Professor Gheorghe RUXANDA, PhD
Department of Economic Cybernetics
Academy of Economic Studies, Bucharest

SUPERVISED PATTERN RECOGNITION WITH POTENTIAL FUNCTIONS METHOD

Abstract. *In this paper, the potential functions method is analyzed, as (nonlinear) supervised pattern recognition method. In this context, the theoretical fundamentals of the potential functions are analyzed, the constructing modality is described and the physical nature of these functions is explained. The classification criterion based on the potential functions method is defined and the practical modality of using the potential functions, for supervised pattern recognition, is illustrated.*

Keywords: *supervised pattern recognition, discriminant analysis, artificial intelligence, potential functions.*

JEL classification: C45, C65, C87

1. INTRODUCTION

The solving of supervised pattern recognition problems by using Bayesian analysis based methods, assumes that a priori probabilities and the conditioned probability densities of classes are known. As, generally, the conditioned probability densities of classes depend upon certain parameters, the methods of supervised pattern recognition of this category are also called *parametric methods*. In this context, a priori probabilities do not cause any problem, as in case they are not known, they may be estimated with the aid of relative frequencies of classes at sample level.

In order the parametric methods to be possible to be used in supervised pattern recognition, it is required that besides knowing the pattern of conditioned probability densities of classes, to be possible to estimate also the parameters of these densities. But, in many cases, the conditioned probability densities of classes are not always known so that the practical use of such methods becomes impossible. When the conditioned probability densities of classes are not known, there is the possibility as their values to be approximated by means of a special method, known in statistical estimation theory as *Kernel functions or potential functions*. This method can be used both as a method of approximating the probability densities of classes with a view to using Bayesian classification techniques and as a proper classification method.

Supervised pattern recognition can be done with the aid of certain non-parametric methods, where the *potential functions method* is one of the most powerful and more frequently met methods of non-parametric nature.

The potential functions method is a non-parametric, simple and very powerful classification method, where the classification of a certain pattern is done according to the value of a certain *potential* , with whose aid, the density probabilities of classes can be approximated, starting from the information included in the training set. In the context of this method, the probability densities are assimilated to what represents in physics, the magnitude of electric field potential in an arbitrary point of its manifestation.

The basic idea of the potential functions method consists in assuming that each pattern of the space of patterns is considered as being associated to a certain potential of another pattern or of another class, similarly to the electric potential of physics, potential generating in the respective pattern neighborhood, a certain field, similar to the electric field. In other words, we assume that in each point of the space of patterns a certain higher or lower potential of a pattern or of a class, in the respective point neighborhood occurs.

The classification of a pattern by using the potential functions method assumes the evaluation of the global or total potential of a class in the point represented by the respective pattern, and the pattern will be assigned to the class exerting the highest force upon the pattern, namely, to the class with the highest potential in the respective point.

The solving of supervised pattern recognition problems by using Bayesian analysis based methods assumes that a priori probabilities and the conditioned probability densities of classes are known. As, generally, the conditioned probability densities of classes depend upon certain parameters, the methods of supervised pattern recognition of this category are also called *parametric methods*. In this context, a priori probabilities do not cause any problem, as in case they are not known, they may be estimated with the aid of relative frequencies of classes at sample level.

In order the parametric methods to be possible to be used in supervised pattern recognition, it is required that besides knowing the pattern of conditioned probability densities of classes, to be possible to estimate also the parameters of these densities. But, in many cases, the conditioned probability densities of classes are not always known so that the practical use of such methods becomes impossible. When the conditioned probability densities of classes are not known, there is the possibility as their values to be approximated by means of a special method, known in statistical estimation theory as *Kernel functions or potential functions*. This method can be used both as a method of approximating the probability densities of classes with a view to using Bayesian classification techniques and as a proper classification method.

Supervised pattern recognition can be done with the aid of certain non-parametric methods, where the *potential functions method* is one of the most powerful and more frequently met methods of non-parametric nature.

The potential functions method is a non-parametric, simple and very powerful classification method, where the classification of a certain pattern is done according to the value of a certain *potential* , with whose aid, the density probabilities of classes can be approximated, starting from the information included in the training set. In the context of this method, the probability densities are assimilated to what represents in physics, the magnitude of electric field potential in an arbitrary point of its manifestation.

The basic idea of the potential functions method consists in assuming that each pattern of the space of patterns is considered as being associated to a certain potential of another pattern or of another class, similarly to the electric potential of physics, potential generating in the respective pattern neighborhood, a certain field, similar to the electric field. In other words, we assume that in each point of the space of patterns a certain higher or lower potential of a pattern or of a class, in the respective point neighborhood occurs.

The classification of a pattern by using the potential functions method assumes the evaluation of the global or total potential of a class in the point represented by the respective pattern, and the pattern will be assigned to the class exerting the highest force upon the pattern, namely, to the class with the highest potential in the respective point.

According to the *superposition* principle, the global potential of a class in an arbitrary point of the space of patterns is represented by the *individual potentials sum* recorded by each pattern of the respective class in that point. In this way, if function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ defines the *individual potential* of pattern $\mathbf{x}^{(i)}$ of class ω_k in point $\mathbf{x}$ and the number of patterns of class ω_k is equal to T_k then, the *global potential* of class ω_k in point $\mathbf{x}$ is given by function:

$$\Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_k} \varphi(\mathbf{x}, \mathbf{x}^{(i)}), \qquad \mathbf{x}^{(i)} \in \omega_k, \quad k=1,2,...,K.$$

Function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ is called *individual potential function* of pattern $\mathbf{x}^{(i)}$, and function $\Psi_{\omega_k}(\mathbf{x})$ is called *global potential function* of class ω_k.

As a rule, individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ may have a different pattern from a pattern $\mathbf{x}^{(i)}$ to another one, and even more, from a class to another one. In order to outline the pattern differences of the individual potential function, we shall use an indexing $\varphi_i(\mathbf{x}, \mathbf{x}^{(i)})$ in order to specify that the function differs from a pattern to another one, respectively, an indexing of type $\varphi_k(\mathbf{x}, \mathbf{x}^{(i)})$ in order to specify that function differs from a class to another one.

As for constructing the global potential function, is it necessary to know the pattern of the individual potential functions, the single difficult problem in this context is to select the pattern for the individual potential functions. Thus, there are

more possible types of individual potential functions, which can be used to the discriminant analysis.

The definition of the potential functions is based on the principle according to which, the force exerted by a certain class upon a certain pattern, is *inversely proportional* to the square of distance where the pattern of the respective class is located.

The potential functions $\Psi(\mathbf{x})$ are functions similar to that of the conditioned probability densities of classes and can be used to define the *classification functions* of classes. If $\Psi_{\omega_k}(\mathbf{x})$ is the potential function of class ω_k then, the classification function corresponding to this class can be defined as follows:

$$D_k(\mathbf{x}) = \Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_k} \varphi_k\left(\mathbf{x}, \mathbf{x}^{(i)}\right), \quad \mathbf{x}^{(i)} \in \omega_k .$$

By assuming that the classification functions are defined for each class, we can determine also the *discriminant functions* for each possible pair of classes, as a difference between the two classification functions of classes of each pair.

2. POTENTIAL FUNCTIONS PHYSICAL RENDITION

From physics point of view, the *interaction* of the forces between two electric charges is described by Coulomb law, according to which, two electric idle charges attract or repel one another with a force *proportional to the product between the magnitude of the two charges and inversely proportional to the square of distance* between them.

The interaction of the two electric charges is determined by a certain field of forces around each electric charge, known under the name of *electric field.* If the scalar values q_1 and q_2 quantify the magnitude and direction of two interacting electric charges, then, according to Coulomb law, the force exerted within the interaction upon the second charge is described by the following vector relation:

$$\mathbf{F}_2 = k_0 \frac{q_2 q_1}{d_{21}^2} \mathbf{u}_{21},$$

where k_0 is a constant whose magnitude depends on the electric permeability of the environment, d_{21}^2 is the square of distance between the two charges and $\mathbf{u}_{21}$ unitarily oriented vector, from the first charge to the second charge, vector having as support the line binding the two charges. Similarly, the force exerted within the interaction upon the first charge is described by the vector relation:

$$\mathbf{F}_1 = k_0 \frac{q_1 q_2}{d_{12}^2} \mathbf{u}_{12}.$$

The previous relation points out the fact that the two charges repel one another if they have the same sign and attract one another if they have opposite sign. Even more, the two relations demonstrate that the forces are Newton type forces, namely:

$$\mathbf{F}_2 = -\mathbf{F}_1 .$$

The mode of calculating the interactions between two charges, by using Coulomb law is valid also for the case when there is a system of more electric charges, as the force, the two charges interact, is not changed in the presence of another charge. This property defines the content of the *superposition principle*, according to which two sources can be united into a single system, by their *superposing* , without changing the sources configuration.

The superposition principle points out the fact that the force exerted on a charge located in an arbitrary point of the system represents the vectorial sum of all forces acting from each individual source in the system. For instance, in case of three individual charges, whose magnitude and direction are represented by q_1, q_2 and q_3, the force exerted on the third charge is:

$$\mathbf{F}_3 = k_0 \frac{q_3 q_1}{d_{31}^2} \mathbf{u}_{31} + k_0 \frac{q_3 q_2}{d_{32}^2} \mathbf{u}_{32} .$$

This relation outlines another property of Coulomb law, namely the *additivity property* of electric charges action. Coulomb law shows that if electric charges of a system and their coordinates within the system are given, we may determine all the electric forces manifested within the system of charges. The magnitude and direction of force with which the charges of a system act upon another charge located in a arbitrary point of the system are described by means of a specific magnitude, called *electric field*.

In order to define the electric field , we shall consider a system of *n* sources distributed into a three-dimensional space, whose coordinates in this space are represented by triplets $\left(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}\right)$, $\left(x_1^{(2)}, x_2^{(2)}, x_3^{(2)}\right)$,..., $\left(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}\right)$ and whose electric charges have the magnitude and direction represented by values $q_1, q_2, ..., q_n$. The force exerted by the *n* charges of the system on a charge q_0 of some source, located in the point of coordinates $\left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right)$ is given by the relation:

$$\mathbf{F}_0 = \sum_{i=1}^{n} k \frac{q_0 q_i}{d_{0i}^2} \mathbf{u}_{0i},$$

where d_{0i}^2 is the square of the distance from point $\left(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}\right)$ to point $\left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right)$, namely:

$$d_{0i}^2\left(x^{(i)}, x^{(0)}\right) = \left(x_1^{(i)} - x_1^{(0)}\right)^2 + \left(x_2^{(i)} - x_2^{(0)}\right)^2 + ... + \left(x_n^{(i)} - x_n^{(0)}\right)^2 .$$

As it can be noticed, the force $\mathbf{F}_0$ is directly proportional to q_0. If we eliminate q_0 from the relation that defines force $\mathbf{F}_0$ it results a vectorial object, depending on coordinates x_1, x_2 and x_3, magnitude called *electric field* and defined by relation:

$$\mathbf{E}(x_1, x_2, x_3) = \sum_{i=1}^{n} k_0 \frac{q_i}{d_{0i}^2} \mathbf{u}_{0i}.$$

This relation of defining the electric field shows that its magnitude in a point is amplified while approaching to the source and is attenuated while moving away from source, the attenuation being directly proportional to the inverse of distance up to source.

From geometrical point of view, the electric field in each point of the three-dimensional space, may be represented under the form of a vector defining the magnitude and the direction of the electric field in the respective point. Although, it is a three-dimensional vectorial object, the electric field can be geometrically represented also into a two-dimensional space, by means of some arrows, properly oriented, whose length is proportional to the field magnitude. The following figure illustrates the representation of the two-dimensional electric field, for the particular case of two opposite sign charges.

