

Министерство образования и науки Российской Федерации

Алтайский государственный технический университет

им. И.И. Ползунова

Отчет

о лабораторной работе №3

по теме: «Построение и оценка регрессионных моделей»

предмет: Компьютерные методы анализа и прогнозирования в  
экономических системах

Выполнил

Студент группы 8ПИЭ-41

Хартов А.Е.

Проверил

Жуковский М.С.

Барнаул 2024

## **Цель работы**

Цель данной лабораторной работы — освоить методы построения и анализа регрессионных моделей для экономических данных с использованием языка программирования R. В рамках работы вы научитесь:

- Разрабатывать линейные модели для экономических данных.
- Анализировать остатки и проводить диагностику моделей.
- Проверять гипотезы о значимости коэффициентов.
- Использовать регрессионные модели для прогнозирования.

## Задание 1: Подготовка данных

### 1. Импорт данных:

```
data("mtcars")
print(head(mtcars))
str(mtcars)
```

Результат:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

'data.frame': 32 obs. of 11 variables:

\$ mpg : num 21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...

\$ cyl : num 6 6 4 6 8 6 8 4 4 6 ...

\$ disp: num 160 160 108 258 360 ...

\$ hp : num 110 110 93 110 175 105 245 62 95 123 ...

\$ drat: num 3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...

\$ wt : num 2.62 2.88 2.32 3.21 3.44 ...

\$ qsec: num 16.5 17 18.6 19.4 17 ...

\$ vs : num 0 0 1 1 0 1 0 1 1 1 ...

\$ am : num 1 1 1 0 0 0 0 0 0 0 ...

\$ gear: num 4 4 4 3 3 3 3 4 4 4 ...

\$ carb: num 4 4 1 1 2 1 4 2 2 4 ...

## Задание 2: Построение линейной регрессии

### 1. Построение простой линейной регрессии:

Модель зависимости `mpg` от `wt`:

```
model_simple <- lm(mpg ~ wt, data = mtcars)
print(summary(model_simple))
```

## Результаты:

Call:

```
lm(formula = mpg ~ wt, data = mtcars)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.5432	-2.3647	-0.1252	1.4096	6.8727

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	37.2851	1.8776	19.858	< 2e-16 ***
wt	-5.3445	0.5591	-9.559	1.29e-10 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.046 on 30 degrees of freedom

Multiple R-squared: 0.7528, Adjusted R-squared: 0.7446

F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10

## Анализ результатов:

### - Коэффициенты модели:

Intercept = 37.2851 - Среднее значение mpg при нулевом wt. Хотя значение "вес = 0" не имеет физического смысла, оно важно для модели как точка отсчета.

wt = -5.3445 - при увеличении веса автомобиля на 1000 фунтов, расход топлива уменьшается в среднем на 5.3445 mpg. Это отрицательная связь.

### - Значимость коэффициентов (p-значения):

	p-value
Intercept	<2e-16
wt	1.29e-10

Оба коэффициента статистически значимы: p-значения значительно меньше 0.05. Вес автомобиля сильно влияет на расход топлива.

### - Коэффициент детерминации ( $R^2$ ):

$R^2 = 0.7528$  - Примерно 75% вариации в значениях mpg объясняется весом автомобиля. Это говорит о сильной линейной зависимости между переменными.

## 2. Построение множественной линейной регрессии:

Модель зависимости `mpg` от `wt` и `hp`:

```
model_multiple <- lm(mpg ~ wt + hp, data = mtcars)
print(summary(model_multiple))
```

Результат:

Call:

```
lm(formula = mpg ~ wt + hp, data = mtcars)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.941	-1.600	-0.182	1.050	5.854

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	37.22727	1.59879	23.285	< 2e-16 ***
wt	-3.87783	0.63273	-6.129	1.12e-06 ***
hp	-0.03177	0.00903	-3.519	0.00145 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.593 on 29 degrees of freedom

