

熵值法和层次分析法在权重确定中的应用

陆添超, 康凯

(中国矿业大学计算机科学与技术学院, 徐州 221116)

摘要: 在实际生产实践中, 解决具体问题时, 往往会遇到权重系数确定的问题, 文中主要介绍了熵值法和基于模糊数学的层析分析法, 就其对这类问题的解决进行了详细介绍。结合具体的实例, 应用以上两种方法分别进行了求解, 并做了比较, 得出了较为理想的结果。

关键词: 熵值法; 模糊评价; 指标; 权重

The Application of Entropy Method and AHP in Weight Determining

ZHENG Weiguo, TIAN Qichong

(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116)

Abstract: The problem of determining the weight coefficients is often encountered in the actual production solving the specific problems. This paper introduced the entropy method and fuzzy mathematics based on chromatography analysis. The solution of these problems were given in detail. At last, the specific examples were carried out to solve applying the above two methods, and a more positive outcome was got after having a comparison between the results.

Key words: Entropy method; Fuzzy evaluation; Indicators; Weight

1 引言

在实际的生活、生产实践中经常会遇到排序或是排名等问题, 往往这些问题的处理都包含不只一个指标, 在问题的处理过程中需要综合这几个指标的作用才能解决问题, 然而这些指标的权重往往是不知道的, 所以这就需要在在解决问题前进行必要的指标权重系数的确定。在解决这一类问题中有许多方法, 本文介绍了熵值法和基于模糊数学的层次分析法。

2 熵值法

熵是系统无序程度的度量, 可以用于度量已知数据所包含的有效信息量和确定权重, 在水质评价中得到了广泛的应用。在水质模糊评价中, 通过对“熵”的计算确定权重, 就是根据各项监测指标值的差异程度, 确定各指标的权重。当各评价对象的某项指标值相差较大时, 熵值较小, 说明该指标提供的有效信息量较大, 其权重也应较大; 反之, 若某项指标值相差较小, 熵值较大, 说明该指标提供的信息量较小, 其权重也应较小。当各被评价对象的某项指标值完全相同时, 熵值达到最大, 这意味着该指标无有用信息, 可以从评价指标体系中去除^[2]。

使用熵权法确定权重主要有以下 3 个步骤:

(1) 原始数据矩阵归一化。设 m 个评价指标 n 个评价对象的原始数据矩阵为 $A = (a_{ij})_{m \times n}$, 对其归一化后得到 $R = (r_{ij})_{m \times n}$, 对大者为优的指标 (如 DO) 而言, 归一化公式为:

$$r_{ij} = \frac{a_{ij} - \min_j \{a_{ij}\}}{\max_j \{a_{ij}\} - \min_j \{a_{ij}\}},$$

而对小者为优的指标 (如 COD) 而言, 归一化公式为:

$$r_{ij} = \frac{\max_j \{a_{ij}\} - a_{ij}}{\max_j \{a_{ij}\} - \min_j \{a_{ij}\}} \quad (1)$$

(2) 定义熵。在有 m 个指标、 n 个被评价对象的评估问

题中, 第 i 个指标的熵为 $h_i = -k \sum_{j=1}^n f_{ij} \ln f_{ij}$, 式中

$f_{ij} = r_{ij} / \sum_{j=1}^n r_{ij}$, 其中, $k = 1 / \ln n$; 当时 $f_{ij} = 0$, 令。

(3) 定义熵权。定义了第 i 个指标的熵之后, 可得到第 i 个指标的熵权:

$$w_i = \frac{1 - h_i}{m - \sum_{i=1}^m h_i} \quad (0 \leq w_i \leq 1, \sum_{i=1}^m w_i = 1) \quad (2)$$

