

基于主成分分析的MLR模型和BP-ANN模型的比较研究

杨 华

(淄博职业学院 山东淄博 255314)

【摘要】 本文分仅使用财务变量和既使用财务变量又使用非财务变量两种情况,使用基于主成分分析的多元线性回归(MLR)模型和误差反向传播神经网络(BP-ANN)模型进行财务危机预警。结果显示,两种情况下的四个模型都可以进行财务危机预警,但预测效果存在差异:既使用财务变量又使用非财务变量的MLR模型和仅使用财务变量的BP-ANN模型的预测效果相同,但考虑总体预测准确率后,仅使用财务变量的BP-ANN模型的预测效果更好。

【关键词】 财务危机预警 主成分分析 MLR模型 BP-ANN模型

一、研究样本设计

Ohlson(1980)的研究表明,根据企业破产后的信息建立模型会高估模型的预测能力。考虑我国的年报披露制度,本文选取上市公司被特别处理的前两年作为研究期间。

由于我国上市公司被特别处理是从1998年开始的,而且研究中需要使用现金流量和成长能力这两个变量,故本文选取2001~2006年的196家被特别处理的上市公司作为财务危机公司。

按照与财务危机公司上市时间相同或相近、行业门类相同或相近的原则,本文还选取了196家非财务危机公司作为配对样本。这样,研究样本就由196家财务危机公司和196家非财务危机公司共同构成。

为保证建立的财务危机预警模型的预测准确性,按照财务危机公司与非财务危机公司的配对,将每年的财务危机公司按公司代码排列后,随机抽取其中的1/4构成测试样本组,作为外部数据用于对构造出来的财务危机预警模型进行检验;剩余的3/4组成构造样本组,作为原始数据用于构建财务危机预警模型。也就是说,构造样本组由294家公司构成,其中

加强管理。比如,每个交易员可以进行的交易类型都应当清楚地用文字表述出来,从而确保所投资的金融工具的性质和所授予的操作权限与战略一致。对柜台衍生品应使用银行制定的标准化合约,严格遵循合约的期限,信息系统管理人员应当阻止黑客入侵、维护系统安全等。

3. 建立完善的内部控制制度。根据BASEL II的要求,银行资本应当包括操作风险监管成本。加强操作风险管理必须从建立完善的内部控制制度入手。管理层和风险管理部应对每个部门、业务线的操作风险进行控制,在操作风险管理体系中进行完整记录。

主要参考文献

1. Basel Committee on Banking Supervision. The New

财务危机公司与非财务危机公司各147家;测试样本组由98家公司构成,其中财务危机公司与非财务危机公司各49家。

二、主成分分析

1. 财务预警变量的选取。上市公司陷入财务危机,可能是财务活动造成的,也可能是非财务活动造成的。借鉴国内外学者的研究成果,本文选取了财务变量和非财务变量,详见表1和表2。

2. 财务预警变量的正态性及显著性检验。本文通过单样本的柯尔莫哥洛夫-斯米诺夫检验来验证选取的37个财务预警变量在总体上是否服从联合正态分布。如果服从,显著性检验时使用T检验法;否则,使用非参数检验法。K-S检验结果显示,仅 X_2 、 X_3 、 X_{23} 、 X_{30} 、 X_{37} 服从正态分布,其余32个变量不服从正态分布,因而可以认定这37个变量总体上不服从联合正态分布,需要使用非参数检验法。

