

# ■ 目 次

■ シリーズの刊行にあたって	iii
■ はじめに	v
<b>Chapter 1</b>	
第 1 章 教師あり学習と正則化	1
1.1 教師あり学習	1
1.1.1 回帰	3
1.1.2 判別	4
1.1.3 過学習	6
1.2 正則化学習法	7
1.3 さまざまなスパース正則化	11
1.3.1 グループ正則化	11
1.3.2 一般化連結正則化	12
1.3.3 トレースノルム正則化	12
1.3.4 正則化関数の組合せ	13
<b>Chapter 2</b>	
第 2 章 凸解析の基本事項	15
2.1 凸関数と凸集合	15
2.2 劣微分と双対関数	19
2.3 フェンシエルの双対定理	26
2.4 近接写像	29
2.5 強凸関数と平滑凸関数の性質	34
<b>Chapter 3</b>	
第 3 章 確率的最適化とは	43
<b>Chapter 4</b>	
第 4 章 オンライン型確率的最適化	47
4.1 オンライン型確率的最適化の枠組み	47
4.2 オンライン学習と確率的最適化の関係	50
4.3 確率的勾配降下法 (SGD)	51
4.3.1 確率的勾配降下法の枠組みとアルゴリズム	52
4.3.2 確率的勾配降下法の収束レート	56
4.3.3 確率的勾配降下法の収束レート (強凸)	58

4.3.4	確率的勾配降下法の収束レートの証明 (一般形)	60
4.3.5	確率的鏡像降下法	65
4.3.6	ネステロフの加速法の適用	71
4.4	確率的双対平均化法 (SDA)	73
4.4.1	確率的双対平均化法のアルゴリズムと収束レート	74
4.4.2	強凸な正則化項における確率的双対平均化法	75
4.4.3	確率的双対平均化法の収束レートの証明 (一般形)	77
4.4.4	確率的双対平均化法の鏡像降下法への拡張	81
4.5	AdaGrad	84
4.6	ミニマックス最適性	89
4.7	オンライン型確率的最適化の汎化誤差について	92

Chapter 5

第 5 章	バッチ型確率的最適化	95
5.1	バッチ型確率的最適化の問題設定	95
5.2	確率的双対座標降下法	97
5.2.1	確率的双対座標降下法のアルゴリズム	97
5.2.2	確率的双対座標降下法の収束証明	103
5.3	確率的分散縮小勾配降下法	108
5.3.1	確率的分散縮小勾配降下法のアルゴリズム	108
5.3.2	確率的分散縮小勾配法の収束証明	111
5.4	確率的平均勾配法	115

Chapter 6

第 6 章	分散環境での確率的最適化	121
6.1	オンライン型確率的最適化の分散処理	122
6.1.1	単純平均	122
6.1.2	同期型・ミニバッチ法	124
6.1.3	非同期型分散 SGD: Hogwild!	127
6.2	バッチ型確率的最適化の分散処理: 確率的座標降下法	132
6.2.1	主問題における並列座標降下法	132
6.2.2	双対問題における並列座標降下法: COCOA	135

Appendix A

付録 A		141
A.1	有用な不等式	141
A.2	正則化学習法の 1 次最適化法 (近接勾配法)	142
A.2.1	平滑でない凸関数の最小化	144
A.2.2	平滑な凸関数の最小化	145
A.2.3	平滑な凸関数の最小化: ネステロフの加速法	148

■ 参考文献	157
■ 索引	163