Detecting Causality between Time Series

based on Convergent Cross-Mapping

Who? Daniel Wäber

Data Scientist at GameDuell

When? September 18, 2014

Background Info

- normally I'm challenging algorithms just behind you
- working and have done my master thesis here

Task: find out more from our server monitoring time series

- identify interactions inside the systems
- detect uncommon behaviour

- now I'm here presenting some of my work

Table of contents

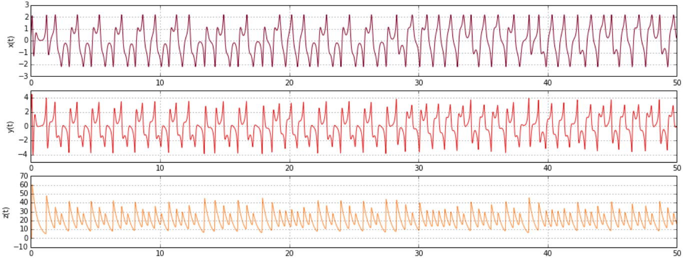
Problem Statement

How to detect coupled and causing components inside systems by only looking at time series?

Problem Statement

How to detect coupled and causing components inside systems by only looking at time series?

Example A system generated three time series X, Y and Z, what interactions can we identify?



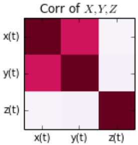
Problem Statement

How to detect coupled and causing components inside systems by only looking at time series?

Example A system generated three time series X, Y and Z, what interactions can we identify?

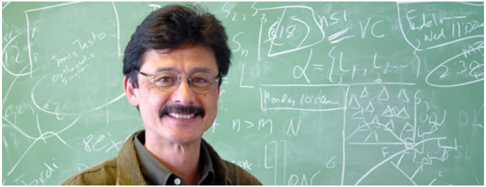
Correlation solves it!?

- only detects linear dependencies
- no directional information



We can do much better!

How - Convergent Cross-Mapping!



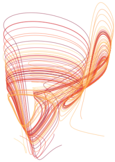
An approach to solve this problem

- was developed by George Sugihara et. al in 2012 at the Scripps Institution of Oceanography [1]
- its called Convergent Cross Mapping (CCM)
- it applies dynamic system theory to identify interactions between time series

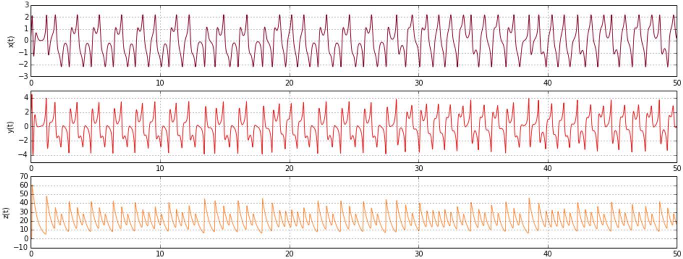
A Dynamic System

Actually the time series are generated by a dynamic system, the Yu-Wang attractor.

$$\begin{aligned}\dot{x} &= \alpha(y-x) && (1)\\ \dot{y} &= \beta x - \gamma x z && (2)\\ \dot{z} &= e^{xy} - \delta z && (3)\end{aligned}$$



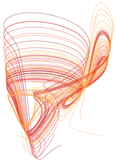
Example



A Dynamic System

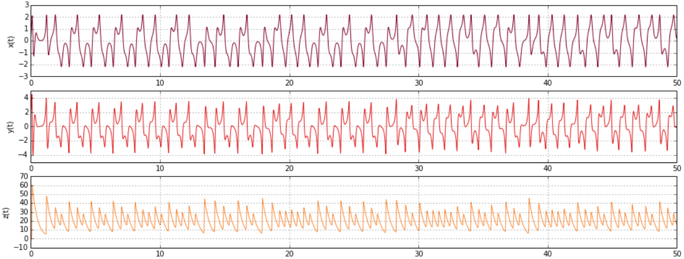
Actually the time series are generated by a dynamic system, the Yu-Wang attractor.

$$\dot{x} = \alpha(y - x) \quad (1)$$
$$\dot{y} = \beta x - \gamma x z \quad (2)$$
$$\dot{z} = e^{xy} - \delta z \quad (3)$$



For our point of view, we only know the generated data and still want to reveal the interactions!

