

# 上市企业盈利能力的主成分聚类分析

罗立宇

(湖南工业大学 计算机与通信学院, 湖南 株洲 412008)

**摘要:** 采用主成分聚类分析方法 (PCA) 对长株潭城市群上市企业盈利能力进行了评价分类。先建立上市企业盈利能力评价指标体系, 然后用主成分分析法对 7 个量化指标进行分析, 提取 2 个主成分, 利用得到的主成分替代原来的评价指标, 对盈利能力进行聚类, 将长株潭城市群上市企业盈利能力分为 3 个级别, 为决策者决策提供依据。

**关键词:** 聚类分析; 主成分分析法; 盈利能力

中图分类号: F224

文献标志码: A

文章编号: 1673-9833(2010)06-0063-04

## Principal Component Cluster Analysis on Listed Companies Profitability

Luo Liyu

(School of Computer and Communication, Hunan University of Technology, Zhuzhou Hunan 412008, China)

**Abstract:** Uses principal component analysis (PCA) and cluster analysis to evaluate and classify the profitability of listed companies in Changsha-Zhuzhou-Xiangtan area in Hunan province. Firstly, sets up the evaluating indicator system of the listed companies, then uses principal component analysis to analyze seven quantize pointers, extracts two principal components to substitute the former evaluating indicators, makes clustering on the profitability, divides the profitability of the listed companies into three levels, and provides scientific basis for the decision maker.

**Keywords:** cluster analysis; principal component analysis; profitability

## 0 引言

长株潭城市群位于湖南省东北部, 包括长沙、株洲、湘潭 3 市, 是我国京广经济带、泛珠三角经济区、长江经济带的接合部, 区位和交通条件优越, 是湖南省经济发展的核心增长极。在现代市场经济条件下, 企业是市场的主体, 对实现城市的可持续发展, 起着重要作用。企业改制上市可以借助资本市场的力量, 迅速发展壮大, 从而实现资本的流动和资源的优化配置。盈利能力是上市公司获取利润能力的简称, 也是衡量企业当前收益水平和今后收益持续性和增长性的主要指标, 反映了公司经营管理水平的高低。通过建

立上市企业盈利能力评价指标体系, 分析上市企业盈利能力影响因素和影响程度, 有利于企业及时发现问题, 改善企业财务结构, 提高企业偿债能力、经营能力, 最终提高企业的盈利能力, 促进企业长期稳定发展, 提高企业竞争能力, 降低企业经营风险和可持续发展后劲, 促进社会经济的发展。

## 1 上市企业盈利能力指标分析

上市公司是一般企业的特殊形式, 因此反映一般企业盈利能力的指标<sup>[1]</sup>都适用于上市公司。但上市公司又不同于一般企业, 具有一些特殊指标<sup>[2-5]</sup>。直接选

收稿日期: 2010-10-05

基金项目: 湖南省教育厅优秀青年科研基金资助项目 (06B103)

通信作者: 罗立宇 (1969-), 女, 湖南茶陵人, 湖南工业大学副教授, 硕士, 主要从事计算机网络与数据库教学与研究,

E-mail: liyu-luo@163.com

取财务数据相关指标进行分析,难以掌握它们的真正含义,通过不同的比率可方便地综合分析公司的财务状况。因此根据分析目标,本文选取如下比率做为分析指标。

股东权益比率  $X_1$ : 是股东权益与资产总额的比率,反映企业资产中有多少是所有者投入的。

股本报酬率  $X_2$ : 反映公司股本总额中平均每百元股本所获得的纯利润。

总资产报酬率  $X_3$ : 是指企业一定时期内获得的报酬总额与资产平均总额的比率,用以评价企业运用全部资产的总体获利能力。

净资产报酬率  $X_4$ : 它反映企业总资产获取收益的能力。

总资产利润率  $X_5$ : 是指利润与资产总额的对比关系,它从整体上反映企业资产的利用效果,可用来说明企业运用其全部资产获取利润的能力。

销售净利率  $X_6$ : 是指企业实现净利润与销售收入的

的对比关系,用以衡量企业在一定时期的销售收入获取的能力。

营业利润率  $X_7$ : 是指企业的营业利润与营业收入的比率。反映在不考虑非营业成本的情况下,企业管理者通过经营获取利润的能力。

## 2 数据来源与研究方法

### 2.1 数据来源

选取长株潭城市群 36 家上市公司的 7 项指标为样本,对 2010 年三季报财务报表数据进行分析,原始数据取自凤凰网财经频道。由于数据差别较大,为了不突出大数据和削弱小数据的作用,消除数量级的影响,对原始指标数据进行总和标准化处理:矩阵中每一个元素除以所在列中所有元素之和,使所有元素在 (0,1) 内。结果见表 1。

