Identifying low-dimensional dynamics from high-dimensional observation

Corbinian Schlosser,
joint work with Francis Bach and Alessandro Rudi

February 13, 2025
L2S

Setting

We consider

$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

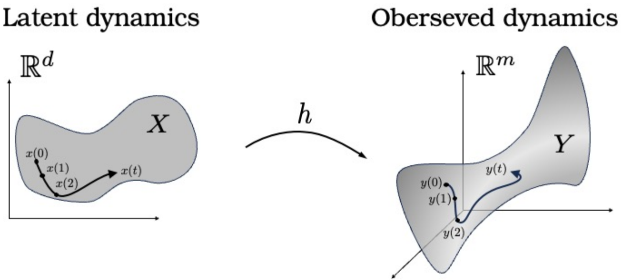
for a continuous vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and a continuous observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.

Setting

We consider

$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

for a continuous vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and a continuous observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.

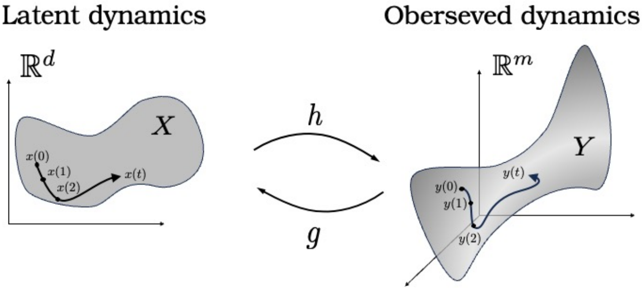


Setting

We consider

$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

for a continuous vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and a continuous observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.



Setting

We consider

$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

for a vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and an observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.

1. We observe trajectories y

Setting

We consider

$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

for a vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and an observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.

1. We observe trajectories y
2. f, h, d and x are unknown and we want to identify them.

Setting

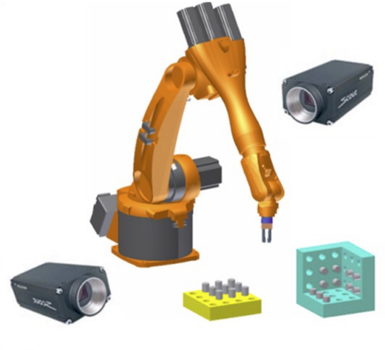
We consider

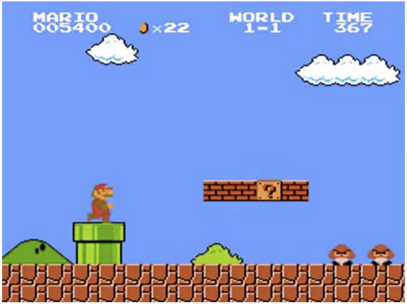
$$x(t+1) = f(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d$$
$$y = h(x)$$

for a vector field $f : \mathbb{R}^d \to \mathbb{R}^d$ and an observation map $h : \mathbb{R}^d \to \mathbb{R}^m$.

1. We observe trajectories y
2. f, h, d and x are unknown and we want to "identify" them.

Possibly possible applications





Assumptions for the rest of the talk

We assume that the maps f, g, h are unknown but satisfy

1. $f : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable
2. $h : \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable and injective (thus $m \geq d$ and typically $m \gg d$)

Assumptions for the rest of the talk

We assume that the maps f, g, h are unknown but satisfy

1. $f : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable
2. $h : \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable and injective (thus $m \geq d$ and typically $m \gg d$)
3. We call $g := h^{-1} : h(\mathbb{R}^d) \to \mathbb{R}^d$ the inverse of h

Assumptions for the rest of the talk

We assume that the maps f, g, h are unknown but satisfy

1. $f : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable
2. $h : \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable and injective (thus $m \geq d$ and typically $m \gg d$)
3. We call $g := h^{-1} : h(\mathbb{R}^d) \to \mathbb{R}^d$ the inverse of h and assume it is of the form
$$g(y) = \theta^T y$$
for some $\theta \in \mathbb{R}^{m \times d}$.

Assumptions for the rest of the talk

We assume that the maps f, g, h are unknown but satisfy

1. $f : \mathbb{R}^d \to \mathbb{R}^d$ is continuously differentiable
2. $h : \mathbb{R}^d \to \mathbb{R}^m$ is continuously differentiable and injective (thus $m \geq d$ and typically $m \gg d$)
3. We call $g := h^{-1} : h(\mathbb{R}^d) \to \mathbb{R}^d$ the inverse of h and assume it is of the form
$$g(y) = \theta^T y$$
for some $\theta \in \mathbb{R}^{m \times d}$.

