

# 知能システム学特論レポート

(DL2 班) Caffe on Ubuntu

2015 年 6 月 29 日

## 1 報告者

15344203	有田 裕太
15344206	緒形 裕太
15344209	株丹 亮
12104125	宮本 和

## 2 進行状況

- 理論研究
- 順伝播型ネットワークについて

## 3 理論研究

### 3.1 ユニットの出力

### 3.2 活性化関数

### 3.3 多層ネットワーク

Fig. 1 に 2 層構造のネットワークを示す. Fig. 1 (a) より各層を  $l = 0, 1, 2$  とすると,  $l = 1$  の層を入力層,  $l = 2$  を中間層, 隠れ層,  $l = 3$  を出力層と呼ぶ. 各層のユニットの入出力を区別するために, 入力を  $\mathbf{u}^{(l)}$ , 出力を  $\mathbf{z}^{(l)}$  と定義すると, 中間層 ( $l = 2$ ) のユニットの出力は以下の式で表される.

$$\mathbf{u}^{(2)} = \mathbf{W}^{(2)}\mathbf{x} + \mathbf{b}^{(2)} \quad (3.1)$$

$$\mathbf{z}^{(2)} = \mathbf{f}(\mathbf{u}^{(2)}) \quad (3.2)$$

$\mathbf{W}^{(2)}$  は入力層と中間層の結合重みであり,  $\mathbf{b}^{(2)}$  は中間層のユニットに与えられたバイアスである. 同様にして  $\mathbf{u}^{(3)}$ ,  $\mathbf{z}^{(3)}$  は

$$\mathbf{u}^{(3)} = \mathbf{W}^{(3)}\mathbf{z}^{(2)} + \mathbf{b}^{(3)} \quad (3.3)$$

$$\mathbf{z}^{(3)} = \mathbf{f}(\mathbf{u}^{(3)}) \quad (3.4)$$

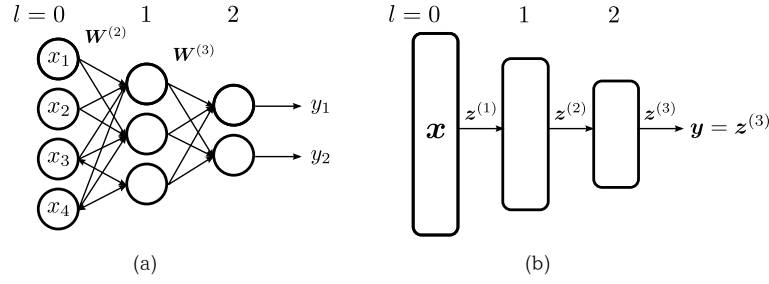


Fig.1 2層のネットワーク

となり、任意の階層  $L$  のネットワークに一般化すると

$$\mathbf{u}^{(l+1)} = \mathbf{W}^{(l+1)} \mathbf{z}^{(l)} + \mathbf{b}^{(l+1)} \quad (3.5)$$

$$\mathbf{z}^{(l+1)} = \mathbf{f}(\mathbf{u}^{(l+1)}) \quad (3.6)$$

と書ける． $l = 1, 2, 3, \dots, L-1$  の順に繰り返していくと最終的な出力  $\mathbf{y}$  を決定することができる．この出力を決定するのは各層間の結合重み  $\mathbf{W}^{(l)}$  ( $l = 2, \dots, L$ ) とユニットのバイアス  $\mathbf{W}^{(l)}$  ( $l = 2, \dots, L$ ) である．これらのパラメータを持つベクトル  $\mathbf{w}$  を定義して， $\mathbf{y}(\mathbf{x}; \mathbf{w})$  と表現する．