Example



Basic Idea

How to detect interactions from time series data?

Basic Idea

How to detect interactions from time series data?

- using a single time series X, we can build a higher dimensional representation, the shadow manifold, that captures the behaviour of the system

Basic Idea

How to detect interactions from time series data?

- using a single time series X, we can build a higher dimensional representation, the shadow manifold, that captures the behaviour of the system
- map representation of different variables X and Y onto each other, so we can test them for interactions

Shadow Manifold

How to detect interactions from time series data?

- soundness of approach based on Taken's Theorem [2]
- "one can capture the state of a dynamic system with a fixed number E of observation of a single variable"

Shadow Manifold

How to detect interactions from time series data?

- soundness of approach based on Taken's Theorem [2]
- "one can capture the state of a dynamic system with a fixed number E of observation of a single variable"
- we call this higher dimensional representation of a single variable X the **shadow manifold** M_X

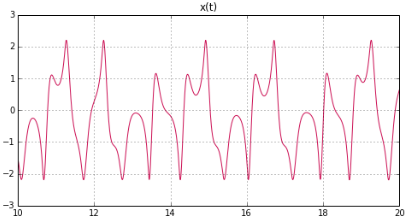
$$M_X : x(t) = \langle x_t, x_{t-\tau}, \cdots, x_{t-(E-1)\tau} \rangle$$

Shadow Manifold

How to detect interactions from time series data?

- soundness of approach based on Taken's Theorem [2]
- "one can capture the state of a dynamic system with a fixed number E of observation of a single variable"
- we call this higher dimensional representation of a single variable X the **shadow manifold** M_X

$$M_X : x(t) = \langle x_t, x_{t-\tau}, \cdots, x_{t-(E-1)\tau} \rangle$$

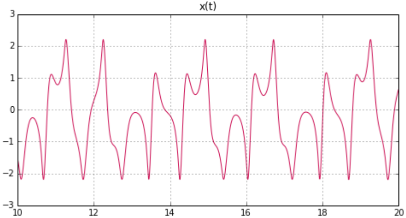


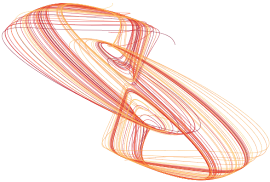
Shadow Manifold

How to detect interactions from time series data?

- soundness of approach based on Taken's Theorem [2]
- "one can capture the state of a dynamic system with a fixed number E of observation of a single variable"
- we call this higher dimensional representation of a single variable X the **shadow manifold** M_X

$$M_X : x(t) = \langle x_t, x_{t-\tau}, \cdots, x_{t-(E-1)\tau} \rangle$$





Cross-Mapping

How to detect interactions from time series data?

- next step: map the high dimensional representation
- build estimation of Y using the information of M_X

$$\hat{Y}|M_X : \hat{y}(t) = \sum_{i=1}^{E+1} w_i(t) \cdot y_{n_i(t)}$$

$w_i(t) = \frac{u_i(t)}{\sum_{j=1}^{E+1} u_j(t)}$, $u_i(t) = \exp(-\frac{d(x(t),n_i(t))}{d(x(t),n_1(t))})$,
$n_i(t)$ index of i-th nearest point to $x(t)$ on M_X

Cross-Mapping

How to detect interactions from time series data?

