Random Value Function Iteration for Dynamic Programming

Jenö Pál John Stachurski

Central European University

Australian National University

July 2011

What's it about?

Topic: Numerical Dynamic Programming

What we do:

- Consider a continuous state stochastic DP problem
- Propose solution method: variation of value function iteration
- Analyze algorithm, show convergence

Objectives:

- Algorithm works for a very general class of stochastic DPs
- Guaranteed convergence

Related Literature

Closest papers:

- Szepesvári and Munos (2008)
- Rust's curse-of-dimensionality paper (1997)

Our setting is somewhat less restrictive

Generic Dynamic Programming Problem

while 1:

- ◦ Agent observes state $x \in \mathbb{X}$ of a given system
- ◦ Responds with action $a \in \mathbb{A}$ from feasible set $\Gamma(x) \subset \mathbb{A}$
- ◦ Receives current reward $r(x, a) \in \mathbb{R}$
- ◦ Current shock U drawn from distribution ϕ
- ◦ New state determined as $x' = F(x, a, U)$

Examples:

- Most stationary DP problems

Assumptions:

- State and action spaces are compact metric spaces
- Functions and correspondences are continuous

Value Function Iteration

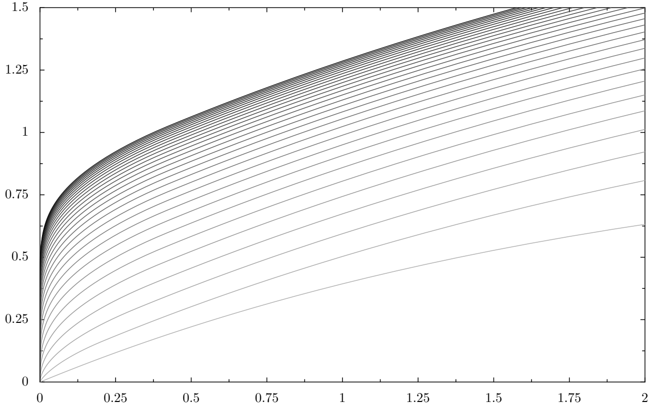


Figure: VFI for the optimal consumption model

Let

- $\mathscr{C}(\mathbb{X})$ be all continuous $w \colon \mathbb{X} \to \mathbb{R}$
- $T \colon \mathscr{C}(\mathbb{X}) \to \mathscr{C}(\mathbb{X})$ be the Bellman operator

$$Tw(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \int w[F(x,a,u)] \phi(du) \right\}$$

- $V_T \in \mathscr{C}(\mathbb{X})$ be the value function

Standard results:

- V_T is the unique fixed point of T
- $\|T^k w - V_T\| = O(\rho^k)$ for all $w \in \mathscr{C}(\mathbb{X})$

Numerical Issue Number 1: Numerical Integration

Consider iterating with Bellman operator

$$Tw(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \int w[F(x,a,u)]\phi(du) \right\}$$

Must approximate integral for many different a, x and w

To compute integrals we use Monte Carlo:

1. Draw $\{U_i\}_{i=1}^n \stackrel{\text{IID}}{\sim} \phi$ using r.n.g.
2. Replace the Bellman operator

$$Tw(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \int w[F(x,a,u)] \phi(du) \right\}$$

with R_n, where

$$R_n w(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \frac{1}{n} \sum_{i=1}^{n} w[F(x,a,U_i)] \right\}$$

Numerical Issue Number 2: Function Approximation

Recall definition of R_n

$$R_n w(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \frac{1}{n} \sum_{i=1}^{n} w[F(x,a,U_i)] \right\}$$

When $\mathbb{X}$ infinite,

- cannot evaluate $R_n w(x)$ at all x in finite time
- cannot store $R_n w(x)$ at all x with finite memory

Function approximation step:

Introduce approximation operator $A \colon \mathscr{C}(\mathbb{X}) \to \mathscr{C}(\mathbb{X})$ s.t.

- $Aw =$ finite parametric representation of w

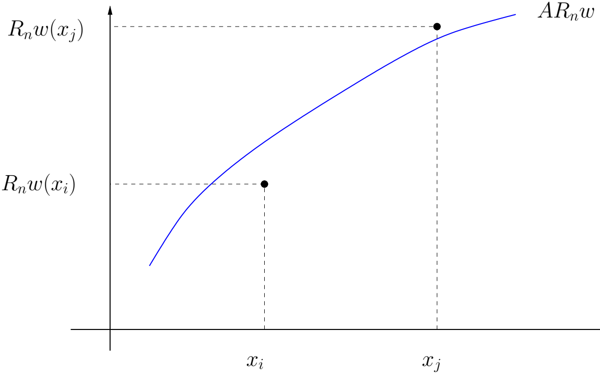
Restriction: A is nonexpansive on $\mathscr{C}(\mathbb{X})$:

$$\|Aw - Aw'\| \leq \|w - w'\| \quad \text{for all } \; w, w' \in \mathscr{C}(\mathbb{X})$$

(Kernel smoothers, continuous piecewise linear interpolation, etc.)

