ACTA UNIVERSITATIS LODZIENSIS

FOLIA OECONOMICA 68, 1987

Krzysztof Jajuga*

ON A GENERALIZATION OF THE PRINCIPAL COMPONENTS ANALYSIS

1. Introduction

Principal components analysis is one of the most frequently used methods of multivariate statistical analysis. Thanks to its simplicity and intuitiveness it is very useful in practical research.

In the course of research of complex phenomena (i.e. the ones described by a multivariate variable) it happens very often that the set of variables describing these phenomena is very numerous, and that these variables describe different, often very loosely connected fragments of a given phenomenon. In such cases there can be many difficulties in the interpretation of the principal components which are determined in such a way as to explain in the best way the variance of the variables composing the set under consideration.

An interesting procedure concerning the avoidance of these interpretational difficulties has been proposed in [5]. It is realized in two stages; in the first stage the set $\{X_1, X_2, \ldots, X_m\}$ of variables is divided into classes of similar variables (in the sense of their correlation): $C_1, C_2, \ldots, C_L$, where $1 \leq L \leq m$, in such a way that

$$C_1 \cup C_2 \cup \ldots \cup C_L = \{X_1, X_2, \ldots, X_m\}$$

$$C_i \cap C_j = \emptyset \quad i \neq j \quad i, j = 1, \ldots, L$$

$$C_i = \{X_{i_1}, X_{i_2}, \ldots, X_{i_{l_i}}\} \quad i = 1, \ldots, L$$

* Lecturer, Economic Academy in Wrocław.

where:

l_i - the number of variables belonging to the class C_i. In order to divide this set any classification method (with regard to variables) can be used.

In the second stage of the considered procedure we have to determine a certain variable for each class, and this variable is a linear combination of the variables belonging to this class, which means that for the j-th class the variable is

$$S_j = a_{1j}X_{i_1} + a_{2j}X_{i_2} + \ldots + a_{l_ij}X_{i_{l_i}} . \quad (1)$$

In determining this variable the principal components analysis is used. Thus the variable S_j is the first principal component for the set of variables belonging to the class C_j. In this sense it is the optimal variable representing this class of similar variables.

What is most important in this method is the determining of the classes of similar variables by means of classification methods. Very often, however, the classification methods have serious weaknesses. The main weakness is the fact that the variables belonging to the same class can be less similar (correlated) to one another than variables from different classes.

Here a method will be proposed that has two features: firstly it eliminates the defects on the classification in the usual sense by means of introducing its generalization, i.e. the fuzzy classification. Secondly, it introduces an optimality of selection of the linear combination representing the class of similar variables, by means of the selection of such linear combination that will satisfy the condition of the maximum correlation with the variables belonging to the initial set.

The fuzzy set (comp. [6]) is a generalization of the set in the usual sense. It is assumed that a set $\mathcal{X}$ (so called "universum") is given and fuzzy subsets are defined on the elements of $\mathcal{X}$. A fuzzy subset A of the universum $\mathcal{X}$ is defined by means of the membership function:

$$w_A : \mathcal{X} \rightarrow [0, 1].$$

The theory of fuzzy sets replaces the notion of "belonging" used with reference to sets in the usual sense and expressed by a zero-one variable, by the notion of "membership", expressed by a continuous variable taking values from the interval [0, 1].

Similarly, the fuzzy class of similar variables is a generalization of the class of similar variables. The degree of membership of variables to fuzzy classes is a number from the interval [0, 1].

The problem of the fuzzy classification of variables can be defined as follows (comp. [2], [4]).

Given the set $\{X_1, X_2, \ldots, X_m\}$ on whose elements (i.e. variables) observations have been made for n studied objects, we have to determine a family of fuzzy subsets: $C_1, C_2, \ldots, C_L$ $(1 \leq L \leq m)$ in such a way that the following conditions hold:

$$(2) \quad 0 \leq w_{ij} \leq 1 \qquad i = 1, \ldots, m \qquad j = 1, \ldots, L,$$

$$(3) \quad \sum_{j=1}^{L} w_{ij} = 1 \qquad i = 1, \ldots, m,$$

where w_{ij} denotes the degree of membership of a variable X_i to a fuzzy class C_j.

The variables whose degrees of membership to the same class are high - are the most similar, and the variables whose degrees of membership to different classes are high - are the least similar.

As one can see, in the fuzzy classification problem, the similarity of variables is defined by the correlation between them. Obviously, the sign of the correlation coefficient is not taken into account, that is, the positive and negative correlation have got the same treatment, and we are interested only in the strength of the correlation.

