ECON8000: Quantitative Skills for Economics
Lecture 9: Optimization 2

Shino Takayama

University of Queensland

March 2020

Stationary Case

We now consider the case $T = +\infty$.

A deterministic stationary discounted dynamic programming problem (henceforth, SDP) is specified by a tuple $\{S, A, \Phi, f, r, \delta\}$, where the following description holds.

1. S is the state space with generic element s. We assume that $S \subset \mathbb{R}^n$ for some n.
2. A is the action space with typical element a. We assume that $A \subset \mathbb{R}^k$ for some k.
3. $\Phi : S \to P(A)$ is the feasible action correspondence that specifies for each $s \in S$ the set $\Phi(s) \subset A$ actions that are available at s.
4. $f : S \times A \to S$ is the transition function for the state, that specifies for each current state-action pair (s, a) the next-period state $f(s, a) \in S$.
5. $r_S \times A \to \mathbb{R}$ is the (one period) reward function that specifies a reward $r(s, a)$ when the action a is taken at the state s.
6. $\delta \in [0, 1)$ is the one-period discount factor.

The interpretation of this framework is similar to that of the finite-horizon model but now time is infinite and the problem is represented as:

$$\begin{aligned}
&\max \sum_{t=0}^{\infty} \delta^t r(s_t, a_t) \\
&s.t: \\
&s_{t+1} = f(s_t, a_t) \quad t = 0, 1, 2, \ldots \\
&a_t \in \Phi(s_t) \quad t = 0, 1, 2, \ldots
\end{aligned}$$

Histories, strategies and the Value Function

- A t-history $h_t = \{s_0, a_0, \dots, s_{t-1}, a_{t-1}, s_t\}$ for the SDP is a list of the state and action in each period up to $t-1$ and thhe period-t state. Let $H_0 = S$, and for $t = 1, 2, \dots$ let H_t denote the set of all possible t-histories. Denote by $s_t[h_t]$ the period-t state in the history h_t.
- A strategy σ is a sequence of functions $\{\sigma_t\}_{t=0}^{\infty}$ such that for each $t = 0, 1, \dots$, $\sigma_t : H_t \to A$ satisfies the feasibility requirement that $\sigma_t(h_t) \in \Phi(s_t[h_t])$. Let Σ denote the space of all strategies.

Fix a strategy σ. From each given initial state $s \in S$, the strategy σ determines a unique sequence of states and actions $\{s_t(\sigma, s), a_t(\sigma, s)\}$ as follows: $s_0(\sigma, s) = s$ and for $t = 0, 1, 2 \ldots,$:

$$\begin{aligned}
h_t(\sigma, s) &= \{s_0(\sigma, s), a_0(\sigma, s), \ldots, s_t(\sigma, s)\} \\
a_t(\sigma, s) &= \sigma_t(h_t(\sigma, s)) \\
s_{t+1}(\sigma, s) &= f(s_t(\sigma, s), a_t(\sigma, s))
\end{aligned}$$

Thus each strategy σ, induces form each initial state s, a period-t reward $r_t(\sigma)(s)$ where:

$$r_t(\sigma)(s) = r[s_t(\sigma, s), a_t(\sigma, s)]$$

Let $W(\sigma)(s)$ denote the total discounted reward from s under strategy σ:

$$W(\sigma)(S) = \sum_{t=0}^{\infty} \delta^t r_t(\sigma)(s)$$

The value function $V : S \to \mathbb{R}$ of the SDP is defined as:

$$V(s) = \sum_{\sigma \in \Sigma} W(\sigma)(s)$$

A strategy σ^* is said to be an optimal strategy for the SDP if:

$$W(\sigma^*)(s) = V(s) \quad , s \in S$$

The Bellman Equation

Simple examples show that several problems arise in the search for an optimal strategy unless we impose some restrictions on the structure of the SDP.

