4 PATTERN CLASSIFICATION USING SECONDARY COMPONENTS PERCEPTRON AND ECONOMIC APPLICATIONS

Daniel CIUIU *

Abstract

In this paper we will classify patterns using a modified Perceptron algorithm (Dumitrache et al., 1999). The generalization uses the eigenvalues and the eigenvectors of the sample covariance matrix, as we did for classifying patterns using PCR (Ciuiu 2007b).We shall also define measurements for the cohesion of the obtained classes and of the separation between them.

The first economic application considered in the paper is a consumer behavior model (Jula 2003), and the second is the same financial application for classifying banks (Ciuiu, 2007a, Ciuiu, 2007b), where we have used regression for classification.

Keywords: Perceptron, principal and secondary components, consumption, banks

JEL Classification: C45, C51, E21, G21

1. Introduction

The Perceptron algorithm (Dumitrache *et al*., 1999) is used for classifying patterns represented by points in R^k in *m* classes. For two classes we consider the hyper-plane:

$$A_0 + \sum_{i=1}^{k} A_i \cdot X_i = 0, \tag{1}$$

and the point $x \in R^k$ is in the first class if in the above relation we have ">" instead of "=", and in the second class if we have "<" instead of "=".

In the classical Perceptron learning algorithm with two classes we consider a sample $X^{(1)},...,X^{(n)}$ and arbitrary starting values of $\left(A_j\right)_{j=\overline{0,k}}$.

Consider the obtained result $y^{(i)} \in \{-1,1\}$ with the signification that $y^{(i)} = 1$ if $X^{(i)}$ is in the first class and $y^{(i)} = -1$ in the contrary case. Setting $A_0 = 0$ and denoting by $t^{(i)}$ the desired result (we know the class of $X^{(i)}$), the Perceptron learning algorithm modifies the values of A_j (Dumitrache *et al.*, 1999) by the formula:

$$A_j \leftarrow A_j + \alpha \cdot \left(t^{(i)} - y^{(i)}\right) X_j^{(i)}, \tag{2}$$

where: $X^{(i)}$ is the current point, $y^{(i)}$ is the obtained result using the current A_j, $t^{(i)}$ is the desired result and $\alpha \in (0,1)$ is the learning factor.

In the case of *m* classes, there is a hyper-plane that separates each class from the other ones. If these hyper-planes are given by $A_j^{(r)}$ with $1 \le j \le k$ and $1 \le r \le m$, and the point $X^{(i)}$ from the class *r* is classified in the class $t \ne r$ we have, with the above significations (Kong and Kosko, 1992),

$$\begin{cases} A_j^{(r)} \leftarrow A_j^{(r)} + \alpha \cdot X_j^{(i)} \\ A_j^{(t)} \leftarrow A_j^{(t)} - \alpha \cdot X_j^{(i)} \end{cases}. \tag{2'}$$

One may see that (2') is the generalization of (2) because at each moment we have $A_j^{(2)} = -A_j^{(1)} = -A_j$. In fact, in (2) and (2') the coefficient of $X_j^{(i)}$ is $\alpha \cdot ERR_r$, where

$$ERR = ERR_1 = t^{(i)} - y^{(i)} \tag{3}$$

in the case of formula (2), and

$$\begin{cases} ERR_r = 1 \\ ERR_t = -1 \end{cases} \tag{3'}$$

in the case of formula (2').

Such artificial neural networks using Perceptron can be used to forecast the exchange rate of euro versus RON (Nastac *et al.*, 2007). In this case, we do not set $A_j^{(i)} = 0$: the biases are modified in the same way as the other coefficients. Another difference in Nastac *et al.*, (2007) is that the error is not discrete as in (3) and (3'): it is continuous, as one may see in the following formula:

$$ERR = ERR_1 = \frac{100}{T} \cdot \sum_{p=1}^{T} \frac{\left|O_{Rp} - O_{Fp}\right|}{\left|O_{Rp}\right|} \cdot f(p), \tag{3''}$$

where: *T* is the number of time steps (days), O_{Rp} is the real output at time step *p*, O_{Fp} is the forecast output at time step *p* and $f(p)=\frac{T}{T+p}$ is a weight function that decreases with the number of time steps *p*.

In the next section, we shall modify this algorithm by reading first all the learning sequence $X^{(1)},...,X^{(n)}$, the values of A_j being computed using the sample covariance matrix. Therefore, we will have the same separators for all classes: the eigenvectors of the covariance matrix.

Let be *n* points in $\mathbf{R}^p$: $X^{(1)},...,X^{(n)}$. The orthogonal linear variety of the dimension *k* (0<*k*<p) is that linear variety with the minimum sum of the squares of Euclidean distances. We know (Saporta, 1990) that this linear variety contains the gravity center of *n* given points and it is generated by the eigenvectors of the sample covariance matrix corresponding to the first maximum *k* eigenvalues. These eigenvectors are called principal components, and for that the orthogonal regression is also called principal components regression (*PCR*).

The principal components analysis is used to simplify the computations in the discriminant analysis by using the Kolmogoroff distance (Saporta and Mahjoub, 1990), and the *PCR* is used to find the eigenvalues end eigenvectors (Costinescu and Ciuiu, 2007) of a symmetric matrix and in for pattern classification (Ciuiu, 2007b). For pattern classification we can also use linear and polynomial regression (Ciuiu, 2007a).

