5.4 損失関数

- 出力と正解の誤差を定義する関数が損失関数.
- 誤差は，あるべき状態との隔離の度合い.
- 誤差の値が大きければ，ニューラルネットワークが望ましい状態から離れていることになる.
- 学習は，この誤差を最小化するようにおこなわれる.
- 損失関数には様々な種類があるが，ディープラーニングでは一般的に二乗和誤差もしくは交差エントロピー誤差がよく用いられる.

5.4.1 二乗和誤差

- 出力値と正解値の差を二乗し，すべての出力層のニューロンで総和をとったものを二乗和誤差と呼ぶ.
- 二乗和誤差は，E を誤差，y_k を出力層の各出力値，t_k を正解値として以下の式で定義される：

$$E = \frac{1}{2} \sum_k (y_k - t_k)^2.$$

- y_k と t_k の差を二乗し，すべての出力層のニューロンで総和をとり$1/2$ をかける.
- $1/2$ をかけるのは微分の際に扱いやすくするため.
- 二乗和誤差を用いることにより，ニューラルネットワークの出力がどの程度正解と一致しているかを定量化することができる.
- 二乗和誤差は，正解や出力が連続的な数値であるケースに向いているため，回帰問題でよく使用される.

5.4.2 交差エントロピー誤差

- 交差エントロピー誤差は，2つの分布の間のズレを表す尺度で，分類問題でよく使用される.
- 交差エントロピー誤差は，次の式のように，出力 y_k の自然対数と正解値の積の総和を，マイナスにしたもので表される：

$$E = -\sum_k t_k(\log(y_k)).$$

- 分類問題における正解値は，1が1つで残りがすべて0となる.
- したがって，右辺の $\sum_k$ 内で t_k が1の項のみ影響を与えることになり，t_k が0の項の影響は無視される.
- その結果，正解値が1のたった1つの項しか誤差に影響を与えないことになる.

- $-\log(y_k)$ は正解に近づくほど小さくなり，正解から離れるほど誤差がどこまでも大きくなる.
- つまり，上の式は，出力が正解から離れるほど誤差がどこまでも大きくなり，出力が正解に近づくほど誤差が0に近づくことを　意味している.