

# 基于复杂网络的时间序列双变量相关性波动研究\*

高湘昀<sup>1)2)</sup> 安海忠<sup>1)2)†</sup> 方伟<sup>2)</sup>

1) (中国地质大学(北京)资源环境管理实验室, 北京 100083)

2) (中国地质大学(北京)人文经管学院, 北京 100083)

(2012年1月13日收到; 2012年2月20日收到修改稿)

为了研究具有时间序列特征的双变量之间相关性的波动规律, 本文选取国际原油期货价格和中国大庆原油现货价格作为样本数据, 借鉴统计物理学的方法进行研究. 运用粗粒化方法建立了相关性波动模态, 并利用复杂网络理论和分析方法对双变量相关性波动模态的统计、变化规律及其演化机理三个问题进行了分析. 结果显示, 双变量相关性波动模态分布具有幂律性、群簇性和周期性, 相关性波动主要通过少数几种模态进行传递和演化. 这些研究成果不仅可以作为双变量间相关性波动研究的方法, 也为不同变量间相关性波动一般规律的研究提供了思路.

**关键词:** 复杂网络, 粗粒化, 相关性, 波动

**PACS:** 89.75.Fb, 89.30.aj

## 1 引言

相关性研究的历史已久, 对于相关性的分析是指两个或多个具备相关性的变量进行分析, 从而衡量两个变量之间的相关密切程度. 相关性基本上涵盖了所有的范围和领域, 然而它不能等同于因果性, 不同学科领域的相关性分析的定义也不尽相同, 但其核心思想是一致的, 即当两个变量之间存在联系时, 一个变量会随着另一个变量发生变化. 通过相关性分析挖掘出变量之间存在的联系, 进而为制定相应的决策提供依据.

在目前宏观经济计量分析中, 双变量相关性的分析比较流行的是线性误差修正模型 (ECM) 刻画了变量之间的线性调整机理<sup>[1]</sup>, 协整方法<sup>[2]</sup>和阈值协整方法 (threshold cointegration), 刻画了经济变量之间的非线性调整机理<sup>[3]</sup>. 通过这些理论方法发现原油期货与现货价格之间的长期关系的偏差构成了均衡误差, 由此利用向量误差修正模型预测原油价格走势<sup>[4]</sup>; 对于相同或不同等级的原油期货与现货价格存在协整关系<sup>[5]</sup>; 利用多元阈值协整方法进行分析时发现当现货价格高于期货价格, 基础值

小于阈值时, 存在因果关系<sup>[6]</sup>. 通过这些研究, 其中一个事实是可以肯定的, 即在长期的原油市场中, 期货价格和现货价格之间具有协整、均衡及因果等相关性关系.

相关系数是衡量双变量之间相关程度的指标, 取值范围为  $[-1, 1]$ , 相关系数的绝对值越大, 表明双变量之间的相关程度越高. 相关性分为正相关和负相关, 相关程度也可分为强相关、弱相关和无相关三个等级.

在当今信息更新加快的时代, 任何一个变量都在时刻发生着变化, 而与其相关的变量同时也在发生着改变. 对于具有时间序列特征的两个变量, 在利用相关性分析时只能静态的停留在某个时间段内, 例如: 对国际原油期货 (WTI) 价格与中国大庆原油现货价格之间的相关性研究, 统计近 9 年的每日价格数据, 两者之间的相关系数高达 0.97, 具有强正相关性. 依据该结果进行决策将可能丢失很多内在的信息, 导致某个时间范围内的定价失误, 因为两者的相关系数并不是一直保持在高度相关的水平上, 而是呈现一种波动的形态, 甚至在某段时间内两者表现为负相关.

\* 国家自然科学基金 (批准号: 71173199), 教育部人文社会科学研究规划基金 (批准号: 10YJA630001) 和中央高校基本科研业务费专项资金 (批准号: 2011PY0215) 资助的课题.

