

Ю. С. Попков

РАНДОМИЗАЦИЯ И ЭНТРОПИЯ В ОБРАБОТКЕ ДАННЫХ, ДИНАМИЧЕСКИХ СИСТЕМАХ, МАШИННОМ ОБУЧЕНИИ

Теория, методы, алгоритмы

- Энтропийные модели неопределенности
- Обработка данных
- Динамические системы с энтропийным оператором
- Макродинамика одного класса нелинейных марковских процессов
- Энтропийно-рандомизированное оценивание функций плотности распределения вероятностей
- Рандомизированное машинное обучение
- Рандомизированное машинное прогнозирование
- Энтропийная кластеризация на основе ЭРО

Прикладные задачи прогнозирования

- Численность мирового населения
- Суточная электрическая нагрузка энергетической системы
- Эволюция площади термокарстовых озер



Ю. С. Попков

**РАНДОМИЗАЦИЯ
И ЭНТРОПИЯ
В ОБРАБОТКЕ ДАННЫХ,
ДИНАМИЧЕСКИХ СИСТЕМАХ,
МАШИННОМ ОБУЧЕНИИ**



**URSS
МОСКВА**

Попков Юрий Соломонович

Рандомизация и энтропия в обработке данных, динамических системах, машинном обучении. — М.: ЛЕНАНД, 2023. — 300 с.

В основе данной книги лежат рандомизация и энтропия. Возникшее в начале века сочетание этих двух подходов оказалось весьма плодотворным для решения разнообразных задач. В настоящей книге собраны результаты исследований последних десяти лет, формулирующие общую концепцию энтропийно-рандомизированного оценивания и ее применение в задачах редукции размерности данных, «жесткого» и «мягкого» рандомизированного машинного обучения, прогнозирования и кластеризации. Значительное внимание уделено сравнительно новому разделу теории динамических (управляемых) систем с энтропийным оператором. Многие теоретические результаты доведены до вычислительных схем и алгоритмов. Несомненный интерес для читателя могут представлять приложения теории энтропийно-рандомизированного оценивания и прогнозирования динамики мирового населения, суточной электрической нагрузки энергетической системы и эволюции площади термокарстовых озер.

Книга может быть полезной для студентов, аспирантов и научных работников, интересующихся концепцией энтропийно-рандомизированного оценивания, а также ее практическими приложениями.

*На 1-й странице обложки использована иллюстрация:
Designed by starline / Freepik*

ООО «ЛЕНАНД».

117312, Москва, пр-т Шестидесятилетия Октября, д. 11А, стр. 11.

Формат 60×90/16. Печ. л. 18,75. Зак. № 184619.

Отпечатано в АО «Т 8 Издательские Технологии».

109316, Москва, Волгоградский проспект, д. 42, корп. 5.

ISBN 978-5-9710-6255-4

© ЛЕНАНД, 2022

34215 ID 295116



Все права защищены. Никакая часть настоящей книги не может быть воспроизведена или передана в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, а также размещение в Интернете, если на то нет письменного разрешения владельца.

Оглавление

| | |
|---|-----------|
| Предисловие | 9 |
| Введение | 11 |
| | |
| Часть I | |
| Теория, методы, алгоритмы | 14 |
| Глава 1. Энтропийные модели неопределенности | 14 |
| 1.1. Неопределенность | 14 |
| 1.2. Рандомизация | 17 |
| 1.3. Энтропия | 22 |
| 1.4. Механизмы формирования стационарных состояний макросистемы | 28 |
| Глава 2. Обработка данных | 33 |
| 2.1. Энтропийные «прямые» и «обратные» проекции | 33 |
| 2.1.1. Энтропийно-оптимальный метод «прямого» и «обратного» проектирования | 34 |
| 2.1.2. Параллельное проектирование с ограничениями информационной емкости | 40 |
| 2.1.3. Энтропийное случайное проектирование | 44 |
| 2.1.4. Случайные матрицы-проекторы с заданными значениями элементов | 47 |

