VAT/iVAT for Preference Data

Thomas A. Runkler
Siemens AG
Otto–Hahn–Ring 6
81739 Munich, Germany
Thomas.Runkler@siemens.com

I. INTRODUCTION

Given a set of n options, a fuzzy preference relation is specified by an $n \times n$ fuzzy preference matrix

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \quad (1)$$

where each matrix element p_{ij} quantifies the degree of preference of option i over option j, $i, j = 1, \ldots, n$. Here we consider *additive* fuzzy preference matrices with $p_{ij} \in [0, 1]$, where $p_{ij} = 0$ indicates absolutely no preference, $p_{ij} = 1$ indicates complete preference, and $p_{ij} = 1/2$ indicates equivalence, so the elements of the main diagonal are $p_{ii} = 1/2$ for all $i = 1, \ldots, n$. Moreover, we consider *reciprocal* additive fuzzy preference matrices, where $p_{ij} + p_{ji} = 1$ for all $i, j = 1, \ldots, n$.

A straighforward way to convert such a reciprocal additive fuzzy preference matrix into a symmetric dissimilarity matrix is by

$$d_{ij} = d_{ji} = \max(p_{ij}, p_{ji}) - 0.5 \quad (2)$$

II. GOOD EXAMPLE

Consider the preference matrix

$$P_1 = \begin{pmatrix} 0.5 & 1 & 1 & 1 \\ 0 & 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 & 0.5 \end{pmatrix} \quad (3)$$

Here, we prefer the first option over all other three, and all other three options are considered equal. This can be visualized using the preference graph shown in Fig. 1, where a directed edge from node i to j corresponds to $p_{ij} = 1, p_{ji} = 0$, and an undirected edge between nodes i and j corresponds to $p_{ij} = p_{ji} = 0.5$. Running VAT and iVAT on P_1 yields the results shown in Fig. 2, with the index vector $i = (1, 2, 3, 4)$. VAT correctly identifies 1 as one cluster and the other three as the other cluster, but iVAT fails.

If we apply eq. (2) to P_1, then we obtain the dissimilarity matrix

$$D_1 = \begin{pmatrix} 0 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \end{pmatrix} \quad (4)$$

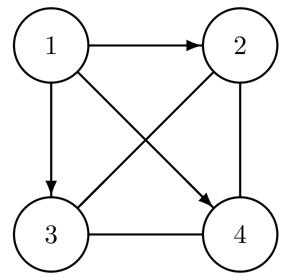

Fig. 1. Preference graph for P_1.

Fig. 2. VAT (left) and iVAT (right) images for P_1.

The VAT and iVAT images for D_1 are shown in Fig. 3, with the index vector $i = (2, 3, 4, 1)$, so options 2, 3, and 4 are in one cluster, and option 1 is in another cluster, as expected.

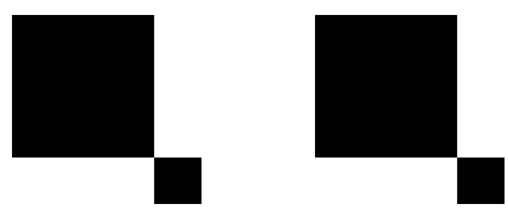

Fig. 3. VAT (left) and iVAT (right) images for D_1.

III. BAD EXAMPLES

Now imagine a soccer tournament with four teams playing against each other. All matches were ties, except team 1 won over team 3, and team 2 won over team 4, so we expect two clusters: the winner teams 1 and 2, and the loser teams 3 and 4. For this case we construct the preference matrix

$$P_2 = \begin{pmatrix} 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \\ 0 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \end{pmatrix} \quad (5)$$

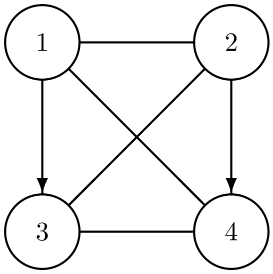

Fig. 4. Preference graph for P_2.

Fig. 5. VAT (left) and iVAT (right) images for P_2.

Fig. 6. VAT (left) and iVAT (right) images for D_2.

visualized in Fig. 4. Applying VAT and iVAT directly to P_2 yields the results shown in Fig. 5, with the index vector $i = (1, 2, 3, 4)$. Neither VAT nor iVAT yields the expected result.

For P_2 eq. (2) yields

$$D_2 = \begin{pmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \end{pmatrix} \tag{6}$$

The VAT and iVAT images for D_2 are shown in Fig. 6, with the index vector $i = (3, 2, 1, 4)$. Again, neither VAT nor iVAT yields the expected result.

Now imagine the same situation but with the reverse result for the match between teams 1 and 3, so team 3 won over team 1 and team 2 won over team 4. Again, we expect two clusters, this time with the winner teams 2 and 3, and the loser teams 1 and 4. This can be expressed by the preference matrix

$$P_3 = \begin{pmatrix} 0.5 & 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \\ 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \end{pmatrix} \tag{7}$$

visualized in Fig. 7. Applying VAT and iVAT directly to P_3 yields exactly the same results as for P_2 (shown in Fig. 5), but now with the index vector $i = (3, 2, 1, 4)$, so again not the expected result.

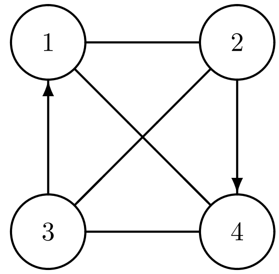

Fig. 7. Preference graph for P_3.

Notice again the difference between P_2 and P_3. However, for P_3 eq. (2) yields the same dissimilarity matrix as for P_2, so

$$P_2 \neq P_3, \quad D_2 = D_3 \tag{8}$$

and so for D_3 VAT and iVAT will yield the same images as for D_2 (shown in Fig. 6), even with the same index vectors. So, again, neither VAT nor iVAT yields the expected result, and moreover we obtain the same VAT and iVAT images for different cluster structures

IV. Conclusions and Open Issues

The two "bad" examples show that "conventional" VAT or iVAT does not yield images that appropriately reflect the preference structure.

I see two ways to overcome this:

1) We may replace (2) by a different formula to convert P to D. As a first requirement this formula needs to yield different dissimilarity matrices for P_2 and P_3, and as a second requirement VAT or iVAT should produce "good" (and different) images for both cases.
2) We may modify VAT and/or iVAT to deal directly with preference matrices. This may require a different sorting scheme for the matrix elements. Since preference matrices are usually non–symmetric, this feature also has to be taken into account. In a more general setup we may even want to consider the non–reciprocal case, where $p_{ij} = p_{ji} \neq 0.5$, so we have to consider both p_{ij} and p_{ji}.

What do you think about this? Comments and suggestions are more than welcome.