

МЕЖГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«БЕЛОРУССКО-РОССИЙСКИЙ УНИВЕРСИТЕТ»

Кафедра «Автоматизированные системы управления»

МОДЕЛИ И МЕТОДЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

*Методические рекомендации к лабораторным работам
для студентов специальности
1-40 80 02 «Системный анализ, управление и обработка
информации»
очной и заочной форм обучения*



Могилев 2019

УДК 004.891
ББК 32.973.202
М74

Рекомендовано к изданию
учебно-методическим отделом
Белорусско-Российского университета

Одобрено кафедрой «Автоматизированные системы управления»
«15» октября 2019 г., протокол № 3

Составители: доц. И. В. Акиншева;
ст. преподаватель Д. А. Денисевич

Рецензент Ю. С. Романович

Даны методические рекомендации по выполнению лабораторных работ по дисциплине «Модели и методы поддержки принятия решений» (1 семестр), а также приведены задания к ним и список литературы для подготовки.

Учебно-методическое издание

МОДЕЛИ И МЕТОДЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Ответственный за выпуск	А. И. Якимов
Редактор	А. А. Подошевка
Компьютерная верстка	Н. П. Полевничая

Подписано в печать . Формат 60×84/16. Бумага офсетная. Гарнитура Таймс.
Печать трафаретная. Усл. печ. л. . Уч.-изд. л. . Тираж 31 экз. Заказ № .

Издатель и полиграфическое исполнение:
Межгосударственное образовательное учреждение высшего образования
«Белорусско-Российский университет».
Свидетельство о государственной регистрации издателя,
изготовителя, распространителя печатных изданий
№ 1/156 от 07.03.2019.
Пр-т Мира, 43, 212022, Могилев.

© Белорусско-Российский
университет, 2019



Содержание

Введение.....	4
Лабораторная работа № 1. Нахождение множества Парето.....	5
Лабораторная работа № 2. Метод анализа иерархий	10
Лабораторная работа № 3. Соотношение суммарных затрат и точности для различных вариантов детализации прикладной модели	14
Лабораторная работа № 4. Построение комплексной модели	18
Лабораторная работа № 5. Реализация теоретико-игрового подхода..	25
Лабораторная работа № 6. Управление рисками	29
Лабораторная работа № 7. Нечеткие экспертные системы	34
Лабораторная работа № 8. Кластеризация данных.....	40
Список литературы	47



Введение

Целью преподавания дисциплины «Модели и методы поддержки принятия решений» у магистрантов является формирование теоретических знаний в области принятия управленческих решений, ознакомление с подходами принятия решений при выполнении практических задач.

Даны методические рекомендации по выполнению лабораторных работ по дисциплине «Модели и методы поддержки принятия решений», а также приведены задания к ним и список литературы для подготовки.



Лабораторная работа № 1. Нахождение множества Парето

Цель работы: освоить способ нахождения множества Парето для определения наилучшей альтернативы.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Получить задание у преподавателя.
- 3 Реализовать задание.
- 4 Дать обоснование полученного решения.
- 5 Сделать выводы по результатам исследований.
- 6 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Поиск решения многокритериальной задачи не представляет особых сложностей, если предпочтение по одному критерию влечет за собой такое же предпочтение по другому критерию, т. е. критерии кооперируются. Например, когда при покупке автомобиля мы преследуем цель купить престижный и красивый автомобиль. Очень часто эти два критерия совпадают и престижный автомобиль является одновременно красивым.

Решение многокритериальной задачи также не представляет особых сложностей, если критерии нейтральны по отношению друг к другу, т. е. поиск решения по одному критерию никаким образом не отражается на поиске решения по другому критерию. Например, когда при покупке автомобиля мы преследуем цель купить надежный и красивый автомобиль.

Приведенные примеры являются частными случаями. Чаше критерии конкурируют друг с другом. В большинстве практических задач поиск более предпочтительного решения по одному критерию приводит к тому, что решение становится менее предпочтительным по другому критерию, т. е. решения несравнимы между собой. Например, рассматривая стоимость и престижность в качестве критериев при покупке автомобиля, можно утверждать, что более дешевый (более предпочтительный по первому критерию) автомобиль является менее престижным (менее предпочтительным по второму критерию).

Анализ таких ситуаций может быть осуществлен при помощи определения множества Парето.



Предположим, что при оценке альтернатив использовались два критерия: стоимость C_1 и надёжность C_2 . Значения критериев для различных альтернатив представлены в таблице 1.

Таблица 1 – Оценка трех альтернатив по двум критериям

Альтернатива	Критерий	
	Стоимость	Надёжность
1	Малая	Низкая
2	Большая	Высокая
3	Малая	Высокая

Здесь области первого и второго критериев соответственно $O_1 = \{\text{Малая, Большая}\}$ и $O_2 = \{\text{Низкая, Высокая}\}$, область альтернатив $A = \{1, 2, 3\}$.

Очевидно, что альтернатива 3 является наиболее предпочтительной, т. к. она не хуже остальных альтернатив по всем критериям.

Альтернатива α_i является доминирующей по отношению к альтернативе α_k , если по всем критериям оценки альтернативы α_i не хуже, чем альтернативы α_k , а хотя бы по одному критерию оценка α_i лучше. При этом альтернатива α_k называется доминируемой. Из определения следует, что альтернатива 3 из вышеприведенного примера является доминирующей по отношению к альтернативе 1 и альтернативе 2. Это можно увидеть из рисунка 1, где альтернатива 3 занимает самое правое и верхнее положение по отношению к другим альтернативам.

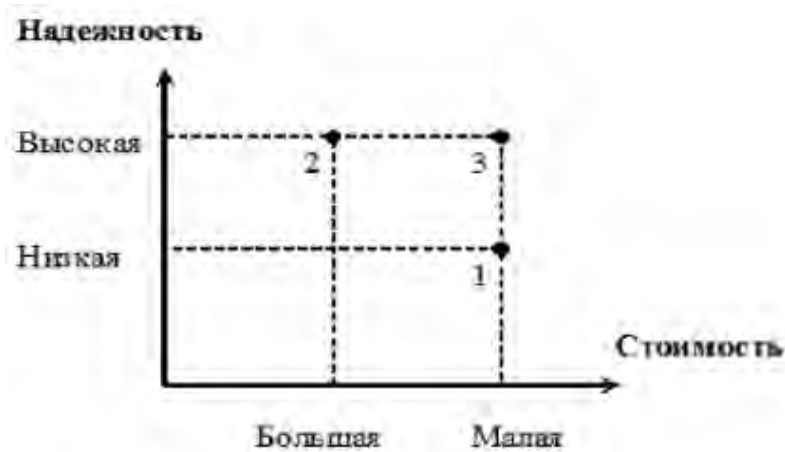


Рисунок 1 – Альтернативы и их положения в соответствии с критериями

Рассмотрим альтернативы 1 и 2. Из рисунка 1 следует, что альтернативы 1 и 2 не находятся в отношении доминирования. По стоимости предпочтительнее альтернатива 1, а по надежности – альтернатива 2. Эти альтернативы являются несравнимыми по отношению предпочтения между векторными оценками, т. к. их невозможно сравнить непосредственно на основе критериальных оценок.

Альтернатива α_i , для которой не существует другой альтернативы α_k , лучшей по всем критериям одновременно, т. е. каждая из них превосходит любую другую по какому-то из критериев, называется недоминируемой, или оптимальной по Парето.

Множество всех альтернатив называется множеством Парето.

Если вернуться к примеру, то оставшиеся альтернативы 1 и 2 принадлежат множеству Парето.

Принцип Парето заключается в том, что оптимальный исход следует искать только среди элементов множества недоминируемых решений $P\alpha C(A)$. Однако это условие справедливо, если выполняется аксиома Парето, устанавливающая «рациональное» поведение лица, принимающего решение, т. е. его стремление получить по возможности большие значения всех компонент векторного критерия C .

Множество Парето также называют множеством не улучшаемых решений. Парето-оптимальность решения α^* означает, что оно не может быть улучшено ни по одному из критериев без ухудшения по какому-нибудь другому критерию.

Так, в примере с оставшимися двумя альтернативами без дополнительной качественной или количественной информации нельзя выделить одну оптимальную альтернативу. Поэтому при поиске одной наиболее предпочтительной альтернативы необходимы дополнительные сведения о критериях или предпочтениях, которые бы смогли уменьшить множество Парето.

Альтернатива называется слабо эффективным решением, если оно не может быть улучшено сразу по всем m критериям.

Рассмотрим пример оценки кандидатов на некоторую должность с использованием трехбалльной системы $O_1 = O_2 = \{1, 2, 3\}$ по двум критериям: образование C_1 и опыт работы C_2 . Результаты оценивания пяти кандидатов представлены в таблице 2.

Таблица 2 – Результаты оценивания кандидатов на должность

Кандидат	Критерий	
	Образование	Опыт
1	1	3
2	2	2
3	2	1
4	1	2
5	3	1

Точки на плоскости на рисунке 2 являются векторными оценками кандидатов. Из этого рисунка видно, что оптимальными по Парето являются кандидаты с номерами 1, 2, 5; кандидаты с номерами 4 и 3 – доминируемые, т. к. угол с вершиной в точке 4 или 3 содержит по две точки из множества недоминируемых точек 1, 2, 5.



Аналогичные углы для точек 1, 2, 5 не содержат точек, соответствующих другим кандидатам. В то же время точки 4 и 3 являются слабо эффективными решениями задачи, так как не существует альтернатив, соответствующих точкам с координатами (2, 3), (3, 3) или (3, 2).

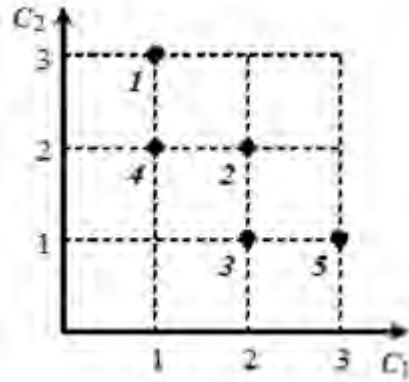


Рисунок 2 – Точки, соответствующие векторным оценкам пяти кандидатов на должность

Из приведенного примера видно, что только привлечением дополнительной информации можно сузить множество Парето.

Сужение множества альтернатив до множества эффективных решений важно не только само по себе, но еще и потому, что на более узком подмножестве могут выполняться различного рода упрощающие дальнейший анализ допущения о предпочтениях, которые заведомо несправедливы для множества решений. Кроме того, эффективные решения могут обладать интересными и практически важными свойствами, не присущими остальным решениям.

