Министерство образования и науки Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский Томский политехнический Университет»



Инженерная школа ядерных технологий

01.03.02 «Прикладная математика и информатика»

**Лабораторная работа №1**

по дисциплине:

**Статистическое моделирование и прогнозирование**

Вариант 17

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Выполнил:** |  | | | | |
| Студент группы | 0В01 |  | Саматов Д.С. |
|  |  |  |
| **Проверил:** | Крицкий О.Л. | | | | |
| преподаватель |  |  |  |  |  | |
|  |  |  |  |  |  | |

Томск 2023 г.

**Задание №1**

По данным о выпуске продукции за десять лет, которые представлены в табл. 1, в соответствии с номером варианта проверить временной ряд на случайность. Вычислить выборочные автоковариации и автокорреляции до пятого порядка включительно. Проверить статистику Дарбина-Уотсона на нулевое значение первой автокорреляции. Выявить наличие тренда и в случае положительного ответа построить трендовую часть модели. Тип трендовой модели выберите самостоятельно. Выявить присутствие сезонности и в случае положительного ответа построить сезонную часть модели.

Построить линейную или нелинейную регрессию по времени (окончательный вид и тип регрессии выбрать самостоятельно), оценив коэффициенты модели и вычислив дисперсию ошибки построенной модели.

Таблица 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № варианта | Годы выпуска продукции (t) | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 13,5 | 12,7 | 12 | 11,9 | 11,5 | 11,2 | 10,8 | 10,7 | 10,6 | 10,5 |
| 2 | 251 | 249 | 248 | 246 | 242 | 239 | 235 | 230 | 228 | 225 |
| 3 | 0,91 | 0,87 | 0,85 | 0,82 | 0,79 | 0,75 | 0,7 | 0,66 | 0,62 | 0,6 |
| 4 | 2,54 | 2,5 | 2,45 | 2,4 | 2,37 | 2,3 | 2,27 | 2,19 | 2,05 | 2 |
| 5 | 6,3 | 6,21 | 6,15 | 6 | 5,8 | -5,45 | 5,05 | 4,85 | 4,5 | 4,2 |
| 6 | 18,2 | 17,5 | 17,1 | 16,8 | 16,1 | 15,7 | 15,2 | 14,5 | 14,3 | 14 |
| 7 | 134 | 130 | 128 | 126 | 122 | 120 | 117 | 112 | 108 | 105 |
| 8 | 64 | 61 | 58 | 52 | 49 | 45 | 40 | 37 | 34 | 30 |
| 9 | 4,25 | 4,2 | 4,18 | 4,11 | 4,05 | 4 | 3,91 | 3,85 | 3,77 | 3,7 |
| 10 | 1,8 | 1,78 | 1,7 | 1,64 | 1,59 | 1,51 | 1,45 | 1,42 | 1,4 | 1,37 |
| 11 | 9,26 | 9,38 | 12,11 | 10,81 | 9,35 | 9,87 | 8,17 | 9,12 | 5,88 | 6,30 |
| 12 | 13,26 | 10,16 | 13,72 | 12,85 | 10,63 | 9,12 | 25,83 | 23,39 | 14,68 | 10,05 |
| 13 | 166,32 | 92,88 | 158,0 | 93,96 | 173,9 | 162,3 | 88,56 | 101,2 | 166,3 | 140,8 |
| 14 | 128,52 | 177,8 | 114,5 | 93,24 | 126,7 | 91,8 | 69,12 | 66,24 | 67,68 | 50,4 |

**Задание №2**

По данным табл. 1 проведите сглаживание данных (методом скользящего среднего или с помощью экспоненциального сглаживания) с длиной полуинтервала *m*=2 и выполните прогноз на следующий период *t*=11 для своего варианта задания, используя при необходимости трендовую (сезонную) части модели или их сумму, а так же линейную и нелинейную регрессии. Сравните погрешности различных методов на исходных данных.

**Задание №3**

В табл. 2 имеются данные об объеме экспорта по кварталам за 2011-2016 гг. Постройте модель временного ряда: трендовую (**четные** варианты), сезонную (**нечетные** варианты) или аддитивную (тренд+сезонность), если первоначальная модель **неполна или невозможна**. Спрогнозируйте экспорт по кварталам на 25 и 26 периоды времени.

