# Bollettino Unione Matematica Italiana

*Sezione A – La Matematica nella Società e nella Cultura*



Annabella Astorino


## Programmazione matematica e classificazione


*Bollettino dell'Unione Matematica Italiana, Serie 8*, Vol. **3-A**—*La Matematica nella Società e nella Cultura* (2000), n.3, p. 263–266.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2000_8_3A_3_263_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2000.

Bollettino U. M. I.
La Matematica nella Società e nella Cultura
Serie 8, Vol. III-A, Dicembre 2000, 263-266

# Programmazione matematica e classificazione.

ANNABELLA ASTORINO

L'obiettivo della tesi è analizzare i diversi approcci di programmazione matematica utilizzati per risolvere problemi di classificazione e proporne di nuovi.

I problemi di classificazione presi in esame consistono nella determinazione di classificatori binari. In tale contesto bisogna trovare un criterio per distinguere gli elementi di due insiemi disgiunti di punti campione (o «pattern» o «training data»). Poiché i pattern sono generalmente rappresentati da punti nello spazio $\mathbb{R}^n$, il problema diventa quello di discriminare tra due insiemi finiti di punti, $\mathcal{A}$ e $\mathcal{B}$, nello spazio $n$-dimensionale $\mathbb{R}^n$, attraverso un iperpiano o una superficie non lineare di separazione.

Se le coperture convesse dei due insiemi di punti hanno intersezione vuota i due insiemi $\mathcal{A}$ e $\mathcal{B}$ sono *linearmente separabili* [3], ossia esiste un iperpiano di separazione tale che tutti i punti di un insieme si trovano da una parte dell'iperpiano e tutti i punti dell'altro insieme si trovano dall'altra parte dell'iperpiano stesso. La Programmazione Lineare può essere usata per costruire tale iperpiano [1].

Se le coperture convesse di $\mathcal{A}$ e $\mathcal{B}$ hanno intersezione non vuota è possibile ottenere un iperpiano $(v, \gamma)$ (con $v \in \mathbb{R}^n$ e $\gamma \in \mathbb{R}$) che minimizza una qualche misura della malclassificazione.

Un possibile criterio è quello di minimizzare la somma delle distanze dall'iperpiano di separazione dei punti malclassificati [5]. Precedenti approcci usano surrogati di distanze che mantengono la linearità del problema da risolvere.

Più in generale ci sono molti settori della ricerca scientifica, quali la statistica e l'informatica (attraverso le basi di dati, il «machine learning», il «pattern recognition», l'intelligenza artificiale e la visualizzazione dei dati), che si sono occupati di problemi di classificazione, proponendo per essi varie metodologie risolutive.

In particolare si ricordano le numerose ed interessanti applicazioni basate sull'uso delle reti neurali.

Un'altra metodologia di programmazione matematica per risolvere problemi di classificazione descritta nella tesi è la Support Vector Machine (SVM), dove la formulazione del problema di classificazione in termini di programmazione quadratica ha l'obiettivo di ottenere il miglior margine di separazione.

La costruzione di alberi di decisione rappresenta un'ulteriore tecnica di ottimizzazione applicata con successo al problemi di classificazione di tipo non lineare.

Il nuovo metodo di separazione poliedrale che viene descritto nella tesi è stato ispirato da una tecnica, messa a punto da Bennett e Mangasarian [1], che consiste nel minimizzare la media delle violazioni del seguente sistema (con $e$ viene indicato un vettore a componenti unitarie di dimensione appropriata):

$$\begin{cases} A^T v \leqslant e(\gamma - 1) \\ B^T v \geqslant e(\gamma + 1). \end{cases}$$

Il sistema (dove le righe delle matrici $A$ e $B$ rappresentano, rispettivamente, i punti di $\mathcal{A}$ e $\mathcal{B}$) ha soluzione, $v \in \mathbb{R}^n$ e $\gamma \in \mathbb{R}$, se e solo se i due insiemi $\mathcal{A}$ e $\mathcal{B}$ sono linearmente separabili.

La funzione obiettivo può essere scritta come somma pesata delle norme $L_1$ dei vettori degli errori di classificazione per entrambi gli insiemi $\mathcal{A}$ e $\mathcal{B}$, di cardinalità, rispettivamente, $m$ e $k$:

$$z^* = \min_{v,\,\gamma} 1/m \|\max\{0, A^T v - e\gamma + e\}\|_1 + 1/k \|\max\{0, -B^T v + e\gamma + e\}\|_1 .$$

Gli insiemi $\mathcal{A}$ e $\mathcal{B}$ sono linearmente separabili se e solo se $z^* = 0$ ed è stato provato che non si ha mai soluzione degenere $v = 0$.

