ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

КАФЕДРА ВТ

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №2

«Анализ временных рядов в среде R. Метод последовательной идентификации составляющих ВР» по дисциплине «Компьютерные технологии анализа и обработки данных»

Выполнили: студенты

гр. АММ2-24

Атласюк Игорь Романович

Ириков Евгений Алексеевич

Проверил: к.т.н., доцент Кафедры ВТ Альсова Ольга Константиновна

Новосибирск 2024

## Содержание

[Постановка задачи 3](#_bookmark0)

[Ход работы 4](#_bookmark1)

[Заключение 23](#_bookmark2)

[Приложение 24](#_bookmark3)

# **Постановка задачи**

Изучить методы и алгоритмы прогнозирования временных рядов на примере решения конкретной задачи ИАД. Исследовать эффективность прогнозирования временных рядов для решения прикладной задачи. Ознакомиться и получить практические навыки работы с языком R для решения задач исследования и прогнозирования временных рядов.

# **Ход работы**

1. **Загрузка данных и построение графика временного ряда и его декомпозиций (рис.1)**

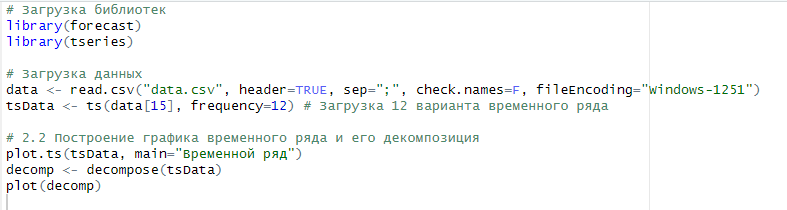


Рисунок 1. Начало работы

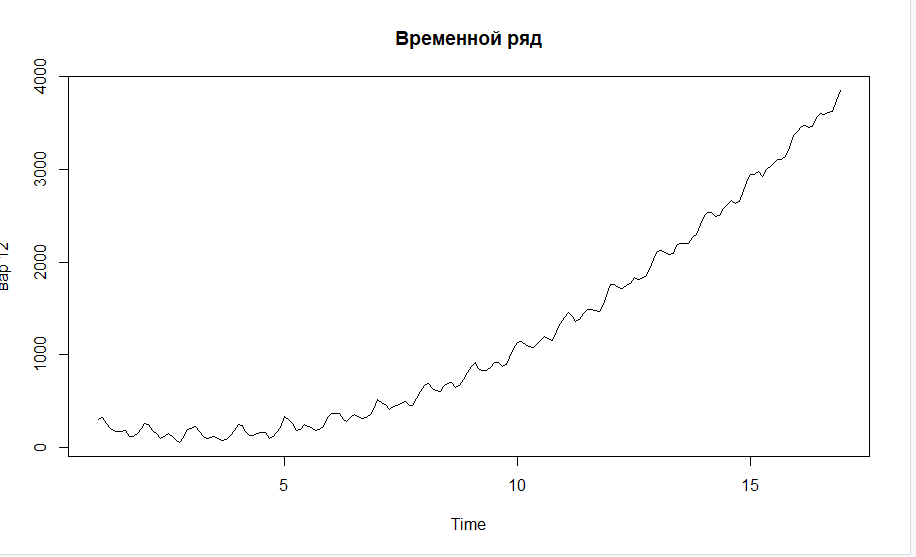


Рисунок 2. Временной ряд

Временной ряд изменяется по параболе, ярко выраженный тренд.

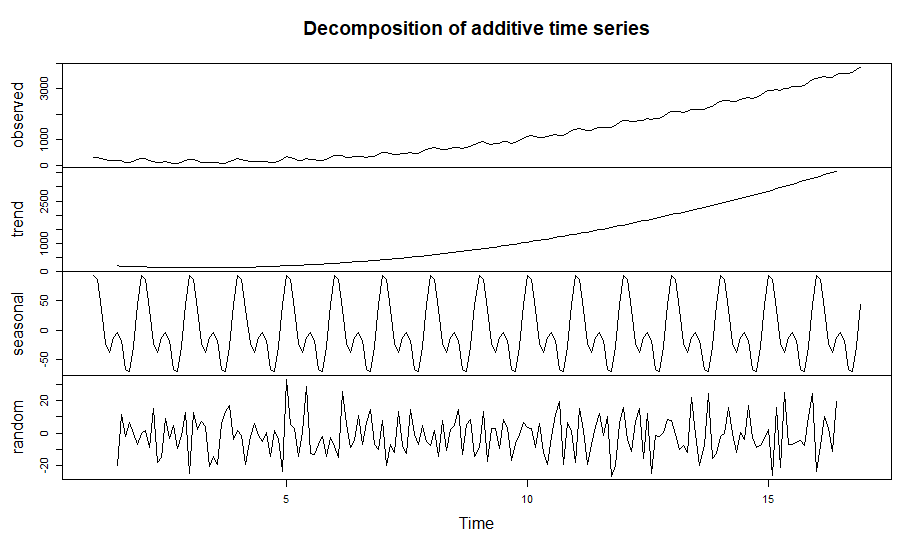


Рисунок 3. Декомпозиция временного ряда

На рисунке 3 изображен график декомпозиции временного ряда. На основании графика можно сделать следующие выводы:

* Трендовая компонента показывает возрастающий тренд. Это свидетельствует о том, что уровни временного ряда растут с течением времени. Тренд, является нелинейным, поскольку график имеет плавное, слегка ускоряющееся повышение.
* Сезонная компонента на графике представлена периодическими колебаниями с четкой периодичностью. Амплитуда колебаний примерно постоянна, что говорит о стабильности сезонной составляющей.
* Случайная компонента (шум) характеризуется небольшими, непостоянными колебаниями, которые не показывают четкой структуры. Это типично для шумовой компоненты, которая включает непредсказуемые отклонения, не объясняемые трендом или сезонностью.
* На графике тренда не наблюдаются горизонтальные или вертикальные асимптоты. Это говорит о том, что в текущем временном интервале тренд растет без замедления.

1. **Построение автокорреляционной и частной автокорреляционной функции временного ряда (рис.4).**



Рисунок 4. Автокорреляционные функции

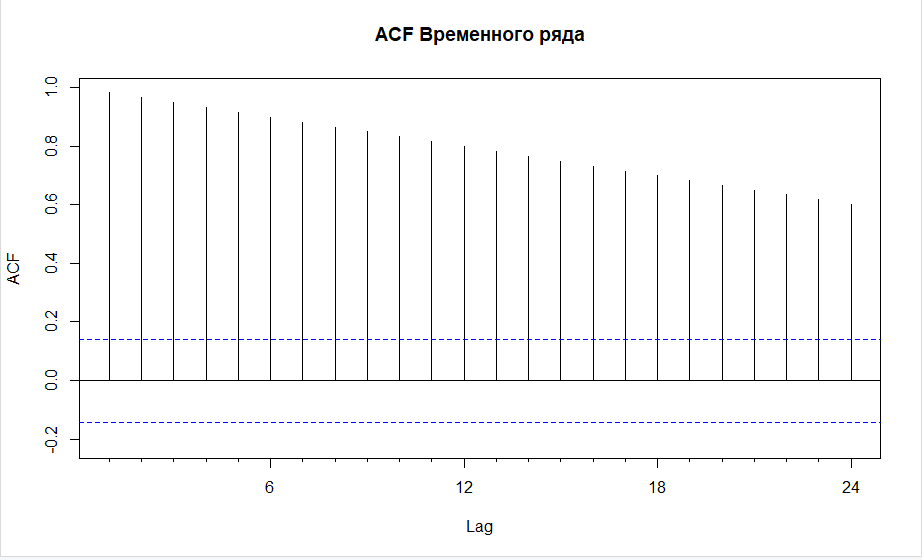


Рисунок 5. Автокорреляционная функция

Автокорреляционная функция убывает очень медленно

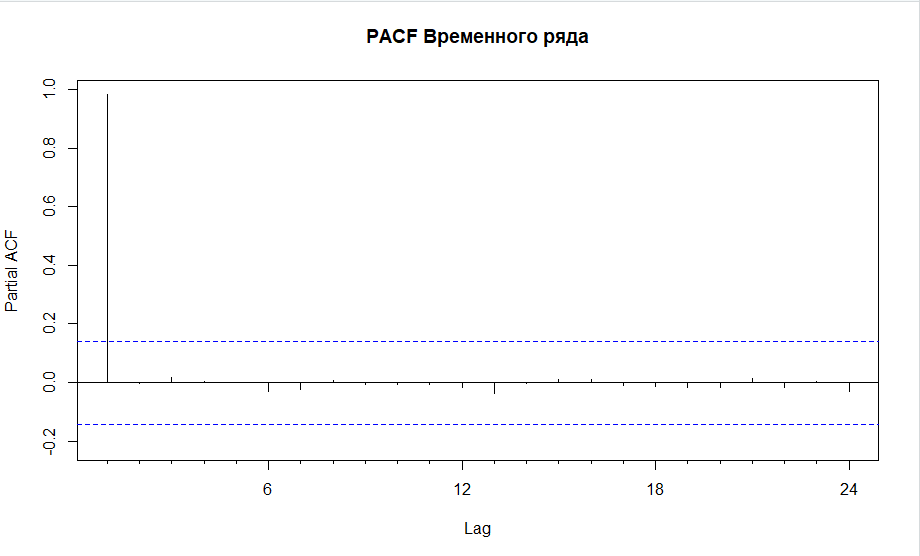


Рисунок 6. Частная автокорреляционная функция

Частная автокорреляционная функция имеет пик при lag = 1;

