ラビットチャレンジ課題　作成者:内山 貴雄

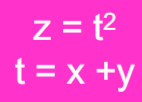
深層学習前編day2の各章の要約及び確認テスト等の考察

●勾配消失問題

本章では、動物の分類問題を例にニューラルネットワーク(NN)による分類問題の学習におけるその入力~中間層の役割や入出力の計算方法を学んだ。また、確認テストでは、前述の動物分類問題をNNにて学習した際を例とし、入力層・中間層の役割概念を学んだ後に、それらをPythonにて実装する手法を学んだ。

〇確認テスト

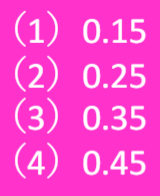
連鎖律の原理を使い、下記式のdz/dxを求めよ。



答案：

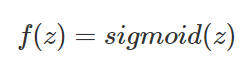
〇確認テスト

シグモイド関数を微分した時、入力値が0の時に最大値をとる。その値として正しいものを選択肢から選べ。

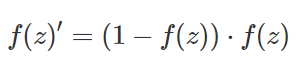


答案：

微分するシグモイド関数を以下とした時、



微分した式は下記となる。



上記式にz=0を代入した結果、微分式の値は0.25となる。

〇確認テスト：

重みの初期値に0を設定すると、どのような問題が発生するか。簡潔に説明せよ。

答案：

重みを0で初期化すると正しい学習が行えない。

全ての重みの値が均一に更新されるため、多数の重みを持つ意味が無くなる。

〇確認テスト

一般的に考えられるバッチ正規化の効果を2点挙げよ。

答案：

1.中間層の重みの更新が安定し、計算速度が向上する。

2.学習データのバラつきが抑えれ、過学習が起こりにくくなる。

●学習率最適化手法

ニューラルネットワークにおいて、誤差を最初にする各パラメータを求める際に勾配勾配法を用いる。本章では、勾配降下法の更新式内の学習率、各パラメータの更新量を値更新毎に変化させ、計算量、計算時間、計算精度を効率化する手法をまなんだ。

〇確認テスト

モーメンタム,AdaGrad,RMSPropの特徴をそれぞれ簡潔に説明せよ。

モーメンタム:

前回計算時の更新量と慣性率の積から誤差を各パラメータで微分した値と学習率の積を引いた値を現在の各パラメータに足し合わせ値を更新する手法

AdaGrad:

誤差を各パラメータで微分した値と、下記の式の様に計算を重ねるとともにあたいが小さくなる学習率との積を現在のパラメータから引き値を更新する手法

RMSProp:

AdaGradと概要は同様の手法であるが、学習率の定義式が下記の様な式となる手法である。

また、学習率の下記定義式について、AdaGradでは、学習率を前回更新時の学習率と誤差の微分値の2乗との単純な和としていたが、本手法では、前述の2つの値の比重を示すパラメーターαを用い、2つの値を比重流動的に操作し学習率を定義できる。

●過学習

ニューラルネットワークの学習において、誤差の大きさとエポック数により描かれる学習曲線が訓練時とテスト時にて乖離する現象を過学習という。その現象の原因は学習するネットワークの自由度(層数、ノード数、パラメータ値)であり、本章では正則化手法を利用してそれを抑制する方法を学んだ。

〇確認テスト

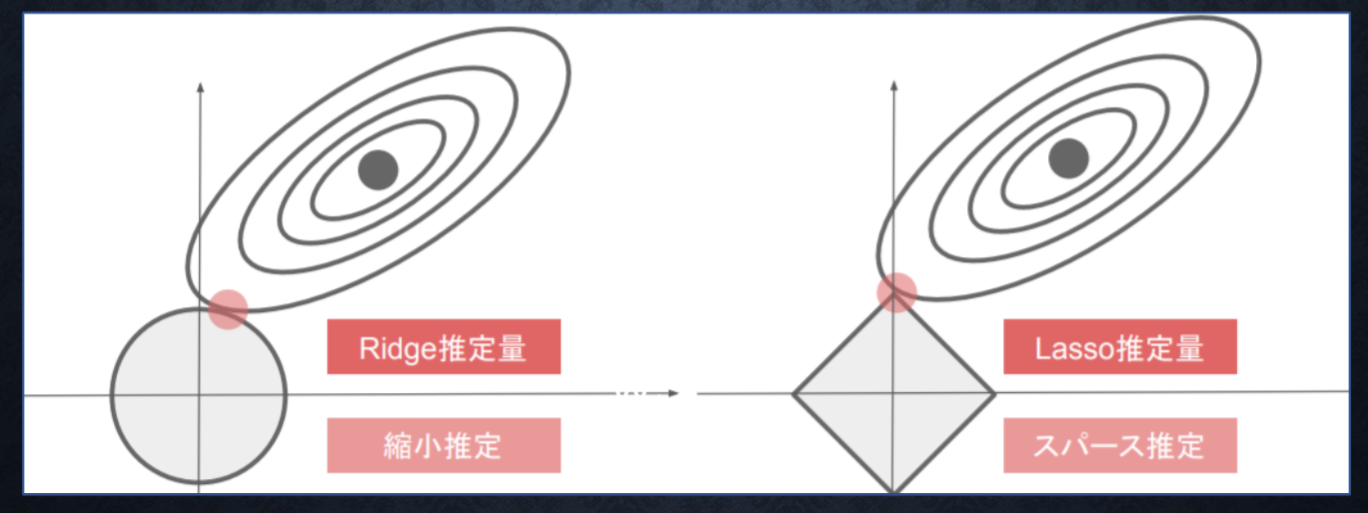
リッジ回帰の特徴として正しいものを選択しなさい。

答案:

(a)ハイパーパラメータを大きな値に設定すると全ての重みが限りなく0に近づく

〇確認テスト

下図においてL1正則化を示しているグラフはどちらか答えよ。



答案:

上の図の左がL1正則化を示している。

●畳み込みニューラルネットワークの概念

畳み込みネットワークとは、画像における空間的情報を取り込み、学習する画像のどの部位にどんな特徴を加味し、カラー画像を例とした複数チャンネルを持つ画像を学習することができるニューラルネットワークである。

　そのCNNにおける各層（畳み込み/プーリング層、全結合層）の概念と計算方法について学習した。

〇確認テスト

サイズ6×6の入力画像を、サイズ2×2のフィルタで畳み込んだ時の出力画像のサイズを答えよ。なおストライドとパディングは1とする。

答案:

出力画像は6 x 6である。

●最新のCNN

本章にて最新のCNNとして学ぶAlexNetとは、2012年にImageNetコンペティションで優勝したトロント大学SuperVisionチームの開発したCNNであり、畳み込み層5層にプーリング層3層という、LeNetと比較するとかなり深い層構造になっている。

またこのCNNでは、過学習対策として、サイズ4096の全結合層の出力において、ドロップアウトを使用している。

