Методы обработки числовых данных

Методы обработки числовых данных

Числовые данные

- **Таблицы**
- **Экспериментальные данные**

Цель

Получение функциональной зависимости $y = f(x)$

2 подхода

1. **Интерполяция - аппроксимирующая функция должна пройти через все точки**

2. **Регрессия - аппроксимирующая функция не обязательно должна проходить через все точки**

Интерполяция

Сущность интерполяции состоит в отыскании значения функции в некоторой промежуточной точке

Виды интерполяции:

- **интерполяция по Лагранжу;**
- **линейная;**
- **квадратичная;**
- **сплайн-интерполяция.**

Интерполяция по Лагранжу

Интерполяционный полином

$$P_n(x) = y_0 b_0(x) + y_1 b_1(x) + \cdots + y_n b_n(x)$$

$b_j(x)$ **– многочлены степени *n***

Система уравнений

$$P_n(x_i) = y_i \qquad i = 0, 1, \ldots, n$$

или

$$y_0 b_0(x_0) + y_1 b_1(x_0) + \cdots + y_n b_n(x_0) = y_0,$$

$$\ldots$$

$$y_0 b_0(x_n) + y_1 b_1(x_n) + \cdots + y_n b_n(x_n) = y_n$$

Интерполяция по Лагранжу

$$b_j(x_i) = \begin{cases} 1 \text{ при } i = j \\ 0 \text{ при } i \neq j \end{cases}$$

$$b_j(x) = C_j(x - x_0)(x - x_1) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n)$$

$$b_j(x_j) = 1$$

$$C_j = 1/((x_j - x_0)(x_j - x_1) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n))$$

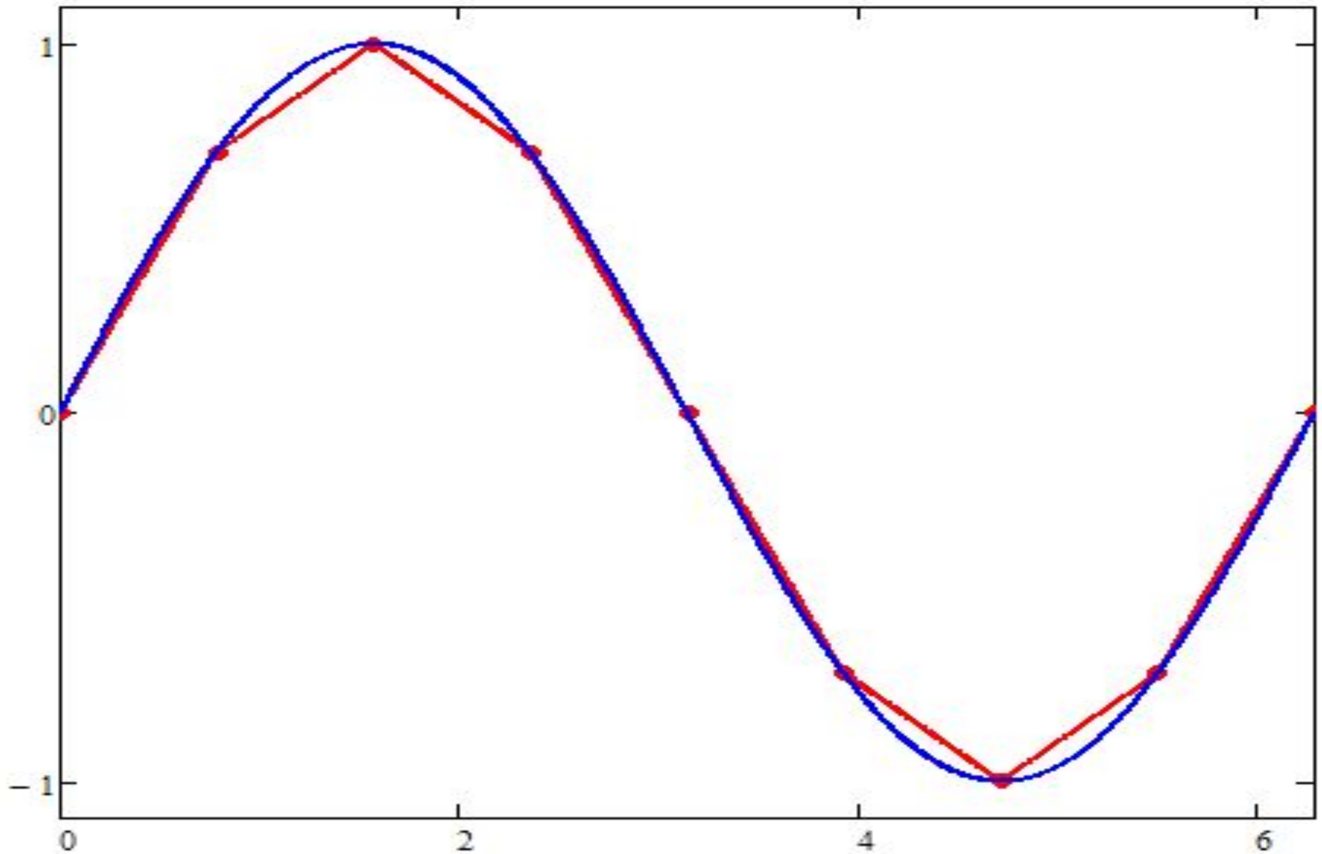
$$P_n(x) = \sum_{j=0}^{n} y_j \frac{(x - x_0)(x - x_1) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n)}{(x_j - x_0)(x_j - x_1) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n)}$$

