

**Программа утверждена на заседании  
Ученого Совета химического факультета  
МГУ имени М.В.Ломоносова  
Протокол № 4 от 27 мая 2016 г.**

### **Рабочая программа дисциплины (модуля)**

1. Наименование дисциплины (модуля): **Введение в MATLAB и регрессионный анализ**

Краткая аннотация:

Программа курса «Введение в MATLAB и регрессионный анализ» предназначена для аспирантов, обучающихся на кафедре физической химии. Курс знакомит слушателей с основами языка программирования MATLAB, в том числе его использовании для оценки погрешностей измерений, проверки статистических гипотез и регрессионного анализа. При этом делается акцент на использование этих знаний для решения физико-химических задач. Курс состоит из трёх связанных разделов: основы программирования на языке MATLAB, оценка погрешностей и проверка статистических гипотез, метод наименьших квадратов.

2. Уровень высшего образования– подготовка научно-педагогических кадров в аспирантуре

3. Направление подготовки: 04.06.01 Химические науки.

4. Место дисциплины (модуля) в структуре ООП: вариативная часть ООП, блок 1 «Дисциплины (модули)». Дисциплина по выбору аспиранта в осеннем семестре второго или третьего года обучения.

5. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников)

Формируемые компетенции (код компетенции)	Планируемые результаты обучения по дисциплине (модулю)
УК-2 способность проектировать и осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного системного научного мировоззрения с использованием знаний в области истории и философии науки	31 (УК-2) <b>Знать</b> методы научно-исследовательской деятельности
ОПК-1 способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий	У1 (ОПК-1) <b>Уметь</b> выбирать и применять в профессиональной деятельности экспериментальные и расчетно-теоретические методы исследования
ПК-4 Способность к самостоятельному проведению научно-исследовательской работы и получению научных результатов, удовлетворяющих установленным требованиям к содержанию диссертаций на соискание ученой степени кандидата наук по направленности (научной специальности) 02.00.04 Физическая химия	35 (ПК-4) <b>Знать</b> современные методы регрессионного анализа и языка матричных вычислений
	36 (ПК-4) <b>Знать</b> современное состояние методов регрессионного анализа для решения физико-химических задач, в т.ч. термодинамического моделирования и использования метода CALPHAD
	У6 (ПК-4) <b>Уметь</b> использовать и программировать современные методы регрессионного анализа при решении практических задач

6. Объем дисциплины (модуля) в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся:

*Объем дисциплины (модуля) составляет 3 зачетные единицы, всего 108 часов, из которых 76 часов составляет контактная работа аспиранта с преподавателем (18 часов занятия лекционного типа, 4 часа мероприятия текущего контроля успеваемости и промежуточной аттестации, 54 часа занятий семинарского типа), 36 часов составляет самостоятельная работа учащегося.*

7. Входные требования для освоения дисциплины (модуля), предварительные условия.

В специалитете или магистратуре должны быть освоены общие курсы «Математический анализ», «Линейная алгебра», «Теория вероятностей», «Информатика», «Физическая химия».

8. Содержание дисциплины (модуля), структурированное по темам.

Наименование и краткое содержание разделов и тем дисциплины (модуля), форма промежуточной аттестации по дисциплине (модулю)	Всего (часы)	В том числе								
		Контактная работа (работа во взаимодействии с преподавателем), часы из них						Самостоятельная работа обучающегося, часы из них		
		Занятия лекционного типа	Занятия семинарского типа	Групповые консультации	Индивидуальные консультации	Текущий контроль успеваемости, промежуточная аттестация	Всего	Выполнение домашних заданий	Работа с оригинальной литературой, подготовка, подготовка	Всего
Тема 1. Основы программирования на языке MATLAB. Основные конструкции языка. Скрипты и функции (в файле, вложенные, анонимные). Векторизация кода. Графика и файловый ввод-вывод. Отладка программ.	36	6	18				24	6	6	12
Тема 2. Оценка погрешностей и проверка стати-	24	4	12				16	4	4	8

