

综合评价中确定权重向量的几种方法比较

王靖¹, 张金锁²

(1. 天津大学 管理学院, 天津 300072; 2. 河北工业大学 管理学院, 天津 300130)

摘要: 以企业技术创新能力评价排序问题为实证范例, 对德尔菲法、层次分析法、熵值法和模糊聚类分析法测定评价指标权重的可信度做了比较研究, 目的是探索各自的适用范围和优缺点。文献通过理论分析特别是实证比较, 认为样本数据完整, 熵值法则较高; 如果缺乏样本数据, 则应采用 AHP 法, 可信度高于德尔菲法; 若样本中含有大量的模糊数据, 且同一层次指标个数较多, 应先采用模糊聚类分析法分类, 再用 AHP 法分别做各个子类的权重。

关键词: 综合评价; 技术创新; 权重向量; 层次分析法; 熵值法

中图分类号: F224.0 **文献标识码:** A

0 引言

近年来, 运用综合评价方法对众多样本做评价排序案例很多, 这一过程涉及 4 项关键工作: 1) 影响因素分析及评价指标体系的设计; 2) 样本数据的获取; 3) 各指标重要程度的辨识; 4) 综合评价方法。目前, 对指标权重的确定存在着不少主观随意性, 严重影响着评价结果的客观性。鉴于此, 本文选取 4 种方法对同一案例做了比较研究, 提出了各自适用的范围及综合运用方法。

1 实证案例

以河北省 532 家大中型工业企业技术创新能力问卷为样本, 根据国内外有代表性的观点, 将企业技术创新能力分解为: 创新决策能力、研究开发能力、生产能力、市场销售能力和组织管理能力 5 个方面^[1]。可用技术创新资源投入能力、创新实施能力、创新实现能力和创新管理能力 4 项一级指标表征。每项一级指标分为两个二级指标, 每项二级指标又细化为若干三级具体指标。为简化讨论, 本文仅取两级指标, 第二级的 8 项指标被分割成 4 个指标域, 如图 1 所示。

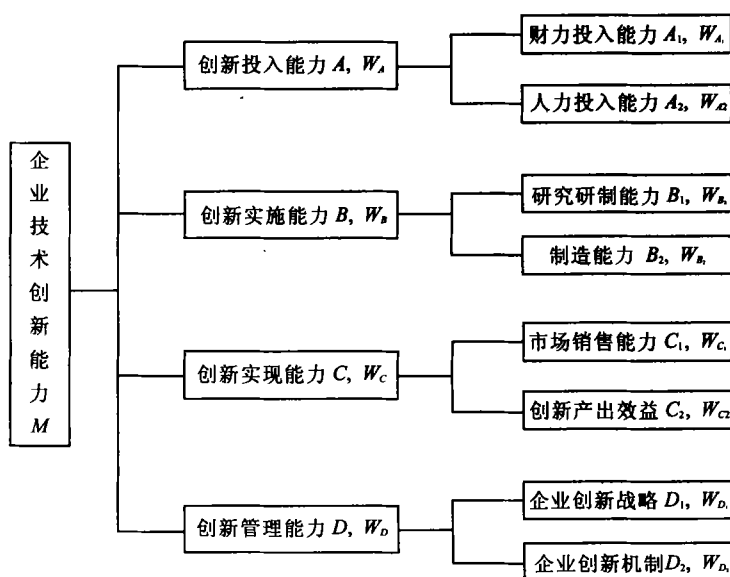


图 1 企业技术创新能力评价指标体系

收稿日期: 2000-11-24

作者简介: 王靖 (1971-), 女 (汉族), 硕士生。

案例综合评价系统的数学模型：设待评价系统由 m 个评价对象（样本）组成，用 n 个指标做综合评价，即论域为：

$$U = \{u_1, u_2, \dots, u_i, \dots, u_m\} \quad (i = 1, 2, \dots, m) \quad (1)$$

每一评价对象 u_i 由 n 个指标的数据（统计数据或问卷答案或专家打分）表征：

$$u_i = \{X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{in}\} \quad (j = 1, 2, \dots, n) \quad (2)$$