3 基于模糊数学的改进层次分析法

模糊综合评价的最终目的就是在论域个方案之间作相对优劣的比较, 从中选择相对最优的方案, 这种优选与论域以外的方案无关, 根据这一优化的相对性可以确定各评价指标值的相对隶属度和论域中相对优等方案与相对次等方案。不失一般性, 设有 n 个评价指标组成对全体 m 个方案的评价指标样本集数据 $\{x(i, j) | i = 1 \sim n, j = 1 \sim m\}$, 各指标值 $x(i, j)$ 均为非负值。为确定单个评价指标的相对隶属度的模糊评价矩阵, 消除各评价指标的量纲效应, 使建模具有通用性, 需对样本数据集 $\{x(i, j)\}$ 进行标准化处理^[3-4]。为了尽可能保持各评价指标值的变化信息, 对越大越优型指标的标准化处理公式可取:

$$(i, p) = x(i, p) / [x_{\max}(i) + x_{\max}(i)]$$

对越小越优型指标的标准化处理公式:

$$r(i, p) = [x_{\max}(i) + x_{\min}(i) - x(i, p)] / [x_{\max}(i) + x_{\min}(i)]$$

根据模糊评价矩阵 $R = (r(i, j))_{n \times m}$ 构造用于确定各评价指标权重的判断矩阵 $B = (b_{ij})$ 。模糊综合评价的实质是一种优选过程, 从综合评价的角度看, 若评价指标 i_1 的样本系列 $\{r(i_1, j) | j = 1 \sim m\}$ 的变化程度比评价指标 i_2 的样本系列 $\{r(i_2, j) | j = 1 \sim m\}$ 的变化程度大, 则评价指标传递的综合评价信息比评价指标 i_2 传递的综合评价信息多。基于此, 可用各评价指标的样本标准差 $s(i) = \left[\sum_{j=1}^m (r(i, p) - \bar{r}_i)^2 / m \right]^{0.5}$ 反映各评价指标对综合评价的影响程度, 并用于构造判断矩阵 B 。其中 $\bar{r}_i = \sum_{j=1}^m r(i, p) / m$ 为各评价指标下样本系列的均值, $i = 1 \sim n$ 。于是, 按照上式可得到 1~9 级判断尺度的判断矩阵:

$$b_{ij} = \begin{cases} \frac{s(i) - s(j)}{s_{\max} - s_{\min}} (b_m - 1) + 1, & s(i) \geq s(j) \\ 1 / \left[\frac{s(j) - s(i)}{s_{\max} - s_{\min}} (b_m - 1) + 1 \right], & s(i) < s(j) \end{cases}$$

式中: s_{\min} 、 s_{\max} 分别为 $\{s(i) | i = 1 \sim n\}$ 的最小值和最大值; 相对重要性程度参数值 $b_m = \min\{9, \text{int}[s_{\max} / s_{\min} + 0.5]\}$, \min 和 int 分别为取小函数和取整函数^[1]。

由于实际评价系统的复杂性、人们认识上的多样性以及主观上的片面性和不稳定性, 判断矩阵 B 的一致性条件不完全满足在实际应用中是客观存在、无法完全消除的, AHP 法只要求判断矩阵 B 具有满意的一致性, 以适应各种复杂系统。若 B 不具有满意的一致性, 则需要修正。设 B 的修正判断矩阵为 $Y = \{y_{ij}\}_{n \times n}$, Y 各要素的权重值仍记为 $\{w_i | i = 1 \sim n\}$, 则称使上式最小的 Y 矩阵为 B 的最优一致性判断矩阵:

$$\min CIC(n) = \sum_{i=1}^n \sum_{j=1}^n |y_{ij} - b_{ij}| / n^2 + \sum_{i=1}^n \sum_{j=1}^n |y_{ij} w_j - w_i| / n^2$$

$$s.t. \begin{cases} y_{ii} = 1, (i = 1 \sim n, j = i + 1 \sim n) \\ 1/y_{ji} = y_{ij} \in [b_{ij} - db_{ij}, b_{ij} + db_{ij}], (i = 1 \sim n, j = i + 1 \sim n) \\ w_i > 0, (i = 1 \sim n) \\ \sum_{i=1}^n w_i = 1 \end{cases}$$

为了度量判断矩阵是否具有满意的一致性, 这里引入判断矩阵的平均随机一致性指标系 $RIC(n)$ 值。用随机模拟方法分别对 3~n 阶各构造 500 个随机判断矩阵, 它们满足判断矩阵的单位性和倒数性, 但不保证判断矩阵满足一致性条件, 计算这些随机矩阵的一致性指标系数值, 然后平均即得 $RIC(n)$ 值, 参见表 1。经大量的计算, 当判断矩阵的一致性指标系数 $RIC(n) < 0.10$ 时, 可认为该判断矩阵具有满意的一致性, 据此计算的各评价指标的权重值 W_i 是可以接受的; 否则需提

