

目次

第1章 序論	1
1.1 例：多項式曲線フィッティング	4
1.2 確率論	11
1.2.1 確率密度	17
1.2.2 期待値と分散	19
1.2.3 ベイズ確率	20
1.2.4 ガウス分布	24
1.2.5 曲線フィッティング再訪	28
1.2.6 ベイズ曲線フィッティング	30
1.3 モデル選択	31
1.4 次元の呪い	33
1.5 決定理論	37
1.5.1 誤識別率の最小化	38
1.5.2 期待損失の最小化	40
1.5.3 棄却オプション	41
1.5.4 推論と決定	42
1.5.5 回帰のための損失関数	45
1.6 情報理論	47
1.6.1 相対エントロピーと相互情報量	54
演習問題	57
第2章 確率分布	65
2.1 二値変数	66
2.1.1 ベータ分布	68
2.2 多値変数	72
2.2.1 デリクレ分布	74

2.3	ガウス分布	76
2.3.1	条件付きガウス分布	82
2.3.2	周辺ガウス分布	85
2.3.3	ガウス変数に対するベイズの定理	88
2.3.4	ガウス分布の最尤推定	91
2.3.5	逐次推定	92
2.3.6	ガウス分布に対するベイズ推論	95
2.3.7	スチューデントの t 分布	100
2.3.8	周期変数	102
2.3.9	混合ガウス分布	107
2.4	指数型分布族	110
2.4.1	最尤推定と十分統計量	113
2.4.2	共役事前分布	114
2.4.3	無情報事前分布	115
2.5	ノンパラメトリック法	117
2.5.1	カーネル密度推定法	119
2.5.2	最近傍法	122
	演習問題	125
第3章	線形回帰モデル	135
3.1	線形基底関数モデル	136
3.1.1	最尤推定と最小二乗法	138
3.1.2	最小二乗法の幾何学	140
3.1.3	逐次学習	141
3.1.4	正則化最小二乗法	142
3.1.5	出力変数が多次元の場合	144
3.2	バイアス-バリエンス分解	146
3.3	ベイズ線形回帰	151
3.3.1	パラメータの分布	151
3.3.2	予測分布	155
3.3.3	等価カーネル	156
3.4	ベイズモデル比較	160
3.5	エビデンス近似	164
3.5.1	エビデンス関数の評価	165
3.5.2	エビデンス関数の最大化	167
3.5.3	有効パラメータ数	169

3.6	固定された基底関数の限界	171
	演習問題	172
第4章	線形識別モデル	177
4.1	識別関数 (判別関数)	179
4.1.1	2クラス	179
4.1.2	多クラス	180
4.1.3	分類における最小二乗	182
4.1.4	フィッシャーの線形判別	185
4.1.5	最小二乗との関連	187
4.1.6	多クラスにおけるフィッシャーの判別	188
4.1.7	パーセプトロンアルゴリズム	190
4.2	確率的生成モデル	195
4.2.1	連続値入力	196
4.2.2	最尤解	199
4.2.3	離散特徴	200
4.2.4	指数型分布族	201
4.3	確率的識別モデル	202
4.3.1	固定基底関数	202
4.3.2	ロジスティック回帰	204
4.3.3	反復再重み付け最小二乗	206
4.3.4	多クラスロジスティック回帰	208
4.3.5	プロビット回帰	209
4.3.6	正準連結関数	211
4.4	ラプラス近似	213
4.4.1	モデルの比較と BIC	215
4.5	ベイズロジスティック回帰	217
4.5.1	ラプラス近似	217
4.5.2	予測分布	218
	演習問題	220
第5章	ニューラルネットワーク	225
5.1	フィードフォワードネットワーク関数	226
5.1.1	重み空間対称性	232
5.2	ネットワーク訓練	233
5.2.1	パラメータ最適化	237

