Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Ивановский государственный энергетический
университет имени В.И. Ленина»

Кафедра программного обеспечения компьютерных систем

Лабораторная работа №1
по курсу
«Методология научных исследований»

# Реализация базовых методов планирования эксперимента

Выполнил студент
группы 2-46м ______________ Ю.М. Овсянников

Проверил
д.т.н., профессор ______________ Е.Р. Пантелеев

Иваново 2023

# Содержание



# Цель лабораторной работы

Изучение и реализация базовых методов планирования эксперимента:

- метод наименьших квадратов (МНК);
- метод Гаусса решения системы линейных алгебраических уравнений.

# Задача лабораторной работы

В ходе лабораторной работы необходимо реализовать и отладить программу аппроксимирующую экспериментальные данные некоторой плоскостью с помощью метода наименьших квадратов (МНК) и метода Гаусса для решения системы линейных алгебраических уравнений.

# Используемые инструменты

Поставленная задача решалась с использованием следующих инструментов:

- операционная система Debain 12;
- компилятор языка с++ из GNU Compiler Collection (gcc);
- сборка проекта производилась с помощью утилиты cmake.
- текстовый редактор vim;

# Теоретические основы

## Постановка задачи

Пусть в результате эксперимента получена серия двумерных массивов, элементы которых можно рассматривать, как точки поверхности и необходимо определить координаты максимума выпуклости поверхности.

Особенность задачи состоит в том, что ограничиться определением индекса элемента, имеющего максимальное значение, нельзя, поскольку, во-первых, некоторые поверхности наклонены к горизонтали, и точки на верхнем краю могут быть выше точек на пике выпуклости, во-вторых, могут присутствовать ложные пики.

Решение состоит в том, чтобы выровнять заданную поверхность, аппроксимировав ее плоскостью, а затем из каждой точки поверхности вычесть значение аппроксимирующей функции в данной точке. Далее, выровненную поверхность можно аппроксимировать некоторой поверхностью вращения, например, параболоидом, и найти координаты его максимума [2].

## Основыне понятия

Методы теории планирования экспериментов (ТПЭ) направлены на разработку оптимальных планов проведения экспериментов с целью сокращения объема проводимых исследований при заданной точности и достоверности получения результатов,извлечения из полученных опытных данных максимума полезных сведений.

Исследуемый объект (реальный объект, модель объекта) рассматривается как «черный ящик», имеющий входы $X$ (управляемые независимые параметры) и выходы $Y$.

Переменные $x$ принято называть *факторами*.

Вектор $Y$ называется *откликом*. Зависимость отклика от факторов носит название *функции отклика*, а геометрическое представление функции отклика – *поверхности отклика*. Функция отклика рассматривается как показатель качества или эффективности объекта [3].

## Метод наименьших квадратов

Рассмотрим простой случай однофакторного эксперимента. Пусть функция отклика в нашем эксперименте (уравенение регрессии) имеет вид:

$$y = b_0 + b_1 x_1 \tag{1}$$

Для нахождения функции отклика необходимо определить коэффициенты $b_0$ и $b_1$ по результатам полученных в эксперименте данным. Поскольку не все экспериментальные точки лежат на одной прямой, то можно говорить о наличии $\xi_i$ – разности между экспериментальным и вычисленным по уравнению регрессии значениями $y$ в $i$-й экспериментальной точке:

$$y_i - b_0 - b_1 x_1 = \xi_i \tag{2}$$

Задача стоит найти такие коэффциенты $b_0$ и $b_1$ при которых отклонения $\xi$ будут минимальны:

$$U = \sum_{i=1}^{N} \xi_i^2 = \min. \tag{3}$$

Уравнение 3 представляет собой метод наименьших квадратов (МНК). Для получения коэффициентов обязательным условием является наличие не меньшего числа экспериментов над числом неизвестных коэффициентов. Преобразуем уравнение 3 к виду:

$$U = \sum_{i=1}^{N} \xi_i^2 = \sum_{i=1}^{N} (y_i - b_0 - b_1 x_{1,i})^2 = \min. \tag{4}$$

