



МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования

«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
ФГБОУ ВО «ИГУ»

Кафедра физико-химической биологии, биоинженерии и биоинформатики



Рабочая программа дисциплины
Наименование дисциплины:

Б1.О.25 «СПЕЦИАЛЬНЫЕ ГЛАВЫ МАТЕМАТИКИ»

Специальность: 06.05.01 «Биоинженерия и биоинформатика»

Специализация: «Биоинженерия и биоинформатика»

Квалификация выпускника: биоинженер и биоинформатик

Форма обучения: очная с элементами электронного обучения и дистанционных образовательных технологий

Согласовано с УМК биолого-почвенного
факультета
Протокол №4 от 20.04.2024
Председатель А. Н. Матвеев

Рекомендовано кафедрой физико-химической
биологии, биоинженерии и биоинформатики
Протокол №15 от 17.04.2024
Зав. кафедрой В.П. Саловарова

Иркутск 2024 г.

Содержание

	стр.
I. Цель и задачи дисциплины	3
II. Место дисциплины в структуре ОПОП	3
III. Требования к результатам освоения дисциплины	3
IV. Содержание и структура дисциплины	7
4.1 Содержание дисциплины, структурированное по темам, с указанием видов учебных занятий и отведенного на них количества академических часов	
4.2 План внеаудиторной самостоятельной работы обучающихся по дисциплине	
4.3 Содержание учебного материала	
4.3.1 Перечень семинарских, практических занятий и лабораторных работ	
4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение в рамках самостоятельной работы студентов	
4.4. Методические указания по организации самостоятельной работы студентов	
4.5. Примерная тематика курсовых работ (проектов)	
V. Учебно-методическое и информационное обеспечение дисциплины	14
а) перечень литературы	
б) периодические издания	
в) список авторских методических разработок	
г) базы данных, поисково-справочные и информационные системы.....	
VI. Материально-техническое обеспечение дисциплины	17
6.1. Учебно-лабораторное оборудование	
6.2. Программное обеспечение	
6.3. Технические и электронные средства обучения	
VII. Образовательные технологии	20
VIII. Оценочные материалы для текущего контроля и промежуточной аттестации	21

I. Цель и задачи дисциплины:

Цель: Изучить основы и методы математического моделирования и математического описания различных биологических процессов и систем для практического применения полученных знаний и навыков для решения профессиональных задач.

Задачи:

- Изучить основы языка программирования R как инструмента для различных комплексных математических вычислений.
- изучить спектр математических функций и их свойств, применяемый для моделирования и описания биологических систем и биологических процессов.
- освоить элементы векторной алгебры, применяемые для описания, сравнения и кластеризации биологических объектов и биологических систем;
- изучить методы теории вероятностей и математической статистики, применимые при анализе биологических данных реализованные в качестве пакетов функция для языка программирования R.
- ознакомиться с различными типовыми методами, применяемыми для численного интегрирования и численного исследования дифференциальных уравнений при описании биологических процессов;

II. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина Б1.О.25 «**Специальные главы математики**» относится к обязательной части образовательной программы.

2.2. Для изучения данной учебной дисциплины необходимы знания, умения и навыки, формируемые предшествующими дисциплинами: «Математика», «Физика», «Информатика», «Иностранный язык», «Современное естествознание».

2.3. Перечень последующих учебных дисциплин, для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной: «Биоинформатика», «Математическая обработка результатов исследований», «Алгоритмы биоинформатики», «Моделирование биологических процессов», «Биофизика», «Молекулярная филогенетика», «Геномный и метагеномный анализ», выполнение ВКР.

III. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций в соответствии с ФГОС ВО и ОП ВО по данному направлению подготовки 06.05.01 «Биология», профиль «Биоинженерия и биоинформатика»:

ОПК-2: Способен использовать специализированные знания фундаментальных разделов математики, физики, химии и биологии для проведения исследований в области биоинженерии, биоинформатики и смежных дисциплин (модулей)

ОПК-3: Способен проводить экспериментальную работу с организмами и клетками, использовать физико-химические методы исследования макромолекул, математические методы обработки результатов биологических исследований

Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ОПК-2 Способен использовать специализированные знания фундаментальных разделов математики, физики, химии и биологии для проведения исследований в области биоинженерии, биоинформатики и смежных дисциплин (модулей)	<i>ИДК ОПК-2.1</i> Демонстрирует специализированные знания в области фундаментальных разделов математики, физики, химии, биологии и перспективы междисциплинарных исследований	Знать: литературу по теме, владеть навыками, анализа информации сети «интернет» для поиска и освоения новых методов анализа данных и информационных технологий. Уметь: выбирать оптимальные методы и программы для решения задач в области анализа биологической информации по разным разделам биологических дисциплин. Владеть: методами построение анализа биологических систем с применением методов анализа функций, векторной алгебры, численных методов.
	<i>ИДК ОПК-2.2</i> Умеет использовать навыки проведения исследований в области биоинженерии, биоинформатики с учетом специализированных фундаментальных знаний	Знать: базовые алгоритмы в языка программирования R программирования, процедуры и функции обработки и визуализации статистических данных и результатов моделирования, реализации численных методов. Уметь: анализировать входные и выходные данные алгоритмов и моделей описания биологических систем, обрабатывать и визуализировать статистические данные и результаты моделирования с помощью базовых средств языка программирования R. Владеть: навыками анализа сложных данных в различных отраслях биологии и биоинформатики.
	<i>ИДК ОПК-2.3</i> Владеет методами химии, физики и математического моделирования для проведения исследований в области биоинженерии, биоинформатики	Знать: классификацию основных типов математических моделей и математических функций для описания и исследования биологических систем и биологических процессов. Уметь: осуществлять интерпретацию результатов математического моделирования и математических расчетов. Владеть: методами анализа комплексных биологических данных с использованием различных вычислительных и численных методов
ОПК-3 Способен проводить экспериментальную работу с организмами и клетками, использовать физико-химические методы исследования макромолекул,	<i>ИДК ОПК-3.1</i> Проводит экспериментальную работу с организмами и клетками с использованием физико-химических методов исследования	Знать: основные математические понятия и методы, применимые для анализа биологических макромолекул. Уметь: адекватно выбрать математический метод поведения биологических систем на молекулярном уровне. Владеть: основными принципами

математические методы обработки результатов биологических исследований	макромолекул	формализации сложных биологических систем в виде математических моделей клеточных и биохимических процессов
	<i>ИДК ОПК-3.2</i> Демонстрирует практические навыки математических методов обработки результатов экспериментальных исследований	Знать: цель, основные задачи и области применения математических методов в рамках направления подготовки. Уметь: формализовать исследуемую биологическую систему и биологический процесс в виде математической модели, использовать биологические данные для проверки и тестирования математических моделей, кластеризации и систематизации биологических данных. Владеть: методами анализа и исследования разработанных математических моделей для описания различных биологических процессов и биосистем, кластеризации и систематизации биологических объектов.
	<i>ИДК ОПК-3.3</i> Владеет опытом применения методов для исследования макромолекул, обработки результатов биологических исследований, прогнозирования перспектив и социальных последствий своей профессиональной деятельности.	Знать: особенности и основные свойства биологических систем, описываемых с помощью математических методов, методов кластеризации и систематизации биологических объектов. Уметь: выбирать адекватные методы для анализа биологических данных систематизации и кластеризации биологических объектов. Владеть: навыками совершенствования своих профессиональных качеств в области построения математических моделей и анализа и систематизации биологических данных.

IV. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Объем дисциплины составляет 5 зачетных единицы, 180 часов.

Из них реализуется с использованием электронного обучения и дистанционных образовательных технологий 30 часов.

Форма промежуточной аттестации: экзамен.

4.1 Содержание дисциплины, структурированное по темам, с указанием видов учебных занятий и отведенного на них количества академических часов

№ п/н	Раздел дисциплины/тема	Семестр	Всего часов	Из них практическая подготовка обучающихся	Виды учебной работы, включая самостоятельную работу обучающихся , практическую подготовку и трудоемкость (в часах)			Самостоятельная работа	Форма текущего контроля успеваемости/ Форма промежуточной аттестации (по семестрам)		
					Контактная работа преподавателя с обучающимися						
					Лекция	Семинар/ Практическое, лабораторное занятие/	Консультация				
1	2	3	4	5	6	7	8	9	10		
1	Тема 1. Введение в язык программирования R как универсальное средство для математических вычислений	3	16	4	4	8		4	KCP		
2	Тема 2. Базовые алгоритмы (ветвления, циклы) и пользовательские функции в языке программирования R.	3	25	4	6	12	1	6	KCP		
3	Тема 3. Понятие функция методы задания функций, элементарные функции, методы исследования функций с применением средств языка программирования R	3	18	4	4	8		6	KCP		
4	Тема 4. Элементы теории вероятности и	3	24	4	6	12		6	KCP		

	математической статики, реализация статистических методов в языке программирования R.								
5	Тема 5. Элементы векторной алгебры реализация алгоритмов векторной алгебры в языке программирования R.	3	16	4	4	8		4	KCP
6	Тема 6. Применение векторной алгебры для анализа сродства, различия и кластеризации объектов, реализация алгоритмов кластеризации в языке программирования R	3	16	4	4	8		4	KCP
7	Тема 7. Теория дифференциальных уравнений, численные методы интегрирования дифференциальных уравнений.	3	18	4	4	8		6	KCP
8	Тема 8. Численное интегрирования и исследование дифференциальных уравнений с помощью средств языка программирования R	3	18	4	4	8		6	KCP

4.2 План внеаудиторной самостоятельной работы обучающихся по дисциплине

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Трудоемкость (час.)		
3	Тема 1. Введение в язык программирования R как универсальное средство для математических вычислений	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по теме разработка линейных алгоритмов для языка программирования R.	1	4	KCP	Раздел 5 а-в

