

変動パターンを考慮した非線形時系列解析

数理情報学専攻 48076217

帯刀一洋

指導教員

鈴木秀幸 准教授

1 はじめに

世の中には複雑な挙動をする時系列が数多く存在する。例えば、脳波、電気回路における電圧や電流、気温、更には株価のような経済活動を表す指標などである。こういった複雑な挙動をする時系列を解析する際に、確率論的立場から解析する手法と、決定論的立場から解析する手法がある。後者を非線形時系列解析手法と呼び、時系列の不規則さは我々が予測することができない確率的な誤差として排除するのではなく、システムに内在する非線形性によって決定論的に生じていると捉える。

本論文では非線形時系列解析手法の中の2つ、局所線形近似予測法と、リカレンスプロット法を使って時系列の解析を行う。これら2つの手法の共通の問題として、時系列の各時点の類似度を測る距離関数をどうやって定義するかという問題がある。通常はユークリッド距離で定義されるが、時系列が短い場合、埋め込みが成り立たない場合などでは不十分な事が多い。そこで、本論文では、新たに変動パターンを考慮した距離関数を定義することで、精度の向上を目指す。

2 時間遅れ座標

非線形時系列解析手法を使う上でまず初めに行わなければならないのは、時間遅れ座標への変換である。時間遅れ座標への変換は観測値を $s(t)$ とした時に、ある正の整数 m と τ を用いて以下のように m 次元ベクトル (式 (1)) を構成する。

$$\mathbf{x}(t) = (s(t), s(t - \tau), \dots, s(t - (m - 1)\tau)), \quad (1)$$

ここで、 m は埋め込み次元、 τ は時間遅れと呼ばれる。

3 局所線形近似予測

局所線形近似予測法は非線形時系列予測手法の中の一つである [1]。一般的な大域線形予測モデルと異なる点は、予測する時点ごとに近傍を使ってパラメータを求めるために、各時点、時点で異なる線形自己回帰モデルが設定されるということである。具体的には以下のように定式化できる。

$$\hat{s}(t + 1) = \mathbf{a}\mathbf{y}(t), \quad (2)$$

ここで、

$$\mathbf{a} = [a_0, a_1, \dots, a_m], \quad (3)$$

$$\mathbf{y}(t) = [1, \mathbf{x}(t)]^T. \quad (4)$$

尚、 Γ は転置を表す。パラメータ \mathbf{a} の値は k 個の近傍 $\mathbf{x}(t_g) (g = 1, 2, \dots, k)$ を使った以下の制約条件を解くことで求まる。

$$\mathbf{a}B = C, \quad (5)$$

ここで、 B, C は各々

$$B = [\mathbf{y}(t_1), \mathbf{y}(t_2), \dots, \mathbf{y}(t_k)], \quad (6)$$

$$C = [\mathbf{y}(t_1 + 1), \mathbf{y}(t_2 + 1), \dots, \mathbf{y}(t_k + 1)]. \quad (7)$$

のように定義された行列である。

4 リカレンスプロット

リカレンスプロット法は、時系列の順序関係に関する情報を捨てずに、時系列の局所的な時間相関を二次元平面上に視覚化する手法である [2]。リカレンスプロット法を用いることで、元の時系列の周期的振る舞いや、非定常な振る舞いなどを視覚的にとらえることができる。

リカレンスプロット法は以下のように定式化できる。

$$R_\epsilon(i, j) = \begin{cases} 1 & D^0(i, j) < \epsilon, \\ 0 & \text{otherwise.} \end{cases}$$

ただし $D^0 = \|\mathbf{x}(i) - \mathbf{x}(j)\|$ で、 ϵ は閾値である。また、 $R_\epsilon(i, j) = 1$ となるとき x 軸、 y 軸、共に時刻を取った二次元平面上 (i, j) にプロットするというを表す。尚、 $j = i$ では $D(i, i) = 0$ となるため右斜め 45° の中央対角線上には常に点がプロットされる。

5 変動パターンを考慮した距離関数

本論文で提案する距離関数はユークリッド距離と変動パターン、両方に着目した距離関数である。具体的にはパラメータ $\alpha (0 \leq \alpha \leq 1)$ を用いて以下の式 (8) のように定義する。

$$D^\alpha(i, j) = (1 - \alpha)D^0(i, j) + \alpha D^1(i, j), \quad (8)$$

ここで、 D^1 (式 (9)) は変動パターンを表す距離である。

$$D^1(i, j) = \|\Delta \mathbf{x}(i) - \Delta \mathbf{x}(j)\|, \quad (9)$$

$$\Delta \mathbf{x}(i) = (s(i) - s(i - \tau), \dots, s(i - (m - 2)\tau) - s(i - (m - 1)\tau)). \quad (10)$$

