

目次

■ シリーズの刊行にあたって	iii
■ まえがき	v

Chapter 1

第1章 序論	1
1.1 機械学習とは	1
1.2 オンライン学習とは	4
1.2.1 オンライン学習の黎明期	5
1.2.2 オンライン学習の再発見	5
1.3 オンライン学習の特徴	6
1.3.1 学習データをすぐ捨てられる	7
1.3.2 学習速度が速い	7
1.3.3 学習結果がいつでも使える	8
1.3.4 実装が簡単である	8
1.3.5 性能解析が容易である	8
1.4 オンライン学習の短所	9

Chapter 2

第2章 準備	11
2.1 数式を読む際の心構え	11
2.2 数式についての約束事	12
2.2.1 数値の表記方法について	12
2.2.2 ベクトル	12
2.2.3 総和記号	13
2.2.4 最小値, 最大値, argmin , argmax	14
2.2.5 絶対値	14
2.2.6 ノルム	14
2.2.7 指数関数	15
2.2.8 偏微分と勾配	15
2.2.9 偏微分に使うテクニック	16
2.2.10 計算量について	16
2.3 凸関数と非凸関数	17

Chapter 3

第3章 基礎	21
3.1 二値分類	21
3.2 線形分類器	22
3.3 パーセプトロン	23

3.4	目的関数と最適化手法	25
3.4.1	目的関数の表記	26
3.4.2	オンライン学習と収束	26
3.5	確率的勾配降下法	27
3.5.1	勾配法とは	28
3.5.2	勾配降下法	28
3.5.3	確率的勾配降下法	29
3.5.4	パーセプトロンの目的関数	30
3.5.5	パーセプトロンの学習アルゴリズムの導出	31
3.6	サポートベクトルマシン	32
3.6.1	線形分離可能とは	33
3.6.2	ハードマージン SVM の導出	33
3.6.3	ソフトマージン SVM の導出	34
3.6.4	SVM の目的関数の解釈	35
3.6.5	確率的勾配降下法による SVM の学習	36
3.7	ロジスティック回帰	38
3.7.1	ロジスティック回帰のパラメータ推定	39
3.7.2	ロジスティック回帰の解釈	40
3.8	正則化の効果	41
3.8.1	過学習を防ぐ	41
3.8.2	疎な解を作る正則化	42
3.8.3	正則化と事前分布	44
3.9	二値分類器の性能評価の方法	46
3.10	二値分類のまとめ	49
3.11	多クラス分類	50
3.11.1	どのような考え方があるのか	50
3.11.2	1 対他法	50
3.11.3	1 対 1 法	51
3.11.4	誤り訂正出力符号法	51
3.11.5	多クラス SVM	53
3.11.6	対数線形モデル	56
3.11.7	対数線形モデルの学習	56
3.12	自然言語処理への応用例	57
3.12.1	文書分類	57
3.12.2	単語分割	59
第 4 章 発展		63
4.1	高精度なオンライン学習	63
4.1.1	パーセプトロン	66
4.1.2	Passive Aggressive (PA)	66
4.1.3	Confidence Weighted Learning (CW)	71
4.1.4	Adaptive Regularization of Weight Vectors (AROW)	72
4.1.5	Soft Confidence-Weighted Learning (SCW)	73
4.1.6	まとめ	74

4.2	オンライン分散並列学習	74
4.2.1	確率的勾配降下法の並列化	74
4.2.2	Parallel SGD	77
4.2.3	Iterative Parameter Mixture (IPM)	78
4.2.4	Stale Synchronous Parallel (SSP)	80
4.3	深層学習で使われるオンライン学習	83
4.3.1	フィードフォワードニューラルネットワーク	83
4.3.2	誤差逆伝播法による勾配計算	85
4.3.3	ミニバッチ確率的勾配降下法	87
4.3.4	モーメンタム法	88
4.3.5	加速勾配法	90
4.3.6	パラメータの不安定性	91
4.3.7	学習率の減衰方法	92
4.3.8	AdaGrad	93
4.3.9	RMSprop	95
4.3.10	vSGD	96

第5章	性能解析	101
5.1	オンライン学習の性能	101
5.2	パーセプトロンの学習定理	101
5.3	線形分離可能でない場合のパーセプトロンの学習定理	104
5.4	リグレット解析	107
5.4.1	Follow The Leader (FTL)	110
5.4.2	Regularized Follow The Leader (RFTL)	111
5.4.3	損失関数が強凸関数のリグレット	116

第6章	実装	119
6.1	ベクトルの実装	119
6.1.1	ベクトルの内部表現	120
6.1.2	内積計算	121
6.1.3	ベクトルの加算	123
6.2	アルゴリズムの実装	126
6.2.1	平均化法の効率的な実装	126
6.2.2	lazy update	127
6.3	浮動小数点数における制約	131
6.3.1	浮動小数点数の仕組み	131
6.3.2	巨大な実数値の扱い	132
6.3.3	桁落ち	134

付録 A	137
A.1 式 (3.35) の導出	137
A.2 ラグランジュ未定乗数法と KKT 条件	138
A.3 CW の変数	140
A.4 AROW の変数	141
A.5 SCW の変数	141
A.6 SSP を用いた並列確率的勾配降下法のリグレット解析	142
■ 参考文献	151
■ 索引	157