Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

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

February 2nd, 2022

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

Exam

Exam Duration: **150 minutes**

Number of Problems: **4**

Permitted aids: One A4 hand written sheet of paper.
No calculators allowed.

Problem 1 **[25 points]**

Consider the dynamics

$$x_{k+1} = x_k + w_k \cdot u_k, \quad k = 0, 1$$

with initial state $x_0 \in \mathbb{R}$. At each stage k, x_k is the state of the system, $u_k \in \mathbb{R}$ the input and $w_k \in \mathbb{R}$ the disturbance of the system. We want to minimize the following cost function:

$$\mathop{\mathbb{E}}_{w_0, w_1} \{-x_2^2 + u_1^2 + u_0^2\}.$$

1. What is the terminal cost $g(x_2)$? *[2 points]*

2. What is the stage cost $g(x_k, u_k, w_k)$, $k = 0, 1$? *[2 points]*

3. The disturbance is normally distributed, $w_1 \sim \mathcal{N}(\alpha, 1)$, where the mean is α and the variance is 1. At each time step, you will receive a forecast y_k that indicates the value of the mean α of the disturbance w_k. If $y_k = 1$, the mean of w_k is $\alpha = 1$, and if $y_k = 2$, the mean of w_k is $\alpha = 2$. A priori information shows that both values of the mean α are equally probable.
   Reformulate your dynamics to take the forecast into account, and calculate the optimal policy $u_1^* = \mu_1^*(\tilde{x}_1)$, as well as the corresponding optimal cost-to-go $J_1^*(\tilde{x}_1)$ at step 1. *[11 points]*

4. Now assume that at time $k = 0$ the input u_0 is constrained between -1 and 1, i.e. $u_0 \in [-1, 1]$. Given the forecast $y_0 = 1$, calculate the optimal policy $u_0^* = \mu_0^*(\tilde{x}_0 | y_0 = 1)$ and the corresponding cost-to-go $J_0^*(\tilde{x}_0 | y_0 = 1)$. *[10 points]*

   If you were not able to solve question 3, assume $J_1^*(\tilde{x}_1) = (1 + y_1)x_1^2$

Solution 1

1. $g(x_2) = -x_2^2$
2. $g_k(x_k, u_k, w_k) = u_k^2 \quad k = 0, 1$
3. There was a typo in this exercise, the final cost should have been x_2^2. However the exercise was still solvable and we will give points to anyone that had a coherent answer and **understood the concept of forecast**.

   first let's define the extended state space: $\tilde{x}_k = (x_k, y_k)$
   the dynamics are:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + w_k \cdot u_k \\ \xi_k \end{bmatrix} \qquad k = 0, 1$$

   ξ_k takes the values 1 and 2 with probability 0.5.

   Then calculate the cost-to-go

   - $J_2^*(x_2) = -x_2^2$
   - <u>**k = 1**</u>

$$\begin{aligned} J_1(x_1, y_1) &= \min_{u_1} \underset{w_1|y_1}{E} \{u_1^2 + 0.5J(x_2, y_2 = 1) + 0.5J(x_2, y_2 = 2)\} \\ &= \min_{u_1} \underset{w_1|y_1}{E} \{u_1^2 + 0.5J(x_1 + u_1 \cdot w_1, y_2 = 1) + 0.5J(x_1 + u_1 \cdot w_1, y_2 = 2)\} \\ &= \min_{u_1} \underset{w_1|y_1}{E} \{u_1^2 - (x_1 + u_1 \cdot w_1)^2\} = \min_{u_1} \underset{w_1|y_1}{E} \{u_1^2 - x_1^2 - u_1^2 \cdot w_1^2 - 2x_1u_1w_1\} \end{aligned}$$

     – <u>$y_1 = 1$</u>

$$\begin{aligned} J_1(x_1, 1) &= \min_{u_1} \underset{w_1|y_1=1}{E} \{u_1^2 - x_1^2 - u_1^2 \cdot w_1^2 - 2x_1u_1w_1\} \\ &= \min_{u_1} [u_1^2 - x_1^2 - 2u_1^2 - 2x_1u_1] = \min_{u_1} [-u_1^2 - 2x_1u_1 - x_1^2] \\ \frac{\partial J_1(x_1, 1)}{\partial u_1} &= -2u_1 - 2x_1 = 0 \quad \Rightarrow \quad u_1^* = \mu(x_1, 1)^* = -x_1 \\ J_1^*(x_1, 1) &= -x_1^2 - x_1^2 + 2x_1^2 = 0 \end{aligned}$$

     – <u>$y_1 = 2$</u>

$$\begin{aligned} J_1(x_1, 2) &= \min_{u_1} \underset{w_1|y_1=2}{E} \{u_1^2 - x_1^2 - u_1^2 \cdot w_1^2 - 2x_1u_1w_1\} \\ &= \min_{u_1} u_1^2 - x_1^2 - 5u_1^2 - 4x_1u_1 = \min_{u_1} -4u_1^2 - 4x_1u_1 - x_1^2 \\ \frac{\partial J_1(x_1, 2)}{\partial u_1} &= -8u_1 - 4x_1 = 0 \quad \Rightarrow \quad u_1^* = \mu(x_1, 2)^* = \frac{-x_1}{2} \\ J_1^*(x_1, 2) &= -x_1^2 - x_1^2 + 2x_1^2 = 0 \end{aligned}$$

     By doing that we solve a **maximization** problem, if the student does not notice it and solve it like shown above, she/he gets full points.

