Supplementary Notes on Chapter 2 of D. Romer's Advanced Macroeconomics Textbook (4th Edition)

Changsheng Xu & Ming Yi

School of Economics, Huazhong University of Science and Technology

This version: March 3, 2025

Calculus of variations (变分法)

- A field of mathematical analysis that deals with maximizing or minimizing **functionals**, which are mappings from a set of functions to the real numbers.
- Functionals are often expressed as definite integrals involving functions and their derivatives. (e.g., the famous shortest (in time) path problem)
- The **Euler-Lagrange equation** provides a **necessary condition** for finding extrema.

Euler-Lagrange equation

Intuition: Finding the extrema of functionals is similar to finding the maxima and minima of functions. This tool provides a link between them to solve the problem. Consider the functional

$$J[y] = \int_{x_1}^{x_2} L\left(x, y(x), y'(x)\right) \mathrm{d}x, \tag{1}$$

where

- x_1, x_2 are constants.
- $y(x)$ is twice continuously differentiable.
- $y'(x) = \frac{\mathrm{d}y}{\mathrm{d}x}$.
- $L(x, y(x), y'(x))$ is twice continuously differentiable with respect to all arguments x, y, and y'.

Euler-Lagrange equation (Continued)

If $J[y]$ attains a local minimum at f, and $\eta(x)$ is an arbitrary function that has at least one derivative and vanishes at the endpoints x_1 and x_2, then for any number $\varepsilon \to 0$, we must have

$$J[f] \leq J[f + \varepsilon\eta]. \tag{2}$$

Term $\varepsilon\eta$ is called the **variation** of the function f. Now define

$$\Phi(\varepsilon) = J[f + \varepsilon\eta]. \tag{3}$$

Since $J[y]$ has a local minimum at $y = f$, it must be the case that $\Phi(\varepsilon)$ has a minimum at $\varepsilon = 0$ and thus

$$\Phi'(0) = \left.\frac{\mathrm{d}\Phi}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} = \int_{x_1}^{x_2} \left.\frac{\mathrm{d}L}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathrm{d}x = 0. \tag{4}$$

Euler-Lagrange equation (Continued)

Now taking total derivative of $L\left[x, f+\varepsilon\eta, (f+\varepsilon\eta)'\right]$, we have:

$$\frac{\mathrm{d}L}{\mathrm{d}\varepsilon} = \frac{\partial L}{\partial y}\eta + \frac{\partial L}{\partial y'}\eta'. \tag{5}$$

Inserting (5) into (4) gives us

$$\begin{aligned} 0 = \int_{x_1}^{x_2} \frac{\mathrm{d}L}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} &= \int_{x_1}^{x_2} \left(\frac{\partial L}{\partial f}\eta + \frac{\partial L}{\partial f'}\eta'\right)\mathrm{d}x \\ &= \int_{x_1}^{x_2} \left(\frac{\partial L}{\partial f}\eta - \eta\frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x}\right)\mathrm{d}x + \frac{\partial L}{\partial f'}\eta\bigg|_{x_1}^{x_2} \\ &= \int_{x_1}^{x_2} \eta\left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x}\right)\mathrm{d}x, \end{aligned}$$

where the last lines uses integration by parts and the fact that η vanishes at x_1 and x_2.

Euler-Lagrange equation (Continued)

Now given

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} \right) \mathrm{d}x = 0, \tag{6}$$

the **fundamental lemma of calculus of variations** makes sure that

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} = 0 \tag{7}$$

must hold!

Euler-Lagrange equation (Continued)

Now given

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} \right) \mathrm{d}x = 0, \tag{6}$$

the **fundamental lemma of calculus of variations** makes sure that

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} = 0 \tag{7}$$

must hold!

- However, it is possible to attain (7) based on (6) without applying the lemma!

Euler-Lagrange equation (Continued)

Now given

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} \right) \mathrm{d}x = 0, \tag{6}$$

the **fundamental lemma of calculus of variations** makes sure that

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} = 0 \tag{7}$$

must hold!

- However, it is possible to attain (7) based on (6) without applying the lemma!
- A special form of η?

Euler-Lagrange equation (Continued)

Now given

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} \right) \mathrm{d}x = 0, \tag{6}$$

the **fundamental lemma of calculus of variations** makes sure that

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} = 0 \tag{7}$$

must hold!

