Utilizing Topological Data Analysis to Detect Periodicity

Elizabeth Munch

University at Albany - SUNY :: Department of Mathematics & Statistics

Oct 2, 2016
Time series in biology

Mitosis

Yeast gene expression
Deckard et al., Bioinformatics 2013

Neuron Spike Trains

ECG
Goldberg et al. 2000
Our definition of time series

Definition
A time series is a function

\[ f : \mathbb{R}_{\geq 0} \rightarrow D \]

for some topological space \( D \).
Our definition of time series

Definition

A time series is a function

\[ f : \mathbb{R}_{\geq 0} \rightarrow D \]

for some topological space \( D \).

Choice for \( D \)

- \( \mathbb{R} \) - Classical time series analysis
- \( \mathbb{R}^{m \times n} \) - \( \mathbb{R} \)-valued \( m \times n \) matrices (movies)
- Pers - Persistence diagram valued time series (vineyards)
Commonly used tools

$\mathbb{R}$-valued TS $\xrightarrow{\text{Takens Embedding}}$ Pers $\xrightarrow{\text{Persistence of persistence}}$ Pers-valued TS

$\mathbb{R}^{m \times n}$-valued TS $\xrightarrow{\text{Sub/Suplevel-set persistence}}$ Pers-valued TS
Common questions

Classification/Clustering
  ▶ Is this signal Type A or Type B?
Common questions

- Classification/Clustering
  - Is this signal Type A or Type B?

- Periodicity
  - Is this signal exhibiting periodic behavior?
Common questions

- Classification/Clustering
  - Is this signal Type A or Type B?

- Periodicity
  - Is this signal exhibiting periodic behavior?

- Forecasting
  - Given this previous signal, what do we expect to have happen next?
Common questions

- Classification/Clustering
  - Is this signal Type A or Type B?
- Periodicity
  - Is this signal exhibiting periodic behavior?
- Forecasting
  - Given this previous signal, what do we expect to have happen next?
- Segmentation
  - Which pieces of this signal come from similar systems?
Common questions

- **Classification/Clustering**
  - Is this signal Type A or Type B?

- **Periodicity**
  - Is this signal exhibiting periodic behavior?

- **Forecasting**
  - Given this previous signal, what do we expect to have happen next?

- **Segmentation**
  - Which pieces of this signal come from similar systems?
Idea:
Persistent homology and other TDA tools can be used to improve time series analysis.
Idea:
Persistent homology and other TDA tools can be used to improve time series analysis.

This talk:
- Mechanical engineering
  - Firas Khasawneh
  - Jose Perea
- Atmospheric science
  - Bill Dong
  - Kristen Corbosiero
  - Jason Dunion
  - Ryan Torn
1. Classification and Machining Dynamics

2. Periodicity and Hurricanes
1 Classification and Machining Dynamics

2 Periodicity and Hurricanes
Machining Dynamics

Images courtesy Firas Khasawneh, SUNYIT; and Boeing.
Deterministic model:

\[ \ddot{y} + 2\zeta\dot{y} + y = K\rho^{\alpha-1}(1 + y(t - \tau) - y(t))^\alpha \]

- Left side: standard linear oscillator
- Right side: input based on cutting forces

Chatter

\begin{figure}[h]
\centering
\includegraphics[width=\columnwidth]{chatter}
\caption{Chatter}
\end{figure}
Takens embedding

Definition

Given a time series $X(t)$, the Takens embedding is

$$\psi^m_{\eta} : t \mapsto (X(t), X(t + \eta), \ldots, X(t + (m - 1)\eta)).$$
Persistent Homology of Point Cloud

Expanding Discs

-1.0 -0.5 0.0 0.5 1.0
-1.0 -0.5 0.0 0.5 1.0
Noise resilience

Original Signals

TSA with TDA

Liz Munch (UAlbany)

Oct 2 ACM-BCB 13 / 30
Noise resilience

Original Signals

Delay Embedding

Persistence Diagrams
Comparing signals using persistence

![Signal, [0.9, 0.07]](image1)

![Signal, [1.42, 0.05]](image2)

![Signal, [1.48, 0.25]](image3)

![Takens Embedding, [0.9, 0.07]](image4)

![Takens Embedding, [1.42, 0.05]](image5)

