

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Тамбовский государственный университет имени Г.Р. Державина»

Институт математики, физики и информационных технологий

Кафедра математического моделирования и информационных технологий



УТВЕРЖДАЮ:

Директор института математики,
физики и информационных
технологий.

Королева Н.Л.

«29» июня 2022 г.

РАБОЧАЯ ПРОГРАММА

по дисциплине

«Методы анализа и обработки данных в научных исследованиях»

Научная специальность:

2.3.1. Системный анализ, управление и обработка информации, статистика

Уровень высшего образования

подготовка кадров высшей квалификации
по программам подготовки научных и
научно-педагогических кадров в аспирантуре

Форма обучения

очная

Год набора

2022

Автор программы: Ковалева Ольга Александровна, доктор технических наук, доцент

Рабочая программа составлена в соответствии с федеральными государственными требованиями к структуре программ подготовки научных и научно-педагогических кадров в аспирантуре (адъюнктуре), условиям их реализации, срокам освоения этих программ с учетом различных форм обучения, образовательных технологий и особенностей отдельных категорий аспирантов (адъюнктов) (приказ Минобрнауки России от 20 октября 2021 г. № 951).

Рабочая программа принята на заседании кафедры математического моделирования и информационных технологий «29» июня 2022 года Протокол № 12

СОДЕРЖАНИЕ

1. Цели и задачи дисциплины
2. Место дисциплины в структуре программы аспирантуры
3. Объем и содержание дисциплины
4. Контроль знаний обучающихся
5. Учебно-методическое и информационное обеспечение дисциплины
6. Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы

1. Цели и задачи дисциплины

1.1 Цель дисциплины – приобретение аспирантами теоретических и практических знаний, умений и навыков в области анализа и обработки экспериментальных данных, подготовка к выбору основных факторов эксперимента, подбору эмпирических зависимостей для экспериментальных данных, оценке коэффициентов регрессионной модели эксперимента, построению оптимальных планов для научно-технических экспериментов.

1.2 Задачи дисциплины:

- получение теоретических знаний и практических навыков по планированию эксперимента, приемов, навыков по выполнению научных и промышленных экспериментальных исследований с помощью методов обработки экспериментальных данных;
- формирование умения демонстрировать базовые знания планирования эксперимента и приобретать новые научные и профессиональные знания по дисциплине;
- формирование навыков применения программных продуктов и информационно-коммуникационных технологий для анализа результатов экспериментальных научных исследований.

1.3 Требования к результатам освоения дисциплины

В результате освоения дисциплины аспирант должен:

Знать:

- современные методы обработки экспериментальных данных;
- сущность задачи аппроксимации экспериментальных данных;
- методы анализа результатов научного эксперимента;
- основные математические пакеты для анализа данных;
- теоретические основы применения статистических методов.

Уметь:

- подбирать конкретный метод анализа экспериментальных данных;
- анализировать процессы в соответствующей профессиональной области;
- получать нужную информацию из результатов научных исследований;
- использовать современное программное обеспечение для ПК с целью анализа процессов в профессиональной деятельности;
- использовать статистические методы обработки и анализа массовых экспериментальных данных в научных исследованиях различной направленности;
- делать точечные оценки параметров регрессионной модели и анализировать их свойства.

Владеть:

- навыками обработки и анализа информации по теме исследования;
- навыками работы в современных математических пакетах для интерполяции и аппроксимации;
- навыками работы в математических пакетах для приближенного поиска минимума ошибки модели с целью определения ее параметров;
- методами подбора эмпирических зависимостей для экспериментальных данных;
- навыками вычисления коэффициентов корреляции и детерминации для оценки применимости рассматриваемой модели процесса;
- методами оценки коэффициентов регрессионной модели эксперимента.

2. Место дисциплины в структуре программы аспирантуры:

Дисциплина «Методы анализа и обработки данных в научных исследованиях» относится к образовательному компоненту «Дисциплины (модули)» программы

аспирантуры по научной специальности 2.3.1. Системный анализ, управление и обработка информации, статистика. Дисциплина является элективной.

Дисциплина «Методы анализа и обработки данных в научных исследованиях» изучается во 2 семестре.

3. Объём и содержание дисциплины

3.1 Объём дисциплины

Очная форма обучения: 2 з.е.