† E-mail: ahz369@163.com

双变量之间相关性的波动是一个复杂的过程, 随着时间推移这种关系的波动构成了一个非线性、非稳定的复杂系统. 现有的传统经济学中的方法很难揭示关系的波动机理, 而在物理经济领域的研究中, 为我们提供了一些新的思路和方法. 首先是对关系波动状态的抽象, 我们可以采用粗粒化的方法. 粗粒化是相空间 (phase space) 中的概念, 相空间中的一个点是最细的细粒, 如果将一组点粗略的当成一个态, 使得对一系列细粒的研究转化为对态的研究 [7,8]. 利用粗粒化方法得到了符号序列, 进而判断地区间温度之间的动力学耦合关系 [9], 对非线性时间序列互信息与 Lempel-Ziv 复杂度的相关性进行了研究 [10]. 还将该方法应用于分子动力学模拟 [11], 石油化学分子力学 [12] 等方面. 通过这些研究发现粗粒化方法可以用来研究复杂性问题, 并且可以用于耦合性、相关性等关系的研究. 然后是对相关性波动复杂性的研究, 可以采用复杂网络分析方法. 复杂性科学的研究是近年来的研究热点, 被认为是 21 世纪的科学, 而复杂网络为复杂性科学的研究提供了一种理论方法, 尤其对于变量间关系演变规律的挖掘可以通过网络的方式进行研究. 其核心思想是将真实系统中各变量之间的联系看成一个复杂网络, 以网络的形式来描述真实复杂系统中各变量之间的关系, 通过对系统结构的分析, 更好的认识其本质特征. 对于复杂网络的研究经历了规则网络、随机网络和一般网络, 近年来随着各种变量间关系的复杂变化, 人们提出了一些更符合实际的网络模型, 小世界网络模型 [13]、NW 网络模型 [14]、BA 无标度网络模型 [15] 等. 并将复杂网络应用于生命科学 [16]、社会科学 [17] 及经济金融 [18] 等领域. 将粗粒化方法与复杂网络相结合对变量进行的研究有: 对温度波动模态的分析 [19]、国际原油价格波动模态的分析 [20], 以及价格间联动性波动模态的分析 [21] 等.

通过以上的研究综述发现, 原油的期货与现货价格之间存在相关性, 通过传统的经济学分析手段, 可以将双变量之间存在的某种关系表述出来, 但对于关系是如何波动的, 以及内在的变动规律和演化机理却无法描述. 要解决该问题可以借鉴统计物理学的思想, 将粗粒化方法和复杂网络理论应用于相关性波动的研究中来. 因此, 本文将借鉴粗粒化方法把双变量之间的相关性的波动关系抽象成为关系模态, 再利用复杂网络理论方法对相关性波动模

态之间的关系进行分析, 进而研究其内在复杂的变动规律和演化机理.

## 2 数据和方法

### 2.1 数据

相关性波动的研究应当具备两个变量, 要求这两个变量之间具有相关性, 且应当具有一定的代表性, 因此, 本文选取 2002 年 11 月 25 日至 2011 年 11 月 29 日 9 年的每日西德克萨斯中质原油期货 (WTI) 价格与中国大庆原油现货价格作为研究样本, 考虑双变量时间序列的一致性, 剔除由于周末或中外节假日的差异所产生的层次不齐的数据, 共获取 2149 组数据.

### 2.2 相关性波动模态的粗粒化处理

相关性波动模态的建立采用粗粒化的方法, 首先要利用相关性系数  $r_{xy}$  对双变量  $x, y$  之间的相关性程度进行量化. 相关性系数定义为

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (1)$$

其中,  $x_i$  为变量  $x$  的时间序列值,  $\bar{x}$  为变量  $x$  的平均值,  $y_i$  为变量  $y$  的时间序列值,  $\bar{y}$  为变量  $y$  的平均值,  $n$  为变量数列的项数.

通过对 2149 组数据相关性系数的计算发现, 双变量的相关性波动并不是平稳的, 而是非平稳, 非线性的波动过程, 如图 1 所示.

