Extending Neural ODEs with Discontinuities

Solutions of ODEs are smooth trajectories. ODEs lack a mechanism for modeling instantaneous interventions.

Example (simulation of a bouncing ball):

Velocity is discontinuous. Position has discontinuous derivative. A Neural ODE fails to model this.

We implement differentiable event handling to build models that learn when and how to apply instantaneous interventions.

Differentiable Event Handling

Define event function \( g(z, t) \). Event defined as when trajectories cross the root, i.e. \( t^* \) s.t. \( g(z(t^*), t^*) = 0 \).

\[
    t^*, z(t^*) = \text{ODESolveEvent}(z_0, f, g, t_0, \theta)
\]

Solves \( \frac{dz}{dt} = f(z, t, \theta) \) with initial state \((z_0, t_0)\) until event \( t^* \).

Gradient of ODESolveEvent can be derived by combining implicit function theorem and adjoint method (see paper for details).

Efficient O(1)-memory gradient implemented in github.com/rtqichen/torchdiffeq.

Neural Event ODE

Repeat: (i) solve until event, (ii) update state instantaneously.

\[
    \text{while } t_i < T \text{ do}
    \begin{align*}
    & (i) \quad t_{i+1}, z_{i+1} = \text{ODESolveEvent}(z_i, f, g, t_i) \\
    & (ii) \quad z_{i+1} = h(t_{i+1}, z'_{i+1})
    \end{align*}
    \text{end while}
\]

Capable of modeling variable number of discontinuities.

Threshold-based Event Functions

Threshold-based event occurs when an integral over a positive function reaches a predetermined threshold.

\[
    t^* \text{ such that } s = \int_{t_0}^{t^*} \lambda(t) \, dt
\]

where \( \lambda(t) > 0 \). Implemented by tracking \( \Lambda(t) = \int_{t_0}^{t} \lambda(s) \, ds \) as part of the ODE state and using \( g(t, z(t)) = s - \Lambda(t) \).

Allows exact gradients for integrate-and-fire spiking neural nets, inverse sampling, temporal point processes (TPP), etc.

Differentiable Sampling for TPPs

Sampling from a temporal point process (repeat):

(i) sample \( s_i \sim \text{Exp}(1) \)
(ii) solve for \( t_i \) such that \( s_i = \int_{t_{i-1}}^{t_i} \lambda(t) \, dt \)

Differentiable event handling through step (ii) provides the reparameterization gradient for temporal point processes.

Modeling Physics with Collision

Moreover, Neural ODE baseline approximates collisions with stiff trajectories, requiring 10× more compute to solve.