于是得到评价系统的初始数据矩阵 X ：

$$X = \{x_{ij}\}_{m \times n} \quad (3)$$

在本案例中，样本数 $m=532$ ，一级指标 $n_1=4$ ，二级指标分为 4 个指标域，每个指标域均由两项指标组成， $n_2=2$ 。现要根据数据矩阵 X 对 m 个样本分别做出综合评价，并做优劣排序。在综合评价中必然涉及指标的并合运算，因此，需要确定各指标的权重。

2 权重确定方法论

2.1 德尔菲法 (Delphi) 确定权重法

这是最常用的一种方法。它是依据若干专家的知识、智慧、经验、信息和价值观，对已拟出的评价指标进行分析、判断、权衡并赋予相应权值的一种调查法。一般需经过多轮匿名调查，在专家意见比较一致的基础上，经组织者对专家意见进行数据处理，检验专家意见的集中程度、离散程度和协调程度，达到要求之后，得到各评价指标的初始权重向量， $w^* = \{w_i^*\}_{1 \times n}$ 再对 w^* 做出归一化处理，获得各评价指标的权重向量：

$$w = \left\{ \frac{w_1^*}{\sum_{i=1}^n w_i^*}, \frac{w_2^*}{\sum_{i=1}^n w_i^*}, \dots, \frac{w_n^*}{\sum_{i=1}^n w_i^*} \right\} = \{w_1, w_2, \dots, w_n\} \quad (4)$$

鉴于读者都熟悉德尔菲法的具体做法，本文不再详述。本项研究聘请 12 位专家，经 3 轮调查，得到图 1 所示评价指标的权重，列于表 1。

2.2 层次分析 (AHP) 确定权重法

层次分析法也是大家熟悉的一种方法，该法的基本原理是将一个复杂的被评价系统，按其内在的逻辑关系，以评价指标（因素）为代表构成一个有序的层次结构，然后针对每一层的指标（或某一指标域），运用专家的知识、经验、信息和价值观，对同一层或同一域的指标进行两两比较对比，并按规定的标度值构造比较判断矩阵 $A = \{a_{ij}\}_{n \times n}$ ，再由组织者计算比较判断矩阵 A 的最大特征根 λ_{\max} ，并由 λ_{\max} 解特征方程：

$$AX = \lambda_{\max} X \quad (5)$$

得到对应 λ_{\max} 的特征向量 $X = \{X_1, X_2, \dots, X_n\}$ ，最后将特征向量归一化，得到各指标的权重向量：

$$w = \left\{ \frac{X_1}{\sum_{i=1}^n X_i}, \frac{X_2}{\sum_{i=1}^n X_i}, \dots, \frac{X_n}{\sum_{i=1}^n X_i} \right\} = \{w_1, w_2, \dots, w_n\} \quad (6)$$

在计算权重向量之前需对 A 作一致性检验。

本文按图 1 请 10 位专家分别构造一级指标的比较判断矩阵 M (4 阶)，二级指标的 4 组二阶比较判断矩阵 A, B, C, D 。经数据处理和检验，获得各级指标的权重，列于表 1。限于篇幅，这里省略了数据运算过程。

2.3 熵值确定权重法

2.3.1 熵与熵值函数

熵 (Entropy) 原是统计物理和热力学中的一个物理概念, 在热力学中熵指一个热力系统在热功转换过程中, 热能有效利用的程度, 一个热力系统的熵值大, 表示系统的能量可利用的程度低; 熵值小, 能量可利用的程度高. 在一个孤立热力系统中, 系统会自发的不可逆的向熵增方向转化, 一个开放的热力系统, 只有外部对系统做功 (输入能量), 其熵才会向熵减方向进行 (俗称负熵过程).

