Data-driven approximation of differential inclusions and application to detection of transportation modes

Pierre-Cyril Aubin-Frankowski, Nicolas Petit

PhD student (École des Ponts ParisTech),
CAS - Centre Automatique et Systèmes, MINES ParisTech

ECC 2020, May 2020

Table of Contents

Problem formulation: approximating sets

Consider $N_0 > 1$ forced/controlled systems (f_i, U_i), with $u(\cdot) \in U_i$

$$q'(t) = f_i(q(t), u(t)) \in \mathbb{R}^n$$

How to identify the type (label i) of a given trajectory $q(\cdot)$? What if we do not know $u(\cdot)$, nor U_i, nor f_i? Consider the couples (q, q')

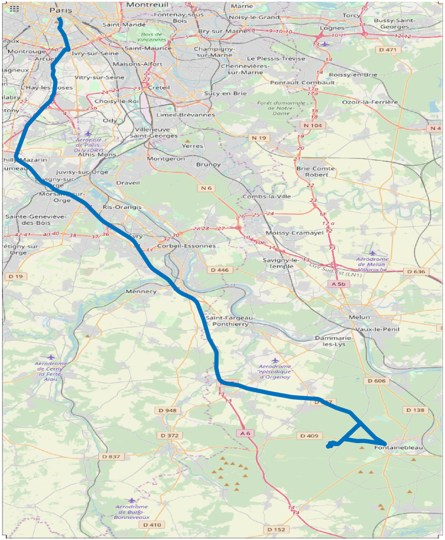
$$q'(t) \in F_i(q(t)) := \{f_i(q(t), u(t)) \,|\, u(\cdot) \in U_i\}$$
$$K_i := \{(q, q') \,|\, q' \in F_i(q)\} \subset \mathbb{R}^n \times \mathbb{R}^n$$

The set-valued map F_i is identified with its graph K_i. Recall that

control systems $\subset$ differential inclusions

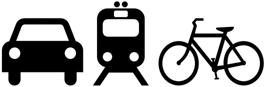
Assume that for each value of i, *labeled* samples of $(q(\cdot), q'(\cdot))$ are available. How to approximate the sets K_i?

Motivation: Detection of transportation modes



Based on smartphone information such as GPS data, what is my mode of travel and when have I changed?

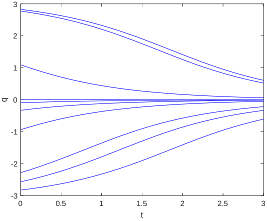
I do not know the controls applied, nor the equations for the vehicles I used.



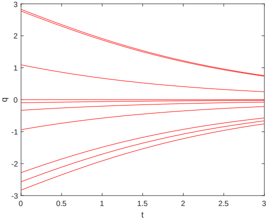
A toy example of nonlinear system without inputs

$$q'(t) = -sin(\omega_i \cdot q(t)) \in \mathbb{R}$$

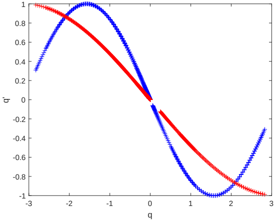
Generate ten trajectories for $\omega_1 = 1$ and $\omega_2 = 1/2$, plot the couples (q, q'),



(t, q) with $\omega_1 = 1$



(t, q) with $\omega_2 = 1/2$

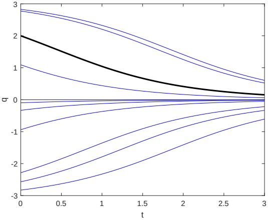


(q, q') phase plane

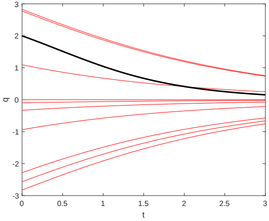
A toy example of nonlinear system without inputs

$$q'(t) = -sin(\omega_i \cdot q(t)) \in \mathbb{R}$$

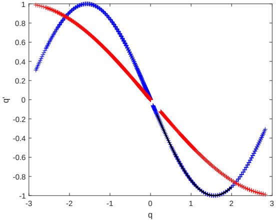
Generate ten trajectories for $\omega_1 = 1$ and $\omega_2 = 1/2$, plot the couples (q, q'), and consider a new trajectory of unknown type, guessing the type is **easy** in phase space



