

Санкт-Петербургский государственный электротехнический университет им. В.И. Ульянова (Ленина)

АППРОКСИМАЦИЯ ВРЕМЕННЫХ РЯДОВ С ПОМОЩЬЮ В-СПЛАЙНОВ

Выполнил: Осипов В.Г., Гр. 3304

Руководитель: д.т.н., проф. Середа В. И.

Цель и задачи

Цель: разработка программного модуля, реализующего методы аппроксимации данных.

Исследуется проблема сглаживания данных.

Актуальность

- Автоматизация обработки данных с помощью В-сплайнов.

Задачи:

- формирование списка существующих методов
- описание методов аппроксимации;
- программная реализация описанных методов;
- постановка вычислительного эксперимента.

Используемая теория

Данные y_i на сетке x_i с постоянным шагом $h = t_i - t_{i-1}$

$$y_i \approx c_0 \phi_0(t_i) + c_1 \phi_1(t_i) + \dots + c_n \phi_n(t_i), \quad i = 0, \dots, m$$

ϕ_i - базисные функции, c_i - коэффициенты, t_i - аргумент (отсчет).

В матричном виде

$$\begin{pmatrix} \phi_0(t_0) & \dots & \phi_n(t_0) \\ \vdots & \ddots & \vdots \\ \phi_0(t_m) & \dots & \phi_n(t_m) \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_n \end{pmatrix} \approx \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix}$$
$$Ac \approx y \text{ или } Ac - y \approx 0$$

Метод наименьших квадратов

$$\min_c \sum_{i=0}^m [(y - Ac)_i]^2$$

Методы аппроксимации

- Кусочно-кубический сплайн

$$S(t) = \sum_{i=1}^n S_i(t),$$

где $S_i(t) = a_i + b_i(t - t_i) + c_i(t - t_i)^2 + d_i(t - t_i)^3$.

- $S_i(t)$ определен на отрезке $t \in [t_i, t_{i+1}]$.
- Необходимо задавать условия на равенство производных в узлах.

Методы аппроксимации (окончание)

- В-сплайны

$$\text{При } k = 0: N_i^0(t) = \begin{cases} 1, & t \in [t_i, t_{i+1}], \\ 0, & t \notin [t_i, t_{i+1}]. \end{cases}$$

$$\text{При } k \geq 1: N_i^k(t) = \frac{t - t_i}{t_{i+k} - t_i} N_i^{k-1}(t) + \frac{t_{i+k+1} - t}{t_{i+k+1} - t_{i+1}} N_{i+1}^k(t)$$

- Сплайн k -й степени

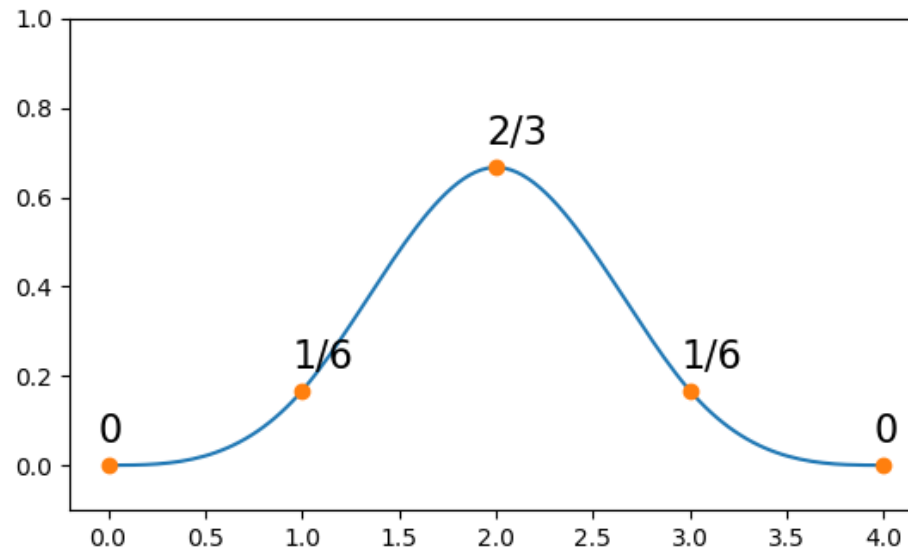
$$S^k(t) = \sum_i \alpha_i N_i^k(t)$$

В-сплайны

Формула В-сплайна 3-й степени

$$N_0^3(t) = \begin{cases} \frac{t^3}{6}, & t \in [0, 1) \\ \frac{4-12t+3t^2-12t^3}{6}, & t \in [1, 2) \\ \frac{-44+60t-24t^2+3t^3}{6}, & t \in [2, 3) \\ \frac{(4-t)^3}{6}, & t \in [3, 4) \\ 0 & \text{Иначе} \end{cases}$$

