ECON-GA 1025 Macroeconomic Theory I

Lecture 14

John Stachurski

Fall Semester 2018

Today's Lecture

This lecture is for reference only

It is **not subject to assessment**

Contents:

- General dynamic programming theory
- Bellman's principle of optimality

Further details on all this material can be found in the course notes

Dynamic Programming: General Theory

Key questions:

- When does Bellman's principle of optimality hold?
- When do optimal policies exist and how can we compute them?

We address these issues in an abstract setting that includes

- All infinite horizon applications covered to date
- Additional applications with nonstandard preferences

An **abstract Markov decision process** (AMDP) is

1. a set X called the **state space**
2. a set A called the **action space**
3. a nonempty correspondence Γ from X to A called the **feasible correspondence**, with **feasible state-action pairs**

$$\mathbb{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}$$

4. a subset $\mathcal{V}$ of $\mathbb{R}^{\mathsf{X}}$ called the set of **candidate value functions** and
5. a **state-action aggregator**

$$Q : \mathbb{G} \times \mathcal{V} \to \mathbb{R}$$

Interpretation:

In each period, controller observes $x \in \mathsf{X}$ and responds with $a \in \mathsf{A}$

$\Gamma(x) =$ all actions available to the controller in state x

Examples.

- all possible consumption choices given wealth w
- stop or continue in an optimal stopping problem
- order stock or don't order (firm inventory problem)

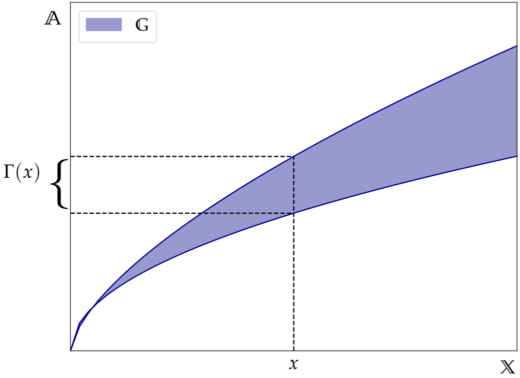

Figure: Γ and $\mathbb{G}$ when $\mathsf{A} = \mathsf{X} = \mathbb{R}_+$

Loosely speaking, $Q(x, a, v) =$ **RHS of the Bellman equation**

In other words, $Q(x, a, v) =$ total lifetime rewards, contingent on

- current action a,
- current state x
- use of v to evaluate future states

Assumption. (Monotonicity) The state-action aggregator Q satisfies

$$v \leqslant v' \implies Q(x, a, v) \leqslant Q(x, a, v') \quad \text{whenever } (x, a) \in \mathbb{G}$$

Example. Consider the generic optimal savings model

- state is $x \in \mathsf{X}$
- the action is $c \in \Gamma(x)$
- $\mathbb{G} = \{(x, c) \in \mathsf{X} \times \mathbb{R}_+ : c \in \Gamma(x)\}$
- Bellman equation is

$$v(x) = \max_{c \in \Gamma(x)} \left\{ u(c) + \beta \int v(g(x, c, z)) \varphi(\mathrm{d}z) \right\} \qquad (x \in \mathsf{X})$$

Maps directly to the AMDP set up with

$$Q(x, c, v) = u(c) + \beta \int v(g(x, c, z)) \varphi(\mathrm{d}z)$$

The monotonicity condition

$$v \leqslant v' \implies Q(x,a,v) \leqslant Q(x,a,v') \quad \text{whenever } (x,a) \in \mathbb{G}$$

holds here

Indeed, with $v \leqslant v'$,

$$\begin{aligned} Q(x,c,v) &= u(c) + \beta \int v(g(x,c,z))\varphi(\mathrm{d}z) \\ &= u(c) + \beta \int v'(g(x,c,z))\varphi(\mathrm{d}z) \\ &= Q(x,c,v') \end{aligned}$$

for all $(x,c) \in \mathbb{G}$

Example. Consider the optimal growth model with IID shocks

- state is $y \in \mathbb{R}_+$
- the action is $c \in \Gamma(y) := [0, y]$
- $\mathbb{G} = \{(y, c) \in \mathbb{R}_+ \times \mathbb{R}_+ : 0 \leqslant c \leqslant y\}$
- Bellman equation is

$$v(y) = \max_{0 \leqslant c \leqslant y} \left\{ u(c) + \beta \int v(f(y - c)z)\varphi(dz) \right\} \qquad (x \in \mathsf{X})$$

