Dynamics and time series: theory and applications

Stefano Marmi
Scuola Normale Superiore

Lecture 9, Dec 19, 2011

Takens theorem

- $\phi : X \to X$ map, $f : X \to R$ smooth observable
- Time-delay map (reconstruction of the dynamics from periodic sampling):
- $F(f,\phi) : X \to R^n$ n is the number of delays
- $F(f,\phi)(x) = (f(x), f(\phi(x)), f(\phi\circ\phi(x)), \ldots, f(\phi^n (x)))$ -1
- Under mild assumptions if the dynamics has an attractor with dimension k and n>2k then for almost any choice of the observable the reconstruction map is injective

Immersions and embeddings

- A smooth map F on a compact smooth manifold A is an immersion if the derivative map DF(x) (represented by the Jacobian matrix of F at x) is one-to-one at every point $x \in A$. Since DF(x) is a linear map, this is equivalent to DF(x) having full rank on the tangent space. This can happen whether or not F is one-to-one. Under an immersion, no differential structure is lost in going from A to F(A).
- An embedding of A is a smooth diffeomorphism from A onto its image F(A), that is, a smooth one-to-one map which has a smooth inverse. For a compact manifold A, the map F is an embedding if and only if ,F is a one- to-one immersion.
- The set of embeddings is open in the set of smooth maps: arbitrarily small perturbations of an embedding will still be embeddings!

Embedology (Sauer, Yorke, Casdagli, J. Stat. Phys. 65 (1991)

Whitney showed that a generic smooth map ,F from a d-dimensional smooth compact manifold M to R^n , n>2d is actually a diffeomorphism on M. That is, M and F(M) are diffeomorphic. We generalize this in two ways:

- first, by replacing "generic" with "probability-one" (in a prescribed sense),
- second, by replacing the manifold M by a compact invariant set A

contained in some R^k that may have noninteger box-counting dimension (boxdim). In that case, we show that almost every smooth map from a neighborhood of A to R^n is one-to-one as long as n>2 * boxdim(A)

We also show that almost every smooth map is an embedding on compact subsets of smooth manifolds. This suggests that embedding techniques can be used to compute positive Lyapunov exponents (but not necessarily negative Lyapunov exponents). The positive Lyapunov exponents are usually carried by smooth unstable manifolds on attractors.

Embedology (Sauer, Yorke, Casdagli, J. Stat. Phys. 65 (1991)

Takens dealt with a restricted class of maps called delay-coordinate maps: these are time series of a single observed quantity from an experiment. He showed (F. Takens, Detecting strange attractors in turbulence, in Lecture Notes in Mathematics, No. 898 (Springer-Verlag, 1981) that if the dynamical system and the observed quantity are generic, then the delay-coordinate map from a d-dimensional smooth compact manifold M to R^n , n>2d is a diffeomorphism on M.

- we replace generic with probability-one
- and the manifold M by a possibly fractal set.

Thus, for a compact invariant subset A under mild conditions on the dynamical system, almost every delay-coordinate map to R^n is one-to-one on A provided that n>2.boxdim(A). Also, any manifold structure within I will be preserved in F(A).

- Only C^1 smoothness is needed.;
- For flows, the delay must be chosen so that there are no periodic orbits with period exactly equal to the time delay used or twice the delay

Embedding method

- Plot $x(t)$ vs. $x(t-\tau)$, $x(t-2\tau)$, $x(t-3\tau)$, …
- $x(t)$ can be any observable
- The embedding dimension is the # of delays
- The choice of τ and of the dimension are critical
- For a typical deterministic system, the orbit will be diffeomorphic to the attractor of the system (Takens theorem)

Choice of Embedding Parameters

Theoretically, a time delay coordinate map yields an valid embedding for any sufficiently large embedding dimension and for any time delay when the data are noise free and measured with infinite precision.

But, there are several problems:

(i) Data are not clean
(ii) Large embedding dimension are computationally expensive and unstable
(iii) Finite precision induces noise

Effectively, the solution is to search for:

(i) Optimal time delay τ
(ii) Minimum embedding dimension d

or

(i) Optimal time window τ_w

There is no one unique method solving all problems and neither there is an unique set of embedding parameters appropriate for all purposes.

Choice of Embedding Parameters

Theoretically, a time delay coordinate map yields an valid embedding for any sufficiently large embedding dimension and for any time delay when the data are noise free and measured with infinite precision.

