

# ЛАБОРАТОРНАЯ РАБОТА № 6

## Тема: Обработка экспериментальных данных

### 1. Кусочно-линейная аппроксимация

При линейной интерполяции аппроксимирующая функция соединяет опытные точки отрезками прямых линий. Для линейной интерполяции используется встроенная функция **linterp**.

**Обращение к функции** `linterp(x,y,t)`,

где  $x$  - вектор опытных значения аргумента,

$y$  - вектор опытных значений функции,

$t$  - значение аргумента, при котором вычисляется интерполирующее значение функции.

#### Пример №1

Провести кусочно-линейную аппроксимацию следующего экспериментального массива данных:

$$X := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \quad Y := \begin{pmatrix} 98 \\ 85 \\ 32 \\ 79 \\ 30 \end{pmatrix}$$

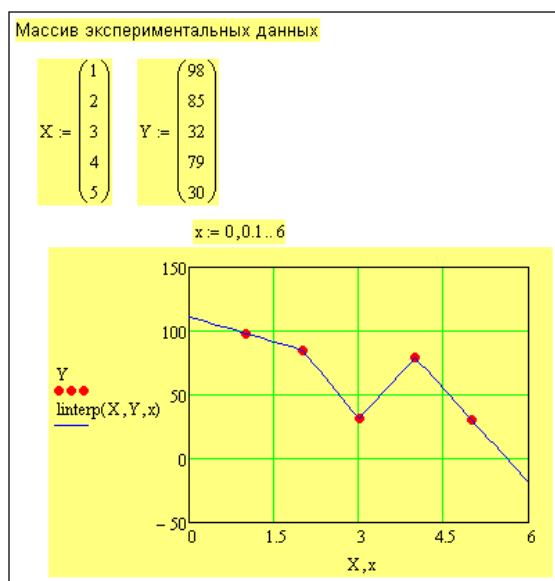


Рис. 1.

### 2. Аппроксимация сплайнами.

При небольшом числе узловых точек (менее 10) линейная **интерполяция** оказывается довольно грубой. При ней даже первая производная функции аппроксимации испытывает резкие скачки в узловых точках. Для целей **экстраполяции** функция `linterp` не предназначена и за пределами области определения может вести себя непредсказуемо.

Гораздо лучшие результаты дает сплайн-аппроксимация. При ней исходная функция заменяется отрезками квадратных или кубических полиномов, проходящих через три смежные узловых точки. Коэффициенты полиномов рассчитываются так, чтобы непрерывными были первая и вторая производные. Линия, которую описывает сплайн-функция, напоминает по форме гибкую линейку, закрепленную в узловых точках (отсюда и название аппроксимации: **splain** - гибкая линейка).

Для осуществления сплайновой аппроксимации система Маткад предлагает четыре встроенные функции. Три из них служат для получения векторов вторых производных сплайн-функций при различном виде интерполяции:

**cspline**(X, Y) - возвращает вектор S вторых производных при приближении в опорных точках к кубическому полиному;

**pspline**(X, Y) - возвращает вектор S вторых производных при приближении к опорным точкам параболической кривой;

**lspline**(X, Y) - возвращает вектор S вторых производных при приближении к опорным точкам прямой. Наконец, четвертая функция **interp**(S, X, Y, x) возвращает значение  $y(x)$  для заданных векторов S, X, Y и заданного значения x.

Таким образом, сплайн-аппроксимация проводится в два этапа. На первом с помощью функций cspline, pspline или lspline отыскивается вектор вторых производных функции  $y(x)$ , заданной векторами X и Y ее значений (абсцисс и ординат). Затем, на втором этапе для каждой искомой точки вычисляется значение  $y(x)$  с помощью функции interp.