显著性检验发现,除公司治理方面的4个变量外,其余变量均在5%的水平上显著,因而在进行主成分分析时剔除公司治理方面的4个变量。

3. 主成分因子的提取。在仅使用财务变量和既使用财务

Basel Capital Accord. Bank for International Settlements, 2003

2. 巴塞尔银行监管委员会发布.中国银行业监督管理委员会译.统一资本计量和资本标准协议:修订框架.北京:中国金融出版社,2004

3. 张吉光.商业银行操作风险识别与管理.北京:中国人民大学出版社,2005

4. 张宏毅.银行操作风险度量方法比较.经济理论与经济管理,2004;11

5. 曾向阳.对商业银行操作风险管理的几点思考.广西金融研究,2005;8

6. 巴曙松.巴塞尔新资本协议研究.北京:中国金融出版社,2003

表1 财 务 变 量

评价内容	变 量	评价内容	变 量
短期偿债能力	速动比率 X ₁	成长能力	主营业务收入增长率 X ₁₅
	营运资本比重 X ₂		资本积累率 X ₁₆
长期偿债能力	资产负债率 X ₃	现金流量	总资产扩张率 X ₁₇
	产权比率 X ₄		净利润增长率 X ₁₈
营运能力	应收账款周转率 X ₅		现金流动负债比 X ₁₉
	存货周转率 X ₆		现金负债总额比 X ₂₀
	总资产周转率 X ₇		主营业务收入现金回收比率 X ₂₁
	固定资产周转率 X ₈		净利润现金比率 X ₂₂
盈利能力	销售净利率 X ₉	市场价值	资产现金回报率 X ₂₃
	资产净利率 X ₁₀		现金流入与现金流出比 X ₂₄
	净资产收益率 X ₁₁		每股收益 X ₂₅
	总资产收益率 X ₁₂		每股净资产 X ₂₆
	主营业务利润率 X ₁₃		每股经营现金净流量 X ₂₇
	主营业务利润比重 X ₁₄		

表2 非 财 务 变 量

评价内容	变 量	对变量的解释
股权结构	是否存在控股股东 X ₂₈	虚拟变量,若第一大股东持股比例大于后九大股东持股比例之和,则认为存在控股股东,设为1,否则为0
	国有股比例 X ₂₉	国家股和国有法人股的数额/股份总额
年报披露	公告日期 X ₃₀	年报披露截止日与年报公告日的时间差
	审计意见类型 X ₃₁	虚拟变量,被出具非标准意见为1,否则为0
公司治理	董事会规模 X ₃₂	董事会成员个数,不包括独立董事和董事长秘书
	监事会规模 X ₃₃	监事会成员个数
	高管持股比例 X ₃₄	高管持股数额占股本总额的比例
	董事长和总经理是否由一人兼任 X ₃₅	虚拟变量,两职合一设为1,否则为0
地域环境	是否为GDP排名前十的省份 X ₃₆	虚拟变量,GDP排名前十的省份设为1,否则为0
公司规模	总资产的对数 X ₃₇	陷入财务危机前两年以10为底的总资产的对数

变量又使用非财务变量的情况下,KMO检验值分别为0.742、0.744,表明这两种情况均适合进行主成分分析。

(1)仅使用财务变量情况下的主成分因子。仅使用财务变量时提取8个主成分因子,分别是现金流量—偿债和营运能力主成分因子F₁(主要由X₁₉、X₂₀、X₂₃解释)、市场价值主成分因子F₂(主要由X₂₅解释)、盈利能力—获利主成分因子F₃(主要由X₉和X₁₁解释)、短期偿债能力主成分因子F₄(主要由X₁和X₂解释)、营运能力—总资产和固定资产主成分因子F₅(主要由X₇和X₈解释)、现金流量—盈利能力主成分因子F₆(主要由X₂₂解释)、盈利能力—主营业务主成分因子F₇(主要由X₁₄解

释)、营运能力—存货主成分因子F₈(主要由X₆解释)。

(2)既使用财务变量又使用非财务变量情况下的主成分因子。既使用财务变量又使用非财务变量时提取10个主成分因子,分别是现金流量—偿债和营运能力主成分因子F₁(主要由X₁₉、X₂₀和X₂₃解释)、市场价值主成分因子F₂(主要由X₂₅解释)、盈利能力主成分因子F₃(主要由X₉、X₁₀、X₁₁、X₁₂和X₁₃解释)、短期偿债能力主成分因子F₄(主要由X₁和X₂解释)、营运能力—总资产和固定资产主成分因子F₅(主要由X₇和X₈解释)、公司规模主成分因子F₆(主要由X₃₇解释)、股权结构主成分因子F₇(主要由X₂₈和X₂₉解释)、营运能力—存货主成分因子F₈(主要由X₆解释)、地域环境主成分因子F₉(主要由X₃₆解释)、现金流量—盈利能力主成分因子F₁₀(主要由X₂₂解释)。