стических гипотез. Погрешности, стандартные отклонения, доверительные интервалы. Распределения и квантили. Погрешности чисел с плавающей запятой и машинный ноль. Проверка статистических гипотез с помощью критериев Стьюдента, Фишера и Пирсона.										
Тема 3. Метод наименьших квадратов. Переопределённые системы уравнений и метод наименьших квадратов. Одно- и многомерная линейная регрессия, погрешности коэффициентов и стандартная ошибка регрессии. Критерии Стьюдента и Фишера применительно к регрессии. Нелинейная регрессия, методы Ньютона, Гаусса-Ньютона, Левенберга-Марквардта. Дуальные числа и автоматическое дифференцирование. Мультиколлинеарность и гребневая регрессия. Методы глобальной оптимизации: метод отжига, генетические алгоритмы, символьная регрессия. Использование метода наименьших квадратов при решении физико-химических задач (построении фазовых диаграмм, расчёта химических равновесий, описания кинетических кривых и т.п.)	44	8	24				32	6	6	12
Промежуточная аттестация <u>зачет</u>	4					4	4			
<b>Итого</b>	108	18	54			4	76	16	16	32

## 8. Образовательные технологии

Лекции с презентациями и интерактивной демонстрацией решения задач на персональном компьютере. Семинарские занятия с решением задач на персональных компьютерах в Microsoft Excel (или LibreOffice) и среде MATLAB (или GNU Octave).

9. Учебно-методические материалы для самостоятельной работы по дисциплине (модулю):

Аспирантам предоставляется программа курса, план занятий, перечень домашних заданий, лекционные материалы и примеры программ на языке MATLAB. По теме каждой лекции указывается материал в источниках из списков основной и вспомогательной литературы.

10. Ресурсное обеспечение:

- Перечень основной и вспомогательной учебной литературы ко всему курсу

#### **Основная литература**

1. Дьяконов В.П. MATLAB. Полный самоучитель. ДМК Пресс, 2012.
2. Ю. Н. Тюрин, А. А. Макаров, «Анализ данных на компьютере». М., «Форум», 2008.
3. Д. Химмельблау, «Анализ процессов статистическими методами». М., «Мир», 1973.
4. «Справочник по прикладной статистике». Под ред. Э. Ллойда и У. Ледермана. М., «Финансы и статистика», 1989.
5. К. Дёрффель, «Статистика в аналитической химии». М., «Мир», 1998.
6. Дж. Дэннис мл., Р. Шнабель. Численные методы безусловной оптимизации и решения нелинейных уравнений. М., «Мир», 1988

#### **Дополнительная литература**

1. F.A. Graybill, H.K. Iyer. Regression analysis. Concepts and applications. Duxbury Pr., 1994.
  2. Heij C., de Boer, P., Franses, P. H., Kloek T., van Dijk. Econometric methods with applications in business and economics. Oxford University Press. 2004.
  3. А.В. Гармаш, Н.В. Сорокина «Метрологические основы аналитической химии».
- Перечень используемых информационных технологий, используемых при осуществлении образовательного процесса, включая программное обеспечение, информационные справочные системы (при необходимости):
    - Программное обеспечение для расчётов: MATLAB или GNU Octave, Microsoft Excel или LibreOffice. Возможно прохождение курса с использованием исключительно свободного (бесплатного) программного обеспечения, в т.ч. операционной системы.
    - Справочные системы: NIST Chemistry Webbook (<http://webbook.nist.gov/chemistry/>), БД «Термические константы веществ» (<http://www.chem.msu.ru/cgi-bin/tkv.pl?show=welcomes.html/welcomes.html>)

11. Язык преподавания – русский

12. Преподаватели:

канд. хим. наук, старший научный сотрудник Восков Алексей Леонидович, [alvoskov@gmail.com](mailto:alvoskov@gmail.com), 8-495-939-22-80

канд. хим. наук, доцент Коваленко Никита Андреевич, [Nik-Kovalenko@yandex.ru](mailto:Nik-Kovalenko@yandex.ru)

### **Фонды оценочных средств, необходимые для оценки результатов обучения**

1. Планируемые результаты обучения для формирования компетенций п.5 и соответствующие им критерии оценивания приведены в Приложении 1.
2. Примеры вопросов к семинарам
  - Погрешности: абсолютная, относительная, случайная, систематическая. Правила округления. Косвенные измерения.
  - Дисперсия и стандартное отклонение. Квантили и доверительные интервалы.
  - Проверка статистических гипотез. Критерии Стьюдента, Фишера, Пирсона. Проверка распределения на нормальность.
  - Числа с плавающей запятой и машинный ноль. Оценка погрешности формул численного дифференцирования.
  - Линейная регрессия. Коэффициенты регрессии, их доверительных интервалов, стандартной ошибки регрессии. Критерий Стьюдента и Фишера.
  - Нелинейная регрессия: решение путём линеаризации. Влияние линеаризации на значение параметров.
  - Нелинейная регрессия: численные методы (Ньютона, Гаусса-Ньютона и Левенберга-Марквардта).
  - Матрица Якоби в нелинейной регрессии и её роль в оценке погрешностей. Способы её расчёта: аналитическое, численное и автоматическое дифференцирование.
  - Регуляризация по Тихонову и гребневая регрессия.
  - Методы глобальной оптимизации. Метод отжига и генетические алгоритмы. Символьная регрессия.
  - Использование численных методов для расчёта фазовых и химических равновесий.

