

TOPSIS法（优劣解距离法）

TOPSIS法（Technique for Order Preference by Similarity to Ideal Solution）可翻译为逼近理想解排序法，国内常简称为优劣解距离法。

TOPSIS 法是一种常用的 综合评价方法，其能 充分利用原始数据的信息，其结果能精确地反映各评价方案之间的差距。利用数据进行说明，而且也对决策因素没有限制。

一、建模步骤

topsris进行建模，大致分为以下四个步骤：

- 1. 将原始矩阵正向化
- 2. 将正向化矩阵标准化
- 3. 计算得分并归一化

二、建模实现

例：小明同宿舍共有四名同学，他们第一学期的成绩、以及与他们吵架的次数（两种因素）如下表所示，请你为这四名同学进行评分，该评分能合理的描述综合成绩的高低。

姓名	成绩	与他人吵架的次数
小明	89	2
小王	60	0
小张	74	1
小李	99	3

第一步：将原始矩阵正向化

在生活中，常见的指标有四种：

指标名称	指标特点	例子
极大型（效益型）指标	越大（多）越好	成绩、GDP增速、企业利润
极小型（成本型）指标	越小（少）越好	费用、坏品率、污染程度
中间型指标	越接近某个值越好	水质质量评估时的PH值
区间型指标	落在某个区间最好	体温、水中植物性营养物量

那么，在 TOPSIS 方法中，就是要将所有指标进行统一正向化，即统一转化为极大型指标。 那么就需要极小型、中间型以及区间型的指标进行转化为极大型指标。

极小型指标 🖱️ 极大型指标

姓名	成绩	与他人吵架的次数	正向化后的争吵次数
小明	89	2	1
小王	60	0	3
小张	74	1	2
小李	99	3	0
指标类型	极大型	极小型	极大型

极小型指标转换为极大型指标的公式： $max - x$
如果所有的元素均为正数，那么也可以使用： $1/x$ ，公式不唯一！

中间型指标 🖱️ 极大型指标

中间型指标： 指标值既不要太大也不要太小，取某特定值最好（如水质质量评估 PH 值）。 $\{x_i\}$ 是一组中间型指标序列，且最佳的数据为 x_{best} ，那么正向化的公式如下：

$$M = \max |x_i - x_{best}|$$
$$\tilde{x}_i = 1 - M |x_i - x_{best}|$$

PH 值（转换之前）	PH 值（转换之后）
6	$1 - \frac{ 6-7 }{2} = \frac{1}{2}$
7	$1 - \frac{ 7-7 }{2} = 1$
8	$1 - \frac{ 8-7 }{2} = \frac{1}{2}$
9	$1 - \frac{ 9-7 }{2} = 0$

区间型指标 🖱️ 极大型指标

区间型指标： 指标值落在某个区间内最好，例如人的体温在36°~37°这个区间比较好。 x_i 是一组区间型指标序列，且最佳的区间为 $[a,b]$, 那么正向化的公式如下：

以下是支持markdown语法的数学公式，类似于 $\$$ ：

$$M = \max\{a - \min\{x_i\}, \max\{x_i\} - b\}, \tilde{x}_i = \begin{cases} 1 - \frac{a-x_i}{M}, & \text{if } x_i < a \\ 1, & \text{if } a \leq x_i \leq b \\ 1 - \frac{x_i-b}{M}, & \text{if } x_i > b \end{cases}$$

那么： $a = 36, b = 37, M = \max\{36 - 35.2, 38.4 - 37\} = 1.4$

体温（转换前）	体温（转换后）
35.2	0.4286
35.8	0.8571
36.6	1
37.1	0.9286
37.8	0.4286
38.4	0

第二步：正向化矩阵标准化

标准化的目的就是消除不同量纲的影响。假设有n个要评价的对象，m个评价指标（已经正向化了）构成的正向化矩阵如下：

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

那么对其标准化后的矩阵记为Z，Z的每一个元素： $z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}}$

即（每一个元素/根号下所在列元素的平方和）

得到标准化矩阵Z:

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nm} \end{bmatrix}$$

注意：标准化的方法不唯一，但目的都是为了去量纲。

那么对题目中的矩阵进行处理：

$$\begin{bmatrix} 89 & 1 \\ 60 & 3 \\ 74 & 2 \\ 99 & 0 \end{bmatrix}$$

经过标准化就变成了

$$\begin{bmatrix} 0.5437 & 0.2673 \\ 0.3665 & 0.8018 \\ 0.4520 & 0.5345 \\ 0.6048 & 0 \end{bmatrix}$$

