



Анализ данных в Mathcad

Математические
вычисления

Аппроксимация (приближение)

- Математический метод, состоящий в замене одних математических объектов другими, близкими к исходным, но более простыми.
- Аппроксимация позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов (например, таких, характеристики которых легко вычисляются, или свойства которых уже известны).



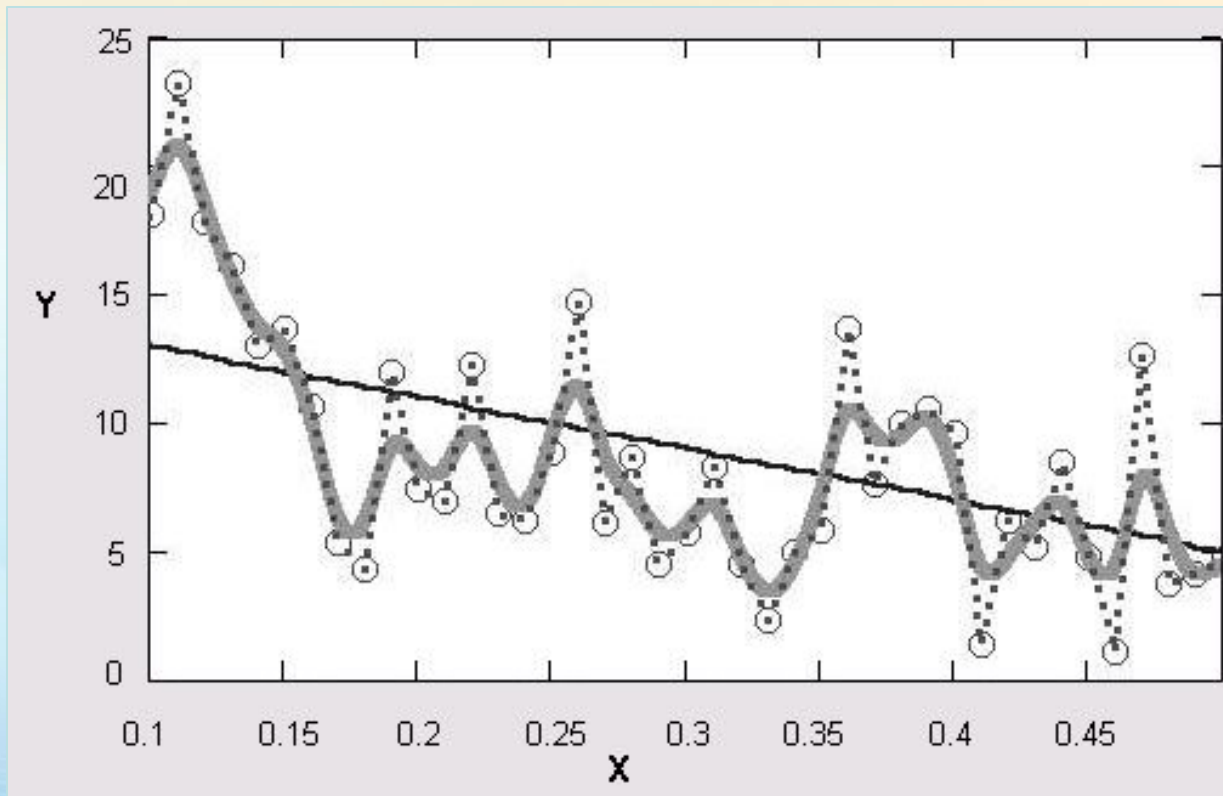
Аппроксимация функций заключается в приближенной замене заданной функции $f(x)$ некоторой функцией $\varphi(x)$ так, чтобы отклонение функции $\varphi(x)$ от $f(x)$ в заданной области было наименьшим.

- Когда имеется выборка экспериментальных данных, то она, чаще всего, представляются в виде массива, состоящего из пар чисел (x_i, y_i) .
- Поэтому возникает задача аппроксимации дискретной зависимости $y(x_i)$ непрерывной функцией $f(x)$. Функция $f(x)$, в зависимости от специфики задачи, может отвечать различным требованиям.

