполяционного полинома, что может приводить к увеличению погрешности в результате возникновения так называемого явления волнистости.

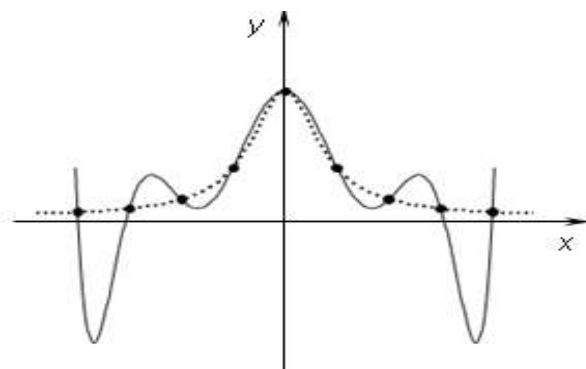


Рис.3

В качестве примера на Рис.3 приведены результаты интерполяции функции  $y(x) = 1/(1 + 25x^2)$  полиномом 8-й степени.

Пунктирной линией показан график исходной функции, сплошная линия показывает график интерполяционного полинома, построенного по заданным точкам. Факт самого наличия погрешности и ее поведение необходимо всегда учитывать в описании при постановке задачи моделирования и разработки самой модели.

### Интерполяционные полиномы Ньютона

Другой формой представления интерполяционного полинома являются **формулы Ньютона**.

$$P_n(x) = \sum_{k=0}^n \alpha_k N(x, k), \quad N(x, k) = \prod_{i=0}^{k-1} (x - x_i). \quad (11)$$

Для получения рабочей формулы Ньютона необходимо определить значения коэффициентов  $\alpha_k$ . При этом для построения интерполяционного многочлена Ньютона вводится рабочий аппарат в виде, так называемых, конечных разностей для системы равноотстоящих интерполяционных узлов и в виде разностных отношений (разделенные разности) для произвольной системы узлов.

Пусть заданы равноотстоящие узлы интерполяции

$$x_i = x_0 + i\Delta x, \Delta x = x_{i+1} - x_i = \text{const} > 0,$$

где  $i = 0, 1, \dots, n$ ;  $\Delta x$  - шаг интерполяции.

Значения  $f(x)$  в них обозначим  $f(x_i) = f_i = y_i$ .

$\Delta^i y_0$  называются (конечными) разностями  $i$ -го порядка. Они определяются так:

$$\Delta y_0 = y_1 - y_0,$$

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 + y_0,$$

$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = y_3 - 3y_2 + 3y_1 - y_0,$$

$u = (x - x_0)/\Delta x$  - нормированный аргумент.

На основе аппарата конечных разностей для определения коэффициентов для искомого многочлена (11) получена формула:

$$\alpha_k = \frac{\Delta^k y_0}{k!(\Delta x)^k}. \quad (12)$$

Используя (12), формулу для полинома Ньютона можно переписать в следующем виде:

**1-я интерполяционная формула Ньютона**, которая используется для интерполирования «вперёд», имеет вид:

$$P_n(x) = y_0 + \frac{u}{1!} \Delta y_0 + \frac{u(u+1)}{2!} \Delta^2 y_0 + \frac{u(u+1)(u+2)}{3!} \Delta^3 y_0 + \dots \quad (13)$$

При  $\Delta x \rightarrow 0$  интерполяционная формула Ньютона превращается в ряд Тейлора.

Выражение (13) может аппроксимировать  $y = f(x)$ . на всем отрезке  $[x_0, x_n]$ . Однако, с точки зрения повышения точности расчетов, скорости вычисления, т.е. уменьшения числа членов ряда в (13), рекомендуется использовать формулу для интервала  $[x_0, x_1]$ . Для других значений аргумента, например, для  $x \in [x_1, x_2]$ , вместо  $x_0$  лучше взять значение  $x_1$ .

1-я интерполяционная формула Ньютона используется для вычисления значений функций в точках левой половины рассматриваемого отрезка.

