Метод наименьших квадратов.

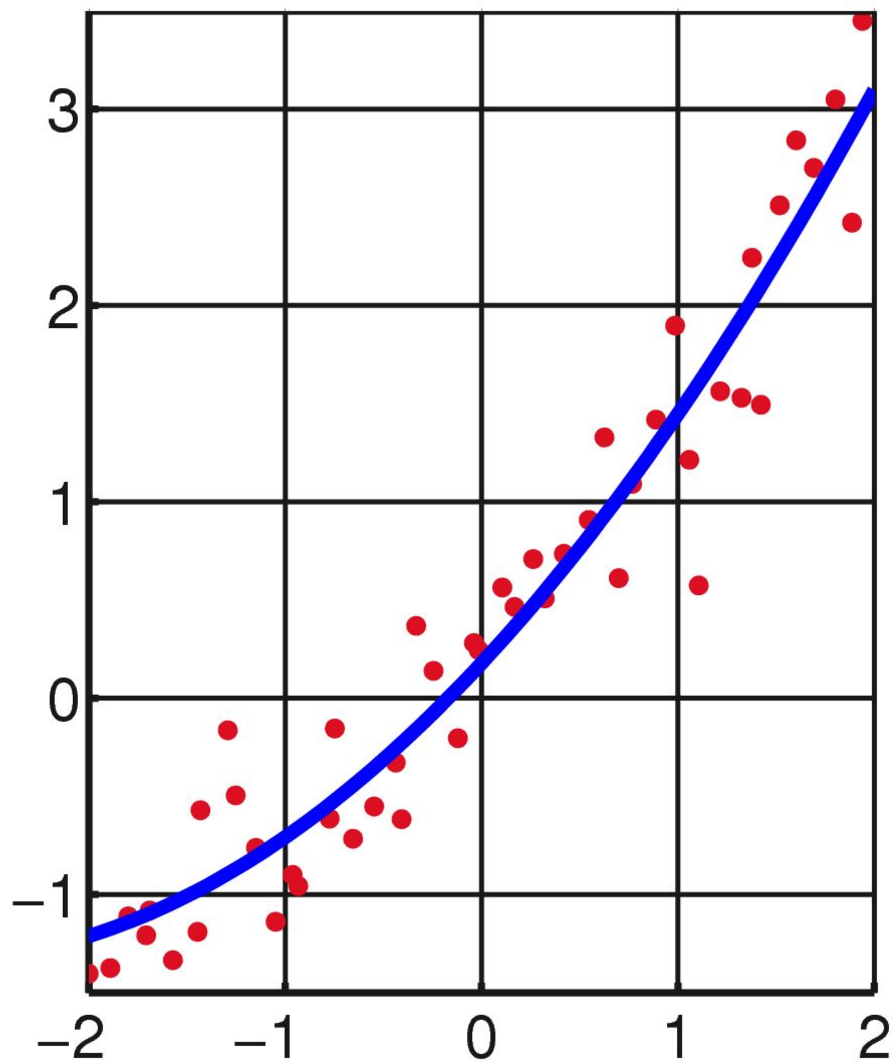
Выполнила

Студентка 2 курса, группа с-15-16

Иняткина Анна

- Метод наименьших квадратов- математический метод, применяемый для решения различных задач, основанный на минимизации суммы квадратов отклонений некоторых функций от искомых переменных.

- МНК является одним из базовых методов регрессионного анализа для оценки неизвестных параметров регрессионных моделей по выборочным данным.
- Применяется также для приближённого представления заданной функции другими (более простыми) функциями и оказывается полезным при обработке наблюдений.


3
2
1
0
–1
–2
–1
0
1
2

Плюсы и минусы данного метода.

\+ он приводит к сравнительно простому математическому способу определения параметров a, b, c, … искомого функционала;

\+ он дает довольно веское теоретическое обоснование с вероятностной точки зрения.

