Notes on Dynamic Programming

Econ720. Lutz Hendricks. September 6, 2017

---

1 Basic Idea

We are given a dynamic optimization problem in discrete time of the form

$$\max \sum_{t=1}^{T} u(k_t,\, c_t, t) + V_{T+1}\left(k_{T+1}\right)$$

subject to $k_{t+1} = g(k_t,\, c_t, t)$ with the initial condition k_0 given and the terminal value of k given, call it $V_{T+1}(k_{T+1})$. We call the variable with the equation of motion the *state variable* and the other choice variable the *control variable.*

One way of solving this problem is to set up a Lagrangean

$$\Gamma = \max \sum_{t=1}^{T} u(k_t,\, c_t, t) + \sum_{t=1}^{T} \lambda_t \left[g(k_t,\, c_t, t) - k_{t+1}\right] + V_{T+1}(k_{T+1}) \tag{1}$$

The first sum is simply the objective function, while the second sum is a convenient way of collecting all constraints. The FOCs are

$$\begin{aligned} u_c(t) &= -\lambda_t\, g_c(t) \\ u_k(t) &= -\lambda_t\, g_k(t) + \lambda_{t-1} \\ \lambda_T &= V'_{T+1}(k_{T+1}) \end{aligned}$$

This is a perfectly good approach, but DP is an alternative that is sometimes more convenient. The basic idea is to restart the problem at some date τ. First note that we can think of the maximized Lagrangean as an indirect utility function. After maximizing out all future values of c and k, Γ becomes a function only of k_t and t which we write as $V_\tau(k_\tau)$. Now move to period $\tau + 1$ and restart the problem again. Given a suitable structure of the problem (whatever that means!), the solution to the period $\tau + 1$ problem will be the same as that of the period τ problem, except for the period τ values of course, which are in the past from the $\tau + 1$ perspective. In other words, if the solution to the date τ problem is $(\hat{c}_t,\, \hat{k}_{t+1}),\ t = \tau,\, \ldots,\, T$, then the solution to the date $\tau + 1$ problem is $(\hat{c}_t,\, \hat{k}_{t+1}),\ t = \tau + 1,\, \ldots,\, T$. The decision maker does not revise his date τ plan at some later time. The problem is *time consistent.*

For this to work, we must be able to rearrange the terms in the Lagrangean so that the problem is divided into two sets of equations. The first set only contains variables prior to date τ and k_τ. The second set contains only variables from date τ onwards. Here:

$$\begin{aligned} \Gamma \;=\; & \max \sum_{t=1}^{\tau-1} u(k_t,\, c_t, t) + \sum_{t=1}^{\tau-1} \lambda_t \left[g(k_t,\, c_t, t) - k_{t+1}\right] \\ & + \sum_{t=\tau}^{T} u(k_t,\, c_t, t) + \sum_{t=\tau}^{T} \lambda_t \left[g(k_t,\, c_t, t) - k_{t+1}\right] \\ & + V_{T+1}(k_{T+1}) \end{aligned}$$

More generally,

$$\Gamma_{\tau*}(k_{\tau*}) = \max \sum_{t=\tau*}^{\tau-1} u(k_t,\, c_t, t) + \sum_{t=\tau*}^{\tau-1} \lambda_t \left[g(k_t,\, c_t, t) - k_{t+1}\right] + \Gamma_\tau(k_\tau) \tag{2}$$

If this is true, then we can write the date τ problem as

$$V_\tau(k_\tau) = \max_{c_\tau,\, k_{\tau+1}} \{u(k_\tau,\, c_\tau, \tau) + \lambda_\tau \left[g(k_\tau,\, c_\tau, \tau) - k_{\tau+1}\right] + V_{\tau+1}(k_{\tau+1})\} \tag{3}$$

You can convince yourself that this is true by expanding $V(.)$ on the right hand side repeatedly. You will recover exactly (1), except that there is a " max" for each date. But since the decision maker does not change his mind, these " max" 's are irrelevant.

The interpretation is simple. Given that behavior is optimal from date $\tau+1$ onwards, the value of $k_{\tau+1}$ is given by the indirect utility function $V_{\tau+1}(k_{\tau+1})$. The date τ decision can then be made by trading off current utility against next period utility.