Maps to the AMDP set up with

$$Q(y, c, v) = u(c) + \beta \int v(f(y - c)z)\varphi(dz)$$

The monotonicity condition

$$v \leqslant v' \implies Q(x, a, v) \leqslant Q(x, a, v')$$

again holds

With $v \leqslant v'$,

$$\begin{aligned} Q(y, c, v) &= u(c) + \beta \int v(f(y - c)z)\varphi(\mathrm{d}z) \\ &= u(c) + \beta \int v'(f(y - c)z)\varphi(\mathrm{d}z) \\ &= Q(y, c, v') \end{aligned}$$

for all $(y, c) \in \mathbb{G}$

Example. Consider an optimal savings problem where

- w_t is current assets
- $\{z_t\}$ is a finite exogenous state process with kernel Π
- labor income is $y_t = y(z_t)$
- The feasible set for consumption is $[0, w]$

Bellman equation is

$$v(w, z) =$$

$$\max_{0 \leqslant c \leqslant w} \left\{ u(c) + \beta \sum_{z' \in \mathsf{Z}} v\left((1+r)(w-c) + y(z'), z'\right) \Pi(z, z') \right\}$$

Map to AMDP:

- State is $x = (w, z)$
- Feasible correspondence is $\Gamma(w, z) = [0, w]$

The aggregator is

$$Q((w,z), c, v) =$$

$$u(c) + \beta \sum_{z' \in \mathsf{Z}} v\left((1+r)(w-c) + y(z'), z'\right) \Pi(z, z')$$

Monotonicity obviously holds

Example. Consider again the job search problem with

- IID wage offers $\{w_t\}$
- unemployment compensation c and discount factor β

Bellman equation is

$$v(w) = \max \left\{ \frac{w}{1-\beta}, c + \beta \int v(w') \varphi(\mathrm{d}w') \right\}$$

Optimal policy is

$$\sigma^*(w) = \mathbb{1} \left\{ \frac{w}{1-\beta} \geqslant c + \beta \int v^*(w') \varphi(\mathrm{d}w') \right\}$$

with $w \in \mathbb{R}_+$

Map to AMDP:

- state is $w \in \mathbb{R}_+$
- action is $a \in \{0, 1\}$ (reject / accept)
- $\Gamma(w) = \{0, 1\}$ for every w
- the aggregator Q is

$$Q(w, a, v) = a \frac{w}{1 - \beta} + (1 - a) \left[c + \beta \int v(w') q(w') \, \mathrm{d} w' \right]$$

Monotonicity holds because

$$v \leqslant v' \implies Q(w, a, v) \leqslant Q(w, a, v') \quad \text{for all } (w, a) \in \mathbb{G}$$

Example. Job search with correlated wage offers

$$w_t = \exp(z_t) + \exp(\mu + \sigma \zeta_t)$$

The value function satisfies the Bellman equation

$$v(w, z) = \max \left\{ \frac{w}{1-\beta}, c + \beta \, \mathbb{E}_z v(w', z') \right\}$$

Optimal policy is

$$\sigma^*(w, z) = \mathbb{1} \left\{ \frac{w}{1-\beta} \geqslant c + \beta \, \mathbb{E}_z v^*(w', z') \right\}$$

Map to AMDP:

- state is $(w, z) \in \mathsf{X} := \mathbb{R}_+ \times \mathbb{R}$
- action is $a \in \mathsf{A} := \{0, 1\}$ (reject / accept)
- $\Gamma(w) = \{0, 1\}$ for every w
- the aggregator Q is

$$Q((w,z)a, v) = a \frac{w}{1-\beta} + (1-a) \left[c + \beta \mathbb{E}_z \, v(w', z') \right]$$