第三步：计算得分并归一化

定义最大值：

定义最小值：

定义第*i* (*i* = 1,2,...,n) 个评价对象与最大值的距离：

$$Z^{+} = (\max\{z_{11}, z_{21}, \cdots, z_{n1}\}, \max\{z_{12}, z_{22}, \cdots, z_{n2}\}, \cdots, \max\{z_{1m}, z_{2m}, \cdots, z_{nm}\})$$

定义第*i* (*i* = 1,2,...,n) 个评价对象与最小值的距离：

$$Z^{-} = (\min\{z_{11}, z_{21}, \cdots, z_{n1}\}, \max\{z_{12}, z_{22}, \cdots, z_{n2}\}, \cdots, \max\{z_{1m}, z_{2m}, \cdots, z_{nm}\})$$

那么，我们可以计算得出第 *i* (*i* = 1,2,...,n) 个评价对象未归一化的得分: $D_i^{+} = \sqrt{\sum_{j=1}^m (Z_j^{+} - z_{ij})^2}$

很明显 $0 \leq S_i \leq 1$ ，且 S_i 越大 D_i^{+} 越小，即越接近最大值。 $D_i^{-} = \sqrt{\sum_{j=1}^m (Z_j^{-} - z_{ij})^2}$

那么对题目进行计算得分： 上一步标准化后的矩阵如下表：

姓名	成绩	正向化后的争吵次数
小明	0.5437	0.2673
小王	0.3665	0.8018
小张	0.4520	0.5345
小李	0.6048	0

最大值：[0.6048, 0.8018]，最小值：[0.3665, 0]

$$D_{\text{小明}}^{+} = \sqrt{(0.6048 - 0.5437)^2 + (0.8018 - 0.2673)^2} = 0.5380$$

$$D_{\text{小明}}^{-} = \sqrt{(0.3665 - 0.5437)^2 + (0 - 0.2673)^2} = 0.3206$$

$$D_{\text{小王}}^{+} = \sqrt{(0.6048 - 0.3665)^2 + (0.8018 - 0.8018)^2} = 0.2382$$

$$D_{\text{小王}}^{-} = \sqrt{(0.3665 - 0.3665)^2 + (0 - 0.8018)^2} = 0.8018$$

...

未归一化的得分: $D_i^{-} = \sqrt{\sum_{j=1}^m (Z_j^{-} - z_{ij})^2}$

得分结果如下：

姓名	D+	D-	未归一化得分	归一化后的得分	排名
小明	0.5380	0.3206	0.3734	0.1857	3
小王	0.2382	0.8018	0.7709	0.3834	1
小张	0.3078	0.5413	0.6375	0.3170	2
小李	0.8018	0.2382	0.2291	0.1139	4

三、模型扩展 (★)

本道题目中默认了各项指标的权重相同，但在实际的评价中指标都是有各自的权重，因此应该用权重对公式进行修正，修正后的公式如下， ω 代表权重。

$$D_i^+ = \sqrt{\sum_{j=1}^m \omega_j (Z_j^+ - z_{ij})^2}, \quad D_i^- = \sqrt{\sum_{j=1}^m \omega_j (Z_j^- - z_{ij})^2}$$

ω 可以由偏主观的层次分析法得出，也可由偏客观的熵权法得出，但是我建议使用综合主观与客观的权重，因为其更具有说服力！

四、模型总结

TOPSIS 法别名优劣解距离法，其主要利用数据的信息，精确的反应评价方案之间的优劣差距。TOPSIS 法多用于解决多指标的决策性问题，其实现原理为通过计算各备选方案与正负理想解之间的相对距离来进行排序并做出选择。其主要步骤如下：

1. 将原始矩阵正向化。（为了统一指标，方便后面计算，因此将指标统一为极大型指标）
2. 将正向化矩阵标准化。（消除量纲的影响）
3. 计算得分并归一化。（统计各指标的最大值，与最小值，并计算得分）

公式： $S_i = \frac{D_i^-}{D_i^+ + D_i^-}$ 实际上是 $\frac{x - \min}{\max - \min}$ 的变形：

$$\frac{x - \min}{\max - \min} \rightarrow \frac{x - \min}{(\max - x) + (x - \min)} \rightarrow \frac{\text{x与最小值的距离}}{\text{x与最大值的距离} + \text{x与最小值的距离}}$$

还有一点，实际上是 欧氏距离，TOPSIS 就是求得当前评价对象各指标与正负理想解相对距离，来最终做出评价的！