(t, q) with $\omega_1 = 1$



(t, q) with $\omega_2 = 1/2$

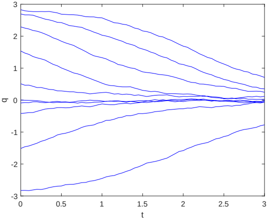


(q, q') phase plane

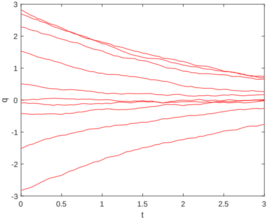
A toy example of nonlinear system with inputs

$$q'(t) = -sin(\omega_i \cdot q(t)) + u(t) \text{ with } |u(t)| \leq \frac{1}{2}$$

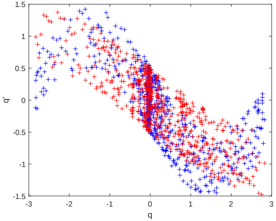
Generate ten trajectories with uniformly random bounded inputs for $\omega_1 = 1$ and $\omega_2 = 1/2$, plot the couples (q, q'),



(t, q) with $\omega_1 = 1$



(t, q) with $\omega_2 = 1/2$

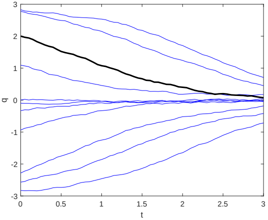


(q, q') phase plane

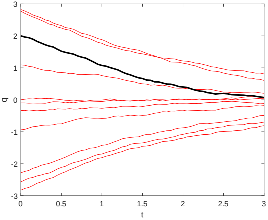
A toy example of nonlinear system with inputs

$$q'(t) = -sin(\omega_i \cdot q(t)) + u(t) \text{ with } |u(t)| \leq \frac{1}{2}$$

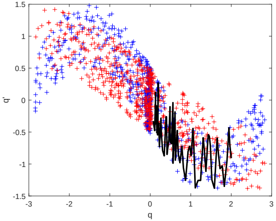
Generate ten trajectories with uniformly random bounded inputs for $\omega_1 = 1$ and $\omega_2 = 1/2$, plot the couples (q, q'), and consider a new trajectory of unknown type, guessing the type is **hard** in phase space



(t, q) with $\omega_1 = 1$



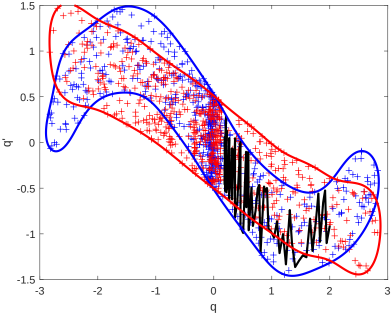
(t, q) with $\omega_2 = 1/2$



(q, q') phase plane

Delineating a graph to characterize a set-valued map

$$q'(t) \in F_i(q(t)) := \{-sin(\omega_i \cdot q(t)) + u(t) \mid u(\cdot) \in L_\infty, |u(t)| \leq 1/2\}$$
$$K_i := \{(q, q') \mid q' \in F_i(q)\} \subset \mathbb{R}^2$$



- Our goal is to approximately delineate the graphs of the sets K_i
- If a $t \mapsto (q(t), q'(t))$ trajectory crosses a boundary i, it cannot be of type i.

Our approach: set approximation then anomaly detection!

How to delineate a set in a plane?

A "Stack Overflow"-like question by J.J. Sylvester (1857)

"It is required to find the least circle which shall contain a given system of points in a plane."[a]

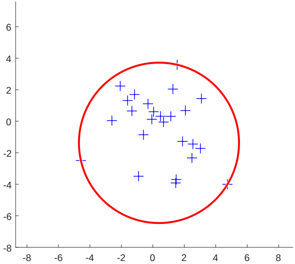
[a] *Quarterly journal of pure and applied mathematics, 1:79, 1857*

An answer ([1]) in 1972 related to quadratic programming

Minimal Enclosing Ball (MEB)

$$\min_{c \in \mathbb{R}^d, R} R^2$$

s.t. $\forall i \leq N, \; \|x_i - c\|_{\mathbb{R}^d} \leq R$



Support Vector Data Description (SVDD) is the (nonlinear) <u>kernelized</u> version of the Minimal Enclosing Ball problem.