Monotonicity holds because

$$v \leqslant v' \implies Q((w,z), a, v) \leqslant Q((w,z), a, v')$$

for all $((w,z), a) \in \mathbb{G}$

Example. Job Search with learning

The Bellman equation is

$$v(w, \pi) = \max \left\{ \frac{w}{1-\beta}, \, c + \beta \int v(w', \kappa(w', \pi)) \, q_\pi(w') \, \mathrm{d}w' \right\}$$

where

$$q_\pi := \pi f + (1-\pi) g$$

and

$$\kappa(w, \pi) := \frac{\pi f(w)}{\pi f(w) + (1-\pi) g(w)}$$

- f and g are densities

Map to AMDP:

- state is $x := (w, \pi) \in \mathbb{R}_+ \times (0, 1)$
- action is $a \in \{0, 1\}$ (reject / accept)
- $\Gamma(w, \pi) = \{0, 1\}$ for every w
- the aggregator Q is

$$Q((w, \pi), a, v) = a \frac{w}{1-\beta} +$$

$$(1-a) \left[c + \beta \int v(w', \kappa(w', \pi)) \, q_\pi(w') \, \mathrm{d}w' \right]$$

Ex. Confirm that monotonicity holds

Example. Firm with adjustment costs, inverse demand function

$$p_t := p(q_t, z_t) = a_0 - a_1 q_t + z_t$$

where

$$z_{t+1} = \rho z_t + \sigma \eta_{t+1}, \qquad \{\eta_t\} \overset{\text{IID}}{\sim} N(0,1)$$

Current profits are given by

$$\pi_t := (p_t - c) q_t - \gamma (q_{t+1} - q_t)^2$$

Bellman equation is

$$v(q,z) = \max_{q'} \left\{ (p(q,z) - c) q - \gamma (q' - q)^2 + \beta \, \mathbb{E}_z v(q', z') \right\}$$

Map to AMDP:

- state is $x := (q, z) \in \mathbb{R}^2$
- action is $q \in \mathbb{R}$
- $\Gamma(q, z) = \mathbb{R}$ for all q, z (unrestricted)
- the aggregator Q is

$$Q((q,z), q', v) = (p(q,z) - c)q - \gamma(q' - q)^2 + \beta\, \mathbb{E}_z v(q', z')$$

Ex. Confirm that monotonicity holds

Example. We studied a finite state Markov decision process with

1. finite state space X and action space A
2. feasible correspondence Γ from X $\to$ A
3. reward function $r \colon \mathbb{G} \to \mathbb{R}$
4. discount factor $\beta \in (0, 1)$ and
5. stochastic kernel Π from $\mathbb{G}$ to X

Bellman equation is

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, a, y) \right\}$$

Maps directly to an AMDP with

$$Q(x, a, v) = r(x, a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, a, y)$$

The monotonicity condition

$$v \leqslant v' \implies Q(x, a, v) \leqslant Q(x, a, v')$$

holds, since $v \leqslant v'$ implies

$$Q(x, a, v) = r(x, a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, a, y)$$

$$= r(x, a) + \beta \sum_{y \in \mathsf{X}} v'(y) \Pi(x, a, y) = Q(x, a, v')$$

for all $(x, a) \in \mathbb{G}$

The Bellman Equation

A function $v \in \mathcal{V}$ is said to satisfy the **Bellman equation** if

$$v(x) = \max_{a \in \Gamma(x)} Q(x, a, v) \qquad \text{for all } x \in \mathsf{X}$$

Example. Suppose that X and A are finite,

$$Q(x, a, v) = r(x, a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, a, y)$$

The Bellman equation is

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, a, y) \right\}$$

Recall the basic IID job search problem, where $\Gamma(w) = \{0, 1\}$ and

$$Q(w, a, v) = a \frac{w}{1-\beta} + (1-a) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right]$$

The Bellman equation is

$$\begin{aligned} v(w) &= \max_{a \in \Gamma(w)} Q(w, a, v) \\ &= \max_{a \in \{0,1\}} \left\{ a \frac{w}{1-\beta} + (1-a) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right] \right\} \\ &= \max \left\{ \frac{w}{1-\beta}, \; c + \beta \int v(w') q(w') \, \mathrm{d}w' \right\} \end{aligned}$$

