ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«07» марта 2024 г.
Протокол №1

Рабочая программа дисциплины (модуля)
«Многомерная теория вероятностей и статистика»

Направление подготовки: *02.03.01 Математика и компьютерные науки*

Направленность (профиль) подготовки: *Программа двух дипломов НИУ ВШЭ и ЦУ «Прикладная математика и информатика»*

Квалификация (степень) выпускника: *бакалавр*

Форма обучения: *очная*

Срок освоения программы: *4 года*

Год набора: *2024*

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 0215B0C000D1B3FEB44A37EE631E0C27D4
Владелец: Ивашкевич Евгений Васильевич
Действителен: с 13.01.2026 до 13.04.2027

Электронный документ

Москва
2024

Содержание

1. **Краткая характеристика дисциплины (модуля)** 3
2. **Перечень планируемых результатов обучения** 5
3. **Тематический план** 7
4. **Содержание дисциплины (модуля)** 8
5. **Учебно-методическое обеспечение** 9
6. **Материально-техническое обеспечение** 9
7. **Методические и оценочные материалы** 11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Многомерная теория вероятностей и статистика» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по направлению 02.03.01 «Математика и компьютерные науки», профиль «Программа двух дипломов НИУ ВШЭ и ЦУ «Прикладная математика и информатика», утвержденный приказом Министерства науки и высшего образования Российской Федерации № 807 от 23.08.2017 года.

Дисциплина (модуль) «Многомерная теория вероятностей и статистика» формирует у студентов фундаментальную математико-статистическую базу для анализа сложных, многопараметрических данных, широко распространённых в науке, экономике, машинном обучении и инженерии. Она обеспечивает понимание структуры и зависимостей в многомерных данных, а также инструменты для их обработки, интерпретации и принятия решений.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 02.03.01 «Математика и компьютерные науки», профиль «Программа двух дипломов НИУ ВШЭ и ЦУ «Прикладная математика и информатика» и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений, как дисциплина по выбору.

Дисциплина (модуль) доступна для изучения на 3 или 4 курсе в 6, 7, 8 семестрах на выбор после успешного освоения дисциплин (модулей): «Математический анализ. Часть 1», «Математический анализ. Часть 2» и «Линейная алгебра и геометрия. Часть 1», «Линейная алгебра и геометрия. Часть 2» или «Математический анализ. Продвинутый уровень. Часть 1», «Математический анализ. Продвинутый уровень. Часть 2» и «Линейная алгебра и геометрия. Продвинутый уровень. Часть 1», «Линейная алгебра и геометрия. Продвинутый уровень. Часть 2».

Цель изучения дисциплины (модуля): формирование у студентов теоретических знаний и практических навыков для анализа многомерных данных с использованием вероятностных и статистических методов, а также подготовка к решению прикладных задач в условиях многомерной неопределённости.

Задачи изучения дисциплины (модуля):

— освоить математический аппарат многомерного анализа (матричная алгебра, случайные векторы, многомерные распределения).

— изучить основные методы описания, визуализации и преобразования многомерных данных.

— овладеть методами статистического вывода в многомерных моделях (оценивание, проверка гипотез).

— научиться применять современные методы анализа многомерных данных (PCA, кластеризация, факторный, канонический и регрессионный анализ).

— выработать способность критически интерпретировать результаты анализа и выбирать адекватные методы в зависимости от задачи и структуры данных.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

— основные типы и способы представления многомерных данных, базовые методы их визуализации и матричную алгебру, необходимую для анализа;

— понятия случайного вектора и матрицы, совместные, маргинальные и условные распределения, независимость компонент и характеристические функции;

— математическое ожидание, ковариационную матрицу, корреляцию и влияние линейных преобразований на моментные характеристики;

— основные многомерные распределения (в первую очередь многомерное нормальное и связанные с ним распределения), многомерные предельные теоремы и принципы многомерного статистического вывода;

— основные методы анализа многомерных данных: многомерная регрессия и MANOVA, проверка многомерных гипотез, PCA, кластеризация, факторный и канонический корреляционный анализ, базовые нелинейные методы снижения размерности;

уметь:

— представлять и визуализировать многомерные данные, вычислять выборочные моменты, ковариационные и корреляционные матрицы и интерпретировать их;

— строить и анализировать линейные и многомерные регрессионные модели, применять условные распределения, множественную и частную корреляцию;

— реализовывать точечное оценивание и проверку гипотез в многомерных моделях, в том числе с использованием критериев нормальности и критериев для вектора средних и ковариационной матрицы;

— применять PCA, методы кластеризации и дискриминантного анализа, факторный и канонический корреляционный анализ для снижения размерности, сегментации и изучения скрытых структур;

— использовать средства программирования (например, Python) для полного цикла анализа многомерных данных: от моделирования и оценивания до проверки гипотез и визуализации результатов;

владеть:

— навыками осмысленного эксплораторного анализа многомерных данных и выбора адекватных статистических методов под задачу;

— навыками построения и критической интерпретации многомерных статистических моделей и результатов проверки гипотез;

— навыками применения методов снижения размерности, кластеризации, факторного и канонического анализа для поддержки решений в задачах машинного обучения и разработки;

— навыками оценки и интерпретации качества моделей машинного обучения на основе статистических метрик, доверительных интервалов и кросс-валидации.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-1.	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1.	Знает методы поиска и анализа информации в области искусственного интеллекта, основные принципы критической оценки источников информации и их релевантности.
		УК-1.2.	Умеет критически оценивать источники информации и синтезировать данные из различных источников для решения задач, применять системный подход к анализу и решению комплексных проблем
		УК-1.3.	Имеет практический опыт работы с современными инструментами и технологиями для обработки информации, формулировании и структурировании задач на основе полученной информации
УК-2.	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	УК-2.1.	Знает действующие правовые нормы, регулирующие деятельность в области решения задач, основные методы и подходы к определению круга задач
		УК-2.2.	Умеет определять круг задач в рамках поставленной цели, выбирать оптимальные способы решения задач, учитывая имеющиеся ресурсы и ограничения
		УК-2.3.	Имеет практический опыт применения знаний о правовых нормах и ресурсах в реальных ситуациях, разработки и реализации решений в соответствии с установленными ограничениями
ОПК-1.	Способен находить, формулировать и решать актуальные и значимые проблемы прикладной и компьютерной математики	ОПК-1.1.	Знает основные методы и подходы к решению задач прикладной и компьютерной математики, включая алгоритмы, математическое моделирование и теорию оптимизации, а также современные инструменты и технологии, используемые в этой области
		ОПК-1.2.	Умеет анализировать и формулировать математические

			задачи, применять соответствующие методы и алгоритмы для их решения, а также интерпретировать и представлять результаты в понятной и доступной форме
		ОПК-1.3.	Имеет практический опыт работы над проектами или исследованиями в области прикладной и компьютерной математики, включая участие в конкурсах, олимпиадах или научных публикациях, где были решены актуальные и значимые задачи
ПК-1.	Способен формулировать задачи с математической точностью, обосновывать утверждения строго и анализировать полученные результаты в области математики и компьютерных наук	ПК-1.1.	Знает методы и подходы к формулированию задач, а также основные принципы математического доказательства и анализа результатов.
		ПК-1.2.	Умеет корректно ставить и формулировать математические задачи, применять строгие методы доказательства и анализировать полученные результаты.
		ПК-1.3.	Имеет опыт работы с задачами в области математики и компьютерных наук, включая применение математических методов для решения практических задач
ПК-2.	Способен решать типовые задачи профессиональной деятельности в области искусственного интеллекта, опираясь на информационную и библиографическую культуру, используя информационно-коммуникационные технологии и учитывая основные требования информационной безопасности	ПК-2.1.	Знает основы информационной и библиографической культуры, а также принципы информационной безопасности и применения информационно-коммуникационных технологий в профессиональной деятельности
		ПК-2.2.	Умеет эффективно использовать информационно-коммуникационные технологии для решения стандартных задач профессиональной деятельности, учитывая требования информационной безопасности
		ПК-2.3.	Имеет опыт работы с информационными ресурсами и технологиями в области искусственного интеллекта, включая соблюдение норм информационной безопасности

