Functional Derivatives

Newtonian, Hamiltonian, Lagrangian mechanics

Principle of least action

Legendre Transform

Noethers Theorem

Lecture 6
2/17

How should we learn physics?

Newtonian Mechanics

$$F = m\ddot{x}$$

- Good for force balance
- Hard to associate w/ invariants
- Coordinate dependent

Hamiltonian Mechanics

$$\dot{x} = J \nabla H$$

- Energy centric
- Good for designing schemes

Legendre Transform

Lagrangian Mechanics

$$x = \underset{\hat{x}}{\operatorname{argmin}}\ S(\hat{x})$$

- optimization centric
- easy constraints, treatment of symmetries

Lagrangian mechanics

Lagrangian

$$L = \underset{\text{kinetic}}{T} - \underset{\text{potential}}{V}$$

$$L = \int_{\Omega} \underbrace{\mathcal{L}[u]}_{\text{Lagrangian density}} \, dx$$

Action

$$S = \int_0^T L \, dt$$

Principle of least action

$$u = \underset{\hat{u}}{\operatorname{argmin}} \int_0^T \int_{\Omega} \mathcal{L}[u] \, dx \, dt$$

- Put informally, the path that a system will evolve through ~~for~~ a fixed initial and final state will be the $u(t)$ which minimizes the action.
- To make sense of this, we need to minimize expression like S. Like standard calculus: take a derivative of a function $f(x)$, set to zero to find extremal pts. Here though, x is a coordinate, f a function. u is a function, S is a functional

$$S: V \to \mathbb{R}, \quad V \text{ a space of functions}$$

i.e. the input space is now <u>infinite dimensional</u>

A few definitions/options for how to do that. We'll show both so you understand what people mean by them, but then get less abstract.

Frechet derivative

- Let V be func. space, $F: V \to \mathbb{R}$. The differential $\delta F[u]$ is defined as, for $\delta u \in V$

$$F[u+\delta u] - F[u] = \delta F[u] + \varepsilon \|\delta u\|$$

such that $\lim_{\varepsilon \downarrow 0} \varepsilon \|\delta u\| = 0$

- This is a challenging definition to work with

Gateaux derivative

<u>Idea</u> Introduce a scalar parameter and use the vector calc def of a directional derivative

$$\delta F[u, \delta u] = \lim_{\varepsilon \downarrow 0} \frac{F[u+\varepsilon \delta u] - F[u]}{\varepsilon}$$

$$= \frac{d}{d\varepsilon} F[u+\varepsilon \delta u] \Big|_{\varepsilon=0}$$

To compute

Given $F[u] = \int_\Omega \mathcal{L}(x, u, D^\alpha u)\, dx$

δF is defined via, $\forall\, \delta u$ vanishing a bndry,

$$(\delta F, \delta u) = \lim_{\varepsilon \downarrow 0} \frac{F[u+\varepsilon \delta u] - F[u]}{\varepsilon}$$

Example Kinetic energy

$$K = \frac{1}{2}\int \rho u^2 \, dx \quad , \quad \text{take var. w.r.t. } u$$

$$(\delta_u K, \delta u) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_\Omega \frac{\rho (u + \varepsilon \delta u)^2}{2} - \frac{\rho u^2}{2} \, dx$$

$$= \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int \rho \left[\frac{u^2}{2} + \varepsilon u \delta u + \frac{\varepsilon^2}{2} \delta u^2 - \frac{\rho u^2}{2} \right] dx$$

$$= \lim_{\varepsilon \downarrow 0} \int \rho u \delta u + \frac{\varepsilon}{2} \rho \delta u^2$$

$$= \int \rho u \, du$$

$$= (\rho u, \delta u)$$

And so we identify

$$\delta_u K = \rho u \quad \text{(the momentum)}$$

Be careful Physicists will derive by setting δu to dirac delta $\rightarrow$ only ok for cont. functionals

