

◎三好 亮介(関西大院) 前田 裕(関西大) 三好 誠司(関西大)

**あらまし** パーセプトロン学習とアダトロン学習の違いを考察し、それを荷重摂動学習に適用することにより適応型荷重摂動学習を提案した。またその汎化能力を統計力学的手法で解析した。その結果、提案する学習則がアダトロン学習と同一の優れた漸近特性を有することが明らかになった

## 背景

- オンライン学習の枠組みにおいて、学習機械のパラメータに摂動を加える学習則として**荷重摂動学習**が提案されている
- 荷重摂動学習の汎化能力は統計力学的手法で解析されており、その結果パーセプトロン学習と同一の漸近特性を有することが明らかにされている
- よく知られている学習則であるアダトロン学習の漸近特性は  $\epsilon_g \sim O(t^{-1})$  であり、パーセプトロン学習の  $\epsilon_g \sim O(t^{-3/2})$  よりも優れている

## 目的

- よく知られている学習則であるパーセプトロン学習とアダトロン学習の違いを考察し、荷重摂動学習に適用することにより**適応型荷重摂動学習**を提案
- 適応型荷重摂動学習の汎化能力を統計力学的手法を用いて解析

## モデルと更新則

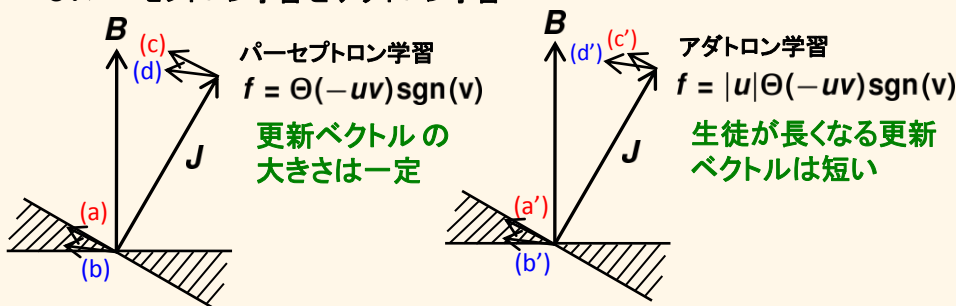
### 1. モデル

教師の出力  $\text{sgn}(v)$ ,  $v = B \cdot x$   
 生徒の出力  $\text{sgn}(ul)$ ,  $ul = J \cdot x$   
 オンライン学習  $J^{m+1} = J^m + f^m x^m$

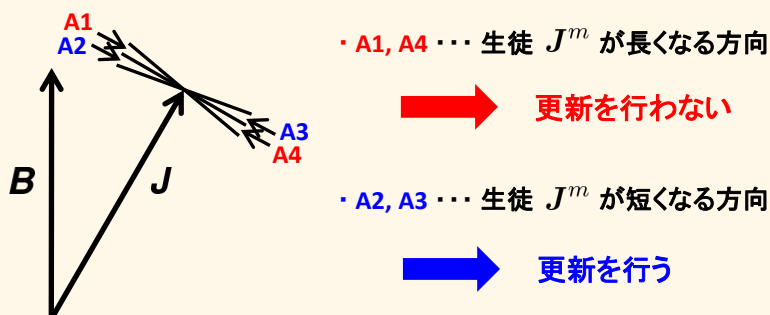
### 2. 荷重摂動学習

$$f^m = -\frac{\eta}{2c} g^m, \quad g^m = \Theta(-v^m(J^m + cx^m) \cdot x^m) - \Theta(-v^m(J^m - cx^m) \cdot x^m)$$

### 3. パーセプトロン学習とアダトロン学習



### 4. 適応型荷重摂動学習の提案



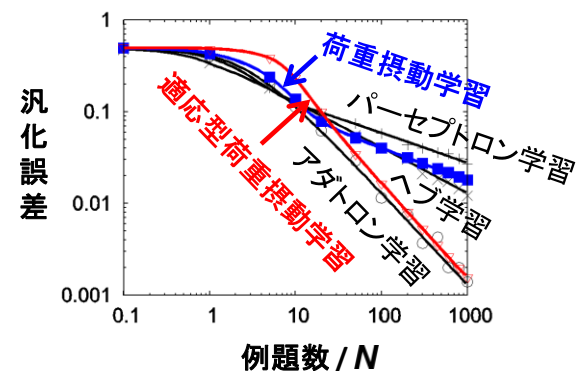
$$f^m = -\frac{\eta}{2c} |u^m| g^m \Theta(u^m g^m)$$

$$g^m = \Theta(-v^m(J^m + cx^m) \cdot x^m) - \Theta(-v^m(J^m - cx^m) \cdot x^m)$$

## 統計力学解析

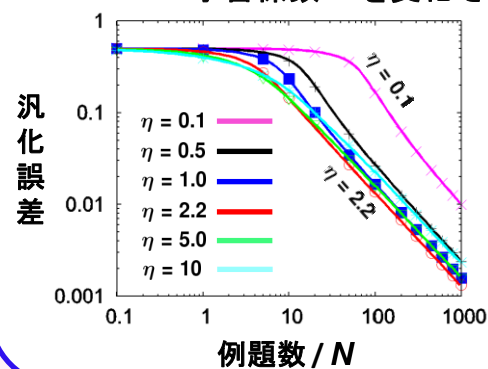
- 次元  $N \rightarrow \infty$
- 巨視的変数に関する微分方程式を自己平均性に基づき決定論的な形で導出

## 結果と考察



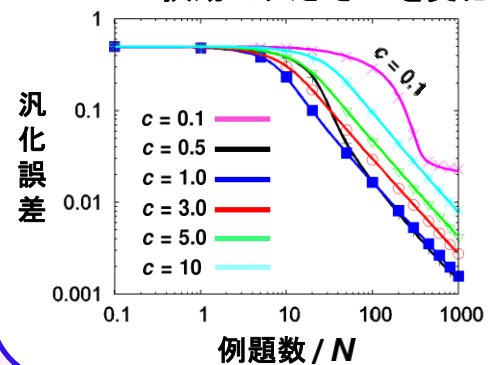
適応型荷重摂動学習はアダトロン学習と同一の優れた漸近特性

### 学習係数 $\eta$ を変化した場合



最適な  $\eta$  が 2.2 付近に存在

### 摂動の大きさ $c$ を変化した場合



$c$  が小さいとき特異な振る舞い

## 参考文献

- J. Werfel, X. Xie, and H. Seung, "Learning curves for stochastic gradient descent in linear feedforward networks", *Neural Computation*, 17, pp.2699-2718, 2005.