Athens University of Economics and Business
Department of Economics

Postgraduate Program - MSc in Economic Theory
Course: Mathematical Economics (Mathematics II)
Bellman Equations and Stationary Dynamic Programming Problems
Prof: Stelios Arvanitis
TA: Dimitris Zaverdas

Semester: Spring 2018-2019

Banach's Fixed Point Theorem and Stationary Dynamic Programming Problems

by Dimitris Zaverdas

It can be shown that if (Y, d_Y) is a compact metric space, $\phi : Y \times Y \to \mathbb{R}$ is an appropriately continuous function, and $\delta \in (0,1)$, then the Bellman equation with respect to ϕ and δ, defined as

$$\forall x \in Y \qquad f(x) = \max_{y \in Y}\{\phi(x,y) + \delta f(y)\}, \qquad \text{with } f \in \mathcal{C}(Y, \mathbb{R})$$

has a solution in $\mathcal{C}(Y, \mathbb{R})$ (say f^*) and it is unique.
A sketch of the proof involves considering a functional function $\Phi : \mathcal{C}(Y, \mathbb{R}) \to \mathbb{R}^Y$ such that

$$\forall x \in Y \qquad (\Phi(f))(x) = \max_{y \in Y}\{\phi(x,y) + \delta f(y)\}, \qquad \text{with } f \in \mathcal{C}(Y, \mathbb{R})$$

and showing that Φ is a self map, it satisfies Blackwell's conditions, and consequently Banach's conditions. Thus, it has a unique fixed point (f^*).

We have yet to establish some relationship between this result and Dynamic Programming Problems. Here we will see that finding the optimal path to a Stationary Dynamic Programming Problem is equivalent to finding a solution to a Bellman equation. For a more detailed walkthrough, study section E4 in Ok's textbook (see syllabus), pages 233-248.

A standard dynamic programming problem

Consider a standard dynamic programming problem in a general metric space, (X, d), with initial state $x_0 \in X$ where the goal is to find the optimal sequence in X, $(x_n^*)_{n \in \mathbb{N}^*}$, that maximizes some real valued objective function, U, of the following form

$$\max_{(x_n)_{n \in \mathbb{N}^*} \text{ in } X} U(x_0, (x_n)_{n \in \mathbb{N}^*}) = \phi(x_0, x_1) + \sum_{i=1}^{\infty} \delta^i \phi(x_i, x_{i+1}), \text{ such that } x_{n+1} \in \Gamma(x_n), \forall n \in \mathbb{N} \quad (1)$$

where ϕ and δ are as above and $\Gamma(x_n)$ is the set of feasible x_{n+1} given x_n.

We will assume that the above objective function is calculatable (but is not necessarily bounded) for all feasible $(x_n)_{n\in\mathbb{N}^*}$ and for all x_0, i.e.

$$\lim_{N\to\infty}\sum_{i=0}^{N}\delta^i\phi(x_i,x_{i+1})\in\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\} \tag{2}$$

We also assume that ϕ is continuous and bounded on its domain (which is the graph of Γ). Thus, the objective is also bounded (because $\delta\in(0,1)$ and $\mathcal{B}(X,\mathbb{R})$ is closed under addition).

We further assume that $\Gamma(x)$ is an appropriately compact subset of X for all $x\in X$.

Bellman's Lemma(s)

For a dynamic programming problem such as the above, define the (unknown) value function, $V : X \to \overline{\mathbb{R}}$, for any $x\in X$ such that for any initial state, $x\in X$, its value is the maximum attainable value of the objective function across all feasible sequences given x

$$V(x) := \sup_{(x_n)_{n\in\mathbb{N}^*}\text{ in }X}\{U(x,(x_n)_{n\in\mathbb{N}^*})\text{ such that }x_1\in\Gamma(x)\text{ and }x_{n+1}\in\Gamma(x_n),\forall n\in\mathbb{N}^*\} \tag{3}$$

(If a solution to our problem exists, then we can replace the sup with a max.)

