伴随法（Adjoint method），拉格朗日乘子法（Lagrange multiplier method），偏微分方程约束优化（PDE-constrained optimization）

卐 Eternally

考虑一个优化问题:

$$\min_p \quad f(x) = \int_0^T x dt \text{ , s.t. } \begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = bx \\ x(0) - a = 0 \end{cases} \text{ , } p = [a, b]^T \text{ , } x = x(p)$$

优化算法需要我们计算 $\frac{\mathrm{d}f}{\mathrm{d}p} = \int_0^T x_p dt$

但是在大多数情况下，计算 x_p 是困难的，而且不能使用ODE求解器

于是我们利用伴随法来避免计算 x_p

第一步，引入优化问题的拉格朗日乘子法

$$L \equiv \int_0^T [x + \lambda(\frac{\mathrm{d}x}{\mathrm{d}t} - bx)]dt + \mu(x(0) - a)$$

其中λ 是一个关于时间的函数，μ 是一个向量。因为 $\frac{\mathrm{d}x}{\mathrm{d}t} - bx = 0$， $x(0) - a = 0$,所以我们可以随意的设置 λ 和 μ 的值，并且 $\frac{\mathrm{d}L}{\mathrm{d}p} = \frac{\mathrm{d}f}{\mathrm{d}p}$， 于是

$$\frac{\mathrm{d}L}{\mathrm{d}p} = \int_0^T x_p + \lambda(-bx_p + \frac{\mathrm{d}x_p}{\mathrm{d}t} + [0, -x])dt + \mu(x(0)_p + [-1, 0])$$

该积分包含 x_p 和 $\frac{\mathrm{d}x_p}{\mathrm{d}t}$ 两项，利用分部积分消除第二项

$$\int_0^T \lambda \frac{\mathrm{d}x_p}{\mathrm{d}t} dt = \lambda x_p|_0^T - \int_0^T \frac{\mathrm{d}\lambda}{\mathrm{d}t} x_p dt$$

$$\begin{aligned} \frac{\mathrm{d}L}{\mathrm{d}p} &= \int_0^T x_p - b\lambda x_p + \lambda[0, -x]dt + \lambda x_p|_0^T - \int_0^T \frac{\mathrm{d}\lambda}{\mathrm{d}t} x_p dt + \mu(x(0)_p + [-1, 0]) \\ &= \int_0^T (1 - b\lambda - \frac{\mathrm{d}\lambda}{\mathrm{d}t}) x_p dt + \int_0^T \lambda[0, -x]dt + \lambda(T)x_p(T) + (\mu - \lambda(0))x_p(0) \\ &\qquad + \mu[-1, 0]) \end{aligned}$$

令 $1 - b\lambda - \frac{\mathrm{d}\lambda}{\mathrm{d}t} = 0$， $\lambda(T) = 0$, $\mu - \lambda(0) = 0$,于是

$$\frac{\mathrm{d}L}{\mathrm{d}p} = \int_0^T \lambda[0, -x]dt + \mu[-1, 0]$$

于是求解 $\frac{\mathrm{d}f}{\mathrm{d}p}$ 需要求解三个ODE方程:

如果是解析解方法，这里
相当于生成函数族了
如果是数值解，就是近似
函数的代入离散计算

1. 求解 $\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = bx \\ x(0) = a \end{cases} \Rightarrow x(t) = ae^{bt}$
2. 求解 $\begin{cases} 1 - b\lambda - \frac{\mathrm{d}\lambda}{\mathrm{d}t} = 0 \\ \lambda(T) = 0 \end{cases} \Rightarrow \lambda(t) = b^{-1}(1 - e^{b(T-t)})$
3. 令 $\frac{\mathrm{d}L}{\mathrm{d}p} = \int_0^T \lambda[0, -x]dt + \mu[-1, 0] \Rightarrow \begin{cases} \frac{\partial f}{\partial a} = b^{-1}(-1 + e^{bT}) \\ \frac{\partial f}{\partial b} = \frac{a}{b} T e^{bT} - \frac{a}{b^2}(e^{bT} - 1) \end{cases}$

作为验证，可以对 $f(x) = \int_0^T x dt = \int_0^T ae^{bt} dt = \frac{a}{b}(e^{bT} - 1)$ 关于 p 求导，会得到相同的结果。