MATH.APP.790 : Topics in Mathematics, Nonlinear time series analysis

Narayan P. Subramaniyam

Tampere University (MET Faculty)

January 2021

Lecture overview

- Phase space reconstruction
- Taken's embedding theorem
- Assumptions and generalizations
- Embedding delay and dimension
- References

1. F. Takens, Detecting Strange Attractors in Turbulence — Dynamical Systems and Turbulence, Lecture Notes in Mathematics 366, Springer (1981).
2. T. Sauer et al., J. Stat. Phys. 65, pp 579 (1991).
3. K. Judd and A. Mees, Physica D 120, pp 273 (1998)

Time series

- **Time series**: Sequence of measurements taken at a fixed (or irregular) sampling interval.

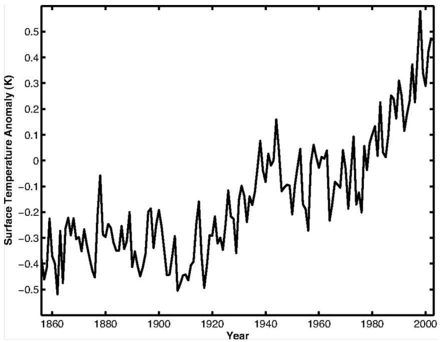
Time series

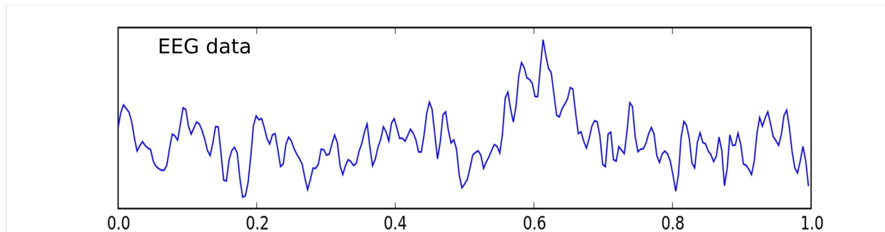
- **Time series**: Sequence of measurements taken at a fixed (or irregular) sampling interval.
- Electroencephalograhy (EEG), Electrocardiography (ECG), paleoclimate time series (irregularly spaced), stock price data etc

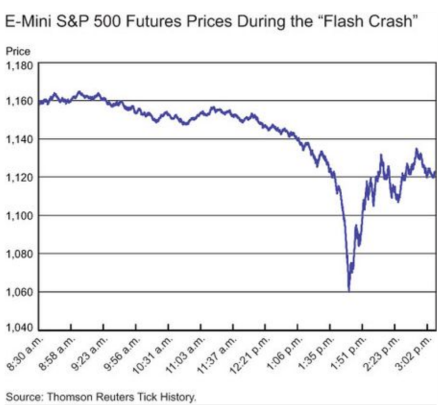
Time series

- **Time series**: Sequence of measurements taken at a fixed (or irregular) sampling interval.
- Electroencephalograhy (EEG), Electrocardiography (ECG), paleoclimate time series (irregularly spaced), stock price data etc
- Data recorded over time may have autocorrelation, trend or seasonal variations etc

Some examples of time series







Why to analyze time series ?

Why to analyze time series ?

- We are surrounded by time series data (Age of big data!)

Why to analyze time series ?

- We are surrounded by time series data (Age of big data!)
- Motivation to analyze time series

Why to analyze time series ?

- We are surrounded by time series data (Age of big data!)
- Motivation to analyze time series
    - Characterization
    - Prediction
    - Modelling

Why to analyze time series ?

- We are surrounded by time series data (Age of big data!)
- Motivation to analyze time series
    - Characterization
    - Prediction
    - Modelling
- Most commonly we have single series of measurements (univariate data), although more than one measurement channel (multivariate) are also widely available.

The need for nonlinear methods

- Linear methods (autocorrelation and power spectrum) is not adequate to distinguish stochastic from deterministic signals.

The need for nonlinear methods

- Linear methods (autocorrelation and power spectrum) is not adequate to distinguish stochastic from deterministic signals.
- Spectral methods make sense only if the system is linear, periodic and stationary.

The need for nonlinear methods

- Linear methods (autocorrelation and power spectrum) is not adequate to distinguish stochastic from deterministic signals.
- Spectral methods make sense only if the system is linear, periodic and stationary.
- Linear approaches do not tell much about the dynamics.