Рисунок 1 – В-сплайн 3-й степени
на равномерной сетке



В-сплайны (продолжение)

Для определения базиса необходимо дополнить основную сетку дополнительными узлами.

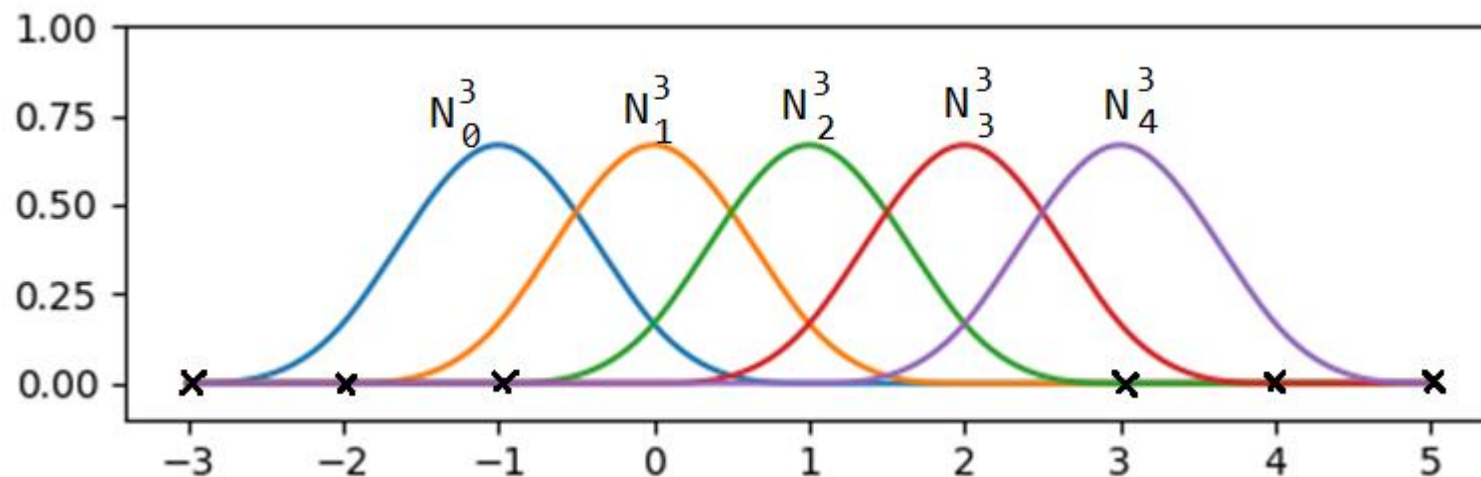


Рисунок 2 – базис В-сплайнов 3-й степени на сетке $x_i = i$, $i = 0, 1, 2$.
Дополнительные узлы помечены крестом.

В-сплайны (окончание)

Решаемая система уравнений $(A^T A)\alpha = A^T y$, $A^{n \times m}, \alpha^{m \times 1}, y^{n \times 1}$.

Шаг сетки сплайна может быть больше, чем шаг сетки данных ($n \geq m$).

Рисунок 3 — схематичное изображение получаемой матрицы $A^T A$ при аппроксимации n точек кубическими В-сплайнами.

$$\begin{pmatrix} e_1 & f_1 & d_1 & c_1 & & & & 0 \\ f_1 & e_2 & f_2 & d_2 & c_2 & & & \\ d_1 & f_2 & e_3 & f_3 & d_3 & . & & \\ c_1 & d_2 & f_3 & e_4 & . & . & . & \\ & c_2 & . & . & . & . & . & c_{m-3} \\ & & . & . & . & . & f_{m-2} & d_{m-2} \\ & & & . & . & f_{m-2} & e_{m-1} & f_{m-1} \\ 0 & & & & c_{m-3} & d_{m-2} & f_{m-1} & e_m \end{pmatrix}$$

Программная реализация

В состав программы, представляющий Python 3-пакет, входят следующие модули:

- графический интерфейс пользователя;
- модуль интерполяции и расчета интерполированных значений;
- модуль аппроксимации.

