

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ  
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С. П. КОРОЛЕВА»  
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»

## МЕТОДЫ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО ПРОГНОЗИРОВАНИЯ

*Утверждено Редакционно-издательским советом университета  
в качестве методических указаний к лабораторным работам*

САМАРА  
Издательство СГАУ  
2011

УДК 33  
ББК СГАУ : У.вбя7

Составители: Е. П. Ростова, Е. Б. Кореева

Рецензент д-р техн. наук, профессор, заведующий каф.  
организации производства В. Г. Засканов

**Методы социально-экономического прогнозирования:**  
метод. указания к лаб. работам/ сост. *Е.П. Ростова,*  
*Е. Б. Кореева.* – Самара: Изд-во Самар. гос. аэрокосм. ун-  
та, 2011. – 35 с.

Рассматриваются основные модели и методы социально-экономического прогнозирования. Охарактеризованы адаптивные методики прогнозирования.

Предназначены для выполнения лабораторных работ по дисциплине «Методы социально-экономического прогнозирования» в рамках специальности «Математические методы экономики».

## Содержание

ЛАБОРАТОРНАЯ РАБОТА 1. ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ.....	4
ЛАБОРАТОРНАЯ РАБОТА 2. ПРИМЕНЕНИЕ МЕТОДОВ И МОДЕЛЕЙ МНОЖЕСТВЕННОЙ РЕГРЕССИИ ПРИ ПРОГНОЗИРОВАНИИ .....	7
ЛАБОРАТОРНАЯ РАБОТА 3. ПРОГНОЗИРОВАНИЕ ДЕМОГРАФИЧЕСКИХ И ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ.....	16
ЛАБОРАТОРНАЯ РАБОТА 4. АНАЛИЗ ВРЕМЕННОГО РЯДА .....	21
ЛАБОРАТОРНАЯ РАБОТА 5. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ С СЕЗОННЫМ КОМПОНЕНТОМ .....	25
ЛАБОРАТОРНАЯ РАБОТА 6. АНАЛИЗ ВРЕМЕННЫХ РЯДОВ С СЕЗОННЫМ КОМПОНЕНТОМ ПО ГАРМОНИКАМ .....	27
ЛАБОРАТОРНАЯ РАБОТА 7. ПРИМЕНЕНИЕ АДАПТИВНЫХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ: МЕТОД ХОЛЬТА.....	29
ЛАБОРАТОРНАЯ РАБОТА 8. ПРОГНОЗИРОВАНИЕ ВРЕМЕННОГО РЯДА С СЕЗОННЫМ КОМПОНЕНТОМ.....	31
ЛАБОРАТОРНАЯ РАБОТА 9. ПОДБОР ПОРЯДКА АППРОКСИМИРУЮЩЕГО ПОЛИНОМА С ПОМОЩЬЮ МЕТОДА ПОСЛЕДОВАТЕЛЬНЫХ РАЗНОСТЕЙ .....	32

**ЛАБОРАТОРНАЯ РАБОТА № 1**  
**ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ**  
**МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ**

**Содержание лабораторной работы**

Дан массив статистических данных:  $x_i, y_i (i = \overline{1, n})$ . Необходимо определить функциональную зависимость  $y=f(x)$  с помощью МНК.

**Порядок выполнения работы**

1. Составить квадратичную функцию  $y^{кв} = b_0 + b_1x + b_2x^2$ .
2. Составить степенную функцию  $y^{см} = ax^b$ .
3. Составить показательную функцию  $y^{нок} = ae^{bx}$ .
4. Определить коэффициенты прогнозируемых зависимостей.
5. Построить графики полученных функций  $y^{кв}, y^{см}, y^{нок}$  и график начальной функции  $y$  в одной системе координат.

**Описание этапов выполнения работы**

1. Составление матриц  $A = \begin{pmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix}, B = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{pmatrix}$ .

2. Вычисление коэффициентов  $b_0, b_1, b_2$ :  $b = A^{-1}B$ .  $b^T = (b_0, b_1, b_2)$ .

3. Составление функции  $y^{кв} = b_0 + b_1x + b_2x^2$  и вычисление ее значений  $y^{кв}_i = y^{кв}(x_i), i = \overline{1, n}$ .

4. Вычисление погрешности  $\delta_{кв} = \sqrt{\sum (y_i^{кв} - y_i)}$ .

5. Составление матриц  $A = \begin{pmatrix} n & \sum \ln x_i \\ \sum \ln x_i & \sum \ln^2 x_i \end{pmatrix}, B = \begin{pmatrix} \sum \ln y_i \\ \sum \ln y_i \ln x_i \end{pmatrix}$ .

6. Вычисление коэффициентов  $a$  и  $b$ :  $a = e^{b_0}, b = b_1$ .  $b = A^{-1}B$ .  $b^T = (b_0, b_1)$ .

7. Составление функции  $y^{см} = ax^b$  и вычисление ее значений  $y^{см}_i = y^{см}(x_i), i = \overline{1, n}$ .

8. Вычисление погрешности  $\delta_{см} = \sqrt{\sum (y_i^{см} - y_i)}$ .

9. Составление матриц  $A = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix}$ ,  $B = \begin{pmatrix} \sum \ln y_i \\ \sum x_i \ln y_i \end{pmatrix}$ .

10. Вычисление коэффициентов  $a$  и  $b$ :  $a = e^{b_0}$ ,  $b = b_1$ .  $b = A^{-1}B$ .  
 $b^T = (b_0, b_1)$ .

11. Составление функции  $y^{нок} = ae^{bx}$  и вычисление ее значений  
 $y^{нок}_i = y^{нок}(x_i)$ ,  $i = \overline{1, n}$ .

12. Вычисление погрешности  $\delta_{нок} = \sqrt{\sum (y_i^{нок} - y_i)}$ .

13. Построение графиков  $y^{кв}$ ,  $y^{см}$ ,  $y^{нок}$ ,  $y$  в одной системе координат.

### Пример выполнения работы

Статистические данные:

Таблица 1

$x_i$	1	1,33	1,67	2	2,33	2,67	3	3,33	3,67	4	4,33	4,67	5
$y_i$	3	2,83	2,67	2,5	2,33	2,17	2	1,25	1,50	1,75	2,00	2,25	2,5

1. Составим матрицы  $A = \begin{pmatrix} 13 & 39 & 137,22 \\ 39 & 137,22 & 533 \\ 137,22 & 533 & 2201,17 \end{pmatrix}$ ,  $B = \begin{pmatrix} 28,75 \\ 81,78 \\ 282,67 \end{pmatrix}$ .

2. Вычислим коэффициенты  $b_0$ ,  $b_1$ ,  $b_2$ :  $b = \begin{pmatrix} 4,69 \\ -1,68 \\ 0,24 \end{pmatrix}$ .

3. Составим функцию  $y^{кв} = 4,69 - 1,68x + 0,24x^2$  и вычислим ее значения:

Таблица 2

$x_i$	1	1,33	1,67	2	2,33	2,67	3	3,33	3,67	4	4,33	4,67	5
$y^{кв}_i$	3,25	2,88	2,56	2,30	2,09	1,93	1,83	1,79	1,79	1,85	1,97	2,14	2,37

4. Вычислим погрешность  $\delta_{кв} = \sqrt{\sum (y_i^{кв} - y_i)} = 0,82$ .

5. Составим матрицы  $A = \begin{pmatrix} 13 & 12,92 \\ 12,92 & 15,88 \end{pmatrix}$ ,  $B = \begin{pmatrix} 9,96 \\ 8,99 \end{pmatrix}$ .

