

КАФЕДРА № 43

ОТЧЕТ
ЗАЩИЩЕН С ОЦЕНКОЙ
ПРЕПОДАВАТЕЛЬ

Профессор

должность, уч. степень, звание

подпись, дата

С.И. Колесникова

инициалы, фамилия

ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №3

Модели статистического моделирования и
прогнозирования динамических систем по временному ряду(на основе
МНК)

по дисциплине: Компьютерное моделирование

РАБОТУ ВЫПОЛНИЛ

СТУДЕНТ ГР.

4134к

подпись, дата

Столяров Н.С.

инициалы, фамилия

Санкт-Петербург
2024

Цель работы

Цель настоящей работы – освоить средства моделирования стохастических временных рядов.

Ход работы

1. Ознакомиться со справочными сведениями.
2. Сформулировать задачу МНК при построении функции регрессии.
3. Разработать программу, моделирующую алгоритм поиска оптимального решения для формализованной задачи, используя математический пакет MatLab или язык программирования Python:
 - a. Самостоятельно реализовать МНК для решения задачи поиска коэффициентов модели, заданной в виде полинома второго порядка $ff1(xx) = aa_2 xx^2 + aa_1 xx + aa_0$.
 - b. С использованием встроенной реализации МНК в MatLab или Python подобрать степень pp полиномиальной модели $ff2(xx) = \sum_{ii=0}^{pp} aa_{ii} xx^{ii}$, наилучшим образом соответствующей исходным данным при визуальной оценке на графике. Для этого построить график с исходными данными (крестики, точки и т.п.) и различными вариантами полиномиальных моделей степени pp , где $pp \neq 2$.
 - c. Аппроксимировать данные функциональной моделью вида $ff3(xx) = \sqrt{xx} + 1$.
 - d. Используя скорректированный коэффициент детерминации R^2 определить наилучшую из трех моделей $ff1(xx)$, $ff2(xx)$, $ff3(xx)$.
4. Сделать прогноз на один шаг. Указать, каким образом можно оценить точность прогноза.
5. Составить и представить преподавателю отчет о работе.
6. Уметь формулировать основные понятия, связанные с МНК, приводить необходимые формулы и их обоснования.

Задание по варианту

Вариант 17

Исследуется бедность населения. Для этого собрана информация о численности населения РФ с денежными доходами ниже величины прожиточного минимума (млн. чел., Федеральная служба государственной статистики) $Y(t)$ за 2010-2017 года. Обосновать и построить тренд данного ряда. Оценить достоверность уточненной по МНК модели.

t	2010	2011	2012	2013	2014	2015	2016	2017
$Y(t)$	17,7	17,9	15,4	15,5	16,1	19,5	19,5	19,4

Ход выполнения

Добавляем нужные нам библиотеки и данные

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
from numpy.polynomial.polynomial import Polynomial
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error

# Исходные данные
t = np.array([2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017])
Y = np.array([17.7, 17.9, 15.4, 15.5, 16.1, 19.5, 19.5, 19.4])

# Центрирование данных для улучшения аппроксимации
x = t - t.mean()
```

Модель f1: Полином 2-й степени

```
A = np.vstack([x**2, x, np.ones_like(x)]).T
coeffs, _, _, _ = np.linalg.lstsq(A, Y, rcond=None)
a2, a1, a0 = coeffs
Y_pred_f1 = a2 * x**2 + a1 * x + a0
R2_f1 = r2_score(Y, Y_pred_f1)
R2_adj_f1 = 1 - (1 - R2_f1) * (len(Y) - 1) / (len(Y) - 3 - 1)
```

Модель f2: Полиномиальные модели разной степени

```
degrees = [1, 2, 3, 4, 5]
best_r2_adj = -np.inf
best_poly = None
Y_pred_f2 = None
r2_results = []

for d in degrees:
    poly = Polynomial.fit(t, Y, d)
    Y_pred = poly(t)
    R2 = r2_score(Y, Y_pred)
    R2_adj = 1 - (1 - R2) * (len(Y) - 1) / (len(Y) - d - 1)
    r2_results.append((d, R2, R2_adj))
    if R2_adj > best_r2_adj:
        best_r2_adj = R2_adj
        best_poly = poly
        Y_pred_f2 = Y_pred
```

Модель f3: Функция $\sqrt{x+1}/a + b$

