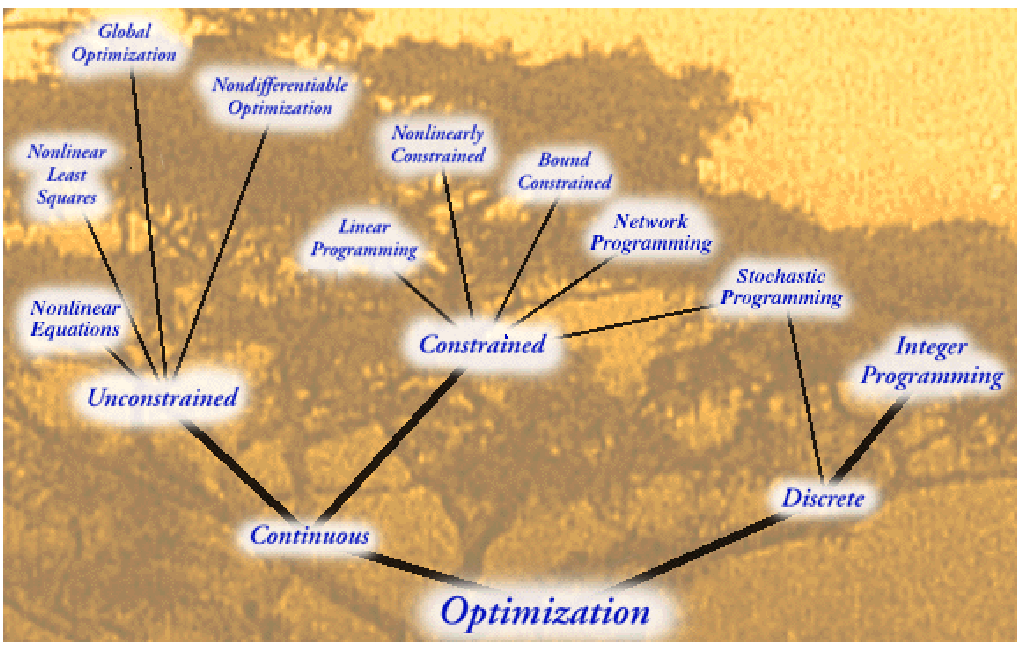
1. An introduction to dynamic optimization -- Optimal Control and Dynamic Programming

AGEC 637 - Summer 2009

I. Overview of optimization

Optimization is the unifying paradigm in almost all economic analysis. So before we start, let's think about optimization. The tree below provides a very 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/

In this course we will use both analytical and numerical methods to solve a certain class of optimization problems. This class focuses on a set of optimization 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:

$$\max_z u(z_a, z_b) \; s.t.$$

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

[pay attention to the notation: z is the vector of choice variables and x is the consumer's exogenously determined income.]

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

$$\max_z U\left(z_{1a}, z_{1b}, z_{2a}, z_{2b}\right) = U\left(z_1, z_2\right) \; s.t.$$

$$\mathbf{p}' z_1 + \mathbf{p}' z_2 \le x_1$$

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 chose z_1, we must take into account how it will affect our choices in period 2.

We're going to make a **huge** (though common) assumption and maintain that assumption throughout the course: utility is additively separable across time. (See Deaton and Muellbauer pp. 137-142 on the negative implications of assuming preferences are additive)

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

Clearly one way to solve this problem would be just as we would a standard static problem: set up a Lagrangian and solve for all optimal choices simultaneously. This may work here, where there are only 2 periods, but if we have 100 periods (or even an infinite number of periods) then this could get really messy. This course will develop methods to solve such problems.

Instead of using brute force to find the solutions of all the z's in one step, we reformulate the problem. Let x_1 be the <u>endowment</u> 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 $x_2 \ge 0$. In this problem x_2 defines the **<u>state</u>** 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 called the **state equation**. This equation is also sometimes referred to as the *equation of motion* or the *transition equation*.