Another interpretation is given by backward induction. Suppose we start at the last date. The decision problem is then trivial (because k_{T+1} is given) and we can derive an indirect utility function $V_T(.)$. Now step back to date $T-1$. Since V_T embodies optimal choice at $T+1$, all we need to do is solve the date T problem treating $V_T(k_T)$ as if it were a given terminal payoff function. This gives us $V_{T-1}(.)$ and we can step back one more period.

Yet another interpretation is to think of the date τ problem as containing a nested problem (the date $\tau+1$ problem), which can be solved first. All that needs to be known to then solve the full problem is the indirect utility function for the nested problem, $V_{\tau+1}(.)$.

2 Optimality Conditions

The necessary conditions for an optimum are derived exactly as in a static Lagrangean problem. In fact, there is nothing special at all about maximizing $V_t(.)$, except that it depends on the unknown function $V_{t+1}(.)$. But this does not matter for the optimality conditions. The only concern is whether $V_{t+1}(.)$ is well-behaved. Optimality therefore requires

$$\begin{aligned} u_c(t) &= -\lambda_t\, g_c(t) \\ \lambda_t &= V'_{t+1}(k_{t+1}) \end{aligned}$$

This looks profoundly useless because we don't know the derivative of the value function. But we can simply differentiate (3) to get it:

$$V'_t(k_t) = u_k(t) + \lambda_t\, g_k(t) \tag{4}$$

This condition is known as the *envelope condition* or the *Benveniste-Scheinkman* condition. Shifting (4) forward by one period, we can obtain an Euler equation

$$\begin{aligned} u_c(t) &= -g_c(t)\left[u_k(t+1) - u_c(t+1)/g_c(t+1)\, g_k(t+1)\right] \qquad (5) \\ &= g_c(t)\left[g_k(t+1)/g_c(t+1) \cdot u_c(t+1) - u_k(t+1)\right] \end{aligned}$$

Additional constraints are handled exactly as in the Lagrangean approach: simply add more terms, each with its own multiplier, to the right hand side of (3).

2.0.1 Under which conditions does DP work?

Consider a general (deterministic) optimization problem:

$$\max \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})$$

subject to

$$\begin{aligned} x_{t+1} &\in \Gamma(x_t) \\ x_0 &\in X \; given \end{aligned}$$

Under which conditions can this sequence problem be represented as a DP of the form

$$v(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta\, v(y)\}$$

Stokey and Lucas (1989, ch. 4) show the following: Assume that

1. The feasible set, $\Gamma(x_t)$, is non-empty for any x_t (Assumption 4.1).
2. For any feasible plan $\{x_t\}$ discounted utility is bounded in the sense that $\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t\, F(x_t, x_{t+1}) < \infty$ (A4.2).
3. The solution to the Bellman equation satisfies $\limsup_{t\to\infty} \beta^t\, v(x_t) \leq 0$ for any feasible plan $\{x_t\}$.

Then solving the sequence problem and solving the DP are equivalent (theorems 4.4 and 4.5). Note how weak these assumptions are. They essentially say:

1. The planner cannot reach a state from which no further choices are feasible (a technicality).
2. Utility is bounded. Otherwise, there is no problem to be solved.
3. The lim sup condition rules out certain degenerate cases where payoffs are delayed forever (see p. 76 of SL).

The problem with the above conditions is that the lim sup condition (#3) is hard to check. Alternative sufficient conditions are (SL p. 78):

- The feasible set is a convex subset of R^l (A4.3).
- The correspondence mapping today's state into next period's state, $\Gamma(x)$, is nonempty, compact-valued, and continuous (A4.3).
- The utility function is bounded and continuous with a discount factor $\beta < 1$ (A4.4).

These are actually a special case of A4.1 and A4.2 for bounded utility. For the stochastic case, see SL, theorem 9.2. For sufficient conditions such that the optimal plan can be derived from a Bellman equation, see SL, theorem 9.4.

