Dynamic programming

Andrew Gimber, European University Institute*

Tuesday, 6 December 2011

The sequence problem

Consider a generic infinite-horizon maximisation problem:

$$V^*(s_0) = \max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \tilde{\sigma}(s_t, c_t) \quad (1)$$

subject to:

$$c_t \in C(s_t) \ \forall t \quad (2)$$

$$s_{t+1} = \tau(s_t, c_t) \ \forall t \quad (3)$$

$$s_0 \text{ given.} \quad (4)$$

The agent is trying to maximise the discounted value of his lifetime payoffs $V^*(s_0)$ by choosing an infinite sequence of **control variables** $\{c_t\}_{t=0}^{\infty}$. (In general, the agent may have control over several variables in each period, so c_t may be a vector.) The payoff $\tilde{\sigma}(s_t, c_t)$ the agent gets in period t may depend on both his choices c_t and the state of the world s_t. (Since there may be multiple **state variables**, s_t may also be a vector.) The agent may face feasibility constraints on his choice of the control variables. Equation 2 says that his choice c_t must be in the feasible set $C(s_t)$, which itself may depend on the state of the world. The state of the world evolves according to a law of motion (equation 3, also known as a **transition equation**), which says that the state tomorrow may depend both on the state today and on the agent's choices. Finally, equation 4 gives us the **initial condition**, which is the state of the world the agent finds himself in when he faces this problem.

Let's consider a specific example of a problem of this kind: a Ramsey growth model with fixed labour supply and full depreciation. The decision-maker here is a social planner who chooses consumption amounts for each period to maximise the discounted lifetime utility of a representative agent:

*These notes draw heavily upon Nicola Pavoni's notes for the Macroeconomics II course taught at the EUI in Spring Term 2011.

$$V^*(k_0) = \max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{5}$$

$$\text{subject to:}$$

$$c_t \in [0, f(k_t)] \ \ \forall t \tag{6}$$

$$k_{t+1} = f(k_t) - c_t \ \forall t \tag{7}$$

$$k_0 \text{ given.} \tag{8}$$

The recursive formulation

What we are going to show is that there is an equivalent (and sometimes more convenient) formulation of the problem. In general, it takes the following form:

$$V(s) = \max_{c \in C(s)} \tilde{\sigma}(s, c) + \beta V(s') \tag{9}$$

$$\text{subject to:}$$

$$s' = \tau(s, c) \tag{10}$$

$V(s)$ is what we refer to as the **value function**. It answers the question: What is the discounted lifetime utility of an agent who faces the sequence problem characterised by equations 1–4 (with $s_0 = s$), given that he will solve the problem optimally? This seems like a strange question to ask, but in the process of answering it we will find the answer to another, much more intuitive, question: What is the optimal choice of c for an agent who finds himself in state of the world s?

Equation 9 is known as the **Bellman (functional) equation**. The essence of dynamic programming is writing down and solving equations of this type. The Bellman equation implicitly defines a policy function $c = \phi(s)$ that maps the current realisations of the state variables into optimal choices for each of the control variables.

Going back to our Ramsey example, we have:

$$V(k) = \max_{c \in [0, f(k_t)]} u(c) + \beta V(k') \tag{11}$$

$$\text{subject to:}$$

$$k' = f(k) - c, \tag{12}$$

which implicitly defines a policy function $c = \phi(k)$.

From the transition equation (10), we know that the agent may be able to affect the future state s' through his current choice c. This means we can think of him as choosing s' directly, and so we can rewrite the Bellman equation as:

$$V(s) = \max_{s' \in \Gamma(s)} \sigma(s, s') + \beta V(s'), \tag{13}$$

where $\Gamma(s)$ is the constraint set that contains all feasible values of s' given s.[1] Since we are thinking of the agent as choosing s', the policy function takes the form $s' = g(s)$.

For our Ramsey example, the rewritten Bellman equation would be:

$$V(k) = \max_{k' \in [0, f(k)]} u\left(f(k) - k'\right) + \beta V\left(k'\right), \tag{14}$$

with policy function $k' = g(k)$.

Showing that the two problems are equivalent

Bellman's Principle of Optimality (which we will use in the proof of Proposition 1 below) is as follows:

> "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." (Bellman, 1957)

Let's relate this to the generic sequence problem characterised by equations 1–4 above. Suppose we have an optimal sequence $\{c_t\}_{t=0}^{\infty}$ that solves the problem. The initial state is s_0 and the initial decision is c_0, so from the transition equation (3) we know that the state resulting from the first decision will be $s_1 = \tau\left(s_0, c_0\right)$. The agent is now faced with a new problem in which the initial condition is s_1 and he must choose a new optimal sequence $\{\tilde{c}_t\}_{t=1}^{\infty}$. The Principle of Optimality states that the remaining sequence $\{c_t\}_{t=1}^{\infty}$ that was chosen optimally in period 0 must solve this new problem, that is, it must be the case that $\{c_t\}_{t=1}^{\infty}=\{\tilde{c}_t\}_{t=1}^{\infty}$.

