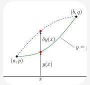
变分法：理解变分

Allan
软件工程师

前两天，看了这篇文章：

烤羚羊：浅谈变分原理
5025 赞同 · 221 评论 文章



这里谈谈我自己的理解。

1. 函数的函数之微分

我们来研究函数的函数的微分形式，设：

$$F = F(y(x))$$

既然研究微分，那么隐含的条件就是：

1）映射 $x \to y$ 是一个可微映射（即连续且可导，所谓可导是指在参数域任意一点上左导数等于右导数）

2）映射 $y \to F$ 是一个可微映射，即y的连续变化同样会引起F的连续变化。

第二点是最容易让人忽视，我们定义函数的函数的导数为：

$$F'_y = \frac{\delta F}{\delta y} = \lim_{y_2 \to y_1} \frac{F(y_2) - F(y_2)}{y_2 - y_1}$$

于是微分为：

$$dF = F'_y \delta y$$

2. 一个简单的抽象问题

按照wiki的定义：泛函（functional）是指以函数构成的向量空间为定义域，**实数为值域**的函数。也就是说泛函是函数的函数的一个特例。比如下面的 I 就是一个泛函：

函数： $y = y(x)$

积分： $I = \int_a^b f(y)dx$

我们现在求解 $y(x)$ 在什么情况下 I 有极值。假设 y 和 x 存在两个不同的函数关系 y_1, y_2 ，则：

$$I_1 = \int_a^b f(y_1)dx$$

$$I_2 = \int_a^b f(y_2)dx$$

两者之差：

$$\Delta I = \int_a^b f(y_2)dx - \int_a^b f(y_1)dx = \int_a^b [f(y_2) - f(y_1)]dx$$

当 $y_1 \to y_2$ 时，不妨令（注：函数的函数的连续可导前提）：

$$y_2 = y_1 + \delta y$$

则：

$\delta I = \int_a^b [f(y_1+\delta y) - f(y_1)]dx$

另外，$\delta y \to 0$ 时，根据微分的定义，显然有：

$f(y_1+\delta y) - f(y_1) = f'|_{y_1} \delta y$

其实就是等于f函数对y函数的微分乘y的变量的微分，近似于求导的链式法则

因此，可以得到一个正式的公式：

$$\delta I = \int_a^b f'(y)\delta y dx \qquad (1)$$

那么 δy 的意义是什么？它是一个常量吗（或者说和x无关的量吗）？

本质上，δy 的定义为：

$\delta y := \lim_{y_1 \to y_2}(y_2 - y_1)$

可见 δy 也是x的函数，更确切的记法是 $\delta y(x)$，因此，并不是一个常量哦！

继续研究公式（1）：

如果 I 处在极值状态，那么意味着 δy 无论任何做什么样的变化（当然是微量变化），δI 都为0，所以只有下面的条件才能保证 I 处在极值状态：

$$f'_y = 0 \qquad (2)$$

2. 欧拉-拉格朗日方程

$I = \int_a^b f(y, y')dx$

依据上面的推理，可以推导出：

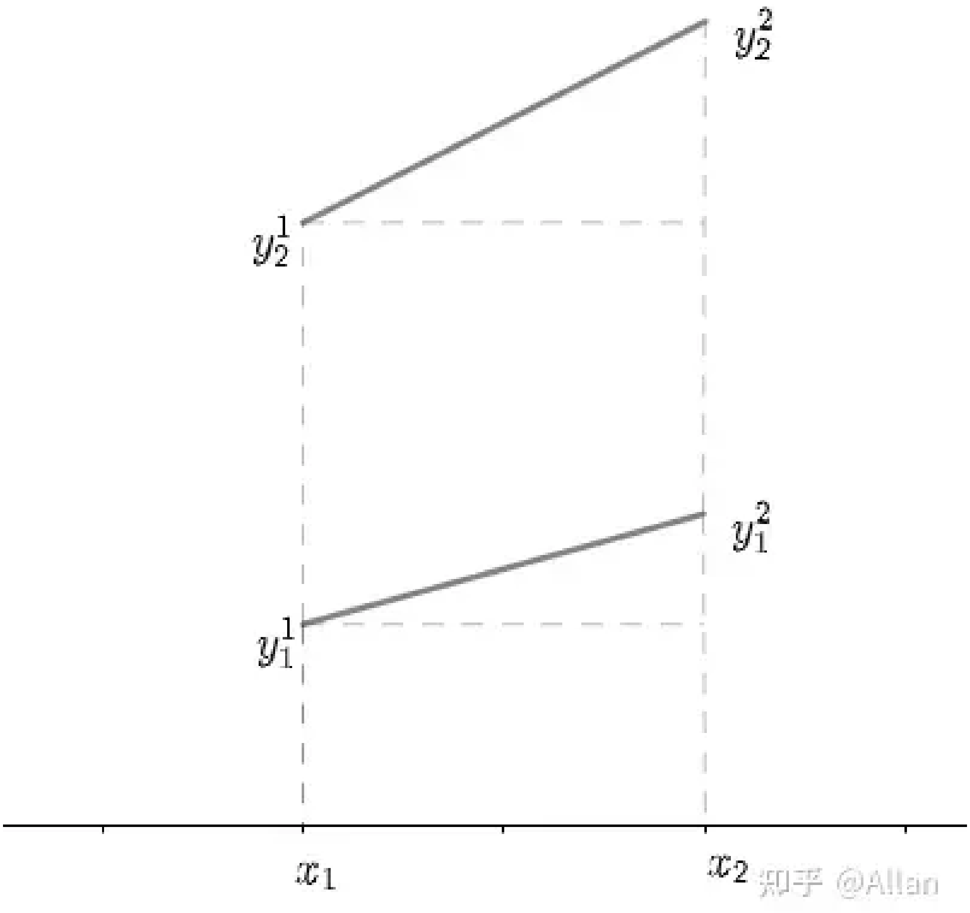
$$\delta I = \int_a^b \delta f(y, y')dx = \int_a^b (\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y')dx \qquad (3)$$

我们先来推导一下 $\delta y'$ 代表的意义：

$\delta y' = \lim_{y_1 \to y_2}(y'_2 - y'_1) = \lim_{y_1 \to y_2}(\frac{dy_2}{dx} - \frac{dy_1}{dx})$

$$= \lim_{y_1 \to y_2} \frac{d(y_2 - y_1)}{dx} = \frac{d(\delta y)}{dx} = (\delta y)' \qquad (4)$$

公式（4）的推导不够直观，我们可以用局部微元线性化的思路来再来阐释一下，如下图，我们可以想象在局部微元区间内，$y_1(x), y_2(x)$ 都是线性变化，于是有：



函数的函数局部微元线性化

$$
\begin{aligned}
\delta y' &= \frac{y_2^2 - y_2^1}{x_2 - x_1} - \frac{y_1^2 - y_1^1}{x_2 - x_1} \\
&= \frac{y_2^2 - y_2^1 - y_1^2 + y_1^1}{x_2 - x_1} \\
&= \frac{y_2^2 - y_1^2}{x_2 - x_1} - \frac{y_2^1 - y_1^1}{x_2 - x_1} \\
&= \frac{(\Delta y)_2 - (\Delta y)_1}{\Delta x} \\
&= \frac{(\delta y)_2 - (\delta y)_1}{dx} \\
&= \frac{d(\delta y)}{dx} \\
&= (\delta y)'
\end{aligned}
$$

它的意义是：**导函数的变分等于变分的导数**。

所以（3）式的第二部分可以表示为：

$$\int_a^b \frac{\partial f}{\partial y'} \delta(y') dx = \int_a^b \frac{\partial f}{\partial y'} (\delta y)' dx = \int_a^b \frac{\partial f}{\partial y'} d(\delta y)$$

然后，我们再用分部积分法进行换元：

$$\int_a^b \frac{\partial f}{\partial y'} d(\delta y) = \int_a^b d(\frac{\partial f}{\partial y'} \delta y) - \int_a^b \delta y d(\frac{\partial f}{\partial y'})$$

$$= \frac{\partial f}{\partial y'} \delta y |_a^b - \int_a^b \frac{d(\frac{\partial f}{\partial y'})}{dx} \delta y dx \qquad (5)$$

因为：$\delta y(a)=\delta y(b)=0$，所以（5）式可以写成：

$$\int_a^b \frac{\partial f}{\partial y'} d(\delta y)=\int_a^b \frac{d(\frac{\partial f}{\partial y'})}{dx}\delta y dx \qquad (6)$$

所以，（3）可以写成：

$$\delta I=\int_a^b [\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})]\delta y dx \qquad (7)$$

看到了没有？（7）式和（1）式是多么的相似！如果 I 取极值，意味着对于任何 δy，都有 $\delta I=0$，所以只能有：

$$\frac{\partial f}{\partial y}-\frac{d}{dx}(\frac{\partial f}{\partial y'})=0 \qquad (8)$$

3. 多元函数的推导

在第2节中推导了 $\delta y'=(\delta y)'$ 公式，它和分部积分法一起构成了欧拉-拉格朗日公式的基础。如果是多元函数的泛函，我们还可以推导出：

$$\delta(\frac{\partial Y}{\partial x_i})=\frac{\partial(\delta Y)}{\partial x_i} \text{ 或 } \delta(Y'_{x_i})=(\delta Y)'_{x_i} \quad (9)$$

上式可以理解为：**偏导数函数的变分等于函数变分的偏导数**。推导如下：

$$\begin{aligned}\delta(Y'_{x_i}) &= \frac{Y_2^2-Y_2^1}{\Delta x_i}-\frac{Y_1^2-Y_1^1}{\Delta x_i}\\ &= \frac{(Y_2^2-Y_1^2)-(Y_2^1-Y_1^1)}{\Delta x_i}\\ &= \frac{(\delta Y)_2-(\delta Y)_1}{\Delta x_i}\\ &= (\delta Y)'_{x_i}\end{aligned}$$

附注：这篇文章还有高维的欧拉-拉格朗日方程的推导过程，对于我们解决实际问题更有用。

说谎的傻子：从欧拉-拉格朗日方程到理论力学和全变分约束降噪
193 赞同 · 20 评论 文章

编辑于 2022-06-02 08:00

变分法 数学