

李守斐,李晗,朱冠旭,等. 基于动态网络方法的中国行业板块联动效应分析[J]. 江西师范大学学报(自然科学版), 2022, 46(1): 25-36.

LI Shoufei, LI Han, ZHU Guanxu, et al. The analysis on linkage effect of Chinese industry sector based on dynamic network method [J]. Journal of Jiangxi Normal University(Natural Science), 2022, 46(1): 25-36.

文章编号:1000-5862(2022)01-0025-12

基于动态网络方法的中国行业板块联动效应分析

李守斐¹, 李 晗², 朱冠旭³, 盛积良^{4*}

(1. 泰康养老保险股份有限公司江西分公司, 江西 南昌 330038; 2. 中国人民银行郴州市中心支行, 湖南 郴州 423000;

3. 南昌三中高中部, 江西 南昌 330029; 4. 江西财经大学统计学院, 江西 南昌 330013)

摘要:该文通过 DCC-MVGARCH 模型和网络分析方法, 建立动态最小生成树网络, 利用 Tucker 分解和 K-均值聚类方法, 将动态网络聚集成 3 个代表网络, 并生成分层结构树图, 对在 2007—2018 年包括次贷危机、“2015 年股市波动”和中美贸易战等多个股市重大事件时间段内的中国行业板块联动性进行实证分析. 研究表明:行业板块联动性长期处于波动状态, 在股价快速上涨、下跌时, 联动性的波动尤为剧烈, 但在代表网络中具有重要影响力的节点变化不大, 分层结构树图聚集状态变化较小, 行业板块联动网络基本稳定; 在中国股票市场的网络结构中生产制造业处于核心地位, 金融业处于边缘地位.

关键词:行业板块联动性; DCC-MVGARCH; 动态最小生成树网络; Tucker 分解

中图分类号: F 830.91 **文献标志码:** A **DOI:** 10.16357/j.cnki.issn1000-5862.2022.01.05

0 引言

行业板块是指证券机构根据某一行业分类标准对股票进行归类所得到的集合. 随着中国金融行业的发展, 金融市场变得更加复杂, 股票之间相依性日益增强, 系统性风险日趋上升, 在这种情况下行业板块受到越来越多的关注. 其原因在于: 一方面是更多的投资者为了规避风险, 减少操作失误, 选择行业配置投资策略来进行资产配置; 另一方面是监管者根据行业板块的划分可以更有效地了解股票市场, 以便拟定与股市相适宜的方针政策. 近年来, 中国经济发展进入新常态, 增速放缓, 尤其是 2015 年股市发生了剧烈波动; 国际经济形势也不稳定, 如 2018 年美国对中国发动了贸易战. 在中国经济下行周期时, 特别是在“2015 年股市波动”和中美贸易战的冲击下, 中国股市的行业板块联动性有何新变化? 这些新变化与次贷危机时期的变化又有什么不同? 在股市风险不断上升的今天, 详细研究股票市场行业板块联动性, 能够深入了解金融市场结构, 掌握各行业

板块在特殊阶段时的相关关系, 进而帮助投资者和监管者预测股市风险的发生及在行业板块之间的传递, 这对投资者的行业配置投资和监管者的风险管理具有重要意义.

在金融行业快速发展的时代, 股票市场已经成为一个典型的复杂系统, 因此复杂网络成为股票市场研究的一个重要方法. R. N. Mantegna^[1]将复杂网络应用在股票领域中, 以股票指数为节点、股票收益率相关系数为边的权重, 生成了标准普尔 500 指数的最小生成树(MST).

此后越来越多的学者采用复杂网络方法对股票市场进行分析. 从国外文献来看, V. Boginask 等^[2]对股票网络进行了统计分析, 发现幂律模型在股票网络中是有效的. J. G. Brida 等^[3]以美国股票市场为例, 提出了构建多维最小生成树的新方法. L. Štefan 等^[4]以标准普尔 100 指数为样本, 比较了动态条件相关(DCC)和滑动窗口这 2 种构建动态网络的方法, 发现基于滑动窗口方法的动态网络更稳健, 但平滑性较差. T. Isogai^[5]提出了以网络聚类算法来处理高维问题, 从而建立高波动金融资产回

收稿日期: 2021-09-08

基金项目: 国家自然科学基金(71973056, 71561011)资助项目.

通信作者: 盛积良(1972—), 男, 江西余干人, 教授, 博士, 博士生导师, 主要从事金融工程与风险管理、金融计量与金融统计等研究. E-mail: shengjiliang@163.com

报率动态网络的新方法. A. Sensoy 等^[6]基于动态条件相关方法构建了亚洲股市关联的动态最小生成树,发现最小生成树随着时间增加而缩小. Wang Gangjin 等^[7]基于网络分析方法对国际房地产证券市场的相关结构进行了研究. E. Kantar 等^[8]通过分层结构方法构建网络,发现金融危机对土耳其股票网络的影响不明显. O. Shirokikh 等^[9]利用网络模型分析了整个美国股票市场的股价波动趋势. A. Nobil 等^[10]研究了金融危机对韩国股票网络的影响,发现在金融危机期间韩国股票网络规模收缩,所选公司股票之间的共生相关系数增大.

国内许多学者采用计量模型对股票市场行业联动性进行了研究. 李俊照等^[11]利用马尔可夫链时序回归模型,选择与房地产板块相关的行业板块,构建房地产板块的回归模型,实现了对房地产板块股价指数的有效预测. 宁建楠等^[12]采用 GARCH-Copula 模型分析了金融危机前后中国股市 4 个行业板块间相依结构的变化,发现各市场间倾向联动,金融危机后相依性增强. 刘井建等^[13]以能源行业、原材料行业和消费行业为例,采用非线性 STR 模型对市场行业风险联动性进行实证分析,发现金融危机后行业风险联动趋势增强. 马丹等^[14]用动态层级模型分析了市场板块的关联效应和传染效应,发现在危机事件爆发时板块之间的联动性增强. 曾裕峰等^[15]基于非对称 MVMQ-CAViaR 模型探究了金融业板块间风险的传导. 上述文献虽然从多个方面研究了行业间联动性的变化规律,但大多数是选取几个代表行业来进行联动性研究,未能反映股市的整体动态规律.

近年来,国内不少学者采用复杂网络方法对股票、行业板块的联动性进行了研究. 庄新田等^[16]通过阈值法建立中国股市网络,从时间和空间 2 个角度对中国股市网络的分形特征进行了研究. 李岸等^[17]利用阈值法、最小生成树法和 DCC-MVGARCH 模型方法分析了中国股票市场国际联动性的变化. 王荣森等^[18]选取了“2015 年股市波动”前后中小板流通市值前 60 的股票,用最小生成树和社区网络方法构建了股票的社区结构图,对“2015 年股市波动”前后的中小板上市公司网络结构进行了对比分析. 黄飞雪等^[19]利用最小生成树和指数分层结构对上证 50 指数的行业风格进行了研究. 谢邦昌等^[20]采用了平面最大过滤图和 InfoMap 算法研究了金融危机对中信行业指数网络的影响,发现在金融危机期间网络联动效应最强.

虽然现有文献从不同角度对中国股票市场的行业板块联动性进行了分析,但是相关研究仍存在一些不足:(i)现有文献主要集中在某一次股市剧烈波

动(如次贷危机、2015 年股市波动)对行业板块联动性影响上,不能精确地得到中国行业板块联动性的变化规律;(ii)大多数文献是通过建立静态网络来对中国行业板块联动性进行研究,这无法反映行业板块联动性的动态变化. 相比静态网络而言,动态网络可以反映股票市场每日的变化,对股市的分析更加精准. 因此,本文在现有文献研究的基础上,首先采用标普中国 500 指数中 24 个行业股票指数历史数据,通过 DCC-MVGARCH 模型计算各行业股指间的每日动态相关系数;然后建立动态最小生成树网络;最后参照文献[5]的方法,利用张量低秩逼近方法对子网络进行聚类,发现股价剧烈波动对行业板块联动性有很大的影响.

