

# リカレンスプロット<sup>ことはじめ</sup>事始

しろまさのり  
城 真範 (産総研 人間情報インタラクション)

第1回リカレンスプロット研究会 / 2022.9.2 / ONLINE

# はじめに&目次

- 本チュートリアルはリカレンスプロット (RP) を初めて知る方向けです
- スライドは、修正ののち、研究会の Web ページに載せる予定です  
→ <https://statdata.pr.aist.go.jp/recplot/>

## ■目次

---

1. リカレンスプロットの作り方と種類
2. リカレンスプロットからの時系列復元
3. データ変換ツールとしての RP

# 1. リカレンスプロットの作り方と種類

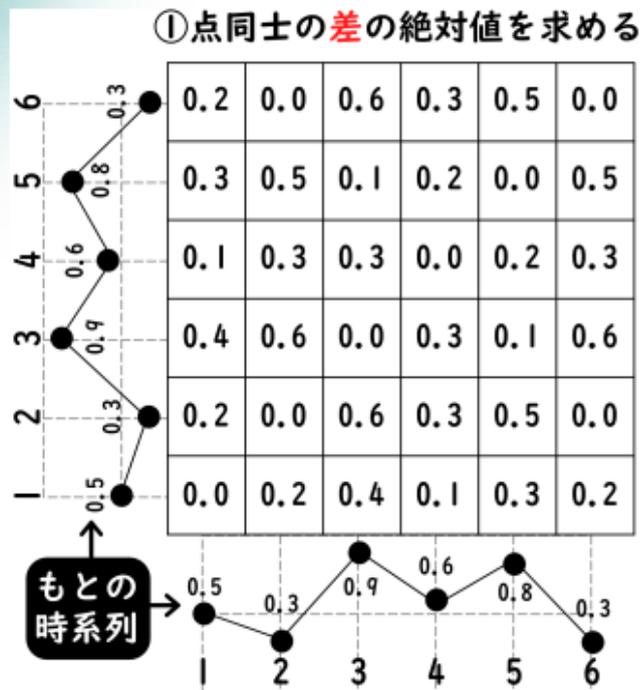
- 時系列を可視化する方法の一つ（カオス・確率過程・周期時系列）  
実は空間系列データでも構わない（蛋白配列の Contact map 等）
- 二次元のプロット

## ■ 定義

リカレンススレートを  $\varepsilon$  とし、時系列  $x$  上の点  $x(i)$  と点  $x(j)$  から、リカレンスプロット  $R$  の要素  $R(i, j)$  を次式で与える。

$$R(i, j) = \begin{cases} 1 & \text{if } \|x(i) - x(j)\| \leq \varepsilon, \\ 0 & \text{otherwise} \end{cases}$$

## ■ 具体的なプロットのしかた



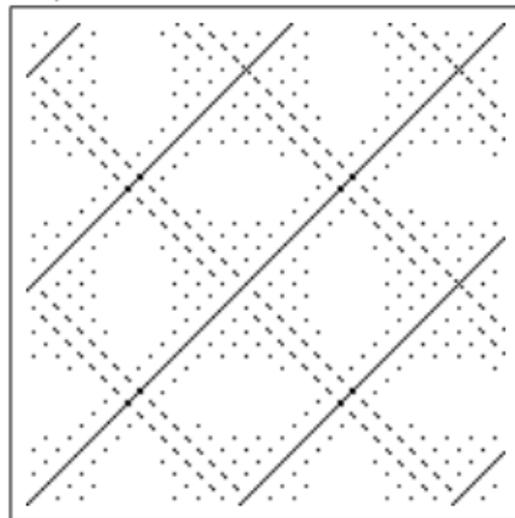
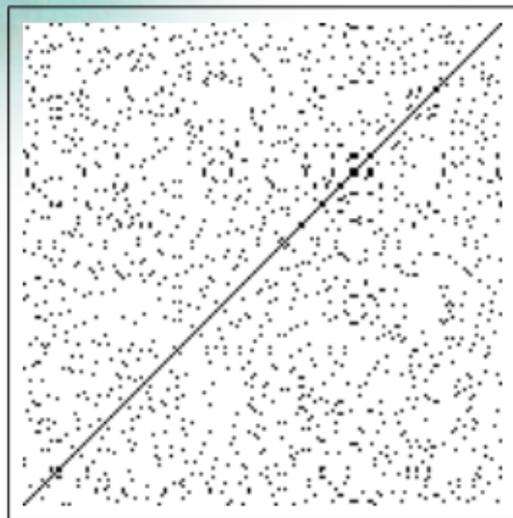
②差  
0.2  
で  
二  
値  
化

