

Список вопросов к экзамену по курсу

«Байесовские методы машинного обучения», осень 2021

1. Байесовский подход к теории вероятностей. Оценка параметров в байесовском и частотном подходе. Примеры байесовских рассуждений.
2. Сопряжённые распределения. Примеры. Экспоненциальный класс распределений, его свойства.
3. Решение задачи выбора модели по Байесу. Обоснованность модели. Полный байесовский вывод.
4. Вероятностная модель линейной регрессии. Метод релевантных векторов для задачи регрессии.
5. Логистическая и мультиномиальная регрессия. Метод релевантных векторов для задачи классификации.
6. ЕМ-алгоритм в общем виде. Примеры применения.
7. Вариационный подход для приближенного байесовского вывода.
8. Задача уменьшения размерности в данных. Вероятностная модель главных компонент, ее обучение с помощью метода максимального правдоподобия и ЕМ-алгоритма.
9. Байесовская модель разделения смеси гауссиан. Вариационный вывод для неё.
10. Методы генерации выборки из одномерных распределений.
11. Методы МСМС для оценки статистик вероятностных распределений. Теоретические свойства марковских цепей.
12. Схема Метрополиса-Хастингса и схема Гиббса. Примеры использования.
13. Гамильтонова динамика и ее применение для задачи сэмплирования. Динамика Ланжевена и ее применение в задаче сэмплирования (без теста Метрополиса).
14. Гауссовские процессы для задачи регрессии. Подбор параметров ковариационной функции.
15. Гауссовские процессы для задачи регрессии. Байесовская оптимизация гиперпараметров.
16. Гауссовские процессы и их применение для задачи классификации.
17. Распределение Дирихле. Свойства маргинального и обусловленного распределения Дирихле. Генерация выборки из Дирихле через гамма-распределения и через stick-breaking.
18. Процессы Дирихле. Представление процесса Дирихле с помощью процесса китайского ресторана. Схема Гиббса для разделения смеси распределений с процессом Дирихле.
19. Процессы Дирихле. Представление процесса Дирихле с помощью stick-breaking. Вариационный вывод для разделения смеси распределений с процессом Дирихле.
20. Тематическая модель LDA. Обучение и вывод в модели.

Теоретический минимум

Вопросы из этой части охватывают базовые математические понятия и алгоритмы, которые активно используются в курсе. Незнание ответа на любой вопрос из данной части автоматически влечёт за собой неудовлетворительную оценку по экзамену.

1. Базовые понятия. Мат.ожидание, мода, медиана, дисперсия и матрица ковариаций случайной величины. Понятие случайного процесса, функция среднего значения и ковариационная функция, стационарность, примеры процессов. Функция правдоподобия, метод максимального правдоподобия, его недостатки.
2. Условная вероятность. Правило суммы и произведения для вероятностей. Формула Байеса. Условная независимость случайных величин.

3. Сопряженные распределения: определение и примеры. Связь сопряженных распределений и полного байесовского вывода.
4. Табличные распределения: одномерное и многомерное нормальное распределение, биномиальное, мультиномиальное, гамма, бета, Уишарта, Дирихле, Стюдента, Гаусс-гамма - функциональная форма плотности распределения, где и почему применяются. Основные свойства нормального распределения (свертка двух нормальных, маргинальное распределение на часть переменных, условное распределение). Представление Стюдента через бесконечную смесь гауссиан.
5. Тожество Вудбери для эффективного обращения матриц.
6. Метод релевантных векторов для задачи регрессии: постановка задачи, зачем используется, с помощью какого метода обучается.
7. Дивергенция Кульбака-Лейблера, её основные свойства и использование для поиска аппроксимирующих распределений.
8. Общая схема ЕМ-алгоритма: какая задача решается и как именно? Формулы для оптимизируемого функционала, для Е и М шагов.
9. Общая схема вариационного вывода для приближённого байесовского вывода: какая задача решается и как именно? Формулы для оптимизируемого функционала, формулы для каждого шага.
10. Алгоритм Метрополиса-Хастингса: какая задача решается и как именно?
11. Схема Гиббса: какая задача решается и как именно?
12. Гауссовские процессы: определение, для чего могут быть использованы.