

アダルトン学習則によるアンサンブル学習の解析

川戸祐介¹ 三好誠司² 岡田真人³ 原一之¹

東京都立工業高専¹ 神戸市立工業高専² 理化学研究所 BSI³

1. はじめに

近年, 単一の学習機械では実現できないようなタスクを, 複数の学習機械を用いることによって実現する試みが行われている[1]. このような学習法をアンサンブル学習, または集団学習と呼ばれている. 三好らは非線型パーセプトロンを用いたアンサンブル学習に対して汎化誤差の解析を行った[2]. 一般に集団の出力は各学習機械の多数決で求めるが, Urbanczik は文献[1]で, 学習機械の結合荷重の平均を結合荷重とするような, 新たな学習機械の出力を集団の出力とする方法を提案している. 文献[2]では Urbanczik らの方法の優位性を示しているが, その理由については明確に示していない. そこで本論文では Urbanczik らの方法が優位となる理由を明らかにすることを目的とする.

2. アンサンブル学習

学習機械は単純パーセプトロンである. 今, 入力を $X=(x_1, \dots, x_N)$, 単純パーセプトロンの結合荷重を $J=(J_1, \dots, J_N)$ とすると, 出力 u は

$$u = \text{sgn}(J \cdot X) \quad (1)$$

で与えられる. ここで $\text{sgn}(\cdot)$ は符号関数である. 入力次元を N とした.

一方, 教師は生徒と同じ入力 X を受けとり, 生徒に対する正解出力 v を出力する. 教師の結合荷重を $B=(B_1, \dots, B_N)$ とすると,

$$v = \text{sgn}(B \cdot X) \quad (2)$$

で与えられる.

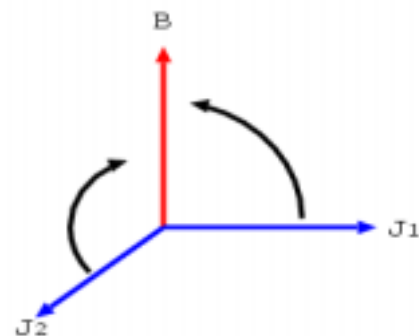
アンサンブル学習では, 複数の生徒を用い, これらを教師に対して個別に学習する. Urbanczik らは集団の出力として, 生徒の結合荷重の平均を結合荷重とする新たな生徒を考え, その出力を集団の出力とする方法を提案している. 式(4)に集団の出力を表す式を示す. U は新たな生徒の出力, \bar{J} が生徒の結合荷重の平均, K は生徒の数である.

$$U = F(\bar{J} \cdot X) = F\left(\frac{1}{K} \sum J_k \cdot X\right) \quad (3)$$

3. 結合荷重を平均する効果

以下, 結合荷重を平均する効果を直感的に説明する. 簡単のために $N=2, K=2$ とする. 教師の結合荷重 B および生徒の結合荷重 J_1 および J_2 はランダムに初期化するものとする. これらは図1に示すように互いに直交する.

図1: 初期の教師と生徒の結合荷重



学習を開始すると図1のように, 生徒 J_1 および J_2 は唯一の教師 B と一致するよう, 結合荷重を変更し, 同じように B に近づく. 学習を開始してある時刻が経った状態を図2

に示す. 図2に示すように, J_1 と J_2 は B と等距離に配置される. この時, J_1 と J_2 の合成ベクトルは J_1 および

Improvement of generalization of on-line ensemble leaning
Yusuke Kawato¹, Miyoshi Seiji², Okada Masato³, Hara Kazuyuki¹
¹Tokyo Metropolitan College of Technology
²Kobe City College of Technology
³Laboratory for Mathematical Neuroscience. Brain Science Institute

