



Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ Робототехники и комплексной автоматизации

КАФЕДРА Системы автоматизированного проектирования (РК-6)

ОТЧЕТ О ВЫПОЛНЕНИИ ЛАБОРАТОРНОЙ РАБОТЫ

Студент Лубянов Александр Дмитриевич

Группа РК6-62Б

Тип задания лабораторная работа

Тема лабораторной работы БПФ и полиномиальная регрессия

Студент _____ Лубянов, А. Д.
подпись, дата фамилия, и.о.

Преподаватель _____ Першин, А. Ю.
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
1. Быстрое преобразование Фурье	6
2. Регрессия.....	9
Заключение	17
Список использованных источников	18

Задание на лабораторную работу

Задача 1 (БПФ)

Даны интегралы

$$I_1 = \int_{-\frac{\pi}{4}}^{\pi} |x| dx,$$
$$I_2 = \int_{-\frac{\pi}{4}}^{\pi} (x \cos x^2 + e^x \cos e^x) dx.$$

Требуется:

1. Вывести общее выражение для формулы численного интегрирования путем аналитического интегрирования тригонометрического ряда, заменяющего подинтегральную функцию.
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 $[-\pi; \pi]$.
3. Написать функцию *spectral_integral(f, N)*, которая вычисляет значение интеграла функции *f*, интерполируемой в *N* узлах с помощью тригонометрического ряда, на интервале $[-\pi/4; \pi]$. Функция *spectral_integral* должна использовать внутри себя функцию *fft_coeff*.
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 $N = 2^n$, где $n \in 1, \dots, 8$.
 - Для каждого *N* найти относительную погрешность вычислений δ и вывести на экран график зависимости δ от *N*, где δ следует отображать в логарифмической шкале.
5. Объяснить, как можно использовать полученные логарифмические графики для оценки порядка точности интегрирования.
6. Ответить, различаются ли порядки точности интегрирования в случае вычисления интегралов I_1 и I_2 и, если различаются, объяснить, с чем это связано.

Задача 9 (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 $[-1; 1]$.

Требуется:

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. Для каждого σ 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 $D_{regr}^{(N)}$ и проверочный набор данных $D_{test}^{(N)}$, где N – число точек в наборе данных.
 - С помощью набора данных $D_{regr}^{(N)}$ и функции *poly_regression* построить многочлен степени p , наилучшим образом приближающийся к данным.
 - Вычислить среднеквадратичную погрешность аппроксимации $\epsilon_{regr}^{(N,p)}$ данных $D_{regr}^{(N)}$ полученным многочленом.
 - Вычислить среднеквадратичную погрешность аппроксимации $\epsilon_{test}^{(N,p)}$ данных $D_{test}^{(N)}$ полученным многочленом.
3. Вывести на экран несколько характерных примеров графиков многочленов вместе с начальными и проверочными данными.
4. Ответить на следующие вопросы:
 - Как влияет увеличение числа начальных и проверочных данных на $\epsilon_{regr}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$? Выведите на экран графики, иллюстрирующие ответ, и сделайте вывод.
 - Как влияет увеличение степени многочлена p на $\epsilon_{regr}^{(N,p)}$ 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 σ ? Что происходит при увеличении σ ? Выведите на экран графики, иллюстрирующие ответ, и сделайте вывод. Как может повлиять изменение функции распределения случайной величины X на сделанный вывод?
5. Исходя из ответов на предыдущие вопросы, сделать общий вывод о свойствах сходимости полиномиальной регрессии в случае зашумленных данных. [1]

Цель выполнения лабораторной работы

Цель выполнения лабораторной работы – овладеть методами Быстрого преобразования Фурье и полиномиальной регрессии. На практике оценить их эффективность и сделать выводы о свойствах методов.

Задачи, выполненные в процессе реализации лабораторной работы

1. Выведено общее выражение для формулы численного интегрирования.
2. Разработаны функции *fft_coeff(y_nodes)*, которая вычисляет и возвращает комплексные коэффициенты тригонометрического полинома и *spectral_integral(f, N)*, которая вычисляет значение интеграла функции *f*, интерполируемой в *N* узлах.
3. Для каждого из интегралов I_1 и I_2 проведен анализ и произведена оценка порядка точности интегрирования.
4. Разработана функция *poly_regression(x_nodes, y_nodes, degree)*, которая возвращает коэффициенты многочлена, наилучшим образом приближающегося к точкам с абсциссами *x_nodes* и ординатами *y_nodes*.
5. Сгенерированы начальный набор данных $D_{regr}^{(N)}$ и проверочный набор данных $D_{test}^{(N)}$. С помощью набора данных $D_{regr}^{(N)}$ построен многочлен степени *p*, наилучшим образом приближающийся к данным, вычислена погрешность. Выведена на экран несколько характерных графиков.
6. Сделаны выводы о свойствах полиномиальной регрессии.

1. Быстрое преобразование Фурье

Вывод общего выражения для формулы численного интегрирования:

Пусть $\int_{-\pi/4}^{\pi} f(x) dx$ – рассматриваемый интеграл. Заменим $f(x)$ тригонометрическим интерполянт $a_0 + \sum_{k=1}^n (a_k \cdot \cos kx + b_k \cdot \sin kx)$.

