

一种识别关联维数无标度区间的新方法

周双 冯勇 吴文渊

A novel method to identify the scaling region of correlation dimension

Zhou Shuang Feng Yong Wu Wen-Yuan

引用信息 Citation: *Acta Physica Sinica*, 64, 130504 (2015) DOI: 10.7498/aps.64.130504

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.130504>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I13>

您可能感兴趣的其他文章

Articles you may be interested in

海杂波 FRFT 域的分形特征分析及小目标检测方法

Fractal property of sea clutter FRFT spectrum for small target detection

物理学报.2015, 64(11): 110502 <http://dx.doi.org/10.7498/aps.64.110502>

基于多重分形去趋势波动分析法的交通流多重分形无标度区间自动识别方法

Multi-fractal detrended fluctuation analysis algorithm based identification method of scale-less range for multi-fractal characteristics of traffic flow

物理学报.2014, 63(20): 200504 <http://dx.doi.org/10.7498/aps.63.200504>

基于小波 leaders 的海杂波时变奇异谱分布分析

Time-varying singularity spectrum distribution of sea clutter based on wavelet leaders

物理学报.2014, 63(15): 150503 <http://dx.doi.org/10.7498/aps.63.150503>

空间交替 Julia 集的反馈控制与线性广义同步

Feedback control and linear generalized synchronization of spatial-alternated Julia sets

物理学报.2014, 63(6): 060503 <http://dx.doi.org/10.7498/aps.63.060503>

非光滑热曲线的分数阶次可微性研究

Fractional differentiability of the non-smooth heat curve

物理学报.2012, 61(19): 190502 <http://dx.doi.org/10.7498/aps.61.190502>

一种识别关联维数无标度区间的新方法*

周双^{1)2)†} 冯勇¹⁾ 吴文渊¹⁾

1)(中国科学院重庆绿色智能技术研究院, 自动推理与认知重庆市重点实验室, 重庆 400714)

2)(中国科学院大学, 北京 100049)

(2014年12月10日收到; 2015年2月9日收到修改稿)

在计算关联维数过程中, 为了减少人为因素识别无标度区间带来的误差, 提出一种基于模拟退火遗传模糊C均值聚类识别无标度区间的新方法. 该方法根据无标度区间对应曲线的二阶导数在零附近波动的变化特征, 利用分类算法进行识别. 首先对双对数关联积分的离散数据进行二阶差分; 然后利用模拟退火遗传模糊C均值聚类方法对该数据进行分类, 选出在零附近波动的数据; 再剔除粗大误差保留有效数据; 最后进行统计分析识别出线性度最好的作为无标度区间. 应用新方法对两个著名的混沌系统Lorenz和Henon进行了仿真, 计算结果与理论值非常符合. 实验表明, 所提出的新方法与主观识别、K-means和2-means方法比较, 可以有效自动识别无标度区间, 减少误差, 计算结果更加精确.

关键词: 关联维数, 无标度区间, 分形, 模糊聚类**PACS:** 05.45.Df, 05.45.-a, 05.45.Pq, 05.45.Tp**DOI:** 10.7498/aps.64.130504

1 引言

分形理论是在20世纪70年代由Mandelbrot创立的, 用于定量描述非线性系统的运动的重要理论. 分形的特点是分维数, 简称分维. 目前, 常见的分维主要有: Hausdorff维数、信息维数、Lyapunov维数、关联维数等. 其中关联维数相对简单易于实现, 已经广泛应用在天体物理^[1]、故障诊断^[2]、信号处理^[3]、水文预测^[4]等等.

自然界的大多数分形只存在于无标度区间内, 一旦超越了这个范围, 分形就不复存在了. 如天空中云彩的边界只在1—1000 km的无标度区间内具有自相似性, 在此范围外, 就不再具有分形特征^[5]. 又如一些多孔材料也是只有在一定的标度范围内才具有自相似性和标度不变性^[6]. 上述例子中的标度就是测量的单位, 也称为尺度. 无标度性是指研究的客体与尺度没有关系, 即无论测量的单位如何

