

我国上市银行经营业绩影响因素分析

吕苏琪

云南财经大学, 云南 昆明

收稿日期: 2022年7月9日; 录用日期: 2022年7月20日; 发布日期: 2022年8月2日

摘要

经营业绩体现银行经营活动所带来的财务状况和经营成果, 它的主要依据是会计报表, 例如, 财务状况主要以资产负债表、现金流量表及相关附注为基础, 经营成果主要以利润表等损益类报表及相关附注为基础。本文首先利用多元线性回归分析法建立银行经营业绩评价分析模型, 然后对回归方程、回归系数进行检验, 并对回归方程进行诊断, 从实验结果可以看出在显著性水平0.05下, 回归方程显著, 但回归系数只有 x_1, x_3 显著, 并且回归方程存在多重共线和异方差。因此, 利用主成分回归对回归方程进行改进。利用主成分分析法选择4个综合指标(包含原始变量92.555%)代替原始变量, 然后对这4个指标做回归分析, 主成分回归方程的所有回归系数均显著, 并且回归方程不存在多重共线性、异方差和自相关性。

关键词

上市银行, 经营业绩, 多元线性回归分析, 主成分回归

Analysis on the Factors Affecting the Operational Performance of China Listed Banks

Suqi Lv

Yunnan University of Finance and Economics, Kunming Yunnan

Received: Jul. 9th, 2022; accepted: Jul. 20th, 2022; published: Aug. 2nd, 2022

Abstract

Operating performances reflect the financial status and operating results brought about by the bank's operating activities. Its main basis is the accounting statements. For example, the financial status is mainly based on the balance sheet, the cash flow statement and related notes, and the

operating results are mainly based on the income statement, etc. profit and loss statements and related notes. In this paper, the multiple linear regression analysis method is used to establish the analysis model of bank business performance evaluation, and then the regression equation and regression coefficient are tested, and the regression equation is diagnosed. It can be seen from the experimental results that at the significance level of 0.05, the regression equation is significant, but the regression coefficient is only x_1, x_8 significant, and the regression equation has multicollinearity and heteroscedasticity. Therefore, the regression equation is improved by using principal component regression. Use principal component analysis to select 4 comprehensive indicators (including 92.555% of the original variables) to replace the original variables, and then perform regression analysis on these 4 indicators, all regression coefficients of the principal component regression equation were significant, and the regression equation did not have multicollinearity, heteroscedasticity and autocorrelation.

Keywords

Listed Bank, Operating Performance, Multiple Linear Regression Analysis, Principal Component Regression

Copyright © 2022 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. 引言

1.1. 研究背景

2020 年以来, 新冠肺炎疫情蔓延对各国经济形成巨大冲击, 全球经济总体上已经陷入衰退, 当前中国经济发展面临前所未有的挑战。国际形势严峻复杂, 我国统筹疫情防控和经济社会发展, 经济运行稳定恢复, 成为全球唯一实现经济正增长的主要经济体, “三大攻坚战”取得决定性成就。各银行主动融入国家发展大局, 始终保持战略定力, 在这个极不平凡的年份取得了极不寻常的成绩。为应对疫情冲击, 我国货币、财政、产业政策多管齐下, 加大逆周期调节力度。随着防范化解金融风险 and 深化金融改革开放的各项措施深入推进, 年初以来快速变化、错综复杂形势带来的问题和挑战得到有效应对。我国银行业整体运行稳健, 风险可控, 服务实体经济能力不断提升。截至 2020 年 6 月末, 我国银行业境内总资产 301.5 万亿元, 同比增长 9.8%。上半年人民币贷款增加 12.09 万亿元, 同比多增 2.42 万亿元。银行业金融机构主要经营和监管指标处于合理区间。银行业金融机构紧扣“六稳”“六保”要求, 把稳企业保就业和服务民营、小微企业更好结合起来, 全力以赴促进经济社会恢复正常循环, 坚定不移推动经济高质量发展。但也必须看到, 受新冠肺炎疫情等因素影响, 在未来一段时期内不良资产上升压力、部分中小金融机构公司治理以及部分市场乱象有所反弹等潜在风险和 challenge 依然较大。在危机中育新机, 于变局中开新局[1]。

