Санкт-Петербургский политехнический университет Петра Великого

Кафедра компьютерных систем и программных технологий

**Отчёт по лабораторной работе**

**Дисциплина**: Теория вероятностей

**Тема**: Аппроксимация результатов измерений зависимых переменных

Выполнил студент гр. 23531/2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ М.А. Хираев

(подпись)

Преподаватель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ К.В.Никитин

(подпись)  
“\_\_\_” \_\_\_\_\_\_\_\_\_\_\_\_ 2018 г.

Санкт-Петербург  
2018

1. **Обработка входных данных**

Имеются k выборок измерений одной величины n раз.

Исходные данные, высчитанные средние значения функций и дисперсии представлены в таблице 1.1.

*Таблица 1.1. Исходные данные и вычисленные оценки*

|  |  |  |  |
| --- | --- | --- | --- |
| k |  |  |  |
| 1 | -2.0 | -0.036478 | 0.000165 |
| 2 | -1.9 | -0.300231 | 1.063396 |
| 3 | -1.8 | -0.665227 | 1.660100 |
| 4 | -1.7 | -0.448739 | 5.851819 |
| 5 | -1.6 | 0.572629 | 9.034255 |
| 6 | -1.5 | 0.415652 | 37.376913 |
| 7 | -1.4 | -2.928039 | 14.869191 |
| 8 | -1.3 | 2.710479 | 35.706984 |
| 9 | -1.2 | -2.864644 | 31.927609 |
| 10 | -1.1 | 3.305898 | 61.869550 |
| 11 | -1.0 | -3.225060 | 42.597944 |
| 12 | -0.9 | -1.950364 | 71.448019 |
| 13 | -0.8 | -4.409246 | 51.126642 |
| 14 | -0.7 | -4.302238 | 66.290990 |
| 15 | -0.6 | -5.498972 | 69.146790 |
| 16 | -0.5 | -7.791506 | 145.893210 |
| 17 | -0.4 | -4.601402 | 64.180584 |
| 18 | -0.3 | -4.015633 | 99.017955 |
| 19 | -0.2 | -4.410198 | 56.063515 |
| 20 | -0.1 | -3.499420 | 101.583467 |
| 21 | 0 | 0.223196 | 48.525925 |
| 22 | 0.1 | 3.270602 | 101.196342 |
| 23 | 0.2 | 7.536747 | 124.027019 |
| 24 | 0.3 | 13.635946 | 40.036494 |
| 25 | 0.4 | 19.478865 | 55.142826 |
| 26 | 0.5 | 24.806730 | 16.774168 |
| 27 | 0.6 | 24.817750 | 13.590551 |
| 28 | 0.7 | 25.305840 | 10.754746 |
| 29 | 0.8 | 21.828090 | 6.555524 |
| 30 | 0.9 | 16.229010 | 6.244343 |
| 31 | 1.0 | 5.844848 | 1.726559 |
| 32 | 1.1 | -6.852843 | 0.081840 |
| 33 | 1.2 | -20.597440 | 0.403911 |
| 34 | 1.3 | -34.986570 | 0.655973 |
| 35 | 1.4 | -49.198010 | 1.843639 |
| 36 | 1.5 | -57.770910 | 12.644222 |
| 37 | 1.6 | -64.804890 | 15.109186 |
| 38 | 1.7 | -65.408150 | 10.863003 |
| 39 | 1.8 | -62.758860 | 27.760981 |
| 40 | 1.9 | -40.126840 | 20.400706 |
| 41 | 2.0 | -24.723520 | 46.864831 |

Средние значения были высчитаны по следующей формуле:

Матрица была посчитана по следующей формуле:

Были высчитаны только диагональные элементы, являющиеся дисперсиями.

