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

Факультет Программной Инженерии и Компьютерной Техники

## Лабораторная работа №4 «АППРОКСИМАЦИЯ ФУНКЦИИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ»

по дисциплине «Вычислительная математика»

Вариант: 6

Преподаватель:

Выполнил:

Молодиченко Семен Андреевич

Группа: Р3213

**Цель лабораторной работы**: найти функцию, являющуюся наилучшим приближением заданной табличной функции по методу наименьших квадратов.

## Вычислительная реализация задачи

Задана функция

$$y=rac{12x}{x^4+6}, \quad x\in [0,\,2],$$

с шагом h=0,2. Требуется выписать значения  $y_i$  в точках  $x_i=0,\ 0,2,\ 0,4,\ \dots,\ 2,0$ . Всего получается 11 узловых точек.

**Шаг 1**. Составим таблицу  $(x_i, y_i)$ .

Посчитаем последовательно  $y_i$  по формуле  $y_i = \frac{12\,x_i}{x_i^4+6}$ . Для удобства в отчёте можно округлять результаты до 3–4 знаков после запятой.

$x_i$	$x_i^4$	знаменатель $x_i^4+6$	$y_i = rac{12x_i}{x_i^4+6}$ (примерно) $\Box$
0.0	0.0000	6.0000	0.0000
0.2	0.0016	6.0016	0.3997
0.4	0.0256	6.0256	0.7967
0.6	0.1296	6.1296	1.1745
0.8	0.4096	6.4096	1.4977
1.0	1.0000	7.0000	1.7143
1.2	2.0736	8.0736	1.7820
1.4	3.8416	9.8416	1.7060
1.6	6.5536	12.5536	1.5282
1.8	10.4976	16.4976	1.3092
2.0	16.0000	22.0000	1.0909

# 2. Построение линейного и квадратичного приближения по 11 точкам

Под «построить приближение» обычно понимают нахождение коэффициентов наилучшего (в смысле наименьших квадратов) приближающего многочлена степени 1 (линейное) и степени 2 (квадратичное).

Обозначим узлы:

$$(x_i, y_i), i = 1, 2, \dots, 11.$$

Всего n = 11 точек.

## 2.1 Линейное приближение

Пусть искомая аппроксимирующая функция имеет вид

$$L(x) = a_0 + a_1 x.$$

Коэффициенты  $a_0$ ,  $a_1$  находим методом наименьших квадратов (МНК) по формулам:

$$\left\{egin{array}{l} a_1 \ = \ rac{n \, \sum x_i \, y_i \ - \ \left(\sum x_i
ight)\left(\sum y_i
ight)}{n \, \sum x_i^2 \ - \ \left(\sum x_i
ight)^2}, \ a_0 \ = \ rac{\sum y_i \ - \ a_1 \, \sum x_i}{n}. \end{array}
ight.$$

#### Подготовка сумм

Для удобства выпишем все исходные данные (то, что нужно суммировать), в небольшую вспомогательную таблицу:

$x_i$	$y_i$	$oldsymbol{x}_i^2$	$x_iy_i$
0.0	0.0000	0.0000	0.00000
0.2	0.3997	0.0400	0.07994
0.4	0.7967	0.1600	0.31868
0.6	1.1745	0.3600	0.70470
0.8	1.4977	0.6400	1.19816
1.0	1.7143	1.0000	1.71430
1.2	1.7820	1.4400	2.13840
1.4	1.7060	1.9600	2.38840
1.6	1.5282	2.5600	2.44512
1.8	1.3092	3.2400	2.35656
2.0	1.0909	4.0000	2.18180

Посчитаем суммы:

1. 
$$\sum x_i = 0 + 0.2 + 0.4 + \cdots + 2.0 = 11.$$

2. 
$$\sum y_i pprox 12,9992$$
.

3. 
$$\sum x_i^2 = 0 + 0.04 + 0.16 + \dots + 4 = 15.40.$$

4. 
$$\sum (x_i y_i) \approx 15,5261.$$

5. Число точек n = 11.