由 (1) 式也可以看出  $r_{xy}$  的取值范围为  $[-1, 1]$ , 因此根据的  $r_{xy}$  值界定双变量间的相关性程度, 并将相关性程度抽象为符号  $cs_i$  表示

$$cs_i = \begin{cases} P & (0.8 \leq r_{xy} \leq 1, \text{强正相关性}), \\ p & (0.3 \leq r_{xy} < 0.8, \text{弱正相关性}), \\ U & (-0.3 < r_{xy} < 0.3, \text{无相关性}), \\ n & (-0.8 < r_{xy} \leq -0.3, \text{弱负相关性}), \\ N & (-1 \leq r_{xy} \leq -0.8, \text{强负相关性}). \end{cases} \quad (2)$$

实验过程中, 我们选取以 10 天的变量数列组为单位, 步长为 1 天作“数据滑动窗 1”, 即每 10 天得出

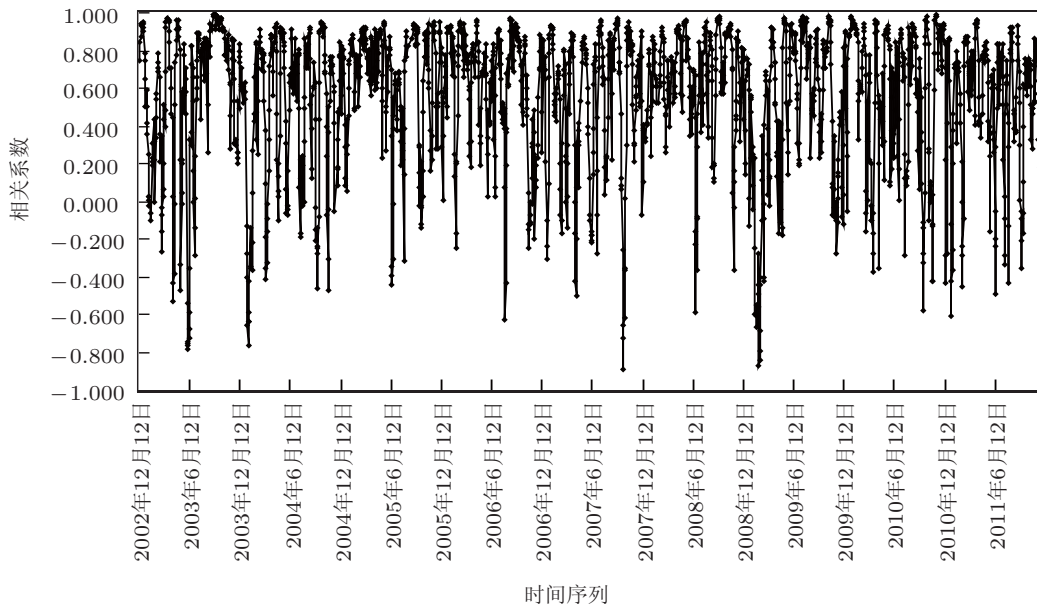


图1 双变量相关性波动序列

表1 相关性波动模态粗粒化处理过程

序号	变量 1	变量 2	相关性系数	相关性符号	粗粒化模态
1	30.55	30.8			
2	30.87	30.77			
3	29.91	30.65			
4	26.65	29.57			
5	26.26	28.73			
6	25.49	28.73			
7	25.24	28.23			
8	25.8	27.43			
9	26.03	28.1			
10	25.67	28.8	0.91		
11	26.49	28.23	0.88		
12	25.72	28.55	0.81		
13	26.23	27.6	0.22		
14	26.98	27.45	-0.47		
15	27.72	28.25	-0.33		
16	27.35	28.7	-0.01		
17	28.5	28.2	0.05	U	U <sub>nnUU</sub>
18	29.17	28.9	0.20	U	mUUU
19	28.74	28.53	0.22	U	nUUUU
20	29.14	29.27	0.58	p	UUUU <sub>p</sub>
...	...	...	...	...	...
2149	99.79	113.61	0.75	p	U <sub>pppp</sub>

数据滑动窗 1

数据滑动窗 2

一个相关性系数, 将得到  $2149 - 10 + 1$  个相关性系数, 由此将抽象出 2140 个符号, 并得到一个相关性符号序列  $cs_i$ , 与联动性波动研究不同的是, 联动性波动只获取了一天的数据建立波动状态 [20], 这样

可能导致变量的时滞性, 而采用相关性的滑动窗便消除了这种时滞效应,

$$CS_i = (cs_1, cs_2, cs_3, \dots, cs_n) \quad (n = 2140). \quad (3)$$

为了能更好的分析双变量间相关性的波动情况,粗粒化处理的符号组合不易过多,本文选取以5个符号为单位,步长为1作“数据滑动窗2”,即每5个相关性符号组合成为一个相关性波动模态,由此将得到  $2140 - 5 + 1$  个模态.通过数据滑动窗的形式进行粗粒化处理,每一个“数据滑动窗2”形成的模态是以前多个模态为基础的,这样使得相关性波动模态具有一定的记忆性与传递性,同时也不乏多元化的特征,粗粒化过程如表1所示.

