第八讲、系统评价

彭勤科
系统工程研究所
E-mail: qkpeng@xjtu.edu.cn
Tel: 82667964

2020年6月3日

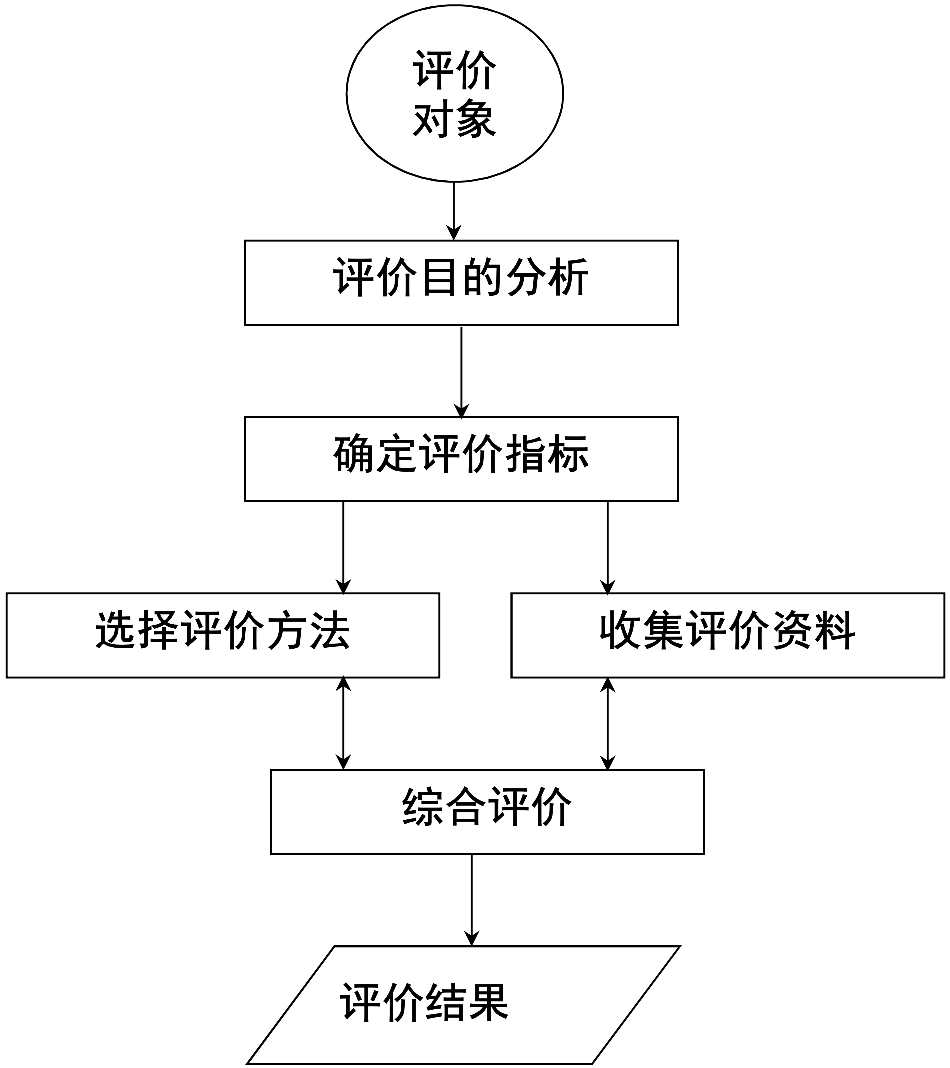
主要内容

- 系统评价概述
- 系统评价的步骤
- 系统评价方法
- 关联矩阵法
- 专家调查法
- 专家打分综合方法
- 模糊综合评价方法
- 层次分析法

系统评价概述

- 概念与意义
  - 意义
  - 系统评价的要素
    - 对象
    - 目的
    - 指标
    - 评价者
- 评价标准
  - 多指标评价
- 评价问题特点
  - 多指标、多因素
  - 评价信息的不完全
  - 多领域知识
- 专家往往不专