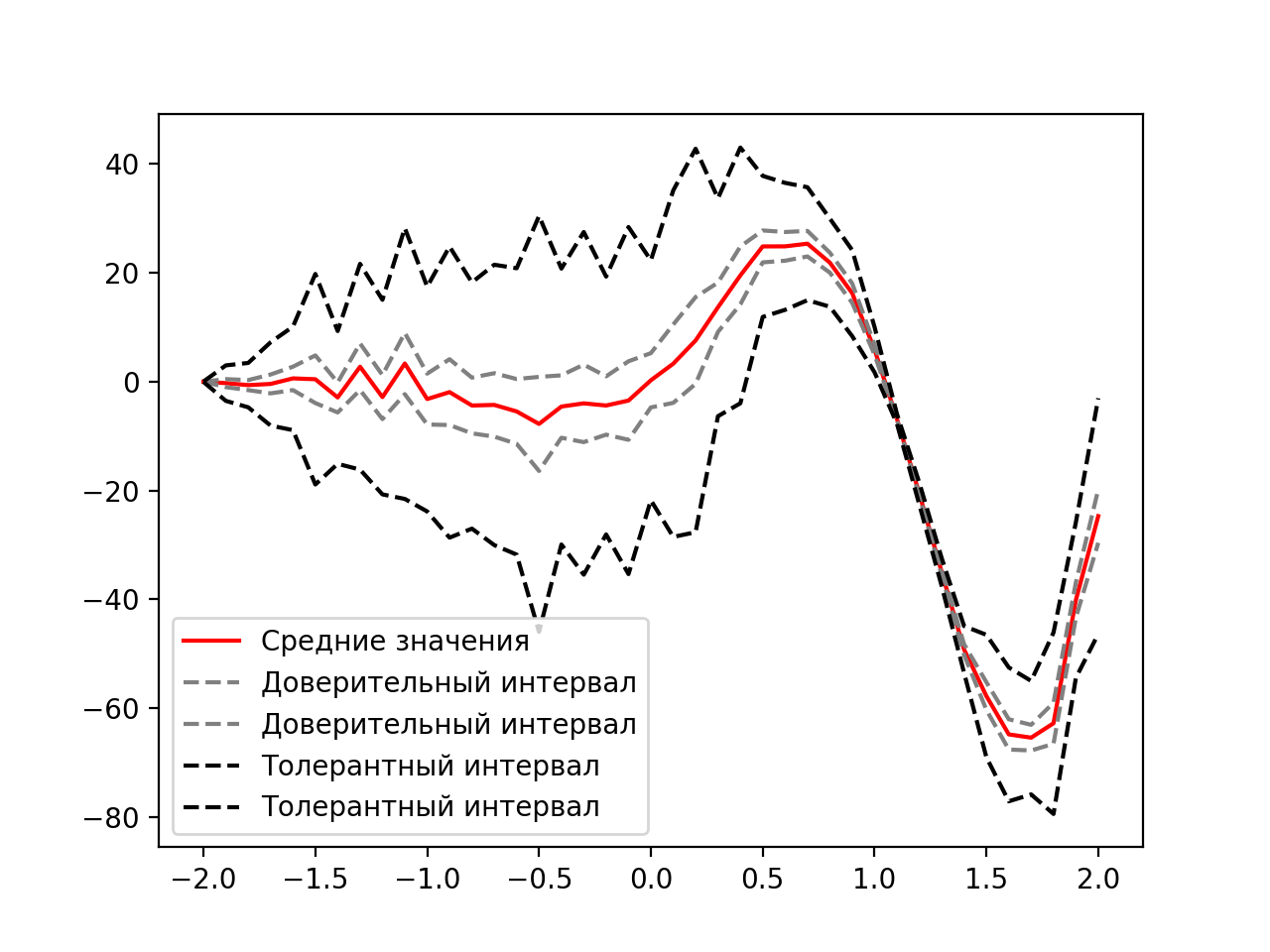
1. **Критерий Кохрена**

Критическое значение критерия Кохрена взято из таблицы и равно (при )

Статистика посчитана по формуле:

В итоге вышло, что

Следовательно, необходимо использовать ОМНК.

Были высчитаны доверительные интервалы и толерантные пределы при заданной доверительной вероятности . Доверительные интервалы и толерантные пределы представлены на рисунке 2.1

*Рисунок 2.1. График доверительных интервалов и толерантных пределов*

1. **Аппроксимация**

В качестве кандидатов были использованы полиномы разной степени от 1 до 39. Ограничение степени 39 является причиной формулы, в которой знаменатель обращается в ноль, когда степень полинома достигает цифры .

Пусть – степень полинома, .

Подсчет коэффициентов полинома

Где

Расчет статистики производится по формуле:

Использование именно этой формулы объясняется тем, что кол-во точек больше кол-ва измерений в каждой точке [Г.П. Солопченко. Теория вероятностей и математическая статистика. Стр. 185]. Именно поэтому нельзя использовать полином степени больше .

**Критическое значение** является квантилью распределения Фишера с количеством степеней свободы и . Использование таких степеней свободы обусловлено тем, что количество измерений в каждой точке меньше количества точек. Критические значения и статистики Фишера представлены в таблице 3.1.

*Таблица 3.1. Критические значения и статистики Фишера*

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 9 | 1.7128429767898328 | 1.4601194231937547 |
| 10 | 1.7140609891666814 | 1.402737563356054 |
| 11 | 1.715355592267296 | 1.44743607586524 |
| 12 | 1.7167342508682548 | 1.3605189126336596 |
| 13 | 1.718205432921014 | 1.3301761413607482 |
| 14 | 1.7197787839648027 | 1.3246119032706432 |
| 15 | 1.721465339230149 | 1.3769830192994135 |
| 16 | 1.7232777832826314 | 1.2189080964439714 |
| 18 | 1.727341320690615 | 1.53849062605748 |
| 19 | 1.7296293208669171 | 1.3653043583728384 |
| 21 | 1.734835483636008 | 1.2349330499639113 |
| 26 | 1.75327999122335 | 1.3851089252432043 |

Коэффициенты некоторых полиномов, которые прошли критерий Фишера представлены в таблице 3.2.

*Таблица 3.2. Коэффициенты полиномов, прошедших критерий Фишера*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | q | | | | | | | |
| 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| a0 | 1.3466 | 0.4785 | 0.5280 | -0.3615 | -0.0961 | 0.4732 | 0.5033 | -0.6922 |
| a1 | 41.0187 | 39.9342 | 39.3005 | 37.0677 | 33.4654 | 35.0827 | 34.6673 | 31.4080 |
| a2 | 41.6020 | 49.1772 | 48.9313 | 60.7130 | 57.3984 | 46.6872 | 46.0746 | 76.0685 |
| a3 | -41.3943 | -39.8939 | -37.2458 | -30.9925 | -9.5690 | -16.2421 | -12.8130 | 6.5276 |
| a4 | -58.0966 | -70.9560 | -70.8847 | -100.723 | -94.4663 | -56.5251 | -54.7770 | -195.321 |
| a5 | 0.5459 | 0.2897 | -3.0738 | -7.7701 | -46.6426 | -38.5565 | -47.1665 | -81.0912 |
| a6 | 18.8685 | 27.0909 | 27.2397 | 56.5633 | 52.5731 | -0.3888 | -2.2029 | 261.1287 |
| a7 | 3.7502 | 3.5912 | 5.4140 | 6.5097 | 37.8382 | 33.9613 | 43.7712 | 67.7706 |
| a8 | -1.7879 | -4.0178 | -4.1016 | -17.5657 | -16.6298 | 19.5543 | 20.4064 | -228.824 |
| a9 | -0.5083 | -0.4707 | -0.9138 | -0.8958 | -13.4395 | -12.7612 | -18.6035 | -25.2106 |
| a10 |  | 0.2163 | 0.2278 | 3.1341 | 3.0987 | -9.8034 | -9.9879 | 120.4302 |
| a11 |  |  | 0.0397 | 0.0177 | 2.4619 | 2.4710 | 4.3559 | 4.3323 |
| a12 |  |  |  | -0.2381 | -0.2467 | 2.0582 | 2.0731 | -36.1754 |
| a13 |  |  |  |  | -0.1848 | -0.1944 | -0.5063 | -0.1954 |
| a14 |  |  |  |  |  | -0.1628 | -0.1628 | 5.7217 |
| a15 |  |  |  |  |  |  | 0.0207 | -0.0174 |
| a16 |  |  |  |  |  |  |  | -0.3695 |

Пример полученного полинома 9 степени:

График зависимости статистики Фишера от степени полинома представлены на рисунке 3.1.