Policies

Recall that $\mathcal{V} \subset \mathbb{R}^{\mathsf{X}}$ is the set of candidate value functions

Let $\Sigma :=$ a family of maps from X to A such that, for each $\sigma \in \Sigma$,

1. $\sigma(x)$ is in $\Gamma(x)$ for all $x \in \mathsf{X}$
2. $\hat{v}(x) := Q(x, \sigma(x), v)$ is in $\mathcal{V}$ for all $v \in \mathcal{V}$

Parts 1 and 2 are called **feasibility** and **consistency** respectively

- Σ is called the **feasible policies**

Example. Consider again the job search problem with

- IID wage offers $\{w_t\}$ taking values in $[0, M]$
- action is $a \in \{0, 1\}$ (reject / accept)
- $\Gamma(w) = \{0, 1\}$ for every w

Set $\mathcal{V}$ = all bounded Borel measurable functions on $[0, M]$

Set Σ = all Borel measurable $\sigma \colon [0, M] \to \{0, 1\}$

Each $\sigma \in \Sigma$ is clearly feasible and also consistent, since

$$Q(w, \sigma(w), v) = \sigma(w) \frac{w}{1 - \beta} + (1 - \sigma(w)) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right]$$

is bounded and Borel measurable in w

Example. Consider the finite state MDP

- X and A finite, feasible correspondence Γ given

Take

- $\mathcal{V}$ to be all of $\mathbb{R}^{\mathsf{X}}$
- Σ be all σ in A^{X} satisfying $\sigma(x)$ is in $\Gamma(x)$ for all $x \in \mathsf{X}$

Obviously each σ in Σ is feasible

Consistency also holds because

$$w(x) := Q(x, \sigma(x), v) = r(x, \sigma(x)) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x, \sigma(x), y)$$

is in $\mathcal{V} = \mathbb{R}^{\mathsf{X}}$ whenever $v \in \mathcal{V}$

Lifetime Value of Policy

Given $\sigma \in \Sigma$ a function $v \in \mathcal{V}$ is called a **σ-value function** if

$$v(x) = Q(x, \sigma(x), v) \quad \text{for all } x \in \mathsf{X}$$

Interpretation: $v = v_\sigma :=$ lifetime value of following σ

- not obvious, but examples given below

Assumption (UNQ). For each $\sigma \in \Sigma$, there is exactly one σ-value function v_σ in $\mathcal{V}$

- essential for our objective function to be well defined

Example. Consider the finite state MDP case we have and suppose that v is a function satisfying, for all $x \in \mathsf{X}$

$$v(x) = Q(x, \sigma(x), v)$$

That is,

$$v(x) = r(x, \sigma(x)) + \beta \sum_{y \in \mathsf{X}} v'(y) \Pi(x, \sigma(x), y)$$

An equivalent statement is $v = r_\sigma + \beta \Pi_\sigma v$

Since $r(\beta \Pi_\sigma) = \beta < 1$, we must have

$$v = v_\sigma := \sum_{t \geqslant 0} \beta^t \Pi_\sigma^t r_\sigma$$

Note that the uniqueness of v_σ in assumption (UNQ) is valid

To see this, pick any $\sigma \in \Sigma$

The statement that

$$v(x) = Q(x, \sigma(x), v) \text{ for all } x \in \mathsf{X}$$

is equivalent to

$$v = r_\sigma + \beta \Pi_\sigma v$$

Since $r(\beta \Pi_\sigma) < 1$, this equation has only one solution

As per the previous slide, this is the lifetime value

$$v_\sigma = \sum_{t \geqslant 0} \beta^t \Pi_\sigma^t r_\sigma$$

Example. In the IID growth model and consumption policy $\sigma \in \Sigma$, suppose v satisfies

