

# 基于数据分析下的中小微企业的信贷决策

曹博文 杨子钰 陈瑞斌

(华北理工大学 河北 唐山 063210)

摘要: 银行向一些缺少抵押资产的中小微企业进行信贷, 既需要考虑中小微企业的资质风险, 又要考虑到银行的利益, 因此建立一套信贷决策方法, 通过主成分分析法对数据进行了分析, 在通过综合评价法对各公司进行了信贷评价。最终帮助银行得出了决策方法。

关键字: 主成分分析; 风险评估模型; 信贷风险;

Credit decision of small, medium and micro enterprises based on data analysis

Cao Bowen Yang Ziyu Chen Ruibin

(North China University of Technology, Tangshan, Hebei 063210)

Abstract: When banks extend credit to some small, medium and micro enterprises that lack mortgage assets, it is necessary to consider not only the qualification risks of small and medium-sized enterprises, but also the interests of the bank. Therefore, a set of credit decision-making methods is established, and the data is analyzed by principal component analysis. The credit evaluation of each company is carried out through the comprehensive evaluation method. Ultimately helped the bank to arrive at a decision-making method.

Keywords: principal component analysis; risk assessment model; credit risk;

## 一、问题分析

我们通过分析指定出四个影响因素, 并运用主成分分析法对数据进行改进, 并建立了构建评价模型的指标体系<sup>[1]</sup>。接着通过建立综合评价模型, 得出四个主要成分的权重。按照求解出的 4 个主成分的权重和中企业与信誉等级的相关数据, 建立了一个风险等级的评估模型。

## 二、分级综合评价指标体系的构建

首先对信贷风险指标进行遴选<sup>[2]</sup>。对于信贷风险而言, 虽然每种指标在成因上互不相同, 但是不同的指标之间往往具有相关性, 根据附件一给出的相关数据, 我们将公司的利润、利润率、供求关系比、交易稳定值四种变量规定为影响信贷风险的主要因素。其次为了找到这些潜在因素以及相应的支配作用, 选用主成分分析法对这些问题做以解决, 步骤如下<sup>[3]</sup>:

Step1 为消除不同变量的量纲的影响, 首先需要对变量进行标准化中处理:

本题涉及到的指标共 4 个, 样本对象 123 个, 设第 j 个样本的第 i 个指标值为  $F_{ij}$ , 将各标准化值按如下方式进行标准化为  $\widetilde{F}_{ij}$ :

$$\widetilde{F}_{ij} = \frac{F_{ij} - \bar{F}_i}{s_i} \quad (1)$$

其中  $\bar{F}_i$  和  $s_i$  分别为 i 指标的均值和标准差。标准化的目标在于消除不同变量的量纲之间的影响, 而且标准化转换并不会影响变量的相关系数。

Step2 计算标准化数据的相关系数矩阵, 求出该相关系数矩阵的特征向量与特征值。

$$r_{ii'} = \frac{\sum_{k=1}^4 \widetilde{F}_{ik} \widetilde{F}_{i'k}}{4-1}, i, i' = 1, 2, 3, 4 \quad (2)$$

设第 i 个指标与第 i' 个指标的相关系数为  $r_{ii'}$ , 选择相关系数矩阵为  $R = (r_{ii'})_{4 \times 4}$ , 其中,  $r_{ii'} = 1, r_{ii'} = r_{i'i}$ 。

Step3 计算特征向量和特征值。

计算相关系数矩阵 R 的特征值  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 \geq 0$ , 以及其相应的特征向量  $\xi_1, \xi_2, \xi_3, \xi_4$ , 其中  $\xi_1 = (\mu_{11}, \mu_{21}, \mu_{31}, \mu_{41})^T$ , 由特征向量组成 4 个信贷指标变量:

$$\begin{cases} Y_1 = \mu_{1,1} \bar{F}_1 + \mu_{2,1} \bar{F}_2 + \mu_{3,1} \bar{F}_3 + \mu_{4,1} \bar{F}_4 \\ Y_2 = \mu_{1,2} \bar{F}_1 + \mu_{2,2} \bar{F}_2 + \mu_{3,2} \bar{F}_3 + \mu_{4,2} \bar{F}_4 \\ Y_3 = \mu_{1,3} \bar{F}_1 + \mu_{2,3} \bar{F}_2 + \mu_{3,3} \bar{F}_3 + \mu_{4,3} \bar{F}_4 \\ Y_4 = \mu_{1,4} \bar{F}_1 + \mu_{2,4} \bar{F}_2 + \mu_{3,4} \bar{F}_3 + \mu_{4,4} \bar{F}_4 \end{cases} \quad (3)$$

其中,  $Y_i$  为第 i 主成分,  $i=1, 2, 3, 4$ .