Lemma: If A is nonexpansive and M is a contraction of mod ρ, then $AM = A \circ M$ is a contraction of mod ρ

The Algorithm: Iterate with AR_n instead of T


$R_n w(x_j)$
$AR_n w$
$R_n w(x_i)$
x_i
x_j

Error Analysis

The following operators are contractions of mod ρ:

- T (obvious)
- AT (because A nonexpansive)
- AR_n (for each $n \in \mathbb{N}$, almost surely)

Corresponding fixed points:

- V_T — fixed point of T (value function)
- V_{AT} — fixed point of AT
- V_{AR_n} — fixed point of AR_n (random)

What we want to compute: V_T (value function)

What we can compute: V_{AR_n} (up to any tolerance)

Error decomposition:

$$\|V_T - V_{AR_n}\| \leq \|V_T - V_{AT}\| + \|V_{AT} - V_{AR_n}\| =: \mathbf{FA} + \mathbf{NI}$$

- $\mathbf{FA}$:= Function approximation error
- $\mathbf{NI}$:= Numerical integration (Monte Carlo) error

Results

1. Regarding the function approximation error,

$$\forall\, \varepsilon > 0,\ \exists\, A \text{ such that } \mathbf{FA} := \|V_T - V_{AT}\| < \varepsilon$$

2. Without any additional conditions,

$$\mathbf{NI} = \|V_{AT} - V_{AR_n}\| \to 0 \qquad (\mathbf{P}^*\text{-a.s.})$$

3. With some additional conditions,

$$\mathbf{NI} = \|V_{AT} - V_{AR_n}\| \leq \mathcal{E}_n = O_P(n^{-1/2})$$

Sketch of proof that $\mathbf{NI} = \|V_{AT} - V_{AR_n}\| \to 0$ $\mathbf{P}^*$-a.s.

We show that

$$\|V_{AT} - V_{AR_n}\| \leq \text{ const. } \times \max_{(x,a)\in\mathbb{G}} \left| \frac{1}{n} \sum_{i=1}^{n} V_{AT}[F(x,a,U_i)] - \int V_{AT}[F(x,a,u)]\phi(du) \right|$$

We can write this as

$$\|V_{AT} - V_{AR_n}\| \leq c \cdot \sup_{h\in\mathcal{H}} \left| \frac{1}{n} \sum_{i=1}^{n} h(U_i) - \int h(u)\phi(du) \right|$$

When does r.h.s. $\to 0$?

A sufficient condition: $\mathcal{H}$ consists of functions $h_\alpha \colon \mathbb{U} \to \mathbb{R}$ with index α in metric space Λ where

1. Λ is compact
2. $\Lambda \ni \alpha \mapsto h_\alpha(u) \in \mathbb{R}$ is continuous for every $u \in \mathbb{U}$
3. $\exists\, H \colon \mathbb{U} \to \mathbb{R}$ s.t. $\int H d\phi < \infty$ and $|h_\alpha| \leq H$ for every $\alpha \in \Lambda$

These conditions easily verified for our set of functions

$$h_{a,x}(\cdot) = V_{AT}[F(x, a, \cdot)] \quad \text{with} \quad (a, x) \in \mathbb{G}$$

Final Comments

Why Not Just Discretize?

Because:

- Analysis of errors more problematic
- Curse of dimensionality

Example: In engineering and other sciences, discrete problems often replaced with continuous ones to make them tractable

Why Bother with Error Analysis?

Errors propagated at each iteration of approx. Bellman operator

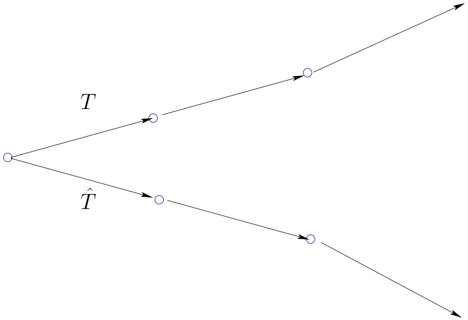


Figure: Iteration with nearby maps