**2-я интерполяционная формула Ньютона** используется для интерполирования «назад»:

$$P_n(x) = y_n + \frac{u}{1!} \Delta y_{n-1} + \frac{u(u+1)}{2!} \Delta^2 y_{n-2} + \frac{u(u+1)(u+2)}{3!} \Delta^3 y_{n-3} + \dots$$

$$P_n(x) = y_n + \frac{u}{1!} \nabla y_0 + \frac{u(u+1)}{2!} \nabla^2 y_0 + \frac{u(u+1)(u+2)}{3!} \nabla^3 y_0 + \dots \quad (14)$$

$$u = (x - x_n) / \Delta x$$

В последней записи вместо разностей  $\Delta^i y_0$  (называемых разностями «вперёд») употребляются разности «назад»  $\nabla^i y_0$ . Существует формула записи и для случая неравно отстоящих узлов, когда рассматриваются так называемые разделённые разности.

При заданном значении  $n$  полиномы Лагранжа и Ньютона соответствуют одинаковому многочлену, представляя просто его различные формы записи.

Однако, в отличие от формулы Лагранжа в силу того, что любой  $k$ -й член полинома Ньютона зависит только от  $k$  первых узлов интерполяции и от значений функции в этих узлах, то добавление новых узлов  $(x_{k+1}, y_{k+1})$  вызывает лишь добавление в формуле новых членов без изменения первоначальных.

В этом состоит существенное, с точки зрения организации вычислений, преимущество полинома Ньютона по сравнению с полиномом Лагранжа.

Это позволяет оценить точность интерполирования. Хотя формулы Ньютона требуют большее количество арифметических действий, чем формулы Лагранжа.

**При  $n = 1$  получаем формулу линейного интерполирования:**

$$P_1(x) = y_0 + \frac{u}{1!} (x - x_0) .$$

**При  $n = 2$  будем иметь формулу параболического интерполирования:**

$$P_2(x) = y_0 + \frac{u}{1!} (x - x_0) + \frac{u(u+1)}{2!} (x - x_0)(x - x_1) .$$

Интерполяция полиномами целесообразна, когда на интервале анализа наблюдается непериодическая функция. В случае периодических зависимостей целесообразно использовать интерполяцию рядом Фурье.

Следует также отметить, что при  $n \gg 1$  (часто  $n > 10$ ) в многочленах Лагранжа и Ньютона может возникать неустойчивость, и поэтому следует проводить дополнительные исследования.

### **Интерполяционные сплайны.**

В алгебраическом интерполировании при увеличении числа узлов увеличивается, как правило, степень интерполяционного полинома. Кроме того, когда интерполирование выполняется для функций, не являющихся достаточно гладкими, интерполирование высокого порядка нецелесообразно.

В таких случаях наиболее точное приближение функции дает интерполяция сплайном. Принципиальное отличие идеи сплайн-интерполяции от интерполяции полиномом состоит в том, что полином один, а сплайн состоит из нескольких полиномов, а именно их количество равно количеству интервалов, внутри которых мы производим интерполяцию.

Слово сплайн (английское слово "spline") означает гибкую линейку, используемую для проведения гладких кривых через заданные точки на плоскости. Изначально подобные линейки использовались в краблестроении при проектировании формы корабля. В современном мире сплайны широко используются в инженерных приложениях, в частности, в компьютерной графике.

#### **Характеристики сплайна:**

**Степенью сплайна** называется максимальная из степеней использованных полиномов.

**Гладкостью сплайна** называется количество непрерывных производных, которые имеются на всем отрезке  $[x_0, x_n]$ .

**Дефектом сплайна** называется разность между степенью и гладкостью сплайна.

Например, непрерывная ломаная линия есть кусочно-линейный сплайн, который имеет степень 1, гладкость 0 и дефект сплайна = 1. Гладкий кусочно-кубический сплайн имеет степень 3, гладкость 2 и дефект = 1.

Другими словами сплайн — это кусочно заданная функция, то есть совокупность нескольких функций, каждая из которых задана на каком-то множестве значений аргумента, причём эти множества попарно непересекающиеся.

Аппроксимирующие полиномы сплайна могут быть второй степени, и тогда мы получим квадратичный сплайн или других степеней.

Но наиболее распространены **кубические сплайны**, у которых в узлах достигается непрерывность первой и второй производных. Кубический сплайн является математической моделью гибкого тонкого стержня, закрепленного в двух точках на концах с заданными углами наклона  $\alpha$  и  $\beta$  (рис. 4).