Известно, что минимум некторой функции, если он существует, достигается при одновременном равенстве нулю частных производных по всем неизвестным, т.е.

$$\frac{\partial U}{\partial b_0} = 0, \qquad \frac{\partial U}{\partial b_1} = 0. \tag{5}$$

Взяв частную производную получаем уравнения для определения коэффициентов $b_0$ и $b_1$:

$$-2 \sum_{i=1}^{N} (y_i - b_0 - b_1 x_{1,i}) = 0, \qquad -2 \sum_{i=1}^{N} (y_i - b_0 - b_1 x_{1,i}) x_{1,i} = 0. \tag{6}$$

Раскрыв скобки и сделав некоторые преобразования получаем:

$$Nb_0 + \sum_{i=1}^{N} x_{1,i} b_1 = \sum_{i=1}^{N} y_i, \qquad \sum_{i=1}^{N} x_{1,i} b_0 + \sum_{i=1}^{N} x_{1,i}^2 b_1 = \sum_{i=1}^{N} y_i x_{1,i}. \tag{7}$$

Из уравнения 7 можно получить выражения для $b_0$ и $b_1$:

$$\begin{gathered} b_0 = \frac{\sum\limits_{i=1}^{N} y_i \sum\limits_{i=1}^{N} x_{1,i}^2 - \sum\limits_{i=1}^{N} y_i x_{1,i} \sum\limits_{i=1}^{N} x_{1,i}}{N \sum\limits_{i=1}^{N} x_{1,i}^2 - \left(\sum\limits_{i=1}^{N} x_{1,i}\right)^2}, \\ b_1 = \frac{N \sum\limits_{i=1}^{N} y_i x_{1,i} - \sum\limits_{i=1}^{N} y_i \sum\limits_{i=1}^{N} x_{1,i}}{N \sum\limits_{i=1}^{N} x_{1,i}^2 - \left(\sum\limits_{i=1}^{N} x_{1,i}\right)^2}. \end{gathered} \tag{8}$$

Обобщение на многофакторный случай не связано с какими-либо принципиальными трудностями, но требуют привлеченеия аппарата агебры матриц [1].

# Описание алгоритма решения задачи

Получим систему уравнений для нахождения коэффициентов регрессии для случая двухфакторного эксперимента с помощью МНК.

Функция отклика для случая двухфакторного эксперимента будет иметь следующий вид:

$$y = b_0 + b_1 x_1 + b_2 x_2, \tag{9}$$

где $x_1$, $x_2$ – факторы в рассматриваемом эксперименте, $b_0$, $b_1$, $b_2$ – искомые коэффициенты.

Минимизируемая функия суммы квадратов отклонений $U$:

$$U = \sum_{i=1}^{N} (y_i - b_0 - b_1 x_{1,i} - b_2 x_{2,i})^2 = \min. \tag{10}$$

Минимум данной функции, если он существует, будет достигаться при одновременном равенстве нулю частных производных по всем неизвестным, т.е.

$$\frac{\partial U}{\partial b_0} = 0, \qquad \frac{\partial U}{\partial b_1} = 0, \qquad \frac{\partial U}{\partial b_2} = 0. \tag{11}$$

$$\frac{\partial U}{\partial b_0} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 x_{1,i} - b_2 x_{2,i}) = 2\sum_{i=1}^{N}(b_0 + b_1 x_{1,i} + b_2 x_{2,i} - y_i) = 0$$

$$\frac{\partial U}{\partial b_1} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 x_{1,i} - b_2 x_{2,i})x_{1,i} = 2\sum_{i=1}^{N}(b_0 + b_1 x_{1,i} + b_2 x_{2,i} - y_i)x_{1,i} = 0 \quad (12)$$

$$\frac{\partial U}{\partial b_2} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 x_{1,i} - b_2 x_{2,i})x_{2,i} = 2\sum_{i=1}^{N}(b_0 + b_1 x_{1,i} + b_2 x_{2,i} - y_i)x_{2,i} = 0$$