Таблица 2

|  |  |  |  |
| --- | --- | --- | --- |
| Номер периода | Экспорт, млн руб. | Номер период | Экспорт, млн руб. |
| 1 | 4087 | 13 | 6975 |
| 2 | 4737 | 14 | 6891 |
| 3 | 5768 | 15 | 7527 |
| 4 | 6005 | 16 | 7971 |
| 5 | 5639 | 17 | 5875 |
| 6 | 6745 | 18 | 6140 |
| 7 | 6311 | 19 | 6248 |
| 8 | 7107 | 20 | 6041 |
| 9 | 5741 | 21 | 4626 |
| 10 | 7087 | 22 | 6501 |
| 11 | 7310 | 23 | 6284 |
| 12 | 8600 | 24 | 6707 |

**Теоретическое содержание**

*Тренды*

Рассмотрим временной ряд вида

где – трендовая составляющая или медленное изменение временного ряда в некотором направлении, которое сохраняется в течение длительного промежутка времени, – сезонная составляющая или изменения, которые происходят регулярно на ежегодной, ежемесячной, еженедельно и т.п. основе, например, выходные дни каждой недели или Новый год, ~ N(0,1) [1].

Существует три вида трендов:

1. Тренд среднего – временной ряд выглядит как колебания около медленно возрастающей или убывающей величины;
2. Тренд дисперсии – временной ряд имеет изменяющиеся во времени амплитуды колебаний (гетероскедастичность процесса);
3. Тренд автоковариации (автокорреляции) – временной ряд обладает изменчивостью корреляции между текущим и предшествующим значением ряда [1].

Тренды среднего бывают:

* полиномиальный тренд:

,

который при n=1 превращается в линейный тренд;

* экспоненциальный тренд:

,

* гармонический тренд:

,

где R– амплитуда колебаний, ω – частота, φ – фаза колебаний.

* тренд, описываемый логистической функцией

.

Для оценки коэффициентов полиномиального или экспоненциального тренда нужно использовать обычный МНК после введения новой переменной .

Далее, для моделирования сезонной составляющей можно использовать фиктивные переменные:

|  |  |
| --- | --- |
| , | (\*) |

где – сезонные фиктивные переменные. =1 в сезон j и нуль в остальное время, p – общее число сезонов [1].

**Замечание**: в моменты времени *s* окончания *j*-го сезона и начала (*j*+1)–го необходимо учитывать фиктивные переменные в (\*) более точно. Для этого требуется полагать , т.е. сумма весов должна быть равна 1 [1].

Отметим, что по построению , т.е. слагаемые в (\*) будут линейно- зависимыми. Значит, не найдется *p* линейно-независимых условий для оценивания неизвестных . Поэтому потребуем дополнительно выполнения условий нормировки:

что означает: сезонная компонента центрирована и влияние эффекта сезонности на уровень ряда оказывается равным нулю [1].

Подставим условие нормировки в (\*):

.

В последнем выражении переменные, стоящие в скобках, будут уже линейно-независимыми. Получили обычную линейную регрессию с (*p*-1) неизвестным коэффициентом. При этом *j*-е слагаемое в означает отклонение от основной динамики временного ряда на величину [1].

**Пример**: используем для оценки исходного временного ряда полиномиальный тренд и сезонность вида (\*):

*,*

где . Это линейная регрессия, для оценки (*n*+*p*+1) неизвестного коэффициента нужно использовать МНК. Вектор неизвестных параметров есть , где *X* – матрица исходных данных, включая первый единичный столбец для определения постоянной , *Y* – вектор-столбец значений . Остальные элементы матрицы *X* задаются по следующим правилам:

1. При учете трендовой составляющей время задается как есть, т.е. *t*=1, 2, …, *n*. Для учета квадратичного тренда *t*2=1, 4, …, *n*2. Переходить к безразмерному времени не требуется;
2. Для определения коэффициентов сезонной составляющей (\*) нужно создать подматрицу *Z*, в которой для *j*-го сезона требуется выделить целый столбец. В этом столбце фиктивные переменные равны единице во все моменты времени, относящиеся к этому сезону, кроме пограничных. В моменты времени сшивки двух сезонов (окончания *j*-го и начала (*j*+1)-го сезонов), фиктивные переменные должны быть равны , чтобы сумма весов по строкам матрицы *Z* была равна 1;
3. Полученную выше матрицу *Z* следует подвергнуть преобразованию и получить матрицу : вычесть из *j*-го столбца *Z* первый ее столбец и оставить результат в (*j*-1)-м столбце , *j* > 1. Присоединить элементы матрицы к матрице, описывающей трендовую составляющую, справа. Это и будет матрица *X* исходных данных для оценки регрессионной модели.