If the student notices it and say that there is a problem there or that the answer is plus or minus infinity she/he gets full points.
But the student needs to show that she/he has understood how a forecast works, if she/he does not make a speration between $y_1 = 1$ and $y_1 = 2$, she/he does not receive points.
However any mention that there is a problem and that the solution is plus or minus infinity she/he gets 2 points

4. If the student chooses to keep its results from the previous part and makes something that make senses, she/he gets full points

4*

$$\begin{aligned}
J_0(x_0, 1) &= \min_{u_0} \underset{w_0|y_0=1}{E} \{u_0^2 + 0.5J_1(x_1, y_1 = 1) + 0.5J_1(x_1, y_1 = 2)\} \\
&= \min_{u_0} \underset{w_0|y_0=1}{E} \{u_0^2 + x_1^2 + 1.5x_1^2\} \\
&= \min_{u_0} \underset{w_0|y_0=1}{E} \{u_0^2 + 2.5x_1^2\} \\
&= \min_{u_0} \underset{w_0|y_0=1}{E} \{u_0^2 + 2.5(x_0 + w_0u_0)^2\} \\
&= \min_{u_0} \underset{w_0|y_0=1}{E} \{u_0^2 + 2.5x_0^2 + 2.5w_0^2u_0^2 + 5x_0u_0w_0\} \\
&= \min_{u_0}\{u_0^2 + 2.5x_0^2 + 5u_0^2 + 5x_0u_0\} = \min_{u_0}\{6u_0^2 + 2.5x_0^2 + 5x_0u_0\} \\
\frac{\partial J_0(x_0, 1)}{\partial u_0} &= 12u_0 + 5x_0 = 0 \quad \Rightarrow \quad u_0 = -\frac{5}{12}x_0
\end{aligned}$$

taking the constraints into account we get:

$$u_0^* = \mu(x_0, 1) = \begin{cases} -1 & x_0 > \frac{12}{5} \\ -\frac{5}{12}x_0 & |x_0| \leq \frac{12}{5} \\ 1 & x_0 < -\frac{12}{5} \end{cases}$$

$$J_0^*(x_0, 1) = \begin{cases} 2.5x_0^2 - 5x_0 + 6 & x_0 > \frac{12}{5} \\ \frac{35}{24}x_0^2 & |x_0| \leq \frac{12}{5} \\ 2.5x_0^2 + 5x_0 + 6 & x_0 < -\frac{12}{5} \end{cases}$$

Problem 2 [25 points]

In New York City, Midtown High School student Peter Parker was bitten by a radioactive spider and acquired the agility and strength of a spider, and the ability to adhere to walls and ceilings. Because of the death of his uncle, he decided to fight against criminals as a masked superhero, Spider-Man.

As he wants to protect the people who are close to him, he must be careful to not expose his true identity. Every morning he can decide whether he will act as Spider-Man to save people in the city (S), or as a normal student (N) during the rest of the day. At the end of the day, he can arrive home in one of three states: not injured (1), injured (2), or exposed (0). The dynamics of the life of Peter Parker or Spider-Man can be described as:

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

Once the true identity of Peter Parker is exposed, it is not possible to cover this up again. So the transition probabilities $P_{i,j}(u_k) := P(w_k = j|x_k = i, u_k)$ between the states are given by:

$$\begin{aligned} &P_{00}(S) = 1.0, \\ &P_{10}(S) = 0.3 \qquad P_{11}(S) = 0.5 \qquad P_{12}(S) = 0.2, \\ &P_{20}(S) = 0.4 \qquad P_{21}(S) = 0.1 \qquad P_{22}(S) = 0.5, \\ &P_{00}(N) = 1.0, \\ &P_{10}(N) = 0.1 \qquad P_{11}(N) = 0.8 \qquad P_{12}(N) = 0.1, \\ &P_{20}(N) = 0.2 \qquad P_{21}(N) = 0.7 \qquad P_{22}(N) = 0.1. \end{aligned}$$

Peter Parker wants to maintain the reputation of the friendly neighborhood Spider-Man, and every action he takes will have some reward. The associated reward is the following:

$$\begin{aligned} &g(0,S,0) = -2 \\ &g(1,S,0) = -5 \qquad g(1,S,1) = 7 \qquad g(1,S,2) = 10 \\ &g(2,S,0) = -5 \qquad g(2,S,1) = 5 \qquad g(2,S,2) = 7 \\ &g(0,N,0) = -1 \\ &g(1,N,0) = -10 \qquad g(1,N,1) = 3 \qquad g(1,N,2) = -4 \\ &g(2,N,0) = -10 \qquad g(2,N,1) = 5 \qquad g(2,N,2) = -5 \end{aligned}$$