1	1	0	0	0	1
0	0	1	0	1	0
1	0	0	1	0	0
0	0	1	0	1	0
1	1	0	0	0	1
1	1	0	1	0	1

1: 黒  
0: 白

## ■ 実例：いくつかの時系列のリカレンスプロット

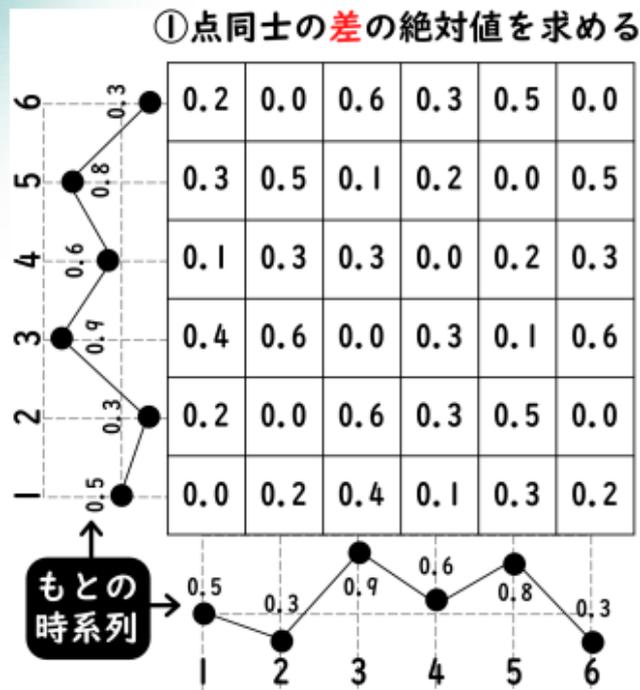
順に、一様乱数、Logistic 写像 ( $x_{n+1} = 3.8x_n(1 - x_n)$ )、複合正弦波



★ Logistic 写像（決定論カオス）では斜めの線が多め★

対称で常に斜めの線が見える（自分自身との距離はゼロ）

## ■ 復習：具体的なプロットのしかた



②差<sup>0.2</sup>で二値化

1	1	0	0	0	1
0	0	1	0	1	0
1	0	0	1	0	0
0	0	1	0	1	0
1	1	0	0	0	1
1	1	0	1	0	1

1: 黒  
0: 白

## ■ 描画時に 2 つの自由度がある

- **差**：距離尺度（点同士の距離の決め方）

$L_n$  ノルム、角度距離（位相差距離）、編集距離（点過程）、対称 KL 等要素同士の距離が決められるデータであれば何でもリカレンスプロットで描ける

- **二値化**：リカレンスレート  $\varepsilon$ （どれくらい近い点を近いとみなすか）

極端な値では情報欠損・情報論的には白黒が半々になるような値がベストだが一般には 0.01~0.1 が使われる

## ■ リカレンスプロットの利点

- 短い時系列でも利用可能：RNN 学習が難しい問題も扱える。
- 時系列の種類を問わない：非線形、非定常、点過程、確率過程
- 時系列が長ければ任意の部分を切り出して解析可能

## ■ リカレンスプロットの弱点

- 二次元なのでたくさんのメモリが必要（それに伴う時間計算量）  
→ GPU との親和性がよいので、並列化と GPU 利用で高速化
- 定数など、一部適さない時系列→逆に複雑な時系列は得意

## ■ リカレンスプロットの拡張

- 基本：1つの時系列に対してユークリッド距離を使い2値化
- Unthreshold RP：2値化しない
- Cross RP：異なる2つの時系列に対してのリカレンスプロット  
同じ時系列の異なる部分にも使える
- Joint RP：2つのリカレンスプロットのANDをとったもの

## ■ なぜ2値化するのか？

- データ量縮減
- **ビット列=記号列として扱える**

## ■ 描いたあと何をするのか？

一般には各種統計量を計算：Recurrence quantification analysis (RQA)  
(Zbilut & Webber 1992)

