From Mechanics to Control

Ying Nian Wu, UCLA Statistics Department

1 Classical mechanics

1.1 Lagrangian

Suppose there are n particles whose positions and velocities are $x(t) = (x_i(t), i = 1, ..., n)$ and $v(t) = (v_i(t), i = 1, ..., n)$, with $v_i(t) = dx_i(t)/dt$. Sometimes, for notional simplicity, we may simply write $x(t)$ as x, and $v(t)$ as v.

The Lagrangian is

$$L(x, v, t) = \frac{1}{2} \sum_{i=1}^{n} m_i v_i^2 - V(x), \tag{1}$$

which is the difference between the kinetic energy and the potential energy. It can be more general.

1.2 Action

The action from 0 to T is

$$S = \int_0^T L(x, v) dt, \tag{2}$$

which is a function (or functional) of the path $x(t), t \in [0, T]$.

1.3 Least action

Nature controls the path $x(t)$ by minimizing the action, i.e., nature does barely enough for the kinetic energy to overcome the potential energy.

In general, we only require the path $x(t)$ to be a stationary point of the action S, i.e., a minimum, a maximum, or a saddle point. But usually it should be a minimum.

1.4 Discretization

A valuable method to understand the continuous process is to always discretize the time into intervals of Δt, very much like making a movie where the time between every two consecutive frames is Δt.

In discrete time, we write S as

$$S = \sum_0^T L(x(t), v(t)) \Delta t, \tag{3}$$

where

$$v(t) = \frac{x(t + \Delta t) - x(t)}{\Delta t}. \tag{4}$$

Then for a fixed t, the part of S that depends on $x(t)$ is

$$L(x(t), v(t))\Delta t + L(x(t - \Delta t), v(t - \Delta t))\Delta t, \tag{5}$$

where

$$v(t - \Delta t) = \frac{x(t) - x(t - \Delta t)}{\Delta t}. \tag{6}$$

1

1.5 Euler-Lagrange

Thus

$$\frac{\partial S}{\partial x(t)} = L^{(1)}(x(t), v(t))\Delta t - L^{(2)}(x(t), v(t)) + L^{(2)}(x(t - \Delta t), v(t - \Delta t)) = 0, \tag{7}$$

which leads to the Euler-Lagrange equation

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\frac{\partial L}{\partial v}. \tag{8}$$

The above derivation of the Euler-Lagrange is unconventional, but is the most direct.

Note: for $f(x, y)$, we write $f^{(1)}(x, y) = \partial f(x, y)/\partial x$, and $f^{(2)}(x, y) = \partial f(x, y)/\partial y$.

1.6 Hamiltonian

Define the momentum

$$p = \frac{\partial L}{\partial v}. \tag{9}$$

Usually $p = mv$.

Define the Hamiltonian $H(x, p)$ to be

$$H = pv - L. \tag{10}$$

H is usually the total energy, i.e., the sum of the kinetic energy and the potential energy.

We will derive the Hamiltonian equation in the next subsection:

$$\frac{\partial p}{\partial t} = -\frac{\partial H}{\partial x}; \tag{11}$$

$$\frac{\partial x}{\partial t} = \frac{\partial H}{\partial p}. \tag{12}$$

1.7 Lagrange multiplier and Hamiltonian

In discretize time.

$$S = \sum_0^T L(x(t), v(t))\Delta t, \tag{13}$$

where the constraint is

$$x(t + \Delta t) = x(t) + v(t)\Delta t. \tag{14}$$

To minimize S under the constraint, we can use Lagrange multiplier, to get

$$\tilde{S} = \sum_0^T \left\{ L(x(t), v(t))\Delta t + \lambda(t)[x(t + \Delta t) - x(t) - v(t)\Delta t] \right\}. \tag{15}$$

At the minimum, we should have

$$\frac{\partial \tilde{S}}{\partial x(t)} = 0; \frac{\partial \tilde{S}}{\partial v(t)} = 0; \frac{\partial \tilde{S}}{\partial \lambda(t)} = 0. \tag{16}$$

2

First,

$$\frac{\partial \tilde{S}}{\partial v(t)} = L^{(2)}(x(t), v(t))\Delta t - \lambda(t)\Delta t = 0. \tag{17}$$

Thus

$$\lambda(t) = L^{(2)}(x(t), v(t)). \tag{18}$$

Compare to the above subsection, λ is the momentum p.

