

目次

まえがき	iii
1章 Python 入門	1
1.1 Python とは	1
1.2 Python のインストール	2
1.2.1 Python のバージョン	2
1.2.2 使用する外部ライブラリ	2
1.2.3 Anaconda ディストリビューション	3
1.3 Python インタプリタ	4
1.3.1 算術計算	4
1.3.2 データ型	5
1.3.3 変数	5
1.3.4 リスト	6
1.3.5 デイクショナリ	7
1.3.6 プーリアン	7
1.3.7 if 文	8
1.3.8 for 文	8
1.3.9 関数	9
1.4 Python スクリプトファイル	9
1.4.1 ファイルに保存	9
1.4.2 クラス	10
1.5 NumPy	11

1.5.1	NumPy のインポート	11	3.1.2	パーセプトロンの復習	41
1.5.2	NumPy 配列の生成	12	3.1.3	活性化関数の登場	42
1.5.3	NumPy の算術計算	12	3.2	活性化関数	44
1.5.4	NumPy の N 次元配列	13	3.2.1	シグモイド関数	45
1.5.5	ブロードキャスト	14	3.2.2	ステップ関数の実装	45
1.5.6	要素へのアクセス	15	3.2.3	ステップ関数のグラフ	47
1.6	Matplotlib	16	3.2.4	シグモイド関数の実装	48
1.6.1	単純なグラフの描画	16	3.2.5	シグモイド関数とステップ関数の比較	49
1.6.2	pyplot の機能	17	3.2.6	非線形関数	51
1.6.3	画像の表示	18	3.2.7	ReLU 関数	51
1.7	まとめ	20	3.3	多次元配列の計算	53
2 章	パーセプトロン	21	3.3.1	多次元配列	53
2.1	パーセプトロンとは	21	3.3.2	行列の内積	54
2.2	単純な論理回路	23	3.3.3	ニューラルネットワークの内積	57
2.2.1	AND ゲート	23	3.4	3 層ニューラルネットワークの実装	58
2.2.2	NAND ゲートと OR ゲート	23	3.4.1	記号の確認	58
2.3	パーセプトロンの実装	25	3.4.2	各層における信号伝達の実装	60
2.3.1	簡単な実装	25	3.4.3	実装のまとめ	64
2.3.2	重みとバイアスの導入	26	3.5	出力層の設計	66
2.3.3	重みとバイアスによる実装	26	3.5.1	恒等関数とソフトマックス関数	66
2.4	パーセプトロンの限界	28	3.5.2	ソフトマックス関数の実装上の注意	68
2.4.1	XOR ゲート	28	3.5.3	ソフトマックス関数の特徴	70
2.4.2	線形と非線形	30	3.5.4	出力層のニューロンの数	71
2.5	多層パーセプトロン	31	3.6	手書き数字認識	72
2.5.1	既存ゲートの組み合わせ	31	3.6.1	MNIST データセット	72
2.5.2	XOR ゲートの実装	33	3.6.2	ニューラルネットワークの推論処理	75
2.6	NAND からコンピュータへ	35	3.6.3	バッチ処理	78
2.7	まとめ	36	3.7	まとめ	81
3 章	ニューラルネットワーク	39	4 章	ニューラルネットワークの学習	83
3.1	パーセプトロンからニューラルネットワークへ	39	4.1	データから学習する	83
3.1.1	ニューラルネットワークの例	39	4.1.1	データ駆動	84
			4.1.2	訓練データとテストデータ	86

