確率的非巡回型神経回路網のための連結相関学習法

5H－4

角田　有希　　三谷　光照　　大堀　隆文　　渡辺　一央
北海道工業大学　電気工学科

1　まえがき

ゆらぎをともなう線形しきい値ニューロン群からなる確率的非巡回型神経回路網（$PFN, Fig.1$）の動作を解析し，新たな学習法として連結相関学習法を提案する．また，簡単な学習課題に適用した結果について報告する．

2　ニューロンモデル

PFNを構成する各ニューロン k（$Fig.2$）は，独立なゆらぎ源 $n_k(t)$ を有し，情報伝達諸量が入力パターン p によって変化するが，一定入力に対してもゆらいでいる，ことを除いては McCulloch-Pitts のニューロンモデルと同じである．ニューロンkの入力パターンpに対する入出力関係は，以下のように記述できる．ただし，$\langle\cdot\rangle$:期待値，$f\{\cdot\}$:線形しきい値関数，^{T}:転置を表し，定義されたベクトルは列ベクトルとする．

シナプス前ニューロン群出力：$\mathbf{q}_k(p,t) \in \{0,1\}$　(1)

入力総和　：$u_k(p,t) = \mathbf{w}_k^T \mathbf{q}_k(p,t)$　(2)

瞬時活性値　：$v_k(p,t) = u_k(p,t) + n_k(t)$　(3)

瞬時出力値　：$r_k(p,t) = f\{v_k(p,t)\}$　(4)

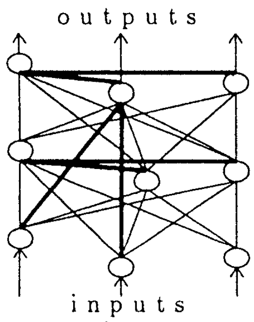

Fig.1
Probabilictic
Feedforward NN

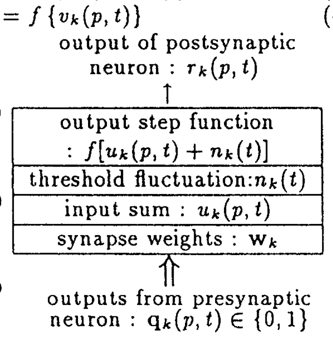

Fig.2
Model of a neuron k
for PFN

3　情報伝達諸量の確率的表現

PFNの情報伝達諸量について離散時間確率系で考察する．ある入力パターンpの入力継続期間におけるニューロンkの伝達諸量 $\mathbf{q}_k(p,t)$, $u_k(p,t)$, $n_k(t)$, $v_k(p,t)$, $r_k(p,t)$ を確率変数として $\mathbf{q}_k$, u_k, n_k, v_k, r_kで表し，各確率変数の分布を $Q(\mathbf{q}_k)$, $U(u_k)$, $N(n_k)$, $V(v_k)$, $R(r_k)$ で表す．このとき伝達諸量の確率分布は，以下の特徴，関係をもつ．以下，時間変数 t，パターン変数 p を省略する．

$$V(v_k) = \textstyle\sum_{\mathbf{q}_k} Q(\mathbf{q}_k) N(v_k - \mathbf{w}_k^T \mathbf{q}_k) \quad (5)$$

$$R(r_k = 1) = \textstyle\sum_{\mathbf{q}_k} Q(\mathbf{q}_k) \int_0^\infty N(v_k - \mathbf{w}_k^T \mathbf{q}_k)\, dv_k \quad (6)$$

$$R(r_k = 0) = 1 - R(r_k = 1) \quad (7)$$

ニューロン出力 r_kは後続のニューロンjの入力ベクトルq_jを構成し，$S(j)$ をjのシナプス前ニューロン群の集合とすると，その確率分布 $Q(\mathbf{q}_j)$ は，次式で表される．

$$Q(\mathbf{q}_j) = Pr\{r_k : k \in S(j)\} \quad (8)$$

一方，瞬時出力誤差 e は，

$$e = \textstyle\sum_{k \in S_{v0}} r_k + \sum_{k \in S_{v1}} (1 - r_k) \quad (9)$$

と表せる．ただし，S_{v0}, S_{v1}は，教師信号が0または1になる可視ニューロン群の部分集合を表す．

4　連結相関学習法

4.1 学習原則

あるパターン入力継続期間における瞬時出力誤差 e の期待値 $\langle e\rangle$ の結合係数 $\mathbf{w}_k$に関する最急降下法によって最小化する．結合係数の修正量$\Delta_p \mathbf{w}_k$は，次式で与えられる．

$$\Delta_p \mathbf{w}_k^T = -\epsilon \frac{\partial\langle e\rangle}{\partial \mathbf{w}_k} = -\epsilon \textstyle\sum_{\mathbf{q}_k} Q(\mathbf{q}_k) \dfrac{\partial\langle e:\mathbf{q}_k\rangle}{\partial \mathbf{w}_k} \quad (10)$$

ただし，$\langle\cdot:\mathbf{q}_k\rangle$ は $\mathbf{q}_k$下での期待値を表す．

4.2 誤差の期待値 $\langle e\rangle$ の結合係数依存性

式 (10) の偏微分量$\partial\langle e:\mathbf{q}_k\rangle/\partial \mathbf{w}_k$ を次式で表す．

$$\frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial \mathbf{w}_k} = \frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial\langle v_k:\mathbf{q}_k\rangle} \cdot \frac{\partial\langle v_k:\mathbf{q}_k\rangle}{\partial \mathbf{w}_k} \quad (11)$$

