

Research on Evaluation of Guizhou Big Data Industry Development Level Based on Intuitionistic Fuzzy Analytic Hierarchy Process

Junmeng Lu, Mu Zhang

School of Finance, Guizhou University of Finance and Economics, Guiyang 550025, China

基于直觉模糊层次分析法的贵州大数据产业发 展水平评价研究

卢俊萌, 张目

贵州财经大学金融学院, 贵阳 550025, 中国

Abstract

There is no reliable and unified standard for the comprehensive evaluation of the development level of the big data industry. The traditional methods for evaluating the level of industrial development cannot objectively and truly reflect the level of development of the big data industry. Based on previous studies, this paper constructs the evaluation index system of Guizhou's big data industry development level in terms of industrial scale and infrastructure, and uses intuitionistic fuzzy analytic hierarchy process to comprehensively evaluate the development level of Guizhou's big data industry. According to the results of the research, Guiyang City and Zunyi City are in the leading position, while the rest of the cities and provinces have little difference in development level. The overall development of Guizhou's big data industry is in good shape.

Keywords: Big Data; Big Data Industry; Intuitionistic Fuzzy Analytic Hierarchy Process (IFAHP); Comprehensive Evaluation

摘要

大数据产业发展水平综合评价尚无可靠统一的标准,传统的产业发展水平评价方法尚不能客观真实反映大数据产业发展水平。在前人研究基础上,本文在产业规模及基础设施两个方面构建贵州大数据产业发展水平评价指标体系,采用直觉模糊层次分析法对贵州大数据产业发展水平进行综合评价。从研究结果可得,贵阳市和遵义市处于领先地位,其余各地州市发展水平差异不大,贵州大数据产业整体发展态势良好。

关键词: 大数据; 大数据产业; 直觉模糊层次分析法(IFAHP); 综合评价

1 引言

大数据产业则是指以挖掘数据分析服务为核心,一切与大数据的产生与集聚、组织与管理、分析与发现、应用与服务相关的经济活动集合。近年来,发达国家纷纷加大对大数据产业发展的支持,将大数据纳入国家发展战略。美国把大数据称为“未来的新石油”。2012年4月,奥巴马政府宣布了“大数据研究和发展倡议”,投入了两亿以上资金^[1]。我国大数据产业亦呈良好发展态势,在国家政策的大力推动下,贵州、广东、上海和重庆等地方政府

先后出台大数据产业的相关发展规划纲要。大数据隐含的巨大经济价值和发展机遇已受到极大的重视，发展大数据产业无疑会给经济、科研、互联网等产生巨大的推动作用。

目前国内已有许多不同产业发展水平测度的研究和方法。王泽宇^[2]等人（2015）运用可变模糊识别模型，从产业的总体规模、产业结构、科技支撑、发展潜力、产业效率五个方面构建评价指标体系，对现代海洋产业发展水平进行综合评价；黄漫宇^[3]等人（2014）则基于钻石模型，从市场需求、资源条件、企业战略、结构与竞争对手、相关产业与支撑产业、政策支持五个方面寻找相应指标，采用灰色关联投影对中国绿色食品产业发展水平进行分析评价。但关于大数据产业发展水平综合评价，尚无可靠且统一的标准。此外，由于缺乏大数据产业发展水平评价的标准，用传统的产业发展水平评价方法只能对大数据产业发展水平进行相对的评价，尚不能客观反映大数据产业发展水平的真实情况。

徐泽水^[4]基于模糊集概念基础上提出的 IFAHP，统一了对定性和定量指标的集结。直觉模糊信息能够较好、较完整地反映决策主体对客观事物评价的不确定性，使得决策更为简单有效。有鉴于此，本文采用直觉模糊层次分析法(IFAHP)，建立贵州大数据产业发展水平评价指标体系，对贵州九个地州市大数据产业发展水平进行综合评价，以期为大数据产业发展水平评价指标的建立及运用的评价方法提供借鉴参考。考虑到大数据产业的实际评价背景，即专家对待评大数据产业的认知较为模糊且不确定，本文采用直觉模糊信息来刻画其评价信息。

2 直觉模糊集理论

直觉模糊集理论的概念由 Atanassov 在模糊集理论基础上推广得出^[5]，直觉模糊集理论常应用于决策领域^[6]，主要定义^[7]有：

定义 1 设 X 为一个非空集合，则

$$A = \{(x, \mu_A(x), \nu_A(x)) | x \in X\}$$

$$\mu_A: X \rightarrow [0,1], x \in X \rightarrow \mu_A(x) \in [0,1]$$

$$\nu_A: X \rightarrow [0,1], x \in X \rightarrow \nu_A(x) \in [0,1]$$

为直觉模糊集，其中 $\mu_A(x)$ 和 $\nu_A(x)$ 分别为 X 中元素 x 属于 A 的隶属度和非隶属度，且满足条件：

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1 \quad x \in X$$

定义 2 X 中元素 x 属于 A 的犹豫度为

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \quad x \in X$$

定义 3 称 $\alpha = (\mu_\alpha, \nu_\alpha)$ 直觉模糊数，其中

$$\begin{cases} \mu_\alpha \in [0,1] \\ \nu_\alpha \in [0,1] \\ \mu_\alpha + \nu_\alpha \leq 1 \end{cases}$$

基于直觉模糊数进行阐述，当决策者对方案的偏好信息存在知识缺乏或者有着一定的犹豫度时，直觉模糊理论能够同时表示“支持”、“反对”和“犹豫”。故在贵州大数据产业发展水平评价中指标权重的确定，将更具有客观性和合理性。

3 贵州大数据产业发展水平评价指标体系构建

要着力发展大数据产业，则需优先发展支撑大数据产业的制造业，从而为大数据产业后续发展奠定良好的经济基础。规模经济反映生产要素投入与产出之间的相关程度，故用于度量大数据产业的资产规模指标亦可用于测度大数据产业的发展水平。同时，基础设施在产业发展中也存在一定的支持作用。基于上述分析并结合专家意见，本文建立贵州大数据产业发展水平的评价指标体系如表 1 所示。

将贵州大数据发展水平设定为目标层，以大数据产业规模和基础设施这两个一级指标作为准则层。指标层则涵盖了工业增加值、主营业务收入、互联网宽带接入用户、交换机总容量等 12 个二级指标，从不同的侧面对贵州大数据产业发展水平进行综合评价。资产规模原始数据来自贵州各地州市 2016 年统计年鉴，基础设施原始数据来自贵州统计局 (<http://www.gz.stats.gov.cn/>)。

表 1 贵州大数据产业发展水平评价指标体系

目标层	准则层	指标层
贵州大数据产业发展水平	产业规模(A ₁)	规模以上企业数(个)(B ₁)
		企业全部从业人员年平均人数(人)(B ₂)
		工业总产值(亿元)(B ₃)
		工业增加值(亿元)(B ₄)
		资产总计(亿元)(B ₅)
		主营业务收入(亿元)(B ₆)
	基础设施(A ₂)	利润总额(亿元)(B ₇)
		固定资产本年累计完成投资(亿元)(B ₈)
		年末移动电话用户(万户)(B ₉)
		互联网宽带接入用户(万户)(B ₁₀)
		交换机总容量(万门)(B ₁₁)
		年末固定电话用户数(万户)(B ₁₂)

4 综合评价步骤

4.1 指标的标准化过程

直觉模糊层次分析一般将指标分为效益型、成本型、固定型、居中型等^[7]。某一指标值越大越好时,则该指标为效益型指标,本文选用指标均为效益型指标。指标标准化方法如下^[8]。

效益型指标的标准化值计算方法为:

$$z_{ij} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}}$$

式中: $i=1,2,\dots,m$, m 为评估对象个数; $j=1,2,\dots,n$, n 为属性个数; x_{ij} 为第 i 个评估对象按属性 j 进行测度得到的属性值; E 为评估对象集合。经标准化后所有指标值均严格从 0 到 1 变化,1 表示在某一属性中某一对象在所有对象中为最优值,0 为最差值。

表 2 定性评价语与直觉模糊数对应表

定性评价语	直觉模糊数	定性评价语	直觉模糊数
极端重要	(0.90,0.10,0.00)	较不重要	(0.40,0.45,0.15)
很重要	(0.80,0.15,0.05)	不重要	(0.30,0.60,0.10)
重要	(0.70,0.20,0.10)	很不重要	(0.20,0.75,0.05)
较重要	(0.60,0.25,0.15)	极端不重要	(0.10,0.90,0.00)
同等重要	(0.50,0.30,0.20)		

4.3 一致性检验及修正

在 IFHAP 中,为得到合理评价结果,需对直觉判断矩阵的一致性进行检验。本文通过构建积型一致性直觉判断矩阵 $\bar{R} = (\bar{r}_{ij})_{n \times n}$ 来建立一致性检验指标^[4]。

