Dynamic Programming

$a \in A$ action space: finite

$y \in Y$ state space: finite

$\pi(y'|y,a)$ transition probability

period utility $u(a,y)$ with discount factor $0 \le \delta < 1$

$\overline{u} = 2\max|u(a,y)|$

Strategies

finite histories $h = (y_1, y_2, \ldots, y_t)$ with $t(h) = t$, $y(h) = y_t$; $h-1$; $y_1(h)$; $h' \geq h$

H space of all finite histories; this is countable

strategies $\sigma : H \to A$

Σ space of all strategies

all maps from a countable set to a finite set

the product topology $\sigma^n \to \sigma$ means that $\sigma^n(h) \to \sigma(h)$ for every h

Theorem: every sequence in the product topology has a convergent subsequence, so the space of strategies is compact

(proven in any elementary topology textbook)

Strong Markov Strategies

define a strong Markov strategy $\sigma(h)=\sigma(h')$ if $y(h)=y(h')$

a strong Markov strategy is equivalent to a map

$$\sigma : Y \to A$$

recursively define

$$\pi(h|y_1,\sigma)\equiv$$

$$\begin{cases} \pi(y(h)|y(h-1),\sigma(h-1))\pi(h-1|y_1,\sigma) & t(h)>1 \\ 1 & t(h)=1 \text{ and } y_1(h)=y_1 \\ 0 & t(h)=1 \text{ and } y_1(h)\neq y_1 \end{cases}$$

calculate the average present value of the objective function

$$V(y_1,\sigma)\equiv(1-\delta)\sum_{h\in H}\delta^{t(h)-1}u(\sigma(h),y(h))\pi(h|y_1,\sigma)$$

Dynamic Programming Problem

(*) maximize $V(y_1, \sigma)$ subject to $\sigma \in \Sigma$

a value function is a map $v\colon Y \to \Re$ bounded by $\bar{u}$

note that in this setting, it is simply a finite vector v_y

Existence

Lemma: a solution to (*) exists

Definition: ***the*** value function

$$v(y_1) \equiv \max_{\sigma \in \Sigma} V(y_1, \sigma)$$

Proof: the maximum exists because in the product topology on Σ $V(y_1, \sigma)$ is continuous in σ and Σ is compact

why is V continuous?

suppose $\sigma^n \to \sigma$

$$V(y_1, \sigma^n) = (1-\delta)\sum_{h \in H} \delta^{t(h)-1} u(\sigma^n(h), y(h))\pi(h|y_1, \sigma^n)$$

$$= (1-\delta)\sum_{t(h)<T} \delta^{t(h)-1} u(\sigma^n(h), y(h))\pi(h|y_1, \sigma^n)$$

$$+(1-\delta)\sum_{t(h)\geq T} \delta^{t(h)-1} u(\sigma^n(h), y(h))\pi(h|y_1, \sigma^n)$$

$$\to (1-\delta)\sum_{t(h)<T} \delta^{t(h)-1} u(\sigma(h), y(h))\pi(h|y_1, \sigma) + O(\delta^T \overline{u})$$

so as $T \to \infty$ we have $V(y_1, \sigma^n) \to V(y_1, \sigma)$

Bellman equation

we define a map $T : \Re^Y \to \Re^Y$ by $w' = T(w)$ if

$$w'(y_1) = \max_{a \in A}(1-\delta)u(a, y_1) + \delta \sum_{y'_1 \in S} \pi(y'_1 \mid y_1, a) w(y'_1)$$

Lemma: the value function is a fixed point of the Bellman equation $T(v) = v$

in other words the most you can get next period is also given by the value function

$$v(y_1) = \max_{a \in A}(1-\delta)u(a, y_1) + \delta \sum_{y'_1 \in S} \pi(y'_1 \mid y_1, a) v(y'_1)$$

Lemma: the Bellman equation is a contraction mapping
$\|T(w)-T(w')\| \leq \delta \|w-w'\|$

