Dynamic Programming & Optimal Control
Advanced Macroeconomics
Doctoral Program in Economics, HUST

Ming Yi (yiming@hust.edu.cn)

School of Economics, Huazhong University of Science and Technology

This version: October 30, 2023

Course information:

http://www.yiming.website/teaching/2023fdmacro/

This lecture note is based mainly on selected materials in Chapter 6 and Chapter 7 of Acemoglu (2008).

(a) Richard E. Bellman (1920-1984) (b) Lev S. Pontryagin (1908-1988)

Figure 1: Two pioneers.

Canonical Discrete-Time Infinite-Horizon Optimization Problem

Canonical form of the problem:

$$\sup_{\{x(t),y(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \tilde{U}(t, x(t), y(t)) \tag{1}$$

$$\text{subject to} \quad y(t) \in \tilde{G}(t, x(t)) \quad \text{for all } t \geq 0, \tag{2}$$

$$x(t+1) = \tilde{f}(t, x(t), y(t)) \quad \text{for all } t \geq 0, \tag{3}$$

$$x(0) \text{ given.} \tag{4}$$

- "sup" interchangeable with "max" within the note. $\beta \in [0, 1)$.
- $x(t) \in X \subset \mathbb{R}^{K_x}$: *state variables* (state vector), $y(t) \in Y \subset \mathbb{R}^{K_y}$: *control variables* (control vector). $K_x, K_y \geq 1$.
- instantaneous payoff function $\tilde{U} : \mathbb{Z}_+ \times X \times Y \to \mathbb{R}$. Objective function: $\sum_{t=0}^{\infty} \beta^t \tilde{U}(t, x(t), y(t))$. Correspondence $\tilde{G} : \mathbb{Z}_+ \times x \rightrightarrows Y$.

Problem A1

The canonical form is rewritten as Problem A1:

$$V^*(0, x(0)) = \sup_{\{x(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(t, x(t), x(t+1)) \quad (5)$$

subject to

$$x(t+1) \in G(t, x(t)) \quad \text{for all} \quad t \geq 0, \quad (6)$$

$$x(0) \quad \text{given.} \quad (7)$$

- Problem A1 is identical to the canonical problem above.
- New expression, why bother?
- $V^*(0, x(0))$ obtained upon *optimal plan* $\{x^*(t+1)\}_{t=0}^{\infty} \in X^{\infty}$.
- What if symbol ∞ is replaced by some $T \in \mathbb{Z}_+$?
- Based on $\tilde{G}, \tilde{U}, \tilde{f}$, the definitions of G, U, and V^* are *trivial*.

Problem A1 (Continued)

Problem A1:

$$V^*(0, x(0)) = \sup_{\{x(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(t, x(t), x(t+1))$$

subject to

$$x(t+1) \in G(t, x(t)) \quad \text{for all} \quad t \geq 0,$$
$$x(0) \quad \text{given}.$$

- Try to define $G(t, x(t))$.

Problem A1 (Continued)

Problem A1:

$$V^*(0, x(0)) = \sup_{\{x(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(t, x(t), x(t+1))$$

$$\text{subject to} \quad x(t+1) \in G(t, x(t)) \quad \text{for all} \quad t \geq 0,$$
$$x(0) \quad \text{given.}$$

- Try to define $G(t, x(t))$.
-

$$G(t, x(t)) = \{\tilde{f}(t, x(t), y(t)) \in X \,|\, y_t \in \tilde{G}(t, x(t))\}$$

Problem A2

In this note, we focus only on Problem A2:

$$V^*(x(0)) = \sup_{\{x(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1)) \tag{8}$$

subject to

$$x(t+1) \in G(x(t)) \quad \text{for all} \quad t \geq 0, \tag{9}$$

$$x(0) \quad \text{given.} \tag{10}$$

- Problem A2 is a *stationary form* of Problem A1: U and G do not explicitly depend on time.
- Stationary dynamic programming.
- Applicable to most economic applications.

Problem A3

Let us consider Problem A3:

$$V(x) = \sup_{y \in G(x)} \{ U(x, y) + \beta V(y) \}, \quad \text{for all } \; x \in X. \tag{11}$$

- The recursively defined $V(x)$ is called **Bellman Equation**.
- The previous problem of finding a *sequence* $\{x^*(t+1)\}_{t=0}^{\infty}$ is replaced by the problem of finding a *function* $V(x)$.
- $V(\cdot)$ is called *value function*.
- Define *policy function* $\pi(\cdot)$ by $y^* = \pi(x)$.
- So $V(x) = U(x, \pi(x)) + \beta V(\pi(x))$.
- Once the value function is known, it is straightforward to induce the policy function.

An Example

The problem is given in the canonical form:

$$\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c(t)),$$

$$\text{subject to} \qquad k(t+1) = f(k(t)) + (1-\delta)k(t) - c(t),$$

where $k(t) \geq 0$, $c(t) \geq 0$, and $k(0) > 0$ is given. $u : \mathbb{R}_+ \to \mathbb{R}$.

Try to transform the canonical problem above to Problem A2 and A3.

V^* satisfies Bellman Equation

Recalling the relationship between $V^*(x(0))$ and $\{x^*(t+1)\}_{t=1}^{\infty}$ in Problem A2:

$$
\begin{aligned}
V^*(x(0)) &= \sum_{t=0}^{\infty} \beta^t U(x^*(t), x^*(t+1)) && (12)\\
&= U(x(0), x^*(1)) + \beta \sum_{s=0}^{\infty} \beta^s U(x^*(s+1), x^*(s+2)) && (13)\\
&= U(x(0), x^*(1)) + \beta V^*(x^*(1)) && (14)\\
&= \sup_{y \in G(x(0))} \{U(x(0), y) + \beta V^*(y)\}, \quad \forall x(0) \in X. && (15)
\end{aligned}
$$

- Equation (14): An optimal plan from $t = 0$ must also be an optimal plan from $t = 1$.
- Equation (15): An optimal plan solves the optimization problem.
- $x^*(t+1) = \pi(x^*(t))$ holds for all t.

