Potsdam Institute for
Climate Impact Research

Leibniz
Association

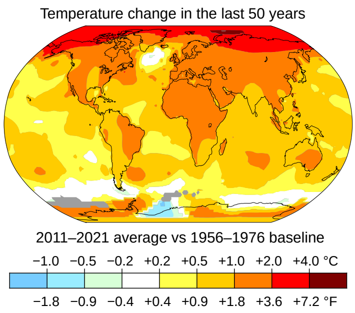
Nonparametric Estimation of Ordinary Differential Equations

Christof Schötz

PIK Future Lab: AI in the Anthropocene

March 9, 2023
German Probability and Statistics Days, Essen

- **Weather** and **climate** models



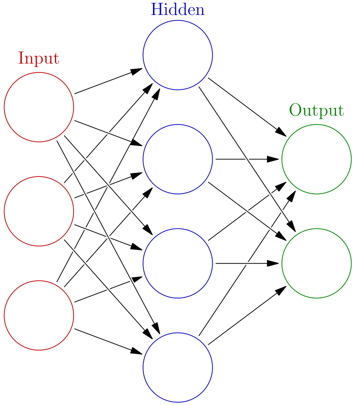
Source: Wikimedia

- **Weather** and **climate** models
- Classical physics-driven models: numerically solve Navier–Stokes partial **differential equations**

Navier–Stokes

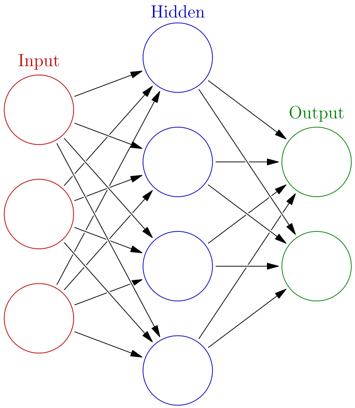
$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} - \nu \nabla^2 \mathbf{u} = -\frac{1}{\rho} \nabla p + \mathbf{g}$$

- **Weather** and **climate** models
- Classical physics-driven models: numerically solve Navier–Stokes partial **differential equations**
- New data-driven models: train **neural network** on weather observations



Source: Wikimedia

- **Weather** and **climate** models
- Classical physics-driven models:
  numerically solve Navier–Stokes
  partial **differential equations**
- New data-driven models:
  train **neural network** on weather observations
- Statisticians view:
  Neural network as **nonparametric** estimator



Source: Wikimedia

- **Weather** and **climate** models
- Classical physics-driven models:
  numerically solve Navier–Stokes
  partial **differential equations**
- New data-driven models:
  train **neural network** on weather observations
- Statisticians view:
  Neural network as **nonparametric** estimator
- **Theory** (convergence rates) of
  nonparametric differential equation estimation?

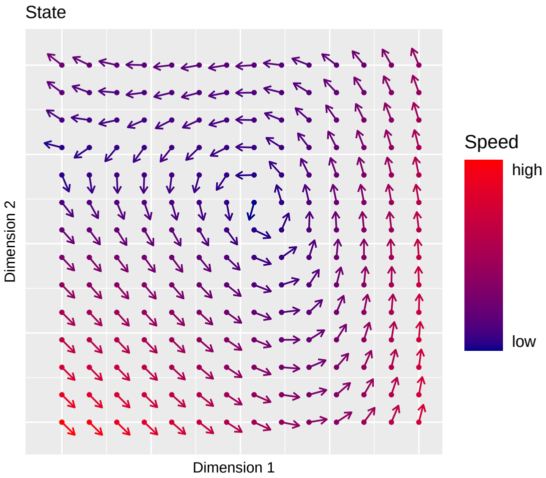
$$\left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(n^{-\frac{2\beta}{2\beta+d}}\right)$$

- **Weather** and **climate** models
- Classical physics-driven models:
  numerically solve Navier–Stokes
  partial **differential equations**
- New data-driven models:
  train **neural network** on weather observations
- Statisticians view:
  Neural network as **nonparametric** estimator
- **Theory** (convergence rates) of
  nonparametric differential equation estimation?
- Start simple:
  Autonomous first-order deterministic
  ordinary differential equation (ODE)
  with measurement noise