系统评价的步骤



系统评价方法

- 费用-效益分析法
- 关联矩阵法
- 关联树方法
- 层次分析法
- 模糊评价法

关联矩阵法

设m个评价指标$P_1,P_2,\cdots,P_m$，

每个评价指标对应的权值为$w_1,w_2,\cdots,w_m$；

设有n个评价对象$A_1,A_2,\cdots,A_n$，

评价对象$A_1,A_2,\cdots,A_n$关于评价指标P_j的评价值为$v_{1j},v_{2j},\cdots,v_{nj}$。

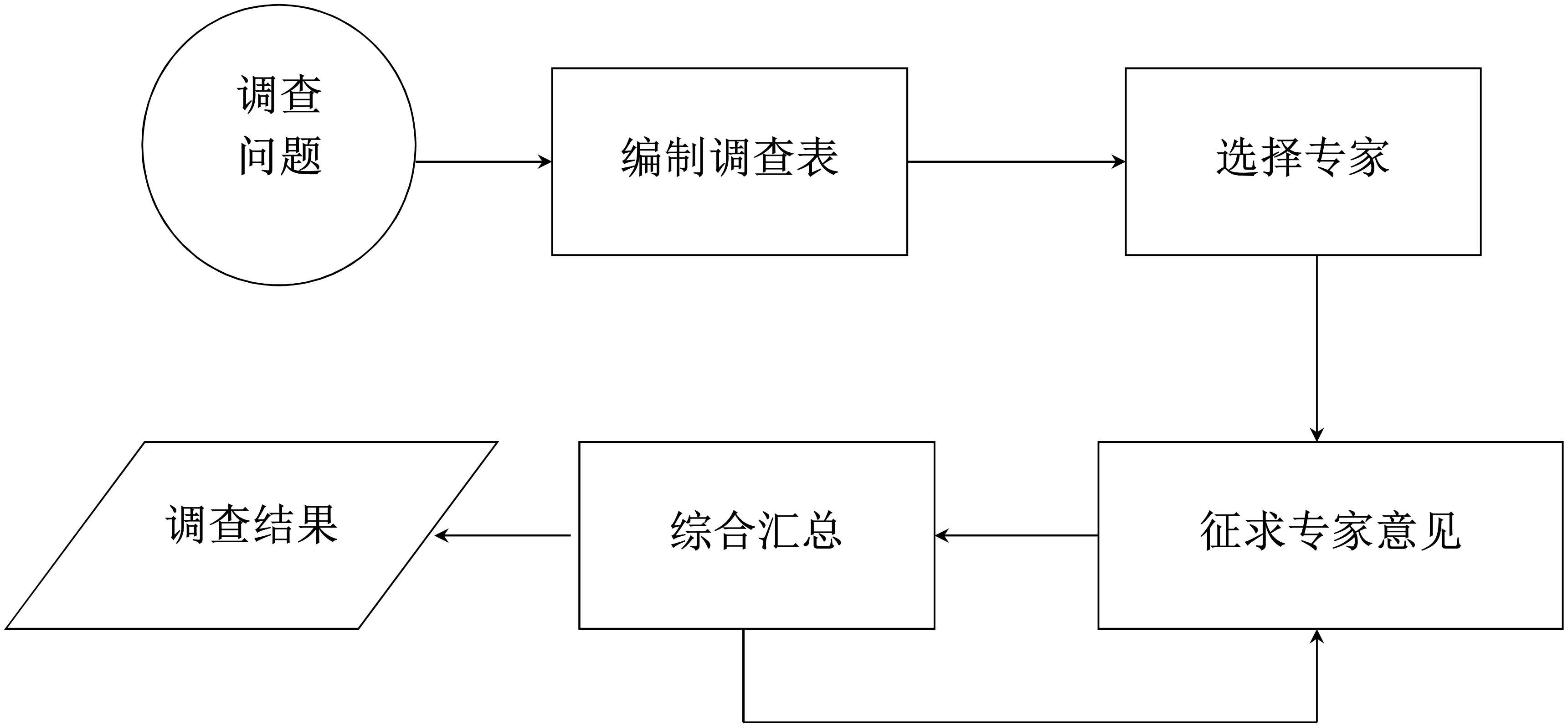
关联矩阵法的关键是确定评价指标的相对权重$w_1,w_2,\cdots,w_m$以及评价对象$A_1,A_2,\cdots,A_n$对应于评价指标P_j的评价值为$v_{1j},v_{2j},\cdots,v_{nj}$。在已知$w_1,w_2,\cdots,w_m$和$v_{1j},v_{2j},\cdots,v_{nj}$的情况下，关联矩阵法可以用下面（关联）矩阵表表示：

关联矩阵法 - 关联矩阵表

指标 P / 权重 W / 对象 A	P_1	P_2	$\cdots$	P_j	$\cdots$	P_m	综合评价
	w_1	w_2	$\cdots$	w_j	$\cdots$	w_m	v_i
A_1	v_{11}	v_{12}	$\cdots$	v_{1j}	$\cdots$	v_{1m}	$v_1=\sum_{j=1}^{m} w_j v_{1j}$
A_2	v_{21}	v_{22}	$\cdots$	v_{2j}	$\cdots$	v_{2m}	$v_2=\sum_{j=1}^{m} w_j v_{2j}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
A_n	v_{n1}	v_{n2}	$\cdots$	v_{nj}	$\cdots$	v_{nm}	$v_n=\sum_{j=1}^{m} w_j v_{nj}$

专家调查法 - DELPHI方法

- **DELPHI方法**（兰德公司，60年代）



DELPHI方法 – 特点

- 原则：独立性、匿名性、客观性。
- 调查过程的要求
  - 调查范围的广泛性：在首轮进行调查时，被调查人员/专家的选取要具有广泛的分布性；
  - 调查结果的收敛性：专家的反馈结果在几次以后可以收敛，达成一致；
  - 不断改进的准确性：专家的反馈结果一轮到另一轮应变得更为准确。

DELPHI方法 –方法指南

- 从为**DELPHI**小组人员那里获得愿意的协议。
- 向小组完整地说明**DELPHI**的手续和步骤。
- 以最简明的形式陈述事件、要素和预测。
- 尽可能使意见询问表容易回答。
- 使用数量适当的提问，少一些问题比太多的问题要好。
- 说明为什么包括了矛盾的事件、要素和预测，避免小组成员认为是在“误导他们”。
- **DELPHI**组负责人不能在小组成员的反馈中加入自己的个人见解，不要干涉小组的审议和商讨。
- 对**DELPHI**小组成员工作要给予适当的补偿（报酬）。
- **DELPHI**小组成员必须具有对所讨论问题的相关知识。

