МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное автономное бюджетное образовательное учреждение высшего образования

Севастопольский государственный университет

кафедра Информационных систем

**Золотарь Дмитрий Сергеевич**

Институт информационных технологий и управления в технических системах

курс 4 группа ИC/м-11(о)

09.03.02. Информационные системы и технологии

КУРСОВАЯ РАБОТА

по дисциплине «Специальные главы математики»

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Листов 19

Отметка о зачёте \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_

(дата)

Руководитель практикума

доцент Карлусов В. Ю.

(должность) (подпись) (инициалы, фамилия)

Севастополь

2016

**АННОТАЦИЯ**

В данной пояснительной записке рассматривается процесс исследования сгенерированных временных рядов на стационарность, построения для них моделей ARIMA и векторной авторегрессии, а также проверки наличия коинтеграции. В приложении приведены разработанные сценарии выполнения Matlab.

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 4](#_Toc469436505)

[1. ПОСТАНОВКА ЗАДАЧИ 5](#_Toc469436506)

[2. СОДЕРЖАТЕЛЬНАЯ ЧАСТЬ ПОЯСНИТЕЛЬНОЙ ЗАПИСКИ ПО РАЗДЕЛАМ ИНДИВИДУАЛЬНОГО ЗАДАНИЯ 7](#_Toc469436507)

[2.1. Определение параметров исследуемых временных рядов 7](#_Toc469436508)

[2.2. Нахождение выборочных АКФ и ЧАКФ 7](#_Toc469436509)

[2.3. Построение модели ARIMA 8](#_Toc469436510)

[**2.3.1. Преобразование ряда к стационарному виду** 8](#_Toc469436511)

[**2.3.2. Подбор параметров модели** 8](#_Toc469436512)

[**2.3.3. Проверка адекватности моделей** 10](#_Toc469436513)

[**2.4. Построение модели векторной авторегрессии** 12](#_Toc469436514)

[**2.5.** **Проверка наличия коинтеграции между временными рядами** 14](#_Toc469436515)

[ЗАКЛЮЧЕНИЕ 15](#_Toc469436516)

[БИБЛИОГРАФИЧЕСКИЙ СПИСОК 16](#_Toc469436517)

[ПРИЛОЖЕНИЕ А 17](#_Toc469436518)

Изм.

Лист

№ докум.

Подпись

Дата

Лист

3

КУРСОВАЯ РАБОТА

Разраб.

Золотарь Д. С.

Провер.

*Карлусов В. Ю.*

Н. Контр.

Утверд.

ПОЯСНИТЕЛЬНАЯ

ЗАПИСКА

Лит.

Листов

Группа ИС/м-11o

# **ВВЕДЕНИЕ**

Временной ряд представляет собой последовательность данных, описывающих объект в последовательные моменты времени. В отличие от анализа случайных выборок, анализ временных рядов основывается на предположении, что последовательные данные наблюдаются через равные промежутки времени.

Существует две основные цели анализа временных рядов: **определение природы ряда и прогнозирование**, т.е. предсказание будущих значений временного ряда по настоящим и прошлым значениям. Обе цели требуют, чтобы модель ряда была определена и более или менее формально описана. Как только модель определена, с ее помощью можно интерпретировать рассматриваемые данные – например, использовать ее для анализа наличия сезонного изменения цен на товары. Затем можно экстраполировать ряд на основе найденной модели, т.е. предсказать его будущие значения.

В ходе работы рассматривается процесс исследования сгенерированных временных рядов при помощи выборочных АКФ и ЧАКФ, построение моделей ARIMA и векторной авторегрессии. Проведена проверка коинтеграции рядов.