![Takens Embedding, [1.48, 0.25]](image6)

![Persistence Diagram, [0.9, 0.07]](image7)

![Persistence Diagram, [1.42, 0.05]](image8)

![Persistence Diagram, [1.48, 0.25]](image9)
Comparing signals using persistence

- Chatter
- Chatter free

$\frac{\Omega}{\omega_n}$

Signals and Takens Embedding

Persistence Diagrams

Liz Munch (UAlbany)
Overview

\[ \mathbb{R} \text{-valued TS} \xrightarrow{\text{Takens Embedding}} \operatorname{Pers} \]

\[ \mathbb{R}^{m \times n} \text{-valued TS} \xrightarrow{\text{Sub/Suplevel-set persistence}} \operatorname{Pers} \text{-valued TS} \]

Liz Munch (UAlbany) TSA with TDA Oct 2 ACM-BCB
Overview

\[ \mathbb{R}^m \times n \text{-valued TS} \rightarrow \mathbb{R}\text{-valued TS} \rightarrow \text{Pers} \]

\[ \text{Takens Embedding} \rightarrow \text{Pers}\text{-valued TS} \]

\[ \text{Max-persistence} \rightarrow \text{Persistence of persistence} \]

\[ \text{Sub/Suplevel-set persistence} \rightarrow \text{Pers} \]
Differentiation by Max Persistence

- Signal, [0.9, 0.07]
- Signal, [1.42, 0.05]
- Signal, [1.48, 0.25]

- Takens Embedding, [0.9, 0.07]
- Takens Embedding, [1.42, 0.05]
- Takens Embedding, [1.48, 0.25]

- Persistence Diagram, [0.9, 0.07]
- Persistence Diagram, [1.42, 0.05]
- Persistence Diagram, [1.48, 0.25]
Turning Model

Maximum Persistence, No Noise

\[ Y(t) \]

\[ Y(t + 2.13) \]

\[ Y(t + 1.62) \]

\[ Y(t + 1.56) \]

\[ \frac{\Omega}{\omega_n} \]

\[ b \]

\[ \frac{\Omega}{\omega_n} \]
Adcock et al. Coordinates

Diagrams 0 and 1-dimensional of the form \{ (x_i, y_i) \}

- \[ \sum x_i (y_i - x_i) \]
- \[ \sum (y_{\text{max}} - y_i)(y_i - x_i) \]
- \[ \sum x_i^2 (y_i - x_i)^4 \]
- \[ \sum (y_{\text{max}} - y_i)^2 (y_i - x_i)^4 \]
- \[ \max \{ (y_i - x_i) \} \]
Machine Learning

Adcock et al. Coordinates

Diagrams 0 and 1-dimensional of the form \( \{(x_i, y_i)\} \)

- \( \sum x_i(y_i - x_i) \)
- \( \sum(y_{\max} - y_i)(y_i - x_i) \)
- \( \sum x_i^2(y_i - x_i)^4 \)
- \( \sum(y_{\max} - y_i)^2(y_i - x_i)^4 \)
- \( \max\{(y_i - x_i)\} \)

Results
(Khasawneh, M, Perea)

- Theoretical stability boundary for training
- Standard logistic classifier
- 97% accuracy
1 Classification and Machining Dynamics

2 Periodicity and Hurricanes
1. Classification and Machining Dynamics

2. Periodicity and Hurricanes
Hurricane Felix

9/2/2007 06:15
Pcolor Hurricane Image
Diurnal cycle

3 hour difference

- $N(t)$ is IR matrix at time $t$
- $N(t) - N(t - 3 \text{ hrs})$
Diurnal cycle

3 hour difference
- $N(t)$ is IR matrix at time $t$
- $N(t) - N(t - 3 \text{ hrs})$

Diurnal cycle
- Sunset: cold ring, “diurnal pulse”
- Starts with radius $\leq 150\text{km}$, spreads outward
- Warm ring forms behind this pulse and spreads outward

Sublevel Set Persistence
Sublevel Set Persistence
Sublevel Set Persistence
Sublevel Set Persistence

1D Persistence - Noise

Birth Time

Death Time

Liz Munch (UAlbany)
Why the obvious thing doesn’t work
Plan B

**Definition**
Let $K_{m \times n} = K$ be the $m \times n$ grid cubical complex.

**Definition**
Given $M \in \mathbb{R}^{m \times n}$, let
- $M : K \to \mathbb{R}$
- $M^{\mu} \subset K$ with function value $\geq \mu$.
- $S : K \to \mathbb{R}$ defined by $S(\sigma) = d(\sigma, M^{\mu})$ for $\dim(\sigma) = 2$
Plan B

Definition

Let $K_{m \times n} = K$ be the $m \times n$ grid cubical complex.