6 数値実験及び考察

本論文では以下2つの方程式から発生させた擬似データを使って数値実験を行う。

- ローレンツ方程式

$$\begin{cases} \frac{dx}{dt} = -\sigma x + \sigma y, \\ \frac{dy}{dt} = -xz + \beta x - y, \\ \frac{dz}{dt} = xy - \gamma z, \end{cases} \quad (11)$$

$$\sigma = 10, \beta = 28, \gamma = 8/3.$$

- レスラー方程式

$$\begin{cases} \frac{dx}{dt} = -y - z, \\ \frac{dy}{dt} = x + \zeta y, \\ \frac{dz}{dt} = \eta - z(x - \iota), \end{cases} \quad (12)$$

$$\zeta = 0.2, \eta = 0.2, \iota = 6.$$

上記の方程式から発生させた様々な長さ (100, 500, 5000, 10000) の時系列に対して局所線形近似予測を用いて10期先の値を予測する。尚、予測期間は200である。10期先の予測の精度を正規化二乗平均誤差 (Normalized root mean squared error, 以下 NRMSE) で評価する。その結果が図1である。 $\alpha = 0$ は通常のユークリッド距離を用いて予測した時の結果で、それ以外は提案距離関数を用いて予測した時の結果である。パラメータ $\alpha = 0$ の時の NRMSE に比べて、 $\alpha > 0$ での NRMSE が低いほど予測精度が改善されているということの意味するのだが、図1を見ると時系列が短い場合には提案距離関数を用いることで予測精度が大幅に向上することがわかる。具体的にはローレンツ方程式では時系列の長さが100の時に、レスラー方程式では、500, 1000の時に予測精度が大幅に向上している。

次にリカレンスプロットでもユークリッド距離を用いた場合と提案距離を用いた場合とで比較を行う。提案距離を用いた場合では、通常のリカレンスプロットから誤り近傍点がうまく除去され、決定論性を示す右斜め45°の方向の対角線が大半を占めるようになっている(図2, 3)。この結果は、提案距離関数を使うことで、局所線形近似予測の予測精度を向上させることができた理由の裏付けとなる。すなわち提案距離関数を用いることで、誤り近傍点を排除できた結果、予測精度が上がったと結論づけられる。

これらの結果を踏まえると提案距離関数を用いることで、元の時系列の特徴をより正確に抽出できるようになり、局所線形近似予測、リカレンスプロット、共に精度が向上することがわかる。

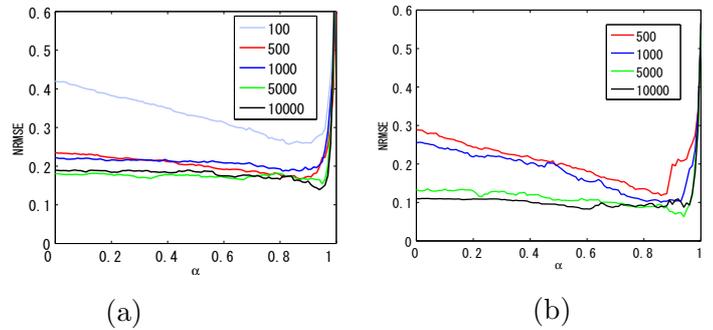


図1. 時系列の長さで予測誤差 NRMSE (a) ローレンツ方程式 (b) レスラー方程式

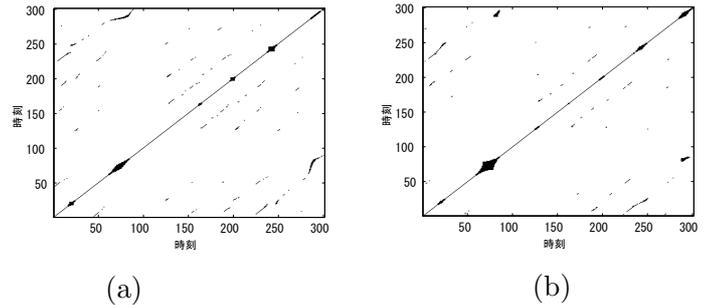


図2. ローレンツ方程式における (a) リカレンスプロット (b) 提案リカレンスプロット

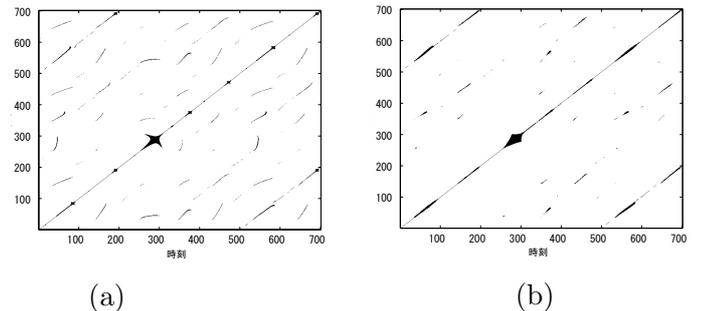


図3. レスラー方程式における (a) リカレンスプロット (b) 提案リカレンスプロット

7 おわりに

提案距離関数のパラメータ α の最適値は予測時点によって異なる。そこで、最適なパラメータ α を予測時点ごとに決定する方法を提案することを今後の課題とする。ユークリッド距離と変動パターン、 α の関係性を学習させればよいのではないかと考えている。

参考文献

- [1] Edward N. Lorenz. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, Vol. 26, pp. 636–646, jan 1969.
- [2] Jean-Pierre Eckmann, Sylvie Oliffson Kamphorst, and David Ruelle. Recurrence plots of dynamical systems. *Europhysics Letters*, Vol. 4, No. 9, pp. 973–977, nov 1987.