改变, 所研究的客体的性质, 如复杂程度、非规则性和形态等, 不随尺度的变化而改变. 因此确定分形客体的无标度区间是准确求得分形维数的重要保证, 进而在计算关联维数算法中, 无标度区间的选择决定着关联维数是否准确的重要因素. 为了找到无标度区间, 往往采用主观识别, 但结果不准确. 为了减少误差, 急需一种客观自动识别无标度区间的方法来得到更加准确的关联维数. 近年来, 很多学者对此进行了研究, 得到了很多成果. 如Yokoya等^[7], Maragos等^[8], Kim等^[9], Harikrishnan等^[10]的方法, 具有一定经验性, 缺乏理论依据; 如汪富泉等的三折线法^[11], S形曲线拟合法^[12]等, 这种方法基于特殊的曲线模型, 但关联积分曲线不能简单看作特殊的曲线; 如Judd算法^[13], 虽然这种方法不需要计算无标度区间, 但针对一些系统具有较好的估计; 如熊杰等^[14]的无标度区间识别只针对交通流的多重分形; 如党建武等^[15]提出基于置信度和相关性的分组递归确定无标度区间, 如

* 国家自然科学基金(批准号: 11301524)、重庆市科技攻关项目(批准号: cstc2012ggB40004)和中国科学院西部之光联合学者资助的课题.

† 通信作者. E-mail: zhoushuang@cigit.ac.cn

巫兆聪^[16]提出的基于标准偏差的自适应方法,但计算繁琐;如杜必强等^[17],王成栋等^[18]和吴虎胜等^[19]的结果基于分段拟合法,但对于大数据计算比较复杂;如K-means方法^[20],2-means方法^[21]等结果,这两种方法的核心过程和所使用的聚类算法可能导致陷入局部最优解.

为了解决这些问题,根据无标度区间对应曲线的二阶导数变化特征,本文提出利用新的聚类算法对双对数曲线点的二阶差分数据进行分类,再结合统计学理论进行无标度区间的识别.首先,判别实际系统是否是混沌系统,这样采用G-P算法计算出的混沌特征更有效^[22];第二步,构造相空间计算关联积分;然后,对双对数曲线点的二阶差分数据,利用模拟退火遗传模糊C均值聚类算法进行分类,识别出零附近波动数据,再剔除粗大误差,接下来利用统计方法识别出无标度区间;最后,对该区间对应的点进行最小二乘法拟合得到关联维数.通过对Lorenz和Henon混沌系统的仿真,验证了该算法的有效性.

2 方法

2.1 G-P 算法

计算关联维的方法是采用Grassberger和Procaccia在Whitney的“嵌入定理”和Packard的重建相空间理论上提出的G-P算法^[23,24].首先,对时间序列 $\{x(n)\}$,采用时间差法重构 m 维相空间中的相点,即 $X(i) = [x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)]$, $i = 1, 2, \dots, N$,其中 m 为嵌入维数; τ 为延迟时间; N 为重构后相空间的点数, $N = n - (m - 1)\tau$.然后统计出相空间中两点距离小于半径的个数,从而得到关联积分

$$C(m, r) = \frac{2}{N(N-1)} \times \sum_{i \neq j}^N H(r - \|X(i) - X(j)\|), \quad (1)$$

其中 $H(\cdot)$ 为Heaviside阶跳函数.当 $x < 0$ 时, $H(x) = 0$.当 $x > 0$ 时, $H(x) = 1$. $C(m, r)$ 是一个积累分布函数,它描写了相空间中的吸引子上两点之间距离小于 r 的概率,刻画了相对于相空间某参考点 $X(i)$ 在 r 内的相点聚焦的程度,而且关联维数与关联积分存在 $C(m, r) \propto r^{D(m)}$,当

关联维数随着嵌入维数的增加,最后达到一饱和值,此时的关联维数为吸引子的关联维数,即 $D_2 = \lim_{m \rightarrow \infty} \lim_{\tau \rightarrow 0} \frac{\ln C(m, r)}{\ln r}$.对于确定的系统,在双对数坐标中绘出 $\ln r - \ln C(m, r)$ 曲线,其中线性度最好的区域所对应的区间就是无标度区间(如图1所示),再利用最小二乘法对该区域的点拟合得到的直线斜率即为关联维数.

