Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования

«Уфимский государственный авиационный технический университет»

Кафедра вычислительной математики и кибернетики

Лабораторная работа №3

по дисциплине: «Компьютерная обработка экспериментальных данных»

«Метод главных компонент»

Выполнили:

студенты группы МО-417

Ярыгин А.Э.

Алимгафаров А.Р.

Проверила:

Харисова Э. А.

Уфа 2022

**Цель:**

Освоить и приобрести практические навыки метода главных компонент для решения задачи снижения размерности экспериментальных данных.

**Задачи**

1. Выполнить задание для самостоятельной работы в соответствии с настоящим руководством по выполнению лабораторной работы;
2. Оформить отчет о выполнении лабораторной работы в соответствии с требованиями к его оформлению.

**Ход работы**

1. В качестве входных данных возьмём матрицу значений различный признаков для областей РФ (9 столбцов исходных данных). Данные будем хранить в excel-файле, можно часть из них рассмотреть на рисунке 1:

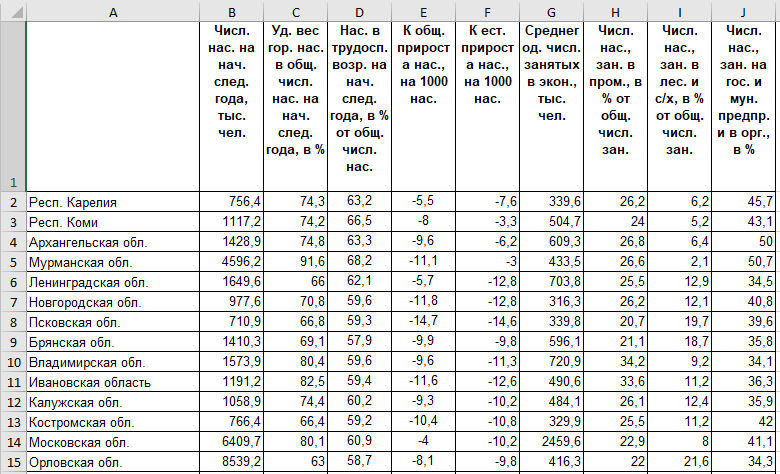
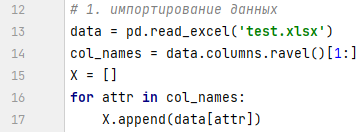
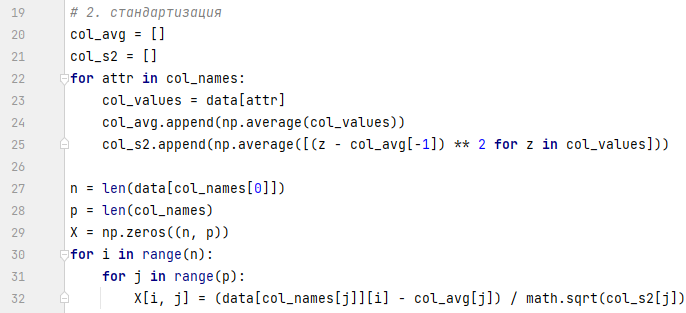


Рисунок 1 - Входные данные

  
Рисунок 2 – Код загрузки исходных данных

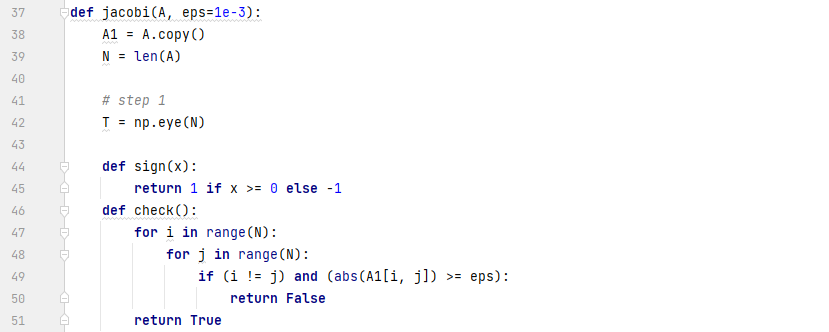
1. Далее для устранения неоднородности исходных данных их необходимо стандартизировать (выполнить нормирование) по столбцам-признакам:

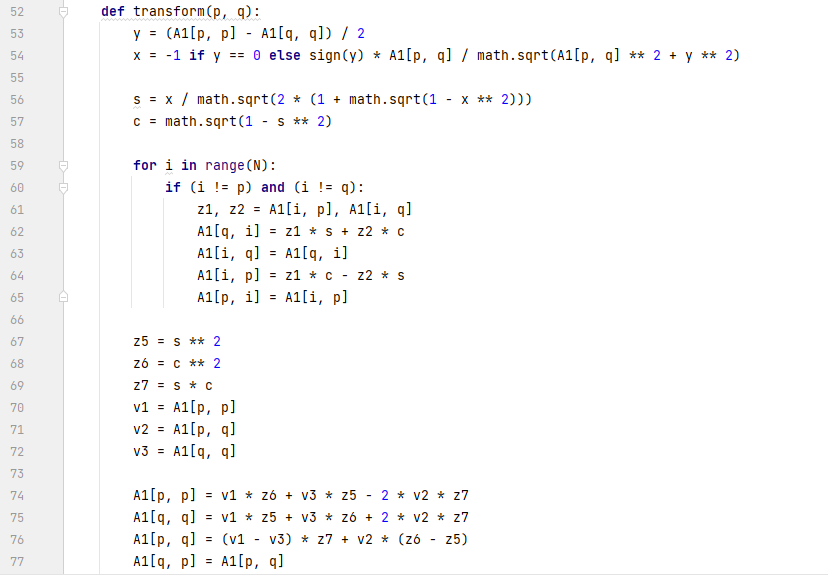
  
Рисунок 3 – Стандартизация исходных данных

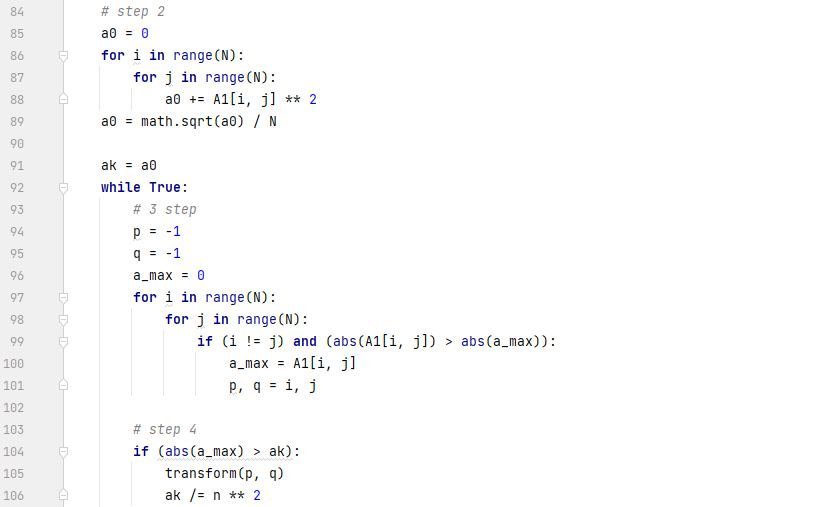
1. Построим матрицу ковариации. Ввиду произведенной стандартизации данных матрица ковариации будет корреляционной матрицей исходных данных *R* порядка *p* на *p*, где *p* – количество столбцов-признаков (*N* – количество данных у каждого признака):
2. Прежде чем начинать анализ главных компонент, целесообразно проверить, значимо ли отличается от единичной матрицы корреляционная матрица исходных стандартизованных данных. В предположении, что исходные данные подчиняются многомерному нормальному распределению, можно воспользоваться статистикой:

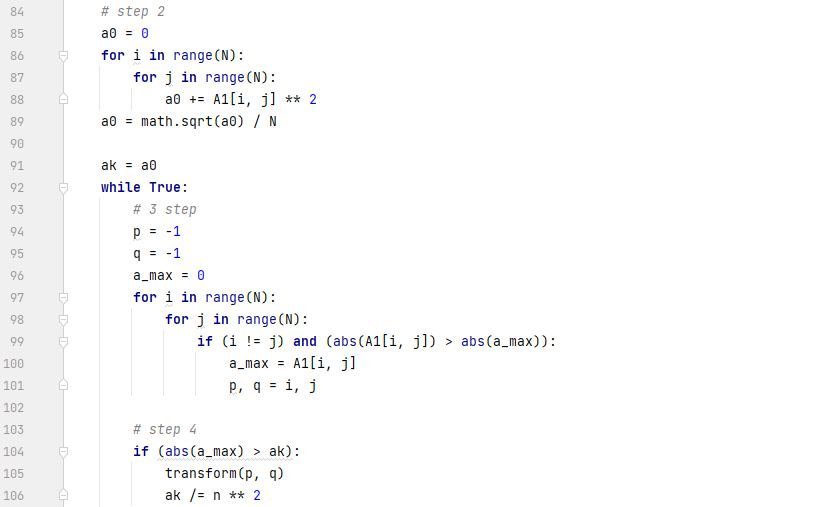
Если *d* ≤ при степенях свободы , то применение метода главных компонент нецелесообразно.