$$L_j(x) = (x - x_0)(x - x_1) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n)$$

$$P_n(x) = \sum_{j=0}^{n} y_j \frac{L_j(x)}{L_j(x_j)}$$

Линейная интерполяция

В основе линейной интерполяции лежит аппроксимация кривой на участке между точками (x_i, y_i) и (x_{i+1}, y_{i+1}) прямой, проходящей через те же точки



Линейная интерполяция

Уравнение прямой, проходящей через 2 точки

$$y - y_i = k(x - x_i), \qquad k = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$$

Линейная интерполяция (Mathcad)

linterp(X,Y,x)

X – вектор табличных значений аргумента;

Y – вектор табличных значений функции;

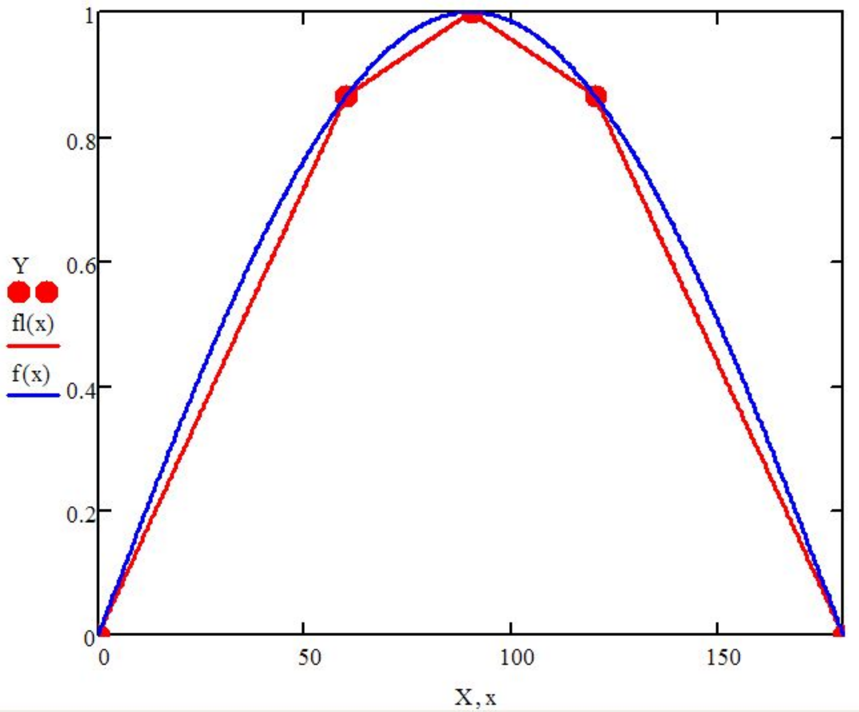
x – значение аргумента, при котором вычисляется интерполирующее значение функции.