But, there are several problems:

(i) Data are not clean
(ii) Large embedding dimension are computationally expensive and unstable
(iii) Finite precision induces noise

Effectively, the solution is to search for:

(i) Optimal time delay τ
(ii) Minimum embedding dimension d

or

(i) Optimal time window τ_w

There is no one unique method solving all problems and
neither there is an unique set of embedding parameters appropriate for all purposes.

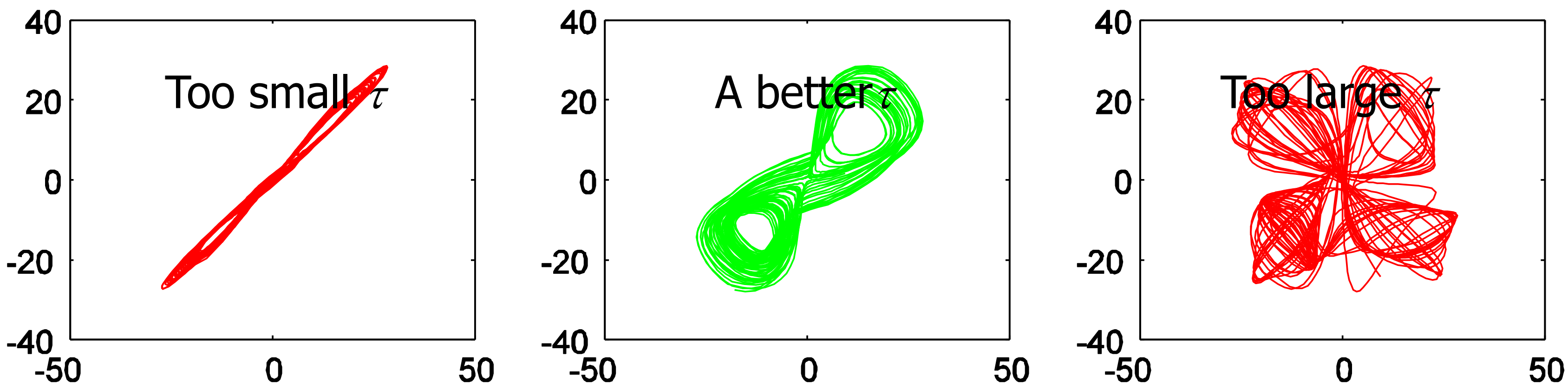
The Role of Time Delay τ

If τ is too small, $x(t)$ and $x(t-\tau)$ will be very close, then each reconstructed vector will consist of almost equal components → *Redundancy (τ_R)*

→ The reconstructed state space will collapse into the main diagonal

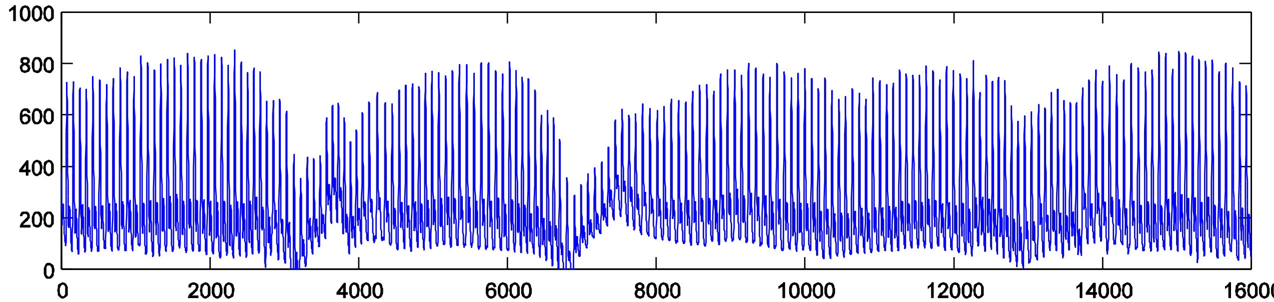
If τ is too large, $x(t)$ and $x(t-\tau)$ will be completely unrelated, then each reconstructed vector will consist of irrelevant components → *Irrelevance (τ_I)*

→ The reconstructed state space will fill the entire state space.



