様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２８ 年　 ５ 月 ３０ 日現在

機関番号： １４４０１

研究種目： 基盤研究(C)（一般）

研究期間： 2012～2015

課題番号： ２４５００２７５

研究課題名（和文） 多クラス識別に対する幾何的マージン最大化ソフトマージンサポートベクトルマシン

研究課題名（英文） Soft-magin support vector machine maximizing geometric margins for multiclass classification

研究代表者

巽　啓司（Tatsimi, Keiji）

大阪大学・工学（系）研究科（研究院）・准教授

研究者番号：３０３０４０１７

交付決定額（研究期間全体）：（直接経費） 3,900,000円

研究成果の概要（和文）：著者がすでに提案している，多クラス識別問題に対して幾何マージンを正確に最大化する多目的多クラスサポートベクトルマシン(MMSVM)に対し，例外データが含まれる教師データに対しても学習可能なソフトマージンモデルへの拡張や，クラスタリングなどを利用した学習時のメモリー使用量・計算時間を削減した定式化，その解法について検討した．その際，制約領域の制限法，多目的最適化問題の求解法，スラック変数に対するペナルティ関数などの選択の組み合わせを様々に検討し，その中で，サポートベクトルに基づくデータ削減，候補集合を広げた一対多に基づくMMSVM，k-means法を用いたMMSVMの有効性を確認した．

研究成果の概要（英文）：We had proposed the multiobjective multiclass support vector machine (MMSVM), which can maximize the geometric margins for multiclass classification problem. In this study, we developed reduction methods of time and space computational complexity of the MMSVM, and extended the MMSVM into soft-margin models which can learn training data including outliers, where we compare the performances of various combinations from some restriction methods of the constraints of the MMSVM, several solving method of the multiobjective optimization, and two kinds of penalty functions. Among them, we verified that some propose methods, a data reduction method based on support vectors for the soft-margin MMSVM, an improved MMSVM based on the one-against-all method, an MMSVM based on the k-means method are more effective than existing methods.

研究分野： 機械学習，最適化

キーワード： 多クラス識別問題　教師有り学習　サポートベクトルマシン　多目的最適化　マージン最大化　クラスタリング　2次錐計画問題

1. 研究開始当初の背景

サポートベクトルマシン (SVM) は，2 クラス識別のための有力な機械学習モデルであり，求解の容易な凸 2 次計画問題を解くことで，識別関数 $f(x) = w^\top x + b$ （w, b が決定変数）を構成する．その際，幾何マージンと呼ばれる，識別超平面と教師データ間の距離の最大化により，汎化性能（未知のデータに対する識別能力）の高い識別器が得られることが統計学的に示されている．

この 2 クラス SVM を，データ $x^i,\ i = 1, \ldots, l$ とその分類先のクラス $p \in M := \{1, \ldots, m\}$ （教師データ）が与えられたもとで未知データ x を正しく分類する問題に拡張する．x がクラス p に分類された場合の確信度を示す関数 $f^p(x) = (w^p)^\top x - b^p$ を作成し，識別関数

$$f^M(x) = \arg\max_{p \in \{1, \ldots, m\}} f^p(x)$$

により確信度が大きいクラスに識別する．この確信度関数の学習方法として，

(i). **一対多手法** (One-against-all:OA) 1 つのクラスとそれ以外を識別する 2 クラス SVM を m 個学習）[1]

(ii). **一括型手法** (All-Together:AT) m 個の確信度関数を一度に学習 [8]

がある．(i) では，各 2 クラス SVM を 2 次計画問題の求解により得る一方，(ii) では，分類先クラスが p であるデータ集合を I^p とすると，正規化された幾何マージン（関数マージン）$1/\|w^p - w^q\|$ の逆数の和の最小化：

$$\begin{aligned}
\min_{w,b}\ & \frac{1}{2} \sum_{p \in M} \sum_{q > p \in M} \|w^p - w^q\|^2 \\
\text{s.t}\ & (w^p - w^q)^\top x^i + (b^p - b^q) \geq 1,\ i \in I^p, \\
& q \neq p,\ p, q \in M.
\end{aligned}$$