专家打分综合方法

设 m 个专家 $P_1, P_2, \cdots, P_m$，其权值为 $w_1, w_2, \cdots, w_m$；有 n 个评价对象 $A_1, A_2, \cdots, A_n$，专家 P_j 对 $A_1, A_2, \cdots, A_n$ 的打分值为 $v_{1j}, v_{2j}, \cdots, v_{nj}$。专家打分综合法见专家打分表

专家 P / 对象 A 权重 W	P_1	P_2	…	P_j	…	P_m	综合打分
	w_1	w_2	…	w_j	…	w_m	v_i
A_1	v_{11}	v_{12}	…	v_{1j}	…	v_{1m}	$v_1 = m\sum_{j=1}^{n} w_j v_{1j}$
A_2	v_{21}	v_{22}	…	v_{2j}	…	v_{2m}	$v_2 = \sum_{j=1}^{m} w_j v_{2j}$
…	…	…	…	…	…	…	…
A_n	v_{n1}	v_{n2}	…	v_{nj}	…	v_{nm}	$v_n = \sum_{j=1}^{m} w_j v_{nj}$
评价结果	根据综合打分 v_i（$i = 1,2,\cdots,n$）的大小排序						

模糊综合评价方法

设有 n 个评价因素 $A_1, A_2, \cdots, A_n$，评价因素集合记为：

$$A = (A_1, A_2, \cdots, A_n)$$

其重要程度由 n 维隶属函数向量 $W = (w_1, w_2, \cdots, w_n)$ 表示,其中 $\sum_{i=1}^{n} w_i = 1$，

$0 \le w_i \le 1$， $i = 1,2,\cdots,n$

再假设每个指标的评语为 $v_1, v_2, \cdots, v_m$，评语集合记为：

$$V = (v_1, v_2, \cdots, v_m)$$

这些评语的隶属函数向量用 m 维模糊向量表示：

$$B = (b_1, b_2, \cdots, b_m)$$

其中 $b_j \in [1,0]$，且 $\sum_{j=1}^{m} b_j = 1$。

模糊综合评价方法（续）

对评价因素 $A_1, A_2, \cdots, A_n$，评语隶属函数向量的取值构成一个 $n \times m$ 矩阵，记为

$$R = \begin{bmatrix} B_{A_1} \\ B_{A_2} \\ \cdots \\ B_{A_n} \end{bmatrix}$$

称矩阵 R 为模糊评价矩阵。

综合评价结果（向量）利用下式计算

$$B = WR$$

假设能给出综合各评语的记分方法，$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_m)$，则评价得到的总分值

$$d = \alpha B^T$$

模糊综合评价方法 – 举例

- 某服装厂对新产品方案的评价

评价因素集 A={款式色彩，穿着舒适，销售价格}，评语集合 V={很好，较好，不太好，不好}，通过专家分析和市场调查，对款式色彩有 20%人很欢迎、70%比较欢迎、10%人不太欢迎，因此，对款式色彩评价的隶书度为

$$B_{款式色彩} = \begin{pmatrix} 0.2 & 0.7 & 0.1 & 0.0 \end{pmatrix}$$

同样有

$$B_{穿着舒服} = \begin{pmatrix} 0.0 & 0.4 & 0.5 & 0.1 \end{pmatrix}$$

$$B_{销售价格} = \begin{pmatrix} 0.2 & 0.3 & 0.4 & 0.1 \end{pmatrix}$$

模糊综合评价方法 – 举例（续）

因此，得到对此新产品评价的模糊矩阵

$$R = \begin{bmatrix} 0.2 & 0.7 & 0.1 & 0.0 \\ 0.0 & 0.4 & 0.5 & 0.1 \\ 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix}$$

如果知道评价指标的权值为 $W = (0.2 \quad 0.5 \quad 0.3)$

则市场对该服装的综合评价为 $\overline{B} = WR = (0.2 \quad 0.4 \quad 0.5 \quad 0.1)$

归一化后得到 $B = (0.17 \quad 0.34 \quad 0.40 \quad 0.09)$

假设能给出综合各评语的记分方法，如 $\alpha = (1 \quad 0.7 \quad 0.4 \quad 0.1)$

则评价得到的总分值 $d = \alpha B^T = 0.69$

多个新产品可以通过各自的总分值进行评价。

层次分析法

- 一个简单例子：通过两两比较确定一组物品重量

用 $w_1, w_2, \cdots, w_n$ 记 n 个物品的重量，假设可以精确得到两两物品之间重量的比值，则可以得到如下重量比较矩阵

$$A=\begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \cdots & \cdots & \cdots & \cdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix}$$

对此物品重量的比较矩阵 $A=(a_{ij})_{n\times n}$ （$a_{ij}=w_i/w_j$）和物品重量向量 $W=[w_1 \quad w_2 \quad \cdots \quad w_n]^T$，有

$$AW=\begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \cdots & \cdots & \cdots & \cdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ \cdots \\ w_n \end{bmatrix}=\begin{bmatrix} nw_1 \\ nw_2 \\ \cdots \\ nw_n \end{bmatrix}=nW$$

