

Research on the Regional Financial Eco-Environments via Multivariate Statistical Analysis

Zhongkui Sun¹, Yifei Liu²

¹Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an Shannxi

²College of Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA

Email: sunzk@nwpu.edu.cn

Received: Feb. 7th, 2015; accepted: Feb. 20th, 2015; published: Feb. 26th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

It is of particular importance to estimate a regional financial eco-environment, based on which the administrators can make their efficient policy, and the investors may estimate the potential risk in investment to make brilliant decisions. For this purpose, the current paper deals with the measurement of a regional financial eco-environment. A mathematical model has been proposed by means of multivariate statistical methods, for example, the factor analysis method, the cluster methods and the discriminant analysis method, and it has been well tested and verified in estimating for regional financial eco-environment through the real data of ten states in United State.

Keywords

Regional Financial Eco-Environment, Evaluating Model, Multivariate Statistical Analysis

基于多元统计分析的地区金融生态环境量化评估模型研究

孙中奎¹, 刘怡菲²

¹西北工业大学理学院应用数学系, 陕西 西安

²乔治亚理工学院理学院, 亚特兰大, 佐治亚州, 美国

Email: sunzk@nwpu.edu.cn

收稿日期：2015年2月7日；录用日期：2015年2月20日；发布日期：2015年2月26日

摘要

合理评估金融生态环境对于改善地区金融环境、提高金融系统的抗风险能力具有十分重要的意义。尤其是针对地区金融生态环境的量化分析和评价既可发挥地区经济特色，也能促进地区经济健康有序发展。有鉴于此，本文研究了地区金融生态环境的评价问题，基于多元统计理论建立了地区金融生态环境的量化评估模型——利用因子分析对指标进行降维、利用聚类分析对样本进行分类、再利用判别分析建立了判别规则。通过所收集的美国10个州的经济、民生数据，验证了本文所建立模型在地区金融生态环境评价中的合理性和有效性。

关键词

地区金融生态环境，评估模型，多元统计分析

1. 引言

金融生态环境是指一个国家(或地区)在一定的政治制度和金融体系下，影响经济主体活动的各种要素的总和，包括政治、经济、文化、地理、人口等一切与金融相互影响、相互作用的因素。因此，合理评估金融生态环境对于改善社会金融环境、提高金融系统的抗风险能力、促进经济的多元健康发展具有十分重要的意义。尤其是针对地区金融环境的全面分析和评价将有助于发挥经济特色，助力地区经济腾飞。

金融生态环境是一个十分复杂的系统，通过对现实表征数据的研究，例如经济基础、金融发展、企业诚信、司法环境等，人们可以挖掘出数据下隐藏的金融规律。目前，借助于数据分析技术，多种统计模型和方法被成功引入金融分析领域，特别是对信用风险的研究之中[1]-[3]。经过几十年的发展，信用风险的研究经历了银行专家体系、基于会计的信用评分系统、现代风险度量模型三个发展阶段[1]。不仅形成了一套较为成熟的信用评价模型，而且诞生了若干知名的信用评级机构[2]。近年来，随着中国经济的飞速发展，我国也步入了“信用经济时代”，社会信用体系的概念也于1999年由中国社会科学院世界经济与政治研究所提出[3]。然而，相较于对信用风险的研究，对于金融生态环境的研究还比较匮乏，相关成果还很少见到。特别地，如何对地区金融生态环境进行量化仍然是国内外研究人员关注的热点[4] [5]。

有鉴于此，本文基于多元统计理论建立了金融生态环境的量化评价模型。在合理选择表征地区金融生态环境的指标数据和样本的基础上，通过因子分析对数据进行降维，提取出具有明显经济学意义的因子；再通过聚类分析得到现实样本的一个划分；最后通过对聚类分析结果分析，利用判别分析建立了判别函数，实现了对地区金融环境信用风险的评估。

2. 指标选取

对地区金融生态环境的研究最早源于瑞士洛桑国际管理学院(IMD)。该项研究的目的是评估一个国家的竞争力——通过将一个国家的竞争力分成8大类290项指标，研究者建议了一套评价国家竞争力的评估方法。其中涉及金融生态环境的指标包括4大类27项指标。然而，这套指标体系侧重于对国家整体竞争力的考量，因此包括了许多类似于汇率、短期实际利率、利息加息率等短期指标。此后，不断有学者对此指标体系进行改善，并针对不同研究对象分别提出了金融竞争力评价指标体系[6]。一般来说，表征