The total reward function that he wants to **maximize** is the following:

$$\lim_{N\to\infty} \mathbb{E} \sum_{k=0}^{N-1} \alpha^k r(x_k, u_k),$$

with:

$$r(x_k, u_k) = \mathop{\mathbb{E}}_{(w|x=x_k, u=u_k)} [g(x,u,w)]$$

where $\alpha = 0.8$, which is the discount factor.

a) Calculate the expected reward $r(x_k, u_k)$ for each state-action pair (x_k, u_k), considering the states $x_k \in \{0, 1, 2\}$, and the actions $u_k \in \{S, N\}$. *[3 points]*

b) Using $V^*(x)$ as the optimal expected value for a certain state x in this Infinite Horizon Problem, write down the Bellman equation for each state $x \in \{0, 1, 2\}$, using $r(x, u)$ calculated from the previous question. *[3 points]*

c) Perform one iteration of the **Value Iteration Algorithm** for the resulting Infinite Horizon Problem. Consider $V^0(0) = 5$, $V^0(1) = 10$, $V^0(2) = 10$ as the initial guesses. *[6 points]*

d) Perform one iteration of the **Policy Iteration Algorithm** for the resulting Infinite Horizon Problem, i.e. compute $\mu^1(0)$, $\mu^1(1)$, $\mu^1(2)$. Consider that Peter Parker wants to act as Spider-Man every day at first, and use $\mu^0(0) = S$, $\mu^0(1) = S$, $\mu^0(2) = S$ as the initial guesses. In the policy evaluation part, round the values to 2 decimal places, i.e. in the form x.xx. *[9 points]*

e) Denote the optimal value vector as $\mathbf{x} = [V(0), V(1), V(2)]^T$. The optimal value vector can be obtained by solving a **Linear Programming** of the generic form

$$\begin{aligned} &\textit{minimize} \quad \mathbf{f}^T\mathbf{x} \\ &\textit{subject to} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$

where $\mathbf{f}$ and $\mathbf{b}$ are row vectors, and $\mathbf{A}$ is a matrix. Write down a choice of $\mathbf{f}$, $\mathbf{b}$ and $\mathbf{A}$ such that the optimal value vector $\mathbf{x}$ is obtained by solving the linear program. *[4 points]*

Solution 2

a) The expected reward is calculated as the following:

$$r(x_k, u_k) = P_{x_k,0}(u_k)g(x_k, u_k, 0) + P_{x_k,1}(u_k)g(x_k, u_k, 1) + P_{x_k,2}(u_k)g(x_k, u_k, 2)$$

Plug in the values:

$$\begin{aligned}
r(0, S) = -2 \qquad & = 1.0 \cdot (-2) \\
r(1, S) = 4 \qquad & = 0.3 \cdot (-5) + 0.5 \cdot 7 + 0.2 \cdot 10 \\
r(2, S) = 2 \qquad & = 0.4 \cdot (-5) + 0.1 \cdot 5 + 0.5 \cdot 7 \\
r(0, N) = -1 \qquad & = 1.0 \cdot (-1) \\
r(1, N) = 1 \qquad & = 0.1 \cdot (-10) + 0.8 \cdot 3 + 0.1 \cdot (-4) \\
r(2, N) = 1 \qquad & = 0.2 \cdot (-10) + 0.7 \cdot 5 + 0.1 \cdot (-5)
\end{aligned}$$

b) The Bellman equation is

$$V^*(x) = \max_{u \in \{S,N\}} [r(x, u) + \alpha P_{x,0}(u)V^*(0) + \alpha P_{x,1}(u)V^*(1) + \alpha P_{x,2}(u)V^*(2)], \forall x \in \{0, 1, 2\}$$

Plug in the values:

$$\begin{aligned}
V^*(0) &= \max_{u \in \{S,N\}} [r(0, u) + \alpha P_{00}(u)V^*(0)] \\
&= \max_{u \in \{S,N\}} \{-2 + 0.8V^*(0), -1 + 0.8V^*(0)\} \\
V^*(1) &= \max_{u \in \{S,N\}} [r(1, u) + \alpha P_{10}(u)V^*(0) + \alpha P_{11}(u)V^*(1) + \alpha P_{12}(u)V^*(2)] \\
&= \max_{u \in \{S,N\}} \{4 + 0.24V^*(0) + 0.4V^*(1) + 0.16V^*(2), 1 + 0.08V^*(0) + 0.64V^*(1) + 0.08V^*(2)\} \\
V^*(2) &= \max_{u \in \{S,N\}} [r(2, u) + \alpha P_{20}(u)V^*(0) + \alpha P_{21}(u)V^*(1) + \alpha P_{22}(u)V^*(2)] \\
&= \max_{u \in \{S,N\}} \{2 + 0.32V^*(0) + 0.08V^*(1) + 0.4V^*(2), 1 + 0.16V^*(0) + 0.56V^*(1) + 0.08V^*(2)\}
\end{aligned}$$

