

КАФЕДРА Системы автоматизированного проектирования (РК-6)

## ОТЧЕТ О ВЫПОЛНЕНИИ ЛАБОРАТОРНОЙ РАБОТЫ

Тема лабораторной работы БПФ, полиномиальная регрессия

Преподаватель \_\_\_\_\_ **Першин Ю.А.**  
подпись, дата фамилия,  
и.о.

Оценка \_\_\_\_\_

Москва, 2019 г.

## Оглавление

Задание на лабораторную работу	3
Цель выполнения лабораторной работы	5
Задачи, выполненные в процессе реализации лабораторной работы	5
Численное интегрирование с помощью тригонометрических интерполянтов	5
Заключение	17
Список использованных источников	17

## Задание на лабораторную работу

### 1. Численное интегрирование с помощью тригонометрических интерполянтов

Даны интегралы:

$$I_1 = \int_{-\pi/4}^{\pi} |x| dx$$

$$I_2 = \int_{-\pi/4}^{\pi} (x \cos x^2 + e \cos e^x) dx$$

1.1. Вывести общее выражение для формулы численного интегрирования путем аналитического интегрирования тригонометрического ряда, заменяющего подинтегральную функцию.

1.2. Используя алгоритм Кули–Тьюки, написать функцию *fft\_coeff(y\_nodes)*, которая вычисляет и возвращает комплексные коэффициенты тригонометрического полинома, интерполирующего узлы *y\_nodes*, равномерно распределенные на отрезке  $[-\pi; \pi]$ .

1.3. Написать функцию *spectral\_integral(f, N)*, которая вычисляет значение интеграла функции *f*, интерполируемой в *N* узлах с помощью тригонометрического ряда, на интервале  $[-\frac{\pi}{4}; \pi]$ . Функция *spectral\_integral* должна использовать внутри себя функцию *fft\_coeff*.

1.4. Для каждого из интегралов  $I_1$  и  $I_2$  провести следующий анализ:

- Найти точное значение интеграла.
- Найти приближенное значение интеграла с помощью функции *spectral\_integral* для  $N = 2^{\tilde{n}}$ , где  $\tilde{n} \in 1, \dots, 8$ .
- Для каждого *N* найти относительную погрешность вычислений  $\delta$  и вывести на экран график зависимости  $\delta$  от *N*, где  $\delta$  следует отображать в логарифмической шкале

1.5. Объяснить, как можно использовать полученные логарифмические графики для оценки порядка точности интегрирования.

1.6. Ответить, различаются ли порядки точности интегрирования в случае вычисления интегралов  $I_1$  и  $I_2$  и, если различаются, объяснить, с чем это связано.

## 2. Полиномиальная регрессия

Дана функция

$$f(x) = -10x^2 + 1.5x + 1 + \sigma X,$$

где  $x \in [-1;1]$  и  $X$  – случайная величина, нормально распределенная на интервале

Требуется:

2.1. Написать функцию `poly_regression(x_nodes, y_nodes, degree)`, которая возвращает коэффициенты многочлена степени `degree`, наилучшим образом приближающегося к точкам с абсциссами `x_nodes` и ординатами `y_nodes`.

2.2. Для каждого  $\sigma$  из множества  $\{10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\}$ ,  $N$  из множества  $\{2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9\}$  и  $p$  из множества  $1, 2, 3, 4, 5$  провести следующий анализ.

- С помощью функции  $f(x)$  сгенерировать начальный набор данных  $D_{reg}^{(N)}$  и проверочный набор данных  $D_{test}^{(N)}$ , где  $N$  – число точек в наборе данных.
- С помощью набора данных  $D_{reg}^{(N)}$  и функции `poly_regression` построить многочлен степени  $p$ , наилучшим образом приближающийся к данным
- Вычислить среднеквадратичную погрешность аппроксимации  $\epsilon_{reg}^{(N,p)}$  данных  $D_{reg}^{(N)}$  полученным многочленом.
- Вычислить среднеквадратичную погрешность аппроксимации  $\epsilon_{test}^{(N,p)}$  данных  $D_{test}^{(N)}$  полученным многочленом.

2.3. Вывести на экран несколько характерных примеров графиков многочленов вместе с начальными и проверочными данными.

2.4. Ответить на следующие вопросы.

- Как влияет увеличение числа начальных и проверочных данных на  $\epsilon_{reg}^{(N,p)}$  и  $\epsilon_{test}^{(N,p)}$ ? Выведите на экран графики, иллюстрирующие ответ, и сделайте вывод
- Как влияет увеличение степени многочлена  $p$  на  $\epsilon_{reg}^{(N,p)}$  и  $\epsilon_{test}^{(N,p)}$ ? Выведите на экран графики, иллюстрирующие ответ, и сделайте вывод
- Как влияет увеличение числа начальных и проверочных данных на относительную погрешность коэффициентов многочлена при сравнении с  $f(x)$  в случае фиксированного  $\sigma$ ? Что происходит при увеличении  $\sigma$ ? Выведите на экран графики, иллюстрирующие ответ, и сделайте вывод. Как может повлиять изменение функции распределения случайной величины  $X$  на сделанный вывод?

2.5. Исходя из ответов на предыдущие вопросы, сделать общий вывод о свойствах сходимости полиномиальной регрессии в случае зашумленных данных.

### **Цель выполнения лабораторной работы**

Реализовать алгоритм БПФ, полиномиальной регрессии и проанализировать результаты.