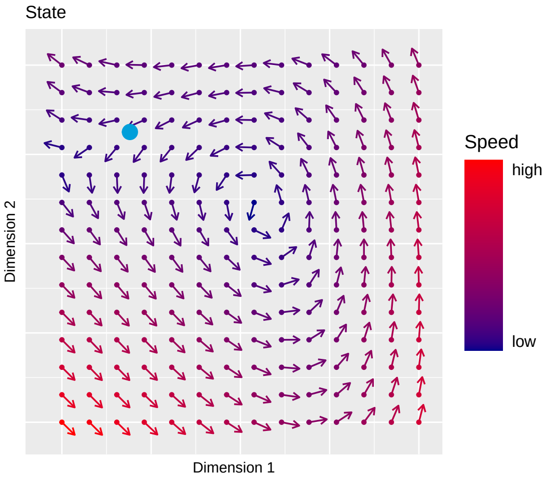
$$\dot{u} = f(u)$$

$$Y_i = u(t_i) + \epsilon_i$$

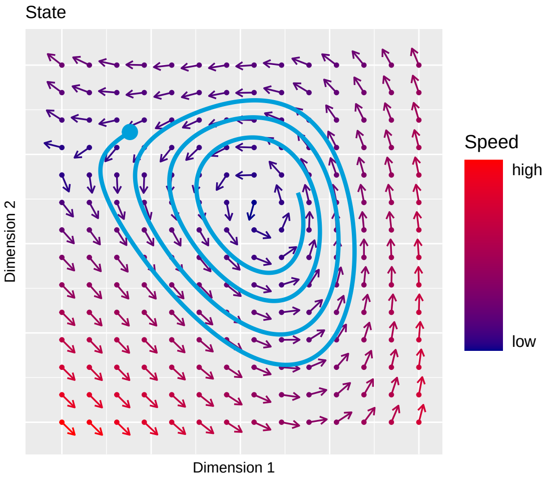
$f : \mathbb{R}^d \to \mathbb{R}^d$



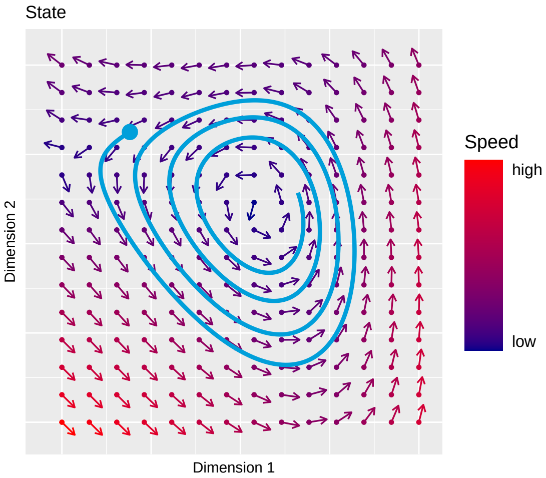
$f : \mathbb{R}^d \to \mathbb{R}^d$ $u_0 \in \mathbb{R}^d$

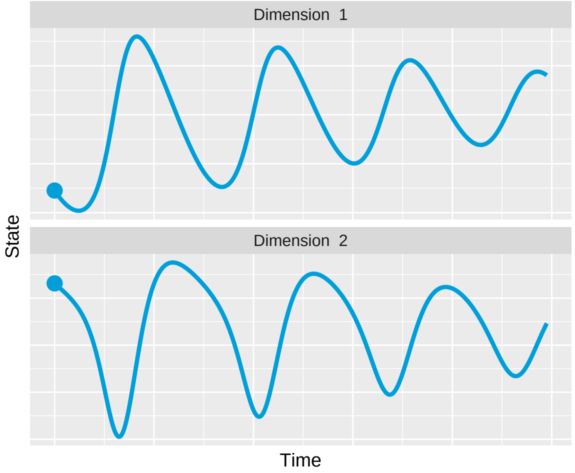


$f \colon \mathbb{R}^d \to \mathbb{R}^d$ $\qquad$ $u_0 \in \mathbb{R}^d$ $\qquad$ $T > 0$ $\qquad$ $u \colon [0, T] \to \mathbb{R}^d,\ \dot{u}(t) = f(u(t)),\ u(0) = u_0$



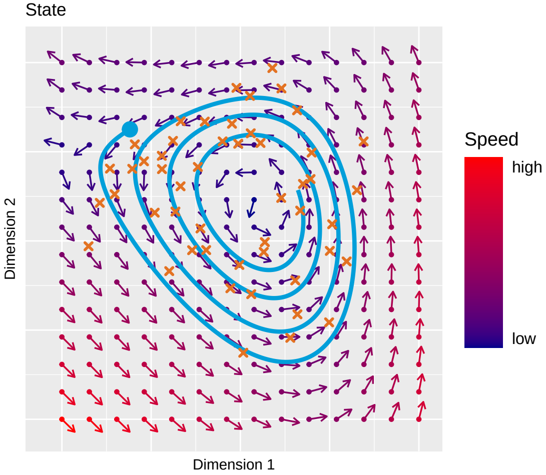
$f\colon \mathbb{R}^d \to \mathbb{R}^d$ $\qquad u_0 \in \mathbb{R}^d$ $\qquad T > 0$ $\qquad u\colon [0, T] \to \mathbb{R}^d,\ \dot{u}(t) = f(u(t)),\ u(0) = u_0$

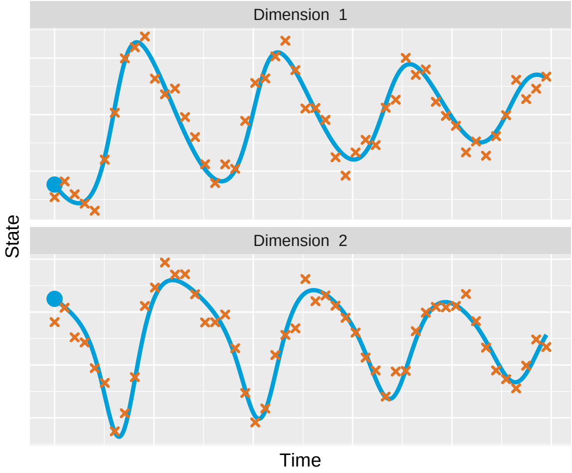




$f \colon \mathbb{R}^d \to \mathbb{R}^d$ $\quad u_0 \in \mathbb{R}^d$ $\quad T > 0$ $\quad u \colon [0, T] \to \mathbb{R}^d,\ \dot{u}(t) = f(u(t)),\ u(0) = u_0$

For $i = 1, \dots, n:$ $\quad t_i = T/n$ $\quad Y_i = u(t_i) + \epsilon_i$ $\quad \epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2 I_d)$





- $\sqrt{n}$-consistency for different estimators when f belongs to a **parametric** class *Qi and Zhao 2010; Gugushvili and Klaassen 2012; Dattner and Klaassen 2015*

- $\sqrt{n}$-consistency for different estimators when f belongs to a **parametric** class *Qi and Zhao 2010; Gugushvili and Klaassen 2012; Dattner and Klaassen 2015*
- **Nonparametric approaches** without rate of convergence:
  - *Chen et al. 2018* (Neural Networks)
  - *Heinonen et al. 2018* (Gaussian Processes)
  - *Gottwald and Reich 2021* (random feature maps + ensemble Kalman filter)

- $\sqrt{n}$-consistency for different estimators when f belongs to a **parametric** class *Qi and Zhao 2010; Gugushvili and Klaassen 2012; Dattner and Klaassen 2015*

- **Nonparametric approaches** without rate of convergence:
  - *Chen et al. 2018* (Neural Networks)
  - *Heinonen et al. 2018* (Gaussian Processes)
  - *Gottwald and Reich 2021* (random feature maps + ensemble Kalman filter)