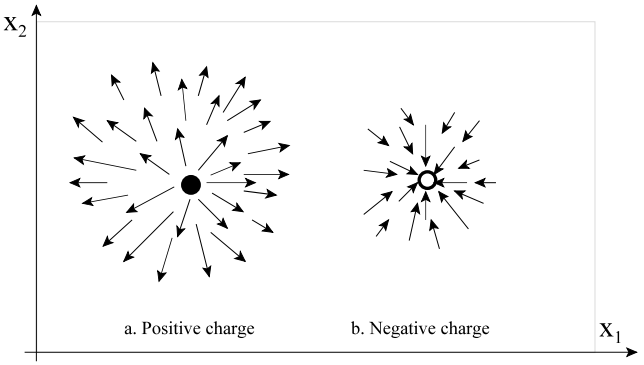

Figure 1: Field representation for two opposite sign charges

To know the vector of the electric field in an arbitrary point of the space is very important to determining the force acting on the charge located in the respective point. At the same time, knowing the electric field in all points of the space, provides a complete description of the whole system of charges, allowing to identify the position and the magnitude of all charges of the system.

In case when the charges are considered not to be of discrete type, but distributed within the space continuously, the distribution of charges in space may be described by means of a charge density function, function defining the charge in each point of the space and which depends on the position of the respective point. If we shall denote by $\rho(x_1, x_2, x_3)$ the charge density in an arbitrary point (x_1, x_2, x_3) then, the magnitude:

$$\rho(x_1, x_2, x_3) dx_1 dx_2 dx_3,$$

where dx_1, dx_2 and dx_3 are positive infinitesimal variations of the three coordinates, represents the charge contained into the infinitesimal volume $dV = dx_1 dx_2 dx_3$, volume located in point (x_1, x_2, x_3). For the continuous case, the value in point (x_1, x_2, x_3) of the electric field determined by the electric charges located in a volume $V = \Delta_{x_1} \Delta_{x_2} \Delta_{x_3}$ is defined by relation:

$$\mathbf{E}(x_1, x_2, x_3) = \int_{\Delta_{x_1}} \int_{\Delta_{x_2}} \int_{\Delta_{x_1}} k_0 \frac{\rho(z_1, z_2, z_3)}{d_{0i}^2} \mathbf{u}_{0i} \, dz_1 \, dz_2 \, dz_3 .$$

Another important magnitude to characterizing the field around an electric charge is the *potential difference* between two arbitrary points of the electric field.

By considering two arbitrary points P_1 and P_2 in a field, the potential difference between the two points is described by the following scalar function:

$$\varphi_{21} = -\int_{P_1}^{P_2} \mathbf{E} \, d\mathbf{s},$$

where $d\mathbf{s}$ is a vector associated to an infinitesimal segment belonging to a path that binds point P_1 with point P_2, vector defined as follows:

$$d\mathbf{s} = \mathbf{x}^{(1)} dx_1 + \mathbf{x}^{(2)} dx_2 + \mathbf{x}^{(3)} dx_3 .$$

In the conditions when point P_1 is set, function φ_{21} depends only on the coordinates of point P_2 so that it may be denoted by:

$$\varphi(x_1, x_2, x_3) = -\int \mathbf{E} \, d\mathbf{s}$$

Function $\varphi(x_1, x_2, x_3)$ is a position scalar function whose value is a number defining the potential associated to the electric field $\mathbf{E}$ in point (x_1, x_2, x_3) and which is called *potential function*.

Bearing in mind the nature of the phenomena, the potential functions can be used in order to characterize the "potential" of some points of the pattern space, in relation to certain classes, which play the role of certain source systems. Accordingly, the potential functions are useful as efficient instruments of pattern classification.

3. CLASSIFICATION USING POTENTIAL FUNCTIONS

The use of potential functions in discriminant analysis relies on the idea that an arbitrary pattern plays the role of a source, with a certain corresponding charge and which generates in a certain neighborhood, a certain field of forces. The following figure illustrates the mode in which a certain potential field is generated around a pattern of the patterns space.

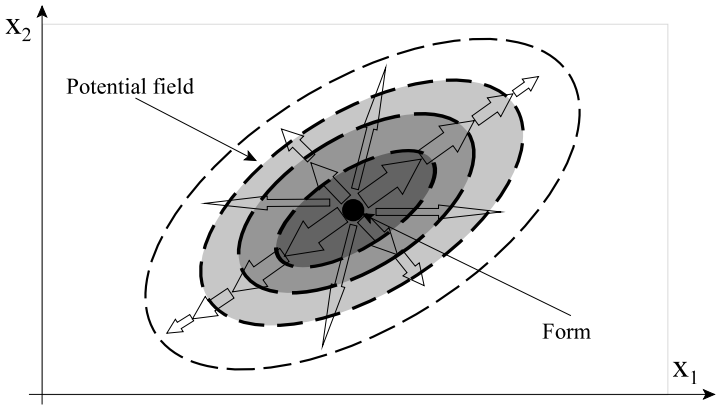

Figure 2. Generating the field around a pattern

By analogy with Coulomb law, the potential magnitude in an arbitrary point of source neighborhood *is inversely proportional to the square of distance* between the respective point and the considered source. The function describing the individual potential magnitude of a source, in an arbitrary point of its neighborhood, is called *individual potential function.* This function has the fundamental property that is *monotonously decreasing* with respect to the square of the distance that separates a point of its definition field from the point defining the charge carrier source.

If we shall denote the individual potential function by $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$, where $\mathbf{x}^{(i)}$ represents the point defining an individual source, then, this function may have one of the following generic patterns:

1. $\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = g\left[\dfrac{1}{d^2(\mathbf{x}, \mathbf{x}^{(i)})}\right]$; 2. $\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = g\left[a^{-d^2(\mathbf{x}, \mathbf{x}^{(i)})}\right]$;

3. $\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = g\left[e^{-d^2(\mathbf{x}, \mathbf{x}^{(i)})}\right]$; 4. $\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = g\left[k - d^2(\mathbf{x}, \mathbf{x}^{(i)})\right]$, $d^2(\mathbf{x}, \mathbf{x}^{(i)}) \le k$.