- next step: map the high dimensional representation
- build estimation of Y using the information of M_X

$$\hat{Y}|M_X : \hat{y}(t) = \sum_{i=1}^{E+1} w_i(t) \cdot y_{n_i(t)}$$

$$w_i(t) = \frac{u_i(t)}{\sum_{j=1}^{E+1} u_j(t)},\ u_i(t) = \exp(-\frac{d(x(t), n_i(t))}{d(x(t), n_1(t))}),$$

$n_i(t)$ index of i-th nearest point to $x(t)$ on M_X

- weights $w_i(t)$ and index $n_i(t)$ only depending on M_X

Cross-mapping applies the structure reveled by one variable on another variable

Convergence of the Cross-Maps

How to detect interactions from time series data?

- we say that X cross-maps Y, iff

$$\lim_{L \to \infty} \hat{Y} | M_X \underset{\ell^2}{\to} Y$$

For a dynamic system one can show:

Convergence of the Cross-Maps

How to detect interactions from time series data?

- we say that X cross-maps Y, iff

$$\lim_{L\to\infty} \hat{Y}|M_X \underset{\ell^2}{\to} Y$$

For a dynamic system one can show:

- iff X cross-maps Y and Y cross-maps X,
  the state of one variable determines the other

Convergence of the Cross-Maps

How to detect interactions from time series data?

- we say that X cross-maps Y, iff

$$\lim_{L\to\infty} \hat{Y}|M_X \underset{\ell^2}{\to} Y$$

For a dynamic system one can show:

- iff X cross-maps Y and Y cross-maps X,
  the state of one variable determines the other

$\Rightarrow$ X and Y are **coupled**

Convergence of the Cross-Maps

How to detect interactions from time series data?

- we say that X cross-maps Y, iff

$$\lim_{L\to\infty} \hat{Y}|M_X \underset{\ell^2}{\to} Y$$

For a dynamic system one can show:

- iff X cross-maps Y and Y cross-maps X, the state of one variable determines the other

$\Rightarrow$ X and Y are **coupled**

- iff only X cross-maps Y, but not the other way round, only the state of X determines the state of Y

Convergence of the Cross-Maps

How to detect interactions from time series data?

- we say that X cross-maps Y, iff

$$\lim_{L\to\infty} \hat{Y}|M_X \underset{\ell^2}{\to} Y$$

For a dynamic system one can show:

- iff X cross-maps Y and Y cross-maps X,
  the state of one variable determines the other

$\Rightarrow$ X and Y are **coupled**

- iff only X cross-maps Y, but not the other way round,
  only the state of X determines the state of Y

$\Rightarrow$ X **causes** Y

Inferring Dependencies

We can detect interactions from time series data:

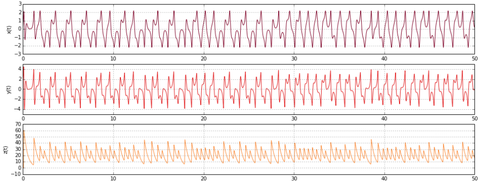
- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour

Inferring Dependencies

We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour

Example

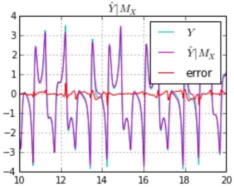


$$\begin{aligned}\dot{x} &= \alpha(y - x)\\ \dot{y} &= \beta x - \gamma x z\\ \dot{z} &= e^{xy} - \delta z\end{aligned}$$