6. Вычислим коэффициенты  $a$  и  $b$ :  $a = e^{b_0}$ ,  $b = b_1$ .  $b = \begin{pmatrix} 1,07 \\ -0,3 \end{pmatrix}$ . Тогда

$a = 2,9$  и  $b = -0,3$ .

7. Составим функцию  $y^{cm} = 2,9x^{-0,3}$  и вычислим ее значения:

Таблица 3

$x_i$	1	1,33	1,67	2	2,33	2,67	3	3,33	3,67	4	4,33	4,67	5
$y^{cm}_i$	2,90	2,66	2,49	2,36	2,25	2,16	2,08	2,02	1,96	1,91	1,87	1,82	1,79

8. Вычислим погрешность  $\delta_k = \sqrt{\sum (y_i^{cm} - y_i)^2} = 1,28$ .

9. Составим матрицы  $A = \begin{pmatrix} 13 & 39 \\ 39 & 137,22 \end{pmatrix}$ ,  $B = \begin{pmatrix} 9,96 \\ 27,91 \end{pmatrix}$ .

10. Вычислим коэффициенты  $a$  и  $b$ :  $a = e^{b_0}$ ,  $b = b_1$ .  $b = \begin{pmatrix} 1,06 \\ -0,1 \end{pmatrix}$ . Тогда

$a = 2,88$  и  $b = -0,1$ .

11. Составим функцию  $y^{nok} = 2,88e^{-0,1x}$  и вычислим ее значения:

Таблица 4

$x_i$	1	1,33	1,67	2	2,33	2,67	3	3,33	3,67	4	4,33	4,67	5
$y^{nok}_i$	2,61	2,53	2,45	2,37	2,30	2,22	2,15	2,08	2,02	1,95	1,89	1,83	1,77

12. Вычислим погрешность  $\delta_{nok} = \sqrt{\sum (y_i^{nok} - y_i)^2} = 1,43$ .

13. Построим графики  $y^{kv}$ ,  $y^{cm}$ ,  $y^{nok}$ ,  $y$  в одной системе координат.

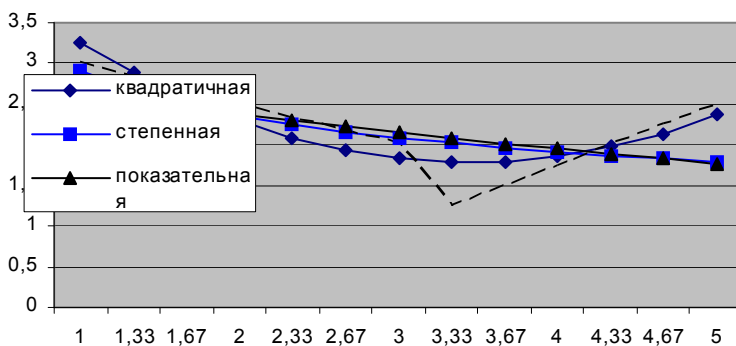


Рис. 1. Сравнение линий трендов

**ЛАБОРАТОРНАЯ РАБОТА № 2**  
**ПРИМЕНЕНИЕ МЕТОДОВ И МОДЕЛЕЙ**  
**МНОЖЕСТВЕННОЙ РЕГРЕССИИ ПРИ ПРОГНОЗИРОВАНИИ**

**Содержание лабораторной работы**

Исследовать зависимость результирующей переменной  $y$  от ряда объясняющих переменных  $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ .

**Порядок работы**

1. Нахождение уравнения связи.
2. Вычисление среднеквадратических ошибок в оценивании коэффициентов регрессии.
3. Проверка гипотезы об отсутствии какой бы то ни было линейной связи между  $y$  и совокупностью объясняющих переменных.
4. Пошаговый отбор наиболее существенных объясняющих переменных.

**Описание**

1. Нахождение уравнения связи.

С помощью метода наименьших квадратов (МНК) определим оценки  $\Theta_{МНК} = (\theta_0, \theta_1, \dots, \theta_p)$  для неизвестных значений параметров функции регрессии соответственно  $\theta_0, \theta_1, \dots, \theta_p$ , при которых сглаженные (регрессионные) значения  $\theta_0 + \theta_1 x_i^{(1)} + \theta_2 x_i^{(2)} + \dots + \theta_p x_i^{(p)}$  результирующего показателя как можно меньше отличались бы от соответствующих наблюдаемых значений  $y_i$  ( $i = \overline{1, n}$ ):

$$\Theta_{МНК} = (X^T X)^{-1} X^T Y, \tag{1}$$

где матрицы  $X = \begin{pmatrix} 1 & x_1^{(1)} & \dots & x_1^{(p)} \\ 1 & x_2^{(1)} & \dots & x_2^{(p)} \\ \dots & \dots & \dots & \dots \\ 1 & x_n^{(1)} & \dots & x_n^{(p)} \end{pmatrix}$  и  $Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ ,

$x_i^{(j)}$  – значение  $j$ -го предиктора (объясняющей переменной)  $i$ -го объекта.

После нахождения значений  $\theta_0, \theta_1, \dots, \theta_p$  можно записать уравнение связи:

$$f(X) = \theta_0 + \theta_1 x^{(1)} + \theta_2 x^{(2)} + \dots + \theta_p x^{(p)}. \quad (2)$$

2. Вычисление среднеквадратических ошибок  $s_l$  в оценивании коэффициентов регрессии  $\theta_l$  ( $l = \overline{0,5}$ ).

Для нахождения среднеквадратических ошибок воспользуемся следующими формулами:

$$\begin{aligned} \sigma^2 &= \frac{1}{n-p-1} \sum (y_i - \theta_0 - \theta_1 x_i^{(1)} - \theta_2 x_i^{(2)} - \dots - \theta_p x_i^{(p)})^2 = \\ &= \frac{1}{n-p-1} (Y - \Theta X)^T (Y - \Theta X) \end{aligned} \quad (3)$$

$$A = (X^T X)^{-1}, \quad (4)$$

$$B = \sigma^2 A, \quad (5)$$

$$s_l = \sqrt{b_{ll}} = \sigma \sqrt{a_{ll}}, \quad (l = \overline{0,5}). \quad (6)$$

3. Проверка гипотезы об отсутствии какой бы то ни было линейной связи между  $y$  и совокупностью объясняющих переменных  $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ .

Выдвигается гипотеза  $H_0: \theta_1 = \theta_2 = \dots = \theta_p = 0$ .

Данную гипотезу можно проверить следующим образом. По заданному уровню значимости критерия  $\alpha$  определяем из Таблицы  $100\alpha\%$ -ную точку  $F(p, n-p-1)$ -распределения  $v_\alpha^2(p, n-p-1)$ . Если окажется, что

$$\gamma = \frac{R_{y.X}^2}{1 - R_{y.X}^2} \cdot \frac{n-p-1}{p} > v_\alpha^2(p, n-p-1), \quad (7)$$

то гипотеза об отсутствии линейной связи между  $y$  и  $X$  отвергается (с вероятностью ошибиться, равной  $\alpha$ ), и принимается – в противном случае.



Множественный коэффициент корреляции  $R_{y.X}^2$  находится по следующей формуле:

$$R_{y.X}^2 = 1 - \frac{\det R}{R_{00}}, \quad (8)$$

где  $R$  – корреляционная матрица размера  $(p+1) \times (p+1)$ ,  $R = [r_{kl}]$ ,  $k = \overline{0, p}$ ,  $l = \overline{0, p}$

$$r(x^{(l)}, y) = \frac{\sum_{i=1}^n (x_i^{(l)} - \bar{x}^{(l)})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i^{(l)} - \bar{x}^{(l)})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (9)$$

$R_{00}$  – алгебраическое дополнение элемента  $r_{00}$  (первая строка и первый столбец матрицы соответствуют результирующей переменной  $y$ ).

4. Пошаговый отбор наиболее существенных объясняющих переменных.

Метод пошагового отбора переменных основывается на методе «всех возможных регрессий», в основе которого лежит критерий

$$R_{y.X^0(k)}^2 = \max_{X(k), k \leq p} R_{y.X(k)}^2. \quad (10)$$

**1 шаг ( $k = 1$ ).** В классе моделей регрессии  $y$  по единственной объясняющей переменной выбирается наиболее информативный (в смысле критерия) предиктор. При  $k = 1$  величина  $R_{y.X}^2(k)$  совпадает с квадратом обычного (парного) коэффициента корреляции  $r(x^{(l)}, y)$ , среди которых находится максимальный. Он и является наиболее информативным предиктором в классе однофакторных регрессионных моделей. Вычислим подправленное (на несмещенность) значение  $r^{*2}(x^{(4)}, y) = R^{*2}(1)$  и его нижнюю доверительную границу  $R_{\min}^2(1)$ :

$$R_{y.X}^{*2}(k) \approx 1 - (1 - R_{y.X}^2(k)) \frac{n-1}{n-p-1}, \quad (11)$$

$$R_{\min}^2(k) = R^{*2}(k) - 2\sqrt{\frac{2k(n-k-1)}{(n-1)(n^2-1)}}(1-R^2(k)). \quad (12)$$

Здесь  $p = k = 1$ .

**2 шаг ( $k=2$ ).** Среди возможных пар объясняющих переменных  $(x^{(l^*)}, x^{(j)})$ ,  $(j = \overline{0,5})$  выбирается наиболее информативная пара предикторов.

Далее по формулам (8), (11) и (12) вычисляем коэффициент детерминации, его подправленное (на несмещенность) значение и его нижнюю доверительную границу соответственно.

Сравнение нижних доверительных границ подтверждает или опровергает целесообразность включения в модель второй переменной.

Если  $R_{\min}^2(1) < R_{\min}^2(2)$ , то включаем в модель вторую переменную, если  $R_{\min}^2(1) > R_{\min}^2(2)$ , то не включаем и метод пошагового отбора закончен.

Оценка коэффициентов уравнения регрессии получена по формуле (1) и по формулам (3) – (6) вычислены их среднеквадратические ошибки.

**3 шаг ( $k=3$ ).** Среди возможных троек объясняющих переменных  $(x^{(l^*)}, x^{(l^{**})}, x^{(j)})$ ,  $(j = \overline{0,5})$ . Аналогично предыдущему шагу, находим:  $R^{*2}(3)$ ,  $R_{\min}^2(3)$ .

Сравнение нижних доверительных границ  $R_{\min}^2(2)$  и  $R_{\min}^2(3)$  говорит о целесообразности включения третьей переменной в модель.

## Пример

$i$	$y$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$
1	9,7	1,59	0,26	2,05	0,32	0,14
2	8,4	0,34	0,28	0,46	0,59	0,66
3	9	2,53	0,31	2,46	0,3	0,31
4	9,9	4,63	0,4	6,44	0,43	0,59
5	9,6	2,16	0,26	2,16	0,39	0,16
6	8,6	2,16	0,3	2,69	0,32	0,17
7	12,5	0,68	0,29	0,73	0,42	0,23
8	7,6	0,35	0,26	0,42	0,21	0,08
9	6,9	0,52	0,24	0,49	0,2	0,08
10	13,5	3,42	0,31	3,02	1,37	0,73
11	9,7	1,78	0,3	3,19	0,73	0,17
12	10,7	2,4	0,32	3,3	0,25	0,14
13	12,1	9,36	0,4	11,51	0,39	0,38
14	9,7	1,72	0,28	2,26	0,82	0,17
15	7	0,59	0,29	0,6	0,13	0,35
16	7,2	0,28	0,26	0,3	0,09	0,15
17	8,2	1,64	0,29	1,44	0,2	0,08
18	8,4	0,09	0,22	0,05	0,43	0,2
19	13,1	0,08	0,25	0,03	0,73	0,2
20	8,7	1,36	0,26	0,17	0,99	0,42

1. Нахождение уравнения связи.

Из условия:  $n=20$ ,  $p=5$ .

Вычислим оценки  $\theta_0, \theta_1, \dots, \theta_p$ :

$$\Theta_{МНК} = \begin{pmatrix} 3,51 \\ -0,01 \\ 15,54 \\ 0,11 \\ 4,47 \\ -2,93 \end{pmatrix}.$$

Составим матрицу  $X$  в виде таблицы:

1	1,59	0,26	2,05	0,32	0,14
1	0,34	0,28	0,46	0,59	0,66
1	2,53	0,31	2,46	0,3	0,31
1	4,63	0,4	6,44	0,43	0,59
1	2,16	0,26	2,16	0,39	0,16
1	2,16	0,3	2,69	0,32	0,17
1	0,68	0,29	0,73	0,42	0,23
1	0,35	0,26	0,42	0,21	0,08
1	0,52	0,24	0,49	0,2	0,08
1	3,42	0,31	3,02	1,37	0,73
1	1,78	0,3	3,19	0,73	0,17
1	2,4	0,32	3,3	0,25	0,14
1	9,36	0,4	11,51	0,39	0,38
1	1,72	0,28	2,26	0,82	0,17
1	0,59	0,29	0,6	0,13	0,35
1	0,28	0,26	0,3	0,09	0,15
1	1,64	0,29	1,44	0,2	0,08
1	0,09	0,22	0,05	0,43	0,2
1	0,08	0,25	0,03	0,73	0,2
1	1,36	0,26	0,17	0,99	0,42

После нахождения значений  $\theta_0, \theta_1, \dots, \theta_p$  можно записать уравнение связи:

$$f(X) = 3,51 - 0,01x^{(1)} + 15,54x^{(2)} + 0,11x^{(3)} + 4,47x^{(4)} - 2,93x^{(5)}.$$

2. Вычисление среднеквадратических ошибок  $s_l$  в оценивании коэффициентов регрессии  $\theta_l$  ( $l = \overline{0,5}$ ).

Для нахождения среднеквадратических ошибок воспользуемся следующими формулами:

$$\sigma^2 = 2,56.$$

Тогда  $s_0 = 5,42$ ,  $s_1 = 0,93$ ,  $s_2 = 21,5$ ,  $s_3 = 0,83$ ,  $s_4 = 1,54$ ,  $s_5 = 3,09$ .

Матрицу  $B = \sigma^2 A$  представим в виде таблицы:

29,36	-1,08	-115,47	2,33	-3,29	9,64
-1,08	0,87	4,55	-0,72	-0,21	-0,75
-115,47	4,55	462,38	-9,61	11,25	-40,36
2,33	-0,72	-9,61	0,69	0,04	0,99
-3,29	-0,21	11,25	0,04	2,38	-2,84
9,64	-0,75	-40,36	0,99	-2,84	9,54

Уравнение связи запишется следующим образом:

$$f(X) = 3,51 - 0,01x^{(1)} + 15,51x^{(2)} + 0,11x^{(3)} + 4,47x^{(4)} - 2,93x^{(5)}.$$

(5,42) (0,93) (21,5) (0,83) (1,54) (3,09)

3. Проверка гипотезы об отсутствии какой бы то ни было линейной связи между  $y$  и совокупностью объясняющих переменных  $x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}$ .

Выдвигается гипотеза  $H_0: \theta_1 = \theta_2 = \dots = \theta_p = 0$ .