### 2.3 相关性波动复杂网络分析方法

通过相关性波动模态的粗粒化处理,将双变量间的相关性波动关系分割为一个一个的模态来表示,这些波动模态互相转换,由此形成了一个相关性波动的复杂系统,而复杂网络分析方法主要揭示的是复杂系统中各要素之间的关系,因此首先应当构建相关性波动的复杂网络模型.

复杂网络包含两个基本要素:网络节点和边.本文将相关性波动模态作为网络的节点,模态间的相互转换关系作为网络的边.如此将对具有时间序列特征的双变量相关性波动的研究转换为了对该复杂网络的研究.双变量相关性复杂网络是一个有向加权的复杂网络,网络中边的方向为模态的转换过程,模态间转换的次数作为网络边的权重.

复杂网络的分析方法提供了很多分析参数,点强度、加权集聚系数、平均最短路径、结构洞等级度、介数集中系数等.通过参数分析来解决双变量间相关性波动的统计、变化规律及其演化机理三个问题.

1) 相关性波动模态统计规律的研究,分析出主要的波动模态是什么.这里的统计不再是传统意义上频次的统计,而是基于模态间转换关系的统计.本文采用点强度来计算,点强度不仅考虑了网络节点的近邻数,又考虑了与近邻节点间的权重,是该节点局域信息的综合体现<sup>[22]</sup>.如与节点*i*连接的节点个数为*n*,网络中所有的节点个数为*N*,节点*i*到节点*j*的权重为*w<sub>ij</sub>*,则节点*i*的点强度*s<sub>i</sub>*定义为

$$s_i = \frac{\sum_{j=1}^n w_{ij}}{\sum_{i,j=1}^N w_{ij}} \quad (4)$$

由(4)式中可知,节点的点强度越大,说明该模

态(节点)与其他模态之间转换的次数越多,在相关性波动系统中该波动状态就越重要.

2) 相关性波动模态变化规律的研究,分析哪些模态间进行了频繁的相互转换,主导的转换形式是什么及转换核心是什么模态.模态间的相互转换将形成一个小型的关系网络,即网络子群.因此对于网络子群的发现,可以采用建立在可达性和直径基础上子群发现的*n*-cliques方法或建立在节点度基础上子群发现的*k*-plex方法.

*n*-cliques方法即要求子群内任何两个节点之间在整个网络中的距离最大不超过*n*,令*d(i, j)*代表节点*i*和*j*的距离,则*d(i, j) ≤ n*,由此将满足该条件的所有节点集合找出,每个节点集合即为一个子群<sup>[23]</sup>.可以看出,*n*的取值越小,子群中节点之间的关系越紧密,模态间相互转换的规模也就越小.*k*-plex方法要求包含了*g*个节点的子群内每个节点都至少与同一子群中的其他节点保持*g-k*条关联,其中*k*为调节系数,当*k*值越小,*g*值越大时,条件越苛刻,模态间的转换形式也就越多<sup>[24]</sup>.

通过子群的发现,可以解决哪些模态间进行了相互转换及转换形式是什么的问题,但对于转换的核心模态的挖掘还需要借助网络的集聚系数来分析.节点的集聚系数反映了该节点的邻接点之间相互关联的程度,由于相关性波动网络为加权网络,因此采用加权集聚系数*C<sup>w</sup>(i)*进行计算,定义为<sup>[25]</sup>

$$C^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,k} \frac{(w_{ij} + w_{ik})}{2} a_{ij} a_{jk} a_{ki}, \quad (5)$$

其中*s<sub>i</sub>*为节点*i*的点强度,*k<sub>i</sub>*为节点*i*的度,*w<sub>ij</sub>*,*w<sub>ik</sub>*分别表示两个节点(*i, j*),(*i, k*)连边的权重,*a<sub>ij</sub>**a<sub>jk</sub>**a<sub>ki</sub>*表示节点*i, j, k*是否构成三角形,值为1表示三个节点之间都有边相连,构成三角形,0表示三个节点之间没有构成三角形.由(5)式可知,节点的加权集聚系数越大,表明与该模态相关的其他模态间的转换越紧密,该模态在自群中越具有核心地位.