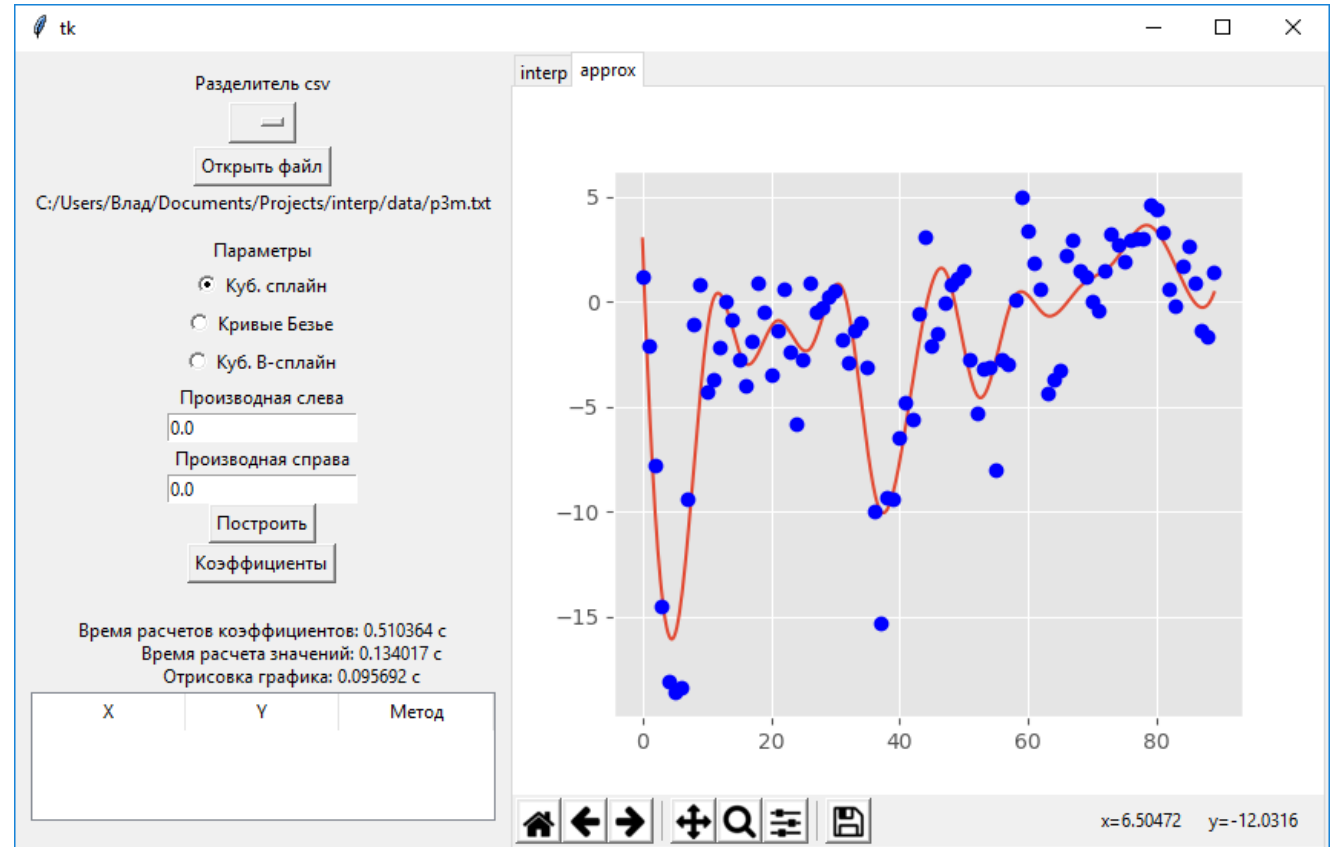


Рисунок 4 – Графический интерфейс

Модули

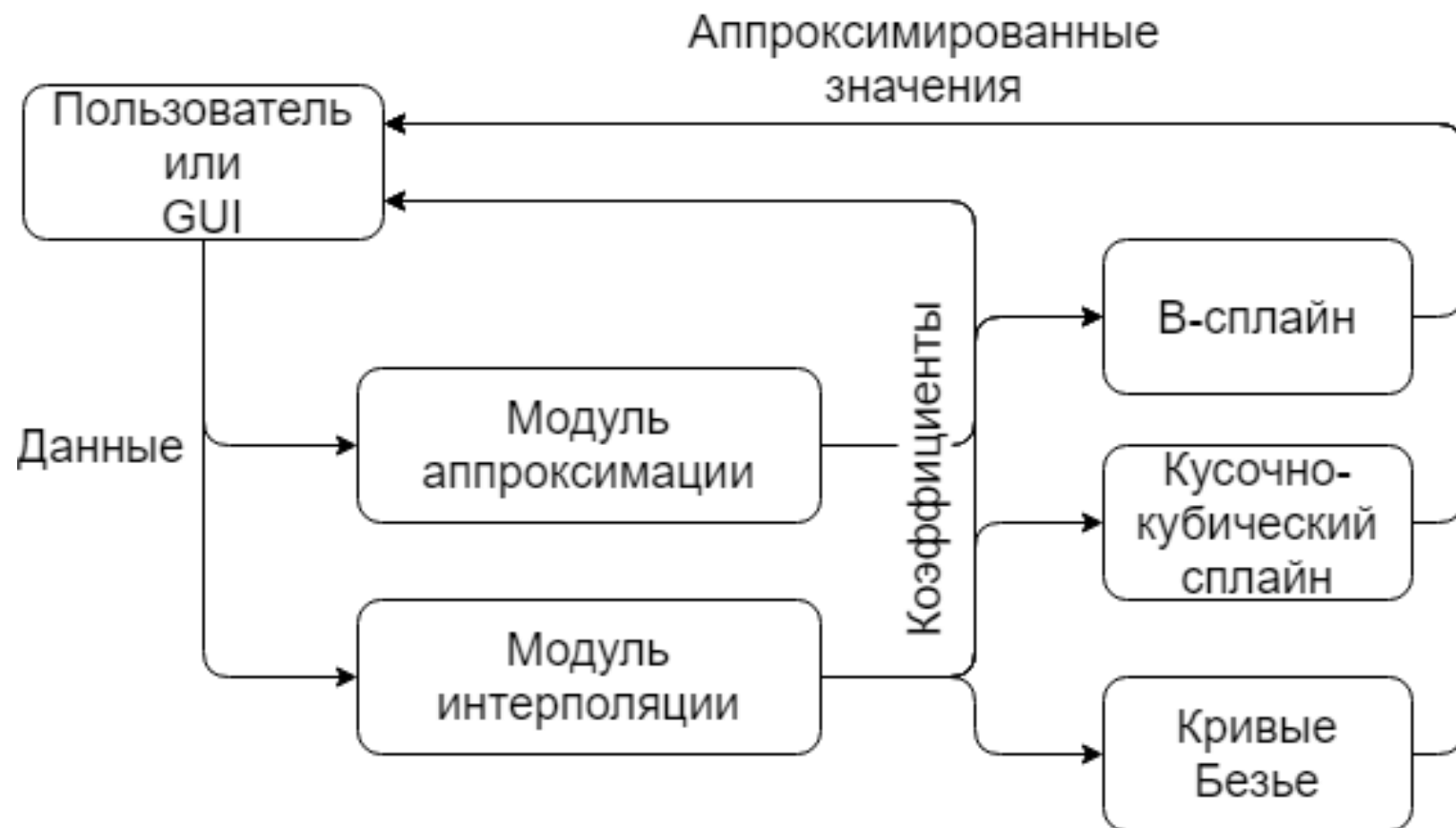


Рисунок 5 – связь модулей программы по данным

Эксперимент

Таблица 1 – Время вычисления коэффициентов при аппроксимации данных методом наименьших квадратов В-сплайнами.