Составим корреляционную матрицу  $R$  и запишем ее в виде таблицы:

1	0,43	0,37	0,40	0,58	0,33
0,43	1	0,85	0,98	0,11	0,34
0,37	0,85	1	0,88	0,03	0,46
0,40	0,98	0,88	1	0,03	0,28
0,58	0,11	0,03	0,03	1	0,57
0,33	0,34	0,46	0,28	0,57	1

Вычислим  $R_{y.X}^2 = 1 - \frac{0,00147}{0,003} = 0,517, \gamma = 3$ .

По заданному уровню значимости критерия  $\alpha = 0,05$  определяем из Таблицы 5, т. е.  $v_{0,05}^2(5, 14) = 2,96$ .

Поскольку  $\gamma > v_{0,05}^2(5, 14)$ , значит гипотеза  $H_0$  об отсутствии линейной связи отвергается. Это говорит о целесообразности продолжения изучения линейной связи между  $y$  и  $x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}, x^{(5)}$ .

4. Пошаговый отбор наиболее существенных объясняющих переменных.

**1 шаг (k = 1).**

Так как  $k = 1$ , то  $R_{y,x}^2(1) = r(x^{(1)}, y)$ ,

значит  $R^2(1) = \max_{1 \leq l \leq 5} r^2(x^{(l)}, y) = r^2(x^{(4)}, y) = 0,58^2 = 0,336$ .

Наиболее информативным предиктором в классе однофакторных регрессионных моделей оказывается переменная  $x^{(4)}$ . Тогда модель будет иметь вид:

$$f(X) = 7,87 + 3,55x^{(4)}.$$

Среднеквадратические ошибки:  $s_0 = 7,412$ ,  $s_4 = 4,16$ .

Оценки коэффициентов вычислены по формуле (1).

Вычислим подправленное (на несмещенность) значение  $r^{*2}(x^{(4)}, y) = R^{*2}(1)$  и его нижнюю доверительную границу  $R_{\min}^2(1)$ :

$$R^{*2}(1) \approx 1 - (1 - 0,336) \frac{20 - 1}{20 - 1 - 1} = 0,3,$$

$$R_{\min}^2(1) = 0,3 - 2 \sqrt{\frac{2 \cdot 1 \cdot (20 - 1 - 1)}{(20 - 1)(20^2 - 1)}} (1 - 0,336) = 0,208.$$

**2 шаг (k=2).** Среди возможных пар объясняющих переменных ( $x^{(4)}, x^{(j)}$ ),  $j = 1, 2, 3, 5$  выбирается наиболее информативная пара предикторов. Поскольку  $\max_{\substack{1 \leq j \leq 5 \\ j \neq 4}} R_{y,(x^{(4)}, x^{(j)})}^2 = R_{y,(x^{(4)}, x^{(3)})}^2 = 0,483$ , то наиболее ин-

формативной парой предикторов оказываются объясняющие переменные  $x^{(4)}$  и  $x^{(3)}$ . Далее по формулам (8), (11) и (12) вычисляем коэффициент детерминации, его подправленное (на несмещенность) значение и его нижнюю доверительную границу соответственно:

$$R_{y,(x^{(4)}, x^{(3)})}^2 = 0,482, \quad R^{*2}(1) \approx 1 - (1 - 0,482) \frac{20 - 1}{20 - 2 - 1} = 0,421 (p=2),$$

$$R_{\min}^2(2) = 0,323.$$

Сравнение нижних доверительных границ  $R_{\min}^2(1) < R_{\min}^2(2)$ , подтверждает целесообразность включения в модель в качестве второй переменной предиктора  $x^{(3)}$ . Тогда уравнение регрессии  $y$  по  $x^{(4)}$  и  $x^{(3)}$  будет иметь вид:

$$f(x^{(3)}, x^{(4)}) = 7,29 + 0,28 x^{(3)} + 3,48x^{(4)}.$$

Коэффициенты получены по формуле (1) и по формулам (3) – (6) получены их среднеквадратические ошибки:  $s_0=0,66$ ,  $s_3=0,13$ ,  $s_4=1,07$ .

3 шаг ( $k=3$ ). Среди возможных троек объясняющих переменных  $(x^{(3)}, x^{(4)}, x^{(j)})$ , ( $j=1, 2, 5$ ). наиболее информативной оказалась тройка  $(x^{(4)}, x^{(3)}, x^{(5)})$ , поскольку  $\max_{j=1,2,5} R^2_{y(x^{(4)}, x^{(3)}, x^{(j)})} = R^2_{y(x^{(4)}, x^{(3)}, x^{(5)})} = 0,513$ .

Аналогично предыдущему шагу, находим:

$$R^{*2}(3) = 0,422, R^2_{\min}(3) = 0,312.$$

Сравнение нижних доверительных границ  $R^2_{\min}(2)$  и  $R^2_{\min}(3)$  говорит о том, что третью переменную  $x^{(5)}$  в модель включать нецелесообразно, т.к.  $R^2_{\min}(2) > R^2_{\min}(3)$ .

Итак, модель имеет вид:  $f(x^{(3)}, x^{(4)}) = 7,29 + 0,28 x^{(3)} + 3,48x^{(4)}$ .

Таблица 5

Значения 100Q-процентных точек  $v^2_Q(v_1, v_2)$  F-распределения с числом степеней свободы числителя  $v_1$  и знаменателя  $v_2$ .

Q	0,1	0,05	0,01
$v^2_Q(5,14)$	2,31	2,96	4,96

## ЛАБОРАТОРНАЯ РАБОТА № 3

### ПРОГНОЗИРОВАНИЕ ДЕМОГРАФИЧЕСКИХ И ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ

#### Содержание лабораторной работы

Даны демографические показатели  $M_t^+$ ,  $M_t^-$ ,  $V_t^+$ ,  $V_t^-$ ,  $L_0$ , объем основного капитала  $K_t$ , объем производимой продукции  $Y_t$  за 30 периодов  $t = 1, 2, \dots, 30$ .

Необходимо спрогнозировать численность населения, объем основного капитала и объем производимой продукции в 35 периоде; определить коэффициенты в производственной функции Кобба-Дугласа на основании ретроспективных данных.

#### Порядок работы

1. Вычислить значения  $L_t$  за периоды  $t=0, 1, \dots, 29$ .
2. Построить графики  $L_t$ ,  $K_t$ .
3. Построить функцию тренда  $L(t)$  и  $K(t)$ .
4. Спрогнозировать значения  $L_{35}$  и  $K_{35}$ .
5. Вычислить параметры производственной функции Кобба-Дугласа.
6. Определить прогнозное значение  $Y_{35}$ .

#### Описание

1. Рассчитать численность населения  $L_t$  за периоды  $t=0, 1, \dots, 29$ :  
$$L_{t+1} = L_t - M_t^- + M_t^+ - V_t^- + V_t^+.$$
2. Построить график  $L_t$ .
3. Построить график  $K_t$ .
4. Получить функцию тренда  $L(t)$ . Среди нескольких возможных функций выбрать ту, которая дает наилучшее приближение. Для сравнения указать результаты других трендовых функций.
5. Аналогично п.4 получить функцию тренда  $K(t)$ .
6. На основе полученной функции  $L(t)$  вычислить прогнозное значение для  $t=35$ .
7. Аналогично п.6 вычислить  $K(t)$  для  $t=35$ .