$$v(y) = Q(y, \sigma(y), v)$$

Expanding out the last expression yields

$$v(y) = u(\sigma(y)) + \beta \int v(f(y - \sigma(y))z)\varphi(\mathrm{d}z)$$

We claim this implies that

$$v(y) = v_\sigma(y) := \mathbb{E} \sum_{t \geqslant 0} \beta^t u(\sigma(y_t))$$

which is the lifetime value of following σ

To see this (using some Banach space theory), observe that

$$v(y) = u(\sigma(y)) + \beta \int v(f(y - \sigma(y))z)\varphi(\mathrm{d}z)$$

is equivalent to

$$v = u \circ \sigma + \beta \, \Pi_\sigma v$$

Here Π_σ is the operator defined at h in $bc\mathbb{R}_+$ by

$$(\Pi_\sigma h)(y) = \int h[f(y - \sigma(y))z]\varphi(\mathrm{d}z)$$

By the Neumann series theorem, the unique solution to (33) is

$$v(y) = \sum_{t \geqslant 0} \beta^t \Pi_\sigma^t (u \circ \sigma) = \mathbb{E} \sum_{t \geqslant 0} \beta^t u(\sigma(y_t))$$

Optimality

A policy σ^* is called **optimal** if $\sigma^* \in \Sigma$ and

$$v_{\sigma^*}(x) \geqslant v_\sigma(x) \quad \text{for all } \sigma \in \Sigma \text{ and all } x \in \mathsf{X}$$

The **value function** associated with our AMDP is defined by

$$v^*(x) = \sup_{\sigma \in \Sigma} v_\sigma(x) \qquad (x \in \mathsf{X})$$

Evidently, a feasible policy σ^* is optimal if and only if

$$v_{\sigma^*}(x) = v^*(x) \quad \text{for all } x \in \mathsf{X}$$

Given v in $\mathcal{V}$, a policy $\sigma \in \Sigma$ is called **v-greedy** if

$$Q(x, \sigma(x), v) = \max_{a \in \Gamma(x)} Q(x, a, v) \qquad \text{for all } x \in \mathsf{X}$$

- treats v as the value function

Equivalent

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} Q(x, a, v) \qquad \text{for all } x \in \mathsf{X}$$

In the IID job search problem, a policy σ is v-greedy if

$$\sigma(w) \in \operatorname*{argmax}_{a \in \{0,1\}} Q(w, a, v)$$

$$= \operatorname*{argmax}_{a \in \{0,1\}} \left\{ a \frac{w}{1-\beta} + (1-a) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right] \right\}$$

This is equivalent to

$$\sigma(w) = \mathbb{1} \left\{ \frac{w}{1-\beta} \geqslant c + \beta \int v(w') \varphi(\mathrm{d}w') \right\}$$

- optimally accept or reject if v is the value function

Example. In the optimal savings model, we can take

$$\Sigma := \{\text{all Borel measurable } \sigma \in \mathsf{A}^{\mathsf{X}} \text{ s.t. } \sigma(x) \in \Gamma(x), \; \forall \, x \in \mathsf{X}\}$$

- Borel measurable so that integrals are well defined

A policy σ is v-greedy if $\sigma \in \Sigma$ and

$$\sigma(x) \in \operatorname*{argmax}_{c \in \Gamma(x)} \left\{ u(c) + \beta \int v(g(x, c, z)) \varphi(\mathrm{d}z) \right\}$$

Fact. If $v \in bc\mathsf{X}$, then at least one v-greedy policy exists

Proof requires a measurable selection theorem — details omitted

Key Optimality Theorem

Let assumption (UNQ) hold

Theorem. If

1. v^* lies in $\mathcal{V}$ and satisfies the Bellman equation
2. at least one v^*-greedy policy exists

then

a. the set of optimal policies is nonempty and
b. σ is optimal if and only if σ is v^*-greedy

In other words, Bellman's principle of optimality holds

Proof:

Suppose $v^* \in \mathcal{V}$ satisfies the Bellman equation

By the definition of greedy policies,

$$\begin{aligned}
\sigma \text{ is } v^*\text{-greedy} &\iff Q(x, \sigma(x), v^*) = \max_{a \in \Gamma(x)} Q(x, a, v^*), \;\; \forall\, x \\
&\iff Q(x, \sigma(x), v^*) = v^*(x), \;\; \forall\, x \\
&\iff v^* = v_\sigma \\
&\iff \sigma \text{ is optimal}
\end{aligned}$$

In other words, Bellman's principle of optimality holds

Existence of an optimal policy follows from $\exists\; v^*$-greedy

Summary

So now we know: <u>If</u>

1. v^* satisfies the Bellman equation
2. we can calculate v^*
3. v^* admits a greedy policy

then finding an optimal policy is trivial: apply Bellman's principle of optimality, compute a v^* greedy policy

Key remaining questions:

- When does v^* satisfy the Bellman equation?
- How can we compute it?

