Dynamic Programming: Infinite State

4.2.3 Job Search with Learning

Thomas J. Sargent and John Stachurski

Introduction

This section we consider a job search model where wage offer distribution is unknown.

The agent learns about φ by starting with a prior and updating it in light of the wage offer sequence.

Learning about the Wage Distribution

- two candidate continuous distributions f and g over wage offers;
- nature chooses φ from f or g, and the entire wage offer sequence is drawn from φ;
- agent does not know the true φ, and starts with an initial guess $\pi_0 f(w) + (1 - \pi_0) g(w)$;
- agent observes a sequence of wage offers $w_1, w_2, \ldots$ and updates her beliefs about φ according to Bayes' rule:

$$\pi_{t+1} = \frac{f(W_{t+1})\pi_t}{f(W_{t+1})\pi_t + g(W_{t+1})(1 - \pi_t)} \tag{1}$$

A bit more on Bayes' rule

The update rule (1) is updating the probability that $\varphi = f$ based on the observation W_{t+1}.

More generally, we can write the update rule as

$$\mathbb{P}\{\varphi = f \mid W_{t+1}\} = \frac{\mathbb{P}\{W_{t+1} \mid \varphi = f\}\mathbb{P}\{\varphi = f\}}{\mathbb{P}\{W_{t+1}\}}$$

where

$$\mathbb{P}\{W_{t+1}\} = \sum_{\psi \in \{f,g\}} \mathbb{P}\{W_{t+1} \mid \varphi = \psi\}\mathbb{P}\{\varphi = \psi\}.$$

Belief State

Now we can formulate the problem as an ADP by incorpreting the belief state.

Assumption 1: Support of Wage Distribution

The densities f and g are positive on $(0, M)$ and zero elsewhere.

let $\varphi_\pi := \pi f + (1 - \pi) g$ represent the estimate of the wage offer distribution given belief π and let

$$\kappa(w, \pi) := \frac{\pi f(w)}{\pi f(w) + (1 - \pi) g(w)} \qquad (w \in (0, M)\ \pi \in (0, 1)).$$

$\kappa(w, \pi)$ is the updated belief π' of π having observed draw w, e.g. $\kappa : (0, M) \times (0, 1) \to (0, 1)$.

ADP Formulation

The state is $(w, \pi) \in (0, M) \times (0, 1) := \mathsf{X}$ and π is referred to as the **belief state**. The action space is $\mathsf{A} = \{0, 1\}$:

The policy operator is given by

$$
\begin{aligned}
(T_\sigma \, v)(w, \pi) = \sigma(w, \pi) \frac{w}{1 - \beta} \\
+ (1 - \sigma(w, \pi)) \left[c + \beta \int v(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \right].
\end{aligned}
$$

Let $\mathbb{T} = \{T_\sigma : \sigma \in \mathsf{A}\}$, and $V := bm\mathsf{X}$. Evidently, T_σ is order preserving selfmap on V. Hence $(V, \mathbb{T})$ is an ADP.

ADP Formulation

Similar to the traditional job search model, it is easy to check that given $v \in V$, show that σ defined by

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

is v-greedy for $(V, \mathbb{T})$.

The Bellman operator is given by

$$v(w, \pi) = \max\left\{\frac{w}{1-\beta},\ c + \beta \int v(w', \kappa(w', \pi))\, \varphi_\pi(w')\, \mathrm{d}w'\right\}. \tag{2}$$

Factor ADP

Similar to the conventional model, we can formulate the problem as a factor ADP to reduce the dimensionality.

$h(\pi)$ be the corresponding reservation wage at belief state π, which is the wage level at which the worker is indifferent between accepting and rejecting. This value satisfies

$$\frac{h(\pi)}{1-\beta} = c + \beta \int v(w', \kappa(w', \pi))\, \varphi_\pi(w')\, \mathrm{d}w'. \tag{3}$$

We combine (2) and (3) to obtain

$$v(w, \pi) = \max\left\{\frac{w}{1-\beta}, \frac{h(\pi)}{1-\beta}\right\}$$

and then use this expression to eliminate v in (2), obtaining

$$h(\pi) = (1-\beta)c + \beta \int \max\left\{w', h[\kappa(w', \pi)]\right\}\, \varphi_\pi(w')\, \mathrm{d}w'. \tag{4}$$

Equation (4) can be understood as a functional equation in h.

Equivalently, the map h is the fixed point of the operator $\hat{T}$ given by

$$(\hat{T}h)(\pi) = (1-\beta)c + \beta \int \max\{w', h[\kappa(w', \pi)]\}\ \varphi_\pi(w')\ \mathrm{d}w'. \tag{5}$$

When this fixed point is well-defined we call it the **optimal reservation wage function**. The value $h(\pi)$ will indicate the smallest wage offer at which the worker is willing to accept, given her current belief state π.

FDP Formulation

We adapt the model to the factored dynamic program (FDP) framework $(V, \hat{V}, \mathbb{G}, F)$.

- State: (w, π), where w is the wage offer and π is the belief state.
- $V = bm\mathsf{X}$
- $\hat{V} = bm\Pi$, where $\pi \in \Pi \coloneqq (0, 1)$
  - Related to reservation wage $h_{\mathsf{res}}(\pi)$ in (3) by $h(\pi) = \frac{h_{\mathsf{res}}(\pi)}{1-\beta}$.

