

УДК 519.81
ББК 22.18
К75

Кохендерфер М., Уилер Т., Рэй К.

К75 Алгоритмы принятия решений / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2023. – 684 с.: ил.

ISBN 978-5-93700-187-0

Книга представляет собой введение в теорию алгоритмов принятия решений в условиях неопределенности, включая формулировки основных математических задач и методы их решения. Рассмотрены современные методы снижения вычислительной нагрузки и поиска оптимальных стратегий в различных сценариях – от простых регуляторов до стохастических многоагентных систем. Основное внимание уделяется планированию и обучению с подкреплением, хотя некоторые из представленных методов основаны на элементах обучения с учителем и оптимизации. Алгоритмы реализованы на языке программирования Julia.

Издание предназначено специалистам в области искусственного интеллекта и систем принятия решений, а также может быть полезно студентам и аспирантам.

УДК 519.81
ББК 22.18

The rights to the Russian-language edition obtained through Alexander Korzhenevski Agency (Moscow).

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

ISBN 978-0-2620-4701-2 (англ.)

ISBN 978-5-93700-187-0 (рус.)

Copyright © 2022 Massachusetts
Institute of Technology
© Перевод, оформление, издание,
ДМК Пресс, 2023

Содержание

От издательства	14
Предисловие	15
Благодарности	16
 1 Введение	 17
1.1. Принятие решений	17
1.2. Области применения	18
1.2.1. Предотвращение столкновения самолетов	19
1.2.2. Автоматизированное вождение.....	19
1.2.3. Скрининг рака молочной железы	19
1.2.4. Доля инвестиций и распределение портфеля.....	20
1.2.5. Распределенное наблюдение за лесными пожарами	20
1.2.6. Исследование Марса	21
1.3. Методы создания агентов.....	21
1.3.1. Явное программирование	22
1.3.2. Обучение с учителем	22
1.3.3. Оптимизация.....	22
1.3.4. Планирование	22
1.3.5. Обучение с подкреплением.....	23
1.4. История автоматизации принятия решений.....	23
1.4.1. Экономика	24
1.4.2. Психология	25
1.4.3. Нейробиология.....	25
1.4.4. Информатика	26
1.4.5. Инженерия.....	26
1.4.6. Математика	27
1.4.7. Исследование операций.....	28
1.5. Воздействие на общество	28

1.6. Краткий обзор содержания книги	30
1.6.1. Вероятностное рассуждение	30
1.6.2. Многостадийные задачи.....	30
1.6.3. Неопределенность модели	31
1.6.4. Неопределенность состояния.....	31
1.6.5. Мультиагентные системы.....	32

Часть I. Вероятностные рассуждения.....33

2 Формальное представление неопределенности34

2.1. Степени доверия и вероятности	34
2.2. Распределения вероятностей.....	35
2.2.1. Дискретные распределения вероятностей.....	35
2.2.2. Непрерывные распределения вероятностей	36
2.3. Совместные распределения	41
2.3.1. Дискретные совместные распределения.....	41
2.3.2. Непрерывное совместное распределение.....	44
2.4. Условные распределения	47
2.4.1. Дискретные модели условных распределений	48
2.4.2. Условные модели Гаусса.....	49
2.4.3. Линейные модели Гаусса	49
2.4.4. Условные линейные модели Гаусса.....	50
2.4.5. Сигмовидные модели	50
2.4.6. Детерминированные переменные	51
2.5. Байесовские сети.....	51
2.6. Условная независимость.....	54
2.7. Заключение	57
2.8. Упражнения	57

3 Вероятностный вывод62

3.1. Вывод в байесовских сетях	62
3.2. Вывод в наивных байесовских моделях	67
3.3. Исключение переменной суммированием-перемножением.....	70
3.4. Распространение доверия	72
3.5. Вычислительная сложность.....	72
3.6. Прямая выборка	73
3.7. Выборка, взвешенная по правдоподобию	76
3.8. Выборка Гиббса	79
3.9. Вывод в гауссовых моделях	81
3.10. Заключение.....	83
3.11. Упражнения	84

4 Параметрическое обучение90

4.1. Обучение по критерию максимального правдоподобия	90
---	----

4.1.1. Оценки максимального правдоподобия для категориальных распределений	91
4.1.2. Оценки максимального правдоподобия для распределений Гаусса	92
4.1.3. Оценки максимального правдоподобия для байесовских сетей.....	93
4.2. Байесовское параметрическое обучение	96
4.2.1. Байесовское обучение для бинарных распределений.....	97
4.2.2. Байесовское обучение для категориальных распределений	99
4.2.3. Байесовское обучение для байесовских сетей	100
4.3. Непараметрическое обучение.....	101
4.4. Обучение с отсутствующими данными	103
4.4.1. Подстановка данных	104
4.4.2. Алгоритм ожидания-максимизации	107
4.5. Заключение.....	109
4.6. Упражнения	110

5 Структурное обучение

5.1. Оценка байесовской сети	116
5.2. Поиск ориентированного графа	119
5.3. Марковские классы эквивалентности	123
5.4. Поиск частично ориентированного графа	124
5.5. Заключение.....	126
5.6. Упражнения	126

6 Простые решения.....

