



Prof. L. Guzzella Prof. R. D'Andrea

Final Examination

January 28th, 2009

Dynamic Programming & Optimal Control (151-0563-01) Prof. R. D'Andrea

Solutions

Exam Duration:	150 minutes
Number of Problems:	4 (25% each)
Permitted Aids:	Textbook Dynamic Programming and Optimal Control by Dimitri P. Bertsekas, Vol. I, 3rd edition, 2005, 558 pages. Your written notes and class handouts.
	No calculators.
Important:	Use only these prepared sheets for your solutions.

General Hints:	$E(\cdot)$ denotes the expected value.
	E(rs) = E(r)E(s), if two random variables r, s are independent.



Figure 1

Find the shortest path from node S to node T for the graph given in Figure 1. Apply the Label Correcting Method. Use Depth-First Search to determine at each iteration which node to remove from OPEN; that is, the node is always removed from the top of OPEN and each node entering OPEN is placed at the top of OPEN (last-in/first-out policy).

Solve the problem by populating a table of the following form:

Iter-	Node exiting	OPEN	d_S	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_T
ation	OPEN												
0	-												
1	S												
2	4												

where the variable d_i denotes the length of the shortest path from node S to node i that has been found so far. Note that the first two nodes exiting OPEN are given.

State the resulting shortest path and its length.

25%

Iter- ation	Node exiting OPEN	OPEN	d_S	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_T
0	-	S	0	∞	∞								
1	S	4,3,1	0	2	∞	1	3	∞	∞	∞	∞	∞	∞
2	4	$5,\!3,\!1$	0	2	∞	1	3	4	∞	∞	∞	∞	∞
3	5	6,3,1	0	2	∞	1	3	4	6	∞	∞	∞	∞
4	6	8,3,1	0	2	∞	1	3	4	6	∞	7	∞	∞
5	8	3,1	0	2	∞	1	3	4	6	∞	7	∞	8
6	3	4,1	0	2	∞	1	2	4	6	∞	7	∞	8
7	4	5,1	0	2	∞	1	2	3	6	∞	7	∞	8
8	5	6,1	0	2	∞	1	2	3	5	∞	7	∞	8
9	6	8,1	0	2	∞	1	2	3	5	∞	6	∞	8
10	8	1	0	2	∞	1	2	3	5	∞	6	∞	7
11	1	-	0	2	∞	1	2	3	5	∞	6	$ \infty$	7

Alternative solution:

Iter- ation	Node exiting OPEN	OPEN	d_S	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_T
0	_	S	0	∞									
1	S	4,1,3	0	2	∞	1	3	∞	∞	∞	∞	∞	∞
2	4	5,1,3	0	2	∞	1	3	4	∞	∞	∞	∞	∞
3	5	6,1,3	0	2	∞	1	3	4	6	∞	∞	∞	∞
4	6	8,1,3	0	2	∞	1	3	4	6	∞	7	∞	∞
5	8	1,3	0	2	∞	1	3	4	6	∞	7	∞	8
6	1	2,3	0	2	7	1	3	4	6	∞	7	∞	8
7	2	3	0	2	7	1	3	4	6	∞	7	∞	8
8	3	4	0	2	7	1	2	4	6	∞	7	∞	8
9	4	5	0	2	7	1	2	3	6	∞	7	∞	8
10	5	6	0	2	7	1	2	3	5	∞	7	∞	8
11	6	8	0	2	7	1	2	3	5	∞	6	∞	8
12	8	-	0	2	7	1	2	3	5	∞	6	∞	7

The shortest path is $S \to 3 \to 4 \to 5 \to 6 \to 8 \to T$ with a total length of 7.

Consider the following system

$$x_{k+1} = f_{i_k}(x_k, u_k, w_k), \quad i_k = 1, 2,$$

where

$$f_1(x_k, u_k, w_k) = w_k x_k + w_k u_k$$
$$f_2(x_k, u_k, w_k) = w_k x_k - 2w_k u_k$$

The disturbance w_k takes the values 0 and 1 with equal probability. The input u_k is restricted to be 1 or -1.

