

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ

«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИМЕНИ
АКАДЕМИКА С.П.КОРОЛЁВА»

ИНСТИТУТ ИНФОРМАТИКИ, МАТМАТИКИ И ЭЛЕКТРОНИКИ

ФАКУЛЬТЕТ ИНФОРМАТИКИ

КАФЕДРА ТЕХНИЧЕСКОЙ КИБЕРНЕТИКИ

СТАТИСТИЧЕСКИЙ АНАЛИЗ И МОДЕЛИРОВАНИЕ ПРОЦЕССОВ
АВТОРЕГРЕССИИ И СКОЛЬЗЯЩЕГО СРЕДНЕГО

курсовая работа по дисциплине «**Теория случайных процессов**»

Вариант № 51

Выполнил: Белоусов А.А.

Группа: 6309

№ зачетной книжки: 146166

Проверил: Храмов А. Г.

Оценка: _____

Дата: _____

Самара 2018

РЕФЕРАТ

Курсовая работа по курсу «Теория случайных процессов» 57 страниц, 6 рисунков, 9 таблиц, 3 источника, 2 приложения.

МОМЕНТНЫЕ ФУНКЦИИ, МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ, ДИСПЕРСИЯ, КОРРЕЛЯЦИОННАЯ ФУНКЦИЯ, МОДЕЛЬ АВТОРЕГРЕССИИ, МОДЕЛЬ СКОЛЬЗЯЩЕГО СРЕДНЕГО, МОДЕЛЬ ARMA

Объектом исследований является выборка из $n = 5000$ последовательных значений стационарного в широком смысле эргодичного случайного процесса с дискретным временем (временной ряд).

Цель работы – оценка моментных функций, построение и исследование моделей авторегрессии, скользящего среднего, смешанных моделей, моделирование случайных процессов.

Были рассчитаны выборочные моментные функции, построены и исследованы модели АР, СС, АРСС. Проведено моделирование процессов АР(3), СС(2), АРСС(3, 3).

Расчеты и построение графиков производились с использованием языка программирования Python 3.6 и библиотек numpy, matplotlib.

СОДЕРЖАНИЕ

1	Задание	4
1.1	Оценивание моментных функций	4
1.2	Построение и исследование моделей авторегрессии	4
1.3	Построение и исследование моделей скользящего среднего	4
1.4	Построение и исследование смешанных моделей авторегрессии – скользящего среднего	4
1.5	Сравнительный анализ построенных моделей	4
1.6	Итоговая таблица сравнения моделей АРСС	5
2	Исходные данные	6
3	Оценивание моментных функций	7
3.1	Моментные функции	7
3.2	Графическое представление	10
4	Построение и исследование моделей авторегрессии	12
4.1	Модель АРСС($M, 0$)	12
4.2	Расчёт теоретической НКФ выходной последовательности	13
4.3	Выбор наилучшей модели случайного процесса в классе моделей АР	14
5	Построение и исследование моделей скользящего среднего	15
5.1	Модель АРСС($0, N$)	15
5.2	Расчёт теоретической НКФ выходной последовательности	16
5.3	Выбор наилучшей модели случайного процесса в классе моделей СС	17
6	Построение и исследование смешанных моделей авторегрессии - скользящего среднего	18
6.1	Модель АРСС(M, N)	18
6.2	Построение модели АРСС	18
6.3	Расчёт теоретической НКФ выходной последовательности	21
6.4	Выбор наилучшей модели случайного процесса в классе моделей АРСС . .	21
7	Уравнения моделирования процессов	22
8	Сравнительный анализ построенных моделей	23
9	Итоговая таблица сравнения моделей АРСС	27
	Заключение	29
	Список литературы	30
	ПРИЛОЖЕНИЕ А Код программы	31
	ПРИЛОЖЕНИЕ Б Метод простых итераций	57

1 ЗАДАНИЕ

1.1 Оценивание моментных функций

Изобразить графически фрагмент исходного случайного процесса (СП). Оценить моментные функции (МФ) исходного случайного процесса, рассчитав выборочные среднее, дисперсию и нормированную корреляционную функцию (НКФ). Оценить интервал корреляции СП. Изобразить графически оценку НКФ исходного СП.

1.2 Построение и исследование моделей авторегрессии

Построить модели авторегрессии $AP(M) = APCC(M, 0)$ порядков $M = 0, 1, 2, 3$ (всего 4 модели) на основе решения системы уравнений Юла–Уокера. Для каждой модели рассчитать теоретические НКФ выходной последовательности. На основе сравнения выборочной НКФ и теоретических НКФ выбрать лучшую модель СП в классе моделей АР.

1.3 Построение и исследование моделей скользящего среднего

Построить модели скользящего среднего $CC(N) = APCC(0, N)$ порядков $N = 0, 1, 2, 3$ (всего 4 модели) на основе решения системы нелинейных уравнений. Для каждой модели рассчитать теоретические НКФ выходной последовательности. На основе сравнения выборочной НКФ и теоретических НКФ выбрать лучшую модель СП в классе моделей СС.

1.4 Построение и исследование смешанных моделей авторегрессии – скользящего среднего

Построить смешанные модели авторегрессии – скользящего среднего $APCC(M, N)$ до третьего порядка включительно ($M = 1, 2, 3; N = 1, 2, 3$) (всего 9 моделей) одним из методов, описанным в приложении А.3. Рассчитать теоретические НКФ выходной последовательности для каждой модели АРСС. На основе сравнения исходной выборочной и теоретических НКФ выбрать лучшую модель СП в классе смешанных моделей АРСС. Исследовать на устойчивость смешанные модели.

1.5 Сравнительный анализ построенных моделей

Для каждой из трёх лучших моделей (АР, СС, АРСС) записать системы уравнений для расчёта параметров модели, записать системы уравнений для расчёта теоретической КФ, смоделировать СП, рассчитать выборочные МФ, сравнить их с выборочными МФ исходного СП и с теоретическими МФ. Для каждой из этих трёх моделей сравнить графически НКФ: (1) выборочную исходного СП, (2) теоретическую, (3) выборочную смоделированного СП.

1.6 Итоговая таблица сравнения моделей АРСС

Изготовить таблицу сравнения МФ и расчёта качества для трёх лучших моделей. Изобразить графически фрагмент реализации СП, сгенерированного по наилучшей модели.

2 ИСХОДНЫЕ ДАННЫЕ

Дана реализация стационарного в широком смысле эргодического случайного процесса с дискретным временем (стационарная случайная последовательность, временной ряд) – выборка из $n = 5000$ последовательных значений (отсчётов) процесса:

-26.364 -31.116 -16.294 -33.269 -12.440 -36.868 -21.972 -29.761 -14.903 -33.905 -13.594 -37.352
-20.000 -32.988 -13.362 -23.077 -14.026 -23.744 -31.442 -39.546 -36.071 -20.985 -21.972 -23.090
-25.598 -20.838 -21.724 -21.249 -27.790 -23.872 -27.504 -27.615 -29.436 -31.517 -17.928
...
-29.333 -23.756 -16.822 -27.577 -20.445 -17.546 -30.783 -18.043 -33.864 -18.692 -35.802 -15.114
-36.468 -17.033 -41.076 -14.229 -36.877 -18.515 -36.372 -10.348 -28.443 -13.548 -38.428 -21.716
-37.438 -13.258 -33.314 -7.763 -33.864 -15.854 -35.060 -24.056 -29.442 -21.433 -25.555

3 ОЦЕНИВАНИЕ МОМЕНТНЫХ ФУНКЦИЙ

3.1 Моментные функции

Пусть дана выборка $X_1 \dots X_n$ из n отсчётов стационарного в широком смысле эргодического дискретного случайного процесса. Оценим моментные функции этого процесса, рассчитав:

1. Выборочное среднее
2. Выборочную дисперсию
3. Нормированную корреляционную функцию (НКФ).

1. Выборочное среднее \bar{X} - это оценка математического ожидания $\mathbb{M}(X)$. Формула для расчётов:

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k, \quad (1)$$

где X_k - элементы выборки при $k \in [0, n]$, а n - размер выборки.

Используя программу из приложения А, получаем $\bar{X} = -25.07220$.

На рисунке 1 показано значение \bar{X} .

2. Выборочная дисперсия S^2 - это оценка дисперсии $D(\bar{X}) = D_x$ распределения на основе выборки. Различают два вида дисперсии – выборочную и несмещённую.

Формула для расчёта выборочной дисперсии:

$$S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2. \quad (2)$$

Формула для расчёта несмещённой или исправленной дисперсии:

$$\hat{S}_n^2 = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2. \quad (3)$$

В данном случае была использована исправленная дисперсия, т.к. она даёт несмещённую оценку дисперсии:

$$M(\hat{S}_n^2) = D_x. \quad (4)$$

Применяя формулу (3), получаем $\hat{S}_n^2 = 45.94676$.

Для вычисления среднеквадратического отклонения (СКО) используется следующая формула:

$$\hat{S}_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2}. \quad (5)$$

Принимая во внимание значение \hat{S}_n^2 , получаем $\hat{S}_n = 6.77840$.

Графически изобразим фрагмент (первые 150 значений) исходного случайного процесса на рисунке 1.

Для нахождения верхней и нижней границ коридора наиболее вероятных значений процесса воспользуемся следующей формулой:

$$\bar{X} \pm \hat{S}_n = [-31.8506; -18.2938] \quad (6)$$

Отобразим найденные границы коридора наиболее вероятных значений случайного процесса на рисунке 1.

3. Формула для вычисления оценки исправленной выборочной корреляционной функции $\hat{R}_\xi(k)$:

$$\hat{R}_\xi(k) = \frac{1}{n-k-1} \sum_{j=1}^{n-k} (X_j - \bar{X})(X_{j+k} - \bar{X}). \quad (7)$$

В таблице 1 представлены значения этой функции для $k = 0, 1, \dots, 10$.

Оценка нормированной корреляционной функции $\hat{r}_\xi(k)$ определяется следующим образом:

$$\hat{r}_\xi(k) = \frac{\hat{R}_\xi(k)}{\hat{R}_\xi(0)} = \frac{\hat{R}_\xi(k)}{\hat{S}_n^2}. \quad (8)$$

Её значения для $k = 0, 1, \dots, 10$ также представлены в таблице 1.

Таблица 1 – Значения выборочных КФ и НКФ

k	$\hat{R}_\xi(k)$	$\hat{r}_\xi(k)$
0	45.94676	1.0
1	-16.83377	-0.36638
2	20.06587	0.43672
3	-20.94726	-0.4559
4	11.22639	0.24433
5	-14.17409	-0.30849
6	8.02627	0.17469
7	-8.94403	-0.19466
8	5.63892	0.12273
9	-5.55074	-0.12081
10	3.49849	0.07614

Оценим интервал корреляции:

$$\tau_k = \min\{\tau : \forall(m > \tau) |r_\eta(m)| < \frac{1}{e}\}, \quad (9)$$

полагая что $m \ll n$.