Is V satisfying Bellman Equation also V^*?

Important assumtions:

1. X is compact in $\mathbb{R}^K$ (closed and bounded), $U(\cdot)$ is continuous.
2. $U(\cdot)$ is concave.
3. $U(\cdot, y)$ is strictly increasing in its first K arguments.
4. U is continuously differentiable in the interior of its domain X_G.

Their functions:

Is V satisfying Bellman Equation also V^*?

Important assumtions:

1. X is compact in $\mathbb{R}^K$ (closed and bounded), $U(\cdot)$ is continuous.
2. $U(\cdot)$ is concave.
3. $U(\cdot, y)$ is strictly increasing in its first K arguments.
4. U is continuously differentiable in the interior of its domain X_G.

Their functions:

- (i) guarantees a $V(\cdot)$ exists. (ii) further ensures the uniqueness of a $V(\cdot)$. (iii) and (iv) add further properties, such as continuity and differentiability, to $V(\cdot)$.

Is V satisfying Bellman Equation also V^*?

Important assumtions:

1. (i) X is compact in $\mathbb{R}^K$ (closed and bounded), $U(\cdot)$ is continuous.
2. (ii) $U(\cdot)$ is concave.
3. (iii) $U(\cdot, y)$ is strictly increasing in its first K arguments.
4. (iv) U is continuously differentiable in the interior of its domain X_G.

Their functions:

- (i) guarantees a $V(\cdot)$ exists. (ii) further ensures the uniqueness of a $V(\cdot)$. (iii) and (iv) add further properties, such as continuity and differentiability, to $V(\cdot)$.
- So, given the uniqueness, we know that $V(\cdot)$ satisfying the Bellman equation is also V^* solving Problem A2.

Bellman Equation

- We have shown that, under pretty weak assumptions, finding the $V^*(\cdot)$ in Problem A2 is equivalent to finding the $V(\cdot)$ in Problem A3.
- We haven't answer the question: Is it easier or more convenient to search for $V(\cdot)$ instead of $V^*(\cdot)$?

Bellman Equation

- We have shown that, under pretty weak assumptions, finding the $V^*(\cdot)$ in Problem A2 is equivalent to finding the $V(\cdot)$ in Problem A3.
- We haven't answer the question: Is it easier or more convenient to search for $V(\cdot)$ instead of $V^*(\cdot)$?
- To answer the question above, as well as to unfold the *beauty* of the Bellman Equation, we should take a detour by spending some (rewarding) time on *contraction mapping*.

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?
- Newton, in 1669, tried this:

$$x = \sqrt{5} \;\Rightarrow\; f(x) = x^2 - 5 = 0 \;\Rightarrow\; x = x - \frac{f(x)}{f'(x)}$$

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?
- Newton, in 1669, tried this:

$$x = \sqrt{5} \;\Rightarrow\; f(x) = x^2 - 5 = 0 \;\Rightarrow\; x = x - \frac{f(x)}{f'(x)}$$

- What if we calculate $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 5}{2x_n}$ iteratively?

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?
- Newton, in 1669, tried this:

$$x = \sqrt{5} \;\Rightarrow\; f(x) = x^2 - 5 = 0 \;\Rightarrow\; x = x - \frac{f(x)}{f'(x)}$$

- What if we calculate $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 5}{2x_n}$ iteratively?
- Given $x_1 = 2$, we have $x_2 = 2.25$, $x_3 = 2.23611111$, $x_4 = 2.2360679779$, $x_5 = 2.2360679774997896964$, $\cdots$

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?
- Newton, in 1669, tried this:

$$x = \sqrt{5} \;\Rightarrow\; f(x) = x^2 - 5 = 0 \;\Rightarrow\; x = x - \frac{f(x)}{f'(x)}$$

- What if we calculate $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 5}{2x_n}$ iteratively?
- Given $x_1 = 2$, we have $x_2 = 2.25$, $x_3 = 2.23611111$, $x_4 = 2.2360679779$, $x_5 = 2.23606797749978969644$, $\cdots$
- this method is obviously applicable to many equation-solving scenarios.

Newton's Method: A Taste

- How to approximate $x = \sqrt{5}$?
- Newton, in 1669, tried this:

$$x = \sqrt{5} \;\Rightarrow\; f(x) = x^2 - 5 = 0 \;\Rightarrow\; x = x - \frac{f(x)}{f'(x)}$$

- What if we calculate $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 5}{2x_n}$ iteratively?
- Given $x_1 = 2$, we have $x_2 = 2.25$, $x_3 = 2.23611111$, $x_4 = 2.2360679779$, $x_5 = 2.23606797749978969644$, $\cdots$
- this method is obviously applicable to many equation-solving scenarios.
- We will show that the method here is a *contraction mapping* on a properly defined domain.

Contraction Mapping

Definition 1

Let (S, d) be a metric space and $T : S \to S$ be an operator mapping S into itself. If for some $\beta \in (0, 1)$,

$$d(Tz_1, Tz_2) \leq \beta d(z_1, z_2) \quad \text{for all} \quad z_1, z_2 \in S,$$

Then T is a contraction mapping (with modulus β).

