

Программа для поступающих в Школу анализа данных

Алгебра

- Подстановки. Определение подстановки, чётность подстановок. Произведение подстановок, разложение подстановок в произведение транспозиций и независимых циклов. Симметрические группы.
- Комплексные числа. Геометрическое изображение, алгебраическая и тригонометрическая форма записи, извлечение корней, корни из единицы.
- Системы линейных уравнений. Прямоугольные матрицы. Приведение матриц и систем линейных уравнений к ступенчатому виду. Метод Гаусса. Численные методы решения СЛАУ (метод Гаусса-Зейделя, метод Якоби).
- 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.
- Определители. Определитель квадратной матрицы, его основные свойства. Критерий равенства определителя нулю. Формула разложения определителя матрицы по строке / столбцу.
- 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.
- 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.
- 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.
- 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.
- Евклидовы пространства. Неравенство Коши-Буняковского. Ортогональные базисы. Ортогонализация Грама-Шмидта. Ортогональные операторы.
- Собственные векторы и собственные значения линейного оператора. Собственные подпространства линейного оператора, их линейная независимость. Условие диагонализируемости оператора.

Источники:

Кострикин А.И. Введение в алгебру: МНЦНО, 2024.

Курош А.Г. Курс высшей алгебры: Лань, 2007.

Винберг Э.Б. Курс алгебры: МЦНМО, 2019.

Сборник задач по алгебре под редакцией Кострикина А.И.: МЦНМО, 2023.

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

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

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

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

Дмитрий Трушин Линейная алгебра и геометрия.

Роман Авдеев Линейная алгебра и геометрия.

Городенцев А.Л. Линейная алгебра и геометрия. http://gorod.bogomolov-lab.ru/ps/stud/geom_ru/2122/list.html

Панов Т. Е. Линейная алгебра и геометрия. <http://higeom.math.msu.su/people/taras/teaching/panov-linalg.pdf>

Математический анализ

- Пределы последовательностей и их свойства. Теорема о промежуточной последовательности. Теорема Вейерштрасса о монотонной ограниченной последовательности.
- 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.
- Непрерывность функции в точке. Односторонняя непрерывность. Свойства функций, непрерывных на отрезке: ограниченность, достижение минимального и максимального значений. Теорема Коши о промежуточном значении непрерывной функции.
- O -символика и асимптотические оценки.
- 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.
- Теоремы Ролля, Лагранжа, Коши. Поиск локальных экстремумов, определение выпуклости и точек перегиба, исследование функций с помощью производной. Формула Тейлора. Правило Лопитала.
- 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.
- Интегрирование. Определённый и неопределённый интегралы, их связь. Методы интегрирования функций.

Первообразные различных элементарных функций. Кратные интегралы (двойные, тройные), замена координат, связь с повторными. Криволинейные и поверхностные интегралы.

- Элементы функционального анализа: нормированные, метрические пространства, непрерывность, ограниченность.
- 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. Степенные ряды и ряды Фурье, интегралы Фурье, преобразование Фурье.

Источники:

Зорич В. А. Математический анализ: МЦНМО, 2021.

Демидович Б. П. Сборник задач и упражнений по математическому анализу: Лань, 2022.

W. J. Kaczor, M. T. Nowak Problems in Mathematical Analysis: AMS, 2000.

Комбинаторика

- Основные правила комбинаторики. Правило подсчёта количества комбинаторных объектов. Принцип Дирихле. Примеры.
- Множества. Круги Эйлера, операции на множествах. Формула включений и исключений. Примеры.
- Размещения, перестановки и сочетания. Бином Ньютона. Треугольник Паскаля. Размещения, перестановки и сочетания с повторениями. Гипергеометрическая схема. Примеры применения комбинаторики в физике (статистика Максвелла-Больцмана, статистика Бозе-Эйнштейна, статистика Ферми-Дирака).
- Графы. Лемма о рукопожатиях. Связность графов. Деревья и их свойства. Эйлеровы и гамильтоновы графы. Планарные графы, формула Эйлера. Ориентированные графы, турниры.

Источники:

Н. Я. Виленкин, А. Н. Виленкин, П. А. Виленкин Комбинаторика: МЦНМО, 2019.

С.А. Генкин, И.В. Итенберг, Д.В. Фомин: Ленинградские математические кружки, 1994.

А. В. Омельченко. Теория графов: МЦНМО, 2018.

Грэхем Р., Кнут Д., Паташник О. Конкретная математика: МЦНМО, 1998.

Алон Н., Спенсер Дж. Вероятностный метод: Лаборатория знаний, 2015.

