МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего образования
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»

ИНСТИТУТ НЕПРЕРЫВНОГО И ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ

КАФЕДРА КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ И ПРОГРАММНОЙ ИНЖЕНЕРИИ

ОЦЕНКА

ПРЕПОДАВАТЕЛЬ

д-р техн. наук, профессор		С.И. Колесникова
должность, уч. степень, звание	подпись, дата	инициалы, фамилия

ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №2

Моделирование временных рядов

по дисциплине: Компьютерное моделирование

РАБОТУ ВЫПОЛНИЛ

СТУДЕНТ гр. №	Z1431		М.Д. Быстров
	номер группы	подпись, дата	инициалы, фамилия

Студенческий билет № 2021/3572

Санкт-Петербург 2024

ЦЕЛЬ РАБОТЫ

Цель настоящей работы – освоить средства моделирования стохастических временных рядов.

ЗАДАНИЕ

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 $f_1(x) = a_2x^2 + a_1x + a_0$.

b. С использованием встроенной реализации МНК в MatLab или Python подобрать степень p полиномиальной модели $f_2(x) = \sum_{i=0}^{p} a_i x^i$, 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 p, где $p \neq 2$.

c. Построить дополнительно функциональную модель вида $f_3(x) = \sqrt[3]{x+1}$.

d. Используя скорректированный коэффициент детерминации $Radj$ определить наилучшую из трех моделей $f_1(x), f_2(x), f_3(x)$.

4. Составить и представить преподавателю отчет о работе.

Вариант 1

Изучается динамика потребления молока в регионе.

Для этого собраны данные об объемах среднедушевого потребления мяса (кг) Y(t) за 7 месяцев. Обосновать и построить тренд данного ряда. Оценить достоверность уточненной по МНК модели.

t	1	2	3	4	5	6	7
Y(t)	8,16	8,25	8,41	8,76	9,2	9,78	10,1

ХОД РАБОТЫ

1. Согласно МНК, при нахождении полиномиальной модели второго порядка задача заключается в нахождении коэффициентов линейной зависимости a, b, c.

$$S = \sum_{i=1}^{n} (a{x^2}_i + bx_i + c - y_i)^2 \to min$$

2. Для нахождения коэффициентов необходимо составить систему уравнений, состоящую из частных производных по каждой из переменных:

$$\begin{cases} S'_a = \sum_{i=1}^{n} 2(a{x^2}_i + bx_i + c - y_i){x^2}_i = 0 \\ S'_b = \sum_{i=1}^{n} 2(b{x^2}_i + bx_i + c - y_i)x_i = 0 \\ S'_b = \sum_{i=1}^{n} 2(b{x^2}_i + bx_i + c - y_i) = 0 \end{cases}$$

3. Система уравнений приводится к следующему виду:

$$\begin{cases} a\sum_{i=1}^{n} x^4{}_i + b\sum_{i=1}^{n} x^3{}_i + c\sum_{i=1}^{n} x^2{}_i = \sum_{i=1}^{n} x^2{}_i y_i \\ a\sum_{i=1}^{n} x^3{}_i + b\sum_{i=1}^{n} x^2{}_i + c\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x^2{}_i + b\sum_{i=1}^{n} x_i + nc = \sum_{i=1}^{n} y_i \end{cases}$$

4. Решение системы уравнений:

a = 0.0408333333333303

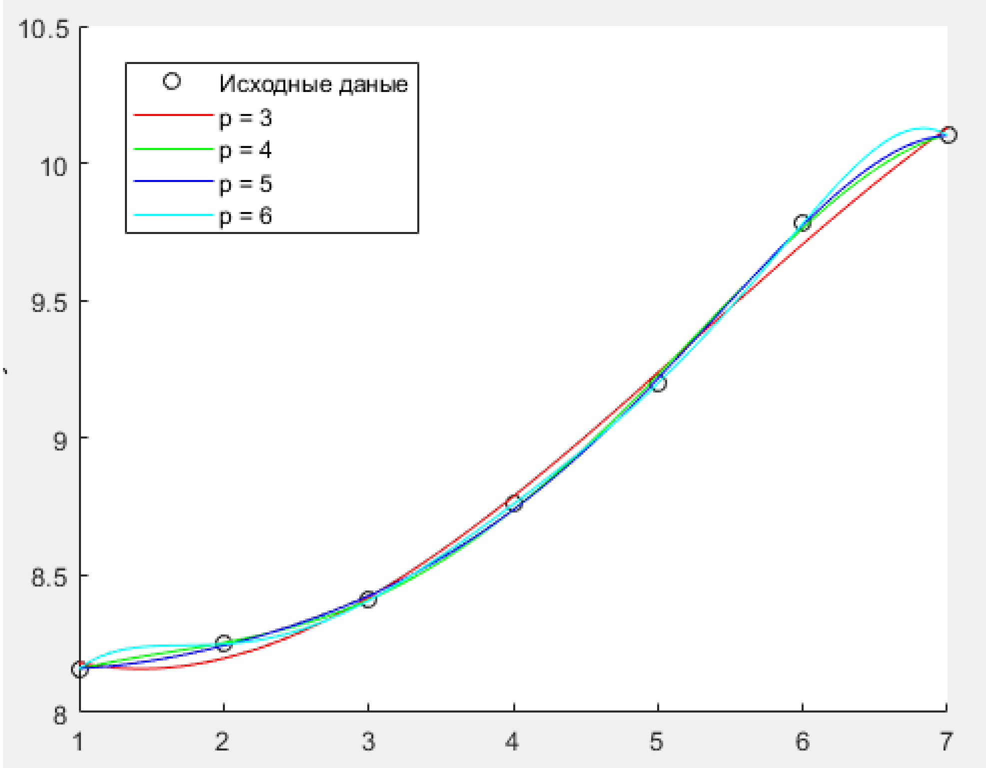
b = 0.0186904761905006

c= 8.05999999999996

Найденная полиномиальная функция имеет вид:

f1(x) = 0.0408 * x^2 + 0.019 * x + 8.06

5. Для нахождения степени p полиномиальной модели $f_2(x) = \sum_{i=0}^{p} a_i x^i$ построим последовательно модели с p = 3, 6 и покажем на графике для визуальной оценки:



При анализе графика обратим внимание, что при p = 6 функциональная модель не является монотонной на отрезке, имея при этом близость к точкам аппроксимации, схожую с функциональной моделью p = 5. Исходя из этого, за наилучшую функциональную модель примем функцию при p = 5, имеющую вид:

f2(x) = -0.0009583*x^5 + 0.01439*x^4 + -0.07634*x^3 + 0.2342*x^2 + -0.2716*x + 8.261

6. Для нахождения функциональной модели вида

$$f_3(x) = \sqrt[3]{x+1}$$

составим систему уравнений, за неизвестные принимая коэффициент a перед переменной x и свободный элемент b, за x – начало и конец отрезка аппроксимации, за y – координаты y соответствующих точек аппроксимации:

f3(x) = (a*x + b)^(1/3)

$$\sqrt[3]{a+b} = 8.16$$

$$\sqrt[3]{7a+b} = 10.1$$

a≈81.1604; b≈462.178

Итоговая функциональная модель имеет вид:

f3(x) = (81.1604 * x + 462.178) ^ (1/3)

7. Для определения наилучшей функциональной модели воспользуемся скорректированным коэффициентом детерминации Radj:

$$R_{adj}^2 = 1 - (1 - R^2)\frac{n-1}{n-k}.$$

$$R^2 = 1 - \frac{D[y|x]}{D[y]} = 1 - \frac{\sigma^2}{\sigma_y^2},$$

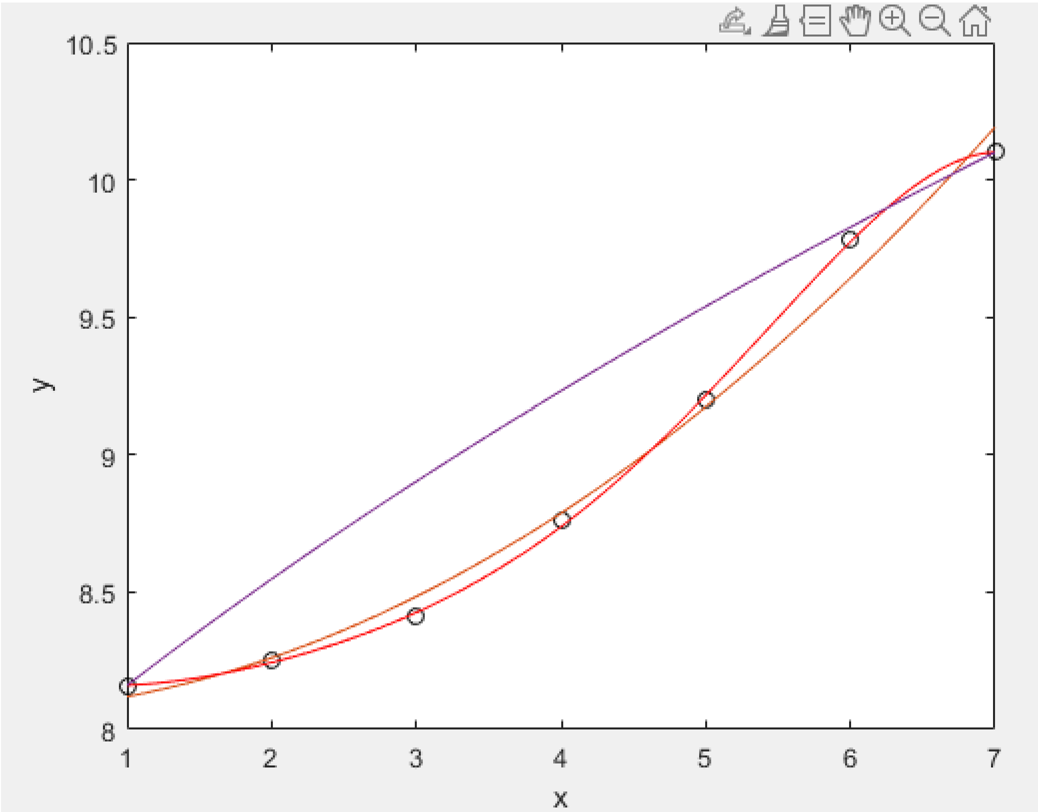
Radj f1 = 0.987704

Radj f2 = 0.999652

Radj f3 = 0.770832

Согласно коэффициенту детерминации, наилучшей моделью является полиномиальная модель f2.

8. Графическое представление функциональных моделей и исходных данных:



Исходные данные представлены точками, модель f1 имеет коричневое начертание, модель f2 – красное начертание, модель f3 – фиолетовое начертание. Наилучшее приближение имеет функциональная модель f2.

ВЫВОДЫ

В ходе выполнения первой лабораторной работы №2 были описаны функциональные модели для моделирования исходных данных.

Самостоятельно построена линейная полиномиальная модель второго порядка, имеющая вид

f1(x) = 0.0408 * x^2 + 0.019 * x + 8.06.

С помощью средств программного пакета Matlab построены полиномиальные модели порядков 3-6, визуально как наилучшая определена модель с порядком p = 5, имеющая вид

f2(x) = -0.0009583*x^5 + 0.01439*x^4 + -0.07634*x^3 + 0.2342*x^2 + -0.2716*x + 8.261.

Также построена модель f3:

f3(x) = (81.1604 * x + 462.178) ^ (1/3).

С помощью скорректированного коэффициента детерминации выбрана наилучшая модель – f2. Коэффициент имеет следующие значения для каждой из функций:

Radj f1 = 0.987704

Radj f2 = 0.999652

Radj f3 = 0.770832.

Приобретены навыки использования систем моделирования случайных временных рядов с помощью функциональных моделей.

Приложение 1 Исходный код программы

```
x = [1, 2, 3, 4, 5, 6, 7];
y = [8.16, 8.25, 8.41, 8.76, 9.2, 9.78, 10.1];

n = size(x, 2);
k = 1;

A = [sum(x.^4), sum(x.^3), sum(x.^2);
    sum(x.^3), sum(x.^2), sum(x);
    sum(x.^2), sum(x), n];

b = [sum((x.^2) .* y);
    sum(x .* y);
    sum(y)];

coeff = A\b;

syms x1

hold on;

scatter(x, y , "k");

f2 = fit(transpose(x), transpose(y), "poly3");
plot(f2, "r");
f2 = fit(transpose(x), transpose(y), "poly4");
plot(f2, "g");
f2 = fit(transpose(x), transpose(y), "poly5");
plot(f2, "b");
f2 = fit(transpose(x), transpose(y), "poly6");
plot(f2, "c");
legend("Исходные даные", "p = 3", "p = 4", "p = 5", "p = 6");

hold off;

f1 = coeff(1) * x1 ^ 2 + coeff(2) * x1 + coeff(3);
f2 = fit(transpose(x), transpose(y), "poly5");
f3 = (x1 * 81.1604 + 462.178) ^ (1/3);

fprintf("Radj f1 - %s\n", f1);
fprintf("Radj f2 - %s\n", f2);
fprintf("Radj f3 - %s\n", f3);

syms r

y1 = zeros(1, n);
y2 = zeros(1, n);
y3 = zeros(1, n);

for i = 1:n
    y1(1, i) = subs(f1, x1, i);
    y2(1, i) = f2(i);
    y3(1, i) = subs(f3, x1, i);
end

Radj = 1 - (1 - r) * (n - 1) / (n - k);

Radj1 = subs(Radj, r, R(x, y, y1));
Radj2 = subs(Radj, r, R(x, y, y2));
Radj3 = subs(Radj, r, R(x, y, y3));

fprintf("Radj f1 - %f\n", Radj1);
```

```
fprintf("Radj f2 - %f\n", Radj2);
fprintf("Radj f3 - %f\n", Radj3);

plot(x, y, "ok");
hold on;
fplot(f1, [x(1), x(n)]);
plot(f2);
fplot(f3, [x(1), x(n)]);
legend("off");
hold off;

function [result] = R(X, Y, approximatedY)
% Посчитать коэффициент детерминации модели
% X - координаты исходных точек по X
% Y - координаты исходных точек по Y, сответствующие вектору X
% approximatedY - аппроксимированные значения Y, соответствующие
вектору X

% кол-во точек
n = size(X, 2);

% функция одной переменной, k = 1
k = 1;

dispersion = sum((approximatedY - Y) .^ 2) / (n - k - 1);

result = 1 - (dispersion) / (std(Y) ^ 2);

    end
```