4.2	損失関数	87	5.4.1	乗算レイヤの実装	137
4.2.1	2乗和誤差	88	5.4.2	加算レイヤの実装	139
4.2.2	交差エントロピー誤差	89	5.5	活性化関数レイヤの実装	141
4.2.3	ミニバッチ学習	91	5.5.1	ReLU レイヤ	141
4.2.4	[バッチ対応版] 交差エントロピー誤差の実装	94	5.5.2	Sigmoid レイヤ	143
4.2.5	なぜ損失関数を設定するのか?	95	5.6	Affine / Softmax レイヤの実装	147
4.3	数値微分	97	5.6.1	Affine レイヤ	147
4.3.1	微分	97	5.6.2	バッチ版 Affine レイヤ	150
4.3.2	数値微分の例	100	5.6.3	Softmax-with-Loss レイヤ	152
4.3.3	偏微分	102	5.7	誤差逆伝播法の実装	156
4.4	勾配	103	5.7.1	ニューラルネットワークの学習の全体図	156
4.4.1	勾配法	106	5.7.2	誤差逆伝播法に対応したニューラルネットワークの実装	157
4.4.2	ニューラルネットワークに対する勾配	109	5.7.3	誤差逆伝播法の勾配確認	160
4.5	学習アルゴリズムの実装	112	5.7.4	誤差逆伝播法を使った学習	162
4.5.1	2層ニューラルネットワークのクラス	113	5.8	まとめ	163
4.5.2	ミニバッチ学習の実装	117			
4.5.3	テストデータで評価	119			
4.6	まとめ	122			
5	章 誤差逆伝播法	123	6	章 学習に関するテクニック	165
5.1	計算グラフ	123	6.1	パラメータの更新	165
5.1.1	計算グラフで解く	124	6.1.1	冒険家の話	166
5.1.2	局所的な計算	126	6.1.2	SGD	166
5.1.3	なぜ計算グラフで解くのか?	127	6.1.3	SGD の欠点	168
5.2	連鎖律	129	6.1.4	Momentum	170
5.2.1	計算グラフの逆伝播	129	6.1.5	AdaGrad	172
5.2.2	連鎖律とは	129	6.1.6	Adam	175
5.2.3	連鎖律と計算グラフ	131	6.1.7	どの更新手法を用いるか?	175
5.3	逆伝播	132	6.1.8	MNIST データセットによる更新手法の比較	176
5.3.1	加算ノードの逆伝播	132	6.2	重みの初期値	177
5.3.2	乗算ノードの逆伝播	134	6.2.1	重みの初期値を 0 にする?	178
5.3.3	リングの例	135	6.2.2	隠れ層のアクティベーション分布	179
5.4	単純なレイヤの実装	137	6.2.3	ReLU の場合の重みの初期値	183
			6.2.4	MNIST データセットによる重み初期値の比較	184
			6.3	Batch Normalization	186
			6.3.1	Batch Normalization のアルゴリズム	187

6.3.2	Batch Normalization の評価	189	7.7.1	LeNet	236
6.4	正規化	189	7.7.2	AlexNet	237
6.4.1	過学習	190	7.8	まとめ	238
6.4.2	Weight decay	193			
6.4.3	Dropout	195			
6.5	ハイパーパラメータの検証	197	8章 ディープラーニング	241	
6.5.1	検証データ	198	8.1	ネットワークをより深く	241
6.5.2	ハイパーパラメータの最適化	199	8.1.1	よりディープなネットワークへ	241
6.5.3	ハイパーパラメータ最適化の実装	201	8.1.2	さらに認識精度を高めるには	244
6.6	まとめ	203	8.1.3	層を深くすることのモチベーション	246
			8.2	ディープラーニングの小歴史	249
7章 畳み込みニューラルネットワーク	205		8.2.1	ImageNet	249
7.1	全体の構造	205	8.2.2	VGG	250
7.2	畳み込み層	206	8.2.3	GoogLeNet	251
7.2.1	全結合層の問題点	207	8.2.4	ResNet	252
7.2.2	畳み込み演算	208	8.3	ディープラーニングの高速化	254
7.2.3	パディング	210	8.3.1	取り組むべき問題	255
7.2.4	ストライド	211	8.3.2	GPU による高速化	256
7.2.5	3次元データの畳み込み演算	214	8.3.3	分散学習	257
7.2.6	ブロックで考える	216	8.3.4	演算精度のビット削減	258
7.2.7	バッチ処理	218	8.4	ディープラーニングの実用例	260
7.3	プーリング層	219	8.4.1	物体検出	260
7.3.1	プーリング層の特徴	220	8.4.2	セグメンテーション	262
7.4	Convolution / Pooling レイヤの実装	221	8.4.3	画像キャプション生成	264
7.4.1	4次元配列	221	8.5	ディープラーニングの未来	265
7.4.2	im2col による展開	222	8.5.1	画像スタイル変換	266
7.4.3	Convolution レイヤの実装	224	8.5.2	画像生成	267
7.4.4	Pooling レイヤの実装	227	8.5.3	自動運転	268
7.5	CNN の実装	229	8.5.4	Deep Q-Network (強化学習)	269
7.6	CNN の可視化	233	8.6	まとめ	272
7.6.1	1層目の重みの可視化	234			
7.6.2	階層構造による情報抽出	235	付録 A Softmax-with-Loss レイヤの計算グラフ	275	
7.7	代表的な CNN	236	A.1	順伝播	276
			A.2	逆伝播	278

A.3 まとめ	285
参考文献	287
Python / NumPy	287
計算グラフ (誤差逆伝播法)	287
Deep Learning のオンライン授業 (資料)	287
パラメータの更新方法	287
重みパラメータの初期値	288
Batch Normalization / Dropout	288
ハイパーパラメータの最適化	289
CNN の可視化	289
代表的なネットワーク	289
データセット	290
計算の高速化	290
MNIST データセットの精度ランキングおよび最高精度の手法	291
ディープラーニングのアプリケーション	291
索引	293