1 行业动态相关系数

本文以行业板块股票指数为节点、行业板块股票指数的相关系数为边的权重来构建动态行业板块联动网络^[1]. 为此利用 DCC-MVGARCH 模型计算各行业股指间的每日动态相关系数.

1.1 DCC-MVGARCH 模型

R. Engle^[21]提出的 DCC-MVGARCH 模型可以很好地刻画股票间的动态相关性,其设定为

$$r_t = u_t + e_t, e_t | \Omega_{t-1} \sim N(0, H_t), \quad (1)$$

$$H_t = D_t R_t D_t, \quad (2)$$

$$Q_t = \left(1 - \sum_{m=1}^M \lambda_{1m} - \sum_{n=1}^N \lambda_{2n}\right) \bar{Q} + \sum_{m=1}^M \lambda_{1m} (\varepsilon_{t-m} \cdot \varepsilon'_{t-m}) + \sum_{n=1}^N \lambda_{2n} Q_{t-n}, \quad (3)$$

其中式(1)的误差扰动项 $\{e_t\}$ 为独立同分布的白噪声过程,它服从均值为0、协方差矩阵为 H_t 的条件多元正态分布, Ω_{t-1} 为时刻 $t-1$ 及以前的信息集;式(2)为方差方程,这里 $R_t = (Q_t^*)^{-1} Q_t (Q_t^*)^{-1}$,且

$$Q_t = \begin{pmatrix} q_{11,t} & q_{12,t} & \cdots & q_{1n,t} \\ q_{21,t} & q_{22,t} & \cdots & q_{2n,t} \\ \cdots & \cdots & \cdots & \cdots \\ q_{n1,t} & q_{n2,t} & \cdots & q_{nn,t} \end{pmatrix},$$

$$Q_t^* = \begin{pmatrix} \sqrt{q_{11,t}} & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22,t}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sqrt{q_{nn,t}} \end{pmatrix},$$

$$D_t = \text{diag}\{\sqrt{h_{ii,t}}\}, h_{ii,t} = \omega_i + \sum_{p=1}^{P_i} \alpha_{ip} e_{i-p}^2 + \sum_{q=1}^{Q_i} \beta_{iq} h_{i-q}.$$

式(3) 中 \overline{Q} 为标准残差的无条件方差矩阵, $\overline{Q} = T^{-1} \sum_{t=1}^T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' = \mathbf{D}_t^{-1} \boldsymbol{\varepsilon}_t$ 为标准化残差向量, λ_{1m} 和 λ_{2n} 是 DCC-MVGARCH 模型的系数。

本文使用的是 DCC-MVGARCH(1,1), 且 $q_{ij,t}$ 为 $q_{ij,t} = (1 - \lambda_1 - \lambda_2) \bar{\rho}_{ij} + \lambda_1 \varepsilon'_{i,t-1} \varepsilon'_{j,t-1} + \lambda_2 q_{ij,t-1}$, 其中 $\bar{\rho}_{ij}$ 为标准化残差 $\boldsymbol{\varepsilon}_t$ 的相关系数, λ_1 和 λ_2 是 DCC 系数, 且 $\lambda_1 > 0, \lambda_2 > 0, \lambda_1 + \lambda_2 < 1$ 。DCC-MVDARCH 模型通过 2 阶段极大似然法求解: 第 1 阶段利用 GARCH 模型对每个股票指数收益率序列进行估计,

计算标准化残差; 第 2 阶段利用极大似然法对 DCC 模型参数进行估计, 计算动态条件相关系数。

1.2 实证数据

本文采用的数据是 2007 年 1 月 10 日—2018 年 9 月 20 日的标普中国股价指数日收盘价数据, 标普中国指数的分类标准与国际上普遍使用的全球行业分类标准 (GICS) 一致, 使得本文的结果具有可比性, 其中每个行业指数有 2 847 个观测值, 数据来自 Wind 数据库。各行业名称及股票指数的描述性统计如表 1 所示。

表 1 标普中国股价指数描述性统计结果

行业	最小值	最大值	均值	标准差	偏度	峰度
半导体	562.2	5 337.7	2 221.0	945.4	0.58	-0.46
保险	388.5	2 286.1	983.0	394.7	1.16	0.66
餐饮休闲	928.0	6 043.1	2 686.8	895.8	0.63	-0.31
电信服务	1 389.5	6 207.7	2 752.5	827.5	1.09	1.38
多样金融	1 772.5	8 386.3	3 435.6	1 221.8	1.30	1.52
房地产	1 420.4	7 855.9	3 694.0	1 275.5	0.46	-0.97
公用事业	1 298.1	5 053.6	2 261.1	637.7	1.04	1.16
技术硬件	838.6	7 560.5	3 000.4	1 197.0	0.69	-0.21
家庭和个人用品	920.3	4 625.8	2 324.9	595.8	0.35	-0.06
健康设备	1 056.7	12 105.1	4 342.4	2 219.7	0.64	-0.64
零售业	1 240.0	7 437.0	3 177.4	880.9	0.69	1.51
媒体	852.4	6 056.6	2 262.2	720.8	1.43	4.04
耐用品	1 204.2	9 982.1	4 227.2	1 905.9	0.71	-0.75
能源	1 238.6	5 288.5	2 246.6	635.7	1.32	2.20
汽车汽配	863.0	8 261.8	3 612.1	1 390.4	0.41	-0.57
软件服务	1 187.8	16 712.9	4 398.7	2 523.7	1.38	2.33
商品服务	1 210.6	10 410.1	3 780.9	1 791.8	1.04	0.30
商业银行	1 420.2	5 020.1	2 851.5	721.7	0.44	-0.74
食品药业	846.2	4 140.6	1 959.3	526.7	0.44	0.35
食品饮料	1 349.4	7 187.9	3 675.7	1 235.4	0.78	0.09
原材料	1 123.1	4 952.4	2 617.4	686.1	0.40	0.07
运输业	891.5	3 810.9	1 775.2	588.7	0.94	0.92
制药生物技术和科学	1 301.8	11 142.8	4 959.8	2 050.6	0.41	-0.86
资本品	1 126.3	7 396.0	3 080.5	944.0	0.88	1.66

图 1 为 2007—2018 年标普中国股价指数日平均收盘价和上证综指价格的变化情况, 表 2 为图 1 中黑色辅助线对应的日期。在 2007 年股权分置改革完成以及资本市场制度进一步完善后, 股指平均收盘价快速上升, 上证综指呈相同的变化趋势, 不同的是上证综指在辅助线 1 (即 2007 年 10 月 16 日) 处达到顶峰, 而股指平均收盘价却在辅助线 2 (即 2008 年 1 月 14 日) 处达到顶峰, 这表明股指平均收盘价和上证综指的 2007 年牛市结束时间不同; 2008 年受次贷危机影响, 股指平均收盘价和上证指数一直下跌; 在 2008 年 9 月雷曼公司破产、次贷危机全面爆发后, 股指平均收盘价和上证综指却从 2008 年 11 月开始

回升, 这是因为 2008 年 11 月中国实施了“四万亿”投资计划, 刺激了经济的增长; 在 2010 年 ~ 2014 年上半年这段时间内, 股市低迷, 股指平均收盘价和上证综指变化不大; 2014 年 4 月中旬股指平均收盘价和上证综指急剧上涨, 2015 年 6 月中旬达到顶峰; 在“2015 年股市波动”后股指平均收盘价和上证综指急剧下跌, 直至 2016 年 2 月才趋于稳定。2016—2017 年, 受“2015 年股市波动”后续影响, 股指平均收盘价和上证综指持续低迷; 2018 年 3 月 23 日, 美国前任总统特朗普在白宫正式签署对中国贸易备忘录, 随着中美贸易战开启, 国际经济形势变得不稳定, 这导致股指平均收盘价和上证综指下跌。