层次分析法(续1)

表明n是矩阵A的特征值，W是对应的特征向量。而且可以证明矩阵A的最大特征值$\lambda_{\max}=n$，其余特征值等于零。

对物品重量的比较矩阵$A=(a_{ij})_{n\times n}$（$a_{ij}=w_i/w_j$），容易证明如下性质：

（1） $a_{ij}>0, a_{ij}=1/a_{ji}$， $i\neq j, i=1,2,\cdots,n, j=1,2,\cdots,n$

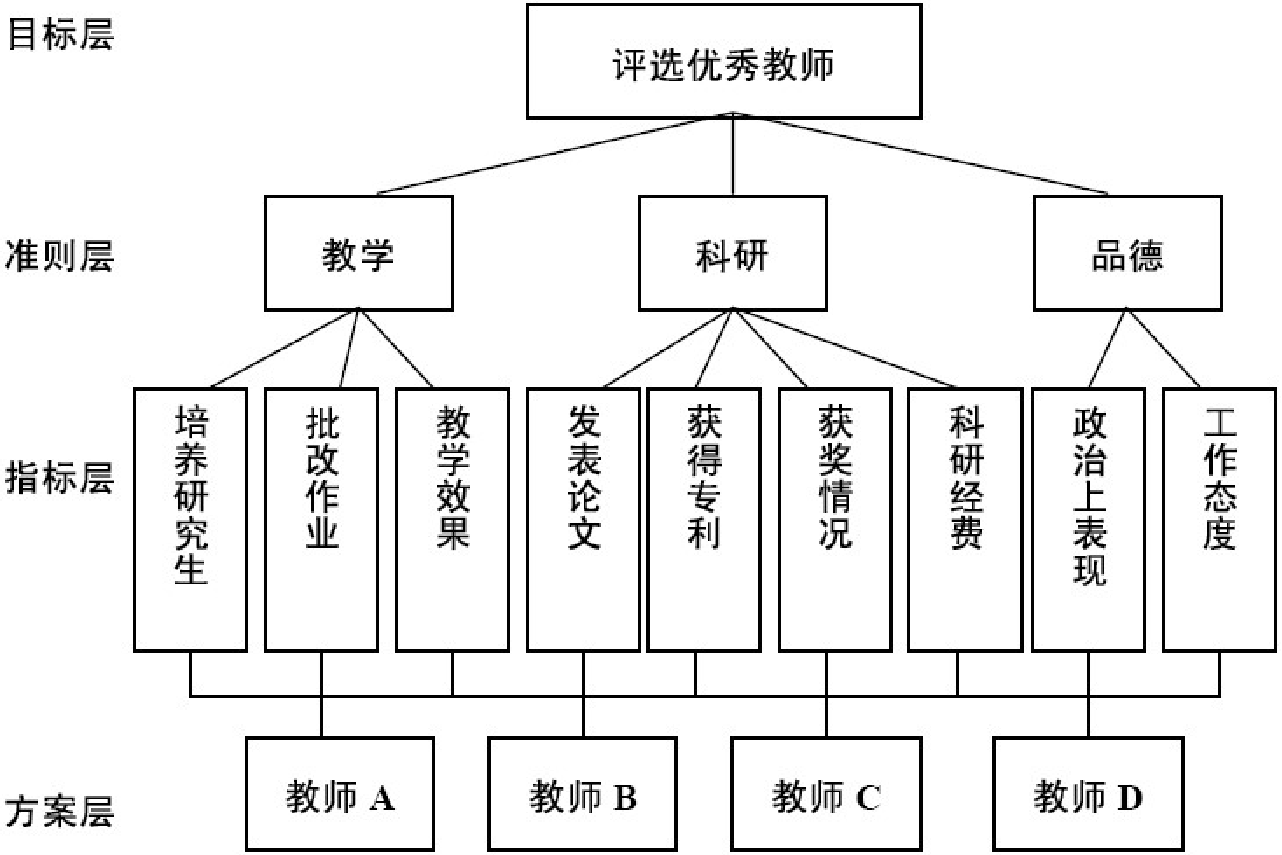
（2） $a_{ii}=1, i=1,2,\cdots,n$

（3） $a_{ij}\times a_{jk}=a_{ik}$， $i=1,2,\cdots,n, j=1,2,\cdots,n$， $k=1,2,\cdots,n$

