春期課題

新B4 福田真悟

2020.3.3

1 "Lorenz 方程式", "Rossler 方程式", "ホワイトガウスノイズ", "Logistic 写像" , "順列エントロピー", "相互情報量", "複雑ネットワーク"について書籍やインターネットで調べる

- Lorenz 方程式

  カオス的なふるまいを示す非線形微分方程式の1つで

$$\frac{dx}{dt} = -px + py \tag{1}$$

$$\frac{dy}{dt} = -xz + rx - y \tag{2}$$

$$\frac{dz}{dt} = xy - bz \tag{3}$$

で表現される [?] .変数は(x, y, z, t)，定数は(p, r, b)である．初期値鋭敏性から初期値によって，それ以降の結果が大きく変化してしまう．そのため初期値の精度には無限大の精度が必要となり，予測が事実上不可能となっている．

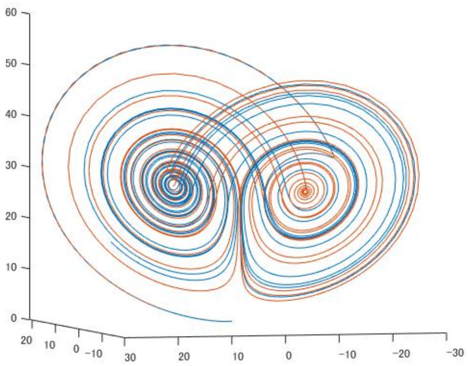

図1 Lorenz方程式の誤差の増幅

図 **??**は初期値の1つを0.1変化させたときの解の変化である．

- Rossler 方程式

  カオス的なふるまいを示す非線形微分方程式の1つで

$$\frac{dx}{dt} = -y - z \tag{4}$$

$$\frac{dy}{dt} = x + ay \tag{5}$$

$$\frac{dz}{dt} = b + xz - cz \tag{6}$$

で表現される [**?**]．変数は(x, y, z, t)，定数は(a, b, c)である．Lorenz方程式と違い，カオスの挙動を把握するために単純化された方程式であり，実際の物理モデルがベースではない．Lorenz方程式と同様に初期値鋭敏性から初期値によって，それ以降の結果が大きく変化してしまう．そのため初期値の精度には無限大の精度が必要となり，予測が事実上不可能となっている．

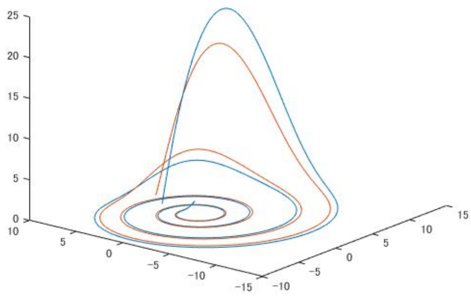

図2 Rossler方程式の誤差の増幅

図 **??**は初期値の1つを0.1変化させたときの解の変化である．

- ホワイトガウスノイズ

  自然界などのランダム過程の効果をモデル化したもの．ホワイトは周波数帯域が均一であることを示している．光の白色から由来している．ガウスは自然界のものはガウス分布に従うという仮定をもとにノイズもガウス分布となっていることを示している．波形の周波数は均一にすべての周波数で，波形の振幅はガウス分布に従う振幅となっていると仮定したノイズ．モデル化されたノイズのためホワイトガウスノイズサンプルの生成を行う関数が作られている [**?**]．

- Logistic 写像

  Logistic方程式を離散化することで得られる漸化式である．

$$x_{n+1} = ax_n(1 - x_n) \tag{7}$$

で表現される [**?**]．定数aによって，値の変化が大きく変わる．aが十分に小さいときは0に収束し，値が大きくなるにつれて，収束値が大きくなり，周期的な振動を示し，最終的にカオス的な挙動を示す．aごとのxの値の変化を下記に示す．

