Chaos analysis of the time series data derived from human activities —classification based on the geometrical structure of an attractor— 人間活動に由来する時系列データのカオス解析 —アトラクタの幾何構造による分類—

Hiroyuki NISHIKAWA, Manabu KURAMOTO, Shigeki OKINO and Fujio SUDA

Department of Resources and Environment Science, Tokai University, 4-1-1 Kitakaname, Hiratsuka, Kanagawa, 259-1292 Japan

TEL: x81-463-58-1211(EX.3440) FAX: x81-463-50-2097 e-mail: nishikawa@vi.cs.is.nagoya-u.ac.jp

(Received 2 October, 2013 Accepted 23 May, 2015)

Various chaos analyses have been applied for the time series of several human activities. As a result of recurrence plot and power spectrum analyses, three rough classes of the transactions between companies, those between companies and individuals, and those between individuals can be refined according to the geometrical structure of an attractor into three more accurate categories: non-stationary, periodic and stochastic changes. In a stochastic one, the slope of the approximate straight line of the log-log graph of power spectrum density: $\beta \geq 1.26$, and the recurrence plot of its attractor shows non-contiguous diagonal lines; in a periodic one, $0.770 \leq \beta \leq 1.25$, equally-spaced diagonals; and in a non-stationary one, $0.108 \leq \beta \leq 0.389$, ambiguous boundaries of domain.

Keywords : Human activities, Power spectrum, Recurrence plots

I. INTRODUCTION

個人の支出や株価などの経済時系列, 伝染病の患者数や種々の流行などの社会的な変化は, それぞれが異なる特徴を持つと同時に, 様々に関連しながら変動している. 例えば, 2008 年のサブプライム・ローンの破綻は多くの企業に経済的な打撃を与えた. 自然災害も, 社会を大きく変化させる要因となる. さらに現在では, 企業や個人の間での情報の交換が社会システムの持続に不可欠な要素となっており, 特にインターネットで扱われる商品の情報・評価は商業的に重要な位置を占めている.¹

最近では, 企業間での取引や社会的な変動に対して, 従来の統計的手法に加え, 系の非線形性や相転移現象, フラクタルの概念などを組み込んだ経済物理学の分野での研究が盛んにおこなわれている.² しかし, 社会的変動は構成要素が多様であり, 変化の仕方も非定常であることが多いため力学の特定は困難である.³ 加えて, 要素間の相互作用は複雑なネットワークを形成する.^{4,5} そのため, 社会構造の把握には通常, 大規模な数値計算が必要とされ, 多くの時間的・経済的コストが掛かる.

そこで本論文では, 異なる種類の人間活動由来の時系

列データにカオス解析を適用し, 平均的特徴の抽出とデータの分類をおこなった. 異なる社会データ間の関係性を評価することで人間活動の中に存在する規則性を明らかにすることを目指す.

収集した時系列データから状態空間にアトラクタを再構成した上で,^{6,7} それらの幾何構造をリカレンスプロット解析⁸⁻¹⁰によって評価した. 特に, パワースペクトル両対数グラフの近似直線の傾きによってリカレンスプロット画像を並べ換えることで, 諸データの変動特性と再構成アトラクタの幾何構造との関係性についての指標を得た.

II. DATA

本論文では, 異なる 16 種類の人間活動由来の時系列データを収集し, それらに対してカオス解析を適用した. 以下の(a)~(p)に箇条書きで, 使用したデータ名と期間, データ数 N および入手先を示す. ここで, (a), (b)は日ごと, (m)は週ごと, 他は月ごとにサンプリングされたデータである.

