

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Московский государственный университет имени М.В. Ломоносова»  
Химический факультет

УТВЕРЖДАЮ

Декан химического факультета,  
акад. РАН, профессор



/В.В. Лунин/

«14» июня 2015 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**  
**Многомерные данные в химическом анализе**

**Уровень высшего образования:**  
Подготовка кадров высшей квалификации

---

**Направление подготовки (специальность):**  
04.06.01 Химические науки

**Направленность (профиль) ОПОП:**  
Аналитическая химия 02.00.02

**Форма обучения:**  
очная

---

Рабочая программа рассмотрена и одобрена  
Учебно-методической комиссией факультета  
(протокол №3 от 03.06.2015)

Москва 2015

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки 04.06.01 «Химические науки» на основе Образовательного стандарта, самостоятельно установленного МГУ имени М.В.Ломоносова (далее – ОС МГУ), утвержденного Приказом № 552 от 23.06.2014 г. по МГУ с учетом изменений в ОС МГУ, внесенных Приказом №831 по МГУ от 31.08.2015 г..

Год (годы) приема на обучение 2014/2015, 2015/2016, 2016/2017, 2017/2018,  
2018/2019, 2019/ 2020

1. **Цели дисциплины** - освоение аспирантами фундаментальных знаний в области хемометрики, изучение основных методов обработки многомерных данных химического анализа.

**Задачи дисциплины:**

- Формирование базовых знаний в области хемометрики как научной дисциплины, интегрирующей математическую и химико-аналитическую подготовку химиков и обеспечивающей методологические основы современных инновационных сфер деятельности;
- Обучение студентов основам хемометрических методов обработки многомерных данных и применения их в химическом анализе;
- Формирование подходов к выбору хемометрических методов, адекватных поставленной прикладной задаче

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, Модуль научной специальности и отрасли наук (спецкурсы по выбору кафедры) (2 семестр 1 год обучения).

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

Компетенция	Планируемые результаты обучения по дисциплине (модулю)
<b>СПК-1</b> Способность разрабатывать методическое, математическое и метрологическое обеспечение химического анализа	<b>Знать:</b> фундаментальные основы хемометрики <b>Знать:</b> основные тенденции развития хемометрики <b>Уметь:</b> обрабатывать многомерные данные химического анализа в соответствии с современными требованиями

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 зачетных единиц, всего 108 часов, из которых 58 часов составляет контактная работа аспиранта с преподавателем (24 часа - занятия лекционного типа, 24 часов - семинары, 8 часов индивидуальные консультации, 2 часа - мероприятия текущего контроля, 6 часов - мероприятия промежуточной аттестации), 50 часов составляет самостоятельная работа аспиранта.

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

В специалитете или магистратуре должны быть освоены общие курсы «Аналитическая химия», «Линейная алгебра», «Элементы прикладной математической статистики», «Основы хемометрики».

8. Образовательные технологии (отметить если применяется электронное обучение и дистанционные технологии).

Используются следующие технологии: лекции-демонстрации и интерактивные лекции. Преподавание дисциплин проводится в форме авторских курсов по программам, которые составлены на основе результатов исследований, полученных научными школами МГУ.

9. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических или астрономических часов и виды учебных занятий

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	Учебные занятия, направленные на проведение текущего контроля успеваемости коллоквиумы, практические контрольные занятия и др)*	Всего	Выполнение домашних заданий	Подготовка рефератов и т.п..	Всего
Раздел 1. Многомерные данные в химическом анализе и основные опе-	27	6	6		5			-	-	10

рации с ними										
Раздел 2. Количественный анализ многокомпонентных систем	35	9	8		4	2		-	-	12
Раздел 3. Методы многомерной классификации и идентификации в химическом анализе	24	6	5		3					10
Раздел 4. Методы многомерного разрешения кривых	16	3	5			2				6
<b>Промежуточная аттестация зачет</b>	6						6			
<b>Итого</b>	<b>108</b>	<b>24</b>	<b>24</b>	<b>-</b>	<b>12</b>	<b>4</b>	<b>64</b>			<b>38</b>

Распределение трудоемкости по разделам и темам, а также формам занятий с указанием форм текущего контроля и промежуточной аттестации :