### **Задачи, выполненные в процессе реализации лабораторной работы**

1. Численное интегрирование с помощью тригонометрических интерполянтов

1.1. Рассчитано общее выражение для формулы численного интегрирования путем аналитического интегрирования тригонометрического ряда, заменяющего подынтегральную функцию. Вывод представлен ниже:

Тригонометрический полином  $k$ -ой степени имеет вид:

$$p(x) = a_0 + \sum_{k=1}^n a_k \cdot \cos(kx) + \sum_{k=1}^n b_k \cdot \sin(kx) \quad (1)$$

Для того, чтобы вывести выражение для формулы численного интегрирования для функции, определенной на промежутке  $[-\frac{\pi}{4}; \pi]$ , подынтегральная функция была заменена тригонометрический полиномом из формулы (1).

$$\int_{-\pi/4}^{\pi} f(x) dx = \int_{-\pi/4}^{\pi} (a_0 + \sum_{k=1}^n (a_k \cdot \cos(kx) + b_k \cdot \sin(kx))) dx =$$

$$a_0 \cdot \int_{-\pi/4}^{\pi} dx + \sum_{k=1}^n (a_k \cdot \int_{-\pi/4}^{\pi} \cos(kx) dx + b_k \cdot \int_{-\pi/4}^{\pi} \sin(kx) dx) \quad (2)$$

Тогда первообразная функции (2):

$$\frac{5\pi}{4} a_0 + \frac{a_k}{k} \cdot (\sin(k\pi) - \sin(-\frac{k\pi}{4})) - \frac{b_k}{k} \cdot (\cos(k\pi) - \cos(-\frac{k\pi}{4})) =$$

$$\frac{5\pi}{4} a_0 + \frac{a_k}{k} \cdot \sin(\frac{k\pi}{4}) - \frac{b_k}{k} \cdot (\cos(k\pi) - \cos(\frac{k\pi}{4})) =$$

$$\frac{5\pi}{4} a_0 + \frac{a_k}{k} \cdot \sin(\frac{k\pi}{4}) + \frac{b_k}{k} \cdot (-\cos(k\pi) + \cos(\frac{k\pi}{4})) \quad (3)$$

Тригонометрический полином (1) часто называют тригонометрическим рядом. Тригонометрический ряд можно записать в экспоненциальной форме. Для экспоненциальной формы  $a_0$ ,  $a_k$ ,  $b_k$  выражаются через комплексный коэффициент  $\hat{a}_k$  через формулу Эйлера:

$$a_0 = \text{Re}(\hat{a}_k) \quad (4)$$

$$a_k = 2\text{Re}(\hat{a}_k) \quad (5)$$

$$b_k = -2\text{Im}(\hat{a}_k) \quad (6)$$

$\text{Im}$  — краткое обозначение мнимой части комплексного числа.

$\text{Re}$  — обозначение для действительной части комплексного числа. [1]

После подставления формул (4-6) искомая формула имеет вид:

$$\frac{5\pi}{4} \text{Re}(\hat{a}_k) + \frac{2\text{Re}(\hat{a}_k)}{k} \cdot \sin(\frac{k\pi}{4}) + \frac{-2\text{Im}(\hat{a}_k)}{k} \cdot (-\cos(k\pi) + \cos(\frac{k\pi}{4})) \quad (7)$$

1.2. Реализована функция `fft_coeff(y_nodes)`, которая вычисляет комплексные коэффициенты тригонометрического полинома на основании алгоритма Кули–Тьюки. Данный алгоритм используется для вычисления дискретного преобразования Фурье (ДПФ) за время  $O(m \log_2 m)$  и основан на следующих принципах:

- ДПФ размерности N выражается через сумму ДПФ более малых размерностей N1 и N2.
- Аналогичным образом N1 и N2 рекурсивно выражаются через ДПФ ещё более малых размерностей.
- Разложение ДПФ в сумму 2 частей можно организовать таким образом, что первая часть будет представлять собой сумму по чётным индексам, а вторая — по нечётным. [2]

Алгоритм вычисляет сумму

$$A_k = \sum_{j=0}^{2m-1} y_j e^{\frac{-ijk\pi}{m}} \quad (8),$$

которая нужна для вычисления комплексного коэффициента  $\hat{a}_k$ , где

$$\hat{a}_k = \frac{(-1)^k}{2m} \cdot A_k \quad (9)$$

1.3. Была написана функция *spectral\_integral(f, N)*, которая вычисляет значение интеграла функции *f*, интерполируемой в N узлах с помощью тригонометрического ряда, на интервале  $[-\frac{\pi}{4}; \pi]$ . Для этого использовались формулы (7) и (9). Функция *spectral\_integral* использует внутри себя функцию *fft\_coeff*. Для вычисления интеграла была использована формула (7).

1.4. Для каждого из интегралов  $I_1$  и  $I_2$  был проведен анализ:

- Найдены точные значения интегралов с помощью библиотеки *sympy*:

$$I_1 = -1.85726$$

$$I_2 = 5.24322$$

- Найдена относительная погрешность вычислений  $\delta$  в зависимости от N. Для нахождения погрешности использовалось приближенное значение интеграла, которое было найдено с помощью функции *spectral\_integral* для  $N = 2^{\tilde{n}}$ , где  $\tilde{n} \in 1, \dots, 8$ . Относительная погрешность приближенного значения  $a^*$  была посчитана по формуле:

$$\delta(a^*) = \left| \frac{a^* - a}{a} \right| \quad (10),$$

$a^*$  – приближенное значение интеграла,

$a$  – точное значение интеграла

Были выведены на экран график зависимости  $\delta$  от  $N$ , где  $\delta$  следует отображается в логарифмической шкале. (рис. 1, рис. 2, рис. 3)

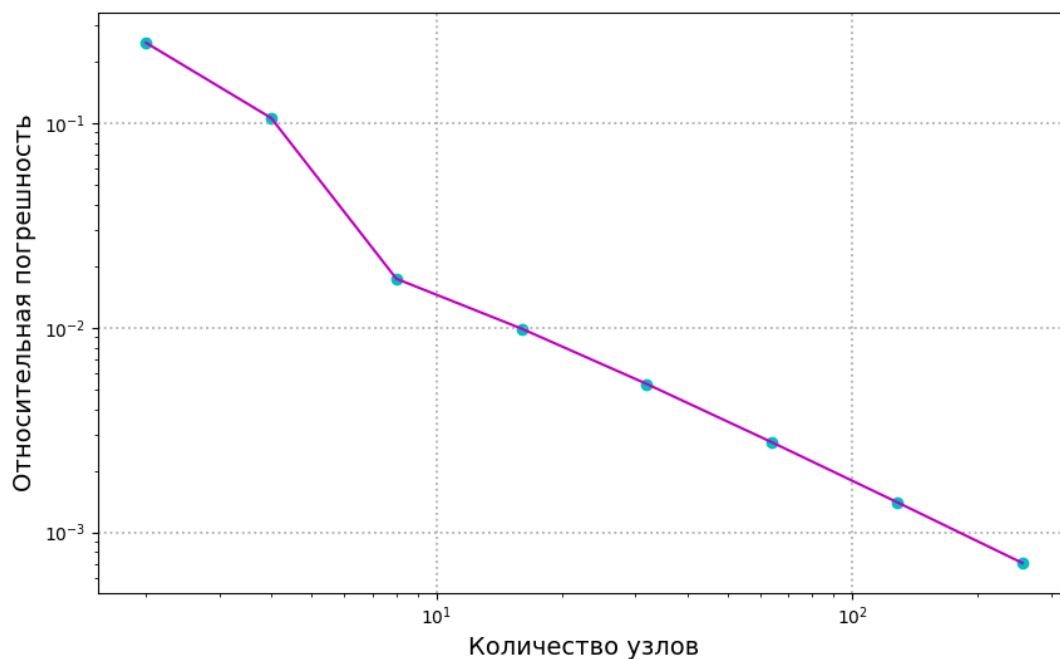


Рис. 1. Зависимость относительной погрешности от количества узла  
для интеграла  $I_1$ , где  $N = 2^{\tilde{n}}$ , где  $\tilde{n} \in 1, \dots, 8$ .

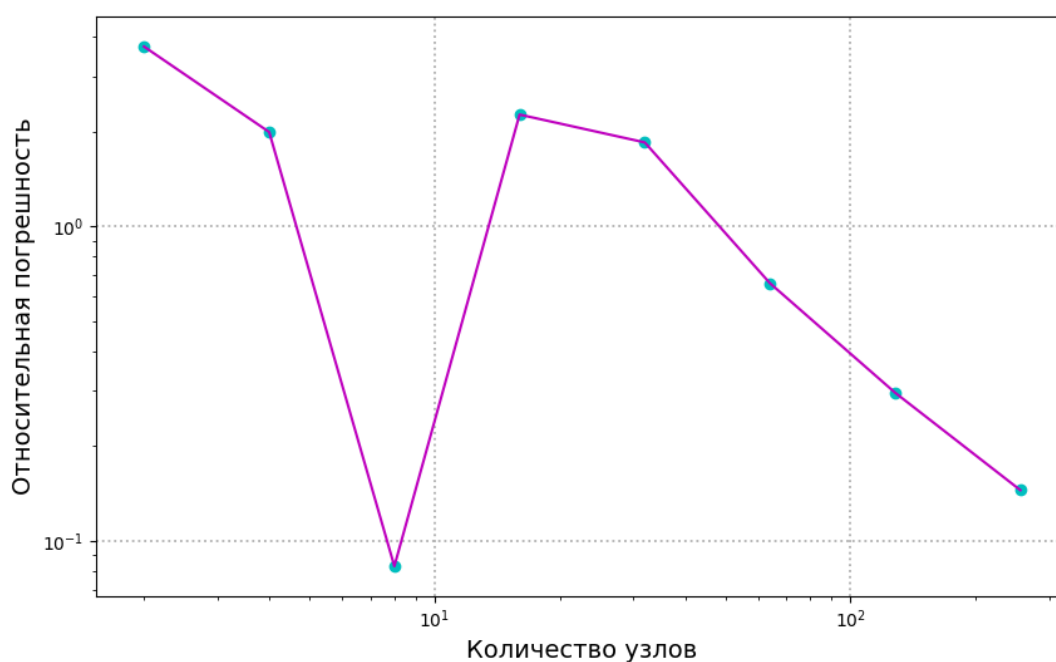


Рис. 2. Зависимость относительной погрешности от количества  
узлов( $N$ ) для интеграла  $I_2$ , где  $N = 2^{\tilde{n}}$ , где  $\tilde{n} \in 1, \dots, 8$ .



1.5. Полученные логарифмические графики можно использовать для оценки порядка точности интегрирования.

Для интеграла  $I_1$  из графика на рис. 1:

- 1) Количество узлов 16, относительная погрешность  $10^{-2}$ .
- 2) Количество узлов 200, относительная погрешность  $10^{-3}$ .

При увеличении количества узлов примерно в 12.5 раз, относительная погрешность уменьшилась в 10 раз. Так как при увеличении количества узлов в  $\approx 10$  раз, уменьшение относительной погрешности происходит в 10 раз, то порядок точности интегрирования  $I_1$  можно считать равным примерно 1.

Для интеграла  $I_2$  из графика на рис. 2, рис. 3:

- 1) Количество узлов  $\approx 56$ , относительная погрешность 10.
- 2) Количество узлов 400, относительная погрешность  $10^{-1}$ .

При увеличении количества узлов примерно в 7 раз, относительная погрешность уменьшилась в 10 раз. По аналогичным утверждениям, точность интегрирования  $I_2$  можно считать равным примерно 1. Также можно отметить, что для интеграла  $I_2$  минимальная относительная погрешность получается при количестве узлов равном 8, а затем резко возрастает, а для интеграла  $I_1$  относительная погрешность уменьшается с возрастанием количества узлов без скачков.

1.6. Был исследован вопрос о том, различаются ли порядки точности интегрирования в случае вычисления интегралов  $I_1$  и  $I_2$ .

