Relationship between Maximum Principle and Dynamic Programming

Toshihiko Mukoyama

August 2023

1 The saving problem

In this note, we demonstrate the relationship between continuous-time Maximum Principle (optimal control) and continuous-time Dynamic Programming using a simple saving problem. The problem is

$$\max_{c(t),a(t)} \int_0^\infty e^{-\rho t} u(c(t))dt$$

subject to

$$\dot{a}(t) = ra(t) - c(t),$$

where $c(t)$ is consumption at time t, $a(t)$ is asset at time t, $\rho > 0$ is the discount rate, $r > 0$ is the interest rate, and $u(\cdot)$ is an increasing and concave utility function. $\dot{a}(t) = da(t)/dt$ represents the time derivative.

2 Maximum Principle

For optimization using Maximum Principle, two alternative formulations can be used. The first method uses the *present-value Hamiltonian*:

$$H^p(t) \equiv e^{-\rho t} u(c(t)) + \lambda(t)(ra(t) - c(t)),$$

where $\lambda(t)$ is the costate variable. The first-order conditions for this formulation are

$$\frac{\partial H^p(t)}{\partial c(t)} = 0$$

and

$$\frac{\partial H^p(t)}{\partial a(t)} + \dot{\lambda}(t) = 0.$$

(The necessary condition for optimality includes the transversality condition, but we omit it here.)

The second uses the *current-value Hamiltonian*:

$$H^c(t) = e^{-\rho t} u(c(t)) + \mu(t)(ra(t) - c(t)),$$

where $\mu(t)$ is the costate variable in this case. The first-order conditions are

$$\frac{\partial H^c(t)}{\partial c(t)} = 0 \tag{1}$$

and

$$\frac{\partial H^c(t)}{\partial a(t)} + \dot{\mu}(t) - \rho\mu(t) = 0. \tag{2}$$

It is easy to check that both approaches are equivalent, with the relationship $\lambda(t) = e^{-\rho t}\mu(t)$. Therefore, we will work with the current-value Hamiltonian below. For the saving problem above, the first-order condition (1) is

$$u'(c(t)) = \mu(t) \tag{3}$$

and (2) is

$$r\mu(t) + \dot{\mu}(t) - \rho\mu(t) = 0. \tag{4}$$

3 Dynamic Programming

We will start by defining the value function

$$V(a(t)) = \int_t^\infty u(c^*(t))dt,$$

where $c^*(t)$ is the optimal value of $c(t)$. To (heuristically) derive the first-order condition, we start by considering a discrete-time formulation with period length Δ. Later we will take the limit of $\Delta \to 0$. The discrete-time Bellman equation is

$$V(a(t)) = \max_{c(t),a(t+\Delta)} u(c(t))\Delta + \frac{1}{1+\rho\Delta}V(a(t+\Delta)) \tag{5}$$

subject to

$$a(t+\Delta) = (1+r\Delta)a(t) - c(t)\Delta.$$

The first-order condition is

$$u'(c(t))\Delta = \frac{1}{1+\rho\Delta}V'(a(t+\Delta))\Delta.$$

Dividing both sides by Δ and take $\Delta \to 0$, we obtain

$$u'(c(t)) = V'(a(t)). \tag{6}$$

The Bellman equation (5) can be rewritten as (with optimally chosen $c(t)$ and $a(t)$)

$$\frac{V(a(t+\Delta)) - V(a(t))}{\Delta} + u(c(t)) - \frac{\rho}{1+\rho\Delta}V(a(t+\Delta)) = 0.$$

Taking the limit of $\Delta \to 0$,

$$V'(a(t))\dot{a}(t) + u(c(t)) - \rho V(a(t)) = 0. \tag{7}$$

This equation is the *Hamilton-Jacobi-Bellman (HJB) equation* and can be interpreted as an "asset equation," that is, rewriting

$$\rho V(a(t))dt = u(c(t))dt + V'(a(t))\dot{a}(t)dt,$$

the left-hand side is the required return of an asset that has the value of $V(a(t))$ over the time period dt, and the right-hand side is the income gain $u(c(t))dt$ plus the capital gain $V'(a(t))\dot{a}(t)dt$.

Replacing $\dot{a}(t)$ in (7) by $ra(t) - c(t)$, we obtain

$$rV'(a(t))a(t) - V'(a(t))c(t) + u(c(t)) - \rho V(a(t)) = 0.$$

This equation has to hold for all $a(t)$, and thus we can take the derivative with respect to $a(t)$ on both sides:

$$rV''(a(t))a(t) + rV'(a(t)) - V''(a(t))c(t) - V'(a(t))\frac{dc(t)}{da(t)} + u'(c(t))\frac{dc(t)}{da(t)} - \rho V'(a(t)) = 0.$$

Rearranging using $\dot{a}(t) = ra(t) - c(t)$ and (6),

$$rV'(a(t)) + V''(a(t))\dot{a}(t) - \rho V'(a(t)) = 0. \tag{8}$$

Now we can see the correspondence between the Maximum Principle equations ((3) and (4)) and the Dynamic Programming equations ((6) and (8)) with $\mu(t) = V'(a(t))$.