Dynamic Programming

Thomas J. Sargent and John Stachurski

Longye Tian
longye.tian@anu.edu.au

Australian National University
School of Economics

March 7th, 2025

Outline

Nonstationary policies

In this section, we will see that under some conditions, the lifetime value of any nonstationary policy will be weakly dominated by the lifetime value of a stationary policy. This ensures that we can focus on the stationary policies without loss of generality.

Comparison

Stationary policy

- Fixed a policy σ
- Lifetime value

$$v_\sigma = \lim_{j \to \infty} T_\sigma^j v$$

Nonstationary policy/Policy Plan

- a policy plan $\overline{\sigma} = (\sigma_t)_{t \geq 0} \in \bigtimes_{t \geq 0} \Sigma$
- Lifetime value of $v_{\overline{\sigma}}$

$$v_{\overline{\sigma}} = \lim_{j \to \infty} T_{\sigma_0} T_{\sigma_1} \cdots T_{\sigma_j} v$$

- Question, why not

$$v_{\overline{\sigma}} = \lim_{j \to \infty} T_{\sigma_j} \cdots T_{\sigma_1} T_{\sigma_0} v$$

Existence of Lifetime value of a policy plan

We want the limit to exist and, ideally, the limit is independent of v.

$$v_{\overline{\sigma}} = \lim_{j \to \infty} T_{\sigma_0} T_{\sigma_1} \cdots T_{\sigma_j} v$$

Set up

- $V = (V, \precsim)$ a partially ordered **space**
- $\mathbb{T} = \{T_\sigma : \sigma \in \Sigma\}$, family of order preserving self-map on V
- Metric d on V
  - d is complete (Every Cauchy sequence converges.)
  - $\exists \lambda \in (0, 1)$ such that

$$d(T_\sigma v, T_\sigma w) \leq \lambda d(v, w) \qquad \text{for all } v, w \in V, \sigma \in \Sigma$$

  - for all $v \in V$, we have

$$\sup_{\sigma \in \Sigma} d(v, T_\sigma v) < \infty$$

  - sup-nonexpansive, for any subsets (v_α) and (w_α) in V such that their supremum exists,

$$d\left(\bigvee_\alpha v_\alpha, \bigvee_\alpha w_\alpha\right) \leq \sup_\alpha d(v_\alpha, w_\alpha)$$

Lemma 1.4.1.(i)

Lemma 1

If the above conditions hold, then for each $v \in V$ and policy plan $\hat{\sigma} := (\sigma_t)_{t\geq 0}$, the limit

$$v_{\hat{\sigma}} = \lim_{n\to\infty} \bigtimes_{t=0}^{n} T_{\sigma_t} v$$

exists in V and is independent of v.

Proof for Existence.

To prove that (v_n) is Cauchy sequence, where $v_n := \bigtimes_{t=0}^{n} T_{\sigma_t} v$ $\square$

Lemma 1.4.1.(i) Continue

Proof.

Fix $v \in V$, $\hat{\sigma} = (\sigma_t)_{t \geq 0}$, $\epsilon > 0$. Let $T_{m,n} := \bigtimes_{t=m}^{t=n} T_{\sigma_t}$, $v_n = T_{0,n} v$.
For $m \in \mathbb{N}$, we have

$$
\begin{aligned}
d(v_m, v_{m+1}) &= d\Big(T_0(T_{1,m} v), T_0(T_{1,m+1} v)\Big) \\
&\leq \lambda d\Big(T_1(T_{2,m} v), T_1(T_{2,m+1} v)\Big) && \text{(contraction)} \\
&\vdots \\
&\leq \lambda^{m+1} d(v, T_{m+1} v) \\
&\leq \lambda^{m+1} b_v && \text{(bounded)}
\end{aligned}
$$

$\square$

Lemma 1.4.1.(i) Continue

Proof.

WLOG, let $m, n, j \in \mathbb{N}, n = m + j, j \geq 0$.

$$
\begin{aligned}
d(v_m, v_{m+j}) &\leq d(v_m, v_{m+1}) + d(v_{m+1}, v_{m+2}) + \cdots + d(v_{m+j-1}, v_{m+j}) && (\Delta \text{ inequality}) \\
&\leq \lambda^{m+1} b_v + \lambda^{m+2} b_v + \cdots + \lambda^{m+j} b_v && (\text{page } 7) \\
&\leq \lambda^{m+1} b_v (1 + \lambda + \cdots + \lambda^{m+j-1}) \\
&\leq \lambda^{m+1} b_v (1 + \lambda + \cdots + \lambda^{m+j-1} + \cdots) \\
&\leq \lambda^{m+1} b_v / (1 - \lambda) && (\text{geom sum})
\end{aligned}
$$

$\implies$ Cauchy. By completeness, we get the limit exists. $\square$

Lemma 1.4.1.(i) Continue

The limit is independent of v.