ゆらぎに，条件1:$\langle n_k\rangle = 0$, を導入すると，式(3)より式(11)は，

$$\frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial \mathbf{w}_k} = \frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial\langle v_k:\mathbf{q}_k\rangle} \cdot \mathbf{q}_k^T \quad (12)$$

となる．ここで,$\langle e:\mathbf{q}_k\rangle$ を r_k下での誤差期待値として表すと，

$$\langle e:\mathbf{q}_k\rangle = R(r_k = 0 \mid \mathbf{q}_k)\langle e:(r_k = 0:\mathbf{q}_k)\rangle + R(r_k = 1 \mid \mathbf{q}_k)\langle e:(r_k = 1:\mathbf{q}_k)\rangle \quad (13)$$

となる．また,$\langle e:(r_k:\mathbf{q}_k)\rangle$ は r_kの状態によってのみ変化し，$R(r_k \mid \mathbf{q}_k)$ は式 (6),(7) の関係をもつ，ことから，式 (12) の偏微分量$\partial\langle e:\mathbf{q}_k\rangle/\partial\langle v_k:\mathbf{q}_k\rangle$ は次式となる．

$$\frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial\langle v_k:\mathbf{q}_k\rangle} = \{\langle e:(r_k = 1:\mathbf{q}_k)\rangle - \langle e:(r_k = 0:\mathbf{q}_k)\rangle\}\, N(-\langle v_k:\mathbf{q}_k\rangle) \quad (14)$$

さらに,$r_k = \{0,1\}$ は $\{v_k < 0, v_k \geq 0\}$ に対応し，式 (14) の誤差期待値を確率的に表すと，次式となる．

$$\frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial\langle v_k:\mathbf{q}_k\rangle} = \textstyle\sum_e e\,\{Pr(e \mid (v_k \geq 0 \mid \mathbf{q}_k)) - Pr(e \mid (v_k < 0 \mid \mathbf{q}_k))\}\, N(-\langle v_k:\mathbf{q}_k\rangle) \quad (15)$$

ここで $Pr(e \mid (v_k \geq 0 \mid \mathbf{q}_k))$, $Pr(e \mid (v_k < 0 \mid \mathbf{q}_k))$ は,$v_k \geq 0$,$v_k < 0$ それぞれの範囲内で一定である．またゆらぎに，条件 2:$N(x)$ は媒介変数 $g(x)$ を介して，任意の x_0で，

$$N(x_0) = \int_0^\infty g(x + x_0) N(x + x_0)\, dx \quad (16)$$

を満たす，を導入すると，式 (15) は次式となる．

$$\frac{\partial\langle e:\mathbf{q}_k\rangle}{\partial\langle v_k:\mathbf{q}_k\rangle} = \textstyle\sum_e \int_{-\infty}^{\infty} e \cdot g(n_k) \cdot Pr(e \mid (v_k \mid \mathbf{q}_k)) N(n_k)\, dn_k \quad (17)$$

よって，式(10) の偏微分量$\partial\langle e\rangle/\partial \mathbf{w}_k$は，式 (12),(17) を式(10) に代入することにより次式を得る．

$$\frac{\partial\langle e\rangle}{\partial \mathbf{w}_k} = \textstyle\sum_{\mathbf{q}_k} \sum_e \int_{-\infty}^{\infty} e \cdot g(n_k) \cdot \mathbf{q}_k^T Pr(e, n_k, \mathbf{q}_k)\, dn_k = \langle e \cdot g(n_k) \cdot \mathbf{q}_k\rangle \quad (18)$$

4.3 学習則 (結合係数修正式)

$N(n_k)$ が標準正規分布のとき，式 (16) から $g(n_k) = n_k$を得る．よって，式 (18) より以下の簡潔な学習則を得る．

$$\Delta_p \mathbf{w}_k = -\epsilon\langle e \cdot n_k \cdot \mathbf{q}_k\rangle \quad (19)$$

5　シミュレーション

各層ユニット数 (4-4-4) 個の 3 層ニューラルネットに対して，4 次元 Random-mapping 課題を適用した．パターン入力継続期間における標準正規分布ゆらぎのサンプル数 1000 個，$\epsilon = 0.1$, 結合係数初期値 10 系列を用いて，学習を行った．その結果，初期値10 系列すべてが収束することを確認した ($Fig.3$)．

Fig.3 Learning curves in Random-mapping problem

6　むすび

線形しきい値ニューロン群からなる PFNのための連結相関学習法を提案した．理論的検討から，本学習法は,(1) 情報伝達諸量の瞬時値 (出力誤差，ゆらぎ，シナプス前ニューロン出力) の積の期待値を用いて結合係数を修正可能，(2) すべての結合係数を並列かつ同時に修正可能，(3) 層構造を意識することなく学習処理可能，などの特徴を有する．また,4 次元 Random-mapping 課題により学習可能であることを確認した．今後，性能支配要因等について検討する．

Probabilistic Feedforward Neural Network
and its High-ordered Correlation Learning

Kakuta YUUKI, Mitsuaki MITANI, Takahumi OOHORI, Kazuhisa WATANABE
Hokkaido Institute of Technology, Teine-ku,Sapporo,JAPAN