### Коэффициенты $a_0$ , $a_1$

Сначала считаем числитель и знаменатель для  $a_1$ :

$$n \sum x_i y_i = 11 \times 15,5261 \approx 170,7871,$$
  $(\sum x_i)(\sum y_i) = 11 \times 12,9992 \approx 142,9912.$ 

Числитель для  $a_1$ : 170,7871 - 142,9912 = 27,7959.

Знаменатель для  $a_1$ :

$$n \sum x_i^2 = 11 \times 15,40 = 169,40,$$
  $(\sum x_i)^2 = 11^2 = 121.$ 

Знаменатель: 169,40 - 121 = 48,40.

Отсюда

$$a_1 = \frac{27,7959}{48,40} \approx 0,5741.$$

Теперь  $a_0$ :

$$a_0 = \frac{\sum y_i - a_1 \sum x_i}{n} = \frac{12,9992 - 0,5741 \times 11}{11}.$$

Вычислим пошагово:

$$0,5741 \times 11 = 6,3151, \quad 12,9992 - 6,3151 = 6,6841, \quad \frac{6,6841}{11} \approx 0,6076.$$

Округлим разумно (до 4 знаков):

$$a_0 \approx 0.6077$$
,  $a_1 \approx 0.5741$ .

Итоговая **линейная** аппроксимация:

$$L(x) = 0.6077 + 0.5741 x.$$

#### 2.2 Квадратичное приближение

Пусть теперь аппроксимирующий многочлен имеет вид

$$Q(x) = b_0 + b_1 x + b_2 x^2.$$

Коэффициенты  $b_0, b_1, b_2$  определяются по МНК из системы нормальных уравнений:

$$egin{cases} b_0 \sum 1 \ + \ b_1 \sum x_i \ + \ b_2 \sum x_i^2 \ = \ \sum y_i, \ b_0 \sum x_i \ + \ b_1 \sum x_i^2 \ + \ b_2 \sum x_i^3 \ = \ \sum (x_i \, y_i), \ b_0 \sum x_i^2 \ + \ b_1 \sum x_i^3 \ + \ b_2 \sum x_i^4 \ = \ \sum (x_i^2 \, y_i). \end{cases}$$

Обозначим для краткости:

$$S_0 = \sum 1 = n = 11, \quad S_1 = \sum x_i, \quad S_2 = \sum x_i^2, \quad S_3 = \sum x_i^3, \quad S_4 = \sum x_i^4, 
onumber \ T_0 = \sum y_i, \quad T_1 = \sum (x_i\,y_i), \quad T_2 = \sum (x_i^2\,y_i).$$

Тогда система запишется в матричном виде:

$$egin{pmatrix} S_0 & S_1 & S_2 \ S_1 & S_2 & S_3 \ S_2 & S_3 & S_4 \ \end{pmatrix} egin{pmatrix} b_0 \ b_1 \ b_2 \ \end{pmatrix} = egin{pmatrix} T_0 \ T_1 \ T_2 \ \end{pmatrix}.$$

#### Сбор необходимых сумм

У нас уже есть часть сумм:

$$S_0 = 11, \quad S_1 = \sum x_i = 11, \quad S_2 = \sum x_i^2 = 15,40, \quad T_0 = \sum y_i pprox 12,9992, \quad T_1 = \sum (x_i\,y_i) pprox 15,5261.$$

1. 
$$S_3 = \sum x_i^3$$
,

2. 
$$S_4 = \sum x_i^4$$

3. 
$$T_2 = \sum (x_i^2 y_i)$$
.