Contraction Mapping

Definition 1

Let (S, d) be a metric space and $T : S \to S$ be an operator mapping S into itself. If for some $\beta \in (0, 1)$,

$$d(Tz_1, Tz_2) \leq \beta d(z_1, z_2) \quad \text{for all} \quad z_1, z_2 \in S,$$

Then T is a contraction mapping (with modulus β).

- A contraction mapping makes *any* couple of elements *closer*.

Contraction Mapping (Continued)

Theorem 1

(Contraction Mapping Theorem) Let (S, d) be a complete metric space and suppose that $T : S \to S$ is a contraction mapping. Then T has a unique $\hat{z} \in S$ such that

$$T\hat{z} = \hat{z}.$$

Contraction Mapping (Continued)

- The formal proof of Theorem 1 is omitted here. The intuition, however, is quite straightforward: Starting from any given point in S, impose T infinitely many times. As the contraction mapping makes the adjacent pair of points closer and closer, the resulting Cauchy sequence must converges to a point in a complete space.

Contraction Mapping (Continued)

- The formal proof of Theorem 1 is omitted here. The intuition, however, is quite straightforward: Starting from any given point in S, impose T infinitely many times. As the contraction mapping makes the adjacent pair of points closer and closer, the resulting Cauchy sequence must converges to a point in a complete space.
- Don't worry about the requirement "complete". The spaces usually dealt with in Economics are all complete: the Euclidean space, the space of continuous real-valued function on a compact set, and so on.

Contraction Mapping (Continued)

Theorem 2

(Applications of Contraction Mappings) Let (S, d) be a complete metric space and $T : S \to S$ be a contraction mapping with $T\hat{z} = \hat{z}$.

a. *If S' is a closed subset of S, and $T(S') \subset S'$, then $\hat{z} \in S'$.*
b. *Moreover, if $T(S') \subset S'' \subset S'$, then $\hat{z} \in S''$.*

Contraction Mapping (Continued)

Theorem 2

(Applications of Contraction Mappings) Let (S, d) be a complete metric space and $T : S \to S$ be a contraction mapping with $T\hat{z} = \hat{z}$.

- a. *If S' is a closed subset of S, and $T(S') \subset S'$, then $\hat{z} \in S'$.*
- b. *Moreover, if $T(S') \subset S'' \subset S'$, then $\hat{z} \in S''$.*

- These results are irrelevant to the main materials in this note.

Contraction Mapping (Continued)

Theorem 2

(Applications of Contraction Mappings) Let (S, d) be a complete metric space and $T: S \to S$ be a contraction mapping with $T\hat{z} = \hat{z}$.

a. *If S' is a closed subset of S, and $T(S') \subset S'$, then $\hat{z} \in S'$.*
b. *Moreover, if $T(S') \subset S'' \subset S'$, then $\hat{z} \in S''$.*

- These results are irrelevant to the main materials in this note.
- Question: If you have the opportunity to draw a pretty precise map of China and spread it out on the square in front of the school building, are you capable of finding a point, if any, on the map, that coincides with its corresponding geographical location on the earth?

Newton's Method Revisited

Define $g(x) = x - \frac{f(x)}{f'(x)}$. Then $f(\hat{x}) = 0 \Rightarrow g(\hat{x}) = \hat{x} = \sqrt{5}$. Furthermore, with $f(x) = x^2 - 5$, we have

$$|g(x) - g(\sqrt{5})| = |g(x) - \sqrt{5}| = |x - \sqrt{5}|| \frac{\sqrt{5}x - 5}{2\sqrt{5}x}|, \quad \text{for } x \in \mathbb{R}_+. \quad (16)$$

Newton's Method Revisited

Define $g(x) = x - \frac{f(x)}{f'(x)}$. Then $f(\hat{x}) = 0 \;\Rightarrow\; g(\hat{x}) = \hat{x} = \sqrt{5}$. Furthermore, with $f(x) = x^2 - 5$, we have

$$|g(x) - g(\sqrt{5})| = |g(x) - \sqrt{5}| = |x - \sqrt{5}|\left|\frac{\sqrt{5}x - 5}{2\sqrt{5}x}\right|, \quad \text{for } x \in \mathbb{R}_+. \tag{16}$$

- With x_1 being close to $\sqrt{5}$, $g(\cdot)$ is a contraction mapping within some subspace of $(\mathbb{R}_+, |\cdot|)$!

Newton's Method Revisited

Define $g(x) = x - \frac{f(x)}{f'(x)}$. Then $f(\hat{x}) = 0 \;\Rightarrow\; g(\hat{x}) = \hat{x} = \sqrt{5}$. Furthermore, with $f(x) = x^2 - 5$, we have

$$|g(x) - g(\sqrt{5})| = |g(x) - \sqrt{5}| = |x - \sqrt{5}|\left|\frac{\sqrt{5}x - 5}{2\sqrt{5}x}\right|, \quad \text{for } x \in \mathbb{R}_+. \tag{16}$$

- With x_1 being close to $\sqrt{5}$, $g(\cdot)$ is a contraction mapping within some subspace of $(\mathbb{R}_+, |\cdot|)$!
- The initialization point chosen in Newton's method is *crucial*.

Bellman Equation and Contraction Mapping

Recall the Bellman equation:

$$V(x) = \sup_{y \in G(x)} \{U(x, y) + \beta V(y)\}, \quad \text{for all} \quad x \in X. \tag{17}$$

$$= TV(x), \qquad \text{for all} \quad x \in X, \tag{18}$$

where the second equality defines operator T.

- We will show that T is a contraction mapping.
- Space for T to operate: all bounded functions defined on X.
- So T is a *functional*: It maps a function to another.
- How should we choose the metric, $d(\cdot)$, of the space?

