

Euler-Lagrange Equation

Yongchao Zhang

NORTHWEST UNIVERSITY

UPDATE: September 10, 2020

1 Euler-Lagrange Equation

参考: [欧拉-拉格朗日方程](#), 以及 [寻找最好 \(2\) – 欧拉-拉格朗日方程](#)

欧拉-拉格朗日方程 (Euler-Lagrange equation) 为变分法中的一条重要方程. 它提供了求泛函的平稳值的一个方法, 其最初的想法是初等微积分理论中的“可导的极值点一定是稳定点 (临界点)”. 当能量泛函包含微分时, 用变分方法推导其证明过程, 简单地说, 假设当前的函数 (即真实解) 已知, 那么这个解必然使能量泛函取全局最小值.

1.1 泛函

我们很清楚函数的概念, 它大致是, 将一个自变量扔到一个黑盒里, 经过黑盒的暗箱操作, 最终得到了一个因变量:

$$x \rightarrow \boxed{f} \rightarrow \mathbb{R},$$

泛函是将一个函数作为自变量, 经过黑盒, 最终得到一个实数的因变量:

$$f(x) \rightarrow \boxed{L} \rightarrow \mathbb{R},$$

可以说, 泛函就是函数的函数, 是更广泛意义上的函数.

1.2 欧拉-拉格朗日方程

1.2.1 最速降线

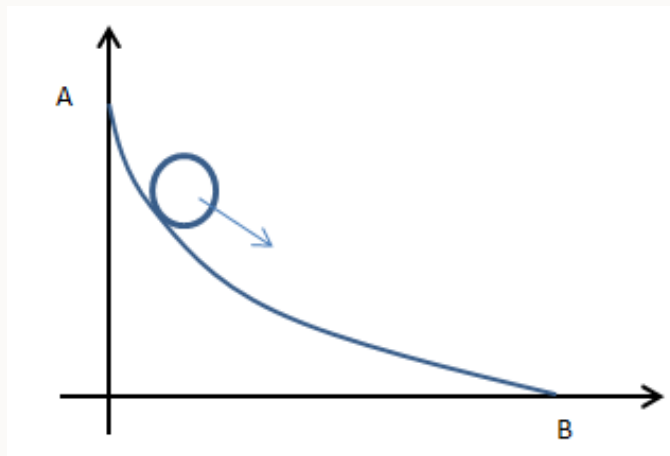
有一种泛函称为简单泛函, 如下所示:

$$\mathcal{A}[f] = \int_{x_1}^{x_2} L(x, f(x), f'(x)),$$
$$f(x) \rightarrow \boxed{\mathcal{A}[f]} \rightarrow \mathbb{R},$$

其中 L 是一个确定的函数, 之所以叫简单函数, 是因为只传递了三个参数, 复杂一点的话还可以继续传递 f 的高阶导数. 现在的问题是, 如果 \mathcal{A} 处于极值点, 它对应的 $f(x)$ 是什么?

这实际上是求一个具体函数, 使得泛函能够取得极值. 一个典型的例子是最速降线问题: 从所有连接不在同一个铅垂线上的定点 A 、 B 的曲线中找出一条曲线, 使得初始速度为 0 的质点, 受重力作用, 由 A 沿着曲线滑下时以最短的时间到达 B , 如图 1:

图 1: A 到 B .

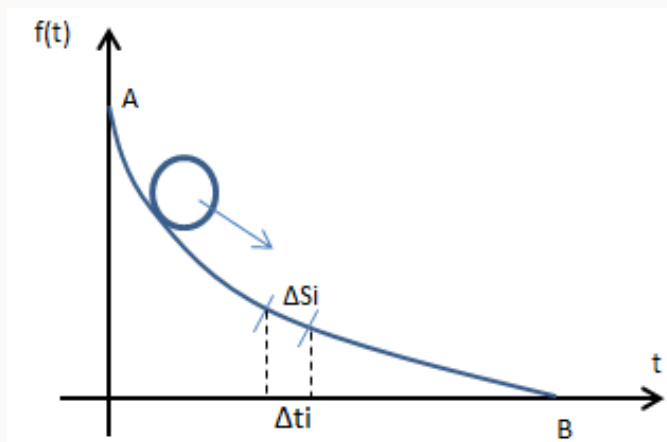


这里我们将曲线看作路径 f 关于时间 t 的函数, 如图 2,

其中 ΔS_i 是在极短时间 Δt_i 内沿着曲线移动的微小弧长, 此时的瞬时速度是 ΔV_i , 距离 = 速度 \times 时间:

$$\Delta S_i \approx \Delta V_i \Delta t_i, \quad \Delta t_i \approx \frac{\Delta S_i}{\Delta V_i},$$