Число заданных точек	С учетом структуры, мс	Без учета структуры, мс
600	33,1	283,9
800	34,7	316,3
1200	38,5	425,6

Эксперимент (продолжение)

Степень сглаживания в зависимости от числа узлов сплайна

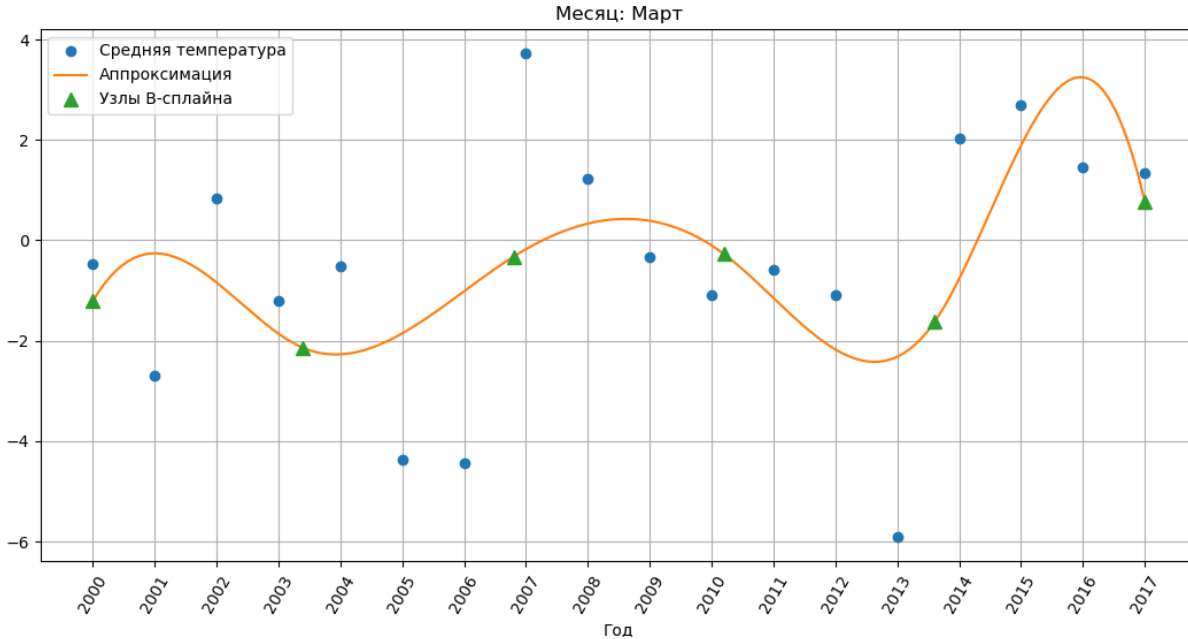


Рисунок 6 – 6 узлов сплайна при 18 заданных точках

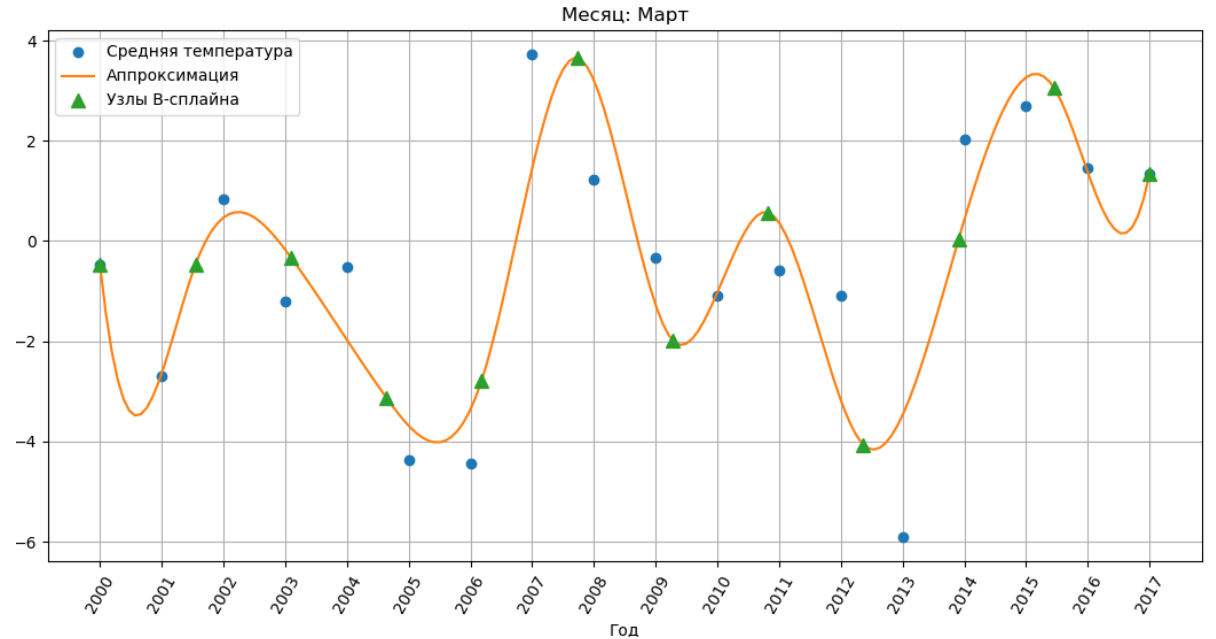


Рисунок 7 – 12 узлов при 18 заданных точках

Эксперимент (окончание)

