Lecture 5: Time-Invariant Decision Rules and Dynamic Programming

Consider an infinite-horizon optimisation problem under certainty. We could search for an optimal sequence of decisions (control variables), where each decision depends upon the entire history of the state variables. Alternatively, we could restrict attention to outcomes where the current decision variable is a function solely of the current state variable.[1]

Suppose that the return function has the form

$$\sum_{t=0}^{\infty} \beta^t r\left(d_t, s_t\right), \tag{1}$$

where d_t is the *decision variable* and s_t is the *state variable*. Note that the function r does not depend on time except through the decision and state variables. The decision variable d_t is constrained to belong to the correspondence $\phi\left(s_t\right)$.

It is often convenient if we can restrict attention to *time-invariant decision rules* h, where $d_t = h\left(s_t\right)$. This rule is feasible if $h\left(s_t\right) \in \phi\left(s_t\right)$ for every possible s_t. Under certain assumptions it will turn out that such rules maximise (1) in the class of all decision rules.

1 Dynamic Programming

Suppose that we have the following problem

$$\begin{aligned} &\max_{\{d_t\}_{t=0}^{\infty}} \sum_{t=0}^{T} \beta^t r\left(d_t, s_t\right) + \beta^{T+1} W^0\left(s_{T+1}\right) \text{ subject to} \\ &g\left(d_t, s_t\right) = s_{t+1},\ t = 0, ..., T \\ &s_0 \text{ given,} \end{aligned} \tag{2}$$

where r is a concave function and the set $\{s_{t+1}, d_t, s_t | g\left(d_t, s_t\right) \geq s_{t+1}\}$ is convex and compact for all admissable d_t and r and d are as smooth as needed.[2]

[1] We can think of the time-t state and control variables as vectors.

[2] This section follows chapter 1 of Sargent, Thomas J., *Dynamic Macroeconomic Theory*, Cambridge, MA, Harvard University Press, 1987.

1.1 Solving as one large problem

The optimisation problem (2) can be solved as one large problem. The Lagrangian is

$$L = \sum_{t=0}^{T} \beta^t r\left(d_t, s_t\right) + \beta^{T+1} W^0\left(s_{T+1}\right) + \sum_{t=0}^{T} \beta^t \lambda_t \left[g\left(d_t, s_t\right) - s_{t+1}\right], \tag{3}$$

where, for convenience, I have written the multipliers in present value form. The first-order conditions for an interior solution are

$$\frac{\partial r_t}{\partial d_t} + \lambda_t \frac{\partial g_t}{\partial d_t} = 0,\ t = 0, ..., T \tag{4a}$$

$$\beta \frac{\partial r_t}{\partial s_t} + \beta \lambda_t \frac{\partial g_t}{\partial s_t} - \lambda_{t-1} = 0,\ t = 1, ..., T \tag{4b}$$

$$\beta \frac{dW^0}{ds_{T+1}} - \lambda_T = 0 \tag{4c}$$

$$s_{t+1} = g_t,\ t = 0, ..., T, \tag{4d}$$

where $r_t = r\left(d_t, s_t\right)$ and $W^0 = W^0\left(s_{T+1}\right)$.

Shifting equation (4b) forward one period yields

$$\lambda_t = \beta \frac{\partial r_{t+1}}{\partial s_{t+1}} + \beta \lambda_{t+1} \frac{\partial g_{t+1}}{\partial s_{t+1}},\ t = 0, ..., T-1.$$

Solving this linear first-order difference equation forward and using (4c) yields

$$\lambda_t = \beta \frac{\partial r_{t+1}}{\partial s_{t+1}} + \sum_{s=2}^{T-t} \beta^s \frac{\partial r_{t+s}}{\partial s_{t+s}} \prod_{u=1}^{s-1} \frac{\partial g_{t+u}}{\partial s_{t+u}} + \beta^{T+1-t} \frac{dW^0}{ds_{T+1}} \prod_{u=1}^{T-t} \frac{\partial g_{t+u}}{\partial s_{t+u}}, \tag{5}$$

$$t = 1, ..., T-2.$$

$$\lambda_{T-1} = \beta \frac{\partial r_T}{\partial s_T} + \beta^2 \frac{dW^0}{ds_{T+1}} \frac{\partial g_T}{\partial s_T}. \tag{6}$$