金融在社会经济不断发展的过程中, 尤其是当今这个通讯、信息高度发达的时代, 重要程度尤为突出。在我国, 银行在金融业占据主导地位。随着实体经济的转型, 我国银行产业的分化越来越明显。近年来, 银行业的门槛进一步降低, 大量竞争对手涌入, 进一步加剧了国内银行业所面临的竞争。我国银行业想实现可持续发展的目标, 就必须进行客观分析, 在评价自身经营状况的前提下, 提升经营业绩。

1.2. 研究意义

业务情况、管理水平的高低、资本结构的完善程度以及抗风险的能力都将影响银行的利润。

随着金融业的不断发展,我国上市银行不仅面临国内银行业的竞争,还要面临外资金业业的巨大挑战。良好的经营业绩,不但关系到股东的切身利益,而且关系到商业银行的竞争与发展[2]。

业绩是一个公司自身盈利水平的保证,业绩的好坏会直接影响公司的发展,同时业绩好的公司也意味着其具备更高的发展潜力。因此本文对 16 家上市银行 2020 年的数据进行分析,建立回归分析,分析对银行业绩有影响的因素。

1.3. 国内外研究现状

曹晓俊选取 13 家上市银行 2014 年 9 月的有关财务指标,利用主成分分析法、因子分析法、聚类分析法分析我国银行业的发展状况,然后将 13 家银行分为高效益型、低效益型、潜在发展型银行,并根据三类银行的经营特点和存在的问题提出相应措施[3];屈磊利用因子分析法比较我国 5 家国有商业银行的经营业绩,指出股权结构的多样化与银行的综合业绩有一定的关联,根据因子得分对五家商业银行的经营业绩进行排名[4];古文浩利用层次分析法构造具体的评价模型和公式,再利用这个评价模型和公式分析具体的商业银行,通过合理的指标评分办法和权重分配来计算并分析银行经营业绩,并且从定性和定量两方面来考虑各个因素对银行业绩的影响[5];杨建海利用指标评价分析法,根据对商业银行经营业绩的影响程度,将其分为基础性业绩和非基础性业绩,并将两者结合得到综合业绩,再建立上市银行的经营业绩评价体系,并且对上市银行和商业银行的评价体系进行比较分析[6];孙红彦、吴书广、赵涛基于因子分析法对我国商业银行综合业绩进行评价,将盈利性、流动性和风险(安全)三方面纳入因子分析中,对我国商业银行经营业绩进行分析[7]。赵萍利用因子分析对我国 25 家上市银行经营业绩进行分析,并提出针对性意见[8]。冯红梅运用向量自回归模型客观评价互联网金融对商业银行经营业绩的短期和长期影响[9]。刘志颖运用聚类分析和判别分析对 16 家上市银行进行业绩上的分类,运用 R 型聚类分析对财务报表各个指标值进行分类研究,运用主成分分析对 16 家上市银行业绩进行排名和对比[10]。

1.4. 研究内容及创新点

为研究 16 家上市银行的经营业绩及影响因素,本文利用多元线性回归进行分析,但多元线性回归方程系数不显著,并且回归方程存在异方差和多重共线性,并且变量所解释的含义与实际意义相违背。因此利用主成分回归首先对数据进行降维,用 4 个综合指标代替原始指标,再对这四个指标做多元线性回归进行分析,此时所有系数通过显著性检验,并且方程不存在多重共线性、异方差和自相关性。

本文的创新点在于利用主成分回归法对模型进行改进,克服了多元线性回归模型的缺陷。

1.5. 指标选择

本文研究的是我国的 16 家上市银行 2020 年的经营业绩,数据来源于新浪财经网各银行 2020 年年报。各指标介绍见表 1。

Table 1. Introduction of indicators

表 1. 指标介绍

符号	指标含义
X_1	营业收入(百万元)
X_2	每股收益(元)
X_3	净资产增长率(%)
X_4	股本报酬率(%)

Continued

X_5	不良贷款率(%)
X_6	股东权益比例(%)
X_7	业务及管理费用(百万元)
X_8	总资产净利润率(%)
X_9	净利润增长率(%)
y	净利润(百万元)

2. 理论知识

2.1. 多元线性回归模型

2.1.1. 多元线性回归模型的一般形式

设随机变量 Y 与一般变量 X_1, X_2, \dots, X_p 的线性回归模型为:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (2-1)$$