在统计物理中, 熵是分子运动无序度的度量, 熵值大, 表示系统分子运动的无序度高. 在孤立系统中, 分子运动的无序度会由低状态向高状态自发进行, 要想使系统由高无序状态向低无序状态转换, 必须有外力作用. 从微观角度, 系统的熵值可从分子排列方式的统计中得出. 设系统内有两种物质 (二元系统), 1 物质有 n_1 个分子, 2 物质有 n_2 个分子, 该系统的熵值可由波尔兹曼 (Boltzman) 公式计算:

$$E = k \ln \Omega$$

Ω 是系统中两种物质分子的微观排列方式, $\Omega = \frac{(n_1+n_2)!}{n_1! n_2!}$. 根据斯梯林公式 $\ln n! = n \ln n - n$, 则:

$$\begin{aligned} E &= k \ln \left[\frac{(n_1+n_2)!}{n_1! n_2!} \right] = k(n_1+n_2) \ln(n_1+n_2) - k(n_1 \ln n_1 + n_2 \ln n_2) \\ &= -k \left[n_1 \ln \frac{n_1}{n_1+n_2} + n_2 \ln \frac{n_2}{n_1+n_2} \right] \end{aligned} \quad (7)$$

E 是系统 (n_1+n_2) 个分子的总熵值, 除以分子总数, 便得到系统的单位熵值:

$$e = \frac{E}{n_1+n_2} = -k \left[\frac{n_1}{n_1+n_2} \ln \frac{n_1}{n_1+n_2} + \frac{n_2}{n_1+n_2} \ln \frac{n_2}{n_1+n_2} \right] \quad (8)$$

令 $y_1 = \frac{n_1}{n_1+n_2}$, $y_2 = \frac{n_2}{n_1+n_2}$ 分别为系统中 1 物质和 2 物质的占分率, 则系统的单位熵值为:

$$e = -k (y_1 \ln y_1 + y_2 \ln y_2)$$

扩展到多元 (m 元) 系统, 其单位熵值函数为:

$$e = -k \sum_{i=1}^m y_i \ln y_i \quad (9)$$

2.3.2 信息系统的熵值函数

在信息系统中的信息熵是信息无序度的度量, 信息熵越大, 信息的无序度越高, 其信息的效用值越小; 反之, 信息的熵越小, 信息的无序度越小, 信息的效用值越大.

在综合评价中, 运用信息熵评价所获系统信息的有序程度及信息的效用值是很自然的, 统计物理中的熵值函数形式对于信息系统应是一致的.

2.3.3 综合评价中信息熵值及权重函数

对于所讨论的系统 (公式(1), (2), (3)), 已获得 m 个样本的 n 个评价指标的初始数据矩阵 $X = \{X_{ij}\}_{m \times n}$, 由于各指标的量纲、数量级及指标优劣的取向均有很大差异, 故需对初始数据做标准化处理:

$$y_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}, \quad 0 \leq y_{ij} \leq 1 \quad (10)$$

由此得数据的标准化矩阵:

$$Y = \{y_{ij}\}_{m \times n} \quad (11)$$

根据 (9) 式可得 j 项指标的信息熵值

$$e_j = -k \sum_{i=1}^m y_{ij} \ln y_{ij} \quad (12)$$

式中常数 k 与系统的样本数 m 有关. 对于一个信息完全无序的系统, 有序度为零, 其熵值最大, $e=1$, m 个样本处于完全无序分布状态时, $y_{ij} = \frac{1}{m}$, 由 (12) 式得:

$$e = -k \sum_{i=1}^m \frac{1}{m} \ln \frac{1}{m} = k \sum_{i=1}^m \frac{1}{m} \ln m = k \ln m = 1$$

于是得到:

$$k = (\ln m)^{-1}, \quad 0 \leq e \leq 1$$

由于信息熵 e_j 可用来度量 j 项指标的信息 (指标的数据) 的效用价值, 当完全无序时, $e_j=1$, 此时, e_j 的信息 (也就是 j 指标的数据) 对综合评价的效用值为零. 因此, 某项指标的信息效用价值取决于该指标的信息熵 e_j 与 1 的差值 h_j

$$h_j = 1 - e_j \quad (13)$$

可见, 利用熵值法估算各指标的权重, 其本质是利用该指标信息的价值系数来计算的, 其价值系数越高, 对评价的重要性就越大 (或称对评价结果的贡献越大), 于是 j 项指标的权重为:

$$w_j = \frac{h_j}{\sum_{j=1}^n h_j} \quad (14)$$

熵值法是根据各指标所含信息有序度的差异性, 也就是信息的效用价值来确定该指标的权重. 为不致使本文篇幅过大, 现省略了所有中间处理过程, 在表 1 中直接给出了图 1 中各指标的权重.

