ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 219-224

The Class of Linear Separability Method

Mohamed Tajine[1], David Elizondo[1,2], Emile Fiesler[2], Jerzy Korczak [1]

[1]LSIIT (CNRS URA 1871) - Université Louis Pasteur, Département d'Informatique
7, rue René Descartes, 67084 Strasbourg Cedex, France

[2]IDIAP, Case Postale 592, CH-1920 Martigny, Switzerland

Abstract. Given a two linearly separable finite sets of points X, Y, we characterize the set of points by which it passes a hyperplane that linearly separates X and Y. Based on this characterization, we propose a novel algorithm for testing linear separability.

1. Introduction

Knowing if two classes in a two class classification problem are linearly separable (LS) can help simplify the topology of a multilayered perceptron neural network. In this case, a single layer perceptron network can be used to classify them. Also, identifying LS sub-sets from a non-LS problem can be used to recursively build a network topology as in the case of the Recursive Deterministic Perceptron (RDP) described in [4].

We characterize the set of points by which it passes a hyperplane that linearly separates a given two linearly separable sets X and Y and we propose a novel algorithm for testing linear separability based on this characterization. This new method is guided because it uses the points of $X \cup Y$ as supporting points to search a hyperplane which linearly separates X, Y.

This paper is divided in four sections. In the second section some standard notations and definitions are given, and some general properties related to these notations and definitions. In the third section several methods for testing linear separability are introduced and a new original method is proposed which is based on computational geometry techniques.

2. Preliminaries

The following standard notions are used: Let $E, F \subset I\!R^d$, $Card(E)$ stands for the cardinality of a set E. $E \setminus F$ is the set of elements which belongs to E and does not belong to F. Let $\vec{p}_1, \vec{p}_2$ be two points in $I\!R^d$, the set $\{t\vec{p}_1 + (1-t)\vec{p}_2 \mid 0 \leq t \leq 1\}$ is called the segment between $\vec{p}_1, \vec{p}_2$ and is denoted by

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 219-224

$[\vec{p}_1, \vec{p}_2]$; if $\vec{u} = (u_1, ..., u_d), \vec{v} = (v_1, ..., v_d)$, then $\vec{u}^T\vec{v}$ stands for $u_1v_1 + ... + u_dv_d$. Two sub-sets X and Y of $\mathbb{R}^d$ are said to be linearly separable if there exists a hyperplane P of $\mathbb{R}^d$ such that the elements of X and those of Y lie on opposite sides of it, this is denoted by $X \parallel Y$ or $X \parallel Y\ (P)$, thus if $X \parallel Y\ \mathcal{P}(\vec{w}, t)$, then $(\forall \vec{x} \in X,\ \vec{w}^T\vec{x} + t > 0$ and $\forall \vec{y} \in Y,\ \vec{w}^T\vec{y} + t < 0)$ or $(\forall \vec{x} \in X,\ \vec{w}^T\vec{x} + t < 0$ and $\forall \vec{y} \in Y,\ \vec{w}^T\vec{y} + t > 0)$. Let $X, Y \subset \mathbb{R}^d$, $\mathbb{P}(X,Y) = \{P \mid X \parallel Y\ (P)\}$. Let $P \in \mathbb{P}$, $\mathcal{C}_Y(X,P)$ is the half space delimited by P and containing X *(i.e.* $\mathcal{C}_Y(X,P) = \{\vec{v} \in \mathbb{R}^d \mid \vec{u}^T\vec{v} + t > 0\}$ *if for some* $\vec{x} \in X$, $\vec{u}^T\vec{x} + t > 0$*)*

3. Methods for Testing Linear Separability

The different methods for testing the linear separability between two sets of points X and Y, can be classified in three groups : the methods based on the Fourier-Kuhn elimination algorithm [2] or on linear programming [1], the methods based on the perceptron algorithm [3], and the methods based on computational geometry techniques.

We now present a new algorithm for testing linear separability based on the notion of class of linear separability.

3.1. The class of linear separability method

In this sub-section, the set of points P of $\mathbb{R}^d$ by which it passes a hyperplane that linearly separates X and Y are characterized.