It may occur, that the variables are highly correlated because they are simultaneously affected by another variable. Then the proper measure of similarity (correlation) for such variables is partial correlation coefficient. In such a case, instead of correlation matrix (or covariance matrix), which is used in the considerations below, the matrix of partial correlation coefficients (or partial covariances) should be used.

2. The problem of fuzzy principal components analysis

Let $X = (X_1, X_2, \ldots, X_m)'$ be an m-dimensional vector of variables (by way of simplification we assume that these variables are standarized). Let R be a correlation matrix, being also a covariance matrix of these variables; and let L ($1 \leq L \leq m$) be a number of fuzzy classes of variables. Moreover, let $W = (w_1, w_2, \ldots, w_L)$ be the degree of membership matrix, where $w_j = (w_{1j}, w_{2j}, \ldots, w_{mj})'$, $j = 1, \ldots, L$, w_{ij} denoting the degree of membership of a variable X_i to the j-th fuzzy class of variables.

Our task is to define a certain variable for the j-th fuzzy class of variables. The variable to be determined is such linear combination $S_j = a_j'X$, where $a_j = (a_{1j}, a_{2j}, \ldots, a_{mj})'$ that the weighted sum of the squares of the correlation coefficients between this variable and the variables X_i, ($i = 1, \ldots, m$) is maximum, and the squares of the degrees of membership of the variables X_i, ($i = 1, \ldots, m$) to the j-th fuzzy class of variables are the weights. The determining of the variable S_j, ($j = 1, \ldots, L$) can also be viewed as a transformation:

$$(4) \quad X \longrightarrow X_j^*$$

where

$$(5) \quad X_j^* = \begin{bmatrix} a_j' X \\ X \end{bmatrix}$$

is an (m+1)-dimensional vector.
Hence the transformation matrix (its dimensions being $(m + 1) \times m$) for the j-th class ($j=1, \ldots, L$) is

$$(6) \quad \begin{bmatrix} a_j' \\ I \end{bmatrix}$$

In light of this the correlation matrix of the vector X calculated for the j-th class (i.e. by using the transformation matrix corresponding to the j-th class) is

$$(7)\quad R_j = \begin{bmatrix} a_j' R a_j & a_j' R \\ R a_j & R \end{bmatrix}$$

for $j=1, \ldots, L$.

Since the weighted sum of the squares of the correlation coefficients should have a maximum, the criterion function for the j-th class is

$$\sum_{i=1}^{m} w_{ij}^{2} r^2(X_i, S_j) = a_j' R W_j W_j R a_j \rightarrow \max$$

where:

$r(X_i, S_j)$ - a correlation coefficient between X_i and S_j,

W_j - a diagonal m x m matrix with the main diagonal elements $w_{1j}, w_{2j}, \ldots, w_{mj}$.

Evidently, in order to solve this maximization problem, we have to normalize the vector a_j; otherwise the maximum is reached when the components of the vector a_j approach infinity. As a normalizing condition we assume

$$(8)\quad a_j' R a_j = 1$$

This condition is derived from the fact that the main diagonal elements of the correlation matrix are equal to 1.

Thus for the j-th fuzzy class of variables ($j = 1, \ldots, L$) we search for the solution to the following mathematical programming problem:

$$(9)\quad a_j' R W_j W_j R a_j \longrightarrow \max$$

under the condition

$$(10)\quad a_j' R a_j = 1$$

Using the Lagrangian multipliers method we can formulate the following unconditional extreme problem:

(11) $\mathcal{L}_j = a_j' R W_j W_j R a_j - \lambda(a_j' R a_j - 1).$

Differentiating (11) with respect to a_j and setting it to zero we obtain

(12) $R W_j W_j R a_j - \lambda R a_j = 0.$

what means

(13) $R^{1/2}(R^{1/2} W_j W_j R^{1/2} - \lambda I) b_j = 0$

where

(14) $b_j = R^{1/2} \cdot a_j$

and

(15) $R^{1/2} = \Gamma \Lambda^{1/2} \Gamma'$

where:

$\Lambda^{1/2}$ - diagonal m x m matrix in which the main diagonal elements are the square roots of the eigenvalues of R arranged in the decreasing order,

Γ - an orthogonal m x m matrix whose columns are the eigenvectors corresponding to the eigenvalues of R arranged in the decreasing order.