3. Тематический план

№п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостоятельная работа	
		Лекции	Семинары			
1	Случайные векторы и распределения	5	5		20	Домашнее задание
2	Предельные теоремы и оценивание	5	5	2	20	Домашнее задание Контрольная работа
3	Проверка гипотез и регрессия	5	5	2	20	Домашнее задание Коллоквиум
4	Снижение размерности и кластеризация	5	5		20	Домашнее задание
5	Классификация, факторный и канонический анализ	5	5		20	Домашнее задание
6	Многомерный статистический анализ и методы снижения размерности	5	5	2	20	Домашнее задание Контрольная работа
	Зачет с оценкой			4		Итоговая работа
	Итого:	***30***	***30***	***10***	***120***	
Объем дисциплины (модуля) (в ак. ч.)		***190***				
Объем дисциплины (модуля) (в зач. ед.)		***5***				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Случайные векторы и распределения	Определение. Совместные, маргинальные, условные распределения Числовые характеристики: матожидание, ковариационная матрица Многомерное нормальное распределение Распределения Уишарта, Хотеллинга
2	Предельные теоремы и оценивание	Многомерная ЦПТ. Дельта-метод. Метод моментов. ММП для многомерной нормальной модели
3	Проверка гипотез и регрессия	Критерий Хотеллинга. Проверка гипотез о векторе средних Модель многомерной линейной регрессии Многомерный дисперсионный анализ (MANOVA).
4	Снижение размерности и кластеризация	Метод главных компонент (PCA) K-средних. Иерархическая кластеризация
5	Классификация, факторный и канонический анализ	Линейный дискриминантный анализ Фишера (LDA) Модель факторного анализа. Отличие от PCA Канонический корреляционный анализ
6	Многомерный статистический анализ и методы снижения размерности	Выбор метода в зависимости от задачи. Применение многомерных методов в анализе данных.

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Попов, А. М. Теория вероятностей : учебник для вузов / А. М. Попов, В. Н. Сотников. — 3-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2026. — 179 с. — (Высшее образование). — ISBN 978-5-534-18266-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/583928

2. Сидняев, Н. И. Теория планирования эксперимента и анализ статистических данных : учебник и практикум для вузов / Н. И. Сидняев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2026. — 495 с. — (Высшее образование). — ISBN 978-5-534-05070-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/582529

Дополнительная литература:

1. Энатская, Н. Ю. Математическая статистика и случайные процессы : учебник для вузов / Н. Ю. Энатская. — Москва : Издательство Юрайт, 2026. — 191 с. — (Высшее образование). — ISBN 978-5-9916-9808-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/583925

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Электронный документ

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование"	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модуле), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное

Электронный документ

Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Многомерная теория вероятностей и статистика» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекция, семинары, контрольные работы, домашние задания, коллоквиум, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удается, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы - получить специальные знания по одной или нескольким темам дисциплины (модуля) и продемонстрировать навыки их практического применения.

Коллоквиум – устные ответы на вопросы, список которых известен студенту заранее.

В процессе подготовки к коллоквиуму необходимо проанализировать учебные материалы, ознакомившись с лекциями, учебниками и дополнительными источниками, акцентируя внимание на ключевых темах. Рекомендуется создать структурированные конспекты, выделяя основные идеи, термины и формулы.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модуле) «Многомерная теория вероятностей и статистика»

Оценивание уровня учебных достижений, обучающихся по дисциплине (модулю), осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме ***зачета с оценкой***, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

Десятибалльная оценка	**Пятибалльная оценка**	**Оценка за зачет**	**Общая характеристика результата обучения по дисциплине (модулю)**
10	Отлично	Зачтено	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину. Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи.
9	Отлично	Зачтено	
8	Отлично	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Зачтено	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	Зачтено	
5	Удовлетворительно	Зачтено	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь
4	Удовлетворительно	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			минимальным набором методов исследования.
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Многомерная теория вероятностей и статистика» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	15%	14	Набор задач по темам недели
Контрольные работы	30%	2	Письменная работа с набором задач, которые нужно решить за ограниченное время
Коллоквиум	20%	1	Устные ответы на вопросы, список которых известен студенту заранее
Зачет с оценкой	35%	1	Итоговая работа

Формула расчёта итоговой оценки по дисциплине (модуле) «Многомерная теория вероятностей и статистика»: «0,15 × среднее за домашние задания + 0,3 × среднее за контрольные работы + 0,2 ×коллоквиум + 0,35 ×зачет с оценкой».