Тогда получим выражение

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

Преобразуем:

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

Раскрываем интегралы:

$$\int_{-\pi/4}^{\pi} f(x) dx = a_0 \cdot \frac{5\pi}{4} + \sum_{k=1}^n (a_k \frac{1}{k} \cdot \sin \frac{k\pi}{4} + b_k \frac{1}{k} \cdot (-\cos k\pi + \cos \frac{k\pi}{4})). \quad (3)$$

Т.к. $a_0 = \text{Re}(\hat{a}_0)$, $a_k = 2\text{Re}(\hat{a}_k)$, $b_k = -2\text{Im}(\hat{a}_k)$ [2], выражение (3) примет вид:

$$\int_{-\pi/4}^{\pi} f(x) dx = \frac{5\pi}{4} \cdot \text{Re}(\hat{a}_0) + \sum_{k=1}^n \left(\frac{2\text{Re}(\hat{a}_k)}{k} \cdot \sin \frac{k\pi}{4} - \frac{2\text{Im}(\hat{a}_k)}{k} \cdot (\cos \frac{k\pi}{4} - \cos k\pi) \right).$$

Реализована функция, которая вычисляет коэффициенты тригонометрического полинома, используя алгоритм Кули-Тьюки, а также функция, вычисляющая значение интеграла некоторой функции с помощью тригонометрического ряда.

Для каждого из интегралов I_1 и I_2 были найдены точные значения с помощью библиотечной функции `sympy.integrate`: $I_1 = 5.2432$, $I_2 = -1.8573$. С помощью разработанной функции `spectral_integral` были найдены приближенные значения интегралов при значениях количества узлов $N = 2^n$, где $n \in 1, \dots, 8$. Например, при $N = 128$ приближенные значения I_1 и I_2 равны соответственно 5.2506 и -1.3081.

Для каждого N найдена относительная погрешность вычислений. На рисунках 1 и 2 представлены log-log графики зависимости этой погрешности от количества узлов интерполяции для интегралов I_1 и I_2 соответственно.

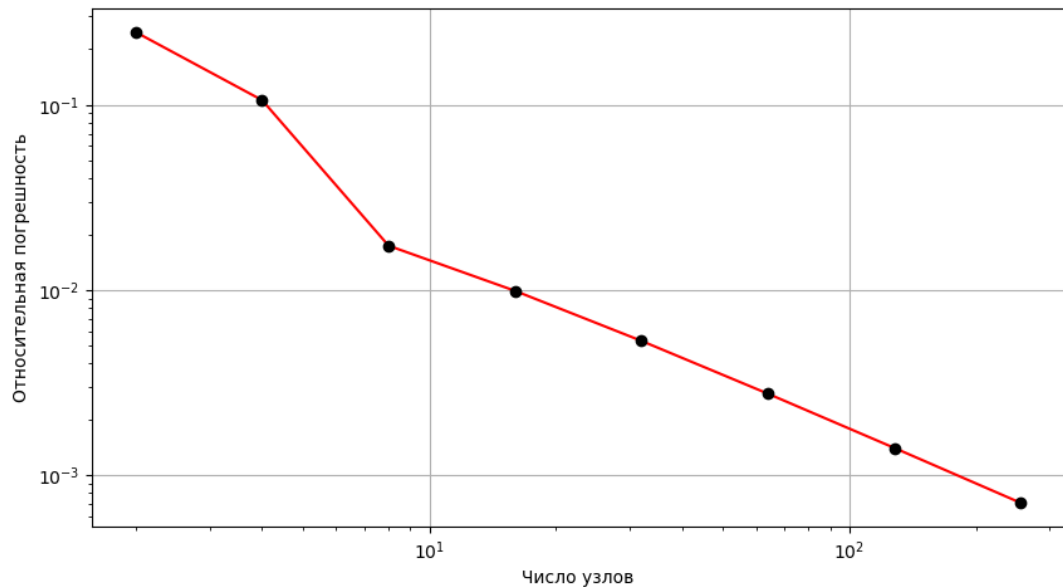


Рисунок 1. График зависимости относительной погрешности от количества узлов интерполяции для I_1

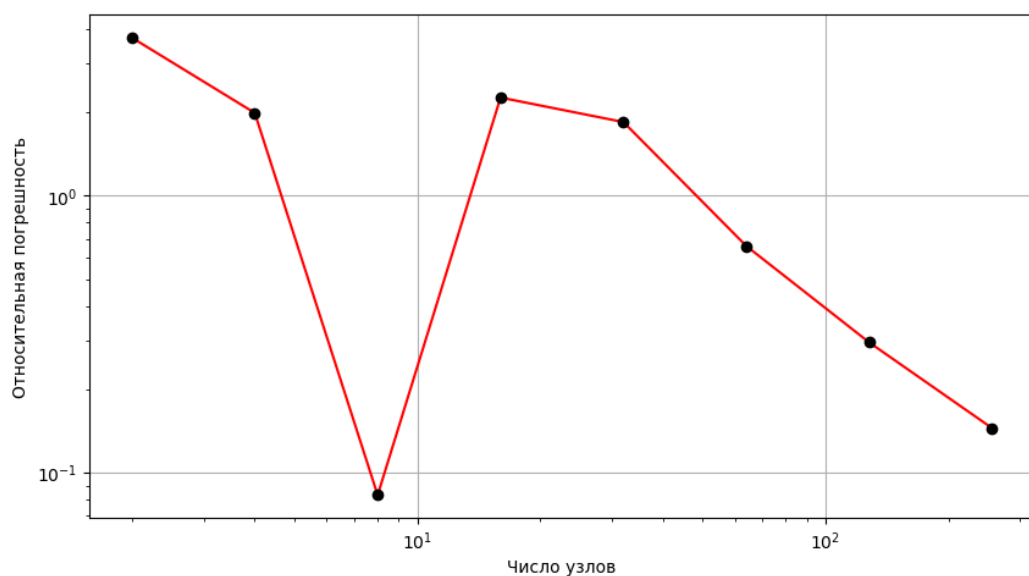


Рисунок 2. График зависимости относительной погрешности от количества узлов интерполяции для I_2

Как видно из графиков, при увеличении количества узлов, относительная погрешность вычислений уменьшается. На графике, изображенном на рисунке 2, имеется скачок при $N = 8$. Это значит, что для интеграла I_2 оптимальное количество узлов интерполирования равно 8.