Then

$$R^{1/2} R^{1/2} = \Gamma \Lambda^{1/2} \Gamma' \Gamma \Lambda^{1/2} \Gamma' = \Gamma \Lambda \Gamma' = R$$

and

(16) $R^{-1/2} = (R^{1/2})^{-1}.$

In order for a non-zero solution to the matrix equation (13) to exist, the following condition must be satisfied:

(17) $| R^{1/2} W_j^2 R^{1/2} - \lambda I | = 0.$

So λ is the eigenvalue of the matrix $R^{1/2} W_j^2 R^{1/2}$.

By virtue of (12)

(18) $R W_j W_j R a_j = \lambda R a_j$.

Premultiplying (17) by a_j' we obtain

(19) $a_j' R W_j W_j R a_j = \lambda a_j' R a_j = \lambda$.

So λ is at the same time equal to the weighted sum of the squares of the correlation coefficients between the variables X_i, $i = 1, 2, \dots, m$, and the variable S_j. In order to maximize this value we have to choose the greatest eigenvalue of the matrix $R^{1/2} W_j^2 R^{1/2}$.

Let b_j be the eigenvector corresponding to that greatest eigenvalue. Obviously,

(20) $b_j' b_j = 1$.

From (14) it follows that

(21) $a_j = R^{-1/2} b_j$.

So in order to determine the variable S_j one has to:

1) determine the eigenvalues of the matrix R as well as their corresponding eigenvectors. They will form the matrices $\Lambda^{1/2}$ and Γ;

2) determine the matrix $R^{-1/2} = (\Gamma \Lambda^{1/2} \Gamma)^{-1}$;

3) determine the greatest eigenvalue of the matrix $R^{1/2} W_j^2 R^{1/2}$ and the corresponding eigenvector, b_j;

4) determine the vector $a_j = R^{-1/2} b_j$.

The coordinates of this vector are the coefficients of the linear combination $S_j = a_j' X$.

The variables $S_1, S_2, \dots, S_L$ can be interpreted as variables representing particular fuzzy classes of variables which are components of the vector X. They will be called the fuzzy principal components.

3. On properties of fuzzy principal components

Note that the determining of the variables $S_1, S_2, \dots, S_L$ can be viewed as a transformation

(21) $X \longrightarrow S$

where $S = (S_1, S_2, \ldots, S_L)'$.

In this case the transformation matrix (its dimensions being Lxm) is

$$(22)\quad \begin{bmatrix} a_1' \\ a_2' \\ \ldots \\ a_L' \end{bmatrix}.$$

Therefore $r(S_i, S_j)$, the correlation coefficient between S_i and S_j, is equal to $a_i'Ra_j$; hence

$$r(S_i, S_j) = a_i'Ra_j = b_i' b_j.$$

Now we shall show that the problem of the fuzzy principal components analysis is a generalization of the classical principal components analysis.

The classical problem of the principal components analysis for standardized variables, i.e. when the covariance matrix is at the same time the correlation matrix, is (cf. [1]): find the linear combination $S = \alpha'X$, where α is the solution to the mathematical programming problem

$$(23)\quad \alpha'R\alpha \longrightarrow \max$$

under the condition

$$(24)\quad \alpha'\alpha = 1.$$

In light of (24):

$$(25)\quad \alpha'R^{-1/2}RR^{-1/2}\alpha = 1 \Rightarrow a'Ra = 1,$$

where $a = R^{-1/2}\alpha$.

Appylying (25) to (23) and (24) we obtain an alternative form of the mathematical programming problem:

(26) $a'R^{1/2} RR^{1/2} a = a'RRa \rightarrow \max$

under the condition

(27) $a'Ra = 1$

Therefore, in order to solve this problem according to the principal components analysis method one has to determine the eigenvalues of the matrix R, and then determine $a = R^{-1/2} \alpha$, where α is the eigenvector corresponding to the greatest eigenvalue of the matrix R.

Now let us consider the problem of the fuzzy principal components analysis, when we have only one fuzzy class of variables. Then $L = 1$ and $W_1 = I$, and the problem is

(28) $\sum_{i=1}^{m} r^2(X_i, S) = a'RRa \rightarrow \max$

under the condition

(29) $a'Ra = 1.$

Using the same methods that were used in the general L-class case (see 2) we conclude that in order to determine the vector a we have to determine the eigenvalues and eigenvectors of the matrix R and then to determine

$$a = R^{-1/2} \alpha,$$

where α is the eigenvector corresponding to the greatest eigenvalue of the matrix R.