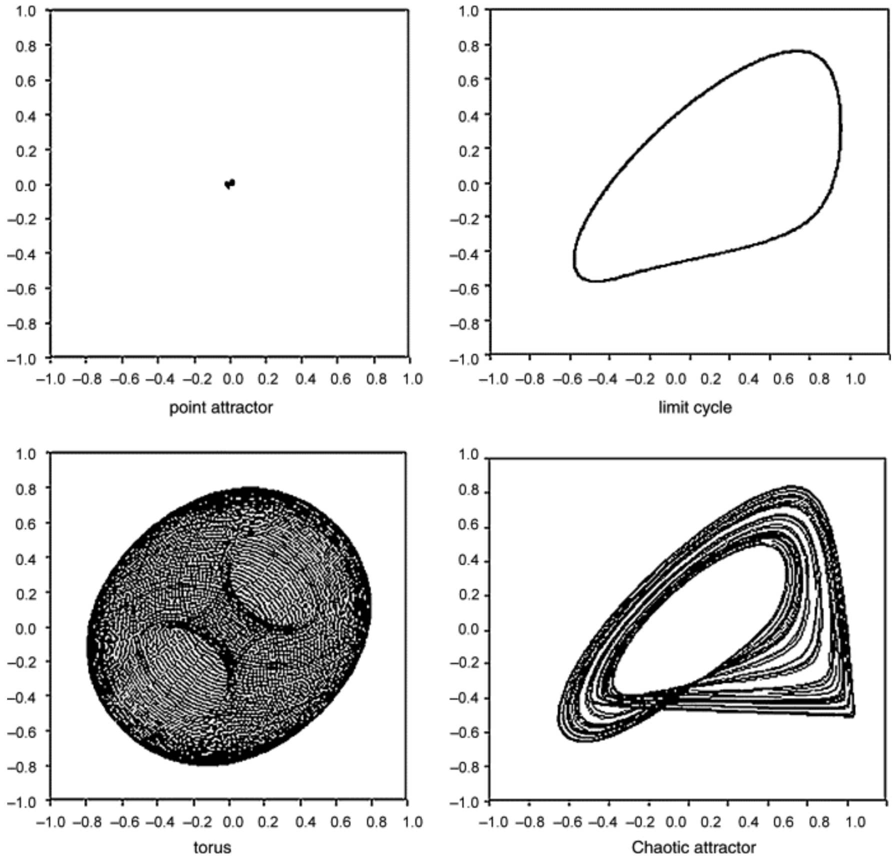
The need for nonlinear methods

- Linear methods (autocorrelation and power spectrum) is not adequate to distinguish stochastic from deterministic signals.
- Spectral methods make sense only if the system is linear, periodic and stationary.
- Linear approaches do not tell much about the dynamics.
- Assumption of Gaussianity must hold for linear methods to be adequate.

Recap : Attractors

- A subset of the phase space to which the trajectories accumulate to as $t \to \infty$
- Dynamics can be deduced from the topology of an attractor
  - Point attractor $\to$ fixed-point dynamics
  - Limit cycle attractor $\to$ periodic dynamics
  - Torus $\to$ quasiperiodic dynamics
  - Strange attractor $\to$ chaotic dynamics

Recap : Attractors



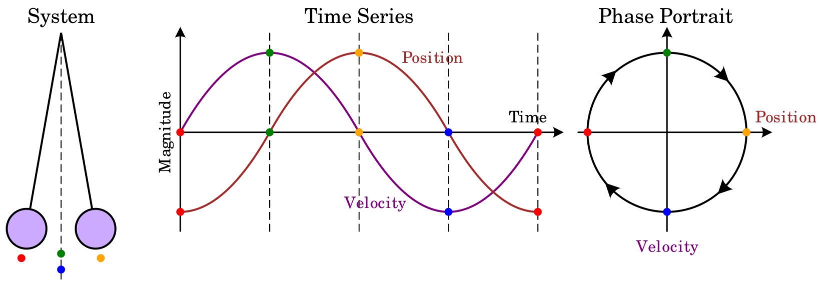
Phase space and time series

Phase space and time series

- A state of a dynamical system is minimum number of variables needed to fully represent the system.

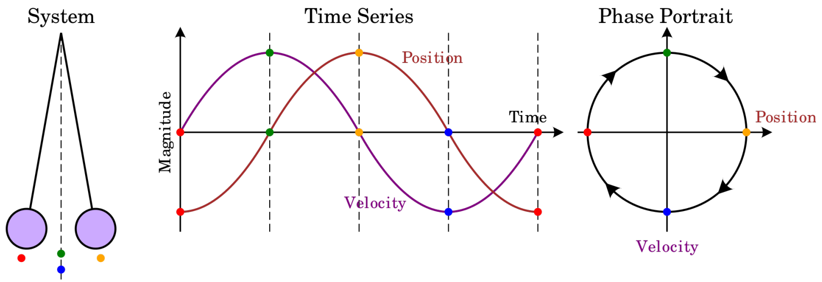
Phase space and time series

- A state of a dynamical system is minimum number of variables needed to fully represent the system.
- Example : A simple pendulum is a dynamical system and its state can represented by **position** and **velocity**.



Phase space and time series

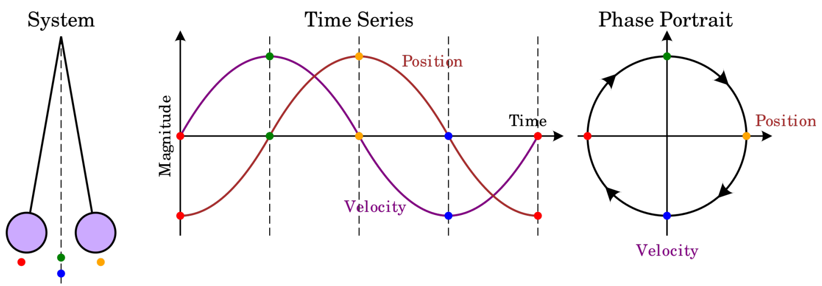
- A state of a dynamical system is minimum number of variables needed to fully represent the system.
- Example : A simple pendulum is a dynamical system and its state can represented by **position** and **velocity**.



