# Assignment 4

(1) $V_0(s_1) = 10 \quad V_0(s_2) = 1 \quad V_0(s_3) = 0$

$$q_1(s_1,a_1) = R(s_1,a_1) + \mathbb{E}[V_0(s')] = 8 + 0.2\cdot 10 + 0.6\cdot 1 + 0.2\cdot 0 = 10.6$$
$$q_1(s_1,a_2) = R(s_1,a_2) + \mathbb{E}[V_0(s')] = 10 + 0.1\cdot 10 + 0.2\cdot 1 + 0.7\cdot 0 = 11.2$$
$$q_1(s_2,a_1) = R(s_2,a_1) + \mathbb{E}[V_0(s')] = 1 + 0.3\cdot 10 + 0.3\cdot 1 + 0.4\cdot 0 = 4.3$$
$$q_1(s_2,a_2) = R(s_2,a_2) + \mathbb{E}[V_0(s')] = -1 + 0.5\cdot 10 + 0.3\cdot 1 + 0.2\cdot 0 = 4.3$$

$\pi_1(s_1) = a_2$ $\quad V_1(s_1) = 11.2$

$\pi_1(s_2) = a_1$ or $a_2$ $\quad V_1(s_2) = 4.3$

$$q_2(s_1,a_1) = R(s_1,a_1) + \mathbb{E}[V_1(s')] = 8 + 0.2\cdot 11.2 + 0.6\cdot 4.3 + 0.2\cdot 0 = 12.82$$
$$q_2(s_1,a_2) = R(s_1,a_2) + \mathbb{E}[V_1(s')] = 10 + 0.1\cdot 11.2 + 0.2\cdot 4.3 + 0.7\cdot 0 = 11.98$$
$$q_2(s_2,a_1) = R(s_2,a_1) + \mathbb{E}[V_1(s')] = 1 + 0.3\cdot 11.2 + 0.3\cdot 4.3 + 0.4\cdot 0 = 5.65$$
$$q_2(s_2,a_2) = R(s_2,a_2) + \mathbb{E}[V_1(s')] = -1 + 0.5\cdot 11.2 + 0.3\cdot 4.3 + 0.2\cdot 0 = 5.89$$

$\pi_2(s_1) = a_1$ $\quad V_2(s_1) = 12.82$

$\pi_2(s_2) = a_2$ $\quad V_2(s_2) = 5.89$

(4) $C_1, C_2, \lambda_1, \lambda_2, h_1, h_2, p_1, p_2$

$N = \{(n_1, m_1, n_2, m_2) \in \mathbb{Z}^4_{\geq 0}$ for $n_1 + m_1 \leq C_1,\ n_2 + m_2 \leq C_2\}$; $T = \emptyset$

$S = N$

$A(s) = \{(m_1', m_2', t)$ s.t. $m_1' + m_1 + n_1 + t \leq C_1,\ m_2' + m_2 + n_2 - t \leq C_2$ and $-n_1 \leq t \leq n_2\}$

I assumed you don't pay holding costs on ~~shipped~~ transferred inventory

(4) I tried running it w/ model params:
$C_1 = 2,\ C_2 = 3$
$\lambda_1 = \lambda_2 = 1$
$h_1 = h_2 = 1$
$p_1 = 10,\ p_2 = 5$
$K_1 = 1;\ K_2 = 2$
$\gamma = 0.9$

Optimal strategy is to always order max from suppliers, but transfer behavior is interesting:
- When state is $(0,0,x,y)$ where $x+y \geq 2$ & $x \geq 1$, it wants to transfer 1 from store 2 to store 1. And when state is $(0,0,x,y)$ where $x \geq 2$ and $x+y=3$, it likes to transfer 2 from 2 to 1.
  - This makes sense since $p_1 >> p_2$, so moving inventory to store 1 is optimal if you can do it w/o running out of inventory in store 2
  - Also makes sense that $\pi(0,0,1,0) = (2,2,0)$ instead of $(1,3,1)$, since you save $5 \cdot e^{-1}$ in expectation in stockout costs, which is less than the cost of the transfer (2)
- You only ever transfer from 1 to 2 in the states $(1,1,0,0)$ and $(2,0,0,0)$, which makes sense.