様式Ｆ－１９

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成　25　年　5　月　31　日現在

機関番号：１２６０１
研究種目：若手研究(B)
研究期間：　2011　～　2012
課題番号：２３７００２６１
研究課題名（和文）　点過程時系列データのための非線形時系列解析

研究課題名（英文）　Nonlinear time series analysis for point processes

研究代表者
平田　祥人（HIRATA YOSHITO）
東京大学・生産技術研究所・特任准教授
研究者番号：４０５１２０１７

研究成果の概要（和文）：本研究は，点過程時系列データを非線形力学系の観点から特徴づけるため，埋め込みを保証するための実用的な方法を提案することを目的とした．点過程時系列データに対して，時間窓の大きさはサンプリング間隔が一定の時系列データの埋め込み次元に対応している．そこで，本研究では予測誤差を用いて点過程の時間窓を定める手法を提案した．埋め込みが保証できることにより，非線形予測，外力の再構成，方向性結合の検定，カオス性の検定の４つの手法を点過程時系列データに対して利用することが可能になった．

研究成果の概要（英文）：This research aimed to propose a practical method for ensuring the embedding to characterize point process data from the viewpoint of nonlinear dynamics. For point process data, the size of time window corresponds to the embedding dimension for a time series sampled at a fixed sampling interval. Thus, I proposed a method to choose the size of time window by using prediction errors. By ensuring the embedding, we enable to apply, to point process data, 4 methods including nonlinear prediction, reconstruction of driving forces, identification of directional couplings, and test of deterministic chaos.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
交付決定額	3,200,000	960,000	4,160,000

研究分野：非線形時系列解析
科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング
キーワード：複雑系，時系列解析，信号処理，非線形現象，非線形理論

１．研究開始当初の背景

(1) 点過程時系列データとは，対象に関するイベントが不規則な時刻に観測されたデータである．典型的な現象は，脳神経の発火，地震の発生系列，経済取引などがある．このような点過程時系列データは大きく分けると，２種類に分類することができる．１種類は，イベントに何らかの付加的な情報を伴う点過程時系列データである．データのことをマーク付き点過程時系列データと呼ぶ．付加的な情報のことをマークと呼ぶ．典型的な例は，地震の発生系列や経済取引である．地震の発生系列では，地震の緯度・経度・深さ・マグニチュードがマークである．それに対して，経済取引では，取引の価格や量などがマークになる．他にも，雨粒や雷などが点過程時系列データの例である．それに対して，もう１種類は，値を伴わない点過程時系列データである．値を伴わない点過程時系列データでは，イベントの起こった時刻のみが観測される．典型的な例は，脳神経の発火系列である．このように，点過程時系列データの例は幅広く，その生成メカニズムに，将来の

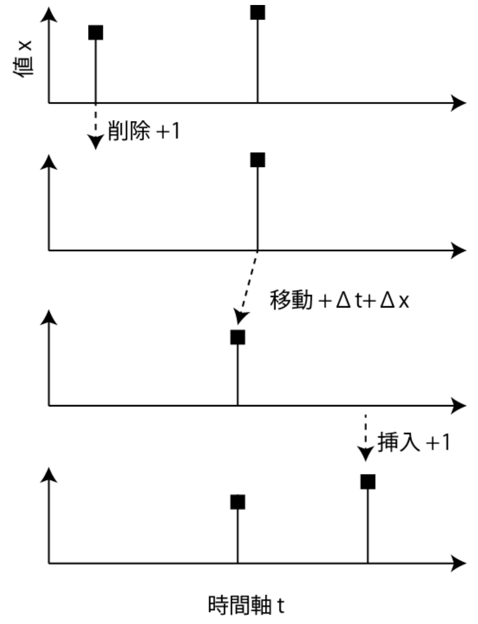

図１:(マーク付き)点過程時系列データ間の距離の計算の模式図.

状態が過去と現在の状態に依存するような力学系の要素を含むことが十分に考えられる．しかし，点過程時系列データを力学系として特徴づける非線形時系列解析手法としては Sauer(Phys. Rev. Lett. (1994))によるイベントの発生間隔を用いた埋め込み定理があるのみで，非線形時系列解析の広範な手法が点過程時系列データに対して使える状況ではなかった.

