ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control
IDSC
Institut für Dynamische Systeme und Regelungstechnik

August 30th, 2021

Dynamic Programming & Optimal Control (151-0563-01)

Prof. R. D'Andrea

Exam

Exam Duration:	**150 minutes**
Number of Problems:	**4**
Permitted aids:	One A4 sheet of paper. No calculators allowed.

Problem 1 **[25 points]**

Consider the system dynamics

$$x_{k+1} = 2x_k + u_k + w_k, \quad k = 0, ..., N-1, \quad N = 2,$$

with the initial state $x_0 \in \mathcal{R}$. At each stage k, u_k is the control input, x_k is the state of the system and the disturbance w_k can assume the values -2, 0, 2 with probabilities $P(w_k = -2) = 1/4$, $P(w_k = 0) = 1/2$, $P(w_k = 2) = 1/4$, respectively. We want to minimize the following cost function

$$E_{w_0, w_1}\{x_2^2 + x_1^2 + u_1^2 + u_0^2\}$$

using the Dynamic Programming Algorithm.

a) What is the terminal cost? *[1 point]*

b) What are $g_2(x_2)$, $g_1(x_1, u_1, w_1)$ and $g_0(x_0, u_0, w_0)$? *[1 point]*

c) Find the optimal policy $u_1^* = \mu_1(x_1)$ and the corresponding optimal cost-to-go $J_1(x_1)$. *[8 points]*

d) Let's now assume that at stage 0 the input u_0 is constrained between -1 and 0, i.e. $-1 \leq u_0 \leq 0$. Find the optimal policy $u_0^* = \mu_0(x_0)$ and the corresponding optimal cost $J_0(x_0)$. *[10 points]*

e) Now assume $w_k = 0$ for each k. In this case, what are the optimal policies and the corresponding costs of the questions c) and d) ? *[5 points]*

SOLUTIONS

Solution 1

a) The terminal cost is $J_2(x_2) = x_2^2$

b) The optimal control problem is considered over a time horizon N = 2 and the cost to be minimized at each stage are $g_2(x_2) = x_2^2$, $g_1(x_1, u_1, w_1) = x_1^2 + u_1^2$, $g_0(x_0, u_0, w_0) = u_0^2$.

c) The dynamic programming algorithm is applied as follows:
STAGE 2:

$$J_2(x_2) = g_2(x_2) = x_2^2$$

STAGE 1:

$$\begin{aligned} J_1(x_1) = &\min_{u_1}[E_{w_1}\{g_1(x_1, u_1, w_1) + J_2(x_2)\}] = && (1)\\ &\min_{u_1}[E_{w_1}\{x_1^2 + u_1^2 + (2x_1 + u_1 + w_1)^2\}] = && (2)\\ &\min_{u_1}[E_{w_1}\{x_1^2 + u_1^2 + (2x_1 + u_1)^2 + w_1^2 + 2(2x_1 + u_1)w_1\}] = && (3)\\ &\min_{u_1}[x_1^2 + u_1^2 + (2x_1 + u_1)^2 + E_{w_1}\{w_1^2\} + 2(2x_1 + u_1)E_{w_1}\{w_1\}] && (4)\end{aligned}$$

Here, using $E_{w_1}\{w_1\} = -2 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 0$ and $E_{w_1}\{w_1^2\} = (-2)^2\frac{1}{4} + (0)^2\frac{1}{2} + (2)^2\frac{1}{4} = 2$, we obtain:

$$J_1(x_1) = \min_{u_1}[x_1^2 + u_1^2 + (2x_1 + u_1)^2 + 2] = \min_{u_1} C(u_1, x_1)$$

