

## 目 次

第 I 部 導 入	1
第 1 章 はじめに	3
1.1 ウェブアプリケーションへの推薦システム導入時の留意点 . . . . .	5
1.1.1 アルゴリズム上の工夫 . . . . .	6
1.1.2 最適化指標 . . . . .	8
1.1.3 探索と活用のトレードオフ . . . . .	9
1.1.4 推薦システムの評価 . . . . .	10
1.1.5 推薦と検索：プッシュとプル . . . . .	11
1.2 シンプルなスコアリングモデル：Most-Popular 推薦 . . . . .	12
1.3 演習 . . . . .	18
第 2 章 古典的手法	19
2.1 アイテム素性ベクトル . . . . .	21
2.1.1 カテゴリ化 . . . . .	21
2.1.2 bag-of-words . . . . .	23
2.1.3 トピックモデリング . . . . .	27
2.1.4 その他のアイテム素性ベクトル . . . . .	28
2.2 ユーザ素性ベクトル . . . . .	29
2.2.1 公表されているユーザプロファイル . . . . .	29

## xii 目次

2.2.2	アイテム素性ベクトルの利用	30
2.2.3	その他のユーザ素性ベクトル	31
2.3	素性ベクトルベースの手法	32
2.3.1	教師なし手法	32
2.3.2	教師あり手法	33
2.3.3	コンテキスト情報	37
2.4	協調フィルタリング	38
2.4.1	ユーザ間の類似度にもとづいた手法	39
2.4.2	アイテム間の類似度にもとづいた手法	41
2.4.3	行列分解	41
2.5	ハイブリッド法	45
2.6	まとめ	47
2.7	演習	48
<b>第3章</b>	<b>推薦問題における探索と活用</b>	<b>49</b>
3.1	探索と活用のトレードオフ	51
3.2	多腕バンディット問題	52
3.2.1	ベイジアンアプローチ	53
3.2.2	ミニマックスアプローチ	57
3.2.3	ヒューリスティックなバンディット戦略	59
3.2.4	備考	60
3.3	推薦システムにおける探索と活用	60
3.3.1	Most-Popular 推薦	60
3.3.2	個別化推薦	61
3.3.3	データスパースネス	62
3.4	スパースデータを用いた探索と活用	62
3.4.1	次元削減手法	63
3.4.2	次元削減を用いた探索と活用	65
3.4.3	オンラインモデル	66
3.5	まとめ	67
3.6	演習	67

<b>第4章 推薦システムの評価</b>	<b>68</b>
4.1 オフライン評価における従来手法	69
4.1.1 データ分割手法	70
4.1.2 精度評価指標	74
4.1.3 ランキング指標	75
4.2 オンラインバケットテスト	81
4.2.1 バケットの構築	81
4.2.2 オンラインパフォーマンス指標	84
4.2.3 テスト結果の解析	85
4.3 オフラインシミュレーション	87
4.4 オフラインリプレイ	90
4.4.1 基本的なリプレイ推定量	91
4.4.2 リプレイの拡張	94
4.5 まとめ	96
4.6 演習	96
<b>第II部 一般的な問題設定</b>	<b>97</b>
<b>第5章 問題設定とシステム構成</b>	<b>99</b>
5.1 問題設定	100
5.1.1 一般的な推薦モジュール	100
5.1.2 アプリケーション設定	104
5.1.3 一般的な統計手法	106
5.2 システム構成	108
5.2.1 主要な構成要素	109
5.2.2 システムの例	110
<b>第6章 Most-Popular 推薦</b>	<b>114</b>
6.1 アプリケーション例：Yahoo!Today モジュール	115
6.2 問題定義	117
6.3 ベイズ的手法	120

