Глобальная и локальная интерполяции

Пусть задана таблица значений (ее называют *сеточной функцией*)

x_i	x_0	x_1	$\ldots$	x_n
y_i	y_0	y_1	$\ldots$	y_n

Интерполяция, использующая сразу все n узлов таблицы, называется *глобальной интерполяцией*.

Начиная с $n \geq 7$ глобальная многочленная интерполяция становиться неустойчивой, поэтому обычную многочленную интерполяцию осуществляют максимум по 3-4 узлам. Интерполяцию по нескольким узлам сеточной функции называют *локальной*.

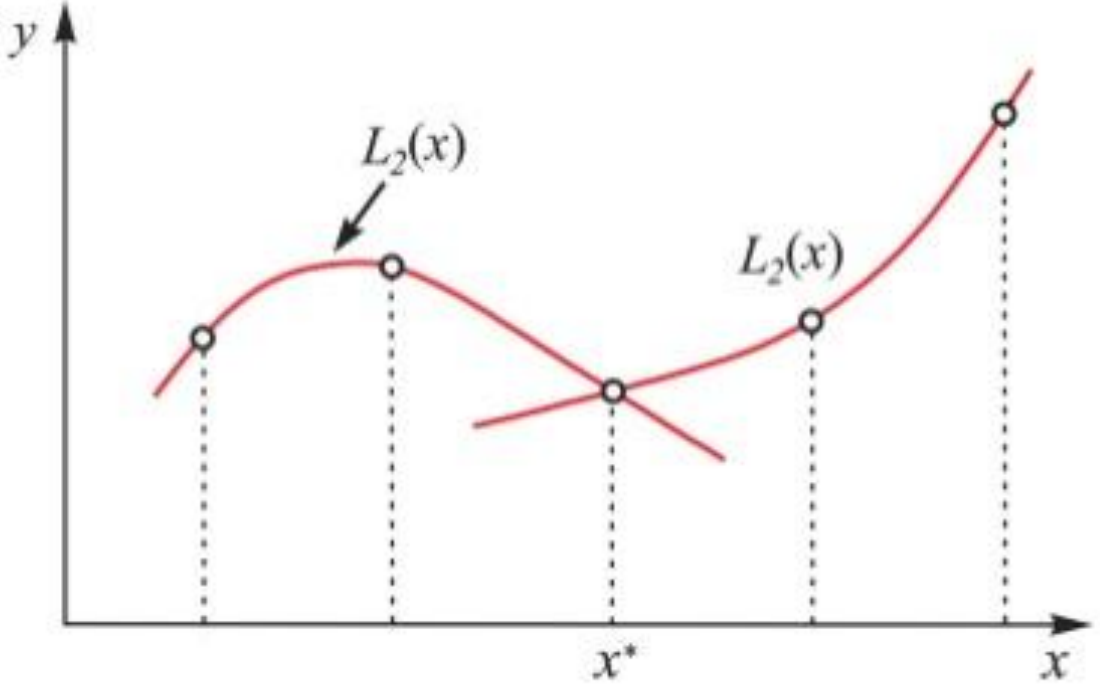
Недостаток локальной интерполяции

Однако локальная интерполяция обладает тем недостатком, что интерполирующая функция в узлах стыковки многочлена имеет непрерывность только нулевого порядка.

От этих недостатков свободна *сплайн-интерполяция*, которая требует непрерывности в узлах стыковки локальных многочленов по производным порядка один, два и т.д.

Определение сплайна

Сплайном степени m дефекта r называется $(m-r)$ раз непрерывно дифференцируемая функция, которая



на каждом отрезке $[x_i - 1, x_i], i = 1, 2 \ldots n$, представляет собой многочлен степени m.

Вывод интерполяционных кубических сплайнов

Сплайны, удовлетворяющие условию интерполяции называются *интерполяционными*.

Выведем интерполяционные кубические сплайны $S_k(x), k = \overline{1, n}$ дефекта один в соответствии с сеточной функцией. Кубический сплайн $S(x)$ на отрезке $[x_i - 1, x_i]$ имеет четыре неизвестных коэффициента. Количество отрезков $[x_i - 1, x_i]$ равно n.

$$S(x_i) = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \qquad i = \overline{1, n}$$

Для определения $4 \times n$ коэффициентов имеются следующие условия в узлах интерполяции:

1. Условие интерполяции $S(x_i) = y_i, i = \overline{0, n}$;
2. Непрерывность сплайнов $S(x_i - 0) = S(x_i + 0), i = \overline{1, n-1}$;
3. Непрерывность производных 1-го порядка $S'(x_i - 0) = S'(x_i + 0), i = \overline{1, n-1}$;
4. Непрерывность производных 2-го порядка $S''(x_i - 0) = S''(x_i + 0), i = \overline{1, n-1}$;
5. $S''(x_0) = 0$;
6. $S''(x_n) = 0$.

