Министерство образования Республики Беларусь

Учреждение образования

БелорусскиЙ государственный университет

информатики и радиоэлектроники

Факультет компьютерных систем и сетей

Кафедра информатики

ОТЧЕТ

по лабораторной работе №1

**Линейная регрессия**

|  |  |  |
| --- | --- | --- |
| Студент |  | И.Ю. Фалько |
| Преподаватель |  | М. В. Стержанов |

Минск 2019

СОДЕРЖАНИЕ

[Условие лабораторной работы 3](#_Toc25711884)

[Ход выполнения 3](#_Toc25711885)

[Заключение 8](#_Toc25711886)

# Условие лабораторной работы

Набор данных **ex1data1.txt** представляет собой текстовый файл, содержащий информацию о населении городов (первое число в строке) и прибыли ресторана, достигнутой в этом городе (второе число в строке). Отрицательное значение прибыли означает, что в данном городе ресторан терпит убытки.

Набор данных **ex1data2.txt** представляет собой текстовый файл, содержащий информацию о площади дома в квадратных футах (первое число в строке), количестве комнат в доме (второе число в строке) и стоимости дома (третье число).

# Ход выполнения

Загрузите набор данных **ex1data1.txt** из текстового файла.

Постройте график зависимости прибыли ресторана от населения города, в котором он расположен.

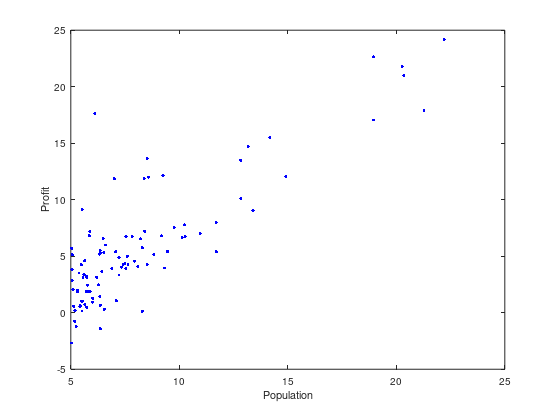
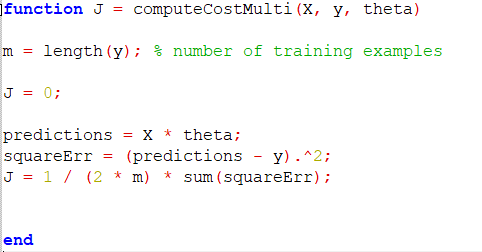


Рисунок 1. – график зависимости прибыли ресторана от населения

Реализуйте функцию потерь J(θ) для набора данных **ex1data1.txt**.

Функция потерь:



Реализуйте функцию градиентного спуска для выбора параметров модели. Постройте полученную модель (функцию) совместно с графиком из пункта 2.

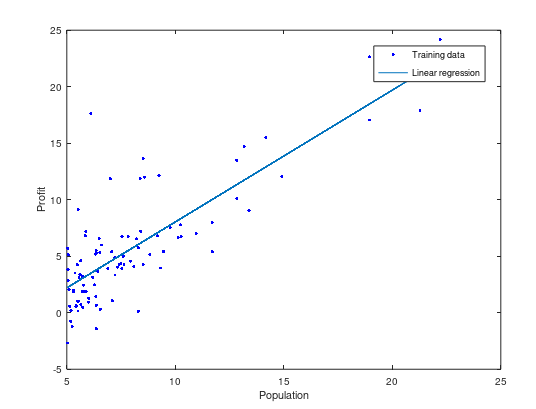


Рисунок 2. – график линейной регрессии и данных

Постройте трехмерный график зависимости функции потерь от параметров модели (θ0 и θ1) как в виде поверхности, так и в виде изолиний (contour plot).

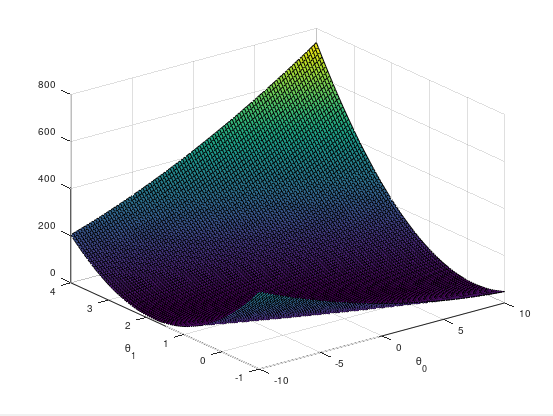


Рисунок 3. – график зависимости функции потерь от параметров модели в виде поверхности

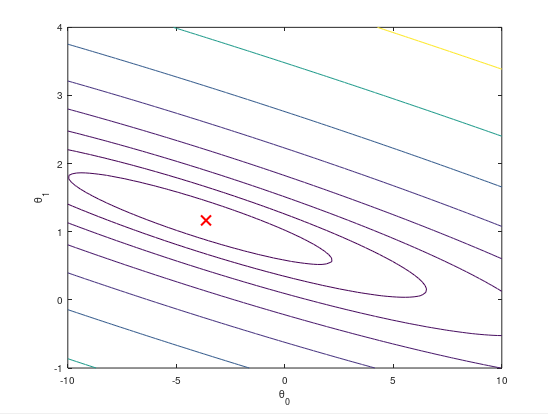


Рисунок 4. – график зависимости функции потерь от параметров модели в виде изолиний

Загрузите набор данных **ex1data2.txt** из текстового файла.

