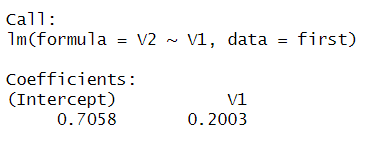
**Задание 1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Средняя мощность тракторов (л. с.) | 4,1 | 6,1 | 7,8 | 9,9 | 12,4 |
| Валовая продукция на одного работника (тыс. руб.) | 1,4 | 2,0 | 2,5 | 2,6 | 3,2 |

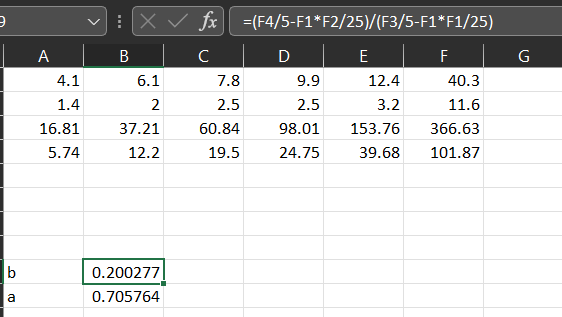
Построим модель линейной регрессии с помощью R, получим следующие коэффициенты:



a = 0.7058

b = 0.2003

Рассчитаем коэффициенты с помощью формул в excel:

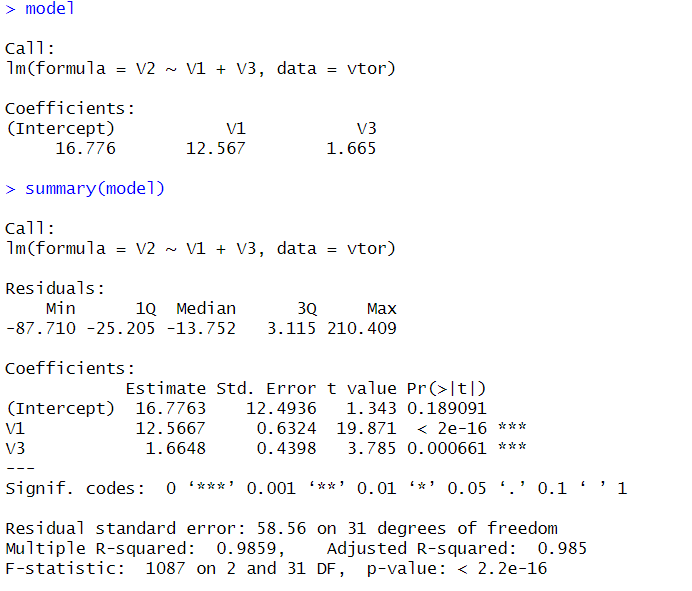


Получили те же результаты, что и в R.

**2 задание**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Страна | Y | *X*1 | *X* 2 | Страна | Y | *X*1 | *X* 2 |
| Люксембург | 0,34 | 5,67 | 0,36 | Швейцария | 5,31 | 101,65 | 6,37 |
| Уругвай | 0,22 | 10,13 | 2,90 | С. Аравия | 6,40 | 115,97 | 8,37 |
| Сингапур | 0,32 | 11,34 | 2,39 | Бельгия | 7,15 | 119,49 | 9,86 |
| Ирландия | 1,23 | 18,88 | 3,44 | Швеция | 11,22 | 124,15 | 8,31 |
| Израиль | 1,81 | 20,94 | 3,87 | Австралия | 8,66 | 140,98 | 14,62 |
| Венгрия | 1,02 | 22,16 | 10,71 | Аргентина | 5,56 | 153,85 | 27,06 |
| Н. Зеландия | 1,27 | 23,83 | 3,10 | Нидерланды | 13,41 | 169,38 | 14,14 |
| Португалия | 1,07 | 24,67 | 9,93 | Мексика | 5,46 | 186,33 | 67,40 |
| Гонконг | 0,67 | 27,56 | 5,07 | Испания | 4,79 | 211,78 | 37,43 |
| Чили | 1,25 | 27,57 | 11,10 | Бразилия | 8,92 | 249,72 | 123,03 |
| Греция | 0,75 | 40,15 | 9,60 | Канада | 18,90 | 261,41 | 23,94 |
| Финляндия | 2,80 | 51,62 | 4,78 | Италия | 15,95 | 395,52 | 57,04 |
| Норвегия | 4,90 | 57,71 | 4,09 | Великобритания | 29,90 | 534,97 | 55,95 |
| Югославия | 3,50 | 63,03 | 22,34 | Франция | 33,59 | 655,29 | 53,71 |
| Дания | 4,45 | 66,32 | 5,12 | ФРГ | 38,62 | 815,00 | 61,56 |
| Турция | 1,60 | 66,97 | 44,92 | Япония | 61,61 | 1040,4 | 116,78 |
| Австрия | 4,26 | 76,88 | 7,51 | США | 181,3 | 2586,4 | 227,64 |