Define the Hamiltonian

$$H(x(t), \lambda(t)) = \lambda(t)v(t) - L(x(t), v(t)). \tag{19}$$

Then

$$\tilde{S} = \sum_{t_1}^{t_2} \lambda(t)(x(t + \Delta t) - x(t)) - H(x(t), \lambda(t))\Delta t. \tag{20}$$

Thus,

$$\frac{\partial \tilde{S}}{\partial \lambda(t)} = 0 \rightarrow \frac{\partial H}{\partial \lambda} = \frac{\partial x}{\partial t} \tag{21}$$

$$\frac{\partial \tilde{S}}{\partial x(t)} = 0 \rightarrow \frac{\partial H}{\partial x} = -\frac{\partial \lambda}{\partial t}. \tag{22}$$

Thus we get the Hamiltonian equation with $\lambda = p$.

2 Quantum mechanics

2.1 Hamiltonian and rotation of belief vector

Let $b(x)$ be the belief vector. In quantum mechanics, $b(x)$ is a complex number for each x. An observer makes a measurement, the probability density (or mass) function of observing x is $p(x) = |b(x)|^2$ for each x, i.e., the sum of squares of the real and imaginary parts of $b(x)$.

For simplicity, let us first assume x can only take a finite number of values (e.g., the spin of the electron). Then $b = (b(x), \forall x)$ forms a finite dimensional vector. b keeps rotating over time, i.e., $b_t(x)$ keeps changing over time, but the total squared length $|b_t|^2 = \sum_x |b_t(x)|^2$ remains the same. We can normalize $|b_t|^2$ to 1, so that $\sum_x p_t(x) = |b_t(x)|^2 = 1$. Thus $p_t(x)$ is always a legitimate probability distribution.

The change of b_t is driven by the Hamiltonian H, according to the Schrodinger equation:

$$ih\frac{\partial b_t}{\partial t} = Hb_t,$$

where h is the Planck constant. In other words,

$$b_{t+\Delta t} = (I - iH\Delta t/h)b_t \approx \exp(-iH\Delta t/h),$$

where we use $e^x = 1 + x + O(x^2)$ for small x, and for matrix M, $e^M = 1 + M + M^2/2 + M^3/3! + \dots$. Because of i, $e^{-ix} = \cos x - i\sin x$ is a rotation in the complex domain, and similarly $\exp(-iH\Delta t/h)$ is a rotation matrix, so that the squared length of b_t is preserved. In fact, $b_t = \exp(-iHt/h)b_0$.

3

Now let us assume that x is continuous, e.g., 1-dimensional, such as the position of the electron. Then $b_t = (b_t(x), \forall x)$ is an infinite dimensional vector, in fact, the dimension is not even countable since x is continuous, and H is an infinite dimensional matrix, or more precisely an operator. But we still have the Schrodinger equation

$$ih\frac{\partial b_t(x)}{\partial t} = Hb_t(x),$$

where $b_t(x)$ is now a wave function, and H is an operator on this function.

How do we get H? The recipe is as follows. Just take the classical Hamiltonian, $H(x, p)$, which is a function of position x and momentum p, and change p to the operator $\hat{p} = ih\partial/\partial x$ (so that p^2 is changed to $\hat{p}^2 = -h^2\partial^2/\partial x^2$). This is inspired by de Broglie's hypothesis that any matter is also a wave whose wavelength is inverse proportional to p.

