

Comprehensive Evaluation of E-Commerce Development Level Based on Factor Analysis and Cluster Analysis

Haijun Jiang

Yunnan University of Finance and Economics—Statistical Sciences, Kunming Yunnan
Email: 2684411018@qq.com

Received: Jul. 16th, 2020; accepted: Jul. 28th, 2020; published: Aug. 5th, 2020

Abstract

In order to understand the development situation in various regions of China's e-commerce, this article, from the Enterprise Environment, network infrastructure, economic level, logistics four, selects 10 indicators: the number of enterprises, the number of websites owned by enterprises, the proportion of enterprises with e-commerce transactions, the number of domain names, the number of websites, the number of web pages, GDP, disposable income, express delivery volume, express delivery business network number. The method of factor analysis is used to analyze the development level of e-business based on the comprehensive score and the relevant policies of various regions. It also points out the policies that should be used for reference and provides reference for the poorer areas. Then, on the basis of 10 selected indexes, the system clustering analysis method is used to classify 31 provinces and regions of the country. According to the classification situation, combined with the ranking of comprehensive scores and regional environment, this paper analyzes the areas with similar or similar development level of electronic commerce, and points out the common problems in order to better target regional development and offer advice. It can be seen that the development of e-commerce is uneven in the region, a large part of the reasons are attributed to the policy and geographical position, the eastern region is a better development, the second is the western region, the development is poor in the central region, and the regional classification is related to the geographical environment and comprehensive scores.

Keywords

E-Business, Development Level, Regional Differences, Factor Analysis, Clustering Analysis

基于因子分析和聚类分析对电子商务发展水平的综合评价

蒋海俊

云南财经大学统数学院—统计学专业, 云南 昆明
Email: 2684411018@qq.com

收稿日期: 2020年7月16日; 录用日期: 2020年7月28日; 发布日期: 2020年8月5日

摘要

为了了解我国各地区电子商务的发展情况, 本文从企业环境、网络基础设施、经济水平、物流方面四个方面选取企业数、企业拥有的网站数、有电子商务交易的企业比重、域名数、网站数、网页数、GDP、城镇居民人均可支配收入、快递量、快递营业网点数10个指标数据, 采用因子分析的方法根据综合得分排名情况, 结合各地区的相关政策分析电子商务的发展水平, 指出值得借鉴的政策, 为发展较差的地方提供参考; 再采用系统聚类分析的方法, 在10个选取的指标基础上, 将全国31个省地区进行分类, 根据分类情况, 结合综合得分的排名和区域环境, 分析电子商务的发展水平相似或相近的地区, 并指出存在的共性问题, 以便好针对区域发展提供建议。可以看出电子商务的发展在地区上表现出不平衡, 一大部分原因归结于政策和地理位置, 东部地区是发展最好的, 其次是西部地区, 发展较差的是中部地区, 且各地区的分类情况与地理环境及综合得分都有关。

关键词

电子商务, 发展水平, 地区差异, 因子分析, 聚类分析

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 绪论

1.1. 研究背景

我国电子商务经过十多年的发展, 已由成长期走向成熟期。几年前伟大的比尔·盖茨曾经说过: 2010年电子商务是转折期, 这个时期后主要的发展倾向于电子商务, 会开创出一片非同凡响的境界, 他这一预言早在我国也基本实现。电子商务是当代的一种商务方式, 数百万企业逐渐将业务推向信息化的市场, 导致越来越多的企业在业务中都或多或少掺杂着电子商务的交易; 对大众来说也是如此。电子商务的发展在不少传统的行业领域都取得相当不错的成绩[1]。

1.2. 研究意义

目前就我国电子商务的发展状况来说, 总体还是很不错的, 增长的趋势是显而易见的, 发展潜力巨大。企业、行业逐渐走向信息化的发展, 也为电商的发展打下坚实的基础。由于网络的普及以及发展, 电子商务迅速进入我们的生活, 发展前景较好, 很多的行业和企业纷纷开展电子商务[2]。但各个地区的发展水平可能不均衡, 同时也存在很多因素影响它的发展, 如消费者经济环境、企业环境、基础设施、物流等因素, 因此评估电子商务的发展水平, 找出主要因素成为当前的重点攻克对象, 了解不同地区电子商务发展存在的不足之处, 在此基础上提出合理性建议, 及时改善管理方法, 以促进电子商务更好的发展, 从而促进经济的全面发展[3]。