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Трудоемкость (час.)		
3	Тема 2. Базовые алгоритмы (ветвления, циклы) и пользовательские функции в языке программирования R.	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по разработке ветвящихся и циклических алгоритмов для языка программирования R.	2	6	KCP	- « -
3	Тема 3. Понятие функция методы задания функций, элементарные функции, методы исследования функций с применением средств языка программирования R	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по исследованию функций средствами языка программирования R.	4	6	KCP	- « -
3	Тема 4. Элементы теории вероятности и математической статистики, реализация статистических методов в языке программирования R.	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по анализа статистических данных средствами языка программирования R.	5	6	KCP	- « -
3	Тема 5. Элементы векторной алгебры реализация алгоритмов векторной алгебры в языке программирования R.	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по анализу и кластеризации объектов с применением методов векторной алгебры средствами языка программирования R.	8	4	KCP	- « -

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Трудоемкость (час.)		
3	Тема 6. Применение векторной алгебры для анализа сродства, различия и кластеризации объектов, реализация алгоритмов кластеризации в языке программирования R	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по анализу и кластеризации объектов с применением методов векторной алгебры средствами языка программирования R использование иерархичной кластеризации и многомерного шкалирования.	10	4	KCP	- « -
3	Тема 7. Теория дифференциальных уравнений, численные методы интегрирования дифференциальных уравнений.	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по теме использование численных методов для интегрирования дифференциальных уравнений.	12	6	KCP	- « -
3	Тема 8. Численное интегрирования и исследование дифференциальных уравнений с помощью средств языка программирования R	1. Разбор темы лекции и практического занятия. 2. Подготовка по контрольным вопросам. 3. Решение домашних задание по теме численное интегрирование дифференциальных уравнений с применением средств языка программирования R.	14	6	KCP	- « -
Общий объем самостоятельной работы по дисциплине (час) – 42						
Из них объем самостоятельной работы с использованием электронного обучения и дистанционных образовательных технологий 30 часов.						

4.3 Содержание учебного материала

Тема 1. Введение в язык программирования R как универсальное средство для математических вычислений.

Данная тема предназначена для обучения студентов навыкам с системой разработки скриптов и конвейеров по анализа данных с помощью языка программирования R. Рассматриваются приёмы работы с командной строкой интерпретатора R и графической среды разработки RStudio. Изучается вопрос установки дополнительных пакетов, реализующих различный набор функций по обработке данных.

Тема 2. Базовые алгоритмы (ветвления, циклы) и пользовательские функции в языке программирования R.

В рамках темы рассматривается базовый синтаксис языка программирования R – алгоритмы ветвлений, циклические алгоритмы, составные операторы. Изучается процесс создания пользовательских функций и элементы объектное ориентированного программирования на языке R.

Тема 3. Понятие функция методы задания функций, элементарные функции, методы исследования функций с применением средств языка программирования R.

Данная тема посвящена вопросу рассмотрения различных элементарных и сложных функций применяемых для описания различных биологических процессов. Рассматриваются вопросы численного исследования функций – нахождение максимумов, минимумов, точек перегиба с помощью средств языка программирования R.

Тема 4. Элементы теории вероятности и математической статистики, реализация статистических методов в языке программирования R.

В рамках темы рассматриваются различные типы распределений случайных величин, встречающихся при описании биологических данных. Разбираются различные задачи на основные теоремы теории вероятностей. Разрабатываются методы описательной статистики, применительно к биологическим данным, реализованные с помощью средств языка программирования R.

Тема 5. Элементы векторной алгебры реализация алгоритмов векторной алгебры в языке программирования R.

В рамках темы рассматривается понятие вектора алгебраического вектора и скалярного произведение векторов. Приводятся примеры использования векторной алгебры для охарактеризации биологических объектов с большим количеством параметров. Рассматриваются вопросы связанные с применением скалярного произведения векторов для сравнения и сопоставления многомерных биологических объектов и многомерных биологических данных. Изучаются функции языка программирования R, применяемые в векторной алгебре и анализе многомерных данных

Тема 6. Применение векторной алгебры для анализа сродства, различия и кластеризации объектов, реализация алгоритмов кластеризации в языке программирования R.

Данная тема посвящена вопросам связанным с изучением методов кластеризации сложных биологических объектов и многомерных биологических данных с применением методов векторной алгебры. Рассматриваются различные методы кластеризации и их реализации наборами функций для языка программирования R. Изучаются методы снижения размерности многомерных данных.

Тема 7. Теория дифференциальных уравнений, численные методы интегрирования дифференциальных уравнений.

В рамках темы рассматривается метод Эйлера и метод Эйлера с пересчетом для численного интегрирования дифференциальных уравнений и систем дифференциальных уравнений. Рассматривается алгоритм реализации метода с помощью языков программирования высокого уровня C, C++ и R. Изучается вопрос о точности численных методов для решения различных практических задач в области анализа биологических систем.

Тема 8. Численное интегрирования и исследование дифференциальных уравнений с помощью средств языка программирования R.

В рамках темы рассматривается возможность библиотеки «deSolve» для языка программирования R содержащий набор функций для численного интегрирования дифференциальных уравнений. Изучаются вопросы, связанные с применением библиотеки «deSolve» для численного интегрирования и исследований различных дифференциальных уравнений и систем дифференциальных уравнений.

Тема 1. Введение. Языки программирования. Классификация языков

4.3.1. Перечень семинарских, практических занятий и лабораторных работ

№ п/н	№ раздела и темы	Наименование семинаров, практических и лабораторных работ	Трудоемкость (час.)		Оценочные средства	Формируемые компетенции (индикаторы)*
			Всего часов	Из них практическая подготовка		
1	2	3	4	5	6	7
1	Тема 1	Изучение работы в консоли интерпретатора языка программирования R и визуальной среде разработки RStudio	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
2	Тема 2	Реализации базовых алгоритмов с использованием языка программирования R	12	12	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
3	Тема 3	Решение задач связанных с исследованием функция с применением библиотек языка	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3

		программирования R.				ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
4	Тема 4	Решение задач связанных с применением методов описательно статистики функция с применением библиотек языка программирования R.	12	12	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
5	Тема 5	Решение задач по нахождение матриц значений различных типов скалярного произведения векторов с применением библиотек языка программирования R.	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
6	Тема 6	Решение задач по кластеризации многомерных данных и применению методов снижения размерности с применением библиотек языка программирования R.	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
7	Тема 7	Решение задач по численному интегрированию дифференциальных уравнений и систем дифференциальных уравнений	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2 ИДК ОПК-3.3
8	Тема 8	Решение задач по численному интегрированию дифференциальных уравнений и систем дифференциальных уравнений с применением	8	8	KCP	ОПК-2 ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3 ОПК-3 ИДК ОПК-3.1 ИДК ОПК-3.2

		библиотеки «deSolve» языка программирования R.				<i>ИДК ОПК-3.3</i>
--	--	---	--	--	--	--------------------

4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение студентами в рамках самостоятельной работы (СРС)

№ п/п	Тема	Задание	Формируемая компетенция	ИДК
1.	Тема 3. Понятие функция методы задания функций, элементарные функции, методы исследования функций с применением средств языка программирования R.	Самостоятельное изучение темы – библиотека «ggplot2» для языка программирования R и ее применение для визуализации сложных функциональных зависимостей. Выполнение посменного самостоятельно задания с предоставлением отчета.	ОПК-2	ИДК ОПК-2.1 ИДК ОПК-2.2 ИДК ОПК-2.3

4.4. Методические указания по организации самостоятельной работы студентов

Самостоятельная работа студентов является составной частью учебного процесса и имеет целью закрепление и углубление полученных знаний и навыков, поиск и приобретение новых знаний, а также выполнение учебных заданий, подготовку к предстоящим занятиям, и экзамену по предмету.

Для организации самостоятельной работы по дисциплине «Основы программирования» используются следующие формы самостоятельной учебной работы:

- Работа по изучение темы с использованием материалов практического занятия.
- Подбор, изучение, анализ рекомендованной литературы.
- Изучения тем занятий, вынесенных на самостоятельное изучение.
- Самостоятельное изучение синтаксических конструкции и дополнительных библиотек для языка программирования R.
- Самостоятельное решения домашних задач по разработке алгоритмов и математических методов по изучаемым темам.
- Подготовка письменных отчетов по решению самостоятельных заданий задач.
- Подготовка к экзамену.

Письменный отчет по решению домашних заданий – это отчет о выполнении домашнего задания по темам дисциплины, содержащий следующую информацию:

- Ф.И.О. номер группы магистранта;
- номер задания;
- формулировка задания;
- описание результат решения задания с приведением файлов алгоритмов и их блок-схем в соответствии с формулировкой задания.

Критерий оценки отчета по решению домашнего задания:

- Оценка «зачтено». Задание выполнено правильно и в полном объеме, все алгоритмы и графики согласно формулировке задания предоставлены в отчете.
- Оценка «не зачтено». Задание выполнено не правильно или не в полном объеме, вопросуется на переделку и доработку.

Подготовка к экзамену в виде тестирования. К экзамену в виде тестирования допускаются студенты, получившие зачеты по всем самостоятельным заданиям.

4.5. Примерная тематика курсовых работ (проектов): не предусмотрены учебным планом.

V. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

а) перечень литературы

1. Мятлев В.Д., Панченко Л.А., Ризниченко Г.Ю., Терехин А.Т. Теория вероятностей и математическая статистика. Математические модели. изд.: Академия, 2009 г. 315 с. (Университетский учебник. Высшая математика и ее приложение в биологии) ISBN 978-5-7695-4704-1.+
2. Дамешек Л.Ю. Высшая математика [Текст]: учебное пособие для биологов. "Иркутский гос. ун-т". - Иркутск : Изд-во ИГУ, 2012. - 213 с.: ил., табл.; 21 см.; ISBN 978-5-9624-0637-4+
3. Шипачев В.С. Высшая математика. Базовый курс [Текст] : учеб. пособие / В. С. Шипачев ; ред. А. Н. Тихонов. - 8-е изд., перераб. и доп. - М. : Изд-во Юрайт, 2011. - 447 с. - ISBN 978-5-9916-0822-0. – ISBN 978-5-9692-0970-1 (10 экз.)+
4. Владимирский Б. М. Математика. Общий курс [Электронный ресурс] : учеб. /Б.М. Владимирский. – М. : Лань, 2008. – 960 с. – (Учебники для вузов. Специальная ли-тература). – Режим доступа: ЭБС "Издательство "Лань". – Неогранич. доступ. – ISBN 978-5-8114-0445-2.

