

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	3.6	固定された基底関数の限界	171
2.3.1	条件付きガウス分布	82	演習問題		172
2.3.2	周辺ガウス分布	85	第4章 線形識別モデル		177
2.3.3	ガウス変数に対するベイズの定理	88	4.1	識別関数 (判別関数)	179
2.3.4	ガウス分布の最尤推定	91	4.1.1	2クラス	179
2.3.5	逐次推定	92	4.1.2	多クラス	180
2.3.6	ガウス分布に対するベイズ推論	95	4.1.3	分類における最小二乗	182
2.3.7	スチューデントのt分布	100	4.1.4	フィッシャーの線形判別	185
2.3.8	周期変数	102	4.1.5	最小二乗との関連	187
2.3.9	混合ガウス分布	107	4.1.6	多クラスにおけるフィッシャーの判別	188
2.4	指数型分布族	110	4.1.7	パーセプトロンアルゴリズム	190
2.4.1	最尤推定と十分統計量	113	4.2	確率的生成モデル	195
2.4.2	共役事前分布	114	4.2.1	連続値入力	196
2.4.3	無情報事前分布	115	4.2.2	最尤解	199
2.5	ノンパラメトリック法	117	4.2.3	離散特徴	200
2.5.1	カーネル密度推定法	119	4.2.4	指数型分布族	201
2.5.2	最近傍法	122	4.3	確率的識別モデル	202
演習問題		125	4.3.1	固定基底関数	202
第3章 線形回帰モデル		135	4.3.2	ロジスティック回帰	204
3.1	線形基底関数モデル	136	4.3.3	反復再重み付け最小二乗	206
3.1.1	最尤推定と最小二乗法	138	4.3.4	多クラスロジスティック回帰	208
3.1.2	最小二乗法の幾何学	140	4.3.5	プロビット回帰	209
3.1.3	逐次学習	141	4.3.6	正準連結関数	211
3.1.4	正則化最小二乗法	142	4.4	ラプラス近似	213
3.1.5	出力変数が多次元の場合	144	4.4.1	モデルの比較と BIC	215
3.2	バイアス-バリエンス分解	146	4.5	ベイズロジスティック回帰	217
3.3	ベイズ線形回帰	151	4.5.1	ラプラス近似	217
3.3.1	パラメータの分布	151	4.5.2	予測分布	218
3.3.2	予測分布	155	演習問題		220
3.3.3	等価カーネル	156	第5章 ニューラルネットワーク		225
3.4	ベイズモデル比較	160	5.1	フィードフォワードネットワーク関数	226
3.5	エビデンス近似	164	5.1.1	重み空間対称性	232
3.5.1	エビデンス関数の評価	165	5.2	ネットワーク訓練	233
3.5.2	エビデンス関数の最大化	167	5.2.1	パラメータ最適化	237
3.5.3	有効パラメータ数	169			

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