

宏观交通量与 GDP 的复杂性测度及其相关分析

刘峰涛

LIU Feng-tao

东华大学 旭日工商管理学院, 上海 200051

Glorious Sun School of Business and Management, Donghua University, Shanghai 200051, China

LIU Feng-tao. Complexity measure and correlation analysis of macroscopic traffic volume & GDP. *Computer Engineering and Applications*, 2008, 44(16): 208-210.

Abstract: It is very important to make a quantificational description for the characters of systematic structure between transportation system and social economic system. The Lempel-Ziv algorithm of information theory is introduced to measure the complexity, one of the characters of systematic structure. Aiming at the limitation of this algorithm, the universal cut-and-try algorithm is put forward. The complexities of macroscopic traffic volume and GDP time series are calculated, respectively. The results show: the complexity of GDP has a positive correlation with the complexity of freight traffic volume; the influence of the complexity of GDP on the complexity of macroscopic traffic volume is long-term with lagged effect; and the correlation between GDP and macroscopic traffic volume has an uncertain relation with the correlation between their complexities.

Key words: complexity; Lempel-Ziv algorithm; universal cut-and-try algorithm; GDP

摘要: 交通运输系统与社会经济系统的系统结构特征需要定量描述。为测度系统的结构特征之一复杂度, 引入信息理论的改进的 Lempel-Ziv 算法——“通用试凑算法”。通过计算宏观交通量与 GDP 时间序列的复杂度, 得到简单结论: GDP 复杂度与货运交通量复杂度具有较强的正相关关系; GDP 复杂度对宏观交通量复杂度的影响是长期的, 具有滞后效应; GDP 与宏观交通量现象的相关程度与二者复杂度的相关程度没有必然的联系。

关键词: 复杂度; Lempel-Ziv 算法; 通用试凑算法; GDP

DOI: 10.3778/j.issn.1002-8331.2008.16.064 文章编号: 1002-8331(2008)16-0208-03 文献标识码: A 中图分类号: N941

交通运输系统是一个多层次、多组份的复杂系统, 度量其复杂性并分析影响其复杂性的因素, 是一个值得深入研究的问题。从宏观层次看, 交通运输系统是整个社会经济系统的一个子系统, 二者的复杂度之间是否也存在某种内在联系呢? 本文试图从复杂度的视角来关注交通运输系统与社会经济系统的系统结构, 进而研究二者的关系。

通常情况下, 对不同经济现象之间的关系的实证研究直接采用经济量数据进行分析, 其好处是能够比较全面地反映不同经济现象关系的信息。但此种做法的缺点是把很多系统结构信息隐藏在其中, 如果想深入挖掘, 则需要运用相应的新方法来实现。复杂度是系统结构的属性, 基于复杂度的视角, 就是要把系统的结构特征表面化, 进而从系统结构层面分析宏观交通运输系统与社会经济系统的关系。具体研究中, 选择宏观交通量作为宏观交通运输系统的指标, 而把社会经济系统交通运输的描述指标确定为 GDP。

1 交通运输系统复杂性刻画研究现状

复杂性科学是还原论和整体论的结合, 被称为 21 世纪的科学。从 1937 年贝塔朗菲(Bertalanffy)提出了一般系统论开

始, 1969 年普利高津(Prigogine)提出的耗散结构理论、哈肯(Haken)构建协同学、1979 年艾根(Eigen)的超循环论——这三种理论构成了自组织理论, 到 1984 年以三位诺贝尔奖获得者盖尔曼(Gell-Mann)、安德森(Anderson)、阿罗(Arrow)为首的一批科学家成立圣菲研究所(SFI), 国际上对复杂性的研究愈发深入。我国在钱学森为代表的多领域学者推动下, 也展开了从系统论到复杂巨系统的研究, 建立了“从定性到定量的综合集成研讨厅”这一复杂性研究体系。近年来, 国际上对复杂性的成因已形成了一个共识——非线性是导致事物复杂行为的唯一和基本的根源。所以, 对关于非线性动力学系统研究中的混沌理论、分形理论、孤立子和拟序结构理论、斑图动力学理论、混沌控制和混沌同步的理论和方法成为了研究和解决复杂性问题的主要理论基础和工具。

具体到交通运输系统的复杂性研究, 目前主要集中在三个领域的研究: (1)集中在交通流系统混沌方面的研究; (2)集中在交通运输系统自组织方面的研究; (3)集中在交通运输系统网络复杂性方面的研究^[1-9]。

这些研究的现状可以总结为表 1, 其难点包括:

(1)刻画理论缺少系统的基础理论和完整的刻画程序。不

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.50478088)。

作者简介: 刘峰涛(1972-), 男, 博士, 副教授, 研究方向: 复杂性与复杂系统。

收稿日期: 2007-11-28 修回日期: 2008-02-25

表1 交通运输系统复杂性刻画研究现状简表

集中领域	方法	进展	难点
交通流混沌	功率谱、相空间重构、Lyapunov 指数、关联维数、替代数据法等	1、离线判据比较成熟 2、实时判据没有解决 3、混沌预测与控制刚起步	1、数据需求大(离线判据) 2、混沌态下复杂程度不能定量比较 3、管控应用困难
自组织	元胞自动机和交通模型仿真; 自组织和耗散结构理论	1、自组织的临界判定 2、构建了一些交通自组织模型 3、管控应用停留在设想阶段	1、临界点内的系统复杂程度并不能实现刻画 2、定义偏于理论化 3、管控应用困难
网络复杂性	复杂网络理论、Lyapunov 指数、相空间重构、关联维数等	1、初步实现了网络复杂性判据 2、复杂网络理论刚被引入交通领域 3、管控应用停留在设想阶段	1、单纯的交通网络对交通运输系统的刻画能力有限 2、离真正的应用还存在较长的距离

但是交通运输系统,一般意义的系统复杂性刻画目前都是立足于具体方法的计算过程,缺少一套从哲学上、数学上关于复杂性刻画前提和假设的理论说明,在此理论基础上也需要建立普适而规范的刻画或者测度程序。

(2)刻画对象需要的数据量较大。前文已经论及,目前交通运输系统复杂性刻画的方法往往需要较长的时间序列数据才能得到稳定的结论,这影响了这些刻画方法的适用性。

(3)刻画方法不能实现可比化度量。目前交通运输系统复杂性刻画的方法只能对系统的复杂性进行判据,并不能够进行可比化度量,这也给交通运输系统的复杂性管控带来了不便。

(4)对交通运输系统的复杂性认识还停留在浅层次的阶段,如何针对系统的复杂特性进行管理和控制尚处空白。

2 Lempel-Ziv 算法及其改进

对于宏观交通量来说,其能采集到的时间序列长度一般不超过 100,前文中的算法都无法实现对这样短序列的复杂性刻画。为了实现对这些超短序列的测度,文献[10]提出了基于 Lempel-Ziv 算法的改进算法——通用试凑算法,其步骤为:

(1)根据文献[11]提出的方法计算 $c(S)$ 。

(2)运用试凑法确定 $c(S)$ 的上界 $b'(S)$,根据

$$n = n_k = \frac{\alpha}{\alpha-1} \left[\alpha^k \left(k - \frac{1}{\alpha-1} \right) + \frac{1}{\alpha-1} \right]^{[12]} \quad (1)$$

对于有限序列, n 已知,可通过试凑法确定 k 值。在试凑时,为了提高逼近速度,可以应用两步逼近法:首先以 1 为步长试凑整数位,然后再以小数步长进行较为精确的试凑。试凑得到 k 后,又根据

$$N_k = \frac{\alpha}{\alpha-1} (\alpha^k - 1)^{[12]} \quad c(S) \leq N_k + 1 = \frac{\alpha}{\alpha-1} (\alpha^k - 1) + 1^{[12]}$$

计算

$$b'(S) = \frac{\alpha}{\alpha-1} (\alpha^k - 1) + 1 \quad (2)$$

(3)根据公式(3)计算得到信息熵 h :

$$h = - \sum_{i=1}^{\alpha} p_i \log_{\alpha} p_i \quad (3)$$

(4)计算归一化复杂度值:

$$C_{LZ}'(S) = \frac{c(S)}{hb'(S)} \quad (4)$$

通用试凑算法的优点很明显:(1)可以用于短序列的复杂性测度;(2)改进了原 Lempel-Ziv 算法存在的一些问题。所以,本文选择此算法来定量刻画宏观交通运输系统和宏观经济系统的系统结构中的复杂程度。

3 复杂性测度

本文测度的对象是我国 1952~2004 年的客、货运交通量和 GDP^[3]。对于这些宏观数据,简单的套用 Lempel-Ziv 算法是存在问题的,因为宏观时间序列数据一般带有趋势项,系统真实的复杂程度信息会被隐藏在趋势项中,得不到表达,所以在需要首先对时间序列进行“平稳化重构”。对宏观交通量时间序列 $I' = \{x_1', x_2', \dots, x_n'\}$ 重构,重构的新序列是宏观交通量的变化率的绝对值。特殊的是,为了更好地反映宏观交通量变化的特性,对时间序列中的拐点(当前变化量与前两期变化量同时是反方向的时点)做特别的处理:当 $x_i' - x_{i-1}' < 0$, 且 $x_{i-1}' - x_{i-2}' > 0$, $x_{i-2}' - x_{i-3}' > 0$ 时,或者当 $x_i' - x_{i-1}' > 0$, 且 $x_{i-1}' - x_{i-2}' < 0$, $x_{i-2}' - x_{i-3}' < 0$ 时,则令 $x_i = |x_i' - x_{i-1}'| / |x_{i-1}' + x_{i-2}'|$, 其他情况则令 $x_i = |x_i' - x_{i-1}'| / |x_{i-1}'|$ 。重构后,得到新序列: $I = \{x_1, x_2, \dots, x_n\}$ 。然后再运用通用试凑算法对这 3 个时间序列进行复杂性测度,得到的结果如表 2。

