

令和 4 年 6 月 3 日現在

機関番号：14501

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K04441

研究課題名(和文) マージン分布の制御による汎化能力識別器の開発

研究課題名(英文) Development of high generalization classifiers by controlling margin distributions

研究代表者

阿部 重夫 (Abe, Shigeo)

神戸大学・工学研究科・名誉教授

研究者番号：50294195

交付決定額(研究期間全体)：(直接経費) 2,300,000円

研究成果の概要(和文)：従来のサポートベクトルマシン(SVM)の汎化能力を向上する方式としてSVMと複雑度最小マシン(MCM)とを融合する複雑度最小SVM(MSVM)を開発した。

- 1) 標準のSVM(L1 SVM)、線形計画SVM(LP SVM)とMCMを組み合わせた5種類のMSVMを開発した。これらのマシンはSVMの最小マージンを最大化するのに加えて最大マージンを最小化するものである。L1 SVMベースのマシンに対しては新たに、最小マージン最大化と最大マージン最小化を交互に行う学習法を開発した。
- 2) ベンチマークデータによる評価の結果、ML1v SVMが平均的にL1 SVMより汎化能力が向上することを示した。

研究成果の学術的意義や社会的意義

パターン認識アルゴリズムは多くの分野で活用されており、未知のデータに対する高い識別能力、すなわち汎化能力が求められている。このため「多くの分野でSVMより格段に汎化能力の高い識別器が存在しうるか」という学術的な問いに肯定的な答えを求めべく研究を行った。今回開発したMSVMはその解となりうるものとして学術的意義も高く、また産業界への応用の上でも貢献しうるものである。

研究成果の概要(英文)：To realize a higher generalization ability than that of the conventional support vector machines (SVMs), we developed minimum complexity support vector machines (M SVMs) that fuse conventional SVMs and minimum complexity machines (MCMs) and obtained the following results:

- 1) Combining standard SVMs (L1 SVMs) and linear programming SVMs (LP SVMs) with MCMs, we developed SL1 SVMs, ML1 SVMs, ML1v SVMs, SLP SVMs, and MLP SVMs. These machines are considered maximizing the minimum margin and minimizing the maximum margin. We developed new training methods for SL1 SVMs, ML1 SVMs, and ML1v SVMs: alternately maximizing the minimum margin and minimizing the maximum margin.
- 2) According to the computer experiments using benchmark problems, ML1v SVMs show the best generalization ability among the developed classifiers and L1 SVMs

研究分野：システム制御

キーワード：パターン認識

## 1. 研究開始当初の背景

パターン認識技術は、電気電子、情報の分野のみならず、医療診断など、さまざまな分野で活用されている。パターン認識システムを開発する上で高い汎化能力（未知のデータに対する識別能力）を実現することが重要であるが、パターン認識方式の1つとしてサポートベクトルマシン(SVM)を中心とするカーネル法の研究が活発に行われている。このような研究の状況を踏まえ、研究代表者は「**多くの分野でSVMより格段に汎化能力の高い識別器が存在しうるか**」という学術的な問いに肯定的な答えを求めて研究を進めている。

SVMでは分離超平面からのデータの距離をマージンと定義して、最小マージンの最大化を図っており、マージン分布には着目していない。このような場合、マージン分布を考慮して識別器を構成すれば汎化能力が向上する可能性がある。汎化能力を推定するVC(Vapnik-Chervonenkis)次元を直接最小化する識別器として最小複雑度マシン(MCM)が開発されている。MCMは線形計画法(LP)で解くことができるという利点があり、また多くの問題でSVMより汎化能力が高いことが示されている。そこで、この方式が研究代表者が求める識別器の原型となりうるか否かの検討を開始した。

その結果、**MCMではregularization項を含まないために収束に問題があることが分かった**。これは、LP SVM、L1 SVM（標準のSVM）などのregularization項を含むSVMと融合することにより、収束の問題が解決し、SVMよりも汎化能力が向上する可能性があることを示している。このため、**MCMとLP SVMあるいはL1 SVMなどのSVMと融合したMLP SVMあるいはML1 SVMを開発することにより、SVMより汎化能力の高い識別器の開発に着手した**。

## 2. 研究の目的

「**多くの分野でSVMより格段に汎化能力の高い識別器が存在しうるか**」という根本的な問いに基づいて、研究を進めてきたが、**MCMとSVMの融合**を根本的な問いを解く新しい切り口として、研究を進める。