| | |
|---|-----------|
| 2.2. Восстановление пропущенных данных | 54 |
| 2.2.1. Структура процедуры рандомизированного восстановления пропущенных данных | 57 |
| 2.2.2. Алгоритмы оптимизации АПД | 59 |
| Глава 3. Динамические системы с энтропийным оператором | 66 |
| 3.1. Математические модели динамических систем с энтропийным оператором | 68 |
| 3.2. Энтропийный оператор | 71 |
| 3.2.1. Качественные свойства энтропийного оператора в состоянии локального равновесия | 71 |
| 3.2.2. Оценивание локальной константы Липшица оператора с больцмановским энтропийным функционалом | 75 |
| 3.2.3. Оценивание локальной константы Липшица оператора с ферми-энтропийным функционалом | 93 |
| 3.3. Устойчивость ДСЭО с сепарабельным потоком | 104 |
| 3.4. Управляемые позитивные ДСЭО | 108 |
| 3.4.1. Математическая модель позитивной управляемой ДСЭО | 108 |
| 3.4.2. Сингулярные точки и их локализация | 111 |
| 3.4.3. Устойчивость «в большом» ненулевой сингулярной точки | 117 |
| 3.4.4. Квазиоптимальное управление одного класса позитивных ДСЭО | 120 |
| 3.5. Колебания в динамических системах с энтропийным оператором | 128 |
| 3.5.1. Математическая модель автономной ДСЭО . . | 128 |
| 3.5.2. Существование τ_0 -периодических колебаний . . | 129 |

| | |
|--|------------|
| 3.5.3. Асимптотический метод определения формы и параметров τ_0 -периодических колебаний . . . | 132 |
| 3.5.4. Асимптотический метод определения формы и параметров τ -периодических колебаний . . . | 134 |
| 3.5.5. Почти-периодические колебания в неавтономных ДСЭО | 137 |
| Глава 4. Макродинамика одного класса нелинейных марковских процессов . . | 141 |
| 4.1. Математическая модель | 141 |
| 4.2. Локально-стационарные макросостояния потоков иммиграционной и эмиграционной вероятностей | 145 |
| 4.3. Уравнение нелинейного марковского процесса | 148 |
| 4.4. Анализ уравнения НМП | 150 |
| Глава 5. Энтропийно-рандомизированное оценивание функций плотности распределения вероятностей | 152 |
| 5.1. Математическая формулировка метода ЭРО | 152 |
| 5.2. Условия оптимальности | 154 |
| 5.3. Существование неявной функции $\lambda(\mathbf{x}^{(r)}, \mathbf{y}^{(r)})$ | 156 |
| 5.4. Асимптотика ЭРО | 160 |
| Глава 6. Рандомизированное машинное обучение | 164 |
| 6.1. Рандомизированная модель и ее характеристика | 164 |
| 6.2. Алгоритм «жесткого» РМО | 166 |
| 6.3. Алгоритм «мягкого» РМО | 169 |

| | |
|---|------------|
| Глава 7. Рандомизированное машинное прогнозирование | 171 |
| 7.1. Рандомизированные временные ряды | 173 |
| 7.2. Рандомизированные динамические модели «вход-выход» | 180 |
| Глава 8. Энтропийная кластеризация на основе ЭРО | 186 |
| 8.1. Принцип рандомизированной кластеризации | 187 |
| 8.2. Числовая характеристика множества объектов | 188 |
| 8.3. Математическая модель задачи рандомизированной энтропийной кластеризации | 190 |
| 8.4. Алгоритмы рандомизированной бинарной кластеризации | 196 |
| 8.5. Иллюстративные примеры | 200 |
| | |
| Часть II | |
| Прикладные задачи | 210 |
| Глава 9. Прогнозирование численности мирового населения | 210 |
| 9.1. Рандомизированная модель численности | 211 |
| 9.2. Рандомизированное оценивание функций ПРВ | 213 |
| 9.3. Тестирование обученной модели численности | 219 |
| 9.4. Рандомизированный прогноз численности | 222 |

| | |
|--|------------|
| Глава 10. Прогнозирование суточной электрической нагрузки энергетической системы | 224 |
| 10.1. Модель динамической регрессии | 224 |
| 10.2. Обучающая коллекция данных | 227 |
| 10.3. Энтропийно-оптимальные функции ПРВ параметров и шумов | 232 |
| 10.4. Обучение моделей | 234 |
| 10.5. Тестирование модели | 238 |
| 10.6. Рандомизированное прогнозирование N -суточной нагрузки | 245 |
| | |
| Глава 11. Прогнозирование эволюции площади термокарстовых озер Западной Сибири | 247 |
| 11.1. Термокарстовые озера Западной Сибири, инструментальные средства и проблемы их исследования | 248 |
| 11.2. Восстановление пропусков данных дистанционного зондирования состояния термокарстовых озер арктической зоны . . . | 251 |
| 11.3. Рандомизированное прогнозирование эволюции площади термокарстовых озер | 259 |
| | |
| Литература | 284 |