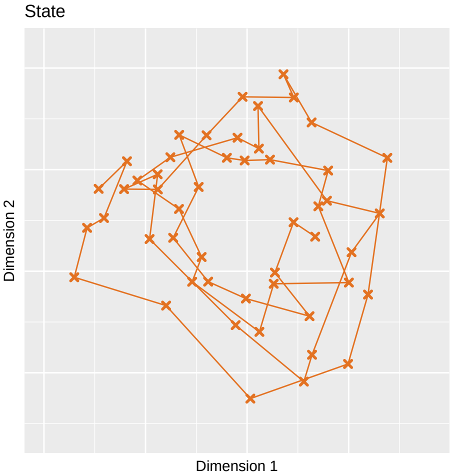
- **Goal:** Upper bound assuming $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is β-times continuously differentiable:

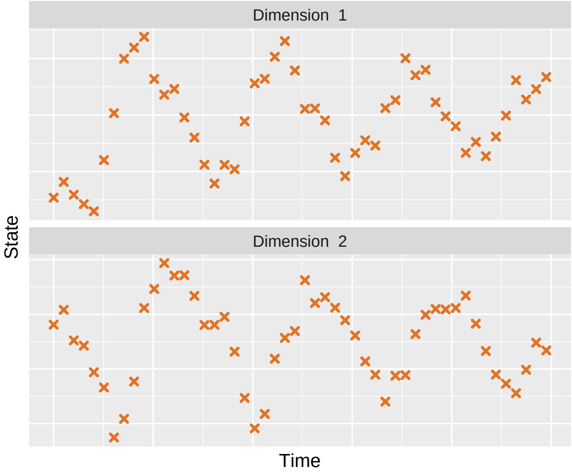
$$\left\| \hat{f}(x) - f(x) \right\| = \mathbf{O}_p\left(n^{-\alpha}\right)$$

$$\alpha = ?$$

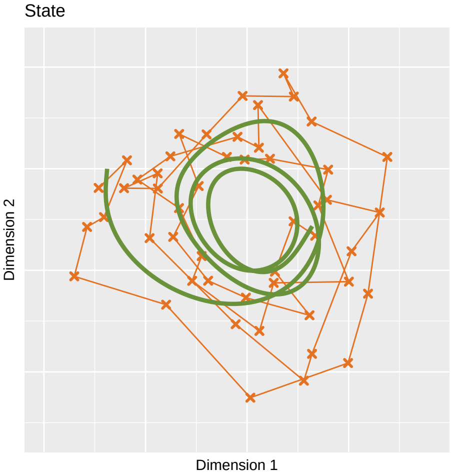
Ansatz 1:
Snake Model and Trajectory Estimation

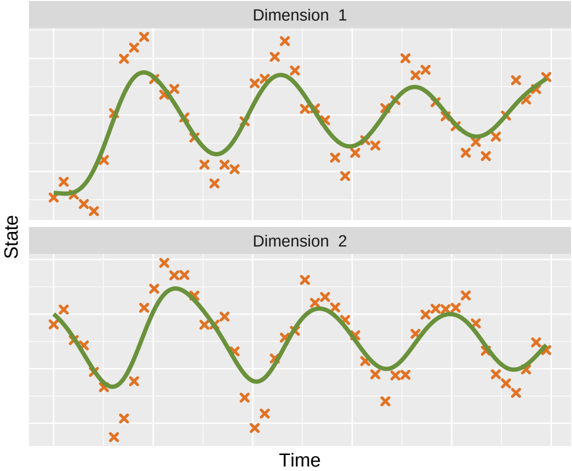
Observations $(Y_i, t_i)_{i=1,\dots,n}$



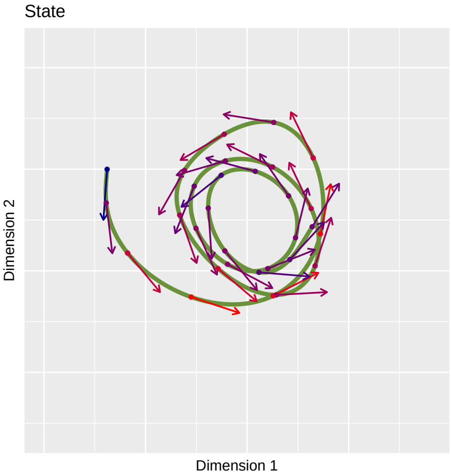


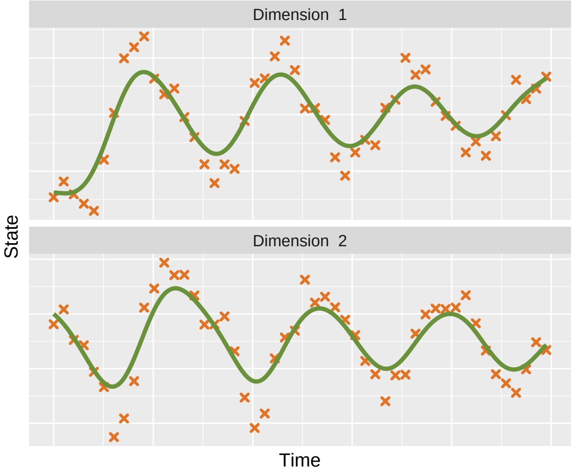
Nonparametric time–state regression, e.g., local linear $\Longrightarrow \hat{u}$



