Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Пермский национальный исследовательский политехнический университет»

Факультет: Прикладной математики и механики
Кафедра: Вычислительной математики, механики и биомеханики

Направление: 09.04.02 Информационные технологии и системная инженерия
Профиль: «Информационные технологии и системная инженерия»

КУРСОВАЯ РАБОТА
по дисциплине
«Специальные главы математики»

Выполнил
студент гр. ИТСИ-24-1м
Слаутин Максим Егорович

Принял Преподаватель кафедры ВММБ
Онискив Владимир Дмитриевич

Пермь 2025

Содержание

Лабораторная работа №1

Задание

Имеется информация о цене товара за год:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12
Цена	6.1	8.1	9.0	8.2	7.8	8.0	8.2	10.0	11.3	11.0	10.2	9.8

Необходимо:

1. Составить уравнение тренда
2. Выделить периодическую составляющую
3. Вычислить sin(t), cost(t), sin(2t), cost(2t)
4. Сделать прогноз на 14 месяцев

Теория

Временной ряд — это последовательность данных, измеренных в последовательные моменты времени, часто через равные интервалы. Основная цель анализа временных рядов — выявление закономерностей, прогнозирование будущих значений и извлечение полезной информации из данных. Ниже представлены основные компоненты временного ряда:

Тренд — долгосрочная тенденция (рост, спад или стабильность).

Сезонность — периодические колебания с фиксированной частотой (например, месячные или квартальные изменения).

Цикличность — долгосрочные колебания без строгой периодичности (например, экономические циклы).

Шум (остатки) — случайные несистематические колебания, не объясняемые моделью.

Решение

```
==================================================
                 ИСХОДНЫЕ ДАННЫЕ
==================================================
+------------+---------------+
```

Месяц (t)	Значение (Y)
1	6.10
2	8.10
3	9.00
4	8.20
5	7.80
6	8.00
7	8.20
8	10.00
9	11.30
10	11.00
11	10.20
12	9.80

```
==================================================
             РАСЧЕТ ЛИНЕЙНОГО ТРЕНДА
==================================================
Сумма t: 78.00
Сумма Y: 107.70
Сумма t²: 650.00
Сумма Y²: 991.91

Среднее t: 6.50
Среднее Y: 8.97

Уравнение линейного тренда:
y(t) = 0.3353 * t + 6.7955

==================================================
            РАСЧЕТ СЕЗОННОЙ КОМПОНЕНТЫ
==================================================

Отклонения (Y - y(t)):
Месяц  1: -1.0308
Месяц  2: +0.6339
Месяц  3: +1.1986
Месяц  4: +0.0633
Месяц  5: -0.6720
Месяц  6: -0.8073
Месяц  7: -0.9427
Месяц  8: +0.5220
Месяц  9: +1.4867
Месяц 10: +0.8514
Месяц 11: -0.2839
Месяц 12: -1.0192

Коэффициенты гармоник:
```

A_1 = 0.1585
A_2 = -2.2320
B_1 = -0.3539
B_2 = -0.2140

==

ПРОГНОЗ НА 14 МЕСЯЦЕВ ВПЕРЕД

==

Прогнозные значения:

Исходный период (12 месяцев):

Месяц	Тренд	Сезонная компонента	Прогноз
1	7.13	-2.07	5.06
2	7.47	0.52	7.99
3	7.80	1.23	9.03
4	8.14	-2.29	5.85
5	8.47	0.28	8.75
6	8.81	2.37	11.18
7	9.14	-1.52	7.63
8	9.48	-0.63	8.85
9	9.81	1.83	11.64
10	10.15	-1.60	8.55
11	10.48	-1.05	9.44
12	10.82	2.59	13.41

Прогноз на следующие 14 месяцев:

Месяц	Тренд	Сезонная компонента	Прогноз
13	11.15	-0.43	10.73
14	11.49	-1.61	9.88
15	11.83	1.76	13.59
16	12.16	-0.48	11.68
17	12.50	-2.03	10.47
18	12.83	2.08	14.91
19	13.17	0.87	14.03
20	13.50	-2.09	11.41
21	13.84	1.07	14.91

```
|       22 |    14.17 |                      0.71 |      14.88 |
+---------+--------+---------------------+----------+
|       23 |    14.51 |                     -2.39 |      12.12 |
+---------+--------+---------------------+----------+
|       24 |    14.84 |                      0.99 |      15.83 |
+---------+--------+---------------------+----------+
|       25 |    15.18 |                      1.98 |      17.16 |
+---------+--------+---------------------+----------+
|       26 |    15.51 |                     -1.89 |      13.62 |
+---------+--------+---------------------+----------+
```

Вывод

В результате выполнения работы была написана программа, которая по исходным значениям временного ряда строит уравнение линейного тренда, выделяет сезонную зависимость с помощью тригонометрических функций, а также прогнозирует будущие наиболее возможные значения.

