JOURNAL OF APPLIED INTELLIGENT SYSTEMS & INFORMATION SCIENCES
VOL. 3, ISSUE 1, PP. 46-62, OCTOBER 2022.
Available at: http://journal.research.fanap.com/
DOI: https://doi.org/10.22034/jaisis.2022.359381.1047

DECISION-MAKING BASED ON INTERRELATED CRITERIA APPLYING AN ADJUSTED FACTOR ANALYSIS APPROACH: FUNDAMENTAL ANALYSIS OF STOCKS

Maghsoud Amiri [1], Siavash Hekmat [2, *], Nematollah Saeidi [3]

[1] *Department of Industrial Management, Faculty of Management and Accounting, Allameh Tabataba'i University, Tehran, Iran*
[2] *Faculty of Industrial and Mechanical Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*
[3] *Department of Artificial Intelligence Engineering, University of Isfahan, Isfahan, Iran*

ABSTRACT

In a typical multi-attribute decision-making (MADM) problem, different alternatives are presented for evaluation depending on multiple criteria. By default, independence or slight interrelation of criteria is an essential prerequisite in most existing MADM techniques to generate appropriate and non-overrated discrimination scores. This research, applying a tailored version of the factor analysis (FA) method, prepares an integrated algorithm for empowering MADM techniques to deal with the kinds of criteria carrying severe interrelation. Accordingly, guidelines for adjusting FA are proposed here to simultaneously eliminate the criteria interrelation and decrease the data volume, so that only the main aspects of data are taken into consideration for decision-making. In the end, the practical case of financial discrimination is investigated for companies listed in the stock exchange applying the proposed algorithm, and the results are validated using ELECTRE and VIKOR techniques. Furthermore, the shortcomings of conventional adjustments for FA are explored through the case study. The proposed approach is also applicable for evaluating alternatives in portfolio management, supply chain management, credit scoring, ranking, etc. It is also helpful in boosting machine learning algorithms and digitization of sectors such as healthcare, manufacturing, marketing, IoT processing, and recommendation systems.

KEYWORDS: Multi-attribute decision-making; Principal component analysis; Factor analysis; Financial management; Fundamental analysis.

1. INTRODUCTION

In a typical multi-attribute decision-making (MADM) problem, several alternatives are presented for evaluation from the perspective of some common criteria. Data acquired through these evaluations form a decision matrix. The purpose of MADM is to calculate the synthetic utility value of existing alternatives using the decision matrix to rank them. Most of the existing MADM techniques could produce appropriate and non-overrated synthetic utility values only when the interrelation of criteria is inconsiderable (Hatami-Marbini et al. 2020; Huang et al., 2021; Hekmat et al. 2021). SAW, AHP, TOPSIS, and VIKOR are among these techniques their additive nature could result in overrated decision assessments. In cases where the criteria are not necessarily

[*] *Corresponding Author, Email: siavashhekmat@gmail.com*

RECEIVED: 27 AUGUST, 2022; **ACCEPTED**: 26 OCTOBER, 2022; **PUBLISHED ONLINE**: 31 OCTOBER, 2022

mutually independent, if an additive aggregating method is used to derive the synthetic utility value, which is the same as the traditional assumption for the independent relationship among criteria, the result would become an overestimate or underestimate in different situations (Chiou et al. 2005; Kuo and Liang 2011). The existence and severity of such interrelation could be discovered using statistical tests based on the statistical correlation of any pair of criteria in the decision matrix. Severe correlations between the criteria may result in synthetic utility values with exaggerated and not properly discriminated values and the consequence is doubtfulness and deviation of the decision maker. This problem makes the decision maker reduce or eliminate the interrelation between criteria as a solution. In addition, the independence of criteria is a useful characteristic by itself, because it denotes that different aspects of data are considered and measured by the criteria (Manly, 2004).

Principal component analysis (PCA) and factor analysis (FA) are appropriate mathematical tools applied to eliminate or reduce the interrelation between criteria or to describe the initial variables according to a fewer number of factors so that a simpler model is found out of the initial model. Thus, the synthesis of these techniques with those of MADM could establish a potent tool for decision-making. However, the output of PCA and FA techniques accompanied by MADM methods, are not favorable all the time and in case of not adjusting them properly, unexpected, and far-fetched conclusions are obtained. Investigating this mismatch and solving it is the main purpose of this research.

Numerous applications of PCA accompanying multi-criteria decision-making (MCDM) methods exist in the literature. For example, Tung and Lee (2009) developed a grey approach to PCA which replaced the correlation matrix with another matrix called GADI. The interrelation of the decision criteria is calculated through a geometrical method in this matrix instead of the usual statistical correlation. Thus, the constraints related to the necessity for the excessive volume of data and conforming to normal distribution are eliminated. Tung and Lee (2010) developed this method for FA and applied it to evaluate the performance of Taiwanese biotechnological companies. Along with the former researchers, Mahbub et al. (2011) used the techniques of PCA and PROMETHEE to explore the process of the build-up of semi and non-volatile organic compounds as the environmental pollutants of the urban roads. In another study, Hatami-Marbini et al. (2020) performed an evaluation process using a multi-attribute efficiency analysis model and a multivariate statistical method, a so-called PCA-Data-Envelopment-Analysis (PCA-DEA) method, to support supplier relationship management under uncertainty. Hekmat et al. (2021) conducted another related study using PCA-DEA for decision-making. This method, classified as a semi-DEA model based on multivariate statistical ranking, provides a complete ranking of suppliers and deals with too many inputs and outputs existing in interrelated datasets. Recently, the model of Stević et al. (2022) is developed based on the integration of DEA, PCA, CRITIC (Criteria Importance Through Inter Criteria Correlation), Entropy, and MARCOS (Measurement Alternatives and Ranking according to the Compromise Solution) methods. This study is conducted for determining the final efficiency of transportation companies based on ten input-output parameters. Also, Dugger et al. (2022) outlined a study to evaluate the effectiveness of PCA as an objective weight assignment method to establish the rank order of pilots in aviation communities. This study intended to reduce subjectivity in the group-based MCDM pilot selection process. In another study in line with the present research, Heydari et al. (2022) considered the multi-criteria global financial development-ranking problem for some Middle Eastern countries. They proposed a solution methodology based on weighted PCA and TOPSIS while considering both linear and nonlinear data relationships.

As an implementation case of PCA in terms of FA, Chiou et al. (2005) used FA to discover the interrelated criteria and classify them as common factors in a MADM problem. This research applied the non-additive fuzzy integral to aggregate the synthetic utility values of the alternatives within the factors. Ultimately, the simple additive weighting (SAW) method is used to gain the final discrimination scores of alternatives. FA, DEMATEL, fuzzy integral, and AHP tools were used together by Tzeng et al. (2007) to prepare a comprehensive evaluation model of electronic training programs. In their model, it was possible to consider the interrelations of criteria and fuzziness of experts' objective judgments to assess the electronic training programs. Mousavi et al. (2009) used FA to classify 10 important criteria as 4 distinct factors and ranked five maintenance strategies by transforming the problem into a multilevel AHP and using the fuzzy TOPSIS method. Huang et al. (2021) developed a DEA-integrated grey factor analysis approach for efficiency evaluation and ranking in uncertain systems. Chen et al. (2021) conducted another related study proposing a random intuitionistic fuzzy FA model to address the problem of complex multi-attribute large group decision-making from three perspectives. This method effectively reduces data dimensionality and considers the underlying random environmental factors that can affect the performance of the alternatives.

Decision-making has always been a problematic issue in exchange marketplaces of bourse for trading stocks, bonds, commodities, and futures. Analytics in this field is considered from two viewpoints: fundamental and technical. The fundamental analysis investigates a company's finances, internal operations, industrial market conditions, and domestic and international policies. Meanwhile, technical analysis assumes that stocks follow certain trends in moving to an equilibrium point. Hence, past prices and volume changes are considered in technical analysis to predict future price trends (Cheng et al., 2021). According to this mindset, the present study intends to perform its decision-making assessments in the stock exchange based on fundamental analysis concepts. The reason behind this is that the long-term profitability of listed companies rather than short or medium-term prediction measures of liquidity or price trends is considered here.