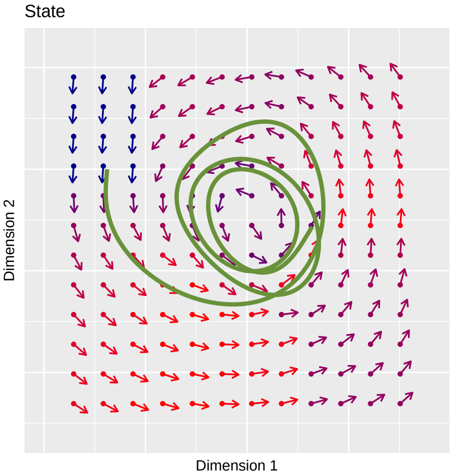


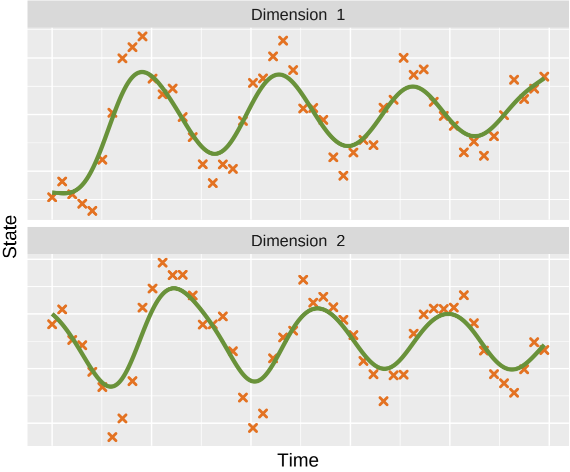
Nonparametric time–state regression, e.g., local linear with **derivative** $\Longrightarrow (\hat{u}, \hat{\dot{u}})$





$(\hat{u}, \hat{\dot{u}}) \Longrightarrow$ nearest neighbor **interpolation** $\Longrightarrow \hat{f}$





Theorem

Let

- $\hat{f}$: *trajectory estimation*
- $\delta \geq 0$

Assume f *is* $(\beta = 1)$*–smooth.*
Then

$$\sup_{x \in \mathrm{B}_\delta(u([0,T]))} \left\| \hat{f}(x) - f(x) \right\| = \mathbf{O}_p\left(\delta + T^{\frac{1}{5}} \left(\frac{n}{\log n} \right)^{-\frac{1}{5}} \right)$$

Theorem

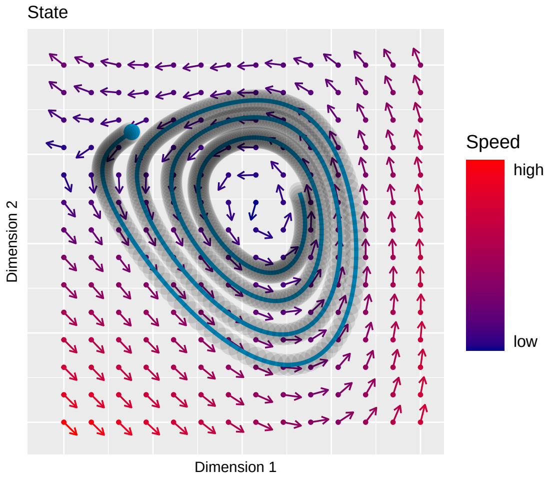
Let

- $\hat{f}$: *trajectory estimation*
- $\delta \geq 0$

Assume f *is* $(\beta = 1)$*–smooth.*
Then

$$\sup_{x \in \mathrm{B}_\delta(u([0,T]))} \left\| \hat{f}(x) - f(x) \right\| = \mathbf{O}_p\left(\delta + T^{\frac{1}{5}} \left(\frac{n}{\log n} \right)^{-\frac{1}{5}} \right)$$

Snake



Corollary (Optimal Trade-Off)

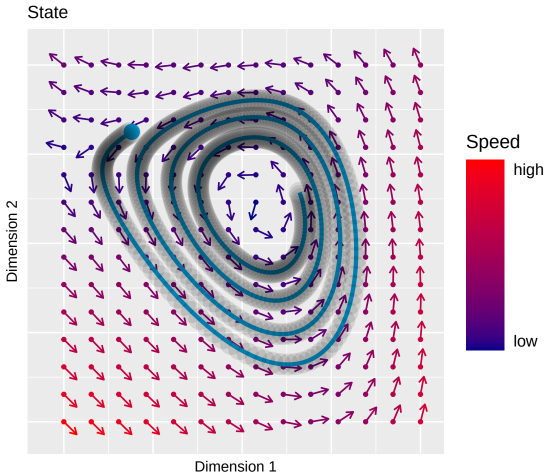
Let $\delta := \sup_{x\in[0,1]^d} \inf_{t\in[0,T]} \|u(t) - x\|$.
Assume

- f *is* $(\beta = 1)$*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{4+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$$

Snake



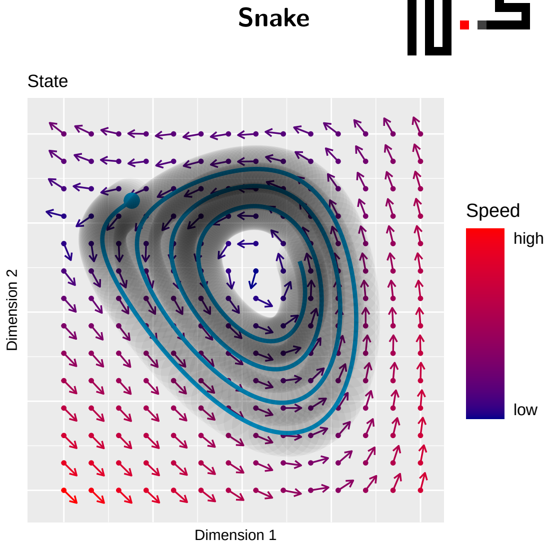
Corollary (Optimal Trade-Off)

Let $\delta := \sup_{x\in[0,1]^d} \inf_{t\in[0,T]} \|u(t) - x\|$. *Assume*

- f *is* $(\beta = 1)$*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{4+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$$



Corollary (Optimal Trade-Off)

Let $\delta := \sup_{x\in[0,1]^d} \inf_{t\in[0,T]} \|u(t) - x\|$.