金融竞争力的指标涉及到经济基础、社会保障、社会基本情况和政府服务四个方面。本文从地区金融生态环境量化评估角度出发，在邓洪中等人针对区域金融生态环境的研究成果基础上[6]，构建了：同步经济指标(Coincident Economic Index)、实际国内生产总值(Real Total Gross Domestic Product)、零售贸易盈余(Retail Trade Earnings)、批发贸易盈余(Wholesale Trade Earnings)、政府税收/税收总额(Government Tax Collection/Total Tax)、个人总收入(Total Personal Income)、健康保险覆盖/不覆盖率(Health Insurance Coverage/Uncovered Rate)、民用劳动力(Civilian Labor Force)、居住人口(Residential Population)以及失业率(Unemployment Rate)等 10 项指标，如表 1 所示。全面涵盖了经济基础、社会保障、社会基本情况和政府服务四个方面。

3. 样本选取

本文选取了美国 Alabama、Massachusetts 等 10 个州作为样本，另外选取 Colorado、Florida 等 6 个州作为检验样本。数据均来自美联储圣路易斯分行(Federal Reserve Bank of ST. LOUIS)发布的 2012 年指标数据[7]，表 2 为所选各州的原始数据。

4. 基于多元统计的地区金融生态环境量化模型

4.1. 对实际指标数据进行降维

在实际问题中，多个指标之间可能会存在信息重叠，通过研究指标之间的相关性和内部依赖关系来探究指标数据的基本结构，并从指标中提取出可反映原指标信息的公因子，如此便实现了对指标数据的降维和简化，这种方法称为因子分析[8]。常用的因子分析类型有 R 型因子分析，其中每一个变量均能表示成公共因子的线性函数与特殊因子之和，公共因子是不可直接观测但又客观存在的共同影响因子，即

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + \varepsilon_i, \quad i = 1, 2, \dots, p \quad (1)$$

公式(1)中的 F_1, F_2, \dots, F_m 称为公共因子， ε_i 称为 X_i 的特殊因子。该模型可用矩阵表示为

$$X = AF + \varepsilon \quad (2)$$

这里，

Table 1. Indexes and Variables

表 1. 指标与变量

| 指标 | 备注 | 变量 | 分类 |
|-------------------------------------------------------|-------|----------|--------|
| 同步经济指标(Coincident Economic Index) | 变化百分比 | X_2 | 经济基础 |
| 实际国内生产总值(Real Total Gross Domestic Product) | 变化百分比 | X_4 | |
| 零售贸易盈余(Retail Trade Earnings) | 变化百分比 | X_6 | |
| 批发贸易盈余(Wholesale Trade Earnings) | 变化百分比 | X_{10} | |
| 政府税收/税收总额(Government Tax Collection/Total Tax) | 变化百分比 | X_7 | 政府公共服务 |
| 个人总收入(Total Personal Income) | 千美元 | X_8 | 社会保障程度 |
| 健康保险覆盖/不覆盖率(Health Insurance Coverage/Uncovered Rate) | 变化百分比 | X_3 | |
| 民用劳动力(Civilian Labor Force) | 变化百分比 | X_1 | 社会基本情况 |
| 居住人口(Residential Population) | 变化百分比 | X_5 | |
| 失业率(Unemployment Rate) | 变化百分比 | X_9 | |