其中, $\beta_0, \beta_1, \dots, \beta_p$ 是 $p+1$ 个未知量, β_0 称为回归常数, β_1, \dots, β_p 称为回归系数, Y 称为被解释变量(因变量), 而 X_1, X_2, \dots, X_p 是 p 个可以精确测量的一般变量, 称为解释变量(自变量) [11]。当 $p=1$ 时, (2-1) 式即为一元线性回归模型, 当 $p \geq 2$ 时, (2-1) 式就称为多元线性回归模型, ε 是随机误差 [11]。对随机误差的假定:

$$\begin{cases} E(\varepsilon) = 0 \\ \text{var}(\varepsilon) = \sigma^2 \end{cases}$$

称

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

为多元线性回归方程。要得到多元线性回归模型, 首先要估计未知参数 $\beta_0, \beta_1, \dots, \beta_p$ 。由于参数估计是基于样本数据的, 由此得到的参数只是参数真值 $\beta_0, \beta_1, \dots, \beta_p$ 的估计值, 记为 $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ [11]。于是有

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_p X_p$$

对于一个实际问题, 如果获得 n 组观测数据 $(x_{i1}, x_{i2}, \dots, x_{ip}, y_i) (i=1, 2, \dots, n)$, 则线性方程(2-1)可以表示为:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p} + \varepsilon_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p} + \varepsilon_2 \\ \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \varepsilon_n \end{cases}$$

其矩阵形式为

$$y = X\beta + \varepsilon$$

其中,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}。$$

2.1.2. 多元线性回归模型的基本假定

对回归方程(2-1)式有如下一些基本假定:

- 1) 所有解释变量之间互不相关[11]。
- 2) 随机误差项具有零均值和同方差, 即

$$\begin{cases} E(\varepsilon_i) = 0, i = 1, 2, \dots, n \\ \text{Cov}(\varepsilon_i, \varepsilon_j) = \begin{cases} \sigma^2, i = j \\ 0, i \neq j \end{cases} i, j = 1, 2, \dots, n \end{cases}$$

这个假定常称为高斯 - 马尔科夫条件。

- 3) 正态分布的假定条件为

$$\begin{cases} \varepsilon_i \sim N(0, \sigma^2), i = 1, 2, \dots, n \\ \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \text{ 相互独立} \end{cases}$$

对于多元线性回归的矩阵形式(2-1)式, 这个条件可以表示为

$$\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$$

随机向量 y 遵从 n 维正态分布, 且

$$\begin{aligned} E(\mathbf{y}) &= \mathbf{X}\boldsymbol{\beta} \\ \text{var}(\mathbf{y}) &= \sigma^2 \mathbf{I}_n \end{aligned}$$

因此,

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n)$$

2.2. 主成分回归

主成分的基本思想

主成分分析的基本思想是降维, 在损失很少信息的前提下把多个指标利用正交旋转变换转化为几个综合指标的多元统计分析方法。通常把转化生成的指标成为主成分, 其中每个主成分都是原始变量的线性组合, 且各个主成分之间不相关[7]。

设对某一事物的研究涉及 p 个指标, 分别用 X_1, X_2, \dots, X_p 表示, 这 p 个指标构成的 p 维随机向量 $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ 。设随机向量 \mathbf{X} 的均值为 $\boldsymbol{\mu}$, 协方差矩阵为 $\boldsymbol{\Sigma}$ [11]。

对 \mathbf{X} 进行线性变换, 可以形成新的综合变量, 用 Y 表示, 也就是说新的综合指标可以由原来的变量线性表示, 即满足下式:

$$\begin{cases} Y_1 = \mu_{11}X_1 + \mu_{12}X_2 + \dots + \mu_{1p}X_p \\ Y_2 = \mu_{21}X_1 + \mu_{22}X_2 + \dots + \mu_{2p}X_p \\ \vdots \\ Y_p = \mu_{p1}X_1 + \mu_{p2}X_2 + \dots + \mu_{pp}X_p \end{cases}$$

由于可以任意的对原始变量进行上述线性变换, 得到的综合变量 Y 的统计特性也不尽相同, 因此为了取得较好的效果, 我们总是希望 $Y_i = \boldsymbol{\mu}_i' \mathbf{X}$ 的方差尽可能大且各 Y_i 之间相互独立[7]。