三、MLR模型的构建及预测效果

因变量y是虚拟变量,表示上市公司是否会出现财务危机,出现财务危机时(即财务危机公司)为1,没有出现财务危机时(即非财务危机公司)为0;自变量是提取的主成分因子。

由于各个主成分因子之间不存在多重共线性问题,因而可以采用全部进入的方法。

根据仅使用财务变量时提取的8个主成分因子建立的MLR模型为:

$$P_{(t-2)} = [\exp(0.720 - 2.025F_1 - 4.29F_2 - 2.635F_3 - 1.048F_4 - 1.332F_5 + 0.766F_6 - 0.281F_7 + 0.756F_8)] / [1 + \exp(0.720 - 2.025F_1 - 4.29F_2 - 2.635F_3 - 1.048F_4 - 1.332F_5 + 0.766F_6 - 0.281F_7 + 0.756F_8)] \quad (1)$$

按照既使用财务变量又使用非财务变量时提取的10个主成分因子建立的MLR模型为:

$$P_{(t-2)} = [\exp(0.661 - 2.449F_1 - 4.937F_2 - 2.488F_3 - 1.105F_4 - 1.356F_5 - 1.141F_6 - 0.646F_7 + 0.707F_8 - 1.308F_9 + 0.521F_{10})] / [1 + \exp(0.661 - 2.449F_1 - 4.937F_2 - 2.488F_3 - 1.105F_4 - 1.356F_5 - 1.141F_6 - 0.646F_7 + 0.707F_8 - 1.308F_9 + 0.521F_{10})] \quad (2)$$

由于采用一一配对方式进行研究,我们将MLR模型的预测分割点默认为0.5。将98个测试样本的主成分的分值分别代入模型(1)和模型(2),得到的预测准确率见表3。

表3

观 测 值	模型(1)		模型(2)				
	是否为财务危机公司	预测准确率(%)	是否为财务危机公司	预测准确率(%)			
	0	1	0	1			
是否为财务危机公司	0	30	19	61.224	32	17	65.306
	1	14	35	71.429	10	39	79.592
总体预测准确率				66.327			72.449

模型(1)和模型(2)对财务危机公司的预测准确率均达到了70%以上,但模型(2)对财务危机公司的预测准确率比模型(1)的高8%左右。可见,在MLR模型中引入非财务变量是非常有必要的。

模型(1)和模型(2)犯I类错误(是指实际为财务危机公司却被误测为非财务危机公司)的次数均少于犯II类错误(是

指实际为非财务危机公司却被误测为财务危机公司)的次数,说明这两个模型对财务危机公司的识别能力均强于对非财务危机公司的识别能力。

可见,本文运用MLR模型进行财务危机预警时,对财务危机公司的预测均取得了较好的效果,而引入非财务变量后,模型的预测效果更好。但MLR模型还是存在一定的局限性,有必要使用同样的变量建立BP-ANN模型进行比较研究。

四、BP-ANN模型的构建及预测效果

设定的学习率 η 为0.15,动量因子 α 为0.8。仅使用财务变量时,输入点 P_1 是8个主成分因子,输出点 P_3 是虚拟变量 y ,隐藏层 P_2 根据式子“(2 P_1+P_3)^{1/2}< P_2 <2 P_1+1 ”并结合实际试算确定为5个,这样得到的模型结构是“8×5×1”,建立的仅使用财务变量的BP-ANN模型为:

$$y_1 = -185.111F_1 - 176.411F_2 - 325.585F_3 - 45.893F_4 - 160.04F_5 - 24.56F_6 + 50.846F_7 + 153.261F_8 \quad (3)$$

