Advanced Macroeconomics 1

Oskari Vähämaa (University of Helsinki)

September 24, 2019

Lecture 6:
Dynamic optimization in continuous time

Optimal control (ref: Acemoglu ch7)

The canonical continuous time problem

- can be written as

$$\max_{\mathbf{x}(t),\mathbf{y}(t)} W(\mathbf{x}(t),\mathbf{y}(t)) \equiv \int_0^{t_1} f(t,\mathbf{x}(t),\mathbf{y}(t))dt \tag{1}$$

subject to

$$\dot{\mathbf{x}}(t) = \mathbf{G}(t,\mathbf{x}(t),\mathbf{y}(t))$$

$$\mathbf{x}(t) \in \mathcal{X}(t) \in \mathbb{R}^{K_x}, \ \mathbf{y}(t) \in \mathcal{Y}(t) \in \mathbb{R}^{K_y}, K_x, K_y \in \mathbb{N}$$

for all t and $\mathbf{x}(0) = \mathbf{x_0}$, where $f : \mathbb{R} \times \mathbb{R}^{K_x} \times \mathbb{R}^{K_y} \to \mathbb{R}$ and $G : \mathbb{R} \times \mathbb{R}^{K_x} \times \mathbb{R}^{K_y} \to \mathbb{R}^{K_x}$.

- Vector x denotes the state variables. The evolution of this is governed by a system of difference equations, given the vector of control variables y.
- The end of the planning horizon can be equal to infinity.
- $W(\mathsf{x}(t), \mathsf{y}(t))$ denotes the value of the objective function when the controls are given by y and the resulting behavior of state variables is summarized by x.
- Note that, now the maximization is with respect to an infinite dimensional object, $y : [t_0, t_1] \to \mathbb{R}$.

Variational arguments

- Consider a special case of (1), where t_1 is finite and where both the state and the control variables are one-dimensional
- We have

$$\max_{\mathbf{x}(t),\mathbf{y}(t),x_1} W(\mathbf{x}(t),\mathbf{y}(t)) \equiv \int_0^{t_1} f(t,x(t),y(t))dt \tag{2}$$

st

$$\dot{x}(t) = g(t,x(t),y(t)) \tag{3}$$

$$x(t) \in \mathcal{X} \subset \mathbb{R}, y(t) \in \mathcal{Y} \subset \mathbb{R} \qquad \forall t,\ x(0) = x_0, x(t_1) = x_1 \tag{4}$$

- $\mathcal{X}$ and $\mathcal{Y}$ are assumed to be non-empty and convex.
- A pair of functions $x(t)$ and $y(t)$ that satisfies (3) and (4) is referred as admissible.
- $W(x(t), y(t)) < \infty$ is assumed for all admissible pair $(x(t), y(t))$.
- Moreover, it is assumed that f and g are continuously differentiable functions of x, y and t.
- Challenges:
  1. We are choosing a function $y : [0, t_1] \to \mathcal{Y}$ rather than a vector or a finite-dimensional object.
  2. Constraint takes the form of a differential equation rather than that of a set of inequalities or equalities.

- Assume that there exists a continuous solution (function) $\hat{y}$
    - that lies everywhere in the interior of the set $\mathcal{Y}$
    - with corresponding state variable, $\hat{x}$, everywhere in the interior of $\mathcal{X}$.
- We have
$$W(\hat{x}(t), \hat{y}(t)) \geq W(x(t), y(t)),$$
  for any admissible pair $(x(t), y(t))$
- Variational argument: there shouldn't be any small changes in controls that increase the value of objective function (analogy: FOCs in standard calculus).
- What is a small deviation here?

- Take an arbitrary fixed continuous function $\eta(t)$
- Let $\varepsilon \in \mathbb{R}$ be a real number
- Then a variation of the function $\hat{y}(t)$

$$y(t, \varepsilon) \equiv \hat{y}(t) + \varepsilon\eta(t),$$

  given $\eta(t)$, $y(t, \varepsilon)$ is obtained by varying ϵ
- Some of these variations may be infeasible.
- However, since $\hat{y}(t) \in int\,\mathcal{Y}$ and continuous function over a compact set $[0, t_1]$ is bounded, for any fixed $\eta(\cdot)$ we can always find ε'_η such that

$$y(t, \varepsilon) = \hat{y}(t) + \varepsilon\eta(t) \in int\,\mathcal{Y}$$

  for all $t \in [0, t_1]$ and for all $\varepsilon \in [-\varepsilon'_\eta, \varepsilon'_\eta]$ so that $y(t, \varepsilon)$ is feasible.