б) дополнительная литература

1. Шипунов А. Б., Балдин Е. М., Волкова П.А., и др. Наглядная статистика. Используем R! Издательство: ДМК Пресс, 2014 – 300 с. Книга доступна в свободном доступе по ссылке: <http://ashipunov.info/shipunov/school/books/rbook.pdf>
2. Мастицкий С.Э., Шитиков В.К. Статистический анализ и визуализация данных с помощью R. Издательство: Хайдельберг –Лондон –Тольятти, 2104 – 401 с. Книга доступна в свободном доступе по ссылке: https://raw.githubusercontent.com/ranalytics/r-tutorials/master/Edition%202014/Book/Mastitsky_and_Shitikov_2014_R_tutorials.pdf
3. Кобзарь А.И. «Прикладная математическая статистика», для инженеров и научных работников. Издательство Москва: ФИЗМАТЛИТ, 2006 – 816 с. Книга доступна по ссылке: http://www.ph4s.ru/books/book_mat/statistika/kobzar.rar

б) периодические издания

1. <https://www.matbio.org/> - сайт журнала «Математическая биология и биоинформатика». Содержит большое количество статей в pdf – формате.
2. <https://journal.r-project.org/> - сайт журнала по статистическим методам на R, «The R Journal».

в) базы данных, информационно-справочные и поисковые системы

1. <http://dmb.biophys.msu.ru> - Информационная система «Динамические модели в биологии», рассчитанная на широкий круг пользователей, включает в себя гипертекстовые документы и реляционные базы данных и обеспечивает унифицированный доступ к разнообразной информации по данной предметной области.
2. <http://www.jcbi.ru/> - сайт объединенного центра вычислительной биологии и биоинформатики

3. <http://mathmod.aspu.ru/> - Сайт совместной лаборатории Института математических проблем биологии Российской академии наук и Астраханского государственного университета
4. <http://www.exponenta.ru/> - образовательный математический сайт
5. <http://www.library.biophys.msu.ru/MathMod/BM.HTML> - книга Г.Ю. Ризниченко «Биология математическая»
6. <http://nature.web.ru/db/msg.html?mid=1156624&uri=index.htm> - Бейли Н.. Математика в биологии и медицине. – М.: Мир, 1970.
7. <http://www.biometrika.tomsk.ru/> - электронный журнал «Биометрика» для медиков и биологов – сторонников доказательной биомедицины. Содержит большое количество статей и иных материалов, посвященных математическим моделям в биологии.
8. <http://www.library.biophys.msu.ru/FominBerk/main.htm> - Фомин С.В., Беркинблит М.Б. Математические проблемы в биологии. - М.: Гаука, 1973. - 200 с.
9. <https://www.elibrary.ru> – электронная библиотека научных статей, монографий и материалов конференций, выпущенных Российскими учеными.
10. <https://pubmed.ncbi.nlm.nih.gov/> - международная база данных научных статей и монографий, посвященная различным вопросам биологии.
11. <https://apps.webofknowledge.com> – международная база данных, индексирующая научные публикации в высокорейтинговых изданиях

VI. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Учебно-лабораторное оборудование:

- Аудитория для проведения занятий лабораторного типа. Компьютерный класс (учебная аудитория). Аудитория оборудована: специализированной (учебной) мебелью на 20 посадочных мест, доской меловой; оборудована техническими средствами обучения: Системный блок PentiumG850, Монитор BenQ G252HDA-1 шт.; Системный блокAthlon 2 X2 250, Монитор BenQ G252HDA – 8 шт.; Системный блок PentiumD 3.0GHz, Монитор Samsung 740N – 3 шт.; Моноблок IRU T2105P – 2 шт.; Системный блок Pentium G3250, Монитор BenQG955 – 1 шт.; Системный блок Pentium G3250, Монитор BenQ GL2250 – 1 шт.; Системный блок Pentium G3250, Монитор Samsung T200 HD – 1 шт.; Системный блок Pentium G3250, Монитор Samsung T190N – 1 шт.; Системный блок Pentium G3250, Монитор Samsung 740N – 1 шт.; Проектор BenQ MX503; экран ScreenVtdiaEcot. С неограниченным доступом к сети Интернет и обеспечением доступа в электронную информационно-образовательную среду организации, учебно-наглядными пособиями, обеспечивающими тематические иллюстрации по дисциплине «Моделирование и программирование биопроцессов» в количестве 8 шт., презентации по каждой теме программы.

- Компьютерный класс (учебная аудитория) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, организации самостоятельной работы. Аудитория оборудована: специализированной (учебной) мебелью на 20 посадочных мест, доской меловой; оборудована техническими средствами обучения: Системный блок PentiumG850, Монитор BenQ G252HDA-1 шт.; Системный блокAthlon 2 X2 250, Монитор BenQ G252HDA – 8 шт.; Системный блок PentiumD 3.0GHz, Монитор Samsung 740N – 3 шт.; Моноблок IRU T2105P – 2 шт.; Системный блок Pentium G3250, Монитор BenQG955 – 1 шт.; Системный блок Pentium G3250, Монитор BenQ GL2250 – 1 шт.; Системный блок Pentium G3250, Монитор Samsung T200 HD – 1 шт.; Системный блок Pentium G3250, Монитор Samsung T190N – 1 шт.; Системный блок Pentium G3250, Монитор Samsung 740N – 1 шт.; с неограниченным доступом к сети

Интернет; Проектор BenQ MX503; экран ScreenVtdiaEcot. Ноутбук Lenovo G580 – 1 шт. С неограниченным доступом к сети Интернет.

- Помещения для хранения и профилактического обслуживания учебного оборудования. Аудитория оборудована: специализированной мебелью на 11 посадочных мест; Шкаф для документов - 3 шт.; Сейф – 1 шт.; Шкаф-купе - 2 шт.; Принтер цвет. Canon LBR-5050 Laser Printer; Принтер Canon LBP-3010; Ноутбук Lenovo G580 – 1 шт.

6.2. Программное обеспечение:

DreamSpark Premium Electronic Software Delivery (3 years) Renewal (Windows 10 Education 32/64-bit (Russian) - Microsoft Imagine, Windows 7 Professional with Service Pack 1 32/64-bit (English) - Microsoft Imagine, Windows Server 2008 Enterprise and Standard without Hyper-V with SP2 32/64-bit (English) - Microsoft Imagine, Access 2016 32/64-bit (Russian) - Microsoft Imagine, Access 2010 32/64-bit (Russian) - Microsoft Imagine). Договор №03-016-14 от 30.10.2014г.

Kaspersky Endpoint Security для бизнеса - Стандартный Russian Edition. 250-499. Форус Контракт №04-114-16 от 14ноября 2016г KES. Счет №РСЦЗ-000147 и АКТ от 23ноября 2016г Лиц.№1B08161103014721370444.

Microsoft Office Enterprise 2007 Russian Academic OPEN No Level. Номер Лицензии Microsoft 43364238.

Microsoft Windows XP Professional Russian Upgrade Academic OPEN No Level. Номер Лицензии Microsoft 41059241.

Office 365 профессиональный плюс для учащихся. Номер заказа: 36dde53d-7cdb-4cad-a87f-29b2a19c463e.

6.3. Технические и электронные средства:

Презентации по всем темам курса.

VII. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации различных видов учебной работы дисциплины используются как стандартные методы обучения, так и интерактивные формы проведения занятий.

Стандартные методы обучения:

1. Информационная лекция.
2. Практические занятия, предназначенные для освоения студентами базовых методов анализа данных и использованию математических методов с помощью методов математического анализа
3. Самостоятельная работа студентов (выполнение домашних заданий, выполнения домашних заданий по теме для самостоятельного изучения, подготовка к экзаменационному тесту).
4. Консультации преподавателя.

Дистанционные образовательные технологии. Под дистанционными образовательными технологиями понимаются образовательные технологии, реализуемые в основном с применением информационно-телекоммуникационных сетей - интернет-технология – задействование образовательного портала ИГУ - educaisu.ru для предоставления письменных отчетов по домашним работам.

Наименование тем занятий с использованием дистанционных образовательных технологий:

№	Тема занятия	Вид занятия	Форма / Методы интерактивного обучения	Кол-во часов
1	Тема 1. Введение в язык программирования R как универсальное средство для математических вычислений	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	3
2	Тема 2. Базовые алгоритмы (ветвления, циклы) и пользовательские функции в языке программирования R.	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
3	Тема 3. Понятие функция методы задания функций, элементарные функции, методы исследования функций с применением средств языка программирования R.	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
4	Тема 4. Элементы теории вероятности и математической статики, реализация статистических методов в языке программирования R.	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
5	Тема 5. Элементы векторной алгебры реализация алгоритмов векторной алгебры в языке программирования R.	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	3
6	Тема 6. Применение векторной алгебры для анализа сродства, различия и кластеризации объектов, реализация алгоритмов кластеризации в языке программирования R	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
7	Тема 7. Теория дифференциальных уравнений, численные методы интегрирования дифференциальных уравнений.	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
8	Тема 8. Численное интегрирования и исследование дифференциальных уравнений с помощью средств языка программирования R	самостоятельная работа	Загрузка задания для контроля на образовательный портал ИГУ educa.isu.ru	4
Итого часов				30

ВIII. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Входной контроль знаний по данной дисциплине не предусмотрен.

Оценочные материалы текущего контроля

Оценочные материалы текущего контроля формируются в соответствии с ЛНА университета.