2.2 Discreteness of Hamiltonian eigenvalues

Where does the name "quantum" come from? It turns out the operator (matrix) H has a discrete list of eigen values, i.e., $H = Q\Lambda Q^\top$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, ...)$ is a diagonal matrix of eigen values, $Q = (q_1, q_2, ...)$ collects the eigen vectors. λ_i are the discrete energy levels. Because of the discreteness of λ_i, we say that the energy is quantized. This assumption was originally made by Planck (as well as Einstein, Bohr) in an ad hoc manner to account for certain experimental results. But in the above formulation, it is a natural consequence.

Why do we take $\hat{p} = ih\partial/\partial x$? We can think of $\hat{p}$ as an infinite matrix (with uncountable entries). Its eigen vectors are $v_p = \exp(-ipx/h)$, so that $\hat{p}v_p = ih\partial \exp(-ipx/h)/\partial x = p\exp(-ipx/h) = pv_p$. Thus p is the eigen value of $\hat{p}$.

As to the position x, its operator version $\hat{x}$ is actually a diagonal matrix whose diagonal elements are different values of x, so that the eigen vectors of $\hat{x}$ are δ_x, i.e., the Dirac one-hot vector for x, and the eigen values are x.

Now we have three sets of basis, position basis $(\delta_x, \forall x)$, momentum basis $(v_p, \forall p)$, and energy basis $(q_i, \forall i)$. For a belief vector b_t, we can expand it in the three sets of basis,

$$b_t = \int_x b_t(x)\delta_x = \int_p c_t(p)v_p = \sum_i a_{ti}q_i.$$

Thus the probability (density) we find the electron at position x is $|b_t(x)|^2$, the probability (density) we find the electron with momentum p is $|c_t(p)|^2$, and the probability we find the electron at energy level i is $|a_{ti}|^2$.

For the uncertainty principle, it can be prove that for any b_t, the variance of the probability distribution $|b_t(x)|^2$ and the variance of the probability distribution $|c_t(p)|^2$ satisfy the constraint that their product must be bigger than a certain constant.

Now you can see that any quantity, be it position, momentum, or energy, is a matrix (or operator), whose eigen values are the possible values of that quantity, and whose eigen vectors form a basis. The belief vector can be expanded in this basis, and the squares of the coordinates become probabilities.

2.3 Lagrangian and Feynman path integral

Suppose at time $t = 0$, the system is in x_0. Then the probability amplitude that the system is in x at time T, i.e., $b_T(x)$, is

$$\int \exp\left(i\frac{S(x(t))}{h}\right) D(x(t)), \tag{23}$$

where $x(t)$ is the path over $[0,T]$ with $x(0) = x_0$ and $x(T) = x$. Intuitively, the above is

$$\sum_{\text{path}} \exp\left(i\frac{S(\text{path})}{h}\right),\tag{24}$$

which is the sum over all possible paths, i.e., sum over histories.

The least action principle can be derived from the above path integral. If $S'(\text{path}) = 0$ at a particular path, then S is locally constant around this path, then the sum of $e^{iS/h}$ around this path is the sum of a constant complex number. If $S'(\text{path}) \neq 0$ at a particular path, then S is locally linear around this path, then the sum of $e^{iS/h}$ around this path is the sum of complex numbers of very different phases if h is small, and they cancel each other. So in the end, only the stationary paths survive.

If we change it to t, then the above path integral becomes the normalizing constant or partition function Z for a statistical mechanical system. This is how MCMC is used for computing the path integral in QM.

Quantum mechanics is simpler than classical mechanics because it is linear. The mathematical language is simply vector matrix multiplication. For instance, in quantum field theory, the creation and annihilation operators act on vectors that correspond to different numbers of particles, so that the creation and annihilation of particles can be expressed by vector matrix multiplication.