Раскрывая скобки и вынося константы за знак суммы, получаем нормальную форму системы линейных уравнений:

$$b_0\sum_{i=1}^{N}1 + b_1\sum_{i=1}^{N}x_{1,i} + b_2\sum_{i=1}^{N}x_{2,i} = \sum_{i=1}^{N}y_i,$$

$$b_0\sum_{i=1}^{N}x_{1,i} + b_1\sum_{i=1}^{N}x_{1,i}^2 + b_2\sum_{i=1}^{N}x_{2,i}x_{1,i} = \sum_{i=1}^{N}y_i x_{1,i}, \quad (13)$$

$$b_0\sum_{i=1}^{N}x_{2,i} + b_1\sum_{i=1}^{N}x_{1,i}x_{2,i} + b_2\sum_{i=1}^{N}x_{2,i}^2 = \sum_{i=1}^{N}y_i x_{2,i}.$$

После подстановки в 13 данных эксперимента $x_{1,i}$, $x_{2,i}$ и $y_i$ для всех $N$ измерений данную систему можно решить методом Гауса для нахождения искомых коэффциентов $b_0$, $b_1$, $b_2$. Алгоритм метода Гаусса представлен на рис. 1.

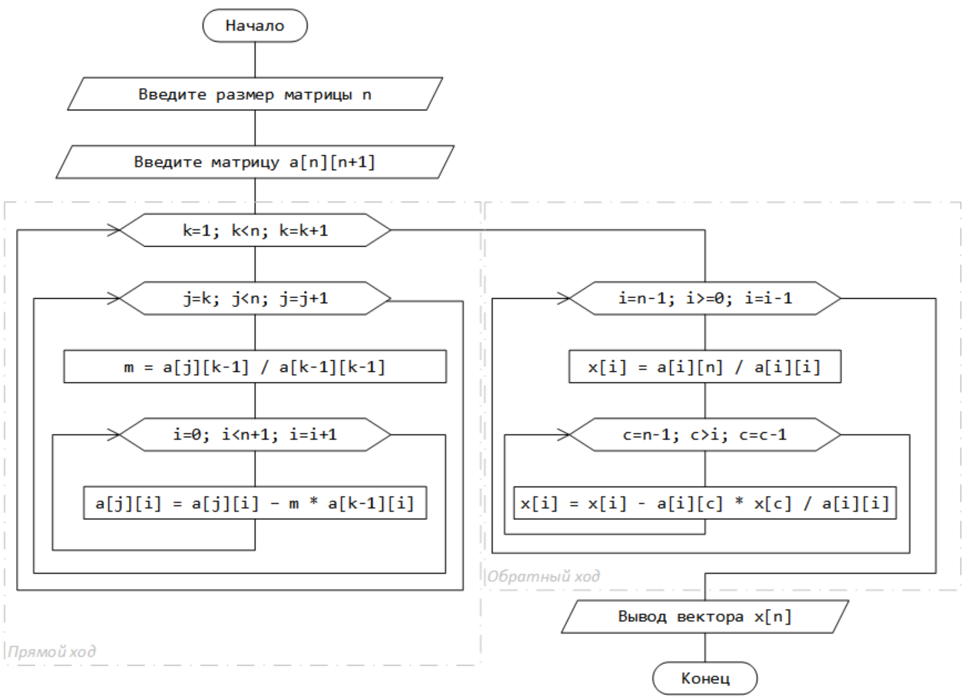


Рис. 1. Алгоритм Гаусса для решения системы линейных алгебраических уравнений

## Метод Гаусса

Метод Гаусса — классический метод решения системы линейных алгебраических уравнений. Рассмотрим систему линейных уравнений с действительными постоянными коэффициентами:

$$\begin{cases} a_{11} \cdot x_{11} + a_{12} \cdot x_2 + \cdots + a_{1n} \cdot x_n = y_1, \\ a_{21} \cdot x_{21} + a_{22} \cdot x_2 + \cdots + a_{2n} \cdot x_n = y_2, \\ \\ a_{n1} \cdot x_{n1} + a_{n2} \cdot x_2 + \cdots + a_{nn} \cdot x_n = y_n. \end{cases} \tag{14}$$

Метод Гаусса решения системы линейных уравнений включает в себя 2 стадии:

- последовательное (прямое) исключение;
- обратная подстановка.