На рис 2 приведен пример аппроксимации сплайнами. Аппроксимация проведена для тех же заданных точек, что и в предыдущем примере. Набраны вектора X, Y и с помощью функций lspline, pspline, cspline получены коэффициенты сплайнов. Чтобы убедиться, что аппроксимирующая кривая проходит через заданные точки, функция interp вычислена для заданных значений x. Поэтому результатом interp явились заданные значения y. Чтобы более точно посмотреть аппроксимирующую кривую для графика x задан в том же интервале, но с шагом 0.1. Как видим из графика, аппроксимация сплайнами совершенно не похожа на кусочно-линейную аппроксимацию.

В функции interp – X,Y- те же векторы заданных значений, а S – вектор коэффициентов уравнений для сплайнов, полученный из функций lspline, pspline или cspline.

## Пример №2

Провести кусочно-линейную аппроксимацию экспериментального массива данных из примера 1.

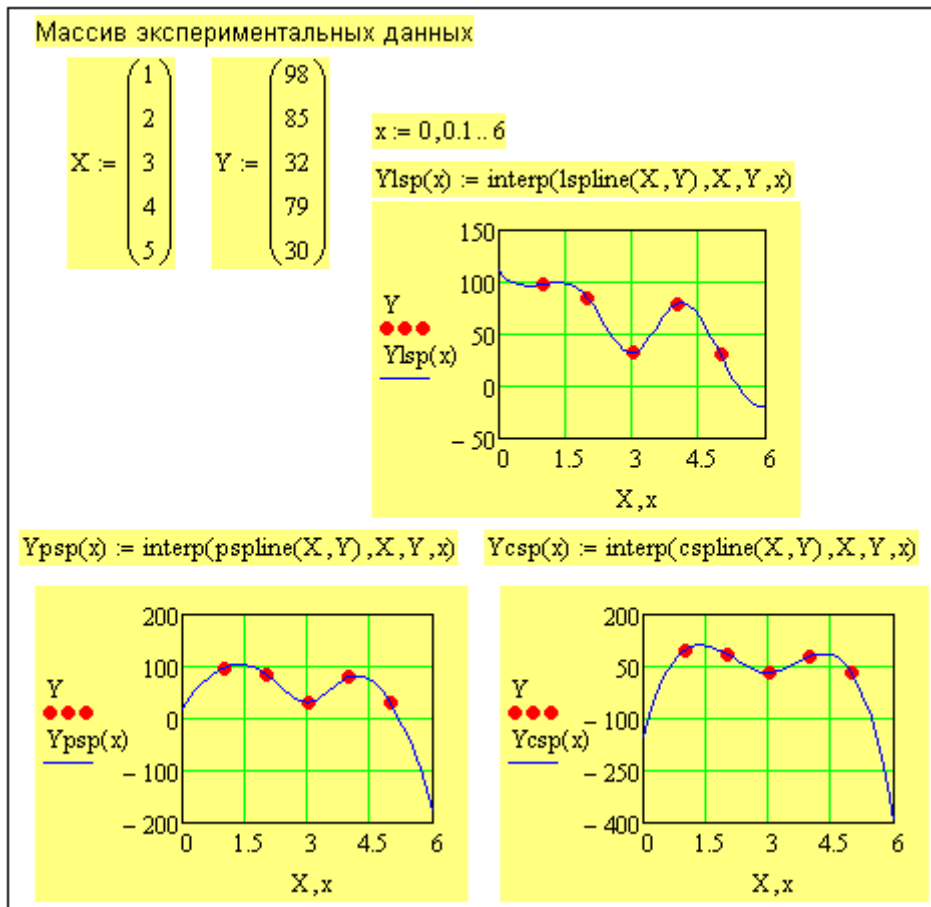


Рис. 2

### 3. Функции регрессии

Смысл регрессии состоит в подборе функции, аппроксимирующей экспериментальные данные.

В отличие от функций интерполяции эти функции не требуют, чтобы аппроксимирующая кривая или поверхность проходила через точки данных.

Рассмотрим две функции, которые не выводят коэффициентов и аппроксимируют массив данных одним степенным полиномом (**regress**) или отрезками нескольких полиномов (**loess**).