Используя программу из приложения А, получаем значение $\tau_k = 3$. На рисунке 2 изображена оценка интервала корреляции исходного случайного процесса.

3.2 Графическое представление

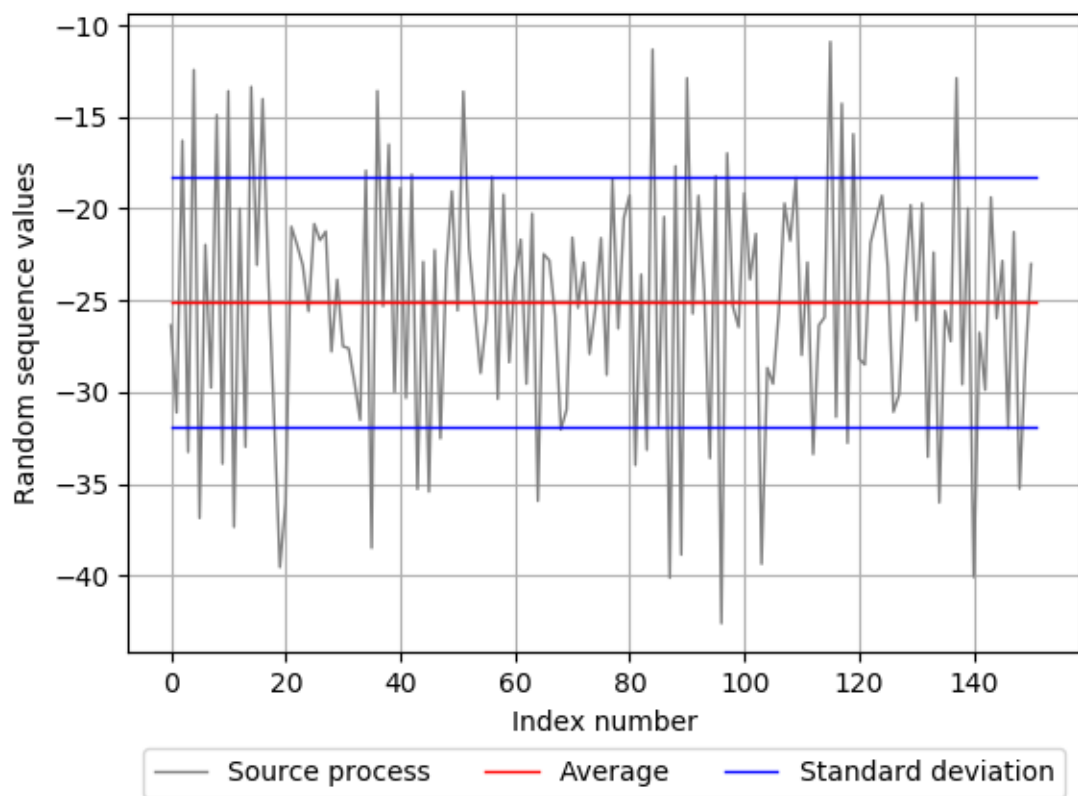


Рисунок 1 – График первых 150 значений исходного случайного процесса

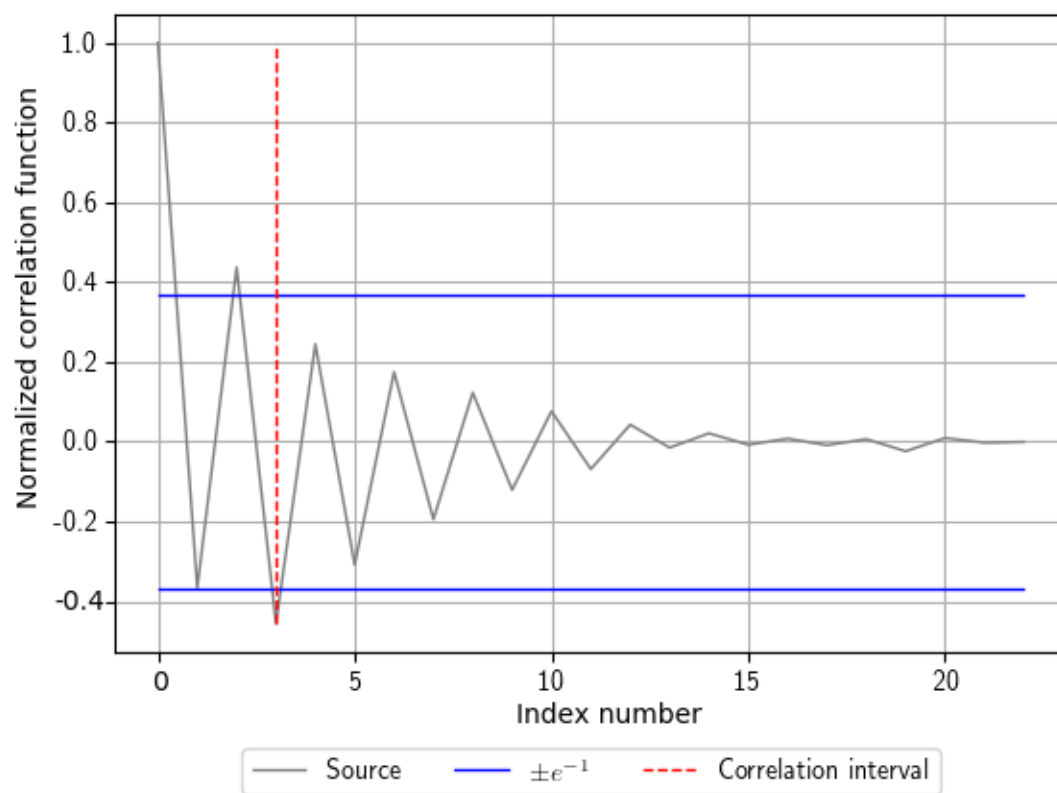


Рисунок 2 – Графическая оценка интервала корреляции

4 ПОСТРОЕНИЕ И ИССЛЕДОВАНИЕ МОДЕЛЕЙ АВТОРЕГРЕССИИ

4.1 Модель АРСС($M, 0$)

Модель авторегрессии порядка M в общем виде может быть записана как:

$$\eta_n = \beta_1 \eta_{n-1} + \beta_2 \eta_{n-2} + \dots + \beta_M \eta_{n-M} + \alpha_0 \xi_n, \quad (10)$$

где ξ – входная некоррелированная случайная последовательность с нулевым математическим ожиданием и единичной дисперсией ($\mathbb{M}(\xi_n) = 0$, $\mathbb{M}(\xi_n^2) = 1$, $\mathbb{M}(\xi_k \xi_n) = \delta_{kn}$), η – выходная случайная, стационарная в широком смысле последовательность с корреляционной функцией $R_\eta(m)$ и нулевым математическим ожиданием, $\{M, \alpha_0, \beta_1, \beta_2, \dots, \beta_M\}$ – параметры модели авторегрессии.

Построим модели авторегрессии для $M = 0, 1, 2, 3$ на основе решения системы уравнений Юла-Уокреа. Для решения линейных систем был использован метод `numpy.linalg.solve` из библиотеки `numpy`. Выпишем системы уравнений.

Для модели АРСС(0, 0) получаем:

$$R_\eta(0) = \alpha_0^2 \quad (11)$$

Решая уравнение (11), получаем $\alpha_0 = 6.77840$.

Для модели АРСС(1, 0) получаем:

$$\begin{cases} R_\eta(0) = \beta_1 R_\eta(1) + \alpha_0^2 \\ R_\eta(1) = \beta_1 R_\eta(0) \end{cases} \quad (12)$$

Решая систему (12), получаем $\alpha_0 = 6.30708$, $\beta_1 = -0.36638$.

Для модели АРСС(2, 0) получаем:

$$\begin{cases} R_\eta(0) = \beta_1 R_\eta(1) + \beta_2 R_\eta(2) + \alpha_0^2 \\ R_\eta(1) = \beta_1 R_\eta(0) + \beta_2 R_\eta(1) \\ R_\eta(2) = \beta_1 R_\eta(1) + \beta_2 R_\eta(0) \end{cases} \quad (13)$$

Решая систему (13), получаем $\alpha_0 = 5.90960$, $\beta_1 = -0.23837$, $\beta_2 = 0.34939$.

Для модели АРСС(3, 0) получаем:

$$\begin{cases} R_\eta(0) = \beta_1 R_\eta(1) + \beta_2 R_\eta(2) + \beta_3 R_\eta(3) + \alpha_0^2 \\ R_\eta(1) = \beta_1 R_\eta(0) + \beta_2 R_\eta(1) + \beta_3 R_\eta(2) \\ R_\eta(2) = \beta_1 R_\eta(1) + \beta_2 R_\eta(0) + \beta_3 R_\eta(1) \\ R_\eta(3) = \beta_1 R_\eta(2) + \beta_2 R_\eta(1) + \beta_3 R_\eta(0) \end{cases} \quad (14)$$

Решая систему (14), получаем $\alpha_0 = 5.64763$, $\beta_1 = -0.13549$, $\beta_2 = 0.27920$, $\beta_3 = -0.29443$.

Проверим полученные модели на устойчивость. Для этого выясним лежат ли все корни z характеристического уравнения (15) внутри единичной окружности $|z| < 1$ на комплексной плоскости:

$$\sum_{k=1}^M \beta_k z^{-k} = 1. \quad (15)$$

Рассмотрим каждую модель:

- Модель $APCC(0, 0)$ устойчива всегда.
- Модель $APCC(1, 0)$ устойчива тогда и только тогда, когда $|\beta_1| < 1$:

$$|-0.36637| < 1 \Rightarrow \text{Модель устойчива.}$$

- Модель $APCC(2, 0)$ устойчива тогда и только тогда, когда $|\beta_2| < 1$, $|\beta_1| < 1 - \beta_2$:

$$\left. \begin{array}{l} |0.34939| < 1 \\ |-0.23837| < 0.65061 \end{array} \right\} \Rightarrow \text{Модель устойчива.}$$

- Модель $APCC(3, 0)$ устойчива тогда и только тогда, когда $|\beta_3| < 1$, $|\beta_1 + \beta_3| < 1 - \beta_2$, $|\beta_2 + \beta_1\beta_3| < |1 - \beta_3^2|$:

$$\left. \begin{array}{l} |-0.29443| < 1 \\ |-0.42993| < 0.72079 \\ |0.31910| < 0.91331 \end{array} \right\} \Rightarrow \text{Модель устойчива.}$$

4.2 Расчёт теоретической НКФ выходной последовательности

Для каждой модели рассчитаем теоретическую НКФ выходной последовательности. Для этого воспользуемся следующей формулой:

$$r_\eta^T(M+k) = \sum_{j=1}^M b_j r_\eta^T(M+k-j), k \geq 1, \quad (16)$$

учитывая что первые $M+1$ значений теоретической НКФ нам уже известны, т.к. они совпадают с первыми значениями выборочной НКФ для исходных данных.