Property 3..1 *Let* $X, Y \subset \mathbb{R}^d$ *be two linearly separable sets and let* $\mathcal{C}(X,Y) = \mathbb{R}^d \setminus (\cup_{P \in \mathbb{P}(X,Y)} P)$*, Then,* $\mathcal{C}(X,Y) = \mathcal{C}_Y(X) \cup \mathcal{C}_X(Y)$ *with* $\mathcal{C}_Y(X) \cap \mathcal{C}_X(Y) = \emptyset$ *and,* $\mathcal{C}_X(Y)$ *and* $\mathcal{C}_Y(X)$*, are sets which are closed, unbounded, and convex.* $\mathcal{C}_Y(X)$ *is called a linear separability class of* X *relatively to* Y*. (i.e, if* $\mathcal{C}(X,Y)$ *is not empty, then for all the points* P *not in* $\mathcal{C}(X,Y)$*, there exists a hyperplane , containing* P*, that linearly separates* X *and* Y*.)*

Proof :
Let $\mathcal{P}(\vec{u}, t) \in \mathbb{P}(X,Y)$, and note $\mathcal{C}_Y(X) = \cap_{P \in \mathbb{P}(X,Y)} \mathcal{C}_X(Y,P)$, $\mathcal{C}_X(Y) = \cap_{P \in \mathbb{P}(X,Y)} \mathcal{C}_Y(X,P)$ then $\mathcal{C}_Y(X)$ and $\mathcal{C}_X(Y)$ are convex because $\forall P \in \mathbb{P}(X,Y)$, $\mathcal{C}_X(Y,P)$ and $\mathcal{C}_Y(X,P)$ are convex.
Moreover, if $\vec{x} \in \mathbb{R}^d \setminus \mathcal{C}(X,Y)$ then $\exists P \in \mathbb{P}(X,Y)$ such that $\vec{x} \in P$. Let $\varepsilon = Min(d(X,P), d(Y,P)$ *(* $\varepsilon > 0$ *because* $X \cap P = Y \cap P = \emptyset$*)*, hence $B(\vec{x}, \varepsilon) = \{\vec{y} \in \mathbb{R}^d \mid \|\vec{x} - \vec{y}\| < \varepsilon\} \subset (\mathbb{R}^d \setminus \mathcal{C}(X,Y))$, thus $\mathcal{C}(X,Y)$ is a closed set in $\mathbb{R}^d$.
Let $\vec{x} \in X$ and $\vec{y} \in Y$ and consider the straight line $D = \{s\vec{x} + (1-s)\vec{y} \mid s \in \mathbb{R}\}$ defined by $\vec{x}, \vec{y}$.
Assume that $\mathcal{C}_Y(X)$ is bounded, then there exists $s_0 < 0$ such that for all $s < s_0, s\vec{x} + (1-s)\vec{y} \notin \mathcal{C}_Y(X)$. Let $s_1 < s_0$, then $s_1\vec{x} + (1-s_1)\vec{y} \in \mathcal{C}(X,Y)$ *(because* $X \parallel Y$ *)*, then there exists $\vec{w}, t$ such that $s_1\vec{x} + (1-s_1)\vec{y} \in \mathcal{P}(\vec{w}, t)$ and $X \parallel Y\ (\mathcal{P}(\vec{w}, t))$, thus, $(\vec{w}^T\vec{x} + t)(\vec{w}^T\vec{y} + t) < 0$, then there exists $0 < s_2 < 1$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 219-224

such that $s_2\vec{x} + (1 - s_2)\vec{y} \in \mathcal{P}(\vec{w}, t)$ hence, $\vec{x}, \vec{y} \in \mathcal{P}(\vec{w}, t))$ which is absurd, thus $\mathcal{C}_Y(X)$ □

Definition 3..1 *Let S be a subset of points in $I\!R^d$ and let $\vec{a} \in S$; then $dimaf(S)$ (dimension affine) is the dimension of the vectorial sub-space generated by $\{\vec{x} - \vec{a} \mid \vec{x} \in S\}$. In other words, $dimaf(S)$ is the dimension of the smallest affine sub-space that contains S.*

Notations :

• Let $Z \subset I\!R^d$ be a set of d affinely independent points, $\mathcal{P}_Z$ is the unique hyperplane of $I\!R^d$ containing Z.

• Let $\mathcal{P}$ a hyperplane of $I\!R^d$, and $X \subset I\!R^d$, $H_{\mathcal{P}}(X)$ is the half of $I\!R^d$ delimited by $\mathcal{P}$ and containing X.