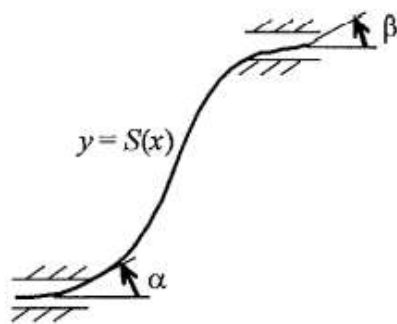


Рис.4

В данной физической модели стержень принимает форму, минимизирующую его потенциальную энергию, и описывается уравнением свободного равновесия. А этому состоянию соответствует многочлен третьей степени между двумя соседними узлами интерполяции.

### Построение кубических сплайнов.

Форма этого сплайна как «универсального лекала» на каждом отрезке описывается кубической параболой. Кубический сплайн строится следующим образом. Пусть заданы  $n$  узлов интерполяции  $(x_0, y_0), (x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)$ , функции  $f(x)$  (в общем случае неэквидистантные). На каждом интервале  $[x_i, x_{i+1}]$  кубический сплайн является алгебраическим полиномом третьей степени:

$$P_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, x_i \leq x \leq x_{i+1}, \quad (15)$$

Причем эти многочлены подобраны так, что при всех  $x_i \leq x \leq x_{i+1}$  непрерывны и сами сплайны и их первая и вторая производные.

Наша цель получить  $(n - 1)$  таких многочленов:

$$P_1(x), \dots, P_i(x), \dots, P_{n-1}(x),$$

т.е. вычислить все неизвестные  $4(n - 1)$  коэффициенты  $a_i, b_i, c_i, d_i$ .

Эти неизвестные коэффициенты определяются из следующих требований:

1. Сплайн-полиномы должны проходить через узловые точки:

$$P_i(x_i) = y_i,$$

$$P_{i+1}(x_{i+1}) = y_{i+1}, \quad i = 1, 2, \dots, (n - 1).$$

2. Первые производные должны быть непрерывны во внутренних узлах:

$$P'_i(x_{i+1}) = P'_{i+1}(x_{i+1}), \quad i = 1, 2, \dots, (n - 2),$$

3. Вторые производные должны быть непрерывны во внутренних узлах:

$$P''_i(x_{i+1}) = P''_{i+1}(x_{i+1}), \quad i = 1, 2, \dots, (n - 2),$$

В 3-х пунктах сформулировано  $(4n - 6)$  ограничений, а число неизвестных составляет  $4(n - 1)$ . Поэтому, для однозначного определения коэффициентов  $a_i, b_i, c_i, d_i$ . необходимы еще два дополнительных условия.

Недостающие два условия для однозначного определения сплайна могут быть заданы краевыми условиями, например, приравниванием нулю вторых производных в точках  $x_0$  и  $x_n$ .

$$P''_1(x_0) = P''_{n-1}(x_n) = 0.$$

Фактически это означает, что недостающие два соотношения получаются из условий закрепления концов сплайна. В частности, при свободном закреплении концов, когда можно приравнять нулю кривизну линии в этих точках.

Эти условия иногда называют условиями “свободного провисания”, показанными на рис.4. Они определяют так называемый **естественный (свободный) кубический сплайн (или чертежный сплайн)**.

Однако краевые условия могут быть весьма разнообразными (разных типов). Можно задавать в граничных точках конкретные наклоны графика сплайна или конкретную кривизну. Иногда при построении сплайнов задаются и другие граничные условия на концах интервала интерполяции для самих функций или первых производных или, например для периодического сплайна возможно равенство значений на концах интервалов. Но, все же чаще всего используются естественные граничные условия:

$$P''_1(x_0) = 0, \quad P''_{n-1}(x_n) = 0,$$

т.е. через крайние точки сплайны должны проходить с нулевой кривизной.