N раздела	Наименование раздела	Трудоёмкость (академических часов) и содержание занятий				Форма текущего контроля
		Аудиторная работа		Самостоятельная работа		
		Лекции	Семинары			
1	<b>Многомерные данные в химическом ана-</b>	№1. 3 часа. Предмет хемометрики,, ее задачи. Полезная информация и шум.	№1. 3 часа. Основные приемы работы с многомерными данными в среде Excel с помощью про-	7 часов. Работа с лекционным материалом. Подготовка ответов на контроль-	ДЗ	

	<b>лизе и основные операции с ними</b>	Многомерные данные в химическом анализе. Аналитический сигнал, аналитический признак, аналитическая позиция. Представление спектральных, хроматографических и других данных в векторной форме. Матрицы "объект-свойство". Программное обеспечение для работы с многомерными данными	граммы MVA	ные вопросы по теме лекции.. Домашнее задание: матричные вычисления с использованием программы MVA	
		№2. 3 часа Линейные преобразования матриц данных: сдвиг (центрирование), масштабирование, поворот, проекция. Их назначение, алгебраическая и геометрическая интерпретация. Пространство главных компонент, матрицы счетов и нагрузок, их свойства, сингулярные числа, формальный и эффективный ранг матрицы	№2. 3 часа. Линейные преобразования матриц многомерных аналитических данных. Представление матриц данных в пространстве главных компонент Расчет матриц счетов и нагрузок, изучение их свойств. Расчет сингулярных чисел матрицы, нахождение ее эффективного ранга.	8 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: линейные преобразования матриц аналитических данных. сингулярное разложение матриц данных, свойства матриц счетов и нагрузок, нахождение эффективного ранга матрицы.	
2	<b>Количественный анализ многокомпонентных систем</b>	№3. 3 часа. Многомерный регрессионный анализ. Математическая формулировка задач градуировки и расчета концентраций в матричной форме. Метод Фирордта и метод множественной линейной регрессии. Метод наименьших квадратов. Прямая и обратная градуировки.	№3. 3 часа. Анализ многокомпонентных систем методами Фирордта и множественной линейной регрессии.	7 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: анализ многокомпонентных систем методами Фирордта и множественной линейной регрессии	<b>ДЗ, КР</b>

		№4. 3 часа. Проекционные методы регрессионного анализа. Проекция на главные компоненты (PCA), проекция на скрытые структуры (PLS1, PLS2). Их использование в анализе многокомпонентных систем	№4. 3 часа. Анализ многокомпонентных систем с использованием проекционных методов	8 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: анализ многокомпонентных систем с использованием проекционных методов	
		№5. 3 часа. Метрологические аспекты многокомпонентного анализа. Описательная и предсказательная способность регрессионной модели. Метод "введено-найденно" в многокомпонентном анализе. Градуировочная и проверочная выборки. Перекрестная проверка на достоверность. Среднеквадратичная погрешность. Оптимизация регрессионной модели	№3. 2 часа. Оценка описательной и предсказательной способности регрессионной модели. Изучение влияния погрешности исходных данных на погрешность результатов. Стратегия выбора оптимальной регрессионной модели и ее параметров для анализа многокомпонентных систем	7 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: изучение влияния погрешности исходных данных на погрешность результатов, выбор оптимальной регрессионной модели и ее параметров в анализе многокомпонентных систем	
<b>3</b>	<b>Методы многомерной классификации и идентификации в химическом анализе</b>	№6. 3 часа. Основные принципы многомерной классификации. Классификационные признаки, принципы их выбора. Расстояние от неизвестного объекта до известного класса как главный классификационный критерий. Способы предварительного преобразования пространства признаков и способы вычисления рас-	№6. 3 часа. Выбор признаков, преобразование данных, вычисление расстояний в методах многомерной классификации, графическое представление результатов	8 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: выбор классификационных признаков, их преобразование, графическое представление, построение дендрограмм	<b>ДЗ</b>