- Phase space (or state space) is really an abstraction to represent the space of such possible states a dynamical system can be in.

Phase space and time series

- A state of a dynamical system is minimum number of variables needed to fully represent the system.
- Example : A simple pendulum is a dynamical system and its state can represented by **position** and **velocity**.



- Phase space (or state space) is really an abstraction to represent the space of such possible states a dynamical system can be in.
- What we observe is the time series!

The Need for Phase-Space Reconstruction

- Access to all dynamical variables is needed to directly observe the attractor.

The Need for Phase-Space Reconstruction

- Access to all dynamical variables is needed to directly observe the attractor.
- Often, only one (or limited number) of dynamical variables are observed.

The Need for Phase-Space Reconstruction

- Access to all dynamical variables is needed to directly observe the attractor.
- Often, only one (or limited number) of dynamical variables are observed.
- The dimension of phase space is not known.

The Need for Phase-Space Reconstruction

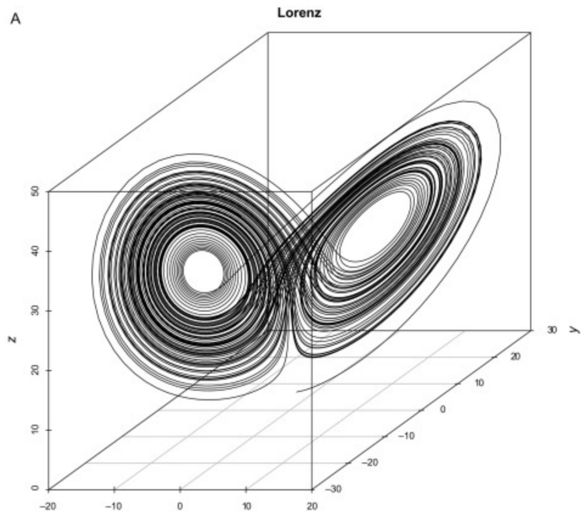
- Time series data measured from a dynamical system are generally some functions of (limited number of) state variables.

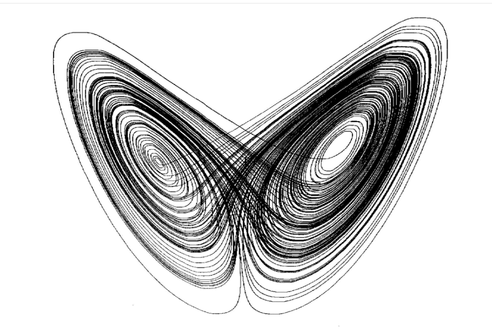
The Need for Phase-Space Reconstruction

- Time series data measured from a dynamical system are generally some functions of (limited number of) state variables.
- Measurement is not precise and is a projection of the full phase space space.

The Need for Phase-Space Reconstruction

- Time series data measured from a dynamical system are generally some functions of (limited number of) state variables.
- Measurement is not precise and is a projection of the full phase space space.
- Projections make trajectories look as if they are crossing each other (violation of the rule!)





Can we obtain important properties of the attractor from just one time series ?

Time series to phase space

- To analyze time-series in terms of dynamical system theory, we must reconstruct the phase space.

Time series to phase space

- To analyze time-series in terms of dynamical system theory, we must reconstruct the phase space.
- Consider a dynamical system, given by the diffeomorphism $\phi\colon M \to M$ on a (compact) manifold M together with smooth measurement function $f\colon M \to \mathbb{R}$.

Time series to phase space

- To analyze time-series in terms of dynamical system theory, we must reconstruct the phase space.
- Consider a dynamical system, given by the diffeomorphism $\phi\colon M \to M$ on a (compact) manifold M together with smooth measurement function $f\colon M \to \mathbb{R}$.
- The trajectories of the dynamical system are given by the sequences $\{x_n = \phi^n(x_0)\}$.

Time series to phase space

- To analyze time-series in terms of dynamical system theory, we must reconstruct the phase space.
- Consider a dynamical system, given by the diffeomorphism $\phi\colon M \to M$ on a (compact) manifold M together with smooth measurement function $f\colon M \to \mathbb{R}$.
- The trajectories of the dynamical system are given by the sequences $\{x_n = \phi^n(x_0)\}$.
- At each time index $n \in \mathbb{Z}$, what is observed is not the whole state x_n, but a scalar $y_n = f(x_n)$.

Time series to phase space

- To analyze time-series in terms of dynamical system theory, we must reconstruct the phase space.
- Consider a dynamical system, given by the diffeomorphism $\phi\colon M \to M$ on a (compact) manifold M together with smooth measurement function $f\colon M \to \mathbb{R}$.
- The trajectories of the dynamical system are given by the sequences $\{x_n = \phi^n(x_0)\}$.
- At each time index $n \in \mathbb{Z}$, what is observed is not the whole state x_n, but a scalar $y_n = f(x_n)$.
- Thus, Time series of successive measurements is given by $\{y_n\}, n = 1 \ldots N$, i.e $\{y_n = f(\phi^n(x_0) = f(x_n)\}$

Time series to phase space : Taken's embedding theorem

Theorem

For a compact $m-$dimensional manifold M, $\ell \geq 1$ and $k \geq 2m$ there is an open and dense subset U in $Diff^{\ell}(M) \times C^{\ell}(M)$, the product of the space of C^{ℓ}- diffeomorphisms on M and the space of C^{ℓ}-functions on M, such that for $(\phi, f) \in U$ the following map is an embedding of M into $\mathbb{R}^k$:

$$M \ni x \mapsto (f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x))) \in \mathbb{R}^k$$

The conclusion of this theorem holds for generic pairs of (ϕ, f)

Time series to phase space : Taken's embedding theorem

- The map from M to $\mathbb{R}^k$ given by, $x \mapsto (f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x))) \in \mathbb{R}^k$, is denoted by Rec_k.