2. Classification using SCP

We consider *n* points $X^{(1)},X^{(2)},...,X^{(n)}$ in $\mathbf{R}^k$. In the classical Perceptron algorithm we have a hyper-plane, and we change the coordinates using the exchange theorem starting from the perpendicular to the hyper-plane. The two obtained classes depend on the sign of this component in the new coordinates. In order to have a good classification, we must have large distances to the hyper-plane.

If we want small ones, we have to use the principal components regression (*PCR*), but the orthogonal regression hyper-plane contains the gravity center of the points and it is generated by the corresponding eigenvectors of the maximum $k-1$ eigenvalues (Saporta, 1990). Therefore, we can build in an analogous manner the hyper-plane used in the Perceptron algorithm: the only difference is that it is generated by the corresponding eigenvectors of the minimum $k-1$ eigenvalues. This idea comes from the fact that the average of the square of distances in *PCR* is given by the smallest eigenvalue. In our case, this average becomes the highest one.

First, we compute the sample covariance matrix Σ, we move the origin in the gravity center of the points, *G*, and next we change the coordinates to the eigenvectors of Σ. Let us suppose the corresponding eigenvalues are ordered increasingly. If the matrix *U* has these eigenvectors in rows the new coordinates are:

$$y^{(i)}=U\cdot X^{(i)}. \tag{4}$$

In the following, we separate the *k* eigenvectors in *sec* secondary components (the first ones) and $\mathrm{princ} = \mathrm{k} - \mathrm{sec}$ principal components. Continuing the analogy to the *PCR*, where the orthogonal linear variety of the dimension *dim* is generated by the corresponding eigenvectors of the highest *dim* eigenvalues (Ciuiu, 2007b), in the algorithm presented in the paper we use the linear variety of the dimension *dim* that contains the gravity center *G* and it is generated by the corresponding eigenvectors of the smallest *dim* eigenvalues. *G* is the new origin of the coordinate system. The reason for the linear variety containing *G* as a new origin is the desired equidistance of the algorithm to all the classes. The secondary components are defined also by analogy with the principal components used in *PCR*.

The classes are built by the signs of the principal components: two points are in the same class if and only if they have the same signs on the principal components. In the following, we present two measurements of our classification.

Definition 1. *Let $\mathbf{C}$ be a class obtained as above. The cohesion of the class $\mathbf{C}$ is the sum of the second sample moments of $\mathbf{C}$ on the principal components minus the sum of the sample variances on the secondary ones.*

Definition 2. *Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two classes obtained as above. The separation between $\mathbf{C}_1$ and $\mathbf{C}_2$ is the sum of the second sample moments of $\mathbf{C}_1 \cup \mathbf{C}_2$ on the principal components for which the sign changes from $\mathbf{C}_1$ to $\mathbf{C}_2$ minus the sum of the sample variances on the secondary ones.*

Remark 1. *The sample moments and variances in the above two definition are computed using the points from the class in the case of cohesion, and from the two classes in the case of separation. For both cases we consider the new coordinates, where the covariances are equal to 0.*

3. Margins for cohesions and separations

For lower and upper margins for cohesions and separations we will denote by $\mathbf{S}\mathrm{ec}$ the set of the secondary components, by $\mathbf{P}\mathrm{rinc}$ the set of the principal components and by $\mathbf{S}\mathrm{ep}(\mathbf{C}_1, \mathbf{C}_2)$ the set of the components that separate $\mathbf{C}_1$ and $\mathbf{C}_2$. We also use the following notations:

$$\begin{cases} S_{\mathrm{sec}}^2 = \sum\limits_{i \in \mathbf{S}\mathrm{ec}} \lambda_i \\ S_{\mathrm{princ}}^2 = \sum\limits_{i \in \mathbf{P}\mathrm{rinc}} \lambda_i \\ S_{\mathrm{sep}}^2(\mathbf{C}_1, \mathbf{C}_2) = \sum\limits_{i \in \mathbf{S}\mathrm{ep}(\mathbf{C}_1, \mathbf{C}_2)} \lambda_i \end{cases}, \quad (5)$$

where: $\mathbf{C}_1$ and $\mathbf{C}_2$ are two distinct classes and the eigenvalues of the sample covariance matrix are $\lambda_1 \le \lambda_2 \le ... \le \lambda_k$.

The cohesion of the class $\mathbf{C}$ with *m* points is:

$$\begin{aligned}\mathrm{coh}(\mathbf{C}) &= \tfrac{1}{m}\cdot\sum_{i\in\mathbf{P}\mathrm{rinc}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)}\right)^2 - \tfrac{1}{m}\cdot\sum_{i\in\mathbf{S}\mathrm{ec}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)} - G_i(\mathbf{C})\right)^2 \\ &= \tfrac{1}{m}\cdot\sum_{i\in\mathbf{P}\mathrm{rinc}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)}\right)^2 - \tfrac{1}{m}\cdot\sum_{i\in\mathbf{S}\mathrm{ec}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)}\right)^2 + \sum_{i\in\mathbf{S}\mathrm{ec}}\left(G_i(\mathbf{C})\right)^2\end{aligned}, \qquad (6)$$

where: $G_i(\mathbf{C})$ is the component *i* of the gravity center of the class $\mathbf{C}$.

It results that

$$\mathrm{coh}(\mathbf{C}) \geq -\frac{1}{m}\cdot\sum_{i\in\mathbf{S}\mathrm{ec}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)}\right)^2 \geq -\frac{n}{m}\cdot S_{\mathrm{sec}}^2 \text{ and}$$

$$\mathrm{coh}(\mathbf{C}) \leq \frac{1}{m}\cdot\sum_{i\in\mathbf{P}\mathrm{rinc}}\sum_{j\in\mathbf{C}}\left(X_i^{(j)}\right)^2 \leq \frac{n}{m}\cdot S_{\mathrm{princ}}^2, \text{ and from here}$$

$$-\frac{n}{m}\cdot S_{\mathrm{sec}}^2 \leq \mathrm{coh}(\mathbf{C}) \leq \frac{n}{m}\cdot S_{\mathrm{princ}}^2 . \qquad (6')$$

If the first "$\leq$" in (6') is, in fact, "=" we must have all the points of the class with the principal coordinates equal to 0 (see the first term in the definition of the cohesion). This cannot be fulfilled because in this case the class can be attached to another class. If the second "$\leq$" in (6') is, in fact, "=" we must have all the points of the other classes with the principal coordinates equal to 0, because the sum of squares on the principal components for the class is equal to those for al the points. This cannot be fulfilled because in this case we can have only one class.

For the separation between $\mathbf{C}_1$ and $\mathbf{C}_2$, $\mathrm{sep}(\mathbf{C}_1, \mathbf{C}_2)$, we obtain the same margins as in (6'), if in this case *m* becomes the number of points in the classes $\mathbf{C}_1$ and $\mathbf{C}_2$. To obtain these margins, the sums from (6) on $i \in \mathbf{P}\mathrm{rinc}$ are in the separation case on $i \in \mathbf{S}\mathrm{ep}(\mathbf{C}_1, \mathbf{C}_2)$, and we take into account that $S_{\mathrm{sep}}^2 \leq S_{\mathrm{princ}}^2$, but $S_{\mathrm{sep}}^2 = S_{\mathrm{princ}}^2$ only if all the principal components separate the classes. Otherwise, we obtain better borders:

$$-\frac{n}{m}\cdot S_{\mathrm{sec}}^2 \leq \mathrm{sep}(\mathbf{C}_1, \mathbf{C}_2) \leq \frac{n}{m}\cdot S_{\mathrm{sep}}^2(\mathbf{C}_1, \mathbf{C}_2). \qquad (7)$$

If the first "$\leq$" in (7) is in fact "=" we must have all the points of the two classes with the coordinates that separate the classes equal to 0, for analogues reasons as for the cohesions. This cannot be fulfilled because in this case the classes can be grouped in only one class.

Suppose that the second "$\leq$" in (7) is, in fact, "=" and we have at least 3 classes. It results that the points from $\mathbf{C}_1 \cup \mathbf{C}_2$ have the same values for each secondary component and for another class $\mathbf{C}_3$ the components that separate $\mathbf{C}_1$ and $\mathbf{C}_2$ must be 0. In this case, we can delete first the components that separate $\mathbf{C}_1$ and $\mathbf{C}_2$ to classify the points using the other components (there exists at least one other principal component to separate, for instance, $\mathbf{C}_1$ and $\mathbf{C}_3$). An obtained class will be $\mathbf{C}_1 \cup \mathbf{C}_2$ and the other classes are the class from the previous classification. Then, we can use the deleted components to separate $\mathbf{C}_1$ and $\mathbf{C}_2$. If we have only two classes and we set to 0 the secondary components (the same value, if we have at least 3 classes, becomes 0 because this is the gravity center) the second "$\leq$" in (7) is "$=$", as we can see from computation.

In the following, we consider as fixed $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_k$, the number of secondary components being *s* with 0<*s*<*k* and the number of principal components being $p = k - s$.

Definition 3. *Let* $\mathbf{C}$ *be a given class with m points. The proximity to border of the cohesion for the class* $\mathbf{C}$ *is the value* $\frac{m \cdot coh(\mathbf{C})}{n \cdot S_{sec}^2}$ *if* $coh(\mathbf{C}) \leq 0$ *and* $\frac{m \cdot coh(\mathbf{C})}{n \cdot S_{princ}^2}$ *in the contrary case.*

Definition 4. *Let* $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *be two given classes with m points together. The proximity to border of the separation for the classes* $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *is the value* $\frac{m \cdot sep(\mathbf{C}_1, \mathbf{C}_2)}{n \cdot S_{sec}^2}$ *if* $sep(\mathbf{C}_1, \mathbf{C}_1) \leq 0$ *and* $\frac{m \cdot sep(\mathbf{C}_1, \mathbf{C}_2)}{n \cdot S_{sep}^2(\mathbf{C}_1, \mathbf{C}_2)}$ *in the contrary case.*

We denote by $prox_1(\mathbf{C})$ the proximity to border of the cohesion for the class $\mathbf{C}$ and by $prox_2(\mathbf{C}_1, \mathbf{C}_2)$ the proximity to border of the separation for the classes $\mathbf{C}_1$ and $\mathbf{C}_2$.

First, we will give an example so that $prox_1(\mathbf{C}) \to -1$ and $\frac{n \cdot S_{sec}^2 \cdot (prox_1(\mathbf{C}) + 1)}{m} \to 0$. For this, we must have $S_{sec}^2 > 0$.