Требования

1. Функция $f(x)$ должна проходить через точки (x_i, y_i) , т. е. $f(x_i) = y_i$, $i = 1 \dots n$. В этом случае говорят об интерполяции данных функцией $f(x)$ во внутренних точках между x_i , или экстраполяции за пределами интервала, содержащего все x_i .
2. Функция $f(x)$ должна некоторым образом приближать $y(x_i)$, не обязательно проходя через точки (x_i, y_i) . Такова постановка задачи регрессии.
3. Функция $f(x)$ должна приближать экспериментальную зависимость $y(x_i)$, учитывая, к тому же, что данные (x_i, y_i) получены с некоторой погрешностью. При этом функция $f(x)$, с помощью того или иного алгоритма уменьшает погрешность, присутствующую в данных (x_i, y_i) . Такого типа задачи называют задачами фильтрации. Сглаживание - частный случай фильтрации.

Различные виды построения аппроксимирующей зависимости $f(x)$



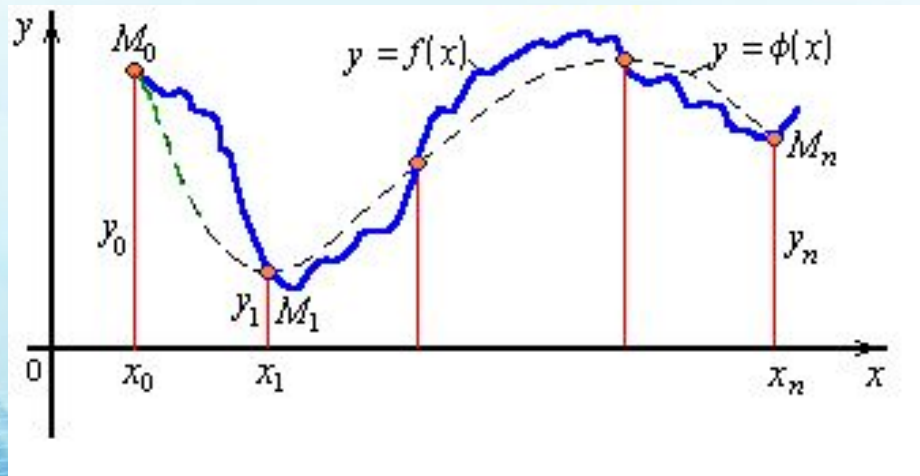
исходные данные, интерполяция, линейная регрессия, сглаживание.

Интерполяция

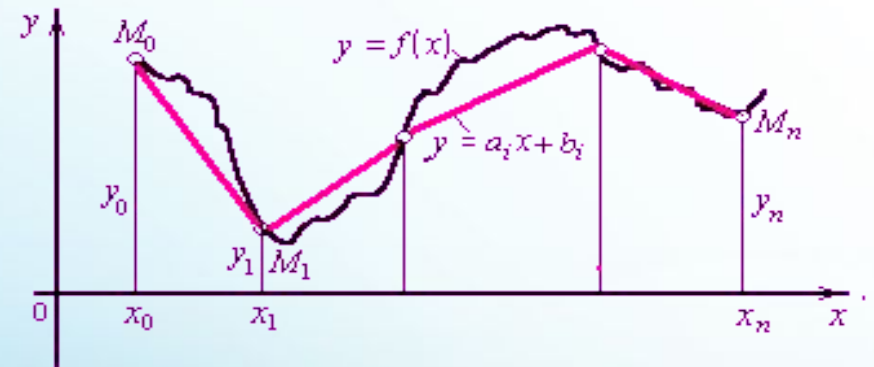
- Задача интерполяции функции одной переменной состоит в замене дискретной зависимости $y(x_i)$, т. е. N пар чисел (x_i, y_i) , или, по-другому, узлов, некоторой непрерывной функцией $y(x)$.
- При этом основным условием является то, что функция $y(x)$ должна проходить через точки (x_i, y_i) , т. е. $y(x_i) = y_i, i = 1 \dots N$, а также возможность вычислить значение $y(x)$ в любой точке, находящейся между узлов.