Погрешность интегрирования зависит от:

1. Вида подынтегральной функции
2. Метода интегрирования
3. Шага сетки [3]

Для вычисления интегралов  $I_1$  и  $I_2$  использовался один и тот же метод интегрирования и одинаковый шаг сетки, поэтому погрешность интегрирования в таком случае зависит от подынтегральной функции. Пусть

подынтегральная функция  $I_1$  будет названа  $f_1$ , а подынтегральная функция  $I_2$ . Тогда

$$f_1(x) = |x| \quad (11)$$

$$f_2(x) = x \cos x^2 + e \cos e^x \quad (12).$$

Функция  $f_1$  периодическая на интервале  $[-\pi; \pi]$ ,  $f_2$  - непериодическая. Непериодическая функция не может быть разложена в ряд Фурье для всех значений  $x$ . Для каждой из функций можно составить новую функцию  $f_1(x) *$ ,  $f_2(x) *$ , выбирая значения в определенном диапазоне и повторяя их вне этого диапазона с интервалом  $2\pi$ . Для функций  $f_1(x) *$ ,  $f_2(x) *$  можно определить ряд Фурье, представляющий функцию в любом диапазоне шириной  $2\pi$ . Для  $f_1$  можно построить такую периодическую функцию  $f_1(x) *$  без разрывов, а  $f_2$  нельзя. Т. к.  $f_1(x) *$  непрерывная на замкнутом отрезке, из аппроксимационной теоремы Вейерштрасса следует, что существует такая последовательность многочленов такая, что она равномерно сходится к  $f_1(x) *$  при  $n \rightarrow \infty$ . Для  $f_2(x) *$  суммы Фурье будут сильно осциллировать в малых окрестностях точек разрыва. Из этого можно сделать вывод, что порядок точности выше у  $I_1$ . [4][5]

## 2. Полиномиальная регрессия

2.1. Была написана функция *poly\_regression(x\_nodes, y\_nodes, degree)*, которая возвращает коэффициенты многочлена степени *degree*, наилучшим образом приближающегося к точкам с абсциссами *x\_nodes* и ординатами *y\_nodes*.

2.2. Для каждого  $\sigma$  из множества  $\{10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\}$ ,  $N$  из множества  $\{2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9\}$  и  $p$  из множества  $1, 2, 3, 4, 5$  был проведен следующий анализ.

- Был сгенерирован начальный набор данных  $D_{reg}^{(N)}$  и проверочный набор данных  $D_{test}^{(N)}$ , где  $N$  - число точек в наборе данных. Для генерации использовалась функция

$$f(x) = -10x^2 + 1.5x + 1 + \sigma X \quad (13)$$

- С помощью набора данных  $D_{regr}^{(N)}$  и функции *poly\_regression* был построен многочлен степени  $p$ , наилучшим образом приближающийся к данным
- Была вычислена среднеквадратичная погрешность аппроксимации  $\epsilon_{regr}^{(N,p)}$  данных  $D_{regr}^{(N)}$  полученным многочленом. Для среднеквадратичной погрешности использовалась формула:

$$\sqrt{\frac{\sum_{i=1}^N (y[i] - P(x_i))^2}{N}} \quad (14)$$

- Была вычислена среднеквадратичная погрешность аппроксимации, также вычисляемая по формуле (14),  $\epsilon_{test}^{(N,p)}$  данных  $D_{test}^{(N)}$  полученным многочленом.

2.3. Несколько характерных примеров графиков многочленов вместе с начальными и проверочными данными показано на рис. 4, рис. 5, рис. 6.

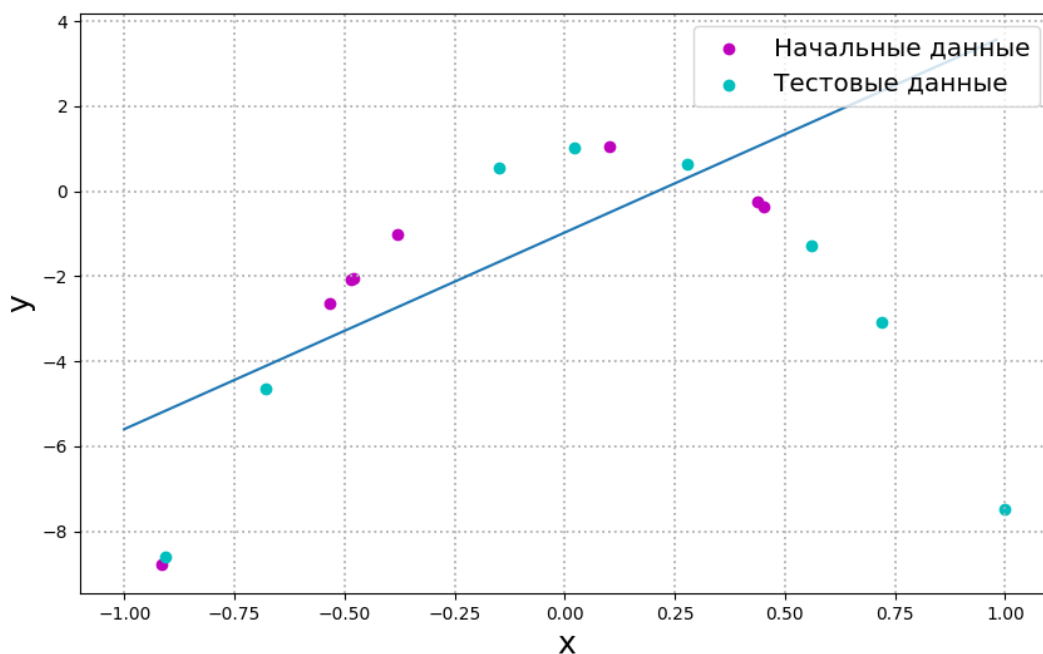


Рис. 3. Многочлен степени  $p = 1$ , полученный из функции  $f(x)$ , где  $\sigma = 10^{-2}$ , количество узлов  $N = 2^3$

