MAE 288A
Assignment 3
Due Wednesday, 22 May

Problems to hand in (Not all problems may be graded.)

All work and explanations must be included for full credit.

1. (5) Consider the normed vector space \( (\mathcal{Y}, \| \cdot \|_2) \) given by \( \mathcal{Y} = L_2(\mathbb{R}) \), \( \| \phi \|_2 = \left[ \int_{\mathbb{R}} |\phi(x)|^2 \, dx \right]^{1/2} \). Consider the operator, \( \mathcal{G} : \mathcal{Y} \rightarrow \mathcal{Y} \) given by \( \mathcal{G}[\phi](x) = \exp(-x^2) + (2/3)\phi(1 + 4x) \) for all \( x \in \mathbb{R} \). Is \( \mathcal{G} \) a contraction? If so, prove it; if not, provide a counterexample. Answer the same question for operator \( \tilde{\mathcal{G}} : \mathcal{Y} \rightarrow \mathcal{Y} \) given by \( \tilde{\mathcal{G}}[\phi](x) = \exp(-x^2) + (2/3)\phi(1 + x/4) \) for all \( x \in \mathbb{R} \).

2. (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 \( U = \{1/2, 3/4\} \). (That is, \( U \) consists of only two possible values.) Let the \( w_t \) be IID normal random variables with mean 0 and variance 1. Let the cost criterion be

\[
J(x, \mu) = E_{w \mid 0,\infty} \left\{ \sum_{t=0}^{\infty} \alpha^t [\xi_t^2 + (\mu(\xi_t))^2] \right\}
\]

where \( \alpha = 2/3 \). Write down the DPE for this problem. Starting with initial guess, \( V_0 \equiv 0 \), perform two steps of the Value Iteration method (a.k.a. successive approximations or Banach Fixed Point Theorem approach) for solving this 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?

3. (5) Indicate an example of a Markov chain with \( \# \mathcal{X} = 4 \) (i.e., four states), and such that there are two equivalence classes where one has period two and the other has period one. Specifically provide both an annotated figure and the transition matrix.
4. (5) Consider the Markov chain with probability transition matrix

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

Obtain the steady-state distribution without propagating forward until steady-state is reached. Indicate why you are able to use the method that you do. Also, indicate a three-state Markov chain, where there may be more than one steady-state distribution, and provide two different steady-state distributions for it.

5. (10) Consider the gambling problem discussed in class with the probability of winning being \( p = 0.46 \) (and the probability of losing being \( q = 1 - p = 0.54 \)). 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 \( X = \{1, 2, 3\} \) and \( U = \{a, b, c\} \). Let the transition probability matrices be

\[
P(a) = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0.75 & 0 & 0.25 \\ 0 & 0.75 & 0.25 \end{bmatrix}, \quad P(b) = \begin{bmatrix} 0.75 & 0 & 0.25 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.25 & 0.75 \end{bmatrix}
\]

\[
P(c) = \begin{bmatrix} 0.75 & 0 & 0.25 \\ 0.33 & 0.33 & 0.34 \\ 0 & 0.33 & 0.67 \end{bmatrix}.
\]
Let the running cost be
\[
I(i, u) = \begin{cases} 
-2i & \text{if } u = a, \\
1 - 2i & \text{if } u = b, \\
4 - 4i & \text{if } u = c.
\end{cases}
\]

Let the discount factor be \(\alpha = 0.75\). Use policy iteration starting with \(\mu(i) = a\) for all \(i\) to solve the problem.

7. (10) Consider a Markov chain control problem with an exit cost criterion. Let the state space be \(X = \{1, 2, 3\}\), with exit state set, \(E = \{3\}\). Let the control set be \(U = \{a, b\}\). Let the transition probability matrices be
\[
P(a) = \begin{bmatrix} 0.7 & 0.3 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0 & 0 & 1 \end{bmatrix}, \quad P(b) = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.875 & 0.125 \\ 0 & 0 & 1 \end{bmatrix}
\]

Let the running cost be
\[
l(i, v) = 3 - i
\]
for \(i \in \{1, 2\}\) and \(v \in \{a, b\}\). Let the exit cost be \(\psi(3) = 5\). The controller wants to minimize the cost. Use value iteration starting with \(V^0 = (0, 0)\) (i.e., \(V^0(1) = 0, V^0(2) = 0\)) to solve the problem. What are the optimal feedback controls? What is the value function? Discuss the behavior of the value iteration in light of the assumptions applied in the development of the algorithm in the case of an exit problem.