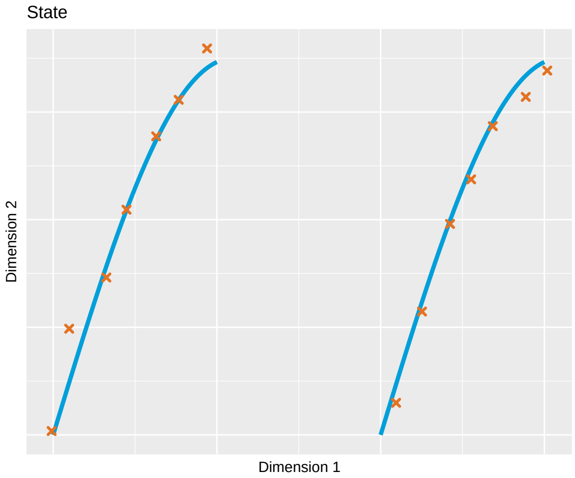
Assume

- f *is* $(\beta = 1)$*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{4+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$$

Snake



Corollary (Optimal Trade-Off)

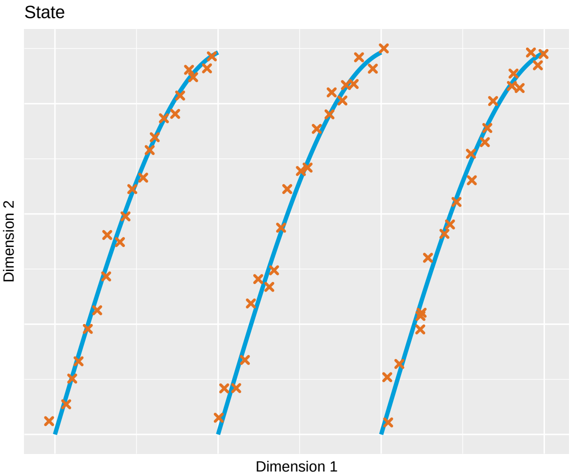
Let $\delta := \sup_{x\in[0,1]^d} \inf_{t\in[0,T]} \|u(t) - x\|$.
Assume

- f *is* $(\beta = 1)$*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{4+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$$

Snake



Corollary (Optimal Trade-Off)

Let $\delta := \sup_{x\in[0,1]^d} \inf_{t\in[0,T]} \|u(t) - x\|$.

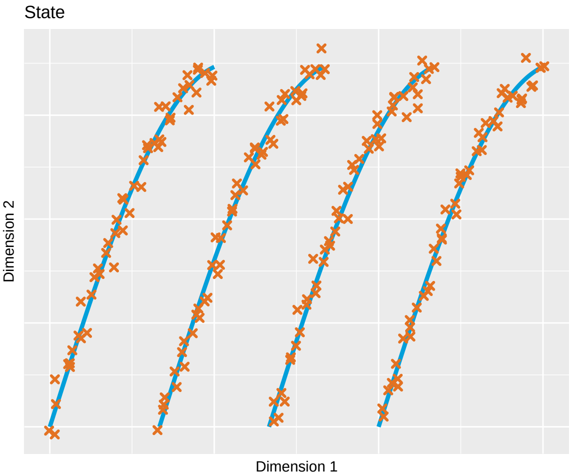
Assume

- f *is* $(\beta = 1)$*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{4+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{4+d}}\right)$$

Snake



Generalization to arbitrary $\beta \in \mathbb{N}$

Theorem

Let $\hat{f}$ be trajectory estimation with rate-optimal nonparametric estimator + polynomial interpolation. Assume

- f *is* β*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{2(\beta+1)+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{2(\beta+1)+d}}\right)$

Then

$$\sup_{x \in [0,1]^d} \left\| \hat{f}(x) - f(x) \right\| \in \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

Generalization to arbitrary $\beta \in \mathbb{N}$

Theorem

Let $\hat{f}$ be trajectory estimation with rate-optimal nonparametric estimator + polynomial interpolation. Assume

- f *is* β*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{2(\beta+1)+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{2(\beta+1)+d}}\right)$

Then

$$\sup_{x \in [0,1]^d} \left\| \hat{f}(x) - f(x) \right\| \in \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

- Same rate as estimating derivative ($r = 1$) of ($b = \beta + 1$)-smooth regression function $g \colon \mathbb{R}^d \to \mathbb{R}$

$$\left\| \hat{g}^{(r)}(x) - g^{(r)}(x) \right\| = \mathbf{O}_p\left(n^{-\frac{b-r}{2b+d}}\right)$$

Generalization to arbitrary $\beta \in \mathbb{N}$

Theorem

Let $\hat{f}$ be trajectory estimation with rate-optimal nonparametric estimator + polynomial interpolation. Assume

- f *is* β*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{2(\beta+1)+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{2(\beta+1)+d}}\right)$

Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| \in \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

- Same rate as estimating derivative ($r = 1$) of ($b = \beta + 1$)-smooth regression function $g \colon \mathbb{R}^d \to \mathbb{R}$

$$\left\|\hat{g}^{(r)}(x) - g^{(r)}(x)\right\| = \mathbf{O}_p\left(n^{-\frac{b-r}{2b+d}}\right)$$

- Awkward assumption for best rate

Generalization to arbitrary $\beta \in \mathbb{N}$

Theorem

Let $\hat{f}$ be trajectory estimation with rate-optimal nonparametric estimator + polynomial interpolation. Assume

- f *is* β*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{2(\beta+1)+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{2(\beta+1)+d}}\right)$

Then

$$\sup_{x \in [0,1]^d} \left\| \hat{f}(x) - f(x) \right\| \in \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

- Same rate as estimating derivative ($r = 1$) of ($b = \beta + 1$)-smooth regression function $g \colon \mathbb{R}^d \to \mathbb{R}$

$$\left\| \hat{g}^{(r)}(x) - g^{(r)}(x) \right\| = \mathbf{O}_p\left(n^{-\frac{b-r}{2b+d}}\right)$$