により求解を行う [8]．なお，線形モデルは，カーネル法を用いて非線形に拡張可能であり，汎化性の議論は同様に行える．

筆者は，上記の (i), (ii) の方法において，汎化性能向上のために重要な幾何マージン，つまり，識別関数 $f^M(x)$ により定まる，各クラス間の識別超平面

$$(w^p - w^q)^\top x + (b^p - b^q) = 0$$

と，この識別超平面に最も近いデータ間の幾何学的距離

$$d_{pq}(w, b) = \min_{i \in I_p \cup I_q} \frac{|(w^p - w^q)^\top x^i + (b^p - b^q)|}{\|w^p - w^q\|}, \quad q > p,\ \ p, q \in M$$

が正確に最大化されてないことを指摘した [7]．(i) の 2 クラス識別時の幾何マージン最大化，(ii) での関数マージン最大化が，必ずしも，幾何マージン最大化を意味しないことを示し，従来の 2 手法に対し，幾何マージンを正確に最大化する，多目的多クラス SVM, Multiobjective Multiclass SVM (MMSVM) および MMSVM-OA, を提案し（以下は, MMSVM の定式化），

$$\begin{aligned}
\text{(M)}\quad \max\quad & \frac{\sigma_{12}}{\|w^1 - w^2\|}, \ldots, \frac{\sigma_{m-1,m}}{\|w^{m-1} - w^m\|} \\
\text{s.t.}\quad & (w^p - w^q)^\top x^i + (b^p - b^q) \geq \sigma_{pq}, \\
& i \in I^p, q \in M \setminus \{p\}, \\
& \sigma_{pq} = \sigma_{qp} \geq 1, q > p \in M,
\end{aligned}$$

そのモデルの最適解に，従来法に比べて汎化性能が高いものが多数存在することを確認した [7]．その際，複数の目的関数（幾何マージン）を同時に最適化する必要があるため，多目的最適化として定式化し，そのパレート最適解が 2 次計画問題とほぼ同等の求解の容易さをもつ **2 次錐計画問題** [5] の求解で得られることを示した．さらに，MMSVM, MMSVM-OA に，非線形識別関数を効率的に構成するカーネル法を適用する方法も提案し，線形識別に比べ，幾何的マージン最大化による識別率の向上がより顕著に見られることも確認した [7].

2. 研究の目的

本研究の目的は，筆者らが既に提案した，幾何マージンを正確に最大化することで汎化性を向上させる多クラス用の新しい SVM, MMSVM に対し，

(1) 例外データが含まれる教師データに対しても学習可能なソフトマージンモデルに拡張

(2) 学習時のメモリー使用量・計算時間を削減したモデルの構築，

(3) 問題に適した厳密解法およびメタヒューリスティックス解法の開発

(4) 幾何マージンに加え，より直接的に汎化性を計量可能な指数をメタヒューリスティック法の適合度に加えた新しいモデルの開発

等を目指す.

3. 研究の方法

研究目的 (1)–(4) を達成するため，MMSVM に基づいた多クラス識別 SVM に対して，以下の①～④のポイントについて，各々の手法を選択したときの組み合わせで得られる識別器の性能について検討する（ただし，必ずしもすべての組み合わせ方が実現可能ではない）．以下の各項目で $*$ がついた手法は，すでに検討済みの組み合わせを表す.

