
IoT eラーニング

ニューラルネットワーク
(概要、仕組み)

国立大学法人 琉球大学

目次

- ニューラルネットワークの概要
 - ニューラルネットワークとは
 - ニューラルネットワークの学習
 - ニューラルネットワークとパーセプトロン
- ニューラルネットワークの仕組み
 - ニューロンの仕組み
 - ニューラルネットワークの仕組み
 - ニューラルネットワークの学習

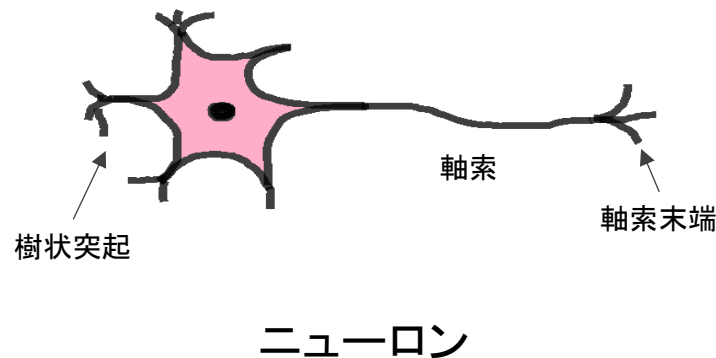
ニューラルネットワークの概要

● ニューラルネットワークとは

動物の脳はニューロンと呼ばれる神経細胞が集まり、ニューロン同士が軸索でシナプス結合して構成されている。

ニューロンは電気信号を樹状突起から軸索末端に軸索に沿って伝達する。軸索末端で他のニューロンとシナプス結合しているとそのニューロンに電気信号が受け取られる。

ニューラルネットワークは、ニューロンをモデル化した人工ニューロンで多層に結合し構成されたネットワークである。



ニューラルネットワークの概要

● ニューラルネットワークの学習

ニューラルネットワークの学習には次の2つがある。

- ・ 教師あり学習：教師データ（正解）の入力によって最適化されていく学習
- ・ 教師なし学習：教師データを必要としない学習

ニューラルネットワークの学習とは、ネットワークとして出力された（出力層からの出力）結果が人間が望む結果になるようパラメータを調整する作業になる。

ニューラルネットワークが学習する際に用いられるアルゴリズムは誤差逆伝播法である。誤差逆伝播法（バックプロパゲーション）は、損失関数の微分を効率よく計算する手法の1つである。

ニューラルネットワークの概要

● ニューラルネットワークとパーセプトロン

1958年にフランク・ローゼンブラットがパーセプトロンを提案した。
これは1943に提案された形式ニューロンをもとにした考えである。

形式ニューロンは、入力が3つある場合、入力それぞれに対して重みを掛け合わせた時、その結果が閾値を超えると1を出力し、越えなければ0を出力するという考えである。
脳などの神経細胞が発火するかしないかの2パターンの動きと似ている。
ローゼンブラットは形式ニューロンを横並びに並べ、出力ニューロンで組み合わせる2層の構造をとり、入力と出力の組み合わせを学習することで正しい解答をすることを示した。

パーセプトロンの考え方が受け入れられ広く研究されるようになると、教師あり学習の線形分離可能問題しか解けないことがわかり廃れていった。

しかし、1986年にラメルハートが誤差逆伝播法（バックプロパゲーション）を提唱したことで、非線形分離可能な問題も解けるようになった。
誤差逆伝播法（バックプロパゲーション）の登場で、ニューラルネットワークは再び注目されることになった。

ニューラルネットワークの仕組み

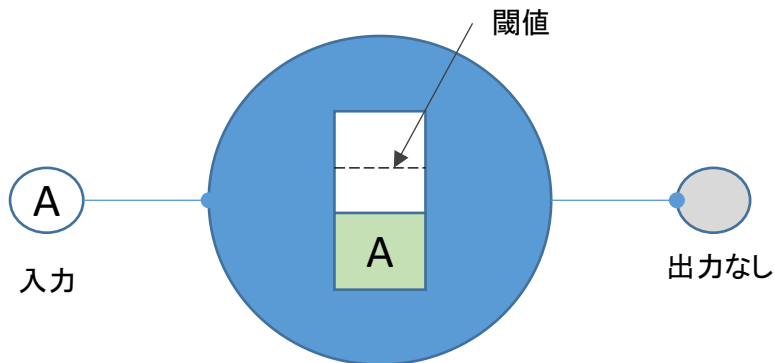
● ニューロンの仕組み

ニューロンを簡単なモデルにすると、樹状突起にあたる入力と軸索末端にあたる出力を持った信号スイッチになる。このスイッチは入力値が一定の値（閾値）を超えないと出力しない構造となっている。