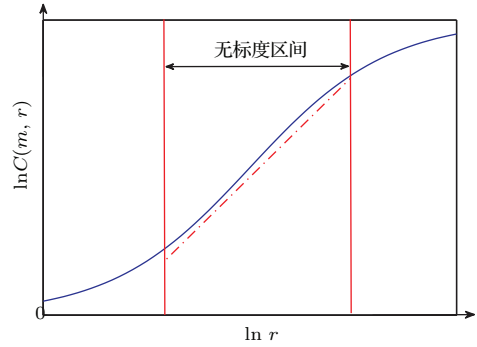


图1 无标度区间示意图

Fig. 1. Schematic diagram of scaling region.

2.2 模拟退火遗传模糊C均值聚类算法

模拟退火算法是Kirkpatrick等在1983年提出一种模拟热力学中物理退火过程的随机优化算法,不论在组合优化问题,还是在连续空间优化问题上都有广泛的应用^[25].模拟退火算法的操作过程如下:

- 1) 初始温度 t (充分大),随机生成初始解 S ,最大迭代次数 L ;
- 2) 以目前解为中心随机产生新的邻近解 S' ;
- 3) 计算增量 $\Delta C = C(S') - C(S)$,其中 $C(S)$ 为评价函数;
- 4) 若 $\Delta C < 0$ 则接受 S' 作为新的当前解,否则以概率 $e^{(-\Delta C/Kt)}$ (其中 K 为物理学中的常数)接受 S' 作为新的当前解;
- 5) 如果满足终止条件则输出当前解作为最优解,结束算法,否则降低温度,继续迭代2)–4).

遗传算法是由Holland在1975年提出的一种模拟生物进化过程的概率搜索算法^[26],其基本运算过程如下:

- 1) 初始化:随机生成 M 个个体作为初始群体;
- 2) 个体评价:计算群体中各个个体的适应度;
- 3) 选择运算:从种群中以一定概率选择适应度好的个体组成新种群;

- 4) 交叉运算: 从3)中得到的新种群中选择两个个体,按一定概率交叉得到新体;
- 5) 变异运算: 从4)中得到的新种群中随机选择一个个体,按一定概率变异到新个体;
- 6) 终止条件判断: 若满足条件,则以进化过程中所得到的具有最大适应度个体作为最优解输出,终止计算.

模糊C均值聚类(FCM)是基于距离的聚类算法,用模糊理论对数据进行分析和建模的方法,建立了数据类属的不确定性描述,能比较客观地反映现实世界,作为K-means的改进算法^[27],具有很高的搜索速度,但是对聚类中心的初值十分敏感,如果初值选择不当,它会收敛到局部极小点.其基本思想是通过优化目标函数 $J = \sum_{j=1}^c \sum_{i=1}^n [u_j(x_i)]^b \|x_i - C_j\|^2$ (其中 $u_j(x_i)$ 是第*i*个数据点 x_i 对于*j*类的隶属度函数,且 $\sum_{j=1}^c u_j(x_i) = 1$; c 为设定的类别数目; $C_j, j = 1, 2, \dots, c$ 为每个聚类的中心; n 为数据点的个数; b 为模糊程度常数)得到每个数据点对所有类别中心的隶属度,从而决定样本点的类属以达到分类的目的.

由于遗传算法(GA)把握总体的能力较强,但局部搜索能力差;而模拟退火算法(SA)具有较强的局部搜索能力.因此可以将遗传算法和模拟退火算法相互结合取长补短用于FCM算法中,这样使该聚类算法更有效、更快速地收敛到全局最优解^[28].模拟退火遗传模糊C均值聚类(SAGAFCM)算法具体详细步骤如下(其流程图如图2所示):

- 1) 初始化参数: 种群个体大小,最大进化次数 MAXGEN,交叉概率,变异概率,退火初始温度 T ,温度冷却系数 k ,终止温度 T_{end} ;
- 2) 随机初始化 c 个聚类中心,并生成初始种群 Chrom,对每个聚类中心计算个样本的隶属度,以及每个个体的适应值 f_i ;
- 3) 设置循环计数变量 $gen = 0$;
- 4) 对群体 Chrom 实施选择、交叉和变异等遗传操作,对新产生的个体利用公式重新计算各个聚类中心、各样本的隶属度,以及每一个体的适应度值 f'_i ;若 $f'_i > f_i$,则以新个体替换旧个体,否则,以概率 $P = \exp((f_i - f'_i)T)$ 接受新个体,舍弃旧个体;

- 5) 若 $gen < MAXGEN$,则 $gen = gen + 1$,转至步骤4),否则,转至步骤6);
- 6) 若 $T_i < T_{end}$,则算法成功结束,返回全局最优解;否则降温操作 $T_{i+1} = kT_i$,转至步骤3).