Таким образом, решение многокритериальной задачи сводится к следующим основным составляющим:

- определению множества не улучшаемых решений Парето;
- получению дополнительной информации о критериях в том или ином виде;
- использованию дополнительной информации о критериях для сужения множества Парето до тех пор, пока это множество не будет содержать только одну альтернативу или группу альтернатив.

Пример выполнения практического задания.

Дано множество (альтернатив) системы «Ноутбук», а также основные их (критерии) характеристики. Исследуемые альтернативы и их характеристики представлены в таблице 3.

Проанализировав данные таблицы 3 и сравнив попарно параметры альтернатив, можно сделать вывод, что паретовское множество составляют две альтернативы: альтернатива Toshiba Satellite A660-10X и альтернатива HP Pavilion dv-4045er.

Таблица 3 – Исследуемые альтернативы и их характеристики

Наименование критерия	Toshiba Satellite A660-10X	HP Pavilion dv-4045er	HP Pavilion dv-7-4120er
1 Количество ядер процессора	4	4	2
2 Тактовая частота, ГГц	1,6	1,6	1,6
3 Диагональ экрана, дюйм	16	17,3	17,3
4 Объем оперативной памяти, Мбайт	4096	4096	4096
5 Ёмкость жесткого диска, Гбайт	640	640	500
6 Вес, кг	2,62	3,06	3,03
7 Количество активных пикселей камеры, Мпикс	1,3	1,3	1,3
8 Стоимость, у. е.	1300	1350	1280

Порядок выполнения лабораторной работы.

- 1 Изучить теоретическую часть данной лабораторной работы.
- 2 Определить существенные критерии для оценки заданных альтернатив.
- 3 Определить и установить заданные величины критериев.
- 4 Парно сравнить исследуемые альтернативы и все худшие по всем критериям отбросить, а все оставшиеся и несравнимые между собой включить в паретовское множество.

Варианты систем для выполнения практического задания:

- | | | |
|-----------------|----------------|---------------------------|
| 1) компьютер; | 5) видеокарта; | 9) фотоаппарат; |
| 2) печь СВЧ; | 6) монитор; | 10) телевизор; |
| 3) самолет; | 7) телефон; | 11) любая другая система. |
| 4) холодильник; | 8) автомобиль; | |

Содержание отчета.

- 1 Титульный лист.
- 2 Цель работы.
- 3 Исходные данные.
- 4 Теоретические сведения.
- 5 Ход выполнения работы.
- 6 Выводы.

Контрольные вопросы

- 1 Дайте определение способа нахождения паретовского множества.
- 2 В чем заключается принцип Парето?
- 3 Что называют неулучшаемыми решениями?
- 4 Что называют альтернативой?
- 5 Что называют слабо эффективным решением?
- 6 В чем суть сужения множества альтернатив?



Лабораторная работа № 2. Метод анализа иерархий

Цель работы: изучить метод анализа иерархий (метод Саати).

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Получить задание у преподавателя.
- 3 Реализовать задание.
- 4 Дать обоснование полученного решения.
- 5 Сделать выводы по результатам исследований.
- 6 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Принятие решений в условиях определенности характеризуется наличием однозначной или детерминированной связи между принятым решением и его исходом. В наиболее простых случаях каждой альтернативе удастся поставить в соответствие определенный исход, оцениваемый некоторым числом – «полезностью» исхода. При этом возможно установить прямую связь: альтернатива – численное значение соответствующего ей исхода, минуя сами исходы. Эта связь описывается целевой скалярной функцией, определенной на множестве альтернатив. Оптимальным решением считают ту альтернативу, которая доставляет целевой функции наибольшее (наименьшее) значение. Иначе говоря, нахождение оптимального решения равно сильно нахождению экстремума функции при некоторых условиях. Рассмотрим подход к принятию решений в ситуациях, когда, например, для идей, чувств, эмоций определяются некоторые количественные показатели, обеспечивающие числовую шкалу предпочтений для возможных альтернативных решений – метод анализа иерархий.

Определение весовых коэффициентов.

Сложность метода анализа иерархий заключается в определении относительных весовых коэффициентов для оценки альтернативных решений. Если имеется n критериев на заданном уровне иерархии, соответствующая процедура создает матрицу A размерности $n \times n$, именуемую матрицей парных сравнений, которая отражает суждение лица, принимающего решение, относительно важности разных критериев. Парное сравнение выполняется таким образом, что критерий в строке i ($i = 1, 2, \dots, n$) оценивается относительно каждого из критериев, представленных n столбцами. Обозначим через a_{ij}



элемент матрицы A , находящийся на пересечении i -й строки и j -го столбца. В соответствии с методом анализа иерархий для описания упомянутых оценок используются целые числа от 1 до 9. При этом $a_{ij} = 1$ означает, что i -й и j -й критерии одинаково важны, $a_{ij} = 5$ отражает мнение, что i -й критерий значительно важнее, чем j -й, а $a_{ij} = 9$ указывает, что i -й критерий чрезвычайно важнее j -го. Другие промежуточные значения между 1 и 9 интерпретируются аналогично. Согласованность таких обозначений обеспечивается следующим условием: если $a_{ij} = k$, то автоматически $a_{ji} = 1/k$. Кроме того, все диагональные элементы матрицы A должны быть равны 1, так как они выражают оценку критерия относительно самих себя.

Относительные веса w (весовые коэффициенты) вычисляются в виде средних значений элементов соответствующих строк нормализованной матрицы N . Для получения нормализованной матрицы делим элементы каждого столбца матрицы сравнений A на сумму элементов этого же столбца.

Оценка альтернативных решений основана на вычислении комбинированного весового коэффициента для каждого из них:

$$F_i = \sum_{j=1}^n w_j w_{ji}; \quad j = \overline{1, n}, i = \overline{1, m},$$

где n – количество критериев;

m – количество альтернативных решений.

Альтернативное решение, имеющее наивысший комбинированный вес, является наиболее оптимальным выбором.

Согласованность матрицы сравнений. Если все столбцы нормализованной матрицы идентичны, то исходная матрица сравнения является *согласованной*. Согласованность означает, что решение будет согласовано с определениями парных сравнений критериев или альтернатив. С математической точки зрения согласованность матрицы A означает, что $a_{ij}a_{jk} = a_{ik}$ для всех i, j и k .

Если матрица парных сравнений не является согласованной, необходимо выяснить, является ли уровень согласованности «допустимым». Для этого необходимо определить соответствующую количественную меру, т. е. коэффициент согласованности.

Коэффициент согласованности находится по формуле

$$CR = \frac{CI}{RI},$$

где CI – коэффициент согласованности матрицы сравнений, $CI = \frac{n_{max} - n}{n - 1}$;
стохастический коэффициент согласованности матрицы, $RI = \frac{1.98(n-2)}{n}$.

Стохастический коэффициент согласованности RI определяется эмпирическим путем как среднее значение коэффициента CI для большой выборки генерированных случайным образом матриц сравнения A .



Коэффициент согласованности CR используется для проверки согласованности матрицы сравнения A следующим образом. Если $CR < 0,1$, уровень несогласованности является приемлемым. В противном случае уровень несогласованности матрицы сравнения A является высоким и лицу, принимающему решение, рекомендуется проверять элементы парного сравнения a_{ij} матрицы A в целях получения более согласованной матрицы.

Значение n_{\max} можно определить путем вычисления вектор-столбца $A \bar{w}$ (\bar{w} – вектор относительных весов) с последующим суммированием его элементов.

Задание

Вариант 1. Абсолютные показатели качества двигателей различных вариантов приведены в таблице 1. Найти оптимальный вариант двигателя.

Таблица 1 – Абсолютные показатели качества двигателей

Вариант двигателя	Показатель качества		
	Мощность, л. с.	Крутящий момент, кгс·м	Масса, кг
1	180	67	850
2	176	70	1000
3	176	68	860
4	181	67	820
5	177	68	860
6	180	66	800

Вариант 2. Показатели эффективности работы предприятий приведены в таблице 2. Выбрать наиболее эффективно работающее предприятие.

Таблица 2 – Показатели эффективности работы предприятий

Номер предприятия	Показатель эффективности работы предприятий		
	Прибыль, д. е.	Себестоимость единицы продукции, д. е.	Доходы, д. е.
1	30	40	20
2	25	20	30
3	40	45	54
4	28	30	35
5	15	12	20
6	50	30	40

Вариант 3. Абсолютные показатели качества двигателей различных вариантов приведены в таблице 3. Найти оптимальный вариант двигателя.

Таблица 3 – Абсолютные показатели качества двигателей

Вариант двигателя	Показатель качества		
	Мощность, л. с.	Крутящий момент, кгс·м	Масса, кг
1	30	40	20
2	25	20	30
3	40	45	54
4	28	30	35
5	15	12	20
6	50	30	40

Контрольные вопросы

- 1 Перечислите основные этапы метода анализа иерархий.
- 2 Опишите процесс попарного сравнения объекта по какому-либо признаку.
- 3 Перечислите основные свойства матрицы попарных сравнений.
- 4 Как происходит формирование вектора локальных приоритетов?
- 5 Можно ли отнести метод анализа иерархий к методам экспертных оценок?



Лабораторная работа № 3. Соотношение суммарных затрат и точности для различных вариантов детализации прикладной модели

Цель работы: освоить основные принципы моделирования и разобрать различные варианты их детализации.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Ответить на контрольные вопросы.
- 3 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Общие свойства моделей.

Адекватность – это степень соответствия модели тому реальному явлению (объекту, процессу), для описания которого она строится.

Конечность – модель отображает оригинал лишь в конечном числе его отношений и, кроме того, ресурсы моделирования конечны.

Упрощенность – модель отображает только существенные стороны объекта.

Полнота – учтены все необходимые свойства.

Приблизительность – действительность отображается моделью грубо или приблизительно.

Информативность – модель должна содержать достаточную информацию о системе – в рамках гипотез, принятых при построении модели.

Потенциальность – предсказуемость модели и ее свойств.

Степень соответствия модели реальному объекту.

Так как моделирование – способ замещения реального объекта его аналогом, то возникает вопрос: насколько она (модель) должна соответствовать исходному объекту?

Рассмотрим предельные варианты.