- The path of the state variable corresponding to the path of the control variable can be defined by

$$\dot{x}(t, \varepsilon) = g(t, x(t, \varepsilon), y(t, \varepsilon)$$

for all $t \in [0, t_1]$, with $x(0, \varepsilon) = x_0$

- Since $\hat{x}(t) \in int\,\mathcal{X}$, for all t, for sufficiently small ε, i.e. $\varepsilon \in [-\varepsilon_\eta, \varepsilon_\eta] \subset [-\varepsilon'_\eta, \varepsilon'_\eta]$ for some $\varepsilon_\eta < \varepsilon'_\eta$, we have that $x(\varepsilon, t) \in int\mathcal{X}$ for all t.
- Thus, when $\varepsilon \in [-\varepsilon_\eta, \varepsilon_\eta]$, $(x(t, \varepsilon), y(t, \varepsilon))$ is feasible.

- Define

$$W(\varepsilon) \equiv W(x(t,\varepsilon), y(t,\varepsilon)) = \int_0^{t_1} f(t,(x(t,\varepsilon), y(t,\varepsilon))dt \qquad (5)$$

- As $\hat{y}(t)$ is optimal and for $\varepsilon \in [-\varepsilon_\eta, \varepsilon_\eta]$, $x(t,\varepsilon)$ and $y(t,\varepsilon)$ feasible, we have

$$W(\varepsilon) \leq W(0) \qquad \forall \varepsilon \in [-\varepsilon_\eta, \varepsilon_\eta]$$

- Next rewrite (3), as

$$g(t, x(t,\varepsilon), y(t,\varepsilon)) - \dot{x}(t) = 0$$

for all t. Thus for any function $\lambda : [0 : t_1] \to \mathbb{R}$

$$\int_0^{t_1} \lambda(t)[g(t, x(t,\varepsilon), y(t,\varepsilon)) - \dot{x}(t)]dt = 0. \qquad (6)$$

- Assume $\lambda(\cdot)$ is continuously differentiable. When chosen suitable, this costate variable is analogous to the Lagrange multiplier.

- Adding (6) to (5) yields

$$W(\varepsilon) = \int_0^{t_1} \{f(t, (x(t,\varepsilon), y(t,\varepsilon)) + \lambda(t)[g(t, x(t,\varepsilon), y(t,\varepsilon)) - \dot{x}(t,\varepsilon)]\}dt \tag{7}$$

- To proceed, consider the integral $\int_0^{t_1} \lambda(t)\dot{x}(t,\varepsilon)dt$. Integrate this by parts

$$\int_0^{t_1} \lambda(t)\dot{x}(t,\varepsilon)dt = \lambda(t_1)x(t_1,\varepsilon) - \lambda(0)x_0 - \int_0^{t_1} \dot{\lambda}(t)x(t,\varepsilon)dt$$

- Plug this into (7)

$$\begin{aligned} W(\varepsilon) = \int_0^{t_1} & [f(t, (x(t,\varepsilon), y(t,\varepsilon)) + \lambda(t)g(t, x(t,\varepsilon), y(t,\varepsilon)) \\ & + \dot{\lambda}(t)x(t,\varepsilon)]dt - \lambda(t_1)x(t_1,\varepsilon) + \lambda(0)x_0 \end{aligned}$$

- Differentiate $W(\varepsilon)$ wrt ε

$$
\begin{aligned}
W'(\varepsilon) \equiv & \int_0^{t_1} [f_x(\cdot) + \lambda(t) g_x(\cdot) + \dot{\lambda}(t)] x_\varepsilon(t, \varepsilon) dt \\
& + \int_0^{t_1} [f_y(\cdot) + \lambda(t) g_y(\cdot)] \eta(t) dt \\
& - \lambda(t_1) x_\varepsilon(t_1, \varepsilon),
\end{aligned}
$$

where we have used the Leibniz Rule and the fact that $y_\varepsilon(t, \varepsilon) = \eta(t)$.

