101 熵值法 (entropy weight)

(1) 熵值法的基本思想

熵的概念源于热力学,是对系统状态不确定性的一种度量。在信息论中,信息是系统有序程度的一种度量,而熵是系统无序程度的一种度量,两者绝对值相等,但符号相反,根据此性质,可以利用多目标决策评价中各方案的固有信息,通过熵值法得到各个指标的信息熵,信息熵越小, 信息的无序度越低,其信息的效用值越大,指标权重越大。反之,信息熵越大,信息无序度越高,其信息的效用值越小,指标权重越小。熵权法的基本思想是: 人们在决策中获得信息的多少和质量是精度和可靠性大小的决定因素之一。

(2) 具体步骤

熵值法是一种根据各指标传输给决策者的信息量大小来确定各指标权数的方法。通过计算偏差度(差异性系数) $G_j = 1 - \mathbb{E}_j (1 \leq j \leq n)$ 从而有 G_j 越大,则指标越重要。

设:有 m 个方案,n 项评价指标, $^{x_{ij}}$ 为第 i 种方案、第 j 项指标的原始数据,对于给定的 j, $^{x_{ij}}$ 的差异越大,则该项指标对方案比较作用就越大:亦即该项指标包含和传输的信息就越多。信息的增加意味着熵的减少,熵可以用来度量这种信息量大小。

设: 评价对象的决策矩阵(original data)为 $X = (x_{ij})_{m \times n}$,用熵值法确定指标权数步骤如下。

① 计算第j项指标下,第i种方案指标值比重 p_{ij} :

$$p_{ij} = \frac{X_{ij}}{\sum_{i=1}^{m} X_{ij}} (i = 1, \dots, m, j = 1, \dots, n)$$

 (x_{ij}) 为经过规范化和无量纲化的数据)

② 计算第 j 项指标熵 E,:

$$E_j = -k \sum_{i=1}^{m} p_{ij} \ln p_{ij}$$
 $0 \le E_j \le 1 (j = 1, \dots, n)$ $k = 1 / \ln n$

(若
$$p_{ii} = 0$$
, 则规定 $p_{ii} \ln p_{ii} = 0$)

③ 计算第j项指标的偏差度(差异性系数) G_j :

$$G_i = 1 - \mathbb{E}_i (1 \le j \le n)$$
 (G_i 越大,则指标越重要)

④ 计算指标j的客观权重 w_i :

$$w_{j} = \frac{G_{j}}{\sum_{i=1}^{n} G_{j}} (1 \le j \le n)$$
 得: 指标的权向量 $W = (w_{1}, w_{2}, \dots, w_{n})$ 。

如果决策者对于指标 f_j 有偏好 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$,则可利用 w_j 来修正 λ_j ,得到较合理的权向量 $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_n^*)$,作为后面步骤中使用的权向量。

其中
$$\lambda^* = \frac{\lambda_j^* w_j}{\sum_{j=1}^n \lambda_j^* w_j} (1 \le j \le n)$$

(3) 熵权法评价

熵权法确定权重,其值是根据各指标所提供信息量的大小和波动情况确定的,并没有考虑评价指标的重要程度。单纯用熵权值解决综合评价问题显然不妥。同样,只用主观赋权值所得权值解决问题,也会因所定权值可能因主观人为因素太强而使结果失之偏颇。

在《矿山地质和环境评价中权值确定方法的探讨》中,提出了"综合权值"的概念,将主观赋权法的权值与客观赋权法的权值有机结合起来,弱化二者的不足。已知有 \mathbf{n} 个评价指标,按客观(objective)权值法求出第 \mathbf{j} 个评价指标的熵权 \mathbf{W}_{sj} ,按照主观(subjective)权值法得到这个指标的权值为 \mathbf{W}_{zj} ,二者综合,即可按下式求得这个评价指标的综合(synthesis)权值 \mathbf{W}_{oj} :

$$W_{oj} = W_{zj}W_{sj} / \sum_{i=1}^{m} W_{zj}W_{sj}$$

陈伟,吴群(2010)就将熵权法和层次分析法结合起来确定权重。