5.2.2	局所二次近似	238
5.2.3	勾配情報の利用	240
5.2.4	勾配降下最適化	241
5.3	誤差逆伝播	242
5.3.1	誤差関数微分の評価	243
5.3.2	単純な例	246
5.3.3	逆伝播の効率	247
5.3.4	ヤコビ行列	248
5.4	ヘッセ行列	250
5.4.1	対角近似	251
5.4.2	外積による近似	252
5.4.3	ヘッセ行列の逆行列	253
5.4.4	有限幅の差分による近似	254
5.4.5	ヘッセ行列の厳密な評価	255
5.4.6	ヘッセ行列の積の高速な計算	256
5.5	ニューラルネットワークの正則化	258
5.5.1	無矛盾なガウス事前分布	259
5.5.2	早期終了	261
5.5.3	不変性	264
5.5.4	接線伝播法	265
5.5.5	変換されたデータを用いた訓練	268
5.5.6	たたみ込みニューラルネットワーク	270
5.5.7	ソフト重み共有	272
5.6	混合密度ネットワーク	274
5.7	ベイズニューラルネットワーク	281
5.7.1	パラメータの事後分布	281
5.7.2	超パラメータ最適化	283
5.7.3	クラス分類のためのベイズニューラルネットワーク	285
	演習問題	287
付録 A データ集合		295
付録 B 確率分布の一覧		303
付録 C 行列の性質		313
付録 D 変分法		321

付録 E ラグランジュ乗数	325
上巻のための参考文献	329
和文索引	334
英文索引	342

目次

第 6 章	カーネル法	1
6.1	双対表現	2
6.2	カーネル関数の構成	4
6.3	RBF ネットワーク	10
6.3.1	Nadaraya–Watson モデル	12
6.4	ガウス過程	14
6.4.1	線形回帰再訪	15
6.4.2	ガウス過程による回帰	17
6.4.3	超パラメータの学習	22
6.4.4	関連度自動決定	23
6.4.5	ガウス過程による分類	25
6.4.6	ラプラス近似	27
6.4.7	ニューラルネットワークとの関係	31
	演習問題	31
第 7 章	疎な解を持つカーネルマシン	35
7.1	最大マージン分類器	35
7.1.1	重なりのあるクラス分布	41
7.1.2	ロジスティック回帰との関係	47
7.1.3	多クラス SVM	48
7.1.4	回帰のための SVM	50
7.1.5	計算論的学習理論	54
7.2	関連ベクトルマシン	56
7.2.1	回帰問題に対する RVM	56
7.2.2	疎性の解析	60
7.2.3	分類問題に対する RVM	64

演習問題	68
第 8 章 グラフィカルモデル	71
8.1 ベイジアンネットワーク	72
8.1.1 例：多項式曲線フィッティング	74
8.1.2 生成モデル	77
8.1.3 離散変数	78
8.1.4 線形ガウスモデル	82
8.2 条件付き独立性	84
8.2.1 3つのグラフの例	85
8.2.2 有向分離 (D分離)	90
8.3 マルコフ確率場	96
8.3.1 条件付き独立性	96
8.3.2 分解特性	98
8.3.3 例：画像のノイズ除去	100
8.3.4 有向グラフとの関係	104
8.4 グラフィカルモデルにおける推論	107
8.4.1 連鎖における推論	108
8.4.2 木	112
8.4.3 因子グラフ	113
8.4.4 積和アルゴリズム	116
8.4.5 max-sum アルゴリズム	126
8.4.6 一般のグラフにおける厳密推論	131
8.4.7 ループあり確率伝播	132
8.4.8 グラフ構造の学習	134
演習問題	134
第 9 章 混合モデルと EM	139
9.1 <i>K</i> -means クラスタリング	140
9.1.1 画像分割と画像圧縮	144
9.2 混合ガウス分布 (Mixtures of Gaussians)	146
9.2.1 最尤推定	149
9.2.2 混合ガウス分布の EM アルゴリズム	151
9.3 EM アルゴリズムのもう一つの解釈	155
9.3.1 混合ガウス分布再訪	157
9.3.2 <i>K</i> -means との関連	159

