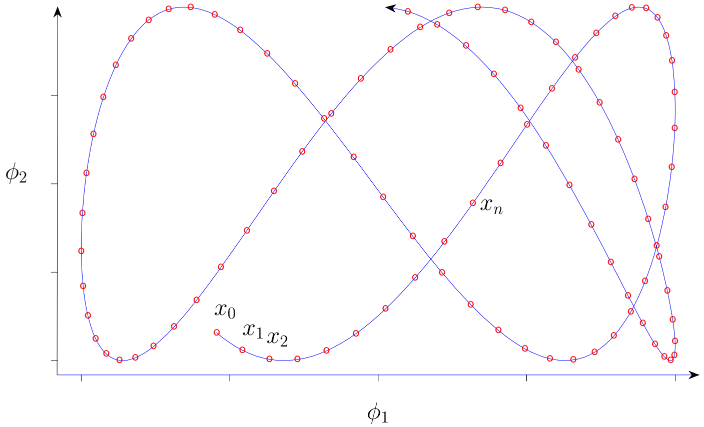
Time Series of Deterministic Dynamic Systems: Celebrated Takens theorem

Sven Laur
swen@math.ut.ee

November 1, 2004

What is a deterministic dynamic system?



Dynamic system evolves in time

$$\boldsymbol{x_{i+1}} = T(\boldsymbol{x_i}),$$

where T is a deterministic rule.

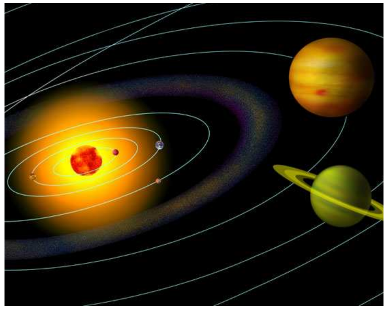
Given initial point $\boldsymbol{x_0}$ and sampling time, we get a positive orbit

$$\boldsymbol{X} = (\boldsymbol{x_0}, \boldsymbol{x_1}, \boldsymbol{x_2}, \ldots) = (\boldsymbol{x_0}, T(\boldsymbol{x_0}), T(T(\boldsymbol{x_0})), \ldots).$$

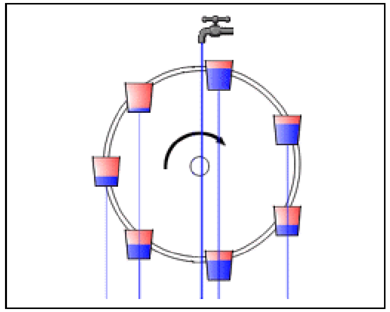
We are interested in long-term properties of the system.