We observe $y_i(t) = h(x_i(t))$, $i \in I$, at times $t \in \mathcal{T}_i \subset \mathbb{N}_0$.
We denote the set of observed points by

$$Y := \{y \in \mathbb{R}^m : y = y_i(t), i \in I, t \in \mathcal{T}_i\}.$$

Problem formulation 1

Problem (1)

For the given data $(y_i(\cdot))_{i\in I}$, *find* $(\tilde{d}, \tilde{f}, (\tilde{x}_i(\cdot))_{i\in I}, \tilde{h}, \tilde{\theta})$ *where*

1. $\tilde{d} \in \mathbb{N}$
2. $\tilde{f} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^{\tilde{d}}$ *is continuously differentiable*
3. $\tilde{x}_i : \mathcal{T}_i \to \mathbb{R}^{\tilde{d}}$ *solving* $\tilde{x}(t+1) = \tilde{f}(x(t))$, *for all* $i \in I$
4. $\tilde{\theta} \in \mathbb{R}^{m\times\tilde{d}}$ *is with* $\tilde{x}_i(\cdot) = \tilde{\theta}^T y_i(\cdot)$ *for all* $i \in I$
5. $\tilde{h} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^m$ *is continuously differentiable and injective on* $\tilde{X} := \{\tilde{\theta}^T y_i(t)) : i \in I, t \in \mathcal{T}_i\}$,

such that it holds

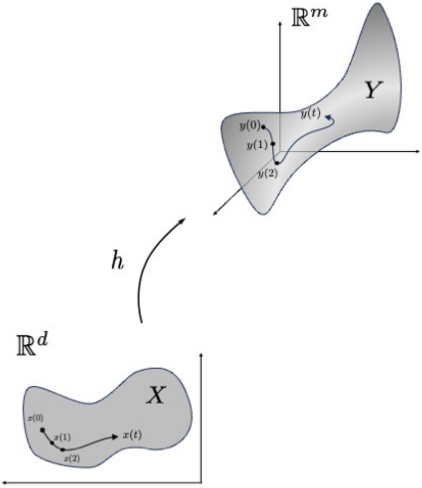
$$y_i(t) = \tilde{h} \circ \tilde{x}_i(t) \qquad \text{for all } i \in I \text{ and } t \in \mathcal{T}_i.$$

Non-uniqueness and pathological solutions

The problem formulation is coordinate-free. A change of coordinates in the latent system induces a “different solution“.

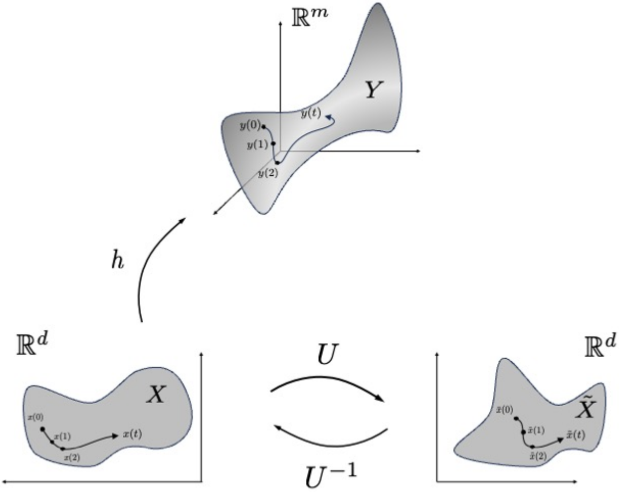
Non-uniqueness and pathological solutions

The problem formulation is coordinate-free. A change of coordinates in the latent system induces a "different solution".



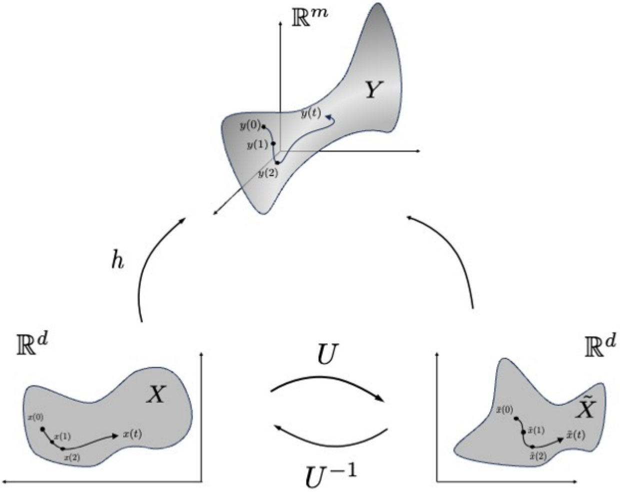
Non-uniqueness and pathological solutions

The problem formulation is coordinate-free. A change of coordinates in the latent system induces a "different solution".



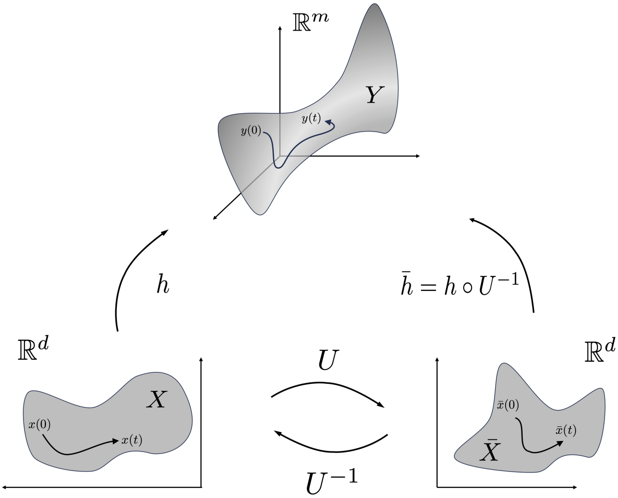
Non-uniqueness and pathological solutions

The problem formulation is coordinate-free. A change of coordinates in the latent system induces a “different solution“.



Non-uniqueness and pathological solutions

The problem formulation is coordinate-free. A change of coordinates in the latent system induces a “different solution“.



Non-uniqueness and pathological solutions

The problem formulation is coordinate free. A change of coordinates in the latent system induces a "different solution".

If I and $\mathcal{T}_i$ are discrete, then, by interpolation, we can always find a solution with $\tilde{d} = 1$. This does not represent the topology of the latent system.

Non-uniqueness and pathological solutions

The problem formulation is coordinate free. A change of coordinates in the latent system induces a “different solution“.

If I and $\mathcal{T}_i$ are discrete, then, by interpolation, we can always find a solution with $\tilde{d} = 1$. This does not represent the topology of the latent system.

Assumption

We assume that the set $X := \{x_i(t) \in \mathbb{R}^d : i \in I, t \in \mathcal{T}_i\}$ *is open in* $\mathbb{R}^d$ *and* $\mathcal{T}_i = [0, T]$ *for all* $i \in I$.

Problem formulation 1'

Problem (1')

For the given data $(y_i(\cdot))_{i\in I}$, *find* $(\tilde{d}, \tilde{f}, (\tilde{x}_i(\cdot))_{i\in I}, \tilde{h}, \tilde{\theta})$ *where*

1. $\tilde{d} \in \mathbb{N}$
2. $\tilde{f} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^{\tilde{d}}$ *is continuously differentiable*
3. $\tilde{x}_i : \mathcal{T}_i \to \mathbb{R}^{\tilde{d}}$ *solving* $\tilde{x}(t+1) = \tilde{f}(x(t))$, *for all* $i \in I$
4. $\tilde{\theta} \in \mathbb{R}^{m\times\tilde{d}}$ *is with* $\tilde{x}_i(\cdot) = \tilde{\theta}^T y_i(\cdot)$ *for all* $i \in I$
5. $\tilde{h} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^m$ *is continuously differentiable and injective on* $\tilde{X} := \{\tilde{\theta}^T y_i(t)) : i \in I, t \in [0,T]\}$ *open*,

such that it holds

$$y_i(t) = \tilde{h} \circ \tilde{x}_i(t) \quad \text{for all } i \in I \text{ and } t \in [0,T].$$

Problem formulation 1'

Problem (1')

For the given data $(y_i(\cdot))_{i\in I}$, *find* $(\tilde{d}, \tilde{f}, (\tilde{x}_i(\cdot))_{i\in I}, \tilde{h}, \tilde{\theta})$ *where*

1. $\tilde{d} \in \mathbb{N}$
2. $\tilde{f} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^{\tilde{d}}$ *is continuously differentiable*
3. $\tilde{x}_i : \mathcal{T}_i \to \mathbb{R}^{\tilde{d}}$ *solving* $\tilde{x}(t+1) = \tilde{f}(x(t))$, *for all* $i \in I$
4. $\tilde{\theta} \in \mathbb{R}^{m\times\tilde{d}}$ *is with* $\tilde{x}_i(\cdot) = \tilde{\theta}^T y_i(\cdot)$ *for all* $i \in I$
5. $\tilde{h} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^m$ *is continuously differentiable and injective on* $\tilde{X} := \{\tilde{\theta}^T y_i(t)) : i \in I, t \in [0,T]\}$ *open*,

such that it holds

$$y_i(t) = \tilde{h} \circ \tilde{x}_i(t) \qquad \text{for all } i \in I \text{ and } t \in [0,T].$$

Lemma

For any solution of Problem 1' it holds $\tilde{d} = d$.

Solving for $(\tilde{d}, \tilde{f}, (\tilde{x}_i(\cdot))_{i \in I}, \tilde{h}, \tilde{\theta})$

How to search for $(\tilde{d}, \tilde{f}, (\tilde{x}_i(\cdot))_{i \in I}, \tilde{h}, \tilde{\theta})$ that solve Problem 1'?

A good θ is "almost" sufficient

Proposition

Let $\tilde{d} \in \mathbb{N}$ and $\tilde{g} : \mathbb{R}^m \to \mathbb{R}^{\tilde{d}}$ be continuously differentiable, injective on Y and such that $\tilde{g}(Y) \subset \mathbb{R}^{\tilde{d}}$ is open. Then $\tilde{d} = d$ and there exists $\tilde{f} : \mathbb{R}^d \to \mathbb{R}^d$, $\tilde{h} : \mathbb{R}^d \to \mathbb{R}^m$ continuous, such that for $i \in I$ and $t \in [0, T]$, $\tilde{x}_i(t) := \tilde{g}(y_i(t))$ it holds

$$\begin{aligned} \tilde{x}_i(t+1) &= \tilde{f}(\tilde{x}_i(t)) \\ y_i(t) &= \tilde{h}(\tilde{x}_i(t)). \end{aligned}$$

A good θ is "almost" sufficient

Proposition

Let $\tilde{d} \in \mathbb{N}$ and $\tilde{g} : \mathbb{R}^m \to \mathbb{R}^{\tilde{d}}$ be continuously differentiable, injective on Y and such that $\tilde{g}(Y) \subset \mathbb{R}^{\tilde{d}}$ is open. Then $\tilde{d} = d$ and there exists $\tilde{f} : \mathbb{R}^d \to \mathbb{R}^d$, $\tilde{h} : \mathbb{R}^d \to \mathbb{R}^m$ continuous, such that for $i \in I$ and $t \in [0, T]$, $\tilde{x}_i(t) := \tilde{g}(y_i(t))$ it holds

$$\begin{aligned} \tilde{x}_i(t+1) &= \tilde{f}(\tilde{x}_i(t)) \\ y_i(t) &= \tilde{h}(\tilde{x}_i(t)). \end{aligned}$$

Proof: *Sketch.* By the invariance of domain it holds $\tilde{d} = d$. For the rest, set $\tilde{h} = \tilde{g}\big|_Y^{-1}$ and extend continuously to $\mathbb{R}^d$ and

$$\tilde{f} := \tilde{g} \circ h \circ f \circ g \circ \tilde{h}.$$

□

A good θ is "almost" sufficient

Proposition

Let $\tilde{d} \in \mathbb{N}$ *and* $\tilde{g} : \mathbb{R}^m \to \mathbb{R}^{\tilde{d}}$ *be continuously differentiable, injective on* Y *and such that* $\tilde{g}(Y) \subset \mathbb{R}^{\tilde{d}}$ *is open. Then* $\tilde{d} = d$ *and there exists* $\tilde{f} : \mathbb{R}^d \to \mathbb{R}^d$, $\tilde{h} : \mathbb{R}^d \to \mathbb{R}^m$ *continuous, such that for* $i \in I$ *and* $t \in [0, T]$, $\tilde{x}_i(t) := \tilde{g}(y_i(t))$ *it holds*

$$\begin{aligned} \tilde{x}_i(t+1) &= \tilde{f}(\tilde{x}_i(t)) \\ y_i(t) &= \tilde{h}(\tilde{x}_i(t)). \end{aligned}$$

Proof: *Sketch.* By the invariance of domain it holds $\tilde{d} = d$. For the rest, set $\tilde{h} = \tilde{g}\big|_Y^{-1}$ and extend continuously to $\mathbb{R}^d$ and

$$\tilde{f} := \tilde{g} \circ h \circ f \circ g \circ \tilde{h}.$$

□

A good θ is "almost" sufficient

Proposition

Let $\tilde{d} \in \mathbb{N}$ and $\tilde{g} : \mathbb{R}^m \to \mathbb{R}^{\tilde{d}}$ be continuously differentiable, injective on Y and such that $\tilde{g}(Y) \subset \mathbb{R}^{\tilde{d}}$ is open. Then $\tilde{d} = d$ and there exists $\tilde{f} : \mathbb{R}^d \to \mathbb{R}^d$, $\tilde{h} : \mathbb{R}^d \to \mathbb{R}^m$ continuous, such that for $i \in I$ and $t \in [0, T]$, $\tilde{x}_i(t) := \tilde{g}(y_i(t))$ it holds

$$\begin{aligned} \dot{\tilde{x}}_i(t) &= \tilde{f}(\tilde{x}_i(t)) \\ y_i(t) &= \tilde{h}(\tilde{x}_i(t)). \end{aligned}$$

Proof: *Sketch.* By the invariance of domain it holds $\tilde{d} = d$. For the rest, set $\tilde{h} = \tilde{g}\big|_Y^{-1}$ and extend continuously to $\mathbb{R}^d$ and

$$\tilde{f} := \tilde{g} \circ h \circ f \circ g \circ \tilde{h}.$$

□

Learning $\tilde{f}, \tilde{h}$

Once we have chosen a candidate $\tilde{\theta} \in \mathbb{R}^{m \times \tilde{d}}$ (i.e. $\tilde{g}(y) = \tilde{\theta}^T y$) estimating $\tilde{f}$ and $\tilde{h}$ is a learning task:

Find $\tilde{h} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^m$ and $\tilde{f} : \mathbb{R}^{\tilde{d}} \to \mathbb{R}^{\tilde{d}}$ with

$$\tilde{h}(\tilde{\theta}^T y(t)) = y(t) \quad \text{and} \quad \tilde{\theta}^T y(t+1) = \tilde{f}(\tilde{\theta}^T y(t))$$

for all observations $y(t)$.