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

- Initial guess:

$$V^0(0) = 5, V^0(1) = 10, V^0(2) = 10$$

- Iteration 1:

$$
\begin{aligned}
V^1(0) &= \max_{u\in\{S,N\}} \{r(0,u) + \alpha P_{00}(u)V^0(0)\} \\
&= \max_{u\in\{S,N\}} \{-2 + 0.8\cdot 5, -1 + 0.8\cdot 5\} \\
&= \max_{u\in\{S,N\}} \{2, 3\} \implies V^1(0) = 3 \\
V^1(1) &= \max_{u\in\{S,N\}} \{r(1,u) + \alpha P_{10}(u)V^0(0) + \alpha P_{11}(u)V^0(1) + \alpha P_{12}(u)V^0(2)\} \\
&= \max_{u\in\{S,N\}} \{4 + 0.24\cdot 5 + 0.4\cdot 10 + 0.16\cdot 10, 1 + 0.08\cdot 5 + 0.64\cdot 10 + 0.08\cdot 10\} \\
&= \max_{u\in\{S,N\}} \{10.8, 8.6\} \implies V^1(1) = 10.8 \\
V^1(2) &= \max_{u\in\{S,N\}} \{r(2,u) + \alpha P_{20}(u)V^0(0) + \alpha P_{21}(u)V^0(1) + \alpha P_{22}(u)V^0(2)\} \\
&= \max_{u\in\{S,N\}} \{2 + 0.32\cdot 5 + 0.08\cdot 10 + 0.4\cdot 10, 1 + 0.16\cdot 5 + 0.56\cdot 10 + 0.08\cdot 10\} \\
&= \max_{u\in\{S,N\}} \{8.4, 8.2\} \implies V^1(2) = 8.4
\end{aligned}
$$

d) Let's consider the initial policy $\mu^0(0) = S$, $\mu^0(1) = S$, $\mu^0(2) = S$ as initialization.

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

$$
\begin{aligned}
V_{\mu^0}(0) &= r(0,S) + \alpha P_{00}(S)V_{\mu^0}(0) \\
V_{\mu^0}(1) &= r(1,S) + \alpha P_{10}(S)V_{\mu^0}(0) + \alpha P_{11}(S)V_{\mu^0}(1) + \alpha P_{12}(S)V_{\mu^0}(2) \\
V_{\mu^0}(2) &= r(2,S) + \alpha P_{20}(S)V_{\mu^0}(0) + \alpha P_{21}(S)V_{\mu^0}(1) + \alpha P_{22}(S)V_{\mu^0}(2)
\end{aligned}
$$

Plug in the values:

$$
\begin{aligned}
V_{\mu^0}(0) &= -2 + 0.8V_{\mu^0}(0) \\
V_{\mu^0}(1) &= 4 + 0.24V_{\mu^0}(0) + 0.4V_{\mu^0}(1) + 0.16V_{\mu^0}(2) \\
V_{\mu^0}(2) &= 2 + 0.32V_{\mu^0}(0) + 0.08V_{\mu^0}(1) + 0.4V_{\mu^0}(2)
\end{aligned}
$$

We can have:

$$
\begin{aligned}
V_{\mu^0}(0) &= -10 \\
V_{\mu^0}(1) &= \frac{480}{217} \approx 2.21 \\
V_{\mu^0}(2) &= -\frac{370}{217} \approx -1.71
\end{aligned}
$$

Policy improvement:

$$
\begin{aligned}
\mu^1(0) &= \arg\max_{u\in\{S,N\}}\{r(0,u)+\alpha P_{00}(u)V_{\mu^0}(0)\} \\
&= \arg\max_{u\in\{S,N\}}\{-2+0.8\cdot(-10),\ -1+0.8\cdot(-10)\} \\
&= \arg\max_{u\in\{S,N\}}\{-10,-9\} \implies \mu^1(0)=N \\
\mu^1(1) &= \arg\max_{u\in\{S,N\}}\{r(0,u)+\alpha P_{00}(u)V_{\mu^0}(0)+\alpha P_{01}(u)V_{\mu^0}(0)+\alpha P_{02}(u)V_{\mu^0}(2)\} \\
&= \arg\max_{u\in\{S,N\}}\{4+0.24\cdot(-10)+0.4\cdot 2.21+0.16\cdot(-1.71), \\
&\qquad\qquad 1+0.08\cdot(-10)+0.64\cdot 2.21+0.08\cdot(-1.71)\} \\
&= \arg\max_{u\in\{S,N\}}\{2.21,1.48\} \implies \mu^1(0)=S \\
\mu^1(2) &= \arg\max_{u\in\{S,N\}}\{r(0,u)+\alpha P_{00}(u)V_{\mu^0}(0)+\alpha P_{01}(u)V_{\mu^0}(0)+\alpha P_{02}(u)V_{\mu^0}(2)\} \\
&= \arg\max_{u\in\{S,N\}}\{2+0.32\cdot(-10)+0.08\cdot 2.21+0.4\cdot(-1.71), \\
&\qquad\qquad 1+0.16\cdot(-10)+0.56\cdot 2.21+0.08\cdot(-1.71)\} \\
&= \arg\max_{u\in\{S,N\}}\{-1.71,0.50\} \implies \mu^1(0)=N
\end{aligned}
$$

The new policy is given by $\mu^1(0)=N$, $\mu^1(1)=S$, $\mu^1(2)=N$

(If we continue doing Policy Iteration, we will find the optimal policy is $\mu^1(0)=N$, $\mu^1(1)=S$, $\mu^1(2)=N$. So Peter Parker will act as Spider-Man when he is not exposed or injured the previous day, and as a normal student when he is injured or exposed.)