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

- 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 determined by the choices made in t. *The state variables in a problem are the variables upon which a decision maker bases his or her choices in each period.* Another important characteristic of state variables is that typically the choices you make in one period will influence the value of the state variable in the next period.
- A **state equation** defines the intertemporal changes in a state variable
- 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 variables** and I will use the terms interchangeably.
- p_a and p_b are **parameters** of the model. They are held constant or change exogenously and deterministically over time.
- Finally, we have what I call **intermediate variables**. These are variables that are really functions of the state and control variables and the parameters. For example, in the problem considered here, one-period utility might be carried as an intermediate variable. In firm problems, production or profit might be other intermediate variables while productivity or profitability (a firm's capacity to generate output or profits) could be state variables. *Do you see the difference? This is very important (see PS#1).* When you formulate a problem it is very important to distinguish state variables from intermediate variables.
- The single-period utility function, $u(z_t)$ will be called the **benefit function**.
- In many problems there are benefits (or costs) that accrue at the end of the planning horizon. This is captured in models by including a **salvage value**.
- 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.
- All of 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.

In sum, the problems that we will study will have the following features. In each period or moment in time the decision maker looks at the state variables (x_t), then chooses the control variables (z_t). The combination of x_t and z_t generates immediate benefits and costs. They also determine the probability distribution over x in the next period or moment.

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

$$\max_{z_t} \sum_{t=1}^{2} u_t\left(z_t\right) \quad s.t.$$
$$\left.\begin{array}{l} x_{t+1} - x_t = -\mathbf{p}' z_t \\ x_{t+1} \geq 0 \end{array}\right\} t = 1,2 \tag{1}$$
$$x_1 \text{ fixed}$$

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 this formulation is quite general in that you could easily write the *n*-period problem by simply replacing the 2's in (1) with *n*.

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.

To see this approach, 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):

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

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

More terminology

In optimal control theory, the variable λ_t is called the **costate variable** and, 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 it is the marginal value of the state variable, x_t. The costate variable plays a critical role in dynamic optimization.

The FOCs for (2) are standard:

$$\partial L / \partial z_{ti} = \partial u / \partial z - \lambda_t p_i = 0, \quad i = a, b, \quad t = 1,2$$
$$\partial L / \partial x_2 = \frac{\partial u}{\partial x_2} - \lambda_1 + \lambda_2 = 0$$
$$\partial L / \partial \lambda_t = \left(x_t - x_{t+1} - \mathbf{p}' z_t\right) = 0 \ \ , t\text{=1,2.}$$

We now use a little notation 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

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

Some things to note about the Hamiltonian:

- the t^{th} Hamiltonian includes only z_t and λ_t,
- Unlike in a Lagrangian, only the RHS of state equation appears in the parentheses.

In the left column of table below we present the first-order conditions of the Lagrangian specification. Then on the right we present the derivative of the Hamiltonian with respect to the same variables. By comparison, we then can see what we would have to place on the right-hand side of the first derivative to obtain the same optimum if using the Hamiltonian that we would reach if we used the Lagrangian approach.

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

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> $\Rightarrow$ 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 4^{th} condition, called the transversality condition (how we *transverse* over to the world beyond t=1,2) we're 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 that purpose. We'll 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.

Although in this class most of the OC problems we'll face are in continuous time, the parallels 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 is now 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 simply

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

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

$$\begin{aligned} &\max_{z_1} u(z_1) + V_2(x_2) \quad s.t. \\ &x_2 = x_1 - \mathbf{p}' z_1 \end{aligned} \tag{3}$$

Note that we've 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. Assume for a moment that the functional form of $V(x_2)$ has been found. We can then write out Lagrangian of the first period problem,

$$L = u(z_1) + V_2(x_2) + \lambda_1(x_1 - \mathbf{p}' z_1 - x_2).$$

Again, we see that the economic meaning of the costate variable, λ_1 is just as in the OC setup, i.e., it is equal to the marginal value of a unit of x_1.

Of course the problem is that we do not have an explicit functional form for $V(\cdot)$ and as the problem becomes more complicated, obtaining a functional form becomes more difficult, even impossible for many problems. Hence, the trick to solving DP problems is to find the function $V(\cdot)$.

V. 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.

VI. Reading for next lecture

Leonard and Van Long, chapter 2.

VII. References

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