- However, it is possible to attain (7) based on (6) without applying the lemma!
- A special form of η?
- How about $\eta(x)$ equals $-(x - x_1)(x - x_2) \left[\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}x} \right]$ for $x \in [x_1, x_2]$ and 0 for $x \notin [x_1, x_2]$?

Euler-Lagrange equation (Continued)

- How does (7) degenerate if y' is not an argument of L?
- Homework: based on equation (2.16) in your textbook, try to derive (2.17).

A simple exercise

Consider the following problem:

$$\max_{[c(t),a(t)]_{t=0}^{1}} \int_0^1 e^{-\rho t} u(c(t)) \mathrm{d}t, \tag{8}$$

$$\text{subject to} \quad \dot{a}(t) = ra(t) + \omega - c(t), \quad a(0) = a_0, \; a(1) = 0. \tag{9}$$

where r and ω are exogenously defined constants.

- Deduce the Euler-Lagrange equation for the problem above.
- Rearrange your result above to give the Euler equation usually used in your textbooks, $\frac{u''(c(t))\dot{c}(t)}{u'(c(t))} = \rho - r$, namely, along the household's optimal path, the growth rate of its marginal utility of consumption should be equal to the gap between the discount rate ρ and interest rate r.

Another exercise: the Brachistochrone Curve

The famous Brachistochrone Problem: given two points (x_0, y_0) and (x_1, y_1) with $x_0 < x_1$ and $y_0 > y_1$ in a two-dimensional world with gravitational acceleration g and without frictions . Find a *smooth path* that connects these points and makes the **travel time** from (x_0, y_0) to (x_1, y_1) the **shortest**.
The problem above is transformed into a mathematical one:

$$\min_{y(x)} J(y) = \int_{x_0}^{x_1} \frac{\sqrt{1 + [y'(x)]^2}}{\sqrt{2g[y_0 - y(x)]}} \mathrm{d}x \tag{10}$$

$$\text{subject to} \quad y(x_0) = y_0, \; y(x_1) = y_1. \tag{11}$$

Check out
`http://mathworld.wolfram.com/BrachistochroneProblem.html` for a detailed introduction to this problem!

Pontryagin's Maximum Principle

- Mainly developed by Pontryagin and his group.
- A Hamiltonian method that generalizes the Euler-Lagrange equation above.

Variational Arguments

Problem A1:

$$\max_{x(t),y(t),x_1} \quad W(x(t), y(t)) = \int_0^{t_1} f(t, x(t), y(t))\,\mathrm{d}t\ , \tag{12}$$

$$\text{subject to} \quad \dot{x}(t) = g(t, x(t), y(t)), \tag{13}$$

$$x(0) = x_0,\ \text{and}\ x(t_1) = x_1. \tag{14}$$

Other settings:

- Continuous differentiability of functions are assumed again.
- the value of the state variable at the terminal of the horizon, $x(t_1)$, is flexible in this problem.
- We ignore here the trivial requirements stating that the values of $x(t)$ and $y(t)$ should always be in some sets $\mathcal{X}, \mathcal{Y} \in \mathbb{R}$ for all t.
- We suppose there exists an interior solution $(\hat{x}(t), \hat{y}(t))$, and focus on the necessary conditions for a solution.

Variational Arguments (Continued)

Take a *variation* of function $\hat{y}(t)$:

$$y(t,\epsilon) = \hat{y}(t) + \epsilon\,\eta(t). \tag{15}$$

Note that given $y(t,\epsilon)$, $x(t)$ is now dependent on ϵ according to evolutionary equation (13), so the resulting $x(t,\epsilon)$ is defined by:

$$\dot{x}(t,\epsilon) = g(t, x(t,\epsilon), y(t,\epsilon)) \text{ for all } t \in [0, t_1], \text{ with } x(0,\epsilon) = x_0. \tag{16}$$

Define:

$$\begin{aligned} \mathcal{W}(\epsilon) &= W(x(t,\epsilon), y(t,\epsilon)) \\ &= \int_0^{t_1} f(t, x(t,\epsilon), y(t,\epsilon))\,\mathrm{d}t. \end{aligned} \tag{17}$$

Variational Arguments (Continued)

Since $\hat{x}(t)$, $\hat{y}(t)$ solve the optimal control problem, we must have:

$$\mathcal{W}(\epsilon) \leq \mathcal{W}(0) \quad \text{for all small engough} \quad \epsilon \to 0. \tag{18}$$

Now recall that, $g(t, x(t, \epsilon), y(t, \epsilon)) - \dot{x}(t, \epsilon) = 0$ holds for all t. Then for any function $\lambda : [0, t_1] \to \mathbb{R}$, we must have:

$$\int_0^{t_1} \lambda(t)[g(t, x(t, \epsilon), y(t, \epsilon)) - \dot{x}(t, \epsilon)] \, \mathrm{d}t = 0. \tag{19}$$

Function $\lambda(t)$ is called the **costate** variable, with an interpretation similar to the Lagrange multipliers in standard (static) optimization problems.

Variational Arguments (Continued)

Combining (17) and (19) lets us redefine $\mathcal{W}(\epsilon)$:

$$\begin{aligned}\mathcal{W}(\epsilon) &= \int_0^{t_1} f(t, x(t,\epsilon), y(t,\epsilon))\,\mathrm{d}t + 0 \\ &= \int_0^{t_1} \{f(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t)[g(t, x(t,\epsilon), y(t,\epsilon)) - \dot{x}(t,\epsilon)]\}\mathrm{d}t \\ &= \int_0^{t_1} \{f(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t) g(t, x(t,\epsilon), y(t,\epsilon)) + \dot{\lambda}(t) x(t,\epsilon)\}\mathrm{d}t \\ &\quad -\lambda(t_1) x(t_1,\epsilon) + \lambda(0) x_0. \end{aligned} \tag{20}$$

The last equality above uses integration by parts.

Variational Arguments (Continued)

Applying Leibniz's Rule to (20) yields:

$$\begin{aligned}\mathcal{W}'(\epsilon) &= \int_0^{t_1}\left[f_x(t,x(t,\epsilon),y(t,\epsilon))+\lambda(t)g_x(t,x(t,\epsilon),y(t,\epsilon))+\dot{\lambda}(t)\right]x_\epsilon(t,\epsilon)\mathrm{d}t \\ &\quad+\int_0^{t_1}\left[f_y(t,x(t,\epsilon),y(t,\epsilon))+\lambda(t)g_y(t,x(t,\epsilon),y(t,\epsilon))\right]\eta(t)\mathrm{d}t \\ &\quad-\lambda(t_1)x_\epsilon(t_1,\epsilon) \end{aligned} \tag{21}$$

Recall that condition (18) can be rewritten as $\mathcal{W}'(0)=0$. We thus have:

$$\begin{aligned} 0 &= \int_0^{t_1}\left[f_x(t,\hat{x}(t),\hat{y}(t))+\lambda(t)g_x(t,\hat{x}(t),\hat{y}(t))+\dot{\lambda}(t)\right]x_\epsilon(t,0)\mathrm{d}t \\ &\quad+\int_0^{t_1}\left[f_y(t,\hat{x}(t),\hat{y}(t))+\lambda(t)g_y(t,\hat{x}(t),\hat{y}(t))\right]\eta(t)\mathrm{d}t \\ &\quad-\lambda(t_1)x_\epsilon(t_1,0) \end{aligned} \tag{22}$$

Variational Arguments (Continued)

Treating $\lambda(t_1)x_\epsilon(t_1, 0)$

- (22) must hold for any continuously differentiable $\lambda(t)$.
- we simply focus on a class of costate variables satisfying

$$\lambda(t_1) = 0 \tag{23}$$

- As a result, $\lambda(t_1)x_\epsilon(t_1, 0) = 0$.

Variational Arguments (Continued)

Treating $\int_0^{t_1}\left[f_x(t,\hat{x}(t),\hat{y}(t))+\lambda(t)g_x(t,\hat{x}(t),\hat{y}(t))+\dot{\lambda}(t)\right]x_\epsilon(t,0)\mathrm{d}t$

- Besides $\lambda(t_1)=0$ as illustrated above, can we add more requirements on the costate variable?
- Again, since (22) must hold for any continuously differentiable $\lambda(t)$, why not focus on the following $\lambda(t)$:

$$\begin{aligned}\dot{\lambda}(t) &= -\left[f_x(t,\hat{x}(t),\hat{y}(t))+\lambda(t)g_x(t,\hat{x}(t),\hat{y}(t))\right], \\ \lambda(t_1) &= 0\end{aligned} \tag{24}$$

Variational Arguments (Continued)

Treating $\int_0^{t_1} \left[f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t))\right] \eta(t) \mathrm{d}t$