Пусть $S''(x) = q(x)$.

Рассмотрим поведение функции $q(x)$ На отрезке $[x_{i-1}, x_i]$.

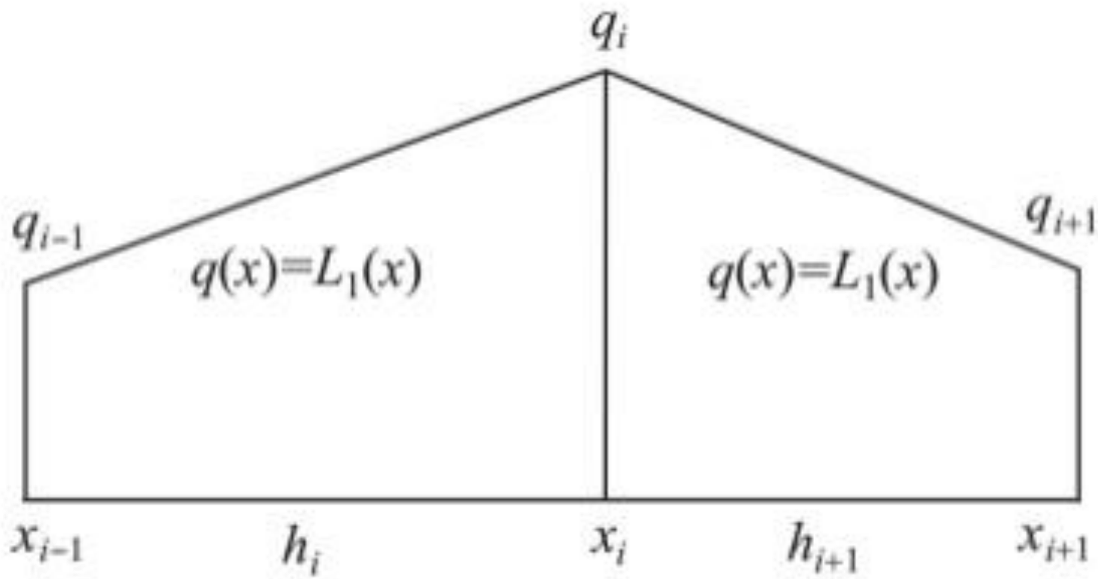


Рисунок 1. Поведение функции $S''(x)$ на элементарных отрезках

- Найдем вторую производную с помощью интерполяционного многочлена Лагранжа 1-й степени $L_1(x)$ на каждом отрезке $[x_{i-1}, x_i]$:

$$q(x) = q_{i-1}\frac{x - x_i}{x_{i-1} - x_i} + q_i\frac{x - x_{i-1}}{x_i - x_{i-1}}. \qquad (1)$$

Выражение (1) уже удовлетворяет условиям непрерывности производных 2-го порядка.

$$x = x_i - 0 \to (1) \to q(x_i - 0) = q_i$$

- Выписывая выражение (1) для отрезка $[x_i, x_{i+1}]$:

$$q(x) = q_i\frac{x_{i+1} - x}{h_{i+1}} + q_{i+1}\frac{x - x_i}{h_{i+1}}, \qquad i = \overline{1, n-1},$$

и подставляя в него $x_i + 0$ вместо x, получим $q(x_i + 0) = q_i$, что и требовалось показать.

Этапы нахождения сплайна

1. Проинтегрируем дважды выражение (1), получим

$$S(x) = q_{i-1}\frac{(x_i - x)^3}{6h_i} + q_i\frac{(x - x_{i-1})^3}{6h_i} + C_1 x + C_2, \qquad (2)$$

где C_1 и C_2 найдем из удовлетворения значений сплайна (2) в узлах x_{i-1}, x_i условиям интерполяции

$$S(x_{i-1}) = y_{i-1} = q_{i-1}\frac{(x_i - x_{i-1})^3}{6h_i} + q_i\frac{(x_{i-1} - x_{i-1})^3}{6h_i} + C_1 x_{i-1} + C_2,$$

$$S(x_i) = y_i = q_{i-1}\frac{(x_i - x_i)^3}{6h_i} + q_i\frac{(x_i - x_{i-1})^3}{6h_i} + C_1 x_i + C_2.$$

2. Решая эту СЛАУ относительно C_1, C_2 и подставляя их в (2), получим:

$$S(x) = q_{i-1}\frac{(x_i - x)^3}{6h_i} + q_i\frac{(x - x_{i-1})^3}{6h_i} + \left(\frac{y_{i-1}}{h_i} - q_{i-1}\frac{h_i}{6}\right) \times$$

$$\times (x_i - x) + \left(\frac{y_i}{h_i} - q_i\frac{h_i}{6}\right)(x - x_{i-1}), \qquad x \in [x_{i-1}, x_i]. \qquad (3)$$

3. Узловые значения для вторых производных q_i будем искать из условий непрерывности первых производных в узлах x_i.