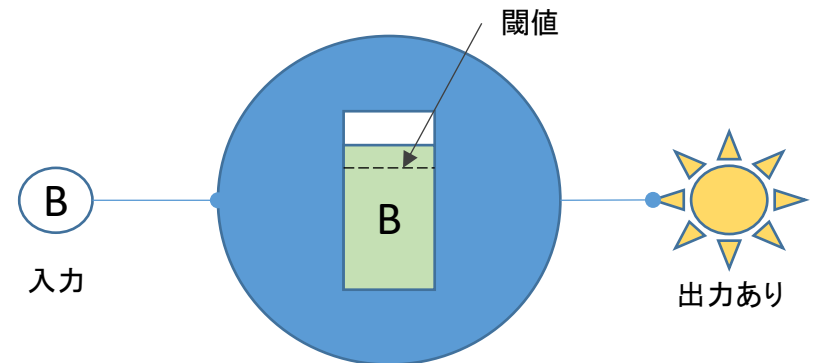
下図の①では入力値Aは閾値を超える値ではないので何も出力していない。

②では入力値Bは閾値を超える値であるので出力された。

このように、ニューロンには入力があっても閾値による出力の制御が行われている。



①ニューロン値が閾値を超えない場合



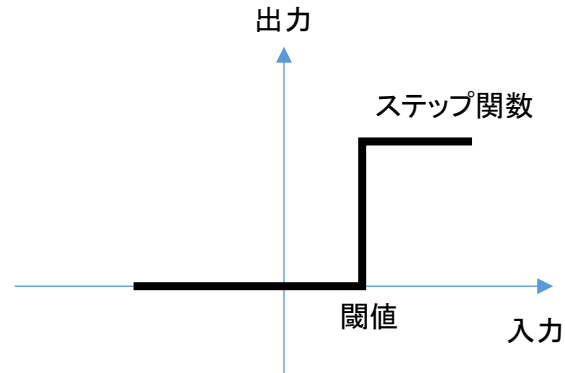
②ニューロン値が閾値を超えた場合

ニューラルネットワークの仕組み

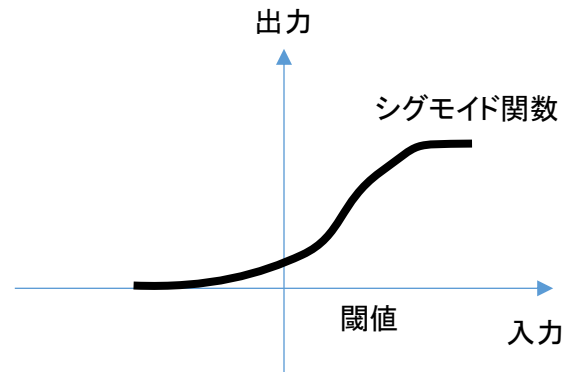
閾値による出力の制御で使われる関数を活性化関数と呼ぶ。

例で示したモデルでは、低い入力値の場合は出力はないが入力値が閾値を超えると出力が発生する。

グラフにすると階段のようになることからステップ関数と呼ばれており、活性化関数の1つとなる。



ステップ関数では出力が跳ね上がる挙動となるが、出力が滑らかな曲線を描く関数にシグモイド関数がある。人工ニューロンのモデル化では出力を制御する関数としてはさまざまあるが、ここではシグモイド関数を使用する。