We now need to compute the value u_1 for which the gradient with respect to u_1 is zero:

$$\frac{\partial C}{\partial u_1} = 4u_1 + 4x_1 = 0,$$

and then $u_1^* = -x_1$.
The optimal cost to go is $J_1(x_1) = 3x_1^2 + 2$

d) STAGE 0:

$$\begin{aligned} J_0(x_0) = &\min_{u_0}[E_{w_0}\{g_0(x_0, u_0, w_0) + J_1(x_1)\}] = && (5)\\ &\min_{u_0}[E_{w_0}\{u_0^2 + 3(2x_0 + u_0 + w_0)^2 + 2\}] = && (6)\\ &\min_{u_0}[E_{w_0}\{u_0^2 + 3((2x_0 + u_0)^2 + w_0^2 + 2(2x_0 + u_0)w_0 + 2)\}] = && (7)\\ &\min_{u_0}[u_0^2 + 3(2x_0 + u_0)^2 + 3E_{w_0}\{w_0^2\} + 6(2x_0 + u_0)E_{w_0}\{w_0\} + 2] = && (8)\\ &\min_{u_0}[u_0^2 + 3(2x_0 + u_0)^2 + 8] = && (9)\\ &\min_{u_0} C(u_0, x_0) && (10)\\ & && (11)\end{aligned}$$

We now compute the value u_0 for which the gradient with respect to u_0 is zero:

$$\frac{\partial C}{\partial u_0} = 2u_0 + 6(2x_0 + u_0) = 0,$$

and then $8u_0 = -12x_0$, bringing us to $u_0 = -\frac{3}{2}x_0$.

Considering the constrained input, we obtain:

1. if $x_0 \leq 0 : u_0^* = 0$ and $J_0(x_0) = 8 + 12x_0^2$
2. if $x_0 > 2/3 : u_0^* = -1$ and $J_0(x_0) = 9 + 3(2x_0 - 1)^2$
3. if $0 < x_0 \leq 2/3 : u_0^* = -\frac{3}{2}x_0$ and $J_0(x_0) = \frac{9}{4}x_0^2 + \frac{3}{4}x_0^2 + 8 = 3x_0^2 + 8$

e) The optimal policies are the same as in the previous 2 cases, while the optimal costs-to-go are $J_1(x_1) = 3x_1^2$, and

1. if $x_0 \leq 0 : u_0^* = 0$ and $J_0(x_0) = 12x_0^2$
2. if $x_0 > 2/3 : u_0^* = -1$ and $J_0(x_0) = 1 + 3(2x_0 - 1)^2$
3. if $0 < x_0 \leq 2/3 : u_0^* = -\frac{3}{2}x_0$ and $J_0(x_0) = \frac{9}{4}x_0^2 + \frac{3}{4}x_0^2 = 3x_0^2$

Problem 2 [25 points]

Consider the following dynamic system:

$$\begin{aligned} x_{k+1} &= w_k, \\ x_k &\in \{0, 1, 2\} \\ u_k &\in \{A, B\} \end{aligned} \tag{12}$$

The transition probabilities $p_{i,j}(u_k) := P(w_k = j | x_k = i, u_k)$ between the states are given by:

$$\begin{aligned}
&p_{00}(A) = 0.2 && p_{01}(A) = 0.6 && p_{02}(A) = 0.2, \\
&p_{10}(A) = 0.1(1-\gamma) && p_{11}(A) = \gamma && p_{12}(A) = 0.9(1-\gamma), \\
&p_{20}(A) = 0.1 && p_{21}(A) = 0.5 && p_{22}(A) = 0.4, \\
&p_{00}(B) = 0.5 && p_{01}(B) = 0 && p_{02}(B) = 0.5, \\
&p_{10}(B) = 0.6(1-\gamma) && p_{11}(B) = \gamma && p_{12}(B) = 0.4(1-\gamma), \\
&p_{20}(B) = 0.7 && p_{21}(B) = 0 && p_{22}(B) = 0.3,
\end{aligned}$$

with $0 < \gamma \leq 1$. The cost function that has to be minimized is the following:

$$\lim_{N\to\infty} \mathrm{E}\left[\sum_{k=0}^{N-1} \alpha^k g(x_k, u_k)\right],$$

with:

$$\begin{aligned}
g(0, A) &= 1 + \beta && g(1, A) = 3\beta && g(2, A) = 3 - \beta, \\
g(0, B) &= 2 && g(1, B) = \beta^2 && g(2, B) = 2.
\end{aligned}$$

where α and $\beta \in \mathcal{R}$ are parameters.

a) Given $\alpha = 1$, state the range of β and γ for which the above problem is a well-defined infinite horizon problem with finite cost. A short explanation is required. Then, draw the probability transition graphs, with the associated probabilities and costs denoted on each arc. *[2 points]*

b) Now let's assume $\beta = 1$, $\gamma = 0.1$. Is there any range of α such that the above problem is a well-defined infinite horizon problem with finite cost? If yes, explain why. *[3 points]*

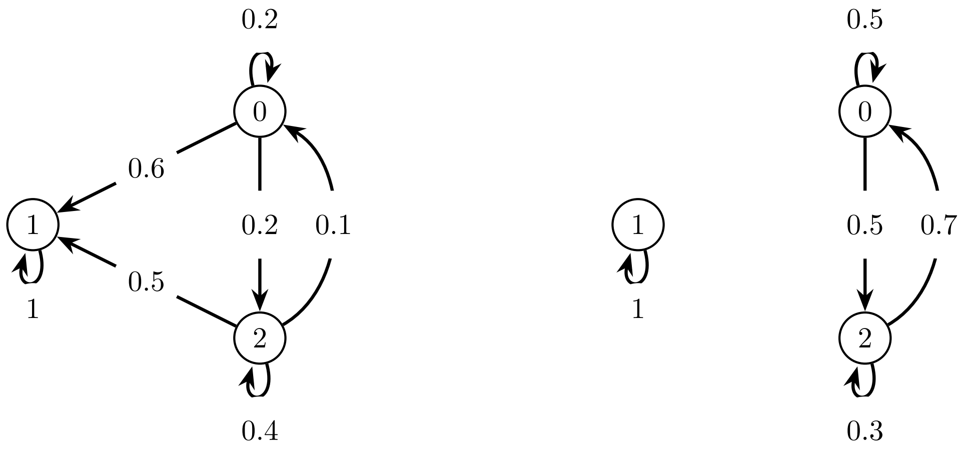
c) Assume $\alpha = 0.1$, $\beta = 1$ and $\gamma = 0$. Perform one iteration of the **Value Iteration Algorithm** for the resulting Discounted Cost Problem. Consider $J_0(0) = 10$, $J_0(1) = 10$, $J_0(2) = 20$ as the initial guesses. *[10 points]*

d) Assume $\alpha = 1$, $\beta = 0$ and $\gamma = 1$. Perform one iteration of the **Policy Iteration Algorithm** for the resulting Stochastic Shortest Path Problem, i.e. compute $\mu^1(0)$, $\mu^1(1)$, $\mu^1(2)$. List all proper policies that can be used as initialization for the **Policy Iteration Algorithm**, and pick as the initial guess one with $\mu^0(0) = B$. *[10 points]*

Solution 2

a) γ must be equal to 1 and β equal to 0. Indeed, the probability of leaving the termination state (1) must be 0 ($\gamma = 1$) and the final cost 0 as well ($\beta = 0$).

The graphs are the following (First is $u = A$, second is $u = B$):



b) $\alpha \in (-1, 1)$. Without a termination stage, the stage cost must be discounted exponentially. (The lecture notes states $\alpha \in (0, 1)$ which is valid as well)

c) Let's initialize the value iteration algorithm and perform one iteration:

- Initial guess:

$$J_0(0) = 10, J_0(1) = 10, J_0(2) = 20$$

- Iteration 1:

$$\begin{aligned}
J_1(0) &= \min_{u\in\{A,B\}} [g(0,u) + \alpha p_{00}(u) J_0(0) + \alpha p_{01}(u) J_0(1) + \alpha p_{02}(u) J_0(2)] = \\
&= \min_{u\in\{A,B\}} [2 + 0.1 \cdot 2 + 0.1 \cdot 6 + 0.1 \cdot 4,\ 2 + 0.1 \cdot 5 + 0.1 \cdot 10] \\
&= \min_{u\in\{A,B\}} [3.2, 3.5] \implies J_1(0) = 3.2 \\
J_1(1) &= \min_{u\in\{A,B\}} [g(1,u) + \alpha p_{10}(u) J_0(0) + \alpha p_{11}(u) J_0(1) + \alpha p_{12}(u) J_0(2)] = \\
&= \min_{u\in\{A,B\}} [3 + 0.1 \cdot 1 + 0.1 \cdot 18,\ 1 + 0.1 \cdot 6 + 0.1 \cdot 8] \\
&= \min_{u\in\{A,B\}} [4.9, 2.4] \implies J_1(1) = 2.4 \\
J_1(2) &= \min_{u\in\{A,B\}} [g(2,u) + \alpha p_{20}(u) J_0(0) + \alpha p_{21}(u) J_0(1) + \alpha p_{22}(u) J_0(2)] = \\
&= \min_{u\in\{A,B\}} [2 + 0.1 \cdot 1 + 0.1 \cdot 5 + 0.1 \cdot 8,\ 2 + 0.1 \cdot 7 + 0.1 \cdot 6] \\
&= \min_{u\in\{A,B\}} [3.4, 3.3] \implies J_1(0) = 3.3
\end{aligned}$$

d) The stationary policies are:

- $\mu^0(0) = A$, $\mu^0(1) = A$, $\mu^0(2) = A$ proper
- $\mu^0(0) = A$, $\mu^0(1) = A$, $\mu^0(2) = B$ proper

- $\mu^0(0) = A$, $\mu^0(1) = B$, $\mu^0(2) = A$ proper
- $\mu^0(0) = A$, $\mu^0(1) = B$, $\mu^0(2) = B$ proper
- $\mu^0(0) = B$, $\mu^0(1) = A$, $\mu^0(2) = A$ proper
- $\mu^0(0) = B$, $\mu^0(1) = A$, $\mu^0(2) = B$ improper
- $\mu^0(0) = B$, $\mu^0(1) = B$, $\mu^0(2) = A$ proper
- $\mu^0(0) = B$, $\mu^0(1) = B$, $\mu^0(2) = B$ improper

Regardless of $\mu^0(1)$, if we choose $\mu^0(0) = B$ and $\mu^0(2) = B$ we remain into the loop and the cost goes to infinity. We can use as the initialization policy both $\mu^0(0) = B$, $\mu^0(1) = B$, $\mu^0(2) = A$ and $\mu^0(0) = B$, $\mu^0(1) = A$, $\mu^0(2) = A$.

Let's consider the first policy $\mu^0(0) = B$, $\mu^0(1) = B$, $\mu^0(2) = A$ as initialization. State 1 is the termination state and has zero cost, independent of the policy. The termination state is not considered during policy evaluation and improvement. For states 0 and 2, we initialize the policy iteration algorithm and perform one iteration.

- Initial guess: $\mu^0(0) = B$, $\mu^0(2) = A$
- Iteration 1:
  1. Policy evaluation:

$$\begin{aligned}
J_{\mu^0}(0) &= g(0, B) + \alpha p_{00}(B) J_{\mu^0}(0) + \alpha p_{02}(B) J_{\mu^0}(2) = \\
&= 2 + 0.5 J_{\mu^0}(0) + 0.5 J_{\mu^0}(2) \\
&\implies J_{\mu^0}(0) = 4 + J_{\mu^0}(2) \\
J_{\mu^0}(2) &= g(2, A) + \alpha p_{20}(A) J_{\mu^0}(0) + \alpha p_{22}(A) J_{\mu^0}(2) = \\
&= 3 + 0.1 J_{\mu^0}(0) + 0.4 J_{\mu^0}(2) \\
&\implies J_{\mu^0}(2) = 3 + 0.4 + 0.5 J_{\mu^0}(2) \\
&J_{\mu^0}(2) = 6.8 \\
&J_{\mu^0}(0) = 10.8
\end{aligned}$$