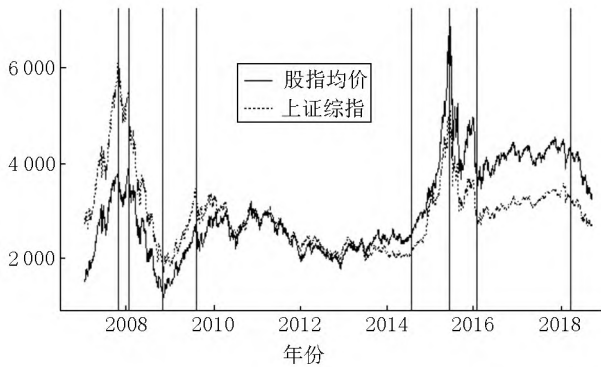


图1 2007—2018年标普中国股价指数日平均收盘价和上证综指

表2 图1中黑色辅助线对应日期

辅助线	日期
1	2007年10月16日
2	2008年1月14日
3	2008年10月28日
4	2009年8月04日
5	2014年7月22日
6	2015年6月12日
7	2016年1月27日
8	2018年3月22日

注:辅助线排序从左往右,辅助线1、3、4、5、6、7根据上证综指股市波动的临界点所设置。

从总体来看,除2007年牛市结束时间外,股指平均收盘价和上证综指的变化趋势基本一致。自2007年后,中国股市有2次巨幅波动,分别是次贷危机前后和“2015年股市波动”前后,由于次贷危机的冲击被2008年11月的“四万亿”投资计划所抵消,因此本文将次贷危机时期分为“2008年股灾”和“2009年牛市”2个时期。另外,2018年的中美贸易战对股市也有不小的影响。本文着重对这3次股市波动时期进行分析,各时期股市波动的时间区间设置参照了上证综指变化的临界点,并根据股指平均收盘价的实际情况进行调整,具体如表3所示。

此外,由于股票指数数据为金融时间序列,较为不稳定,因此将原始数据按对数差分收益率进行处理,对数差分收益率计算公式为

$$R_{i,t} = \ln(p_{i,t}/p_{i,t-1}),$$

其中 $p_{i,t}$ 为股票 i 在 t 时刻处的收盘价。

表3 股市波动的时间

股市波动时期	时间
2007年牛市	2007年1月10日—2008年1月14日
2008年股灾	2008年1月14日—2008年10月28日
2009年牛市	2008年10月28日—2009年8月4日
2014年牛市	2014年7月22日—2015年6月12日
2015年股市波动	2015年6月12日—2016年1月27日
2015年股市波动后	2016年1月27日—2018年3月23日
中美贸易战	2018年3月23日—2018年9月20日

注:数据时间区间为2007年1月10日—2018年9月20日。

1.3 DCC-MVGARCH 模型实证结果

应用DCC-MVGARCH模型,对各行业股票指数收益率进行估计,得到了收益率序列单变量GARCH系数、DCC系数和每日行业动态相关系数。

表4为各行业股票指数收益率序列单变量GARCH模型系数,其中 α 为ARCH项系数, β 为GARCH项系数,所有股票指数的 $\alpha + \beta$ 均高于0.99且接近于1,这说明行业股票指数波动趋势未来将得到持续。

表4 股票指数收益率序列单变量GARCH模型系数

行业	α	β	$\alpha + \beta$
半导体	0.051 4	0.946 7	0.998 1
耐用品	0.053 3	0.942 3	0.995 6
保险	0.054 9	0.943 0	0.997 9
能源	0.045 9	0.951 4	0.997 3
餐饮休闲	0.047 0	0.949 6	0.996 6
汽车汽配	0.043 3	0.953 8	0.997 1
电信服务	0.054 7	0.937 3	0.992 0
软件服务	0.048 0	0.942 7	0.990 7
多样金融	0.042 4	0.955 7	0.998 1
商品服务	0.043 8	0.949 9	0.993 7
房地产	0.051 3	0.945 7	0.997 0
商业银行	0.063 6	0.934 4	0.998 0
公用事业	0.062 2	0.936 4	0.998 6
食品药业	0.057 9	0.936 1	0.994 0
技术硬件	0.051 3	0.940 7	0.992 0
食品饮料	0.060 4	0.933 0	0.993 4
家庭和个人用品	0.069 1	0.923 6	0.992 7
原材料	0.054 3	0.937 2	0.991 4
健康设备	0.053 6	0.941 3	0.994 9
运输业	0.047 0	0.949 5	0.996 5
零售业	0.058 3	0.933 1	0.991 4
制药生物技术和科学	0.052 4	0.946 2	0.998 7
媒体	0.051 9	0.944 6	0.996 4
资本品	0.050 2	0.944 7	0.995 0

DCC-MVGARCH模型的 λ_1 系数为0.0134, λ_2 系数为0.9763,且模型系数显著性检验 p 值均为0.00,这说明行业板块之间存在联动性。

图2(a)、(b)、(c)、(d)分别为整个时期动态相关系数均值、标准差、偏度和峰度的变化情况,表5显示的是股价波动时期动态相关系数均值、标准差、偏度和峰度的变化趋势。从图2(a)、(b)来看,在次贷危机、“2015年股市波动”期间,均值、标准差呈现相同的变化趋势,在股价快速下跌时,相关系数均值上升,标准差下降,这说明股指相关性增强,离散程度下降。在股价快速下跌结束后,相关系数均值下降,标准差上升,这说明股指相关性减弱,离散程度上升。中美贸易战期间相关系数均值、标准差的变化说明了中美贸易的摩擦使得中国股指相关性增强。同时发现:在中国股市的历次股价快速下跌中,“2015年股市波动”对股指相关性的影响最大;在

“2015 年股市波动”结束时,相关系数均值达到了历史最高点.

值得注意的是:在 2007 年牛市期间,相关系数均值下降,标准差上升,股指相关性减弱,离差程度上升;而在 2014 年牛市期间,相关系数均值先上升后下降,标准差先下降后上升,且变化幅度均比较小,相关性变化不明显.从相关系数均值、标准差来

看,2007 年牛市和 2014 年牛市对股指相关性的影响不同.但相关系数均值和标准差的变化只能反映各股指之间相关性的变化,这难以准确地反映行业板块网络联动性的变化,其原因是未经过处理的相关系数矩阵反映的信息是有噪声的^[17].为解决这个问题,本文用最小生成树网络对行业板块联动性进行进一步分析.