Substituting (4c), (5) and (6) into (4a) yields

$$\frac{\partial r_T}{\partial d_T} + \beta \frac{dW^0}{ds_{T+1}} \frac{\partial g_T}{\partial d_T} = 0 \tag{7a}$$

$$\frac{\partial r_{T-1}}{\partial d_{T-1}} + \beta \frac{\partial r_T}{\partial s_T} \frac{\partial g_{T-1}}{\partial d_{T-1}} + \beta^2 \frac{dW^0}{ds_{T+1}} \frac{\partial g_{T-1}}{\partial d_{T-1}} \frac{\partial g_T}{\partial s_T} = 0 \tag{7b}$$

$$\frac{\partial r_t}{\partial d_t} + \frac{\partial g_t}{\partial d_t} \left(\beta \frac{\partial r_{t+1}}{\partial s_{t+1}} + \sum_{s=2}^{T-t} \beta^s \frac{\partial r_{t+s}}{\partial s_{t+s}} \prod_{u=1}^{s-1} \frac{\partial g_{t+u}}{\partial s_{t+u}} + \beta^{T+1-t} \frac{dW^0}{ds_{T+1}} \prod_{u=1}^{T-t} \frac{\partial g_{t+u}}{\partial s_{t+u}} \right) \tag{7c}$$

$$= 0,\ t = 0, ..., T-2.$$

We are now left with equations (4d) and (7). By (4d) (evaluated at T) and (7a) we can find

$$s_{T+1} = g\left(h_T\left(s_T\right), s_T\right) \text{ and } d_T = h_T\left(s_T\right). \tag{8}$$

Likewise, by (4d) (evaluated at $T-1$) and (7b) we can find

$$s_T = g\left(h_{T-1}\left(s_{T-1}\right), s_{T-1}\right) \text{ and } d_{T-1} = h_{T-1}\left(s_{T-1}\right). \tag{9}$$

Continuing in this fashion we can use (4d) and (7c) to find

$$s_{t+1} = g\left(h_t\left(s_t\right), s_t\right) \text{ and } d_t = h_t\left(s_t\right),\ t = 0, ..., T. \tag{10}$$

1.2 Solving the problem by backwards recursion

Once we derived the first-order conditions for problem (2) we were able to solve them recursively. This suggests that it might be sensible to set up the problem in a way that exploits this. Consider the one-period problem of maximising $r\left(d_T, s_T\right) + \beta W_o\left(s_{T+1}\right)$ subject to $s_{T+1} = g\left(d_T, s_T\right)$ and s_T given. Define the *value function* for this problem as

$$\begin{aligned} W^1\left(s_T\right) &= \max_{d_T}\left\{r\left(d_T, s_T\right) + \beta W^0\left(s_{T+1}\right)\right\} \text{ subject to} \\ s_{T+1} &= g\left(d_T, s_T\right) \\ & \quad s_T \text{ given.} \end{aligned}$$

The Lagrangian for this problem is

$$r\left(d_T, s_T\right) + \beta W_o\left(s_{T+1}\right) + \lambda_T\left[g\left(d_T, s_T\right) - s_{T+1}\right].$$

After eliminating the Lagrangian the first-order conditons are (4d) (evaluated at T) and (7a). Consequently, equation (8) holds and

$$W^1\left(s_T\right) = r\left(h_T\left(s_T\right), s_T\right) + \beta W^0\left(g\left(h_T\left(s_T\right), s_T\right)\right). \tag{11}$$

If we knew that h was differentiable then differentiating and using (11), (7a) and (8) would yield[3]

$$\frac{dW^1}{ds_T} = \frac{\partial r_T}{\partial s_T} + \beta\frac{dW^0}{ds_{T+1}}\frac{\partial g_T}{\partial s_T}. \tag{12}$$

[3] I am being very casual in this heuristic approach. See Benveniste, Lawrence and Jose Scheinkman, "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica* 47, 1979, 727-732.