This shows that the fuzzy principal components analysis for the case when $L = 1$ is a problem analogous to the classical principal components analysis. Therefore the fuzzy principal components analysis can be treated as a logical generalization of the principal components analysis.

Such a generalization of principal components analysis may be useful in certain problems solved by multivariate statistical methods (for example, by ordering methods), particularly, when the variables describing studied phenomenon can be, on the merits of the case, divided into several classes, which describe different aspects of the phenomenon.

In addition, different classes may contain the variables highly correlated, for which the assumption of orthogonality of components is not necessary.

4. An iterative algorithm of determining fuzzy principal components

In practical problems the degrees of membership of variables to particular classes are unknown. In such cases two alternative procedures can be proposed; firstly, we can determine the degrees of membership by means of the fuzzy classification methods (cf. [2], [3]). As it has been mentioned these methods enable us to determine a fuzzy classification of objects characterized by the values of a vector of variables. Also dual approach can be considered, i.e. the classification of variables characterized by their values observed in a set of objects (units of investigation). By applying the fuzzy classification method within the dual approach we obtain the degrees of membership of the variables to particular fuzzy classes.

An alternative method will be proposed here. In this algorithm the degrees of membership and the fuzzy principal components are determined simultaneously. The algorithm proceeds as follows: Let R be the correlation matrix of the vector of standardized variables $X=(X_1, X_2, \ldots, X_m)'$; and let L be the number of the fuzzy classes. The initial values of the degrees of membership matrix will be $\bar{W}^o = (w_{lj}^o)$, $l = 1, \ldots, m$; $j = 1, \ldots, L$; where

$0 < w_{lj}^o < 1$, $\sum_{j=1}^{L} w_{lj}^o = 1$, $l = 1, \ldots, m$ and w_{lj}^o denoting the initial value of the degree of membership of the variable X_l to the j-th fuzzy class.

First of all one has to determine eigenvalues and eigenvectors of the matrix R, and then determine the matrix $R^{-1/2} = (\Gamma\Lambda^{1/2}\Gamma')^{-1}$.

Then the following iteration process is applied. In the i-th iteration we determine for the j-th class ($j=1, \ldots, L$):

1) the greatest eigenvalue of the matrix $R^{1/2} W_j^{i-1} W_j^{i-1} R^{1/2}$, the corresponding eigenvector b_j^i, where W_j^{i-1} is a diagonal $m \times m$ matrix, the main diagonal elements being $w_{1j}^{i-1}, w_{1j}^{i-1}, \ldots, w_{mj}^{i-1}$ (these are diagonal entries of the matrix $\overline{W}^{\,i-1}$);

2) vectors $a_j^i = R^{-1/2} b_j^i$;

3) correlation coefficients between the variables $S_j^i = a_j^{i'} X$, $j = 1, \ldots, L$ and variables X_l ($l = 1, \ldots, m$), according to the formula: $r(X_l, S_j^i) = r_{lj}^i$, where r_{lj}^i is the l-th coordinate of the vector $a_j^{i'} R$;

4) new values of the degrees of membership (they will form the matrix $\overline{W}^i$) according to the formula:

$$w_{lj}^i = \frac{(r_{lj}^i)^2}{\left(\sum_{t=1}^{L} (r_{lt}^i)^2\right)} \tag{30}$$

for $l = 1, \ldots, m$.

The iterative procedure is continued until the values of the degrees of membership cease to change in a significant degree, e.g. in the r-th stage, when

$$\max_{l,j} |w_{lj}^{r+1} - w_{lj}^r| < \varepsilon \tag{31}$$

where ε is a small postive number.

Then the fuzzy principal components are the variables $S_1^r, S_2^r, \ldots, S_L^r$. In our numerical examples the algorithm was convergent, and it was resistant to the assumed initial values of the degrees of membership included in the matrix $\overline{W}^o$.

5. Examples

In the examples presented below we give the determined values of the fuzzy principal components depending on different forms of the matrix R.

