

Экзаменационные вопросы по курсу “Байесовские методы в анализе данных”

Формулы, которые будут на проекторе/доске во время экзамена:

- Формулы условной и маргинальной плотности нормального распределения
- Формула для вычисления свертки двух нормальных распределений
- Формулы среднего и матрицы ковариаций апостериорного распределения для метода релевантных векторов
- Формулы градиентов нормы, квадратичной функции, скалярного произведения

На все вопросы ожидаются **развернутые** и **подробные** ответы с **формулами**. Во всех вопросах необходимо уметь расшифровывать все обозначения.

В экзаменационной работе формулировки вопросов могут отличаться.

Теория вероятностей, математическая статистика и точный байесовский вывод

1. Запишите формулы условной вероятности, полной вероятности и формулу Байеса. Запишите определения математического ожидания, моды, медианы, дисперсии и матрицы ковариаций случайной величины.
2. Запишите формулу KL-дивергенции. Для чего она используется в байесовском машинном обучении? Каковы ее свойства?
3. Запишите плотность многомерного нормального распределения.
4. В чем состоят отличия оценки параметров в байесовском и частотном подходе?
5. Что такое метод максимального правдоподобия? В чем состоят его недостатки?
6. Поясните, почему поиск максимума апостериорной вероятности соответствует регуляризации модели.
7. Запишите определение сопряженного семейства распределений. Приведите два примера сопряженных семейств распределений.
8. Как сопряженные распределения используются в точном байесовском выводе?
9. Запишите определение экспоненциального класса распределений. Каковы его свойства? Приведите два примера распределений из этого класса.
10. Какие распределения необходимо задать, чтобы задать байесовскую модель машинного обучения?
11. Какое распределение необходимо искать на этапе обучения байесовской модели машинного обучения? А на этапе предсказания?
12. Что такое обоснованность модели? Как она связана с апостериорным распределением на параметры модели? Для чего она используется? Приведите пример.

Точный байесовский вывод для линейной регрессии

13. Запишите модель вероятностной линейной регрессии.
14. Запишите алгоритм обучения модели вероятностной линейной регрессии. Каковы гиперпараметры этой модели?
15. По какой формуле будут вычисляться среднее и дисперсия предсказания для нового объекта в модели вероятностной линейной регрессии (с выводом)?
16. Запишите модель метода релевантных векторов для задачи регрессии. Каким образом в этой модели производится отбор признаков? Можно ли точно рассчитать обоснованность в этой модели (с обоснованием)?

ЕМ-алгоритм для моделей со скрытыми переменными

17. Запишите, какой функционал оптимизируют в ЕМ-алгоритме. Нижней оценкой на какую величину он является? Запишите формулы Е- и М-шагов и поясните, как они выводятся.
18. Запишите модель вероятностного метода главных компонент. Что является наблюдаемыми, скрытыми переменными и параметрами в этой модели?
19. Запишите модель смеси гауссиан. Что является наблюдаемыми, скрытыми переменными и параметрами в этой модели?
20. Запишите формулу Е-шага для модели смеси гауссиан. Запишите, какой функционал необходимо минимизировать на М-шаге, и укажите, по каким величинам его необходимо минимизировать.
21. Выведите формулу М-шага для обновления центров гауссиан в модели смеси гауссиан.

Приближенный байесовский вывод: вариационный вывод и методы MCMC

22. Выведите нижнюю оценку $L(q)$ на логарифм правдоподобия $\log p(X)$ для произвольного распределения q на множестве параметров модели.
23. Запишите, по какой формуле ищут приближенное апостериорное распределение в факторизованном вариационном выводе. Выведите эту формулу.
- 24.

Рассмотрим следующую вероятностную модель:

$$\begin{aligned} p(X, \mu, \lambda) &= \left[\prod_{n=1}^N p(x_n | \mu, \lambda) \right] p(\mu, \lambda), \\ p(x_n | \mu, \lambda) &= \mathcal{N}(x_n | \mu, \lambda^{-1}), \\ p(\mu, \lambda) &= \mathcal{NG}(\mu, \lambda | m_0, \beta_0, a_0, b_0) = \mathcal{N}(\mu | m_0, (\beta_0 \lambda)^{-1}) \mathcal{G}(\lambda | a_0, b_0). \end{aligned}$$

Для поиска в рамках вариационного подхода факторизованного приближения вида

$$q(\mu)q(\lambda) \simeq p(\mu, \lambda | X)$$

требуется выписать формулы для итерационного пересчёта факторов $q(\mu)$, $q(\lambda)$, а также значение функционала $\mathcal{L}(q)$.

25. Каким образом семплирование используется в приближенном байесовском выводе?
26. Как работают методы rejection sampling и importance sampling? Позволяют ли эти методы семплировать из распределения, известного с точностью до константы? Если да, то каким образом?
27. Что такое распределение, инвариантное относительно марковской цепи? Что такое эргодичная марковская цепь? Запишите и докажите достаточное условие инвариантности.
28. Запишите схему Метрополиса-Хастингса семплирования выборки из распределения, известного с точностью до нормировочной константы.
29. Запишите схему Гиббса семплирования выборки.

Байесовское глубинное обучение

30. Что такое байесовская нейронная сеть? Какой функционал оптимизируют при обучении этой модели? Как вычислять/оценивать слагаемые в этом функционале?
31. Как работает reparametrization trick? Для чего он нужен?
32. Запишите модель SparseVD (sparse variational dropout) для нейронной сети. Какие величины необходимо настроить в процессе обучения? Как обучают эту модель? Как работает модель в режиме предсказания? Почему использование этой модели приводит к разреженным весам нейронной сети?
33. Какую задачу решает вариационный автокодировщик? Запишите модель вариационного автокодировщика. Какой функционал оптимизируют при обучении этой модели?
34. Как работает метод REINFORCE? Укажите задачу, которую этот метод решает. Каковы проблемы этого метода? Приведите пример использования этого метода в вариационном автокодировщике.
35. Как работает метод Gumbel Softmax? Укажите задачу, которую этот метод решает. Каковы проблемы этого метода? Приведите пример использования этого метода в вариационном автокодировщике.
36. Что такое нормализующие потоки и для чего они используются? Как обучать модель нормализующего потока (подробно)?

Непараметрические методы

37. Что такое ядро? Как оно связано с переходом в признаковое пространство? Для чего ядра используются в машинном обучении?
38. Что такое гауссовский процесс?
39. В чем отличие параметрических и непараметрических методов машинного обучения? Приведите по два примера из каждой группы.

40. В чем состоит модель гауссовского процесса для задачи регрессии? Запишите формулу для среднего и дисперсии предсказания на новом объекте в этой модели.
41. В чем состоит обучение гауссовского процесса для задачи регрессии? Какие величины необходимо настраивать?
42. Какую задачу решает байесовская оптимизация? Запишите алгоритм байесовской оптимизации.

На экзамене могут быть задачи с такими формулировками (не исчерпывающий список):

- Найти сопряженное распределение
- Вычислить среднее апостериорного распределения в модели с сопряженными распределениями
- Показать, что семейство распределений принадлежит экспоненциальному классу
- Посчитать обоснованность модели
- Вывести формулы E- и M-шага
- Выполнить факторизованный вариационный вывод
- Записать схему МСМС или схему Гиббса для выданного распределения
- Какой метод байесовского вывода применим к модели?