Next, it is intended to discuss a brief review of the PCA method in section 2. In section 3, first, the exploratory FA method is illustrated as the structural basis of the model. Then, the integration issue of the existing FA method into the MADM approach is analyzed in part 3.2. Also, the development of the method and necessary adjustments are reasoned in this part. The last part of section 3 is dedicated to illustrating the step-by-step algorithm of the approach. The discriminative study and fundamental analysis of 41 listed companies in the Tehran stock exchange are explained in part 4.1. The other part of section 4 tries to verify the proposed approach in a more explicit case using the results of the former part.

2. Principal Component Analysis

PCA technique was first introduced by Karl Pearson (1901) and its practical calculations in the presence of two or more variables were illustrated by Hotelling (1933 cited Manly 2004). PCA is a statistical and multivariate approach for data reduction and is applied to gain a smaller set of variables explaining a significant proportion of the variance of initial data (Bolch and Huang, 1974). The influence of PCA as a pioneer dimensionality reduction technique is evident in making predictions and uncovering patterns of big data in terms of machine learning (Reddy et al., 2020). PCA is also a favorable technique for ranking in multivariate analysis (Slottje et al., 1991 cited Zhu, 1998). PCA method tries to find compositions of p variables (criteria) of $\mathbf{X}_1$, $\mathbf{X}_2$, …, $\mathbf{X}_p$ to form p independent variables of $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_p$. These independent variables are known as principal components of initial variables and can explain the whole amount of the variance of initial variables. In other words, principal components are some independent criteria that could be substituted for initial severe interrelated criteria. The lack of interrelation between the criteria is a useful characteristic because it denotes that the criteria are measuring different aspects of data. PCA starts with p variables as criteria for n alternatives. The 1st principal component is a composition of the variables $\mathbf{X}_1$, $\mathbf{X}_2$, …, $\mathbf{X}_p$ such that: $\mathbf{Z}_1=v_{11}\mathbf{X}_1+v_{12}\mathbf{X}_2+\ldots+v_{1p}\mathbf{X}_p$, with the conditions of:

i. $v_{11}^2+v_{12}^2+\cdots+v_{1p}^2=1$

ii. Variance of $\mathbf{Z}_1$ is maximized.

Note that each variable $\mathbf{X}_j$ (j= 1, …, p) is a column vector of n elements denoting the scores of n alternatives for the jth criteria. The 2nd principal component is calculated as $\mathbf{Z}_2=v_{21}\mathbf{X}_1+v_{22}\mathbf{X}_2+\ldots+v_{2p}\mathbf{X}_p$, in order that:

i. $v_{21}^2+v_{22}^2+\cdots+v_{2p}^2=1$

ii. Setting $\mathbf{Z}_1$ aside, the variance of $\mathbf{Z}_2$ is the maximum accessible amount explaining the remaining variance of data.

iii. $\mathbf{Z}_2$ is independent of $\mathbf{Z}_1$.

Just like in the previous step, the 3rd principal component is described, considering $\mathbf{Z}_3$ independent of the previous components. Thus, p independent principal components are obtained out of the initial variables explaining different aspects of data, with the feature:

$$\mathrm{Var}(\mathbf{Z}_1) \geq \mathrm{Var}(\mathbf{Z}_2) \geq \cdots \geq \mathrm{Var}(\mathbf{Z}_p) \quad (1)$$

A more practical procedure to construct the principal components is explained in the literature. This procedure is initiated by replacing each variable with its standardized format (by subtracting the variable mean value and dividing the result by its standard deviation). Then, the correlation matrix is calculated for the variables matrix, whose columns stand for the variables. The correlation matrix of an $n \times p$ matrix is a symmetric $p \times p$ matrix, whose element of position (i,j) stands for the statistical correlation coefficient between columns i and j $(i,j = 1,\ldots,p)$. Finally, the eigenvectors of the correlation matrix are used instead of the mentioned coefficients of $\mathbf{V}_i = (v_{i1}, v_{i2}, \ldots, v_{ip})$ to construct the principal components (Manly, 2004). The nonzero column vector V of dimension p is identified as an eigenvector of matrix $\mathbf{C}_{p \times p}$, if a nonzero scalar of λ exists such that: $\mathbf{CV}=\lambda\mathbf{V}$. Also, λ is called an eigenvalue of $\mathbf{C}$ corresponding to $\mathbf{V}$.

3. Proposed Approach

3.1. *Structural basis of the model*

According to the PCA and factor analysis literature indicated in Section 1 and the beginning of Section 2, initial data could be demonstrated in a fewer number of classes to form a simpler model of the initial one. It means that there exist various techniques to abbreviate p initial variables in form of m ($m \leq p$) common factors; this is the foundation of FA. FA is a method of dimension reduction in multivariate statistics, which is applied to extract latent variables among manifest variables, and PCA is one of the main techniques to accomplish it (Chiou et al. 2005). Also, there are other approaches such as a *maximum likelihood* or *Little–Jiffy* to accomplish the FA.

As it was mentioned before, this research tries to employ FA to propose a practical method of data reduction for the elimination or reduction of criteria interrelation in MADM problems. This helps the decision maker to obtain more differentiated and realistic responses in case of complex and interrelated situations. Next, the adjusted procedure of data reduction for MADM applications is presented, which is integrated into FA using PCA (for a detailed illustration of the general FA method see (Manly 2004).

First of all, to construct a uniform procedure, the data should be classified in a decision matrix with n rows of alternatives evaluated on p columns of criteria. These criteria should all have beneficial nature, i.e., the bigger the better. In order to accomplish this purpose, the columns of cost criteria (the smaller the better) must be subtracted from their maximum value. Thereafter, the columns of the decision matrix should be standardized by subtracting their mean value and dividing the result by the standard deviation of each column. Next, to exert the influence of different criteria levels of importance, each column should be multiplied by the corresponding criterion weight. Just as mentioned before, the columns of the non-interrelated matrix are calculated using the standardized and weighted matrix as below:

$$\mathbf{Z}_i = v_{i1}\mathbf{X}_1 + v_{i2}\mathbf{X}_2 + \cdots + v_{ip}\mathbf{X}_p, i = 1,\ldots, p \qquad (2)$$

The vectors $\mathbf{V}_i = (v_{i1}, v_{i2}, \ldots, v_{ip})$ stand for the eigenvectors calculated for the covariance matrix of the standardized and weighted decision matrix. The calculation of the covariance matrix is similar to the correlation matrix, which was explained before. The eigenvectors calculated here must satisfy the properties of being united and orthogonal. $\mathbf{V}_i$ is a unit vector when $v_{i1}^2+v_{i2}^2+\ldots+v_{ip}^2=1$. $\mathbf{V}_i$ and $\mathbf{V}_j$ are orthogonal vectors concerning each other when $v_{i1}v_{j1}+v_{i2}v_{j2}+\ldots+v_{ip}v_{jp}=1$ $(i, j= 1, \ldots, p)$. The point here is that each principal component is uniquely constructed using one of the eigenvectors. In addition, if $\mathbf{Z}_i$ is constructed using $\mathbf{V}_i$, then the variance explained by it would be equal to the corresponding eigenvalue, λ_i. Therefore, the eigenvectors have to be sorted by descending order of eigenvalues so that the property of Eq. (1) is held. Note that the eigenvalues here are all positive real scalars.

Here, it is important to say that the covariance matrix is substituted for the correlation matrix in the context of general PCA, to activate the influence of criteria weights. It means that the use of a correlation matrix here would have neutralized the effect of diverse criteria weights considering the criteria equal. In addition, note that the covariance matrix of the standardized but not weighted decision matrix is equal to the correlation matrix.