(2) そこで，点過程時系列データに対しても距離が定義できることに着目し(Hirata and Aihara, J. Neurosci. Methods (2009); Suzuki, Hirata, and Aihara, Int. J. Bifurcat. Chaos (2010))，距離を使って行う非線形時系列解析手法，例えば，リカレンスプロットを用いる手法を発展させてきた(Hirata and Aihara, Phys. Rev. E (2010a); (2010b); Hirata and Aihara, Int. J. Bifurcat. Chaos (2011)). 点過程の距離は，2 つの点過程時系列データの時間窓間に対して，一方の時間窓を編集しつつ，もう一方の時間窓を作る時の合計の手間の最小値として定義される (Victor and Purpura, Network (1997); Suzuki, Hirata, and Aihara, Int. J. Bifur. Chaos (2010))(図１)．許す操作は, イベントの削除・挿入・移動である．削除と挿入に対しては，手間１を割り当てる．イベントの移動に対しては，移動させるイベントの時間差・マークの値の差に比例する量を手間として割り当てる．このような距離を使って非線形時系列解析の手法を構成しようとする中で問題点として明らかになってきたことは，点過程時系列データに対して埋め込みの性質を利用できれば広範な非線形時系列解析手法が適用可能になることである．そのような解析の例としては，非線形予測，時間的にゆっくりと変化する外力の再構成，ネットワークシステムにおける方向性の結合の検定，決定論的カオスの検定等がある．そのため，もしも点過程時系列データに対して埋め込みを保証することができれば，点過程時系列データを力学系の視点から特徴づける有効な手段となり得る.

(3) しかし，本研究開始以前の点過程時系列データの解析の研究では，点過程時系列データに対して，任意の時刻に対して埋め込みであることを保証するような実用的な手法は存在しなかった.

(4) 点過程時系列データに対して，時間窓の大きさが，サンプリング間隔が一定の時系列データに対する埋め込み次元に相当する量になると考えられる．しかし，埋め込みの意味で時間窓の大きさを適切に選ぶ手法は存在していなかった.

２．研究の目的

そこで，本研究では点過程時系列データに対して様々な非線形時系列解析の手法を応用することが可能になるように，埋め込みを保証し時間窓の大きさを実用的に決める手法を提案することを目的とした．この手法により，非線形予測，外力の再構成，ネットワーク構造の推定，決定論的カオスの検定等が点過程時系列データに対しても使えるようになることを目指した.

３．研究の方法

(1) １年目には，点過程時系列データの時間窓が埋め込みになることを保証できる実用的な手法の研究を行った．３つの手法を試みた．１つ目の手法はサンプリング間隔が一定の時系列データで標準的に用いられる誤り近傍法(Kennel, Brown, and Abarbanel, Phys. Rev. A (1992))の拡張，２つ目の手法はリカレンスプロット (Marwan *et al.*, Phys. Rep. (2007))の斜めの線の長さのヒストグラムが指数関数的に減少する性質を利用する手法，３つ目の手法は予測誤差を使う手法である．最も単純な非線形予測としては，過去か

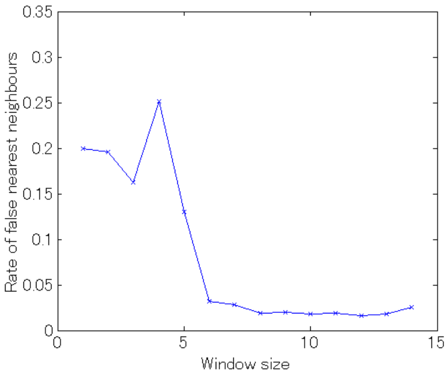

図２：誤り最近傍法の結果(レスラーモデルの極大値系列の例).

