Lecture 05

Saturday, 14 January 2023 11:09 AM

<u>Proposition 1</u>: For every initial state x_0, the optimal cost $J^*(x_0)$ of the basic problem is equal to $J_0(x_0)$ where the function J_0 is given by the last step of the algorithm, which proceeds backward in time from period $N-1$ to period 0.:

$$J_N(x_N) = g_N(x_N) \quad —① $$

$$J_k(x_k) := \min_{u_k \in U_k(x_k)} \mathbb{E}_{w_k} \left[g_k(x_k, u_k, w_k) + J_{k+1}\left(f_k(x_k, u_k, w_k)\right) \right]$$

$$k = 0, 1, \ldots, N-1 \quad —②$$

where the expectation is taken w.r.t. the prob. dist. of w_k, which depends on x_k and u_k. Furthermore, if $u_k^* = \mu_k^*(x_k)$ minimizes the r.h.s of (2) for each x_k and k, the policy $\pi^* = \{\mu_0^*, \ldots, \mu_{N-1}^*\}$ is optimal.

<u>Proof</u>: For any admissible policy $\pi = \{\mu_0, \mu_1, \ldots, \mu_{N-1}\}$ and each $k = 0, 1, \ldots, N-1$. denote $\pi^k = \{\mu_k, \mu_{k+1}, \ldots, \mu_{N-1}\}$. For $k = 0, 1, \ldots, N-1$, let $J_k^*(x_k)$ be the optimal cost for the $(N-k)$ stage problem that starts at state x_k and time k, and ends at time N; that is

$$J_k^*(x_k) = \min_{\pi^k} \mathbb{E}_{w_k, \ldots, w_{N-1}} \left[g_N(x_N) + \sum_{i=k}^{N-1} g_i(x_i, \mu_i(x_i), w_i) \right]$$

For $k = N$, we define $J_N^*(x_N) = g_N(x_N)$. We will show by induction that $J_k^* = J_k \;\forall\, k = 0, 1, \ldots, N$, so that for $k = 0$ we get the desired result.

Indeed for $k = N$, $J_N^* = J_N = g_N$. Assume that for some k and all x_{k+1}, we have

$$J_{k+1}^*(x_{k+1}) = J_{k+1}(x_{k+1})$$

Then since $\pi_k = (\mu_k, \pi^{k+1})$, we have for all x_k

$$J_k^*(x_k) = \min_{(\mu_k, \pi^{k+1})} \mathbb{E} \left[g_k(x_k, \mu_k(x_k), w_k) + g_N(x_N) \right.$$

$$J_k^*(x_k) = \min_{(\mu_k,\pi^{k+1})} \mathbb{E}_{w_k,\dots,w_{N-1}} \Big[g_k(x_k,\mu_k(x_k),w_k) + g_N(x_N) + \sum_{i=k+1}^{N-1} g_i(x_i,\mu_i(x_i),w_i) \Big]$$

$$= \min_{\mu_k} \mathbb{E}_{w_k} \Big[g_k(x_k,\mu_k(x_k),w_k) + \min_{\pi^{k+1}} \mathbb{E}_{w_{k+1},\dots,w_{N-1}} \Big[g_N(x_N) + \sum_{i=k+1}^{N-1} g_i(x_i,\mu_i(x_i),w_i) \Big]\Big]$$

$$= \min_{\mu_k} \mathbb{E}_{w_k} \Big[g_k(x_k,\mu_k(x_k),w_k) + J_{k+1}^*(x_{k+1}) \Big]$$

$$= \min_{\mu_k} \mathbb{E}_{w_k} \Big[g_k(x_k,\mu_k(x_k),w_k) + J_{k+1}(x_{k+1}) \Big]$$

$$= \min_{u_k \in U(x_k)} \mathbb{E}_{w_k} \Big[g_k(x_k,u_k,w_k) + J_{k+1}(f_k(x_k,u_k,w_k)) \Big]$$

$$= J_k(x_k)$$

We converted the minimization over μ_k to a minimization over u_k, using the fact that for any function F of x and u,

$$\min_{\mu \in M} F(x,\mu(x)) = \min_{u \in U(x)} F(x,u)$$

where M is the set of all functions μ s.t. $\mu(x) \in U(x)$. ▨

Example: Consider an inventory control problem that is slightly different from the previous one.

Assume that u_k and the demand w_k are nonnegative integers and the excess demand $(w_k - x_k - u_k)$ is lost. So the system dynamics become

$$x_{k+1} = \max(0, x_k + u_k - w_k)$$

Assume that the store can keep 2 unit of goods at a time. The holding/shortage cost for k^{th} period is given by

$$(x_k + u_k - w_k)^2$$

Unit price is 1. So the cost per period is

$$g_k(x_k,u_k,w_k) = u_k + (x_k+u_k-w_k)^2$$

$$g_k(x_k, u_k, w_k) = u_k + (x_k + u_k - w_k)^2$$

The terminal cost is assumed to be 0, i.e., $g_N = 0$

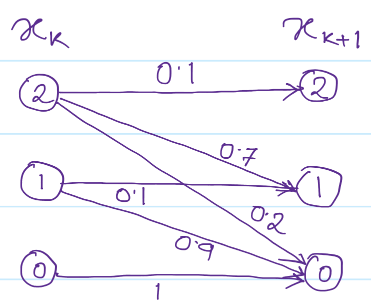
The planning horizon, $N = 3$. Initial stock, $x_0 = 0$.

Demand w_k has the following distribution for $k = 0, 1, 2$

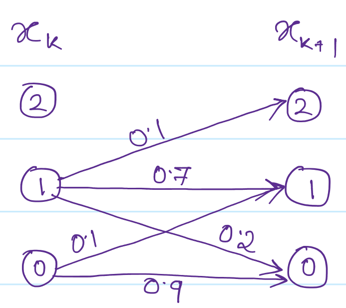
$$P(w_k = 0) = 0.1, \quad P(w_k = 1) = 0.7, \quad P(w_k = 2) = 0.2$$

* The system can also be presented in terms of transition probabilities which will help us to compute expectations.

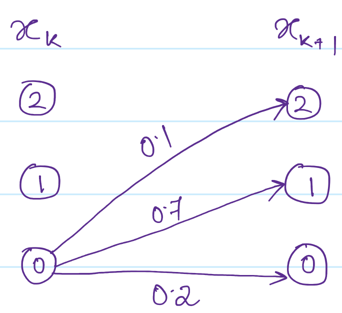
- When $u_k = 0$



- When $u_k = 1$



- When $u_k = 2$



The starting equation for the DP algorithm is

$$J_3(x_3) = 0$$

Since the terminal cost is 0. The algorithm takes the form

$$J_k(x_k) = \min_{0 \le u_k \le 2 - x_k} \mathbb{E}\left[u_k + (x_k + u_k - w_k)^2 + J_{k+1}\left(\max(0, x_k + u_k - w_k)\right) \right]$$

where $k = 0, 1, 2$, and x_k, u_k, w_k can take value in $\{0, 1, 2\}$.

where $k = 0, 1, 2$, and x_k, u_k, w_k can take value in $\{0, 1, 2\}$.

Period 2:

$$J_2(0) = \min_{u_2 = 0,1,2} \mathop{\mathbb{E}}_{w_2} \left[u_2 + (u_2 - w_2)^2 \right]$$
$$= \min_{u_2 = 0,1,2} \left[u_2 + 0.1 \times u_2^2 + 0.7 (u_2 - 1)^2 + 0.2 (u_2 - 2)^2 \right]$$

For $u_2 = 0$, $\mathbb{E}[\cdot] = 0 + 0 + 0.7 + 0.2 \times 4 = 1.5$

For $u_2 = 1$, $\mathbb{E}[\cdot] = 1 + 0.1 + 0 + 0.2 = 1.3$

For $u_2 = 2$, $\mathbb{E}[\cdot] = 2 + 0.1 \times 4 + 0.7 + 0 = 3.1$

So, $J_2(0) = 1.3$ and $\mu_2^*(0) = 1$

Similarly,

$$J_2(1) = \min_{u_2 = 0,1} \mathop{\mathbb{E}}_{w_2} \left[u_2 + (1 + u_2 - w_2)^2 \right]$$
$$= \min_{u_2 = 0,1} \left[u_2 + 0.1 (1 + u_2)^2 + 0.7 \times u_2^2 + 0.2 \times (u_2 - 1)^2 \right]$$

For $u_2 = 0$, $\mathbb{E}[\cdot] = 0 + 0.1 + 0 + 0.2 = 0.3$

For $u_2 = 1$, $\mathbb{E}[\cdot] = 1 + 0.1 \times 4 + 0.7 + 0 = 2.1$

So, $J_2(1) = 0.3$ and $\mu_2^*(1) = 0$

for $x_2 = 2$ there is only one feasible u_2 so, $\mu_2^*(2) = 0$.

and $J_2(2) = \mathop{\mathbb{E}}_{w_2} \left[0 + (2 + 0 - w_2)^2 \right]$

$$= 0.1 \times 2^2 + 0.7 \times 1^2 + 0 = 1.1$$

Period 1: Again we compute $J_1(x_1)$ for each of the three posible state $x_1 = 0, 1, 2$ using values of $J_2(0)$, $J_2(1)$ and $J_2(2)$.

$$J_1(0) = \min_{u_1 = 0,1,2} \mathop{\mathbb{E}}_{w_1} \left[u_1 + (u_1 - w_1)^2 + J_2(\max(0, u_1 - w_1)) \right]$$

For $u_1 = 0$: $\mathbb{E}[\cdot] = 0.1 \cdot J_2(0) + 0.7 (1 + J_2(0)) + 0.2 (4 + J_2(0))$
$= 2.8$

For $u_1 = 1$: $\mathbb{E}[\cdot] = 1 + 0.1 (1 + J_2(1)) + 0.7 \times J_2(0) + 0.2 (1 + J_2(0))$
$= 2.5$

For $u_1 = 1$: $\mathbb{E}[\cdot] = 1 + 0.1\,(1 + J_2(1)) + 0.7 \times J_2(0) + 0.2\,(1 + J_2(0))$

$= 2.5$

For $u_1 = 2$: $\mathbb{E}[\cdot] = 2 + 0.1\left(4 + J_2(2)\right) + 0.7 \times \left(1 + J_2(1)\right) + 0.2 \times J_2(0)$

$= 3.68$

$J_1(0) = 2.5, \quad \mu_1^*(0) = 1$

For $x_1 = 1$, we have

$$J_1(1) = \min_{u_1 = 0,1} \mathop{\mathbb{E}}_{w_1}\left[u_1 + (1 + u_1 - w_1)^2 + J_2\left(\max(0, 1 + u_1 - w_1)\right)\right]$$

$u_1 = 0$: $\mathbb{E}[\cdot] = 0.1\left(1 + J_2(1)\right) + 0.7 \cdot J_2(0) + 0.2\left(1 + J_2(0)\right)$

$= 1.2$

$u_1 = 1$: $\mathbb{E}[\cdot] = 1 + 0.1\left(4 + J_2(2)\right) + 0.7\left(1 + J_2(1)\right) + 0.2 \times J_2(0)$

$= 2.68$

$J_1(1) = 1.2, \quad \mu_1^*(1) = 0$

For $x_1 = 2$, the only admissible control is $u_1 = 0$, so we have

$$J_1(2) = \mathop{\mathbb{E}}_{w_1}\left[(2 - w_1)^2 + J_2\left(\max(0, 2 - w_1)\right)\right]$$

$$= 0.1\left(4 + J_2(2)\right) + 0.7\left(1 + J_2(1)\right) + 0.2\, J_2(0)$$

$$= 1.68$$

and $\mu_1^*(2) = 0$

<u>Period 0</u> : Here we need to compute $J_0(0)$ as $x_0 = 0$.

$$J_0(0) = \min_{u_0 = 0,1,2} \mathop{\mathbb{E}}_{w_0}\left[u_0 + (u_0 - w_0)^2 + J_1\left(\max(0, u_0 - w_0)\right)\right]$$

For $u_0 = 0$: $\mathbb{E}[\cdot] = 0.1 \times J_1(0) + 0.7 \times \left(1 + J_1(0)\right) + 0.2\left(4 + J_1(0)\right)$

$= 4.0$

For $u_0 = 1$: $\mathbb{E}[\cdot] = 1 + 0.1\left(1 + J_1(1)\right) + 0.7 \times J_1(0) + 0.2\left(1 + J_1(0)\right)$

$= 3.67$

For $u_0 = 2$: $\mathbb{E}[\cdot] = 2 + 0.1\left(4 + J_1(2)\right) + 0.7\left(1 + J_1(1)\right) + 0.2\, J_1(0)$

$= 5.108$

$\therefore \; J_0(0) = 3.67 \;, \quad \mu_0^*(0) = 1 .$