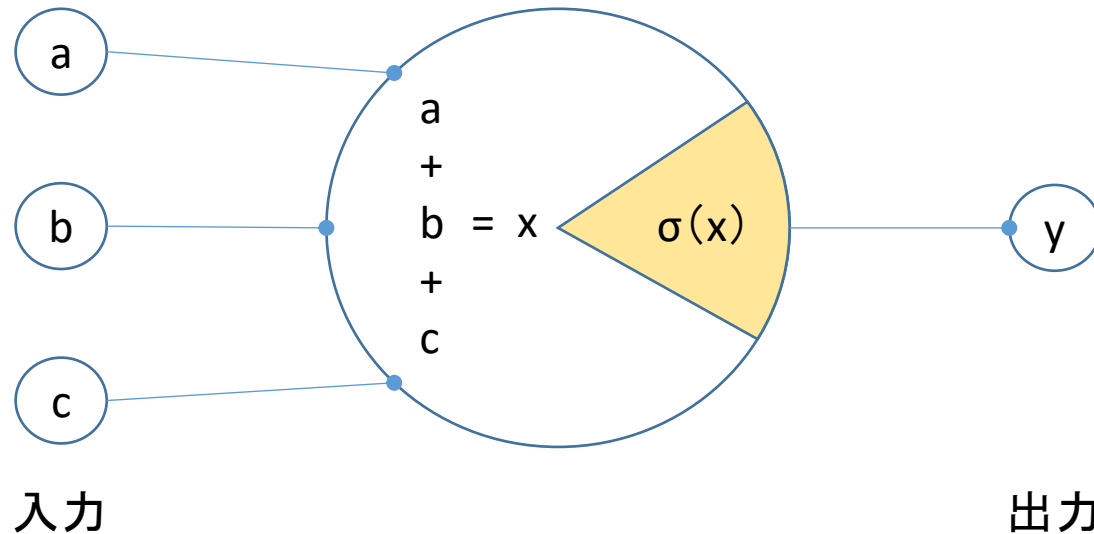


ニューラルネットワークの仕組み

シグモイド関数は数式にすると次のようになる。eはネイピア数と呼ばれる数学定数である。自然対数の底であり、2.71828…と続く超越数である。

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

ニューロンの入力はいくつだけではなく複数の入力を持っているため、各入力値の合計を人工ニューロンが受けた入力値とし、この入力値による出力値の演算をシグモイド関数でモデル化すると下図のようになる。