Further properties

Denote $\frac{\delta}{\delta u} F[u] = \delta_u F[u]$

- Linearity $\frac{\delta}{\delta u}\left(\lambda F[u] + \mu G[u]\right) = \lambda \delta_u F[u] + \mu \delta_u G[u]$
- Productrule $\frac{\delta}{\delta u}\left(F[u] G[u]\right) = \delta_u F \, G + F \, \delta_u G$
- Chain rule $\frac{\delta}{\delta u}\left(F[g(u)]\right) = \frac{\delta F[g(u)]}{\delta g(u)} \frac{d g(u)}{du}$

The Euler-Lagrange equations

Assume a generic functional density

$$F[u] = \int_\Omega f(x, u, \nabla u)\, dx$$

$$(\delta_u F, \delta u) = \left[\frac{d}{d\varepsilon} \int f(x, u + \varepsilon \delta u, \nabla u + \varepsilon \nabla \delta u) dx\right]_{\varepsilon = 0}$$

Chain rule

$$= \int_\Omega \frac{\partial f}{\partial u} \delta u + \frac{\partial f}{\partial \nabla u} \cdot \nabla \delta u \quad dx$$

$$= \int_\Omega \left(\frac{\partial f}{\partial u} - \nabla \cdot \frac{\partial f}{\partial \nabla u}\right) \cdot \delta u + \cancel{\int_{\partial\Omega} \frac{\partial f}{\partial \nabla u} \delta u \cdot dA}$$

Assume vanishing

So $$\delta_u F = \cancel{\frac{\partial f}{\partial u} - \sum_i \frac{\partial_{x_i}}{\partial x_i} \frac{\partial f}{\partial x_i u}}$$

$$= \frac{\partial f}{\partial u} - \sum_i \frac{\partial}{\partial x_i} \frac{\partial f}{\partial (\partial_{x_i} u)}$$

- We can just apply this as a formula and skip all the limits
- For different form f, get a different equ

ex

$$F[u] = \int_0^1 f(x, u, \partial_x u, \partial_{xx} u)\, dx \Rightarrow$$

$$\delta_u F = \frac{\partial f}{\partial u} - \frac{d}{dx} \frac{\partial f}{\partial (\partial_x u)} + \frac{d^2}{dx^2} \frac{\partial f}{\partial (\partial_{xx} u)}$$

Returning to the least action principle

$$S[u] = \int_a^b \int_\Omega \mathcal{L}(x, u, \dot{u})\, dx\, dt$$

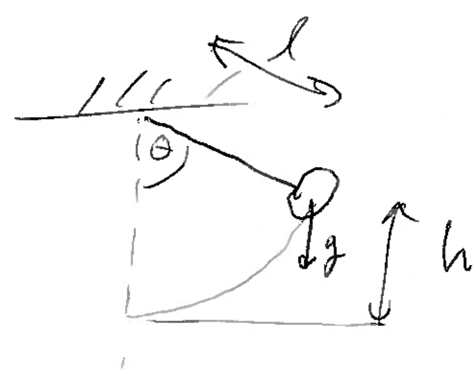
Finally we can state that the path which admits an extremal value of S satisfies

$$\delta_u S[u] = 0$$

Applying the E-L eqns

$$\boxed{\frac{\partial \mathcal{L}}{\partial u} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{u}} = 0}$$

Again, depending on $\mathcal{L}$ we can get different versions of this



<u>Ex</u> Linear pendulum

$$T = \frac{1}{2} m v^2 = \frac{1}{2} m l^2 \dot{\theta}^2$$

$$U = mgh$$

$$h = l(1 - \cos\theta) \approx \frac{1}{2} l \theta^2$$

$$L = \frac{1}{2} m l^2 \dot{\theta}^2 - \frac{1}{2} mgl\theta^2$$

$$\partial_\theta L = -mgl\theta$$

$$\partial_{\dot{\theta}} L = ml^2\dot{\theta}$$

<u>E-L eqn</u> $\partial_\theta L - \frac{d}{dt}\partial_{\dot{\theta}} L = 0 \Rightarrow -mgl\theta - \frac{d}{dt} ml^2\dot{\theta} = 0$