Table 2. The values of indexes
表 2. 指标原始数据表

| States | X ₁ | X ₂ | X ₃ | X ₄ | X ₅ | X ₆ | X ₇ | X ₈ | X ₉ | X ₁₀ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------------|----------------|-----------------|
| Alabama | -1.2 | 1.9 | 1.8 | 1.2 | 0.3 | 1.1 | 4.8 | 5,449,607 | -1.4 | 1.9 |
| Massachusetts | 0.1 | 3.6 | 0.7 | 2.2 | 0.6 | 2.0 | 3.3 | 13,808,256 | -0.5 | 0.0 |
| North Carolina | 1.0 | 3.2 | 0.9 | 2.7 | 1.0 | 5.2 | 1.4 | 17,248,510 | -0.9 | 5.8 |
| Pennsylvania | 1.1 | 1.9 | 1.2 | 1.7 | 0.2 | 1.7 | 1.8 | 17,079,509 | -0.1 | 3.0 |
| California | 0.5 | 3.5 | -1.8 | 3.5 | 0.9 | 4.7 | 4.9 | 84,835,581 | -1.4 | 4.9 |
| New Jersey | 0.6 | 2.6 | -1.4 | 1.3 | 0.4 | 2.4 | -3.7 | 16,249,428 | 0.0 | 4.8 |
| Washington | 0.3 | 4.0 | -0.9 | 3.6 | 1.1 | 6.9 | 1.2 | 14,486,873 | -1.1 | 5.7 |
| Illinois | -0.1 | 3.1 | -1.1 | 1.9 | 0.1 | 3.1 | 12.6 | 22,897,228 | -0.7 | 5.9 |
| Nebraska | 1.5 | 2.8 | 1.0 | 1.5 | 0.7 | 5.1 | 1.0 | 3,100,678 | -0.5 | 7.5 |
| Texas | 1.0 | 4.6 | 0.8 | 4.8 | 1.6 | 5.5 | 19.1 | 57,558,626 | -1.0 | 9.4 |
| Colorado | 0.7 | 3.5 | -2.0 | 2.1 | 1.4 | 4.2 | 8.3 | 11,429,578 | -0.7 | 5.7 |
| Florida | 1.1 | 2.2 | 1.7 | 2.4 | 1.2 | 4.2 | 1.3 | 30,952,154 | -1.5 | 6.6 |
| Iowa | -0.2 | 2.7 | 0.1 | 2.4 | 0.4 | 4.0 | 8.2 | 4,932,736 | -0.7 | 7.3 |
| New York | 0.8 | 2.6 | -0.9 | 1.3 | 0.4 | 3.8 | 5.3 | 29,524,898 | 0.2 | 6.8 |
| Utah | 1.7 | 4.5 | -0.2 | 3.4 | 1.4 | 6.8 | 6.1 | 4,987,599 | -1.4 | 6.8 |
| Virginia | 0.3 | 1.7 | -0.9 | 1.1 | 1.0 | 3.1 | 4.2 | 14,074,820 | -0.6 | 0.4 |

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \cdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{bmatrix}, X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}, F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}$$

且满足:

- 1) $m \leq p$;
- 2) $\text{cov}(F, \varepsilon) = 0$, 即公共因子与特殊因子是不相关的;
- 3) $D_F = D(F) = I_m$, 即各个公共因子不相关且方差为 1;

$$4) D_\varepsilon = D(\varepsilon) = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_p^2 \end{bmatrix}, \text{即各个特殊因子不相关, 方差不同。}$$

其中, a_{ij} 是第 i 个变量在第 j 个因子上的“载荷”, 表示变量 X_i 在 m 维空间中坐标轴 F_j 上的投影, A 称为因子载荷矩阵。对标准化的各州指标数据进行因子分析。

表 3 显示前四个公因子的累积方差贡献率已经达到 87.464%, 故应提取四个因子。在实际分析中, 有些变量可能在多个公因子上都有较大的载荷, 说明它对多个公因子都有较明显的影响作用, 这样很难对所提取的公共因子的实际含义进行合理的解释, 故不可取。通过因子旋转法, 使每个变量仅在一个公共因子上有较大载荷。本文采用正交旋转法, 即给载荷矩阵 A 右乘一个正交矩阵 Γ , 使得 AF 具有更鲜明的实际意义。旋转以后的公共因子向量为 $F^* = \Gamma'F$, 它的各个分量 $F_1^*, F_2^*, \dots, F_m^*$ 也是互不相关的公共因子。旋转后的因子载荷矩阵如表 4。

Table 3. Characteristic root and variance contribution rate
表 3. 特征根与方差贡献率

| 初始特征值 | | | 提取平方和载入 | | | 旋转平方和载入 | | |
|-----------|-----------|---------|---------|--------|--------|---------|--------|--------|
| 合计 | 方差百分比 | 累积百分比 | 合计 | 方差百分比 | 累积百分比 | 合计 | 方差百分比 | 累积百分比 |
| 4.684 | 46.844 | 46.844 | 4.684 | 46.844 | 46.844 | 4.402 | 44.021 | 44.021 |
| 1.680 | 16.797 | 63.642 | 1.680 | 16.797 | 63.642 | 1.650 | 16.496 | 60.517 |
| 1.271 | 12.715 | 76.357 | 1.271 | 12.715 | 76.357 | 1.460 | 14.602 | 75.119 |
| 1.111 | 11.107 | 87.464 | 1.111 | 11.107 | 87.464 | 1.234 | 12.345 | 87.464 |
| 0.581 | 5.813 | 93.277 | | | | | | |
| 0.512 | 5.116 | 98.392 | | | | | | |
| 0.106 | 1.057 | 99.449 | | | | | | |
| 0.054 | 0.542 | 99.991 | | | | | | |
| 0.001 | 0.009 | 100.000 | | | | | | |
| 1.413E-16 | 1.413E-15 | 100.000 | | | | | | |