The value of the potential function in point **x** represents the potential magnitude of source $\mathbf{x}^{(i)}$ in the respective point. As it can be noticed, the individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ is a *nonlinear function* with respect to source **x**.

If we shall denote by a scalar variable z the distance between source **x** and source $\mathbf{x}^{(i)}$, namely $z = d(\mathbf{x}, \mathbf{x}^{(i)})$ then, the individual potential function appears as being a scalar variable function of pattern $\varphi(\mathbf{z})$ and the previous generic patterns become:

1. $\varphi(z) = g\left(\dfrac{1}{z^2}\right)$; 2. $\varphi(z) = g\left(a^{-z^2}\right)$; 3. $\varphi(z) = g\left(e^{-z^2}\right)$; 4. $\varphi(z) = g[k - z]$, $z \le k$.

By taking into account the specific of phenomena behavior which it models, it is necessary as an individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ to verify a series of important properties, such as:

to be continuous, non-negative and monotonously decreasing with respect to the square of the distance towards source, respectively:

$$\varphi(\mathbf{x}, \mathbf{x}^{(i)}) \geq 0; \quad \frac{d\varphi(\mathbf{x}, \mathbf{x}^{(i)})}{d\left[d^2(\mathbf{x}, \mathbf{x}^{(i)})\right]} < 0;$$

· to admit a maximum in the source point, respectively:

$$\varphi(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}) = \max_{\mathbf{x} \in \mathbf{X}} \varphi(\mathbf{x}, \mathbf{x}^{(i)}),$$

where we denoted by $\mathbf{X}$ the domain where can take values $\mathbf{x}$;

- to tend to zero, when the distance towards source increases indefinitely, respectively:

$$\lim_{d^2(\mathbf{x}, \mathbf{x}^{(i)}) \to +\infty} \varphi(\mathbf{x}, \mathbf{x}^{(i)}) = 0;$$

· if the function value is equal for two arbitrary patterns $\mathbf{x}'$ and $\mathbf{x}''$, namely:

$$\varphi(\mathbf{x}', \mathbf{x}^{(i)}) = \varphi(\mathbf{x}'', \mathbf{x}^{(i)}),$$

then, the two points are *equivalent* between them with respect to source $\mathbf{x}^{(i)}$.

The potential function associated to an arbitrary pattern $\mathbf{x}^{(i)}$ of the patterns space may be defined as follows:

Definition: Given a pattern $\mathbf{x}^{(i)}$ of the patterns space, we call *potential function* of pattern $\mathbf{x}^{(i)}$ the *nonlinear* function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ defined as follows:

$$\varphi : \mathbf{X}_\omega \to \Re,$$

with the property that it is *concave* and *monotonously decreasing* with respect to the square of the distance between pattern $\mathbf{x}$ and pattern $\mathbf{x}^{(i)}$, respectively:

$$\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = g\left(\frac{a}{d^2(\mathbf{x}, \mathbf{x}^{(i)})}\right),$$

where $d(\mathbf{x}, \mathbf{x}^{(i)})$ is the distance between pattern $\mathbf{x}$ and pattern $\mathbf{x}^{(i)}$.

In case when the patterns are of two-dimensional or three-dimensional nature, an individual potential function verifying the above-mentioned properties, has the graphs similar to those ones illustrated in the following figures:

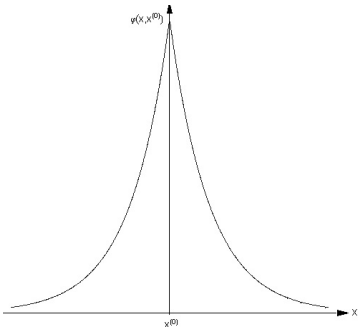

Figure 3: Graph of a one-dimensional potential function

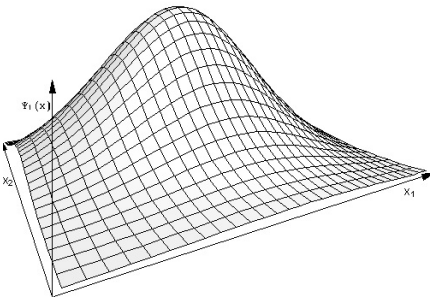

Figure 4: Graph of a two-dimensional potential function

The individual potential functions may be assimilated to some symmetric probability densities of *local* type, where the distribution mean is represented by the pattern whose potential is represented by this function.

Having in view these properties, it follows that potential in an arbitrary point **x**, with respect to the individual source $\mathbf{x}^{(i)}$ may be described, up to a certain proportionality factor, with the aid of an *individual potential function* , as follows:

$$\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = \frac{a}{b + c\, d^2(\mathbf{x}, \mathbf{x}^{(i)})},$$

where $d^2(\mathbf{x}, \mathbf{x}^{(i)})$ is the square of the distance between point **x** and source $\mathbf{x}^{(i)}$ and *a*, *b* and *c* are specific constants.

In the context of supervised pattern recognition, source $\mathbf{x}^{(i)}$ is a known pattern of the training set, and as a distance $d(\mathbf{x}, \mathbf{x}^{(i)})$ we can take any of the known distances, respectively the Euclidian distance, Manhatan distance, Mahalanobis distance etc.

Parameters *a*, *b* and *c* are estimated on the basis of the information supplied by the training set, so that the individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ to allow to get a global potential function, which to describe, as well as possible, the membership to classes of the of the training set patterns, namely of a priori known membership patterns.

Besides the previously presented pattern, there are also another particular possible patterns of the individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ under restriction that all these patterns to verify the previously mentioned properties. For instance, another pattern of the individual potential function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ that is convenient from several points of view, is the following:

$$\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = e^{-\frac{1}{c} d^2(\mathbf{x}, \mathbf{x}^{(i)})},$$

where *c* is a parameter giving shape or allure to function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$. As it can be noticed, this function verifies all the properties previously mentioned.

Excepting a proportionality factor equal to $1/\sqrt{2\pi c}$ this last pattern of function $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$ is identical to the normal probability density, where parameter *c* is the *variance*[1] of the distribution and $\mathbf{x}^{(i)}$ is the vector of the distribution *means*.

