Solutions Final ECRMMAT (open book), 9-11-2011

Problem 1 [35 pts.] Consider the following optimal control problem: minimize $\int_0^1 (u^2(t) - at^4 x(t))dt$ over all piecewise continuous functions $u : [0,1] \to U := \mathbb{R}$ such that $x(0) = 3$ and $x(1) \geq -1$. Here the dynamical system is $\dot{x} = u$ and a is a fixed parameter in $\mathbb{R}$.

a. Determine for the above optimal control problem the candidate-optimal control function(s) for general a in $\mathbb{R}$ by means of the MP. *Hints:* 1. This problem is similar to homework exercise 3 in week 39! 2. Section 21.5 in Dowling can be applied here but watch out for the differences in notation between Dowling and Bertsekas. 3. Make sure to *verify* that your candidate-optimal trajectory/trajectories indeed satisfy the inequality constraint.

b. The Sufficiency Theorem, as given in class (but see also section 21.3 in Dowling), needs to be modified so as to take care of the inequality constraint, as used in section 21.5 and the above problem (a similar modification, namely for the free endpoint problem, is sketched in Example 4 on p. 498 of Dowling). Sketch such a modification of the Sufficiency Theorem (note: no proof is required) and show that it can be applied to ensure the optimality of the function(s) found in part b.

c. As a final check, consider the special case $a = 0$. First state what the obvious optimal control function is for that problem (without using any MP). Next, check if your solution of part a, when specialized to $a = 0$, agrees with it.

Summary of the main part of the solution. As mentioned in the problem statement, the solution method is exactly the one used for homework exercise 3, certainly if one uses the sign trick. Of course, there is now also a parameter a; it can be processed routinely, for instance just as the price parameter in the worked-out example for week 39.

Solution. a. To conform completely to Dowling's maximization format, it is safest to apply the sign trick.[1] By this trick, the original problem is equivalent to maximizing $\int_0^1 at^4 x(t) - u^2(t))dt$ over all piecewise continuous functions $u : [0,1] \to U := \mathbb{R}$ such that $x(0) = 3$ and $x(1) \geq -1$. Therefore, the Hamiltonian is $H(t, x, u, p(t)) = at^4 x - u^2 + p(t)u$, so $H_x = at^4$ and $H_u = -2u + p(t)$. Observe that Dowling's conditions 1., 2. and 3. on p. 494 apply, because the "heroic interiority assumption" holds ($U = \mathbb{R}$).[2] First, the adjoint equation is $\dot{p} = -at^4$ and this gives $p(t) = -\frac{a}{5}t^5 + c_1$. Then Dowling's condition 1. gives $u^*(t) = p(t)/2 = -\frac{a}{10}t^5 + \frac{c_1}{2}$, so the dynamical system leads to $x^*(t) = -\frac{a}{60}t^6 + \frac{c_1}{2}t + c_2$, with $c_2 = 3$ by the initial condition $x(0) = 3$. Now the only possibilities are as follows.

Case 1: $x^*(1) > -1$. In this case the transversality condition $p(1) = 0$ holds (see section 21.5), so the above gives $0 = p(1) = -\frac{a}{5} + c_1$, i.e., $c_1 = a/5$. This causes

[1] Warning: if this is not done, then the inequality $p(T) \geq 0$, as used in Dowling's case 2, reverses sign: it turns into $p(T) \leq 0$! Symbolically, this reversal is understood as follows: $(-g) + p(-f)$ for the Hamiltonian, as section 21.5 wants it, corresponds to $g + (-p)f$.

[2] As in class, the incorrect statement preceding those three conditions should be ignored.

the candidate-optimal solution to be $u^*(t) = -\frac{a}{10}t^5 + \frac{a}{10}$, with associated trajectory $x^*(t) = -\frac{a}{60}t^6 + \frac{a}{10}t + 3$. However, in the present case 1 this candidate-optimal solution can only be accepted if $-1 < x^*(1) = -\frac{a}{60} + \frac{a}{10} + 3$, i.e., if $a > -48$.