Multiple R-squared: 0.8268, Adjusted R-squared: 0.8148

F-statistic: 69.21 on 2 and 29 DF, p-value: 9.109e-12

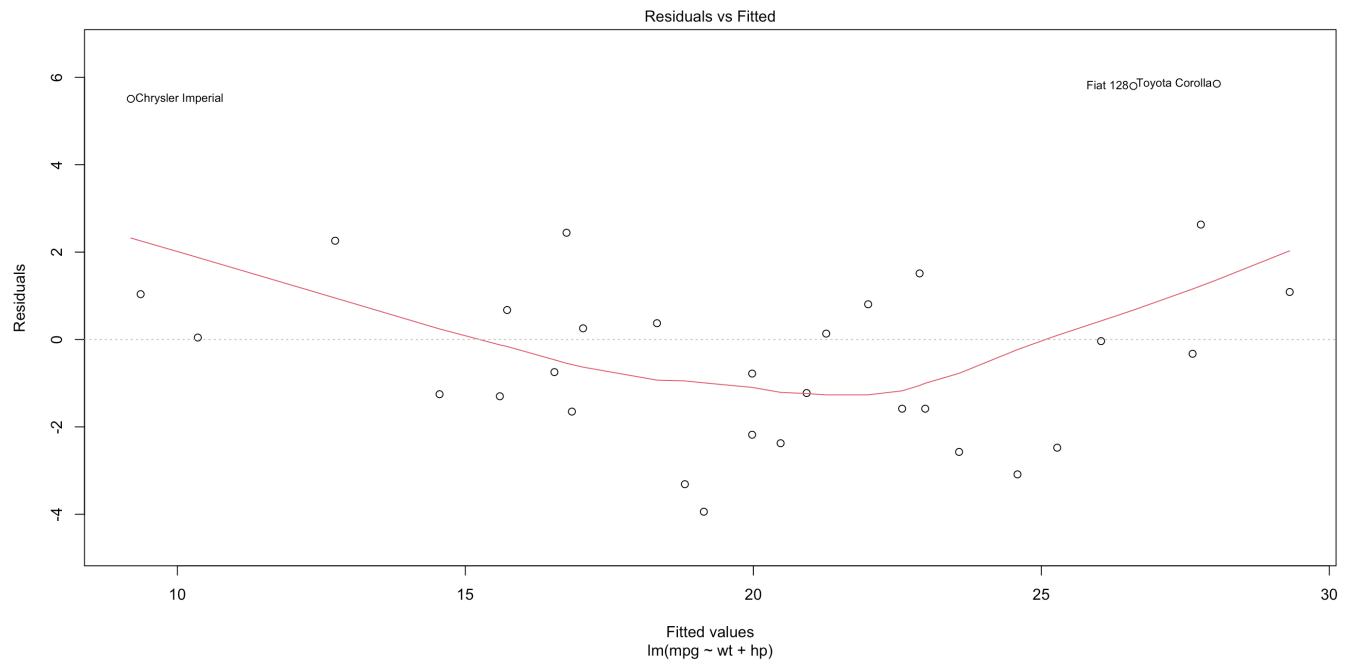
Сравнение результатов с простой регрессией:

Значение Intercept осталось таким же, значение wt уменьшилось по модулю. После учёта мощности (hp) влияние веса на расход ослабло, так как часть вариации в mpg объясняется теперь hp.

Коэффициент  $R^2$  увеличился до 0.8268, видно что модель улучшилась. Модель теперь объясняет больше вариации в mpg.

