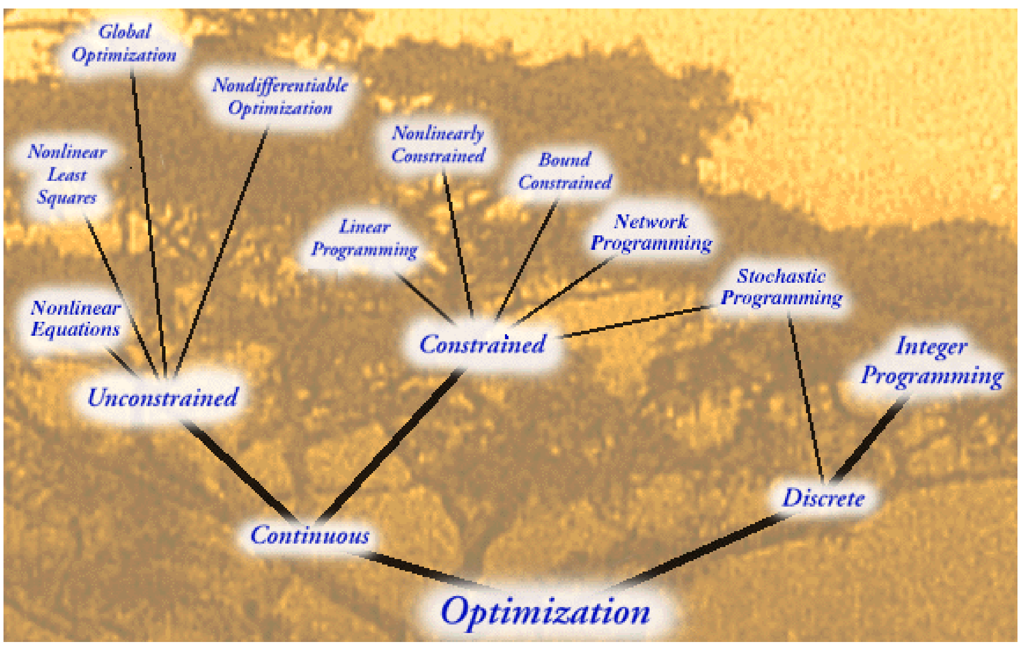
1. An Introduction to Dynamic Optimization -- Optimal Control and Dynamic Programming

AGEC 642 - 2024

I. Overview of optimization

Optimization is a unifying paradigm in most economic analysis. So, before we start, let's think about optimization. The tree below provides a nice general representation of the range of optimization problems that you might encounter. There are two things to take from this. First, all optimization problems have a great deal in common: an objective function, constraints, and choice variables. Second, there are lots of different types of optimization problems and how you solve them will depend on the branch on which you find yourself.

In terms of the entire tree of all optimization problems, the ones that could be solved analytically would represent a couple of leaves at best – numerical methods must be used to solve the rest. Fortunately, a great deal can be learned about economics by studying those problems that can be solved analytically.



Source: The Optimization Technology Center: http://www.ece.northwestern.edu/OTC/ (very old!)

In this course we will use both analytical and numerical methods to solve dynamic optimization problems, problems that have two common features: the objective function is a linear aggregation over time, and a set of variables called the state variables are constrained across time. And so, we begin …

II. Introduction – A simple 2-period consumption model

Consider the simple consumer's optimization problem:

$$\max_{z} u(z_a, z_b) \ s.t.$$

$$p_a z_a + p_b z_b \leq x$$

[pay attention to the notation: z is the vector of choice variables and x is the consumer's exogenously determined income. This use of z and x will be used throughout this course.]

Solving the one-period problem should be familiar to you. Now consider what happens if the consumer lives for two periods but must survive off the income endowment provided at the beginning of the first period. That is, what happens if her problem is

$$\max_{z} U\left(z_{1a}, z_{1b}, z_{2a}, z_{2b}\right) = U\left(z_1, z_2\right) \quad \text{s.t. } \mathbf{p}' z_1 + \mathbf{p}' z_2 \leq x_1,$$

where the constraint uses matrix notation with $\mathbf{p} = \left[p_a, p_b\right]$ refers to a price vector and $z_1 = \left[z_{1a}, z_{1b}\right]$? We now have a problem of *dynamic* optimization. When we choose z_1, we must consider how it will affect our choices in period 2.