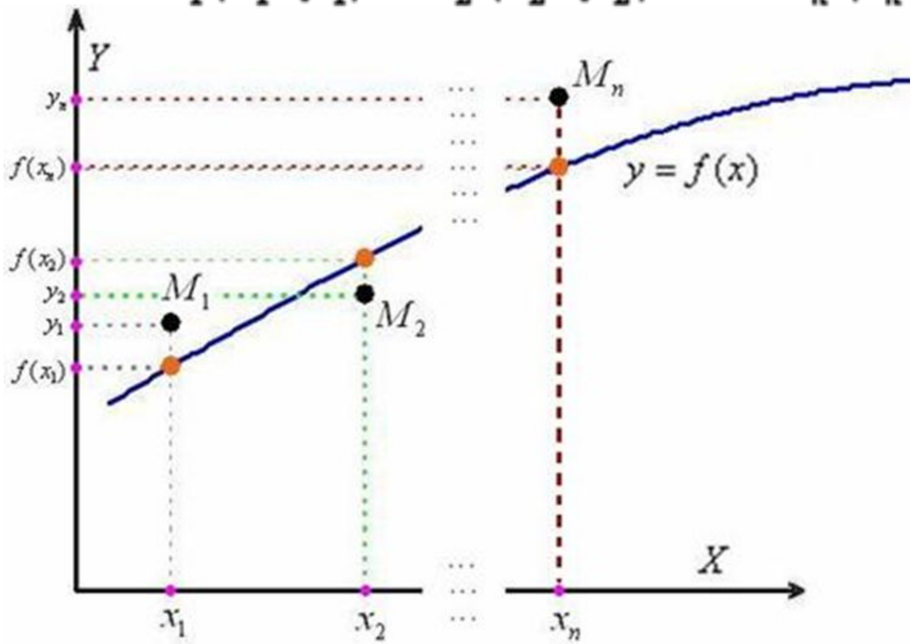
\- основным недостатком МНК является чувствительность оценок к резким выбросам, которые встречаются в исходных данных.

X	x_1	x_2	$\cdots$	x_n
Y	y_1	y_2	$\cdots$	y_n

- нам нужно подобрать функцию , $y = f(x)$ график которой проходит как можно ближе к точкам. Такую функцию называют ***аппроксимирующей*** *(аппроксимация – приближение)* или ***теоретической функцией***.
-

Пусть некоторая функция приближает экспериментальные данные

: $M_1(x_1; y_1), M_2(x_2; y_2), \ldots, M_n(x_n; y_n)$



$$f(x_1), f(x_2), \ldots, f(x_n)$$

$$e_1 = y_1 - f(x_1), e_2 = y_2 - f(x_2), \ldots, e_n = y_n - f(x_n)$$

$$e_1^2 + e_2^2 + \ldots + e_n^2$$

$$\sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} \left(y_i - f\left(x_i\right)\right)^2$$

Иными словами, **задача состоит в нахождении таких коэффициентов a и b – чтобы сумма квадратов отклонений**

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

была наименьшей.

$$y = f(x) = ax + b$$

$$y = f(x) = \frac{a}{x} + b$$

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

$$F(a\,;b) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \left(\frac{a}{x_i} + b \right) \right)^2$$

Вывод формул для нахождения коэффициентов.

$$\frac{\partial F}{\partial a} = \left(\sum_{i=1}^{n} (y_i - (ax_i + b))^2 \right)'_a = \sum_{i=1}^{n} \left[2(y_i - (ax_i + b)) \cdot (y_i - (ax_i + b))'_a \right] =$$

$$= 2\sum_{i=1}^{n} \left[(y_i - ax_i - b) \cdot (0 - (x_i + 0)) \right] = 2\sum_{i=1}^{n} \left[(y_i - ax_i - b) \cdot (-x_i) \right] = 2\sum_{i=1}^{n} (ax_i^2 + bx_i - x_i y_i)$$

$$\frac{\partial F}{\partial b} = \left(\sum_{i=1}^{n} (y_i - (ax_i + b))^2 \right)'_b = \sum_{i=1}^{n} \left[2(y_i - (ax_i + b)) \cdot (y_i - (ax_i + b))'_b \right] =$$

$$= 2\sum_{i=1}^{n} \left[(y_i - ax_i - b) \cdot (0 - (0 + 1)) \right] = 2\sum_{i=1}^{n} (ax_i + b - y_i)$$