Reproducing kernel Hilbert spaces (RKHS) in one slide

A RKHS $(\mathcal{F}_k(X), \langle\cdot,\cdot\rangle_{\mathcal{F}_k})$ is a Hilbert space of real-valued functions over a set X if one of the following is satisfied (Aronszajn, 1950 [2])

$\exists\, k : X \times X \to \mathbb{R}$ s.t. $k_x(\cdot) = k(x,\cdot) \in \mathcal{F}_k(X)$ and $f(x) = \langle f, k_x\rangle_{\mathcal{F}_k}$

k is s.t. $\exists\, \Phi_k : X \to \mathcal{F}_k(X)$ s.t. $k(x,y) \;=\; \langle \Phi_k(x), \Phi_k(y)\rangle_{\mathcal{F}_k}$

k is s.t. $\mathbf{G} = [k(x_i,x_j)]_{i,j=1}^n \succcurlyeq 0$ and $\mathcal{F}_k(X) := \overline{\mathrm{span}(\{k_x(\cdot)\}_{x\in X})}$, i.e. the completion for the pre-scalar product $\langle k_x, k_y\rangle_{k,0} = k(x,y)$

Classical kernels for $X = \mathbb{R}^d$ include the Gaussian and linear kernels

$$k_\sigma(x,y) = \exp\left(-\|x-y\|^2_{\mathbb{R}^d}/(2\sigma^2)\right) \qquad k_{\mathrm{lin}}(x,y) = \langle x, y\rangle_{\mathbb{R}^d}$$

There is a one-to-one correspondence between positive definite kernels k and RKHSs $\mathcal{F}_k(X)$.

Convex optimization perspective: MEB vs SVDD

MEB [1]

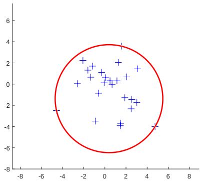
$$\min_{c \in \mathbb{R}^d, R} R^2$$

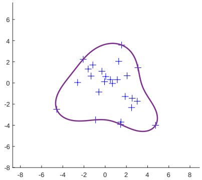
s.t. $\forall\, i \leq N,\ \|x_i - c\|_{\mathbb{R}^d} \leq R$

SVDD [3]

$$\min_{f \in \mathcal{H}_k, R} R^2$$

s.t. $\forall\, i \leq N,\ \|k(x_i, \cdot) - f(\cdot)\|_{\mathcal{H}_k} \leq R$





Dual problem of SVDD and MEB

$$\min_{\{\alpha \in \mathbb{R}_+^N \mid \sum_{i=1}^N \alpha_i = 1\}} \alpha^T G \alpha - \alpha^T \mathrm{diag}(G) \quad \text{with } G := (k(x_i, x_j))_{i,j \leq N}$$

Gaussian SVDD is an orthogonal projection

$$k_\sigma(x,y) = \exp\left(-\|x-y\|^2_{\mathbb{R}^d}/(2\sigma^2)\right) \text{ and } X_N = \{x_i\}_{i\le N} \subset \mathbb{R}^d$$

Lemma (SVDD with Gaussian kernels k_σ)

The center f_σ of the minimal enclosing ball B_σ^{SVDD} in the RKHS $\mathcal{H}_\sigma(\mathbb{R}^d)$ of $\Phi_\sigma(X_N) := \{k_\sigma(x_i,\cdot)\}_{i\le N}$ is the orthogonal projection of 0 onto $\mathrm{co}(\Phi_\sigma(X_N))$ (convex hull). Its radius R_σ satisfies:

$$R_\sigma = \sqrt{1-\|f_\sigma\|^2_\sigma} \text{ where } f_\sigma(\cdot) = \sum_{i=1}^N \overline{\alpha}_i k_\sigma(x_i,\cdot) = \operatorname*{arg\,min}_{f\in\mathrm{co}(\Phi_\sigma(X_N))} \|f\|^2_\sigma$$

$x \in K_\sigma^{SVDD} := \Phi_\sigma^{-1}(B_\sigma^{SVDD})$ iff a simple testing criterion holds:

$$\sum_{i,j\le N} \overline{\alpha}_i\overline{\alpha}_j k_\sigma(x_i,x_j) = \|f_\sigma\|^2_\sigma \le f_\sigma(x) = \sum_{i\le N} \overline{\alpha}_i k_\sigma(x_i,x)$$