- Awkward assumption for best rate
- Extension to multiple trajectories

Generalization to arbitrary $\beta \in \mathbb{N}$

Theorem

Let $\hat{f}$ be trajectory estimation with rate-optimal nonparametric estimator + polynomial interpolation. Assume

- f *is* β*-smooth*
- $T = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{\frac{d-1}{2(\beta+1)+d}}\right)$
- $\delta = \mathbf{O}\left(\left(\frac{n}{\log n}\right)^{-\frac{1}{2(\beta+1)+d}}\right)$

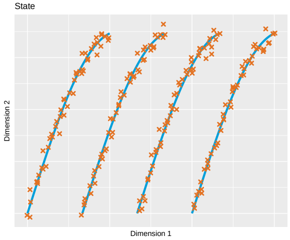
Then

$$\sup_{x\in[0,1]^d} \left\|\hat{f}(x) - f(x)\right\| \in \mathbf{O}_p\left(\left(\frac{n}{\log n}\right)^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

- Same rate as estimating derivative ($r = 1$) of ($b = \beta + 1$)-smooth regression function $g\colon \mathbb{R}^d \to \mathbb{R}$

$$\left\|\hat{g}^{(r)}(x) - g^{(r)}(x)\right\| = \mathbf{O}_p\left(n^{-\frac{b-r}{2b+d}}\right)$$

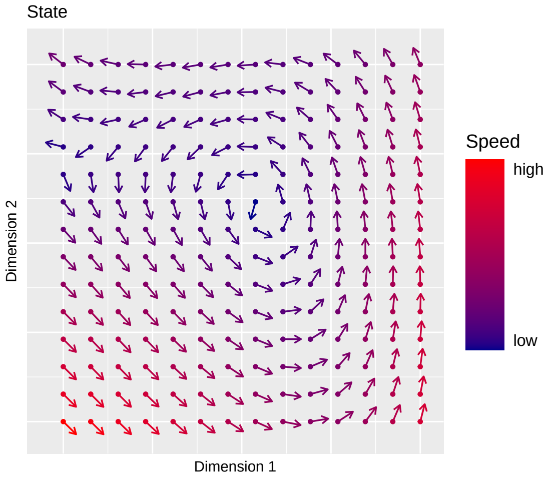
- Awkward assumption for best rate
- Extension to multiple trajectories
- State smoothness is ignored



Ansatz 2:
Stubble Model and Increment Estimation

Stubble Model:

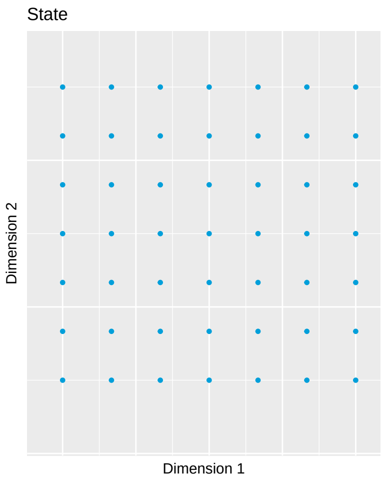
- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0,1]^d$, $i = 1, \ldots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$



Stubble Model:

- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \dots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

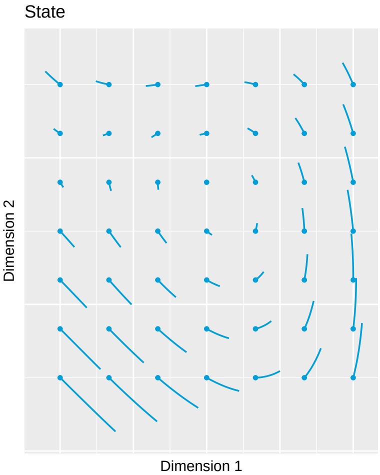
Stubble



Stubble Model:

- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \ldots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

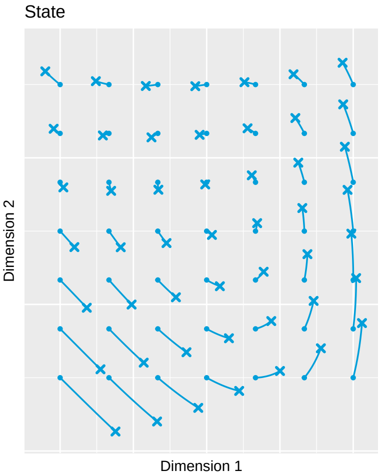
Stubble



Stubble Model:

- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \ldots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

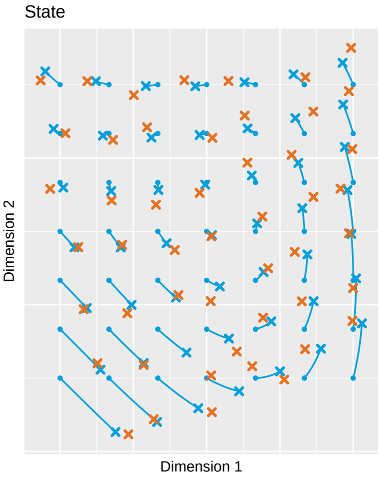
Stubble



Stubble Model:

- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \ldots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

Stubble



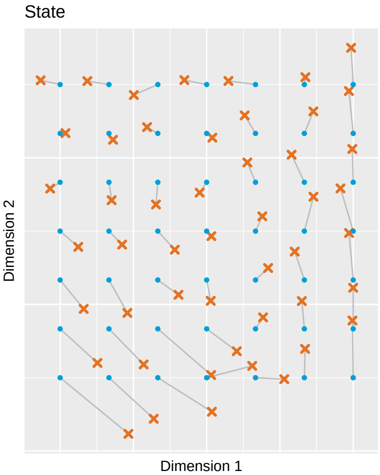
Stubble Model:

- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \ldots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

Increment Estimation:

- Estimate Increment $x \mapsto U(f, x, s) - x$
- Nadaraya–Watson estimator $\hat{\imath} \colon \mathbb{R}^d \to \mathbb{R}^d$ on data $(Y_i - x_i,\ x_i)_{i=1,\ldots,n}$