1.3. 国内外研究情况

何红用因子分析的方法,从综合得分和 GDP 的排名情况,得出电子商务的发展与经济水平之间的相关性很高,经济发达为电子商务的发展提供好的发展条件,而电子商务又带动经济的发展[4]。颜妍在因子分析的方法中,根据综合得分我国城镇居民的消费受各地区的经济发展状况影响比较明显,发达地区相对而言比较高,沿海地区和中西部地区的消费也存在很大的差异;在系统聚类的方法中根据聚类情况得出沿海地区由于实行改革开放,拉近了与北京、上海的消费水平[5]。

2. 理论知识

2.1. 因子分析理论知识

2.1.1. 因子分析法的基本思想及原理

1904 年查尔斯·斯皮尔曼通过对学生考试成绩的研究得出因子分析的思想。因子分析的核心是用较少的互相独立的因子反映原有变量的绝大部分信息[6]。基本思想是通过所选取的指标变量间的相关性大小,将原始变量进行分组,使得同组内的变量间的相关性普遍高,而不同组变量间的相关性则比较低[7]。每组变量代表一种基本结构,用一个不可观测的综合变量表示这种基本结构,这个基本结构称为公因子[8]。

设有 n 个样本,每个样本有 p 个变量,这 p 个变量之间有较强的相关性, $X = (X_1, X_2, \dots, X_p)'$ 是进行标准化的变量[9]。如果每个变量有 m 个潜变量 (F_1, F_2, \dots, F_m)

模型的矩阵形式为: $X = AF + \varepsilon$

其中

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix} = (A_1, A_2, \dots, A_m) \quad (2-1)$$

由模型(2-1)及其假设前提知,我们称 F_1, F_2, \dots, F_m 为 X 的公因子, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ 为 X 的特殊因子。

因子分析的基本假设为:

- (1): $m < p$
- (2): $E(X) = 0$
- (3): $E(F) = 0, \text{cov}(F) = I$
- (4): $E(\varepsilon) = 0, \text{cov}(F, \varepsilon) = 0$, 且 ε 的各分量之间相互独立

2.1.2. 因子分析的步骤

1) 因子分析的适合度检验

因子分析适合度主要有 Bartlett 球形检验、KMO 统计量和相关系数矩阵,本文采用 Bartlett 球形检验和 KMO 统计量[9]。

Bartlett 在 1950 年提出了球形检验,它是用于检验相关系数是否为单位阵的一种方法,SPSS 操作后根据 Bartlett 球形检验的卡方统计量的值和显著性值,判断相关系数是否为单位阵,若显著性 p 值接近 0,小于显著性水平,则认为因子分析适合[10]。

KMO 统计量是比较样本相关系数和偏相关系数,看 KMO 统计量的取值,KMO 统计量的定义如下[10]:

$$KMO_j = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} a_{ij}^{*2}}$$

$$KMO = \frac{\sum_{i \neq j} \sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} \sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} \sum_{i \neq j} a_{ij}^{*2}}$$

Kaiser (1974 年)给出的经验原则为[10] (见表 1):

Table 1. KMO decision table
表 1. KMO 判定表

KMO统计量	适合性
0.9以上	很好
0.8~0.9	良好
0.7~0.8	中度
0.6~0.7	一般
0.5~0.6	不好
0.5以下	不能接受

2) 确定提取因子个数

提取因子方法有主成份分析法、主轴因子法、极大似然法等, 本文采用主成份分析法[9]。

确定公因子个数的方法有以下三种方法:

第一种方法: 因子累积方差贡献率, 一般要求累积方差率大于 80%, 也就是才损失 20%的信息, 基本保留大部分信息[9];

第二种方法: 因子的特征值, 特征值以 1 为界线, 大于 1 的公因子就提取, 反之不提取;

第三种方法: 碎石图; 一般选取碎石图曲线变得平缓的点对应的值就是公因子的个数, 该方法不十分准确, 可能存在误差[9]。

本文选择累积方差贡献率。

3) 因子的命名解释

从载荷矩阵的系数判断提取的因子与变量之间的相关性, 若没有经过旋转的载荷矩阵各因子的系数与变量间相关性较弱, 则采用旋转后的矩阵进行分析。因子旋转主要是为了让载荷矩阵中的系数更加明显, 加强解释力度, 旋转的方法有正交和斜交, 本文选择最大方差法进行正交旋转。

