STUDIA UNIV. BABEŞ-BOLYAI, MATHEMATICA, XXXIX, 3, 1994

A STUDY OF THE PROPERTIES OF THE FUZZY RELAXATION ALGORITHM

Horia F. POP*

Dedicated to Professor Sever Groze on his 65th anniversary

Received: February 1, 1995
AMS subject classification: 68P99

REZUMAT. - Un studiu asupra proprietăţilor Algoritmului de Relaxare Fuzzy. Una dintre problemele cele mai dificile ale instruirii supervizate este tratarea datelor neseparabile liniar. Problema a fost doar parţial rezolvată prin utilizarea algoritmilor de instruire nuanţată. Una din posibilităţile de abordare a problemei este generalizarea pentru cazul neseparabil a unor tehnici de instruire nuanţată care funcţionează bine în cazul separabil. In cele ce urmează vom studia prorietăţile Algoritmului de Relaxare Nuanţată [1,2]. Acesta permite o generalizare favorabilă datelor neseparabile liniar.

1. Introduction

In [1,2] it has been proposed a new training method that allows the use of fuzzy sets in order to develop the training. Based on this method a series of algorithms representing generalizations of some well known classical algorithms have been given.

There have also been proposed robust variants of the fuzzy training algorithm. These robust algorithms are capable of learning a training set consisting of two fuzzy linearly non-separable classes.

In this paper we study the fuzzy relaxation algorithm proposed in [2]. We approach here the convergence of the algorithm for the case when the constant b from this algorithm is a certain real number.

* *"Babeş-Bolyai" University, Faculty of Mathematics and Computer Science, 3400 Cluj-Napoca, Romania*

By this modification the fuzzy relaxation algorithm becomes capable of separating two fuzzy linearly non-separable classes.

2. The Fuzzy Relaxation Algorithm

Let $X = \{x^1,\ldots,x^p\}$, $x^j \in \mathbf{R}^s$ a data set and $\{A_1,A_2\}$ a fuzzy binary partition on X. We will consider the vectors y^j, obtained from x^j by adding a (s+1)-th component equal to 1.

We consider the sign normalization [2]

$$z = \begin{cases} y & \text{if } A_1(y) \geq 0.5, \\ -y & \text{if } A_2(y) > 0.5. \end{cases}$$

We will denote by Z the set of normalized vectors and will consider A_1 and A_2 as fuzzy sets on Z.

The Fuzzy Relaxation Algoritm [1,2] produces (in certain given conditions) a unitary separation vector v satisfying

$$v^T z \geq b > 0, \tag{1}$$

where b is a real positive number.

The correction rule used by the algorithm is (see [2])

$$v^{k+1} = \begin{cases} v^k + c(A_i(z^k))^2 \dfrac{b - v^{kT} z^k}{\|z^k\|^2} z^k & \text{if } A_i(z^k) > 0.5 \text{ and } v^{kT} z^k \leq b, \\ v^k & \text{otherwise}. \end{cases} \tag{2}$$

Instead of using the separation condition (1), as it appears in [2], we will use here a slightly different separation condition, namely:

$$v^T z \geq b,\ b \in \mathbf{R}. \tag{1'}$$

As we will see in the next section, the separation condition (1') allows the algorithm to work in the non-separable case.

3. Study of the properties of this algorithm

Firstly, let us introduce some important notations:

$$E(v,A) = \min \{ v^T z \mid z \in Z \}$$
$$E(A) = \sup \{ E(v, A) \mid \|v\| = 1, v \in \mathbb{R}^{s+1} \} \quad (3)$$
$$V^*(b) = \{ v \mid \|v\| = 1, v^T z > b \ \forall z \in Z \}$$

The following theorem shows a link between E(A) and the linear separability of the fuzzy sets A and $\tilde{A}$:

Theorem 1. Let X be a data set and A a fuzzy set on X. The following statements are equivalent:

(i) $E(A) > 0$;

(ii) A and $\tilde{A}$ are linearly separable fuzzy sets.

Proof. For the first part of the proof, let us see that $E(A)>0$ implies that there exists a certain v so that $E(v, A) > 0$. Let us denote the quantity $E(v, A)$ by b. From the definitions (3) we may deduce that $v^T z \geq b > 0$ for all $z \in Z$, and from here, the linear separability of A and $\tilde{A}$.