В Mathcad регрессия с использованием одного полинома реализуется комбинацией встроенных функций регрессии и интерполяции

**interp(s,x,y,t)** и **regress(x,y,n)**, где

x - вектор значений аргумента, элементы которого расположены в порядке возрастания,

y - вектор значений функции того же размера,

s - вектор коэффициентов для построения аппроксимирующего полинома, создаваемый функцией **regress**,

t - значение аргумента, при котором вычисляется интерполирующая функция,

n - степень аппроксимирующего полинома.

Степень аппроксимирующего полинома может быть любой.

Обращение к указанным функциям:

**s=regress(X,Y,n)**      **YY(t)=interp(s,X,Y,t)**

или

**YY(t)=interp(regress(X,Y,n),X,Y,t)**

Регрессия с использованием нескольких отрезков полинома реализуется комбинацией встроенных функций регрессии и интерполяции

**interp(s,X,Y,t)** и **loess(X,Y,L)**, где

**s=loess(X,Y,L)** - вектор коэффициентов для построения аппроксимирующего полинома второй степени, требуемый функцией **interp**,

L>0 - параметр, определяющий размер отрезков полиномов.

Параметр L задает степень сглаженности данных.

На практике  $0.2 < L < 2$ .

Насколько точно функция регрессии отображает массив данных можно установить с помощью коэффициента детерминированности  $R^2$ , который вычисляется по формуле

$$R^2 = 1 - \frac{S_{\text{ост}}}{S_{\text{общ}}}, \text{ где } S_{\text{ост}} = \sum_{i=1}^n (y_i - y_i^{\text{T}})^2, \quad S_{\text{общ}} = \sum_{i=1}^n (y_i - y_i^{\text{T}})^2 + \sum_{i=1}^n (y_i^{\text{T}} - \bar{y})^2.$$

Здесь  $y_i$  экспериментальные значения функции;  $y_i^{\text{T}}$  - теоретические значения функции;  $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$

- среднее арифметическое значение по y.

$R^2$  может изменяться в пределах от 0 до 1. Чем ближе  $R^2$  к 1, тем лучше функция регрессии отображает экспериментальные данные.

#### Пример №3

Аппроксимировать массив данных из примера 1:

1. С помощью функции **regress**, используя полиномы 1, 2, 3 и 4 степени. Установить полином какой степени наиболее точно аппроксимирует заданный массив.

2. С помощью функции **loess**. Подобрать параметр L, который наиболее точно аппроксимирует заданный массив.

Массив экспериментальных данных

ORIGIN := 1

$X := \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$

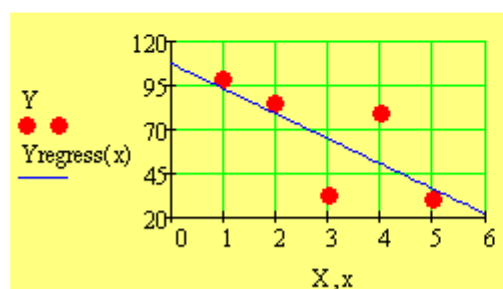
$Y := \begin{pmatrix} 98 \\ 85 \\ 32 \\ 79 \\ 30 \end{pmatrix}$

$x := 0, 0.1..6$

$Ysr := \text{mean}(Y) = 64.8$

N := 1    степень полинома

Yregress(x) := interp(regress(X,Y,N),X,Y,x)



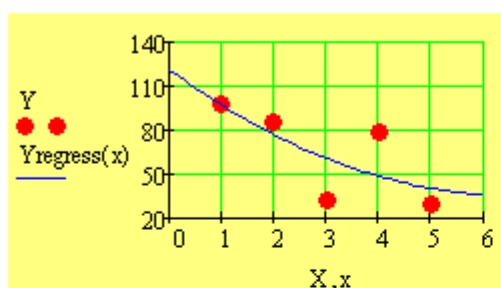
$$S_o := \sum_{m=1}^5 (Y_m - Y_{\text{regress}}(X_m))^2 = 1.982 \times 10^3$$

$$S_f := \sum_{m=1}^5 (Y_{\text{regress}}(X_m) - Y_{sr})^2 = 2.016 \times 10^3$$

$$S_x := S_o + S_f = 3.999 \times 10^3 \quad K_d := 1 - \frac{S_o}{S} = 0.504$$

N := 2    степень полинома

Yregress(x) := interp(regress(X,Y,N),X,Y,x)



