On Clustering the criteria in an MCDA

Raymond Bisdorff

University of Luxembourg, FSTC/CSC

Paris, October, 2008

Motivation A bipolar-valued correlation index Bipolar-valued clustering from $\widetilde{T}$ Conclusion

Outline

Motivation A bipolar-valued correlation index Bipolar-valued clustering from $\widetilde{T}$ Conclusion

Motivation

Example (Ronda decision problem)

A family, staying during their holidays in Ronda (Andalusia), is planning the next day's activity.

The set of alternatives

Identifier	Name	Comment
ant	Antequerra	An afternoon excursion to Antequerra and surroundings.
ard	Ardales	An afternoon excursion to Ardales and El Chorro.
be	beach	Sun, fun and more.
crd	Cordoba	A whole day visit by car to Cordoba.
dn	fa niente	Doing nothing.
lw	long walk	A whole day hiking.
mal	Malaga	A whole day visit by car to Malaga.
sev	Sevilla	A whole day visit by car to Sevilla.
sw	short walk	Less than a half day hiking.

Motivation A bipolar-valued correlation index Bipolar-valued clustering from $\widetilde{T}$ Conclusion

Example (Ronda example – continued)

The family members agree to measure their preferences with respect to the following set of criteria:

The family of criteria

Identifier	Name	Comment
cult	*Cultural Interest*	Andalusian heritage.
dis	*Distance*	Minutes by car to go to and come back from the activity.
food	*Food Quality*	Quality of the expected food opportunities.
sun	*Sun, Fun, & more*	No comment.
phy	*Physical Investment*	Contribution to physical health care.
rel	*Relaxation*	Anti-stress support.
tour	*Tourist Attraction*	How many stars in the guide ?

Example (Ronda example – continued)

The common evaluation of the performances of the nine alternatives on all the criteria results in the performance table shown here:

The performance table

Criteria	ant	ard	be	crd	dn	lw	mal	sev	sw
cult	7	3	0	10	0	0	5	10	0
dis	120	100	90	360	0	90	240	240	0
phy	3	7	0	5	0	10	5	5	5
rel	1	5	8	3	10	5	3	3	6
food	8	10	4	8	10	1	8	10	1
sun	0	3	10	3	1	3	8	5	5
tour	5	7	3	10	0	8	10	10	5

Example (Ronda example – continued)

- All performances on the qualitative criteria are marked on a same ordinal scale going from 0 (lowest) to 10 (highest).
- On the quantitative *Distance* criterion (to be minimized), the required travel time to go to and return from the activity is marked in minutes.
- In order to model only effective preferences, an indifference threshold of 1 point and a preference threshold of 2 points is put on the qualitative performance measures.
- On the *Distance* criterion, an indifference threshold of 20 min., and a preference threshold of 45 min. is considered.
- Furthermore, a difference of more than two hours ($\geqslant 121$ min.) to attend the activity's place is considered raising a veto.

Example (Ronda example – continued)

Preference discrimination thresholds

Criterion	Preference direction	Discrimination thresholds indifference	preference	veto
cult	max	1 pt	2 pts	-
dis	min	20 min.	45 min.	121 min.
food	max	1 pt	2 pts	-
sun	max	1 pt	2 pts	-
phy	max	1 pt	2 pts	-
rel	max	1 pt	2 pts	-
tour	max	1 pt	2 pts	-

The individual criteria each reflect one or the other member's preferential point of view. Therefore they are judged equi-significant for the best action to be eventually chosen.

Ronda example – continued

- How do the criteria express their preferential view point on the set of alternatives ?
- For instance the *Tourist Attraction* criterion appears to be in its preferential judgments somehow positively correlated with both the *Cultural Interest* and the *Food Quality* criteria.
- It is also apparent that the *Distance* criterion is somehow negatively correlated to these latter criteria.
- How can we explore and illustrate these intuitions ?

Notations

- We consider a finite set A of n alternatives and denote by x and y any two alternatives.
- We consider also a set F of outranking criteria denoted by variables i or j,
- with $k = 0, 1, \ldots$ discrimination thresholds.
- The performance of an alternative x on criterion i is denoted by x_i.