State constraints are given as follows:

$$-2k \le x_k \le 2k.$$

Starting from an initial state $x_0 = 0$, the goal is to minimize the cost

$$E_{w_0,w_1}\left\{x_0^2 + x_1^2 + x_2^2\right\}.$$

The control inputs are u_k and i_k .

Apply the Dynamic Programming algorithm to find the optimal control policy and the optimal cost $J_0(0)$.

25%

The optimal control problem is considered over a time horizon N = 2 and the cost, to be minimized, is defined by

$$g_2(x_2) = x_2^2$$
 and $g_k(x_k, u_k, w_k) = x_k^2$, $k = 0, 1$.

Note that the state x_k takes on only integer values since $w_k \in \{0, 1\}$, $u_k \in \{-1, 1\}$, and $x_0 = 0$.

The DP algorithm proceeds as follows:

2nd stage:

The recursion is started with

$$J_2(x_2) = x_2^2$$

for all feasible $x_2 \in \{-4, -3, \dots, 2, 3, 4\}$.

1st stage:

Proceeding backwards, we get:

$$J_{1}(x_{1}) = \min_{\substack{i_{1} \in \{1,2\}\\u_{1} \in \{-1,1\}}} E\left\{x_{1}^{2} + J_{2}\left(f_{i_{1}}\left(x_{1},u_{1},w_{1}\right)\right)\right\}$$
$$= \min_{\substack{i_{1} \in \{1,2\}\\u_{1} \in \{-1,1\}}} E\left\{x_{1}^{2} + \left(f_{i_{1}}\left(x_{1},u_{1},w_{1}\right)\right)^{2}\right\}.$$

The functions $f_1(x_k, u_k, w_k)$ and $f_2(x_k, u_k, w_k)$ are rewritten as

$$f_1(x_k, u_k, w_k) = w_k (x_k + u_k) = w_k h_1(x_k, u_k)$$

$$f_2(x_k, u_k, w_k) = w_k (x_k - 2u_k) = w_k h_2(x_k, u_k).$$

Using these definitions, we get

$$J_{1}(x_{1}) = \min_{\substack{i_{1} \in \{1,2\}\\u_{1} \in \{-1,1\}}} E_{w_{1}} \left\{ x_{1}^{2} + w_{1}^{2} \left(h_{i_{1}} \left(x_{1}, u_{1} \right) \right)^{2} \right\}$$
$$= \min_{\substack{i_{1} \in \{1,2\}\\u_{1} \in \{-1,1\}}} \left\{ x_{1}^{2} + 0.5 \cdot 0^{2} + 0.5 \cdot 1^{2} \cdot \left(h_{i_{1}} \left(x_{1}, u_{1} \right) \right)^{2} \right\}$$
$$= x_{1}^{2} + 0.5 \cdot \min_{\substack{i_{1} \in \{1,2\}\\u_{1} \in \{-1,1\}}} \left\{ \left(h_{i_{1}} \left(x_{1}, u_{1} \right) \right)^{2} \right\}.$$

Now, by evaluating $h_{i_1}(x_1, u_1)$ for all feasible $x_1 \in \{-2, -1, 0, 1, 2\}$ considering possible input pairs,

$$(u_1, i_1) \in \left\{ (1, -1), (1, 1), (2, -1), (2, 1) \right\},\$$

the state-dependent minimizing input and the corresponding cost-to-go $J_1(x_1)$ are found:

$$\begin{split} J_1(-2) &= 4 + 0 = 4, & \text{with} & \mu^*(-2) = (2, -1) \\ J_1(-1) &= 1 + 0 = 1, & \text{with} & \mu_1^*(-1) = (1, 1) \\ J_1(0) &= 0 + 0.5 = 0.5, & \text{with} & \mu_1^*(0) = (1, \pm 1) \\ J_1(1) &= 1 + 0 = 4, & \text{with} & \mu_1^*(1) = (1, -1) \\ J_1(2) &= 4 + 0 = 4, & \text{with} & \mu_1^*(2) = (2, 1). \end{split}$$

0th stage:

Finally with the initial condition $x_0 = 0$, the optimal cost is calculated by

$$J_{0}(0) = \min_{\substack{i_{0} \in \{1,2\}\\u_{0} \in \{-1,1\}}} E_{w_{0}} \left\{ 0^{2} + J_{1} \left(f_{i_{0}} \left(0, u_{0}, w_{0} \right) \right) \right\}$$

$$= \min_{\substack{i_{0} \in \{1,2\}\\u_{0} \in \{-1,1\}}} \left\{ 0.5 \cdot J_{1} \left(f_{i_{0}} \left(0, u_{0}, 0 \right) \right) + 0.5 \cdot J_{1} \left(f_{i_{0}} \left(0, u_{0}, 1 \right) \right) \right\}$$

$$= \min_{\substack{i_{0} \in \{1,2\}\\u_{0} \in \{-1,1\}}} \left\{ 0.5 \cdot J_{1} \left(0 \right) + 0.5 \cdot J_{1} \left(f_{i_{0}} \left(0, u_{0}, 1 \right) \right) \right\}.$$

The input,

$$\mu_0^*(0) = (1, \pm 1) \,,$$

minimizes the cost function and results in

$$J_0(0) = 0.5 \cdot J_1(0) + 0.5 \cdot J_1(\pm 1)$$

= 0.25 + 0.5 = 0.75.

Consider the following discrete-time system:

$$x_{k+1} = x_k + u_k + w_k, \qquad k = 0, 1, \dots, \infty$$
$$y_k = x_k,$$

where x_k , u_k , w_k , and y_k are real numbers. The initial condition is $x_0 = 1$. The w_k are independent random numbers with $E(w_k) = 0$ and $E(w_k^2) = 1$. The cost function is the following:

$$J = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} E\left(\frac{x_k^2}{2} + u_k^2\right).$$

We restrict ourselves to the following class of feedback laws:

$$u_k = F y_k \,,$$

where F is a constant gain.

- a) Find the gain F that minimizes the cost J.
- b) Now consider the measurement corrupted by noise:

$$y_k = x_k + v_k \,,$$

where the v_k are independent random numbers with $E(v_k) = 0$ and $E(v_k^2) = 1$. What is the cost J if the **same** feedback gain that you found in part a) is used?

a) This is an infinite horizon, perfect information problem for a linear system with a quadratic cost function. The optimal feedback strategy is a constant gain, which coincides with the class of feedback laws that we are considering in this problem. We therefore just have to find the optimal LQR gain for

$$A = 1$$
, $B = 1$, $Q = \frac{1}{2}$, and $R = 1$.

Solve the Riccati equation:

$$K = K - \frac{K^2}{K+1} + \frac{1}{2} \quad \Leftrightarrow \quad \frac{K^2}{K+1} = \frac{1}{2}$$

The positive solution therefore is K = 1. The optimal feedback gain is thus

$$F = \frac{-K}{K+1} = -\frac{1}{2}.$$

b) Consider the closed loop with $u_k = -\frac{1}{2}y_k = -\frac{1}{2}x_k - \frac{1}{2}v_k$:

$$x_{k+1} = x_k - \frac{1}{2}x_k - \frac{1}{2}v_k + w_k = \frac{1}{2}x_k - \frac{1}{2}v_k + w_k.$$
 (1)

The cost becomes

$$J = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(E\left(\frac{x_k^2}{2} + u_k^2\right) \right)$$

= $\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{1}{2} E(x_k^2) + E\left(\frac{1}{4} x_k^2 + \frac{1}{2} x_k v_k + \frac{1}{4} v_k^2\right) \right)$
= $\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{1}{2} E(x_k^2) + \frac{1}{4} E(x_k^2) + \frac{1}{4} E(v_k^2) \right)$
= $\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{3}{4} E(x_k^2) + \frac{1}{4} \right),$

where we used $E(x_k v_k) = E(x_k)E(v_k) = 0$, which holds because v_k is an independent random variable and $E(v_k) = 0$.