Three possible types of dynamic systems

Catastrophic



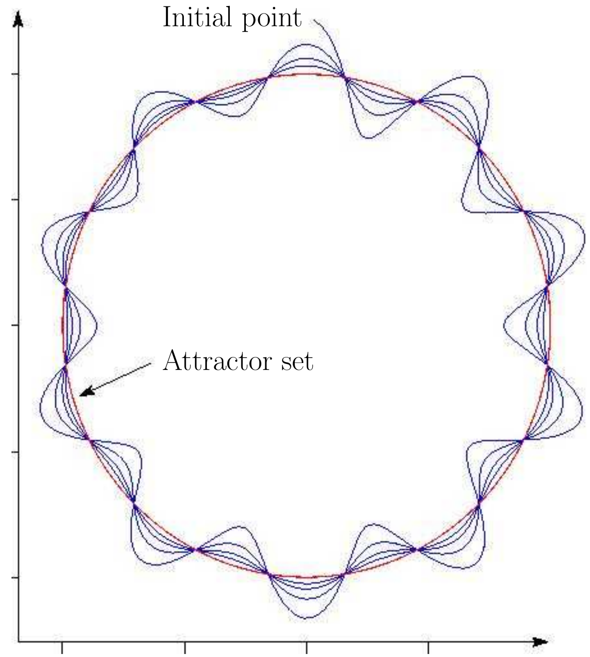
Stable



Chaotic

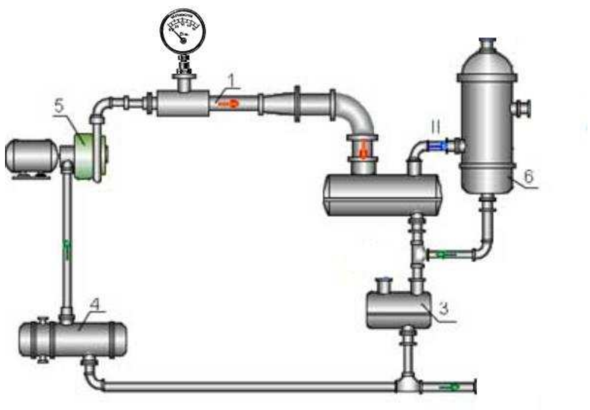
- Catastrophic—the trajectory in the phase space is unbounded.
- Stable—trajectory is periodic or quasi-periodic.
- Chaotic—trajectory jumps randomly between different sub-paths.

Long-term properties of dynamical system



- If the orbit is bounded then there exists an attractor set A such that if n is large enough $d(A, \boldsymbol{x_n}) < \epsilon$. The value $\epsilon > 0$ can be arbitrary.
- We require that A is stationary $T(A) = A$ and in some sense minimal.
- The geometrical shape of the attractor set A determines complexity of dynamical system.

Observations. Measurement scheme



- We cannot directly observe the state $\boldsymbol{x_k} \in \mathcal{X}$ of the system.
- System states completely determine measurements via read-out function

$$f : \mathcal{X} \to \mathbb{R}$$

For each orbit $\boldsymbol{X}$ there is a corresponding time serie

$$\boldsymbol{Y} = (y_0, y_1, \ldots) = (f(\boldsymbol{x_0}), f(\boldsymbol{x_1}), \ldots)$$

Can we reconstruct the internal state of the system?

Delay maps. Extended observation orbits

Single measurement cannot describe internal state of the complex system.

Consider k-tuples $(y_i, y_{i+1}, \ldots, y_{i+k-1})$ and denote

$$\mathrm{Rec}_k(\boldsymbol{x}) = (f(\boldsymbol{x}), f(T(\boldsymbol{x})), \ldots, f(T^{k-1}(\boldsymbol{x}))).$$

We would like to

- distinguish $\mathrm{Rec}_k(\boldsymbol{X_1})$ and $\mathrm{Rec}_k(\boldsymbol{X_2})$ if orbits $\boldsymbol{X_1} \neq \boldsymbol{X_2}$.
- detect “critical” points, where external forces cause change of orbit.
  (a) jumps
  (b) angle-points

Takens theorem (1981)

Let $\mathcal{X}$ be a bounded set. In the Cartesian product space of C^1-mappings on $\mathcal{X}$ and the space of C^1-functions from $\mathcal{X}$ to $\mathbb{R}$ there exists a open and dense subset U such that if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

Takens theorem (1981)

on $\mathcal{X}$ and the space of C^1-functions from $\mathcal{X}$ to $\mathbb{R}$ there exists a open and dense subset U such that if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

Takens theorem (1981)

dense subset U such that if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

Takens theorem (1981)

if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

- We have a deterministic system with rule $T : \mathcal{X} \to \mathcal{X}$.
- We have a read-out function $f : \mathcal{X} \to \mathbb{R}$.

Ideal regressor

$$
\begin{tikzcd}
\text{Current state} \arrow[r, "Det.rule", "T"'] \arrow[d, shift right, "\mathsf{Rec}_k"'] & \text{Next state} \arrow[d, "\mathsf{Rec}_k"'] \\
(y_i, \ldots, y_{i+k-1}) \arrow[u, shift right, "\mathsf{Rec}_k^{-1}"'] \arrow[r, "Prediction", "\mathsf{Rec}_k \circ T \circ \mathsf{Rec}_k^{-1}"'] & (y_{i+1}, \ldots, y_{i+k})
\end{tikzcd}
$$

- Explicitly stated, if $k > 2\dim(\mathcal{X})$ there exists a precise deterministic rule g for predicting the next state of the time serie!
- But g might be missing $g \notin \mathcal{F}$ from regression functions.