Порядок точности интегрирования I_1 больше 1, что видно из графика, представленном на рисунке 1 (при увеличении количества узлов в n раз, погрешность уменьшается в $k > n$ раз). Из графика 2 следует, что порядок точности интегрирования I_2 меньше, чем у I_1 .

Различие порядков точности интегрирования I_1 и I_2 связано с тем, что функция интеграла I_1 $f(x) = |x|$ - четная и непрерывная на отрезке $[-\pi; \pi]$. Согласно теореме Вейерштрасса для функции, обладающей такими свойствами, существует тригонометрический многочлен, который равномерно сходится к исходной функции.[3] Функция интеграла I_2 не обладает этими свойствами, к ней теорема Вейерштрасса не применима, поэтому точность её интегрирования ниже.

2. Регрессия

В ходе работы была разработана функция, которая возвращает коэффициенты многочлена заданной степени, наилучшим образом приближающегося к данным. Для каждого из заданных σ , N и p сгенерирован начальный набор данных $D_{regr}^{(N)}$ и проверочный набор данных $D_{test}^{(N)}$. С помощью наборов $D_{regr}^{(N)}$ получены многочлены, наилучшим образом приближающиеся к данным. На рисунках 3-6 показаны некоторые из получившихся многочленов вместе с начальными и проверочными данными.

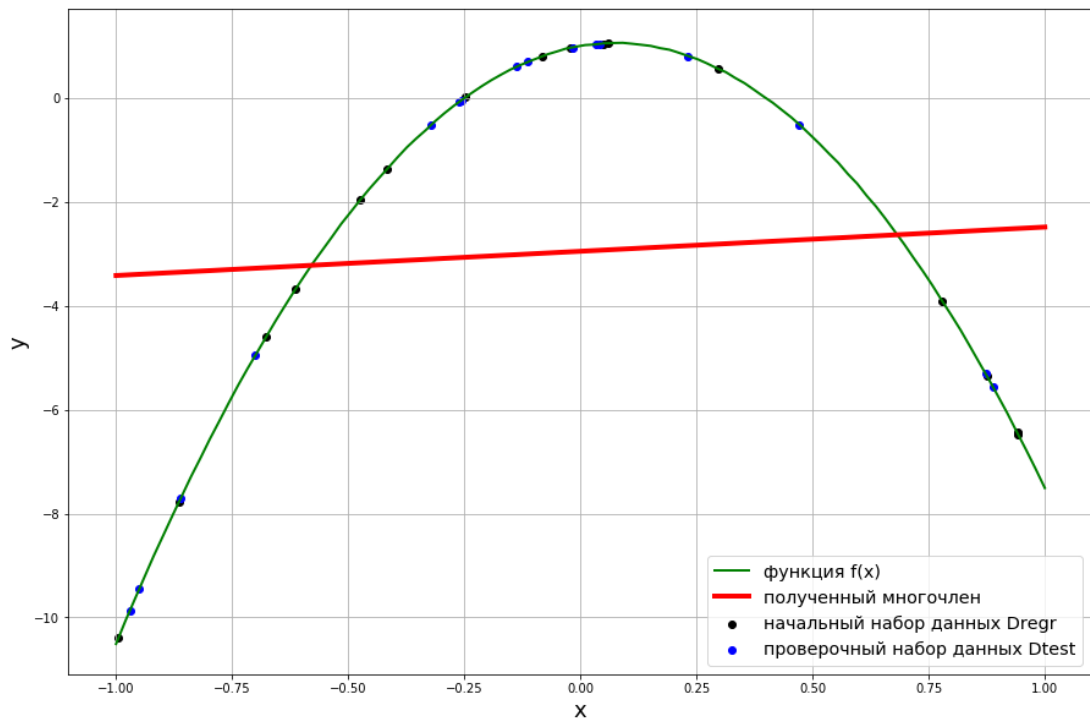


Рисунок 3. График многочлена степени $p=1$, построенный по количеству узлов $N=2^4$, $\sigma=0.01$