# **ПОСТАНОВКА ЗАДАЧИ**

В ходе выполнения курсовой работы необходимо выполнить следующие задания:

* Определить коэффициенты корреляции для каждой пары параметров. Сделать заключение о связях параметров.
* Провести первичную обработку данных, по возможности представив их в наглядном виде. Для этого используются методы описательной статистики: группировка данных, их графическое представление, вычисление различных показателей, описывающих положение данных на числовой оси, степень их разброса, симметрии и т.п.
* Построить графики выборочных АКФ и ЧАКФ.
* Для анализа поведения временного ряда надо выполнить:
  + преобразование ряда к стационарному виду;
  + идентификацию модели, т.е. подбор порядка модели p и q;
  + проверку адекватности модели.
* Построить векторную модель авторегрессии исследуемых временных рядов.
* Проверить наличие коинтеграции между нестационарными временными рядами.

Исследование проводилось для двух рядов, сгенерированных при помощи заданного сценария Matlab. Первый из представленных сигналов использовался при выполнении всех пунктов работы. Второй – при построении векторной модели авторегрессии и проверке наличия коинтеграции. Ниже представлены графики исследуемых рядов.

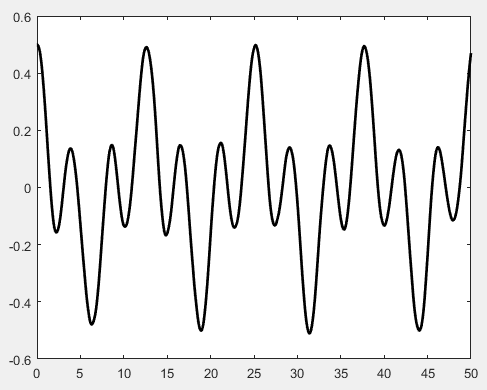


Рисунок 1 – Первый сгенерированный ряд

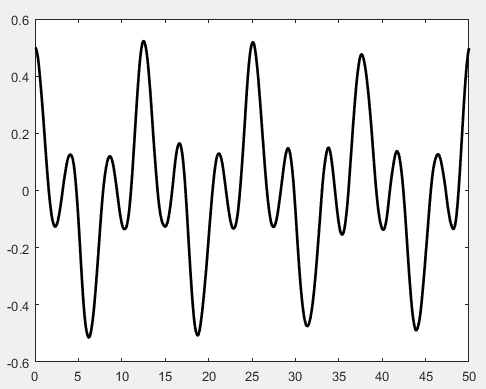


Рисунок 2 – Второй сгенерированный ряд

# **СОДЕРЖАТЕЛЬНАЯ ЧАСТЬ ПОЯСНИТЕЛЬНОЙ ЗАПИСКИ ПО РАЗДЕЛАМ ИНДИВИДУАЛЬНОГО ЗАДАНИЯ**

* 1. Определение параметров исследуемых временных рядов

Для каждого из рядов найдено математическое ожидание и среднеквадратическое отклонение.

Для первого ряда:

μ = -0.0018

σ = 0.0621

Для второго ряда:

μ = -0.0036

σ = 0.0617

* 1. **Нахождение выборочных АКФ и ЧАКФ**

В этом и следующем разделе будем использовать только первый ряд.

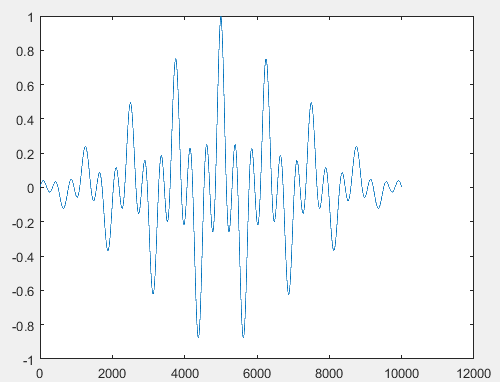


Рисунок 3 – График выборочной АКФ

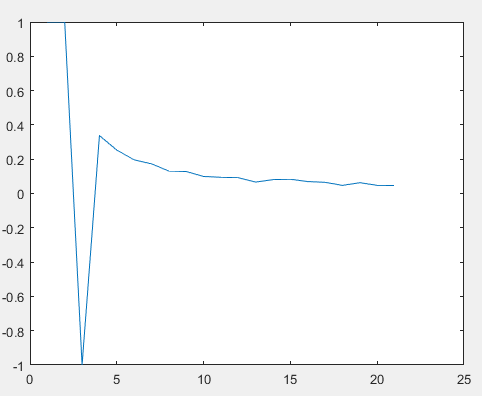


Рисунок 4 – График выборочной ЧАКФ (отрезок)

* 1. Построение модели ARIMA
     1. **Преобразование ряда к стационарному виду**

Произведена проверка ряда на стационарность при помощи теста Дики-Фулера, который дал отрицательный результат.

Переход от нестационарных процессов к стационарным осуществляется при помощи взятия разностей порядка d[1].

Взята первая разность ряда и также протестирована на стационарность. Получен положительный результат.

Из этого можно сделать вывод что параметр d для модели ARIMA должен быть равен единице.

* + 1. **Подбор параметров модели**

Как было сказано выше параметр d=1. Для подбора параметров p и q построим выборочные АКФ и ЧАКФ для первой разности ряда.

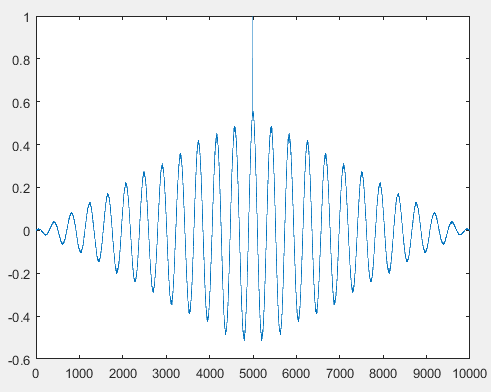


Рисунок 5 – Выборочная АКФ первой разности ряда

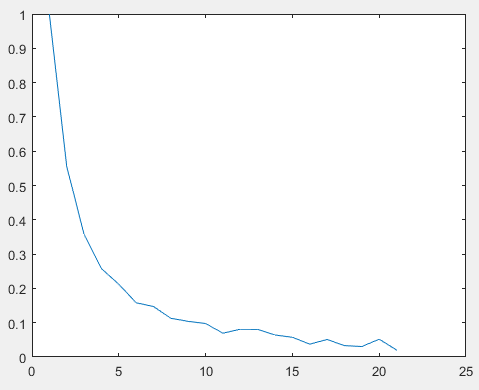


Рисунок 6 – Выборочная ЧАКФ первой разности ряда

В то время как автокорреляционная функция процесса авторегрессии порядка р спадает плавно, ее частная автокорреляционная функция имеет обрыв после р-й задержки. Обратно, автокорреляционная функция процесса скользящего среднего порядка q обрывается после задержки q, в то время как ее частная автокорреляция плавно спадает с ростом задержки. Далее, автокорреляционная функция смешанного процесса, содержащая компоненту авторегрессии порядка р и компоненту скользящего среднего порядка q, после первых q - р задержек представляется в виде суммы экспонент и затухающих синусоид. Обратно, частная автокорреляционная функция смешанного процесса приближенно представляется суммой экспонент и затухающих синусоид после р - q задержек[2].

Выборочная АКФ экспоненциально спадает без задержек, а ЧАКФ имеет ярко выраженный спад с задержкой 1. Следовательно p=q≠0.

Далее будем исследовать модели ARIMA(1, 2, 1) и ARIMA(2, 2, 2), так как модели более высоких порядков неоправданно сложны.

* + 1. **Проверка адекватности моделей**

После идентификации модели и оценки параметров подгоняемая модель подвергается диагностической проверке. Нужно ответить на вопрос, адекватна ли модель. Ели будут обнаружены свидетельства серьезной неадекватности, возникнет необходимость узнать, как нужно изменить модель на следующем итеративном цикле[3].

Для проверки адекватности моделей была выбрана последняя 50-ю часть ряда и выполнена попытка спрогнозировать её значения при помощи оставшейся. Ниже приведены графики, но который синим обозначен исследуемый ряд, зелёным – средний прогноз, красным доверительный интервал.

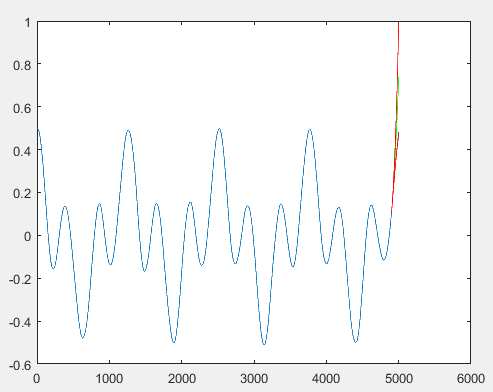


Рисунок 7 – Прогноз ряда для модели ARIMA(1, 2, 1)

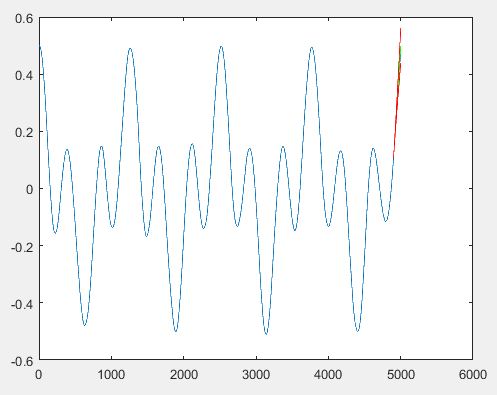


Рисунок 8 – Прогноз ряда для модели ARIMA(2, 2, 2)

Для оценки адекватности модели вычислим коэффициенты автокорреляции остатков для обоих моделей.

Автокорреляция остатков модели ARIMA(1, 2, 1):

cov(r1) = 0.0066

Автокорреляция остатков модели ARIMA(2, 2, 2):

cov(r2) = 0.0016

Корреляция отсутствует, следовательно модели адекватны. Тем не менее предпочтительнее использовать модель ARIMA(1, 2, 1), так как она обладает меньшим количеством параметров и, соответственно, вычислительной сложностью.

* 1. Построение модели векторной авторегрессии

Модель векторной авторегрессии описывает авторегрессию двух рядов. Её порядок был взят равным двум. Параметры были рассчитаны при помощи системы Matlab.

Вектор констант равен:

Матрицы авторегрессии:

Построим прогноз последней 50-й части рядов по оставшейся.

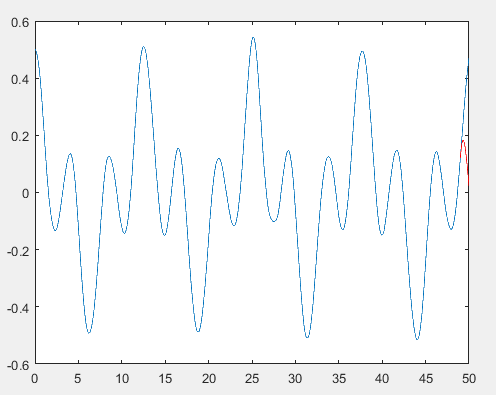


Рисунок 9 – Прогноз первого ряда

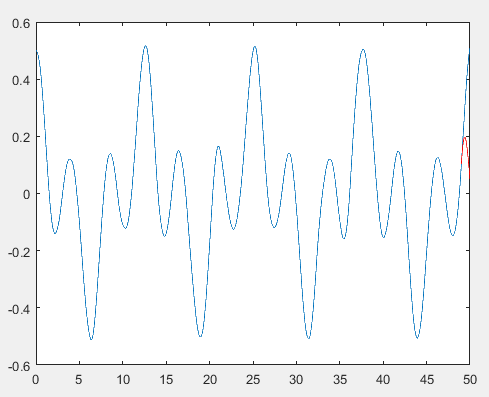


Рисунок 10 – Прогноз второго ряда

* 1. **Проверка наличия коинтеграции между временными рядами**

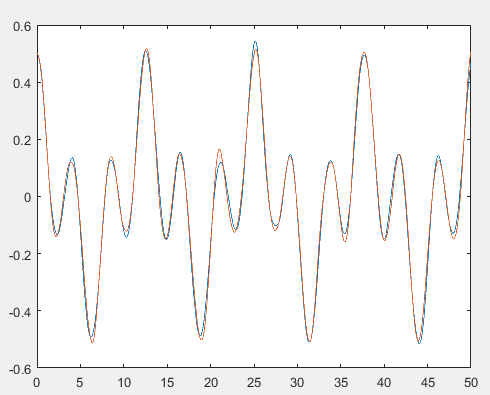


Рисунок 11 – Графики рядов, проверяемых на коинтеграцию

Для проверки наличия коинтеграции использовался тест Энгла-Грейнджера. Расчёты проводились в среде Matlab. Тест показал отсутствие коинтеграции между рядами.

# **ЗАКЛЮЧЕНИЕ**

При выполнении курсовой работы был проведён анализ сгенерированных временных рядов.

Были найдены математическое ожидание и среднеквадратическое отклонение.

Построены выборочные автокоррелционная и частная автокорреляционная функции.

Также было выполнено построение моделей ARIMA. Было определено, что исходные ряды не являются стационарными, поэтому для анализа пришлось взять разности первого порядка. Ряды разностей первого порядка получились стационарными и прошли тест Дики-Фуллера на стационарность. Это означает, что оба данных ряда являются интегрированными рядами первого порядка.

Анализ АКФ и ЧАКФ рядов первых разностей показал, что для исследуемых рядов целесообразно использовать модели ARIMA(1, 2, 1) и ARIMA(2, 2, 2), однако предпочтительнее использовать первую из-за малого количества параметров.

Также была построена модель векторной авторегрессии для исследуемых рядов. Был выбран порядок модели, равный 2. Полученная модель хорошо интерполирует ряды, а также в своем прогнозе угадывает общий тренд обоих рядов.

Было определено, что ряды не обладают свойством коинтеграции.

Таким образом выполнены все задачи курсового проектирования.

**БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

1. Многомерные модели нестационарных временных рядов. Методические указания и индивидуальные задания к курсовой работе по дисциплине «Специальные главы математики» / Сост. Первухина Е.Л., Голикова В.В. – Севастополь: Изд-во СевНТУ, 2015. – 21 с.
2. Специальные статистические приложения и модели. Методические указания и индивидуальные задания к курсовой работе по дисциплине «Специальные главы математики» / Сост. Первухина Е.Л., Голикова В.В. – Севастополь: Изд-во СевНТУ, 2014. – 14 с.
3. Компьютерные методы анализа данных и прогнозирования. Методические указания и индивидуальные задания к курсовой работе по дисциплине «Интеллектуальный анализ данных» / Сост. Первухина Е.Л., Токарев А.И., Голикова В.В. – Севастополь: Изд-во СевНТУ, 2015. – 22 с.

**ПРИЛОЖЕНИЕ А**

**Разработанные сценарии выполнения Matlab.**

Текст файла generateProcess.m:

function [y\_sum, t\_model] = generateProcess

n\_max=5000; % число отсчётов анализируемого процесса

t0=0.01; % шаг квантования входного процесса

t\_max=t0\*n\_max; % время моделирования

t\_model=[0.0:t0:t\_max]; % временная шкала

% параметры фильтра, формирующего полезный процесс 1

k\_us=1.0;% коэффициент усиления фильтра-прототипа

wf=2.5; %собственная частота фильтра-прототипа (аналогового)

psi=0.05; % коэффициент затухания

tf=2.0\*pi/wf;% период колебаний

wft0=wf\*t0;