Define the Metric

For functions f, g defined on X, the *supremum norm* is used for metric:

$$d(f,g) = \|f - g\| = \sup_{x \in X} |f(x) - g(x)| \tag{19}$$

- The "distance" between two functions defined on X is determined by the greatest "gap" between the two functions on X.
- You can, of course, adopt other kinds of metrics. But any well defined norm will prove the same result as the supremum norm does.

Prove T is a Contraction Mapping

Proposition 1

Let $X \in \mathbb{R}^K$ and $B(X)$ be the space of bounded functions $f : X \to \mathbb{R}$ defined on X, equipped with the supremum norm $\|\cdot\|$, Define mapping $T : B(X) \to B(X)$ by

$$(Tf)(x) = \sup_{y \in G(x)} \{U(x, y) + \beta f(y)\}, \quad \forall x \in X, \quad \forall f \in B(X).$$

Then T is a contraction mapping.

Prove T is a Contraction Mapping (Continued)

Proof: Given $f, g \in B(X)$,

$$f(x) - g(x) \le |f(x) - g(x)| \le \|f - g\| \quad \forall x \in X \tag{20}$$

$$\Rightarrow f(x) \le g(x) + \|f - g\| \quad \forall x \in X \tag{21}$$

$$\Rightarrow (Tf)(x) = \sup_{y \in G(x)} \{U(x, y) + \beta f(y)\} \quad \forall x \in X$$

$$\le \sup_{y \in G(x)} \{U(x, y) + \beta g(y) + \beta \|f - g\|\} \quad \forall x \in X$$

$$\Rightarrow (Tf)(x) \le (Tg)(x) + \beta \|f - g\| \quad \forall x \in X \tag{22}$$

$$\text{Analogously,} \quad (Tg)(x) \le (Tf)(x) + \beta \|f - g\| \quad \forall x \in X \tag{23}$$

Combining (22) and (23) yields

$$\|Tf - Tg\| = \sup_{x \in X} |(Tf)(x) - (Tg)(x)| \le \beta \|f - g\| \tag{24}$$

T is thus a contraction mapping. □

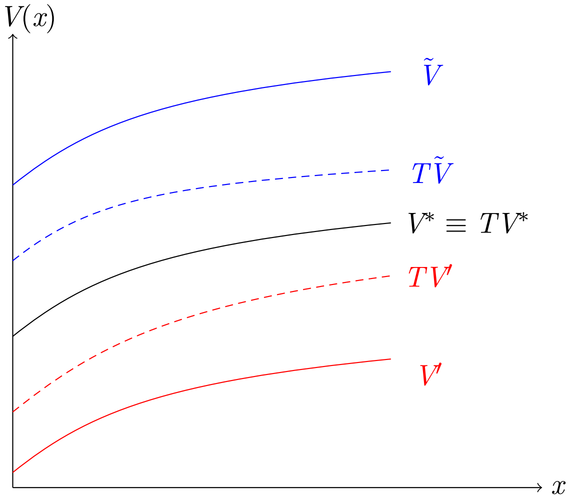


Figure 2: $T: B(X) \to B(X)$ defined in the Bellman equation is a contraction mapping.

Utilizing Contraction Mapping T

- We have shown that T is a contraction mapping in space $B(X)$.

Utilizing Contraction Mapping T

- We have shown that T is a contraction mapping in space $B(X)$.
- So, T admits a *unique* fixed point, i.e., a unique $V(\cdot)$ such that $V \equiv TV$.

Utilizing Contraction Mapping T

- We have shown that T is a contraction mapping in space $B(X)$.
- So, T admits a *unique* fixed point, i.e., a unique $V(\cdot)$ such that $V \equiv TV$.
- Recall that, the unique solution to the Bellman equation, V, is also the function V^* in Problem A2.

Utilizing Contraction Mapping T

- We have shown that T is a contraction mapping in space $B(X)$.
- So, T admits a *unique* fixed point, i.e., a unique $V(\cdot)$ such that $V \equiv TV$.
- Recall that, the unique solution to the Bellman equation, V, is also the function V^* in Problem A2.
- Then, as long as the fixed point in $B(X)$, function V, is found, we have the value function, and can thus deduce the policy function. Problem solved!

Utilizing Contraction Mapping T

- We have shown that T is a contraction mapping in space $B(X)$.
- So, T admits a *unique* fixed point, i.e., a unique $V(\cdot)$ such that $V \equiv TV$.
- Recall that, the unique solution to the Bellman equation, V, is also the function V^* in Problem A2.
- Then, as long as the fixed point in $B(X)$, function V, is found, we have the value function, and can thus deduce the policy function. Problem solved!
- Question: How to find the fixed point V?

Utilizing Contraction Mapping T (Continued)

- Answer: Just like how function g is used iteratively in the Newton's method example, we can guess here any kind of bounded function, denoted by V^1 in $B(X)$, and use T iteratively to generate functions $V^2, V^3, V^4, \cdots$. The contraction mapping T will make sure that the resulting sequence of function converges to the true V (or say, V^*).

Utilizing Contraction Mapping T (Continued)

- Answer: Just like how function g is used iteratively in the Newton's method example, we can guess here any kind of bounded function, denoted by V^1 in $B(X)$, and use T iteratively to generate functions $V^2, V^3, V^4, \cdots$. The contraction mapping T will make sure that the resulting sequence of function converges to the true V (or say, V^*).
- Complexity. Usually, it is impossible to track the iterative process in a analytical way. We instead use *numerical approximations*.