Линейная интерполяция

$X := (0\ \ 60\ \ 90\ \ 120\ \ 180)$ $\qquad X := X^T$

$Y := (0\ \ 0.866\ \ 1\ \ 0.866\ \ 0)$ $\qquad Y := Y^T$

$fl(x) := \mathrm{linterp}(X, Y, x)$ $\qquad f(x) := \sin\left(\frac{x \cdot \pi}{180}\right)$



$\mathrm{linterp}(X, Y, 30) = 0.433$

$fl(30) = 0.433$

$f(30) = 0.5$

Сплайн-интерполяция

Сплайн – это группа сопряженных кубических многочленов, в местах сопряжения которых первая и вторая производные непрерывны.

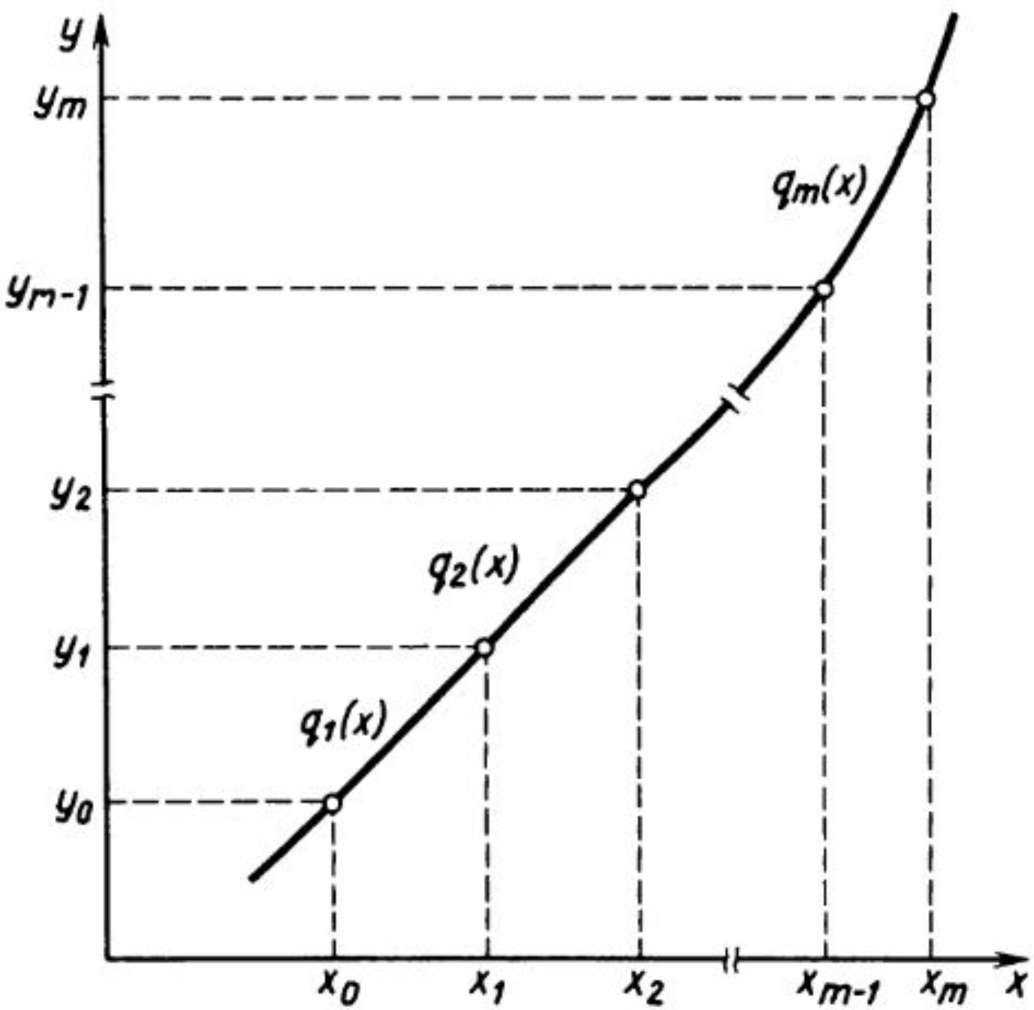
Такие функции называют кубическими сплайнами.