Текущий контроль успеваемости обучающихся по дисциплине (модуле)

Примерные домашние задания

Тема 1: Случайные векторы и многомерные распределения

1. Дан случайный вектор $\mathbf{X} = (X_1, X_2)^\top$ с $\mathbb{E}[\mathbf{X}] = (2, -1)^\top$ и

$$\mathrm{Cov}(\mathbf{X}) = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}.$$

   Найдите $\mathbb{E}[X_1 + 2X_2]$ и $\mathrm{Var}(X_1 + 2X_2)$.

2. Постройте выборочную ковариационную и корреляционную матрицы для выборки из 5 наблюдений:

$$(1, 2), (2, 3), (3, 1), (4, 4), (5, 2).$$

3. Докажите, что ковариационная матрица всегда симметрична и положительно полуопределённа.

Тема 2: Многомерные предельные теоремы и оценивание

4. Сформулируйте и докажите многомерную ЦПТ для независимых одинаково распределённых векторов.

5. Пусть $\mathbf{X}_1, \ldots, \mathbf{X}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Выведите оценку максимального правдоподобия для $\boldsymbol{\mu}$ и $\boldsymbol{\Sigma}$.
6. Смоделируйте в Python выборку из двумерного нормального распределения с заданными $\boldsymbol{\mu}$ и $\boldsymbol{\Sigma}$, постройте диаграмму рассеяния и эллипс равной плотности.

Тема 3: Проверка гипотез и регрессия

7. Проверьте гипотезу $H_0 : \boldsymbol{\mu} = (0, 0)^\top$ по выборке из 20 наблюдений с $\bar{\mathbf{x}} = (0.8, -0.5)^\top$,

[

\mathbf{S} = \begin{pmatrix} 2 & 0.5 \ 0.5 & 3 \end{pmatrix}.

]

Используйте статистику Хотеллинга T^2, уровень значимости $\alpha = 0.05$.

8. Постройте модель многомерной регрессии $\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$ для данных:

$$\mathbf{Y} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

   Найдите оценку $\hat{\mathbf{B}}$.

Тема 4: Снижение размерности и кластеризация

9. Примените метод главных компонент к выборке из задания 2. Найдите первую главную компоненту и долю объяснённой дисперсии.

10. Реализуйте алгоритм K-средних (k=2) вручную для точек: (1,1), (1,2), (3,4), (4,5). Проведите один шаг итерации.
11. Постройте би-график (biplot) для PCA в Python (используя `matplotlib` и `sklearn`).

Примерные задания по контрольной работе

1. Дан случайный вектор $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, где

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}.$$

   Найдите распределение величины $Y = X_1 - X_2$, а также её математическое ожидание и дисперсию.

2. По выборке из 25 наблюдений построен вектор средних $\bar{\mathbf{x}} = (1.2, 0.8)^\top$, выборочная ковариационная матрица

$$\mathbf{S} = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}.$$

   Проверьте гипотезу $H_0 : \boldsymbol{\mu} = (0, 0)^\top$ на уровне значимости $\alpha = 0.05$. Укажите используемую статистику, её распределение при нулевой гипотезе и сделайте вывод.

3. Для выборки из четырёх точек: (1,3), (2,1), (3,4), (4,2):
   - Вычислите выборочную ковариационную матрицу.
   - Найдите направление первой главной компоненты.
   - Определите долю общей дисперсии, объяснённую первой компонентой.

4. Даны две группы переменных: $\mathbf{X} = (X_1, X_2)$ и $\mathbf{Y} = (Y_1)$.
   Опишите, как находятся канонические корреляции между этими группами. Каково максимальное число ненулевых канонических корреляций в этом случае?