Вид учебной работы	Очная форма обучения (всего часов)
Общая трудоёмкость дисциплины	72
<i>Контактная работа (по учебным занятиям)</i>	22
Лекции (Л)	10
Практические (семинарские) занятия (ПЗ)	12
Лабораторные занятия (ЛЗ)	-
<i>Самостоятельная работа (СР)</i>	50
<i>Зачет</i>	

3.2 Содержание дисциплины:

№ те мы	Название раздела/темы	Вид учебной работы, час. (очная форма)				Формы текущего контроля
		Л	ПЗ	ЛЗ	СР	
1.	Тема 1. Основные понятия	2	2	-	10	опрос
2.	Тема 2. Методы интерполяции данных. Примеры программ в математических пакетах	2	4	-	15	опрос
3.	Тема 3. Методы построения аппроксимирующих функций по экспериментальным данным. Использование математических пакетов	4	4	-	15	опрос
4.	Тема 4. Генераторы случайных процессов	2	2	-	10	опрос

Тема 1. Основные понятия

Лекция.

Понятия интерполяции и экстраполяции данных. Место анализа данных в различных областях человеческой деятельности.

Постановка задачи интерполяции и аппроксимации. Применение на практике.

Языки программирования для анализа данных: Python, математические пакеты Mathcad и MATLAB.

Практическое занятие.

Обработка экспериментальных данных в среде MathCAD (интерполяция)

Задания для самостоятельной работы:

По основной литературе [1] ознакомиться с постановкой задачи интерполяции и аппроксимации.

Тема 2. Методы интерполяции данных. Примеры программ в математических пакетах

Лекция.

Кусочно-линейная и квадратичная интерполяция. Интерполяционный многочлен Лагранжа. Интерполяционная формула Ньютона. Недостатки и достоинства интерполяционных формул.

Использование сплайнов. Квадратичные и кубические сплайны. Построение интерполяционных многочленов в аналитическом виде в пакете Mathcad. Анализ результатов для разных видов интерполяционных формул.

Практическое занятие.

Построение интерполяционных многочленов в аналитическом виде в пакете Mathcad.

Задания для самостоятельной работы:

1. По основной литературе [1] изучить основные методы интерполяции.
2. В математическом пакете Mathcad или MATLAB произвести построение разных видов интерполяционных формул для данных в исследуемой области, построить графики в одной системе координат и сравнить результаты.

Тема 3. Методы построения аппроксимирующих функций по экспериментальным данным. Использование математических пакетов

Лекция.

Понятие математической модели процесса.

Характер экспериментальных данных и подбор эмпирических функций. Переопределенные системы уравнений. Сущность задачи аппроксимации экспериментальных данных.

Минимизация сумм модулей и квадратов остатков модели. Построение функции ошибки для заданного вида функции-модели, описывающей процесс в исследуемой области. Метод наименьших квадратов (МНК): историческая справка, реализация в математических пакетах MATLAB и Mathcad, нахождение параметров модели исследуемого процесса.

Вывод системы нормальных уравнений. Решение в простейшем случае для линейной регрессии $y(x) = ax + b$. Коэффициенты корреляции и детерминации. Вычисление коэффициентов корреляции и детерминации в пакете Mathcad.

Нелинейная регрессия. Линеаризация. Примеры моделей, не сводящихся к линейным. Обзор численных методов решения экстремальных задач применительно к минимизации суммы квадратов остатков модели. Примеры использования численных методов в пакетах Mathcad и MATLAB для приближенного поиска минимума ошибки модели с целью определения ее параметров. Взвешенный МНК.

Основные понятия теории искусственных нейронных сетей: уравнение нейрона, функция активации, однослойные и многослойные сети. Формирование архитектуры сети в пакете MATLAB. Функция ошибки, характеризующая качество обучения (МНК). Обучение нейронных сетей для аппроксимации экспериментальных данных. Примеры в пакете MATLAB.

Практическое занятие.

Решение задач анализа данных в случае линейной регрессии.

Вычисление коэффициентов корреляции и детерминации в пакете Mathcad.

Примеры использования численных методов в пакетах Mathcad и MATLAB для приближенного поиска минимума ошибки модели с целью определения ее параметров.

Формирование архитектуры искусственной нейронной сети в пакете MATLAB.