3) 相关性波动模态演化规律的研究,分析模态间是通过何种形态进行转换的,转换周期是多少.模态间的相互转换随着时间推移是一种演化过程,因此不同的模态间以及子群间的转换需要经历一些过程,本文将借助网络结构洞的等级度和平均最短路径进行分析.结构洞的存在使得处于中间位置的波动状态居于重要的联络位置,很大程度上控制着相关性波动的传递.通过对结构洞等级度的测量,

将每个节点的结构洞等级度算出, 等级度越高, 说明该模态在相关性波动的传递过程中越居于核心地位 [24]. 节点  $i$  在网络中的等级度计算公式为

$$H_i = \frac{\sum_j \left( \frac{C_{ij}}{C/N} \right) \ln \left( \frac{C_{ij}}{C/N} \right)}{N \ln(N)}, \quad (6)$$

其中  $N$  是节点  $i$  的个体网络中的节点数,  $C/N$  是每个节点的限制度均值. 节点的限制度  $C_{ij}$  表示为

$$C_{ij} = \left( p_{ij} + \sum_q p_{iq} p_{qj} \right)^2, \quad (7)$$

其中  $p_{iq}$  是节点  $i$  的全部关系, 与节点  $q$  的关系占总关系的比例.

相关性波动具有周期性的规律, 而波动模态间的转换也同样具有周期性, 因此我们将通过平均最短路径的计算, 最短路径是网络中任意两点之间经历的边数, 通过计算可知相关性波动模态间的转换天数. 由此来确定相关性波动模态转换周期的大小, 为相关性波动状态的预测提供参考依据.

### 3 分析与结果

#### 3.1 相关性波动模态的统计

通过对两个变量相关系数的计算发现, 两者的相关性系数高达 0.97, 整体呈现强正相关性. 然而通过对相关性波动模态的粗粒化处理, 我们得到了相关性符号序列, 对该符号序列进行传统的统计, 发现符号  $P$  整个序列中占 31.78%, 符号  $p$  在占 46.50%. 说明原油期货与现货价格的相关性主要为正相关性, 与事实相符, 但是更多的却是呈现为弱正相关性, 如图 2 所示.

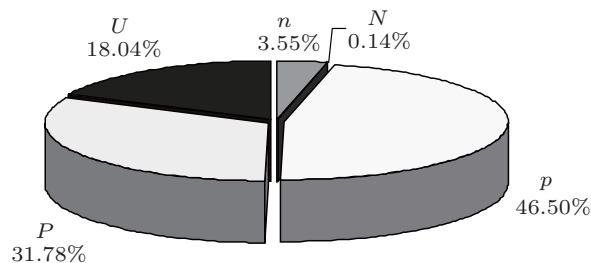


图 2 相关性符号传统统计占比

由于粗粒化处理, 理论上模态应该有  $5^5$  种类型, 但通过实验发现相关性波动系统中只出现了其中的 164 种模态, 说明双变量的相关性波动模态的

分布具有幂律的特征. 为了更进一步验证其幂律性, 我们对相关性波动复杂网络中节点强度的计算并排序, 发现  $ppppp$  和  $PPPPP$  两种模态的点强度为最高, 两种模态的点强度累计分布达到 30%, 对点强度  $S$  与排序序号  $R$  作双对数计算, 得到线性回归方程  $y = -1.4071x - 0.5714$ , 回归系数为 0.9629, 因此相关性波动模态的分布具有幂律性, 如图 3 所示.