Виды интерполяции

● Глобальная



● Локальная



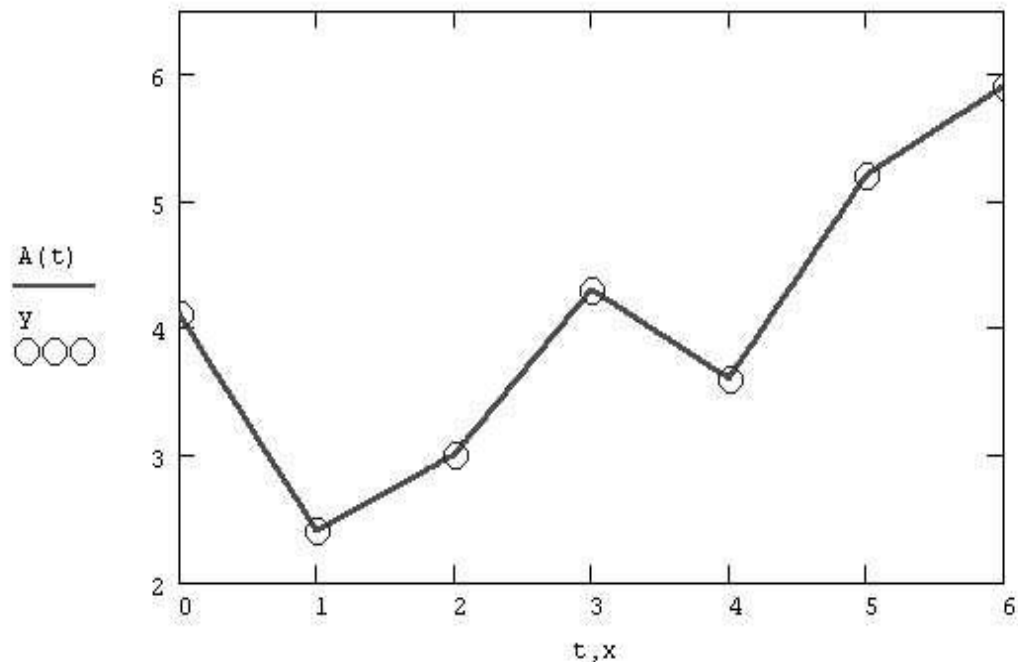
Локальная интерполяция

- При локальной интерполяции между различными узлами выбираются различные многочлены невысокой степени.
- В среде Mathcad есть для этого инструментарий: средства линейной интерполяции (функция `linterp`) и интерполяции сплайном – линейным (`lspline`), параболическим (`pspline`) и кубическим (`cspline`).



Линейная интерполяция (кусочно-линейная)

- Самый простой вид интерполяции, которая представляет искомую зависимость в виде ломаной линии. Интерполирующая функция $y(x)$ состоит из отрезков прямых, соединяющих точки (x_i, y_i)



Сплайн-интерполяция

- В большинстве практических приложений лучше соединить экспериментальные точки (x_i, y_i) не ломаной линией, а гладкой кривой.
- Лучше всего для этих целей подходит интерполяция $y(x)$ квадратичными или кубическими сплайнами, т. е. отрезками квадратичных или кубических парабол.
- Сплайн-интерполяция обеспечивает равенство в узлах не только самих соседних параболических интерполирующих функций (сплайнов), но и их производных. Благодаря этому сплайн-интерполяция выглядит как очень гладкая функция.

Сплайн – это математическая модель гибкого, тонкого стержня из упругого материала. Стержень закрепляется в двух соседних узлах с заданными углами наклона. Стержень длиннее, чем расстояние между двумя точками. Линия, которую описывает сплайн-функция, напоминает по форме гибкую линейку, закреплённую в узловых точках (откуда и название: spline – гибкая линейка).



Локальная интерполяция

Линейная интерполяция

`linterp(vx, vy, x)` Использует векторы данных `vx` и `vy`, чтобы вернуть линейно интерполируемое значение `y`, соответствующее третьему аргументу `x`.