We are going to make a **huge**[1] (though common) assumption and maintain that assumption throughout the course: utility is additively separable across time:

$$u(\mathbf{z}) = u(z_1) + u(z_2) \ .$$

Clearly one way to solve this problem would be just as we would a standard static problem: set up a Lagrangian and take four first-order conditions for the z variables and a fifth for the Lagrange multiplier, then solve for all optimal choices simultaneously. This may work here, when there are only 2 periods, but if we have 100 periods (or even an infinite number of periods) then this would get really messy. This course will develop methods to solve such problems.

This is a good point to introduce some *very important* terminology:

- All dynamic optimization problems have a **time step** and a **time horizon**. In the problem above time is indexed with t. The time step is 1 period, and the time horizon is from 1 to 2, i.e., t={1,2}. However, the time step can also be continuous, so that t takes on every value between t_0 and T, and we can even solve problems where $T \to \infty$.
- x_t is what we call a **state variable** because it is *the state that the decision-maker faces in period t*. Note that x_t is parametric (i.e., it is taken as given) to the decision-maker's problem in t, and x_{t+1} is parametric to the choices in period t+1. However, x_{t+1} is affected by the choices made in t. The state variables in a problem are those that a decision maker takes as given when making his or her choices in each period, but future values are either determined by current choices or unknown at t.
- A **state equation** defines the intertemporal changes in a state variable. This equation is sometimes referred to as the equation of motion or the transition equation.

[1] See Deaton and Muellbauer (137-142) on the negative implications of assuming preferences are additive.

- z_t is the vector of t^{th} period **choice variables**. Choice variables determine the (expected) payoff in the current period and the (expected) state next period. These variables are also referred to as **control or action variables** and I will use all these terms interchangeably.

  To distinguish state & control variables, I like to say, "You wake up in the morning, look at your state variables, make decisions about your control variables, then go back to sleep."

- p_a and p_b are **parameters** of the model. They are held constant or change exogenously and deterministically over time. [Note that the difference between parameters and state variables is subtle; parameters are not controllable at all, state variables are not controllable at time t but decisions at t can affect future values].

- Finally, we have what I call **intermediate variables**. These are variables that are actually *functions* of the state and control variables and the parameters. For example, in the problem considered here, the utility at a moment in time might be carried as an intermediate variable for analytical convenience when you solve the problem. In firm problems, production might be an intermediate variable. When you formulate a problem, it is very important, but often difficult, to distinguish intermediate variables so that you do not treat them incorrectly (see PS#1). Intermediate variables are never critical to the solution, i.e., the problem could be specified and solved without any intermediate variables.

- The **benefit function** tells the instantaneous or single period ***net*** benefits that accrue to the planner during the planning horizon. In our problem $u(z_t)$ is the benefit function. The benefit function should probably be called the net benefit function (benefits minus costs) and can be positive or negative. For example, in a problem in which the goal is to minimize the costs over the time horizon, the cost in each period would be the (net) benefit function.

- In many problems there are benefits (or costs) that accrue after the planning horizon. This is captured in models by including a **salvage value**, which is usually a function of the terminal stock. Since the salvage value occurs after the planning horizon, it cannot be a function of the control variables, though it can be a separate optimization problem in which choices are made.

- The sum (or integral) over the planning horizon plus the salvage value determines the **objective function**. We usually use discounting when we sum up over time. Pay close attention to this – the objective function is ***not*** the same as the benefit function.

- All the problems that we will study in this course fall into the general category of **Markov decision processes** (MDP). In an MDP the probability distribution over the states in the next period is wholly determined by the current state and current actions. One important implication of limiting ourselves to MDPs is that, typically, history does not matter, i.e., x_{t+1} depends on z_t and x_t, irrespective of the value of x_{t-1}. When history is important in a problem, then the relevant historical variables must be explicitly included as state variables. We will consider stochastic MDPs in much more depth in Lecture 9.

