

Notes on Hamiltonian Monte Carlo

Junsong Tang

July 8, 2025

1 Notations

- $\mathbf{q} = (q_1, \dots, q_n)$: generalized coordinates
- $\dot{\mathbf{q}} = (\dot{q}_1, \dots, \dot{q}_n)$: generalized velocity
- $\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) = T - V = \frac{1}{2}M \|\dot{\mathbf{q}}\|^2 - V(\mathbf{q})$: Lagrangian, where M is the mass matrix, T is kinetic energy and V is potential energy
- \mathbf{p} : generalized momentum, i.e., $\mathbf{p} = M \cdot \dot{\mathbf{q}}$

2 Euler–Lagrange Equation

Generalized version: given a function $\mathbf{y}(x)$ and its derivative: $\mathbf{y}'(x)$ and $f(x, \mathbf{y}, \mathbf{y}')$. Define the functional:

$$F[\mathbf{y}] = \int_a^b f(x, \mathbf{y}, \mathbf{y}') dx \quad (1)$$

As the topic in calculus of variation, we want the functional $F[\mathbf{y}]$ to obtain its local min for some function \mathbf{y} . Perturbate $\mathbf{y}(x)$ with $\mathbf{y}(x) + \varepsilon \mathbf{u}(x)$ for any function \mathbf{u} and $\varepsilon \in \mathbb{R}$ with small enough $|\varepsilon|$. and put

$$g(\varepsilon) = \int_a^b f(x, \mathbf{y} + \varepsilon \mathbf{u}, (\mathbf{y} + \varepsilon \mathbf{u})') dx$$

To make $F[\mathbf{y}]$ obtaining its local min for some \mathbf{y} satisfying the initial conditions: $\mathbf{y}(a)$ and $\mathbf{y}(b)$ being fixed, i.e., $\mathbf{u}(a) = \mathbf{u}(b) = 0$, it is equivalent to require $g'(0) = 0$, hence:

$$\begin{aligned} g'(0) &= \frac{d}{d\varepsilon} \left(\int_a^b f(x, \mathbf{y} + \varepsilon \mathbf{u}, (\mathbf{y} + \varepsilon \mathbf{u})') dx \right) \\ &= \int_a^b \left(\frac{\partial f}{\partial \mathbf{y}} \mathbf{u} + \frac{\partial f}{\partial \mathbf{y}'} \mathbf{u}' \right) dx \quad \text{By total derivative} \\ &= \int_a^b \mathbf{u} \frac{\partial f}{\partial \mathbf{y}} dx + \frac{\partial f}{\partial \mathbf{y}'} \mathbf{u} \Big|_a^b - \int_a^b \mathbf{u} \frac{d}{dx} \left(\frac{\partial f}{\partial \mathbf{y}'} \right) dx \quad \text{integration by parts} \\ &= \int_a^b \mathbf{u} \left(\frac{\partial f}{\partial \mathbf{y}} - \frac{d}{dx} \frac{\partial f}{\partial \mathbf{y}'} \right) dx = 0 \quad \text{By initial condition: } \mathbf{u}(a) = \mathbf{u}(b) = 0 \end{aligned}$$

By the fundamental lemma of calculus of variation, since $\mathbf{u}(x)$ is arbitrary, hence we must have:

$$\frac{\partial f}{\partial \mathbf{y}} - \frac{d}{dx} \frac{\partial f}{\partial \mathbf{y}'} = 0 \quad (2)$$

Equation (2) is called the **Euler–Lagrangian Equation** in calculus of variation, and for \mathbf{y} satisfying Equation (2) is a sufficient condition for $F[\mathbf{y}]$ to have its local min.

In a dynamical system, with Lagrangian $\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}})$, we replace f with \mathcal{L} , \mathbf{y} with generalized coordinates \mathbf{q} , and x with time t , we get the **Lagrangian Equation** for dynamical system:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = 0 \quad (3)$$

3 Legendre Transform

4 Hamiltonian Equation

Hamiltonian is the Legendre transform of Lagrangian, hence we have:

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) \quad (4)$$

Since $\mathbf{p} \cdot \dot{\mathbf{q}} = 2T$, so $\mathcal{H}(\mathbf{q}, \mathbf{p}) = 2T - (T - V) = T + V$ and hence Hamiltonian can be interpreted as the total energy of a dynamical system.

We can derive Hamiltonian equations from Lagrange equation. Note that:

$$\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = M\dot{\mathbf{q}} = \dot{\mathbf{p}} \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} = \dot{\mathbf{p}}$$

so with (4), we must have:

$$\frac{\partial \mathcal{H}}{\partial \mathbf{p}} = \dot{\mathbf{q}} \quad (5)$$

$$\frac{\partial \mathcal{H}}{\partial \mathbf{q}} = -\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\dot{\mathbf{p}} \quad (6)$$

(5) and (6) are called **Hamiltonian equations**

Or equivalently, we can take the total differentiation on both sides of (4):

$$\begin{aligned} d\mathcal{H} &= d(\mathbf{p} \cdot \dot{\mathbf{q}}) - d\mathcal{L}(t, \mathbf{q}, \dot{\mathbf{q}}) \\ &= \dot{\mathbf{q}} \cdot d\mathbf{p} + \mathbf{p} \cdot d\dot{\mathbf{q}} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} d\mathbf{q} - \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}} d\dot{\mathbf{q}} - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \dot{\mathbf{q}} \cdot d\mathbf{p} - \frac{\partial \mathcal{L}}{\partial \mathbf{q}} d\mathbf{q} - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \frac{\partial \mathcal{H}}{\partial \mathbf{p}} d\mathbf{p} + \frac{\partial \mathcal{H}}{\partial \mathbf{q}} d\mathbf{q} + \frac{\partial \mathcal{H}}{\partial t} dt \end{aligned}$$

Hence we can correspond the coefficients to get (5) and (6).

5 Hamiltonian Monte Carlo