The mentioned reforming of variables (**X**s) to principal components (**Z**s) has an orthogonal mode, such that the reverse equalities are held as:

$$\mathbf{X}_i = v_{1i}\mathbf{Z}_1 + v_{2i}\mathbf{Z}_2 + \cdots + v_{pi}\mathbf{Z}_p, i = 1, \ldots, p \qquad (3)$$

In decomposition to factors, just the principal components with their accumulative variance amounts exceeding 90 percent of the whole variance of data are kept. These components are selected among the ones with greater amounts of variances. Consider Eq. (4) to obtain the weight of principal component *j*:

$$w'_j = \lambda_j \Big/ \sum_{k=1}^{p} \lambda_k, j = 1, \ldots, p \qquad (4)$$

Thus, only the first *m* out of *p* principal components are chosen with their aggregate weight value reaching 0.90. Thus, Eq. (3) is reformed as:

$$\mathbf{X}_i = v_{1i}\mathbf{Z}_1 + v_{2i}\mathbf{Z}_2 + \cdots + v_{mi}\mathbf{Z}_m + \mathbf{E}_i, i = 1, \ldots, p \qquad (5)$$

Here, the vector $\mathbf{E}_i$ stands for the residual known as the special factor. The criterion mentioned here is based on achieving a specified cumulative percentage of the total variance which has been defined by successive factors. This amount is not constant everywhere and aims to ensure that the extracted factors can explain at least a specified amount of variance. Practically, to be satisfactory, the total amount of variance defined by factors should be at least 95 percent in the natural sciences, and 60 percent in the social sciences. However, no absolute threshold has been adopted for all applications (Hair et al. 1998).

There is also another criterion in which only the factors corresponding to eigenvalues greater than 1 are considered significant. But this criterion is only applicable when the correlation matrix is used instead of the covariance matrix, or when the covariance matrix is calculated upon the standardized and not weighted decision matrix, which is equal to the correlation matrix in this case. It should be noted that the first criterion is defined in the mentioned way because in the case of using the covariance matrix some eigenvalues might become so large that the ones neighboring 1 are considered insignificant.

Now, the next step is to rescale the principal components $\mathbf{Z}_1$, $\mathbf{Z}_2$, …, $\mathbf{Z}_m$ to have variances equal to 1, so that appropriate factors are made. For this rescaling, $\mathbf{Z}_j$ (j= 1, …, m) must be divided by its standard deviation, $\sqrt{\lambda_j}$. Considering the substitution of $\mathbf{Z}_j \big/ \sqrt{\lambda_j} = \mathbf{F}_j$, Eq. (5) is revised as:

$$\mathbf{X}_i = \sqrt{\lambda_1} v_{1i}\mathbf{F}_1 + \sqrt{\lambda_2} v_{2i}\mathbf{F}_2 + \cdots + \sqrt{\lambda_m} v_{mi}\mathbf{F}_m + \mathbf{E}_i, i = 1, \ldots, p \qquad (6)$$

Assuming $\sqrt{\lambda_j} v_{ji} = a_{ij}$, the latter equation may be displayed as:

$$\mathbf{X}_i = a_{i1}\mathbf{F}_1 + a_{i2}\mathbf{F}_2 + \cdots + a_{im}\mathbf{F}_m + \mathbf{E}_i, i = 1, \ldots, p \; (7)$$

As a complementary explanation, it should be indicated that:

$$\lambda_j = a_{1j}^2 + a_{2j}^2 + \cdots + a_{pj}^2, j = 1, \ldots, m \qquad (8)$$

So far, data was abbreviated in the form of *m* factors. Most of the time, it is necessary to extract the factors in a way that the absolute statistical covariance coefficient between each initial variable and only a small number of factors (preferably only one factor) is maximized, and the rest are minimized. This can be conducted through a rotation of factors. Considering that the a_{ij} loading is equal to the statistical covariance coefficient of i^{th} initial criteria and j^{th} factor ($\mathbf{X}_i$ and $\mathbf{F}_j$), rotation of factors intends to minimize the absolute value of all but just one (or a small) number of a_{ij}s for each *i* (in case of using the correlation matrix instead of the covariance matrix, a_{ij} would be the statistical correlation coefficient of $\mathbf{X}_i$ and $\mathbf{F}_j$). Thus, each initial variable is correlated severely with only one (or a small) number of factors. As a result, each initial variable is never represented by more than one factor (aspect), so the reduction of data is accomplished properly.

The rotation of factors could be orthogonal or oblique. Orthogonal rotation does not allow the factors to be interrelated, just as the original factors do. In contrast, rotated factors using oblique rotation are interrelated to some extent and this makes the oblique rotation inappropriate for the current approach. However, loadings are more polarized in oblique rotation and $\mathbf{X}_i$s are related to $\mathbf{F}_j$s more distinctly. As an orthogonal rotation, Varimax rotation is widely used as a standard method. This method was first proposed by H.F. Kaiser and later was modified by normalizing to reach presumably better results. So, Varimax rotation could be applied using Kaiser's normalization method or without it.

Now, if the loadings are rotated using one of the existing rotation techniques (such as Varimax loading), then the model would turn to:

$$\mathbf{X}_i = g_{i1}\mathbf{F}_1^* + g_{i2}\mathbf{F}_2^* + \cdots + g_{im}\mathbf{F}_m^* + \mathbf{E}_i, i = 1, \ldots, p \qquad (9)$$

Which $\mathbf{F}_i^*$ stands for the i^{th} factor after rotation and g_{ij} denote the rotated loading expressing the covariance coefficient of $\mathbf{X}_i$ and $\mathbf{F}_j^*$. It is useful to mention that the rotation of factors also transforms the corresponding eigenvalues. Considering Eq. (8), the eigenvalues are substituted as:

$$\lambda_j = g_{1j}^2 + g_{2j}^2 + \cdots + g_{pj}^2, j = 1, \ldots, m \qquad (10)$$

Because of complex and excessive amounts of calculations, it is better to utilize computational software programs like MATLAB or PSAW (SPSS) to accomplish the rotation of loadings. To estimate the factor scores after rotation, the following equation could be applied:

$$\mathbf{F}^* = \mathbf{XG}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1} \qquad (11)$$

Where $\mathbf{F}_{n\times m}^* = [\mathbf{F}_1^*, \mathbf{F}_2^*, \ldots, \mathbf{F}_m^*]$, $\mathbf{X}_{n\times p} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p]$, and $\mathbf{G}$ is the $p\times m$ matrix of rotated factor loadings. This equation estimates the factor scores as linear compositions of initial variables. For these factors, the weight values are calculated by Eq. (12) to be used for MADM applications:

$$w_j'' = \lambda_j \Big/ \sum_{k=1}^{m} \lambda_k, j = 1, \ldots, m \qquad (12)$$

3.2. Analysis and development

Here, some necessary statements would be discussed to analyze and adjust the vulnerable points of the mentioned structure.

Proposition 1. If $\mathbf{V}_i$ is an eigenvector of matrix $\mathbf{C}$ and λ_i is the corresponding eigenvalue, then $-\mathbf{V}_i$ is an eigenvector too ($i = 1, \ldots, p$).

Proof. If $\mathbf{CV}_i = \lambda_i\mathbf{V}_i$, then $\mathbf{C}(-\mathbf{V}_i) = \lambda_i(-\mathbf{V}_i)$. □

Proposition 2. If $\mathbf{V}_i$ holds the property of unity, then $-\mathbf{V}_i$ does too (i =1, …, p).

Proof. If $\mathbf{V}_i^2=v_{i1}^2+v_{i2}^2+\ldots+v_{ip}^2=1$, then $(-\mathbf{V}_i)^2= (-v_{i1})^2+(-v_{i2})^2+\ldots+(-v_{ip})^2=1$. □

Proposition 3. If $\mathbf{V}_i$ and $\mathbf{V}_j$ are orthogonal vectors, then $-\mathbf{V}_i$ and $\mathbf{V}_j$ are so too (i, j =1, …, p).