因此，在已知重量比较矩阵A的情况下，通过求解A的最大特征值$\lambda_{\max}$对应的特征向量W，此特征向量的分量就是物品的重量。

层次结构模型

- **最高层**：表示解决问题的目的；
- **中间层**：表示实现总体目标的各种政策、评价指标和评价准则等；
- **最底层**：措施和方案层。



构造判断矩阵

定义 8.1 设有 n 个对象，若它们两两比较值构成的矩阵 $A=(a_{ij})_{n\times n}$ 满足条件：

（1） $a_{ij}>0, a_{ij}=1/a_{ji}$， $i\neq j, i=1,2,\cdots,n, j=1,2,\cdots,n$

（2） $a_{ii}=1, i=1,2,\cdots,n$

则称矩阵 A 为**判断矩阵**，也称满足这两个条件的矩阵为**正互反矩阵**

比较值取值的规定

重要程度	a_{ij}
相等（当）	1
i 比 j 较强	3
i 比 j 强	5
i 比 j 很强	7
i 比 j 绝对强	9
介于上述之间	2，4，6，8

判断矩阵的一致性

定义：如果正互反矩阵 $A=(a_{ij})_{n\times n}$ 满足

$$a_{ij}\times a_{jk}=a_{ik} \quad i=1,2,\cdots,n, j=1,2,\cdots,n，k=1,2,\cdots,n$$

则称 A 为**一致矩阵**。

备注：一致矩阵的定义是类比物体重量的比较关系。

一致矩阵 A 的性质：

（1） $a_{ij}=1/a_{ji}$， $a_{ii}=1$， $i=1,2,\cdots,n$， $j=1,2,\cdots,n$ 。

（2） A 的转置矩阵 A^T 也是一致矩阵

（3） $rank(A)=1$

（4） A 的最大特征值 $\lambda_{\max}=n$，其余特征值等于零。

（5） 若最大特征值 $\lambda_{\max}$ 对应的特征向量为 $W=(w_1,w_2,\cdots,w_n)^T$，则

$$a_{ij}=w_i/w_j，\quad i=1,2,\cdots,n，\quad j=1,2,\cdots,n$$

一致性判别指标

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

当判断矩阵完全一致时，$CI = 0$，$\lambda_{\max} - n$ 越大，CI 越大，判断矩阵的一致性越差，因此，CI 值要求越小越好。

随着 n 的增大，判断的误差也在变大，为了合理验证判断矩阵的一致性，需要考虑 n 的影响，因此，使用随机一致性指标 CR。

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

式中 RI 为平均随机一致性指标，下表是 500 个样本的平均值

矩阵阶数	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49	1.52	1.54	1.56	1.58	1.59

当 $CR < 0.1$ 时，判断矩阵的一致性是可以接受的。

最大特征值及其特征向量的计算

(1)幂法　步骤一：　任取与判断矩阵 A 同阶的归一化初始向量 W^0；

步骤二：　利用迭代公式计算（w_i^{k+1} 表示 W^{k+1} 的第 i 个分量）

$$\overline{W}^{k+1} = AW^k$$

$$\alpha = \sum_{i=1}^{n} \overline{w}_i^{k+1}$$

$$W^{k+1} = \frac{1}{\alpha}\overline{W}^{k+1} \qquad k = 0,1,2,\cdots$$

步骤三：　对预先给定的精度 $\varepsilon(>0)$，当

$$\max_{1\le i\le n} |\overline{w}_i^{k+1} - w_i^k| < \varepsilon$$

时迭代停止，令 $W = W^{k+1}$

$$\lambda_{\max} = \frac{1}{n}\sum_{i=1}^{n}\frac{w_i^{k+1}}{w_i^k}$$

则 $\lambda_{\max}$ 和 W 为所求的最大特征值和特征向量。

最大特征值及其特征向量的计算 - 和积法（近似方法）

步骤一： 对判断矩阵 A 的每一列进行归一化

$$\bar{a}_{ij} = \frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}} \qquad i, j = 1,2,\cdots,n$$

步骤二： 把正规化的判断矩阵每列相加

$$\bar{w}_i = \sum_{j=1}^{n} \bar{a}_{ij} \qquad i = 1,2,\cdots,n$$

步骤三： 对向量 $\overline{W} = (\bar{w}_1, \bar{w}_2, \cdots, \bar{w}_n)^T$ 进行归一化

$$w_i = \frac{\bar{w}_i}{\sum_{j=1}^{n} \bar{w}_j} \qquad i = 1,2,\cdots,n$$

步骤四： 令 $W = (w_1, w_2, \cdots, w_n)^T$，则 W 为所求的特征向量。

步骤五： 计算最大特征值 $\lambda_{\max} = \sum_{i=1}^{n} \frac{(AW)_i}{w_i}$，式中 $(AW)_i$ 为 AW 的 i 个分量。

最大特征值及其特征向量的计算 -方根法（近似方法）

步骤一：　把判断矩阵 A 的元素按行相乘

$$\overline{\overline{w}}_i = \prod_{j=1}^{n} a_{ij} \qquad i = 1,2,\cdots,n$$

步骤二：　把所得的乘积开 n 次方

$$\overline{w}_i = \sqrt[n]{\overline{\overline{w}}_i} \qquad i = 1,2,\cdots,n$$

步骤三：　对向量 $\overline{W} = (\overline{w}_1, \overline{w}_2, \cdots, \overline{w}_n)^T$ 进行归一化

$$w_i = \frac{\overline{w}_i}{\sum_{j=1}^{n} \overline{w}_j} \qquad i = 1,2,\cdots,n$$

步骤四：　令 $W = (w_1, w_2, \cdots, w_n)^T$，则 W 为所求的特征向量。

步骤五：　计算最大特征值 $\lambda_{\max} = \sum_{i=1}^{n} \frac{(AW)_i}{w_i}$，式中 $(AW)_i$ 为 AW 的 i 个分量。