Solving for θ first

Solving for θ first

From $x(t+1) = f(x(t))$ and $y = h(x)$ and $x = \theta^T y$ we infer

$$y(t+1) = h(f(x(t))) = h(f(\theta^T y(t))) =: A(y(t)).$$

Solving for θ first

From $x(t+1) = f(x(t))$ and $y = h(x)$ and $x = \theta^T y$ we infer

$$y(t+1) = h(f(x(t))) = h(f(\theta^T y(t))) =: A(y(t)).$$

We formulate the question

Can we identify θ from A?

Solving for θ first

From $x(t+1) = f(x(t))$ and $y = h(x)$ and $x = \theta^T y$ we infer

$$y(t+1) = h(f(x(t))) = h(f(\theta^T y(t))) =: A(y(t)).$$

We formulate the question

Can we identify θ from A?

No! The vector field A is not unique.

Non-uniqueness of A; an example

Consider the following one-dimensional linear system

$$x(t+1) = -\frac{1}{2}x(t)$$

with observations $y = (x, x)^T \in \mathbb{R}^2$. Then we have

$$\begin{pmatrix} \frac{1}{2}x(t) \\ \frac{1}{2}x(t) \end{pmatrix} = \quad y(t{+}1)$$

Non-uniqueness of A; an example

Consider the following one-dimensional linear system

$$x(t+1) = -\frac{1}{2}x(t)$$

with observations $y = (x, x)^T \in \mathbb{R}^2$. Then we have

$$\begin{aligned} A_1(y(t)) &:= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y(t) \\ &= y(t+1) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} y =: A_2(y(t)) \end{aligned}$$

Non-uniqueness of A; an example

Consider the following one-dimensional linear system

$$x(t+1) = -\frac{1}{2}x(t)$$

with observations $y = (x, x)^T \in \mathbb{R}^2$. Then we have

$$\begin{aligned} A_1(y(t)) &:= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} y(t) \\ &= y(t+1) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} y =: A_2(y(t)) \end{aligned}$$

And only one of the vector fields A_1, A_2 has a low-dimensional structure.

A closer look at $y(t+1)$

We have

$$y(t+1) = h(f(x(t)) = h(f(\theta^T y(t))) = A(y) = \bar{A}(\theta^T y)$$

for

$$\bar{A}(z) := h(f(z)).$$

A closer look at $y(t+1)$

We have

$$y(t+1) = h(f(x(t)) = h(f(\theta^T y(t))) = A(y) = \bar{A}(\theta^T y)$$

for

$$\bar{A}(z) := h(f(z)).$$

For this choice of A, we have

$$\mathrm{D}A(y) = \mathrm{D}\bar{A}(\theta^T y)\theta^T.$$

A closer look at $y(t+1)$

We have

$$y(t+1) = h(f(x(t)) = h(f(\theta^T y(t))) = A(y) = \bar{A}(\theta^T y)$$

for

$$\bar{A}(z) := h(f(z)).$$

For this choice of A, we have

$$\mathrm{D}A(y) = \mathrm{D}\bar{A}(\theta^T y)\theta^T.$$

We can access $\tilde{\theta}$ from $\mathrm{D}A(y)$!

Proposition

$A \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^m)$ *and* $\tilde{\theta} \in \mathbb{R}^{m \times \tilde{d}}$ *such that*

1. *For all* $y \in Y$ *it holds* $y(t+1) = A(y(t))$
2. *For all* $z \in \mathrm{conv}(Y)$*:* $\mathrm{D}A(z) = M(z)\tilde{\theta}^T$ *with* $M(z) \in \mathbb{R}^{m \times d}$
3. *For* $y_1, y_2 \in Y$ *we have*

$$y_1(t) \neq y_2(t) \text{ implies } y_1(t+1) \neq y_2(t+1).$$

Then the map $y \mapsto \tilde{\theta}^T y$ *is injective on* Y*.*

Proposition

$A \in \mathcal{C}^1(\mathbb{R}^m; \mathbb{R}^m)$ and $\tilde{\theta} \in \mathbb{R}^{m\times\tilde{d}}$ such that

1. *For all $y \in Y$ it holds $y(t+1) = A(y(t))$*
2. *For all $z \in \operatorname{conv}(Y)$: $\mathrm{D}A(z) = M(z)\tilde{\theta}^T$ with $M(z) \in \mathbb{R}^{m\times d}$*
3. *For $y_1, y_2 \in Y$ we have*

$$y_1(t) \neq y_2(t) \text{ implies } y_1(t+1) \neq y_2(t+1).$$

Then the map $y \mapsto \tilde{\theta}^T y$ is injective on Y.

Proof: Let $y_1, y_2 \in Y$ with $\theta^T y_1 = \tilde{\theta}^T y_2$. We have

$$\begin{aligned} y_2(t+1) - y_1(t+1) &= A(y_2(t)) - A(y_1(t)) \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{ds}} A(y_1 + s(y_2 - y_1))\, ds \\ &= \int_0^1 M(y_1 + t(y_2 - y_1)) \underbrace{\tilde{\theta}^T (y_2 - y_1)}_{=0}\, dt = 0. \end{aligned}$$

By the third condition in the statement, we conclude $y_1 = y_2$. □

Finding θ: A convex optimization approach

Recall

$$\mathrm{D}A(y) = \mathrm{D}\bar{A}(\theta^T y)\theta^T \quad \text{for all } y,$$

Finding θ: A convex optimization approach

Recall

$$\mathrm{D}A(y) = \mathrm{D}\bar{A}(\theta^T y)\theta^T \quad \text{for all } y,$$

and, hence,

$$\mathrm{ran}((\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T) \subset \mathrm{ran}\ \theta$$

for points $w_1, \ldots, w_N \in \mathbb{R}^m$.

