Separation by ellipsoids using semidefinite programming

François Glineur

Faculté Polytechnique de Mons
Journée du 3ème cycle F.N.R.S.
Han-sur-Lesse

8^{th} May 1998.

Plan

- Introduction
- Semidefinite Programming
- Model
- Results
- Conclusions

$\boxed{\text{Introduction}}$

Problem

Let $\mathcal{C}$ be a set of n classes. Each class C_i is a homogeneous set of entities. Each entity has d numerical characteristics.

$$\mathcal{C} = \{C_i \in \mathbb{R}^d,\ i = 1, 2, \ldots, n\}$$

We want to separate these classes, i.e. partition $\mathbb{R}^d$ such that each class belongs to a different set of the partition.

Classification

1. Group some well-known entities into homogeneous classes (learning set)

2. Separate those classes

3. Use this partition to classify other entities (generalization)

Examples : Character recognition, weather prediction, medical diagnostic, etc.

Approach

To separate the classes, we compute n d-dimensional ellipsoids E_i such that

- $C_i \subset E_i \ \forall i$

- $C_j \cap E_i = \emptyset \ \forall i \neq j$

These ellipsoids can be computed one at a time $\Rightarrow n$ *independent* subproblems

The problem of computing the *best* ellispoid can be modelled as a semidefinite program

N.B. We do not require these ellipsoids to be disjoint.

Optimization objective

Feasible ellipsoid $\Leftrightarrow$ Separating ellipsoid

Best ellipsoid $\Leftrightarrow$ Maximal separation

Semidefinite Programming

Definitions

$\mathbb{S}^n$ = space of $n \times n$ symmetric matrices

Inner product on $\mathbb{S}^n$: $A \bullet B = \text{trace}(\mathsf{A}^\mathsf{T}\mathsf{B})$

A is positive definite $\Leftrightarrow A \succ 0$
$\Leftrightarrow$ all eigenvalues $> 0 \Leftrightarrow u^T A u > 0 \; \forall u \neq 0$

A is positive semidefinite $\Leftrightarrow A \succeq 0$
$\Leftrightarrow$ all eigenvalues $\geq 0 \Leftrightarrow u^T A u \geq 0 \; \forall u$

Löwner order : $A \succeq B \Leftrightarrow A - B \succeq 0$ and $A \succ B \Leftrightarrow A - B \succ 0$

Formulation

Min $C \bullet X$ s.t. $A_i \bullet X = b_i$ and $X \succeq 0$
with $X \in \mathbb{S}^n$ and $i = 1, 2, \ldots, m$

Data : $b \in \mathbb{R}^m$, $C, A_i \in \mathbb{S}^n$, n^2 variables, m constraints, one objective. Linear programming is a special case.

Dual

Max $b^T y$ s.t. $\sum_{i=1}^{m} y_i A_i \preceq C$
with $y \in \mathbb{R}^m$

Algorithms

Convex problem : interior points methods, very efficient (low parameter self-concordant barrier, Nesterov & Nemirovsky)

Primal-dual scheme (generalization of infeasible primal-dual path-following methods for linear programming $\Leftrightarrow$ same concepts : logarithmic barrier function, central path, duality gap, etc.)

No feasible start needed.

Complexity : ϵ-solution reached in $O(\sqrt{n} \ln \frac{1}{\epsilon})$ iterations (same for linear programming *with* n *variables !*)

Applications

- Exact models : Eigenvalue optimization. Convex quadratic optimization. Robust control. Structural optimization. Statistics.

- Relaxations : Nonconvex quadratic optimization. Combinatorial optimization (with bounds !) : max-cut, boolean quadratic optimization. General 0-1 linear programming.

- Very active field

$\boxed{\text{Model}}$

Ellipsoid

$$E \subset \mathbb{R}^n \equiv \{x \in \mathbb{R}^n | (x-c)^T D(x-c) \leq 1\}$$

where $c \in \mathbb{R}^n$ and $D \succ 0 \in \mathbb{S}^n$

Constraints

Let $\{x_i^+\}$ be the points to be included in the ellipsoid and $\{x_i^-\}$ the points to be excluded. We ask for $(x_i^+ - c)^T D(x_i^+ - c) \leq 1$ and $(x_i^- - c)^T D(x_i^- - c) \geq 1 + t$ and try to maximize t in order to obtain the best separation.

- If the points are separable by an ellipsoid, we find an optimal solution with $t > 0$

- If the points are not separable by an ellipsoid, we find an optimal solution with $t < 0$ but minimize the uncertain zone

Formulation

Use a trick : letting

$$Z = \begin{pmatrix} s & z^T \\ z & D \end{pmatrix}$$

it is possible to prove that the problem is equivalent to

min $-t$
s.t. $Z \bullet Y_i^+ + r_i^+ = 1$
and $Z \bullet Y_i^- - r_i^- = 1 + t$
with $Z \succeq 0, r_i^+ \geq 0, r_i^- \geq 0$

which is a semidefinite program. We have $Y_i^\pm = (-1; x_i^\pm)^T(-1; x_i^\pm)$ and we can recover c with $D^{-1}z$.

N.B. There is an additional constraint to avoid the case $D = 0$ when points are not separable

Results

Implementation

We used Matlab 5.2 and a semidefinite package, SDPPack 0.9 by F. Alizadeh, J.-P. Haeberly, M.V. Nayakkankuppam, M.L. Overton and S. Schmieta

Benchmark

Data	Cl.	Pts	dim	10%	50%	Ref
Iris	3	150	4	85.6	92.7	
Breast	2	683	9	94.9	96.2	96.9
Housing	2	506	14	85.3	95.6	84.0
Diabetes	2	768	8	67.8	64.7	71.9

Conclusions

- Main idea : use ellipsoids to separate classes

- Compute ellipsoids with semidefinite programming

- Very encouraging results

- Further research
  - Better handling of infeasibility
  - Better exploitation of ellipsoids
  - Variable subset selection
  - Clustering (nonconvex)