3 An Example

Consider the following simple example:

$$\max \sum_{t=1}^{T} \beta^t u(c_t) + \beta^{T+1} V_{T+1}(k_{T+1})$$

subject to $k_{t+1} = f(k_t) - c_t$, k_1 given.

Mapping this into the general notation: $g(k, c) = f(k) - c$. The Euler equation is then the familiar one: $u_c(t) = f\prime(k_{t+1}) \beta u_c(t+1)$.

4 Stationary DP

Note that DP converts the problem of finding sequences of real variables into the much harder looking problem of finding a sequence of value functions and policy functions $c_t = h_t(k_t)$. It is not too surprising that this is often not at all simpler than directly dealing with the Lagrangean.

However, if the problem has an infinite horizon and is stationary (independent of t), DP becomes very powerful because it can make use of the fact that $V_{t+1}(.) = V_t(.)$, i.e. the value function is the same in every period.

Assume that u and g do not depend on t. The problem is then

$$\max \sum_{t=1}^{\infty} \beta^t u(k_t, c_t)$$

subject to $k_{t+1} = g(k_t, c_t)$ with the initial condition k_0 given (there is of course no terminal condition). Note that we allow for a discount factor in front of u. The Bellman equation is then

$$V(k) = \max \ u(k, c) + \lambda [g(k, c) - k\prime] + \beta V(k\prime)$$

It is conventional to drop the time subscripts and use a prime to denote next period variables. Note the β in front of next period's V.

4.1 Existence of the optimal policy function

It can be shown that the following conditions are sufficient to guarantee that an optimal policy function exists (Sundaram 12.5; SL):

- u is bounded and continuous.
- g is continuous.
- The set of feasible actions is a continuous, compact-valued correspondence.

Boundedness of the current reward function is typically violated in economic problems, but in ways that do not matter. For example, in the neoclassical growth model the utility function is typically unbounded, but it is known that attainable utility (given technology and endowments) is bounded.

4.2 Properties of the Value Function

This is a reader's guide to SL ch. 4 and ch. 9, which establish conditions under which the value function has nice properties, such as concavity, continuity, etc. To establish notation, consider the following general problem:

$$v(x,z) = \sup_{y \in \Gamma(x,z)} \left\{ F(x,y,z) + \beta \int_Z v(y,z')\, Q(z,dz') \right\}$$

Here $x \in X$ denotes the endogenous state variable (e.g. capital), y is the control (e.g. next period's capital stock), and $z \in Z$ is an exogenous shock which evolves according to the transition function Q.

Example: For a standard growth model, $k = x$, $k' = y$, z is the technology shock. $F(x,y,z) = u\left(f(k,z) - k'\right)$. The Bellman equation is

$$v(k,z) = \sup_{k' \in \Gamma(k,z)} \left\{ u(f(k,z) - k') + \beta \int_Z v(k',z')\, Q(z,dz') \right\}$$

where the feasible set of controls is determined by the feasibility constraint: $k' \in \Gamma(k,z) = \{k' \mid k' \leq f(k,z)\}$.

The argument underlying SL argument may be outlined as follows.

1. Prove that the Bellman equation, viewed as an operator, is a contraction mapping. Contraction mappings have unique fixed points. Hence, the value function is unique.

2. Prove that the Bellman operator maps the space of continuous and bounded functions into itself. Then *one* value function must be continuous and bounded (the operator has a fixed point in that space). But since the value function is unique, *the* value function must be continuous and bounded.

The details are unfortunately more complicated (see SL ch. 4 for the deterministic case and ch. 9 for the stochastic case). The properties of the value function are established using a sequence of theorems with ever stricter assumptions and sharper characterizations.

Next, I list a set of *assumptions* that are sufficient (in some combinations) for establishing properties of the value function:

1. Let X denote the set of possible values for the endogenous state variable. X is a convex Borel set in R^l (SL assumption 9.4). In the determinstic case, X is a convex subset of R^l.

2. Let Z denote the set of possible values for the exogenous shock. Z is either countable. Or Z is a compact Borel set in R^l and the transition function Q has the Feller property (assumption 9.5).