В рамках дисциплины «Спецглавы математики» используются следующие формы текущего контроля:

- письменная работа по решению самостоятельных заданий (все формулировки заданий для самостоятельного решения с необходимыми сопроводительными материалами выложены на образовательном портале ИГУ в темах курса «Спецглавы математики»);

Перечень посменных работ для самостоятельного выполнения по разделам – темам дисциплины.

Задание по теме 1:

Задание по теме 2:

Задание по теме 3:

Задание по теме 3. Раздел для самостоятельного изучения.

Задание по теме 4:

Задание по теме 5:

Задание по теме 6:

Задание по теме 7:

Задание по теме 8:

Оценочные средства для промежуточной аттестации

Промежуточная аттестация проходит в форме экзамена (3 семестр), к которому допускаются студенты, выполнившие в полном объеме аудиторную нагрузку, самостоятельную работу. Студенты, имеющие задолженность, должны выполнить все обязательные виды деятельности.

Фонд оценочных средств для промежуточной аттестации включает:

- тестовые задания для экзамена.

Назначение оценочных средств: выявить сформированность компетенций ОПК-2, ОПК-3 (см. п. III).

Тестовое задание включает два варианта по 20 вопросов по всем темам курса. К тесту допускаются студенты, задавшие все домашние заданий и получившие по каждому заданию зачет.

Критерий оценивания тестового экзаменационного задания

№	Тип задания	Критерии оценки	Результат оценивания
1	Задание закрытого типа на установление соответствие	Считается верным, если правильно установлены все соответствия (позиции одного столбца верно соотнесены с позициями другого столбца)	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
2	Задание закрытого типа на установление последовательности	Считается верным, если правильно указана вся последовательность цифр	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
3	Задание комбинированного типа с выбором одного верного ответа из четырех предложенных и обоснованием выбора	Считается верным, если правильно указана цифра (буква) правильного ответа и приведены корректные аргументы, используемые при выборе ответа	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
4	Задание комбинированного типа с выбором нескольких верных ответов из четырех предложенных и обоснованием выбора	Считается верным, если правильно указаны цифры (буквы) правильного ответа и приведены корректные аргументы, используемые при выборе ответа	Полное совпадение с верным ответом – 1 балл Все остальные случаи – 0 баллов
5	Задание открытого типа с развернутым ответом	Считается верным, если ответ совпадает с эталонным ответом по содержанию и полноте	Полное соответствие эталонному ответу – 1 балл Все остальные случаи – 0 баллов

Система получения баллов за тестирование

Оценка	критерий
отлично	18 и более баллов
хорошо	16 – 17 баллов
удовлетворительно	15 – 13 баллов
неудовлетворительно	12 баллов и менее

Оценочные материалы для промежуточной аттестации (экзамена)

Тестирование (Вариант 1).

Индекс и содержание формируемой компетенции	Индикаторы компетенций	Тестовые задания для промежуточной аттестации
ОПК-2 Способен использовать специализированные знания фундаментальных разделов математики, физики, химии и биологии для проведения исследований в области биоинженерии, биоинформатики и смежных дисциплин (модулей)	<p><i>ИДК ОПК-2.1</i> Демонстрирует специализированные знания в области фундаментальных разделов математики, физики, химии, биологии и перспективы междисциплинарных исследований</p> <p><i>ИДК ОПК-2.2</i> Умеет использовать навыки проведения исследований в области биоинженерии, биоинформатики с учетом специализированных фундаментальных знаний</p>	<p>Задание комбинированного типа с выбором одного или нескольких верных ответов из четырех предложенных с аргументацией выбора</p> <p>Вопрос 1. Какой из следующих способов является наиболее рекомендуемым для установки пакета dplyr из CRAN в R? а) source("http://cran.r-project.org/package=dplyr") б) install.packages("dplyr") в) library("dplyr") г) update.packages("dplyr") Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: б) Обоснование: <ul style="list-style-type: none"> • а) source() используется для выполнения R-скриптов, а не для установки пакетов. • б) install.packages("dplyr") - это стандартный и рекомендуемый способ установки пакетов из CRAN. Он скачивает и устанавливает пакет dplyr и все необходимые зависимости. • в) library("dplyr") - загружает установленный пакет в текущую сессию R, а не устанавливает его. Пакет должен быть установлен до использования library(). • г) update.packages("dplyr") - обновляет уже установленный пакет dplyr до последней версии. </p> <p>Вопрос 2. Какой фрагмент кода с циклом while в R корректно выводит числа от 1 до 5 (включительно)? а) i <- 1; while (i <= 5) { print(i); i <- i + 1 } б) i <- 1; while (i < 5) { print(i); i <- i + 1 } в) i <- 1; while (i <= 5) { print(i) } г) i <- 1; while (TRUE) { print(i); i <- i + 1; if (i > 5) break } Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: а) и г) Обоснование: <ul style="list-style-type: none"> • а) Это верный код. Цикл while (i <= 5) будет выполняться, пока i меньше или равно 5. Внутри цикла print(i) выводит текущее значение i, а i <- i + 1 увеличивает i на 1. • б) Этот код выводит числа только до 4, поскольку условие i < 5 означает, что цикл остановится, когда i достигнет значения 5. • в) Этот код создает бесконечный цикл, так как значение i никогда не изменяется, и условие i <= 5 всегда будет истинным. </p>
ОПК-3 Способен проводить	<i>ИДК ОПК-3.1</i> Проводит	

<p>экспериментальную работу с организмами и клетками, использовать физико-химические методы исследования макромолекул, математические методы обработки результатов биологических исследований</p>	<p>экспериментальную работу с организмами и клетками с использованием физико-химических методов исследования макромолекул</p> <p><i>ИДК ОПК-3.2</i> Демонстрирует практические навыки математических методов обработки результатов экспериментальных исследований</p>	<ul style="list-style-type: none"> • г) Это тоже верный код. Здесь используется бесконечный цикл while (TRUE), но внутри цикла есть проверка if ($i > 5$) break. Когда i становится больше 5, оператор break прерывает выполнение цикла. <p>Вопрос 3. Вопрос: Что делает оператор break внутри цикла в R? а) Прерывает текущую итерацию цикла и переходит к следующей. б) Завершает выполнение всего цикла. в) Возвращает значение из цикла. г) Переходит к началу текущей итерации цикла. Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: б) Обоснование:<ul style="list-style-type: none"> • а) Это описание оператора next (см. следующий вопрос), а не break. • б) Это верное описание оператора break. break полностью завершает выполнение цикла, независимо от того, сколько итераций еще осталось. • в) Оператор break не возвращает значение. • г) Это также описание оператора next </p> <p>Вопрос 4. Что произойдет, если вы попытаетесь создать вектор в R, объединив числовые и символьные значения (например, c(1, "a", 2, "b"))? а) Будет создана ошибка, так как вектор не может содержать элементы разных типов. б) Будет создан числовой вектор, в котором символьные значения будут автоматически преобразованы в числа (NA). в) Будет создан символьный вектор, в котором все числовые значения будут преобразованы в символьные строки. г) Будет создан список, содержащий как числовые, так и символьные значения. Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: в) Обоснование:<ul style="list-style-type: none"> • а) Ошибка не возникнет. R автоматически преобразует элементы к общему типу. • б) Преобразование в NA не произойдет. • в) Это верное утверждение. Векторы в R могут содержать только элементы одного типа. Когда вы пытаетесь объединить числовые и символьные значения, R приводит все элементы к наиболее общему типу - в данном случае, символьному (character). • г) Создается символьный вектор, а не список. </p> <p>Вопрос 5. Какой из следующих синтаксисов корректно определяет функцию в R, которая принимает два аргумента x и y, и возвращает их сумму? а) sum_function(x, y) { return(x + y) } б) sum_function <- function(x, y) { x + y } в) function sum_function(x, y) { return</p>
---	---	---

`x + y } г) sum_function = x, y -> { return(x + y) }`

Выберите один верный ответ и обоснуйте свой выбор:

Правильный ответ: б)

Обоснование:

- а) Отсутствует оператор присваивания `<-` или `=`. Это создает функцию, но она никуда не сохраняется.
- б) Это верный синтаксис. `sum_function <- function(x, y) { x + y }` правильно определяет функцию с именем `sum_function`, принимающую аргументы `x` и `y`, и возвращающую их сумму (последнее выражение в блоке кода функции автоматически возвращается). В R `return()` не всегда обязателен.
- в) `function sum_function(x, y) - Неправильный порядок слов и отсутствие оператора присваивания. return x + y также не использует скобки, что не является распространенной практикой.`
- г) `sum_function = x, y -> { return(x + y) }` - Это синтаксис, похожий на лямбда-функции в других языках программирования (например, Python или JavaScript), но он не является корректным синтаксисом для определения функций в R.

Вопрос 6.

Какие из перечисленных функций в R используются для вычисления основных показателей описательной статистики для числовых данных? (Выберите все верные ответы)

а) `mean()` б) `median()` в) `mode()` г) `sd()`

Выберите все верные ответы и обоснуйте свой выбор:

Правильный ответ: а), б), г)

Обоснование:

- а) `mean()` - Это верная функция. Она вычисляет среднее арифметическое значение вектора числовых данных.
- б) `median()` - Это верная функция. Она вычисляет медиану вектора числовых данных.
- в) `mode()` - Эта функция в R определяет тип данных переменной (например, "numeric", "character", "logical"), а не моду (наиболее часто встречающееся значение).
- г) `sd()` - Это верная функция. Она вычисляет стандартное отклонение вектора числовых данных.

Вопрос 7.

Какая функция в R используется для создания гистограммы?

а) `plot()` б) `hist()` в) `boxplot()` г) `density()`

Выберите один верный ответ и обоснуйте свой выбор:

Правильный ответ: б)

Обоснование:

- а) `plot()` - `plot()` - это универсальная функция для построения графиков, но она не предназначена специально для гистограмм. Её можно использовать, но потребуются дополнительные параметры.
- б) `hist()` - Это верный ответ. `hist(x)` создаст гистограмму для числового вектора `x`.
- в) `boxplot()` - `boxplot()` используется для создания ящиков с усами (диаграмм размаха), а не гистограмм.

