

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ОБРАЗОВАНИЯ  
«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ НЕФТЕГАЗОВЫЙ УНИВЕРСИТЕТ»  
ТОБОЛЬСКИЙ ИНДУСТРИАЛЬНЫЙ ИНСТИТУТ (ФИЛИАЛ)

Е.С. Чижикова

**ИСПОЛЬЗОВАНИЕ СТАТИСТИЧЕСКИХ МЕТОДОВ  
ДЛЯ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ**

*Методические указания к лабораторным работам по дисциплине  
«Моделирование систем и процессов»*

Тобольск  
2018

Использование статистических методов для построения математических моделей: методические указания к лабораторным работам / сост. Е.С. Чижикова; Филиал ТИУ в г. Тобольске, 2018г.

Методические указания рассмотрены и рекомендованы к изданию на заседании кафедры естественно научных и гуманитарных наук «31» августа 2018 года, протокол № 1.

### **Аннотация**

Методические указания содержат рекомендации по выполнению слушателями лабораторных работ по дисциплине «Моделирование систем и процессов» программы профессиональной переподготовки «Автоматизация технологических процессов и производств».

Методические указания предназначены для выполнения четырех лабораторных работ, направленных на получение практических навыков использования статистических методов для построения математических моделей. В них описана процедура выполнения лабораторных работ, последовательность действий и порядок обработки результатов.

---

## СОДЕРЖАНИЕ

### ВВЕДЕНИЕ

#### ЛАБОРАТОРНАЯ РАБОТА № 1 Тема: ИЗУЧЕНИЕ СРЕДЫ MATLAB 6.5

##### 1 Общие сведения о среде MATLAB

##### 2 Работа с векторами и матрицами

##### 3 Работа с числовыми последовательностями. Построение графиков функций

##### 4 Содержание отчета

#### ЛАБОРАТОРНАЯ РАБОТА № 2 Тема: ПОЛУЧЕНИЕ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ ПАРНОЙ РЕГРЕССИИ И ВЫЧИСЛЕНИЕ ПОГРЕШНОСТИ

##### 1 Нахождение суммы элементов вектора

##### 2 Решение систем алгебраических уравнений

##### 3 Применение метода наименьших квадратов для получения коэффициентов парной линейной регрессии

###### 3.1 Краткие сведения из теории

###### 3.2 Задание

###### 3.3 Варианты исходных данных

###### 3.4 Построение графика линейной функции по данным наблюдения, полученной МНК

###### 3.5 Оценка погрешности парной линейной регрессии

##### 4 Содержание отчета

#### ЛАБОРАТОРНАЯ РАБОТА № 3 Тема: НЕЛИНЕЙНАЯ РЕГРЕССИЯ. ПОДГОНКА КРИВЫХ И ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ

##### 1 Краткие сведения из теории

##### 2. Порядок выполнения работы

##### 3 Варианты исходных данных

##### 4 Содержание отчета

ЛАБОРАТОРНАЯ РАБОТА № 4 Тема: ПОЛУЧЕНИЕ КОЭФФИЦИЕНТОВ МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ И ВЫЧИСЛЕНИЕ ПОГРЕШНОСТИ

1 Краткие сведения из теории

2. Задание

3 Варианты матрицы  $y$

4 Построение трехмерных графиков

5 Порядок выполнения работы

6 Содержание отчета

ЛАБОРАТОРНАЯ РАБОТА № 5 Тема: ФАКТОРНЫЙ АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

1 Краткие сведения из теории

2 Задание к лабораторной работе

3 Порядок выполнения работы

4 Содержание отчета

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

---

## ВВЕДЕНИЕ

Одной из основных задач современной науки и техники является разработка методов исследования динамики функционирования сложных систем и процессов. К классу таких систем относятся средства управления крупными производственными, энергетическими и техническими комплексами, создаваемые на базе автоматики, телемеханики, электроники и вычислительной техники.

При проектировании и создании сложных систем, их испытаниях и эксплуатации возникают задачи, требующие знания количественных и качественных закономерностей, свойственных этим системам.

Средством изучения процессов, протекающих в различных системах и устройствах без их физического создания, является построение и исследование математической модели этих систем и процессов. Их применение позволяет сделать из электронной вычислительной машины экспериментальную установку, что ведет к удешевлению и ускорению этапа проектирования систем и устройств.

Одной из основных задачи математического моделирования является исследование и прогнозирование поведения объекта моделирования при различных внешних воздействиях и изменениях его параметров.

Данные методические указания предназначены для выполнения четырех лабораторных работ, направленных на получение практических навыков использования статистических методов для построения математических моделей.

### ЛАБОРАТОРНАЯ РАБОТА № 1 ТЕМА: ИЗУЧЕНИЕ СРЕДЫ MATLAB 6.5

**Цель работы:** Получение навыков выполнения элементарных вычислений и построения графиков в среде MATLAB 6.5.

MATLAB – одна из систем автоматизации математических расчетов, основанная на применении матричных операций.

## Порядок выполнения работы.

- 1 Изучите общие сведения о среде MATLAB 6.5.
- 2 Войдите в среду MATLAB 6.5.
- 3 Наберите приведенные в методических указаниях примеры (команды, которые необходимо ввести, выделены жирным шрифтом) и изучите их.
- 4 Выполните задания по каждому пункту.

### 1. Общие сведения о среде MATLAB

Основное окно системы, позволяющее проводить вычисления в командном режиме, показано на рис. 1.

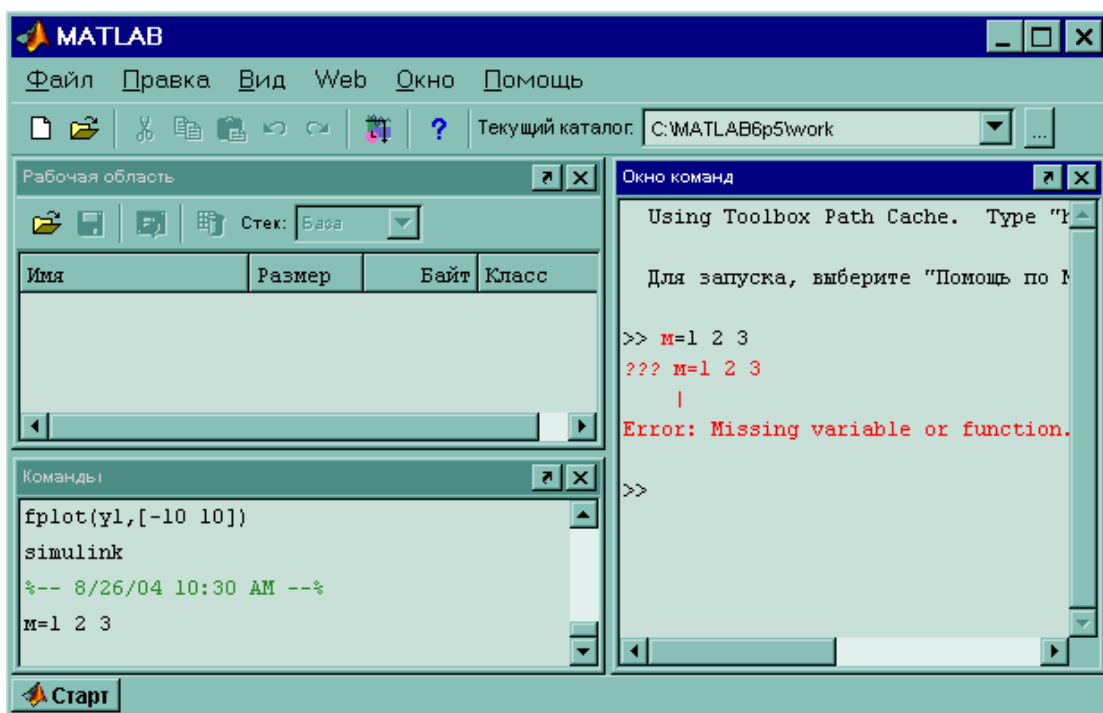


Рис 1 Вид окна MATLAB в командном режиме

Окно MATLAB включает в себя три основные зоны – «окно команд», «рабочая область» и «команды».

В поле «окно команд» после знака «>>» вводятся команды. После нажатия кнопки «Enter» команда выполняется и результаты отображаются ниже в этом же окне.

В поле «команды» отражается «история» сеанса работы – все команды, введенные в процессе сеанса, а в «рабочей области» – все вводимые переменные.

В простейших вычислениях могут участвовать переменные, константы и функции (как стандартные, так и вводимые пользователем).