Observe that if for a given x_0, $(x_n^*)_{n\in\mathbb{N}^*}$ is an optimal path for (1), (3) is equivalent to

$$V(x_0)=U(x_0,(x_n^*)_{n\in\mathbb{N}^*})=\phi(x_0,x_1^*)+\sum_{i=1}^{\infty}\delta^i\phi(x_i^*,x_{i+1}^*) \tag{4}$$

Bellman's Lemma 1

Bellman's first lemma states the following:

For a dynamic programming problem such as (1), consider its value function $V : X \to \overline{\mathbb{R}}$ such that (3) holds. For any initial state x_0, if $(x_n^*)_{n\in\mathbb{N}^*}$ is a solution to (1) for this given state, then

$$V(x_0)=\phi(x_0,x_1^*)+\delta V(x_1^*) \tag{5}$$

and

$$V(x_n^*)=\phi(x_n^*,x_{n+1}^*)+\delta V(x_{n+1}^*),\ \forall n\in\mathbb{N}^* \tag{6}$$

If ϕ is continuous and bounded, the converse is also true, i.e. if a pair $(x_0,(x_n^*)_{n\in\mathbb{N}^*})$ satisfies (5) and (6), the sequence $(x_n^*)_{n\in\mathbb{N}^*}$ is a solution to (1) with initial state x_0.

We will prove the direct statement of the lemma via mathematical induction and its converse via substitution.

For $n = 0$, if $(x_n^*)_{n\in\mathbb{N}^*}$ is the optimal path among all feasible paths given x_0, then observe that

$$\begin{aligned} V(x_0) &= \phi(x_0, x_1^*) + \sum_{i=1}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \\ &= \phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \\ &\geqslant \phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2) + \sum_{i=2}^{\infty} \delta^i \phi(x_i, x_{i+1}) \end{aligned}$$

for all feasible $(x_n)_{n\in\mathbb{N}^*}$ given x_0. Consider the last inequality and see that

$$\phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \geqslant \phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2) + \sum_{i=2}^{\infty} \delta^i \phi(x_i, x_{i+1})$$

$$\delta\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \geqslant \delta\phi(x_1^*, x_2) + \sum_{i=2}^{\infty} \delta^i \phi(x_i, x_{i+1})$$

$$\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^{i-1} \phi(x_i^*, x_{i+1}^*) \geqslant \phi(x_1^*, x_2) + \sum_{i=2}^{\infty} \delta^{i-1} \phi(x_i, x_{i+1})$$

then set $j = i - 1$ to get

$$\phi(x_1^*, x_2^*) + \sum_{j=1}^{\infty} \delta^j \phi(x_{j+1}^*, x_{j+2}^*) \geqslant \phi(x_1^*, x_2) + \sum_{j=1}^{\infty} \delta^j \phi(x_{j+1}, x_{j+2})$$

But the left hand side can be seen as the value of the objective function of a dynamic programming problem with initial state x_1^* calculated at an optimal sequence $(x_{n+1}^*)_{n\in\mathbb{N}^*}$. The value function of this "sub-problem" is identical to the starting problem. So at the solution

$$V(x_1^*) = \phi(x_1^*, x_2^*) + \sum_{j=1}^{\infty} \delta^j \phi(x_{j+1}^*, x_{j+2}^*)$$

Thus,

$$\begin{aligned}
V(x_0) &= \phi(x_0, x_1^*) + \sum_{i=1}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \\
&= \phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*) \\
&= \phi(x_0, x_1^*) + \delta \left(\phi(x_1^*, x_2^*) + \sum_{i=2}^{\infty} \delta^{i-1} \phi(x_i^*, x_{i+1}^*) \right) \\
&= \phi(x_0, x_1^*) + \delta \left(\phi(x_1^*, x_2^*) + \sum_{j=1}^{\infty} \delta^{j} \phi(x_{j+1}^*, x_{j+2}^*) \right) \\
&= \phi(x_0, x_1^*) + \delta V(x_1^*)
\end{aligned}$$

Now, assume that for $n = k > 0$ the desired equation holds

$$\begin{aligned}
V(x_k^*) &= \phi(x_k^*, x_{k+1}^*) + \delta V(x_{k+1}^*) \\
\phi(x_k^*, x_{k+1}^*) + \sum_{i=1}^{\infty} \delta^i \phi(x_{i+k+1}^*, x_{i+k+2}^*) &= \phi(x_k^*, x_{k+1}^*) + \delta V(x_{k+1}^*) \\
\sum_{i=1}^{\infty} \delta^i \phi(x_{i+k+1}^*, x_{i+k+2}^*) &= \delta V(x_{k+1}^*) \\
\delta\phi(x_{k+1}^*, x_{k+2}^*) + \sum_{i=2}^{\infty} \delta^i \phi(x_{i+k+1}^*, x_{i+k+2}^*) &= \delta V(x_{k+1}^*) \\
\phi(x_{k+1}^*, x_{k+2}^*) + \sum_{i=2}^{\infty} \delta^{i-1} \phi(x_{i+k+1}^*, x_{i+k+2}^*) &= V(x_{k+1}^*) \\
\phi(x_{k+1}^*, x_{k+2}^*) + \sum_{j=1}^{\infty} \delta^{j} \phi(x_{j+k+2}^*, x_{j+k+3}^*) &= V(x_{k+1}^*)
\end{aligned}$$

which means that the equation also holds for $n = k + 1$.

