Proceedings of Machine Learning Research vol xxx:1–4, 2020

Learning Dynamical Systems with Side Information”

Amir Ali Ahmadi AAA@PRINCETON.EDU
Bachir El Khadir BKHADIR@PRINCETON.EDU

1. Proof of proposition 3

Fix a set $\Omega \in \mathbb{R}^n$ and two positive scalars T and L. We prove that there exists a constant $C_{\Omega,T}$ such that

$$\|f - g\|_{\Omega,T} \leq C_{\Omega,T}\|f - g\|_{\Omega}$$

for every two L-Lipchitz vector fields f and g that leave the set Ω invariant.

Lemma 1 (Gronwall) *Let $I = [a, b]$ denote a non-empty interval on the real line. Let α, β, and u be continuous, real-valued functions defined on I and satisfying*

$$u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)\, \mathrm{d}s \quad \forall t \in I.$$

If α is nondecreasing and β is nonnegative, then

$$u(t) \leq \alpha(t) \exp\left(\int_a^t \beta(s)\, \mathrm{d}s\right) \quad t \in I.$$

Proof Consider two trajectories x and y both starting from $x_0 \in \Omega$ and following f and g respectively. The proof is divided into two parts. In the first, we bound $\|y(t) - x(t)\|$, and in the second part, we bound $\|\frac{\partial}{\partial t}y(t) - \frac{\partial}{\partial t}x(t)\|$.

For the first part of the proof, we will bound $\|y(t) - x(t)\|$. By definition of y and x, for every $t \in [0, T]$ we have that

$$\begin{aligned} x(t) - y(t) &= \int_0^t f(x(s)) - g(y(s))\mathrm{d}s \\ &= \int_0^t f(x(s)) - g(x(s))\, \mathrm{d}s + \int_0^t g(x(s)) - g(y(s))\, \mathrm{d}s. \end{aligned}$$

Using the triangular inequality, we get

$$\|y(t) - x(t)\| \leq \int_0^t \|f(x(s)) - g(x(s))\|\, \mathrm{d}s + \int_0^t \|g(x(s)) - g(y(s))\|\, \mathrm{d}s. \quad (1)$$

Since f leaves Ω invariant, we know that for all $s \in [0, t]$, $x(s) \in \Omega$ and therefore $\|f(x(s)) - g(x(s))\| \leq \|f - g\|_\Omega$. Furthermore, because the function g is L-Lipchiz, $\|g(x(s)) - g(y(s))\| \leq L\|x(s) - y(s)\|$. From (1) we conclude therefore that

$$\|y(t) - x(t)\| \leq t\|f - g\|_\Omega + L\int_0^t \|x(s) - y(s)\|\, \mathrm{d}s,$$

and by Gronwall's lemma,

$$\|y(t) - x(t)\| \leq t \exp(Lt)\|f - g\|_\Omega \; \forall t \in [0, T].$$

For the second part of the proof, we will bound $\|\frac{\partial}{\partial t}y(t) - \frac{\partial}{\partial t}x(t)\| = \|f(x(s)) - g(y(s))\|$. note that since f and g leave Ω invariant, we know that for all $s \in [0, t]$, $x(s), y(s) \in \Omega$. By triangular inequality again, we know that $\|f(x(s))-g(y(s))\| \leq \|f(x(s))-g(x(s))\|+\|g(x(s))-g(y(s))\|$. Using the fact that g is L-Lipchiz on Ω, $\|g(x(s))-g(y(s))\| \leq L\|y(t)-x(t)\|$. Therefore $\|f(x(s)) - g(y(s))\| \leq \|f - g\|_\Omega + L\|y(t) - x(t)\|$.

