

Белорусский государственный университет

УТВЕРЖДАЮ

Проректор по учебной работе  
и образовательным инновациям

Чуприс О. И.

«17» 05 2018 г.

Регистрационный № УД- 5603 /уч.

## **ЧИСЛЕННЫЕ МЕТОДЫ И ПРОГРАММНЫЕ СРЕДСТВА КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ И АНАЛИЗА**

**Учебная программа учреждения высшего образования  
по учебной дисциплине для специальности**

1-31 81 12

Прикладной компьютерный анализ данных

2018 г.

Учебная программа составлена на основе образовательного стандарта высшего образования ОСВО 1-31 81 12-2015; учебного плана G31-251/уч., утв. 26.05.2017.

## **СОСТАВИТЕЛЬ**

Б. В. Фалейчик, доцент кафедры вычислительной математики Белорусского государственного университета, кандидат физико-математических наук, доцент.

## **РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:**

Кафедрой вычислительной математики БГУ (протокол № 14 от 19.04.2018);  
Научно-методическим Советом БГУ (протокол № 5 от 04.05.2018)

## ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

**Цель** учебной дисциплины «Численные методы и программные средства компьютерного моделирования и анализа» – обучение студентов базовым численным методам и алгоритмам интеллектуального анализа данных, а также способам их реализации на языке Wolfram в системе технических вычислений Mathematica.

В рамках поставленной цели **задачи** учебной дисциплины состоят в следующем:

- изучение основ программирования на языке Wolfram;
- ознакомление с основными методами численными методами, используемыми в интеллектуальном анализе данных;
- изучение прикладных задач интеллектуального анализа данных;
- формирование у студентов навыков программной реализации численных методов на языке Wolfram.

Рассматриваемая учебная дисциплина относится к циклу специальных дисциплин (компонент учреждения высшего образования).

Основой для изучения этой дисциплины являются курсы первой степени, «Геометрия и алгебра», «Программирование», а также курс «Математическое моделирование и оптимизация сложных систем» второй степени высшего образования.

Учебный материал, излагаемый в учебной дисциплине, связан с учебной дисциплиной «Методы визуализации в анализе больших данных» из компонента учреждения высшего образования (4 семестр).

В результате изучения дисциплины студент магистратуры должен:

### **знать**

- синтаксис, основные принципы и функции языка Wolfram
- математическую постановку базовых задач интеллектуального анализа данных
- численные методы линейной алгебры, применяющиеся в интеллектуальном анализе данных

### **уметь**

- реализовывать вычислительные алгоритмы на языке Wolfram;
- решать задачи интеллектуального анализа данных численными методами
- анализировать результаты, полученные с помощью численных методов

### **владеть**

- основными терминами и понятиями интеллектуального анализа данных
- навыками построения и реализации численных методов.

Освоение рассматриваемой учебной дисциплины должно обеспечить формирование следующих академических, социально-личностных и профессиональных компетенций:

**академические компетенции:**

АК-1. Способность к самостоятельной профессиональной деятельности (анализ, сопоставление, систематизация, абстрагирование, моделирование, проверка достоверности данных, принятие решений и др.), готовность генерировать и использовать новые идеи.

**социально-личностные компетенции:**

СЛК-1. Учитывать социальные и нравственно-этические нормы в социально-профессиональной деятельности;

СЛК-7. Проявлять инициативу и креативность, в том числе в нестандартных ситуациях;

**профессиональные компетенции:**

ПК-3. Самостоятельно разрабатывать эффективные численные методы и алгоритмы, а также интегрировать их в компьютерные системы анализа данных;

ПК-4. Обосновывать выбор методов и инструментов для решения прикладных задач.

В соответствии с учебным планом специальности 1-31 81 12 «Прикладной компьютерный анализ данных» учебная программа предусматривает для изучения дисциплины всего 180 часов, из них 64 аудиторных часа, в том числе лекций – 34 часа, практических занятий – 34 часа. Трудоемкость учебной дисциплины составляет 5 зачетных единиц.

Форма обучения – дневная.

Форма текущей аттестации – экзамен в 3 семестре.

# СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

## Раздел 1. Основы языка Wolfram

**Тема 1.1. Базовые принципы языка Wolfram.** Выражения: атомы и нормальные выражения. Полная форма выражений. Древоподобное представление выражений. Заголовок выражения, уровни выражения. Доступ к частям выражений с помощью функций Part, Level, Extract. Шаблоны и правила подстановки. Глобальная база правил. Принцип вычисления выражений.

**Тема 1.2. Элементарные операции языка Wolfram.** Именованное символом, получение информации о символах. Значения символов: OwnValues, DownValues, SubValues. Мгновенные и отложенные присваивания: Set и SetDelayed. Функции проверки выражений на равенство: Equal, SameQ, TrueQ. Условные выражения: If, Which, Switch. Циклы: For, While, Do. Недостатки процедурного подхода.