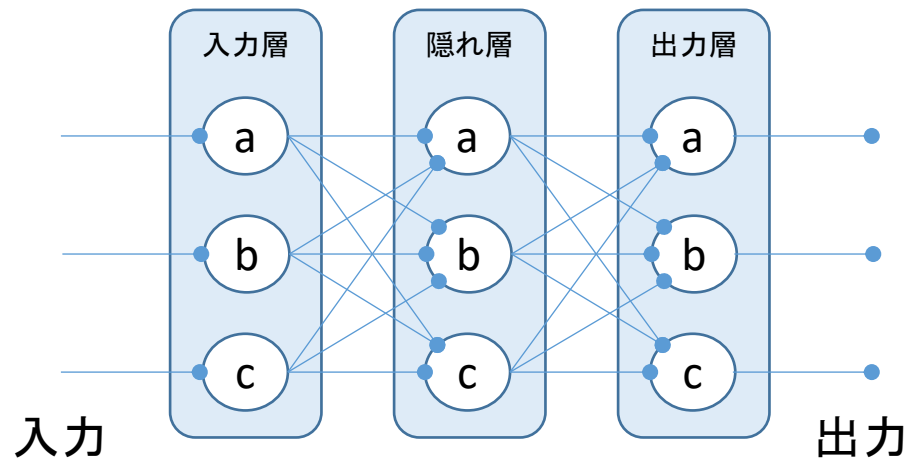


ニューラルネットワークの仕組み

● ニューラルネットワークの仕組み

動物の脳はニューロンが多層に接続されており、多くのニューロンから電気信号を受け取り発火した場合には多くのニューロンに信号を伝達している。

まず、1層に複数のニューロン（ノード）を持ち、層は3層（入力層、隠れ層、出力層）で構成されている構造をモデル化すると下図のようになる。



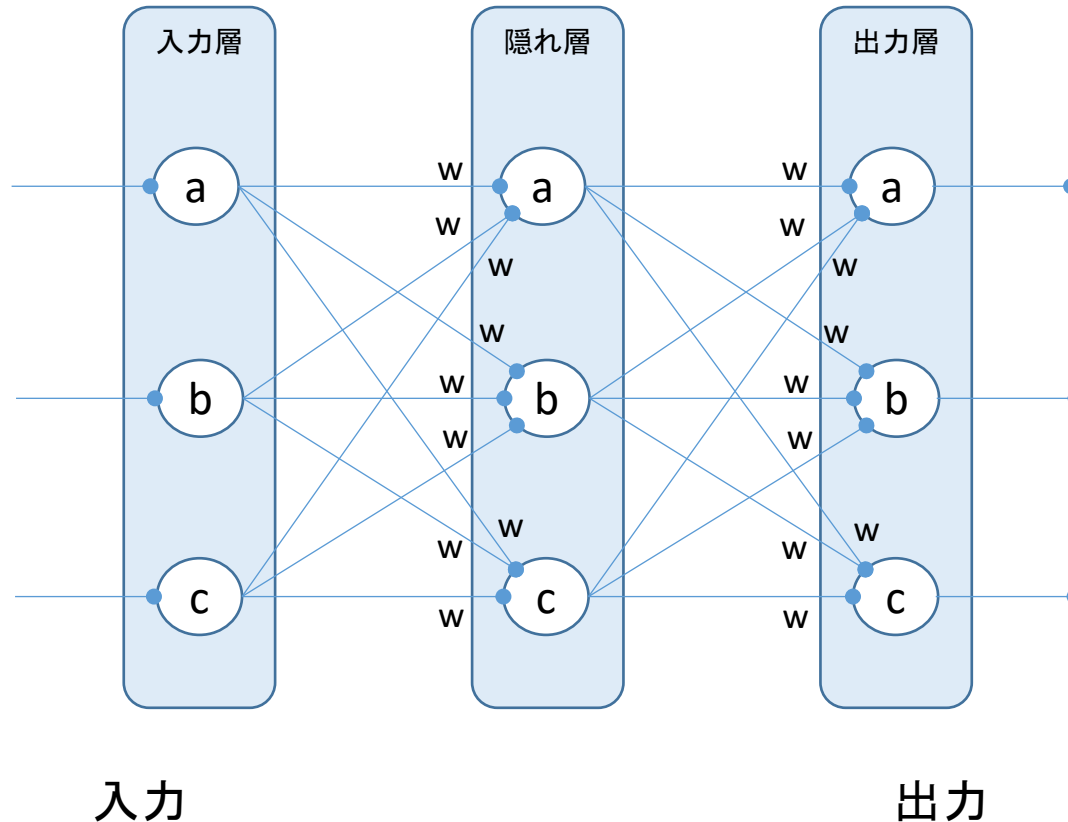
しかし、これでは学習などによる調整を行うパラメータがない。

これを解消するには、ノード間の接続の強さを調整することでネットワークの出力を変化させる必要がある。対応として、各ノードを接続している接続ごとに「重み」を持たせることで調整する。

ニューラルネットワークの仕組み

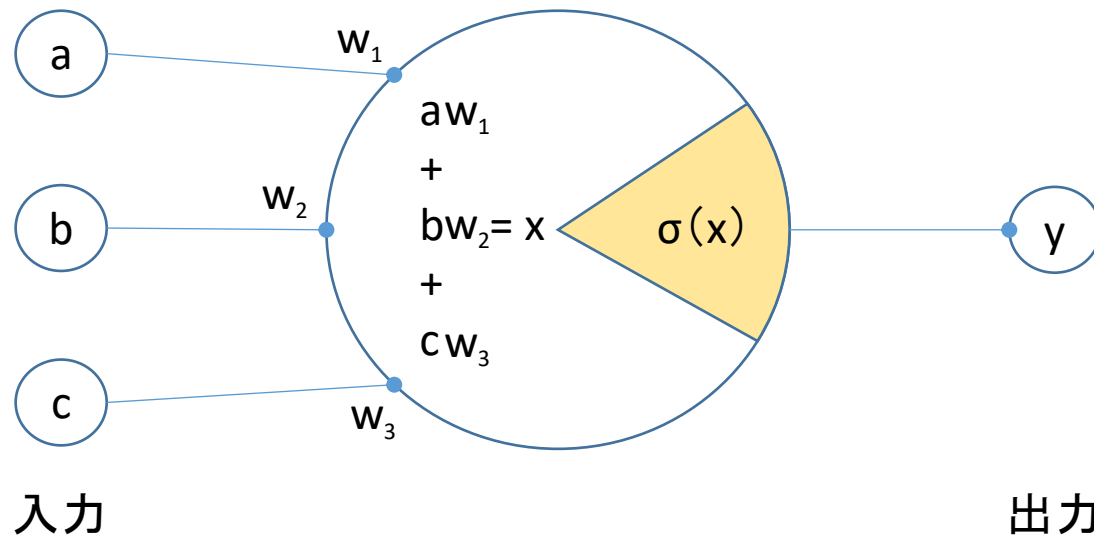
低い重みの場合、信号を減少させる。高い重みの場合、信号を増幅させる。これにより、重みの調整でネットワークの出力を調整できるようにする。

