MATH4406 (Control Theory), HW3 (Unit 3)
Finite Horizon MDP.

Prepared by Yoni Nazarathy, Last Updated: August 31, 2014

This homework project is mostly about dynamic programming of finite horizon MDP.

- Use Chapters 3 and 4 of [Put94]. You can get the chapter on-line from the library.
- For some of you the programming aspect may still be a bit of a challenge. Make sure to allow enough time for this. Break up each programming task into well-defined sub-tasks.
- Please make sure to present your results in a clear and organised manner. Numerical output results should always be well explained and documented. Labels on graphs, diagrams, tables etc...
- Hand-in all code (preferably as an appendix).

Problem 1: Inventory Control

- If needed reread Section 3.2; this was part of HW2, but perhaps this was a question you skipped.
- Read the numerical example of Section 4.6.1. That examples illustrates how to find the optimal policy using "backward induction " for N=3 and other data as in Section 3.2.2: $K = 4$, $c(u) = 2u$, $g(u) = 0$, $h(u) = u$, $M = 3$, $f(u) = 8u$ and $p_j = (1/4,\ 1/2,\ 1/4)$ for $j = 0, 1, 2$ respectively.

Now repeat this example for $N = 5$ with other problem data of your choice (e.g. modify $C(\cdot)$ to be $c(u) = 3u$ and take $M = 5$ etc...). The idea in this exercise is NOT to implement the algorithm in software, but rather to work through its steps by hand. Nevertheless, if you wish to verify your answer using software, feel free to do so.

Problem 2: Threshold policy in inventory control
Continuing from the previous problem, you will notice that on page 96, it is mentioned that (σ, Σ) policies are known to be optimal when the ordering cost is

$$O(u) = c\,u + \mathbf{1}\{u > 0\}\,K. \tag{1}$$

Implement (in software) the backward induction algorithm for solving an inventory control problem. Use your implementation to solve 10 different problems of your choice with a time horizon of $N = 15$, each satisfying (1). Verify that the resulting policy is indeed a threshold (σ, Σ) policy for each of the 10 problems that you chose. Please present the results in a compact manner.

Now do one of the following (your choice which one):

1. Find some ordering cost $O(\cdot)$, different from (1), that yields a non threshold policy. Demonstrate that the output of your backward induction algorithm yields a non-threshold policy.
2. Argue and demonstrate the benefits of knowing (in case (1) holds) that a threshold (σ, Σ) policy is optimal. Be as creative (or not) as you wish in this demonstration. E.g. illustrate how knowing this can yield faster algorithms for the optimal policy, or anything else. Do computation if needed.

Problem 3: Optimal Markov Deterministic Policies
Look at Theorem 4.4.2 on page 89 of together with part (a) of Proposition 4.4.3 on the next page. Package these into a single theorem that states that when $\mathcal{S}$ is finite or countable and $\mathcal{A}_s$ are all finite then there exists a deterministic Markovian policy that is optimal. Write out the proof. Be precise and neat.

Problem 4: The Secretary Problem
Read the analysis of "The Secretary Problem" in Section 4.6.4.

1. Reproduce Figure 4.6.2 (supply your code).
2. Fill in any missing details in the analysis on page 102, yielding the Ne^{-1} rule. This is a good "rule of thumb" for life. How would you use it elsewhere?
3. Think (or look up) generalisations or modifications of the Secretary Problem (there has been much research on this). Briefly (in one paragraph) present one such modified problem.