- г) density() - density() оценивает плотность вероятности, но сама по себе не строит гистограмму. Обычно используется с plot() для визуализации оценки плотности.

Вопрос 8.

Какая функция в R используется для вычисления матрицы расстояний между строками в таблице данных?

- а) dist() б) distance() в) cor() г) scale()

Выберите один верный ответ и обоснуйте свой выбор:

Правильный ответ: а)

Обоснование:

- а) Это верный ответ. dist() - функция в R, предназначенная для вычисления матрицы расстояний.
- б) distance() - Такой функции в базовом R нет. Она может быть доступна в каких-либо пакетах, но не является стандартной.
- в) cor() - Вычисляет матрицу корреляций, а не расстояний.
- г) scale() - Стандартизирует данные (вычитает среднее и делит на стандартное отклонение), но не вычисляет расстояния.

Вопрос 9.

Какие типы расстояний можно вычислить с помощью функции dist() в R? (Выберите все верные ответы)

- а) Евклидово расстояние (Euclidean distance) б) Манхэттенское расстояние (Manhattan distance) в) Косинусное расстояние (Cosine distance) г) Расстояние Чебышева (Chebyshev distance)

Выберите все верные ответы и обоснуйте свой выбор:

Правильный ответ: а), б), г)

Обоснование:

- а) Это верный ответ. Евклидово расстояние является стандартной метрикой, используемой по умолчанию функцией dist().
- б) Это верный ответ. Манхэттенское расстояние (также известное как расстояние городских кварталов) можно вычислить с помощью dist(method = "manhattan").
- в) Косинусное расстояние не поддерживается непосредственно функцией dist(). Для вычисления косинусного расстояния потребуются другие функции или пакеты (например, lsa или написание собственной функции).
- г) Это верный ответ. Расстояние Чебышева можно вычислить с помощью dist(method = "maximum"). Метод "maximum" соответствует расстоянию Чебышева.

Вопрос 10.

Какие из перечисленных алгоритмов кластеризации доступны в базовом R (без необходимости установки дополнительных пакетов)? (Выберите все верные ответы)

- а) K-средних (K-means) б) Иерархическая кластеризация (Hierarchical clustering) в) DBSCAN г) Gaussian Mixture Models (GMM)

	<p>Выберите все верные ответы и обоснуйте свой выбор:</p> <p>Правильный ответ: а), б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Это верный ответ. Алгоритм К-средних реализован в функции kmeans() в базовом R. • б) Это верный ответ. Иерархическая кластеризация реализована в функции hclust() в базовом R. • в) DBSCAN (Density-Based Spatial Clustering of Applications with Noise) не доступен в базовом R. Он обычно реализуется с помощью пакетов, таких как dbscan. • г) Gaussian Mixture Models (GMM) не доступны в базовом R. Они обычно реализуются с помощью пакетов, таких как mclust <p>Вопрос 11.</p> <p>Каково основное назначение многомерного шкалирования (MDS)?</p> <p>а) Уменьшение размерности данных, сохраняя при этом наиболее важные переменные. б) Визуализация схожести (или несходства) между объектами в низкоразмерном пространстве. в) Кластеризация объектов на основе их сходства. г) Прогнозирование значений одной переменной на основе значений других переменных.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Уменьшение размерности данных - Это функция многих методов, включая PCA, но не является <i>основной</i> целью MDS. MDS может уменьшать размерность, но в первую очередь фокусируется на визуализации схожести. • б) Это верный ответ. MDS стремится представить объекты в низкоразмерном пространстве (обычно 2D или 3D) таким образом, чтобы расстояния между объектами в этом пространстве максимально соответствовали их исходным мерам несходства (dissimilarity). • в) Кластеризация объектов - Это функция кластерного анализа, хотя результаты MDS могут быть использованы для кластеризации, это не основное назначение MDS. • г) Прогнозирование значений одной переменной - Это задача регрессии, а не MDS. <p>Вопрос 12.</p> <p>Что такое дифференциальное уравнение?</p> <p>а) Алгебраическое уравнение, содержащее неизвестные переменные и коэффициенты. б) Уравнение, связывающее функцию с ее производными. в) Уравнение, описывающее геометрическую фигуру в пространстве. г) Уравнение, представляющее собой систему линейных алгебраических уравнений.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Алгебраическое уравнение - это базовый тип уравнений, но не определяет дифференциальное уравнение.
--	--

- б) Это верный ответ. Дифференциальное уравнение по определению содержит функцию и ее производные.
- в) Геометрическую фигуру описывают другие типы уравнений (например, уравнение окружности).
- г) Система линейных алгебраических уравнений - это отдельный класс уравнений.

Вопрос 13.

Какое утверждение лучше всего описывает метод Эйлера для численного решения дифференциальных уравнений?

- а) Это точный метод, который всегда дает решение дифференциального уравнения. б) Это метод первого порядка, использующий значение производной в текущей точке для оценки значения в следующей точке. в) Это метод высокого порядка, требующий вычисления нескольких производных. г) Это метод, который применим только к линейным дифференциальным уравнениям.

Выберите один верный ответ и обоснуйте свой выбор:

Правильный ответ: б)

Обоснование:

- а) Метод Эйлера - это *приближенный* метод и, как правило, *не* точный. Точность зависит от размера шага.
- б) Это верный ответ. Метод Эйлера является методом *первого* порядка, поскольку он использует только значение производной в *текущей* точке для оценки следующего значения. Это простейший и наименее точный метод.
- в) Метод Эйлера - это метод *низкого* порядка (первого). Методы высокого порядка (например, Рунге-Кутты) требуют вычисления нескольких производных или промежуточных значений.
- г) Метод Эйлера применим как к *линейным*, так и к *нелинейным* дифференциальным уравнениям.

Задание закрытого типа на установление соответствия

Вопрос 14.

Установите соответствие между задачей и наиболее подходящим типом циклического оператора в R для ее решения.

Задачи:

1. Вычислить сумму элементов вектора.
2. Читать строки из файла до тех пор, пока файл не закончится.
3. Повторять попытки подключения к серверу, пока соединение не будет установлено (максимум 10 попыток).
4. Выполнять итерацию, пока не будет введено корректное значение (проверка ввода).

Операторы (могут быть использованы несколько раз):

A. for B. while C. repeat

Правильный ответ:

- 1 - A
- 2 - B

- 3 - В
- 4 - С

Вопрос 15.

Установите соответствие между методом кластеризации в R и его ключевой характеристикой.

Методы кластеризации:

1. k-means
2. Иерархическая кластеризация (Hierarchical Clustering)
3. DBSCAN
4. PAM (Partitioning Around Medoids)

Характеристики:

A. Основан на поиске областей высокой плотности точек, игнорируя выбросы. В. Разбивает данные на k кластеров, минимизируя сумму квадратов расстояний до центроидов. Требует предварительного задания количества кластеров. С. Строит иерархию кластеров, объединяя или разделяя кластеры на каждом шаге. D. Разбивает данные на k кластеров, минимизируя сумму расстояний до *медоидов* (представителей кластеров). Требует предварительного задания количества кластеров.

Правильный ответ:

- 1 - В
- 2 - С
- 3 - А
- 4 - Д

Задание закрытого типа на установление последовательности

Вопрос 16.

Какая из следующих последовательностей действий наиболее точно отражает процесс выполнения иерархического кластерного анализа в R?

- a) Масштабирование данных -> Расчет матрицы расстояний -> Выбор метода кластеризации -> Визуализация дендрограммы -> Определение оптимального числа кластеров. b) Выбор метода кластеризации -> Расчет матрицы расстояний -> Масштабирование данных -> Визуализация дендрограммы -> Определение оптимального числа кластеров. c) Расчет матрицы расстояний -> Масштабирование данных -> Выбор метода кластеризации -> Определение оптимального числа кластеров -> Визуализация дендрограммы. d) Масштабирование данных -> Визуализация дендрограммы -> Расчет матрицы расстояний -> Выбор метода кластеризации -> Определение оптимального числа кластеров.

Правильный ответ: a)

Вопрос 17.

Предположим, у нас есть дифференциальное уравнение первого порядка: $dy/dx = f(x, y)$, с начальным условием $y(x_0) = y_0$. Метод Эйлера предлагает простой способ приближенного решения этого уравнения численно. Суть метода заключается в дискретизации области определения x на небольшие шаги размера h и последовательном приближении значения y в каждой точке, используя значение y в предыдущей точке и значение производной в этой точке.

Рассмотрим конкретный пример: $dy/dx = x + y$, с начальным условием $y(0) = 1$. Мы хотим найти приближенное значение $y(0.2)$ используя метод Эйлера с шагом $h = 0.1$. Это значит, что мы выполним два шага: от $x_0 = 0$ до $x_1 = 0.1$ и от $x_1 = 0.1$ до $x_2 = 0.2$.

Вычислите значения y на каждом шаге, используя формулу метода Эйлера: $y_{i+1} = y_i + h * f(x_i, y_i)$. Определите последовательность действий, необходимых для нахождения $y(0.2)$ и укажите приближенное значение $y(0.2)$, полученное в результате этих вычислений.

Варианты ответа (перемешаны):

- A) Вычислить $y(0.1)$ по формуле $y(0.1) = y(0) + h * f(0, y(0))$, где $f(x, y) = x + y$. B) Вычислить $y(0.2)$ по формуле $y(0.2) = y(0.1) + h * f(0.1, y(0.1))$, где $f(x, y) = x + y$. C) Подставить $y(0) = 1$, $x_0 = 0$ и $h = 0.1$ в формулу для вычисления $y(0.1)$. D) Подставить найденное значение $y(0.1)$, $x_1 = 0.1$ и $h = 0.1$ в формулу для вычисления $y(0.2)$. E) Значение $y(0.2)$ примерно равно 1.22. F) Значение $y(0.2)$ примерно равно 1.21. G) Значение $y(0.2)$ примерно равно 1.23.