高参数, 直到具有满意的一致性为止。

表 1 判断矩阵平均随机一致性指标系数 $RIC(n)$ 值

| 阶数 n | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|-------|-----|-----|-----|-----|-----|-----|
| $RIC(n)$ | 0.578 | 0.4 | 0.4 | 0.3 | 0.3 | 0.3 | 0.2 |

现对上述 AHP_FCE 法再作初步的理论分析。AHP 法的主要内容, 就是将待评价的复杂系统各要素按其关联隶属关系建立递阶层次结构模型, 构造两两比较的判断矩阵, 并据此求解各要素权重和检验、修正判断矩阵的一致性。AHP_FCE 法直接从原判断矩阵 B 的一致性程度出发构造上式的修正判断矩阵的准则函数, 根据上式, 原判断矩阵 B 具有完全一致性的充要条件是上式取全局最小 $CIC(n) = 0$, 该修正准则较为直观和简便; AHP_FCE 法通过原判断矩阵 B 各要素的调整来修正, 因此该法的修正具有全局性; AHP_FCE 的灵敏度分析的结果表明了 AHP_FCE 的计算结果是稳定的。

4 实例求解

4.1 用熵值法求解四个安全指标的权重系数

利用熵值法, 确定各指标的权重。原始数据矩阵归一化, 4 个指标对 35 个评价对象的原式数据矩阵为 $(a_{ij})_{m \times n}$ 。

对大者为优的指标 (如 DO) 而言, 应根据公式 (1) 进行归一化。 m 个指标中第 i 个指标的熵由公式 $h_i = -k \sum_{j=1}^n f_{ij} \ln f_{ij}$ 可以求得。其中 $f_{ij} = r_{ij} / \sum_{j=1}^n r_{ij}$, $k = 1 / \ln n$; 当 $f_{ij} = 0$ 时, 令 $f_{ij} \ln f_{ij} = 0$, 求得 h_i 的值。

最后根据公式 (2) 定义熵权, 定义了第个指标的熵之后, 可得到第个指标的熵权:

$$W_i [0.27881223 \quad 0.2878283 \quad 0.17071736 \quad 0.2626421]$$

4.2 评价法评价四个安全指标的权重系数

先求出样本标准差 $s(i)$, 构造判断矩阵 B , 得到样本系列的均值 \bar{r}_i 和判断矩阵 b_{ij} 。

经过以上计算, 求解各影响因素的权重。最后进行归一化处理, 归一化后得到的数据如表 2 所示:

表 2 各影响因素的权重

| 电导率 | 菌落总数 | 大肠菌群 | 霉菌和酵母 |
|--------|--------|--------|--------|
| 0.4673 | 0.2772 | 0.1601 | 0.0954 |

在得到各指标的权重后, 对模糊综合评价进行检验。通过计算机编程求解得到判断矩阵的一致性指标系数 $CIC(n) = 0.0025$ 。

当 $CIC(n) < 0.10$ 时, 可认为该判断矩阵具有满意的一致性, 因此, 可以认为用模糊综合评价法求得的各指标权重是有效的。

参考文献

- [1] 宋兆基, 等. Matlab6.5 在科学计算中的应用 [M]. 北京: 清华大学出版社, 2005.
- [2] 李志萍. 熵权法在农村安全饮水水质评价中的应用 [J].