Proof:

key observation $\|\max_\alpha f(\alpha) - \max_\alpha g(\alpha)\| \leq \max_\alpha \|f(\alpha) - g(\alpha)\|$

$$\left\| \max_{\alpha\in A}(1-\delta)u(\alpha,y_1) + \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w(y'_1) - \max_{\alpha\in A}(1-\delta)u(\alpha,y_1) + \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w'(y'_1) \right\|$$

$$\leq \max_{\alpha\in A} \left\| (1-\delta)u(\alpha,y_1) + \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w(y'_1) - (1-\delta)u(\alpha,y_1) + \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w'(y'_1) \right\|$$

$$= \max_{\alpha\in A} \left\| \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w(y'_1) - \delta \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha)w'(y'_1) \right\|$$

$$\leq \delta \max_{\alpha\in A} \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha) \mid w(y'_1) - w'(y'_1) \mid$$

$$\leq \delta \max_{\alpha\in A} \sum_{y'_1\in S} \pi(y'_1|y_1,\alpha) \|w-w'\|$$

$$= \delta \|w-w'\|$$

Summing Up

Corollary: the Bellman equation has a unique solution

Proof: Let w be another solution

$$\|v - w\| = \|T(v) - T(w)\| \leq \delta \|v - w\| \Rightarrow \|v - w\| = 0$$

Conclusion: the unique solution to the Bellman equation is the value function

since the value function is *a* solution to the Bellman equation, and the solution is unique

Existence of Strong Markov

Lemma: there is a strong Markov optimum and it may be found from the Bellman equation

Proof:

define a strong Markov plan by

$$\sigma(y_1) \in \arg\max_{\alpha \in A} (1-\delta) u(\alpha, y_1) + \delta \sum_{y'_1 \in S(y_1)} \pi(y'_1 | y_1, \alpha) v(y'_1)$$

work the value function forward recursively to find

$$\begin{aligned} v(y_1(h)) = (1-\delta) & \sum\nolimits_{t(h)<T} \delta^{t(h)-1} \pi(y(h) | y_1(h), \sigma) u(\sigma(y(h)), y(h)) \\ & +(1-\delta) \sum\nolimits_{t(h)=T} \delta^{T} \pi(y(h) | y_1(h), \sigma) v(h) \end{aligned}$$

and observe that v is bounded by $\overline{u}$ so that the final terms disappears asymptotically

Application – job search

three states: unemployed (u), have a bad job (b), have a good job (g)

the only choice: whether or not to quit a bad job and become unemployed

pr(g|g) = 1 (good job is absorbing)

pr(g|u) = a > b = pr(g|b, not quit) (chance of getting a good job)

pr(b|u) = c (chance of getting a bad job when unemployed)

u(g) = d

u(b) = 1

u(u) = 0

procedure: find the value function

$$v(g) = d$$

$$v(u) = (1-\delta)0 + \delta(av(g) + cv(b) + (1-a-c)v(u))$$

$$v(b) = \max\begin{cases}(1-\delta) + \delta(bv(g) + (1-b)v(b)) \\ (1-\delta) + \delta v(u)\end{cases}$$

step 0: substitute out v(g)

$$v(u) = (1-\delta)0 + \delta(ad + cv(b) + (1-a-c)v(u))$$

$$v(b) = \max\begin{cases}(1-\delta) + \delta(bd + (1-b)v(b)) \\ (1-\delta) + \delta v(u)\end{cases}$$

case 1: optimum is to quit a bad job

$$v(b) = (1-\delta) + \delta v(u)$$

substitute:

$$v(u) = \delta(ad + c((1-\delta) + \delta v(u)) + (1-a-c)v(u))$$

$$(1 - \delta(1-a-c) - \delta^2 c)v(u) = \delta ad + \delta(1-\delta)c$$

$$v(u) = \frac{\delta ad + \delta(1-\delta)c}{(1 - \delta(1-a-c) - \delta^2 c)}$$

verify the Bellman equation

$$(1-\delta)+\delta v(u) \geq (1-\delta)+\delta(bd+(1-b)v(b))$$
$$= (1-\delta)+\delta(bd+(1-b)((1-\delta)+\delta v(u)))$$

$$v(u) \geq \frac{bd+(1-b)(1-\delta)}{1-\delta(1-b)}$$

$$\frac{\delta ad+\delta(1-\delta)c}{(1-\delta(1-a-c)-\delta^2 c)} \geq \frac{bd+(1-b)(1-\delta)}{1-\delta(1-b)}$$

for example when $b=0, c=0, a=1$, $\delta d \geq 1$