После нахождения оценок коэффициентов требуется вычислить ,так как он не входит в число уже найденных параметров [1].

*Сглаживание временного ряда*

При построении тренда вместо оценивания его коэффициентов можно использовать метод скользящих средних, когда значения временного ряда *xt* заменяются последовательностью вычисленных на перемещаемом отрезке средних его величин. Пусть (2*m*+1) – длина некоторого отрезка времени. Подберем полином

, *p*<*m*,

к группе первых (2*m*+1) членов ряда. Этот полином будем использовать в дальнейшем при определении значения тренда в точке *t*=(*m*+1), т.е. в середине выбранного отрезка. Далее временной отрезок сдвигается на единицу вправо, т.е. рассматриваются моменты 2, 3, …, (2*m*+2) и вычисления повторяются: находится значение *τm*+2. Проводим процедуру до тех пор, пока не будет достигнута последняя группа из (2*m*+1) точек [1].

Поиск коэффициентов полинома осуществляем с помощью МНК по первым (2*m*+1) точкам:

причем начало расчетов сдвинуто в точку *t*=*m*. Дифференцирование *F* по параметрам приводит к СЛАУ из (*p*+1) уравнения вида:

|  |  |
| --- | --- |
| ,  j = 0, 1, …, p. | (\*) |

Заметим, что в этом равенстве все суммы с нечетными степенями *j* равны нулю (в силу симметрии относительно начала суммирования и из-за наличия нечетных индексов):

Кроме того, полином, проходящий через (2*m*+1) точку, используется только для вычисления значения в точке *t*=*m*, или в новых обозначениях индекса суммирования, в точке *t*=0. Поставляя *t*=0 в , получаем значение *a*0. Система (\*) разбивается на две подсистемы: в первой участвуют только коэффициенты с четными индексами , а во второй – только с нечетными . Решение (\*) относительно зависит от численный значений и линейных функций типа . В итоге, – некоторое среднее арифметическое с весами , зависящими от m и p. Найденное выражение для применяется в дальнейшем для всех последующих временных отрезков скольжения без пересчета , так как процедура определения значений тренда в середине нового отрезка эквивалентна нахождению значения в новой точке середины t [1].

*Случайность данных*

Используем несколько выборочных автокорреляций , для доказательства гипотезы о случайности значений исходного временного ряда . Рассмотрим Q-статистику Льюнга-Бокса в предположении о нормальности распределения значений :

Было показано, что статистический критерий работает даже в условиях отсутствия нормального закона распределения для исходного временного ряда (выполнена ЦПТ, т.е. дисперсия D( ) < ∞). Нулевая гипотеза состоит в том, что ряд является винеровским процессом, т.е. процессом с независимыми приращениями и нулевым средним. Если значение статистики Q(r) больше критического (табличного) значения функции распределения при заданном уровне значимости и числе степеней свободы, то признается наличие ненулевых автокорреляций до порядка m включительно [1].

*Некоторые статистические критерии, используемые при анализе временных рядов*

Выборочной автоковариацией *k*-го порядка временного ряда называется число:

Выборочной автокорреляцией *k*-го порядка временного ряда называется число:

Известно, что выборочные автокорреляции имеют нормальное асимптотическое распределение [1].

*Стационарность временного ряда*

Автокорреляции удобны и для проверки временного ряда на стационарность. В целом можно заметить, что для выявления стационарности нужно вычислять автокорреляции до некоторого порядка и заметить, что коррелограмма быстро убывает после нескольких первых значений (первое значение может быть любым). Если же автокорреляция первого порядка близка к единице, а коррелограмма медленно убывает по экспоненте, то это свидетельствует о нестационарности [1].

*Статистическая значимость временного ряда*

Пусть для временного ряда вычислены автокорреляции до k-го порядка включительно. Проверим статистическую гипотезу о величине первой автокорреляции ρ1, вычисляя статистику Дарбина-Уотсона (DW-test):

где – погрешность модели (разность между наблюдаемым и модельным значением).