4) 计算因子得分

因子得分是指所选取的公因子在每个样品指标上的值, 用回归思想建立公共因子与原始变量的线性组合, 即

$$F_j = \beta_{j1}X_1 + \beta_{j2}X_2 + \cdots + \beta_{jp}X_p, \quad j = 1, 2, \dots, m \quad (2-2)$$

本文计算综合得分采用综合评价法, 即以所提取的各公因子的方差贡献率占总方差贡献率的比重进行加权汇总[8], 具体公式如下:

$$F = \frac{w_1F_1 + w_2F_2 + \cdots + w_mF_m}{w_1 + w_2 + \cdots + w_m} \quad (2-3)$$

其中 w 为旋转前或旋转后的各因子的方差贡献率[12], 本文是旋转后的

2.2. 系统聚类分析理论知识

2.2.1. 系统聚类分析基本思想

系统聚类主要是对个体或者对象进行分类，使得分类后同一类中相似性比较接近或相似，分为 Q 型和 R 型聚类，本文是 Q 型聚类[9]。

系统聚类的基本思想可以归结为“由多到一，层层集聚”即将每个个体都置入聚类空间，然后将相近程度最高的两类合并组成一个新类，之后将新类与相近程度最高合并，不断重复，直到所有个体都归为一类[11]。

2.2.2. 系统聚类分析的步骤

系统聚类的具体步骤如下：

第一步：将每个数据作为一类(n 个个体共有 n 类)，按照定义的距离方法计算距离，并形成一距离阵，其中聚类方法主要有最近邻元素法、最远邻元素法、组间链接法、质心聚类法等方法[9]，本文主要采用组间链接法即类平均法；第二步：将距离最近的和并为一类，每次减少一个类别，再计算新类别和其余类别间的距离或相似度，形成新的距离阵；第三步：重复第二步，再将距离最接近的两个类别合并，如果类的个数大于 1，则继续重复这一步骤，直到所有数据合并为一类为止，最后将聚类过程用一张谱系图直观地表示，从中可以大致看出分类[9]。

3. 因子分析与系统聚类分析的综合评价

3.1. 评价指标的选择与数据来源

为了充分体现我国各地区电子商务的发展情况，本文从企业环境、网络基础设施、经济水平、物流方面选取企业数、企业拥有的网站数、有电子商务交易的企业比重、域名数、网站数、网页数、GDP、城镇居民人均可支配收入、快递量、快递营业网点数 10 个指标数据，数据来自 2014 年、2015 年以及 2016 年的《中国统计年鉴》相关指标数据。

3.2. 因子分析评价

3.2.1. 选择因子分析缘由

因子分析主要有以下两个方面的应用：一是将具有复杂关系的对象经过因子分析后综合为少数几个因子，用这几个因子的信息包含大多数因子的信息，二是用于分类，对变量或样品进行分类，本文主要采用第一方面的应用[8]。

3.2.2. 因子分析操作步骤

- 1) 标准化：分析 - 描述统计 - 描述将标准化得分另存为变量
 - 2) 因子分析：分析 - 降维 - 因子分析 σ 描述 - 相关矩阵 - KMO 和 Bartlett 检验
- σ 抽取 - 主成份
- 旋转 - 最大方差法
 - 得分 - 显示系数矩阵

3.2.3. 适合性检验

因子分析的主要目的是用来降维，从表 2 可以看出 KMO 测量值为 0.804，根据 KMO 指标值的判断标准，处于 0.8~0.9 之间，适合性为良好，从 Bartlett 球形度检验来看， P 值接近 0，即认为变量之间存在一定的关系，因此结合 KMO 和 Bartlett 球形度检验，采用因子模型是适宜的。

Table 2. KMO and Bartlett tests
表 2. KMO 和 Bartlett 检验

KMO和Bartlett的检验		
取样足够度的Kaiser-Meyer-Olkin度量		0.804
Bartlett的球形度检验	近似卡方	448.574
	Df	45
	Sig.	0.000

3.2.4. 提取的因子个数

从表 3 可以看出特征值大于 1 的有两个因子，且累计方差贡献率符合条件，第一个包含 10 个指标变量的 68.054% 的信息，第二公因子个包含 17.301% 的信息，提取的两个公因子的累计方差贡献率为 85.355%，大于 85%，说明这两个公因子代表了 10 个指标变量的大多数信息，已经达到降维的目的。

Table 3. Explains the total variance
表 3. 县解释总方差

成份	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差的%	累积%	合计	方差的%	累积%	合计	方差的%	累积%
1	6.805	68.054	68.054	6.805	68.054	68.054	5.195	51.952	51.952
2	1.730	17.301	85.355	1.730	17.301	85.355	3.340	33.403	85.355
3	0.541	5.407	90.762	—	—	—	—	—	—
4	0.360	3.596	94.358	—	—	—	—	—	—
5	0.230	2.300	96.658	—	—	—	—	—	—
6	0.156	1.556	98.214	—	—	—	—	—	—
7	0.127	1.270	99.484	—	—	—	—	—	—
8	0.032	0.317	99.801	—	—	—	—	—	—
9	0.013	0.133	99.934	—	—	—	—	—	—
10	0.007	0.066	100.00	—	—	—	—	—	—

3.2.5. 因子的命名解释

从表 4 可以看出，提取的两个因子中，第一个因子与 10 个变量的值都较高；第二个因子与 10 个变量的载荷普遍较低，解释作用不显著，因此采用最大方差法对载荷矩阵进行正交旋转。

Table 4. Matrix of unrotated components
表 4. 未旋转成份矩阵

	成份	
	1	2
Zscore(企业数)	0.897	-0.371
Zscore(有电子商务交易的企业比重)	0.425	0.710
Zscore(企业拥有网站数)	0.922	-0.312
Zscore(域名数)	0.855	0.036
Zscore(网站数)	0.906	0.272
Zscore(网页数)	0.774	0.501

Continued

Zscore(GDP)	0.877	-0.447
Zscore(城镇居民人均可支配收入)	0.698	0.549
Zscore(快递量)	0.914	0.010
Zscore(快递营业网点)	0.853	-0.405

从表 5 可以看出经过旋转后,第二因子与变量间的相关性升高,能更好解释变量,公因子 F_1 在企业数、企业拥有网站数、GDP、快递营业网点上的载荷量比较大,都在 0.9 以上。企业数、企业拥有网站数的载荷量分别为 0.95、0.938,对于企业来说,电子商务的快速发展,带动更多的人目睹电子商务的市场存在的巨大的发展潜力,越来越多的企业间纷纷开展电子商务的业务往来,导致拥有互联网的企业越来越多,同时也让电子商务在企业间得到快速发展,二者反映的是企业环境方面的指标;GDP 的载荷量达到 0.977,它是国民经济的指标,可以近似衡量地区的经济发展状况和发展水平,在一定程度上反映经济环境方面的指标;对于快递营业网点来说,载荷量达到 0.933,接近 1,电子商务的发展带动物流行业快速发展起来,消费方式也在此基础上发生变化,比如我们平时网购的东西,都离不开物流,该指标表示的是物流方面的指标,所以公因子 F_1 主要反映企业环境、经济环境、物流方面的指标,将 F_1 命名为带动发展的因子[12]。

Table 5. Rotation component matrix
表 5. 旋转成份矩阵

	成份	
	1	2
Zscore(企业数)	0.950	0.199
Zscore(有电子商务交易的企业比重)	-0.049	0.826
Zscore(企业拥有网站数)	0.938	0.261
Zscore(域名数)	0.686	0.511
Zscore(网站数)	0.595	0.735
Zscore(网页数)	0.357	0.850
Zscore(GDP)	0.977	0.125
Zscore(城镇居民人均可支配收入)	0.268	0.846
Zscore(快递量)	0.749	0.523
Zscore(快递营业网点)	0.933	0.145

公因子 F_2 在网站数、网页数、城镇居民人均可支配收入上的载荷量比较大。城镇居民人均可支配收入上的载荷量达到 0.846,由于经济快速发展,城镇居民人均可支配收入不断提高,越来越多的人跟上潮流,居民收入与电商的发展存在相辅相成的关系;网站数和网页数的载荷量分别达到 0.735、0.85,二者主要反映基础设施方面的指标,因而公因子 F_2 主要反映经济水平方面和基础设施方面的指标,将 F_2 命名为硬性条件因子。

根据表 3,由式 2-1 原始变量可由公因子 F_1 、 F_2 表示为:

$$\begin{cases} x_1 = 0.59F_1 + 0.199F_2 \\ x_2 = -0.049F_1 + 0.826F_2 \\ \vdots \\ x_{10} = 0.933F_1 + 0.145F_2 \end{cases} \quad (3-1)$$