ら最も距離が近い時間窓を探してきて，その次の時間窓を予測として返す予測手法が考えられる(最近傍予測)．予測誤差は，やはり点過程時系列データの距離を使って与えることができる．最近傍予測との比較には，現在の時間窓が繰り返されると思う持続予測を用いた．つまり，最近傍予測の予測誤差の平均を持続予測の予測誤差の平均で割った値が最も小さくなるところで時間窓の最適な大きさを決めた．

本手法の妥当性を評価するために，積分発火ニューロンモデルから生成される，値を伴わない点過程時系列データやレスラーモデルの極大値系列(マーク付き点過程時系列データ，Yabuta and Ikeguchi, Proc. NOLTA (2007))に応用して手法の妥当性を評価するとともに，脳神経・為替取引・地震などの実データで手法の有効性を検証した．手法の有効性の評価には予測誤差を用いた．つまり，現在の点過程の時間窓が繰り返されると思う持続予測に対して過去から良く似ている時間窓を探しその次の時間窓を予測とする最近傍予測を用いる時，予測誤差が小さくなるかどうかを基準とした．

(2) 2 年目には, 1 年目に開発した手法を外力の再構成，ネットワーク構造の推定，決定論的カオスの検定に応用し，応用可能性を吟味した．これらの応用の過程で，1 年目に提案した手法が妥当であるかどうか検討し，必要に応じて修正を加えた．

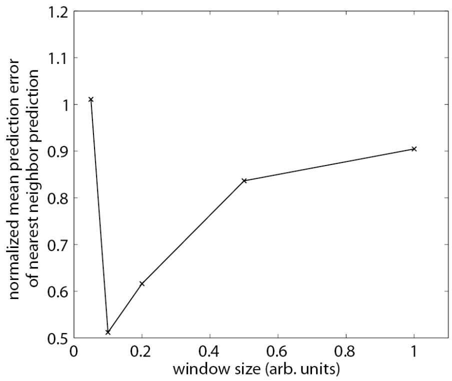

図３：ローレンツモデルを積分発火ニューロン入力として加えて生成した点過程時系列データに対する，時間窓の長さと正規化した予測誤差の関係.

４．研究成果

(1) ３. (1)で候補に挙げた 3 つの手法のうち，1 つ目の誤り近傍法の拡張は適切な方法ではないことが分かった．時間窓を大きくしていって新しいイベントが時間窓に加わったところで距離がジャンプしてしまうため，誤り近傍点の割合が不自然に変化してしまうという問題点があったためである(図２)．

３. (1)で候補に挙げた 2 つ目の方法も適切な方法でないことが分かった．十分大きな時間窓を用いると，リカレンスプロット上の斜めの線の長さに関する分布が常に一定に指数関数的に減少するという性質を示すと予想したが予想と合わない結果が得られた(後の部分で議論)．この結果も時間窓を大きくしていった時に新たなイベントを含むところで距離が離散的にジャンプしてしまう性質に基づくものと思われる．

３. (1)で挙げた 3 つ目の，予測誤差を用いる手法では，最近傍予測の方が持続予想よりも良い予測をより高い確率で与えた．例えば，ローレンツモデルの出力を積分発火ニューロンモデルに加えて，値を伴わない点過程時系列データを生成した時，図３のように長さ 0.1 の時間窓が最適として選ばれた．この時間窓を用いると, 最近傍予測は確率 0.80 で持続予測よりも良い予測を与えた．

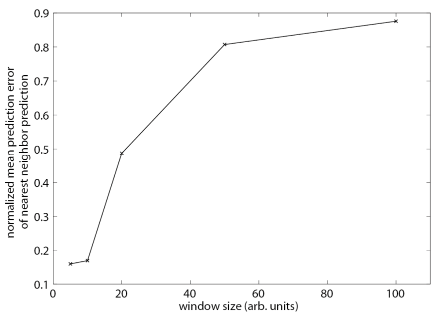

図４：レスラーモデルの極大値系列に対する，時間窓と正規化した予測誤差の関係．

マーク付き点過程時系列データに対しても同様に最適な時間窓を選ぶことができる．ここでは，レスラーモデルの極大値の時刻と x の値の系列を用いた．この時，図４に示すように最適な時間窓の大きさは５であった．この時間窓を用いて予測すると，最近傍予測は，持続予測に対して勝率 0.99 であった．