Лабораторная работа №2

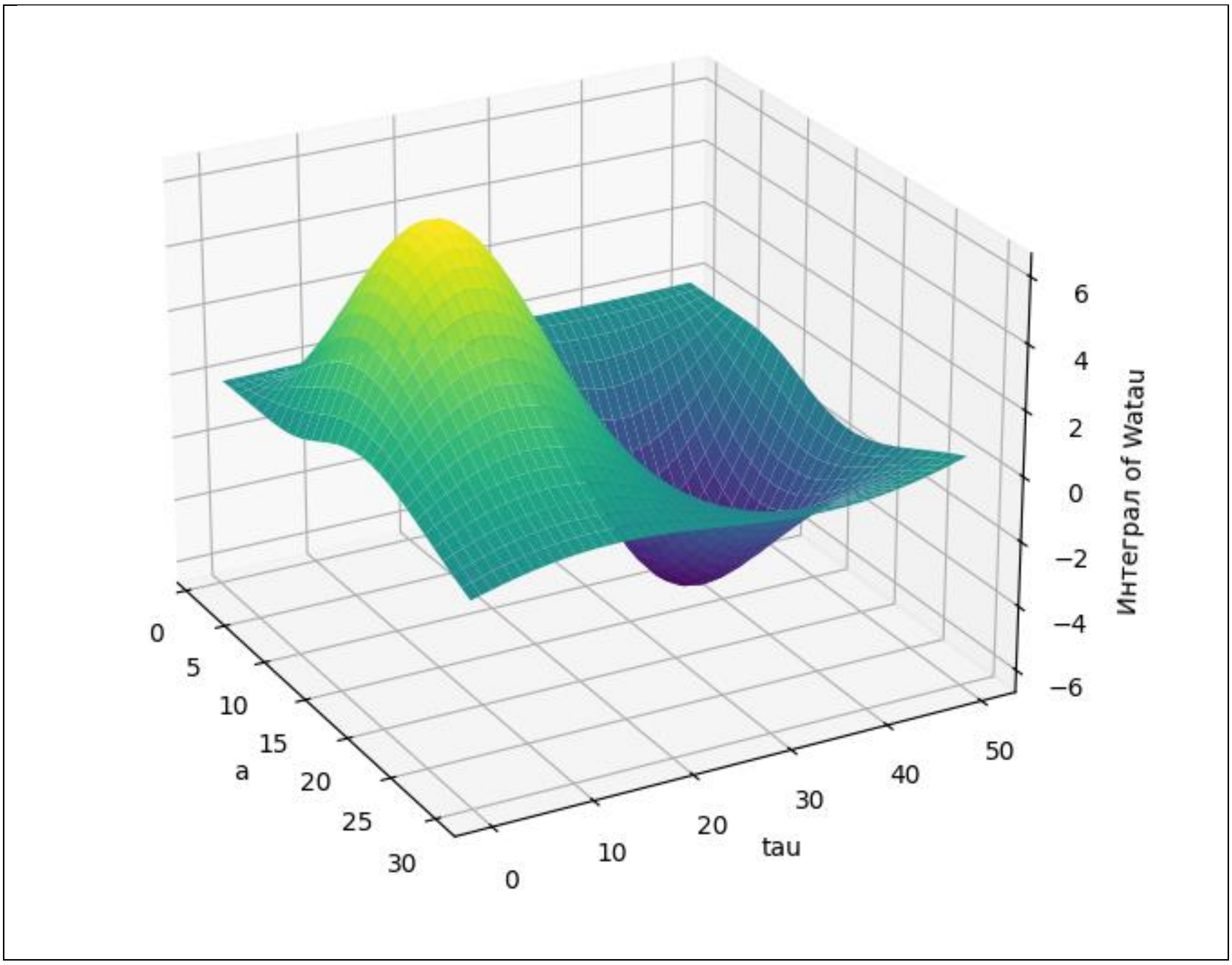
Задание

Построить изолинию W(a, τ, t) для a=[1,30] и τ=[0,50], на промежутке от -25 до 75 (t=[-25,75]).

Теория

Изолиния (или изоповерхность) — это множество точек в трёхмерном пространстве, в которых функция W(a, τ, t) принимает постоянное значение. Изолиния (или изоповерхность) представляет собой геометрическое место точек в трёхмерном пространстве, где функция W(a, τ, t)сохраняет постоянное значение. Это означает, что для любого набора параметров a, τ и t, удовлетворяющих уравнению W(a, τ, t)=C, где C — фиксированная константа, соответствующая точка принадлежит изолинии. Такие поверхности широко используются в математическом моделировании, физике и инженерных расчётах для визуализации скалярных полей и анализа распределения значений функции в пространстве. Построение изолиний позволяет наглядно представить структуру данных, выделяя области с одинаковыми характеристиками, что особенно полезно при исследовании сложных многомерных зависимостей.

Решение



Вывод

В ходе выполнения лабораторной работы была написана программа, которая визуализируют в 3D-модель изолинию на выбранных ограничениях по переменным, на которой хорошо видно к a, τ и t влияют на функцию W(a, τ, t).

Лабораторная работа №3

Задание

Исследуемый сигнал задается рядом дискретных значений:

N	1	2	3	4	5	6	7	8	9	10
t	0	0.5	1	1.6	2.1	2.6	3.1	3.4	3.6	3.9
f	0	0.5	0.9	0.98	0.8	0.4	-0.8	1.6	-1.4	-1.6
N	11	12	13	14	15	16	17	18	19	
t	4.2	4.45	4.7	4.97	5.24	5.5	5.7	6.02	6.28	
f	1	-1.4	-2	1.1	-1.1	-2.6	1.8	0	-1.05	

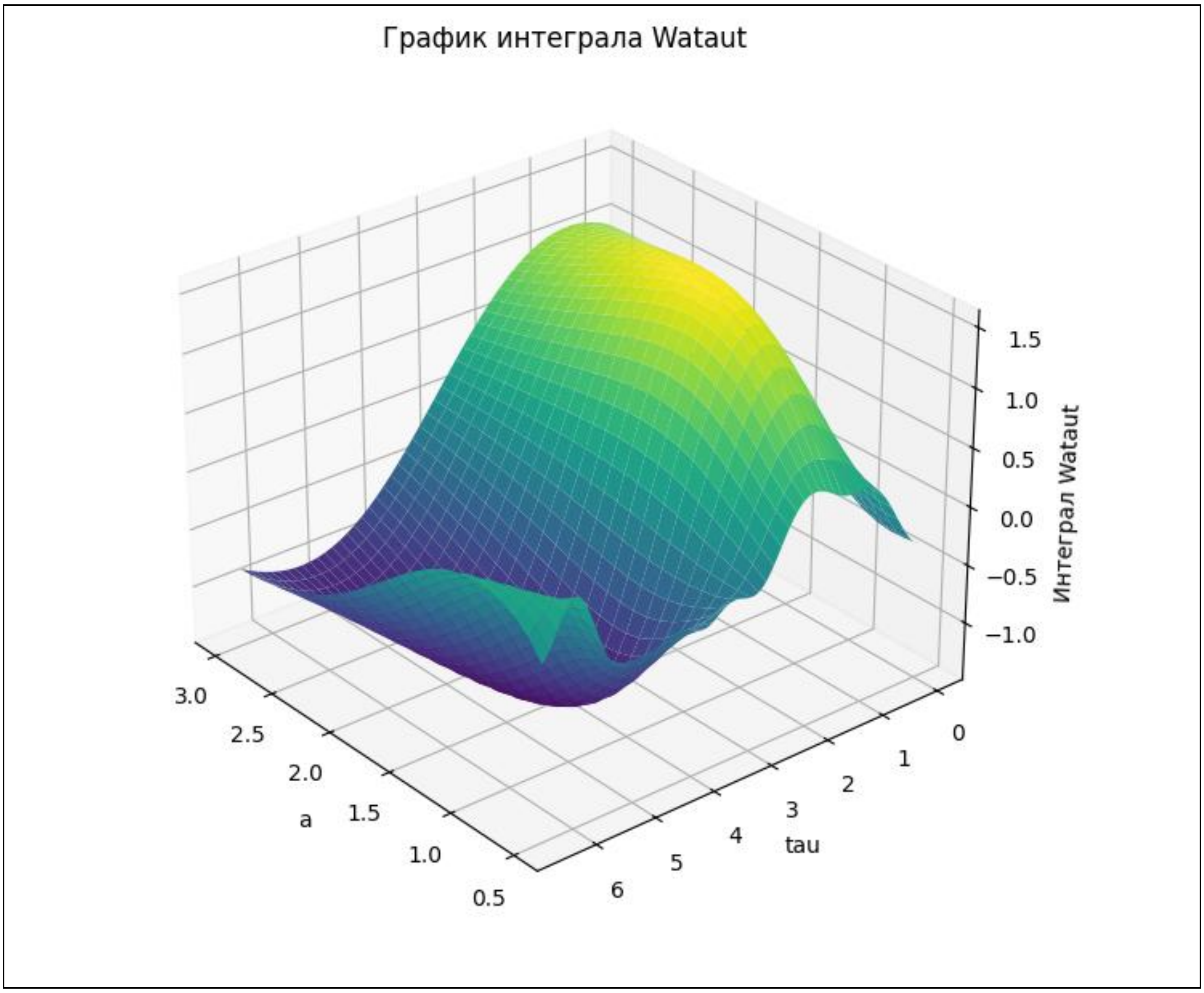
На основе заданного дискретного временного сигнала построить непрерывное вейвлет-преобразование с использованием вейвлет-функции «Мексиканская шляпа».

