

# ■ 目 次

■ シリーズの刊行にあたって	iii
■ まえがき	v
<b>Chapter 1</b>	
第 1 章 バンディット問題とは	1
1.1 はじめに	1
1.2 バンディット問題の例	3
1.3 確率的バンディットと敵対的バンディット	4
1.4 プレイヤー方策の評価法	5
1.5 バンディット問題の歴史	7
1.6 関連分野	10
1.7 本書の構成	11
<b>Chapter 2</b>	
第 2 章 確率的バンディット問題の基礎知識	13
2.1 中心極限定理による確率近似	13
2.2 裾確率の評価	15
2.3 大偏差原理	17
<b>Chapter 3</b>	
第 3 章 確率的バンディット問題の方策	21
3.1 定式化	21
3.2 理論限界	22
3.3 $\epsilon$ -貪欲法	27
3.4 尤度に基づく方策	30
3.4.1 UCB 方策	30
3.4.2 MED 方策	34
3.5 確率一致法とトンプソン抽出	36
3.5.1 確率一致法の特徴と解釈	37
3.5.2 トンプソン抽出	38
3.5.3 トンプソン抽出と UCB 方策の関係	42
3.6 最悪時の評価	42
3.6.1 最悪時の評価例	43
3.6.2 最悪時での最適方策	44

<b>Chapter 4</b>	<b>第 4 章 確率的バンディット問題のリグレット解析</b> . . . . .	47
	4.1 リグレットの分解 . . . . .	47
	4.1.1 収束後の挙動 . . . . .	48
	4.1.2 収束前の挙動 . . . . .	49
	4.2 累積分布関数と期待値 . . . . .	51
	4.3 UCB 方策の性能解析 . . . . .	52
	4.4 トンプソン抽出の性能解析 . . . . .	56
	4.4.1 事後分布の裾確率 . . . . .	56
	4.4.2 リグレット解析 . . . . .	58
<b>Chapter 5</b>	<b>第 5 章 敵対的バンディット問題</b> . . . . .	65
	5.1 問題設定 . . . . .	65
	5.2 オンライン学習理論と Hedge アルゴリズム . . . . .	68
	5.3 Exp3 方策 . . . . .	73
	5.4 Exp3.P 方策 . . . . .	78
	5.5 敵対的多腕バンディット問題のリグレット下界 . . . . .	85
	5.6 最適オーダーの方策 . . . . .	90
<b>Chapter 6</b>	<b>第 6 章 最適腕識別と A/B テスト</b> . . . . .	95
	6.1 定式化 . . . . .	96
	6.1.1 累積報酬最大化との違い . . . . .	97
	6.1.2 $\epsilon$ -最適腕識別 . . . . .	99
	6.1.3 単純リグレット . . . . .	100
	6.2 標本複雑度 . . . . .	101
	6.3 最適腕識別の方策 . . . . .	102
	6.3.1 一様選択に基づく方法 . . . . .	103
	6.3.2 スコアに基づく方法 . . . . .	107
	6.4 固定予算の設定 . . . . .	108
<b>Chapter 7</b>	<b>第 7 章 線形モデル上のバンディット問題</b> . . . . .	111
	7.1 線形バンディット . . . . .	111
	7.2 文脈付きバンディット . . . . .	115
	7.3 LinUCB 方策 . . . . .	117
	7.4 線形モデル上のトンプソン抽出 . . . . .	120
	7.4.1 正規分布モデルでの事後確率の計算 . . . . .	120
	7.4.2 多変量正規分布からの乱数生成 . . . . .	121
	7.4.3 誤差項が正規分布でない場合 . . . . .	122
	7.5 ロジスティック回帰モデル上のバンディット . . . . .	122

Chapter 8	第 8 章 連続腕バンディットとベイズ最適化 .....	127
	8.1 定式化と観測モデル .....	127
	8.2 リグレットの設定 .....	129
	8.3 期待値関数のクラス .....	130
	8.3.1 滑らかさの制約 .....	130
	8.3.2 ベイズ最適化 .....	130
	8.4 連続腕バンディットの方策 .....	134
	8.4.1 GP-UCB 方策 .....	134
	8.4.2 トンプソン抽出 .....	135
	8.4.3 期待改善量方策 .....	136
	8.4.4 多項式時間で実行可能な方策 .....	138
	8.5 共分散関数のパラメータ推定 .....	145
Chapter 9	第 9 章 バンディット問題の拡張 .....	147
	9.1 時間変化のあるバンディット問題 .....	147
	9.1.1 文脈付きバンディットに基づく方法 .....	147
	9.1.2 敵対的バンディットに基づく方法 .....	148
	9.1.3 有限回の時間変化がある場合 .....	149
	9.1.4 その他の手法 .....	151
	9.2 比較バンディット .....	151
	9.2.1 定式化 .....	152
	9.2.2 比較バンディットの方策 .....	154
	9.3 部分観測問題 .....	158
	9.3.1 部分観測問題の例 .....	159
	9.3.2 分類と理論限界 .....	164
	9.3.3 部分観測問題の方策 .....	166
	9.4 その他の拡張 .....	169
Chapter 10	第 10 章 バンディット手法の応用 .....	171
	10.1 モンテカルロ木探索 .....	171
	10.2 インターネット広告 .....	176
	10.3 推薦システム .....	180

付録 <b>A</b>	逆行列の更新 .....	189
付録 <b>B</b>	ベータ分布の裾確率 .....	191
■ 参考文献 .....		195
■ 索引 .....		203