表2 交通量与 GDP 的复杂度

年度	GDP	客运	货运	年度	GDP	客运	货运
		交通量	交通量			交通量	交通量
1952-1972	0.653 4	0.672 9	0.653 4	1952-1989	0.628 8	0.656 8	0.580 7
1952-1973	0.725 7	0.661 3	0.635 0	1952-1990	0.681 2	0.639 9	0.504 5
1952-1974	0.707 2	0.650 4	0.527 2	1952-1991	0.666 6	0.620 3	0.551 7
1952-1975	0.690 6	0.640 3	0.512 3	1952-1992	0.657 8	0.676 2	0.541 1
1952-1976	0.680 6	0.611 4	0.583 5	1952-1993	0.647 3	0.670 5	0.535 2
1952-1977	0.666 5	0.686 5	0.567 0	1952-1994	0.579 2	0.660 2	0.521 3
1952-1978	0.724 3	0.680 0	0.553 7	1952-1995	0.520 3	0.644 2	0.515 4
1952-1979	0.702 4	0.673 8	0.542 0	1952-1996	0.566 4	0.638 7	0.509 8
1952-1980	0.683 2	0.643 5	0.607 3	1952-1997	0.560 5	0.633 3	0.504 4
1952-1981	0.670 8	0.636 7	0.596 3	1952-1998	0.549 4	0.628 2	0.499 3
1952-1982	0.659 5	0.686 6	0.586 2	1952-1999	0.590 0	0.623 1	0.489 2
1952-1983	0.712 2	0.678 5	0.576 7	1952-2000	0.582 7	0.618 3	0.479 8
1952-1984	0.693 7	0.652 4	0.559 8	1952-2001	0.625 0	0.602 3	0.474 5
1952-1985	0.677 1	0.643 9	0.550 4	1952-2002	0.617 1	0.641 2	0.469 4
1952-1986	0.661 8	0.635 8	0.534 8	1952-2003	0.614 1	0.578 6	0.467 9
1952-1987	0.652 4	0.616 6	0.595 7	1952-2004	0.603 6	0.634 6	0.509 6
1952-1988	0.637 8	0.661 5	0.585 8				

4 相关性分析

4.1 GDP 与客运量

GDP 与客运量原序列如图 1, 其线性相关系数 $R=0.92$, 说明 GDP 与客运量存在很强的线性正相关。为了进一步挖掘二者系统结构的信息, 需要分析 GDP 的复杂度与客运量的复杂度之间的关系(如图 2)。二者即期复杂度的线性相关系数 $R=0.14$; 进一步, 客运量复杂度与滞后 1~5 期 GDP 复杂度的线性相关系数 $R=0.41$ 。说明 GDP 与客运量的复杂度相关性不强, GDP 复杂度对客运量复杂度的影响是长期的。

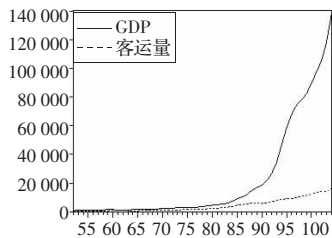


图1 GDP与客运量

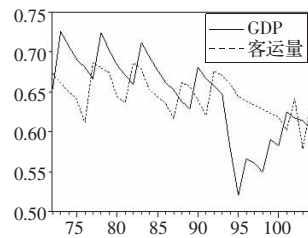


图2 GDP复杂度与客运量复杂度

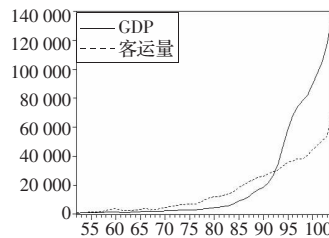


图3 GDP与货运量

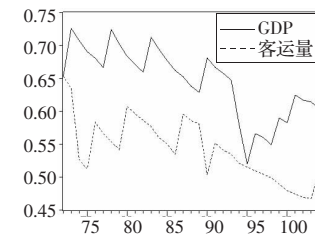


图4 GDP复杂度与货运量复杂度

4.2 GDP与货运量

GDP与货运量原序列如图3,其线性相关系数 $R=0.89$,说明GDP与货运量也存在很强的线性正相关。GDP的复杂度与货运量的复杂度如图4,二者即期复杂度的线性相关系数 $R=0.29$;进一步,客运量复杂度与滞后2~5期GDP复杂度的线性相关系数 $R=0.75$ 。说明GDP与货运量的复杂度还是存在很强的线性正相关关系的。同样,GDP复杂度对货运量复杂度的影响是长期的,具有滞后效应。