Example: Rubis discrimination thresholds

The four discrimination thresholds we may observe, for instance, on each criterion i in the Rubis choice method are:

- "weak preference" wp_i $(0 < wp_i)$,
- "preference" p_i $(wp_i \leqslant p_i)$,
- "weak veto" wv_i $(p_i < wv_i)$, and
- "veto" v_i $(wv_i \leqslant v_i)$.

Example: Rubis discrimination thresholds

Each performance difference $(x_i - y_i)$ may thus be classified into one and only one of the following nine categories:

($\ggg$) "*veto against $x \leqslant y$*"	$\Leftrightarrow$	$v_i \leqslant (x_i - y_i)$
($\gg$) "*weak veto against $x \leqslant y$*"	$\Leftrightarrow$	$wv_i \leqslant (x_i - y_i) < v_i$
($>$) "*x better than y*"	$\Leftrightarrow$	$p_i \leqslant (x_i - y_i) < wv_i$
($\geqslant$) "*x better than or equal y*"	$\Leftrightarrow$	$wp_i \leqslant (x_i - y_i) < p_i$
($=$) "*x indifferent to y*"	$\Leftrightarrow$	$-wp_i < (x_i - y_i) < wp_i$
($\leqslant$) "*x worse than or indifferent to y*"	$\Leftrightarrow$	$-p_i < (x_i - y_i) \leqslant -wp_i$
($<$) "*x worse than y*"	$\Leftrightarrow$	$-wp_i < (x_i - y_i) \leqslant -p_i$
($\ll$) "*weak veto against $x \geqslant y$*"	$\Leftrightarrow$	$-v_i < (x_i - y_i) \leqslant -wp_i$
($\lll$) "*veto against $x \geqslant y$*"	$\Leftrightarrow$	$(x_i - y_i) \leqslant -v_i$

Homogenous semiorders

In general, let us consider on each criterion i, supporting a set of discrimination thresholds p_i^k $(r = 1, .., k)$ such that $0 < p_i^1 \leqslant \ldots \leqslant p_i^k$, the Kendall vector (Degenne 1972) gathering the classification of all possible differences $(x_i - y_i)$ into one of the following $2k + 1$ categories:

$$(x_i - y_i) \in \begin{cases} (>_k) & \text{if} \quad p_i^k \leqslant (x_i - y_i) \\ (>_r) & \text{if} \quad p_i^r \leqslant (x_i - y_i) < p_i^{r+1}, \text{ for } r = 1, \ldots k-1 \\ (=) & \text{if} \quad -p_i^1 < (x_i - y_i) < p_i^1 \\ (<_r) & \text{if} \quad -p_i^{r+1} < (x_i - y_i) \leqslant -p_i^r, \text{ for } r = 1, \ldots k-1 \\ (<_k) & \text{if} \quad (x_i - y_i) \leqslant -p_i^k \end{cases} \tag{1}$$

A bipolar-valued ordinal correlation index

- Considering criteria i and j, we say that x and y are *concordantly* (resp. *discordantly*) compared if $(x_i - y_i)$ and $(x_j - y_j)$ are classified into the same category (resp. different categories) on both criteria.
- There are $n(n-1)$ distinct ordered pairs of performances and each pair (x, y) is thus either concordantly or discordantly classified.
- Denoting by S_{ij} the number c_{ij} of concordantly classified minus the number d_{ij} of discordantly classified ordered pairs, the *ordinal criteria correlation index* $\widetilde{T}$ is defined on $F \times F$ as

$$\tilde{T}(i,j) = \frac{c_{ij} - d_{ij}}{c_{ij} + d_{ij}} = \frac{S_{ij}}{n(n-1)}.$$

A bipolar-valued correlation index

Property (1)

The ordinal criteria correlation index $\tilde{T}$ is symmetrically valued in the rational bipolar credibility domain $[-1, 1]$.

Proof.

- If $d_{ij} = 0$ (resp. $c_{ij} = 0$), $\widetilde{T}(i,j) = 1.0$ (resp. -1.0).
- If $\widetilde{T}(i,j) > 0$ (resp. < 0) both criteria are more *similar than dissimilar* (resp. *dissimilar than similar*) in their preferential judgments. When $\widetilde{T}(i,j) = 0$, no conclusion can be drawn.
- A performance difference $(x_i - y_i)$ is classified in one and only one category.
- The category of $(x_i - y_i)$ corresponds bijectively to a unique symmetric category of $(y_i - x_i)$.