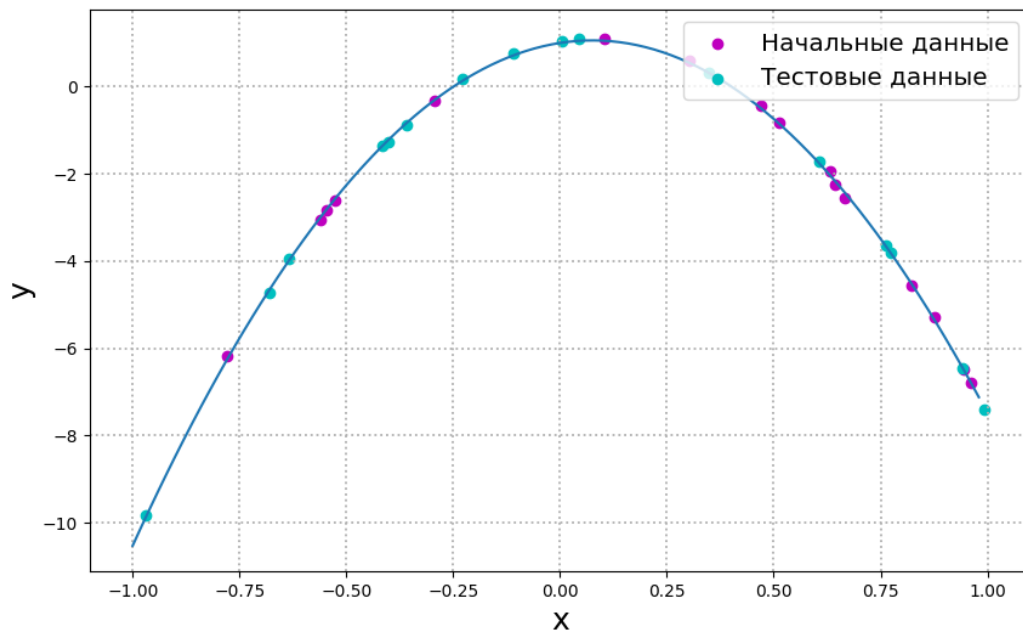


Рис. 4. Многочлен степени  $p = 2$ , полученный из функции  $f(x)$ , где  $\sigma = 10^{-1}$ , количество узлов  $N = 2^4$ .

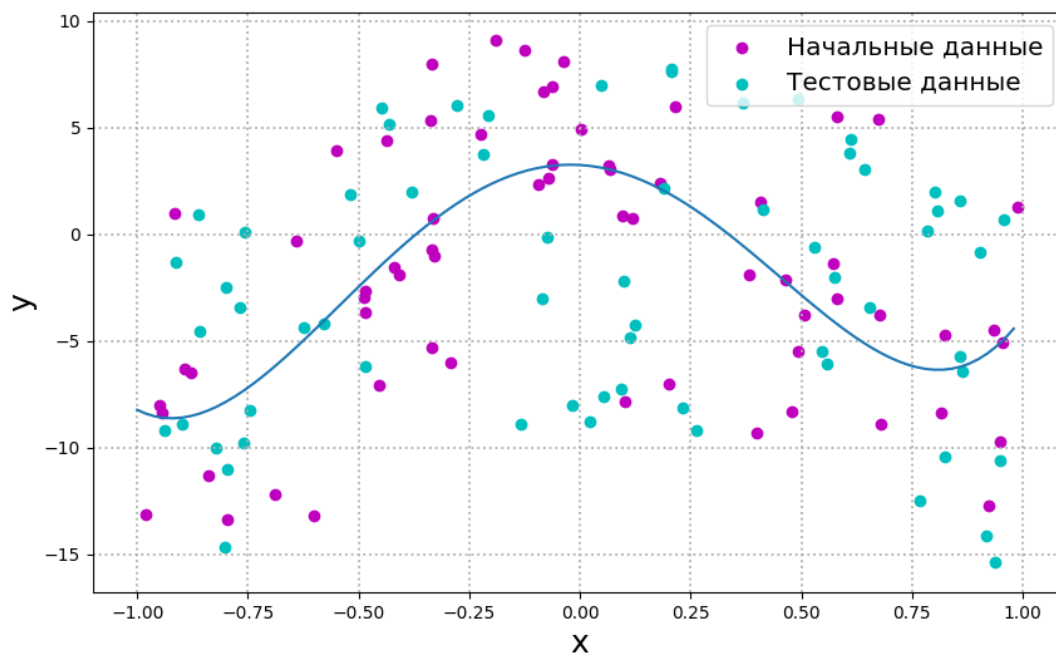


Рис. 5. Многочлен степени  $p = 4$ , полученный из функции  $f(x)$ , где  $\sigma = 10^1$ , количество узлов  $N = 2^6$ .

2.4. Была исследована среднеквадратичная погрешность.

- Из рис. 7, рис. 8 следует, что при увеличении числа начальных и проверочных данных,  $\epsilon_{reg}^{(N,p)}$  и  $\epsilon_{test}^{(N,p)}$  существенно различны только при небольшом количестве узлов (до 64 узлов), после этого погрешности начинают сходиться к  $\sigma$ .