Instead of *n,* we take the number of points as $n_1 \cdot n \cdot 2^{p-1}$. The first class, $\mathbf{C}_1$, has $m = n_1 \cdot (n-1)$ points so that for this class we have $\overline{X_i} = \overline{X_j} = \overline{X_i \cdot X_j} = 0$ for any *i,j* with $0 < i < j \leq s$, the sample variance of X_i is $S_i^2 = \left(\frac{n \cdot 2^{p-1}}{n-1} - \frac{1}{(n-1)^2}\right) \cdot \lambda_i$ for $i = \overline{1, s}$, and $X_i = \alpha_i$ for $i = \overline{s+1, k}$.

If $p > 1$, we choose one of the principal components, say X_t, and we build other $2^{p-1} - 1$ classes with *m* points such the first condition is further fulfilled, the sample variances on the secondary components become $S_i^2 = \frac{\lambda_i}{(n-1)(n \cdot 2^{p-1} - n + 1)}$, and the third condition is fulfilled only for $i = t$: the other principal components have the same absolute value, but all the possible signs.

For each class from the already 2^{p-1} defined classes we define a new class with n_1 points so that the first two conditions of the last $2^{p-1}-1$ classes are fulfilled, and we have $X_i = -(n-1)\beta_i$ if $i > s$ and $X_i = \beta_i$ in the corresponding class.

By computation, we can prove that for all the $n_1 \cdot n \cdot 2^{p-1}$ points the first condition is fulfilled, and for any secondary component we have $S_i^2 = \lambda_i$. If we consider the last condition for the principal components we obtain $\alpha_i^2 = \frac{\lambda_i}{n-1}$. The cohesion of the first class is $\mathrm{coh}(\mathbf{C}_1) = -\frac{n \cdot 2^{p-1}}{n-1} \cdot S_{sec}^2 + \frac{S_{sec}^2}{(n-1)^2} + \frac{S_{princ}^2}{n-1}$. The desired condition for this cohesion can be checked by computation.

If we want to give an analogue example for separation, we must have at least two principal components. The only conditions from the example of cohesions that we modify are those regarding the sample variances. We choose the first two from the first 2^{p-1} classes and for both we set the sample variances for $i = \overline{1,s}$ as $S_i^2 = \left(\frac{n \cdot 2^{p-2}}{n-1} - \frac{1}{2(n-1)^2}\right) \cdot \lambda_i$. For the other classes we set for the secondary components $S_i^2 = \frac{\lambda_i}{(n-1)(n \cdot 2^{p-1} - 2 \cdot n + 2)}$.

We can prove that in this case, for all the $n_1 \cdot n \cdot 2^{n-1}$ points, we have $\overline{X_i} = \overline{X_j} = \overline{X_i \cdot X_j} = 0$ for $0 < i < j \le k$ and $S_i^2 = \lambda_i$ for $i = \overline{1,k}$. The separation between these classes is $\mathrm{sep}(\mathbf{C}_1, \mathbf{C}_2) = -\frac{n \cdot 2^{p-2}}{n-1} \cdot S_{sec}^2$

$+ \frac{S_{sec}^2}{2(n-1)^2} + \frac{S_{sep}^2(\mathbf{C}_1, \mathbf{C}_2)}{n-1}$. The desired conditions can be also checked by computation.

If we take in the above example with $\mathrm{prox}_1(\mathbf{C}) \to -1$ only one principal component we have only two classes, and we obtain by computation $\mathrm{coh}(\mathbf{C}_2) = (n-1) \cdot S_{princ}^2 - \frac{S_{sec}^2}{n-1}$. It results that $\mathrm{prox}_1(\mathbf{C}_2) \to 1$ and $n \cdot S_{princ}^2 \cdot (1 - \mathrm{prox}_1(\mathbf{C}_2)) \to S_{princ}^2$. We notice that the second limit in the case of $\mathbf{C}_1$ is for the difference between the proximity and -1 multiplied by the minimum cohesion. In the case of $\mathbf{C}_2$, the second limit is for the difference between the proximity and 1 multiplied by the maximum cohesion.

For the example with $\mathrm{prox}_2(\mathbf{C}_1, \mathbf{C}_2) \to -1$ we take $p = 2$, and we obtain analogously 4 classes and $\mathrm{sep}(\mathbf{C}_3, \mathbf{C}_4) = (n-1) \cdot S_{sep}^2(\mathbf{C}_3, \mathbf{C}_4) - \frac{S_{sec}^2}{2(n-1)}$. It results that $\mathrm{prox}_2(\mathbf{C}_3, \mathbf{C}_4) \to 1$ and $n \cdot \mathrm{sep}(\mathbf{C}_3, \mathbf{C}_4) \cdot$

$\left(1-\text{prox}_2\left(\mathbf{C}_3,\mathbf{C}_4\right)\right)\rightarrow \text{sep}\left(\mathbf{C}_3,\mathbf{C}_4\right)$.

From (6'), it results that for each class the minimum cohesion is increasing by the number of points, and the maximum one is decreasing. Both borders have the same values for the same number of points.

The results from the examples in the next section are obtained by our C++ program called "percepDlg.cpp".