これをモデルにすると下図のようになる。重みは w で表している。



ニューラルネットワークの仕組み

ニューロンモデルに重みを追加すると下図のようになる。
入力元からの入力値に対して重みを掛け合わせることで接続の強弱を表す。ノードの入力値としては各接続の入力値に重みを掛け合わせた値の合計値となる。

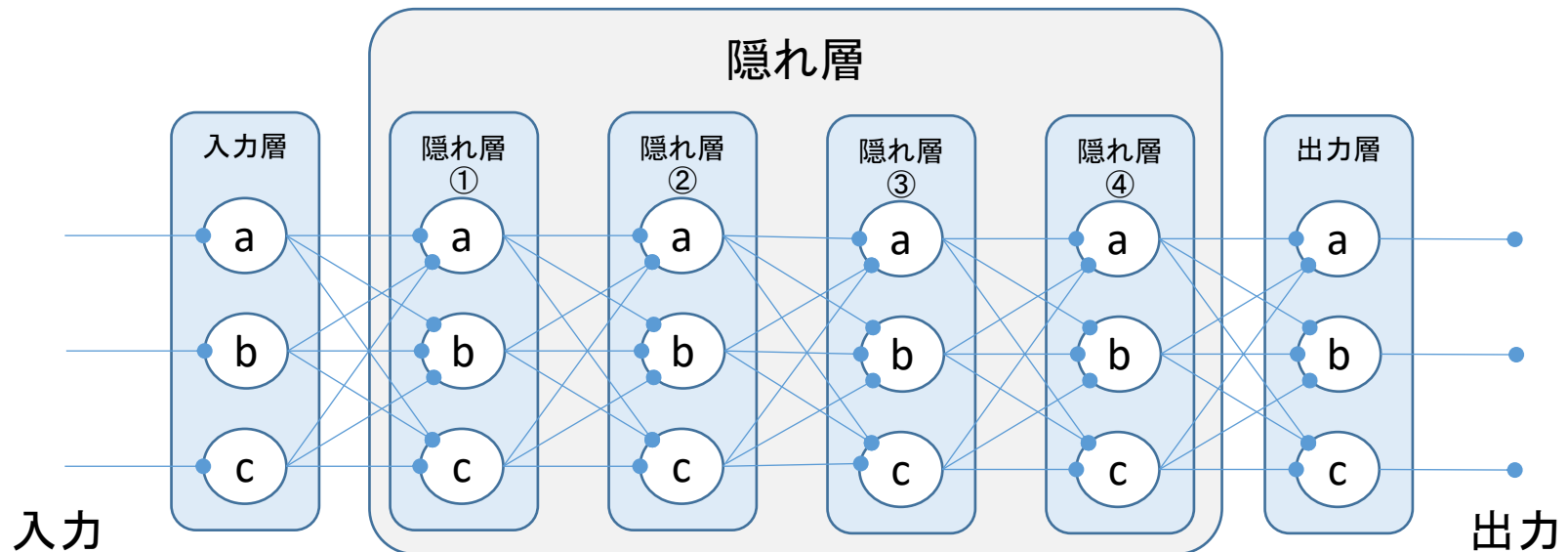


ニューラルネットワークの仕組み

ニューラルネットワークをモデル化した際に3層でモデル化したが、動物などの脳は多層となっている。