□

A bipolar-valued correlation index

Property (2)

Let Δ_ϵ denote the smallest observable difference in a given performance table. If i and j admit single preference thresholds $p^1_i, p^1_j \leqslant \Delta_\epsilon$ and we don't observe ties in the performance table, then $\widetilde{T}(i,j)$ is identical with the classical rank correlation index τ of Kendall (1938).

Proof.

- Either $(>_1)$: $(x_i - y_i) \geqslant \Delta_\epsilon$, or $(<_1)$: $(x_i - y_i) \leqslant \Delta_\epsilon$.
- Let p_{ij} be the number of pairs (x, y) in $A \times A$ such that conjointly $(x_i - y_i) \geqslant \Delta_\epsilon$ and $(x_j - y_j) \geqslant \Delta_\epsilon$:

$$\widetilde{T}(i,j) = (2 \times \frac{2p_{ij}}{n(n-1)}) - 1, \quad \forall (i,j) \in F \times F,$$

i.e. Kendall's original τ definition.

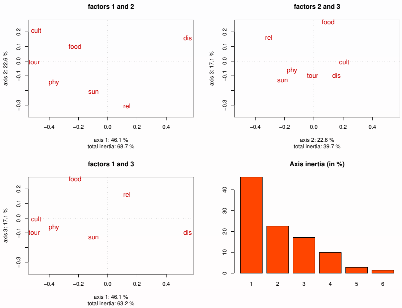
□

Example (The Ronda example – continued)

The ordinal criteria correlation table

$\widetilde{T}$	dis	food	phy	rel	sun	tour
cult	-0.83	0.00	-0.17	-0.78	-0.39	$+0.28$
dis		-0.78	-0.83	-0.61	-0.67	-0.83
food			-0.39	-0.22	-0.56	-0.17
phy				-0.17	-0.28	$+0.33$
rel					-0.17	-0.50
sun						0.00

Example (The Ronda example: Principal Component Analysis)



A bipolar-valued similarity digraph

- $\widetilde{\mathrm{T}}$ represents a bipolar-valued characteristic denotation of the propositional statement *"criteria i and j express similar preferential statements on A"*.
- We consider indeed this statement to be more or less validated if both criteria are concordant on a *majority* of pairwise comparisons and discordant on a minority ones.
- $\widetilde{\mathrm{T}}$ is characterising a bipolar-valued *similarity* graph, we denote by $\widetilde{\mathrm{S}}(F, \widetilde{\mathrm{T}})$ or $\widetilde{\mathrm{S}}$ for short.
- Following from the logical denotation of the bipolar valuation, we say that there is an arc between i and j if $\widetilde{\mathrm{T}}(i,j) > 0$.

A bipolar-valued similarity digraph

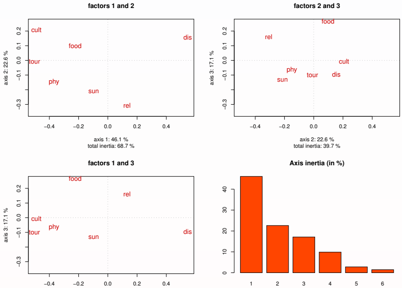
- In general, we may associate a crisp graph $S(F, T)$ with $\widetilde{\mathrm{S}}$, where $T = \{(i,j) | \widetilde{\mathrm{T}}(i,) > 0\}$.
- All properties of S are canonically transferred to $\widetilde{\mathrm{S}}$.
- For instance, S is a symmetric digraph, so is $\widetilde{\mathrm{S}}$.
- Similarly, a clique C in $\widetilde{\mathrm{S}}$ is a subset of criteria such that for all i and j in C, we have $\widetilde{\mathrm{T}}(i,j) \geqslant 0$.

Example (The Ronda example – continued)

The criteria similarity graph in the Ronda example contains only three edges:

- between *Physical Investment* and *Tourist Attraction* ($\widetilde{\mathrm{T}}$(phy,tour) $= 0.33$),
- between *Tourist Attraction* and *Cultural Interest* ($\widetilde{\mathrm{T}}$(tour,cult) $= 0.28$), and
- the weak (or potential) similarity between criteria *Food* and *Cultural Interest* ($\widetilde{\mathrm{T}}$(food,cult) $= 0.0$).
- Notice that the similarity relation is not transitive.

Example (The Ronda example: Principal Component Analysis)



Maximal bipolar-valued cliques

What we are looking for are maximal cliques, i.e. subsets C of criteria which verify both the following properties:

1. *Internal stability*:
   all criteria in C are similar, i.e. the subgraph $(C, \widetilde{\mathrm{T}}_{|C})$ is a clique;
2. *External stability*:
   if a criteria i is not in C, there must exist a criteria j in C such that $\widetilde{\mathrm{T}}(i,j) < 0$ and $\widetilde{\mathrm{T}}(j,i) < 0$.

Bipolar-valued maximal cliques

Definition (Stabilities' credibility)

For any $C \in F$, we denote by $\Delta^{int}(C)$ (resp. $\Delta^{ext}(C)$) its credibility of being internally (resp. externally) stable:

$$\Delta^{int}(C) = \begin{cases} 1.0 & \text{if} \quad |C| = 1, \\ \min_{i \in C} \min_{j \in C}^{j \neq i} \left(\widetilde{\mathrm{T}}(i,j)\right) & \text{otherwise.} \end{cases}$$

$$\Delta^{ext}(C) = \begin{cases} 1.0 & \text{if} \quad C = F, \\ \min_{i \in F}^{i \notin C} \max_{j \in C} \left(- \widetilde{\mathrm{T}}(i,j)\right) & \text{otherwise.} \end{cases}$$

Bipolar-valued maximal cliques

Property (3)

A subset C of criteria is a maximal clique of the similarity graph $\widetilde{\mathrm{S}} \equiv (F, \widetilde{\mathrm{T}})$ if and only if both $\Delta^{int}(C) \geqslant 0$ and $\Delta^{ext}(C) > 0$.

Proof.

- $\Delta^{int}(C) \geqslant 0$ implies that $(C, \widetilde{\mathrm{T}}_{|C})$ is a clique
- $\Delta^{ext}(C) > 0$ implies that, for any criterion i not in C, there exists at least one criterion j in C such that $\widetilde{\mathrm{T}}(i,j) < 0$.

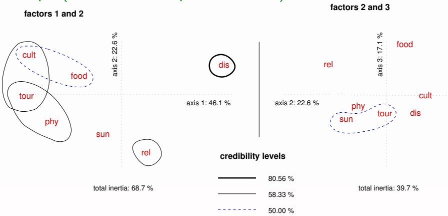
□

Example (The Ronda example – continued)

Clustering the criteria

Maximal cliques	credibility level (in%[3])	stability external	stability internal
{dis}	80.56	+0.083	+1.00
{rel}	58.33	+0.17	+1.00
{phy,tour}	58.33	+0.17	+0.33
{tour,cult}	58.33	+0.17	+0.28
{sun,tour}	50.00	+0.28	0.0
{cult,food}	50.00	+0.17	0.0

Example (The Ronda example – continued)



The bipolar-valued criteria clusters

Concluding Remarks

In this communication we have presented:

- a generalisation of Kendall's rank correlation τ measure to homogenous semiorders;
- a bipolar ordinal criteria correlation index;
- a graphical illustration of oppositions and agreements between criteria with the help of a PCA;
- a bipolar-valued criteria similarity digraph;
- a bipolar-valued clustering of the criteria.

References I

Kendall M G (1938)
A New Measure of Rank Correlation.
Biometrica (OUP) 30:81–86

Kendall M G (1970)
Rank Correlation Methods. Charles Griffin (Edition: Fourth)

Degenne A (1972)
Techniques ordinales en analyse des données statistique.
Hachette, Paris

Benzecri JP (1976)
L'analyse des données – Tome 2: *L'Analyse des correspondances*. Dunod Paris

References II

Bisdorff R (2002)
Logical Foundation of Multicriteria Preference Aggregation.
In: Bouyssou D et al (eds) *Essay in Aiding Decisions with Multiple Criteria*. Kluwer Academic Publishers 379–403

Bisdorff R (2002)
Electre-like clustering from a pairwise fuzzy proximity index.
Eur Jour Oper Res 138:320–331

Bidorff R, Meyer P, Roubens M (2008)
RUBIS: a bipolar-valued outranking method for the choice problem. *4OR* 6(2):143–165