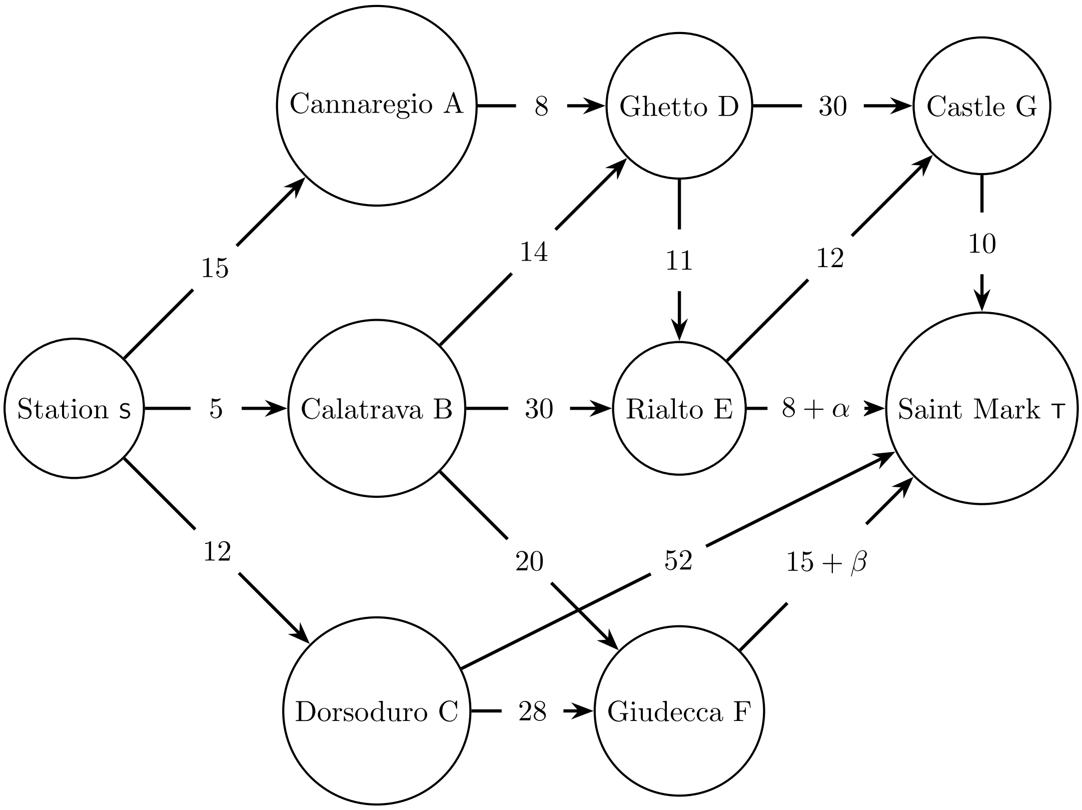
  2. Policy improvement:

$$\begin{aligned}
\mu^1(0) &= \arg\min_{u \in \{A,B\}} [g(0, u) + \alpha p_{00}(u) J_{\mu^0}(0) + \alpha p_{02}(u) J_{\mu^0}(2)] = \\
&= \arg\min_{u \in \{A,B\}} [1 + 0.2 \cdot 10.8 + 0.2 \cdot 6.8,\ 2 + 0.5 \cdot 10.8 + 0.5 \cdot 6.8] \\
&= \arg\min_{u \in \{A,B\}} [1 + 2.16 + 1.36,\ 2 + 5.4 + 3.4] \\
&= \arg\min_{u \in \{A,B\}} [4.52, 10.8] \implies \mu^1(0) = A \\
\mu^1(2) &= \min_{u \in \{A,B\}} [g(2, u) + \alpha p_{20}(u) J_{\mu^0}(0) + \alpha p_{22}(u) J_{\mu^0}(2) = \\
&= \arg\min_{u \in \{A,B\}} [3 + 0.1 \cdot 10.8 + 0.4 \cdot 6.8,\ \ 2 + 0.7 \cdot 10.8 + 0.3 \cdot 6.8] \\
&= \arg\min_{u \in \{A,B\}} [3 + 1.08 + 2.72,\ 2 + 7.56 + 2.04] \\
&= \arg\min_{u \in \{A,B\}} [6.8, 11.6] \implies \mu^1(2) = A
\end{aligned}$$

