

目录

第一章 最优化简介	1
1.1 最优化问题概括	1
1.1.1 最优化问题的一般形式	1
1.1.2 最优化问题的类型与应用背景	2
1.2 实例：稀疏优化	3
1.3 实例：低秩矩阵恢复	7
1.4 实例：深度学习	8
1.5 最优化的基本概念	12
1.5.1 连续和离散优化问题	13
1.5.2 无约束和约束优化问题	14
1.5.3 随机和确定性优化问题	14
1.5.4 线性和非线性规划问题	15
1.5.5 凸和非凸优化问题	15
1.5.6 全局和局部最优解	16
1.5.7 优化算法	16
1.6 总结	21
习题 1	22
第二章 基础知识	23
2.1 范数	23
2.1.1 向量范数	23
2.1.2 矩阵范数	24
2.1.3 矩阵内积	25
2.2 导数	26

2.2.1	梯度与海瑟矩阵	26
2.2.2	矩阵变量函数的导数	30
2.2.3	自动微分	32
2.3	广义实值函数	34
2.3.1	适当函数	35
2.3.2	闭函数	35
2.4	凸集	38
2.4.1	凸集的相关定义	38
2.4.2	重要的凸集	41
2.4.3	保凸的运算	43
2.4.4	分离超平面定理	44
2.5	凸函数	46
2.5.1	凸函数的定义	46
2.5.2	凸函数判定定理	48
2.5.3	保凸的运算	53
2.5.4	凸函数的性质	56
2.6	共轭函数	58
2.6.1	共轭函数的定义和例子	58
2.6.2	二次共轭函数	60
2.7	次梯度	61
2.7.1	次梯度的定义	61
2.7.2	次梯度的性质	64
2.7.3	凸函数的方向导数	66
2.7.4	次梯度的计算规则	68
2.8	总结	75
	习题 2	76
第三章	优化建模	79
3.1	建模技术	80
3.1.1	目标函数的设计	80
3.1.2	约束的设计	84
3.2	回归分析	86
3.2.1	概述	86
3.2.2	线性回归模型	87

3.2.3 正则化线性回归模型	88
3.3 逻辑回归	91
3.4 支持向量机	93
3.5 概率图模型	95
3.6 相位恢复	98
3.7 主成分分析	101
3.8 矩阵分离问题	103
3.9 字典学习	104
3.10 K-均值聚类	106
3.11 图像处理中的全变差模型	108
3.12 小波模型	111
3.13 强化学习	113
3.14 总结	115
习题 3	117
第四章 典型优化问题	121
4.1 线性规划	121
4.1.1 基本形式和应用背景	121
4.1.2 应用举例	122
4.2 最小二乘问题	125
4.2.1 基本形式和应用背景	125
4.2.2 应用举例	125
4.3 复合优化问题	130
4.3.1 基本形式和应用背景	130
4.3.2 应用举例	132
4.4 随机优化问题	134
4.4.1 基本形式和应用背景	134
4.4.2 应用举例	134
4.5 半定规划	136
4.5.1 基本形式和应用背景	136
4.5.2 应用举例	137
4.6 矩阵优化	140
4.6.1 基本形式和应用背景	140
4.6.2 应用举例	142

4.7	整数规划	145
4.7.1	基本形式和应用背景	145
4.7.2	应用举例	145
4.8	典型优化算法软件介绍	148
4.9	优化模型语言	149
4.9.1	CVX	149
4.9.2	AMPL	150
4.10	总结	151
	习题 4	152
第五章	最优性理论	157
5.1	最优化问题解的存在性	157
5.2	无约束可微问题的最优性理论	160
5.2.1	一阶最优性条件	160
5.2.2	二阶最优性条件	161
5.2.3	实例	163
5.3	无约束不可微问题的最优性理论	164
5.3.1	凸优化问题一阶充要条件	164
5.3.2	复合优化问题的一阶必要条件	165
*5.3.3	非光滑非凸问题的最优性条件	166
5.3.4	实例	168
5.4	对偶理论	169
5.4.1	拉格朗日函数与对偶问题	169
5.4.2	带广义不等式约束优化问题的对偶	171
5.4.3	实例	174
5.5	一般约束优化问题的最优性理论	179
5.5.1	一阶最优性条件	179
5.5.2	二阶最优性条件	189
5.6	带约束凸优化问题的最优性理论	191
5.6.1	Slater 约束品性与强对偶原理	192
5.6.2	一阶充要条件	195
*5.6.3	一阶充要条件: 必要性的证明	196
5.7	约束优化最优性理论应用实例	203
5.7.1	仿射空间的投影问题	203