4. Economic applications

Example 1. Consider the following consumer behavior model with 25 customers, where X_1 represents the advertisement, X_2 represents the prices and X_3 represents the sales (Jula 2003):

X_1	3	2	0.8	2.5	2	1.4	2.5	2.5	3	1.4	1	1.2	1.6
X_2	1.3	2.8	1.5	0.2	1.8	4	1.8	2	0.5	2.8	3.2	2.5	1.3
X_3	2	0.5	1.5	3	1	0	2.1	1.8	3	0.7	0.5	1	1.4

X_1	1.8	1	2.8	3.5	2.6	2.4	3.4	1.6	1.9	3.5	1.6	3
X_2	2.2	3.5	1.1	0	0.2	2	1.2	3	3	0.6	3.2	0.3
X_3	1.2	0.8	2.3	3.5	3.8	1.8	2.6	0.8	1.2	4.2	0.8	2.5

If we consider two secondary components and one principal component we obtain the following two classes:

$\mathbf{C}_1 = \{1,4,7,8,9,16,17,18,19,20,23,25\}$ with 12 customers, and

$\mathbf{C}_2 = \{2,3,5,6,10,11,12,13,14,15,21,22,24\}$ with 13 customers.

The cohesions of the two classes are 2.97208 and 2.14801, respectively, the minimum cohesions are -0.63964 and -0.59044, respectively, the maximum cohesions are 5.90636 and 5.4511, respectively, and the proximities to border are 0.50329 and 0.39405, respectively.

The separation between the classes is 2.52754, the minimum separation between the classes is -0.30703, the maximum separation between the classes is 2.83457, and the proximity to border is 0.89168.

If we consider one secondary component and two principal components we obtain the following 4 classes:

$\mathbf{C}_1 = \{1,7,8,16,17,19,20,23\}$ with 8 customers,

$\mathbf{C}_2 = \{2,6,21,22,24\}$ with 5 customers,

$\mathbf{C}_3 = \{3,5,10,11,12,13,14,15\}$ with 8 customers and

$\mathbf{C}_4 = \{4,9,18,25\}$ with 4 customers.

The cohesions of the four classes are 2.66277, 3.54305, 2.09161 and 4.45726, respectively, the minimum cohesions of the classes are -0.31024, -0.49638, -0.31024 and -0.62047, respectively, the maximum cohesions of the classes are 9.50726, 15.21162, 9.50726 and 19.01453, respectively, and the proximities to border are 0.28088, 0.23292, 0.22 and 0.23441, respectively.

The separations between classes alphabetically ordered (between the classes 1 and 2, between the classes 1 and 3,..., between the classes 2 and 3, and so on) are: 2.85538, 2.37717, 0.03508, 0.17624, 3.94918 and 2.60534, respectively, the minimum separations between classes alphabetically ordered are -0.19091, -0.15512, -0.20682, -0.19091, -0.27577 and -0.20682, respectively, the maximum separations between classes alphabetically ordered are 5.4511, 4.75363, 0.43282, 0.39952, 8.4509 and 5.90536, respectively, and the proximities to borders alphabetically ordered are 0.52382, 0.50007, 0.08105, 0.44112, 0.446731 and 0.44118, respectively.

Example 2. We have the following 29 banks, where X_1 is the annual interest for an account without term, X_2 is the annual interest for an account with one-month term, X_3 is the annual interest for an account with three-month term, X_4 is the annual interest for an account with six-month term, X_5 is the annual interest for an account with nine-month term and X_6 is the annual interest for an account with one-year term (Ciuiu, 2007a, Ciuiu, 2007b).

Bank	X_1	X_2	X_3	X_4	X_5	X_6
ABN-Amro Romania	0.25%	3.5%	3.75%	3.75%	0%	3.75%
Alpha Bank	0.1%	6.25%	6.5%	7%	7%	7.25%
Banc Post	0%	7.25%	7.25%	7.15%	0%	7.15%

Banca Comercialã Carpatica	1%	7.5%	7.55%	7.6%	7.75%	7.8%
BCR	0.25%	6%	6.25%	6.5%	6.75%	7.5%
Banca Italo-Romena	0%	5.5%	5.75%	6%	6.15%	6.25%
Banca Româneascã	0.75%	7.3%	7.75%	8.05%	8.1%	8.1%
Banca Transilvania	0.25%	7.5%	7.5%	7.5%	7.75%	7.75%
Bank Leumi Romania	0.25%	7.5%	7.5%	7.75%	7.75%	8%
Blom Bank Egypt	0.1%	6%	6.5%	6.5%	6.75%	7%
BRD-Groupe Société Générale	0.25%	5.5%	5.6%	5.65%	5.65%	5.75%
C.R. Firenze Romania	0.1%	6.5%	6.75%	7%	7.25%	7.5%
CEC	0.25%	7%	7%	7.25%	0%	7.25%
Citibank Romania	1%	4.28%	4.28%	4.28%	3.87%	3.46%
Emporiki Bank	0.5%	6.75%	7%	7.25%	7%	7%
Finansbank	0.1%	7.5%	8%	8%	8%	8.5%
HVB-Þiriac Bank	0.1%	6.4%	6.3%	6.2%	6.1%	6.1%
ING Bank	6.85%	5.5%	5.75%	6%	6.25%	6.5%
Libra Bank	0%	8%	8.1%	7.6%	7.6%	8.5%
Mind Bank	0.25%	7%	7%	7.25%	7.5%	7.75%
OTP Bank	0.25%	6.25%	6.5%	7%	7%	7.25%
Piraeus Bank	0.5%	7%	7.1%	7.25%	7.1%	7.35%
Pro Credit Bank	7%	7.5%	7.65%	7.7%	0%	7.85%
Raiffeisen Bank	0.25%	4%	4.25%	4.5%	4.6%	4.75%
Romanian International Bank	0.25%	6.5%	6.75%	7%	7.5%	7.75%
Romexterra	0.25%	7.5%	7.75%	7.75%	8.1%	8.1%
San Paolo IMI Bank	0.1%	6.5%	6.7%	6.8%	7%	7.2%
Uni Credit Romania	0.1%	5%	5%	5.25%	5.5%	5.5%
Volksbank	0.1%	4.5%	4.75%	4.5%	3.5%	3.25%

In the above table the null values have the signification that we cannot open such accounts with those banks.