In case when in order to evaluate the distance, we use the Euclidian distance, the previous individual potential function gets the following form:

$$\varphi(\mathbf{x}, \mathbf{x}^{(i)}) = e^{-\frac{1}{c} (\mathbf{x} - \mathbf{x}^{(i)})^t (\mathbf{x} - \mathbf{x}^{(i)})}.$$

The most representative category of potential functions is represented by symmetric probability densities, as for instance the normal or exponential density which verifies all the properties of a potential function.

[1] We assume that the homoskedasticity hypothesis is verified, namely the variance is identical according to all axes of the space of patterns.

Generally, we recommend to use as potential function, any simple probability density, which eventually does not depend on any parameter. Such a function may be, for instance, the normal probability density with mean zero and the unit standard deviation.

All *Kernel* functions may be used as individual potential functions to estimating the probability densities. We mention, among them, the most important ones:

• "*rectangle*" function: $\varphi(z)=\begin{cases}\frac{1}{2}, \text{ dacă } |z|\le 1;\\ 0, \text{ dacă } |z|>1;\end{cases}$ • *Gauss* function: $\varphi(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$;

• "*triangle*" function: $\varphi(\mathbf{z})=\begin{cases}1-|z|, \text{ dacă } |z|\le 1;\\ 0, \text{ dacă } |z|>1;\end{cases}$ • *Cauchy* function: $\varphi(z)=\frac{1}{\pi(1+z^2)}$;

• *negative exponential* function: $\varphi(z)=\frac{1}{2}e^{-|z|}$; • *Square Sinus* function: $\varphi(z)=\frac{1}{\pi(1+z^2)}$;

• *Epanechnikov* function: $\varphi(z) = \frac{1}{\pi(1+z^2)}$.

According to the *superposition* principle, if there exists a system composed of several sources denoted by $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(n)}$ the *global potential* in an arbitrary point $\mathbf{x}$ in the neighborhood of this system, is represented by a function that is the sum of functions $\varphi(\mathbf{x}, \mathbf{x}^{(i)})$, respectively:

$$\Psi(\mathbf{x}) = \sum_{i=1}^{n} \varphi(\mathbf{x}, \mathbf{x}^{(i)}),$$

that is called the *potential function* of the system of sources $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(n)}$.

Each term of the sum defining the potential function $\Psi(\mathbf{x})$ of the system of sources $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(n)}$ represents the *contribution* of the individual potential of source $\mathbf{x}^{(i)}$ to forming the global potential of the system of sources in point $\mathbf{x}$.

The *n* sources $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(n)}$ may be considered as representing *n* forms of the training set which belong to a certain given class ω_k so that the potential function $\Psi(\mathbf{x})$ defines the potential of class ω_k in point $\mathbf{x}$ or the potential of class ω_k with respect to pattern $\mathbf{x}$.

Accordingly, it follows that the *potential function* of *class* ω_k may be defined under the following form:

$$\Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_k} \varphi_k(\mathbf{x}, \mathbf{x}^{(i)}), \quad \mathbf{x}^{(i)} \in \omega_k,$$

where T_k is the number of patterns of the training set, belonging to class ω_k.

The next figure presents the graphical representation of three individual potential functions and of the potential function $\Psi(\mathbf{x})$ corresponding to the case of three sources of one-dimensional type.

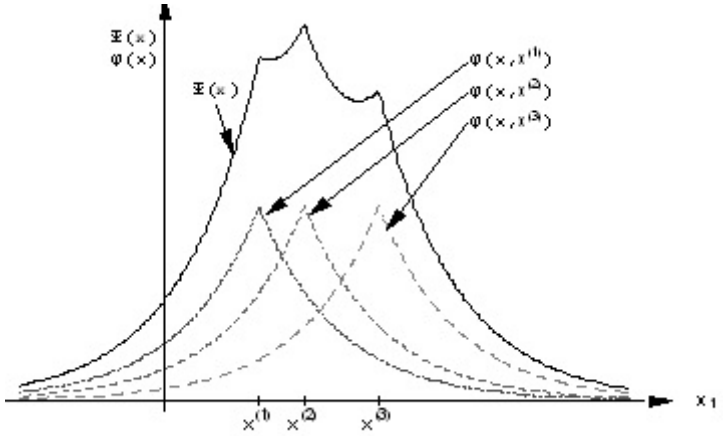

Figure 5: Graph of a global potential function constructed on the basis of three individual potential functions

We want to remark that in most cases, the functions $\varphi_k(\mathbf{x}, \mathbf{x}^{(i)})$ defining the potential function $\Psi(\mathbf{x})$, are assumed to have the *same general pattern* for all the patterns of a class and for all classes, what means that the space of patterns is supposed to have a certain homogeneity from the view point of potentials manifestation. But, on the other side, it is possible as the potential functions $\varphi_k(\mathbf{x}, \mathbf{x}^{(i)})$ to be different from a class to another one, from the view point of the values of the possible parameters which these functions may include.

As a rule, the values of the parameters of a potential function are settled down by taking into consideration the nature of the regions of the decisional space associated to classes and the a priori probabilities of classes, so that to get a classification rate as high as possible.

For the two particular patterns taken into consideration previously, in order to describe the potential in a point, the *global potential function* of class ω_k has the form:

$$\Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_k} \varphi_k(\mathbf{x}, \mathbf{x}^{(i)}) = \sum_{i=1}^{T_k} \frac{a_k}{b_k + c_k d^2(\mathbf{x}, \mathbf{x}^{(i)})}, \qquad \mathbf{x}^{(i)} \in \omega_k,$$

respectively, the form:

$$\Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_i} \varphi_k(\mathbf{x}, \mathbf{x}^{(i)}) = \sum_{i=1}^{T_i} e^{-\frac{1}{c_k} d^2(\mathbf{x}, \mathbf{x}^{(i)})}, \quad \mathbf{x}^{(i)} \in \omega_k.$$

The global potential functions may be used to defining a specific criterion of patterns classification, by considering them as playing the role of classification functions. To this end, the patterns of a certain class are regarded as being charge carrier generating a certain field of forces in the neighborhood of the respective class.