  The new policy is given by $\mu^1(0) = A$, $\mu^1(1) = B$ or A, $\mu^1(2) = A$

Problem 3 **[25 points]**

You just reached the main station in Venice (Venezia Santa Lucia) and your goal is to go to Saint Mark's Square as fast as possible through the *calli* (streets) or the canals of Venice, that is, you need to find the shortest path from the main station to the square. The following graph represents a map of Venice showing the travelling time between the Station (node S), Saint Mark's square (node T), and major neighbourhoods (nodes A to G). The numbers on the edges represent the time (in minutes) to travel between the nodes. $\alpha \in [0, 10]$ and $\beta \in [0, 10]$ are constant delays that apply when Venice is overcrowded.



a) Recall that $\mathcal{V}$ is the vertex space of the graph and $|\mathcal{V}|$ is the number of elements in the space. How many elements are in the space? What is the time horizon N of the equivalent Deterministic Finite State System? *[2 points]*

b) Consider solving the SP problem using the Label Correcting Algorithm with the best-first search method on the equivalent DFS system. Complete Table 1 with all iterations and find the optimal path, as α and β vary. *[8 points]*

c) Complete Table 2 for three iterations using a depth-first search to determine at each iteration which node to remove from the OPEN bin. In this case consider α and β as equal to 0. *[5 points]*

d) Complete Table 3 for three iterations using a breadth-first search to determine at each iteration which node to remove from the OPEN bin. In this case consider α and β as equal to 0. *[5 points]*

e) Perform the initialization and one iteration step of the Dynamic Programming Algorithm on the equivalent DFS system. In this case consider α and β as equal to 0. *[5 points]*

P.S. When more than one node enter OPEN in one iteration, consider alphabetic order (as example: A and B enter OPEN at the same iteration, A enter first, B enter second)

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
1	s										

Table 1: Label Correcting Algorithm Table for question b).

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	15	5	12	∞	∞	∞	∞	∞
2											
3											
4											

Table 2: Label Correcting Algorithm Table for question c). (LAST IN FIRST OUT)

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
						...					
3	B	C,D,E,F	0	15	5	12	19	35	25	∞	∞
4											
5											
6											

Table 3: Label Correcting Algorithm Table for question d). (FIRST IN FIRST OUT)

Solution 3

a) Cardinality is 9, time horizon is 8

b) Table filled in for question 2:

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	15	5	12	∞	∞	∞	∞	∞
2	B	A,C,D,E,F	0	15	5	12	19	35	25	∞	∞
3	C	A,D,E,F	0	15	5	12	19	35	25	∞	64
4	A	D,E,F	0	15	5	12	19	35	25	∞	64
5	D	E,F,G	0	15	5	12	19	30	25	49	64
6	F	E,G	0	15	5	12	19	30	25	49	40+β*
7	E	G	0	15	5	12	19	30	25	42	$\min\{38+\alpha, 40+\beta\}$
8	G		0	15	5	12	19	30	25	42	$\min\{38+\alpha, 40+\beta\}$**
9	-		0	15	5	12	19	30	25	42	$\min\{38+\alpha, 40+\beta\}$

Table 4: Label Correcting Algorithm Table for question ii).
**Note that 64 is always more than 40 +β*
*** Note that 52 is always more than 40 + β and 38 + α*

If $\alpha - 2 < \beta$ the optimal path is $\mathtt{S}, B, D, E, \mathtt{T}$, and the optimal cost is 38 + α. If $\alpha - 2 > \beta$, then the optimal path is $\mathtt{S}, B, F, \mathtt{T}$ with optimal cost 40 +β. If $\alpha - 2 = \beta$ the two paths are equivalent.

c) Table 5 filled in for question c) (ALL ITERATIONS).

