Lecture Notes 5: Continuous Time Optimal Control*

Ivan Rudik

ECON 509: Computational Methods

April 18, 2016

Many problems are better described in continuous time rather than finite time. For example, stock prices can change rapidly and massively in milliseconds. Physical climate systems are not discrete but actually change continuously. Here we will analyze deterministic continuous time optimal control.[1] When considering the optimal control of a deterministic dynamic system, we can solve the system as a sequence of many individual optimization problems: an open-loop solution. One alternative is to transform the problem into a form where it looks more like the static problems you are familiar with, where we will have differential equations that govern the optimal controls and trajectories.

1 The Fundamentals of Continuous Time Optimal Control

Consider a problem where each period an agent obtains flow utility $J(x(t), u(t))$, where x is our state and u is our control. Suppose there is a finite horizon with a terminal time T. The agent's objective is to maximize the total payoff, subject to the transitions of the states,

$$\max_{u,x_T} \int_0^T J(x(t), u(t))\, dt \tag{1}$$

$$\text{subject to: } \dot{x}(t) = g(x(t), u(t)),\ x(0) = x_0,\ x(T) = x_T$$

Dot notation indicates a time derivative. See that our control is a *function of time*, it is not simply a scalar value that is selected. This is an open-loop solution so we optimize our entire policy trajectory from time $t = 0$. Moreover, we are selecting a function that maximizes a function, so

*These notes are based on those of Derek Lemoine's.

[1]Stochastic continuous time problems are also common, but we will not discuss them in class.

our total payoff is a functional. This is the conventional set up of an optimal control problem.

Consider a thought experiment. Suppose we have a discrete time constrained dynamic optimization problem, and we want to solve it by selecting the sequence of controls, shadow values and states. We can form the Lagrangian for this problem as,

$$\max_{\{u_t,x_t,\lambda_t\}_{t=0}^T,\, a_T} \sum_{t=0}^{T} J(x_t,u_t) + \sum_{t=1}^{T} \lambda_t[g(x_{t-1},u_{t-1}) - (x_t - x_{t-1})] + \mu_0(a_0 - x_0) + \mu_T(a_T - x_T)$$

Where we can think of the transition equations as a constraint with their own shadow value, and of the initial and terminal conditions (now with notation a) as constraints as well. When solving this dynamic problem we have,

- T first-order conditions of the form $J_u(x_t,u_t) + \lambda_{t+1}\, g_u(x_t,u_t) = 0$ for $t \in [0, T-1]$
- T-1 first-order conditions of the form $J_x(x_t,u_t) + \lambda_{t+1}\, g_x(x_t,u_t) + \lambda_{t+1} - \lambda_t = 0$ for $t \in [1, T-1]$
- One first-order condition of the form $J_x(x_0,u_0) + \lambda_1\, g_x(x_0,u_0) + \lambda_1 - \mu_0 = 0$
- T first-order conditions of the form $g(x_{t-1},u_{t-1}) - (x_t - x_{t-1}) = 0$ for $t \in [1, T]$
- One first-order condition of the form $J_u(x_T,u_T) = 0$
- One first-order condition of the firm $J_x(x_T,u_T) - \lambda_T - \mu_T = 0$
- One first-order condition of the form $a_0 = x_0$
- One first-order condition of the form $a_T = x_T$
- One first-order condition of the form $\mu_T = 0$[2]

Three different sets of T conditions (the first four bullet points) tell us how the optimal controls, $u^*(t)$, evolve over time, how the shadow values evolve over time, and how the states evolve over time. Also, since the transition constraint is actually a definition, it does not pin down λ_t: λ_t is a free variable to be selected. Therefore we need to control λ_t to determine how the controls and states evolve over time. By shifting these conditions to continuous time, we can translate these conditions into differential equations. If we have a two-stated problem, we will need four more conditions to pin down a unique solution. Otherwise, we just know the optimal shape of the trajectory but not where it is located in the state-space. These conditions come from the last four bullet points. We need that the marginal payoff of increasing the control in the final period must

[2] This arises because there is no fixed, exogenous end point for the state.

be zero, we cannot leave any value on the table. We must also have that the marginal value of the state in the final period must equal its shadow cost, and the state variable must be pinned down at its initial and final points.