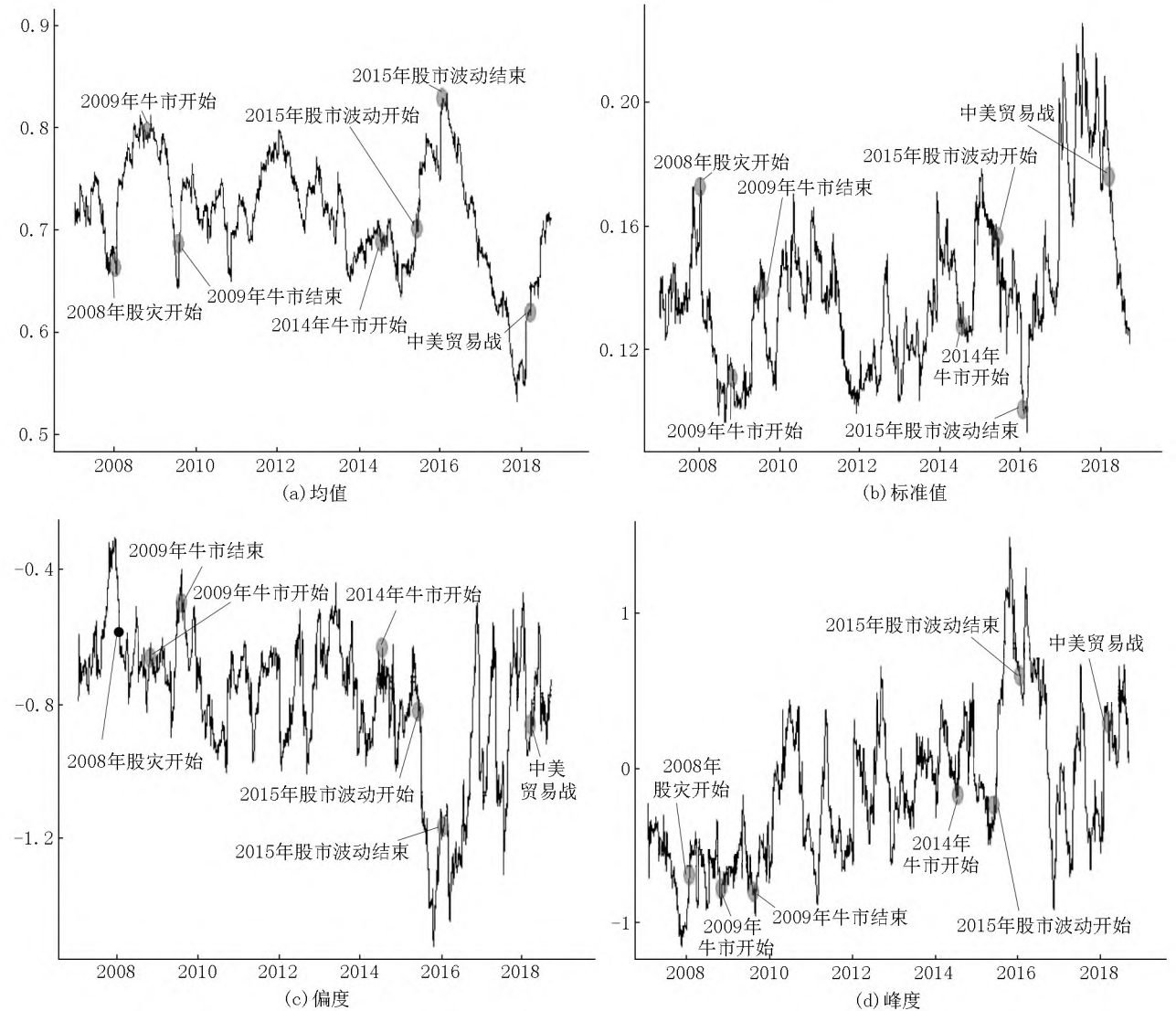


图 2 动态相关系数均值、标准差、偏度和峰度变化

表 5 股市波动时期动态相关系数均值、标准差、偏度、峰度变化趋势

时期	统计量变化趋势			
	均值	标准差	偏度	峰度
2007 年牛市	下降	上升	不明显	不明显
2008 年股灾	上升	下降	不明显	不明显
2009 年牛市	下降	上升	不明显	不明显
2014 年牛市	先下降后上升	先上升后下降	不明显	不明显
2015 年股市波动	上升	下降	下降	上升
2015 年股市波动后	下降	上升	上升	下降
中美贸易战	上升	下降	不明显	不明显

从图 2(c)、(d) 来看,次贷危机和中美贸易战对相关系数偏度和峰度没有明显的影响.而在“2015 年股市波动”期间,相关系数偏度呈下降趋势,峰度呈上升趋势;在“2015 年股市波动”结束后,偏度呈上升趋势,峰度呈下降趋势.为了更好地观察相关系数分布的变化,图 3 通过 Jarque-Bera 系数来描述相关系数分布的变化.

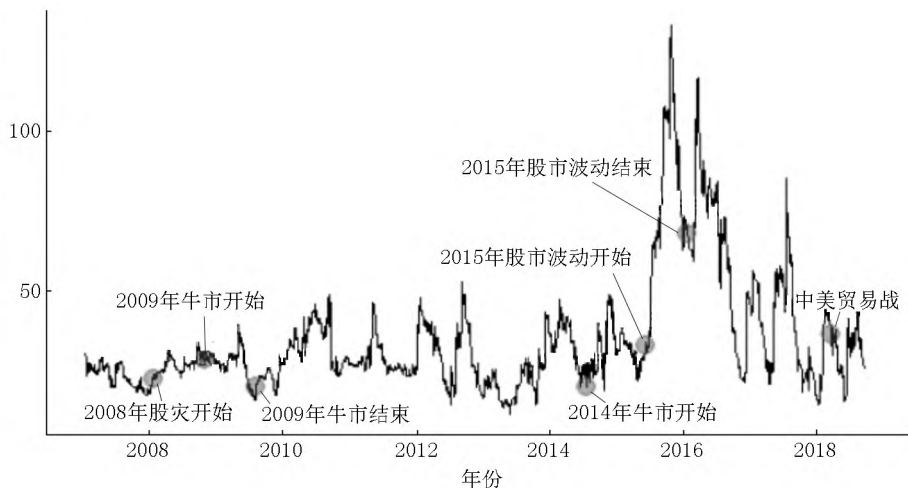


图 3 Jarque-Bera 系数

2 动态最小生成树网络

2.1 网络分析方法

2.1.1 最小生成树 根据动态相关系数构建的 $N \times N$ 相关系数矩阵为 $A = (c_{ii})_{N \times N}$, 由 A 构建的初始关联网络为全连接网络, 此时复杂网络反映的信息大部分是无用的. 因此要通过最小生成树法, 在相关系数的基础上, 滤去冗余的信息, 建立真正的关联网络^[17].

仿照文献[1]的方法建立最小生成树网络. 首先将相关系数转化为欧氏距离, 用它来表示 2 个股票指数之间的相对距离:

$$d_{ij} = \sqrt{2(1 - c_{ij})}.$$

再使用 Kruskal 算法将欧氏距离矩阵转化为最小生成树. Kruskal 算法具体如下:

(i) 将每个指数作为节点, 指数间的关联性作为边. 视每一个节点为独立的, 将指数间的关联系数作为边的权重.

(ii) 在生成网络后, 从最小边开始, 并检测当前所选边的加入是否会构成回路, 若不构成回路, 则将该边并入最小生成树中.

(iii) 依次类推, 直至在网络图中含有 $n - 1$ 条边为止.

由图 3 可知, 股票指数相关系数在大部分时期内拒绝正态分布. Jarque-Bera 系数在“2015 年股市波动”前后时间段内波动剧烈, 而在其他时间段内, 波动较小, 比较平稳, 这进一步说明“2015 年股市波动”对行业板块联动性的影响与其他股市波动是不同的.

2.1.2 网络拓扑属性 本文采用网络拓扑属性对动态网络进行分析. 节点度是指与节点相连边的条数, 又称关联度. 一个节点的节点度越大意味着这个节点的度中心性越高, 该节点在网络中就越重要.

节点强度是指与节点相连边的权重之和, 被用来衡量该节点与网络中其他节点之间的相关程度.

接近中心性度量的思想是: 如果一个节点与许多其他节点都很“接近”, 那么该节点处于网络的中心位置. 其定义为该节点到其他所有节点距离之和的倒数, 即

$$c_{Cl}(v) = 1 / \sum_{u \in V} f(v, u),$$

其中 $f(v, u)$ 是节点 $u, v \in V$ 的捷径距离.

中介中心性度量是某个节点在多大程度上“介于”其他节点对之间. 对于给定节点 v , 其定义为

$$c_B(v) = \sum_{s \neq t \neq v \in V} \delta(s, t | v) / \delta(s, t),$$

其中 $\delta(s, t | v)$ 是 s 与 t 之间通过 v 的最短路径数量, $\delta(s, t)$ 是 s 与 t 之间的最短路径总数.

网络的平均路径长度是任意 2 个节点之间距离的平均值, 其计算公式为

$$\bar{l} = \frac{2}{N(N-1)} \sum_{i > j} l(i, j),$$

其中 $l(i, j)$ 是网络中 2 个节点 i 和 j 的距离, 其定义为连接这 2 个节点的最短路径的边数.