Теория вероятностей

- Основные понятия теории вероятностей. Определение вероятностного пространства, простейшие дискретные случаи (выборки с порядком и без него, упорядоченные и неупорядоченные), классическая вероятностная модель. Аксиоматическое определение вероятности Колмогорова. Геометрическая вероятность.
- Условные вероятности. Определение условной вероятности, формула полной вероятности, формула Байеса. Независимость событий на вероятностном пространстве. Парная независимость и независимость в совокупности.
- Схема Бернулли. Формула Бернулли. Формула Пуассона. Формулы Муавра-Лапласа и их применение. Полиномиальная схема
- Случайные величины как измеримые функции. Функция распределения. Функция плотности. Независимость случайных величин. Случайные векторы. Функции от случайной величины.
- 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. Условное математическое ожидание. Регрессия. Другие числовые характеристики случайных величин. Моменты старших порядков.
- 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. Усиленный закон больших чисел. Закон повторного логарифма.
- Распределения. Стандартные дискретные и непрерывные распределения, их математические ожидания, дисперсии и свойства: биномиальное; равномерное; нормальное и многомерное нормальное; пуассоновское; показательное; геометрическое; отрицательно-биномиальное; распределение Паскаля; гипергеометрическое; распределение Гнеденко-Вейбулла; гамма-распределение.

Источники:

Гнеденко, Б. В. Курс теории вероятностей: URSS, 2022.

Гнеденко Б. В., Хинчин А. Я. Элементарное введение в теорию вероятностей: URSS, 2022.

Ширяев, А. Н. Вероятность: МЦНМО, 2021.

Севастьянов Б. А., Курс теории вероятностей и математической статистики: URSS, 2022.

Севастьянов, Б. А., Чистяков, В. П., Зубков, А. М. Сборник задач по теории вероятностей: Наука, 1986.

Шабанов Д. А. Теория вероятностей.

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

- Простейшие конструкции языка программирования. Циклы, ветвления, рекурсия.
- Анализ алгоритмов. Понятие о сложности по времени и по памяти. Асимптотика, O- символика. Инварианты, пред- и пост-условия. Доказательство корректности алгоритмов.
- Простейшие структуры данных. Массивы, стеки, очереди, связанные списки. Сравнение временных затрат при различных типах операций.
- Строки и операции над ними. Представление строк. Вычисление длины, конкатенация.

- Простейшие конструкции языка программирования. Циклы, ветвления, рекурсия. Указатели и динамическое управление памятью.
- Сортировки. Нижняя теоретико-информационная оценка сложности задачи сортировки. Алгоритмы сортировки вставками, пузырьком, быстрая сортировка, сортировка слиянием. Оценка сложности.
- 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

- Время и память как основные ресурсы. Модели вычислений: RAM, разрешающие деревья. Нижняя оценка на число сравнений при сортировке в модели разрешающих деревьев.
- 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.
- Жадные алгоритмы: общая идея, критерии применимости. Примеры: построение остовов минимального веса в графах, жадные алгоритмы в задачах планирования.
- Порядковые статистики. Рандомизированный алгоритм Quick-Select. Детерминированный алгоритм поиска (метод «медианы медиан»).
- 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.
- Хеш-функции. Коллизии. Разрешение коллизий методом цепочек. Гипотеза простого равномерного хеширования, оценка средней длины цепочки.
- Графы: проверка на ацикличность и топологическая сортировка.
- Деревья поиска. Вставка и удаление элементов. Inorder-обход дерева.
- Сильно связные компоненты и топологическая сортировка конденсации.
- Дерево отрезков.
- Кратчайшие пути в графах. Оценки расстояний и их релаксация. Алгоритмы Беллмана–Форда, Флойда и Дейкстры.
- Остовы минимального веса. Лемма о минимальном ребре в разрезе. Алгоритмы Краскала и Прима.
- Системы непересекающихся множеств. Реализация с использованием леса.

Источники:

Шень А. Программирование: теоремы и задачи: МЦМНО, 2021.

Клейнберг Дж., Тардос Е. Алгоритмы: разработка и применение: Питер, 2016.

Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ: Вильямс, 2020.

http://en.wikipedia.org/wiki/Code_Complete

http://en.wikipedia.org/wiki/Design_Patterns

<http://www.amazon.es/Effective-Specific-Programs-Professional-Computing/dp/0321334876>

Анализ данных

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

Источники:

Учебник ШАДа по ML. <https://academy.yandex.ru/handbook/ml>

Конспекты лекций К. В. Воронцова.

Материалы курса по машинному обучению Евгения Соколова: ФКН ВШЭ. <https://github.com/esokolov/ml-course-hse>

Курс по Машинному обучению Радослава Нейчева и Владислава Гончаренко: ФИВТ МФТИ. <https://github.com/girafe-ai/ml-course>