1 Finite Horizon, No Uncertainty

Consider a general finite horizon maximization problem in discrete time. Let y_t be the stock or state variable, that summarizes the state of the economy at each time t (all current information relevant to the decision maker). Let z_t be the flow or control variable, which is under the optimizer's choice and affects changes in the stock variable y_t. I refer to y as "the" state and z as "the" control as if there is only one of each, but these could very well be vectors of states and controls respectively, and the analysis here continues to hold.

The criterion or objective function faced by the decision maker is,

$$\sum_{t=0}^{T} F(y_t, z_t, t) \tag{1}$$

The implicit assumption in the objective function is that it is additively separable in the period functions $F(.)$. Each period $t = 0, 1, 2, ...T$ the decision maker faces a dynamic constraint describing the relationship between the control and state variables,

$$y_{t+1} = y_t + Q(y_t, z_t, t) \tag{2}$$

and is often called the law of motion, as it describes the evolution of the state variable. There could also be some time t constraints on time t control and state variables,

$$G(y_t, z_t, t) \leq 0 \tag{3}$$

for each $t = 0, 1, 2, ...T$. The initial value of the stock variable $y_0 > 0$ is taken to be given as an initial condition (the result of some unspecified history). Next we discuss solution methods for this intertemporal optimization problem.

1.1 Lagrange Method

One approach to solving this problem is the standard Lagrange method. This is similar to the standard Lagrange method used in static optimization as long as we are cautious of the intertemporal nature of the problem. For example we have to recognize that the constraints (2) and (3) faced by the decision maker are sequences of constraints, one for each $t = 0, 1, ...T$. To apply the Lagrange method let λ_t be the multiplier on the time t constraint (3), and π_{t+1} the multiplier on the time t law of motion. These multipliers have the standard interpretation as shadow prices. Consider for example the interpretation of π_{t+1}, the multiplier on the dynamic constraint, which is less standard. This is the shadow price on the time $t + 1$ stock: it tells us how much the value of the objective would increase if the constraint of increasing the stock next period was relaxed, i.e., if we were given a gift of a little bit extra y_{t+1}. Next, set up the Lagrangian,

$$L = \sum_{t=0}^{T} F(y_t, z_t, t) + \sum_{t=0}^{T} \pi_{t+1} \{y_t + Q(y_t, z_t, t) - y_{t+1}\} - \sum_{t=0}^{T} \lambda_t G(y_t, z_t, t)$$

We can re-arrange the terms in the second summation term of the Lagrangian as follows,

$$\sum_{t=0}^{T} \pi_{t+1} \{y_t - y_{t+1}\} = \pi_1(y_0 - y_1) + \pi_2(y_1 - y_2) + ... + \pi_{T+1}(y_T - y_{T+1}) =$$

$$= \pi_1 y_0 + y_1(\pi_2 - \pi_1) + y_2(\pi_3 - \pi_2) + ... + y_T(\pi_{T+1} - \pi_T) - \pi_{T+1} y_{T+1} =$$

$$= \sum_{t=1}^{T} y_t(\pi_{t+1} - \pi_t) + \pi_1 y_0 - \pi_{T+1} y_{T+1}$$

Then the Lagrangian can be re-written as,

$$L = \sum_{t=1}^{T} F(y_t, z_t, t) + \sum_{t=1}^{T} \pi_{t+1} Q(y_t, z_t, t) + \sum_{t=1}^{T} y_t(\pi_{t+1} - \pi_t) - \sum_{t=1}^{T} \lambda_t G(y_t, z_t, t) +$$

$$+ F(y_0, z_0, 0) + \pi_1 Q(y_0, z_0, 0) - \lambda_0 G(y_0, z_0, 0) + \pi_1 y_0 - \pi_{T+1} y_{T+1}$$

The first order conditions with respect to z_t for $t = 0, 1, ...T$ are,

$$F_z(y_t, z_t, t) + \pi_t Q_z(y_t, z_t, t) - \lambda_t G_z(y_t, z_t, t) = 0$$

The first order conditions with respect to y_t for $t = 1, ...T$ are,

$$F_y(y_t, z_t, t) + \pi_{t+1} Q_y(y_t, z_t, t) + \pi_{t+1} - \pi_t - \lambda_t G_y(y_t, z_t, t) = 0$$

What about the choice of y_{T+1}? If any positive stock is left then it must be worthless, i.e.,

$$y_{T+1} \geq 0, \pi_{T+1} \geq 0,$$

with complementary slackness. Another way of writing this is, $y_{T+1}\pi_{T+1} = 0$, and it is called a transversality condition.