Из необходимых условий экстремума следует:

$$\begin{cases} \dfrac{\partial F}{\partial a} = 0 \\ \dfrac{\partial F}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} 2\sum\limits_{i=1}^{n}(ax_i^2 + bx_i - x_i y_i) = 0 \\ 2\sum\limits_{i=1}^{n}(ax_i + b - y_i) = 0 \end{cases}$$

$$\begin{cases} a\sum\limits_{i=1}^{n} x_i^2 + b\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i \\ a\sum\limits_{i=1}^{n} x_i + bn = \sum\limits_{i=1}^{n} y_i \end{cases}$$

Алгоритм МНК:

1) Находим суммы $\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i, \sum_{i=1}^{n} x_i^2, \sum_{i=1}^{n} x_i y_i$
2) Составляем систему уравнений с определённым кол-вом неизвестных.
3) Решаем систему. (метод Крамера)
4) В результате получаем стационарную точку S (искомые коэффициенты a;b)
5) Вычисляем сумму квадратов отклонений между эмпирическими и теоретическими значениями.

$$\sum e_i^2 = \sum (y_i - f(x_i))^2$$

Пример:

x_i	1	2	3	4	5
y_i	5,3	6,3	4,8	3,8	3,3

- Методом наименьших квадратов найти линейную функцию, которая наилучшим образом приближает эмпирические (опытные) данные.

- Коэффициенты a, b оптимальной функции y=ax+b найдём как решение системы:

$$\begin{cases} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i \end{cases}$$

x_i	1	2	3	4	5	$\sum x_i =$	15
y_i	5,3	6,3	4,8	3,8	3,3	$\sum y_i =$	23,5
x_i^2	1	4	9	16	25	$\sum x_i^2 =$	55
$x_i y_i$	5,3	12,6	14,4	15,2	16,5	$\sum x_i y_i =$	64

- Таким образом, получаем следующую систему:

$$\begin{cases} 55a + 15b = 64 \\ 15a + 5b = 23,5 \end{cases}$$

$$\Delta = \begin{vmatrix} 55 & 15 \\ 15 & 5 \end{vmatrix} = 55 \cdot 5 - 15 \cdot 15 = 275 - 225 = 50 \neq 0$$

$$\Delta_a = \begin{vmatrix} 64 & 15 \\ 23,5 & 5 \end{vmatrix} = 64 \cdot 5 - 23,5 \cdot 15 = 320 - 352,5 = -32,5$$

$$a = \frac{\Delta_a}{\Delta} = \frac{-32,5}{50} = -0,65$$

$$\Delta_b = \begin{vmatrix} 55 & 64 \\ 15 & 23,5 \end{vmatrix} = 55 \cdot 23,5 - 15 \cdot 64 = 1292,5 - 960 = 332,5$$

$$b = \frac{\Delta_b}{\Delta} = \frac{332,5}{50} = 6,65$$

$$55 \cdot (-0,65) + 15 \cdot 6,65 = -35,75 + 99,75 = 64$$

$$15 \cdot (-0,65) + 5 \cdot 6,65 = -9,75 + 33,25 = 23,5$$

Таким образом, искомая аппроксимирующая функция:

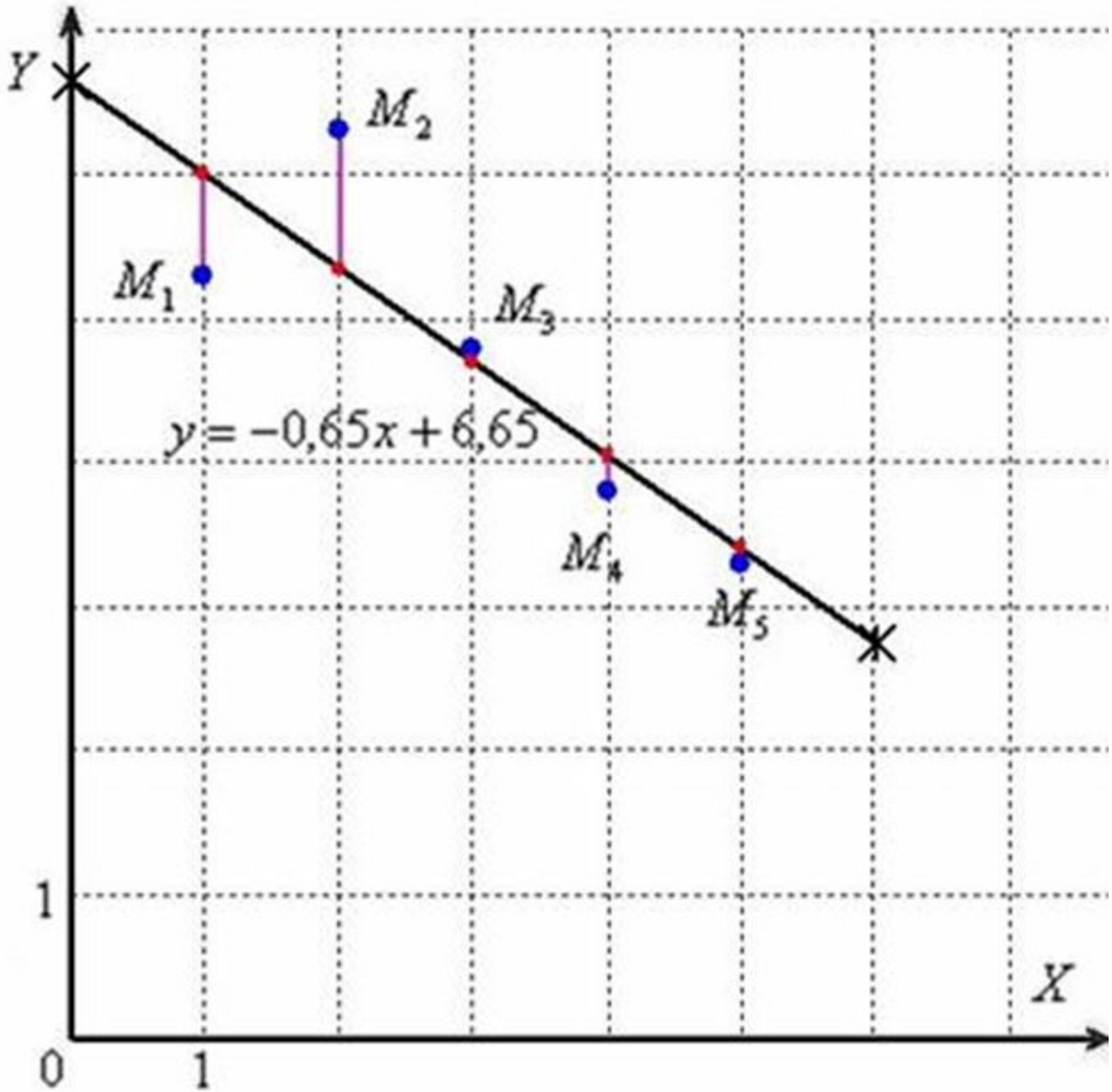
$$y = f(x) = -0,65x + 6,65$$

Для построения графика аппроксимирующей функции найдём два её значения:

$$f(0) = -0,65 \cdot 0 + 6,65 = 6,65$$

$$f(6) = -0,65 \cdot 6 + 6,65 = -3,9 + 6,65 = 2,75$$

и выполним чертёж:



Y
M_2
M_1
M_3
$y = -0,65x + 6,65$
M_4
M_5
1
X
0
1

$$\sum e_i^2 = \sum (y_i - f(x_i))^2$$

x_i	1	2	3	4	5		
y_i	5,3	6,3	4,8	3,8	3,3		
$f(x_i)$	6	5,35	4,7	4,05	3,4		
$(y_i - f(x_i))^2$	0,49	0,9025	0,01	0,0625	0,01	$\sum e_i^2 =$	1,475

$$f(x_1) = f(1) = -0,65 \cdot 1 + 6,65 = 6$$

$$(y_1 - f(x_1))^2 = (5,3 - 6)^2 = (-0,7)^2 = 0,49$$

СПАСИБО ЗА ВНИМАНИЕ