8. Для определения параметров функции Кобба-Дугласа, решить систему уравнений:

$$\left\{ \begin{array}{l} n \ln A + \alpha \sum_{i=0}^{29} \ln K_i + \beta \sum_{i=0}^{29} \ln L_i = \sum_{i=0}^{29} \ln Q_i, \\ \ln A \sum_{i=0}^{29} \ln K_i + \alpha \sum_{i=0}^{29} \ln^2 K_i + \beta \sum_{i=0}^{29} \ln L_i \ln K_i = \sum_{i=0}^{29} \ln Q_i \ln K_i, \\ \ln A \sum_{i=0}^{29} \ln L_i + \alpha \sum_{i=0}^{29} \ln K_i \ln L_i + \beta \sum_{i=0}^{29} \ln^2 L_i = \sum_{i=0}^{29} \ln Q_i \ln L_i. \end{array} \right.$$

9. Подставить полученные значения  $A^*$ ,  $\alpha^*$ ,  $\beta^*$  в функцию  $Y_{35} = A^* L_{35}^{\beta^*} K_{35}^{\alpha^*}$ .

### Пример

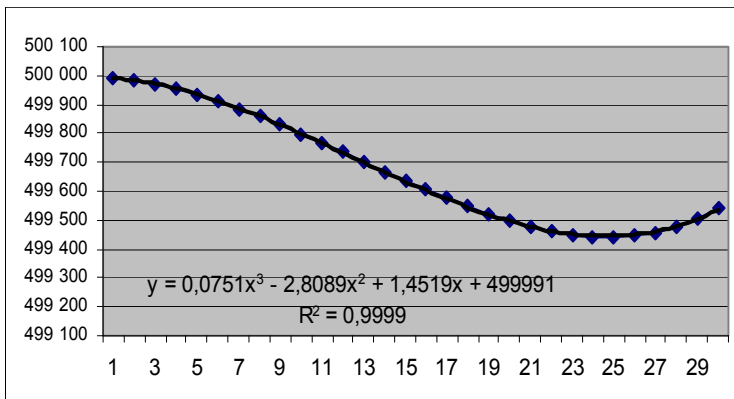
Даны демографические показатели  $M_t^+$ ,  $M_t^-$ ,  $V_t^+$ ,  $V_t^-$ ,  $L_0$ , объем основного капитала  $K_t$ , объем производимой продукции  $Y_t$  за 30 периодов  $t = 1, 2, \dots, 30$ .

Необходимо спрогнозировать численность населения, объем основного капитала и объем производимой продукции в 35 периоде; определить коэффициенты в производственной функции Кобба-Дугласа на основании ретроспективных данных.

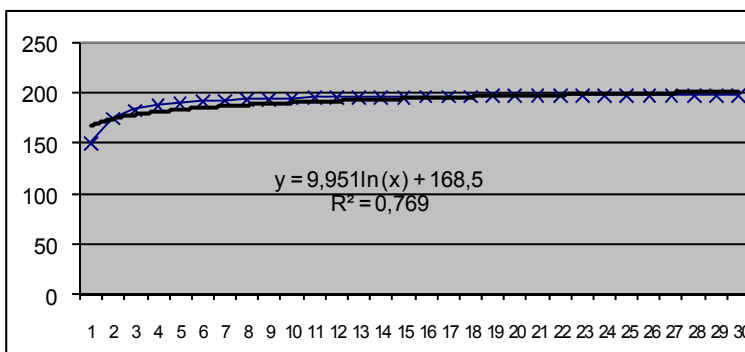
$i$	$V^+$	$V^-$	$M^+$	$M^-$	$K$	$Y$
0	22,1	20	2	10	150	102,383384
1	20	18,09675	2,210342	13,9	175	108,894638
2	18,1	16,37462	2,442806	17,6	183,3333	110,939324
3	16,4	14,81636	2,699718	21,1	187,5	111,940316
4	14,9	13,4064	2,983649	24,4	190	112,534063
5	13,6	12,13061	3,297443	27,5	191,6667	112,926858
6	12,5	10,97623	3,644238	30,4	192,8571	113,205758
7	11,6	9,931706	4,027505	33,1	193,75	113,413864
8	10,9	8,986579	4,451082	35,6	194,4444	113,574962
9	10,4	8,131393	4,919206	37,9	195	113,703253
10	10,1	7,357589	5,436564	40	195,4545	113,807747
11	10	6,657422	6,008332	41,9	195,8333	113,894435
12	10,1	6,023884	6,640234	43,6	196,1538	113,967461
13	10,4	5,450636	7,338593	45,1	196,4286	114,029784
14	10,9	4,931939	8,1104	46,4	196,6667	114,083575
15	11,6	4,462603	8,963378	47,5	196,875	114,130466
16	12,5	4,03793	9,906065	48,4	197,0588	114,171707
17	13,6	3,65367	10,94789	49,1	197,2222	114,208276
18	14,9	3,305978	12,09929	49,6	197,3684	114,240948
19	16,4	2,991372	13,37179	49,9	197,5	114,270348
20	18,1	2,706706	14,77811	50	197,619	114,296988
21	20	2,449129	16,33234	49,9	197,7273	114,321288
22	22,1	2,216063	18,05003	49,6	197,8261	114,343603
23	24,4	2,005177	19,94836	49,1	197,9167	114,364233
24	26,9	1,814359	22,04635	48,4	198	114,383434
25	29,6	1,6417	24,36499	47,5	198,0769	114,401428
26	32,5	1,485472	26,92748	46,4	198,1481	114,41841
27	35,6	1,34411	29,75946	45,1	198,2143	114,434551
28	38,9	1,216201	32,88929	43,6	198,2759	114,450006
29	42,4	1,100464	36,34829	41,9	198,3333	114,464912

1. Рассчитаем численность населения  $L_t$  за периоды  $t=0, 1, \dots, 29$ , построим график и функцию тренда.

2, 4. Построим график и функцию тренда численности населения.



3, 5. Проведем аналогичные действия для объема основного капитала.



6. На основе полученной функции  $L(35) = L_{35} = 499820,8$ .

7. Аналогично п.6 вычислить  $K(35) = K_{30} = 203,96$ .

8. Выявим зависимость объема производимой продукции от численности населения и объема основного капитала:

$$Y = AK^{\alpha}L^{\beta}$$

Вспользуемся, например, методом наименьших квадратов.

Для определения параметров функции Кобба-Дугласа, решить систему уравнений:

$$\left\{ \begin{array}{l} n \ln A + \alpha \sum_{i=0}^{29} \ln K_i + \beta \sum_{i=0}^{29} \ln L_i = \sum_{i=0}^{29} \ln Q_i, \\ \ln A \sum_{i=0}^{29} \ln K_i + \alpha \sum_{i=0}^{29} \ln^2 K_i + \beta \sum_{i=0}^{29} \ln L_i \ln K_i = \sum_{i=0}^{29} \ln Q_i \ln K_i, \\ \ln A \sum_{i=0}^{29} \ln L_i + \alpha \sum_{i=0}^{29} \ln K_i \ln L_i + \beta \sum_{i=0}^{29} \ln^2 L_i = \sum_{i=0}^{29} \ln Q_i \ln L_i. \end{array} \right.$$

Для этого составим матрицу коэффициентов:

$$\begin{array}{ccc} 30 & 162,9035 & 406,7734 \\ 162,9035 & 856,1985 & 2137,574 \\ 406,7734 & 2137,574 & 5337,568 \end{array}$$

и столбец свободных членов:

$$\begin{array}{c} 146,5159 \\ 769,9932 \\ 1922,541 \end{array}$$

Вычислим обратную матрицу

$$\begin{array}{ccc} -1 & 0,00017 & 0,076142 \\ 0,00017 & 6,701413 & -2,68378 \\ 0,076142 & -2,68378 & 1,069175 \end{array}$$

и умножим ее на столбец свободных членов:

$$\begin{array}{c} 0 \\ 0,4 \\ 0,2 \end{array}$$

Значит,  $A^* = 1$ ,  $\alpha^* = 0,4$ ,  $\beta^* = 0,2$ . Тогда  $Y = K^{0,4} L^{0,2}$ .