3. Let Γ denote the correspondence describing the constraints. That is, $\Gamma(x,z)$ is the set of feasible values for the endogenous state next period. Γ is nonempty, compact-valued, and continuous (assumption 9.6).

4. The period return function F is bounded, continuous, and $\beta \in (0,1)$ (assumption 9.7).

5. For each (y, z), the period return function $F(., y, z)$ is strictly increasing (assumption 9.8).

6. For each z , the set of feasible choices $\Gamma(x, z)$ is increasing in the sense that $x' \geq x$ implies $\Gamma(x, z) \subseteq \Gamma(x', z)$. In words, larger values of the endogenous state (e.g. more capital) increase the choice set (assumption 9.9).

7. For given z, the period utility function is strictly concave (jointly) in (x, y). That is $F(\hat{x}, \hat{y}, z) \geq \theta\, F(x, y, z) + (1 - \theta)\, F(x', y', z)$ for any $\theta \in (0, 1)$, (x, y), and (x', y'), where $(\hat{x}, \hat{y}) \equiv \theta\,(x, y) + (1 - \theta)\,(x', y')$ (assumption 9.10).

8. For given z, the choice set $\Gamma(x, z)$ is convex in the following sense: Pick any $\theta \in (0, 1)$, $y \in \Gamma(x, z)$ and $y' \in \Gamma(x', z)$. Then $\theta\, y + (1 - \theta)\, y' \in \Gamma\,(\theta\, x + (1 - \theta)\, x', z)$. Note that y and y' are chosen for different x values (assumption 9.11).

9. For given z, the period utility function $F(., ., z)$ is continuously differentiable in (x, y) (assumption 9.12).

10. For given (x, y), $F(x, y, .)$ is strictly increasing in the shock (assumption 9.13).

11. For given x, $\Gamma(x, .)$ is increasing in z in the sense that $z' > z$ implies $\Gamma(x, z) \subseteq \Gamma(x, z')$ (assumption 9.14).

12. The transition function Q is monotone. That is, if $f : Z \to R$ is nondecreasing, then the "expectation" function $(Mf)(z) = \int f(z')\, Q\,(z, dz')$ is also nondecreasing (assumption 9.15).

A sequence of theorems in SL provide sufficient conditions, drawn from the assumptions above, for various properties of the value function.

Theorem 9.6: If assumptions 9.4-9.7 hold, then the value function is *unique*.

Theorem 9.7: If assumptions 9.4-9.9 hold, then the value function $v(., z)$ is strictly *increasing* in the endogenous state x for any given value of the shock z.

Theorem 9.8: If assumptions 9.4-9.7 and 9.10-9.11 hold, then the value function is strictly *concave* in x for each z. The policy correspondence $G(., z)$ is a *single-valued* function and *continuous*.

Theorem 9.10: If assumptions 9.4-9.7 and 9.10-9.12 hold, then the value function is continously *differentiable* in x at interior points (x_0, z_0). Interior points satisfy $x_0 \in int\, X$ and $g(x_0, z_0) \in int\, \Gamma(x_0, z_0)$.

Theorem 9.11: If assumptions 9.4-9.7 and 9.13-9.15 hold, then $v(x, .)$ is strictly *increasing* in z.

4.2.1 Transition Functions

Transition functions are used to describe the evolution of stochastic shocks over time. Let (Z, Ξ) be a measurable space. A transition function $Q : Z \times \Xi \to [0, 1]$ takes as inputs a value of the shock $z \in Z$ and a set $A \in \Xi$. It returns the probability that next period's shock $z' \in A$, given z: $\Pr(z' \in A \,|\, z) = Q(z, A)$. Formally, the function Q must satisfy:

a. For each $z \in Z$, $Q(z, .)$ is a probability measure on (Z, Ξ).

b. For each $A \in \Xi$, $Q(., A)$ is a Ξ-measurable function.

The associated conditional expectations operator for any measurable function f is defined by

$$(Tf)(z) = \int f(z')\, Q(z, dz')$$

This is also called the Markov operator associated with Q.

A transition function Q has the *Feller property*, if the operator T maps the space of bounded continuous functions into itself. In other words: The conditional expectation $E\{f(z') \mid z\}$ is a continuous and bounded function, if f is a continuous and bounded function.