Построим модель линейной регрессии в R

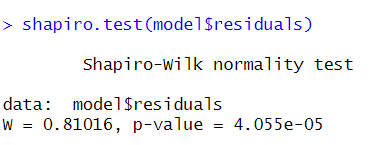


У регрессоров маленькое значение p-value, значит мы отвергаем нулевую гипотезу о том, что их коэффициенты равны нулю. У интерцепта значение p-value больше уровня значимости 0.05, следовательно можно заключить, что скорее всего он равен 0. Коэффициенты детерминации близки к 1 и p-value критерия очень мал, значит, модель хорошо аппроксимирует данные.

Проверим теорему Гаусса-Маркова:

1. Проверка нормальности распределения остатков

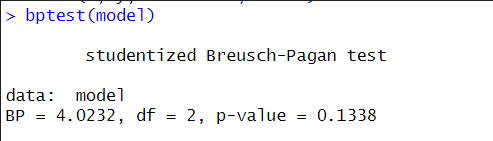
Проведем тест Шапиро-Уилка



p-value меньше 0.05, значит можно отвергнуть гипотезу о том, что остатки распределены нормально.

1. Проверка на гетероскедастичность

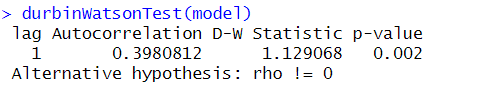
Проведем тест Бройша-Пагана



p-value больше 0.05, значит принимаем нулевую гипотезу о том, что остатки гомоскедастичны

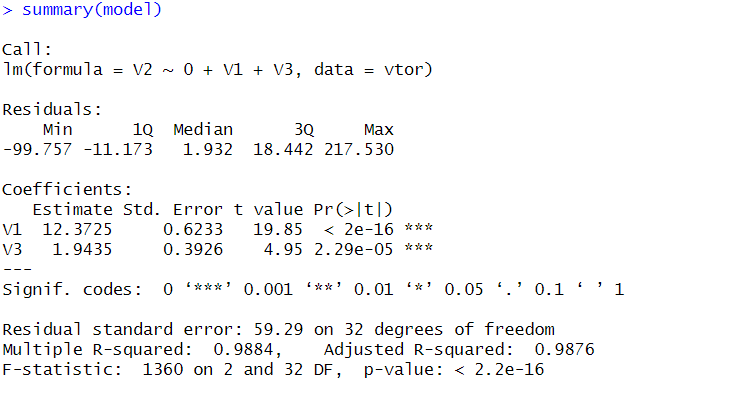
1. Проверка коррелированности остатков

Проведем тест Дарбина-Уотсона



p-value меньше 0.05, значит отвергаем гипотезу о том, что корреляции в остатках нет

Попробуем убрать интерцепт из модели так как у него большое значение p-value.

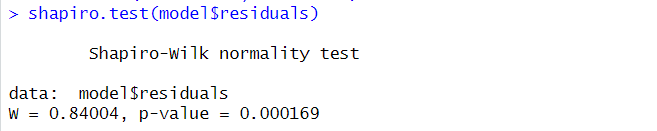


У регрессоров маленькое значение p-value, значит мы отвергаем нулевую гипотезу о том, что их коэффициенты равны нулю. Коэффициенты детерминации близки к 1 и p-value критерия очень мал, значит, модель хорошо аппроксимирует данные.

Проверим теорему Гаусса-Маркова:

1. Проверка нормальности распределения остатков

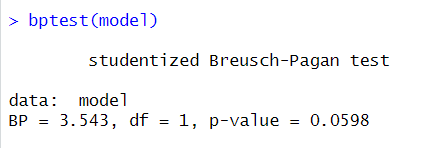
Проведем тест Шапиро-Уилка



p-value меньше 0.05, значит можно отвергнуть гипотезу о том, что остатки распределены нормально.

1. Проверка на гетероскедастичность

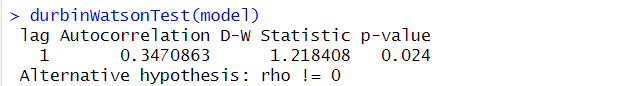
Проведем тест Бройша-Пагана



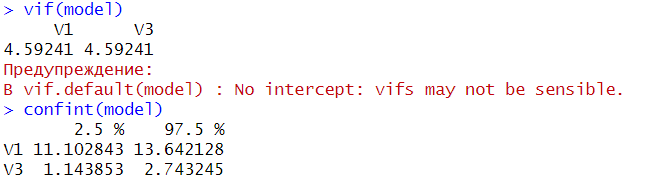
p-value больше 0.05, значит принимаем нулевую гипотезу о том, что остатки гомоскедастичны

1. Проверка коррелированности остатков

Проведем тест Дарбина-Уотсона

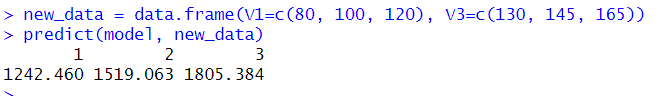


p-value меньше 0.05, значит отвергаем гипотезу о том, что корреляции в остатках нет



Коэффициенты инфляции - умеренные

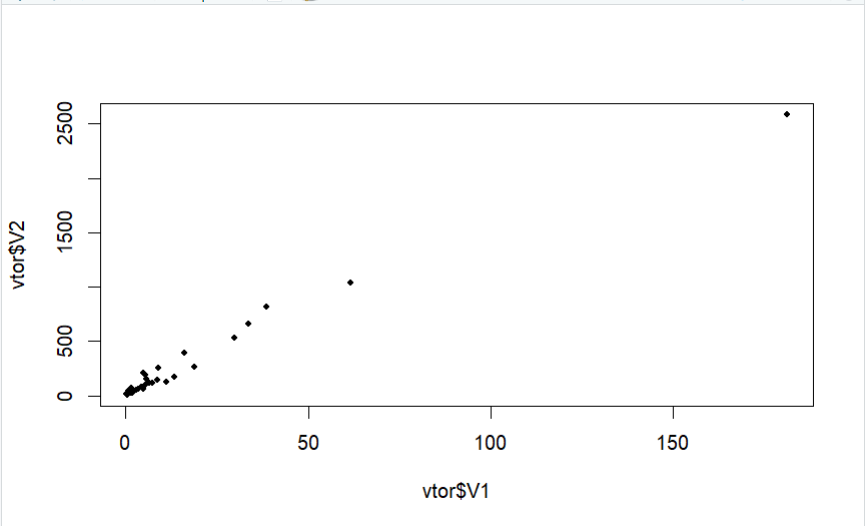
Обе модели не прошли проверку теоремой Гаусса-Маркова, значит они обе плохо будут аппроксимировать данные. Но т.к. у нас нет других регрессоров, то приходится работать с тем, что есть. Сделаем прогноз для 3 новых компонент регрессоров.



