

# Программа для поступающих в Школу анализа данных

## Алгебра

1. Подстановки. Определение подстановки, чётность подстановок. Произведение подстановок, разложение подстановок в произведение транспозиций и независимых циклов.
2. Комплексные числа. Геометрическое изображение, алгебраическая и тригонометрическая форма записи, извлечение корней, корни из единицы.
3. Системы линейных уравнений. Прямоугольные матрицы. Приведение матриц и систем линейных уравнений к ступенчатому виду. Метод Гаусса.
4. Линейная зависимость и ранг. Линейная зависимость строк / столбцов. Основная лемма о линейной зависимости, базис и ранг системы строк / столбцов. Ранг матрицы. Критерий совместности и определённости системы линейных уравнений в терминах рангов матриц. Фундаментальная система решений однородной системы линейных уравнений.
5. Определители. Определитель квадратной матрицы, его основные свойства. Критерий равенства определителя нулю. Формула разложения определителя матрицы по строке / столбцу.
6. Операции над матрицами и их свойства. Теорема о ранге произведения двух матриц. Определитель произведения квадратных матриц. Обратная матрица, ее явный вид (формула), способ выражения с помощью элементарных преобразований строк.
7. Векторные пространства. Базис. Векторное пространство, его базис и размерность. Преобразования координат в векторном пространстве. Подпространства как множества решений систем однородных линейных уравнений. Связь между размерностями суммы и пересечения двух подпространств. Линейная независимость подпространств. Базис и размерность прямой суммы подпространств.
8. Линейные отображения и линейные операторы. Линейные отображения, их запись в координатах. Образ и ядро линейного отображения, связь между их размерностями. Сопряжённое пространство и сопряжённые базисы. Изменение матрицы линейного оператора при переходе к другому базису.
9. Билинейные и квадратичные функции. Билинейные функции, их запись в координатах. Изменение матрицы билинейной функции при переходе к другому базису. Ортогональное дополнение к подпространству относительно симметрической билинейной функции. Связь между симметрическими билинейными и квадратичными функциями. Существование ортогонального базиса для симметрической билинейной функции. Нормальный вид вещественной квадратичной функции. Закон инерции.
10. Евклидовы пространства. Неравенство Коши-Буняковского. Ортогональные базисы. Ортогонализация Грама-Шмидта. Ортогональные операторы.
11. Собственные векторы и собственные значения линейного оператора. Собственные подпространства линейного оператора, их линейная независимость. Условие диагонализуемости оператора.

Источники:

- [1] Кострикин А.И. Введение в алгебру: МНЦНО, 2020.
- [2] Курош А.Г. Курс высшей алгебры: Лань, 2007.
- [3] Винберг Э.Б. Курс алгебры: МЦНМО, 2019.
- [4] Сборник задач по алгебре под редакцией Кострикина А.И / И. В. Аржанцев,

В. А. Артамонов, Ю. А. Бахтурин и др.: МЦНМО, 2009.

[5] Сборник задач по аналитической геометрии и линейной алгебре под редакцией Смирнова Ю.М.: МЦНМО, 2016.

[6] Смирнов С. В., Гайфуллин А. А., Пенской А. В. Задачи по линейной алгебре и геометрии: МЦНМО, 2014.

[7] Прасолов В.В. Задачи и теоремы линейной алгебры: МЦНМО, 2015.

[8] Городенцев А.Л. Алгебра. Учебник для студентов-математиков. Часть 1: МЦНМО, 2013.

[9] Городенцев А.Л. Линейная алгебра и геометрия.

[http://gorod.bogomolov-lab.ru/ps/stud/geom\\_ru/2122/list.html](http://gorod.bogomolov-lab.ru/ps/stud/geom_ru/2122/list.html)

[10] Панов Т. Е. Линейная алгебра и геометрия.