(a) 日経平均株価終値 (1985.01.04~2008.11.14), $N=6697$,

- 日経平均プロフィール(<http://indexes.nikkei.co.jp/>)
- (b) WTI 原油価格 (1986.01.02~2010.05.25), $N=6155$,
Energy Information Administration (<http://www.eia.gov/>)
- (c) 企業物価指数 (1986.01~2011.04), $N=304$,
日本銀行日本銀行時系列統計データ検索サイト
(<http://www.stat-search.boj.or.jp/>)
- (d) 日本のコメ輸入量 (1988.01~2011.08), $N=284$,
財務省貿易統計 (<http://www.customs.go.jp/>)
- (e) 全国企業倒産件数 (1998.05~2011.07), $N=159$,
帝国データバンク (<http://www.tdb.co.jp/>)
- (f) 東京都のコメ日別支出 (2000.01.01~2011.08.31),
 $N=4261$, 政府統計の総合窓口(<http://www.e-stat.go.jp/>)
- (g) 米国天然ガス家庭用価格 (1981.01~2011.05), $N=365$,
石油情報センター (<http://oil-info.iej.or.jp/>)
- (h) 全国現金給与総額 (1997.08~2011.06), $N=167$,
厚生労働省 (<http://www.mhlw.go.jp/>)
- (i) 東京電力需要量 (1998.04~2010.12), $N=153$,
電気事業連合会 (<http://www.fepec.or.jp/>)
- (j) 沖縄電力需要量 (1998.04~2010.12), $N=153$,
電気事業連合会 (<http://www.fepec.or.jp/>)
- (k) 東京都自動車保有台数 (2001.07~2011.08), $N=122$,
一般社団法人日本自動車工業 (<http://www.jama.or.jp/>)
- (l) 東京都 23 区人口 (1956.01~2010.12), $N=660$,
東京都の統計 (<http://www.toukei.metro.tokyo.jp/>)
- (m) 東京都 AIDS 感染者数 (2001.01~2011.09), $N=560$,
東京都感染症情報センター
(<http://idsc.tokyo-eiken.go.jp/>)
- (n) 東京都交通事故死者(2010.08.01~2011.07.31), $N=365$,
交通財団法人 交通事故総合分析センター
(<http://www.itarda.or.jp/>)
- (o) 全国ガス事故件数 (1986.01~2011.04), $N=304$,
資源エネルギー庁 (<http://www.enecho.meti.go.jp/>)
- (p) 東京都仕送り金支出 (2000.01~2011.08), $N=140$,
政府統計の総合窓口 (<http://www.e-stat.go.jp/>)

これら 16 種類のデータを, Table 1 に示すように, まず大まかに 3 つのクラスに分類した. すなわち, (I) 企業間での取引, (II) 企業と個人の間での取引, そして (III) 個人間での取引である. さらに, 情報の流れは国から企業, 企業から個人へと (逆も然り) 段階的に伝わると仮定し, これらを階層構造として捉えた (Fig. 1). 但し, これらの分類は作業仮説的な位置づけであるため, 時系列データに対する周波数解析およびカオス解析結果をもとに再度データを分類し, Table 1 のクラスとの比較をおこなった.

Table 1. Classification of the acquired data. These data were classified into three rough classes: (I) the transactions between companies, (II) those between companies and individuals, and (III) those between individuals.

Classes	Data
(I) transactions between companies	(a), (b), (c), (d), (e)
(II) transactions between companies and individuals	(f), (g), (h), (i), (j), (k)
(III) transactions between individuals	(l), (m), (n), (o), (p)

Fig. 1. Hierarchical structure of information transactions on human activities. The dot line represents the area of data treated in this paper.

III. RESULTS AND DISCUSSION

III-A. Observation of time series

16 種類の時系列データを Figs. 2 (a)~(p) に示す. これらのグラフから, 以下のような特徴が確認できる.

(a)~(c) および (e) などから観察できる, 細かな変化に加えて, 急激な上昇や下降がある変動.

(d) や (f)~(j) に見られる, 周期性の強い変動.

(k), (l) に見られる, 緩やかに上昇・下降する変動.

(m)~(p) の乱雑な変動.