If we consider five secondary components and one principal component we obtain the following 2 classes:

$\mathbf{C}_1 = \{$*ABN-Amro Romania, Banc Post, Banca Italo-Romena, BRD-Groupe Société Générale, CEC, Citibank Romania, HVB-Țiriac Bank, ING Bank, Pro Credit Bank, Raiffeisen Bank, Uni Credit Romania, Volksbank*$\}$ with 12 banks, and $\mathbf{C}_2 = \{$*Alpha Bank, Banca* Comercială Carpatica, BCR, Banca Românească, Banca Transilvania, Bank Leumi Romania, Blom Bank Egypt, C.R. Firenze Romania, Emporiki Bank, Finansbank, Libra Bank, Mind Bank, OTP Bank, Piraeus Bank, Romanian International Bank, Romexterra, San Paolo IMI Ba*nk*} with 17 banks.

The cohesions of the 2 classes are -1.11317 and 5.06516, respectively, the minimum cohesions are -15.91748 and -11.23587, respectively, the maximum cohesions are 21.87761 and 15.44302, respectively, and the proximities to border are -0.06993 and 0.32799, respectively.

The separation between the classes is 2.46626, the minimum separation between the classes is -6.58654, the maximum separation between the classes is 9.05281, and the proximity to border is 0.27243.

If we consider four secondary components and two principal components we obtain the following 4 classes:

$\mathbf{C}_1 =$ {*ABN-Amro Romania, Banca Italo-Romena, BRD-Groupe Société Générale, Citibank Romania, HVB-Țiriac Bank, Raiffeisen Bank, Uni Credit Romania, Volksbank*} with 8 banks, $\mathbf{C}_2 =$ {*Alpha Bank, BCR, Blom Bank Egypt, C.R. Firenze Romania, Emporiki Bank, Mind Bank, OTP Bank, Romanian International Bank, San Paolo IMI Bank*} with 9 banks, $\mathbf{C}_3 =$ {*Banc Post, CEC, ING Bank, Pro Credit Bank*} with 4 banks and $\mathbf{C}_4 =$ {*Banca Comercială Carpatica, Banca Românească, Banca Transilvania, Bank Leumi Romania, Finansbank, Libra Bank, Piraeus Bank, Romexterra*} with 8 banks.

The cohesions of the 4 classes are 19.06344, 2.47738, 16.92442 and 9.32004, respectively, the minimum cohesions of the classes are -8.65729, -7.69537, -17.31458 and -8.65729, respectively, the maximum cohesions of the classes are 48.03535, 42.69809, 96.0707 and 48.03535, respectively, and the proximities to border are 0.39686, 0.05802, 0.17617 and 0.19402, respectively.

The separations between alphabetically ordered classes are: 8.22855, 4.34508, 14.13648, 6.90053, 0.07626 and 4.48739, respectively, the minimum separations between alphabetically ordered classes are: -4.07402, -5.77153, -4.32805, -5.32756, -4.07402 and -5.77153, respectively, the maximum separations between alphabetically ordered classes are: 15.44302, 10.14595, 24.01767, 29.56021, 7.16185 and 21.87761, respectively, and the proximities to alphabetically ordered borders are: 0.53283, 0.42826, 0.58859, 0.23344, 0.01065 and 0.20511, respectively.

If we consider three secondary components and three principal components, we obtain the following eight classes:

$\mathbf{C}_1 =$ {*ABN-Amro Romania, HVB-Țiriac Bank*} with 2 banks, $\mathbf{C}_2 =$ {*Alpha Bank, Blom Bank Egypt, C.R. Firenze Romania, Mind Bank, San Paolo IMI Bank*} with 5 banks, $\mathbf{C}_3 =$ {*Banc Post, CEC*} with 2 banks, $\mathbf{C}_4 =$ {*Banca Comercială Carpatica, Banca*

Românească} with 2 banks, $\mathbf{C}_5 =${*BCR, Emporiki Bank, OTP Bank, Romanian International Bank*} with 4 banks, $\mathbf{C}_6 =${*Banca Italo-Romena, BRD-Groupe Société Générale, Citibank Romania, Raiffeisen Bank, Uni Credit Romania, Volksbank*} with 6 banks, $\mathbf{C}_7 =${*Banca Transilvania, Bank Leumi Romania, Finansbank, Libra Bank, Piraeus Bank, Romexterra*} with 6 banks and $\mathbf{C}_8 =${*ING Bank, Pro Credit Bank*} with 2 banks.

The cohesions of the eight classes are 33.53253, 2.74579, 35.22195, 9.7789, 2.18437, 15.09169, 9.61425 and 58.47459, respectively, the minimum cohesions of the classes are: -2.48642, -0.99497, -2.48742, -2.48742, -1.24371, -0.82914, -0.82914 and -2.48642, respectively, the maximum cohesions of the classes are: 224.28314, 89.71326, 224.28314, 224.28314, 112.14157, 74.76105, 74.76105 and 224.28314, respectively, and the proximities to borders are: 0.14951, 0.03061, 0.15704, 0.0436, 0.01948, 0.20187, 0.1286 and 0.26072, respectively.

