

基于稀疏因子分析及熵值法的股票综合评价

摘要：本文采用稀疏因子分析及熵值法对多元金融板块的股票进行评价，通过因子得分排序筛选出 5 只未来走势良好的潜力股票。并利用 4 月 16 至 4 月 28 日后的数据查看趋势，采用平均增长率来测度趋势。

一、改进措施

在因子分析模型中，传统的估计方法是用极大似然法对模型进行估计，然后利用旋转法求稀疏因子载荷。张等^[1]采用非凸惩罚似然稀疏估计的方法对因子分析进行分析，通过选择正则化参数，产生比最大似然估计旋转技术更稀疏的解，从而增强其解释性。

张等在进行因子得分计算的时候采用的是对各公因子的方差贡献率占方差累积贡献率的比重作为权重，可以构造综合评模型。在实际评价工作中，这样容易导致各公共因子权重两极分化，高估某一公共因子对评价结果的影响。

而本文考虑采用熵值法计算权重进而构造综合评价模型。根据熵的特性，可以通过计算熵值来判断某个指标的离散程度，指标的离散程度越大，该指标对综合评价的影响（权重）越大，测度不同公共因子对综合评价的影响程度，因此运用此方法赋权更加客观。

熵值法计算步骤：

1、异质指标同质化

$$\text{正向指标: } x'_{ij} = \frac{x_{ij} - \min\{x_{ij}, \dots, x_{nj}\}}{\max\{x_{ij}, \dots, x_{nj}\} - \min\{x_{ij}, \dots, x_{nj}\}}$$

$$\text{负向指标: } x'_{ij} = \frac{\max\{x_{ij}, \dots, x_{nj}\} - x_{ij}}{\max\{x_{ij}, \dots, x_{nj}\} - \min\{x_{ij}, \dots, x_{nj}\}}$$

2、计算第 j 项指标下第 i 个样本值占该指标的比重 $p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}, i = 1, 2, \dots, n, j = 1, 2, \dots, m$

3、计算第 y 项指标熵值 $e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln p_{ij}$

4、计算第 y 项指标权重 $w_j = \frac{1 - e_j}{\sum_{j=1}^m (1 - e_j)}$

本文模型的求解步骤：1、通过 KMO 及 Bartlett 检验是否适合进行因子分析；2、对数据进行标准化处理；3、通过特征值排序图确定因子个数，此时特征值划分位置设置位 0；4、进行稀疏因子分析，得到因子载荷矩阵；5、回归法计算各个稀疏因子得分；6、通过熵值法计算各个稀疏因子的权重；7、计算最后得分进行排名。

二、数据来源与采集时间

为了能在相对长的一段时间内持有股票获得收益，本文选取企业财务指标作为因子分析中的变量。数据来自西南金点子财富终端的 2021 年 4 月 16 日多元金融板块的股票共 28 只。选取总资产、流动负债、营业利润、净利润、利润同比、收入同比、每股收益、权益比、净收益率。由于 4 月 16 日前有企业更新了财务指标，且在 4 月 28 日有企业更新了第一季度的财务指标，所以本模型仅用作 4 月 16 日至 4 月 28 日期间的潜力股评价。以 4 月 16 日至 4 月 28 日收盘价的评价增长率作为增长趋势的衡量。

三、R 程序分析结果

表 1 KMO 和 Bartlett 检验

检验方法		检验值
KMO 检验		0.617
Bartlett	近似卡方	299.239

自由度	36
p 值	0.000

KMO 检验的值大于 0.5，Bartlett 检验的 p 值小于 0.05，该数据适合做因子分析。

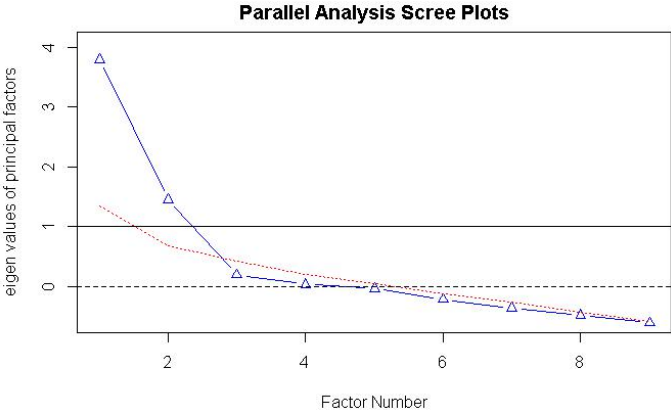


图 1 特征值排序图

以 0 为限，将因子个数确定为 4 个。

表 2 因子载荷

变量	F1	F2	F3	F4
总资产	-0.059	-0.231	0	0.927
流动负债	0	0	0.041	0.978
营业利润	0.008	0.0169	0.496	0.798
净利润	0	0	0.706	0.606
利润同比	0	0	0.654	0
收入同比	-0.132	0	0	0
每股收益	-0.191	0	0.847	0
权益比	-0.151	0.519	0	-0.146
净益率	-0.854	0	0.346	0

通过稀疏因子分析得到因子载荷，此处使用 BIC 准则，参数 $\gamma = 2$ ， $\rho = 0.281$ 。

表 3 不同方法的因子得分及平均增长率排名前 5 的股票

排名	熵值法		方差贡献率		平均增长率	
	股票名称	得分	股票名称	得分	股票名称	值
1	中油资本	0.738	中油资本	2.437	ST 熊猫	0.002478
2	五矿资本	0.579	吉艾科技	0.768	南华期货	0.001832
3	中航资本	0.573	中航资本	0.506	宏盛科技	0.001628
4	越秀金控	0.536	越秀金控	0.375	越秀金控	0.001430
5	江苏租赁	0.439	五矿资本	0.275	江苏租赁	0.000746

熵权法得到的潜力股排名前五有两个实际表现不错，传统方法的选出的潜力股仅有一个实际表现排名前五。

四、结论与展望

采用本文提出的稀疏化因子分析及熵权法效果并未达到预期。下一步可以选取新的指标，如涨幅、量比等，从时间维度上进行动态对比与分析。

参考文献：

[1]张夏伟,刘惠篮.基于稀疏因子分析的贵州省城市通信水平综合评价[J].数学理论与应用,2019,39(Z1):78-89.