H-015

マージン最大化学習を用いた連想記憶の耐雑音性改善

Robust Associative Memory Using the Maximal Margin Classifier

秋本仁志†　　服部元信†
Hitoshi Akimoto　Motonobu Hattori

あらまし

Support Vector Machines(SVMs) の学習法として知られるマージン最大化学習を拡張して，連想記憶の 1 つであるホップフィールドモデルの学習に適応することで非常に高い耐雑音性を持つ連想記憶を構築した．そして従来法と比較することで本手法の優位性を検証した．

1. まえがき

近年，あいまいな情報処理を可能とするために人間の脳を工学的に模倣した連想記憶の研究が盛んに行われている．しかし，従来の学習法は記憶容量や学習速度の改善のみに着目しているものが多く，あいまいさの吸収や耐雑音性を考慮した学習法はあまり多くないのが現状である．一方，パターン認識の分野で非常に優れた成果を挙げている手法として Support Vector Machines(SVMs)[1] がある．SVMs はマージン最大化学習を用いることで一意に最適な重みを得ることを可能としている．この学習を連想記憶の学習に適用できれば，マージン最大化の原理により，耐雑音性に優れた連想記憶を構築できると考えられる．そこで本論文ではマージン最大化学習を連想記憶モデルの 1 つであるホップフィールドモデルの学習に用いるために拡張する方法を提案する．そして，シミュレーションを行い従来の学習法と耐雑音性について比較を行う．

2. マージン最大化学習

学習パターンとその識別を行うための識別面が与えられたとき，SVMs におけるマージンとは，サポートベクターと呼ばれる識別面に最も近いパターンと識別面との距離で表される．マージン最大化とは識別すべき 2 クラスどちらのサポートベクターとも距離が最大になる，つまり 2 クラスの真ん中を識別面が通ることを意味している．その様子を 2 次元のパターン空間で表したものを図 1 に示す．ここで□が class A の，○が class B の学習パターンを表している．SVMs は線形識別器なので学習パターン $\boldsymbol{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$ の識別関数は次のように表すことができる．

$$f\left(\boldsymbol{x}^{(k)}\right) = \sum_{j=0}^{n} w_{ij} x_j^{(k)} \quad \begin{cases} f(\cdot) < 0 \; class\,A \\ f(\cdot) \geq 0 \; class\,B \end{cases} \tag{1}$$

このときマージン γ は点と直線の距離より式 (2) のように書ける．

$$\gamma = \min \frac{\left|f\left(\boldsymbol{x}^{(k)}\right)\right|}{\|\boldsymbol{w}\|} \quad (k = 1, \ldots, m) \tag{2}$$

ここで重み $\boldsymbol{w}$ は何倍しても意味する識別面は変わらないため，式 (3) の制約を加えるとマージンは $1/\|\boldsymbol{w}\|$ で表すことができ，このマージンを最大化することで最適な重み $\boldsymbol{w}$ を得ることができる．

$$\min\left|f\left(\boldsymbol{x}^{(k)}\right)\right| = 1 \quad (k = 1, \ldots, m) \tag{3}$$

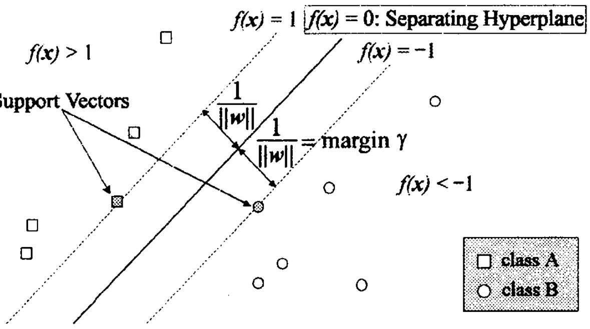

図 1: 特徴空間上の識別面とマージン

3. ホップフィールドモデル

ホップフィールドモデルは 1 層からなるニューラルネットワークであり，ネットワークを構成するニューロンの出力を他の全てのニューロンにフィードバックする相互結合型の構造をしている．想起時には 1 時刻につき 1 つのニューロンのみが式 (1) と同様な式を用いて状態更新を行い，ネットワーク全体のニューロンの状態が変化しなくなったとき，その状態を出力とする．代表的な学習法には擬似緩和学習法 [2] などがある．ここで，想起時のニューロン 1 つの動作は，SVMs と同じように線形識別を行っていることに注目し，マージン最大化学習を適用できると考えた．

4. 従来の学習法

4.1 擬似緩和学習法

Oh らによって提案された擬似緩和学習法は緩和学習法を改良して少ない更新回数で重みを解領域に収束させる方法である．擬似緩和学習法ではホップフィールドモデルの各ニューロンにおける全ての学習パターンの想起を保証する条件を次の不等式で表す．

$$x_i^{(k)} f\left(\boldsymbol{x}^{(k)}\right) - \xi \geq 0 \quad (i = 1, \ldots, n,\; k = 1, \ldots, m) \tag{4}$$

このとき，重みは次式によって更新される．

$$w_{ij}^{(new)} = w_{ij} - \Delta w_{ij} \tag{5}$$

$$\Delta w_{ij} = \frac{\lambda}{1+n}\left(f\left(\boldsymbol{x}^{(k)}\right) - \xi x_j^{(k)}\right) x_i^{(k)} \tag{6}$$

ここで，$\xi > 0$, $1 < \lambda < 2$ である．擬似緩和学習法は，ξ によって遠ざけられ解領域に対して重みの更新を行うことで高速な収束を可能としている．

†山梨大学大学院医学工学総合教育部, 甲府市
Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, 4-3-11 Takeda, Kofu 400-8511 Japan