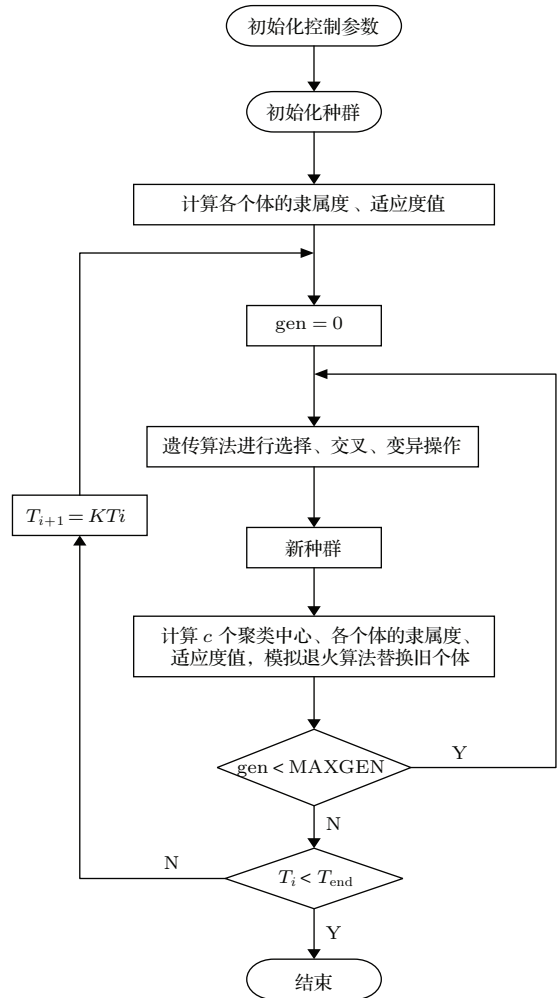


图2 模拟退火遗传模糊C均值聚类算法流程图
Fig. 2. Flow chart of the simulated annealing genetic fuzzy C-means clustering algorithm.

2.3 无标度区间识别新方法

无标度区间对应的曲线是一段相对波动较小的直线段,直线段的一阶导数是常数,因而无标度区间对应点的一阶导数会在常数附近微小波动,其余点会波动较大.但如果存在多个直线段,并不在同一条水平直线上,满足条件的点会在不同的常数附近波动,从图3中可知很难进行分类识别直线段(无标度区间对应的直线段).因此,考虑无标度区间对应曲线的二阶导数等于零或近似为零,这样满足条件的点都会为零附近做微小波动,而其他点会在零以上或零以下做较大的正波动和负波动,进而

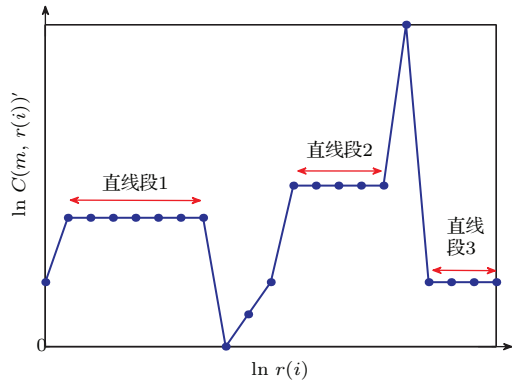


图3 双对数关联积分的一阶差分曲线示意图
Fig. 3. The first order differential of the double logarithm correlation integral curve.