実際には、ニューラルネットワークでも隠れ層を複数持たせ、またノード接続の重みの調整（学習）により、ネットワークに入力された問題に対する解答を出力する。

隠れ層が複数になった例を下図に示す。

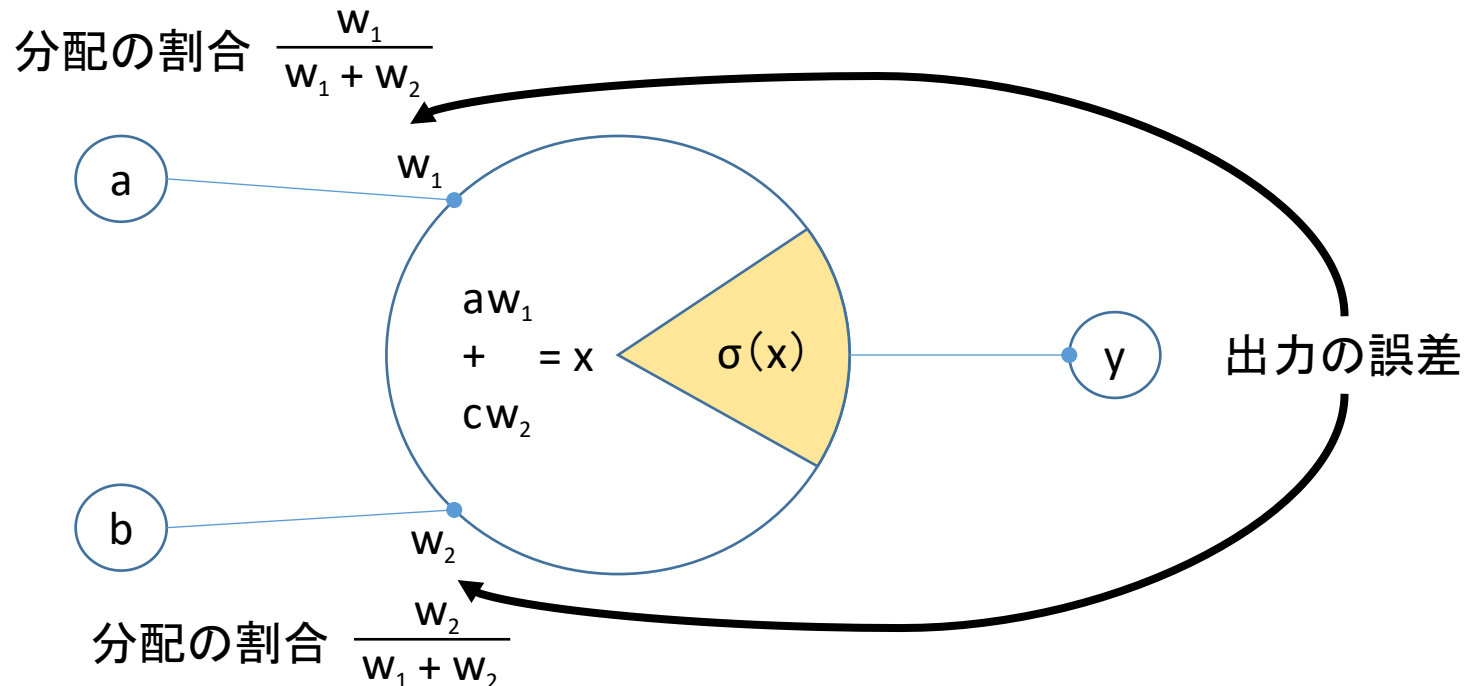


ニューラルネットワークの仕組み

● ニューラルネットワークの学習

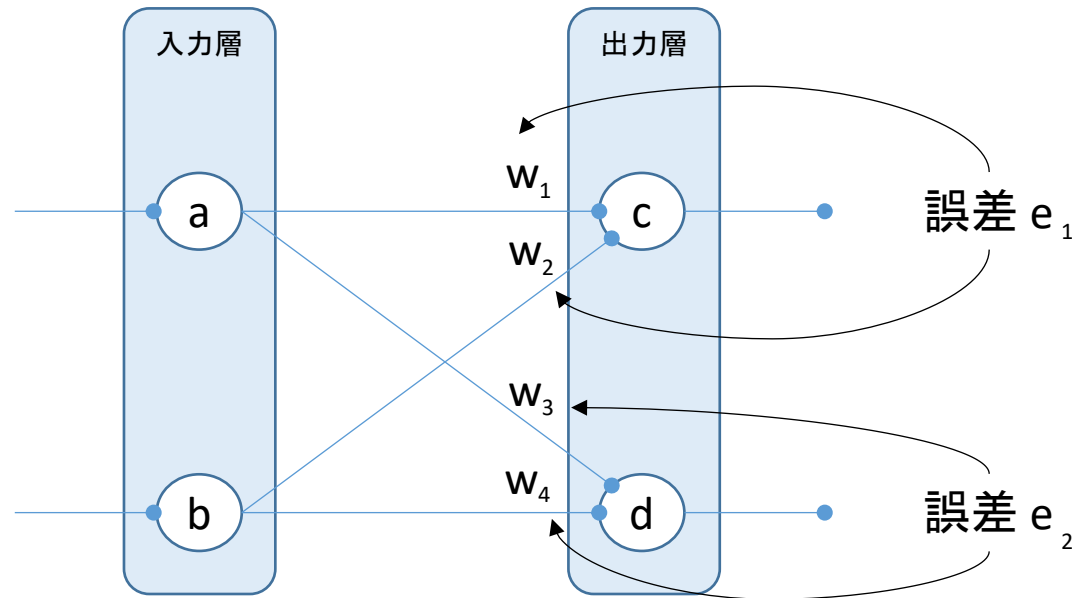
ニューラルネットワークが出力した結果に対し、人間が望む結果との差が誤差になる。この誤差を改善する方法として、ノードの入力値の重みに誤差を伝達することにより調整する方法がある。