Исключения Гаусса основаны на идее последовательного исключения переменных по одной до тех пор, пока не останется только одно уравнение с одной переменной в левой части. Затем это уравнение решается относительно единственной переменной. Таким образом, систему уравнений приводят к треугольной (ступенчатой) форме. Для этого среди элементов первого столбца матрицы выбирают ненулевой (а чаще максимальный) элемент и перемещают его на крайнее верхнее положение перестановкой строк. Затем нормируют все уравнения, разделив его на коэффициент $a_{i1}$, где $i$– номер столбца.

$$\begin{cases} x_{11} + \frac{a_{12}}{a_{11}} \cdot x_2 + \cdots + \frac{a_{1n}}{a_{11}} \cdot x_n = \frac{y_1}{a_{11}}, \\ x_{11} + \frac{a_{22}}{a_{21}} \cdot x_2 + \cdots + \frac{a_{2n}}{a_{21}} \cdot x_n = \frac{y_2}{a_{21}}, \\ \\ x_{11} + \frac{a_{n2}}{a_{n1}} \cdot x_2 + \cdots + \frac{a_{nn}}{a_{n1}} \cdot x_n = \frac{y_1}{a_{n1}}, \end{cases} \tag{15}$$

Затем вычитают получившуюся после перестановки первую строку из остальных строк:

$$\begin{cases} x_{11} + \frac{a_{12}}{a_{11}} \cdot x_2 + \cdots + \frac{a_{1n}}{a_{11}} \cdot x_n = \frac{y_1}{a_{11}}, \\ 0 + \left(\frac{a_{22}}{a_{21}} - \frac{a_{12}}{a_{11}}\right) \cdot x_2 + \cdots + \left(\frac{a_{2n}}{a_{21}} - \frac{a_{1n}}{a_{11}}\right) \cdot x_n = \frac{y_2}{a_{21}}, \\ \\ 0 + \left(\frac{a_{n2}}{a_{n1}} - \frac{a_{12}}{a_{11}}\right) \cdot x_2 + \cdots + \left(\frac{a_{nn}}{a_{n1}} - \frac{a_{1n}}{a_{11}}\right) \cdot x_n = \frac{y_1}{a_{n1}}, \end{cases} \tag{16}$$

Получают новую систему уравнений, в которой заменены соответствующие коэффициенты:

$$\begin{cases} x_{11} + a'_{12} \cdot x_2 + \cdots + a'_{1n} \cdot x_n = y'_1, \\ 0 + a'_{22} \cdot x_2 + \cdots + a'_{2n} \cdot x_n = y'_2, \\ \\ 0 + a'_{n2} \cdot x_2 + \cdots + a'_{nn} \cdot x_n = y'_n. \end{cases} \tag{17}$$

После того, как указанные преобразования были совершены, первую строку и первый столбец мысленно вычёркивают и продолжают указанный процесс для всех последующих уравнений пока не останется уравнение с одной неизвестной:

$$\begin{cases} x_{11} + a'_{12} \cdot x_2 + \cdots + a'_{1n} \cdot x_n = y'_1, \\ 0 + x_2 + \cdots + a''_{2n} \cdot x_n = y''_2, \\ \ldots \\ 0 + 0 + \cdots + x_n = y_n^{n'}. \end{cases} \tag{18}$$

Обратная подстановка предполагает подстановку полученного на предыдущем шаге значения переменной $x_n$ в предыдущие уравнения [4].

