Multi-criteria hierarchical clustering

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

1 Definitions

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

2 Input data

- Evaluation table:

a	$f_1(\cdot)$	$f_2(\cdot)$	$\ldots$	$f_j(\cdot)$	$\ldots$	$f_m(\cdot)$
a_1	$f_1(a_1)$	$f_2(a_1)$	$\ldots$	$f_j(a_1)$	$\ldots$	$f_m(a_1)$
a_2	$f_1(a_2)$	$f_2(a_2)$	$\ldots$	$f_j(a_2)$	$\ldots$	$f_m(a_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
a_i	$f_(a_i)$	$f_2(a_i)$	$\ldots$	$f_j(a_i)$	$\ldots$	$f_m(a_i)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
a_n	$f_1(a_n)$	$f_2(a_n)$	$\ldots$	$f_j(a_n)$	$\ldots$	$f_m(a_n)$

- $\mathcal{W}$: the set of m weights for the criteria: $\mathcal{W} = \{w_1, w_2, \ldots, w_m\}$ (notation: w_k, $k = 1, 2, \ldots, m$)
- Big M (large value, see Note 2)

3 Decision variables

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

4 Equations

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

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

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

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

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

$$\beta_{ijk} \geq \frac{f_k(a_i) - f_k(a_j)}{M} \qquad \text{(linearization Note 1)} \tag{6}$$

$$\beta_{ijk} < \frac{f_k(a_i) - f_k(a_j)}{M} + 1 \tag{7}$$

$$\beta_{ijk} \in \{0, 1\} \tag{8}$$

$$c_{ik} \in \{0, 1\} \qquad \text{(decision variables)} \tag{9}$$

5 Notes

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

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