Proposition 1. *The "true" value function V^* generated by solving the sequence problem solves the Bellman functional equation.*

Proof.

$$\begin{aligned}
V^*(s_0) &= \max_{c_0 \in C(s_0)} \left\{ \max_{\{c_t \in C(s_t)\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t \tilde{\sigma}\left(s_t, c_t\right) \right\} \text{ subject to } s_{t+1} = \tau\left(s_t, c_t\right) \ \forall t \\
&= \max_{c_0 \in C(s_0)} \left\{ \max_{\{c_t \in C(s_t)\}_{t=1}^{\infty}} \tilde{\sigma}\left(s_0, c_0\right) + \beta \sum_{t=0}^{\infty} \beta^t \tilde{\sigma}\left(s_{t+1}, c_{t+1}\right) \right\} \text{ subject to } s_{t+1} = \tau\left(s_t, c_t\right) \ \forall t \\
&= \max_{c_0 \in C(s_0)} \tilde{\sigma}\left(s_0, c_0\right) + \beta \max_{\{c_t \in C(s_t)\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t \tilde{\sigma}\left(s_{t+1}, c_{t+1}\right) \text{ subject to } s_{t+1} = \tau\left(s_t, c_t\right) \ \forall t \\
&= \max_{c_0 \in C(s_0)} \tilde{\sigma}\left(s_0, c_0\right) + \beta V^*\left(s_1\right) \text{ subject to } s_1 = \tau\left(s_0, c_0\right)
\end{aligned}$$

This is exactly the same as (9) and (10), just with different notation (s_0, c_0 and s_1 instead of s, c and s' respectively). □

[1] We switch the notation for the objective function from $\tilde{\sigma}$ to σ to reflect the change in its arguments.

Proposition 2. *If V solves the Bellman functional equation and for all feasible sequences $\{s_{t+1} \in \Gamma(s_t)\}_{t=0}^{\infty}$ we have*

$$\lim_{T\to\infty} \beta^T V(s_T) = 0, \tag{15}$$

then V is the "true" value function, that is, $V = V^*$, and any sequence generated by the associated policy function $s' = g(s)$ starting from s_0 is optimal.

Proof.

$$\begin{aligned}
V(s_0) &= \max_{s_1 \in \Gamma(s_0)} \sigma(s_0, s_1) + \beta V(s_1) \\
&= \max_{\{s_{t+1}\in\Gamma(s_t)\}_{t=0}^{1}} \sigma(s_0, s_1) + \beta\sigma(s_1, s_2) + \beta^2 V(s_2) \\
&\vdots \\
&= \max_{\{s_{t+1}\in\Gamma(s_t)\}_{t=0}^{T-1}} \sum_{t=0}^{T-1} \beta^t \sigma(s_t, s_{t+1}) + \beta^T V(s_T) \\
&\overset{T\to\infty}{=} \max_{\{s_{t+1}\in\Gamma(s_t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \sigma(s_t, s_{t+1}) \\
&= \max_{\{c_t\in C(s_t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \tilde{\sigma}(s_t, c_t) \text{ subject to } s_{t+1} = \tau(s_t, c_t) \ \forall t \\
&= V^*(s_0)
\end{aligned}$$

□

The contraction mapping theorem

Consider our simplified generic Bellman equation (13). As shown in Russell Cooper's "Overview of Dynamic Programming" notes, there are some specific cases (such as the cake-eating problem and the Ramsey model with fixed labour supply, log utility and full depreciation) where we can use the "guess and verify" approach to find closed-form solutions for the value function V and the policy function $s' = g(s)$. However, the usefulness of the dynamic programming approach is not limited to these special cases.

For cases where we cannot find a closed-form solution for V, it would be nice to have a fool-proof procedure for approximating it. (Having approximated V, we can then infer an approximate policy function.) Fortunately, such a procedure exists. It involves making an initial guess V_0 for the value function and repeatedly applying (usually with the help of a computer) something called the **Bellman operator** (not to be confused with the similar-looking Bellman equation) to generate an iterative sequence $\{V_n\}_{n=1,2,\ldots}$. The Bellman operator, which we will label T, maps functions into functions as follows:

$$TV_n = \max_{s'\in\Gamma(s)} \sigma(s, s') + \beta V_n(s') \tag{16}$$

Starting with V_0 and applying T repeatedly, we get:

$$\begin{aligned} V_1 &= TV_0 = \max_{s' \in \Gamma(s)} \sigma(s, s') + \beta V_0(s') \\ V_2 &= TV_1 = T^2 V_0 \\ &\vdots \\ V_n &= TV_{n-1} = T^n V_0 \end{aligned}$$

We refer to this procedure as **value function iteration**. What we would like is for this iterative procedure to converge to the "true" value function V that solves the Bellman equation (and, by Proposition 2, is the value function of the original sequence problem). It would be even better if this were to happen regardless of how wrong our initial guess V_0 was. The contraction mapping theorem tells us the conditions we need for this to be the case. Before we get to the theorem itself, let's recall some definitions (all of which should have featured in Antonio Villanacci's Background Course in Mathematics).