<http://higeom.math.msu.su/people/taras/teaching/panov-linalg.pdf>

## Математический анализ

1. Пределы последовательностей и их свойства. Теорема о промежуточной последовательности. Теорема Вейерштрасса о монотонной ограниченной последовательности.
2. Пределы функций в точке и на бесконечности, их свойства. Теорема о промежуточной функции. Критерий Коши существования конечного предела функции. Существование односторонних пределов у монотонной функции. Первый и второй замечательные пределы.
3. Непрерывность функции в точке. Односторонняя непрерывность. Свойства функций, непрерывных на отрезке: ограниченность, достижение минимального и максимального значений. Теорема Коши о промежуточном значении непрерывной функции.
4. О-символика и асимптотические оценки.
5. Производная функции одной переменной. Односторонние производные. Непрерывность функции, имеющей производную. Дифференцируемость функции в точке. Механический и геометрический смысл производной и дифференциала. Свойства производных. Производные элементарных функций. Производные высших порядков.
6. Теоремы Ролля, Лагранжа, Коши. Поиск локальных экстремумов, определение выпуклости и точек перегиба, исследование функций с помощью производной. Формула Тейлора. Правило Лопиталя.
7. Функции многих переменных, их непрерывность и дифференцируемость. Частные производные. Градиент и его геометрический смысл. Производная по направлению. Гессиан. Метод градиентного спуска. Поиск экстремумов функций от многих переменных. Поиск условных экстремумов функций нескольких переменных, метод множителей Лагранжа. Теорема о неявной функции.
8. Интегрирование. Определённый и неопределённый интегралы, их связь. Методы интегрирования функций. Первообразные различных элементарных функций. Кратные интегралы (двойные, тройные), замена координат, связь с повторными.
9. Элементы функционального анализа: нормированные, метрические пространства, непрерывность, ограниченность.
10. Ряды. Числовые и функциональные ряды. Признаки сходимости (Даламбера, Коши, интегральный, Лейбница). Абсолютно и условно сходящиеся ряды. Признаки Абеля и Дирихле. Сходимость степенных рядов. Круг и радиус сходимости. Формула Коши-Адамара для радиуса сходимости.

Источники:

[1] Зорич В. А. Математический анализ: МЦНМО, 2021.

[2] Кудрявцев, Л.Д. Курс математического анализа. Учебник. В 3 томах: Юрайт, 2017.

- [3] Демидович, Б. П., Сборник задач и упражнений по математическому анализу: Лань, 2022.
- [4] W. J. Kaczor, M. T. Nowak Problems in Mathematical Analysis: AMS, 2000.

## Комбинаторика

1. Основные правила комбинаторики. Правило подсчёта количества комбинаторных объектов. Принцип Дирихле. Примеры.
2. Множества. Круги Эйлера, операции на множествах. Формула включений и исключений. Примеры.
3. Сочетания. Размещения, перестановки и сочетания. Бином Ньютона. Треугольник Паскаля. Размещения, перестановки и сочетания с повторениями.
4. Графы. Лемма о рукопожатиях. Связность графов. Деревья и их свойства. Эйлеровы и гамильтоновы графы. Планарные графы, формула Эйлера. Ориентированные графы, турниры.

Источники:

- [1] Виленкин Н.Я. Комбинаторика: Наука, 1969.
- [2] С.А. Генкин, И.В. Итенберг, Д.В. Фомин: Ленинградские математические кружки, 1994.
- [3] А. В. Омельченко. Теория графов: МЦНМО, 2018.
- [4] Грэхем Р., Кнут Д., Паташник О. Конкретная математика: МЦНМО, 1998.
- [5] Алон Н., Спенсер Дж. Вероятностный метод: Лаборатория знаний, 2015.

## Теория вероятностей

1. Основные понятия теории вероятностей. Определение вероятностного пространства, простейшие дискретные случаи (выборки с порядком и без него, упорядоченные и неупорядоченные), классическая вероятностная модель.
2. Условные вероятности. Определение условной вероятности, формула полной вероятности, формула Байеса. Независимость событий на вероятностном пространстве. Парная независимость и независимость в совокупности.
3. Случайные величины как измеримые функции. Функция распределения. Функция плотности. Независимость случайных величин. Случайные векторы.
4. Математическое ожидание в дискретном и абсолютно непрерывном случае, дисперсия, ковариация и корреляция. Их основные свойства. Дисперсия суммы независимых случайных величин. Математическое ожидание и матрица ковариаций случайного вектора. Симметричность и неотрицательная определенность матрицы ковариаций. Математическое ожидание случайной величины в общем виде.
5. Неравенство Маркова. Неравенство Чебышёва. Виды сходимости случайных величин: с вероятностью 1 (почти наверное), по вероятности, в среднем порядка  $p > 0$ , по распределению. Закон больших чисел. Центральная предельная теорема.

6. Распределения. Стандартные дискретные и непрерывные распределения, их математические ожидания, дисперсии и свойства:
- биномиальное;
  - равномерное;
  - нормальное и многомерное нормальное;
  - пуассоновское;
  - показательное;
  - геометрическое.

Источники:

- [1] Гнеденко, Б. В. Курс теории вероятностей: УРСС, 2001.
- [2] Гнеденко Б. В., Хинчин А. Я. Элементарное введение в теорию вероятностей, 1970.
- [3] Ширяев, А. Н. Вероятность: Наука, 1989.
- [4] Севастьянов Б. А., Курс теории вероятностей и математической статистики: Наука, 1982.
- [5] Севастьянов, Б. А., Чистяков, В. П., Зубков, А. М. Сборник задач по теории вероятностей: Наука, 1986.
- [6] Шабанов Д. А. Теория вероятностей.  
[http://wiki.cs.hse.ru/Теория\\_вероятностей\\_2018/2019\\_\(пилотный\\_поток\)](http://wiki.cs.hse.ru/Теория_вероятностей_2018/2019_(пилотный_поток))