(下转到 53 页)

```
dataGridView1.DataSource = ds.Tables [ 0] .Default~
View;
}
```

“输出到 Word” 按钮的 Click 事件的实现代码如下:

```
private void button2_Click (object sender, EventArgs
e)
{
    ExportDataGridview (dataGridView1, true);
}
```

定义方法 ExportDataGridview () 用来实现将 DataGridView 控件中数据导出到 Word, 该方法主要代码如下:

```
public bool ExportDataGridview (DataGridView dgv, bool
isShowWord)
{
    Word.Document mydoc = new Word.Document ();
    Word.Table mytable;
    Word.Selection mysel;
    Object myobj;
    if (dgv.Rows.Count == 0)
        return false;
    //建立 Word 对象
    Word.Application word = new Word.Application ();
    myobj = System.Reflection.Missing.Value;
    mydoc = word.Documents.Add (ref myobj, ref myobj,
ref myobj, ref myobj);
    word.Visible = isShowWord;
    mydoc.Select ();
    mysel = word.Selection;
    //将数据生成 Word 表格文件
    mytable = mydoc.Tables.Add (mysel.Range, dgv.Row~
Count, dgv.ColumnCount, ref myobj, ref myobj);
    //设置列宽
    mytable.Columns.SetWidth ( 30, Word.WdRulerStyle.
wdAdjustNone);
    //输出列标题数据
    for (int i = 0; i < dgv.ColumnCount; i++)
    {
        mytable.Cell ( 1, i + 1) .Range.InsertAfter ( dgv.
Columns [i] .HeaderText);
    }
    //输出控件中的记录
    for (int i = 0; i < dgv.RowCount - 1; i++)
    {
        for (int j = 0; j < dgv.ColumnCount; j++)
        {
            mytable.Cell ( i + 2, j + 1) .Range.InsertAfter
```

```
(dgv [j, i] .Value.ToString ());
}
}
return true;
}
```

本实例通过使用 Microsoft Word 自动对模型中的 Cell 对象, 将 DataGridView 控件中的数据导出到 Word 文档表格中, 本实例主要用到了 Cell 对象的 InsertAfter () 方法。

InsertAfter () 方法用来指定文本插入某区域或选定内容的后面, 其语法格式如下:

```
Public void InsertAfter ( string text)
```

参数 text 为要插入的文本。

另外, 需要特别注意的是在程序中对 Word 进行操作控制时, 需要引用 Word 动态链接库 (Microsoft Word 11.0 Object Library), 添加方法如下:

在 Visual Studio 2005 中选中当前项目, 单击右键, 选择“添加引用”选项, 在弹出的“添加引用”对话框中选择“COM”选项卡 (如图 2), 然后找到要引用的 Word 动态链接库, 单击“确定”按钮即可。

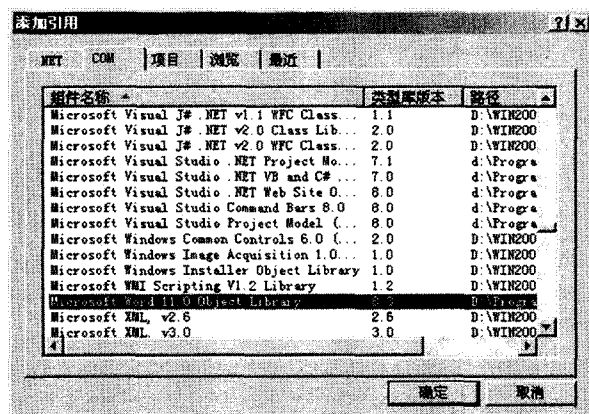


图 2 添加 Word 动态链接库

通过 C# 语言的一个 ASP.NET 实例, 来讨论解决在 ASP.NET 开发信息系统中如何将信息输出到 Microsoft Word 中进行查看或打印, 这些方法和技巧在开发基于 C# 语言的 ASP.NET 信息系统中有很高的实用价值。

作者简介

樊宁, 男 (1975), 硕士, 讲师, 研究领域为: 数据库, 计算机网络。

(上接第 20 页)

人民黄河, 2007, 29 (5): 35-36.

- [3] 刘德尚. 地方矿区生活饮用水卫生质量模糊综合评价 [J]. 安徽预防医学杂志, 2000, 6 (4): 309-310.
- [4] 张小咪, 邓旭东. 基于模糊层次分析评价的物流中心选址研究 [J]. 科学与管理, 2009 (1): 17-18.

- [5] 任珺, 等. 国内外饮用水水质标准的综合评价 [J]. 中国环境监测, 2007, 23 (6): 20-24.
- [6] 李满, 等. 饮用水生物稳定性的影响因素及评价体系 [J]. 水科学与工程技术, 2008, (1): 38-40.