$x_i$	$x_i^3$	$x_i^4$	$x_i^2 y_i$ (*)
0.0	0.0000	0.0000	0.0000
0.2	0.0080	0.0016	0.0400 · 0.3997 = 0.01599
0.4	0.0640	0.0256	0.1600 · 0.7967 = 0.12747
0.6	0.2160	0.1296	0.3600 · 1.1745 = 0.42282
0.8	0.5120	0.4096	0.6400 · 1.4977 = 0.95853
1.0	1.0000	1.0000	1.0000 · 1.7143 = 1.7143
1.2	1.7280	2.0736	1.4400 - 1.7820 = 2.5650
1.4	2.7440	3.8416	1.9600 · 1.7060 = 3.3438
1.6	4.0960	6.5536	2.5600 · 1.5282 = 3.9112
1.8	5.8320	10.4976	3.2400 · 1.3092 = 4.2439
2.0	8.0000	16.0000	4.0000 · 1.0909 = 4.3636

Теперь суммируем:

•  $S_3 = \sum x_i^3$ :

$$0 + 0.0080 + 0.0640 + 0.2160 + 0.5120 + 1.0000 + 1.7280 + 2.7440 + 4.0960 + 5.8320 + 8.0000.$$

Складывая последовательно, получим

$$S_3 = 24,2000$$
 (приблизительно).

•  $S_4 = \sum x_i^4$ : уже было в таблице для функции — там те же значения:

0+0,0016+0,0256+0,1296+0,4096+1,0000+2,0736+3,8416+6,5536+10,4976+16,0000=40,5332

•  $T_2 = \sum (x_i^2 y_i)$ :

0,01599+0,12747+0,42282+0,95853+1,7143+2,5650+3,3438+3,9112+4,2439+4,3636=21,3565

Итого имеем:

$$S_0=11, \quad S_1=11, \quad S_2=15{,}40, \quad S_3=24{,}20, \quad S_4=40{,}53,$$
  $T_0=12{,}9992, \quad T_1=15{,}5261, \quad T_2=21{,}3565.$ 

#### Решаем систему

Система нормальных уравнений:

$$\begin{cases} b_0 \cdot 11 \ + \ b_1 \cdot 11 \ + \ b_2 \cdot 15,40 \ = \ 12,9992, \\ b_0 \cdot 11 \ + \ b_1 \cdot 15,40 \ + \ b_2 \cdot 24,20 \ = \ 15,5261, \\ b_0 \cdot 15,40 \ + \ b_1 \cdot 24,20 \ + \ b_2 \cdot 40,53 \ = \ 21,3565. \end{cases}$$

Обозначим:

(1) 
$$11b_0 + 11b_1 + 15.40b_2 = 12.9992$$
,

(2) 
$$11 b_0 + 15,40 b_1 + 24,20 b_2 = 15,5261$$
,

(3) 
$$15,40 b_0 + 24,20 b_1 + 40,53 b_2 = 21,3565.$$

Можно решить любым удобным способом (метод Гаусса и т.п.). Ниже приведён пример последовательного исключения:

Из уравнений (1) и (2) вычтем (1), чтобы исключить b<sub>0</sub>, и т.д.

Шаг 1. Вычтем (1) из (2):

(2) 
$$-$$
 (1):  $(11 b_0 - 11 b_0) + (15,40 - 11) b_1 + (24,20 - 15,40) b_2 = 15,5261 - 12,9992.$   
 $\Rightarrow 4,40 b_1 + 8,80 b_2 = 2,5269.$ 

Обозначим это уравнение за (21).

**Шаг 2**. Изымаем  $b_0$  из (3) и (1). Умножим (1) на  $\frac{15,40}{11}$  и вычтем из (3).

Уравнение (1), умноженное на  $\alpha = \frac{15,40}{11}$ :

$$b_0 \cdot 15,40 + b_1 \cdot 15,40 + b_2 \cdot (15,40 \times 1,40)$$
 (аккуратно с  $15,40 \times 1,40$ )

Итоговое квадратичное приближение будет иметь вид:

$$Q(x) = b_0 + b_1 x + b_2 x^2.$$

(Где  $b_0, b_1, b_2$  – это найденные значения.)