e) Be careful about the signs and inequalities.

$$
\begin{aligned}
&\mathit{minimize} \quad \sum_{i=0}^{2} V(i) \\
&\mathit{subject\,to} \quad V(i)\geq r(i,u)+\alpha\sum_{j=0}^{2}P_{ij}(u)V(j), \forall u\in\{S,N\}, i\in\{0,1,2\}
\end{aligned}
$$

$$
\mathbf{P}_S = \begin{bmatrix} P_{00}(S) & P_{01}(S) & P_{02}(S) \\ P_{10}(S) & P_{11}(S) & P_{12}(S) \\ P_{20}(S) & P_{21}(S) & P_{22}(S) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{bmatrix}
$$

$$
\mathbf{P}_N = \begin{bmatrix} P_{00}(N) & P_{01}(N) & P_{02}(N) \\ P_{10}(N) & P_{11}(N) & P_{12}(N) \\ P_{20}(N) & P_{21}(N) & P_{22}(N) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0.2 & 0.7 & 0.1 \end{bmatrix}
$$

$$\mathbf{A}_S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \alpha \mathbf{P}_S = \begin{bmatrix} 0.2 & 0 & 0 \\ -0.24 & 0.6 & -0.16 \\ -0.32 & -0.08 & 0.6 \end{bmatrix}$$

$$\mathbf{A}_N = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \alpha \mathbf{P}_N = \begin{bmatrix} 0.2 & 0 & 0 \\ -0.08 & 0.36 & -0.08 \\ -0.16 & -0.56 & 0.92 \end{bmatrix}$$

$$\mathbf{b}_S = \begin{bmatrix} r(0,S) \\ r(1,S) \\ r(2,S) \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \\ 2 \end{bmatrix}, \mathbf{b}_N = \begin{bmatrix} r(0,N) \\ r(1,N) \\ r(2,N) \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{f} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{A} = -\begin{bmatrix} \mathbf{A}_S \\ \mathbf{A}_N \end{bmatrix} = \begin{bmatrix} -0.2 & 0 & 0 \\ 0.24 & -0.6 & 0.16 \\ 0.32 & 0.08 & -0.6 \\ -0.2 & 0 & 0 \\ 0.08 & -0.36 & 0.08 \\ 0.16 & 0.56 & -0.92 \end{bmatrix}$$

$$\mathbf{b} = -\begin{bmatrix} \mathbf{b}_S \\ \mathbf{b}_N \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ -2 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

Problem 3 [25 points]

You have just reached Notting Hill in London, and your goal is to go to London Bridge as fast as possible by taking the Tube (London's underground). Therefore, you need to find the shortest path from Notting Hill to London Bridge. The following graph represents a map of the Tube in London showing the travelling time between Notting Hill (node S), London bridge (node T), and major tube stops (nodes A to F). The numbers on the edges represent the time (in minutes) to travel between the nodes (Figure 1).