① 制約条件の制限方法：計算量削減や高い識別率をもつ解を効率的に求解できるよう，提案した MMSVM の実行可能解を限定する方法を検討

(a) 一対多手法 [1] で得られた重み $\bar{w}^p, p \in M$ を用いた制限その 1*: $w^p = \alpha \bar{w}^p, p \in M$

(b) 一対多手法で得られた重み $\bar{w}^p, p \in M$ を用いた制限その 2: $w^p = \sum_{q \in M \setminus \{p\}} \alpha_p^q \bar{w}^q, p \in M$

(c) 各クラス間で 2 クラス SVM を構成する方法である一対一手法 [2] で得られた重みを用いた制限

(d) 1 クラスを囲い込む境界を SVM を用いて構成する方法である 1 クラス SVM(SVDD)[3] を用いて，そこで得られた情報を用いた制限

(e) クラスタリング（k-means 法）[4] による前処理による制限

② 多目的最適化問題の求解方法：多目的問題 (M) をスカラー化手法に基づいて求解する方法や，汎化性を考慮した目的関数をさらに導入し求解する方法を検討

(I) ε 制約法 $*$：対象とする多目的最適化問題の複数の目的関数から，一つ選択したものを目的関数，残りの目的関数を ① で得られた情報に基づいて選択した定数を用いて制約とした 1 目的 2 次錐最適化問題を求解する方法

(II) クラス間距離などを用いた参照点法：一対一手法やいくつかの手法で得られる各クラスペア間のマージン距離や重心間距離をもとにして，実現可能な（また理想的な）幾何マージン比を推定し，その値を参照点として用い，その比に近い距離を最適解に持つ，近似 1 目的 2 次錐最適化問題を導出し求解する方法.

(III) マージン差最小化を考慮する方法：各クラス間の幾何マージンは，識別超平面と 2 つのクラスデータ中の最近のデータとの距離であたえられ，その超平面と各クラス間の距離は等しいとは限らない．そこで，各クラス間距離の差もより小さくする指標としてマージン差を導入しメタヒューリスティック手法等で求解する方法

③ ソフトマージンモデル：例外データ用のスラック変数のためのペナルティ関数を，多目的最適化問題の目的関数として構成する方法の検討

(i) ソフトマージンを導入しない方法 $*$（ハードマージンのもとでの他の方法の組合せのチェック用）

(ii) マージン最大化と侵入量最小化をすべて個別の目的関数にした方法

(iii) クラスペアごとにマージンと侵入量の逆数の和の最大化を行う方法

④ データの削減法：ソフトマージン法の場合，データ数だけのスラック変数が必要になるため，全データを用いた多目的問題の求解には膨大な計算量が必要となる可能性が高い．そのための前処理 ① で得られた情報を用いてデータ数を削減する方法の検討

(A) データ削減を導入しない方法 $*$

(B) 前処理で得られたサポートベクトル以外のデータもしくは，識別平面の近傍に入らないとみなせるデータの削除

の様々な組み合せの各々を検討する.

4. 研究成果

3. で述べた組み合わせを検討した結果の概要を (1) で示し，(2) では，その中でも良質な性質をもつ 3 つの組み合わせを報告する．また，主に線形化モデルでの定式化を示すが，容易にカーネル法の適用可能な非線形モデルに拡張が可能であり，数値実験では非線形モデルを用いて検証した.

(1) 各組み合わせについての概要

ポイント①について：

(a), (b) 一対多，(c) 一対一，(d)SVDD, (e)k-means 法による実行可能領域限定のための前処理にかかる計算時間・必要なメモリーサイズについては，概ね，(e) $<$ (d) $<$ (c) $<$ (a),(b) の順に小さく，また，得られた識別器の性能は，(b) $>$ (a),(c) $>$ (e) $>>$ (d) の順に高いことを確認した．ただし，一対多の (a),(b)，一対一の (c) については，各々有効となる問題が異なっている.

ポイント②について：

データ数が少ない問題に対し，(III) によるマージン差最小化を加えた定式化により得られた識別器には，高い識別能力をもつものが得られることを確認．データ数を増やすにつれその指標の有効性は明確には確認できなかった．(I), (II) については，(I) により得られる識別器の方が若干識別率が高いものの，①での一対多手法と一対一手法での傾向と同様に，2 つの方法でそれぞれ有効な問題が異なる結果が見られた.

ポイント③ について：

得られる識別率の性能に関して (i) と (ii) での大きな差はなかったが，スラック変数の和を各クラスペアごとに目的関数として扱う (ii) のほうが，パラメータ選定が容易であることを確認した.