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	15	5	12	∞	∞	∞	∞	∞
2	C	A,B,F	0	15	5	12	∞	∞	40	∞	64
3	F	A,B	0	15	5	12	∞	∞	40	∞	55
4	B	A,D,E,F	0	15	5	12	19	35	25	∞	55
5	F	A,D,E	0	15	5	12	19	35	25	∞	40
6	E	A,D,G	0	15	5	12	19	35	25	47	40
7	G	A,D	0	15	5	12	19	35	25	47	40
8	D	A,E,G	0	15	5	12	19	30	25	47	40
9	G	A,E	0	15	5	12	19	30	25	47	40
10	E	A,G	0	15	5	12	19	30	25	42	38
11	G	A	0	15	5	12	19	30	25	42	38
12	A	D	0	15	5	12	19	30	25	42	38
13	D	E,G	0	15	5	12	19	30	25	42	38
14	G	E	0	15	5	12	19	30	25	42	38
15	E		0	15	5	12	19	30	25	42	38
16	-		0	15	5	12	19	30	25	42	38

Table 5: Label Correcting Algorithm Table for question c). (LAST IN FIRST OUT)

d) Table 6 filled in for question d) (ALL ITERATIONS).

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_G	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	15	5	12	∞	∞	∞	∞	∞
2	A	B,C,D	0	15	5	12	23	∞	∞	∞	∞
3	B	C,D,E,F	0	15	5	12	19	35	25	∞	∞
4	C	D,E,F	0	15	5	12	19	35	25	∞	64
5	D	E,F,G	0	15	5	12	19	30	25	49	64
6	E	F,G	0	15	5	12	19	30	25	42	38
7	F	G	0	15	5	12	19	30	25	42	38
8	G		0	15	5	12	19	30	25	42	38
9	-		0	15	5	12	19	30	25	42	38

Table 6: Label Correcting Algorithm Table for question d). (FIRST IN FIRST OUT)

e) Start DP Algorithm by initializing with one move to the end:

i	$J_{N-1}(i)$
S	∞
A	∞
B	∞
C	52
D	∞
E	8
F	15
G	10

Recursion: (2 moves to the end)

i	$J_{N-2}(i)$
S	$12 + 52 = 64$
A	∞
B	$\min\{30 + 8, 20 + 15\} = 35$
C	$\min\{28 + 15, 52\} = 43$
D	$\min\{11 + 8, 30 + 10\} = 19$
E	$\min\{8, 12 + 10\} = 8$
F	15
G	10

Problem 4 **[25 points]**

Consider a control maneuver for a cart on a road. The cart can move in a straight line by applying a force to the wheels. The simplified dynamics are:

$$\ddot{x}(t) = u(t),$$

where $\ddot{x}(t)$ is the acceleration, $\dot{x}(t)$ the velocity, $x(t)$ the position, and $u(t)$ the normalized thrust of the cart with $-1 \leq u(t) \leq 2$.

At time 0, the cart is at rest, thus the initial position is 0 and the initial velocity is 0. At time T, the cart is required to be at position 4 with velocity equal to 0.

The goal is to compute a control trajectory that allows the cart to reach the terminal state while minimizing the time T. Pontryagin's Minimum Principle is here required.
The co-state of this problem is defined as $\mathbf{p}(t) = [p_1(t), p_2(t)]$ for all $t \in [0, T]$.

a) Write the system dynamics $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), u(t))$ and all boundary conditions. *[3 points]*

b) Write the Hamiltonian $H(\mathbf{x}(t), u(t), \mathbf{p}(t))$ of the problem. *[3 points]*

c) Write the co-state derivative $\dot{\mathbf{p}}(t)$ of the problem. Provide a guess for the optimal control input, too. *[6 points]*

d) Compute the optimal time T and optimal control input $u^*(t)$ *[13 points]*

Solution 4

a) We define $\mathbf{x}(t) = [x_1(t), x_2(t)] = [x(t), \dot{x}(t)]$.
The dynamics are $\dot{x}_1(t) = x_2(t)$, $\dot{x}_2(t) = u(t)$

Boundary conditions:

- $x_1(0) = 0$, $x_1(T) = 4$
- $x_2(0) = 0$, $x_2(T) = 0$

b) The Hamiltonian is

$$\begin{aligned} H(\mathbf{x}, u, \mathbf{p}) &= 1 + p_1(t)\dot{x}_1(t) + p_2(t)\dot{x}_2(t) \\ &= 1 + p_1(t)x_2(t) + p_2(t)u(t) \end{aligned}$$

c) The co-state derivatives are defined by

$$\dot{\mathbf{p}}(t) = -\frac{\partial H(\mathbf{x}, u, \mathbf{p})}{\partial \mathbf{x}},$$

hence we have $\dot{p}_1(t) = 0$, $\dot{p}_2(t) = -p_1(t)$,

Since the Hamiltonian is linear in u, The optimal control input $u^*(t)$ is attained on the boundaries of the control space C= $[-1, 2]$. The solution is the following (Bang bang Solution):

- $u^*(t) = -1$, if $p_2(t) > 0$
- $u^*(t) = 2$, if $p_2(t) < 0$
- undefined if $p_2(t) = 0$

d) The adjoint equations are integrated resulting in the following equations for the co-states:

$$\begin{aligned} p_1(t) &= C_1 \\ p_2(t) &= -C_1 t + C_2 \end{aligned}$$

where C_1, C_2 are constants.

The intuitive behaviour is the following: the cart first accelerates towards the goal, then decelerates to reach the target velocity at the goal. Combining this intuition and using point c), we have:

1. $0 \le t \le t_1$
2. $t_1 \le t \le T$

and $t_1 \le T$. We have to find t_1 and T that minimize the cost function.

- First arc $t \in [0, t_1]$
  We have $\dot{x}_2(t) = 2$ with the boundary conditions $x_1(0) = 0$, $x_2(0) = 0$
  As a result we have

  $$\begin{aligned} x_1(t) &= t^2 + c_0 t + c_1 \\ x_2(t) &= 2t + c_0 \end{aligned}$$

  with $c_0 = 0$ and $c_1 = 0$ (given the boundary conditions), hence:

  $$x_1(t) = t^2 \tag{13}$$
  $$x_2(t) = 2t \tag{14}$$

- Second arc $t \in [t_1, T]$
  We have: $\dot{x}_2(t) = -1$; $\dot{x}_1(t) = x_2(t) = -t + c_2$. We obtain:

$$x_1(t) = -t^2/2 + c_2 t + c_3 \tag{15}$$
$$x_2(t) = -t + c_2 \tag{16}$$

with c_2 and c_3 two constants. To ensure continuity of the states x (a discontinuity in position or velocity would not make physical sense), we require that the equations $(15, 16)$ satisfy $(13, 14)$.
We obtain:

$$x_1(t_1) = t_1^2 = -t_1^2/2 + c_2 t_1 + c_3 \tag{17}$$
$$x_2(t_1) = 2t_1 = -t_1 + c_2 \tag{18}$$

It follows that $c_2 = 3t_1$ and $c_3 = -3t_1^2/2$. Furthermore, at time T, the terminal conditions $x_1(T) = 4$ and $x_2(T) = 0$ need to be satisfied. As a result we have

$$x_1(T) = -T^2/2 + 3t_1 T - 3t_1^2/2 = 4 \tag{19}$$
$$x_2(T) = -T + 3t_1 = 0 \tag{20}$$

from equation (20) we obtain $t_1 = \frac{T}{3}$; inserting in (19) we obtain a quadratic equation for T:

$$T^2/3 = 4 \tag{21}$$

which is solved by $T = \pm 2\sqrt{3}$. Since $T > 0$, the minimum terminal time is $T = 2\sqrt{3}$. Using (20), we obtain the switching time $t_1 = 2\sqrt{3}/3$ and thus, the optimal input is:

$$u(t) = \begin{cases} 2 & \text{for } t \in [0, 2\sqrt{3}/3] \\ -1 & \text{for } t \in [2\sqrt{3}/3, 2\sqrt{3}] \end{cases} \tag{22}$$