これらの観察結果を Table 1 で示した分類に照らしてまとめると, 「企業間取引」に当てはまるデータは, 日々の変化は細かく乱雑だが, 急激な上昇・下降が見られることが特徴である. すなわち, 短い時間スケールの変化に対応する差分時系列は, ほぼ平均が 0 および分散が 1 で, 各データがほぼ無相関の確率的な変動である. しかし, 実時系列には単調増加・単調減少として近似できる領域があることから, 変動には持続性がある. これは, 確率的変動の和として表されるブラウン運動の特徴を示している.¹¹ このような変動の持続性が現れることは, Fig. 2 (a) の株価や Fig. 2 (b) の原油価格のように, 取引情

Fig. 2. Time series data of various human activities. (a) Nikkei average stock price (1985.01.04~2008.11.14, daily), (b) WTI crude oil price (1986.01.02~2010.05.25, daily), (c) the corporate goods price index in Japan (1986.01~2011.04, monthly), (d) import volume of the rice in Japan (1988.01~2011.08, monthly), (e) the number of bankruptcies in Japan (1998.05~2011.07, monthly), (f) daily expenditure of the rice in Tokyo (2000.01.01~2011.08.31, monthly), (g) natural gas price in U.S.A (1981.01~2011.05, monthly), (h) monthly gross pay in Japan (1997.08~2011.06, monthly), (i) amount of the electricity demand in Tokyo (1998.04~2010.12, monthly), (j) amount of the electricity demand in Okinawa (1998.04~2010.12, monthly), (k) the number of cars in Tokyo (2001.07~2011.08, monthly), (l) population in Tokyo (1956.01~2010.12, monthly), (m) the number of infected people of AIDS in Tokyo (2001.01~2011.09, weekly), (n) the number of deaths of traffic accidents in Japan (2010.08.01~2011.07.31, monthly), (o) The number of gas accidents (1986.01~2011.04, monthly), (p) daily expenditure in Tokyo (2000.01~2011.08, monthly).

報が公開されているために、前日の値が時間的な遅れをもって後日に反映される傾向があることに起因する。

一方、Table 1 で「企業・個人間取引」に分類したデータには周期的な変動が観察できる。この結果は、Figs. 2 (i), (j) の電力需要量に代表されるように、定期的な売買によって取引が継続していることを反映している。

そして「個人間取引」に属するデータは t 日目と $t+1$ 日目の値がほぼ独立であることから、確率的変動とみなせるが、Fig. 2 (n) や 2 (p) に見られるように、突発的な変化を示すパルス状の振動が現れる特徴も持つ。

III-B. Recurrence plot analysis

本セクションでは、リカレンスプロット解析を用いて再構成アトラクタの幾何構造を評価した。

まず、時系列データを Takens の埋め込み法⁷を用いて時間遅れ座標系に変換し、状態空間にアトラクタを再構成した。すなわち、観測時系列を $y(t)$ としたとき時間遅れ値を τ として、以下の m 次元ベクトルを構成する。

$$\mathbf{v}(t) = \{y(t), y(t+\tau), y(t+2\tau), \dots, y(t+(m-1)\tau)\} \quad (1)$$

次に、一辺がデータ数 N の 2 次元画像を考え、画像の x 方向、 y 方向をそれぞれインデックス i および j とし、この点の組 $\{i, j\}$ の 2 点間距離を計算する。

$$\mathbf{R}(i, j) = H(\varepsilon - |\mathbf{v}(i) - \mathbf{v}(j)|), \quad i, j = 1, \dots, N, \quad (2)$$

ここで H は Heaviside のステップ関数 ($x < 0$ で $H(x) = 0$, $x \geq 0$ で $H(x) = 1$ を与える) であり、 $\mathbf{v}(i)$ および $\mathbf{v}(j)$ は (1) 式における i 番目および j 番目の再構成ベクトルである。 $|\cdot|$ はノルムで、本研究ではユークリッドノルムを採用した。 ε は閾値であり、2 点間距離が $\mathbf{R}(i, j) < \varepsilon$ を満たすときに第 (i, j) 画素を描写する。本解析では、 ε は 2 点間距離の最大値の 0.1 倍とした。¹⁰ この $\mathbf{R}(i, j)$ に基づいて作成される $N \times N$ 画素がリカレンスプロット (Recurrence plot: RP) である。⁸⁻¹⁰ 状態空間の次元 (埋め込み次元) m は 3 に設

定した。

対象とする系が明確な周期性や決定論的力学を有する場合、RP 画像には等間隔な点または対角線が並び、点間または線間の距離が周期に対応する。一方、ランダムデータの場合は全体において一様な分布をする。縦・横の直線はデータの定常性およびアトラクタの漸近安定性を反映している。¹⁰