Finding θ: An optimization approach

Finding θ: An optimization approach

For points $w_1, \ldots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \mathrm{rank}((\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)) \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

where $\mathcal{F}$ is a class of candidate functions from $\mathbb{R}^m$ to $\mathbb{R}^m$.

Finding θ: An optimization approach

For points $w_1, \dots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{aligned} \min_{A \in \mathcal{F}} \quad & \mathrm{rank}((\mathrm{D}A(w_1)^T, \dots, \mathrm{D}A(w_N)^T)) \\ \text{s.t.} \quad & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{aligned}$$

where $\mathcal{F}$ is a class of candidate functions from $\mathbb{R}^m$ to $\mathbb{R}^m$.

Choosing $\tilde{d}, \tilde{\theta}$

1. Let A^* be a minimizer
2. (U, Σ, V) be the singular value decomposition of $(\mathrm{D}A(w_1)^T, \dots, \mathrm{D}A(w_N)^T)$
3. Choose $\tilde{d} \in \mathbb{N}$ such that $\sigma_{\tilde{d}} \gg \sigma_{\tilde{d}+1}$ and $\sigma_{\tilde{d}+1} \ll 1$
4. Choose $\tilde{\theta}$ to be the first $\tilde{d}$ columns of U.

Finding θ: An optimization approach

For points $w_1, \ldots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \mathrm{rank}((\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)) \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

where $\mathcal{F}$ is a class of candidate functions from $\mathbb{R}^m$ to $\mathbb{R}^m$.

Choosing $\tilde{d}, \tilde{\theta}$

1. Let A^* be a minimizer
2. (U, Σ, V) be the singular value decomposition of $(\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)$
3. Choose $\tilde{d} \in \mathbb{N}$ such that $\sigma_{\tilde{d}} \gg \sigma_{\tilde{d}+1}$ and $\sigma_{\tilde{d}+1} \ll 1$
4. Choose $\tilde{\theta}$ to be the first $\tilde{d}$ columns of U.

Non-convex!

Finding θ: An optimization approach

For points $w_1, \ldots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \mathrm{rank}((\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)) \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

where $\mathcal{F}$ is a class of candidate functions from $\mathbb{R}^m$ to $\mathbb{R}^m$.

Choosing $\tilde{d}, \tilde{\theta}$

1. Let A^* be a minimizer
2. (U, Σ, V) be the singular value decomposition of $(\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)$
3. Choose $\tilde{d} \in \mathbb{N}$ such that $\sigma_{\tilde{d}} \gg \sigma_{\tilde{d}+1}$ and $\sigma_{\tilde{d}+1} \ll 1$
4. Choose $\tilde{\theta}$ to be the first $\tilde{d}$ columns of U.

Non-convex!

We use a convex relaxation.

Finding θ: A convex optimization approach

For points $w_1, \ldots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \mathrm{rank}((\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)) \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

Finding θ: A convex optimization approach

For points $w_1, \ldots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \|(\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T)\|_* \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

where $\|B\|_* = \sum_i \sigma_i(B)$ the trace norm of a matrix B.

Finding θ: A convex optimization approach

For points $w_1, \dots, w_N \in \mathrm{conv}(Y)$ consider

$$\begin{array}{ll} \min\limits_{A \in \mathcal{F}} & \|(\mathrm{D}A(w_1)^T, \dots, \mathrm{D}A(w_N)^T)\|_* \\ \text{s.t.} & y(t+1) = A(y(t)) \text{ for all } y \in Y \text{ and } t \in \mathcal{T} \end{array}$$

where $\|B\|_* = \sum_i \sigma_i(B)$ the trace norm of a matrix B.

Convex!

Soft constraining the problem

For $(y_1, y_1^+), \ldots, (y_M, y_M^+)$ and $w_1, \ldots, w_N \in \mathrm{conv}\{y_1, \ldots, y_M\}$ we solve

$$\min_{A \in \mathcal{F}} \quad \frac{1}{M} \sum_{i=1}^{M} \|y_i^+ - A(y_i)\|_2^2 \quad + \mu \|(\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T))\|_* + R(A)$$

where μ is a penalty parameter and $R(A)$ describes a regularization penalty.