(2) 有効性が確認できた代表的な結果

① ソフトマージンへの拡張およびデータ選択法

[組み合わせ (a)-(I)-(ii)-(B)]

(M) をソフトマージンモデルへと拡張する方法として，スラック変数 ξ_{pqi} を用いてすべてのデータを正しく識別する制約を緩和し，マージンを最大化しつつも，スラック変数に対するペナルティ関数

$$\zeta_{pq}(\sigma, \xi) = \frac{\sum_{i \in I_p} \xi_{pqi} + \sum_{i \in I_q} \xi_{qpi}}{\sigma_{pq}}$$

を導入し，それを最小化する問題として定式化し，

$$
\begin{array}{ll}
\text{(SMOA)} & \\
\max\limits_{\alpha,\beta,\sigma} & \dfrac{\sigma_{12}}{\|\alpha^1\bar{w}^1-\alpha^2\bar{w}^2\|},\dots,\dfrac{\sigma_{m-1,m}}{\|\alpha^{m-1}\bar{w}^{m-1}-\alpha^m\bar{w}^m\|} \\
& \qquad -\zeta_{12}(\sigma,\xi),\dots,-\zeta_{pq}(\sigma,\xi) \\
\text{s.t.} & (\alpha^p\bar{w}^p-\alpha^q\bar{w}^q)^\top x^i+(\beta^p-\beta^q)\geq\sigma_{pq}-\xi_{pqi}, \\
& \qquad i\in I_p,\ q>p,\ p,q\in M, \\
& (\alpha^q\bar{w}^q-\alpha^p\bar{w}^p)^\top x^i+(\beta^q-\beta^p)\geq\sigma_{pq}-\xi_{qpi}, \\
& \qquad i\in I_q,\ q>p,\ p,q\in M, \\
& \sigma_{pq}\geq 1,\ q>p,\ p,q\in M, \\
& \xi_{pqi}\geq 0,\ p,q\in M.
\end{array}
$$

とした．この場合，変数の数が $m(5m-2)/2+|\sum_p I_p|(m-1)$ となり，第 2 項が全データ数を含むため計算量が増加する．そこで，教師信号として，一対多手法での学習時，いずれかの 2 クラス SVM でのサポートベクトルとなっているデータのみを用いる方法を提案した．つまり，$S^p := \{i\in I^p|\ x^i$がサポートベクトル $\}$ とし，(SMOA) における I^p を S^p で置き換えた定式化 (RSMOA) を導出した．

このとき，変数の数は $m(5m-2)/2+\sum_p|S^p|(m-1)$ となる．また，削減したデータ数分，制約条件数も減少する．この 2 つの定式化に対して，カーネル法を適用することで非線形モデルも導出可能であり，(I) の ε 制約法によりパレート解を求める．

数値実験として，ベンチマークデータ [6] を用いて，最も安定性のある Cross-Validation により 10 試行での平均識別率を比較した．(RSMOA) は，(SMOA) とほぼ同等の識別率を維持する一方，(RSMOA) の学習時間は，比較的多数データをからなる識別問題の場合，(SMOA) に比して 50 %以下となった．つまり，提案したデータ削減法が，識別率を維持しながら計算時間を大幅に削減できていることを確認した．

② 候補集合を広げた一対多手法に基づく MMSVM，MMSVM-IOA [組み合わせ (b)-(I)-(i)-(A)]