Time series to phase space : Taken's embedding theorem

- The map from M to $\mathbb{R}^k$ given by, $x \mapsto (f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x))) \in \mathbb{R}^k$, is denoted by Rec_k.
- The vector $(f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x)))$ is called the reconstruction vector.

Time series to phase space : Taken's embedding theorem

- The map from M to $\mathbb{R}^k$ given by, $x \mapsto (f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x))) \in \mathbb{R}^k$, is denoted by Rec_k.
- The vector $(f(x), f(\phi(x)), \ldots, f(\phi^{k-1}(x)))$ is called the reconstruction vector.
- Taken's theorem is an adaptation of Whitney's embedding theorem which states any m-dimensional manifold can be smoothly embedded in $\mathbb{R}^{2m}$.

Time series to phase space : Taken's embedding theorem

- Taken's embedding theorem says that for a time series $\{y_n\}$, we can consider the sequence $\{(y_n, y_{n+1}, \cdots, y_{n+k-1}) \in \mathbb{R}^k\}_n$, which are termed reconstruction vectors.
- The reconstruction vectors are diffeomorphic to the trajectories of the underlying dynamical system generating the time series sequence $\{y_n\}$ if,
  - k is sufficiently large ($k \geq 2m + 1$).
  - The pair (ϕ, f) is generic.

Generalizations

For the Taken's embedding theorem the following assumptions were made

Generalizations

For the Taken's embedding theorem the following assumptions were made

- Discrete-time dynamics

Generalizations

For the Taken's embedding theorem the following assumptions were made

- Discrete-time dynamics
- Measurement function produces a scalar for each time index (i.e., univariate signal)

Generalizations

For the Taken's embedding theorem the following assumptions were made

- Discrete-time dynamics
- Measurement function produces a scalar for each time index (i.e., univariate signal)
- Manifold is compact

Continuous time

Continuous time

- Consider that $x(t) \in \mathbb{R}^m$ is a m-dimensional vector and $\phi\colon \mathbb{R}^m \to \mathbb{R}^m$ is a **vector field**.

Continuous time

- Consider that $x(t) \in \mathbb{R}^m$ is a m-dimensional vector and $\phi \colon \mathbb{R}^m \to \mathbb{R}^m$ is a **vector field**.
- Two approaches to define reconstruction : 1) Discretization and 2) Derivatives

Continuous time

- Consider that $x(t) \in \mathbb{R}^m$ is a m-dimensional vector and $\phi\colon \mathbb{R}^m \to \mathbb{R}^m$ is a **vector field**.
- Two approaches to define reconstruction : 1) Discretization and 2) Derivatives
- Discretization : Define a small interval $h > 0$ and obtain Rec_k in terms of the diffeomorphism ϕ^h.
- For $k > 2m$ and generic (ϕ, f, h), it can be shown that Rec_k is an embedding.

Continuous time

- For the derivative approach, the embedding can be defined as

$$Rec_k(x) = \left(f(\phi^t(x)), \frac{\partial}{\partial t} f(\phi^t(x)), \cdots, \frac{\partial^{k-1}}{\partial t^{k-1}} f(\phi^t(x)) \right) \Big|_{t=0}.$$

Continuous time

- For the derivative approach, the embedding can be defined as

$$Rec_k(x) = \left(f(\phi^t(x)), \frac{\partial}{\partial t} f(\phi^t(x)), \cdots, \frac{\partial^{k-1}}{\partial t^{k-1}} f(\phi^t(x)) \right)\Big|_{t=0}.$$

- For example, in case of the pendulum, if we know (or measure) the position, we can simply compute the derivative to get the velocity.

Continuous time

- For the derivative approach, the embedding can be defined as

$$Rec_k(x) = \left(f(\phi^t(x)), \frac{\partial}{\partial t} f(\phi^t(x)), \cdots, \frac{\partial^{k-1}}{\partial t^{k-1}} f(\phi^t(x)) \right) \Big|_{t=0}.$$

- For example, in case of the pendulum, if we know (or measure) the position, we can simply compute the derivative to get the velocity.
- Use difference methods!

Continuous time

- For the derivative approach, the embedding can be defined as

$$Rec_k(x) = \left(f(\phi^t(x)), \frac{\partial}{\partial t} f(\phi^t(x)), \cdots, \frac{\partial^{k-1}}{\partial t^{k-1}} f(\phi^t(x)) \right) \Big|_{t=0}.$$

- For example, in case of the pendulum, if we know (or measure) the position, we can simply compute the derivative to get the velocity.
- Use difference methods!
- Two practical issues with the derivative approach
  - It magnifies noise
  - We do not know what all the state variables are or its dimension.