- A **Formal Statement of the Optimization Problem** is a set of mathematical expressions including the objective function and all the constraints. The constraints include the state equation(s), any conditions that must be satisfied at the beginning and end of the time horizon, and any constraints that restrict choices between the beginning and end. At a minimum, dynamic optimization problems must include the objective function, the state equation(s) and initial conditions for the state variables.

In sum, the problems that we will study will have the following features. In each period or moment in time the decision maker takes as given the state variables and parameters, then makes optimal choices for the control variables considering the objective function and state equations. Instead of using brute force to the solve for the optimal value of all the z's in the two-period consumer problem in one step, we reformulate the problem. Let x_1 be the endowment, which is available in period 1, and x_2 be the endowment that remains in period 2. Following from the budget constraint, we can see that $x_2 = x_1 - \mathbf{p}'z_1$, with $x_2 \geq 0$. In this problem x_2 defines the **state** that the decision maker faces at the start of period 2. The equation which describes the change in the x from period 1 to period 2, $x_2 - x_1 = -\mathbf{p}'z_1$, is the **state equation**.

We now rewrite our consumer's problem, this time making use of the state equation:

$$\max_{z_t} \sum_{t=1}^{2} u_t(z_t) \quad s.t.$$

$$\left.\begin{array}{l} x_{t+1} - x_t = -\mathbf{p}' z_t \\ x_{t+1} \geq 0 \end{array}\right\} t = 1,2 \tag{1}$$

$$x_1 \text{ fixed}$$

We now have a nasty little optimization problem with four constraints, two of them inequality constraints. Not fun. This course will help you solve and understand these kinds of problems. Note that while (1) looks complicated, it is quite general since you could easily write the T-period problem by simply replacing the 2's with T.

III. The OC (optimal control) way of solving the problem

We will solve dynamic optimization problems using two related methods. The first of these is called optimal control. Optimal control makes use of Pontryagin's maximum principle.

First, note that for most specifications, economic intuition tells us that $x_2>0$ and $x_3=0$. Hence, for $t=1$ ($t+1=2$), we can suppress inequality constraint in (1). We'll use the fact that $x_3=0$ at the very end to solve the problem.

Write out the Lagrangian of (1):

$$L = \sum_{t=1}^{2} \left[u_t(z_t, x_t) + \lambda_t \left(x_t - x_{t+1} - \mathbf{p}' z_t \right) \right] \tag{2}$$

where we include x_t in $u(\cdot)$ for completeness, though in this case $\partial u / \partial x = 0$.

Sometimes it is helpful to write out all the terms of a summation, to make sure that you know what is being said. In this case

$$\begin{aligned} L &= \sum_{t=1}^{2}\left[u_t\left(z_t, x_t\right) + \lambda_t\left(x_t - x_{t+1} - \mathbf{p}' z_t\right)\right] \\ &= \left[u_1\left(z_1, x_1\right) + \lambda_1\left(x_1 - x_2 - \mathbf{p}' z_1\right)\right] + \left[u_2\left(z_2, x_2\right) + \lambda_2\left(x_2 - x_3 - \mathbf{p}' z_2\right)\right]. \end{aligned}$$

More terminology

In optimal control theory, the variable λ_t is called the **costate variable**. Following the standard interpretation of Lagrange multipliers, at its optimal value λ_t is equal to the marginal value of relaxing the constraint. In this case, that means that λ_t is equal to the marginal value of the state variable, x_t. The costate variable plays a critical role in dynamic optimization and has important economic meaning.