To answer these questions we introduce two operators

Operators

For each $\sigma \in \Sigma$, we define the **σ-value operator**

$$T_\sigma v(x) = Q(x, \sigma(x), v) \qquad (x \in \mathsf{X}) \tag{1}$$

- Maps $\mathcal{V}$ to itself (by the definition of Σ)
- constructed s.t. fixed points of T_σ coincide with σ-value functions

By assumption, T_σ has exactly one fixed point in $\mathcal{V}$

Lemma The operator T_σ is isotone on $\mathcal{V}$ when paired with the pointwise partial order $\leqslant$

- why?

Our second operator is the **Bellman operator**, defined on $\mathcal{V}$ by

$$Tv(x) = \sup_{a \in \Gamma(x)} Q(x, a, v) \tag{2}$$

Constructed such that

1. any solution to the Bellman equation is a fixed point of T and
2. a fixed point v of T in $\mathcal{V}$ is a solution to the Bellman equation <u>if</u> the sup in (2) can be replaced with max

Greedy policies can now be characterized as follows:

$$\sigma \text{ is } v\text{-greedy} \iff Tv = T_\sigma v \tag{3}$$

Theorem. If

1. T has at least one fixed point $\bar{v}$ in $\mathcal{V}$,
2. there exists at least one $\bar{v}$-greedy policy in Σ, and
3. for all $\sigma \in \Sigma$ and all $x \in \mathsf{X}$,

$$\lim_{k \to \infty} T_\sigma^k \, \bar{v}(x) \geqslant v_\sigma(x) \tag{4}$$

then

1. $\bar{v} = v^*$ and
2. v^* is the unique solution to the Bellman equation in $\mathcal{V}$

$\implies$ existence of an optimal policy and Bellman's principle of optimality

Key Sufficient Conditions

Let $\mathcal{V}$ be endowed with a metric ρ such that

$$\lim_{n \to \infty} \rho(v_n, v) = 0 \implies \lim_{n \to \infty} v_n(x) = v(x) \text{ for all } x \in \mathsf{X}$$

- example?

Stable AMDP assumptions:

S1. Given any $\sigma \in \Sigma$, the system $(\mathcal{V}, T_\sigma)$ is globally stable

S2. There exists a subset $\hat{\mathcal{V}}$ of $\mathcal{V}$ such that

a. Each $v \in \hat{\mathcal{V}}$ has at least one v-greedy policy in Σ and

b. $(\hat{\mathcal{V}}, T)$ is globally stable

Key Theorem for Applications

Theorem. If the stable AMDP conditions S1–S2 hold, then

1. Assumption UNQ is satisfied
2. v^* lies in $\hat{\mathcal{V}}$ and is the unique solution to the Bellman equation in $\mathcal{V}$
3. $T^n v \to v^*$ whenever $v \in \hat{\mathcal{V}}$
4. Bellman's principle of optimality is valid and at least one optimal policy exists

This is all we need for applications

Proof is in course notes

Example. Recall the finite state MDP

- X, A finite, feasible correspondence Γ given
- $\mathcal{V}$ = all of $\mathbb{R}^{\mathsf{X}}$
- Σ = all σ in A^{X} satisfying $\sigma(x)$ is in $\Gamma(x)$ for all $x \in \mathsf{X}$

and

$$T_\sigma v(x) = r(x, \sigma(x)) + \beta \sum_{y \in \mathsf{X}} v'(y) \Pi(x, \sigma(x), y)$$

Claim: Condition S1 holds

Proof: T_σ is a contraction of modulus β on $(\mathcal{V}, d_\infty)$

(See lecture 10)