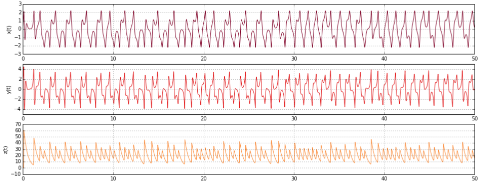
Inferring Dependencies

We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour



Example

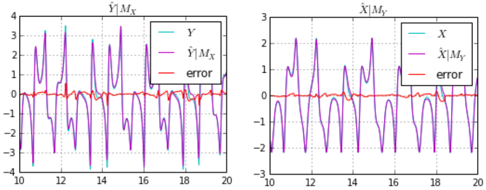


$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

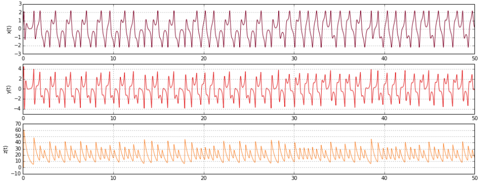
Inferring Dependencies

We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour



Example

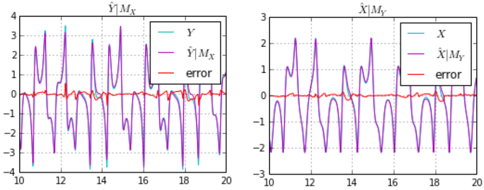


$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

Inferring Dependencies

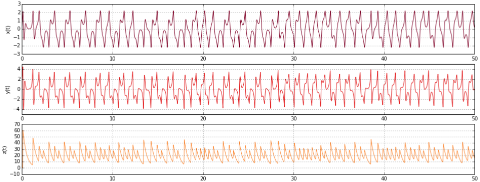
We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour



X and Y are coupled

Example

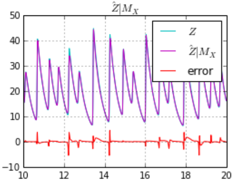


$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

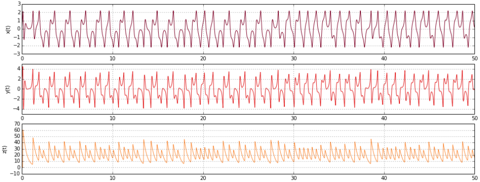
Inferring Dependencies

We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour



Example

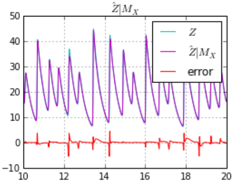


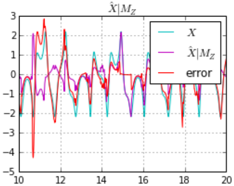
$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

Inferring Dependencies

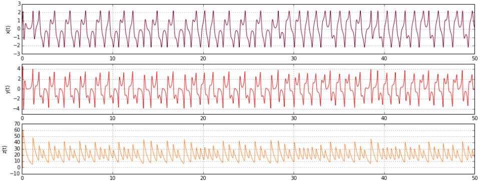
We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour





Example

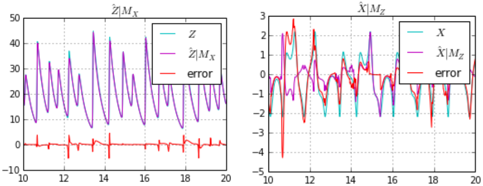


$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

Inferring Dependencies

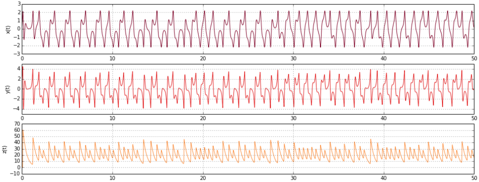
We can detect interactions from time series data:

- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour



X causes Z

Example



$$\begin{aligned} \dot{x} &= \alpha(y - x) \\ \dot{y} &= \beta x - \gamma x z \\ \dot{z} &= e^{xy} - \delta z \end{aligned}$$

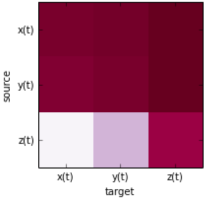
Inferring Dependencies

We can detect interactions from time series data:

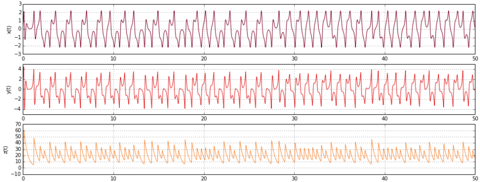
- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour

Calculating this for all combinations

- we get a cross-mapping fit matrix
- revealing the variable determinations



Example



$$\begin{aligned} \dot{x} &= \alpha(y-x) \\ \dot{y} &= \beta x - \gamma x z \\ \dot{z} &= e^{xy} - \delta z \end{aligned}$$

