(In)consistency of the Selection of the Method of Multiplicative Decision-Making

Summary

The paper analyzes the main causes of the (in)consistency of the selected method of multiplicative decision-making: data normalization, weight coefficients and the application of the Likert scale for the purpose of measuring quantitative attributes. Normalized data in the methods of multiatributive decision making represents the substitute for a subjective attribute ratings by decision makers. Since they are calculated on the basis of mathematical transformations of empirical data, one gains the impression that the choices based on normalized values are „objective“. Therefore, the sensitivity analysis of the results has dealt exclusively with effects of weight coefficients on the final choices so far, while the potential impact of normalization is completely ignored; meanwhile, the deformations caused by the normalization of data have been attributed to the effects of weight coefficients and their inevitable subjectivism. We intent to point out at the deformations of empirical values that are the result of normalization and which call into question the application of normalized values as a deci-
sion base. It can be proven that the normalized values are an unreliable information base for decision-making. In addition, the (in)consistency of selection methods of multi-attributive decision-making is also influenced by changes in the method of measuring and formulating attributes.

**Keywords:** Multi-attributive decision-making, data normalization, weight coefficients, Likert scale

**JEL Classification:** C44.

**Introduction**

Within the prescriptive decision theory, a number of methods have been defined to support us in solving various groups of problems. Multi Attribute Decision Making Methods (MADM) are defined to solve the problem of choosing between complex options in terms of certainty, uncertainty and risk. The problem is formally presented by choosing one of the \( m \) options \( A_i, (i=1,2,...,m) \) which we evaluate and compare among ourselves on the basis of \( n \) characteristics, attributes, \( X_j, (j=1,2,...,n) \) whose values we know. We display the options by vectors \( A_i=(x_{i1},x_{i2},...,x_{ij},...,x_{in}) \) where \( x_{ij} \) value of the \( i \) option is \( j \) by the corresponding attribute. Since the attributes influence the final estimates of the variables to varying degrees, we attribute to each attribute the weight coefficient (weight or ponder), \( j=1,2,...,n \), (where \( \sum w_j = 1 \)), which reflects its relative importance in the evaluation of the options. By applying the MADM method to this formulated selection problem, we form a ranking list of options by priority or determine an optimal, ie, set of optimal options.

Within operational research, there are a number of methods applicable for choosing between complex options. Some of them are based on the concept of utility, while other methods are applied to values obtained by simple mathematical transformations of empirical data (the so-called normalizations). The empirical values of the attributes are mapped to the scale \([0,1]\). Options are then evaluated using different procedures that, due to different approaches to the problem, can also have different solutions. Although normalization procedures greatly facilitate and speed up the application of different methods (since they allow the application of computer programs), they unfortunately, significantly distort data and are not an adequate substitute for utility.

The problem of precise calculation of the assessment of each option is very complex, for at least two reasons: the first is that the attributes are incomparable in size (these are the properties of the options we express in different units of measure), and the other problem is that the weight coefficients (ie, relative significance of certain attributes) are difficult to pinpoint and phrase numerically.
Due to the problem of precise quantification (cardinal utility and / or weight coefficients) of the solution, the methods of multatributive decision-making should be followed by their “checking”, and testing the sensitivity of the results. If obtained result is stable and, to what extent, the changes in the values of the weight coefficients and the cardinal utility of the attribute will be examined by analyzing the sensitivity. But, on the basis of all this, it would be wrong to conclude that the present problems arise solely because of the imperfections of ordinary decision makers.

The reasons for the inconsistency of the solution can be found in the logic of selecting certain methods, which is why there are criteria in this field for testing the logical consistency of their solutions. Although we will not stay on them, they undoubtedly emphasize the importance of the decision-making, i.e. a deliberate choice of procedures by which we make a choice.

Each MADM method is characterized by its specific selection criteria, which is why, by applying different methods to the same problem of choice, different results are obtained as a rule.

Even more comparative analyzes have been used by authors in recent past in order to find out the characteristics of the problem of choices that condition equality, that is, the differences in the solutions of certain MADM methods. Not disputing their importance we must say, however, that these analyses and their results do not say anything about the rationality of the solutions offered; namely, the fact that several methods suggest the same choice is not yet sufficient guarantee of their “quality”. This indicates the need for the phase of selection of the MADM method to be deprived by the arbitrariness and to objectify the selection of the decision making method.