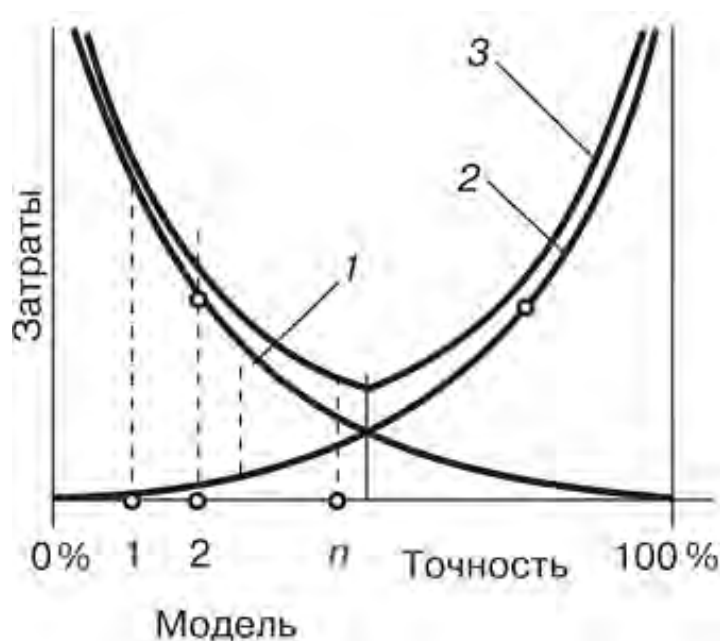
Вариант 1 – соответствие 100 %. Очевидно, что точность решения в этом случае максимальна, а ущерб от применения модели минимален. Но затраты на построение такой модели бесконечно велики, т. к. объект повторяется во всех



своих деталях; фактически создается точно такой же объект путем копирования его до атомов (что само по себе не имеет смысла).

Вариант 2 – соответствие 0 %. Модель совсем не похожа на реальный объект. Очевидно, что точность решения минимальна, а ущерб от применения модели максимален, бесконечен. Но затраты на построение такой модели нулевые.

Реальная модель — компромисс между затратами на ее построение и ущербом от неточности ее применения. Это точка между двумя бесконечностями. То есть, моделируя, следует иметь в виду, что исследователь должен стремиться к оптимуму суммарных затрат, включающих ущерб от применения и затраты на изготовление модели (рисунок 1).



1 – ущерб от использования модели; 2 – затраты на построение модели; 3 – общие затраты

Рисунок 1 – Соотношение суммарных затрат и точности для различных вариантов детализации прикладной модели

Из графика видно, что неточные модели не нужны, но и абсолютная точность невозможна и не нужна. Частое и распространенное заблуждение при построении моделей – требовать «как можно точнее». Модель – компромисс между точностью модели и затратами на ее реализацию. В модель включаются только существенные аспекты, представляющие объект, и отбрасываются все остальные (бесконечное большинство). Существенный или несущественный аспект описания определяют согласно цели исследования. То есть каждая модель составляется с какой-то целью. Начиная моделирование, исследователь должен определить цель, отделив ее от всех возможных других целей, число которых, по-видимому, бесконечно.

Однако построение кривой, показанной на рисунке 1, до начала моделирования невозможно. Поэтому на практике действуют таким образом:

двигаются по шкале точности слева направо, т. е. от простых моделей («Модель 1», «Модель 2», ...) ко все более сложным («Модель 3», «Модель 4», ...). Процесс моделирования имеет циклический спиралевидный характер: если построенная модель не удовлетворяет требованиям точности, то ее детализируют, дорабатывают на следующем цикле (рисунок 2).

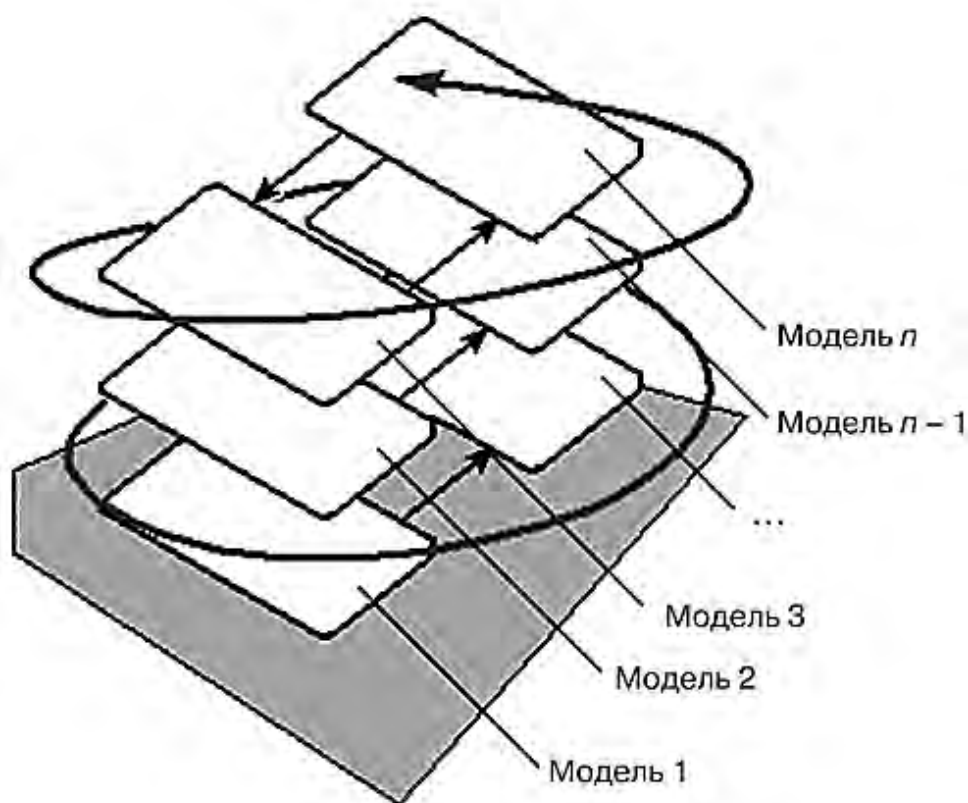


Рисунок 2 – Спиралевидный характер процесса проектирования и уточнения прикладных моделей

При улучшении модели необходимо следить за балансом «затраты – эффект». Как только затраты начинают превалировать над эффектом, следует остановиться, поскольку точка оптимума достигнута. Такой подход всегда гарантирует окупаемость вложений.

Из всего сказанного следует, что моделей может быть несколько: приближенная, более точная, еще точнее и т. д. Модели как бы образуют ряд. Двигаясь от варианта к варианту, исследователь совершенствует модель. Для построения и совершенствования моделей необходима их преемственность, средства отслеживания версий и т. д., т. е. моделирование требует инструмента и опирается на технологию.

Однако из всего вышесказанного следует, что создание ряда физических моделей далеко не всегда экономически оправдано. Поэтому в настоящее время исследователи все чаще прибегают к помощи электронно-вычислительной техники, создают и исследуют компьютерные модели.

Задание

Привести два-три собственных примера моделей. Объяснить, для какой задачи они могут быть использованы. Аргументировать выбранные критерии для моделей.

Контрольные вопросы

- 1 Дайте определения основным свойствам модели.
- 2 Опишите основные степени соответствия модели реальному объекту.
- 3 Дайте определение понятию «реальная модель».



Лабораторная работа № 4. Построение комплексной модели

Цель работы: построить информационную комплексную модель информационной системы с помощью методологии IDEF1X.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Выполнить практическое задание.
- 3 Сделать выводы по полученным результатам.
- 4 Ответить на контрольные вопросы.
- 5 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Метод IDEF1, разработанный Т. Рэмей (Т. Ramey), также основан на подходе П. Чена и позволяет построить модель данных, эквивалентную реляционной модели в третьей нормальной форме. В настоящее время на основе совершенствования методологии IDEF1 создана ее новая версия – методология IDEF1X. IDEF1X разработана с учетом таких требований, как простота изучения и возможность автоматизации. IDEF1X-диаграммы используются рядом распространенных CASE-средств (в частности, ERwin, Design/IDEF).

Сущность в методологии IDEF1X является независимой от идентификаторов или просто независимой, если каждый экземпляр сущности может быть однозначно идентифицирован без определения его отношений с другими сущностями.

Сущность называется зависимой от идентификаторов или просто зависимой, если однозначная идентификация экземпляра сущности зависит от его отношения к другой сущности.

Каждой сущности присваивается уникальное имя и номер, разделяемые косой чертой «/» и помещаемые над блоком.

Связь может дополнительно определяться с помощью указания степени или мощности (количества экземпляров сущности-потомка, которое может существовать для каждого экземпляра сущности-родителя). В IDEF1X могут быть выражены следующие мощности связей:

– каждый экземпляр сущности-родителя может иметь ноль, один или более связанных с ним экземпляров сущности-потомка;



- каждый экземпляр сущности-родителя должен иметь не менее одного связанного с ним экземпляра сущности-потомка;
- каждый экземпляр сущности-родителя должен иметь не более одного связанного с ним экземпляра сущности-потомка;
- каждый экземпляр сущности-родителя связан с некоторым фиксированным числом экземпляров сущности-потомка.

Если экземпляр сущности-потомка однозначно определяется своей связью с сущностью-родителем, то связь называется идентифицирующей, в противном случае – неидентифицирующей.

Связь изображается линией, проводимой между сущностью-родителем и сущностью-потомком с точкой на конце линии у сущности-потомка.

Идентифицирующая связь между сущностью-родителем и сущностью-потомком изображается сплошной линией (рисунок 1). Сущность-потомок в идентифицирующей связи является зависимой от идентификатора сущностью. Сущность-родитель в идентифицирующей связи может быть как независимой, так и зависимой от идентификатора сущностью (это определяется ее связями с другими сущностями).



Рисунок 1 – Изображение идентифицирующей связи

Пунктирная линия изображает неидентифицирующую связь (рисунок 2). Сущность-потомок в неидентифицирующей связи будет независимой от идентификатора, если она не является также сущностью-потомком в какой-либо идентифицирующей связи.

Атрибуты изображаются в виде списка имен внутри блока сущности. Атрибуты, определяющие первичный ключ, размещаются наверху списка и отделяются от других атрибутов горизонтальной чертой (рисунок 3).