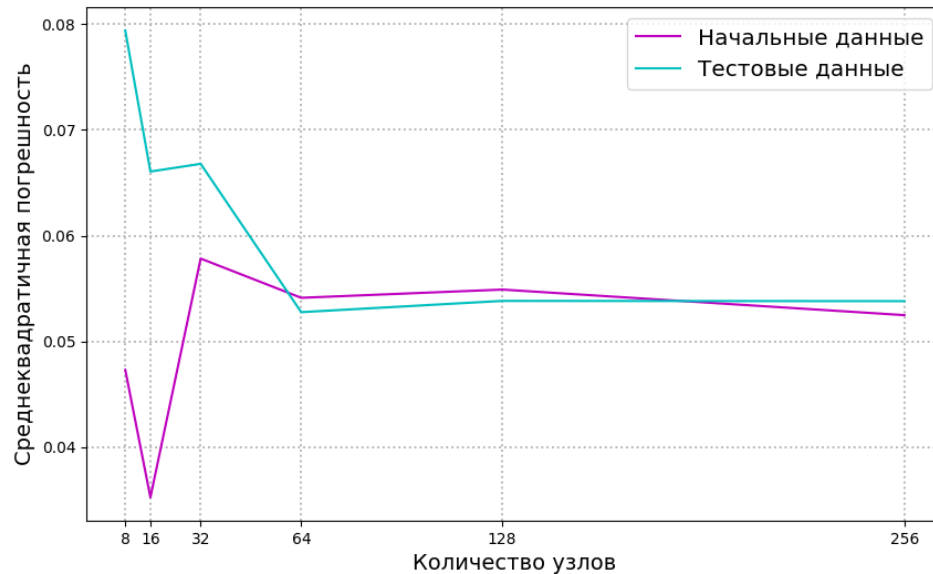


Рис. 6. Среднеквадратическая погрешность полинома ( $p = 3$ ,  $\sigma = 10^{-1}$ ) для разного количества узлов.

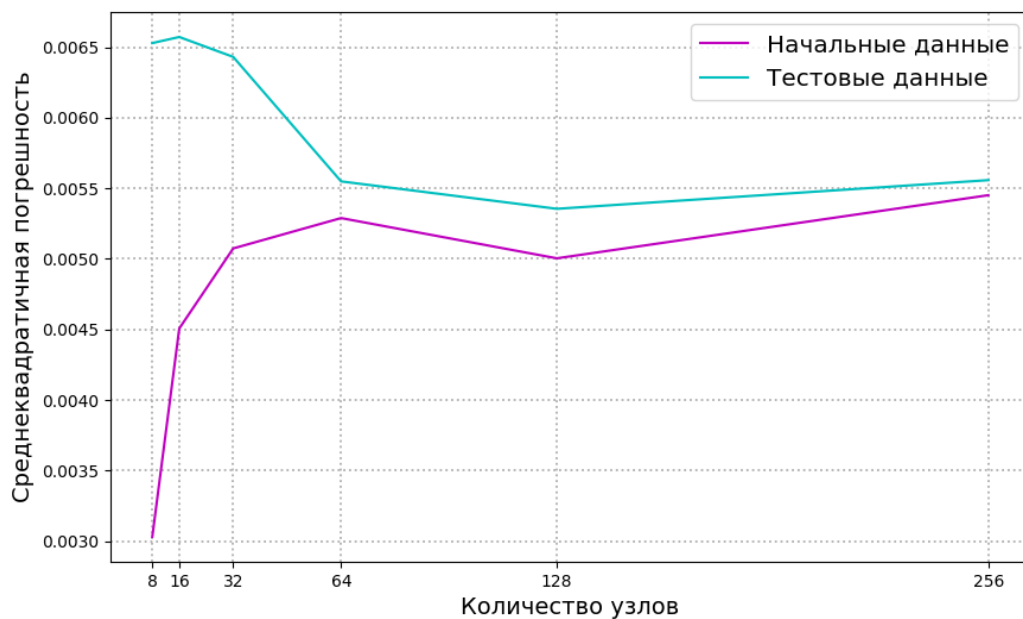


Рис. 7. Среднеквадратическая погрешность полинома ( $p = 2$ ,  $\sigma = 10^{-2}$ ) для разного количества узлов.

- Было исследовано, как влияет увеличение степени многочлена  $p$  на  $\epsilon_{reg}^{(N,p)}$  и  $\epsilon_{test}^{(N,p)}$ . Из рис. 8-10 следует, что в целом графики погрешностей

сходятся к  $\sigma$ . Можно заметить, что  $\epsilon_{reg}^{(N,p)}$  и  $\epsilon_{test}^{(N,p)}$  имеют большое значение при степени полинома  $p = 1$ , для полиномов большей степени, погрешность меньше.

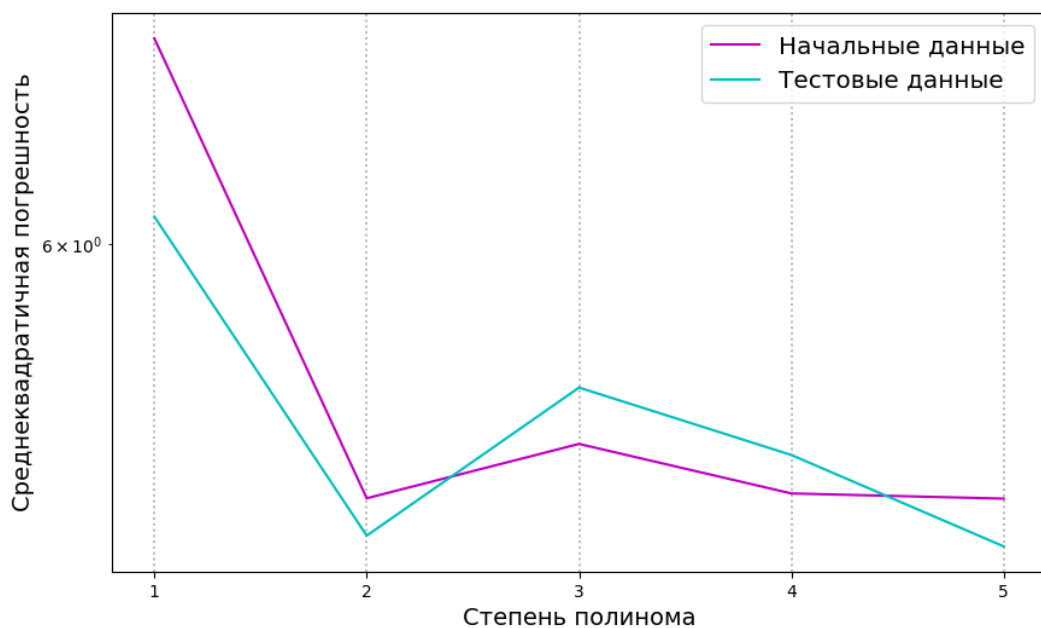


Рис. 8. Среднеквадратическая погрешность полинома ( $N = 2^8$ ,  $\sigma = 10^1$ ) для разной степени полинома.