Значения статистики γ лежат в интервале [0,4]. Распределение статистики известно и имеет два критических значения: dl и du, их значения приведены в таблицах. Выдвигаем нулевую гипотезу H0: ρ1=0, где ρ1 - первая автокорреляция (остальные автокорреляции статистически не проверяются). Нулевая гипотеза о нулевой автокорреляции подтверждается, если du < γ < 4-du. H0 отклоняется в пользу альтернативы о наличии положительной автокорреляции, если γ < dl. H0 отклоняется в пользу альтернативы о наличии отрицательной автокорреляции, если 4-dl < γ. Зона неопределенности критерия, когда нельзя ни принять основную гипотезу, ни принять альтернативную, состоит из двух интервалов, описываемых неравенствами: dl < γ < du и 4-du < γ < 4-dl [1].

**Ход работы**

**Задание 1**

Для начала вычислим первые пять значений выборочные автоковариации и автокорреляции до пятого порядка включительно.

Значения первых пяти автоковариаций: 0.0073, 0.0043, 0.0016, -0.0008, -0.0027;

Значения первых пяти автокорреляций: 1, 0.59, 0.21, -0.11, -0.36, -0.53.

Значение статистики Люнга-Бокса составила – 31.45, а значение квантиля распределения хи-квадрат на уровне значимости 0.05 – 12.59. Следовательно, гипотезу о случайности исходных данных отвергаем.

Для начала построим распределение исходных данных рис.1.



Рисунок 1 – Распределение исходных данных.

Из данного распределения видно, что линейная модель без учета сезонности хорошо подойдет для описания исходных данных, рис. 2.

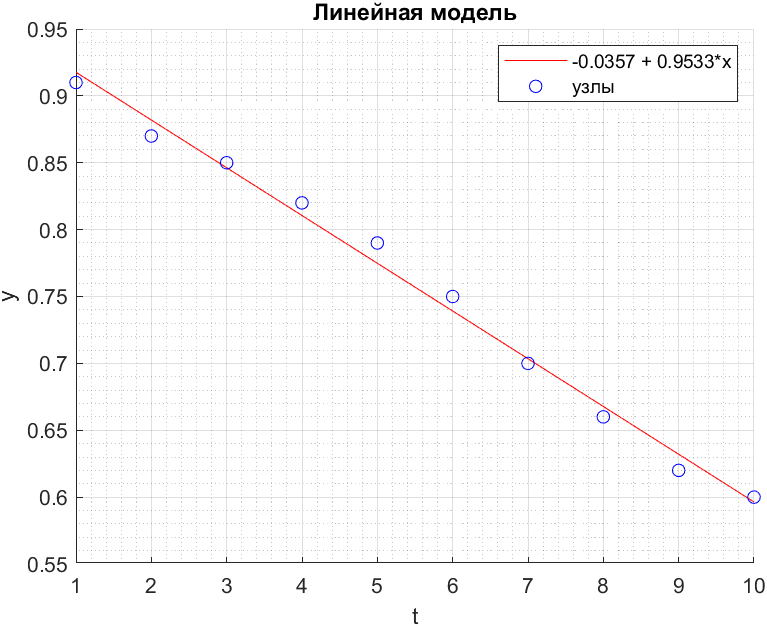


Рисунок 2 – Линейная модель.

Дисперсия ошибки построенной модели составила 0.0001. Использовать полином более высокого порядка не имеет смысла, так как данные монотонно убывают. Кроме этого, между минимальным и максимальным значениями очень малый разброс (т.е. 0.31), следовательно, не имеет смысла использовать показательную регрессию.

Проверим статистику Дарбина-Уотсона на нулевое значение первой автокорреляции.

Для выбранной линейной модели получим следующее значение статистики: 0.947. Соответствующие значения критических точек критерия Дарбина-Уотсона – 0.879 и 1.320. Таким образом, получим:

Значение статистики находится в зоне неопределенности, следовательно, нельзя ни принять основную гипотезу, ни принять альтернативную.

**Задание 2**

Проведем сглаживание данных методом скользящего среднего с длиной полуинтервала *m*=2.