4.2 Hasegawa らの改良法

擬似緩和学習法では，重みが実際の解領域に到達した時点で学習が終了する．Hasegawa らは，耐雑音性を改善することを目的として，重みが解領域に到達した後も学習を続ける方法を提案した [3]．具体的には式 (6) の更新を $0 < \lambda < 1$ として行い，式 (4) の不等式で表される領域の境界をなす超平面と重みの距離の変化が十分に小さくなったときに学習を終了する．これにより解領域の境界をなす全ての超平面から同程度の距離に重みが収束するため，全ての学習パターンに対して想起率が改善される．

5. マージン最大化学習の拡張

先に述べたようにホップフィールドモデルのニューロン 1 つの動作は一般的な SVMs と同じことを行っているためマージン最大化学習を適用できると考えられるが，ただ SVMs をニューロンの数だけ組み合わせるわけにはいかない．何故ならホップフィールドモデルの想起時にネットワークの状態がエネルギー関数の安定点に達するためには，以下の 2 つの制約条件を満たすように定式化しなければならないからである．

1. 対称な重みを用いる
2. 自己結合は 0 である

5.1 学習定式化

そこでまず，n 次元からなる m 個の学習パターン $\boldsymbol{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$ が与えられたとき，目的関数としてマージン $1/||\boldsymbol{w}||$ を最大化するために次の式を最小化することを考える．ただし，各ニューロンの閾値は重みの先頭に埋め込んで計算を行うものとする．

$$L(\boldsymbol{w}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=0}^{n}(w_{ij})^2 \tag{7}$$

さらに，制約条件を考慮するためにラグランジュの未定乗数法を用いると目的関数は次式になる．

$$\begin{aligned} L(\boldsymbol{w}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=0}^{n}(w_{ij})^2 \\ &\quad - \sum_{i=1}^{n}\sum_{k=1}^{m}\alpha_i^{(k)}\left\{x_i^{(k)}\left(\sum_{j=0}^{n} w_{ij}x_j^{(k)}\right) - 1\right\} \\ &\quad - \sum_{i=1}^{n}\sum_{\substack{j=0 \\ (j\neq i)}}^{n}\beta_{ij}(w_{ij} - w_{ji}) - \sum_{i=1}^{n}\gamma_i w_{ii} \end{aligned} \tag{8}$$

式 (8) 中で制約条件は正の変数であるラグランジュ乗数を用いて表されており α の項は「全てのニューロンが全ての学習パターンを識別できる」という制約を，β の項は先に述べた 1 の制約を，γ の項は 2 の制約をそれぞれ表している．この式 (8) の目的関数を最適化することで初期値によらず一意に最適な重みを得ることができる．シミュレーション中では通常 SVMs でよく用いられるようにガウスカーネル関数を用いてパターンを高次の空間に写像して識別を行った．

6. 計算機シミュレーション

シミュレーションは図 2 に示す 7 × 7bit からなる 26 個のアルファベットパターン"A-Z"を用いて学習を行い，学習パターンにノイズを加えたものを入力として与えて想起を行った．試行回数は 100 回とし，このときの提案法，Hasegawa 法，擬似緩和学習法を用いた場合の想起率を図 3 に示した．また，同様に 100bit のランダムパターン 50 個を用いてシミュレーションを行った結果を図 4 に示した．ここで，想起率とは入力に対応する学習パターンが正しく想起された確率を示している．

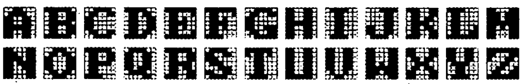

図 2: シミュレーションに用いたアルファベットパターン

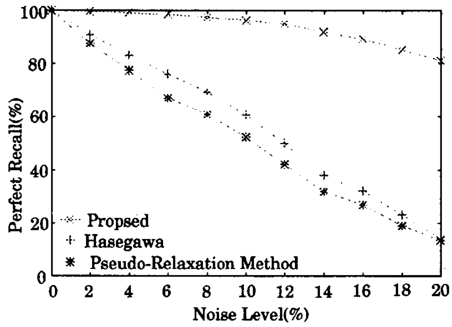

図 3: 各学習法の想起率 (アルファベットパターン)

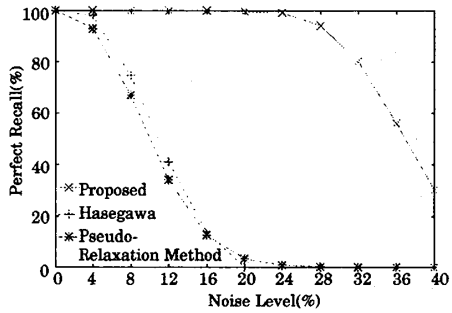

図 4: 各学習法の想起率 (ランダムパターン)

図 3，4 の結果を見ると提案法は従来法のどちらと比べても極めて高い耐雑音性を得ることができている．また図 4 の結果から学習パターンの特性に依存せず，汎用性があることが確認できた．

7. むすび

連想記憶の 1 つであるホップフィールドモデルに SVMs の学習法で知られるマージン最大化学習を拡張して適用する方法を提案し，その有効性を示した．提案法を用いることで高い耐雑音性を持つ，優れた連想記憶を構築することができることが分かった．

参考文献

[1] Cristianini, N. and Shawe-Taylor, J., "An Introduction to Support Vector Machine and other kernel-based learning methods," Cambridge University Press, 2000.

[2] Oh, H and Kothari, S.C., "A new learning approach to enhance the storage capacity of the hopfield model," IJCNN, Singapore, pp.2056-2062, 1991.

[3] Hasegawa, K. and Hattori, M., "Improved Pseudo-Relaxation Learning Algorithm for Robust Bidirectional Associative Memory," Artificial Neural Nets and Genetic Algorithms, pp288-293, 1999.