Chapter 8
Selection of the Model

April 25, 2007

Vicenç Torra, Yasuo Narukawa (2007) Modeling Decisions: Information Fusion and Aggregation Operators, Springer. http://www.springer.com/3-540-68789-0; http://www.mdai.cat/ifao

Selection of the model

- Construction of the appropriate model
  1. Selection of the aggregation operator
  2. Determination of its parameters

- The process should take into account several factors
  - Mathematical properties.
  - Interpretability.
    * An user needs to understand the model.
  - Adaptability.
    * The environment changes w.r.t. time.
  - Simplicity.
    * Other factors being equal, the simpler the model, the better.

Selection of the model

- Parameter determination
  - Methods based on an expert.
    - The expert (almost) directly supply the required parameters
      - Analytic Hierarchy Process (AHP)
    - The expert supply some relevant information used for parameter determination
      - Parameter determination from orness or compensation
  - Methods based on data.
    - having preferences (or partial order) on examples
    - having intended outputs from examples

Methods based on an expert

Analytic Hierarchy Process

- Designed to derive ratio scales (Saaty)
  - Weights in OWA and WM are in ratio scales
- Based on pairwise comparison of the objects (e.g. criteria) M
  (AHP permits us to obtain weights even in the case of inconsistent matrices)
- Procedure:
  - Compute the principal eigenvector of the matrix M
  - Normalize the principal eigenvector
- Consistency Index:

$$CI = \frac{\lambda_{\max} - N}{N - 1}.$$

the principal eigenvalue of the matrix is $\lambda_{\max}$

OWA Weights from Orness

- An expert supplies the orness (compensation degree)

- **Example.** OWA with Yager quantifier ($Q_\alpha(x) = x^\alpha$), and orness equal to 0.2 (i.e., $\delta = 0.2$)
  - According to Proposition 7.28:
    $$orness(Q_\alpha) = \frac{1}{\alpha+1}.$$
  - Therefore, the problem is to find α s.t.
    $$\delta = \frac{1}{\alpha+1}.$$
  - So, $\alpha = (1-\delta)/\delta = (1-0.2)/0.2 = 4$.

- In general,
  - ... different weighting vectors with the same orness
    Example. $(0, 1/2, 1/2, 0)$ and $(1/4, 1/4, 1/4, 1/4)$
    $\rightarrow$ same orness, equal to 0.5

OWA Weights from Orness

- In this case, it is usual to add another restriction.
  E.g., maximize dispersion (entropy) of the weights
  - Formally,

$$
\begin{array}{l}
Maximize\ dispersion \\
Subject\ to \\
\qquad \delta = orness \\
\\
\qquad w\ is\ a\ weighting\ vector
\end{array}
$$

OWA Weights from Orness

- In this case, it is usual to add another restriction.
  E.g., maximize dispersion (entropy) of the weights
  - And still more formally,

$$\begin{aligned} &\text{Minimize } -\sum_{i=1}^N w_i \ln w_i \\ &\text{Subject to} \\ &\qquad \delta = \frac{1}{N-1}\sum_{i=1}^N (N-i) w_i \\ &\qquad \sum_{i=1}^N w_i = 1 \\ &\qquad w_i \geq 0 \end{aligned}$$

- Proposition 8.5. Gives expressions for the solution (single solution)
  An OWA with these weights corresponds to a ME-OWA (Maximum Entropy OWA).

Methods Based on Data

Expected Outcome

Formulation of the problem

- Examples defined by ($input$, $output$) pairs.

$$\begin{array}{cccc|c} a_1^1 & a_2^1 & \ldots & a_N^1 & b^1 \\ a_1^2 & a_2^2 & \ldots & a_N^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^M & a_2^M & \ldots & a_N^M & b^M \end{array}$$

- The best model:

$$\begin{array}{l} \text{Minimize } D_{\mathbb{C}}(P) = \sum_{j=1}^{M} (\mathbb{C}_P(a_1^j, \ldots, a_N^j) - b^j)^2 \\ \text{Subject to logical constraints on } \; P \end{array} \tag{1}$$

Formulation of the problem

- In the case of the weighted mean:

$$\text{Minimize } D_{WM}(\mathbf{p} = (p_1, \ldots, p_N)) = \sum_{j=1}^{M} (WM_{\mathbf{p}}(a_1^j, \ldots, a_N^j) - b^j)^2$$

$$\textit{Subject to}$$

$$\sum_{i=1}^{N} p_i = 1$$

$$p_i \geq 0$$

- WM: a quadratic program with linear inequality constraints
  $\longrightarrow$ algorithms to obtain an optimal solution

- Case of multiple solutions
  $\longrightarrow$ we can *add* e.g. "Maximize dispersion"

Formulation of the problem

- Example.

Student	ML	P	M	L	G	Subjective evaluation
s_1	0.8	0.9	0.8	0.1	0.1	0.7
s_2	0.7	0.6	0.9	0.2	0.3	0.6
s_3	0.7	0.7	0.7	0.2	0.6	0.6
s_4	0.6	0.9	0.9	0.4	0.4	0.8
s_5	0.8	0.6	0.3	0.9	0.9	0.8
s_6	0.2	0.4	0.2	0.8	0.1	0.3
s_7	0.1	0.2	0.4	0.1	0.2	0.1
s_8	0.3	0.3	0.3	0.8	0.3	0.4
s_9	0.5	0.2	0.1	0.2	0.1	0.3
s_{10}	0.8	0.2	0.2	0.5	0.1	0.5

- Solution for the WM (optimal):

  $p_{ML} = 0.4244$, $p_P = 0.4108$, $p_M = 0.0000$, $p_L = 0.1249$ and $p_G = 0.0399$.

- Solution for the OWA (optimal):

  $w_1 = 0.1245$, $w_2 = 0.6385$, $w_3 = 0.0531$, $w_4 = 0.0000$, $w_5 = 0.1839$

Methods Based on Data

Preferences

Formulation of the problem

- Example.

Student	ML	P	M	L	G	Subjective evaluation
s_1	0.8	0.9	0.8	0.1	0.1	3rd
s_2	0.7	0.6	0.9	0.2	0.3	4th
s_3	0.7	0.7	0.7	0.2	0.6	4th
s_4	0.6	0.9	0.9	0.4	0.4	1st
s_5	0.8	0.6	0.3	0.9	0.9	1st
s_6	0.2	0.4	0.2	0.8	0.1	8th
s_7	0.1	0.2	0.4	0.1	0.2	10th
s_8	0.3	0.3	0.3	0.8	0.3	7th
s_9	0.5	0.2	0.1	0.2	0.1	8th
s_{10}	0.8	0.2	0.2	0.5	0.1	6th

Formulation of the problem

- Formulation for the Weighted Mean:
  - S is defined by the pairs such that $s_r > s_t$ (student s_r is preferred to student s_t)
  - $y_{(s,t)}$ a variable (degree of violation of $s_r > s_t$)

$$
\begin{array}{l}
\text{Minimize } \sum_{(r,t)\in S} y_{(r,t)} \\
\text{Subject to} \\
\qquad \sum_{i=1}^{N} p_i(f^r(x_i) - f^t(x_i)) + \quad y_{(r,t)} > 0 \\
\qquad y_{(r,t)} \geq 0 \\
\qquad \sum_{i=1}^{N} p_i = 1 \\
\qquad p_i \geq 0
\end{array}
\qquad (2)
$$