层次单排序

（1）问题

确定n个对象（候选方案）$B_1, B_2, \cdots, B_n$关于指标A的排序/最优决策/评价结果。

（2）方法

- 构造判断矩阵A
- 一直性验证；
- 计算判断矩阵对应于最大特征值$\lambda_{\max} = n$的特征向量

$$W = (w_1, w_2, \cdots, w_n)^T$$

- 把W标准化（归一化），根据各个w_i的大小确定排序。

层次总排序

设层次 A 有 $A_1, A_2, \cdots, A_m$ 个因素，其总排序权值为 $\alpha_1, \alpha_2, \cdots, \alpha_m$；与 α_i 对应的本层次 B 有 $B_1, B_2, \cdots, B_n$ 个因素，B 层对应上层次因素 A_i 的单层排序的权值为 $b_{1i}, b_{2i}, \cdots, b_{ni}$。

层次总排序表

层次 A / 层次 B	A_1	A_2	…	A_i	…	A_m	B 层总排序 β_i
	α_1	α_2	…	α_i	…	α_m	
B_1	b_{11}	b_{12}	…	b_{1i}	…	b_{1m}	$\beta_1 = \sum_{j=1}^{m} \alpha_j b_{1j}$
B_2	b_{21}	b_{22}	…	b_{2i}	…	b_{2m}	$\beta_2 = \sum_{j=1}^{m} \alpha_j b_{2j}$
…	…	…	…	…		…	…
B_n	b_{n1}	b_{n2}	…	b_{ni}	…	b_{nm}	$\beta_n = \sum_{j=1}^{m} \alpha_j b_{nj}$

层次总排序中一致性指标的计算

在层次总排序中，两个层次的一致性检验指标之间有相应的计算关系，设

- CI 为层次 A 总排序的一致性指标
- RI 为层次 A 总排序的平均随机一致性指标
- CR 为层次 A 总排序的随机一致性指标

用 CI_i 表示与 α_i 对应的 B 层中判断矩阵的一致性指标，用 RI_i 表示与 α_i 对应的 B 层中判断矩阵的平均随机一致性指标，则 A 层一致性指标与 B 层的一致性指标之间的关系为

$$CI = \sum_{i=1}^{m} \alpha_i CI_i , \quad RI = \sum_{i=1}^{m} \alpha_i RI_i , \quad CR = \frac{CI}{RI}$$