Putting the first and second part of the proof together, we have

$$\|y(t) - x(t)\| \leq T \exp(LT)\|f - g\|_\Omega$$

and

$$\begin{aligned} \|\frac{\partial}{\partial t}y(t) - \frac{\partial}{\partial t}x(t)\| &\leq \|f - g\|_\Omega + L\|y(t) - x(t)\| \\ &\leq \|f - g\|_\Omega + Lt \exp(Lt)\|f - g\|_\Omega. \end{aligned}$$

Therefore

$$\|f - g\|_{\Omega,T} \leq C(\Omega, T, L)\|f - g\|_\Omega$$

with $C(\Omega, T, L) = \max\{T \exp(LT), 1 + LT \exp(LT)\}$. ■

2. Proof of Theorem 4

Fix a compact set $\Omega \subset R^n$, a time horizon $T > 0$, and a desired accuracy $\varepsilon > 0$. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous differentiable vector field that satisfies any one of the following constraints

(i) equilibria at a given finite set of points,

(ii) invariance of a full-dimensional, star-shaped basic semialgebraic set,

(iii) directional monotonicity,

(iv) nonnegativity,

(v) gradient or Hamiltonian structure,

(vi) symmetry.

In this section, we prove that there exists a polynomial vector field $p : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\|f - g\|_{\Omega,T} \leq \varepsilon,$$

and and p satisfies the same side information as f. Before we give separate proofs depending on which side information ((i), ...) f satisfies, we present the following universal-approximation result that will be invoked frequently.

Theorem 2 *(Weirstrass approximation theorem) If $f : \Omega \to \mathbb{R}^n$ is continuously-differentiable, then for any $\varepsilon > 0$, there exists a polynomial $p : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$\|f(x) - p(x)\| \leq \varepsilon \quad \forall x \in \Omega.$$

2.1. If f satisfies (i)

Suppose f satisfies

$$f(x_i) = \alpha_i \quad i = 1, \ldots, m. \tag{2}$$

Let p be a polynomial vector field uniformly approximating f as in Theorem 2. Let $v_i = (\alpha_i - (f(x_i) - p(x_i)))$, and note that $\|v\| \leq \varepsilon$.

The proposition below, shows that there exists a polynomial q such that

$$(p + q)(x_i) = \alpha_i \quad i = 1, \ldots, m,$$

and $\|q\|_\Omega \leq C\varepsilon$, where C is a constant depending only on the x_i. The last inequality implies in particular that

$$\|f - (p + q)\| \leq \varepsilon(1 + C).$$

Proposition 3 *For any positive integers d and n, for any d points $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^d$ in a compact set Ω, there exists a constant $C(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d)$ such that the following holds:*
For any vector $v \in \mathbb{R}^d$, there exists a polynomial $p \in \mathbb{R}_d[\boldsymbol{x}]$ such that

$$p(\boldsymbol{x}_i) = v_i, \quad i = 1, \ldots, d \text{ and } \max_{\boldsymbol{x} \in \Omega} |p(\boldsymbol{x})| \leq C(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d)\|v\|_2$$

Lemma 4 (Proposition 4.3 in Comon et al.) ***(Multivariate Polynomial Interpolation)*** *If $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^d$ are d different points of $\mathbb{R}^n$, then the vectors $m_d(\boldsymbol{x}^1), \ldots, m_d(\boldsymbol{x}^d)$ are linearly independent. (Here $m_d(\boldsymbol{x}^i)$ is the vector of monomials in $\boldsymbol{x}_i$ up to degree d.)*

Proof [Proof of Theorem 3] By Theorem 4, the system of linear equations

$$p(\mathbf{x}_i) = v_i \quad i = 1, \ldots, n \tag{3}$$

in the variable $p \in \mathbb{R}_d[\mathbf{x}]$ are independent. If we identify p by its coefficients in the monomial basis, then we can rewrite the system in Equation (3) as

$$Ap = v,$$

where $A = (m_d(\mathbf{x}^1), \ldots, m_d(\mathbf{x}^d))^T$ is a matrix depending only on $\mathbf{x}_1, \ldots, \mathbf{x}_d$ that has independent rows. The proof follows by taking $C(\mathbf{x}_1, \ldots, \mathbf{x}_d)$ to be the operator norm of the matrix $M := (A^T A)^{-1} A^T$ and $p = Mv$. ■

