Deriving rankings from incomplete preference information: A comparison of different approaches

Rudolf Vetschera
University of Vienna

Universidad Complutense de Madrid
October 30, 2014
• Motivation: Incomplete information models
• Research questions
• Models
• Computational experiments
• Conclusions
• Decision models contain many parameters (probabilities, attribute weights etc.)
• Parameters are often not know precisely
  – Lack of information
  – Elicitation too complex for decision makers
  – Different opinions in group decisions...
• Models needed to make decisions in the presence of incompletely specified parameters
Additive multi-attribute utility with unknown weights:

\[ u(X) = \sum_{k} w_k u_k(x_k) \]

Note: most concepts can also be applied to

- other uncertain parameters (e.g. partial value functions)
- other preference models (e.g. outranking models)
- other domains (e.g. group decisions)
Forms of incomplete information

- **Intervals:**
  weight is between....
  \[ w_k \leq w_k \leq \overline{w_k} \]

- **Rankings:**
  attribute \( k \) is more important than attribute \( m \)
  \[ w_m \leq w_k \]

- **Ratios:**
  attribute \( k \) is at least twice as important as \( m \)
  \[ 2w_m \leq w_k \]

- **Comparison of alternatives:**
  \( A_i \) is better than \( A_j \)
  \[ \sum_k w_k u_k(a_{ik}) \geq \sum_k w_k u_k(a_{jk}) \]

  In general: Linear constraints on \( w_k \)
Admissible parameters

$W$ = set of parameters which fulfill all constraints

$\sum_k w_k \leq 1$

No information: all parameter vectors fulfilling scaling conditions are admissible

More information $\rightarrow$ additional constraints $\rightarrow W$ becomes smaller
Decisions with incomplete information

**Approaches**

**Dominance:**
Establish relations that hold for all possible parameters

- Kmietowicz/Pearman 1984
- Kirkwood/Sarin 1985
- ROR: Greco et al 2008

**Single parameter:**
Identify one "best" parameter vector

- Srinivasan/Shocker 1973
- UTA: Jacquet-Lagreze/Siskos 1982
- Representative value functions: Greco et al. 2011

**Volume-based:**
Relative size of regions in parameter space

- Domain criterion: Starr 1962
- Charnetski/Soland 1978
- VIP: Climaco/Dias 2000
Dominance based approach

A is necessarily better than B iff
\[ \forall w \in W : u(A, w) > u(B, w) \]

C is possibly better than D iff
\[ \exists w \in W : u(C, w) > u(D, w) \]

Necessarily better is usually **incomplete** relation on set of alternatives
Dominance based approach: LP models

**Necessarily better**

\[
\max z = \sum_{k} w_k (b_k - a_k)
\]
\[
\text{s.t.}
\]
\[
w \in W
\]
Optimal \( z < 0 \):
\( A \) necessarily better than \( B \)

**Possibly better**

\[
\max z = \sum_{k} w_k (a_k - b_k)
\]
\[
\text{s.t.}
\]
\[
w \in W
\]
Optimal \( z > 0 \):
\( A \) possibly better than \( B \)
Single parameter approach

"Representative" parameter vector (center of W)
Example: constraints from pairwise comparison of alternatives

\[
\begin{align*}
\text{max } z \\
\text{s.t.} \\
\sum_k w_k (a_{ik} - a_{jk}) - z & \geq 0 \quad \forall i, j : A_i > A_j \\
z & \leq 0
\end{align*}
\]
Volume based approach

\[ p(A < B) = \frac{\text{Vol}(W_2)}{\text{Vol}(W)} \]

\[ p(A > B) = \frac{\text{Vol}(W_1)}{\text{Vol}(W)} \]
Results from volume-based methods (SMAA)

- **Rank acceptability index:**
  Probability $r_{ik}$ that alternative $A_i$ obtains rank $k$

- **Pairwise winning index:**
  Probability $p_{ij}$ that alternative $A_i$ is preferred to $A_j$
Decisions with incomplete information

**Approaches**

**Dominance:**
Establish relations that hold for all possible parameters

- Kmietowicz/Pearman 1984
- Kirkwood/Sarin 1985
- ROR: Greco et al 2008

Usually incomplete relation

**Single parameter:**
Identify one "best" parameter vector

- Srinivasan/Shocker 1973
- UTA: Jacquet-Lagreze/Siskos 1982
- Representative value functions: Greco et al. 2011