この最適な時間窓を用いて，リカレンスプロット(Marwan *et al.*, Phys. Rep. (2007))を求め，斜めの線の数を数えた．サンプリング間隔が一定の時系列データでは，埋め込み次元が十分大きい時には斜めの線の数の分布は指数関数的に減少するが，レスラーモデルの極大値系列のケースでは，長さ１から長さ２の間での数の減少率がそれ以降の部分に比べて大きかった(図５)．そのため，２つ目のリカレンスプロットの斜めの線の数の減少率を使う手法は時間窓の大きさを決める手法として有効ではない．

予測誤差が埋め込みを実用的に決める上で良い指標となる理論的な説明は，以下のように与えることができる．現在の力学系の状態を x_{wt} とする．x_{wt} は，時間 w の間にダイナミクス f^w によって，状態 $x_{w(t+1)} = f^w(x_{wt})$ に変化するとする．状態 x_{wt} から，時間 w の間に点過程 $G_w(x_{wt})$ が観測されるとする．同様に $x_{w(t+1)}$ からも時間 w の間に点過程 $G_w(x_{w(t+1)})$ が観測されるとする．この時，図６のようなダイアグラムが書ける：もし，点過程の観測 G_w が埋め込み，つまり，1 対 1 であるならば，観測の逆写像 G_w^{-1} が存在し，$h^w = G_w \circ f^w \circ G_w^{-1}$ と書ける．

$$\begin{array}{ccc} x_{wt} & \xrightarrow{f^w} & x_{w(t+1)} \\ \downarrow G_w & & \downarrow G_w \\ G_w(x_{wt}) & \xrightarrow{h^w} & G_w(x_{w(t+1)}). \end{array}$$

図６：点過程埋め込みのダイアグラム．

これは，G_w が埋め込みならば，点過程 $G_w(x_{wt})$ から点過程 $G_w(x_{w(t+1)})$ を予測するような予測方法が作れることを意味する．

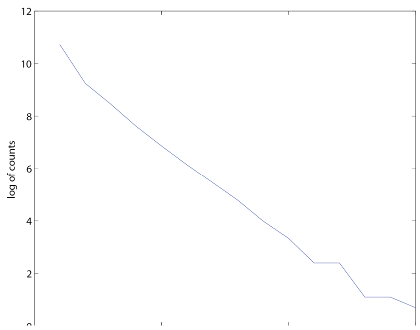

図５：レスラーモデルの極大値系列のリカレンスプロット(時間窓のサイズ５)に含まれる斜めの線の数の分布．

実際に用いたい性質は，この性質の対偶である．つまり，点過程 $G_w(x_{wt})$ から点過程 $G_w(x_{w(t+1)})$ を予測するような予測方法が作れないならば，G_w は埋め込みでない．この意味で，本研究で提案する手法は，実用的に G_w の埋め込みを保証する手法になっている．

実際に実データに対しても応用した．窓の長さを上記の手法で決めると，コオロギの気流細胞(Suzuki *et al.*, Biol. Cyber. (2000))と地震のデータ(気象庁提供，日本周辺のマグニチュード４以上のデータ．2000 年から 2010 年)では，最近傍予測により，持続予測よりも良い予測が得られる傾向にあった．為替のデータに対しては，解析した期間に結果が依存した．

上記の考察の結果，予測誤差を用いる手法を基にして，埋め込みを保証するような時間窓の大きさを決定する手法を構成していくことにした．

(2) 2 年目の応用研究に進むにあたり，時間窓の適切な大きさをより正確に求める必要性が出てきた．当初，指定した時間窓の大きさの範囲で，持続予測の予測誤差で正規化した時の最近傍予測による予測誤差が最も小さくなる時間窓を選ぶことにしていた．しかしこの基準をより厳密に使うと，為替取引のデータで，小さな時間窓を使うほど良いということになり，究極的には，1 秒の時間窓が選ばれることになった．しかし，この場合，この時