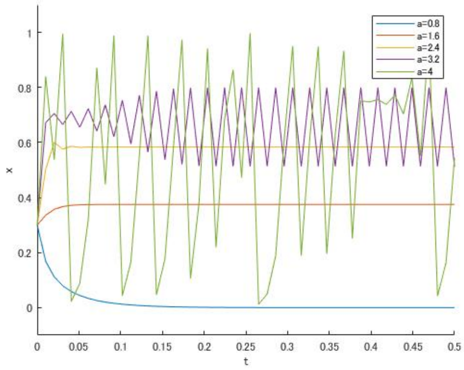

図3 aの変化によるLogistic写像の値の変化

- 順列エントロピー

　時系列の乱雑さを定量評価する指標であり，時系列の値ではなく，大小を比較する手法である．長さがLの時系列に対して遅れ時間τの間隔があるデータD個をD次元の遅れ時間座標$\mathbf{X}_j = \{x_j, x_{j+\tau}, \cdots, x_{j+(D-1)\tau}\}$とし，この座標を考える．ただし，$j = 1, 2, \cdots, L-(D-1)\tau$とする．この$D$次元の座標のそれぞれの値の大小をパターン化すると$D!$個のランクオーダーパターンに分類される．あるランクオーダーパターンπ_iの相対度数$p(\pi_i)$は，

$$p(\pi_i) = \frac{\sum_{j=1}^{N} \chi_i(X_j)}{L-(D-1)\tau} \tag{8}$$

$\chi_i(X_j)$は指示関数であり，

$$\chi_i(X_j) = \begin{cases} 1 & (\pi_j = \pi_i) \\ 0 & (\pi_j \neq \pi_i) \end{cases} \tag{9}$$

となる．そしてこの相対度数の情報エントロピー(Shannonエントロピー)で計算したものが順列エントロピーである．さらに定量的に様々な時系列を比較するために正規化を行ったものを順列エントロピーと定義する．その式を下記に示す．

$$H_p = \frac{-\sum_{i=1}^{D!} p(\pi_i)\log_2 p(\pi_i)}{\log_2 D!} \tag{10}$$

- 相互情報量

　2つの情報源の依存度を表した値である．一方の情報をもっているとき，他方の情報を得たらどの程度の不確かさ(情報エントロピー)が減少するのかを示している．2つの情報源X, Yからそれぞれから情報x, yを得られる確率を$P_X(x), P_Y(y)$とし，同時に得られる確率を$P_{X,Y}(x, y)$とすると，相互情報量の式は，

$$H_p = \sum_{x\in X}\sum_{y\in Y} P_{X,Y}(x,y)\log_2 \frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)} \tag{11}$$

となる[?].

- 複雑ネットワーク[?]
  ノード(点)間をリンク(線)で結んで構成されているものをネットワークという．その中で「スケールフリー性」，「スモールワールド性」などの特徴をもつものを複雑ネットワークという．　あるノードに結ばれているリンクの数を次数といい，その分布を次数分布という．複雑ネットワークにおいては，次数分布がべき分布となっている．べき分布は確率変数Xの要素xに対して，確率密度関数が，

$$f(x) = ax^k \tag{12}$$

となる．式(??)の確率変数が定数倍(c倍)になったとき，確率密度関数は，

$$f(x) = a(cx)^k \tag{13}$$

となり，式(??)のc^k倍すれば表現できるため，べき分布はスケール不変性を持っている．そのため複雑ネットワークは「スケールフリー性」という特徴を持つ．　2つのノードを結ぶ最短のパスの長さの平均を平均ノード間距離，あるノードに対して，隣接ノードの内2つのノードを選んだときのその2つのノードがリンクされてるときの確率の平均を平均クラスター係数とする．複雑ネットワークでは，平均ノード間距離が小さく，平均クラスター係数が大きい．このことを「スモールワールド性」という．

2 相互情報量とLorenz 方程式に関するMATLAB ファイル(mutual.m) を実行し，図を出力する

出力された結果を下記に示す.