图 2: f 关于时间 t 的函数



重力加速的推论, 在 t 时间处的速度 $V^2 = 2gh$:

$$\Delta t_i \approx \frac{\Delta S_i}{\Delta V_i} = \frac{\Delta S_i}{\sqrt{2gf(t_i)}},$$

质点从 A 点到 B 点的总时间:

$$\begin{aligned} T &\approx \sum_{i=1}^n \frac{\Delta S_i}{\Delta V_i} \sum_{i=1}^n \frac{\Delta S_i}{\sqrt{2gf(t_i)}} \\ &= \int_{t_A}^{t_B} \frac{dS}{\sqrt{2gf(t)}} dt. \end{aligned}$$

根据弧长公式, 可以将 dS 化简, 进一步写成 (关于弧长公式, 可参考: [数学笔记 25--弧长和曲面面积](#)):

$$ds = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_a^b \sqrt{1 + (y')^2} dx,$$

因此

$$T = \int_{t_A}^{t_B} \frac{dS}{\sqrt{2gf(t)}} dt = \int_{t_A}^{t_B} \frac{\sqrt{1 + f'(t)^2}}{\sqrt{2gf(t)}} dt.$$

把结论和简单泛函做个对比, 可以看到二者形式相吻合:

$$\begin{aligned}\mathcal{A}[f] &= \int_{x_1}^{x_2} L(x, f(x), f'(x)) dx, \\ T[f] &= \int_{t_A}^{t_B} \frac{\sqrt{1 + f'(t)^2}}{\sqrt{2gf(t)}} dt, \\ T[f] \rightarrow \mathcal{A}[f], \quad t \rightarrow x, \quad \frac{\sqrt{1 + f'(t)^2}}{\sqrt{2gf(t)}} &\rightarrow L(x, f(x), f'(x)).\end{aligned}$$

现在回到最初的问题, AB 间有无数条曲线, 每条曲线都可以求得时间 $f(t)$, 在众多的曲线中, 有一条唯一的曲线能够使得 $T[f]$ 取得最小值, 这个 $f(t)$ 应该长成什么样?

1.2.2 EL 方程的推导

这里先抛开具体的速降问题, 只看 $\mathcal{A}[f]$, 并且假设 $f_0(x)$ 就是符合条件的最优函数. 现在, 将 $f_0(x) + k(x)$ 定义为是有别于最优曲线 $f_0(x)$ 的其它函数, 其中 $k(x)$ 可以是任意函数. 如果如下定义 $f(x, k)$:

$$f(x, k) = (f_0 + k)(x), \quad f'(x, k) = (f_0 + k)'(x).$$

因为 $f_0(x)$ 是最优函数, 此时 $\mathcal{A}[f_0]$ 有最小值, 则一定有:

$$\begin{aligned}\mathcal{A}[f(x, k)] &\geq \mathcal{A}[f_0], \\ \int_{x_1}^{x_2} L(x, f(x, k), f'(x)) dx &\geq \int_{x_1}^{x_2} L(x, f_0(x), f'_0(x)) dx.\end{aligned}$$

由于任意函数 $k(x)$ 不好定义, 为了能够使 $k(x)$ 任意小, 令:

$$k(x) = \epsilon \eta(x) \quad \forall \epsilon \in \mathbb{R},$$

其中, η 是任意函数, 当 ϵ 取极小值时, 可以看作是对 $f_0(x)$ 的轻微扰动; 还需要额外定义的是, 在端点处是不能扰动的, 即 $\epsilon \eta(A) = \epsilon \eta(B) = 0$, 这对于任意 ϵ 都适用, 所以 $\eta(A) = \eta(B) = 0$, 如图 3. 注意 ϵ 是对 $f_0(x)$ 的扰动程度, $\epsilon \eta(x)$ 是扰动后的增量, $\epsilon \eta(x) = 0$ 说明扰动为 0, 也就是无扰动, $f_0(x) + \epsilon \eta(x)$ 才是扰动后的函数.