- Оптимизация параметров физико-химических моделей по экспериментальным данным

### 3. Примеры домашних заданий

- Построить график функции  $f(x,y) = ue^{x^2-y^2}$  в виде поверхности.
- Рассчитать объём полусферы с  $r=2$  с помощью метода Монте-Карло, используя векторизацию кода.
- Построить изображение множества Мандельброта, используя векторизацию кода.
- Одномерная линейная регрессия.
- Многомерная линейная регрессия.
- Нелинейная регрессия с линеаризацией (описание экспериментальных данных по теплоёмкости и химической кинетике)
- Нелинейная регрессия с использованием метода Левенберга-Марквардта (оптимизация параметров термодинамической модели)
- Линейная гребневая регрессия.
- Найти глобальный минимум двумерной функции Растригина с помощью генетического алгоритма.
- Найти глобальный минимум двумерной функции Растригина с помощью метода отжига.
- Рассчитать кривую титрования  $H_3PO_4$ , рассматривая диссоциацию одновременно по всем ступеням и полагая коэффициенты активности равными единице.
- Построить  $p$ - $x$  и  $T$ - $x$  диаграммы системы «ацетон-хлороформ», найдя нужные параметры стабильности в справочнике и оптимизировав параметр модели регулярных растворов, используя экспериментальные данные.

### 4. Вопросы к зачету:

#### A1. Теоретические вопросы по погрешностям

Абсолютная и относительная погрешность. Случайная и систематическая погрешность. Погрешность косвенных измерений. Правила округления.

Смещённая и несмещённая оценка дисперсии. Стандартное отклонение. Квантили и доверительные интервалы (на примере нормального распределения и распределения Стьюдента)

Проверка статистических гипотез. Критерии Стьюдента, Фишера, Пирсона. Проверка распределения на нормальность.

Числа с плавающей запятой и машинный ноль. Оценка погрешности формул численного дифференцирования с выбором оптимального шага.

## A2. Теоретические вопросы по методу наименьших квадратов

Метод наименьших квадратов. Линейная регрессия. Вывод формул для коэффициентов регрессии, их доверительных интервалов, стандартной ошибки регрессии. Критерий Стьюдента и Фишера.

Метод наименьших квадратов. Нелинейная регрессия и способы линеаризации (полиномы, функции типа уравнения Аррениуса, уравнений ферментативной кинетики и т.п.). Влияние линеаризации на значение параметров. Преимущества и недостатки линеаризации по сравнению с прямым решением задачи.

Метод наименьших квадратов. Нелинейная регрессия и численные методы. Методы Ньютона, Гаусса-Ньютона и Левенберга-Марквардта, их сравнение друг с другом. Матрица Якоби и оценка доверительных интервалов.

Матрица Якоби в нелинейной регрессии и её роль в оценке погрешностей. Способы её расчёта: аналитическое, численное и автоматическое дифференцирование.

Мультиколлинеарность и обусловленность задачи. Регуляризация по Тихонову и гребневая регрессия (ридж-регрессия). Смещённые и несмещённые оценки параметров модели.

Методы глобальной оптимизации. Метод отжига и генетические алгоритмы. Символьная регрессия. Выбор варианта метода наименьших квадратов для решения конкретной задачи.

## B1. Практическая задача по основам MATLAB

Проверка нормальности распределения

Практическая реализация генетического алгоритма

Практическая реализация метода Монте-Карло для численного интегрирования объёма фигуры

## B2. Практическая задача по физико-химическим расчётам

Реализация метода Ньютона с векторизацией кода на примере расчёта рН раствора (кривые титрования)

Построение  $p$ - $x$  и  $T$ - $x$  диаграмм «жидкость-пар» двухкомпонентной системы с азеотропом

Построение диаграммы плавкости двухкомпонентной системы с эвтектикой. Оптимизация параметра модели.