この方法では，一対多手法で得られたクラス p と残りのクラスを識別する識別平面の法線ベクトル $\bar{w}^p,\ p\in M$ を用いて，

$$w^p=\bar{W}\alpha^p,\ \bar{W}=[w^1,\dots,w^m]$$

と限定する．ここで，$\alpha^p\in R^m, p\in M$ は決定変数である．この変数は以下の問題を解くことにより決定する．

$$
\begin{array}{ll}
\text{(IMOA)} & \\
\max\limits_{\alpha,b,\sigma} & \dfrac{\sigma_{12}}{\|\bar{W}\alpha^1-\bar{W}\alpha^2\|},\ \dots,\ \dfrac{\sigma_{m-1,m}}{\|\bar{W}\alpha^{m-1}-\bar{W}\alpha^m\|} \\
\text{s.t.} & (\bar{W}\alpha^p-\bar{W}\alpha^q)^\top\phi(x^i)+(b^p-b^q)\geq\sigma_{pq}, \\
& \qquad i\in I_p,\ q>p,\ p,q\in M, \\
& (\bar{W}\alpha^p-\bar{W}\alpha^q)^\top\phi(x^i)+(b^p-b^q)\geq\sigma_{pq}, \\
& \qquad i\in I_q,\ q>p,\ p,q\in M.
\end{array}
$$

この多目的最適化問題のパレート最適解は，(I) の ε-制約法を用いて導出した，1 目的 2 次錐計画問題

$$
\begin{array}{ll}
(\varepsilon\text{-IMOA}) & \\
\max\limits_{\alpha,b,\sigma} & f_1(\alpha,\sigma)=\dfrac{\sigma_{rs}}{\|\bar{W}\alpha^r-\bar{W}\alpha^s\|} \\
\text{s.t.} & \dfrac{\sigma_{pq}}{\|\bar{W}\alpha^p-\bar{W}\alpha^q\|}\geq\varepsilon_{pq}, \\
& \qquad q>p,\ (p,q)\neq(r,s),\ p,q\in M, \\
& (\bar{W}\alpha^p-\bar{W}\alpha^q)^\top\phi(x^i)+(b^p-b^q)\geq\sigma_{pq}, \\
& \qquad i\in I_p,\ q>p,\ p,q\in M, \\
& (\bar{W}\alpha^q-\bar{W}\alpha^p)^\top\phi(x^i)+(b^q-b^p)\geq\sigma_{pq}, \\
& \qquad i\in I_q,\ q>p,\ p,q\in M.
\end{array}
$$

を解くことで得られる．

①と同様に，ベンチマーク問題を用いた数値実験により，MMSVM-IOA と MMSVM-OA を比較すると，多くの問題に対し，MMSVM-IOA により MMSVM-OA よりも高い識別率をもつ識別器が得られた．計算時間については，少数のクラス数の識別問題に対して 2 つの手法の計算量に大きな差はない一方，4 クラス以上の識別問題に対しては IMMSVM-OA のほうが計算量が多いものの，6 クラスの識別問題においても MMSVM-OA の 2 倍未満であり，計算量に余裕がある場合には比較的識別率の高い MMSVM-IOA が有効であることを確認した．

③ k-means 法を用いた MMSVM，MMSVM-KM [組み合わせ (e)-(II)-(i)-(A)]

手法 (a)–(c) での一対多，一対一手法を前処理に使用する場合，全データを用いる 2 クラス SVM を複数個学習する，もしくは，比較的小さな 2 クラス SVM をクラスペアの組み合わせだけ学習することになり，大規模データに用する場合に必要なメモリー，計算時間が膨大になりうる．そこで，各クラスごとに代表点を選択し，その点の線形和として w^p に制限を加える方法を提案した．

クラス p に対応するデータ $\phi(x^i),\ i\in I_p$ に代表的なクラスタリング手法である k-means 法 [4] を適用し，以下の関数を最小化するようなクラスタ（分割）$I_p^k\subset I_p, k=1,\dots,n^p, I_p=\cup_{k=1}^{n_p}I_p^k$ を求める．

$$E^p=\sum_{k=1}^{n_p}\sum_{i\in I_p^k}\|\phi(x^i)-\psi^{p,k}\|^2$$

ここで，n_p は，各クラスごとに設定するクラスタ数を表し，各クラスタ k の 重心 $\psi^{p,k}$ は，$\psi^{p,k}=\sum_{i\in I_p^k}\phi(x^i)/|I_p^k|$ で与えられる．また，カーネル法も容易に適用可能であり，以下では非線形モデルでの記述を行う．k-means 法は，初期値依存性を持ち，必ずしも E^p を最小化する解が得られるとは限らな