4. Для нахождения производной $S'(x_i+0)$ запишем (3) для отрезка $[x_i, x_{i+1}]$

$$S(x) = q_i \frac{(x_{i+1}-x)^3}{6h_{i+1}} + q_{i+1}\frac{(x-x_i)^3}{6h_{i+1}} + \left(\frac{y_i}{h_{i+1}} - q_i\frac{h_{i+1}}{6}\right) \times$$

$$\times (x_{i+1}-x) + \left(\frac{y_{i+1}}{h_{i+1}} - q_{i+1}\frac{h_{i+1}}{6}\right)(x-x_i), \qquad x \in [x_i, x_{i+1}]. \quad (4)$$

5. Вычисляя производные первого порядка от (3) и (4) и подставляя в них значение $x = x_i$, получим

$$S'(x_i-0) = q_{i-1}\frac{h_i}{6} + q_i\frac{h_i}{3} + \frac{y_i - y_{i-1}}{h_i},$$

$$S'(x_i+0) = -q_i\frac{h_{i+1}}{3} - q_{i+1}\frac{h_{i+1}}{6} + \frac{y_{i+1}-y_i}{h_{i+1}}.$$

6. Приравнивая эти выражения в соответствии с условиями непрерывности первых производных в узлах интерполяции x_i, получим

$$q_{i-1}\frac{h_i}{6}+q_i\frac{h_i+h_{i+1}}{3}+q_{i+1}\frac{h_{i+1}}{6}=\frac{y_{i+1}-y_i}{h_{i+1}}-\frac{y_i-y_{i-1}}{h_i}, \qquad (5)$$

$$i=\overline{1,n-1};$$

$$q_0=q_n=0. \qquad (6)$$

7. Подставляя найденные $q_i, i=\overline{0,n}$ в (3), получим кубические сплайны дефекта один на каждом отрезке $x\in[x_{i-1},x_i], i=\overline{1,n}$.

Таким образом, определяющими выражениями для нахождения кубических сплайнов дефекта один являются выражения $(4),(5),(6)$.

Пример

Для заданной таблицы с $h = x_i - x_{i-1} = 1 = const$ выписать интерполяционные кубические сплайны дефекта один на каждом отрезке $x \in [x_{i-1}, x_i], i = \overline{1,4}$. Проверить непрерывность сплайнов и их производных до второго порядка включительно в узле $x^* = 2$.

i	0	1	2	3	4
x_i	$x_0 = 1$	$x_1 = 2$	$x_2 = 3$	$x_3 = 4$	$x_4 = 5$
y_i	$y_0 = 1$	$y_1 = 3$	$y_2 = 6$	$y_3 = 9$	$y_4 = 21$

Решение. Для узлов $x_1 = 2; x_2 = 3; x_3 = 4$ с учетом $q_0 = q_4 = 0$ составляется СЛАУ (5) относительно неизвестных q_1, q_2, q_3:

$$i = 1: \qquad \frac{2}{3}q_1 + \frac{1}{6}q_2 = \frac{y_2 - y_1}{1} - \frac{y_1 - y_0}{1} = 1,$$

$$i = 2: \qquad \frac{1}{6}q_1 + \frac{2}{3}q_2 + \frac{1}{6}q_3 = \frac{y_3 - y_2}{1} - \frac{y_2 - y_1}{1} = 0,$$

$$i = 3: \qquad \frac{1}{6}q_2 + \frac{2}{3}q_3 = \frac{y_4 - y_3}{1} - \frac{y_3 - y_2}{1} = 9.$$

Вычисляются прогоночные коэффициенты по формулам

$$A_i = \frac{-c_i}{b_i + a_i A_{i-1}}; \quad B_i = \frac{d_i - a_i B_{i-1}}{b_i + a_i A_{i-1}}, \quad i = 1, 2, 3;$$

$$a_1 = c_3 = 0, \quad A_1 = -\frac{1}{4}; \quad B_1 = \frac{3}{2}; \quad A_2 = -\frac{4}{15}; \quad B_2 = -\frac{2}{5};$$

$$A_3 = 0; \quad B_3 = \frac{102}{7}$$

и значения

$$q_i = A_i q_{i+1} + B_i, \quad i = 3, 2, 1: \quad q_3 = A_3 q_4 + B_3 = B_3 = 102/7;$$

$$q_2 = A_2 q_3 + B_2 = -\frac{4}{15} \cdot \frac{102}{7} - \frac{2}{5} = -\frac{30}{7};$$

$$q_1 = A_1 q_2 + B_1 = -\frac{1}{4} \cdot \left(-\frac{30}{7}\right) + \frac{3}{2} = \frac{18}{7}.$$