Conversely, if A and $\tilde{A}$ are linear separable, then there exists a vector v so that $v^T z > 0$ for all $z \in Z$. Let us denote

$$v' = v / \|v\|.$$

Thus, we have that $E(v', A) > 0$, and from here it occurs that $E(A) > 0$. □

The following theorem shows a condition for the existence of the set $V^*(b)$ of solution vectors:

Theorem 2. Let X be a data set, A a fuzzy set on X and b a real number. The following statements are equivalent:

(i) $b \geq E(A)$;

(ii) $V^*(b) = \emptyset$.

Proof. For the first part of the proof let us suppose that $b \geq E(A)$. It implies that for all the unitary vectors v, $E(v,A) \leq b$. Thus, for each unitary vector v, there exists at least a $z \in Z$ such that $v^T z \leq b$, and from here we deduce that $V^*(b) = \emptyset$.

Conversely, let us consider the unitary vector v as fixed. Thus, there exists at least a $z \in Z$ such that $v^T z \leq b$, and from here we have that $E(v, A) \leq b$. As the vector v was previously considered fixed, this propriety is valid for all the unitary vectors v. Thus, we conclude that $E(A) \leq b$. □

The following proposition gives a few properties of the se $\sqrt{}^*(b)$:

Proposition. Let X be a data set, A a fuzzy set on X and b a real number. The following statements are valid (Int V denotes the interior part of the set V, and Fr V denotes the border of the set V):

(i) $\text{Int } V^*(b) = \{v \mid \|v\| = 1, \forall z, v^T z > b\} = V^*(b)$;

(ii) $\text{Fr } V^*(b) = \{v \mid \|v\| = 1, \forall z, v^T z \geq b \text{ and } \exists z: v^T z = b\}$.

Proof. (i) Let us consider the family of sets

$$M_z = \{v \mid \|v\| = 1, v^T z > b\}, z \in Z.$$

Thus it is clear that

$$V^*(b) = \cap \{M_z \mid z \in Z\}.$$

But, M_z is an open set, and thus, $V^*(b)$, being a finite intersection of open sets, is an open set.

(ii) Let us denote

$$M = \{v \mid \|v\| = 1, \forall z, v^T z \geq b \text{ and } \exists z : v^T z = b\}.$$

Let us consider a v^* in M. So, we may split the set Z into Z_1 and Z_2 so that

$$(v^*)^T z > b, \forall z \in Z_1$$

and

$$(v^*)^T z = b, \forall z \in Z_2.$$

Let us chose an ε so that the sphere $S(v^*, \varepsilon)$ verifies the property

$$\forall\, v \in S(v^*, \varepsilon),\ \forall\, z \in Z_1,\ v^T z > b. \qquad (4)$$

Due to the first part of this proposition, such an · does certainly exist.

The hyperplanes $(v^*)^T z = b$ for $z \in Z_2$ split the sphere into a finite number of distinct regions. Important for us are only two of these regions, let us denote them R_1 and R_2, that verify

$$\forall\, v \in R_1,\ \forall\, z \in Z_2,\ v^T z > b \qquad (5)$$

and

$$\forall\, v \in R_2,\ \forall\, z \in Z_2,\ v^T z < b. \qquad (6)$$

From the relations (4), (5) and (6) it is clear that every v in R_1 is inside $V^*(b)$ and every v in R_2 is outside $V^*(b)$, and that proves that $M = \mathrm{Fr}\ V^*(b)$. This concludes the proof. □

From this proposition we may deduce the following

Corolary. Let X be a data set, A a fuzzy set on X and b a real number. The following equivalences are valid:

(i) $v \in \mathrm{Int}\ V^*(b) \Leftrightarrow E(v, A) > b$;

(ii) $v \in \mathrm{Fr}\ V^*(b) \Leftrightarrow E(v, A) = b$.

Proof. Is very easy as is based on the definition of $E(v, A)$ and the Proposition above. □

Let us denote by $R(b)$ the separation vector produced by the Fuzzy Relaxation Algorithm (as modified in this paper) under the condition $v^T z > b$, when this separation vector does exist.

The following theorem presents a convergence condition of the sequence (v^n) prodced by the Fuzzy Relaxation Algorithm. It represents a generalization of the convergence theorem given in [4]:

Theorem 3. Let X be a data set, A a fuzzy set on X and b and c two real numbers. Let (v^n) be the sequence of the vectors produced by the Fuzzy Relaxation Algorithm under the condition $v^T z > b$. If $0<c<2$ and $b<E(A)$, then the sequence (v^n) is convergent.

