**Министерство науки и высшего образования Российской Федерации**

федеральное государственное автономное образовательное учреждение высшего образования

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»**

**Отчёт**

по лабораторной работе номер 2

по дисциплине «Методы оптимизации»

Авторы: Лихачев Александр Викторович

Кислов Данил Александрович

Хадыров Алмасгали Аделгалиевич

Факультет: ФИТиП

Группа: M32351

Преподаватель: Ким С. Е.



Санкт-Петербург 2023

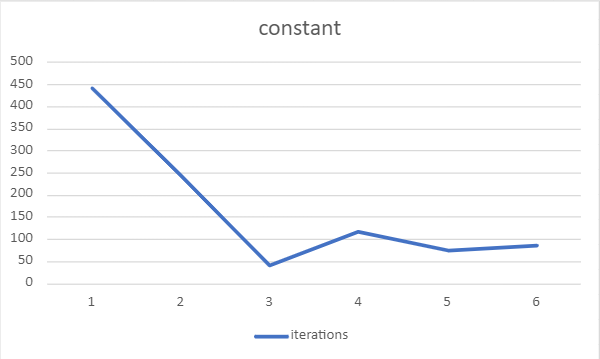
**Цель работы:** реализовать алгоритм стохастического градиентного спуска, использовав его для решения задачи линейной регрессии и протестировать его в различных конфигурациях алгоритма.

**Задачи работы:**

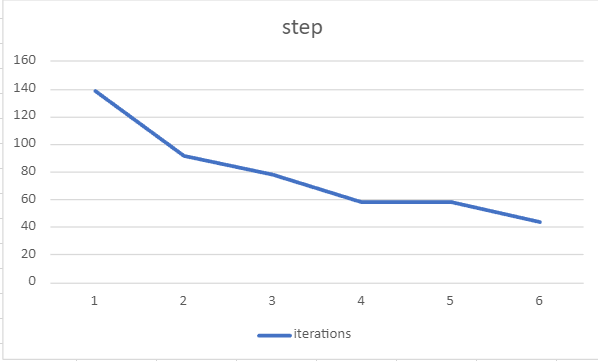
* Реализовать метод стохастического градиентного спуска
* Провести исследования его работы с разными размером батча
* Подобрать функцию изменения шага
* Исследовать различные его модификации
* Исследовать сходимость и эффективность разных конфигураций алгоритма, использовав в качестве метрик такие показатели как время работы, затраченная память, количество арифметических операций
* Построить траекторию спуска
* Проверить работу алгоритма на задаче полиномиальной регрессии
* Исследовать влияние регуляризации на результат полиномиальной регрессии

**Исследование работы стохастического градиентного спуска в зависимости от размера батча**

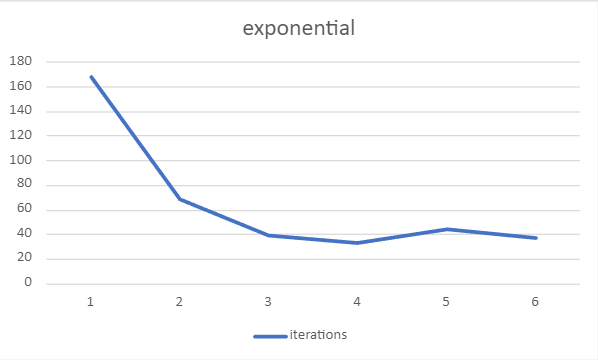
Константная функция изменения шага (lr = 0.01)



Ступенчатая функция изменения шага (initial\_lr = 0.01, drop = 1.2, frequency = 10)



Экспоненциальная функция изменения шага (initial\_lr = 0.1, k = 0.01)

