

文章编号: 1002- 1566(1999) 05- 0001- 04

聚类分析在股票市场板块分析中的应用^{*}

邓秀勤

(广东韶关大学计算机系, 韶关, 512005)

摘 要

邓秀勤. 聚类分析在股票市场板块分析中的应用.

本文将聚类分析应用于股票市场的研究, 研究实例表明, 聚类分析方法是股市板块分析中的一种有效、实用的方法。

关键词: 聚类分析, 板块, 投资, 系统聚类

中图分类号: O212 F830.9

文献标识码: A

一、问题的提出

随着我国市场经济建设的高速发展, 人们的金融意识和投资意识日益增强, 而作为市场经济的组成部分—股票市场, 正逐步走向成熟与规范, 越来越多的投资者把眼光投向了股票, 历史已经证明股票是一种不仅在过去已提供了投资者可观的长期利益, 并且在将来也将提供良好机遇的投资媒体。然而, 股价涨跌无常, 股市变幻莫测, 投资者要想在股市投资中赢取丰厚的投资回报, 成为一个成功的投资者, 就得认真研究上市公司的历史、业绩和发展前景, 详细分析上市公司的财务状况, 树立以基本分析为主, 技术分析为辅的投资理念, 找出真正具有投资价值的股票, 进行长期投资。

股票投资的基本分析分为宏观分析、中观分析、微观分析三大部分, 宏观分析指对国家的国民经济以及政治、文化的分析, 微观分析指公司分析, 而中观分析主要指行业分析和地区分析等, 板块分析主要归属于中观分析, 兼有微观分析。中国股市从无到有, 发展至今已颇具规模, 前些年, 在中国股市发展的初生阶段, 由于市场规模小, 上市公司数量不多, 加上股民的投资思维和方法不太成熟, 因此, 投机性特强, 这时用不上多少板块分析。但是, 随着股市发展, 投资手法和证券监管方法的成熟, 以及上市公司数量的不断增多, 如果再和以往一样, 面对上千种股票胡乱抓一气, 碰运气, 甚至受各种股评和谣言所左右, 则很难走向理性化, 进而难以最终取得投资成功。因此, 在成熟股市中, 一个股民若想成功, 必须学会板块分析, 习惯理性操作, 树立板块投资理念。

板块是指具有共同特征的股票群。股市中的板块可以行业和产业、地域、时间、特殊题材等多种角度来划分, 而在每一板块中又有几十种甚至上百种股票。面对众多股票及每个公司的众多财务数据, 怎样才能客观、全面、准确地分析并选出各板块及板块内的绩优龙头股和潜力股

* 收稿日期: 1998- 6- 4

呢?笔者用多元统计分析中的聚类分析方法,对此问题作一些探讨。

二、聚类分析在股市板块分析中的应用

系统聚类分析的基本思想是首先将每个样本当作一类,然后根据样本之间的相似程度并类,并计算新类与其它类之间的距离,再选择相近者并类,每合并一次减少一类,继续这一过程,直到所有样本都并成一类为止。在聚类过程中,我们选用欧氏距离来度量类与类之间的相似程度,聚类方法采用类平均法。

