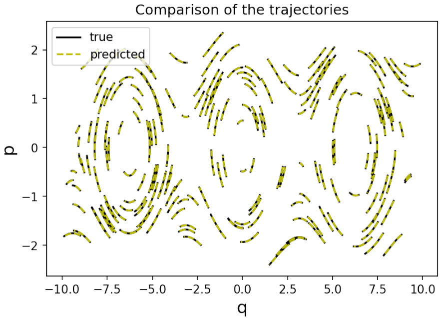
Learning the Hamiltonian of some constrained mechanical systems

Davide Murari
davide.murari@ntnu.no
Norwegian University of Science and Technology, Trondheim

Computational Mathematics and Machine Learning
Lorentz Center, 04-10-2021

Part of an ongoing project with Elena Celledoni, Ergys Çokaj, Andrea Leone, Brynjulf Owren.

Definition of the problem



Problem:

- How can we approximate the Hamiltonian H of a vector field $X_H \in \mathfrak{X}(M)$, where M is a symplectic manifold, starting from a set of given trajectories?
- And how can we approximate the solutions of this Hamiltonian system with a neural network?

A Neural Network as the Hamiltonian

- **Assumption on the Hamiltonian to learn:**

$$H(q,p) = \frac{1}{2} p^T M^{-1} p + U(q), \ \ (q,p) \in M \subset \mathbb{R}^{2n}$$

A Neural Network as the Hamiltonian

- **Assumption on the Hamiltonian to learn:**

$$H(q,p) = \frac{1}{2} p^T M^{-1} p + U(q), \;\; (q,p) \in M \subset \mathbb{R}^{2n}$$

- **Approximation of the kinetic energy**

$$K(p) \approx K_A(p) = \frac{1}{2} \|Ap\|^2 = \frac{1}{2} p^T (A^T A) p$$

A Neural Network as the Hamiltonian

- **Assumption on the Hamiltonian to learn:**

$$H(q,p) = \frac{1}{2}p^T M^{-1} p + U(q), \;\; (q,p) \in M \subset \mathbb{R}^{2n}$$

- **Approximation of the kinetic energy**

$$K(p) \approx K_A(p) = \frac{1}{2}\|Ap\|^2 = \frac{1}{2}p^T(A^T A)p$$

- **Approximation of the potential energy**

$$U(q) \approx U_\theta(q) = f_{\theta_m} \circ \ldots \circ f_{\theta_1}(q),$$

$$\theta_i = (W_i, b_i) \in \mathbb{R}^{n_i \times n_{i-1}} \times \mathbb{R}^{n_i}, \; \theta := [\theta_1, \ldots, \theta_m]$$

$$f_{\theta_i}(q) := \Sigma(qW_i^T + b_i), \; \mathbb{R}^n \ni z \mapsto \Sigma(z) = [\sigma(z_1), \ldots, \sigma(z_n)] \in \mathbb{R}^n,$$

and for example $\sigma(x) = \tanh(x)$.

Mechanical systems with holonomic constraints

We focus on the case $M = T^*Q \subset \mathbb{R}^{2n}$, where

$$Q = \{q \in \mathbb{R}^n : \;\; g(q) = 0 \in \mathbb{R}^m\} \subset \mathbb{R}^n.$$

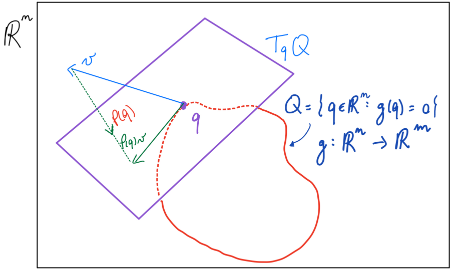
Mechanical systems with holonomic constraints

We focus on the case $M = T^*Q \subset \mathbb{R}^{2n}$, where

$$Q = \{q \in \mathbb{R}^n : \; g(q) = 0 \in \mathbb{R}^m\} \subset \mathbb{R}^n.$$

Let $H : M \to \mathbb{R}$ be the Hamiltonian. Then the vector field $X_H \in \mathfrak{X}(M)$ can be written in the form

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} 0_n & P(q) \\ -P(q)^T & M(q,p) \end{bmatrix} \nabla H(q,p) = X_H(q,p)$$



Case M is homogeneous

Homogeneous manifold

A smooth manifold M is homogeneous if there is a transitive Lie group action $\psi : G \times M \to M$.

Case M is homogeneous

Homogeneous manifold

A smooth manifold M is homogeneous if there is a transitive Lie group action $\psi : G \times M \to M$.

- We restrict to the case $T^*Q = M$ is homogeneous.
- Here the Δt flow of $X \in \mathfrak{X}(M)$, reads

$$\Phi^{\Delta t}(x) = \psi(exp(\sigma_x(\Delta t)), x)$$

$$\dot{\sigma}_x = dexp_{\sigma_x}^{-1} \circ f \circ \psi(exp(\sigma_x), x),\ \sigma_x(0) = 0 \in \mathfrak{g},$$

for some $f : M \to \mathfrak{g} = T_e G$.

