

5.ニューラルネットワーク(Part.1)

序論

脳のはたらきに例えなくても、ニューラルネットワークを導入することは可能です。線形分類の項では、画像に含まれるさまざまな視覚的カテゴリーの判別スコアを、

$s = Wx$ (ただし、 W は行列で、 x は、画像のすべてのピクセルデータを含む入力列ベクトル)

CIFAR-10の場合、 x は $[3072 \times 1]$ の列ベクトルであり、 W は $[10 \times 3072]$ の行列であり、出力スコアは10個の分類スコアのベクトルとなります。

ニューラルネットワークの例では、線形分類の時とは変わり、 $s = W_2 \max(0, W_1 x)$ を計算します。

ここで、 W_1 は、例えば、画像を100次元の中間ベクトルに変換する $[100 \times 3072]$ 行列などが考えられます。関数 $\max(0, -)$ は、要素ごとに適用される非線形性です。

非線形性にはいくつかの選択肢がありますが（以下で検討します）、この選択肢は一般的なもので、単純に0以下のアクティベーションをすべて0にきき値化します。最後に、行列 W_2 は $[10 \times 100]$ の大きさになり、クラススコアと解釈される10個の数字が再び得られます。この非線形性は、計算上非常に重要であることに注意してください。もし、この非線形性を残しておけば、2つの行列は1つの行列に折りたたまれ、予測されるクラススコアは再び入力の線形関数となります。非線形性は、私たちが微動だにしないところです。パラメータ W_2, W_1 は確率的勾配降下法で学習され、その勾配は連鎖律で導かれます（バックプロパゲーションで計算されます）。

3層構造のニューラルネットワークは、数式で表現するならば $s = W_3 \max(0, W_2 \max(0, W_1 x))$ のようになり、 W_3, W_2, W_1 のすべてが学習されるパラメータとなります。中間の隠れベクトルの大きさは、ネットワークのハイパーパラメータで、その設定方法は後ほど説明します。では、これらの計算をニューロンやネットワークの観点からどのように解釈するかを考えてみましょう。

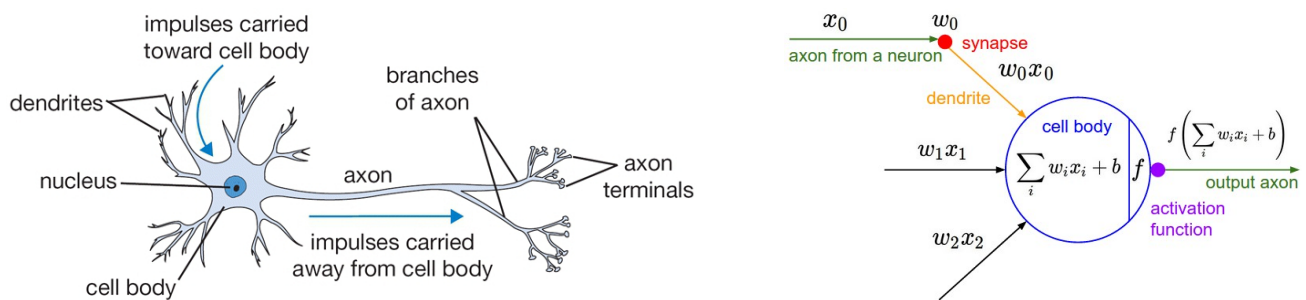
ニューロンのモデリング

ニューラルネットワークの分野は、元々は生物学的な神経システムをモデル化するという目的に触発されていましたが、その後、エンジニアリングの問題となり、機械学習のタスクで良い結果を出すことができるようになりました。それでも、この分野の大部分が触発されてきた生物学的システムについて、非常に簡単にハイレベルな説明から議論を始めます。

生物学的刺激とつながり

脳が構成されている基本的な単位はニューロンです。人間の神経系には約860億個のニューロンが存在し、それらは約 10^{14} ~ 10^{15} 個のシナプスで接続されています。下の図は、生物学的なニューロンの概略図（左）と、一般的な数学的モデル（右）を示しています。各ニューロンは、樹状突起から入力信号を受け取り、その（単一の）軸索に沿って出力信号を生成します。軸索は最終的に分岐し、シナプスを介して他のニューロンの樹状突起に接続される。ニューロンの計算モデルでは、軸索を伝わる信号（例えば x_0 ）は、そのシナプスでのシナプス強度（例えば w_0 ）に基づいて、相手のニューロンの樹状突起と乗算的に相互作用をします。（例えば $w_0 x_0$ ）。

考え方としては、シナプスの強度（重み w ）は学習可能であり、あるニューロンが他のニューロンに与える影響の強さ（およびその方向：興奮性（正の重み）または抑制性（負の重み））を制御するという考え方です。基本的なモデルでは、樹状突起が信号を細胞体に伝え、そこですべての信号が合計されます。最終的な和がある閾値を超えると、そのニューロンは発火し、軸索に沿ってスパイクを送ることができます。計算モデルでは、スパイクの正確なタイミングは重要ではなく、発火の頻度のみが情報を伝達すると仮定しています。このレートコードの解釈に基づいて、ニューロンの発火率を、軸索に沿ったスパイクの頻度を表す活性化関数 f でモデル化します。これまで、活性化関数の一般的に用いられている関数はシグモイド関数 σ です。これは、実数値の入力（和の後の信号強度）を受け取り、0と1の間の範囲になるようにそれを正規化するからです。これらの活性化関数の詳細については、このセクションの後半で見えていきます。



生物学的なニューロン（左）とその数学的モデル（右）を図式した