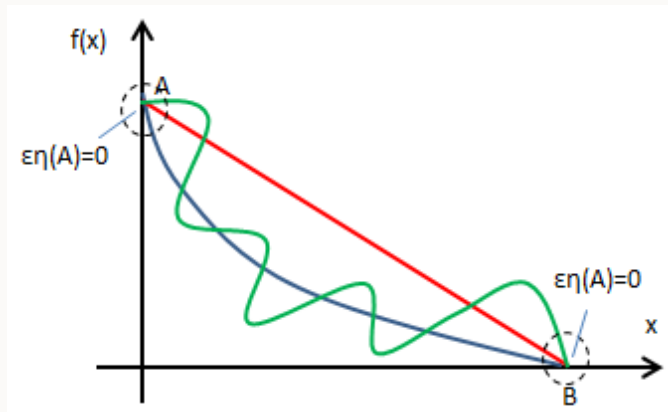
由于假设 $f_0(x)$ 是最优函数, 所以可以将 $f_0(x)$ 看作已经确定的函数, 如果再将任意函数 $\eta(x)$ 看作一个确定的函数, 那么 \mathcal{A} 可以看成 ϵ 的函数, 若 $\mathcal{A}' = 0$, 函数存在极值, 这已经变成了极值点的求解:

$$\begin{aligned}f(x, \epsilon) &= (f_0 + \epsilon \eta)(x), \quad f'(x, \epsilon) = (f_0 + \epsilon \eta)'(x), \\ \mathcal{A}[\epsilon] &= \int_{x_A}^{x_B} L(x, (f_0 + \epsilon \eta)(x), (f_0 + \epsilon \eta)'(x)) dx = \int_{x_A}^{x_B} L(x, f(x, \epsilon), f'(x, \epsilon)) dx.\end{aligned}$$

根据链式法则 (可参考[多变量微积分笔记 4--全微分与链式法则](#)):

$$\text{if } f = f(x, y, z), \quad x = x(t), \quad y = y(t), \quad z = z(t),$$

图 3: f 关于时间 x 的函数



then $\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial t},$

hence,

$$\frac{d\mathcal{A}}{d\epsilon} = \int_{x_A}^{x_B} \frac{dL}{d\epsilon} dx = \int_{x_A}^{x_B} \frac{\partial L}{\partial x} \underbrace{\frac{\partial x}{\partial \epsilon}}_{=0} + \frac{\partial L}{\partial f} \frac{\partial f}{\partial \epsilon} + \frac{\partial L}{\partial f'} \frac{\partial f'}{\partial \epsilon} dx = \int_{x_A}^{x_B} \frac{dL}{d\epsilon} dx = \int_{x_A}^{x_B} \frac{\partial L}{\partial f} \frac{\partial f}{\partial \epsilon} + \frac{\partial L}{\partial f'} \frac{\partial f'}{\partial \epsilon} dx,$$

where,

$$\begin{aligned} \frac{\partial f}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} (f_0 + \epsilon\eta)(x) = \frac{\partial}{\partial \epsilon} (f_0(x) + \epsilon\eta(x)) = \eta(x), \\ \frac{\partial f'}{\partial \epsilon} &= \frac{\partial}{\partial \epsilon} (f_0 + \epsilon\eta)'(x) = \frac{\partial}{\partial \epsilon} (f_0'(x) + \epsilon\eta'(x)) = \eta'(x), \\ &\implies \end{aligned}$$

$$\frac{d\mathcal{A}}{d\epsilon} = \int_{x_A}^{x_B} \frac{\partial L}{\partial f} \eta(x) + \frac{\partial L}{\partial f'} \eta'(x) dx = \int_{x_A}^{x_B} \frac{\partial L}{\partial f} \eta(x) dx + \int_{x_A}^{x_B} \frac{\partial L}{\partial f'} \eta'(x) dx,$$