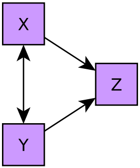
Inferring Dependencies

We can detect interactions from time series data:

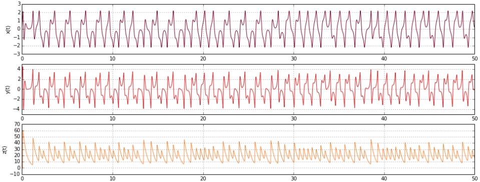
- build cross-mapping $\hat{Y}|M_X$ and $\hat{X}|M_Y$
- test for there convergence behaviour

Calculating this for all combinations

- we get a cross-mapping fit matrix
- revealing the variable determinations
- can be visualized as directed graph



Example

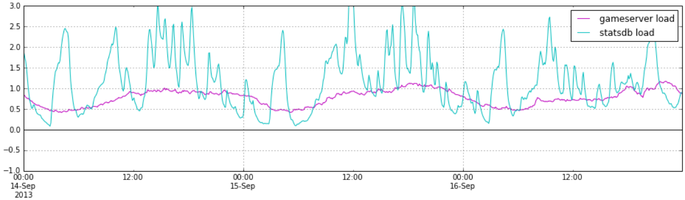


$$
\begin{aligned}
\dot{x} &= \alpha(y - x) \\
\dot{y} &= \beta x - \gamma x z \\
\dot{z} &= e^{xy} - \delta z
\end{aligned}
$$

Manifolding real world data

now it's time to apply CCM to real-world time series

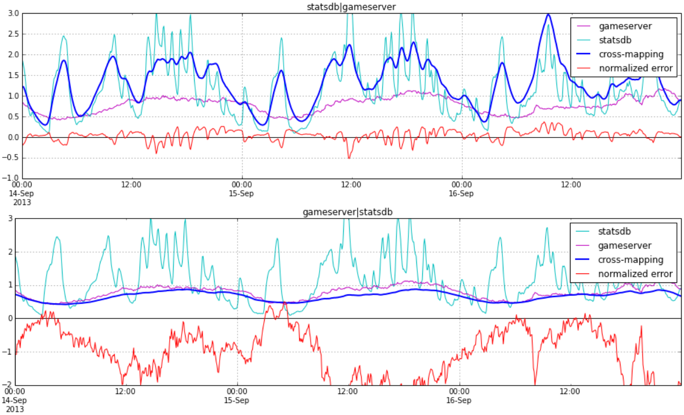
- here are two monitoring time series



Manifolding real world data

now it's time to apply CCM to real-world time series

- calculating the cross-mapping

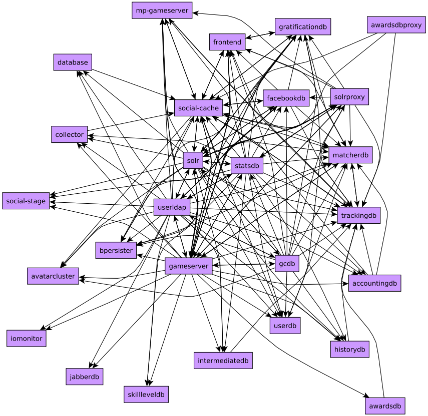


reveals that gameserver load determines statsdb load

Manifolding real world data

now it’s time to apply CCM to real-world time series

- graph showing causing and coupled components



Outlier Detection with CCM

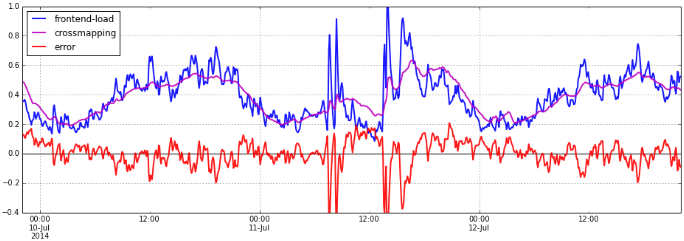
- we assume that platform with monitoring time series can be modeled as (perturbed) dynamic system
- if assumption is true,

$$\lim_{L \to \infty} \hat{X} | M_X \to X + \text{noise}$$

Outlier Detection with CCM

- we assume that platform with monitoring time series can be modeled as (perturbed) dynamic system
- if assumption is true,

