Solution 4

Exercise 6.2

a) Consider the CEC applied to this problem. At stage 1 we solve the deterministic problem:

$$\min_{u_1} (0 + J_2(x_2)) = \min_{u_1} \|x_2\| = \min_{u_1} \|x_1 + bu_1 + d\bar{w}_1\|$$
$$= \min_{u_1} \|x_1 + bu_1\|.$$

Thus, $\mu_1^*(x_1) = -x_1^1$ (i.e. the first coordinate of x_1), and the optimal cost to go is $J_1(x_1) = \left\| \begin{matrix} 0 \\ x_1^2 \end{matrix} \right\|$ (where x_1^2 is the second coordinate of x_1).

At stage 0 we solve the deterministic problem:

$$\min_{u_0} (0 + J_1(x_1)) = \min_{u_0} \left\| \begin{matrix} 0 \\ x_1^2 \end{matrix} \right\| = \min_{u_0} \left\| \begin{matrix} 0 \\ x_0^2 + (\sqrt{2}/2)\bar{w}_0 \end{matrix} \right\| = 0;$$

so μ_0 is unconstrained. We choose $\mu_0^*(0) = 0$

The corresponding cost of the CEC is:

$$E\{\|x_2\|\} = E\{\|x_0 + b\mu_0^*(0) + dw_0 + b\mu_1^*(x_0 + b\mu_0^*(0) + dw_0) + dw_1\|\}$$

$$= E\left\{\left\| \begin{bmatrix} 0 \\ \sqrt{2}/2 \end{bmatrix} w_0 + \begin{bmatrix} 1/2 \\ \sqrt{2}/2 \end{bmatrix} w_1 \right\|\right\}.$$

Using the probability distribution of w_0 and w_1, it is straightforward to obtain $E\{\|x_2\|\} = 1$.

b) Consider the open-loop optimal policy. Here we solve the problem:

$$\min_{u_0,u_1} E\{\|x_2\|\} = \min_{u_0,u_1} E\{\|x_0 + b(u_0 + u_1) + d(w_0 + w_1)\|\}.$$

For the given values of x_0, b and d, and the probability distribution of w_0 and w_1, the problem is written as:

$$\min_{u_0,u_1} \frac{1}{4} \left\{ \left\| \begin{matrix} u_0 + u_1 + 1 \\ \sqrt{2} \end{matrix} \right\| + \left\| \begin{matrix} u_0 + u_1 \\ 0 \end{matrix} \right\| + \left\| \begin{matrix} u_0 + u_1 \\ 0 \end{matrix} \right\| + \left\| \begin{matrix} u_0 + u_1 - 1 \\ -\sqrt{2} \end{matrix} \right\| \right\}.$$

It is straightforward to check that the minimum is attained when $u_0 + u_1 = 0$, in which case we obtain the optimal open loop cost as:

$$\frac{1}{4} \left\| \begin{matrix} 1 \\ \sqrt{2} \end{matrix} \right\| + \left\| \begin{matrix} -1 \\ -\sqrt{2} \end{matrix} \right\| = \frac{\sqrt{3}}{2}.$$

Therefore, the CEC is strictly suboptimal.

c) Consider the closed-loop optimal policy. Here:

$$J_0(0) = \min_{u_0} \left\{ \frac{1}{2} \min_{u_1} \left\{ \frac{1}{2} \left\| \begin{matrix} u_0 + u_1 + 1 \\ \sqrt{2} \end{matrix} \right\| + \frac{1}{2} \left\| \begin{matrix} u_0 + u_1 \\ 0 \end{matrix} \right\| \right\} + \frac{1}{2} \min_{u_1} \left\{ \frac{1}{2} \left\| \begin{matrix} u_0 + u_1 - 1 \\ -\sqrt{2} \end{matrix} \right\| + \frac{1}{2} \left\| \begin{matrix} u_0 + u_1 \\ 0 \end{matrix} \right\| \right\} \right\}.$$

Using the figure from part b., it is seen that the optimal value is to take u_1 so that $u_0 + u_1 = 0$ and the same optimal value as in the open-loop case is obtained.