1. Previous research

Although theory and decision-making in conditions of uncertainty have existed for a long time, a critical review of some of its parts has been omitted. In particular, we mean two major problems, which are: normalization of empirical values and weighting, ie. assigning weight coefficients to individual attributes. There are not many authors who have dealt with these problems. However, we can distinguish several critical thoughts.

Classical methods of normalization do not always consider situations in which a different nature of data can affect the final outcomes. New methods of normalization are proposed (Acuña, Liern and Pérez-Gladish, 2018), which are based on the similarity of options with respect to the attributes with an ideal solution. Similarly, some authors (Mathew, Sahu and Upadhyay, 2017) suggest the use of a new normalization method, WASPAS (weighted aggregated sum product as-
essment method) as the best normalization method. New methods eliminate all mistakes and problems that have been observed in commonly accepted methods.

The application and effects of different methods of normalization on the well-known method of multi-criteria optimization-weighted averages, WA (Weighted Average), or simply additive weighting, SAW (Simple Additive Weighting), show that out of six most common normalization techniques, even four do not fulfill basic conditions of consistency (Vafaei, Ribeiro and Camarinha-Matos, 2018).

The study of the effects of weight variation of the most important and most critical criteria for the stability of the ranking of six considered methods of multi-criteria optimization is performed in both one-dimensional and multidimensional analysis of the sensitivity to weight. For this purpose, the use of the MOOR method (The Multiobjective Optimization on the basis of Ratio Analysis methods) is suggested. The optimization of weightings between individual attributes (Karande, Zavadskas and Chakraborty, 2016) optimizes to a large extent.

Some authors also consider that the choice of the scale has a direct influence on the selection of the appropriate statistical technique of data analysis (Soldić-Aleksić and Krasavac, 2009). This problem is discussed in following lines of this paper.

2 Measurement scales and statistical techniques

Relationship between the measurement scales and the statistical techniques that can be applied to a particular type of data is shown in Table 1.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Descriptive measures</th>
<th>Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Percents, mode</td>
<td>Chi-squared, Binomial test</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Percentiles, median</td>
<td>Rank correlation, Friedman\ANOVA</td>
</tr>
<tr>
<td>Interval</td>
<td>Distance, weighted average, standard deviation</td>
<td>Correlation coefficient, t-test, ANOVA, regression, Factor analysis</td>
</tr>
<tr>
<td>Ratio</td>
<td>All descriptive measures for mean value and variability, shape of distribution</td>
<td>All statistical techniques</td>
</tr>
</tbody>
</table>

Source: Soldić-Aleksić, J., Chroneos Krasivac, B., Quantitative techniques in market research, Faculty of Economics, Belgrade 65, 2009.

In addition to the already mentioned data type division, often in practical research, division into categorical and non-categorical data, or metric and non-metric data, is used. For the above-mentioned types of primary scales, data measured on a nominal or ordinal scale are categorical data, while data measured on an interval or ratio scale are designated as metric data.
3 Scale selection, Likert scale

The scale selection on which one attribute is measured is not arbitrary. It does not depend on the good will or the needs of the decision maker, but is determined by the nature of the observed attribute. It is possible, however, that attributes that are measurable on more precise scales are measured on less precise scales, but not vice versa. For example, instead of showing income in a monetary amount (on a ratio scale), we can rank them in size, i.e. use an ordinal scale, or we can group them into related subgroups, e.g. by different sources, that is, apply a nominal scale. In doing so, with each transition from a more precise toward a less precise scale, we lose part of the previously available information. However, the reversed procedure is not possible.

Also, we can not use an ordinal scale to measure qualitative attributes. It is meaningless to say, for example, that one candidate's knowledge is twice the knowledge of another (ratio scale), or that candidate A's knowledge compared to candidate B is three times higher than the candidate B's knowledge compared to candidate C (interval scale) and the like. In other words, the division of attributes into quantitative and qualitative is the consequence of their different metrics; although quantitative attributes can be measured on less precise scales, we can not use ordinal or interval scale to measure qualitative attributes.