い．そのため，数値実験では，k-means 法を 20 試行し，その中で最小の E^p が得られたクラスタリングでのクラスタ重心を用いた．

クラス p のデータに k-means 法を適用して得られるクラスタの重心 $\psi^{p,k}$, $k=1,\ldots,n_c$ と新たな決定変数 $\gamma^p_{q,k}$, $k=1,\ldots,n_p$, $p,q\in M$ を用いて，MMSVM のクラス $p\in M$ に対する重みベクトル w^p を

$$w^p=\sum_{q=1}^{m}\sum_{k=1}^{n_c}\gamma^p_{q,k}\psi^{q,k}=\Psi\gamma^p$$

と限定する．このとき，識別関数は，

$$f(x)=\operatorname*{argmax}_{p\in M}\left\{(\Psi\gamma^p)^\top x+b^p\right\}$$

と表され，多目的最適化問題は，

$$\begin{array}{ll}
\text{(KM)} & \\
\displaystyle\max_{\gamma,b,\sigma} & \dfrac{\sigma_{12}}{\|\Psi\gamma^1-\Psi\gamma^2\|},\ldots,\dfrac{\sigma_{m-1,m}}{\|\Psi\gamma^{m-1}-\Psi\gamma^m\|}\\
\text{s.t.} & (\Psi\gamma^p-\Psi\gamma^q)^\top\phi(x^i)+(b^p-b^q)\geq\sigma_{pq},\\
 & \qquad i\in I_p,\ q>p,\ p,q\in M,\\
 & (\Psi\gamma^q-\Psi\gamma^p)^\top\phi(x^i)+(b^q-b^p)\geq\sigma_{pq},\\
 & \qquad i\in I_q,\ q>p,\ p,q\in M,\\
 & \sigma_{pq}\geq 1,\ q>p,\ p,q\in M.
\end{array}$$

で与えられる．この問題のパレート解を得るために行列 $\Psi^\top\Psi$ を対角化 $\Psi^\top\Psi=T_\psi\Lambda_\psi T_\psi^\top$ し，その正の固有値に対応する対角行列 $\bar{\Lambda}_\psi$ と行列 $\bar{T}_\psi$ を用いて新たな決定変数を $\zeta^p=\bar{\Lambda}_\psi^{\frac{1}{2}}\bar{T}_\psi^\top\gamma^p$ とおくと以下の (KM2) が得られる．

$$\begin{array}{ll}
\text{(KM2)} & \\
\displaystyle\max_{\zeta,b,\sigma} & \dfrac{\sigma_{12}}{||\zeta^1-\zeta^2||},\ldots,\dfrac{\sigma_{m-1,m}}{||\zeta^{m-1}-\zeta^m||},\\
\text{s.t.} & (\zeta^p-\zeta^q)^\top T_\psi^\top\Lambda_\psi^{-\frac{1}{2}}\bar{\kappa}(x^i)+(b^p-b^q)\geq\sigma_{pq},\\
 & \qquad i\in I_p,\ q>p,\ p,q\in M,\\
 & (\zeta^q-\zeta^p)^\top T_\psi^\top\Lambda_\psi^{-\frac{1}{2}}\bar{\kappa}(x^i)+(b^q-b^p)\geq\sigma_{pq},\\
 & \qquad i\in I_q,\ q>p,\ p,q\in M,\\
 & \sigma_{pq}\geq 1,\ q>p,\ p,q\in M.
\end{array}$$