可以分出两类(即零附近波动和正波动或零附近波动和负波动),或者三类(即零附近波动、正波动和负波动).具体分类数目根据曲线二阶差分图变化特征而定(注:由于离散数据不能求导数,因此用一阶差分近似一阶导数,二阶差分近似二阶导数).不妨假设数据分为三类,识别无标度区具体步骤如下:

- 1) 判别实际系统是否是混沌系统;
- 2) 对时间序列 $\{x(i), i = 1, 2, \dots, n\}$ 进行处理,初始嵌入维数 $m \in Z^+(Z^+ \text{ 为正整数})$,利用互信息计算出延迟时间 τ ,采用时间差法重构相空间 $X(i) = [x(i), x(i+\tau), x(i+2\tau), \dots, x(i+(m-1)\tau)]$, $i = 1, 2, \dots, n - (m - 1)\tau$;
- 3) 利用G-P算法对相空间中的点计算得到双对数点集 $\{(\ln r(i), \ln C(m, r(i))), i = 1, 2, \dots, N'\}$,并对其进行二阶差分得到集合 $S_1 = \{\ln C(m, r(i))'', i = 1, 2, \dots, M\}$ (这里 $\ln C(m, r(i))''$ 为 $\ln C(m, r(i))$ 的二阶差分符号),其中 $r(i) \in [r_{\min}, r_{\max}]$, r_{\min} 为相空间中两点最小距离, r_{\max} 为相空间中两点最大距离;
- 4) 将数据 S_1 利用模拟退火遗传模糊C均值聚

类算法分为正波动、负波动及零波动三类数据,由于无标度区间对应曲线的二阶导数会在零附近波动,因此保留零波动数据 S_2 ,见图4;

5) 在分类过程中,可能会把异常点分类到零波动数据中,因此根据统计学中的 3σ 准则[29],设定误差范围,剔除 S_2 中粗大误差点,保留有效数据,记为 S_3 ;

6) 在图4中,可以看出存在多个直线段,也就是对应多个可能的无标度区间,要从中选择出对应的数据线性度最好的,因此对 S_3 中指标为连续自然数的区间(图4所示的直线段对应的区间,通常这里需要区间内所含点的个数大于等于3)对应的双对数曲线点集分别进行线性回归分析,如果通过相关系数显著性检验,选择波动最平缓(即方差最小)的数据,进行最小二乘法线性拟合,计算出拟合直线的斜率,此斜率为关联维数,否则,转到步骤4),此时令 $S_1 = S_3$.

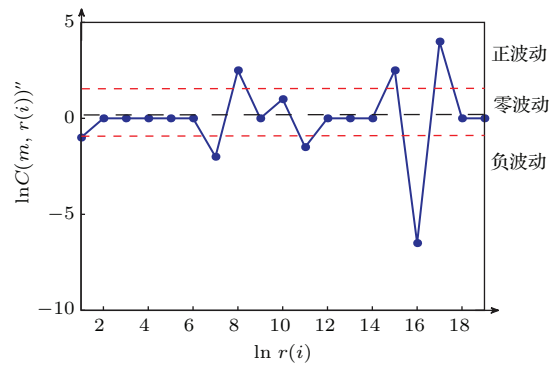


图4 双对数关联积分的二阶差分曲线示意图
Fig. 4. The second order differential of the double logarithm correlation integral curve.

2.4 实验仿真

Lorenz系统是1963年洛伦兹在讨论天气预报和大气湍流现象时提出的第一个确定方程得出混

表1 不同算法比较
Table 1. Comparison of different algorithms

模型	参数	主观识别法	K-means方法	2-means方法	新算法	理论值
Lorenz	$\begin{cases} dx/dt = 10(y - x) \\ dy/dt = 28x - xz - y \\ dz/dt = xy - (8/3)z \end{cases}$ $[x_0, y_0, z_0] = [12, 2, 9]$ $N = 10000$ 嵌入维数 $m = 12$	2.100 ± 0.061	2.381	2.064	2.052	2.05 ± 0.01
Henon	$\begin{cases} x_{n+1} = y_n + 1 - 1.4x_n^2 \\ y_{n+1} = 0.3x_n \end{cases}$ 随机生成初始值 $[x_0, y_0]$ $N = 10000$ 嵌入维数 $m = 12$	1.334 ± 0.056	1.407	1.240	1.269	1.25 ± 0.02

沌解的实例. Henon 系统是 1976 年法国天文学家伊依从研究球状星团时提出一种最简单的混沌吸引子系统, 并用它建立的理论成功解释了太阳系的稳定性问题. 为了验证算法的有效性和准确性, 通过对这两个系统进行仿真, 并与文献 [21] 中的方法进行比较, 结果如表 1 所示. 从此表中可以看出新方法是有效可行的, 而且比其他方法计算更准确.