The separations between alphabetically ordered classes are: 10.71519, 5.89402, 21.62663, 11.91814, 0.03937, 15.32456, 26.0841, 8.68152, 0.32871, -0.09373, 6.72446, 0.18815, 18.66331, 16.99085, 13.185, 9.07563, 10.12034, 14.95782, 0.10486, 13.48898, 0.12341, 8.57439, 6.76371, 0.1971, 14.82534, 12.35024, 11.33962, and 13.60276, respectively, the minimum separations between alphabetically ordered classes are: -0.71069, -1.24371, -1.24371, -0.82914, -0.62186, -0.62186, -1.24371, -0.71069, -0.71069, -0.55276, -0.45226, -0.45226, -0.71069, -1.24371, -0.82914, -0.62186, -0.62186, -1.24371, -0.82914, -0.62186, -0.62186, -1.24371, -0.49748, -0.49748, -0.82914, -0.41457, -0.62186, and -0.62186, respectively, the maximum separations between alphabetically ordered classes are: 37.50448, 30.43786, 112.14157, 54.46914, 8.03544, 48.03535, 46.50873, 54.89754, 26.57642, 7.14261, 29.71044, 11.06831, 64.0809, 81.70371, 74.76105, 23.25437, 32.81642, 16.07087, 20.29191, 48.03535, 8.03544, 65.63284, 26.25314, 18.60349, 64.04713, 37.38052, 15.21893, and 40.85186, respectively, and the proximities to alphabetically ordered borders are: 0.2857, 0.19364, 0.19285, 0.21881, 0.0049, 0.31903, 0.56084, 0.15814, 0.01237, -0.16956, 0.22633, 0.017, 0.29125, 0.20796, 0.17636, 0.39028, 0.30839, 0.93074, 0.00517, 0.28081, 0.01536, 0.13064, 0.25763, 0.01059, 0.23148, 0.33039, 0.7451 and 0.33298, respectively.

If we consider two secondary components and four principal components we obtain the following 14 classes, and the 2 “ignored” classes to 16 are, with the signs from the last component (corresponding to the maximum eigenvalue) to the first principal component (corresponding to the fourth eigenvalue in decreasing order), $+--+$ and $+++-$.

If we consider one secondary component and five principal components we obtain the following 20 classes, and the 12 “ignored” classes to 32 are, with the signs from the last component to the first principal component, $----+$, $---++$, $-+---$, $-+-++$, $-++-+$, $-+++-$, $+--+-$, $+-++-$, $++---$, $+++--$ and $+++-+$. In this case, we have “ignored” classes for objective reasons (the number of points is less than the maximum number of classes: $29 < 32$), but as one may see for two secondary components, we have $14 < 16 < 29$, and we have 2 “ignored” classes. One may see that both signs codes for the two “ignored” classes ($+--+$ and $+++-$) in the case of two secondary components are each a prefix for two “ignored” classes in the case of one secondary component: between the “ignored” classes in the last case we have $+--+-$, $+--++$, $+++--$ and $+++-+$.

5. Conclusions

The method presented in this paper can be connected to the methods starting from *PCR* and k-means (Ciuiu, 2007b). It works in each model where we can use regression, or the classical Perceptron algorithm, including economic applications (Nastac *et al*., 2007).

If we put together the results from this paper and those where we use *PCR* we can conclude that the principal components group the points in the same class, and the secondary ones separate the points in different classes. The analogy is that in both papers (this and Ciuiu, 2007b), we start from known algorithms for neural networks (Perceptron, respectively k-means). The differences are that in contrast to the results obtained using regression (Ciuiu, 2007a, Ciuiu, 2007b) we can have classes with only one point, and we have not different axes for different classes: the classes depend on the signs of principal components. In fact, there exists also a common starting point for Perceptron, k-means and Bayes (Kong and Kosko, 1992): this is the discriminant surface, which is a hyper-plane in the case of Perceptron, the Euclidean distance to the gravity center of the classes multiplied by -1 in the case of k-means, and the posterior probability to have a point in a given class in the case of Bayes.

For the algorithm, we must have at least one secondary component. Of course, if we want to have only principal components, we can increase the dimension of the space by 1, and this new component is set to a constant value. The new component is for the new higher dimension space the only secondary component, and it is called "bias" in neural networks (Dumitrache *et al*., 1999, Kong and Kosko, 1992). If we set the new component, bias, as the only secondary component in the examples from the previous section, we see that we have no “ignored” class in the first example, and in the second example there are only 5 new classes. The prefix property found in the second example can be also checked in the case of setting bias as the only secondary component.

The cohesion of the class measures the power of grouping the points in the same class. If it is negative, the points in that class are closer to the gravity center (the origin in the new coordinates system) on the principal components than the variance on the secondary ones. The proximity to border is useful for comparing classifications of two sets of points with different measure orders or with different numbers of secondary components.

The separation between two classes measures the power of separating the classes. If it is negative, the classes are closer to each other than the variance on the secondary components. The proximity of separations to borders was introduced for the same reason as for cohesions.