1. Two or more strategies may yield infinite total rewards from some initial states, yet may not appear equivalent from an intuitive stand point.
2. No optimal strategy may exist even if $\sum_{t=0}^{\infty} \delta^t r(s_t, a_t)$ converges for each feasible sequence $\{a_t\}$ from each initial state s.

The first problem can be overcome if we assume S and A to be compact sets and r to be continuous on $S \times A$. In this case, r is bounded on $S \times A$, so for any feasible action sequence, total reward is finite since $\delta < 1$. More generally, we may directly assume, as we shall, that r is bounded on $S \times A$.

Assumption 1. There is a real K such that $|r(s, a)| \leq K$ for all $(s, a) \in S \times A$.

Lemma. Under Assumption 1, the value function V is well defined as a function from S to $\mathbb{R}$, even if an optimal strategy does not exist. Moreover, given any $s \in S$ and any $\epsilon > 0$ it is the case that there exists a strategy σ, depending possibly on ϵ and s, such that $W(\sigma(s)) \geq V(s) - \epsilon$.

Lemma: The Bellman Equation

Lemma. The value function V satisfies the following equation (the "Bellman equation") at each $s \in S$:

$$V(s) = \sup_{a \in \Phi(s)} \{r(s,a) + \delta V[(f(s,a)]\}.$$

Contraction

Definition. Let (X, d) be a metric space and $T : X \to X$. The mapping T is a contraction if there is $\rho \in [0, 1)$ such that $d(T(x), T(y)) \leq \rho d(x, y) \quad x, y \in X$.

Theorem. Let (X, d) be a complete metric space and $T : X \to X$ a contraction. Then T has a unique fixed point.

Review from Lec 4: The Contraction Mapping Theorem

- Economic equilibria are typically fixed points.
- For this reason the result we describe now, which is also known as the Banach fixed point theorem, is especially significant for us.
- Let (X, d) be a metric space, and let $f : X \to X$ be a function.
- A point x is a fixed point of f if $f(x) = x$.
- We say that f is a contraction if there is a constant $\alpha \in (0, 1)$, called the modulus of contraction, such that $d(f(x), f(y)) \leq \alpha d(x, y)$ for all $x, y \in X$.

Review: The Contraction Mapping Theorem

Theorem: If X is nonempty and complete and $f : X \to X$ is a contraction, then f has a unique fixed point.

Proof. Let α be the modulus of contraction. Uniqueness is easy: if x and x' are fixed points, then

$$d(x, x') = d(f(x), f(x')) \leq \alpha d(x, x'),$$

so $d(x, x') = 0$ and thus $x = x'$.

Proof: Existence

To prove existence, let x be any point in X, and define the sequence $x_0, x_1, x_2, \ldots$ be setting $x_0 = x$ and $x_i = f(x_{i-1})$ for $i \geq 1$. By induction $d(x_i, x_{i+1}) \leq \alpha^i d(x_0, x_1)$, so when $m < n$ we have

$$d(x_m, x_n) \leq \left(\sum_{i=m}^{n-1} \alpha_i \right) d(x_0, x_1) \leq \frac{\alpha^m}{1 - \alpha} d(x_1, x_2).$$

Thus (x_i) is Cauchy, and has a limit x^* because X is complete. We have

$$d(f(x^*), x_{i+1}) \leq \alpha d(x^*, x_i) \to 0,$$

so $x_i \to f(x^*)$ and thus $f(x^*) = x^*$.

Blackwell's conditions

Let $S \subset \mathbb{R}^\ell$. Let $B(S)$ denote the set of all bounded functions from S to $\mathbb{R}$.

Lemma. $B(S)$ is a complete metric space when endowed with the sup norm: for each $v, w \in B(S)$,
$d(v, w) = \sup_{y \in S} |v(y) - w(y)|$.

Lemma (Blackwell's conditions). Let $L : B(S) \to B(S)$ be any mapping that satisfies

- (1) Monotonicity: $w \geq v \to Lw \geq Lv$
- (2) Discounting: There is $\beta \in [0, 1)$ such that $L(w + c) = Lw + \beta c$ for all $w \in B(S)$ and $c \in \mathbb{R}$.

