3A303-12

入力が相関を有するオンライン学習に関する統計力学的解析

中尾 健人（関西大） 鳴川 雄太（ダイヘン） 三好 誠司（関西大）

THINK × ACT KANSAI UNIVERSITY

あらまし 互いに相関がある入力を毎回K個ずつ使うオンライン学習の汎化能力について統計力学的手法を用いて解析する.その結果，ヘブ学習とアダトロン学習では入力の相関が大きいほど学習が遅くなり，汎化誤差に下限が存在するのに対し，パーセプトロン学習では学習速度が入力の相関の影響を受けず，汎化誤差の下限も存在しないことが明らかになった. また，学習の初期段階で学習が急速に進む特異なダイナミクスを有することが明らかになった.

背 景

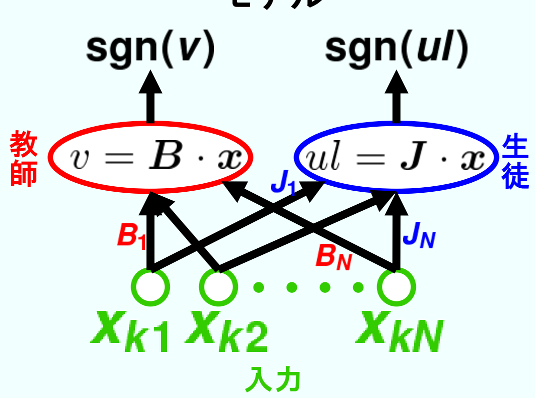
- 最近，時間的あるいは空間的な観点で興味深いいくつかのモデルがオンライン学習の枠組みで解析されているが，それらはいずれも入力が独立に生成される場合を扱っている
- 実際のパターン認識の応用の場面においては入力が相関を有する場合が多くみられる

目 的

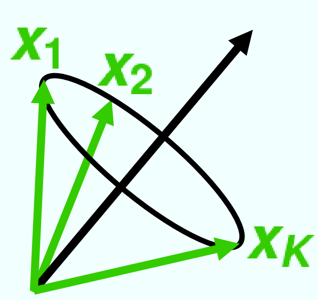
- 互いに相関がある入力を毎回 K個ずつ使うオンライン学習の汎化能力について統計力学的手法を用いて解析する

手 法

モデル



相関がある入力



生徒の更新

$$J^{m+1} = J^m + \sum_{k=1}^{K} f_k^m x_k^m$$

汎化誤差

$$\epsilon_g = \frac{1}{\pi}\cos^{-1} R$$

次元N→∞において決定論的に導出される巨視的連立微分方程式

$$\frac{dl^2}{dt} = K\langle f_k^2\rangle + a^2 K(K-1)\langle f_k f_{k'}\rangle + 2Kl\langle f_k u_k\rangle \qquad \frac{d(Rl)}{dt} = K\langle f_k v_k\rangle$$

三つの学習則

ヘブ学習 $f = \mathrm{sgn}(v)$

$$\langle f_k u_k\rangle = \frac{2R}{\sqrt{2\pi}} \qquad \langle f_k v_k\rangle = \sqrt{\frac{2}{\pi}} \qquad \langle f_k^2\rangle = 1 \qquad \langle f_k f_{k'}\rangle = \left(1 - \frac{2}{\pi}\cos^{-1} a^2\right)$$

パーセプトロン学習 $f = \Theta(-uv)\mathrm{sgn}(v)$

$$\langle f_k u_k\rangle = \frac{R-1}{\sqrt{2\pi}} \qquad \langle f_k v_k\rangle = -\langle f_k u_k\rangle \qquad \langle f_k^2\rangle = \frac{1}{\pi}\cos^{-1} R$$

$$\langle f_k f_{k'}\rangle = \int du_k du_{k'} dv_k dv_{k'} P(u_k, u_{k'}, v_k, v_{k'})\Theta(-u_k v_k)\mathrm{sgn}(v_k)\Theta(-u_{k'} v_{k'})\mathrm{sgn}(v_{k'})$$