In the case of separation, we need at least two principal components in the examples of the section 3 because if we have only one principal component we have only two classes, and the separation between them is $\mathrm{sep}(\mathbf{C}_1, \mathbf{C}_2) = S^2_{\mathrm{princ}} - S^2_{\mathrm{sec}} = S^2_{\mathrm{sep}}(\mathbf{C}_1, \mathbf{C}_2) - S^2_{\mathrm{sec}}$. In this case, it results that $|\mathrm{sep}(\mathbf{C}_1, \mathbf{C}_2)|$ does not tend to 1 if $S^2_{\mathrm{princ}} \neq 0$ and $S^2_{\mathrm{sec}} \neq 0$. If $S^2_{\mathrm{sec}} = 0$, it is obvious that $\mathrm{prox}_2(\mathbf{C}_1, \mathbf{C}_2) = 1$. In the first example in section 4 we still have the proximity to border $\mathrm{prox}_2(\mathbf{C}_1, \mathbf{C}_2) = 0.89168$, which is very close to 1 even if we have $p = 1$. This can be explained by the eigenvalues of the sample covariance matrix, which are $\lambda_1 = 0.09928$, $\lambda_2 = 0.20775$ and $\lambda_3 = 2.83457$. By computation, we can check that $\frac{\lambda_3 - \lambda_1 - \lambda_2}{\lambda_3} = 0.89168$.

It is an open problem to find theoretical examples for the maximum proximities more general than those from the third section (with more than one principal component in the case of cohesions, and more than two in the case of separations).

If we increase the number of principal components, the number of classes also increases. In this case, one may see that the number of negative cohesions decreases, and the number of proximities to borders of the cohesions between 0 and 0.1 increases. In none of the examples in section 4 we have proximities to borders of cohesions between 0.9 and 1. However, from the theoretical consideration of section 3, we can only have one such proximity to borders and the involving class has the main part of the sum of squares on principal components. In this case, the proximity between 0.9 and 1 disappears when a new principal component separates the above class.

The minimum of proximities to borders of cohesions between 0.1 and 0.9 generally decreases, but it can increase when it begins to “ignore” classes. Their maximum generally decreases, but it can increase when a negative cohesion disappears. The difference between the above maximum and minimum generally decreases, but it can also increase when a negative cohesion disappears. Another increase in this difference is when the number of classes increases from 14 to 20, but the increase in this case is only 0.0123, and it can be explained by the same number of proximities to borders of cohesions between 0.1 and 0.9: 3.

The negative separations appear only in the second example, and one may say the same thing about the proximities to borders of the separations between 0.9 and 1. The first appearances can be seen when the number of principal components is maximum, so that we have no “ignored” classes. The first number swings, and the second number turns 0 from 1 immediately. The number of the proximities to borders of the separations between 0 and 0.1 increases, the minimum of the proximities to borders of the separations between 0.1 and 0.9 generally decreases, the only increase is from 14 to 20 classes, which is very small: 0.00001. The maximum generally decreases, but it can increase when negative separations and proximity to borders between 0.9 and 1 appear. The same thing can be said about the difference between maximum and minimum.

When we set the bias as the only secondary component, all the cohesions and separations are positive. In both cases, we have no proximity to borders between 0.9 and 1, and in the first example we have no proximity to borders of separation between 0 and 0.1. In the second example, there are 222 such proximities, but in this case we have almost one point in each class: the 29 banks are in 25 classes. In the first example, one may remark the decreases in the above minimums and maximums, and the increase in the difference. The significant increase is in the case of separations. In the second example, the above maximums, minimums and differences are close to the case of two secondary components (only one without bias).

References

Ciuiu, D., (2007a), "Pattern Classification Using Polynomial and Linear Regression", *Proceedings of the International Conference Trends and Challenges in Applied Mathematics, 20-23 June 2007*, Technical University of Civil Engineering, Bucharest, Romania, pp. 153-156.

Ciuiu, D., (2007b), "Pattern Classification using Principal Components Regression", *Proceedings of the International Conference Trends and Challenges in Applied Mathematics, 20-23 June 2007*, Technical University of Civil Engineering, Bucharest, Romania, pp. 149-152.

Costinescu, C. and Ciuiu, D., (2007), "The Monte Carlo Method to Find Eigenvalues and Eigenvectors", *Proceedings of the International Conference Trends and Challenges in Applied Mathematics, 20-23 June 2007*, Technical University of Civil Engineering, Bucharest, Romania, pp. 157-160.

Dumitrache, I., Constantin, N. and Dragoigea, M., (1999), *Reţele Neurale*, Matrix Rom, Bucureşti.

Jula, D., (2003), *Introducere în econometrie*, Ed. Professional Consulting, Bucureşti.

Kong, S.G. and Kosko, B., (1992), "Differential Competitive Learning for Phoneme Recognition", *Neural Networks for Signal Processing*, Kosko, B. (Ed.), Prentice Hall, New Jersey, pp. 1-36.

Nastac, I. and Pelinescu E., Dobrescu, E., (2007), "Neuro-Adaptive Model for Financial Forecasting", *Romanian Journal of Economic Forecasting*, (3): 19-41.

Saporta, G., (1990), *Probabilités, Analyse des Donées et Statistique*, Editions Technip, Paris.

Saporta, G. and Mahjoub S., (1994), "Une méthode de discrimination non paramétrique", *Revue de Statistique Appliquée*, Vol. XLII, (2): 99-113.