Multivariate measurements

Multivariate measurements

- Measurement function has values in $\mathbb{R}^v$, i.e., v-dimensional measurement vector.

Multivariate measurements

- Measurement function has values in $\mathbb{R}^v$, i.e., v-dimensional measurement vector.
- The embedding Rec_k then has values in $\mathbb{R}^{vk}$.

Multivariate measurements

- Measurement function has values in $\mathbb{R}^v$, i.e., v-dimensional measurement vector.
- The embedding Rec_k then has values in $\mathbb{R}^{vk}$.
- Conclusions of the theorem holds for $vk > 2m$.

Non-compact sets

Non-compact sets

- Assuming M to be compact , we can say that after **some** time, every new state $\{x_n\}$ is approximately equal (or upto some arbitrary distance ϵ) to one of the past states.

Non-compact sets

- Assuming M to be compact , we can say that after **some** time, every new state $\{x_n\}$ is approximately equal (or upto some arbitrary distance ϵ) to one of the past states.
- In other words, compactness guarantees **recurrence**.

Non-compact sets

- Assuming M to be compact , we can say that after **some** time, every new state $\{x_n\}$ is approximately equal (or upto some arbitrary distance ϵ) to one of the past states.
- In other words, compactness guarantees **recurrence**.
- Due this property, the evolution of the states and the sequence of reconstructed vectors are metrically equal up to a bounded distortion, i.e.,

$$\frac{d(x_n, x_m)}{||(y_n, \cdots, y_{n+k-1}) - ((y_m, \cdots, y_{m+k-1})||} \tag{1}$$

are bounded away from zero.

Non-compact sets

- Assuming M to be compact , we can say that after **some** time, every new state $\{x_n\}$ is approximately equal (or upto some arbitrary distance ϵ) to one of the past states.
- In other words, compactness guarantees **recurrence**.
- Due this property, the evolution of the states and the sequence of reconstructed vectors are metrically equal up to a bounded distortion, i.e.,

$$\frac{d(x_n, x_m)}{||(y_n, \cdots, y_{n+k-1}) - ((y_m, \cdots, y_{m+k-1})||} \tag{1}$$

  are bounded away from zero.
- For a non-compact set, the theorem still holds but not this property of bounded distortion

More on compactness of phase space

For any evolution $\{x_n\}_{n\geq 0}$ of a dynamical system with compact manifold, one can define its limit set as

$$\omega(x_0) = \{x | \exists n_i \to \infty \quad \text{such that} \quad x_{n_i} \to x\}$$

- Such limit sets are often attractors.
- Due to compactness of the manifold, we can prove that for $\epsilon > 0$, there is some time segment $N'(\epsilon)$ such that for each $n > N'(\epsilon)$, any for any metric d, $d(x_n, \omega(x_0)) < \epsilon$
- In many cases, it is enough that the map Rec_k be an embedding of the limit set instead of the whole manifold M.
- Limit sets such as attractors often have much lower dimension than M ($\mathcal{A} \subset M$).

Summary

- We only observe (some transformation) of limited state variables with time series
- Phase space reconstruction enable us to study the unobserved state variables
- Often enough to reconstruct the attractor (lower dimension)!
- Attractors contain geometrical and dynamical properties of the original phase space
- The reconstructed attractor is diffeomorphic with the original attractor!.

Reconstructing the Lorenz attractor

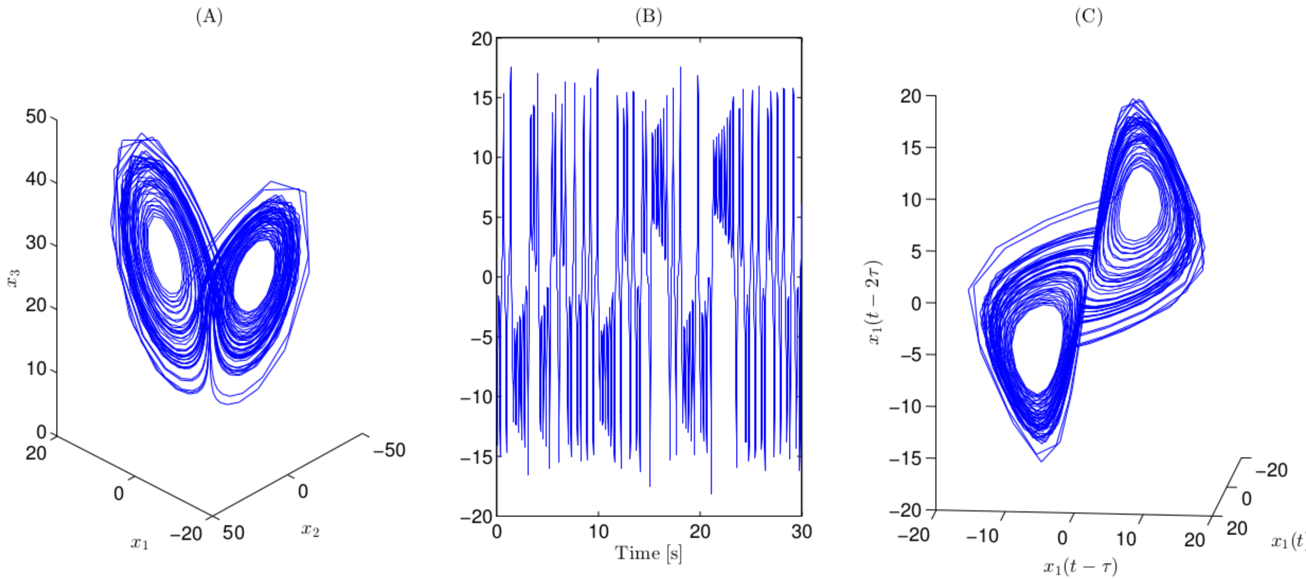


Figure: (A) The original Lorenz attractor. (B) The x_1-component of the Lorenz attractor, which is observed. (C) The Lorenz attractor in the reconstructed space after applying embed- ding theorem using only the x_1-component.

