

## タスクの類似性を利用したコスト考慮型学習

学籍番号 23413561 氏名 本郷 辰哉

指導教員名 竹内一郎 准教授

## 1 まえがき

ガンの診断やメールフィルタリングなどの現実の 2 クラス分類問題では、分類を誤った時のコストが等価でない場合が多く存在する。コストが等価でない場合を考慮した学習法としてコスト考慮型学習が提案されており、本研究ではコスト考慮型 SVM(Support Vector Machine)[1] を用いる。

多くの現実問題では、分類器を作成する時点で誤分類のコストが不確定なため、様々なコストに対するコスト考慮型分類器を学習する必要がある。ある学習データに対して複数のコスト考慮型分類器を学習する場合、それらは互いに類似したものとなることが考えられるため、本研究では、この類似性をうまく取り入れて学習することで汎化性能の向上を目指す。

本研究ではコスト考慮型学習の問題にマルチタスク学習のアプローチを導入する。従来のマルチタスク学習では離散的な有限個のタスクを学習する方法が考察されてきたが、本研究ではパラメータ表現された無限個のタスクを学習する新たな方法を提案する。具体的には、文献 [2] のアプローチをパラメータ表現されたタスクに適用できるように拡張する。また、この手法に変更を加え、パラメータ表現されたタスクに共通の変数を選択する方法も導入する。

提案法ではコスト比が連続的に変わるタスクについて、指定した範囲内の分類器に共通の低次元空間を見つけることができる。計算機実験により提案法が従来法よりも高性能であることを示す。

## 2 縮小ランクモデル

タスク間の類似性を学習に利用するために縮小ランクモデルを適用する。

複数のコスト考慮型 SVM の学習に縮小ランクモデル [2] を適用することで、複数のコスト考慮型分類器のパラメータが共通の低ランク空間内に存在するようにできる。これは、各分類器のパラメータを並べた行列が低ランクとなることに相当する。

以下で問題を定式化するが、 $\mathbb{N}_n := \{1, \dots, n\}$  と表記する。同時に学習するタスク数を  $T$  個とし、あ

るタスク  $t \in \mathbb{N}_T$  の識別関数を  $f_t$  とする。各タスク  $t \in \mathbb{N}_T$  はサイズ  $n$  の学習サンプル  $\{(\mathbf{x}_{ti}, y_{ti}) \in \mathbb{R}^p \times \{-1, +1\}\}_{i \in \mathbb{N}_n}$  を持つとする。

タスク  $t$  の識別関数 (この関数の符号により分類する) を

$$f_t(\mathbf{x}) = v_t + (\mathbf{w} + \mathbf{b}_t)^\top \mathbf{x},$$

と表す。ここで、 $v_t$  は定数項、 $\mathbf{w} \in \mathbb{R}^p$  は全タスクで共有する係数ベクトル、 $\mathbf{b}_t \in \mathbb{R}^p$  はタスクごとに異なる値をとる係数ベクトルである。全タスクで共有する  $\mathbf{w}$  は適当なコスト比を用いてあらかじめ求めておくものとする。一方、 $\{\mathbf{b}_t\}_{t \in \mathbb{N}_T}$  は各タスクの違いを特徴づけるもので、縮小ランクモデルでは、これらが低ランク空間内に存在するような制約を課すことで異なるタスク間の共通性を表現する。縮小ランクモデルの最適化問題は 2 つの最適化変数について交互に最適化するようなアルゴリズムとなる。

2 クラス分類問題では、一方のクラスを他のクラスに誤分類するコストを  $\theta$ 、その逆を  $1 - \theta$  と表わし、コスト比  $\theta : 1 - \theta$  のコスト考慮型学習を行う。

## 3 パラメトリック縮小ランクモデル

前節で述べた縮小ランクモデルは有限個のタスクに相当する離散的な  $\{\theta_t\}_{t \in \mathbb{N}_T}$  を与え、これらのタスクで同時に学習するためのモデルであった。本研究では、縮小ランクモデルのアプローチを拡張し、連続値でパラメトリック表現されたタスクを同時に学習するアルゴリズムを提案する。提案法では、コスト比  $\theta : 1 - \theta, \theta \in (0, 1)$  が一定の範囲内で不確実な場合にこの範囲内の全タスクで共通な低ランク部分空間を同定できる。提案法の利点は、範囲内の全てのタスクに対する分類器を学習できることや、よりタスク間の類似性を学習に活かせる事などである。提案法では、パラメトリック計画法 [3] を利用する。コスト比を表す  $\theta \in (0, 1)$  を連続パラメータとみなすとコスト考慮型学習は、パラメータ表示された連続的なタスクとみなすことができる。これらのタスクに対し、パラメトリック計画法を用いて、各タスクに対する連続的な解を得ることができる。