**MCMとSVMの融合**の研究により、SVM以上の汎化能力をSVMと同等の学習時間で実現できる可能性があり、研究が実現できれば、学術的価値とともに、応用展開も加速することができる。

## 3. 研究の方法

**MSVMアーキテクチャの開発**：MCMとLP SVMあるいはL1 SVMと融合して、複雑度が最小のMLP SVMあるいはML1 SVMのアーキテクチャを開発する。

**ML1 SVMの高速学習方式の開発**：MLP SVMは線形計画法で学習できるが、ML1 SVMはL1 SVMの学習法では学習できない。L1 SVMではSMO (Sequential Minimal Optimization)とNewton法を融合したFusing Sequential Minimal Optimization and Newton's Method for Support Vector Training, Shigeo Abe, Int. J. Machine Learning and Cybernetics, 7 (3), 345-364, 2016 の方法で高速に学習できるように、ML1 SVMで上記の学習法を拡張すべく検討を進め、高速学習方式を開発する。

**MSVMの性能評価**：ML1 SVMの高速学習方式をプログラミングして、これとLPでプログラミングしたMLP SVMを種々のベンチマークデータで、汎化能力および学習時間の観点から評価する。これと従来のLP SVM, L1 SVMなどと比較を行う。このとき、複数の教師、テストデータのペアで教師データを用いた学習とテストデータによる評価を行い、**ML1 SVM、MLP SVMが従来のSVMよりも統計的有意差があるかの検定を行う**。

**MSVMの理論的な汎化能力の解析**：既発表の論文ではMCMの汎化能力の解析が不十分であったので、理論的な解析を行う。VC次元は超球でデータの存在範囲を近似することにより評価してい

るが、MCMは決定平面方向が開いた超直方体であるため、そのままの形では汎化能力の解析が行えない。**VC次元をベースに汎化能力の評価が行えるか検討し、行えない場合は新しい評価方法を導出する。**

**MSVMの拡張：** の評価で、MLP SVM、ML1 SVMの利点、欠点が明らかになるので、これにより、改良点の洗い出しを行い、必要な改良を加える。さらに、文献調査により汎化能力の向上法の調査を精力的に行い、MSVMをさらに拡張する方式、あるいはこれとは異なる方式の検討を進める。

#### 4. 研究成果

**MSVMアーキテクチャの開発：** MCMとLP SVMあるいはL1 SVMと融合して、複雑度が最小のMLP SVMあるいはML1 SVMの以下のアーキテクチャを開発した。

- a) MLP SVM: 分離超平面からの教師データへの距離に上限を設けて、上限の最小化(すなわち最大マージンの最小化)の条件をLP SVMに導入したMLP SVMを開発した。
- b) SLP SVM: 外れ値が教師データに含まれるときに、最大マージンを最小化すると汎化能力に影響を及ぼすことが考えられるので、最大マージンが上限を超えることを許すSoft MLP SVM (SLP SVM)を開発した。
- c) ML1 SVM: L1 SVMに最大マージンを最小化する制約を加えたML1 SVMを開発した。
- d) ML1v SVM: ML1 SVMにおいて、最大マージンを最小化する項に関する不等号制約を等号化することにより、ML1 SVMを簡略化したML1v SVMを開発した。
- e) SL1 SVM: ML1 SVMの最大マージンを最小化する制約をsoft化するSoft ML1 SVM (SL1 SVM)を開発した。

**ML1 SVMの高速学習方式の開発：** MLP SVM、SLP SVMは線形計画法で学習できる。ML1 SVM、ML1v SVM、SM1 SVMはL1 SVMの学習法を拡張する必要がある。このため、学習を最小マージンを最大化する部分問題と最大マージンを最小化する部分問題の2つに分割して、この部分問題を交互に学習する方法を開発した。各々の部分問題はL1 SVMの学習で用いられるSMO (Sequential Minimal Optimization)を用いればよいが、さらに学習の高速化を図るために、SMOとNewton法を融合した方法で高速に学習できるようにした。

**MSVMの性能評価：** 開発した5種類のMSVMを2クラス問題と多クラス問題で、汎化能力および学習時間の観点から評価した。2クラス問題に対しては**MSVMと従来手法の認識率で統計的検定を行ったが、ほぼ同程度かよい結果が得られた。**以下では、多クラス問題に対する結果を示す。

識別器の比較をより公正なものとするために、教師データとテストデータとを用意し、教師データを用いて、各識別器のパラメータ値は5重交叉検定により決めた。さらに決められたパラメータを設定して、教師データを用いて識別器を学習し、テストデータで認識率を計算した。開発識別器との比較には、標準のSVMであるL1 SVM、線形計画SVM(LP SVM)、複雑度最小マシン(MCM)、最小自乗SVM(LS SVM)および非制約大マージン分布マシン(ULDM)を用いた。