*Константа* представляет собой непосредственно заданное число, целое или дробное, действительное или комплексное.

Константы могут быть представлены в одной из следующих форм:

2

-5

2.0007

-12.234e-5

2+3i

12.5-1.03e5i

*Переменная* – имеющий имя объект, способный хранить значение. Значением переменной может быть как число, так и целый массив. Переменным можно задавать значения с помощью операции присваивания, имеющей следующий формат:

Имя переменной = выражение

Например:

$V = 5$

$S = [1\ 2\ 3\ 4]$

$Mat = [1\ 4\ 5; 2\ 3\ 2]$

*Имя переменной* может содержать символы латинского алфавита, цифры и знак «\_», но должно начинаться с буквы.

*Выражение* может состоять из констант, переменных, функций и знаков операций.

*Знаки операций* – «+» – сумма, «-» – вычитание, «\*» – умножение, «/» – деление, «^» – возведение в степень и т.д.

*Функции* – это имеющие имена объекты, выполняющие преобразование своих аргументов и возвращающие результаты этих преобразований. Аргументы функции указываются в скобках после имени функции.

Например:

$\sin(x)$  – синус переменной  $x$ ,

$\cos(1)$  – косинус 1,

$\text{sqrt}(2+x^2)$  – корень квадратный из выражения  $2+x^2$ .

Примеры арифметических выражений:

$2+3$

$4.5+\sin(v)$

$\text{sqrt}(x+3)/2$

Если значение выражения явно не присваивается никакой переменной, значение результата получает системная переменная *ans*.

**Пример** вычислений (после знака «>>» вводимая команда, в следующей строке – ответ системы):

```
>>s=6
```

```
s=6
```

```
>>s=s/2
```

```
s=3
```

```
>>3+2
```

```
ans=5
```

```
>>s*ans
```

```
ans=15
```

### **Варианты заданий к п. 1.**

Вычислите значения выражений:

Вариант 1.  $\frac{\sqrt{x-a^2} + \sin(2x-3)}{\cos^2(x+a)}$ ;  $x = 12; a = 3$ .



Вариант 2.  $\frac{2\sqrt{x+1} + \sin(2x)}{\sqrt{a+5} + \cos^2(x+3a)}$ ;  $x=10; a=13$ .

Вариант 3.  $\frac{\cos(\sqrt{x+3a^3}) + \sin^2(2x-3)}{\cos^a(x)}$ ;  $x=4; a=2$ .

Вариант 4.  $\frac{\sqrt{\sin(2x-3)+1}}{2 + \cos^x(x+a)}$ ;  $x=5; a=-3$ .

Вариант 5.  $\frac{\sqrt{\sin(2a+x^2-3)}-1}{2 + \sqrt{x/2+6}}$ ;  $x=3; a=-3$ .

Вариант 6.  $\frac{\sqrt{\sin(2x-3)+1}}{2 + \cos^x(x+a)}$ ;  $x=5; a=-3$ .

## 2 Работа с векторами и матрицами

Для задания значений векторам и матрицам их элементы необходимо перечислить в квадратных скобках *через пробел*. Например, вектор из четырех элементов может быть задан следующим образом:

```
>>v= [1 2 3 4]
```

При задании матрицы элементы каждой строки должны отделяться «;». Например, матрица, состоящая из двух строк и трех столбцов должна быть задана следующим образом:

```
>>matr=[1 1 1;3 5 6]
```

При выполнении арифметических операций, в которых участвуют матрицы и константы, эти операции производятся над всеми элементами.

Например, в результате команды

```
>>matr=matr.*2
```

все элементы матрицы будут увеличены вдвое.

Если имеется еще одна матричная переменная

```
>>n=[1 2 3; 4 5 6]
```

результатом выполнения команды

```
>>matr+n
```

будет матрица, являющаяся суммой этих матриц (это значение получит системная переменная ans).

Некоторые особенности при работе с матрицами имеют операции умножения и деления. Если перед знаком операции ставится точка, то выполняется поэлементное умножение и деление, в противном случае – умножение и деление по правилам умножения матриц.

### Пример

```
>>a=[1 2;3 4]
```

```
>>b=[5 6;7 8]
```

```
>>c=[1 1 1;2 2 2]
```

```
>>a.*b
```

```
ans=5 12
```

```
21 32
```

```
>>a*c
```

```
ans=5 5 5
```

```
11 11 11
```

Для обращения к отдельному элементу матрицы в круглых скобках указывается номер элемента (для векторов) или номера строки и столбца, в которых расположен элемент.

Примеры.

```
>>v(2)=1
```

```
v=1 1 3 4
```

```
>>matr(1,2)=10
```

```
matr = 1 10 1
```

```
3 5 6
```

## Задания к п. 2

Задайте следующие матрицы:

$$a = \begin{pmatrix} 1 & 4 & 6 & 7 \\ 2 & 3 & 11 & 9 \end{pmatrix} \quad b = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 4 & 4 \\ 2 & 1 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad c = \begin{pmatrix} 2 & 3 & 2 & 1 \\ 4 & 6 & 1 & 9 \end{pmatrix}$$

Вычислите:

- 1 Поэлементное произведение матриц  $a$  и  $c$ .
- 2 Произведение матриц  $a$  и  $b$ .
- 3 Получите матрицу, элементы которой равны квадратным корням элементов матрицы  $c$ .
- 4 Получите матрицу, элементы которой равны сумме элементов полученной матрицы и матрицы  $a$ .

## 3 Работа с числовыми последовательностями

### Построение графиков функций

Для создания числовых последовательностей используется оператор «:», имеющий следующий формат:

Начальное значение : шаг : конечное значение.

Например, оператор  $1 : 0.1 : 10$  создаст числовую последовательность от 1 до 10 с шагом 0.1.

Для вычисления значений функции на заданном интервале необходимо применить эту функцию к полученной последовательности.

### Пример

Вычислить значение функции  $y = \sin^2 x - \cos(x)/2$  на интервале от 0 до  $2\pi$  с шагом  $\pi/10$ . Последовательность команд должна быть следующей:

```
>>x=0:pi/10:2*pi
```

```
>>y=sin(x).^2-cos(x)/2
```

*Замечание.* Наличие точек перед знаками операций обязательно, так как значения должны быть вычислены для всех элементов последовательности.

### Построение графиков функций

Для построения графика необходимо выполнить следующие 2 операции:

- построить последовательность в заданном интервале с заданным шагом;
- использовать команду построения графика plot (функция).

Таким образом, для построения графика функции  $y=\sin^2(x)-\cos(x)/2$  нужно задать команду

```
>>plot(x, sin(x).^2-cos(x)/2)
```

Более гладкие графики получаются при использовании функции `fplot('f(x)',[xmin xmax])`, которая строит график функции  $f(x)$ , вид которой задается в апострофах в диапазоне  $[xmin\ xmax]$  без фиксированного шага.

### Пример

```
>> fplot('sin(x).^2-cos(x)/2',[0 2*pi])
```

### Варианты заданий к п. 3

1 Вычислить значение функции  $y=\sin^3(x)+\cos^2(x)$  на интервале от 0 до  $\pi$  с шагом  $\pi/10$ . Построить график.

2 Вычислить значение функции  $y=2\sin^2(x)-\sin(x)$  на интервале от  $-\pi$  до  $\pi$  с шагом  $\pi/12$ . Построить график.

3 Вычислить значение функции  $y=3\sin^2(x)\cos(x)$  на интервале от 0 до  $2\pi$  с шагом  $\pi/8$ . Построить график.

4 Вычислить значение функции  $y=7\sin^2(x)/\cos(x)$  на интервале от 0 до  $\pi/3$  с шагом  $\pi/32$ . Построить график.

5 Вычислить значение функции  $y=3\sin(x)-\cos^3(x)$  на интервале от  $-\pi$  до  $2\pi$  с шагом  $\pi/6$ . Построить график.

6 Вычислить значение функции  $y=\sin^4(x)-2\cos^2(x)$  на интервале от 0 до  $2\pi$  с шагом  $\pi/10$ . Построить график.

#### 4 Содержание отчета

Отчет должен содержать решения заданий по каждому пункту и сами задания.

