

# 目次

概要	i
第 I 部 理論	1
第 1 章 数学的準備	2
1.1 線形代数	2
1.1.1 スカラー、ベクトル、行列、テンソルとそれらの記法	2
1.1.2 積の計算	2
1.1.3 行列の微分	3
1.1.4 ノルム	4
1.2 確率論	4
1.2.1 確率論を使うことの意義	4
1.2.2 確率変数と確率分布	4
1.2.3 条件付き確率と確率の連鎖律	6
1.2.4 ベイズの定理	6
1.2.5 期待値と分散	6
1.3 情報理論	7
1.3.1 自己情報量とシャノン・エントロピー	7
1.3.2 KL ダイバージェンスと交差エントロピー	8
第 2 章 機械学習の基礎	9
2.1 機械学習の基礎	9
2.2 最適化	9
2.2.1 最急降下法	10
2.2.2 確率的勾配降下法	11
2.2.3 ミニバッチ勾配降下法	11
2.3 ベイズ推定	11
2.3.1 ベイズの定理の意味とその例	12
2.3.2 最尤推定	12
2.3.3 最大事後確率推定	12
2.4 過学習の問題	12
2.4.1 過学習・過小学習	12

2.4.2	交差検証	12
2.4.3	正則化	13
第3章	フィードフォワード・ニューラルネットワーク	14
3.1	アーキテクチャとニューラルネットワークの働き	14
3.2	入力層	15
3.3	隠れ層	16
3.3.1	万能近似定理	16
3.3.2	活性化関数	16
3.3.2.1	ReLU	17
3.3.2.2	$\tanh(x)$	17
3.3.2.3	シグモイド関数	17
3.3.2.4	ソフトマックス関数	17
3.4	出力層	18
3.4.1	コスト関数	18
3.4.1.1	平均二乗誤差	19
3.4.1.2	交差エントロピー	19
3.5	誤差逆伝搬法	19
3.5.1	微分の連鎖律	19
3.5.2	誤差逆伝搬法を用いた学習	20
3.5.3	勾配消失問題	21
第4章	畳み込みニューラルネットワーク	22
4.1	畳み込み処理	22
4.2	ゼロパディングとストライド	23
4.3	フィルターの効果	23
4.4	プーリング	24
4.5	ドロップアウト	25
第II部	実装	26
第5章	方法	27
5.1	MNIST データセット	27
5.2	動作環境	27
5.3	フィードフォワード・ニューラルネットワークによる学習	28
5.3.1	ベースとなるフィードフォワード・ニューラルネットワークの構築	28
5.3.2	訓練データをシャッフルすることの効果	29
5.3.3	活性化関数の比較	29
5.3.4	コスト関数の比較	29
5.3.5	最適化アルゴリズムの比較	29
5.4	畳み込みニューラルネットワークによる学習	29

第 6 章	結果	30
6.1	. . . . .	30
6.1.1	. . . . .	30
6.1.1.1	. . . . .	30
第 7 章		31
7.1	. . . . .	31
7.1.1	. . . . .	31
7.1.1.1	. . . . .	31
付録 A	証明	32
謝辞		35