Now define the value function for a two-period problem as

$$\begin{aligned} W^2(s_{T-1}) &= \max_{d_{T-1}} \left\{ r(d_{T-1}, s_{T-1}) + \beta W^1(s_T) \right\} \text{ subject to} \\ s_T &= g(d_{T-1}, s_{T-1}) \\ & s_{T-1} \text{ given.} \end{aligned}$$

Proceeding as before we can find the first order conditions for the maximisation problem as (4d) (evaluated at $T-1$) and (7b). Consequently equation (9) holds and

$$W^2(s_{T-1}) = r(h_{T-1}(s_{T-1}), s_{T-1}) + \beta W^1(g(h_{T-1}(s_{T-1}), s_{T-1})).$$

Differentiating and substituting in (12) yields

$$\frac{dW^2}{ds_{T-1}} = \frac{\partial r_{T-1}}{\partial s_{T-1}} + \beta \frac{\partial g_{T-1}}{\partial s_{T-1}} \left(\frac{\partial r_T}{\partial s_T} + \beta \frac{dW^0}{ds_{T+1}} \frac{\partial g_T}{\partial s_T} \right).$$

Continuing we find

$$\frac{dW^{s+1}}{ds_{T-s}} = \frac{\partial r_{T-s}}{\partial s_{T-s}} + \beta \frac{\partial g_{T-s}}{\partial s_{T-s}} \frac{dW^s}{ds_{T-s+1}}. \tag{13}$$

Looking at (4b) and (4c) we see that these derivatives are the multipliers.

This T-period problem has a particular structure that allows us to solve it by backwards recursion this way: the time-t decision variable d_t affects the present and future but not the past. In economics the assumption of perfect foresight sometimes makes this assumption invalid. Suppose for a moment that $T = 3$. Then we are saying that

$$W^4(s_0) = \max_{\{d_0, d_1, d_2\}} \left\{ r(d_0, s_0) + \beta r(d_1, s_1) + \beta^2 r(d_2, s_2) + \beta^3 W^0(s_3) \right\}$$

subject to $g(d_t, s_t) = s_{t+1}$, $t = 0, 1, 2$ and s_0 is equivalent to

$$W^4(s_0) = \max_{d_0} \left\{ r(d_0, s_0) + \beta \max_{d_1} \left\{ r(d_1, s_1) + \beta \max_{d_2} \left\{ r(d_2, s_2) + \beta W^0(s_3) \right\} \right\} \right\}.$$

Bellman's Principle of Optimality: "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

This solution method guarantees that the outcome is time consistent: there is no incentive to deviate from the optimal plan. Note also that this outcome is related to the way we find perfect equilibria in games.

1.3 An Example

Suppose that we want to solve[4]

$$\max_{\{d_t\}_{t=0}^{2}} -\sum_{t=0}^{2} d_t^2 \text{ subject to} \sum_{t=0}^{2} d_t = 100,\ d_t \geq 0.$$

In period 2 the problem is

$$\max_{d_2} -d_2^2 \text{ subject to} \sum_{t=0}^{2} d_t = 100,\ d_t \geq 0.$$

The trivial solution is

$$d_2 = 100 - d_0 - d_1$$

In period 1 the problem is

$$\max_{d_1} -d_1^2 - (100 - d_0 - d_1)^2 .$$

The first-order conditions are

$$d_1 = d_2 = \frac{100 - d_0}{2}$$

In period 0 the problem is

$$\max_{d_0} -d_0^2 - \frac{1}{2}(100 - d_0)^2 .$$

The first-order conditions are

$$d_0 = d_1 = d_2 = \frac{100}{3}.$$

1.4 The Value Function

We have that Bellman's equation says that

$$V^{j+1}(s_{T-j}) = \max_{d_{T-j}} \left\{ r(d_{T-j}, s_{T-j}) + \beta V^j(s_{T-j+1}) \right\} .$$

More neatly we can write this as

$$V^{j+1}(s) = \max_{d} \left\{ r(d, s) + \beta V^j(s') \right\} , \tag{14}$$

[4] This is from Intriligator, Michael D., *Mathematical Optimization and Economic Theory*, Englewood Cliffs, NJ, Prentice-Hall, Inc., 1971, ch. 13.

where a variable without a prime is a current-period variable and a variable with a prime is a next-period variable.

The 19th century approach to this problem was to guess $V^0(\cdot)$. Then, equation (14) gives us

$$V^1(s) = \max_d \left\{ r(d,s) + \beta V^0(s') \right\}.$$

Proceeding in this fashion, does one converge to a steady state? In this case

$$V(s) = \max_d \{ r(d,s) + \beta V(s') \}.$$

Is this steady state the same for any initial guess $V^0(\cdot)$? Does

$$V(s_0) = \max_{\{d_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(d_t, s_t)\,?$$

