

**Метод наименьших квадратов
оценки параметров
функциональной зависимости**

Выполнила: Гнездилова Алена

Группа: С-1841

Что такое МНК?

- **Метод наименьших квадратов (МНК)** — математический метод, применяемый для решения различных задач, основанный на минимизации суммы квадратов отклонений некоторых функций от искомым переменных.

Почему МНК?

- Название свое метод наименьших квадратов получил, исходя из основного принципа, которому должны удовлетворять полученные на его основе оценки параметров: **сумма квадратов ошибки модели должна быть минимальной.**

- $$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2 \rightarrow \min$$

Общий вид модели

- Модель в общем виде может быть представлена уравнением:

$$y_t = a_0 + a_1 x_{1t} + \dots + a_n x_{nt} + \varepsilon_t .$$

- Исходными данными при оценке параметров a_0, a_1, \dots, a_n является вектор значений зависимой переменной $y = (y_1, y_2, \dots, y_T)'$ и матрица значений независимых переменных

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{n1} \\ 1 & x_{12} & x_{22} & \dots & x_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1T} & x_{2T} & \dots & x_{nT} \end{bmatrix}$$

- в которой первый столбец, состоящий из единиц, соответствует коэффициенту модели .

Применение

Метод наименьших квадратов используется для оценки параметров уравнение регрессии.

Регрессионный анализ представляет собой вывод уравнения регрессии, с помощью которого находится средняя величина случайной переменной (признака-результата), если величина другой (или других) переменных (признаков-факторов) известна. Он включает следующие этапы:

- выбор формы связи (вида аналитического уравнения регрессии);
- оценку параметров уравнения;
- оценку качества аналитического уравнения регрессии.

Пример использования

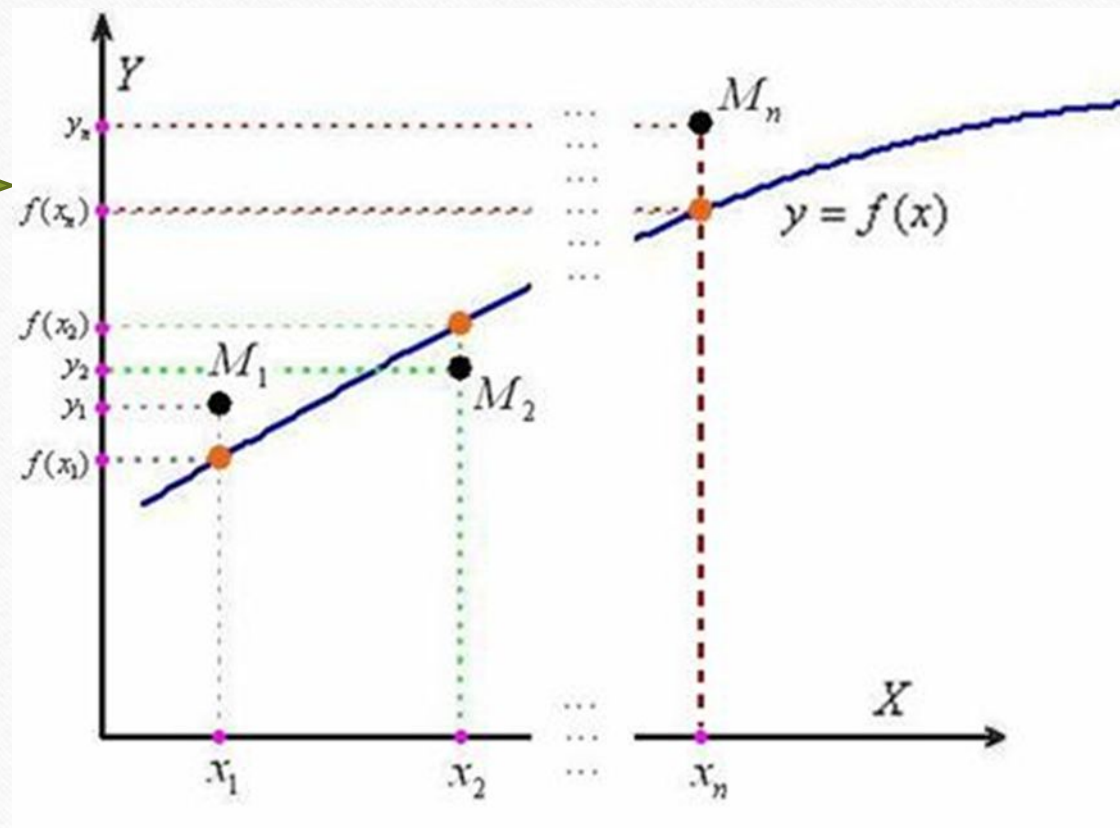
- X – торговую площадь продовольственного магазина, кв.м.,
- Y – годовой товароборот продовольственного магазина, млн. руб.