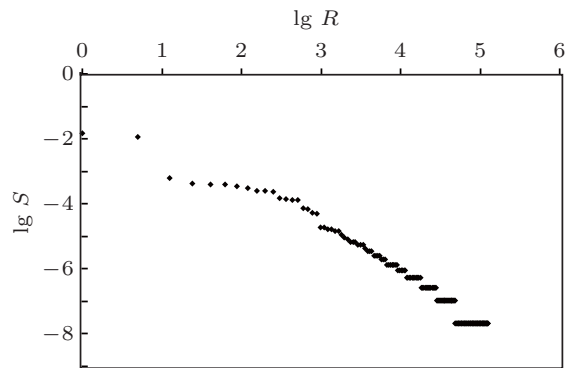


图 3 点强度与排序序号双对数关系

对相关性波动连续状态出现的频次进行统计, 发现强正相关性组成的字符串长度达 37, 出现频次随着长度的增加而降低. 综合考虑字符长度与出现频次, 发现连续 14—16 天的弱正相关性为主要的波动形态, 如图 4 所示.

通过以上的分析可知, 原油价格相关性波动的主要模态为连续多天的弱正相关性和强正相关性的形态, 其中连续 14—16 天的弱正相关性波动比较突出.

#### 3.2 相关性波动模态的变化规律

利用  $n$ -cliques 方法来发现相关性波动网络中子群, 由于该方法是建立在可达性和直径基础上的, 因此没有考虑网络边的权重. 当  $n$  选取 2, 节点规模不小于 8 时, 一共发现了 8 组子群, 由表 2 所示.

这些模态之间的相互转化都至多经过一个过程, 在整个相关性波动系统中, 这 8 组模态子群的相互转化比较快捷. 从表 2 中也可以分为 4 类, 1—3 组为弱正相关性波动模态类, 去掉重复的模态后只有 11 种; 第 4 组为弱负相关性波动模态类; 5, 6 组为强正相关性波动模态类, 去掉重复的模态后有 9 种; 7, 8 组为无相关性波动模态类, 去掉重复的模态后有 9 种.



### 3.3 相关性波动模式的演化机理

从图 1(双变量相关性波动序列)中随机的选择一段时间序列,发现相关性波动具有相似的波动形态,并且具有一定的周期性,如图 5 所示.

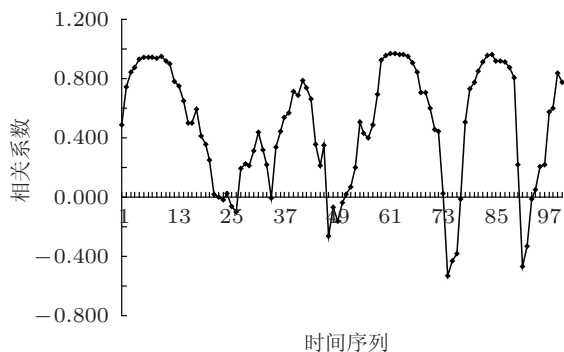


图 5 随机选取一段时间序列的相关性波动

从图 5 中可以看出,在波峰时期的点比较密集,波动过程中的点比较稀疏,说明相关性波动的演化是通过弱正相关性或无相关性进行传递的.波峰处的点密集而平缓,而波谷处的点稀疏且急促,说明相关性的波动总是处于一段较强的正相关性后随即发生较短时间内相关程度的下降.

对相关性波动的演化,我们通过网络结构洞的等级度对模态在相关性波动网络中起到中介传递的作用进行度量.通过实验发现  $ppPpp$ ,  $PPppp$ ,  $ppUUU$  和  $pppPp$  四种模态的等级度都在 0.50 以上,具有较高的等级度,说明这四种模态在相关性波动网络的拓扑结构中具有重要作用.尤其是  $ppPpp$  模态的等级度达到了 0.53,说明双变量在强、弱正相关性的转换过程中,该模态起到重要的中介传递作用.

通过对相关性网络中节点之间最短路径的计算可知,节点间的路径最长为 23,平均最短路径为 8.74,即模态间的转换周期最长为 23 天,平均为 8.74 天.由图 5 中也可以看出,一个完整的相关性波动周期经历了波峰、波谷及中间过程,通过对强正相关性波峰间时差的计算得知,相关性波动系统中,一个完整的波动周期为 18.55 天.