		стояний в нем. Расстояние "метрическое" (Минковского) и "статистическое" (Махаланобиса). Графическое представление результатов классификации. Кластерный анализ, дендрограммы			
		№7. 3 часа. Основные методы многомерной классификации: линейный дискриминантный анализ, SIMCA, PLS-классификация, метод К ближайших соседей. Оптимизация классификационной модели.	№3. 2 часа. Классификация химических объектов с использованием различных методов и оценка качества классификации	7 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: классификация химических объектов с использованием различных методов и оценка качества классификации	
4	Методы многомерного разрешения кривых	№8. 3 часа. Постановка задачи разрешения кривых. Регрессионный анализ с ограничениями. Метод чередующихся наименьших квадратов (ALS), итерационного целевого (ITTFA), эволюционного (EFA) и оконного (WFA) факторного анализа	№8. 3 часа. Разрешение спектральных, хроматографических, кинетических данных. Получение спектров чистых компонентов из спектров смесей, нахождение констант сложных равновесий	8 часов. Работа с лекционным материалом. Подготовка ответов на контрольные вопросы по теме лекции. Домашнее задание: выделение спектров чистых компонентов из спектров смесей, расчеты констант кислотности многоосновных кислот, констант устойчивости комплексов в сложных системах	ДЗ, КР

10. Учебно-методические материалы для самостоятельной работы аспирантов.

11. Ресурсное обеспечение:

- Перечень основной и дополнительной учебной литературы

#### А. Основная литература

№ п/п	Автор	Название книги/статьи	Отв. редактор	Место издания	Издательство	Год издания	Название журнала (сборника)	Том, номер, страницы
1	Померанцев А.Л	Хеометрика в Excel: учебное пособие		Томск	Издательство ТПУ	2014		
2	Эсбенсен К.	Анализ многомерных данных. Избранные главы	Родионова О.Е.	Черноголовка	ИПХФ РАН	2005		
3		Аналитическая химия. Проблемы и подходы. Т. 2. Гл. 12	Кельнер Р., Мерме Ж.-М., Отто М., Видмер Г.М.	Москва	Мир, АСТ	2004		
4	Родионова О.Е, Померанцев А.Л.	Хеометрика: достижения и перспективы		Москва		2006	Успехи химии	Т.75. № 4. С.302-321

#### Дополнительная литература

№ п/п	Автор	Название книги/статьи	Отв. редактор	Место издания	Издательство	Год издания	Название журнала (сборника)	Том, номер, страницы
1	Ким Дж., Мюллер Ч. и др.	Факторный, дискриминантный и кластерный анализ		Москва	Финансы и статистика	1989		

2	Шараф М.А., Ил-мэн Д.Л., Ковальски Б.Р.	Хемометрика		Ленинград	Химия	1989		
3	Brereton R.G.	Chemometrics. Data Analysis for the Laboratory and Chemical Plant		Москва	John Wiley & Sons	2003		

### **Программное обеспечение и Интернет-ресурсы**

Решение задач на семинарских занятиях проводится в среде Excel с использованием программы MVA, разработанной на кафедре аналитической химии МГУ. Домашние задания аспиранты могут выполнять также с использованием этой программы либо, по их желанию, в среде Excel с использованием надстройки Chemometrics Add-In (доступна на сайте <http://rscs.chemometrics.ru> и в приложении к книге [1] из списка основной литературы) или в среде Matlab.

### **Материально-техническое обеспечение**

Для чтения курса необходимо наличие в аудитории исправного электрооборудования, классной доски и персональных компьютеров (по одному на каждого слушателя). Возможно проведение занятий в аудитории, не оборудованной компьютерами, при условии, что переносные персональные компьютеры слушатели приносят с собой.

Б. Переносной компьютер (ноутбук), мультимедийный проектор, экран.

12. Язык преподавания - русский

13. Автор курса, лектор: доцент, канд.хим.наук. **ГАРМАШ АНДРЕЙ ВИКТОРОВИЧ**

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

1. Планируемые результаты обучения приведены в п.5.

## 2. Вопросы для подготовки к текущей и промежуточной аттестации

**Текущая аттестация** проводится еженедельно, контрольные работы - после изучения 2 и 4 разделов курса. Критерии формирования оценки – выполнение домашних заданий, посещаемость занятий, активность студентов на лекциях и семинарах, уровень подготовки к лекциям, уровень знаний пройденной части курса.