2.2 动态网络

先根据每日动态相关系数计算出距离矩阵,然后利用 Kruskal 算法将距离矩阵转化为动态最小生

成树网络(即动态行业板块联动网络),再运用网络拓扑属性对动态网络进行特征化,最后得到的结果如图 4 和图 5 所示。

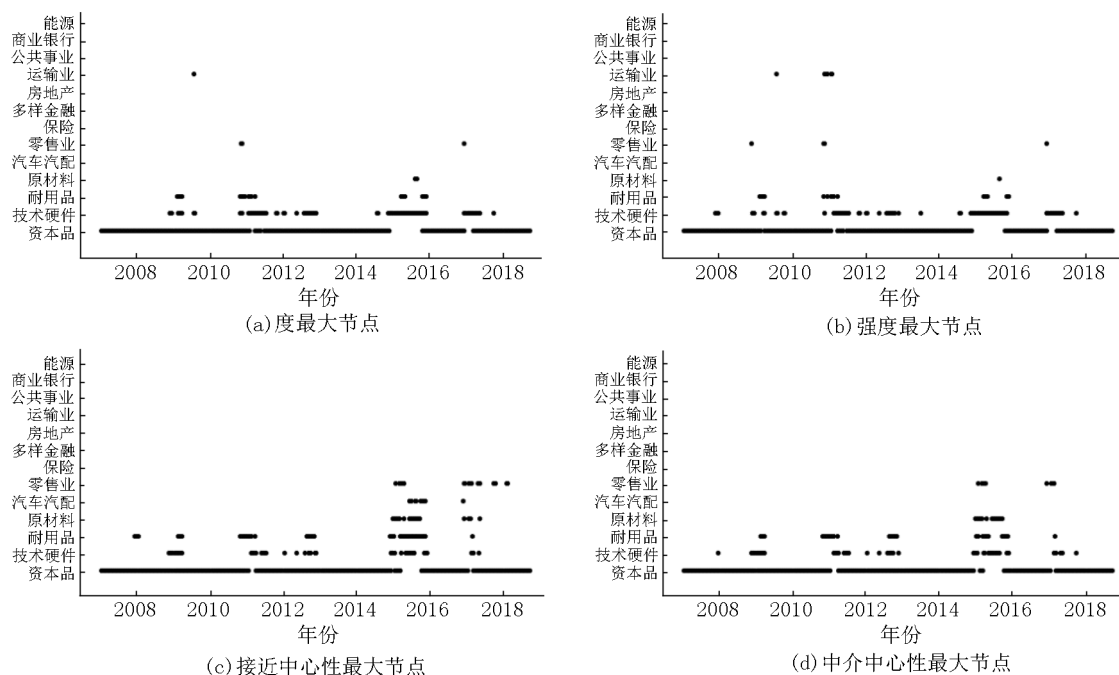


图 4 动态网络 4 个中心性最大节点散点图

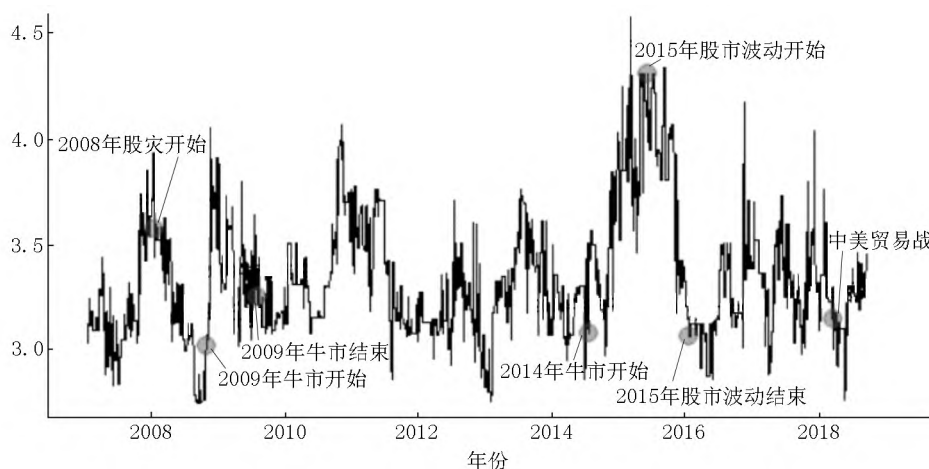


图 5 平均路径长度

图 4(a)、(b)、(c)、(d) 分别显示的是节点度、节点强度、接近中心性和中介中心性最大节点的变化,散点的含义是其对应的行业股指为在对应日期网络中的中心性最大节点。表 6 显示的是在网络中重要股指的子行业。节点度、节点强度、接近中心性和中介中心性被用于衡量节点中心性的大小。从图 4 可以看出:一方面,除 2015 年外,资本品一直是联动网络的核心节点,而股市剧烈波动会导致中心性最大节点发生变化,如在次贷危机中,技术硬件、耐用品和资本品一起成为联动网络的中心性最大节点,尤其是在 2015 年,技术硬件取代了资本品成为了联动网络的核心节点,整个动态联动网络的核心

节点是资本品或技术硬件,这表明在中国股票市场中生产制造业一直处于核心地位;另一方面,在中国经济中日益繁荣的金融行业在股票网络中却处于边缘地位,这是因为金融板块股票的市盈率一直处于低位,如银行股市盈率长期处于 10 左右,涨跌幅度很小,对于投资者来说,波动很小的金融板块是不具有投资价值的。这与其他国家的研究结果不同,在股市平稳期间,在韩国股票网络中金融证券类公司是核心节点^[10],美国股票市场由金融部门所主导^[22],在德国股票网络中银行业是核心节点^[23],这些发达国家平稳时期的股票网络核心节点均为金融行业。这意味着中国与发达国家的金融市场结构不同。

表 6 重要股指的子行业

股指	子行业
资本品	航空航天与国防、建筑产品、建筑与工程、电气及机械等
技术硬件	通信设备、电脑与外围设备、电子设备及元器件
耐用品	家庭耐用消费品、休闲设备与用品、纺织品、服装和奢侈品
原材料	化学制品、建筑材料、容器与包装、金属与采矿、纸类与林业产品

对于复杂网络而言,相关文献经常使用平均路径长度和聚类系数这 2 个特征来衡量网络的小世界性,若一个网络图平均路径长度较短、聚类系数较大,则该网络图具有小世界性.网络图具有小世界性意味着网络更加紧密,节点之间的联动性更强.然而,最小生成树网络属于二分图,其聚类系数为 0,因此最小生成树网络不能用聚类系数来衡量. A. Sensoy 等^[6]认为在最小生成树网络中,平均路径长度的减少意味着网络变得更紧密以及联动性增强.本文采用平均路径长度^[6]来衡量最小生成树网络联动性的大小.

表 7 相关系数均值、网络平均路径长度及联动性的变化趋势

时期	变化趋势			
	相关系数均值	股指相关性	网络平均路径长度	板块联动性
2007 年牛市	下降	减弱	上升	减弱
2008 年股灾	上升	增强	下降	增强
2009 年牛市	下降	减弱	先上升后下降(整体表现为上升)	减弱(整体趋势)
2014 年牛市	先下降后上升	不明显	上升	减弱
2015 年股市波动	上升	增强	下降	增强
2015 年股市波动后	下降	减弱	上升	减弱
中美贸易战	上升	增强	先下降后上升,再稳定于某个程度	不明显

综上所述,中国行业板块联动性长期处于波动状态,在股价快速上涨、下跌时,联动性的波动尤为剧烈,具体表现为:股价快速上涨会使联动网络更稀疏,行业板块联动性减弱;而股价快速下跌会使联动网络更紧密,行业板块联动性增强,但在股价快速下跌结束后,联动网络会变得稀疏,行业板块联动性减弱.

3 动态网络聚类

3.1 聚类方法

参照文献[5]的方式:第 1 步,将动态网络邻接矩阵看作一个 3 阶张量,利用 Tucker 分解方法分解为矩阵;第 2 步,将含有时间信息的矩阵转化为向量,再通过 K-均值聚类方法对含有时间信息的向量进行聚类.

张量是向量或者矩阵的高阶泛化.如向量是 1

图 5 显示的是整个时期动态网络平均路径长度的变化,表 7 显示的是动态相关系数均值、股指相关性、动态网络平均路径长度及联动性的变化趋势,除 2014 年牛市和中美贸易战时期外,动态相关系数均值所反映的股指相关性的变化和行业板块联动性的变化相一致,这说明动态相关系数均值对行业板块联动性有一定的预测价值.

与动态相关系数均值变化相似,“2015 年股市波动”期间平均路径长度波动最大,这说明“2015 年股市波动”期间行业板块联动性的变化最剧烈.值得注意的是:在 2010 年末至 2011 年初期间,平均路径长度波动较大,而这段时间股市相对次贷危机来说却比较平稳,对于这个问题,本文在网络聚类方法中做进一步的分析;在中美贸易战发生后,平均路径长度先下降后上升,再稳定于某个程度,这表明中美贸易战对中国行业板块联动性造成了一定的影响,但由于中美贸易战发生在 2018 年,所能得到与之相关数据的时间较短,所以无法明确中美贸易战对行业板块联动性的影响.

阶张量,矩阵是 2 阶张量,本文的动态网络邻接矩阵是 3 阶张量,每日网络的邻接矩阵是 2 阶张量.

首先,Tucker 分解.以 3 阶张量为例,假设 \mathbf{x} 是大小为 $n_1 \times n_2 \times n_3$ 的张量,进行 Tucker 分解后的表达式可以写成:

$$\mathbf{x} \approx \mathbf{g} \mathbf{x}_1 \mathbf{U} \mathbf{x}_2 \mathbf{V} \mathbf{x}_3 \mathbf{W},$$

其中张量 \mathbf{g} 的大小为 $r_1 \times r_2 \times r_3$ (也称为核心张量),矩阵 \mathbf{U} 的大小为 $n_1 \times r_1$,矩阵 \mathbf{V} 的大小为 $n_2 \times r_2$,矩阵 \mathbf{W} 的大小为 $n_3 \times r_3$, \mathbf{x}_1 、 \mathbf{x}_2 和 \mathbf{x}_3 是对应模块的张量积.

然后,使用 Tucker 分解的结果进行动态网络邻接矩阵的子空间聚类. \mathbf{W} 包含着动态网络邻接矩阵的时间信息,所以 \mathbf{x} 对子空间的投影定义为

$$\mathbf{Y} = \mathbf{g} \mathbf{x}_1 \mathbf{W}.$$

最后,将 \mathbf{Y} 转化为向量,使用 K-均值聚类方法对其聚类.

3.2 动态网络聚类结果

标普中国股价指数共有24个行业指数,所选时间共有2847个交易日,因此动态网络的邻接矩阵可以看作是一个 $24 \times 24 \times 2847$ 的张量,核心向量 \mathbf{g} 的秩数设置参照文献[5],为了减少邻接矩阵的维度,令 $r_1 = r_2 = 3$,为了保证张量有关时间的信息尽可能地保存,令 $r_3 = n_3 = 2847$.在K-均值聚类中令 $k=3$,将动态网络聚集成3个代表网络.

为了直观地展示网络聚类分期结果,将图5中网络平均路径长度的变化进行分期标记.图6把

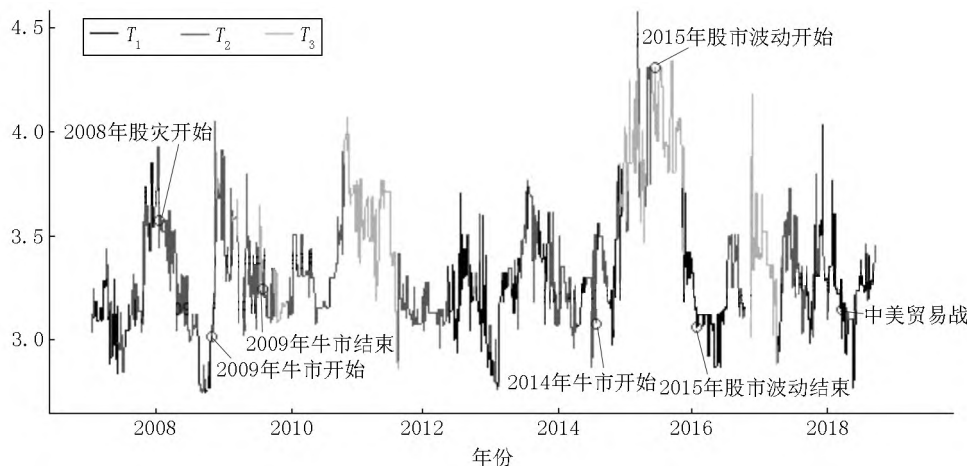


图6 网络聚类结果

图7显示的是2007—2018年 T_1 、 T_2 和 T_3 时期在一年中所占的百分比,不同的年份的每个子时期的占比有较大的差异。 T_1 在2011年之前占比较高,2011年后有所下降, T_2 在2011年后占比增加。而 T_3 在2015年占比最大,这说明“2015年股市波动”对行业板块联动性的影响比次贷危机更深刻; T_3 在2011年占比超过了50%,在2010年缓慢牛市结束后,2011年前期股市震荡上涨,4月份相关部门实施货币紧缩政策,导致股价下跌,在这段时间内,虽然股价相对次贷危机来说波动较小,但是由于股市多处于震荡状态,所以主力可以借此洗盘和震仓,这使得股票市场处于高应激期,这也是从2010年末至2011年初期间平均路径长度波动较大的原因。

通过张量分解和K-均值聚类能够将动态网络邻接矩阵的维数从2847(交易日)减少到3.这种降维简化了在时间轴上动态联动网络的比较,可以便于总结出与代表网络相对应的子时期的差异.具体地,首先对每个子时期的动态相关系数进行加权平均,得到3个子时期的相关系数;再通过最小生成树方法创建子时期的代表网络;最后通过使用3个代表网络来总结整个中国股票市场的动态相关结构。

图8(a)、(b)、(c)分别是 T_1 、 T_2 、 T_3 时期代表网

2007年1月10日—2018年9月20日共2847个交易日分在 T_1 、 T_2 和 T_3 3个子时期中.由图6可知, T_1 和 T_2 时期的网络分散在整个期间,而 T_3 时期的网络(浅灰色折线)多集中在平均路径长度剧烈波动的时期.通过对各个时期平均路径长度波动的比较发现: T_3 处于高应激期, T_1 和 T_2 时期处于普通应激期.而在T. Isogai^[5]对日本股市的研究中,次贷危机时期属于最高级别的波动时期,这是因为相比日本的救市措施来说,中国的“四万亿”计划更有力度,使得次贷危机对中国股市的影响较小。

络,表8是各时期代表网络的重要节点度.由图8发现 T_1 和 T_2 时期的最小生成树网络变化不大,表8中 T_1 和 T_2 重要节点的变化仅有排名第5的制药生物技术和科学变为了耐用品,而 T_3 与 T_1 、 T_2 相比,网络图核心节点发生了变化,技术硬件取代了资本品成为核心节点;但 T_1 、 T_2 和 T_3 时期前5的节点始终有资本品、技术硬件、原材料这3个节点,而且在 T_1 、 T_2 时期技术硬件是网络图第2大节点,在 T_3 时期资本品是网络图第2大节点,虽然行业板块联动性一直有不小的波动,但是在网络中发挥重要作用的节点的改变很小,这说明中国行业板块联动网络比较稳定.为了进一步分析联动网络的稳定性,本文建立了指数分层结构对联动网络进行研究。