1. **Исследование и модельное описание временного ряда на основе метода последовательной идентификации: Структурная модель тренда временного ряда: полином второй степени (не линейная)**



Рисунок 7. Идентификация тренда

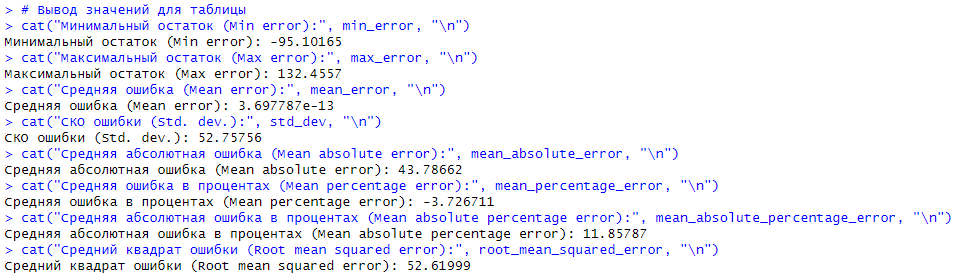


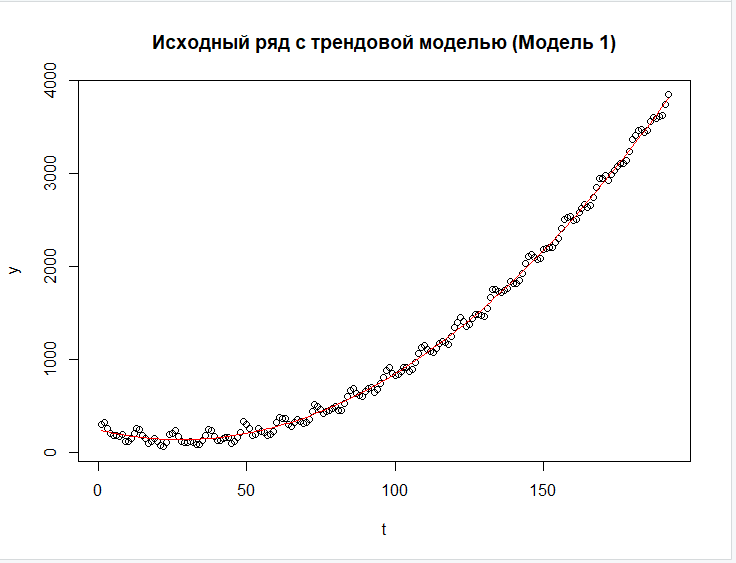
Рисунок 8. Значения для таблицы

Рисунок 9. График исходного ряда с наложением тренда

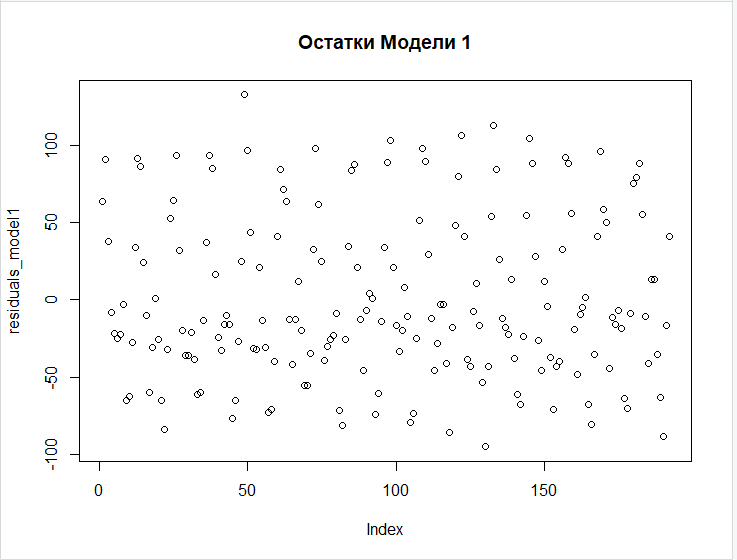


Рисунок 10. Остатки модели

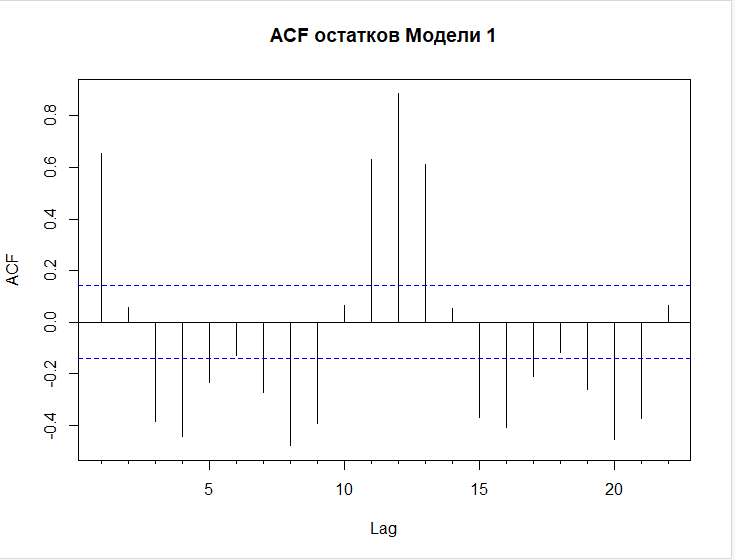


Рисунок 11. Автокорреляционная функция остатков модели

При наложении тренда все равно остаются пики на автокорреляционной функции.