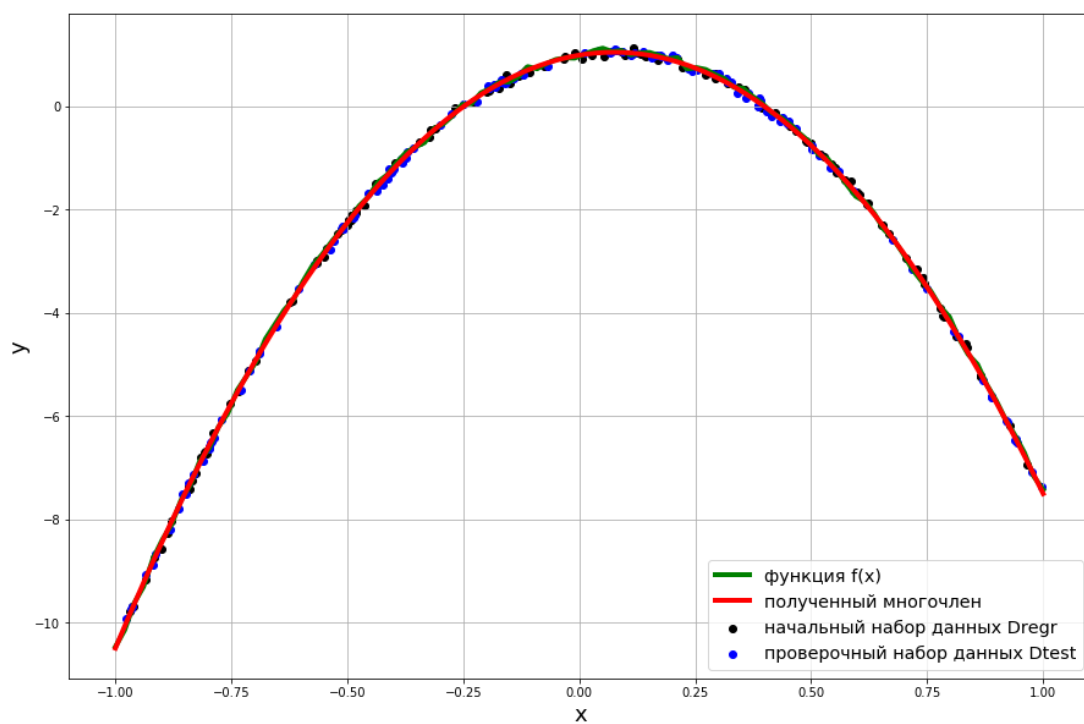


Рисунок 4. График многочлена степени $p=2$, построенный по количеству узлов $N=2^7$, $\sigma=0.1$

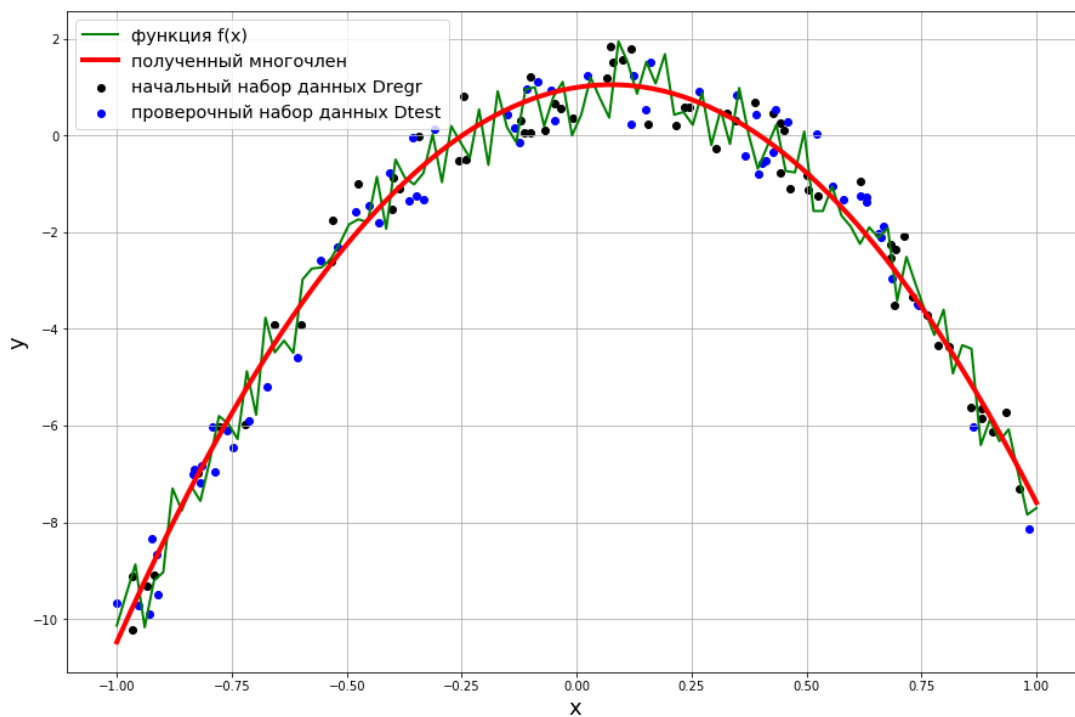


Рисунок 5. График многочлена степени $p=3$, построенный по количеству узлов $N=2^6$, $\sigma=1$

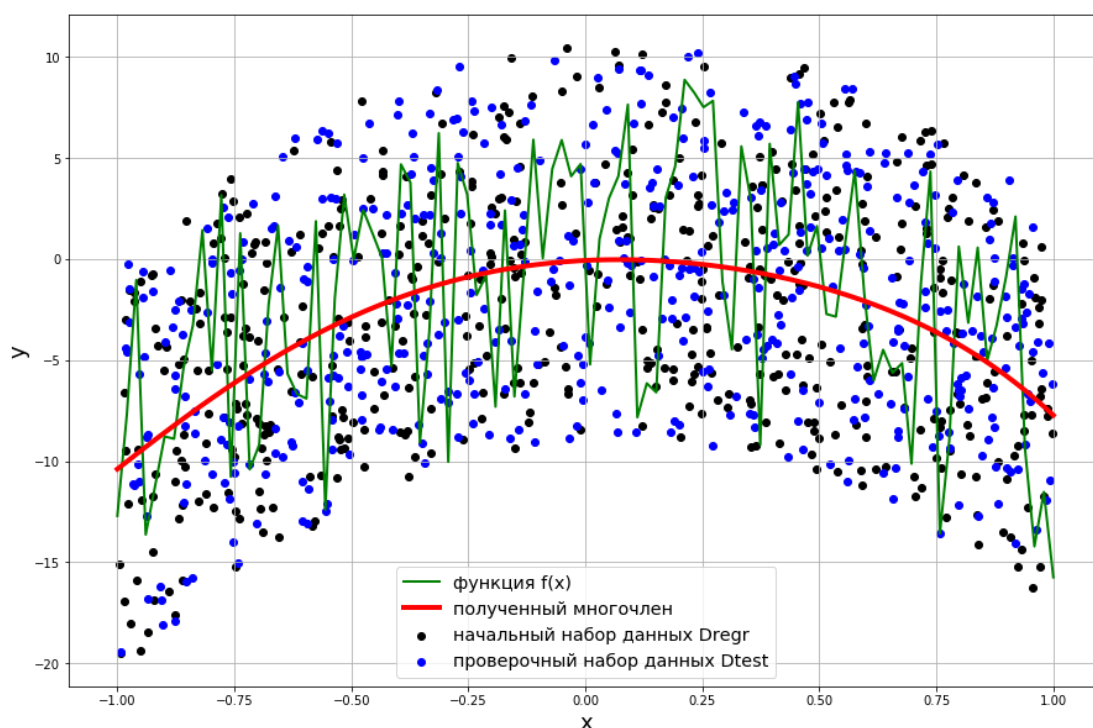


Рисунок 6. График многочлена степени $p=5$, построенный по количеству узлов $N=2^9$, $\sigma=10$

Вычислена среднеквадратичная погрешность аппроксимации $\epsilon_{regr}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ по формуле:

$$\epsilon_{regr}^{(N,p)} = \sqrt{\frac{\sum_{i=1}^N \left(y_i(D_{regr}^{(N)}) - P(x_i) \right)^2}{N}}, \quad (4)$$

где $y_i(D_{regr}^{(N)})$ – i -ое значение функции в наборе $D_{regr}^{(N)}$, $P(x_i)$ – значение полинома в i -ом узле. Погрешность для проверочных данных находится аналогично.

На рисунке 7 представлен график зависимости среднеквадратичных погрешностей $\epsilon_{regr}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ от количества начальных и проверочных данных N .

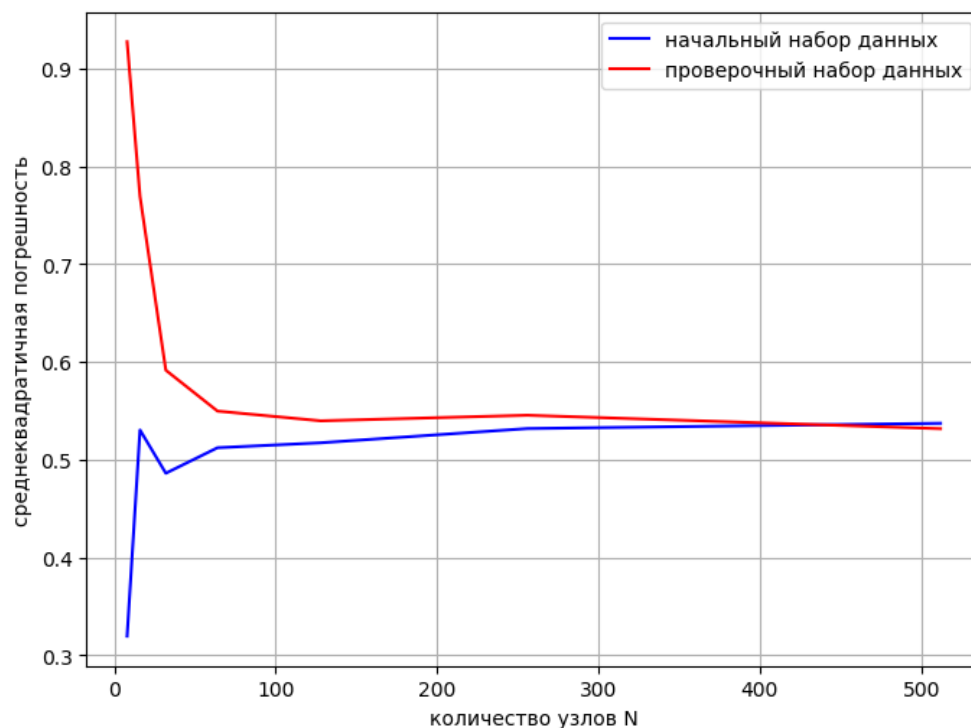


Рисунок 7. Графики зависимости среднеквадратичных погрешностей $\epsilon_{reg}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ от количества начальных и проверочных данных N при многочлене степени $p=3$, $\sigma=1$

Из графика следует, что при малых значениях N $\epsilon_{reg}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ различны, но при увеличении числа данных N значения $\epsilon_{reg}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ стремятся к одному числу.

Влияние степени многочлена p на значения среднеквадратичных погрешностей можно оценить по графикам, изображенным на рисунках 8 и 9.

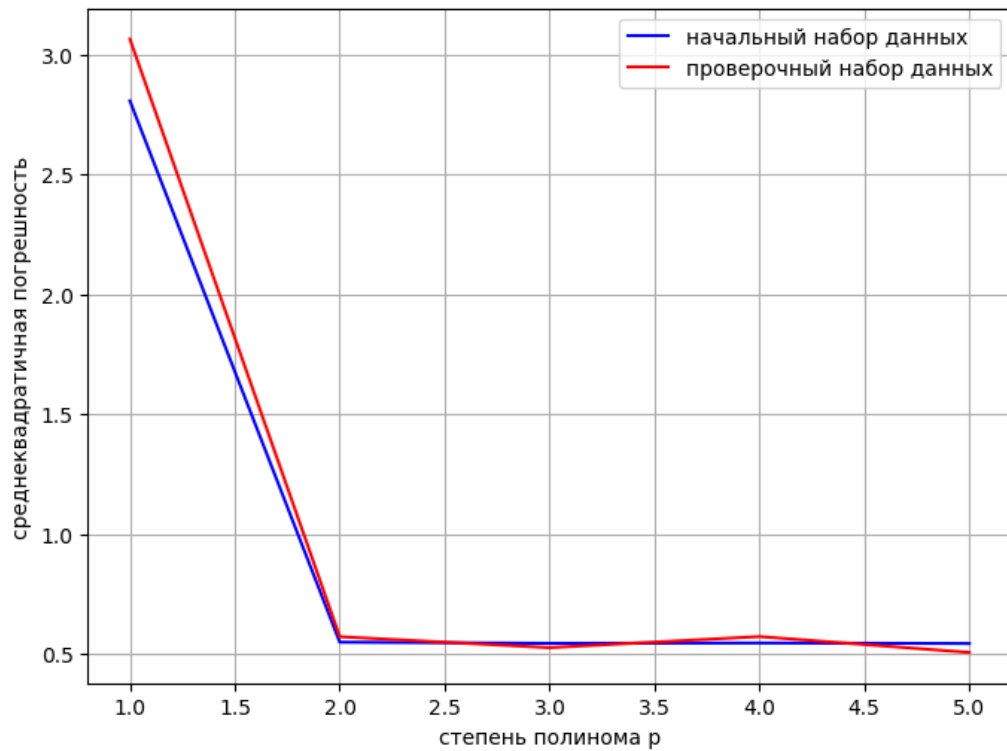


Рисунок 8. Графики зависимости среднеквадратичных погрешностей $\epsilon_{reg}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ от степени полинома p при количестве данных $N=2^7$, $\sigma=1$

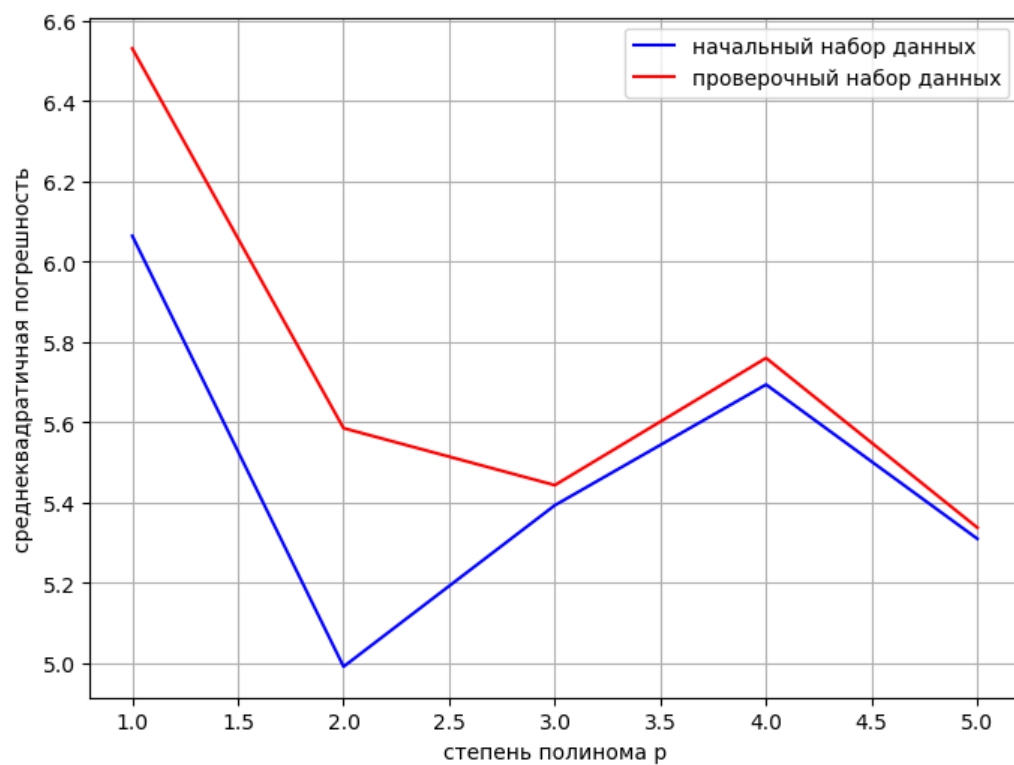


Рисунок 9. Графики зависимости среднеквадратичных погрешностей $\epsilon_{reg}^{(N,p)}$ и $\epsilon_{test}^{(N,p)}$ от степени полинома p при количестве данных $N=2^7$, $\sigma=10$

При малых σ наблюдается уменьшение среднеквадратичной погрешности при увеличении степени полинома. При этом значительное уменьшение погрешности происходит только при изменении степени p с 1 до 2. Поэтому при небольших σ следует выбирать полином 2 степени.

На рисунке 10 представлен график зависимости относительной погрешности коэффициентов многочлена от N при фиксированном σ , а на рисунке 11 – при увеличении σ (в 10 раз при увеличении N в 2 раза).