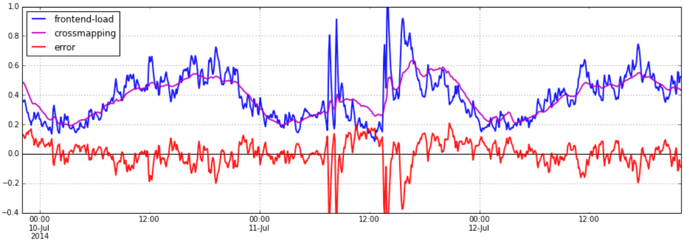
$$\lim_{L\to\infty} \hat{X}|M_X \to X + \text{noise}$$



Outlier Detection with CCM

- we assume that platform with monitoring time series can be modeled as (perturbed) dynamic system
- if assumption is true,

$$\lim_{L \to \infty} \hat{X}|M_X \to X + \text{noise}$$

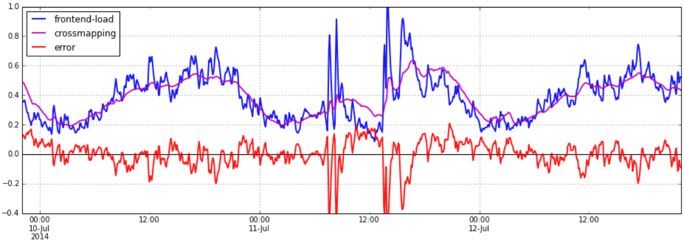


- generally this model seems to fit, but not for uncommon behavior

Outlier Detection with CCM

- we assume that platform with monitoring time series can be modeled as (perturbed) dynamic system
- if assumption is true,

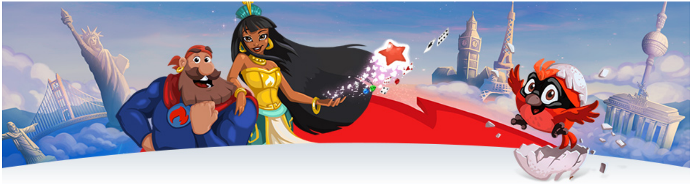
$$\lim_{L \to \infty} \hat{X} | M_X \to X + \text{noise}$$



⇒ so we can detect outliers with this

- even works for nonlinear systems

Q&A



- Thank you for listening
  - further details in my soon to be published thesis
- Input and discussion appreciated

  Just join me for a beer afterwards!

Further Reading

George Sugihara et. al
Detecting Causality in Complex Ecosystems
Science Express

F. Takens
Detecting strange attractors in turbulence
Springer-Verlag

Taken's Theorem

Appendix

Theorem *Given a dynamic system with a strange attractor A of boxcounting dimensionality d_A, one can build a reconstruction of A with $E = 2d_A + 1$ observations of just a single generic variable.*

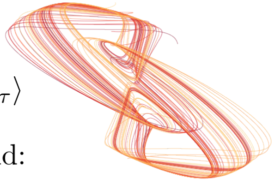
boxcounting $d_A = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log 1/\varepsilon}$,
with $N(\varepsilon)$ # ε-sized boxes needed to cover A

reconstruction $\exists$ one to one mapping, that maps different points on A to different ones on the reconstruction

generic conditions for variable
(# low period orbits, eigenvalues)

Shadow Manifolds

Appendix



$$M_X : x(t) = \langle x_t, x_{t-\tau}, \cdots, x_{t-2d_A\tau} \rangle$$

For generic variables of a manifold:

- the shadow maps one to one to the original manifold
- two shadows of a system map one to one
- the shadow manifold preserves the local structure