Обучение нейронных сетей для аппроксимации экспериментальных данных. Примеры в пакете MATLAB.

Задания для самостоятельной работы:

1. По основной литературе [1, 2] освоить метод наименьших квадратов (МНК), решить задачи о построении уравнений линейной регрессии.

2. В математическом пакете Mathcad найти численные решения задачи определения параметров нелинейных моделей процессов в исследуемой области по экспериментальным данным, используя реализацию МНК в пакете, а также свести к решению экстремальной задачи на минимум суммы квадратов остатков модели.

3. По основной литературе [3] ознакомиться с основными понятиями теории искусственных нейронных сетей.

4. Используя многослойные нейронные сети, в пакете MATLAB построить аппроксимацию экспериментальных данных в исследуемой области.

Тема 4. Генераторы случайных процессов

Лекция.

Генераторы псевдослучайных чисел на ЭВМ: основные алгоритмы. Генерация случайных процессов с заданными характеристиками.

Практическое занятие.

Генерация случайных процессов с заданными характеристиками.

Задания для самостоятельной работы:

По дополнительной литературе [1] изучить моделирование непрерывных случайных величин с заданным законом распределения в MATLAB.

4. Контроль знаний обучающихся

4.1 Формы текущего контроля работы аспирантов: опрос.

4.2 Задания текущего контроля

Вопросы для опроса

1. Постановка задачи аппроксимации функций.
2. Когда применяется интерполяция функций?
3. Постановка задачи интерполяции.
4. Что является исходными данными в задаче аппроксимации функций?
5. Что такое линейная интерполяция и как она реализуется в системе MathCAD?
6. Получить формулу интерполяционного полинома Лагранжа.
7. Реализация вычислений значений интерполяционного полинома
8. Лагранжа в системе MathCAD.
9. Что такое сплайн-интерполяция и как она реализуется в системе
10. MathCAD?
11. Когда применяется линейная интерполяция, интерполяция полиномом
12. Лагранжа, сплайн-интерполяция?
13. Как влияет число узлов на абсолютную погрешность интерполяции?
14. Постановка задачи аппроксимации функции методом наименьших
15. квадратов.
16. Когда применяется метод наименьших квадратов?
17. Получить решение задачи о вычислении коэффициентов
18. аппроксимирующего полинома в методе наименьших квадратов.
19. Сущность аппроксимации методом наименьших квадратов;

20. Принципиальное отличие метода интерполирования от метода
21. наименьших квадратов.
22. Какие группы функций пакета Neural Networks Toolbox вы знаете?
23. Назовите функции, с помощью которых можно создать сеть.
24. Как получить справочную информацию по интересующей функции?
25. Опишите «соревновательную» функцию NNT. Какие параметры можно задавать в соответствующем запросе?
26. Назовите различия использования функций `hardlim(X)` и `hardlims(X)`.
27. Что может быть задано в качестве параметра X? Привести примеры.
28. Охарактеризуйте аргументы `net`, `Pd`, `Tl`, `Ai`, `Q`, `TS`, `VV`, `TV` функции `trainbfg()`.
29. Как можно задать максимальное количество циклов обучения сети и начальные входные условия?
30. Какую информацию возвращает функция `learncon(code)`?
31. Назовите функции, выполняющие одномерную оптимизацию НС.

4.3 Промежуточная аттестация по дисциплине проводится в форме зачета.