ベンチマークデータとしては、車のナンバープレートから収集した数字データ、ひらがなデータ(H50, H13, H105)、UCI RepositoryからのThyroid(甲状腺)、Satimage(衛星画像)、Letter(英文字)のデータ、白血球の分類問題の血球データ、米国郵政省の郵便番号データ(USPS)、手書き数字データMNISTを用いた。ただしMNIST(r)はMNISTの教師データのデータ数が大きいために、教師データとテストデータを逆転したものである。MNISTは識別器によっては、交叉検定の時間がかかりすぎるために、結果を得られていない。

表1に11個のベンチマークデータに対する各識別器のテストデータの認識率を示す。表において太字は各問題における最大の認識率、下線は最低の認識率を示す。また、平均と全体平

均は、各々MNISTを除いた、および含めた平均である。

表よりML1v SVMとULDMが最良の認識率を4回示し、ULDMの平均の認識率はML1v SVMより劣っていた。これは、数字と甲状腺問題で認識率が低いためで、ML1v SVMよりも問題依存性が大きいことが分かる。

LP SVMは評価した識別器で最も悪い識別器であった。LPベースの識別器でMCMはLP SVMより認識率は向上したが、MLP SVM, SLP SVMよりも平均性能で劣った。MLP SVM, SLP SVMは平均値が同じで同等の性能といえ、LP SVMに対して最大マージンを最小化することにより、汎化能力が一般的に向上することが示された。また最大マージンの最小化をsoft化したSLP SVMの効果はMLP SVMに対してあまり見られなかった。

L1 SVMベースの識別器では、平均値からはML1v SVMが最良でML1 SVMがそれに続き、SL1 SVMは3番目であったが、L1 SVMよりもよかった。これよりML1 SVM、ML1v SVMをsoft化したSL1 SVMの効果はあまり見られなかった。しかし、最大マージンの最小化の効果はSL1 SVMでも確かめられた。以上の結果より、**ML1v SVMがL1 SVMの認識性能を平均的に上回る可能性がある識別器の一つと結論できる。**

表1 ベンチマークデータに対する各識別器の認識率 (%)

ベンチマーク	SL1	ML1	ML1v	L1	SLP	MLP	MCM	LP	LS	ULDM
数字	99.39	<b>99.76</b>	<b>99.76</b>	<b>99.76</b>	99.51	99.27	99.51	99.27	<u>99.15</u>	99.39
甲状腺	97.26	97.23	97.23	97.26	97.64	<b>97.81</b>	94.02	97.52	<u>95.39</u>	95.57
血球	93.65	93.32	93.65	93.19	93.39	93.36	94.07	<u>92.52</u>	94.23	<b>94.61</b>
H50	98.92	99.20	99.11	98.98	99.02	99.22	98.76	<u>98.11</u>	<b>99.48</b>	98.92
H13	99.86	99.88	<b>99.92</b>	99.76	99.82	99.90	99.76	<u>99.50</u>	99.87	99.90
H105	<b>100.00</b>	<u>99.99</u>	<b>100.00</b>	<b>100.00</b>						
Satimage	91.55	91.80	91.85	91.90	90.95	91.00	<u>90.55</u>	90.95	91.95	<b>92.25</b>
USPS	<b>95.67</b>	95.47	95.37	95.27	95.47	95.32	95.52	<u>94.32</u>	95.47	95.42
MNIST(r)	96.90	96.96	96.95	96.55	96.59	96.40	96.86	<u>94.96</u>	96.99	<b>97.03</b>
Letter	97.85	97.98	<b>98.03</b>	97.70	97.78	97.88	97.65	<u>96.03</u>	97.88	97.75
MNIST	<b>98.81</b>	98.78	98.78	98.56	---	---	---	---	---	---
平均	97.10	97.16	<b>97.18</b>	97.04	97.02	97.02	96.67	96.32	97.04	97.08
全体平均	97.26	97.31	<b>97.33</b>	97.17	---	---	---	---	---	---

SL1 SVM、ML1 SVM、ML1v SVMではL1 SVMに比べて学習の収束が遅くなるのが問題である。図1にSL1 SVMとL1 SVMの血球問題のクラス2と3のデータに対する収束の様子を示す。縦軸は目的関数の値で横軸は収束回数を示す。L1 SVMが収束した後もSL1 SVMの収束が続き、しかも目的関数の増加が少なく、収束のスピードが遅くなっていることが分かる。サポートベクトルの増加が原因の一つであるが、SMOとニュートン法を融合しても高速化が十分達成できなかった。高速化に関しては今後の課題としたい。