Proof. If $\mathbf{V}_i.\mathbf{V}_j=v_{i1}v_{j1}+v_{i2}v_{j2}+\ldots+v_{ip}v_{jp}=0$, then $(-\mathbf{V}_i).\mathbf{V}_j= (-v_{i1})v_{j1}+(-v_{i2})v_{j2}+\ldots+(-v_{ip})v_{jp}=0$. □

In conclusion, the additive inverse vector of each eigenvector is an eigenvector by itself, holding the properties of unity and orthogonality with respect to the rest, and thus could be substituted for the original vector. The intrinsic problem with the identified FA method is expressed in this way: if the eigenvector $\mathbf{V}_j$ is replaced by its additive inverse, then the respective principal component $\mathbf{Z}_j$, becomes additively inversed too (multiplied by –1). Then, if this principal component is intended to be transformed as a factor, again the extracted factor and even its rotated vector become additively inversed, and this leads to a quite contradictory factor. The main point here is to select the correct eigenvectors so that the extracted factors become reasonable. Considering the point that the first step of PCA is spent to transform the cost criteria into benefit criteria, it is expected that the determined principal components and also the factors using benefit criteria possess the property of bigger the better. This property may be inherited by the existing positive covariance coefficients between each factor and the main criteria from which it is generated. It was already mentioned that the covariance coefficient between the j^{th} factor and the i^{th} criterion is equal to a_{ij} loading in Eq. (7) for on-rotated factors and g_{ij} loading in Eq. (9) for rotated factors. Thus, for each j if most of the g_{ij} loadings with large absolute values gain positive amounts, the property of having a benefit nature is transferred to the rotated factor j. In case of no factor rotation, this condition could be explored by considering amounts of g_{ij} coefficients to be the same as those of a_{ij}. Thus, the condition could be satisfied by one of the following criteria:

i. Making the sum of g_{ij} coefficients for each j positive. In other words, the sum of the columns of matrix **G** should be positive values. It is concluded that between an eigenvector and its additive inverse, the one which makes the sum of the elements of the corresponding column of matrix **G** positive is chosen (remember that each column of this matrix is derived from an eigenvector). In a more precise notation, the eigenvector $\mathbf{V}_j$ is chosen if:

$$\sum_{i=1}^{p} g_{ij} \geq 0,\ j=1,\ldots,m \quad (13)$$

Otherwise, its additive inverse is appropriate to be chosen.

ii. As a second adjusting method, it is suggested to choose an eigenvector if the following criterion is held:

$$\sum_{i=1}^{p} w_i g_{ij} \geq 0,\ j=1,\ldots,m \qquad (14)$$

Here, w_i stands for the weight of i^{th} initial criterion (variable). This inequality tries to make the most of absolute covariance coefficients between each factor and the initial criteria become positive values considering the importance of the initial criteria.

It was already mentioned that if the additive inverse of an eigenvector is substituted for itself, the corresponding principal component, the corresponding factor, and also the rotated factor become additively inversed. Thus, the corresponding loadings of a_{ij} and g_{ij} must be additively inversed so that Eq. (7) and (9) are held. In conclusion, for the indicated sum of Eq. (13) or the weighted sum of Eq. (14) to become additively inversed, it is enough to additively inverse the corresponding eigenvector.

Another approach is presented by Slottje et al. (1991) cited and applied by Zhu (1998) and Premachandra (2001) which tries to calculate the final synthetic utility value of each alternative using a simple additive weighting method upon the initial criteria. In order to overcome the distorting effect of existing severe interrelations, this approach exerts the influence of the PCA method by allocating new weights to the criteria. Considering Eq. (2) and the weights of principal components expressed in Eq. (4), the final synthetic utility values of n alternatives are calculated by the following equation:

$$\mathbf{Y}_{n\times 1} = \sum_{i=1}^{p} w_i' \mathbf{Z}_i = \sum_{j=1}^{p} \left(\sum_{i=1}^{p} w_i' v_{ij} \right) \mathbf{X}_j \qquad (15)$$

It is assumed in the literature that negative or positive signs may be assigned to the weights w_i', depending on the corresponding eigenvectors. This issue could be interpreted as classifying the principal components as benefit or cost criteria. The rule is expressed in this way: when the i^{th} eigenvector consists of nonnegative elements, the weight of principal component i is assumed to be positive. Otherwise, when the corresponding eigenvector consists of non-positive elements, the weight is assumed to be negative. However, the results of the eigenvectors' orthogonality feature make these criteria useless in general. Considering eigenvectors in which both positive and negative elements are involved simultaneously, the rule determines the sign w_i' so that all of the aggregated weights identified by Eq. (16) become nonnegative.

$$\widetilde{w}_j = \sum_{i=1}^{p} w_i' v_{ij}, j = 1, \dots, p \quad (16)$$

Again, the existence of cases lacking possible solutions makes this criterion, not a well-defined one.

Based on the discussed justifications of the proposed adjusting criteria, this rule may be modified to go with the FA method. Thus, when the i^{th} column of matrix **G**, as corresponding loadings of factor i, is consisted of nonnegative elements, the related factor weight is assigned to be positive; and when the mentioned column consists of non-positive elements, the corresponding weight is considered to be negative. Otherwise, considering $\mathbf{H}=[h_{ij}]_{p\times m}=\mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}$ in Eq. (11), the sign of factors' weight obtained by Eq. (12) are determined in a way that makes the aggregated weights defined by Eq. (17) all become nonnegative.

$$\widetilde{\widetilde{w}}_i = \sum_{j=1}^{m} w_j'' h_{ij}, i = 1, \dots, p \; (17)$$

Thus, the same as Eq. (15), the synthetic utility values of the alternatives are acquired using:

$$\boldsymbol{Y}_{n\times 1} = \sum_{j=1}^{m} w_j'' \boldsymbol{F}_j = \sum_{i=1}^{p} \widetilde{\widetilde{w}}_i \boldsymbol{X}_i$$

However, through the empirical analysis, it is demonstrated that in some cases, any permutation of the weights with negative or positive signs does not lead to proper nonnegative weights. This problem invalidates the generality of the recent approach and values the proposed adjustments.

Next, the precise procedure of the proposed approach is illustrated through a step-by-step algorithm.

3.3. Algorithm of the model

Step 1. Classify the data in a decision matrix with n rows allocated to alternatives scored on p columns of criteria and entitle it $\mathbf{X}_{n\times p} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p]$.

Step 2. Convert the cost criteria (smaller the better) to benefit criteria (bigger the better) by subtracting the corresponding columns from their maximum values.

Step 3. Standardize each column of the decision matrix by subtracting its mean value and dividing the result by the standard deviation.

Step 4. Multiply each column by its respective criterion weight and replace the resultant matrix with **X**. In case of equal weights for all criteria, skip this step.

Step 5. Calculate the covariance matrix of **X** and entitle it $\mathbf{C}_{p\times p}$.

Step 6. Calculate the eigenvalues and eigenvectors of **C** and sort them considering the descending order of eigenvalues. Note that the eigenvalues (expressed by λ_is, i=1, …, p) are all positive real scalars. Constitute the matrix $\mathbf{V}_{p\times p}=[\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_p]^{\mathrm{T}}$ whose rows consist of the eigenvectors in the mentioned order.

Step 7. Form the principal components matrix ($\mathbf{Z}_{n\times p}=[\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_p]$) using the equation: $\mathbf{Z}=\mathbf{X}\mathbf{V}^{\mathrm{T}}$.

Step 8. Considering $\lambda_i \big/ \sum_{k=1}^{p} \lambda_k$ the weight of principal component i, select the first m out of p principal components, with their aggregate weight value reaching 0.90. In case of equal weights for all criteria in step 4, select the principal components corresponding to eigenvalues greater than 1. Substitute the shrunk matrices of principal components and eigenvectors for **Z** and **V** respectively.

Step 9. Rescale the principal components to have variances equal to 1. For this, set $\mathbf{Z}_j \big/ \sqrt{\lambda_j} = \mathbf{F}_j$ and $\sqrt{\lambda_j}\, v_{ji} = a_{ij}$ (i= 1, …, p and j= 1, …, m). Thus, consider $\mathbf{F}_{n\times m}=[\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_m]$ as the on-rotated factors matrix and $\mathbf{A}_{p\times m}=[a_{ij}]$ as the on-rotated loadings matrix.

Step 10. Rotate the loadings matrix of **A** to form $\mathbf{G}_{p\times m}$ using orthogonal rotation methods. Computational software programs like MATLAB or PSAW (SPSS) could be utilized in this step. Calculate the sum of the squares of rotated factor loadings in each column of **A** to form the new eigenvalues after rotation (λ'_js).

Step 11. If the following criterion is not held for a j (j= 1, …, m), then multiply the corresponding row(s) in matrix **V** of step 6 by –1 and repeat the algorithm.

i. $\sum_{i=1}^{p} g_{ij} \geq 0,\ j = 1, \ldots, m$

The criterion (ii) below may be substituted for (i). Be aware to choose only one of these criteria to continue the algorithm in all iterations.

ii. $\sum_{i=1}^{p} w_i g_{ij} \geq 0,\ j = 1, \ldots, m$

Step 12. Estimate the rotated factor scores by $\mathbf{F}^*=\mathbf{X}\mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}$. $\mathbf{F}^*_{n\times m} = [\mathbf{F}^*_1, \mathbf{F}^*_2, \ldots, \mathbf{F}^*_m]$ stands for the final decision matrix with independent criteria expressed in columns and rows representing scores of alternatives.