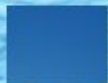
Сплайновая интерполяция (проходит в два этапа)

`lspline(vx, vy)` Все эти функции возвращают
`pspline(vx, vy)` вектор коэффициентов вторых

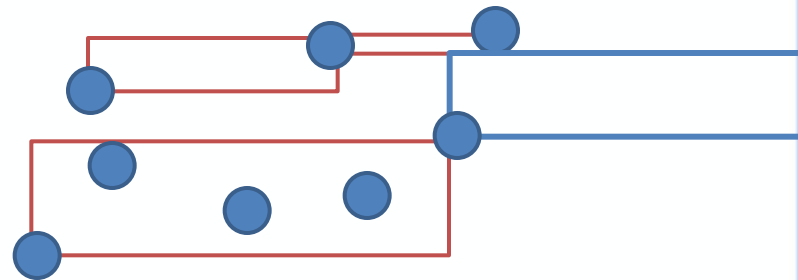
`cspline(vx, vy)` полиномов (или кубов). Возвращают

Линейное предсказание

определение значений вне набора данных



Экстраполяция



- Интерполяция дает возможность по значениям табличной функции находить значения в промежуточных точках. Однако бывают случаи, когда необходимо оценить табличную функцию за пределами ее области данных.
- Такие предсказания позволяют осуществлять отдельные численные методы. Принцип их работы основывается на анализе поведения зависимости в нескольких ее точках.
- В Mathcad функцией, реализующей один из алгоритмов предсказания (метод линейного предсказания Берга), является встроенная функция **predict**.



Формат функции: `predict(vy,m,n)`

- **vy** – вектор табличных значений функции (элементы вектора должны быть взяты через равные интервалы);
- **m** – число последних исходных значений табличной функции, по которым выполняется прогноз;
- **n** – число предсказанных значений.



Пример

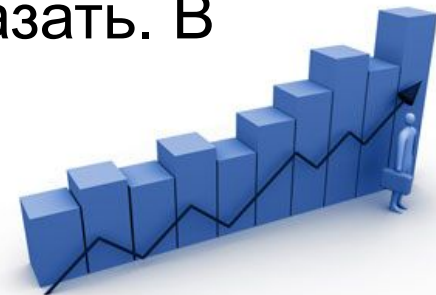
- Пусть получен вектор данных из таблицы

$Y :=$

0.86
0.77
0.7
0.63
0.58
0.52
0.47
0.39

Определение m , число известных значений Y по которым будет построена экстраполяция, в нашем случае $m = 8$ (функция `length`);

Определение n , число значений Y , по которым строится экстраполяция + количество точек которое необходимо предсказать. В нашем случае $n = 13$.



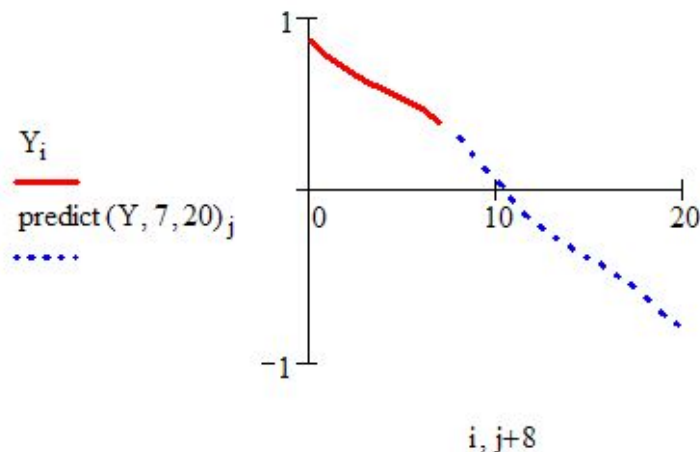
Построим график функции предсказания

- Для наглядности построим на одном графике исходную функцию (по точкам) и функцию предсказания, причем таким образом, чтобы последняя являлась продолжением первой функции.
- Для этого необходимо:
 - Так как мы имеем 8 значений исходных данных, то будем строить исходную функцию в 8 точках:
 $i:=0..7$



Построим график функции предсказания

- Так как мы собираемся строить экстраполяцию для 13-ти точек, то $j:=0..13$
- Построим график функций. Причем необходимо обратить внимание, что мы сдвинули начало координат для функции предсказания на 8 точек вправо (потому что исходная функция кончается на 8-й точке) с помощью выражения $j+8$.



Применение

- Функция предсказания обеспечивает достаточно высокую точность для аналитических зависимостей, при монотонных исходных функциях или исходных функциях, представляемых полиномом невысокой степени при достаточно большом числе исходных точек.
- Для хорошего прогноза необходимо тщательно подбирать число m , иначе качество прогноза может сильно ухудшиться.