9. Подставим полученные значения  $A^*$ ,  $\alpha^*$ ,  $\beta^*$ ,  $L_{35}$ ,  $K_{35}$  в функцию  $Y_{35} = 203,96^{0,4} 499820,8^{0,2} = 115,77$ .

## ЛАБОРАТОРНАЯ РАБОТА № 4

### АНАЛИЗ ВРЕМЕННОГО РЯДА

#### Содержание

Дан временной ряд. Требуется проверить гипотезу о неизменности среднего значения временного ряда и выделить его неслучайные составляющие.

#### Порядок работы

1. Построение графика
2. Проверка гипотеза с помощью критерия, основанного на медиане
3. Проверка гипотезы с помощью критерия «восходящих» и «нисходящих» серий
4. Выделение неслучайной составляющей временного ряда

#### Описание

1. Введем начальные данные
2. Построим график по таблице начальных данных
3. Сформулируем гипотезу  $H_0: Mx(t) = a = const.$
4. Вычислим медиану

$$x_{med}^{(n)} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{если } n \text{ нечетно,} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right), & \text{если } n \text{ четно.} \end{cases}$$

5. Составим вспомогательную таблицу для критерия, основанного на медиане.

$$y_i = \begin{cases} -, & \text{если } x(i) < x_{med}^{(n)}, \\ +, & \text{если } x(i) > x_{med}^{(n)}. \end{cases}$$

6. Определим параметры  $\nu(n)$  и  $\tau(n)$ .

$\nu(n)$  – общее число «серий»

$\tau(n)$  – протяженность самой длинной «серии»

«Серия» – это последовательность подряд идущих плюсов и минусов.

7. Проверим неравенства:

$$\nu(n) > \left[ \frac{1}{2}(n+2) - 1,96\sqrt{n-1} \right],$$

$$\tau(n) < [1,43 \ln(n+1)].$$

8. Сделаем вывод о принятии/непринятии гипотезы  $H_0$ .

Если одно из неравенств не выполняется, то гипотеза  $H_0$  отвергается с вероятностью ошибки  $\alpha$ ,  $0,05 < \alpha < 0,0975$ .

9. Составим вспомогательную таблицу для критерия «восходящих» и «нисходящих» серий.

$$z_i = \begin{cases} -, & \text{если } x(i+1) - x(i) < 0, \\ +, & \text{если } x(i+1) - x(i) > 0. \end{cases}$$

10. Определим параметры  $\nu(n)$  и  $\tau(n)$ .

11. Проверим неравенства

$$\nu(n) > \left[ \frac{1}{3}(2n-1) - 1,96\sqrt{\frac{16n-29}{90}} \right],$$

$$\tau(n) < \tau_0(n).$$

Здесь  $\tau_0(n)$  находится из таблицы:

$n$	$n \leq 26$	$26 < n \leq 153$	$153 < n \leq 1170$
$\tau_0(n)$	$\tau_0=5$	$\tau_0=6$	$\tau_0=7$

12. Сделаем вывод о принятии/непринятии гипотезы  $H_0$ .

Если хотя бы одно из неравенств п.11 нарушено, то гипотезу  $H_0$  следует отвергнуть (и, соответственно, признать, что в разложении

анализируемого временного ряда присутствует неслучайная, зависящая от времени  $t$  компонента).

13. Введем новую переменную, описывающую время

$$t' = t - (m+1),$$

тогда  $\sum_{t'=-m}^{t'=m} (t')^k = 0$  для нечетных  $k$ .

14. Для линейной функции тренда  $f_{\text{ТР}}(t) = \theta_0 + \theta_1 t'$  вычислим коэффициенты

$$\theta_0 = \frac{1}{2m+1} \sum_{t'=-m}^{t'=m} x(t'),$$

$$\theta_1 = \frac{\sum_{t'=-m}^{t'=m} t'x(t') - \theta_0 \sum_{t'=-m}^{t'=m} t'}{\sum_{t'=-m}^{t'=m} (t')^2}.$$

15. Для квадратичной функции тренда  $f_{\text{ТР}}(t) = \theta_0 + \theta_1 t' + \theta_2 (t')^2$  вычислим коэффициенты

$$\theta_0 = \frac{\sum_{t'=-m}^{t'=m} (t')^4 \sum_{t'=-m}^{t'=m} x(t') - \sum_{t'=-m}^{t'=m} (t')^2 \sum_{t'=-m}^{t'=m} (t')^2 x(t')}{(2m+1) \sum_{t'=-m}^{t'=m} (t')^4 - \left( \sum_{t'=-m}^{t'=m} (t')^2 \right)^2},$$

$$\theta_1 = \frac{\sum_{t'=-m}^{t'=m} t'x(t')}{\sum_{t'=-m}^{t'=m} (t')^2},$$

$$\theta_2 = \frac{(2m+1) \sum_{t'=-m}^{t'=m} (t')^2 x(t') - \sum_{t'=-m}^{t'=m} (t')^2 \sum_{t'=-m}^{t'=m} x(t')}{(2m+1) \sum_{t'=-m}^{t'=m} (t')^4 - \left( \sum_{t'=-m}^{t'=m} (t')^2 \right)^2}.$$

16. Вычисляем среднее квадратическое отклонение для каждой линии тренда.

17. Для «лучшей» линии тренда строим спектральную плотность.

Для квадратичной линии тренда:

$$p(\tilde{\omega}) = 2(1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2) \cos(2\pi\tilde{\omega}) - 2\theta_2 \cos(4\pi\tilde{\omega})),$$

$$0 \leq \tilde{\omega} \leq \frac{1}{2}.$$

Для линейной линии тренда:

$$p(\tilde{\omega}) = 2(1 + \theta_1 - 2\theta_1 \cos(2\pi\tilde{\omega})),$$

$$0 \leq \tilde{\omega} \leq \frac{1}{2}.$$

$\tilde{\omega}$  подбираем таким образом, чтобы среднее квадратическое отклонение (СКО) было наименьшим. В отчете по лабораторной работе представить данные по СКО для анализируемых  $\tilde{\omega}$ .

18. Построить график полученной спектральной плотности совместно с графиком начальных данных.



**ЛАБОРАТОРНАЯ РАБОТА № 5**  
**АНАЛИЗ ВРЕМЕННЫХ РЯДОВ**  
**С СЕЗОННЫМ КОМПОНЕНТОМ**

**Содержание лабораторной работы**

Дан временной ряд  $x(t)$ ,  $t = 1, 2, \dots, N$ . Требуется определить функцию тренда  $f_{TP}(t)$  и сезонную составляющую  $S(t)$  с периодом  $T$ , затем вычислить прогнозное значение.

**Порядок работы**

1. Определить «наилучшую» функцию тренда  $f_{TP}(t)$ .
2. Определить функцию, описывающую сезонные колебания  $S(t)$ .
3. Вычислить прогнозные значения  $x(t+k)$ ,  $k=1, 2, \dots, T$ .
4. Построить итоговый график.

**Описание**

1. Подобрать функцию, описывающую линию тренда, с наименьшим отклонением.

Критерием отбора может служить среднее квадратическое отклонение.