Soft constraining the problem

For $(y_1, y_1^+), \ldots, (y_M, y_M^+)$ and $w_1, \ldots, w_N \in \operatorname{conv}\{y_1, \ldots, y_M\}$ we solve

$$\min_{A \in \mathcal{F}} \quad \frac{1}{M} \sum_{i=1}^{M} \|y_i^+ - A(y_i)\|_2^2 \quad + \mu \|(\mathrm{D}A(w_1)^T, \ldots, \mathrm{D}A(w_N)^T))\|_* \\ + R(A)$$

where μ is a penalty parameter and $R(A)$ describes a regularization penalty.

Solving the optimization problem: We use a primal-dual formulation and use projected accelerated gradient descent.

Numerical examples

Example (Linear systems)

Consider

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned}$$

with $C \in \mathbb{R}^{m \times d}$ with $\ker C = \{0\}$.

Example (Linear systems)

Consider

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned}$$

with $C \in \mathbb{R}^{m\times d}$ with $\ker C = \{0\}$. $C : \mathbb{R}^d \to \operatorname{ran}(C)$ is an isomorphism.

Example (Linear systems)

Consider

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned}$$

with $C \in \mathbb{R}^{m\times d}$ with $\ker C = \{0\}$. $C : \mathbb{R}^d \to \operatorname{ran}(C)$ is an isomorphism.
Proper orthogonal decomposition: Find $V := \operatorname{ran}(C) = \operatorname{Span}\{y : y \in Y\}$, select an ONB $\tilde{\theta}_1, \ldots, \tilde{\theta}_d \in \mathbb{R}^m$ of V and set $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_d)$. This is done by SVD for the data matrix $(y_1, \ldots, y_N)$.

Example (Linear systems)

Consider

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned}$$

with $C \in \mathbb{R}^{m\times d}$ with $\ker C = \{0\}$. $C : \mathbb{R}^d \to \mathrm{ran}(C)$ is an isomorphism.

Proper orthogonal decomposition: Find $V := \mathrm{ran}(C) = \mathrm{Span}\{y : y \in Y\}$, select an ONB $\tilde{\theta}_1, \ldots, \tilde{\theta}_d \in \mathbb{R}^m$ of V and set $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_d)$. This is done by SVD for the data matrix $(y_1, \ldots, y_N)$.

Our approach: The solution of our optimization problem does something similar, except that $\tilde{\theta}$ does not consist of singular vectors.

De-Koopmanizing

Consider

Example

$$\begin{aligned} x_1(t+1) &= \frac{1}{2}x_1(t) \\ x_2(t+1) &= \frac{1}{2}x_2(t) + x_1^2(t). \end{aligned}$$

and

$$h(x_1, x_2) = (x_1, x_2, x_1^2, x_1x_2, x_1^3, x_2^2, x_1^2x_2, x_1^4).$$

De-Koopmanizing

Consider

Example

$$x_1(t+1) = \frac{1}{2}x_1(t)$$
$$x_2(t+1) = \frac{1}{2}x_2(t) + x_1^2(t).$$

and

$$h(x_1, x_2) = (x_1, x_2, x_1^2, x_1x_2, x_1^3, x_2^2, x_1^2x_2, x_1^4).$$

Our approach recovers the 2-dimensional structure of the problem and that there is only one asymptotically stable fixed point. However, the system we obtain and the original system are not diffeomorphically conjugated.

Example

$$\begin{aligned}\dot{x}_1 &= x_1 x_2\\ \dot{x}_2 &= x_2 + x_1^2\\ y &= (x_1, x_2, x_1^2, x_2^2, x_1 x_2, 3x_1 - x_2^2).\end{aligned}$$

We generated 250 random samples of initial conditions $x(0) \in [0, 2]^2$.

Singular values

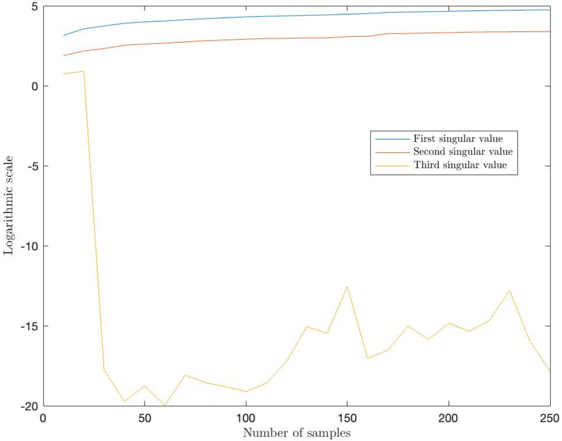
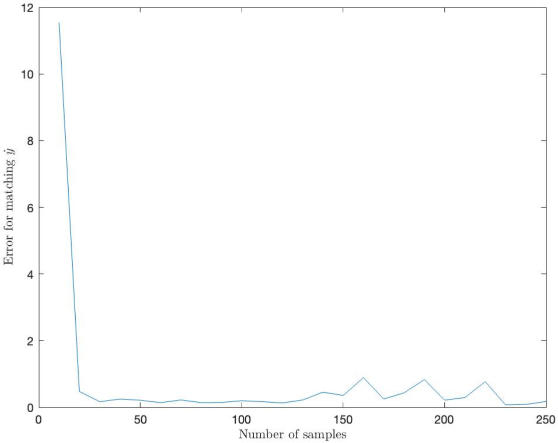


Figure: First three singular values of the matrix W for different number of sample points.