### ЛАБОРАТОРНАЯ РАБОТА № 2 ТЕМА: ПОЛУЧЕНИЕ КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ ПАРНОЙ РЕГРЕССИИ И ВЫЧИСЛЕНИЕ ПОГРЕШНОСТИ

Цель работы. **Получение навыков построения линейной модели по данным экспериментов.**

#### **Порядок выполнения.**

1. Войдите в среду MATLAB 6.5 (см. лаб. работу № 1)
2. Изучите примеры и выполните задания из п.1 и п. 2.
3. Введите вектора  $x$  и  $y$  для своего варианта задания п. 3.
4. Вычислите коэффициенты системы уравнений для получения коэффициентов линейной парной регрессии.
5. Решите полученную систему уравнений.
6. Выведите на график точки корреляционного поля, полученную прямую и прямую, полученную системой MATLAB 6.5. Убедитесь, что прямые совпали, в противном случае найдите ошибку в вычислениях или исходных данных.
7. Проведите анализ достоверности модели, рассчитав коэффициент корреляции и стандартное квадратичное отклонение.

8. Сделайте вывод об адекватности полученной модели.

### 1 Нахождение суммы элементов вектора

Для нахождения суммы элементов вектора используется функция `sum(<переменная-вектор>)`.

Например, найти  $\sum_{j=1}^m (x_j^2 - y_j^2)$ , где  $x, y$  – векторные переменные:

```
>>sum(x.*x-y.*y)
```

**Задание.** Вычислить:

$$1 \quad \sum_{j=1}^5 x_j^2, x = (1,1,2,3,6)$$

$$2 \quad \sum_{j=1}^5 x_j * y_j, y = (2,3,2,1,2)$$

### 2 Решение систем алгебраических уравнений

В среде MATLAB можно выполнять символьные вычисления и преобразования, то есть формулировать задачу и получать решения не только в численной, но и в аналитической форме (например, символьное дифференцирование и интегрирование, упрощение математических выражений, получение решения систем уравнений в аналитической форме).

Для реализации символьных вычислений необходимо создать символьные переменные. В отличие от обычных переменных, символьным переменным не нужно присваивать никаких значений. Они получают свои значения в результате решения задачи в символьном виде. Объявляются символьные переменные следующим образом:

```
>>syms x y z
```

Для решения алгебраических уравнений и систем уравнений служит функция *solve*(уравнение 1, уравнение 2 ,..., уравнение N, переменная 1, переменная 2 ,..., переменная N), где переменные являются символьными переменными, относительно которых решается система уравнений.

Например, решить систему уравнений:

$$\begin{cases} 2x - 3y = 4; \\ y - 5x = 7. \end{cases}$$

```
>>syms x y
```

```
>>S=solve('2*x-3*y=4','y-5*x=7',x,y)
```

Ответ будет:

S=x: [1x1 sym]

y: [1x1 sym]

Для получения решения в аналитической форме необходимо обратиться к S.x, S.y:

```
>>S.x
```

```
>>S.y
```

Для вычисления числовых значений с командной строки ввести полученные в аналитической форме выражения.

**Задание.** Получить аналитическое и численное решения системы уравнений:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = 1; \\ a_{21}x + a_{22}y + a_{23}z = 3; \\ a_{32}y + a_{33}z = 1. \end{cases}$$

$a_{11}=2, a_{12}=-1, a_{13}=-3, a_{21}=-1, a_{22}=4, a_{23}=5, a_{32}=-1, a_{33}=2$

### 3 Применение метода наименьших квадратов для получения коэффициентов парной линейной регрессии

#### 3.1 Краткие сведения из теории

**Регрессионным анализом** называется метод построения модели на основе экспериментальных данных.

Применение регрессионного анализа для обработки результатов наблюдений позволяет получить *теоретическую зависимость (регрессию)* переменных состояния системы от ее параметров и входных воздействий.

Совокупность точек в пространстве, соответствующих экспериментальным данным (каждая из осей соответствует входному воздействию или параметру системы, одна из осей представляет собой отклик системы), называется **корреляционным полем**.

Различают *парную* и *множественную* регрессию.

Регрессия называется *парной*, если она представляет собой зависимость переменной  $y$  от единственной переменной  $x$ .

Регрессия называется *множественной*, если получена зависимость переменной  $y$  от нескольких переменных  $x_1, x_2, \dots, x_n$ .

*Линейная регрессия* ищется в виде линейной функции, *нелинейная* – в виде некоторой разновидности нелинейных функций.

Регрессионный анализ состоит из двух основных этапов.

1 Определяется вид зависимости (общий вид функции)  $y = f(x)$ , характер поведения которой близок к описанию поведения точек корреляционного поля.

2 Определяются параметры этой функции, при которых она наилучшим образом описывает (*аппроксимирует*) поведение точек корреляционного поля.

Для нахождения теоретической линии регрессии (параметров выбранной функции) по данным экспериментальных замеров применяется **метод наименьших квадратов**.



Суть его состоит в том, что отыскивается теоретическая линия регрессии  $y$  по  $x$ , такая, что сумма квадратов расстояний от этой линии до каждой экспериментальной точки с такими же  $x$ - координатами минимальна, то есть, функция регрессии  $y$  по  $x$  строится таким образом, чтобы соблюдался принцип наименьших квадратов:

$$S^2 = \sum_{j=1}^m (y_j - y'_j)^2 \rightarrow \min, \quad (2.1)$$

где  $j$  – порядковый номер точки в экспериментальном числовом ряду:

$y_j$  – экспериментальное значение  $y$  для определенного значения аргумента  $x_i$ ;

$y'_j$  – расчетное значение  $y$  при заданной величине аргумента  $x_i$  в соответствии с их теоретической взаимосвязью (то есть полученное путем подстановки  $x_i$  в уравнение теоретической зависимости).

Для нахождения значений параметров функции, соответствующих принципу наименьших квадратов, находятся частные производные функции по этим параметрам и приравниваются к нулю. Решая полученную систему уравнений, получаем параметры функции, при которых она максимально приближена к точкам корреляционного поля.

### **Линейная парная регрессия**

Линейная парная регрессия ищется в виде:

$$y = a + bx. \quad (2.2)$$

Задача сводится к отысканию коэффициентов регрессии  $a$  и  $b$  уравнения (2.2), при которых сумма квадратов отклонений точек корреляционного поля от прямой будет минимальна.

Величина  $\Delta y_j$ , представляющая собой расстояние от каждой точки корреляционного поля до теоретической линии регрессии, определяется из уравнения

$$\Delta y_j = y_j - (a + bx_j), \quad (2.3)$$

где  $x_j$  – параметр  $x$ , соответствующий измеренному значению  $y_j$ .

Для определения численных значений коэффициентов регрессии  $a$  и  $b$ , исходя из принципа минимизации квадратов отклонений, нужно приравнять нулю частные производные функции  $S^2$  уравнения (2.1) по  $a$  и  $b$ :

$$\frac{\partial S^2}{\partial a} = \sum_{j=1}^m 2(a + bx_j - y_j) = 2(ma + b \sum_{j=1}^m x_j - \sum_{j=1}^m y_j) = 0 \quad ; \quad (2.4)$$

$$\frac{\partial S^2}{\partial b} = \sum_{j=1}^m 2x_j(a + bx_j - y_j) = 2(a \sum_{j=1}^m x_j + b \sum_{j=1}^m x_j^2 - \sum_{j=1}^m x_j y_j) = 0 \quad . \quad (2.5)$$

Выполнив необходимые преобразования уравнений (2.4), (2.5), получим систему двух уравнений с двумя неизвестными для определения  $a$  и  $b$ :

$$\begin{pmatrix} m & \sum_{j=1}^m x_j \\ \sum_{j=1}^m x_j & \sum_{j=1}^m x_j^2 \end{pmatrix} * \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m y_j \\ \sum_{j=1}^m x_j y_j \end{pmatrix} . \quad (2.6)$$

Решая систему уравнений (2.6) относительно  $a$  и  $b$ , находим численные значения коэффициентов регрессии.

### 3.2 Задание

Получить в аналитической форме и вычислить значения коэффициентов парной линейной регрессии с помощью МНК для следующих рядов данных

экспериментов (первое число в скобке – входное значение  $x$ , второе – соответствующее ему значение  $y$ ).

### 3.3 Варианты исходных данных

**1** (1, 9.8); (1, 9.6); (1,10.6); (2,11.8); (2,13.0); (3,15.4); (3,13.8); (3,12.4); (4,15.4); (5,19.2); (5,18.2); (6,19.4); (6,18.6); (6,18.8); (7,22.0); (8,23.6); (9,25.4); (9,26.0); (10,26.8); (10,26.4).