Complete order relation

**Volume-based:**
Relative size of regions in parameter space

- Domain criterion: Starr 1962
- Charnetski/Soland 1978
- VIP: Climaco/Dias 2000

Probabilistic information about ranks and relations

Complete order
• How to derive a complete order relation among alternatives from the probabilistic information obtained by volume-based approaches?
• How are relations obtained in such way different from those obtained by other methods?
Assignment problem

\[
\sum_{k=1}^{N_{\text{alt}}} x_{ik} = 1 \quad \forall \ i
\]

each alternative is assigned to one rank

\[
\sum_{i=1}^{N_{\text{alt}}} x_{ik} = 1 \quad \forall \ k
\]

to each rank, one alternative is assigned

\[x_{ik} \in \{0, 1\}\]

\[x_{ik} : \text{Alternative } A_i \text{ is assigned to rank } k\]
Objective functions

Average probability of assignments

$$\max \sum_{i, k : x_{ik} = 1} r_{ik} = \sum_{i=1}^{N_{alt}} \sum_{k=1}^{N_{alt}} r_{ik} x_{ik}$$

Joint probability

$$\max \prod_{i, k : x_{ik} = 1} r_{ik} = \sum_{i=1}^{N_{alt}} \sum_{k=1}^{N_{alt}} \log(r_{ik}) x_{ik}$$

Minimum probability of assignment

$$\max z$$

$$z \leq r_{ik} + (1 - x_{ik}) \quad \forall i, k$$
Construct complete order relation:

Complete and asymmetric
\[ y_{ij} + y_{ji} = 1 \quad \forall \; i \neq j \]

Irreflexive
\[ y_{ii} = 0 \quad \forall \; i \]

Transitive
\[ y_{ij} \geq y_{ik} + y_{kj} - 1.5 \quad \forall \; k \neq i, j \]

\[ y_{ik} : \text{Alternative } A_i \text{ is preferred to } A_j \]
Rank from assignment model

\[ R_i = \sum_k k x_{ik} \]

Rank from relation model

\[ R_i = 1 + \sum_j y_{ji} \]
Computational study

Generate problem
- Marginal utilities
- "True" weights

Generate information levels (pairwise comparisons)

Perform SMAA

Solve models and benchmarks

Rank correlation to "true" ranking

All levels

N

Y
Computational study

- Additive value function
- Unknown weights
- Incomplete information provided via pairwise comparisons of alternatives
  - Neighboring alternatives in true ranking (sorted by utility difference)
  - Neighboring alternatives by numbers (1-2, 3-4 etc., then 2-3, 4-5)
- Results: Rank correlation to "true" ranking of alternatives
- Benchmarks:
  - Barycenter approach (average of all weights in simulation)
  - Distance based approach: LP model to find parameter vector in center of polyhedron defined by comparisons (similar to UTA)
Information levels

**Alternatives**

- $A_1: u(A_1)$
- $A_2: u(A_2)$
- $A_3: u(A_3)$
- $A_{n-1}: u(A_{n-1})$
- $A_n: u(A_n)$

**Method 1: by numbers**

1. $A_1: A_2$
2. $A_2: A_3$
3. $A_3: A_4$
4. $A_4: A_5$

**Method 2: by preference**

1. $A_{[1]}: u(A_{[1]})$
2. $A_{[2]}: u(A_{[2]})$
3. $A_{[3]}: u(A_{[3]})$

Largest diff

$u(A_{[1]}) > u(A_{[2]}) > \ldots > u(A_{[n]})$
Measuring information

No information: unit simplex

More Info (W)

True weights
Compared models

• **Distance** (Benchmark)
• **Barycenter:** Average of all admissible weight vectors generated during simulation

**Rank acceptability indices:**
- **RankSum:** Objective sum of probabilities
- **RankProd:** Objective joint probability
- **RankMM:** Objective minimum probability