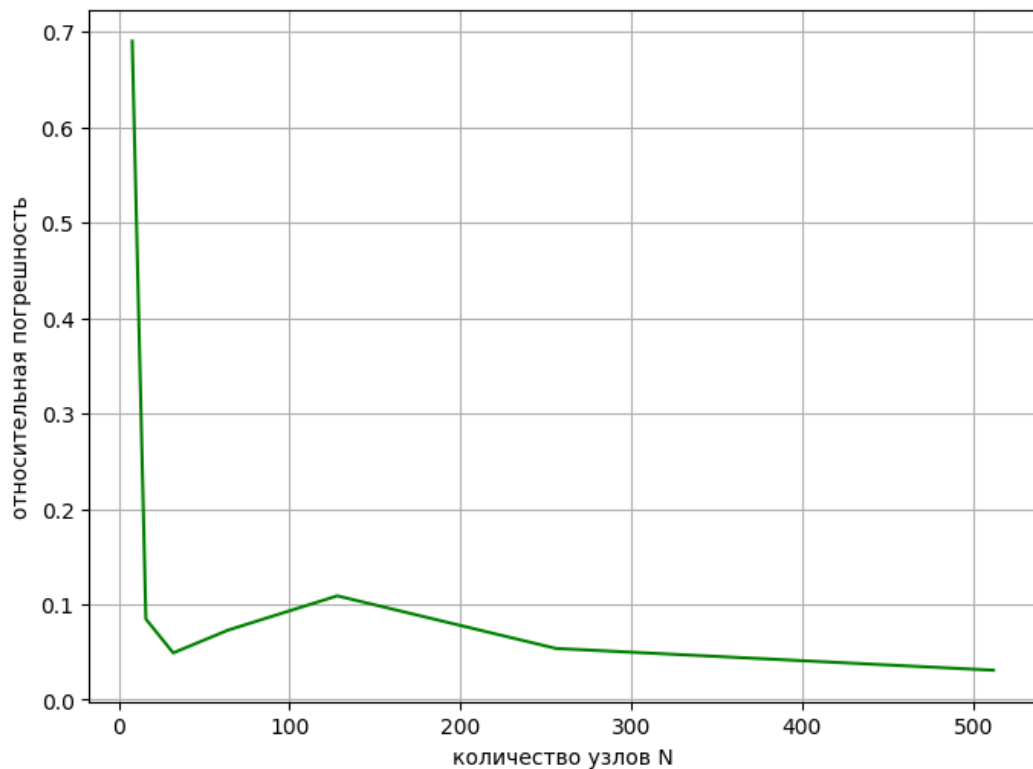


Рисунок 10. График зависимости относительной погрешности коэффициентов многочлена от N при фиксированном $\sigma=1$, $p=2$

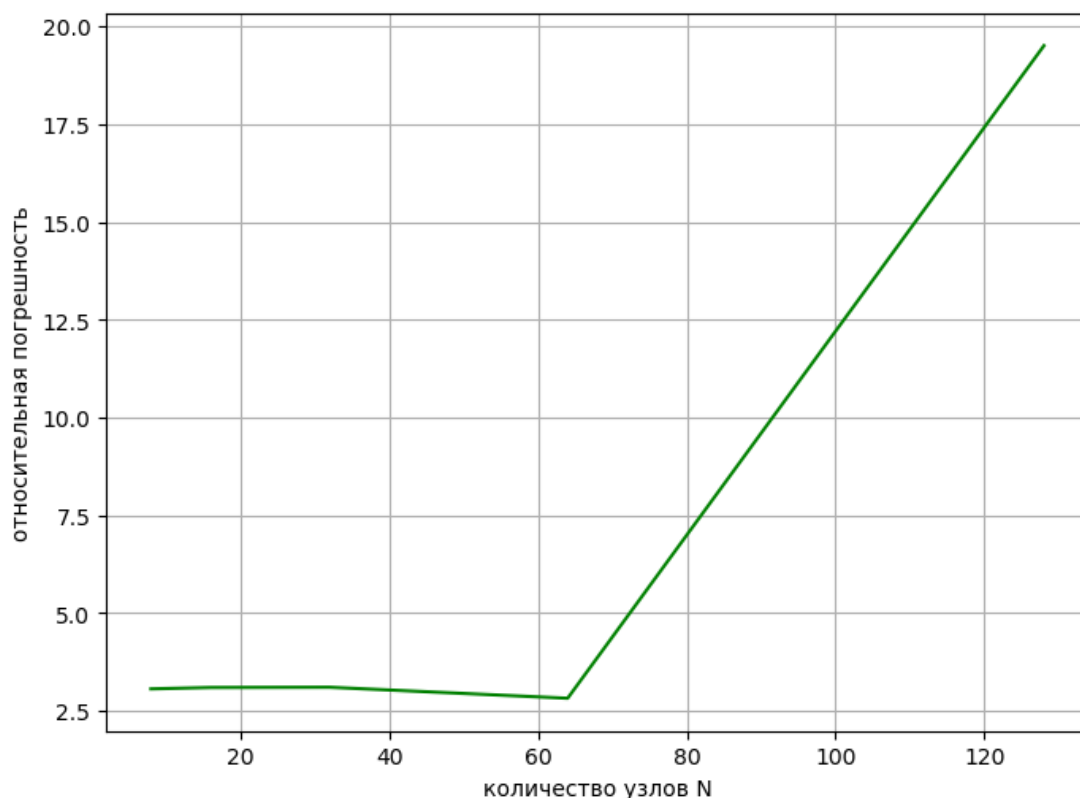


Рисунок 11. График зависимости относительной погрешности коэффициентов многочлена от N при увеличении σ , $p=2$

При фиксированном σ относительная погрешность коэффициентов многочлена степени 2 не превышает 1 и уменьшается с увеличением количества узлов. При увеличении σ погрешность становится значительно больше (от 2.5 до 20) и увеличивается при увеличении числа данных N . Поэтому при использовании полиномиальной регрессии величина σ должна быть постоянна, чтобы обеспечить точность вычислений.