Глобальная интерполяция

- Кубические сплайны - это мощное и удобное средство, но необходимо учитывать влияние направления и величины касательных векторов, указывать все точки кривой до ее изображения, невозможна локальная коррекция кривой. Расчет кубического сплайна требует обращения большой матрицы, зависящей от всех элементов сплайна; т. е. изменение любого сегмента затрагивает все остальные сегменты. Воздействие уменьшается при удалении от точки возмущения, но полностью пренебречь им нельзя.
- Параболическая интерполяция разрешает большинство этих проблем за счет того, что она только непрерывна, т. е. в точках соединения сегментов сохраняется непрерывность лишь первой производной, причем параболическая интерполяция не требует больших расчетов.



$P_n(x)$ степени n , значения которого в заданных точках x_0, x_1, \dots, x_n совпадают со значениями y_0, y_1, \dots, y_n функции f в этих точках. Многочлен $P_n(x)$ определяется единственным образом, но в зависимости от задачи его удобно записывать различными по виду формулами.

- Интерполяционная формула Лагранжа

• **Интерполяционные формулы**, дающие приближённое выражение функции $y = f(x)$ при помощи интерполяции, т. е. через интерполяционный многочлен $P_n(x)$ степени n , значения которого в заданных точках x_0, x_1, \dots, x_n совпадают со значениями y_0, y_1, \dots, y_n функции f в этих точках. Многочлен $P_n(x)$ определяется единственным образом, но в зависимости от задачи его удобно записывать различными по виду формулами.

- Интерполяционная формула Лагранжа

- Интерполяционная формула Ньютона

- Интерполяционная формула Стирлинга

- Интерполяционная формула Бесселя



Интерполяция функций по Лагранжу

- Приблизить таблично заданную функцию можно не привязываясь к конкретной точке (локальная интерполяция), а использовать все узловые точки (глобальная).
 - Пусть некоторая функция $f(x)$ определена рядом своих узловых точек (x_i, y_i) на некотором отрезке $[a, b]$.
Под *интерполяцией* подразумевается вычисление значений $f(x)$ в любом промежутке $[x_i, x_{i+1}]$ в пределах отрезка $[a, b]$.
Соответственно, любое вычисление $f(x)$ вне отрезка $[a, b]$ является *экстраполяцией*.
 - Значения $f(x)$ вычисляются с помощью аппроксимирующего полинома: $f_a(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_i x^i + \dots + a_2 x^2 + a_1 x + a_0 = \sum_{i=0}^n a_i x^i$
- Реализация полиномиальной аппроксимации сводится к вычислению коэффициентов полинома $a_n, a_{n-1}, \dots, a_1, a_0$ так, чтобы точки $f_a(x_i)$ точно совпадали с узловыми точками.

Решение задачи

ИНТЕРПОЛЯЦИЯ И ЭКСТРАПОЛЯЦИЯ ПО ФОРМУЛЕ ЛАГРАНЖА

$$x_i := \begin{pmatrix} 1 \\ 7 \\ 5 \\ 2 \\ 10 \end{pmatrix} \quad y_i := \begin{pmatrix} 10 \\ 42 \\ 35 \\ 4 \\ 95 \end{pmatrix}$$

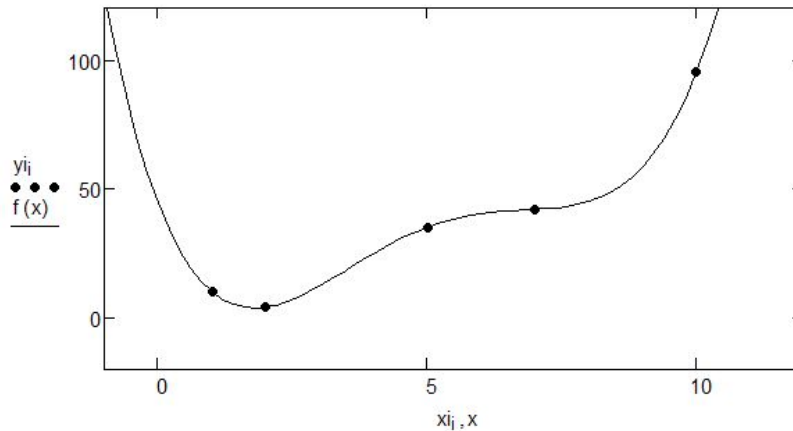
Векторы x_i и y_i задают таблицу интерполируемой функции для последующей интерполяции методом Лагранжа с применением общей формулы интерполяции