- L = 2, and the matrix $\mathbb{R}$ is

$$\begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

a=0

$S_1 = X_1$

$S_2 = X_2$

$a = \pm 0.1$

$S_1 = 0.99X_1 \pm 0.05X_2$

$S_2 = \pm 0.05X_1 + 0.99X_2$

$a = \pm 0.2$

$S_1 = 0.98X_1 \pm 0.10X_2$

$S_2 = \pm 0.10X_1 + 0.98X_2$

$a = \pm 0.3$

$S_1 = 0.94X_1 \pm 0.16X_2$

$S_2 = \pm 0.16X_1 + 0.94X_2$

$a = \pm 0.4$

$S_1 = 0.88X_1 \pm 0.23X_2$

$S_2 = \pm 0.23X_1 + 0.88X_2$

$a = \pm 0.5$

$S_1 = 0.75X_1 \pm 0.39X_2$

$S_2 = \pm 0.39X_1 + 0.75X_2$

$a = \pm 0.6$

$S_1 = 0.56X_1 \pm 0.55X_2$

$S_2 = \pm 0.55X_1 + 0.56X_2$

$a = \pm 0.7$

$S_1 = 0.54X_1 \pm 0.54X_2$

$S_2 = \pm 0.54X_1 + 0.54X_2$

$a = \pm 0.8$

$S_1 = 0.53X_1 \pm 0.53X_2$

$S_2 = \pm 0.53X_1 + 0.53X_2$

$a = \pm 0.9$

$S_1 = 0.51X_1 \pm 0.51X_2$

$S_2 = \pm 0.51X_1 + 0.51X_2$

- the matrix R is

$$\begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$$

$L = 2$

$S_1 = S_2 = 0.41X_1 + 0.41X_2 + 0.41X_3$

$L = 3$

$S_1 = S_2 = S_3 = 0.41X_1 + 0.41X_2 + 0.41X_3$

- the matrix R is

$$\begin{bmatrix} 1 & 0.1 & 0.2 \\ 0.1 & 1 & 0.1 \\ 0.2 & 0.1 & 1 \end{bmatrix}$$

So the form of the matrix suggests the occurence of three fuzzy classes of variables.

$L = 2$

$S_1 = 0.64X_1 + 0.06X_2 + 0.64X_3$

$S_2 = 0.05X_1 + 0.99X_2 + 0.05X_3$

$L = 3$

$S_1 = 0.97X_1 + 0.05X_2 + 0.10X_3$

$S_2 = 0.05X_1 + 0.99X_2 + 0.05X_3$

$S_3 = 0.10X_1 + 0.05X_2 + 0.97X_3$

- the matrix R is

$$\begin{bmatrix} 1 & 0.9 & 0.1 \\ 0.9 & 1 & 0.2 \\ 0.1 & 0.2 & 1 \end{bmatrix}$$

So the form of the matrix suggests the occurence of two fuzzy classes of variables

$L = 2$

$S_1 = 0.03X_1 + 0.10X_2 + 0.97X_3$

$S_2 = 0.52X_1 + 0.50X_2 + 0.05X_3$

$L = 3$

$S_1 = S_2 = 0.51X_1 + 0.50X_2 + 0.05X_3$

$S_2 = 0.03X_1 + 0.10X_2 + 0.97X_3$

- the matrix R is

$$\begin{bmatrix} 1 & 0.9 & 0.8 \\ 0.9 & 1 & 0.9 \\ 0.8 & 0.9 & 1 \end{bmatrix}$$

in this case the form of the matrix suggests the occurence of one class.

$L = 2$

$S_1 = S_2 = 0.35X_1 + 0.36X_2 + 0.35X_3$

$L = 3$

$S_1 = S_2 = S_3 = 0.34X_1 + 0.36X_2 + 0.34X_3$

- the matrix R is

$$\begin{bmatrix} 1 & 0.9 & 0.1 & 0.1 \\ 0.9 & 1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 1 & 0.9 \\ 0.1 & 0.1 & 0.9 & 1 \end{bmatrix}$$

The form of the matrix suggests the occurence of two fuzzy classes.

$L = 2$

$S_1 = 0.03X_1 + 0.03X_2 + 0.51X_3 + 0.51X_4$

$S_2 = 0.51X_1 + 0.51X_2 + 0.03X_3 + 0.03X_4$

$L = 3$

$S_1 = 0.05X_1 + 0.05X_2 + 0.05X_3 + 0.98X_4$

$S_2 = 0.51X_1 + 0.51X_2 + 0.03X_3 + 0.04X_4$

$S_3 = 0.05X_1 + 0.05X_2 + 0.98X_3 + 0.05X_4$

- the matrix R is

$$\begin{bmatrix} 1 & 0.9 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 & 0.9 \\ 0.9 & 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 0.9 & 1 \end{bmatrix}$$

Here the form of the matrix suggests the occurence of one fuzzy class.