3 Statistical mechanics

Consider a system $(x,v) = (x_i, v_i, i = 1,...,n)$ evolves over time. According to the classical mechanics, we get a deterministic video $(x(t), v(t), t \in [0,T])$. Suppose T is large, and we randomly sample a frame $t \sim \text{Unif}[0,T]$. Then $(x(t), v(t))$ follows a probability distribution $p(x,v)$. Marginally, $x(t) \sim p(x)$. This is the conceptual basis of statistical mechanics. For large system, the statistical properties of $x(t)$ tend to have little fluctuations, i.e., $\text{Var}[\phi(x(t))]$ tends to be small for a summary ϕ. Thus when we measure ϕ, it is close to $\mathbb{E}[\phi(x(t))]$. The closeness of $\phi(x(t))$ to its expectation is often characterized by concentration inequality or large deviation result, e.g., $\Pr(|\phi(x(t)) - \mathbb{E}[\phi(x(t))]| > \varepsilon) \leq c_1 \exp(-nc_2\varepsilon^2)$, where c_1 and c_2 are constants.

In the above, the video is deterministic, and the randomness comes from the random sampling of t. This is very different from quantum mechanics, where the randomness is intrinsic.

In Boltzmann's formation, $\mathbb{E}[\phi(x(t))]$ is considered time average, which is what we did. In Gibbs' thinking, $\mathbb{E}[\phi(x(t))]$ is considered ensemble average, i.e., one can imagine an ensemble (or population) of systems evolve in parallel, and we average the ensemble. Regardless, one needs concentration of the probability (i.e., fluctuations diminish) to explain the success of statistical mechanics.

For Hamiltonian Monte Carlo (or hybrid Monte Carlo), we assume $p(x,v) = p(x)q(v) \propto \exp(-H(x,v))$, where q is Gaussian. Given the current x, we randomly sample v from $q(v)$. Then we let the system evolve according to Hamiltonian equation based on $H(x,v)$. This is like randomly pick a short video segment.

4 Optimal control

The optimal control for deterministic system, in particular, the Pontryagin's principle, follows the treatment of classical mechanics.

Here we continue to use the notation in mechanics following the Russian school, so that state and action are denoted by (x,u). In reinforcement learning, we use (s,a).

4.1 Deterministic dynamics

The optimal control problem has a dynamics

$$\dot{x} = F(x, u), \tag{25}$$

where the Newtonian notation $\dot{x} = dx/dt$, $u(t)$ is the control variable. In discrete time, it means

$$x(t + \Delta t) - x(t) = F(x(t), u(t))\Delta t. \tag{26}$$

We want to find $u(t), t \in [0, T]$, to minimize the cost

$$S = \phi(x(T)) + \int_0^T l(x, u)dt. \tag{27}$$

where $\phi(x)$ is the cost of the final state and $l(x, u)$ is the loss rate. In discrete time,

$$S = \phi(x(T)) + \sum_0^T l(x(t), u(t))\Delta t. \tag{28}$$

We can plug in $x(t + \Delta t) = x(t) + F(x(t), u(t))\Delta t$, so that S is a function of $u(t), t \in [0, T]$. We can minimize S by gradient descent over $u(t), t \in [0, T]$, where the gradient can be computed by back-propagation through time.

4.2 Lagrange multiplier and Pontryagin

Instead of back-propagation through time after plugging in $x(t + \Delta t) = x(t) + F(x(t), u(t))\Delta t$, we can also treat it as a constraint, and use Lagrange multiplier.

$$\tilde{S} = \phi(x(T)) + \sum_0^T \left\{ l(x(t), u(t))\Delta t + \lambda(t)[x(t + \Delta t) - x(t) - F(x(t), u(t))\Delta t] \right\}. \tag{29}$$

Define the Hamiltonian

$$H(x(t), u(t), \lambda(t)) = \lambda(t)F(x(t), u(t)) - l(x(t), u(t)), \tag{30}$$

then

$$\tilde{S} = \phi(x(T)) + \sum_0^T \left\{ \lambda(t)[x(t + \Delta t) - x(t)] - H(x(t), u(t), \lambda(t))\Delta t \right\}. \tag{31}$$