Then L is a contraction.

These conditions are also known as *Blackwell's conditions*.

Proof: Blackwell's Conditions

- Let $v, w \in B(S)$.
- Then for any $s \in S$,
  $w(s) - v(s) \leq \sup_{y \in S} |w(y) - v(y)| = \|w - v\|$.
- Hence $w(s) \leq v(s) + \|w - v\|$.
- By applying monotonicity and discounting in order, we obtain

$$Lw(s) \leq L(v + \|w - v\|)(s) \leq Lv(s) + \beta\|w - v\|.$$

- Overall, we obtain
  $Lw(s) - Lv(s), Lv(s) - Lw(s) \leq \beta\|w - v\|$, so
  $|Lw(s) - Lv(s)| \leq \beta\|w - v\|$.
- Hence, $\sup_{s \in S} |Lw(s) - Lv(s)| = \|Lw - Lv\| \leq \beta\|w - v\|$.

Stationary Equilibrium

Definition. A Markovian strategy σ for the SDP is defined to be a strategy where for each T, σ_t depends on h_t only through t and the period-t state under h_t, $s_t[h_t]$.
Definition. A stationary strategy is a Markovian strategy $\{\pi_t\}$ which satisfies the further condition that $\pi_t = \pi_\tau = \pi$ for all t and τ.
Thus, in a stationary strategy, the action taken in any period t depends only on the state at the beginning of that period, and not even on the value of t. It is usual to denote such a strategy by $\pi^{(\infty)}$, but for notational simplicity we denote it by function π.
Definition. A stationary optimal strategy is a stationary strategy that is also an optimal strategy.

Assumptions

We have already assumed that:

1. r is bounded on $S \times A$
2. r is continuous on $S \times A$.
3. f is continuous on $S \times A$.
4. Φ is a continuous, compact-value correspondence on S.

Theorem. Suppose the SDP $\{S, A, \Phi, f, r\delta\}$ satisfies assumptions 1 to 4. Then there exists a stationary optimal policy π^*. Furthermore, the value function $V = W(\pi^*)$ is continuous on S, and is the unique bounded function that satisfies the Bellman equation at each $s \in S$:

$$\begin{aligned} W(\pi^*)(s) &= \max_{a \in \Phi(s)} \{r(s,a) + \delta W(\pi^*)(f(s,a))\} \\ &= r(s, \pi^*(s)) + \delta W(\pi^*)[f(s, \pi^*(s))] \end{aligned}$$

Optimal Growth Model

1. There is a single good which may be consumed or invested.
2. The conversion of investment to output takes one period and is achieved through a production function $f : \mathbb{R}_+ \to \mathbb{R}_+$.
3. x_t denotes period-t investment
4. y_{t+1} denote the output available on period-$(t+1)$, given $f(x_t)$.
5. Agents beings with an initial endowment of $y = y_0 \in \mathbb{R}_{++}$.
6. In each period, agent observes available stock y_t and decides on the division of this stock between period-t consumption c_t and period-t investment x_t.
7. Consumption of c_t in period t gives utility $u(c_t)$: $u : \mathbb{R}_+ \to \mathbb{R}$.
8. Agent discounts future utility by discount factor $\delta \in [0, 1)$ and wishes to maximize total discounted utility from lifetime consumption.