Utilizing Contraction Mapping T (Continued)

- Answer: Just like how function g is used iteratively in the Newton's method example, we can guess here any kind of bounded function, denoted by V^1 in $B(X)$, and use T iteratively to generate functions $V^2, V^3, V^4, \cdots$. The contraction mapping T will make sure that the resulting sequence of function converges to the true V (or say, V^*).
- Complexity. Usually, it is impossible to track the iterative process in a analytical way. We instead use *numerical approximations*.
- The homework questions give you some basic ideas on how to realize iterations in a functional space using numerical approximations. Please do spend enough time on it!

Another Approach: Euler Equation

Recall the Bellman equation:

$$V(x) = \sup_{y \in G(x)} \{ U(x, y) + \beta V(y) \} \quad \text{for all } x \in X. \tag{25}$$

Question: If all the assumptions regarding the functions and sets related to the optimization problem are taken, can we find out the fixed point V in an analytical way?

First, let us denote by $D_x U$ the gradient with respect to the first K arguments, and by $D_y U$ the gradient with respect to the last K arguments. The gradient DV is naturally defined.

Another Approach: Euler Equation (Continued)

- Suppose for a given x, $y^*(x) \in G(x)$ solves the problem, then we must have:

$$D_y U(x, y^*(x)) + \beta D V(y^*(x)) = 0. \tag{26}$$

Another Approach: Euler Equation (Continued)

- Suppose for a given x, $y^*(x) \in G(x)$ solves the problem, then we must have:

$$D_y U(x, y^*(x)) + \beta DV(y^*(x)) = 0. \tag{26}$$

- The *Euler equations* above should hold for all $x \in X$.

Another Approach: Euler Equation (Continued)

- Suppose for a given x, $y^*(x) \in G(x)$ solves the problem, then we must have:

$$D_y U(x, y^*(x)) + \beta D V(y^*(x)) = 0. \tag{26}$$

- The *Euler equations* above should hold for all $x \in X$.
- Also notice that $V(x) = U(x, y^*(x)) + \beta V(y^*(x))$ holds for all $x \in X$. Differentiating both sides with respect to x and inserting (26) into it yield (it could also be interpreted as an application of the Envelope Theorem with an interior solution assumed):

$$\begin{aligned} DV(x) =& D_x U(x, y^*(x)) + \left[\frac{\partial y^*(x)}{\partial x}\right]' [D_y U(x, y^*(x)) + \beta D V(y^*(x))] \\ =& D_x U(x, y^*(x)) \end{aligned} \tag{27}$$

Another Approach: Euler Equation (Continued)

- Recall y^* maps X into X. Applying recursively Equation (27), $DV(x) = D_x U(x, y^*(x))$, yields:

$$DV(y^*(x)) = D_x U\left(y^*(x), y^*\left(y^*(x)\right)\right) \tag{28}$$

- Denote $y^*(x) = \pi(x)$, after inserting Equation (28) into (26), the Euler equations appears:

$$D_y U(x, \pi(x)) + \beta D_x U(\pi(x), \pi(\pi(x))) = 0, \qquad \forall x \in X. \tag{29}$$

Another Approach: Euler Equation (Continued)

When both x and y are variables (vectors with dimension 1), the Euler equations are:

$$\frac{\partial U(x(t), x^*(t+1))}{\partial y} + \beta \frac{\partial U(x^*(t+1), x^*(t+2))}{\partial x} = 0 \tag{30}$$

Another Approach: Euler Equation (Continued)

When both x and y are variables (vectors with dimension 1), the Euler equations are:

$$\frac{\partial U(x(t), x^*(t+1))}{\partial y} + \beta \frac{\partial U(x^*(t+1), x^*(t+2))}{\partial x} = 0 \tag{30}$$

- The Euler equations themselves are only *necessary* condtions for the problem, combining with the *transversality condition*

$$\lim_{t \to \infty} \beta^t D_x U(x^*(t), x^*(t+1)) \cdot x^*(t) = 0$$

makes them both necessary and sufficient conditions.

Another Approach: Euler Equation (Continued)

When both x and y are variables (vectors with dimension 1), the Euler equations are:

$$\frac{\partial U(x(t), x^*(t+1))}{\partial y} + \beta \frac{\partial U(x^*(t+1), x^*(t+2))}{\partial x} = 0 \tag{30}$$

- The Euler equations themselves are only *necessary* condtions for the problem, combining with the *transversality condition*

$$\lim_{t \to \infty} \beta^t D_x U(x^*(t), x^*(t+1)) \cdot x^*(t) = 0$$

  makes them both necessary and sufficient conditions.
- If you are **lucky enough**, you can solve the problem using Euler equations in a analytical and beautiful way, by correctly guessing the form of the policy function.

Example 6.4

Now do the following exercise:

$$\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c(t)$$

$$\begin{aligned} \text{subject to} \qquad & k(t+1) = k(t)^{\alpha} - c(t), \\ & k(0) > 0, \beta \in (0,1) \end{aligned}$$

Example 6.4

Now do the following exercise:

$$\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c(t)$$

$$\text{subject to} \quad k(t+1) = k(t)^{\alpha} - c(t),$$

$$k(0) > 0, \beta \in (0,1)$$

- Method 1: Guess the policy function as $\pi(x) = \gamma x^{\alpha}$, and verify your guess by determining the value of γ. (economic intuition?)

Example 6.4

Now do the following exercise:

$$\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c(t)$$

$$\text{subject to} \quad k(t+1) = k(t)^\alpha - c(t),$$

$$k(0) > 0, \beta \in (0,1)$$

- Method 1: Guess the policy function as $\pi(x) = \gamma x^\alpha$, and verify your guess by determining the value of γ. (economic intuition?)
- Method 2: Guess the value function as $V(x) = \lambda + \xi \log x$, and verify your guess by determining the values of λ and ξ.