Есть тенденция на возрастание среднегодовой температуры

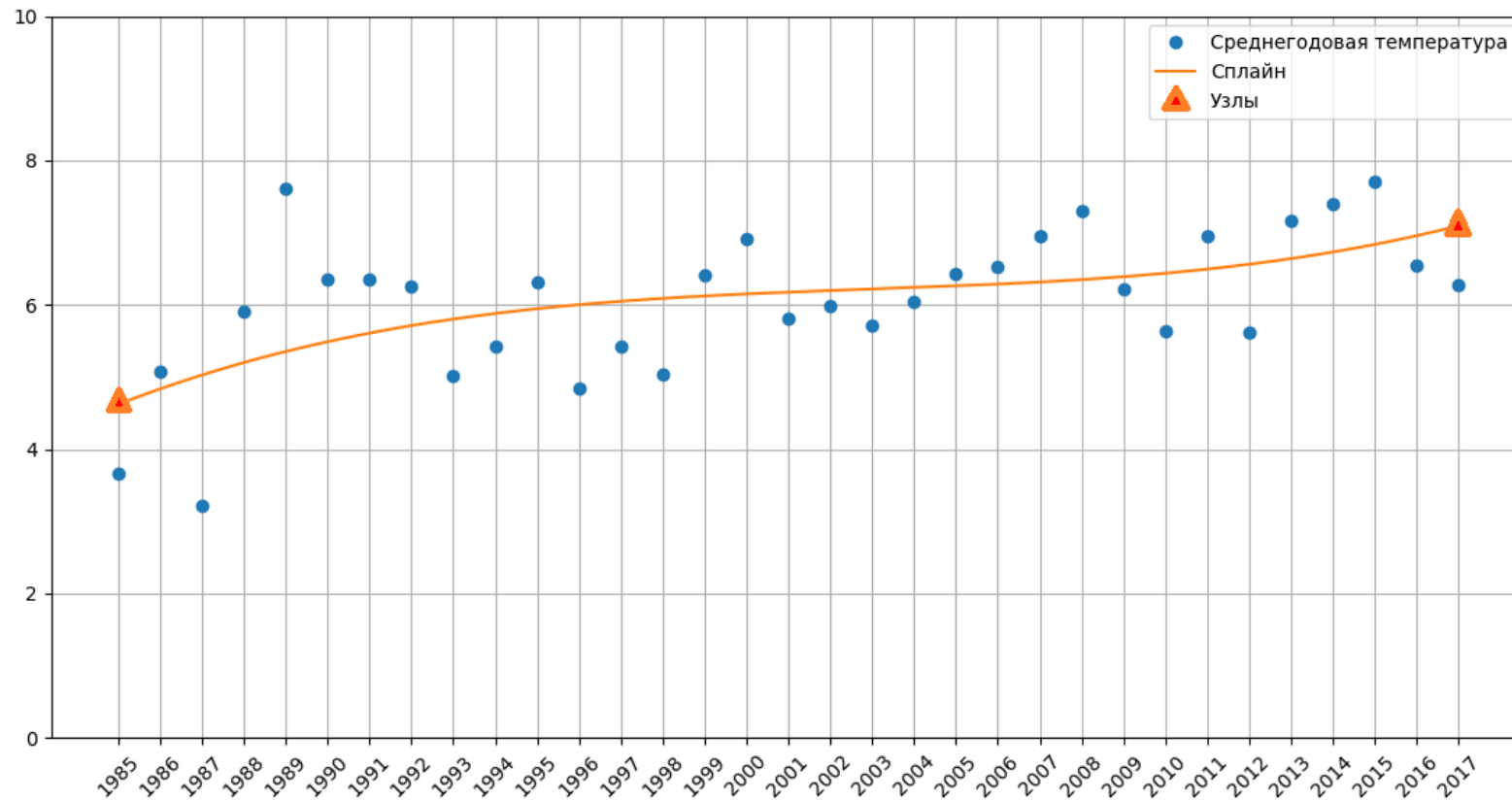


Рисунок 8 – Сглаживание среднегодовой температуры

Заключение

- Сделан обзор существующих методов аппроксимации.
- Спроектировано и разработано приложение для аппроксимации данных.
- Эксперимент показал практическую эффективность разработки.
- Дальнейшие направления исследований включают в себя расширение списка методов аппроксимации, обобщения методов аппроксимации на многомерные ряды.

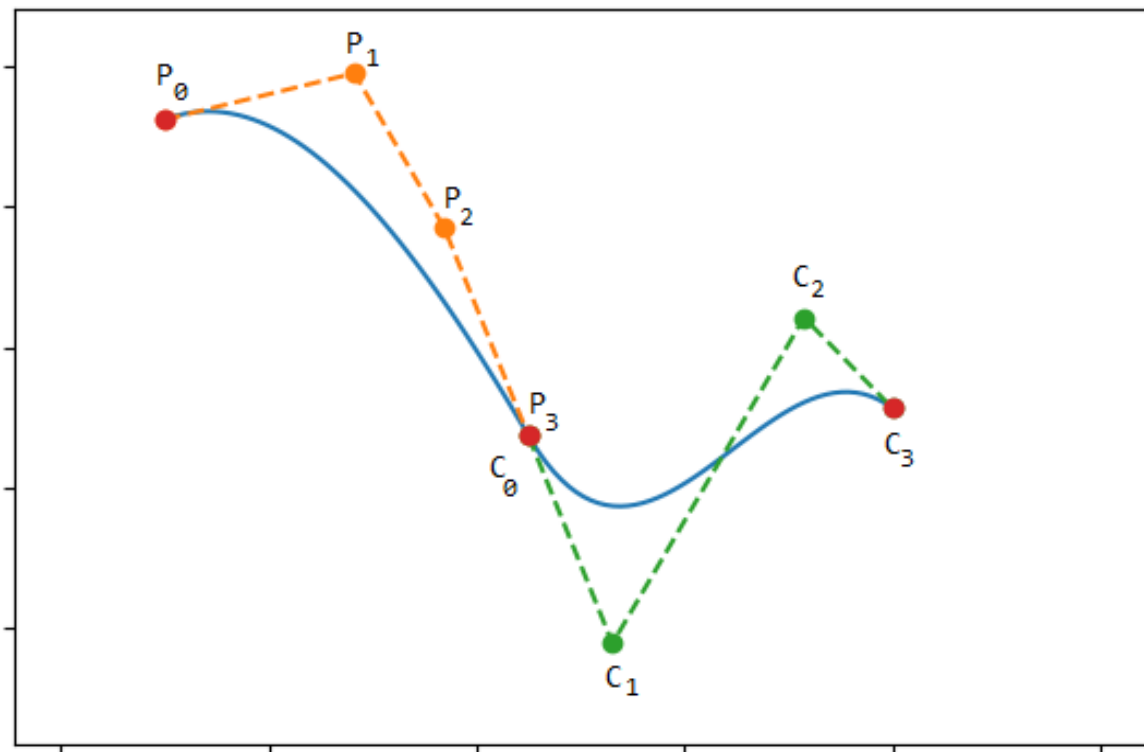
Репозиторий проекта

<https://github.com/Veelz/interp>

Кривые Безье

$$B(t) = P_0(1 - t)^3 + 3P_1(1 - t)^2t + 3P_2(1 - t)t^2 + P_3t^3$$

Рисунок — пример кривой, полученной стыковкой двух кривых Безье



Эксперимент

Таблица – зависимость среднеквадратического отклонения (СКО) от числа точек сплайна (по рис. 6, 7).

Число заданных точек	Число узлов сплайна	СКО
18	6	2
	12	1,3
	18	0

Эксперимент

Высокая коэффициент корреляции: температура и осадки; давление и осадки.