Table 4. The factor loading matrix
表 4. 因子载荷阵

| | F_1 | F_2 | F_3 | F_4 |
|----------|--------|--------|--------|--------|
| X_1 | 0.411 | 0.831 | -0.266 | 0.102 |
| X_2 | 0.884 | -0.076 | 0.116 | -0.074 |
| X_3 | -0.32 | 0.156 | 0.503 | 0.737 |
| X_4 | 0.935 | -0.167 | 0.039 | 0.076 |
| X_5 | 0.908 | 0.024 | 0.041 | 0.358 |
| X_6 | 0.875 | 0.192 | -0.010 | 0.101 |
| X_7 | 0.042 | 0.088 | 0.847 | -0.500 |
| X_8 | 0.661 | -0.350 | -0.322 | -0.251 |
| X_9 | -0.449 | 0.760 | -0.190 | -0.235 |
| X_{10} | 0.725 | 0.431 | 0.270 | -0.199 |

表 4 显示经过旋转后的载荷系数已经明显地两极分化了，并且每个公共因子都具有明显的经济学意义，例如第一个公共因子在指标 $X_2, X_4, X_5, X_6, X_{10}$ 上有较大载荷，对应于前文所述的经济基础。可以看出这四个公共因子分别对应经济基础、社会基本情况、社会保障情况以及政府公共服务情况。同时结合方差贡献率来分析，可知第一公共因子，即经济基础对金融环境的贡献是最高的，其次是社会基本情况，最后是政府服务情况。由此看来一个地区的金融环境的好坏受到当地经济基础的极大影响。

利用汤姆孙回归法可以得到各因子得分[8] [9]:

$$\hat{F} = A'R^{-1}X \tag{3}$$

其中 R 是 X 的相关系数矩阵。再根据因子得分和原始变量可以计算得到每个因子的得分。

如此，我们便通过因子分析方法将原本的 10 个指标简化为了 4 个因子，达到了降低数据维数的目的。表 5 给出了这 4 个因子的实际数值。

4.2. 地区金融生态环境的聚类分析

将指标数据降维后，再利用聚类分析对各因子所表征的各州金融环境进行分类，使金融环境相似的州分为一类。本文采用 K 均值聚类分析方法[8]，每个样品的 p 个指标构成一个 p 维的向量，将样品看成 p 维空间中的点，通过马氏距离来测度样品之间的距离。令 d_{ij} 表示样品 X_i 与 X_j 的距离，设 X_i 与 X_j 是来自均值为 μ 、协方差为 $\Sigma (>0)$ 的总体 G 中的样品，则两个样品间的马氏距离为

$$d_{ij}^2(M) = (X_i - X_j)' \Sigma^{-1} (X_i - X_j) \tag{4}$$

这种方法的基本思想是将每一个样品分配给最近中心均值的类中，具体算法包括以下三个步骤：

- 1) 将所有的样品分为 K 个初始类；
- 2) 计算某个样品与各类中心的距离，将样品划入距离最近的类中，重新计算获得样品与失去样品的类的中心坐标；
- 3) 重复步骤 2)，直到所有的样品都不能再分配时为止。