The higher is the field potential of a class in an arbitrary point in the neighborhood of a class, the more justified is the pattern membership to the respective class. So, the potential function $\Psi_{\omega_k}(\mathbf{x})$ may be used also as *classification function* corresponding to class ω_k namely:

$$D_k(\mathbf{x}) = \Psi_{\omega_k}(\mathbf{x}) = \sum_{i=1}^{T_k} \varphi_k(\mathbf{x}, \mathbf{x}^{(i)}), \quad \mathbf{x}^{(i)} \in \omega_k,$$

and pattern **x** will be classified in the class for which the highest value of the classification function $D_i(\mathbf{x})$ is obtained , namely, the highest value of the potential function $\Psi_{\omega_k}(\mathbf{x})$.

We shall suppose that we get a training set including a number of *T* patterns, grouped under the form of *K* classes and T_i is the number of objects of the second *i-th* class.

If we shall consider that $\mathbf{x}^{(k_i)}$ is a pattern of class ω_i and $\mathbf{x}^{(k_j)}$ is a pattern of class ω_j, namely $\mathbf{x}^{(k_i)} \in \omega_i$ and $\mathbf{x}^{(k_j)} \in \omega_j$, the classification criterion of an arbitrary pattern **x** may be stated in terms of *potential functions* , as follows:

• *if*: $\Psi_{\omega_i}(\mathbf{x}) = \sum_{k_i=1}^{T_i} \varphi_i(\mathbf{x}, \mathbf{x}^{(k_i)}) > \Psi_{\omega_j}(\mathbf{x}) = \sum_{k_j=1}^{T_j} \varphi_j(\mathbf{x}, \mathbf{x}^{(k_j)}); \ \forall j = 1,2,...,K; \ i \neq j$, *then*: $\mathbf{x} \in \omega_i$,

and in terms of *classification functions*, as follows:

• *if*: $D_i(\mathbf{x}) > D_j(\mathbf{x}), \quad \forall j = 1,2,...,K; \ i \neq j$, *then*: $\mathbf{x} \in \omega_i$.

As according to the superposition principle, the global potential of a point is defined as sum of the sources potential, it follows that the value of the potential function in a point is higher so as the sources number is higher.

In these conditions, in case when classes are different from the view point of the number of patterns, the highest values of the potential function will be automatically recorded for classes with higher number of objects, so that the classification criterion defined previously can not provide any longer a right classification of patterns.

In order to eliminate this shortcoming, the potential in an arbitrary point of the field induced by a system of *n* sources, may be evaluated under the form of an *average potential*, by normalizing the global potential with respect to the sources number, as follows:

$$\overline{\Psi}(\mathbf{x}) = \frac{1}{n} \Psi(\mathbf{x}).$$

This average potential, evaluated for every class, has a nature similar to the nature of the probability densities of classes. Thus, the classification functions of classes may be defined as average potentials, respectively:

$$D_i(\mathbf{x}) = \overline{\Psi}_{\omega_i}(\mathbf{x}) = \frac{1}{T_i}\Psi_{\omega_i}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i}\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right), \quad i=1,2,...,K,$$

and the classification criterion based on the *potential functions* gets the following form:

• *if:* $\overline{\Psi}_{\omega_i}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i}\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) > \overline{\Psi}_{\omega_j}(\mathbf{x}) = \frac{1}{T_j}\sum_{k_j=1}^{T_j}\varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right), \ \forall j=1,...,K;\ i\neq j$ *,then:* $\mathbf{x}\in\omega_i$.

It follows that the classification method based on potential functions may be defined by means of the following classification criterion:

Potential functions criterion: Given two classes ω_i şi ω_j, whose potential functions are $\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right)$ and $\varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right)$, some pattern $\mathbf{x}\in\mathbf{X}_\omega$ may be classified according to the following rule:

• pattern **x** belong to class ω_i, if:

$$\overline{\Psi}_{\omega_i}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i}\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) > \overline{\Psi}_{\omega_j}(\mathbf{x}) = \frac{1}{T_j}\sum_{k_j=1}^{T_j}\varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right);$$

• pattern **x** belong to class ω_i, if:

$$\overline{\Psi}_{\omega_i}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i}\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) < \overline{\Psi}_{\omega_j}(\mathbf{x}) = \frac{1}{T_j}\sum_{k_j=1}^{T_j}\varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right).$$

As it can be noticed, the potential functions are used as classification functions. It follows that the discriminant functions corresponding to the previously defined criterion may be defined as the differences of the associated potential functions. Therefore, for the case of two classes ω_i and ω_j the discriminat function may be defined as follows:

$$D_{ij}(\mathbf{x}) = \overline{\Psi}_{\omega_i}(\mathbf{x}) - \overline{\Psi}_{\omega_j}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i}\varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) - \frac{1}{T_j}\sum_{k_j=1}^{T_j}\varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right) = D_i(\mathbf{x}) - D_j(\mathbf{x}).$$

From physical point of view, this statement of the discriminant function outlines the fact that the patterns belonging to class ω_i are regarded as being associated to a *positive type charges* , and the patterns belonging to class ω_j are regarded as being associated to *negative type charges*.

The previous classification criterion, stated with the aid of the discriminant functions, gets the following form:

• *if:* $D_{ij}(\mathbf{x}) > 0$, *then:* $\mathbf{x}\in\omega_i$.

The decisional surface corresponding to the discriminant function $D_{ij}(\mathbf{x})$ is constituted of the set of patterns that cancels this discriminant function, namely, of the set of patterns that verifies the equation:

$$D_{ij}(\mathbf{x}) = 0.$$

From the view point of physical rendition, the patterns belonging to the decisional surface define those points where the potential differences between the two classes is null, as:

$$D_{ij}(\mathbf{x}) = 0 \quad \Rightarrow \quad \overline{\Psi}_{\omega_i}(\mathbf{x}) - \overline{\Psi}_{\omega_j}(\mathbf{x}) = 0;$$

$$\Rightarrow \quad \frac{1}{T_i}\sum_{k_i=1}^{T_i} \varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) - \frac{1}{T_j}\sum_{k_j=1}^{T_j} \varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right) = 0.$$

As it can be seen from the nature of the potential functions and from the mode in which the classification criterion is stated, the potential functions method is equivalent to the classification method of patterns, where the classification is done on the basis of evaluating the average distances from the respective patterns to the objects of each of the existing classes.

