対称錐計画法を用いた判別問題の解法

北原　知就

(東京工業大学大学院社会理工学研究科経営工学専攻　現所属・同大学院社会理工学研究科経営工学専攻)

指導教員　水野　眞治 教授

1.　はじめに

本研究では，最近 Lanckriet ら[1]によって提案された判別問題に対するミニマックスアプローチを発展させ，実用性を向上させることを目的とした．彼らのミニマックスアプローチについては2節で概説するが，これが二つの群を一次の関数を用いて判別する問題を対象にしているのに対し，本研究では

1. 二つの群を二次の関数で判別
2. 三つ以上の群を一次の関数で判別
3. 三つ以上の群を二次の関数で判別

という三つの問題を扱った．

Lanckriet らのミニマックスアプローチを本研究で扱うような広い問題に対して拡張することは容易ではなく，既存の結果はほとんど知られていない．本研究では様々な工夫により，上記の問題を数理計画問題として定式化することに成功した．得られた問題はそのままでは解くのが困難であるが，問題の特殊な性質に注目し，対称錐計画法を利用して効率的に解くことができることを示した．このことにより，上記の各問題に対して実際に判別関数を求めることができるようになった．さらに，計算実験により，本研究の判別手法の有効性を示した．

2.　ミニマックスアプローチ

平均ベクトルと分散共分散行列のみがわかっている二つの群を，一次の関数の符号によって判別する問題を考える．このとき，なるべく精度が高い判別関数を決定したいというのは自然な考え方である．しかし，各群の平均ベクトルと分散共分散行列しかわかっていない場合，判別の精度は各群がどのような分布形を取るかによって変わってくる．したがって，ある分布形のもとで高い精度を示す判別関数が，別の分布形のもとでは非常に低い精度を示すということが起こりうる．このような分布形による判別精度の不確実性に対し，Lanckriet ら[1]は，様々に分布形を動かしたときの最も悪い場合の（マックス）誤判別率が最小（ミニ）になるような判別関数を採用することを提案した．Lanckriet らは，このような判別関数を求める問題が二次錐計画問題として表され，解が効率的に計算できることを示した．彼らのアルゴリズムは現在ミニマックス確率マシンと呼ばれている．Lanckriet らは，ミニマックス確率マシンによる判別関数が現実の問題に対して有効に機能することを明らかにした．

3.　二群二次判別

いま，平均ベクトル，分散共分散行列が既知である二群を二次関数の符号によって判別する問題を考える．本研究では様々な分布を考えたときの最悪の場合の誤判別率を最小にする二次の判別関数を求める問題を，非線形計画問題として定式化した．

3.1　パラメトリック実行可能性問題の解法

本研究では，得られた問題が次の三つの特殊な性質を持つことを明らかにした．

1. 目的関数は判別関数の最悪の場合の正判別率 α であり，ミニマックスアプローチはこれを最大化とすることと等価である．また，$\alpha\in[0,1)$ である．
2. 制約条件は α を一つ固定すると半正定値制約になる．
3. ある α' に対して制約が実行不能であれば，任意の $\alpha\in[\alpha',1)$ に対して制約は実行不能となる．

このような問題を，本研究ではパラメトリック半正定値計画実行可能性問題と呼ぶ．問題を直接解くことは困難であるが，上記の三つの性質により，二分探索法を用いて簡単に近似最適解を求めることができる．

4.　三つ以上の群の判別

平均ベクトル，分散共分散行列が既知である $m(\geq 3)$ 群があり，新たに得られたサンプルを m 個の関数 $f_i(i=1,\cdots,m)$ によって，その値が最も小さい群に判別する問題を考える．ここでは，$f_i(i=1,\ldots,m)$ とし

て，一次または二次の関数を考える．このとき，各群の分布形による判別精度の変動を考慮に入れ，最も悪い場合でも精度が悪くならないような判別関数を求めたいとする．

紙数の都合上，詳細は割愛するが，本研究の結果，

1．一次関数による判別問題⇒パラメトリック二次錐計画実行可能性問題
2．二次関数による判別問題⇒パラメトリック半正定値計画実行可能性問題

へとそれぞれ帰着させることができた．それぞれの問題は3.1節で述べたように，二分探索法により効率的に近似最適解を求めることができる．

5．数値実験

数値実験により得られた主な知見は次の二点である．

1．二群二次判別

ミニマックス確率マシンによる一次の判別関数と，本研究による二次の判別関数が一致する，という性質が成り立つことが示唆された．このことの一般性の検証は今後の課題である．

2．三つ以上の群の判別

UCI Machine learning Repository[1] より，iris, wine というベンチマーク問題を採用し，本研究の判別手法を適用した．それぞれの問題の詳細を表1に示す．各問題に対して，本研究の判別手法によって判別した結果を表2に示す．表2は，例えばirisにおいて本研究による一次の判別関数を用いると，第1群のデータはすべて正しく判別され，正判別率が1であることを表している．表2から今回採用した問題に対して本研究の判別手法が高い判別率を示していることがわかる．

表1 ベンチマーク問題の詳細

問題	データ数	群の数	説明変数の数
iris	150	3	4
wine	178	3	13

表2 判別結果

	1群	2群	3群
iris(一次)	1	0.94	0.98
iris(二次)	1	0.94	0.98
wine(一次)	1	1	1
wine(二次)	1	1	1

6．まとめ

平均ベクトルと分散共分散行列がわかっている二つの群があるとき，Lanckriet ら[1]はそれらを最悪の場合の（マックス）誤判別率を最小に（ミニ）するような一次の関数によって判別することを提案した．

本研究はLanckrietらのミニマックスアプローチをより広い問題に対して広げることを目的とした．本研究ではそのような問題を数理計画問題として定式化することに成功した．さらにその特殊な性質と対称錐計画法を利用した問題の効率的な解法を示した．このことによって，上記の問題に対して実際に判別関数を求めることができるようになった．さらに，計算実験により，本研究の判別手法が有効であることを確認した．

参考文献

[1] G. R. G. Lanckriet, L. El Ghaoui, C. Bhattacharyya and M. I. Jordan. A Robust Minimax Approach to Classification. *Journal of Machine Learning Research*, 3: 555-582, 2002.

[1] http://www.ics.uci.edu/~mlearn/MLRepository.html