It is also important to note that the results of most mathematical operations do not make sense if they are done on numbers from an ordinal scale, and that mathematical operations on numbers with a nominal scale are meaningless.

The application of most of the MADM methods assumes that the qualitative attribute modalities are also given in the numerical form. In order to display the options as value vectors, we need to associate the levels of the ordinally measurable attributes and verbal descriptions of strictly qualitative attributes with numbers.

For these purposes, we use the so-called Likert scale with 5 (1, 2, 3, 4, 5) or 9 (1, 2, 3, 4, 5, 6, 7, 8, 9) points. The Likert scale originates from psychology where it is used to measure attitudes. It represents the so-called well-ordered ordinal scale, which differs from the classical ordinal scale in that each number is accompanied by a verbal description, so that the entire domain of the attribute is covered by the scale. In psychology, the numbers as a rule have the following meaning: 1 - I completely disagree, 2 - I do not partially agree, 3 - I am neutral, 4 - I partially agree, 5 - I completely agree. These descriptions cover all the manifest forms of the degree of agreement of the respondents with the stated viewpoint.

Within the MADM analysis, the verbal description of each number from the Likert scale is adapted to the observed attribute. If the attribute is measurable on an ordinal scale, then the numeric values attributed to individual modalities reflect the different intensities of that attribute. For example, when comparing
candidates for knowledge, numbers on scale 1, 2, 3, 4, 5 represent the following levels of knowledge, respectively: weak, sufficient, good, very good, excellent. Sometimes instead of the scale: 1, 2, 3, 4, 5, its transformation: \( y = 2x - 1 \) is used, i.e. scale: 1, 3, 5, 7, 9, whose numbers are interpreted in the same way. In the case of the application of a 9-point scale, even numbers allow us to compare the levels of attributes more subtly among them.

If the attribute is measurable on a nominal scale, then numbers from the Likert scale reflect our tastes and preferences. For example, when comparing apartments by location, numbers 1, 2, 3, 4, 5, we express our preferences to different locations: extremely unfavorable, unfavorable, average, favorable, very favorable. Since it is about individual preferences, it is possible that different individuals (according to their different tastes and needs) assign different numbers to the same locations, and that the individual differentiates the locations of the same apartments depending on their intended purpose, for example, whether they are observed by them as a residential or commercial space.

We have seen that, regardless of the type of qualitative attribute (ordinally measurable or strictly qualitative), number 1 is always attributed to his worst modality, and number 5 (that is, 9) is the best mode. Since a larger number indicates a more favorable modality, we conclude that between the two options we always choose one with a higher value of a qualitative attribute. This means that between the numbers on the Likert scale and our usefulness there is a positive correlation, that is, qualitative attributes belong to a group of income attributes (which is important to keep in mind in the normalization phase of all values).

In what way does the Likert scale differ from the classical ordinal scale and why can not we accept it as an interval scale? The Lycert scale is in a certain sense more informative than the classical ordinal scale, but at the same time it is less precise than it is. As we have already said, an ordinal scale reveals only the relative order of options per given attribute. Based on ranks, we do not receive information about the “quality” of individual modalities, or their position within the domain of the observed attribute.

Options with their “values” can cover the entire domain of attributes, or just a small segment. For example, everything can be very good or very bad for a given attribute, which can not be detected on the basis of ranks. It is also possible that the first-ranked ones are much more or less different from the other, than the other one is different from the third, etc. The ranking list does not contain this information because the unit difference between the individual ranks does not indicate the difference between the two levels of attributes directly ranked one below the other.

The Likert scale is more informative because the verbal description we attach to each number indicates a certain level of attribute; therefore, based on the
number, we can determine the position of the modalities within the domain of the observed attribute. As we said, number 1 corresponds to the lowest level, and the number 5 (or 9) is the highest attribute level. But this trait hides the danger of incorrect interpretation of numbers and their differences, and the acceptance of the scale as interval. With a superficial look at the numbers and their descriptions, we might conclude that the difference between the numbers reflects precise differences between the individual levels, which is not true. Not only does the Likert scale have no interval scale characteristics, but it is in a certain sense even less precise than the classical ordinal scale. Its inaccuracy is especially noticeable if the attribute levels of some options are assigned the same number. Equality of numbers does not in any way imply a strict equality of options for a given attribute, but above all reflects the limit of the scale to only 5 or 9 points.