3. 多元线性回归分析法

3.1. 建立回归方程

本文选取 9 个解释变量研究上市银行 2020 年的利润, 解释变量为: X_1 (营业收入)、 X_2 (每股收益)、

X_3 (净资产增长率)、 X_4 (股本报酬率)、 X_5 (不良贷款率)、 X_6 (股东权益比例)、 X_7 (业务及管理费用)、 X_8 (总资产净利润率)、 X_9 (净利润增长率)。本文选取 2020 年新浪财经网中我国 16 家上市银行 2020 年的数据, 以银行净利润作为因变量, 以如上 9 个变量作为因变量做多元线性回归分析。

由 SPSS25.0 得到如下结果:

Table 2. Analysis of variance table

表 2. 方差分析表

模型	平方和	自由度	均方	F	显著性
回归	147438147512.741	9	16382016390.305	343.630	0.000
残差	286040422.889	6	47673403.815		
总计	147724187935.630	15			

由表 2 方差分析表知, 显著性即 P 值小于 0.05 (本文的显著性水平取为 0.05), 则在显著性水平 α 下, y 与 x_1, x_2, \dots, x_9 有显著的线性关系, 即回归方程显著。

Table 3. Regression equation coefficients

表 3. 回归方程系数

模型	未标准化系数		标准化系数	t	显著性
	B	标准错误	Beta		
(常量)	-92332.580	53574.644		-1.723	0.136
x_1	0.392	0.050	1.057	7.880	0.000
x_2	6519.938	16001.811	0.062	0.407	0.698
x_3	-514.734	584.971	-0.031	-0.880	0.413
x_4	-105.491	130.511	-0.126	-0.808	0.450
x_5	5021.615	13705.596	0.016	0.366	0.727
x_6	3813.421	5900.353	0.020	0.646	0.542
x_7	-0.238	0.181	-0.158	-1.317	0.236
x_8	85032.367	29803.769	0.151	2.853	0.029
x_9	180.079	267.592	0.018	0.673	0.526

由表 3 回归方程系数可以得到, y 对 9 个自变量的回归方程为:

$$\hat{y} = -92332.580 + 0.392 * x_1 + 6519.938 * x_2 - 514.734 * x_3 - 105.491 * x_4 + 5021.615 * x_5 + 3813.421 * x_6 - 0.238 * x_7 + 85032.367 * x_8 + 180.079 * x_9$$

由于自变量 x_1, x_2, \dots, x_9 的量纲不相同, 数据差异往往很大, 不利于在同一标准上进行比较, 因此为消除量纲和数量级的不同所带来的影响一般对数据进行标准化处理。则标准化的回归方程为:

$$\hat{y}^* = 1.057 * x_1^* + 0.62 * x_2^* - 0.031 * x_3^* - 0.126 * x_4^* + 0.016 * x_5^* + 0.020 * x_6^* - 0.158 * x_7^* + 0.151 * x_8^* + 0.018 * x_9^*$$

从标准化回归方程可以看出, 对经营业绩影响最大的是营业收入 (X_1), 其次是业务及管理费用 (X_7)。

且 $x_1, x_2, x_5, x_6, x_8, x_9$ 对银行经营业绩起正影响, x_3, x_4, x_7 对银行经营业绩起负影响。这与定性分析的结果不一致, 可能是变量之间存在相关关系。

从表 3 可以看出其中一些变量的回归系数的显著性检验不能通过, 回归方程并不理想, 因此对模型进行诊断和优化。

3.2. 回归方程的诊断

3.2.1. 多重共线性的诊断

由 SPSS25.0 得到如下结果:

Table 4. Collinearity diagnostics

表 4. 共线性诊断

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准错误	Beta			容差	VIF
(常量)	-92332.580	53574.644		-1.723	0.136		
x_1	0.392	0.050	1.057	7.880	0.000	0.018	55.764
x_2	6519.938	16001.811	0.062	0.407	0.698	0.014	71.398
x_3	-514.734	584.971	-0.031	-0.880	0.413	0.254	3.933
x_4	-105.491	130.511	-0.126	-0.808	0.450	0.013	75.793
x_5	5021.615	13705.596	0.016	0.366	0.727	0.173	5.779
x_6	3813.421	5900.353	0.020	0.646	0.542	0.336	2.979
x_7	-0.238	0.181	-0.158	-1.317	0.236	0.022	44.487
x_8	85032.367	29803.769	0.151	2.853	0.029	0.115	8.677
x_9	180.079	267.592	0.018	0.673	0.526	0.450	2.220