Предположим, что после проведения наблюдений/опытов/подсчётов в нашем распоряжении оказываются числовые данные:

X				...	
Y				...	

Такую функцию называют *аппроксимирующей* (аппроксимация — приближение) или *теоретической функцией*.

Табличные данные также можно записать в виде точек $M_1(x_1; y_1)$; $M_2(x_2; y_2); \dots; M_n(x_n; y_n)$ и изобразить в декартовой системе ХОУ.



Как оценить точность данного приближения?

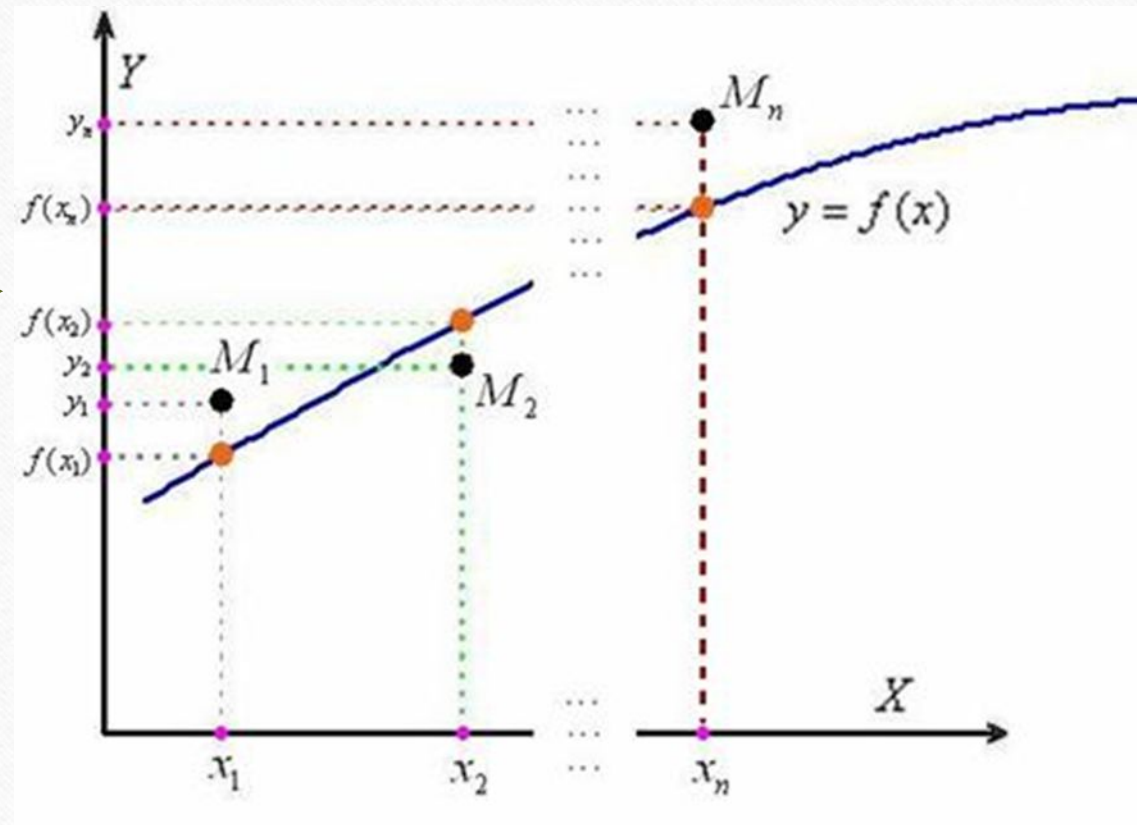


В качестве оценки точности приближения принимаем сумму квадратов отклонений:

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

Из чего следует:

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

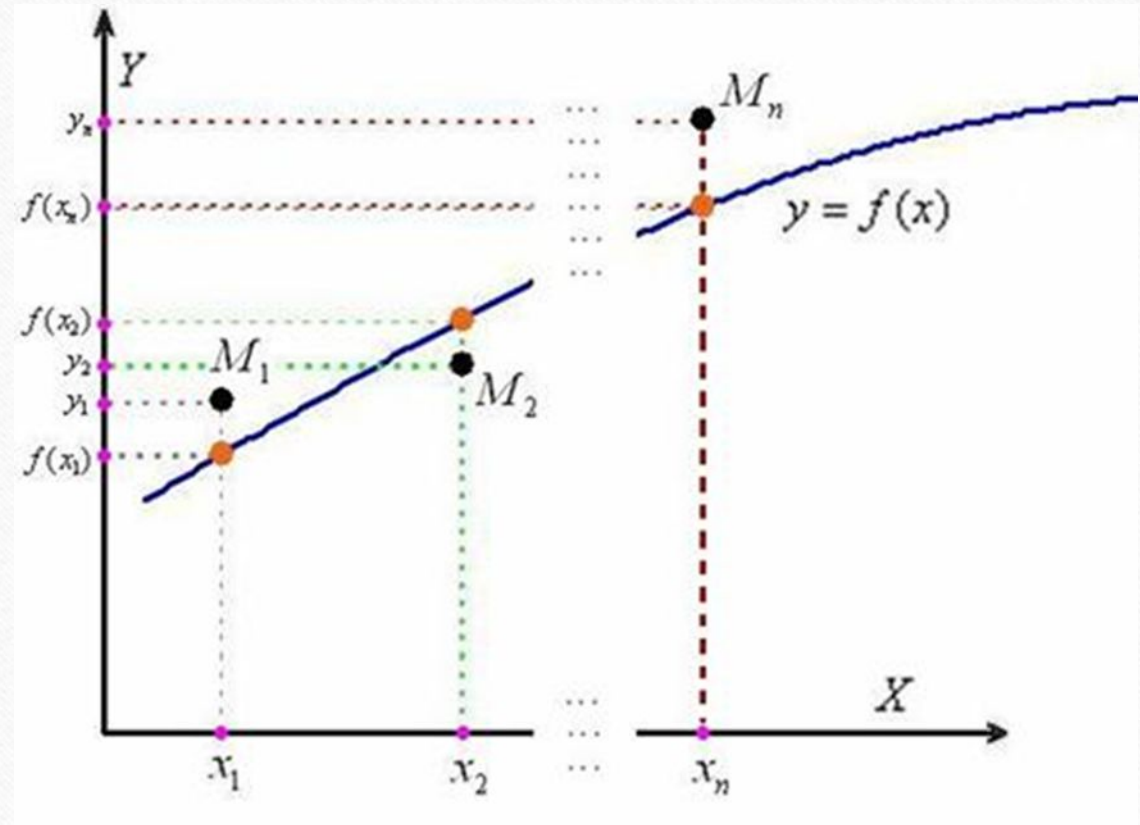


Какой класс функций выбрать для исследования?

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

Если точки M_1, M_2, \dots, M_n имеют тенденцию располагаться по прямой, то следует искать уравнение прямой $y=f(x) = ax + b$ с оптимальными значениями a и b .

Задача найти ТАКИЕ коэффициенты a, b – чтобы сумма квадратов отклонений была наименьшей.

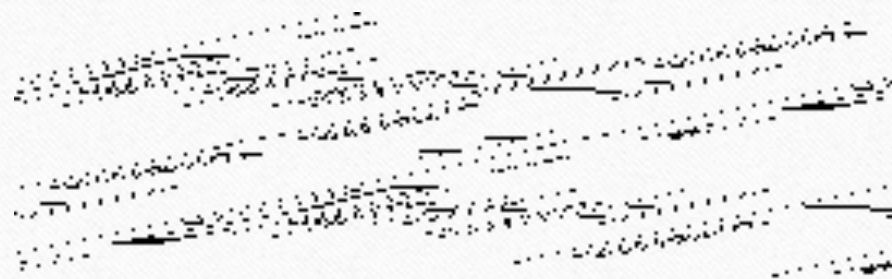


Сначала найдем частные производные 1-го порядка.

Согласно правилу линейности дифференцировать можно прямо под значком суммы:

The image shows a handwritten mathematical derivation. It consists of two main parts, each starting with a dot. The first part shows a sum of terms being differentiated term-by-term. The second part shows a similar sum of terms, also being differentiated term-by-term. The handwriting is somewhat messy and includes some underlines and arrows indicating the differentiation process.