Сглаживание данных будем проводить помощью следующей модели:

где *m* – полуинтервал для «окна данных» длины . Нумерация в каждом интервале будет внутренняя. Для оценки значения коэффициентов так же используется метод наименьших квадратов. При оценки коэффициентов модели следующие:

В таб. 1 приведем значения коэффициентов модели для каждого из «окон данных».

Таблица 1

Найденные коэффициенты выбранной модели

|  |  |  |  |
| --- | --- | --- | --- |
| Номер окна |  |  |  |
| 1 | 0.8466 | -0.0290 | 0.0007 |
| 2 | 0.8217 | -0.0300 | -0.0029 |
| 3 | 0.7891 | -0.0370 | -0.0036 |
| 4 | 0.7483 | -0.0410 | -0.0021 |
| 5 | 0.7026 | -0.0430 | 0.0007 |
| 6 | 0.6574 | -0.0380 | 0.0043 |

Предсказанное значение данной модели при равняется 0.58.

**Задание 3**

Для начала построим распределение исходных данных рис. 3.

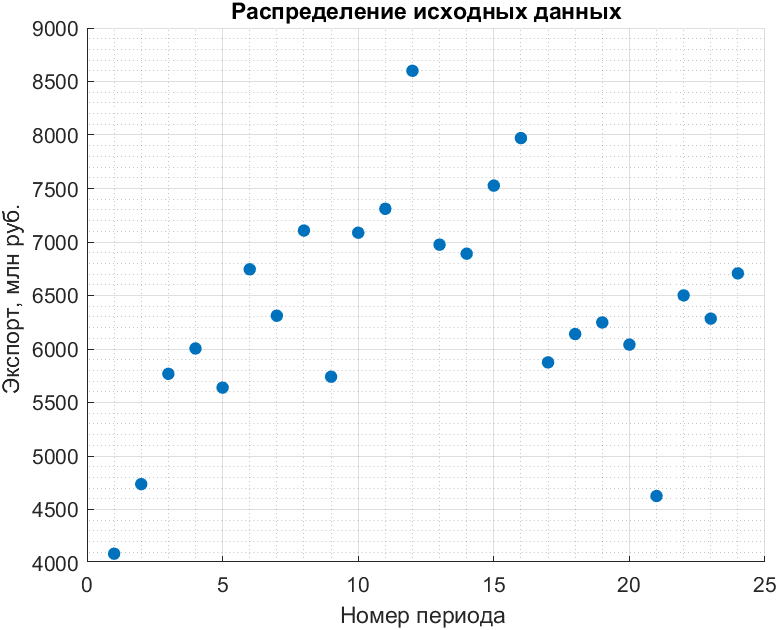
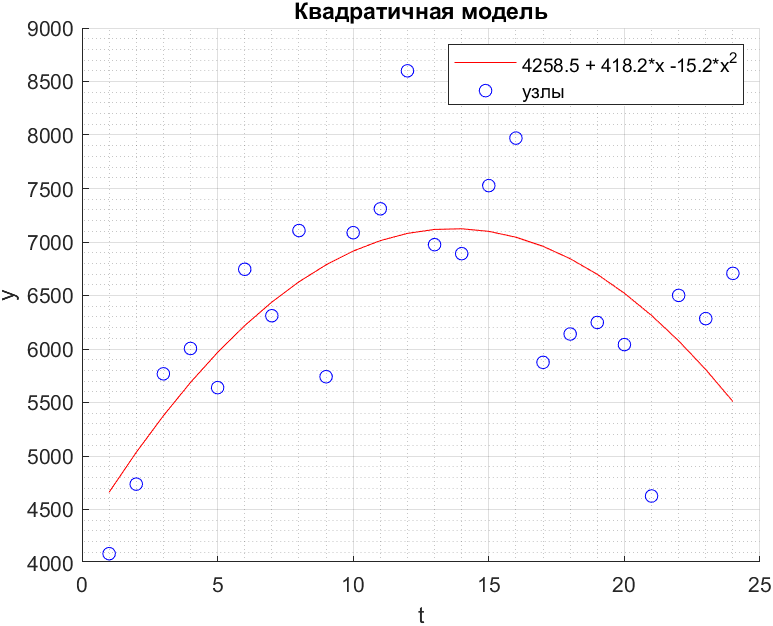
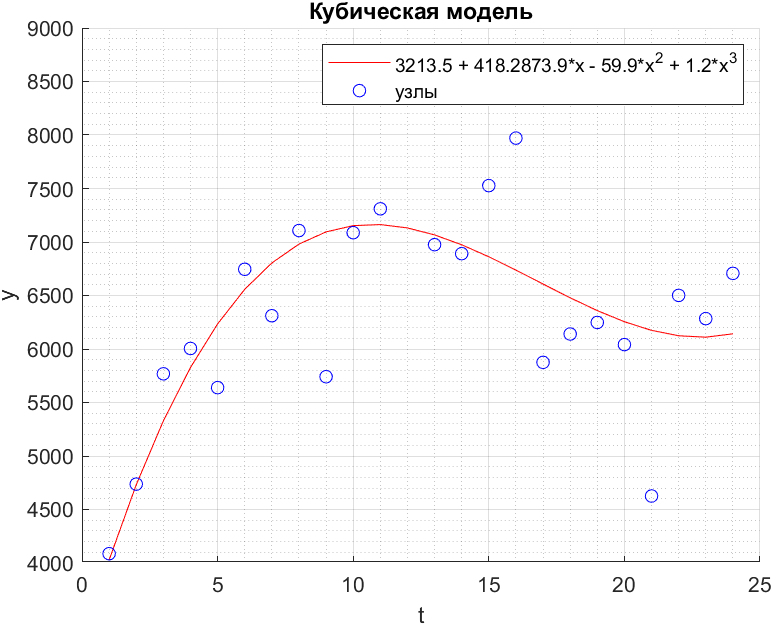
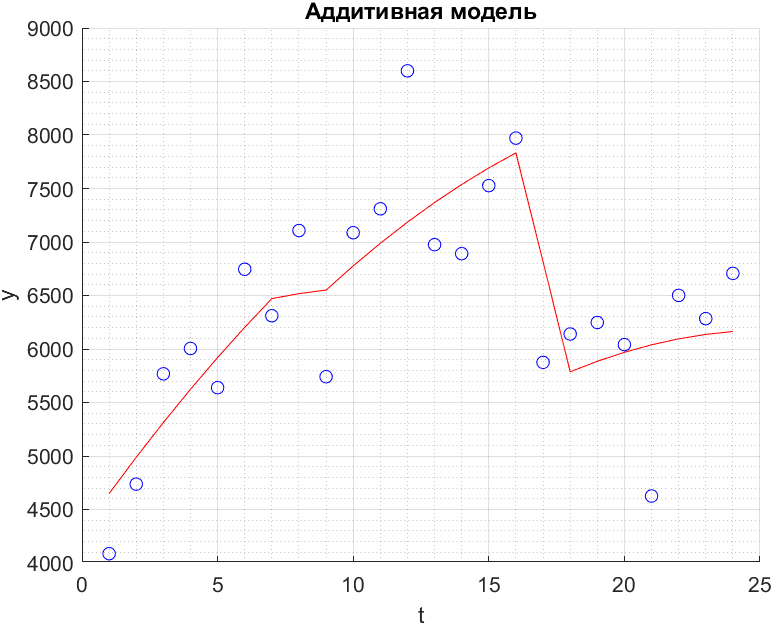