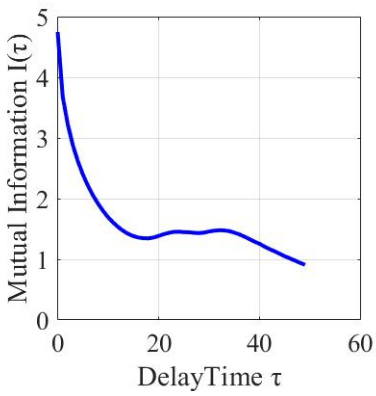

図4 課題2の出力結果

また，時間遅れの極小値は18となった．このことからアトラクタの構築するために最適な時間遅れは18だと考えられる．x軸をMATファイルの時系列データ$f(t)$とし，y軸をMATファイルの時系列データ$f(t+\tau)$とし，z軸をMATファイルの時系列データ$f(t+2\tau)$としたときのプロットを下記に示す.

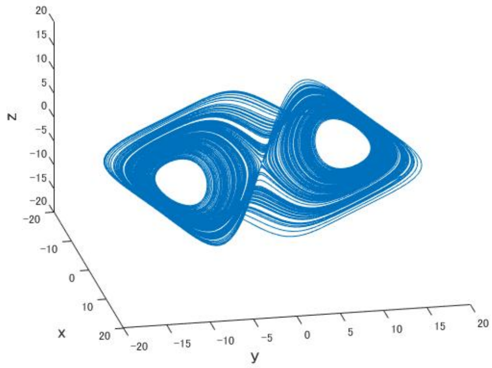

図5 最適な時間遅れのアトラクタ

また，これに対して不適切な時間遅れである$\tau = 2, 1000$のパターンも下記に示す．

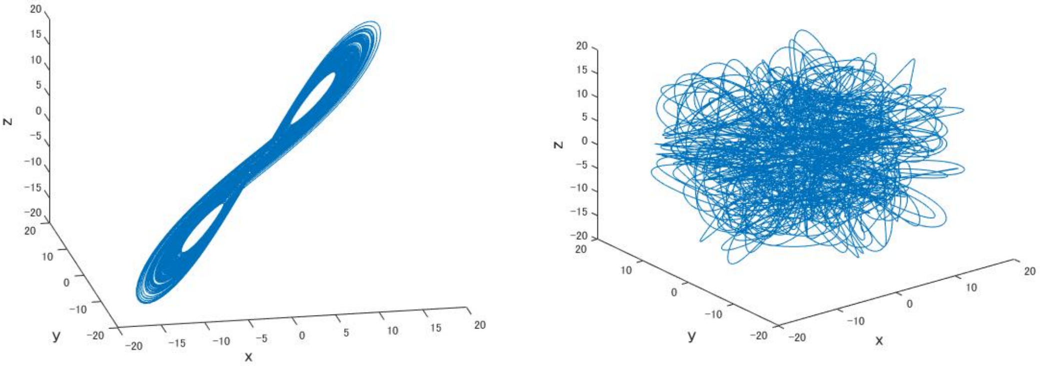

図6 不適切な時間遅れのアトラクタ(左が$\tau = 2$，右が$\tau = 1000$)

これは，時間遅れが小さいと時系列と遅れ時系列の相関が高くなり，傾き45°の直線近傍のプロットとなってしまう．時間遅れが大きいと無相関となり，不安定な軌道を示してしまう．

3 ホワイトガウスノイズと順列エントロピーに関するMATLABファイル(B4kadai_2020.m)を実行し，図を出力し考察を加える．

まず出力された結果を下記に示す．このときのホワイトガウスノイズの生成数は10000個で次元数は5，遅れ時間は1である．

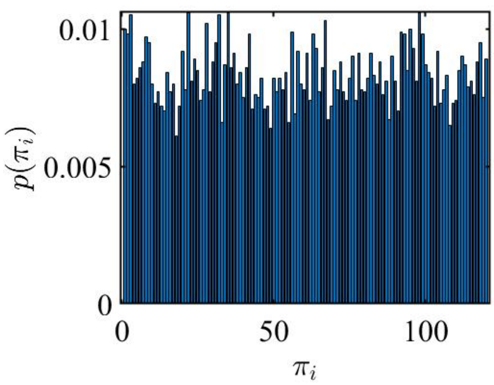

図7 課題3の出力結果(生成数は10000個で次元数は5，遅れ時間は1)