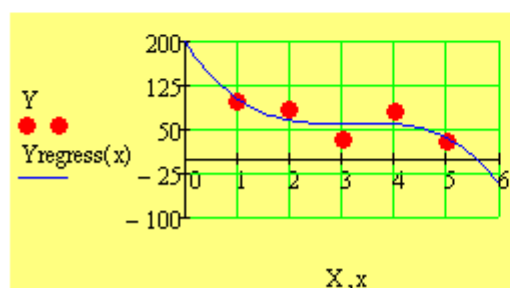
$$S_o := \sum_{m=1}^5 (Y_m - Y_{\text{regress}}(X_m))^2 = 1.926 \times 10^3$$

$$S_f := \sum_{m=1}^5 (Y_{\text{regress}}(X_m) - Y_{sr})^2 = 2.072 \times 10^3$$

$$S_x := S_o + S_f = 3.999 \times 10^3 \quad K_d := 1 - \frac{S_o}{S} = 0.518$$

N := 3    степень полинома

Yregress(x) := interp(regress(X,Y,N),X,Y,x)



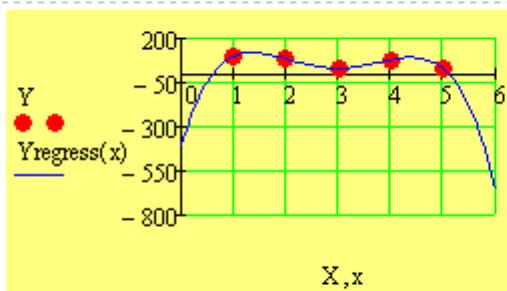
$$S_o := \sum_{m=1}^5 (Y_m - Y_{\text{regress}}(X_m))^2 = 1.613 \times 10^3$$

$$S_f := \sum_{m=1}^5 (Y_{\text{regress}}(X_m) - Y_{sr})^2 = 2.386 \times 10^3$$

$$S_x := S_o + S_f = 3.999 \times 10^3 \quad K_d := 1 - \frac{S_o}{S} = 0.597$$

$N := 4$  степень полинома

$Y_{\text{regress}}(x) := \text{interp}(\text{regress}(X, Y, N), X, Y, x)$



$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{regress}}(X_m))^2 = 0$$

$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{regress}}(X_m) - Y_{\text{sr}})^2 = 3.999 \times 10^3$$

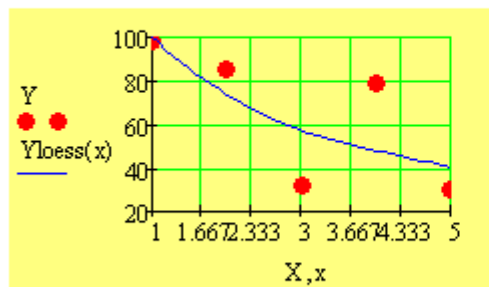
$$S_{\text{а}} := S_{\text{о}} + S_{\text{f}} = 3.999 \times 10^3$$

$$K_{\text{д}} := 1 - \frac{S_{\text{о}}}{S} = 1$$

Установили: полином четвертой степени наиболее точно аппроксимирует заданный массив данных

$L := 2$  характеристика длины отрезков полинома

$Y_{\text{loess}}(x) := \text{interp}(\text{loess}(X, Y, L), X, Y, x)$



$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{loess}}(X_m))^2 = 1.829 \times 10^3$$

