

# LP,DP,SVM を用いた時系列識別関数学習の検討

野口 顕彦\* 内田 誠一\* 迫江 博昭\*  
\*九州大学大学院 システム情報科学府

## 1 はじめに

音声やオンライン文字などの時系列パターンでは時間軸方向の伸縮が発生する．伸縮変形に対して頑強な識別関数を得るには変形したものを含む多数のパターンを用いて学習をおこなう必要があるが，十分な量のパターンを収集することは困難である．そこで，線形計画法(LP)による線形分離を基本とし，動的計画法(DP)による非線形ワープを用いて学習に有効な伸縮パターンを自動的に生成し，これを学習パターンに追加して線形分離学習するというサイクルを識別関数の精度が十分に高まるまで繰り返す手法を考える(LPDP[1])．

LPDP は少数の学習パターンから出発して高性能な識別関数が得られることが確認されているが，LPDP サイクルの進行にともない学習パターン集合が成長し学習にかかる計算量が増大して実行が困難な状況になる．また，LP による線形分離学習における係数変域制限の影響で解が歪むという問題が存在する．本研究では，これらの既知の問題に対する解決手法を組み込んだ LPDPSVM を提案する．

## 2 改良点

### 2.1 計算量の削減

LP による線形分離学習は，全学習パターンの線形分離を制約条件とし，分離マージンを最大化することを目的関数としてシンプレックス法で解く．線形計画問題を双対型に変形することで，制約条件数に対するタブローのサイズを標準型の場合より大幅に抑えることができる．

新しい制約条件が増えるとそれに対応する列がタブローに追加される．DP による伸縮パターンが学習パターン集合に追加されるとき，従来法ではシンプレックスタブローを初期状態から再度計算しなおしている．本手法では最適解状態タブローの基底変数に対応する部分の初期状態行列  $B$  から現在の状態(単位行列  $I$ ) への変形，すなわち  $B^{-1}$  を新規に追加される列  $p$  にかき  $B^{-1}p$  として前サイクルでの最適解状態タブローに追加し，これを基底解状態タブローとして最適化を続行する．この基底解は初期状態からある程度最適化が進行しているものであるから，最適解に収束するまでの掃き出し回数が従来法に比して大幅に削減される．

### 2.2 解の歪みの除去

LP による線形分離学習において，目的関数値が無限大になることを防ぐために識別関数の係数の変域制限が制約条件として課される．この制限が各変数に対して独立であるため，最適化の結果が各変数の変域の限界に引き寄せられて解が歪み，未知パターンに対する頑健性が落ちる．

本手法では LPDP サイクルが終了したあとで解の歪みが発生しないアルゴリズムで新たに学習を行うことでより精度の高い識別関数を得ることを考える．繰り返し処理をおこなう LPDP サイクルにおける学習アルゴリズムは比較的高速な LP で済ませ，学習パターン集合が十分に成長してから複雑なアルゴリズムを用いて学習することで最終的な性能を稼ぐことを考える．本研究では LPDP 収束後に SVM[2] を適用して歪みの問題を解決した識別関数の学習をはかる．

## 3 実験

ETL 文字データベースの ETL1 に含まれる '2' と '3' の手書き数字データを用いて本稿で提案した手法の効果を確認した．手書き数字画像の周辺から文字領域までの距離を時系列的に並べたもの(周辺パターン)を対象のパターンとし，両クラスから 50 個のパターンを取り出して初期学習用とし，LPDP サイクルを繰り返した．

図 1 は LPDP サイクル進行にともなうサイクルあたりの計算量の変化の比較である．従来法では毎回初期状態タブローから計算するのでタブローの拡大とともに計算量が増加するのに対し，本手法では前回の最適解状態から最適化を続行するので計算量は減少してほぼ一定となる．累積計算量で考えると，サイクルの進行にともない全体としての高速化効果が向上することがわかる．以上から，提案手法により学習パターン集合の成長に影響されずに繰り返し学習を行うことが可能となる．

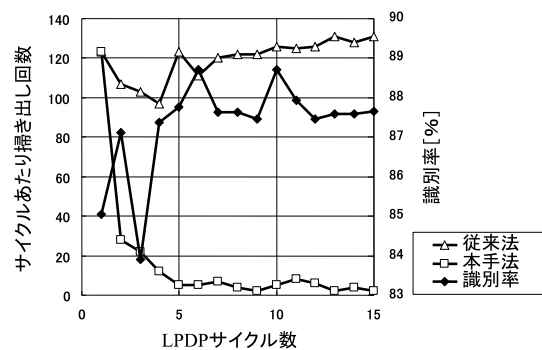


図 1: LPDP サイクルあたりの計算量

また，同手書き文字データを用いて学習パターン数や伸縮許容範囲などの条件を変化させながら SVM 適用前後の性能比較を行った結果，多くの場合において識別率が向上することを確認した．一方で識別率が低下する場合も存在し，原因追究が今後の課題となる．

## 4 まとめ

本稿では，時系列パターンの非線形伸縮に対して頑健な識別関数学習のための伸縮パターン自動生成追加機能を持つ反復学習法における問題点(計算量の増加と解の歪み)に対する改良を提案し，その有効性を確認した．

### 参考文献

- [1] 廣瀬 智貴, 片山 嬉規, 迫江 博昭, "線形計画法と動的計画法によるワープに対して安定な識別関数の学習法," 信学技報, PRMU99-169(1999-12), pp.27-34.
- [2] Christopher J.C. Burges, "A Tutorial on Support-Vector Machines for Pattern Recognition," Data Mining and knowledge discovery vol 2, No. 2, pp121-167, 1998.