Чтобы построить кубический сплайн, необходимо задать коэффициенты, которые единственным образом определяют кубический многочлен в промежутке между данными точками.
Необходимо задать все кубические функции $q_1(x)$, $q_2(x)$, … , $q_m(x)$.

$$q_i(x) = k_{i1} + k_{i2}x + k_{i3}x^2 + k_{i4}x^3$$

$$i = 1,2, \dots, m$$

Сплайн-интерполяция



Сплайн-интерполяция

ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТОВ СПЛАЙНОВ

Условие непрерывности сплайна

$$q_i(x_i) = y_i, \quad i = 1, \dots, m$$

$$q_{i+1}(x_i) = y_i, \quad i = 0, \dots, m-1$$

Условие непрерывности первых производных

$$q'_{i+1}(x_i) = q'_i(x_i), \quad i = 1, \dots, m-1$$

Условие непрерывности вторых производных

$$q''_{i+1}(x_i) = q''_i(x_i), \quad i = 1, \dots, m-1$$

Два недостающих уравнения получают из ограничений на значение сплайна и его производных на концах промежутка $[a, b]$; их называют краевыми условиями.

«Естественные» краевые условия

$$q''_1(x_0) = 0 \text{ и } q''_m(x_m) = 0$$

Сплайн-интерполяция

$$q_i(x) = ty_i + (1-t)y_{i-1} + \Delta x_i[(k_{i-1} - d_i)t(1-t)^2 - (k_i - d_i)t^2(1-t)]$$

$$i = 1, \dots, m$$

$$\Delta x_i = x_i - x_{i-1} \qquad t = \frac{x - x_{i-1}}{\Delta x_i} \qquad d_i = \frac{y_i - y_{i-1}}{\Delta x_i}$$

При таком выборе кубических многочленов автоматически удовле творяются все условия, кроме условий, налагаемых на вторые производные

Условие непрерывности вторых производных

для внутренних точек

$$k_{i-1}\Delta x_{i+1} + 2k_i(\Delta x_i + \Delta x_{i+1}) + k_{i+1}\Delta x_i = 3(d_i \Delta x_{i+1} + d_{i+1}\Delta x_i)$$

для двух внешних $\quad 2k_0 + k_1 = 3d_1$

$$k_{m-1} + 2k_m = 3d_m$$

Сплайн-интерполяция

Система уравнений

$$\begin{bmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \Delta x_2 & 2(\Delta x_1 + \Delta x_2) & \Delta x_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \Delta x_3 & 2(\Delta x_2 + \Delta x_3) & \Delta x_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \Delta x_m & 2(\Delta x_{m-1} + \Delta x_m) & \Delta x_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 2 \end{bmatrix} \begin{pmatrix} k_0 \\ k_1 \\ k_2 \\ \vdots \\ k_{m-1} \\ k_m \end{pmatrix} = \begin{pmatrix} d_1 \\ d_1 \Delta x_2 + d_2 \Delta x_1 \\ d_2 \Delta x_3 + d_3 \Delta x_2 \\ \vdots \\ d_{m-1} \Delta x_m + d_m \Delta x_{m-1} \\ d_m \end{pmatrix}$$

Система уравнений – трехдиагональная

Для решения таких систем используется метод прогонки

Существуют и другие сплайны, получающиеся при других условиях на концах или использовании многочленов более высоких степеней

Сплайн-интерполяция (Mathcad)

`interp(vs,X,Y,x)`

`vs` - **вектор вторых производных, созданный функцией `lspline(X,Y)`, `spline(X,Y)` или `cspline(X,Y)`;**

`X` – вектор табличных значений аргумента;

`Y` – вектор табличных значений функции;

`x` – значение аргумента, при котором вычисляется интерполирующее значение функции.

