

3.6. Тренды

Тренд – это основная тенденция изменения чего-либо: например, в математике – временного ряда. Тренды могут быть описаны различными уравнениями: линейными, логарифмическими, степенными и так далее.

Рассмотрим пример. На рисунке 29 показано распределение температур по месяцам в некотором городе за 2021 год.

Для данных температур необходимо подобрать математическое описание (функцию) так, чтобы эти точки были как можно ближе к линии, построенной по выбранной функции. Эта линия называется линией **тренда**. Для наших точек линию тренда можно описать с помощью полинома второй степени:

$$f(x) = -0.78 \cdot x^2 + 9.03 \cdot x + 0.95.$$

Например, возьмем месяц май, который на графике имеет индекс 4, то есть функция тренда вычисляется для этого значения ($x = 4$):

$$f(4) = 24.5733.$$

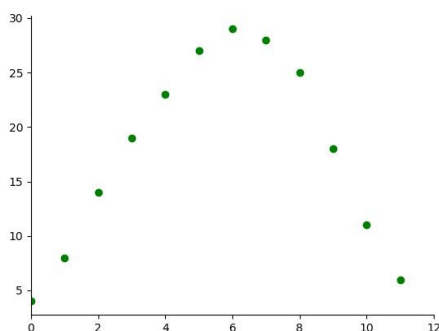


Рисунок 29. График распределения температур по месяцам

А точное значение температуры в мае – 23 градуса. Ошибка составляет 1.6 градус. Можно посчитать погрешность вычисления по формуле:

$$\delta = \left| \frac{\text{точное_значение} - \text{приближенное}}{\text{точное_значение}} \right| \cdot 100\%.$$

Для температуры в мае погрешность составит 6.84%. Если посчитать значения для других месяцев, погрешность вычисления будет другой.

Добавив линию тренда на график, можно проанализировать, насколько тренд хорошо отражает тенденцию изменения температур (рисунок 30):

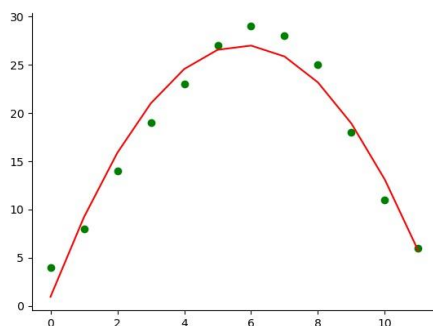


Рисунок 30. График распределения температур и тренд

В качестве функций при построении тренда можно использовать линейную функцию:

$$f(x) = a_0 \cdot x + a_1$$

и полиномиальную:

$$f(x) = a_0 \cdot x^n + a_1 \cdot x^{n-1} + \dots + a_{n-1} \cdot x + a_n.$$

Линейную функцию можно рассматривать как частный случай полиномиальной, степень полинома $n = 1$.

В модуле **numpy** для построения линейного и полиномиального тренда используется функция:

```
polyfit(x_array, y_array, deg)
```

где `x_array`, `y_array` – массивы координат точек по оси ОХ и ОУ должны иметь одинаковый размер;

`deg` – степень полинома (линейная функция – 1, квадратичная функция – 2 и т.д.)

Результатом функции является массив, в котором перечислены значения коэффициентов полинома $a_0, a_1, \dots, a_n$.

Пример. Вычислить коэффициенты линии тренда для температур, вычислить погрешность в каждой точке. Погрешность округлить до 1 знака после запятой.

Код программы:

```
import numpy as np

month = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11])
temperature = np.array([4, 8, 14, 19, 23, 27, 29, 28, 25, 18, 11, 6])

k_poly = np.polyfit(month, temperature, 2)
print(k_poly)
```

Результат:

```
[-0.78171828  9.03246753  0.95054945]
```

Чтобы посчитать значение тренда в некоторой точке, необходимо создать функцию, параметрами которой будут искомая точка и коэффициенты полинома:

```
def get_trend(x, a):
    y = a[0] * x **2 + a[1]* x + a[2]
    return y
```

Теперь можно вычислить погрешность построения тренда для каждой точки:

```
import numpy as np

def get_trend(x, a):
    y = a[0] * x **2 + a[1]* x + a[2]
    return y

month = np.array([0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11])
temperature = np.array([4, 8, 14, 19, 23, 27, 29, 28, 25, 18, 11, 6])

k_poly = np.polyfit(month, temperature, 2)
print(k_poly)
delta = np.abs((temperature - get_trend(month, k_poly))/temperature)*100
print("погрешность:", np.round(delta, 1))
```

Результат:

```
[-0.78171828  9.03246753  0.95054945]
погрешность: [76.2 15.  13.5 10.6  6.8  1.6  6.9  7.6  7.3  5.1 19.1  4.7]
```

Задание для самопроверки

Получить коэффициенты линий тренда 3 и 4 степени для задачи с температурами, вычислить погрешность тренда в каждой точке для полиномов 3 и 4 степени. Вывести средние значения погрешностей для полиномов 3 и четвертой степени.

Результат:

```
n = 3 [-0.03703704 -0.17060717  6.45839346  2.78388278]
средняя погрешность: 11.3
n = 4 [ 0.01343969 -0.33271011  1.86646513  1.93881928  4.30448718]
средняя погрешность: 4.33
```