Можно записать полиномы в форме, которая значительно упростит все операции с ними:

$$P_i(x) = \bar{z}y_1 + zy_{i+1} + h_i[(l_i - t_i)z\bar{z}^2 - (l_{i+1} - t_i)z^2\bar{z}],$$

где  $h_i = x_{i+1} - x_i$  - расстояние между соседними узлами,

$z = (x - x_i)/h_i$ ,  $\bar{z} = (1 - z) = (x_{i+1} - x)/h_i$  - относительные расстояния от левого  $i$ -ого и правого  $(i + 1)$  ого узла,

$t_i = (y_{i+1} - y)/h_i$  - тангенс угла наклона прямой,

$l_i$  - неизвестные константы, подлежащие определению.

При таком представлении сплайна получим, что при  $x = x_i \rightarrow z = 0, \bar{z} = 1$ , а при  $x = x_{i+1} \rightarrow z = 1, \bar{z} = 0$ , а для производных -  $dz/dx = 1/h$  и  $d\bar{z}/dx = -1/h$ . Тогда вычисление последнего полинома в точках  $x_i$ ,  $x_{i+1}$  показывает, что остальные  $2(n - 1)$  ограничений удовлетворяются автоматически.

Первая производная такого измененного полинома будет равна

$$P'_i(x) = t_i + [(l_i - t_i)\bar{z}(\bar{z} - 2z) - (l_{i+1} - t_i)z(2\bar{z} - z)],$$

Откуда видно, что необходимое требование при  $x = x_i, x = x_{i+1}$  также выполняется.

Наконец, вычисляя вторую производную измененного полинома, получим

$$P''_i(x) = \frac{2}{h_i} [(l_i - t_i)(z - 2\bar{z}) - (l_{i+1} - t_i)z(\bar{z} - 2z)].$$

Тогда во внутренних узлах слева и справа получаем

$$P''_i(x_{i+1}) = \frac{2}{h_i} [l_i + 2l_{i+1} - 3t_i], \quad P''_{i+1}(x_{i+1}) = \frac{2}{h_{i+1}} [-2l_{i+1} - l_{i+2} - 3t_{i+1}]$$

Приравнявая эти производные друг к другу в соответствии с требованием, приходим к следующей системе уравнений:

$$g_i l_i + 2(g_{i+1} + g_i)l_{i+1} + g_{i+1}l_{i+2} = 3(g_i t_i + g_{i+1} t_{i+1}),$$

где  $g_i = 1/h_i$

Из оставшегося условия получаем еще два уравнения

$$2g_1 l_1 + g_1 l_2 = 3g_1 t_1 \quad g_{n-1} l_{n-1} + 2g_{n-1} l_n = 3g_{n-1} t_{n-1},$$

А вместе взятые они могут быть записаны в форме симметричной трехдиагональной системы уравнений:



$$\begin{bmatrix} 2g_1 & g_1 & & & & & \\ g_1 & 2(g_1 + g_2) & g_2 & & & & \\ & g_2 & 2(g_2 + g_3) & g_3 & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \\ & & & g_{n-1} & 2g_{n-1} & & \end{bmatrix} \times \begin{bmatrix} l_1 \\ l_2 \\ l_3 \\ \dots \\ l_n \end{bmatrix} = 3 \begin{bmatrix} g_1 t_1 \\ g_1 t_1 + g_2 t_2 \\ g_2 t_2 + g_3 t_3 \\ \dots \\ g_{n-1} t_{n-1} \end{bmatrix}$$

Эта система уравнений обычно хорошо обусловлена и может быть легко решена относительно констант  $l_1, \dots, l_n$ . После нахождения коэффициентов  $l_i$  можно определить коэффициенты  $a_{0i}, a_{1i}, a_{2i}, a_{3i}$  по формулам:

$$\begin{aligned} a_i &= y_i, \\ b_i &= l_i, \\ c_i &= g_i(3t_i - 2l_i - l_{i+1}), \\ d_i &= g_i^2(-2t_i + l_i + l_{i+1}), \quad i = 1, 2, \dots, (n-1). \end{aligned} \tag{16}$$

Для решения системы уравнений можно воспользоваться и методом исключения Гаусса. Однако его целесообразно адаптировать под специальный вид трехдиагональной системы уравнений. Нужно учитывать, что в этой системе уравнений заполняются только главная диагональ и две дополнительные диагонали (наддиагональ и поддиагональ). Поэтому коэффициенты системы уравнений целесообразно запоминать не в двумерном массиве, а в двух одномерных: (наддиагональ и поддиагональ). Решение системы можно получать на месте столбца свободных членов.

