**Лабораторная работа № 1**

**Линейная регрессия.**

**Решить задачу парной регрессии на данных alligators.txt и cygage.txt в среде R.**

**ЧАСТЬ I Изучение метода решения**

1. Ответьте на вопросы: **Таблица 1**

|  |  |  |
| --- | --- | --- |
| **№** | **Вопрос** | **Ответ** |
|  | Как называется модель, изучаемая в лабораторной работе 1? |  |
|  | К какой категории методов ML она относится? |  |
|  | Что может прогнозировать алгоритм, обученный данной моделью? |  |
|  | В файле исходных данных ***alligators.txt***: |  |
| a) | Что является признаком, что ответом? |  |
| b) | Сколько ответов? |  |
|  | Запишите функцию ***Q(w0, w1, X)*** ошибки алгоритма для данных в файле ***alligators.txt***. |  |
|  | Запишите градиент этой функции, т.е. ***∇ Q(w0, w1, X).*** | Здесь д.б. скрин текста от руки с выводом координат вектора градиента от функции качества. |
|  | Выпишите **в развернутом виде** от руки итерационную формулу для метода градиентного спуска **в парной регрессии.** | Здесь д.б. скрин текста от руки с выводом итерационной формулы для метода градиентного спуска в парной регрессии.. |

**ЧАСТЬ II Выполнение задания**

Выполните следующие задания:

1. Для данных **alligators.txt** постройте регрессионную модель, отражающую зависимость веса аллигатора от его длины:
2. Загрузите данные из файла alligators.txt.
3. Сравните свое решение задачи с решением с посредством стандартного пакета для парной регрессии методом градиентного спуска.
4. Выведите диаграмму рассеяния в RStudio, используя примеры [1]. В этих же осях выведите результат обученной модели (см. рис.1) [2]
5. Повторите шаги с 1) по 3) для данных **cygage.txt**, построив регрессию, выражающую зависимость возраста исследуемых отложений от глубины залегания, используя веса наблюдений.

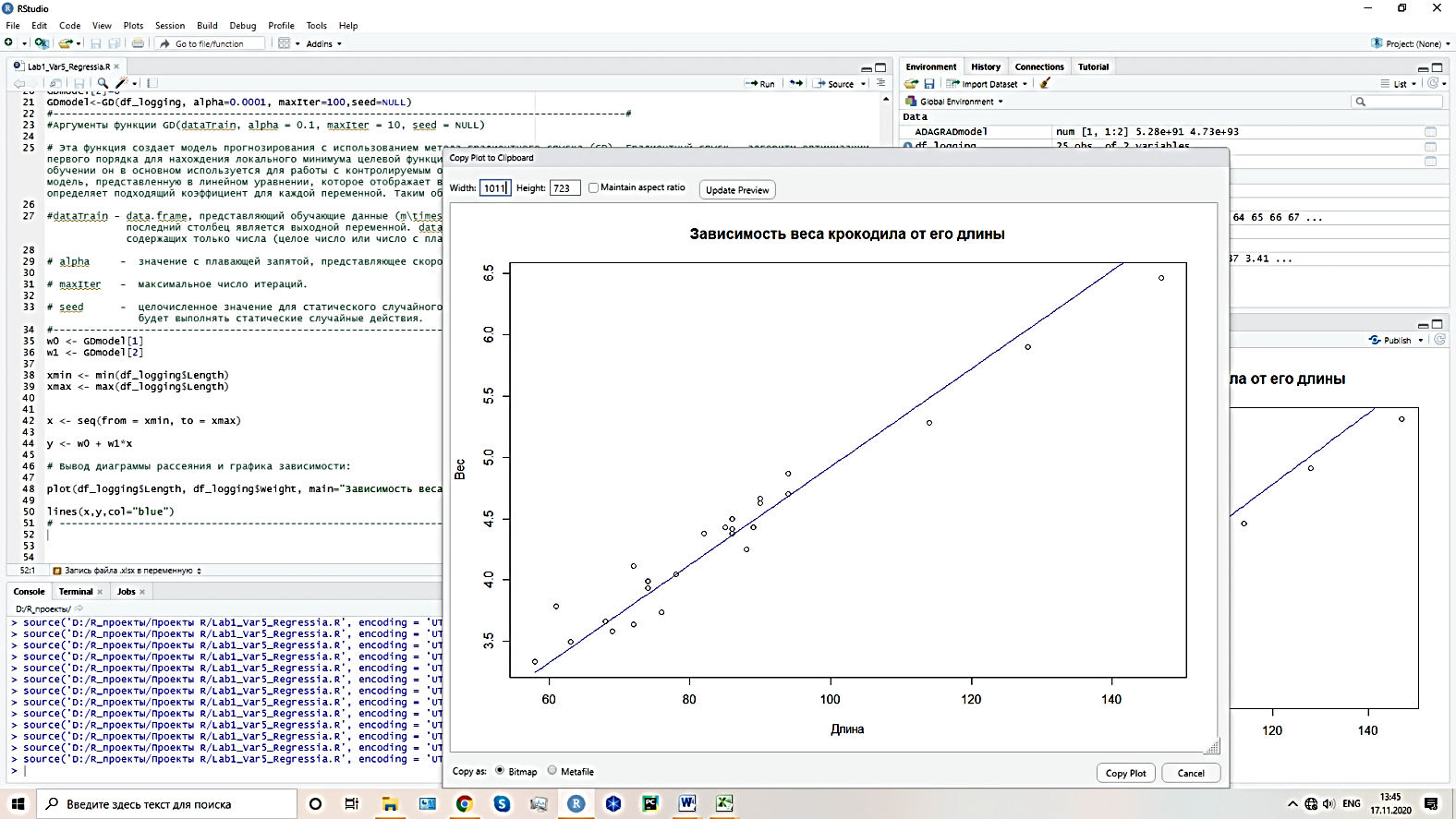


Рис. 1 Диаграмма рассеяния и график модели парной регрессии.

**ЧАСТЬ III Оформление работы**

Оформить отчет по работе надо вследующем виде:

* 1. Титульный лист (см. приложение)
  2. Таблица 1 с ответами на вопросы.
  3. Формулировка задачи об аллигаторах.
     1. Программный код.
     2. Результат вычислений в виде графиков.
     3. Выводы о качестве построенного алгоритма для обработанных данных **alligators.txt**.
  4. Формулировка задачи о возрасте исследуемых отложений.
     1. Программный код.
     2. Результат вычислений в виде графиков.
     3. Выводы о качестве построенного алгоритма для обработанных данных **cygage.txt**.

Литература

1. Построение графиков

<https://tsamsonov.github.io/r-geo-course/graphics.html>

1. Диаграмма рассеяния

<https://tsamsonov.github.io/r-geo-course/graphics.html>

<http://www.inp.nsk.su/~baldin/DataAnalysis/R/R-02-data.pdf>

1. Графическая библиотека ***ggplot2()*** и ее возможности

<https://rstudio-pubs-static.s3.amazonaws.com/338694_0241a111721b4846b6fee38fba64302d.html>

**Приложение.**

Министерство сельского хозяйства Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«Пермский государственный аграрно-технологический университет

имени академика Д.Н. Прянишникова»

Кафедра информационных технологий и программной инженерии

**ЛАБОРАТОРНАЯ РАБОТА №1**

Классические методы обучения с учителем.

Регрессия.

|  |  |
| --- | --- |
| Выполнил: | группа ПИНб-4 |
| Проверил: | доцент каф. ИТиПИ,  Т.А. Казаченко |

Пермь-2020 г.