既使用财务变量又使用非财务变量时,输入点 P_1 是10个主成分,输出点 P_3 也是虚拟变量 y ,隐藏层 P_2 仍确定为5个,得到的模型结构是“10×5×1”,建立的既使用财务变量又使用非财务变量的BP-ANN模型为:

$$y_2 = -163.008F_1 - 492.725F_2 - 58.456F_3 - 57.285F_4 - 113.502F_5 - 132.819F_6 + 46.572F_7 + 16.77F_8 - 218.935F_9 + 136.447F_{10} \quad (4)$$

对预测结果进行详细测算,选取0为分割点。将98个测试样本的主成分的分值分别代入模型(3)和模型(4),得到的预测准确率见表4。

表 4

观测值	模型(3)			模型(4)			
	是否为财务危机公司	预测准确率(%)		是否为财务危机公司	预测准确率(%)		
		0	1		0	1	
是否为财务危机公司	0	39	10	79.592	27	22	55.102
	1	10	39	79.592	12	37	75.510
总体预测准确率				79.592			63.265

模型(3)和模型(4)对财务危机公司的预测准确率均达到了75%以上,模型(3)的预测准确率高于模型(4),说明仅使用财务变量的BP-ANN模型的预测效果更好,非财务变量的加入并没有使模型的预测准确率提高。

模型(3)和模型(4)犯 I 类错误的次数也少于犯 II 类错误的次数,说明这两个模型对财务危机公司的识别能力均强于对非财务危机公司的识别能力。

五、四个模型的预测效果比较

以上建立的四个模型对财务危机公司的预测准确率都较

高,都在70%以上。另外,四个模型犯 I 类错误的次数都少于犯 II 类错误的次数,体现出它们对财务危机公司具有较强的识别能力,适用于财务危机预警。模型(1)对财务危机公司的预测准确率最低,为71.429%;模型(2)和模型(3)对财务危机公司的预测准确率相同,均为79.592%。但从模型的总体预测准确率来看,模型(3)的最高,模型(2)的次之。具体情况见表5。

表 5

两种情况	模型	财务危机公司的预测准确率(%)	总体预测准确率(%)
仅使用财务变量	(1)	71.429	66.327
	(3)	79.592	79.592
既使用财务变量又使用非财务变量	(2)	79.592	72.449
	(4)	75.510	63.265

六、结论

本文以主成分分析法为基础,分仅使用财务变量和既使用财务变量又使用非财务变量两种情况,使用MLR模型和BP-ANN模型进行财务危机预警,发现仅使用财务变量的BP-ANN模型的预测效果最佳。通过本文的定量分析,我们可以得出如下结论:

第一,财务预警变量总体上不服从联合正态分布,所以不满足多元判别分析法的适用前提。考虑到模型的预测效果,适宜使用MLR模型,显著性检验适宜使用非参数检验法。

第二,除公司治理变量外,财务危机公司与非财务危机公司在偿债能力、营运能力、盈利能力、成长能力、现金流量、市场价值、股权结构、年报披露、地域环境、公司规模方面均存在显著差异,从而证实本文在财务预警变量的选择上是成功的。

第三,对于MLR模型来说,既使用财务变量又使用非财务变量时的预测效果较好;而对于BP-ANN模型来说,仅使用财务变量时的预测效果较好。可见,并非引入非财务变量就能提高模型的预测准确率。

第四,无论是MLR模型还是BP-ANN模型,其预测准确率都不可能为100%,这说明定量方法存在一定的缺陷,而要想提高预测准确率,必须辅之以定性方法。

主要参考文献

1. Beaver W. H.. Financial Ratios as Predictors of Failure. Journal of Accounting Research,1966;4
2. Ohlson J. A.. Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research,1980;1
3. 姚靠华, 蒋艳辉. 静态财务预警与动态财务预警系统比较研究. 长沙铁道学院学报(社会科学版),2005;2