For example, between the two “average” options, to which we have assigned number 3, there may be a small, but still noticeable difference, which we cannot numerically display due to an insufficiently precise scale. With these options, we will join different ranks in the ranking, which suggests that a classical ordinal scale allows a more subtle comparison of the modalities of the ordinally measurable attributes.

Since the same number on the Likert scale does not mean a strict equality of options for a given attribute, it follows that even the same difference between numbers (eg 5-4 = 4-3 = 2-1) does not reflect identical differences between the corresponding modalities. Therefore, these differences cannot be compared with each other precisely. For example, if we have assigned the numbers to the three options: A1 - 5, A2 - 3 and A3 - 2, based on the relationship between the intervals (5-3): (3-2) = 2: 1, we cannot conclude that A1 compared to A2 is twice as good as A2 compared to A3 (that “excellent” knowledge compared to “good” is twice as good as “good” compared to “enough”). The differences between the modalities of the ordinally measurable attributes cannot be treated in the same way as the differences and relationships between the values of the quantitative, cardinally measurable attributes.

A similar situation exists with nominally “measurable” attributes. Numerical values attributed to different modalities of these attributes reflect our subjective attitudes, that is, the structure of our preferences. In this case, we can accept the numbers on the Likert scale only as roughly determined ordinal utilities that allow us to rank the options according to priority. They can be less precise than classical ordinal utilities, because we do not need to be indifferent between the options to which we have assigned the same number; on a classical ordinal scale, we would rank them differently based on the subtle but present differences in preferences.
Moreover, numbers on the Likert scale cannot be treated as cardinal utilities, which are measurable on the interval scale. Recall that the methods for calculating cardinal utilities (for example, the standard method of game theory of Neumann and Morgenstern) are very complex, and that they are based on a set of rigorous assumptions concerning the preferences of a decision maker.

The simple procedure by which numbers from the Likert scale are attributed to the modalities of strictly quantitative attributes cannot be accepted as a procedure of calculating cardinal utility.

Beware that in the literature devoted to multi-attributive analysis, the Likert scale is primarily treated as an ordinal scale. However, the problem arises in the application phase of the Likert scale, when it is implicitly accepted as an interval scale. The reason for this is the fact that the selection criteria of the MADM method are calculated using mathematical operations whose results make sense only if the data are measurable on cardinal scales.

4 Normalization of empirical values

Normalization ensures that by aggregating values of all attributes we can calculate the criterium on the basis of which the options are being evaluated, compared among themselves and the best is being selected.

The available literature pays little attention to the normalization steps, and based on the information we have, the side effects of the normalization have not been more seriously analyzed. We believe that the reason for that is the specific treatment that normalization procedures have in MADM methods; they are, namely, considered a contribution to the more objective evaluation of options.

There are a huge number of normalization procedures, the most common of which are the following three: Simple Normalization (SN), Linear Normalization (LN) and Vector Normalization (VN).

Table 2 contains formulas for normalization of income and expense attributes according by all three methods.

Normalized data in MADM methods represent a substitute for subjective attribute ratings by DO. Since they are calculated on the basis of mathematical transformations of empirical data, the impression is that the choices based on normalized values are “objective”.

Therefore, the sensitivity analysis of the results has so far dealt exclusively with the effects of weight coefficients on the final choices, while the potential impact of normalization is completely ignored; according to this, deformations caused by the normalization of data are attributed to the effects of weight coefficients and their inevitable subjectivism.
### Table 2
Normalization of income and expense attributes