ここで，$k^p(x)=(\sum_{j\in I_p^1}k(x^j,x)/|I_p^1|,\ \ldots,\ \sum_{j\in I_p^{n_p}}k(x^j,x)/|I_p^{n_p}|)$, である．(MK2) を解くことで，(MK) のパレート最適解が得られる．(MK2) を解くために，手法 (II) の参照点法を用いた 1 目的 2 次錐計画問題を導出する．各クラス p の重心 $g^p=\sum_{i\in I_p}\phi(x^i)/|I_p|, p\in M$ に着目し，各クラスペアの重心間距離を用いて参照点を

$$r_{pq}=\|g^p-g^q\|,\ p,q\in M,$$

と設定する．ここで，この重心間距離の比が，最適な識別関数における幾何マージンの比に比較的近いものと仮定する．次に，参照点とパレート最適解との距離を l_1 ノルムの意味で最小化する 1 目的最適化問題を導き，さらに，その問題に容易に求解可能となるための変形を加え，以下の 2 次錐計画問題 (KMR) を導出する．

$$\begin{array}{ll}
\text{(KMR)} & \\
\displaystyle\min_{\zeta,b,\sigma,l} & \displaystyle\sum_{p\in M}\sum_{q>p}l_{pq}\\
\text{s.t.} & l_{pq}>0,\ q>p,\ p,q\in M,\\
 & r_{pq}\|\ \zeta^p-\zeta^q\ \|-\sigma_{pq}\leq l_{pq},\\
 & \qquad q>p,\ p,q\in M\\
 & \bar{\kappa}(x^i)^\top\Lambda_\psi^{-\frac{1}{2}}T_\psi(\zeta^p-\zeta^q)+(b^p-b^q)\geq\sigma_{pq},\\
 & \qquad i\in I_p,\ q>p,\ p,q\in M,\\
 & \bar{\kappa}(x^i)^\top\Lambda_\psi^{-\frac{1}{2}}T_\psi(\zeta^q-\zeta^p)+(b^p-b^q)\geq\sigma_{pq},\\
 & \qquad i\in I_q,\ q>p,\ p,q\in M,\\
 & \sigma_{pq}\geq 1,\ \ q>p,\ \ p,q\in M.
\end{array}$$

この問題は凸計画問題であり，内点法による効率的な求解か可能である．一方で，この変形のため，(KM) のパレート最適解を近似的に求めていることになる．

表 1 ：4 つの方法での平均識別率（%）

問題	(M)	(MOA)	(IMOA)	(MK)
Car	**99.48**	99.36	99.36	98.21
Balance	97.72	99.68	99.52	**99.84**
Glass	**96.69**	95.28	**96.69**	**96.69**

表 2 ：4 つの方法での平均学習時間（秒）

問題	(M)	(MOA)	(IMOA)	(MK)
Car	1019.6	34.869	35.494	**7.702**
Balance	15.727	3.155	2.973	**0.789**
Glass	5.414	0.731	0.802	**0.530**

①，②と同様にベンチマーク問題に対する 10 試行平均識別率を比較する数値実験を行い，MMSVM (M), MMSVM-OA (MOA), MMSVM-IOA (IMOA), MMSVM-KM (MK) を比較した．表 1 ，2 は，その中で，RBF カーネルを用い，クラスタ数として $n_p=\lfloor 0.1|I_p|\rfloor$ と設定した場合の 3 つの問題に対する平均識別率と平均学習時間を比較したものであり，MMSVM-KM が学習時間を削減しながら，識別率をほぼ維持もしくはよりよい値をもつ識別器が得られていることが分かる．この結果から，k-means 法による前処理やその結果を用いた識別超平面の法線ベクトルの限定，クラスタ重心間距離を用いた参照点法による (KM) の近似的求解が有効であることを示している．

④ 成果の位置づけと今後の展望

ハードマージン MMSVM の識別能力をほぼ維持しつつ変数を削減する方法や，データを削減しつつソフトマージンモデルに拡張する方法の開発により，より大規模なデータに提案多クラス SVM が適用可

能になり，SVM そのものの適用先をより広げることに貢献したといえる．今後の展望として，この研究で得られた結果を生かした大規模な実際の識別問題への適用や，幾何マージン自身の近似により，パレート解集合からより効率的に解を求めることで更なる求解の効率化をはかる研究などが考えられる．