**Pairwise winning indices** (relation)
- **RelSum:** Objective sum of probabilities
- **RelProd:** Objective joint probability
- **RelMM:** Objective minimum probability
• Nested linear regressions
• Dependent variable: correlation between obtained ranking and “true ranking”
• Independent variables:
  – Problem characteristics (M0):
    • number of alternatives $N_{Alt}$
    • number of criteria $N_{Crit}$
  – Information provided (M1):
    Volume of admissible parameter set relative to set of all possible parameters ($Volume$)
  – Method used (M2)
  – Interaction effects (M3)
<table>
<thead>
<tr>
<th></th>
<th>M0</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>*** 0.8268</td>
<td>*** 0.9870</td>
<td>*** 0.9808</td>
<td>*** 0.9803</td>
</tr>
<tr>
<td>NAlt=9</td>
<td>*** 0.0032</td>
<td>*** -0.0351</td>
<td>*** -0.0351</td>
<td>*** -0.0351</td>
</tr>
<tr>
<td>NAlt=12</td>
<td>*** 0.0189</td>
<td>*** -0.0421</td>
<td>*** -0.0421</td>
<td>*** -0.0421</td>
</tr>
<tr>
<td>NAlt=15</td>
<td>*** 0.0302</td>
<td>*** -0.0458</td>
<td>*** -0.0458</td>
<td>*** -0.0459</td>
</tr>
<tr>
<td>NCrit=5</td>
<td>*** -0.0623</td>
<td>*** -0.0856</td>
<td>*** -0.0856</td>
<td>*** -0.0856</td>
</tr>
<tr>
<td>NCrit=7</td>
<td>*** -0.1005</td>
<td>*** -0.1300</td>
<td>*** -0.1300</td>
<td>*** -0.1300</td>
</tr>
<tr>
<td>Volume</td>
<td>*** -0.3530</td>
<td>*** -0.3530</td>
<td>*** -0.3530</td>
<td>*** -0.3510</td>
</tr>
<tr>
<td>Bary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RankSum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RankProd</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RankMM</td>
<td></td>
<td>*** -0.0067</td>
<td></td>
<td>0.0016</td>
</tr>
<tr>
<td>RelSum</td>
<td></td>
<td>*** 0.0132</td>
<td>*** 0.0101</td>
<td></td>
</tr>
<tr>
<td>RelProd</td>
<td></td>
<td>*** 0.0132</td>
<td>*** 0.0101</td>
<td></td>
</tr>
<tr>
<td>RelMM</td>
<td></td>
<td>*** 0.0132</td>
<td>*** 0.0100</td>
<td></td>
</tr>
<tr>
<td>Vol:Bary</td>
<td>*** 0.0118</td>
<td>*** 0.0089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vol:RankSum</td>
<td>*** 0.0030</td>
<td>*** 0.0069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vol:RankProd</td>
<td>*** 0.0023</td>
<td>*** 0.0066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vol:RankMM</td>
<td>*** -0.0154</td>
<td>*** -0.0171</td>
<td>*** -0.0331</td>
<td></td>
</tr>
<tr>
<td>Vol:RelSum</td>
<td>*** 0.0125</td>
<td>*** 0.0125</td>
<td>*** 0.0126</td>
<td></td>
</tr>
<tr>
<td>Vol:RelProd</td>
<td>*** 0.0125</td>
<td>*** 0.0125</td>
<td>*** 0.0126</td>
<td></td>
</tr>
<tr>
<td>Vol:RelMM</td>
<td>*** 0.0114</td>
<td>*** 0.0154</td>
<td>*** -0.0171</td>
<td>*** -0.0331</td>
</tr>
<tr>
<td>adj. R2</td>
<td>0.0524</td>
<td>0.3181</td>
<td>0.3195</td>
<td>0.3202</td>
</tr>
</tbody>
</table>
Interaction effects method and volume

15 alternatives, 3 criteria

15 alternatives, 7 criteria

Correlation: Difference to distance-based

Volume
"Returns to information"
Information in one experiment
Information anomalies

True parameter vector

Initial admissible set

Additional constraint

Revised estimate

Initial estimate
Effects on information anomalies

![Graph showing the effect of information anomalies on volume. The graph includes lines for Bary, RankSum, RankProd, RankMM, RelSum, RelProd, and RelMM.](image)
Conclusions

- Complete ranking can be obtained from stochastic indices
- Models are independent of underlying preference model
- Differences between methods are small, but significant
- Barycenter and methods based pairwise on winning indices usually outperform benchmark (distance-based)
- Max-min probability objective function gives worse results
- Information anomalies occur quite frequently
- Models based on rank acceptability indices and average/joint probability lead to fewer information anomalies
• Parallel application of methods
  – Mutual confirmation vs. differences
  – Differences could guide elicitation process

• Uncertainty about utilities (not just weights)

• Other preference models
Thank you for your attention!