4.3 简单的结论

基于以上分析,对于我国的GDP和宏观交通量可以得出以下简单结论:

(1)GDP复杂度与货运交通量复杂度具有较强的正相关关系。说明货运交通运输子系统的结构变化趋势与宏观经济系统的系统结构的变化趋势具有一致性。

(2)GDP复杂度对宏观交通量复杂度的影响是长期的,具有滞后效应。说明宏观经济系统的系统结构变化对交通运输系统的结构变化具有长期的影响。

(3)GDP与宏观交通量现象的相关程度与二者复杂度的相关程度没有必然的联系。

参考文献:

- [1] Disbro J E,Frame M.Traffic flow theory and chaotic behavior[C]//Transportation Research Record 1225,TRB,National Research Council,Washington D C,1989:109-115.
- [2] Nagatani T.Stability analysis and stabilization strategies for linear supply chains[J].Physica A,2004,335:644-660.
- [3] Lan L W,Lin F Y,Kuo A Y.Identification for chaotic phenomena in short-term traffic flows;a parsimony procedure with surrogate

data[J].Journal of the Eastern Asia Society for Transportation Studies,2005,6:1518-1533.

- [4] Biham O,Middleton A,Levine D.Self-organization & a dynamic transition in traffic flow models[J].Phys Rev A,1992,46(10):6124-6127.
- [5] Muramatsu M,Nagatani T.Jamming transition of pedestrian traffic at a crossing with open boundaries[J].Physica A:Statistical Mechanics and its Applications,2000,286:377-390.
- [6] Herczeg L M.Co-evolutionary agent self-organization for city traffic congestion modeling[J].Lecture Notes in Computer Science,2004,3103:993-1004.
- [7] Thibault S,Marchard A.Reseaux et topologie[M].France:Institute National Des Sciences Appliquees de Lyon,Villeurbanne,1987.
- [8] Sen P,Dasgupta S,Chatterjee A,et al.Small-world properties of the Indian Railway network[J].Phys Rev E,2003,67.
- [9] Jiang B,Claramunt C.Topological analysis of urban street networks[J].Environment and Planning B,2004,31:151-162.
- [10] Liu F T,Tang Y.Improved Lempel-Ziv algorithm based on complexity measurement of short time series[C]//The 4rd International Conference on Fuzzy Systems and Knowledge Discovery,IEEE,2007.
- [11] Kasper F,Schuster H G.Easily calculable measure for the complexity of spatiotemporal patterns[J].Physical Review A,1987,36(2):842-848.
- [12] Lempel A,Ziv J.On the complexity of finite sequences[J].IEEE Transactions on Information Theory,1976,IT-22(1):75-81.
- [13] 中国经济信息网[EB/OL].[2005-07-01].http://www.cei.gov.cn/index/Transform.asp?cedb=7&ThreeBlockCode=030701&Template=dbjjnj-027&blockcode=DBjjnj_ys.

(上接204页)

误 DT 为538 min,受波及的航班总数 S 为7,航班延误指数 DM 为5.483。当 ID 取180 min时,统计量值分别为 $ID=180$ 、 $S=12$ 、 $DT=1107$ 、 $DM=7.15$ 。

容易得出如下结论:航班计划实施过程中,应特别关注一天中早期航班延误。换言之,应尽可能减少早期航班延误时间长度;同时在航班计划中,一天中早些时候的航班之间应尽可能多的给出松弛时间,从而能够尽量吸收延误,减少波及。

5 结论

本文研究了初始航班延误在相关航班资源影响下的延误波及情况,分析了飞机、机组航班计划与延误波及之间的关系。通过延误波及树给出了航班初始延误、影响范围和波及程度,计算了综合统计量值,为航空公司策略地使用松弛时间、适当调整航班计划、减少航班延误提供了切实可行的依据。

参考文献:

- [1] Shervin AhmadBeygi,Cohn A,Guan Yihan.Analysis of the potential for delay propagation in passenger aviation flight networks[R].2007 Sloan Industry Studies Working Paper Series,April 1,2007.
- [2] 马正平,催德光.机场航班延误优化模型[J].清华大学学报:自然科学版,2004,44(4):474-477.
- [3] Barnhart C,Cohn A,Johnson E,et al.Airline crew scheduling[M].Handbook of Transportation Science,Kluwer's International Series,2003:279-279.
- [4] Beatty R.preliminary evaluation of flight delay propagation through an airline schedule [C]//2nd USA/Europe Air Traffic Management R&D Seminar Orlando,USA,2000:1-9.
- [5] 民航空发(2003)96号 民航航班正常统计办法[S].中国民用航空总局,2003.