- 斜めの線の長さの分布、縦の線の長さの分布  
そこから計算されるエントロピー
- 斜めの白い線の長さの分布、縦の白い線の長さの分布  
そこから計算されるエントロピー
- 縦の線の部分的なリカレンスレート (Trend)、その推移

など。時系列を特徴量の組で表現する。※より一般的には 3. で。

## 2. リカレンスプロットからの時系列復元

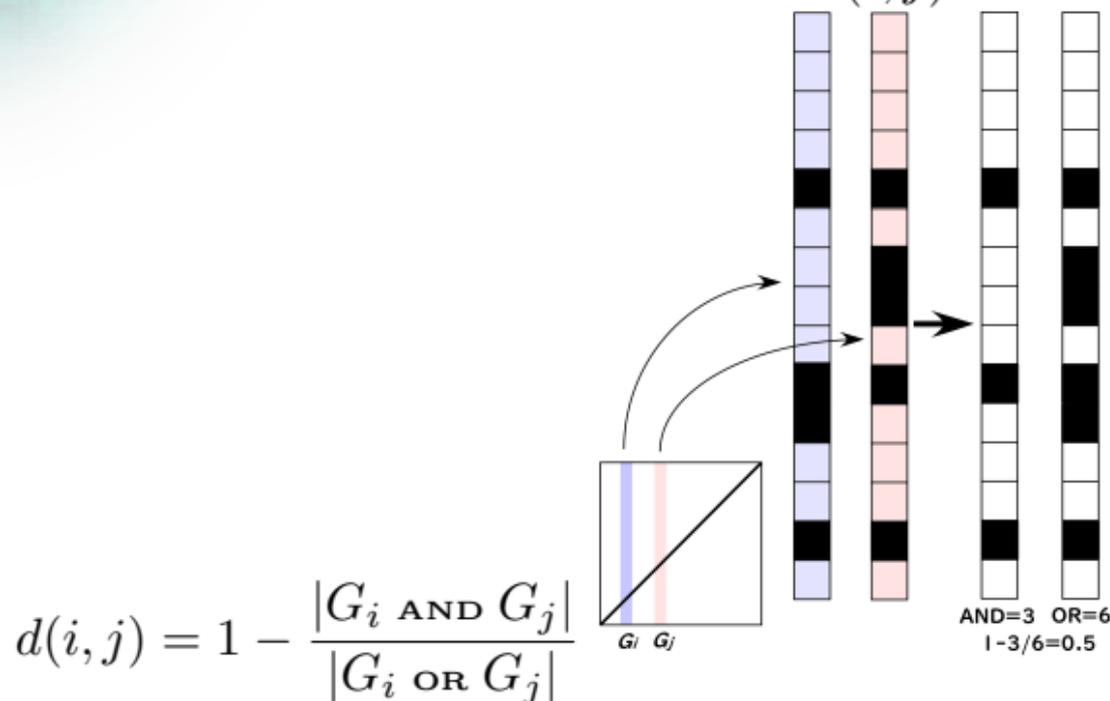
連結で十分長いリカレンスプロットのみから（定数倍を除いた）元の時系列を復元できる。おそらくリカレンスプロット最大の特長

(Hirata, Horai & Aihara (2008))

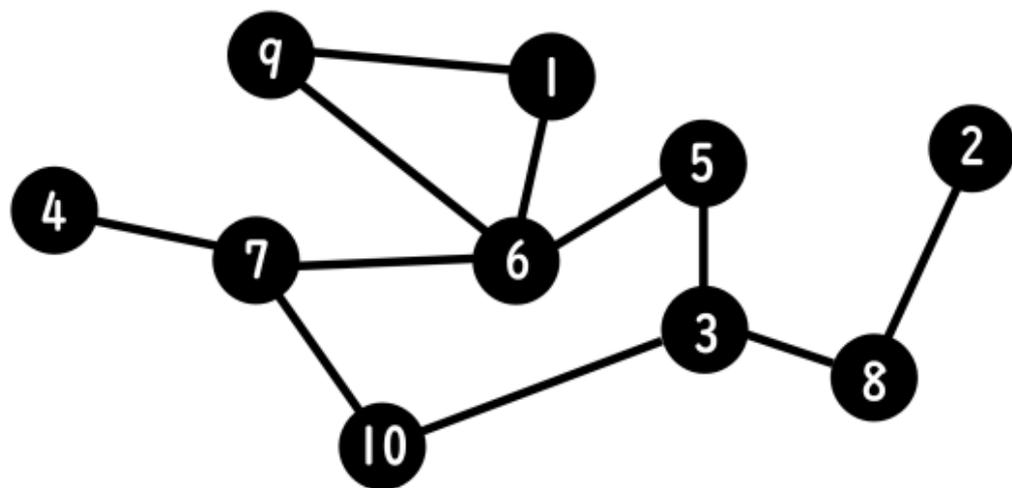
- 条件 1：時系列が十分に長いこと
- 条件 2：ネットワークとして見たとき連結になっていること