Теория

Непрерывное вейвлет-преобразование (НВП) — это метод анализа сигналов, который позволяет изучать их частотно-временные характеристики. Вейвлет «Мексиканская шляпа»— это вторая производная функции Гаусса, имеющая характерную форму, напоминающую сомбреро.

Основная идея — свернуть сигнал с масштабированными и сдвинутыми копиями вейвлета, получая коэффициенты, показывающие наличие сигналов с похожей структурой на каждом масштабе.

Решение



Вывод

В результате выполнения лабораторной работы была визуализирована функция W(a, τ, t) с зависимостью от a и τ. Как видно из графика от τ зависит знак графика и то куда она будет стремится, в то время как от a зависит то насколько неровным будет данный график.

Лабораторная работа №4

Задание

Выдан сигнал x(t). Продолжительность сигнала 1 секунда. Необходимо:

- Определить значение коэффициентов C_n, где n=0,1,2,…,7 ряда Хаара.

- Построить функцию $x^*(t) = \sum_{n=0}^{7} C_n\, h_n(t)$.
- Оценить ошибки $\mathcal{E}(t) = |x(t) - x^*(t)|$ и интеграл квадрата ошибки с среднеквадратичным отклонением.

Для функции $x(t) = Asin^2(\pi t)$ для t от 0 до 1, а в других случаях 0. Для A=3.

Теория

Алгоритм Хаара — это простой и эффективный метод вейвлет-преобразования, который разбивает сигнал на грубую аппроксимацию и детализирующие коэффициенты. Он работает по принципу усреднения и разности значений, что позволяет анализировать данные на разных масштабах. Алгоритм работы ряда Хаара:

Разбиение сигнала: исходный сигнал (например, массив чисел) делится на пары соседних значений.

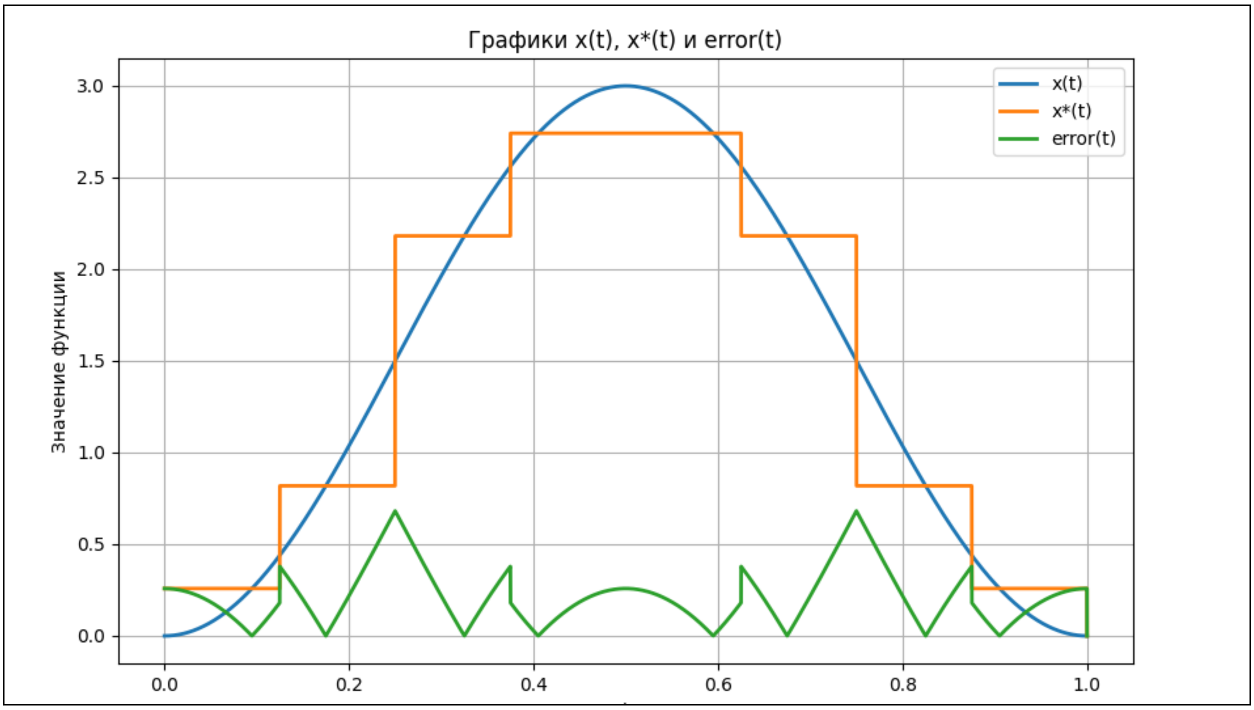
Вычисление средних: для каждой пары находится среднее значение — это аппроксимирующие коэффициенты, которые дают общее представление о сигнале.

