複数の準最適教師を用いた線形パーセプトロン学習

吉川 恭平†　　原 一之‡
東京都立工業高等専門学校†　　東京都立産業技術高等専門学校‡

1　研究背景

オンライン学習に於いて、真の教師が明示的に与えられず、複数の準最適な教師を生徒が学習する問題をでは、生徒は準教師群より優れた能力を獲得できることが報告されている[1]。しかし、準教師の性能が異なると、性能の低い教師に引っ張られてしまい、学習の効果が劣化してしまうことがある。そこで、生徒が準教師の性能に比例して教師を信用するという学習法を提案し、計算機シミュレーションにより、その有効性を示す。

2　定式化

2.1　従来の方法

真の教師、準教師、生徒は、線形パーセプトロンとし、それぞれの結合荷重を $\boldsymbol{A}$、$\boldsymbol{B}_k$、$\boldsymbol{J}$ とする。線形パーセプトロンとは、図1のような複数の入力と一つの出力、それをつなぐ結合荷重のあるネットワークのことである。V_k は図では一つしか描かれていないが、実際は複数存在する。

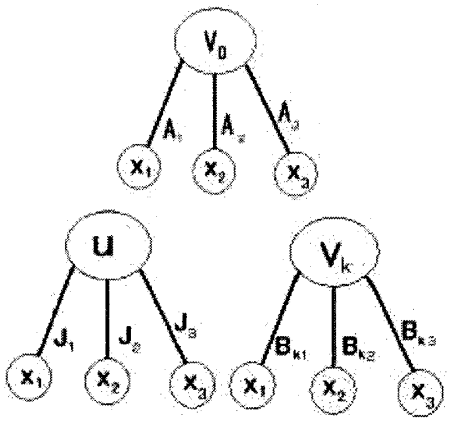

図1.線形パーセプトロン

Linear perceptron leaning using many quasi-optial teachers

†Kyohei Yoshikawa・Tokyo Metroporitan College of Technology
‡Kazuyuki Hara・Tokyo Metroporitan College of Industrial Technology

学習はまず、それぞれに同じ入力 x を与え、複数の準教師からランダムに選択した一つの教師を選択し、学習する。これを一定の回数になるまで繰り返す。真の教師、準教師、生徒の出力の式は以下である。

$$v_o = \sum_{i=0}^{N} A_i x_i = \mathbf{A}\cdot\mathbf{x} \qquad (1)$$

$$v_k = \sum_{i=0}^{N} B_{ki} x_i = \mathbf{B}_k\cdot\mathbf{x} \qquad (2)$$

$$u = \sum_{i=0}^{N} J_i x_i = \mathbf{J}\cdot\mathbf{x} \qquad (3)$$

ここで、V_k はk番目の準教師を表わす。
学習則には、勾配法を用いる。

$$\mathbf{J}(m+1) = \mathbf{J}(m) + \eta(v_k - u)\mathbf{x} \qquad (4)$$

ここで N は入力の次元である。結合荷重の各要素は平均0、分散1のガウス分布からサンプルする。また、入力の各要素は平均0、分散 $1/N$ のガウス分布からサンプルする。学習の評価には、真の教師と生徒の二乗誤差を使用する。

$$E_g = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{2}(v_0 - u)^2 \qquad (5)$$

今回はM=Nとする。
この式は、未知の入力に対し、生徒が出力を間違える確率である。

2.2　提案手法1

今回の学習法には(4)式に改良を加える。そのために、各準教師の信頼度を考える。考え方としては、生徒が各順教師に対して、テスト行い、その結果を見て、生徒が各準教師に対してどの程度信頼するかを決める。今回は、真の教師と、各準教師との方向余弦を信頼度とした。

$$R_k = \frac{\mathbf{B}_k\cdot\mathbf{A}}{\|\mathbf{B}_k\|\cdot\|\mathbf{A}\|} \qquad (6)$$

R_k に学習係数をかけたものを新たな学習係数とする。

$$\eta' = \eta\cdot R_k \qquad (7)$$

したがって、学習則は

Copyright © 2010 Information Processing Society of Japan.
All Rights Reserved.