Consider the continuous time analog of the discrete time objective above,

$$\int_0^T J(x(t), u(t)) + \lambda(t)[g(x(t), u(t)) - \dot{x}(t)]dt + \mu_0(x_0 - x(0)) + \mu_T(x_T - x(T)),$$

and that with integration by parts, we have the following definition,

$$\int_0^T \lambda(t)\dot{x}(t)dt = \lambda(T)x(T) - \lambda(0)x(0) - \int_0^T \dot{\lambda}(t)x(t)dt.$$

If we substitute the definition into the Lagrangian we have that,

$$\int_0^T \left[J(x(t), u(t)) + \lambda(t)g(x(t), u(t)) + \dot{\lambda}(t)x(t)\right] dt - \lambda(T)x(T) + \lambda(0)x(0) + \mu_0(x_0 - x(0)) + \mu_T(x_T - x(T))$$

In a dynamic constrained optimization problem, we will have a function, similar to the Lagrangian (but not a dynamic Lagrangian!) for static problems, that yields the first-order conditions. This function is called the *Hamiltonian*. We will use the following definition,

$$H(x(t), u(t), \lambda(t)) \equiv J(x(t), u(t)) + \lambda(t)g(x(t), u(t)).$$

If we substitute the Hamiltonian into the continuous time problem we have,

$$\int_0^T [H(x(t), u(t), \lambda(t)) + \dot{\lambda}(t)x(t)]dt - \lambda(T)x(T) + \lambda(0)x(0) + \mu_0(x_0 - x(0)) + \mu_T(x_T - x(T))$$

If we recognize that $\dot{\lambda}(t) \approx \lambda_{t+1} - \lambda_t$ and $\dot{x}(t) \approx x_{t+1} - x_t$, we can recover something similar to the first-order conditions for the large discrete time problem by requiring that $\partial H/\partial u = 0$, $\partial H/\partial x + \dot{\lambda}(t) = 0$, and $\partial H/\partial \lambda - \dot{x}(t) = 0$. Indeed, Pontryagin's Maximum Principle states that the following conditions are necessary for an optimal solution to problem (1),

$$\frac{\partial H(x(t), u(t), \lambda(t))}{\partial u} = 0 \; \forall t \in [0, T] \qquad \text{(Maximality condition)}$$
$$\frac{\partial H(x(t), u(t), \lambda(t))}{\partial x} = -\dot{\lambda}(t) \qquad \text{(Co-state condition)}$$
$$\frac{\partial H(x(t), u(t), \lambda(t))}{\partial \lambda} = \dot{x}(t) \qquad \text{(State transitions)}$$
$$x(0) = x_0 \qquad \text{(Initial condition)}$$
$$\lambda(T) = 0 \qquad \text{(Transversality condition)}$$

What do these conditions mean? First, we must know that the Hamiltonian yields the contribution of that instant t to overall utility via the change in flow utility and the change in the state (which affects future flow utilities). The decisionmaker can use her control to increase the contemporaneous flow utility and reap immediate rewards, or to alter the state variable to increase future rewards. The maximality condition states that in every instant, we select the control so that we can no longer increase our total payoff. This is a very familiar expression and effectively sets the *net* marginal benefits of the control to zero. The state transition expression is also very natural since it is effectively just a definition. λ is the shadow value of our of our state transition, or transition constraint. Taking the derivative of the Hamiltonian with respect to the shadow value, just like a Lagrangian, yields this constraint back.[3]

This leaves what is called the co-state condition. This defines how the shadow value of our state transition, called the *co-state variable*, evolves over time. The co-state variable tells us the additional future value of having one more unit of our state variable (e.g. CO_2, knowledge, capital). Suppose we increase today's stock of x by one unit and this increases the instantaneous change in our value (H). Then the shadow value of that stock (λ) must decrease! Why? Because the state transition constraint $\dot{x} = g(\cdot)$, is now slightly more slack. We can re-write the co-state equation as,

$$\frac{\partial J}{\partial x} + \lambda(t)\frac{\partial g}{\partial x} + \dot{\lambda}(t) = 0,$$

so we can see that there are three effects that must precisely cancel out along an optimal path. We must have that the a unit of the stock's value must change (third term), so that is exactly offsets the change in value from increasing the stock in the immediate instant of time. The immediate value is made up of the actual utility payoff (first term), and the future utility payoff payoff from

[3] Do not confuse a Hamiltonian as being the dynamic version of a Lagrangian. If there are constraints on controls a dynamic Lagrangian can be formed which is composed of the Hamiltonian and the constraints.

how increasing the stock today affects the stock in the future (second term).[4]

These three conditions tell us how the controls, states, and shadow values must evolve along the optimal path. They give us the shape of the optimal path but they do not tell us what the optimal path *is*. Many different paths are consistent with these differential equations, i.e. under some conditions we may follow the three conditions and achieve a capital stock of 20 at $t = 50$, but other conditions may achieve a capital stock of 26 at $t = 50$. These three conditions are just a set of ordinary differential equations. Therefore the solution, without any additional information, yields an improper integral.

We will need additional constraints in order to select the optimal path from the family of possible paths. Technically, we are just pinning down the constants of integration. The three differential equations can be reduced to two differential equation, so we will need two conditions to pin them down. Our first condition is the initial condition. This pins down where the state path is. We also need to pin down our co-state path. This is what our transversality condition does. The transversality condition recognizes that along an optimal path, small changes in the stock cannot change the value of the problem. At the terminal time, if this value was positive, then we are leaving value on the table and are clearly not on an optimal path.