Вычисление разностей: для каждой пары также вычисляется разность между значениями и их средним — это детализирующие коэффициенты, которые показывают локальные изменения.

Рекурсивное применение: процесс повторяется для новых усреднённых значений, пока не останется один коэффициент или не будет достигнут нужный уровень детализации.

Решение

```
c_values: [1.50, 0.00, -0.68, 0.68, -0.14, -0.14, 0.14, 0.14]
Интеграл квадрата ошибки: 0.07010449674703255
Среднеквадратичное отклонение: 0.2646661120731577
```



Вывод

В результате выполнения лабораторной работы были найдены значения C_n, благодаря которым был построен график $x^*(t)$ и $\mathcal{E}(t)$ рядом с графиком x(t). Опираясь на график, интеграл квадрата ошибки и среднеквадратичное отклонение можно сказать, что построенный график почти полностью повторяет изначальный график, одна при увеличении количества C_n значения можно было бы и улучшить.

Лабораторная работа №5

Задание

Исследуется подобие временного ряда с использованием вейвлет-функций. Заданы три временных ряда на отрезке [0, 1]:

1) $y_1(\mathrm{t}) = \mathrm{Asin}(2\pi\mathrm{t})$

2) $y_2(\mathrm{t}) = \mathrm{Asin}(2\pi\mathrm{t}) + \mathrm{t}$

3) $y_3(\mathrm{t}) = \mathrm{Asin}(\pi\mathrm{t})$

Задание: оценить степень близости (подобия) временных рядов на отрезке [0, 1].

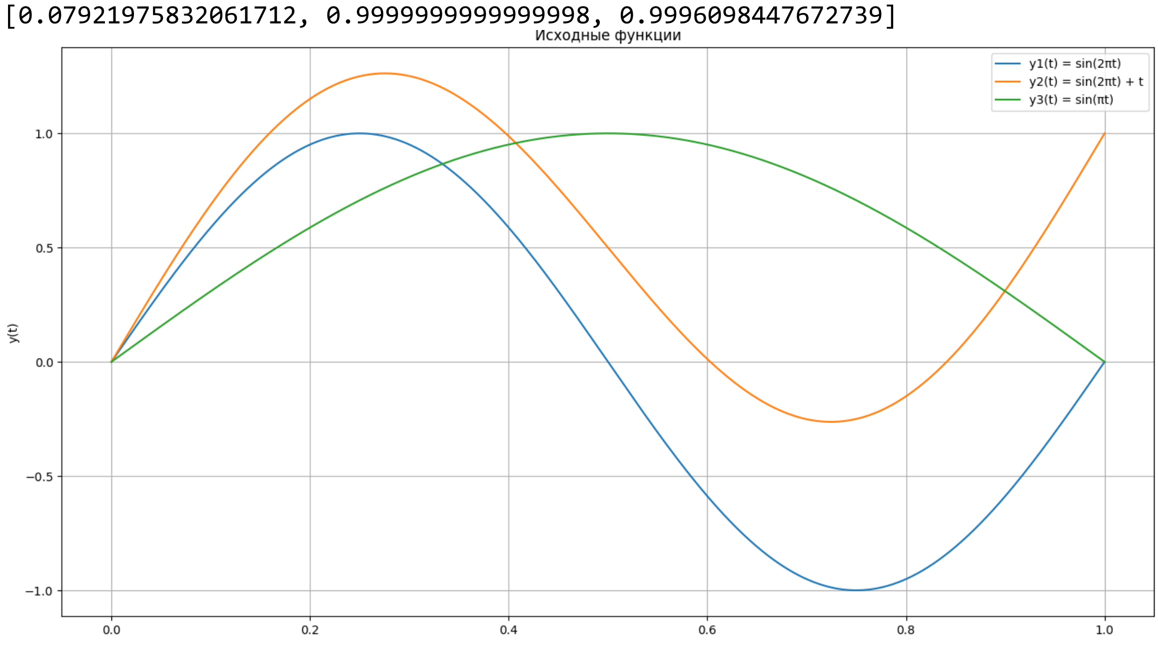
Теория

Вейвлет-анализ — это математический метод изучения сигналов и функций, позволяющий разлагать их на компоненты разной частоты и анализировать в различных временных или пространственных масштабах. В отличие от классического преобразования Фурье, которое дает только частотную информацию, вейвлеты показывают, как меняются частоты во времени.

Основная идея — использование локализованных волнообразных функций (вейвлетов), которые можно масштабировать и сдвигать. Это помогает выявлять резкие изменения, скрытые периодичности и особенности структуры данных.

Решение

```
cos
[0.9207802416793829, 2.220446049250313e-16, 0.0003901552327261193]
MS1
[0.07921975832061712, 0.9999999999999998, 0.9996098447672739]
```



Вывод

В результате выполнения лабораторной работы была написана программа, которая применяет вейвлет преобразование к временным рядам и сравнивает их. Из полученных данных можно сказать, что угол между y_1 и y_2 небольшой, это говорит о том, что функции довольно похожи в терминах их разложения по базисным функциям.

Лабораторная работа №6

Задание

Исследуется подобие зашумленных временных рядов с использованием вейвлет-преобразования. Заданы три временных ряда:

1) $y_1(t) = a\sin(2\pi t) + \xi_1(t)$

2) $y_2(t) = a\sin(2\pi t) + \xi_2(t) + t$

3) $y_3(t) = a\sin(\pi t) + \xi_3(t)$

где $\xi_1(t), \xi_2(t), \xi_3(t)$ – это случайные функции времени:

ξ_1(t)=RND(ξ)*k_1, где $k_1 \in [0,1;0,2]$

ξ_2(t)=RND(ξ)*k_2, где $k_2 \in [0,1;0,3]$

ξ_3(t)=RND(ξ)*k_3, где $k_3 \in [0,1;0,5]$

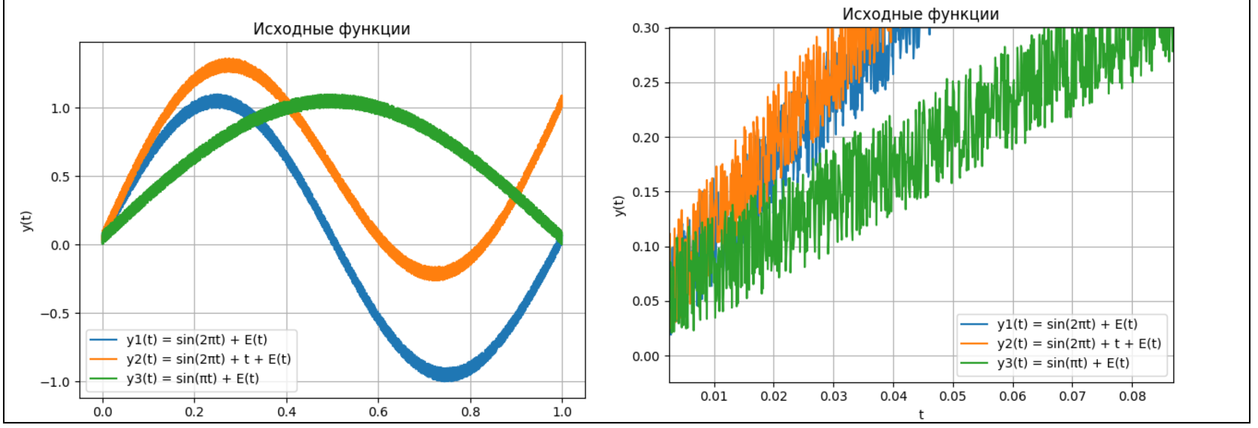
Необходимо сравнить, есть ли у рядов закономерность.

Теория

Временные ряды часто содержат шум — случайные или систематические искажения, мешающие анализу полезного сигнала. Шум может возникать из-за ошибок измерений, внешних помех или несовершенства оборудования. Шум в данных может носить различный характер: высокочастотные случайные колебания, низкочастотные дрейфы или единичные импульсные выбросы. Источниками таких помех часто становятся измерительные приборы, внешние воздействия или особенности самого изучаемого процесса. Для очистки данных применяют различные подходы, включая статистические методы сглаживания, частотную фильтрацию и современные алгоритмы машинного обучения. Особенно эффективными оказываются вейвлет-методы, позволяющие разделить сигнал на составляющие и точечно устранить шумовые компоненты без существенного искажения полезной информации.

Решение

```
cos
[0.9381313680105847, 0.0002958302625621148, 0.0004577572638232025]
MS1
[0.06186863198941528, 0.9997041697374379, 0.9995422427361768]
```



Вывод

В ходе выполнения лабораторной работы была модернизирована лабораторная работа №6 добавлением шума. Из-за шума видно, что график стал более массивным, а значение углов стали удаляться от крайних значений (как и от похожих углов, так и от ортогональных).

Лабораторная работа №7

Задание

Аппроксимируется функция с помощью алгоритма CART. Дана дискретная функция y(x):

В целях аппроксимации функции разбить на контейнеры, при этом:

$$\frac{1}{N}\sum\left(y_i - y_{cp}\right)^2$$

Должно быть минимальным (N – количество элементов в контейнере,y_{cp}– среднее значение по контейнеру). В контейнере должно быть не менее трех точек.

Необходимо построить дерево решения для определения y: $\forall x \in [1,18]$. Необходимо также определить: $y|_{x=4.5}, y|_{x=13.5}, y|_{x=17.2}$.