Step 13. Calculate the weight values of factors by $w''_j = \lambda'_j \big/ \sum_{k=1}^{m} \lambda'_j$.

Step 14. Apply a MADM technique to rank the alternatives conveniently.

According to the discussed algorithm, the steps are briefly presented in Fig. 1.

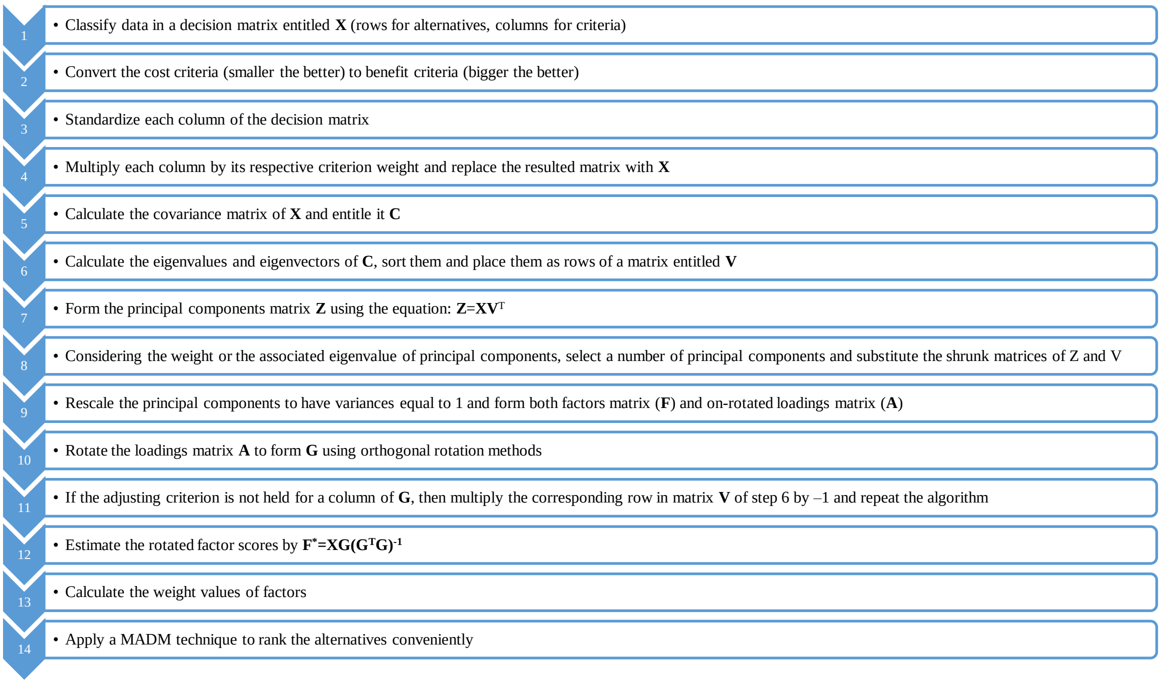

Fig. 1. The proposed approach in a step-by-step algorithm

4. Empirical Analysis

4.1. Case study: Ranking listed companies

Fundamental analysis and financial discrimination of companies have assumed great importance recently since it is beneficially applicable to evaluate the alternatives in portfolio management, supply chain management, credit scoring, ranking, etc. Here, it is intended to discriminate and rank 41 Iranian listed companies of the Tehran stock exchange regarding several impressing criteria. Consider the data displayed in Table 1 as a report of several financial indices obtained for the companies chosen from different industries. Data is extracted from balance sheets and income statements of 1393 and 1394 financial years in the solar Hijri calendar (starting from 2014 to 2016) and is arranged as illustrated in the 1st step. Since these financial indices are calculated upon some common data, the criteria are expected to be interrelated together. This issue is tested through the two following statistical tests based upon the correlations of criteria. The value of the *Kaiser–Meyer–Olkin* (*KMO*) measure of sampling accuracy is 0.663, which is much higher than 0.5 and thus is considered acceptable. Also, *Bartlett's test of sphericity* reached the value of 389.042, and considering 55 degrees of freedom, the associated level of significance is 0.000, indicating that the population correlation matrix is not an identity matrix. The test results indicated that the sample data was suitable for FA.

Table 1. Comparison of 41 Iranian listed companies considering 11 impressing financial indices

Index	Profit margin	Debt / Equity	Debt / Assets	ROA	ROE	Operating leverage	Current ratio	Operating earnings / Net sale	Interest coverage ratio	Asset turnover ratio	Cost of goods sold / Net sales
Type	Benefit	Cost	Cost	Benefit	Benefit	Benefit	Benefit	Benefit	Benefit	Benefit	Cost
Weight	**2**	**1**	**1**	**2**	**2**	**2**	**1**	**2**	**2**	**1**	**1**
1	0.05	4.38	0.81	0.05	0.31	21.53	1.19	0.12	1.85	1.14	0.85
2	0.03	15.23	0.94	0.01	0.19	1.74	0.93	0.11	1.50	0.41	0.83
3	0.17	3.40	0.77	0.16	0.69	0.73	0.87	0.26	4.72	1.11	0.65
4	0.10	1.44	0.59	0.11	0.31	0.69	1.27	0.16	2.49	1.17	0.69
5	0.08	2.96	0.75	0.04	0.17	-0.61	0.54	0.19	1.95	0.53	0.69
6	0.15	2.01	0.67	0.09	0.27	-5.99	1.30	0.20	6.85	0.58	0.72
7	0.25	0.78	0.44	0.10	0.20	0.10	1.36	0.34	5156.07	0.48	0.43
8	0.02	3.26	0.77	0.03	0.10	-9.62	0.93	0.03	1.83	1.44	0.97
9	0.07	4.05	0.80	0.05	0.17	0.13	1.09	0.19	1.55	0.56	0.74
10	0.08	3.59	0.78	0.08	0.40	-1.61	1.17	0.10	2.44	0.99	0.88
11	0.08	2.39	0.71	0.10	0.37	1.17	2.00	0.07	5.99	1.33	0.86
12	0.26	1.29	0.56	0.21	0.56	1.37	1.60	0.34	13.70	0.99	0.61
13	-0.05	4.25	0.81	-0.06	-0.27	5.10	0.46	0.03	0.38	1.01	0.93
14	0.15	1.17	0.54	0.14	0.31	21.67	0.83	0.09	6.24	0.99	0.84
15	0.38	0.65	0.39	0.44	0.82	0.91	1.52	0.41	67.14	1.22	0.51
16	-0.05	8.58	0.90	-0.01	-0.06	-6.11	0.48	0.00	0.72	0.14	0.93
17	0.13	1.81	0.64	0.12	0.35	5.22	1.30	0.21	6.22	0.91	0.74
18	0.27	0.92	0.48	0.14	0.25	2.49	1.09	0.28	12.74	0.53	0.65
19	0.07	3.52	0.78	0.09	0.47	-33.55	1.03	0.13	2.55	1.32	0.80
20	-0.06	9.66	0.91	-0.09	-0.63	1.97	0.70	0.01	0.04	1.39	0.94
21	0.11	1.81	0.64	0.11	0.34	-3.39	1.65	0.17	3.99	0.95	0.78
22	0.02	1.63	0.62	0.01	0.02	6.76	1.08	-0.01	1.39	0.39	0.75
23	0.22	1.00	0.50	0.19	0.39	0.05	1.37	0.30	8.23	0.90	0.61
24	-0.05	2.86	0.74	-0.02	-0.07	3.95	0.66	0.04	0.46	0.34	0.89
25	0.23	1.47	0.59	0.14	0.36	0.48	1.38	0.34	5.85	0.71	0.63
26	0.22	1.27	0.56	0.09	0.21	1.30	1.18	0.23	3.43	0.45	0.86
27	0.52	1.73	0.63	0.16	0.42	2.02	0.48	0.56	206.75	0.35	0.39
28	0.19	0.94	0.48	0.10	0.21	7.16	1.90	0.20	24.58	0.56	0.69
29	0.07	1.52	0.60	0.14	0.37	-9.76	1.14	0.09	6.11	2.06	0.85
30	0.01	1.31	0.57	0.02	0.04	4.79	1.37	0.00	1.31	1.11	0.89
31	0.15	2.35	0.70	0.11	0.39	16.41	0.92	0.27	3.17	0.78	0.63
32	0.19	2.55	0.72	0.06	0.28	0.84	1.16	0.19	13.92	0.45	0.80
33	0.01	3.34	0.77	0.04	0.14	0.50	0.72	0.02	1.92	3.08	0.96
34	0.03	3.62	0.78	0.02	0.10	31.06	0.96	0.07	1.82	0.90	0.92
35	0.43	0.47	0.32	0.32	0.51	1.14	1.71	0.46	31.89	0.81	0.47
36	0.05	7.46	0.88	0.06	0.57	1.71	1.05	0.11	4.65	1.45	0.73
37	-0.04	1.58	0.61	-0.01	-0.04	-2.89	1.82	0.02	0.51	0.38	0.85
38	0.27	0.88	0.47	0.39	0.94	1.39	1.61	0.26	16.51	1.71	0.69
39	0.03	5.99	0.86	0.02	0.12	-2.51	0.53	0.12	1.30	0.92	0.81
40	0.03	1.95	0.66	0.04	0.12	2.05	0.91	0.07	2.12	1.37	0.91
41	0.04	3.36	0.77	0.06	0.23	-1.22	0.94	0.06	2.42	1.41	0.91

