

变分优化方法

CHINHUA

一、知识准备：变分

我要首先谈变分问题，一个是它很重要，因为它表示的是函数改变的数学后果；而一般微分仅仅是讨论自变量改变但函数不变的数学后果。而经济学研究往往又研讨的是函数改变问题。更重要的是，学经济的不少人又往往把函数的改变，与自变量改变所导致的点在函数上的运动，混为一谈。所以弄清楚变分的数学意义，对于理解经济意义，也是极为重要的。

另一个是，变分远不像一般人想象的那么复杂。只要弄清楚了变分的数学意义，运用变分就跟运用微分一样简单。经济学上纷繁复杂的跨期最优化，就成为简单的微分一样，用傻瓜方法也可以做了。

所以在本变分阐述中，我不准备分析哈密顿方程那些东西。我愿意让读者不用哈密顿方程，而用最基本最简单的数学法则，直接求解出结果——换句话说，读者只要掌握了这些基本的数学法则，也能轻松推导出哈密顿方程。

给读者的建议：推荐使用变分方法。因为在变分推导的每一个步骤，你都知道自己推导的经济意义所在，而且转换成哈密顿并不复杂。很多学生虽然会用哈密顿，但知其然不知其所以然，只会套公式，却不知道自己优化的经济含义，这样，解书本上的题还可以，一到现实问题的优化，基本上都会出错误。

(1) 变分的性质

已知函数 $y_0(x)$ ，则其附近的函数曲线可以表示成

$$y(x) = y_0(x) + \delta y(x) \quad (\text{A.20})$$

$\delta y(x)$ 称为 $y(x)$ 的变分。

$$\delta y = \lim_{y(x) \rightarrow y_0(x)} (y(x) - y_0(x)) = \lim_{y(x) \rightarrow y_0(x)} \Delta y \quad (\text{A.21})$$

变量的微分是变量；函数的变分是函数。

因此要完全按照函数来对待 $\delta y(x)$ 。

譬如 $\delta y(x + dx) = \delta y(x) + \frac{d}{dx} \delta y(x) dx$ （事实上应该写成泰勒展开，但我

为了简便而且容易让读者看懂，只写了一阶泰勒展开。况且在微小量下，这已足够）。可以看到，这里完全是把函数变分当作函数来处理的。

变量的微分是不确定的微小变量；函数的变分是不确定的微小函数。

这对于以后求最优化时有用。由于变分是不确定的微小函数，所以只要有一个可以设计的微小函数不符合最优化要求，最优化也不能成立；同时，如果只有某些设计好的微小函数符合最优化要求，最优化自然也不能成立。

由于函数的变分和函数中变量的微分是在完全相互独立的空间进行，所以函数的变分与函数中变量的微分、积分可以相互交换。即：

$$\delta \int f(x) dx = \int \delta f(x) dx \quad (\text{A.22})$$

$$\delta \frac{df(x)}{dx} = \frac{d}{dx} \delta f(x) \quad (\text{A.23})$$

变分的其它运算性质与微分一样，例如：

$$\delta(f_1 + f_2) = \delta f_1 + \delta f_2 \quad (\text{A.24})$$

$$\delta(f_1 f_2) = f_1 \delta f_2 + f_2 \delta f_1 \quad (\text{A.25})$$

由于变分是函数，所以函数取极值时的充分条件是变分为 0。二次变分大于 0 时为极小值，小于 0 时为极大值。（微积分取极值是导数为 0；二次导数大于 0 时为极小值，小于 0 时为极大值。其与变分一样的，都是函数一阶变化为 0，二阶变化判断极大或极小）。即：

$$\delta^2 J(y(x)) = 0, \quad \begin{array}{l} \delta^2 J > 0, \text{极小} \\ \delta^2 J < 0, \text{极大} \end{array} \quad (\text{A.26})$$

引理：如果 $\delta u(a) = \delta u(b) = 0$ ，且对任何 δu 都有 $\int_a^b f(x) \delta u(x) dx = 0$ ，则

在 $[a, b]$ 上， $f(x) \equiv 0$ 。

证明：用反证法。设若在某个区间 $x_0 \in [x_1, x_2]$ 有 $f(x_0) > 0$ ，则取：

$$\delta u(x) = \begin{cases} (x - x_1)(x_2 - x)^2, & x \in [x_1, x_2] \\ 0, & x \notin [x_1, x_2] \end{cases}$$

此时 δu 满足题设条件，但是却有：

$$\int_a^b f(x) \delta u(x) dx = \int_{x_1}^{x_2} f(x) \delta u(x) dx > 0$$

这与命题的前提不符合，所以应该有 $f(x) \equiv 0$ 。证毕。

(2) 变分降阶的关键运用步骤——分部积分法

分部积分法是微积分中很常见的方法。在变分中也是一样的运用。只是由于在变分运算中经常用到，成为降阶的重要步骤，所以这里单独提出。其重要性远超过欧拉方程（欧拉方程本书不讨论，因为分部积分法很自然的就会推出欧拉方程）

对于积分中含有变分的导数的，可以通过分部积分法，去掉导数。

即：

$$\begin{aligned} \int_a^b f(x) \delta u_x dx &= \int_a^b f(x) d\delta u \\ &= f(x) \delta u \Big|_a^b - \int_a^b \delta u \frac{df(x)}{dx} dx \end{aligned} \quad (\text{A.27})$$

(A.27) 式中，就通过分部积分法，把高阶 δu_x （即 $\delta \frac{du}{dx}$ 或 $\frac{d}{dx} \delta u$ ）转化为低阶 δu ，于是就可以求解了。

二、利用变分推导哈密顿方程

简单演示one control variable and one state variable 的模型的变分推导过程。

$$\max \int_0^{\infty} U(x) e^{-\rho t} dt \quad (1)$$

$$\text{subject to: } \dot{z} = g(x, z) \quad (2)$$

$$f(x, z) \geq 0, \quad z(0) \text{ 给定} \quad (4)$$

写出拉格朗日式子：

$$L = \int_0^{\infty} [U(x) e^{-\rho t} + \mu (g - \dot{z}) e^{-\rho t} + \lambda f e^{-\rho t}] dt \quad (5)$$

$$\text{令 } H = U(x) + \mu g + \lambda f \quad (6)$$

则(5)式成为：

$$L = \int_0^{\infty} (H e^{-\rho t} - \mu \dot{z} e^{-\rho t}) dt \quad (7)$$

以 z 为自变量对(7)式优化：

$$\delta L = \int_0^{\infty} (H_z e^{-\rho t} \delta z - \mu e^{-\rho t} \delta \dot{z}) dt = 0 \quad (8)$$

交换第2项中变分和微分的顺序并整理得：

$$\delta L = \int_0^{\infty} H_z e^{-\rho t} \delta z dt - \int_0^{\infty} \mu e^{-\rho t} \frac{d(\delta z)}{dt} dt = 0 \quad (9)$$

对第2项分部积分法降阶，得：

$$\delta L = \int_0^{\infty} H_z e^{-\rho t} \delta z dt - \left(\mu e^{-\rho t} \delta z \Big|_0^{\infty} - \int_0^{\infty} \frac{d(\mu e^{-\rho t})}{dt} \delta z dt \right) = 0 \quad (10)$$

计算并整理得：

$$\delta L = \int_0^{\infty} (H_z e^{-\rho t} + \dot{\mu} e^{-\rho t} - \rho \mu e^{-\rho t}) \delta z dt - \mu e^{-\rho t} \delta z \Big|_0^{\infty} = 0 \quad (11)$$

根据变分引理得：

$$H_z e^{-\rho t} + \dot{\mu} e^{-\rho t} - \rho \mu e^{-\rho t} = 0$$

$$\mu e^{-\rho t} \delta z \Big|_0^{\infty} = 0$$

由上面的第1个式子可以得到：

$$\dot{\mu} = \rho \mu - H_z \quad (12)$$

证毕。