Составим стандартную систему:



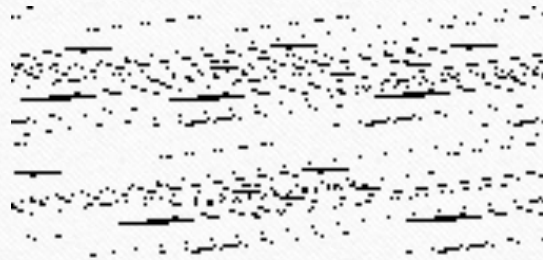
Сокращаем каждое уравнение на «двойку» и, кроме того, «разваливаем» суммы:



Примерчик:

$$= b * (1 + 1 + \dots + 1) = bn$$
$$\sum_{i=1}^n \sum_{j=1}^n 1$$

Перепишем систему в «прикладном» виде:



После чего начинает прорисовываться алгоритм решения нашей задачи:

1) Составляем простейшую систему двух линейных уравнений с двумя неизвестными («а» и «бэ»). Систему решаем, в результате чего получаем стационарную точку 

2) Проверив достаточное условие экстремума, можно убедиться, что в данной точке функция достигает именно минимума.

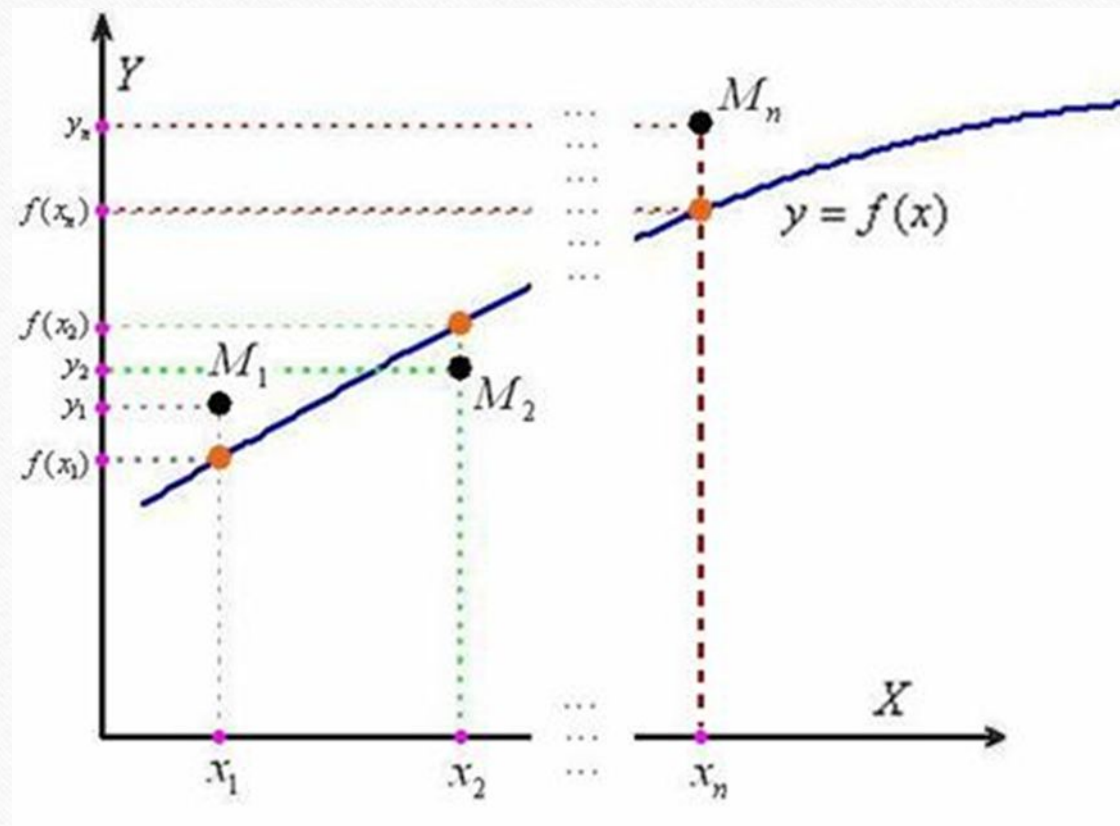
Делаем
ОКОНЧАТЕЛЬНЫЙ
ВЫВОД:

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

наилучшим образом приближает
экспериментальные точки

$$M_1; M_2; \dots; M_n.$$

Её график проходит максимально
близко к этим точкам.



Задача 1

- В результате исследования взаимосвязи двух показателей, получены следующие пары чисел:

X	1	2	3	4	5
Y	5,3	6,3	4,8	3,8	3,3

Методом наименьших квадратов найти линейную функцию, которая наилучшим образом приближает эмпирические (опытные) данные. Сделать чертеж, на котором в декартовой прямоугольной системе координат построить экспериментальные точки $M_i(X_i; Y_i)$ и график аппроксимирующей функции $y=f(x) = ax + b$. Найти сумму квадратов отклонений между эмпирическими y_i и теоретическими $f(x_i)$ значениями.