由表 4 可以看出, X_1, X_2, X_4, X_7 的方差扩大因子较大, 分别为 $VIF_1 = 55.764$, $VIF_2 = 71.398$, $VIF_4 = 75.793$, $VIF_7 = 44.487$, 大于 10, 说明回归方程存在多重共线性。

3.2.2. 异方差性的检验

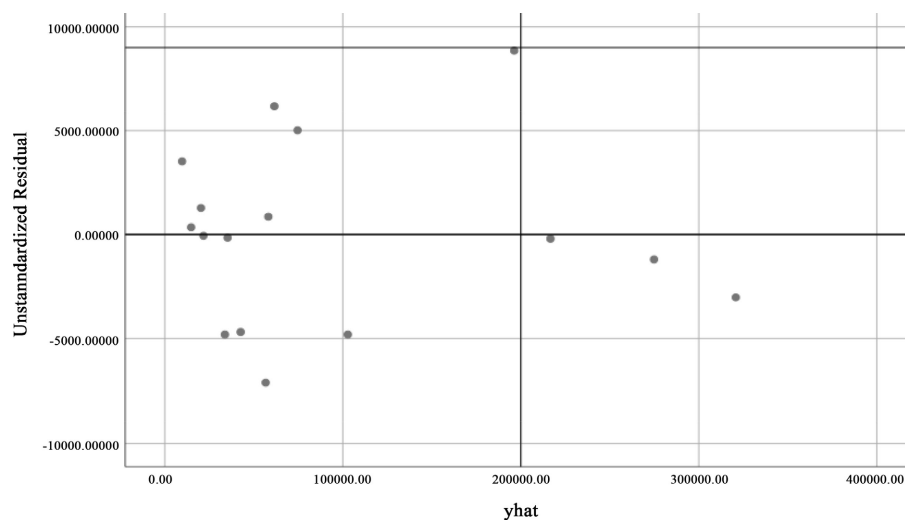


Figure 1. Residual plot

图 1. 残差图

以因变量回归值 \hat{y} 为横轴，以残差 e_i 为纵轴画散点图记为残差图，由图 1 残差图可以看出，残差有扩张的趋势，不满足基本假设，从而可以判断模型存在异方差。

3.2.3. 自相关性的检验

Table 5. DW inspection
表 5. DW 检验

模型	R	R方	调整后R方	标准估算的错误	德宾 - 沃森
1	0.999	0.998	0.995	6904.593	1.754

由表 5 可以看出， $DW = 1.754$ ，自相关系数 $\hat{\rho} = 1 - \frac{1}{2}DW = 0.123$ 接近于 0，因此误差项可能不存在自相关性。

3.3. 回归方程的改进

为避免不同变量的量纲不同所产生的影响，先将数据标准化。

由 SPSS25.0 得到如下结果：

Table 6. Total variance explained
表 6. 总方差解释

成分	初始特征值			提取载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%
1	3.892	43.241	43.241	3.892	43.241	43.241
2	2.774	30.817	74.058	2.774	30.817	74.058
3	1.172	13.025	87.083	1.172	13.025	87.083
4	0.492	5.471	92.555	0.492	5.471	92.555
5	0.417	4.639	97.194	0.417	4.639	97.194
6	0.147	1.635	98.829	0.147	1.635	98.829
7	0.084	0.928	99.757	0.084	0.928	99.757
8	0.016	0.183	99.940	0.016	0.183	99.940
9	0.005	0.060	100.000	0.005	0.060	100.000

由表 6 可以看出，特征值大于 1 的是前三个主成分，方差百分比反映主成分所能解释数据变异的的比例，即包含原始数据的信息比例。前三个主成分累计包含了原始 9 个变量 87.083% 的信息量，前四个主成分累计包含了原始 9 个变量 92.555% 的信息量。

Table 7. Extract four common factors
表 7. 提取四个公因子

	初始	提取
X_1^*	1.000	0.986
X_2^*	1.000	0.984

Continued

X_3^*	1.000	0.821
X_4^*	1.000	0.988
X_5^*	1.000	0.888
X_6^*	1.000	0.950
X_7^*	1.000	0.974
X_8^*	1.000	0.935
X_9^*	1.000	0.803