2. Выделить сезонные колебания

$$S(t) = x(t) - f_{TP}(t).$$

3. Определить функцию, описывающую сезонную компоненту

$$S(t) = A_3 \sin(\omega t \Delta + \psi) = A_4 \sin(\omega t \Delta) + A_5 \cos(\omega t \Delta),$$

где  $A_3$  – амплитуда гармоники,  $\omega = \frac{2\pi}{T}$  – её частота,  $T$  – период,

$\psi$  – начальная фаза,  $A_4 = A_3 \cos \psi$ ,  $A_5 = A_3 \sin \psi$ .

4. Составить функцию  $S(t)$ .
5. Составить результирующую функцию аппроксимации

$$x(t) = f_{TP}(t) + S(t).$$

6. Вычислить прогнозные значения на  $T$  тактов.

7. Построить графики по начальным данным временного ряда  $x(t)$  и для аппроксимирующей функции  $x(t)$ .

8. Определить характеристики точности построенной модели и точности прогнозирования.

Точность модели оценивается с помощью коэффициента детерминации:

$$R^2 = \frac{\sum_{k=0}^N (Y_k^o - M[Y_k])^2}{\sum_{k=0}^N (Y_k - M[Y_k])^2} = 1 - \frac{\sum_{k=0}^N (Y_k^o - Y_k)^2}{\sum_{k=0}^N (Y_k - M[Y_k])^2},$$

где  $Y_k^o$  – модельные значения (вычисленные по модели),  $Y_k$  – реальные (фактические) значения.

Точность прогнозирования оценивается либо с помощью средней относительной ошибки прогноза (или среднеабсолютной процентной), так называемой MAPE-оценки:

$$\bar{\gamma} = \frac{1}{l} \sum_{k=1}^l \frac{|Y_k - Y_k^*|}{Y_k} \cdot 100\%,$$

здесь  $Y_k^*$  – прогнозное значение показателя на  $k$ -е наблюдение;

либо с помощью коэффициента несоответствия (коэффициенты Тейла):

$$K_{T1} = \sqrt{\frac{\sum_{k=1}^l (Y_k - Y_k^*)^2}{\sum_{k=1}^l Y_k^2 + \sum_{k=1}^l (Y_k^*)^2}}.$$

**ЛАБОРАТОРНАЯ РАБОТА № 6**  
**АНАЛИЗ ВРЕМЕННЫХ РЯДОВ**  
**С СЕЗОННЫМ КОМПОНЕНТОМ ПО ГАРМОНИКАМ**

**Содержание лабораторной работы**

Дан временной ряд  $x(t)$ ,  $t = 1, 2, \dots, N$ . Требуется определить функцию тренда  $f_{\text{тр}}(t)$  и сезонную составляющую  $\varphi(t)$  с периодом  $T$ , затем вычислить прогнозное значение.

**Порядок работы**

1. Определить «наилучшую» функцию тренда  $f_{\text{тр}}(t)$ .
2. Определить функцию, описывающую сезонные колебания  $\varphi(t)$ .
3. Вычислить прогнозные значения  $x(t+k)$ ,  $k=1, 2, \dots, T$ .
4. Построить итоговый график.

**Описание**

1. Подобрать функцию, описывающую линию тренда, с наименьшим отклонением.

Критерием отбора может случить среднее квадратическое отклонение.

2. Выделить сезонные колебания

$$\varphi(t) = x(t) - f_{\text{тр}}(t).$$

3. Вычислить коэффициенты

$$a(\omega) = \frac{1}{\sqrt{\pi N}} \sum_{t=1}^N \varphi(t) \cos(\omega t),$$

$$b(\omega) = \frac{1}{\sqrt{\pi N}} \sum_{t=1}^N \varphi(t) \sin(\omega t).$$

Здесь описывается гармоническая функция с периодом  $T = \frac{2\pi}{\omega}$ .

4. Составить функцию

$$\varphi(t) = \lambda \left( \frac{a(0)}{2} + a(\omega) \cos(\omega t) + b(\omega) \sin(\omega t) \right).$$

5. Подобрать значение корректирующего множителя  $\lambda$  таким, чтобы среднее квадратическое отклонение было наименьшим.

6. Составить результирующую функцию аппроксимации

$$x(t) = f_{TP}(t) + \varphi(t).$$

7. Вычислить прогнозные значения на  $T$  тактов.

8. Построить графики по начальным данным временного ряда  $x(t)$  и для аппроксимирующей функции  $x(t)$ .

**ЛАБОРАТОРНАЯ РАБОТА № 7**  
**ПРИМЕНЕНИЕ АДАПТИВНЫХ**  
**МЕТОДОВ ПРОГНОЗИРОВАНИЯ: МЕТОД ХОЛЬТА**

**Содержание лабораторной работы**

Дан временной ряд  $x(t)$ ,  $t = 1, \dots, N$ . Требуется с помощью метода Хольта аппроксимировать временной ряд и вычислить прогнозное значение.

**Порядок работы**

1. Аппроксимировать анализируемый временного ряда  $x(t)$  методом Хольта.
2. Вычислить прогнозное значение  $x(t+10)$ .
3. Построить итоговый график.

**Описание**

1. За основу можно взять первые  $k$  значений временного ряда  $x(t)$ ,  $t = 1, 2, \dots, k$ . С помощью метода наименьших квадратов найдем значения для  $a_0(1; \lambda_1; \lambda_2)$  и  $a_1(1; \lambda_1; \lambda_2)$ :

$$\begin{cases} 10a_0(1) + a_1(1) \sum_{i=1}^k i = \sum_{i=1}^k x(i), \\ a_0(1) \sum_{i=1}^k i + a_1(1) \sum_{i=1}^k i^2 = \sum_{i=1}^k ix_i. \end{cases}$$

2. Поставим полученные значения коэффициентов  $a_0(1; \lambda_1; \lambda_2)$  и  $a_1(1; \lambda_1; \lambda_2)$  в формулу пересчета:

$$a_0(t+1; \lambda_1; \lambda_2) = \lambda_1 x(t) + (1 - \lambda_1)(a_0(t; \lambda_1, \lambda_2) + a_1(t; \lambda_1, \lambda_2)),$$

$$a_1(t+1; \lambda_1; \lambda_2) = \lambda_2 (a_0(t+1; \lambda_1, \lambda_2) - a_1(t; \lambda_1, \lambda_2)) + (1 - \lambda_2) a_1(t; \lambda_1, \lambda_2).$$

То есть

$$a_0(2; \lambda_1; \lambda_2) = \lambda_1 x(1) + (1 - \lambda_1)(a_0(1; \lambda_1, \lambda_2) + a_1(1; \lambda_1, \lambda_2)),$$
$$a_1(2; \lambda_1; \lambda_2) = \lambda_2(a_0(2; \lambda_1, \lambda_2) - a_1(1; \lambda_1, \lambda_2)) + (1 - \lambda_2)a_1(1; \lambda_1, \lambda_2).$$

Получим выражения с двумя неизвестными  $\lambda_1$  и  $\lambda_2$ . Подставим эти выражения в формулу метода Хольта:

$$x(t; l) = a_0(t; \lambda_1; \lambda_2) + l a_1(t; \lambda_1; \lambda_2), \text{ где } t = 1, 2, \dots, N - l.$$

То есть

$$x(2; l) = a_0(2; \lambda_1; \lambda_2) + l a_1(2; \lambda_1; \lambda_2) =$$
$$= (1 + l\lambda_2)(\lambda_1 x(1) + (1 - \lambda_1)(a_0(1; \lambda_1, \lambda_2) + a_1(1; \lambda_1, \lambda_2))) +$$
$$+ (1 - 2\lambda_2)l a_1(1; \lambda_1, \lambda_2).$$

Подставляя значения  $l$ , например,  $l = 1, 2$ , найдем неизвестные параметры  $\lambda_1$  и  $\lambda_2$ .