Geometrical perspective: Gaussian kernel embedding

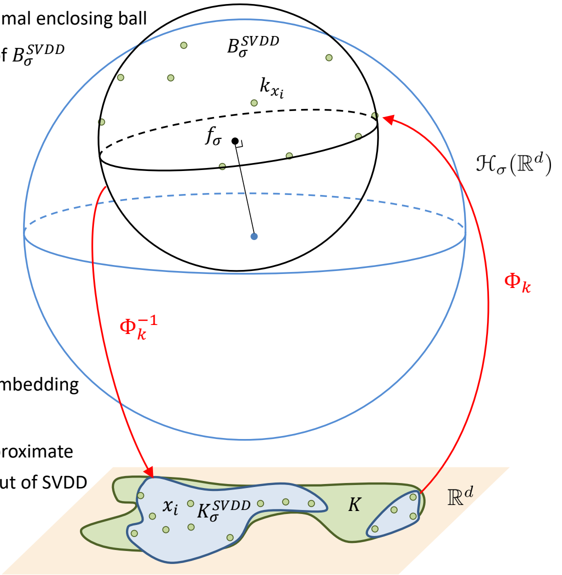
B_σ^{SVDD}: minimal enclosing ball

f_σ: center of B_σ^{SVDD}

Φ_k: kernel embedding

K: set to approximate

K_σ^{SVDD}: output of SVDD



Gaussian SVDD is set-consistent

The representation is usually sparse (only a few coefficients of f_σ are not 0), it is also consistent.

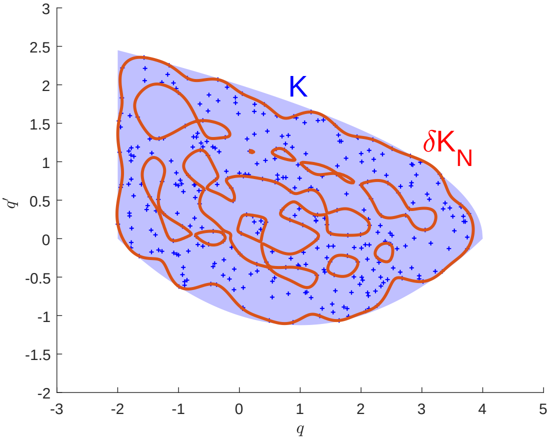
Proposition (Set-consistency of SVDD)

The estimate K_σ^{SVDD} of X_N by the SVDD algorithm for Gaussian kernels satisfies the following two properties

- $\exists\, M > 0,\ \forall\, \sigma > 0,\ K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0, M)$
- $\forall\, \epsilon > 0,\ \exists\, \sigma_0 > 0,\ \forall\, 0 < \sigma \leq \sigma_0,\ K_\sigma^{SVDD} \subset X_N + B_{\mathbb{R}^d}(0, \epsilon)$

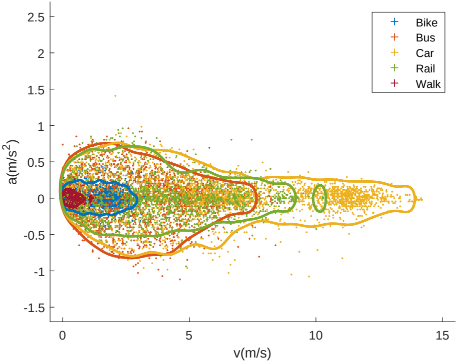
i.e. the sequence $(K_\sigma^{SVDD})_{\sigma > 0}$ is bounded and, for σ small enough, it lies in a neighborhood of X_N for the norm of $\mathbb{R}^d$.

In the limit case, when N tends to ∞, if X_∞ is dense in a given compact $K \subset \mathbb{R}^d$, then K_σ^{SVDD} is dense as well and lies in a neighborhood of K.



Motivation: Detection of transportation modes

We apply the SVDD algorithm to draw boundaries around the training sets in phase space (speed-acceleration).



- Five modes of transportation
- (Differentiated) GPS data ([4])
- Overlapping SVDD envelopes
- Various σ_i used

Numerical simulation on two vehicles - 1

Toy model of car and bike, forced to asymptotically track a reference velocity signal $v_{req}(\cdot)$ stemming from an urban-part of the NEDC cycle (New European Driving Cycle):

$$m\dot{v}(t) = -kv^2(t) + u(t)$$

$$\text{where } u(t) := \begin{cases} -F_{max} \text{ if } k_p(v_{req}(t) - v(t)) < -F_{max} \\ F_{max} \text{ if } k_p(v_{req}(t) - v(t)) > F_{max} \\ k_p(v_{req}(t) - v(t)) \text{ otherwise} \end{cases}$$

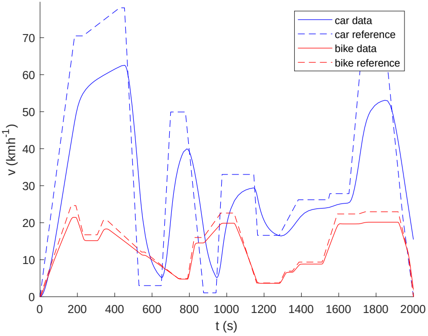
Table: List of parameters of the NEDC simulation

	m	k	k_p	F_{max}	$\max(v_{req})$
Car	1 T	0.27	20	2 kN	80 km/h
Bike	100 kg	0.5	20	30 N	30 km/h