**Раздел 1. Название раздела** Многомерные данные в химическом анализе и основные операции с ними

**Тема 1. Название темы** Многомерные данные в химическом анализе

**Содержание темы** Предмет хемометрики, ее задачи. Полезная информация и шум. Многомерные данные в химическом анализе. Аналитический сигнал, аналитический признак, аналитическая позиция. Представление спектральных, хроматографических и других данных в векторной форме. Матрицы "объект-свойство". программное обеспечение для работы с многомерными данными.

### **Примеры заданий для семинарских занятий**

1. Задачи для овладения основными приемами расчетов в среде Excel с помощью программы MVA: разметка рабочего поля, маркировка данных, запись формул, выполнение вычислений.
2. Задачи на выполнение основных матричных операций с помощью программы MVA: сложение, умножение, масштабирование, транспонирование, обращение матриц, генерация случайного шума с заданными параметрами и его наложение на матрицу данных.

### **Примеры заданий для самостоятельной работы**

1. Что такое многомерные данные в химическом анализе? Какова их возможная структура, какую информацию о веществе они могут содержать?
2. Что такое аналитический сигнал, аналитический признак, аналитическая позиция?
3. Приведите примеры одномерных (скалярных) и многомерных аналитических сигналов и признаков применительно к конкретному методу анализа (в соответствии со специализацией аспиранта).
4. Как представить в виде многомерных данных электронный спектр поглощения, масс-спектр, хроматограмму, вольтамперограмму, масс-хроматограмму?
5. Что такое матрицы "объект-свойство", какова их структура?

6. Практические задания для овладения приемами работы с программой MVA: выполнить сложение, умножение двух матриц, проверить, соблюдается ли для умножения матриц свойство коммутативности, сгенерировать модельный спектр (в виде гауссова пика или суперпозиции таких пиков) и наложить на него случайный шум, вычесть сигнал фона из матрицы спектров образцов и т.д.

***В дальнейшем все расчетные упражнения на семинарских занятиях выполняются в среде Excel с помощью программы MVA. Домашние задания аспиранты могут выполнять также с помощью программы MVA либо, по их желанию, в среде Excel с помощью надстройки Chemometrics Add-In или в среде Matlab.***

**Тема 2. Название темы** *Линейные преобразования матриц данных*

**Содержание темы** *Линейные преобразования матриц данных: сдвиг (центрирование), масштабирование, поворот, проекция. Их назначение, алгебраическая и геометрическая интерпретация. Пространство главных компонент, матрицы счетов и нагрузок, их свойства, сингулярные числа, полный и эффективный ранг матрицы*

### **Примеры заданий для семинарских занятий**

1. Для матриц спектральных (хроматографических, вольтамперометрических) данных выполнить следующие операции: центрирование, масштабирование на размах или стандартное отклонение, поворот, проекцию, произвольное линейное преобразование. Представить полученные результаты графически.
2. Представить заданную матрицу данных полного ранга в виде произведения матрицы счетов на матрицу нагрузок.
3. Проверить, что все столбцы матрицы счетов ортогональны друг другу.
4. Проверить, что все строки матрицы нагрузок ортогональны друг другу и имеют единичную норму.
5. Вычислить сингулярные числа для этой матрицы.
5. Вычислить сингулярные числа для матрицы данных неполного ранга. На основании полученных значений определить ее ранг.
7. Наложить на эту матрицу случайные шумы различного уровня и снова вычислить ее сингулярные числа. Представить графически зависимость сингулярных чисел матрицы от уровня шума и проанализировать эту зависимость.
8. Представить графически векторы нагрузок и проанализировать характер их изменения с увеличением номера нагрузки

### **Примеры заданий для самостоятельной работы**

1. В чем заключаются линейные преобразования матрицы данных? Как они записываются в матричной форме?