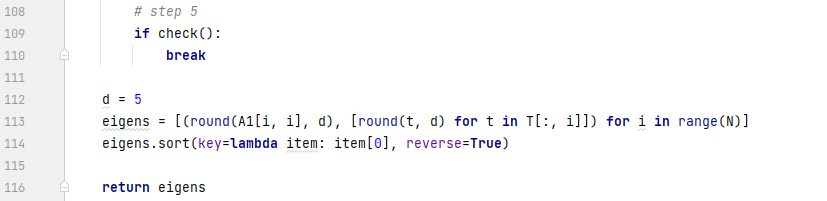
1. Если проверочное условие выполняется, то к полученной матрице *R* применим итерационный метод Якоби с преградами для нахождения собственных чисел и собственных векторов. Код алгоритма представлен на рисунках 4 – 8:

  
Рисунок 4 – Метод Якоби: часть 1

  
Рисунок 5 – Метод Якоби: часть 2

  
Рисунок 6 – Метод Якоби: часть 3

  
Рисунок 7 – Метод Якоби: часть 4

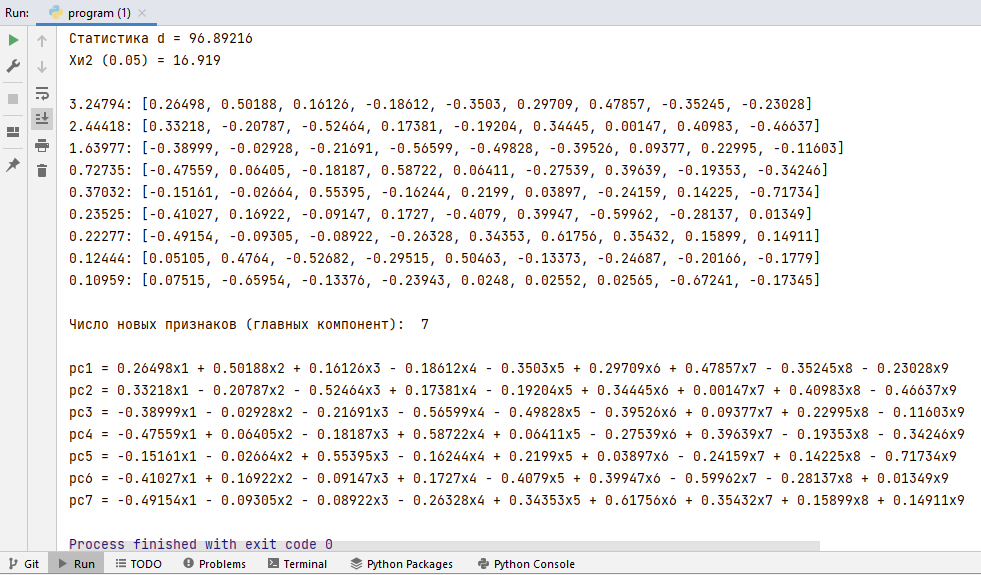
  
Рисунок 8 – Метод Якоби: часть 5

Функция *jacobi(A, eps)* возвращает отсортированный в порядке убывания список пар , где λ – число собственное значение, *с –* соответствующий собственному числу собственный вектор (нормированный); упорядочивание списка производится по собственным числам.

1. Определим количество главных компонент. Вычислим величину:

причем *p’* – число новых признаков, выбрать минимальным, удовлетворяющим условию *I( p′) > 0.95*.

1. Проверим основные свойства главных компонент:
2. Относительная доля разброса, приходящаяся на *j*-ую главную компоненту:
3. Относительная доля разброса, приходящаяся на p ′ первых компонент:
4. Рассчитаем проекции объектов на главные компоненты:
5. Протестируем программу на экспериментальных данных и выведем на экран полученную информацию. Результат работы программы можно рассмотреть на рисунке 9:

  
Рисунок 9 – Пример работы программы

**Вывод**

В ходе выполнения лабораторной работы мы реализовали программу для нахождения главных компонент с помощью итерационного метода Якоби.