Now, using (1) and by independence of v_k and w_k , we have

$$E(x_{k+1}^2) = E\left(\frac{1}{4}x_k^2 + \frac{1}{4}v_k^2 + w_k^2 - \frac{1}{2}x_kv_k + x_kw_k - v_kw_k\right)$$

= $\frac{1}{4}E(x_k^2) + \frac{1}{4}E(v_k^2) + E(w_k^2)$
= $\frac{1}{4}E(x_k^2) + \frac{1}{4} + 1$
= $\frac{1}{4}E(x_k^2) + \frac{5}{4}.$ (2)

Let $\alpha = \lim_{k \to \infty} E(x_k^2)$, which exists since the recursion (1) is stable. Substituting this in (2), we can solve for α ,

$$\alpha = \frac{1}{4}\alpha + \frac{5}{4} \quad \Leftrightarrow \quad \alpha = \frac{5}{3}.$$

Therefore, the cost is

$$J = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{3}{4} \cdot \frac{5}{3} + \frac{1}{4} \right)$$
$$= \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \left(\frac{3}{2} \right)$$
$$= \frac{3}{2}.$$

Consider the following nonlinear differential equations

$$\dot{x}_1(t) = x_2(t)u_1(t)$$

 $\dot{x}_2(t) = x_1(t)u_2(t)$

with initial conditions

$$x_1(0) = -1$$

 $x_2(0) = 0$,

where all variables are real numbers.

The control inputs have the following constraints:

$$|u_1(t)| \le 1$$
 and $|u_2(t)| \le 1$.

- a) Find control inputs $u_1(t)$, $u_2(t)$ which drive the system to $(x_1, x_2) = (1, 0)$ as quickly as possible.
- **b)** Does it take longer or shorter to drive the system to $(x_1, x_2) = (0, 0)$?

a) The goal is to find an input trajectory $(u_1(t), u_2(t)), t \in [0, T]$ which accomplishes the transfer from

$$(x_1(0), x_2(0)) = (-1, 0)$$
 to $(x_1(T), x_2(T)) = (1, 0)$

in *minimum time*. Thus, we want to

minimize
$$T = \int_0^T 1 \, dt$$
.

Using the standard notation of the cost functional,

$$h(x(T)) + \int_0^T g(x(t), u(t)) dt \,,$$

that is

$$g(x) = 1$$
 and $h(x(T)) = 0$.

Apply the Minimum Principle.

• The Hamiltonian is given by

$$H(x, u, p) = g(x, u) + p^T f(x, u)$$

= 1 + p_1 x_2 u_1 + p_2 x_1 u_2.

• The adjoint equations follow from the equation above:

$$\dot{p}_1(t) = -\frac{\partial H}{\partial x_1} \left(x^*(t), u^*(t), p(t) \right) = -p_2(t) u_2^*(t)$$
$$\dot{p}_2(t) = -\frac{\partial H}{\partial x_2} \left(x^*(t), u^*(t), p(t) \right) = -p_1(t) u_1^*(t) \,.$$

Since we have a Fixed Terminal State Problem, there are no boundary conditions on p.

• In addition, for optimal time problems, the Hamiltonian satisfies

$$H(x^{*}(t), u^{*}(t), p(t)) = 1 + p_{1}(t)x_{2}^{*}(t)u_{1}^{*}(t) + p_{2}(t)x_{1}^{*}(t)u_{2}^{*}(t) = 0$$
(3)

for all $t \in [0, T]$.

• The optimal input $u^*(t)$ is obtained by minimizing the Hamiltonian along the optimal trajectory

$$u^{*}(t) = \underset{\substack{|u_{1}| \leq 1 \\ |u_{2}| \leq 1}}{\arg\min} \left\{ H\left(x^{*}(t), u, p(t)\right) \right\} = \underset{\substack{|u_{1}| \leq 1 \\ |u_{2}| \leq 1}}{\arg\min} \left\{ 1 + p_{1}x_{2}^{*}u_{1} + p_{2}x_{1}^{*}u_{2} \right\}$$

Since the Hamiltonian is linear in u_1 and u_2 , it follows that the optimal strategy is bang-bang; that is,

$$u_1^*(t) = \begin{cases} -1, & \text{if } p_1 x_2^* \ge 0\\ 1, & \text{if } p_1 x_2^* < 0 \end{cases}$$
(4)

and

$$u_2^*(t) = \begin{cases} -1, & \text{if } p_2 x_1^* \ge 0\\ 1, & \text{if } p_2 x_1^* < 0, \end{cases}$$
(5)

assuming that there are no intervals $\mathcal{T} = [t_1, t_2], \ 0 \le t_1 < t_2 \le T$, where

$$p_1(t)x_2^*(t) = 0 \quad \text{or} \quad p_2(t)x_1^*(t) = 0, \qquad \forall \ t \in \mathcal{T}.$$
 (6)

Prove bang-bang strategy.