パラメトリックタスク変数選択についても同じような手順で最適化できる。

## 4 計算機実験

複数の誤分類のコスト比の異なったコスト考慮型 SVM を同時に学習する問題を考える。この問題に対し、パラメトリックタスク縮小ランク学習 (PTRR), パラメトリックタスク変数選択 (PTVS) を用いることでコスト考慮型分類の性能が個別学習より向上することを実データを用いた実験により確認する。実データは UCI から 10 のデータセットを利用した。比較の対象とするタスク (コスト比) はコスト比  $((1 - \theta) : \theta) = (0.9 : 0.1)$  から  $(0.1 : 0.9)$  までの 8 つとし、その合計を比較した。PTRR, PTVS は  $\theta_t = 0.05$  から  $\theta_t = 0.95$  間の全タスクでマルチタスク学習した。データはそれぞれ学習, 評価, テストの 3 つにほぼ等しい割合でランダムに 3 分割し, 10 回の試行における平均と標準偏差を表 1 に示す。分類器の評価値としてコスト付き誤差関数:

$$L_t(f) = (1 - \theta_t) \sum_{i:y_i=1} I(f(\mathbf{x}_i) \leq 0) \quad (1)$$

$$+ \theta_t \sum_{i:y_i=-1} I(f(\mathbf{x}_i) > 0),$$

を用いる。ここで関数  $I$  は引数が正なら 1 を, 負なら 0 を返すものである。また, コスト考慮型 SVM のハイパーパラメータは, 評価データに対して, 個別学習においては式 (1) を, PTRR, PRVS においてはコストの積分式:

$$L_t(f) = \sum_{i:y_i=1} \int_{0.05}^{0.95} (1 - \theta_t) I(f(\mathbf{x}_i) \leq 0) d\theta_t$$

$$+ \sum_{i:y_i=-1} \int_{0.05}^{0.95} \theta_t I(f(\mathbf{x}_i) > 0) d\theta_t,$$

を用いて決定した。

### 4.1 結果と考察

表 1 より, PTRR は 10 のデータセットの内 9 のデータセットに対して, PTVS は 10 のデータセットの内 7 のデータセットに対して, タスク合計での平均評価値が個別学習より良い性能の分類器を得た。また, 10 のデータセットに対し PTRR, PTVS の方が個別に学習するより標準偏差を小さくすることができている。これらから提案法である PTRR,

表 1: 実験結果。表内の値は全タスク合計でのコスト付き誤差。結果が良い値を太字で示す。括弧内の値は標準偏差を示す。

データ名	手法		
	個別	PTRR	PPVS
Park.	32.3	<b>30.21</b>	30.25
	(10.6)	(9.09)	(8.531)
Bre.Can.D	20.36	<b>18.49</b>	19.46
	(7.77)	(6.15)	(5.889)
Bre.Can.P	48.97	49.28	<b>48.68</b>
	(12.92)	(9.827)	(9.875)
Aust	117.97	<b>106.25</b>	111.22
	(22.965)	(12.656)	(15.951)
Dia	185.9	179.802	<b>175.95</b>
	(21.135)	(16.309)	(16.261)
Four	181.69	179.3	<b>178.67</b>
	(22.128)	(14.246)	(19.241)
Germ	242.21	<b>219.66</b>	237.2
	(18.346)	(16.216)	(15.781)
Spli	179.8	<b>151.69</b>	183.54
	(24.218)	(18.022)	(21.272)
svmgui	175.7	<b>170.16</b>	179.76
	(15.55)	(9.986)	(14.757)
vowe	<b>175.16</b>	175.74	175.5
	(13.775)	(9.366)	(7.384)

PTVS を用いることで性能のよいコスト考慮型分類器が安定して得られている事が分かる。

## 5 まとめ

本研究では, 連続値のコスト比パラメータで表現されたタスクでマルチタスク学習する手法として PTRR, PTVS を提案した。また実験において提案法の有効性を示した。

## 参考文献

- [1] Y. Lin and Y. Lee and G. Wahba. Support vector machines for classification in nonstandard situations. *Machine Learning*, 2002.
- [2] A. Argyriou and T. Evgeniou and M. Pontil. Convex multitask feature learning. *Machine Learning*, 2008.
- [3] T. Gal. Postoptimal Analysis, Parametric Programming, and Related Topics. *Walter de Gruyter*, 1995.