*Рисунок 3.1. График зависимости статистики Фишера от степени полинома*

Для первого принятого вектора коэффициентов полинома со степенью 9 была построена ковариационная и корреляционная матрицы оценок коэффициентов.

Ковариационная матрица была получена по следующей формуле:

Ковариационная матрица представлена в таблице 3.3.

*Таблица 3.3. Ковариационная матрица для полинома степени 9*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.7966 | -0.5724 | -1.9949 | 1.3928 | 1.5106 | -0.9877 | -0.4515 | 0.2700 | 0.0467 | -0.0252 |
| -0.5724 | 4.2411 | 0.2774 | -8.5493 | 0.6644 | 5.8596 | -0.4400 | -1.6621 | 0.0676 | 0.1669 |
| -1.9949 | 0.2774 | 7.8694 | -3.0110 | -7.1534 | 2.7571 | 2.3827 | -0.8184 | -0.2654 | 0.0774 |
| 1.3928 | -8.5493 | -3.0110 | 19.9906 | 0.7350 | -14.8057 | 0.3274 | 4.3923 | -0.0896 | -0.4534 |
| 1.5106 | 0.6644 | -7.1534 | 0.7350 | 7.1607 | -1.1105 | -2.5513 | 0.3326 | 0.2988 | -0.0265 |
| -0.9877 | 5.8596 | 2.7571 | -14.8057 | -1.1105 | 11.4655 | -0.0774 | -3.5019 | 0.0526 | 0.3689 |
| -0.4515 | -0.4400 | 2.3827 | 0.3274 | -2.5513 | -0.0774 | 0.9557 | 0.0320 | -0.1163 | -0.0072 |
| 0.2700 | -1.6621 | -0.8184 | 4.3923 | 0.3326 | -3.5019 | 0.0320 | 1.0930 | -0.0182 | -0.1171 |
| 0.0467 | 0.0676 | -0.2654 | -0.0896 | 0.2988 | 0.0526 | -0.1163 | -0.0182 | 0.0146 | 0.0025 |
| -0.0252 | 0.1669 | 0.0774 | -0.4534 | -0.0265 | 0.3689 | -0.0072 | -0.1171 | 0.0025 | 0.0127 |

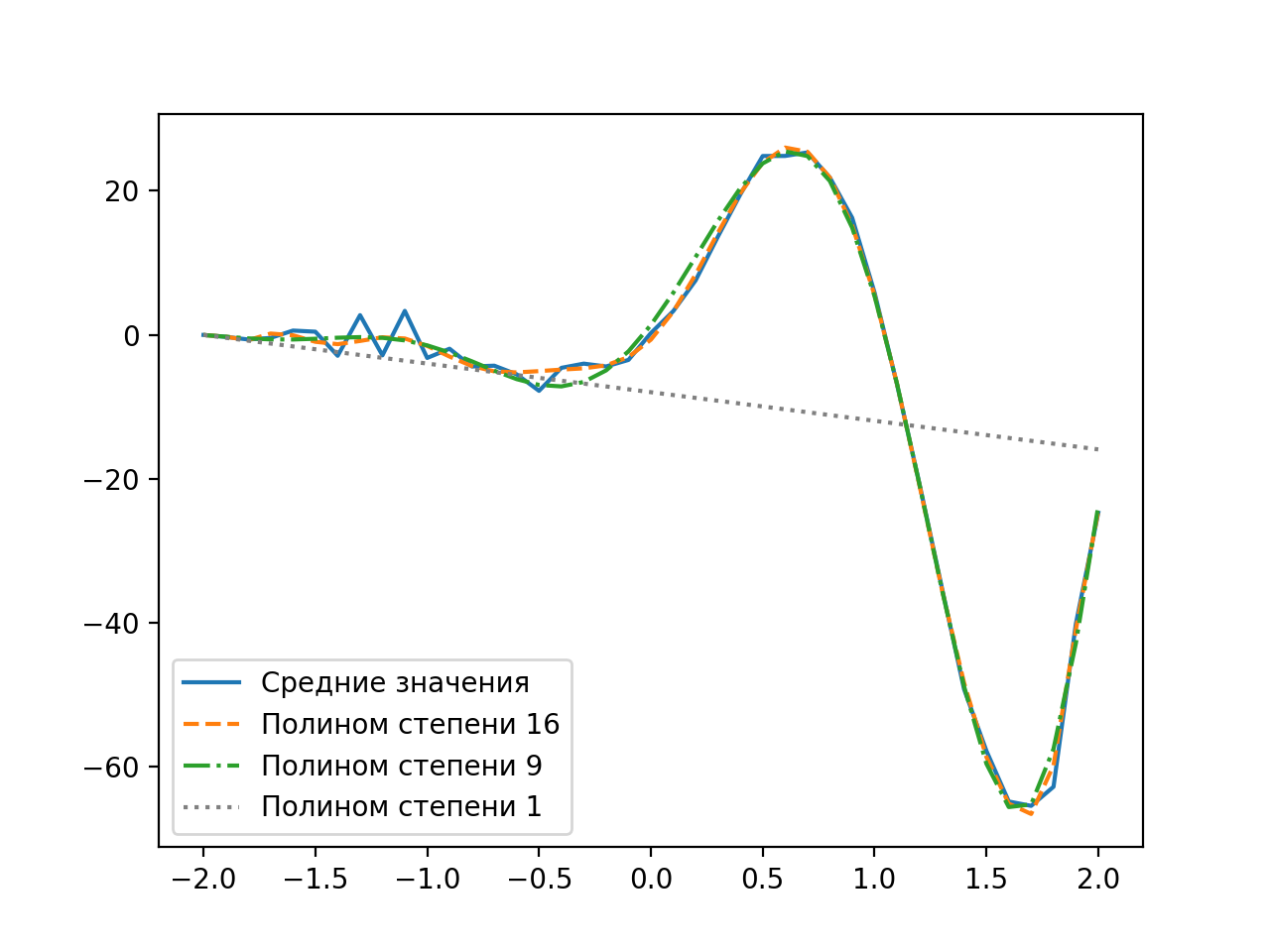
Построим корреляционную матрицу по формуле:

Корреляционная матрица представлена в таблице 3.4.