根据分部积分 (可参考 [单变量微积分笔记 24--分部积分](#)):

$$\begin{aligned} \int_a^b uv' dx &= uv|_a^b - \int_a^b u'v dx \\ \Rightarrow \int_{x_A}^{x_B} \frac{\partial L}{\partial f'} \eta'(x) dx &= \frac{\partial L}{\partial f'} \eta(x) \Big|_{x_A}^{x_B} - \int_{x_A}^{x_B} \left(\frac{\partial L}{\partial f'} \right)' \eta(x) dx \\ \frac{\mathcal{A}}{d\epsilon} &= \int_{x_A}^{x_B} \frac{\partial L}{\partial f} \eta(x) dx + \frac{\partial L}{\partial f'} \eta(x) \Big|_{x_A}^{x_B} - \int_{x_A}^{x_B} \left(\frac{\partial L}{\partial f'} \right)' \eta(x) dx, \end{aligned}$$

from $\eta(x_A) = \eta(x_B) = 0$, we get $\frac{\partial L}{\partial f'} \eta(x) \Big|_{x_A}^{x_B} = 0$,

$$\begin{aligned} \frac{\mathcal{A}}{d\epsilon} &= \int_{x_A}^{x_B} \frac{\partial L}{\partial f} \eta(x) dx - \int_{x_A}^{x_B} \left(\frac{\partial L}{\partial f'} \right)' \eta(x) dx, \\ &= \int_{x_A}^{x_B} \left[\frac{\partial L}{\partial f} - \left(\frac{\partial L}{\partial f'} \right)' \right] \eta(x) dx \\ &= \int_{x_A}^{x_B} \left[\frac{\partial L}{\partial f} - \frac{d}{dx} \left(\frac{\partial L}{\partial f'} \right) \right] \eta(x) dx, \end{aligned}$$

if $\mathcal{A}[\epsilon]$ has min value, then $\frac{d\mathcal{A}}{d\epsilon} = 0$,

means, $\int_{x_A}^{x_B} \left[\frac{\partial L}{\partial f} - \frac{d}{dx} \left(\frac{\partial L}{\partial f'} \right) \right] \eta(x) dx = 0$,

hence, $\frac{\partial L}{\partial f} - \frac{d}{dx} \left(\frac{\partial L}{\partial f'} \right) = 0$ or $\eta(x) = 0$,

$\eta(x) = 0$ 说明对 $f_0(x)$ 无扰动时, \mathcal{A} 能取得极值, 但它对 f_0 的具体形式无任何帮助; 因此最优函数 $f_0(x)$ 的具体形式由第一个解确定:

$$\frac{\partial L}{\partial f} - \frac{d}{dx} \left(\frac{\partial L}{\partial f'} \right) = 0.$$

这就是欧拉-拉格朗日方程 (Euler-Lagrange equation), 可以帮助我们求解泛函下的极值, 这里 L 是已知的. 这个方程是泛函中非常重要的方程, 也是非常经典的能量泛函极小化的方法, 不论在物理还是计算机领域, 应用非常广泛. 所谓能量泛函, 是指微分的范数平方再积分.

它的最初的思想来源于微积分中“可导的极值点一定是稳定点 (临界点)”. 它的精髓思想在于: 假定当前泛函的解已知, 那么这个解必然使得泛函取得最小值 (假定是最小值). 换言之, 只要在泛函中加入任何扰动, 都会使泛函的值变大, 所以扰动为 0 的时候, 就是泛函关于扰动的一个极小值. 所以当扰动的能量趋近于 0, 泛函关于扰动的导数也是 0. 关键是扰动如何表示. 答案是扰动用一个很小的数 ϵ 乘上一个连续函数. 当 ϵ 趋近于 0, 意味着扰动也趋近于 0. 所以当 ϵ 为 0 的时候, 泛函对 ϵ 的导数也为 0. 这就非常巧妙的把对函数求导的问题转化成了一个单因子变量求导的问题. 就是这个思想的伟大之处.