Suppose that a planner is constrained to have welfare of V^0 in the final period but can pick the optimal strategy given that in the period before, receiving V^1. Now suppose that he is constrained to get V^0 in the final period but can pick the optimal strategies subject to this for two periods before. One would expect that $V^2 \geq V^1$ and so forth as the planner gets more and more flexibility. So it turns out that uncertain certain assumptions the answers to these questions is yes. One set of sufficient conditions (Sargent (1987)) is that r is concave and bounded and that $\{s_{t+1}, d_t, s_t | g(d_t, s_t) \geq s_{t+1}\}$ for all admissable d_t.[5] There is a time-invariant policy rule $d_t = h(s_t)$ and we have that, as in equation (13), away from corners

$$V'(s) = \frac{\partial r(h(s), s)}{\partial s} + \beta \frac{\partial g(h(s), s)}{\partial s} V(g(h(s), s)). \tag{15}$$

This suggests two ways of solving the problem

$$\max_{\{d_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(d_t, s_t).$$

The first is to guess a solution V and verify that it is correct or solve for V numerically. The second is to iterate backward. Consider the problem

$$\max_{\{d_t\}_{t=0}^{T}} \sum_{t=0}^{T} \beta^t r(d_t, s_t).$$

[5] See Lucas and Stokey (1989) for a formal analysis.

Solve first

$$\max_{d_T} \sum_{t=T}^{T} \beta^t r\left(d_t, s_t\right).$$

Then solve

$$\max_{d_{T-1}} \sum_{t=T-1}^{T} \beta^t r\left(d_t, s_t\right)$$

after substituting in the solution to the proceeding problem. Proceed in this fashion to find

$$\max_{d_{T-s}} \sum_{t=T-s}^{T} \beta^t r\left(d_t, s_t\right).$$

Continue until you notice a pattern and then let $s \to \infty$. Or, if no pattern emerges (the usual case) solve these problems numerically until $d_t = h_t\left(s_t\right)$ converges.

1.5 An optimal growth problem

Consider the special case of the optimal growth problem:

$$\begin{aligned}
&\max \sum_{t=0}^{\infty} \beta^t \ln c_t \text{ subject to} \\
&c_t = Ak_t^\alpha - k_{t+1}
\end{aligned}$$

Along with the transversality condition the necessary and sufficient conditions are

$$\begin{aligned}
c_t &= \beta A \alpha k_t^{\alpha-1} c_{t-1} \\
c_t &= Ak_t^\alpha - k_{t+1}.
\end{aligned}$$

We could rewrite this as

$$\begin{aligned}
h\left(Ak^\alpha - h\left(k\right)\right) &= \frac{\beta A \alpha h\left(k\right)}{\left[Ak^\alpha - h\left(k\right)\right]^{1-\alpha}} \\
k_t &= Ak_{t-1}^\alpha - h\left(k_{t-1}\right)
\end{aligned}$$

and solve numerically for h. Or, one might guess that

$$\begin{aligned}
c_t &= \gamma k_t^\alpha \\
k_t &= \delta k_{t-1}^\alpha.
\end{aligned}$$

Substituting into the optimality equations yields

$$\begin{aligned}
\gamma\left(\delta k_{t-1}^{\alpha}\right)^{\alpha} &= \frac{\beta A\alpha\gamma k_{t-1}^{\alpha}}{\left(\delta k_{t-1}^{\alpha}\right)^{1-\alpha}} \\
\gamma k_t^{\alpha} &= Ak_t^{\alpha} - \delta k_t^{\alpha}
\end{aligned}$$

So, $\delta = \beta A\alpha$ and $\gamma = A(1-\beta A)$. Thus,

$$\begin{aligned}
c_t &= A(1-\beta A)\,k_t^{\alpha} \\
k_t &= \beta A\alpha k_{t-1}^{\alpha}.
\end{aligned}$$

Or, try solving this recursively as a T period problem. We have

$$\begin{aligned}
k_{T+1} &= 0,\ c_T = Ak_T^{\alpha} \\
k_T &= \frac{\alpha\beta Ak_{T-1}^{\alpha}}{1+\alpha\beta},\ c_{T-1} = \frac{Ak_{T-1}^{\alpha}}{1+\alpha\beta} \\
k_{T-1} &= \frac{\left(\alpha\beta+\alpha^2\beta^2\right)Ak_{T-2}^{\alpha}}{1+\alpha\beta+\alpha^2\beta^2},\ c_{T-2} = \frac{Ak_{T-2}^{\alpha}}{1+\alpha\beta+\alpha^2\beta^2}
\end{aligned}$$