计算出特征向量为如下表:

表一、各因子的特征向量

因子 1	因子 2	因子 3	因子 4
-0.4577	0.4075	-0.0366	0.7276
-0.2109	0.7386	-0.0665	-0.3172
-0.4605	-0.0667	0.8309	-0.2702
0.4885	0.5228	0.1222	-0.2493

Step4 确定 x 个主成分, 进行统计分析。

根据上述步骤, 本文利用 python 编程软件以及 SPSS 统计软件<sup>[4]</sup>, 求得各公司的收入总和、支出总和、利润率、供求关系比、交易稳

定值等一系列数据, 其中  $\text{供求关系比} = \frac{\text{进货商家总量}}{\text{出货商家总量}}$ , 当供求关系比=1 时, 则证明该公司供求关系最稳定;

$$\text{交易稳定值} = \frac{\text{有效发票数}}{\text{总发票数量}}$$
，交易稳定值越接近 1，则证明

该公司交易失败的可能性越小；  

$$\text{利润率} = \frac{\text{净利润}}{\text{总利润}}$$
，利润率越高，说明该公司发展潜力更大。

做完上述准备后，我们利用 matlab 编程软件，求得各指标的相关性系数表：

表二、各指标相关性系数

	因子 1	因子 2	因子 3	因子 4
因子 1	1	1.165489	1.657066	1.91592
因子 2	0.858009	1	1.421778	1.643877
因子 3	0.603476	0.703345	1	1.156212
因子 4	0.521942	0.608318	0.864893	1

从表中不难发现，某些指标具有很强的相关性，如果直接用这些个指标对信贷风险进行分级，不仅会使得计算量变大，同时还会发生信息的重叠，影响分级的结果客观性。而主成分分析可以把多个指标转化成少数的几个不相关的指标。例如下表：

表三、信贷风险的特征值以及累计贡献率

主成分	特征值	方差	累计贡献率
1	1.7874	0.3575	0.3575
2	1.3501	0.2700	0.6275
3	0.8071	0.1614	0.7889
4	0.6919	0.1384	0.9273

在累计方差为 92.73%的前提下得到了 4 个主成分，这 4 个主成分满足了主成分分析原则。从上表还可以看出，主成分 1 和 2 的累计贡献率较大，这就可以结束为主成分 1 和主成分 2 可能是信贷风险分级的最重要指标。

利用 matlab 软件计算得到的主成分分析正交解如下表所示：

表四、信贷风险主成分分析正交解

	主成分			
	1	2	3	4
净利润	-0.9311	0.6434	0.0010	1.1199
利润率	-0.0877	0.8606	0.2676	0.9367
供求关系比	1.6796	-3.4701	0.3548	0.5198
交易稳定值	-1.7584	0.1900	0.3864	0.4725

### 三、结果分析

1. 四个主要成分的权重从大到小依次为：净利润>供求关系比>交易稳定值>利润率。

2. 纯利润值越大说明该公司承担能力越强，信贷风险更低，当信贷总值固定时，可以给予其较高的信贷额度；反之则给予其较低的信贷额度。

3. 客户流失率与信誉等级呈负相关关系，与银行年利率成正相关。

从而建立一个风险评价指标，对要进行信贷的中小微企业的公

司情况进行一系列的评估，得到风险评价模型，并对企业的相关成分进行评级。

风险评级一共分为 4 等，对应的表示分别为 A-D。根据求解出的 4 个主成分的权重，可以建立一个风险等级的评估模型，如下式：

$$\Phi_1 = 23.86 + 3.42x_1 + 1.66x_2 + 3.29x_3 + 0.51x_4 \quad (4)$$

其中， $x_1, x_2, x_3, x_4$  分别对应：纯利润、供求关系比、交易稳定值、利润率。

信誉等级的评定与中小微企业的四项指标有着密切的关系，其中，纯利润可以反应这个公司的盈利情况；利润率反映了这个公司的潜在竞争力；供求关系比可以反应这个公司的供销是否平衡，会不会出现供大于求或者供小于求的情况；交易稳定值可以反应这个公司交易稳定程度进而反应公司是否存在破产风险。

为了更直观的表现出某一个公司的信贷风险情况，建立了信贷风险评价模型，该模型的四个等级可以用所得到的分数来直观的表现出来。将满分记为 100 分，最低分记为 0 分，可以根据 4 个等级划分出得分的具体情况，即为：

$$\begin{cases} A, 95 \leq L \leq 100 \\ B, 80 \leq L < 95 \\ C, 60 \leq L < 80 \\ D, L < 60 \end{cases} \quad (5)$$

其中 L 为该企业信誉等级评分。

为检验以上模型的准确性，将附件一中的企业的各指标值代入公式(4)，得到  $\Phi_1$ ；再将  $\Phi_1$  代入公式(5)，得到正确率 P。以 60 分为分界值，根据计算结果判定该企业属于低风险组还是高风险组。

该银行在年度信贷总额为 1 亿元时对这些企业的信贷策略为：

(1) 当纯利润极高且信誉评分为 A 时，可考虑贷款额为 80-100 万，年利率为 11%-15%；

(2) 当纯利润高且信誉评分为 B 时，可考虑贷款额为 80-90 万，年利率为 6%-12%；

(3) 当纯利润一般且信誉等级为 C 时，可考虑贷款额为 10-60 万，年利率为 4%-6%。

(4) 当纯利润为负或信誉等级为 D 时，不考虑贷款。

### 参考文献：

- [1]李帅辰.A 农村商业银行信贷风险控制研究[D].南昌大学,2019.
- [2]朱苗苗.商业银行小微企业信贷风险与防范管理研究[J].现代营销(信息版),2020,(07):76-77.
- [3]梁佳荣.商业银行小微企业信贷风险防范管理研究[J].全国流通经济,2020,(06):150-151.DOI:10.16834/j.cnki.issn1009-5292.2020.06.069
- [4]李帅鹏.基于贝叶斯决策规则的商业银行信贷风险研究[D].淮北师范大学,2020.