Matching $\dot{y}$



,

Figure: Error of the estimation of the state $\dot{y}$ via the constructed latent system for different number of sample points.

Matching y

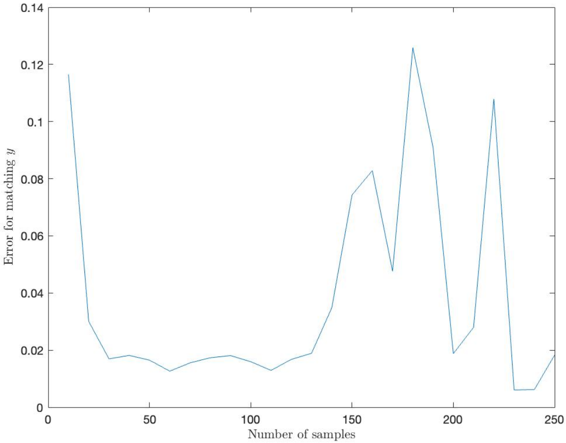


Figure: Error of the estimation of the state y via the constructed latent system for different number of sample points.

Computation time

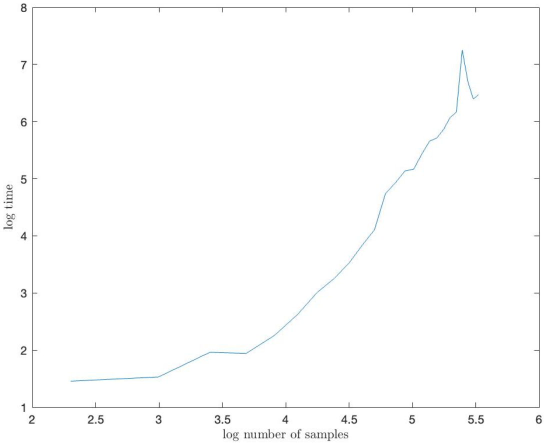


Figure: Log-log plot of computation time; solved in Matlab with CVX on a laptop.

Merci !

Merci !

Questions?

Example: Linear reconstruction

Example

Consider the system

$$\begin{aligned} \dot{x} &= Ax \\ y &= H(x) := (Cx, \bar{h}(x)) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \end{aligned}$$

for an invertible matrix $A \in \mathbb{R}^{d \times d}$, a matrix $C : \mathbb{R}^d \to \mathbb{R}^{m_1}$ with $\ker C = \{0\}$ and a differentiable map $\bar{H} : \mathbb{R}^d \to \mathbb{R}^{m_2}$.

Example: Two-dimensional spiral

Example (Two dimensional spiral)

In this example, we treat the "two-dimensional spiral":

$$\begin{aligned} \dot{x}_1 &= 1, \quad x_1(0) \in \mathbb{R} \\ \dot{x}_2 &= 1, \quad x_1(0) \in \mathbb{R} \\ y &= h(x_1, x_2) := (x_1, x_2, \cos(x_1 + x_2), \sin(x_1 + x_2)). \end{aligned}$$

For $x_1, x_2 \in \mathbb{R}$ and $(y_1, \ldots, y_4) = y = h(x_1, x_2)$ it holds

$$\dot{y} = (1, 1, -\sin(x_1 + x_2), \cos(x_1 + x_2)) = (1, 1, -y_4, y_3).$$

This motivates to define the (affine) vector field $A : \mathbb{R}^4 \to \mathbb{R}^4$ by $A(y_1, \ldots, y_4) := (1, 1, -y_4, y_3)$. But A forgets about the "height".

Lifting

When the recovery map $g : Y \to \mathbb{R}^d$ is non-linear but known to live in a function space $\mathcal{F}$ with basis $\phi = (\phi_1, \ldots, \phi_p)$, we can perform a lifting.

Lemma

Assume it holds $x = g(y)$ *for some* $g = \theta^T \phi(y)$. *Consider the extended observation map* $\bar{h} := \phi \circ h$, *i.e. we observe* $z(x) := (\phi \circ h)(x) = \phi(y)$ *system it holds*

$$x = \theta^T z.$$

Proof: It follows immediately that $x = \theta^T \phi(y) = \theta^T z$. □

Unique velocity condition

To find a good candidate θ, we imposed the "unique velocity condition":

$$y_1, y_2 \in Y \text{ with } y_1 \neq y_2 \text{ implies } y_1(t+1) \neq y_2(t+1).$$

Considering time t as an additional state $y_{m+1}(t) = t$, we can always guarantee the "unique velocity condition" through lifting.

Lemma

Let $\phi : \mathbb{R}^{m+1} \to \mathbb{R}^{3m+1}$ be given by

$$\phi(y) := (y_1, \ldots, y_{m+1}, y_1^2, \ldots, y_{m+1}^2, y_1 y_{m+1}, \ldots, y_m y_{m+1}).$$

Then the "unique velocity condition" holds for the extended observation $z = \phi(y)$.