6.1. Ограничения рациональных предпочтений.....	129
6.2. Функции полезности.....	131
6.3. Выявление полезности	132
6.4. Принцип максимальной ожидаемой полезности.....	134
6.5. Сети принятия решений.....	136
6.6. Полезность информации.....	139
6.7. Иррациональность	141
6.8. Заключение.....	143
6.9. Упражнения	143

Часть II. Задачи последовательного принятия решений.....

7 Методы точного решения.....

7.1. Марковские процессы принятия решений.....	149
7.2. Оценка стратегии	153
7.3. Нахождение стратегии через функцию полезности.....	156
7.4. Итерация по стратегиям	157
7.5. Итерация по критерию	159

7.6. Асинхронная итерация по критерию.....	162
7.7. Представление задачи в виде линейной программы	164
7.8. Линейные системы с квадратичным вознаграждением	166
7.9. Заключение	170
7.10. Упражнения.....	171

8 Приближенное вычисление функции

полезности	179
8.1. Параметрические представления	179
8.2. Аппроксимация по ближайшему соседу	181
8.3. Ядерное сглаживание.....	183
8.4. Линейная интерполяция	185
8.5. Симплексная интерполяция	188
8.6. Линейная регрессия.....	191
8.7. Регрессия на основе нейронной сети.....	195
8.8. Заключение.....	196
8.9. Упражнения	196

9 Онлайн-планирование

9.1. Планирование с отступающим горизонтом.....	201
9.2. Стратегия развертывания.....	203
9.3. Прямой поиск	204
9.4. Метод ветвей и границ	206
9.5. Разреженная выборка	207
9.6. Поиск по дереву Монте-Карло	209
9.7. Эвристический поиск.....	218
9.8. Эвристический поиск с разметкой	219
9.9. Планирование с открытым контуром	224
9.9.1. Прогнозирующее управление с детерминированной моделью	226
9.9.2. Робастное прогностическое управление.....	228
9.9.3. Многовариантное прогностическое управление.....	229
9.10. Заключение.....	231
9.11. Упражнения	231

10 Поиск стратегии

10.1. Приблизительная оценка стратегии.....	236
10.2. Локальный поиск	238
10.3. Генетические алгоритмы	241
10.4. Метод перекрестной энтропии	242
10.5. Эволюционные стратегии	244
10.6. Изотропные эволюционные стратегии	248
10.7. Заключение	250
10.8. Упражнения	251

11 Нахождение градиента стратегии	255
11.1. Конечная разность	255
11.2. Градиент регрессии	258
11.3. Отношение правдоподобия	260
11.4. Предстоящее вознаграждение	263
11.5. Вычитание базисного значения	266
11.6. Заключение	270
11.7. Упражнения	270
12 Оптимизация методом градиентного спуска по стратегиям	273
12.1. Обновление стратегии методом градиентного подъема	273
12.2. Ограниченное обновление градиента	275
12.3. Метод натурального градиента	277
12.4. Метод поиска в доверительной области	280
12.5. Зажатие замещенной цели	285
12.6. Заключение	288
12.7. Упражнения	289
13 Методы «актор–критик»	292
13.1. Определение актора и критика	292
13.2. Обобщенная оценка преимуществ	294
13.3. Градиент детерминированной стратегии	298
13.4. Метод «актор–критик» с поиском по дереву Монте-Карло	301
13.5. Заключение	303
13.6. Упражнения	304
14 Проверка стратегии	306
14.1. Оценка показателей качества стратегии	306
14.2. Моделирование редких событий	312
14.3. Анализ робастности системы	315
14.4. Анализ компромиссов	317
14.5. Состязательный анализ	319
14.6. Заключение	322
14.7. Упражнения	322
Часть III. Неопределенность модели	325
15 Исследование среды и использование знаний	326
15.1. Задача однорукого бандита	326
15.2. Оценка байесовской модели	328
15.3. Стратегии ненаправленного исследования	330
15.4. Стратегии направленного исследования	332

15.5. Оптимальные стратегии исследования	336
15.6. Исследование с несколькими состояниями	338
15.7. Заключение	338
15.8. Упражнения	339

16 Методы на основе моделей

16.1. Модели максимального правдоподобия	343
16.2. Схемы обновления модели.....	346
16.2.1. Полное обновление	346
16.2.2. Рандомизированное обновление.....	347
16.2.3. Приоритетный механизм обновления	347
16.3. Исследование.....	349
16.4. Байесовские методы	352
16.5. Адаптивные по Байесу марковские процессы принятия решений.....	355
16.6. Апостериорная выборка	356
16.7. Заключение	358
16.8. Упражнения	359