http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Blood Pressure Signal

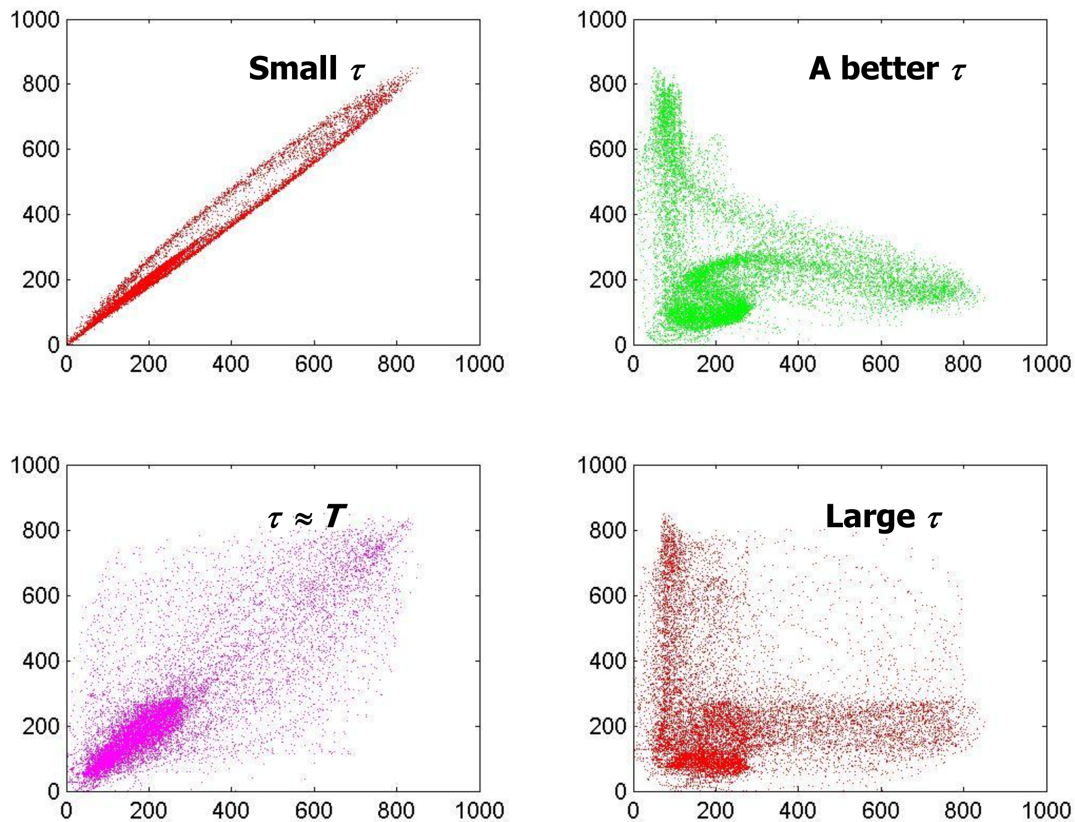


A better choice is:

$\tau_R < \tau_W < \tau_I$

Caution: τ should not be close to main period

Collapsing of state space



http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Some Recipes to Choose τ

Based on Autocorrelation

Estimate autocorrelation function: $C(\tau) = \frac{1}{N-\tau-1}\sum_{t=0}^{N-\tau-1} x(t)x(t+\tau) = \langle x(t)x(t+\tau)\rangle$

Then, $\tau_{opt} \approx C(0)/e$

or

first zero crossing of $C(\tau)$

Modifications:

1. Consider minima of higher order autocorrelation functions, $<x(\tau)x(t+\tau)x(t+2\tau)>$ and then look for time when these minima for various orders coincide.

Albano et al. (1991) *Physica D*

2. Apply nonlinear autocorrelation functions: $<x^2(\tau)x^2(t+2\tau)>$

Billings, Tao (1991) *Int. J. Control.*

http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Based on Time delayed Mutual Information

The information we have about the value of $x(t+\tau)$ if we know $x(t)$.

1. Generate the histogram for the probability distribution of the signal $x(t)$.

2. Let p_i is the probability that the signal will be inside the i-th bin and $p_{ij}(t)$ is the probability that $x(t)$ is in i-th bin and $x(t+\tau)$ is in j-th bin.

3. Then the mutual information for delay τ will be

$$I(\tau) = \sum_{i,j} p_{ij}(\tau) \log p_{ij}(\tau) - 2\sum_{i} p_i \log p_i$$

For $\tau \to 0$, $I(\tau) \to$ Shannon's Entropy

$\tau_{opt} \approx$ First minimum of $I(\tau)$

http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt