Takens Embedding Theorem

Content

Introduction

Description	Notation	Remarks
State space	Compact manifold M of dimension m	unobservable
Dynamic	Vector field X with global flow ϕ_t	unknown
Measurements	Smooth functions $y_i, 1 \leq i \leq L$	$y_i : M \to \mathbb{R}$ at least C^1
Experiments	Set of initial states $\{x_j \in M, 1 \leq j \leq K\}$	
Dataset	$\{y_i(\phi_{kT}(x_j))\}$ $0 \leq k \leq N-1$	Timetstep $T > 0$

Q: Is there enough information in our dataset to predict $y_i(\phi_{NT}(x_j))$ for some measurement function y_i and some initial state x_j?

Introduction

Q: Is there enough information in our dataset to predict $y_i(\phi_{NT}(x_j))$ for some measurement function y_i and some initial state x_j?
A: Yes if:

- $N > 2m$
- The experiments are "representative" of M
  (i.e., K is large enough and the x_j are well spread across M)

Moreover, we only need one measurement function (i.e., $L = 1$).

Takens embedding theorem [1]

Theorem 1: Takens

Let M be a compact manifold of dimension m.
Let $\phi \in \mathit{Diff}^1(M)$ and $y \in C^1(M, \mathbb{R})$.
Define the map $f_{\phi,y} : M \to \mathbb{R}^{2m+1}$ by
$f_{\phi,y}(x) = (y(x), y(\phi(x)), \ldots, y(\phi^{2m}(x)))$.
Then, for pairs (ϕ, y), it is a generic property that $f_{\phi,y}$ is an embedding.

A property is *generic* if it holds for a set which is *open and dense* (in the C^1 topology).

Illustration (Sauer et al. [2])

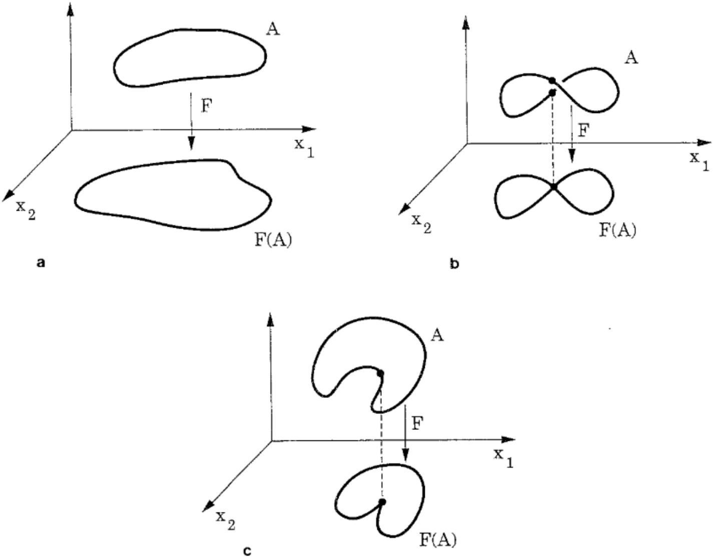

Fig. 1. (a) An embedding F of the smooth manifold A into R^2. (b) An immersion that fails to be one-to-one. (c) A one-to-one map that fails to be an immersion.

Summary of a proof [3], [4], [5]

Theorem 2: Whitney embedding theorem

$Emb(M, \mathbb{R}^{2m+1})$ is open and dense in $C^1(M, \mathbb{R}^{2m+1})$.

Openness

- The map $(\phi, y) \mapsto f_{\phi,y}$ from $Diff^1(M) \times C^1(M, \mathbb{R})$ to $C^1(M, \mathbb{R}^{2m+1})$ is continuous.
- By Whitney theorem, the set of (ϕ, y) such that $f_{\phi,y}$ is an embedding is **open**.

It remains to show that this set is **dense**.

Summary of a proof

- Let $x \in M$ **not** on a periodic orbit of ϕ of period less or equal than $2m$.
- Then the point $x, \phi(x), \ldots, \phi^{2m}(x)$ are distinct.
- We can perturb y independently in the neighbourhood of each of these points, so that the perturbed $f_{\phi,y}$ agrees with any map of $C^1(M, \mathbb{R}^{2m+1})$ on a neighbourhood of a x.
- By Whitney theorem, the set of y for which $f_{\phi,y}$ is an embedding on some neighbourhood of x is dense.

Summary of a proof

In order to piece together these local embeddings, we have to show that $f_{\phi,y}(x) \neq f_{\phi,y}(x')$ for all $x \neq x'$. However, if $x' = \phi^k(x)$ for some $2m < k \leq 4m$ then $f_{\phi,y}(x)$ and $f_{\phi,y}(x')$ share some components and we cannot perturb ϕ and y to affect independently $f_{\phi,y}(x)$ and $f_{\phi,y}(x')$.