Для каждого из четырех интервалов выписываем сплайны (3):

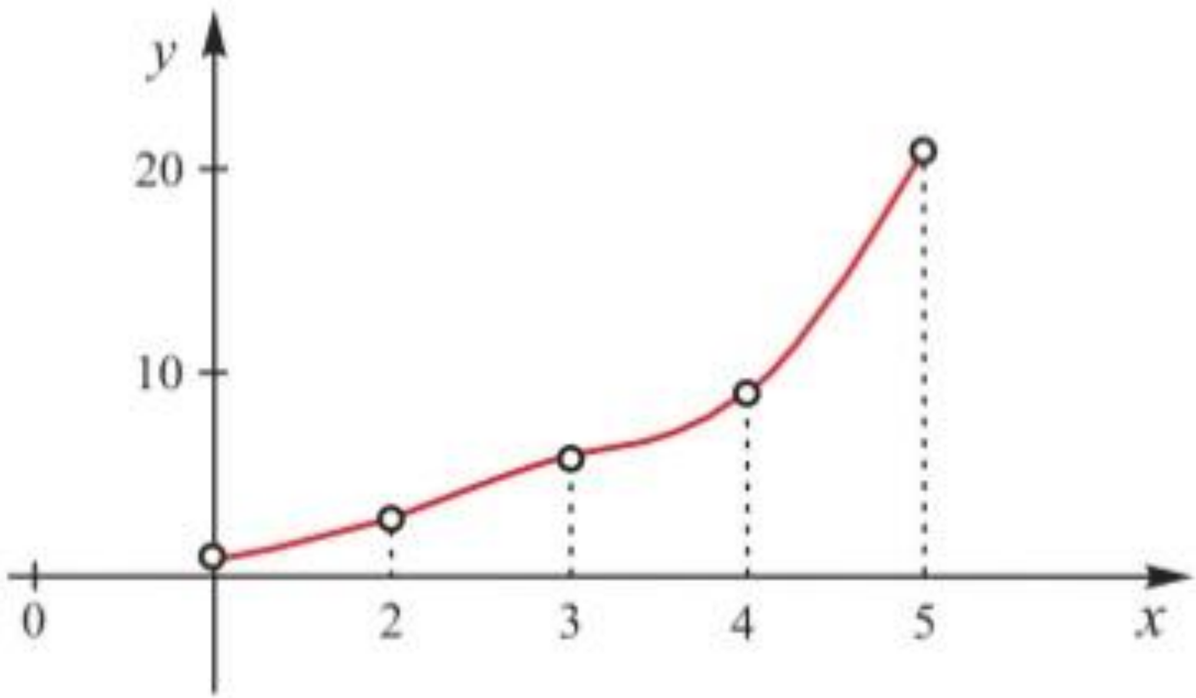
$$i=1: \quad S_1(x) = \frac{18}{42}(x-1)^3 + (2-x) + \frac{108}{42}(x-1), \quad x \in [1;2];$$

$$i=2: \; S_2(x) = \frac{18}{42}(3-x)^3 + \left(-\frac{30}{42}\right)(x-2)^3 + \\ + \frac{108}{42}(3-x) + \frac{282}{42}(x-2), \quad x \in [2;3];$$

$$i=3: \; S_3(x) = -\frac{30}{42}(4-x)^3 + \frac{102}{42}(x-3)^2 + \\ + \frac{282}{42}4 - x + \frac{276}{42}(x-3), \quad x \in [3;4];$$

$$i = 4: \ S_4(x) = \frac{102}{42}(5-x)^3 + \frac{276}{42}(5-x) + 21(x-4), \quad x \in [4;5].$$

Построим полученные сплайны:



Проверим правильность построения сплайнов для узла $x^* = 2$.К нему примыкают сплайны $S_1(x)$ и $S_2(x)$:

$$S_1(2-0) = 3; \qquad S_2(2+0) = 3;$$

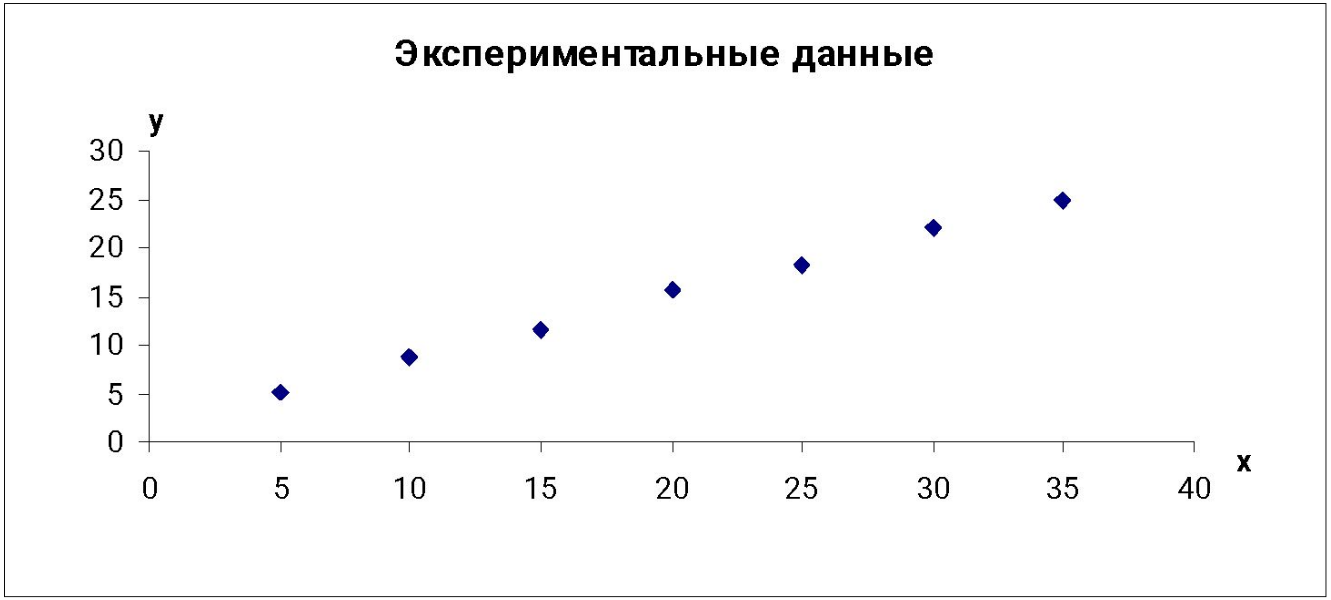
$$S_1'(2-0) = \frac{120}{42}; \qquad S_2'(2+0) = \frac{120}{42};$$

$$S_1''(2-0) = \frac{108}{42}; \qquad S_2''(2+0) = \frac{108}{42}.$$

МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

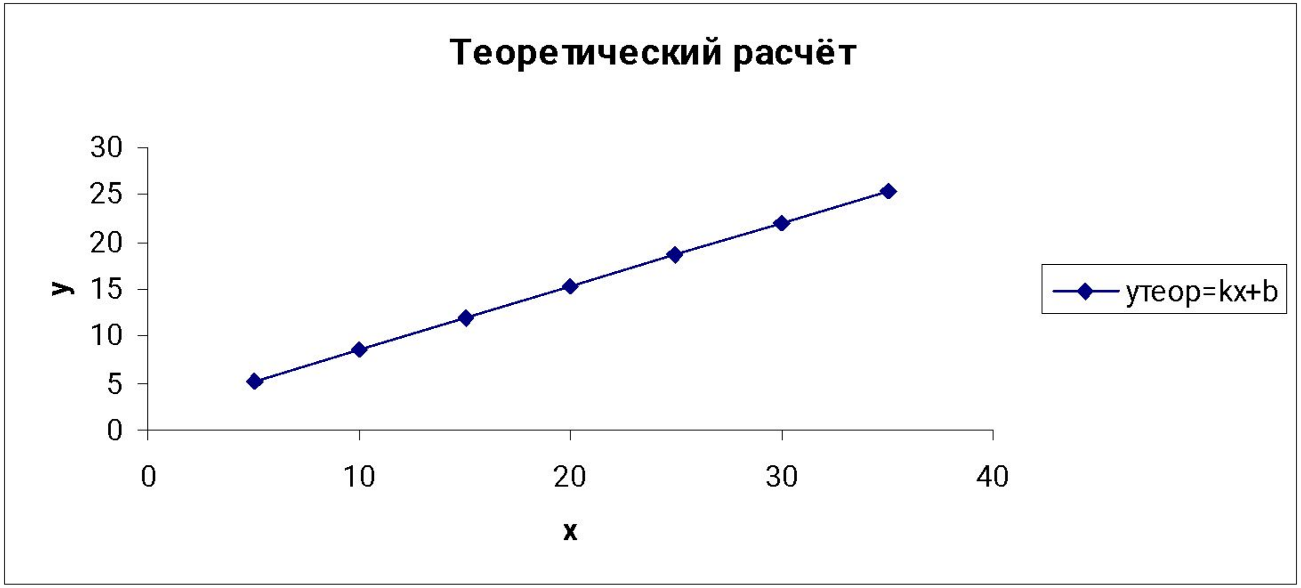
Количественный анализ

- Проведение количественного анализа, как правило, включает в себя построение графика по данным, найденным в ходе эксперимента