ノードに2つの接続がある場合、出力で発生した誤差をどのように分配するかにも方法があり、均等に分配する方法や重みの比率で分配する方法などがある。下図では重みによる比率で誤差を分配したイメージになる。



ニューラルネットワークの仕組み

複数のノードを持つ2層の場合のイメージは下図のようになる。

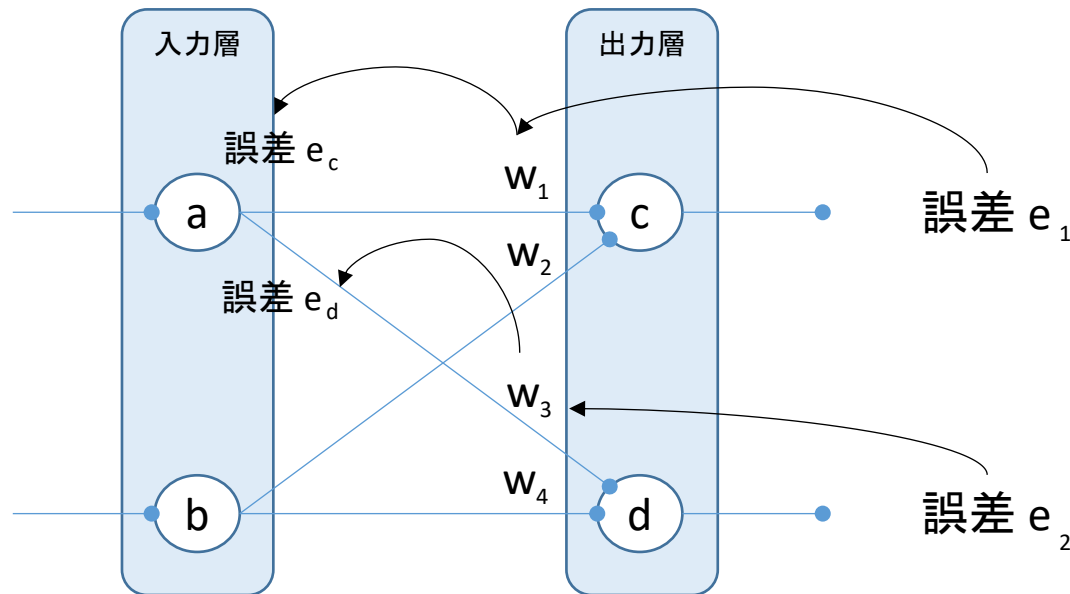


出力結果から出力層への誤差の伝播は、複数のノードがあったとしても、ノードごとの出力値に対する誤差を自身に接続している重みに分配することになる。

これは、ある出力ノード接続しているリンクは他の出力ノードへ影響を与えないからである。

ニューラルネットワークの仕組み

出力層で分配された誤差を入力層のノードに伝播するイメージは下図のようになる。



入力層のaノードには出力層のcノードとdノードへの接続がある。

このことから、aノードには、

- ◆ cノードの重み w_1 に伝播した誤差 e_c
- ◆ dノードの重み w_3 に伝播した誤差 e_d

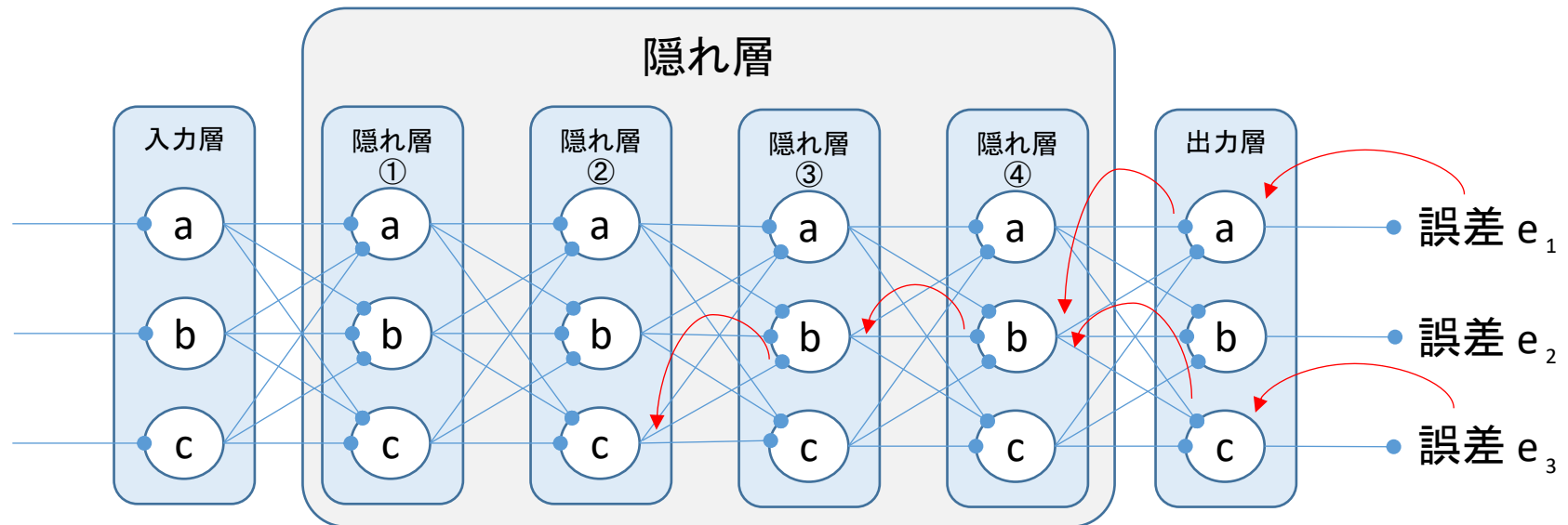
が伝播されてくることになり、aノードの誤差はこの合計と考えることができる。

$$\text{aノードの誤差 } e_c = e_1 \frac{w_1}{w_1 + w_2} + e_2 \frac{w_3}{w_3 + w_4}$$

ニューラルネットワークの仕組み

ニューラルネットワークのモデルでは、隠れ層に複数の層を持たせている。