Definition (Metric space)**.** A metric space (X, d) is a set X, together with a metric (or distance function) $d : X \times X \to \mathbb{R}$ such that for all $x, y, z \in X$, we have:

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) \leq d(x, y) + d(y, z)$

Definition (Convergence)**.** Let (X, d) be a metric space. The sequence $\{x_n\}_{n=0}^{\infty}$ converges to the limit $y \in X$ (with respect to the metric space (X, d)) if for each real number $\epsilon > 0$ there exists a natural number N such that for any $n \geq N$ we have $d(x_n, y) < \epsilon$.

Less formally, if a sequence converges to a point, then you can always get arbitrarily close to that point by going far enough along the sequence.

Definition (Cauchy)**.** A sequence $\{x_n\}_{n=0}^{\infty}$ is Cauchy if for each real number $\epsilon > 0$ there exists a natural number N such that for any $l, m \geq N$ we have $d(x_l, x_m) < \epsilon$.

In a Cauchy sequence, the elements of the sequence are getting closer and closer together as you move along the sequence. Any convergent sequence is necessarily Cauchy, but the reverse is not always true. This brings us to our next definition.

Definition (Complete metric space)**.** A metric space (X, d) is complete if every Cauchy sequence is a convergent sequence.

Definition (Contraction mapping)**.** Let (X, d) be a metric space and let $\tilde{T} : X \to X$ be a function mapping X into itself. $\tilde{T}$ is a contraction mapping (with modulus $\tilde{\beta}$) on (X, d) if for some $\tilde{\beta} \in [0, 1)$, we have

$$d(\tilde{T}x, \tilde{T}y) \leq \tilde{\beta} d(x, y)$$

for all $x, y \in X$.

Now we are ready to present the theorem itself.

Theorem 1 (Contraction mapping theorem (Banach fixed-point theorem))**.** *If (X, d) is a complete metric space and $\tilde{T} : X \to X$ is a contraction mapping with modulus $\tilde{\beta}$, then:*

1. *$\tilde{T}$ has exactly one fixed point in X, that is, $\exists!\ x^* \in X$ such that $\tilde{T}x^* = x^*$.*

2. *For any $x_0 \in X$, the sequence $\{x_n\}_{n=1}^{\infty}$ where $x_n = \tilde{T}x_{n-1}$ converges to x^*.*

Proof. See Section 3.2 of Stokey, Lucas and Prescott (1989), or the Wikipedia entry for "Banach fixed-point theorem". □

So now we know the conditions under which the value function iteration approach "works", that is, converges to the true value function regardless of the initial guess. Notice the similar notation used for the Bellman operator T and the generic contraction mapping $\tilde{T}$; and for the discount factor β and the generic modulus $\tilde{\beta}$. To be able to apply the contraction mapping theorem, we need to show that we have a complete metric space and that the Bellman operator T is indeed a contraction mapping (with modulus β). The following theorems will help.

Theorem 2. *Let $X \subset \mathbb{R}^n$. Then the set $C(X)$ of bounded and continuous functions $f : X \to \mathbb{R}$, together with the supremum metric $d_\infty(f, g) \equiv \sup_t |f(t) - g(t)|$ is a complete linear metric space (Banach space).*

Theorem 3 (Blackwell's sufficient conditions)**.** *Let $X \subset \mathbb{R}^n$, let $B(X)$ be the set of bounded functions $f : X \to \mathbb{R}$ and let $(B(X), d_\infty)$ be the metric space composed of this set $B(X)$ and the supremum metric $d_\infty(f, g) \equiv \sup_t |f(t) - g(t)|$. Let $\tilde{T}$ be an operator satisfying:*

1. *Monotonicity: For any pair of functions $f, g \in B(X)$ such that $f(x) \leq g(x)\ \forall x \in X$, we have $\left(\tilde{T}f\right)(x) \leq \left(\tilde{T}g\right)(x)\ \forall x \in X$.*

2. *Discounting: There exists some $\tilde{\beta} \in [0, 1)$ such that $\forall f \in B(X), a \geq 0, x \in X$ we have:*
$$\left[\tilde{T}(f + a)\right](x) \leq \left(\tilde{T}f\right)(x) + \beta a.$$
(Note that $(f + a)(x) = f(x) + a$.)

Then $\tilde{T}$ is a contraction mapping with modulus $\tilde{\beta}$.

So if we can argue that the value functions $\{V_n\}_{n=0,1,2,...}$ we are considering are bounded and continuous, and that the Bellman operator T associated with our problem satisfies both monotonicity and discounting, then we know we can apply the contraction mapping theorem and therefore that we can solve our problem using value function iteration.