Let $v, w \in V$. Then

$$\begin{aligned} d(v_n, w_n) &= d\Big(T_0(T_{1,n}v), T_0(T_{1,n}w)\Big) \\ &\leq \lambda d\Big(T_1(T_{2,n}v), T_1(T_{2,n}w)\Big) && \text{(contraction)} \\ &\vdots \\ &\leq \lambda^{n+1} d(v, w) \end{aligned}$$

So (v_n) and (w_n) have the same limit. $\square$

Lemma 1.4.1.(ii)

Lemma 2

If the conditions in page 5 holds, every $T_\sigma \in \mathbb{T}$ is continuous, globally stable on V, with unique fixed point v_σ satisfying

$$v_\sigma = \lim_{j \to \infty} T_\sigma^j v \qquad \text{for all } v \in V$$

Proof.

From contraction and completeness. □

Lemma 1.4.1.(iii)

Lemma 3

If the conditions in page 5 holds, there exists a $v \in V$ *such that* $v := \bigvee_{\sigma\in\Sigma} T_\sigma v$

Proof.

If T is well-defined on V, then for $v, w \in V$, we have

$$\begin{aligned} d(Tv, Tw) &= d\left(\bigvee_{\sigma\in\Sigma} T_\sigma v, \bigvee_{\sigma\in\Sigma} T_\sigma w\right) \\ &\leq \sup_{\sigma\in\Sigma} d(T_\sigma v, T_\sigma w) && \text{(sup-nonexpansionary)} \\ &\leq \lambda d(v, w) && \text{(contraction)} \end{aligned}$$

Hence, T is a contraction, therefore, it has at least one fixed point in V. □

Question: Do we need $v \in V_G$? Is T continuous, globally stable with unique fixed point?

Review Theorem 1.3.3.

Theorem 4

Let V be a pospace, $(V, \mathbb{T})$ be regular, globally stable and T has a fixed point in V, then

- *the fundamental optimality properties hold*
- *VFI, HPI, OPI converge.*

Review of the Fundamental Optimality Properties

Let $(V, \mathbb{T})$ be regular and well-posed. We say **the fundamental optimality properties hold** for $(V, \mathbb{T})$ if

(B1) at least one optimal (stationary) policy exists

(B2) $v^* := \bigvee_\sigma v_\sigma$ is the unique solution to the Bellman equation

(B3) Bellman's principle of optimality holds (optimal policy is v^*-greedy)

Review on Pospace

A partial order $\precsim$ on topological space V is called **closed** if, given any two nets $(u_\alpha)_{\alpha\in\Lambda}$ and $(v_\alpha)_{\alpha\in\Lambda}$ contained in V,

$$u_\alpha \to u, v_\alpha \to v \qquad \text{and } u_\alpha \precsim v_\alpha \text{ for all } \alpha \in \Lambda \implies u \precsim v$$

A **partially ordered space**, is a Hausdorff topological space **endowed with a closed partial order**.

Proposition 1.4.2. Any policy plan is weakly dominated by a stationary policy

Proposition 1.4.2.

If $(V, \mathbb{T})$ is regular and conditions in page 5 holds, then

- the fundamental optimality properties hold
- Given any policy plan $\overline{\sigma}$, there exists a stationary policy plan σ such that $v_{\overline{\sigma}} \precsim v_\sigma$

□

Proof.

Part One from Theorem 1.3.3. Part Two: Fix a policy plan $\overline{\sigma}$ and let σ be the optimal policy(B1). Then, for all $j \in \mathbb{N}$, we have

$$T_{\sigma_0} T_{\sigma_1} \cdots T_{\sigma_j} v_\sigma \to v_{\overline{\sigma}}, T^j v_\sigma \to v_\sigma, \qquad T_{\sigma_0} T_{\sigma_1} \cdots T_{\sigma_j} v_\sigma \precsim T^j v_\sigma = v_\sigma$$

The partial order is closed implies $v_{\overline{\sigma}} \precsim v_\sigma$. □

Minimization Problem

For a given ADP $(V, \mathbb{T})$, a minimization problem can be converted to a maximization problem by reversing the partial order on V. Hence, we can focus on solving the maximization problem without loss of generality.