Problem to Solve

Thus, the problem is to solve:

$$\max \sum_{t=0}^{\infty} \delta^t u(c_t)$$

$$s.t \quad y_0 = y$$

$$y_{t+1} = f(x_t) \quad t = 0, 1, 2 \ldots$$

$$c_t + x_t = y_t \quad t = 0, 1, 2 \ldots$$

$$c_t, x_t \geq 0 \quad t = 0, 1, 2 \ldots$$

Environment

- $S^* = \mathbb{R}_+$ is the state space.
- $A^* = \mathbb{R}_+$ action space.
- $\Phi(y) = [0, y]$ is the feasible action correspondence taking states $y \in S^*$ into the set of feasible action $[0, y] \subset A^*$ at y.
- $r(y, c) = u(c)$, the reward from taking action $c \in \Phi(y)$ at the state $y \in S^*$.
- $F(y, c) = f(y - c)$ is the transition function taking current state-action pairs (y, c) into future states $F(y, c)$.
- The tuple $\{S^*, A^*, \Phi, r, F, \delta\}$ now defines a stationary discounted programming problem, which represents the optimal growth model.

Existence of Optimal Strategies

- Notice that u may be unbounded on $\mathbb{R}_+$.
- Rather than imposing unboundedness, we consider more natural and plausible restriction which will ensure that we may restrict S^* and A^* to compact intervals in $\mathbb{R}_+$, thereby obtaining boundedness of u from its continuity.

Suppose that production function f satisfies the following conditions:

1. $f(0) = 0$ (no free production)
2. f is continuous and nondecreasing on $\mathbb{R}_+$ (continuity and monotonicity)
3. There is $\bar{x} > 0$ such that $f(x) \leq x \quad \forall x \geq \bar{x}$ (unproductivity at high investment levels)

1. $u : \mathbb{R}_+ \to \mathbb{R}$ is continuous on $\mathbb{R}_+$
2. The tuple $\{S, A, \Phi, r, F, \delta\}$ now meets the requisite compactness and continuity to guarantee the existence of an optimal strategy.

Theorem. There is a stationary optimal strategy $g : S \to A$ in the optimal growth problem under Assumptions 1 and 2. The value function V is continuous on S and satisfies the Bellman equation at each $y \in S$:

$$\begin{aligned} V(y) &= \max_{c \in [0,y]} \{u(c) + \delta V[f(y-c)]\} \\ &= u(g(y)) + \delta V[f(y - g(y))] \end{aligned}$$

Example

An economy is composed by a continuum of identical agents with preferences:

$$\mathbb{E}\left(\sum_{t=0}^{\infty} \beta^t u(c_t)\right)$$

where c_t is the level of consumption in t, with $u_c > 0$ and $u_{cc} < 0$. The consumption good (also used as capital asset) is produced according to the function $F(K_t, N_t)$, where K_t is capital, which depreciates at a rate $\delta < 1$, N_t is total employment and F exhibits constant returns to scale (Cobb-Douglas). Each agent has an initial amount of capital K_0 and a unit of time to devote to work. Consumers are the owners of capital which lease it to the representative firm.

Example: Recursive Formulation

(a) Define the recursive problem of the central planner and derive the first order conditions.

Answer. $V(k) = \max_{0 \leq y \leq f(k)} u(f(k) - y) + \beta V(y).$

The first order condition with respect to y is

$$u'(K(1-\delta) + F(K, N) - y) = \beta V'(y)$$

Example

Assume now that in the same economy a labour tax τ is imposed to provide liquidity to the investment I of a public good, g. This public good depreciates each period at a rate $\lambda < 1$ such that $g_t = g_t(1-\lambda) + I$. Agents value this public good, so the utility function per period is $u(c_t; g_t)$.

(b) Define the recursive problem of the central planer. The wage is given by w. (Hint: $I = \tau w = \tau F_2(K, N) = \tau \alpha F(K, N)$ if Cobb-Douglas $K^{1-\alpha} \cdot N^{\alpha}$ is assumed and $N = 1$).

(c) What happen if $\lambda = 1$. Is g still a state variable?