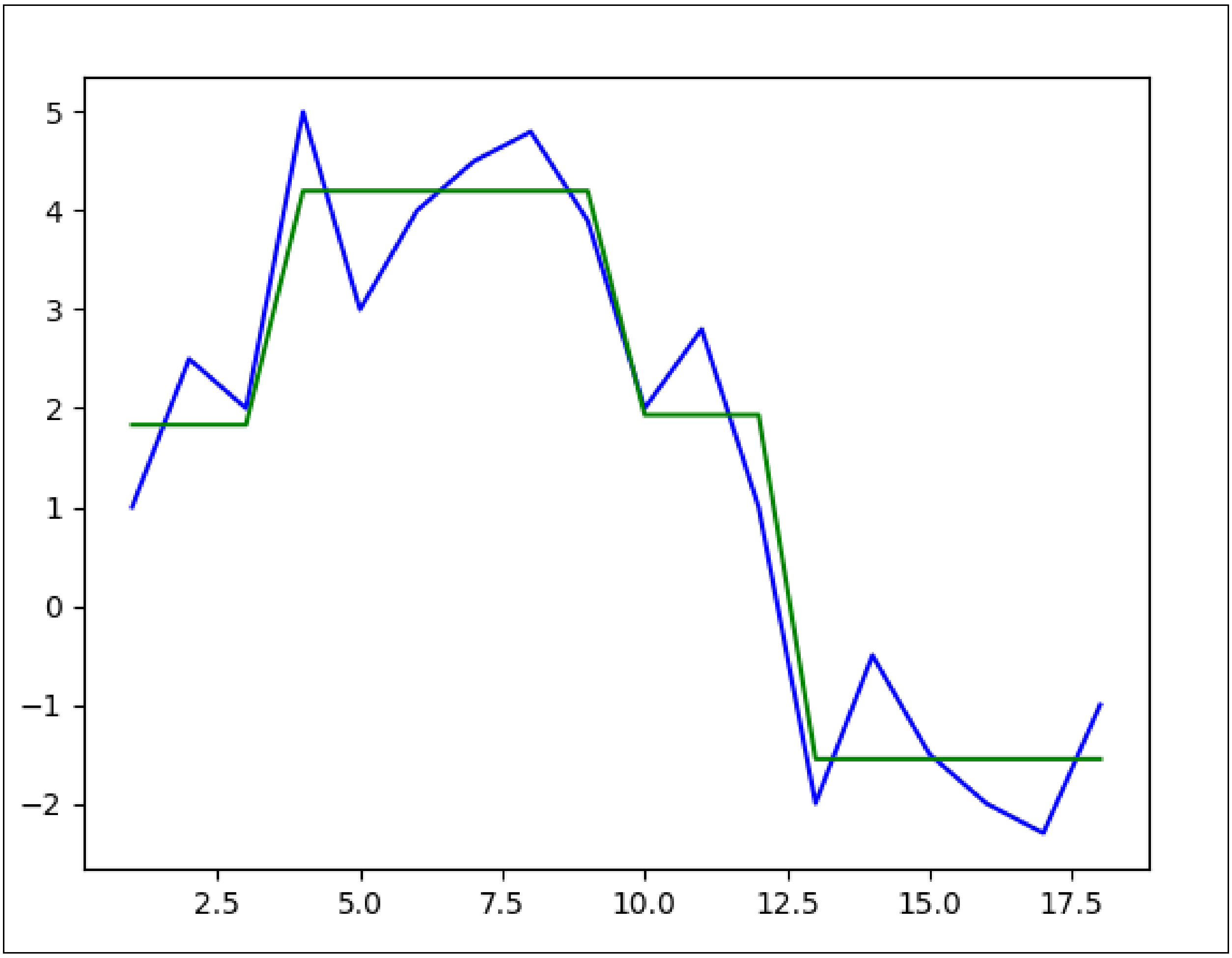
Теория

Алгоритм CART– это метод машинного обучения, который строит деревья решений для задач классификации и регрессии. Его ключевая особенность – бинарное разбиение данных на подмножества с помощью простых условий вида если признак X ≤ порогового значения, то идем в левую ветку, иначе – в правую. Алгоритм начинает работу с корневого узла, содержащего все данные, и последовательно делит их, выбирая на каждом шаге оптимальный признак и порог разбиения, которые максимально уменьшают неоднородность. Процесс продолжается рекурсивно, пока не выполнится критерий остановки – например, достижение максимальной глубины дерева или минимального количества объектов в листе.

Решение

```
Дисперсия - Контейнер:
0.3888888888888889 - [Point(1, 1.0), Point(2, 2.5), Point(3, 2.0)]
0.4433333333333333 - [Point(4, 5.0), Point(5, 3.0), Point(6, 4.0), Point(7, 4.5), Point(8, 4.8), Point(9, 3.9)]
0.5422222222222222 - [Point(10, 2.0), Point(11, 2.8), Point(12, 1.0)]
0.375 - [Point(13, -2.0), Point(14, -0.5), Point(15, -1.5), Point(16, -2.0), Point(17, -2.3), Point(18, -1.0)]

{1.8333333333333333: [Point(1, 1.0), Point(2, 2.5), Point(3, 2.0)], 4.2: [Point(4, 5.0), Point(5, 3.0), Point(6, 4.0), Point(7, 4.5), Point(8, 4.8), Point(9, 3.9)], 1.9333333333333333: [Point(10, 2.0), Point(11, 2.8), Point(12, 1.0)], -1.55: [Point(13, -2.0), Point(14, -0.5), Point(15, -1.5), Point(16, -2.0), Point(17, -2.3), Point(18, -1.0)]}
Предсказания:
[Point(4.5, 4.2), Point(13.5, -1.55), Point(17.2, -1.55)]
```



Вывод

В результате выполнения работы была написана программа, которая строит дерево регрессии методом CART, разбивает функцию на оптимальные контейнеры и вычисляет аппроксимационные значения для заданных точек.

Лабораторная работа №8

Задание

Имеется 20 исходов в виде последовательности шаров (синие и желтые). Выглядит следующим образом

ж	с	с	с	с	ж	ж	ж	ж	с	с	с	с	ж	ж	ж	ж	ж	ж	с

Необходимо разбить на 2 контейнера таким образом, чтобы энтропия и индекс Джинни имели минимальные значения.

Теория

Энтропия и индекс Джинни – это два ключевых критерия, используемые для оценки неоднородности данных в задачах классификации

Энтропия измеряет степень хаоса или неопределенности в системе: чем выше энтропия, тем более разнородны данные. Она вычисляется как сумма произведений вероятностей классов на логарифмы этих вероятностей с отрицательным знаком, стремясь к нулю для чистых узлов (где все объекты одного класса) и достигая максимума при равномерном распределении классов.

Индекс Джинни, напротив, отражает вероятность неправильной классификации случайно выбранного элемента, если его метку предсказывать согласно распределению классов в узле. Он варьируется от нуля (полная однородность) до максимального значения при равновероятных классах.

Решение

```
# Желтый – 1
# Синий - 0

Стандартные значения
Энтропия 0.9927744539878084
Индекс Джинни 0.49499999999999999

Лучшие значения для разбиения на 2 части:
Энтропия: 0.831889265573684
Индекс Джинни: 0.39340659340659345
Левый массив (13 элементов): [1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0]
Правый массив (7 элементов): [1, 1, 1, 1, 1, 1, 0]
Разница Энтропии: 0.16088518841412436
Разница индекса Джинни: 0.10159340659340643

Лучшие значения для разбиения на 3 части:
Энтропия: 0.6248037930698694
Индекс Джинни: 0.3076923076923077
Левый массив (13 элементов): [1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0]
Центральный массив (6 элементов): [1, 1, 1, 1, 1, 1]
Правый массив (1 элементов): [0]
Разница Энтропии: 0.36797066091793895
Разница индекса Джинни: 0.18730769230769218
```

Вывод

В результате выполнения лабораторной работы были найдены лучшее разделение на 2 и 3 контейнера 20 шаров двух цветов, где энтропия и индекс Джинни стремятся к 0.

Лабораторная работа №9

Задание

Для заданного ряда шаров необходимо найти оптимальное разбиение по индексу Джинни на 3 контейнера

к	с	ж	ж	с	с	с	ж	к	к	к	с	с	ж	ж	ж	к	с	к	к

Решение

```
# Красный – 2
# Желтый – 1
# Синий - 0

Стандартные значения
Энтропия 1.5812908992306927
Индекс Джини 0.665

Лучшие значения для разбиения на 3 части:

Лучшие значения для индекса Джини:
Энтропия: 1.1543083163590162
Индекс Джини: 0.49038461538461536
Левый массив (13 элементов): [2, 0, 1, 1, 0, 0, 0, 1, 2, 2, 2, 0, 0]
Центральный массив (3 элементов): [1, 1, 1]
Правый массив (4 элементов): [2, 0, 2, 2]
Разница Энтропии: 0.42698258287167645
Разница индекса Джини: 0.17461538461538467
```

Вывод

В результате выполнения лабораторной работы была доработана лабораторная работа №8, путём добавления большего количества цветов шаров.

Лабораторная работа №10

Задание

Даны два вектора X и Y. Необходимо разбить их на 3 контейнера таким образом, чтобы минимизировать суммарную остаточную сумму квадратов (RSS) по всему вектору и отдельно для каждого контейнера.