Definitions

Let $(V, \mathbb{T})$ be an ADP with policy set Σ. We define

- **Bellman min-operator** $T_\perp$ by

$$T_\perp v = \bigwedge_{\sigma \in \Sigma} T_\sigma v \qquad \text{whenever the infimum exists}$$

- $\sigma \in \Sigma$ is **v-min-greedy** if $T_\sigma v \precsim T_\tau v$ for all $\tau \in \Sigma$
- $(V, \mathbb{T})$ is **min-regular** if, for each $v \in V$, at least one v-min-greedy policy exists (V_G or V_G^{min})
- v satisfies the **Bellman min-equation** if $v = T_\perp v$

Definitions

Suppose $(V, \mathbb{T})$ is well-posed. We define

- **min-value function** by

$$v_\perp^* = \bigwedge_{\sigma \in \Sigma} v_\sigma \quad \text{whenever the infimum exists}$$

- $\sigma \in \Sigma$ is **min-optimal** for $(V, \mathbb{T})$ if $v_\sigma = v_\perp^*$
- $(V, \mathbb{T})$ obeys **Bellman's principle of min-optimality** if

$$\sigma \in \Sigma \text{ is min-optimal for } (V, \mathbb{T}) \iff \sigma \text{ is } v_\perp^*\text{-min-greedy}$$

Definitions

We say that the **fundamental min-optimality properties hold** if

(B1') at least one min-optimal policy exists

(B2') $v^*_\perp$ is the unique solution to the Bellman min-equation in V

(B3') Bellman's principle of min-optimality holds.

Definitions

When $(V, \mathbb{T})$ is min-regular, we define the **Howard policy min-operator** corresponding to $(V, \mathbb{T})$ via

$$H_\perp : V_G \to V_\Sigma, \quad H_\perp v = v_\sigma \quad \text{where } \sigma \text{ is } v\text{-min-greedy}$$

For each $m \in \mathbb{N}$, the **optimistic policy min-operator** via

$$W_\perp : V_G \to V, \quad W_\perp v = T_\sigma^m v \quad \text{where } \sigma \text{ is } v\text{-min-greedy}$$

Let V_D be all $v \in V$ with $T_\perp v \precsim v$. We say that

- **min-VFI converges** if $T_\perp^n \downarrow v_\perp^*$ for all $v \in V_D$
- **min-OPI converges** if $W_\perp^n v \downarrow v_\perp^*$ for all $v \in V_D$ and all $m \in \mathbb{N}$
- **min-HPI converges** if $H_\perp^n v \downarrow v_\perp^*$ for all $v \in V_D$.

Dual ADPs

How minimization problems can be converted to maximization problem in this abstract setting?

Order dual and Dual ADP

Definition 5

Given partially ordered set V, let $V^\partial = (V, \precsim^\partial)$ be the **order dual** (also called the **dual**), so that, for $u, v \in V$, we have

$$u \precsim^\partial v \iff v \precsim u$$

Definition 6

For ADP $(V, \mathbb{T})$, we call $(V, \mathbb{T})^\partial := (V^\partial, \mathbb{T})$ the **dual** of $(V, \mathbb{T})$. In other words, the dual ADP is created by replacing the poset V with its order dual V^∂.

Exercise 1.4.1

Show that $(V, \mathbb{T})^\partial$ is an ADP.

Proof.

We need to show that V^∂ is a poset. And T_σ is order-preserving self-map on V^∂ for any $\sigma \in \Sigma$. Let $u, v, w \in V$, we have

- (Reflexivity) $u \precsim u \implies u \precsim^\partial u$
- (Antisymmetry) $u \precsim^\partial v, v \precsim^\partial u \implies v \precsim u, u \precsim v \implies u = v$
- (Transitivity) $u \precsim^\partial v, v \precsim^\partial w \implies w \precsim v, v \precsim u \implies w \precsim u \implies u \precsim^\partial w$.

Hence V^∂ is a poset. Let $u, v \in V$ and $u \precsim^\partial v$. We have $v \precsim u \implies T_\sigma v \precsim T_\sigma u \implies T_\sigma u \precsim^\partial T_\sigma v$ for any $\sigma \in \Sigma$. Hence, T_σ is order-preserving self map on V^∂. □

Notation for the dual ADP

For the dual ADP $(V, \mathbb{T})^\partial$,

- the Bellman max-operator will be denoted by T^∂
- the Bellman min-operator will be denoted by $T_\perp^\partial$
- the max-value function will be denoted by $(v^*)^\partial$

Self-Dual

Each ADP is a self-dual, i.e.,

$$((V, \mathbb{T})^{\partial})^{\partial} = (V, \mathbb{T})$$

This follows from the fact that all partially ordered sets are order self-dual.

