Athens University of Economics and Business
Department of Economics

Postgraduate Program - MSc in Economic Theory
Course: *Mathematical Economics (Mathematics II)*
Prof: Stelios Arvanitis
TA: Dimitris Zaverdas*

Semester: Spring 2023-2024

Bellman Equations, Banach's Fixed Point Theorem, and Stationary Dynamic Programming Problems

by Dimitris Zaverdas

Let (X, d) be a compact metric space, $\phi : X \times X \to \mathbb{R}$ an appropriately continuous function, and $\delta \in (0, 1)$. The Bellman equation with respect to ϕ and δ is a functional equation, defined as

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

where the function, $f \in \mathcal{C}(X, \mathbb{R})$, plays the role of the equation's "variable"[1].

It can be shown that this has a unique solution, f^*, in $\mathcal{C}(X, \mathbb{R})$. A sketch of the proof is the following. Consider a functional function $\Phi : \mathcal{C}(X, \mathbb{R}) \to \mathbb{R}^X$ such that

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

Observe that Φ is a self map. Show that it satisfies Blackwell's conditions, and consequently Banach's conditions. Thus, the functional equation has a unique fixed point, f^*, which can be shown to coincide with the solution to the Bellman equation.

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), pp. 233-248.

*Please report any typos, mistakes, or even suggestions at zaverdasd@aueb.gr.

[1]Same as the real number, $x \in \mathbb{R}$, plays the role of the variable in the following real equation with respect to α and β

$$x = \alpha + \beta x$$

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 an optimal sequence of future states in X, $(x_n^*)_{n\in\mathbb{N}^*}$, which maximizes some real valued objective function, U, under a set of feasibility constraints, of the following form

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

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

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

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

We also assume that ϕ is continuous and bounded on its domain. 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)

Let us define the value function. For a dynamic programming problem such as (1), **a** value function, $V : X \to \overline{\mathbb{R}}$ (which may have an unknown functional form), is defined as **any** function such that for any $x \in X$ its value is the maximum attainable value of the objective function across all feasible sequences of future states given x as the initial state, i.e.,

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

(If a solution to (1) exists, then we can replace the sup with a max.)

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

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

We now proceed to stating and proving Bellman's lemmata.

Bellman's Lemma 1

For a dynamic programming problem such as (1), consider a 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.

Proof

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,

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

and we have arrived at equation (5).

Now assume that the equation in condition (6) holds for some $n = k > 0$

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

which means that the equation also holds for $n = k+1$, and by induction the condition in (6) holds for all $n \in \mathbb{N}^*$.

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

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

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

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

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

Bellman's Lemma 2 (The Principle of Optimality)

For any dynamic programming problem such as (1) and a function $V \in \mathcal{B}(X, \mathbb{R})$, which satisfies the following Bellman equation

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

it holds that V is a value function of (1) and

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

Proof

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

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

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.

Discussion

Let us examine what we have so far.

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 value function and initial state, x_0, we can find optimal sequences of future states that maximize the objective by leveraging (5) and (6).

Bellman's second lemma (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 space 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 regards the uniqueness of the optimal. 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 have 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, $0 \leqslant k_{t+1} \leqslant y_t$ 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 $\Gamma(k_t) = [0, \sqrt{k_t}]$.

For the remainder of this solution, to simplify notation, define $x \coloneqq k_t$ and $y \coloneqq k_{t+1}$.

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

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

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

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

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 solution does exist and it is unique. It is $\alpha = \frac{2}{3}$ and $\beta = ln(9) - \frac{8}{3} ln(4)$, 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

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

e.g. if $k_0 = 1$, then

$$(k_t^*)_{t \in \mathbb{N}^*} = \left(\frac{1}{4}, \quad \frac{1}{8}, \quad \frac{1}{8}\sqrt{\frac{1}{2}}, \quad \dots\right)$$