### **Методические материалы для проведения процедур оценивания результатов обучения**

Зачет проводится по билетам, каждый из которых включает один теоретический вопрос и одно практическое контрольное задание (ПКЗ). Уровень знаний аспиранта оценивается на «зачтено», «незачтено». Оценка «зачтено» выставляется, если по шкале оценивания учащийся демонстрирует знания умения и владения, соответствующие категориям 3, 4 и 5. В ходе зачета, проводимого в форме индивидуального собеседования, оценивается степень сформированности «знаниевой» компоненты компетенций УК-2 и ПК-1 (знание современного состояния науки в области методов исследования полупроводников). Частично сформированность умения выбирать и применять в профессиональной деятельности экспериментальные и расчетно-теоретические методы исследования (ОПК-1) проверяется при выполнении ПКЗ, их оценка учитывается как одна из составляющих при выставлении зачета.

**Оценочные средства для промежуточной аттестации по дисциплине «Введение в MATLAB и регрессионный анализ» на основе карт компетенций выпускников**

РЕЗУЛЬТАТ ОБУЧЕНИЯ по дисциплине (модулю)	КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю) и ШКАЛА оценивания					ПРОЦЕДУРЫ ОЦЕНИВАНИЯ*
	1	2	3	4	5	
<p align="center">В1 (УК-1)</p> <p><b>Владеть</b> навыками анализа методологических проблем, возникающих при решении исследовательских и практических задач, в том числе в междисциплинарных областях</p>	Отсутствие владения	Неполное владение отдельными навыками численного анализа и обработки больших массивов разнородных данных	Владение стандартными приемами численного анализа и обработки больших массивов разнородных данных	Довольно полное, но содержащее отдельные пробелы владение численного анализа и обработки больших массивов разнородных данных	Сформированное систематическое владение численного анализа и обработки больших массивов разнородных данных	Выполнение практического контрольного задания в ходе зачета
<p align="center">З1 (УК-2)</p> <p><b>Знать</b> методы научно-исследовательской деятельности</p>	Отсутствие знаний	Фрагментарные знания о методах численного анализа и обработки больших массивов разнородных данных	Несистематические знания о методах численного анализа и обработки больших массивов разнородных данных	Сформированные, но содержащие отдельные пробелы представления о методах численного анализа и обработки больших массивов разнородных данных	Сформированные и систематические знания о методах численного анализа и обработки больших массивов разнородных данных	Устный опрос в ходе зачета

1	2	3	4	5	6	7
У1(ОПК-1) <b>Уметь</b> выбирать и применять в профессиональной деятельности экспериментальные и расчетно-теоретические методы исследования	Отсутствие умений	Умение использовать современные методы регрессионного анализа и языки матричных вычислений	Шаблонное умение использовать современные методы регрессионного анализа и языки матричных вычислений	В целом достаточно полное, но содержащее отдельные пробелы умение использовать современные методы регрессионного анализа и языки матричных вычислений	Сформированное умение грамотно выбирать и использовать современные методы регрессионного анализа и языки матричных вычислений	Выполнение практического контрольного задания в ходе зачета; выполнение домашних заданий
35 (ПК-4) <b>Знать</b> современные методы регрессионного анализа и языки матричных вычислений	Отсутствие знаний	Фрагментарные представления о регрессионном анализе и языках матричных вычислений	Ограниченные представления о регрессионном анализе и языках матричных вычислений	Сформированные, но содержащие отдельные пробелы знания о регрессионном анализе и языках матричных вычислений	Сформированные и систематические знания о регрессионном анализе и языках матричных вычислений	Устный опрос в ходе зачета
36 (ПК-4) <b>Знать</b> современное состояние методов регрессионного анализа для решения физико-химических задач, в т.ч. термодинамического модели-	Отсутствие знаний	Фрагментарные представления о применении методов регрессионного анализа для решения физико-химических задач	Ограниченные представления о применении методов регрессионного анализа для решения физико-химических задач	Сформированные, но содержащие отдельные пробелы знания о применении методов регрессионного анализа для решения физико-химических задач	Сформированные и систематические знания о применении методов регрессионного анализа для решения физико-химических задач	Устный опрос в ходе зачета

рования и использования метода CALPHAD						
У6 (ПК-4) <b>Уметь</b> использовать и программировать современные методы регрессионного анализа при решении практических задач	Отсутствие умений применять программное обеспечение	Разрозненные умения использовать современные методы регрессионного анализа при решении практических физико-химических задач	Умение использовать современные методы регрессионного анализа при решении практических физико-химических задач, но сопряженное с частыми ошибками в выборе оптимального метода	В большинстве случаев успешные умения использовать современные методы регрессионного анализа при решении практических физико-химических задач, но выбор программ не всегда является оптимальным и их возможности не всегда полностью используются	хорошо развитые способности использовать современные методы регрессионного анализа при решении практических физико-химических задач	Выполнение практического контрольного задания в ходе зачета