当 $j > i + 1$ 时, $\bar{r}_{ij} = (\bar{\mu}_{ij}, \bar{\nu}_{ij})$, 其中

4.2 直觉判断矩阵的建立

判断矩阵是决策者表达偏好关系的一种有效方式,以表达决策者偏好值的直觉模糊数来构建直觉判断矩阵,可更全面直观地描述决策者对所有指标的偏好信息。

决策者通过对指标间重要性进行比较后给出的定性评价语,再转化为直觉模糊数,由此构造直觉判断矩阵 $R = (r_{ij})_{n \times n}$, 其中 $r_{ij} = (\mu_{ij}, \nu_{ij})(i, j = 1, 2, \dots, n)$, $\mu_{ij} \in [0, 1]$, $\nu_{ij} \in [0, 1]$, $\mu_{ij} = \nu_{ji}$, $\nu_{ij} = \mu_{ji}$, $\mu_{ii} = \nu_{ii} = 0.5$, $\mu_{ij} + \nu_{ij} \leq 1$ 。 μ_{ij} 表示决策者对属性 i 和 j 的重要性比较时,偏爱 i 的程度, ν_{ij} 表示决策者对偏爱 j 的程度, $\pi_{ij} = 1 - \mu_{ij} - \nu_{ij}$, 表示决策的犹豫度。定性评价语与直觉模糊数的对应关系如表 2 所示^[9]。

$$\bar{\mu}_{ij} = \frac{\sqrt[j-i-1]{\prod_{t=i+1}^{j-1} \mu_{it} \mu_{tj}}}{\sqrt[j-i-1]{\prod_{t=i+1}^{j-1} \mu_{it} \mu_{tj} + \sqrt[j-i-1]{\prod_{t=i+1}^{j-1} (1 - \mu_{it})(1 - \mu_{tj})}}$$

$$\bar{\nu}_{ij} = \frac{\sqrt[j-i-1]{\prod_{t=i+1}^{j-1} \nu_{it} \nu_{tj}}}{\sqrt[j-i-1]{\prod_{t=i+1}^{j-1} \nu_{it} \nu_{tj} + \sqrt[j-i-1]{\prod_{t=i+1}^{j-1} (1 - \nu_{it})(1 - \nu_{tj})}}$$

当 $j = i + 1$ 或 $j = i$ 时, $\bar{r}_{ij} = r_{ij}$; 当

$j < i$ 时, $\tilde{r}_{ij} = (\bar{v}_{ij}, \bar{\mu}_{ij})$ 。

若 \mathbf{R} 与 $\bar{\mathbf{R}}$ 满足下述条件, 则认为直觉判断矩阵 \mathbf{R} 通过一致性, 否则未通过:

$$d(\bar{\mathbf{R}}, \mathbf{R}) < 0.1$$

其中 d 为 \mathbf{R} 与 $\bar{\mathbf{R}}$ 之间的测度距离, 有

$$d(\bar{\mathbf{R}}, \mathbf{R}) = \frac{1}{2(n-1)(n-2)} \sum_{i=1}^n \sum_{j=1}^n (|\bar{\mu}_{ij} - \mu_{ij}| + |\bar{v}_{ij} - v_{ij}| + |\pi_{ij} - \bar{\pi}_{ij}|)$$

若未通过一致性检验, 则需设置参数 σ 进行迭代, $\sigma \in [0, 1]$, 通过对迭代参数 σ 进行调整使直觉判断矩阵通过一致性检验。调整过程如下:

$$\bar{\mu}_{ij} = \frac{(\mu_{ij})^{1-\sigma} (\bar{\mu}_{ij})^\sigma}{(\mu_{ij})^{1-\sigma} (\bar{\mu}_{ij})^\sigma + (1 - \mu_{ij})^{1-\sigma} (1 - \bar{\mu}_{ij})^\sigma}$$

$$\bar{v}_{ij} = \frac{(v_{ij})^{1-\sigma} (\bar{v}_{ij})^\sigma}{(v_{ij})^{1-\sigma} (\bar{v}_{ij})^\sigma + (1 - v_{ij})^{1-\sigma} (1 - \bar{v}_{ij})^\sigma}$$

得到调整后直觉判断矩阵 $\bar{\mathbf{R}} = (\bar{r}_{ij})_{n \times n}$, 其中 $\bar{r}_{ij} = (\bar{\mu}_{ij}, \bar{v}_{ij})$ 。再进行一致性检验直至通过。