Exercise 6.10

Consider the example in which $r(x(t)) = 1$, $x(0) = (0, 0)$, and $x(T) = (a, b)$. Then minimizing

$$\int_0^T r(x(t))\, dt$$

over the control constraint $\|u(t)\| = 1$ corresponds to finding the shortest trajectory from $x(0)$ to $x(T)$. The solution to this problem is clearly a straight line from $(0, 0)$ to (a, b), which yields a distance $\sqrt{a^2 + b^2}$. However, the discretization provided does not approach this distance if a and b are both nonzero. The discretization provided only allows moves in vertical and horizontal directions, and thus the shortest distance becomes $a + b$, regardless of the discretization size Δ.

Exercise 6.16

By substituting $D_k = p^k$ for $G_k = (p(2-p))^k$ into the derivation on pps. 319-320, we have $R_k = p(2-p)R_{k-1} + p^2 D_{k-1}(1 - R_{k-1})$, with $R_0 = 1$. Dividing both sides by $D_k = pD_{k-1}$, we have:

$$\frac{R_k}{D_k} = (2-p)\frac{R_{k-1}}{D_{k-1}} + p(1 - R_{k-1})$$

As $k \to \infty$, $D_k \to 0$, meaning $R_k \to 0$ also. So we obtain for large N:

$$\frac{R_N}{D_N} = O((2-p)^N)$$

Because $2 - p > 1$, $\frac{R_k}{D_k}$ increases exponentially with k.

Exercise 6.20

(a) **Prop.6.3.1**: Assume that for all x_k and k, we have

$$\min_{u_k \in \overline{U}_k(x_k)} \max_{w_k \in W_k(x_k, u_k)} \left[g_k(x_k, u_k, w_k) + \widetilde{J}_{k+1}(f_k(x_k, u_k, w_k)) \right] \leq \widetilde{J}_k(x_k). \tag{1}$$

Then the cost-to-go functions corresponding to a one-step lookahead policy that uses $\widetilde{J}_k$ and $\overline{U}_k(x_k)$ satisfy for all x_k and k

$$\overline{J}_k(x_k) \leq \min_{u_k \in \overline{U}_k(x_k)} \max_{w_k \in W_k(x_k, u_k)} \left[g_k(x_k, u_k, w_k) + \widetilde{J}_{k+1}(f_k(x_k, u_k, w_k)) \right]. \tag{2}$$

We define

$$\widehat{J}_k(x_k) = \min_{u_k \in \overline{U}_k(x_k)} \max_{w_k \in W_k(x_k, u_k)} \left[g_k(x_k, u_k, w_k) + \widetilde{J}_{k+1}(f_k(x_k, u_k, w_k)) \right]$$

and through a backward induction approach similar to that in Prop.6.3.1, the above conclusion in (2) can be proved.

Prop.6.3.2: Let $\widetilde{J}_k(x), k = 0, 1, ..., N$, be functions of x_k with $\widetilde{J}_k(x_N) = g_N(x_N)$ for all x_N, and let $\pi = \{\overline{\mu}_0, \cdots, \overline{\mu}_{N-1}\}$ be a policy such that for all x_k and k, we have

$$\max_{w_k \in W_k(x_k, u_k)} \left[g_k(x_k, u_k, w_k) + \widetilde{J}_{k+1}(f_k(x_k, u_k, w_k)) \right] \leq \widetilde{J}_k(x) + \delta_k, \tag{3}$$

where $\delta_0, \delta_1, \cdots, \delta_{N-1}$ are some scalars. Then for all x_k and k, we have

$$J_{\pi,k}(x_k) \leq \widetilde{J}_k(x) + \sum_{i=k}^{N-1} \delta_i, \tag{4}$$

where $J_{\pi,k}(x_k)$ is the cost-to-go of π starting from state x_k at stage k. Through a backward induction approach similar to that in Prop.6.3.2, the above conclusion in (4) can be proved.

(b) In a rollout algorithm, since for all x_k and k we have$\mu_k(x_k) \in \overline{U}_k(x_k)$, the assumption in (1) is satisfied and the desired result directly follows (2).

6.231 Dynamic Programming and Stochastic Control
Fall 2015