Полученные значения, для $k = 0, 1, 2, \dots, 10$, представлены в таблице 2.

Таблица 2 – Значения теоретических НКФ для моделей $AP(M)$

$r_{\eta}^T(k)$	AP(0)	AP(1)	AP(2)	AP(3)
$r_{\eta}^T(0)$	1.00000	1.00000	1.00000	1.00000
$r_{\eta}^T(1)$	0.00000	-0.36638	-0.36638	-0.36638
$r_{\eta}^T(2)$	0.00000	0.13423	0.43672	0.43672
$r_{\eta}^T(3)$	0.00000	-0.04918	-0.23211	-0.4559
$r_{\eta}^T(4)$	0.00000	0.01802	0.20791	0.29158
$r_{\eta}^T(5)$	0.00000	-0.0066	-0.13066	-0.29538
$r_{\eta}^T(6)$	0.00000	0.00242	0.10379	0.25567
$r_{\eta}^T(7)$	0.00000	-0.00089	-0.07039	-0.20297
$r_{\eta}^T(8)$	0.00000	0.00032	0.05304	0.18586
$r_{\eta}^T(9)$	0.00000	-0.00012	-0.03724	-0.15713
$r_{\eta}^T(10)$	0.00000	0.00004	0.02741	0.13294

4.3 Выбор наилучшей модели случайного процесса в классе моделей AP

Выбор лучшей модели проводится на основе анализа нормированных корреляционных функций с использованием критерия среднего квадратичного отклонения по первым десяти отсчётам:

$$\epsilon^2 = \sum_{k=1}^{10} (r^T(k) - \hat{r}(k))^2, \quad (17)$$

где $\hat{r}(k)$ – выборочная НКФ исходного процесса, r^T – теоретическая НКФ. Результаты представлены в таблице 3. По результатам делаем следующий вывод: среди данных моделей, процесс лучше всего описывает модель AP(3).

Таблица 3 – Результаты построения моделей $AP(M)$

Порядок модели	Параметры модели				Погрешность модели
M	β_1	β_2	β_3	α_0	ϵ^2
0	-	-	-	6.77840	0.79153
1	-0.36637	-	-	6.30708	0.50184
2	-0.23837	0.34939	-	5.90960	0.11772
3	-0.13549	0.27920	-0.29443	5.64763	0.01756

5 ПОСТРОЕНИЕ И ИССЛЕДОВАНИЕ МОДЕЛЕЙ СКОЛЬЗЯЩЕГО СРЕДНЕГО

5.1 Модель АРСС(0, N)

Модель скользящего среднего порядка N в общем виде может быть записана как:

$$\eta_n = \alpha_0 \xi_n + \alpha_1 \xi_{n-1} + \dots + \alpha_N \xi_{n-N}, \quad (18)$$

где ξ – входная некоррелированная случайная последовательность с нулевым математическим ожиданием и единичной дисперсией ($\mathbb{M}(\xi_n) = 0$, $\mathbb{M}(\xi_n^2) = 1$, $\mathbb{M}(\xi_k \xi_n) = \delta_{kn}$), η – выходная случайная, стационарная в широком смысле последовательность с корреляционной функцией $R_\eta(m)$ и нулевым математическим ожиданием, $\{N, \alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N\}$ – параметры модели скользящего среднего.

Построим модели скользящего среднего для $N = 0, 1, 2, 3$ на основе решения системы нелинейных уравнений. Решение будем получать исходя из известных значений выборочной корреляционной функции $R_\eta(m)$ и следующей системы нелинейных уравнений:

$$R_\eta(m) = \begin{cases} \sum_{n=1}^{N-m} \alpha_n \alpha_{n+m} & \text{если } m \leq N \\ 0 & \text{иначе} \end{cases} \quad (19)$$

Выпишем системы для СС(0), СС(1), СС(2) и СС(3). Решим эти системы используя метод простых итераций из приложения Б.

- Для СС(0) получаем:

$$R_\eta(0) = \alpha_0^2 \rightarrow \alpha_0^2 = 45.94676 \rightarrow \alpha_0 = 6.77840.$$

- Для СС(1):

Условие существования решения:

$$R_\eta(0) - \alpha_0^2 \geq 0$$

Система:

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 \\ R_\eta(1) = \alpha_0 \alpha_1 \end{cases} \rightarrow \alpha_0 = -2.70925, \alpha_1 = 6.21343.$$

- Для СС(2):

Условие существования решения:

$$R_\eta(0) < \alpha_1^2 + \alpha_2^2$$

Система:

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 \\ R_\eta(1) = \alpha_0\alpha_1 + \alpha_1\alpha_2 \rightarrow \alpha_0 = 5.317996, \alpha_1 = -1.85165, \alpha_2 = 3.77320. \\ R_\eta(2) = \alpha_0\alpha_2 \end{cases}$$

- Для СС(3):

Условие существования решения:

$$R_\eta(0) < \alpha_1^2 + \alpha_2^2 + \alpha_3^2$$

Система:

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \\ R_\eta(1) = \alpha_0\alpha_1 + \alpha_1\alpha_2 + \alpha_2\alpha_3 \\ R_\eta(2) = \alpha_0\alpha_2 + \alpha_1\alpha_3 \\ R_\eta(3) = \alpha_0\alpha_3 \end{cases} \rightarrow \text{Условие не выполняется, система не имеет решения.}$$

При решении систем с помощью программы из приложения А было установлено: модели СС(3) в данном случае не существует.

5.2 Расчёт теоретической НКФ выходной последовательности

Для каждой модели рассчитаем теоретическую НКФ выходной последовательности. Для этого воспользуемся следующим фактом: первые $N+1$ значений теоретической НКФ совпадают со значениями выборочной НКФ исходного процесса, а остальные равны нулю (это следует из некоррелированности входного процесса ξ).

Значения, для $k = 0, 1, 2, \dots, 10$, представлены в таблице 4.

Таблица 4 – Значения теоретических НКФ для моделей $CC(N)$

$r_{\eta}^T(k)$	CC(0)	CC(1)	CC(2)
$r_{\eta}^T(0)$	1.00000	1.00000	1.00000
$r_{\eta}^T(1)$	0.00000	-0.366375	-0.366375
$r_{\eta}^T(2)$	0.00000	0.00000	0.43672
$r_{\eta}^T(3)$	0.00000	0.00000	0.00000
$r_{\eta}^T(4)$	0.00000	0.00000	0.00000
$r_{\eta}^T(5)$	0.00000	0.00000	0.00000
$r_{\eta}^T(6)$	0.00000	0.00000	0.00000
$r_{\eta}^T(7)$	0.00000	0.00000	0.00000
$r_{\eta}^T(8)$	0.00000	0.00000	0.00000
$r_{\eta}^T(9)$	0.00000	0.00000	0.00000
$r_{\eta}^T(10)$	0.00000	0.00000	0.00000

5.3 Выбор наилучшей модели случайного процесса в классе моделей CC

Выбор лучшей модели проводится на основе анализа нормированных корреляционных функций с использованием критерия среднего квадратичного отклонения по первым десяти отсчётам (17).

Результаты представлены в таблице 5. По результатам делаем следующий вывод: среди данных моделей, процесс лучше всего описывает модель $CC(2)$.

Таблица 5 – Результаты построения моделей $CC(N)$

Порядок модели	Параметры модели				Погрешность модели
M	α_0	α_1	α_2	α_3	ϵ^2
0	6.77840	-	-	-	0.79153
1	-2.70925	6.21343	-	-	0.65730
2	5.317996	-1.85165	3.77320	-	0.46657
3	Модель не существует				

6 ПОСТРОЕНИЕ И ИССЛЕДОВАНИЕ СМЕШАННЫХ МОДЕЛЕЙ АВТОРЕГРЕССИИ - СКОЛЬЗЯЩЕГО СРЕДНЕГО

6.1 Модель АРСС(M, N)

Модель АРСС(M, N) в общем виде может быть записана как:

$$\eta_n = \beta_1 \eta_{n-1} + \beta_2 \eta_{n-2} + \cdots + \beta_M \eta_{n-M} + \alpha_0 \xi_n + \alpha_1 \xi_{n-1} + \cdots + \alpha_N \xi_{n-N}, \quad (20)$$

где ξ – входная некоррелированная случайная последовательность с нулевым математическим ожиданием и единичной дисперсией ($\mathbb{M}(\xi_n) = 0, \mathbb{M}(\xi_n^2) = 1, \mathbb{M}(\xi_k \xi_n) = \delta_{kn}$), η – выходная случайная, стационарная в широком смысле последовательность с корреляционной функцией $R_\eta(m)$ и нулевым математическим ожиданием, $\{N, \alpha_0, \alpha_1, \alpha_2, \dots, \alpha_N, M, \beta_1, \beta_2, \dots, \beta_M\}$ – параметры модели авторегрессии - скользящего среднего.

Для построения модели АРСС существует несколько методов, рассмотрим один из них.

6.2 Построение модели АРСС

Уравнение (20) можно переписать в виде:

$$\zeta_n = \alpha_0 \xi_n + \alpha_1 \xi_{n-1} + \cdots + \alpha_N \xi_{n-N}, \quad (21)$$

где

$$\zeta_n = \eta_n - \beta_1 \eta_{n-1} - \beta_2 \eta_{n-2} - \cdots - \beta_M \eta_{n-M}. \quad (22)$$

Уравнение (21) – уравнение модели СС(N) для промежуточной последовательности ζ , которая может быть получена из исходной последовательности η по уравнению (22).

Шаг 1. Из системы M линейных уравнений (23) находятся оценки M неизвестных коэффициентов $\beta_1, \beta_2, \dots, \beta_M$.

$$\left\{ \begin{array}{l} R_\eta(N+1) = \beta_1 R_\eta(N) + \beta_2 R_\eta(N-1) + \cdots + \beta_M R_\eta(N-M+1) \\ R_\eta(N+2) = \beta_1 R_\eta(N+1) + \beta_2 R_\eta(N) + \cdots + \beta_M R_\eta(N-M+2) \\ \dots \\ R_\eta(N+M) = \beta_1 R_\eta(N+M-1) + \beta_2 R_\eta(N+M-2) + \cdots + \beta_M R_\eta(N) \end{array} \right. \quad (23)$$

Шаг 2. Из исходной последовательности η строится промежуточная последовательность ζ с использованием уравнения (22).

