

# MACS SG12 外れ値で見る理学

教員：宮路智行(京大, 代表教員), 友枝明保(関西大), 林邦好(京都女子大)  
学生：中森聡(数学専攻M1, 発表者), 岡本和也(早稲田大D3), 竹田航太(数学専攻D3),  
松田凌(数学専攻D3), 下村顕士(物理専攻D1), 岸広登(数学専攻M1)

## 概要

時系列データの変化点を検出する方法として、 $\left\{ \begin{array}{l} \text{統計モデルに基づく外れ値検出手法} \\ \text{力学系理論に基づく Early-warning signal の手法} \end{array} \right\}$  を比較する。

## 背景・目的

- 外れ値：標本集団の中で他のデータから著しく外れたデータ
- 外れ値を検知するという事は、何らかの意味で正常な大部分とそうでない一部を区別するという事。逆に言えば、データにパターンを見出す事。
- 外れ値から理学へ：外れ値にまつわる数理的・統計的手法を理学の問題や経済・交通流などの実データに応用したい。
- 理学から外れ値へ：力学系理論や統計物理といった理学の視点から外れ値検知の理論を検討したい。

## 時系列解析

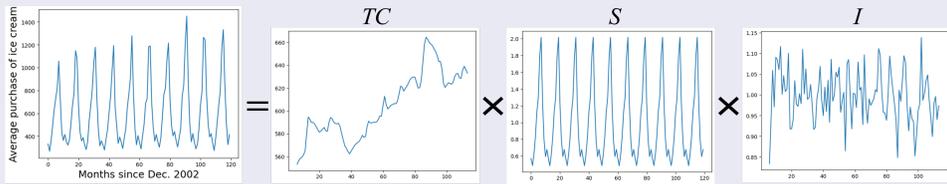
Ref. 島田直希『時系列解析』共立出版 (2019)

### 時系列データの分解

- 時系列データ  $y_t$  を次の3つの成分に分解できると仮定。
  - 傾向変動(トレンド)  $TC$ : 増加または減少傾向を持続する長期的な変動
  - 季節変動  $S$ : 循環を繰り返すような周期的変動
  - 不規則変動  $I$ : 上記以外の不規則かつ短期間の変動

$$y_t = TC_t \times S_t \times I_t$$

- 不規則変動には突発的に生じた外れ値の効果が入っている！
- 例：アイスクリームの月間平均購入額[円/世帯]の時系列データ  
データの出自：<https://okumuralab.org/okumura/stat/160118.html>

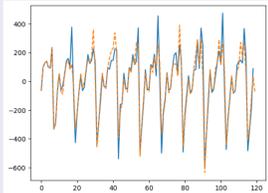


### ARモデル

- 観測値を確率過程に基づいて直接モデル化
- ある時点の観測値  $y_t$  を過去の観測値  $y_{t-p}, \dots, y_{t-1}$  で説明

$$y_t = c + \phi_1 y_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{W.N.}(\sigma^2)$$

- $\epsilon_t$  はホワイトノイズを仮定
- 例：アイスクリームの月間平均購入額[円/世帯]の時系列データ  
データの1次階差系列をAR(11)モデルで予測する
  - 青実線: 1次階差系列
  - オレンジ破線: AR(11)モデルによる予測
  - 予測には最初の11個の観測値を用いる



## 外れ値検出

### 教師あり学習に基づく外れ値検出

説明変数  $X_1, \dots, X_n$  に対し、  
応答変数  $Y = 0$  (正常) or  $1$  (異常) を教師として学習を行う。

### 統計モデルに基づく外れ値検出

正常なデータの従う確率分布を仮定し、  
その分布の下で出現確率が極めて低いデータを異常と見なす。

### 影響関数に基づく外れ値検出

観測値  $x_1, \dots, x_n$  に基づいて何らかの推定値  $\hat{\theta}(x_1, \dots, x_n)$  を定めたとき、各  $x_i$  が  $\hat{\theta}$  に与える「影響」を評価する。

## Early-warning signal (EWS)

Ref. Scheffer, M et al. Early-warning signals for critical transitions Nature 461,53-59(2009)

- ある安定状態から別の安定状態へ移行する現象(critical transition)の前兆となる指標のこと。
  - 摂動からの回復率の低下
  - 自己相関の増加
  - 分散の増加
- 例: 大規模な気候変動
  - Greenhouse-icehouse transition
  - 氷河期の終わり
  - アフリカでの砂漠化
- 例: 生理機能の急激な変化
  - 喘息とてんかん

## 変化点検出とEWSの比較

Mayモデルに対して変化点検出とEWSを適用

- Mayモデル

Ref. Robert M. May Thresholds and breakpoints in ecosystems with a multiplicity of stable states Nature 269, pages471- 477 (1977)

$$\dot{x} = -F'(x), \quad F(x) = -\frac{1}{2}x^2 + \frac{1}{3k}x^3 + c(x - \arctan x), \quad \dot{c} = \epsilon$$

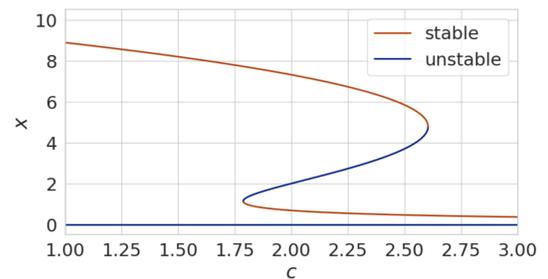
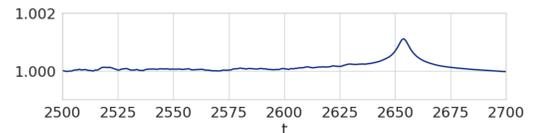


Figure: 平衡点の分岐図

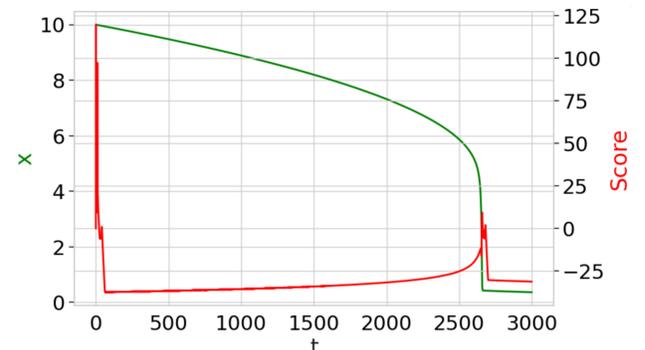
- EWS
  - 自己相関係数の増加



- 変化点検出手法

- 統計モデルとしてAR(1)を採用し、予測値と観測値とのずれをもとに変化点を検出する。(ChangeFinder)

Ref. 島田直希『時系列解析』共立出版 (2019)



## 今後の展望

- 交通流における一様流と渋滞流との遷移の予兆を数理モデルおよび実データから検出できるか? data-drivenで変化の予兆を検出できるか?
- 力学系の分岐と変化点の関係は?