Taken's embedding theorem in practice

Theorem

Consider a compact, m-dimensional manifold M, with $\phi : M \to M$, a smooth diffeomorphism (at least in the sense of C^2). Consider a smooth (again in the sense of C^2) observation function, $f : M \to \mathbb{R}$. It is a generic property that,

$$Rec_k : M \to \mathbb{R}^{2m+1},$$

is an embedding and it is given as

$$Rec_k(x) = (f(x), f(\phi^1(x)), f(\phi^2 x)), \cdots, f(\phi^{2m}(x))$$

Taken's embedding theorem in practice

Theorem

Consider a compact, m-dimensional manifold M, with $\phi : M \to M$, a smooth diffeomorphism (at least in the sense of C^2). Consider a smooth (again in the sense of C^2) observation function, $f : M \to \mathbb{R}$. It is a generic property that,

$$Rec_k : M \to \mathbb{R}^{2m+1},$$

is an embedding and it is given as

$$Rec_k(x) = (f(x), f(\phi^1(x)), f(\phi^2 x)), \cdots, f(\phi^{2m}(x))$$

The components of Rec_k correspond to the time lagged observations of the dynamics on M (or $\mathcal{A}$), as defined by the smooth observation function f.

Taken's embedding theorem in practice : Embedding parameters

Taken's embedding theorem in practice : Embedding parameters

- To perform the embedding in practice, we need to know the embedding lag (τ) and dimension (k).

Taken's embedding theorem in practice : Embedding parameters

- To perform the embedding in practice, we need to know the embedding lag (τ) and dimension (k).
- In theory, the theorem will work for any arbitrary value of τ and the embedding theorem assumes $\tau = 1$.

Taken's embedding theorem in practice : Embedding parameters

- To perform the embedding in practice, we need to know the embedding lag (τ) and dimension (k).
- In theory, the theorem will work for any arbitrary value of τ and the embedding theorem assumes $\tau = 1$.
- The embedding dimension is $k > 2m$, but we do not know what is m in practice!

In general, any value of $\tau \in \mathbb{R}_+$ can be chosen. The reconstructed embedding vector, for an arbitrary τ can be given as,

$$x_n \in M \Longleftrightarrow (y_n, y_{n+\tau}, y_{n+2\tau}, \cdots y_{n+(k-1)\tau}).$$

Illustration

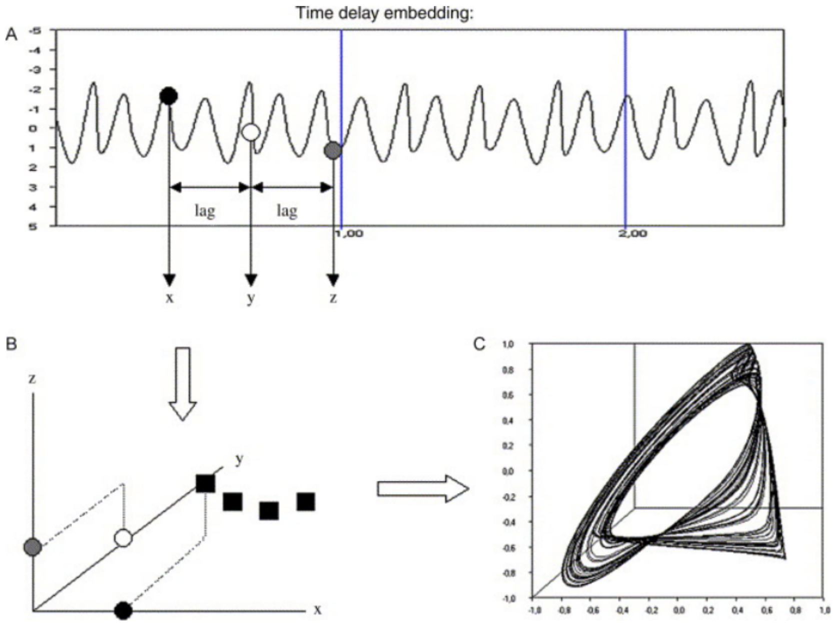
n	y
0	0.3
1	0.6
2	0.9
3	0.3
4	0.1
5	0.9
.	.
.	.
.	.