% коэффициенты ЦФ

koef\_a=[1.0+2.0\*psi\*wft0+wft0^2,-2.0\*(1.0+psi\*wft0),1.0];

koef\_b=[k\_us\*(t0^2)\*(2.0\*psi\*tf^2)];

% Белый шум

x\_s=rand(1,length(t\_model))-0.5;

y\_s1=filter(koef\_b,koef\_a,x\_s); % процесс1

figure

plot(t\_model,y\_s1,'k-','LineWidth',2);

y\_s2=diric(t\_model,4); % процесс 2, накладываемый

figure

plot(t\_model,y\_s2,'k-','LineWidth',2);

% веса процессов

alfa=2.0;

beta=0.5;

y\_sum=alfa\*y\_s1+beta\*y\_s2; %итоговый процесс

figure

plot(t\_model,y\_sum,'k-','LineWidth',2);

end

Текст файла arimaPrediction.m:

function [y\_pred, residuals] = arimaPrediction(p, d, q, y)

model = arima(p, d, q);

modelSize = size(y);

n = modelSize(2);

predictSize = round(n / 50);

est = estimate(model, y(1:(n - predictSize))');

y\_pred = simulate(est, predictSize, 'NumPaths', 1000, 'Y0', y(1:(n - predictSize))');

lower = prctile(y\_pred, 2.5, 2);

upper = prctile(y\_pred, 97.5, 2);

mn = mean(y\_pred,2);

figure

plot(y)

hold on

plot((n - predictSize + 1):n, mn, 'g')

plot((n - predictSize + 1):n, lower, 'r')

plot((n - predictSize + 1):n, upper, 'r')

residuals = y((n - predictSize + 1):n) - mn';

end

Текст файла vectorAutoregressionPrediction.m:

function [EstSpec1, FY] = vectorAutoregressionPrediction(t\_model, Y)

modelSize = size(Y);

count = round(modelSize(1));

n = modelSize(2);

predictSize = round(n / 50);

dt = logical(eye(count));

VAR2diag = vgxset('ARsolve',repmat({dt},count,1),'asolve',logical(ones(1,count)));

EstSpec1 = vgxvarx(VAR2diag,Y(:,1:(n - predictSize))');

FY = vgxpred(EstSpec1,predictSize,[],Y(:,1:(n - predictSize))');

vgxdisp(EstSpec1)

figure

plot(t\_model, Y(1,:))

hold on

plot(t\_model((n - predictSize + 1):n), FY(:,1)', 'r')

figure

plot(t\_model, Y(2,:))

hold on

plot(t\_model((n - predictSize + 1):n), FY(:,2)', 'r')

end

Текст файла kurs.m:

clear all

close all

clc

[y\_sum, t\_model] = generateProcess;

y\_sum2 = generateProcess;

modelSize = size(t\_model);

n = modelSize(2);

m1 = mean(y\_sum)

sigma1 = var(y\_sum)

m2 = mean(y\_sum2)

sigma2 = var(y\_sum2)

acf = crosscorr(y\_sum, y\_sum, (n - 1));

figure

plot(acf)

pacf = parcorr(y\_sum);

%parcorr(y\_sum, (n - 1));

figure

plot(pacf)

i = 0;

y\_integr = y\_sum;

while not(adftest(y\_integr))

y\_integr = diff(y\_integr);

i = i + 1;

end

acfi = crosscorr(y\_integr, y\_integr, (n - i - 1));

figure

plot(acfi)

pacfi = parcorr(y\_integr);

figure

plot(pacfi)

p = 1;

q = 1;

[y\_pred, residuals] = arimaPrediction(p, i, q, y\_sum);

p1q1cov = cov(residuals)

p=2;

q=2;

[y\_pred, residuals] = arimaPrediction(p, i, q, y\_sum);

p2q2cov = cov(residuals)

Y = [y\_sum; y\_sum2];

vectorAutoregressionPrediction(t\_model, Y);

figure

plot(t\_model,Y(1,:),t\_model,Y(2,:))

cointegration = egcitest(Y')