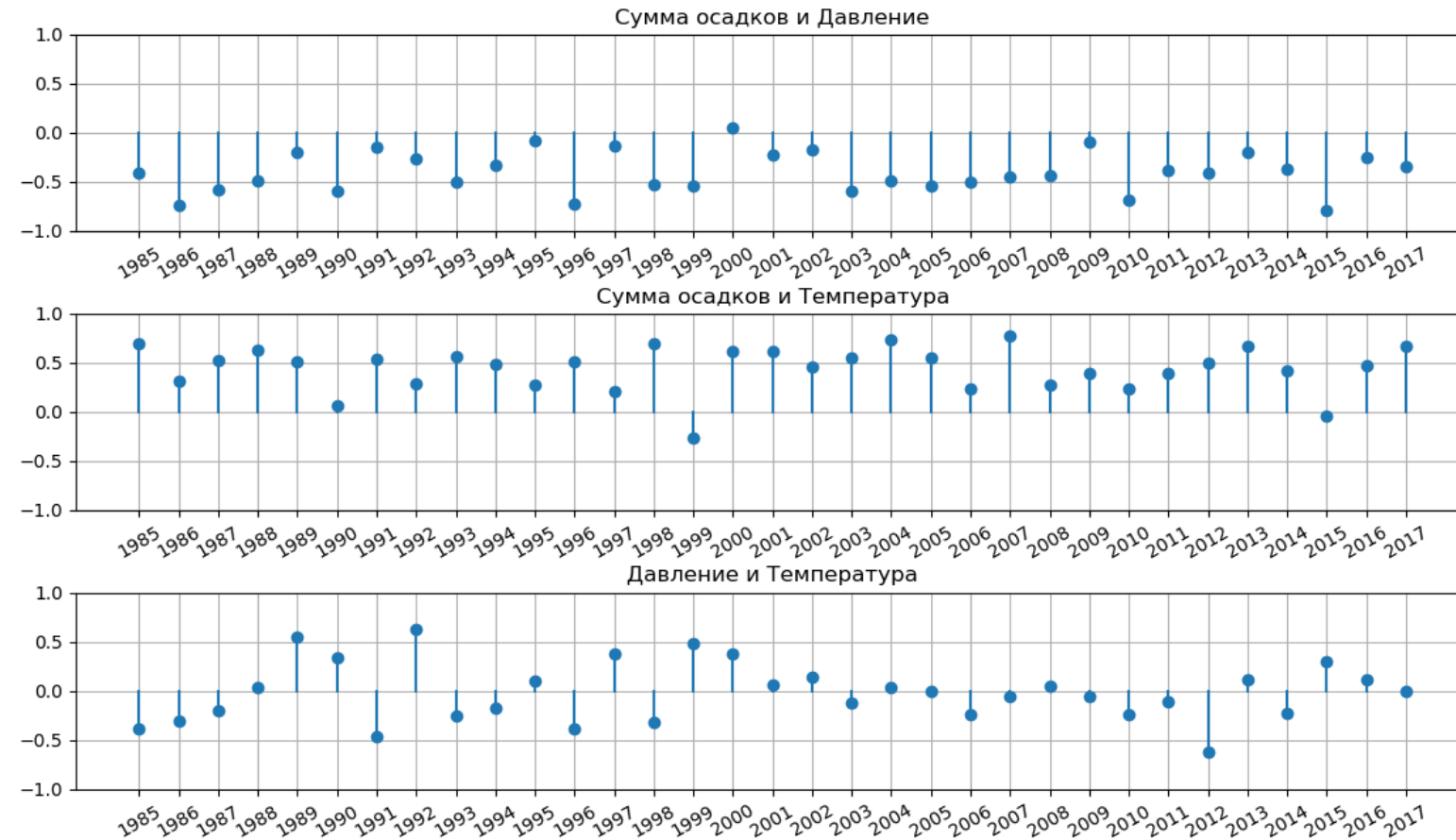


Рисунок – Корреляция между показателями

Тестирование

- в узлах сетки значения интерполирующей функции равны заданным значениям;
- вблизи узлов сетки интерполирующая функция монотонная;
- результат многократного вычисления значений аппроксимирующей или интерполирующей функции от одного и того же аргумента не меняется;
- сумма базисных функций, необходимых для кривых Безье и В-сплайнов, в любой точке определения равна 1;
- базисные функции В-сплайнов имеют один максимум в известной заранее точке и равны нулю на отрезке, где они не определены;
- вблизи узлов сетки В-сплайна при аппроксимации функция монотонная.