We can re-write the time t FOC with respect to y_t as,

$$[F_y(y_t, z_t, t) - \lambda_t G_y(y_t, z_t, t)] + \pi_{t+1} Q_y(y_t, z_t, t) = -(\pi_{t+1} - \pi_t)$$

This is the standard intertemporal no-arbitrage condition. The left-hand-side of this expression is the "dividend" that you would get from having an extra unit of y_t (first term is the current period return, while the second term is the extra value from having one more unit in the following period). The right-hand-side is the change in the shadow price of y_{t+1}, and can be thought of as the "capital gain." These conditions characterize the optimal sequence $\left\{z_t^*, y_{t+1}^*\right\}_{t=0}^T$ for a given initial $y_0 > 0$. Plugging the optimal sequence into the objective function we get the resulting maximum value function, as a function of the initial state y_0

$$V(y_0) \equiv \sum_{t=0}^{T} F(y_t^*, z_t^*, t)$$

Then the derivative of the value function with respect to the initial stock, $V_y(y_0)$, is the shadow price of initial y_0.

1.2 Dynamic Programming (DP)

DP is an alternative way to solve dynamic optimization problems. Consider again the problem above,

$$\sum_{t=0}^{T} F(y_t, z_t, t) \tag{4}$$

s.t.

$$y_{t+1} = y_t + Q(y_t, z_t, t) \tag{5}$$

$$G(y_t, z_t, t) \leq 0 \tag{6}$$

for $t = 0, 1, 2, ...T$. Let the resulting maximum value function as a function of the initial stock be,

$$V(y_0) \equiv \sum_{t=0}^{T} F(y_t^*, z_t^*, t)$$

Suppose now that instead of starting at $t = 0$ we started the optimization at some other point in time $t = \tau > 0$. Then for decisions starting at time τ the only relevant information about the past is summarized in the stock variable at time τ, that is y_τ. Taking now y_τ as given we can start the whole problem at $t = \tau$, that is, solve,

$$max \sum_{t=\tau}^{T} F(y_t, z_t, t) \tag{7}$$

s.t.

$$y_{t+1} = y_t + Q(y_t, z_t, t) \tag{8}$$

$$G(y_t, z_t, t) \leq 0 \tag{9}$$

for $t = \tau, \tau + 1, \tau + 2, ...T$. Let the maximum value function of this problem be,

$$V(y_\tau, \tau) \equiv \sum_{t=\tau}^{T} F(\widehat{y}_t, \widehat{z}_t, t)$$

where $\{\widehat{z}_t, \widehat{y}_{t+1}\}_{t=\tau}^{T}$ is the optimal sequence chosen from the point of view of time τ. The shadow price on initial stock y_τ is $V_y(y_\tau, \tau)$.

As an example, suppose that $\tau = 1$. Then we could write the value function of the sub-problem starting at time $t = 1$ as $V(y_1, 1)$. Consider the choice of decision maker about z at time $t = 0$. The choice of z_0 affects the current value of the objective $F(y_0, z_0, 0)$ but also affects next period's stock y_1 through the law of motion: $y_1 = y_0 + Q(y_0, z_0, 0)$, which in turn implies, if you are optimizing from $t = 1$ onwards, a maximum value of $V(y_1, 1)$. Then the total value from choosing z_0 at $t = 0$ can be broken into two terms, $F(y_0, z_0, 0)$ which accrues at once and $V(y_1, 1)$ which accrues thereafter. The choice of z_0 at $t = 0$ should maximize the sum of these two terms,

$$V(y_0, 0) = max_{z_0} \{F(y_0, z_0, 0) + V(y_1, 1)\}$$

s.t.

$$y_1 = y_0 + Q(y_0, z_0, 0)$$

$$G(y_0, z_0, 0) \leq 0$$

More generally when contemplating the choice of z_t for any two consecutive periods $(t, t+1)$ the decision maker should maximize,

$$V(y_t, t) = max_{z_t} \{F(y_t, z_t, t) + V(y_{t+1}, t+1)\}$$

s.t.

$$y_{t+1} = y_t + Q(y_t, z_t, t)$$

$$G(y_t, z_t, t) \leq 0$$

This suggests an algorithm to solve the original optimization problem by starting at the end and then working backward recursively.

Start at the last period T for any y_T. Given that there is no future beyond T, the continuation value $V(y_{T+1}, T+1) = 0$. Then the problem in the last period T is,

$$max_{z_T} \{F(y_T, z_T, T)\}$$

s.t.

$$y_{T+1} = y_T + Q(y_T, z_T, T)$$

$$G(y_T, z_T, T) \leq 0$$

This is a static optimization problem. The solution is a *policy function* $z(y_T, T)$ (that is a rule that gives the optimal value of the control for any given value of the state), which in turn yields the maximum value function $V(y_T, T) = F(y_T, z(y_T, T), T)$.