Exercise 1.4.2 (i)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

$$\sigma \text{ is } v\text{-min-greedy for } (V, \mathbb{T}) \text{ if and only if } \sigma \text{ is } v\text{-max-greedy for } (V, \mathbb{T})^\partial$$

Proof ($\iff$)

$T_\sigma v \precsim T_\tau v$ for all $\tau \in \Sigma \iff T_\tau v \precsim^\partial T_\sigma v$ for all $\tau \in \Sigma$. $\square$

Exercise 1.4.2 (ii)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

$$(V, \mathbb{T}) \text{ is min-regular if and only if } (V, \mathbb{T})^\partial \text{ is max-regular}$$

Proof.

By Exercise 1.4.2 (i) $\square$

Exercise 1.4.2.(iii)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

If $T^\partial v$ exists then so does $T_\perp v$, and, moreover, $T_\perp v = T^\partial v$

Proof.

By the definition of Bellman max-operator, we have $T_\sigma v \precsim^\partial T^\partial v$ for all $\sigma \in \Sigma$, i.e., $T^\partial v \precsim T_\sigma v$ for all $\sigma \in \Sigma$. Hence, $T_\perp v$ exists and equals to $T^\partial v$ by definition. $\square$

Exercise 1.4.2.(iv)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

If $W^\partial v$ exists then so does $W_\perp v$, and, moreover, $W_\perp v = W^\partial v$

Proof.

By definition, we have $W^\partial v = T_\sigma v$ where σ is v-max-greedy for $(V, \mathbb{T})^\partial$. Hence by Exercise 1.4.2.(i), σ is v-min-greedy for $(V, \mathbb{T})$. Hence, we have $W^\partial v = W_\perp v$. $\square$

Exercise 1.4.2.(v)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

If $H^\partial v$ exists then so does $H_\perp v$, and, moreover, $H_\perp v = H^\partial v$

Proof.

By definition, we have $H^\partial v = v_\sigma$ where σ is v-max-greedy for $(V, \mathbb{T})^\partial$. Hence by Exercise 1.4.2.(i), σ is v-min-greedy for $(V, \mathbb{T})$. Hence, we have $H^\partial v = H_\perp v$. $\square$

Exercise 1.4.2.(vi)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that: If the max-value function $(v^*)^\partial$ exists for $(V, \mathbb{T})^\partial$ then the min-value function $v^*_\perp$ exists for $(V, \mathbb{T})$, and, moreover,$v^*_\perp = (v^*)^\partial$.

Proof.

By definition,

$$(v^*)^\partial = \bigvee_{\sigma \in \Sigma}^{\partial} v_\sigma = \bigwedge_{\sigma \in \Sigma} v_\sigma = v^*_\perp$$

following Exercise A.1.15. □

Exercise 1.4.2.(vii)

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. Fix $v \in V$ and verify that:

$$\sigma \in \Sigma \text{ is min-optimal for } (V, \mathbb{T}) \text{ if and only if } \sigma \text{ is max-optimal for } (V, \mathbb{T})^\partial$$

Proof.

σ is min-optimal for $(V, \mathbb{T})$ if and only if $v_\sigma = v_\perp^* = (v^*)^\partial$, i.e., σ is max-optimal for $(V, \mathbb{T})^\partial$ from Exercise 1.4.2.(vi). $\square$

Optimality and Convergence

Theorem 7

Let $(V, \mathbb{T})$ be a well-posed ADP with dual $(V, \mathbb{T})^\partial$. The fundamental max-optimality properties hold for $(V, \mathbb{T})^\partial$ if and only if the fundamental min-optimality properties hold for $(V, \mathbb{T})$. Moreover,

(i) *max-VFI converges for $(V, \mathbb{T})^\partial$ if and only if min-VFI converges for $(V, \mathbb{T})$*

(ii) *max-OPI converges for $(V, \mathbb{T})^\partial$ if and only if min-OPI converges for $(V, \mathbb{T})$*

(iii) *max-HPI converges for $(V, \mathbb{T})^\partial$ if and only if min-HPI converges for $(V, \mathbb{T})$*