Homework for WISB372/ECRMMAT, week 42

Exercise 1 Consider the following discrete-time optimal control problem: minimize $\sum_{k=0}^{N-1}(u_k^2 + x_k)$ over all $(u_0, u_1, \ldots, u_{N-1}) \in \mathbb{R}^N$ such that $x_0 = 0 = x_N$, where $x_{k+1} = x_k + u_k$ is the dynamical system. Find all candidate-optimal solutions for this problem for general N. Note that this problem is not precisely of the type "mod 2b" as used in Bertsekas, p. 131. However, one can adapt Proposition 3.3.2 as follows: replace the transversality condition $p_n = \nabla g_N(x_N^*)$ by $x_N^* = 0$. *Hint 1:* Recall that $\sum_{i=1}^{m} i = \frac{1}{2}m(m+1)$. *Hint 2:* When solving difference equations it is enough to discover a general pattern – its validity need not be proven formally (say by induction).

Exercise 2 Consider the following discrete-time optimal control problem: *maximize* $\sum_{k=0}^{N-1} u_k^2 x_k$ over all $(u_0, u_1, \ldots, u_{N-1}) \in [-1, 1]^N$ such that $x_0 = 0$, where $x_{k+1} = x_k + 3u_k$ is the dynamical system.

a. Find, by simple heuristic principles, the optimal control function for this problem.

b. Show that the solution found un part a satisfies Proposition 3.3.2 on p. 131. *Hint:* note that it is (3.43) that must be verified and not (3.44), even though H_k is convex in u_k.

Exercise 3 On 19-10 in class the problem in Example 1.3.1 was also solved by means of Proposition 3.3.2. This gives the candidate-optimal control function

$$u_1^* = \frac{ar[T - (1-a)^2 x_0]}{a^2 r(1-a)^2 + a^2 r + 1} \text{ and } u_0^* = (1-a)u_1^*.$$

Show that this solution has the following connection with μ_1^* and μ_0^* obtained on pp. 26-27: $\mu_1^*(x_1^*) = u_1^*$ and $\mu^*(x_0) = u_0^*$. Here x_1^* is the state at time $k = 1$ that is obtained by applying the above candidate-optimal control function.

Exercise 4 Formulate the initial "cheapest route" example of 19-10 as a DP-problem in standard form and solve it by means of the DPA. *Hint:* Consider the first pages of chapter 2 in Bertsekas.

Exercise 5 You play N rounds of a certain betting game. Your starting capital is B. In each round you are allowed to stake any amount, but not more than what you own at that time (so you cannot borrow, etc.). In each round of betting the chance of winning is $p > 1/3$, in which case you receive three times the amount staked in that round (and, of course, if you lose the bet, you lose the the amount staked). Your aim is to maximize the expected value of the *logarithm* of the capital that is left after the N-th round.

a. Formulate this problem as a standard form DP-problem.

b. Use the DPA to solve this gambling problem. Show in particular that the optimal policy prescribes that a certain fraction of the capital should be staked in each round and that this fraction stays the same for the different rounds. *Check your answer:* the fraction is $\frac{1}{2}(3p - 1)$.

c. Verify the correctness of your expression for the fraction obtained in part b by considering the extreme case where $p = 1$.

Exercise 6 You can roll a standard dice at most six times. After each roll you must decide to stop or to go on (but, of course, going on is no option after the sixth roll). Your reward is the number reached at the last roll. Your goal is to maximize the expectation of that reward.

a. Formulate this problem as a standard form DP-problem.

b. Use the DPA to solve this gambling problem. Show that you should only stop after your first roll if the outcome of that roll is "six", but that this changes into "six or five" for stopping after your second roll.