Рисунок 2 – Изображение неидентифицирующей связи

Имя_Сущности/Номер_Сущности

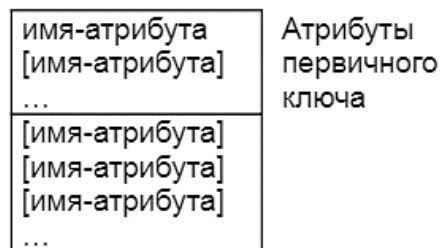


Рисунок 3 – Изображение атрибутов на сущности

Сущности могут иметь также внешние ключи (Foreign Key), которые могут использоваться в качестве части или целого первичного ключа или неключевого атрибута. Внешний ключ изображается с помощью помещения внутрь блока сущности имен атрибутов, после которых следуют буквы FK в скобках (рисунок 4).

Работа с пакетом Computer Associates ERWin.

Запуск программы осуществляется через меню Пуск -> Программы -> Computer Associates ERWin. После загрузки программы появится главное окно программы (рисунок 5).

Интерфейс программы во многом схож с BPWin. Однако поскольку ERWin воплощает другую методологию, то, соответственно, его инструменты будут отличаться.



Рисунок 4 – Пример изображения внешнего ключа

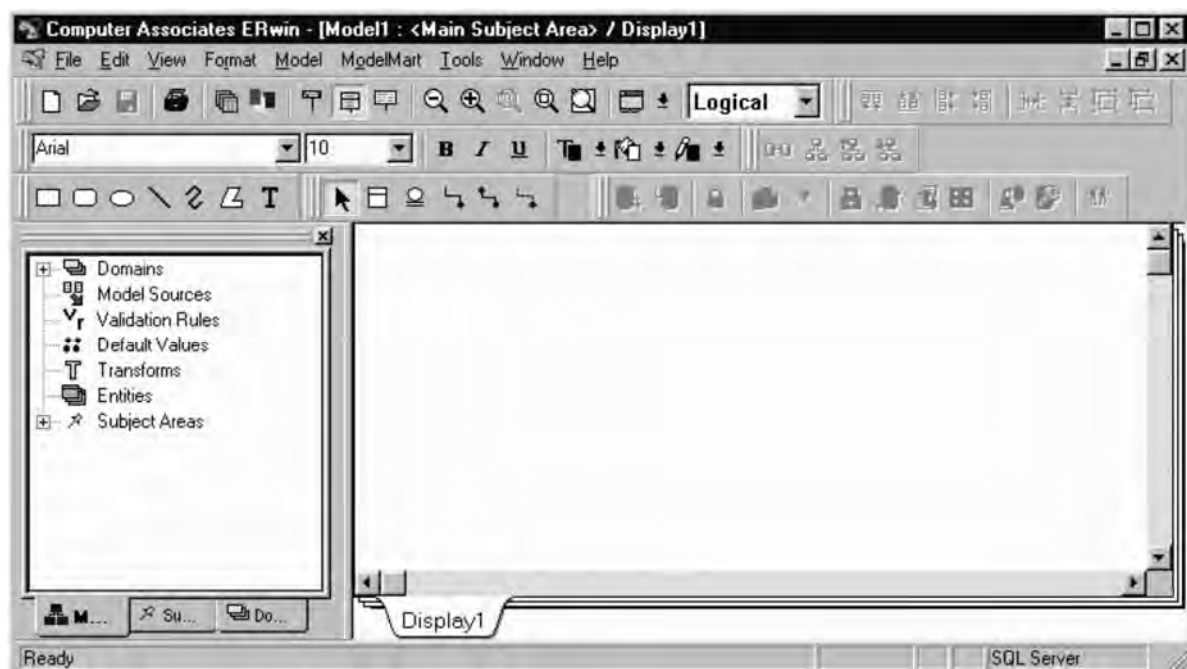


Рисунок 5 – Главное окно программы

Внимание! Пакет ERWin также использует по умолчанию шрифты без поддержки кириллицы. Для исправления этого выполните следующее.

1 На свободном участке поля модели щёлкните правой кнопкой мыши. Выберите пункт Default Fonts&Colors.

2 В появившемся окне (рисунок б) смените все упоминания шрифта Arial на Arial CYR во всех закладках.

3 Нажмите кнопку ОК.



Рисунок 6 – Окно изменения шрифтов

Панель инструментов. Панель инструментов ERWin приведена на рисунке 7.



Рисунок 7 – Панель инструментов ERWin

Здесь:



– режим выделения. Аналогичен BWin;



– создание новой сущности;



– обозначение полного включения;



– идентифицирующая связь (один-ко-многим);




– связь «многие-ко-многим»;



– неидентифицирующая связь.

Импорт сущностей из BWin. Для того чтобы импортировать файл с сущностями, созданный вами в BWin, создайте новую модель в ERWIN (File->New), а затем дайте команду на импорт файла (File->Import->BWin...). Появится стандартное окно Windows открытия файла. Найдите свой файл и нажмите «Открыть». Сущности будут импортированы в новую модель ERWin.

Создание новой сущности. Для того чтобы создать новую сущность, войдите в режим создания новой сущности (нажав левую кнопку мыши на пиктограмме ) и щёлкните на свободном участке рабочей области экрана. На этом месте появится новая сущность.

Редактирование сущности. Для того чтобы отредактировать существующую сущность, выделите её левой кнопкой мыши, затем нажмите правую кнопку мыши и в появившемся окне выберите Attributes (Атрибуты). Появится окно атрибутов (рисунок 8).

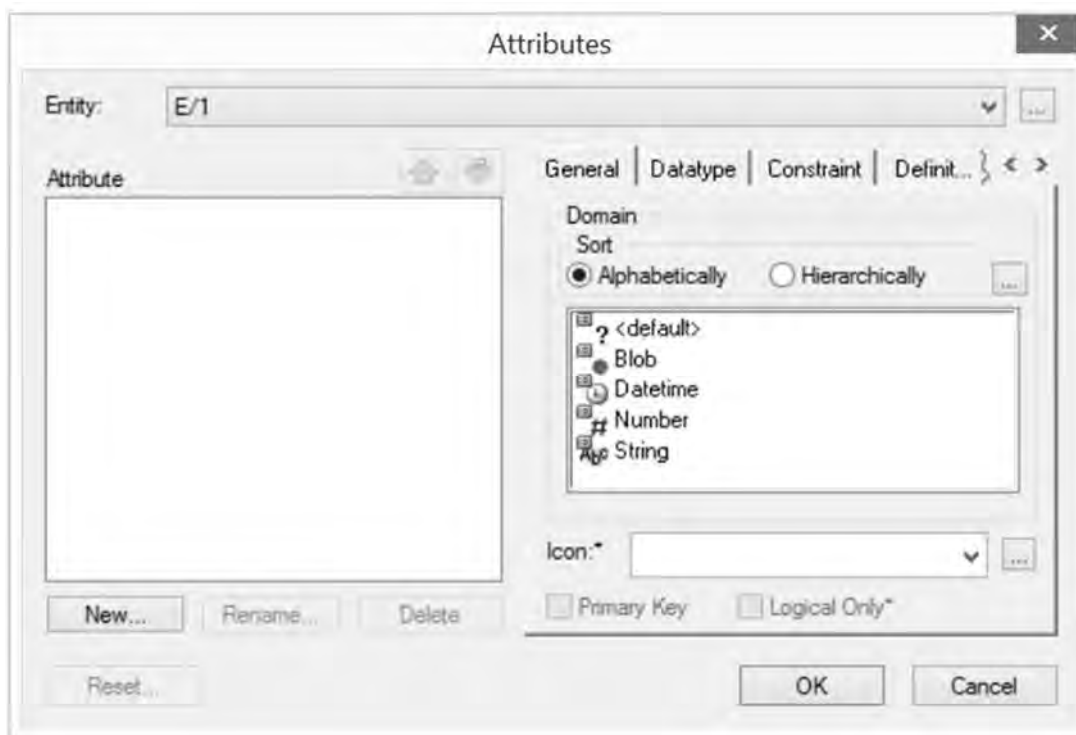


Рисунок 8 – Окно атрибутов

В этом окне вы можете создавать/редактировать/удалять атрибуты для данной сущности. Для того чтобы создать новый атрибут, нажмите кнопку New. Появится окно создания атрибута (рисунок 9).

В этом окне вы даёте атрибуту имя (поле Attribute Name) и определяете его тип (поле по центру). Определив имя и тип атрибута, нажмите кнопку ОК.

В окне атрибутов появится новый атрибут. Для того чтобы редактировать атрибут, воспользуйтесь закладками на правой части окна (рисунок 8). Здесь вы можете определить, является ли атрибут первичным (Primary Key), уточнить тип данных атрибута (вкладка Datatype) и т. д.

Пример информационной модели. После добавления всех атрибутов, расстановки связей между отношениями информационная модель выглядит следующим образом (рисунок 10).



Рисунок 9 – Окно создания атрибута

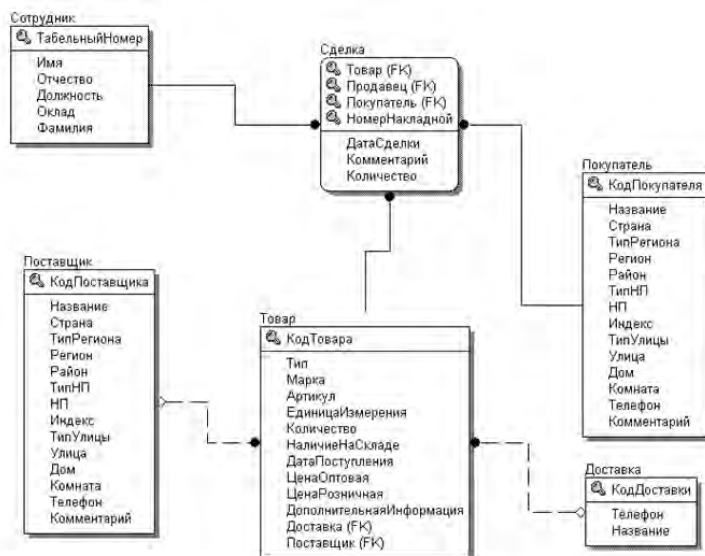


Рисунок 10 – Пример информационной модели

Задание

- 1 Изучить представление методологии IDEF1X в пакете ERWin.
- 2 Построить информационную модель выбранной предметной области.

Контрольные вопросы

- 1 Что такое информационная модель? Для чего она используется?
- 2 Назовите основные элементы информационной модели.
- 3 Что такое сущность? Атрибут? Связь?
- 4 Как осуществляется импорт из BPWin?