間窓の間に取引が１つあるかないかの状態になり，為替取引の大局的な状態，つまり埋め込みに相当する状態が反映されているとは言えない状況になった．

そこで，Uzal, Grinblat, and Verdes (Phys. Rev. E (2011))の考え方を導入し，予測誤差の評価を１ステップで行うのではなく，５ステップ先の時間窓までの予測誤差の合計で評価することにした．その結果，為替取引のデータでも選ばれる時間窓が数時間のオーダーの窓になり，妥当な結果になった．そこで，今後は５ステップ先までの予測誤差の合計を使って時間窓を決定することにした．

予測誤差を用いた時間窓の大きさの決定手法を，まずは外力の再構成に用いた．外力の再構成では，まず点過程の距離を用いてリカレンスプロットを用いて，そのリカレンスプロットを疎視化し，Hirata *et al.* (Eur. Phys. J. Spec. Top. (2008))の手法を使って，疎視化したリカレンスプロットを連続値の時系列データに変換することによって行った．この手法を，ゆっくりとした周期的な外力によって駆動されるレスラーモデルから生成した x の極大値系列に対して適用し，手法の妥当性を検討した．その結果，予測誤差に最適な時間窓よりも，その時間窓の２倍の大きさの時間窓を用いる方が，より正確に外力が再構成できることが分かった．この結果はサンプリング間隔が一定の時系列データで，埋め込み次元を大きめにとると，背後に隠れた外力の情報を再構成できる(Casdagli, Physca D (1997); Stark, J. Nonlinear Sc. (1999); Hegger *et al.*, Phys. Rev. Lett. (2000))という先行研究の結果と一致する．また，為替取引のデータに応用した結果，１週間の周期性が明らかになった．これは，人間の活動に周期性があるところから来ており，手法の妥当性を示す証拠であると考えられる．

ネットワーク構造の推定の問題に関しても，ネットワークの距離と Hirata and Aihara (Phys. Rev. E (2010a))の，リカレンスプロットを用いる手法により行った．Hirata and Aihara (Phys. Rev. E (2010a))の手法では，x から y への方向性の結合がある時，遅れ座標により求めた x のリカレンスプロットが，同じく遅れ座標によって求めた y のリカレンスプロットを覆うことができるという性質を使って，方向性の結合を検定する．Hirata and Aihara (Phys. Rev. E (2010a))で例として取り上げられている結合レスラー方程式から生成した極大値系列を解析した結果，ネットワーク推定に関しても，予測に対して最適な時間窓を使うよりも，その２倍の大きさの時間窓を使う方がより正確にネットワーク構造を推定できることが分かった．また，コオロギの気流細胞をレスラーカオスの風で振動させた時の気流細胞の応答(Suzuki *et al.*, Biol. Cyber. (2000))を解析した結果，気流細胞が風によって駆動されているという，実験の設定と符合する結果を得た．

また，点過程の距離と Hirata and Aihara (Phys. Rev. E (2010b))の手法を組み合わせて，Devaney(An Introduction to Chaotic Dynamical Systems, 1989)の意味での決定論的カオスの検定が構成できるかどうか検討した．Hirata and Aihara (Phys. Rev. E (2010b))の論文では，Devaney のカオスの 3 つの条件である位相推移性，周期解の稠密性，初期値鋭敏性の中に現れる「任意の近傍」の部分を「リカレンスプロットを求める時に用いた近傍」というように条件を緩和することで，決定論的カオスの検定を構成している．緩和した条件が満たされなければ，元々の条件が満たされないという論理による．位相推移性とは，任意の開近傍が任意の開近傍の時間発展と空でない積集合を持ち得るという条件である．周期解の稠密性は，どんな開近傍を取って来てもその中に周期解を見つけることができるという性質である．初期値鋭敏性は，どんな近傍点の２点を取って来ても，距離がある値以上大きくなるような時間が存在するという性質である．