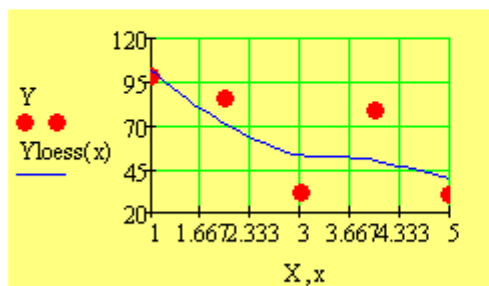
$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{loess}}(X_m) - Y_{\text{sr}})^2 = 2.249 \times 10^3$$

$$S_{\text{а}} := S_{\text{о}} + S_{\text{f}} = 4.079 \times 10^3$$

$$K_{\text{д}} := 1 - \frac{S_{\text{о}}}{S} = 0.551$$

$L := 1.5$  характеристика длины отрезков полинома

$Y_{\text{loess}}(x) := \text{interp}(\text{loess}(X, Y, L), X, Y, x)$



$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{loess}}(X_m))^2 = 1.585 \times 10^3$$

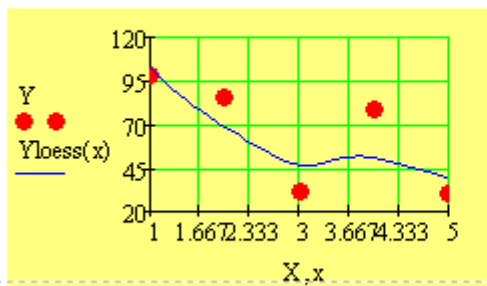
$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{loess}}(X_m) - Y_{\text{sr}})^2 = 2.371 \times 10^3$$

$$S_{\text{а}} := S_{\text{о}} + S_{\text{f}} = 3.956 \times 10^3$$

$$K_{\text{д}} := 1 - \frac{S_{\text{о}}}{S} = 0.599$$

$L := 1.3$  характеристика длины отрезков полинома

$$Y_{\text{loess}}(x) := \text{interp}(\text{loess}(X, Y, L), X, Y, x)$$



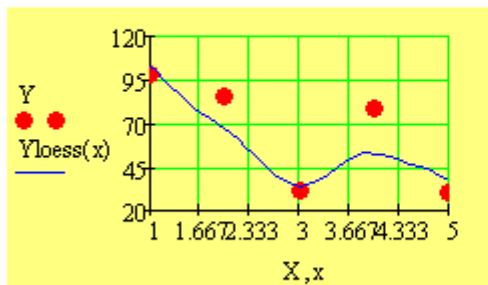
$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{loess}}(X_m))^2 = 1.375 \times 10^3$$

$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{loess}}(X_m) - Y_{\text{sr}})^2 = 2.612 \times 10^3$$

$$S := S_{\text{о}} + S_{\text{f}} = 3.987 \times 10^3 \quad K_{\text{d}} := 1 - \frac{S_{\text{о}}}{S} = 0.655$$

$L := 1.1$  характеристика длины отрезков полинома

$$Y_{\text{loess}}(x) := \text{interp}(\text{loess}(X, Y, L), X, Y, x)$$



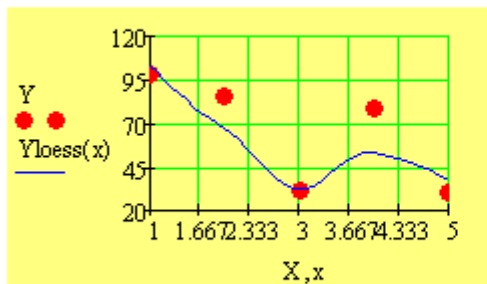
$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{loess}}(X_m))^2 = 1.117 \times 10^3$$

$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{loess}}(X_m) - Y_{\text{sr}})^2 = 3.305 \times 10^3$$

$$S := S_{\text{о}} + S_{\text{f}} = 4.421 \times 10^3 \quad K_{\text{d}} := 1 - \frac{S_{\text{о}}}{S} = 0.747$$