Case 2: $x^*(1) = -1$. In this case we must only check that $p(1)$ is nonnegative. For the original expression for $x^*(\cdot)$ we now have $-1 = x^*(1) = -\frac{a}{60} + \frac{c_1}{2} + 3$, i.e., $c_1 = \frac{a}{30} - 8$ and then it must be that $0 \leq p(1) = -\frac{a}{5} + c_1 = -\frac{a}{6} - 8$, which is to say $\frac{a}{6} \leq -8$, i.e., $a \leq -48$. In this case the optimal solution is $u^*(t) = -\frac{a}{10}t^5 + \frac{a}{60} - 4$ by the original expression for $u^*(\cdot)$.

Summary: the optimal solution is $u^*(t) = -\frac{a}{10}t^5 + \frac{a}{10}$ if $a > -48$ and $u^*(t) = -\frac{a}{10}t^5 + \frac{a}{60} - 4$ if $a \leq -48$.

b. In view of a pattern seen during the first part of the course, the proper modification should be the addition of extra concavity: if for every $t \in [0, T]$ the Hamiltonian function $H(t, x, u, p(t))$ is concave in (x, u), then the necessary conditions – in this case these are the conditions stated in section 21.5 – are also sufficient.

All that remains to be done is to check the joint concavity in (x, u) of $H(t, x, u, p(t)) = at^4x - u^2 + p(t)u$. Note that $H_{xx} = H_{xu} = H_{ux} = 0$ and $H_{uu} = -2$. Hence, the Hessian matrix H_H is evidently negative semi-definite, so the desired concavity property follows.

c. For $a = 0$ the obvious optimal solution is $u^* \equiv 0$, which gives the lowest possible value for $J(u)$, namely 0, and has an associated trajectory $x^* \equiv 3$ that meets the initial and end time conditions. Because $0 > -48$, the formula found in part a gives $u^* \equiv 0$ as well.

Problem 2 [35 pts.] In the two-rounds chess tournament in Bertsekas with sudden death possibility (see Example 1.1.5 on p. 11), the objective is to maximize the player's probability of winning the tournament (e.g., see the lines following (2) in Example 1.1.5). This problem is completely solved in Example 1.3.3 (pp. 32-33).[3] Now consider precisely the same tournament, but with the following objective: to *maximize the player's expected net score at the end of the tournament.* Here "net score" is as defined in Example 1.3.3 (p. 32).

a. Formulate the associated maximization problem as a standard dynamical programming problem. *Hint:* be careful about the net end-score in the sudden death possibility.

b. Find the optimal policy for maximizing the expected net end-score if $p_d = 1/2$ and $p_w = 1/5$.

Summary of the solution: For $N = 2$ in Example 1.3.3 a positive net score x_2 (winning the tournament) was valued by 1 and a negative net score (losing) by 0; this was done so as to deal with maximizing $P(E)$ = expectation of the characteristic fuction 1_E = probability of winning the tournament (here E is the event of winning the tournament). The sudden death mode required a small adaptation. The only difference with the present problem is that now the net score x_2 itself must be counted, again with a similar adaptation for the sudden death mode.

Solution. a. The model of pp. 32-33 can be copied (of course for $N = 2$, as in Example 1.1.5), but now $g_N = g_2$, which is the expression on the right in formula (1.10) of Bertsekas, must be modified: it is $g_2(x_2) = x_2$ if $x_2 \neq 0$ and $2p_w - 1$ if $x_2 = 0$. Indeed, if $x_2 \neq 0$ then it is the net end-score x_2 which counts, and if there is

[3] Recall: "timid" or "bold" can be played in each round, with "timid" resulting in a draw [loss] with probability p_d $[1 - p_d]$ and "bold" in a win [loss] with probability p_w $[1 - p_w]$.

a draw at the end ($x_2 = 0$) the same reasoning about the sudden death extension can be followed as in Bertsekas: it makes only sense to play "bold" and this results in a net end-score of 1 (probability p_w) or of -1 (probability $1 - p_w$), causing the expected net score to be $2p_w - 1$. Otherwise, nothing changes, so the DPA-algorithm is as in (1.8) (with $N = 2$ and $J_2 =$ new g_2 above).