• $h(X,Y) = \{Z \subset X \cup Y \mid Card(Z) = d,\ Z$ is a set of affinely independent points and $(X \setminus \mathcal{P}_Z) \mid\mid (Y \setminus \mathcal{P}_Z)\ (\mathcal{P}_Z) and (X \cap \mathcal{P}_Z) \mid\mid (Y \cap \mathcal{P}_Z)\}$.

Theorem 3..1 *Let $X, Y \subset I\!R^d$, such that $dimaf(X \cup Y) = d$, then $X \mid\mid Y$ iff there exists $\mathcal{P}(\vec{w}, t)$ such that $dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}, t)) = d - 1$
$[(X \setminus \mathcal{P}(\vec{w}, t)) \mid\mid (Y \setminus \mathcal{P}(\vec{w}, t))\ (\mathcal{P}(\vec{w}, t))$ and $(X \cap \mathcal{P}(\vec{w}, t)) \mid\mid (Y \cap \mathcal{P}(\vec{w}, t)))]$ or
$[(X \cup Y) \subset \mathcal{P}(\vec{w}, t))$ and $X \mid\mid Y]$.*

Proof :

1) Assume that there exists $\mathcal{P}(\vec{w}, t)$ such that $dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}, t)) = d - 1$; $(X \setminus \mathcal{P}(\vec{w}, t)) \mid\mid (Y \setminus \mathcal{P}(\vec{w}, t))\ (\mathcal{P}(\vec{w}, t))$ and $(X \cap \mathcal{P}(\vec{w}, t)) \mid\mid (Y \cap \mathcal{P}(\vec{w}, t))(\mathcal{P}(\vec{w}', t'))$ and assume that $\forall \vec{x} \in (X \setminus \mathcal{P}(\vec{w}, t))$, $\vec{w}^T\vec{x} + t > 0$ and $\forall \vec{z} \in (X \cap \mathcal{P}(\vec{w}, t))$, $\vec{w}'^T\vec{z} + t' > 0$.
Let $\alpha = Max(\{|\vec{w}'^T\vec{z} + t'| \ ;\ \vec{z} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}, t))\})$ and $\varepsilon = Min(\{|\vec{w}^T\vec{z} + t| \ ;\ \vec{z} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}, t))\})$. Let $0 < \delta \leq \frac{\varepsilon}{2\alpha}$, $\vec{w}^\delta = \vec{w} + \delta\vec{w}'$ and $t^\delta = t + \delta t'$, then by construction $X \mid\mid Y\ (\mathcal{P}(\vec{w}^\delta, t^\delta)$.
2) Assume that $X \mid\mid Y\ (\mathcal{P}(\vec{w}, t)$ and assume by induction that there exists $\mathcal{P}(\vec{w}_i, t_i)$ such that $(X \setminus \mathcal{P}(\vec{w}_i, t_i)) \mid\mid (Y \setminus \mathcal{P}(\vec{w}_i, t_i))\ (\mathcal{P}(\vec{w}_i, t_i))$, $(X \cup Y) \setminus \mathcal{P}(\vec{w}_i, t_i) \neq \emptyset$ and $i \leq k = dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}_i, t_i)) < d - 1$. Let $\vec{z} \in (X \cup Y) \cap \mathcal{P}(\vec{w}_i, t_i)$ and $\{\vec{v}_1, ..., \vec{v}_{d-1}\}$ an orthogonal family *(i.e. $\{\vec{v}_1, ..., \vec{v}_{d-1}\} \cup \{\vec{w}_i\}$ is an orthogonal basis of $I\!R^d$)* such that
$\mathcal{P}(\vec{w}_i, t_i) = \{\vec{z} + \lambda_1\vec{v}_1 + ... + \lambda_{d-1}\vec{v}_{d-1} \ ;\ \lambda_1, ..., \lambda_{d-1} \in I\!R\}$ and $(X \cup Y) \cap \mathcal{P}(\vec{w}_i, t_i) \subset \{\vec{z} + \lambda_1\vec{v}_1 + ... + \lambda_k\vec{v}_k \ ;\ \lambda_1, ..., \lambda_k \in I\!R\}(k < d - 1)$. Two vectors $\vec{X} = \{x_1, ..., x_d\}$ and $\vec{Y}\{y_1, ..., y_d\}$ are orthogonal if $\vec{X}^T\vec{Y} = 0$
Assume that $(X \cup Y) \setminus \mathcal{P}(\vec{w}_i, t_i) = \{\vec{z}_1, ..., \vec{z}_m\}$ and let $f_1, ..., f_m$ be the continuous functions such that $f_j(\theta) = (cos(\theta)\vec{w}_i + sin(\theta)\vec{v}_{d-1})^T(\vec{z}_j - \vec{z})$ for $1 \leq j \leq m$. Then $\forall j \leq m$, $\exists \theta_j \in]-\frac{\pi}{2}, \frac{\pi}{2}[\ f_j(\theta_j) = 0$, actually if $\vec{z}_j - \vec{z} = \lambda_1\vec{v}_1 + ... + \lambda_{d-1}\vec{v}_{d-1} + \lambda_d\vec{w}_i$, then $\theta_j = arctan(-\frac{\lambda_d|\vec{w}_i|^2}{\lambda_{d-1}|\vec{v}_{d-1}|^2})$. Let $\theta_n = Min(\{\theta_j; 1 \leq j \leq m\})$, $\vec{w}_{i+1} = cos(\theta_n)\vec{w}_i + sin(\theta_n)\vec{v}_{d-1}$ and $t_{i+1} = -w_{i+1}^T\vec{z}$, then $(X \setminus \mathcal{P}(\vec{w}_{i+1}, t_{i+1})) \mid\mid (Y \setminus \mathcal{P}(\vec{w}_{i+1}, t_{i+1}))\ (\mathcal{P}(\vec{w}_{i+1}, t_{i+1}))$ and $dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}_i, t_i)) < dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}_{i+1}, t_{i+1}))$ because $(X \cup Y) \cap \mathcal{P}(\vec{w}_i, t_i) \subset (X \cup Y) \cap \mathcal{P}(\vec{w}_{i+1}, t_{i+1})$, $z_n \in \mathcal{P}(\vec{w}_{i+1}, t_{i+1})$ and $z_n \notin \mathcal{P}(\vec{w}_i, t_i)$.
Hence, $\exists \vec{w} \in I\!R^d, \exists t \in I\!R$, $((X \cup Y) \subset \mathcal{P}(\vec{w}, t))$ or $(dimaf((X \cup Y) \cap \mathcal{P}(\vec{w}, t)) =$