$\tau = 2$

$k = 3$

$$\begin{bmatrix} 0.3 & 0.9 & 0.1 \\ 0.6 & 0.3 & 0.9 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Illustration



Illustration

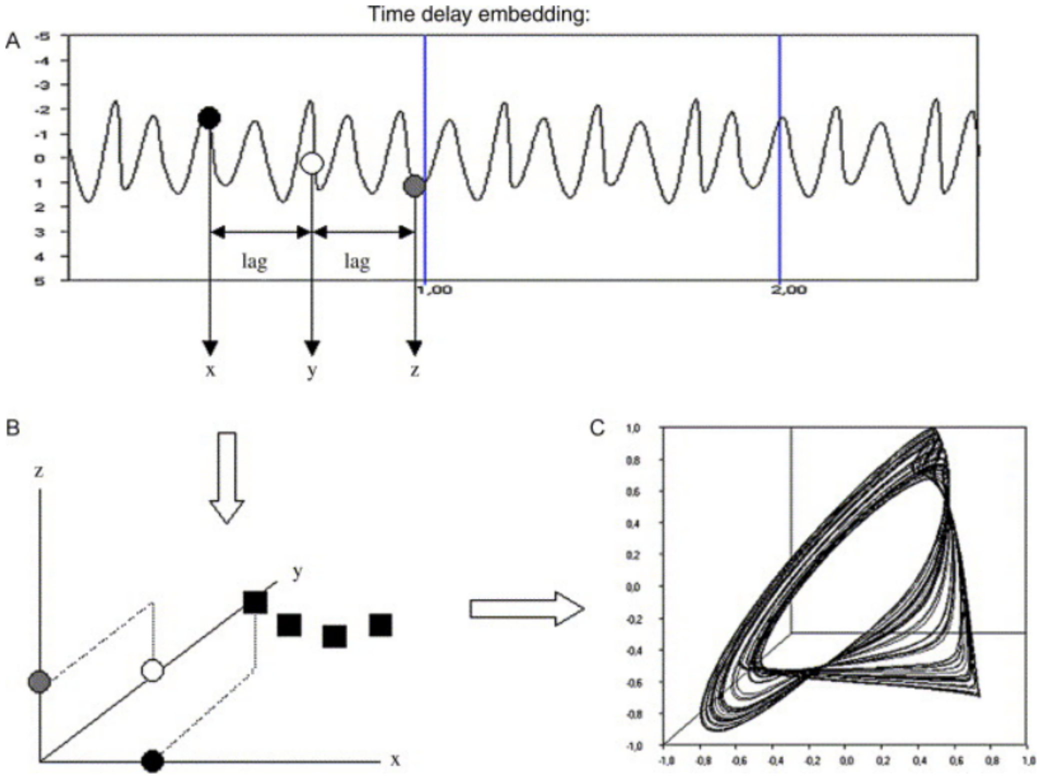


Figure: How would the attractor look if lag (τ) is set to 0 ?

Example of attractor reconstruction

The EEG time series below shows the transition from inter-ictal to ictal dynamics

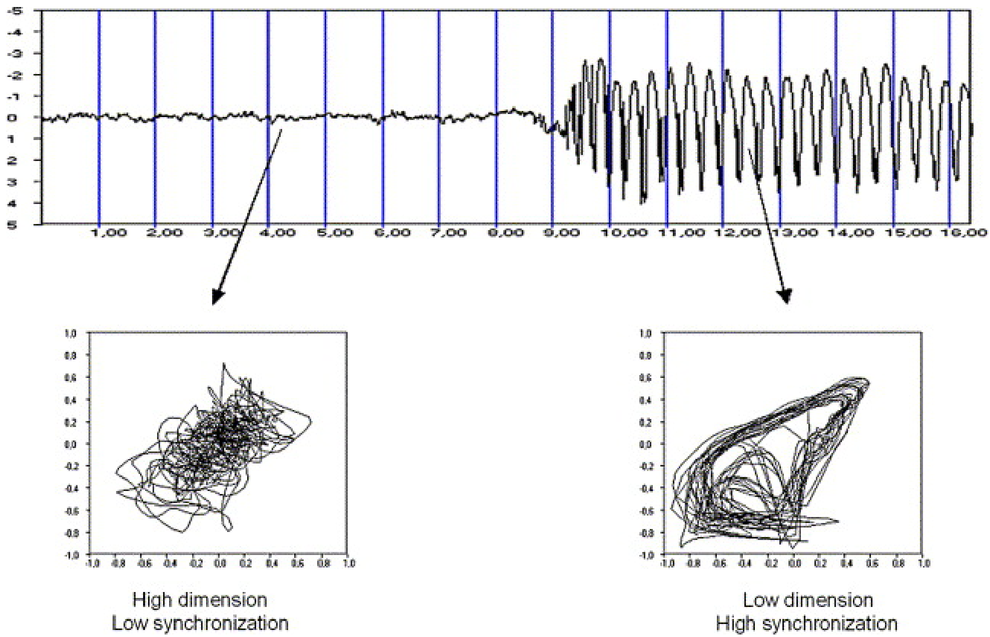


Figure: The geometry of the attractor gives insight into the dynamics!