The first-order conditions (FOCs) for (2) are standard:

$$\partial L/\partial z_{ti} = \partial u/\partial z_{ti} - \lambda_t p_i = 0, \quad i = a, b, \quad t = 1, 2$$

$$\partial L/\partial x_2 = \frac{\partial u}{\partial x_2} - \lambda_1 + \lambda_2 = 0$$

[note that x_1 is not a choice variable since it is fixed at the outset and x_3 is equal to zero]

$$\partial L/\partial \lambda_t = \left(x_t - x_{t+1} - \mathbf{p}' z_t\right) = 0 \ , \ t=1,2.$$

We now use a little notational change that simplifies this problem and adds some intuition (we'll see how the intuition arises in later lectures). That is, we define a function known as the **Hamiltonian** where

$$H\left(z_t, x_t, \lambda_t\right) = u\left(z_t, x_t\right) + \lambda_t\left(-\mathbf{p}' z_t\right).$$

Some things to note about the Hamiltonian:

- the t^{th} Hamiltonian only includes current variables: z_t, x_t and λ_t,
- unlike in a Lagrangian, only the right-hand side of state equation appears after λ_t.

In the left column of the table below we present the familiar FOCs of the Lagrangian. On the right we present the derivative of the Hamiltonian with respect to the same variables. Comparing the two sides, we can see what we need to put on the right-hand side of the derivatives of the Hamiltonian to obtain the same result as when using the Lagrangian.

Lagrangian		Hamiltonian
$L=\sum_{t=1}^{2}\left[u_t(z_{ta},z_{tb})+\lambda_t(x_t-x_{t+1}-(p_a z_{ta}+p_b z_{tb}))\right]$		$H=u\left(z_t,x_g\right)+\lambda_t\left(-\mathbf{p}'z_t\right)$
Standard FOCs		**$\partial H/\partial$__**
$\frac{\partial L}{\partial z_{ti}}=\frac{\partial u_t}{\partial z_{ti}}-\lambda_t p_i=0$, t=1,2, i=a,b	z_{ti}	$\frac{\partial H}{\partial z_{ti}}=\frac{\partial u_t}{\partial z_{ti}}-\lambda_t p=0$
$\frac{\partial L}{\partial x_2}=\frac{\partial u(\cdot)}{\partial x_2}-\lambda_1+\lambda_2=0$	x_2	$\frac{\partial H}{\partial x_2}=\frac{\partial u\left(z_2,x_2\right)}{\partial x_2}=\lambda_2-\lambda_1$
$\frac{\partial L}{\partial \lambda_t}=x_t-x_{t+1}-\mathbf{p}'z_t=0$, t=1,2, i=a,b	λ_t	$\frac{\partial H}{\partial \lambda_t}=-\mathbf{p}'z_t=x_2-x_1$

Hence, we see that for the solution using the Hamiltonian to yield the same maximum the following conditions must hold:

1. $\frac{\partial H}{\partial z_t}=0$ $\Rightarrow$ The Hamiltonian should be maximized w.r.t. the control variable at every point in time.
2. $\frac{\partial H}{\partial x_t}=\lambda_{t-1}-\lambda_t$ for t>1 $\Rightarrow$ The costate variable changes over time at a rate equal to minus the marginal value of the state variable to the Hamiltonian.
3. $\frac{\partial H}{\partial \lambda_t}=x_{t+1}-x_t$ $\Rightarrow$ The state equation must always be satisfied.

When we combine these with a 4th condition, called the *transversality condition* (how we *transverse* over to the world beyond t=1,2) we are able to solve the problem. In this case the condition that x_3 =0 (which for now we will assume to hold without proof) serves as the transversality condition. We will discuss the transversality condition in more detail in a few lectures.

These four conditions are the starting points for solving most optimal control problems and sometimes the FOCs alone are sufficient to understand the economics of a problem. However, if we want an explicit solution, then we would solve this system of equations.

In this class most of the OC problems we will face are in continuous time. The parallels between the discrete time case presented here and the continuous time case should be obvious when we get there.

IV. The DP (Dynamic Programming) way of solving the problem

The second way that we will solve dynamic optimization problems is using Dynamic Programming. DP is about **backward induction** – thinking backwards about problems. Let's see how this is applied in the context of the 2-period consumer's problem.