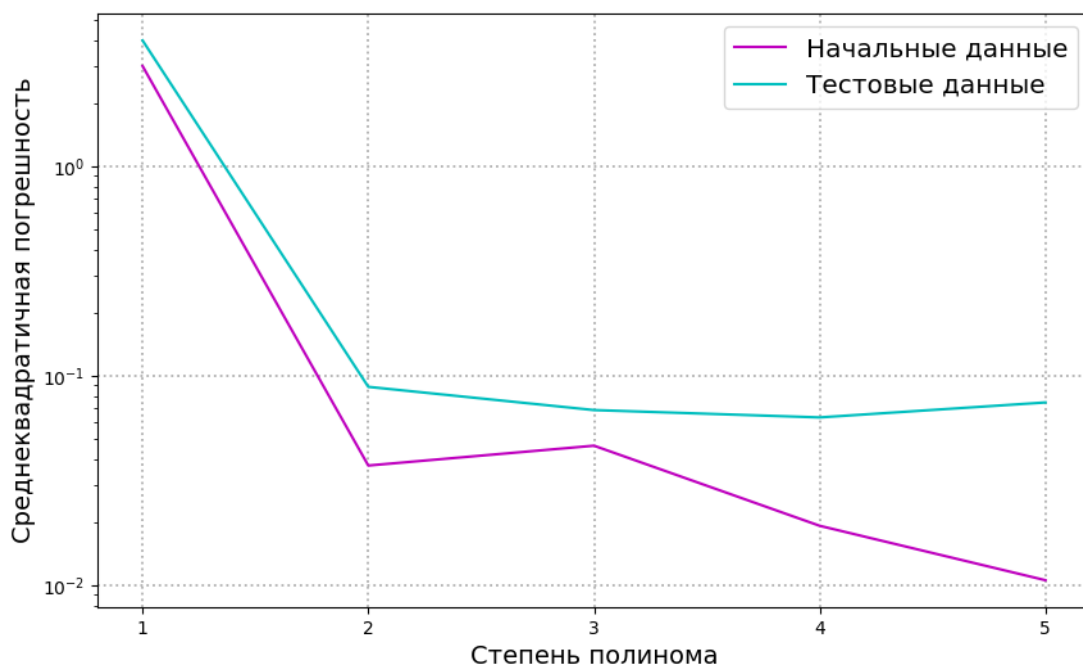


Рис. 9. Среднеквадратическая погрешность полинома ( $N = 2^3$ ,  $\sigma = 10^{-1}$ ) для разной степени полинома.

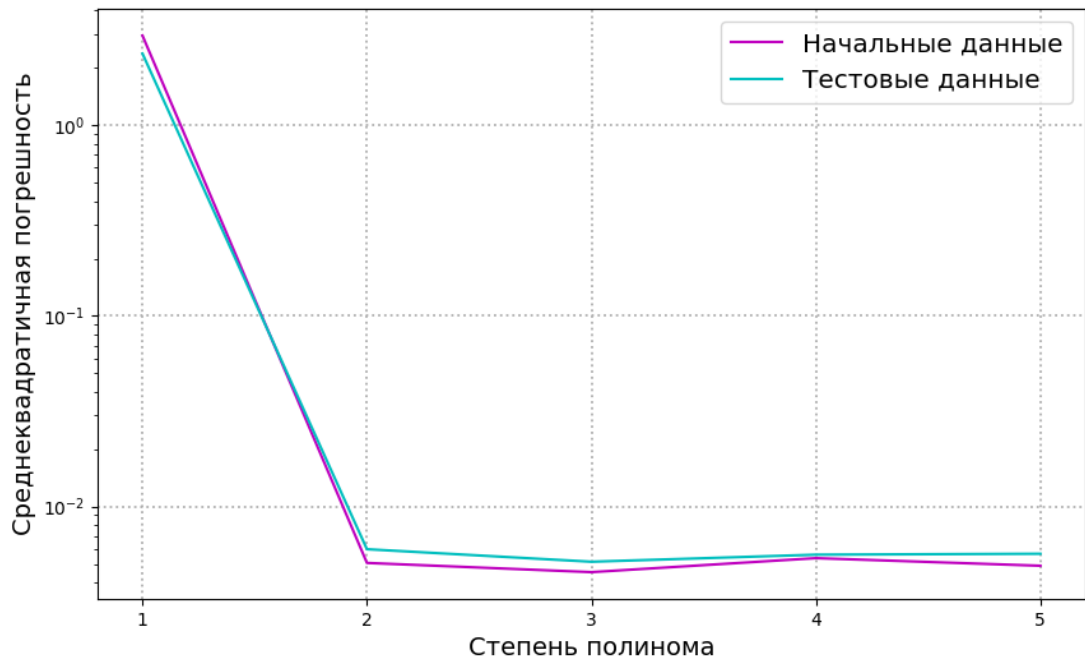


Рис. 10. Среднеквадратическая погрешность полинома ( $N = 2^6$ ,  $\sigma = 10^{-2}$ ) для разной степени полинома.

- Было исследовано, как влияет увеличение числа начальных и проверочных данных на относительную погрешность коэффициентов многочлена при сравнении с  $f(x)$  в случае фиксированного  $\sigma$ . Для подсчета погрешности использовалась формула:

$$\frac{1}{p+1} \sum_{i=0}^p \left| \frac{a^i - b^i}{a^i} \right|,$$

$a^i$  - коэффициенты исходной функции,

$b^i$  - коэффициенты полинома,

$p = 2$  из исходной функции

График для  $\sigma = 10^{-2}$  представлен на рис.12. Из графика можно сделать вывод, что относительная погрешность коэффициентов уменьшается при увеличении количества узлов. То есть, чтобы полином был максимально близок к исходной функции, нужно брать большое количество точек.