Numerical simulation on two vehicles - 2

Example of tracking a reference velocity signal $v_{req}(\cdot)$

$$m\dot{v}(t) = -kv^2(t) + \text{sat}_{F_{max}}(k_p(v_{req}(t) - v(t)))$$



- Piecewise affine v_{req}
- Random breakpoints
- First order response v

Numerical simulation on two vehicles - 3

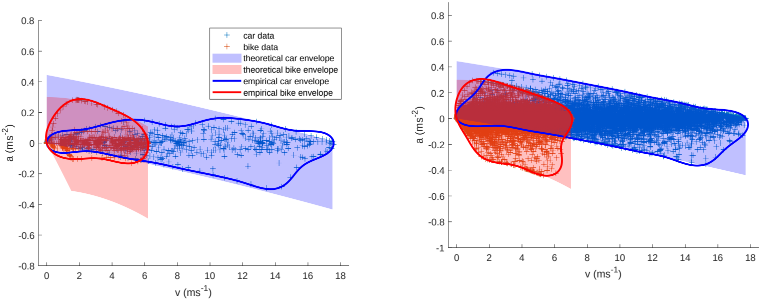


Figure: Estimate sets K_{σ}^{SVDD} of theoretical dynamical limits (filled areas) by SVDD on simulation data, when varying the number of training points

Numerical simulation on two vehicles - 3

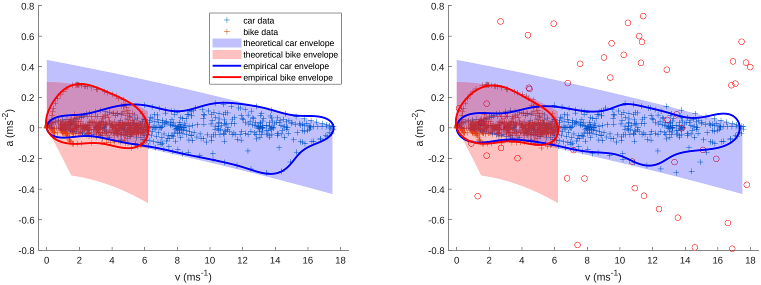


Figure: Estimate sets K_{σ}^{SVDD} of theoretical dynamical limits (filled areas) by SVDD on simulation data, when adding 10% uniform noise with a modification of SVDD to mitigate noise (see article)

Numerical simulation on two vehicles - 3

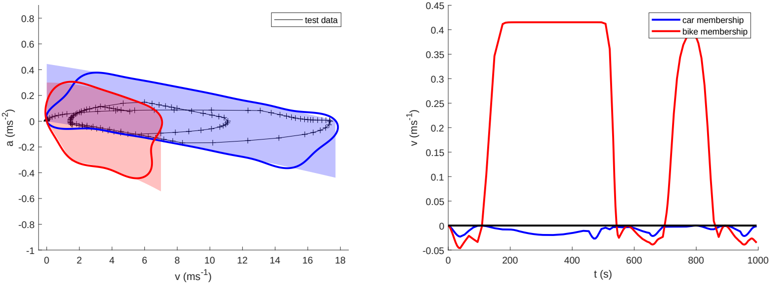


Figure: Estimate sets K_σ^{SVDD} of theoretical dynamical limits (filled areas) by SVDD on simulation data, when testing membership of a given trajectory to a class, by computing $\varphi_i(t) := f_{\sigma_i,i}(x(t)) - \|f_{\sigma_i,i}\|^2_{\sigma_i}$ (positivity means the trajectory crossed the boundary)

Conclusion

Through differential inclusions and kernel methods we

- reformulated an identification task as a problem of learning sets
- presented the SVDD algorithm [3], proving it was consistent w.r.t to the sampled set for Gaussian kernels and small σ
- applied it to both simulated and real data for detection of transportation modes

Thank you for your attention!

Not seen in the talk, to be found in the article:

- formulation of SVDD to mitigate noise in the samples
- stability of SVDD to variations of σ for the Gaussian kernel

References I

D. J. Elzinga and D. W. Hearn, "The minimum covering sphere problem," *Management Science*, vol. 19, pp. 96–104, sep 1972.

N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68, pp. 337–337, mar 1950.

D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, pp. 45–66, jan 2004.

B. Martin, V. Addona, J. Wolfson, G. Adomavicius, and Y. Fan, "Methods for real-time prediction of the mode of travel using smartphone-based GPS and accelerometer data," *Sensors*, vol. 17, p. 2058, sep 2017.