表 1 标准化后上市企业指标值

Table 1 The standardized indicator values of the listed companies

编号	上市企业	评价指标						
		$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$
1	中联重科	0.017 180	0.090 697	0.069 023	0.159 900	0.059 458	0.053 834	0.049 742
2	通程控股	0.015 706	0.018 852	0.021 616	0.048 531	0.020 803	0.012 183	0.012 641
3	华天酒店	0.019 788	0.019 426	0.025 141	0.043 801	0.023 138	0.046 055	0.033 743
4	湖南投资	0.037 459	-0.011 604	0.025 884	0.024 470	0.026 777	0.146 150	0.162 660
5	金德发展	0.037 989	-0.002 470	-0.008 999	-0.008 059	-0.007 483	-0.018 768	-0.038 351
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	株冶集团	0.015 034	-0.024 807	-0.018 647	-0.039 012	-0.014 212	-0.006 667	-0.006 254
35	新五丰	0.032 637	0.000 507	0.001 113	0.001 168	0.004 051	0.000 740	-0.020 608
36	广汽长丰	0.022 362	0.030 111	0.024 214	0.039 596	0.021 353	0.014 158	0.013 201

### 2.2 研究方法

#### 2.2.1 主成分分析法

在系统分析中使用主成分分析方法通常有 2 个目的<sup>[6-8]</sup>: 一是利用降维的思想消除各指标之间相关性所带来的信息重叠,从而降低分析问题的复杂性;二是克服综合评价中人为确定各指标权重系数的问题,从而能适用于各领域的综合评价中,实现对系统的定量分析。

将该方法应用于长株潭上市企业盈利能力分析中,主要是为了消除指标相关性导致的信息重叠,为下一步的聚类分析做数据处理准备,基本步骤是:

- 1) 对原始数据矩阵进行总和标准化,计算相关系数矩阵;
- 2) 求相关系数矩阵的特征根和相应的特征向量;
- 3) 计算主成分贡献率及累计贡献率,通过累计贡献率确定主成分个数;
- 4) 计算主成分载荷,确定各主成分得分。根据主

成分个数,确定新的分析指标,得到新指标系统下的样本数据,为聚类分析提供样本。

#### 2.2.2 聚类分析方法

聚类分析是一种建立分类的方法,根据数据样本的不同特征,依据一定算法将样本划分为不同的类别。该方法可以从样本数据出发,客观地决定分类标准,通常一个类就是一个具有相似的个体的集合,不同类之间具有明显的非相似性。在分类过程中,不必事先给出一个分类标准,根据样本数据特征,客观地决定分类结果。从数学意义上讲,聚类的目的就是使属于同一类别的个体之间的距离尽可能小,而不同类别上的个体间的距离尽可能大。通过对上市企业样本数据进行聚类分析,将其划分为不同的类别进行研究,为决策者提供决策依据。常用的聚类分析方法有模糊聚类分析法、系统聚类分析法和动态聚类分析法,本文采用系统聚类分析法,类间亲疏程度度量方法采用最短距离法。

2.2.3 分析步骤

- 1) 选定分析指标, 并对其进行总和标准化处理;
- 2) 进行主成分分析, 剔除信息重叠, 得到新的指标系统及样本数据;
- 3) 按新主成分组成的样本数据进行聚类分析, 将上市企业按盈利能力划分类别。

3 结果分析

3.1 主成分选取

对表1数据进行总和标准化处理后, 运用主成分分析方法, 首先求得各个主成分的贡献率, 见表2。

表2 特征值、贡献率和累积贡献率

Table 2 Eigenvalues, contribution rate and cumulative contribution

主成分	特征值	贡献率/%	累积贡献率/%
$F_1$	5.101 6	72.88	72.88
$F_2$	1.062 7	15.18	88.06
$F_3$	0.554 4	7.92	95.98
$F_4$	0.182 4	2.61	98.59
$F_5$	0.083 5	1.19	99.78
$F_6$	0.012 7	0.18	99.96
$F_7$	0.002 6	0.04	100.00