Правильная последовательность действий и ответ: C -> A -> D -> B -> F.

Задание открытого типа с развернутым ответом

Вопрос 18.

Циклический оператор for в языке программирования R с примером кода для для вычисления квадратов чисел от 1 до 5 с выводом результата на экран

Ответ:

Циклический оператор for в R – это фундаментальный инструмент для повторения блока кода заданное количество раз. Он позволяет эффективно обрабатывать массивы, списки и другие структуры данных, выполняя одни и те же операции над каждым элементом. Синтаксис for цикла в R достаточно прост и понятен: `for (переменная in последовательность) { действия }`. Здесь переменная принимает значение каждого элемента из последовательности на каждой итерации цикла, а действия – это блок кода, который будет выполнен.

Последовательность, по которой происходит итерация, может быть представлена различными способами: вектором чисел, строковым вектором, списком и т.д. Ключевым моментом является то, что цикл for пробегает по *каждому* элементу этой последовательности, что позволяет нам организовать логику повторения, адаптированную к конкретным данным. Внутри блока кода действия мы можем обращаться к текущему значению переменной, использовать его для вычислений, условных проверок и других операций.

Рассмотрим пример использования for цикла для вычисления квадратов чисел от 1 до 5:

```
for (i in 1:5) {
    квадрат <- i * i
```

```
print(paste("Квадрат числа", i, "равен", квадрат))
}
```

Вопрос 19.

Евклидово расстояние в кластерном анализе?

Ответ:

Евклидово расстояние, или евклидова метрика, представляет собой один из наиболее фундаментальных и широко используемых методов измерения расстояния между двумя точками в многомерном пространстве. В контексте кластерного анализа, где задача состоит в группировке схожих объектов в кластеры, евклидово расстояние служит мощным инструментом для определения степени похожести между этими объектами. Основываясь на теореме Пифагора, евклидово расстояние вычисляет кратчайшее расстояние между двумя точками, представляя собой прямую линию, соединяющую эти точки.

Математически, евклидово расстояние между двумя точками p и q в n -мерном пространстве определяется как квадратный корень из суммы квадратов разностей между соответствующими координатами этих точек. Формула выглядит следующим образом: $d(p, q) = \sqrt((p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2)$. Простота и интуитивность этой формулы делают евклидово расстояние легко понимаемым и реализуемым в различных алгоритмах кластеризации.

В кластерном анализе, евклидово расстояние используется для оценки сходства между объектами на основе их характеристик, представленных в виде векторов в многомерном пространстве. Объекты, расположенные ближе друг к другу в евклидовом пространстве, считаются более похожими и, следовательно, более подходящими для объединения в один кластер. Алгоритмы, такие как k -средних (k -means) и иерархическая кластеризация, часто используют евклидово расстояние для определения оптимального расположения центроидов кластеров или для последовательного объединения наиболее близких кластеров.

Преимуществом евклидова расстояния является его простота и вычислительная эффективность, особенно для данных с небольшим количеством измерений. Однако, стоит учитывать, что евклидово расстояние может быть чувствительно к масштабу данных. Если разные признаки имеют разные единицы измерения или значительно различаются по диапазону значений, признаки с большими значениями могут доминировать в расчете расстояния, искажая результаты кластеризации. Поэтому, перед использованием евклидова расстояния, часто рекомендуется нормализовать или стандартизировать данные, чтобы все признаки имели одинаковую значимость.

Вопрос 20.

Метод локтя (Elbow method) для выбора наиболее оптимального количества кластеров при кластерном анализе методом К-средних?

Ответ:

Метод локтя (Elbow Method) – это эвристический подход, используемый в кластерном анализе для определения оптимального числа кластеров в наборе данных. Он основан на идеи оценки суммарного внутрикластерного расстояния (within-cluster sum of squares, WCSS) при различных значениях k (количество кластеров) и

		<p>визуального поиска точки перегиба на графике, напоминающей локоть.</p> <p>Суть метода заключается в следующем: по мере увеличения количества кластеров, WCSS, как правило, уменьшается, поскольку каждый кластер становится более компактным и однородным. Однако, на определенном этапе добавление новых кластеров начинает давать все меньший эффект в плане снижения WCSS. Точка, в которой уменьшение WCSS замедляется и происходит перегиб на графике, считается оптимальным количеством кластеров. Это происходит потому, что после этой точки добавление дополнительных кластеров приводит лишь к незначительному улучшению внутренней однородности, при этом усложняя модель и потенциально увеличивая риск переобучения.</p> <p>Для применения метода локтя необходимо выполнить кластеризацию данных с использованием алгоритма, например, K-means, для различных значений k. После каждой кластеризации вычисляется WCSS – сумма квадратов расстояний от каждой точки данных до центра своего кластера. Полученные значения WCSS затем отображаются на графике в зависимости от количества кластеров. Оптимальное значение k обычно находится визуально, путем определения точки "локтя" на графике.</p>
--	--	--

Тестирование (Вариант 2).

Индекс и содержание формируемой компетенции	Индикаторы компетенций	Тестовые задания для промежуточной аттестации
ОПК-2 Способен использовать специализированные знания фундаментальных разделов математики, физики, химии и биологии для проведения исследований в области биоинженерии, биоинформатики и смежных дисциплин (модулей)	<p><i>ИДК ОПК-2.1</i> Демонстрирует специализированные знания в области фундаментальных разделов математики, физики, химии, биологии и перспективы междисциплинарных исследований</p> <p><i>ИДК ОПК-2.2</i> Умеет использовать навыки проведения исследований в области биоинженерии,</p>	<p>Задание комбинированного типа с выбором одного или нескольких верных ответов из четырех предложенных и аргументацией выбора</p> <p>Вопрос 1. Что произойдет, если при использовании <code>install.packages()</code> в R, пакет, который вы пытаетесь установить, имеет неустановленные зависимости?</p> <p>а) Установка пакета завершится с ошибкой, и R предложит установить зависимости вручную. б) <code>install.packages()</code> автоматически попытается установить все необходимые зависимости из CRAN. в) <code>install.packages()</code> проигнорирует неустановленные зависимости и установит только запрошенный пакет. г) R запросит подтверждение на установку зависимостей перед продолжением установки основного пакета.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: б) Обоснование:</p> <ul style="list-style-type: none"> • а) Хотя ошибки могут возникнуть при проблемах с доступом к репозиториям, <code>install.packages()</code> автоматически разрешает и устанавливает зависимости, если они доступны в репозитории (например, CRAN). • б) Это верное описание поведения <code>install.packages()</code>. По умолчанию, функция автоматически загружает и

	<p>биоинформатики с учетом специализированных фундаментальных знаний</p> <p><i>ИДК ОПК-2.3</i> Владеет методами химии, физики и математического моделирования для проведения исследований в области биоинженерии, биоинформатики</p>	<p>устанавливает все необходимые зависимости для выбранного пакета.</p> <ul style="list-style-type: none"> в) <code>install.packages()</code> не игнорирует зависимости. Пакет может работать неправильно или вообще не работать, если его зависимости не установлены. г) <code>install.packages()</code> по умолчанию не запрашивает подтверждение на установку зависимостей, а делает это автоматически. <p>Вопрос 2. Какой из следующих вариантов кода корректно использует цикл <code>for</code> в R для вывода квадратов чисел от 1 до 5 (включительно)?</p> <p>а) <code>for (i in 1:5) { print(i^2) }</code> б) <code>for (i = 1, i <= 5, i++) { print(i^2) }</code> в) <code>for (i in range(1, 6)) { print(i^2) }</code> г) <code>for i = 1 to 5 { print(i^2) }</code></p> <p>Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: а)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> а) Это верный код. Цикл <code>for (i in 1:5)</code> корректно итерируется по последовательности чисел от 1 до 5, и <code>print(i^2)</code> выводит квадрат каждого числа. б) Синтаксис <code>for (i = 1, i <= 5, i++)</code> не является корректным синтаксисом цикла <code>for</code> в R. Это больше похоже на синтаксис C-подобных языков. в) Функция <code>range()</code> не существует в базовом R для создания последовательности чисел. В R для этого используется оператор <code>:</code>. г) Синтаксис <code>for i = 1 to 5</code> не является корректным синтаксисом цикла <code>for</code> в R. Это синтаксис, используемый в некоторых других языках программирования, но не в R.
	<p>ОПК-3 Способен проводить экспериментальную работу с организмами и клетками, использовать физико-химические методы исследования макромолекул, математические методы обработки результатов биологических исследований</p>	<p>Вопрос 3. Что делает оператор <code>next</code> внутри цикла в R?</p> <p>а) Завершает выполнение всего цикла. б) Прерывает текущую итерацию цикла и переходит к следующей. в) Возвращает значение из цикла. г) Выполняет следующую итерацию цикла, пропуская оставшуюся часть текущей итерации.</p> <p>Выберите один верный ответ и обоснуйте свой выбор: Правильный ответ: б) и г)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> а) Это описание оператора <code>break</code>. б) и г) Оба варианта по сути говорят об одном и том же, просто разными словами. Оператор <code>next</code> пропускает оставшуюся часть кода в текущей итерации цикла и сразу переходит к следующей итерации. Поэтому можно сказать, что он "прерывает текущую итерацию и переходит к следующей". в) Оператор <code>next</code> не возвращает значение. <p>Вопрос 4.</p>
	<p><i>ИДК ОПК-3.2</i> Демонстрирует практические навыки математических методов обработки результатов</p>	