Thus, solutions for the dynamic programming problem satisfy (5) and (6).

For the converse, consider a pair $(x_0, (x_n^*)_{n\in\mathbb{N}^*})$ such that (5) and (6) hold. Then

$$\begin{aligned}
V(x_0) &= \phi(x_0, x_1^*) + \delta V(x_1^*) \\
&= \phi(x_0, x_1^*) + \delta\phi(x_1^*, x_2^*) + \delta^2 V(x_2^*) \\
&= \ldots \\
&= \phi(x_0, x_1^*) + \sum_{i=1}^{k} \delta^i \phi(x_i^*, x_{i+1}^*) + \delta^k V(x_k^*)
\end{aligned}$$

By our assumptions above (ϕ is continuous and bounded) V is bounded, so

$$\lim_{k\to\infty} \delta^k V(x_k^*) = 0$$

and by letting k approach infinity we get

$$V(x_0) = \phi(x_0, x_1^*) + \sum_{i=1}^{\infty} \delta^i \phi(x_i^*, x_{i+1}^*)$$

which implies that $(x_n^*)_{n\in\mathbb{N}^*}$ is a solution to the dynamic programming problem with initial state x_0.

Bellman's Lemma 2

Bellman's second lemma, known as the Principle of Optimality, states that:
For any dynamic programming problem such as (1) and a function $V \in \mathcal{B}(X, \mathbb{R})$ such that

$$V(x) = \max_{y\in\Gamma(x)} \{\phi(x,y) + \delta V(y)\},\ \forall x \in X \tag{7}$$

it holds that

$$V(x) = \max_{(x_n)_{n\in\mathbb{N}^*} \text{ in } X \text{ with } x_{n+1}\in\Gamma(x_n)} \{U(x, (x_n)_{n\in\mathbb{N}^*})\},\ \forall x \in X \tag{8}$$

To show this, consider a bounded real function on X, V, that satisfies (7). For any $x \in X$ and any feasible sequence $(x_n)_{n\in\mathbb{N}}$, observe that

$$\begin{aligned} V(x) &\geqslant \phi(x, x_1) + \delta V(x_1) \\ &\geqslant \phi(x, x_1) + \delta\phi(x_1, x_2) + \delta^2 V(x_2) \\ &\geqslant ... \\ &\geqslant \phi(x, x_1) + \sum_{i=1}^{k} \delta^i \phi(x_i, x_{i+1}) + \delta^{k+1} V(x_{k+1}) \end{aligned}$$

by the definition of (7).
Since by our assumptions V is bounded, by letting k approach infinity, we get

$$V(x) \geqslant \phi(x, x_1) + \sum_{i=1}^{\infty} \delta^i \phi(x_i, x_{i+1}) = U(x, (x_n)_{n\in\mathbb{N}})$$

over all $x \in X$ and feasible sequences.

Now for any $x \in X$ consider a feasible sequence, $(x_n^*)_{n\in\mathbb{N}}$, such that

$$V(x) = \phi(x, x_1^*) + \delta V(x_1^*)$$

and

$$V(x_n^*) = \phi(x_n^*, x_{n+1}^*) + \delta V(x_{n+1}^*),\ \forall n \in \mathbb{N}^*$$

Such a sequence exists for all x, as maximizers of the right hand side of (7) with the appropriate x. By iterative substitution

$$V(x) = \phi(x, x_1) + \sum_{i=1}^{k} \delta^i \phi(x_k, x_{k+1}) + \delta^{k+1} V(x_{k+1})$$

and because V is bounded

$$V(x) = \phi(x, x_1) + \sum_{i=1}^{\infty} \delta^i \phi(x_i, x_{i+1}) = U(x, (x_n^*)_{n\in\mathbb{N}^*})$$

Thus, for all $x \in X$, $V(x)$ is the maximum value attainable for $U(x, (x_n)_{n\in\mathbb{N}^*})$ at that x across all feasible consequent sequences.

Let us dissect what all this entails.

The converse of Bellman's first lemma tells us that if we have the functional forms of the value functions of a dynamic programming problem such as (1) and ϕ is continuous and bounded, then the value functions are also bounded and for each initial state, x_0, and value function we can find optimal sequences that maximize the objective by leveraging (5) and (6).

The principle of optimality says that bounded solutions to Bellman equations are value functions of the associated dynamic programming problem.