## ■ そのアルゴリズム

### 1. 隣接する点同士のパターン $G$ の近さ $d(i, j)$ を求める



2. 隣接している点同士をたどってすべての点同士の最短距離を求める (ダイクストラ法など)



3. すべての点同士の距離 (距離行列) ができたら MDS (多次元尺度法) で点の順番を決める = (定数倍を除いた) 時系列の復元が完成

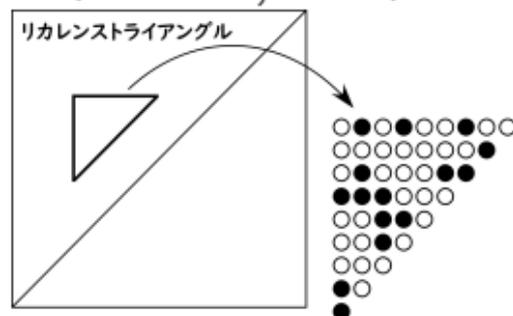
### 3. データ変換ツールとしての RP

同じリカレンスプロットに複数の解釈がある＝リカレンスプロットを経由したデータ変換に使える（複雑な情報における橋渡しツール！）

- 斜めの線と統計的解釈（RQA 周辺）
- 縦（横）の列とビット列解釈
- 対称行列としての解釈

## ■ 斜めの線と統計的解釈 (RQA 周辺)

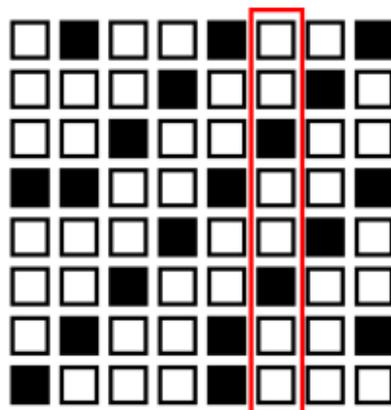
- 時系列の決定論性を示す (Eckmann,1987) : 似た値からは似た値に推移する = ひとたび黒い点が現れたらそれが続く傾向がある → RQA
- リカレンストライアングル (Hirata,2021)



長さ  $L$  のトライアングルに含まれる点のパターン数が  $L$  の増加に対して指数関数的増大以下  
→ 決定論系、指数関数的増大より速い → 確率論的系

## ■ 縦（横）の列とビット列解釈

- 時系列を単語の並び（言語）に変換（Graben,2013）
- 記号力学系



## ■ 実対称行列としての解釈

- 隣接行列：複雑ネットワーク (Zhang,2006)
- (Unthreshold RP) 行列分解：固有値分解 ( $A = X\Lambda X'$ ) など
- (Unthreshold RP) リカレンスプロット同士の加算・転置でリカレンスプロットを構成： $A + B^T$
- (Unthreshold RP) 分散共分散行列

# まとめ

1. リカレンスプロットの作り方と種類：定義・ $\epsilon$  と距離尺度の自由度・事例
2. リカレンスプロットからの時系列復元
3. データ変換ツールとしての RP：斜めの線と統計・ビット列・対称行列



リカレンスプロットは複雑な情報における橋渡しツール！

## ■ 参考文献

- 最初の論文

Eckmann *et al.*. “Recurrence plots of dynamical systems.” *Europhys. Lett.* 4 973(1987).

- 網羅的 Review

Marwan *et al.*. “Recurrence plots for the analysis of complex systems.” *Physics reports* 438.5-6 (2007): 237-329.

- リカレンスプロットから時系列を再構成する論文

Hirata *et al.*. “Reproduction of distance matrices and original time series from recurrence plots and their applications.” *The European Physical Journal Special Topics* 164.1 (2008): 13-22.