Шаг 3. Находится выборочная корреляционная функция $R_\zeta(m)$ промежуточной последовательности ζ .

Шаг 4. Из системы $N + 1$ нелинейных уравнений (24) находится оценка $N + 1$ неизвестных коэффициентов $\alpha_0, \alpha_1, \dots, \alpha_N$.

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 + \dots + \alpha_N^2 \\ R_\eta(1) = \alpha_0\alpha_1 + \alpha_1\alpha_2 + \dots + \alpha_{N-1}\alpha_N \\ \dots \\ R_\eta(N) = \alpha_0\alpha_N \end{cases} \quad (24)$$

Шаг 5. Проверить модель на устойчивость так же, как и в разделе (4.1). Выясним лежат ли все корни z характеристического уравнения (15) внутри единичной окружности $|z| < 1$ на комплексной плоскости.

Разберём подробно построение модели АРСС(1, 3):

Шаг 1. Из линейного уравнения (25) находим оценку неизвестного коэффициента β_1 .

$$R_\eta(4) = \beta_1 R_\eta(3) + \beta_2 R_\eta(2) + \beta_3 R_\eta(1) \quad (25)$$

Решая данное уравнение, получаем $\beta_1 = -0.53593$.

Шаг 2. Из исходной последовательности η строим промежуточную последовательность ζ с использованием уравнения (26).

$$\zeta_n = \eta_n - \beta_1 \eta_{n-1} - \beta_2 \eta_{n-2} - \beta_3 \eta_{n-3}. \quad (26)$$

Шаг 3. Вычислим выборочную корреляционную функцию $R_\zeta(m)$ промежуточной последовательности ζ , используя формулу (7). Результаты представлены в таблице 6.

Таблица 6 – Выборочная корреляционная функция промежуточной последовательности ζ

m	$R_\zeta(m)$
0	41.12153
1	13.72420
2	5.58428
3	-10.18949
4	-4.37808
5	-7.92605
6	-2.06989
7	-4.19016
8	-0.51507
9	-2.24310
10	-0.16676

Шаг 4. Из системы четырех нелинейных уравнений (27) находим оценку неизвестных коэффициентов $\alpha_0, \alpha_1, \alpha_2, \alpha_3$.

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \\ R_\eta(1) = \alpha_0\alpha_1 + \alpha_1\alpha_2 + \alpha_2\alpha_3 \\ R_\eta(2) = \alpha_0\alpha_1 + \alpha_1\alpha_3 \\ R_\eta(3) = \alpha_0\alpha_3 \end{cases} \quad (27)$$

Используем метод простых итераций из приложения Б.

Начальное приближение: $\alpha_0 = \sqrt{R_\eta(0)}, \alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$.

Схема итераций:

1. $\alpha_3 \leftarrow \frac{R_\eta(3)}{\alpha_0}, \alpha_2 \leftarrow \frac{R_\eta(2) - \alpha_1\alpha_3}{\alpha_0}, \alpha_1 \leftarrow \frac{R_\eta(1) - \alpha_1\alpha_2 - \alpha_2\alpha_3}{\alpha_0}$
2. Если $R_\eta(0) < \alpha_1^2 + \alpha_2^2 + \alpha_3^2$, то считаем, что решение не существует, и вычисление завершается с выводом соответствующей ошибки, иначе $\alpha_0 \leftarrow \sqrt{R_\eta(0) - \alpha_1^2 - \alpha_2^2 - \alpha_3^2}$.
3. Проверка на завершение итераций – уменьшение скорости сходимости итераций до заданного значения ϵ .
4. Переход к пункту 1.

В результате итерационного процесса было найдено одно из возможных решений системы $\alpha_0 = 5.25709, \alpha_1 = 2.42818, \alpha_2 = 1.95748, \alpha_3 = -1.93823$, что является достаточным.

Шаг 5. Проверим модель на устойчивость так же, как и в разделе (4.1).

Выясним лежат ли все корни z характеристического уравнения (15) внутри единичной окружности $|z| < 1$ на комплексной плоскости.

Модель АРСС(1, 3) устойчива тогда и только тогда, когда $|\beta_1| < 1$:

$$|-0.53593| < 1 \Rightarrow \text{Модель устойчива.}$$

После проверки на устойчивость можно записать уравнение имитационного моделирования случайного процесса по этой модели. Учтем, что для соответствия математического ожидания моделируемого процесса исходному необходимо прибавить к результату значение математического ожидания исходного процесса $\bar{X} = -25.07220$. Итоговый вид уравнения имитационного моделирования для модели АРСС(1,3):

$$\eta_n = -0.53593\eta_{n-1} + 5.25709\xi_n + 2.42818\xi_{n-1} + 1.95748\xi_{n-2} - 1.93823\xi_{n-3} - 25.07220 \quad (28)$$

Используя данный метод построим модели $APCC(M, N)$, для $M = 1, 2, 3$ и $N = 1, 2, 3$. Значения коэффициентов, полученных с помощью программы из приложения А, представлены в таблице 6.

6.3 Расчёт теоретической НКФ выходной последовательности

Для каждой модели рассчитаем теоретическую НКФ выходной последовательности. Для этого воспользуемся следующей формулой:

$$r_{\eta}^T(N + M + k) = \beta_1 r_{\eta}^T(N + M + k - 1) + \beta_2 r_{\eta}^T(N + M + k - 2) + \dots + \beta_M r_{\eta}^T(N + k), k \geq 1, \quad (29)$$

учитывая что первые $N + M + 1$ значений теоретической НКФ нам уже известны, т.к. они совпадают с первыми значениями выборочной НКФ для исходных данных.

6.4 Выбор наилучшей модели случайного процесса в классе моделей APCC

Выбор лучшей модели проводится на основе анализа нормированных корреляционных функций с использованием критерия среднего квадратичного отклонения по первым десяти отсчётам (17).

Результаты представлены в таблице 6. По результатам делаем следующий вывод: среди данных моделей, процесс лучше всего описывает модель $APCC(3, 3)$.

Таблица 7 – Результаты построения моделей $APCC(M, N)$

Порядок модели		Параметры модели							Погрешность модели
M	N	β_1	β_2	β_3	α_0	α_1	α_2	α_3	ϵ^2
1	1	Модель не существует							
1	2	Модель не существует							
1	3	-0.53593	-	-	5.25708	2.42818	1.95748	-1.93824	0.00896
2	1	Модель не существует							
2	2	Модель не существует							
2	3	0.23067	0.80028	-	Модель неустойчива				
3	1	0.07562	0.32953	-0.36819	-1.30849	5.63670	-	-	0.00004
3	2	0.14318	0.39562	-0.37349	5.63176	-1.70196	-0.35664	-	0.00005
3	3	0.18594	0.59337	-0.19097	5.56886	-2.03828	-1.50236	-1.05953	$4 \cdot 10^{-8}$

7 УРАВНЕНИЯ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ

Напишем уравнения для моделирования процессов на основе лучших моделей в каждом классе, а именно $AP(3)$, $CC(2)$ и $APCC(3,3)$. Для соответствия математического ожидания моделируемого процесса исходному, прибавим к правой части каждого уравнения математическое ожидание исходного процесса $\bar{X} = -25.07220$.

- Для модели $AP(3)$ имеем:

$$\eta_n = -0.13549\eta_{n-1} + 0.27920\eta_{n-2} - 0.29443\eta_{n-3} + 5.64763\xi_n - 25.07220 \quad (30)$$

- Для модели $CC(2)$ имеем:

$$\eta_n = 5.31799\xi_n - 1.85165\xi_{n-1} + 3.77320\xi_{n-2} - 25.07220 \quad (31)$$

- Для модели $APCC(3,3)$ имеем:

$$\begin{aligned} \eta_n = & 0.18593\eta_{n-1} + 0.59336\eta_{n-2} - 0.19097\eta_{n-3} + \\ & + 5.56886\xi_n - 2.03828\xi_{n-1} - 1.50236\xi_{n-2} - 1.05952\xi_{n-3} - 25.07220 \end{aligned} \quad (32)$$

8 СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПОСТРОЕННЫХ МОДЕЛЕЙ

В таблице 8 представлены погрешности моделей $APCC(M, N)$, для $M = 0, 1, 2, 3$ и $N = 0, 1, 2, 3$.

Таблица 8 – Погрешности моделей $APCC(M, N)$

$M \backslash N$	0	1	2	3
0	0.791530	0.65730	0.46657	-
1	0.50184	-	-	0.00896
2	0.11772	-	-	-
3	0.01756	0.00004	0.00005	$4 \cdot 10^{-8}$

Из таблицы видно, что исходный процесс лучше всего описывается моделями $AP(3)$, $CC(2)$ и $APCC(3,3)$. Смоделируем эти процессы, используя формулы (30), (31), (32) соответственно. Принимая во внимание то, что сгенерированная случайная последовательность приобретает свойство стационарности по истечении интервала времени большего, чем интервал корреляции сгенерируем 6000 отсчётов и отбросим первые 1000 элементов. Код генерации представлен в приложении А.

Важно заметить то, что у смоделированного случайного процесса нулевое математическое ожидание. Чтобы исправить это, прибавим ко всем значениям математическое ожидание исходного процесса.

На рисунках 3, 4 и 5 изображены: выборочные НКФ исходного процесса, теоретические НКФ моделей и выборочные НКФ смоделированных процессов.