X= {1,2,3,4,5}

Y= {1,2,9,10,10}

Теория

RSS – ключевая метрика качества регрессионных моделей, измеряющая совокупное отклонение предсказанных значений от реальных данных. Рассчитывается как сумма квадратов разностей между фактическими и прогнозируемыми значениями целевой переменной, где каждое отклонение возводится в квадрат, чтобы исключить взаимопогашение положительных и отрицательных ошибок и усилить влияние крупных расхождений. Чем меньше значение RSS, тем точнее модель описывает данные, при нулевом значении достигая идеальной аппроксимации.

Решение

```
Минимальный RSS: 0.5
Контейнеры:
Левый: [1, 2]
Центральный: [9]
Правый: [10, 10]
```

Вывод

В результате выполнения лабораторной работы была написана программа, которая находит наименьшее значений RSS для 3 контейнеров.

Лабораторная работа №11

Задание

Имеются следующие статистические данные по представителям физлиц. Были выделены 4 фактора (X_1 – возраст, X_2 – доход, X_3 – семья, X_4 – имеющаяся собственность). Y -просрочка.

Просрочка – это ситуация, когда заёмщик не вовремя вернул деньги.

№	X_1 лет	X_2 т. р.	X_3	X_4 млн. р.	Y
1	46	50	2	2	Нет
2	36	42	2	1.5	Нет
3	34	40	3	1.4	Нет
4	28	45	1	0.5	Да
5	24	30	0	0.3	Да
6	21	25	0	0.1	Да
7	21	25	0	0.1	Да
8	35	40	1	0.5	Нет
9	48	45	2	0.35	Нет
10	37	54	2	0.45	Нет
11	18	25	0	5	Да
12	24	30	1	0.4	Да
13	33	45	1	3	Нет
14	45	50	1	4	Да
Тестовое множество					
15	38	50	2	2.5	Нет
16	24	25	1	5	Да
17	42	30	3	4	Нет
18	23	35	0	3	Нет
19	32	35	0	1.5	Да
20	40	45	2	4.5	Нет

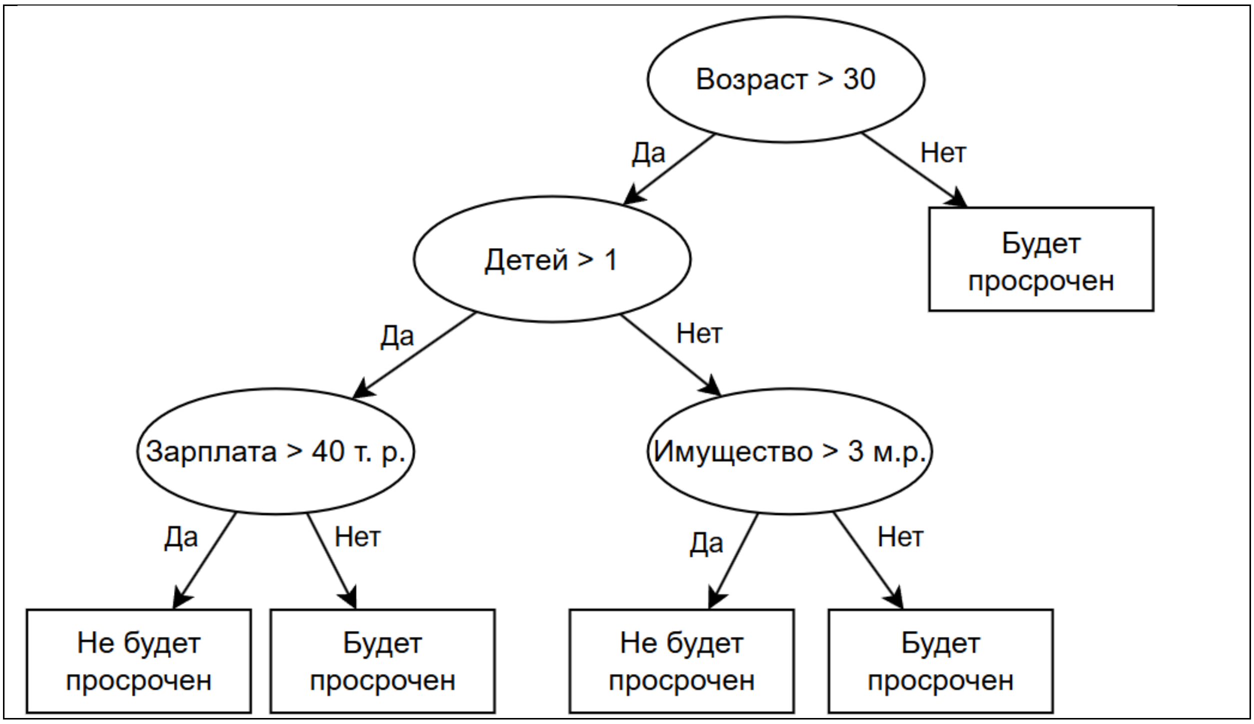
Необходимо:

1. Составить дерево на основе статистики
2. Проверить дерево на основе тестового множества
3. Узнать стоит ли выдавать кредит человеку с таким показателями:
   a. Возраст – 35 лет;
   b. Семья – 2 ребенка;
   c. Доход – 30 тысяч рублей;
   d. Собственность – 5 миллионов рублей.

Теория

Дерево решений — это алгоритм классификации и регрессии, который строит модель в виде древовидной структуры, где каждый внутренний узел соответствует проверке одного из признаков, ветви — исходам проверки, а листья — конечным решениям (классам или числовым значениям). Этот метод широко применяется в кредитном скоринге, медицинской диагностике и бизнес-аналитике благодаря своей интерпретируемости.