Лабораторная работа № 5. Реализация теоретико-игрового подход

Цель работы: ознакомиться с методами решения задач с помощью теоретико-игрового подхода.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Выполнить практическое задание.
- 3 Сделать выводы по полученным результатам.
- 4 Ответить на контрольные вопросы.
- 5 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Теория игр – совокупность математических методов решения конфликтных ситуаций (столкновений интересов). В теории игр игрой называется математическая модель конфликтной ситуации. Предмет особого интереса теории игр – исследование стратегий принятия решений участников игры в условиях неопределённости. Неопределённость связана с тем, что две или более стороны преследуют противоположные цели, а результаты любого действия каждой из сторон зависят от ходов партнёра. При этом каждая из сторон стремится принимать оптимальные решения, которые реализуют поставленные цели в наибольшей степени.

Наиболее последовательно теория игр применяется в экономике, где конфликтные ситуации возникают, например, в отношениях между поставщиком и потребителем, покупателем и продавцом, банком и клиентом. Применение теории игр можно найти и в политике, социологии, биологии, военном искусстве.

Матричной игрой в математической теории игр называется игра двух лиц с нулевой суммой, в которой в распоряжении каждого из них имеется конечное множество стратегий. Правила матричной игры определяет платёжная матрица, элементы которой – выигрыши первого игрока, которые являются также проигрышами второго игрока.

Матричная игра является антагонистической игрой. Первый игрок получает максимальный гарантированный (не зависящий от поведения второго игрока) выигрыш, равный цене игры, аналогично, второй игрок добивается минимального гарантированного проигрыша.



Под стратегией понимается совокупность правил (принципов), определяющих выбор варианта действий при каждом личном ходе игрока в зависимости от сложившейся ситуации.

В матричной игре её правила определяет платёжная матрица.

Рассмотрим игру, в которой имеются два участника: первый игрок и второй игрок. Пусть в распоряжении первого игрока имеется m чистых стратегий, а в распоряжении второго игрока – n чистых стратегий. Поскольку рассматривается игра, естественно, что в этой игре есть выигрыши и есть проигрыши.

В платёжной матрице элементами являются числа, выражающие выигрыши и проигрыши игроков. Выигрыши и проигрыши могут выражаться в пунктах, количестве денег или в других единицах.

Составим платёжную матрицу:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

Если первый игрок выбирает i -ю чистую стратегию, а второй – j -ю чистую стратегию, то выигрыш первого игрока составит a_{ij} единиц, а проигрыш второго игрока – также a_{ij} единиц.

Так как $a_{ij} + (-a_{ij}) = 0$, то описанная игра является матричной игрой с нулевой суммой.

Простейшим примером матричной игры может служить бросание монеты. Правила игры следующие. Первый и второй игроки бросают монету и в результате выпадает «орёл» или «решка». Если одновременно выпали «орёл» и «орёл» или «решка» и «решка», то первый игрок выиграет одну единицу, а в других случаях он же проиграет одну единицу (второй игрок выиграет одну единицу). Такие же две стратегии и в распоряжении второго игрока. Соответствующая платёжная матрица будет следующей:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Задача теории игр – определить выбор стратегии первого игрока, которая гарантировала бы ему максимальный средний выигрыш, а также выбор стратегии второго игрока, которая гарантировала бы ему максимальный средний проигрыш.

Как происходит выбор стратегии в матричной игре?

Вновь посмотрим на платёжную матрицу:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Сначала определим величину выигрыша первого игрока, если он использует i -ю чистую стратегию. Если первый игрок использует i -ю чистую стратегию, то логично предположить, что второй игрок будет использовать такую чистую стратегию, благодаря которой выигрыш первого игрока был бы минимальным. В свою очередь первый игрок будет использовать такую чистую стратегию, которая бы обеспечила ему максимальный выигрыш. Исходя из этих условий выигрыш первого игрока, который обозначим как v_1 , называется максиминным выигрышем или нижней ценой игры.

При решении задач на цену игры и определение стратегии для этих величин у первого игрока следует поступать следующим образом. Из каждой строки выписать значение минимального элемента и уже из них выбрать максимальный. Таким образом, выигрыш первого игрока будет максимальным из минимальных. Отсюда и название – максиминный выигрыш. Номер строки этого элемента и будет номером чистой стратегии, которую выбирает первый игрок.

$$\max_i \min_j a_{ij} = \underline{a} \text{ – нижняя цена игры.}$$

Теперь определим величину проигрыша второго игрока, если он использует j -ю стратегию. В этом случае первый игрок использует такую свою чистую стратегию, при которой проигрыш второго игрока был бы максимальным. Второй игрок должен выбрать такую чистую стратегию, при которой его проигрыш был бы минимальным. Проигрыш второго игрока, который обозначим как v_2 , называется минимаксным проигрышем или верхней ценой игры.

При решении задач на цену игры и определение стратегии для определения этих величин у второго игрока следует поступать следующим образом. Из каждого столбца выписать значение максимального элемента и уже из них выбрать минимальный. Таким образом, проигрыш второго игрока будет минимальным из максимальных. Отсюда и название – минимаксный выигрыш. Номер столбца этого элемента и будет номером чистой стратегии, которую выбирает второй игрок. Если второй игрок использует «минимакс», то независимо от выбора стратегии первым игроком он проиграет не более v_2 единиц.

$$\min_j \max_i a_{ij} = \bar{a} \text{ – верхняя цена игры.}$$

Седловая точка в матричных играх.

Если верхняя и нижняя цены игры одинаковы, то считается, что матричная игра имеет седловую точку. Верно и обратное утверждение: если матричная игра имеет седловую точку, то верхняя и нижняя цены матричной игры одинаковы. Соответствующий элемент одновременно является наименьшим в строке и наибольшим в столбце и равен цене игры.

Таким образом, если $\underline{a} = \bar{a} = v$, то игра имеет седловую точку (такую ситуацию называют равновесной), а пара оптимальных стратегий (A_{ionm}, B_{jonm}) – седловая точка матрицы. В этом случае матричная игра имеет решение в чистых стратегиях.

1 Если $v > 0$, то игра выгодна для игрока А.

2 Если $v < 0$ – для игрока В.

3 Если $v = 0$, то игра справедлива, т. е. является одинаково выгодной для обоих участников.

Практическое задание.

Найти цену игры, седловую точку и оптимальные стратегии игроков:

$$1) \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix};$$

$$5) \quad A = \begin{pmatrix} -2 & 1 & 2 \\ 1 & -1 & 0 \\ -1 & -3 & 1 \end{pmatrix};$$

$$2) \quad A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix};$$

$$6) \quad A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix};$$

$$3) \quad A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 2 \\ -1 & -2 & 1 \end{pmatrix};$$

$$7) \quad A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & -2 & 1 \\ -1 & 0 & 2 \end{pmatrix}.$$

$$4) \quad A = \begin{pmatrix} -1 & -2 & -3 \\ 0 & 1 & -2 \\ -1 & 2 & 0 \end{pmatrix};$$

Контрольные вопросы

1 Что называется антагонистической игрой?

2 Дайте определение матричной игры.

3 Как реализуется матричная игра?

4 Что представляет собой чистая стратегия 1-го (2-го) игрока?

5 Что такое максиминная и минимаксная стратегии игроков?

6 Что называется ситуацией равновесия?



Лабораторная работа № 6. Управление рисками

Цель работы: ознакомиться с методами решения задач при возникновении рисков.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Выполнить практическое задание.
- 3 Сделать выводы по полученным результатам.
- 4 Ответить на контрольные вопросы.
- 5 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

При принятии управленческих решений в условиях неопределенности и риска необходимо проводить анализ рисков. Анализ рисков подразделяется на два взаимно дополняющих друг друга вида: качественный, главная задача которого состоит в определении факторов риска и обстоятельств, приводящих к рисковому ситуациям, и количественный, позволяющий вычислить величину отдельных рисков и риска проекта в целом. Исследование риска целесообразно проводить в следующей последовательности:

- выявление объективных и субъективных факторов, влияющих на конкретный вид риска;
- анализ выявленных факторов;
- оценка конкретного вида риска с финансовых позиций, определяющая либо финансовую состоятельность проекта, либо его экономическую целесообразность;
- установка допустимого уровня риска;
- анализ отдельных операций по выбранному уровню риска;
- разработка мероприятий по снижению риска при принятии управленческого решения.

После проведения анализа рисков в процессе разработки управленческого решения используются специальные приемы управления риском.

Вопросами теории управления риском занимается риск-менеджмент.

Риск-менеджмент – специальная форма предпринимательской деятельности. Осуществляют ее профессиональные институты специалистов, страховые компании, финансовые менеджеры.



Одна из основных сфер риск-менеджмента – страховой рынок, где объектом купли-продажи выступают страховые услуги, предоставляемые организациям и отдельным гражданам преимущественно страховыми компаниями и негосударственными пенсионными фондами.

Основные приемы риск-менеджмента при принятии управленческих решений:

- избежание риска – уклонения от мероприятия, связанного с риском;
- удержание риска – оставление риска за инвестором (предполагая покрытие возможных убытков за счет резервных средств инвестора);
- передача риска – передача ответственности за риск, например, страховой компании;
- снижение степени риска – уменьшение вероятности потерь и сокращение ожидаемого их объема.

Наиболее распространенные приемы для снижения степени риска:

- диверсификация;
- получение дополнительной информации о ситуации принятия решения;
- лимитирование за счет установления предельных сумм расходов, продажи, кредита;
- самострахование за счет создания натуральных и денежных резервных (страховых) фондов;
- страхование.

Таким образом, в процессе разработки и принятия управленческих решений в условиях неопределенности и риска менеджер сталкивается с необходимостью проведения анализа существующих рисков, а также осуществления мероприятий, связанных с избеганием, удержанием, передачей рисков или снижения их степени. Кроме того, в условиях неопределенности и риска менеджеру необходимо использовать специальные приемы и методы разработки и принятия решений.

Правила и критерии принятия решений в условиях неопределённости.

Приведем несколько общих критериев рационального выбора вариантов решений из множества возможных. Критерии основаны на анализе матрицы возможных состояний окружающей среды и альтернатив решений.

Матрица, приведенная в таблице 1, содержит: A_j – альтернативы, т. е. варианты действий, один из которых необходимо выбрать; S_i — возможные варианты состояний окружающей среды; a_{ij} — элемент матрицы, обозначающий значение стоимости капитала, принимаемое альтернативой j при состоянии окружающей среды i .