この場合も層毎に誤差の伝播を確定し次の層へと伝播をしていくことで、全体の重みに誤差を伝播させることができる。



ニューラルネットワークの仕組み

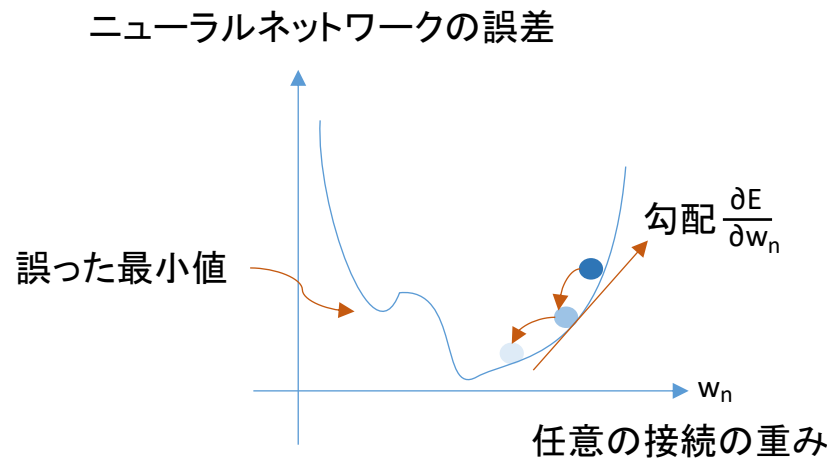
重みに対して誤差の伝播ができたので、この誤差を使い重みを更新する必要がある。

ここでは勾配降下法を利用した重みの計算を使用する。

次のグラフでは、ニューラルネットワークが出力する誤差に対して重みを勾配で調整するイメージである。

勾配が正の値の場合は重みの値を減らせば誤差が小さくなり、負の場合だと増やせば小さくなる。

また、誤った最小値に落ち込んだままとならないよう考慮する必要がある。



ニューラルネットワークの仕組み

先のグラフのイメージから重みの更新に使用する式を表すとすると次のような式になる。

$$\text{新しい}w_n = \text{古い}w_n - \alpha \frac{\partial E}{\partial w_n}$$

α は誤った最小値へ落ち込まないようにする学習率になる。この値を調整することで、誤った最小値への落ち込み回避を図るようにする。

このように、ニューラルネットワークでは、出力された結果において発生した誤差を、出力層から入力層に向かって誤差を伝播することにより重みを調整し、最適な答えを導くための学習を行う。

またこの方法を誤差を遡って伝えていくことから、誤差逆伝播法（バックプロパゲーション）と呼ばれている。