<table>
<thead>
<tr>
<th>Type of attribute</th>
<th>Normalization</th>
<th>Revenue, $X_j^+$</th>
<th>Expense, $X_j^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>$r_{ij} = \frac{x_{ij}}{x_j^*}$, $x_{ij} &lt; 0$</td>
<td>$r_{ij} = \frac{x_{ij}^-}{x_j^*}$, $x_{ij} &lt; 0$</td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>$r_{ij} = 1 - \frac{x_j^* - x_{ij}}{x_j^* - x_j^-}$</td>
<td>$r_{ij} = 1 - \frac{x_{ij}^- - x_j^-}{x_j^* - x_j^-}$</td>
<td></td>
</tr>
<tr>
<td>Vector</td>
<td>$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}$</td>
<td>$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} \left(\frac{1}{x_{ij}}\right)^2}}$, $x_{ij} &lt; 0$</td>
<td></td>
</tr>
</tbody>
</table>

Where:

$$x_j^* = \max_i x_{ij},$$
$$x_j^- = \min_i x_{ij}.$$

Another reason may be the application of computer programs that greatly accelerates and facilitates the application of the MADM method, but, as a rule, interrupts the user’s insights into the intermediate results, or makes it superficial, reducing the chances of unacceptable deformation of empirical data.

Deformations of empirical values that are the result of normalization call into question the application of normalized values as a decision base.

### 5 Weight coefficients of attributes

The weight coefficients reflect the relative significance of the attribute in an ideal case, i.e. when the set of options is complete. Since this theoretical assumption is not usually fulfilled in practice, weighted values should also depend on empirical data, or the extent to which values of the attributes domain of the observed options are different from their potential domains.

For example, if the options among themselves differ slightly in a very significant attribute, then its weight coefficient is reduced; thereby preventing slight
differences in this attribute to determine the final choice and at the same time increasing the weight coefficients of one or more attributes according to which the differences between the options are very conspicuous. Thus, we increase their relative impact on the final choice.

Any change in the domain of the attribute value should not be accompanied by a change in its weighted coefficient. If, by switching on or switching off an option, there would be a significant change in the domain of the attribute, then it would be justified and desirable to correct the weights. But if the change is caused exclusively by the unit of measure in which we express objectively the same size, then the domain of the attribute does not actually change, therefore there is no valid reason for changing its weight coefficient.

The results of the MADM method are in favour of our position. Namely, if the attribute is measurable on the ratio scale, then, regardless of the selected unit of measure, we will obtain the same results only if the value of the weights remains unchanged. The same applies if the attribute is measurable on an interval scale and the MADM method uses LN normalization. But if the method uses SN or VN normalization, then its results would be affected by the change in the unit of measure of the interval measurable attribute. It is difficult to find rational arguments in which, in the first case when the attributes are measurable on the ratio scale, with the change of the unit of measure, we would support the application of the existing weight coefficients, while in the second case, when the attributes were interval measurable, we would advocate for their modification, but only if the MADM method is based on SN or VN normalization (evidence is given in: French, 1988, for the condition of independence from the measurement scale; Kahneman and Tversky, 2000, for the condition of independence from the attribute formulation; Pavličić, 2001, for the instability of the normalized values on the displayed empirical data transformation).

The same logic is applied in the case of qualitative attributes. By changing the scale 1, 2, 3, 4, 5, with the scale 1, 3, 5, 7, 9, at first glance, seems that the domain of the attribute has expanded from the interval to the interval, but essentially, only the numbers by which we show the same modalities of a qualitative attribute have changed.

However, it is a widely accepted viewpoint that we perceive the differences between the numbers on two scales as different (for example, differences 5-4 and 9-7 are not treated as equal differences between the two modalities of the same attribute) and therefore, by changing the scale, we would change the weight of the given attribute. In this case, we will probably agree that on the scales 1, 2, 3, 4, 5 and 6, 7, 8, 9, 10 the differences between the corresponding levels are equal, i.e. 5-4 = 10-9, etc., so we would use the same weight coefficient for a given attribute; on the other hand, for example, scales 1, 2, 3, 4, 5 and 10, 20, 30, 40, 50 will be
treated as different, so depending on their application, we will assign the different weights to the same attribute. In the first case, the scales are interconnected by a positive affine transformation: which causes changes in values, and therefore we will obtain inconsistent results using the same weights.

In the second case, the scales are interconnected with a linear function: \( y = 10x \) which does not affect the values of \( r_{ij}^p \) and \( r_{ij}^v \), so the results of the MADM method will remain the same only in the case of an unchanged weight of the given attribute.

Therefore, we think that the change in the unit of measure in which we express the values of the attribute should not be accompanied by changes in the weight coefficients.

A slightly different problem arises from the re-formatting of the attributes. The results of behavioral decision theory clearly show that our preferences are sensitive to changes in the “framework”.

For this reason it is possible that the weighting depends on the formulation of the attribute we opt for. For example, the rise in quality from 93% to 97% of the correct products seems to us less significant than the reduction of the scrap from 7% to 3%. But if we assign a smaller weight to the income form of this attribute than to the expenditure form, we will merely demonstrate our irrationality. In order to remove it, it is necessary to re-examine the differences in weights for both attribute formulations and try to determine its unique value.

Since we use MADM methods to support rational decision-making, in the phase of determining weight coefficients, we must endeavor to eliminate or at least minimize possible subjective errors, which include those caused by the effects of the “framework”.

The inconsistency of the selection of the MADM method does not occur due to the fixed attribute weights, but depends on the type of functional link between the two forms of attributes, the revenue or expense, and the type of normalization applied (BN, LN or VN).

If the MADM method uses LN normalization and if the expense and revenue form of one attribute are connected with the function \( X_j^+ + X_j^- = C \), then changing the attribute formulation does not cause a change of result only if the attribute weight remains the same.

Similarly, if the MADM method uses BN or VN normalization, and if among the attributes there is one whose forms are linked by function \( X_j^- = \frac{C}{X_j^+} \), where \( C = \text{const.} \), then for both formulations of the attribute we obtain consistent results only with unchanged weight coefficients.

These results clearly show that the inconsistency of the selection of the MADM method does not come from fixed weight coefficients.
Conclusion

Normalization processes are given little attention, and based on the information we have, the undesirable effects of normalization have not been seriously analyzed so far. We believe that the reason for that is the specific treatment that normalization procedures have in MADM methods; they are, namely, considered a contribution to the more objective evaluation of options.

Recent results show that normalization of empirical values significantly influences final choices and that one of the main causes of inconsistency of results is the method of multi-attributive decision-making. Analyzing the scales on which the values of the qualitative attributes (Likert scale) are measured, we can notice that their determination also influences the consistency of the selection of the MADM methods. MADM methods can only be applied provided that all attributes are measurable on the ratio scale. The weight coefficients reflect the relative significance of the attributes in an ideal case, i.e. when the options set is complete.

Since this theoretical assumption is not usually fulfilled in practice, weighted values should also depend on empirical data, or the extent to which values of the attributes domain of the observed options are different from their potential domains.

Any change in the domain of the attribute value should not be accompanied by a change in its weight coefficient. Similarly, the change in the unit of measure in which we express the values of the attribute should not be accompanied by changes in the weight coefficients. Since we use MADM methods to support rational decision-making, in the phase of determining weight coefficients, we must endeavor to eliminate or at least minimize possible subjective errors, which include those caused by the effects of the “framework”.

List of References


Резиме

У раду се анализирају основни узроци (не)конзистентности избора метода вишеатрибутивног одлукивања: нормализација података, тежински коефицијенти и примјена Ликертових скала за мјерење ква­литативних атрибута. Нормализовани подаци у методама вишеатри­бутивног одлукивања представљају замјену за субјективне оцјене атрибута од стране доносилаца одлука. Пошто их израчунавамо на основу математичких трансформација емпиријских података, стиче се утисак да су избори засновани на нормализованим вриједностима „објективни”. Због тога се анализира осјетљивост резултата до сада бавила искључиво утицајима тежинских коефицијената на конечне изборе, док је могући утицај нормализације у потпуности занемарен; при томе су, по свему судећи, деформације изазване нормализацијом података приписиване утицајима тежинских коефицијената и
(In)consistency of the Selection of the Method of Multiplicative Decision-Making

BSc Katarina Račić et al.

their inevitable subjectivism. Our intention is to highlight the deformations of empirical values which are the result of normalization and which question the application of normalized values as a basis for decision-making. In addition, the (in)consistency of the choice of the method of multi-attribute decision-making affects changes in the manner of measurement and formulation of attributes.

Keywords: multi-attribute decision-making, normalization of data, weight coefficients, Likert scale.