Для выбора оптимальной стратегии в ситуации неопределённости используются различные правила и критерии.



Таблица 1 – Матрица решений

Альтернатива	S (состояние среды)					
	S_1	S_2	...	S_i	...	S_m
A	A_{11}	A_{12}	...	A_{1i}	...	A_{1m}
A_1	A_{11}	A_{12}	...	A_{1i}	...	A_{1m}
...
A_j	A_{j1}	A_{j2}	...	A_{ji}	...	A_{jm}
A_n	A_{n1}	A_{n2}	...	A_{jn}	...	A_{nm}

Правило максимин (критерий Ваальда).

В соответствии с этим правилом из альтернатив a_j выбирают ту, которая при самом неблагоприятном состоянии внешней среды имеет наибольшее значение показателя. С этой целью в каждой строке матрицы фиксируют альтернативы с минимальным значением показателя и из отмеченных минимальных выбирают максимальное. Альтернативе α с максимальным значением из всех минимальных даётся приоритет.

Принимающий решение в этом случае минимально готов к риску, предполагая максимум негативного развития состояния внешней среды и учитывая наименее благоприятное развитие для каждой альтернативы.

По критерию Ваальда лица, принимающие решения, выбирают стратегию, гарантирующую максимальное значение наихудшего выигрыша (критерия максимина), которое рассчитывается по формуле

$$\alpha = \max_i \min_j a_{ij} .$$

Правило максимакс (правило оптимизма).

В соответствии с этим правилом выбирается альтернатива с наивысшим достижимым значением оцениваемого показателя. При этом ЛПР (лицо принимающее решение) не учитывает риска от неблагоприятного изменения окружающей среды. Альтернатива находится по формуле

$$\alpha = \max_i \max_j a_{ij} .$$

Используя это правило, определяют максимальное значение для каждой строки и выбирают наибольшее из них.

Большой недостаток правил максимакса и максимина – использование только одного варианта развития ситуации для каждой альтернативы при принятии решения.

Правило минимин (правило пессимизма).

Предназначено для выбора наименьшего элемента матрицы доходности из её минимально возможных элементов:



$$\alpha = \min_i \min_j a_{ij} .$$

Правило пессимизма предполагает, что развитие ситуации будет неблагоприятным для лица, принимающего решение. При использовании этого правила лицо, принимающее решение, ориентируется на возможную потерю контроля над ситуацией и поэтому старается исключить все потенциальные риски и выбрать вариант с минимальной доходностью.

Правило минимакс (критерий Севиджа).

В отличие от максимина, минимакс ориентирован на минимизацию не столько потерь, сколько сожалений по поводу упущенной прибыли. Данный принцип позволяет выбрать вариант стратегии с меньшей величиной риска по сравнению с более высоким, первоначально ожидаемым уровнем риска. Данный критерий ориентирует лицо, принимающее решение, на более благоприятное развитие ситуации по сравнению с наихудшим состоянием, на которое то рассчитывало в начале. Правило допускает разумный риск ради получения дополнительной прибыли. Критерий Севиджа рассчитывается по формуле

$$\alpha = \min_i \max_j a_{ij} .$$

Расчёт минимакса состоит из двух этапов:

- 1) для каждой строки сожалений находим максимальное значение;
- 2) выбираем решение, при котором максимальное решение будет меньше других.

Правило Гурвица.

В соответствии с этим правилом правила максимакс и максимин сочетаются связыванием максимума минимальных значений альтернатив. Это правило называют ещё правилом оптимизма – пессимизма. Оптимальную альтернативу можно рассчитать по формуле

$$\alpha = k \max_j a_{ij} + (1 - k) \min_j a_{ij} ,$$

где k – коэффициент оптимизма, $k = 1 \dots 0$; при $k = 1$ альтернатива выбирается по правилу максимакс, при $k = 0$ – по правилу максимин. Учитывая боязнь риска, целесообразно задавать $k = 0,3$. Наибольшее значение целевой величины и определяет необходимую альтернативу.

Правило Гурвица применяют, учитывая более существенную информацию, чем при использовании правил максимин и максимакс.

Таким образом, при принятии управленческого решения в общем случае необходимо:

- спрогнозировать будущие условия, например, уровни спроса;



- разработать список возможных альтернатив;
- оценить окупаемость всех альтернатив;
- определить вероятность каждого условия;
- оценить альтернативы по выбранному критерию решения.

Задание

Игроки, независимо друг от друга, записывают целые числа от 3 до 6 включительно. Если первый игрок записал число x , а второй – y , то первый получает $x \cdot y$ единиц выигрыша. Если $\frac{x}{y} > 1$, и платит $x + y$ единиц, если $\frac{x}{y} < 1$. Если же $\frac{x}{y} = 1$, то игра заканчивается вничью. Необходимо построить платёжную матрицу и найти решение игры.

Контрольные вопросы

- 1 Опишите пять принципов обоснования решений.
- 2 Объясните экономический смысл понятия «максимин».
- 3 Объясните экономический смысл понятия «минимакс».
- 4 Как происходит принятие решений в условиях неопределенности и риска (классификация решений)?
- 5 Как осуществляется измерение рисков при принятии решений в условиях неопределенности?



Лабораторная работа № 7. Нечеткие экспертные системы

Цель работы: изучить метод экспертного ранжирования, методы парных сравнений, последовательных сопоставлений, изучить способы оценки согласованности между ранжировками экспертов.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Выполнить практическое задание.
- 3 Сделать выводы по полученным результатам.
- 4 Ответить на контрольные вопросы.
- 5 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Метод экспертного ранжирования.

Пусть имеется m экспертов $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_m$ и n целей Z_1, Z_2, \dots, Z_n . Каждый эксперт проводит оценку целей, пользуясь 10-балльной шкалой, причем оценки могут быть как целыми, так и дробными. В этих условиях веса целей определяются следующим образом:

- 1) составляется матрица оценок экспертов (таблица 1);
- 2) $0 \leq p_{ji} \leq 10$ ($j = 1, m; i = 1, n$);
- 3) составляется матрица нормированных оценок:

$$\omega = p_{ji} / \sum p_{ji} \quad (j = 1, m; i = 1, n);$$

- 4) вычисляются искомые веса целей:

$$\omega_i = \sum \omega_{ij} / \sum \sum \omega_{ij} \quad (i = 1, n); \quad \sum \omega_i = 1.$$

Таблица 1 – Матрица оценок экспертов

\mathcal{E}_j/Z_i	Z_1	Z_2	...	Z_n
\mathcal{E}_1	S_{11}	S_{12}	...	S_{1n}
\mathcal{E}_2	S_{21}	S_{22}	...	S_{2n}
...
\mathcal{E}_m	S_{m1}	S_{m2}	...	S_{mn}



Пример 1 – Найдем веса целей для случая $m = 2$ и $n = 6$.

1 Составим матрицу оценок экспертов (таблица 2).

Таблица 2 – Матрица оценок экспертов

\mathcal{E}_j/Z_i	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
\mathcal{E}_1	10	7	9	3	4	5
\mathcal{E}_2	8	6	10	4	2	7

2 Составим матрицу нормированных оценок (таблица 3).

Таблица 3 – Матрица нормированных оценок

\mathcal{E}_j/Z_i	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
\mathcal{E}_1	10/38	7/38	9/38	3/38	4/38	5/38
\mathcal{E}_2	8/38	6/38	10/38	4/38	2/38	7/38

3 Получаем искомые веса целей:

$$\omega_1 = (10/38 + 8/37)/2 = 0,239;$$

$$\omega_2 = \dots = 0,173;$$

$$\omega_3 = \dots = 0,254;$$

$$\omega_4 = \dots = 0,093;$$

$$\omega_5 = \dots = 0,079;$$

$$\omega_6 = \dots = 0,16.$$

Метод последовательных сравнений

Одна из возможных модификаций метода состоит в следующем.

1 Все цели располагаются в виде массива в порядке убывания их важности и назначаются предварительные оценки целей. При этом первая цель массива получает оценку 100, а остальным целям ставятся в соответствие оценки, отражающие их важность.

2 Первая цель массива сравнивается со всеми возможными комбинациями нижестоящих целей (с парами). В случае необходимости оценка первой цели корректируется. Вторая цель массива сравнивается со всеми возможными комбинациями нижестоящих целей (с парами). В случае необходимости оценка 2-й цели корректируется и т. д.

3 Производится запись скорректированных оценок и расчет на их основе весов целей.

Пример 2 – Эксперт проводит оценку четырех целей, которые связаны с решением транспортной проблемы:

Z_1 – построить метрополитен;

Z_2 – ввести 2-этажный автобус;

Z_3 – расширить существующую транспортную сеть;



Z_4 – ввести скоростной трамвай.

Расположим цели в виде массива и назначим предварительные оценки Z_1, Z_3, Z_4, Z_2 . Выставляем баллы: $p_1 = 100; p_3 = 60; p_4 = 40; p_2 = 10$.

Выполним сравнение целей и корректировку их оценок:

$$\begin{aligned} &Z_1(Z_3 \& Z_4); \\ &Z_1(Z_3 \& Z_2); \\ &Z_1(Z_4 \& Z_2); \\ &Z_3(Z_4 \& Z_2) \end{aligned}$$

(т. е. цель Z_1 сравниваем с комбинацией Z_3 и Z_4 и т. д.).

Возможно, что построить метрополитен лучше, чем 3 и 4, но 3 + 4 дают 100, поэтому корректируем оценки: $p_1 = 125; p_3 = 60$.

Запишем скорректированные оценки и вычислим веса целей:

$$p_1 = 125; p_3 = 60; p_4 = 40; p_2 = 10;$$

$$v_1 = 125/\text{сумма всех оценок} = 0,54; v_3 = 0,25; v_4 = 0,17; v_2 = 0,04.$$

Проверка: сумма v_i должна равняться 1.

Получаем порядок предпочтения целей: Z_1, Z_3, Z_4, Z_2 .

Согласованность между ранжировками двух экспертов можно определить с помощью **коэффициента ранговой корреляции Спирмэна**:

$$\rho = 1 - \frac{\sum_{i=1}^n (x_{ij} - x_{ik})^2}{n(n^2 - 1)} = 1 - \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

где x_{ij} – ранг, присвоенный i -му объекту j -м экспертом;

x_{ik} – ранг, присвоенный i -му объекту k -м экспертом;

d_i – разница между рангами, присвоенными i -му объекту;

n – количество факторов.