In order to have a bang-bang time-optimal solution, we have to show that assumption (6) holds.

First, assume $x_1^*(t) \equiv 0 \quad \forall t \in \mathcal{T}$. Then,

$$\dot{x}_2^*(t) = x_1^*(t)u_2^*(t) \equiv 0 \quad \forall t \in \mathcal{T}.$$

It follows that

$$x_2^*(t) \equiv const \quad \forall t \in \mathcal{T}.$$

That is, we do not move anywhere! Since the differential equations for x_1 and x_2 are symmetric, we can proceed in the same way for $x_2^*(t) \equiv 0 \quad \forall t \in \mathcal{T}$. So, the case, where x_1 or x_2 are 0 for any interval \mathcal{T} , is not interesting, since nothing happens. We do not move. Now, assume $p_1(t) \equiv 0 \quad \forall t \in \mathcal{T}$. Then,

$$\dot{p}_1(t) = -p_2(t)u_2^*(t) \stackrel{!}{=} 0 \quad \forall t \in \mathcal{T}.$$

In order to satisfy the equation above, we have two options:

- **a)** $p_2(t) \equiv 0 \quad \forall t \in \mathcal{T};$ however, this is a contradiction to (3).
- **b)** $u_2^*(t) \equiv 0 \quad \forall t \in \mathcal{T}; \text{ but then, } (3) \text{ does not hold either.}$

Because of symmetry, we can do the same for $p_2(t) \equiv 0 \quad \forall t \in \mathcal{T}$.

To sum up, above's considerations show that the optimal input $u^*(t)$ has to be bang-bang.

Consider zero-switch solution.

We first consider the case, where $u_1^*(t) = const$ and $u_2^*(t) = const$ over the whole time horizon $0 \le t \le T$.

Distinguish two different cases:

I) $\mathbf{u_1^*} = \mathbf{u_2^*} = \mathbf{u} = \pm \mathbf{1}$: In this case,

$$\dot{x}_1^* = x_2^* u$$

 $\dot{x}_2^* = x_1^* u$.

That is,

$$\ddot{x}_1^* = u^2 x_1^* = x_1^*$$
 and (7)

$$x_2^* = \frac{\dot{x}_1^*}{u} \,. \tag{8}$$

It follows that,

$$x_1^*(t) = A \cosh t + B \sinh t$$
$$x_2^*(t) = \frac{A}{u} \sinh t + \frac{B}{u} \cosh t,$$

where A, B are constant and cosh and sinh denote the hyperbolic cosines and sines:

$$\cosh t := \frac{e^t + e^{-t}}{2}$$
 and $\cosh t := \frac{e^t - e^{-t}}{2}$

The constants A and B are chosen, such that the initial conditions are satisfied:

$$x_1^*(0) = A = -1$$

 $x_2^*(0) = \frac{B}{u} = 0.$

Finally, we get

$$x_1^*(t) = -\cosh t$$
 and $x_2^*(t) = -\frac{1}{u}\sinh t$.

However, this solution does not solve our problem. It never reaches our final destination $(x_1, x_2) = (1, 0)$; that is, $x_1^*(t) \neq 1$ and $x_2^*(t) \neq 0$ for any t > 0.

 $II) \quad u_1=-u_2=u=\pm 1:$

In this case,

$$\dot{x}_1^* = x_2^* u$$

 $\dot{x}_2^* = -x_1^* u$.

That is,

$$\ddot{x}_1^* = -u^2 x_1^* = -x_1^*$$
 and
 $x_2^* = \frac{\dot{x}_1^*}{u}$.

