ご購入はこちら

第5章 損失関数と勾配降下法を使ってパラメータを更新する バックプロパゲーションのからくりを解き明かす

ニューラル・ネットワークが学習を行う仕組み

サポート・ページに説明動画あり

新井 正敏

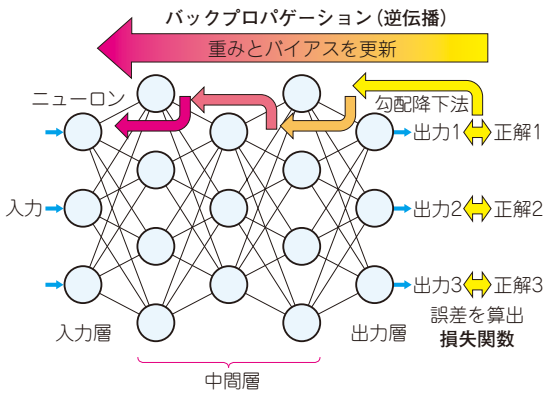

図1　全結合層のバックプロパゲーション（逆伝播）
出力層から入力層に向かって順次勾配を計算してパラメータを更新していく

● 全結合層の学習のメカニズムを解き明かす

第1章では，ニューロンがどのように学習しているかを説明しました．本章では，**図1**のように，複数のニューロンが結合した全結合層（全結合層モデル）で，出力層や中間層のニューロンがどのように学習をしていくのかをひも解いていきます．

全結合層モデルの出力層の出力（順伝播した予測値）と，あるべき値（正解値，ターゲット値）との間の誤差（エラー：E）を算出するための関数が**損失関数**（コスト関数）です．

損失関数の値（損失）を最小化することで，全結合層モデルの出力があるべき値に近づきます．最小化の代表的な手法の1つが**勾配降下法**（Gradient Descent）です．勾配降下法は，損失関数の**勾配**（偏微分）を計算し，その勾配に従ってパラメータ（重みやバイアス）を更新することで損失を最小化していきます．

● 学習のアルゴリズム…バックプロパゲーション

ニューラル・ネットワークの学習プロセスで，損失関数の勾配を計算し，勾配に基づいてパラメータを更新するアルゴリズムを**バックプロパゲーション**（Backpropagation；逆伝播）と呼びます．

バックプロパゲーションは，**図1**のように出力層から入力層に向かって順次勾配を計算してパラメータを更新していきます．このため，バックプロパゲーションを行うことで，ニューラル・ネットワーク（全結合層モデル）があるべき値に近い良い予測を行うように学習されます．

第1章では，シンプルな重みやバイアスの更新ルールで学習しましたが，バックプロパゲーションも基本的には同じです．本章では具体的にどのように学習するのか，からくりを示していきます．

5-1　学習に使う数学の道具

1 損失関数 …正解値と出力値との誤差を算出

プロローグ1で示したように，ディープ・ラーニングは回帰問題と分類問題を扱います．第4章の図14で示したように，全結合モデルの出力値（予測値）と正解値（あるべき値）の特性により，**表1**のような損失関数を使います．

本節の出力層では**表2**のようなニューロン数やインデックスを使います．

● 回帰の損失関数…平均二乗誤差

▶式と概要

回帰問題では，**図2**のように全結合層モデルが連続的な数値を予測（出力）するので，出力値と正解値の差を評価することが重要です．このため，回帰問題は