3.2.6. 计算因子得分

根据表 6，由式(2-3)得出因子得分函数：

$$F_1 = 0.23x_1 - 0.18x_2 + 0.214x_3 + 0.092x_4 + 0.021x_5 \\ - 0.069x_6 + 0.252x_7 - 0.094x_8 + 0.108x_9 + 0.236x_{10}$$

$$F_2 = -0.103x_1 + 0.374x_2 - 0.073x_3 + 0.088x_4 + 0.205x_5 \\ + 0.303x_6 - 0.141x_7 + 0.32x_8 + 0.081x_9 - 0.1236x_{10}$$

Table 6. Scoring coefficient matrix

表 6. 得分系数矩阵

	成份	
	1	2
Zscore(企业数)	0.230	-0.103
Zscore(有电子商务交易的企业比重)	-0.180	0.374
Zscore(企业拥有网站数)	0.214	-0.073
Zscore(域名数)	0.092	0.088
Zscore(网站数)	0.021	0.205
Zscore(网页数)	-0.069	0.303
Zscore(GDP)	0.252	-0.141
Zscore(城镇居民人均可支配收入)	-0.094	0.320
Zscore(快递量)	0.108	0.081
Zscore(快递营业网点)	0.236	-0.123

最后综合得分的计算，采用综合评价法，对加权形成一维综合评估标准，累计方差贡献率为权重，由式(3-1)可得

$$F = \frac{51.592 * F_1 + 33.403 * F_2}{85.355}$$

2015 年各公因子得分、综合得分以及 2014 年综合得分比较见附录 B，可以看出西藏的硬性条件虽然已经提升很多，排名第七，这是值得鼓励的，但是它能带动的发展还是比较差的，主要原因是由于西藏属于高寒地区，自然环境恶劣，交通不便利；山东带动的发展是比较可喜的，排名第三，但是与广东还是有很大，综合得分相差 1.39，主要是由于山东发展的硬性条件较差。综合多的分排在前六的地区分别是广东、江苏、浙江、北京、山东、上海，且广东、江苏、浙江、北京的综合得分都在 1 以上，得分最高的广东得分为 2.317，比第二名高出 0.924。上海的得分与浙江、广东、北京存在差别，主要是由于上海的经济发展趋势不如广东[12]，从得分可以看出，电子商务的发展最好的是广东，发展较差的是青海地区。综合得分前五的地区都是属于东部经济发展较好的地区，得分靠后的新疆、甘肃、宁夏、西藏、青海都属于西部地区，其中青海和西藏地区是由于地理环境多的限制才导致电子商务和经济发展较差[13]。

不过东部地区与中部和西部地区的发展也没有绝对优劣之分，比如，属于西部的河南和四川的综合得分都排在前十，分别为第八名和第九名。而对于东部的海南来说，该地区的综合得分相对较后，排在 26 名，和西部地区发展较差的青海相差不是很大，再次说明西部地区的发展空间在逐渐扩大，有些地区甚至已经超过发展优势比较大的东部地区，西部地区正朝着赶上或者超越东部一部分地区的趋势发展。西部的大部分地区应该向四川多借鉴一下电子商务的发展模式以及政策。对比 2014 年可以明显看出，排

名前十的地区都没有发生改变，并且这些地区中，大多数都是增长的状态，只有山东和上海呈现下降趋势，但是下降的趋势并不大，10名到31名的地区也或多或少出现上升或者下降，程度都不大，说明各个地区都在加强电子商务的发展[13]。

在发展水平较好的几个地区中，以北京、广东为例。广东在电子商务发展方面独占鳌头，和经济水平有密不可分的关系，当然也离不开国家近年来对互联网的大力支持，广东地区的电子商务发展得到很好地发挥，以至于在短短10多年的时间里，就创造了广东经济遥遥领先的地位，这是其他地区应该学习的。对北京来说，优势在于大部分电子商务交易平台企业都聚集在北京，当然离不开国家政策的支持与鼓励。电子商务发展水平较低的地区中，选择宁夏、青海、西藏为例。这几个地区电子商务的发展相对来说是比较差的，原因方面有自身地理位置恶劣程度，导致与其他地区的发展难以维持，也有政策问题及对电子商务的认识等原因。首先说一下宁夏，在政策上虽然该地区的政府强调要加快该地区的电子商务的发展，也成立了电子商务协会，并且网民数量也在增长，但是相对其他发达地区还是相当落后的，主要是由于当地的经济水平较低，还不能够快速跟上时代的步伐，加之对电子商务的观念意识不够强，导致宁夏电子商务的发展不容乐观。对于青海来说，虽然已经出台很多政策，但海拔较高，难以发展[14]。对于西藏来说，政府也开始重视农牧区的通讯建设，开展了宽带通讯工程，建立一些电子信息服务平台，物流平台，让西藏地区的人也能慢慢享受电子商务所带来的便利。