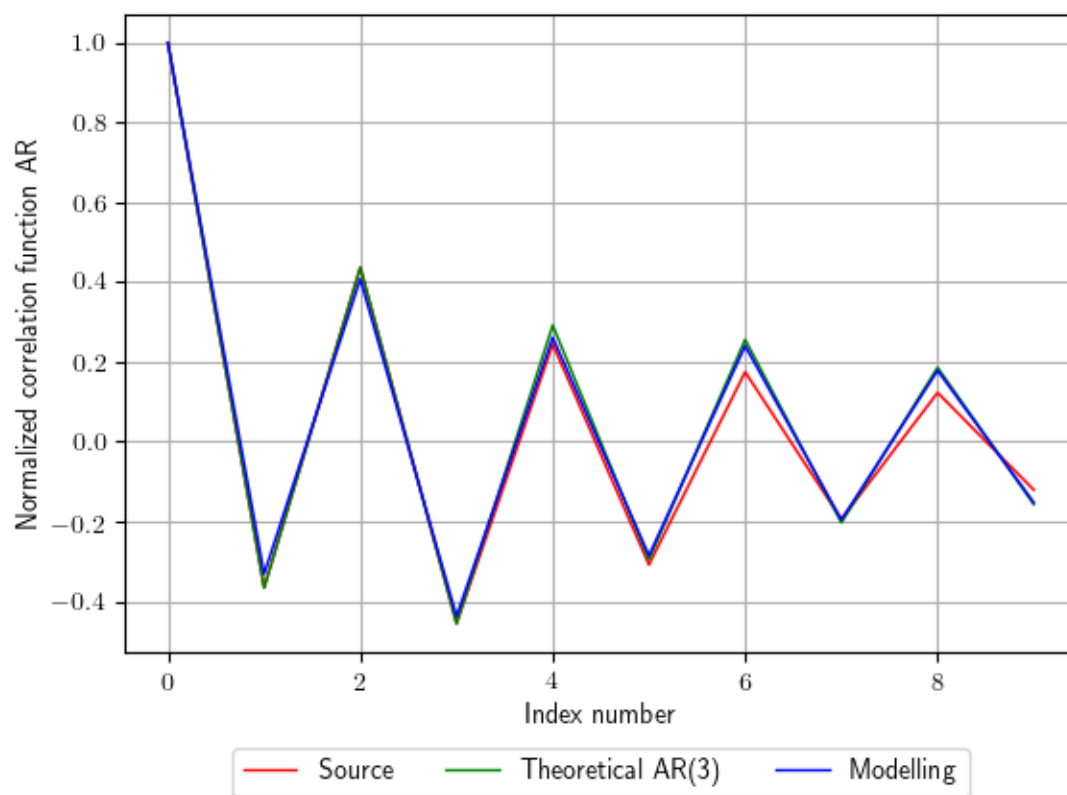


Рисунок 3 – Графическое сравнение НКФ для модели $AR(3)$

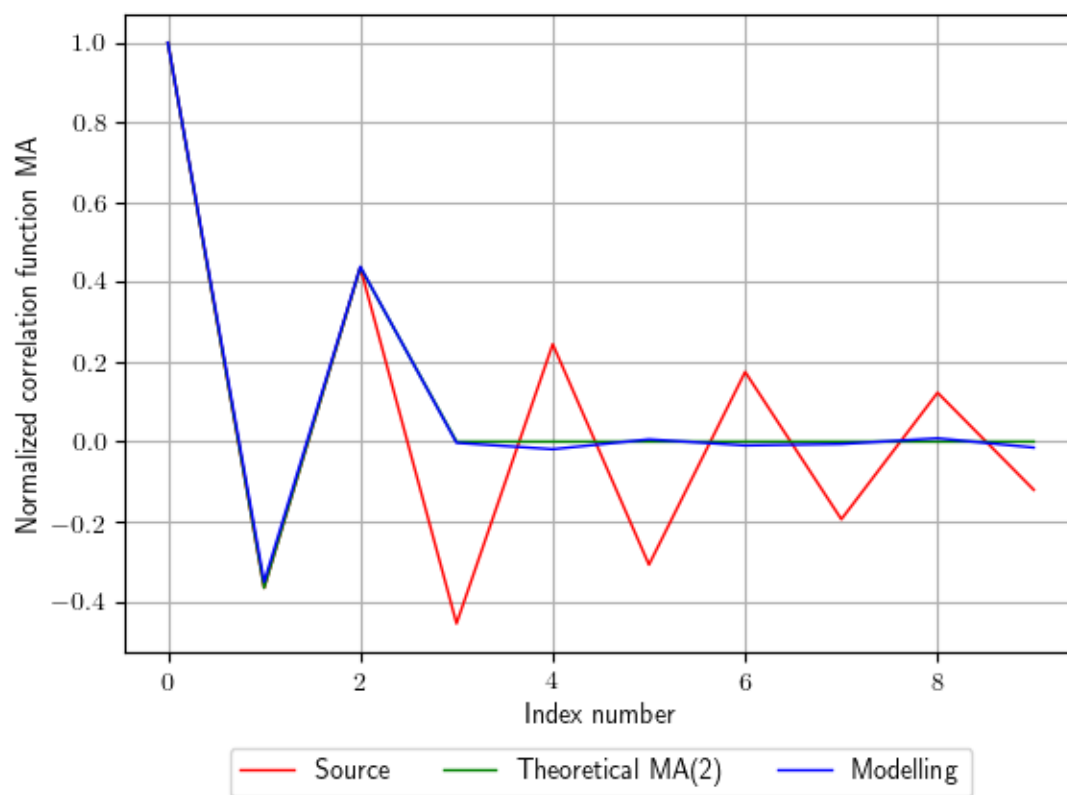


Рисунок 4 – Графическое сравнение НКФ для модели $CC(2)$

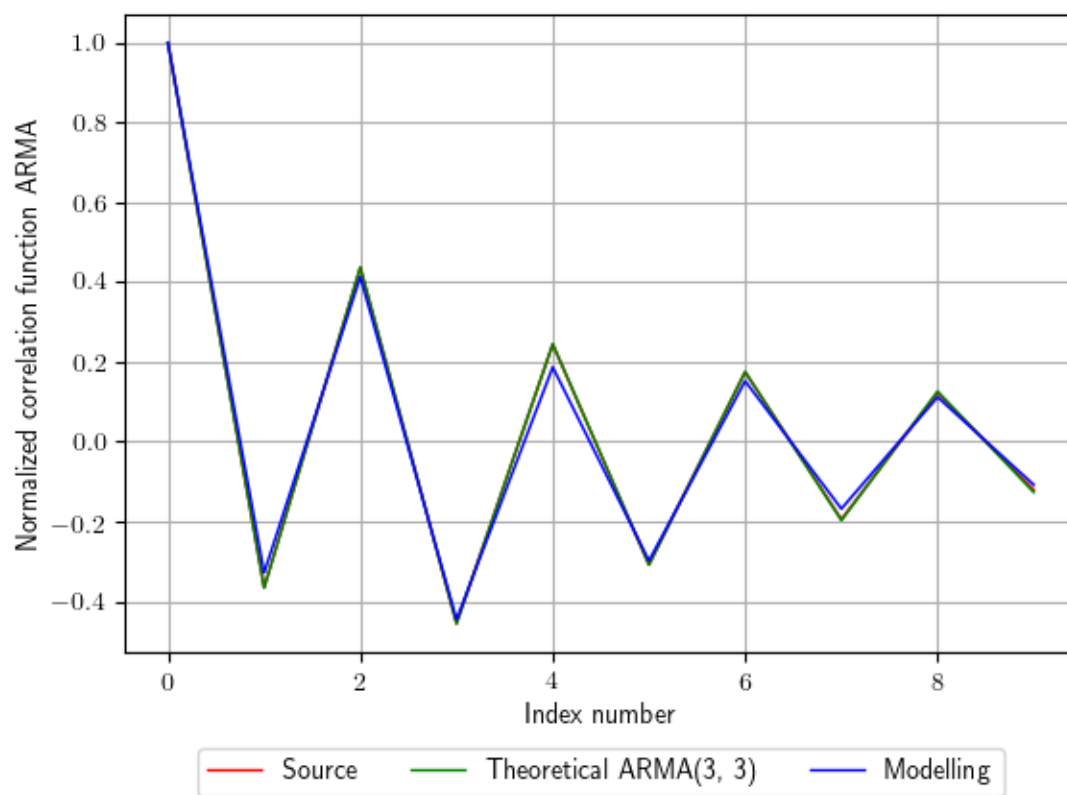


Рисунок 5 – Графическое сравнение НКФ для модели $APCC(3, 3)$

9 ИТОГОВАЯ ТАБЛИЦА СРАВНЕНИЯ МОДЕЛЕЙ АРСС

В таблице 9 представлены итоговые результаты.

Таблица 9 – Итоговые результаты статистического анализа СП и моделирования

Параметры процесса	Исходный процесс	АР(3) теория	АР(3) модель	СС(2) теория	СС(2) модель	АРСС(3,3) теория	АРСС(3,3) модель
Среднее	-25.07220	-25.07220	-25.11893	-25.07220	100.18219	-25.07220	-25.11893
Дисперсия	45.94676	45.94676	42.97016	45.94676	251.76423	45.94676	42.97016
СКО	6.77840	6.77840	6.55516	6.77840	15.86708	6.77840	6.55516
Нормированная корреляционная функция							
r(0)	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
r(1)	-0.36638	-0.36638	-0.33168	-0.36637	-0.35185	-0.36637	-0.32881
r(2)	0.43672	0.43672	0.40778	0.43672	0.43846	0.43672	0.41313
r(3)	-0.4559	-0.4559	-0.43777	0.00000	-0.00294	-0.45590	-0.4458
r(4)	0.24433	0.29158	0.261	0.00000	-0.01892	0.24433	0.18681
r(5)	-0.30849	-0.29538	-0.2856	0.00000	0.00572	-0.30848	-0.29844
r(6)	0.17469	0.25567	0.2401	0.00000	-0.0095	0.17468	0.15196
r(7)	-0.19466	-0.20297	-0.19613	0.00000	-0.00593	-0.19723	-0.16836
r(8)	0.12273	0.18586	0.17833	0.00000	0.00845	0.12589	0.11161
r(9)	-0.12081	-0.15713	-0.15285	0.00000	-0.01484	-0.12698	-0.10818
r(10)	0.07614	0.13294	0.12863	0.00000	0.00479	0.08875	0.07494
Погрешность модели		0.01756	0.01433	0.46657	0.47234	$4 \cdot 10^{-8}$	0.00697

Для нахождения погрешности использовался критерий среднего квадратичного отклонения по первым десяти отсчётам:

$$\epsilon^2 = \sum_{k=1}^{10} (r^M(k) - \hat{r}(k))^2, \quad (33)$$

где $\hat{r}(k)$ – выборочная НКФ исходного процесса, r^M – выборочная НКФ смоделированного процесса.

На рисунке 6 представлен смоделированный процесс АРСС(3, 3).