# Листинг программы МНК

```cpp
#include <iostream>
using namespace std;

// Вывод системы уравнений
void sysout(double **a, double *y, int n)
{

  for (int i = 0; i < n; i++)
  {
    for (int j = 0; j < n; j++)
    {
      cout << a[i][j] << "*x" << j;
      if (j < n - 1)
        cout << " + ";
    }
    cout << " = " << y[i] << endl;
  }
  return;
}

double * gauss(double **a, double *y, int n)
{
  double *x, max;
  int k, index;
  const double eps = 0.00001;  // точность
  x = new double[n];
  k = 0;
  while (k < n)
  {
    // Поиск строки с максимальным a[i][k]
    max = abs(a[k][k]);
    index = k;

    for (int i = k + 1; i < n; i++)
    {
      if (abs(a[i][k]) > max)
      {
        max = abs(a[i][k]);
        index = i;
      }
    }
    // Перестановка строк
    if (max < eps)
    {
      // нет ненулевых диагональных элементов
      cout << "Решение получить невозможно изза- нулевого столбца ";
      cout << index << " матрицы A" << endl;
      return 0;
    }

    for (int j = 0; j < n; j++)
    {
      double temp = a[k][j];
      a[k][j] = a[index][j];
      a[index][j] = temp;
    }

```

```
    double temp = y[k];
    y[k] = y[index];
    y[index] = temp;

    // Нормализация уравнений
    for (int i = k; i < n; i++)
    {
      double temp = a[i][k];
      if (abs(temp) < eps) continue; // для нулевого коэффициента пропустить
                                     //
      for (int j = 0; j < n; j++)
        a[i][j] = a[i][j] / temp;

      y[i] = y[i] / temp;
      if (i == k)  continue; // уравнение не вычитать само из себя
                             //
      for (int j = 0; j < n; j++)
        a[i][j] = a[i][j] - a[k][j];

      y[i] = y[i] - y[k];
    }
    k++;
  }

  // обратная подстановка
  for (k = n - 1; k >= 0; k--)
  {
    x[k] = y[k];
    for (int i = 0; i < k; i++)
      y[i] = y[i] - a[i][k] * x[k];
  }

  return x;
}

int main()
{
  double **a, *y, *x;
  int n;
  system("chcp 1251");
  system("cls");
  cout << "Введите количество уравнений: ";
  cin >> n;
  a = new double*[n];
  y = new double[n];

  for (int i = 0; i < n; i++)
  {
    a[i] = new double[n];
    for (int j = 0; j < n; j++)
    {
      cout << "a[" << i << "][" << j << "]= ";
      cin >> a[i][j];
    }
  }

  for (int i = 0; i < n; i++)
  {
    cout << "y[" << i << "]= ";
```

```
    cin >> y[i];
  }

  sysout(a, y, n);
  x = gauss(a, y, n);

  for (int i = 0; i < n; i++)
    cout << "x[" << i << "]=" << x[i] << endl;

  cin.get(); cin.get();
  return 0;
}
```

# Список литературы


1. **Адлер Ю. П.** Планирование эксперимента при поиске оптимальных условий. — Наука, 1976. — С. 278.

2. Аппроксимация эмпирически полученной поверхности методом наименьших квадратов. — URL: http://www.delphikingdom.com/asp/viewitem.asp?catalogid=1368 (дата обр. 01.12.2023).

3. Методы планирования экспериментов с использованием пакета Minitab: учебное пособие студентам магистратуры направления 09.04.02 «Информационные системы и технологии» для выполнения практических и самостоятельных работ по дисциплине «Модели и методы планирования экспериментов» / И. В. Макарова, Р. Г. Хабибуллин, А. И. Беляев, К. А. Шубенкова. — Набережные Челны: Набережночелнинский институт КФУ, 2016. — URL: https://kpfu.ru/staff_files/F430820888/Uchebnoe_posobie_MiMPE.pdf.

4. Решение систем линейных уравнений методом Гаусса. — URL: https://prog-cpp.ru/gauss/ (дата обр. 01.12.2023).