Вопросы зачета

1. Понятия интерполяции и экстраполяции данных.
2. Место анализа данных в различных областях человеческой деятельности.
3. Постановка задачи интерполяции и аппроксимации. Применение на практике.
4. Языки программирования для анализа данных: Python, математические пакеты Mathcad и MATLAB.
5. Кусочно-линейная и квадратичная интерполяция.
6. Интерполяционный многочлен Лагранжа.
7. Интерполяционная формула Ньютона.
8. Использование сплайнов. Квадратичные и кубические сплайны.
9. Анализ результатов для разных видов интерполяционных формул.
10. Понятие математической модели процесса.
11. Характер экспериментальных данных и подбор эмпирических функций.
12. Сущность задачи аппроксимации экспериментальных данных.
13. Минимизация сумм модулей и квадратов остатков модели.
14. Построение функции ошибки для заданного вида функции-модели, описывающей процесс в исследуемой области.
15. Метод наименьших квадратов (МНК): историческая справка, реализация в математических пакетах MATLAB и Mathcad, нахождение параметров модели исследуемого процесса.
16. Вывод системы нормальных уравнений. Решение в простейшем случае для линейной регрессии $y(x) = ax + b$.
17. Коэффициенты корреляции и детерминации. Вычисление коэффициентов корреляции и детерминации в пакете Mathcad.
18. Нелинейная регрессия. Линеаризация. Примеры моделей, не сводящихся к линейным.
19. Обзор численных методов решения экстремальных задач применительно к минимизации суммы квадратов остатков модели.
20. Примеры использования численных методов в пакетах Mathcad и MATLAB для приближенного поиска минимума ошибки модели с целью определения ее параметров. Взвешенный МНК.
21. Основные понятия теории искусственных нейронных сетей: уравнение нейрона, функция активации, однослойные и многослойные сети.
22. Формирование архитектуры сети в пакете MATLAB. Функция ошибки, характеризующая качество обучения (МНК).

23. Обучение нейронных сетей для аппроксимации экспериментальных данных. Примеры в пакете MATLAB.
24. Генераторы псевдослучайных чисел на ЭВМ: основные алгоритмы. Генерация случайных процессов с заданными характеристиками.

Задания для зачета

1. Построить интерполяционный многочлен Лагранжа, произвести кусочно-линейную и сплайн-интерполяцию в аналитическом виде в пакете Mathcad для данных по заданию преподавателя. Построить графики в одной системе координат и сравнить результаты.
2. В математическом пакете Mathcad найти численные решения задачи определения параметров нелинейных моделей процессов в исследуемой области по заданию преподавателя, используя реализацию МНК в пакете, а также сведя к решению экстремальной задачи на минимум суммы квадратов остатков модели.
3. Используя многослойные нейронные сети, в пакете MATLAB построить аппроксимацию данных в исследуемой области по заданию преподавателя.

4.4 Шкала оценивания промежуточной аттестации

Оценка	Основные показатели достижения результата
«зачтено»	Демонстрирует высокий уровень знаний современных методов обработки экспериментальных данных; сущности задачи аппроксимации экспериментальных данных; методов анализа результатов научного эксперимента; основных математических пакетов для анализа данных; теоретических основных применения статистических методов.
	Свободно ориентируется в методах анализа экспериментальных данных. Анализирует процессы в соответствующей профессиональной области. Получает нужную информацию из результатов научных исследований. Использует современное программное обеспечение для ПК с целью анализа процессов в профессиональной деятельности; Уверенно использует статистические методы обработки и анализа массовых экспериментальных данных в научных исследованиях различной направленности; делать точечные оценки параметров регрессионной модели и анализировать их свойства.
	В полном объеме владеет навыками обработки и анализа информации по теме исследования; навыками работы в современных математических пакетах для интерполяции и аппроксимации; навыками работы в математических пакетах для приближенного поиска минимума ошибки модели с целью определения ее параметров; методами подбора эмпирических зависимостей для экспериментальных данных; навыками вычисления коэффициентов корреляции и детерминации для оценки применимости рассматриваемой модели процесса; методами оценки коэффициентов регрессионной модели эксперимента. На вопросы отвечает кратко, аргументировано, уверенно, по существу
«не зачтено»	Демонстрирует слабый уровень знаний современных методов обработки экспериментальных данных; сущности задачи аппроксимации экспериментальных данных; методов анализа результатов научного эксперимента; основных математических пакетов для анализа данных; теоретических основных применения статистических методов.
	Не может анализировать процессы в соответствующей

	<p>профессиональной области</p> <p>Не ориентируется в методах анализа экспериментальных данных.</p> <p>Не может получить нужную информацию из результатов научных исследований. Не может использовать современное программное обеспечение для ПК с целью анализа процессов в профессиональной деятельности;</p> <p>Не может использовать статистические методы обработки и анализа массовых экспериментальных данных в научных исследованиях различной направленности; делать точечные оценки параметров регрессионной модели и анализировать их свойства.</p>
	<p>Не владеет навыками обработки и анализа информации по теме исследования; навыками работы в современных математических пакетах для интерполяции и аппроксимации; навыками работы в математических пакетах для приближенного поиска минимума ошибки модели с целью определения ее параметров; методами подбора эмпирических зависимостей для экспериментальных данных; навыками вычисления коэффициентов корреляции и детерминации для оценки применимости рассматриваемой модели процесса; методами оценки коэффициентов регрессионной модели эксперимента.</p> <p>Неуверенно и логически непоследовательно излагает материал.</p> <p>Неправильно отвечает на поставленные вопросы или затрудняется с ответом.</p>