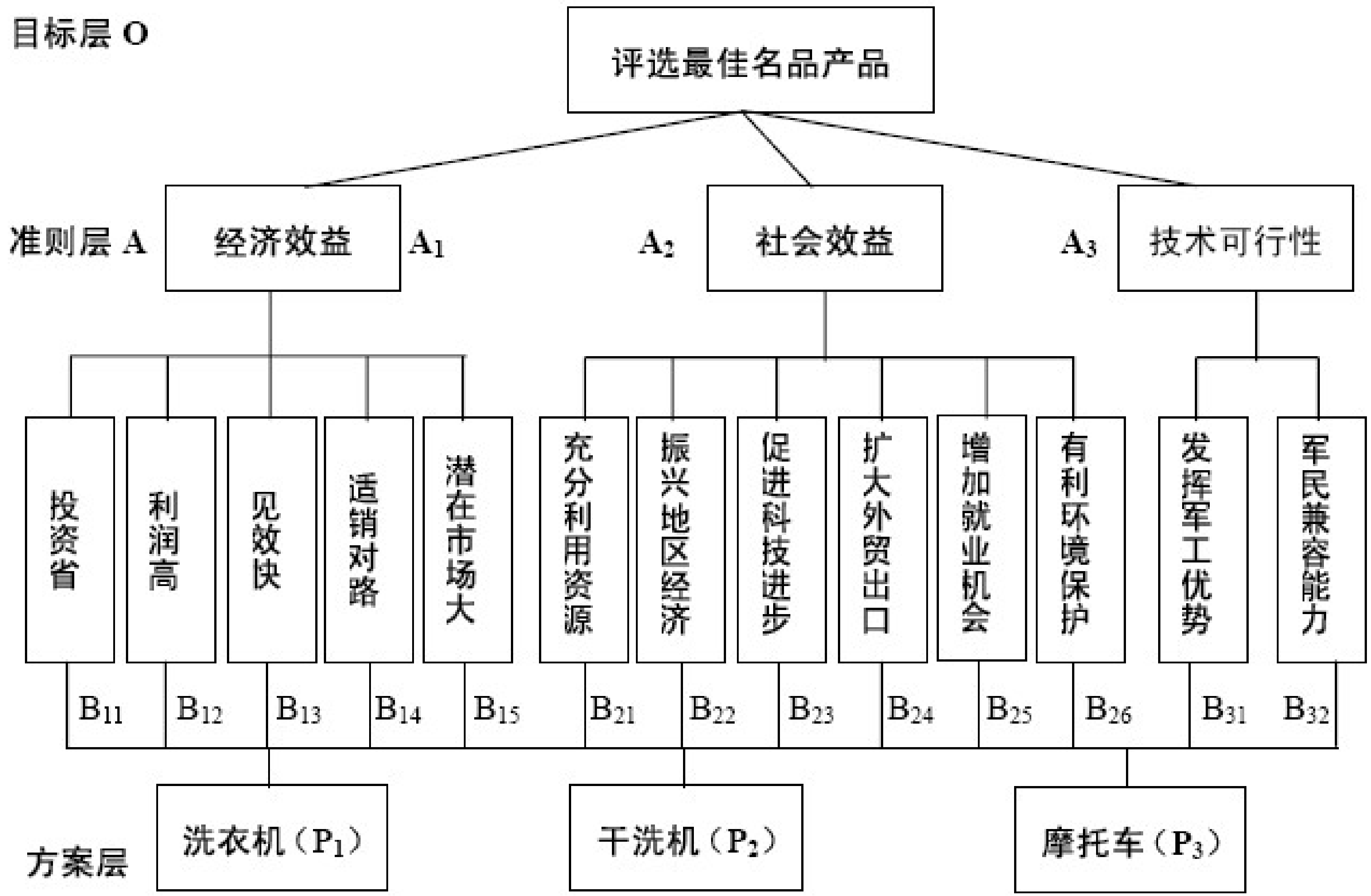
同样当 $CR < 0.1$ 时，层次总排序的计算结果具有满意的一致性。

层次分析法的步骤

- 分析与建立层次模型
- 利用单层排序方法计算每一层相对于上一层准则排序的加权值向量
- 从底层到高层，利用综合排序方法计算方案层对于各层直至最高层的总排序。

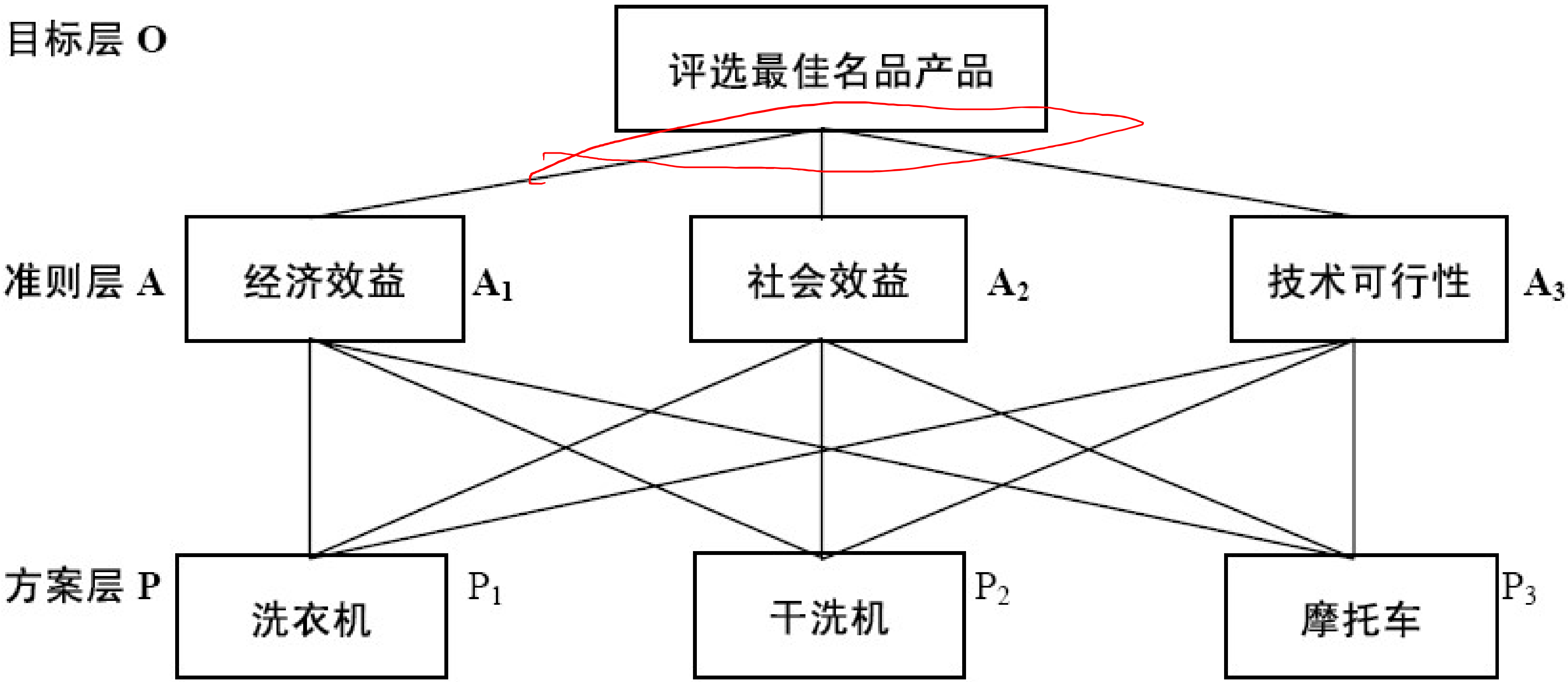
层次分析法举例

- 某军工企业名品生产方案评价



层次分析法举例(续1)

- 一个简化后的情况（两层）



层次分析法举例(续2)