ホワイトガウスノイズは周波数帯域が均一のため理論的には，ランクオーダーパターンも均一になると考えられる．つまりランクオーダーパターンのエントロピーである順列エントロピーは理論的には，1となると考えられる．実際に順列エントロピーの計算すると0.9989となった．予想の結果との誤差は0.11[%]となるので今回の計算結果は妥当であると考えられる．

結果からホワイトガウスノイズのランクオーダーパターンにはバラつきがあることがわかる．今回のシミュレーションでは，次元数5($5! = 120$個のパターン)に対して，ノイズの生成数が10000となっているので平坦に見えるにはデータ数が少ない．実際に$p(\pi_i)$の平均値は約0.083($= 1/120$)となっており，各ランクオーダーパターンの相対度数のヒストグラムを制作すると約0.083($= 1/120$)を中心にガウス分布に近いものとなっている．ノイズの生成を増やせば，各ランクオーダーパターンの相対度数は均一に近づいていくと考えられる．下記にノイズの生成数が100000とノイズの生成数が1000000に増やしたときの出力結果を示す．

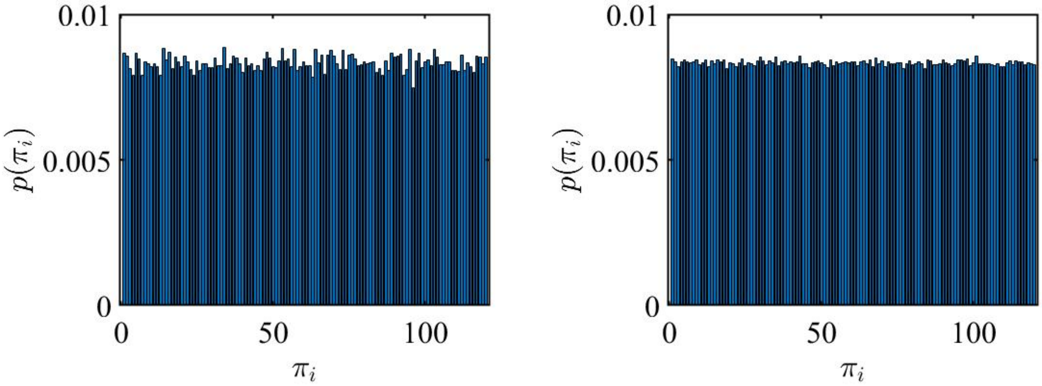

図8 課題3の出力結果(左が生成数は100000，右が生成数は1000000)

この結果からデータ数を増やすことでバラつきが小さくなることがわかった．また，ノイズの生成数が100000のときのヒストグラムも下記に示す．

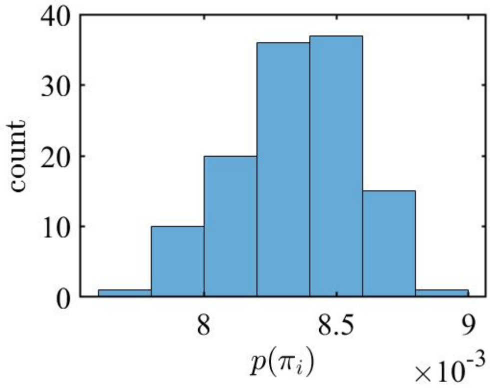

図9 課題3のヒストグラム(生成数は100000)

参考文献

[1] ローレンツ方程式，http://www.isc.meiji.ac.jp/~random/lecture/2015-comp2/Lorentz_equ.html

[2] カオス・フラクタル 講義ノート #8，https://ocw.hokudai.ac.jp/wp-content/uploads/2016/01/ChaosFractal-2011-Note-08.pdf

[3] ホワイトガウスノイズサンプルの生成-MATLAD wgn，https://jp.mathworks.com/help/comm/ref/wgn.html

[4] ロジスティック写像の時系列 - 工学院大学，https://brain.cc.kogakuin.ac.jp/~kanamaru/Chaos/Logits/

[5] 相互情報量の意味とエントロピーとの関係 — 高校数学の美しい物語,https://mathtrain.jp/mutualinfo

[6] 複雑ネットワーク：統計物理学の視点，http://mercury.yukawa.kyoto-u.ac.jp/~bussei.kenkyu/pdf/03/1/9999-031210.pdf