## xiv 目次

6.3.1	2 × 2 問題：2つのアイテム，2つの時点	121
6.3.2	K × 2 問題：K 個のアイテム，2つの時点	125
6.3.3	一般解	127
6.4	非ベイズ的手法	131
6.5	実証的評価	133
6.5.1	比較分析	133
6.5.2	各戦略の特徴	137
6.5.3	セグメンテーション分析	138
6.5.4	バケットテストの結果	141
6.6	巨大なコンテンツプール	143
6.7	まとめ	144
6.8	演習	144
<b>第7章</b>	<b>素性ベクトルベースの回帰による個別化</b>	<b>145</b>
7.1	高速オンライン双線形因子モデル (FOBFM)	148
7.1.1	概要	148
7.1.2	モデルの詳細	150
7.2	オフライン学習	153
7.2.1	EM アルゴリズム	153
7.2.2	E ステップ	155
7.2.3	M ステップ	156
7.2.4	スケーラビリティ	158
7.3	オンライン学習	158
7.3.1	ガウシアンオンラインモデル	159
7.3.2	ロジスティックオンラインモデル	159
7.3.3	探索-活用戦略	160
7.3.4	オンラインモデル選択	161
7.4	Yahoo! データセットの実例	162
7.4.1	My Yahoo! データセット	164
7.4.2	Yahoo! フロントページデータセット	167

7.4.3	オフライン双線形項なしの FOBFM	170
7.5	まとめ	170
7.6	演習	171
<b>第8章</b>	<b>因子モデルによる個別化</b>	<b>172</b>
8.1	回帰ベース潜在因子モデル (RLFM)	172
8.1.1	行列分解から RLFM	173
8.1.2	モデルの詳細	175
8.1.3	RLFM の確率過程	181
8.2	学習アルゴリズム	182
8.2.1	ガウス応答のための EM アルゴリズム	183
8.2.2	ロジスティック応答のための ARS ベース EM アルゴリズム	190
8.2.3	ロジスティック応答のための変分 EM アルゴリズム	194
8.3	コールドスタートの実例	197
8.4	時間依存するアイテムの大規模な推薦	202
8.4.1	オンライン学習	202
8.4.2	並列学習アルゴリズム	204
8.5	大規模問題の実例	207
8.5.1	MovieLens-1M データ	209
8.5.2	小規模な Yahoo! フロントページデータ	211
8.5.3	大規模な Yahoo! フロントページデータ	213
8.5.4	結果	217
8.6	まとめ	219
8.7	演習	219
<b>第 III 部</b>	<b>高度な話題</b>	<b>221</b>
<b>第9章</b>	<b>潜在ディリクレ分配による因子分解</b>	<b>223</b>
9.1	はじめに	223
9.2	モデル	225
9.2.1	概要	225

## xvi 目次

9.2.2	モデルの詳細	227
9.3	学習と予測	231
9.3.1	モデルの当てはめ	231
9.3.2	予測	238
9.4	実験	238
9.4.1	MovieLens データ	239
9.4.2	Yahoo! Buzz への適用	240
9.4.3	BookCrossing データセット	242
9.5	関連研究	245
9.6	まとめ	246
<b>第10章</b>	<b>コンテキスト依存推薦</b>	<b>248</b>
10.1	テンソル分解モデル	250
10.1.1	モデル	250
10.1.2	モデルの当てはめ	252
10.1.3	考察	253
10.2	階層的縮小	254
10.2.1	モデル	255
10.2.2	モデルの当てはめ	257
10.2.3	局所拡張テンソルモデル	260
10.3	多面的なニュース記事推薦	262
10.3.1	探索的データ分析	263
10.3.2	実証的評価	272
10.4	関連アイテム推薦	279
10.4.1	意味的な関連性	280
10.4.2	応答予測	281
10.4.3	予測応答と関係性の統合	282
10.5	まとめ	282

<b>第 11 章 多目的最適化</b>	<b>284</b>
11.1 アプリケーション設定	285
11.2 セグメントアプローチ	287
11.2.1 問題設定	287
11.2.2 目的関数の最適化	288
11.3 個別化アプローチ	291
11.3.1 主問題による定式化	292
11.3.2 ラグランジュ双対性	295
11.4 近似手法	299
11.4.1 クラスタリング	299
11.4.2 サンプリング	300
11.5 実験	301
11.5.1 実験設定	302
11.5.2 結果	304
11.6 関連研究	312
11.7 まとめ	314
<b>参考文献</b>	<b>317</b>
<b>索引</b>	<b>325</b>