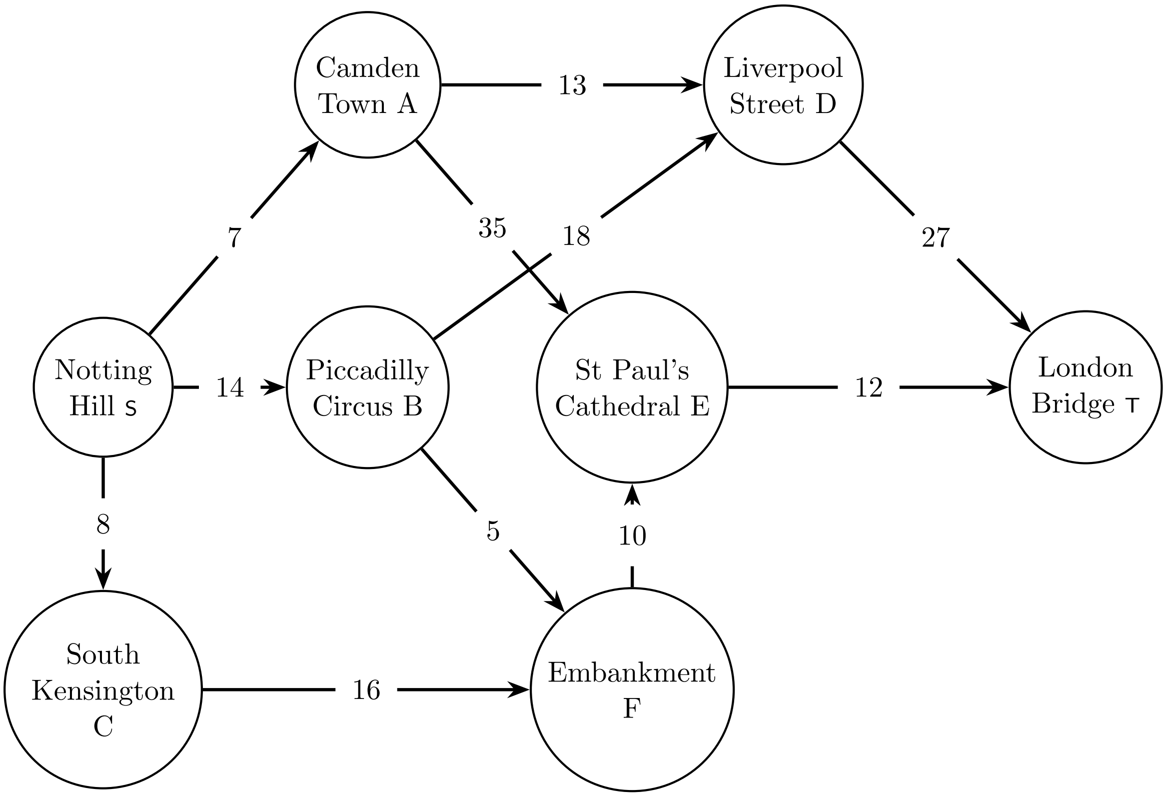

Figure 1

a) $\mathcal{V}$ 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) Apply the Dynamic Programming Algorithm to compute for each node the minimum time required to travel to the terminal node London Bridge T. *[8 points]*

c) 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.
Instructions: Recall that only one instance of a node can be in OPEN at any time. If a node that is already in the OPEN bin enters the OPEN bin again, treat this node as if it would enter the OPEN bin at the current iteration. If two nodes enter the OPEN bin in the same iteration, follow alphabetic order. Example: OPEN bin: A, B, F; Node exiting OPEN F (nodes entering OPEN: A, C); new OPEN bin: B, A, C; *[7 points]*

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

Table 1: Label Correcting Algorithm Table for question d). Best-first search

d) Due to delayed trains, the true travel times are higher than in Figure 1 and are shown by arrows between the nodes in Figure 2. Calculate the minimum travel time from node S to T and the corresponding path using the A^*-Algorithm and the results from b) as a lower bound. Use the depth-first (last-in/first-out) method to determine at each iteration which node exits the OPEN bin. Solve the problem by populating a table of the form given in Table 2. State the resulting path from node S to T and its travel time. *[8 points]*

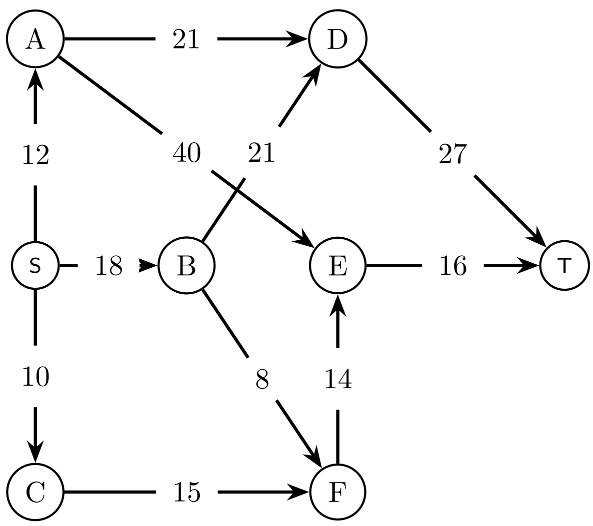

Figure 2

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

Table 2: A^* Algorithm Table for question d). Depth-First search

Solution 3

a) Cardinality is 8, time horizon is 7

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

i	$J_{N-1}(i)$
S	∞
A	∞
B	∞
C	∞
D	27
E	12
F	∞

Recursion: (2 moves to the end)

i	$J_{N-2}(i)$
S	∞
A	$\min\{27+13, 35+12\} = 40$
B	$18+27 = 45$
C	∞
D	27
E	12
F	$10+12 = 22$

Recursion: (3 moves to the end)

i	$J_{N-3}(i)$
S	$\min\{7+40, 14+45\} = 47$
A	40
B	$\min\{45, 5+22\} = 27$
C	$16+10+12 = 38$
D	27
E	12
F	22

Recursion: (4 moves to the end)

i	$J_{N-4}(i)$
S	$\min\{47, 14+27, 8+38\} = 41$
A	40
B	27
C	38
D	27
E	12
F	22

$J_{N-4}(i) = J_{N-5}(i)$. END!
The minimum time $h(i)$ required to travel from any node i to the terminal node 7 is:

i	$h(i)$
S	41 min
A	40 min
B	27 min
C	38 min
D	27 min
E	12 min
F	22 min

c) Table 1 filled in for question c).

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	7	14	8	∞	∞	∞	∞
2	A	B,C,D,E	0	7	14	8	20	42	∞	∞
3	C	B,D,E,F	0	7	14	8	20	42	24	∞
4	B	D,E,F	0	7	14	8	20	42	19	∞
5	F	D,E	0	7	14	8	20	29	19	∞
6	D	E	0	7	14	8	20	29	19	47
7	E	-	0	7	14	8	20	29	19	41
8	-		0	7	14	8	20	29	19	41

Table 3: Label Correcting Algorithm Table for question c). (Best-First Search)

The optimal path is S, B, F, E, T. The travel time is 41 minutes.

d) Table 2 filled in for question d).

Iteration	Remove	OPEN	d_s	d_A	d_B	d_C	d_D	d_E	d_F	d_t
0	-	s	0	∞	∞	∞	∞	∞	∞	∞
1	s	A,B,C	0	12	18	10	∞	∞	∞	∞
2	C	A,B,F	0	12	18	10	∞	∞	25	∞
3	F	A,B,E	0	12	18	10	∞	39	25	∞
4	E	A,B	0	12	18	10	∞	39	25	55
5	B	A	0	12	18	10	∞^1	39	25	55
6	A	-	0	12	18	10	∞^2	39	25	55

Table 4: A Algorithm Table for question d). (Last-In, First-Out)*

1) D does not enter the open bin as $d_B + h_D + a_{BT} > d_T$
2) D does not enter the open bin as $d_A + h_D + a_{AT} > d_T$
The optimal path is S, C, F, E, T; the travel time is 55 minutes.

Problem 4 **[25 points]**

Consider a control maneuver for a rocket in space. The rocket can move in a straight line by applying a force with its thrusters. The simplified dynamics are

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

where $x(t)$ is the position and $u(t)$ the normalized thrust of the spacecraft, with $|u(t)| \leq \frac{3}{2}$.

At time 0, the spacecraft is at rest, thus the initial position is 0 and the initial velocity is 0. At time $T = 2$, the spacecraft is required to be at a velocity equal to 2.

The goal is to compute a control trajectory that allows the spacecraft to reach the terminal state while minimizing the following cost function:

$$\int_0^T (\dot{x}(t) - \alpha)^2 dt$$

The co-state of this problem is defined as $p(t) = [p_1(t), p_2(t)]^T$ for all $t \in [0, T]$.

a) Write the system dynamics $\dot{x}(t) = f(x(t), u(t))$ and all boundary conditions. *[2 points]*

b) Write the Hamiltonian $H(x, u, p)$ of the problem. *[2 points]*

c) Write the co-state derivative $\dot{p}(t)$ of the problem. *[2 points]*

d) By considering $0 < \alpha < 2$, compute the control input $u^*(t)$ that satisfies the Pontryagin's Minimum Principle. Discuss what happens if $\alpha = 0$ or if $\alpha = 2$. *[7 points]*

e) By considering $\alpha > 2$, compute the control input $u^*(t)$ that satisfies the Pontryagin's Minimum Principle. Find the value of α for which the resulting velocity trajectory $x_2(t)$ derived from the Minimum Principle never intersects α. *[6 points]*

f) Let's now consider $\alpha = -1$. compute the control input $u^*(t)$ that satisfies the Pontryagin's Minimum Principle. What is the final position reached by the rocket? *[6 points]*

Solution 4

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

Boundary conditions:

- $x_1(0) = 0$;
- $x_2(0) = 0$, $x_2(T) = 2$.

b) The Hamiltonian is

$$H(x, u, p) = (x_2 - \alpha)^2 + p_1 x_2 + p_2 u$$

c) The co-state derivatives are defined by

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

hence, by substituting x(t), p(t), u(t), we have $\dot{p}_1(t) = 0$, $\dot{p}_2(t) = -p_1(t) - 2x_2(t) + 2\alpha$,

Since the Hamiltonian is linear in u, The optimal control input $u^*(t)$ is attained on the boundaries of the control space C= $[-\frac{3}{2}, \frac{3}{2}]$. The solution is the following (Bang-Bang Solution):

- $u^*(t) = -\frac{3}{2}$, if $p_2(t) > 0$
- $u^*(t) = \frac{3}{2}$, if $p_2(t) < 0$
- $u_s(t)$ if $p_2(t) = 0$

d) The intuitive behaviour is the following: the rocket first accelerates to reach the reference velocity, then stops applying thrust, and finally accelerates again to reach the velocity at the goal. Combining this intuition, we must first have $p_2(t) < 0$, then $p_2(t) = 0$, and at the end again $p_2(t) < 0$. Anything else would result in not being able to meet the terminal conditions. We then have three regular arcs:

$$u(t) = \begin{cases} \frac{3}{2} & \text{for } t \in [0, t_1] \\ 0 & \text{for } t \in [t_1, t_2] \\ \frac{3}{2} & \text{for } t \in [t_2, T] \end{cases} \tag{2}$$

and $t_1 \leq t_2$. We have to find t_1 and t_2.

- First arc $t \in [0, t_1]$
  We have $\dot{x}_2(t) = \frac{3}{2}$ with the boundary conditions $x_1(0) = 0$ and $x_2(0) = 0$.
  As a result we have

$$x_1(t) = \frac{3}{4}t^2$$
$$x_2(t) = \frac{3}{2}t$$

- Second arc $t \in [t_1, t_2]$
  We have $\dot{x}_2(t) = 0$. Because of continuity between arcs 1 and 2, the boundary conditions are $x_1(t_1) = \frac{3}{4}t_1^2$ and $x_2(t_1) = \frac{3}{2}t_1$. But $\frac{3}{2}t_1 = \alpha$; hence as a result we have

$$\begin{aligned} x_1(t) &= \alpha(t - \frac{1}{2}t_1) \\ x_2(t) &= \alpha \end{aligned}$$

  and $t_1 = \frac{2}{3}\alpha$
- Third arc $t \in [t_2, T]$
  We have $\dot{x}_2(t) = \frac{3}{2}$ and thus

$$\begin{aligned} x_1(t) &= \frac{3}{4}t^2 + C_3 t + C_4 \\ x_2(t) &= \frac{3}{2}t + C_3. \end{aligned}$$