FDP Operators

- Operator $F\colon V \to \hat{V}$: Maps full value v to continuation value $h = Fv$.

$$(Fv)(\pi) = c + \beta \int v(w', \kappa(w', \pi))\, \varphi_\pi(w')\, \mathrm{d}w'. \quad \text{(from (3))}$$

- Family $\mathbb{G} = \{G_\sigma \mid \sigma \in \Sigma\}$: Maps $h \in \hat{V}$ to value $v = G_\sigma h$ under policy σ.

$$(G_\sigma h)(w, \pi) = \frac{\sigma(w, \pi) w}{1 - \beta} + (1 - \sigma(w, \pi)) h(\pi)$$

  where $\sigma(w, \pi) \in \{0, 1\}$ (0=reject, 1=accept).

- $G\colon \hat{V} \to V$: Defined as $G = \sup_{\sigma \in \Sigma} G_\sigma$. This yields:

$$(Gh)(w, \pi) = \max\left\{\frac{w}{1 - \beta}, h(\pi)\right\}.$$

  This corresponds to using the policy σ^* that is greedy w.r.t. h.

$T_\sigma := G_\sigma F \colon V \to V$:

$$
\begin{aligned}
(T_\sigma v)(w, \pi) &= (G_\sigma(Fv))(w, \pi) \\
&= \frac{\sigma(w, \pi) w}{1 - \beta} + (1 - \sigma(w, \pi))(Fv)(\pi) \\
&= \frac{\sigma(w, \pi) w}{1 - \beta} \\
&\quad + (1 - \sigma(w, \pi)) \left[c + \beta \int v(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \right]
\end{aligned}
$$

This is the standard policy operator for policy σ in the original state space V.

$\hat{T}_\sigma := FG_\sigma \colon \hat{V} \to \hat{V}$:

$$
\begin{aligned}
(\hat{T}_\sigma h)(\pi) &= (F(G_\sigma h))(\pi) \\
&= c + \beta \int (G_\sigma h)(w', \kappa(w', \pi)) \, \varphi_\pi(w') \, \mathrm{d}w' \\
&= c + \beta \int \left[\frac{\sigma(w', \kappa(w', \pi)) w'}{1 - \beta} \right. \\
&\quad \left. + (1 - \sigma(w', \kappa(w', \pi))) h(\kappa(w', \pi)) \right] \varphi_\pi(w') \, \mathrm{d}w'
\end{aligned}
$$

This operator updates the continuation value function h under policy σ.

- It is easy to verify that F and G_σ are order preserving, and
- according to the discussion on v-greedy policies in the previous section, the set $\{G_\sigma \hat{v}\}$ has the greatest element for each $\hat{v} \in \hat{V}$.
- Hence, $(V, \hat{V}, \mathbb{G}, F)$ is indeed a factored dynamic program.

Recall the operator $\hat{T}$ from (5) acting on the reservation wage function h:

$$(\hat{T}h)(\pi) = (1-\beta)c + \beta \int \max\{w', h[\kappa(w',\pi)]\}\ \varphi_\pi(w')\ \mathrm{d}w'.$$

Goal: Show $\hat{T}$ maps $bc(0,1)$ (bounded, continuous functions on $(0,1)$) to itself. If $h \in bc(0,1)$, is $\hat{T}h$ also continuous?

Applying Dominated Convergence

Let $\pi_n \to \pi$ in $(0,1)$. We need to show $(\hat{T}h)(\pi_n) \to (\hat{T}h)(\pi)$. Requires showing convergence of the integral term:

$$\int \underbrace{\max\{w', h[\kappa(w', \pi_n)]\}\ \varphi_{\pi_n}(w')}_{H_n(w')}\, \mathrm{d}w' \to$$
$$\int \max\{w', h[\kappa(w', \pi)]\}\ \varphi_{\pi}(w')\ \mathrm{d}w'$$

We use the Dominated Convergence Theorem (DCT). Check conditions:

1. Pointwise Convergence: $H_n(w') \to H(w')$ a.e.
2. Domination: $\sup_n |H_n(w')| \leqslant H(w')$ for some integrable H.

Checking DCT Conditions

Assume $h \in bc(0,1)$. Let $\|h\|_\infty = \sup |h(\pi)|$.

1. Pointwise Convergence:
   - $\kappa(w', \pi)$ and $\varphi_\pi(w')$ are continuous in $\pi \in (0,1)$ for $w' \in (0, M)$.
   - Since h is continuous, $h[\kappa(w', \pi_n)] \to h[\kappa(w', \pi)]$.
   - Thus $H_n(w') \to \max\{w', h[\kappa(w', \pi)]\}\, \varphi_\pi(w')$ pointwise for $w' \in (0, M)$.
2. Domination:
   - $\max\{w', h[\kappa(w', \pi_n)]\} \leqslant \max\{M, \|h\|_\infty\}$.
   - $\varphi_{\pi_n}(w') = \pi_n f(w') + (1 - \pi_n) g(w') \leqslant f(w') + g(w')$.
   - $|H_n(w')| \leqslant \underbrace{\max\{M, \|h\|_\infty\}\,(f(w') + g(w'))}_{H(w')}$.
   - $\int H(w')\,\mathrm{d}w' = \max\{M, \|h\|_\infty\}(\int f + \int g) = 2\max\{M, \|h\|_\infty\} < \infty$.

Conclusion: By DCT, the integral converges. Thus $\hat{T}h$ is continuous.