Effect of τ on embedding

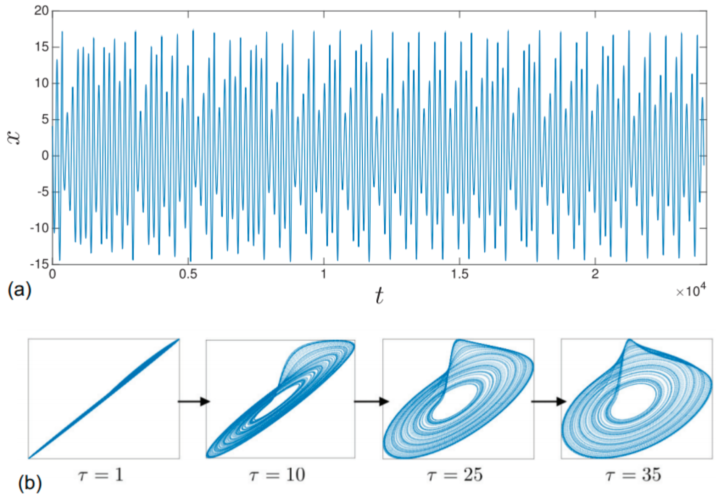
- If τ is too small, then there is almost no new information between the successive observations and the coordinates in the phase space will be too close to each other

Effect of τ on embedding

- If τ is too small, then there is almost no new information between the successive observations and the coordinates in the phase space will be too close to each other
- A very large τ can make the successive observations completely disconnected, relating each other in a random fashion.

Effect of τ on embedding

- If τ is too small, then there is almost no new information between the successive observations and the coordinates in the phase space will be too close to each other
- A very large τ can make the successive observations completely disconnected, relating each other in a random fashion.



How to choose delay and dimension ?

How to choose delay and dimension ?

- Embedding delay : Compute the (auto) mutual-information of the measurement time series,

$$I(\tau) = \sum_{n=1}^{N} p(y_n, y_{n+\tau}) \log\left(\frac{p(y_n, y_{n+\tau})}{p(y_n)p(y_{n+\tau})}\right).$$

How to choose delay and dimension ?

- Embedding delay : Compute the (auto) mutual-information of the measurement time series,

$$I(\tau) = \sum_{n=1}^{N} p(y_n, y_{n+\tau}) \log \left(\frac{p(y_n, y_{n+\tau})}{p(y_n)p(y_{n+\tau})} \right).$$

- Optimal value for the embedding delay is the value of τ at which $I(\tau)$ reaches the first local minimum.

How to choose delay and dimension ?

- Embedding delay : Compute the (auto) mutual-information of the measurement time series,

$$I(\tau) = \sum_{n=1}^{N} p(y_n, y_{n+\tau}) \log\left(\frac{p(y_n, y_{n+\tau})}{p(y_n)p(y_{n+\tau})}\right).$$

- Optimal value for the embedding delay is the value of τ at which $I(\tau)$ reaches the first local minimum.
- Embedding dimension : False nearest neighbour (FNN) principle.

How to choose delay and dimension ?

- Embedding delay : Compute the (auto) mutual-information of the measurement time series,

$$I(\tau) = \sum_{n=1}^{N} p(y_n, y_{n+\tau}) \log \left(\frac{p(y_n, y_{n+\tau})}{p(y_n) p(y_{n+\tau})} \right).$$

- Optimal value for the embedding delay is the value of τ at which $I(\tau)$ reaches the first local minimum.
- Embedding dimension : False nearest neighbour (FNN) principle.
- The idea behind the FNN method is to examine if the points along a trajectory in dimension k are also neighbors in dimension $k+1$.

How to choose delay and dimension ?

- Embedding delay : Compute the (auto) mutual-information of the measurement time series,

$$I(\tau) = \sum_{n=1}^{N} p(y_n, y_{n+\tau}) \log \left(\frac{p(y_n, y_{n+\tau})}{p(y_n)p(y_{n+\tau})} \right).$$

- Optimal value for the embedding delay is the value of τ at which $I(\tau)$ reaches the first local minimum.
- Embedding dimension : False nearest neighbour (FNN) principle.
- The idea behind the FNN method is to examine if the points along a trajectory in dimension k are also neighbors in dimension $k+1$.
- In dimensions lower than the actual dimension, many points on the trajectory will be close to each other (false neighbors) due to projection

FNN algorithm

- Starting with an initial dimension of $k = 1$, increment k by one and for each reconstructed state vector, the nearest neighbor is computed.
- For every increase in k, the percentage of false neighbors is also computed.
- The value of k at which the percentage of false neighbors becomes zero (or arbitrarily small due to the effect of noise) is considered as an appropriate choice for k.

Summary

Summary

- We only observe the time series generated by a dynamical system.

Summary

- We only observe the time series generated by a dynamical system.
- Taken's embedding theorem provides a way to reconstruct the phase space from single (or multiple) observations.

Summary

- We only observe the time series generated by a dynamical system.
- Taken's embedding theorem provides a way to reconstruct the phase space from single (or multiple) observations.
- Such reconstructions are diffeomorphic to the original phase space.

Summary

- We only observe the time series generated by a dynamical system.
- Taken's embedding theorem provides a way to reconstruct the phase space from single (or multiple) observations.
- Such reconstructions are diffeomorphic to the original phase space.
- Often one just needs to reconstruct the attractor, which has lower dimension than the manifold.

Summary

- We only observe the time series generated by a dynamical system.
- Taken's embedding theorem provides a way to reconstruct the phase space from single (or multiple) observations.
- Such reconstructions are diffeomorphic to the original phase space.
- Often one just needs to reconstruct the attractor, which has lower dimension than the manifold.
- Embedding requires two parameters - delay and dimension.