由表 7 可以看出, 提取 4 个公因子时, 每个变量的共同度更加接近 1, 说明提取 4 个公因子更合理。各个变量的信息损失都较小。

Table 8. Composition matrix

表 8. 成分矩阵

	成分			
	1	2	3	4
X_1^*	-0.367	0.897	-0.085	-0.200
X_2^*	0.848	0.067	0.507	-0.058
X_3^*	0.807	0.000	-0.375	0.170
X_4^*	0.796	0.119	0.583	-0.002
X_5^*	-0.868	0.025	0.324	0.171
X_6^*	-0.209	0.763	0.273	0.500
X_7^*	-0.328	0.909	-0.077	-0.185
X_8^*	0.686	0.654	-0.011	-0.189
X_9^*	0.613	0.338	-0.492	0.265

对表 8 成分矩阵的第 i 列的每个元素分别除以第 i 个特征值的平方根 $\sqrt{\lambda_i}$, 就得到主成分分析第 i 个主成分的系数。结果如下:

Table 9. Principal component coefficients

表 9. 主成分系数

	主成分 1	主成分 2	主成分 3	主成分 4
X_1^*	-0.18604	0.538611	-0.07851	-0.28501
X_2^*	0.429861	0.040231	0.468264	-0.08265
X_3^*	0.409078	0.000213	-0.34635	0.242258
X_4^*	0.403502	0.071454	0.538457	-0.00285
X_5^*	-0.44	0.015011	0.299245	0.243683
X_6^*	-0.10594	0.458149	0.252142	0.712524

Continued

X_7^*	-0.16627	0.545816	-0.07112	-0.26363
X_8^*	0.347741	0.392699	-0.01016	-0.26933
X_9^*	0.310737	0.202955	-0.45441	0.377638

由表 9 可以得到 4 个主成分的线性组合如下:

$$y_1 = -0.186 * x_1^* + 0.430 * x_2^* + 0.409 * x_3^* + 0.404 * x_4^* - 0.44 * x_5^* \\ - 0.106 * x_6^* - 0.166 * x_7^* + 0.348 * x_8^* + 0.311 * x_9^*$$

$$y_2 = 0.539 * x_1^* + 0.040 * x_2^* + 0.409 * x_3^* + 0.0002 * x_4^* + 0.071 * x_5^* \\ + 0.458 * x_6^* + 0.546 * x_7^* + 0.393 * x_8^* + 0.203 * x_9^*$$

$$y_3 = -0.079 * x_1^* + 0.468 * x_2^* - 0.364 * x_3^* + 0.538 * x_4^* + 0.299 * x_5^* \\ + 0.252 * x_6^* - 0.071 * x_7^* - 0.010 * x_8^* - 0.454 * x_9^*$$

$$y_4 = -0.285 * x_1^* - 0.083 * x_2^* + 0.242 * x_3^* - 0.003 * x_4^* + 0.244 * x_5^* \\ + 0.713 * x_6^* - 0.264 * x_7^* - 0.269 * x_8^* + 0.378 * x_9^*$$

Table 10. Factor scores

表 10. 因子得分

	FAC1_1	FAC2_1	FAC3_1	FAC4_1
工商银行	-1.39597	3.075277	-0.31072	-0.40435
建设银行	-1.03055	2.545015	-0.09888	-0.61892
招商银行	3.946919	2.00712	1.995874	0.148642
农业银行	-1.34879	1.650741	-1.11301	-0.48947
兴业银行	2.128304	-0.23342	1.264216	-0.26752
华夏银行	-1.24671	-1.03427	0.960854	0.926302
中国银行	-1.27686	1.857993	-0.52776	0.496801
交通银行	-0.68553	-0.19236	0.115886	0.591915
中信银行	-1.1961	-1.31715	-0.32123	-0.43591
民生银行	-3.32931	-2.23725	1.808189	-1.07407
浦发银行	0.222218	-0.48077	0.655462	0.652218
宁波银行	3.360197	-0.99704	-0.68088	-0.78566
平安银行	0.596975	-0.79791	-0.57582	0.690797
南京银行	2.223144	-1.68548	-1.91428	-0.65324
北京银行	-0.7399	-1.59815	-0.18685	-0.09686
光大银行	-0.22803	-0.56232	-1.07106	1.319307