Stubble



Stubble Model:

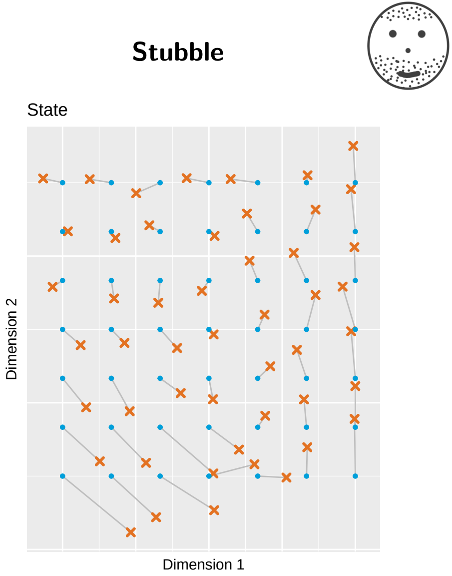
- Let $U(f, x, t)$ solve $\dot{u} = f(u), u(0) = x$
- Grid $x_i \in [0, 1]^d$, $i = 1, \dots, n$
- Step size $s > 0$
- Observation $Y_i = U(f, x_i, s) + \epsilon_i$

Increment Estimation:

- Estimate Increment $x \mapsto U(f, x, s) - x$
- Nadaraya–Watson estimator $\hat{\iota}\colon \mathbb{R}^d \to \mathbb{R}^d$ on data $(Y_i - x_i,\ x_i)_{i=1,\dots,n}$

Scaled Increment Estimator of f:

$$\hat{f}(x) := \frac{1}{s}\hat{\iota}(x)$$



Theorem

Assume f *is* $(\beta = 1)$*-smooth.*
Then, for $x_0 \in [0,1]^d$,

$$\left\| \hat{f}(x_0) - f(x_0) \right\| = \mathbf{O}_p\left(s + s^{-\frac{2}{2+d}} n^{-\frac{1}{2+d}} \right) .$$

Corollary

Assume

- f *is* $(\beta = 1)$*-smooth*
- $s = \mathbf{O}\left(n^{-\frac{1}{4+d}} \right)$

Then, for $x_0 \in [0,1]^d$,

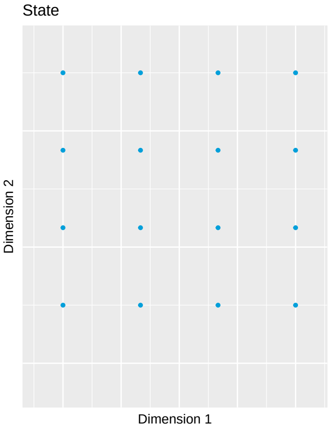
$$\left\| \hat{f}(x_0) - f(x_0) \right\| = \mathbf{O}_p\left(n^{-\frac{1}{4+d}} \right) .$$

Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$i = 1, .., n, j = 1, .., \beta$

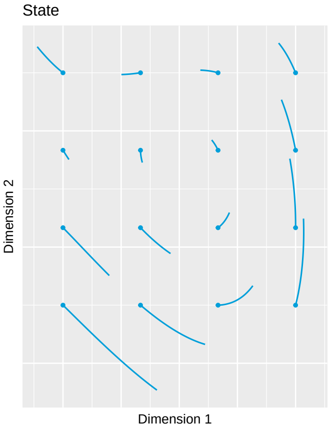


Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n, j = 1, .., \beta$$

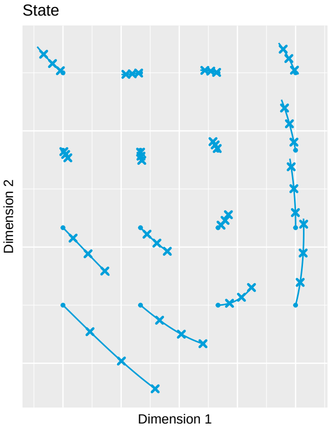


Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n, j = 1, .., \beta$$

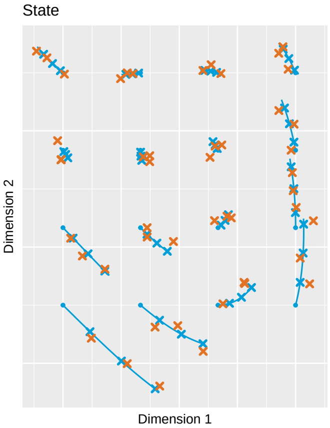


Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n \,, j = 1, .., \beta$$

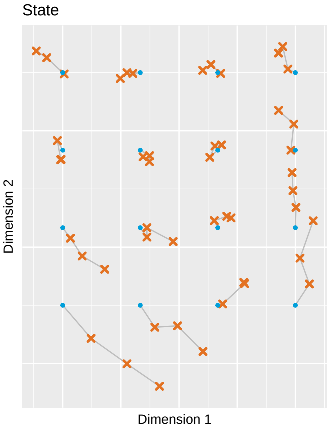


Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n, j = 1, .., \beta$$



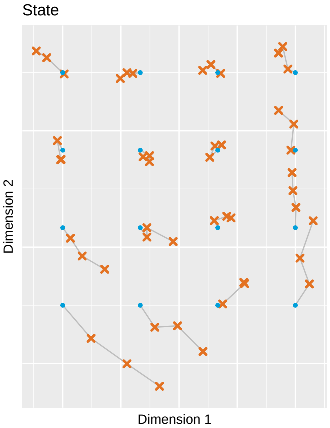
Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$i = 1, .., n, j = 1, .., \beta$

- For each $j = 1, .., \beta$:
  Estimate increment $U(f, x, js) - x$
  $\implies \quad \hat{\iota}_j$