**Тема 1.3. Работа со списками.** Принцип эффективной работы со списками. Генерация списков: функции Table, Array, Range. Доступ к частям списка: функции Take, Drop, Span, Length, Part, ReplacePart, Position. Добавление элементов: AppendTo, PrependTo, Insert, Delete. Работа с вложенными списками: функции Partition, Transpose, Flatten, Split. Операции с несколькими списками: функции Join, Intersection, Union, Complement. Ассоциативные списки.

**Тема 1.4. Правила и шаблоны.** Принципы работы с шаблонами. Функции Rule, RuleDelayed, Replace, ReplaceAll. Простейшие шаблоны, шаблоны с условиями, шаблоны с необязательными и повторяющимися аргументами. Функции Cases, DeleteCases, MemberQ, Position, Count, FreeQ.

**Тема 1.5. Функции и средства локализации переменных.** Принципы разработки функций. Функции как правила. Префиксная и постфиксная форма записи выражений. Рекурсивные функции. Наложение ограничений на аргументы функций. Локализация переменных с помощью функций Block, Module, With, основные отличия между ними и практики использования. Атрибуты функций: Listable, Orderless, Flat, Protected, Hold\*. Анонимные функции (Pure functions). Функции с аргументами по умолчанию. Функции с опциями.

**Тема 1.6. Функциональное программирование.** Принципы функционального программирования. Функции Map, MapAt, MapAll, MapIndexed, Scan и примеры их использования. Спецификация уровня применения функции. Функции Thread, MapThread, Inner, Outer. Функции Nest, NestWhile, Fold и примеры их использования. Функции FixedPoint и FixedPointList.

**Тема 1.7. Средства повышения производительности.** Компиляция функций: понятие скомпилированной функции, ограничения на компилируемый код. Функция Compile: синтаксис, опции. Виды компиляции: “WVM” и

“С”. Параллельные вычисления: принципы организации. Функции Parallelize, ParallelTable, ParallelMap.

**Тема 1.8. Средства разработки интерактивных приложений.** Функция Dynamic и принципы динамического отображения выражений. Быстрая разработка интерактивных форм с помощью функции Manipulate. Основные динамические объекты управления: функции Button, Slider, Setter, ClickPane, LocatorPane, Функция DynamicModule и ее отличия от Module.

## **Раздел 2. Численные методы интеллектуального анализа данных**

**Тема 2.1. Линейная задача наименьших квадратов и численные методы ее решения.** Постановка линейной задачи наименьших квадратов, ее разрешимость. Метод нормальных уравнений. Решение методом QR-разложения: преобразования отражений и вращений, алгоритм построения QR-разложения. Случай матрицы неполного ранга. Приближенное решение задачи наименьших квадратов

**Тема 2.2. Методы подпространств Крылова.** Подпространства Крылова и их свойства. Форма Хессенберга. Алгоритм ортогонализации Арнольди. Случай симметричной матрицы: алгоритм Ланцоша. Крыловские методы решения СЛАУ: обобщенный метод минимальных невязок, метод полной ортогонализации, метод сопряженных градиентов. Крыловские методы решения проблемы собственных значений для больших матриц.

**Тема 2.3. Неотрицательные матричные разложения.** Постановка задачи неотрицательного матричного разложения, корректность. Сведение к линейной задаче условной среднеквадратичной минимизации. Метод попеременных наименьших квадратов. Алгоритм мультипликативного уточнения. Связь с методом градиентного спуска.

**Тема 2.4. Сингулярное разложение матрицы.** Определение сингулярного разложения (СР). Теорема существования, геометрический смысл. Приближение матриц с помощью СР. Решение задачи наименьших квадратов с помощью СР. Численные методы построения СР, bidiagonalization Голуба-Кахана.

## **Раздел 3. Интеллектуальный анализ данных в приложениях**

**Тема 3.1. Классификация рукописных цифр.** Постановка задачи распознавания. Матричное представление данных. Простейший алгоритм на векторных подпространствах. Классификация с помощью сингулярного разложения.

**Тема 3.2. Метод ранжирования ссылок PageRank.** Постановка задачи ранжирования ссылок. Случайные блуждания и марковские цепи, стохастические матрицы и их свойства. Алгоритм PageRank на основе степенного метода.

**Тема 3.3. Методы интеллектуального анализа текстов.** Интеллектуальный анализ текстов: постановка задачи, предварительная обработка данных, модель векторного пространства. Латентное семантическое индексирование, кластеризация, использование неотрицательного матричного разложения. Автоматическое выделение ключевых слов и предложений: оценка заметности, матрица «термин-предложение». Выделение ключевого предложения с помощью аппроксимации низкого ранга.

## УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов		Количество часов УСР	Форма контроля знаний
		Лекции	Практические занятия		
<b>1</b>	<b>Раздел 1. Основы языка Wolfram</b>				Контрольная работа
1.1	Базовые принципы языка Wolfram	2	2		Лабораторная работа №1
1.2	Элементарные операции языка Wolfram	2	2		
1.3	Работа со списками	2	2		
1.4	Правила и шаблоны	2	2		
1.5	Функции и средства локализации переменных	2	2		Лабораторная работа №2
1.6	Функциональное программирование	2	2		
1.7	Средства повышения производительности	2	2		
1.8	Средства разработки интерактивных приложений	2	2		
<b>2</b>	<b>Раздел 2. Численные методы интеллектуального анализа данных</b>				Коллоквиум
2.1	Линейная задача наименьших квадратов	2	2		Лабораторная работа №3
2.2	Методы подпространств Крылова	2	2		
2.3	Неотрицательные матричные разложения	2	2		
2.4	Сингулярное разложение матрицы	4	4		
<b>3</b>	<b>Раздел 3. Интеллектуальный анализ данных в приложениях</b>				Коллоквиум
3.1	Классификация рукописных цифр	2	2		Лабораторная работа №4
3.2	Метод ранжирования ссылок PageRank	2	2		
3.3	Методы интеллектуального анализа текстов	4	4		
	<b>Всего</b>	<b>34</b>	<b>34</b>		



## ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

### Рекомендуемая литература

#### Основная

1. Maeder, R. Computer Science with MATHEMATICA. – Cambridge University Press, 2000. – 389 p.
2. Wolfram, S. The MATHEMATICA book, Version 4. – Cambridge University Press. – 1470 p.
3. Eldén, L. Matrix Methods in Data Mining and Pattern Recognition. – SIAM, 2007. – 216 p.
4. Саад Ю. Итерационные методы для разреженных линейных систем. В 2-х томах. Том 1. – М.: Издательство Московского университета, 2013. – 306 с.

#### Дополнительная

1. Деммель, Дж. Вычислительная линейная алгебра: теория и приложения. – М.: Мир, 2001. – 432 с.
2. Shifrin, L. Mathematica programming: an advanced introduction : [www.mathprogramming-intro.org](http://www.mathprogramming-intro.org)

## Перечень рекомендуемых средств диагностики

Для текущего контроля качества усвоения знаний студентами рекомендуется использовать следующий диагностический инструментарий:

- лабораторные работы;
- письменные контрольные работы
- коллоквиум;
- устные вопросы.

**Лабораторные работы** как правило представляют собой задания, включающие программную реализацию указанного численного метода (язык программирования обычно выбирается самим студентом), проведение вычислительного эксперимента и комментарии по его итогам. Рекомендуемая форма отчетности по лабораторной работе – письменный отчет. Лабораторная работа оценивается по стандартной 10-балльной шкале. Оценка за лабораторную работу может быть снижена в случае несвоевременного выполнения.

**Письменные контрольные работы** проводятся для контроля знаний по одному или нескольким разделам курса. Они включают, как правило, 4-5 заданий и оцениваются по 10-балльной шкале. В случае неудовлетворительной оценки контрольная работа может быть переписана.

**Коллоквиум** представляет собой персональную устную беседу преподавателя со студентом с целью определения уровня знаний по пройденным темам. Для более точной оценки коллоквиум может включать дополнительный письменный этап. По результатам коллоквиума выставляется оценка по 10-балльной шкале.

**Устный опрос** студентов проводится в свободной форме в течение лабораторных занятий. Его результаты учитываются преподавателем при выставлении рейтинговой оценки в конце семестра.

## Методика формирования итоговой оценки

Итоговая оценка формируется на основе:

1. Правил проведения аттестации студентов (Постановление Министерства образования Республики Беларусь №53 от 29 мая 2012 г.);
2. Положения о рейтинговой системе оценки знаний по дисциплине в БГУ (Приказ ректора БГУ от 18.08.2015)
3. Критериев оценки знаний студентов (письмо Министерства образования от 22.12.2003 г.)

## ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу
Методы визуализации в анализе больших данных	Кафедра вычислительной математики	Нет	Именений не требуется, протокол № 14 от 19.04.2018

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО  
ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на \_\_\_\_ / \_\_\_\_ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры  
\_\_\_\_\_ (протокол № \_\_\_\_ от \_\_\_\_\_ 201\_ г.)

Заведующий кафедрой

\_\_\_\_\_

УТВЕРЖДАЮ  
Декан факультета

\_\_\_\_\_

\_\_\_\_\_