表 1 企业技术创新能力评价指标权重系数表

指标 代号	A	B	C	D	A1	A2	B1	B2	C1	C2	D1	D2
德尔 菲法	0.213	0.206	0.305	0.276	0.416	0.584	0.814	0.186	0.342	0.658	0.277	0.723
	0.294*	0.285*	0.421*	-								
AHP 法	0.211	0.199	0.321	0.269	0.422	0.578	0.786	0.214	0.318	0.682	0.304	0.696
	0.289*	0.272*	0.439									
熵值法	0.258	0.241	0.501	-	0.373	0.627	0.787	0.213	0.214	0.786	-	-

注: 带 “*” 数据是将 D 项数据按比例分配给 A、B、C 的结果.

2.4 模糊聚类分析确定指标权重法

当若干评价指标具有模糊性时, 可采用模糊聚类分析法^[1], 对指标做模糊分类, 并给出分类的权重和排序. 其方法如下:

取评价指标体系为论域 $U = \{u_1, u_2, \dots, u_i, \dots, u_m\}$, 其中任一指标 u_i 由 n 个样本的隶属度表征,

即:

$$u_i = \{x_{i1} = \mu_1(u_i), x_{i2} = \mu_2(u_i), \dots, x_{ij} = \mu_j(u_i), \dots, x_{in} = \mu_n(u_i)\} \quad (15)$$

于是得到评价系统的初始数据矩阵 X :

$$X = \{x_{ij} = \mu_j(u_i)\}_{m \times n} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (16)$$

并进一步得到标准化数据矩阵 Y :

$$Y = \{y_{ij}\}_{m \times n}, \quad \text{其中}, \quad y_{ij} = \frac{x_{ij}}{\sum_{j=1}^n x_{ij}} \quad (17)$$

第二步, 列出论域 U 中所有指标的序偶 (u_i, u_k) , 采用数量积法计算所有序偶的模糊相似系数, 得到论域 U 上的一个模糊相似关系矩阵 \tilde{R} :

$$\tilde{R} = \{r_{u_i u_k}\} = \{r_{ik}\}_{m \times m} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{m1} \\ r_{21} & r_{22} & \dots & r_{m2} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mm} \end{bmatrix} \quad (18)$$

其中:

$$r_{ij} = \begin{cases} 1 & \text{当 } i=j \\ \frac{1}{M} \sum_{l=1}^n y_{il} \times y_{kl} & \text{当 } i \neq j \end{cases} \quad (19)$$

式中正常数 M 满足 $M \geq \max_{ik} (y_{ij} | i=1, 2, \dots, m; j=1, 2, \dots, n)$.

第3步, 计算的模糊等价矩阵, 使其具备自反性、对称性和传递性. 为此对 \tilde{R} 做自乘运算:

$\tilde{R}^2 = \tilde{R} \circ \tilde{R}$, $\tilde{R}^4 = \tilde{R}^2 \circ \tilde{R}^2$, 直到出现 $\tilde{R}^k = \tilde{R}^{2k}$, 这时, $\tilde{R}^* = \tilde{R}^k$, 为模糊等价关系矩阵.

第4步, 计算 \tilde{R}^* 在不同阈值 (λ) 下的截矩阵 (布尔矩阵), 对一系列截矩阵做分析可得到评价指标重要程度的分类. 一般情况下, 随 λ 由小到大变化, 对指标的分类由少到多, 为尽可能区分出各评价指标的重要程度, 应选取较大的 λ 值, 以便对评价指标的重要程度做分类和排序, 作为进一步确定各指标权重的基础.

本文选取河北省 5 个城市的企业技术创新能力调查数据, 对图 1 的一级指标做了模糊聚类分析, 结果见图 2.