3.3. 系统聚类分析评价

3.3.1. 选择系统聚类分析缘由

聚类分析中比较常用的方法主要是系统聚类和K均值聚类，对于K均值聚类法来说，优点是计算速度快，但对分类数没有统一的标准，并且当样本的变量超过3个时，K均值聚类法的可行性较差[15]。因此本文选择系统聚类。

3.3.2. 系统聚类分析操作步骤

分析 - 分类 - 系统聚类 σ^2 方法 - 组间连接
 σ^2 区间 - 平方 Euclidean 距离

3.3.3. 类别的确定

经过SPSS软件操作后得出聚类结果，并在excel中用聚类距离与聚类步数得出下图：

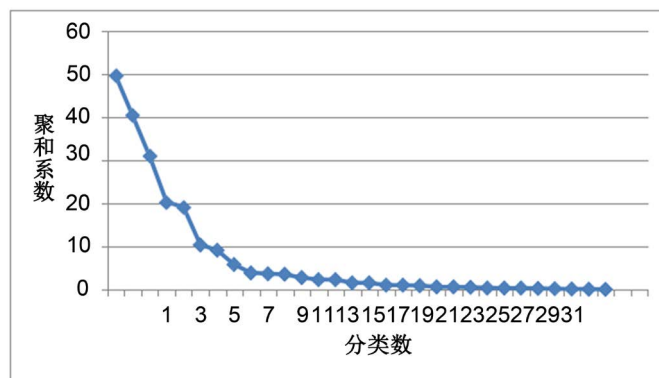


Figure 1. Graph of aggregation coefficient versus classification number
 图 1. 聚合系数随分类数变化的曲线图