## 4 讨论和结论

双变量关系的研究在传统的经济学中已然有了很多的研究,通过计量经济学可以描述出双变量

之间存在的相关关系、协整关系、因果关系等.然而在实际中,这些关系并不是处于稳定的状态,而是一个非线性、非平稳的波动系统,传统经济学很难进行这些关系波动的深入研究.因此,为了更透彻的认识到双变量之间存在的关系,挖掘其内在的波动规律,物理经济学尤其是统计物理学为我们提供了新的思路.

本文选取原油期货与现货价格作为样本数据,利用粗粒化处理方法将双变量的相关性波动抽象成为一个个的模态,使得对相关性波动的研究转化为了对波动模态的研究.由于模态的数量和种类繁多,并且模态之间具有相互转化,传递等特征,构成了一个非线性、非平稳的复杂系统.因此借鉴复杂网络理论思想,构建出了双变量相关性波动复杂网络模型,通过对该网络模型的分析,解决了双变量间相关性波动的统计、变化规律及其演化机理三个问题.

1) 双变量相关性波动的模态主要表现为连续多天的弱正相关性和强正相关性的形态,其中连续 14—16 天的弱正相关性波动比较突出.  $ppppp$  和  $PPPPP$  两种模态尤为突出,两种模态基于关系的累计分布达到 30%,通过对点强度与排序序号作双对数计算发现相关性波动模态的分布具有幂律性.

2) 双变量相关性波动网络中发现弱正相关性、强正相关性和无相关性三类波动子群,且每类子群中都围绕一个核心模态进行波动.每类子群中波动模态的数量在 11 种以内,说明在双变量相关性波动系统中模态之间的相互转换并不是杂乱无章的,而是围绕着几种模态进行.弱正相关性波动子群趋向于由弱正相关性向强正相关性转换,强正相关性波动子群趋向于由强正相关性向弱正相关性转换,而无相关性波动子群中则是由弱正相关性向无相关性转换.

3) 相关性波动主要通过  $ppPpp$ ,  $PPppp$ ,  $ppUUU$  和  $pppPp$  等少数几种模态进行传递和演化,这些模态在相关性波动网络的拓扑结构中具有重要的中介作用.相关性的波动具有周期性,模态间的转换平均需要 8.74 天,而一个完整的波动周期平均需要 18.55 天.

本文针对的是具有时间序列特征双变量相关性波动的研究,因此要研究的双变量首先应当具有相关性.然而实际中,很多变量之间都存在相关性,

下一步将收集更多的数据, 研究不同类型的变量之间相关性的波动情况, 以探寻是否存在相关性波动

的一般规律. 同时, 多变量之间相关性波动的研究也将是今后要研究的内容.