Thus we get the Hamiltonian equation:

$$\frac{\partial \tilde{S}}{\partial \lambda(t)} = 0 \rightarrow \frac{\partial H}{\partial \lambda} = \frac{\partial x}{\partial t} \tag{32}$$

$$\frac{\partial \tilde{S}}{\partial x(t)} = 0 \rightarrow \frac{\partial H}{\partial x} = -\frac{\partial \lambda}{\partial t}. \tag{33}$$

The first equation is just the dynamics. We can run it forward to get $x(t)$. The second equation enables us to update

$$\lambda(t) = \lambda(t + \Delta t) - \frac{\partial H}{\partial x}\Delta t, \tag{34}$$

backward, starting from

$$\lambda(T) = \phi'(x(T)). \tag{35}$$

With $x(t)$ and $\lambda(t)$, we can then minimize $\tilde{S}$ with respect to $u(t)$. This is an unconstrained minimization, so that we only need to maximize each $H(x(t), u(t), \lambda(t))$ with respect to $u(t)$ separately, without the need for back-propagation through time.

The above is the Pontryagin maximum principle. One can see that the above optimal control method follows the classical mechanics closely. The classical mechanics can be considered as nature controls the trajectory, where the control variable is $v(t)$.

4.3 Dynamic programming and Bellman

The dynamic programming approach is based on the value function $v(x, t)$, with

$$v(x, T) = \phi(x), \tag{36}$$

and

$$
\begin{align}
v(x,t) &= \max_u [l(x,u)\Delta t + v(x(t+\Delta t), t+\Delta t) \tag{37}\\
&= \max_u [l(x,u)\Delta t + v(x+F(x,u)\Delta t, t+\Delta t). \tag{38}
\end{align}
$$

The above is the Bellman equation for value iteration. Directly applying the above equation to continuous x can be inconvenient because we need to memorize $v(x,t)$ for all x, and if x is high dimensional, there is what Bellman called "curse of dimensionality."

Bellman is more convenient for finite discrete state space. It underlies reinforcement learning and in particular Q-learning.

4.4 Hamilton-Jacobi-Bellman

Applying Taylor expansion to Bellman,

$$
\begin{align}
v(x,t) &= \max_u [-l(x,u)\Delta t + v(x+F(x,u)\Delta t, t+\Delta t) \tag{39}\\
&= \max_u [-l(x,u)\Delta t + v(x, t+\Delta t) + v_x(x, t+\Delta t)F(x,u)\Delta t]. \tag{40}
\end{align}
$$

Thus

$$-v_t(x,t) = \max_u [-l(x,u) + v_x(x,t)F(x,u)], \tag{41}$$

where v_x and v_t are derivatives of $v(x,t)$ with respect to x and t respectively.

Compared to Pontryagin, $\lambda(t) = v_x(x,t)$, and the right hand side of the above equation is the Hamiltonian $H(x, u, v_x)$, i.e.,

$$-v_t = H(x, v_x, t). \tag{42}$$

Again the above equation follows the Hamiton-Jacobi equation of classical mechanics. This equation was used by Schrodinger to justify his wave equation by making analogy to optics, where there are both wave and particle treatments.

The LQR (linear quadratic regulation) can be derived from HJB.

4.5 Stochastic dynamics

For stochastic dynamics, we need to use expectation with respect to the random transition. Then we get into the Markov decision process. In continuous time, the dynamics can also be modeled by adding Brownian motion to dynamics, so that

$$x(t + \Delta t) = x(t) + F(x(t), u(t))\Delta t + \sigma(x(t), u(t))\sqrt{\Delta t}\varepsilon_t, \tag{43}$$

where $\varepsilon_t \sim N(0, I)$ independently over t, and $\sigma(x, u)$ accounts for diffusion. Optimal control in this situation can be made parallel to Kalman filter.