数理经济学 (变分原理和应用)

王鸣晖

西南财经大学

数学学院

wangmh@swufe.edu.cn

2023 年 03 月 21 日

二阶条件

在之前的研究中，我们将目标泛函 $V[y]$ 视为扰动 ϵ 的函数 $V[\epsilon]$，进而得到极值的一阶必要条件，即：

$$\left.\frac{dV[\epsilon]}{d\epsilon}\right|_{\epsilon=0} = 0 \Rightarrow \text{E-L 方程及对应的横截条件.}$$

我们通过上述方法虽然找到了极值路径，但并不知道其为极大值还是极小值路径．故需要研究二阶条件，来判断其为极大值还是极小值路径，即判断条件

$$\left.\frac{d^2V[\epsilon]}{d\epsilon^2}\right|_{\epsilon=0}$$

的符号．

二阶条件

来自微积分的定理:

定理 5-1

设函数 $f(x)$ 在点 $x = x_0$ 二阶导存在, 且 $f'(x_0) = 0$, $f''(x_0) < 0$, 则 x_0 为 $f(x)$ 的极大值点. 若 $f''(x_0) > 0$, 则 x_0 为 $f(x)$ 的极小值点.

我们将依据定理 5-1, 来讨论变分问题的极值点.

二阶条件

由扰动假设可知

$$y(t) = y^*(t) + \epsilon p(t) \Rightarrow \frac{dy(t)}{d\epsilon} = p(t).$$

以及

$$y'(t) = y^{*'}(t) + \epsilon p'(t) \Rightarrow \frac{dy'(t)}{d\epsilon} = p'(t).$$

由此，可得:

$$\begin{aligned}\frac{d^2V}{d\epsilon^2} &= \frac{d}{d\epsilon}\left(\frac{dV}{d\epsilon}\right) = \frac{d}{d\epsilon}\int_0^T [L_y p(t) + L_{y'} p'(t)]\, dt \\ &= \int_0^T \left[p(t)\frac{d}{d\epsilon}L_y + p'(t)\frac{d}{d\epsilon}L_{y'}\right] dt. \quad \text{（交换求导和积分顺序）}\end{aligned}$$

二阶条件

又由于

$$\frac{d}{d\epsilon}L_y = L_{yy}\frac{dy}{d\epsilon} + L_{y'y}\frac{dy'}{d\epsilon} = L_{yy}p(t) + L_{y'y}p'(t).$$

以及:

$$\frac{d}{d\epsilon}L_{y'} = L_{yy'}p(t) + L_{y'y'}p'(t).$$

可得二阶条件如下:

$$\frac{d^2V}{d\epsilon^2} = \int_0^T \left[L_{yy}p^2(t) + 2L_{yy'}p(t)p'(t) + L_{y'y'}p'^2(t)\right] dt. \tag{C-5-1}$$

二阶条件

注记-5-1

二阶条件(C-5-1)中的被积函数可以看做是一个关于向量 $(p(t),p'(t))$, $\forall t\in[0,T]$ 的二次型，即:

$$L_{yy}p^2(t)+2L_{yy'}p(t)p'(t)+L_{y'y'}p'^2(t)=(p(t),p'(t))D(p(t),p'(t))^T,$$

其中

$$D=\begin{pmatrix} L_{yy} & L_{y'y} \\ L_{y'y} & L_{y'y'} \end{pmatrix}.$$

由此可知，当矩阵 D 正定时，条件(C-5-1)大于 0，原问题取得极小值. 反之，当矩阵 D 负定时，条件(C-5-1)小于 0，原问题取得极大值.

二阶条件

本质上，矩阵 D 的正定情况其实等价于函数 $L(t,y,y')$ 关于 (y,y') 的凹凸性，故有如下定理成立：

充分性条件 (固定端点问题)

对于基本的变分问题 (固定端点问题)，如果被积函数 $L(t,y,y')$ 关于向量 (y,y') 为凹函数，则 E-L 方程取得的极值路径 $y^*(t)$ 是极大值的充分性条件. 类似地，如果被积函数 $L(t,y,y')$ 关于向量 (y,y') 为凸函数，则 E-L 方程取得的极值路径 $y^*(t)$ 是目标泛函 $V[y]$ 的极小值的充分性条件.

注记-5-2

这里函数 $L(t,y,y')$ 是关于向量 (y,y') 的凹凸函数，而不是指关于 y 或者 y' 分别为凸函数或者凹函数.

二阶条件

只以凹函数为例证明这个定理，凸函数的情况类似.

证明

假设函数 $L(t,y,y')$ 为一个可微的凹函数，即函数 $L(t,y,y')$ 关于 (y,y') 为凹函数. 则根据凹函数的定义，对定义域内任意两个不同的点 $(t,y^*,y^{*'})$ 和 (t,y,y')，有下式成立：

$$\begin{aligned}
&L(t,y,y') - L(t,y^*,y^{*'}) \\
&\qquad \le L_y(t,y^*,y^{*'})(y-y^*) + L_{y'}(t,y^*,y^{*'})(y'-y^{*'}) \\
&\qquad = L_y(t,y^*,y^{*'})\epsilon p(t) + L_{y'}(t,y^*,y^{*'})\epsilon p'(t)
\end{aligned}$$

这里，$y^*(t)$ 表示最优路径，$y(t)$ 表示扰动路径（和 $y^*(t)$ 不同）.

二阶条件

证明（接上）

对上式两边关于时间 t 在区间 $[0,T]$ 上积分可得:

$$
\begin{aligned}
V[y]-V[y^*] &\le \epsilon\int_0^T \left[L_y(t,y^*,{y^*}^{'})p(t)+L_{y'}(t,y^*,{y^*}^{'})p'(t)\right]dt \\
&= \epsilon\int_0^T p(t)\left[L_y(t,y^*,{y^*}^{'})-\frac{d}{dt}L_{y'}(t,y^*,{y^*}^{'})\right]dt = 0.
\end{aligned}
\tag{E-5-1}
$$

上式中，第一个不等式来自于对于变分问题的定义. 第二个等式来自分部积分:

$$
\begin{aligned}
&\int_0^T L_{y'}(t,y^*,{y^*}^{'})p'(t)dt \\
&\quad = p(t)L_{y'}(t,y^*,{y^*}^{'})\Big|_0^T - \int_0^T \frac{d}{dt}L_{y'}(t,y^*,{y^*}^{'})p(t)dt.
\end{aligned}
\tag{E-5-2}
$$

二阶条件

证明（接上）

结合 E-L 方程，可以得到上述结果. 由此可以得到:

$$V[y] \le V[y^*].$$

其中，y 是任意路径. 由此可知，如果函数 $L(t,y,y^*)$ 是关于 (y,y^*) 的凹函数，则由 E-L 方程得到的路径 y^* 是极大路径.

注记-5-3

如果函数 $L(t,y,y^*)$ 是关于 (y,y^*) 的严格凹函数 $\Rightarrow V[y] < V[y^*]$. 此时，可以得到 y^* 是 $V[y]$ 的唯一极大值. 类似地，如果函数 $L(t,y,y^*)$ 是一个严格凸函数，则由 E-L 方程得到的 y^* 是 $V[y]$ 的唯一极小值.

带有垂直终止线的二阶条件

上述充分性条件也可以推广到带有垂直终止线的最优变分问题. 由于在上节课推导横截条件时, 我们知道在研究可变终值问题时, $p(T)=0$ 条件不再成立. 故在上述条件证明过程中式(E-5-1)变为:

$$\epsilon\int_0^T p(t)\left[L_y(t,y^*,y^{*'})-\frac{d}{dt}L_{y'}(t,y^*,y^{*'})\right]dt+\left[L_{y'}(y^*-y^{*'})\right]\Big|_{t=T}=0.$$

上式中第三项来自于分部积分(E-5-2)中的第一项, 即:

$$p(t)L_{y'}(t,y^*,y^{*'})\Big|_0^T=\left[p(t)L_{y'}(t,y^*,y^{*'})\right]\Big|_{t=T}\neq 0.$$

此时, 可以得到:

$$V[y]\le V[y^*]+\left[L_{y'}(y^*-y^{*'})\right]\Big|_{t=T}.$$

带有垂直终止线的二阶条件

此时, 可以得到如下三种情况:

- 如果 $[L_{y'}(y^* - y^{*'})]\Big|_{t=T} = 0$, 那么原来的结论依然是成立的.
- 如果 $[L_{y'}(y^* - y^{*'})]\Big|_{t=T} < 0$, 此时 $V[y^*]$ 肯定也是最大值.
- 如果 $[L_{y'}(y^* - y^{*'})]\Big|_{t=T} > 0$, 此时 $V[y^*]$ 不一定为最大值(充分性条件不一定成立).

回忆带有垂直终止线的最优变分问题, 可以知道其对应的横截条件为:

$$L_{y'}\Big|_{t=T} = 0.$$

此时, 可以得到 $[L_{y'}(y^* - y^{*'})]\Big|_{t=T} = 0$ 成立. 由此可知, 带有固定时间水平的最优变分问题的充分性条件不变.

Legendre 条件

提出原因

函数 $L(t, y, y')$ 关于 (y, y') 为凹（凸）函数的要求是非常困难的，在大多数情况下，函数 $L(t, y, y')$ 都不为凹（凸）函数. 故需要找到一些更弱的条件.

Legendre 条件

设 $y^{*'}$ 最大变分问题的最优解，则成立:

$$L_{y'y'}(t, y^*, y^{*'}) \leq 0. \tag{L}$$

类似地，如果是最小变分问题，则条件(L)的符号反过来.

理论依据

考虑二阶条件(C-5-1)，设 $v = L_{yy'}$ 和 $u = p^2(t)$，则有：

$$dv = \frac{dL_{yy'}}{dt}dt, \quad du = 2p(t)p'(t)dt.$$

此时，二阶条件(C-5-1)的中间项可以改写成：

$$\begin{aligned}\int_0^T vdu &= uv\Big|_0^T - \int_0^T udv = L_{yy'}p^2(t)\Big|_0^T - \int_0^T p^2(t)\frac{dL_{yy'}}{dt}dt \\ &= 0 - \int_0^T p^2(t)\frac{dL_{yy'}}{dt}dt. \quad (\text{此时假设 } p(0) = p(T) = 0.)\end{aligned} \tag{E-5-3}$$

将上式(E-5-3)中得到的结果代入回(C-5-1)中，可以得到：

$$\frac{d^2V}{d\epsilon^2} = \int_0^T \left[\left(L_{yy} - \frac{dL_{yy'}}{dt}\right)p^2(t) + L_{y'y'}p'^2(t)\right]dt. \tag{E-5-4}$$

为了得到 Legendre 条件, 还需要如下引理:

引理-1

设 $P(t)$ 和 $Q(t)$ 是区间 $[t_0, t_1]$ 上给定的两个连续函数, 又设如下二次泛函

$$\int_{t_0}^{t_1} \{P(t)[h'(t)]^2 + Q(t)[h(t)]^2\} dt, \qquad (*)$$

对于在区间 $[t_0, t_1]$ 上满足 $h(t_0) = h(t_1) = 0$ 的所有连续函数 $h(t)$ 都有定义. 则对于所有这样的函数 $h(t)$, 式$(*)$非正的必要条件为:

$$P(t) \leq 0, \quad t \in [t_0, t_1].$$

Legendre 条件 (接上)

理论依据

取 $P(t) = L_{y'y'}$ 且 $Q(t) = L_{yy} - \frac{dL_{yy'}}{dt}$, 由引理-1 可得:

$$L_{y'y'} \leq 0.$$

由此得到了 Legendre 条件.

例-5-1

给定泛函

$$V[y]=\int_a^z y\sqrt{1+y'^2}dt$$

及其边界条件 $y(a)=A$ 且 $y(z)=Z$, 求其极值路径并给出对应的 Legendre 条件.

Weierstrass 条件

提出原因

对 Legendre 条件进行推广.

Weierstrass 条件

若 $y^*(\cdot)$ 是一个连续可微的极小路径, 则对于 $\forall x \in R$ 和 $\forall t \in [t_0, t_1]$ 都成立:

$$L(t, y^*(t), x) \geq L(t, y^*(t), y^{*'}(t)) + \frac{\partial L}{\partial x}(t, y^*(t), y^{*'}(t))(x - y^{*'}(t)). \quad \text{(W)}$$

注记-5-4

条件(W)主要用于判别函数的凸性. 所以如果函数 L 是一个凸函数, 此时条件显然成立. 此外, Weierstrass 条件可以看做判断函数凸性的一阶条件的直接应用, 故 Weierstrass 条件可以看做是 Legendre 条件的推广.

在之前的讨论中，我们主要研究了如下两个导数:

$$\left.\frac{dV[\epsilon]}{d\epsilon}\right|_{\epsilon=0} \quad (\text{一阶条件}) \quad \text{和} \quad \left.\frac{d^2V[\epsilon]}{d\epsilon^2}\right|_{\epsilon=0} \quad (\text{二阶条件}).$$

而上述两个导数本质上是研究对极值路径 y^* 的扰动所得到的微分条件.
而上述两个导数完全可以由目标泛函 $V[y]$ 的扰动来描述，即变分:

$$\triangle V = V[y] - V[y^*] = \int_{t_0}^{t_1} L(s, y(s), y'(s))ds - \int_{t_0}^{t_1} L(s, y^*(s), y'^*(s))ds.$$

为了更好的理解变分 $\triangle V$, 我们将函数 $L(t,y,y')$ 在极值点 $(t,y^*,y^{'*})$ 进行 Taylor 展开.

Taylor 展开

$$\begin{aligned} L(t,y,y') = & L(t,y^*,y^{*'}) + [L_t(t-t) + L_y(y-y^*) + L_{y'}(y'-y^{*'})] \\ & + \frac{1}{2}[L_{tt}(t-t)^2 + L_{yy}(y-y^*)^2 + L_{y'y'}(y'-y^{*'})^2 \\ & + 2L_{ty}(t-t)(y-y^*) + 2L_{ty'}(t-t)(y'-y^{*'}) \\ & + 2L_{yy'}(y-y^*)(y'-y^{*'})] + L_n. \end{aligned} \tag{T-1}$$

其中 L_n 是余项.

此时，取 $y-y^*=\epsilon p$ 和 $y-{y^*}'=\epsilon p'$，并将它们代入(T-1)，可以得到：

$$\begin{aligned} L(t,y,y') = & L(t,y^*,{y^*}') + L_y\epsilon p + L_{y'}\epsilon p' \\ & + \frac{1}{2}[L_{yy}(\epsilon p)^2 + L_{y'y'}(\epsilon p')^2 + 2L_{yy'}\epsilon^2 pp'] + L_n. \end{aligned} \tag{T-2}$$

由此可得变分 $\triangle V$ 的展开

$\triangle V$ 的展开

$$\begin{aligned} \triangle V = & \epsilon\int_{t_0}^{t_1}(L_y p + L_{y'}p')dt \\ & + \frac{\epsilon^2}{2}\int_{t_0}^{t_1}[L_{yy}p^2 + L_{y'y'}(p')^2 + 2L_{yy'}pp'] + L_n^*. \end{aligned} \tag{T-3}$$

其中 L_n^* 为高阶余项.

一阶变分和二阶变分

一阶变分

$$\delta V = \int_{t_0}^{t_1} (L_y p + L_{y'} p') dt = \frac{dV[\epsilon]}{d\epsilon}.$$

取上述一阶变分 $\delta V = 0$ 可以直接得到 E-L 方程.

二阶变分

$$\delta^2 V = \int_{t_0}^{t_1} [L_{yy} p^2 + L_{y'y'} (p')^2 + 2L_{yy'} pp'] dt = \frac{d^2 V[\epsilon]}{d\epsilon^2}.$$

研究上述二阶变分 $\delta^2 V$ 的符号，可以得到二阶充分性条件.