最終的な配列に並べ替えた RP 画像描出結果を、後に Fig. 4 に示す。RP 画像の色は状態空間における点間距離に依存しており、距離の増加につれて白色(白)~暗色(黒)へと変化している(カラーでは黄色~紺色に対応する)。そのため、アトラクタの軌道周期が不安定である場合や急激に拡大する軌道が含まれる場合には、それに応じて RP 画像の色も変化する。16 種類のデータに対して出力したリカレンスプロットは、アルファベット順に並べても前節の仮説的分類 (Table 1) における(a)~(e): 企業間取引, (f)~(k): 企業と個人間の取引, (l)~(p): 個人間取引の 3 クラスに、きれいに対応しない。すなわち、相似の構造を持つ RP が異なるクラスにまたがって存在する。そこで、データの分類に定量的指標を導入するため、パワースペクトル密度関数を計算した。

III-C. Power spectrum analysis

パワースペクトル密度関数(PSD)を推定し、各時系列データの周期性および変動の特性を評価した。¹²

一般に、パワーが周波数 f に対して一定の減衰率を示すときには $PSD \propto 1/f^\beta$ となる。 β が $1 \leq \beta \leq 3$ をとる場合、系は非整数ブラウン運動の力学を持つとされる。¹¹ 特に、典型的なブラウン運動の場合は $\beta \cong 2$ を示す。また白色雑音のように、どの周波数領域でもパワーが一様に分布している場合には $\beta=0$ になる。ここでの β 値は傾きの絶対値を表している。

本研究で扱うデータの多くは確率過程とみなせることから、PSD の推定にはウィナー=ヒンチンの定理に基づく間接法であるブラックマン・テューキー法を採用した。具体的には、T. Niizuma *et al.* (2007) のプログラムを利用した。¹² 最小 2 乗法による近似直線の計算においては、対数軸上での成分数の偏りによる高周波帯域の過大評価を避けるため、各プロットからのズレの 2 乗和に x 座標の間隔を重みづけしている。近似直線の傾きから、そのデータの β を求めた。

Figure 3 にパワースペクトルの例を示す。Figures 3 の (i), (ii), (iii) はそれぞれ、企業間取引の(b) WTI 原油価格、企業と個人間の取引の(j)沖縄電力需要量、そして個人間取引の(n)交通事故死者数に対応する (Table 1)。各図には

パワースペクトルの近似直線も示した。¹²

Figure 3 (i) のスペクトルは、傾き -2.06 の直線によく近似できる。傾き β の絶対値がほぼ 2 であることから、原油価格の変動が非整数ブラウン運動の特徴を有することを示している。Figure 3 (ii) には時系列データ [Fig. 2 (j)] に観察される 12 ヶ月周期に対応するピークが現れている。 $|\beta|$ は 1.25 である。Figure 3 (iii) からはいくつかのピークが観察できるが、 $|\beta| = 0.180$ と比較的ゼロに近く、変動が確率過程に従っているとみなせる。また、 $|\beta|$ が小さいことは時系列変動の持続性が短いことを示している。以下では、 $\beta \equiv |\beta|$ として記述する。

Fig. 3. Examples of the power spectra and their approximate straight lines. The slope β of a power spectrum was determined by the least mean square method. (i) WTI crude oil price, (ii) electricity demand of Okinawa electric company, and (iii) the number of gas accidents.

推定された β 値は大別すると、 $\beta \cong 2$ 、 $\beta \cong 1$ および $\beta \cong 0$ であり、それぞれブラウン・ノイズ、ピンク・ノイズ、そしてホワイト・ノイズが有する特徴である。

そこで、 β 値の大小関係によるデータの分類を試みた。特に、時系列の変動特徴と、データの動力学的特徴を併せて捉えるために、リカレンスプロット解析の結果を用いた。すなわち、 β 値によって RP 画像を並べ替えることで、分類上の本質的特徴が見えてくるかどうかを、次に検証した。

III-D. Recurrence plot analysis reconsidered

各々のデータから推定した β 値に基づいてリカレンスプロットを昇順に並べた結果が Fig. 4 である. 紙面の都合上 4×4 行列で配置してあり, 行列(1, 1), (1, 2), (1, 3), ..., (4, 3), (4, 4) の順に β が大きくなる. すなわち, 行列(1, 1)にあたる Fig. 4 (n) がもっとも β が小さく, 行列(4, 4)にあたる Fig. 4 (c) がもっとも β 値が大きい.