Величина коэффициента ранговой корреляции ρ может изменяться в диапазоне от -1 при наибольшем расхождении в мнениях экспертов до $+1$ при полном совпадении оценок.

Когда необходимо определить согласованность в ранжировках большого (более двух) числа экспертов, рассчитывается так называемый **коэффициент конкордации (коэффициент Кендалла)** – общий коэффициент ранговой корреляции для группы, состоящей из m экспертов:

$$W = \frac{12 S}{m^2(n^3 - n)},$$

где n – количество объектов;

m – количество экспертов;

S – квадрат отклонения,



$$S = \sum_{i=1}^n \left(\sum_{j=1}^m x_{ij} - \frac{1}{2} m(n+1) \right)^2,$$

где x_{ij} – ранг, присвоенный i -му объекту j -м экспертом.

Заметим, что вычитаемое в скобках представляет собой не что иное, как среднюю сумму рангов (при суммировании для каждого объекта), полученных i -ми объектами от экспертов.

Коэффициент W изменяется в диапазоне от нуля при полной несогласованности оценок экспертов до единицы в случае, если все эксперты присвоили объектам одинаковые ранги. Согласованность считается достаточной, если коэффициент конкордации больше 0,5.

Далее приведем примеры расчета коэффициентов ρ и W .

Пример 3 – Пусть два эксперта приписали четырем факторам, влияющим на успешность реализации проекта, ранги, показанные в таблице 4. На основе приведенных данных рассчитать коэффициент ранговой корреляции Спирмэна.

Решение

Промежуточные результаты расчетов (d_i и d_i^2) приведены в таблице 4.

Таблица 4 – Исходные данные и промежуточные результаты расчетов

Фактор	Ранг		d_i	d_i^2
	Первый эксперт x_{i1}	Второй эксперт x_{i2}		
Z_1	4	4	0	0
Z_2	3	1	2	4
Z_3	1	2	-1	1
Z_4	2	3	-1	1
Σ				6

Подставляя значения в формулу для вычисления коэффициента Спирмэна, получим

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \cdot 6}{4(4^2 - 1)} = 0,4.$$

Такое значение коэффициента Спирмэна свидетельствует о некоторой согласованности оценок экспертов.

Пример 4 – Три эксперта проранжировали четыре варианта капиталовложений (соответствующие оценки приведены в таблице 5). Проверить



согласованность ранжировок, используя коэффициент конкордации (*Кендалла*).

Решение

Количество экспертов $m = 3$, количество объектов $n = 4$.
В таблице 5 приведены промежуточные результаты расчетов.

Таблица 5 – Исходные данные и промежуточные результаты расчетов

Вариант	Эксперт			Сумма рангов (по строкам)	Квадрат отклонения S	$\left(\sum_{j=1}^m x_{ij} - \frac{1}{2}m(n+1) \right)^2$
	1	2	3			
Z_1	1	2	3	6	2,25	
Z_2	2	3	1	6	2,25	
Z_3	4	4	4	12	20,25	
Z_4	3	1	2	6	2,25	
Σ					$\Sigma S = 27$	$S = \sum_{i=1}^n \left(\sum_{j=1}^m x_{ij} - \frac{1}{2}m(n+1) \right)^2$

Подставляя вычисленное значение в формулу коэффициента конкордации, получим

$$W = \frac{12 S}{m^2 (n^3 - n)} = \frac{12 \cdot 27}{3^2 (4^3 - 4)} = 0,6.$$

Такая величина коэффициента конкордации W позволяет сделать вывод о том, что существует неслучайная согласованность в мнениях экспертов.

Указания к выполнению задания

К задаче по варианту необходимо составить матрицу попарного сравнения альтернатив.

Задачи, имеющие четные номера, решаются с применением метода экспертного ранжирования.

Задачи, имеющие нечетные номера, решаются с применением метода последовательных сопоставлений.

Необходимо оценить согласованность между ранжировками экспертов с помощью *коэффициента ранговой корреляции Спирмэна* либо с помощью *коэффициента конкордации (коэффициента Кендалла)*.

Задание

Вариант 1. Группа экспертов из трех человек проводит оценку четырех вариантов по строительству торгового центра:



- 1) достроить одноэтажное неиспользуемое помещение в центральном районе города;
 - 2) построить новый супермаркет, требующий крупных капиталовложений, с выгодным расположением;
 - 3) построить супермаркет за чертой города, с небольшими затратами;
 - 4) построить торговый центр на окраине города, район оснащен развитой транспортной сетью и паркингом.
- Найти оптимальный вариант проекта.

Вариант 2. Компания устраивает рекламную акцию по поводу открытия нового магазина. Группа экспертов из двух человек собрана для выбора наилучшей программы для привлечения потенциальных покупателей:

- 1) пригласить популярную поп-группу;
- 2) предложить скидку каждому покупателю;
- 3) раздавать призы;
- 4) устроить лотерею.

Вариант 3. Анализ результатов экономической деятельности предприятия показал его неспособность функционировать на рынке. Пригласили четырех экспертов для помощи руководству принять решение о выходе из сложившейся ситуации. Рассматриваются следующие варианты:

- 1) ликвидировать предприятие;
- 2) выставить на продажу;
- 3) объявить банкротом;
- 4) провести санацию.

Выяснить оптимальный путь дальнейшего развития предприятия.

Контрольные вопросы

- 1 Классификация и общая характеристика методов экспертных оценок.
- 2 Сущность метода экспертного ранжирования.
- 3 Сущность метода парных сравнений.
- 4 Когда применяется коэффициент ранговой корреляции Спирмэна?
А коэффициент конкордации?
- 5 Для чего нужно оценивать согласованность между ранжировками экспертов?



Лабораторная работа № 8. Кластеризация данных

Цель работы: получить навыки практического применения методов кластерного анализа.

Порядок выполнения работы

- 1 Изучить теоретические сведения.
- 2 Выполнить практическое задание.
- 3 Сделать выводы по полученным результатам.
- 4 Ответить на контрольные вопросы.
- 5 Оформить отчет.

Требования к отчету

- 1 Цель работы.
- 2 Постановка задачи.
- 3 Результаты задания.
- 4 Выводы.

Основные теоретические положения

Кластеризация (сегментация) – это разделение множества объектов на группы, которые обладают схожими характеристиками. Кластерный анализ – широкий класс процедур многомерного статистического анализа, позволяющих произвести автоматизированную группировку наблюдений в однородные классы – кластеры.

Кластер имеет следующие математические характеристики: центр; радиус; дисперсия кластера; среднеквадратическое отклонение.

Центр кластера – это среднее геометрическое место точек в пространстве переменных.

Радиус кластера – максимальное расстояние точек от центра кластера.

Дисперсия кластера – это мера рассеяния точек в пространстве относительно центра кластера.

Среднеквадратичное отклонение (СКО) объектов относительно центра кластера – квадратный корень из дисперсии кластера.

Методы кластерного анализа

Методы кластерного анализа можно разделить на две группы:

- 1) иерархические;
- 2) неиерархические.

Каждая группа включает множество подходов и алгоритмов.

Используя различные методы кластерного анализа, аналитик может получить различные решения на одних и тех же данных. Это считается нормальным явлением.



1 Иерархические методы кластерного анализа.

Суть иерархической кластеризации состоит в последовательном объединении меньших кластеров в большие или разделении больших кластеров на меньшие.

Иерархические агломеративные методы (Agglomerative Nesting, AGNES). Эта группа методов характеризуется последовательным объединением исходных элементов и соответствующим уменьшением числа кластеров.

В начале работы алгоритма все объекты являются отдельными кластерами. На первом шаге наиболее похожие объекты объединяются в кластер. На последующих шагах объединение продолжается до тех пор, пока все объекты не будут составлять один кластер.

Иерархические дивизимные (делимые) методы (DIvisive ANAlysis, DIANA). Эти методы являются логической противоположностью агломеративным методам. В начале работы алгоритма все объекты принадлежат одному кластеру, который на последующих шагах делится на меньшие кластеры, в результате образуется последовательность расщепляющих групп.

Иерархические методы кластеризации различаются правилами построения кластеров. В качестве правил выступают критерии, которые используются при решении вопроса о «схожести» объектов при их объединении в группу.

2 Меры сходства.

Для вычисления расстояния между объектами используются различные меры сходства (меры подобия), называемые также метриками или функциями расстояний.

Евклидово расстояние является геометрическим расстоянием в многомерном пространстве и вычисляется по формуле

$$R(x, y) = \sqrt{\sum_i (x_i - y_i)^2}.$$

Евклидово расстояние (и его квадрат) вычисляется по исходным, а не по стандартизованным данным.

Квадрат евклидова расстояния вычисляется по формуле

$$R(x, y) = \sum_i (x_i - y_i)^2.$$

Манхэттенское расстояние (расстояние городских кварталов), также называемое «хэмминговым» или «сити-блок» расстоянием, рассчитывается как среднее разностей по координатам. В большинстве случаев эта мера расстояния приводит к результатам, подобным расчетам евклидова расстояния. Однако для данной меры влияние отдельных выбросов меньше, чем при использовании евклидова расстояния, поскольку здесь координаты не возводятся в квадрат. Манхэттенское расстояние вычисляется по формуле



$$R(x, y) = \sum_i |x_i - y_i|.$$

Расстояние Чебышева. Это расстояние стоит использовать, когда необходимо определить два объекта как «различные», если они отличаются по какому-то одному измерению. Расстояние Чебышева вычисляется по формуле

$$R(x, y) = \max |x_i - y_i|.$$

Степенное расстояние используется в тех случаях, когда желают прогрессивно увеличить или уменьшить вес, относящийся к размерности, для которой соответствующие объекты сильно отличаются. Степенное расстояние вычисляется по формуле

$$R(x, y) = \left(\sum_i |x_i - y_i|^p \right)^{\frac{1}{r}},$$

где r и p – параметры, определяемые пользователем. Параметр p отвечает за постепенное взвешивание разностей по отдельным координатам, параметр r – за прогрессивное взвешивание больших расстояний между объектами. Если оба параметра r и p равны двум, то это расстояние совпадает с расстоянием Евклида.

Процент несогласия используется в тех случаях, когда данные являются категориальными. Это расстояние вычисляется по формуле

$$R(x, y) = \frac{\text{Количество } x_i \neq y_i}{n}.$$

Метрика Левенштейна для данных в виде слов используется в поисковиках.

3 Методы объединения или связи.