从图1可以看出在聚合系数为3.944时，该曲线逐渐趋于平坦，所以将31个地区分为4类。

3.3.4. 系统聚类结果的分析

经统计软件 SPSS 运行结果如下图所示：

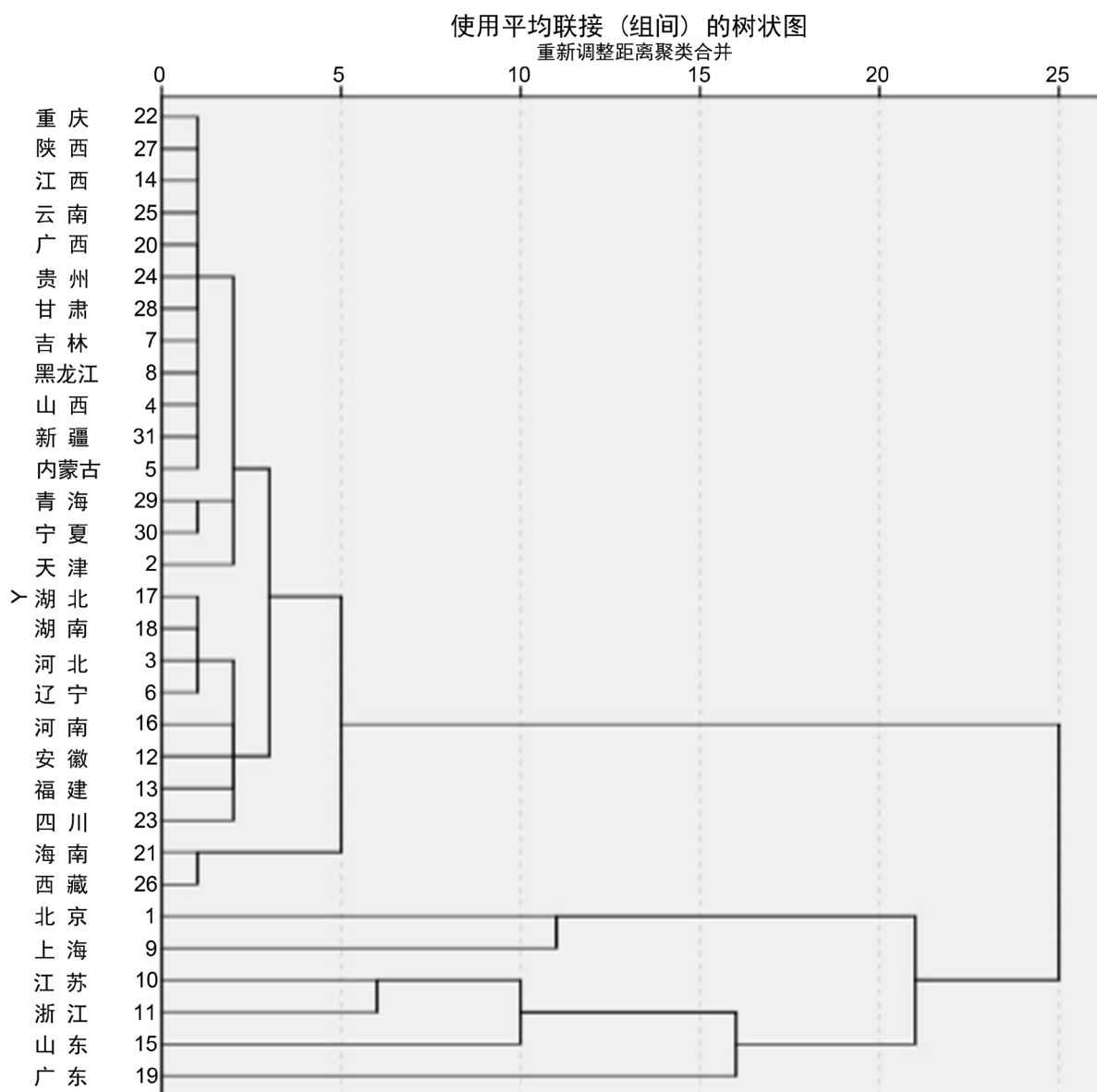


Figure 2. System clustering tree

图 2. 系统聚类树状图

从图 2 可以得到的分类结果如下：第一类：{广东}
 第二类：{山东、浙江、江苏} 第三类：{上海、北京}
 第四类：其他地区

从分类结果可以看出，我国电子商务发展水平极不平衡，可以看出区位条件对地区的经济的发展有着重要的影响[16]。前三类的综合得分相对而言较高，都是排在前十，其中广东、江苏、浙江、北京的综合得分分别是前四名，值都在 1 以上，说明综合得分较高的地区，电子商务的发展水平也相对比较高，结合综合得分可以看出，电子商务发展水平相似或相近的地区会被分成一类，在第二类中山东、江苏、

浙江的综合得分分别为 0.92、1.38、1.25 比较接近,在第三类中上海、北京的综合得分分别为 0.64、1.12,虽然相差有点大,但是从地理位置的划分上都是属于东部地区。这些地区的地理条件都比较优越,广东地处沿海,与香港和澳门相邻,加之港口较多,经济实力一直都很雄厚,经济技术方面也是数一数二的,因此发展电子商务比较有优势;浙江相邻地区较多,并且这些地区的经济实力都排在前面,比如上海、江苏;北京是交通要道,自然条件较好,地质比较稳定;上海有我国最大的海港,加之国家政策支持,海运在上海的发展是较好的。在第四类得分类中,属于东部地区的福建、河北、天津、海南综合得分排名分别为第七名、第十二名、第十五名、二十六名,属于西部的四川综合得分排名为第九名,而对于综合得分相对较低的地区当然被分为该类中。再次说明分类要考虑综合得分和地理位置的划分。对于综合得分较低的地区其电子商务的发展水平比较低,主要有:贵州、吉林、海南、新疆、甘肃、宁夏、西藏、青海这几个地区,其中西藏和青海的综合得分比较低。其主要原因是由于地理环境比较恶劣,西藏的海拔大多在 5000 米以上,属于高寒地区,并且该地区远离海洋,高山比较多,阻碍与其他地区经济的联系,比较难以扩展经济;交通也不便利,出行大多靠牛和马;教育比较落后,对电子商务的认识非常淡薄,西藏实行政教合一制,农牧民把大多时间都用来参加宗教活动,没有更多的时间来进行达电子方面经济的发展;对于青海地区来说,地理位置也是非常不容乐观,海拔都在 3000~5000 米,地形复杂多样,同样阻隔与相邻地区经济的联系。以上这些原因导致西藏和青海的电子商务发展水平是很低的[16]。