In general there are four types of transversality conditions.[5] Two are for pinning down the initial or terminal time if they are free (e.g. if T was not set exogenously, but something to be selected by the agent), and two are for pinning down the initial or terminal state variables if they are free. Typically initial conditions are exogenously given, which are the effective transversality conditions for those end points. Terminal conditions however, are often free. If the agent can select at which time T the program ends (e.g. see Lemoine and Rudik (2014) for a setting where the policymaker does so), then at that time T the Hamiltonian must be zero. Recall the Hamiltonian give the instantaneous change in value. If this were positive, then the policymaker could profitably deviate by ending the program at some future time $\bar{T} > T$. Similarly, if the terminal state is free, its shadow value (change in value from one more unit) must be zero. If it were positive the policymaker could deviate by altering the level of the stock. Finally, these are all necessary conditions of the problem.

In the basic optimal control problem we discount the future using exponential discounting.[6] Now time can directly affect value. Assume that time does not directly affect instantaneous payoffs or the transitions equations. Then our value is $J(x(t), u(t), t) = e^{-rt} V(x(t), u(t))$. J yields the present, time 0 value of the change in value at time t. J is the *present value* and V is the *current value*. Present value refers to the value with respect to a specific period that we call the present.

[4]This has the flavor of an Euler equation, but where we are balancing the marginal benefits of the stock over time.

[5]See Caputo (2005) for a nice description.

[6]This is the continuous time analog of having a constant discount factor.

Current value is continually being updated as time moves along. Our previous necessary conditions apply to present value Hamiltonians, but let us analyze a current value Hamiltonian to avoid including time terms,

$$H^{cv}(x(t), u(t), \mu(t)) \equiv e^{rt} H(x(t), u(t), \lambda(t), t) = e^{rt} J(x(t), u(t), t) + e^{rt} \lambda(t) g(x(t), u(t)), \tag{CV Hamiltonian}$$

where $\mu(t)$ is the shadow value λ brought into current value terms: $\mu(t) = e^{rt}\lambda(t)$. We can then re-write our necessary conditions in current value by substituting in for the shadow value (which implies that $\dot{\lambda}(t) = -re^{-rt}\mu(t) + e^{-rt}\dot{\mu}(t)$), and for $\partial H/\partial x = e^{-rt}\,\partial H^{cv}/\partial x$ into our co-state condition,

$$e^{-rt}\frac{\partial H^{cv}(x(t), u(t), \mu(t))}{\partial x} = e^{-rt}\left[r\mu(t) - \dot{\mu}(t)\right] \tag{CV Co-state condition}$$

Recall before that the present value[7] form of the co-state condition required that the change in the present shadow value precisely equal the effect of the state variable on instantaneous value. In its current value form, the co-state condition recognizes that the change in the present shadow value is comprised of two parts. The change in the current shadow value, but also the reduction in present value purely from the passage of time. The current value co-state condition must adjust for this depreciation. If discounting is high (large r), then the current shadow value must change quicker in order to compensate the policymaker for leaving stock for the future.

Just like with discrete time problems, we can obtain an Euler equation which describes the optimal trajectory. The way to obtain the Euler equation is by taking two steps:

1. Differentiate the maximality condition with respect to time. This will yield a condition for the time derivative of our co-state variable.

2. Solve the maximality condition and its time derivative for the co-state and the time derivative of the co-state. Plug this into the co-state condition.

This tells us how our control variable evolves over time, while our transition equation tells us how our state variable evolves over time.

2 Applications: Hotelling Extraction

The Hotelling model has conventionally been solved in continuous time. Consider an oil extracting firm with a known stock $x(t)$ of oil at time t, and it decides on how much to extract, $q(t)$ at time t

[7] The previous derivation of the conditions was in both present and current value since it was not discounted.

to maximize the discounted stream of profits over some time horizon T. The initial amount in the oil reservoir is $x_0 > 0$. The stock at any point in time evolves according to,

$$x(t) = x_0 - \int_0^t q(s)ds \Rightarrow \dot{x}(t) = -q(t).$$

At time T, the firm can decide how much oil to leave in the ground, so x_T is a free variable, decided by the firm. However, there is no value to leaving oil in the ground. The firm has a profit function $\pi(q(t), x(t))$ that depends on the amount of extraction and how much oil is left in the ground, $\pi_q(q, x) > 0$, $\pi_x(q, x) > 0$ (oil is more costly to extract as the reservoir is depleted). Assume that the marginal cost of extraction is also increasing, $\pi_{qq}(q, x) < 0$. The discount rate is r. The firm's optimal control problem is,

$$\begin{aligned} &\max_{q(t), x_T} \int_0^T \pi(q(t), x(t))e^{-rt}dt \\ s.t. \quad &\dot{x}(t) = -q(t), x(0) = x_0, x(T) = x_T \end{aligned}$$