Рисунок 3 – Распределение исходных данных.

Для исходных данных из третьего задания построим три модели временного ряда: квадратичную, кубическую и квадратичную с учетом сезонности. Результаты простроенных моделей представлены на рис. 4.

(а) (б)



(в)

Рисунок 4 – (а) Квадратичная модель; (б) Кубическая модель; (в) Аддитивная модель.

Дисперсия ошибки построенных моделей представлены в табл. 2.

Таблица 2

Дисперсия ошибки моделей

|  |  |
| --- | --- |
| Модель | Дисперсия ошибки |
| Квадратичная | 606433.5 |
| Кубическая | 523403.2 |
| Аддитивная | 451673.5 |

Как видно из полученных результатов лучше всего показала себя аддитивная модель (т.е. квадратичная модель с тремя сезонами).

Значения, предсказанные данной моделью в моменты времени 25 и 26 получились равными: 6176.9 и 6175.9.

**Вывод**

В ходе выполнения лабораторной работы временной ряд был исследован на случайность при помощи критерия Льюнга-Бокса, проверены некоторые статистические критерии, используемые при анализе временных рядов. Также проверена статистическая значимость временного при помощи DW-test. К тому же исследованы различные модели для описания значений временного ряда: сезонные, трендовые, аддитивные. Кроме этого, был проведен анализ временного ряда при помощи сглаживания значений методом скользящего среднего.

**Список использованной литературы**

1. Крицкий О.Л. Эконометрика: монография О.Л. Крицкий; Томский политехнический университет. – Томск, 2022. – 37 с.

**Приложение А**

*main.m*

clc, clearvars, close all, format compact

y\_1 = [0.91, 0.87, 0.85, 0.82, 0.79, 0.75, 0.7, 0.66, 0.62, 0.6];

y\_2 = [4087, 4737, 5768, 6005, 5639, 6745, 6311, 7107, 5741, 7087, 7310, 8600, 6975, 6891, 7527, 7971, 5875, 6140, 6248, 6041, 4626, 6501, 6284, 6707];

n\_1 = length(y\_1);

n\_2 = length(y\_2);

t\_1 = 1:n\_1;

t\_2 = 1:n\_2;

X\_1 = horzcat(ones(n\_1, 1), transpose(t\_1));

X\_2 = horzcat(ones(n\_2, 1), transpose(t\_2), transpose(t\_2.^2));

X\_3 = horzcat(ones(n\_2, 1), transpose(t\_2), transpose(t\_2.^2), transpose(t\_2.^3));

theta\_1 = find\_parametrs(X\_1, y\_1);

disp\_1 = create\_regression\_model(t\_1, y\_1, theta\_1, 0);

[disp\_1, cov, cor, gamma] = find\_characteristics(n\_1, y\_1, t\_1, theta\_1);

Q = Q\_statistics(cor, length(y\_1));

chi = chi2inv(0.95, 6);

y\_pred = moving\_average\_method(y\_1);

theta\_2 = find\_parametrs(X\_2, y\_2);

disp\_2 = create\_regression\_model(t\_2, y\_2, theta\_2, 1);

center = [1, 9, 17, 24];

X\_season = zeros(24, 3);

theta\_3 = find\_parametrs(X\_3, y\_2);

disp\_3 = create\_regression\_model(t\_2, y\_2, theta\_3, 2);

for i = 1:3

for j = center(i):center(i+1)

X\_season(j, i) = 1;

end

end

X\_season(9,1) = 1/2;

X\_season(9,2) = 1/2;

X\_season(17,2) = 1/2;

X\_season(17,3) = 1/2;

X\_season = horzcat(X\_season(:,2) - X\_season(:,1), X\_season(:,3) - X\_season(:,1));

X\_4 = horzcat(X\_2, X\_season);

theta\_4 = find\_parametrs(X\_4, y\_2);

disp\_4 = create\_season\_regression\_model(t\_2, y\_2, theta\_4);

[y\_pred\_25, y\_pred\_26] = predict(theta\_4, [25, 26]);

create\_season\_regression\_model\_pred(t\_2, y\_2, theta\_4, [y\_pred\_25, y\_pred\_26]);

*create\_regression\_model.m*

function disp = create\_regression\_model(t, y, params, flag)

if flag == 0

f = @(x) params(1) + params(2)\*x;

elseif flag == 1

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2;

else

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + + params(4)\*x.^3;

end

figure('Color', 'w')

hold on

plot(t, f(t), Color='red')

plot(t, y, 'bo')

hold off

grid on

grid minor

if flag == 0

title('Линейная модель');

legend('-0.0357 + 0.9533\*x', 'узлы');

elseif flag == 1

title('Квадратичная модель');

legend('4258.5 + 418.2\*x -15.2\*x^2', 'узлы');

else

title('Кубическая модель');

legend('3213.5 + 418.2873.9\*x - 59.9\*x^2 + 1.2\*x^3', 'узлы');

end

xlabel('t')

ylabel('y')

disp = disp\_error(f(t), y, length(t), length(params));

end

*create\_season\_regression\_model.m*

function disp = create\_season\_regression\_model(t, y, params)

data = zeros(1, 24);

figure('Color', 'w')

for i = 1:24

if i < 8

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + (-params(4)-params(5));

data(i) = f(t(i));