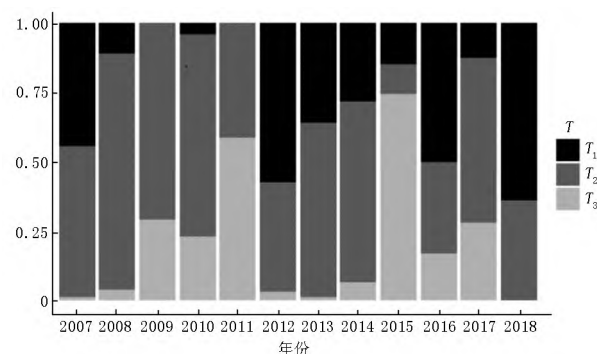


图7 2007—2018年 T_1 、 T_2 和 T_3 时期百分比

表 8 T_1 、 T_2 和 T_3 时期网络重要节点度

时段	标号	1	2	3	4	5
T_1	节点	资本品	技术硬件	原材料	零售业	制药生物技术和科学
	节点度	12	5	2	2	2
T_2	节点	资本品	技术硬件	原材料	零售业	耐用品
	节点度	10	6	2	2	2
T_3	节点	技术硬件	资本品	耐用品	原材料	制药生物技术和科学
	节点度	8	6	4	2	2

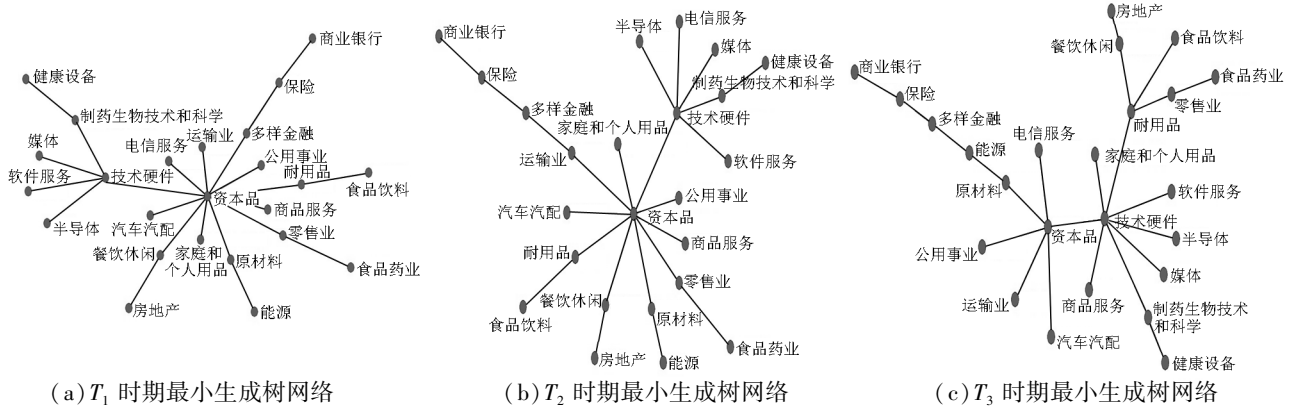


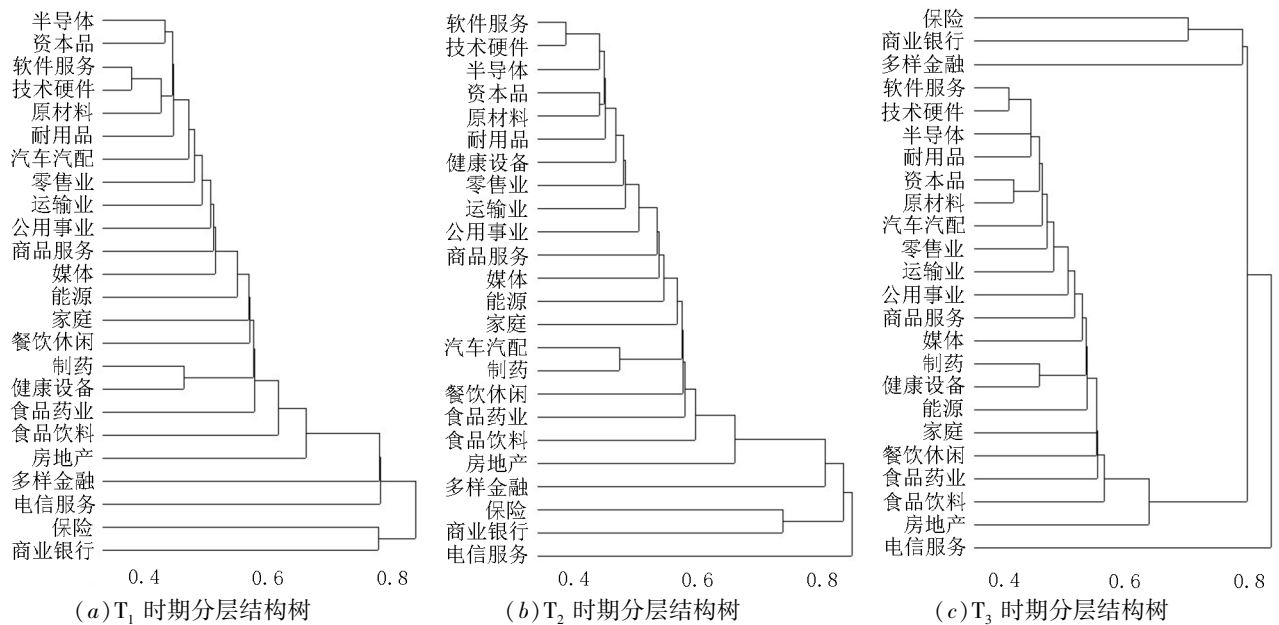
图 8 T_1 、 T_2 和 T_3 时期的最小生成树

3.3 指数分层结构

在建立最小生成树网络后,可通过最小生成树网络确定超度量距离 \tilde{c}_{ij} 矩阵,在该矩阵中的每个元素 \tilde{c}_{ij} 等于当沿着连接起始对象和终止对象的最短 MST 路径移动时,经过的任意 2 个相邻目标之间的

最大距离,再由该矩阵生成指数分层结构树图^[19].

图 9(a)、(b)、(c)分别为 T_1 、 T_2 、 T_3 时期的行业股票指数分层结构树图,横坐标为行业股票指数名称,纵坐标为行业股票指数距离,指数分层结构树图展现了行业股票指数间的聚集状态.



注:“制药”为“制药生物技术和科学”,“家庭”为“家庭和个人用品”.

图 9 T_1 、 T_2 和 T_3 时期分层结构树

通过图 9(a)、(b)、(c)对比发现:在 3 个时期中资本品、耐用品、原材料、技术硬件等生产制造业的指数距离均比较小,商业银行、保险、多样金融等

金融行业的指数距离均比较大,3 个时期分层结构树的变化更多地表现为股票指数距离齐升同降,并且指数间的相对距离变化较小,这表明 T_1 、 T_2 和 T_3

3 个时期的指数聚集状态基本相似,也说明中国行业板块联动网络基本稳定. A. Nobi 等^[10]在研究金融危机对韩国股票市场的影响中发现:在金融危机期间,大型制造业公司通过成为股票网络核心节点来稳定由金融危机给股票市场所带来的震荡. 中国股票网络的核心节点始终是生产制造业,这在一定程度上使得中国股票市场趋于稳定.

在 3 个时期中由商业银行和保险所组成的聚集层次与邻近层次的距离一直在 0.7 左右,这说明在行业联动网络中金融行业一直处于边缘地位. 这与前面动态网络中得出的结论一致. 值得注意的是: T_3 时期银行、保险和多样金融组成了一个聚类层次. 这说明 T_3 时期的股市高波动加强了金融行业指数之间的联动性.