### Задание 3: Анализ остатков и диагностика модели

#### 1. График остатков:

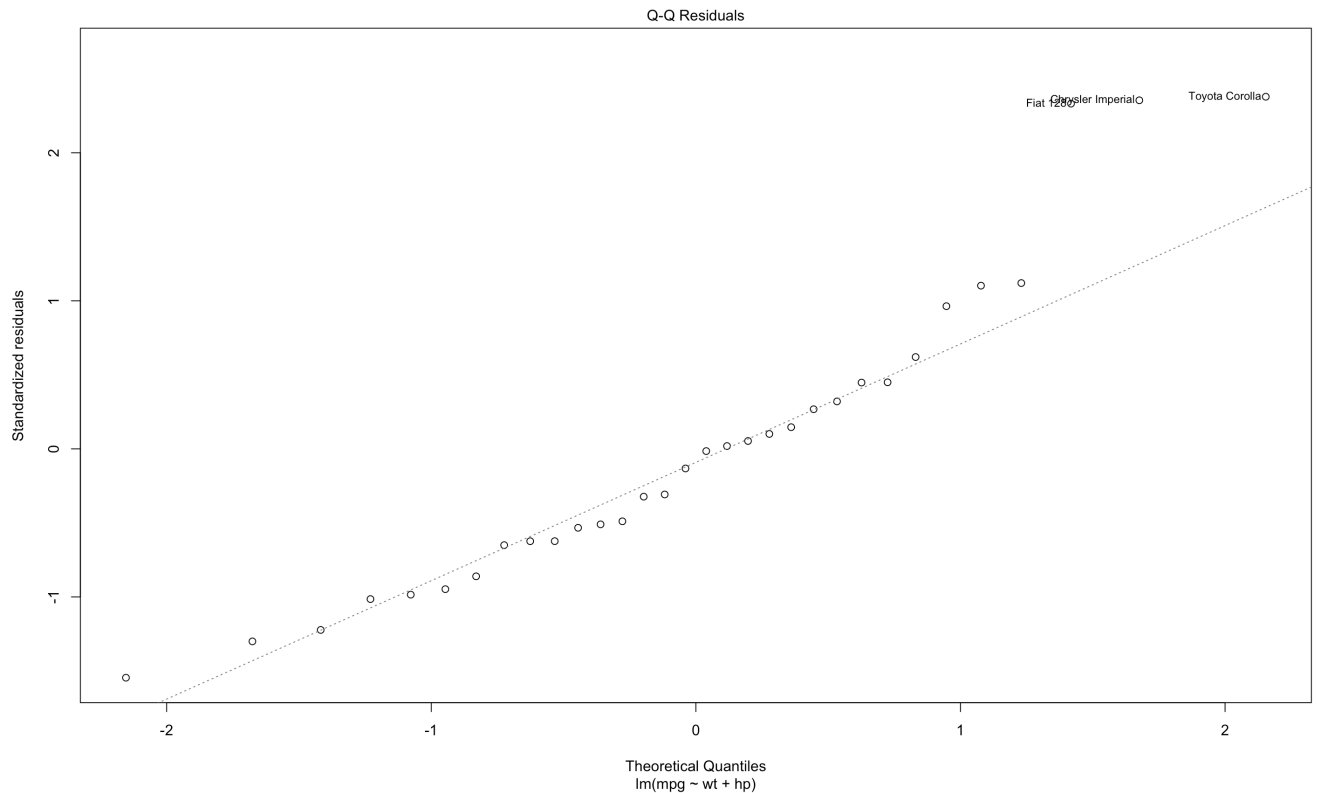


- Есть ли систематические отклонения?

Да, систематические отклонения есть, так как распределение остатков формируют U-образную форму

## 2. Q-Q график:

```
plot(model_multiple, which = 2)
```



- Выполняется ли предположение о нормальности остатков?

Большинство точек расположено вдоль прямой линии, что соответствует нормальному распределению остатков. Предположение о нормальности остатков выполняется

## 3. Тест на гетероскедастичность:

тест Бройша-Пагана:

```
print(bptest(model_multiple))
```

studentized Breusch-Pagan test

data: model\_multiple

BP = 0.88072, df = 2, p-value = 0.6438

## 4. Мультиколлинеарность:

- Проверьте корреляцию между предикторами:

```
print(cor(mtcars$wt, mtcars$hp))
```

Корреляция равна 0.6587479. Корреляция меньше 0.7 значит она умеренно высокая.

#### Задание 4: Проверка гипотез о значимости коэффициентов

##### 1. Гипотеза о значимости коэффициентов:

```
print(summary(model_multiple))
```

Call:

```
lm(formula = mpg ~ wt + hp, data = mtcars)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.941	-1.600	-0.182	1.050	5.854

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	37.22727	1.59879	23.285	< 2e-16 ***
wt	-3.87783	0.63273	-6.129	1.12e-06 ***
hp	-0.03177	0.00903	-3.519	0.00145 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.593 on 29 degrees of freedom

Multiple R-squared: 0.8268, Adjusted R-squared: 0.8148

F-statistic: 69.21 on 2 and 29 DF, p-value: 9.109e-12

p-value для всех коэффициентов (Intercept, wt, hp) меньше 0.05. Это означает, что мы отвергаем нулевую гипотезу для каждого коэффициента.

Следовательно, вес машины (wt) и мощность двигателя (hp) оказывают статистически значимое влияние на расход топлива (mpg).

##### 2. Общая значимость модели:

```
print(anova(model_multiple))
```

Analysis of Variance Table

Response: mpg

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
wt	1	847.73	847.73	126.041	4.488e-12 ***
hp	1	83.27	83.27	12.381	0.001451 **
Residuals	29	195.05	6.73		

Видно, что для wt и для hp очень маленькие p-value (< 0.05).



Так как p-value для всей модели меньше 0.05, мы отвергаем нулевую гипотезу. Следовательно, хотя бы один предиктор (wt или hp) существенно связан с расходом топлива (mpg).

Задание 5: Прогнозирование на основе регрессионной модели 1.

Прогнозирование новых значений:

```
new_data <- data.frame(wt = c(3.5, 4.0), hp = c(150, 200))
predictions <- predict(model_multiple, newdata = new_data, interval =
                        "confidence")
print(predictions)
```

Результат:

	fit	lwr	upr
1	18.88892	17.89528	19.88256
2	15.36136	14.11210	16.61062

2. Визуализация прогнозов:

График фактических и прогнозируемых значений:

