Visualizing the Inner-Workings of TOPSIS

R. Susmaga[a] **and I. Szczech** [a;*]

[a]Institute of Computing Science, Poznan University of Technology, Poland

CONTEXT AND MOTIVATIONS. The use of Multi-Criteria Decision Making (MCDM) is widespread across a wide range of industries. MCDM techniques support decision makers in dealing with complex real-world issues by, among others, assessing and ranking alternatives described by multiple criteria. From variety of MCDM methods tackling the task of ranking alternatives from the most preferred to the least preferred, a commonly chosen approach is TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [1]. TOPSIS calculates distances between the considered alternatives and two predefined ones, namely the ideal and the anti-ideal, and creates a ranking of the alternatives according to a chosen aggregation of these distances. Over the decades, numerous versions and modifications of the method have been proposed, leaving, however, its core based on calculating and aggregating distances unchanged.

The bulk of the research on TOPSIS is focused on practical use cases and different ways of criteria weighting and normalization, lacking attempts to formally describe the systematic relations between the properties of alternatives and the effects of aggregations. As a result, aggregations are compared on a use case basis rather than globally, i.e. with respect to the space of all possible alternatives. Finally, no approaches exist that are capable of visualizing such general, dataset-independent properties of aggregations. Below, we briefly present our recent paper [2] that has addressed those issues by formalizing and visualizing the inner workings of TOPSIS.

DATASET-INDEPENDENT ANALYSES. The input data for the method is represented in a $m \times n$ matrix of values, usually called a decision matrix, with m being the number of considered alternatives and n the number of criteria characterizing each alternative. An exemplary decision matrix X containing four alternatives (students) described by three criteria (final grades obtained in three subjects) is depicted in Fig. 1A. The data in the decision matrix usually undergo some normalization and weighting first. Next, TOPSIS determines two reference points: ideal and anti-ideal, and calculates distances from each alternative's representation to one or to both of those points. Finally, the alternatives are ranked according to an assumed function that aggregates distances between the alternatives and the reference points.

To ensure most general, i.e dataset-independent, analyses of TOPSIS, our approach does not focus on any particular decision matrix, but considers sets of all possible alternative representations, defining them as three consecutive spaces, i.e. exhaustive sets (Fig. 1): *criteria space* (domain space of the original alternatives), *utility space* (rescaled criteria space to unify the intervals and the preference types of all criteria), and *MSD-space* (newly proposed space).

Within this presentation, we assume no additional criteria normalization and no criteria weighting.

* Corresponding Author. Email: izabela.szczech@cs.put.poznan.pl.

TOPSIS AGGREGATIONS IN MSD-SPACE. To measure the distance of an alternative's representation $\mathbf{u}$ in the utility space to the ideal point (denoted as $\mathbf{1}$) or the anti-ideal point ($\mathbf{0}$) we use a rescaled Euclidean distance ($\delta_2^{01}(\cdot)$) that is simply a Euclidean distance divided by $\sqrt{n}$, to make the maximal distance n-independent. This ensures that our results and visualizations are easily comparable regardless of the number of criteria. The three commonly considered aggregations are denoted by I, A and R, standing for the distance to the ideal point, distance to the anti-ideal point, and the relative distance, respectively. They are defined as:

$$\mathsf{I}(\mathbf{u}) = 1 - \delta_2^{01}(\mathbf{u},\mathbf{1}), \qquad \mathsf{A}(\mathbf{u}) = \delta_2^{01}(\mathbf{u},\mathbf{0}),$$

$$\mathsf{R}(\mathbf{u}) = \frac{\delta_2^{01}(\mathbf{u},\mathbf{0})}{\delta_2^{01}(\mathbf{u},\mathbf{1}) + \delta_2^{01}(\mathbf{u},\mathbf{0})}.$$

Let $\overline{\mathbf{u}} = [mean(\mathbf{u}), mean(\mathbf{u}), ..., mean(\mathbf{u})]$ be the vector consisting of repeated alternative's mean. Employing the fact that for every $\mathbf{u} \in US$ vectors $\overline{\mathbf{u}} - \mathbf{0}$ and $\mathbf{u} - \overline{\mathbf{u}}$ as well as $\mathbf{u} - \overline{\mathbf{u}}$ and $\mathbf{1} - \overline{\mathbf{u}}$ are orthogonal, one can apply the Pythagorean theorem to relate these vectors (Fig. 2A). Moreover, as shown in our work [2], the lengths of the above-mentioned vectors can be expressed as follows: $\delta_2^{01}(\overline{\mathbf{u}},\mathbf{0}) = mean(\mathbf{u})$; $\delta_2^{01}(\overline{\mathbf{u}},\mathbf{1}) = 1 - mean(\mathbf{u})$ and $\delta_2^{01}(\mathbf{u},\overline{\mathbf{u}}) = std(\mathbf{u})$. These characteristics of US allowed us to formulate the *IA-MSD property*.

Definition 1 (IA-MSD Property)**.**

$$\delta_2^{01}(\mathbf{u},\mathbf{0}) = \sqrt{mean(\mathbf{u})^2 + std(\mathbf{u})^2},$$
$$\delta_2^{01}(\mathbf{u},\mathbf{1}) = \sqrt{(1 - mean(\mathbf{u}))^2 + std(\mathbf{u})^2}.$$

This interesting dependency between the distances of an alternative to the ideal and anti-ideal points, inspired us to introduce a new space called *MSD-space*, which uses the alternatives' means (M) and standard deviations (SD) as its coordinates (see Figs 1D; 2B,C) .

Definition 2 (MSD). MSD-space $=\{[mean(\mathbf{u}), std(\mathbf{u})] | \mathbf{u} \in US\}$.

Given any $\mathbf{u} \in US$, the IA-MSD property makes it possible to express all the aggregations with $mean(\mathbf{u})$ and $std(\mathbf{u})$:

$$\mathsf{I}(\mathbf{u}) = 1 - \sqrt{(1 - mean(\mathbf{u}))^2 + std(\mathbf{u})^2},$$

$$\mathsf{A}(\mathbf{u}) = \sqrt{mean(\mathbf{u})^2 + std(\mathbf{u})^2},$$

$$\mathsf{R}(\mathbf{u}) = \frac{\sqrt{mean(\mathbf{u})^2 + std(\mathbf{u})^2}}{\sqrt{(1 - mean(\mathbf{u}))^2 + std(\mathbf{u})^2} + \sqrt{mean(\mathbf{u})^2 + std(\mathbf{u})^2}}.$$

As a result, each point in MSD-space representing a particular alternative can be color-coded with respect to the values of a particular aggregation function (see Fig. 2C for aggregation $\mathsf{R}(\mathbf{u})$). This allows swift visual analysis and comparison of different aggregations: e.g., identification of the isoline characteristics, possible to obtain values, position of reference points.

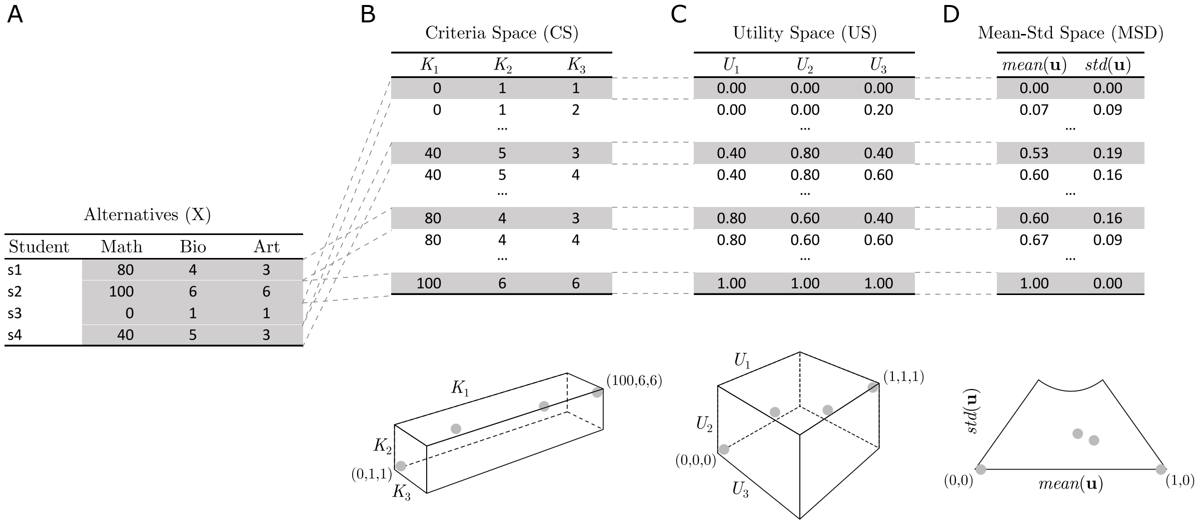

Figure 1. A schematic representation of alternatives in different spaces. (A) The original dataset (decision matrix) describing four students (alternatives) using final grades in three subjects (criteria). (B) The alternatives depicted as a subset of the criteria space, i.e., the space of all possible alternatives within the given criteria. (C) The alternatives presented as a subset of utility space, the re-scaled equivalent of criteria space. (D) The alternatives represented in the MSD-space, defined by the mean and standard deviation of the utilities assigned to the alternatives.

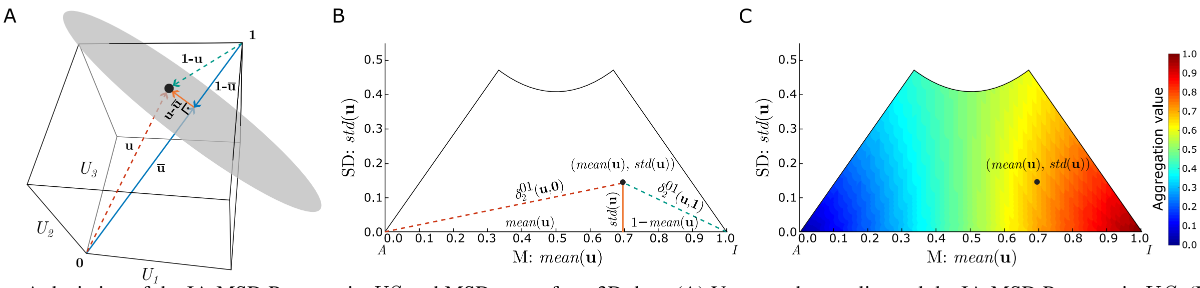

Figure 2. A depiction of the IA-MSD Property in US and MSD-space for a 3D data. (A) Vector orthogonality and the IA-MSD Property in US. (B) Vector orthogonality and the IA-MSD Property in MSD-space. The re-scaled δ_2^{01} lengths of vectors $\bar{\mathbf{u}}$ and $\mathbf{u} - \bar{\mathbf{u}}$ in (A) correspond to $mean(\mathbf{u})$ and $std(\mathbf{u})$ in (B). (C) Color encoding of the aggregation $\mathsf{R}(\mathbf{u})$, with blue representing the least preferred and red the most preferred values.

Exemplary application. Trade-offs between $mean(\mathbf{u})$ and $std(\mathbf{u})$ summarized in Table 1 and visualized in Fig. 2(C) show how carefully designed changes to the criteria may positively influence the final result. For example, under $\mathsf{R}(\mathbf{u})$, with $mean(\mathbf{u}) > 0.5$, an alternative can improve its rating just by lowering its standard deviation of utilities, even when the mean remains unchanged. This is clearly demonstrated by isolines of $\mathsf{R}(\mathbf{u})$ and may be additionally exemplified as follows.

The alternative depicted as the point in Fig. 2 is described by $[0.50, 0.75, 0.85]$ in US (A) and $[0.70, 0.15]$ in MSD-space (B,C). To increase its rating, one may simply increase the value of one utility while retaining the values of all the other ones (effectively increasing their mean). But such a ploy may be not easy, and only increasing one utility while decreasing another (enough to preserve the mean) may prove possible. Notice that methods that rate alternatives by the mean of utilities will not change their results after such 'compensatory' modifications to the utilities. TOPSIS, however, will update its rating even then: after 0.5 is increased to 0.6, while 0.85 is decreased to 0.75, the mean stays the same (0.70), but the standard deviation drops from 0.15 to 0.07. In result, the point moves straight down, towards higher isolines of $\mathsf{R}(\mathbf{u})$ (see Fig. 2(B,C)). Since the standard deviation is of type 'cost' for $mean(\mathbf{u}) > 0.5$ (see Table 1), the rating of the alternative actually grows.

Note that the effect could be opposite in other regions the MSD-space, so opposite modifications to the utilities (i.e. ones that actually increase the standard deviation) might be required.

Table 1. Preference-related interplay of $mean(\mathbf{u})$ and $std(\mathbf{u})$

aggregation	$mean(\mathbf{u})$	$std(\mathbf{u})$
$\mathsf{I}(\mathbf{u})$	gain	cost
$\mathsf{A}(\mathbf{u})$	gain	gain
$\mathsf{R}(\mathbf{u})$	gain	$mean(\mathbf{u}) < 0.5$: gain $mean(\mathbf{u}) = 0.5$: neutrality $mean(\mathbf{u}) > 0.5$: cost

Conclusions. Investigating the algebraic, dataset-independent aspects of TOPSIS, we have demonstrated that the alternative's ratings calculated as their distances to ideal and anti-ideal points can be actually expressed with the mean value ($mean(\mathbf{u})$) and the standard deviation ($std(\mathbf{u})$) of their utilities. The two easily interpretable features directly influence the final results, and provide a 2D space that can be productively visualized. The space is thus a practical tool that can be used to: inform decision makers about the properties of alternatives, highlight the consequences of using particular aggregation, and potentially suggest actions that will improve alternatives' ratings.

Acknowledgments. Research partly funded by the National Science Centre, Poland, grant number: 2022/47/D/ST6/01770.

References

[1] Ching-Lai Hwang and Kwangsun Yoon, *Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, 1981.

[2] R. Susmaga, I. Szczech, P. Zielniewicz, and D. Brzezinski, 'MSD-space: Visualizing the inner-workings of TOPSIS aggregations', *European Journal of Operational Research*, **308**(1), 229–242, (2023).