Next we move one period backwards, and solve the following maximization problem for each value of the state y_t,

$$V(y_t, t) = max_{z_t} \{F(y_t, z_t, t) + V(y_{t+1}, t+1)\} \tag{10}$$

s.t.

$$y_{t+1} = y_t + Q(y_t, z_t, t)$$
$$G(y_t, z_t, t) \leq 0$$

where we replace $V(y_{t+1}, t+1)$ on the RHS above with $V(y_T, T)$ that we derived in the first step of the algorithm. The solution to this problem is a policy function $z(y_t, t)$ and a value function $V(y_t, t) = F(y_t, z(y_t, t), t) + V(y_t + Q(y_t, z(y_t, t), t), t+1)$. This step is repeated all the way back to $t = 0$. The outcome of this algorithm is a sequence of policy functions $\{z(y_t, t)\}_{t=0}^{T}$ and a sequence of value functions $\{V(y_t, t)\}_{t=0}^{T}$.

This approach of optimization over time, of solving a succession of static optimization problems, is called dynamic programming, and was pioneered by Richard Bellman. The underlying idea is that whatever the decision at time t the subsequent decisions should proceed optimally for the subproblem starting at $t+1$ is called the Bellman Principle of Optimality. Equation (10) is called the Bellman Equation.

To characterize the solution to the problem at time t look at the maximization problem on the RHS of (10),

$$max_{z_t} \{F(y_t, z_t, t) + V(y_t + Q(y_t, z_t, t), t+1)\}$$

s.t.

$$G(y_t, z_t, t) \leq 0$$

where I have substituted the law of motion into the period $t+1$ value function. Letting λ_t be the multiplier on the time t constraint yields the FOC for z_t,

$$F_z(y_t, z_t, t) + V_y(y_{t+1}, t+1) Q_z(y_t, z_t, t) - \lambda_t G_z(y_t, z_t, t) = 0$$

Notice that this is the same condition as that derived from the Lagrange method after you recognize that V_y and π are both shadow prices. So the two methods lead to the same rule for setting z.

When z_t is chosen optimally then the Bellman equation holds with equality,

$$V(y_t, t) = \{F(y_t, z_t, t) + V(y_{t+1}, t+1)\}$$

Differentiate this with respect to y_t and using the Envelope theorem on the RHS we get,

$$V_y(y_t, t) = F_y(y_t, z_t, t) + V_y(y_{t+1}, t+1)\left[1 + Q_y(y_t, z_t, t)\right] - \lambda_t G_y(y_t, z_t, t)$$

2 Infinite Horizon, No Uncertainty

Suppose that $T \to \infty$, as is the case in many economic problems. Since there is no last period anymore we cannot use backward induction to solve the dynamic programming problem. We focus here on stationary economies. Stationarity means that conditional on the state, the value and policy functions are independent of time, i.e., they are not indexed by t. Dropping the time index from the value function and policy function they become $z(y)$ and $V(y)$ respectively. We now denote the next period variables with a prime, e.g. y_t and y_{t+1} would be y and y' respectively. The Bellman equation is now,

$$V(y) = max_z \{F(y, z) + V(y')\}$$

s.t.

$$y^{'} = y + Q(y, z)$$

$$G(y, z) \leq 0$$

The optimal policy is a time-invariant function of the current state $z(y)$.

The value function $V(y)$ enters both the RHS and the LHS of the Bellman (although evaluated at different points, $V(y')$ on RHS and $V(y)$ on LHS). Recall this is different from the finite horizon case where V was indexed by a different time index on the RHS and LHS. The same value function enters on both sides of the Bellman in the infinite horizon case and this function is unknown. Thus in the infinite horizon case the Bellman equation is a *functional equation* (FE) because the unknown is a function rather than a number. Thus here we obtain the value and policy functions by simultaneously solving the maximization problem above to get $z(y)$ and by solving the functional equation to get $V(y)$. This is conceptually more complicated than backward induction. Important questions that have to be answered are: does a solution exist to the FE? if yes, is it unique? What are the properties of the value and policy functions? How do you find a solution of the FE? These questions are left for 1st year Ph.D courses! See applications from class about how to use first order conditions and the envelope theorem (Benveniste-Scheinkman) when the value function is differentiable, to derive optimality conditions (Euler equations) from dynamic programs.