我们以高科技板块中的 31 个上市公司为研究对象,分析中选取了这 31 个上市公司 1997 年的每股收益、每股净资产、股东权益率、净资产收益率、净利润率等五个反映上市公司综合盈利能力的指标,数据取自 [4](略)。应用 SAS 软件中的系统聚类过程 CLUSTER 对 31 个样本进行聚类,得到表 1 所示的聚类过程。表 1 中 NCL 为聚类数; Clusters Joined 为每次聚成一个新类的 2 个样品(标有 OB)或旧类(标有 CL); FREQ 为新类中所含有的样品数; SPRSQ 为半偏 R^2 , 它表示每一次合并对信息的损失程度,看这一列的数值可知:从 4 类合并成 3 类时信息损失(为 0.1042)最多,此统计量表明聚成 4 类较合适; CCC 在 $NCL=4$ 时达到唯一的峰值 -2.49, 它支持分 4 类; PSF 为伪 F 统计量, PSF 出现峰值时所对应的分类数较合适,从这一列的数值可知 PSF 在 $NCL=17$ $NCL=12$ $NCL=9$ 和 $NCL=4$ 时 4 次达到峰值,但在 $NCL=4$ 时峰更陡些; PST2 为 t^2 统计量, PST2 出现峰值的前一行所对应的分类数较合适,从这一列的数值可知 $NCL=3$ 时出现峰值 9.8, 它也支持分 4 类。综合这四个统计量可知:将 31 个样本分成 4 类较合适。根据表 1, 作出图 1 所示的聚类谱系图。当我们取分类阈值为 1.0 时, 31 个样本被分成了四类, 各类包含的样本如下:

第一类: OB18(天津磁卡), OB29(燕化高新);

第二类: OB22(深科技), OB15(实达电脑), OB10(清华同方), OB12(东大阿派), OB14(长城电脑), OB31(风华高科), OB17(东方通信);

第三类: OB1(华光科技), OB26(佛山照明), OB4(国脉通信), OB8(工大高新), OB11(振华科技), OB21(彩虹股份);

第四类: OB2(冰箱压缩), OB19(同济科技), OB20(华东电脑), OB9(长安信息), OB23(中科健), OB16(湘计算机), OB30(倍特高新), OB6(南华西), OB28(深圳华强), OB7(厦门信达), OB25(华意压缩), OB27(粤 TCL), OB3(复华实业), OB5(南洋实业), OB13(中国高新), OB24(深华源)。

第一、二类公司在经营规模、经营实力、技术水平等方面具有一定优势, 竞争能力强, 经营业绩优良, 综合财务状况良好, 属高科技板块的绩优龙头股, 颇具发展潜力和长期投资价值, 是高科技板块中投资者的首选投资对象。其中第一类的天津磁卡、燕化高新 1997 年净利润率分别为 45.86% 和 44.32%, 净利润增长率分别是 93.6% 和 96.93%, 其获利能力远远大于其它公司。第四类公司业绩一般, 投资者应谨慎介入, 可多加观望。

三、结束语

聚类分析能综合多项财务指标来反映上市公司的盈利能力和水平, 所得聚类结果与公司的实际财务状况和经营状况相吻合。我们还可以对所选出的各个板块的龙头潜力股再进行聚

类分析 ,找出最具实力的板块龙头股。

表 1 系 统 聚 类 过 程

NCL	Clusters	Joined	FREQ	SPRSQ	CCC	PSF	PST2	DIST
30	OB2	OB19	2	0.0018	.	19.0	.	0.2330
29	OB16	OB30	2	0.0026	.	16.1	.	0.2792
28	OB9	OB23	2	0.0026	.	15.7	.	0.2804
27	OB25	OB27	2	0.0027	.	15.7	.	0.2839
26	OB6	OB28	2	0.0041	.	14.3	.	0.3516
25	CL30	OB20	3	0.0056	.	12.6	3.1	0.3751
24	OB3	CL27	3	0.0058	.	11.7	2.2	0.3880
23	OB1	OB26	2	0.0063	.	11.2	.	0.4340
22	CL25	CL28	5	0.0116	.	9.5	3.4	0.4492
21	OB17	OB31	2	0.0070	.	9.5	.	0.4599
20	CL24	OB5	4	0.0087	.	9.2	2.1	0.4654
19	OB10	OB15	2	0.0075	.	9.4	.	0.4744
18	OB4	OB8	2	0.0082	.	9.5	.	0.4952
17	OB11	OB21	2	0.0082	.	9.7	.	0.4965
16	CL26	CL29	4	0.0133	.	9.4	4.0	0.4998
15	CL20	OB7	5	0.0111	.	9.5	1.9	0.5228
14	CL23	CL18	4	0.0143	.	9.5	2.0	0.5680
13	OB14	CL21	3	0.0129	.	9.7	1.8	0.5857
12	OB13	OB24	2	0.0141	.	9.9	.	0.6512
11	CL22	CL16	9	0.0427	.	8.5	7.2	0.6546
10	OB12	CL13	4	0.0185	.	8.8	1.9	0.6848
9	CL19	OB22	3	0.0224	.	9.1	3.0	0.7485
8	CL11	CL15	14	0.0868	.	7.0	9.2	0.7943
7	OB18	OB29	2	0.0213	.	7.8	.	0.7989
6	CL14	CL17	6	0.0492	- 3.12	7.8	5.3	0.8501
5	CL8	CL12	16	0.0573	- 2.85	8.0	3.8	0.9007
4	CL9	CL10	7	0.0703	- 2.49	8.4	5.1	0.9533
3	CL6	CL5	22	0.1745	- 3.55	6.2	9.8	1.0341
2	CL3	CL4	29	0.1656	- 3.49	4.8	6.7	1.0618
1	CL2	CL7	31	0.1428	- 0.00	.	4.8	1.3179