图 1 和图 2 比较了第一类各州的经济基础，通过表 6 可以看出第一类各州的经济基础均在温和上升

Table 5. The real value of factors to the 10 states in United State

表 5. 美国各州因子实际数值表

| | F_1 | F_2 | F_3 | F_4 |
|----------------|-------------|-------------|-------------|-------------|
| Alabama | -1.10690411 | -2.18296589 | -1.242474 | -0.03500807 |
| Massachusetts | -0.58746834 | -0.6535674 | -0.35917688 | -0.57031158 |
| North Carolina | 0.63001654 | 0.25249944 | -0.82691901 | -0.65425517 |
| Pennsylvania | -1.03004267 | 0.82484784 | -0.52307878 | -0.70780369 |
| California | 0.57836426 | -0.8712835 | 2.08008237 | -1.62648292 |
| New Jersey | -1.09375418 | 0.9339093 | 1.05135666 | -0.36947303 |
| Washington | 0.95710077 | -0.34842169 | 0.21100746 | -0.4699598 |
| Illinois | -0.84037152 | -0.44268151 | 1.03979483 | 2.65076271 |
| Nebraska | 0.14060162 | 1.31344839 | -1.00135703 | 0.14741414 |
| Texas | 1.92972963 | -0.17269051 | -0.20708572 | 1.22869841 |
| Colorado | 0.17723531 | 0.3225885 | 0.74193016 | 0.54201162 |
| Florida | 0.7232965 | -0.28350012 | -1.33674093 | -0.91799056 |
| Iowa | -0.313884 | -0.24920001 | -0.17416538 | 1.06413294 |
| New York | -0.7844275 | 1.37291081 | 0.98498395 | 0.39450921 |
| Utah | 1.55522022 | 0.38429765 | -0.65885504 | 0.07124817 |
| Virginia | -0.93471703 | -0.20019178 | 0.22069283 | -0.74747996 |

Table 6. The results from cluster analysis

表 6. 聚类结果表

| 类别 | States |
|----|----------------------------------------------------------------|
| 1 | Alabama, Massachusetts, North Carolina, Pennsylvania, Nebraska |
| 2 | California, Washington, Texas |
| 3 | New Jersey, Illinois |

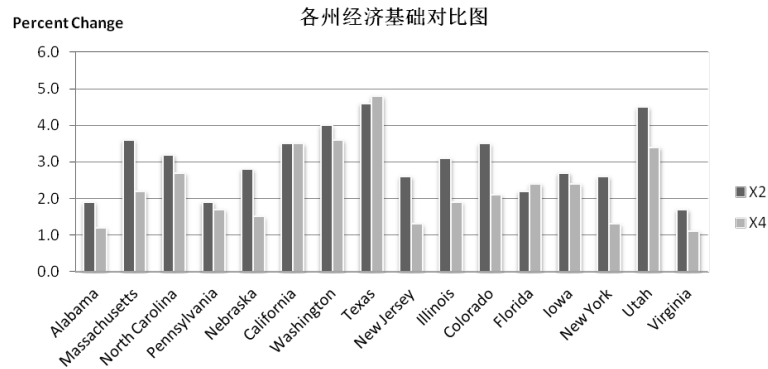


Figure 1. Economic foundation of the 10 state in United State: X_2 and X_4

图 1. 各州经济基础: X_2 和 X_4

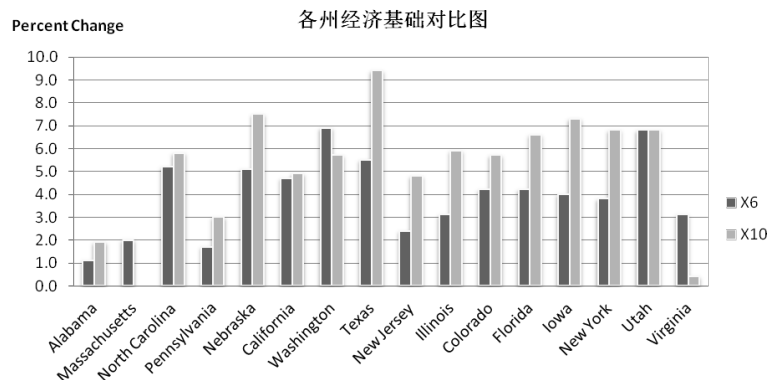


Figure 2. Economic foundation of the 10 state in United State: X_6 and X_{10}

图 2. 各州经济基础: X_6 和 X_{10}

中，例如 Massachusetts，该州的一致经济指数活动指数较高，说明该州的经济活动较为活跃，与此同时该州的 GDP 也呈现出较快的增长速度。该州的零售贸易收入增幅较小，而批发贸易的增幅基本为零。Massachusetts 的优势产业为高等教育产业、健康医疗以及金融服务产业，虽在 2008 年的金融危机中，该州的金融服务产业受到了冲击，然而该州的经济整体趋势是回暖的。第二类中所有的州经济基础增长都十分迅猛，所有的州经济活动活跃程度都很高，一致经济活动指数的增长率都在平均水平以上。各州的 GDP 增速显著，名列十个州的前三名。零售批发贸易收入增速也均在平均水平以上。其中 Texas 的各项指标数据增长十分强劲，该州的经济呈显著上升趋势，这种情况与该州作为全美经济复苏的领头羊的地位是相符的，该州农牧业发达，作为传统优势产业，农牧业在金融危机中称为了该州经济的坚强后盾，该州经济多元化发展，高科技产业兴起，作为新的经济增长点，显然这些产业推动了该州的经济的发展。第三类各州的经济也处于温和回暖中，各州经济活动活跃程度较为一般，在均值附近，地区生产总值和零售贸易收入也增长缓慢，批发贸易收入增长尚可观，在均值附近。值得注意的是 New Jersey 和 Illinois 的金融服务业都较为发达。