4 结论与建议

本文基于标普中国 24 个行业股票指数 2007 年 1 月 10 日—2018 年 9 月 20 日日收盘价的数据,通过 DCC-MVGARCH 模型计算各行业股指间的每日动态相关系数,建立动态最小生成树网络,利用 Tucker 张量分解和 K -均值聚类方法,将动态网络聚集成 3 个代表网络,并在此基础上生成分层结构树图,研究了行业板块联动性的变化,得到如下结论:

1) 通过标普行业股指动态相关系数的描述性统计和动态网络的特征化发现:中国行业板块联动性长期处于波动状态,在股价快速上涨、下跌时,联动性的波动尤为剧烈. 具体表现为:股价快速上涨会使联动网络更稀疏,行业板块联动性减弱;而股价快速下跌会使联动网络更紧密,行业板块联动性增强,但在股价快速下跌结束后,联动网络会变得稀疏,行业板块联动性减弱.

2) 股市的剧烈波动会导致行业板块联动网络中心节点的变化,资本品和技术硬件是行业板块网络不同时期的核心节点,金融相关行业没有在行业板块联动网络中发挥重要作用. 这说明在中国股票市场的网络结构中生产制造业处于核心地位,金融业处于边缘地位.

3) 通过动态网络聚类发现:“2015 年股市波动”对行业板块联动性的影响比次贷危机更深刻;在代表网络中具有重要影响力的节点改变不大,并且指数分层结构树图聚集状态变化较小. 这说明行业板块联动网络基本稳定.

根据上述结论,研究建议:

1) 根据股指价格和行业板块联动性的变化,投资者和监管者可以在一定程度上预测股市的系统性风险. 如当股指价格快速下降时,行业板块联动性迅速上升,这是股市出现系统性风险的征兆,投资者可以及时出仓以避免可能出现的损失,监管者可以采取相应的措施来防范系统性风险.

2) 根据相应时期最小生成树和指数分层结构树图,投资者在进行行业配置投资时所投资的行业类型之间的指数相对距离要保证在一定范围内,以避免由所选股票的行业关联性过强而导致整个资产组合风险上升;监管者在股价波动时期可以预测风险传递路径,做好相关准备,减少由风险传递所带来的损失.

5 参考文献

- [1] MANTEGNA R N. Hierarchical structure in financial markets [J]. The European Physical Journal B: Condensed Matter and Complex Systems, 1999, 11(1): 193-197.
- [2] BOGINSKI V, BUTENKO S, PARDALOS P M. Statistical analysis of financial networks [J]. Computational Statistics and Data Analysis, 2005, 48(2): 431-443.
- [3] BRIDA J G, RISSO W A. Multidimensional minimal spanning tree; the Dow Jones case [J]. Physica A: Statistical Mechanics and Its Applications, 2008, 387(21): 5205-5210.
- [4] ŠTEFAN L, VÝROST T, BAUMÖHL E. Stock market networks; the dynamic conditional correlation approach [J]. Physica A: Statistical Mechanics and Its Applications, 2012, 391(16): 4147-4158.
- [5] ISOGAI T. Dynamic correlation network analysis of financial asset returns with network clustering [J]. Applied Network Science, 2017, 2(1): 1-10.
- [6] SENSOY A, TABAK B M. Dynamic spanning trees in stock market networks; the case of Asia-Pacific [J]. Physica A: Statistical Mechanics and Its Applications, 2014, 414: 387-402.
- [7] WANG Gangjin, XIE Chi. Correlation structure and dynamics of international real estate securities markets: a network perspective [J]. Physica A: Statistical Mechanics and Its Applications, 2015, 424: 176-193.
- [8] KANTAR E, KESKIN M, DEVIREN B. Analysis of the effects of the global financial crisis on the Turkish economy, using hierarchical methods [J]. Physica A: Statistical Mechanics and Its Applications, 2012, 391(7): 2342-2352.
- [9] SHIROKIKH O, PASTUKHOV G, BOGINSKI V, et al.

- Computational study of the US stock market evolution; a rank correlation-based network model [J]. Computational Management Science, 2013, 10(2/3): 81-103.
- [10] NOBI A, MAENG S E, HA G G, et al. Structural changes in the minimal spanning tree and the hierarchical network in the Korean stock market around the global financial crisis [J]. Journal of the Korean Physical Society, 2015, 66(8): 1153-1159.
- [11] 李俊照, 郭坤, 姚宏亮, 等. 基于马尔可夫链时序回归模型的房地产板块指数预测 [J]. 系统工程理论与实践, 2014, 34(4): 817-825.
- [12] 宁建楠, 易文德. 金融危机对中国股市各行业板块间相依结构的影响 [J]. 系统工程, 2015, 33(11): 10-17.
- [13] 刘井建, 焦怀东, 南晓莉. 危机冲击背景下股票市场风险联动非线性 [J]. 系统工程, 2015, 33(12): 16-22.
- [14] 马丹, 刘丽萍, 陈坤. 关联效应还是传染效应 [J]. 统计研究, 2016, 33(2): 99-106.
- [15] 曾裕峰, 简志宏, 彭伟. 中国金融业不同板块间风险传导的非对称性研究: 基于非对称 MVMQ-CAViaR 模型的实证分析 [J]. 中国管理科学, 2017, 25(8): 58-67.
- [16] 庄新田, 张鼎, 苑莹, 等. 中国股市复杂网络中的分形特征 [J]. 系统工程理论与实践, 2015, 35(2): 273-282.
- [17] 李岸, 粟亚亚, 乔海曙. 中国股票市场国际联动性研究: 基于网络分析方法 [J]. 数量经济技术经济研究, 2016, 33(8): 114-128.
- [18] 王荣森, 陈阳. 股灾前后中小板上市公司网络结构对比研究: 基于最小生成树和聚类分析方法 [J]. 哈尔滨工业大学学报(社会科学版), 2016, 18(5): 133-140.
- [19] 黄飞雪, 赵昕, 侯铁珊. 基于最小生成树的上证 50 指数分层结构 [J]. 系统工程, 2009, 27(1): 71-76.
- [20] 谢邦昌, 游涛. 金融危机前后中信行业指数联动效应及其社团结构比较 [J]. 商业经济与管理, 2015(1): 80-87.
- [21] ENGLE R. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models [J]. Journal of Business and Economic Statistics, 2002, 20(3): 339-350.
- [22] TSE C K, LIU Jing, LAU F C M. A network perspective of the stock market [J]. Journal of Empirical Finance, 2010, 17(4): 659-667.
- [23] BRIDA J G, RISSO W A. Hierarchical structure of the German stock market [J]. Expert Systems with Applications, 2010, 37(5): 3846-3852.

The Analysis on Linkage Effect of Chinese Industry Sector Based on Dynamic Network Method

LI Shoufei¹, LI Han², ZHU Guanxu³, SHENG Jiliang⁴

(1. Jiangxi Branch Taikang Endowment Insurance Corporation Limited, Jiangxi Nanchang 330038, China;

2. Chenzhou Central Sub Branch, the People's Bank of China, Chenzhou Hunan 423000, China;

3. High School Department, No. 3 Middle School of Nanchang, Jiangxi Nanchang 330029, China;

4. School of Statistics, Jiangxi University of Finance and Economics, Jiangxi Nanchang 330013, China)

Abstract: The dynamic minimum spanning tree network is established through DCC-MVGARCH model and network analysis method, and the dynamic network is integrated into three representative networks by using Tucker decomposition and K-means clustering method and hierarchical tree diagram is generated. The empirical analysis of China's industry sector linkages is made during the period from 2007 to 2018, including the subprime mortgage crisis, the "2015 stock market fluctuations" and the trade war. The results show that the linkage of industry sector has been in a state of volatility for a long time, when the stock price rises and falls rapidly, the linkage fluctuation is particularly severe, but the nodes with important influence in the linkage network have not changed much, hierarchical tree graph aggregation state change less, the linkage network between the industry sectors is basically stable. In the network structure of China's stock market, manufacturing industry is at the core, and the financial industry is at the marginal position.

Key words: industry sector linkage; DCC-MVGARCH; dynamic mst network; Tucker decomposition

(责任编辑: 曾剑锋)