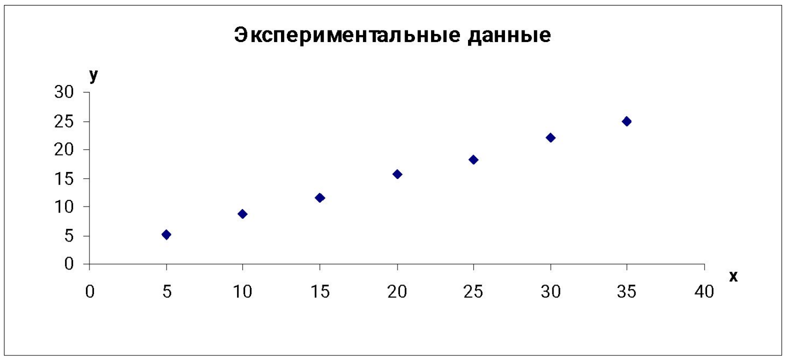


КОЛИЧЕСТВЕННЫЙ АНАЛИЗ

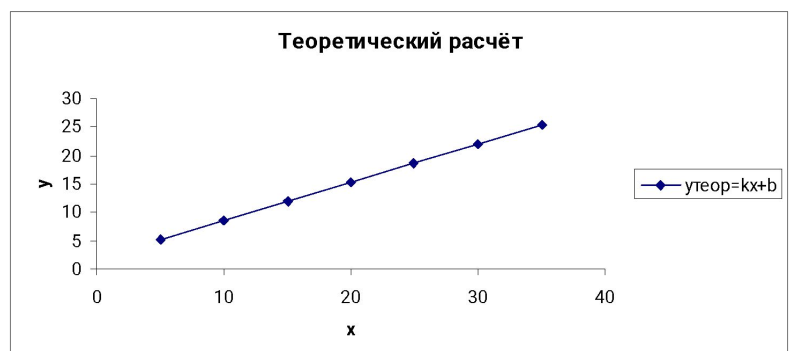
- Теоретически результаты эксперимента должны укладываться в некоторую зависимость, которую можно выразить *формулой*.



КОЛИЧЕСТВЕННЫЙ АНАЛИЗ



- Но на практике это не так



ОШИБКА!

- Причины:
  - Погрешность измерений
  - Недостигаемость условий (идеальный газ, стандартное давление и т.д.)
  - Ошибка в расчете

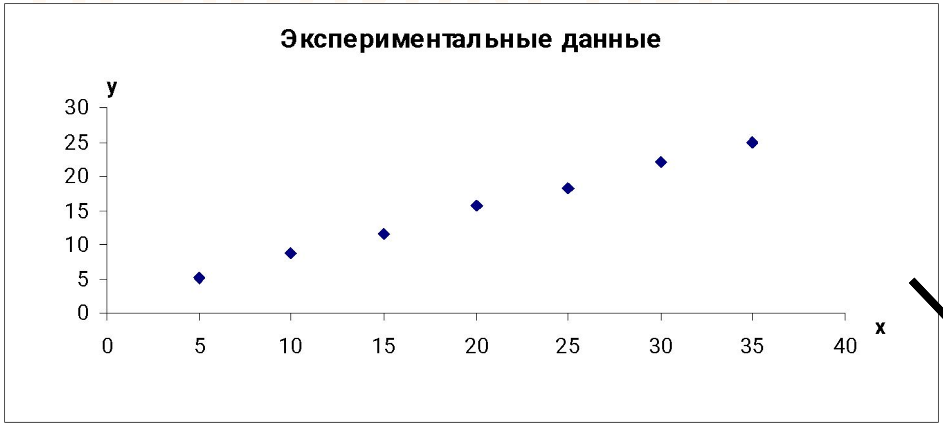
МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

- Это один из методов регрессионного анализа для оценки неизвестных величин по результатам измерений, содержащих случайные ошибки.
- Метод наименьших квадратов применяется также для приближенного представления заданной функции другими (более простыми) функциями и часто оказывается полезным при обработке наблюдений.