**Программирование, алгоритмы и структуры данных**  
 (предполагается владение одним из основных языков программирования,  
 предпочтительным является C/C++)

1. Простейшие конструкции языка программирования. Циклы, ветвления, рекурсия.
2. Анализ алгоритмов. Понятие о сложности по времени и по памяти. Асимптотика, О-символика. Инварианты, пред- и пост- условия. Доказательство корректности алгоритмов.
3. Простейшие структуры данных. Массивы, стеки, очереди, связные списки. Сравнение временных затрат при различных типах операций.
4. Строки и операции над ними. Представление строк. Вычисление длины, конкатенация.
5. Простейшие конструкции языка программирования. Циклы, ветвления, рекурсия. Указатели и динамическое управление памятью.
6. Сортировки. Нижняя теоретико-информационная оценка сложности задачи сортировки. Алгоритмы сортировки вставками, пузырьком, быстрая сортировка, сортировка слиянием. Оценка сложности.
7. Графы — неориентированные и ориентированные — и их представление в памяти (списками смежности и матрицей инцидентности). Деревья. Бинарные деревья. Обход графа в ширину. Обход графа в глубину.

**Задачи на следующие темы могут быть заданы на второй части второго этапа отбора, а также на письменном экзамене нового трека или на собеседованиях**

8. Время и память как основные ресурсы. Модели вычислений: RAM, разрешающие деревья. Нижняя оценка на число сравнений при сортировке в модели разрешающих деревьев.
9. Динамическое программирование: общие принципы, свойства задач, эффективно решаемых при помощи динамического программирования. Рекурсивная реализация с мемоизацией и итеративная реализация. Примеры: вычисление редакционного расстояния, решение задачи о рюкзаке.

10. Жадные алгоритмы: общая идея, критерии применимости. Примеры: построение остовов минимального веса в графах, жадные алгоритмы в задачах планирования.
11. Порядковые статистики. Рандомизированный алгоритм Quick-Select. Детерминированный алгоритм поиска (метод «медианы медиан»).
12. Кучи: основные определения и свойства. Операции Sift-Down и Sift-Up. Бинарные и k-ичные кучи. Построение кучи за линейное время. Алгоритм сортировки Heap-Sort.
13. Хеш-функции. Коллизии. Разрешение коллизий методом цепочек. Гипотеза простого равномерного хеширования, оценка средней длины цепочки.
14. Графы: проверка на ацикличность и топологическая сортировка.
15. Деревья поиска. Вставка и удаление элементов. Inorder-обход дерева.
16. Сильно связанные компоненты и топологическая сортировка конденсации.
17. Дерево отрезков.
18. Кратчайшие пути в графах. Оценки расстояний и их релаксация. Алгоритмы Беллмана–Форда, Флойда и Дейкстры.
19. Остовы минимального веса. Лемма о минимальном ребре в разрезе. Алгоритмы Краскала и Прима.
20. Системы непересекающихся множеств. Реализация с использованием леса.

Источники:

- [1] Шень А. Программирование: теоремы и задачи: МЦМНО, 2007.
- [2] Вирт Н. Алгоритмы и структуры данных: Невский диалект, 2005.
- [3] Керниган Б., Ритчи Д. Язык программирования C: Вильямс, 2008.
- [4] Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ: Вильямс, 2005.
- [5] [http://en.wikipedia.org/wiki/Code\\_Complete](http://en.wikipedia.org/wiki/Code_Complete)
- [6] [http://en.wikipedia.org/wiki/Design\\_Patterns](http://en.wikipedia.org/wiki/Design_Patterns)
- [7] <https://www.amazon.es/Effective-Specific-Programs-Professional-Computing/dp/0321334876>

## Анализ данных

1. Основные задачи машинного обучения: классификация, регрессия, ранжирование, кластеризация. Обучение с учителем и без учителя.
2. Предобработка и очистка данных. Работа с пропущенными значениями.
3. Feature Engineering. Работа с категориальными признаками.
4. Переобучение: как его обнаружить и как с ним бороться. Разделение на обучающую и тестовую выборки. Методы регуляризации.
5. Сравнение моделей. Метрики в задачах классификации и регрессии. Методология подбора гиперпараметров.
6. Основные модели классификации и регрессии: линейные модели, решающие деревья. Ансамбли алгоритмов.

Источники:

- [1] Учебник ШАД.  
<https://ml-handbook.ru/>
- [2] Конспекты Воронцова.  
[http://www.machinelearning.ru/wiki/index.php?title=Машинное обучение](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение)  
(курс лекций, К.В.Воронцов)

[3] Лекции Жени Соколова.

<https://github.com/esokolov/ml-course-hse>

[4] Курс ML на Физтехе.

<https://ml-mipt.github.io/>