- 
- [1] James E H, Davidson, David F H, Frank S, Yeo S 1978 *Econ. J.* **88** 661
- [2] Robert F E, Granger C W J 1987 *Econometrica* **55** 251
- [3] Nathan S B, Thomas B F 1997 *Int. Econ. Rev.* **38** 627
- [4] Coppola A 2008 *J. Futures Mark.* **28** 34
- [5] Svetlana M, Russell S 2009 *Energ Policy* **37** 1687
- [6] Huang B N, Yang C W, Hwang M J 2009 *Energy Econ.* **31** 91
- [7] Hao B L 1999 *Science* **51** 3
- [8] Wackerbauer R, Witt A, Atmanspacher H, Kurths J, Scheingraber H 1994 *Chaos Solitons Fractals* **4** 133
- [9] Feng G L, Hou W, Dong W J 2006 *Acta Phys. Sin.* **55** 962 (in Chinese) [封国林, 侯威, 董文杰 2006 物理学报 **55** 962]
- [10] Zhang D Z 2007 *Acta Phys. Sin.* **56** 3152 (in Chinese) [张佃中 2007 物理学报 **56** 3152]
- [11] Xu P J, Tang Y Y, Zhang J, Zhang Z B, Wang K, Shao Y, Shen H J, Mao Y C 2011 *Acta Phys. Chim. Sin.* **27** 1839 (in Chinese) [许佩军, 唐媛媛, 张静, 张知博, 王昆, 邵颖, 沈虎峻, 毛英臣 2011 物理化学学报 **27** 1839]
- [12] Zhang H Y, Wang Y Y, Tao G Q, Gui B, Yin C L, Chai Y M, Que G H 2011 *Acta Chim. Sin.* **69** 2053 (in Chinese) [张宏玉, 王艳艳, 陶国强, 桂彬, 殷长龙, 柴永明 2011 化学学报 **69** 2053]
- [13] Watts D J, Strogatz S H 1998 *Nature* **393** 6684
- [14] Newman M E J 1999 *Phys. Lett. A* **263** 341
- [15] Barabási A L, Albert R 1999 *Science* **286** 509
- [16] Janssen M A, Walker B H, Langridge J, Abel N 2000 *Ecol. Modell.* **131** 249
- [17] Zhang L, Liu Y 2008 *Acta Phys. Sin.* **57** 5419 (in Chinese) [张立, 刘云 2008 物理学报 **57** 5419]
- [18] Keles A, Kolcak M, Keles A 2008 *Knowl. Based Syst.* **21** 951
- [19] Zhou L, Gong Z Q, Zhi R, Feng G L 2008 *Acta Phys. Sin.* **57** 7380 (in Chinese) [周磊, 龚志强, 支蓉, 封国林 2008 物理学报 **57** 7380]
- [20] Chen W D, Xu H, Guo Q 2010 *Acta Phys. Sin.* **59** 4514 (in Chinese) [陈卫东, 徐华, 郭琦 2010 物理学报 **59** 4514]
- [21] Gao X Y, An H Z, Liu H H, Ding Y H 2011 *Acta Phys. Sin.* **60** 0689021 (in Chinese) [高湘响, 安海忠, 刘红红, 丁颖辉 2011 物理学报 **60** 0689021]
- [22] Yook S H, Jeong H, Barabási A L, Tu Y 2001 *Phys. Rev. Lett.* **86** 5835
- [23] Wasserman S, Faust K 1994 *Social network analysis: Methods and applications* (Cambridge: Cambridge University Press) p275
- [24] Ronald S B 1992 *Structural Holes: the Social Structure of Competition* (Cambridge: Harvard University Press) p66
- [25] Barrat A, Barthelemy M, Pastor S R, Vespignani A, 2004 *Proc. Natl. Acad. Sci. U.S.A* **101** 3747



# Research on fluctuation of bivariate correlation of time series based on complex networks theory\*

Gao Xiang-Yun<sup>1)2)</sup> An Hai-Zhong<sup>1)2)†</sup> Fang Wei<sup>2)</sup>

1) (*Lab of Resources and Environmental Management, China University of Geosciences, Beijing 100083, China*)

2) (*School of Humanities and Economic Management, China University of Geosciences, Beijing 100083, China*)

(Received 13 January 2012; revised manuscript received 20 February 2012)

## Abstract

In order to study the fluctuation of bivariate correlation which had time series characters, this paper selected International crude oil futures prices and Chinese Daqing crude oil spot prices as the sample data, using the method of statistical physics to study. The modes of fluctuation of correlation were defined by coarse graining process. Then three problems modes' statistics, law of variation and evolution mechanism were analyzed by complex network theory and analytical method. The results indicated that forms of modes showed that consecutive days of weak or strong positive correlation, and modes obeyed the power-law distribution. There were three kinds of sub-groups appearing in the network of fluctuation of bivariate correlation. These sub-groups were fluctuation of weak positive correlation, strong positive correlation and unrelated, and a core mode existed in each category of sub-groups. Transmission and evolution of fluctuation of bivariate correlation were a few modes. The fluctuation of bivariate correlation was transmitted and evolved by a few modes. The fluctuation had periodicity that the transmission among modes need average 8.74 days and a whole volatility cycle need about 18.55 days. These results not only can be the analyze method between two variables but also provides idea for researching a general law in different variables.

**Keywords:** complex networks, coarse-grained, correlation, fluctuations

**PACS:** 89.75.Fb, 89.30.aj

---

\* Project supported by the National Natural Science Foundation of China (Grant No. 71173199), the Humanities and Social Sciences Planning Funds Project under the Ministry of Education of the PRC (Grant No. 10YJA630001) and The Fundamental Research Funds for the Central Universities (Grant No.2011PY0215).

† E-mail: ahz369@163.com