elseif i == 8

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + (-params(4)-params(5)) \* 1/2 + params(4) \* 1/2;

data(i) = f(t(i));

elseif i > 8 && i < 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(4);

data(i) = f(t(i));

elseif i == 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(4) \* 1/2 + params(5) \* 1/2;

data(i) = f(t(i));

elseif i > 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(5);

data(i) = f(t(i));

end

end

hold on

plot(data, 'red')

plot(t, y, 'bo')

hold off

grid on

grid minor

title('Аддитивная модель');

xlabel('t')

ylabel('y')

disp = disp\_error(data, y, 24, length(params));

end

*create\_season\_regression\_model\_pred.m*

function create\_season\_regression\_model\_pred(t, y, params, y\_pred)

t\_new = horzcat(t, [25, 26]);

data = zeros(1, 26);

figure('Color', 'w')

for i = 1:26

if i < 8

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + (-params(4)-params(5));

data(i) = f(t\_new(i));

elseif i == 8

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + (-params(4)-params(5)) \* 1/2 + params(4) \* 1/2;

data(i) = f(t\_new(i));

elseif i > 8 && i < 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(4);

data(i) = f(t\_new(i));

elseif i == 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(4) \* 1/2 + params(5) \* 1/2;

data(i) = f(t\_new(i));

elseif i > 17

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(5);

data(i) = f(t\_new(i));

end

end

hold on

plot(data, 'red')

plot(t, y, 'bo')

plot([25, 26], y\_pred, '\*')

hold off

grid on

grid minor

title('Аддитивная модель');

xlabel('t')

ylabel('y')

end

*disp\_error.m*

function disp = disp\_error(y\_pred, y, n, p)

disp = 0;

for i = 1:n

disp = disp + (y\_pred(i) - y(i))^2;

end

disp = disp / (n-p);

end

*find\_characteristics.m*

function [disp, cov, cor, gamma] = find\_characteristics(n, y, t, params)

model = @(x) params(1) + params(2) \* x;

m = 5;

y\_pred = ones(1, n);

cov = zeros(6, 1);

cor = zeros(5, 1);

e = zeros(1, n);

for i = 1:n

y\_pred(i) = model(t(i));

end

disp = disp\_error(y\_pred, y, n, length(params));

y\_mean = mean(y\_pred);

for i = 1:m+1

sum = 0;

for j = 1:n-i

sum = sum + (y\_pred(j) - y\_mean) \* (y\_pred(j+i) - y\_mean);

end

cov(i) = sum / n;

end

for i = 1:m+1

cor(i) = cov(i) / cov(1);

end

e = y\_pred - y;

gamma = 0;

sum\_gamma = 0;

for i = 2:n

gamma = gamma + (e(i) - e(i-1))^2;

sum\_gamma = sum\_gamma + e(i)^2;

end

gamma = gamma / (sum\_gamma + e(1)^2);

end

*find\_parametrs.m*

function params = find\_parametrs(X, y)

params = inv(transpose(X) \* X) \* transpose(X) \* transpose(y);

end

*moving\_average\_method.m*

function y\_pred = moving\_average\_method(y)

a = zeros(6, 3);

t = 3;

for i = 1:6

a(i,1) = 1/35 \* (-3 \* y(i) + 12 \* y(i+1) + 17 \* y(i+2) + 12 \* y(i+3) - 3 \* y(i+4));

a(i,2) = 1/10 \* (-2 \* y(i) - y(i+1) + y(i+3) + 2 \* y(i+4));

a(i,3) = 1/14 \* (2 \* y(i) - y(i+1) - 2 \* y(i+2) - y(i+3) + 2 \* y(i+4));

end

y\_pred = a(6,1) + a(6,2) \* t + a(6,3) \* t^2;

end

*predict.m*

function [y\_pred\_25, y\_pred\_26] = predict(params, x\_test)

f = @(x) params(1) + params(2)\*x + params(3)\*x.^2 + params(5);

y\_pred\_25 = f(x\_test(1));

y\_pred\_26 = f(x\_test(2));

end

*Q\_statistics.m*

function q = Q\_statistics(r, T)

sum = 0;

for k = 1:length(r)

sum = sum + r(k) ^ 2 / (T-k);

end

q = sum \* (T + 2) \* T;

end