Applied Math 226

Solutions for Assignment 1.

1. We introduce the value functions

$$V(x,i,M) = \inf E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]$$

where $E_{x,i}$ denotes expected value given $X_i = x$, and the infimum is over feedback controls (note that these controls can depend on the time to go).

We claim that $V(x,i,M) = \bar{V}(x, M-i)$. The proof is by backward induction on i. Let $\{u(x,i)\}$ be any control scheme, where we stop if $u(X_i,i) = 1$, and continue if $u(X_i,i) = 0$. Let N denote the corresponding stopping time. Let $\{\bar{u}(x,i)\}$ be the control defined in terms of $\bar{V}$. Thus $\bar{u}(x,i) = 1$ if $g(x) \le c(x) + \sum_{y\in S} p(x,y)\bar{V}(x, M-i-1)$ and $\bar{u}(x,i) = 0$ otherwise. Our inductive assumption is that $V(x,j,M) = \bar{V}(x,M-j)$ and that $\bar{u}(x,j)$ is optimal for $i < j \le M$. By the Markov property and the definition of V as the minimal cost,

$$\begin{aligned}
&E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right] \\
&= g(x)1_{\{u(x,i)=1\}} + E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right] 1_{\{u(x,i)=0\}} \\
&= g(x)1_{\{u(x,i)=1\}} + E_{x,i}\left[E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]\Bigg| X_{i+1}\right] 1_{\{u(x,i)=0\}} \\
&\ge g(x)1_{\{u(x,i)=1\}} + \left[c(x) + \sum_{y\in S} p(x,y)\ V(y,i+1)\right] 1_{\{u(x,i)=0\}} \\
&= g(x)1_{\{u(x,i)=1\}} + \left[c(x) + \sum_{y\in S} p(x,y)\bar{V}(y, M-i-1)\right] 1_{\{u(x,i)=0\}} \\
&\ge \min\left[g(x), c(x) + \sum_{y\in S} p(x,y)\bar{V}(y,M-i-1)\right] \\
&= \bar{V}(x, M-i).
\end{aligned}$$

Since the control scheme is arbitrary, this shows that $V(x,i,M) \geq \bar{V}(x,M-i)$. Next consider the particular scheme $\{\bar{u}(x,i)\}$. Since $V(x,i,M)$ is the minimal cost,

$$\begin{aligned}
&V(x,i,M)\\
&\leq \quad E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]\\
&= \quad g(x)1_{\{\bar{u}(x,i)=1\}} + E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]1_{\{\bar{u}(x,i)=0\}}\\
&= \quad g(x)1_{\{\bar{u}(x,i)=1\}} + E_{x,i}\left[E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]\middle| X_{i+1}\right]1_{\{\bar{u}(x,i)=0\}}\\
&= \quad g(x)1_{\{\bar{u}(x,i)=1\}} + \left[c(x) + \sum_{y\in S} p(x,y)\bar{V}(y,M-i-1)\right]1_{\{\bar{u}(x,i)=0\}}\\
&= \quad \min\left[g(x), c(x) + \sum_{y\in S} p(x,y)\bar{V}(y,M-i-1)\right]\\
&= \quad \bar{V}(x,M-i).
\end{aligned}$$

Thus $V(x,i,M) \leq \bar{V}(x,M-i)$. Combining and letting $i=0$ we get $V(x,M) = V(x,0,M) = \bar{V}(x,M)$.

2. Let us expand the state space by adding the absorbing state Δ. Let $p(x,y) = r(x,y)$, $p(x,\Delta) = 1 - \sum_{y\in S} r(x,y)$, and $p(\Delta,y) = 0$ if $x,y \in S$. We also define $f(\Delta) = 0$.

Let X_i denote the associated chain, and define N to be the first time Δ is reached. We claim that

$$W_n(x) = E_x\left[\sum_{i=0}^{(N\wedge n)-1} c(X_i) + f(X_{(N\wedge n)})\right].$$

The condition $W_0(x) = f(x)$ holds, and by the Markov property

$$\begin{aligned}
&W_{n+1}(x)\\
&= \quad E_x\left[\sum_{i=0}^{(N\wedge(n+1))-1} c(X_i) + f(X_{(N\wedge(n+1))})\right]
\end{aligned}$$

$$
\begin{aligned}
&= E_x\left[c(x) + E_x\left[\left.\sum_{i=1}^{(N\wedge(n+1))-1} c(X_i) + f(X_{(N\wedge(n+1))})\right| X_1\right]\right] \\
&= c(x) + \sum_{y\in S} p(x,y)E_x\left[\left.\sum_{i=1}^{(N\wedge(n+1))-1} c(X_i) + f(X_{(N\wedge(n+1))})\right| X_1 = y\right] + p(x,\Delta)\cdot 0 \\
&= c(x) + \sum_{y\in S} r(x,y)W_n(y).
\end{aligned}
$$

This proves the representation.