3 结 论

1) 通过对两个经典混沌系统的仿真, 新算法计算出的关联维数与理论值符合度高, 验证了该方法的有效性; 2) 本文根据无标度区间对应的曲线的二阶导数变化特征, 利用模糊 C 均值算法的快速收敛性, 又利用了模拟退火和遗传算法的全局搜索性, 不至于陷入局部最优解; 3) 相对于主观识别, K -means 和 2-means 方法更能有效地用于无标度区间识别; 4) 新方法涉及高等数学、模式识别理论和统计学知识, 因此无标度区间的确定是有理论依据的, 而且计算过程易于计算机实现, 减少了人为干预, 能够实现无标度区间的自动识别, 这也为其他分维的无标度区间识别提供了参考.

参考文献

- [1] Ostryakov V M, Usoskin I G 1990 *Solar Physics* **127** 405
- [2] Wang W J, Wu Z T 2000 *Journal of Shanghai Jiaotong* **34** 1265 (in Chinese) [汪慰军, 吴昭同 2000 上海交通大学学报 **34** 1265]
- [3] Casaleggio A, Corana A 2000 *Chaos, Solitons & Fractals* **11** 2017
- [4] Huang G R, Rui X F 2004 *Advances in Water Science* **15** 255 (in Chinese) [黄国如, 芮孝芳 2004 水科学进展 **15** 255]
- [5] Huang R S, Huang H 2007 *Chaos and Its Applications* (Second Edition) (Wuhan: Wuhan University Press) p217 (in Chinese) [黄润生, 黄浩 2007 混沌及其应用 (第二版) (武汉: 武汉大学出版社) 第 217 页]
- [6] Sheng Y G, Xu Y, Li Z S, Wu D, Sun Y H, Wu Z H 2005 *Acta Phys. Sin.* **54** 221 (in Chinese) [盛永刚, 徐耀, 李志宏, 吴东, 孙予罕, 吴中华 2005 物理学报 **54** 221]
- [7] Yokoya N, Yamamoto K, Funakubo N 1989 *Computer Vision, Graphics, and Image Processing* **46** 284
- [8] Maragos P, Sun F K 1993 *IEEE Transactions on signal Processing* **41** 108
- [9] Kim H S, Eykholt R, Salas J D 1999 *Physica D* **127** 48
- [10] Harikrishnan K P, Misra R, Ambika G, Kembhavi A K 2006 *Physica D* **215** 137
- [11] Wang F Q, Luo C S, Cheng G X 1993 *Chinese Journal of Computational Physics* **10** 345 (in Chinese) [汪富泉, 罗朝盛, 陈国先 1993 计算物理 **10** 345]
- [12] Bolea J, Laguna P, Remartínez J M, Rovira E, Navarro A, Bailón R 2014 *Computational and Mathematical Methods in Medicine* **2014** 1
- [13] Judd K 1994 *Physica D* **71** 421
- [14] Xiong J, Chen S K, Wei W, Liu S, Guan W 2014 *Acta Phys. Sin.* **63** 200504 (in Chinese) [熊杰, 陈绍宽, 王伟, 刘爽, 关伟 2014 物理学报 **63** 200504]
- [15] Dang J W, Shi Y, Huang J G 2003 *Computer Engineering and Applications* **23** 35 (in Chinese) [党建武, 施怡, 黄建国 2003 计算机工程与应用 **23** 35]
- [16] Wu Z C 2002 *Acta Geodaetica et Cartographica Sinica* **31** 240 (in Chinese) [巫兆聪 2002 测绘学报 **31** 240]
- [17] Du B Q, Jia Z W, Tang G J 2013 *Journal of Vibration and Shock* **32** 40 (in Chinese) [杜必强, 贾子文, 唐贵基 2013 震动与冲击 **32** 40]
- [18] Wang C D, Ling D, Miao Q 2012 *Computer Engineering and Applications* **48** 9 (in Chinese) [王成栋, 凌丹, 苗强 2012 计算机工程与应用 **48** 9]
- [19] Wu H S, Ni L P, Zhang F M, Zhou X, Du J Y 2014 *Control and Decision* **29** 455 (in Chinese) [吴虎胜, 倪丽萍, 张凤鸣, 周漩, 杜继勇 2014 控制与决策 **29** 455]
- [20] Yang H Y, Ye H, Wang G Z, Pan G D 2008 *3rd IEEE Conference on Industrial Electronics and Applications* Singapore, Jun 3–5, 2008 p1383
- [21] Ji C C, Zhu H, Jiang W 2010 *Chinese Science Bulletin* **31** 3069 (in Chinese) [姬翠翠, 朱华, 江炜 2010 科学通报 **31** 3069]
- [22] Wang A L, Yang C X 2002 *Acta Phys. Sin.* **51** 2719 (in Chinese) [王安良, 杨春信 2002 物理学报 **51** 2719]
- [23] Grassberger P, Procaccia I 1983 *Physical Review Letters* **50** 346
- [24] Grassberger P, Procaccia I 1983 *Physica D* **9** 189
- [25] Kirkpatrick S, Gelatt J C D, Vecchi M P 1983 *Science* **220** 671
- [26] Holland J H 1975 *Adaptation in Natural and Artificial Systems* (Ann Arbor: The University of Michigan Press)
- [27] Jain A K 2010 *Pattern Recognition Letters* **31** 651
- [28] Bai L Y, Hu S Y, Liu S H 2005 *Computer Engineering and Applications* **41** 56 (in Chinese) [白莉媛, 胡声艳, 刘素华 2005 计算机工程与应用 **41** 56]
- [29] Liang J W, Chen L C, He G 2001 *Error Theory and Data Processing* (Revised Edition) (Beijing: China Metrology Press) p57 (in Chinese) [梁晋文, 陈林才, 何贡 2001 误差理论与数据处理 (修订版) (北京: 中国计量出版社) 第 57 页]