图 3 显示的是第一类各州得社会保障情况。就社会保障情况而言，第一类各州的社会保障程度并不令人满意，健康保险未覆盖率增长均较快而个人收入的增长幅度很小。说明这些州有很多人并未享受到健康保险福利，并且收入增幅小，生活上经济压力可能较大。第二类的州健康保险未覆盖率减小并且个人收入增加，社会保障程度较为理想。例如 California 的健康保险未覆盖率显著减小而个人收入情况增长幅度很大。第三类的州健康保险未覆盖率显著下降的同时保持了较低的个人收入增长。第三类各州的社会保障程度虽不及第二类，但较第一类更好一些。

对比各州的社会基本情况(见图 4)，我们不难发现第一类中各州的劳动力水平上升，居民人口增加，失业率下降，这与第一类各州的经济基础温和回暖这一情况相匹配。第二类中各州因为经济情况十分可观，不断吸引外来劳动力，并且创造了许多就业，例如 Texas 劳动力的增长则十分显著，Texas 作为经济复苏的领头羊，吸引了许多劳动力在该州寻找就业机会。第三类中各州居民人口总数缓慢增长的同时失业率缓慢减少。

对比各州的政府公共服务情况(见图 5)，第一类各州的政府税收总体在温和上涨；第二类中 California 温和下降，Washington 轻微上涨，Texas 明显上涨。第三类 New Jersey 轻微上涨，Illinois 明显上涨。可

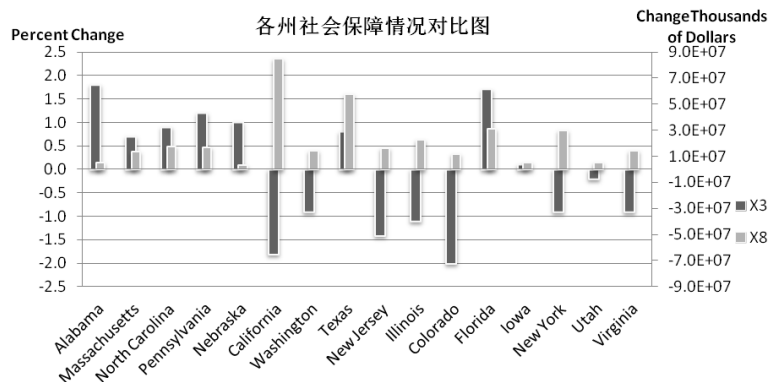


Figure 3. Social security of the 10 state in United State: X₃ and X₈
图 3. 各州社会保障: X₃ 和 X₈

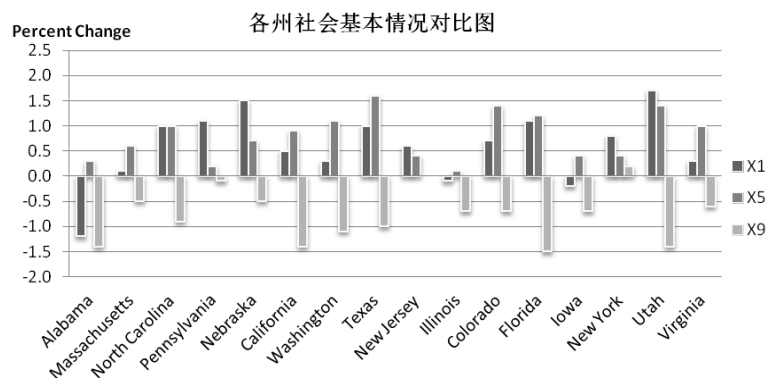


Figure 4. Basic condition of social of the 10 state in United State: X₁, X₅ and X₉
图 4. 各州社会基本情况: X₁、X₅ 和 X₉