b. The new details give $g_2(x_2) = x_2$ if $x_2 \neq 0$, i.e., if $x_2 = 2, 1, -1$ or -2 and $g_2(0) = \frac{2}{5} - 1 = -\frac{3}{5}$. For $k = N - 1 = 1$ and $x_1 = -1, 0$ or 1, formula (1.8) implies

$$J_1(x_1) = \max[\frac{1}{2}J_2(x_1) + \frac{1}{2}J_1(x_1 - 1), \frac{1}{5}J_2(x_1 + 1) + \frac{4}{5}J_2(x_1 - 1)]$$

Concretely, for $x_1 = -1$ this gives

$$J_1(-1) = \max[\frac{1}{2} * (-1) + \frac{1}{2} * (-2), \frac{1}{5} * (-\frac{3}{5}) + \frac{4}{5} * (-2)] = -1.5 \text{ for } u_1^* = \text{"timid"}$$

(apparently this situation gives so little hope that the player simply concentrates on preventing the net end-score to be -2!). For $x_1 = 0$ and $x_1 = 1$ respectively the same formula gives

$$J_1(0) = \max[\frac{1}{2} * (-\frac{3}{5}) + \frac{1}{2} * (-1), \frac{1}{5} * 1 + \frac{4}{5} * (-1)] = -0.6 \text{ for } u_1^* = \text{"bold"}$$

$$J_1(1) = \max[\frac{1}{2} * 1 + \frac{1}{2} * (-\frac{3}{5}), \frac{1}{5} * 2 + \frac{4}{5} * (-\frac{3}{5})] = 0.2 \text{ for } u_1^* = \text{"timid"}$$

Finally, $x_0 = 0$ gives, still using (1.8) in Bertsekas,

$$J_0(0) = \max[\frac{1}{2} * J_1(0) + \frac{1}{2} * J_1(-1), \frac{1}{5} * J_1(1) + \frac{4}{5} * J_1(-1)] =$$

$$= \max[\frac{1}{2} * (-0.6) + \frac{1}{2} * (-1.5), \frac{1}{5} * 0.2 + \frac{4}{5} * (-1.5)] = -1.05 \text{ for } u_0^* = \text{"timid"}$$

Problem 3 [35 pts.] Consider the following optimal control problem, which is a modification of the resource allocation problem studied and solved in Examples 3.1.2 (p. 107) and 3.3.2 (pp. 121-122) in Bertsekas. Maximize $x(T) + \int_0^T (1 - u(t))x(t)dt$ subject to $u(t)$, the portion of the production rate used for reinvestment, being in $U := [0, 1]$ for all $t \in [0, T]$ and $x(0) = x_0$. Here the dynamical system is: $\dot{x}(t) = \gamma u(t)x(t)$ and $\gamma > 0$ and $x_0 > 0$ are given parameters.

a. Let $u(\cdot) : [0, T] \to [0, 1]$ be any control function. Using (16.1) in Dowling, demonstrate that the associated trajectory $x(\cdot)$ in the above problem is such that $x(t) \geq x_0 > 0$ for all $t \in [0, T]$.

b. Imitate Example 3.3.2, including the use of illustrative figures such as those on p. 122, as much as possible to obtain the candidate-optimal control function(s) for the above optimal control problem. Among other things, show that the co-state function $p(t)$ has $\dot{p}(t) < 0$ for all t.

c. Similar to part b, obtain the the candidate-optimal control function(s) for the following much simpler problem: maximize $x(T)$ subject to precisely the same conditions as before (i.e., $U = [0, 1]$, $\dot{x} = \gamma ux$ and $x(0) = x_0$).

d. Give a brief *economic* interpretation of the candidate-optimal solution(s) found in part c; this should certainly include a short sketch of what the optimal control problem in part c means economically, when compared to Examples 3.1.2 and 3.3.2 in Bertsekas.

Summary of the main part of the solution. Compared to p. 121, only the transversality condition changes: it is now $p(T) = 1$. One can preserve the reasoning in Bertsekas (see figures 3.3.1-3.3.3) if $1/\gamma > 1 = p(T)$ and if $1/\gamma \leq 1 = p(T)$ the reasoning is trivial, leading to $u^* \equiv 1$.