Definition

Given $M \in \mathbb{R}^{m \times n}$, let

- $M : K \rightarrow \mathbb{R}$
- $M^\mu \subset K$ with function value $\geq \mu$.
- $S : K \rightarrow \mathbb{R}$ defined by $S(\sigma) = d(\sigma, M^\mu)$ for $\dim(\sigma) = 2$
Plan B

**Definition**
Let $K_{m \times n} = K$ be the $m \times n$ grid cubical complex.

**Definition**
Given $M \in \mathbb{R}^{m \times n}$, let

- $M : K \rightarrow \mathbb{R}$
- $M^\mu \subset K$ with function value $\geq \mu$.
- $S : K \rightarrow \mathbb{R}$ defined by $S(\sigma) = d(\sigma, M^\mu)$ for $\dim(\sigma) = 2$
Resulting persistence diagrams


Pcolor Hurricane Image

Perseus diagram

pixel values

birth function value

death function value

Liz Munch (UAlbany)

TSA with TDA

Oct 2 ACM-BCB
Overview

- \( \mathbb{R} \text{-valued TS} \) → \( \text{Takens Embedding} \) → \( \text{Pers} \)
  - Persistence of persistence

- \( \mathbb{R}^{m \times n} \text{-valued TS} \) → \( \text{Sub/Suplevel-set persistence} \) → \( \text{Pers-valued TS} \)

Liz Munch (UAlbany)
TSA with TDA
Oct 2 ACM-BCB 27 / 30
Fourier spectrum of threshold

threshold = 35

max persistence


time
Fourier spectrum of threshold
Fourier spectrum of threshold

Threshold = 35

Maximum persistence

Time:
- 9/12:15 PM
- 9/21:15 AM
- 9/21:15 AM
- 9/2/15 PM
- 9/3:15 AM
- 9/3:15 AM
- 9/4:15 PM
- 9/4:15 PM
- 9/5/15 PM
- 9/5/15 AM

Coefficient

Frequency:
- -0.4
- -0.2
- 0.0
- 0.2
- 0.4

Period approx 23 hrs
Fourier spectrum of threshold

Results

- 23 hour day?
General tools for TSA with TDA

- **Takens embedding → persistence**
  - Real-valued time series
  - Can do classification, segmentation using persistence diagrams

- Structures and behaviors that are easy to tease out
  - Circles/holes
  - Periodicity

Liz Munch (UAlbany)
General tools for TSA with TDA

- **Takens embedding → persistence**
  - Real-valued time series
  - Can do classification, segmentation using persistence diagrams

- **Image → sublevelset persistence**
  - Get a time series of persistence diagrams
  - Pick out information from each diagram (max pers) to use standard TSA methods
  - Analyze speed
  - Persistence of persistence
    (Kramar, Levanger, et al. 2015 arXiv:1505.06168)
General tools for TSA with TDA

- **Takens embedding → persistence**
  - Real-valued time series
  - Can do classification, segmentation using persistence diagrams

- **Image → sublevelset persistence**
  - Get a time series of persistence diagrams
  - Pick out information from each diagram (max pers) to use standard TSA methods
  - Analyze speed
  - Persistence of persistence
  (Kramar, Levanger, et al. 2015 arXiv:1505.06168)

- **Structures and behaviors that are easy to tease out**
  - Circles/holes
  - Periodicty
Thank you!

**Hurricanes**
Kristen Corbosiero (Albany)
Jason Dunion (Albany)
Bill Dong (Guilderland High School)
Ryan Torn (Albany)

**Machining Dynamics**
Firas Khasawneh (SUNY Poly)
Jose Perea (MSU)

FK and EM. *Chatter detection in turning using persistent homology*. Mechanical Systems and Signals Processing, 2016.
elizabethmunch.com
emunch@albany.edu