Day 6. 拉格朗日乘子法

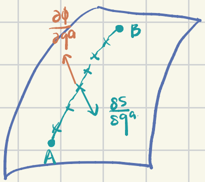
之前的做法：直接观察约束 → 选择独立广义坐标。而在很多情形下，难以直接约束求出广义坐标 ⇒ 带着约束求问题，此时应使用乘子方法。

先从函数的乘子法来看一看，求 $F(x,y)$ 在约束 $\phi(x,y)=0$ 下的极值。一个直观点：若约束曲线与等高线不平行，则沿着约束线的函数值还有增加或减少。在约束极值处，函数梯度必与约束法向平行。

$\Rightarrow \nabla F=-\lambda\nabla\phi$ ⇒ $\frac{\partial F}{\partial x}+\lambda\frac{\partial\phi}{\partial x}=0$，$\frac{\partial F}{\partial y}+\lambda\frac{\partial\phi}{\partial y}=0$。这相当于要求一个扩展的函数的无约束极值：$\tilde{F}(x,y,\lambda)=F(x,y)+\lambda\phi(x,y)$。同样在求泛函极值时，也可使用类似方法。

简单起见，我们考虑有个完整约束的情形。有约束情形：$\delta S=-\int dt\left(\frac{d}{dt}\frac{\partial L}{\partial\dot q^a}-\frac{\partial L}{\partial q^a}\right)\delta q^a=0$。设有一个完整约束 $\phi(t,q)=0$，它使得广义坐标变分不独立，⇒ $\delta\phi=\frac{\partial\phi}{\partial q^a}\cdot\delta q^a=0$。

此时轨道只可在约束面上变动。$\frac{\delta S}{\delta q^a}$：若设坐标 q^a 以广义坐标各一点作用变换方向 $\left(\frac{\delta S}{\delta q^a}\right)$ 这个量是"调整轨道上某一点、作用量变化最大的方向"，而 $\frac{\partial\phi}{\partial q^a}$ 则是约束曲面的法向。



一个直观的理解：在轨道上有许多"控制点"，我们要推动这些点，使它们点点处 $\frac{\delta S}{\delta q^a}$ 与 $\frac{\partial\phi}{\partial q^a}$ 平行，从而使作用量有约束下的极值。⇒ $\frac{\delta S}{\delta q^a}=-\lambda\frac{\partial\phi}{\partial q^a}$

从而我们可以使用扩展的作用量：$\tilde{S}[q,\lambda]=\int dt\,\tilde{L}=\int dt\,[L(t,q,\dot q)+\lambda(t)\cdot\phi(t,q)]$　$\lambda(t)$ 是时变的，因为 $\frac{\delta S}{\delta q^a}$ 与 $\frac{\partial\phi}{\partial q^a}$ 在轨道上逐点平行，各点 λ 可不同。

从上给出：
$$\begin{cases}\frac{d}{dt}\frac{\partial L}{\partial\dot q^a}-\frac{\partial L}{\partial q^a}-\lambda\frac{\partial\phi}{\partial q^a}=0\\ \frac{d}{dt}\frac{\partial\tilde L}{\partial\dot\lambda}-\frac{\partial\tilde L}{\partial\lambda}=-\phi=0\end{cases}$$

从几何上可以直接看出：乘子法成立的条件是约束被表示为位形空间中的流形，因此不可推广任意的非完整约束情形，只有个别例外。

→ 若非完整约束有如下形式：$\phi(t,q,\dot q)=A_a(t,q)\dot q^a+B(t,q)=0$（只有 $\dot q^a$ 的线性项），从而有 $d\phi=A_a\cdot dq^a+B\cdot dt=0$ ⇒ 在任一瞬时有 $A_a\delta q^a=0$，类比上面完整约束情形有 $\frac{\delta S}{\delta q^a}=-\lambda A_a$

从而有 $\frac{d}{dt}\frac{\partial L}{\partial\dot q^a}-\frac{\partial L}{\partial q^a}=\lambda\cdot A_a$。

→ 若为积分给出的"等周约束"：$\int dt\cdot\phi(t,q,\dot q)=0$，它也可以用乘子法的：$\tilde{S}[q]=\int dt\,(L+\lambda\phi)=0$。直观解释：约束这个泛函，相当于在位形空间中可行轨道族，则在轨道上 L 的导数必平行于 ϕ 的导数。

回到最简单的完整约束的情形：我们发现 λ 只有自身出现在 L 中，从而广义坐标可分两类：
- 动力学变量：广义速度出现在 L 中。
- 辅助(auxiliary)变量：只有自身出现，广义速度不出现。（λ 是辅助变量特例）

从而将 L 写成如下形式：$L=L(t,q,\dot q,x)$，简单起见，考虑一个 q 与一个 x。

$$S[q;x]=\int dt\cdot L(t,q,\dot q,x)\Rightarrow\begin{cases}\frac{\partial L(t,q,\dot q,x)}{\partial q}-\frac{d}{dt}\frac{\partial L(t,q,\dot q,x)}{\partial\dot q}=0\\ \frac{\partial L(t,q,\dot q,x)}{\partial x}=0\end{cases}$$

⇒ 为 x 的代数方程，可解出 $x=x(t,q,\dot q)$，从而它完全由 t,q 决定，没有独立的动力学演化方程！

可给出所谓"等效作用量"：$S_{eff}[q]=\int dt\cdot L(t,q,\dot q,x(t,q,\dot q))=\int dt\,L_{eff}(t,q,\dot q)$

下面给一些辅助变量的小技巧：→ 广义速度线性化：若 $\frac{\partial^2 L}{\partial\dot q^2}\neq 0$（$L$ 为 $\dot q$ 的非线性函数），引入乘子 λ 与辅助 x。$S[q,\lambda,x]=\int dt\,[L(t,q,x)+\lambda(x-\dot q)]$

可求得有：$\lambda=-\frac{\partial L}{\partial x}$，从而作用量被写成等价形式：$S[q,x]=\int dt\,[L(t,q,x)-\frac{\partial L}{\partial x}(x-\dot q)]$，这里广义速度变成线性的。

→ 高阶导的降阶，设系统的拉氏量中有高阶导数，$S[q]=\int dt\cdot L(t,q,\dot q,\ddot q)$，可给出含辅助变量的等价作用量：$S[q,\xi,\lambda]=\int dt\,[L(t,q,\dot q,\dot\xi)-\lambda(\xi-\dot q)]$