See SL, ch. 8.1 for more detail.

5 Guess and Verify

It is sometimes possible to solve explicitly for the value function and the policy function in a DP problem. A common approach is to simply guess the value function and to verify that guess by plugging it into the right-hand-side of Bellman's equation. Then solve the maximization problem, which is now simply a static problem, and verify that the solution, the value function on the left-hand-side, has the conjectured form.

Of course, it is usually impossible to exactly guess the value function, but sometimes its general form can be guessed. For example, one might guess that $V(k) = A + B\,\ln(k)$ without being able to guess A and B. That is enough. The correct values of A and B can be determined as we go along verifying the general form of V.

5.0.1 Example (Dixit, ch. 11)

A consumer maximizes

$$\sum_{t=0}^{\infty} \beta^t\, c_t^{1-\sigma}/(1-\sigma)$$

subject to W_0 given and $W_{t+1} = R\,(W_t - c_t)$. The gross interest rate R is constant. Bellman's equation is

$$V(W) = \max\; c^{1-\sigma}/(1-\sigma) + \beta\, V(R\,[W - c])$$

The FOC is

$$c^{-\sigma} = \beta\, R\, V\prime(W\prime)$$

and the envelope condition is

$$V\prime(W) = \beta\, R\, V\prime(W\prime)$$

which leads to the Euler equation

$$c^{-\sigma} = \beta\, R\,(c\prime)^{-\sigma}$$

To solve this explicitly, guess that the value function inherits the functional form of the period utility function:

$V(W) = A\,W^{1-\sigma}/(1-\sigma)$. The right-hand-side of Bellman's equation is then $c^{1-\sigma}/(1-\sigma) + \beta\, A\, R^{1-\sigma}\,(W - c)^{1-\sigma}/(1-\sigma)$. The first-order condition becomes $c^{-\sigma} = \beta\, A\, R^{1-\sigma}\,(\,W - c)^{-\sigma}$, which can be solved for the policy function

$$c(W) = W/(1 + [\beta\, A\, R^{1-\sigma}]^{1/\sigma})$$

Substituting this into the right-hand-side of Bellman's equation yields

$$AW^{1-\sigma}/(1-\sigma) = [\,W/(1+B)]^{1-\sigma}/(1-\sigma) + \beta\, AR^{1-\sigma}\;\left(W\,(1 - 1/[1+B])\right)^{1-\sigma}/(1-\sigma) \tag{6}$$

where $B = (\beta\, A\, R^{1-\sigma})^{1/\sigma}$. Since this has the correct functional form, the guess for the value function is verified. Moreover, we can use (6) to solve for A.

$$A = (1 + B)^{\sigma-1} + B^{\sigma}\,(1 - 1/[1 + B])^{1-\sigma}$$

This implies

$$\begin{aligned} A\,(1+B)^{1-\sigma} &= 1 + B^{\sigma}\,B^{1-\sigma} = 1 + B \\ A^{1/\sigma} &= 1 + (\beta\, A\, R^{1-\sigma})^{1/\sigma} \\ 1 - \beta^{1/\sigma}\,R^{(1-\sigma)/\sigma} &= A^{-1/\sigma} \end{aligned}$$

Instead of guessing the value function, we can also guess the policy function: $c(W) = X\,W$. The Euler equation is then
$X\,W = (\beta\,R)^{-1/\sigma}\,X\,R\,(W - X\,W)$, which can be solved for X: $\beta^{1/\sigma}\,R^{(1-\sigma)/\sigma} = 1 - X$. Again, since the conjectured functional form is recovered, the guess is verified.

6 Reading

- Stokey and Lucas (with Prescott, SL) is an entire book on DP; an excellent reference. Ch. 2 is very readable and contains a self-contained statement of the method together with an economic example.
- Acemoglu, "Introduction to Modern Economic Growth," ch. 6.
- Sargent (DMT ch. 1) and Dixit ch. 11 both explain the " guess and verify" method.
- Sargent and Ljungvist (2000, ch. 20).
- Sundaram (1996).