*Таблица 3.4. Корреляционная матрица для полинома степени 9*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1.0000 | -0.3114 | -0.7968 | 0.3490 | 0.6325 | -0.3268 | -0.5175 | 0.2894 | 0.4336 | -0.2504 |
| -0.3114 | 1.0000 | 0.0480 | -0.9285 | 0.1206 | 0.8403 | -0.2185 | -0.7720 | 0.2719 | 0.7186 |
| -0.7968 | 0.0480 | 1.0000 | -0.2401 | -0.9529 | 0.2903 | 0.8689 | -0.2791 | -0.7839 | 0.2446 |
| 0.3490 | -0.9285 | -0.2401 | 1.0000 | 0.0614 | -0.9780 | 0.0749 | 0.9396 | -0.1660 | -0.8989 |
| 0.6325 | 0.1206 | -0.9529 | 0.0614 | 1.0000 | -0.1226 | -0.9753 | 0.1189 | 0.9251 | -0.0878 |
| -0.3268 | 0.8403 | 0.2903 | -0.9780 | -0.1226 | 1.0000 | -0.0234 | -0.9892 | 0.1287 | 0.9657 |
| -0.5175 | -0.2185 | 0.8689 | 0.0749 | -0.9753 | -0.0234 | 1.0000 | 0.0313 | -0.9854 | -0.0656 |
| 0.2894 | -0.7720 | -0.2791 | 0.9396 | 0.1189 | -0.9892 | 0.0313 | 1.0000 | -0.1443 | -0.9928 |
| 0.4336 | 0.2719 | -0.7839 | -0.1660 | 0.9251 | 0.1287 | -0.9854 | -0.1443 | 1.0000 | 0.1839 |
| -0.2504 | 0.7186 | 0.2446 | -0.8989 | -0.0878 | 0.9657 | -0.0656 | -0.9928 | 0.1839 | 1.0000 |

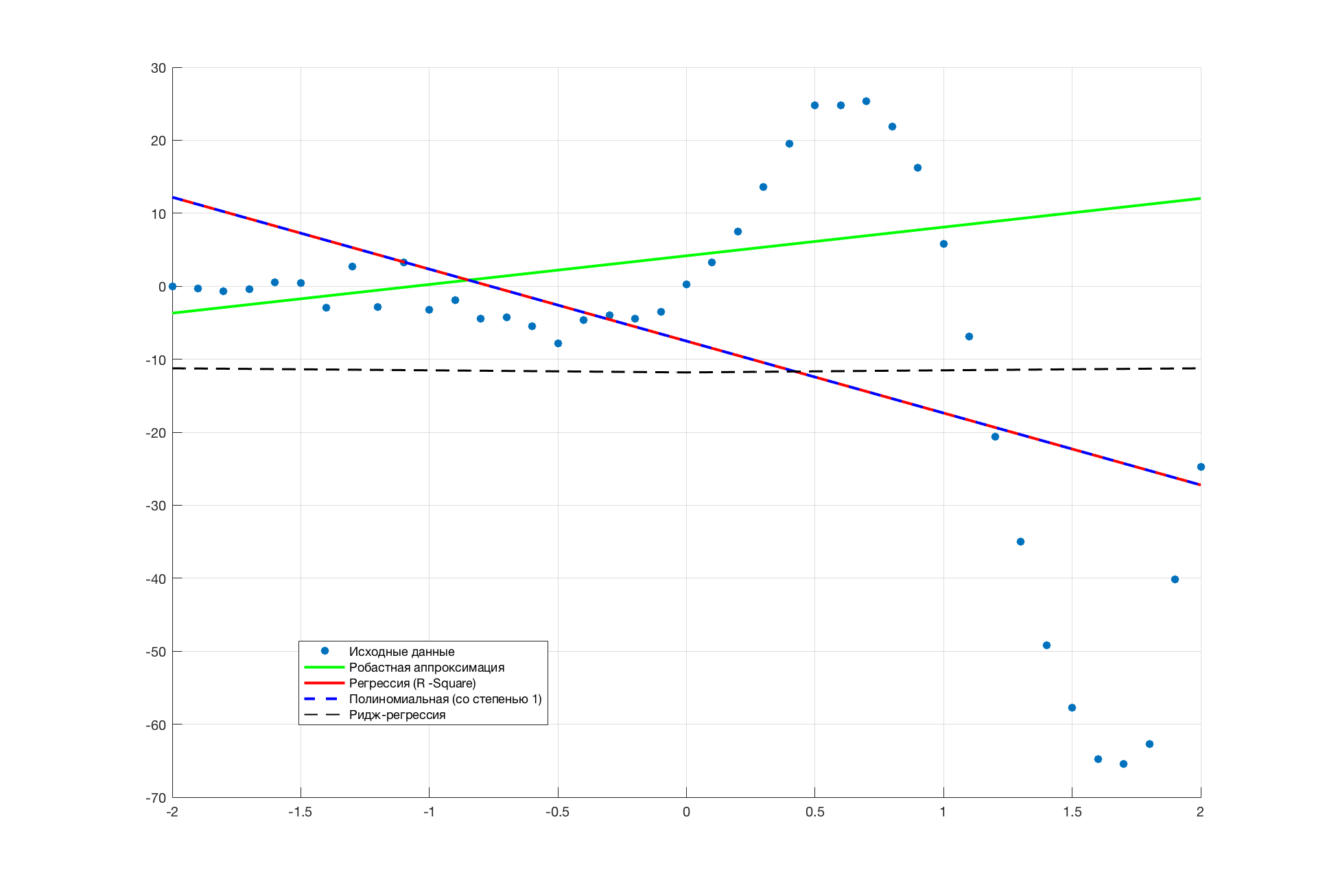
Для сравнения были построены графики для полиномов различных степеней. Графики представлены на рисунке 3.2.



*Рисунок 3.2. Полиномы различных степеней*

1. **Аппроксимация другими методами и интерполяция**

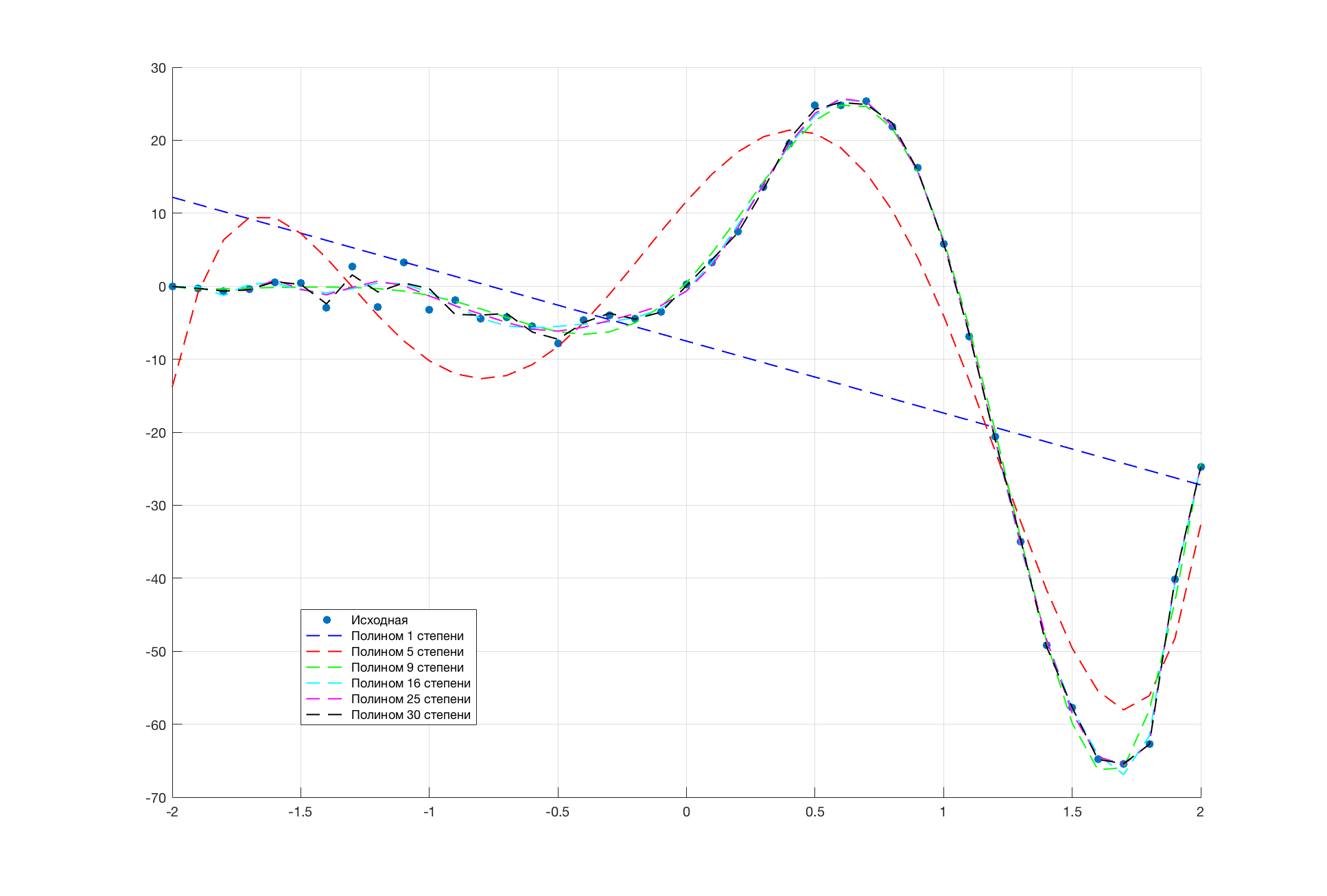
**Линейная аппроксимация** с помощью встроенных функций *MATLAB* (*robustfit*, *regress*, *polyfit* и *polyval*, *ridge*). Результаты представлены на рисунке 4.1.



*Рисунок 4.1. Графики линейной аппроксимации*

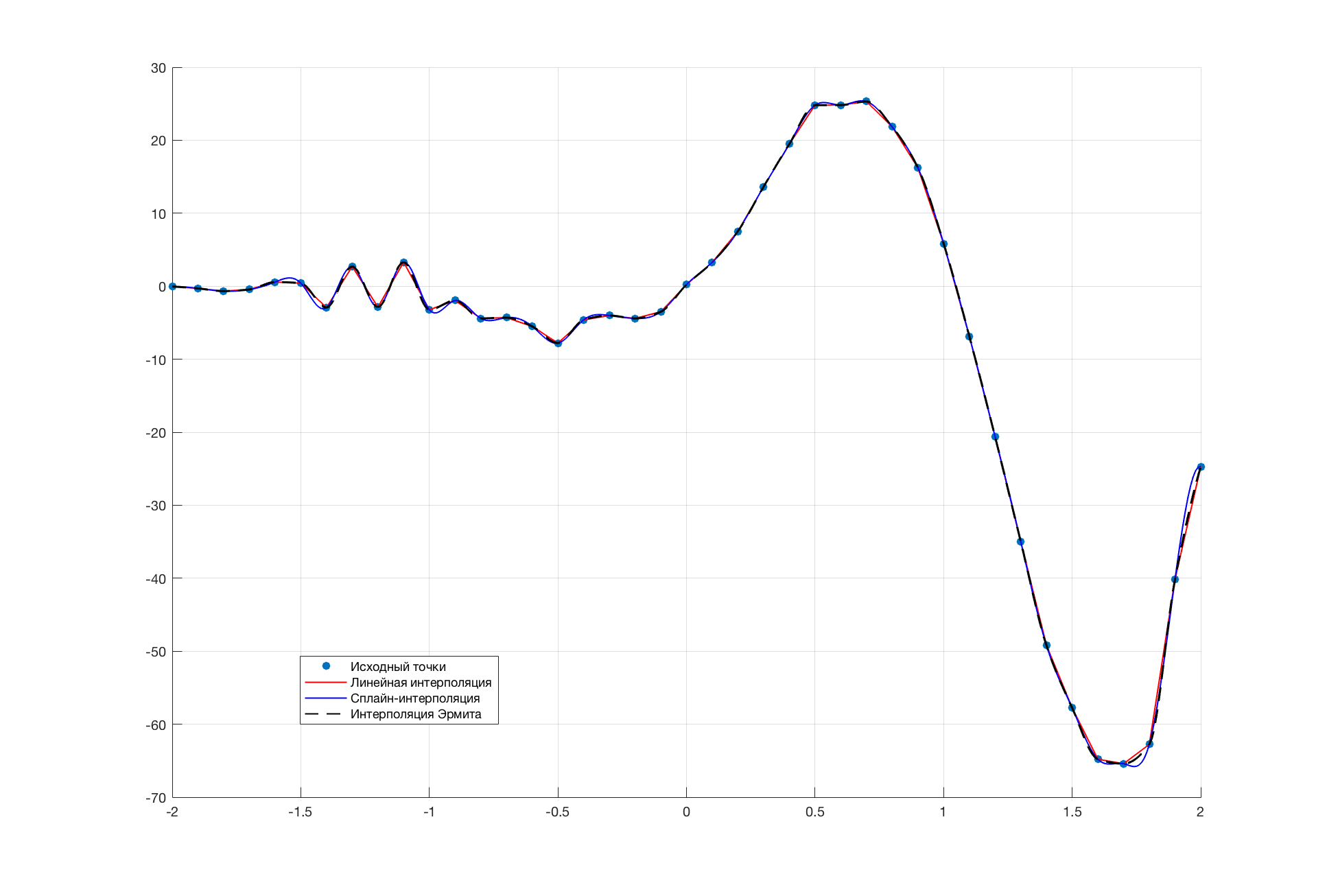
В данном случае нет смысла сравнивать линейные аппроксимации, так как исходная функция нелинейная и попытки аппроксимировать ее прямой не имеют никакого смысла.

**Полиномиальная аппроксимация** с разными степенями с помощью встроенных функций *polyfit* и *polyval*. Результаты представлены на рисунке 4.2.

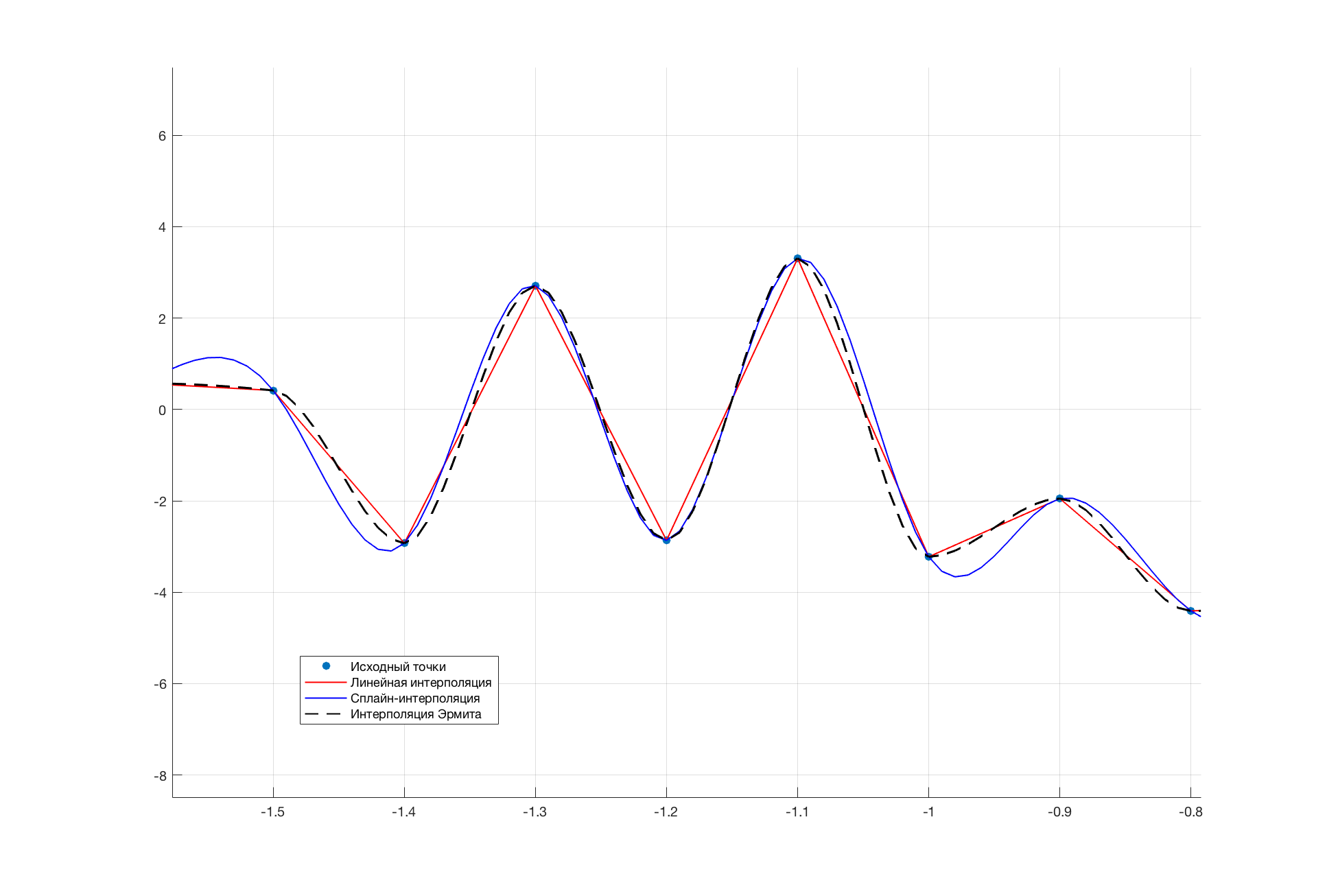
*Рисунок 4.2. Графики полиномиальной аппроксимации*

Полиномиальная аппроксимация хорошо приближает функцию, результаты такие же, какие и при самостоятельном поиске коэффициентов в предыдущих пунктах.

**Интерполяция** разными методами (линейная, сплайн и кубическая) с помощью встроенной функции *interp1* в *MATLAB*. Результаты представлены на рисунках 4.3 и 4.4.



*Рисунок 4.3. Графики интерполяции*



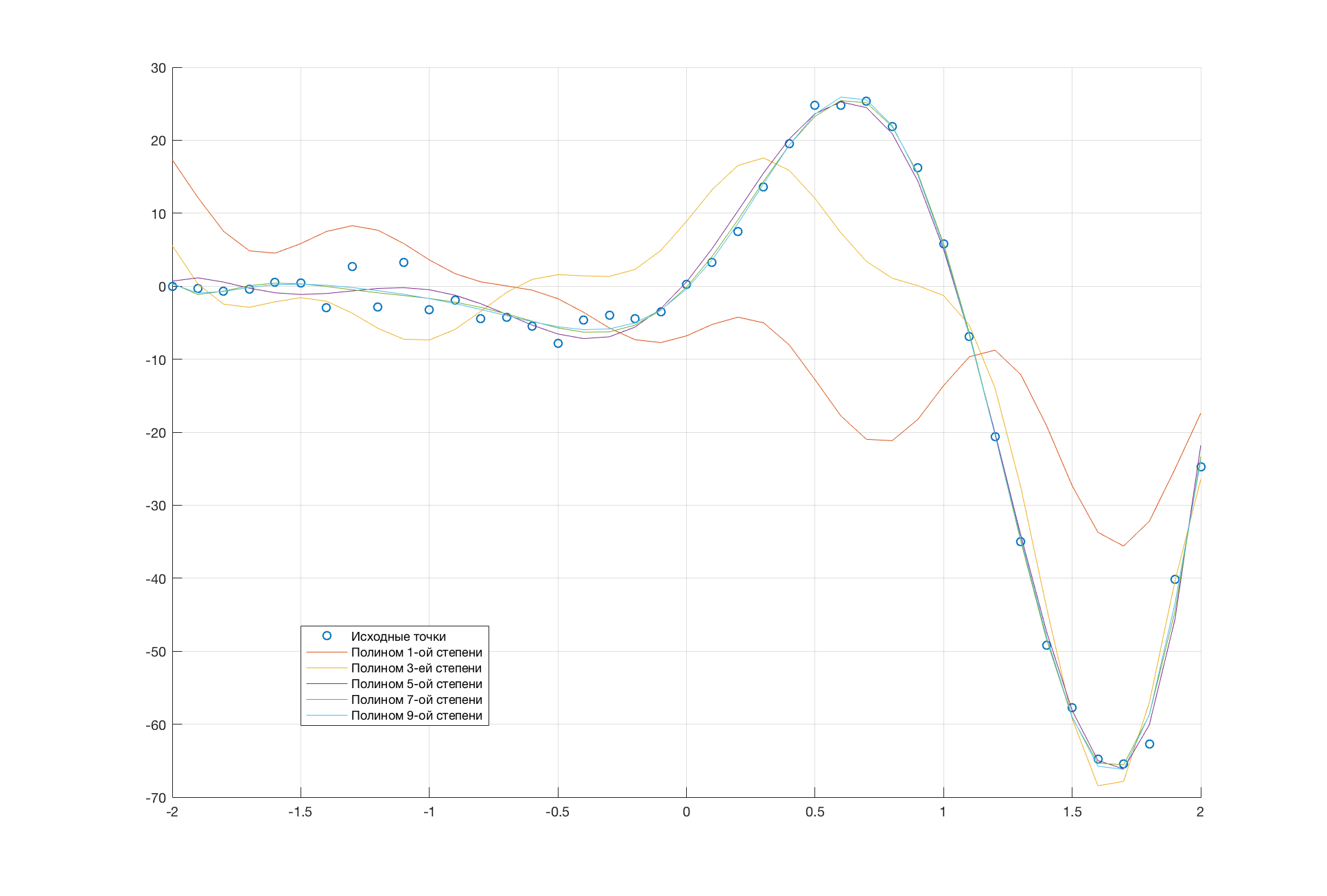
*Рисунок 4.4. Графики интерполяции в масштабе*

Интерполяция не очень хорошо подходит для приближения функции, так как интерполяция чувствительна к погрешностям исходных данных, которые присутствуют в наших данных.