The severe interrelation of criteria makes most MADM techniques incapable of reaching proper discriminative rankings. Therefore, the FA method is suggested to be used to eliminate the existing interrelation of criteria.

As is observable in the decision matrix, 11 impressing criteria are taken into consideration to evaluate the financial situation of the companies. Profit margin (I/S), as the first index, is the ratio of net income to net sales in a financial year. It indicates the profitability generated from revenue and hence is an important measure of operating performance. It also provides clues to a company's pricing, cost structure, and production efficiency. Debt ratio (D/A) is another index obtained by dividing total liabilities by total assets. This index is a cost criterion and compares total debt to total assets. The well-defined index of debt/equity ratio (D/E) is considered the next criterion having a cost nature. The index, calculated by dividing total liabilities by stockholders' equity, is a significant measure of solvency since a high degree of debt in the capital structure may make it difficult for the company to meet interest charges and principal payments at maturity. Thus, the smallest values are considered as best for these two criteria. The succeeding criterion, return on assets (ROA), is defined by the ratio of net income to average total assets and indicates the efficiency with which management has used its available resources to generate income. Also, return on equity (ROE) is the ratio of earnings available to common stockholders to the average stockholders' equity and measures the rate of return earned on the common stockholders' investment. Operating leverage (OL) is the ratio of the percentage change in operating earnings (or EBIT) to the percentage change in sales and is a measure of operating risk which arises from fixed operating costs. A simple indication of OL is the effect that a change in sales has on earnings. The current ratio (CR), applied as the next criterion, is equal to current assets divided by current liabilities. This ratio, which is subject to seasonal fluctuations, is used to measure the ability of an enterprise to meet its current liabilities out of current assets. The operating earnings to sale ratio (OE/S) are considered a financial index, gained by dividing operating profit by net sales. The ratio of earnings before interest and taxes (EBIT) to interest expense leads to the interest coverage ratio (ICR), as another benefit criterion. It is a safety margin indicator in the sense that it shows how much of a decline in earnings a company can absorb. Total asset turnover ratio (ATR) stands for a financial index, obtained by calculating the ratio of net sales to the average total assets. This index helps evaluate a company's ability to use its asset base efficiently to generate revenue. Finally, the cost of goods sold to sale ratio (CGS/S), as the last index, is a cost-natured criterion with the smallest values considered as best calculated by dividing the cost of goods sold by net sales (Neveu 1989; Shim and Siegel 2007).

As it is obvious, the indices dealing with debt are considered cost criteria and thus must be treated based on the 2nd step of the algorithm. Moreover, the indices dealing directly with profit are intended to gain doubled importance in this research. The criteria weights resulting from this intent are presented in Table 1.

After calculating the covariance matrix of the standardized and weighted data according to steps 3, 4, and 5, the corresponding eigenvalues and eigenvectors of the covariance matrix are obtained based on step 6. Applying steps 7 to 10 results in the rotated factor loadings using Varimax rotation with Kaiser's normalization method. According to the 8th step, the first m=5 principal components with larger eigenvalues and an aggregate weight value reaching 0.90 should be chosen. Thus, 5 factors are extracted. By selecting the first criterion in the 11th step, the 1st, 2nd, and 3rd eigenvectors are to be multiplied by –1, so that the criterion is satisfied. Thereafter, the algorithm is repeated once more from step 6 to the end. This results in the correlation matrix between the criteria and adjusted rotated factors. Considering matrix **G** consisting of covariance coefficients, the recent matrix is calculated simply by dividing the element (i,j) (i=1, …,11, j=1, …, 5) of **G** by multiplication of the standard deviations of i^{th} standardized and weighted criteria and j^{th} factor, in which the second standard deviation value is always equal to 1. Due to the rather small number of alternatives, correlation coefficients greater than 0.65 are considered meaningful relations. In consistency with the result of rotation, this shows that each criterion is correlated severely by only one of the factors. According to step 12, the adjusted and rotated factor scores extracted from the financial indices, are estimated, and the importance of the recent factors is acquired using step 13. Now, the factors could be applied practically to rank the companies using MADM methods in a reliable context and without interrelation.

Now, to investigate the other adjusting approach applied by Zhu (1998), which is illustrated in section 3.2, various positive or negative signs are assigned to the weights of principal components and also factors obtained respectively by Eq. (4) and (12). Thus, 32 different permutations of weights are generated in each case. Therefore, according to Eq. (16) and (17), 32 sets of aggregated weights are generated for the financial indices in each method of PCA and FA. Nevertheless, for both PCA and FA methods, there is not any case among these sets with all the aggregated weights having nonnegative values. So, the adjusting criterion is not satisfied in any case, and that's why the problem of ranking remains unsolved. Finally, as was claimed before, it is concluded that this criterion ought not to be applied as a general rule.

Next, considering the special specifications of the ELECTRE method, it is applied to specify the preference of companies with respect to each other. Its first idea concerning concordance, discordance, and outranking concepts originates from real-world applications. Since any direct summation of the scores of each alternative upon the criteria is not adopted by ELECTRE, it has a particular ability in dealing with cases carrying severe interrelations. Thus, the initial decision matrix, the final unadjusted one, and also the adjusted one are explored by the ELECTRE I method, and three preference matrices are summarized in Table 2. The estimated factor scores without performing the adjustments of step 11 are used to establish the final unadjusted decision matrix. Also, the final adjusted decision matrix is constructed considering the proposed approach.

Table 2. Number of alternatives being preferred and preferred to by each alternative considering three decision matrices based on the results of ELECTRE I method for Iranian listed companies

	Initial data			Unadjusted factors			Adjusted factors		
	Being preferred	Preference	Preference Indicator	Being preferred	Preference	Preference Indicator	Being preferred	Preference	Preference Indicator
1	2	15	13	18	6	-12	7	15	8
2	28	1	-27	25	2	-23	12	8	-4
3	5	25	20	31	1	-30	4	30	26
4	15	16	1	10	20	10	11	14	3
5	19	9	-10	20	10	-10	17	8	-9
6	16	7	-9	12	16	4	15	11	-4
7	0	33	33	31	0	-31	0	31	31
8	30	2	-28	1	31	30	31	2	-29
9	19	9	-10	16	12	-4	16	9	-7
10	18	11	-7	15	20	5	13	10	-3
11	15	14	-1	2	28	26	9	11	2
12	4	29	25	28	4	-24	3	32	29
13	31	1	-30	3	20	17	34	1	-33
14	1	28	27	12	7	-5	7	16	9
15	0	33	33	26	1	-25	0	38	38
16	36	1	-35	9	23	14	31	2	-29
17	10	21	11	19	11	-8	10	22	12
18	5	24	19	21	5	-16	7	23	16
19	23	0	-23	1	20	19	19	1	-18
20	39	0	-39	1	27	26	38	0	-38
21	14	11	-3	6	23	17	14	11	-3
22	22	5	-17	2	27	25	24	4	-20
23	6	25	19	17	8	-9	5	26	21
24	28	3	-25	4	21	17	31	2	-29
25	7	24	17	25	4	-21	6	23	17
26	13	20	7	18	11	-7	11	14	3
27	0	32	32	39	0	-39	0	22	22
28	6	23	17	10	21	11	8	11	3
29	17	5	-12	5	29	24	15	7	-8
30	24	6	-18	1	35	34	22	2	-20
31	1	27	26	30	2	-28	2	27	25
32	12	16	4	19	9	-10	12	15	3
33	11	9	-2	0	29	29	24	4	-20
34	1	9	8	16	7	-9	7	12	5
35	0	32	32	31	1	-30	1	36	35
36	12	16	4	19	4	-15	8	15	7
37	30	2	-28	0	39	39	35	0	-35
38	1	32	31	21	3	-18	1	35	34
39	28	3	-25	18	11	-7	20	8	-12
40	21	7	-14	6	26	20	22	7	-15
41	22	6	-16	11	25	14	23	10	-13