$L := 1.04$  характеристика длины отрезков полинома

$$Y_{\text{loess}}(x) := \text{interp}(\text{loess}(X, Y, L), X, Y, x)$$



$$S_{\text{о}} := \sum_{m=1}^5 (Y_m - Y_{\text{loess}}(X_m))^2 = 1.078 \times 10^3$$

$$S_{\text{f}} := \sum_{m=1}^5 (Y_{\text{loess}}(X_m) - Y_{\text{sr}})^2 = 3.434 \times 10^3$$

$$S := S_{\text{о}} + S_{\text{f}} = 4.512 \times 10^3 \quad K_{\text{d}} := 1 - \frac{S_{\text{о}}}{S} = 0.761$$

Установили: при  $L=1.04$  функция loess наиболее точно аппроксимирует заданный массив данных

Однако функция regress при аппроксимации полиномом четвертой степени давала еще более лучшую аппроксимацию

## ЗАДАНИЕ.

Выполнить для экспериментального массива данных следующие действия:

1. Провести кусочно-линейную аппроксимацию.
2. Провести аппроксимацию сплайнами.
3. Аппроксимировать:
  - а) с помощью функции regress, используя полиномы различной степени. Установить полином какой степени наиболее точно аппроксимирует заданный массив;
  - б) с помощью функции loess. Подобрать параметр L, который наиболее точно аппроксимирует заданный массив.

### ВАРИАНТЫ ЗАДАНИЙ

1.	X	3	5	6	7	8	10	12	14
	Y	-0,9	5,4	8,8	7,8	3,3	-0,2	8,2	4,7

2.	X	-2	0	1	3	4	6	7	9	10
	Y	-6,5	-2,0	2,2	-1,3	-5,8	-3,4	1,3	0,1	-4,7

3.	X	-3	-1	2	4,5	7	10	12	14,5	17
	Y	7,0	7,7	7,8	8,0	7,4	7,3	7,3	7,9	7,9

4.	X	-3,3	-2,1	-0,9	0,5	1,5	2,7	4
	Y	-661,2	-93,0	-4,3	-0,8	-9,3	15,3	464

5.	X	0,5	2,8	5	7,4	10	12	14,3	16,6	19
	Y	4,0	9,4	13,7	11,2	15,0	15,5	13,0	17,0	15,5

6.	X	-4	-2,4	-1	1	2,4	4	5,6	7,2	9	11
	Y	9,6	11,0	8,7	5,3	8,3	12,7	13,4	11,3	12,8	16,2

7.	X	-4	-2,4	-0,2	1	2,4	3,6	6	7,2	9	12
	Y	-4,8	2,8	-5,0	-8,0	-5,3	2,2	2,3	-3,8	-1,1	5,2

8.	X	-3	-2,4	-2	-1	0,5	1	1,7	3	3,4
	Y	-6,9	-5,1	-2,8	4,1	7,8	6,5	4,6	13,2	23,2

9.	X	-4	-3	-2	-1	0,5	1	2,2	3	4
	Y	184,3	38,9	4,5	0,1	0,0	-0,1	-2,9	-9,7	-20,5

10.	X	-4	0,4	2	5	8,3	11	14,7	17	20
	Y	0,5	-2,6	0,3	-4,5	1,5	-3,9	1,9	-3,4	2,3

<b>11.</b>	X	1	3	4	7	9	10	12	14	20	21
	Y	4,1	-6,9	-4,6	5,3	-6,4	-5,9	5,9	1,0	2,9	-3,8

<b>12.</b>	X	1	2	5	8	9	11	13	14	15	16
	Y	1,1	2,5	0,9	1,6	0,0	2,0	0,0	1,6	0,9	0,6

<b>13.</b>	X	-6	-1	0	6	11	14	18	21	26	28
	Y	0,7	1,1	-1,0	0,7	2,0	1,6	2,4	1,3	2,4	0,4

<b>14.</b>	X	0,3	1,8	3,3	4,8	6,3	7,8	9,3	10,8	12,3	13,8
	Y	0,2	0,2	2,0	1,7	1,8	1,8	2,0	0,8	0,9	1,9

<b>15.</b>	X	-7,5	-4,2	-0,9	2,4	5,7	9	12,3	15,6	18,9
	Y	1,6	0,9	1,3	2,0	0,2	1,8	2,0	1,9	0,7