Should we care about U?

if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

Should we care about U?

exists a open and dense subset U such that if $(T, f) \in U$, then the reconstruction map Rec_k is an embedding, whenever $k > 2 \cdot \dim(\mathcal{X})$. Moreover, the embedding is continuously differentiable and has also continuously differentiable inverse.

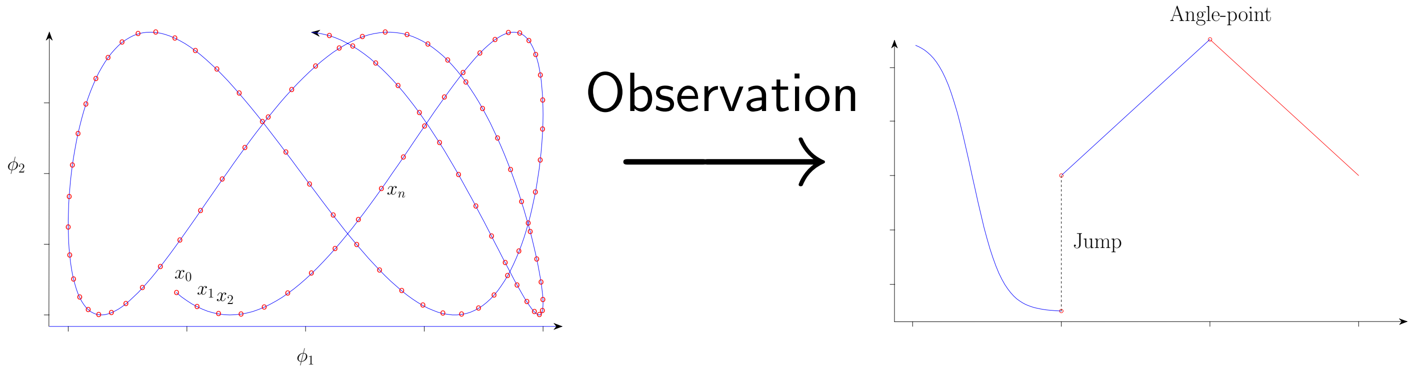
The latter means

- Embedding exists almost for all function pairs (f, T)
- If $(f, T) \notin U$, then exists close function pairs $(\hat{f}, \hat{T}) \in U$. More precisely, for any $\epsilon > 0$ we have $\hat{f}$ and $\hat{T}$ such that

$$\forall \boldsymbol{x} \in \mathcal{X} \quad |f(\boldsymbol{x}) - \hat{f}(\boldsymbol{x})| + |f'(\boldsymbol{x}) - \hat{f}'(\boldsymbol{x})| < \epsilon,$$
$$\forall \boldsymbol{x} \in \mathcal{X} \quad |T(\boldsymbol{x}) - \hat{T}(\boldsymbol{x})| + |T'(\boldsymbol{x}) - \hat{T}'(\boldsymbol{x})| < \epsilon.$$

Assumptions of the Takens theorem

- The set $\mathcal{X}$ is bounded—the system in non-catastrophic.
- The rule $T : \mathcal{X} \to \mathcal{X}$ must be continuously differentiable—most physical systems satisfy it by default.
- The read-out function $f : \mathcal{X} \to \mathbb{R}$ is continuously differentiable.

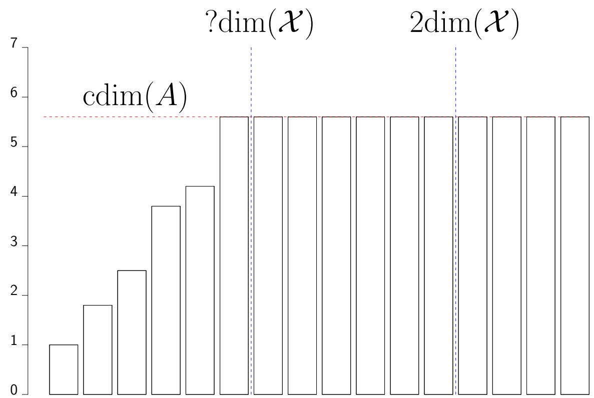


When the delay map is large enough?

