

Программа для поступления в Школу анализа данных

Алгебра

1. Подстановки. Определение подстановки, четность 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 определенности 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. 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. Преобразование 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. Сопряженное пространство и сопряженные 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. Собственные векторы и собственные значения линейного оператора. Собственные подпространства линейного оператора, их линейная независимость. Условие диагонализруемости оператора.

Литература

- [1] Кострикин А.И. Введение в алгебру, 1977, Наука.
- [2] Кострикин А.И. Введение в алгебру, ч. I,II, 2000, Физматлит.
- [3] Курош А.Г. Курсы высшей алгебры, 1975, Наука.

[4] Винберг Э.Б. Курс алгебры, 1999, 2001, Факториал.

[5] Сборник задач по алгебре по редакцией Кострикина А.И / И. В. Аржанцев, В. А. Артамонов, Ю. А. Бахтурин и др. – МЦНМО Москва, 2009.

Математический анализ

1. Пределы и непрерывность. Пределы последовательностей и функций. Непрерывные функции.
2. Ряды. Числовые и функциональные ряды. Признаки сходимости (Даламбера, Коши, интегральный, Лейбница). Абсолютно и условно сходящиеся ряды.
3. Дифференцирование. Дифференцирование функций. Применение производной для нахождения экстремумов функций. Формула Тейлора.
4. Функции многих переменных. Частные производные. Градиент и его геометрический смысл. Гессиан. Метод градиентного спуска. Поиск экстремумов функций от многих переменных.
5. Интегрирование. Определенный и неопределенный интегралы. Методы интегрирования функций. Первообразные различных элементарных функций. Кратные интегралы (двойные, тройные), замена координат, связь с повторными.
6. Элементы функционального анализа: нормированные, метрические пространства, непрерывность, ограниченность.

Литература

[1] Архипов Г. И., Садовничий В. А., Чубариков В. Н. Лекции по мат. анализу. Изд-во Университет, 1999.

[2] Зорич В. А. Математический анализ. Часть I. М.: Наука, 1981. 544 с. Часть II. М.: Наука, 1984. 640 с.

[3] Кудрявцев, Л.Д., Курс математического анализа (в трех томах). Т. 1. Дифференциальное и интегральное исчисления (функций одной переменной). Т. 2. Ряды. Дифференциальное и интегральное исчисления (функций многих переменных). Т. 3. Гармонический анализ. Москва, Изд-во Высшая школа, 1981.

[4] Демидович, Б. П, Сборник задач и упражнений по математическому анализу. Изд-во АСТ, 2007.

Комбинаторика

1. Основные правила комбинаторики. Правило подсчета 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. Сочетания с повторениями.

Литература

- [1] Виленкин Н.Я. Комбинаторика. М., Наука, 1969
 [2] С.А. Генкин, И.В. Итенберг, Д.В. Фомин. Ленинградские математические кружки, 1994

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. Случайная величина, функция распределения.
2. Условные вероятности. Определение условной вероятности, формула полной вероятности, формула Байеса.
3. Математическое ожидание, дисперсия, корреляция. Определение математического ожидания, дисперсии, ковариации и корреляции, их свойства.
4. Независимость событий. Попарная независимость и независимость в совокупности.
5. Основные теоремы теории вероятностей. Неравенство Чебышева. Закон больших чисел. Центральная предельная теорема.
6. Распределения. Стандартные дискретные и непрерывные распределения, их математические ожидания, дисперсии и свойства:
 - биномиальное;
 - равномерное;
 - нормальное;
 - пуассоновское;
 - показательное;
 - геометрическое.

Литература

- [1] Гнеденко, Б. В. Курс теории вероятностей, УРСС. М.: 2001
 [2] Гнеденко Б. В., Хинчин А. Я. Элементарное введение в теорию вероятностей, 1970
 [3] Ширяев, А. Н. Вероятность, Наука. М.: 1989
 [4] Севастьянов Б. А., Курс теории вероятностей и математической статистики, Ч М.: Наука, 1982
 [5] Севастьянов, Б. А., Чистяков, В. П, Зубков, А. М. Сборник задач по теории вероятностей, М.: Наука, 1986

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, связанные списки. Сравнение временных затрат при различных типах операций.
4. Строки и операции над ними. Представление строк. Вычисление длины, конкатенация.

5. Сортировки. Нижняя теоретико-информационная оценка сложности задачи сортировки. Алгоритмы сортировки вставками, пузырьком, быстрая сортировка, сортировка слиянием. Оценка сложности.
6. Указатели. Указатели и динамическое управление памятью.

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

7. Время и память как основные ресурсы. Модели вычислений: RAM, разрешающие деревья. Нижняя оценка на число сравнений при сортировке в модели разрешающих деревьев.
8. 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.
9. Жадные алгоритмы: общая идея, критерии применимости. Примеры: построение остовов минимального веса в графах, жадные алгоритмы в задачах планирования.
10. Порядковые статистики. Рандомизированный алгоритм Quick-Select. Детерминированный алгоритм поиска (метод "медианы медиан").
11. Кучи: основные определения и свойства. Операции Sift-Down и Sift-Up. Бинарные и к-ичные кучи. Построение кучи за линейное время. Алгоритм сортировки Heap-Sort.
12. Хеш-функции. Коллизии. Разрешение коллизий методом цепочек. Гипотеза простого равномерного хеширования, оценка средней длины цепочки.
13. Графы: проверка на ацикличность и топологическая сортировка.
14. Деревья поиска. Вставка и удаление элементов. Inorder-обход дерева.
15. Сильно связанные компоненты и топологическая сортировка конденсации.
16. Дерево отрезков.
17. Кратчайшие пути в графах. Оценки расстояний и их релаксация. Алгоритмы Беллмана-Форда, Флойда и Дейкстры.
18. Остовы минимального веса. Лемма о минимальном ребре в разрезе. Алгоритмы Краскала и Прима.
19. Системы непересекающихся множеств. Реализация с использованием леса.

Литература

- [1] Шень А. Программирование: теоремы и задачи. МЦМНО, 2007
- [2] Вирт Н. Алгоритмы и структуры данных. Изд-во Невский диалект, 2005
- [3] Керниган Б., Ритчи Д. Язык программирования С. Изд-во Вильямс, 2008
- [4] Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ, 2-е издание. - М. Издательский дом Вильямс, 2005
- [5] http://en.wikipedia.org/wiki/Code_Complete
- [6] 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.