Изменим значения величины X : новый интервал $[-15;3]$, математическое ожидание = 1, среднеквадратичное отклонение = 5. Графики зависимости относительной погрешности коэффициентов многочлена от N при постоянном σ и увеличении σ показаны на рисунках 12 и 13 соответственно.

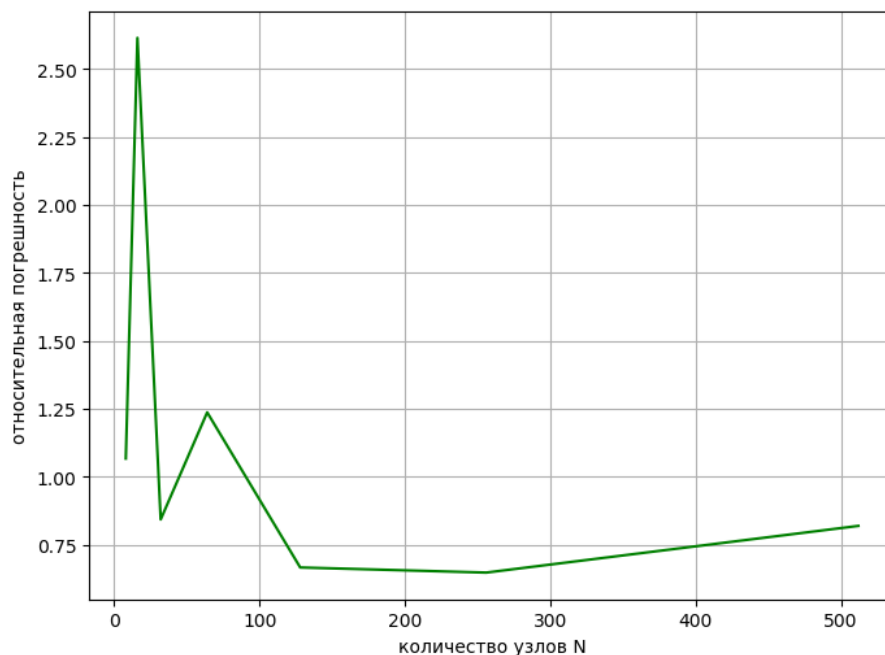


Рисунок 12. График зависимости относительной погрешности коэффициентов многочлена от N при фиксированном $\sigma=1$, $p=2$ при новых значениях X

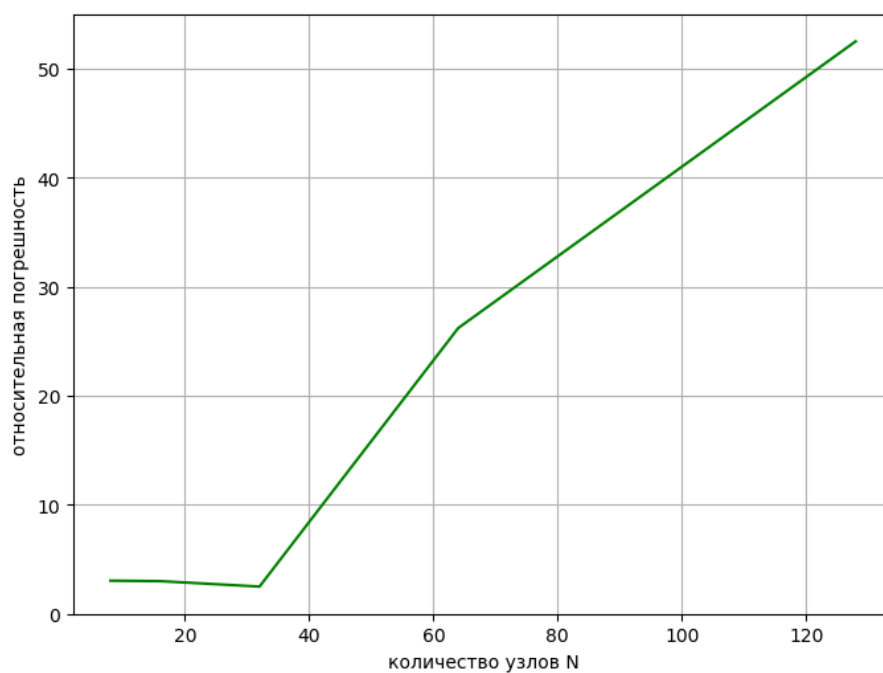


Рисунок 13. График зависимости относительной погрешности коэффициентов многочлена от N при увеличении σ , $p=2$ при новых значениях X

Относительная погрешность значительно возросла (примерно в 2.5 раза), график при фиксированном значении σ стал более хаотичен, т.к. при увеличении значений X данные становятся более зашумленными, что влияет на точность приближения

Заключение

В случае зашумленных данных, которые имеют нормальное распределение, с помощью полиномиальной регрессии можно получить приближение, погрешность которого уменьшается с увеличением числа данных. Метод регрессии следует использовать при большом количестве зашумленных данных.

Список использованных источников

[1] Соколов, А.П., Першин, А.Ю. Инструкция по выполнению лабораторной работы (общая). Соколов, А.П., Першин, А.Ю., Москва, 2018.

[2] Першин, А.Ю. Лекции по вычислительной математике. Першин А.Ю., Москва, 2019.

[3] Теорема Вейерштрасса – Стоуна [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Теорема_Вейерштрасса_—_Стоуна – (Дата обращения: 29.04.2019).