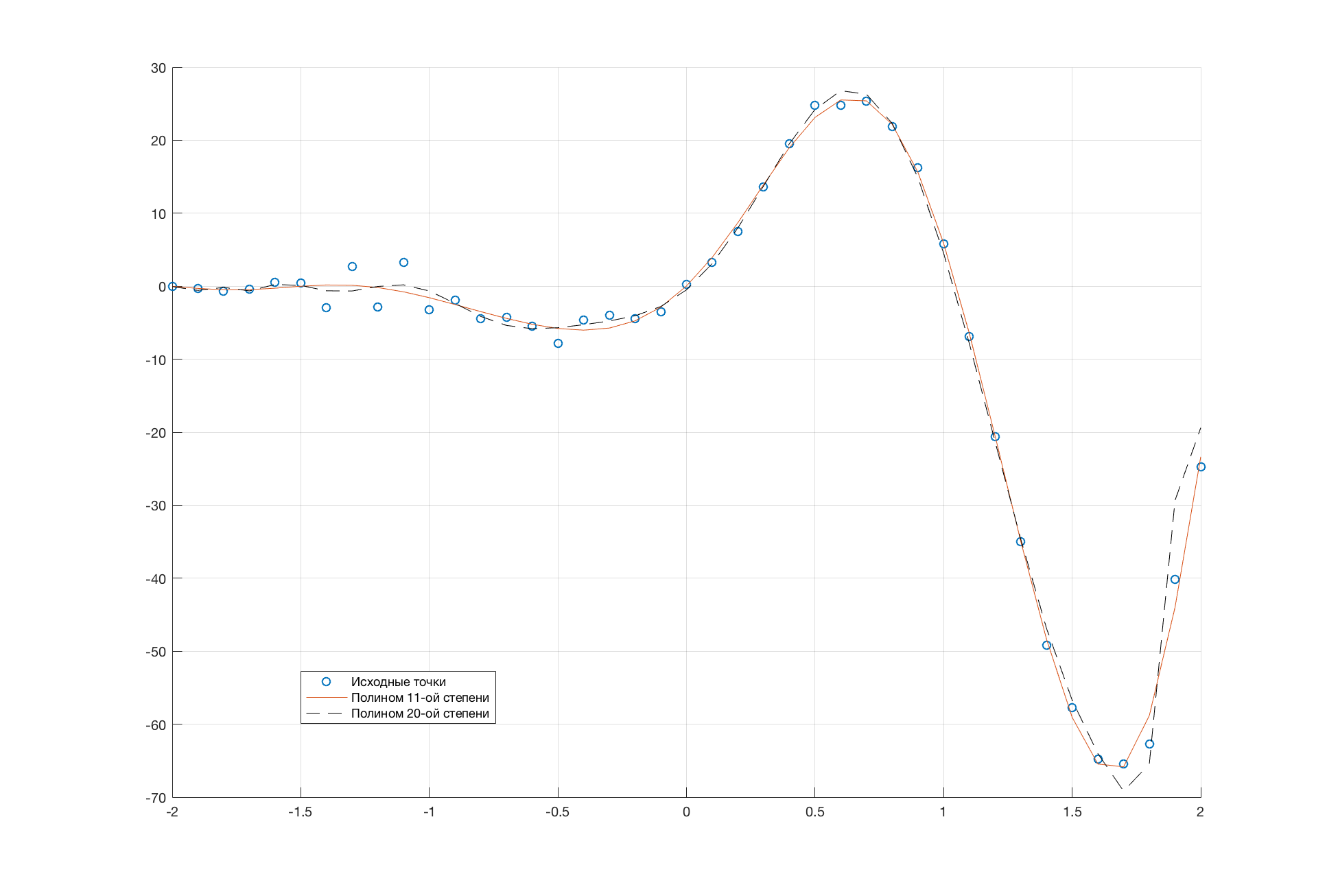
**Нелинейная аппроксимация** с помощью встроенной функции *nlinfit* в *MATLAB*.

В качестве модели функции использовалась следующая функция:

Степень полинома бралась выборочно из допустимого диапазона. Полученные графики представлены на рисунках 4.5 и 4.6, а найденные параметры в таблице 4.1.



*Рисунок 4.5. Графики нелинейной аппроксимации*



*Рисунок 4.6. Графики нелинейной аппроксимации*

*Таблица 4.1. Параметры функции*

|  |  |  |  |
| --- | --- | --- | --- |
| Параметр | Степень полинома | | |
| 1 | 3 | 5 |
|  | 6.6331 | 5.3092 | 1.5294 |
|  | -1.8772 | 1.3269 | 0.0218 |
| a0 | 3.6233 | 6.6888 | 26.8250 |
| a1 | 5.2466 | 6.3159 | 31.1030 |
| a2 |  | -9.6875 | -20.7348 |
| a3 |  | -5.8956 | -31.8070 |
| a4 |  |  | -3.3719 |
| a5 |  |  | 2.9288 |

Результаты примерно такие же, как и при аппроксимации обычным полиномом.

**Вывод**

В результаты выполнения данной работы были проанализированы выборки значений функции, посчитанные с некоторой неизвестной погрешностью. Для средних значений были построены доверительные и толерантные интервалы.

Выборка из средних значений была аппроксимирована полиномами различной степени. Качество аппроксимации оценивалось с помощью критерия Фишера. Наилучшими оказались аппроксимации полиномами степени 9 и 16 (у одной самая маленькая степень, другая дает наименьшую статистику). Полученные аппроксимации хорошо приближают исходную функцию. Высокая степень полинома чувствительна к погрешностям исходных данных, поэтому аппроксимации с высокой степенью не прошли критерий Фишера.

В работе также были применены различные аппроксимации с помощью функций в MATLAB:

* Линейные аппроксимации. Все одинаково плохо приближают функцию, так как исходная функция нелинейная.
* Полиномиальная аппроксимация. Такая же, какая и при самостоятельном поиске подходящей степени. Слишком маленькие и слишком большие степени полинома плохо приближают исходную функцию, однако полином с правильно выбранной степенью хорошо справляется с поставленной задачей.
* Интерполяция. Неверно применять интерполяцию к исходным данным так как они имеют довольно высокую погрешность.
* Нелинейная регрессия. Очень хорошо приближает функцию. Однако по трудоёмкости сильно уступает полиномиальной аппроксимации. При высоких степенях полинома вычисления занимают очень много времени.

В итоге, попробовав все перечисленные методы аппроксимаций (и интерполяции) я отдаю предпочтение полиномиальной аппроксимации, так как она хорошо приближает функцию и не очень сложна в реализации.