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 219-224

$ClassSepar(X, Y, d, \vec{w}, t)$
– data: two data set vectors, X and Y representing
two classes
– result: a weight vector, $\vec{w}$ and a threshold t such that which $\mathcal{P}(\vec{w}, t)$
linearly separate the two classes if $X \parallel Y$
Begin
  If $h(X, Y) = \emptyset$ then $not(X \parallel Y)$
  else
  begin
    Select $Z \in h(X, Y)$;
    $ClassSepar(X \cap P_z, Y \cap P_z, d-1, \vec{w}', t')$;
    $\varepsilon := Min(\{|\vec{w}_Z^T\vec{a} + t_Z| \ \vec{a} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}_Z, t_Z))\})$;
    $(\ \mathcal{P}_Z = \mathcal{P}(\vec{w}_Z, t_Z)\ )$;
    $\alpha := Max(\{|\vec{w}'^T\vec{a} + t'| \ \vec{a} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}_Z, t_Z))\})$;
    $\vec{w} := \vec{w}_Z + \frac{\varepsilon}{2\alpha}\vec{w}'$;
    $t := t_Z + \frac{\varepsilon}{2\alpha}t'$;
  end
End

Table 1: The class of linear separability method

$d-1$ and $(X \setminus \mathcal{P}(\vec{w}, t)) \parallel (Y \setminus \mathcal{P}(\vec{w}, t))\ (\mathcal{P}(\vec{w}, t))$ □

Class of linear separability method :
Given $X, Y \subset \mathbb{R}^d$, the class of linear separability procedure, presented in the algorithm 1, computes recursively $\vec{w} \in \mathbb{R}^d$ and $t \in \mathbb{R}$ *(if there exists)* such that $X \parallel Y\ (\mathcal{P}(\vec{w}, t))$

Theorem 3..2 *Let* $X, Y \subset \mathbb{R}^d$, *such that* $dimaf(X \cup Y) = d$ *then,* $\mathcal{C}_Y(X) = \cap_{Z \in h(X,Y)} H_{\mathcal{P}_Z}(X)$.