### 3.4 出力層の設計と誤差関数

#### 3.4.1 学習の枠組み

順伝播型ネットワークが表現する関数  $\mathbf{y}(\mathbf{x}; \mathbf{w})$  をネットワークのパラメータ  $\mathbf{w}$  を変えることで変化させ、望みの関数を与えることを考える．入力  $\mathbf{x}$  と望みの出力  $\mathbf{d}$  のペアを次のように与える．

$$\{(\mathbf{x}_1, \mathbf{d}_1), (\mathbf{x}_2, \mathbf{d}_2), \dots, (\mathbf{x}_N, \mathbf{d}_N)\} \quad (3.7)$$

これらのペア  $(\mathbf{x}, \mathbf{d})$  1つ1つを訓練サンプル (training samples) といい、その集合を訓練データ (training data) という．ネットワーク  $\mathbf{w}$  を調整することで訓練データの入出力ペアをできるだけ再現すること学習という．

この場合、ネットワークが表す関数と訓練データとの近さ  $(\mathbf{y}(\mathbf{x}_n; \mathbf{w}))$  を誤差関数 (error function) で定義する．誤差関数は問題の種別や活性化関数によって異なる．Tab.1 に問題の種別ごとの活性化関数と誤差関数の一覧を示す．

Tab.1 問題の種別ごとの活性化関数と誤差関数

問題の種別	出力層の活性化関数	誤差関数
回帰	正接双曲線関数や恒等写像	二乗誤差 式 (3.8)
二値分類	ロジスティック関数	式 (3.9)
多クラス分類	ソフトマックス関数	交差エントロピー 式 (3.10)

### 3.4.2 回帰

回帰 (regression) とは出力連続値をとる関数を対象に訓練データを良く再現する関数を求めることをいう。回帰では活性化関数に正接双曲線関数や恒等写像を用い、評価関数は次式が良く用いられる。

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{d}_n - \mathbf{y}(\mathbf{x}_n; \mathbf{w})\|^2 \quad (3.8)$$

### 3.4.3 二値分類

二値分類では入力  $\mathbf{x}$  に応じて2種類に区別する問題を考える。すなわち、 $d \in \{0, 1\}$  とする。このとき、活性化関数はロジスティック関数  $y = 1/(1 + \exp(-u))$  とし、誤差関数は次式で与える。

$$E(\mathbf{w}) = - \sum_{n=1}^N [d_n \log y(\mathbf{x}_n; \mathbf{w}) + (1 - d_n) \log \{1 - y(\mathbf{x}_n; \mathbf{w})\}] \quad (3.9)$$

### 3.4.4 多クラス分類

多クラス分類とは入力  $\mathbf{x}$  に応じて有限個のクラスに分類する問題である。一例として Fig.2 に手書き文字認識の例を示す。この問題では活性化関数にはソフトマックス関数 (softmax function) が良く用いられる。また、誤差関数は次式で与える。

$$E(\mathbf{w}) = - \sum_{n=1}^N \sum_{k=1}^N d_{nk} \log y_k(\mathbf{x}_n; \mathbf{w}) \quad (3.10)$$

なお、この関数は交差エントロピー (cross entropy) と呼ばれる。

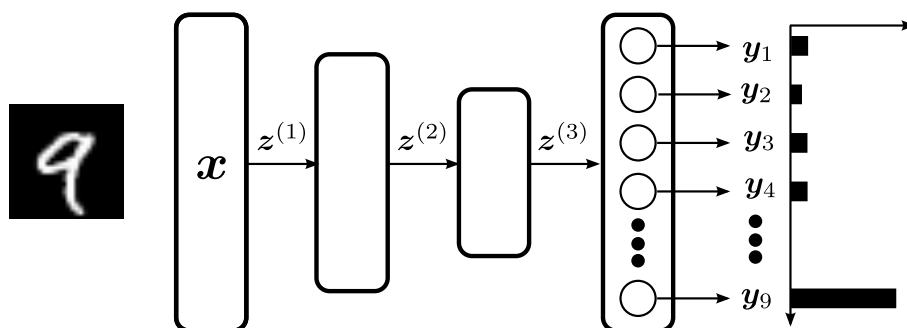


Fig.2 手書き文字認識の例

## 4 今後の課題

- 理論研究を進める。
- Caffe を使いこなす