Example 6.4

Now do the following exercise:

$$\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log c(t)$$

subject to
$$k(t+1) = k(t)^{\alpha} - c(t),$$
$$k(0) > 0, \beta \in (0,1)$$

- Method 1: Guess the policy function as $\pi(x) = \gamma x^{\alpha}$, and verify your guess by determining the value of γ. (economic intuition?)
- Method 2: Guess the value function as $V(x) = \lambda + \xi \log x$, and verify your guess by determining the values of λ and ξ.
- You should find that the two methods above are equivalent.

Guess and Verify

- Question: If you have a guess that is successfully verified, can it be an incorrect one?

Guess and Verify

- Question: If you have a guess that is successfully verified, can it be an incorrect one?
- Answer: usually not in economic applications, especially after assumptions $(i) - (iv)$ have been made. The tough question here is how would you know the specific form of the policy function (value function) without any clue, for any given utility function?

Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.

Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.
- What if there are noises in the problem, e.g., noises in state-evolution function $\tilde{f}$?

Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.
- What if there are noises in the problem, e.g., noises in state-evolution function $\tilde{f}$?
- Answer: $E(\cdot)$ appears in the objective function.

Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.
- What if there are noises in the problem, e.g., noises in state-evolution function $\tilde{f}$?
- Answer: $E(\cdot)$ appears in the objective function.
- What if the function $\tilde{f}$ is totally unknown?

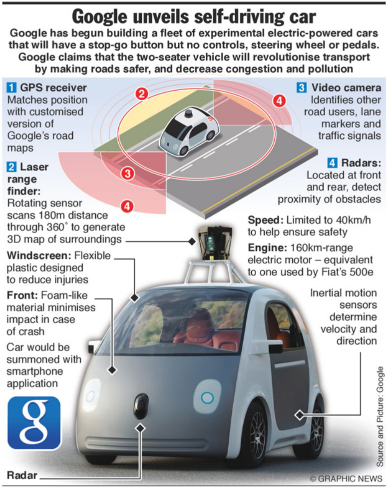
Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.
- What if there are noises in the problem, e.g., noises in state-evolution function $\tilde{f}$?
- Answer: $E(\cdot)$ appears in the objective function.
- What if the function $\tilde{f}$ is totally unknown?
- Learn it from experience!

Miscellaneous Notes

- There are also tools for *non-stationary* dynamic programming problems.
- What if there are noises in the problem, e.g., noises in state-evolution function $\tilde{f}$?
- Answer: $E(\cdot)$ appears in the objective function.
- What if the function $\tilde{f}$ is totally unknown?
- Learn it from experience!
- *Reinforcement Learning*. Everybody is talking about Artificial Intelligence and Machine Learning!

(a) AlphaGo by DeepMind



(b) Goole Self-Driving Car

Figure 3: The A.I. mania relies intensively on dynamic programming tools.

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

$$J[x] = \int_{t_1}^{t_2} L\left(t, x(t), x'(t)\right) \mathrm{d}t \,, \tag{31}$$

where

- t_1, t_2 are constants.
- $x(t)$ is twice continuously differentiable.
- $x'(t) = \frac{\mathrm{d}x}{\mathrm{d}t}$.
- $L\left(t, x(t), x'(t)\right)$ is twice continuously differentiable with respect to all arguments.

Euler-Lagrange equation (Continued)

If $J[x]$ attains a local maximum at f, and $\eta(t)$ is an arbitrary function that has at least one derivative and vanishes at the endpoints t_1 and t_2, then for any number $\varepsilon \to 0$, we must have

$$J[f] \geq J[f + \varepsilon\eta]. \tag{32}$$

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

$$\Phi(\varepsilon) = J[f + \varepsilon\eta]. \tag{33}$$

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

$$\Phi'(0) = \left.\frac{\mathrm{d}\Phi}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} = \int_{t_1}^{t_2} \left.\frac{\mathrm{d}L}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathrm{d}t = 0. \tag{34}$$

Euler-Lagrange equation (Continued)

Now taking total derivative of $L\left[t, f+\varepsilon\eta, (f+\varepsilon\eta)'\right]$, we have:

$$\frac{\mathrm{d}L}{\mathrm{d}\varepsilon} = \frac{\partial L}{\partial x}\eta + \frac{\partial L}{\partial x'}\eta'. \tag{35}$$

Inserting (35) into (34) gives us

$$\begin{aligned} 0 = \int_{t_1}^{t_2} \left.\frac{\mathrm{d}L}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathrm{d}t &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial f}\eta + \frac{\partial L}{\partial f'}\eta'\right) \mathrm{d}t \\ &= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial f}\eta - \eta\frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t}\right) \mathrm{d}t + \left.\frac{\partial L}{\partial f'}\eta\right|_{t_1}^{t_2} \\ &= \int_{t_1}^{t_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t}\right) \mathrm{d}t, \end{aligned}$$

where the last two equalities use integration by parts and the fact that η vanishes at t_1 and t_2.

Euler-Lagrange equation (Continued)

Now given

$$\int_{t_1}^{t_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} \right) \mathrm{d}t = 0, \tag{36}$$

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

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} = 0 \quad , \forall t \in (t_1, t_2) \tag{37}$$

must hold!

Euler-Lagrange equation (Continued)

Now given

$$\int_{t_1}^{t_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} \right) \mathrm{d}t = 0, \tag{36}$$

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

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} = 0 \quad , \forall t \in (t_1, t_2) \tag{37}$$

must hold!