Ignoring constraints for ease of exposition, the firm's Hamiltonian is,

$$H(t, x, q, \lambda) = \pi(q, x)e^{-rt} - \lambda_q \tag{2}$$

The necessary conditions for an optimal trajectory are,

$$\begin{aligned} H_q(t, x, q, \lambda) = \pi_q(q, x)e^{-rt} - \lambda = 0 &\qquad \text{(maximality condition)} \\ -\pi_x(q, x)e^{-rt} = \dot{\lambda}(t) &\qquad \text{(co-state condition)} \\ \dot{x}(t) = -q(t) &\qquad \text{(state transition)} \\ x(0) = x_0 &\qquad \text{(initial condition)} \\ \lambda(T) = 0 &\qquad \text{(transversality condition)} \end{aligned}$$

$\lambda(t)$ yields the present value of the shadow price of the oil at time t. Along an optimal path, the value of any remaining oil in the reservoir is worthless: there's no scrap value associated with it (otherwise we would have a scrap value function at the right hand side). The maximality condition tells us that along an optimal path, the present value of the marginal profit of extraction must equal the shadow price: the firm must be indifferent between extracting another unit (increasing profit at time t) or leaving it in the ground (with value $\lambda(t)$ from future extraction).

Let us consider a current value expression. Define $\mu(t) = e^{rt}\lambda(t)$ as the current value shadow price, and that $\dot{\mu}(t) = r\lambda(t)e^{rt} + \dot{\lambda}(t)e^{rt}$. If we substitute in the definition of $\mu(t)$ and the co-state condition above, we have that,

$$\frac{\dot{\mu}(t)}{\mu(t)} = r - \frac{\pi_x(q(t), x(t))}{\mu(t)} < r \qquad \text{(generalized Hotelling rule)}$$

If the extraction cost depends on the stock, the shadow price of the oil can be rising or falling. Just like with Euler equations before, we can intepret this as a no-arbitrage condition. The firm can extract an additional unit of oil today and invest the revenues to earn $(1 + r)\pi_q(t) = (1 + r)\mu(t)$ tomorrow, but by depleting the stock, the firm loses π_x tomorrow. On the other hand, the firm can leave that unit of oil in the ground and extract tomorrow, yielding a marginal profit tomorrow of $\pi_q(t + 1) = \mu(t + 1)$. Along an optimal trajectory the firm cannot arbitrage so we must have that $\mu(t + 1) = (1 + r)\mu(t) - \pi_x$. We can rearrange this expression to obtain the generalized Hotelling rule above. If $\pi_x(q, x) = 0$, then the right hand side simplifies to just r. This is the conventional Hotelling rule stating that the shadow price, and thus marginal profit, must rise at the rate of interest. The firm must be indifferent between extracting today or tomorrow, so profits must rise at the rate of interest to offset gains from extracting today and investing at rate r.

3 Numerical Solutions

How do we solve these problems? The equations that describe the trajectories of the system are just ODEs, so the simple answer is to just integrate them. However, we still need to determine where these trajectories are, i.e. pin them down with the transversality conditions. Most languages have ODE solvers that can integrate systems of equations.

Returning to the Hotelling problem above, it seems like we can simply plug into a solver: the equations of motion, the control trajectory which we can back out from the maximality condition, $x(0) = x_0$, and then just search over $\lambda(0)$ to find a value where $\lambda(T) = 0$. Then all of our necessary conditions are satisfied and if the problem is concave, we have found a unique solution. The problem with unique solutions is that they lie on saddle paths: any initial guess of $\lambda(0)$ off our saddle path is repelled away from the optimal trajectory, the stable manifold is infinitely thin. In other words, the values of the terminal conditions we calculate are incredibly sensitive to the initial values we select. However this gives us something we can exploit: if terminal conditions are very sensitive to initial conditions, then the initial conditions must *not* be sensitive to the terminal conditions. If we reverse time when integrating out the problem, it will become much more stable. Indeed, this is the most common way to solve continuous time optimal control problems under the methods of backwards integration or reverse shooting (Judd, 1998).

References

Caputo, Michael Ralph (2005) *Foundations of dynamic economic analysis: optimal control theory and applications*: Cambridge University Press.

Judd, Kenneth L. (1998) *Numerical Methods in Economics*, Cambridge, MA: MIT Press.

Lemoine, Derek and Ivan Rudik (2014) "Steering the Climate System: Using Inertia to Lower the Cost of Policy," *University of Arizona Working Paper 14-03*.