



**Федеральное государственное автономное образовательное учреждение  
высшего образования «Национальный исследовательский университет  
ИТМО»**

**Факультет программной инженерии и компьютерной техники**

**Вычислительная математика**

**Лабораторная работа №1**

**Аппроксимация по методу наименьших модулей**

**Выполнила:**

**Гафурова Фарангиз Фуркатовна**

**Группа Р3220**

**Преподаватель:**

**Перл Ольга Вячеславовна**

**г. Санкт-Петербург**

**2025**

## Задача

Title: Аппроксимация по методу наименьших модулей Description:

[ДИСКЛЕЙМЕР: Это не всё, что вам нужно сделать для сдачи лабораторной работы по курсу, но этого достаточно, чтобы получить вариант на следующую работу.] Дан набор точек, по которым необходимо построить линейную аппроксимацию по методу наименьших модулей. Необходимо найти значение модуля наибольшего отклонения среди заданных точек относительно полученной линейной аппроксимации. [Обратите внимание, что при подготовке отчёта по лабораторной работе, желательно иметь несколько аппроксимирующих функций, кроме линейной.] Формат входных данных:  $x_1$   $x_2$   $x_3$  ...  $y_1$   $y_2$   $y_3$  ... где  $x_1 \dots x_n$  - список значений аргумента для узлов интерполяции,  $y_1 \dots y_n$  - список значений функции для соответствующего значения аргумента для узлов интерполяции. В тестах также вначале задаётся количество задаваемых точек, однако, в функцию этот параметр не передаётся. Формат выходных значений: вещественное число, являющееся значением модуля наибольшего отклонения исходных данных от полученной линейной аппроксимации.

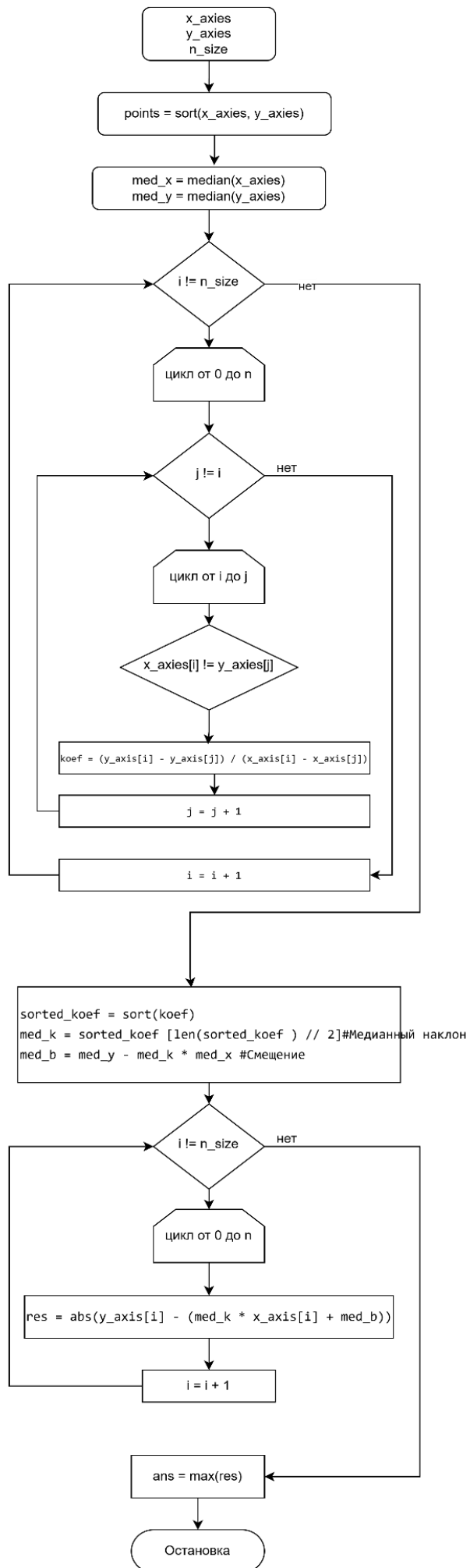
## Описание Метода

Метод наименьших модулей предназначен для построения линейной аппроксимации набора точек. В отличие от метода наименьших квадратов, он минимизирует наибольшее отклонение, обеспечивая большую устойчивость к выбросам.

Процедура метода:

1. Упорядочить точки по значениям  $x$ .
2. Найти медианные значения  $x$  и  $y$ .
3. Вычислить медианный наклон  $k = \text{median}(y_i - y_j / x_i - x_j) \mid i \neq j$ .
4. Определить свободный коэффициент  $b = y_m - kx_m$ .
5. Найти максимальное отклонение:

# Блок-схема



## Код численного метода (Python)

```
#
# Complete the 'approximate_linear_least_modules' function below.
#
# The function is expected to return a DOUBLE.
# The function accepts following parameters:
# 1. DOUBLE_ARRAY x_axis
# 2. DOUBLE_ARRAY y_axis
#

def med(ar):
    return ar[len(ar) // 2]

def koefNak(x_axis, y_axis, n_size):
    return [(y_axis[i] - y_axis[j]) / (x_axis[i] - x_axis[j])
            for i in range(n_size) for j in range(i) if x_axis[i] !=
x_axis[j]]

def approximate_linear_least_modules(x_axis, y_axis):
    n_size = len(x_axis)

    points = sorted(zip(x_axis, y_axis))
    x_axis, y_axis = zip(*points)

    med_x = med(x_axis)
    med_y = med(y_axis)

    slopes = koefNak(x_axis, y_axis, n_size)

    slopes.sort()
    med_k = slopes[len(slopes) // 2] # Медианный наклон
    med_b = med_y - med_k * med_x # Смещение

    # Вычисляем отклонения
    res = [abs(y_axis[i] - (med_k * x_axis[i] + med_b)) for i in
range(n_size)]

    return max(res)
```

## Примеры работы программы

y	res
1 2 3 4 5	0
-1 -0.5 0 0.5 1	0
3 3.1 2.9 3.2 3.0	0.2
2 4 6 8 10	0
2 3 4 5 6	-
1 1 1 1 1 10 1 1 1 1	9
10 -10 10 -10 10	10
2 5 -1 6 3	3

## Корректность сделанных выводов

- Метод позволяет строить линейную аппроксимацию, минимизируя максимальное отклонение.
- В сравнении с методом наименьших квадратов, он менее чувствителен к выбросам.
- Метод применим к любому набору точек, если значения различны.
- Алгоритмическая сложность: из-за вычисления всех парных наклонов.
- Численные ошибки минимальны, так как используется медианная оценка вместо средних значений.

## Вывод

В ходе лабораторной работы мной был исследован метод наименьших модулей для линейной аппроксимации данных. Данный подход проявил свою надежность в условиях наличия выбросов, что делает его особенно привлекательным для анализа реалистичных данных. Однако стоит учитывать, что выполнение этого метода требует квадратичного времени из-за необходимости сортировки и вычисления медианных значений. Это может ограничить его применение при работе с большими объемами данных. В связи с этим, важно учитывать баланс между устойчивостью метода к аномальным значениям и его вычислительной сложностью.