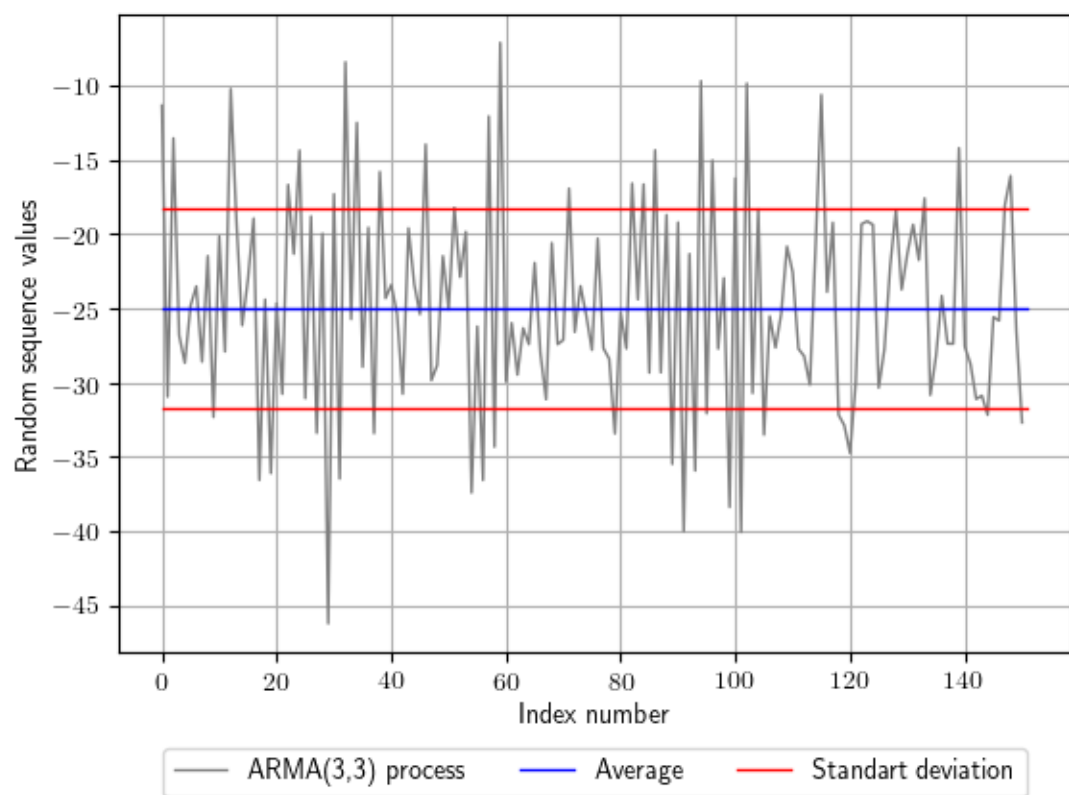


Рисунок 6 – Фрагмент реализации $APCC(3, 3)$

ЗАКЛЮЧЕНИЕ

Было проведено исследование и построение моделей АРСС для некоторого исходного неизвестного эргодического процесса. В ходе работы была проанализирована выборка из отсчётов исходного процесса, построены модели АР с помощью уравнений Юла-Уокера и модели СС с помощью систем нелинейных уравнений. Лучшими моделями в каждом классе оказались АР(3) и СС(2) с погрешностями 0.01756 и 0.46657 соответственно. Были построены смешанные модели АРСС до третьего порядка включительно, лучшей моделью среди них оказалась АРСС(3,3) с погрешностью равной $4 \cdot 10^{-8}$.

Для каждой из трёх лучших моделей были смоделированы случайные процессы, рассчитаны выборочные моментные функции, а также был произведен сравнительный анализ построенных моделей. Наилучшей оказалась модель АРСС(3,3) с погрешностью равной 0.00697.

СПИСОК ЛИТЕРАТУРЫ

1. **Тараскин, А.Ф.** Статистический анализ временных рядов авторегрессии и скользящего среднего: учебное пособие [Текст] // Самара: СГАУ, 1998. – 56с.
2. **Тараскин, А.Ф.** Статистическое моделирование и метод Монте–Карло: учебное пособие [Текст] // Самара: СГАУ, 1997. – 62с.
3. **Храмов, А.Г.** Анализ и моделирование процессов АРСС: интернет-ресурс к курсовой работе [Электронный ресурс] // Самара: СГАУ, 2009.

ПРИЛОЖЕНИЕ А КОД ПРОГРАММЫ

correlation.py:

```
import numpy as np
from functools import lru_cache

def corr_func(data):

    data = np.array(data)

    data = data - data.mean()

    @lru_cache(maxsize=None)
    def func(n):
        if n < 0:
            return func(-n)
        if n == 0:
            return data.std() ** 2

        x = data[:n] * data[n:]

        return np.sum(x) / (np.size(x) - 1)

    return func

def normalized_corr_func(data):

    data = np.array(data)

    corr = corr_func(data)

    dispersion = corr(0)

    @lru_cache(maxsize=None)
    def func(n):
        return corr(n) / dispersion

    return func

def corr_error(corr1, corr2, n=11):
```

```
return sum((corr1(n) - corr2(n))**2 for n in range(1, n))
```


autoregression.py:

```
import numpy as np
from correlation import *

def solve(a, b):

    a = np.matrix(a)
    b = np.array(b)

    b = np.reshape(b, (-1, 1))

    x = np.linalg.solve(a, b)
    x = np.array(x)

    return x.flatten()

def make_walker_matrix(corr, n):

    R = np.zeros((n, n))

    for i in range(n):
        for j in range(n):
            R[i, j] = corr(abs(i - j))

    return np.matrix(R)

def make_walker_vector(corr, n):

    R = np.zeros((n, 1))

    for i in range(n):
        R[i, 0] = corr(i + 1)

    return R

def calc_model_ar(n, data):

    data = np.array(data)
```

```

corr = corr_func(data)

if n == 0:
    return np.sqrt([corr(0)])

A = make_walker_matrix(corr, n)
b = make_walker_vector(corr, n)

betas = solve(A, b)

tmp = corr(0) - betas.dot(b)

if tmp < 0:
    raise ArithmeticError('No solution for model ar({})'.format(n))

alpha = np.sqrt(corr(0) - betas.dot(b))

if n == 1:
    if abs(betas[0]) >= 1:
        raise ArithmeticError('Model ar(1) is unstable')

if n == 2:
    if abs(betas[1]) >= 1 \
        or abs(betas[0]) >= 1 - betas[1]:
        raise ArithmeticError('Model ar(2) is unstable')

if n == 3:
    if abs(betas[2]) >= 1 \
        or abs(betas[0] + betas[2]) >= 1 - betas[1] \
        or abs(betas[1] + betas[0] * betas[2]) >= abs(1 - betas[2]**2):
        raise ArithmeticError('Model ar(3) is unstable')

return np.append(alpha, betas)

def theory_ncf_ar(n, data):
    coeffs = calc_model_ar(n, data)
    R = corr_func(data)

    @lru_cache(maxsize=None)
    def func(m):
        if abs(m) <= n:
            return R(m)

```

```

    if n == 0:
        return 0

    rs = np.array([func(m - x) for x in range(1, n + 1)])

    return rs.dot(coeffs[1:])

def res(m):
    return func(m) / R(0)

return res

def model_ar(n, data):
    coeffs = calc_model_ar(n, data)

    alpha = coeffs[0]

    if n > 0:
        betas = coeffs[1:]

def get_model(xs, x0):
    if len(xs) != n:
        raise ValueError('Length of 'xs' must be equal to {}'.format(n))

    w = np.random.normal()

    if n > 0:
        xs = np.array(xs)
        x = x0

    while True:
        yield x

        if n > 0:
            x = xs.dot(betas) + w * alpha

            xs[1:] = xs[:-1]
            xs[0] = x
        else:
            x = w * alpha

```

```
w = np.random.normal()

return get_model
```

moving_average.py:

```
import numpy as np
from correlation import *

def calc_model_ma0(data):
    """
         $R(0) = a_0^{**2}$ 

        Exact solution:
         $a_0 = \text{sqrt}(R(0))$ 

        R - correlation function
        a_i - model coef.
    """
    data = np.array(data)

    R = corr_func(data)

    R0 = R(0)

    if R0 < 0:
        raise ArithmeticError(
            'Dispersion should be non-negative, maybe bug in corr_func')

    return np.sqrt([R0])

def theory_ncf_ma0(data):
    r = normalized_corr_func(data)

    def func(n):
        if n == 0:
            return r(0)
        return 0

    return func

def model_ma0(data):
    """
         $x(t) = w(t) * a_0$ 
    """
```

```

        w - white noise
        a_i - coefs. from calc_model_ma0
        """
a = calc_model_ma0(data)

def get_model(x0):
    w = np.random.normal()
    x = x0

    while True:
        yield x

        x = w * a
        w = np.random.normal()

    return get_model

def calc_model_ma1(data, eps=1e-5, maxiter=100):
    """
         $R(0) = a_0^{**2} + a_1^{**2}$ 
         $R(1) = a_0 * a_1$ 

        Iterative solution:
         $a_0 = R(1) / a_1$ 
         $a_1 = \text{sqrt}(R(0) - a_0^{**2})$ 

        R - correlation function
        a_i - model coefs.
    """
    data = np.array(data)

    R = corr_func(data)

    R0 = R(0)
    R1 = R(1)

    a0 = 0
    a1 = R0

    for _ in range(maxiter):
        if abs(R0 - a0**2 - a1**2) < eps \
            and abs(R1 - a0 * a1) < eps:

```

```

        return np.array([a0, a1])

a0 = R1 / a1

if (R0 - a0**2) < 0:
    raise ArithmeticError('Solution doesn\'t exist [ma(1)]')

a1 = np.sqrt(R0 - a0**2)

if abs(R0 - a0**2 - a1**2) > 10 * eps \
    or abs(R1 - a0 * a1) > 10 * eps:
    raise ArithmeticError('Didn\'t coverage to solution')

return np.array([a0, a1])

def theory_ncf_ma1(data):
    r = normalized_corr_func(data)

    def func(n):
        if abs(n) <= 1:
            return r(n)
        return 0

    return func

def model_ma1(data, eps=1e-5, maxiter=100):
    """
         $x(t) = a_0 * w(t) + a_1 * w(t - 1)$ 

        w - white noise
        ai - coefs. from calc_model_ma1
    """
    a = calc_model_ma1(data, eps, maxiter)

    def get_model(x0):
        w = np.random.normal(size=2)

        x = x0

        while True:
            yield x

```

```

        x = a.dot(w)

        w[1:] = w[:-1]
        w[0] = np.random.normal()

    return get_model

def calc_model_ma2(data, eps=1e-5, maxiter=100):
    """
         $R(0) = a_0^{**2} + a_1^{**2} + a_2^{**2}$ 
         $R(1) = a_0 * a_1 + a_1 * a_2$ 
         $R(2) = a_0 * a_2$ 

        Iterative solution:
         $a_2 = R(2) / a_0$ 
         $a_1 = (R(1) - a_1 * a_2) / a_0$ 
         $a_0 = \text{sqrt}(R(0) - a_1^{**2} - a_2^{**2})$  if  $R(0) \geq a_1^{**2} + a_2^{**2}$ 

        R - correlation function
        a_i - model coeffs.
    """
    data = np.array(data)

    R = corr_func(data)

    R0 = R(0)
    R1 = R(1)
    R2 = R(2)

    a0 = np.sqrt(R0)
    a1 = 0
    a2 = 0

    for _ in range(maxiter):
        if abs(R0 - a0**2 - a1**2 - a2**2) < eps \
            and abs(R1 - a0 * a1 - a1 * a2) < eps \
            and abs(R2 - a0 * a2) < eps:
            return np.array([a0, a1, a2])

        a2 = R2 / a0
        a1 = (R1 - a1 * a2) / a0