5. Учебно-методическое и информационное обеспечение дисциплины

5.1 Основная литература:

1. Петров, И.Б. Введение в вычислительную математику: учебное пособие [Электронный ресурс] / И.Б. Петров, А.И. Лобанов. — Электрон. дан. — М.: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 352 с. — Режим доступа: <http://www.iprbookshop.ru/94848.html> — Загл. с экрана.
2. Бояршинов, М.Г. Прикладные задачи вычислительной математики и механики: учебное пособие [Электронный ресурс] / М. Г. Бояршинов. — Электрон. дан. — Саратов: Вузовское образование, 2020. — 344 с. — Режим доступа: <http://www.iprbookshop.ru/93067.html> — Загл. с экрана.
3. Далингер, В. А. Теория вероятностей и математическая статистика с применением Mathcad : учебник и практикум для вузов / В. А. Далингер, С. Д. Симонженков, Б. С. Галюкшов. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2020 — 145 с. — (Высшее образование). — ISBN 978-5-534-10080-8. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/452017>.

5.2 Дополнительная литература:

1. Афонин, В.В. Моделирование систем: учебное пособие [Электронный ресурс] / В.В. Афонин, С.А. Федосин. — Электрон. дан. — М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2020. — 269 с. — Режим доступа: <http://www.iprbookshop.ru/52179.html> — Загл. с экрана.
3. Маккинли, У. Python и анализ данных [Электронный ресурс] / У. Маккинли. — Электрон. дан. — Саратов: Профобразование, 2019. — 482 с. — Режим доступа: <http://www.iprbookshop.ru/64058.html> — Загл. с экрана.
4. Барский, А.Б. Введение в нейронные сети: учебное пособие [Электронный ресурс] / А.Б. Барский. — Электрон. дан. — М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2020. — 357 с. — Режим доступа: <http://www.iprbookshop.ru/52144.html> — Загл. с экрана.

6. Материально-техническое обеспечение дисциплины, программное обеспечение, профессиональные базы данных и информационные справочные системы

Для проведения занятий по дисциплине необходимо следующее материально-техническое обеспечение: помещения для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, помещения для самостоятельной работы.

Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Для проведения занятий лекционного типа используются наборы демонстрационного оборудования, обеспечивающие тематические иллюстрации (проектор, ноутбук, экран/ интерактивная доска).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Электронная информационно-образовательная среда

<http://moodle.tsutmb.ru>

Взаимодействие преподавателя и аспиранта в процессе освоения дисциплины осуществляется посредством мультимедийных, гипертекстовых, сетевых, телекоммуникационных технологий, используемых в электронной информационно-образовательной среде университета.

Лицензионное программное обеспечение:

Операционная система Microsoft Windows 10 Home x64

Microsoft Office Профессиональный плюс 2007 12

Kaspersky Endpoint Security для бизнеса – Стандартный Russian Edition. 1500-2499 Node 1 year Educational Renewal Licence

Информационные справочные системы и профессиональные базы данных:

ЭБС «Университетская библиотека онлайн»	http://www.biblioclub.ru
ЭБС «Консультант студента»: Медицина. Здравоохранение, Комплект Гуманитарные науки	http://www.studentlibrary.ru
ЭБС «IPRSMART» (старое название « IPR books»)	http://iprbookshop.ru
ЭБС «Юрайт»	http://www.urait.ru
Сетевая электронная библиотека педагогических вузов	https://e.lanbook.com/
Научная электронная библиотека eLIBRARY.RU	http://elibrary.ru
Государственная информационная система «Национальная электронная библиотека»	https://нэб.пф
Президентская библиотека имени Б.Н. Ельцина	http://www.prilib.ru
Электронный справочник «Информо»	www.informio.ru
Справочная правовая система «Консультант Плюс»	http://www.consultant.ru
Архив научных журналов зарубежных издательств	https://arch.neicon.ru