Takens theorem assures that if $k > 2\dim(\mathcal{X})$ then

$$\mathrm{cdim}(A) = \mathrm{cdim}(\mathsf{Rec}_k(A)).$$

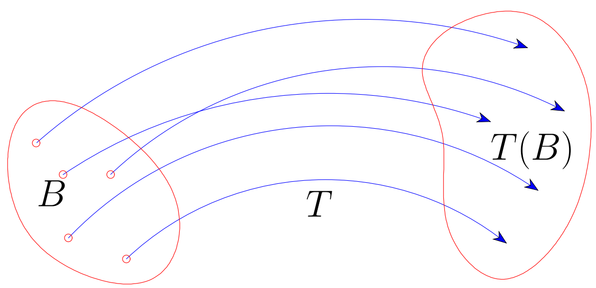
In other words the $\mathrm{cdim}(\mathsf{Rec}_k(A))$ stops growing.



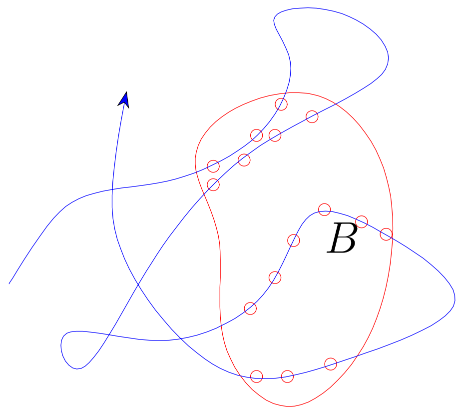
Stationary probability distribution

To get grasp over correlation dimension of the attractor set A, we need a stationary probability distribution

$$\Pr\left[\boldsymbol{x_i} \in B\right] = \Pr\left[\boldsymbol{x_{i+1}} \in T(B)\right] = \Pr\left[\boldsymbol{x_i} \in T(B)\right]$$



Average presence time



- Average presence time counts how long on average orbit stays in the set B

$$\Pr\left[\boldsymbol{x} \in B\right] = \lim_{n\to\infty} \frac{1}{n+1} \sum_{i=0}^{n} \Pr\left[T^i \boldsymbol{x}_{\mathbf{0}} \in B\right]$$

- Average presence time is stationary probability distribution.
- We can sample points from it.

Correlation dimension

- The fractional dimension of an attractor set A is defined via

$$C(r) = \Pr\left[\|X - Y\|_\infty \leq r\right],$$

where X and Y are independently drawn from the stationary distribution.

- The correlation dimension is a limit

$$\operatorname{cdim}(A) = \lim_{r \to 0^+} \frac{\log C(r)}{\log r} \qquad \operatorname{cdim}(A) \in [0, \dim(\mathcal{X})]$$

Monte-Carlo integration. Grassberger-Proccacia

If $\boldsymbol{z_i}$ and $\boldsymbol{z_j}$ are drawn from our distribution

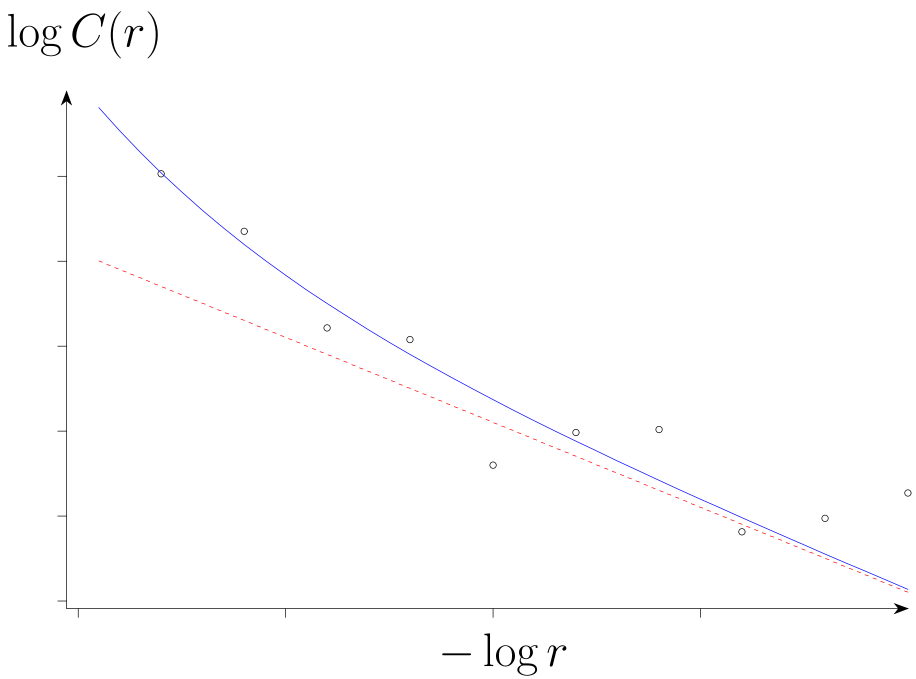
$$C_n(r) = \frac{2}{n(n-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} ?\mathrm{Is}[\|\boldsymbol{z_i} - \boldsymbol{z_j}\|_\infty \leq r] \approx \Pr\left[\|X - Y\|_\infty \leq r\right]$$