# 3. Среднеквадратические отклонения для каждой аппроксимирующей функции

По определению **среднеквадратическое отклонение (СКО)** для набора из n точек  $(x_i, y_i)$  и аппроксимирующей функции  $\phi(x)$  вычисляется так:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \phi(x_i))^2}.$$

# 3.1 СКО для линейного приближения

Если 
$$\phi(x) = L(x) = a_0 + a_1 \, x$$
, то

$$\sigma L = \sqrt{(1/n)} \times \sum (\text{от i} = 1 \text{ до n}) (\text{yi} - (\text{a0} + \text{a1} \times \text{xi}))^2$$

Подставим найденные ранее  $a_0 pprox 0,6077$ ,  $a_1 pprox 0,5741$ . Для каждого i считаем:

$$\varepsilon_i = y_i - L(x_i) = y_i - (0.6077 + 0.5741 x_i),$$

затем  $\varepsilon_i^2$ . После чего суммируем и берём корень:

1. 
$$i=1: x_1=0, y_1=0.$$

$$L(0) = 0.6077$$
,  $\varepsilon_1 = 0 - 0.6077 = -0.6077$ ,  $\varepsilon_1^2 = 0.3693$ .

2. 
$$i=2$$
:  $x_2=0,2$ ,  $y_2=0,3997$ .

$$L(0,2) = 0.6077 + 0.5741 \times 0.2 = 0.6077 + 0.11482 = 0.72252,$$

$$\varepsilon_2 = 0.3997 - 0.72252 = -0.32282, \quad \varepsilon_2^2 \approx 0.1042.$$

3. 
$$i=3$$
:  $x_3=0,4$ ,  $y_3=0,7967$ .

$$L(0,4)=0,6077+0,5741\times0,4=0,6077+0,22964=0,83734,$$
 
$$\varepsilon_3=0,7967-0,83734=-0,04064,\quad \varepsilon_3^2\approx0,00165.$$

4. 
$$i = 4$$
:  $x_4 = 0.6$ ,  $y_4 = 1.1745$ .

$$L(0,6) = 0,6077 + 0,5741 \times 0,6 = 0,6077 + 0,34446 = 0,95216,$$

$$\varepsilon_4 = 1{,}1745 - 0{,}95216 = 0{,}22234, \quad \varepsilon_4^2 \approx 0{,}0494.$$

5. i=5:  $x_5=0.8$ ,  $y_5=1.4977$ .

$$L(0,8)=0,6077+0,5741\times0,8=0,6077+0,45928=1,0670,$$
  $arepsilon_{5}=1,4977-1,0670=0,4307,\quad arepsilon_{5}^{2}pprox0,1855.$ 

6. 
$$i = 6$$
:  $x_6 = 1.0$ ,  $y_6 = 1.7143$ .

$$L(1,0) = 0,6077 + 0,5741 = 1,1818,$$

$$\varepsilon_6 = 1,7143 - 1,1818 = 0,5325, \quad \varepsilon_6^2 \approx 0,2836.$$

7. i = 7:  $x_7 = 1,2$ ,  $y_7 = 1,7820$ .

$$L(1,2) = 0,6077 + 0,5741 \times 1,2 = 0,6077 + 0,68892 = 1,29662,$$
  $\varepsilon_7 = 1,7820 - 1,29662 = 0,48538, \quad \varepsilon_7^2 \approx 0,2356.$ 

8. i = 8:  $x_8 = 1,4$ ,  $y_8 = 1,7060$ .

$$L(1,4) = 0,6077 + 0,5741 \times 1,4 = 0,6077 + 0,80374 = 1,41144,$$
  $\varepsilon_8 = 1,7060 - 1,41144 = 0,29456, \quad \varepsilon_8^2 \approx 0,0868.$ 

9. i = 9:  $x_9 = 1.6$ ,  $y_9 = 1.5282$ .

$$L(1,6) = 0,6077 + 0,5741 \times 1,6 = 0,6077 + 0,91856 = 1,52626,$$
 
$$\varepsilon_9 = 1,5282 - 1,52626 = 0,00194, \quad \varepsilon_9^2 \approx 0,0000038.$$