2. Какие основные виды линейных преобразований матриц данных используются в химическом анализе? Для чего они применяются? Какими свойствами обладают матрицы этих преобразований?
3. Что такое сингулярное разложение матрицы, пространство главных компонент, матрица счетов, матрица нагрузок, сингулярные числа, ранг и эффективный ранг матрицы?
4. Какими свойствами обладают матрица счетов и матрица нагрузок?
5. Как можно геометрически интерпретировать пространство главных компонент матрицы данных?
6. Как осуществить переход от исходных координат данных к их координатам в системе главных компонент? Как осуществить обратный переход?
6. По каким критериям можно определить эффективный ранг матрицы?
7. Для заданной или предварительно сгенерированной матрицы спектральных, хроматографических или вольтамперометрических данных выполнить операции центрирования, масштабирования, поворота, проекции. Представить полученные результаты графически.
8. Представить эту матрицу в виде произведения матрицы счетов на матрицу нагрузок, проверить свойства полученных матриц. Вычислить сингулярные числа матрицы, определить ее ранг.

## **Раздел 2. Название раздела** *Количественный анализ многокомпонентных систем*

### **Тема 1. Название темы** *Многомерный регрессионный анализ*

**Содержание** *Многомерный регрессионный анализ. Математическая формулировка задач градуировки и расчета концентраций в матричной форме. Метод Фирордта и метод множественной линейной регрессии. Метод наименьших квадратов. Прямая и обратная градуировки.*

### **Примеры заданий для семинарских занятий**

1. Для модельных (сгенерированных) и реальных спектральных данных провести градуировку и анализ многокомпонентной системы по методу Фирордта и множественной линейной регрессии. Сравнить погрешности полученных результатов.
2. Проанализировать зависимость погрешности результатов от выбора аналитических позиций (длин волн) и состава градуировочных образцов.

### **Примеры заданий для самостоятельной работы**

1. В чем заключается метод Фирордта для анализа многокомпонентных систем?
2. В чем сходство и в чем различие метода Фирордта и метода множественной линейной регрессии?
3. В чем сущность метода наименьших квадратов?
5. Что такое прямая и обратная градуировка?
6. Как в матричной форме записываются уравнения, лежащие в основе анализа многокомпонентных систем методом множественной линейной регрессии, при использовании прямой градуировки? Обратной градуировки? Как решить эти уравнения?
7. Каковы ограничения на число аналитических позиций при прямой и обратной градуировке? В чем причина этих ограничений?
8. Какую предварительную обработку данных проводят при анализе многокомпонентных систем методами Фирордта и множественной линейной регрессии? С какой целью?
9. Практические задания по теме семинарского занятия.

### **Тема 2. Название темы *Проекционные методы регрессионного анализа***

**Содержание темы *Проекционные методы регрессионного анализа. Проекция на главные компоненты (PCA), проекция на скрытые структуры (PLS1, PLS2). Их использование в анализе многокомпонентных систем***

### **Примеры заданий для семинарских занятий**

Применение проекционных методов для количественного анализа многокомпонентных систем (на примере модельных и реальных спектральных данных).

### **Примеры заданий для самостоятельной работы**

1. В чем сущность проекционных методов регрессионного анализа?
2. Каковы преимущества проекционных методов регрессионного анализа перед методом множественной линейной регрессии?
3. В чем сходство и различие методов проекции на главные компоненты и на скрытые структуры?
4. В чем сходство и различие в свойствах матриц счетов и нагрузок в этих методах?

5. В чем различие методов PLS1 и PLS2?
6. Какой способ градуировки - прямой или обратный - целесообразнее применять в проекционных методах? Почему?
7. В чем сходство и различие в свойствах матриц счетов и нагрузок при проекции на главные компоненты и на скрытые структуры?
8. Как записать в матричном виде уравнения для градуировки и анализа многокомпонентных систем при использовании методов PCA, PLS1, PLS2? Как решить эти уравнения?
9. Какую предварительную обработку данных проводят при анализе многокомпонентных систем проекционными методами? С какой целью?
10. Как выбрать число счетов и нагрузок при анализе многокомпонентных систем проекционными методами?
11. Практические задания по теме семинарского занятия.

### **Тема 3. Название темы *Метрологические аспекты многокомпонентного анализа***

**Содержание темы *Метрологические аспекты многокомпонентного анализа. Описательная и предсказательная способность регрессионной модели. Метод "введено-найдено" в многокомпонентном анализе. Градуировочная и проверочная выборки. Перекрестная проверка на достоверность. Среднеквадратичная погрешность. Оптимизация регрессионной модели***

### **Примеры заданий для семинарских занятий**

На примере модельных (сгенерированных) и реальных спектральных данных - оценка среднеквадратичной погрешности результатов многокомпонентного анализа в зависимости от:

- 1) уровня шума исходных данных;
- 2) степени сходства спектров индивидуальных компонентов;
- 3) способа предварительной обработки данных;
- 4) выбора метода регрессионного анализа;
- 5) для проекционных методов - выбора числа счетов и нагрузок.