Generalization to arbitrary $\beta \in \mathbb{N}$

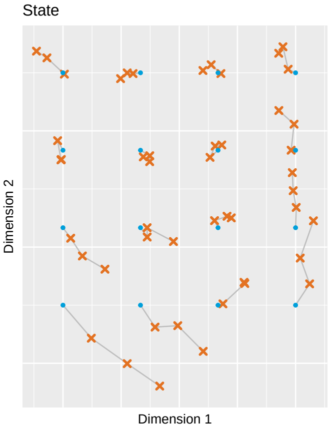
- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n\,, j = 1, .., \beta$$

- For each $j = 1, .., \beta$:
  Estimate increment $U(f, x, js) - x$
  $\implies \quad \hat{\iota}_j$

- $t \mapsto \hat{p}(x, t)$ interpolation polynomial of $(0, 0), (s, \hat{\iota}_1(x)), \ldots, (\beta s, \hat{\iota}_\beta(x))$



Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

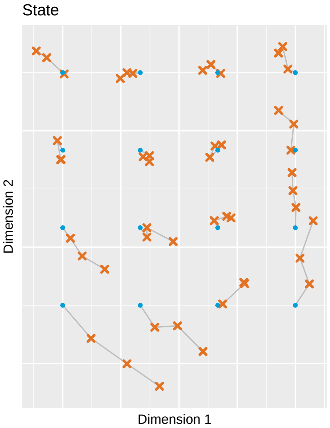
$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n\,, j = 1, .., \beta$$

- For each $j = 1, .., \beta$:
  Estimate increment $U(f, x, js) - x$
  $\implies \quad \hat{\iota}_j$

- $t \mapsto \hat{p}(x, t)$ interpolation polynomial of
  $(0, 0), (s, \hat{\iota}_1(x)), \ldots, (\beta s, \hat{\iota}_\beta(x))$

- $\hat{f}(x) := \dot{\hat{p}}(x, 0)$



Generalization to arbitrary $\beta \in \mathbb{N}$

- Observe β points at for each x_i,

$$Y_{i,j} = U(f, x_i, js) + \epsilon_{i,j}$$

$$i = 1, .., n\,, j = 1, .., \beta$$

- For each $j = 1, .., \beta$:
  Estimate increment $U(f, x, js) - x$
  $\implies \quad \hat{\iota}_j$

- $t \mapsto \hat{p}(x, t)$ interpolation polynomial of $(0, 0), (s, \hat{\iota}_1(x)), \dots, (\beta s, \hat{\iota}_\beta(x))$

- $\hat{f}(x) := \dot{\hat{p}}(x, 0)$

Theorem

Assume

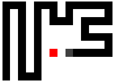
- f *is* β*-smooth*
- $s = \mathbf{O}\left(n^{-\frac{1}{2(\beta+1)+d}}\right)$

Then, for $x_0 \in [0, 1]^d$,

$$\left\|\hat{f}(x_0) - f(x_0)\right\| = \mathbf{O}_p\left(n^{-\frac{\beta}{2(\beta+1)+d}}\right) .$$

Conclusion

Conclusion

Snake 

- few long trajectories
- cover area of interest by trajectory
- estimate trajectory (time $\rightarrow$ state)
- interpolate in state space (state $\rightarrow$ state)

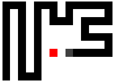
Stubble 

- many short trajectories
- cover area of interest by initial conditions
- estimate increments (state $\rightarrow$ state)
- interpolate in time (time $\rightarrow$ state)

$$\left\| \hat{f}(x) - f(x) \right\| = \mathbf{O}_p\left(n^{-\frac{\beta}{2(\beta+1)+d}} \right)$$

Conclusion

Snake



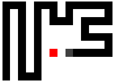
- few long trajectories
- cover area of interest by trajectory
- estimate trajectory (time $\to$ state)
- interpolate in state space (state $\to$ state)

Stubble

- many short trajectories
- cover area of interest by initial conditions
- estimate increments (state $\to$ state)
- interpolate in time (time $\to$ state)

$$\left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(n^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

Conjecture: This is the Minimax-Rate for β-Hölder smooth f.

Snake 

- few long trajectories
- cover area of interest by trajectory
- estimate trajectory (time $\to$ state)
- interpolate in state space (state $\to$ state)

Stubble 

- many short trajectories
- cover area of interest by initial conditions
- estimate increments (state $\to$ state)
- interpolate in time (time $\to$ state)

$$\left\|\hat{f}(x) - f(x)\right\| = \mathbf{O}_p\left(n^{-\frac{\beta}{2(\beta+1)+d}}\right)$$

Conjecture: This is the Minimax-Rate for β-Hölder smooth f.

Thanks!

Slides: https://github.com/chroetz/GPSD23slides/

Chen, R. T. Q., Y. Rubanova, J. Bettencourt, and D. Duvenaud (2018). "Neural Ordinary Differential Equations". In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. NIPS'18. Montréal, Canada: Curran Associates Inc., pp. 6572–6583.

Dattner, I. and C. Klaassen (Jan. 2015). "Optimal rate of direct estimators in systems of ordinary differential equations linear in functions of the parameters". In: *Electronic Journal of Statistics* 9.2, pp. 1939–1973.

Gottwald, G. and S. Reich (2021). "Supervised learning from noisy observations: Combining machine-learning techniques with data assimilation". In: *Physica D: Nonlinear Phenomena* 423, p. 132911.

Gugushvili, S. and C. Klaassen (2012). "$\sqrt{n}$-consistent parameter estimation for systems of ordinary differential equations: bypassing numerical integration via smoothing". In: *Bernoulli* 18.3, pp. 1061–1098.

Heinonen, M., C. Yildiz, H. Mannerström, J. Intosalmi, and H. Lähdesmäki (2018). "Learning unknown ODE models with Gaussian processes". In: *International Conference on Machine Learning*. PMLR, pp. 1959–1968.

Qi, X. and H. Zhao (Feb. 1, 2010). "Asymptotic efficiency and finite-sample properties of the generalized profiling estimation of parameters in ordinary differential equations". In: *The Annals of Statistics* 38.1. arXiv: `0903.3400[math,stat]`.