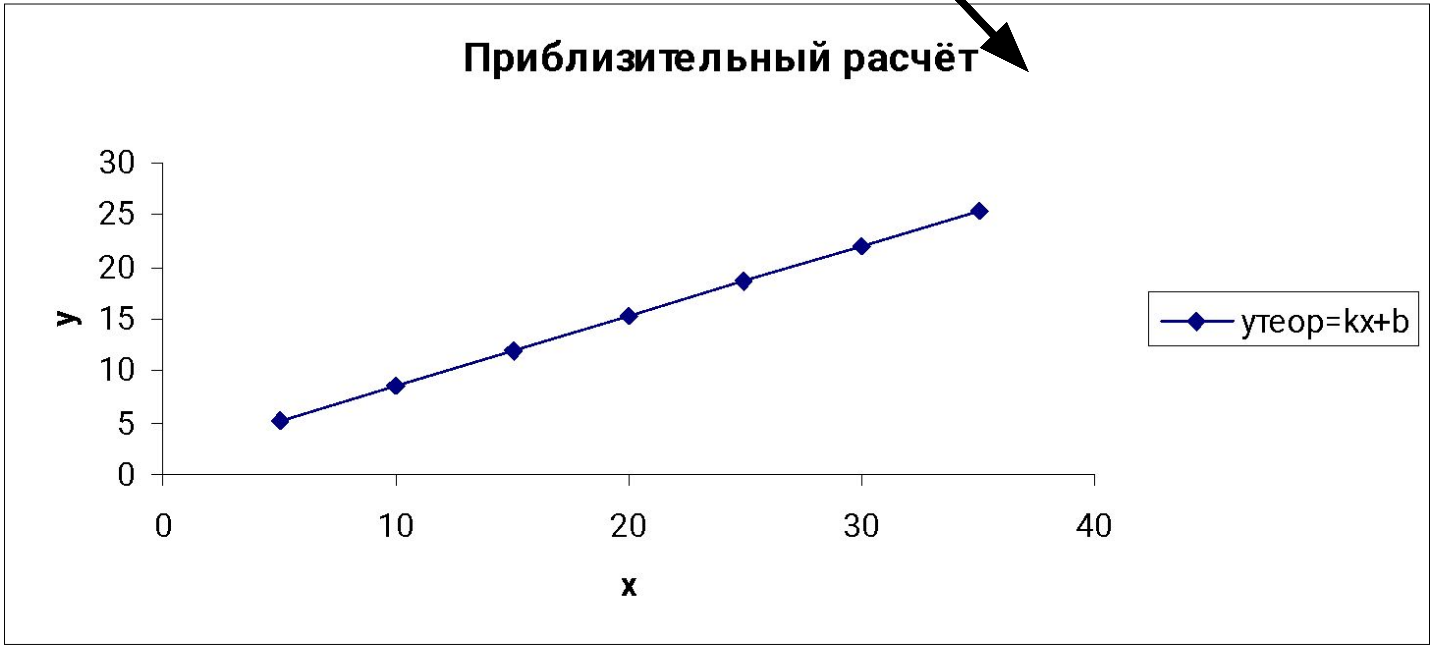
ОСНОВНОЙ ПРИНЦИП МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

- При замене точного (неизвестного) параметра модели приблизительным значением необходимо минимизировать разницу между экспериментальными данными и теоретическими (вычисленными при помощи предложенной модели).