**2** (1, 7.8); (1, 7.0); (1, 7.6); (2,14.2); (2,13.6); (3,20.6); (3,17.4); (3,20.6); (4,25.0); (5,30.6); (5,31.4); (6,36.8); (6,38.8); (6,38.8); (7,44.0); (8,48.6); (9,54.4); (9,54.0); (10,62.4); (10,60.2).

**3** (1, 6.6); (1, 6.8); (1, 7.4); (2,11.8); (2,11.8); (3,13.0); (3,14.2); (3,13.4); (4,14.8); (5,18.8); (5,20.2); (6,22.0); (6,20.4); (6,21.4); (7,26.6); (8,27.6); (9,32.8); (9,32.6); (10,35.2); (10,34.0).

**4** (1,14.0); (1,12.6); (1,13.8); (2,20.8); (2,18.6); (3,22.2); (3,24.4); (3,23.8); (4,27.6); (5,32.2); (5,33.2); (6,40.0); (6,40.8); (6,40.4); (7,44.4); (8,48.6); (9,52.2); (9,52.6); (10,57.6); (10,60.0).

**5** (1,19.4); (1,18.6); (1,19.6); (2,26.8); (2,27.8); (3,34.8); (3,35.8); (3,34.2); (4,41.4); (5,51.2); (5,50.4); (6,58.8); (6,58.6); (6,59.0); (7,66.6); (8,75.8); (9,81.6); (9,84.6); (10,90.4); (10,92.4).

**6** (1,10.0); (1, 7.8); (1, 8.2); (2,16.6); (2,14.4); (3,21.2); (3,23.4); (3,22.2); (4,29.2); (5,36.0); (5,37.0); (6,44.2); (6,43.8); (6,45.0); (7,51.8); (8,58.6); (9,63.2); (9,63.8); (10,70.4); (10,72.0);

### 3.4 Построение графика линейной функции по данным наблюдения, полученной МНК

График линейной функции, полученной МНК, можно получить с помощью функции *lsline*. Для этого сначала нужно построить график, соответствующий рядам данных экспериментов (для вывода точек корреляционного поля используются «+», для вывода полученной прямой – «\*», системная прямая выведется синей линией), затем воспользоваться этой функцией. Например:

```
>>plot(x,y,'+', x, a+b*x,'*'); lsline
```

### **Задание.**

На один график выведите ряды данных эксперимента, полученную вами прямую и прямую, полученную встроенными средствами MATLAB. Сравните результаты.

### **3.5 Оценка погрешности парной линейной регрессии**

После построения модели необходимо выполнить проверку ее адекватности.

Оценка адекватности модели производится на основании анализа значения коэффициента корреляции  $r$ , вычисляемого по формуле:

$$r = b \frac{\sigma_x}{\sigma_y}, \quad (2.9)$$

Где

$$\sigma_y = \frac{1}{m} \sqrt{m \sum_{j=1}^m y_j^2 - \left(\sum_{j=1}^m y_j\right)^2}; \quad (2.10)$$

$$\sigma_x = \frac{1}{m} \sqrt{m \sum_{j=1}^m x_j^2 - \left(\sum_{j=1}^m x_j\right)^2}. \quad (2.11)$$

Его величина может изменяться в пределах отрезка от  $-1$  до  $1$ . Чем ближе к единице модуль значения коэффициента корреляции, тем теснее линейная связь между  $x$  и  $y$ . При полном отсутствии связи  $r = 0$ , при  $|r| < 0.5$  гипотеза о наличии линейной связи отвергается, и линейная модель признается неадекватной.

Что касается диапазона значений модуля коэффициента от 0.5 до 1, то минимальный предел значения модуля  $r$ , при котором модель признается адекватной, определяется требованиями к точности модели – чем выше требования, тем ближе этот предел к единице.

Помимо оценки тесноты связи, коэффициент корреляции позволяет судить о характере зависимости между величинами  $x$  и  $y$ : если  $r > 0$ , между  $x$  и  $y$  имеет место положительная корреляционная связь, т. е. с ростом параметра  $x$  увеличивается параметр  $y$ ; если  $r < 0$ , между  $x$  и  $y$  имеет место отрицательная связь.

Мерой ошибки регрессионной модели служит среднее квадратичное отклонение (*стандартное отклонение*), которое рассчитывается по формуле:

$$\sigma = \left[ \frac{\sum_{j=1}^m (a + bx_j - y_j)^2}{m - 2} \right]^{1/2} \quad (2.12)$$

**Задание.** Вычислить значения всех приведенных параметров, оценить достоверность полученных коэффициентов.

#### 4 Содержание отчета

В отчете должны быть представлены:

- решения задач к п. 1 и п. 2;
- понятие регрессии, суть метода наименьших квадратов;
- последовательность команд для нахождения коэффициентов линейной регрессии и полученные коэффициенты;
- полученное уравнение линейной регрессии;
- график полученной прямой, системной прямой и точек корреляционного поля;

- команды для вычисления коэффициента корреляции, среднего квадратичного отклонения и полученные значения;
- анализ достоверности и точности полученной линейной модели.

**ЛАБОРАТОРНАЯ РАБОТА № 3**  
**ТЕМА: НЕЛИНЕЙНАЯ РЕГРЕССИЯ.**  
**ПОДГОНКА КРИВЫХ И ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ**

**Цель работы.** Получение вида и параметров нелинейной зависимости между величинами  $x$  по  $y$  экспериментальным данным.

**1 Краткие сведения из теории**

Связь между аргументом и функцией может иметь нелинейный характер. Аппроксимация кривой выполняется тем же путем с использованием метода наименьших квадратов, что и в случае прямой линии. Линия регрессии должна удовлетворять условию минимума суммы квадратов расстояний до каждой точки корреляционного поля. В данном случае в уравнении (1)  $y$  представляет собой расчетное значение функции, определенное при помощи уравнения выбранной криволинейной связи по фактическим значениям  $x_j$ .

В общем случае нелинейной зависимости

$$y'(x) = a_0 + a_1x + a_2x^2 + \dots + a_kx^k .$$

По методу наименьших квадратов находим частные производные по коэффициентам регрессии и приравниваем их к нулю. Получаем систему уравнений:

$$\begin{cases} \frac{\partial S^2}{\partial a_0} = -2 \sum_{j=1}^m (y_j - a_0 - a_1 x_j - a_2 x_j^2 - \dots - a_k x_j^k) = 0; \\ \frac{\partial S^2}{\partial a_1} = -2 \sum_{j=1}^m (y_j - a_0 - a_1 x_j - a_2 x_j^2 - \dots - a_k x_j^k) x_j = 0; \\ \dots \\ \frac{\partial S^2}{\partial a_k} = -2 \sum_{j=1}^m (y_j - a_0 - a_1 x_j - a_2 x_j^2 - \dots - a_k x_j^k) x_j^k = 0. \end{cases}$$

Оценкой тесноты связи при криволинейной зависимости служит теоретическое корреляционное отношение  $\eta_{xy}$ , представляющее собой корень квадратный из соотношения двух дисперсий: среднего квадрата  $\sigma_p^2$  отклонений расчетных значений  $y'_j$  функции по найденному уравнению регрессии от среднеарифметического значения  $Y$  величины  $y$  к среднему квадрату отклонений  $\sigma_y^2$  фактических значений функции  $y_j$  от ее среднеарифметического значения :

$$\eta_{xy} = \sqrt{\frac{\sigma_p^2}{\sigma_y^2}} = \sqrt{\frac{\sum_{j=1}^m (y'_j - Y)^2}{\sum_{j=1}^m (y_j - Y)^2}}.$$

## 2. Порядок выполнения работы

- 1 Загрузите ряды данных – вектора  $x$  и  $y$ .
- 2 Командой *cftool* откройте основное окно *Curve Fitting Tool*.
- 3 Нажмите кнопку *DATA*.
- 4 Задайте *XData* (вектор  $x$ ) и *YData* (вектор  $y$ ).
- 5 Нажмите *Create data set* и *OK* – окно закроется.
- 6 Нажмите кнопку *Fitting*. Появится окно редактора подгонки кривых.
- 7 Для выполнения нового сеанса подгонки нажмите *New Fit*.
- 8 В списке *Type of fit* выберите вид функции регрессии. Поочередно устанавливайте следующие виды функции регрессии (для выполнения подгонки

нажмите *APPLY*), результаты (вид функции и значения коэффициентов) увидите в окне *Results*. Зафиксируйте их в отчете.):

– *Polynomial* – полиномиальная. Поочередно попробуйте линейную, второго, третьего и четвертого порядка;

– *Fourier* – приближение рядом Фурье;

– *Gaussian* – приближение кривой гауссиана;

– *Power* – степенная функция;

– *Rational* – рациональная функция с выбором числителя и знаменателя;

– *Custom equation* – функция, задаваемая пользователем. Попробуйте задать следующие функции:

а)  $y=ax^2+bx^3+c$ ,

б)  $y=asin(2x)-bexp(x-3)+cx+d$ ,

в) такую функцию, которая, по вашему мнению, наилучшим образом приближена к экспериментальным данным.