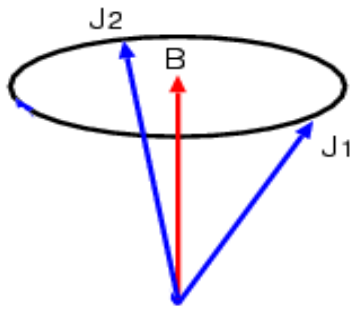


図 2：ある時刻の教師と生徒の結合荷重

J_2 より B に近づくことに注意する．したがって，生徒の結合荷重を平均することによって，教師の結合荷重のよい近似が得られる．

一方，生徒の出力の平均を集団の出力とする場合，式(1)からわかるように， J_1 と J_2 は符号関数を作用させた後に加算されるため，直接加算するような効果が期待できない．

4．結果

オンライン学習で単純パーセプトロンを学習する．学習則は AdaTron を用いる．結合荷重の更新式は

$$J(m+1) = J(m) - u\Theta(-uv)x \quad (5)$$

で与えられる． m は学習回数を示す．計算機シミュレーションでは入力の次元は $N=1000$ ，汎化誤差は学習に用いる入力と別に用意した 1000 個の入力に対する 2 乗誤差の平均により求めた．図 3，図 4 に生徒数 $K=1, K=3, K=5$ の場合の生徒の出力平均と，結合荷重の平均を結合荷重とする新しい生徒の汎化誤差をそれぞれ示す．

図のように，結合荷重の平均を結合荷重とする新しい生徒の汎化誤差が，出力平均の汎化誤差を取った時に比べ，小さくなっていることがわかる．以上より，第 3 節で述べた議論の正当性を裏付ける結果が得られた．

5．まとめ

本研究では，結合加重平均における汎化誤差が出力平均における汎化誤差よりも小さくなることをシミュレーションし，明らかにした．今後

は，マージンを加えるなどさらなる改善を検討する．

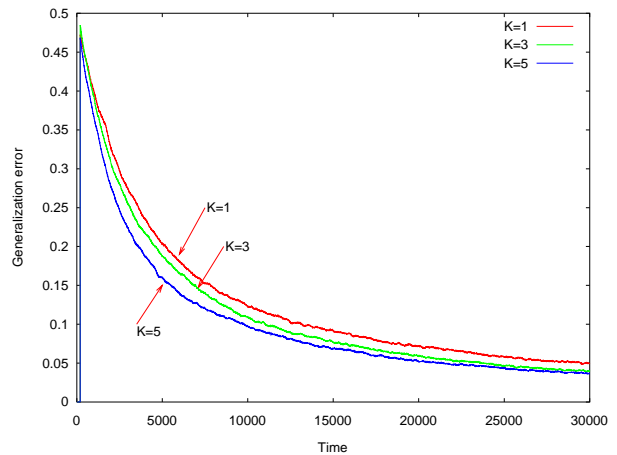


図 3：出力平均の汎化誤差

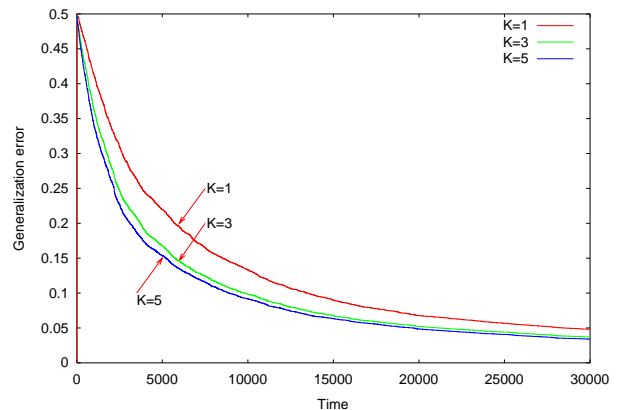


図 4：結合荷重の平均を結合荷重とする新しい生徒の汎化誤差

参考文献

- [1] Urbanczik R., "online learning with ensembles", Phys. Rev. E, **62**, pp. 1448 (2000).
- [2] 三好, 原, 岡田, "オンライン学習理論による非線型パーセプトロンのアンサンブル学習の解析", IBIS, pp. 283 (2003).