Figure 4 を行列の順に見ていくと, β が大きくなるにつれて, RP 画像は細かく乱雑な線から成る構造から, 縦・横の直線の間隔あるいは対角線の間隔が一様な構造へと変化している. さらに大きな β では, 部分的な「島」構造が現れ, かつ曲線の境界が不明確なグラデーションと

なる構造を示している. β 値とアトラクタ構造の関係を詳細に見ていくと, $0.1 < \beta < 0.4$ の RP 画像には間隔のはっきりした直線は少なく, 全体として細かく不連続な線で覆われている. これは, 元の時系列に確率的な変動が多く含まれ, 周期的な成分や急激な値の変化が少ないことを反映している. 状態空間における軌道で捉えると, 大半がランダムに動き, かつ方向が変化するまでの時間ステップが短いことに対応する. 次に, $0.7 < \beta < 1.26$ の範囲では対角線や直線の間隔が明確であり, その間隔は再構成アトラクタの回帰軌道の周期に対応する. すなわち, 時系列に明確な周期があることに対応して, アトラクタには周期軌道が存在する. そして $1.26 \leq \beta$ の RP 画像に

Fig. 4. Recurrence plot images of various social time series data. These data have been sorted by the magnitude of the inclination of power spectrum.

は、比較的均一に分布している領域（「島」）があり、それらが境界の不明確な曲線に囲まれている。これは、状態空間における軌道は短いステップ数ではランダム的だが、同一方向に比較的長く進む傾向があることを反映している。RP 画像に現れる「島」構造は、一般的に非定常過程を示すデータが有する特徴である。¹⁰

当初の 3 分類である、企業間取引(a)~(e)、企業・個人間取引(f)~(k)、個人間取引(l)~(p)は、リカレンスプロット解析による 3 分類、すなわち非定常過程 ($1.26 \leq \beta \leq 2.70$)、周期性が強い変動 ($0.770 \leq \beta \leq 1.25$)、そして確率過程 ($0.108 \leq \beta \leq 0.389$) に大まかに対応する。しかし、一致しない部分もある。例えば、Figs. 4 (o), (l), (k) に対応するガス事故件数、東京都 23 区人口、自動車保有台数は、一見、個人間での情報の流れや共有が想定されない確率的変動と考えられるが、それらのリカレンスプロットには株価[Fig. 4 (a)] や原油価格 [Fig. 4 (b)]の RP 画像に観察される、非定常過程の特徴が現れている。加えて、Figs. 2 (o), (l), (k)の時系列には、一定の期間継続する緩やかな上昇・下降が観察される。この特徴は RP 画像において、色の境界が不明確な直線構造として反映されている。この事実は、Figs. 4 (o), (l), (k)の変動が地政学的影響を反映しやすい性質を持ち、ガス事故件数のような一見単発的に発生する現象においても、社会的共通性を表すことを示唆している。

このように、任意の社会データに対して、 β の大きさに配置した RP 画像を観察することで、他の社会的変化との関連も含めた変動の特徴を抽出することが可能となる。したがって、当初の 3 分類は、よりの確な 3 分類である非定常過程[(o)(e)(b)(a)(g)(l)(k)(c)], 周期性が強い変動[(f)(d)(j)], 確率過程[(n)(i)(p)(h)(m)] へと精緻化される。