A novel method to identify the scaling region of correlation dimension*

Zhou Shuang^{1)2)†} Feng Yong¹⁾ Wu Wen-Yuan¹⁾

1) (Chongqing Key Laboratory of Automated Reasoning and Cognition, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China)

2) (University of Chinese Academy of Sciences, Beijing 100049, China)

(Received 10 December 2014; revised manuscript received 9 February 2015)

Abstract

A random fractal exhibits self-similarity over the scaling region, this is different from the regular fractal. The scaling region obtained by the proper method for the exact fractal dimension is very important. And the correlation dimension is one of the fractal dimensions which is used widely in many fields. Therefore, it is necessary and timely to identify the scaling region that plays a critical role in calculating the correlation dimension accurately in various chaotic systems. Visual identification is widely used to determine the scaling region as a quick and simple subjective method. However, this method may lead to an inaccurate result in Grassberger Procaccia algorithm. In order to reduce the error caused by human factors from computing the correlation dimension, a novel method of identifying the scaling region based on simulated annealing genetic fuzzy C-means clustering algorithm is proposed. This new method is based on the fluctuating characteristics that the second-order derivative of the curve within the scaling region is zero or nearly zero. Firstly, the second-order differential of the double logarithm correlation integral discrete data is calculated. Secondly, the simulated annealing genetic fuzzy C-means clustering method is used for dividing the data into three groups: positive fluctuation data, zero fluctuation data, and negative fluctuation data. The zero fluctuation data are selected to retain, the rest is excluded. Thirdly, the 3σ criteria are used for excluding gross errors to retain those valid from the zero fluctuation data. Fourthly, the data of the consecutive nature point interval are chosen from the retained data. Finally, the regression analysis and statistical test are used to identify the scaling region from the data valid. In order to verify the effectiveness of the proposed method, it is used to simulate the Lorenz and Henon systems. The calculated results are in good agreement with the theoretical values. Experimental results show that the proposed new approach is easy to operate, more efficient and more accurate than the subjective recognition, K -means method, and 2-means method in identifying the scaling region.

Keywords: correlation dimension, scaling region, fractal, fuzzy clustering

PACS: 05.45.Df, 05.45.-a, 05.45.Pq, 05.45.Tp

DOI: [10.7498/aps.64.130504](https://doi.org/10.7498/aps.64.130504)

* Project supported by the National Natural Science Foundation of China (Grant No. 11301524), the Chongqing Science and Technology Key Project, China (Grant No. cstc2012ggB40004), and the CAS western light program.

† Corresponding author. E-mail: zhoushuang@cigit.ac.cn