- However, it is possible to attain (37) based on (36) without applying the lemma!

Euler-Lagrange equation (Continued)

Now given

$$\int_{t_1}^{t_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} \right) \mathrm{d}t = 0, \tag{36}$$

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

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} = 0 \quad , \forall t \in (t_1, t_2) \tag{37}$$

must hold!

- However, it is possible to attain (37) based on (36) without applying the lemma!
- A special form of η?

Euler-Lagrange equation (Continued)

Now given

$$\int_{t_1}^{t_2} \eta \left(\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} \right) \mathrm{d}t = 0, \tag{36}$$

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

$$\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} = 0 \quad , \forall t \in (t_1, t_2) \tag{37}$$

must hold!

- However, it is possible to attain (37) based on (36) without applying the lemma!
- A special form of η?
- How about $\eta(t)$ equals $-(t - t_1)(t - t_2) \left[\frac{\partial L}{\partial f} - \frac{\mathrm{d}(\frac{\partial L}{\partial f'})}{\mathrm{d}t} \right]$ for $t \in [t_1, t_2]$ and 0 for $t \notin [t_1, t_2]$?

A simple exercise

Consider the following problem:

$$\max_{[c(t),a(t)]_{t=0}^{1}} \int_0^1 e^{-\rho t} u(c(t))\mathrm{d}t, \tag{38}$$

$$\text{subject to} \quad \dot{a}(t) = ra(t) + \omega - c(t), \quad a(0) = a_0, \ a(1) = 0. \tag{39}$$

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:

$$\min_{y(x)} J(y) = \int_{x_0}^{x_1} \frac{\sqrt{1 + [y'(x)]^2}}{\sqrt{2g[y_0 - y(x)]}} \mathrm{d}x \tag{40}$$

$$\text{subject to} \quad y(x_0) = y_0, \; y(x_1) = y_1. \tag{41}$$

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 B1:

$$\max_{x(t),y(t),x_1} \quad W(x(t),y(t)) = \int_0^{t_1} f(t,x(t),y(t))\,\mathrm{d}t\ , \tag{42}$$

$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{43}$$

$$x(0) = x_0,\ \text{and}\ x(t_1) = x_1. \tag{44}$$

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 $\hat{y}(t)$:

$$y(t, \epsilon) = \hat{y}(t) + \epsilon\, \eta(t). \tag{45}$$

Note that given $y(t, \epsilon)$, $x(t)$ is now dependent on ϵ according to evolutionary equation (43), so the resulting $x(t, \epsilon)$ is defined by:

$$\dot{x}(t, \epsilon) = g(t, x(t, \epsilon), y(t, \epsilon)) \text{ for all } t \in [0, t_1], \text{ with } x(0, \epsilon) = x_0. \tag{46}$$

Define:

$$\begin{aligned} \mathcal{W}(\epsilon) &= W(x(t, \epsilon), y(t, \epsilon)) \\ &= \int_0^{t_1} f(t, x(t, \epsilon), y(t, \epsilon))\, \mathrm{d}t. \end{aligned} \tag{47}$$

Variational Arguments (Continued)

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

$$\mathcal{W}(\epsilon) \leq \mathcal{W}(0) \quad \text{for all small engough} \quad \epsilon \to 0. \tag{48}$$

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:

$$\int_0^{t_1} \lambda(t) [g(t, x(t, \epsilon), y(t, \epsilon)) - \dot{x}(t, \epsilon)] \, \mathrm{d}t = 0. \tag{49}$$

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 (47) and (49) lets us redefine $\mathcal{W}(\epsilon)$:

$$\begin{aligned}
\mathcal{W}(\epsilon) &= \int_0^{t_1} f(t, x(t,\epsilon), y(t,\epsilon))\,\mathrm{d}t + 0 \\
&= \int_0^{t_1} \{f(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t)[g(t, x(t,\epsilon), y(t,\epsilon)) - \dot{x}(t,\epsilon)]\}\mathrm{d}t \\
&= \int_0^{t_1} \{f(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t) g(t, x(t,\epsilon), y(t,\epsilon)) + \dot{\lambda}(t) x(t,\epsilon)\}\mathrm{d}t \\
&\quad -\lambda(t_1) x(t_1,\epsilon) + \lambda(0) x_0. \qquad (50)
\end{aligned}$$

The last equality above uses integration by parts.

Variational Arguments (Continued)

Applying Leibniz's Rule to (50) yields:

$$\begin{aligned} \mathcal{W}'(\epsilon) &= \int_0^{t_1} \left[f_x(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t) g_x(t, x(t,\epsilon), y(t,\epsilon)) + \dot{\lambda}(t) \right] x_\epsilon(t,\epsilon) \mathrm{d}t \\ &\quad + \int_0^{t_1} \left[f_y(t, x(t,\epsilon), y(t,\epsilon)) + \lambda(t) g_y(t, x(t,\epsilon), y(t,\epsilon)) \right] \eta(t) \mathrm{d}t \\ &\quad - \lambda(t_1) x_\epsilon(t_1, \epsilon) \end{aligned} \tag{51}$$

Recall that condition (48) can be rewritten as $\mathcal{W}'(0) = 0$. We thus have:

$$\begin{aligned} 0 &= \int_0^{t_1} \left[f_x(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_x(t, \hat{x}(t), \hat{y}(t)) + \dot{\lambda}(t) \right] x_\epsilon(t, 0) \mathrm{d}t \\ &\quad + \int_0^{t_1} \left[f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t)) \right] \eta(t) \mathrm{d}t \\ &\quad - \lambda(t_1) x_\epsilon(t_1, 0) \end{aligned} \tag{52}$$