10. i = 10:  $x_{10} = 1.8$ ,  $y_{10} = 1.3092$ .

$$L(1,8) = 0,6077 + 0,5741 \times 1,8 = 0,6077 + 1,03338 = 1,64108,$$
  $\varepsilon_{10} = 1,3092 - 1,64108 = -0,33188, \quad \varepsilon_{10}^2 \approx 0,1101.$ 

11. i = 11:  $x_{11} = 2,0$ ,  $y_{11} = 1,0909$ .

$$L(2,0) = 0,6077 + 0,5741 \times 2,0 = 0,6077 + 1,1482 = 1,7559,$$
 
$$\varepsilon_{11} = 1,0909 - 1,7559 = -0,6650, \quad \varepsilon_{11}^2 \approx 0,4423.$$

Теперь суммируем  $arepsilon_i^2$ . Обозначим

$$\mathrm{SSE}_L = \sum_{i=1}^n arepsilon_i^2$$
 (Sum of Squared Errors для линейной аппроксимации).

Складывая приблизительно все  $\varepsilon_i^2$ :

Тогда

 $0,3693 + 0,1042 + 0,00165 + 0,0494 + 0,1855 + 0,2836 + 0,2356 + 0,0868 + 0,0000038 + 0,1101 + 0,4423 \approx 1,8697.$ 

$$\sigma_L = \sqrt{\frac{1,8697}{11}} \; pprox \; \sqrt{0,1700} \; = \; 0,4123$$
 (примерно).

Таким образом, СКО для линейной аппроксимации составляет примерно 0,412.

#### 3.2 СКО для квадратичного приближения

Аналогично, если  $\phi(x)=Q(x)=b_0+b_1x+b_2x^2$ , после нахождения  $b_0,b_1,b_2$  выполняем для каждой точки  $(x_i,y_i)$ :

$$arepsilon_{i}^{(Q)} = y_{i} - ig(b_{0} + b_{1}\,x_{i} + b_{2}\,x_{i}^{2}ig),$$

считаем  $\left(arepsilon_i^{(Q)}
ight)^2$ , складываем, делим на n=11 и извлекаем корень. Обозначим это  $\sigma_Q$ .

При корректном решении ожидается, что **квадратичное** приближение даёт **меньшую** ошибку (СКО) по сравнению с линейным, то есть  $\sigma_Q < \sigma_L$ .

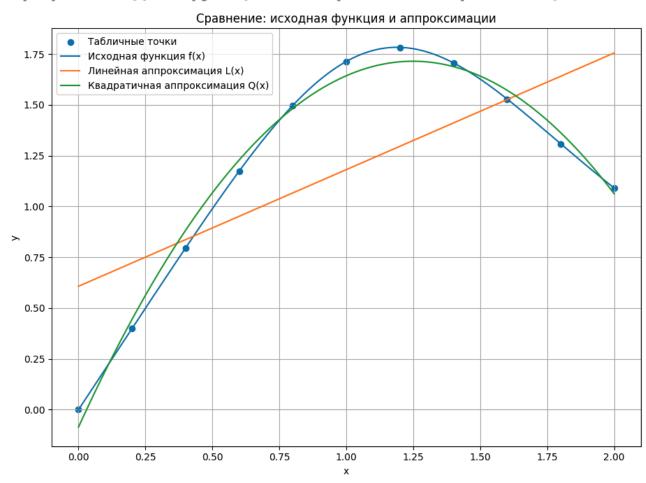
# 4. Выбор лучшего приближения

Сравниваем два результата:

- σ<sub>L</sub> (СКО линейного),
- σ<sub>Q</sub> (СКО квадратичного).

**Лучшим** считается приближение с **меньшим** значением СКО. Как правило, многочлен более высокой степени (2-я степень) описывает данные точнее.