As a result, obtained from the preference matrix of the initial data show the companies of number 7, 15, 27, and 35 are not preferred by any of the alternatives and are preferred by most of them. This result is almost held considering the adjusted factors. However, the preference matrix of the unadjusted factors demonstrates a quite contradictory result for the mentioned alternatives, i.e., they are not preferred to any, or only preferred to one of the alternatives, whereas numerous numbers of the alternatives are preferred to them. Thus, these companies are considered as best by the initial data and adjusted factors whereas they have been classified as worst according to the unadjusted approach. This indicates that the unadjusted approach is incapable of making decision matrices with similar specifications to the initial data.

Next, it is intended to apply the VIKOR method to obtain a more explicit comparison of the decision matrices. The VIKOR method was developed as a MADM method to solve discrete decision problems with non-commensurable and conflicting criteria (Opricovic and Tzeng, 2004). It introduces the multi-criteria synthetic utility value based on the particular measure of "closeness" to the "ideal" solution (Opricovic, 1998). The multi-criteria measure for compromise ranking is developed from the L_p-*metric* used as an aggregating function in a compromise programming method.

Thus, three sets of synthetic utility values are obtained using the VIKOR method (with parameter 0.5) upon the initial decision matrix, the unadjusted factors, and also the adjusted factors. Considering the VIKOR synthetic utility values varying in the range between 0 to 1, and regarding their cost nature of them (smaller the better), three sets of comparable scores are acquired after subtracting the synthetic utility values from 1 and normally standardizing them. The result is demonstrated in Fig. 2. As it is clear, and in conformity with the ELECTRE results, the scores of the unadjusted decision matrix are in apparent contradiction with the other ones in most cases. According to the diagram, despite similar scores of the initial data and the adjusted factors, the second one can discriminate the alternatives more distinctly, and this is a quite favorable result in multiple interrelated criteria decision-making.

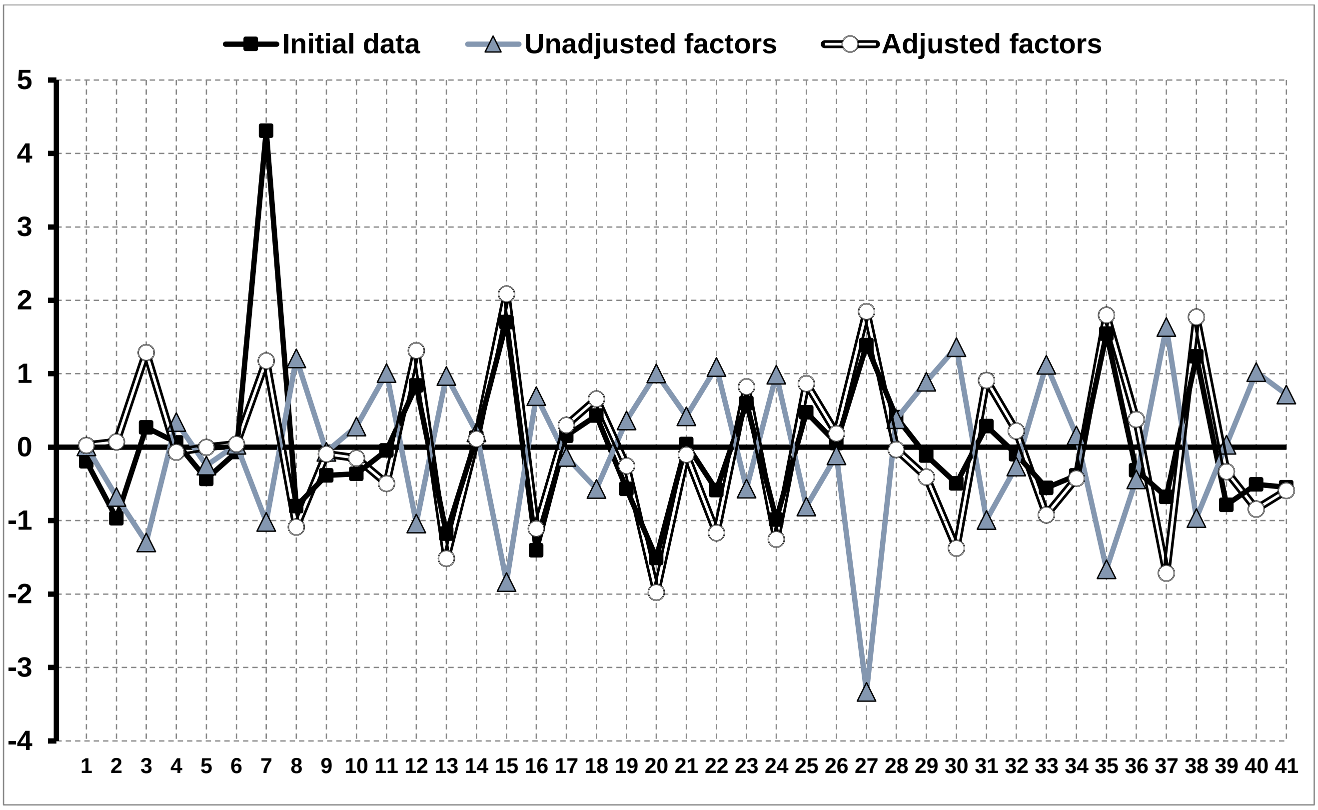

Fig. 2. Discrimination of 41 listed companies based on different decision matrices applying the VIKOR method

To investigate the validity of the proposed adjustments in PCA or FA, it is necessary to compare the evaluation results before and after applying adjustments with the evaluation results of the initial dataset. Since the proposed method serves to reduce or eliminate the underestimations or overestimations, it is expected that the proper after-results comply with the before-results in an approximation. The Euclidean distance is used here to compare between results. As is observable in Table 3, the Euclidean distance of the results of the initial dataset from the adjusted dataset is significantly less than the unadjusted one. As was expected, this result is true for the utility values provided by both ELECTRE and VIKOR methods. This indicates the necessity of adjustments in the proposed evaluation process. It is noted that the final utility value (the preference indicator) is computed in ELECTRE by subtracting the 'being preferred' measure from 'preference' in Table 2.

Table 3. The Euclidean distance between utility values of different datasets

	Initial utility values - ELECTRE	Initial utility values - VIKOR
Unadjusted utility values	253	3.46
Adjusted utility values	45	1.88

To validate the capability of the proposed method in its core functionality, i.e., in reducing or eliminating the underestimations or overestimations, descriptive statistics of VIKOR results are clarified. According to Table 4, the mean utility value is reduced in the adjusted dataset compared to the initial dataset, as was expected. This is considered the evidence of removing overestimations made by applying the addition operation in VIKOR computations. In the meantime, the variance of utility values is considerably increased for the adjusted dataset compared to the initial dataset. This indicates that the distinguishing power of the proposed process is enhanced as a result of removing or reducing underestimations and overestimations. Thus, the proposed approach is applicable to complement VIKOR or other MADM techniques such as SAW, AHP, and TOPSIS. ELECTRE is not considered here since it is not regarded as a technique significantly exposed to underestimation or overestimation effects. This is because alternatives' preference is specified in ELECTRE without aggregating 'components of decision matrix' (criteria-based evaluations of alternatives).