3 方法比较

3.1 方法分类

上述 4 种方法, 从原理角度可分为 3 类: 德尔菲法与 AHP 法基本属于一类, 都是基于专家群体的知识、经验和价值判断, 只是 AHP 法对专家的主观判断进一步做了数学处理, 使之更科学, 但专家经验、知识的局限性并未消除. 熵值法是根据样本数据自身的特征做出的权重判断. 模糊聚类分析法是基于样本模糊数据的相似性, 对评价指标群体做出相对重要程度分类.

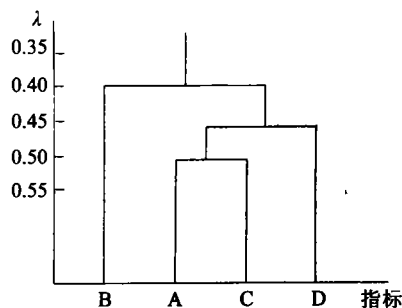


图 2 模糊聚类分析动态变化图

3.2 适用范围与优缺点比较

德尔菲法和 AHP 法与另两种方法比较,优点是:不需要具备样本数据,专家仅凭对评价指标内涵与外延的理解即可做出判断。因此,适用范围较广,特别对一些定性的模糊指标,仍可做出判断,且在判断过程中可以吸纳更多的信息。例如,对财力投入与人力投入的比较,专家结合中国企业既缺乏创新资金又缺乏创新人才的现状,德尔菲法给出的两项指标权重的差距比熵值法要小 28%,对其他指标也有类似情况(见表 1)。AHP 法与德尔菲法比较,适用范围相同,由于 AHP 法对各指标之间相对重要程度的分析更具逻辑性,刻划的更细,再加上数学处理,其可信度高于德尔菲法。这两种方法的缺点是:在一定程度上都存在主观性,如专家选择不当则可信度更低。

熵值法由于深刻地反映了指标信息熵值的效用价值,其给出的指标权值比德尔菲法和 AHP 法有较高的可信度,但它缺乏各指标之间的横向比较,又需要样本数据,在应用上受到限制,如本文由于缺乏“创新管理能力”的样本数据,而无法计算 D 指标(D1, D2)的权重。

模糊聚类分析法适用于模糊指标的重要程度分类,特别适用于同一层次有多项指标时。该法的缺点是只能给出指标分类的权重,而不能确定单项指标的权重,应在此基础上,结合采用其他方法做到底。

4 建议

如果有较完整的样本数据,应采用熵值法,并将其结果通过指标之间的横向比较做适当修正。如果缺乏样本数据,特别是含有大量定性指标时,建议采用 AHP 法,当含有大量模糊指标时,建议将模糊聚类分析法与德尔菲法结合使用。

参考文献:

- [1] 贺仲雄. 模糊数学及其应用 [M]. 天津: 天津科学技术出版社, 1982.
- [2] 高建. 中国企业技术创新分析 [M]. 北京: 清华大学出版社, 1999.
- [3] Negoita C V. Applications of Fuzzy Sets to Systems Analysis [J]. Applied Economics Letters, 1997, 4: 497-501.

Comparing Several Methods of Assuring Weight Vector in Synthetical Evaluation

WANG Jing¹, ZHANG Jin-suo²

(1. School of Management, Tianjin University, Tianjin 300072, China; 2. School of Management, Hebei University of Technology, Tianjin 300130, China)

Abstract: Based on the evaluation and sorting of enterprise's technology innovation ability, this paper compares the reliability of measuring weight of evaluating index by Delphi method, Analytic Hierarchy Process (AHP), Entropy method and Fuzzy Cluster Analysis, to research applicable area and advantage and disadvantage of each method. The paper draws conclusion through theoretical analysis especially practical compare that if complete sample datum can be available, the reliability of Entropy method is higher. If sample datum is scare, the reliability of AHP is higher than Delphi method. If there are a lot of fuzzy indexes among sample, we should adopt Fuzzy Cluster Analysis, then conclude the weights of indexes by AHP.

Key words: synthetical evaluation; technology innovation; weight vector; Analytic Hierarchy Process (AHP); Entropy method