- Given the costate $\lambda(t)$ defined above, equality $\int_0^{t_1} \left[f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t))\right] \eta(t) \mathrm{d}t = 0$ must hold for arbitrary $\eta(t)$.
- Applying the fundamental lemma of calculus of variations yields:

$$f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t)) = 0 \quad \text{for all} \quad t \in [0, t_1] \tag{25}$$

Variational Arguments (Continued)

Theorem

Suppose Problem A1 has an interior continuous solution $(\hat{x}(t), \hat{y}(t))$, *then there exists a continuously differentiable costate* $\lambda(t)$ *defined on* $[0, t_1]$, *such that* (13), (23), (24), *and* (25) *hold.*

Problem A2

$$
\begin{aligned}
\max_{x(t),y(t)} \; W(x(t), y(t)) = \int_0^{t_1} f(t, x(t), y(t))\,\mathrm{d}t\;, \\
\text{subject to} \quad \dot{x}(t) = g(t, x(t), y(t)), \\
x(0) = x_0, \;\text{and}\; x(t_1) = x_1.
\end{aligned}
\tag{26}
$$

Theorem

Suppose Problem A2 has an interior continuous solution $(\hat{x}(t), \hat{y}(t))$*, then there exists a continuously differentiable costate* $\lambda(t)$ *defined on* $[0, t_1]$*, such that* (26)*,* (24)*, and* (25) *hold.*

Can you figure out how and why Theorem 2 differs from Theorem 1?

Revisiting the simple exercise

Consider the following problem:

$$\max_{[c(t),a(t)]_{t=0}^{1}} \int_0^1 e^{-\rho t} u(c(t)) \mathrm{d}t, \tag{27}$$

$$\text{subject to} \quad \dot{a}(t) = ra(t) + \omega - c(t), \quad a(0) = a_0, \ a(1) = 0. \tag{28}$$

where r and ω are exogenously defined constants.

- Use the Pontryagin's Maximum Principle (Theorem 4 above) to get the same results (Euler equation) as before.
- Given $u(c) = \log(c)$, can you ***solve*** the problem above? What if $u(c) = \left[\theta - e^{-\beta c(t)}\right]$?

Problem A3

$$\max_{x(t),y(t)} \; W(x(t),y(t)) = \int_0^{t_1} f(t,x(t),y(t))\,\mathrm{d}t\;,$$
$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{29}$$
$$x(0) = x_0, \text{ and } x(t_1) \geq x_1.$$

Theorem

Suppose Problem A3 has an interior continuous solution $(\hat{x}(t),\hat{y}(t))$, *then there exists a continuously differentiable costate* $\lambda(t)$ *defined on* $[0,t_1]$, *such that* (29), (24), (25), *and* $\lambda(t_1)[x(t_1)-x_1] = 0$ *hold.*

Can you figure out how and why Theorem 3 differs from Theorem 1?

Pontryagin's Maximum Principle

Revisit Problem A1. Define the **Hamiltonian**:

$$H(t, x(t), y(t), \lambda(t)) \equiv f(t, x(t), y(t)) + \lambda(t) g(t, x(t), y(t)). \tag{30}$$

Theorem

Suppose Problem A1 has an interior continuous solution $(\hat{x}(t), \hat{y}(t))$, then there exists a continuously differentiable function $\lambda(t)$ such that the following necessary conditions hold:

$$\dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{y}_(t), \lambda(t)) \quad \text{for all} \quad t \in [0, t_1], \tag{31}$$

$$\dot{x}(t) = H_\lambda(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \quad \text{for all} \quad t \in [0, t_1], \tag{32}$$

$$H(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \geq H(t, \hat{x}(t), y, \lambda(t)) \text{ for all feasible } y, t, \tag{33}$$

$$x(0) = x_0, \quad \lambda(t_1) = 0. \tag{34}$$

Question: relationship between (25) and (33)?

Pontryagin's Maximum Principle (Continued)

- The results of equations (31)-(34) are straightforward to understand given the variational arguments illustrated earlier.
- Can you figure out the Pontryagin's Maximum Principle for Problems A2 and A3?
- (important) In all problems so far, we have both x and y one-dimensional variables. The Pontryagin's maximum Principle also applies to scenarios where $\mathbf{x}$ and $\mathbf{y}$ are actually vectors of state and control variables, respectively. In these cases, it may be necessary to introduce more than one costate variables, e.g., $\lambda_1(t), \cdots, \lambda_k(t)$ into the Hamiltonian. For instance, if the evolutionary equation becomes $\dot{x} = g_1(\cdot)$ and $\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = g_2(\cdot)$. We actually have two state variables here: $\mathbf{x} = (x, \dot{x})$.
- So far only necessary conditions discussed. Sufficient conditions? It suffices to have some degree of concavity of $H(t, x, y, \lambda)$ in (x, y).

Infinite Horizon Problem

Consider Problem A4:

$$\max_{x(t),y(t)} \quad W(x(t), y(t)) = \int_0^{\infty} f(t, x(t), y(t)) \, \mathrm{d}t \, , \tag{35}$$

$$\text{subject to} \quad \dot{x}(t) = g(t, x(t), y(t)), \tag{36}$$

$$x(0) = x_0, \text{ and } \lim_{t\to\infty} b(t)x(t) \geq x_1.$$

Notice: When the integral is defined on an unbounded interval, we need more assumptions (related to the dominated convergence theorem) on the integrability of some functions for the Leibniz's Rule to apply. However, these details are trivial and almost always satisfied in economic applications.

Infinite Horizon Problem (continued)

Theorem

Suppose Problem A4 has an interior continuous solution $(\hat{x}(t), \hat{y}(t))$*, then there exists a continuously differentiable function* $\lambda(t)$ *such that the following necessary conditions hold:*

$$\dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{y}_(t), \lambda(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{37}$$

$$\dot{x}(t) = H_\lambda(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{38}$$

$$H(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \geq H(t, \hat{x}(t), y, \lambda(t)) \text{ for all feasible } y, t, \tag{39}$$

$$x(0) = x_0, \quad \lim_{t\to\infty} b(t)x(t) \geq x_1. \tag{40}$$

Current-Value Hamiltonian

Consider Problem A5:

$$\max_{x(t),y(t)} \; W(x(t),y(t)) = \int_0^\infty e^{-\rho t} f(x(t),y(t))\,\mathrm{d}t \,, \tag{41}$$

$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{42}$$

$$x(0) = x_0, \text{ and } \lim_{t\to\infty} b(t)x(t) \geq x_1.$$

The Hamiltonian is:

$$H(t,x(t),y(t),\lambda(t)) = e^{-\rho t}\left[f(x(t),y(t)) + e^{\rho t}\lambda(t)g(t,x(t),y(t))\right] \tag{43}$$

Define function $\mu(t) \equiv e^{\rho t}\lambda(t)$. The **current-value Hamiltonian** is thus defined as

$$\hat{H}(t,x(t),y(t),\mu(t)) \equiv f(x(t),y(t)) + \mu(t)g(t,x(t),y(t)). \tag{44}$$

Current-Value Hamiltonian(continued)

Theorem

Suppose Problem A5 has an interior continuous solution $(\hat{x}(t), \hat{y}(t))$, *then there exists a continuously differentiable function* $\mu(t)$ *such that the following conditions hold:*

$$\hat{H}_y(t, \hat{x}(t), \hat{y}(t), \mu(t)) = 0 \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{45}$$

$$\rho\mu(t) - \dot{\mu}(t) = \hat{H}_x(t, \hat{x}(t), \hat{y}(t), \mu(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{46}$$

$$\dot{x}(t) = \hat{H}_\mu(t, \hat{x}(t), \hat{y}(t), \mu(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{47}$$

$$\lim_{t \to \infty} [e^{-\rho t}\mu(t)\hat{x}(t)] = 0 \tag{48}$$

$$x(0) = x_0, \quad \lim_{t \to \infty} b(t)\hat{x}(t) \geq x_1. \tag{49}$$

(48) is a simplified version of the **Transversality Condition** for optimization problems with an infinite horizon.

Homework

Try your best to understand:

- Example 7.1 (page 233), Example 7.3 (page 252), and Section 7.8: The q-theory (page 269) in Acemoglu (2008).
- Or examples between pages 641-643, and Section 20.5 (page 649) in Chiang and Wainwright (2005). Note that the notations used in these books are different.

References

[1] Acemoglu, D. (2008). *Introduction to modern economic growth*, Princeton University Press.

[2] Chiang, A. and Wainwright, K. (2005). *Fundamental Methods of Mathematical Economics*, McGraw-Hill higher education, McGraw-Hill.