＜引用文献＞

1) L. Bottou, C. Cortes and etc., Comparison of classifier methods: A case study in handwriting digit recognition, *in Proc. Int. Conf. Pattern Recognition* (1994) 77–87.
2) U. Kressel, Pairwise classification and support vector machines, in *Advances in kernel methods – Support vector learning*, (Eds) B. Schölkopf, C. Burges, and A. J. Smola, MIT Press, Cambridge, pp. 255–268, (1999)
3) D. Lee and J. Lee, Domain description support vector classifier for multi-classification problems, *Pattern Recognition*, Vol. 40, pp. 41–51 (2007)
4) J. B. MacQueen, Some methods for classification and analysis of multivariate observations, *Fifth Berkeley symposium on mathematics, statistics and probability*, 281–297 (1967)
5) H.D. Mittelmann, An independent benchmarking of SDP and SOCP solvers, *Mathematical Programming*, Ser. B, 95 (2003) 407–430. *benchmarkprob.*
6) J. K. Martin and D. S. Hirschberg, Small sample statistics for classification error rates I: Error rate measurements, Technical Report no. 96–21, Department of Information and Computer Science, University of California, Irvine (1996)
7) K. Tatsumi, T. Tanino and K. Hayashida: Multiobjective Multiclass Support Vector Machines Maximizing Geometric Margins, *Pacific Journal of Optimization*, Vol. 6, No. 1, pp. 115-140 (2010)
8) V. N. Vapnik, *Statistical Learning Theory*, Wiley, NewYork (1998)

5. 主な発表論文等

【雑誌論文】(計 1 件)

① Keiji Tatsumi and Tetsuzo Tanino, Support vector machines maximizing geometric margins for multi-class classification, TOP: An Official Journal of the Spanish Society of Statistics and Operations Research，査読有，Vol. 22, 2014, pp. 815 － 840, doi:10.1007/s11750-014-0338-8

【学会発表】(計 7 件)

① 巽 啓司, クラスタリングにより識別器候補を限定した多目的マルチクラスサポートベクトルマシンの計算量削減法, 2016 年 3 月 11 日, 計測自動制御学会 第 43 回知能システムシンポジウム, 室蘭工業大学（北海道，室蘭市）

② Keiji Tatsumi and Tetsuzo Tanino, A Complexity Reduction Method for the Multiobjective Multiclass Support Vector Machine, 2014 年 7 月 14 日, The 20th Conference of the International Federation of Operational Research Societies, Centre de Convencions Internacional de Barcelona - CCIB, Barcelona(Spain).

③ 巽 啓司，一対多手法に基づいた幾何マージン最大化手法のソフ トマージンモデルへの拡張と学習時間削, 2014 年 3 月 7 日, 日本オペレーションズ・リサーチ学会 2014 年春季研究発表会，大阪大学豊中キャンパス（大阪，豊中市）

④ 巽 啓司，一対一手法に基づく多クラス多目的サポートベクトルマシンに対する参照点を利用した学習法，2013 年 9 月 12 日, 日本オペレーションズ・リサーチ学会 2013 年秋季研究発表会，徳島大学 常三島キャンパス（徳島，徳島市）

⑤ 巽 啓司, 海老田 祐司，谷野 哲三, 一対一手法に基づいた多クラス多目的サポートベ クトルマシンの学習, 2013 年 5 月 15 日, 第 57 回システム制御情報学会研究発表講演会, 兵庫県民会館（兵庫，神戸市）

⑥ Keiji Tatsumi, Multiclass Support Vector Machines Based on Multiobjective Optimization, 2012 年 10 月 24 日, 西安交通大学電子情報工学部（招待講演）西安 (中華人民共和国)

⑦ Keiji Tatsumi, A soft-margin multiobjective multiclass support vector machine with unified objective functions for maximizing geometric margins and minimizing slack variables, 2012 年 09 月 26 日, The 15th Czech-Japan Seminars on Data Analysis and Decision Making under Uncertainty, 大阪大学待兼山会館（大阪, 豊中市）

6．研究組織

(1) 研究代表者

巽　啓司 (Tatsumi Keiji)

大阪大学・大学院工学研究科・准教授

研究者番号：30304017