To overcome this issue and deal with periodic orbits $\leq 2m$

First, prove the result for diffeomorphisms ϕ with particular assumptions concerning periodic orbits with period $\leq 4m$.
Then show that the set of these diffeomorphisms is dense.

Summary of a proof

Let $\mathcal{D} = \{\phi \in \mathit{Diff}^1(M)$ such that all periodic points of ϕ of period $q \leq 4m$ are hyperbolic and has distinct eigenvalues$\}$.

- $\mathcal{D}$ is open and dense in $\mathit{Diff}^1(M)$ (using Kupka-Smale theorem).
- Given $\phi \in \mathcal{D}$, let $\mathcal{P}_\phi$ be the set of periodic points of ϕ of period less or equal than $4m$. Since hyperbolic orbits are isolated from any other periodic orbit, $\mathcal{P}_\phi$ consists of a finite number of points.

There exists an open and dense set $\mathcal{A}_\phi \subset C^1(M, \mathbb{R})$ such that $f_{\phi,y}$ is an embedding on $\mathcal{P}_\phi$.

- The injectivity comes from the fact that the set of y that are injective on $\mathcal{P}_\phi$ is open and dense (since $\mathcal{P}_\phi$ consists of a finite number of points).
- The immersivity is trickier so we assume it (it relies on a Vandermonde determinant).

Summary of a proof

Let $\mathcal{E} = \{(\phi, y) : \phi \in \mathcal{D}, y \in \mathcal{A}_\phi\}$.

- It can be shown that $\mathcal{E}$ is open and dense in $\mathit{Diff}^1(M) \times C^1(M, \mathbb{R})$.
- In the previous slide, we have seen that for $(\phi, y) \in \mathcal{E}$, the map $f_{\phi,y}$ is an embedding of $\mathcal{P}_\phi$.
- To complete the proof, we have to show that for an open dense subset of $\mathcal{E}$, the map $f_{\phi,y}$ is an embedding of M. Both the immersivity and the injectivity are proved with a **transversality** argument.
- We only see the injectivity part and assume the immersivity.

Summary of a proof

Let the map $\sigma : \mathcal{E} \to C^1(M \times M \setminus \Delta, \mathbb{R}^d \times \mathbb{R}^d)$ defined by $\sigma(\phi, y)(x, x') = (f_{\phi,y}(x), f_{\phi,y}(x'))$, where Δ is the diagonal in $M \times M$. We will show that, for $f_{\phi,y}$ to be injective, it is necessary that $d \geq 2m + 1$. If $\hat{\Delta}$ is the diagonal in $\mathbb{R}^d \times \mathbb{R}^d$, then $f_{\phi,y}$ is injective if and only if its image does not intersect $\hat{\Delta}$.

- Using Abraham's parametric transversality theorem, it can be shown that the set of (ϕ, y) such that $\sigma(\phi, y)$ is transversal to $\hat{\Delta}$ is residual in $\mathcal{E}$.
- This means that, for generic $(\phi, y) \in \mathcal{E}$, $\sigma(\phi, y)_*(T_{x,x'}(M \times M \setminus \Delta)) + T_{z,z}\hat{\Delta} = T_{z,z}(\mathbb{R}^d \times \mathbb{R}^d)$ for all $(x, x') \in M \times M$ such that $\sigma(\phi, y)(x, x') = (z, z) \in \hat{\Delta}$.

Summary of a proof

For generic $(\phi, y) \in \mathcal{E}$,
$\sigma(\phi, y)_*(T_{x,x'}(M \times M \setminus \Delta)) + T_{z,z}\hat{\Delta} = T_{z,z}(\mathbb{R}^d \times \mathbb{R}^d)$
for all $(x, x') \in M \times M$ such that $\sigma(\phi, y)(x, x') = (z, z) \in \hat{\Delta}$.
On the other hand, we have that:

- $\dim(T_{x,x'}(M \times M \setminus \Delta)) = 2m$,
  hence $\dim(\sigma(\phi, y)_*(T_{x,x'}(M \times M \setminus \Delta))) \leq 2m$.
- $\dim(T_{z,z}\hat{\Delta}) = d$.
- $\dim(\mathbb{R}^d \times \mathbb{R}^d) = 2d$.

Hence, if $2m + d < 2d$ then there cannot exist any $(x, x') \in M \times M$ such that $\sigma(\phi, y)(x, x') \in \hat{\Delta}$, i.e., if $d \geq 2m + 1$, $f_{\phi,y}$ is injective.

Variants

- Prevalence
- Forced system
- Stochastic system

Definiton 1: Prevalence

A Borel subset S of a normed linear space V is prevalent if there is a finite-dimensional subspace E of V such that for each v in V, $v + e$ belongs to S for (Lebesgue) almost every e in E.

The space E is called a probe space.