```



```

    if R0 < a1 ** 2 + a2 ** 2:
        raise ArithmeticError('Solution doesn\'t exist [ma(2)]')

    a0 = np.sqrt(R(0) - a1**2 - a2**2)

    if abs(R0 - a0**2 - a1**2 - a2**2) > 10 * eps \
        or abs(R1 - a0 * a1 - a1 * a2) > 10 * eps \
        or abs(R2 - a0 * a2) > 10 * eps:
        raise ArithmeticError('Didn\'t coverage to solution')

    return np.array([a0, a1, a2])

def theory_ncf_ma2(data):
    r = normalized_corr_func(data)

    def func(n):
        if abs(n) <= 2:
            return r(n)
        return 0

    return func

def model_ma2(data, eps=1e-5, maxiter=100):

    a = calc_model_ma2(data, eps, maxiter)

    def get_model(x0):
        x = x0
        w = np.random.normal(size=3)

        while True:
            yield x

            x = a.dot(w)

            w[1:] = w[:-1]
            w[0] = np.random.normal()

    return get_model

```

```

def calc_model_ma3(data, eps=1e-5, maxiter=100):

    data = np.array(data)

    R = corr_func(data)

    R0 = R(0)
    R1 = R(1)
    R2 = R(2)
    R3 = R(3)

    a0 = np.sqrt(R0)
    a1 = 0
    a2 = 0
    a3 = 0

    for _ in range(maxiter):
        if abs(R0 - a0**2 - a1**2 - a2**2 - a3**2) < eps \
            and abs(R1 - a0 * a1 - a1 * a2 - a2 * a3) < eps \
            and abs(R2 - a0 * a2 - a1 * a3) < eps \
            and abs(R3 - a0 * a3) < eps:
            return np.array([a0, a1, a2, a3])

        a3 = R3 / a0
        a2 = (R2 - a1 * a3) / a0
        a1 = (R1 - a1 * a2 - a2 * a3) / a0

        if R0 < a1**2 + a2**2 + a3**2:
            raise ArithmeticError('Solution doesn\'t exist [ma(3)]')

        a0 = np.sqrt(R0 - a1**2 - a2**2 - a3**2)

    if abs(R0 - a0**2 - a1**2 - a2**2 - a3**2) > 10 * eps \
        or abs(R1 - a0 * a1 - a1 * a2 - a2 * a3) > 10 * eps \
        or abs(R2 - a0 * a2 - a1 * a3) > 10 * eps \
        or abs(R3 - a0 * a3) > 10 * eps:
        raise ArithmeticError('Didn\'t coverage to solution')

    return np.array([a0, a1, a2, a3])

def theory_ncf_ma3(data):

```

```

r = normalized_corr_func(data)

def func(n):
    if abs(n) <= 3:
        return r(n)
    return 0

return func

def model_ma3(data, eps=1e-5, maxiter=100):

    a = calc_model_ma3(data, eps, maxiter)

    def get_model(x0):
        x = x0
        w = np.random.normal(size=4)

        while True:
            yield x

            x = a.dot(w)

            w[1:] = w[:-1]
            w[0] = np.random.normal()

    return get_model

def calc_model_ma(n, data, eps=1e-5, maxiter=100):

    if n == 0:
        return calc_model_ma0(data)
    if n == 1:
        return calc_model_ma1(data, eps, maxiter)
    if n == 2:
        return calc_model_ma2(data, eps, maxiter)
    if n == 3:
        return calc_model_ma3(data, eps, maxiter)

    raise ValueError('Calculating coeffs. for 'n > 3' not implemented')

```

```

def theory_ncf_ma(n, data):

    if n == 0:
        return theory_ncf_ma0(data)
    if n == 1:
        return theory_ncf_ma1(data)
    if n == 2:
        return theory_ncf_ma2(data)
    if n == 3:
        return theory_ncf_ma3(data)

    raise ValueError('NCF for 'n > 3' not implemented')


def model_ma(n, data, eps=1e-5, maxiter=100):

    if n == 0:
        return model_ma0(data)
    if n == 1:
        return model_ma1(data, eps, maxiter)
    if n == 2:
        return model_ma2(data, eps, maxiter)
    if n == 3:
        return model_ma3(data, eps, maxiter)

    raise ValueError('Models with 'n > 3' not implemented')

```

arma.py:

```
from correlation import *
from autoregression import *
from moving_average import *

def calc_model_arma(m, n, data, eps=1e-5, maxiter=100):

    data = np.array(data)

    matrix = np.zeros((m, m))
    vector = np.zeros((m, 1))

    R = corr_func(data)

    for i in range(m):
        vector[i] = R(n + i + 1)
        for j in range(m):
            matrix[i, j] = R(n + i - j)

    betas = solve(matrix, vector)

    def gen_seq():
        for i in range(m + 1, len(data)):
            xs = data[i - m - 1:i]
            y = xs[0] - xs[1:].dot(betas)

            yield y

    seq = gen_seq()
    seq_data = [x for x in seq]

    cf_seq = corr_func(seq_data)

    print('Start cf for temp. seq.')
    for i in range(11):
        print(i, cf_seq(i))
    print('End cf for temp. seq.')

    alphas = calc_model_ma(n, seq_data, eps, maxiter)

    if m == 1:
        if abs(betas[0]) >= 1:
```

```

        raise ArithmeticError('Solution unstable, betas: {}'.format(betas))
if m == 2:
    if abs(betas[1]) >= 1 \
        or abs(betas[0]) >= 1 - betas[1]:
        raise ArithmeticError('Solution unstable, betas: {}'.format(betas))
if m == 3:
    if abs(betas[2]) >= 1 \
        or abs(betas[0] + betas[2]) >= 1 - betas[1] \
        or abs(betas[1] + betas[0] * betas[2]) >= 1 - betas[2]**2:
        raise ArithmeticError('Solution unstable, betas: {}'.format(betas))

return betas, alphas

def theory_ncf_arma(m, n, data, eps=1e-5, maxiter=100):

    from functools import lru_cache

    coeffs = calc_model_arma(m, n, data, eps, maxiter)

    ar_coeffs, ma_coeffs = coeffs

    R = corr_func(data)

    @lru_cache(maxsize=None)
    def func(x):
        if x < 0:
            return func(-x)
        if x <= m + n:
            return R(x)

        rs = np.array([func(x - i) for i in range(1, m + 1)])

        return rs.dot(ar_coeffs)

    def res(x):
        return func(x) / R(0)

    return res

def model_arma(m, n, data, eps=1e-5, maxiter=100):

```

```

if m == 0:
    return model_ma(n, data, eps, maxiter)
if n == 0:
    return model_ar(m, data)

coeffs = calc_model_arma(m, n, data, eps, maxiter)

ar_coeffs, ma_coeffs = coeffs

def get_model(xs, x0):
    if len(xs) != m:
        raise ValueError('Length of 'xs' must be equal to {}'.format(m))

    xs = np.array(xs)
    w = np.random.normal(size=(n + 1))

    x = x0

    while True:
        yield x

        x = xs.dot(ar_coeffs) + ma_coeffs.dot(w)

        w[1:] = w[:-1]
        w[0] = np.random.normal()

        xs[1:] = xs[:-1]
        xs[0] = x

    return get_model

```

main.py:

```
from math import exp
import numpy as np
import matplotlib.pyplot as plt
from correlation import *
from moving_average import *
from autoregression import *
from arma import *

data = np.array(list(map(float, open('./tsp.txt'))))

if __name__ == '__main__':
    maxiter = 15000
    np.random.seed(0)

    mean = data.mean()
    std = data.std()

    print('Mean: {}'.format(mean))
    print('Dispersion.: {}'.format(std ** 2))
    print('STD: {}'.format(std))

    x = range(0, 151)
    y = data[0:151]

    fig = plt.figure()
    ax = plt.subplot(111)

    plt.rc('lines', linewidth=1)

    source, = ax.plot(x, y, 'gray')
    avg, = ax.plot([0, 151], [mean] * 2, 'red')
    mn, = ax.plot([0, 151], [mean + std] * 2, 'blue')
    mn, = ax.plot([0, 151], [mean - std] * 2, 'blue')

    plt.xlabel('Index number')
    plt.ylabel('Random sequence values')

    box = ax.get_position()
    ax.set_position([box.x0, box.y0 + box.height * 0.1,
                    box.width, box.height * 0.9])

    ax.legend([source, avg, mn], ['Source process', 'Average', 'Standard deviation'],
```



```

        loc='upper center', bbox_to_anchor=(0.5, -0.13), ncol=3, fancybox=True)

plt.grid(True)

plt.savefig('plot1.png')
plt.show()

cnst = exp(-1)
ncf = normalized_corr_func(data)
cf = corr_func(data)

limit = 300

for n in range(limit):
    ncf(n)

interval = limit

while abs(ncf(interval)) < exp(-1):
    interval -= 1

print('Correlation interval: {}'.format(interval))

print('Corr.f., ncf - first 10 values')
for n in range(11):
    print(n, ' & ', round(cf(n), 5), '& ', round(ncf(n), 5), ' \\\\' \\hline')

print('Corr.f, Norm.c.f in interval')
for n in range(interval + 1):
    print(round(cf(n), 5), round(ncf(n), 5))

fig = plt.figure()
ax = plt.subplot(111)

plt.rc('text', usetex=True)

x = range(interval + 20)
y = list(map(ncf, x))

source, = ax.plot(x, y, 'gray')
cnst_plt, = ax.plot([min(x), max(x)], [cnst] * 2, 'blue')
cnst_plt, = ax.plot([min(x), max(x)], [-cnst] * 2, 'blue')
interval, = ax.plot([interval, interval], [