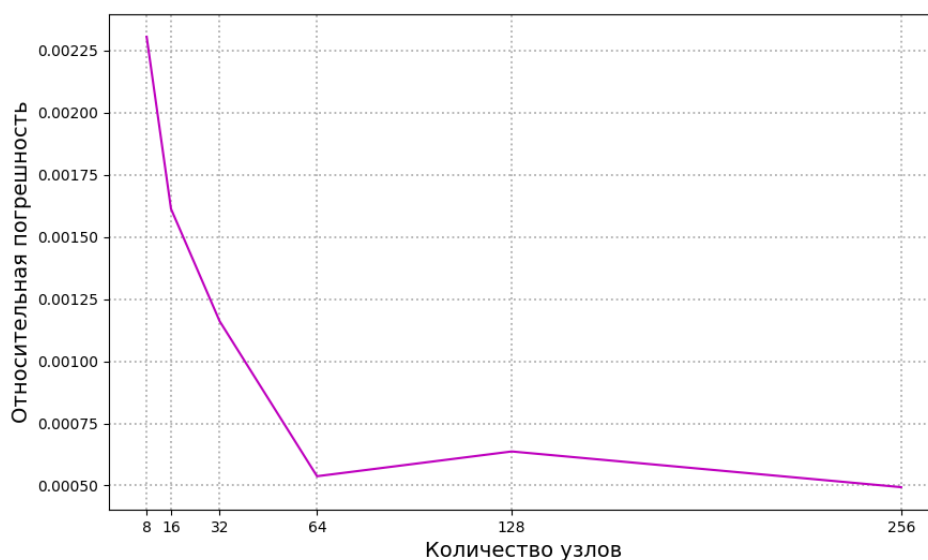


Рис. 11. Зависимость относительной погрешности коэффициентов от количества узлов полинома для  $\sigma = 10^{-2}$ .

- Было исследовано, что произойдет при увеличении  $\sigma$ .  $\sigma$  изменялась от  $10^{-2}$  до  $10^2$ ,  $N$  от  $2^3$  до  $2^7$ . График относительной погрешности коэффициентов показан на рис. 12. Из него можно сделать вывод, что при увеличении количества узлов при возрастающем  $\sigma$  относительная погрешность растет. Из этого можно сделать вывод, что постоянная  $\sigma$  в таком случае будет эффективней.

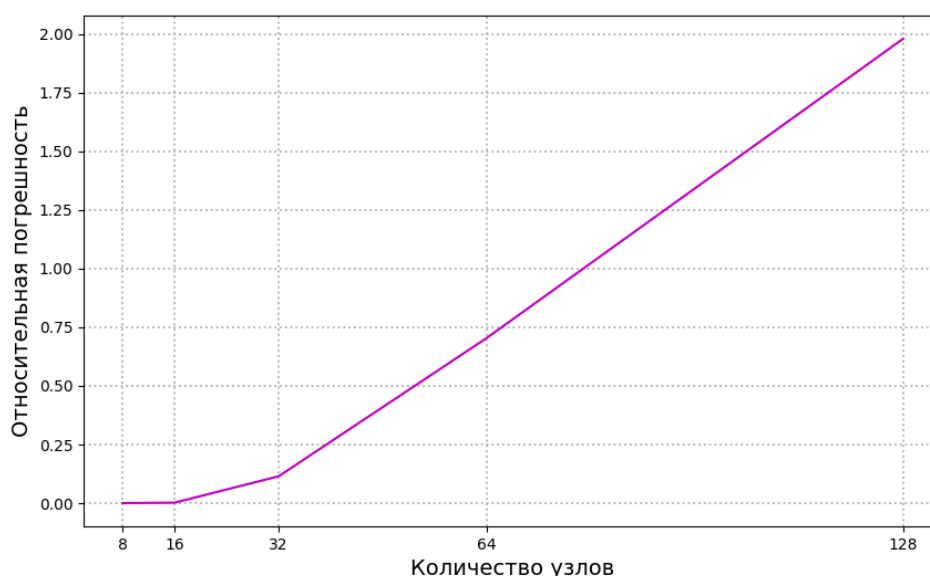


Рис. 12. Зависимость относительной погрешности коэффициентов от количества узлов полинома для изменяющейся сигма  $\sigma$  от  $10^{-2}$  до  $10^2$ .



2.5. Сделан вывод о свойствах сходимости полиномиальной регрессии в случае зашумленных данных: зашумленные данные ухудшают точность приближения полиномом, также необходимо брать постоянную  $\sigma$  для большей точности.

### **Заключение**

- 1) При использовании тригонометрических полиномов при численном интегрировании получается наиболее точный результат, если подынтегральная функция периодическая.
- 2) Полиномиальная регрессия может использоваться для зашумленных данных, для этого следует выбирать степень полинома, при которой погрешность будет минимальна.

### **Список использованных источников**

1. Першин А.Ю., Соколов А.П., Вычислительная математика, Лабораторные работ. Учебное пособие. Москва, 2018.
2. Першин А.Ю. Лекции по вычислительной математике. Москва, 2019, 143 с.
3. Быстрое преобразование Фурье [Электронный ресурс] // URL: <https://ru.bmstu.wiki/>
4. Численное интегрирование [Электронный ресурс] // URL: <http://portal.tpu.ru:7777/SHARED/1/LOPATKIN/Students/DG/8-Integration.pdf>
5. Ряд Фурье [Электронный ресурс] // URL: <https://tehtab.ru/Guide/GuideMathematics/SeriesOfTaylorMaklorenFourier/FourierSeries/>
6. Р. В. Голованов, Н. Н. Калиткин, К. И. Луцкий, Нечетное продолжение для фурье-аппроксимации непериодических функций, Матем. моделирование, 2013, том 25, номер 5, 67–84