$$\mathbf{J}\cdot(m+1)=\mathbf{J}(m)+\eta'\cdot(V_k-u)\mathbf{x} \qquad (8)$$

となる。

2.3 提案手法 2

式(8)の学習則は、R_k が極端に小さいと、η'が小さくなりすぎるため、η'を全準教師の R_k で正規化した。次の学習則に対しても、実験を行う。

$$\eta''=\eta\cdot t\cdot\frac{R_k}{\sum_{i=1}^{i=t}R_t} \qquad (9)$$

t は準教師の数である。この学習係数は他の準教師の信頼度も見ているので、より全体の中の選択された準教師を見ることができる。これにより、学習則は以下になる。

$$\mathbf{J}\cdot(m+1)=\mathbf{J}(m)+\eta''(V_k-u)\mathbf{x} \qquad (10)$$

3 実験

3.1 実験 1

N=1000,η=0.1 とした。準教師は 1,3,5,7,10 生徒数は 1 である。真の教師と各準教師との方向余弦の平均は 0.6 である。

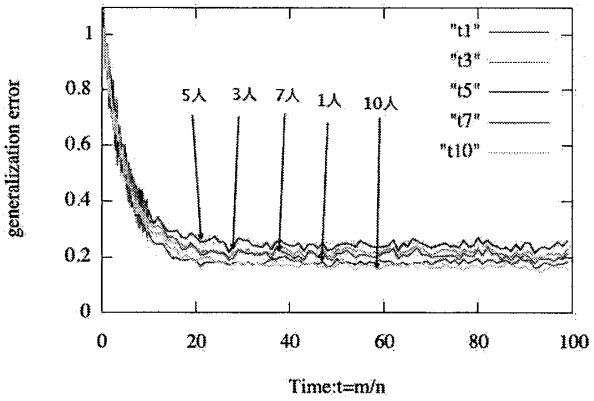

図 2.従来の学習則の場合

図 2 は従来の学習則を用いた結果である。また、5 人時より 3 人時の方が性能が良くなる等、逆転現象も発生している。

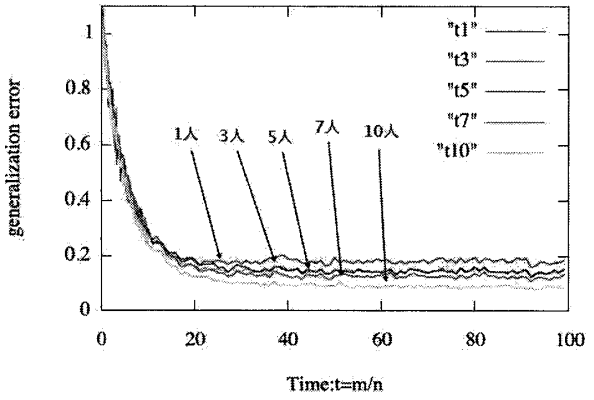

図 3..提案した学習則式(8)の場合

図 3 の提案した学習則では、全体的な学習では、全体的な学習効果の向上に加えて、逆転現象の改善もみられる。これは、従来の学習則に比べ、信頼度をおくことにより、準教師の信頼度によって学習の度合いを変えることで、学習効果を高めるという狙いが成功したと言えるだろう。これにより提案手法の有効性が示された。

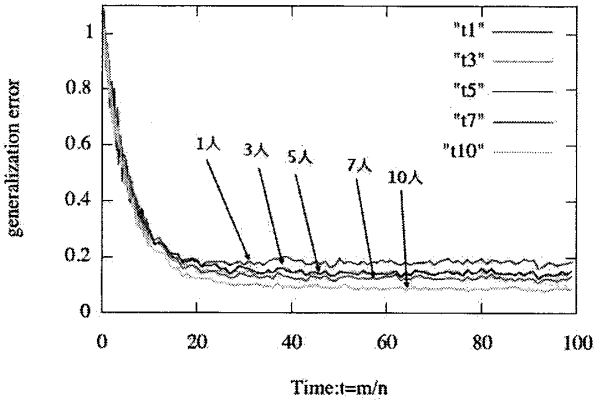

図 4..提案した学習則式(10)の場合

図 4 も、ほぼ図 3 の傾向と同じになった。

3.2 実験 2

式(8),(10)の学習則に対して次の実験を行った。N=1000,η=0.1 とし、準教師の数を 3,各準教師の性能は、R_{10}=0.01,R_2=0.02,R_3=0.03 とした。

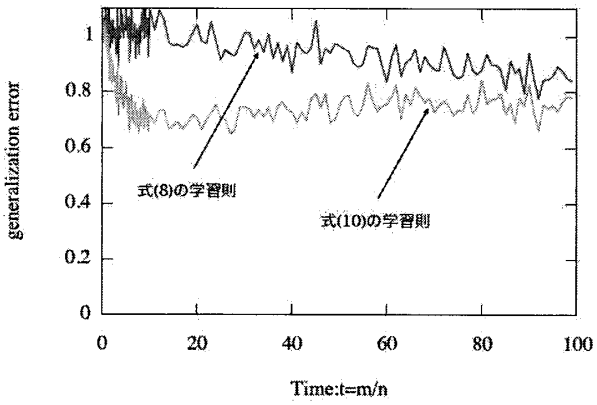

図 5.各準教師の性能が極端に低い場合

実験 2 は提案した二つの学習則に対して、各準教師の性能を極端に下げて実験を行った。結果を見て分かるとおり、式(10)の学習則は、式(8)の学習則よりも、準教師の性能が極端に低いことに対する耐性が高いことがわかる。

参考文献

[1]**吉川恭平,原一之,三好誠司,岡田真人,"複数の準最適教師を用いた線形パーセプトロン学習",電子情報通信学会学生会**,(2009)

Copyright © 2010 Information Processing Society of Japan.
All Rights Reserved.