Imagine that the decision-maker in our consumer choice problem is in period 2, having already used up part of her endowment in period 1, leaving x_2 to be spent. In period 2, her problem is simple:

$$V_2\left(x_2\right)=\max_{z_2} u_2\left(z_2\right) \quad \text{s.t. } \mathbf{p}' z_2 \le x_2 .$$

If we solve this problem, we can easily obtain the function $V(x_2)$, which tells us the maximum utility that can be obtained if she arrives in period 2 with x_2 dollars remaining. The function $V(\cdot)$ is equivalent to the indirect utility function with p_a and p_b suppressed. The period 1 problem can then be written

$$V_1\left(x_1\right)=\max_{z_1} u\left(z_1\right)+V_2\left(x_2\right) \;\; s.t. \;\; x_2=x_1-\mathbf{p}' z_1 . \tag{3}$$

The value of having x_1 in period one is the solution to this problem. This equation is known as a Bellman's equation, and it is the cornerstone of dynamic programming.

Note that we have implicitly assumed an interior solution so that the constraint requiring that $x_3 \ge 0$ is assumed to hold with an equality and can be suppressed.

Once we know the functional form of $V(\cdot)$, (3) becomes a simple static optimization problem and its solution is straightforward. If the functional form of $V(x_2)$ has been found, then we can use the state equation to eliminate x_2 to obtain:

$$V_1\left(x_1\right)=\max_{z_1} u\left(z_1\right)+V_2\left(x_1-\mathbf{p}' z_1\right).$$

A major challenge with the DP approach, however, is that we do *not* have *a priori* a functional form for $V(\cdot)$, and as problems become more complicated, obtaining a functional form becomes more difficult, even impossible for many problems. Hence, the trick to solving DP problems is to find or approximate the function $V(\cdot)$.

V. Are OC and DP equivalent? Yes.[2]

As we will see throughout this course, either of these approaches can be used to solve a dynamic optimization problem. In this section we will quickly show that the first-order conditions for a simple problem are equivalent.

Consider the continuous-time dynamic optimization problem,

$$\max_{z_t} \int_{t=0}^{T} u\left(x_t, z_t\right) dt \;\; \text{s.t. } \; \dot{x}_t=f\left(x_t, z_t\right),$$

where, as we will discuss in Lecture 2, $\dot{x}_t \equiv \partial x_t / \partial t$. The discrete-time analog of this problem is

$$\max_{z_j} \sum_{j=0}^{T/\Delta} u\left(x_j, z_j\right)\Delta \;\; \text{s.t. } \; x_{t+\Delta}-x_t=\Delta f\left(x_t, z_t\right),$$

where Δ is some fraction of a period and $u\left(x_j, z_j\right)$ is the *rate* at which utility is generated per full period. For example, if Δ=0.5, then there are two increments per period that go from j=0 to j=2T=T/0.5 and in each of these intervals the utility obtained is $\Delta \cdot u\left(x_j, z_j\right)$.

The Hamiltonian and Bellman equations for these two problems are:

---

[2] This discussion may be difficult to follow based only on the discussion above. After reading later lectures, especially Lectures 3 and 5, you should review this section.

$$H\left(x_t, z_t, \lambda_t\right) = \Delta u\left(x_t, z_t\right) + \lambda_t \Delta f\left(x_t, z_t\right), \text{ and} \tag{4}$$

$$V\left(x_t, t\right) = \max_{z_t} \Delta \cdot u\left(x_t, z_t\right) + V\left(x_{t+\Delta}, t+\Delta\right) \text{ s.t. } x_{t+\Delta} = x_t + \Delta \cdot f\left(x_t, z_t\right). \tag{5}$$

Equivalence for optimal choice

First, we show the equivalence of the FOCs w.r.t the control variable as $\Delta \rightarrow 0$. The first order condition for the Hamiltonian is, as above, $\frac{\partial H}{\partial z_t} = \Delta u_z + \lambda_t \Delta f_z = 0$. Dividing by Δ, this yields $\frac{\partial H}{\partial z_t} = u_z + \lambda_t f_z = 0$.

For the Bellman's equation, we know that at the optimum value for z, $\partial V / \partial z = 0$, i.e. $\frac{\partial V}{\partial z_t} = \Delta \cdot u_z + \frac{\partial V\left(x_{t+\Delta}, t+\Delta\right)}{\partial x_{t+\Delta}}\left(\Delta \cdot f_z\right) = 0$. Dividing by Δ and taking the limit as $\Delta \rightarrow 0$ so that $t+\Delta \rightarrow t$, we have $u_z + \frac{\partial V\left(x_t, t\right)}{\partial x_t} f_z = 0$.