La precedente formulazione è equivalente al seguente problema di programmazione matematica:

$$\begin{aligned} z^* = \quad & \min_{v,\,\gamma,\,y,\,z} \; e^T y/m + e^T z/k \\ \text{s.v.} \quad & y \geqslant A^T v - e\gamma + e \\ & z \geqslant -B^T v + e\gamma + e \\ & y \geqslant 0 \\ & z \geqslant 0 \end{aligned}$$

dove $y_i$ è non negativo e rappresenta l'errore per il punto $a_i \in \mathcal{A}$ e $z_l$ è non negativo e rappresenta l'errore per il punto $b_l \in \mathcal{B}$.

Quando i due insiemi $\mathcal{A}$ e $\mathcal{B}$ non possono essere separati da un iperpiano è naturale cercare superfici di separazione non lineari.

In [4] viene descritto un metodo per ottenere una separazione stretta attraverso uno o più iperpiani o superfici non lineari. La programmazione Lineare gioca un ruolo rilevante.

In [2] il problema di separare i due insiemi $\mathcal{A}$ e $\mathcal{B}$ attraverso due iperpiani viene affrontato come un problema di programmazione bilineare.

L'essenza del nuovo metodo descritto nella tesi è quella di ottenere una separazione con più di un iperpiano, usando una funzione di errore derivata dall'approccio descritto precedentemente [1].

Quando sussiste

$$\operatorname{conv}(A) \cap \operatorname{conv}(B) \neq \emptyset ,$$

i due insiemi $\mathcal{A}$ e $\mathcal{B}$ non possono essere separati da un iperpiano. Se però la copertura convessa di $\mathcal{A}$ e l'insieme $\mathcal{B}$ hanno intersezione vuota

$$\operatorname{conv}(\mathcal{A}) \cap B = \emptyset ,$$

i due insiemi sono *h-poliedralmente separabili* [6], ossia esistono $h$ iperpiani tali che i punti di $\mathcal{A}$ sono contenuti in un poliedro convesso (intersezione di $h$ semispazi) e i punti di $\mathcal{B}$ sono fuori da tale poliedro.

L'obiettivo è stato quello di mettere a punto un algoritmo in grado di trovare un insieme di $h$ iperpiani (con $h \geqslant 1$ fissato), $\{v^{(j)}, \gamma_j\}$ $(j = 1, \ldots, h)$, che separano poliedralmente l'insieme $\mathcal{A}$ dall'insieme $\mathcal{B}$, se tali iperpiani realmente esistono, oppure che minimizzano una misura di mal classificazione. A tale scopo è stata introdotta una funzione di errore affine a tratti, né concava né convessa, e si è messo a punto un algoritmo di discesa basato sulla risoluzione iterativa di sottoproblemi lineari. Il problema della separazione poliedrale, infatti, non può essere trasformato in un problema di Programmazione Lineare, come, invece, è possibile nel caso della separazione lineare.

La tesi si conclude con la descrizione dei risultati ottenuti dall'algoritmo di separazione poliedrale proposto, nell'applicazione ad una serie di database di dati reali (UCI Repository of Machine Learning Databases), la maggior parte di tipo medico, utilizzati in letteratura per testare nuove tecniche di classificazione e disponibili in rete.

Un'altra applicazione presentata nella tesi è quella relativa allo studio del decorso post operatorio dei pazienti sottoposti a trapianto renale. Tale decorso può evolvere verso differenti stati (buon trapianto, rigetto acuto, forme più o meno gravi di intossicazione). Il problema consiste nel prevedere tale decorso a partire da dati clinici rilevati nei periodi immediatamente successivi all'operazione. La ricerca è stata effettuata in collaborazione con il Reparto di Nefrologia dell'Ospedale Civile dell'Annunziata di Cosenza.

## BIBLIOGRAFIA

[1] Bennett K. P. e Mangasarian O. L., *Robust Linear Programming Discrimination of Two Linearly Inseparable Sets*, Optimization Methods and Software, **1** (1992), 23-34.

[2] Bennett K. P. e Mangasarian O. L., *Bilinear Separation of Two Sets in n-Space*, Computational Optimization and Applications, **2** (1993), 207-227.

[3] Mangasarian O. L., *Linear and nonlinear separation of patterns by linear programming*, Operations Research, **13** (1965), 444-452.

[4] Mangasarian O. L., *Multisurface Method of Pattern Separation*, IEEE Transactions on Information Theory, **IT-14 6** (1968), 801-807.
[5] Mangasarian O. L., *Arbitrary-Norm Separating Plane*, Operations Research Letters, **24** (1999), 15-23.
[6] Megiddo N., *On the Complexity of Polyhedral Separability*, Discrete Comput. Geom., **3** (1988), 325-337.

Istituto per la Sistemistica e l'Informatica, ISI-CNR, c/o DEIS
Università della Calabria, e-mail: astorino@deis.unical.it
Dottorato di Ricerca in Ricerca Operativa
(sede amministrativa: Università di Roma «La Sapienza») - Ciclo XII
Direttore di Ricerca: Prof. Manlio Gaudioso, Università della Calabria