$$n := \text{length}(x_i) - 1 \quad i := 0..n \quad j := 0..n$$
$$f(x) := \sum_{i=0}^n y_i \cdot \frac{\prod_{j=0, j \neq i}^n \text{if}(i = j, 1, x - x_j)}{\prod_{j=0}^n \text{if}(i = j, 1, x_i - x_j)}$$

Общая формула интерполяции Лагранжа

Примеры интерполяции и экстраполяции: $f(-1) = 125.067$ $f(2) = 4$ $f(6) = 40.315$ $f(10.5) = 126.874$

$i := 0..n$ $x := -1, -0.8.. 12$



Построение по данным интерполяции графика функции с нанесенными на него узловыми точками - темными кружками. График проходит точно через узловые точки, которые расположены неравномерно.



Параболическая интерполяция

- Кривая полинома точно должна пройти через все узловые точки.
- Особенностью глобальной полиномиальной интерполяции (параболической интерполяции) является однозначное соответствие между числом узловых точек N аппроксимируемой функции и степенью полинома $n=N-1$.
 - На практике можно нередко задать функцию множеством точек, но тогда степень полинома станет очень большой, его вычисления займут много времени, а точность вычислений резко ухудшается. Максимальная степень полинома не должна превышать 8-10.



Пример

$$x := \begin{pmatrix} 0 \\ 0.2 \\ 0.4 \\ 0.6 \\ 0.8 \\ 1 \end{pmatrix} \quad y := \begin{pmatrix} 0 \\ 52 \\ 23 \\ 2 \\ 4 \\ 12 \end{pmatrix}$$

ПАРАБОЛИЧЕСКАЯ ИНТЕРПОЛЯЦИЯ

Векторы x и y задают узловые точки (x, y)

$n := \text{length}(x) - 1$ $n = 5$ Вычисление степени полинома

$i := 0..n$ $j := 0..n$ $D_{i,j} := (x_i)^j$ Формирование матрицы D :

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.2 & 0.04 & 0.008 & 0.0016 & 0.00032 \\ 1 & 0.4 & 0.16 & 0.064 & 0.0256 & 0.01024 \\ 1 & 0.6 & 0.36 & 0.216 & 0.1296 & 0.07776 \\ 1 & 0.8 & 0.64 & 0.512 & 0.4096 & 0.32768 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Вычисление коэффициентов a полинома из решения системы линейных уравнений (можно использовать любой способ для решения СЛАУ, например $a := D^{-1} * y$):

$$a := \text{Isolve}(D, y)$$

$P(x) := \sum_{i=0}^n (a_i \cdot x^i)$ Определение полинома $P(x)$

Определитель Вандермонда

$$a = \begin{pmatrix} 0 \\ 745.333 \\ -3.41 \times 10^3 \\ 5.698 \times 10^3 \\ -4.115 \times 10^3 \\ 1.094 \times 10^3 \end{pmatrix}$$

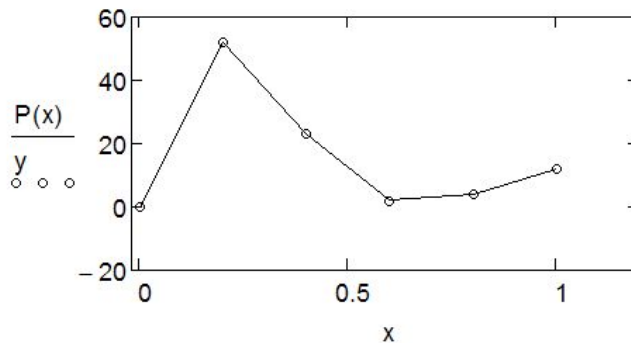


График полинома $P(x)$ и нанесенные на него узловые точки. Для этого вида интерполяции характерно, что график точно проходит через узловые точки.