4. NUMERICAL EXAMPLES

We shall assume that the training set is constituted of a number of seven three-dimensional forms, grouped in two classes ω_1 and ω_2. The information contained in the training set is found in the following table.

Training Set

Table 1

Form	Class	Coordinates		
		x_1	x_2	x_3
$\mathbf{x}^{(1)}$	ω_1	3	1	2
$\mathbf{x}^{(2)}$	ω_1	4	2	1
$\mathbf{x}^{(3)}$	ω_1	3	2	2
$\mathbf{x}^{(4)}$	ω_2	5	3	7
$\mathbf{x}^{(5)}$	ω_2	6	8	4
$\mathbf{x}^{(6)}$	ω_2	4	9	5
$\mathbf{x}^{(7)}$	ω_2	7	5	6

In order to evaluate the individual potential of patterns of the training set, we shall use a Gaussian type potential function of the following form:

$$\varphi\left(\mathbf{x},\mathbf{x}^{(i)}\right) = e^{-\frac{1}{100}d^2\left(\mathbf{x},\mathbf{x}^{(i)}\right)},$$

for which we shall consider that the distance among points is an Euclidian distance. We shall suppose that we have to classify pattern **x** whose coordinates are the followings:

$$\mathbf{x} = (6 \quad 10 \quad 5)^t,$$

for which the average potentials of the two classes are:

$$\overline{\Psi}_{\omega_1}(\mathbf{x}) = \frac{1}{3}\Psi_{\omega_1}(\mathbf{x}) = \frac{1}{3}\sum_{k_i=1}^{3} \varphi\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) = \frac{1}{3}\sum_{k_i=1}^{3} e^{-d^2\left(\mathbf{x},\mathbf{x}^{(k_i)}\right)} = e^{-0,99} + e^{-0,84} + e^{-0,82} = 0,415;$$

$$\overline{\Psi}_{\omega_2}(\mathbf{x}) = \frac{1}{4}\Psi_{\omega_2}(\mathbf{x}) = \frac{1}{4}\sum_{k_j=1}^{4} \varphi\left(\mathbf{x},\mathbf{x}^{(k_j)}\right) = \frac{1}{4}\sum_{k_j=1}^{4} e^{-d^2\left(\mathbf{x},\mathbf{x}^{(k_j)}\right)} = e^{-0,54} + e^{-0,05} + e^{-0,052} + e^{-0,27} = 0,812.$$

As the value of the average potential in point **x** is higher for the second class, respectively:

$$D_2(\mathbf{x}) = \overline{\Psi}_{\omega_2}(\mathbf{x}) = 0,812 > 0,415 = D_1(\mathbf{x}) = \overline{\Psi}_{\omega_1}(\mathbf{x}),$$

pattern **x** will be assigned to class ω_2. At the same time, as the individual potential functions $\varphi\left(\mathbf{x},\mathbf{x}^{(i)}\right)$ of patterns are nonlinear functions, it follows that the average potential function $\overline{\Psi}_{\omega_k}(\mathbf{x})$ of class ω_k is a nonlinear function, what means that also the discriminant functions of classes, defined by relation:

$$D_{ij}(\mathbf{x}) = \overline{\Psi}_{\omega_i}(\mathbf{x}) - \overline{\Psi}_{\omega_j}(\mathbf{x}) = \frac{1}{T_i}\sum_{k_i=1}^{T_i} \varphi_i\left(\mathbf{x},\mathbf{x}^{(k_i)}\right) - \frac{1}{T_j}\sum_{k_j=1}^{T_j} \varphi_j\left(\mathbf{x},\mathbf{x}^{(k_j)}\right),$$

are *nonlinear functions*. Consequently, we may say that the classification method based on potential functions is a *nonlinear type method.*

In order to illustrate the effective significance of the concepts of individual potential function, of average potential function and of discriminant function corresponding to them, we shall consider an example where we shall suppose the existence of a training set with a number of 17 two-dimensional patterns, grouped into two classes ω_1 and ω_2, having the following structure:

$$\omega_1 = \left\{\begin{pmatrix}3\\8\end{pmatrix}; \begin{pmatrix}2\\1\end{pmatrix}; \begin{pmatrix}1\\10\end{pmatrix}; \begin{pmatrix}2\\9\end{pmatrix}; \begin{pmatrix}2\\3\end{pmatrix}; \begin{pmatrix}0\\9\end{pmatrix}; \begin{pmatrix}3\\10\end{pmatrix}\right\};$$

$$\omega_2 = \left\{\begin{pmatrix}4\\3\end{pmatrix}; \begin{pmatrix}7\\12,5\end{pmatrix}; \begin{pmatrix}5,5\\4\end{pmatrix}; \begin{pmatrix}7\\3\end{pmatrix}; \begin{pmatrix}4\\12\end{pmatrix}; \begin{pmatrix}8\\6\end{pmatrix}; \begin{pmatrix}9\\8\end{pmatrix}; \begin{pmatrix}3\\3\end{pmatrix}; \begin{pmatrix}10\\3\end{pmatrix}; \begin{pmatrix}9\\1\end{pmatrix}\right\}.$$

In order to evaluate the individual potential of patterns of the training set, we shall use a Gaussian type potential function of the following form:

$$\varphi\left(\mathbf{x},\mathbf{x}^{(i)}\right) = e^{-\frac{1}{100}d^2\left(\mathbf{x},\mathbf{x}^{(i)}\right)}.$$