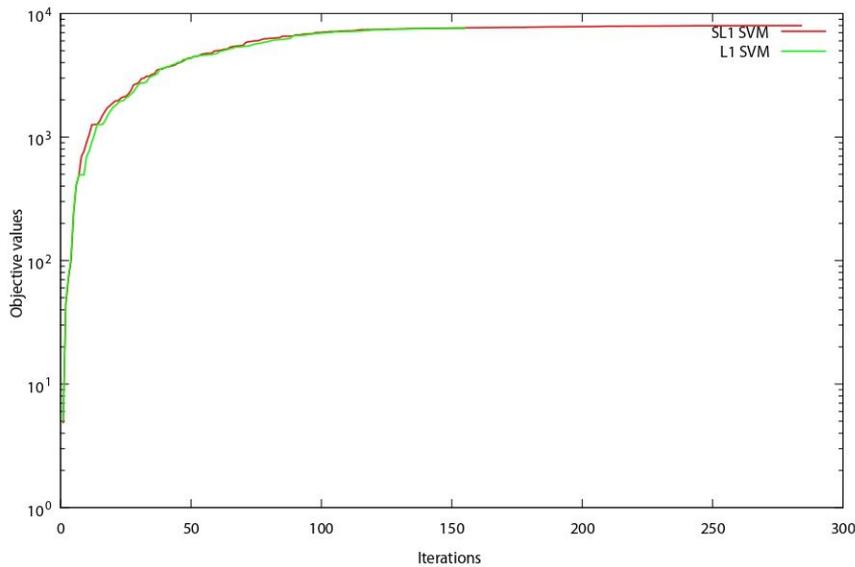


図1 血球問題のSL1 SVMとL1 SVMの収束の状況

**MSVMの理論的な汎化能力の解析**：MCMでは、教師データを覆う半径最小の超球の中心が原点にあると仮定して、VC次元の最小化によりMCMを導出しているが、通常、超球の中心は原点に存在しないのでこの仮定は成り立たない。そのため、この仮定が問題であると考えてきたが、結論が分離超平面からのデータの距離を最小にすることであるから、超球の中心が原点という仮定に依存しない結論であり、中心が原点いなくても成り立つように思える。今後もう少し検討を加えるが、実験結果でも超平面からの距離を最小化することにより、汎化能力が向上することが多くみられるために、仮定がなくても成り立つことを補強していると考えられる。

**MSVMの拡張**： の評価により、MLP SVM、ML1 SVMによりLP SVM、L1 SVMより汎化能力が向上することが確かめられた。しかしながら、学習が場合により遅くなることが明らかになった。学習の高速化に関しては今後の課題としたい。

SVMの変形としてよく用いられるLS SVMは最小自乗法により、分離超平面の周りに教師データが集まるように学習するために最小マージンの最大化とともに最大マージンの最小化と同等な考え方に基づいていると考えられる。しかしながら、表1で示したように、L1 SVMと比較して問題依存性が強い。このために、LS SVMに最大マージンの最小化をさらに導入するアーキテクチャを開発した。この方式で汎化能力が上がるかどうかの評価は今後の課題としたい。

また、SVMの変形として複数のカーネルの線形和により汎化能力の向上を図る多重カーネルSVMがある。このSVMに対しても最大マージンの最小化でさらに汎化能力が向上するかどうかは今後の課題としたい。

今後、以上の課題の検討を進め、L1 SVMに対して有効な汎化能力の向上方式の確立を進めたい。

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Shigeo Abe	4. 巻 1
2. 論文標題 Soft Upper-bound Minimal Complexity LP SVMs	5. 発行年 2021年
3. 雑誌名 Proc. 2021 International Joint Conference on Neural Networks (IJCNN)	6. 最初と最後の頁 1-7
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IJCNN52387.2021.9533540	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shigeo Abe	4. 巻 1
2. 論文標題 Minimal Complexity Support Vector Machines	5. 発行年 2020年
3. 雑誌名 IAPR Workshop on Artificial Neural Networks in Pattern Recognition	6. 最初と最後の頁 89-101
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-030-58309-5_7	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shigeo Abe	4. 巻 9
2. 論文標題 Minimal Complexity Support Vector Machines for Pattern Classification	5. 発行年 2020年
3. 雑誌名 Computers	6. 最初と最後の頁 88
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/computers9040088	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Abe Shigeo	4. 巻 1
2. 論文標題 Analyzing Minimal Complexity Machines	5. 発行年 2019年
3. 雑誌名 2019 International Joint Conference on Neural Networks (IJCNN)	6. 最初と最後の頁 1-8
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IJCNN.2019.8852084	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------