	<p>экспериментальных исследований</p> <p>ИДК ОПК-3.3 Владеет опытом применения методов для исследования макромолекул, обработки результатов биологических исследований, прогнозирования перспектив и социальных последствий своей профессиональной деятельности.</p>	<p>Какой из следующих способов корректно создает числовой вектор в R, содержащий элементы 1, 2, 3, 4 и 5?</p> <p>а) vector <- c(1, 2, 3, 4, 5) б) vector = list(1, 2, 3, 4, 5) в) vector <- 1:5 г) vector = seq(1, 5)</p> <p>Выберите все верные ответы и обоснуйте свой выбор:</p> <p>Правильный ответ: а), в), г)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) vector <- c(1, 2, 3, 4, 5) - Это верный способ. Функция c() используется для объединения элементов в вектор. • б) vector = list(1, 2, 3, 4, 5) - Это неверный способ. list() создает список, а не вектор. Списки могут содержать элементы разных типов, в то время как векторы обычно содержат элементы одного типа. • в) vector <- 1:5 - Это верный способ. Оператор : создает последовательность чисел от 1 до 5 с шагом 1. • г) vector = seq(1, 5) - Это верный способ. Функция seq() создает последовательность чисел от 1 до 5 (по умолчанию с шагом 1). <p>Вопрос 5.</p> <p>Как определить функцию в R, которая принимает аргумент x и необязательный аргумент y со значением по умолчанию равным 10?</p> <p>а) my_function <- function(x, y = 10) { ... } б) my_function <- function(x, optional y = 10) { ... } в) my_function <- function(x, y = NULL) { if(is.null(y)) y <- 10; ... } г) my_function <- function(x = 10, y) { ... }</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: а)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Это верный способ. my_function <- function(x, y = 10) { ... } задает значение по умолчанию для аргумента y равным 10. Если при вызове функции аргумент y не будет указан, он автоматически примет значение 10. • б) optional y = 10 - Некорректный синтаксис. Слово optional не используется в R для определения аргументов по умолчанию. • в) y = NULL и проверка на is.null(y) - Это рабочий способ, но он более громоздкий, чем использование аргумента по умолчанию напрямую. y = 10 при определении функции - более лаконичный и предпочтительный вариант. • г) x = 10, y - Нельзя, чтобы аргумент без значения по умолчанию шел после аргумента со значением по умолчанию. Порядок имеет значение. В R сначала должны идти обязательные аргументы, а затем - аргументы со значениями по умолчанию. <p>Вопрос 6.</p> <p>Как в R можно обработать пропущенные значения (NA) при вычислении среднего значения с помощью функции mean()?</p> <p>а) Функция mean() автоматически игнорирует NA значения. б) Необходимо удалить NA значения из данных перед вычислением среднего значения. в) Нужно установить аргумент na.rm = TRUE в функции mean(). г) Нужно заменить NA значения нулями перед вычислением среднего значения.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p>
--	--	--

	<p>Правильный ответ: в)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Функция mean() по умолчанию возвращает NA, если в данных есть пропущенные значения. • б) Удаление NA значений может быть необходимо в некоторых случаях, но не является единственным или всегда лучшим решением. Удаление данных может привести к потере информации. • в) Это верный способ. Установка аргумента na.rm = TRUE указывает функции mean() игнорировать NA значения при вычислении среднего значения. • г) Замена NA нулями может исказить результаты, особенно если NA представляют собой отсутствие данных, а не реальные нулевые значения. <p>Вопрос 7.</p> <p>Какая функция в R используется для создания диаграммы размаха (боксплота)?</p> <p>а) hist() б) plot() в) boxplot() г) pie()</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Ваш ответ: в)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) hist() - Используется для создания гистограмм. • б) plot() - Можно использовать, но требует дополнительных параметров для создания боксплота. • в) Это верный ответ. Функция boxplot() создает диаграмму размаха. • г) pie() - Используется для создания круговых диаграмм. <p>Вопрос 8.</p> <p>Какой пакет в R обычно используется для расчета расстояния Махalanобиса?</p> <p>а) Базовый R (не требует установки пакетов) б) stats в) MASS г) cluster</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: в)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Базовый R - функция для вычисления расстояния Махalanобиса требует вычисления обратной ковариационной матрицы, что обычно делается с помощью функций из других пакетов. • б) stats - Пакет stats содержит базовые статистические функции, но не предоставляет специализированную функцию для расчета расстояния Махalanобиса. • в) MASS - Это верный ответ. Пакет MASS (который часто загружается вместе с базовым R, но все равно является отдельным пакетом) содержит функцию mcd.cov() для оценки ковариационной матрицы и может быть использован для вычисления расстояния Махalanобиса. В самом пакете MASS нет функции напрямую названной как расстояние Махalanобиса, но он предоставляет необходимые инструменты для его расчета. • г) cluster - Пакет cluster содержит функции для кластеризации данных, включая некоторые метрики расстояния, но обычно не используется напрямую для расчета расстояния Махalanобиса.
--	--

Вопрос 9.

В каких случаях рекомендуется нормализовать данные перед вычислением расстояний (например, евклидова расстояния)? (Выберите все верные ответы)

- а) Когда переменные имеют разные единицы измерения.
- б) Когда переменные имеют сильно различающиеся масштабы.
- в) Когда все переменные имеют одинаковые единицы измерения и масштабы.
- г) Нормализация всегда необходима перед вычислением расстояний.

Выберите все верные ответы и обоснуйте свой выбор:

Правильный ответ: а), б)

Обоснование:

- а) Это верный ответ. Если переменные измерены в разных единицах (например, метры и килограммы), нормализация необходима, чтобы одна переменная не доминировала при вычислении расстояния.
- б) Это верный ответ. Если переменные имеют сильно различающиеся масштабы (например, возраст от 0 до 100 и доход от 0 до 1,000,000), переменная с большим масштабом будет оказывать непропорционально большое влияние на вычисление расстояния.
- в) Когда переменные имеют одинаковые единицы измерения и масштабы, нормализация не всегда необходима, но может быть полезна в некоторых случаях.
- г) Нормализация не всегда необходима. Если переменные имеют сопоставимые масштабы и единицы измерения, можно вычислять расстояния без нормализации. Однако, нормализация может улучшить результаты в некоторых случаях, даже если масштабы похожи.

Вопрос 10.

Какие методы можно использовать для определения оптимального количества кластеров (k) при использовании алгоритма К-средних? (Выберите все верные ответы)

- а) Метод локтя (Elbow method)
- б) Силуэтный анализ (Silhouette analysis)
- в) Метод главных компонент (Principal Component Analysis - PCA)
- г) Интуитивный выбор на основе знания предметной области

Выберите все верные ответы и обоснуйте свой выбор:

Правильный ответ: а), б), г)

Обоснование:

- а) Это верный ответ. Метод локтя предполагает построение графика суммы внутригрупповых квадратов (Within-Cluster Sum of Squares, WCSS) для разного количества кластеров и выбор "локтя" на графике, где добавление дополнительных кластеров перестает значительно уменьшать WCSS.
- б) Это верный ответ. Силуэтный анализ измеряет, насколько объект похож на объекты в своем кластере (сплоченность) по сравнению с объектами в других кластерах (разделенность). Высокий силуэт указывает на хорошее соответствие объекта своему кластеру.
- в) Метод главных компонент (PCA) используется для уменьшения размерности данных, а не для определения количества кластеров. PCA можно использовать *перед* кластеризацией, чтобы упростить данные, но сам по себе

	<p>PCA не определяет оптимальное k.</p> <ul style="list-style-type: none"> г) Это верный ответ. Знание предметной области и понимание данных могут помочь определить разумное количество кластеров, даже если формальные методы не дают однозначного ответа. <p>Вопрос 11.</p> <p>Какие типы многомерного шкалирования существуют? (Выберите все верные ответы)</p> <p>а) Метрическое MDS (Metric MDS) б) Неметрическое MDS (Non-metric MDS) в) Линейное MDS (Linear MDS) г) Полиномиальное MDS (Polynomial MDS)</p> <p>Выберите все верные ответы и обоснуйте свой выбор:</p> <p>Правильный ответ: а), б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> а) Это верный ответ. Метрическое MDS (Metric MDS) предполагает, что исходные меры несходства являются метрическими (т.е. удовлетворяют аксиомам метрики) и стремится сохранить точные расстояния в низкоразмерном пространстве. б) Это верный ответ. Неметрическое MDS (Non-metric MDS) предполагает, что важен только ранг (порядок) мер несходства, а не их точные значения. Он стремится сохранить порядок расстояний в низкоразмерном пространстве. в) Линейное MDS (Linear MDS) - Такого общепринятого типа MDS не существует. MDS может быть реализован с использованием линейных алгебраических методов, но сама классификация не является "линейным MDS". г) Полиномиальное MDS (Polynomial MDS) - Такого общепринятого типа MDS не существует. <p>Вопрос 12.</p> <p>Что определяет порядок дифференциального уравнения?</p> <p>а) Количество неизвестных функций в уравнении. б) Наивысший порядок производной, входящей в уравнение. в) Количество членов в уравнении. г) Степень, в которой функция возведена в уравнении.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ: б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> а) Количество неизвестных функций - это влияет на тип системы дифференциальных уравнений, но не определяет порядок <i>отдельного</i> уравнения. б) Это верный ответ. Порядок дифференциального уравнения определяется наивысшим порядком производной, присутствующей в уравнении. в) Количество членов - не имеет отношения к порядку уравнения. г) Степень функции - это влияет на линейность уравнения, но не на его порядок. <p>Вопрос 13.</p> <p>В чем основное отличие метода Эйлера с пересчетом (Heun's method) от обычного метода Эйлера?</p> <p>а) Метод Эйлера с пересчетом требует меньшего размера шага. б) Метод Эйлера с пересчетом использует среднее</p>
--	---

