

变分法简介Part 1. (Calculus of Variations)

Dr.Stein  
Ph.D. 计算力学 物理仿真

1,522 人赞同了该文章

1. 泛函数 (Functionals)

简而言之，泛函数是函数的函数，即它的输入是函数，输出是实数。而这个输出值取决于一个或多个函数（输入）在一整个路径上的积分而非像一般函数一样取决于离散的变量。这样说可能还是比较抽象，不过坚持看到下文的*Example 1*就可以更好理解了。

通常在变分法中，泛函数是一个积分，记做 $I$ 。

$$I(y) = \int_{x_1}^{x_2} F dx$$

其中我们想要通过选择被积函数 $F$ 来最大化或最小化泛函数 $I$ 的值。同时我们称 $F$ 为拉格朗日函数 (Lagrange function)。 $F$ 可以是函数 $y(x)$ 和 $y(x)$ 各阶导数的函数(以下 $y(x)$ 均简写成 $y$ )。为了说明方便，我们先姑且设 $F$ 是 $y$ 和 $y'$ 的函数，所以我们可以进一步将泛函数 $I$ 写成：

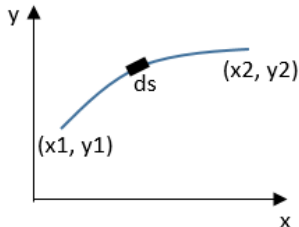
$$I(y) = \int_{x_1}^{x_2} F(y, y'; x) dx$$

积分里面我用分号;将 $x$ 和前面的 $y$ 隔开代表 $y$ 和 $y'$ 是 $x$ 的函数。一般 $F$ 和 $y$ 的函数关系是已知的，所以想要最大或最小化泛函数，实际上是通过选择适当的函数 $y(x)$ 。

为了透彻理解这个概念，我们可以来看一个简单的例子。

Example 1.

一个最简单直观的例子是求两个固定点之间的最短路径。当然大家都知道两点之间直线最短，这里可以用变分法做出解释。



如上图所示路径是一任意路径，我们取区中一小段微元 $ds$ ，可以容易计算微元断的长度为：

$$ds^2 = dx^2 + dy^2 = [1 + (y')^2]dx^2,即：$$

$$ds = \sqrt{1 + (y')^2}dx$$

积分得到总的路径长度为：

$$L = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + (y')^2} dx$$

这个例子中， $L$ 是泛函数， $\sqrt{1 + (y')^2}$ 是拉格朗日函数 $F$ ，我们想要找一个函数 $y(x)$ 使得泛函数 $L$ 的值最小。这次Part 1.的任务就是为解决这个问题做准备。Part 2.中我们会用变分法证明这个 $y(x)$ 确实是直线的方程。

2. 泛函数的极值

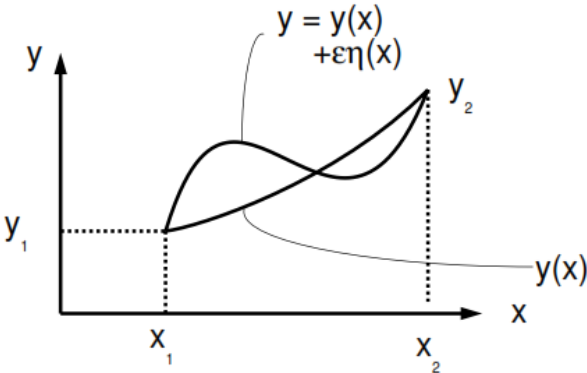
这里重申下，泛函数 $I$ 在区间 $[x_1, x_2]$ 上的值取决于积分路径的选择，即取决于函数 $y(x)$ 的选择。我们有理由假设存在一个这样的 $u(x)$ ，可以使得泛函数 $I$ 取到极值。而在这个 $u(x)$ 附近的任意路



$$\tilde{y}(x) = y(x) + \epsilon \eta(x)$$

这样做的好处是对于一个给定的 $\eta(x)$ ，我们可以通过改变 $\epsilon$ 的值来得到无穷多的路径，同时对于任何 $\eta(x)$ ，当 $\epsilon = 0$ 的时候， $\tilde{y}(x)$ 和 $y(x)$ 重合。

图像直观表示如下图：



由于在边界条件的限制， $\eta(x_1) = \eta(x_2) = 0$ 。这样就能保证 $\tilde{y}(x)$ 可以通过两个固定端点。

这时我们可以说， $y(x)$ 所对应的泛函数 $I$ 的值是泛函数 $\tilde{I} = \int_{x_1}^{x_2} F(\tilde{y}, \tilde{y}'; x) dx$ 的极值。我们可以进一步用 $\epsilon$ 表示 $\tilde{I}$ ：

$$\tilde{I} = \int_{x_1}^{x_2} F(\tilde{y}, \tilde{y}'; x) dx = \int_{x_1}^{x_2} F(y + \epsilon \eta, y' + \epsilon \eta'; x) dx$$