3. Используя подобранные данные, вычислить прогнозное значение  $x(N; 10) = a_0(N; \lambda_1; \lambda_2) + 10 a_1(N; \lambda_1; \lambda_2)$  методом Хольта.

**ЛАБОРАТОРНАЯ РАБОТА № 8**  
**ПРОГНОЗИРОВАНИЕ ВРЕМЕННОГО РЯДА**  
**С СЕЗОННЫМ КОМПОНЕНТОМ**

**Содержание лабораторной работы**

Дан временной ряд  $x(t)$ ,  $t = 1, \dots, N$ . Требуется аппроксимировать временной ряд и вычислить прогнозные значения.

**Порядок работы**

1. Аппроксимировать анализируемый временного ряда  $x(t)$ .
2. Вычислить прогнозное значение на число временных тактов, содержащихся в полном сезонном цикле.
3. Построить итоговый график.

**Описание**

1. Определить функцию тренда.

Для определения функции тренда  $x(t; l) = f_{TP}(t; l) = f_{TP}(t + l)$  используем метод Хольта:  $x(t; l) = a_0(t) + l a_1(t)$ ,

$$a_0(t+1; \lambda_1; \lambda_2) = \lambda_1 x(t) + (1 - \lambda_1)(a_0(t; \lambda_1, \lambda_2) + a_1(t; \lambda_1, \lambda_2)),$$

$$a_1(t+1; \lambda_1; \lambda_2) = \lambda_2 (a_0(t+1; \lambda_1, \lambda_2) - a_1(t; \lambda_1, \lambda_2)) + (1 - \lambda_2) a_1(t; \lambda_1, \lambda_2).$$

2. Вычислить коэффициенты сезонности  $\omega(k)$   $k = 1, 2, \dots, 12$ .

$$x(t; l) = f_{TP}(t; l) \omega\left(\left\{\frac{t+l}{T}\right\}T\right),$$

здесь  $\{\cdot\}$  – дробная часть.

Коэффициенты сезонности определяем по формуле

$$\omega(k) = \frac{T}{N} \sum_{i=1}^{N/T} \frac{x(k+iT)}{f_{TP}(k+iT)}, \quad k = 1, 2, \dots, 12.$$

Подставим полученные значения в формулу для вычисления прогнозного значения:

$$x(t; l) = f_{TP}(t; l) \omega\left(\left\{\frac{t+l}{T}\right\}T\right),$$

где  $t=N$ ,  $l=1, 2, \dots, T$ .

## ЛАБОРАТОРНАЯ РАБОТА № 9

### ПОДБОР ПОРЯДКА АППРОКСИМИРУЮЩЕГО ПОЛИНОМА С ПОМОЩЬЮ МЕТОДА ПОСЛЕДОВАТЕЛЬНЫХ РАЗНОСТЕЙ

#### Содержание лабораторной работы

Дан временной ряд  $x(t)$ ,  $t = 1, \dots, N$ . Требуется определить порядок аппроксимирующего полинома с помощью метода последовательных разностей.

#### Порядок работы

1. Вычислить конечные разности  $\Delta^p x(t)$  ( $t = 2, \dots, N - p + 1$ ) и подсчитать значение  $\sigma^2(p)$ .
2. Определить порядок  $k$  аппроксимирующего полинома  $P_k(t)$ .
3. Составить аппроксимирующие полиномы  $P_{k-2}(t)$ ,  $P_{k-1}(t)$ ,  $P_k(t)$ ,  $P_{k+1}(t)$ ,  $P_{k+2}(t)$ .
4. Вычислить для каждого из полиномов достоверность аппроксимации  $R^2$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma$ .
5. Построить графики полиномов и график данных начального временного ряда.

#### Описание

1. Вычислить последовательные разности 1-го порядка:

$$\Delta x(t) = x(t) - x(t-1), \quad t = 2, \dots, N.$$

2. Определить  $\sigma^2(1)$ :

$$\sigma^2(1) = \frac{\sum_{t=1}^{N-1} (\Delta x(t))^2}{(N-1)C_2^1}.$$

3. Вычислить последовательные разности 2-го порядка:

$$\Delta^2 x(t) = \Delta x(t) - \Delta x(t-1), \quad t = 2, \dots, N-1.$$

4. Определить  $\sigma^2(2)$ :



$$\sigma^2(2) = \frac{\sum_{t=1}^{N-2} (\Delta^2 x(t))^2}{(N-2)C_4^2}.$$

....

2p-1. Вычислить последовательные разности p-го порядка:

$$\Delta^p x(t) = \Delta^{p-1} x(t) - \Delta^{p-1} x(t-1), \quad t = 2, \dots, N-p+1.$$

2p. Определить  $\sigma^2(p)$ :

$$\sigma^2(p) = \frac{\sum_{t=1}^{N-p} (\Delta^p x(t))^2}{(N-p)C_{2p}^p}.$$

Вычисления продолжать до тех пор, пока  $\sigma^2(p)$  не перестанет убывать. Значит, порядок аппроксимирующего полинома должен быть  $k = p - 1$ .

2p+1. Составить аппроксимирующий полином степени  $k - 2$ :

$$P_{k-2}(t) = \sum_{i=0}^{k-2} a_i t^i.$$

2p+2. Вычислить для полинома  $P_{k-2}(t)$  достоверность аппроксимации  $R_{k-2}^2$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma_{k-2}$ .

2p+3. Составить аппроксимирующий полином степени  $k - 1$ :

$$P_{k-1}(t) = \sum_{i=0}^{k-1} b_i t^i.$$

2p+4. Вычислить для полинома  $P_{k-1}(t)$  достоверность аппроксимации  $R_{k-1}^2$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma_{k-1}$ .

2p+ 5. Составить аппроксимирующий полином степени  $k$ :

$$P_k(t) = \sum_{i=0}^k c_i t^i .$$

2p+ 6. Вычислить для полинома  $P_k(t)$  достоверность аппроксимации  $R^2_k$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma_k$ .

2p+ 7. Составить аппроксимирующий полином степени  $k + 1$ :

$$P_{k+1}(t) = \sum_{i=0}^{k+1} d_i t^i .$$

2p+ 8. Вычислить для полинома  $P_{k+1}(t)$  достоверность аппроксимации  $R^2_{k+1}$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma_{k+1}$ .

2p+ 9. Составить аппроксимирующий полином степени  $k + 2$ :

$$P_{k+2}(t) = \sum_{i=0}^{k+2} e_i t^i .$$

2p+ 10. Вычислить для полинома  $P_{k+2}(t)$  достоверность аппроксимации  $R^2_{k+2}$  (встроенный параметр аппроксимации Excel), либо среднее квадратическое отклонение  $\sigma_{k+2}$ .

2p+ 11. Построить графики полиномов и график данных начального временного ряда.

*Учебное издание*

**МЕТОДЫ СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО  
ПРОГНОЗИРОВАНИЯ**

*Методические указания  
к лабораторным работам*

Составители: ***Ростова Елена Павловна,  
Кореева Екатерина Борисовна***

Редактор И.И. Спиридонова  
Доверстка И.И. Спиридонова

Подписано в печать 20.06.2011 г. Формат 60 × 84 1/16.  
Бумага офсетная. Печать офсетная. Печ.л. 2,25.  
Тираж 100 экз. Заказ \_\_\_\_ . Арт. С-С2/2011.

Самарский государственный аэрокосмический университет.  
443086, Самара, Московское шоссе, 34.

---

Изд-во Самарского государственного аэрокосмического университета.  
443086, Самара, Московское шоссе, 34.

