

层次分析和主成分分析理论和实现代码

层次分析法 (Analytic Hierarchy Process, 简称AHP) 是对一些较为复杂、较为模糊的问题作出决策的简易方法, 它特别适用于那些难于完全定量分析的问题。它是美国运筹学家T. L. Saaty 教授于上世纪70年代初期提出的一种简便、灵活而又实用的多准则决策方法。

§ 1 层次分析法的基本原理与步骤

人们在进行社会的、经济的以及科学管理领域问题的系统分析中, 面临的常常是一个由相互关联、相互制约的众多因素构成的复杂而往往缺少定量数据的系统。层次分析法为这类问题的决策和排序提供了一种新的、简洁而实用的建模方法。

运用层次分析法建模, 大体上可按下面四个步骤进行:

- (i) 建立递阶层次结构模型;
- (ii) 构造出各层次中的所有判断矩阵;
- (iii) 层次单排序及一致性检验;
- (iv) 层次总排序及一致性检验。

下面分别说明这四个步骤的实现过程。

1.1 递阶层次结构的建立与特点

应用 AHP 分析决策问题时, 首先要把问题条理化、层次化, 构造出一个有层次的结构模型。在这个模型下, 复杂问题被分解为元素的组成部分。这些元素又按其属性及关系形成若干层次。上一层次元素作为准则对下一层次有关元素起支配作用。

这些层次可以分为三类:

(i) 最高层: 这一层次中只有一个元素, 一般它是分析问题的预定目标或理想结果, 因此也称为目标层。

(ii) 中间层: 这一层次中包含了为实现目标所涉及的中间环节, 它可以由若干个层次组成, 包括所需考虑的准则、子准则, 因此也称为准则层。

(iii) 最底层: 这一层次包括了为实现目标可供选择的各种措施、决策方案等, 因此也称为措施层或方案层。递阶层次结构中的层次数与问题的复杂程度及需要分析的详尽程度有关, 一般地层次数不受限制。每一层次中各元素所支配的元素一般不要超过9个。这是因为支配的元素过多会给两两比较判断带来困难。

下面结合一个实例来说明递阶层次结构的建立。

例1 假期旅游有 P_1 、 P_2 、 P_3 个旅游胜地供你选择, 试确定一个最佳地点。

在此问题中, 你会根据诸如景色、费用、居住、饮食和旅途条件等一些准则去反复比较3个候选地点。可以建立如图1的层次结构模型。

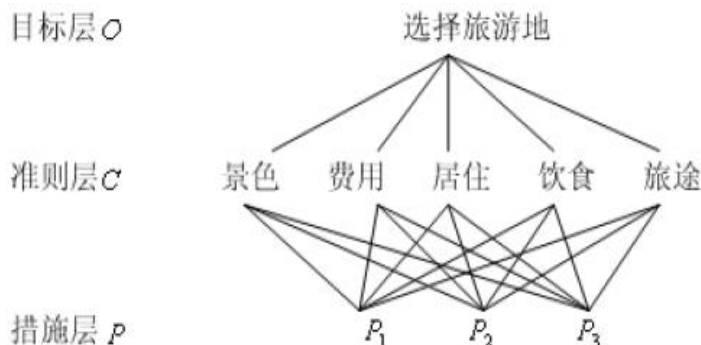


图1 层次结构模型

1.2 构造判断矩阵

层次结构反映了因素之间的关系,但准则层中的各准则在目标衡量中所占的比重并不一定相同,在决策者的心目中,它们各占有一定的比例。

在确定影响某因素的诸因子在该因素中所占的比重时,遇到的主要困难是这些比重常常不易量化。此外,当影响某因素的因子较多时,直接考虑各因子对该因素有多大程度的影响时,常常会因考虑不周全、顾此失彼而使决策者提出与他实际认为的重要性程度不相一致的数据,甚至有可能提出一组隐含矛盾的数据。为看清这一点,可作如下假设:将一块重为1千克的石块砸成 n 小块,你可以精确称出它们的重量,设为 w_1, \dots, w_{n-1} ,现在,请人估计这 n 小块的重量占总重量的比例(不能让他知道各小石块的重量),此人不仅很难给出精确的比值,而且完全可能因顾此失彼而提供彼此矛盾的数据。

设现在要比较 n 个因子 $\{x_1 \dots x_n\}$ 对某因素 Z 的影响大小,怎样比较才能提供可信的数据呢? Saaty 等人建议可以采取对因子进行两两比较建立成对比较矩阵的办法。即每次取两个因子 x_i 和 x_j , 以 a_{ij} 表示 x_i 和 x_j 对 Z 的影响大小之比。

定义 1 若矩阵 $A = (a_{ij})_{n \times n}$ 满足

$$(i) a_{ij} > 0, (ii) a_{ji} = \frac{1}{a_{ij}} \quad (i, j = 1, 2, \dots, n)$$

则称之为正互反矩阵(易见 $a_{ii} = 1, i = 1, \dots, n$)。

关于如何确定 a_{ij} 的值, Saaty 等建议引用数字 1~9 及其倒数作为标度。表 1 列出了 1~9 标度的含义:

表 1 标度的含义

标度	含 义
1	表示两个因素相比,具有相同重要性
3	表示两个因素相比,前者比后者稍重要
5	表示两个因素相比,前者比后者明显重要
7	表示两个因素相比,前者比后者强烈重要
9	表示两个因素相比,前者比后者极端重要
2, 4, 6, 8	表示上述相邻判断的中间值
倒数	若因素 i 与因素 j 的重要性之比为 a_{ij} , 那么因素 j 与因素 i 重要性之比为 $a_{ji} = 1/a_{ij}$ 。

