MAE 288A
Assignment 3
Due Thursday, 12 Nov., 2009

Note: You may find it helpful to write code in order to solve some of these problems. You may work in Matlab, C, C++, Java, Fortran, Maple, or Mathematica. Please include copies of any codes used.

1. (5) Define $F[\phi]$ for $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$ by
   
   $$F[\phi](x) = \min_{u \in U} \{ l(x, u) + \alpha E_w [\phi(f(x, u, w))] \}$$

   where one assumes sufficient conditions so that this expression is well-defined. For two functions, $\phi$ and $\psi$, we say $\phi \leq \psi$ if $\phi(x) \leq \psi(x)$ for all $x$. Show that if $\phi \leq \psi$, then $F[\phi] \leq F[\psi]$.

2. (5) Consider the iteration
   
   $$x_{n+1} = \alpha x_n - 0.25 e^{1-x_n^2}.$$ 

   For what values of $\alpha$, do the conditions of the Banach Fixed-Point Theorem hold? When they hold, what equation will the iteration limit solve?

3. (10) Consider the discounted-cost, infinite time-horizon control problem with dynamics
   
   $$\xi_{t+1} = (1 - u_t)\xi_t + w_t$$

   where $\xi_t, u_t, w_t$ are scalar-valued. Suppose $U = \{1/2, 3/4\}$. Let the $w_t$ be IID normal random variables with mean 0 and variance 1. Let the cost criterion be
   
   $$J(x, \bar{u}) = E_{\xi_0=x} \left\{ \sum_{t=0}^{\infty} \alpha^t \left[ \frac{1}{2} \xi_t^2 + \frac{1}{2} (\bar{u}_t(\xi_t))^2 \right] \right\}$$ 

   where $\alpha = 3/4$. Write down the DPE for this problem. Starting with initial guess, $V_0 \equiv 0$, perform two steps in a value iteration/successive
approximations/ Banach Fixed Point Theorem approach to solving the DPE. That is, compute \( V_1 \) and \( V_2 \). If you stopped after only these two steps, what would be your estimate of the optimal feedback control?

4. (5) Give examples of Markov chains (with state space no smaller than 3 and no larger than 4) which are:

(a) Irreducible and aperiodic.

(b) Decomposable into two equivalence classes of communicating states.

5. (10) Consider the gambling problem discussed in class with the probability of winning being \( p = 0.45 \) (and the probability of losing being \( q = 1 - p = 0.55 \)). Suppose the gambler bets \$1 at each time, that the gambler starts with \( g_0 = \$2 \), and that the opponent starts with \( o_0 = \$4 \). If either person goes bankrupt, they stay bankrupt for all time; otherwise the bet remains at \$1 at each time. Estimate the probabilities of the player eventually going bankrupt and of the opponent eventually going bankrupt. What is the expected payoff to the gambler? Obtain the same information in the case where the gambler bets the maximum allowable (such that neither player can ever have a negative balance) at each step. Again estimate the probabilities of each player going bankrupt and the expected payoff to the gambler. Should the gambler bet \$1 each time, or the maximum allowable amount?

Now, vary \( o_0 \) from \$1 to \$50, and compare the expected payoffs to the gambler for each of the two control policy options.

6. (10) Consider a Markov chain control problem where \( \mathcal{X} = \{1, 2, 3\} \) and \( U = \{a, b, c\} \). Let the transition probability matrices be

\[
P(a) = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 3/4 & 1/4 \end{bmatrix}, \quad P(b) = \begin{bmatrix} 3/4 & 0 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/4 & 3/4 \end{bmatrix}
\]

\[
P(c) = \begin{bmatrix} 3/4 & 0 & 1/4 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}.
\]

Let the running cost be

\[
l(i, u) = \begin{cases} -2i & \text{if } u = a \\ 1 - 2i & \text{if } u = b, c. \end{cases}
\]
Let the discount factor be $\alpha = 0.8$. Use policy iteration starting with $\bar{u}(i) = c$ for all $i$ to solve the problem.

7. (5) Suppose $\mathcal{X}$ and $\mathcal{U}$ are finite, and that $f, g : \mathcal{X} \times \mathcal{U} \rightarrow \mathbb{R}$. Show that

$$\left| \min_{u \in \mathcal{U}} f(i, u) - \min_{v \in \mathcal{U}} g(i, v) \right| \leq \max_{u \in \mathcal{U}} | f(i, u) - g(i, u) |$$

for all $i \in \mathcal{X}$. 