由表 10 因子得分, 可以通过因子得分的关系算的主成分得分。用主成分法做因子分析得到的第 i 个样本的因子得分为 $f_{(i)} = (A'A)^{-1} A'X_{(i)}$ ($i=1,2,\dots,n$), 用主成分法得到的第 i 个样本的因子得分为

$z_{(i)} = UX_{(i)}$ ，可以看到， $z_{(ik)} = \sqrt{\lambda_k} f_{(ik)} (i = 1, 2, \dots, n; k = 1, 2, \dots, p)$ 。因此可以得到前四个主成分的系数如下表 11。

Table 11. Principal component method principal component coefficients
表 11. 主成分法主成分系数

	Prin1	Prin2	Prin3	Prin3
工商银行	-2.75387	5.121554	-0.33642	-0.28374
建设银行	-2.033	4.238458	-0.10706	-0.43432
招商银行	7.786207	3.34265	2.16098	0.104307
农业银行	-2.6608	2.749138	-1.20508	-0.34348
兴业银行	4.19857	-0.38874	1.368797	-0.18773
华夏银行	-2.45942	-1.72247	1.040339	0.650015
中国银行	-2.5189	3.094294	-0.57142	0.34862
交通银行	-1.35237	-0.32036	0.125473	0.415365
中信银行	-2.35958	-2.19358	-0.3478	-0.30589
民生银行	-6.56783	-3.72591	1.957769	-0.75371
浦发银行	0.438376	-0.80067	0.709684	0.457681
宁波银行	6.628763	-1.66047	-0.7372	-0.55132
平安银行	1.177671	-1.32884	-0.62345	0.484753
南京银行	4.385664	-2.80699	-2.07264	-0.4584
北京银行	-1.45962	-2.66155	-0.20231	-0.06797
光大银行	-0.44984	-0.93649	-1.15966	0.925798

用 y^* 对前四个主成分 prin1, prin2, prin3, prin4 做普通最小二乘回归。

Table 12. Coefficients of principal component regression equation
表 12. 主成分回归方程系数

模型	未标准化系数		标准化系数	t	显著性	共线性统计	
	B	标准错误	Beta			容差	VIF
(常量)	-1.203E-6	0.034		0.000	1.000		
1 prin1	-0.080	0.009	-0.313	-8.803	0.000	1.000	1.000
prin2	0.329	0.013	0.914	25.707	0.000	1.000	1.000
prin3	-0.092	0.030	-0.108	-3.039	0.011	1.000	1.000
prin4	-0.411	0.072	-0.202	-5.689	0.000	1.000	1.000

由表 12 可以得到主成分回归方程为：

$$\hat{y}^* = -1.203 \times 10^{-6} - 0.080 * \text{prin1} + 0.329 * \text{prin2} - 0.092 * \text{prin3} - 0.411 * \text{prin4}$$

由从表 12 可以看出四个主成分的回归系数的显著性检验均通过，并且多重共线性诊断方差扩大因子 VIF 均为 1，则不存在多重共线性。

Table 13. Regression equation test
表 13. 回归方程检验

模型	R	R方	调整后R方	标准估算的错误	德宾 - 沃森
1	0.993	0.986	0.981	0.13768497	2.026

由表 13 可知, 此回归方程的样本决定系数 $R^2 = 0.986$, 调整后的样本决定系数 $R^2 = 0.981$ 。由调整后的决定系数来看, 回归方程高度显著。且 $DW = 2.026$, 查 DW 表, $n = 16$, $k = 5$, 显著性水平 $\alpha = 0.05$, 得 $d_L = 0.74$, $d_U = 1.93$, 由于 $d_U < 2.026 < 4 - d_U = 2.07$, 因而 DW 落入无自相关区域, 因此误差项不存在自相关性。