4. 结论及建议

4.1. 因子分析的结论

对于电子商务的发展,好政策的提出和政府的重视以及民众的认识和支持是非常重要的,当然也离不开创新。对于电子商务发展水平较高的地区,例如:北京、上海、广东,这三个地区的发展是非常快速的,这些地区都是在政府提出相关政策后,民众和企业纷纷响应,才致使电商的快速发展,但不仅仅是这些原因,还有其他方面原因,比如各地区的经济水平,经济水平决定消费水平,而对于发展水平较低的地区,企业和民众缺乏对电子商务的认知及辨识意识,因此对政府提出的相关政策没有做出积极响应。

4.2. 聚类分析结论

在因子分析中综合得分高的地区,分类会比较接近,当然分类情况还与地理位置等因素有关。

对于东部地区来说,综合得分较高,地理位置也比较有优势,大多属于沿海地区,地势比较平坦,交通便利,相邻地区较多,与这些相邻地区的相互合作与发展会拓宽发展方向的范围。相对中西部地区来说基础设施较完善,经济实力较好,对于发展电子商务来说,发展空间较大[16]。

而对于西部地区来说,很明显综合得分不高,这些地区的电子商务的发展由于地理环境的影响,相对东部比较欠缺。经济的发展影响着电子商务的发展,西部地区生态环境比较差,水土流失、草原退化、森林减少等问题比较严重,在加上人类的肆意破坏,使得西部地区的生态环境已愈发糟糕,地理位置的影响导致经济发展的方式比较局限,最终导致了西部地区的电子商务比较匮乏。

4.3. 建议

各地区的发展,都应该互相合作,取长补短,争取把经济最大化发展,发挥 $1+1>2$ 的协同效应[16]。对于西部发展水平较低的地区来说,首先应该提高他们对电子商务观念的认识,开展相关的活动,让他们体验电子商务所带来的便利,其次对待制度和政策都要创新,政府应该重视电子商务的发展。对于东部、中部地区来说,虽然发展水平大多高于西部,但也有很大的区别,比如广东明显比其他地区发展较好,因此,其他地区应该多向广东借鉴相关政策和模式。

电子商务现在处于发展的成长阶段, 为了更好更快的发展互联网, 在发展的同时需要采取不同的模式去创新, 并实现快速可持续发展, 让更多的企业在寻求新的商业模式基础上, 能充分结合传统产业发展的特色, 从而更好的发挥互联网经济增长的优势。

参考文献

- [1] 赵艳莉, 李瑶. 电子商务基础与应用[M]. 合肥: 安徽科学技术出版社, 2016: 4-21.
- [2] 杜江萍. 电子商务概[M]. 上海: 上海财经大学出版社, 2013: 23-43.
- [3] 金江军, 刘古权, 杨汉东. 信息化与区域经济发展[M]. 北京: 经济管理出版社, 2011: 223-224.
- [4] 何红. 我国电子商务发展及区域差异的初步分析[D]: [硕士学位论文]. 兰州: 兰州大学, 2016.
- [5] 颜妍. 我国城镇居民可支配收入与消费结构的实证分析[D]: [硕士学位论文]. 大连: 辽宁师范大学, 2016.
- [6] 王毓慧. 我国区域物流发展水平差异的实证研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨商业大学, 2016.
- [7] 庞雅莉. 电子商务情境下服装品牌竞争力的影响因素研究——基于顾客价值的视角[D]: [硕士学位论文]. 重庆: 西南大学, 2014.
- [8] 何晓群. 多元统计分析第四版[M]. 北京: 中国人民大学出版社, 2015: 142-169.
- [9] 冯岩松. SPSS22.0 统计分析应用教程[M]. 北京: 清华大学出版社, 2015: 287-355.
- [10] 李洪成, 姜宏华. 数据分析教程[M]. 北京: 人民邮电出版, 2012: 251-253.
- [11] 吴广, 刘荣, 牛玲. SPSS 分析与应用[M]. 北京: 电子工业出版社, 2013: 292-296.
- [12] 谢龙汉, 上涛, 蔡明京. SPSS 统计分析与数据挖掘[M]. 北京: 电子工业出版社, 2014: 281-289.
- [13] 徐文静. 沿海港口城市物流竞争力评价与聚类分析[J]. 商业经济研究, 2017(24): 85-88.
- [14] 刘奇洪. 区域中国: 把脉中国区域发展[M]. 北京: 中国经济出版社, 2009: 77-86.
- [15] 胡雷芳. 五种常用系统聚类方法及其比较[J]. 浙江统计, 2007(4): 11-13.
- [16] 黎雨. 大格局: 变动中的中国区域发展战略布局[M]. 北京: 国家行政学院出版社, 2013: 1-3.