需要注意的是, 欧拉-拉格朗日方程的前提条件是端点不会扰动, 也就是说需要固定两个端点.

2 最速降线的解

有了欧拉-拉格朗日方程, 终于可以计算最速降线的最优解:

$$\begin{aligned} T[f] &= \int_{x_A}^{x_B} \frac{\sqrt{1 + f'(t)^2}}{\sqrt{2gf(t)}} dt, \\ L(t, f(t), f'(t)) &= L(f, f') = \frac{\sqrt{1 + f'^2}}{\sqrt{2gf}}, \\ \text{the EL-Equation is: } &\frac{\partial L}{\partial f} - \frac{d}{dt} \left(\frac{\partial L}{\partial f'} \right) = 0, \end{aligned} \tag{1}$$

其中 $f = f(t)$. 由于:

$$\begin{aligned} \frac{dL}{dt} &= \frac{d}{dt} L(f, f') = \frac{\partial L}{\partial f} \frac{\partial f}{\partial t} + \frac{\partial L}{\partial f'} \frac{\partial f'}{\partial t} = f' \frac{\partial L}{\partial f} + f'' \frac{\partial L}{\partial f'}, \\ \frac{d}{dt} \left(f' \frac{\partial L}{\partial f'} \right) &= f'' \frac{\partial L}{\partial f'} + f' \frac{d}{dt} \left(\frac{\partial L}{\partial f'} \right) \\ \frac{d}{dt} \left(L - f' \frac{\partial L}{\partial f'} \right) &= \frac{dL}{dt} + \frac{d}{dt} \left(f' \frac{\partial L}{\partial f'} \right) \\ &= \left(f' \frac{\partial L}{\partial f} + f'' \frac{\partial L}{\partial f'} \right) - \left(f'' \frac{\partial L}{\partial f'} + f' \frac{d}{dt} \left(\frac{\partial L}{\partial f'} \right) \right) \\ &= f' \frac{\partial L}{\partial f} - f' \frac{d}{dt} \left(\frac{\partial L}{\partial f'} \right) \end{aligned}$$

$$= f' \underbrace{\left[\frac{\partial L}{\partial f} - \frac{d}{dt} \left(\frac{\partial L}{\partial f'} \right) \right]}_{EL-Equation=0} = 0$$

hence, $L - f' \frac{\partial L}{\partial f'} = C.$

将 (1) 代入:

$$L - f' \frac{\partial L}{\partial f'} = \frac{\sqrt{1+f'^2}}{\sqrt{2gf}} - f' \frac{\partial}{\partial f'} \left(\frac{\sqrt{1+f'^2}}{\sqrt{2gf}} \right) = C,$$

by chain rule:

$$\begin{aligned} \frac{\partial L}{\partial f'} &= \frac{\partial}{\partial f'} \left(\frac{\sqrt{1+f'^2}}{\sqrt{2gf}} \right) \\ &= \frac{f'}{\sqrt{2gf} \sqrt{1+f'^2}}, \end{aligned}$$

因此

$$\begin{aligned} L - f' \frac{\partial L}{\partial f'} &= \frac{\sqrt{1+f'^2}}{\sqrt{2gf}} - f' \frac{f'}{\sqrt{2gf} \sqrt{1+f'^2}} \\ &= \frac{1}{\sqrt{2gf} \sqrt{1+f'^2}} \\ &= C, \end{aligned}$$

\implies

$$\frac{1}{2gf(1+f'^2)} = C \implies f(f+f'^2) = \frac{1}{2gC^2}.$$

现在, 使用参数方程:

$$\text{let, } \frac{1}{1gC^2} = 2r, \quad t = t(\theta), \quad f' = \cot \frac{\theta}{2} = \frac{\cos(\theta/2)}{\sin(\theta/2)},$$

之所以令 $f = \cot(\theta/2)$, 是因为在 θ 的定义域 $[0, 2\pi]$ 上, f 可以取任意值, 如图 4.

所以有

$$\begin{aligned} (f)(1+f'^2) &= (f)\left(1+\cot^2\frac{\theta}{2}\right) \\ &= \end{aligned}$$

图 4: $y = \cot(x/2)$, $0 \leq x \leq 2\pi$