重要な特徴として、Fig. 4 (a)の株価や Fig. 4 (b)の原油価格などの比較的大きなデータ数を持つ RP 画像では、画像の一部分を拡大しても全体と相似な構造が現れる。これは再構成アトラクタが自己相似性を有することを示唆している。特に、拡大スケールを縦・横で変化させることで、相似な構造を得やすい。また、Fig. 4 において「島」構造が観察され、曲線の境界が不明確なグラデーションとなるような非定常過程の特徴を示したのは、 β 値が $1.26 \leq \beta \leq 2.70$ の範囲にあるデータであった。これは、系が非整数ブラウン運動の力学を持つとされる β の範囲内 ($1 \leq \beta \leq 3$) に収まっている。その場合の時系列データは、一部分を抜き出して縦・横の軸のスケールを異方的に拡大すると全体のパターンと相似になる、自己アフィン性を有する。¹¹ これらの結果は、時系列の自己アフィン性が再構成アトラクタの自己相似性に帰着することを示唆

している。

IV. CONCLUSION

本論文では、種々の人間活動由来の時系列データに対してスペクトル解析およびリカレンスプロット解析を適用し、データの分類法を提案するとともに、変動の特徴を抽出した。その結果、以下のことが得られた。

(1) リカレンスプロット解析の結果、各々のデータから再構成したアトラクタの幾何構造を特徴づけることができた。特に、パワースペクトル両対数グラフの近似直線の傾き β の増加につれて、RP 画像は短い線が乱雑に分布した構造から、周期性が明確な構造を経て、非定常性を反映した構造へと変化した。この結果は状態空間におけるアトラクタの構造が β に依存して変化することを示している。

(2) 扱った社会データについて、確率過程、周期性が強い変動、そして非整数ブラウン運動という 3 種の特徴ないし分類が得られた。これらは、仮説的分類として提案した個人間取引、企業・個人間取引、および企業間取引と大まかに対応する。

本研究で提案したカオス解析手法による分類方法は、任意の社会データに対しても適用が可能である。特に、リカレンスプロットは力学系が共通している場合には同種の構造を示すことから、データの特徴を容易に識別し、かつ比較することが可能である。¹³ ただし本稿で扱ったデータは 16 種類のみであるため、より正確な結果を得るためには、更に多くの人間活動データに対して、時系列データ同士の相関性も考慮した解析手法を適用する必要がある。加えて、リカレンスプロットの定量化手法を用いた、より詳細な評価も課題である。

REFERENCES

- ¹ T. Mizuno, and T. Watanabe, *The European Physical Journal B, Condensed Matter and Complex Systems*, 76.4, pp. 501-505(2010).
- ² M. Takayasu, and H. Takayasu, *Complex Systems in Finance and Econometrics*, pp. 444-463(2010).
- ³ T. Terasaki, T. Ikeguchi, K. Aihara, and S. Tanaka, *IEICE*, 12, 1601(1995).
- ⁴ K. Kamehama, T. Iino, H. Iyetomi, Y. Ikeda, T. Ohnishi, H. Takayasu and M. Takayasu, *Journal of Physics: Conference Series*. Vol. 221, No. 1, 012013 (2010.)
- ⁵ R. Albert and A-L. Barabasi, *Reviews of modern physics*, 74(1), 47 (2002).
- ⁶ K. Aihara edit. “CHAOS JIKIETSU KAISEKI NO

KISO TO OYO”, Sangyo Tosyo, 2002. Chapter 1~5, in Japanese.

- ⁷ F. Takens, in: *Lecture notes in mathematics*, Vol. 898. *Dynamical systems and turbulence*, eds. D. Rand and L.S. Young (Springer, Berlin, 1981) p. 366.
- ⁸ M. C. Casdagli, *Physica D*, 108, 1-2(1997).
- ⁹ N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdwan, and J. Kurths, *Phys. Rev. E*, 66, 026702 (2002).
- ¹⁰ N. Marwan, M.C. Romano, M. Thiel and J. Kurths, *Physics Reports*, 438, pp.237-329(2007).
- ¹¹ J. FEDER, *Fractals*. Japanese Translation by Matsushita, Honjo and Hayakawa., Keigaku Publishing Co., 1991.
- ¹² T. Niizuma, A. Toba, Y. Yajima, K. Hisada, S. Okino and F. Suda, *J. Adv. Sci.* Vol. **19**, No.3 and 4, 41(2007).
- ¹³ G. Robinson and M. Thiel, *Chaos* 19, 023104 (2009).