Так как  $g_i > 0$ , то наибольшие по модулю коэффициенты уже стоят на главной диагонали и перестановка уравнений, которая обычно производится в алгоритме Гаусса, здесь не нужна. Следует учитывать также, что при прямой подстановке на каждом шаге исключаются только коэффициенты, стоящие над главной диагональю, а остальные коэффициенты равны нулю изначально.. В ходе прямой подстановки изменяться будут только диагональные и свободные члены, при обратной подстановке из последнего уравнения находится  $l_n$ , из предпоследнего -  $l_{n-1}$  и т.д.

Значения сплайна и его производные в процедуре можно вычислять по известной схеме Горнера:

$$\begin{aligned} P_i(x) &= [(d_i \bar{x} + c_i) \bar{x} + b_i] \bar{x} + a_i, \\ P'_i(x) &= (3d_i \bar{x} + 2c_i) \bar{x} + b_i, \quad \bar{x} = x - x_i. \end{aligned}$$

Метод сплайн-интерполяции приводит к удовлетворительным результатам в процессе интерполяции непрерывных функций с гладкими первой и второй производными. При этом кубический интерполяционный сплайн, построенный по узлам  $y_i = f(x_i)$  будет обладать минимальной кривизной по сравнению с кривизной любой интерполяционной функции, имеющей непрерывные первую и вторую производные.

Сплайн-интерполяция функций с резким изменением производных сплайна может давать большие ошибки. Сплайны более высокого порядка, чем третий, используются редко, так как при вычислении большого числа коэффициентов может накапливаться ошибка, приводящая к значительным отклонениям.

Для квадратичного и кубического сплайна имеет место равномерная сходимость. Причем скорость сходимости повышается вместе с порядком сплайна и гладкостью  $f(x)$ .

### **Типы граничных условий.**

Как уже было сказано выше, недостающие два условия для однозначного определения сплайна могут быть заданы приравнением нулю вторых производных в точках  $x_0$  и  $x_n$

$$P_1''(x_0) = 0, P_{n-1}''(x_n) = 0.$$

Фактически это означает, что недостающие два соотношения получаются из условий закрепления концов сплайна. В частности, при свободном закреплении концов можно приравнять нулю кривизну линии в этих точках.

Однако краевые условия могут быть весьма разнообразными (разных типов). Можно задавать в граничных точках конкретные наклоны графика сплайна или конкретную кривизну. Иногда при построении сплайнов задаются и другие граничные условия на концах интервала интерполяции для самих функций или первых производных, например, для периодического сплайна возможно равенство значений на концах интервалов.

#### **Краевые условия 1-ого типа:**

На концах промежутка  $[a, b]$  задаются значения первой производной искомой функции.

#### **Краевые условия 2-ого типа:**

На концах промежутка  $[a, b]$  задаются значения второй производной искомой функции.

#### **Краевые условия 3-ого типа:**

Их называют периодическими и определяют период. Выполнение подобных условий естественно требовать, когда интерполируемая функция является периодической с периодом  $T = b - a$ . Условия могут накладываться и на первую и на вторую производные.

#### **Краевые условия 4-ого типа:**

Во внутренних узлах сетки третья производная может быть разрывна. Число разрывов третьей производной можно уменьшить или исключить, задав условия четвертого типа. В этом случае построенный сплайн будет трижды непрерывно дифференцируем на внутренних промежутках.

### **Многомерные сплайны.**

В рамках алгебраического подхода многомерные сплайны или сплайн-поверхности чаще всего строятся из одномерных посредством **тензорного произведения или смешивания**. Однако, каким бы удобным ни оказывался иногда этот подход, он страдает двумя присущими ему недостатками:

1. ограничение областями и сетками узлов прямоугольного вида,
2. использование в конечном счете довольно неестественных функциональных пространств и норм.

Для произвольно размещенных узлов в многомерной задаче восстановления функции более плодотворным оказывается первоначальный вариационный подход, описываемый общими условиями вариационного исчисления.