Решение



```
Для 1: Предсказано 0, Факт 0
Для 2: Предсказано 0, Факт 0
Для 3: Предсказано 1, Факт 0
Для 4: Предсказано 1, Факт 1
Для 5: Предсказано 1, Факт 1
Для 6: Предсказано 1, Факт 1
Для 7: Предсказано 0, Факт 0
Для 8: Предсказано 0, Факт 0
Для 9: Предсказано 1, Факт 1
Для 10: Предсказано 0, Факт 0
Для 11: Предсказано 1, Факт 1
Для 12: Предсказано 1, Факт 1
Для 13: Предсказано 0, Факт 0
Для 14: Предсказано 1, Факт 1

Точность: 13/14 = 92.9%


Для 15: Предсказано 0, Факт 0
Для 16: Предсказано 1, Факт 1
Для 17: Предсказано 1, Факт 0
Для 18: Предсказано 0, Факт 0
Для 19: Предсказано 1, Факт 1
Для 20: Предсказано 0, Факт 0

Точность: 5/6 = 83.3%


Для 21: Предсказано 1, Факт 1

Точность: 1/1 = 100.0%
```

Вывод

В результате выполнения лабораторной работы было построено дерево, которое по существующим критериям выдаёт максимальную точность на основном множестве и на тестовом. По созданному дереву человеку не стоит выдавать кредит.

Лабораторная работа №12

Задание

Используя KNN необходимо оценить вероятность кредита для человека с такими параметрами:

Возраст – 35 лет;

Семья – 2 ребенка;

Доход – 30 тысяч рублей;

Собственность – 5 миллионов рублей.

Другие данные брать из лабораторной работы №11.

Теория

KNN – один из методов, который широко практикуется в Data Science. Он может быть использован как в задачах классификации, так и в задачах регрессии.

Особенность: не смотря на необходимость иметь Dataset определённого объёма. Процедура обучения в нём практически отсутствует. По этой причине алгоритм называют ленивым (Lazy).

Решение

```
Степень метрики Минковского: 2
Количество соседей: 7

Без весовой KNN
Кредит будет просрочен
Вероятность возврата кредита без задержек: 0.43

Весовой линейный KNN
Кредит будет сдан вовремя
Коэффициент сдачи вовремя: 2.857142857142857
Коэффициент просрочки: 2.142857142857143
```

Вывод

В результате выполнения лабораторной работы была написана метрика Минковского (степень = 2, что соответствует Евклидовой метрикой) и метода k-ближайших соседей показал противоречивые прогнозы в зависимости от типа KNN:

Без весовой функции модель предсказывает просрочку кредита с вероятностью возврата без задержек всего 43%, что указывает на высокий риск.

С весовой функцией (линейные веса) прогноз меняется на своевременный возврат, так как суммарный вес соседей с классом вовремя превышает вес просрочки

Лабораторная работа №13

Задание

Компания в целях минимизации логистических точек планирует распределить 200 торговых точек по кластерам таким образом, чтобы каждая из точек продаж управлялся из офиса этого кластера.

Исходные данные:

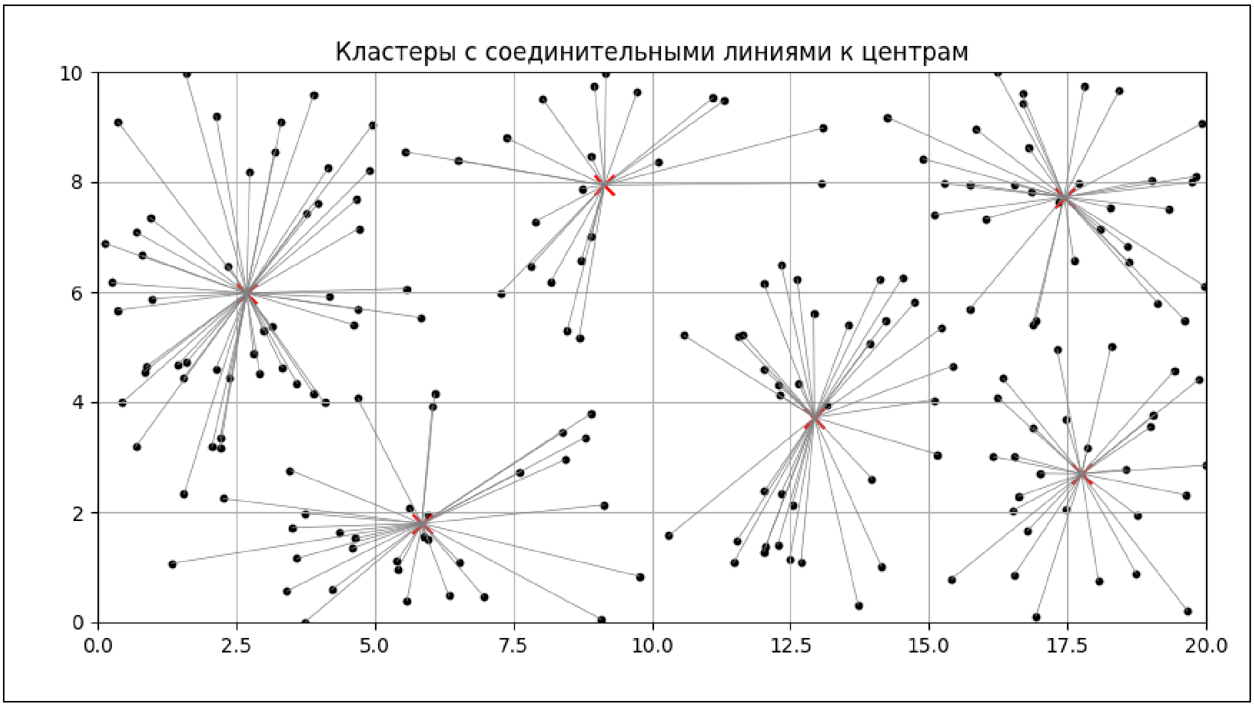
200 точек, 6 центров офисов. Размещены на поле 10 на 20.

Необходимо определить координаты офисов таким образом, чтобы расстояние до офисов в каждом классе была минимальна.

Теория

Алгоритм k-средних — это метод кластеризации, который разделяет данные на k кластеров, минимизируя суммарные квадраты расстояний между точками и центроидами кластеров. Процесс начинается с случайного выбора k начальных центроидов, после чего итеративно выполняются два шага: на этапе назначения каждая точка приписывается к ближайшему центроиду, а на этапе обновления центроиды пересчитываются как средние значения всех точек кластера. Алгоритм завершается, когда центроиды стабилизируются или достигается максимальное число итераций. Основные преимущества включают простоту и скорость работы, но метод чувствителен к начальному выбору центроидов и может сходиться к локальным оптимумам.

Решение



Вывод

В ходе выполнения лабораторной работы была написана программа, которая находит лучшие значения для минимизации логистических расходов путём расстановки центров в наиболее лучшие позиции относительно точек продаж.