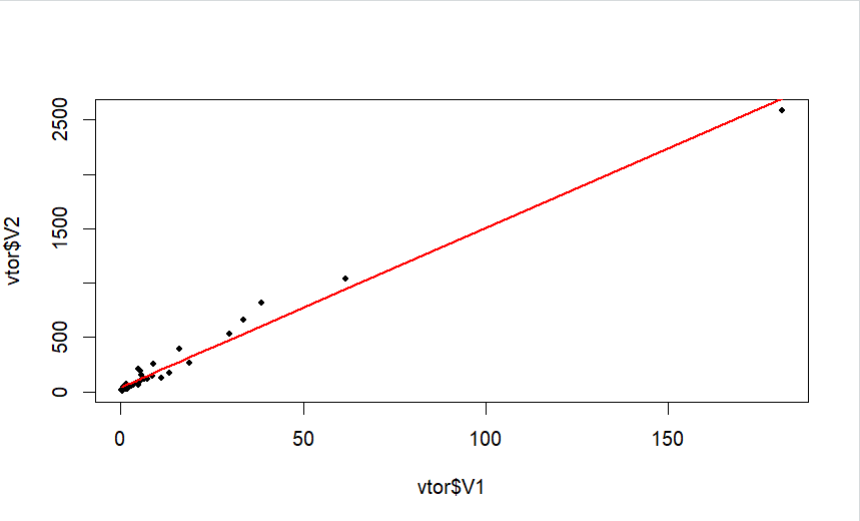
**3 задание**

В качестве регрессора возьмем Y.

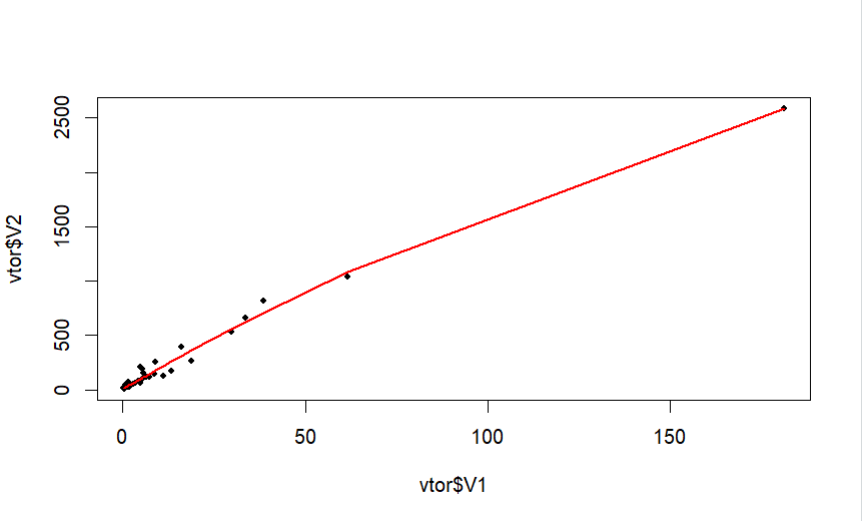
Построим график входных данных



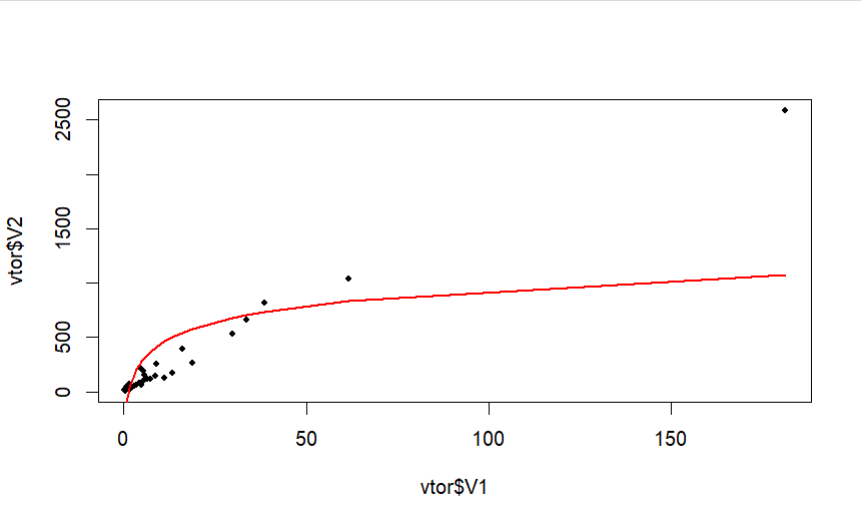
Аппроксимируем линейной моделью



Аппроксимируем полиномиальной моделью 2 степени



Аппроксимируем экспоненциальной моделью



Из трех моделей лучше всех данные аппроксимирует полиномиальная модель