分别用四个主成分 prin1, prin2, prin3, prin4 做因变量, 以 9 个原始自变量为自变量做线性回归, 得到:

$$\begin{aligned} \text{prin1} &= 8.630 * 10^{-7} - 0.367 * x_1 + 0.848 * x_2 + 0.807 * x_3 + 0.796 * x_4 \\ &\quad - 0.868 * x_5 - 0.209 * x_6 - 0.328 * x_7 + 0.686 * x_8 + 0.613 * x_9 \\ \text{prin2} &= 2.706 * 10^{-6} + 0.897 * x_1 + 0.067 * x_2 + 0.0004 * x_3 + 0.009 * x_4 \\ &\quad + 0.025 * x_5 + 0.763 * x_6 + 0.909 * x_7 + 0.654 * x_8 + 0.338 * x_9 \\ \text{prin3} &= -6.091 * 10^{-7} - 0.085 * x_1 + 0.507 * x_2 - 0.375 * x_3 + 0.583 * x_4 \\ &\quad + 0.324 * x_5 + 0.273 * x_6 - 0.077 * x_7 - 0.011 * x_8 - 0.492 * x_9 \\ \text{prin4} &= -7.894 * 10^{-7} - 0.200 * x_1 - 0.058 * x_2 + 0.170 * x_3 - 0.002 * x_4 \\ &\quad + 0.171 * x_5 + 0.500 * x_6 - 0.185 * x_7 - 0.189 * x_8 + 0.265 * x_9 \end{aligned}$$

还原后的主成分回归方程为:

$$\begin{aligned} \hat{y}^* &= -1.191 * 10^{-9} + 0.415 * x_1^* - 0.069 * x_2^* - 0.100 * x_3^* - 0.077 * x_4^* \\ &\quad - 0.022 * x_5^* + 0.037 * x_6^* + 0.408 * x_7^* + 0.239 * x_8^* - 0.002 * x_9^* \end{aligned}$$

由主成分回归方程可以看到, 对经营业绩影响最大的是营业收入 (X_1), 其次是业务及管理费用 (X_7), 营业收入每增加 1%, 银行净利润平均增加 0.415%, 而业务及管理费用每增加 1%, 银行净利润平均增加 0.408%。

4. 小结

本文通过建立多元线性回归模型, 并且利用主成分回归对模型进项改进, 可以看出对经营业绩影响最大的是营业收入 (X_1), 其次是业务及管理费用 (X_7), 营业收入每增加 1%, 银行净利润平均增加 0.415%, 而业务及管理费用每增加 1%, 银行净利润平均增加 0.408%。

本文存在的不足:

- 1) x_1, x_6, x_7, x_8 对银行净利起正影响, x_2, x_3, x_4, x_5, x_9 起负影响。这与定性分析的结果不一致, 可能是主成分回归不适用于对该模型进行改进, 或者样本量太小。
- 2) 指标选取的局限性, 影响银行经营业绩的因素有很多, 很多因素没有考虑在内。
- 3) 研究方法的单一性。

致 谢

衷心感谢这一路师兄师姐给予我帮助, 给我提意见, 帮助我解决写作过程中遇到的各种问题, 耐心讲解。感谢家人一路以来的支持和鼓励。感恩所有的相遇。

参考文献

- [1] 中国银行业协会行业发展研究委员会. 2020 年度中国银行业发展报告[M]. 北京: 中国金融出版社, 2020.
- [2] 张平. 关于上市商业银行经营业绩的分析与对策研究[D]: [硕士学位论文]. 青岛: 中国海洋大学, 2007.
- [3] 曹晓俊. 对我国上市银行经营业绩的分析——基于主成分分析、因子分析和聚类分析的方法[J]. 宿州学院学报, 2016(7): 25-29.
- [4] 屈磊. 国有银行业绩分析[J]. 枣庄学院学报, 2012(5): 36-40.
- [5] 古文浩. 我国商业银行经营业绩评价指标分析[D]: [硕士学位论文]. 上海: 东华大学, 2004.
- [6] 杨建海. 我国商业银行业绩评价研究[D]: [硕士学位论文]. 苏州: 苏州大学, 2006.
- [7] 孙红彦, 吴书广, 赵涛. 基于因子分析法的我国商业银行综合业绩评价[J]. 金融发展研究, 2010(2): 58-61.
- [8] 赵萍. 我国上市银行经营业绩分析——基于因子分析的研究[J]. 金融经济: 下半月, 2018(4): 140-142.
- [9] 冯红梅. 互联网金融对商业银行经营业绩影响分析[D]: [硕士学位论文]. 济南: 山东大学, 2018.
- [10] 刘志颖. 2016 年上市银行业绩的多元统计分析[J]. 中国市场, 2017(5): 63-64.
- [11] 何晓群, 刘文卿. 应用回归分析[M]. 北京: 中国人民大学出版社, 2015.