### **Примеры заданий для самостоятельной работы**

1. Что такое описательная и предсказательная способность регрессионной модели? Как их можно охарактеризовать количественно?
2. Как влияют на описательную и предсказательную способность модели число ее параметров?

3. Что такое градуировочная и проверочная выборки?
4. В чем заключается метод перекрестной проверки на достоверность?
5. Из каких основных этапов состоит процесс оптимизации регрессионной модели для анализа многокомпонентных систем?
6. Практические задания по теме семинарского занятия.

### **Раздел 3. Название раздела** *Методы многомерной классификации и идентификации в химическом анализе*

#### **Тема 1. Название темы** *Основные принципы многомерной классификации*

**Содержание** *Основные принципы многомерной классификации. Классификационные признаки, принципы их выбора. Расстояние от неизвестного объекта до известного класса как главный классификационный критерий. Способы предварительного преобразования пространства признаков и способы вычисления расстояний в нем. Расстояние "метрическое" (Минковского) и "статистическое" (Махаланобиса). Кластерный анализ, дендрограммы.*

#### **Примеры заданий для семинарских занятий**

Кластеризация химических объектов. Зависимость результатов кластеризации от выбора классификационных признаков, способа предварительного преобразования данных, сокращения их размерности, выбора метрики. Графическое представление результатов кластеризации: дендрограммы, графики счетов.

#### **Примеры заданий для самостоятельной работы**

1. В чем состоит основной принцип многомерной классификации?
2. Что такое обучающая и контрольная выборки, каковы требования к ним?
3. Каковы требования к аналитическим признакам, используемым для классификации химических объектов?
4. Какие предварительные преобразования данных и с какой целью применяют в многомерной классификации?
5. Каковы основные способы вычисления расстояний между объектами в пространстве аналитических признаков?
6. В чем состоит принцип группировки данных в кластеры? Как можно графически представить результаты такой группировки?
7. Практические задания по материалам семинарского занятия.

## **Тема 2. Название темы *Основные методы многомерной классификации***

**Содержание *Основные методы многомерной классификации: линейный дискриминантный анализ, SIMCA, PLS-классификация, метод К ближайших соседей. Оптимизация классификационной модели.***

### **Примеры заданий для семинарских занятий**

Классификация химических объектов различными методами. Оценка качества классификации в зависимости от применяемого метода, его параметров, способа предварительного преобразования данных. Выбор оптимального классификационного правила.

### **Примеры заданий для самостоятельной работы**

1. В чем сходство и различие задач регрессионного анализа и классификации?
2. Какие основные принципы лежат в основе линейного дискриминантного анализа, метода SIMCA, PLS-классификации, метода К ближайших соседей?
3. Охарактеризуйте основные достоинства и недостатки этих методов.
4. Какие из этих методов позволяют оценить относительную дискриминирующую способность отдельных признаков? Каким образом?
5. Практические задания по материалам семинарского занятия.

## **Раздел 4. Название раздела *Методы многомерного разрешения кривых***

### **Тема 1. Название темы *Методы многомерного разрешения кривых***

**Содержание *Постановка задачи разрешения кривых. Регрессионный анализ с ограничениями. Метод чередующихся наименьших квадратов (ALS), итерационного целевого (ITTFA), эволюционного (EFA) и оконного (WFA) факторного анализа***

### **Примеры заданий для семинарских занятий**

Решение задач на разрешение спектральных, хроматографических, кинетических данных с использованием различных ограничений.

### **Примеры заданий для самостоятельной работы**

1. В чем состоит задача разрешения кривых?

2. Что такое спектральный и концентрационный профиль?
3. В чем заключается неопределенность задачи разрешения кривых? Каким способом можно устранить эту неопределенность?
4. Какие основные виды ограничений используют при решении задачи разрешения кривых? Как эти ограничения формулируются математически?
5. Охарактеризуйте суть методов чередующихся наименьших квадратов, итерационного целевого, эволюционного и оконного факторного анализа.
6. Практические задания по теме семинарского занятия.

