Dynamic Programming Lecture #10

Outline:

- Termination problems
- Monotonicity property

Termination Problems

- Two possibilities:
  - Play until end:

$$\text{realized cost } = g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, \mu_k(x_k), w_k)$$

  - Or terminate at stage k^*:

$$\text{realized cost } = T + \sum_{k=0}^{k^*-1} g_k(x_k, \mu_k(x_k), w_k)$$

- Augment state space: x^T & u^T
- New state dynamics:

$$\begin{aligned} x_{k+1} &= \tilde{f}_k(x_k, u_k, w_k) \\ &= \begin{cases} f_k(x_k, u_k, w_k) & x_k \neq x^T \text{ and } u_k \neq u^T; \\ x^T & x_k = x^T \text{ or } u_k = u^T \end{cases} \end{aligned}$$

- New stage cost:

$$\tilde{g}_k(x_k, u_k, w_k) = \begin{cases} g_k(x_k, u_k, w_k) & x_k \neq x^T \text{ and } u_k \neq u^T; \\ T & x_k \neq x^T \text{ and } u_k = u^T; \\ 0 & x_k = x^T \end{cases}$$

- New terminal cost:

$$\tilde{g}_N(x_N) = \begin{cases} g_N(x_N) & x_N \neq x^T; \\ 0 & x_N = x^T \end{cases}$$

- We will drop "~" notation.

Asset Selling

- Problem: Determine whether to take offer on asset.
- Original formulation:
  - Dynamics:

$$x_{k+1} = w_k$$

    * x_k = current bid.
    * w_k = next bid (uncertain).
  - Payoff functions:

$$g_k(x_k) = \begin{cases} (1+r)^{N-k} x_k & \text{if } u_k = u^T \\ 0 & \text{otherwise} \end{cases}$$

  Invest sale until end of horizon or...

$$g_N(x_N) = x_N$$

  Must accept final offer, w_{N-1}, at end of horizon
- Can convert original formulation to explicitly show termination.

Asset Selling DP Iterations

- Final stage
  - $x_N \neq x^T$: $J_N(x_N) = x_N$
  - $x_N = x^T$: $J_N(x_N) = 0$
- Prior stage: Expected immediate plus future.
  - $x_{N-1} \neq x^T$

    $$J_{N-1}(x_{N-1}) = \max_{\text{sell/don't sell}} \left\{(1+r)x_{N-1} + E\left\{J_N(x^T)\right\}, E\left\{J_N(x_N)\right\}\right\}$$

    Accept offer if

    $$(1+r)x_{N-1} > E\left\{J_N(x_N)\right\} = E\left\{x_N\right\} = E\left\{w_{N-1}\right\}$$

    or

    $$x_{N-1} > E\left\{w_{N-1}\right\}/(1+r)$$

    (a threshold)
  - $x_{N-1} = x^T$

    $$J_{N-1}(x_{N-1}) = 0$$

    i.e., there is no gain if property is already sold.
- Stage $N-2$, $x_{N-2} \neq x^T$:

  $$J_{N-2}(x_{N-2}) = \max_{\text{sell/don't sell}} \left\{(1+r)^2 x_{N-2}, E\left\{J_{N-1}(w_{N-2})\right\}\right\}$$

  but

  $$E\left\{J_{N-1}(w_{N-2})\right\} = E\left\{\max\left\{(1+r)w_{N-2}, E\left\{w_{N-1}\right\}\right\}\right\}$$
- Can continue to construct thresholds $(\alpha_0, \ldots, \alpha_{N-1})$ for optimal policy: Sell if $x_k > \alpha_k$.

Increasing Thresholds

- Assume i.i.d offers (independent, identical probability distributions).
- EXPECT: $\alpha_0 \geq \alpha_1 \geq \ldots \alpha_{N-1}$, i.e., in early stages hold out for high offer...panic in later stages.
- PROOF:
  - FACT: $J_{N-1}(x) \geq (1+r)J_N(x)$.

$$\begin{aligned} J_{N-1}(x) &= \max\{(1+r)x, E\{w\}\} \\ &= \max\{(1+r)J_N(x), E\{w\}\} \\ &\geq (1+r)J_N(x) \end{aligned}$$

  - FACT: $J_{N-2}(x) \geq (1+r)J_{N-1}(x)$.

$$\begin{aligned} J_{N-2}(x) &= \max\left\{(1+r)^2 x, E\{J_{N-1}(w)\}\right\} \\ &\geq \max\left\{(1+r)^2 x, (1+r)E\{J_N(w)\}\right\} \\ &= (1+r)\max\{(1+r)x, E\{J_N(w)\}\} \\ &= (1+r)J_{N-1}(x) \end{aligned}$$

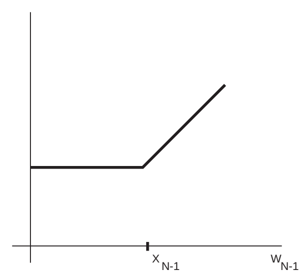
  - Now suppose $\alpha_{N-2} \leq x < \alpha_{N-1}$. Then

$$J_{N-2}(x) = (1+r)^2 x \geq (1+r)J_{N-1}(x) > (1+r)^2 x?!$$

    since $x < \alpha_{N-1}$.
  - Similar analysis for remaining stages.

Extension: Keep Old Offers

- Stage-invariant dynamics:
$$x^+ = \max(x, w)$$
- DP iterations:
  - Stage N: $J_N(x_N) = x_N$
  - Stage $N-1$:
$$J_{N-1}(x_{N-1}) = \max_{\text{sell/don't sell}} \{(1+r)x_{N-1}, E\{\max(x_{N-1}, w_{N-1})\}\}$$
  - Note for later:
$$J_{N-1}(z) \geq (1+r)z = (1+r)J_N(z)$$
- Q: What is $E\{\max(x_{N-1}, w_{N-1})\}$?
  - For convenience, assume w is an RV over $\mathcal{R}$.
$$P(w \in [\underline{w}, \overline{w}]) = \int_{\underline{w}}^{\overline{w}} p(w)dw \text{ vs } \sum_{w_i \in [\underline{w}, \overline{w}]} p_i$$



  - $\max(x_{N-1}, w_{N-1})$ either equals x_{N-1} or w_{N-1}, so
$$\begin{aligned} E\{\max(x_{N-1}, w_{N-1})\} &= P(w_{N-1} \leq x_{N-1})x_{N-1} + \int_{x_{N-1}}^{\infty} wp(w)dw \\ &= \int_0^{x_{N-1}} p(w)dw x_N + \int_{x_{N-1}}^{\infty} wp(w)dw \end{aligned}$$
- Q: What is sell/don't sell threshold?
$$(1+r)x_{N-1} = \int_0^{x_{N-1}} p(w)dw x_N + \int_{x_{N-1}}^{\infty} wp(w)dw$$

Keep Old Offers, cont (2)

- Let's inspect
$$\phi(\alpha) \stackrel{\text{def}}{=} (1+r)\alpha - \int_0^\alpha p(w)dw\alpha - \int_\alpha^\infty wp(w)dw$$
  - For $\alpha = 0$:
  $$\phi(\alpha) = -\int_0^\infty wp(w)dw = -E\{w\} < 0$$
  - For $\alpha \to \infty$:
  $$\phi(\alpha) \to r\alpha > 0$$
  - In between
  $$\begin{aligned} \frac{d\phi}{d\alpha} &= (1+r) - \alpha p(\alpha) - \int_0^\alpha p(w)dw + \alpha p(\alpha) \\ &= (1+r) - \int_0^\alpha p(w)dw \geq r > 0 \end{aligned}$$

  So $\phi(\alpha) = 0$ at only 1 point.

- Let α^* satisfy
$$(1+r)\alpha^* = \alpha^* \int_0^{\alpha^*} p(w)dw + \int_{\alpha^*}^\infty wp(w)dw$$
Optimal policy: Sell if $x_{N-1} \geq \alpha^*$.
Optimal cost:
$$J_{N-1}(x_{N-1}) = \begin{cases} (1+r)x_{N-1} & x_{N-1} \geq \alpha^* \\ P(w_{N-1} \leq x_{N-1})x_{N-1} + \int_{x_{N-1}}^\infty wp(w)dw & x_{N-1} < \alpha^* \end{cases}$$

Keep Old Offers, cont (3)

- Step back again...

$$J_{N-2}(x_{N-2}) = \max\left\{(1+r)^2 x_{N-2}, E\left\{J_{N-1}(\max(x_{N-2}, w_{N-2}))\right\}\right\}$$

- FACT: $x_{N-2} \geq \alpha^*$ then should sell.
- PROOF:
  - $x_{N-2} \geq \alpha^* \Rightarrow$
  - $x_{N-1} \geq \alpha^* \Rightarrow$
  - $J_{N-1}(x_{N-1}) = (1+r)x_{N-1} \Rightarrow$

$$\begin{aligned} J_{N-2}(x_{N-2}) &= \max\left\{(1+r)^2 x_{N-2}, (1+r)E\left\{\max(x_{N-2}, w_{N-2})\right\}\right\} \\ &= (1+r)\max\left\{(1+r)x_{N-2}, E\left\{\max(x_{N-2}, w_{N-2})\right\}\right\} \\ &= (1+r)^2 x_{N-2} \end{aligned}$$

  since this is same analysis for J_{N-1}.

- This shows non-intuitive result that

$$\alpha_{N-2} \leq \alpha_{N-1} = \alpha^*$$

i.e., don't hold out for better offer at earlier stage?!

Keep Old Offers, cont (4)

- FACT: $J_{N-2}(z) \geq (1+r)J_{N-1}(z)$
- PROOF:

$$J_{N-1}(z) = \max\left\{(1+r)z, E\left\{J_N(\max(z,w))\right\}\right\}$$

$$J_{N-2}(z) = \max\left\{(1+r)^2 z, E\left\{J_{N-1}(\max(z,w))\right\}\right\}$$

$$(\text{ first term })_{N-2} \geq (1+r)\,(\text{ first term })_{N-1}$$

$$(\text{ second term })_{N-2} \geq (1+r)\,(\text{ second term })_{N-1}$$

$$\Rightarrow$$

$$J_{N-2}(z) \geq (1+r)J_{N-1}(z)$$

- Now suppose $\alpha_{N-2} \leq x_{N-2} < \alpha^*$:

$$\begin{aligned} J_{N-2}(x_{N-2}) &= (1+r)^2 x_{N-2} \\ &\geq (1+r)J_{N-1}(x_{N-2}) \\ &= (1+r)\max\left\{(1+r)x_{N-2}, E\left\{J_N(\max(x_{N-2},w))\right\}\right\} \\ &= (1+r)E\left\{J_N(\max(x_{N-2},w))\right\} \quad (\text{because } x_{N-2} < \alpha^*) \\ &> (1+r)^2 x_{N-2}?! \end{aligned}$$

- This shows $\alpha_{N-2} = \alpha_{N-1} = \alpha^*$, i.e., same threshold.
- Similar analysis shows same threshold for ANY stage. Another example where myopic policy is optimal.

Monotonicity

- Previous analysis closely related to theoretically important "monotonicity" property of DP.
- Setup: Time invariant system & costs:

$$x^+ = f(x, u, w)$$

$$\text{cost} = g_N(x_N) + \sum_{k=0}^{N-1} g(x_k, u_k, w_k)$$

  i.e., no explicit dependence on stage.

- FACT: If for all x:

$$J_{N-1}(x) \leq J_N(x)$$

  then

$$J_k(x) \leq J_{k+1}(x), \quad \text{for all } x, k$$

- NOTE: We are not comparing $J_k(x_k)$ vs $J_{k+1}(x_{k+1})$!
- PROOF: Assume $J_k(x) \leq J_{k+1}(x)$. Then

$$\begin{aligned} J_{k-1}(x) &= \min_u E\{g(x,u,w) + J_k(f(x,u,w))\} \\ &\leq \min_u E\{g(x,u,w) + J_{k+1}(f(x,u,w))\} \\ &= J_k(x) \end{aligned}$$

- Likewise

$$J_{N-1}(x) \geq J_N(x)$$

$$\Rightarrow$$

$$J_k(x) \geq J_{k+1}(x), \quad \text{for all } x, k$$