アダトロン学習 $f = -u\Theta(-uv)$

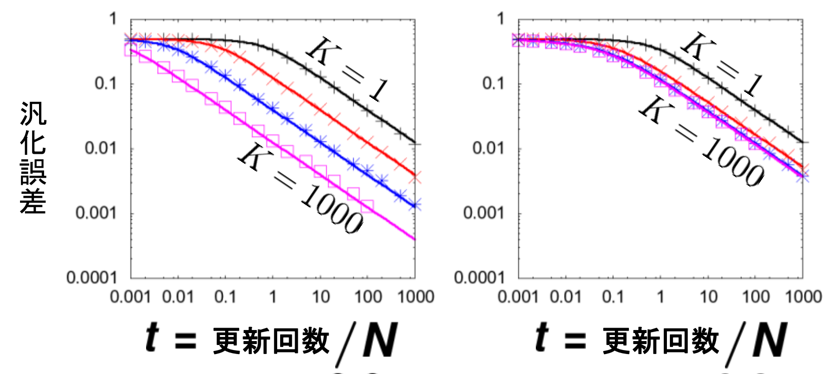
$$\langle f_k u_k\rangle = \frac{1}{\pi}\left(R\sqrt{1-R^2} - \cos^{-1} R\right) \quad \langle f_k v_k\rangle = \frac{1}{\pi}(1-R^2)^{\frac{3}{2}} + R\langle f_k u_k\rangle \quad \langle f_k^2\rangle = -\langle f_k u_k\rangle$$

$$\langle f_k f_{k'}\rangle = \int du_k du_{k'} dv_k dv_{k'} P(u_k, u_{k'}, v_k, v_{k'}) u_k\Theta(-u_k v_k) u_{k'}\Theta(-u_{k'} v_{k'})$$

ヘブ学習についてはサンプル平均を代入して得られる連立微分方程式は解析的に解くことができる
パーセプトロン学習とアダトロン学習については数値的に解く必要がある

結果と考察

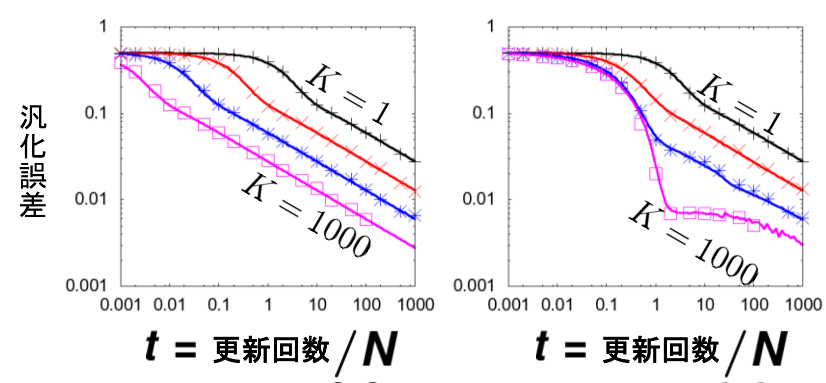
ヘブ学習



相関 a = 0.0 相関 a = 0.6

- 入力の相関が大きいほど学習が遅くなる
- 入力に相関がある場合，汎化誤差に下限が存在

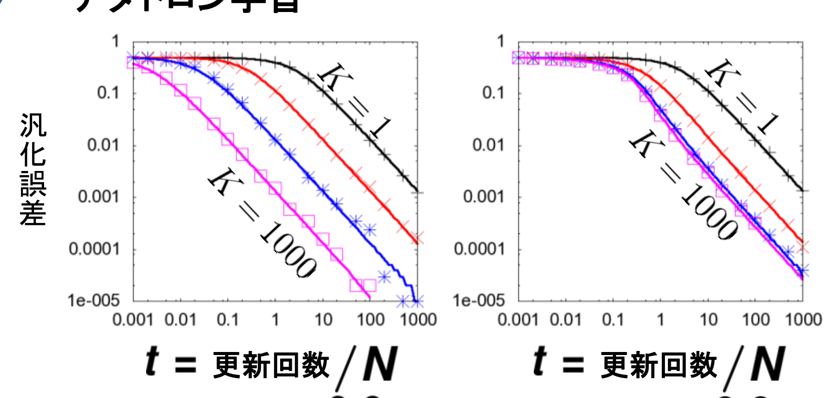
パーセプトロン学習



相関 a = 0.0 相関 a = 0.6

- 漸近領域における汎化誤差が入力の相関の影響を受けない
- 入力に相関がある場合，汎化誤差の下限も存在しない
- t = 1付近で学習が急速に進む特異なダイナミクス

アダトロン学習



相関 a = 0.0 相関 a = 0.6

- 入力の相関が大きいほど学習が遅くなる
- 入力に相関がある場合，汎化誤差に下限が存在