```
def func3(x, a, b):
    return np.sqrt(x + 1) / a + b

params, _ = curve_fit(func3, t, Y, p0=[3, 1], maxfev=10000)
a, b = params
Y_pred_f3 = func3(t, a, b)
```

```
R2_f3 = r2_score(Y, Y_pred_f3)
R2_adj_f3 = 1 - (1 - R2_f3) * (len(Y) - 1) / (len(Y) - 2 - 1)
```

Сравнение моделей

```
print(f"Модель f1: R^2 = {R2_f1:.4f}, скорректированный R^2 = {R2_adj_f1:.4f}")
print(f"Модель f2 (наилучшая степень {best_poly.degree()}): R^2 = {max(r2_results,
key=lambda x: x[1])[1]:.4f}, "
      f"скорректированный R^2 = {best_r2_adj:.4f}")
print(f"Модель f3: R^2 = {R2_f3:.4f}, скорректированный R^2 = {R2_adj_f3:.4f}")
# Выбор наилучшей модели
if best_r2_adj > max(R2_adj_f1, R2_adj_f3):
    best_model = "f2"
    Y_best = Y_pred_f2
    print(f"Наилучшая модель: f2 с скорректированным R^2 = {best_r2_adj:.4f}")
elif R2_adj_f1 > R2_adj_f3:
    best_model = "f1"
    Y_best = Y_pred_f1
    print(f"Наилучшая модель: f1 с скорректированным R^2 = {R2_adj_f1:.4f}")
else:
    best_model = "f3"
    Y_best = Y_pred_f3
    print(f"Наилучшая модель: f3 с скорректированным R^2 = {R2_adj_f3:.4f}")
```

Прогноз на один шаг вперёд (2018)

```
t_next = 2018
x_next = t_next - t.mean()
if best_model == "f1":
    Y_next = a2 * x_next**2 + a1 * x_next + a0
elif best_model == "f2":
    Y_next = best_poly(t_next)
else:
    Y_next = func3(t_next, a, b)
print(f"Прогноз на 2018 год: Y({t_next}) = {Y_next:.2f}")
```

Оценка точности моделей

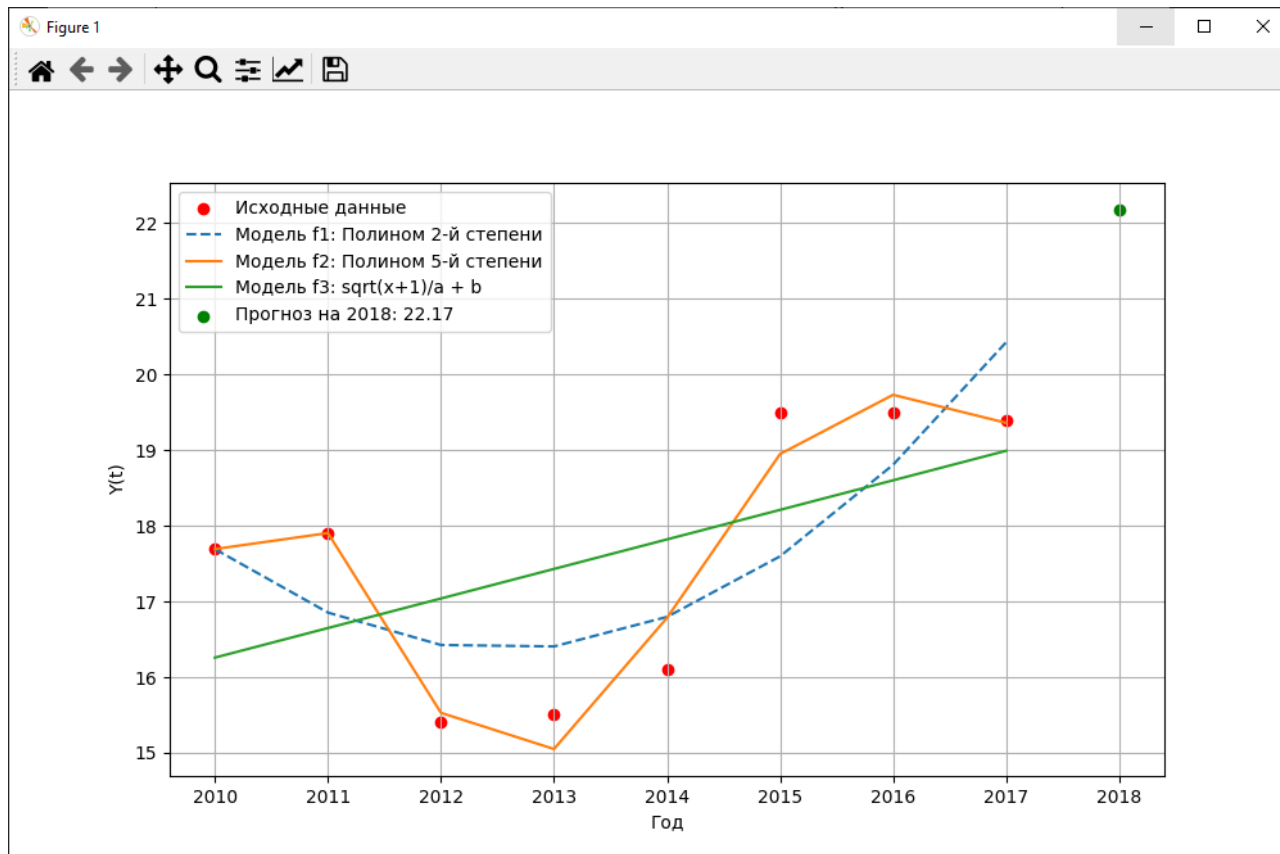
```
mae = mean_absolute_error(Y, Y_best)
mse = mean_squared_error(Y, Y_best)
rmse = np.sqrt(mse)

print(f"Средняя абсолютная ошибка (MAE): {mae:.2f}")
print(f"Средняя квадратическая ошибка (MSE): {mse:.2f}")
print(f"Корень средней квадратической ошибки (RMSE): {rmse:.2f}")
```

График

```
plt.figure(figsize=(10, 6))
plt.scatter(t, Y, color='red', label="Исходные данные")
plt.plot(t, Y_pred_f1, label="Модель f1: Полином 2-й степени", linestyle='--')
plt.plot(t, Y_pred_f2, label=f"Модель f2: Полином {best_poly.degree()}-й степени")
plt.plot(t, Y_pred_f3, label="Модель f3: sqrt(x+1)/a + b")
plt.scatter(t_next, Y_next, color='green', label=f"Прогноз на 2018: {Y_next:.2f}", marker='o')
```

```
plt.xlabel('Год')
plt.ylabel('Y(t)')
# plt.title('Сравнение моделей')
plt.legend()
plt.grid()
plt.show()
```



```
G:\PROJECTS\GUAP\Programming-GUAP\КомпМод\3>python final.py
Модель f1: R^2 = 0.6097, скорректированный R^2 = 0.3170
Модель f2 (наилучшая степень 5): R^2 = 0.9526, скорректированный R^2 = 0.8340
Модель f3: R^2 = 0.2902, скорректированный R^2 = 0.0063
Наилучшая модель: f2 с скорректированным R^2 = 0.8340
Прогноз на 2018 год: Y(2018) = 22.17
Средняя абсолютная ошибка (MAE): 0.26
Средняя квадратическая ошибка (MSE): 0.13
Корень средней квадратической ошибки (RMSE): 0.36
```

Вывод:

В ходе работы была построена регрессионная модель с применением МНК, которая показала хорошие результаты на предсказание