- **O—A**判断矩阵（对目标**O**而言，评价准则**A1**、**A2**、**A3**相对重要性比较）

O	A_1	A_2	A_3	W
A_1	1	1	3	0.4286
A_2	1	1	3	0.4286
A_3	1/3	1/3	1	0.1428

其中 $\lambda_{\max} = 3, CI = 0, CR = 0$

层次分析法举例(续3)

- **A1—P**判断矩阵（对准则**A1**而言，方案P1、**P2**、**P3**相对重要性比较）

A_1	P_1	P_2	P_3	W
P_1	1	3	5	0.6370
P_2	1/3	1	3	0.2583
P_3	1/5	1/3	1	0.1047

其中 $\lambda_{\max} = 3.0385, CI = 0.0193, CR = 0.58$，$CR = 0.0332 < 0.1$

层次分析法举例(续4)

- **A2—P**判断矩阵（对准则**A2**而言，方案P1、**P2**、**P3** 相对重要性比较）

A_2	P_1	P_2	P_3	W
P_1	1	2	3	0.5499
P_2	1/2	1	1	0.2402
P_3	1/3	1	1	0.2098

其中 $\lambda_{max} = 3.0184, CI = 0.0092, CR = 0.58$，$CR = 0.0159 < 0.1$

层次分析法举例(续5)

- **A3—P**判断矩阵（对准则**A3**而言，方案P1、**P2**、**P3**相对重要性比较）

A_3	P_1	P_2	P_3	W
P_1	1	5	7	0.7207
P_2	1/5	1	3	0.1957
P_3	1/7	1/3	1	0.0835

其中 $\lambda_{max}=3.090, CI=0.045, CR=0.58$，$CR=0.078<0.1$

层次分析法举例(续6)

- 层次总排序（对对目标O而言，方案P1、**P2**、**P3**相对重要性比较）

O / P	A_1	A_2	A_3	层次P总排序权值	方案排序
	0.4268	**0.4268**	**0.1428**		
P_1	0.6347	0.5499	0.7207	0.6116	1
P_2	0.2583	0.2402	0.1957	0.2416	2
P_3	0.1047	0.2098	0.0835	0.1467	3

其中 $\lambda_{\max} = 3.090, CI = 0.045, CR = 0.58$， $CR = 0.078 < 0.1$

层次分析法举例(续7)

- 总排序的一致性验证

$$CI=\sum_{i=1}^{3}\alpha_i CI_i=0.4268\times0.0193+0.4268\times0.0092+0.1248\times0.045=0.0186$$

$$RI=\sum_{i=1}^{3}\alpha_i RI_i=0.4268\times0.58+0.4268\times0.58+0.1248\times0.58=0.58$$

$$CR=\frac{CI}{RI}=\frac{0.0186}{0.58}=0.032<0.1$$

优化问题、决策问题、评价问题的区别

- 优化问题
  - 掌握问题所有信息
  - 完全可以用数学形式描述
  - 客观性：问题解与求解人无关，完全可以计算机。
- 决策问题
  - 部分掌握问题有关信息，决策环境具有不确定性。
  - 数据工具使用受到限制
  - 主观性：问题解与决策者有关。
- 评价问题
  - 问题有关的许多信息不完整、相关知识有限，甚至存在争议。
  - 数学工具使用受到很大限制。
  - 高度主观性：依赖于来自不同领域的多个专家。

系统工程

资料阅读

- 阅读钱学森等提出的综合研讨厅及其应用方面的文章，关注国际上与此方法相近的研究领域。 ?
  - 基于综合集成的研讨厅体系与系统复杂性
  - 综合集成研讨厅中的研讨信息组织模型
  - 基于综合集成方法的网上舆论倾向分析与评估系统方案
  - 武器装备论证综合集成研讨厅系统

资料阅读2

- 选读下面（不限于此，鼓励自查文献）文章，加深对系统评价方法及应用方面知识的认识。
  - 产品数据管理应用系统评价方法的研究
  - 多层次系统的综合评价方法研究
  - 面向指挥自动化效能评估的**C3IEEE**系统设计
  - 模糊综合评判在物流系统评价中的应用
  - 企业集成的评价准则
  - 无人机电磁环境效应评估及其准则研究
  - **A Quantitative Method for Evaluating Machine Translation Systems**

资料阅读2（续）

- Design and evaluation of a generic software architecture for on-demand video servers
- Experimental evaluation of a COTS system for space applications
- Experimenting with quantitative evaluation tools for monitoring operational security
- Neural networks for short-term load forecasting-a review and evaluation
- Performance evaluation of storage systems based on network-attached disks
- Petri-net based performance-evaluation of distributed homogeneous task systems
- SCI evaluation in multinode environments for computing and data-processing
- Simulation-based performance evaluation of routing protocols for mobile ad hoc networks

作业

- 阅读钱学森等提出的综合研讨厅及其应用方面的文章，关注国际上与此方法相关的研究领域，写出一篇利用群体/专家智能进行评价或决策的简要综述。
- 查阅最新相关学术论文，并结合自己研究方向，说明层次分析法（或最新的其它评价方法）在系统评价中的应用。
- 简述优化问题、决策问题和评价问题的区别与联系。