`lspline(X,Y)` – создает вектор коэффициентов кривой, которая приближается к прямой линии в граничных точках;

`pspline(X,Y)` – создает вектор коэффициентов кривой, которая приближается к квадратичной параболе в граничных точках;

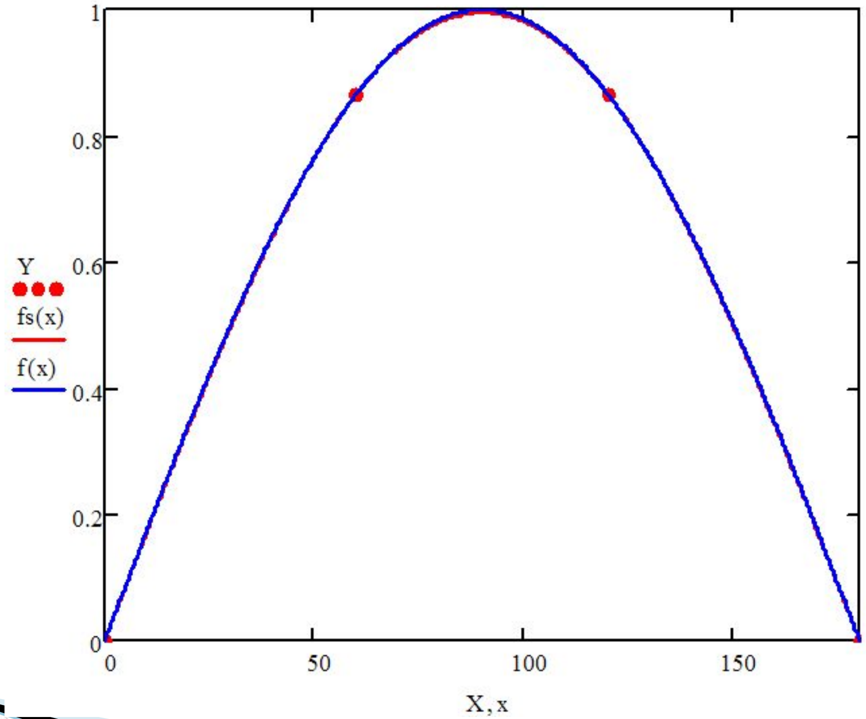
`cspline(X,Y)` – создает вектор коэффициентов кривой, которая приближается к кубической параболе в граничных точках.

Сплайн-интерполяция (Mathcad)

$X := (0\ \ 60\ \ 120\ \ 180)\qquad X := X^T$

$Y := (0\ \ 0.866\ \ 0.866\ \ 0)\qquad Y := Y^T$

$fs(x) := \mathrm{interp}(\mathrm{lspline}(X,Y),X,Y,x)\qquad f(x) := \sin\left(\frac{x\cdot\pi}{180}\right)$



$fs(30) = 0.498$

$fs(90) = 0.996$

Обработка экспериментальных данных
Метод наименьших квадратов

$$y = a_0 + a_1 x$$

$$y = a_0 + a_1 x + a_2 x^2$$

$$S = \sum_{i=1}^{n} (f(x_i) - y_i)^2 = \sum_{i=1}^{n} (a_0 + a_1 x_i - y_i)^2$$

$$\frac{\partial S}{\partial a_0} = 2\sum_{i=1}^{n} (a_0 + a_1 x_i - y_i) \cdot 1 = 0$$

$$\frac{\partial S}{\partial a_1} = 2\sum_{i=1}^{n} (a_0 + a_1 x_i - y_i) \cdot x_i = 0$$

$$\begin{cases} n \cdot a_0 + \sum_{i=1}^{n} x_i \cdot a_1 = \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i \cdot a_0 + \sum_{i=1}^{n} x_i^2 \cdot a_1 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