Proof :
Let $Z \in h(X, Y)$ and $\vec{w}, \vec{w}', t, t'$ such that $\mathcal{P}_Z = \mathcal{P}(\vec{w}, t)$ and $(X \cap \mathcal{P}_Z) \parallel (Y \cap \mathcal{P}_Z)\mathcal{P}(\vec{w}', t')$. Set $\alpha = Max(\{|\vec{w}'^T\vec{c} + t'|\ ;\ \vec{c} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}, t))\})$ and $\varepsilon = Min(\{|\vec{w}^T\vec{c} + t|\ ;\ \vec{c} \in ((X \cup Y) \setminus \mathcal{P}(\vec{w}, t))\})$. Let $\delta \in]0, \frac{\varepsilon}{2\alpha}]$ and $\vec{w}^\delta = \vec{w} + \delta\vec{w}'$ and $t^\delta = t + \delta t'$, then by construction $X \parallel Y\ (\mathcal{P}(\vec{w}^\delta, t^\delta))$, thus $\mathcal{C}_Y(X) \subset \cap_{\delta \in]0, \frac{\varepsilon}{2\alpha}]} H_{\mathcal{P}(\vec{w}^\delta, t^\delta)}(X)$ hence, $\mathcal{C}_Y(X) \subset H_{\mathcal{P}_Z}(X)$ because $\mathcal{C}_Y(X)$ is closed set.
Assume that $\cap_{Z \in h(X,Y)} H_{\mathcal{P}_Z}(X) \setminus \mathcal{C}_Y(X) \neq \emptyset$, and let $\vec{c} \in (\cap_{Z \in h(X,Y)} H_{\mathcal{P}_Z}(X) \setminus \mathcal{C}_Y(X))$, then there exists $\vec{w} \in \mathbb{R}^d$ and $t \in \mathbb{R}$ such that $X \parallel Y\ (\mathcal{P}(\vec{w}, t))$ and $\vec{c} \in \mathcal{P}(\vec{w}, t)$.
By using the construction proposed in the theorem 3.2 there exists a hyperplane $\mathcal{P}(\vec{w}', t') = \{\vec{z} + \lambda_1\vec{v}_1 + ... + \lambda_{d-1}\vec{v}_{d-1}\ ;\ \lambda_1, ..., \lambda_{d-1} \in \mathbb{R}\}$ with $\vec{v}_{d-1} = \vec{c} - \vec{z}$, $\{\vec{z} + \lambda_1\vec{v}_1 + ... + \lambda_{d-2}\vec{v}_{d-2}\ ;\ \lambda_1, ..., \lambda_{d-2} \in \mathbb{R}\}$ contain a subset I of $X \cup Y$ of $d-1$ affinely independent points, and $\forall \vec{x} \in X\ \vec{w}'^T\vec{x} + t' \geq 0$, then by the theorem 3.2 there exists $\vec{z}" \in (X \cup Y) \setminus \mathcal{P}(\vec{w}', t')$ and $\theta_{n"}$ such that $\vec{w}"^T(\vec{z}" - \vec{z}) = 0$ and $\vec{w}"'^T(\vec{z}"' - \vec{z}) = 0$ where $\vec{w}" = cos(\theta_{n"})\vec{w}' - sin(\theta_{n"})\vec{v}_{d-1}$.

Let $t" = -\vec{w}"^T\vec{z}$, then $I \cup \{\vec{z}"\}$ is a set of d affinely independent points contained in $\mathcal{P}(\vec{w}", t")$, thus $\mathcal{P}(\vec{w}", t") \in h(X, Y)$, $(X \backslash \mathcal{P}(\vec{w}", t") \parallel (Y \backslash \mathcal{P}(\vec{w}", t")\ (\mathcal{P}(\vec{w}", t"))$ and $\forall \vec{x} \in X\ \vec{w}"^T\vec{x} + t" \geq 0$, but $\vec{w}"^T \vec{v} + t" < 0$, then $\vec{v} \notin H_{\mathcal{P}(\vec{w}", t")}(X)$ which absurd. So, $\mathcal{C}_Y(X) = \cap_{Z \in h(X,Y)} H_{\mathcal{P}_Z}(X)$ □

Remark :
Let $X, Y \subset I\!R^d$ such that $dimaf(X \cup Y) = l < d$, then the separability problem for X, Y in $I\!R^d$ can be transformed into a separating problem in $I\!R^l$.