Solution. a. In formula (16.1) of Dowling we substitute $z = 0$ and $v(t) := -\gamma u(t)$. Then (16.1) gives $x(t) = e^{-V(t)}(A+0)$, where $V(t)$ is a fixed primitive function for the function $v(t) = -\gamma u(t) \leq 0$. For $V(t)$ we may choose $V(t) := \int_0^t -\gamma u(t')dt' \leq 0$ by what is said in Dowling (see Example 3 on p. 363). Now $-V(t) \geq 0$ for all $t \in [0, T]$ implies $x(t) = Ae^{-V(t)} \geq Ae^0 = A = x(0) = x_0$.

b. The Hamiltonian is the same as on p. 121, namely $H(x, u, p(t)) = (1 - u)x + p(t)\gamma ux$. The adjoint equation is also the same: $\dot{p}(t) = -\gamma u^*(t)p(t) - 1 + u^*(t)$, but the transversality condition is different: this time it is $p(T) = 1$. The maximum principle is again the same: $u^*(t) = 0$ if $p(t) < 1/\gamma$ and $u^*(t) = 1$ if $p(t) > 1/\gamma$. Substitution in the adjoint equation of the latter two expressions for $u^*(t)$ yields $\dot{p}(t) = -1 < 0$ if $p(t) < 1/\gamma$ and $\dot{p}(t) = -\gamma p(t) < 0$ if $p(t) > 1/\gamma$. Moreover, for $p(t) = 1/\gamma$ the adjoint equation gives $\dot{p}(t) = -u^*(t) - 1 + u^*(t) = -1 < 0$. So $p(t)$ is strictly decreasing in t.

Case 1: $\gamma < 1$. In this case $p(T) = 1 < 1/\gamma$, so the reasoning in Bertsekas can be repeated: close to T we will have $p(t) < 1/\gamma$ and this causes $u^*(t) = 0$ by the maximum principle and in turn $\dot{p} = -1$, i.e., $p(t) = -t + c_1$. Then $1 = p(T) = -T + c_1$ implies $p(t) = 1 + T - t > 0$ for those t near T. A switch will occur if $p(t) = 1/\gamma$, which corresponds to the switch time $t_s = T + 1 - \frac{1}{\gamma} < T$. If $t_s < 0$, i.e., if $T + 1 \leq 1/\gamma$, then $u^* \equiv 0$ and no switch will actually occur. On the other hand, if $T + 1 > 1/\gamma$, then $t_s \in (0, T)$ and a switch occurs at t_s.

Case 2: $\gamma \geq 1$. By the above, $p(t)$ decreases strictly to $p(T) = 1$. So for every $t < T$ we have $p(t) > p(T) = 1 \geq 1/\gamma$. By the maximum principle, this implies that $u^*(t) = 1$ for all $t \leq T$ if $\gamma > 1$ and if $\gamma = 1$ then it certainly implies that $u^*(t) = 1$ for al $t < T$, while, as argued in class, the value of $u^*(\cdot)$ in the single point T can be set equal to 1 as well. So the candidate-optimal control function in this case is $u^* \equiv 1$.

c. This time, we have $H(t, x, u, p(t)) = p(t)\gamma xu$, so $H_x = p(t)u$. Hence, the adjoint equation is $\dot{p} = -\gamma p(t)u(t)$. By (16.1) again, its solution is $p(t) = Ae^{-W(t)}$, where $W(t) := \int_0^t \gamma u(t')dt'$, similar to what was done in part a (the reversal of the sign comes from the extra minus sign in the adjoint equation). The maximum principle gives that $p(t) > 0$ implies $u^*(t) = 1$ and $p(t) < 0$ implies $u^*(t) = 0$. Transversality gives $p(T) = 1$, so above it follows that $A > 0$. Because $\gamma u(\cdot)$ is nonnegative, it follows that $W(t)$ is nondecreasing in t, so $p(t)$ is nonincreasing in t. This gives $p(t) \geq p(T) = 1$ for all t. Therefore, the maximum principle gives $u^* \equiv 1$, with its associated trajectory $x^*(t) = t + x_0$.

d. To maximize $x(T)$ means that one only attaches value to the capital at the end time, and that no value at all is attached to the amount stored. So clearly the obvious thing to do is not to contribute to storage at all and to concentrate entirely on raising the end capital. This means choosing $u^* \equiv 1$.