

БАКАЛАВРИАТ

А. В. Макшанов, А. Е. Журавлев

ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ



www.e.lanbook.com



ЭБС
ЛАНЬ® **ЛАНЬ**

ОГЛАВЛЕНИЕ

Введение	7
1. Информация и данные	8
1.1. Три уровня анализа информации. Информация и данные	10
1.2. Информатика и информационные системы	11
1.3. Интеллектуальный анализ данных (ИАД).....	12
1.4. Data Mining.....	13
1.5. Подсистемы Data Mining.....	16
1.6. Контрольные вопросы.....	16
2. Системы и модели	17
2.1. Аналитические модели.....	17
2.2. Имитационные модели.....	17
2.3. Большие системы.....	18
2.4. Модели вместо законов.....	18
2.5. Многозначность понятий.....	19
2.6. Человек и машина.....	20
2.7. Теорема Гёделя	20
2.8. Генератор случая	21
2.9. Живое и неживое	21
2.10. Социо-кибер-физические системы.....	21
2.11. Контрольные вопросы	22
3. Методы матричного анализа	23
3.1. Системы нелинейных уравнений и матрица Якоби	23
3.2. Определение типа экстремума и матрица Гесса	26
3.3. Функции матричного аргумента	27
3.4. Системы дифференциальных уравнений	29
3.5. Скалярные функции матричного аргумента	32
3.6. Контрольные вопросы.....	33
4. Моделирование многомерных данных.....	34
4.1. Многомерный нормальный закон	34
4.1.1. Оценки параметров многомерного нормального закона	37
4.2. Моделирование многомерных случайных данных.....	38
4.3. Контрольные вопросы.....	41
5. Множественный регрессионный анализ	42
5.1. Оптимальный среднеквадратический линейный прогноз.....	42

5.2. Множественная регрессия	47
5.3. Восстановление пропусков в данных	50
5.4. Контрольные вопросы	50
6. Метод наименьших квадратов (МНК).....	51
6.1. Точечное оценивание параметров.....	51
6.2. Проверка гипотез и доверительное оценивание	56
6.3. Контрольные вопросы.....	58
7. Целенаправленное проектирование и редукция размерности	59
7.1. Главные компоненты и факторный анализ	59
7.2. Нелинейные главные компоненты (Проектирование с контрастированием)	64
7.3. Линейный дискриминантный анализ.....	64
7.4. Контрольные вопросы.....	69
8. Проверка статистических гипотез и информационные расстояния	70
8.1. Критерии значимости. Задача о выборе из двух гипотез.....	70
8.2. Основные критерии нормальной теории	73
8.3. Аппроксимация распределений статистик критериев.....	78
8.4. Контрольные вопросы.....	80
9. Классификация многомерных измерений.....	81
9.1. Дискриминантные информанты и классификация	81
9.2. Оценка вероятностей ошибочной классификации	86
9.3. Классификация на линейных дискриминантных формах	88
9.4. Контрольные вопросы.....	92
10. Кластерный анализ.....	93
10.1. Кластеризация. Выбор метрики	93
10.2. Метод k средних и EM -алгоритм.....	94
10.3. Иерархическая кластеризация на основе дендрограммы.....	96
10.4. Оценка качества разделения	98
10.5. Контрольные вопросы.....	99
11. Рекуррентные алгоритмы	100
11.1. Рекуррентное вычисление среднего.....	100
11.2. Одномерный линейный фильтр Калмана – Бьюси	100
11.3. m -мерный линейный фильтр Калмана.....	102
11.4. Варианты определения исходных данных	104
11.5. Контрольные вопросы.....	105
12. Генетические алгоритмы и алгоритмы прямого поиска	106
12.1. Назначение и возможности пакета.....	106

12.2. Краткие сведения о генетических алгоритмах.....	107
12.2.1. Естественный отбор в природе	107
12.2.2. Что такое генетический алгоритм.....	109
12.2.3. Особенности генетических алгоритмов	114
12.2.4. Структура генетического алгоритма пакета	118
12.3. Графический интерфейс генетического алгоритма	122
12.3.1. Общие правила работы с интерфейсом.....	122
12.4. Графические возможности интерфейса	125
12.4.1. Опции алгоритма.....	125
12.4.2. Экспорт и импорт результатов.....	129
12.5. Использование генетического алгоритма в режиме CLI.....	130
12.5.1. Использование алгоритма с опциями по умолчанию ..	130
12.6. Описание алгоритма поиска по образцу.....	131
12.7. Графический интерфейс алгоритма поиска по образцу	133
12.7.1. Общие правила работы с интерфейсом.....	133
12.7.2. Экспорт и импорт результатов.....	138
12.8. Использование алгоритма поиска по образцу в CLI	138
12.8.1. Использование алгоритма с опциями по умолчанию ..	138
12.8.2. Установка требуемых опций.....	140
12.9. Контрольные вопросы.....	141
13. Непараметрическая статистика.....	142
13.1. Сглаживание данных	142
13.2. Локально-полиномиальное сглаживание	143
13.2.1. Экспоненциальный фильтр	147
13.3. Контрольные вопросы.....	149
14. Нечеткие вычисления	150
14.1. Нечеткая информация и выводы	150
14.2. Нечеткая аппроксимация	154
14.3. Нечеткая кластеризация	156
14.4. Контрольные вопросы.....	158
15. Сокращение размерности и визуализация	159
15.1. Главные компоненты и факторный анализ	159
15.2. Нелинейные главные компоненты	163
15.3. Линейный дискриминантный анализ.....	164
15.4. Контрольные вопросы.....	169
16. Прикладной анализ временных рядов.....	170
16.1. Построение динамических стохастических моделей	170

16.2. Разложение Юла	175
16.3. Тренд и сезонные составляющие	175
16.4. Корреляционный и спектральный анализ	180
16.4.1. Преобразование Фурье	182
16.5. Анализ случайных процессов	184
16.5.1. Процессы в линейных системах	187
16.6. Частота Найквиста и теорема Котельникова.....	189
16.7. Быстрое преобразование Фурье (БПФ).....	190
16.8. Дискретные модели стационарных временных рядов.....	193
16.9. Авторегрессия: процессы Маркова и Юла	196
16.10. Уравнения Юла – Уокера.....	197
16.11. Оценивание коэффициентов параметрических моделей ...	199
16.11.1. Оценивание порядка модели	201
16.12. Прогнозирование	202
16.13. Контрольные вопросы	204
Заключение	205
Список литературы	207