Considering the Euclidian distance as the distance among points, the global average potential functions for the two classes are the followings:

$$\bar{\Psi}_{\omega_1}(\mathbf{x}) = \frac{1}{7}\sum_{i=1}^{7}\varphi_1\left(\mathbf{x},\mathbf{x}^{(1_i)}\right) = \frac{1}{7}\sum_{k_1=1}^{7} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)};$$

$$\bar{\Psi}_{\omega_2}(\mathbf{x}) = \frac{1}{10}\sum_{k_2=1}^{10}\varphi_2\left(\mathbf{x},\mathbf{x}^{(k_2)}\right) = \frac{1}{10}\sum_{k_2=1}^{10} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)}.$$

If we shall denote the classification functions of the two classes by $D_1(\mathbf{x})$ and $D_2(\mathbf{x})$ these have the following form:

$$D_1(\mathbf{x}) = \bar{\Psi}_{\omega_1}(\mathbf{x}) = \frac{1}{7}\sum_{k_1=1}^{7} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)}; \quad D_2(\mathbf{x}) = \bar{\Psi}_{\omega_2}(\mathbf{x}) = \frac{1}{10}\sum_{k_2=1}^{10} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)},$$

and the discriminant function corresponding to the two classes is:

$$D_{12}(\mathbf{x}) = D_1(\mathbf{x}) - D_2(\mathbf{x}) = \bar{\Psi}_{\omega_1}(\mathbf{x}) - \bar{\Psi}_{\omega_2}(\mathbf{x}) = \frac{1}{7}\sum_{k_1=1}^{7} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_1)}\right)} - \frac{1}{10}\sum_{k_2=1}^{10} e^{-\frac{1}{100}\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)^t\left(\mathbf{x}-\mathbf{x}^{(k_2)}\right)}.$$

The next figure illustrates graphically the two average global potential functions of classes, in the space of patterns.

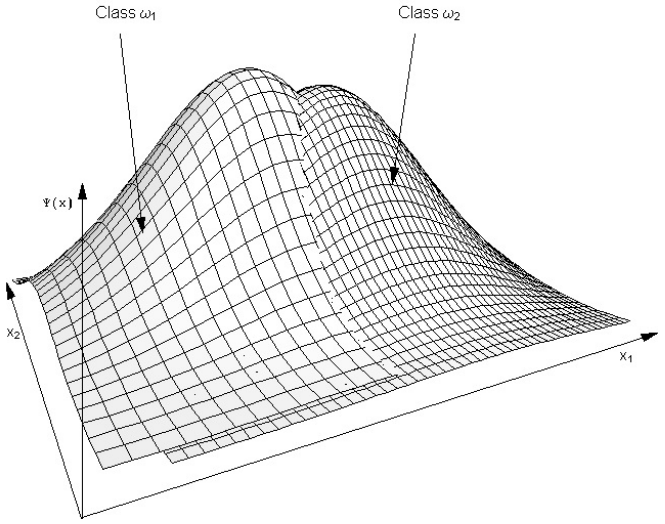

Figura 6: Simultaneous graphical representation of global potential functions $\bar{\Psi}_{\omega_1}(\mathbf{x})$ and $\bar{\Psi}_{\omega_2}(\mathbf{x})$

The same average global potential functions are represented in the next figure by means of level curves.

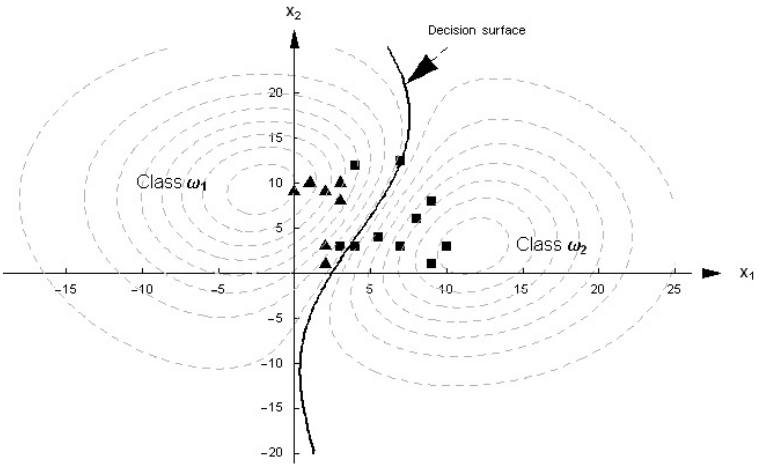

Figura 7: Graphical representation of the level curves of the global potential functions $\overline{\Psi}_{\omega_1}(\mathbf{x})$ and $\overline{\Psi}_{\omega_2}(\mathbf{x})$

As we notice, the decision curve is a nonlinear curve , certifying the fact that the potential functions method is a nonlinear type method. The decision curve presented above provides a relatively good separation of the two classes, specifying that it incorrectly assigns to class ω_1 two patterns of class ω_2.

REFERENCES

[1] **Bakir G., Hofmann T., Schölkopf B., Smola A., Taskar B., Vishwanathan S. V. N.**, ***Predicting Structured Data***. MIT Press, Cambridge, MA, 2007

[2] **Bishop C.M.**, ***Neural Networks for Pattern Recognition***, Oxford Univ. Press, 1995

[3] **Duda O. R., Hart E. P., Stork G. D.**, ***Pattern Classification***, 2nd Edition, Wiley-Interscience, 2001

[4] **Ripley B. D.,** ***Pattern Recognition and Neural Networks,*** Cambridge University Press, Cambridge, UK, 1996

[5] **Ruxanda G.**, ***Supervised pattern recognition with iterative methods*****,** Studii şi Cercetări de Calcul Economic şi Cibernetică Economică, nr. 1/2007, Vol. 47, Academia de Studii Economice, Bucureşti

[6] **Schölkopf B., Smola A. J.**, ***Learning with Kernels***, MIT Press, Cambridge, MA, 2002

[7] **Steinwart I., Christmann A.**, ***Support Vector Machines***, Springer-Verlag, New York, 2008

[8] **Shawe-Taylor J., Cristianini N.**, ***Kernel Methods for Pattern Analysis***, Cambridge Univ. Press, 2004