$L = 2$

$S_1 = S_2 = 0.26X_1 + 0.26X_2 + 0.26X_3 + 0.26X_4$

$L = 3$

$S_1 = S_2 = S_3 = 0.26X_1 + 0.26X_2 + 0.26X_3 + 0.26X_4$

$L = 4$

$S_1 = S_2 = S_3 = S_4 = 0.26X_1 + 0.26X_2 + 0.26X_3 + 0.26X_4$

– the matrix R is

$$\begin{bmatrix} 1 & 0.9 & 0.1 & 0.1 & 0.2 \\ 0.9 & 1 & 0.1 & 0.1 & 0.2 \\ 0.1 & 0.1 & 1 & 0.9 & 0.1 \\ 0.1 & 0.1 & 0.9 & 1 & 0.1 \\ 0.2 & 0.2 & 0.1 & 0.1 & 1 \end{bmatrix}$$

$L = 2$

$S_1 = 0.49X_1 + 0.49X_2 + 0.03X_3 + 0.03X_4 + 0.15X_5$

$S_2 = 0.03X_1 + 0.03X_2 + 0.51X_3 + 0.51X_4 + 0.03X_5$

$L = 3$

$S_1 = 0.50X_1 + 0.50X_2 + 0.03X_3 + 0.03X_4 + 0.07X_5$

$S_2 = 0.08X_1 + 0.08X_2 + 0.04X_3 + 0.04X_4 + 0.94X_5$

$S_3 = 0.03X_1 + 0.03X_2 + 0.51X_3 + 0.51X_4 + 0.03X_5$

$L = 4$

$S_1 = 0.09X_1 + 0.09X_2 + 0.04X_3 + 0.04X_4 + 0.94X_5$

$S_2 = 0.50X_1 + 0.50X_2 + 0.03X_3 + 0.03X_4 + 0.06X_5$

$S_3 = S_4 = 0.03X_1 + 0.03X_2 + 0.51X_3 + 0.51X_4 + 0.03X_5$

$L = 5$

$S_1 = S_2 = 0.50X_1 + 0.50X_2 + 0.03X_3 + 0.03X_4 + 0.06X_5$

$S_3 = S_4 = 0.03X_1 + 0.03X_2 + 0.51X_3 + 0.51X_4 + 0.03X_5$

$S_5 = 0.08X_1 + 0.08X_2 + 0.04X_3 + 0.04X_4 + 0.94X_5$

in this case the form of the matrix suggests the occurence of three fuzzy classes.

As we can see, in the majority of cases the algorithm allows us to find out the proper number of fuzzy classes,and this number is equal to the number of various fuzzy principal components, regardless of the number L assumed in the algorithm.

In the above examples we can notice quite good "adjustment" of the coefficients in the linear combinations to the similarities between the variables represented by the elements of the correlation matrix.

Bibliography

[1] A n d e r s o n T. W. (1958): Introduction to Multivariate Statistical Analysis, New York, Wiley.

[2] B e z d e k J. C. (1974): Numerical Taxonomy with Fuzzy Sets, "Journal of Mathematical Biology", 1, p. 57-71.

[3] J a j u g a K. (1981): Metody analizy wielowymiarowej w ilościowych badaniach przestrzennych, Akademia Ekonomiczna, Wrocław.

[4] J a j u g a K., Zbiory rozmyte w zagadnieniu klasyfikacji, "Przegląd Statystyczny", in press.

[5] P l u t a W. (1979): Metody wielowymiarowej analizy porównawczej w modelowaniu informacji ekonomicznej kombinatu przemysłowego, "Prace Naukowe Akademii Ekonomicznej", 156, Wrocław.

[6] Z a d e h L. A. (1965); Fuzzy Sets, "Information and Control", 8 p. 338-353.

Krzysztof Jajuga

O UOGÓLNIENIU ANALIZY GŁÓWNYCH SKŁADOWYCH

Artykuł przedstawia pewne uogólnienia analizy głównych składowych. Idea polega na tym, że zbiór zmiennych jest zastąpiony przez rodzinę podzbiorów rozmytych określonych ze względu na te zmienne, z różnymi stopniami przynależności poszczególnych zmiennych do tych klas. Wyznaczone rozmyte główne składowe są optymalnymi reprezentami szczególnych klas rozmytych.

Jest również prezentowany iteracyjny algorytm otrzymywania rozmytych głównych składowych. Rozważania są ilustrowane prostymi przykładami.