从表2可看出, 前2个主成分累积贡献率已达到88.06%, 说明前2个主成分基本包括了全部指标所具有的信息, 根据主成分选取标准, 取前2个主成分作为分析依据, 据此确定这2个主成分的载荷, 见表3。

表3 各主成分载荷

Table 3 The principal components load

$F_1$	$F_2$	$F_1$	$F_2$
0.495 55	0.704 84	0.960 16	-0.195 08
0.883 78	-0.384 07	0.873 20	0.365 83
0.961 37	-0.179 05	0.846 42	0.348 06
0.866 01	-0.305 37		

根据表3, 确定前2个主成分表达式如下:

$$F_1 = -0.4956 * X_1 + 0.8838 * X_2 + 0.9637 * X_3 + 0.8660 * X_4 + 0.9602 * X_5 + 0.8732 * X_6 + 0.8464 * X_7;$$

$$F_2 = 0.7048 * X_1 - 0.3841 * X_2 - 0.1791 * X_3 - 0.3054 * X_4 - 0.1951 * X_5 + 0.3658 * X_6 + 0.3481 * X_7。$$

根据这2个主成分表达式, 重新计算样本数据, 得表4。从表4中可看出, 综合得分在长株潭上市企业平均得分之上的(大于0.12)有16家, 得分大于0的上市企业有14家, 还有6家得分小于0, 得分较少的企业其盈利能力较弱。由以上讨论可知, 直接通过主成分得分结果判定企业盈利能力, 结论过于粗放, 因此, 本文选择聚类分析方法对结果作进一步分析。

表4 盈利能力主成分得分及其排序

Table 4 Profitability of principal component scores and their ranks

编号	上市企业	评价指标			
		$F_1$	$F_2$	综合得分	位次
1	中联重科	0.439 700	-0.058 505	0.311 570	4
2	通程控股	0.128 570	-0.010 062	0.092 171	21
3	华天酒店	0.180 070	0.012 688	0.133 160	15
4	湖南投资	0.365 900	0.114 690	0.284 080	5
5	金德发展	-0.055 023	0.013 043	-0.038 121	31
⋮	⋮	⋮	⋮	⋮	⋮
34	株冶集团	-0.090 947	0.033 533	-0.061 192	33
35	新五丰	0.005 797	0.014 561	0.006 435	30
36	广汽长丰	0.139 300	-0.006 621	0.100 520	18

3.2 聚类分析

取表4中 $F_1$ 和 $F_2$ 列数据, 应用最短距离法进行系统聚类分析。本文采用MATLAB7.0编程实现聚类分析的过程, 得到聚类分析谱系图, 见图1。