(1) Compute pairs $(r_1, \alpha_1), \ldots (r_m, \alpha_m)$.

(2) Fit a line through $(\log r_1, \log \alpha_1), \ldots, (\log r_m, \log \alpha_m)$.

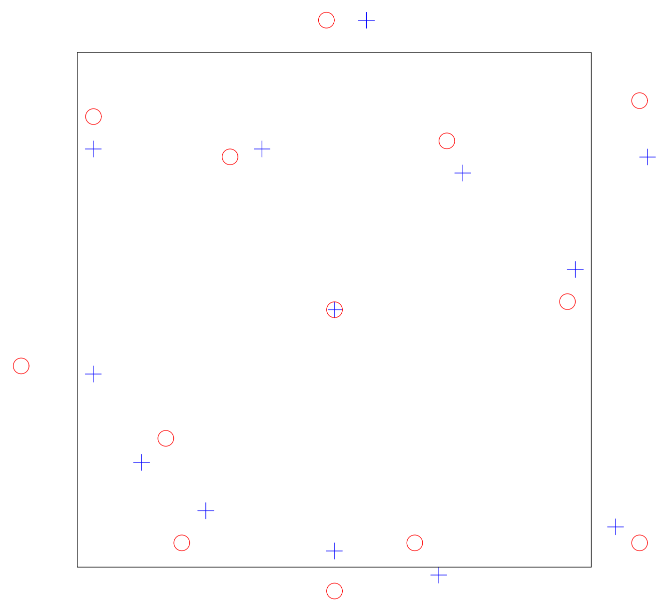
(3) Slope is the estimator of correlation dimension.

Systematical error versus statistical error



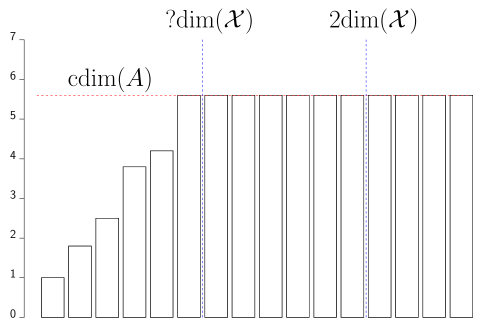
If we decrease r then there will be smaller number of samples close enough.

The effect of noise on Grassberger-Proccacia estimator



Noise adds a bias to counting—it is more probable to move points apart than bring together.

Hidden rocks in shallow water



- Early stop due to the errors in the correlation dimension estimate.
- Just a bad luck

$$\mathrm{cdim}(A) \ll \dim(\mathcal{X}) \Rightarrow \dim(\mathcal{X}) \not\leq k$$

- System is chaotic—the ideal regressor

$$\mathsf{Rec}_k \circ T \circ \mathsf{Rec}_k^{-1}.$$

is no better than random guessing.

Interpretation of Takens Theorem

IF the correlation dimensions of $\mathsf{Rec}_k(\boldsymbol{Y})$ and $\mathsf{Rec}_{k+1}(\boldsymbol{Y})$

- are equal
- or close enough

THEN

- $1^\star \cdot \dim(\mathcal{X}) \leq k \leq 2 \cdot \dim(\mathcal{X}) + 1$
- and the optimal regressor size is between k and $2^\star \cdot k + 1$.