Example (Dynamic routing to two exponential server queues)

Consider a system with two service stations, each consisting of a queue plus server, as shown in Figure 1. Customers arrive according to a Poisson process of rate $\lambda > 0$. A customer is routed to the first station with probability u, and to the second station with probability $1-u$, where u is a variable to be controlled, and can depend on the state of the system. Service times at the first station are exponentially distributed with parameter μ_1, and service times at the second station are exponentially distributed with paramenter μ_2. The holding cost per unit time per customer in station 1 is c_1 and in station 2 is c_2, where $c_i > 0$ for $i = 1, 2$. This example fits into the dynamic programming framework we have described, with the following choices:

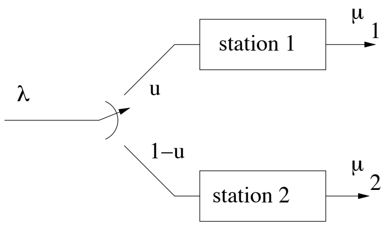

Figure 1: Dynamic routing to two service stations

$U = [0, 1]$

$\mathcal{S} = \mathbb{Z}_+^2$, where state $x = (x_1, x_2)$ denotes x_i customers in station i, for $i = 1, 2$.

$g(x) = c_1x_1 + c_2x_2$

F corresponds to the exponential distribution with parameter $\gamma = \lambda + \mu_1 + \mu_2$.

$p_{x,y}(u) = \frac{1}{\gamma}\left(\mu_1 I_{\{y=D_1x\}} + \mu_2 I_{\{y=D_2x\}} + u\lambda I_{\{y=A_1x\}} + (1-u)\lambda I_{\{y=A_2x\}}\right)$ where
$A_1x = (x_1 + 1, x_2)$, $A_2x = (x_1, x_2 + 1)$, $D_1x = ((x_1 - 1)_+, x_2)$, and $D_2x = (x_1, (x_2 - 1)_+)$.

The backwards equation of dynamic programming becomes

$$V_{n+1}(x) = g(x) + \min_{0\le u\le 1} \frac{\beta}{\gamma}\left\{\mu_1 V_n(D_1x) + \mu_2 V_n(D_2x) + u\lambda V_n(A_1x) + (1-u)\lambda V_n(A_2x)\right\}$$

or, after plugging in the optimal value for u,

$$V_{n+1}(x) = g(x) + \frac{\beta}{\gamma}\left\{\mu_1 V_n(D_1x) + \mu_2 V_n(D_2x) + \lambda \min\{V_n(A_1x), V_n(A_2x)\},\right\}$$

with the intial condition $V_0 \equiv 0$. Furthermore, the optimal control is given in feedback form as follows. Given there are n steps to go and the state is x, the optimal control action is $u_n^*(x)$, given by

$$u_n^*(x) = \begin{cases} 1 & \text{if } V_n(A_1x) \le V_2(A_2x) \\ 0 & \text{else.} \end{cases} \tag{1}$$

That is, if the current state is x and an arrival occurs, the optimal control routes the arrival to whichever station yields the lower expected cost.

Consider the symmetric case: $\mu_1 = \mu_2$ and $c_1 = c_2$. We will prove that $u_n^*(x) = I_{\{x_1 \le x_2\}}$ for all $n \ge 1$. That is, the send to shorter policy is optimal. It is easy to check that the control specification (1) is equivalent to $u_n^*(x) = I_{\{x_1 \le x_2\}}$, if V_n has the following three properties:

1. (symmetry) $V(x_1, x_2) = V(x_2, x_1)$ for all $(x_1, x_2) \in \mathcal{S}$
2. (monotonicity) $V(x_1, x_2)$ is nondecreasing in x_1 and in x_2.
3. $V(x_1, x_2) \le V(x_1 - 1, x_2 + 1)$ whenever $0 \le x_1 \le x_2$.

Together properties 1 and 3 mean that when restricted to the states along a line segment of the form $\{x \in \mathcal{S} : x_1 + x_2 = l\}$ for $l \ge 1$, the cost function is a monotonically nondecreasing function of the distance from the midpoint.

Thus, it remains to prove that V_n has the three properties stated, for all $n \ge 0$. This is done by induction. Trivially, the function V_0 satisfies all three properties. So for the sake of argument by induction, suppose that V_n satisfies all three properties. We show that V_{n+1} satisfies all three properties. Property 1 for V_{n+1} follows from property 1 for V_n and the symmetry of the backwards equations in x_1 and x_2. Property 2 for V_{n+1} follows from property 2 for V_n. It remains to show that V_{n+1} has property 3. Since property 3 is closed under summation, it suffices to prove that $g(x)$, $V_n(D_1x)$, $V_n(D_2x)$, and $\min\{V_n(A_1x), V_n(A_2x)\}$ have property 3. It is easily checked that the first three of these functions has property 3, so it remains to prove that $\min\{V_n(A_1x), V_n(A_2x)\}$ has property 3. So suppose that $1 \le x_1 \le x_2$, and refer to Figure 2. It must be shown that $a \wedge b \le b \wedge c$ where $a = V_n(x_1 + 1, x_2)$, $b = V_n(x_1, x_2 + 1)$, and $c = V_n(x_1 - 1, x_2 + 2)$. We argue first that $a \le b$, by considering two cases: If $x_1 = x_2$ then the states $(x_1, x_2 + 1)$ and $(x_1 + 1, x_2)$ are obtained from each other by swapping coordinates, so that $a = b$ by property 1 for V_n. If $x_1 > x_2$ then $x_1 \ge x_2 + 1$ so that $a \le b$ by property 3 for V_n. Thus, in general $a \le b$. Similarly, $b \le c$ by property 3 for V_n. Thus, $a \le b \le c$, which immediately implies $a \wedge b \le b \wedge c$, as desired. The proof of the induction step is complete. Therefore, V_n has properties 1-3 for all $n \ge 0$, and the send to shorter queue policy is optimal, in the symmetric case.

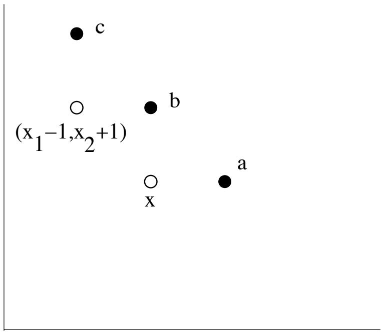

Figure 2: V_n is sampled to get values a, b, and c.

In the general case, the optimal control is given by $u_n^*(x) = I_{\{x_2 > s_n(x_1)\}}$, where s_n is a nondecreasing function. This fact can be established by proving, by induction on n, that the cost-to-go functions V_n have a more elaborate set of properties.