Для добавления нового слагаемого нажмите *Add a term*.

9 Выберите три наиболее подходящих функции, выпишите их вид и коэффициенты.

10 Вернитесь в командную строку и для каждой функции вычислите значение теоретического корреляционного отношения.

11 Сделайте вывод о наиболее подходящей функции.

### 3 Варианты исходных данных

1. (1, 32.1); (1, 29.5); (1, 30.5); (2, 110.9); (2, 110.5); (3, 272.9); (3, 273.1); (3, 273.3); (4, 560.5); (5, 999.1); (5, 996.9); (6, 1627.3); (6, 1629.1); (6, 1630.1); (7, 2482.7); (8, 3597.1); (9, 5004.7); (9, 5006.7); (10, 6741.9); (10, 6743.5).

2. (1, -49.2); (1, -49.0); (1, -50.0); (2, -110.7); (2, -110.3); (3, -210.0); (3, -208.8); (3, -208.0); (4, -349.3); (5, -541.6); (5, -543.2); (6, -799.3); (6, -801.3); (6, -800.9); (7, -1128.2); (8, -1542.7); (9, -2045.4); (9, -2045.8); (10, -2648.9); (10, -2651.9).



3. (1, -7.1); (1, -6.3); (1, -5.3); (2, 32.2); (2, 29.8); (3, 111.5); (3, 111.1); (3, 112.9); (4, 244.8); (5, 442.5); (5, 442.5); (6, 712.8); (6, 713.8); (6, 715.2); (7, 1073.5); (8, 1525.4); (9, 2090.1); (9, 2088.5); (10, 2768.4); (10, 2768.6).

4. (1, -33.1); (1, -33.3); (1, -32.5); (2, -103.4); (2, -104.4); (3, -215.9); (3, -212.7); (3, -215.1); (4, -364.6); (5, -551.9); (5, -552.7); (6, -779.4); (6, -780.6); (6, -781.6); (7, -1047.5); (8, -1353.8); (9, -1696.1); (9, -1695.5); (10, -2078.4); (10, -2081.2).

5. (1, -2.6); (1, -3.0); (1, -4.6); (2, -49.0); (2, -50.4); (3, -132.0); (3, -132.6); (3, -132.0); (4, -262.0); (5, -445.6); (5, -445.6); (6, -696.2); (6, -696.6); (6, -698.6); (7, -1025.0); (8, -1438.4); (9, -1946.0); (9, -1944.6); (10, -2554.6); (10, -2554.4).

6. (1, 1.7); (1, 3.3); (1, 2.7); (2, -54.9); (2, -53.9); (3, -216.3); (3, -216.9); (3, -216.7); (4, -525.3); (5, -1017.1); (5, -1017.1); (6, -1730.5); (6, -1729.7); (6, -1730.5); (7, -2710.9); (8, -3993.7); (9, -5619.3); (9, -5618.5); (10, -7625.7); (10, -7624.5).

#### 4 Содержание отчета

Отчет должен содержать:

- общий вид функции нелинейной регрессии, метод получения значений ее коэффициентов;
- вид и коэффициенты трех наиболее подходящих функций, их графики;
- последовательность команд для получения значения теоретического корреляционного отношения для каждой из трех функций и полученные значения;
- выводы на основе анализа полученных значений теоретического корреляционного отношения.

#### ЛАБОРАТОРНАЯ РАБОТА № 4 ТЕМА: ПОЛУЧЕНИЕ КОЭФФИЦИЕНТОВ МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ И ВЫЧИСЛЕНИЕ ПОГРЕШНОСТИ

**Цель работы.** Получение навыков построения линейной модели для множественной зависимости по данным экспериментов.

### 1 Краткие сведения из теории

*Множественной регрессией* называется взаимосвязь трех и более переменных, или влияние двух и более аргументов на функцию

$$y = f(x_1, x_2, \dots, x_n).$$

Для простоты рассмотрим случай, когда ищется зависимость  $y$  от двух аргументов  $x_1$  и  $x_2$ . Такую зависимость графически можно представить в трехмерном пространстве  $\{y, x_1, x_2\}$ . Совокупность всех  $m$  точек представляет собой корреляционное пространство. Задача определения зависимости  $y$  от  $x_1$  и  $x_2$  состоит в том, чтобы подобрать такую плоскость, которая наилучшим, в смысле принципа наименьших квадратов, образом вписалась бы в данное корреляционное пространство. Зависимость ищется в виде:

$$y = a + b_1 x_1 + b_2 x_2.$$

Расстояние от точек корреляционного пространства до плоскости определяется выражением:

$$\Delta y_j = y_j - (a + b_1 x_1 + b_2 x_2).$$

Требуется найти значения коэффициентов  $a$ ,  $b_1$  и  $b_2$ .

Продифференцировав функцию суммы квадратов отклонений по коэффициентам  $a$ ,  $b_1$  и  $b_2$  и приравняв производные к нулю, получим систему трех уравнений с тремя неизвестными:

$$\begin{cases} \sum_{j=1}^m y_j = ma + b_1 \sum_{j=1}^m x_{1j} + b_2 \sum_{j=1}^m x_{2j}; \\ \sum_{j=1}^m y_j x_{1j} = a \sum_{j=1}^m x_{1j} + b_1 \sum_{j=1}^m x_{1j}^2 + b_2 \sum_{j=1}^m x_{1j} x_{2j}; \\ \sum_{j=1}^m y_j x_{2j} = a \sum_{j=1}^m x_{2j} + b_1 \sum_{j=1}^m x_{1j} x_{2j} + b_2 \sum_{j=1}^m x_{2j}^2. \end{cases} \quad (4.1)$$

Решение системы уравнений относительно коэффициентов  $a$ ,  $b_1$  и  $b_2$ , позволяет определить их численные значения. Величины  $\Sigma y$ ,  $\Sigma x_1$ ,  $\Sigma x_1^2$ ,  $\Sigma ux_1$ ,  $\Sigma y x_2$ ,  $\Sigma x_2$ ,  $\Sigma x_2^2$ ,  $\Sigma x_1 x_2$  находятся непосредственно по данным измерений.

Таким образом, найденное уравнение регрессии описывает совместное влияние  $x_1$  и  $x_2$  на функцию  $y$ . Коэффициенты  $a$ ,  $b_1$  и  $b_2$  при этом имеют следующий математический смысл.

Коэффициент  $a$  равен значению функции  $y$  при нулевых значениях аргументов  $x_1$  и  $x_2$ . В геометрической интерпретации он соответствует ординате точки пересечения плоскости регрессии с осью  $y$ .

Коэффициент  $b_1$  равен изменению функции  $y$  при изменении первого аргумента  $x_1$  на единицу при неизменном втором аргументе  $x_2$ . Аналогично, коэффициент регрессии  $b_2$  равен изменению функции  $y$  при изменении второго аргумента  $x_2$  на единицу при неизменном первом аргументе  $x_1$ .

Из уравнения множественной линейной регрессии могут быть получены уравнения *частной регрессии* аргументов  $x_1$  и  $x_2$  на функцию  $y$ :

$$y = a'_1 + b_1 x_1; \quad (4.2)$$

$$y = a'_2 + b_2 x_2. \quad (4.3)$$

При этом угловые коэффициенты регрессии  $b_1$  и  $b_2$  сохраняют те же числовые значения, что и в уравнении множественной регрессии. Свободные члены уравнений для  $y$  можно подсчитать следующим образом:

$$a'_1 = a + b_2 X_2; \quad (4.4)$$

$$a'_2 = a + b_1 X_1 \quad (4.5)$$

где  $a$ — свободный член в уравнении множественной регрессии ;  
 $X_1, X_2$ —средние значения соответствующих аргументов.

Оценкой тесноты связи при множественной линейной регрессии служит коэффициент множественной корреляции  $R$ , определяемый по формуле:

$$R = \sqrt{(b_1 \frac{\sigma_{x1}}{\sigma_y})^2 + (b_2 \frac{\sigma_{x2}}{\sigma_y})^2} \quad (4.6)$$

Величина коэффициента множественной корреляции всегда положительна и может меняться от 0 (при отсутствии связи) до 1 (при функциональной связи). С помощью коэффициента множественной корреляции оценивают совместное влияние на зависимую переменную всех включенных в расчет аргументов.

## 2. Задание

Решив систему уравнений (4.1), получить в аналитической форме, вычислить значения коэффициентов множественной линейной регрессии, получить с помощью формул (4.2) – (4.5) уравнения частной регрессии, вычислить по формуле (4.6) коэффициент множественной корреляции для заданных рядов

данных экспериментов. Построить трехмерные графики корреляционного поля и полученной плоскости, соответствующей уравнению множественной линейной регрессии.

Матрица значений  $x_1$  задается следующим образом:

$$x_1 = \begin{matrix} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 \\ & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 \\ & 7 & 7 & 7 & 7 & 7 & 7 & 7 & 7 & 7 & 7 \\ & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 \\ & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \end{matrix}$$

Матрица значений  $x_2$ :

$$x_2 = \begin{matrix} & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{matrix}$$

### 3 Варианты матрицы $y$ .

Вариант 1

3.7 -3.3 -6.9 -14.5 -19.7 -26.1 -32.7 -37.7 -43.7 -46.5  
 3.0 -2.8 -6.2 -9.2 -18.2 -26.8 -33.0 -38.4 -41.4 -49.2  
 3.3 -2.5 0.1 -10.5 -19.3 -20.9 -30.3 -34.9 -39.5 -50.7

5.2 3.4 -9.6 -4.2 -22.0 -14.2 -25.6 -26.6 -39.6 -46.2  
1.9 5.1 -8.7 -14.5 -22.3 -19.1 -22.9 -23.7 -27.5 -48.3  
22.2 -2.8 -6.2 -8.4 -22.6 -21.2 -25.8 -20.8 -45.8 -49.2  
1.7 -1.3 8.3 -3.1 -21.5 -17.5 -28.9 -30.5 -39.1 -49.1  
14.4 -2.6 1.2 -4.6 -12.0 -14.6 -23.6 -23.0 -46.4 -33.0  
28.5 1.1 16.9 -15.9 -18.1 -0.5 -15.3 -33.7 -21.5 -30.9  
13.4 3.6 23.8 14.0 -13.8 2.4 -17.4 -35.2 -21.0 -40.8

### Вариант 2

-31.2 -36.2 -43.2 -45.6 -53.4 -56.6 -63.6 -67.6 -73.6 -80.4  
-43.3 -48.7 -52.5 -59.5 -65.3 -73.1 -73.3 -84.3 -87.3 -90.7  
-57.6 -63.6 -72.0 -72.0 -74.4 -80.4 -87.0 -93.6 -101.4 -100.2  
-72.3 -80.1 -80.7 -81.3 -93.1 -100.9 -100.7 -112.5 -115.5 -116.9  
-74.8 -86.2 -91.6 -93.0 -101.4 -114.8 -121.2 -124.6 -121.0 -125.4  
-82.5 -102.3 -98.1 -119.1 -104.1 -131.1 -135.3 -132.3 -135.3 -152.7  
-112.8 -100.0 -123.6 -126.2 -124.6 -127.2 -138.2 -153.4 -161.6 -154.4  
-132.7 -136.5 -141.9 -136.1 -131.9 -142.1 -160.3 -146.5 -147.1 -173.3  
-118.2 -147.0 -138.0 -152.4 -163.2 -159.6 -172.2 -170.4 -157.8 -190.2  
-129.3 -158.7 -154.1 -169.5 -180.9 -176.3 -189.7 -185.1 -198.5 -187.9

### Вариант 3

-1.6 7.0 14.4 20.8 25.6 32.6 40.2 49.8 56.8 60.2  
-7.6 2.6 5.2 16.2 22.4 29.0 33.6 42.2 50.4 57.8  
-12.8 -2.8 -1.8 9.4 14.6 22.8 28.6 32.6 42.6 50.2  
-9.6 -13.8 -0.4 1.0 16.8 19.0 26.0 29.8 40.0 39.8  
-24.0 -21.0 -14.0 3.0 15.0 16.0 28.0 34.0 38.0 46.0  
-17.6 -17.8 -3.6 -9.8 2.0 16.2 26.8 29.0 21.6 37.0  
-39.4 -11.4 -29.6 -4.4 -15.6 -0.2 12.4 19.4 27.8 32.0  
-36.0 -32.2 -33.2 -21.4 -16.0 0.6 12.4 -1.4 5.6 27.0  
-32.0 -43.0 -27.0 -20.0 -29.2 -11.4 17.2 20.6 9.6 22.0

-64.0 -49.0 -34.0 -9.0 -8.0 1.0 -8.0 9.0 10.0 25.0

#### Вариант 4

9.7 15.1 17.5 21.5 24.7 26.7 30.1 33.5 37.9 40.9  
23.2 23.4 31.2 31.4 38.0 37.0 42.0 44.2 46.4 51.8  
35.1 38.5 37.1 37.5 39.7 46.1 52.5 48.7 58.1 63.9  
36.6 40.8 45.8 50.8 52.6 56.0 56.2 57.2 60.6 76.0  
45.3 45.7 53.1 68.5 59.9 72.3 66.7 69.1 67.5 72.9  
58.8 63.4 56.0 58.2 65.2 77.0 81.6 79.0 92.0 90.6  
72.5 61.9 66.7 67.3 81.9 79.7 88.7 90.7 82.9 98.9  
76.2 66.8 84.6 86.4 105.8 91.6 103.0 98.4 105.0 106.8  
87.3 90.7 101.3 90.3 115.3 102.5 107.7 94.9 116.3 130.5  
88.8 86.2 121.6 123.0 96.4 115.8 113.2 102.6 140.0 123.4

#### Вариант 5

-23.1 -31.1 -39.3 -45.7 -51.7 -62.9 -67.5 -77.7 -83.5 -90.7  
-32.7 -41.3 -49.1 -57.7 -59.5 -67.7 -80.3 -87.7 -95.9 -98.9  
-44.9 -53.3 -53.9 -64.7 -71.9 -82.1 -86.9 -97.1 -107.3 -110.3  
-44.3 -56.9 -68.7 -66.9 -76.3 -84.9 -91.9 -106.9 -107.5 -112.9  
-52.1 -65.9 -64.7 -81.5 -97.3 -90.1 -99.9 -111.7 -126.5 -124.3  
-70.7 -74.9 -79.1 -98.9 -101.9 -98.9 -110.3 -124.1 -125.9 -148.1  
-78.9 -72.7 -93.1 -100.9 -104.5 -115.1 -121.5 -140.5 -137.1 -151.9  
-85.1 -100.9 -107.1 -121.3 -117.9 -124.1 -141.5 -130.1 -157.1 -163.3  
-85.1 -107.3 -115.1 -115.7 -134.3 -143.9 -153.5 -139.7 -165.5 -166.1  
-109.1 -102.9 -130.7 -144.5 -130.3 -146.1 -135.9 -139.7 -155.5 -181.3

#### Вариант 6

-5.4 -4.0 -5.6 -4.6 -5.8 -6.0 -5.0 -6.4 -8.2 -7.6  
3.6 2.2 8.0 1.4 3.6 6.6 3.6 2.2 7.6 4.6  
17.8 17.6 17.4 11.2 11.6 12.0 11.2 19.4 12.0 11.2

21.2 30.6 23.2 34.2 20.4 29.0 21.6 25.4 32.4 18.6  
43.6 33.4 38.2 32.0 37.8 44.6 28.4 28.2 37.0 36.8  
49.6 49.4 39.6 43.0 44.0 52.2 41.2 37.4 44.4 39.4  
57.8 64.6 47.6 65.6 51.4 48.4 49.6 67.6 53.4 51.8  
82.8 71.4 66.4 74.2 69.2 78.6 57.6 79.8 76.4 66.6  
88.0 69.8 96.6 76.6 89.0 65.4 74.2 81.2 75.6 80.8  
75.6 101.4 83.2 107.0 88.8 90.6 100.4 90.2 86.0 81.8

#### 4 Построение трехмерных графиков

- 1 Введите вектора  $x1$  и  $x2$  (например, последовательности от 1 до 10).
- 2 Создайте из них матрицу:  $[x1, x2] = \text{meshgrid}(x1, x2)$ .
- 3 Задайте функцию  $z = x1 .* \sin(x1 + x2)$  (примечание: функция может быть задана не формулой, а матрицей значений  $z$ ).
- 4 Постройте график:  $\text{meshc}(x1, x2, z)$ .
- 5 Сохраните график для возможности вставки в документ, для этого выберите меню File→export. Выберите тип файла, который можно вставить в документ и дайте имя файлу.

#### 5 Порядок выполнения работы

- 1 Введите матрицы  $x1$  и  $x2$ .
- 2 Вычислите коэффициенты регрессии по методу наименьших квадратов.
- 3 Постройте графики по исходным данным и по полученному уравнению регрессии. Сравните результаты.
- 4 Вычислите значение коэффициента множественной корреляции.
- 5 Постройте уравнения частной регрессии  $y$  по  $x1$  и по  $x2$ .
- 6 Сделайте вывод об адекватности линейной модели.

#### 6 Содержание отчета



Отчет должен содержать:

- вид полученной функции множественной линейной регрессии, метод получения значений ее коэффициентов (последовательность команд);
- последовательность команд для получения значения коэффициента множественной корреляции и его значение;
- уравнения частной регрессии;
- выводы на основе анализа значения коэффициента множественной корреляции.

**ЛАБОРАТОРНАЯ РАБОТА № 5**  
**ТЕМА: ФАКТОРНЫЙ АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ**

**Цель работы.** Ознакомление с методом факторного анализа.

**1 Краткие сведения из теории**

Факторный анализ – это статистический метод анализа результатов наблюдений, зависящих от различных, одновременно действующих факторов, выбор наиболее важных факторов и оценка их влияния.

Идея факторного анализа заключается в разложении общей дисперсии случайной величины на независимые случайные слагаемые, каждое из которых характеризует влияние того или иного фактора.

Пусть  $X$  – случайная величина,  $A$  и  $B$  – влияющие на нее факторы,  $\bar{x}$  – среднее значение величины  $X$ . Допустим, что отклонение  $X$  от  $\bar{x}$  при действии факторов  $A$  и  $B$  можно представить в виде суммы

$$(X - \bar{x}) = \alpha + \beta + \gamma,$$

где  $\alpha$  – отклонение, вызываемое фактором  $A$ ;

$\beta$  – отклонение, вызываемое фактором  $B$ ;

$\gamma$  – отклонение, вызываемое другими неучтенными факторами.

Обозначим дисперсии  $X$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$  соответственно  $\sigma_X^2, \sigma_\alpha^2, \sigma_\beta^2, \sigma_\gamma^2$ .

Тогда имеет место равенство:

$$\sigma_X^2 = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2.$$

Сравнивая значения дисперсий, можно установить степень влияния факторов на величину  $X$ .

Пусть имеется  $r$  значений фактора  $A$  и  $s$  значений фактора  $B$ ;  $x_{ij}$  – значение величины  $X$  при  $i$ -м значении фактора  $A$  и  $j$ -м значении фактора  $B$ . Тогда наблюдаемые значения величины  $X$  можно представить в виде матрицы, строки которой соответствуют уровням фактора  $A$ , а столбцы – уровням фактора  $B$ . Вычислим средние значения по каждому столбцу, строке, а также общее среднее:

$$\bar{x}_{i*} = \frac{1}{s} \sum_{j=1}^s x_{ij}, \quad \bar{x}_{*j} = \frac{1}{r} \sum_{i=1}^r x_{ij}, \quad \bar{x} = \frac{1}{rs} \sum_{j=1}^s \sum_{i=1}^r x_{ij}.$$

Тогда основное тождество факторного анализа имеет вид:

$$Q = \sum_{j=1}^s \sum_{i=1}^r (x_{ij} - \bar{x})^2 = s \sum_{i=1}^r (\bar{x}_{i*} - \bar{x})^2 + r \sum_{j=1}^s (\bar{x}_{*j} - \bar{x})^2 + \sum_{i=1}^r \sum_{j=1}^s (x_{ij} - \bar{x}_{i*} - \bar{x}_{*j} + \bar{x})^2 = Q_1 + Q_2 + Q_3$$

Слагаемое  $Q_1$  характеризует изменение признака по фактору  $A$ ,  $Q_2$  – по фактору  $B$ ,  $Q_3$  – влияние неучтенных факторов.

Произведем оценку дисперсий:

$$s^2 = \frac{1}{rs-1} \sum_{j=1}^s \sum_{i=1}^s (x_{ij} - \bar{x})^2 = \frac{Q}{rs-1};$$

$$s_1^2 = \frac{1}{r-1} s \sum_{i=1}^r (\bar{x}_{i*} - \bar{x})^2 = \frac{Q_1}{r-1};$$

$$s_2^2 = \frac{1}{s-1} r \sum_{j=1}^s (\bar{x}_{*j} - \bar{x})^2 = \frac{Q_2}{s-1};$$

$$s_3^2 = \frac{1}{(r-1)(s-1)} \sum_{i=1}^r \sum_{j=1}^s (x_{ij} - \bar{x}_{i*} - \bar{x}_{*j} + \bar{x})^2 = \frac{Q_3}{(r-1)(s-1)}.$$

Известно, что если случайная величина распределена нормально, то отношение выборочных дисперсий имеет  $F$  - распределение.

$$F_A = \frac{s_1^2}{s_3^2}, \quad F_B = \frac{s_2^2}{s_3^2}.$$

Полученные значения  $F_A$  и  $F_B$  сравнивают с табличными при выбранном уровне значимости. Если значения меньше табличных, влияние факторов считается незначительным.

## **2 Задание к лабораторной работе**

В матрице даны значения величины  $y$  при различных значениях факторов  $x_1$  (изменяется по строкам) и  $x_2$  (изменяется по столбцам). Выполнить факторный анализ: определить существенность влияния каждого из факторов на величину  $y$  при уровне значимости 0.01.

## **3 Порядок выполнения работы**

1 Введите заданную матрицу (ее можно скопировать) и вставить в MATLAB, оформив в соответствующем формате.

2 Вычислите среднее арифметическое элементов матрицы.

3 Вычислите средние по столбцам.

Вычисление суммы элементов столбцов матрицы  $a$ :

```
>>sum(a)
```

Результатом является вектор, элементы которого являются суммами элементов по каждому столбцу матрицы. Для вычисления среднего арифметического разделите этот вектор на количество элементов в столбце.

4 Вычислите средние по строкам.

Для вычисления суммы элементов строки матрицы  $a$  можно воспользоваться той же функцией, предварительно выполнив транспонирование матрицы:

```
>>sum(a')
```

(при транспонировании строки превращаются в столбцы, а столбцы – в строки, операция транспонирования – «'»).

5 Из полученных средних по строкам и по столбцам сформируйте матрицы таким образом, чтобы из элементов каждой строки исходной матрицы можно было вычесть ее среднее значение (по строке), а из элементов каждого столбца – его среднее (то есть нужно получить 2 матрицы  $10 \times 10$ , в одной из которых все элементы строки равны среднему по этой строке, а во второй – все элементы столбца равны среднему по этому столбцу).

Прежде чем переходить к операциям с большими матрицами, выполните следующие действия с матрицами  $2 \times 2$ .

а) Вычесть из 1-й строки матрицы  $a$  размером  $2 \times 2$  единицу, из 2-й строки – двойку.

**Решение.** Для этого необходимо сформировать матрицу такого же размера, в каждой строке которой содержатся единицы, а во второй – двойки.

**Пример.** Пусть матрица  $a$  задана следующим образом:

```
>>a=[1 2; 3 4]
```

Зададим матрицу b:

```
>>b=[1 1; 2 2]
```

```
>> a=a-b
```

б) Вычесть из 1-го столбца матрицы a размером 2x2 единицу, из 2-го столбца – двойку.

**Решение.** Для этого необходимо сформировать матрицу такого же размера, содержащую в каждом столбце число, которое необходимо вычесть.

**Пример.** Для той же матрицы a сформируем новую матрицу b:

```
>>b=[1 2; 1 2]
```

в) Операции объединения матриц.

Можно получить матрицу, являющуюся объединением других матриц.

Матрицы при этом задаются так же , как элементы матриц.

**Пример.** Пусть вектор b=[1 2].

Тогда операция

```
>> b=[b b]
```

даст вектор b=[ 1 2 1 2],

а операция

```
>>b=[b;b]
```

даст матрицу b=[1 2; 1 2].

Операция

```
>> b=[b' b']
```

в результате даст матрицу b=[1 1; 2 2]

6 Вычислите значения  $s_1^2$ ,  $s_2^2$ ,  $s_3^2$ .

7 Вычислите значения F–распределения для обоих факторов.

8 Имея в виду, что табличное значение F–распределения равно 5.35, сделайте вывод о необходимости учета влияния каждого фактора.

## Варианты заданий

Вариант 1

29.90 32.29 36.07 41.19 47.62 55.32 64.26 74.44 85.83 98.42  
29.95 32.34 36.12 41.24 47.67 55.37 64.31 74.49 85.88 98.47  
30.02 32.40 36.19 41.31 47.74 55.44 64.38 74.56 85.95 98.54  
30.10 32.49 36.27 41.40 47.82 55.52 64.47 74.64 86.03 98.63  
30.20 32.58 36.37 41.49 47.92 55.62 64.56 74.74 86.13 98.72  
30.31 32.69 36.48 41.60 48.03 55.73 64.67 74.85 86.24 98.83  
30.43 32.81 36.60 41.72 48.15 55.85 64.79 74.97 86.36 98.95  
30.56 32.95 36.73 41.85 48.28 55.98 64.93 75.10 86.49 99.08  
30.70 33.09 36.87 42.00 48.42 56.12 65.07 75.24 86.64 99.23  
30.85 33.24 37.02 42.15 48.57 56.27 65.22 75.40 86.79 99.38

Вариант 2

21.26 21.58 21.88 22.16 22.44 22.72 22.99 23.26 23.52 23.78  
23.96 24.27 24.57 24.86 25.14 25.42 25.69 25.96 26.22 26.48  
28.45 28.77 29.07 29.36 29.64 29.91 30.18 30.45 30.71 30.98  
34.75 35.06 35.36 35.65 35.93 36.21 36.48 36.74 37.01 37.27  
42.84 43.15 43.45 43.74 44.02 44.30 44.57 44.83 45.10 45.36  
52.73 53.04 53.34 53.63 53.91 54.18 54.46 54.72 54.99 55.25  
64.41 64.73 65.03 65.31 65.60 65.87 66.14 66.41 66.67 66.93  
77.90 78.21 78.51 78.80 79.08 79.36 79.63 79.89 80.16 80.42  
93.18 93.50 93.79 94.08 94.36 94.64 94.91 95.18 95.44 95.70  
110.26 110.58 110.88 111.16 111.44 111.72 111.99 112.26 112.52  
112.78

Вариант 3

3.43 9.86 20.06 33.87 51.19 71.94 96.06 123.49 154.19 188.13  
3.59 10.02 20.22 34.03 51.35 72.10 96.22 123.65 154.35 188.29  
3.69 10.13 20.33 34.14 51.46 72.21 96.32 123.75 154.45 188.39  
3.77 10.21 20.41 34.22 51.54 72.28 96.40 123.83 154.53 188.47  
3.84 10.27 20.47 34.28 51.60 72.35 96.47 123.90 154.60 188.53  
3.89 10.33 20.53 34.34 51.66 72.40 96.52 123.95 154.65 188.59  
3.94 10.37 20.57 34.39 51.70 72.45 96.57 124.00 154.70 188.64  
3.98 10.42 20.62 34.43 51.75 72.50 96.61 124.04 154.74 188.68  
4.02 10.46 20.66 34.47 51.79 72.53 96.65 124.08 154.78 188.72  
4.06 10.49 20.69 34.50 51.82 72.57 96.69 124.12 154.82 188.75

Вариант 4

23.71 24.72 25.98 27.41 29.00 30.72 32.56 34.50 36.55 38.69  
23.96 24.98 26.23 27.66 29.25 30.97 32.81 34.76 36.80 38.95  
24.38 25.40 26.65 28.09 29.67 31.39 33.23 35.18 37.23 39.37  
24.97 25.99 27.24 28.68 30.26 31.98 33.82 35.77 37.82 39.96  
25.73 26.75 28.00 29.44 31.02 32.74 34.58 36.53 38.58 40.72  
26.66 27.68 28.93 30.37 31.95 33.67 35.51 37.46 39.51 41.65  
27.76 28.77 30.03 31.46 33.05 34.77 36.61 38.56 40.60 42.74  
29.02 30.04 31.29 32.73 34.32 36.03 37.87 39.82 41.87 44.01  
30.46 31.48 32.73 34.16 35.75 37.47 39.31 41.26 43.30 45.44  
32.06 33.08 34.33 35.77 37.35 39.07 40.91 42.86 44.91 47.05

Вариант 5

27.08 28.94 31.34 34.18 37.40 40.97 44.84 49.01 53.44 58.12  
27.17 29.02 31.42 34.27 37.49 41.05 44.93 49.09 53.53 58.21  
27.27 29.12 31.52 34.36 37.59 41.15 45.03 49.19 53.62 58.31  
27.37 29.23 31.63 34.47 37.69 41.26 45.13 49.30 53.73 58.41  
27.48 29.34 31.74 34.58 37.80 41.37 45.24 49.41 53.84 58.52  
27.60 29.45 31.85 34.69 37.92 41.48 45.36 49.52 53.95 58.64  
27.71 29.57 31.97 34.81 38.03 41.60 45.48 49.64 54.07 58.76  
27.84 29.69 32.09 34.93 38.16 41.72 45.60 49.76 54.19 58.88  
27.96 29.82 32.22 35.06 38.28 41.85 45.72 49.89 54.32 59.00  
28.09 29.94 32.34 35.19 38.41 41.97 45.85 50.01 54.45 59.13

Вариант 6

11.42 11.64 11.90 12.18 12.49 12.82 13.16 13.52 13.89 14.27  
12.93 13.15 13.40 13.69 14.00 14.32 14.66 15.02 15.39 15.78  
16.46 16.68 16.94 17.22 17.53 17.86 18.20 18.56 18.93 19.31  
22.70 22.92 23.18 23.46 23.77 24.10 24.44 24.80 25.17 25.55  
32.25 32.47 32.73 33.01 33.32 33.65 33.99 34.35 34.72 35.10  
45.67 45.89 46.15 46.43 46.74 47.07 47.41 47.77 48.14 48.52  
63.49 63.71 63.97 64.25 64.56 64.89 65.23 65.59 65.96 66.34  
86.21 86.43 86.69 86.98 87.28 87.61 87.95 88.31 88.68 89.06  
114.32 114.54 114.80 115.08 115.39 115.71 116.05 116.41 116.78

148.27 148.49 148.74 149.03 149.34 149.66 150.00 150.36 150.73  
151.12

#### 4 Содержание отчета

- понятие и назначение факторного анализа, основная идея и вычисление дисперсий;
- листинг сессии вычисления дисперсий;
- выводы о значимости факторов.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. **Максимей, И.В.** Имитационное моделирование на ЭВМ / И.В. Максимей. – М : «Радио и связь», 1988. – 227 с.
2. **Петрович, М.Л.** Регрессионный анализ и его математическое обеспечение на ЕС ЭВМ / М.Л. Петрович. – М : «Финансы и статистика», 1982. – 194 с.
3. **Бусленко, Н.П.** Моделирование сложных систем / Н.П. Бусленко.– М: «Наука», 1968. – 355 с.
4. **Адаптивные системы автоматического управления:** учеб. пособие / под ред. В.Б. Яковлева. – Л. : Изд-во Ленинградского университета, 1984. – 202 с.
5. **Хорафас, Д.Н.** Системы и моделирование / Д.Н. Хорафас. – М : «Мир», 1967. – 410 с.
6. **Перегудов, Ф.И.** Введение в системный анализ / Ф.И. Перегудов, Ф.И. Тарасенко. – М : «Высш. школа», 1989. – 367 с.
7. Системный анализ и структуры управления / под ред. В.Г. Шорина. – М : «Знание», 1975. – 303 с.
8. **Бусленко, Н.П.** Лекции по теории сложных систем / Н.П. Бусленко, В.В. Калашников, И.Н. Коваленко. – М : «Сов. радио», 1973. – 439с.



9. **Ермаков, С.М.** Статистическое моделирование / С.М. Ермаков, Г.А. Михайлов. – М : «Наука», 1982. – 296 с.
10. **Новгородцев, А.** Расчет электрических цепей в MATLAB / А. Новгородцев. – СПб : «ПИТЕР», 2004. – 250 с.
11. **Дьяконов, В.П.** MATLAB 6.5 в математике и моделировании / В.П. Дьяконов. – М : «СОЛОН–Пресс», 2003. – 560 с.