На первом шаге, когда каждый объект представляет собой отдельный кластер, расстояния между этими объектами определяются выбранной мерой. Однако когда связываются вместе несколько объектов, необходимо использовать другие методы определения расстояния между кластерами. Существует множество методов объединения кластеров.

Одиночная связь (метод ближайшего соседа) – расстояние между двумя кластерами определяется расстоянием между двумя наиболее близкими объектами (ближайшими соседями) в различных кластерах.

Полная связь (метод наиболее удаленных соседей) – расстояния между кластерами определяются наибольшим расстоянием между любыми двумя объектами в различных кластерах (т. е. «наиболее удаленными соседями»).



Невзвешенное попарное среднее – расстояние между двумя различными кластерами вычисляется как среднее расстояние между всеми парами объектов в них.

Взвешенное попарное среднее – метод идентичен методу *невзвешенного попарного среднего*, за исключением того, что при вычислениях размер соответствующих кластеров (т. е. число объектов, содержащихся в них) используется в качестве весового коэффициента.

Невзвешенный центроидный метод – расстояние между двумя кластерами определяется как расстояние между их центрами тяжести.

Взвешенный центроидный метод (медиана) – метод идентичен невзвешенному центроидному методу, за исключением того, что при вычислениях используются веса для учёта разницы между размерами кластеров (т. е. числами объектов в них).

Метод Варда – расстояние между кластерами определяется как прирост суммы квадратов расстояний объектов до центров кластеров, получаемый в результате их объединения. Метод отличается от всех других методов, поскольку он использует методы дисперсионного анализа для оценки расстояний между кластерами. Метод минимизирует сумму квадратов для любых двух (гипотетических) кластеров, которые могут быть сформированы на каждом шаге.

Метод ближайшего соседа.

Расстояние между двумя классами определяется как расстояние между ближайшими их представителями.

Перед началом работы алгоритма рассчитывается *матрица расстояний* между объектами. Согласно критерию классификации, объединение происходит между кластерами, расстояние между ближайших представителей которых наименьшее: выбираются два объекта с наименьшим расстоянием в один кластер. После этого необходимо произвести перерасчет матрицы расстояний с учетом нового кластера. На каждом шаге в матрице расстояний ищется минимальное значение, соответствующее расстоянию между двумя наиболее близкими кластерами. Найденные кластеры объединяются, образуя новый кластер. Эта процедура повторяется до тех пор, пока не будут объединены все кластеры.

При использовании метода ближайшего соседа особое внимание следует уделять выбору меры расстояния между объектами. На основе нее формируется начальная матрица расстояний, которая и определяет весь дальнейший процесс классификации.

4 Итеративные методы.

При большом количестве наблюдений иерархические методы кластерного анализа не пригодны. В таких случаях используют неиерархические методы, основанные на разделении исходных кластеров на другие кластеры, и которые представляют собой итеративные методы дробления исходной совокупности. В процессе деления новые кластеры формируются до тех пор, пока не будет выполнено правило остановки.



Такая неиерархическая кластеризация состоит в разделении набора данных на определенное количество отдельных кластеров. Существует два подхода. Первый заключается в определении границ кластеров как наиболее плотных участков в многомерном пространстве исходных данных, т. е. определение кластера там, где имеется большое «сгущение точек». Второй подход заключается в минимизации меры различия объектов.

В отличие от иерархических методов классификации итеративные методы могут привести к образованию пересекающихся кластеров, когда один объект может одновременно принадлежать нескольким кластерам.

К итеративным методам относятся, например, метод k -средних, метод поиска сгущений и другие. Итеративные методы относятся к быстродействующим, что позволяет использовать их для обработки больших массивов исходной информации.

Алгоритм k -средних (k -means)

Среди итеративных методов наиболее популярным методом является метод k -средних Мак-Кина. В отличие от иерархических методов в большинстве реализаций этого метода сам пользователь должен задать искомое число конечных кластеров, которое обычно обозначается « k ». Алгоритм k -средних строит k кластеров, расположенных на возможно больших расстояниях друг от друга. Основное требование к типу задач, которые решает алгоритм k -средних, – наличие предположений (гипотез) относительно числа кластеров, при этом они должны быть различны настолько, насколько это возможно. Выбор числа k может базироваться на результатах предшествующих исследований, теоретических соображениях или интуиции.

Как и в иерархических методах кластеризации, пользователь при этом может выбрать тот или иной тип меры сходства. Разные алгоритмы метода k -средних отличаются и способом выбора начальных центров задаваемых кластеров. В некоторых вариантах метода сам пользователь может (или должен) задать такие начальные точки, либо выбрав их из реальных наблюдений, либо задав координаты этих точек по каждой из переменных. В других реализациях данного метода выбор заданного числа k -начальных точек производится случайным образом, причем эти начальные точки (центры кластеров) могут в последующем уточняться в несколько этапов. Можно выделить четыре основных этапа таких методов:

- 1) выбираются или назначаются k -наблюдений, которые будут первичными центрами кластеров;
- 2) при необходимости формируются промежуточные кластеры приписыванием каждого наблюдения к ближайшим заданным кластерным центрам;
- 3) после назначения всех наблюдений отдельным кластерам производится замена первичных кластерных центров на кластерные средние;
- 4) предыдущая итерация повторяется до тех пор, пока изменения координат кластерных центров не станут минимальными.

Общая идея алгоритма: заданное фиксированное число k кластеров наблюдения сопоставляется кластерам так, что средние в кластере (для всех переменных) максимально возможно отличаются друг от друга.

Описание алгоритма

1 Первоначальное распределение объектов по кластерам.

Выбирается число k и k -точек. На первом шаге эти точки считаются «центрами» кластеров. Каждому кластеру соответствует один центр.

Выбор начальных центроидов может осуществляться следующим образом:

- выбор k -наблюдений для максимизации начального расстояния;
- случайный выбор k -наблюдений;
- выбор первых k -наблюдений.

Затем каждый объект назначается определенному наиболее близкому кластеру.

2 Итеративный процесс.

Вычисляются центры кластеров, которыми затем и далее считаются покоординатные средние кластеров. Объекты опять перераспределяются.

Процесс вычисления центров и перераспределения объектов продолжается до тех пор, пока не выполнено одно из условий:

- кластерные центры стабилизировались, т. е. все наблюдения принадлежат кластеру, которому принадлежали до текущей итерации. В некоторых вариантах этого метода пользователь может задать числовое значение критерия, трактуемого как минимальное расстояние для отбора новых центров кластеров. Наблюдение не будет рассматриваться как претендент на новый центр кластера, если его расстояние до заменяемого центра кластера превышает заданное число. Такой параметр в ряде программ называется «радиусом». Кроме этого параметра, возможно задание обычно достаточно малого числа, с которым сравнивается изменение расстояния для всех кластерных центров. Этот параметр обычно называется «конвергенцией», т. к. отражает сходимость итерационного процесса кластеризации;

- число итераций равно максимальному числу итераций.

Проверка качества кластеризации

После получения результатов кластерного анализа методом k -средних следует проверить правильность кластеризации (т. е. оценить, насколько кластеры отличаются друг от друга). Для этого рассчитываются средние значения для каждого кластера. При хорошей кластеризации должны быть получены сильно отличающиеся средние для всех измерений или хотя бы большей их части.

Достоинства алгоритма k -средних:

- простота использования;
- быстрота использования;
- понятность и прозрачность алгоритма.



Недостатки алгоритма k -средних:

- алгоритм слишком чувствителен к выбросам, которые могут искажать среднее. Возможным решением этой проблемы является использование модификации алгоритма – алгоритма k -медианы;
- алгоритм может медленно работать на больших базах данных. Возможным решением данной проблемы является использование выборки данных.

Задание

Разработать алгоритмы методов ближнего соседа и k -средних и реализовать их в виде компьютерных программ. С помощью ДСЧ сгенерировать 50 реализаций $x = (x_1, x_2)$ – случайной 2-мерной величины, координаты которой распределены равномерно в интервале (3, 8). Распределить их с помощью разработанных программ на минимальное число кластеров, каждый из которых помещается в сфере радиусом 0,15.

Контрольные вопросы

- 1 В чем состоит цель кластеризации? Приведите формальную постановку задачи кластеризации.
- 2 Приведите перечень и особенности методов кластеризации.
- 3 Какие метрики применяют в кластеризации?
- 4 Приведите математические характеристики кластеров и меры качества результатов кластеризации.
- 5 Охарактеризуйте этапы кластерного анализа выбранного метода.
- 6 Сформулируйте выводы и объясните полученные результаты лабораторной работы.



Список литературы

- 1 **Баранова, Е. К.** Моделирование системы защиты информации. Практикум: учебное пособие / Е. К. Баранова, А. В. Бабаш. – Москва: РИОР; ИНФРА-М, 2015. – 120 с.
- 2 **Часовников, С. Н.** Портфель финансовых инструментов: оптимальность по Эджварту – Парето / С. Н. Часовников, И. В. Кирьянов // Сибирская финансовая школа. – 2012. – № 3 (92). – С. 84–90.
- 3 **Шелобаев, С. И.** Математические методы и модели в экономике, финансах, бизнесе: учебное пособие для вузов / С. И. Шелобаев. – Москва: ЮНИТИ-ДАНА, 2000. – 367 с.
- 4 **Ларичев, О. И.** Теория и методы принятия решений: учебник / О. И. Ларичев. – Москва : Логос, 2002. – 296 с.
- 5 **Афанасьева, Т. В.** Введение в проектирование систем интеллектуального анализа данных: учебное пособие / Т. В. Афанасьева, А. Н. Афанасьев. – Ульяновск: УлГТУ, 2017. – 64 с.
- 6 **Маклаков, С. В.** BPWin и ERWin. CASE-средства разработки информационных систем / С. В. Маклаков. – Москва: ДИАЛОГ-МИФИ, 2001. – 256 с.
- 7 **Князь, Д.** Анализ основных алгоритмов кластеризации многомерных данных / Д. Князь. – Саарбрюккен: LAP Lambert Academic Publishing, 2014. – 796 с.
- 8 **Авдошин, С.** Информатизация бизнеса. Управление рисками / С. Авдошин, Е. Песоцкая. – Москва: ДМК, 2011. – 176 с.
- 9 Математические и инструментальные методы экономики: учебное пособие для студентов высших учебных заведений, обучающихся по специальности 080801 «Прикладная информатика» / П. В. Акинин [и др.]. – Москва: КноРус, 2012. – 232 с.

