階層型のネットワーク

浅川伸一 <asakawa@twcu.ac.jp>

パーセプトロンに代表されるフィードフォワード型の結合を持つ階層型のネットワークは、パターン認識 (pattern recognition)、情報圧縮 (data compression)、運動制御 (motion control)、雑音除去 (noise reduction)、および時系列予測 (time series prediction) などへの理論的、または応用的研究が試みられている。ここでは、パーセプトロンの学習について、幾何学的表現を用いて詳しく解説し、線形分離可能な問題について紹介する。続いてバックプロパゲーション法を導入し、階層型のネットワークについての幾つかの話題を解説する。

1 階層型ネットワークの行列表現

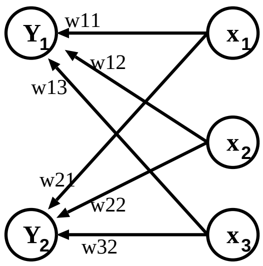

図 1: ネットワークの行列表現

図 1 には、 3 つの入力ユニットと 2 つの出力ユニットの活性値 x_1, x_2, x_3 と y_1, y_2 および入力ユニットと出力ユニットの結合係数を表す $w_{11}, w_{12}, \cdots,$ w_{32} が示されている。これらの記号をベクトル $\boldsymbol{x}, \boldsymbol{y}$ と行列 $\boldsymbol{W}$ を使って表すと $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}$ となる。図 1 の場合、ベクトルと行列の各要素を書き下せば、

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{1}$$

のようになる。

行列の積は、左側の行列の i 行目の各要素と右側の行列（ベクトルは 1 列の行列でもある）の i 列目の各要素とを掛け合わせて合計することなので、

以下のような、加算記号を用いた表記と同じことである。

$$\begin{array}{lllll} y_1 & = & w_{11}x_1 + w_{12}x_2 + w_{13}x_3 & = & \sum_i w_{1i}x_i \\ y_2 & = & w_{21}x_1 + w_{22}x_2 + w_{23}x_3 & = & \sum_i w_{2i}x_i \end{array} \tag{2}$$

これを、m 個の入力ユニットと n 個の出力ユニットの場合に一般化すれば、

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{3}$$

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x} \tag{4}$$

と表現できる。しきい値の扱いについては、常に 1 を出力する仮想的なユニット $x_0 = 1$ を考えて W に組み込むことが多い。

実際の出力は $\boldsymbol{y}$ の各要素に対して

$$f(y) = \frac{1}{1 + e^{-y}} \tag{5}$$

のような非線型変換を施すことが行なわれる。

階層型のネットワークにとっては、(5) 式の非線型変換が本質的な役割を果たす。なぜならば、こうした非線形変換がなされない場合には、ネットワークの構造が何層になったとしても、この単純なシナプス結合係数を表す行列を $\boldsymbol{W}_i$ ($i = 1, \cdots, p$) としたとき、$\boldsymbol{W} = \prod_{i=1}^{p} \boldsymbol{W}_i$ と置くことによって本質的には 1 層のネットワークと等価になるからである。

$$\boldsymbol{y} = \boldsymbol{W}_p \boldsymbol{W}_{p-1} \cdots \boldsymbol{W}_1 \boldsymbol{y} = \left(\prod_{i=1}^{p} \boldsymbol{W}_i \right) \boldsymbol{y} \tag{6}$$