虽然 $y(x)$ 未知，但是根据之前的合理假设， $y(x)$ 是一个存在的确定函数。所以根据上式，如果给定一个特定的 $\eta(x)$ ， $\tilde{I}$ 的变化只取决于 $\epsilon$ 的变化。所以我们现在可以把 $\tilde{I}$ 看做是 $\epsilon$ 的函数。用泰勒展开公式将 $\tilde{I}$ 在 $\epsilon = 0$ 处展开得到：

$$\tilde{I}(\epsilon) = \tilde{I}|_{\epsilon=0} + \left(\frac{d\tilde{I}}{d\epsilon}\right)\Big|_{\epsilon=0} \cdot \epsilon + \left(\frac{d^2\tilde{I}}{d\epsilon^2}\right)\Big|_{\epsilon=0} \cdot \frac{\epsilon^2}{2!} + \cdots = \tilde{I}_0 + \tilde{I}_1\epsilon + \tilde{I}_2\epsilon^2 + \cdots$$

很明显，当 $\epsilon = 0$ 时， $\tilde{I}|_{\epsilon=0} = I$ ,带上式可得到：

$$\tilde{I} - I = \tilde{I}_1\epsilon + \tilde{I}_2\epsilon^2 + \cdots$$

这里我们记 $\delta I = \tilde{I}_1\epsilon = \frac{d\tilde{I}}{d\epsilon}\Big|_{\epsilon=0} \cdot \epsilon$ ，并称之为**一阶变分**。同理二阶变分为 $\delta I^2 = \tilde{I}_2\epsilon^2$ 。

（这里插一句变分和微分的区别。变分在上图的直观解释是 $\tilde{y}$ 和 $y$ 在竖直方向上的距离，称之为 $\delta y$ ，所以这个差是在同一个x上计算的。而微分则是由于x的微小变动引起的y的变动。）

然后我们可以类比求函数极值时的做法。求函数极值时，我们会令函数的一阶导数为零。这里同样，为了求泛函数 $\tilde{I}$ 的极值，我们令一阶变分 $\delta I = 0$ 。现在我们计算化简 $\delta I$ :

$$\delta I = \left(\int_{x_1}^{x_2} \frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0} dx\right) \cdot \epsilon$$

$$\frac{d\tilde{F}}{d\epsilon}\Big|_{\epsilon=0} \cdot \epsilon = \left(\frac{\partial \tilde{F}}{\partial \tilde{y}} \cdot \frac{d\tilde{y}}{d\epsilon} + \frac{\partial \tilde{F}}{\partial \tilde{y}'} \cdot \frac{d\tilde{y}'}{d\epsilon}\right)\Big|_{\epsilon=0} \cdot \epsilon$$

因为  $\tilde{y}(x) = y(x) + \epsilon \eta(x)$ ，不难得到： $\frac{d\tilde{y}}{d\epsilon} = \eta, \frac{d\tilde{y}'}{d\epsilon} = \eta'$  ,另外我们有 $\delta y = \epsilon \eta, \delta y' = \epsilon \eta'$

又因为当 $\epsilon \rightarrow 0$ 时， $\tilde{F} \rightarrow F, \tilde{y} \rightarrow y, \tilde{y}' \rightarrow y'$  ,将这些式子带入原式可以得到：

$$\delta I = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y} \delta y + \frac{\partial F}{\partial y'} \delta y'\right) dx$$

终于到最后一步啦，分部积分一下得到：

$$\delta I = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'}\right)\right) \delta y dx + \frac{\partial F}{\partial y'} \delta y \Big|_{x_1}^{x_2}$$

另 $\delta I = 0$ 就可以解得最小化泛函数的y啦。我们注意到 $\delta I$ 有两个部分。对于第一个积分部分，由于 $\delta y$ 是任意的，所以要想使这个部分等于零，需要保证<sup>[1]</sup>：

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'}\right) = 0 \quad (x_1 \leq x \leq x_2)$$

这就是传说中的欧拉-拉格朗日方程(E-L equation)。

而第二部分等于零则是边界条件。

在Part 2.，我们会以用这次介绍的内容和上述方程解决两点之间直线最短的问题为开头，继续介绍变分法。

-----  
-----

注[1]:

假设 $x_1$ 和 $x_2$ 是给定的常数， $\phi(x)$ 是一个特定的在 $x_1 \leq x \leq x_2$ 上连续的函数，那么如果对于任意连续可微的函数 $\eta(x)$ 都成立 $\int_{x_1}^{x_2} \phi(x) \eta(x) dx = 0$ ，则 $\phi(x) = 0 \quad (x_1 \leq x \leq x_2)$ 。

（任意函数和一个非零的特定函数的乘积仍是任意函数，由于无法保证任意函数的积分是零，所以这个特定函数必须在这个区间上恒等于零使得乘积为零，这样可以保证积分为零。）