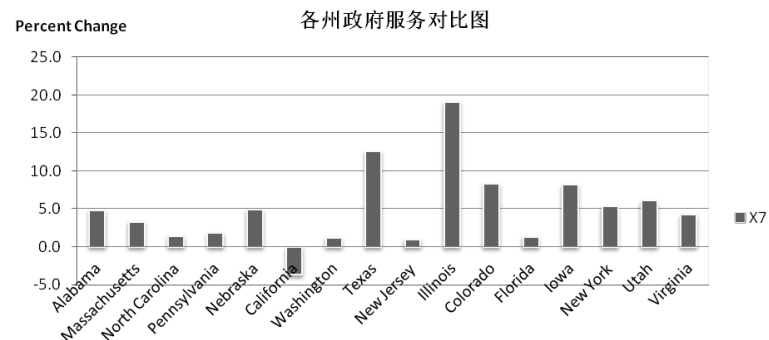


Figure 5. Government service of social of the 10 state in United State: X₇
图 5. 各州政府服务: X₇

以看出每一类各州的情况存在一定的差异。

考虑到四个公因子的方差贡献率各异，故综合四方面情况考虑各州的金融环境时，首先考虑每一类中各州经济基础情况的相似程度，其次社会基本情况，社会保障情况，最后是政府服务情况。再结合专家意见，就可以实现对各州金融生态环境的量化评价：将第二类各州评为 A 级，第三类各州为 B 级，第一类各州为 C 级。A 级的各州经济上升迅猛，州内就业情况可观，劳动力水平上升，社会保障情况乐观。B 级各州的经济底子良好，经济温和回暖，州内就业率上升，劳动力水平上升，政府的服务能力较好，社会保障程度较好。C 级各州的经济也温和回暖，然而社会保障薄弱，劳动力水平和就业率温和上涨。

4.3. 地区金融生态环境判断规则的建立

为了进一步实现对其它各州金融生态环境的量化评价，本小节将利用 Fisher 判别法建立金融生态环境的统一判断规则，并以 Colorado、Iowa、Florida、New York、Utah 和 Virginia 州为例，对所建立的判断规则进行检验。

建模原理：假设有 n 个样本，每个样本有 p 项指标，这些样本分属于 k 个类别 G_1, G_2, \dots, G_k ，每个类别的分布函数为 $F_1(x), F_2(x), \dots, F_k(x)$ 。根据这些信息，利用 Fisher 判别法[8]确定判别函数，进而得到统一判断规则。不妨假设有 k 个总体 G_1, G_2, \dots, G_k ，其均值和协方差矩阵分别为 μ_i 和 $\Sigma_i (>0) (i=1, 2, \dots, k)$ 。考虑线性判别函数 $u'X$ ，在 $X \in G_i$ 的条件下，有

$$\begin{aligned} E(u'X) &= E(u'X|G_i) = u'E(X|G_i) = u'\mu_i, \quad i=1, 2, \dots, k \\ D(u'X) &= D(u'X|G_i) = u'D(X|G_i) = u'\Sigma_i u, \quad i=1, 2, \dots, k \end{aligned} \quad (5)$$

令

$$\begin{aligned} b &= \sum_{i=1}^k (u'\mu_i - u'\bar{\mu})^2 \\ e &= \sum_{i=1}^k u'\Sigma_i u = u' \left(\sum_{i=1}^k \Sigma_i \right) u = u'Eu \end{aligned} \quad (6)$$

其中， $\bar{\mu} = \frac{1}{k} \sum_{i=1}^k \mu_i, E = \sum_{i=1}^k \Sigma_i$ 。这里 b 表示组间差， e 表示组内差。

根据方差分析思想：选择 u 使得函数 $\Phi(u) = b/e$ 达到极大，就可以得到判别规则，若

$$|u'X - u'\mu_j| = \min_{1 \leq i \leq k} |u'X - u'\mu_i| \quad (7)$$

则判定 X 来自总体 G_j 。

将 Colorado、Iowa、Florida、New York、Utah 和 Virginia 州的数据代入判别规则(7)，则可以得到对这 6 个州金融生态环境的判别分析结果，如表 7 所示。

Florida 的一致经济活动指数、GDP 和零售贸易收入增长都位于全美国的平均值左右，批发贸易收入增幅较大，健康保险未覆盖率增长显著，同时个人收入水平温和上涨。该州劳动力水平和居民人口增长的同时失业率显著下降，政府税收收入轻微上涨。这与第一类的各州情况较为接近。Utah 的经济基础各项指标增长均十分迅猛，劳动力水平和居民人口增长的同时失业率显著下降，政府税收收入增幅基本位于均值水平，这些与第二类中各州情况比较接近。New York 经济基础各项指标增长比较迅猛，个人收入水平温和上涨的同时保持较为明显的健康保险未覆盖率下降，这与 New Jersey 和 Illinois 的情况十分相似。