- Evaluate the previous derivative at $\varepsilon = 0$, the optimality requires that $W'(\varepsilon) = 0$ for all $\eta(t)$.
- This is possible in general if all the terms are zero

$$\dot{\lambda}(t) = -[f_x(\cdot) + \lambda(t)g_x(\cdot)], \tag{8}$$

$$f_y(\cdot) + \lambda(t)g_y(\cdot) = 0 \tag{9}$$

$$\lambda(t_1) = 0 \tag{10}$$

- The condition $\lambda(t_1) = 0$ is similar to the one looked at we in the finite discrete time problems.

Necessary conditions

- Consider the problem of maximizing (2) subject to (3) and (4), with f and g continuously differentiable. Suppose that this problem has an interior continuous solution $\hat{y}(t) \in int\mathcal{Y}$ with a corresponding path of state variable $\hat{x} \in int\mathcal{X}$.
- Then there exists a continuously differentiable costate function $\lambda(\cdot)$ defined on $t \in [0, t_1]$ such that (3), (8), (9) and (10) hold.

The Hamiltonian

- By analogy with the Lagrangian, there is a more economical way of expressing the previous optimality conditions.
- Let's construct the Hamiltonian:

$$H(t, x(t), y(t), \lambda(t)) \equiv f(t, x(t), y(t)) + \lambda(t) g(t, x(t), y(t)). \tag{11}$$

- The necessary conditions

$$H_y(t, \hat{x}(t), \hat{y}(t), \lambda(t)) = 0 \qquad \forall t \in [0, t_1], \tag{12}$$

$$\dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \qquad \forall t \in [0, t_1] \tag{13}$$

and

$$\dot{x}(t) = H_\lambda(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \qquad \forall t \in [0, t_1], \tag{14}$$

with $x(0) = x_0$ and $\lambda(t_1) = 0$.

- Notation: $\dot{x}$ means $\dot{\hat{x}}$.
- As with the discrete time optimization the solution is characterized by a set of "multipliers", $\lambda(t)$, and the optimal path of state, $\hat{x}(t)$ and control variables, $\hat{y}(t)$.
- Moreover, as with the Lagrange multipliers, the costate variable, $\lambda(t)$, is informative about the value of relaxing the constraint.
- With this interpretation, it makes sense that $\lambda(t_1) = 0$ is part of the solution.

Sufficient conditions

- Necessary conditions are also sufficient when either
  - Mangasarian's sufficiency conditions: $H(t, x(t), y(t), \lambda(t))$ is joint concave in $(x, y) \in \mathcal{X} \times \mathcal{Y}$ (if f and g are concave we are typically fine),
  - Arrow's sufficiency conditions: the maximized Hamiltonian, $\max_{y \in \mathcal{Y}} H(t, x(t), y(t), \lambda(t))$ is concave in x.

  hold.

Exponential discounting

- Typically in economics, the utility is discounted exponentially.

$$V(x_0) = \max_{x(t), y(t), x_1} \int_0^{t_1} e^{-\rho t} f(x(t), y(t)) dt \qquad \textit{with } \rho > 0, \tag{15}$$

st

$$\dot{x}(t) = g(x(t), y(t)) \tag{16}$$

$$x(t) \in \mathcal{X} \subset \mathbb{R}, y(t) \in \mathcal{Y} \subset \mathbb{R} \qquad \forall t,\ x(0) = x_0, x(t_1) = x_1 \tag{17}$$

- Note that, we also assumed autonomous differential equation for $g(\cdot)$.
- We often also need a terminal value condition $\lim_{t\to\infty} b(t)x(t) \geq x_1$ for some $x_1 \in \mathbb{R}$, where $\lim_{t\to\infty} b(t) < \infty$ (think about the no Ponzi-game condition).

- The Hamiltonian is

$$H(\cdot) = e^{-\rho t} f(x(t), y(t)) + \lambda(t) g(x(t), y(t)) = e^{-\rho t}[f(\cdot) + \mu(t) g(\cdot)] \tag{18}$$

- Rather than working with the standard Hamiltonian,$H(\cdot)$, we can work with the current-value Hamiltonian $\hat{H}$.

$$H(\cdot) = e^{-\rho t}[f(x(t), y(t)) + \mu(t) g(x(t), y(t))] = e^{-\rho t} \hat{H} \tag{19}$$