5. Пусть $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, и $\mathbf{X} \sim N(\mathbf{0}, \mathbf{I})$. Найдите распределение вектора $\mathbf{Y} = \mathbf{AX}$, укажите его математическое ожидание и ковариационную матрицу.

6. Опишите основные шаги алгоритма K-средних. Почему результат кластеризации может зависеть от выбора начальных центров кластеров?

Примерные задания для коллоквиума

1. Дайте определение случайного вектора и его математического ожидания.
2. Что такое ковариационная матрица случайного вектора? Каковы её свойства?
3. Выведите формулу для ковариационной матрицы линейного преобразования **Y**=**AX**+**b**.
4. Чем выборочная ковариационная матрица отличается от теоретической? Как она вычисляется?
5. Почему выборочная ковариационная матрица становится сингулярной, если число наблюдений меньше числа признаков?
6. Сформулируйте многомерную центральную предельную теорему.
7. Как выглядит плотность многомерного нормального распределения? Перечислите его основные свойства.
8. Что такое совместное, маргинальное и условное распределение в контексте многомерного нормального закона?
9. Выведите формулу условного математического ожидания $E[X1|X2=x2]$ для двумерного нормального распределения.
10. Выведите формулу условной дисперсии $\mathrm{Var}(X1|X2=x2)$ в случае двумерного нормального распределения.
11. Как строится выборочная корреляционная матрица? Чем она отличается от ковариационной?
12. Что такое статистика Хотеллинга $T2$? Для чего она используется?
13. Как проверяется гипотеза о равенстве вектора средних нулю с помощью статистики $T2$?
14. Каково распределение статистики Хотеллинга $T2$ при справедливости нулевой гипотезы?

Электронный документ

15. Запишите модель многомерной линейной регрессии в матричной форме.

16. Как оцениваются коэффициенты в многомерной регрессии? Приведите формулу оценки.

17. В чём разница между маргинальным и условным распределением компонент случайного вектора?

18. Объясните идею метода главных компонент (PCA).

19. Почему перед применением PCA данные центрируют (и, при необходимости, нормируют)?

20. Как находятся главные компоненты с помощью собственных векторов ковариационной матрицы?

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1	Укажите основное применение неравенства Хёфдинга в контексте медианы средних.	Оценка отклонений	УК-1
2	Найдите верхнюю оценку вероятности для субгауссовой случайной величины с параметром σ=1 при t=2.	2e^{-2}	УК-1
3	Определите, что такое субэкспоненциальное распределение в одном слове.	Тяжелый хвост	УК-1
4	Укажите количество независимых величин, необходимое для применения неравенства Чернова с ε=0.1.	100	УК-1
5	Найдите минимальное значение n для концентрации нормы случайного вектора в высоких размерностях.	1	УК-2
6	Укажите применение PCA в анализе субгауссовых векторов.	Кластеризация	УК-2
7	Определите, что такое сетки в теории случайных матриц в одном слове.	Покрытия	УК-2
8	Найдите верхнюю оценку операторной нормы субгауссовой матрицы при t=3.	2d e^{-9/2}	УК-2
9	Укажите неравенство, используемое для концентрации на сфере.	Липшицевых функций	ОПК-1
10	Найдите размерность после применения леммы Джонсона-Линденштрауса для ε=0.1.	100	ОПК-1
11	Определите, что такое декуплинг в одном слове.	Разделение	ОПК-1
12	Укажите применение неравенства Хэнсона-Райта.	Завершение матриц	ОПК-1
13	Найдите гауссову ширину единичного шара в R^2.	π	ПК-1
14	Укажите неравенство для эмпирических процессов.	Дадли	ПК-1
15	Определите, что такое VC-размерность в одном слове.	Сложность	ПК-1
16	Найдите оценку отклонения случайной матрицы на множествах.	Матричное неравенство	ПК-1
17	Укажите применение теоремы Дворецкого-Милмана.	Геометрия норм	ПК-2
18	Найдите матричное неравенство для зависимых данных.	Бернштейна	ПК-2
19	Определите, что такое принцип контракции в одном слове.	Сжатие	ПК-2
20	Укажите комбинацию методов для задач машинного обучения.	Концентрация и матрицы	ПК-2