Prevalence [2]

Theorem 3: Fractal Takens embedding prevalence theorem

Let ϕ be a diffeomorphism on an open subset U of $\mathbb{R}^k$, A be a compact subset of U of box-counting dimension d.
Assume that for every $p \leq 2m+1$, the set of periodic points of period p has box-dimension $< p/2$ and that each of these points has distinct eigenvalues.
Then, for a prevalent subset of functions y on U, the map $f_{\phi,y}$ is injective on A, and an immersion on each compact subset of a smooth manifold contained in A.

Forced system [5]

$$x_{t+1} = \alpha(x_t, u_t),$$
$$u_{t+1} = \beta(u_t).$$

Theorem 4: Forced Takens theorem

Let M and N be a compact manifolds of dimension $m \geq 1$ and n.
Let $\alpha \in \mathit{Diff}^1(M)$, $\beta \in \mathit{Diff}^1(N)$ and $y \in C^1(M, \mathbb{R})$.
Suppose that the periodic orbits of period $\leq 4(m+n)+1$ of β are isolated and have distinct eigenvalues.
Define the map $f_{\alpha,\beta,y} : M \to \mathbb{R}^{2m+1}$ by
$f_{\alpha,\beta,y}(x,u) = (y(\mu_0(x,u)), y(\mu_1(x,u)), \ldots, y(\mu_{2(m+n)}(x,u)))$
where $\mu_{i+1}(x,u) = \alpha(\mu_i(x,u), \beta^i(u))$ and $\mu_0(x,u) = x$.
Then, for pairs (α, y), it is a generic property that $f_{\alpha,\beta,y}$ is an embedding.

Applications

Interpretation

Takens theorem says that the image $f(M)$ of M under $f = f_{\phi,y}$ is completely equivalent to M apart from a smooth invertible change of coordinates given by f.
On $f(M) \subset \mathbb{R}^{2m+1}$, we can define a conjugate dynamics $\psi = f \circ \phi \circ f^{-1}$.
It is the same dynamical system as ϕ but in the new coordinates given by the coordinate change f.

The applications of Takens theorem include:

- Estimation of fractal dimensions
- Estimation of Lyapunov exponents
- Prediction
- Noise reduction

Applications

Prediction

Let π be the last coordinate function of $\mathbb{R}^{2m+1}$ and let $g = \pi \circ \psi$.
Let $\{y_i\}$ be a time series of observations.
Then $g(y_0, y_1, \ldots, y_{2m}) = y_{2m+1}$.
Takens theorem ensures the existence of such function g which can then be estimated from a sufficiently large sample of observations $\{y_i\}$, for example with (recurrent) neural networks.

Generalisation

Let P be a distribution of dimension p ($1 \leq p \leq m$) on M.
Assume that P is integrable.
Let X be a controlled vector field tangent to P. We see X as a vector field of $M \times [0,1]$ whose flow ϕ verifies:

- $\forall x \in M, \forall u \in [0,1], \phi_0(x,u) = (x,u)$.
- $\forall x \in M, \forall u \in [0,1], \forall t \in \mathbb{R}, \dot{\phi}_t(x,u) = X_{\phi_t(x,u)} \in P_{(\pi \circ \phi_t)(x,u)}$.
- $\forall x \in M, \forall u \in [0,1], \forall s,t \in \mathbb{R}, \phi_{s+t}(x,u) = \phi_s(\phi_t(x,u))$.

Generalisation

Let $u \in C^\infty(\mathbb{R}, [0,1])$ and $T > 0$.
Let $f_{y,\phi,u} : x \mapsto (y(x), y(\phi_T(x, u(T))), \ldots, y(\phi_{NT}(x, u(NT))))$.
Denote $f = f_{y,\phi,u}$ and define the distribution Q on $f(M) \subset \mathbb{R}^N$ by $Q_{f(x)} = f_*(P_x)$.
By Frobenius theorem: Q is integrable if and only if for every Y and Z tangent to Q we have $[Y, Z]$ is also tangent to Q.
We are looking for conditions on u and N such that, for generic (ϕ, y), the distribution Q is integrable ...

Next step: study of the input space of neural networks from the information geometry perspective, for data with a sequential structure.

References

F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence*, vol. 898, pp. 366–381, 1981.

T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *Journal of Statistical Physics*, vol. 65, pp. 579–616, 1991.

L. Noakes, "The takens embedding theorem," *International Journal of Bifurcation and Chaos*, vol. 1, no. 4, pp. 867–872, 1991.

J. P. Huke, "Embedding nonlinear dynamical systems: A guide to takens' theorem," tech. rep., Manchester Institute for Mathematical Sciences, 1993.

J. Stark, "Delay embeddings for forced systems. i. deterministic forcing," *Journal of Nonlinear Science*, vol. 9, pp. 255–332, 1999.