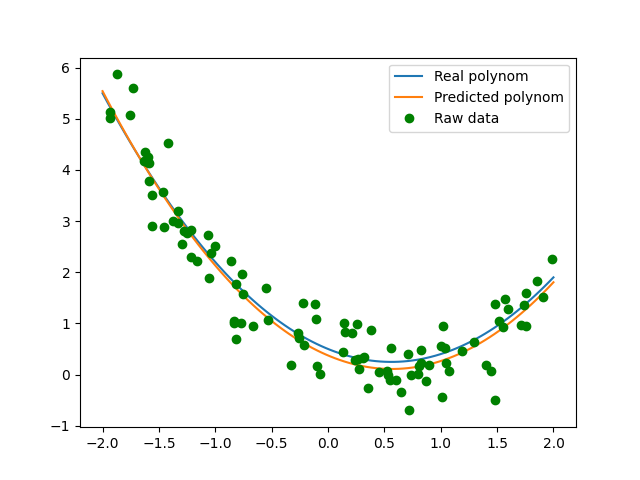


Можно заметить, что экспоненциальная функция изменения шага показывает наилучший результат в плане количества итераций. Так же можно заметить, что чем больше размер батча, там меньше нужно итерация для схождения

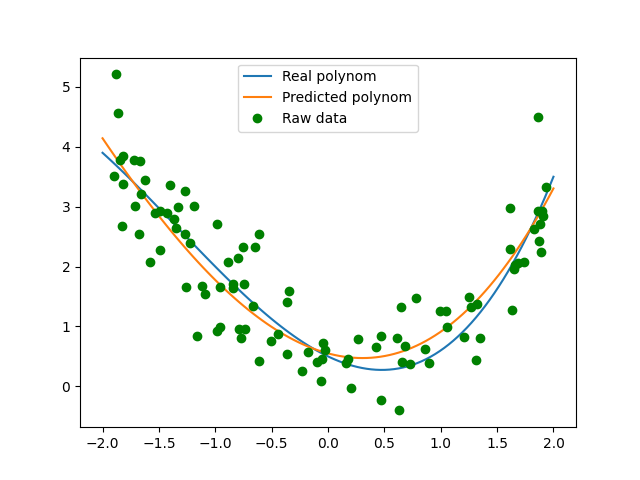
**Применение алгоритма к задаче полиномиальной регрессии**

В качестве точек для функции потерь возьмём 100 случайных значений x на отрезке от -2 до 2 и соответствующие им значения функции. Для усложнения задачи добавим шум (вектор нормально распределённых случайных значений) к значениям y. В качестве алгоритма возьмём стохастический градиентный спуск с модификацией Adam.

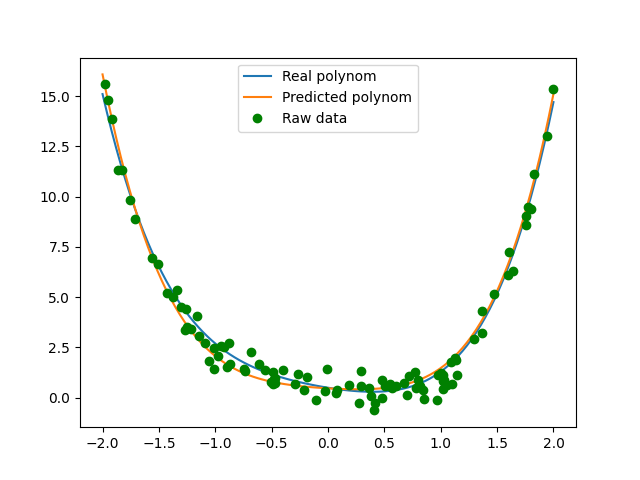
**Работа алгоритма на полиноме 3 степени**