从心理学观点来看,分级太多会超越人们的判断能力,既增加了作判断的难度,又容易因此而提供虚假数据。Saaty 等人还用实验方法比较了在各种不同标度下人们判断结果的正确性,实验结果也表明,采用 1~9 标度最为合适。

最后,应该指出,一般地作 $\frac{n(n-1)}{2}$ 次两两判断是必要的。有人认为把所有元素都和某个元素比较,即只作 $n-1$ 次比较就可以了。这种作法的弊病在于,任何一个判断的失误均可导致不合理的排序,而个别判断的失误对于难以定量的系统往往是难以避免的。进行 $\frac{n(n-1)}{2}$ 次比较可以提供更多的信息,通过各种不同角度的反复比较,从而导出一个合理的排序。

1.3 层次单排序及一致性检验

判断矩阵 A 对应于最大特征值 λ_{\max} 的特征向量 W ，经归一化后即为一层次相

应因素对于上一层次某因素相对重要性的排序权值，这一过程称为层次单排序。

上述构造成对比较判断矩阵的办法虽能减少其它因素的干扰，较客观地反映出一对因子影响力的差别。但综合全部比较结果时，其中难免包含一定程度的非一致性。

如果比较结果是前后完全一致的，则矩阵 A 的元素还应当满足：

$$a_{ij}a_{jk} = a_{ik}, \quad \forall i, j, k = 1, 2, \dots, n \quad (1)$$

定义 2 满足关系式 (1) 的正互反矩阵称为一致矩阵。

需要检验构造出来的 (正互反) 判断矩阵 A 是否严重地非一致，以便确定是否接受 A 。

定理 1 正互反矩阵 A 的最大特征根 λ_{\max} 必为正实数，其对应特征向量的所有分量均为正实数。 A 的其余特征值的模均严格小于 λ_{\max} 。

定理 2 若 A 为一致矩阵，则

(i) A 必为正互反矩阵。

(ii) A 的转置矩阵 A^T 也是一致矩阵。

(iii) A 的任意两行成比例，比例因子大于零，从而 $\text{rank}(A) = 1$ (同样， A 的任意两列也成比例)。

(iv) A 的最大特征值 $\lambda_{\max} = n$ ，其中 n 为矩阵 A 的阶。 A 的其余特征根均为零。

(v) 若 A 的最大特征值 λ_{\max} 对应的特征向量为 $W = (w_1, \dots, w_n)^T$ ，则 $a_{ij} = \frac{w_i}{w_j}$ ，

$\forall i, j = 1, 2, \dots, n$ ，即

$$A = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \dots & \frac{w_2}{w_n} \\ \dots & \dots & \dots & \dots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \dots & \frac{w_n}{w_n} \end{bmatrix}$$

定理 3 n 阶正互反矩阵 A 为一致矩阵当且仅当其最大特征根 $\lambda_{\max} = n$ ，且当正互反矩阵 A 非一致时，必有 $\lambda_{\max} > n$ 。

根据定理 3，我们可以由 λ_{\max} 是否等于 n 来检验判断矩阵 A 是否为一致矩阵。由

于特征根连续地依赖于 a_{ij} ，故 λ_{\max} 比 n 大得越多， A 的非一致性程度也就越严重，

λ_{\max} 对应的标准化特征向量也就越不能真实地反映出 $X = \{x_1, \dots, x_n\}$ 在对因素 Z 的影响中所占的比重。因此，对决策者提供的判断矩阵有必要作一次一致性检验，以决定是否接受它。

对判断矩阵的一致性检验的步骤如下：

(i) 计算一致性指标 CI

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

(ii) 查找相应的平均随机一致性指标 RI 。对 $n = 1, \dots, 9$ ，Saaty 给出了 RI 的值，如表 2 所示。

表 2 RI 的值

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

RI 的值是这样得到的，用随机方法构造 500 个样本矩阵：随机地从 1~9 及其倒数中抽取数字构造正互反矩阵，求得最大特征根的平均值 λ'_{\max} ，并定义

$$RI = \frac{\lambda'_{\max} - n}{n - 1}。$$

(iii) 计算一致性比例 CR

$$CR = \frac{CI}{RI}$$

当 $CR < 0.10$ 时，认为判断矩阵的一致性是可以接受的，否则应对判断矩阵作适当修正。

1.4 层次总排序及一致性检验

上面我们得到的是一组元素对其上一层中某元素的权重向量。我们最终要得到各元素，特别是最低层中各方案对于目标的排序权重，从而进行方案选择。总排序权重要自上而下地将单准则下的权重进行合成。

表 3 层次总排序合成表

层 A \ 层 B	A_1	A_2	...	A_m	B 层总排序权值
B_1	a_{11}	a_{12}	...	a_{1m}	$\sum_{j=1}^m b_{1j}a_j$
B_2	b_{21}	b_{22}	...	b_{2m}	$\sum_{j=1}^m b_{2j}a_j$
\vdots	\vdots
B_n	b_{n1}	b_{n2}	...	b_{nm}	$\sum_{j=1}^m b_{nj}a_j$

具体程序见附件 AHP_MATLAB.txt 中¹

¹ AHP_MATLAB.txt