#### **В. Перечень вопросов к зачету:**

##### **1. Многомерные данные в химическом анализе и основные операции с ними**

Предмет и задачи хемометрики. Полезная информация и шум. Структура данных. Математические модели данных.

Аналитический сигнал, аналитический признак, аналитическая позиция. Представление аналитических данных в векторной и матричной форме.

Линейные преобразования матриц данных: сдвиг (центрирование), масштабирование, поворот, проекция. Их назначение, алгебраическая запись, геометрическая интерпретация.

Пространство главных компонент. Матрицы счетов и нагрузок, их свойства. Сингулярные числа, формальный и эффективный ранг матрицы. Критерии определения эффективного ранга матрицы.

##### **2. Количественный анализ многокомпонентных систем**

Многомерный регрессионный анализ. Формулировка задач градуировки и расчета концентраций в матричной форме. Метод Фирордта и метод множественной линейной регрессии. Метод наименьших квадратов. Прямая и обратная градуировка.

Проекционные методы регрессионного анализа. Проекция на главные компоненты, проекция на скрытые структуры, их использование в регрессионном анализе. Анализ многокомпонентных систем с использованием проекционных методов регрессионного анализа.

Описательная и предсказательная способность регрессионной модели. Градуировочная и проверочная выборки. Перекрестная проверка на достоверность. Среднеквадратичная погрешность. Оптимизация регрессионной модели.

##### **3. Многомерная классификация в химическом анализе**

Постановка задач классификации и идентификации. Основные принципы многомерной классификации. Классификационные признаки, принципы их выбора. Применение корреляционного анализа для выбора классификационных признаков. Пространство классификационных признаков.

Расстояние от неизвестного объекта до известного класса как главный классификационный критерий. Способы предварительного преобразования пространства признаков и способы вычисления расстояний в нем. Расстояния "метрическое" (Минковского) и "статистическое" (Махаланобиса). Критерий Стьюдента как простейший пример использования расстояния Махаланобиса. Графическое представление (визуализация) результатов классификации.

Группировка объектов в классы. Кластерный анализ. Дендрограммы и способы их построения.

Основные методы многомерной классификации. Линейный дискриминантный анализ, SIMCA, PLS-классификация, метод К ближайших соседей: способы вычисления расстояний, классификационные критерии.

Оптимизация классификационной модели. Обучающая и контрольная выборки.

Применение многомерной классификации для контроля качества образцов, установления их чистоты, источника происхождения.

#### 4. Многомерное разрешение кривых

Постановка задачи разрешения кривых. "Спектральный" и "концентрационный" профили.

Неоднозначность решения задачи разрешения кривых и способы ее преодоления. Регрессионный анализ с ограничениями. Основные виды ограничений, используемые в задачах разрешения кривых: неотрицательность, унимодальность, монотонность. Их математическая формулировка.

Методы чередующихся наименьших квадратов, итерационный целевой, эволюционный и оконный факторный анализ: алгоритмы, основные формулы.

Применение методов разрешения кривых в спектроскопии, хроматографии, термодинамических и кинетических исследованиях.

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

Зачет проводится по билетам. В ходе сдачи зачета проверяется, в первую очередь, формирование «знаниевой» компоненты компетенций, перечисленных в п.5, а также сформированность перечисленных в п.5 умений. Уровень знаний аспиранта по каждому вопросу оценивается на «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». В случае, если на все вопросы был дан ответ, оцененный не ниже чем «удовлетворительно», аспирант получает общую оценку «зачтено». Ведомость приема зачета подписывается членами комиссии.

Шкала оценивания знаний, умений и навыков является единой для всех дисциплин (приведена в таблице ниже)

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТА ОБУЧЕНИЯ по дисциплине (модулю)				
Оценка	2	3	4	5
Результат				

Знания	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания
Умения	Отсутствие умений	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности непринципиального характера)	Успешное и систематическое умение
Навыки (владения)	Отсутствие навыков	Наличие отдельных навыков	В целом, сформированные навыки, но не в активной форме	Сформированные навыки, применяемые при решении задач