**Работа алгоритма на полиноме 4 степени**

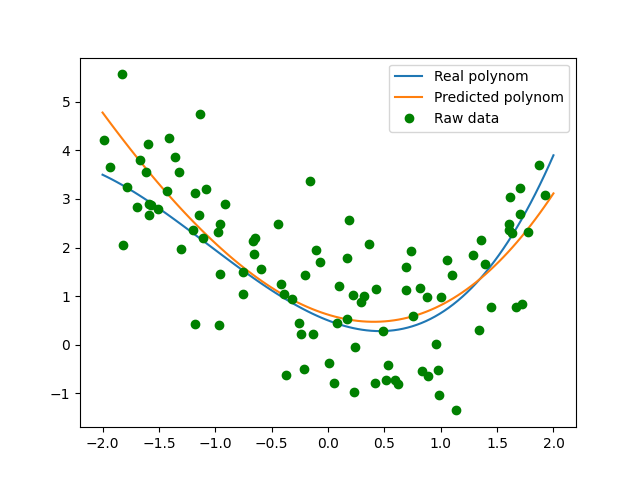


**Работа алгоритма на полиноме 5 степени**

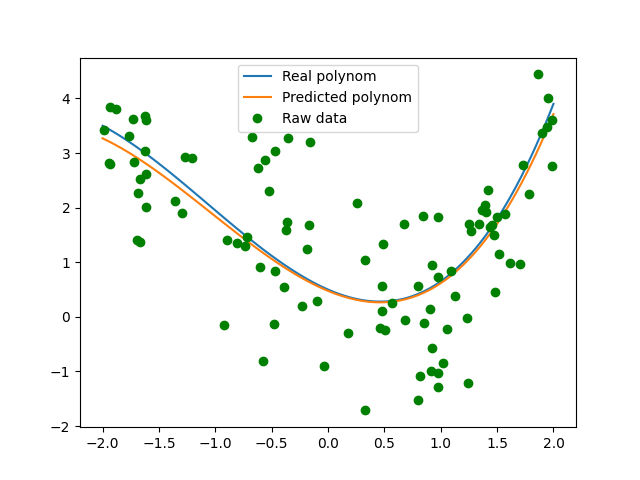


При высоком зашумлении данных имеет смысл использовать регуляризации чтобы улучшить результат.

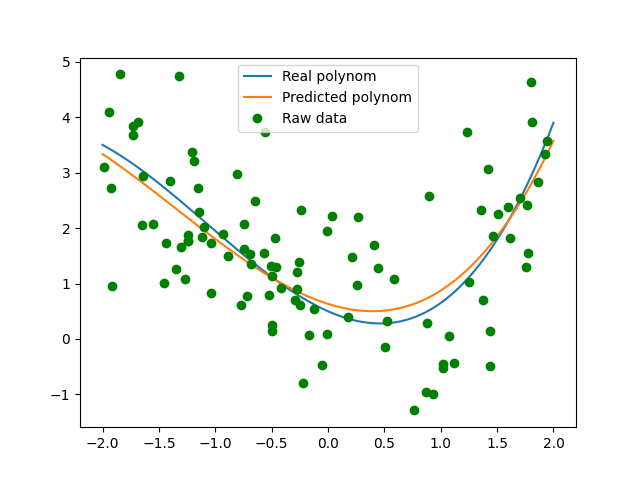
**Алгоритм без регуляризации**



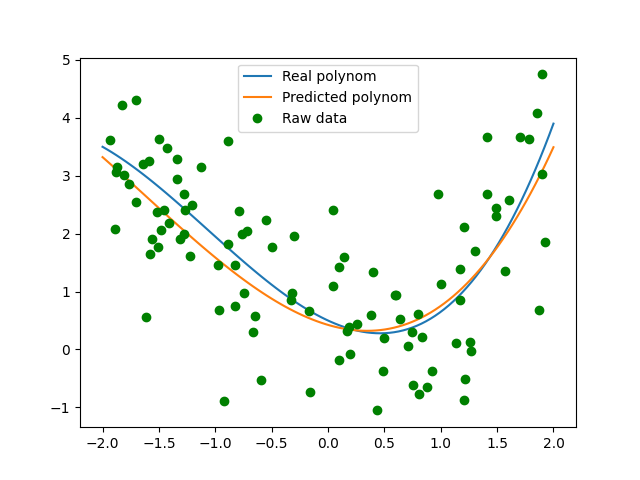
**Алгоритм с регуляризацией L1**



**Алгоритм с регуляризацией L2**



**Алгоритм регуляризации ElasticNet**



Алгоритм регуляризации ElasticNet объединяет в себе L1 и L2 и, зачастую, работает не хуже каждого из них, из-за чего по-умолчанию рекомендуется к использованию. Положительное влияние регуляризации на качество решения задачи полиномиальное регрессии особенно ощущается при высоком уровне зашумления и практически незаметно при невысоком уровне шума.