17 Свободные методы обучения с подкреплением

17.1. Инкрементное вычисление среднего значения распределения.....	362
17.2. Q-обучение.....	365
17.3. Алгоритм SARSA	367
17.4. Следы приемлемости	369
17.5. Формирование вознаграждения	371
17.6. Аппроксимация функции полезности действия	371
17.7. Воспроизведение опыта	375
17.8. Заключение	378
17.9. Упражнения.....	378

18. Имитационное обучение

18.1. Поведенческое копирование.....	383
18.2. Агрегация наборов данных	386
18.3. Итеративное обучение путем стохастического смешивания	389
18.4. Обратное обучение с подкреплением с максимальной разницей	392
18.5. Обратное обучение с подкреплением с максимальной энтропией	396
18.6. Генеративно-сопоставительное имитационное обучение	399
18.7. Заключение	400
18.8. Упражнения	400

Часть IV. Неопределенность состояния.....

19 Убеждения

19.1. Начальные убеждения	406
19.2. Фильтр дискретных состояний	407

19.3. Фильтр Калмана	412
19.4. Расширенный фильтр Калмана	414
19.5. Сигма-точечный фильтр Калмана	415
19.6. Парциальный фильтр	418
19.7. Внесение частиц	422
19.8. Заключение	425
19.9. Упражнения	426

20 Точное планирование с использованием убеждений-состояний

436

20.1. MDP убеждений-состояний	436
20.2. Условные планы	437
20.3. Альфа-векторы	441
20.4. Сокращение	444
20.5. Итерация по полезности	447
20.6. Линейные стратегии	449
20.7. Заключение	451
20.8. Упражнения	451

21 Офлайн-планирование с использованием убеждений-состояний

455

21.1. Аппроксимация полностью наблюдаемой полезности	455
21.2. Метод быстрой инфограницы	458
21.3. Методы быстрой оценки снизу	459
21.4. Точечная итерация по полезности	461
21.5. Рандомизированная точечная итерация по полезности	464
21.6. Пилообразная оценка сверху	465
21.7. Выбор точек в наборе убеждений	469
21.8. Пилообразный эвристический поиск	472
21.9. Триангулированные функции полезности	474
21.10. Заключение	477
21.11. Упражнения	478

22 Онлайн-планирование с использованием убеждений-состояний

483

22.1. Предпросмотр с развертываниями	483
22.2. Прямой поиск	483
22.3. Метод ветвей и границ	486
22.4. Разреженная выборка	486
22.5. Поиск по дереву Монте-Карло	487
22.6. Поиск по детерминированному разреженному дереву	490
22.7. Эвристический поиск на основе разности границ	494
22.8. Заключение	496

22.9. Упражнения	497
------------------------	-----

23 Понятие контроллера..... 500

23.1. Контроллеры.....	500
23.2. Итерация по стратегиям.....	504
23.3. Нелинейное программирование.....	509
23.4. Градиентный подъем.....	512
23.5. Заключение.....	518
23.6. Упражнения	519

Часть V. Многоагентные системы..... 521

24 Логический вывод в многоагентных системах..... 522

24.1. Простые игры	522
24.2. Модели откликов.....	525
24.2.1. Наилучший отклик.....	526
24.2.2. Отклик softmax.....	526
24.3. Равновесие доминирующей стратегии	527
24.4. Равновесие Нэша.....	528
24.5. Согласованное равновесие	530
24.6. Итеративный поиск лучшего отклика	533
24.7. Иерархическая форма модели softmax	534
24.8. Фиктивная игра	536
24.9. Градиентный подъем.....	539
24.10. Заключение.....	542
24.11. Упражнения	542

25 Последовательные задачи..... 548

25.1. Марковские игры.....	548
25.2. Модели отклика.....	550
25.2.1. Наилучший отклик.....	551
25.2.2. Стратегия отклика softmax.....	551
25.3. Равновесие Нэша.....	552
25.4. Фиктивная марковская игра.....	553
25.5. Градиентный подъем.....	557
25.6. Q-обучение Нэша	559
25.7. Заключение	561
25.8. Упражнения	561

26 Неопределенность состояния..... 564

26.1. Частично наблюдаемые марковские игры	564
26.2. Оценка стратегии.....	566
26.2.1. Оценка условных планов	566
26.2.2. Оценка стохастических контроллеров	568

26.3. Равновесие Нэша.....	569
26.4. Динамическое программирование.....	571
26.5. Заключение.....	574
26.6. Упражнения	575
27 Совместные действия агентов	577
27.1. Децентрализованные частично наблюдаемые марковские процессы принятия решений	577
27.2. Подклассы	578
27.3. Динамическое программирование	582
27.4. Итерация по наилучшим откликам	582
27.5. Эвристический поиск.....	584
27.6. Нелинейное программирование	587
27.7. Заключение	588
27.8. Упражнения.....	589
A Основные математические понятия.....	592
B Распределения вероятностей	604
C Вычислительная сложность	606
D Представление функций в форме нейронных сетей.....	610
E Алгоритмы поиска	628
F Задачи принятия решений.....	637
G Язык программирования Julia.....	655
Предметный указатель.....	677