Now, recall that λ_t is a shadow price, its economic interpretation is the marginal value of relaxing the state equation constraint or a marginal increase in x_t, i.e., $\lambda_t = \frac{\partial V\left(x_t, t\right)}{\partial x_t}$.

Hence, it follows that $\lim_{\Delta \to 0} \frac{\partial V}{\partial z_t} = \frac{\partial H}{\partial z_t} = 0$, so the FOC's with respect to z of the optimal control and dynamic programming specifications are equivalent.

Equivalence for optimal value of future *x*

Next, we can show the equivalence of the FOC w.r.t. the state variable, x_t, as $\Delta \rightarrow 0$. Again, we stated above that the FOC of the Hamiltonian for the state variable is

$$\frac{\partial H}{\partial x_t} = \Delta \frac{\partial u}{\partial x_t} + \lambda_t \Delta \frac{\partial f}{\partial x_t} = \lambda_t - \lambda_{t+\Delta}. \tag{6}$$

We can then divide the middle and last part of this equality by Δ to obtain

$$\frac{\partial u}{\partial x_t} + \lambda_t \frac{\partial f}{\partial x_t} = \frac{\lambda_t - \lambda_{t+\Delta}}{\Delta}$$

and then take the limit as $\Delta \rightarrow 0$ so that, using the definition of a partial derivative, this becomes

$$\frac{\partial u}{\partial x_t} + \lambda_t \frac{\partial f}{\partial x_t} = -\frac{\partial \lambda_t}{\partial t}. \tag{7}$$

For the Bellman's equation, x_t is not a choice variable at time t – it is fixed at time t. However, we can take the derivative of the Bellman's equation (5) with respect to x_t using the chain rule to obtain:

$$\frac{\partial V\left(x_t, t\right)}{\partial x_t} = \Delta \frac{\partial u}{\partial x_t} + \frac{\partial V\left(x_{t+\Delta}, t+\Delta\right)}{\partial x_{t+\Delta}} \frac{\partial x_{t+\Delta}}{\partial x_t}. \tag{8}$$

Notice that there is some nice intuition in (8): the marginal value of the state variable is equal to the sum of what you get out of it in the first interval of length Δ, $\Delta \cdot u_x$, plus what you get in the future because you have more x_t: $\frac{\partial V^{t+\Delta}}{\partial x_{t+\Delta}} \frac{\partial x_{t+\Delta}}{\partial x_t} = \lambda_{t+\Delta}\left(1 + f_x \Delta\right)$, where we are using the equation $x_{t+\Delta} = x_t + \Delta f\left(x_t, z_t\right)$. Hence, (8) can be rewritten $\lambda_t = \Delta u_x + \lambda_{t+\Delta}\left(1 + f_x \Delta\right)$, subtracting $\lambda_{t+\Delta}$ from both sides and dividing by Δ we obtain $\frac{\lambda_t - \lambda_{t+\Delta}}{\Delta} = u_x + \lambda_{t+\Delta}\left(f_x\right)$. Again we take the limit at $\Delta \rightarrow 0$, which in this case gives us $-\frac{\partial \lambda}{\partial t}$ on the LHS, to obtain $-\frac{\partial \lambda}{\partial t} = u_x + \lambda_t f_x$, which is the same as FOC for the Hamiltonian, (7).

Finally, it is obvious that the state equation in both formulations must hold, regardless of the length of Δ. Hence, we have shown that the two approaches are equivalent.

VI. Summary

- OC problems are solved using the vehicle of the Hamiltonian, which must be maximized at each point in time.
- DP is about backward induction.
- Both techniques are equivalent to standard Lagrangian techniques and the interpretation of the shadow price, λ, is the same.

VII. Additional reading for next lecture

Leonard and Van Long, chapter 2.

VIII. References

Deaton, Angus and John Muellbauer. 1980. *Economics of Consumer Behavior.* New York: Cambridge University Press.