

Multi-criteria hierarchical clustering

N.A.V. Doan, J. Rosenfeld, Y. De Smet

1 Definitions

- \mathcal{A} : the set of n alternatives $\mathcal{A} = \{a_1, a_2, \dots, a_n\}$ (notation: a_i or a_j , $i, j = 1, 2, \dots, n$)
- \mathcal{F} : the set of m criteria $\mathcal{F} = \{f_1, f_2, \dots, f_m\}$ (notation: f_k , $k = 1, 2, \dots, m$)
- \mathcal{R} : the set of l clusters $\mathcal{R} = \{r_1, r_2, \dots, r_l\}$ (notation: r_h , $h = 1, 2, \dots, l$)

2 Input data

- Evaluation table:

a	$f_1(\cdot)$	$f_2(\cdot)$	\dots	$f_k(\cdot)$	\dots	$f_m(\cdot)$
a_1	$f_1(a_1)$	$f_2(a_1)$	\dots	$f_k(a_1)$	\dots	$f_m(a_1)$
a_2	$f_1(a_2)$	$f_2(a_2)$	\dots	$f_k(a_2)$	\dots	$f_m(a_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
a_i	$f_1(a_i)$	$f_2(a_i)$	\dots	$f_k(a_i)$	\dots	$f_m(a_i)$
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
a_n	$f_1(a_n)$	$f_2(a_n)$	\dots	$f_k(a_n)$	\dots	$f_m(a_n)$
- \mathcal{W} : the set of m weights for the criteria: $\mathcal{W} = \{w_1, w_2, \dots, w_m\}$ (notation: w_k , $k = 1, 2, \dots, m$)
- Big M (large value, see Note 1)

3 Decision variables

- $c_{ih} = \begin{cases} 1 & \text{if } a_i \in r_h \\ 0 & \text{otherwise} \end{cases}$, $c_{ih} \in \{0, 1\}$
- r_{kh} : position of the cluster h on the criterion k , $f_k(r_h)$

4 Equations

$$\max z = \pi - \pi \tag{1}$$

$$\text{s.t.} \tag{2}$$

$$\phi(a_i) = \phi^+(a_i) - \phi^-(a_i) \tag{netflow} \tag{3}$$

$$\phi^+(a_i) = \frac{1}{n-1} \sum_{j=1, j \neq i}^n \sum_{k=1}^m w_k \beta_{ijk} \tag{positive flow} \tag{4}$$

$$\phi^-(a_i) = \frac{1}{n-1} \sum_{j=1, j \neq i}^n \sum_{k=1}^m w_k \beta_{jik} \tag{negative flow} \tag{5}$$

$$\beta_{ijk} \geq \frac{f_k(a_i) - f_k(a_j)}{M} \quad \forall i, j = 1, 2, \dots, n; \quad k = 1, 2, \dots, m \tag{linearization, see Note 2} \tag{6}$$

$$\beta_{ijk} < \frac{f_k(a_i) - f_k(a_j)}{M} + 1 \quad \forall i, j = 1, 2, \dots, n; \quad k = 1, 2, \dots, m \tag{7}$$

$$\beta_{ijk} \in \{0, 1\} \quad \forall i, j = 1, 2, \dots, n; \quad k = 1, 2, \dots, m \tag{8}$$

$$c_{ih} \in \{0, 1\} \quad \forall i = 1, 2, \dots, n; \quad h = 1, 2, \dots, l \tag{decision variables} \tag{9}$$

5 Notes

1. Big M chosen so that $\frac{f_k(a_i) - f_k(a_j)}{M} \in]-1; 1[$

2. $\beta_{ijk} = \begin{cases} 1 & \text{if } f_k(a_i) > f_k(a_j) \\ 0 & \text{otherwise} \end{cases}, \quad \beta_{ijk} \in \{0, 1\}$