5.7.2	线性规划问题	204
5.7.3	基追踪	205
5.7.4	最大割问题的半定规划松弛及其非凸分解模型	207
5.8	总结	209
	习题 5	210
第六章	无约束优化算法	215
6.1	线搜索方法	215
6.1.1	线搜索准则	216
6.1.2	线搜索算法	221
6.1.3	收敛性分析	222
6.2	梯度类算法	224
6.2.1	梯度下降法	225
6.2.2	Barzilar-Borwein 方法	229
6.2.3	应用举例	231
6.3	次梯度算法	237
6.3.1	次梯度算法结构	237
6.3.2	收敛性分析	237
6.3.3	应用举例	242
6.4	牛顿类算法	245
6.4.1	经典牛顿法	245
6.4.2	收敛性分析	245
6.4.3	修正牛顿法	247
6.4.4	非精确牛顿法	249
6.4.5	应用举例	250
6.5	拟牛顿类算法	252
6.5.1	割线方程	253
6.5.2	拟牛顿矩阵更新方式	255
6.5.3	拟牛顿法的全局收敛性	258
6.5.4	有限内存 BFGS 方法	260
6.5.5	应用举例	264
6.6	信赖域算法	266
6.6.1	信赖域算法框架	267
6.6.2	信赖域子问题求解	269

6.6.3	收敛性分析	276
6.6.4	应用举例	280
6.7	非线性最小二乘问题算法	280
6.7.1	非线性最小二乘问题	281
6.7.2	高斯-牛顿算法	282
6.7.3	Levenberg-Marquardt 方法	285
6.7.4	大残量问题的拟牛顿法	288
6.7.5	应用举例	290
6.8	总结	291
	习题 6	294
第七章	约束优化算法	297
7.1	罚函数法	297
7.1.1	等式约束的二次罚函数法	297
7.1.2	收敛性分析	300
7.1.3	一般约束问题的二次罚函数法	303
7.1.4	应用举例	305
7.1.5	其他类型的罚函数法	308
7.2	增广拉格朗日函数法	311
7.2.1	等式约束优化问题的增广拉格朗日函数法	311
7.2.2	一般约束优化问题的增广拉格朗日函数法	317
7.2.3	凸优化问题的增广拉格朗日函数法	320
7.2.4	基追踪问题的增广拉格朗日函数法	323
7.2.5	半定规划问题的增广拉格朗日函数法	330
7.3	线性规划内点法	332
7.3.1	原始-对偶算法	333
7.3.2	路径追踪算法	335
7.4	总结	338
	习题 7	339
第八章	复合优化算法	343
8.1	近似点梯度法	343
8.1.1	邻近算子	344
8.1.2	近似点梯度法	349

8.1.3	应用举例	351
8.1.4	收敛性分析	353
*8.1.5	非凸函数的邻近算子与近似点梯度法	357
8.2	Nesterov 加速算法	359
8.2.1	FISTA 算法	359
8.2.2	其他加速算法	364
8.2.3	应用举例	366
8.2.4	收敛性分析	369
8.3	近似点算法	373
8.3.1	近似点算法	374
8.3.2	与增广拉格朗日函数法的关系	375
8.3.3	应用举例	377
8.3.4	收敛性分析	382
8.3.5	Moreau-Yosida 正则化	384
8.4	分块坐标下降法	387
8.4.1	问题描述	388
8.4.2	算法结构	390
8.4.3	应用举例	393
*8.4.4	收敛性分析	400
8.5	对偶算法	412
8.5.1	对偶近似点梯度法	413
8.5.2	原始-对偶混合梯度算法	418
8.5.3	应用举例	420
8.5.4	收敛性分析	424
8.6	交替方向乘法	430
8.6.1	交替方向乘法	430
8.6.2	Douglas-Rachford Splitting 算法	436
8.6.3	常见变形和技巧	440
8.6.4	应用举例	443
*8.6.5	收敛性分析	455
8.7	随机优化算法	461
8.7.1	随机梯度下降算法	463
8.7.2	应用举例	470

8.7.3	收敛性分析	473
8.7.4	方差减小技术	482
8.8	总结	489
	习题 8	491
附录 A	符号表	495
附录 B	数学基础	499
B.1	线性代数基础	499
B.1.1	矩阵内积与迹	499
B.1.2	正交矩阵与 (半) 正定矩阵	499
B.1.3	矩阵的秩	500
B.1.4	像空间和零空间	500
B.1.5	行列式	501
B.1.6	特征值与特征向量	501
B.1.7	广义逆	502
B.1.8	Sherman-Morrison-Woodbury 公式	503
B.1.9	Schur 补	504
B.2	数值代数基础	505
B.2.1	解线性方程组	505
B.2.2	系数矩阵为特殊矩阵的方程组解法	508
B.2.3	特征值分解与奇异值分解	511
B.2.4	数值代数软件	513
B.3	概率基础	518
B.3.1	概率空间	518
B.3.2	随机变量	519
B.3.3	条件期望	525
B.3.4	随机变量的收敛性	529
B.3.5	随机过程	529
B.3.6	概率不等式	531
	参考文献	535
	索引	557