5. 结论

本文基于多元统计理论研究了地区金融生态环境的量化评价问题。首先利用因子分析对现实经济指

Table 7. The results from discriminant analysis
表 7. 判别分析结果

| | 实际分组 | 预测分组 | 判别函数1 | 判别函数2 |
|----------------|------|------|---------------|---------------|
| Alabama | 1 | 1 | -3.028 | -0.697 |
| Massachusetts | 1 | 1 | -1.415 | -0.356 |
| North Carolina | 1 | 1 | -1.129 | -22.067 |
| Pennsylvania | 1 | 1 | -2.671 | 1.118 |
| California | 2 | 2 | 3.634 | -1.635 |
| New Jersey | 3 | 3 | 0.384 | 2.935 |
| Washington | 2 | 2 | 1.446 | -1.937 |
| Illinois | 3 | 3 | 2.328 | 4.599 |
| Nebraska | 1 | 1 | -1.786 | 0.167 |
| Texas | 2 | 2 | 2.236 | -2.127 |
| Colorado | 未分组 | 3 | 1.908 | 1.105 |
| Florida | 未分组 | 1 | -2.062 | -3.352 |
| Iowa | 未分组 | 3 | -0.021 | 1.454 |
| New York | 未分组 | 3 | 0.773 | 3.447 |
| Utah | 未分组 | 2 | 0.700 | -2.565 |
| Virginia | 未分组 | 1 | -0.903 | 0.686 |

标进行了降维，剔除冗余信息，提取出四个具有明显经济学意义的因子；然后运用聚类分析模型得到现实样本的一个划分；最后利用判别分析建立了判别函数，实现了对地区金融环境信用风险的评估。为了验证本文所建立模型的有效性，我们选取了美国十六个州作为研究样本，运用所建立的量化评价模型将这 16 个州分成了 3 类：第一类各州的经济温和回暖，劳动力水平和就业率温和上涨，然而社会保障薄弱；第二类各州经济上升迅猛，州内就业情况可观，劳动力水平上升，社会保障情况乐观；第三类各州经济底子良好，经济温和回暖，州内就业率上升，劳动力水平上升，政府的服务能力较好，社会保障程度较好。结合各州金融环境的实际情况，我们将第二类各州评价为 A 级，第三类各州评价为 B 级，第一类各州评价为 C 级。最后利用判别分析生成了判别函数，建立了一般性的判别规则。

基金项目

感谢陕西省青年科技新星项目(批准号：2013KJXX-19)对本研究的支持。

参考文献 (References)

- [1] Altman, E.I. and Saunders, A. (2011) Credit risk measurement: Development over the last 20 years. *Journal of Banking & Finance*, **21**, 1721-1742.
- [2] 许文彬, 张亦春, 黄瑾轩 (2009) 美国三大评级公司主权评级模式评析. *国际金融研究*, **10**, 26-33.
- [3] 林钧跃 (2012) 社会信用体系理论的传承脉络与创新. *征信*, **1**, 1-13.
- [4] 李扬, 王国刚, 刘煜辉 (2005) 中国城市金融生态环境评价. 人民出版社, 北京.
- [5] 黄国平, 刘煜辉 (2007) 中国金融生态环境评价体系设计与分析. *系统工程理论与实践*, **6**, 7-14.
- [6] 邓淇中, 李鑫, 陈瑞 (2012) 区域金融生态环境指标体系构建及竞争力评价研究. *湖南科技大学学报*, **6**, 75-80.
- [7] Federal Reserve Bank of ST. LOUIS (2012) Economic research. <https://research.stlouisfed.org/>
- [8] 朱建平 (2006) 应用多元统计分析. 科学出版社, 北京.
- [9] 高惠璇 (2004) 应用多元统计分析. 北京大学出版社, 北京.

汉斯出版社为全球科研工作者搭建开放的网络学术中文交流平台。自2011年创办以来，汉斯一直保持着稳健快速发展。随着国内外知名高校学者的陆续加入，汉斯电子期刊已被450多所大中华地区高校图书馆的电子资源采用，并被中国知网全文收录，被学术界广为认同。

汉斯出版社是国内开源（Open Access）电子期刊模式的先行者，其创办的所有期刊全部开放阅读，即读者可以通过互联网免费获取期刊内容，在非商业性使用的前提下，读者不支付任何费用就可引用、复制、传播期刊的部分或全部内容。