We have seen that Bellman equations always have a unique continuous function as a solution, when the domain of the function is appropriately compact.

Thus, for any dynamic programming problem such as (1) with a continuous and bounded ϕ and X and $\Gamma(X)$ compact, there exists a unique continuous value function, V, which is related to the solutions of the problem through (5) and (6).

An interesting question that remains is when is the optimal path unique. For this, it would suffice that V be a strictly concave function on X. Then there exists a unique maximizer, x_1^*, of (5) given x_0, and a unique maximizer of (6) given x_n^*.

To end up with a strictly concave value function, the following assumptions are sufficient:

- the graph of Γ (which is a subset of $X \times \Gamma(X)$) is a convex set, i.e. $\forall \lambda \in (0,1)$ and $\forall x, x' \in X$, for any $y \in \Gamma(x)$ and $y' \in \Gamma(x')$

$$\lambda \cdot y + (1-\lambda) \cdot y' \in \Gamma(\lambda \cdot x + (1-\lambda) \cdot x')$$

- ϕ is concave on the graph of Γ, i.e. $\forall \lambda \in (0,1)$ and $\forall x, x' \in X$, for any $y \in \Gamma(x)$ and $y' \in \Gamma(x')$

$$\phi(\lambda \cdot (x,y) + (1-\lambda) \cdot (x',y')) > \lambda \cdot \phi(x,y) + (1-\lambda) \cdot \phi(x',y')$$

(We need the graph of Γ to be convex so that we can talk about the concavity of ϕ without worrying about missing points.)

We will not show here how the above assumptions give a strictly concave value function, but you probably have some general intuition as to why that is. For the derivation, see section E4.3 in Ok's textbook.

An Optimal Growth Problem

The following is an applied example of an optimal growth problem (based on example 5 in Ok's section E4.2).

Consider a dynamic economy with a representative agent.

In each period, t, the agent has the output from last period's production, $y_t = f(k_t) = \sqrt{k_t}$, and decides how much of it will be saved as an input for next period's production, k_{t+1}, and the rest of it is consumed yielding utility $u(c_t) = ln(c_t)$. The starting capital, k_0, and resulting output, $y_0 = f(k_0)$, are given endowments with $k_0 \in [0,1]$.

Naturally, the amount of capital saved cannot exceed the available output, $0 \leqslant k_{k+1} \leqslant y_t$ and we have that $\forall t \in \mathbb{N}\ y_t \geqslant c_t + k_{t+1}$. Of course, the agent being rational, they consume all of the output that is not saved for future production and $\forall t \in \mathbb{N}\ c_t = y_t - k_{t+1} = f(k_t) - k_{t+1}$.

The agent's discount factor is $\frac{1}{2}$ and their objective is to maximize their intertemporal utility subject to the constraint discussed above

$$\max_{\{k_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \left(\frac{1}{2}\right)^t u(c_t) \text{ such that } 0 \leqslant k_{t+1} \leqslant y_t\ \forall t \in \mathbb{N} \tag{9}$$

To reduce notation, we can rewrite the above as

$$\max_{\{k_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \left(\frac{1}{2}\right)^t ln(\sqrt{k_t} - k_{t+1}) \text{ such that } k_{t+1} \in [0, \sqrt{k_t}]\ \forall t \in \mathbb{N} \tag{10}$$

We can draw parallels to the discussion that preceded and see that this is a standard dynamic programming problem. The initial state is k_0 and the sequence over which we maximize is $(k_t)_{t\in\mathbb{N}^*}$. $\phi(k_t, k_{t+1}) = ln(\sqrt{k_t} - k_{t+1})$ and $\delta = \frac{1}{2}$. Finally, $X = [0,1]$ and $\Gamma(k_t) = [0, \sqrt{k_t}]$.
By the Principle of Optimality, it suffices to find the solutions, V, to the following functional equation

$$g(x) = \max_{y\in\Gamma(x)} \{ln(\sqrt{x} - y) + \frac{1}{2}g(y)\}, \, \forall x \in [0,1] \tag{11}$$