ПРИБЛИЖЕНИЕ

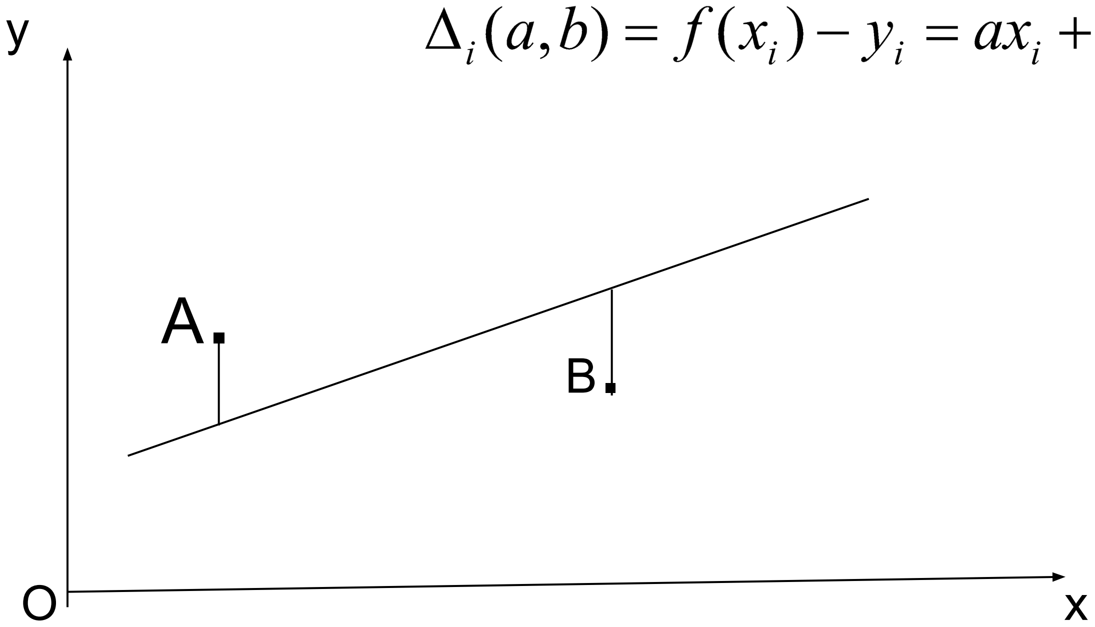


МНК



ОТКЛОНЕНИЕ ТОЧКИ ОТ ПРЯМОЙ

$$\Delta_i(a,b) = f(x_i) - y_i = ax_i + b - y_i$$



КАК УЧЕСТЬ ОТКЛОНЕНИЕ ВСЕХ ТОЧЕК?

- В рамках метода наименьших квадратов минимизируется величина:

$$S(a,b)=\sum_{i=1}^{n}(ax_i+b-y_i)^2$$

Суммарное отклонение всех точек

МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

- Пусть нам известно оптимальное значение *a*. Тогда S зависит только от b. Для того, чтобы найти минимум, надо приравнять производную к нулю.

$$S'_b = \sum_{i=1}^{n} 2\cdot\Delta_i\cdot(\Delta_i)'_b = \sum_{i=1}^{n} 2\cdot(ax_i + b - y_i)\cdot 1 = 2\cdot(a\sum_{i=1}^{n} x_i + bn - \sum_{i=1}^{n} y_i) = 0$$

$$b = \frac{\sum_{i=1}^{n} y_i}{n} - a\cdot\frac{\sum_{i=1}^{n} x_i}{n} = \overline{y} - a\overline{x}$$

ИТОГИ

- Вычисление коэффициентов прямой по формулам:

$$a = \frac{\sum_{i=1}^{n} [(y_i - \overline{y})(x_i - \overline{x})]}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

$$\overline{b} = \overline{y} - a\overline{x}$$

МЕТОД НАИМЕНЬШИХ КВАДРАТОВ В MICROSOFT EXCEL 2003

- По формулам.
- Функция ЛИНЕЙН

x_i	y_i
5	5,05
10	8,62
15	11,59
20	15,45
25	18,41
30	22,04
35	25,00

$\xrightarrow{\text{ЛИНЕЙН}}$ a = 0,664357

\+

$\xrightarrow{\text{Ctrl+Shift+Enter}}$ b = 1,878571

ПРОВЕРКА АДЕКВАТНОСТИ УРАВНЕНИЯ РЕГРЕССИИ

Компоненты дисперсии	Сумма квадратов	Число степеней свободы	Дисперсия на одну степень свободы
Общая	$\sum(y-\bar{y})^2$	$n-1$	$S^2_{общ}=\dfrac{\sum(y-\bar{y})^2}{n-1}$
Факторная	$\sum(\hat{y}_x-\bar{y})^2$	m	$S^2_{факт}=\dfrac{\sum(\hat{y}_x-\bar{y})^2}{m}$
Остаточная	$\sum(y-\hat{y}_x)^2$	$n-m-1$	$S^2_{ост}=\dfrac{\sum(y-\hat{y}_x)^2}{n-m-1}$

(n –число наблюдений, m–число параметров при переменной x)

ПРОВЕРКА АДЕКВАТНОСТИ УРАВНЕНИЯ РЕГРЕССИИ

Определение дисперсии на одну **степень свободы** приводит дисперсии к сравнимому виду. Сопоставляя факторную и остаточную дисперсии в расчете на одну степень свободы, получим величину **F-критерия Фишера**. Фактическое значение F -критерия Фишера сравнивается с табличным значением $F_{табл.}$ (α, k_1, k_2) при заданном уровне значимости α и степенях свободы k_1= m и k_2=n-m-1. При этом, если фактическое значение F-критерия больше табличного **$F_{факт}$ > $F_{теор}$**, то признается статистическая значимость уравнения в целом. Для парной линейной регрессии m=1 , поэтому:

$$F = \frac{S^2_{\text{факт}}}{S^2_{\text{ост}}} = \frac{\sum\left(\hat{y}_x - \bar{y}\right)^2}{\sum\left(y - \hat{y}_x\right)^2} \cdot (n-2)$$

ВЫВОДЫ:

- Метод наименьших квадратов, а также его различные модификации широко используется при анализе экспериментальных данных.
- В рамках метода наименьших квадратов минимизируется величина сумма квадратов отклонений действительных (экспериментальных) значений от теоретических.