Case M is homogeneous

Homogeneous manifold

A smooth manifold M is homogeneous if there is a transitive Lie group action $\psi : G \times M \to M$.

- We restrict to the case $T^*Q = M$ is homogeneous.
- Here the Δt flow of $X \in \mathfrak{X}(M)$, reads

$$\Phi^{\Delta t}(x) = \psi(exp(\sigma_x(\Delta t)), x)$$

$$\dot{\sigma}_x = dexp_{\sigma_x}^{-1} \circ f \circ \psi(exp(\sigma_x), x), \ \sigma_x(0) = 0 \in \mathfrak{g},$$

  for some $f : M \to \mathfrak{g} = T_e G$.
- **Examples:** systems on S^2, $SO(3)$, $SE(3)$, their (co)tangent bundles and even their cartesian products.

Case M is homogeneous

Homogeneous manifold

A smooth manifold M is homogeneous if there is a transitive Lie group action $\psi : G \times M \to M$.

- We restrict to the case $T^*Q = M$ is homogeneous.
- Here the Δt flow of $X \in \mathfrak{X}(M)$, reads

$$\Phi^{\Delta t}(x) = \psi(exp(\sigma_x(\Delta t)), x)$$

$$\dot{\sigma}_x = dexp_{\sigma_x}^{-1} \circ f \circ \psi(exp(\sigma_x), x), \; \sigma_x(0) = 0 \in \mathfrak{g},$$

  for some $f : M \to \mathfrak{g} = T_e G$.
- **Examples:** systems on S^2, $SO(3)$, $SE(3)$, their (co)tangent bundles and even their cartesian products.
- If $X = X_H$ is Hamiltonian, then f just depends on H and M, i.e. $f = F[H]$ for some F depending on M.

Learning a better Hamiltonian

- **Approximation of the Hamiltonian:** $Net(q,p) = K_A(p) + U_\theta(q)$,

Learning a better Hamiltonian

- **Approximation of the Hamiltonian:** $Net(q,p) = K_A(p) + U_\theta(q)$,
- **Training trajectories:**

$$\{(x_i, y_i^1, ..., y_i^M)\}_{i=1,...,N},\ y_i^j = \Phi_{X_H}^{j\Delta t}(x_i)$$

Learning a better Hamiltonian

- **Approximation of the Hamiltonian:** $Net(q, p) = K_A(p) + U_\theta(q)$,
- **Training trajectories:**

$$\{(x_i, y_i^1, ..., y_i^M)\}_{i=1,\dots,N},\ y_i^j = \Phi_{X_H}^{j\Delta t}(x_i)$$

- **Approximation of the Δt flow with some numerical method** $\hat{y}_i^{j+1} = \psi(exp(\bar{\sigma}), \hat{y}_i^j)$ where $\bar{\sigma}$ is the Δt solution of

$$\begin{cases} \dot{\sigma} = dexp_\sigma^{-1} \circ F[Net] \circ \psi(exp(\sigma), \hat{y}_i^j), \\ \sigma(0) = 0 \in \mathfrak{g} \end{cases}$$

approximated with some Runge-Kutta method.

Learning a better Hamiltonian

- **Approximation of the Hamiltonian:** $Net(q, p) = K_A(p) + U_\theta(q)$,
- **Training trajectories:**

$$\{(x_i, y_i^1, ..., y_i^M)\}_{i=1,...,N},\ y_i^j = \Phi_{X_H}^{j\Delta t}(x_i)$$

- **Approximation of the Δt flow with some numerical method** $\hat{y}_i^{j+1} = \psi(exp(\bar{\sigma}), \hat{y}_i^j)$ where $\bar{\sigma}$ is the Δt solution of

$$\begin{cases} \dot{\sigma} = dexp_\sigma^{-1} \circ F[Net] \circ \psi(exp(\sigma), \hat{y}_i^j), \\ \sigma(0) = 0 \in \mathfrak{g} \end{cases}$$

approximated with some Runge-Kutta method.

Optimize the weights: $\mathcal{L}(A, \theta) := \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M d(\hat{y}_i^j, y_i^j)^2$

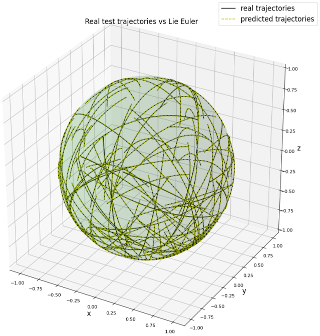
Case $M = T^*S^2$

- A transitive action on M is

$$\psi : SE(3) \times M \to M, \quad \psi((R, r), (q, p)) = (Rq, Rp + r \times Rq),$$

- The Hamiltonian is a function $H : M \to \mathbb{R}$, and

$$f = F[H] = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \nabla H.$$



Some questions for future work

- How can we preserve not just the geometry of the phase space, but even symplecticity of the flow map and the Hamiltonian energy?
- How does this global approach compare to one based on intrinsic formulation of the dynamics typical of geometric mechanics?
- How can we extend this approach to Hamiltonian systems defined on generic symplectic manifolds?

Thanks for the attention