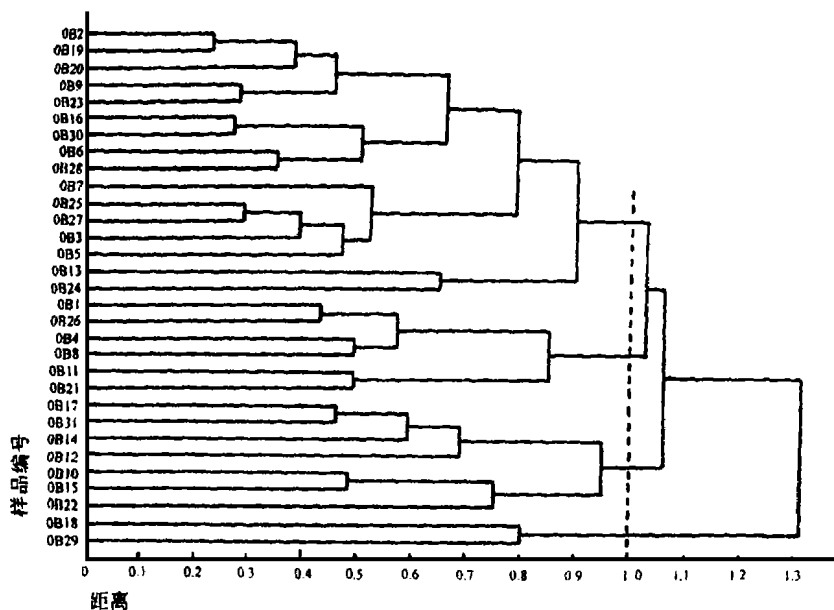


图 1 样本聚类分析的动态过程

参考文献

- [1] 何剑著 (1997),《深沪股市板块分析》,广州出版社,广州.
- [2] 武冠卓等编著 (1997),《深沪股市公司分析》,广州出版社,广州.
- [3] 方开泰 (1989),《实用多元统计分析》,华东师范大学出版社,北京.
- [4] 信息早报社编 (1998),《98选股投资手册: 上市公司 97年报大全》,中国档案出版社,北京.
- [5] 胡良平主编 (1996),《现代统计学与 SAS应用》,军事医学科学出版社,北京.
- [6] 高惠璇等编 (1995),《SAS系统与基础统计分析》,北京大学出版社,北京.

Application of Cluster analysis in stock market board analysis

DENG Xiu-qin

(Dept. of Computer, Shaoguan University, 512005)

Abstract

The cluster analysis has been applied in study of stock market. It point out that cluster analysis is an effective and practical way in stock market analysis.

Key words Cluster Analysis; Board; Investment; System.