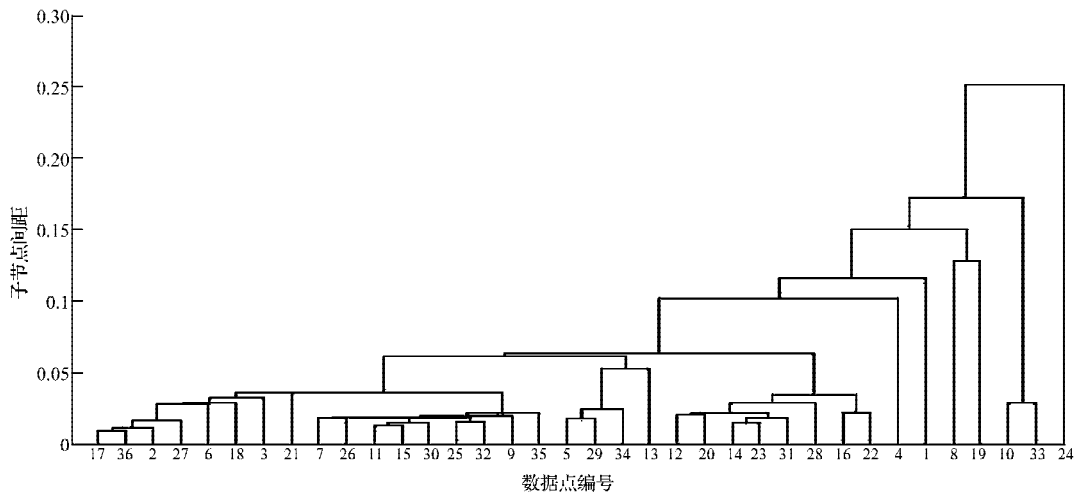


图1 聚类谱系图

Fig. 1 Clustering dendrogram

从图1中可看出,当类间距离小于0.05时,全部上市企业可以分为10类。这种分类比较细腻,便于投资者仔细鉴别,确定投资方向。但对于管理者来讲,这种分类还不够细化,不利于管理决策。本文以上市企业综合得分及位次为依据,做如下分类:当类间距离大于0.15小于0.20时,综合得分在平均得分之上的16家,可以分为2类,即序号为1,3,4,8,12,14,16,18,19,20,21,22,23,28,31为一类,企业盈利能力较强,24为一类,企业盈利能力强;综合得分在平均得分之下,但其值为正的14家上市企业,分为2类,即序号为2,6,17,27,36为一类,企业盈利能力稍强,7,9,11,15,25,26,30,32,35为一类,企业盈利能力一般;综合得分为负分的6家上市企业,5,13,29,34为一类,企业盈利能力较差,10,33为另一类,企业盈利能力差。通过这种分类,为决策者决策提供依据,如确定哪些企业为重点扶持企业,确定财政投资方向等,使其为城市建设带来更好的经济效益,更好地服务于社会。

#### 4 结语

主成分聚类分析可以有效降低分析数据维数,从而降低分析问题的复杂度。应用主成分聚类分析结果建立新的指标体系,对长株潭上市企业盈利能力进行分类,并对结果进行分析,为决策者提供依据,这种方法所得结论在理论上是可行的,在实际应用中有一定参考价值。

#### 参考文献:

- [1] 龙筱刚. 企业盈利能力指标分析[J]. 审计与理财, 2003(12): 76-77.  
Long Xiaogang. The Pointer Analyse of Enterprise's Earning Capacity[J]. Auditing & Finance, 2003(12): 76-77.
- [2] 郭蕊. 上市公司财务报表分析刍议[J]. 山西煤炭科技, 2006(3): 35-36.  
Guo Rui. The First Analysis about the Report of Financial Affairs in Companies which is on Sale[J]. Shanxi Coking Coal Science & Technology, 2006(3): 35-36.
- [3] 张新广. 合理利用财务报表对上市公司进行财务分析[J]. 商业会计, 2005(11): 14-15.  
Zhang Xinguang. Rational Use of Financial Statements for the Financial Analysis of Listed Companies[J]. Commercial Accounting, 2005(11): 14-15.
- [4] 王广斌. 上市公司财务报表分析基本原理和方法[J]. 山西高等学校社会科学学报, 2004, 16(10): 46-47.  
Wang Guangbin. Rationle and Methods of Listed Companies Finance Report Analysis[J]. Social Sciences Journal of Colleges of Shanxi, 2004, 16(10): 46-47.
- [5] 郭宁, 沈剑飞. 上市公司财务报表分析研究[J]. 今日科苑, 2006(4): 10.  
Guo Ning, Shen Jianfei. Analysis of the Financial Statements of Listed Companies[J]. Today Panorama of Modern Sciences, 2006(4): 10.
- [6] 陆璇, 叶俊. 实用多元统计分析[M]. 6版. 北京: 清华大学出版社, 2008.  
Lu Xuan, Ye Jun. Applied Multivariate Statistical Analysis [M]. 6th ed. Beijing: Tsinghua University Press, 2008.
- [7] 徐建华. 现代地理学中的数学方法[M]. 2版. 北京: 高等教育出版社, 2004.  
Xu Jianhua. Mathematical Method in Contemporary Geography[M]. 2nd ed. Beijing: Higher Education Press, 2004.
- [8] 周光辉, 鲍立威. 主成分分析法在卷烟品牌销售系统中的应用[J]. 计算机应用与软件, 2005, 22(5): 57-59.  
Zhou Guanghui, Bao Liwei. The Application of Main-Ingredient Analysis in the System of Cigarette Sales[J]. Computer Applications and Software, 2005, 22(5): 57-59.

(责任编辑: 邓光辉)