```

```

        min(y), max(y)], 'red', linestyle='--')

plt.xlabel('Index number')
plt.ylabel('Normalized correlation function')

box = ax.get_position()
ax.set_position([box.x0, box.y0 + box.height * 0.1,
                 box.width, box.height * 0.9])

ax.legend([source,
          cnst_plt,
          interval],
          ['Source',
           $r \pm e^{-1}$ ,
          'Correlation interval'],
          loc='upper center',
          bbox_to_anchor=(0.5,
                        -0.13),
          ncol=3,
          fancybox=True)

plt.grid(True)

plt.savefig('plot2.png')
plt.show()

# Autoregression
print('~' * 25)
print('Autoregression')
best_ar = 0
best_ar_err = 1e6
for m in range(4):
    try:
        print('Autoregression {}:'.format(m))
        print(' ' * 4 + str(calc_model_ar(m, data)))

        ncf_th = theory_ncf_ar(m, data)

        for x in range(1, 11):
            print(' ' * 4, x, round(ncf_th(x), 5))

        err = corr_error(ncf, ncf_th)
        print(' ' * 4 + str(err))

```

```

        if best_ar_err > err:
            best_ar = m
            best_ar_err = err
    except Exception as e:
        print(e)
print('Best ar err: {}, model ar({})'.format(best_ar_err, best_ar))

print('~' * 25)
print('Moving average')
best_ma = 0
best_ma_err = 1e6
for n in range(4):
    try:
        print('Moving average {}'.format(n))
        print(' ' * 4 + str(calc_model_ma(n, data, maxiter=maxiter)))

        ncf_th = theory_ncf_ma(n, data)

        for x in range(1, 11):
            print(' '*4, x, ncf_th(x))

        err = corr_error(ncf, ncf_th)
        print(' ' * 4 + str(err))

        if best_ma_err > err:
            best_ma = n
            best_ma_err = err
    except Exception as e:
        print(e)
print('Best ma err: {}, model ma({})'.format(best_ma_err, best_ma))

print('~' * 25)
print('ARMA')
best_arma_ar = 0
best_arma_ma = 0
best_arma_err = 1e6
for m in range(1, 4):
    for n in range(1, 4):
        try:
            print('ARMA({}, {})'.format(m, n))
            print(' ' * 4 + str(calc_model_arma(m, n, data, maxiter=maxiter)))

```

```

ncf_th = theory_ncf_arma(m, n, data)

for x in range(1, 11):
    print(' '*4, x, ncf_th(x))

err = corr_error(ncf, ncf_th)
print(' '*4 + str(err))

if best_arma_err > err:
    best_arma_ar = m
    best_arma_ma = n
    best_arma_err = err
except Exception as e:
    print(e)
print('Best arma err: {}, model arma({}, {})'.format(
    best_arma_err, best_arma_ar, best_arma_ma))

print('Data')
print('Mean: {}'.format(data.mean()))
print('Std: {}'.format(data.std()))

ar = model_ar(best_ar, data)([1] * best_ar, 1)
data_ar = np.array([next(ar) for _ in range(6000)][1000:]) + data.mean()

print('Modeled ar')
print('Mean: {}'.format(data_ar.mean()))
print('Mean err: {}'.format(abs(data_ar.mean() - data.mean())))
print('Std: {}'.format(data_ar.std()))
print('Std err: {}'.format(abs(data_ar.std() - data.std())))
ncf_ar = normalized_corr_func(data_ar)
for i in range(1, 11):
    print(i, round(ncf_ar(i), 5))
print('Error: {}'.format(round(corr_error(ncf, ncf_ar), 5)))

ma = model_ma(best_ma, data, maxiter=maxiter)(1)
data_ma = np.array([next(ma) for _ in range(6000)][1000:]) + data.mean()

print('Modeled ma')
print('Mean: {}'.format(data_ma.mean()))
print('Mean err: {}'.format(abs(data_ma.mean() - data.mean())))
print('Std: {}'.format(data_ma.std()))
print('Std err: {}'.format(abs(data_ma.std() - data.std())))
ncf_ma = normalized_corr_func(data_ma)

```

```

for i in range(1, 11):
    print(i, round(ncf_ma(i), 5))
print('Error: {}'.format(round(corr_error(ncf, ncf_ma), 5)))

arma = model_arma(
    best_arma_ar,
    best_arma_ma,
    data,
    maxiter=maxiter)(
    [1] *
    best_arma_ar,
    1)
data_arma = np.array([next(arma)
                      for _ in range(6000)][1000:]) + data.mean()

print('Modeled arma')
print('Mean: {}'.format(data_arma.mean()))
print('Mean err: {}'.format(abs(data_arma.mean() - data.mean())))
print('Std: {}'.format(data_arma.std()))
print('Std err: {}'.format(abs(data_arma.std() - data.std())))
ncf_arma = normalized_corr_func(data_arma)
for i in range(1, 11):
    print(i, round(ncf_arma(i), 5))
print('Error: {}'.format(round(corr_error(ncf, ncf_arma), 5)))

# AR plot

x = range(10)
y1 = list(map(ncf, x))
y2 = list(map(theory_ncf_ar(best_ar, data), x))
y3 = list(map(normalized_corr_func(data_ar), x))

fig = plt.figure()
ax = plt.subplot(111)

ncf_plot, = ax.plot(x, y1, 'red')
th_ncf_plot, = ax.plot(x, y2, 'green')
mdl_ncf_plot, = ax.plot(x, y3, 'blue')

plt.xlabel('Index number')
plt.ylabel('Normalized correlation function AR')

box = ax.get_position()

```

```

ax.set_position([box.x0, box.y0 + box.height * 0.1,
                box.width, box.height * 0.9])

ax.legend([ncf_plot, th_ncf_plot, mdl_ncf_plot], ['Source', 'Theoretical AR({})'.format(
    best_ar), 'Modelling'], loc='upper center', bbox_to_anchor=(0.5, -0.13), ncol=3,
    ↪ fancybox=True)

plt.grid(True)

plt.savefig('plot_ar_ncf.png')
plt.show()

# MA plot

x = range(10)
y1 = list(map(ncf, x))
y2 = list(map(theory_ncf_ma(best_ma, data), x))
y3 = list(map(normalized_corr_func(data_ma), x))

fig = plt.figure()
ax = plt.subplot(111)

ncf_plot, = ax.plot(x, y1, 'red')
th_ncf_plot, = ax.plot(x, y2, 'green')
mdl_ncf_plot, = ax.plot(x, y3, 'blue')

plt.xlabel('Index number')
plt.ylabel('Normalized correlation function MA')

box = ax.get_position()
ax.set_position([box.x0, box.y0 + box.height * 0.1,
                box.width, box.height * 0.9])

ax.legend([ncf_plot, th_ncf_plot, mdl_ncf_plot], ['Source', 'Theoretical MA({})'.format(
    best_ma), 'Modelling'], loc='upper center', bbox_to_anchor=(0.5, -0.13), ncol=3,
    ↪ fancybox=True)

plt.grid(True)

plt.savefig('plot_ma_ncf.png')
plt.show()

# ARMA plot

```

```

x = range(10)
y1 = list(map(ncf, x))
y2 = list(map(theory_ncf_arma(best_arma_ar, best_arma_ma, data), x))
y3 = list(map(normalized_corr_func(data_arma), x))

fig = plt.figure()
ax = plt.subplot(111)

ncf_plot, = ax.plot(x, y1, 'red')
th_ncf_plot, = ax.plot(x, y2, 'green')
mdl_ncf_plot, = ax.plot(x, y3, 'blue')

plt.xlabel('Index number')
plt.ylabel('Normalized correlation function ARMA')

box = ax.get_position()
ax.set_position([box.x0, box.y0 + box.height * 0.1,
                 box.width, box.height * 0.9])

ax.legend(
    [
        ncf_plot, th_ncf_plot, mdl_ncf_plot], [
        'Source', 'Theoretical ARMA({}, {}){}.format(
            best_arma_ar, best_arma_ma), 'Modelling'], loc='upper center', bbox_to_anchor=(
            0.5, -0.13), ncol=3, fancybox=True)

plt.grid(True)

plt.savefig('plot_arma_ncf.png')
plt.show()

# best would be arma

x = range(0, 151)
y = data_arma[0:151]

mean = data_arma.mean()
std = data_arma.std()

fig = plt.figure()
ax = plt.subplot(111)

```

```

data_plot, = ax.plot(x, y, 'gray')
mean_plot, = ax.plot([0, 151], [mean, mean], 'blue')
std_plot, = ax.plot([0, 151], [mean - std, mean - std], 'red')
_ = ax.plot([0, 151], [mean + std, mean + std], 'red')

plt.xlabel('Index number')
plt.ylabel('Random sequence values')

box = ax.get_position()
ax.set_position([box.x0, box.y0 + box.height * 0.1,
                box.width, box.height * 0.9])

ax.legend(
    [
        data_plot, mean_plot, std_plot], [
        'ARMA({},{}) process'.format(
            best_arma_ar, best_arma_ma), 'Average', 'Standart deviation'], loc='upper center
        ↪ ', bbox_to_anchor=(
            0.5, -0.13), ncol=3, fancybox=True)

plt.grid(True)

plt.savefig('plot_arma_modeled.png')
plt.show()

```


ПРИЛОЖЕНИЕ Б МЕТОД ПРОСТЫХ ИТЕРАЦИИ

Итерационный метод решения системы нелинейных алгебраических уравнений рассмотрим на примере модели СС(2). Известно, что рассматриваемый ниже метод простых итераций для устойчивых моделей АРСС всегда позволяет найти одно из решений системы, если это решение существует.

Модель СС(2)

$$\eta_n = \alpha_0 \xi_n + \alpha_1 \xi_{n-1} + \alpha_2 \xi_{n-2}$$

Система нелинейных уравнений для расчёта $\alpha_0, \alpha_1, \alpha_2$ имеет вид:

$$\begin{cases} R_\eta(0) = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 \\ R_\eta(1) = \alpha_0 \alpha_1 + \alpha_1 \alpha_2 \\ R_\eta(2) = \alpha_0 \alpha_2 \end{cases}$$

Значения $R_\eta(0), R_\eta(1), R_\eta(2)$ известны.

Начальное приближение: $\alpha_0 = \sqrt{R_\eta(0)}, \alpha_1 = 0, \alpha_2 = 0$.

Схема итераций:

1. $\alpha_2 \leftarrow \frac{R_\eta(2)}{\alpha_0}$.
2. $\alpha_1 \leftarrow \frac{R_\eta(1) - \alpha_1 \alpha_2}{\alpha_0}$.
3. Если $R_\eta(0) < \alpha_1^2 + \alpha_2^2$, то решений нет, иначе $\alpha_0 \leftarrow \sqrt{R_\eta(0) - \alpha_1^2 - \alpha_2^2}$.
4. Проверка на завершение итераций – уменьшение скорости сходимости итераций до заданного значения ϵ .
5. Переход к пункту 1.

В результате итерационного процесса будет найдено одно из возможных решений системы, что является достаточным в данной работе, или же $R_\eta(0) < \alpha_1^2 + \alpha_2^2$, что говорит о том, что решений нет.