Cross-Mapping Details

Appendix

$\hat{Y}|M_X$: mapping the local structure of M_X to Y

Cross-Mapping Details

Appendix

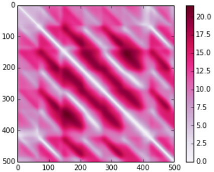
$\hat{Y}|M_X$: mapping the local structure of M_X to Y

- using distances of points on M_X

$$d^t_{t'} = d(x(t), x(t')) = \sqrt{\sum_{i=0}^{E} (x_{t-i\tau} - x_{t'-i\tau})^2}$$

Cross-Mapping Details

Appendix



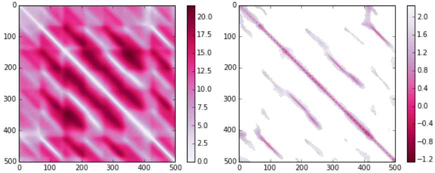
$\hat{Y}|M_X$: mapping the local structure of M_X to Y

- using distances of points on M_X

$$d^t_{t'} = d(x(t), x(t')) = \sqrt{\sum_{i=0}^{E}(x_{t-i\tau} - x_{t'-i\tau})^2}$$

Cross-Mapping Details

Appendix



$\hat{Y}|M_X$: mapping the local structure of M_X to Y

- using distances of points on M_X

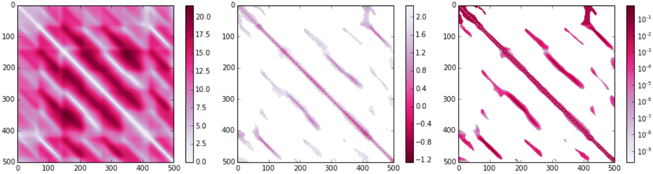
  $$d^t_{t'} = d(x(t), x(t')) = \sqrt{\sum_{i=0}^{E} (x_{t-i\tau} - x_{t'-i\tau})^2}$$

- for each t get index for $E+1$ nearest points of $x(t)$

  $n_i(t)$ for $i = 1, \cdots, E+1$

Cross-Mapping Details

Appendix



$\hat{Y}|M_X$: mapping the local structure of M_X to Y

- using distances of points on M_X

$$d^t_{t'} = d(x(t), x(t')) = \sqrt{\sum_{i=0}^{E}(x_{t-i\tau} - x_{t'-i\tau})^2}$$

- for each t get index for $E+1$ nearest points of $x(t)$

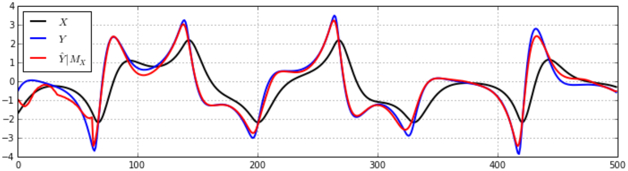
  $n_i(t)$ for $i = 1, \cdots, E+1$

- weights from distances inside the near environment

$$w_i(t) = \frac{u_i(t)}{\sum u_j(t)}, \text{ with } u_i(t) = \exp\left(-\frac{d^t_{n_i(t)}}{d^t_{n_1(t)}}\right)$$

Cross-Mapping Details

Appendix



$\hat{Y}|M_X$: mapping the local structure of M_X to Y

- using distances of points on M_X

  $d^t_{t'} = d(x(t), x(t')) = \sqrt{\sum_{i=0}^{E}(x_{t-i\tau} - x_{t'-i\tau})^2}$
- for each t get index for $E+1$ nearest points of $x(t)$

  $n_i(t)$ for $i = 1, \cdots, E+1$
- weights from distances inside the near environment

  $w_i(t) = \frac{u_i(t)}{\sum u_j(t)}$, with $u_i(t) = \exp(-\frac{d^t_{n_i(t)}}{d^t_{n_1(t)}})$
- estimate $\hat{Y}|M_X$ by applying weights to Y

  $\hat{Y}|M_X : \hat{y}(t) = \sum_{i=1}^{E+1} w_i(t) \cdot y_{n_i(t)}$