- Rewritten necessary conditions

$$\hat{H}_y(x(t), y(t), \mu(t)) = 0 \tag{20}$$

$$\rho\mu(t) - \dot{\mu}(t) = \hat{H}_x(x(t), y(t), \mu(t)) \tag{21}$$

$$\dot{x}(t) = \hat{H}_\mu(x(t), y(t), \mu(t)) \tag{22}$$

$$e^{-\rho t_1} \mu(t_1) = 0 \tag{23}$$

with $x(0)$ given

Discounted infinite-horizon problems

- In growth theory (and in Macro more generally), the time horizon is usually infinite.
- The previously stated optimality conditions (20)-(22) are still valid for the infinite-horizon problem.
- Condition (23) is, however, replaced with a transversality condition.
- Here we just state the problem typically encountered in economics contexts and give the ”cookbook” solution method.
- Under certain relatively mild technical conditions (that typically hold in the macro/growth context) it gives the necessary conditions for (interior) optima (see theorem 7.13 in Acemoglu’s book for details).

Infinite-horizon and the current value Hamiltonian

- The current value Hamiltonian is defined as before

$$\hat{H}(x(t), y(t), \mu(t)), = f(x(t), y(t)) + \mu(t) g(x(t), y(t)) \tag{24}$$

- The necessary conditions

$$\hat{H}_y(x(t), y(t), \mu(t)) = 0 \tag{25}$$

$$\rho\mu(t) - \dot{\mu}(t) = \hat{H}_x(x(t), y(t), \mu(t)) \tag{26}$$

$$\dot{x}(t) = \hat{H}_\mu(x(t), y(t), \mu(t)) \tag{27}$$

  for all t and

$$\lim_{T\to\infty} e^{-\rho T}\mu(T)x(T) = 0 \tag{28}$$

  with $x(0)$ given.
- Sufficiency: for any feasible pair $(x(t), y(t))$ $lim_{t\to\infty}[e^{-\rho t}\mu(t)x(t)] \geq 0$ and $\max_{y(t)\in\mathcal{Y}} \hat{H}(x(t), y(t), \mu(t))$ is concave in $x(t) \in \mathcal{X}$ for all t (see theorem 7.14 for technical details).

The Hamilton-Jacobi-Bellman equation

- Suppose that $(\hat{x}(t), \hat{y}(t))$ is a solution to

$$V(x(0)) = \max_{x(t), y(t)} \int_0^{\infty} e^{-\rho t} f(x(t), y(t)) dt \tag{29}$$

  subject to (16) and (17).
- Then,

$$V(x(0)) = \int_0^{t_1} e^{-\rho t} f(\hat{x}(t), \hat{y}(t)) dt + e^{-\rho t_1} V(\hat{x}(t_1)) \tag{30}$$

- Hamilton-Jacobi-Bellman equation

$$\rho V(\hat{x}(0)) = f(x(\hat{0}), \hat{y}(0)) + \dot{V}(\hat{x}(0)) \tag{31}$$

- Interpretation: no arbitrage asset value equation.

Heuristic derivation

- Start with the value function

$$\begin{aligned} V(x_0) &= \int_0^{\Delta t} e^{-\rho t} f(\hat{x}(t), \hat{y}(t)) dt + e^{-\rho \Delta t} \int_{\Delta t}^{\infty} e^{-\rho(t-\Delta t)} f(\hat{x}(t), \hat{y}(t)) dt \\ &= f(\hat{x}(0), \hat{y}(0)) \Delta t + o(\Delta t) + e^{-\rho \Delta t} V(\hat{x}(\Delta t)) \end{aligned}$$

  where we have used eq (30) and $o(\Delta t)$ is the residual from approximation $\int_0^{\Delta t} e^{-\rho t} f(\hat{x}, \hat{y}) \approx f(\hat{x}(0), \hat{y}(0)) \Delta t$
- The $o(\Delta t)$ term is second order in the sense that $\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$.

- Subtracting $V(\Delta t)$ from both sides and dividing both sides by Δt yields

$$\frac{V(\hat{x}(0)) - V(\hat{x}(\Delta t))}{\Delta t} = f(\hat{x}(0), \hat{y}(0)) + \frac{o(\Delta t)}{\Delta t} + \frac{e^{-\rho t_1} - 1}{\Delta t} V(\hat{x}(\Delta t)) \tag{32}$$

- Taking limits as $\Delta t \to 0$,

$$-\dot{V}(\hat{x}(0)) = f(\hat{x}(0), \hat{y}(0)) - \rho V(\hat{x}(0))$$

or

$$\rho V(\hat{x}(0)) = f(\hat{x}(0), \hat{y}(0)) + \dot{V}(\hat{x}(0)) \tag{33}$$