Произведите нормализацию признаков. Постройте график сходимости градиентного спуска

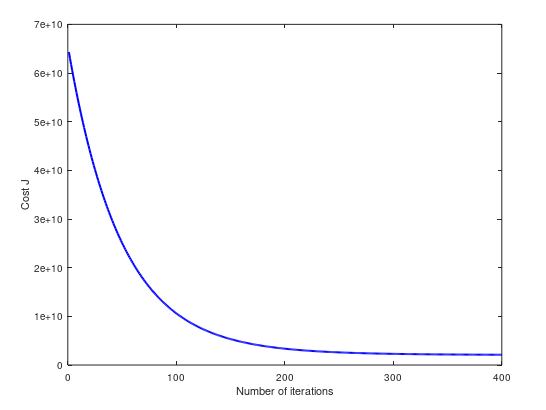
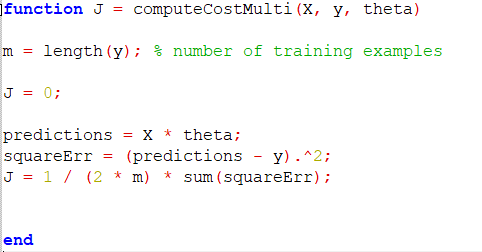


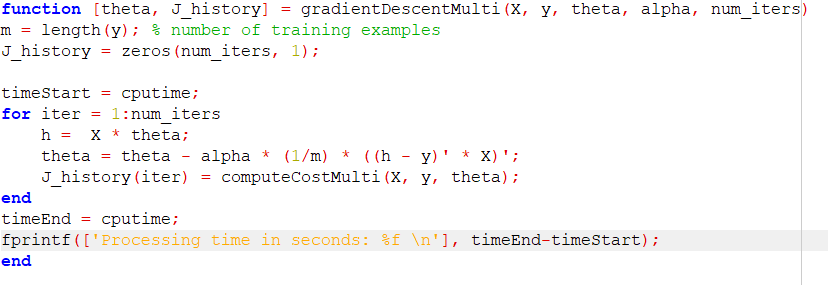
Рисунок 5. – график сходимости градиентного спуска

Реализуйте функции потерь J(θ) и градиентного спуска для случая многомерной линейной регрессии с использованием векторизации.

Функция потерь:



Градиентный спуск:



Покажите, что векторизация дает прирост производительности.

До векторизации



После векторизации:



Попробуйте изменить параметр ɑ (коэффициент обучения). Как при этом изменяется график функции потерь в зависимости от числа итераций градиентного спуск? Результат изобразите в качестве графика.

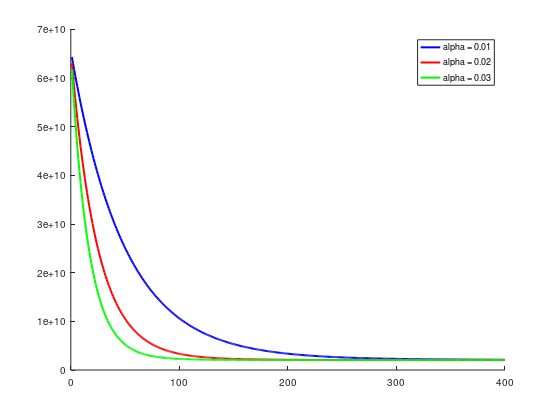
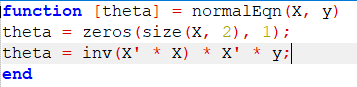


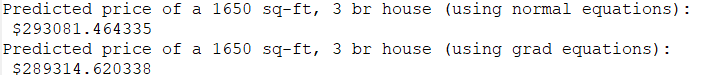
Рисунок 6. – график функции потерь в зависимости от числа итераций градиентного спуска

Постройте модель, используя аналитическое решений. Сравните результаты данной модели с моделью, полученной с помощью градиентного спуска.

Метод нормальных уравнений:



Сравнение результатов:



# ЗАКЛЮЧЕНИЕ

В данной лабораторной работе были освоены навыки логической регрессии с помощью градиентного спуска. Также было изучено аналитическое решение для нахождения минимального значения функции потерь – метод нормальных уравнений.

Сравнение этих методов показало, что при использовании метода нормальных уравнений не нужно выбирать коэффициент обучения. Он не итерируемый. Однако недостатком является низкая эффективность при большом количестве n.