Noticing a pattern here, I guess that

$$\begin{aligned}
k_{T-s} &= \frac{\left[\sum_{j=0}^{s+1}(\alpha\beta)^j - 1\right]Ak_{T-s-1}^{\alpha}}{\sum_{j=0}^{s+1}(\alpha\beta)^j} = \frac{\left[\alpha\beta-(\alpha\beta)^{s+2}\right]Ak_{T-s-1}^{\alpha}}{1-(\alpha\beta)^{s+2}} \to \alpha\beta Ak_{T-s-1}^{\alpha} \\
c_{T-s} &= \frac{Ak_{T-s-1}^{\alpha}}{\sum_{j=0}^{s+1}(\alpha\beta)^j} = \frac{(1-\alpha\beta)\,Ak_{T-s-1}^{\alpha}}{1-(\alpha\beta)^{s+2}} \to (1-\alpha\beta)\,Ak_{T-s-1}^{\alpha}.
\end{aligned}$$

This is an unusual problem in that it has a simple solution.

2 Optimisation Under Uncertainty

Consider an infinite horizon optimisation problem under uncertainty. We could search for an optimal sequence of actions where each action depends upon the entire history of the stochastic shocks. But, this might not be a particularly useful way of going about things. An alternative, that is often possible in economic applications, is to search for decision rules which are functions of a limited number of state variables.

The canonical stationary statistical decision problem was studied by David Blackwell in 1965.[6] It is

[6] David Blackwell was an African American mathematician who died a year and a half ago. He received his PhD from Illinois in 1941 at age 22. He went to the Institute of Advanced Studies as a post grad but, because of his race, he was not allowed to attend lectures at Princeton. Initially he could only be hired at black colleges, but finally in 1955 became the first tenured black professor at UC Berkeley.

assumed that the return function has the form

$$\sum_{t=0}^{\infty} \beta^t r(d_t, s_t), \tag{16}$$

where d_t is the *decision variable* and s_t is the *state variable.* It is assumed that the time-t state variable is either directly observable at time t or is a function of observable variables and can be inferred at time t. The cumulative distribution of s_{t+1}, conditional on s_t and d_t, denoted by $F(s_{t+1}|d_t, s_t)$ is time invariant. The decision variable d_t is constrained to belong to the correspondence $\phi(s_t)$.

A *stationary decision policy* δ is a decision rule $d_t = \delta(s_t)$. The key feature of δ is that it does not depend on time except through s_t. This rule is feasible if $\delta(s_t) \in \phi(s_t)$ for every possible s_t. Blackwell showed that if an optimal policy existed in a more general class of policies where $d_t = \delta(s_1, ..., s_t)$ and if $d_t \in \phi(s_t)$ almost surely then an optimal stationary policy exists.

Under some fairly weak assumptions that we usually assume are satisfied, there exists a unique value function for our problem that satisfies Bellman's optimality equation

$$v(s) = \sup_{d \in \phi(s)} \left\{ r(d, s) + \beta \int v(y)\, dF(s|d, s) \right\}. \tag{17}$$

The function v is the supremum of the expected discounted return (in equation (1)) over all feasible policies, given s. A feasible stationary policy $\delta(\cdot)$is optimal over all feasible policies if and only if this same supremum is obtained for $d = \delta(s)$. Compactness of the feasible sets $\phi(s)$ continuity assumptions ensure existence.

Which variables are part of s depends upon the model. In general, everything that is necessary to summarise the relevant part of the current environment and everything that is useful in predicting the future goes into s. In a simple model s might be the capital stock. If utility today depends on how much you consumed last period, then last period's consumption is typically part of s. If the money supply follows a second-order autoregressive process then both the current and past value of the money supply may be part of s.[7]

[7] For an intuitive treatment of this see"Equilibrium under Uncertainty: Multi-Agent Statistical Decision Theory." Robert M. Townsend and Edward C. Prescott, "Equilibrium under Uncertainty: Multi-Agent Statistical Decision Theory," in *Studies in Bayesian Econometrics and Statistics in Honor of Harold Jeffreys*, Arnold Zellner, ed., North-Holland, 1980: 169-194, http://www.robertmtownsend.net/sites/default/files/files/papers/published/EquilibriumUnderUncertainty1980.pdf. For a rigourous treatment see Stokey, Nancy L. and Robert E. Lucas, *Recursive Methods in Dynamic Economics*, Cambridge, MA., Harvard University Press, 1989.