4.4 计算权重

对用过一致性检验的直觉判断矩阵, 可计算出各 2 级指标结合上一层指标权重的总权重, 计算公式如下, 其中 $i = 1, 2, \dots, n$:

$$\omega_i = \left(\frac{\sum_{j=1}^n \mu_{ij}}{\sum_{i=1}^n \sum_{j=1}^n (1 - v_{ij})}, 1 - \frac{\sum_{j=1}^n (1 - \mu_{ij})}{\sum_{i=1}^n \sum_{j=1}^n v_{ij}} \right)$$

表 3 归一化后的指标值

	贵阳市	六盘水市	遵义市	安顺市	毕节市	铜仁市	黔西南州	黔东南州	黔南州
B ₁	1.00	0.13	0.94	0.19	0.19	0.44	0.25	0.25	0.25
B ₂	0.95	0.03	1.00	0.22	0.51	0.27	0.28	0.29	0.32
B ₃	0.89	0.14	1.00	0.12	0.04	0.03	0.16	0.16	0.18
B ₄	0.82	0.22	1.00	0.09	0.27	0.14	0.15	0.15	0.16
B ₅	0.24	1.00	0.09	0.08	0.01	0.00	0.10	0.10	0.11
B ₆	0.74	0.33	1.00	0.12	0.02	0.00	0.16	0.16	0.18
B ₇	0.54	1.00	0.84	0.08	0.02	0.01	0.10	0.10	0.11
B ₈	1.00	0.04	0.41	0.12	0.28	0.15	0.15	0.15	0.17
B ₉	1.00	0.31	0.66	0.19	0.43	0.26	0.26	0.32	0.29
B ₁₀	1.00	0.21	0.58	0.19	0.30	0.24	0.23	0.34	0.27
B ₁₁	1.00	0.36	0.98	0.25	0.59	0.40	0.42	0.46	0.43
B ₁₂	1.00	0.24	0.66	0.18	0.24	0.17	0.15	0.28	0.29

据步骤 4.2、4.3 及 4.4, 通过由专家打分得出 12 个二级指标的直觉模糊偏好关系,

有表 1 得, 综合评价指标体系有 $n_1 = 2$ 个一级指标, 各指标间权重向量为 w_A ; 有 $n_2 = 12$ 个二级指标, 各指标间权重向量为 w_B 。本文采用直觉模糊数运算算子, 来计算指标层各因素相对于目标层的组合权重, 其中 $i = 1, 2, \dots, n_2$ [7]:

$$W_i = \bigoplus_{j=1}^{n_1} (w_{Bij} \otimes w_{Aj})$$

4.5 综合评价

将 m 个待评价对象归一化后指标值 $\mathbf{Z} = (Z_{ij})_{m \times n}$, 结合各指标相对于目标层的组合权重 \mathbf{W} , 得各指标综合评价 [10], 其中 $i = 1, 2, \dots, m$:

$$F_i = \bigoplus_{j=1}^n (Z_{ij} \oplus w_j)$$

$$kw = \langle 1 - (1 - \mu_w)^k, v_w^k \rangle$$

运用得分函数打分排序方法, 得出贵州大数据产业发展水平评价的得分函数:

$$\rho(F) = 1 - \frac{(1 + \pi_F)(1 - \mu_F)}{2}$$

将直觉模糊数 F 转化为可比较实数 $\rho(F)$, $\rho(F)$ 越大, 贵州大数据产业发展水平越高。

5 实证分析

根据步骤 4.1, 将各地州市评价指标值进行标准化, 得归一化指标值如表 3 所示。

构建资产规模下二级指标的直觉判断矩阵 R_1 及基础设施下二级指标的直觉判断矩阵 R_2 , 计算 R_1 与 \bar{R}_1 的距离和 R_2 与 \bar{R}_2 的距离, 经过调整后得 $d(\bar{R}_1, R_1) = 0.0981 < 0.1$,

$d(\bar{R}_2, R_2) = 0.0896 < 0.1$, 均通过一致性检验。计算各二级指标的总权重, 结果如表 4 所示:

表 4 总权重表

一级指标	二级指标	总权重
(0.3488, 0.5135)	(0.0582, 0.9063)	(0.0203, 0.9544)
	(0.0582, 0.9045)	(0.0203, 0.9535)
	(0.1122, 0.8333)	(0.0391, 0.9189)
	(0.1236, 0.8194)	(0.0431, 0.9122)
	(0.0938, 0.8559)	(0.0327, 0.9299)
	(0.1122, 0.8316)	(0.0391, 0.9181)
	(0.1577, 0.7865)	(0.0550, 0.8961)
	(0.1023, 0.8403)	(0.0357, 0.9223)
(0.5116, 0.3243)	(0.1685, 0.7183)	(0.0862, 0.8097)
	(0.2416, 0.6408)	(0.1236, 0.7573)
	(0.1629, 0.7394)	(0.0834, 0.8239)
	(0.2247, 0.6479)	(0.1150, 0.7621)

根据步骤 4.5, 将归一化后的指标值与二级指标对应的总权重通过直觉模糊运算,

得到贵州省九个地州市大数据产业发展水平的综合评价价值, 如表 5 所示。

表 5 贵州大数据产业发展水平的综合评价价值

地区	综合评价价值	排序
贵阳市	0.784	1
六盘水市	0.555	3
遵义市	0.761	2
安顺市	0.510	9
毕节市	0.533	4
铜仁市	0.514	8
黔西南州	0.518	7
黔东南州	0.532	5
黔南州	0.529	6

6 结论

本文通过运用直觉模糊层次分析法对贵州大数据产业发展水平进行实证测评, 并用其对九个地州市的金融集聚程度进行了综合评价。结果表明, 贵阳市及遵义市综合评价价值分别为 0.784 和 0.761, 远高于其余七个地州市, 大数据产业发展水平处于贵州省领先地位, 安顺市、铜仁市和黔西南州的综合评价价值分别为 0.510、0.514 和 0.518, 大数据产业发展水平处于全省最低, 黔东南州、黔南州、六盘水市和毕节

市四个地州市的大数据产业发展水平较接近分别为 0.532、0.529、0.555 和 0.533, 略高于铜仁市。可见, 贵州省各地州市的大数据产业发展水平较不均衡, 贵阳市和遵义市明显高于其余地州市, 推断存在区位和政策影响。评价结果与实际基本相符, 表明贵州大数据产业整体发展水平仍需进一步提高, 应采取相应政策, 在保证贵阳市和遵义市的发展优势下, 使其整体上得到更好的发展。

致谢

本文为国家自然科学基金地区项目《贷款风险补偿资金对科技型中小企业信

贷配给的影响机理研究》(71263011)、2017 年度第二批贵州省基础研究计划(软科学类别)项目《贵州大数据产业、生态文明建设与金融集聚耦合协调发展研究》

(黔科合基础(2017)1516-1)、贵州财经大学 2018 年度在校学生科研资助项目《贵州省大数据产业基金绩效评价体系研究》的阶段性成果之一。

参考文献

- [1] 汪应洛, 黄伟, 朱志祥. 大数据产业及管理问题的一些初步思考. 科技促进发展, 2014(1):15-19.
- [2] 王泽宇, 郭萌雨, 孙才志, 等. 基于可变模糊识别模型的现代海洋产业发展水平评价. 资源科学, 2015, 37(3):534-545.
- [3] 黄漫宇, 彭虎锋. 中国绿色食品产业发展水平的地区差异及影响因素分析. 中国农业科学, 2014, 47(23):4745-4753.
- [4] Xu Z, Liao H. Intuitionistic Fuzzy Analytic Hierarchy Process. IEEE Transactions on Fuzzy Systems, 2014, 22(4):749-761.
- [5] Atanassov K T. Intuitionistic fuzzy sets. Fuzzy Sets & Systems, 1986, 20(1):87-96.
- [6] He Y, He Z, Chen H. Intuitionistic Fuzzy interaction bonferroni means and its application to multiple attribute decision making. IEEE Transactions on Cybernetics, 2014, 45(1):116-128.
- [7] 徐泽水. 直觉模糊信息集成理论及应用. 科学出版社, 2008.
- [8] Zhang M, Zhou Z. A credit rating model for enterprises based on projection pursuit and k-means clustering algorithm. Journal of Risk Analysis & Crisis Response, 2012, 2(2):131.
- [9] 徐玖平, 吴巍. 多属性决策的理论与方法. 清华大学出版社, 2006.
- [10] Zeshui Xu, Ronald R. Yager. Some geometric aggregation operators based on intuitionistic fuzzy set. International Journal of General Systems, 2006, 35(4):417-433.