  Due to the terminal conditions, we have $x_2(T) = 2$; moreover, $x_2(t_2) = \alpha$.
  Hence:

$$x_2(T) = \frac{3}{2}T + C_3 = 2.$$

  where C_3=-1. By considering that $\frac{3}{2}t_2 - 1 = \alpha$, we get that $t_2 = \frac{2}{3}(\alpha + 1)$

Solution:

$$u(t) = \begin{cases} \frac{3}{2} & \text{for } t \in [0, \frac{2}{3}\alpha] \\ 0 & \text{for } t \in [\frac{2}{3}\alpha, \frac{2}{3}(\alpha+1)] \\ \frac{3}{2} & \text{for } t \in [\frac{2}{3}(\alpha+1), 2] \end{cases} \tag{3}$$

If $\alpha = 0$, the solution is in two arcs. During the first arc, the rocket maintains the rest condition until $t_1 = \frac{2}{3}$, then applies the input $\frac{3}{2}$ to satisfy the terminal constraint. Conversely, if $\alpha = 2$, the rocket accelerates to reach the reference until $t_1 = \frac{4}{3}$, and then stops applying thrust to satisfy the terminal condition.

e) In this case the intuitive behaviour is the following: the rocket first accelerates to reach the reference velocity, then stops applying thrust, and finally decelerates again to reach the velocity at the goal. Combining this intuition, we must first have $p_2(t) < 0$, then $p_2(t) = 0$, and at the end $p_2(t) > 0$. Anything else would result in not being able to meet the terminal conditions. We then have three regular arcs:

$$u(t) = \begin{cases} \frac{3}{2} & \text{for } t \in [0, t_1] \\ 0 & \text{for } t \in [t_1, t_2] \\ -\frac{3}{2} & \text{for } t \in [t_2, T] \end{cases} \tag{4}$$

and $t_1 \leq t_2$. We have to find t_1 and t_2. $t_1 = \frac{2}{3}\alpha$, same as before. The difference is in the third arc $t \in [t_2, T]$:

We have $\dot{x}_2(t) = -\frac{3}{2}$ and thus

$$x_1(t) = -\frac{3}{4}t^2 + C_3 t + C_4$$
$$x_2(t) = -\frac{3}{2}t + C_3.$$

Due to the terminal conditions, we have $x_2(T) = 2$; moreover, $x_2(t_2) = \alpha$.
Hence:

$$x_2(T) = -\frac{3}{2}T + C_3 = 2.$$

where C_3=5. By considering that $-\frac{3}{2}t_2 + 5 = \alpha$, we get that $t_2 = \frac{2}{3}(5 - \alpha)$

Solution:

$$u(t) = \begin{cases} \frac{3}{2} & \text{for } t \in [0, \frac{2}{3}\alpha] \\ 0 & \text{for } t \in [\frac{2}{3}\alpha, \frac{2}{3}(5-\alpha)] \\ -\frac{3}{2} & \text{for } t \in [\frac{2}{3}(5-\alpha), 2] \end{cases} \tag{5}$$

$t_1 = t_2 \Rightarrow \frac{2}{3}\alpha = \frac{2}{3}(5 - \alpha)$ we get $\alpha = \frac{5}{2}$. If $\alpha > \frac{5}{2}$, the trajectory of the velocity never intersects α, and the solution is in two arcs, with $t_1 = 5/3$.

f) In this case the intuitive behaviour is the following: the rocket first decelerates to reach the reference velocity, and finally accelerates again to reach the velocity at the goal. Combining this intuition, we must first have $p_2(t) < 0$ and at the end $p_2(t) > 0$. Anything else would result in not being able to meet the terminal conditions. We then have two regular arcs:

- First arc $t \in [0, t_1]$
  We have $\dot{x}_2(t) = \frac{3}{2}$ with the boundary conditions $x_1(0) = 0$ and $x_2(0) = 0$.
  As a result we have

$$x_1(t) = -\frac{3}{4}t^2$$
$$x_2(t) = -\frac{3}{2}t$$

- second arc $t \in [t_1, T]$
  We have $\dot{x}_2(t) = \frac{3}{2}$ and thus

$$x_1(t) = \frac{3}{4}t^2 + C_3 t + C_4$$
$$x_2(t) = \frac{3}{2}t + C_3.$$

  Due to the terminal conditions, we have $x_2(T) = 2$.Hence

$$x_2(T) = \frac{3}{2}T + C_3 = 2.$$

  where C_3=-1.
  By combining the two arcs we get $-\frac{3}{2}t_1 = \frac{3}{2}t_1 - 1 \Rightarrow t_1 = \frac{1}{3}$.
  The terminal position after t_1 is $x_1(t_1) = -\frac{1}{12} \Rightarrow \frac{3}{4}t_1^2 - t_1 + C_4 = -\frac{1}{12} \Rightarrow C_4 = \frac{1}{6}$.
  Hence $x_1(T) = \frac{7}{6}$