Answer (b)

Normalizing with $N = 1$, we have:

$$I = \tau w \rightarrow \tau F_2(K, 1) = \tau \cdot \alpha F(K, 1) \text{(assuming Cobb-Douglas)}$$

Therefore, the recursive problem of the central planner is defined as:

$$V(K, g) = \max_{K'} \left\{ u\left(K(1-\delta) + F(K,1) - K' - \tau \cdot \alpha F(K,1); g\right) \right.$$
$$\left. + \beta V(K', g') \right\}$$
$$s.t$$
$$g' = g(1-\lambda) + \tau \cdot \alpha F(K,1)$$

Answer (c)

If $\lambda = 1$, then $g' = \tau\alpha F(K, 1)$ and we don't need g to project g'. However, g is still a state variable given that enters in the utility function and its value is determined by the level of capital in the last period.

Example 2 (Q2 in page 278 Sundaram)

Consider the problem of optimal harvesting of a natural resource. A firm (say, a fishery) begins with a given stock of $y > 0$ of a natural resource (fish). In each period $t = 1, 2, \cdots, T$ of a finite horizon, the firm must decide how much of the resource to sell on the market that period. If the firm decides to sell x units of the resource, it receives a profit of $\pi(x)$ where $\pi : \mathbb{R} \to \mathbb{R}$. The amount $(y - x)$ of the resource left unharvested grows to an available amount of $f(x)$ at the beginning of the next period where $f : \mathbb{R} \to \mathbb{R}$. The firm wishes to choose a strategy that will maximize the sum of its profits over the model's T-period horizon.

Q & A (a)

Question. Set up firm's optimization problem as a finite-horizon dynamic programming problem, i.e., describe precisely the state space S, the action space A, the period-t reward function feasible action r_t, etc.
Answer. $S = [0, f_T(y)]$ and state is the stock of fish y that firm have. $A = [0, f_T(y)]$. $r_t(y, x) = \pi(x)$. $f_t(y, x) = f(y - x)$. $\Phi_t(y) = [0, y]$.

Q & A (b)

Question. Describe sufficient conditions on f and π under which an optimal strategy exists in this problem.
Answer. They need to satisfy:

- **A1**: π has to continuous on $[0, f_T(y)]$ + bounded on $S \times A$ (which is the case given that $S \times A$ is compact).
- **A2**: f has to be continuous on $\mathbb{R}$.
- **A3**: Φ_t is continuous, compact-value correspondence on S.

Q & A (c)

Question. Assuming $\pi(x) = \log x$ and $f(x) = x^\alpha$ $(0 < \alpha \leq 1)$, solve this problem for the firm's optimal strategy using backwards induction.

Answer. Start in T, where optimal is $g_T(y) = y$ (that is, consume all the remaining stock). Therefore, $V_T(y) = \ln(y)$. Then, at period $T - 1$,

$$\max_{x \in [0,y]} \ln x + ln(y - x)^\alpha.$$

Answer (c) Continued

Given that problem consist in maximize a logarithm function, FOC are necessary and sufficient. Finding FOC we obtain:

$$g_{T-1}(y) = \frac{y}{1+\alpha}$$

which implies that
$V_{T-1}(y) = (1+\alpha)\ln y - (1+\alpha)\ln(1+\alpha) = (1+\alpha)\ln y + \psi$.
Then, doing the same in $T-2$ gives:

$$g_{T-2} = \frac{y}{1+\alpha+\alpha^2}$$

$$V_{T-2}(y) = (1+\alpha+\alpha^2)\ln y + \Omega(\alpha, t)$$

Answer (c) continued

So, a guess arises:

$$g_t(y) = \frac{y}{1 + \alpha + \ldots + \alpha^{T-t}}$$

$$V_t(y) = (1 + \alpha + \ldots \alpha^{T-t}) \ln y + \Omega(\alpha, t)$$

We have to verify the guess:

$$\max_{x \in [0,y]} \ln x + (1 + \ldots + \alpha^{T-t-1}) \ln (y - x)^{\alpha}$$

with solution:

$$x = \frac{y}{1 + \alpha + \ldots \alpha^{T-t}}$$

so guess of g is confirmed. Then, plugging $x = \frac{y}{1+\alpha+\ldots\alpha^{T-t}}$ confirms that our guess was indeed the Value Function.