Because u is continuous and bounded on the graph of Γ and $\Gamma(x)$ is compact for all x, we have seen that by Banach's Fixed Point Theorem, V is the unique fixed point of the following function, Φ, defined by

$$(\Phi(g))(x) = \max_{y\in\Gamma(x)} \{ln(\sqrt{x} - y) + \frac{1}{2}g(y)\}, \, \forall x \in [0,1] \tag{12}$$

with $g \in \mathcal{B}(\Gamma(x), \mathbb{R})$ and can be found as

$$V = \lim_{n\to\infty} \Phi^{(n)}(g) \tag{13}$$

for any starting $g \in \mathcal{B}(\Gamma(x), \mathbb{R})$.
Start with $g : [0, \sqrt{x}] \to [0, \sqrt{x}]$ such that $\forall x$, $g(y) = 0 \, \forall y \in [0, \sqrt{x}]$. Thus,

$$\begin{aligned}
(\Phi(g))(x) &= \max_{y\in[0,\sqrt{x}]} \left\{ ln(\sqrt{x} - y) + \frac{1}{2} \cdot 0 \right\}, && \forall x \in [0,1] \\
&= \max_{y\in[0,\sqrt{x}]} \{ln(\sqrt{x} - y)\}, && \forall x \in [0,1] \\
&= ln(\sqrt{x}), && \forall x \in [0,1] \\
&= \frac{1}{2} ln(x), && \forall x \in [0,1]
\end{aligned}$$

Then use $\Phi(g)$ to get

$$\begin{aligned}
(\Phi(\Phi(g)))(x) = (\Phi^{(2)}(g))(x) &= \max_{y\in[0,\sqrt{x}]} \left\{ ln(\sqrt{x} - y) + \frac{1}{2} \cdot \frac{1}{2} ln(y) \right\}, && \forall x \in [0,1] \\
&= ... \\
&= \frac{5}{8} ln(x) + \left(2ln(2) - \frac{5}{4} ln(5) \right), && \forall x \in [0,1]
\end{aligned}$$

Continuing indefinitely we can see that as n increases it has the following form

$$(\Phi^{(n)}(g))(x) = \alpha ln(x) + \beta, \, \forall x \in [0,1] \tag{14}$$

A "wise" guess (though not certain) would be that the limit of $\Phi^{(n)}(g)$ as n approaches infinity - which is the fixed point of the functional equation and the value function of the dynamic programming problem - is also of the same form. So we substitute to get

$$\begin{aligned} V(x) &= \max_{y\in[0,\sqrt{x}]} \left\{ ln(\sqrt{x}-y) + \frac{1}{2}\cdot V(y) \right\}, & \forall x \in [0,1] \\ \alpha ln(x) + \beta &= \max_{y\in[0,\sqrt{x}]} \left\{ ln(\sqrt{x}-y) + \frac{1}{2}\cdot (\alpha ln(y)+\beta) \right\}, & \forall x \in [0,1] \end{aligned}$$

It can be found that, given $x \in [0,1]$, the right hand side is maximized at $y^* = \dfrac{\alpha}{2+\alpha}\sqrt{x}$. By substituting and rearranging we get

$$\alpha ln(x) + \beta = \left(\frac{1}{2}+\frac{\alpha}{4}\right) ln(x) + ln\left(1-\frac{\alpha}{2+\alpha}\right) + \frac{\alpha}{2} ln\left(\frac{\alpha}{2+\alpha}\right) + \frac{\beta}{2},\ \forall x \in [0,1] \tag{15}$$

which implies the system of equations on α and β

$$\left\{ \begin{array}{l} \alpha = \dfrac{1}{2} + \dfrac{\alpha}{4} \\ \beta = ln\left(1-\dfrac{\alpha}{2+\alpha}\right) + \dfrac{\alpha}{2} ln\left(\dfrac{\alpha}{2+\alpha}\right) + \dfrac{\beta}{2} \end{array} \right\} \tag{16}$$

If a solution to this system exists, then we were right to assume that the limit is of this form and we will have found the value function in question. A unique solution does exist, it is $\alpha = \dfrac{2}{3}$ and $\beta = ln(9) - \dfrac{8}{3} ln(4)$, and the value function is

$$V(x) = \frac{2}{3} ln(x) + ln(9) - \frac{8}{3} ln(4),\ \forall x \in [0,1] \tag{17}$$

and as already seen, given x, the optimal $y \in \Gamma(x)$ is

$$y^* = \frac{\alpha}{2+\alpha}\sqrt{x} = \frac{1}{4}\sqrt{x}$$

Thus, we have also found the optimal policy function.

So the optimal path for the representative agent, given k_0, is

$$(k_t^*)_{t\in\mathbb{N}^*} = \left(\frac{1}{4}\sqrt{k_0},\quad \frac{1}{4}\sqrt{\frac{1}{4}\sqrt{k_0}},\quad \frac{1}{4}\sqrt{\frac{1}{4}\sqrt{\frac{1}{4}\sqrt{k_0}}},\quad ... \right) \tag{18}$$