Метод наименьших квадратов
Реализация в Mathcad (способ 1)

$ORIGIN := 1 \qquad n := 6$

$y = 1 + 2x$

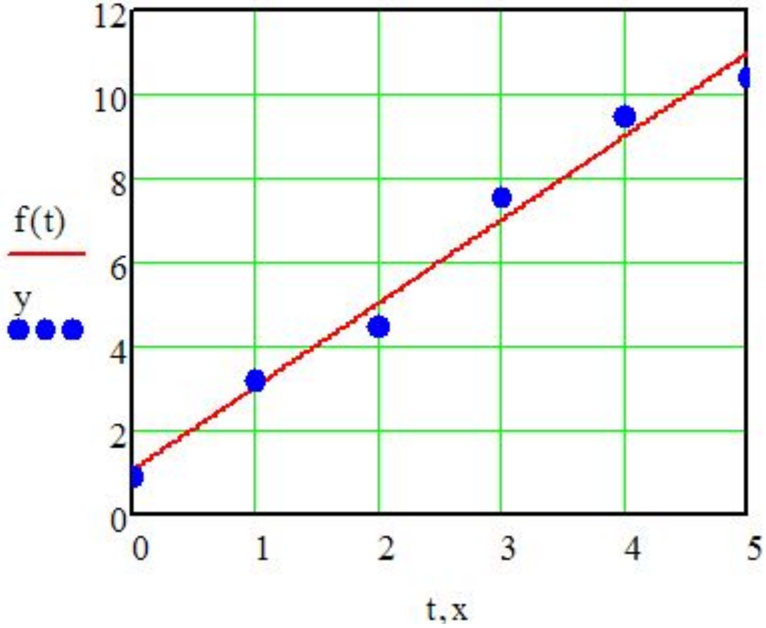
$$x := \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \qquad y := \begin{pmatrix} 0.9 \\ 3.2 \\ 4.5 \\ 7.6 \\ 9.5 \\ 10.4 \end{pmatrix}$$

$$A := \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} \left(x_i\right)^2 \end{bmatrix} \qquad B := \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} \left(x_i \cdot y_i\right) \end{bmatrix}$$

Метод наименьших квадратов
Реализация в Mathcad

$X := A^{-1}B = \begin{pmatrix} 1.052 \\ 1.986 \end{pmatrix}$ $a_0 := X_1$ $a_1 := X_2$

$f(x) := a_0 + a_1 x$



$corr(f(x), y) = 0.991$

Метод наименьших квадратов
Реализация в Mathcad (способ 2)

$ORIGIN := 1 \qquad n := 6 \qquad\qquad y = 1 + 2x$

$$x := \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \qquad y := \begin{pmatrix} 0.9 \\ 3.2 \\ 4.5 \\ 7.6 \\ 9.5 \\ 10.4 \end{pmatrix}$$

$i := 1 .. n$

$X_{i,1} := 1$

$X^{\langle 2 \rangle} := x \qquad X^{\langle 3 \rangle} := x^2$

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{pmatrix}$$

$A := X^T X \qquad B := X^T y \qquad X := A^{-1} B = \begin{pmatrix} 0.779 \\ 2.396 \\ -0.082 \end{pmatrix}$

Задание

1. Написать функцию с использованием C++Builder, аналогичную функции linterp в MathCAD (линейная интерполяция) с построением графиков.
2. Написать функцию с использованием C++Builder, реализующую интерполяцию по Лагранжу с построением графиков.
3. Написать функцию с использованием C++Builder, аналогичную функции interp в MathCAD (сплайн-интерполяция) с построением графиков.
4. Написать функцию с использованием C++Builder для обработки экспериментальных данных методом наименьших квадратов с построением графиков.

Контрольные вопросы

1. Использование линейной и сплайн-интерполяции в MathCAD.
2. Реализация метода наименьших квадратов в MathCAD
3. Интерполяция по Лагранжу.
4. Линейная интерполяция.
5. Сплайн-интерполяция.
6. Метод наименьших квадратов.

Благодарю
за внимание!