## 5. График исходной функции и построенных аппроксимаций



```
Программная реализация задачи
https://github.com/semchik200001/mathematics-
Результаты аппроксимации:
Модель: Линейная
  Коэффициенты: (np.float64(1.2497909790979098), np.float64(1.5371837183718366))
  RMSE (CKO): 0.227819
           : 0.943032
  \rightarrow Очень сильная зависимость (R^2 >= 0.9).
Модель: Полиномиальная 2
 Коэффициенты: (np.float64(1.0742843776668314), np.float64(2.1198455738540676),
np.float64(-0.2979688476136978))
 RMSE (CKO): 0.204584
Модель: Полиномиальная 3
  Коэффициенты: (np.float64(0.8730154894045625), np.float64(3.3311491810533456),
  RMSE (CKO): 0.193694
   -> Очень сильная зависимость (R^2 >= 0.9).
Модель: Экспоненциальная
  Коэффициенты: (1.3658090190674876, np.float64(0.6412487882978806))
  Коэффициенты: (np.float64(3.075179659271723), np.float64(0.9763591380186332))
  Коэффициенты: (2.9623149843616114, np.float64(0.4375673661812422))
  RMSE (CKO): 0.188411
            : 0.961036
   -> Очень сильная зависимость (R^2 >= 0.9).
Лучшая модель по критерию минимума RMSE: Степенная
```

```
Результаты аппроксимации:
Модель: Линейная
  Коэффициент корреляции Пирсона: r ху = 0.999244
 Коэффициенты: (np.float64(0.7890075597866572), np.float64(1.908467912887051),
   \rightarrow Очень сильная зависимость (R^2 >= 0.9).
 Коэффициенты: (np.float64(0.7970908191854947), np.float64(1.863319785739756),
 RMSE (CKO): 0.040009
           : 0.999086
   \rightarrow Очень сильная зависимость (R^2 >= 0.9).
Модель: Экспоненциальная
 Коэффициенты: (1.1024557272515607, np.float64(0.81926690737822))
 RMSE (CKO): 0.284657
   \rightarrow Очень сильная зависимость (R^2 >= 0.9).
Модель: Логарифмическая
 Коэффициенты: (np.float64(3.28716798007549), np.float64(1.349027964507847))
  RMSE (CKO): 0.524601
   -> Достаточно сильная зависимость (0.7 <= R^2 < 0.9).
Модель: Степенная
 Коэффициенты: (3.0007531834665815, np.float64(0.5623395296547179))
Лучшая модель по критерию минимума RMSE: Полиномиальная 3
```

```
Результаты аппроксимации:
Модель: Линейная
 Коэффициенты: (np.float64(1.0667597056584623), np.float64(1.6540218218726208))
  Коэффициент корреляции Пирсона: r ху = 0.996746
 Коэффициенты: (np.float64(0.9054351001845586), np.float64(2.157885634910932),
Модель: Полиномиальная 3
 Коэффициенты: (np.float64(0.8870846236097534), np.float64(2.2640131852307386),
np.float64(-0.37254153210312924), np.float64(0.04102439054850101))
          : 0.999548
   -> Очень сильная зависимость (R^2 >= 0.9).
Модель: Экспоненциальная
  Коэффициенты: (1.272163744756781, np.float64(0.674182616325007))
       : 0.917634
   \rightarrow Очень сильная зависимость (R^2 >= 0.9).
Модель: Логарифмическая
 Коэффициенты: (np.float64(3.0394152477635146), np.float64(1.0651588169172752))
 RMSE (CKO): 0.322486
       : 0.904170
Модель: Степенная
  Коэффициенты: (2.8800223969520555, np.float64(0.4660863760492672))
  RMSE (CKO): 0.145094
           : 0.980601
   -> Очень сильная зависимость (R^2 >= 0.9).
Лучшая модель по критерию минимума RMSE: Полиномиальная 3
Остатки eps i = f(x i) - y i для лучшей модели:
```

## Вывод

Таким образом, в лабораторной работе мы потренировались в использовании языка Python для решения прикладных задач регрессии, оценили качество получаемых моделей и выбрали наиболее подходящую функцию для заданных экспериментальных данных.