	<p>значение производной в начале и конце шага. в) Метод Эйлера с пересчетом может применяться только к уравнениям второго порядка. г) Метод Эйлера с пересчетом всегда дает точное решение.</p> <p>Выберите один верный ответ и обоснуйте свой выбор:</p> <p>Правильный ответ ответ: б)</p> <p>Обоснование:</p> <ul style="list-style-type: none"> • а) Оба метода используют размер шага, определяемый пользователем. Однако, для достижения <i>одной и той же</i> точности, метод Эйлера с пересчетом может позволить себе <i>больший</i> размер шага, чем обычный метод Эйлера. • б) Это верный ответ. Метод Эйлера с пересчетом (также называемый методом Хойна или улучшенным методом Эйлера) использует <i>среднее</i> значение производной в начале и конце шага, что повышает точность. • в) Метод Эйлера с пересчетом может применяться к уравнениям <i>любого</i> порядка (преобразованным в систему уравнений первого порядка). • г) Метод Эйлера с пересчетом - это <i>приближенный</i> метод и <i>не всегда</i> дает точное решение. Он точнее обычного метода Эйлера, но все же является приближением. <p>Задание закрытого типа на установление соответствие</p> <p>Вопрос 14.</p> <p>Установите соответствие между циклическим оператором в R и его описанием.</p> <p>Операторы:</p> <ol style="list-style-type: none"> 1. for 2. while 3. repeat 4. break <p>Описания:</p> <p>А. Выполняет блок кода до тех пор, пока условие истинно. Проверяет условие <i>перед</i> каждой итерацией. В. Немедленно прекращает выполнение цикла и переходит к следующей строке кода после цикла. С. Выполняет блок кода заданное количество раз. Д. Выполняет блок кода бесконечно, пока не будет прерван оператором break.</p> <p>Правильный ответ:</p> <ul style="list-style-type: none"> • 1 - C • 2 - A • 3 - D • 4 - B <p>Вопрос 15.</p> <p>Установите соответствие между функциями в R и их назначением при выполнении кластерного анализа:</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: center; padding: 5px;">Функция в R</th> <th style="text-align: center; padding: 5px;">Назначение</th> </tr> </thead> <tbody> <tr> <td style="text-align: center; padding: 5px;"></td> <td style="text-align: center; padding: 5px;"></td> </tr> </tbody> </table>	Функция в R	Назначение		
Функция в R	Назначение				

1. kmeans()	A. Вычисление матрицы расстояний между объектами.
2. hclust()	B. Выполнение иерархической кластеризации.
3. dist()	C. Визуализация результатов кластеризации.
4. plot()	D. Выполнение кластеризации методом K-means.

Правильные ответы: 1-D, 2-B, 3-A, 4-C

Задание закрытого типа на установление последовательности

Вопрос 16.

Какие шаги необходимо выполнить для определения оптимального количества кластеров после применения K-means в R? Расположите их в правильной последовательности:

1. Визуализация внутригрупповой суммы квадратов (within-cluster sum of squares, WCSS) для различного количества кластеров.
 2. Выбор количества кластеров на основе графика WCSS (метод "локтя").
 3. Применение алгоритма K-means с разным числом кластеров.
- a) 3 -> 1 -> 2 b) 1 -> 2 -> 3 c) 2 -> 3 -> 1 d) 3 -> 2 -> 1

Правильный ответ: a)

Вопрос 17.

Предположим, у вас есть матрица диссимилярностей dist_matrix, представляющая различия между объектами. Какова правильная последовательность шагов для применения метрического многомерного шкалирования (Metric Multidimensional Scaling - MDS) с использованием функции cmdscale() в R для визуализации данных в двумерном пространстве?

- (a) Создать матрицу диссимилярностей -> Применить cmdscale(dist_matrix, k = 2) -> Визуализировать точки с помощью plot()
 (b) Визуализировать точки с помощью plot() -> Создать матрицу диссимилярностей -> Применить cmdscale(dist_matrix, k = 2)
 (c) Применить cmdscale(dist_matrix, k = 2) -> Создать матрицу диссимилярностей -> Визуализировать точки с помощью plot()
 (d) Применить cmdscale(dist_matrix, k = 2) -> Визуализировать точки с помощью plot() -> Создать матрицу диссимилярностей
- Правильный ответ: (a)*

Задание открытого типа с развернутым ответом

Вопрос 18.

	<p>Циклический оператор while в языке программирования R с примером кода для вычисления квадратов чисел от 1 до 5 с выводом результата на экран.</p> <p>Ответ:</p> <p>Циклический оператор while в языке программирования R представляет собой фундаментальную конструкцию для организации повторяющихся вычислений, пока определенное условие остается истинным. Он позволяет выполнять блок кода многократно, что незаменимо в ситуациях, когда количество итераций заранее неизвестно и зависит от динамически изменяющихся данных или состояний программы. Синтаксис оператора while достаточно прост: сначала указывается ключевое слово while, за которым в круглых скобках следует логическое выражение, выступающее в качестве условия продолжения цикла. В фигурных скобках заключен блок кода, который будет выполняться на каждой итерации цикла, пока условие истинно.</p> <p>Для вычисления квадратов чисел от 1 до 5 с использованием цикла while и выводом результатов на экран, можно создать следующий код на R:</p> <pre>i <- 1 while (i <= 5) { square <- i^2 cat("Квадрат числа", i, "равен", square, "\n") i <- i + 1 }</pre> <p>Вопрос 19.</p> <p>Бинарное расстояние в кластерном анализе?</p> <p>Ответ:</p> <p>Бинарное расстояние в кластерном анализе — это метрика, используемая для измерения сходства или различия между объектами, представленными в виде бинарных векторов. Бинарный вектор состоит из значений 0 и 1, где 0 обычно означает отсутствие признака, а 1 — его наличие. Примерами таких объектов могут быть документы, представленные как набор слов (наличие/отсутствие слова в документе), или пользователи, описанные набором интересов (интересуется/не интересуется).</p> <p>Существует несколько способов вычисления бинарных расстояний, каждый из которых акцентирует внимание на различных аспектах сходства или различия. Один из самых распространенных — расстояние Хэмминга, которое просто считает количество позиций, в которых два бинарных вектора отличаются. Например, если два вектора [1, 0, 1, 1] и [0, 0, 1, 0], то расстояние Хэмминга между ними равно 2 (отличия в первой и четвертой позициях).</p> <p>Другой популярной метрикой является коэффициент Жаккара, который определяет сходство между двумя наборами, деля размер пересечения на размер объединения этих наборов. В контексте бинарных векторов это означает количество признаков, которые присутствуют в обоих объектах, деленное на общее количество уникальных признаков, присутствующих хотя бы в одном из объектов. Коэффициент Жаккара принимает значения от 0 (нет общих признаков) до 1 (полное совпадение).</p> <p>Выбор конкретного бинарного расстояния зависит от задачи кластеризации и интерпретации результатов. Например, если важно учитывать только общие признаки, а отсутствие общих признаков не имеет большого значения, то</p>
--	---

коэффициент Жаккара может быть более подходящим, чем расстояние Хэмминга. В другом случае, если все признаки одинаково важны, то расстояние Хэмминга может быть предпочтительнее.

Использование бинарных расстояний в кластерном анализе позволяет группировать объекты на основе их бинарных характеристик, выявляя скрытые структуры и закономерности в данных. Это особенно полезно в областях, где данные естественным образом представлены в бинарном формате например в биоинформатике и вычислительной экологии.

Вопрос 20.

Силуэтный анализ (Silhouette analysis) для выбора наиболее оптимального количества кластеров при кластерном анализе методом K-средних?

Ответ:

Силуэтный анализ - это мощный инструмент оценки качества кластеризации, особенно полезный при использовании алгоритма K-средних для выявления оптимального числа кластеров в данных. Он предоставляет визуальное и количественное представление о том, насколько хорошо каждая точка данных "подходит" к своему кластеру, а также насколько она далека от соседних кластеров. В основе силуэтного анализа лежит вычисление силуэтного коэффициента для каждой точки, который представляет собой меру сходства точки со своим кластером по сравнению с ее сходством с другими кластерами.

Силуэтный коэффициент для каждой точки рассчитывается на основе двух расстояний: среднего расстояния от точки до всех других точек в ее кластере (a) и минимального среднего расстояния от точки до всех точек в любом другом кластере (b). Силуэтный коэффициент (s) вычисляется по формуле: $s = (b - a) / \max(a, b)$. Этот коэффициент принимает значения от -1 до 1. Значение, близкое к 1, указывает на то, что точка хорошо кластеризована, так как она близка к точкам в своем кластере и далека от других кластеров. Значение, близкое к -1, говорит о том, что точка, вероятно, была отнесена к неправильному кластеру. Значение около 0 указывает на то, что точка находится на границе между двумя кластерами.

Для применения силуэтного анализа при выборе оптимального числа кластеров необходимо выполнить кластеризацию K-средних для различных значений K (количества кластеров). Для каждого значения K вычисляются силуэтные коэффициенты для всех точек данных. Затем вычисляется средний силуэтный коэффициент для каждого значения K. Значение K, для которого средний силуэтный коэффициент является наибольшим, считается оптимальным числом кластеров. Этот подход позволяет выбрать структуру кластеризации, которая обеспечивает наилучшее разделение точек данных между кластерами.

Визуализация результатов силуэтного анализа осуществляется с помощью силуэтных графиков. Силуэтный график отображает силуэтные коэффициенты для каждой точки, отсортированные по убыванию внутри каждого кластера. Кластеры располагаются на графике по вертикали, а длина каждой полосы соответствует силуэтному коэффициенту соответствующей точки. Визуальный анализ силуэтного графика позволяет оценить однородность кластеров и наличие точек, которые могут быть неправильно кластеризованы. Хороший силуэтный график демонстрирует кластеры с широкими полосами (высокие силуэтные коэффициенты) и минимальным количеством точек с отрицательными значениями.

Разработчик:

Букин

(подпись)

доцент Букин Ю.С.

Программа составлена в соответствии с требованиями ФГОС ВО по направлению 06.05.01 «Биоинженерия и биоинформатика»

Программа рассмотрена на заседании кафедры физико-химической биологии, биоинженерии и биоинформатики 17.04.2024 г. протокол № 15.

Зав. кафедрой, д.б.н., профессор В.П. Саловарова Баев

Настоящая программа, не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.