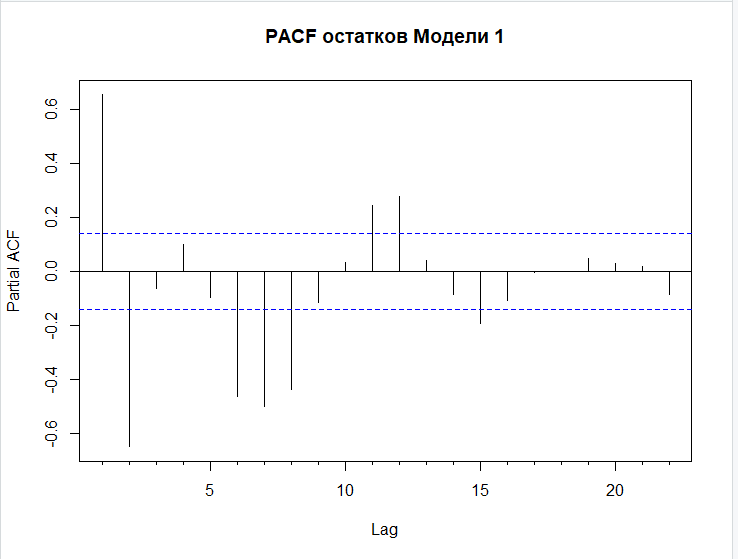


Рисунок 12. PACF остатков модели

При наложении тренда все равно остаются пики на частной автокорреляционной функции

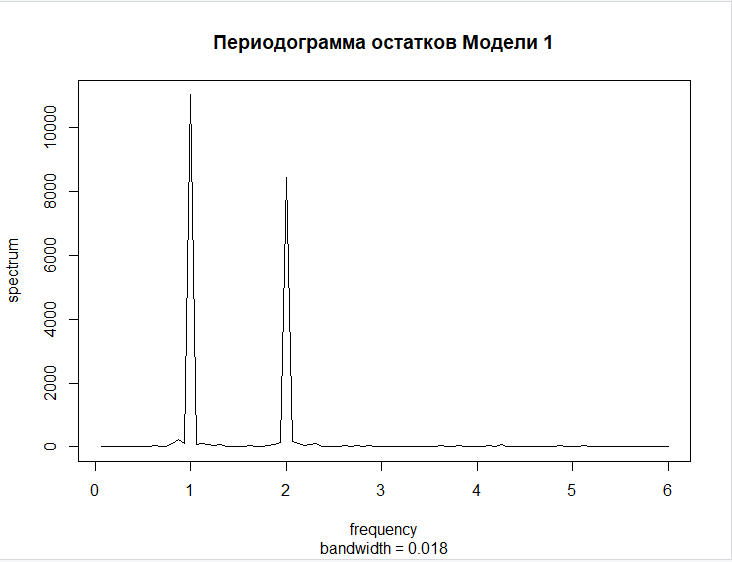


Рисунок 13. Периодограмма остатков модели

1. **Идентификация сезонной составляющей временного ряда**

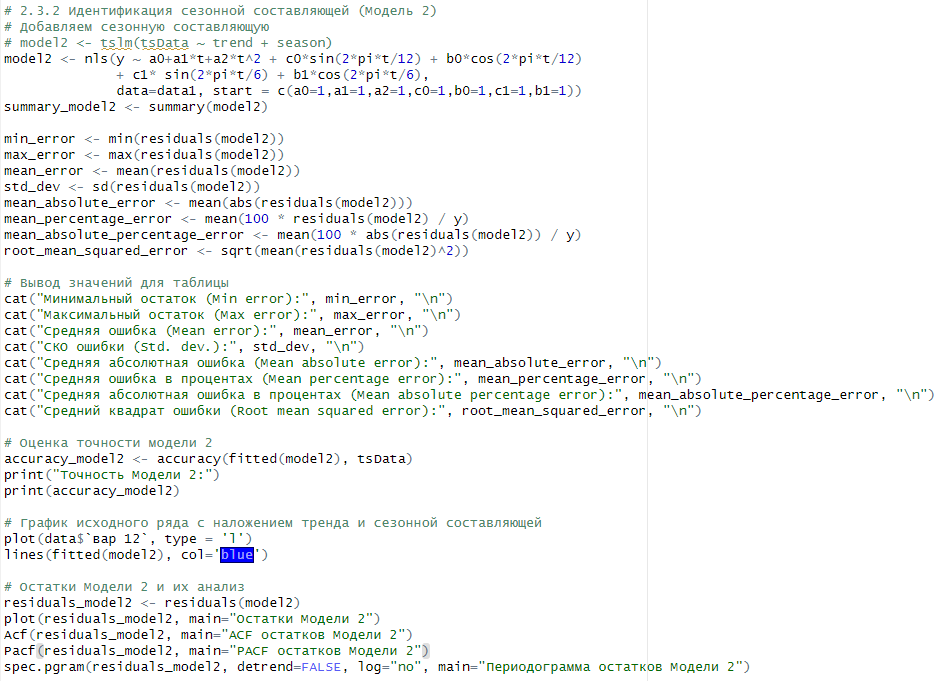


Рисунок 14. Идентификация сезонной составляющей

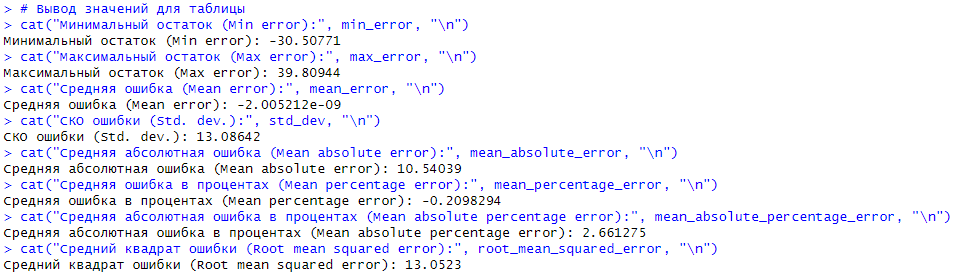


Рисунок 15. Значения для таблицы

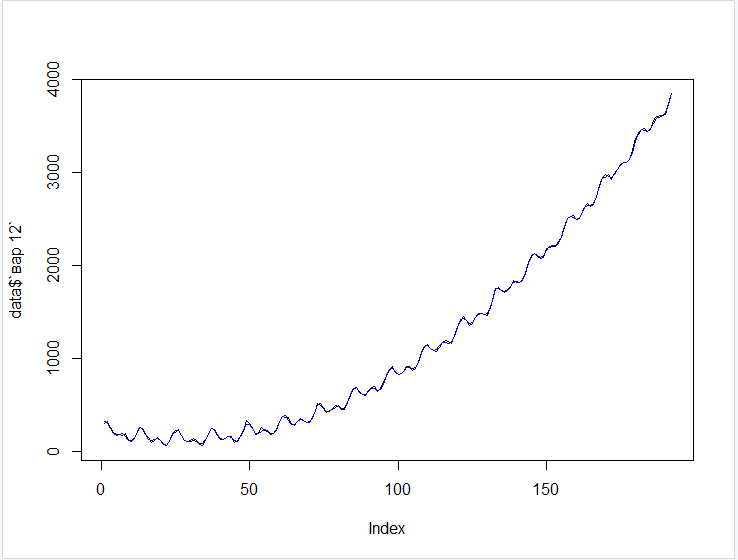


Рисунок 16. График исходного ряда с наложением тренда и сезонной составляющей

При добавлении сезонности количество остатков заметно уменьшилось.

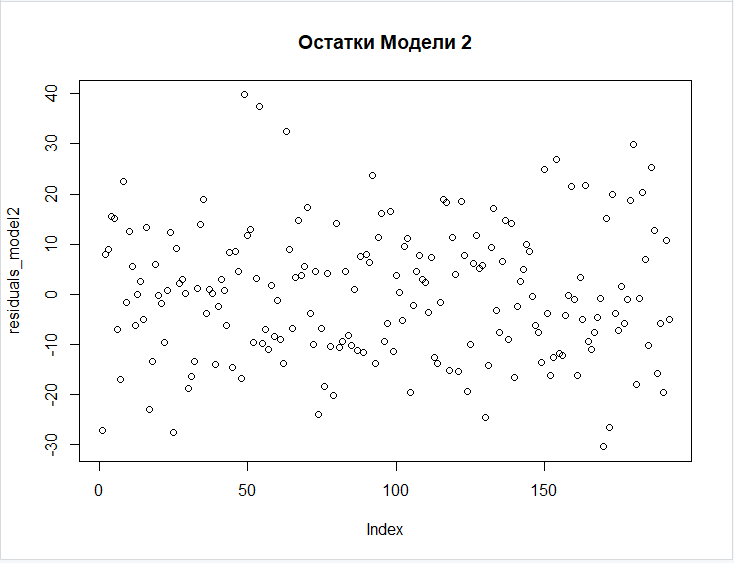


Рисунок 17. Остатки модели

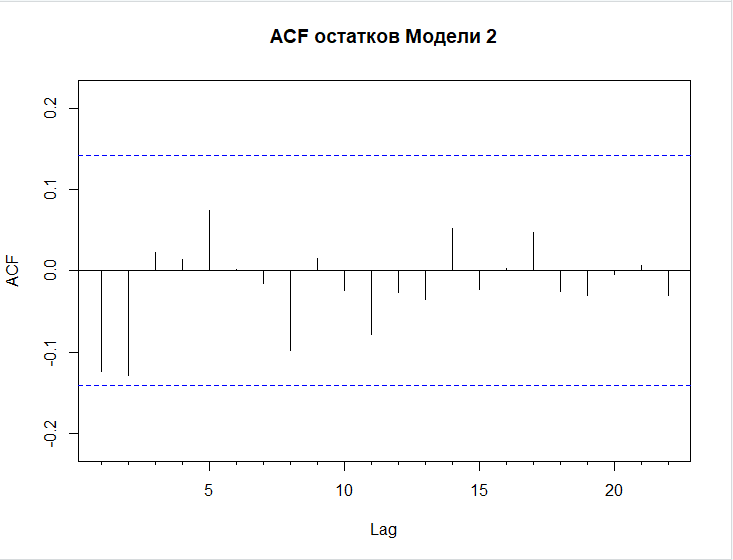


Рисунок 18. ACF остатки модели

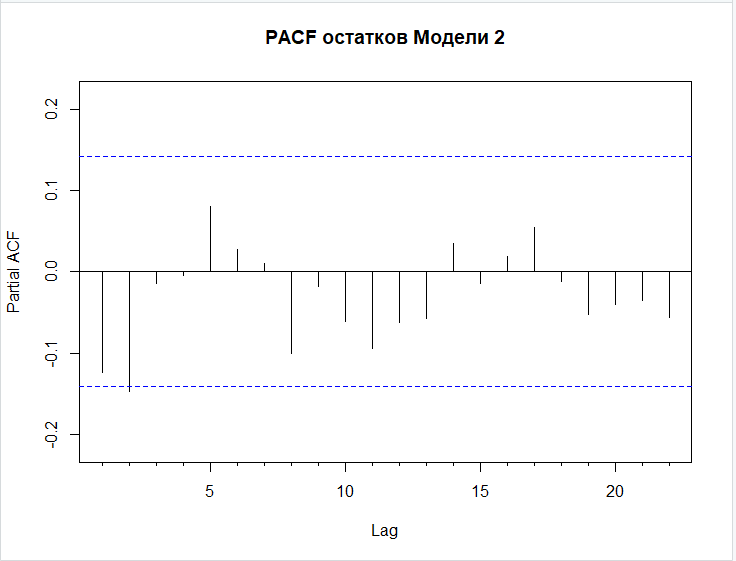


Рисунок 19. PACF остатки модели

Пики автокорреляции не выходят за пределы, это значит, что автокорреляции на этих лагах статистически незначимы. В остатках нет значимой автокорреляции.



Рисунок 20. Периодограмма модели

Таблица 1. Характеристики точности прогнозных моделей

|  |  |  |  |
| --- | --- | --- | --- |
|  | Модель1 | Модель2 | Модель3 |
| 1.Прогнозная модель | (Полином второй степени) модель тренда | (Полином второй степени) тренд + сезонная составляющая | Не была взята |
| 2. Минимальный остаток (Min error) | -95.10165 | -30.50771 |  |
| 3.Максимальный остаток (Max error) | 132.4557 | 39.80944 |  |
| 4. Средняя ошибка (Mean error) | 3.697787e-13 | -2.005212e-09 |  |
| 5. СКО ошибки (Std. dev.) | 52.75756 | 13.08642 |  |
| 6. Средняя абсолютная ошибка (Mean absolute error) | 43.78662 | 10.54039 |  |
| 7. Средняя ошибка в процентах (Mean percentage error) | -3.726711 | -0.2098294 |  |
| 8. Средняя абсолютная ошибка в процентах (Mean absolute percentage error) | 11.85787 | 2.661275 |  |
| 9. Средний квадрат ошибки (Root mean squared error) | 52.61999 | 13.0523 |  |
| 10. Коэффициент детерминации | 9.977248e-01 | 9.998600e-01 |  |

Модель 1 имеет более широкий диапазон остатков, что указывает на большую изменчивость ошибок модели по сравнению с моделью 2. Модель, демонстрирует гораздо лучшую точность по сравнению с моделью 1 по всем показателям точности. Она имеет меньшую изменчивость ошибок, меньшие средние абсолютные ошибки и более низкие значения RMSE, что указывает на её более высокую стабильность и предсказательную способность.

**5. Прогноз значений ВР на 3 шага вперед на основе полученной итоговой модели.**

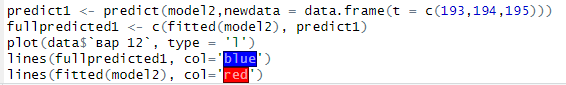
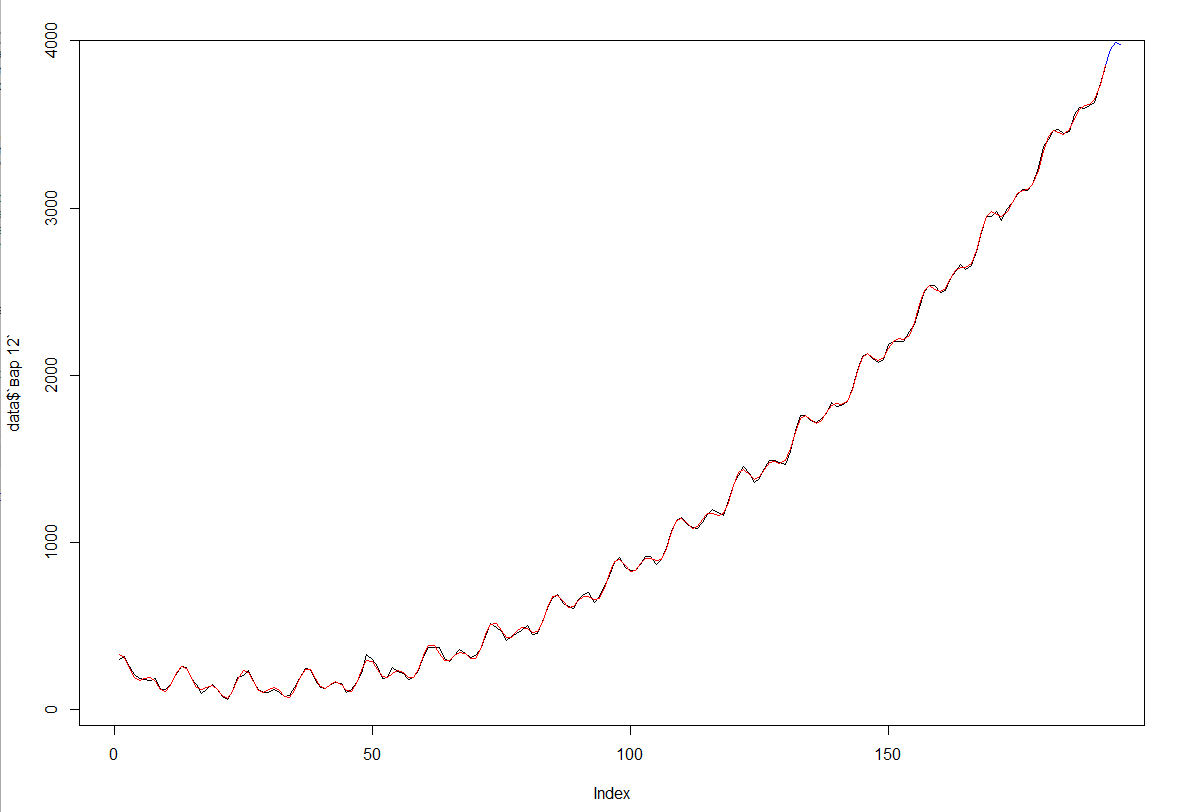


Рисунок 26. Прогноз.



Сделали прогноз, и ряд аппроксимируется практически идеально

**Заключение**

В ходе выполнения лабораторной работы были изучены методы и алгоритмы прогнозирования временных рядов, исследована эффективность прогнозирования временных рядом.

Вывод: при добавлении тренда и сезонности, остатки не содержат значимой автокорреляции.

**Приложение**

# Загрузка библиотек

library(forecast)

library(tseries)

# Загрузка данных

data <- read.csv("data.csv", header=TRUE, sep=";", check.names=F, fileEncoding="Windows-1251")

tsData <- ts(data[15], frequency=12) # Загрузка 12 варианта временного ряда

# 2.2 Построение графика временного ряда и его декомпозиция

plot.ts(tsData, main="Временной ряд")

decomp <- decompose(tsData)

plot(decomp)

# Построение автокорреляционной и частной автокорреляционной функций

Acf(tsData, main="ACF Временного ряда")

Pacf(tsData, main="PACF Временного ряда")

# 2.3.1 Идентификация тренда (Модель 1)

# Используем модель тренда

t <- 1:length(tsData)

y <- as.numeric(tsData)

data1 <- data.frame(y=y, t=t)

model1 <- nls(y ~ a0+a1\*t+a2\*t^2, data=data1, start = c(a0=1,a1=1,a2=1))

summary\_model1 <- summary(model1)

# Оценка точности модели

accuracy\_model1 <- accuracy(fitted(model1), tsData)

print("Точность Модели 1:")

print(accuracy\_model1)

min\_error <- min(residuals(model1))

max\_error <- max(residuals(model1))

mean\_error <- mean(residuals(model1))

std\_dev <- sd(residuals(model1))

mean\_absolute\_error <- mean(abs(residuals(model1)))

mean\_percentage\_error <- mean(100 \* residuals(model1) / y)

mean\_absolute\_percentage\_error <- mean(100 \* abs(residuals(model1)) / y)

root\_mean\_squared\_error <- sqrt(mean(residuals(model1)^2))

# Вывод значений для таблицы

cat("Минимальный остаток (Min error):", min\_error, "\n")

cat("Максимальный остаток (Max error):", max\_error, "\n")

cat("Средняя ошибка (Mean error):", mean\_error, "\n")

cat("СКО ошибки (Std. dev.):", std\_dev, "\n")

cat("Средняя абсолютная ошибка (Mean absolute error):", mean\_absolute\_error, "\n")

cat("Средняя ошибка в процентах (Mean percentage error):", mean\_percentage\_error, "\n")

cat("Средняя абсолютная ошибка в процентах (Mean absolute percentage error):", mean\_absolute\_percentage\_error, "\n")

cat("Средний квадрат ошибки (Root mean squared error):", root\_mean\_squared\_error, "\n")

# График исходного ряда с наложением тренда

plot(t, y, main="Исходный ряд с трендовой моделью (Модель 1)")

lines(t, predict(model1), col="red")

# Остатки Модели 1 и их анализ

residuals\_model1 <- residuals(model1)

plot(residuals\_model1, main="Остатки Модели 1")

Acf(residuals\_model1, main="ACF остатков Модели 1")

Pacf(residuals\_model1, main="PACF остатков Модели 1")

residuals\_model12 <- ts(residuals\_model1, frequency = 12)

spec.pgram(residuals\_model12, detrend=FALSE, log="no", main="Периодограмма остатков Модели 1")

# 2.3.2 Идентификация сезонной составляющей (Модель 2)

# Добавляем сезонную составляющую

# model2 <- tslm(tsData ~ trend + season)

model2 <- nls(y ~ a0+a1\*t+a2\*t^2 + c0\*sin(2\*pi\*t/12) + b0\*cos(2\*pi\*t/12)

+ c1\* sin(2\*pi\*t/6) + b1\*cos(2\*pi\*t/6),

data=data1, start = c(a0=1,a1=1,a2=1,c0=1,b0=1,c1=1,b1=1))

summary\_model2 <- summary(model2)

min\_error <- min(residuals(model2))

max\_error <- max(residuals(model2))

mean\_error <- mean(residuals(model2))

std\_dev <- sd(residuals(model2))

mean\_absolute\_error <- mean(abs(residuals(model2)))

mean\_percentage\_error <- mean(100 \* residuals(model2) / y)

mean\_absolute\_percentage\_error <- mean(100 \* abs(residuals(model2)) / y)

root\_mean\_squared\_error <- sqrt(mean(residuals(model2)^2))

# Вывод значений для таблицы

cat("Минимальный остаток (Min error):", min\_error, "\n")

cat("Максимальный остаток (Max error):", max\_error, "\n")

cat("Средняя ошибка (Mean error):", mean\_error, "\n")

cat("СКО ошибки (Std. dev.):", std\_dev, "\n")

cat("Средняя абсолютная ошибка (Mean absolute error):", mean\_absolute\_error, "\n")

cat("Средняя ошибка в процентах (Mean percentage error):", mean\_percentage\_error, "\n")

cat("Средняя абсолютная ошибка в процентах (Mean absolute percentage error):", mean\_absolute\_percentage\_error, "\n")

cat("Средний квадрат ошибки (Root mean squared error):", root\_mean\_squared\_error, "\n")

# Оценка точности модели 2

accuracy\_model2 <- accuracy(fitted(model2), tsData)

print("Точность Модели 2:")

print(accuracy\_model2)

# График исходного ряда с наложением тренда и сезонной составляющей

plot(data$`вар 12`, type = 'l')

lines(fitted(model2), col='blue')

# Остатки Модели 2 и их анализ

residuals\_model2 <- residuals(model2)

plot(residuals\_model2, main="Остатки Модели 2")

Acf(residuals\_model2, main="ACF остатков Модели 2")

Pacf(residuals\_model2, main="PACF остатков Модели 2")

spec.pgram(residuals\_model2, detrend=FALSE, log="no", main="Периодограмма остатков Модели 2")

# 2.4 Сравнение моделей и прогноз на 3 шага вперед на основе Модели 2

print("Сравнение точности моделей:")

print(accuracy\_model1)

print(accuracy\_model2)

# Прогноз на 3 шага вперед на основе модели 2

#forecast\_model2 <- forecast(model2, h=3)

#print("Прогноз на 3 шага вперед:")

#print(forecast\_model2)

#plot(forecast\_model2)

predict1 <- predict(model2,newdata = data.frame(t = c(193,194,195)))

fullpredicted1 <- c(fitted(model2), predict1)

plot(data$`вар 12`, type = 'l')

lines(fullpredicted1, col='blue')

lines(fitted(model2), col='red')