Variational Arguments (Continued)

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

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

$$\lambda(t_1) = 0 \tag{53}$$

- 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 (52) must hold for any continuously differentiable $\lambda(t)$, why not focus on the following $\lambda(t)$:

$$\begin{aligned}\dot{\lambda}(t) &= -\left[f_x(t,\hat{x}(t),\hat{y}(t))+\lambda(t)g_x(t,\hat{x}(t),\hat{y}(t))\right], \\ \lambda(t_1) &= 0\end{aligned} \tag{54}$$

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:

$$f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t)) = 0 \quad \text{for all} \quad t \in [0, t_1] \tag{55}$$

Variational Arguments (Continued)

Theorem 3

Suppose Problem B1 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* (43), (53), (54), *and* (55) *hold.*

Problem B2

$$\max_{x(t),y(t)} \quad W(x(t), y(t)) = \int_0^{t_1} f(t, x(t), y(t))\, \mathrm{d}t\ ,$$
$$\text{subject to} \quad \dot{x}(t) = g(t, x(t), y(t)), \tag{56}$$
$$x(0) = x_0, \text{ and } x(t_1) = x_1.$$

Theorem 4

Suppose Problem B2 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* (56), (54), *and* (55) *hold.*

Can you figure out how and why Theorem 4 differs from Theorem 3?

Revisiting the simple exercise

Consider the following problem:

$$\max_{[c(t),a(t)]_{t=0}^{1}} \int_0^1 e^{-\rho t} u(c(t)) \mathrm{d}t, \tag{57}$$

$$\text{subject to} \quad \dot{a}(t) = ra(t) + \omega - c(t), \quad a(0) = a_0, \ a(1) = 0. \tag{58}$$

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 B3

$$\max_{x(t),y(t)} W(x(t),y(t)) = \int_0^{t_1} f(t,x(t),y(t))\,\mathrm{d}t \,,$$
$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{59}$$
$$x(0)=x_0, \text{ and } x(t_1) \geq x_1.$$

Theorem 5

Suppose Problem B3 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* (59), (54), (55), *and* $\lambda(t_1)[x(t_1)-x_1]=0$ *hold.*

Can you figure out how and why Theorem 5 differs from Theorem 3?

Pontryagin's Maximum Principle

Revisit Problem B1. Define the **Hamiltonian**:

$$H(t, x(t), y(t), \lambda(t)) \equiv f(t, x(t), y(t)) + \lambda(t) g(t, x(t), y(t)). \qquad (60)$$

Theorem 6

Suppose Problem B1 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:

$$\dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{y}_(t), \lambda(t)) \quad \text{for all} \quad t \in [0, t_1], \qquad (61)$$

$$\dot{x}(t) = H_\lambda(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \quad \text{for all} \quad t \in [0, t_1], \qquad (62)$$

$$H(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \geq H(t, \hat{x}(t), y, \lambda(t)) \text{ for all feasible } y, t, \qquad (63)$$

$$x(0) = x_0, \quad \lambda(t_1) = 0. \qquad (64)$$

Question: relationship between (55) and (63)?

Pontryagin's Maximum Principle (Continued)

- The results of equations (61)-(64) are straightforward to understand given the variational arguments illustrated earlier.
- Can you figure out the Pontryagin's Maximum Principle for Problems B2 and B3?
- (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 B4:

$$\max_{x(t),y(t)} \; W(x(t),y(t)) = \int_0^\infty f(t,x(t),y(t))\,\mathrm{d}t\,, \tag{65}$$

$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{66}$$

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

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 7

Suppose Problem B4 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:*

$$\dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{y}_(t), \lambda(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{67}$$

$$\dot{x}(t) = H_\lambda(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{68}$$

$$H(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \geq H(t, \hat{x}(t), y, \lambda(t)) \text{ for all feasible } y, t, \tag{69}$$

$$x(0) = x_0, \quad \lim_{t \to \infty} b(t) x(t) \geq x_1. \tag{70}$$

Current-Value Hamiltonian

Consider Problem B5:

$$\max_{x(t),y(t)} \; W(x(t),y(t)) = \int_0^{\infty} e^{-\rho t} f(x(t),y(t))\,\mathrm{d}t\,, \tag{71}$$

$$\text{subject to} \quad \dot{x}(t) = g(t,x(t),y(t)), \tag{72}$$

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

The Hamiltonian is:

$$H(t,x(t),y(t),\lambda(t)) = e^{-\rho t}\left[f(x(t),y(t)) + e^{\rho t}\lambda(t)g(t,x(t),y(t))\right] \tag{73}$$

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

$$\hat{H}(t,x(t),y(t),\mu(t)) \equiv f(x(t),y(t)) + \mu(t)g(t,x(t),y(t)). \tag{74}$$

Current-Value Hamiltonian(continued)

Theorem 8

Suppose Problem B5 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:

$$\hat{H}_y(t, \hat{x}(t), \hat{y}(t), \mu(t)) = 0 \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{75}$$

$$\rho\mu(t) - \dot{\mu}(t) = \hat{H}_x(t, \hat{x}(t), \hat{y}(t), \mu(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{76}$$

$$\dot{x}(t) = \hat{H}_\mu(t, \hat{x}(t), \hat{y}(t), \mu(t)) \quad \text{for all} \quad t \in \mathbb{R}_+, \tag{77}$$

$$\lim_{t\to\infty} [e^{-\rho t}\mu(t)\hat{x}(t)] = 0 \tag{78}$$

$$x(0) = x_0, \quad \lim_{t\to\infty} b(t)\hat{x}(t) \geq x_1. \tag{79}$$

(78) 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.