2.2. If f satisfies (ii)

Let B be the set defined by the inequalities $h_1(x) \geq 0, \ldots, h_m(x) \geq 0$, where each of the h_i is concave, and suppose there exists a point $x_0 \in B$ such that $h_i(x) > 0$ for all i. Suppose f leaves B invariant, i.e.,

$$h_j(x) \geq 0 \ \forall j \text{ and } h_i(x) = 0 \implies \langle f(x), \nabla h_i(x) \rangle \geq 0 \quad \forall x \in \mathbb{R}^n, i = 1, \ldots, m. \tag{4}$$

Let $f_\varepsilon(x) = f(x) - \varepsilon(x - x_0)$, and notice that $\|f_\varepsilon - f\| \leq |\Omega|\varepsilon$. (Here $|\Omega| := \sup_{x,y \in \Omega} \|x - y\|$). Moreover, if $h_j(x) \geq 0 \ \forall j$ and $h_i(x) = 0$, then

$$\langle f_\varepsilon(x), \nabla h_i(x) \rangle = \langle f(x), \nabla h_i(x) \rangle + \varepsilon \langle x - x_0, \nabla h_i(x) \rangle \geq \varepsilon h_i(x_0).$$

Let p be a polynomial vector field such that $\|p - f_\varepsilon\|_\Omega \leq \varepsilon$, then by continuity, p satisfies Equation (4). Moreover,

$$\|f - p\| \leq \|f - f_\varepsilon\| + \|f_\varepsilon - p\| \leq \varepsilon(|\Omega| + 1).$$

2.3. If f satisfies (iii)

Suppose f satisfies $\frac{\partial f_i}{x_j}(x) \trianglerighteq_{i,j} 0 \; \forall x \in B_{i,j}$ for $i, j \in \{1, \ldots, n\}$. Let $f_\varepsilon(x)$ the vector field defined component wise by

$$f_{\varepsilon,j}(x) = f_j(x) + \varepsilon \alpha_{ij} x_i,$$

where $\alpha_{ij} = 1$ if $\trianglerighteq_{i,j} = \geq$, and $\alpha_{ij} = -1$ otherwise.

2.4. If f satisfies (iv)

Suppose f satisfies $f_i(x) \trianglerighteq_i 0 \; \forall x \in B_i$ for $i, j \in \{1, \ldots, n\}$. Let $f_\varepsilon(x)$ the vector field defined component wise by

$$f_{\varepsilon,j}(x) = f_j(x) + \varepsilon \alpha_i,$$

where $\alpha_i = 1$ if $\trianglerighteq_i = \geq$, and $\alpha_i = -1$ otherwise.

2.5. If f satisfies (v)

Suppose f is a gradient, i.e., $f(x) = -\nabla V(x)$. Let W be polynomial approximation of V, i.e. $\|V - W\| \leq \varepsilon$ and $\|\nabla V - \nabla W\| \leq \varepsilon$. Then $p = -\nabla W$ works.

2.6. If f satisfies (vi)

Suppose f satisfies a symmetry

$$f(Ax) = Bf(x). \tag{5}$$

For a vector field h, let ψh denote the vector field $\frac{h(Ax)+Bf(x)}{2}$, and note that ψ is linear and that the equality in Equation (5) translates to $\psi f = f$.

Let p be a polynomial vector field uniformly approximating f as in Theorem 2, i.e., $\|p - f\|_\Omega \leq \varepsilon$, and let $q = \lim \psi^n p$. Note that q is a polynomial, and that $\psi q = q$.

References

Pierre Comon, Gene Golub, Lek-Heng Lim, and Bernard Mourrain. Symmetric tensors and symmetric tensor rank. URL http://arxiv.org/abs/0802.1681.