$$\ddot{\theta} = -\frac{g}{l}\theta$$

Legendre transform

A technique to switch between the Lagrangian and Hamiltonian description

Given a Lagrangian

$$L(q_1, \dots, q_N, \dot{q}_1, \dots \dot{q}_N)$$

The legendre transform induces the <u>conjugate momenta</u> $p_1, \dots p_N$

$$\begin{cases} H(q_1, \dots q_N, p_1, \dots p_N) = \sum_i p_i \dot{q}_i - L(q_1, \dots q_N, \dot{q}_1 \dots \dot{q}_N) \\ p_i = \partial_{\dot{q}_i} \mathcal{L} \end{cases}$$

Ex Getting Hamiltonian from Lagrangian for pendulum

$$q = \theta \quad , \quad \mathcal{L} = \frac{1}{2} m l^2 \dot{q}^2 - \frac{1}{2} m g l q^2$$

$$p_q = \partial_{\dot{q}} \mathcal{L} = m l^2 \dot{q} \Rightarrow \dot{q} = \frac{1}{m l^2} p$$

$$\mathcal{L} = \frac{1}{2} (m l^2)^{-1} p^2 - \frac{1}{2} m g l q^2$$

$$H = p \dot{q} - \mathcal{L}$$

$$= \frac{1}{m l^2} p^2 - \frac{1}{2} (m l^2)^{-1} p^2 - \frac{1}{2} m g l q^2$$

$$= \frac{p^2}{2 m l^2} - \frac{m g l q^2}{2}$$

Constrained Lagrangians

<u>def</u> If constraint is equality and a function of coordinate only (no $\dot{q}$'s) it is <u>holonomic</u>

$$f_i(q, t) = 0$$

We can augment the Lagrangian w/ a Lagrange multiplier

$$\hat{L}(q, \dot{q}, \lambda) = L(q, \dot{q}) + \sum_i \lambda_i f_i(q, t)$$

Where λ is a <u>field</u> (in contrast to KKT we considered previously)

Applying Euler-Lagrange we obtain equations of motion

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_j}\right) = \dfrac{\partial L}{\partial q_j} + \lambda(t)\dfrac{\partial f_i}{\partial q_j} \\[2ex] f_i(q, t) = 0 \end{cases}$$

Noether's theorem

Informally - symmetries which leave the Lagrangian unchanged each have a corresponding conserved quantity.

Consider the maps
$$t \to t' = t + \delta t$$
$$q \to q' = q + \delta q$$

Assume N of such maps
$$\delta t = \sum_r \varepsilon_r T_r$$
$$\delta q = \sum_r \varepsilon_r Q_r$$

Then
$$\left(\frac{\partial L}{\partial \dot{q}} \cdot \dot{q} - L \right) T_r - \frac{\partial L}{\partial \dot{q}} \cdot Q_r$$

Are conserved quantities.

Examples

(1) Time Invariance
$$t \to t + \delta t$$
$$N = 1, \quad T = 1, \quad Q = 0$$

$$\frac{\partial L}{\partial \dot{q}} \cdot \dot{q} - L$$
is conserved (this is actually the Legendre transform providing H!)

(2) Translation invariance
$$q \to q + \delta q$$
$$N = 1, \quad T = 0, \quad Q = 1$$

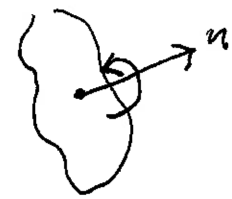
$$\frac{\partial L}{\partial \dot{q}} = \text{const}$$
(this is the definition of conjugate momentum!)

③ Rotational invariance

$L = r \times p$ is angular momentum

Consider $r \to r + \delta\theta \, n \times r$

Rotation of $\delta\theta$ about a given axis n.



$N=1, \; T=0 \qquad Q = n \times r$

$$\text{const} = \frac{\partial L}{\partial \dot{q}} \cdot Q = p \cdot (n \times r)$$
$$= n \cdot (r \times p)$$
$$= n \cdot L$$

When machine learning dynamics, we will design architectures which respect this set of symmetries