It follows that

$$x_1^*(t) = A\cos t + B\sin t$$
$$x_2^*(t) = -\frac{A}{u}\sin t + \frac{B}{u}\cos t$$

with A, B being constant. From the initial conditions, we determine

$$x_1^*(0) = A = -1$$

 $x_2^*(0) = \frac{B}{u} = 0.$

Finally, we get

$$x_1^*(t) = -\cos t$$
 and $x_2^*(t) = \frac{1}{u}\sin t$. (9)

Note that for $T = \pi$, the final state is reached, $x_1^*(T) = 1$ and $x_2^*(T) = 0$. What still has to be checked, is that (9) and the corresponding costates $p_1(t)$ and $p_2(t)$ satisfy the necessary conditions of the Minimum Principle, namely Equations (5), (3), and (4). In order to verify these conditions, we have to calculate the costates:

$$\dot{p}_1 = p_2 u$$
$$\dot{p}_2 = -p_1 u$$

That is,

$$\ddot{p}_1 = -u^2 p_1 = -p_1$$
$$\dot{p}_2 = \frac{\dot{p}_1}{u} \,.$$

It follows that

$$p_1(t) = C\cos t + D\sin t \tag{10}$$

$$p_2(t) = -\frac{C}{u}\sin t + \frac{D}{u}\cos t.$$
(11)

With these results, Equations (10) and (11), and the state trajectory (9), Equation (3) reads as

$$H(x^{*}(t), u^{*}(t), p(t)) = 1 + p_{1}(t)x_{2}^{*}(t)\mathbf{i}?^{*}\frac{1}{2}, u - p_{2}(t)x_{1}^{*}(t)u$$

= 1 + C cos t sin t + D sin² t - C cos t sin t + D cos² t
= 1 + D
$$\stackrel{!}{=} 0.$$

That is D = -1. Moreover, in order to satisfy Equations (4) and (5), the constant C has to be chosen: C = 0.

To conclude, we found a trajectory (9) transferring the state from

$$(x_1(0), x_2(0)) = (-1, 0)$$
 to $(x_1(T), x_2(T)) = (1, 0)$, in time $T = \pi$,

and satisfying the necessary conditions of the Minimum Principle. The trajectory describes a circle which, depending on the choice of $u \in \{1, -1\}$, lies above or under the x_1 -axis.

Important to note is that the Minimum Principle is just a necessary condition. There might be a multiple-switch solution which is faster; i.e., with $T < \pi$. Proving that there is actually no other trajectory satisfying the necessary conditions is quite cumbersome and not required here. Nevertheless, intuition tells us that for a fast transition from the initial to the final state, $\dot{x}_1(t) = x_2(t)u_1(t)$ should be chosen large. Therefore, for an optimal solution, $|x_2(t)|$ should increase quickly during the first half of the trajectory and decrease at the end. This behavior can be observed when carefully looking at (9).

 \Rightarrow Grading: Required for full grade was the derivation of the zero switch solution II), which results in a circular trajectory from the initial state to the final destination, and the verification of the necessary conditions for this trajectory.

b) Longer. In fact, it takes an infinite amount of time to drive the system to (0,0). Why? Assume that we can drive the system to (0,0) in a finite time $T, T < \infty$. Define

$$z_1(t) = x_1(T-t),$$
 $v_1(t) = u_1(T-t),$
 $z_2(t) = x_2(T-t),$ $v_2(t) = u_2(T-t).$

Then,

$$\dot{z}_1(t) = -\dot{x}_1(T-t) = -x_2(T-t) u_1(T-t) = -z_2(t) v_1(t)$$

$$\dot{z}_2(t) = -\dot{x}_2(T-t) = -x_1(T-t) u_2(T-t) = -z_1(t) v_2(t)$$

with initial conditions

$$z_1(0) = 0$$

 $z_2(0) = 0$.

The inputs u_1 and u_2 are bounded and, therefore, v_1 and v_2 are bounded. From the initial conditions, it follows

$$\dot{z}_1(t) = \dot{z}_2(t) = 0 \quad \forall t > 0,$$

which is a contradiction to $z_1(T) = -1$, our initial condition in the original problem definition.