Table 4. Descriptive statistics of VIKOR utility values

	Mean	Variance
Initial utility values	0.757	0.026
Adjusted utility values	0.507	0.059

In conclusion, the unrealistic results acquired by the unadjusted FA, and the conforming results obtained from the initial criteria and the uncorrelated factors, implying that the proposed adjustments are reasonable and essential for the algorithm.

4.2. Adjustments in practice: Investigating the validity

Here, 5 sets of information drawn out as the adjusted factors in the previous section, are taken into consideration as the inputs of the proposed algorithm. All of these factors are assumed to have equal importance and thus there is no need to make this decision matrix weighted. Each set compares 41 different alternatives from the viewpoint of a common factor and has a variance of 1 and a mean value of zero. Also, the correlation coefficient for each pair of sets is equal to zero and thus there is not any interrelation between the criteria. Therefore, the covariance and correlation matrices of this decision matrix are equal to the identity matrix. So, the same decision matrix is expected to be acquired through the implementation of the FA method one more time. Note that there is no need to implement steps 1 through 4 of the algorithms here. Thus, the unadjusted new factor loadings are calculated. The structure of the loadings makes the rotation unnecessary and unreasonable. The loading matrix may be considered as matrix **G** in the calculation process of new factors. Thus, a decision matrix would be obtained in which the 2nd, 3rd, and 5th columns become additively inversed (multiplied by –1) concerning the former dataset including adjusted and rotated factor scores. But this is quite an unreasonable result because it is expected that the same decision matrix is acquired through FA in the absence of interrelation. However, implementing one of the adjustments of step 11 makes the corresponding eigenvectors additively inversed, and this modification results in the desired output.

5. Conclusions and Further Research

The problem of ranking some facing alternatives was discussed through this research considering numerous severely interrelated criteria in the framework of MCDM. The research intended to describe the disadvantages of routine decision-making methods in the absence of criteria independence and attempted to propose an integrated step-by-step algorithm to solve the problem by preparing a dependable solution. Prerequisite subjects were discussed, and necessary justifications were reasoned through the text. Eventually, the real-world case study of discriminating against some listed companies of the Tehran stock exchange was applied considering numerous financial indices to investigate the results. Since all of the indices were extracted from common information describing the financial situation of companies, it was expected to deal with a practical situation of the criteria interrelation problem. This issue was tested through two statistical tests and the expected conclusions were obtained. Thus, two sets of data were extracted from the initial one using the algorithm in the absence and presence of the proposed adjustments of the research. Against other MADM methods, particular specifications of the ELECTRE method such as the avoidance of direct summation between the interrelated criteria scores was a proper reason to apply it as a performance evaluator of the research approach. The evaluations based on the ELECTRE method and displayed distinctly by the VIKOR method verified the proper adjustments of the proposed approach. Finally, the adjustments were confirmed more explicitly by a second example.

As a suggestion for further research, it is useful to investigate the extension of the method in non-deterministic contexts which can prepare a more practical and potent situation for decision-making in the real world. Also, considering the limitations of the PCA method as an initial solution, it is worthwhile to apply more advanced FA approaches such as maximum likelihood, to produce more precise results out of data reduction. Comparison of the research approach with other combinations of FA and discrimination approaches such as the fuzzy integral method could also be a good idea to investigate the performance of the proposed method. Since excluding nonessential data attributes is impressive in boosting machine learning algorithms in the presence of high data dimensionality, it would be beneficial to conduct studies on the influence of adjusted factor analysis in this regard. This will help in uncovering patterns for the digitization of sectors such as healthcare, manufacturing, marketing, IoT processing, and recommendation systems.

REFERENCES

Bolch, B. W., & Huang, C. J. (1973). Multivariate statistical methods for business and economics. Prentice Hall.

Chen, X., Wu, M., Tan, C., & Zhang, T. (2021). A random intuitionistic fuzzy factor analysis model for complex multi-attribute large group decision-making in dynamic environments. Fuzzy Optimization and Decision Making, 20(1), 101-127.

Cheng, K. C., Huang, M. J., Fu, C. K., Wang, K. H., Wang, H. M., & Lin, L. H. (2021). Establishing a multiple-criteria decision-making model for stock investment decisions using data mining techniques. Sustainability, 13(6), 3100.

Chiou, H. K., Tzeng, G. H., & Cheng, D. C. (2005). Evaluating sustainable fishing development strategies using fuzzy MCDM approach. Omega, 33(3), 223-234.

Dugger, Z., Halverson, G., McCrory, B., & Claudio, D. (2022). Principal Component Analysis in MCDM: An exercise in pilot selection. Expert Systems with Applications, 188, 115984.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Statistical significance versus statistical power. Multivariate Data Analysis, 5th Edn. New Jersey, Prentice-Hall, 10-13.

Hatami-Marbini, A., Hekmat, S., & Agrell, P. J. (2020). A strategy-based framework for supplier selection: A grey PCA-DEA approach. Operational Research, 1-35.

Hekmat, S., Amiri, M., & Madraki, G. (2021). Strategic supplier selection in payment industry: A multi-criteria solution for insufficient and interrelated data sources. International Journal of Information Technology & Decision Making, 20(06), 1711-1745.

Heydari, A., Niroomand, S., & Garg, H. (2022). An improved weighted principal component analysis integrated with TOPSIS approach for global financial development ranking problem of Middle East countries. Concurrency and Computation: Practice and Experience, 34(13), e6923.

Huang, M., Xia, T., Chen, Z., Pan, E., & Xi, L. (2021). A DEA integrated grey factor analysis approach for efficiency evaluation and ranking in uncertain systems. Computers & Industrial Engineering, 162, 107681.

Kuo, M. S., & Liang, G. S. (2011). A novel hybrid decision-making model for selecting locations in a fuzzy environment. Mathematical and Computer Modelling, 54(1-2), 88-104.

Mahbub, P., Ayoko, G. A., Goonetilleke, A., & Egodawatta, P. (2011). Analysis of the build-up of semi and non-volatile organic compounds on urban roads. Water Research, 45(9), 2835-2844.

Manly, B. F., & Alberto, J. A. N. (2016). Multivariate statistical methods: a primer. Chapman and Hall/CRC.

Mousavi, S. S., Nezami, F. G., Heydar, M., & Aryanejad, M. B. (2009, July). A hybrid fuzzy group decision-making and factor analysis for selecting maintenance strategy. In 2009 International Conference on Computers & Industrial Engineering (pp. 1204-1209). IEEE.

Neveu, R. P. (1989). Fundamentals of managerial science, 3rd edn. Cincinnati: South-Western Pub Co.

Opricovic, S. (1998). Multicriteria optimization of civil engineering systems. Faculty of civil engineering, Belgrade, 2(1), 5-21.

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European journal of operational research, 156(2), 445-455.

Premachandra, I. M. (2001). A note on DEA vs principal component analysis: An improvement to Joe Zhu's approach. European Journal of Operational Research, 132(3), 553-560.

Reddy, G. T., Reddy, M. P. K., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). Analysis of dimensionality reduction techniques on big data. IEEE Access, 8, 54776-54788.

Shim, J. K., & Siegel, J. G. (2007). Schaum’s outline of financial management, 3rd edn. New York: McGraw–Hill.

Stević, Ž., Miškić, S., Vojinović, D., Huskanović, E., Stanković, M., & Pamučar, D. (2022). Development of a Model for Evaluating the Efficiency of Transport Companies: PCA–DEA–MCDM Model. Axioms, 11(3), 140.

Tung, C. T., & Lee, Y. J. (2009). A novel approach to construct grey principal component analysis evaluation model. Expert Systems with Applications, 36(3), 5916-5920.

Tung, C. T., & Lee, Y. J. (2010). The innovative performance evaluation model of grey factor analysis: A case study of listed biotechnology corporations in Taiwan. Expert Systems with Applications, 37(12), 7844-7851.

Tzeng, G. H., Chiang, C. H., & Li, C. W. (2007). Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. Expert systems with Applications, 32(4), 1028-1044.

Zhu, J. (1998). Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities. European journal of operational research, 111(1), 50-61.