How about S2, which requires a subset $\hat{\mathcal{V}}$ of $\mathcal{V}$ such that

a. Each $v \in \hat{\mathcal{V}}$ has at least one v-greedy policy in Σ and
b. $(\hat{\mathcal{V}}, T)$ is globally stable

This works with $\hat{\mathcal{V}} := \mathcal{V} =$ all of $\mathbb{R}^{\mathsf{X}}$

Existence of greedy policies is trivial in a finite setting

Moreover

$$Tv(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \sum_{y \in \mathsf{X}} v(y) \Pi(x,a,y) \right\}$$

is a contraction of modulus β on $(\mathbb{R}^{\mathsf{X}}, d_\infty)$

Example. Consider again the job search problem with

- IID wage offers $\{w_t\}$ taking values in $[0, M]$
- action is $a \in \{0, 1\}$ (reject / accept)
- $\Gamma(w) = \{0, 1\}$ for every w

Set

- $\mathcal{V}$ = all bounded Borel measurable functions on $[0, M]$
- $\hat{\mathcal{V}} = \mathcal{V}$
- Σ = all Borel measurable $\sigma \colon [0, M] \to \{0, 1\}$

S1 requires that, given any $\sigma \in \Sigma$, the system $(\mathcal{V}, T_\sigma)$ is globally stable

To see this is true, observe that, given σ, we have

$$T_\sigma v(x) = \sigma(w) \frac{w}{1-\beta} + (1 - \sigma(w)) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right]$$

Ex. Fix v_1 and v_2 in $\mathcal{V}$ and $w \in [0, M]$

1. Show that

$$|T_\sigma v_1(w) - T_\sigma v_2(w)| \leqslant \beta \|v_1 - v_2\|_\infty$$

2. Conclude that T_σ is a contraction of modulus β on $\mathcal{V}$

S2 requires a subset $\hat{\mathcal{V}}$ of $\mathcal{V}$ such that

a. Each $v \in \hat{\mathcal{V}}$ has at least one v-greedy policy in Σ and
b. $(\hat{\mathcal{V}}, T)$ is globally stable

This works with $\hat{\mathcal{V}} = \mathscr{C} :=$ all continuous functions on $[0, M]$

The Bellman operator T is

$$Tv(w) = \max_{a \in \{0,1\}} \left\{ a \frac{w}{1-\beta} + (1-a) \left[c + \beta \int v(w') q(w') \, \mathrm{d}w' \right] \right\}$$

$$= \max \left\{ \frac{w}{1-\beta}, \; c + \beta \int v(w') q(w') \, \mathrm{d}w' \right\}$$

We have already shown that greedy policies always exist, T is a contraction map on $(\mathscr{C}, d_\infty)$

Example. Recall the generic optimal savings model

- the action is $c \in \Gamma(x)$
- utility function u is bounded
- $\mathbb{G} = \{(x, c) \in \mathsf{X} \times \mathbb{R}_+ : c \in \Gamma(x)\}$
- state-action aggregator is

$$Q(x, c, v) = u(c) + \beta \int v(g(x, c, z)) \varphi(\mathrm{d}z)$$

- $\Sigma =$ all Borel measurable $\sigma \in \mathsf{A}^{\mathsf{X}}$ s.t. $\sigma(x) \in \Gamma(x)$, $\forall\, x \in \mathsf{X}$

S1 requires that, given any $\sigma \in \Sigma$, the system $(\mathcal{V}, T_\sigma)$ is globally stable

S2 requires existence of a subset $\hat{\mathcal{V}}$ of $\mathcal{V}$ such that

a. Each $v \in \hat{\mathcal{V}}$ has at least one v-greedy policy in Σ and
b. $(\hat{\mathcal{V}}, T)$ is globally stable

We have already checked these conditions when

- $\mathcal{V} = bm\mathsf{X} :=$ all Borel measurable functions in $b\mathsf{X}$
- $\hat{\mathcal{V}} = bc\mathsf{X}$

In particular,

- T_σ is a contraction of modulus β on $bm\mathsf{X}$ for all σ
- T is a contraction of modulus β on $bc\mathsf{X}$