

# 目次

まえがき

<b>1</b>	<b>はじめに</b> .....	<b>1</b>
1.1	現象のモデル化 .....	1
1.2	識別・判別 .....	6
1.3	次元圧縮 .....	9
1.4	分類 .....	10
<b>2</b>	<b>線形回帰モデル</b> .....	<b>13</b>
2.1	2変数間の関係を捉える .....	13
2.1.1	データとモデル .....	13
2.1.2	モデルの推定—最小2乗法 .....	16
2.1.3	モデルの推定—最尤法 .....	18
2.2	多変数間の関係を捉える .....	20
2.2.1	データとモデル .....	21
2.2.2	モデルの推定 .....	23
2.2.3	モデル選択 .....	30
2.2.4	幾何学的考察 .....	33
	<b>第2章演習問題</b> .....	<b>36</b>
<b>3</b>	<b>非線形回帰モデル</b> .....	<b>39</b>
3.1	現象のモデル化 .....	39
3.1.1	実際例 .....	41
3.2	基底関数に基づくモデル .....	43
3.2.1	スプライン .....	44
3.2.2	<i>B</i> -スプライン .....	46

3.2.3	動径基底関数	50	5.2	情報量規準	106
3.3	基底展開法	52	5.2.1	カルバック・ライブラー情報量	106
3.3.1	基底関数展開	52	5.2.2	情報量規準 AIC	109
3.3.2	モデルの推定	53	5.2.3	マルチモデル推測	116
3.3.3	モデルの評価と選択	57	5.3	ベイズ型モデル評価基準	117
3.4	正則化法	61	5.3.1	モデルの事後確率と BIC の導出	117
3.4.1	正則化最小 2 乗法	62	5.3.2	ベイズ推測とモデル平均化法	121
3.4.2	正則化最尤法	64	第 5 章演習問題		122
3.4.3	モデルの評価と選択	66	6	判別分析	127
3.4.4	リッジ回帰と lasso	70	6.1	フィッシャーの線形判別	127
第 3 章演習問題		76	6.1.1	基本的考え方	127
4	ロジスティック回帰モデル	79	6.1.2	線形判別関数	131
4.1	リスク予測モデル	79	6.1.3	多変数 2 群判別	137
4.1.1	比率データとモデル	79	6.1.4	事前確率と損失	139
4.1.2	2 値反応データ	83	6.1.5	誤判別率	140
4.2	複合リスク予測モデル	86	6.2	マハラノビス距離に基づく判別法	142
4.2.1	モデルの推定	87	6.2.1	線形判別	145
4.2.2	モデルの評価と選択	90	6.2.2	2 次判別	146
4.3	非線形ロジスティック回帰モデル	91	6.3	多群判別	147
4.3.1	モデルの推定	93	6.3.1	多群判別法	147
4.3.2	モデルの評価と選択	94	6.3.2	分析例	149
第 4 章演習問題		96	6.4	変数選択	154
5	モデル評価基準	99	6.4.1	予測誤差と変数選択	154
5.1	予測誤差に基づく評価基準	99	6.4.2	逐次選択法	156
5.1.1	予測 2 乗誤差	100	6.4.3	分析例	157
5.1.2	クロス・バリデーション	102	6.5	正準判別	159
5.1.3	$C_p$ 基準	104	6.5.1	正準判別による次元圧縮	160
			6.5.2	分析例	163
			第 6 章演習問題		166

<b>7</b>	<b>ベイズ判別</b>	169
7.1	ベイズの定理	169
7.2	ベイズ判別法	172
7.2.1	確率分布モデルと尤度	172
7.2.2	判別関数	175
7.3	ロジスティック判別	178
7.3.1	ロジスティック線形判別	179
7.3.2	非線形ロジスティック判別	184
7.3.3	多群の非線形判別	187
	<b>第7章演習問題</b>	190
<b>8</b>	<b>サポートベクターマシン</b>	193
8.1	分離超平面の構成	193
8.1.1	線形分離可能	193
8.1.2	マージン最大化	196
8.1.3	2次計画問題から双対問題へ	198
8.2	線形分離可能でない場合のテクニック	204
8.2.1	ソフトマージン	204
8.2.2	主問題から双対問題へ	208
8.3	線形から非線形へ	212
8.3.1	高次元特徴空間への写像	212
8.3.2	カーネル法	216
8.3.3	非線形判別	218
	<b>第8章演習問題</b>	221
<b>9</b>	<b>主成分分析</b>	225
9.1	主成分の構成	225
9.1.1	基本的考え方	226
9.1.2	主成分導出のプロセス	231

9.1.3	次元圧縮と情報損失	235
9.1.4	分析例	236
9.1.5	画像データの圧縮復元	240
9.1.6	特異値分解	244
9.2	カーネル主成分分析	247
9.2.1	データの中心化と固有値問題	247
9.2.2	高次元空間への写像	250
9.2.3	カーネル法	254
	<b>第9章演習問題</b>	258

## 10 クラスター分析

10.1	階層的分類法	261
10.1.1	個体間の類似度	262
10.1.2	クラスター間距離	263
10.1.3	クラスター形成のプロセス	265
10.1.4	ウォード法	269
10.2	非階層的分類法	274
10.2.1	$k$ -平均法	274
10.2.2	自己組織化マップ	275
10.3	混合分布モデル	278
10.3.1	混合分布モデルと分類	278
10.3.2	モデルの推定—EMアルゴリズム	280
	<b>第10章演習問題</b>	283

<b>付録A</b>	<b>ブートストラップ法</b>	285
A.1	ブートストラップ誤差推定	285
A.2	回帰モデル	287
A.3	ブートストラップ選択確率	288

<b>付録B</b>	<b>ラグランジュの未定乗数法</b>	289
B.1	等式制約条件下での最適化問題	289

目次

B.2	不等式制約条件下での最適化問題 .....	290
B.3	等式制約と不等式制約条件下での最適化問題 .....	291
<b>付録 C</b>	<b>EM アルゴリズム .....</b>	<b>294</b>
C.1	EM アルゴリズムの実行プロセス .....	294
C.2	混合分布モデルの推定 .....	295
参考文献		299
索引		303