3.2. Example

To illustrate how the Class of Linear Separability method works, we will apply it to a 2-class 2 dimensional classification problem. Let $X = \{(4,5), (8,6), (2,7), (6,7), (8,7), (5,8), (7,8), (6,9), (1,8)\}$ and $Y = \{(1,1), (3,1), (5,1), (7,2), (1,3), (6,3), (5,4)\}$ represent the input patterns for the two classes which define our problem. We want to find out if $X \parallel Y$. Figure 1-a shows a plot of the 2 classes. Following the algorithm, we want to identify all the hyperplanes that (a) pass by one point of each of the two classes and (b) linearly separate the remaining points.

In this example there is only one such hyperplane as illustrated in figure 1-b. This hyperplane is represented by $\mathcal{P}((1,1), -9)$. We now recursively reduce the original dimension of the problem to 1 dimension. Once this is done, we calculate the middle point between the two original points belonging to each of the classes. Next we calculate a hyperplane in the original dimension which passes by this middle point and is different than the first hyperplane. This is illustrated in figure 1-c. The selected two points are highlighted with a circle (⊙ and ⊗). They correspond to the points (4,5), and (5,4) respectively. Thus the middle point has a value equal to (4.5,4.5). This second hyperplane is represented by $\mathcal{P}((1,-1), 0)$. With these two hyperplanes we now compute the values of ε and α in the following way:

$$\begin{array}{lllll} \varepsilon & := & Min(\{|\vec{w}_Z^T\vec{a} + t_Z| \ ; \ \vec{a} \in (A \cup B)\}) & := & 6 \\ \alpha & := & Max(\{|\vec{w}'^T\vec{a} + t'| \ ; \ \vec{a} \in (A \cup B)\}) & := & 4 \end{array}$$

were A and B correspond to the set of points by which the first hyperplane does not pass as illustrated in figure 1-d. We use these values for ε and α to calculate the final hyperplane which linearly separates the two classes in the following way:

$$\begin{array}{lllll} \vec{w} & := & \vec{w}_Z + \frac{\varepsilon}{2\alpha}\vec{w}' & := & (-0.25, -1.75) \\ t & := & t_Z + \frac{\varepsilon}{2\alpha}t' & := & (9) \end{array}$$

This is illustrated in figure 1-e. The final hyperplane is represented by a dotted line and corresponds to $\mathcal{P}((-0.25, -1.75), 9)$. This final hyperplane is illustrated in figure 1-f.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 219-224

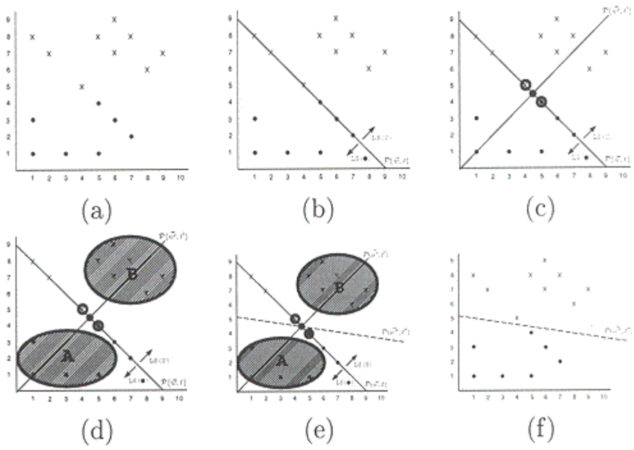

Figure 1: Steps followed by the Class of Linear Separability for finding the hyperplane which linearly separates classes x and $\bullet$

4. Discussion and Concluding Remarks

A new algorithm for testing linear separability has been presented. This new algorithm transforms recursively the problem of linear separability from d dimensions to $d-1$. This transformation can require $O(n^d)$ operations in the worse case. Thus, in order to simplify the research of this hyperplane, any hyperplane can be used to begin with, which can be move in function of the points until it contains a d affinely independent set of points of the two classes, and linearly separates the rest of the points.

References

[1] M. S. Bazaraa and J. J. Jarvis. *Linear Programming and Network Flow.* John Wiley and Sons, London, 1977.

[2] H. W. Kuhn. Solvability and consistency for linear equations and inequalities. *American Mathematical Monthly*, 63:217–232, 1956.

[3] F. Rosenblatt. *Principles of Neurodynamics.* Spartan, Washington D.C., 1962.

[4] M. Tajine and D. Elizondo. Enhancing the perceptron neural network by using functional composition. Technical Report 96-07, Computer Science Department, Université Louis Pasteur, Strasbourg, France, 1996.