9.3.3	混合ベルヌーイ分布	160
9.3.4	ベイズ線形回帰に関する EM アルゴリズム	164
9.4	一般の EM アルゴリズム	165
	演習問題	171
第 10 章	近似推論法	175
10.1	変分推論	176
10.1.1	分布の分解	177
10.1.2	分解による近似のもつ性質	180
10.1.3	例：一変数ガウス分布	184
10.1.4	モデル比較	187
10.2	例：変分混合ガウス分布	187
10.2.1	変分事後分布	189
10.2.2	変分下限	195
10.2.3	予測分布	196
10.2.4	混合要素数の決定	197
10.2.5	導出された分解	199
10.3	変分線形回帰	200
10.3.1	変分分布	201
10.3.2	予測分布	203
10.3.3	変分下限	203
10.4	指数型分布族	204
10.4.1	変分メッセージパッシング	206
10.5	局所の変分推論法	207
10.6	変分ロジスティック回帰	212
10.6.1	変分事後分布	213
10.6.2	変分パラメータの最適化	215
10.6.3	超パラメータの推論	216
10.7	EP 法	219
10.7.1	例：雑音データ問題	225
10.7.2	グラフィカルモデルと EP 法	227
	演習問題	231
第 11 章	サンプリング法	237
11.1	基本的なサンプリングアルゴリズム	239
11.1.1	標準的な分布	240

11.1.2	棄却サンプリング	242
11.1.3	適応的棄却サンプリング	244
11.1.4	重点サンプリング	246
11.1.5	SIR	249
11.1.6	サンプリングとEM アルゴリズム	250
11.2	マルコフ連鎖モンテカルロ	252
11.2.1	マルコフ連鎖	253
11.2.2	Metropolis–Hastings アルゴリズム	255
11.3	ギブスサンプリング	257
11.4	スライスサンプリング	261
11.5	ハイブリッドモンテカルロアルゴリズム	263
11.5.1	力学系	263
11.5.2	ハイブリッドモンテカルロアルゴリズム	267
11.6	分配関数の推定	269
	演習問題	271
第 12 章	連続潜在変数	275
12.1	主成分分析	277
12.1.1	分散最大化による定式化	277
12.1.2	誤差最小化による定式化	279
12.1.3	主成分分析の応用	282
12.1.4	高次元データに対する主成分分析	285
12.2	確率的な主成分分析	286
12.2.1	最尤法による主成分分析	290
12.2.2	EM アルゴリズムによる主成分分析	294
12.2.3	ベイズ的主成分分析	298
12.2.4	因子分析	302
12.3	カーネル主成分分析	304
12.4	非線形潜在変数モデル	308
12.4.1	独立成分分析	309
12.4.2	自己連想ニューラルネットワーク	310
12.4.3	非線形多様体のモデル化	313
	演習問題	318
第 13 章	系列データ	323
13.1	マルコフモデル	324

13.2 隠れマルコフモデル	328
13.2.1 HMM の最尤推定	333
13.2.2 フォワード-バックワードアルゴリズム	336
13.2.3 HMM の積和アルゴリズム	343
13.2.4 スケーリング係数	345
13.2.5 Viterbi アルゴリズム	347
13.2.6 隠れマルコフモデルの拡張	349
13.3 線形動的システム	353
13.3.1 LDS における推論	356
13.3.2 LDS の学習	360
13.3.3 LDS の拡張	362
13.3.4 粒子フィルタ	363
演習問題	365
第 14 章 モデルの結合	371
14.1 ベイズモデル平均化	372
14.2 コミッティ	373
14.3 ブースティング	374
14.3.1 指数誤差の最小化	377
14.3.2 ブースティングのための誤差関数	379
14.4 木構造モデル	380
14.5 条件付き混合モデル	384
14.5.1 線形回帰モデルの混合	384
14.5.2 ロジスティックモデルの混合	387
14.5.3 混合エキスパートモデル	390
演習問題	392
 下巻のための参考文献	 395
訳者あとがき	405
和文索引	408
英文索引	421