点過程の距離を使って求めたリカレンスプロット上での決定論的カオスの検定では，イベントの削除や挿入に伴う距離の離散的な変化を克服する必要があった．この問題を克服するために，Hirata and Aihara (J. Neurosci. Methods (2009))で行ったように，それぞれの時間窓で時間窓前後のイベントを考慮に入れる場合と入れない場合の組み合わせ，合計 16 通りの距離を計算し，その最小値を改めて局所的な距離として定義し，ISOMAP(Tenembaum *et al.*, Science (2000))のようにこの局所的な距離を貼り合わせて行って大局的な距離を定義することにした．この手法を用いると，非常に周期解に近い決定論的カオスのパラメータのレスラーモデルから生成した極大値系列では正しい判別はできなかった

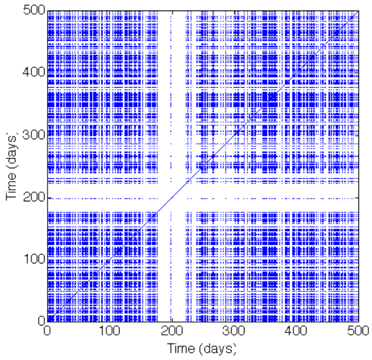

図７：地震のリカレンスプロット.

が，典型的な周期解のパラメータのレスラーモデルから生成した極大値系列，典型的な決定論的カオスにパラメータを選んだ時のレスラーモデルの極大値系列で，それぞれ正しく判別できた．また，この手法の妥当性を積分発火ニューロンモデルにローレンツモデルの x を入力として加えたケースと，コオロギの気流細胞をレスラーモデルをベースに生成した風で振動させた時の応答(Suzuki *et al.*, Biol. Cyber. (2000))でも評価した．加えて，日本周辺の地震のデータ(気象庁提供，2000 年から 2010 年のマグニチュード 4 以上の地震を使用)について評価した結果，位相推移性と周期解の稠密性を満たさず，Devaney のカオスとは整合的でないという結果を得た(リカレンスプロットを図７に示す)．

(3) 本研究の結果として，点過程時系列データに対して，非線形時系列解析の様々な手法，特に，非線形予測，外力の再構成，ネットワーク構造の推定，決定論的カオスの検定が使えるようになった．これらの手法を自然現象や工学的なシステムに適用することにより，点過程時系列データの背後にある力学系の性質が今後明らかになっていくであろう．

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 2 件）

① 平田祥人, 点過程時系列データの非線形時系列解析, システム/制御/情報 56, 355-360 (2012). 査読有 http://ci.nii.ac.jp/naid/110009480292

② Yoshito Hirata and Kazuyuki Aihara, Timing matters in foreign exchange markets, Physica A 391, 760-766 (2012). 査読有 doi:10.1016/j.physa.2011.09.013

〔学会発表〕（計 4 件）

① 平田祥人，合原一幸，点過程時系列データとその非線形時系列解析，第 2 回社会神経科学研究会「社会の中で生きる心の理解」，2013 年 1 月 31 日，岡崎

② Yoshito Hirata, Practically ensuring embedding for analyzing point process data, 2012 International Symposium on Nonlinear Theory and its Applications, 2012 年 10 月 24 日，パルマ(スペイン)

③ 平田祥人，近江崇宏，尾形良彦，合原一幸，点過程の距離を使った地震予測，第 3 回研究集会「地震活動の評価に基づく地震発生予測システム：東北地方太平洋沖地震前後の地震発生予測」，2012 年 7 月 12 日，東京

④ Yoshito Hirata and Kazuyuki Aihara, Recurrence plots for point processes and their application to earthquakes, The 4th International Symposium on Recurrence Plots, 2011 年 12 月 7 日，香港(中国)

〔図書〕（計 1 件）

① Yoshito Hirata, Eric J. Lang, and Kazuyuki Aihara, Analyzing multiple spike trains using distance measures and recurrence plots, edited by N. Kasabov, Springer Handbook of Bio- and Neuroinformatics, Springer, 2013 in press.

〔その他〕

ホームページ等

http://www.sat.t.u-tokyo.ac.jp/~yoshito

６．研究組織

(1)研究代表者

平田　祥人 (HIRATA YOSHITO)

東京大学・生産技術研究所・特任准教授

研究者番号：４０５１２０１７

(2)研究分担者

（　　）

研究者番号：

(3)連携研究者

（　　）

研究者番号：