The proof of the convergence theorem as it has been stated in [3,4] is applicable even for the supplementary conditions imposed to b. Moreover, the proof of the theorem from [3,4] is based, even if not clearly specified, on a condition similar with $b < E(A)$.

Theorem 4. On the notation from the Theorem 3, if $b < E(A)$ and the sequence (v^n) is finite, then $E(R(b), A) > b$.

Proof. Since (v^n) is finite, there exists a certain $N \geq 1$ such that for all the i's with $i \geq N$, $v^i = v^N$. So, $v^N = v^{N+1} = \ldots = R(b)$, and that implies $(R(b))^T z^i > b$ for every i, and thus $R(b) \in V^*(b)$. Finally, we have that $E(R(b), A) > b$, and that concludes the proof. □

Theorem 5. On the notations from the Theorem 3, if $b < E(A)$, $0<c<2$ and the sequence (v^n) is infinite, then $E(R(b), A) = b$.

Proof. Let us remember that

$$R(b) = \lim_{n\to\infty} v^n.$$

We must show the following:

(i) for every i, $R(b)^T z^i \geq b$;

(ii) there exists at least an i such that $R(b)^T z^i = b$.

For the first part, let us consider the correction rule (2). Since R(b) is the limit of the sequence (v^n), it results that the correction rule do not modifies its value. So, we have the cases:

(a) There exists a certain i such that $R(b)^T z^i \geq b$. It results that

$$R(b) = R(b) + c((A_j(z^i))^2 \frac{b - R(b)^T z^i}{\|z^i\|^2} z^i,$$

where $A_j(z^i) > 0.5$, and, from here,

$$c\,(b - R(b)^T z^i) = 0$$

and, finally,

$$R(b)^T z^i = b.$$

(b) For the rest of the i's, we have that $R(b)^T z^i > b$, and the correction rule lets R(b) unmodified.

So, for all the i's $R(b)^T z^i \geq b$.

For the second part, if we had $R(b)^T z^i > b$ for all the i's, we would have that $R(b)^T z^i \in$ Int $V^*(b) = V^*(b)$. Since every vicinity of R(b) contains at least an element v^j of the sequence (v^n), it results that there exists a $v^j \in V^*(b)$. Thus, v^j is a stop point and (v^n) is a finite sequence, and that contradicts the hypothesis.

Finally, we have that $R(b) \in$ Fr $V^*(b)$ and that $E(R(b), A) = b$. That concludes the proof. □

3. Concluding remarks

It is certainly interesting to study what happens in the case $b \geq E(A)$. Even if we haven't proved yet, the experience enables us to consider the following

Conjecture. On the notations from the Theorem 3, if $b \geq E(A)$ the sequence (v^n) is convergent and $E(R(b), A) = E(A)$.

Let us notice that we did not introduce any restriction with respect to $b > 0$. Consequently, the theorems presented above are valid for the case $b < 0$, with the single condition that $b < E(A)$. So, we may assure the output of a 'separation' hyperplane for the case of linear non-separability, i.e. when $b < E(A) < 0$. This is a remarkable property of the Fuzzy

Relaxation Algorithm.

Other interesting problem is whether there exists a modality to compute directly E(A) and whether there exists a certain v such that E(v,A) = E(A). Thus, the Fuzzy Relaxation Algorithm would be able to produce the optimal separation hyperplane with respect to E(A).

REFERENCES

1. D. Dumitrescu. A fuzzy training algorithm. Studia Universitatis "Babeş-Bolyai", Ser. Math. 35, (1990), 7-12.
2. D. Dumitrescu. Fuzzy training procedures, I. Fuzzy Sets and Systems 56 (1993), 155--169.
3. D. Dumitrescu. Principiile matematice ale teoriei clasificării, Editura Academiei Române, Bucureşti, 1995.
4. D. Dumitrescu. Fuzzy sets and their applications for clustering and training, to appear.
5. D. Dumitrescu, Horia F. Pop. Convex decomposition of fuzzy partitions, I,II. Fuzzy Sets and Systems (1995), to appear.
6. J. Sklansky, and G. N. Wassel. Pattern Classifiers and Trainable Machines. Springer Verlag, New York, 1981.