Under the given condition, all states save Δ are transient. In fact an explicit upper bound for the probability to still be in S can be given in terms of

$$
\varepsilon = \min_{x\in S}\left(1 - \sum_{y\in S} r^k(x,y)\right) > 0,
$$

which shows that the probability that the chain is *not* absorbed by time n decays faster than $K(1-\varepsilon/k)^n$. Thus the exit time N is integrable. By Lebesque Dominated Convergence,

$$
W_n(x) = E_x\left[\sum_{i=0}^{(N\wedge n)-1} c(X_i) + f(X_{(N\wedge n)})\right] \to W(x) = E_x\left[\sum_{i=0}^{N-1} c(X_i)\right].
$$

3. Define $N = \min\{i : U_i(X_i) = 1\}$. Let G be an arbitrary continuous function of the argument $(X_1,\ldots,X_n,N1_{\{N\le n\}})$. Using the definition of conditional expectation, it is enough to show that

$$
\begin{aligned}
&E\left[E\left[F(X_{n+1},\ldots,X_m)|\, X_n\right]G(X_1,\ldots,X_n,N1_{\{N\le n\}})\right] \\
&\quad = E\left[F(X_{n+1},\ldots,X_m)G(X_1,\ldots,X_n,N1_{\{N\le n\}})\right].
\end{aligned}
$$

We condition on $F_n = \sigma(X_1,\ldots,X_n,U_1(\cdot),\ldots,U_n(\cdot))$. By the Markov property of X_n and the independence of the random vector fields from this chain,

$$
E\left[F(X_{n+1},\ldots,X_m)|\, X_n\right] = E\left[F(X_{n+1},\ldots,X_m)|\, X_1,\ldots,X_n,U_1(\cdot),\ldots,U_n(\cdot)\right].
$$

Thus

$$
\begin{aligned}
&E\left[E\left[F(X_{n+1},\ldots,X_m)|\, X_n\right]G(X_1,\ldots,X_n,N1_{\{N\le n\}})\right] \\
&\quad = E\left[E\left[F(X_{n+1},\ldots,X_m)|\, X_1,\ldots,X_n,U_1(\cdot),\ldots,U_n(\cdot)\right]G(X_1,\ldots,X_n,N1_{\{N\le n\}})\right]
\end{aligned}
$$

$$
\begin{aligned}
&= E\left[E\left[F(X_{n+1},\dots,X_m)G(X_1,\dots,X_n,N1_{\{N\le n\}})\middle| X_1,\dots,X_n,U_1(\cdot),\dots,U_n(\cdot)\right]\right]\\
&= E\left[F(X_{n+1},\dots,X_m)G(X_1,\dots,X_n,N1_{\{N\le n\}})\right].
\end{aligned}
$$

For $M<\infty$ and a given admissible stopping time N, let

$$
W_M(x,N) = E_x\left[\sum_{i=0}^{(N\wedge M)-1} c(X_i) + g(X_N)1_{\{N\le M\}} + \bar V(X_M)1_{\{M<N\}}\right]
$$

and

$$
W_M(x,i,N) = E_{x,i}\left[\sum_{j=i}^{(N\wedge M)-1} c(X_j) + g(X_N)1_{\{N\le M\}} + \bar V(X_M)1_{\{M<N\}}\right].
$$

Then

$$
\begin{aligned}
&E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j) + g(X_{M\wedge N})\right]\\
&= E_{x,i}\left[E_{x,i}\left[g(x)1_{\{u(x,i)=1\}} + E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j)+g(X_{M\wedge N})\right]1_{\{u(x,i)=0\}}\middle| 1_{\{u(x,i)=1\}}\right]\right]\\
&= g(x)E_{x,i}\left[1_{\{u(x,i)=1\}}\right] + E_{x,i}\left[E_{x,i}\left[\sum_{j=i}^{(M\wedge N)-1} c(X_j)+g(X_{M\wedge N})\middle| 1_{\{u(x,i)=1\}}\right]1_{\{u(x,i)=0\}}\right]\\
&= g(x)E_{x,i}\left[1_{\{u(x,i)=1\}}\right] + E_{x,i}\left[E_{x,i}\left[c(x)+W_M(X_{i+1},i+1,N)\middle| 1_{\{u(x,i)=1\}}\right]1_{\{u(x,i)=0\}}\right]\\
&= g(x)E_{x,i}\left[1_{\{u(x,i)=1\}}\right] + E_{x,i}\left[E_{x,i}\left[c(x)+W_M(X_{i+1},i+1,N)\right]1_{\{u(x,i)=0\}}\right]\\
&= g(x)P_{x,i}\left\{u(x,i)=1\right\} + E_{x,i}\left[c(x)+W_M(X_{i+1},i+1,N)\right]P_{x,i}\left\{u(x,i)=0\right\}\\
&\ge \min\left[g(x), c(x)+\sum_{y\in S} p(x,y)W_M(y,i+1,N)\right].
\end{aligned}
$$

By induction, this shows

$$
W_M(x,N)\ge W_M(x,M,N)\ge \bar V(x).
$$

We now let $M\to\infty$. Since $c>0$, the cost is ∞ if N is not integrable. When N is integrable, LDCT gives

$$
W(x,N) = \inf E_x\left[\sum_{i=0}^{N-1} c(X_i) + g(X_N)\right] \ge \bar V(x).
$$

Since the stopping time is arbitrary, $V(x)\ge \bar V(x)$. The reverse inequality is exactly as in class for the case where we restricted to feedback controls.